ANALYSIS I-III

2011/2012

16. April 2020

GEORGE MARINESCU

Inhaltsverzeichnis

13.10.2011

1. Reelle und komplexe Zahlen

Die Zahlen sind freie Schöpfungen des menschlichen Geistes, sie dienen als ein Mittel, um die Verschiedenheit der Dinge leichter und schärfer aufzufassen.

R. Dedekind

Was sind die reellen Zahlen? Wir folgen der axiomatisch-deduktiven Methode, die uns gestattet, dieser Frage auszuweichen. Wir nutzen aus, dass die uns vertrauten reellen Zahlen einige Strukturen tragen, die den Bedingungen eines sogenannten *vollständig angeordneten Körpers* genügen. Diese Bedingungen betrachten wir als Axiomensystem für die reellen Zahlen, und aus diesem Axiomensystem lassen sich alle Aussagen über reelle Zahlen deduzieren, die in der Mathematik benötigt werden. Wir betonen also die Eigenschaften der reellen Zahlen und nicht ihre Natur.

Eine zusätzliche Rechtfertigung erfährt dieses Vorgehen durch ein Resultat, demzufolge je zwei Strukturen, die den Bedingungen dieses Axiomensystems genügen, *isomorph* sind und daher die gleichen für die Mathematik bedeutsamen Eigenschaften haben. Wir *definieren* also die reellen Zahlen als einen vollständig angeordneten Körper.

Man kann sich nun fragen, ob ein vollständig angeordneter Körper überhaupt *existiert*. Zur Beruhigung ängstlicher Gemüter sei sogleich gesagt, dass es verschiedene Konstruktionen solch eines Körpers gibt. Man konstruiert zunächst die Menge der natürlichen Zahlen, dann die Menge der ganzen Zahlen und schließlich die Menge der rationalen Zahlen. Da die Menge der rationalen Zahlen „lückenhaft“ ist, konstruiert man die reellen Zahlen durch einen Prozess der „Vervollständigung“. Es gibt dafür drei Methoden: durch Dedekindsche Schnitte, durch Fundamentalfolgen (Cauchy-Folgen) oder durch Intervallschachtelungen. Diesen Weg zu gehen, kostet aber mehr Zeit, als wir zur Verfügung haben. Eine sehr gute Quelle dazu ist das Buch [6] (oder der Klassiker von E. Landau [14]).

Wieso heißen die reellen Zahlen *reell*? Die reellen Zahlen sind idealisierte Modelle von reellen Objekten. Sie sind entstanden aus dem Wunsch, Längen mit absoluter Exaktheit zu berechnen, z.B. die Länge der Diagonale eines Einheitsquadrates. (Man beachte, dass auch ein Einheitsquadrat eine Idealisierung ist.) Die intuitive Vorstellung ist, dass jede reelle Zahl einem Punkt auf der Zahlengeraden entspricht. Wir möchten eine präzise und explizite mathematische Formulierung dieser Intuition geben.

Einen Platonischen Dialog[1] über die Notwendigkeit der reellen Zahlen findet man auf der Webseite von T. Gowers.

1.1. Körperaxiome.

1.1.1. **Definition.** Sei K eine Menge. Wir setzen voraus, dass für $a,b\in K$ die Summe $a+b\in K$ und das Produkt $a\cdot b=ab\in K$ erklärt sind. Die Abbildungen $+:K\times K\longrightarrow K$, $(a,b)\mapsto a+b$, $\cdot:K\times K\longrightarrow K$, $(a,b)\mapsto ab=a\cdot b$ heißen Addition und Multiplikation.

Es gelten die ***Körperaxiome***:

1. *Axiome für die Addition*:

(A1) Für alle $a,b,c\in K$ gilt $(a+b)+c=a+(b+c)$. (Assoziativgesetz)

(A2) Es gibt ein Element $0\in K$, so dass für alle $a\in K$ gilt: $a+0=a$. (Neutrales Element der Addition oder Nullelement)

(A3) Für alle $a\in K$ gibt es $(-a)\in K$ mit $a+(-a)=0$. (Additives Inverses)

(A4) Für alle $a,b\in K$ ist $a+b=b+a$. (Kommutativgesetz)

[1] http://www.dpmms.cam.ac.uk/~wtg10/reals.html

2. *Axiome für die Multiplikation*:

(M1) Für alle $a, b, c \in K$ gilt $(ab)c = a(bc)$. (Assoziativgesetz)

(M2) Es gibt ein Element $1 \in K \setminus \{0\}$, so dass für alle $a \in K$ gilt: $a \cdot 1 = a$. (Neutrales Element der Multiplikation oder Einselement)

(M3) Für alle $a \neq 0$ gibt es $a^{-1} \in K$ mit $a a^{-1} = 1$. (Multiplikatives Inverses)

(M4) Für alle $a, b \in K$ ist $ab = ba$. (Kommutativgesetz)

3. *Distributivgesetz*:

(D) Für alle $a, b, c \in K$ gilt $(a + b)c = ac + bc$.

Dann heißt K zusammen mit der Addition und der Multiplikation ein ***Körper***, bezeichnet mit $(K, +, \cdot)$.

Körpern werden intensiv in der Algebra und Zahlentheorie studiert und sind zentral z.B. in der Galois-Theorie (siehe der Klassiker van der Waerden [22] oder Lang [15]). Beispiele von Körpern: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ mit p Primzahl (Restklassenkörper), $\mathbb{Q}(\sqrt{2}) := \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ (oder allgemeiner $\mathbb{Q}(\sqrt{m})$, mit m quadratfreie ganze Zahl, sog. quadratische Zahlkörper), Körper der rationalen Funktionen $K(x) = \{P/Q : P, Q$ Polynome in der Variabel $x\}$ usw. Körpern sind für die Lineare Algebra grundlegend, da Vektorräume mit Hilfe eines Grundkörpers definiert werden (siehe z.B. [12]). 17.10.2011

1.1.2. **Satz.**

(a) *Das Nullelement und das Einselement sind eindeutig bestimmt.*

(b) *Für alle $a, b \in K$ hat die Gleichung $a + x = b$ genau eine Lösung $x_0 = b + (-a) =: b - a$. Für alle $a, b \in K$, $a \neq 0$ hat die Gleichung $ax = b$ genau eine Lösung $x_0 = a^{-1}b =: \frac{b}{a}$. Insbesondere ist das zu a (bzw. $a \neq 0$) additive (bzw. multiplikative) Inverse eindeutig bestimmt.*

(c) *Es gelten die folgenden Rechenregeln:*

$$
\begin{aligned}
&-(-a) = a\,, \quad (-a) + (-b) = -(a + b) \\
&(a^{-1})^{-1} = a\,, \quad a^{-1}b^{-1} = (ab)^{-1} \quad (a, b \neq 0) \\
&a \cdot 0 = 0\,, \quad a \cdot (-b) = -ab\,, \quad (-a)(-b) = ab \\
&a(b - c) = ab - ac
\end{aligned}
$$

(d) *Aus $ab = 0$ folgt $a = 0$ oder $b = 0$.*

1.2. **Anordnungsaxiome.**

1.2.1. **Definition.** Ein Körper $(K, +, \cdot)$ heißt ***total angeordnet***, falls eine Teilmenge $P \subset K$ mit den folgenden Eigenschaften existiert:

(O1) Für jedes $x \in K$ gilt genau eine der Beziehungen $x \in P$, $-x \in P$, $x = 0$.

(O2) Für alle $x, y \in P$ gilt $x + y, xy \in P$.

Ist $x \in P$ (bzw. $-x \in P$), so heißt x ***positiv*** (bzw. ***negativ***). In einem total angeordneten Körper K führen wir eine Kleiner-Relation ein:

Seien $x, y \in K$. Wir sagen „x ist kleiner als y“, geschrieben $x < y$, oder „y ist größer als x“, geschrieben $y > x$, falls $y - x \in P$.

$$x < y :\Leftrightarrow y - x \in P\,; \text{also} x \in P \Leftrightarrow x > 0 \text{und} -x \in P \Leftrightarrow x < 0\,.$$

1.2.2. **Satz** (Rechenregeln für Ungleichungen)**.** *Sei K ein total angeordneter Körper. Für alle $x, y, z, t \in K$ gilt:*

(a) *$x < y$ und $y < z \Rightarrow x < z$ (Transitivität)*

(b) *$x < y \Rightarrow x + z < y + z$*

(c) *$x < y \Rightarrow -y < -x$*

(d) *$x < y$ und $z > 0 \Rightarrow xz < yz$; $x < y$ und $z < 0 \Rightarrow xz > yz$*

(e) *$x \neq 0 \Rightarrow x^2 > 0$. Insbesondere $1 > 0$.*

(f) $x > 0 \Rightarrow \frac{1}{x} > 0$; $x < 0 \Rightarrow \frac{1}{x} < 0$
(g) $0 < x < y \Rightarrow 0 < \frac{1}{y} < \frac{1}{x}$, $\frac{x}{y} < 1$, $\frac{y}{x} > 1$
(h) $x < y$ *und* $z < t \Rightarrow x + z < y + t$
(i) $0 < x < y$, $0 < z < t \Rightarrow 0 < xz < yt$
(j) $x < y$ *und* $0 < \lambda < 1 \Rightarrow x < \lambda x + (1-\lambda)y < y$.

Jetzt können wir zeigen, dass ein total angeordneter Körper außer 0 und 1 noch weitere Elemente hat! Denn $0 < 1 \Rightarrow 0 + 1 < 1 + 1 =: 2 \Rightarrow 0 < 1 < 2$ usw. In einem beliebigen Körper (z.B. in $\mathbb{Z}/2\mathbb{Z}$) kann es passieren, dass $1 + 1 = 0$!

Wenn wir $\lambda = \frac{1}{2}$ in (j) betrachten, ergibt sich $x < \frac{x+y}{2} < y$. Für zwei (nicht unbedingt verschiedene) Elemente $x, y \in K$ heißt $\frac{x+y}{2}$ das ***arithmetische Mittel*** von x und y.

1.2.3. **Definition.** Sei K ein total angeordneter Körper. Ein Element $x \in K$ wird ***nichtnegativ*** genannt, geschrieben $x \geqslant 0$ (x größer-gleich 0), wenn $x > 0$ oder $x = 0$. Wir definieren eine zweistellige Relation $\leqslant$ auf K durch $x \leqslant y :\Leftrightarrow y - x \geqslant 0$ (x kleiner-gleich y). $\leqslant$ ist eine ***Ordnungsrelation***, d.h.

(a) $x \leqslant x$ (Reflexivität)
(b) $x \leqslant y$ und $y \leqslant z \Rightarrow x \leqslant z$ (Transitivität)
(c) $x \leqslant y$ und $y \leqslant x \Rightarrow x = y$ (Antisymmetrie)

Eine ganz wichtige (aber einfache) Ungleichung ist

$$\frac{x^2 + y^2}{2} \geqslant xy, \quad \text{für alle } x, y \in K, \tag{1.1}$$

wobei die Gleichheit genau dann auftritt, wenn $x = y$. Das folgt aus $(x-y)^2 \geqslant 0$ mit Gleichheit genau dann, wenn $x = y$. Die Ungleichung (1.1) ist ein Spezialfall der AGM-Ungleichung, die wir später beweisen (siehe (1.8)).

1.2.4. **Definition** (Vorzeichen und Absolutbetrag)**.**
Sei K ein total angeordneter Körper. Sei $x \in K$. Man nennt

$$\operatorname{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

das ***Vorzeichen*** von x. Der ***Betrag*** von x ist

$$|x| := x \cdot \operatorname{sgn}(x) = \begin{cases} x & x > 0 \\ 0 & x = 0 \\ -x & x < 0 \end{cases} \quad \text{d.h.} \quad |x| = \begin{cases} x & , x \geqslant 0 \\ -x & , x < 0 . \end{cases}$$

1.2.5. **Satz** (Rechenregeln)**.** *Sei K ein total angeordneter Körper. Für alle $x, y \in K$ gilt:*

(a) $|x| \geqslant 0$ *and* $x \neq 0 \iff |x| > 0$,
(b) $||x|| = |x|$, $|-x| = |x|$,
(c) $|x| \leqslant y \iff \pm x \leqslant y \iff -y \leqslant x \leqslant y$,
(d) $|x + y| \leqslant |x| + |y|$, $\big||x| - |y|\big| \leqslant |x - y|$ *(Dreiecksungleichungen),*
(e) $|x| \leqslant |y| \Leftrightarrow x^2 \leqslant y^2$,
(f) $x = y \Leftrightarrow |x| = |y|$ *und* $\operatorname{sgn}(x) = \operatorname{sgn}(y)$,
20.10.2011 (g) $\operatorname{sgn}(xy) = \operatorname{sgn}(x) \cdot \operatorname{sgn}(y)$, $|xy| = |x| \cdot |y|$,

Die Dreiecksungleichungen sind ein wichtiges Mittel in der Analysis. Die Gleichheit in der Dreiecksungleichung $|x + y| \leqslant |x| + |y|$ gilt genau dann, wenn x und y dasselben Vorzeichen haben oder einer von x, y Null ist, d. h. es gibt ein $\lambda \geqslant 0$ mit $x = \lambda y$. Ansonsten ist die Ungleichung strikt, z. B. für $x = 2$, $y = -1$ gilt $|x + y| = |2 - 1| = 1$, $|x| + |y| = |2| + |-1| = 3$. Folgende Varianten der Dreiecksungleichung sind oft benutzt:

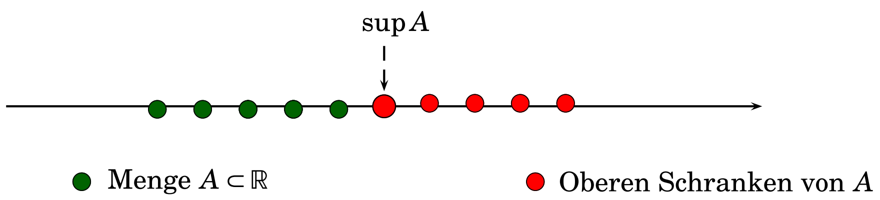
$$|x| \geqslant |y| - |x - y|, \quad \big||x| - |y|\big| \leqslant |x + y|$$

Die Dreiecksungleichung gilt für Vektoren in höherer Dimension. Für zwei Vektoren x und y bilden x, y und $x+y$ die Seiten eines Dreiecks und die Dreiecksungleichung besagt, daß zwei Seiten eines Dreiecks zusammen immer mindestens so lang wie die dritte sind. In der Dimension eins degeneriert das Dreieck zu drei Punkten auf der Geraden, aber die Aussage bleibt richtig.

1.3. **Vollständigkeitsaxiom.** Dieses Axiom bringt die geometrische Anschauung zum Ausdruck, dass die Zahlengerade ein Kontinuum ist. Die Vollständigkeit ist eine der wichtigsten Eigenschaften von $\mathbb{R}$. Ohne sie könnten wir z.B. keine Wurzeln ziehen, keine Reihen summieren und nicht integrieren.

1.3.1. **Definition.** Sei K ein total angeordneter Körper, $A \subset K$, $A \neq \emptyset$.

- A heißt ***nach oben beschränkt***, falls es ein $c \in K$ gibt mit $a \leqslant c$ für alle $a \in A$; c heißt dann eine ***obere Schranke*** von A. Ist c eine obere Schranke von A mit $c \in A$, so heißt c ***Maximum von*** A, bezeichnet mit $\max A$.
- A heißt ***nach unten beschränkt***, falls es ein $c \in K$ gibt mit $a \geqslant c$ für alle $a \in A$; c heißt dann eine ***untere Schranke*** von A. Ist c eine untere Schranke von A mit $c \in A$, so heißt c ***Minimum von*** A, bezeichnet mit $\min A$.
- Ist A nach unten und oben beschränkt, so heißt A ***beschränkt***.
- Ein Element $c \in K$ heißt ***kleinste obere Schranke*** oder ***Supremum*** von A, bezeichnet mit $\sup A$, falls die folgenden Eigenschaften erfüllt sind:
  (i) $a \leqslant c$ für alle $a \in A$.
  (ii) Ist b eine obere Schranke von A, so folgt $c \leqslant b$.
- Ein Element $c \in K$ heißt ***größte untere Schranke*** oder ***Infimum*** von A, bezeichnet mit $\inf A$, falls die folgenden Eigenschaften erfüllt sind:
  (i) $a \geqslant c$ für alle $a \in A$
  (ii) Ist b eine untere Schranke von A, so folgt $c \geqslant b$.



1.3.2. **Bemerkung.**

(i) Hat eine Menge ein Maximum (bzw. Minimum, Supremum, Infimum), so ist dieses eindeutig bestimmt.
(ii) Falls A ein Maximum hat, so hat A ein Supremum, und es gilt $\sup A \in A$.
In diesem Falle ist $\max A = \sup A$.
(iii) Es gibt nach oben beschränkte Mengen, die ein Supremum besitzen, aber kein Maximum, z.B. $K_- = \{x \in K : x < 0\}$. Dann ist die Menge der oberen Schranken von K_- die Menge $\{y : y \geqslant 0\}$ also $\sup K_- = 0$, aber K_- hat kein Maximum, weil $0 \notin K_-$.
(iv) $\sup A = \min B$, wobei $B = \{b : b \text{ obere Schranke von } A\}$,
$\inf A = \max B$, wobei $B = \{b : b \text{ untere Schranke von } A\}$.
(v) Genau dann gilt $c = \sup A$, wenn
  (a) $a \leqslant c$ für alle $a \in A$, und
  (b) es für alle $\varepsilon > 0$ ein $a \in A$ mit $c - \varepsilon < a$ gibt.
(vi) c ist keine obere Schranke für $A \Leftrightarrow \exists a \in A : a > c$.
(vii) A ist nicht nach oben beschränkt $\Leftrightarrow \forall c \in K\, \exists a \in A : a > c$.

1.3.3. **Definition.** Ein total angeordneter Körper K heißt ***vollständig***, falls die folgende Eigenschaft (Vollständigkeitsaxiom) erfüllt ist:

(V) Jede nichtleere nach oben beschränkte Teilmenge von K besitzt ein Supremum.

Die Aussage (V) ist äquivalent zur folgenden Aussage:

(V') Jede nichtleere nach unten beschränkte Teilmenge von K besitzt ein Infimum.

In der Tat, A ist nach unten beschränkt g.d.w. $-A := \{-a : a \in A\}$ nach oben beschränkt ist. Gegebenfalls ist $\inf A = -\sup(-A)$.

1.3.4. **Bemerkung.** Der Körper $\mathbb{Q}$ der rationalen Zahlen ist ein total angeordneter Körper, der nicht vollständig ist. In der Tat, die Menge $A = \{x \in \mathbb{Q} : x^2 < 2\}$ ist eine nichtleere nach oben beschränkte Menge (z. B. 2 ist eine obere Schranke), die kein Supremum besitzt. Nehmen wir an, es gibt $c = \sup A \in \mathbb{Q}$. Da $1 \in A$, so ist $c \geq 1 > 0$. Es gibt nun drei Fälle für c^2: $c^2 < 2$, $c^2 > 2$ oder $c^2 = 2$.

Sei $c^2 < 2$. Wir suchen ε, so dass $b = c + \varepsilon \in A$, d. h. $b^2 = c^2 + 2c\varepsilon + \varepsilon^2 < 2$ oder $2c\varepsilon + \varepsilon^2 < 2 - c^2$ und $c < b$. Wir suchen $\varepsilon < 2 - c$ und dann $2c\varepsilon + \varepsilon^2 < \varepsilon(2c + \varepsilon) < \varepsilon(2 + c)$. Wähle also

$$\varepsilon = \frac{2 - c^2}{c + 2}, \quad b = c + \varepsilon = \frac{2c + 2}{c + 2} \cdot$$

Dann gilt

$$b > 0, \quad b - c = \varepsilon, \quad b^2 - 2 = \frac{2(c^2 - 2)}{(c + 2)^2} \cdot$$

Es gilt $\varepsilon > 0$, da $c^2 < 2$. Es folgt $c < b$ und $b^2 < 2$, also $b \in A$, d. h. c ist keine obere Schranke von A. Dies ist ein Widerspruch zu $c = \sup A$, dieser Fall ist ausgeschlossen.

Ist $c^2 > 2$, sei b wie oben. Es folgt $c > b$ und $b^2 > 2$. Für jedes $a \in A$ gilt also $a^2 < 2 < b^2$. Ist $a \leq 0$ so ist $a < b$. Ist $a > 0$ so ist $a + b > 0$ und wegen $b^2 - a^2 = (b - a)(a + b) > 0$ gilt $b - a > 0$, also $b > a$. Somit ist b eine obere Schranke von A und $b < c$. Dies ist aber falsch, da c die *kleinste* obere Schenke von A ist. Dieser Fall ist also auch ausgeschlossen.

Es bleibt also nur die Möglichkeit, dass $c^2 = 2$. Das ist aber ein Widerspruch, da es keine rationale Zahl c gibt, mit $c^2 = 2$ (siehe Satz A.3.10).

Im Satz 1.5.8 über Existenz der k-ten Wurzel wird auch gezeigt, dass $c^2 = 2$.

Wir werden bald viele Anwendungen des Vollständigkeitsaxioms erhalten, unter anderem die Existenz der Quadratwurzel.

1.3.5. **Definition.** Ein vollständig angeordneter Körper K heißt ***Körper der reellen Zahlen***, bezeichnet mit $\mathbb{R}$.

Um diese Definition zu rechtfertigen müssen wir zwei Fragen beantworten. Existiert ein Körper mit diesen Eigenschaften? Inwieweit ist er eindeutig bestimmt?

1.3.6. **Satz** (Existenz von $\mathbb{R}$; Cantor, Dedekind)**.** *Es gibt mindestens einen Körper der reellen Zahlen.*

Zum Beweis und Aufbau des Zahlsystems $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ aus den natürlichen Zahlen siehe [6] oder den Klassiker von E. Landau [14]. Die Konstruktion von Dedekind basiert sich auf Dedekindschen Schnitte, die Konstruktion von Cantor auf Cauchy-Folgen.

1.3.7. **Satz** (Eindeutigkeit der reellen Zahlen)**.** *Seien K und K' zwei Körper der reellen Zahlen. Dann gibt es eine bijektive Abbildung $\varphi : K \longrightarrow K'$, so dass für alle $x, y \in K$ gilt:*

(i) $\varphi(x + y) = \varphi(x) + \varphi(y)$
(ii) $\varphi(xy) = \varphi(x)\varphi(y)$
(iii) $\varphi(1) = 1$
(iv) $x < y \Rightarrow \varphi(x) < \varphi(y)$.

Der Satz besagt, dass es eine bijektive Abbildung $\varphi : K \longrightarrow K'$ gibt, die verträglich mit der algebraischen Strukturen und Ordnungstruktur ist. Eine solche Abbildung heißt Isomorphismus von total angeordneter Körpern. Wir sagen, dass $\mathbb{R}$ bis auf Isomorphie (von total angeordneter Körpern) eindeutig bestimmt ist. Da für uns nur diese Strukturen wichtig sind, sprechen wir von *dem* Körper $\mathbb{R}$.

1.4. **Natürliche Zahlen und vollständige Induktion.** Bisher haben wir ein naives Verständnis der natürlichen Zahlen. Wir *definieren* nun die Menge der natürlichen Zahlen. Wir möchten zum einen, dass 1 eine natürliche Zahl ist, und zum zweiten, dass jede natürliche Zahl n einen Nachfolger $n+1$ hat. Wir präzisieren diesen Gedanken so:

1.4.1. **Definition.** Eine Menge $M \subset \mathbb{R}$ heißt ***induktiv***, falls
(a) $1 \in M$,
(b) $x \in M \Rightarrow x+1 \in M$.

Die Menge $\mathbb{R}$ oder die Menge der positiven Zahlen $\mathbb{R}_+$ sind induktiv. Es ist klar, dass eine induktive Menge mindestens die Elemente $1, 2 := 1+1, 3 := 2+1, \ldots$ enthalten muss: Eben die natürlichen Zahlen nach unserem naiven Verständnis. Es gibt viele induktive Mengen, z.B. $M = \{x \geqslant 1\}$. Uns interessiert die kleinste solche Menge. Die ***Menge der natürlichen Zahlen*** ist definiert als Durchschnitt aller induktiven Teilmengen von $\mathbb{R}$:

$$\mathbb{N} := \bigcap \{M : M \subset \mathbb{R}, M \,\text{induktiv}\}$$

$\mathbb{N}$ ist auch induktiv, ist also die kleinste induktive Menge, d. h. für jede induktive Teilmenge M von $\mathbb{R}$ gilt $\mathbb{N} \subset M$.

1.4.2. **Definition.** Die ***Menge der ganzen Zahlen*** ist definiert durch

$$\mathbb{Z} := \mathbb{N} \cup \{0\} \cup (-\mathbb{N}) = \{0, \pm 1, \pm 2, \ldots\} \subset \mathbb{R}.$$

$(\mathbb{Z}, +, \cdot)$ ist ein Ring, wobei die Addition und Multiplikation von der entsprechenden Operationen von $\mathbb{R}$ induziert sind. Wir setzen

$$\mathbb{N}_0 := \mathbb{N} \cup \{0\}.$$

Die ***Menge der rationalen Zahlen*** ist definiert durch

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

Zusammen mit der Addition und Multiplikation der reellen Zahlen und mit der Ordnung der reellen Zahlen ist $\mathbb{Q}$ ein total angeordneter Körper. Eine Zahl $x \in \mathbb{R} \setminus \mathbb{Q}$ heißt ***irrational***.

1.4.3. **Satz** (Induktionsprinzip)**.** *Hat $M \subset \mathbb{N}$ die Eigenschaften:*
(a) $1 \in M$,
(b) $x \in M \Rightarrow x+1 \in M$,

so folgt $M = \mathbb{N}$.

1.4.4. **Satz** (Beweis durch vollständige Induktion)**.** *Seien $A(n)$ Aussagen, die für alle $n \in \mathbb{N}$ definiert sind. Wenn*
(a) *$A(1)$ richtig ist und*
(b) *$\forall n \in \mathbb{N}\; A(n) \Rightarrow A(n+1)$ gilt,*

dann ist $A(n)$ richtig für alle $n \in \mathbb{N}$.

Der ***Induktionsanfang*** besteht aus $A(1)$ und dessen Beweis. Die ***Induktionsvoraussetzung*** ist „$A(n)$ ist wahr“, die ***Induktionsbehauptung*** ist „$A(n+1)$ ist wahr“. Der ***Induktionsschritt*** ist der Beweis von „$A(n) \Longrightarrow A(n+1)$“. Induktionsbeweise strukturiert man so:

(i) Induktionsanfang ($n = 1$): Beweis von $A(1)$.
(ii) Induktionsschritt ($n \rightsquigarrow n+1$): Beweis von „$A(n) \Longrightarrow A(n+1)$“ d. h., dass aus der Wahrheit von $A(n)$ die Wahrheit von $A(n+1)$ folgt.

Einen Induktionsbeweis könnte man sich als eine unendliche Reihe von Dominosteinen vorstellen, die man zu Fall bringen möchte. Jeder Dominostein steht für ein $A(n)$. Um alle Steine umzustoßen, muss man den ersten Stein umschubsen (Beweis von $A(1)$), und der Stein $A(n+1)$ muss nah genug zu $A(n)$ stehen, so dass jeder Stein durch seinen Vorgänger mit umgerissen wird (Beweis von „$A(n) \Longrightarrow A(n+1)$").

Beispiel: Beweise mit vollständiger Induktion, dass für alle $n \in \mathbb{N}$ gilt:

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

Induktionsanfang ($n = 1$): $1^3 = 1 = \frac{1}{4} 1^2 (1+1)^2$ ist eine wahre Aussage.

Induktionsschritt ($n \rightsquigarrow n+1$):

Induktionsvoraussetzung (IV): Für ein $n \in \mathbb{N}$ gilt $1^3 + 2^3 + \ldots + (n-1)^3 + n^3 = \frac{n^2(n+1)^2}{4}$.

Induktionsbehauptung: Dann gilt $1^3 + 2^3 + \ldots + (n-1)^3 + n^3 + (n+1)^3 = \frac{(n+1)^2(n+2)^2}{4}$.

Beweis des Induktionsschrittes:

$$\begin{aligned} 1^3 + 2^3 + \ldots + (n-1)^3 + n^3 + (n+1)^3 &= [1^3 + 2^3 + \ldots + (n-1)^3 + n^3] + (n+1)^3 \\ &\overset{\text{IV}}{=} \frac{n^2(n+1)^2}{4} + (n+1)^3 = (n+1)^2 \Big[\frac{n^2}{4} + (n+1) \Big] \\ &= \frac{1}{4}(n+1)^2(n^2 + 4n + 4) = \frac{1}{4}(n+1)^2(n+2)^2 . \end{aligned}$$

24.10.2011

1.4.5. **Bemerkung.** Einen Beweis durch Induktion kann man mit einer beliebigen ganzen Zahl anfangen. Sei $q \in \mathbb{Z}$ und $\mathbb{Z}(\geqslant q) = \{n \in \mathbb{Z} : n \geqslant q\}$. Seien $A(n)$ Aussagen, die für $n \in \mathbb{Z}(\geqslant q)$ definiert sind. Wenn

1) $A(q)$ richtig ist und
2) $A(n) \Rightarrow A(n+1)$ für alle $n \in \mathbb{Z}(\geqslant q)$ gilt,

dann gilt $A(n)$ für alle $n \in \mathbb{Z}(\geqslant q)$. Dies folgt leicht aus Satz 1.4.4 durch Indexverschiebung; die Abbildung $f : \mathbb{Z}(\geqslant q) \to \mathbb{N}$, $f(n) = n + q - 1$ ist bijektiv und $f(n+1) = f(n) + 1$.

1.4.6. **Bemerkung.** Das Induktionsprinzip kann auch als Definitionsmittel benutzt werden. Wir definieren zum Beispiel die ganzzahligen Potenzen einer reellen Zahl. Mehr zur Definition durch Induktion finden Sie in §1.9 (Rekursionssatz).

1.4.7. **Definition** (natürlichen Potenzen)**.** Sei $a \in \mathbb{R}$ eine reellen Zahl. Wir definieren die ***natürlichen Potenzen*** a^n, $n \in \mathbb{N}$, durch

$$a^1 = a\,, \quad a^{n+1} = a^n \cdot a\,, \quad n \in \mathbb{N}. \tag{1.2}$$

In der Tat, die Menge $M = \{n \in \mathbb{N} : a^n \text{ ist definiert}\}$ hat die Eigenschaften: $1 \in M$, $n \in M \Rightarrow n+1 \in M$, also $M = \mathbb{N}$, nach Satz 1.4.3.

1.4.8. **Definition** (ganzzahligen Potenzen)**.** Sei $a \in \mathbb{R}$ eine reellen Zahl. Wir definieren die ***ganzzahligen Potenzen*** a^n, $n \in \mathbb{Z}$, durch (1.2) falls $n \in \mathbb{N}$ und

$$a^0 = 1 \,(\text{auch für } a = 0) \tag{1.3}$$

$$a^n = (a^{-1})^{-n} = \left(\frac{1}{a}\right)^{-n}, \quad \text{für } a \neq 0,\ n \in \mathbb{Z},\ n < 0 . \tag{1.4}$$

Man beweist durch Induktion die ***Potenzgesetze***:

$$a^{m+n} = a^m \cdot a^n\,, \quad a^n \cdot b^n = (ab)^n\,, \quad (a^m)^n = a^{mn}$$

für $a, b \in \mathbb{R}$, $m, n \in \mathbb{N}_0$ oder $a, b \in \mathbb{R}^*$, $m, n \in \mathbb{Z}$.

Eine sehr nüztliche Ungleichung, die durch Induktion bewiesen wird, ist die folgende:

1.4.9. **Satz** (Bernoulli-Ungleichung)**.** *Sei $x \in \mathbb{R}$, $x \geqslant -1$ und $n \in \mathbb{N}_0$. Dann gilt*

$$(1+x)^n \geqslant 1 + nx .$$

Ist $n \geqslant 2$, $x \neq 0$ so ist die Ungleichung strikt (>).

Im folgenden führen wir mit Hilfe der rekursiven Definition die Addition und Multiplikation für eine beliebige Anzahl von Summanden bzw. Faktoren ein. Dies erlaubt es, $n!$ und die Binomialkoeffizienten zu definieren.

Zunächst definieren wir den Begriff der Familie.

1.4.10. **Definition.** Seien I und X zwei Mengen. Eine Abbildung $I \to X$, $i \mapsto x_i$ heißt auch eine ***Familie von Elementen*** von X, geschrieben $(x_i)_{i\in I}$. Die Menge I heißt ***Indexmenge***. Ist $I = \{1,2,\dots,n\}$ so heißt $(x_i)_{i\in I} =: (x_1,\dots,x_n) =: (x_i)_{1\leqslant i\leqslant n}$ ein n-Tupel.

1.4.11. **Definition.** Für alle $n \in \mathbb{N}$ und alle n-Tupel $(x_1,\dots,x_n)$ von $\mathbb{R}$ gibt es Elemente $\sum_{i=1}^n x_i$, $\prod_{i=1}^n x_i$ von $\mathbb{R}$, rekursiv definiert durch

$$\sum_{i=1}^{1} x_i := x_1\,,\ \sum_{i=1}^{n+1} x_i := \big(\sum_{i=1}^{n} x_i\big) + x_{n+1}\,, \quad \prod_{i=1}^{1} x_i := x_1\,,\ \prod_{i=1}^{n+1} x_i := \big(\prod_{i=1}^{n} x_i\big) + x_{n+1}\,.$$

Wir schreiben auch $\sum_{i=1}^n x_i = x_1 + \ldots + x_n$.

Sei I eine Indexmenge mit n Elementen, $(x_i)_{i\in I}$ eine Familie von Elementen von $\mathbb{R}$. Sei $I = \{i_1,\dots,i_n\}$ eine Aufzählung von I. Wir setzen

$$\sum_{i\in I} x_i := \sum_{k=1}^{n} x_{i_k} = x_{i_1} + \ldots + x_{i_k}\,, \quad \prod_{i\in I} x_i := \prod_{k=1}^{n} x_{i_k} = x_{i_1} + \ldots + x_{i_k} <,.$$

Die Definition ist unabhängig von der Wahl der Aufzählung.

Eine nützliche Vereinbarung, die es uns oft erlaubt, Fallunterschiedungen zu vermeiden, ist die folgende: Die leere Summe ist 0 und das leere Produkt ist 1, d.h.

$$\sum_{i\in\emptyset} x_i := 0\,, \quad \prod_{i\in\emptyset} x_i := 1. \tag{1.5}$$

Als Beispiel betrachten wir $m,n \in \mathbb{Z}$ und $I = \{k \in \mathbb{Z} : m \leqslant k \leqslant n\} = \{m,\dots,n\}$. Dann setzen wir

$$\sum_{k=m}^{n} x_i := \sum_{i\in I} x_i\,, \quad \prod_{k=m}^{n} x_i := \prod_{i\in I} x_i.$$

Für $m > n$ ist $I = \emptyset$, also gilt nach (1.5): $\sum_{i=m}^n x_i = 0$, $\prod_{i=m}^n x_i = 1$.

1.4.12. **Definition.** Für $n \in \mathbb{N}_0$ wird $n!$ (n ***Fakultät***) definiert durch

$$n! := \prod_{k=1}^{n} k$$

also $0! = 1$, $1! = 1$, und $n! = 1\cdot 2\cdot\ldots\cdot n$ für $n > 1$. Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}_0$ definiert man die ***Binomialkoeffizienten*** „α über k“ durch

$$\binom{\alpha}{k} := \frac{1}{k!}\prod_{j=0}^{k-1}(\alpha - j), \quad \text{also} \quad \binom{\alpha}{0} = \prod_{j=0}^{-1}(\alpha - j) = 1\,, \binom{\alpha}{1} = \alpha,$$

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{k!} \quad \text{für} \quad k > 1\,.$$

Eigenschaften der Binomialkoeffizienten:

(1) Rekursionsformel:

$$\binom{\alpha}{k+1} = \binom{\alpha}{k}\frac{\alpha-k}{k+1}$$

(2) Additionsformel:

$$\binom{\alpha}{k} + \binom{\alpha}{k+1} = \binom{\alpha+1}{k+1};$$

(3) Für $n,k \in \mathbb{N}_0$ gilt

$$\binom{n}{k} = \begin{cases} \dfrac{n!}{k!(n-k)!}\,, & k \leqslant n \\ 0\,, & k > n \end{cases}$$

(4)
$$\binom{n}{k} = \binom{n}{n-k} \quad \text{für} \quad n,k \in \mathbb{N}_0, \quad n \geqslant k\,.$$

1.4.13. **Satz** (Binomischer Lehrsatz)**.** *Für alle $x, y \in \mathbb{R}$, $n \in \mathbb{N}$ gilt*
$$\begin{aligned}(x+y)^n &= \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \ldots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n \\ &= \sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k\,.\end{aligned}$$

Für $n = 2$ und $n = 3$ erhalten wir die wohlbekannten Formel $(x+y)^2 = x^2 + 2xy + y^2$, $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$.

1.5. **Folgerungen des Vollständigkeitsaxioms.**

1.5.1. **Satz** (Satz von Archimedes)**.** $\mathbb{N} \subset \mathbb{R}$ *ist nicht nach oben beschränkt, d.h. zu jedem $x \in \mathbb{R}$ gibt es $n \in \mathbb{N}$ mit $n > x$.*

Dies zeigt auch, dass $\mathbb{N}$ ist unendlich (da jede endliche Menge in $\mathbb{R}$ ein Maximum besitzt). Eine Umformulierung lautet: for alle $x \in \mathbb{R}$, $y \in \mathbb{R}_+$ gibt es $n \in \mathbb{N}$ mit $ny > x$.

27.10.2011

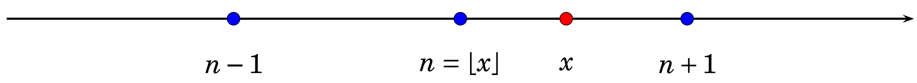
1.5.2. **Satz** (Satz von Eudoxus)**.** *Zu jedem $\varepsilon > 0$ gibt es $n \in \mathbb{N}$ mit $\frac{1}{n} < \varepsilon$.*

Eine Umformulierung des Satzes ist $\inf\{\frac{1}{n} : n \in \mathbb{N}\} = 0$. Der Aussage ist klar für $\varepsilon \geq 1$ (man kann $n = 1$ nehmen). Sie ist interessant für ε "klein". Die griechischen Buchstaben ε und δ (epsilon und delta) werden in der Mathematik oft für kleine positive Zahlen verwendet.

1.5.3. **Satz** (Wohlordnungsprinzip)**.** *Jede nichtleere, nach oben (bzw. unten) beschränkte Teilmenge von $\mathbb{Z}$ besitzt ein Maximum (bzw. Minimum). Jede nichtleere Teilmenge von $\mathbb{N}$ besitzt ein Minimum.*

In der Tat, sei $A \subset \mathbb{Z}$ nach oben beschränkt. Laut Axiom (V) gibt es $s = \sup A \in \mathbb{R}$. Da $s-1$ keine obere Schranke von A ist, existiert $n \in A$ mit $s-1 < n \leq s$, also $s < n+1$. Daher gilt $s < m$ und somit $m \notin A$ für alle $m \geq n+1$. Folglich gilt $a \leq n$ für alle $a \in A$ also n ist das Maximum von A (und natürlich $s = n$).

1.5.4. **Definition** (Gauß-Klammer)**.** Zu jedem $x \in \mathbb{R}$ heißt die ganze Zahl
$$\lfloor x \rfloor := \max\{k \in \mathbb{Z} : k \leqslant x\}$$
die ***Gauß-Klammer*** oder ganzzahliger Anteil von x (d. h. die grösste ganze Zahl, die kleiner oder gleich x ist). Wir setzen $\operatorname{frac}(x) = x - \lfloor x \rfloor$ (fractional part). Ist $x \geq 0$, so ist $\operatorname{frac}(x)$ der Nachkommaanteil von x. Zum Beispiel $\lfloor 3{,}2 \rfloor = 3$, $\operatorname{frac}(3{,}2) = 0{,}2$ und $\lfloor -3{,}2 \rfloor = -4$, $\operatorname{frac}(-3{,}2) = 0{,}8$.



1.5.5. **Satz** (Dichtheit von $\mathbb{Q}$ in $\mathbb{R}$)**.** *Seien $a, b \in \mathbb{R}$, $a < b$. Dann gibt es $q \in \mathbb{Q}$ mit $a < q < b$.*

Seien $A \subset B \subset \mathbb{R}$. Wir sagen, dass A liegt dicht in B, falls für alle $x, y \in B$, $x < y$, gibt es $z \in A$ mit $x < z < y$. Satz 1.5.5 besagt, dass $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt.

Die Idee des Beweises kann man folgendermassen beschreiben. Ein Männchen läuft auf der reellen Gerade mit konstanter Schrittweite $\frac{1}{n}$ mit Startpunkt links von a und durchläuft Punkte der Form $\frac{p}{n}$. Ist nun die Schrittweite kleiner als der Länge des Intervalls (a,b) d. h. $\frac{1}{n} < b-a$, dann wird das Männchen irgendwann in (a,b) fallen, d. h. es gibt $m \in \mathbb{N}$, so dass nach m Schritte $a < \frac{m}{n} < b$.

Formal argumentieren wir so: Wähle $n \in \mathbb{N}$ mit $\frac{1}{n} < b - a$ (Satz 1.5.2). Dann gilt $na + 1 < nb$. Aber $\lfloor na \rfloor \le na < \lfloor na \rfloor + 1$, also $na < \lfloor na \rfloor + 1 \le na + 1 < nb$. Setze $m = \lfloor na \rfloor + 1$.

Das zeigt auch , dass eine reelle Zahl x lässt sich mit rationalen Zahlen so approximieren:

$$\left| x - \frac{\lfloor nx \rfloor}{n} \right| < \frac{1}{n} \cdot$$

Der Fehler bei dieser Approximation ist $\frac{1}{n}$, also kann beliebig klein gewählt werden. Dafür muss man aber den Nenner n entsprechend gross wählen. Kann man reelle Zahlen mit rationale Zahlen (Brüche) gut approximieren indem man den Nenner möglichst klein wählt? Zum Beispiel für $\pi = 3.1415926\ldots$ gilt $|\pi - \frac{22}{7}| < \frac{1}{100}$ ·

1.5.6. **Satz** (Approximationssatz von Dirichlet). *Sei $x \in \mathbb{R}$ und sei $N \in \mathbb{N}$, $N > 1$. Dann existieren teilfremde ganze Zahlen p und q mit $1 \le q < N$, so dass*

$$\left| x - \frac{p}{q} \right| \le \frac{1}{qN} \cdot$$

Beweis: Betrachte die $N+1$ Zahlen $0, \mathrm{frac}(x), \mathrm{frac}(2x), \ldots, \mathrm{frac}((N-1)x), 1$ in Intervall $[0,1]$. Teilen wir $[0,1]$ in N Intervalle der Länge $\frac{1}{N}$. Nach dem Schubfachprinzip (Satz 1.9.2) müssen zwei der obigen Zahlen in derselben Intervall liegen. Da $kx = \lfloor kx \rfloor + \mathrm{frac}(kx)$ ist ihre Differenz der Form $qx - p$, wobei $1 \le q < N$. Also $|qx - p| \le \frac{1}{N}$. Wenn wir mit $\mathrm{ggT}(p,q)$ dividieren, erreichen wir auch die Bedingung, dass p und q teilerfremd sind. □

1.5.7. **Satz** (Satz über Division mit Rest). *Seien $a, b \in \mathbb{Z}$, $b \neq 0$. Dann gibt es $q, r \in \mathbb{Z}$, $0 \leqslant |r| < |b|$ so dass $a = bq + r$, nämlich $q := \lfloor a/b \rfloor$ und $r := a - bq$.*

Mehr zum Wohlordnungsprinzip finden Sie in §1.9.

Eine weitere wichtige Anwendung des Vollständigkeitsaxioms ist die Existenz der Wurzeln.

1.5.8. **Satz** (Existenz der k-ten Wurzel). *Sei $a \in \mathbb{R}$, $a > 0$, und sei $k \in \mathbb{N}$. Dann gibt es genau ein $x \in \mathbb{R}$ mit $x > 0$ und $x^k = a$, genannt k-te Wurzel von a, geschrieben $x =: \sqrt[k]{a} =: a^{1/k}$.* 31.10.2011

In (1.2) und (1.3) haben wir die ganzzahligen Potenzen definiert.

1.5.9. **Definition** (rationale Potenzen). Sei $a > 0$. Für $n \in \mathbb{N}_0$ sei $a^{-n} := (a^{-1})^n$, und sei

$$a^{\frac{k}{m}} := \sqrt[m]{a^k}, \quad \text{für } k \in \mathbb{Z},\, m \in \mathbb{N} \tag{1.6}$$

definiert. Damit ist also a^q für jedes $q \in \mathbb{Q}$ erklärt.

Die Definition (1.6) hängt von der Darstellung $q = k/m$ der rationalen Zahl q nicht ab (siehe Aufgabe 1.8.12). Dann gelten die ***Potenzgesetze***:

$$a^q a^r = a^{q+r}, \quad (a^q)^r = a^{qr}, \quad \text{für } a > 0 \text{ und } q, r \in \mathbb{Q}. \tag{1.7}$$

Sei $(x_1, \ldots, x_n)$ ein n-Tupel von reellen Zahlen. Die Zahl

$$\frac{1}{n} \sum_{k=1}^{n} x_k = \frac{1}{n}\big(x_1 + \ldots + x_n\big)$$

heißt das ***arithmetische Mittel*** von $(x_1, \ldots, x_n)$. Sind alle $x_k \geqslant 0$, so heißt

$$\Big(\prod_{i=1}^{n} x_i\Big)^{\frac{1}{n}} = \sqrt[n]{x_1 \cdots x_n}$$

das ***geometrische Mittel*** von $(x_1, \ldots, x_n)$. Es gilt die ***AGM-Ungleichung*** (wobei A=arithmetisches, G=geometrisches, M=Mittel):

$$\frac{1}{n}\big(x_1 + \ldots + x_n\big) \geqslant \sqrt[n]{x_1 \cdots x_n},, \quad \text{für alle } x_1, \ldots, x_n \geqslant 0, \tag{1.8}$$

wobei die Gleichheit genau dann auftritt, wenn $x_1 = x_2 = \ldots = x_n$. Es gibt viele Beweise dieser wichtigen Ungleichung; für ein Beweis durch Induktion siehe Aufgabe 1.8.5 (in der Großübung behandelt) oder [24, 3.7]. Wir werden (1.8) im Satz 8.1.10 beweisen, als Konsequenz der Jensenschen Ungleichung für konvexe Funktionen. Für $n = 2$ lautet (1.8): $x + y \geqslant 2\sqrt{xy}$ für alle $x, y \geqslant 0$, woraus man (1.1) wiederfindet.

Eine Anwendung der AGM-Ungleichung ergibt:

$$\left(1+\frac{x}{n}\right)^n < \left(1+\frac{x}{n+1}\right)^{n+1}, \quad \text{für } n \in \mathbb{N},\, x \geqslant -n,\, x \neq 0\,. \tag{1.9}$$

1.5.10. **Definition** (Intervalle)**.** Eine Teilmenge $I \subset \mathbb{R}$ heißt ***Intervall***, wenn I mit je zwei Elementen $x \leqslant y$ auch alle Elemente z mit $x \leqslant z \leqslant y$ enthält.

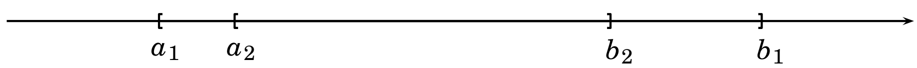
1.5.11. **Beispiel.** Seien $a, b \in \mathbb{R}$. Dann ist $[a,b] := \{x \in K : a \leqslant x \leqslant b\}$ ein Intervall. Falls $a = b$ ist $[a,b] = \{a\}$, falls $b < a$ ist $[a,b] = \emptyset$. Andere Beispiele:

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

Intervalle der Form (a,b) heißen ***offen***, $(a,b]$, $[a,b)$ halboffen. Die Intervalle $[a,b]$ heißen ***abgeschlossen***. Das Intervall $[a,b]$ heißt auch ***kompakt***. Später werden wir allgemeine offene, abgeschlossene und kompakte Mengen definieren.

1.5.12. **Definition** (Intervallschachtelung)**.** Für ein Intervall I mit Endpunkten $a, b \in \mathbb{R}$ definieren wir die ***Länge*** von I als $|I| := |b-a|$. Eine Familie $(I_n)_{n\in\mathbb{N}}$ von kompakten Intervallen heißt ***Intervallschachtelung*** , falls gilt:

(i) $I_{n+1} \subset I_n$ für alle $n \in \mathbb{N}$, und
(ii) für alle $\varepsilon > 0$ gibt es $n \in \mathbb{N}$ mit $|I_n| < \varepsilon$.



1.5.13. **Satz** (Intervallschachtelungsprinzip)**.** *Zu jeder Intervallschachtelung $(I_n)_{n\in\mathbb{N}}$ gibt es $x \in \mathbb{R}$ mit $\bigcap_{n\in\mathbb{N}} I_n = \{x\}$.*

Der Satz gilt nicht, wenn wir nicht fordern, dass die I_n abgeschlossen sind. Für $I_n = (0, \frac{1}{n})$ gelten beide Eigenschaften (i) und (ii) aus 1.5.12, aber $\cap_{n\in\mathbb{N}} I_n = \emptyset$.

1.5.14. **Satz.** *Sei $I_n = \left[\left(1+\frac{1}{n}\right)^n, \left(1+\frac{1}{n}\right)^{n+1}\right]$, $n \in \mathbb{N}$. Dann ist $(I_n)_{n\in\mathbb{N}}$ eine Intervallschachtelung. Die Zahl $e \in \mathbb{R}$ mit mit $\cap_{n\in\mathbb{N}} I_n = \{e\}$ heißt **Eulersche Zahl**.*

Es gilt $e = 2{,}71828\ldots$; e ist eine irrationale Zahl. Die Zahl e spielt eine wichtige Rolle in der Mathematik, sie ist die Basis des natürlichen Logarithmus und der natürlichen Exponentialfunktion. Man kann die Endpunkte der Intervalle in einer Intervallschachtelung als Approximationen des Punktes in der Durchschnitt. In diesen Fall gilt $I_1 = [2,4]$, $I_2 = [2\frac{1}{4}, 3\frac{3}{8}]$, $I_3 = [\frac{64}{27}, \frac{256}{81}] = [2\frac{10}{27}, 3\frac{16}{81}], \ldots$

1.6. **Die Überabzählbarkeit von $\mathbb{R}$.**

1.6.1. **Definition.**

(i) Zwei Mengen X, Y heißen ***gleichmächtig*** (geschrieben $X \sim Y$), wenn es eine bijektive Abbildung $\varphi : X \longrightarrow Y$ gibt. Die Relation $X \sim Y$ ist eine Äquivalenzrelation.

(ii) Eine Menge X heißt ***endlich***, falls $X = \emptyset$ oder es ein $n \in \mathbb{N}$ gibt mit $X \sim \{1, \ldots, n\}$. Die ***Kardinalzahl*** $|X|$ (Anzahl der Elemente von X) ist definiert durch $|X| = 0$ falls $X = \emptyset$ und $|X| = n$ falls $X \sim \{1, \ldots, n\}$.

(iii) Eine Menge heißt ***unendlich***, wenn sie nicht endlich ist.

(iv) Eine Menge X heißt ***abzählbar***, falls $X \sim \mathbb{N}$ und ***höchstens abzählbar***, wenn sie endlich oder abzählbar ist. Ist X höchstens abzählbar, so heißt eine Bijektion $f : \{1,\dots,n\} \to X$ oder $f : \mathbb{N} \to X$ ***Aufzählung*** von X; dann ist $X = \{x_1,\dots,x_n,\dots\}$, $x_k = f(k)$. Ist X nicht höchstens abzählbar, so heißt X ***nicht abzählbar*** oder ***überabzählbar***.

1.6.2. **Beispiele.** Die Funktion $f : \mathbb{Z} \to \mathbb{N}$, $f(z) = 2z$ falls $z \geqslant 0$ und $f(z) = -2z-1$ falls $z < 0$ ist bijektiv. Daraus folgt, dass die Menge $\mathbb{Z}$ abzählbar ist; dies ist schon erstaunlich, weil es bedeutet, dass $\mathbb{N}$ und $\mathbb{Z}$ gleich viele Elemente haben, obwohl $\mathbb{N}$ eine echte Teilmenge von $\mathbb{Z}$ ist! Die Mengen $\mathbb{Q}_+$, $\mathbb{Q}$ sind auch abzählbar!

Wir können dagegen die reellen Zahlen nicht abzählen:

1.6.3. **Satz** (Cantor)**.** *Die Menge $\mathbb{R}$ ist nicht abzählbar.*

7.11.2011

1.7. **Der Körper der komplexen Zahlen.** In $\mathbb{R}$ gibt es keine Zahl x mit $x^2 = -1$, da $x^2 \geqslant 0 > -1$ für alle $x \in \mathbb{R}$. Wir konstruieren einen Körper, in dem die Gleichung $x^2 = -1$ eine Lösung hat. Auf $\mathbb{R}^2$ führen wir eine Addition und Multiplikation wie folgt ein:

$$\begin{aligned}(x,y)+(u,v) &:= (x+u, y+v)\\ (x,y)\cdot(u,v) &:= (xu-yv, xv+yu).\end{aligned} \tag{1.10}$$

1.7.1. **Satz.** $(\mathbb{R}^2,+,\cdot)$ *ist ein Körper mit Nullelement* $(0,0)$ *und Einselement* $(1,0)$*. Dieser Körper heißt **Körper der komplexen Zahlen**, bezeichnet mit* $\mathbb{C} := (\mathbb{R}^2,+,\cdot)$.

1.7.2. **Satz.** *Die Abbildung $\varphi : \mathbb{R} \longrightarrow \mathbb{C}$, $\varphi(x) = (x,0)$ hat die Eigenschaften $\varphi(x+y) = \varphi(x)+\varphi(y)$, $\varphi(xy) = \varphi(x)\varphi(y)$, $\varphi(1) = (1,0)$, d.h. φ ist ein **Körper-Homomorphismus**.*

Die komplexen Zahlen $\{(x,0) : x \in \mathbb{R}\}$ bilden einen Körper mit der induzierten Addition und Multiplikation (1.10). Wir sagen, dass $\{(x,0) : x \in \mathbb{R}\}$ ein ***Unterkörper*** von $\mathbb{C}$ ist. Der Homomorphismus $\varphi : \mathbb{R} \longrightarrow \{(x,0) : x \in \mathbb{R}\}$ ist bijektiv, d.h. ein ***Körper-Isomorphismus***. Wir *identifizieren* deshalb $\mathbb{R}$ mit $\{(x,0) : x \in \mathbb{R}\}$ und sagen, dass $\mathbb{R}$ ein Unterkörper von $\mathbb{C}$ ist.

Wir schreiben für $(x,0)$ kurz x, also 0 für $(0,0)$, 1 für $(1,0)$, usw.

1.7.3. **Definition.** Die (nicht-reelle) Zahl $i = (0,1)$ heißt ***imaginäre Einheit***. Es gilt

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

Für $z = (x,y)$ schreiben wir nun

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

Dann heißt x ***Realteil*** von z, und y heißt ***Imaginärteil*** von z, geschrieben $\operatorname{Re} z := x$, $\operatorname{Im} z := y$. Man beachte, dass der Imaginärteil y reell ist. Zahlen der Form iy mit $y \in \mathbb{R}$ heißen auch (rein) imaginär.

1.7.4. **Definition.** Die ***konjugierte Zahl*** zu $z = x+iy$ ist $\overline{z} := x-iy$.

1.7.5. **Satz** (Rechenregeln)**.** *Für alle $z,w \in \mathbb{C}$ gilt:*

(i) $\overline{z+w} = \overline{z}+\overline{w}$, $\overline{z\cdot w} = \overline{z}\cdot\overline{w}$.
(ii) $z+\overline{z} = 2\operatorname{Re} z$, $z-\overline{z} = 2i\operatorname{Im} z$.
(iii) $z = \overline{z} \Leftrightarrow z \in \mathbb{R}$, $z = -\overline{z} \Leftrightarrow z \in i\mathbb{R}$.
(iv) $z\cdot\overline{z} = x^2+y^2 \geqslant 0$ *für* $z = x+iy$.
(v) $\overline{\overline{z}} = z$.

1.7.6. **Definition.** Für $z \in \mathbb{C}$ heißt $|z| := \sqrt{z\cdot\overline{z}} = \sqrt{x^2+y^2}$ der ***Betrag*** von z. Für $z \in \mathbb{R}$ ist $|z| = \sqrt{z^2}$ der übliche Betrag von reellen Zahlen.

1.7.7. **Satz** (Rechenregeln für den Betrag)**.** *Für alle $z,w \in \mathbb{C}$ gilt:*

(i) $|z| \geqslant 0$; $|z| = 0 \Leftrightarrow z = 0$.
(ii) *Ist $z \neq 0$, so gilt* $z^{-1} = \overline{z}/|z|^2$.

(iii) $|\overline{z}| = |z|$.

(iv) $|\mathrm{Re}\, z| \leqslant |z|$, $|\mathrm{Im}\, z| \leqslant |z|$.

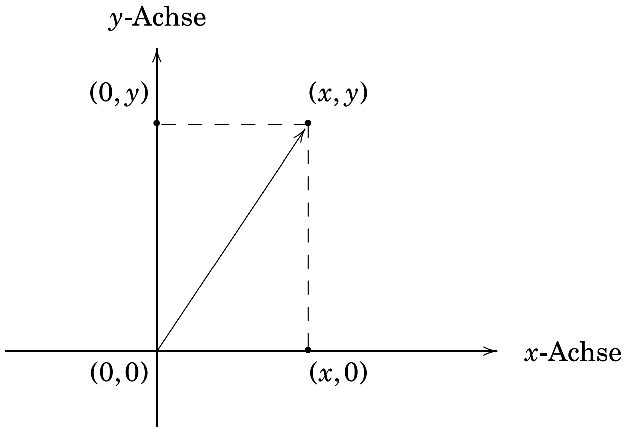
(v) $|z \cdot w| = |z| \cdot |w|$.

(vi) $|z + w| \leqslant |z| + |w|$ (***Dreiecksungleichung***).
Die Gleichheit gilt genau dann, wenn $z = 0$ (*bzw.* $w = 0$) *oder* $w/z \in \mathbb{R}_+$ (*bzw.* $z/w \in \mathbb{R}_+$).

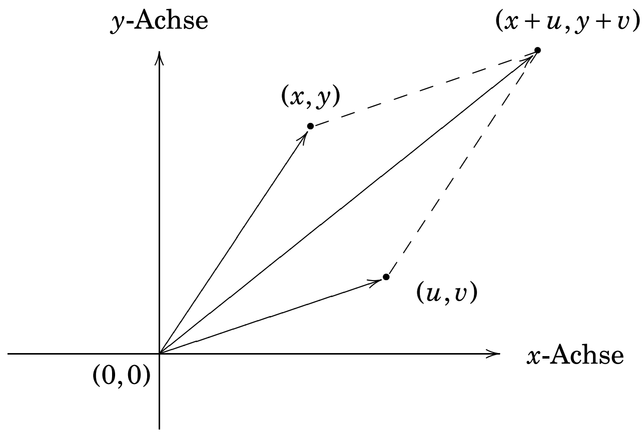
(vii) $\big||z| - |w|\big| \leqslant |z - w|$ (***umgekehrte Dreiecksungleichung***).

Geometrische Deutung der komplexen Zahlen:

Man veranschaulicht sich seit Gauß die komplexen Zahlen als Punkte in der ***Gaußschen Zahlenbene*** mit rechtwinkligen Koordinaten (oder als Vektoren mit Ursprung im Nullpunkt $(0,0)$ und Endpunkt in (x, y)).



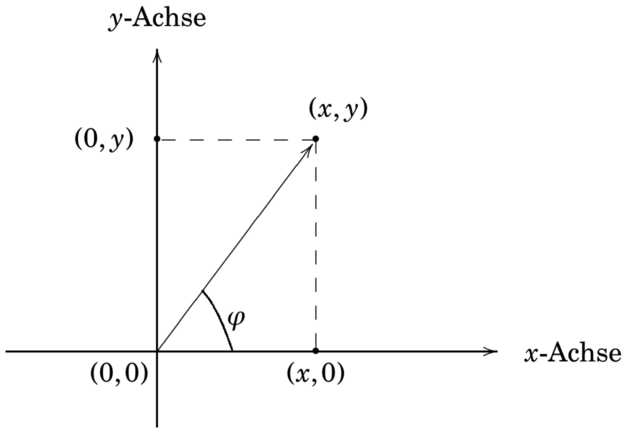
Die Addition komplexer Zahlen ist dann die übliche Vektoraddition nach der Parallelogrammregel.



Um die Multiplikation zu interpretieren, führen wir Polarkoordinaten ein. Das ist nicht ganz elementar und wird erst später gerechtfertigt. (Wir benötigen Eigenschaften der Cosinus- und Sinusfunktion, die zwar wohlbekannt sind, deren Beweise aber tiefer liegen; siehe §5.5 insbesondere die Sätze 8.4.9 und 8.4.11.) Wir schreiben jeden Punkt (x, y) als $r(\cos\varphi, \sin\varphi)$, wobei $r = \sqrt{x^2 + y^2}$ die Länge des Vektors (x, y) ist und φ der Winkel zwischen der positiven x-Achse und dem Vektor (x, y). Die Gleichung

$$\boxed{z = r(\cos\varphi + i \sin\varphi)} \tag{1.11}$$

heißt eine ***Polarkoordinatendarstellung*** oder *trigonometrische Form* von z, die Zahlen r, φ heißen ***Polarkoordinaten von*** z. Diese Darstellung ergibt sich leicht, wenn man das rechtwinklige Dreieck mit den Ecken $(0,0)$, $(x,0)$, (x, y) betrachtet. Wir nennen φ ein ***Argument*** von (x, y). Nach der Definition ist $\varphi \in [0, 2\pi)$, wir können aber auch alle $\varphi + 2k\pi$ mit $k \in \mathbb{Z}$ als Argument von (x, y) betrachten.



Nach der Definition der Multiplikation gilt

$$\begin{aligned}(x,y)\cdot(u,w) &= rs(\cos\varphi\cos\psi - \sin\varphi\sin\psi, \sin\varphi\cos\psi + \cos\varphi\sin\psi)\\ &= rs(\cos(\varphi+\psi), \sin(\varphi+\psi))\end{aligned}$$

oder

$$\big[r(\cos\varphi + i\sin\varphi)\big]\cdot\big[s(\cos\psi + i\sin\psi)\big] = rs\big(\cos(\varphi+\psi) + i\sin(\varphi+\psi)\big),$$

wobei man die Additionssätze für cos und sin benutzt. Die Multiplikationsregel für komplexe Zahlen lautet also: „Die Längen werden multipliziert, die Argumente werden addiert“. Insbesondere erhalten wir die ***Formel von de Moivre***:

$$\boxed{\big[r(\cos\varphi + i\sin\varphi)\big]^n = r^n(\cos n\varphi + i\sin n\varphi), \quad n\in\mathbb{Z}.}$$

1.8. **Übungen.**

1.8.1. ***Aufgabe*** (Endliche geometrische Reihe). Sei $a\neq 1$. Dann gilt

$$1 + a + \ldots + a^n = \frac{1-a^{n+1}}{1-a}$$

1.8.2. ***Aufgabe***. Sei $n\in\mathbb{N}$. Dann gelten die Identitäten:

$$\begin{aligned}1+2+\ldots+n &= \frac{n(n+1)}{2}\\ 1^2+2^2+\ldots+n^2 &= \frac{n(n+1)(2n+1)}{6}\\ 1^3+2^3+\ldots+n^3 &= \frac{n^2(n+1)^2}{4}\end{aligned}$$

Sei $S_{n,k} = 1^k + 2^k + \ldots + n^k$, $k\in\mathbb{N}_0$. Dann gilt:

$$(n+1)^k - n^k = \binom{k}{1}S_{n,k-1} + \binom{k}{2}S_{n,k-2} + \ldots + \binom{k}{k}S_{n,0}.$$

Diese Formel stellt eine Rekursionsformel für $S_{n,k}$ dar, die erlaubt, alle $S_{n,k}$ zu berechnen.

1.8.3. ***Aufgabe*** (Teleskopsummen). Sei $n\in\mathbb{N}$. Dann gilt:

$$\begin{aligned}&\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \ldots + \frac{1}{n(n+1)} = \frac{n}{n+1},\\ &\frac{1}{1\cdot 2\cdot 3} + \frac{1}{2\cdot 3\cdot 4} + \ldots + \frac{1}{n(n+1)(n+2)} = \frac{1}{2}\left[\frac{1}{2} - \frac{1}{(n+1)(n+2)}\right],\\ &\frac{1}{2!} + \frac{2}{3!} + \ldots + \frac{n-1}{n!} = 1 - \frac{1}{n!}\cdot\end{aligned}$$

1.8.4. ***Aufgabe***. Beweise:

a) $\sum\limits_{k=0}^{n}\binom{n}{k} = 2^n$, b) $\sum\limits_{k=0}^{n} k\binom{n}{k} = n\cdot 2^{n-1}$, c) $\sum\limits_{k=0}^{n}(-1)^k\binom{n}{k} = 0$, d) $\sum\limits_{k=0}^{n}\frac{(-1)^k}{k+1}\binom{n}{k} = \frac{1}{n+1}\cdot$

1.8.5. ***Aufgabe***. Beweise die AGM-Ungleichung (1.8).

1.8.6. ***Aufgabe***. (a) Seien $a_1, \ldots, a_n > 0$. Zeige:

$$\Big(\sum_{k=1}^{n} a_k\Big)\Big(\sum_{k=1}^{n} \frac{1}{a_k}\Big) \geq n^2 \text{ mit Gleichheit genau dann, wenn } a_1 = \ldots = a_n$$

(b) Sei $a \in \mathbb{R}_+ \setminus \{1\}$, $n \in \mathbb{N}$, $n \geq 2$. Zeige

$$\sqrt[n]{a} < 1 + \frac{a-1}{n} \qquad \text{und} \qquad \sqrt[n]{a^p} < 1 + \frac{p}{n}(a-1) \quad \text{für} \quad p \in \mathbb{N}, 1 \leq p < n\,.$$

1.8.7. ***Aufgabe***. Seien $n \in \mathbb{N}$ und $a_1, \ldots, a_n, b_1 \ldots, b_n$ reelle Zahlen.
(a) Beweise die Lagrange-Identität:

$$\sum_{1 \leqslant i < j \leqslant n} (a_i b_j - a_j b_i)^2 = \left(\sum_{i=1}^{n} a_i^2\right) \cdot \left(\sum_{i=1}^{n} b_i^2\right) - \left(\sum_{i=1}^{n} a_i b_i\right)^2. \tag{1.12}$$

(b) Beweise die Cauchy-Schwarzsche Ungleichung:

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leqslant \left(\sum_{i=1}^{n} a_i^2\right) \cdot \left(\sum_{i=1}^{n} b_i^2\right).$$

Zeige, daß die Gleichheit genau dann auftritt, wenn es $\lambda \in \mathbb{R}$ gibt, so daß $b_i = \lambda a_i$ für alle $i = 1, \ldots, n$ oder $a_i = \lambda b_i$ für alle $i = 1, \ldots, n$. Soll bedeuten:
Gleichheit genau dann, wenn $(a_1, \ldots, a_n)$ und $(b_1, \ldots, b_n)$ *linear abhängig* sind.

1.8.8. ***Aufgabe***. Zeige durch direkte Rechnung, daß für $m, n \in \mathbb{N}$ mit $m < n$ gilt:
(a) $\binom{m}{k} < \binom{n}{k}$ für $k \in \{1, \ldots, n\}$,
(b) $\frac{1}{m^k}\binom{m}{k} < \frac{1}{n^k}\binom{n}{k} < \frac{1}{k!} \leq \frac{1}{2^{k-1}}$ für $k \in \{2, \ldots, n\}$,
(c) $\left(1 + \frac{1}{m}\right)^m < \left(1 + \frac{1}{n}\right)^n$. Tip: Benutze den binomischen Lehrsatz und (b).

1.8.9. ***Aufgabe***. (a) Wegen der Bernoulli-Ungleichung gilt $\left(1 + \frac{1}{n}\right)^n \geqslant 2$ für alle $n \in \mathbb{N}$.
Zeige, daß außerdem gilt: $\left(1 + \frac{1}{n}\right)^n < 3$ für alle $n \in \mathbb{N}$. (Tip: Binomischer Lehrsatz, endliche geometrische Reihe und Aufgabe 1.8.8.)
(b) Zeige für alle $n \in \mathbb{N}$ mit $n \geqslant 2$: $\prod_{k=1}^{n-1}\left(1 + \frac{1}{k}\right)^k = \frac{n^n}{n!}$.
(c) Folgere aus (a) und (b): $3\left(\frac{n}{3}\right)^n \leqslant n! \leqslant 2\left(\frac{n}{2}\right)^n$ für alle $n \in \mathbb{N}$.

1.8.10. ***Aufgabe***. Beweise durch Induktion nach k, dass:

$$\left(1 + \frac{1}{n}\right)^k < 1 + \frac{k}{n} + \frac{k^2}{n^2}, \quad n, k \in \mathbb{N}, k \leq n.$$

Folgere, dass $\left(1 + \frac{1}{n}\right)^n < 3$ für alle $n \in \mathbb{N}$.

1.8.11. ***Aufgabe***. Bestimme für jede der folgenden Mengen, ob sie ein Supremum bzw. Infimum besitzt. Falls ja, berechne es und entscheide, ob es in der jeweiligen Menge enthalten ist.

(a) $M_1 = \left\{(-1)^n \left(2 + \frac{3}{n}\right) \mid n \in \mathbb{N}\right\}$;
(b) $M_2 = \left\{\left(-\frac{1}{3}\right)^m + \frac{5}{n} \mid m, n \in \mathbb{N}\right\}$;
(c) $M_3 = \{x \in \mathbb{R} \mid (x+a)(x+b)(x+c) > 0\}$, wobei $a < b < c$ fest.

1.8.12. ***Aufgabe***. (a) Zeige: Für $a > 1$ ist $a > \sqrt{a} > \sqrt[3]{a} > \ldots > 1$, und für $0 < a < 1$ ist $a < \sqrt{a} < \sqrt[3]{a} < \ldots < 1$.
(b) Sei $x > 0$. Für $n \in \mathbb{N}_0$ sei $x^{-n} := (x^{-1})^n$, und für $k \in \mathbb{Z}$, $m \in \mathbb{N}$ sei $x^{\frac{k}{m}} := \sqrt[m]{x^k}$ definiert. Damit ist also x^q für jedes $q \in \mathbb{Q}$ erklärt. (warum eindeutig?)
Zeige für $x > 0$ und $q, r \in \mathbb{Q}$: $\quad x^q x^r = x^{q+r}$ und $(x^q)^r = x^{qr}$.

1.8.13. ***Aufgabe***. Für positive Zahlen a, b definiert man das arithmetische, geometrische und harmonische Mittel durch

$$A(a,b) := \frac{a+b}{2}\,, \quad G(a,b) := \sqrt{ab}\,, \quad H(a,b) := \frac{1}{A(\frac{1}{a}, \frac{1}{b})} = \frac{2ab}{a+b}\,.$$

(i) Beweise die Ungleichungen

$$H(a,b) \leqslant G(a,b) \leqslant A(a,b)$$

und zeige, daß eine Gleichheit der Mittel nur für $a = b$ eintritt.
(ii) Es sei $0 < a < b$. Man definiere Intervalle $[a_n, b_n]$ für alle $n \in \mathbb{N}$, rekursiv durch $[a_1, b_1] := [a, b]$ sowie durch $a_{n+1} := G(a_n, b_n)$ und $b_{n+1} := A(a_n, b_n)$. Man zeige, daß sie eine Intervallschachtelung bilden. Man zeige ferner die Abschätzung

$$b_{n+1} - a_{n+1} \leqslant \frac{1}{8a}(b_n - a_n)^2\,.$$

Die in allen Intervallen $[a_n, b_n]$ liegende Zahl heißt **arithmetisch–geometrisches Mittel** der Zahlen a und b und wird mit $M(a,b)$ bezeichnet.

1.8.14. ***Aufgabe***. Seien $a, b > 0, a^2 \neq b$. Setze

$$a_0 = \frac{b}{a}\,, \quad b_0 = a\,, \quad b_n = \frac{a_{n-1} + b_{n-1}}{2}\,, \quad a_n = \frac{b}{b_n}\,, \quad \text{für } n \in \mathbb{N}\,.$$

(i) Zeige induktiv: $a_n < b_n$, $n \in \mathbb{N}$.
(ii) Zeige, dass $([a_n, b_n])_{n \in \mathbb{N}}$ eine Intervallschachtelung ist, in deren Durchschnitt $\sqrt{b}$ liegt.

1.8.15. ***Aufgabe***. (a) Bestimme Betrag, Real- und Imaginärteil von

$$\text{(i)} \quad \frac{3+4i}{2-i} \quad , \quad \text{(ii)} \quad (1+i)^8 \quad , \quad \text{(iii)} \quad \left(\frac{1+i}{1-i}\right)^n, \; n \in \mathbb{Z} \quad .$$

(b) Seien $z, w \in \mathbb{C}$. Dann gilt
(i) $|z+w|^2 + |z-w|^2 = 2(|z|^2 + |w|^2)$;
(ii) $z, w \neq 0$: $|z+w| = |z| + |w| \iff z/w > 0$.

1.8.16. ***Aufgabe***. Sei $g := \frac{1}{2}(1+\sqrt{5})$ (die Zahl des goldenen Schnittes) und $h := g^{-1}$. Man rechnet leicht nach: $g : 1 = (1+g) : g$, $g^2 = 1 + g$, $h^2 = 1 - h$, $g = 1 + h$.
(a) Sei $\zeta := \frac{1}{2}\left(h + i\sqrt{4-h^2}\right)$. Zeigen Sie, dass ζ eine Lösung der Gleichung $z^5 - 1 = 0$ (also eine *5. Einheitswurzel*) ist. Tipp: Es gilt $z^4 + z^3 + z^2 + z + 1 = (z^2 + gz + 1)(z^2 - hz + 1)$.
(b) Zeigen Sie, dass $1, \zeta, \zeta^2, \zeta^3, \zeta^4$ in $\mathbb{C}$ die Ecken eines regelmäßigen Fünfecks mit Mittelpunkt 0 bilden, und dass $|\zeta^2 - 1| : |\zeta - 1| = g$ gilt. Was bedeutet letzteres geometrisch?

1.9. **Notizen.** Wir sammeln in diesen Abschnitt Kommentare über die Definition durch Induktion, Wohlordnungsprinzip und Kardinalzahlen.
Notiz zum Rekursionssatz. Ein weiteres Beispiel von Definition durch Induktion ist die Definition der Fibonacci-Folge $(u_n)_{n \in \mathbb{N}_0}$, wobei $u_0 = u_1 = 1$ und $u_2 = u_0 + u_1$, $u_3 = u_1 + u_2$ und so weiter. Dieses „und so weiter“ basiert auf einem allgemeinen Satz, dem *Rekursionssatz*, welcher es erlaubt, gewisse Objekte induktiv (rekursiv) zu definieren. Vergleiche dazu [9, Kap. 16], [6, p. 15].

Satz (Rekursionssatz). *Sei a ein Element einer Menge X und f eine Funktion von X in X. Dann gibt es eine Funktion u von $\mathbb{N}$ in X so, dass $u(0) = a$ und $u(n+1) = f(u(n))$ für alle n in $\mathbb{N}$.*

Eine Anwendung des Rekursionssatzes nennen wir *rekursive Definition* oder *Definition durch vollständige Induktion*. Für die Fibonacci-Folge müssen wir allerdings eine allgemeinere Version des Satzes anwenden [24, § 2.8]. Wir betrachten $X = \mathbb{N}$ und $f : X \times X \to X$, $f(x,y) = x + y$ und setzen $u_0 = u_1 = 1$, $u_{n+2} = f(u_n, u_{n+1})$ für $n \geqslant 2$.

Notiz zum Wohlordungsprinzip. Das Wohlordnungsprinzip hat viele anderen Anwendungen. Für $a, b \in \mathbb{Z}$ sagen wir, dass a teilt b falls ein $c \in \mathbb{Z}$ existiert, mit $b = ac$. Wir schreiben $a \mid b$. Seien $a, b \in \mathbb{Z}$, nicht beide Null. Setze

$$\mathrm{ggT}(a,b) := (a,b) := \max\{d \in \mathbb{N} : d \mid a,\, d \mid b\} \quad \text{(größter gemeinsamer Teiler von } a,b)$$

Seien $a, b \in \mathbb{Z}$, keine Null. Setze

$$\mathrm{kgV}(a,b) := [a,b] := \min\{v \in \mathbb{N} : a \mid v,\, b \mid v\} \quad \text{(kleinstes gemeinsames Vielfaches von } a,b)$$

Die Existenz ist von dem Wohlordnungsprinzip gesichert. Wir beschreiben kurz das *Euklidische Algorithmus*. Es geht darum, den größten gemeinsamen Teiler zweier ganzen Zahlen zu bestimmen. Er basiert auf der Division mit Rest.

Satz (Euklidischer Algorithmus). *Seien $a, b \in \mathbb{Z}$, $b \neq 0$. Betrachte die Folge $r_0 = a$, $r_1 = b$, $r_{i+1} =$ Rest von r_{i-1} bei Division durch r_i, falls $r_i \neq 0$, 0 sonst. Dann gibt es einen kleinsten Index $n \in \mathbb{N}$ mit $r_{n+1} = 0$ und es gilt $r_n = \mathrm{ggT}(a,b)$:*

$$\begin{aligned}
a &= b\,q_1 + r_2\,, & |r_2| &< |b| = |r_1|\,,\\
b &= r_2\,q_2 + r_3\,, & |r_3| &< |r_2|\,,\\
r_2 &= r_3\,q_3 + r_4\,, & |r_4| &< |r_3|\,,\\
&\;\vdots & &\;\vdots\\
r_{n-2} &= r_{n-1}\,q_{n-1} + r_n\,, & |r_n| &< |r_{n-1}|\,,\\
r_{n-1} &= r_n\,q_n\,.
\end{aligned}$$

Eine Zahl $p \in \mathbb{Z}$ heißt ***Primzahl***, falls aus $p|ab$ stets $p|a$ oder $p|b$ folgt. Eine Zahl $p \in \mathbb{Z}$ heißt ***irreduzibel*** wenn $|p| > 1$ und p besitzt außer ± 1 und $\pm p$ keine weiteren Teiler. Man kann zeigen, dass eine Zahl prim genau dann ist, wenn sie irreduzibel ist.

Satz (Primfaktorzerlegung). *Jede ganze Zahl $n \in \mathbb{Z} \setminus \{0, \pm 1\}$ kann als Produkt endlich vieler Primzahlen, der Primfaktoren, dargestellt werden. Diese Primfaktorzerlegung ist bis auf die Reihenfolge der auftretenden Zahlen eindeutig.*

Wir nehmen an, die Behauptung sei falsch. Nach dem Wohlordnungsprinzip gibt es dann eine kleinste natürliche Zahl n_0, die nicht in Primfaktoren zerlegt werden kann. Insbesondere kann n_0 keine Primzahl sein. Somit gibt es $n, m \in \mathbb{N}$ mit $n_0 = nm$ und $n, m > 1$. Dies impliziert aber $n < n_0$ und $m < n_0$. Aus der Definition von n_0 ergibt sich nun, dass n und m als Produkte endlich vieler Primzahlen dargestellt werden können, also auch $n_0 = nm$, was wir aber ausgeschlossen haben. Damit haben wir die Existenzaussage bewiesen.

Notiz zur endlichen Mengen und Kombinatorik.

1.9.1. **Bemerkung.** Seien X, Y endlich.

(1) $|X|$ ist wohldefiniert: $X \sim \{1, \ldots, n\}$ und $X \sim \{1, \ldots, m\} \Rightarrow n = m$,
(2) $X \subset Y \Rightarrow |X| \leqslant |Y|$,
(3) $\varphi : X \longrightarrow Y$ injektiv $\Rightarrow |X| \leqslant |Y|$,
(4) $\varphi : X \longrightarrow Y$, $|X| > |Y| \Rightarrow \exists x_1 \neq x_2 \in X$ mit $\varphi(x_1) = \varphi(x_2)$ (oder $|X| > |Y| \Rightarrow \exists y \in Y$ mit $|\varphi^{-1}(y)| \geqslant 2$).
(5) $|X \times Y| = |X| \times |Y|$.
(6) Sei $X = \cup_{1 \leqslant i \leqslant k} X_i$ eine Zerlegung von X. Dann gilt $|X| = \sum_{i=1}^{k} |X_i|$.
(Unter einer ***Zerlegung*** einer Menge X verstehen wir eine Familie $(X_i)_{i \in I}$ von paarweise disjunkten Teilmengen von X (d.h. $X_i \cap X_j = \emptyset$ für $i \neq j$), so dass $X = \cup_{i \in I} X_i$.)

Die Eigenschaft (4) ist nichts anderes als das Schubfachprinzip, das in der Zahlentheorie und in der diskreten Mathematik oft benutzt wird.

1.9.2. **Satz.** *[Dirichletsches Schubfachprinzip] Falls man n Objekte auf m Mengen* $(n, m \in \mathbb{N})$ *verteilt, und n größer als m ist, dann gibt es mindestens eine Menge, in der mehr als ein Objekt landet.*

Aufgaben. Beispiele für die Anwendung des Schubfachprinzips:

1. Auf jeder Party gibt es mindestens zwei Leute, die gleich vielen anderen Leuten die Hand geschüttelt haben.

2. In jeder Gruppe von 6 Menschen gibt es entweder 3, die alle miteinander befreundet sind oder 3, die alle nicht miteinander befreundet sind. (Die Freundschaftsbeziehung ist symmetrisch: Wenn A mit B befreundet ist, dann auch B mit A.)

3 (Erdössche Einstiegsfrage zur Mathematik). Sei $X \subset \{1, 2, \ldots, 2n\}$ eine Teilmenge mit $n+1$ Elementen. Dann gibt es immer zwei Zahlen in X so, dass eine die andere teilt. (Tipp: Schreibe jede Nummer $x \in X$ in der Form $x = 2^k m$, wobei m eine ungerade Zahl zwischen 1 und $2n-1$ sei.)

4 (Satz von Erdös und Szekeres). Sei $F = (a_1, \ldots, a_n)$ eine Folge von n verschiedenen reellen Zahlen. Wenn $n \geqslant sr + 1$, dann besitzt F entweder eine steigende Teilfolge aus $s+1$ Zahlen oder eine fallende Teilfolge aus $r+1$ Zahlen (oder beides). (Tipp: Ordne jeder Zahl a_i aus der Folge F ein Paar (x_i, y_i) zu: x_i ist die Länge der längsten steigenden Teilfolge mit der Endung a_i und y_i ist die Länge der längsten fallenden Teilfoge mit dem Anfang a_i.)

5 (Gitterpunkte). Unter 5 Gitterpunkten der Ebene sind stets zwei, deren Verbindungslinie einen Gitterpunkt als Mittelpunkt hat. Unter 9 Gitterpunkten der Ebene gibt es drei, deren Schwerpunkt ebenfalls ein Gitterpunkt ist.

6. Zehn Ecken eines regelmäßigen 100-Ecks seien rot und zehn andere blau gefärbt. Man beweise: Unter den Verbindungsstrecken zweier roter Punkte gibt es mindestes eine, die genauso lang ist wie eine der Verbindungsstrecken zweier blauer Punkte.

7 (Sylvesters Problem). Eine endliche Menge von Punkten habe die Eigenschaft, dass jede Gerade durch zwei Punkte durch einen dritten Punkt verläuft. Zeige, dass alle Punkte auf einer Geraden liegen.

1.9.3. **Definition.** Seien $n, k \in \mathbb{N}$. Sei X eine n-elementige Menge.

- Eine ***k*-Permutation ohne Wiederholung** von n Elementen ist eine Auswahl von k verschiedenen Elementen von X, bei der es auf die Reihenfolge ankommt.
- Eine n-Permutation ohne Wiederholung von n Elementen heißt einfach ***Permutation*** von n Elementen.
- Eine ***k*-Kombination ohne Wiederholung** von n Elementen ist eine Auswahl von k verschiedenen Elementen von X, bei der es auf die Reihenfolge nicht ankommt.
- Eine ***k*-Permutation mit Wiederholung** von n Elementen ist eine Auswahl von k nicht unbedingt verschiedenen Elementen von X, bei der es auf die Reihenfolge ankommt.
- Eine ***k*-Kombination mit Wiederholung** von n Elementen ist eine Auswahl von k nicht unbedingt verschiedenen Elementen von X, bei der es auf die Reihenfolge nicht ankommt.

Anstelle von k-Permutation (k-Kombination) ohne Wiederholung sagt man einfach k-Permutation (k-Kombination).

- Eine k-Permutation ist eine Familie $(x_1, \ldots, x_k)$ von verschiedenen Elementen von X, d.h. eine injektive Abbildung $\{1, \ldots, k\} \to X$. Dabei muss $k \leqslant n$ sein.
- Eine Permutation ist eine bijektive Abbildung $\{1, \ldots, n\} \to X$.
- Eine k-Kombination ist eine k-elementige Teilmenge $\{x_1, \ldots, x_k\}$ von X.
- Eine k-Permutation mit Wiederholung ist eine Abbildung $\{1, \ldots, k\} \to X$ (also ein k-Tupel).
- Eine k-Kombination mit Wiederholung ist eine Teilmenge $\{x_1, \ldots, x_s\}$ von X zusammen mit Vielfachheiten $j_{x_1}, \ldots, j_{x_s} \in \mathbb{N}$ mit $j_{x_1} + \ldots + j_{x_s} = k$.

Ein Beispiel für eine Kombination ohne Wiederholung ist das Zahlenlotto „6 aus 49“. Dabei zieht man 6 Kugeln aus einer Urne mit 49 nummerierten Kugeln, wobei es auf die Reihenfolge nicht ankommt. Eine gezogene Kugel wird nicht zurückgelegt. Würde man jede Kugel nach ihrer Ziehung wieder zurücklegen, so wäre das Ergebnis eine Kombination mit Wiederholung.

Sei $A = \{1, 2, 3\}$. Dann lauten alle

- 2-Permutationen ohne Wiederholung: $(1,2)$, $(2,1)$, $(1,3)$, $(3,1)$, $(2,3)$, $(3,2)$.

- 2-Kombinationen ohne Wiederholung: $\{1,2\}$, $\{2,3\}$, $\{1,3\}$.
- 2-Permutationen mit Wiederholung: $(1,1)$, $(2,2)$, $(3,3)$, $(1,2)$, $(2,1)$, $(1,3)$, $(3,1)$, $(2,3)$, $(3,2)$.
- 2-Kombinationen mit Wiederholung: (11), (12), (13), (22), (23), (33).

1.9.4. **Satz.**

- *Die Anzahl von k-Permutationen von n Elementen ist* $\dfrac{n!}{(n-k)!}$.
- *Die Anzahl von Permutationen von n Elementen ist* $n!$.
- *Die Anzahl von k-Kombinationen von n Elementen ist* $\dbinom{n}{k}$.
- *Die Anzahl von k-Permutationen mit Wiederholung von n Elementen ist* n^k.

Aufgaben.

8. Für zwei Mengen X, Y bezeichnet man mit $\mathcal{B}(X,Y)$ die Menge der bijektiven Abbildungen von X nach Y. Seien X, Y, Z drei gleichmächtige Mengen. Zeigen Sie, dass $\mathcal{B}(X,Y)$ und $\mathcal{B}(X,Z)$ gleichmächtig sind.
(Tipp: Sei $\psi : Y \to Z$ bijektiv. Dann ist $\mathcal{B}(X,Y) \to \mathcal{B}(X,Z)$, $\varphi \mapsto \psi \circ \varphi$ bijektiv.)
Das zeigt, dass die Ausdrücke in Definition 1.9.3 wohldefiniert sind, d.h. sie hängen von der Menge X nicht ab.

9. Die Anzahl von k-Kombinationen mit Wiederholung von n Elementen ist $\binom{n+k-1}{k}$.

Notiz zur Kardinalzahlen.

Satz.

(i) *Jede Teilmenge einer abzählbaren Menge ist höchstens abzählbar.*
(ii) *Jede unendliche Menge besitzt eine abzählbare Teilmenge.*
(iii) *A,B sind abzählbar $\Rightarrow A \times B$ ist abzählbar.*
(iv) *A_k ist (höchstens) abzählbar für jedes $k \in \mathbb{N} \Rightarrow \cup_{k\in\mathbb{N}} A_k$ ist (höchstens) abzählbar.*
(v) *Seien $A \subset B$ zwei Mengen, so dass A abzählbar ist und $B \setminus A$ unendlich ist. Dann gilt $B \sim B \setminus A$.*

Eine lustige Formulierung einiger Eigenschaften von abzählbaren Mengen findet man in dem Artikel[2] Hilberts Hotel.

Die Menge $\mathbb{R}$ hat mehr Elemente als $\mathbb{Q}$. Wir haben also auch eine Hierarchie zwischen unendlichen Mengen erhalten, die wir mathematisch so präzisieren können: Zu jeder Menge X assoziieren wir das Symbol $|X|$ (Kardinalzahl von X), so dass $|X| = |Y| \Leftrightarrow X \sim Y$. Wenn X endlich ist und n Elemente hat, so setzen wir wie zuvor $|X| = n$. Wenn X abzählbar ist, so setzen wir $|X| = \aleph_0$ ($\aleph$, gelesen aleph, ist der erste Buchstabe des hebräischen Alphabets). Wenn $X \sim \mathbb{R}$, so sagen wir, dass X die Mächtigkeit des Kontinuums hat, und setzen $|X| = \aleph_c$. Beispiel: $\mathbb{R} \setminus \mathbb{Q}$ oder nichtentartete Intervalle $I \subset \mathbb{R}$ haben die Mächtigkeit des Kontinuums.

Wir sagen, dass Y höchstens mächtiger als X ist und schreiben $|X| \leqslant |Y|$, falls X gleichmächtig zu einer Teilmenge von Y ist (wenn es also eine Injektion von X in Y gibt).

Wir sagen, dass Y mächtiger als X ist und schreiben $|X| < |Y|$, falls X gleichmächtig zu einer echten Teilmenge von Y, aber nicht gleichmächtig zu Y ist (wenn es also eine Injektion von X in Y, aber keine Bijektion von X nach Y gibt). Beispiel: Für X endlich gilt $|X| < \aleph_0 < \aleph_c$.

Gibt es Kardinahlzahlen zwischen $\aleph_0$ und $\aleph_c$? Die Cantorsche Kontinuumshypothese besagt, dass $\aleph_0 < |X| \leqslant \aleph_c$ impliziert: $|X| = \aleph_c$. Gödel (1938) und Cohen (1963) haben gezeigt, dass auf der Basis der Zermelo-Fraenkel-Axiomatik der Mengenlehre weder die Kontinuumshypothese noch ihr Gegenteil beweisbar sind [4].

Können wir beliebig große Kardinalzahlen konstruieren? Wenn wir das Produkt $X \times X$ betrachten, erhalten wir keine mächtigere Menge als X, falls X unendlich ist:

Satz. *Sei X eine unendliche Menge. Dann gilt $|X \times X| = |X|$.*

Die richtige Methode ist es, die Potenzmenge $\mathcal{P}(X)$ einer Menge X (siehe A.4.11) zu nehmen.

Satz (Cantor)**.** *Für jede Menge X gilt $|X| < |\mathcal{P}(X)|$.*

Der folgende Satz zeigt, dass $\leqslant$ eine Ordnung für Kardinalzahlen ist:

[2]http://de.wikipedia.org/wiki/Hilberts_Hotel

Satz (Bernstein-Schröder)**.** *Sind $\aleph_a$ und $\aleph_b$ zwei Kardinalzahlen, so gibt es nur drei Möglichkeiten: $\aleph_a < \aleph_b$, $\aleph_a = \aleph_b$, $\aleph_b < \aleph_a$.*

Als Anwendung des Satzes von Bernstein-Schröder zeigen wir:

Satz. $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$.

Zunächst gilt $\mathbb{R} \sim I := (0,1)$ und $\mathcal{P}(\mathbb{N}) \sim \{0,1\}^{\mathbb{N}} := \{f : \mathbb{N} \to \{0,1\}\}$ (eine Teilmenge $A \subset \mathbb{N}$ definiert eindeutig eine Folge $f \in \{0,1\}^{\mathbb{N}}$, indem man $f(n) = 1$ für $n \in \mathbb{N}$ mit $n \in A$ und $f(n) = 0$ für $n \in \mathbb{N}$ mit $n \notin A$ setzt). Wir definieren nun $T : \{0,1\}^{\mathbb{N}} \to \mathbb{R}$ durch

$$T(f) = \begin{cases} \sum\limits_{n=1}^{\infty} \dfrac{f(n)}{2^n}, & \text{falls} \quad f(n) = 0 \quad \text{für unendlich viele} \, n, \\ 1 + \sum\limits_{n=1}^{\infty} \dfrac{f(n)}{2^n}, \text{sonst.} \end{cases}$$

Da jede reelle Zahl aus $(0,1)$ eine eindeutig bestimmte Dualbruchentwicklung der Gestalt $\sum_{n=1}^{\infty} \frac{f(n)}{2^{n+1}}$ besitzt, bei der $f(n) = 0$ für unendlich viele n ist, sieht man leicht, dass $T : \{0,1\}^{\mathbb{N}} \to \mathbb{R}$ die Menge $\{0,1\}^{\mathbb{N}}$ bijektiv auf eine Obermenge $J \subset \mathbb{R}$ von I abbildet. Daher ist $\{0,1\}^{\mathbb{N}} \sim J$, also $|\mathcal{P}(\mathbb{N})| = |\{0,1\}^{\mathbb{N}}| = |J|$. Nun gilt $I \subset J \subset \mathbb{R}$, also $|I| \leqslant |J| \leqslant |\mathbb{R}|$ und $|I| = |\mathbb{R}|$, und der Satz von Bernstein-Schröder ergibt $|I| = |J| = |\mathbb{R}|$. Wir erhalten also $|\mathcal{P}(\mathbb{N})| = |J| = |\mathbb{R}|$.

Mehr zum Thema Kardinalzahlen findet man in [9].

3.11.2011

2. Folgen und Konvergenz

2.1. Definition und Beispiele.

2.1.1. Definition.
(i) Sei $A \neq \emptyset$. Sei $q \in \mathbb{Z}$, $\mathbb{Z}(\geqslant q) := \{n \in \mathbb{Z} : n \geqslant q\}$.
Eine Abbildung $f : \mathbb{Z}(\geqslant q) \longrightarrow A$ heißt ***Folge*** in A. Ist $a_n := f(n)$, so schreibt man f als $(a_n)_{n\geqslant q}$. Ist $q = 1$, so bezeichnet man $(a_n)_{n\geqslant 1}$ auch mit $(a)_{n\in\mathbb{N}}$. Die Elemente a_n heißen ***Glieder*** der Folge. Der Einfachheit halber betrachten wir in der Theorie nur Folgen mit Indexmenge $\mathbb{N}$, in der Praxis treten aber auch Folgen mit Indexmenge $\mathbb{Z}(\geqslant q)$ auf.

(ii) Eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ heißt ***konvergent***, wenn es $a \in \mathbb{R}$ gibt mit der Eigenschaft: Zu jedem $\varepsilon > 0$ gibt es $n_0 = n_0(\varepsilon) \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geqslant n_0$ gilt: $|a_n - a| < \varepsilon$. Man beachte, dass $n_0 = n_0(\varepsilon) \in \mathbb{N}$ im Allgemeinen von ε abhängt.
Die Definition lautet formal:

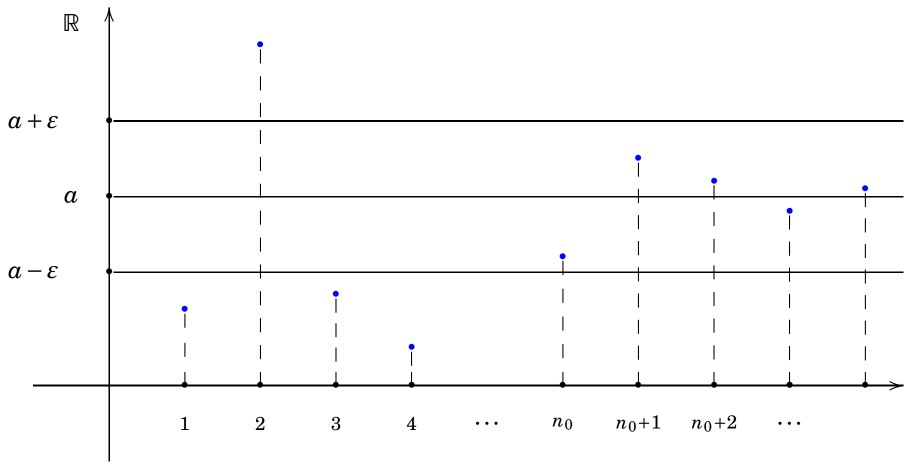
$$(a_n)_{n\in\mathbb{N}} \text{ konvergent} :\Longleftrightarrow \exists a \in \mathbb{R}\; \forall \varepsilon > 0\; \exists n_0\; \forall n \geqslant n_0 : |a_n - a| < \varepsilon\,.$$

Eine Zahl $a \in \mathbb{R}$ wie in der Definition heißt ***Grenzwert*** oder ***Limes*** der Folge. Wir schreiben „$\lim_{n\to\infty} a_n = \lim a_n = a$“ oder „$a_n \to a$ für $n \to \infty$“ und sagen, dass $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert.

(iii) Konvergiert eine Folge gegen 0, so heißt sie ***Nullfolge***.

(iv) Eine nicht konvergente Folge heißt ***divergent***.

Die blauen Punkte stellen den Graph $\{(n, a_n) : n \in \mathbb{N}\}$ einer konvergenten Folge dar:



2.1.2. Definition. Sei $a \in \mathbb{R}$, $\varepsilon > 0$. Die Menge $B_\varepsilon(a) = \{x \in \mathbb{R} : |x - a| < \varepsilon\} = (a - \varepsilon, a + \varepsilon)$ heißt ***ε-Umgebung*** von a. Jede Teilmenge $U \subset \mathbb{R}$, zu der es ein $\varepsilon > 0$ mit $B_\varepsilon(a) \subset U$ gibt, heißt ***Umgebung*** von a.

Umformulierung der Definition 2.1.1 der Konvergenz:

$$\lim_{n\to\infty} a_n = a \iff \text{Zu jeder Umgebung } U \text{ von } a \text{ gibt es } n_0 = n_0(U) \in \mathbb{N}, \text{ so dass für alle } n \geqslant n_0 \text{ gilt: } a_n \in U.$$

Es ist zweckmäßig, die folgende *Sprechweise* zu benutzen: Wir sagen, dass *fast alle* Elemente einer Menge eine Eigenschaft haben, wenn höchstens endlich viele Elemente der Menge die betreffende Eigenschaft nicht haben.

Neue Umformulierung der Definition 2.1.1 der Konvergenz:

$$\lim_{n\to\infty} a_n = a \iff \text{Für jede Umgebung } U \text{ von } a \text{ gilt } a_n \in U \text{ für fast alle } n \in \mathbb{N}.$$

Durch Negation erhalten wir auch:

$$(a_n)_{n\in\mathbb{N}} \text{ divergent} \iff \text{Jedes } a \in \mathbb{R} \text{ hat eine Umgebung } U, \text{ so dass } a_n \notin U \text{ für unendlich viele } n \in \mathbb{N}.$$

Beispiel: Die Folge $1,0,1,0,\dots$, d.h.

$$a_n = \begin{cases} 1, n \,\text{ungerade} \\ 0, n \,\text{gerade} \end{cases}$$

ist divergent. In der Tat, für jedes $a \in \mathbb{R}$ wähle $U = B_{1/2}(a) = (a - \frac{1}{2}, a + \frac{1}{2})$. Dann ist entweder $1 \notin B_{1/2}(a)$ oder $0 \notin B_{1/2}(a)$, also $a_n \notin B_{1/2}(a)$ für n ungerade oder $a_n \notin B_{1/2}(a)$ für n gerade.

2.1.3. **Satz.** *Es gilt*

$$\lim_{n\to\infty} \frac{1}{n^s} = 0, \quad \textit{für jedes } s \in \mathbb{Q}_+, \tag{2.1}$$

$$\lim_{n\to\infty} \sqrt[n]{a} = 1, \quad \textit{für jedes } a > 0, \tag{2.2}$$

$$\lim_{n\to\infty} \sqrt[n]{n} = 1, \tag{2.3}$$

$$\lim_{n\to\infty} a^n = 0, \quad \textit{für jedes } a \in \mathbb{R} \textit{ mit } |a| < 1, \tag{2.4}$$

$$\lim_{n\to\infty} \frac{n^s}{a^n} = 0, \quad \textit{für jedes } s \in \mathbb{Q} \textit{ und jedes } a \in \mathbb{R} \textit{ mit } |a| > 1. \tag{2.5}$$

Beweis: Wegen der Wichtigkeit des Konvergenzbegriffes geben wir den Beweis. Um die Konvergenz von (a_n) gegen a zu beweisen, muss man für beliebig vorgegebenes $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon)$ angeben, so dass $|a_n - a| < \varepsilon$ für alle $n \geqslant n_0$ gilt. Im Prinzip muss man also die Ungleichung $|a_n - a| < \varepsilon$ nach der Unbekannten $n \in \mathbb{N}$ auflösen und eine Menge $\{n \in \mathbb{N} : n \geqslant n_0\}$ von Lösungen finden. Falls a_n eine einfache Funktion von n ist, kann die Lösung der Ungleichung leicht sein. Falls nicht, muss man zunächst eine Abschätzung $|a_n - a| \leqslant b_n$ finden, wobei b_n eine einfache Form hat. Danach löst man die Ungleichung $b_n < \varepsilon$.

Wir gehen in zwei Etappen vor. Zunächst kommt eine heuristische Phase (Laborphase), in der man eine geeignete Vorausabschätzung (Ansatz) für $|a_n - a|$ sucht und die Ungleichung $|a_n - a| (\leqslant b_n) < \varepsilon$ löst. Diese Etappe erscheint normalerweise nicht in Schriftform, und wenn, dann nur in Büchern über das Lösen von Problemen (ein berühmtes Beispiel ist das Buch [18] von G. Polya). Man benutzt aber die Ergebnisse der Laborphase, um anschließend einen formalen Beweis niederzuschreiben.

Zu (2.1). Laborphase: Wir lösen $|a_n| = \frac{1}{n^s} < \varepsilon$ für gegebenes $\varepsilon > 0$. Es gilt

$$\frac{1}{n^s} < \varepsilon \Leftrightarrow \frac{1}{\varepsilon} < n^s \Leftrightarrow \frac{1}{\varepsilon^{1/s}} < n \Leftrightarrow n \geqslant \left\lfloor \frac{1}{\varepsilon^{1/s}} \right\rfloor + 1.$$

Wir könnten es also mit einem $n_0 > \frac{1}{\varepsilon^{1/s}}$ versuchen. Und in der Tat, für $n \geqslant n_0$ gilt $n^s \geqslant n_0^s$ (weil $s > 0$), also $\frac{1}{n^s} \leqslant \frac{1}{n_0^s} < \varepsilon$.

Formaler Beweis: Sei $\varepsilon > 0$ vorgegeben. Wähle $n_0 \in \mathbb{N}$ mit $n_0 > \frac{1}{\varepsilon^{1/s}}$ (das ist möglich nach dem Satz von Archimedes 1.5.1 oder Eudoxus 1.5.2). Dann gilt für alle $n \geqslant n_0$:

$$\left| \frac{1}{n^s} - 0 \right| = \frac{1}{n^s} \leqslant \frac{1}{n_0^s} < \varepsilon.$$

Nach der Definition folgt $\lim_{n\to\infty} \frac{1}{n^s} = 0$.

Zu (2.2). Sei zunächst $a \geqslant 1$ (1. Fall).

Laborphase: Die Ungleichung $|\sqrt[n]{a} - 1| < \varepsilon$ lässt sich nicht direkt lösen. Wir brauchen eine a-priori-Abschätzung. Dazu betrachten wir $x_n := \sqrt[n]{a} - 1 \geqslant 0$ und formen um: $1 + x_n = \sqrt[n]{a}$, $(1 + x_n)^n = a$. Die Form $(1 + x_n)^n$ lässt uns an die Bernoulli-Ungleichung oder die Binomialformel denken; nach Bernoulli gilt $(1 + x_n)^n \geqslant 1 + n x_n$; außerdem $1 + n x_n > n x_n$. Zusammengefasst:

$$a = (1 + x_n)^n \geqslant 1 + n x_n > n x_n \,.$$

Also ist die gewünschte Abschätzung $x_n < \frac{a}{n}$. Wir lösen die Ungleichung $\frac{a}{n} < \varepsilon$; die Lösungsmenge ist $\{n \in \mathbb{N} : n \geqslant \lfloor \frac{a}{\varepsilon} \rfloor + 1\}$.

Formaler Beweis: Setze $x_n := \sqrt[n]{a} - 1$. Wegen $a \geqslant 1$ gilt $x_n \geqslant 0$. Nach Bernoulli gilt $a = (1 + x_n)^n \geqslant 1 + n x_n > n x_n$, also $x_n < \frac{a}{n}$. Sei $\varepsilon > 0$ vorgegeben. Wähle $n_0 > \frac{a}{\varepsilon}$. Dann gilt für alle $n \geqslant n_0$:

$$|\sqrt[n]{a} - 1| = x_n < \frac{a}{n} \leqslant \frac{a}{n_0} < \varepsilon\,.$$

2. Fall: $0 < a < 1$.

Laborphase: Wir versuchen, den 1. Fall zu benutzen. Ist $0 < a < 1$, so ist $\frac{1}{a} > 1$. Nach dem 1. Fall gilt $\sqrt[n]{\frac{1}{a}} \to 1$, also $\frac{1}{\sqrt[n]{a}} \to 1$ für $n \to \infty$. Kann man daraus schließen, dass $\sqrt[n]{a} \to 1$? Das ist der Fall nach einem der Konvergenzsätze, die wir bald beweisen: Gilt $x_n \to x \neq 0$ für $n \to \infty$, so gibt es $N \in \mathbb{N}$ mit $x_n \neq 0$ für alle $n \geqslant N$, und die Folge $(\frac{1}{x_n})_{n \geqslant N}$ konvergiert gegen $\frac{1}{x}$.

Formaler Beweis: Ist $0 < a < 1$, so $\frac{1}{a} > 1$. Nach dem 1. Fall gilt $\frac{1}{\sqrt[n]{a}} = \sqrt[n]{\frac{1}{a}} \to 1 \neq 0$ für $n \to \infty$. Nach den Konvergenzsätzen folgt $\sqrt[n]{a} \to \frac{1}{1} = 1$, $n \to \infty$.

Zu (2.3). Laborphase: Wir setzen unsere erfolgreiche Strategie fort und betrachten $x_n := \sqrt[n]{n} - 1 \geqslant 0$. Dann ist $n = (1 + x_n)^n$. Die Abschätzung $n = (1 + x_n)^n \geqslant 1 + n x_n > n x_n$ liefert nur $x_n < 1$ und hilft nicht weiter. Wir brauchen also ein x_n^2 auf der rechten Seite, um eine ähnliche Abschätzung wie in (2) zu erhalten. Glücklicherweise ist das möglich mit Hilfe der Binomialformel:

$$(1 + x_n)^n = 1 + \binom{n}{1} x_n + \binom{n}{2} x_n^2 + \ldots + \binom{n}{n} x_n^n \geqslant 1 + \binom{n}{2} x_n^2 .$$

Hier benutzt man, dass wegen $x_n \geqslant 0$ alle Glieder $\binom{n}{k} x_n^k \geqslant 0$ sind. Also ist

$$n = (1 + x_n)^n \geqslant 1 + \frac{n(n-1)}{2} x_n^2 ,$$

und daraus folgt $x_n < \sqrt{\frac{2}{n}}$. Das ist die gewünschte Abschätzung. Nun löst man die Ungleichung $\sqrt{\frac{2}{n}} < \varepsilon$:

$$\sqrt{\frac{2}{n}} < \varepsilon \Leftrightarrow \frac{2}{n} < \varepsilon^2 \Leftrightarrow \frac{2}{\varepsilon^2} < n \Leftrightarrow n \geqslant \left\lfloor \frac{2}{\varepsilon^2} \right\rfloor + 1 .$$

Formaler Beweis: Setze $x_n := \sqrt[n]{n} - 1$. Es ist $\sqrt[n]{n} \geqslant 1$ (sonst $n = (\sqrt[n]{n})^n < 1$). Daraus folgt $x_n \geqslant 0$. Weiter wird die Binomialformel benutzt:

$$n = (1 + x_n)^n = 1 + \tbinom{n}{1} x_n + \tbinom{n}{2} x_n^2 + \ldots + \tbinom{n}{n} x_n^n \geqslant 1 + \tbinom{n}{2} x_n^2 .$$

Hieraus erhalten wir $n - 1 \geqslant \frac{n(n-1)}{2} x_n^2$ und somit $x_n^2 \leqslant \frac{2}{n}$, d.h. $x_n \leqslant \sqrt{\frac{2}{n}}$.
Zu vorgegebenem $\varepsilon > 0$ wähle $n_0 > \frac{2}{\varepsilon^2}$. Dann gilt für alle $n \geqslant n_0$:

$$|\sqrt[n]{n} - 1| = x_n < \sqrt{\tfrac{2}{n}} \leqslant \sqrt{\tfrac{2}{n_0}} < \varepsilon .$$

Zu (2.4). Schreibe $\dfrac{1}{|a|} = 1 + x$, wobei $x > 0$. Es folgt nach Bernoulli:

$$\frac{1}{|a|^n} = (1 + x)^n > 1 + n x > n x .$$

Zu vorgegebenem $\varepsilon > 0$ wähle $n_0 > \frac{1}{\varepsilon x}$. Dann gilt für alle $n \geqslant n_0$:

$$|a^n - 0| = a^n x_n < \frac{1}{n x} \leqslant \frac{1}{n_0 x} < \varepsilon .$$

Zu (2.5) 1. Fall: $s > 0$.

Laborphase: Hier sind zwei entgegengesetzt wirkende Kräfte am Werk: n^s wächst über alle Grenzen (siehe (1); später formulieren wir dieses Sachverhältnis als $n^s \to \infty$) und $\frac{1}{a^n}$ konvergiert gegen Null (siehe (4)). Wer gewinnt die Oberhand?

Wie oben schreiben wir $|z| = 1 + x$ mit $x > 0$. Um den Einfluss von n^s zu dämpfen, müssen wir $\frac{1}{a^n}$ nach oben abschätzen durch eine Potenz $\frac{1}{n^k}$ mit $k > s$. Äquivalent dazu: Wir müssen $a^n = (1 + x)^n$ nach unten abschätzen durch eine Potenz n^k mit $k > s$. Wir legen ein $k \in \mathbb{N}$, $k > s$ fest, z.B. $k = \lfloor s \rfloor + 1$. Wir versuchen es nochmals mit der Binomialformel, da die Binomialkoeffizienten Potenzen von n enthalten. Es gilt $(1 + x)^n > \binom{n}{k} x^k$ mit $\binom{n}{k} = \frac{n(n-1)\ldots(n-k+1)}{k!}$. Wenn n groß wird, dann wird auch $n - k + 1$ sehr groß, nämlich $n - k + 1 > \frac{n}{2}$ für alle $n > 2k$. Dann ist $n > \frac{n}{2}, n - 1 > \frac{n}{2}, \ldots, n - k + 1 > \frac{n}{2}$, und durch Multiplizieren: $\binom{n}{k} > \frac{1}{k!} \left(\frac{n}{2}\right)^k$. Also gilt für $n > 2k$:

$$\left| \frac{n^s}{a^n} - 0 \right| = \frac{n^s}{a^n} = \frac{n^s}{(1 + x)^n} < \frac{n^s}{\binom{n}{k} x^k} < \frac{n^s}{\frac{1}{k!}(\frac{n}{2})^k x^k} = \frac{2^k k!}{x^k} \cdot \frac{1}{n^{k-s}} .$$

Nun wissen wir nach (1), dass $\frac{1}{n^{k-s}} \to 0$ gilt, und weil $\frac{2^k k!}{x^k}$ eine Konstante ist, folgt $\frac{n^s}{a^n} \to 0$. Bei der Wahl von k können wir großzügiger sein; wenn wir $k \geqslant s + 1$ festlegen, wird $\frac{1}{n^{k-s}} \leqslant \frac{1}{n}$. Dann wird die Wahl von $n_0(\varepsilon)$ zu vorgegebenem ε einfacher.

Formaler Beweis: Schreibe $|a| = 1 + x$, wobei $x > 0$. Wähle $k > s + 1$ fest. Aus der Binomialformel ergibt sich leicht die Abschätzung

$$|a|^n = (1 + x)^n > \tbinom{n}{k} x^k = n(n-1) \cdot \ldots \cdot (n - k + 1) \tfrac{1}{k!} x^k .$$

Für $n > 2k$ gilt $n-k+1 > \frac{n}{2}$, $n-k+2 > \frac{n}{2}, \dots, n-1 > \frac{n}{2}$, $n > \frac{n}{2}$, also

$$|a|^n = (1+x)^n > \left(\tfrac{n}{2}\right)^k \tfrac{1}{k!} x^k .$$

Daraus folgt

$$\frac{n^s}{|a|^n} < n^s \left(\frac{2}{n}\right)^k \frac{k!}{x^k} = \frac{2^k}{n^{k-s}} \frac{k!}{x^k} < \frac{2^k k!}{x^k} \frac{1}{n},$$

wobei in der letzten Ungleichung benutzt wurde, dass $n^{k-s} \geqslant n$ wegen $k - s \geqslant 1$ gilt.

Zu vorgegebenem $\varepsilon > 0$ wähle $n_0 > \max\{\frac{2^k k!}{x^k}, 2k\}$. Dann gilt für alle $n \geqslant n_0$:

$$\left|\frac{n^s}{a^n} - 0\right| = \frac{n^s}{|a|^n} < \frac{2^k k!}{x^k} \frac{1}{n} \leqslant \frac{2^k k!}{x^k} \frac{1}{n_0} < \varepsilon .$$

2. Fall: $s \leqslant 0$. Es gilt $\frac{n^s}{a^n} = \frac{1}{n^{-s}} \frac{1}{a^n}$. Weil $-s \geqslant 0$ ist, gilt $n^{-s} \geqslant n^0 = 1$, also $|\frac{1}{n^{-s}}| \leqslant 1$. Außerdem ist $\frac{1}{a^n}$ eine Nullfolge nach (4). Daraus folgt, dass auch $\frac{n^s}{a^n}$ eine Nullfolge ist. □

2.1.4. **Definition.** Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt ***beschränkt***, wenn es $M > 0$ gibt, mit $|a_n| \leqslant M$ für alle $n \in \mathbb{N}$. Eine reelle Folge $(a_n)_{n\in\mathbb{N}}$ heißt ***monoton wachsend (steigend)*** (bzw. ***fallend***) wenn $a_n \leqslant a_{n+1}$ (bzw. $a_n \geqslant a_{n+1}$) für alle $n \in \mathbb{N}$ gilt.

Eine reelle Folge $(a_n)_{n\in\mathbb{N}}$ heißt ***streng monoton wachsend (steigend)*** (bzw. ***fallend***) wenn $a_n < a_{n+1}$ (bzw. $a_n > a_{n+1}$) für alle $n \in \mathbb{N}$ gilt.

2.1.5. **Satz** (Monotonieprinzip). *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$, $A := \{a_n : n \in \mathbb{N}\}$.*

(i) *$(a_n)_{n\in\mathbb{N}}$ monoton wachsend und (nach oben) beschränkt $\Rightarrow a_n \longrightarrow \sup A$, $n \to \infty$.*
(ii) *$(a_n)_{n\in\mathbb{N}}$ monoton fallend und (nach unten) beschränkt $\Rightarrow a_n \longrightarrow \inf A$, $n \to \infty$.*

2.1.6. **Beispiele.**

(i) Sei (I_n) eine Intervallschachtelung mit $I_n = [a_n, b_n]$ und $\cap I_n = \{x\}$. Dann sind (a_n) und (b_n) konvergent und $a_n \to x$, $b_n \to x$, $n \to \infty$. In der Tat, wegen der Definition der Intervallschachtelung, ist (a_n) monoton steigend und (b_n) monoton fallend. Im Satz 1.5.13 (Intervallschachtelungsprinzip) haben wir gesehen, dass (a_n) nach oben beschränkt ist (jedes b_n ist eine obere Schranke) und x wurde als $\sup\{a_n\}$ definiert. Da auch $x = \inf\{b_n\}$, die Aussage folgt aus dem Monotonieprinzip. Aus Satz 1.5.14 erhalten wir

$$\left(1+\frac{1}{n}\right)^n \longrightarrow e, \quad \left(1+\frac{1}{n}\right)^{n+1} \longrightarrow e, \qquad n \longrightarrow \infty. \tag{2.6}$$

(ii) Sei $a_n = \sum\limits_{k=0}^{n} \dfrac{1}{n!} = 1 + \dfrac{1}{1!} + \dfrac{1}{2!} + \ldots + \dfrac{1}{n!}$, für $n \in \mathbb{N}_0$; (a_n) ist monoton wachsend und beschränkt, $2\frac{1}{2} < a_n < 3$ für $n \geqslant 2$, also konvergent gegen eine Zahl in $(2,3]$. Der Grenzwert ist die Eulersche Zahl $e = \lim a_n$.

(iii) Sei $a > 0$, $k \in \mathbb{N}$, $k \geqslant 2$. Wähle $x_1 > 0$ mit $x_1^k > a$; definiere rekursiv für $n \in \mathbb{N}$:

$$x_{n+1} := x_n - \frac{x_n^k - a}{k x_n^{k-1}} = \frac{(k-1)x_n^k + a}{k x_n^{k-1}} > 0 .$$

(x_n) ist monoton fallend, nach unten durch $\sqrt[k]{a}$ beschränkt $\Rightarrow (x_n)$ konvergent.

2.1.7. **Satz.** *Ist $(a_n)_{n\in\mathbb{N}}$ konvergent so ist $(a_n)_{n\in\mathbb{N}}$ beschränkt.*

Bemerkung:

(i) $(a_n)_{n\in\mathbb{N}}$ unbeschränkt $\Longrightarrow (a_n)_{n\in\mathbb{N}}$ divergent. Z.B. $a_n = n$ ist divergent, ebenso

$$a_n = (-1)^n n = \begin{cases} n, & n \text{ gerade} \\ -n, & n \text{ ungerade} \end{cases}$$

(ii) $(a_n)_{n\in\mathbb{N}}$ beschränkt $\not\Longrightarrow (a_n)_{n\in\mathbb{N}}$ konvergent. Z.B. ist die Folge $1, 0, 1, 0, \ldots$ beschränkt, aber divergent.

14.11.2011

2.2. Rechnen mit konvergenten Folgen.

2.2.1. **Satz.** *Seien (a_n), (b_n) Folgen mit $a_n \to a$, $b_n \to b$ für $n \to \infty$. Dann gilt*

(i) $a_n + b_n \longrightarrow a + b$, $\lambda a_n \longrightarrow \lambda a$ *für* $\lambda \in \mathbb{R}$
(ii) $a_n b_n \longrightarrow ab$
(iii) $b \neq 0 \Rightarrow \exists n_0$ *mit* $b_n \neq 0$ *für* $n \geqslant n_0$, *und* $(\frac{a_n}{b_n})_{n \geqslant n_0}$ *ist konvergent mit* $\frac{a_n}{b_n} \longrightarrow \frac{a}{b}$
(iv) $|a_n| \longrightarrow |a|$
(v) (c_n) *beschränkt,* $a_n \longrightarrow 0 \Rightarrow a_n c_n \longrightarrow 0$.

2.2.2. **Satz** (Vergleichsprinzip). *Seien (a_n) und (b_n) konvergente Folgen mit $a_n \longrightarrow a$, $b_n \longrightarrow b$ für $n \to \infty$. Es gelte $a_n \leqslant b_n$ für fast alle $n \in \mathbb{N}$. Dann gilt $a \leqslant b$.*

Bemerkung: Auch wenn $a_n < b_n$ ist, so folgt im Allgemeinen nur $a \leqslant b$. Z.B. $a_n = -\frac{1}{n}$, $b_n = \frac{1}{n}$ und $a_n \longrightarrow 0$, $b_n \longrightarrow 0$.

2.2.3. **Folgerung.** *Sei (a_n) eine konvergente Folgen mit $a_n \longrightarrow a$ und $a_n \in [b,c]$ für fast alle $n \in \mathbb{N}$. Dann gilt $a \in [b,c]$.*

2.2.4. **Satz** (Einschließungsprinzip). *Seien (a_n) und (b_n) konvergente Folgen mit $a_n \longrightarrow a$, $b_n \longrightarrow a$ für $n \to \infty$. Sei (c_n) eine weitere Folge, und es gelte $a_n \leqslant c_n \leqslant b_n$ für fast alle $n \in \mathbb{N}$. Dann ist (c_n) konvergent, und es gilt $c_n \to a$ für $n \to \infty$.*

2.3. Der Satz von Bolzano-Weierstraß und das Cauchy-Kriterium.

2.3.1. **Definition.** Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$.

(i) $h \in \mathbb{R}$ heißt ***Häufungswert*** (HW) von (a_n), wenn für jedes $\varepsilon > 0$ gilt: $a_n \in B_\varepsilon(a)$ für unendlich viele $n \in \mathbb{N}$.

(ii) Ist $(n_k)_{k\in\mathbb{N}}$, $n_1 < n_2 < n_3 < \ldots < n_k < \ldots$ eine streng monoton wachsende Folge in $\mathbb{N}$, so heißt $(a_{n_k})_{k\in\mathbb{N}}$, d.h. die Folge $(a_{n_1}, a_{n_2}, a_{n_3}, \ldots)$ ist eine ***Teilfolge*** von (a_n).

2.3.2. **Beispiel.** (i) Ist (a_n) konvergent, so ist $a = \lim a_n$ ein Häufungswert von (a_n), und zwar der einzige. (Beweis identisch zum Beweis von Übung 2.5.1.)

(ii) Die Folge

$$a_n = \begin{cases} 0, n & \text{gerade} \\ 1, n & \text{ungerade} \end{cases}$$

hat 0 und 1 als Häufungswerte.

(iii) $\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \ldots, \frac{1}{2n}, \ldots$ und $1, \frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{2^n}, \ldots$ sind Teilfolgen von $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$. Aber die Folge $\frac{1}{4}, \frac{1}{2}, \frac{1}{6}, \frac{1}{8}, \ldots, \frac{1}{2n}, \ldots$ ist keine Teilfolge von $(\frac{1}{n})_{n\in\mathbb{N}}$, weil die Reihenordnung geändert ist.

17.11.2011

2.3.3. **Satz.** *$h \in \mathbb{R}$ ist Häufungswert von (a_n) $\Longleftrightarrow$ es gibt eine Teilfolge von (a_n), die gegen h konvergiert.*

2.3.4. **Satz** (Satz von Bolzano-Weierstraß). *Sei $(a_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$. Dann hat $(a_n)_{n\in\mathbb{N}}$ einen Häufungswert. Genauer: Sie hat einen größten Häufungswert h^* und einen kleinsten Häufungswert h_* definiert durch*

$$h^* := \inf_{n\in\mathbb{N}} \sup \{a_k : k \geqslant n\}, \quad h_* := \sup_{n\in\mathbb{N}} \inf \{a_k : k \geqslant n\}$$

Diese sind dadurch charakterisiert, dass für alle $\varepsilon > 0$ gilt:

- $h_* - \varepsilon < a_n < h^* + \varepsilon$ *für fast alle* $n \in \mathbb{N}$, *und*
- $a_n < h_* + \varepsilon$, *bzw.* $a_n > h^* - \varepsilon$, *für jeweils unendlich viele* $n \in \mathbb{N}$.

*Man nennt h^* den oberen Limes oder **Limes superior**, h_* den unteren Limes oder **Limes inferior** der Folge $(a_n)_{n\in\mathbb{N}}$ und schreibt dafür*

$$h^* = \limsup_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n\,, h_* = \liminf_{n\to\infty} a_n = \underline{\lim}_{n\to\infty} a_n\,.$$

2.3.5. **Folgerung.** (a_n) *konvergent* $\Longleftrightarrow$ (a_n) *ist beschränkt und* $\liminf a_n = \limsup a_n$.
Gegebenenfalls ist $\lim a_n = \liminf a_n = \limsup a_n$.

2.3.6. **Definition.** Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt ***Cauchy-Folge*** (CF), wenn gilt: Zu jedem $\varepsilon > 0$ gibt es $n_0 = n_0(\varepsilon) \in \mathbb{N}$ mit $|a_m - a_n| < \varepsilon$ für alle $n, m \geqslant n_0$.

2.3.7. **Satz** (Cauchy-Kriterium). *Eine Folge in $\mathbb{R}$ konvergiert genau dann, wenn sie eine Cauchy-Folge ist.*

2.3.8. **Definition.** (i) Wir definieren $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$, wobei $-\infty \notin \mathbb{R}$, $+\infty \notin \mathbb{R}$ zwei (ideale) Elemente sind. $\overline{\mathbb{R}}$ heißt die ***erweiterte Zahlengerade***. Wir erweitern die Ordnungsrelation in $\mathbb{R}$ auf $\overline{\mathbb{R}}$ durch $-\infty < x < \infty$ für alle $x \in \mathbb{R}$.

(ii) Für $A \subset \mathbb{R}$ beliebig, definieren wir $\sup A \in \overline{\mathbb{R}}$ als die kleinste obere Schranke von A in $\overline{\mathbb{R}}$ und $\inf A \in \overline{\mathbb{R}}$ als die größte untere Schranke von A in $\overline{\mathbb{R}}$. Dann gilt

$$\sup A = \infty \Longleftrightarrow A \text{ ist nach oben unbeschränkt}$$
$$\inf A = -\infty \Longleftrightarrow A \text{ ist nach unten unbeschränkt}$$

(iii) Sei $a \in \mathbb{R}$. Setze

$$(a,\infty) := \{x \in \overline{\mathbb{R}} : a < x < \infty\}, \quad (a,\infty] := \{x \in \overline{\mathbb{R}} : a < x \leqslant \infty\}$$

und analog $[a,\infty)$, $[a,\infty]$, $(-\infty,a)$, $[-\infty,a)$, $[-\infty,a]$ usw. Eine Menge $U \subset \overline{\mathbb{R}}$ heißt ***Umgebung von*** ∞ ***in*** $\overline{\mathbb{R}}$ (bzw. $-\infty$ in $\overline{\mathbb{R}}$) falls es $M \in \mathbb{R}$ gibt, so dass $(M,\infty] \subset U$ (bzw. $[-\infty,M) \subset U$).

(iv) Eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ heißt ***divergent in*** $\mathbb{R}$ ***gegen*** ∞ (bzw. $-\infty$) wenn jede Umgebung von ∞ (bzw. $-\infty$) fast alle Folgenglieder a_n enthält, d. h. zu jedem $M > 0$ gibt es ein $n_0 \in \mathbb{N}$, so dass für alle $n \geqslant n_0$ gilt: $a_n > M$ (bzw. $a_n < -M$). Wir schreiben

$$\lim_{n\to\infty} a_n = \infty \quad \text{oder } a_n \to \infty,\, n \to \infty$$

und analog für $-\infty$. Die Folge heißt in diesem Fall auch ***bestimmt divergent in*** $\mathbb{R}$. Man nennt ∞ (bzw. $-\infty$) den ***uneigentlichen Grenzwert*** von $(a_n)_{n\in\mathbb{N}}$.

(v) Eine folge heißt ***konvergent in*** $\overline{\mathbb{R}}$ wenn sie konvergent in $\mathbb{R}$ ist oder divergent gegen ∞ oder $-\infty$ ist (bzw. $-\infty$) , geschrieben $\lim_{n\to\infty} a_n \in \overline{\mathbb{R}}$. Konvergiert eine Folge nicht in $\overline{\mathbb{R}}$, so heißt sie ***divergent in*** $\overline{\mathbb{R}}$ oder ***unbestimmt divergent***.

(vi) Eine Folge (a_n) in $\mathbb{R}$ hat ∞ (bzw. $-\infty$) als Häufungswert in $\overline{\mathbb{R}}$, falls für alle $M > 0$ gilt: $a_n \in (M,\infty)$ (bzw. $a_n \in (-\infty,-M)$) für unendlich viele $n \in \mathbb{N}$. Äquivalent dazu ist: (a_n) besitzt eine Teilfolge, die gegen ∞ (bzw. gegen $-\infty$) divergiert. Dann schreiben wir $\limsup a_n = \infty$ (bzw. $\liminf a_n = -\infty$).

2.3.9. **Beispiel.** (i) Ist $a_n > 0$ und $a_n \to 0$, $n \to \infty$, so gilt $a_n^{-1} \to \infty$, $n \to \infty$.

(ii) $\lim\limits_{n\to\infty} n^s = \infty$ für $s \in \mathbb{Q}_+$.

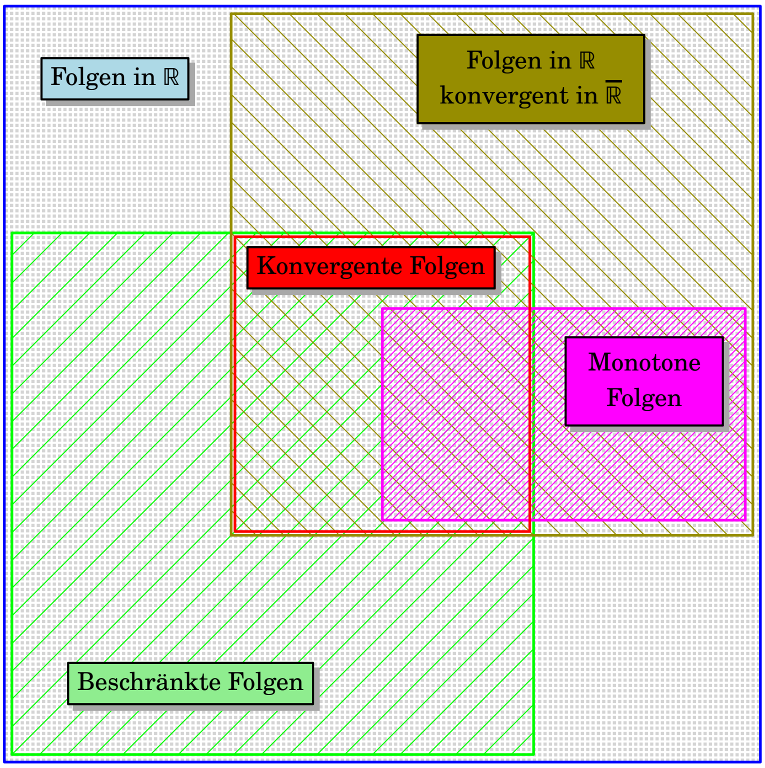
(iii) $\lim\limits_{n\to\infty} a^n = \infty$ für $a > 1$.

(iv) $\lim\limits_{n\to\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right) = \infty$.

(v) Für die Folge (a^n) mit $a < -1$ gilt $\limsup a_n = \infty$ (weil $a^{2n} \longrightarrow \infty$), $\liminf a_n = -\infty$ (weil $a^{2n+1} \longrightarrow -\infty$), also ist (a^n) mit $a < -1$ ist unbestimmt divergent.

Wir sehen, dass bei der Formulierung mit Hilfe des Umgebungsbegriffes die Definition 2.1.1 der Konvergenz in $\mathbb{R}$ und die Definition 2.3.8 der Konvergenz in $\overline{\mathbb{R}}$ übereinstimmen. Das ist kein Zufall und wird sich auch für die Konvergenz in $\mathbb{C}$ und $\overline{\mathbb{C}}$ wiederholen (siehe §2.4).

Wir erläutern noch die Beziehungen zwischen den verschiedenen Mengen von Folgen. Seien $\mathcal{F}$ = {Folgen in $\mathbb{R}$}, $\mathcal{K}_{\overline{\mathbb{R}}}$ = {Folgen in $\mathbb{R}$ konvergent in $\overline{\mathbb{R}}$}, $\mathcal{K}_{\mathbb{R}}$ = {konvergente Folgen in $\mathbb{R}$}, $\mathcal{B}$ = {beschränkte Folgen in $\mathbb{R}$}, $\mathcal{M}$ = {monotone Folgen in $\mathbb{R}$}, $\mathcal{C}$ = {Cauchy-Folgen in $\mathbb{R}$}. Dann gilt: $\mathcal{K}_{\mathbb{R}} \subset \mathcal{B}$, $\mathcal{K}_{\mathbb{R}} \subset \mathcal{K}_{\overline{\mathbb{R}}}$, $\mathcal{K}_{\mathbb{R}} = \mathcal{B} \cap \mathcal{K}_{\overline{\mathbb{R}}}$, $\mathcal{M} \subset \mathcal{K}_{\overline{\mathbb{R}}}$, $\mathcal{B} \cap \mathcal{M} \subset \mathcal{K}_{\mathbb{R}}$, $\mathcal{C} = \mathcal{K}_{\mathbb{R}}$.



2.4. **Folgen komplexer Zahlen.** Wir erhalten die Definition der Konvergenz in $\mathbb{C}$ durch Übertragung der reellen Definition.

2.4.1. **Definition.**
(i) Eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ heißt ***konvergent***, wenn es $a \in \mathbb{C}$ gibt mit der Eigenschaft: Zu jedem $\varepsilon > 0$ gibt es $n_0 = n_0(\varepsilon) \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geqslant n_0$ gilt: $|a_n - a| < \varepsilon$.
Eine Zahl $a \in \mathbb{C}$ wie in der Definition heißt ***Grenzwert*** oder ***Limes*** der Folge. Wir schreiben „$\lim_{n\to\infty} a_n = \lim a_n = a$" oder „$a_n \to a$ für $n \to \infty$" und sagen, dass $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert.
(ii) Eine nicht konvergente Folge heißt ***divergent***.

Sei $a \in \mathbb{C}$, $\varepsilon > 0$. Die Menge $B_\varepsilon(a) = \{z \in \mathbb{C} : |z - a| < \varepsilon\}$ heißt ***ε-Umgebung*** von a. Jede Teilmenge $U \subset \mathbb{C}$, zu der es ein $\varepsilon > 0$ mit $B_\varepsilon(a) \subset U$ gibt, heißt ***Umgebung*** von a.
Umformulierung der Definition 2.4.1 der Konvergenz:

$$\lim_{n\to\infty} a_n = a \iff \text{Zu jeder Umgebung } U \text{ von } a \text{ gibt es } n_0 = n_0(U) \in \mathbb{N}, \text{ so dass für alle } n \geqslant n_0 \text{ gilt: } a_n \in U.$$

2.4.2. **Beispiel.** (i) Ist $z \in \mathbb{C}$, $|z| < 1$, so gilt $z^n \to 0$, $n \to \infty$.
(ii) Ist $z \in \mathbb{C}$, $|z| < 1$, so gilt $1 + z + \ldots + z^n \to \dfrac{1}{1-z}$, $n \to \infty$.

21.11.2011

Die Ungleichung $|\operatorname{Re} z|, |\operatorname{Im} z| \leqslant |z| \leqslant |\operatorname{Re} z| + |\operatorname{Im} z|$ (1.7.7(iii)&(v)) liefert:

2.4.3. **Satz.** *Sei $(a_n)_{n\in\mathbb{N}}$ Folge in $\mathbb{C}$. Dann ist $(a_n)_{n\in\mathbb{N}}$ konvergent genau dann, wenn $(\operatorname{Re} a_n)_{n\in\mathbb{N}}$ und $(\operatorname{Im} a_n)_{n\in\mathbb{N}}$ konvergent sind. In diesem Fall gilt:*

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \operatorname{Re} a_n + i \lim_{n\to\infty} \operatorname{Im} a_n .$$

Die Definitionen von beschränkten Folgen, von Häufungswerten in $\mathbb{C}$ und Cauchy-Folgen übertragen sich wörtlich aus der jeweiligen reellen Version (Definitionen 2.1.4, 2.3.1, 2.3.6).

2.4.4. **Definition.**
(i) $h \in \mathbb{C}$ heißt ***Häufungswert*** (HW) von $(a_n)_{n\in\mathbb{N}}$, wenn für jedes $\varepsilon > 0$ gilt $a_n \in B_\varepsilon(a)$ für unendlich viele $n \in \mathbb{N}$.

(ii) Eine Menge $A \subset \mathbb{C}$ heißt ***beschränkt*** wenn es $M > 0$ gibt, mit $|z| \leqslant M$ für alle $z \in A$. Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt ***beschränkt***, wenn die Menge $\{a_n : n \in \mathbb{N}\}$ beschränkt ist, d. h. wenn es $M > 0$ gibt, mit $|a_n| \leqslant M$ für alle $n \in \mathbb{N}$.

2.4.5. **Satz** (Cauchy-Kriterium in $\mathbb{C}$)**.** *Eine Folge in $\mathbb{C}$ konvergiert genau dann, wenn sie eine Cauchy-Folge ist.*

2.4.6. **Satz** (Bolzano-Weierstraß in $\mathbb{C}$)**.** *Jede beschränkte Folge in $\mathbb{C}$ besitzt einen Häufungswert.*

2.4.7. **Definition.**

(i) Wir ergänzen $\mathbb{C}$ durch ein (ideales) Element $\infty \notin \mathbb{C}$ und setzten $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. Die Menge $\overline{\mathbb{C}}$ heißt die ***erweiterte Zahlenebene*** . Geometrisch handelt es sich um die Menge, die sich aus der Hinzunahme eines Punktes in der Unendlichkeit zu der komplexen Ebene ergibt.

(ii) Eine Menge $U \subset \overline{\mathbb{C}}$ heißt ***Umgebung von*** ∞ ***in*** $\overline{\mathbb{C}}$, wenn es $M > 0$ gibt, so dass $U \supset \{|z| > M\}$.

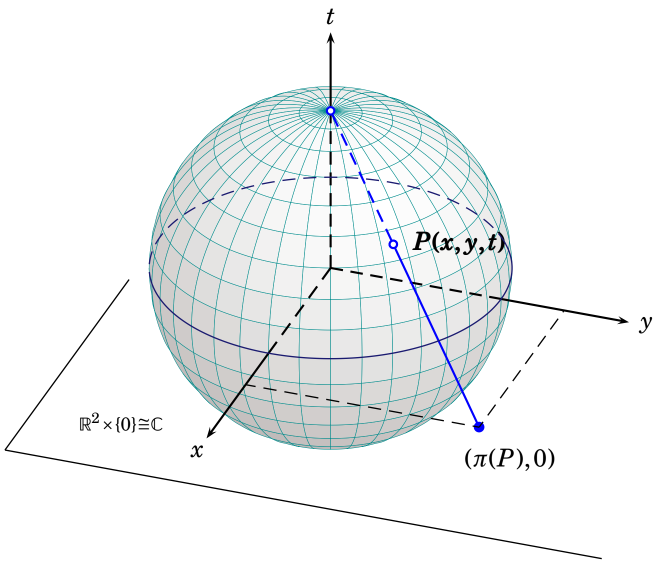
(iii) Eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ ***konvergiert in*** $\overline{\mathbb{C}}$ ***gegen*** ∞ : $\Longleftrightarrow$ Für jede Umgebung U von ∞ gilt $a_n \in U$ für fast alle $n \in \mathbb{N}$, $\Longleftrightarrow \lim_{n\to\infty} |a_n| = \infty$.
Schreibweise: $a_n \longrightarrow \infty$ oder $\lim_{n\to\infty} a_n = \infty$ in $\overline{\mathbb{C}}$.

2.4.8. **Beispiel.** (i) Ist $z \in \mathbb{C}$, $|z| > 1$, so gilt $z^n \to \infty \in \overline{\mathbb{C}}$, $n \to \infty$.

(ii) Vorsicht: Falls $a < -1$, dann $a^n \longrightarrow \infty \in \overline{\mathbb{C}}$, aber (a^n) ist unbestimmt divergent in $\overline{\mathbb{R}}$ (siehe Definition 2.3.8(v) und Beispiel 2.3.9(4)). Wenn aber eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\overline{\mathbb{R}}$ gegen $\infty \in \overline{\mathbb{R}}$ oder $-\infty \in \overline{\mathbb{R}}$ konvergiert, dann konvergiert sie auch in $\overline{\mathbb{C}}$ gegen $\infty \in \overline{\mathbb{C}}$. Wenn aber eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ gegen $\infty \in \overline{\mathbb{R}}$ oder $-\infty \in \overline{\mathbb{R}}$ divergiert, dann divergiert sie auch gegen $\infty \in \overline{\mathbb{C}}$.

Ein Modell für $\overline{\mathbb{C}}$ ist die Sphäre $S^2 = \{(x,y,t) \in \mathbb{R}^3 : x^2 + y^2 + t^2 = 1\}$.

Sei $\pi : S^2 \setminus \{(0,0,1)\} \longrightarrow \mathbb{C}$ die ***stereographische Projektion*** bezüglich des „Nordpols“ $(0,0,1)$: Die Abbildung π ordnet dem Punkt $(x,y,t) \in S^2 \setminus \{(0,0,1\}$ den Schnittpunkt der Geraden durch $(0,0,1)$ und (x,y,t) mit der Ebene $xOy \cong \mathbb{R}^2 \cong \mathbb{C}$ zu; π ist bijektiv, und daher ist auch die Fortsetzung $\overline{\pi} : S^2 \to \overline{\mathbb{C}}$ von π durch $\overline{\pi}(0,0,1) = \infty$ bijektiv und erlaubt es uns, S^2 mit $\overline{\mathbb{C}}$ zu identifizieren.



2.5. Übungen.

2.5.1. ***Aufgabe***. Sei $(a)_n$ eine konvergente Folge. Zeige, dass $\lim a_n$ eindeutig bestimmt ist.

2.5.2. ***Aufgabe***. Zeige:

$$\text{(i)} \quad \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k(k+1)} = 1,$$

$$\text{(ii)} \quad \lim_{n\to\infty} \frac{n^k}{n!} = 0, \quad k \in \mathbb{N} \text{ fest},$$

$$\text{(iii)} \quad \lim_{n\to\infty} \frac{a^n}{n!} = 0, \quad a \geqslant 0 \text{ fest},$$

(iv) $\lim\limits_{n\to\infty} \sqrt{n}(\sqrt{n+1}-\sqrt{n}) = \dfrac{1}{2}$.

2.5.3. ***Aufgabe***. Die Folge $(x_n)_{n\in\mathbb{N}}$ sei rekursiv definiert durch

$$x_1 = 1 \quad , \qquad x_{n+1} = \sqrt{2+x_n}\ .$$

(a) Beweise, dass die Folge $(x_n)_{n\in\mathbb{N}}$ streng monoton wachsend ist.
(b) Beweise, dass $(x_n)_{n\in\mathbb{N}}$ durch 2 nach oben beschränkt ist.
(c) Beweise, dass $(x_n)_{n\in\mathbb{N}}$ konvergent ist, und bestimme ihren Grenzwert.

2.5.4. ***Aufgabe***. Die Folge $(x_n)_{n\in\mathbb{N}}$ sei rekursiv definiert durch

$$x_1 = 2 \quad , \qquad x_{n+1} = 1+\sqrt{x_n}\ .$$

(a) Beweise, dass die Folge $(x_n)_{n\in\mathbb{N}}$ monoton wachsend ist.
(b) Beweise, dass $2 \leqslant x_n \leqslant 4$ für alle $n\in\mathbb{N}$ gilt.
(c) Beweise, dass $(x_n)_{n\in\mathbb{N}}$ konvergent ist, und bestimme ihren Grenzwert.

2.5.5. ***Aufgabe***. Sei $(a_n)_{n\in\mathbb{N}}$ die durch $a_1 := 1$, $a_{n+1} = \sqrt{1+a_n}$ rekursiv definierte Folge. Zeige, dass $(a_n)_{n\in\mathbb{N}}$ konvergiert und bestimme $\lim\limits_{n\longrightarrow\infty} a_n$.

2.5.6. ***Aufgabe***. Sei $k\in\mathbb{N}$ und seien $a>0$ und $x_1>0$ reelle Zahlen. Die Folge $(x_n)_{n\in\mathbb{N}}$ werde rekursiv definiert durch

$$x_{n+1} = \frac{1}{k}\left((k-1)x_n + \frac{a}{x_n^{k-1}}\right)\ ,\ n\in\mathbb{N}.$$

Zeige, dass die Folge konvergent ist und berechne ihren Grenzwert.

2.5.7. ***Aufgabe***. Sei $x_n = 1+\dfrac{1}{1!}+\ldots+\dfrac{1}{n!}$, $n\in\mathbb{N}$. Zeige:

(i) $x_n \geq \left(1+\dfrac{1}{n}\right)^n$ und $\lim x_n \geq e$. (Benutze Übung 1.8.8.)
(ii) Für $m\in\mathbb{N}$ mit $2\leq m<n$ gilt

$$\left(1+\frac{1}{n}\right)^n > 1+\frac{1}{1!}+\frac{1}{2!}\left(1-\frac{1}{n}\right)+\frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)+\ldots$$
$$+\frac{1}{m!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{m-1}{n}\right).$$

(iii) Folgere $e\geq x_m$ für alle $m\in\mathbb{N}$ und schließe $\lim x_n = e$.

2.5.8. ***Aufgabe***. Sei $x_n = 1+\dfrac{1}{1!}+\ldots+\dfrac{1}{n!}$.

(a) Zeige, dass für $m>n$ gilt $0<x_m-x_n<\dfrac{1}{n!\,n}$.
(b) Folgere $0<e-x_n\leq\dfrac{1}{n!\,n}$.
(c) Berechne e bis auf 10^{-3} genau.
(d) Zeige, dass e irrational ist. (Benutze (b) für einen Widerspruchsbeweis.)

2.5.9. ***Aufgabe***. (a) Für die Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n=\frac{1}{n}$ berechne $\sup_{k\geq n}x_k$, $\inf_{k\geq n}x_k$, $\limsup_{n\to\infty}x_n$, $\liminf_{n\to\infty}x_n$.

(b) Sei $(x_n)_{n\in\mathbb{N}}$ eine beschänkte Folge in $\mathbb{R}$. Zeige:

(i) $\liminf_{n\to\infty}x_n := \sup_{n\geq1}\inf_{k\geq n}x_k$ ist ein Häufungswert von $(x_n)_{n\in\mathbb{N}}$ und zwar der kleinste.
(ii) Die Folge $(x_n)_{n\in\mathbb{N}}$ ist konvergent genau dann wenn $\limsup_{n\to\infty}x_n=\liminf_{n\to\infty}x_n$. Ist das der Fall, so ist $\lim_{n\to\infty}x_n=\limsup_{n\to\infty}x_n=\liminf_{n\to\infty}x_n$.

2.5.10. ***Aufgabe***. Berechne die Häufungswerte der Folgen $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ definiert durch

$$a_n=\frac{n}{n+1}\sin\left(\frac{n\pi}{2}\right),\quad b_n=\left(1+\frac{\cos(n\pi)}{n}\right)^n.$$

2.5.11. ***Aufgabe***. Es seien (x_n) und (y_n) Folgen in $\mathbb{R}$ und

$$x_* := \liminf x_n,\; x^* := \limsup x_n,\; y_* := \liminf y_n,\; y^* := \limsup y_n.$$

Man verifiziere folgende Aussagen:

(a) $\limsup(-x_n) = -x_*$.

(b) $\limsup(x_n + y_n) \leqslant x^* + y^*$, $\liminf(x_n + y_n) \geqslant x_* + y_*$, falls (x_*, y_*) und (x^*, y^*) verschieden sind von $(\infty, -\infty)$ und $(-\infty, \infty)$.

(c) Aus $x_n \geqslant 0$ und $y_n \geqslant 0$ für $n \in \mathbb{N}$ folgt

$$0 \leqslant x_* y_* \leqslant \liminf(x_n y_n) \leqslant x_* y^* \leqslant \limsup(x_n y_n) \leqslant x^* y^*,$$

falls $(x_*, y_*) \notin \{(0,\infty),(\infty,0)\}$, $(x_*, y^*) \neq (\infty, 0)$, $(x^*, y^*) \neq (0,\infty)$ gelten.

(d) Konvergiert (y_n) in $\mathbb{R}$ gegen y, so folgen

$$\liminf(x_n + y_n) = x_* + y, \quad \limsup(x_n + y_n) = x^* + y,$$

und

$$\limsup(x_n y_n) = y x^*,\; y > 0, \quad \liminf(x_n y_n) = y x_*,\; y < 0.$$

(e) Aus $x_n \leqslant y_n$ für $n \in \mathbb{N}$ folgen $\liminf x_n \leqslant \liminf y_n$ und $\limsup x_n \leqslant \limsup y_n$.

2.5.12. ***Aufgabe*** (Satz von Cesaro-Stolz). (a) Sei (x_n) eine beliebige Folge und (y_n) eine monoton steigende Folge mit $\lim\limits_{n\to\infty} y_n = \infty$. Zeige: Existiert $\lim\limits_{n\to\infty} \dfrac{x_{n+1} - x_n}{y_{n+1} - y_n}$ in $\overline{\mathbb{R}}$, dann existiert $\lim\limits_{n\to\infty} \dfrac{x_n}{y_n}$ in $\overline{\mathbb{R}}$ und $\lim\limits_{n\to\infty} \dfrac{x_n}{y_n} = \lim\limits_{n\to\infty} \dfrac{x_{n+1} - x_n}{y_{n+1} - y_n}$.

(b) Studiere mit Hilfe von (a) die Konvergenz der folgenden Folgen (z_n):

(i) $z_n = \dfrac{a^n}{n}$, $a \geqslant 0$ fest.

(ii) $z_n = \dfrac{1^p + 2^p + \ldots + n^p}{n^{p+1}}$, $p \in \mathbb{N}$ fest.

(iii) $z_n = \dfrac{x_1 + x_2 + \ldots + x_n}{n}$, wobei $(x_n)_{n\in\mathbb{N}}$ einen Grenzwert in $\overline{\mathbb{R}}$ besitzt.

2.5.13. ***Aufgabe***. (i) Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge positiver Zahlen. Zeige

$$\liminf \frac{a_{n+1}}{a_n} \leqslant \liminf \sqrt[n]{a_n} \leqslant \limsup \sqrt[n]{a_n} \leqslant \limsup \frac{a_{n+1}}{a_n}\ .$$

(ii) Zeige mit Hilfe von (i), daß (a) $\lim\limits_{n\to\infty} \sqrt[n]{n!} = \infty$, (b) $\lim\limits_{n\to\infty} \dfrac{n}{\sqrt[n]{n!}} = e$, (c) $\lim\limits_{n\to\infty} \dfrac{\sqrt[n+1]{(n+1)!}}{\sqrt[n]{n!}} = 1$.

3. Reihen

3.1. Definitionen und Beispiele.

3.1.1. Definition.

(i) Sei $(a_n)_{n\geqslant q}$ eine Folge in $\mathbb{C}$. Dazu definiere $(s_n)_{n\geqslant q}$ durch $s_q = a_q$, $s_{q+1} = a_q + a_{q+1}, \ldots, s_n = \sum_{j=q}^n a_j$. Das Folgenpaar $((a_n)_{n\geqslant q}, (s_n)_{n\geqslant q})$ heißt ***Reihe*** mit ***Gliedern*** a_n und wird mit $\sum_{n\geqslant q} a_n$ bezeichnet. Die Zahl s_n heißt n-te ***Partialsumme*** der Reihe.

(ii) $\sum_{n\geqslant q} a_n$ heißt ***konvergent***, wenn $(s_n)_{n\geqslant q}$ konvergent ist. Gilt $s_n \longrightarrow s \in \mathbb{C}$, so schreibt man $s = \sum_{n=q}^\infty a_n$; die Zahl s heißt die ***Summe*** der Reihe.

(iii) Eine nicht konvergente Reihe heißt ***divergent***. Sind die a_n reell, und gilt $s_n \longrightarrow \infty$ (bzw. $-\infty$), so ist $\sum_{n\geqslant q} a_n$ bestimmt divergent mit $\sum_{n=q}^\infty a_n = \infty$ (bzw. $-\infty$).

3.1.2. Beispiele.

(i) Die ***geometrische Reihe*** $1 + z + z^2 + \ldots = \sum_{n\geqslant 0} z^n$ mit $z \in \mathbb{C}$. Für $z \neq 1$ gilt $s_n = 1 + z + z^2 + \ldots + z^n = \dfrac{1-z^{n+1}}{1-z}$. Für $|z| < 1$ ist $\lim_{n\to\infty} z^{n+1} = 0$, also $\sum_{n=0}^\infty z^n = \dfrac{1}{1-z}$.

(ii) $\sum_{n\geqslant 1} \dfrac{1}{n(n+1)} = \dfrac{1}{1\cdot 2} + \dfrac{1}{2\cdot 3} + \dfrac{1}{3\cdot 4} + \ldots$ ist konvergent, und es gilt $\sum_{n=1}^\infty \dfrac{1}{n(n+1)} = 1$.

(iii) $\sum_{n\geqslant 0} \dfrac{1}{n!}$ ist konvergent und $\sum_{n=0}^\infty \dfrac{1}{n!} = e$.

(iv) Die ***harmonische Reihe*** $\sum_{n\geqslant 1} \dfrac{1}{n} = 1 + \dfrac{1}{2} + \dfrac{1}{3} + \ldots$ ist divergent; $\sum_{n=1}^\infty \dfrac{1}{n} = \infty$.

(v) Die Reihe $\sum_{n\geqslant 0} (-1)^n = 1 - 1 + 1 - 1 \ldots$ ist unbestimmt divergent, da $s_n = 0$ für n ungerade und $s_n = 1$ für n gerade.

3.1.3. Satz (Rechenregeln). *Seien $\sum_{n=0}^\infty a_n = a$, $\sum_{n=0}^\infty b_n = b$ in $\mathbb{C}$. Dann gilt:*

(i) $\sum_{n=0}^\infty (a_n + b_n) = a + b$,
(ii) $\sum_{n=0}^\infty \mathrm{Re}(a_n) = \mathrm{Re}\, a$, $\sum_{n=0}^\infty \mathrm{Im}(a_n) = \mathrm{Im}\, a$,
(iii) $\sum_{n=0}^\infty \overline{a}_n = \overline{a}$.

Vorsicht: Aus der Konvergenz von $\sum_{n\geqslant 1} a_n$, $\sum_{n\geqslant 1} b_n$ folgt nicht die Konvergenz von $\sum_{n\geqslant 1} a_n b_n$! Beispiel: Die Reihe $\sum_{n\geqslant 1} \frac{(-1)^n}{\sqrt{n}}$ konvergiert nach Leibniz-Kriterium, aber $\sum_{n\geqslant 1} \frac{(-1)^n}{\sqrt{n}} \cdot \frac{(-1)^n}{\sqrt{n}} = \sum_{n\geqslant 1} \frac{1}{n}$ divergiert.

3.2. Konvergenzkriterien.

3.2.1. Satz (Cauchy-Kriterium). *$\sum_{n\geqslant q} a_n$ ist konvergent genau dann, wenn es zu jedem $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{Z}(\geqslant q)$ gibt, so dass für alle $n, m \in \mathbb{N}$ mit $n > m \geqslant n_0$ gilt: $|a_{m+1} + a_{m+2} + \ldots + a_n| < \varepsilon$.*

3.2.2. Folgerung.

(i) *$\sum_{n\geqslant 0} a_n$ konvergent $\Rightarrow \lim_{n\to\infty} a_n = 0$. (Notwendiges Kriterium für Konvergenz)*
(ii) *Das Ändern endlich vieler a_n ändert nicht das Konvergenzverhalten von $\sum_{n\geqslant 0} a_n$ (wohl aber ggf. die Summe).*

3.2.3. Satz (Majorantenkriterium). *Ist $|a_n| \leqslant b_n$ für alle $n \geqslant q$ und konvergiert $\sum_{n\geqslant q} b_n$, so konvergieren auch $\sum_{n\geqslant q} |a_n|$ und $\sum_{n\geqslant q} a_n$, und es gilt:*

$$\Big|\sum_{n=q}^\infty a_n\Big| \leqslant \sum_{n=q}^\infty |a_n| \leqslant \sum_{n=q}^\infty b_n.$$

Die Reihe $\sum_{n\geqslant q} b_n$ heißt Majorante für $\sum_{n\geqslant q} a_n$.
Beispiel: Sei $a_n \leqslant 1$ für alle $n \in \mathbb{N}_0$; $z \in \mathbb{C}$, $|z| < 1 \Rightarrow \sum_{n\geqslant 0} z^n$ ist Majorante von $\sum a_n z^n \Rightarrow \sum_{n\geqslant 0} a_n z^n$ konvergiert und $|\sum_{n=0}^\infty a_n z^n| \leqslant \sum_{n=0}^\infty |z|^n = \frac{1}{1-|z|}$.

3.2.4. Folgerung (Minorantenkriterium). *Wenn $a_n \leqslant b_n$ für alle $n \geqslant q$ und $\sum_{n=q}^\infty a_n = \infty$, dann $\sum_{n=q}^\infty b_n = \infty$.*

3.2.5. Satz (Monotoniekriterium). *Sei $\sum_{n\geqslant q} a_n$ mit $a_n \geqslant 0$ für alle $n \in \mathbb{N}$. Dann ist $\sum_{n\geqslant q} a_n$ konvergent genau dann, wenn die Folge (s_n) beschränkt ist.*

3.2.6. **Beispiel.** Sei $s \in \mathbb{Q}$, $s > 1$ und $a_n = \frac{1}{n^s} > 0$ für $n \geqslant 1$. Dann gilt $s_n \leqslant \frac{1}{1-(\frac{1}{2})^{s-1}}$. Aus dem Monotoniekriterium folgt, dass $\sum_{n\geqslant 1} \frac{1}{n^s}$ konvergiert. Die Funktion

$$\boxed{\sum_{n=1}^{\infty} \frac{1}{n^s} =: \zeta(s)} \tag{3.1}$$

heißt die ***Riemannsche Zeta-Funktion***.
Sei $s \in \mathbb{Q}$, $s \leqslant 1$. Aus $\frac{1}{n^s} \geqslant \frac{1}{n}$ und $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$ folgt nach dem Minorantenkriterium $\sum_{n=1}^{\infty} \frac{1}{n^s} = \infty$. Die Zeta-Funktion ist also auf $(1, \infty)$ definiert.

Eine Reihe der Form $\sum_{n\geqslant 0}(-1)^n a_n$ oder $\sum_{n\geqslant 0}(-1)^{n+1} a_n$ mit $a_n > 0$ heißt ***alternierend***. Die Glieder haben abwechselndes Vorzeichen.

3.2.7. **Satz** (Leibniz-Kriterium)**.** *Sei (a_n) eine monoton fallende Nullfolge. Dann ist die Reihe $\sum_{n\geqslant 0}(-1)^n a_n = a_0 - a_1 + a_2 + \ldots$ konvergent gegen $s \in \mathbb{R}$.*

Restabschätzung: $|s - s_n| = |\sum_{k=n+1}^{\infty}(-1)^k a_k| \leqslant a_{n+1}$ für alle $n \geqslant 0$.

Zusatz: $([s_{2k+1}, s_{2k}])_{k\geqslant 0}$ bildet eine Intervallschachtelung mit $\cap_{k\geqslant 0}[s_{2k+1}, s_{2k}] = \{s\}$, d.h. $(s_{2k})_{k\geqslant 0}$ bildet eine monoton fallende Folge, $(s_{2k+1})_{k\geqslant 0}$ bildet eine monoton wachsende Folge und beide konvergieren gegen s; es gilt $s_1 \leqslant s_3 \leqslant \ldots \leqslant s \leqslant \ldots \leqslant s_2 \leqslant s_0$.

3.2.8. **Beispiel.** Die ***alternierende harmonische Reihe*** $\sum_{n\geqslant 1} \frac{(-1)^{n+1}}{n}$ ist konvergent. Ihre Summe ist $\log 2$ (siehe (7.5)).

24.11.2011

3.3. Absolute Konvergenz.

3.3.1. **Definition.** $\sum_{n\geqslant q}^{\infty} a_n$ heißt ***absolut konvergent***, falls $\sum_{n\geqslant q}^{\infty} |a_n|$ konvergiert.

3.3.2. **Bemerkung.**

(i) $\sum_{n\geqslant q}^{\infty} |a_n|$ ist konvergent genau dann, wenn $(\sum_{k=q}^{n} |a_k|)_{n\geqslant q}$ beschränkt ist (Monotoniekriterium 3.2.5).

(ii) Ist $\sum_{n\geqslant q}^{\infty} |a_n|$ konvergent, so konvergiert nach Majorantenkriterium 3.2.3 auch $\sum_{n\geqslant q} a_n$, und $|\sum_{n=q}^{\infty} a_n| \geqslant \sum_{n=q}^{\infty} |a_n|$.

(iii) Aus der Konvergenz von $\sum_{n\geqslant q} a_n$ folgt nicht die Konvergenz von $\sum_{n\geqslant q} |a_n|$.

3.3.3. **Satz** (Wurzelkriterium)**.**

(i) *Ist $l = \limsup \sqrt[n]{|a_n|} < 1$, so konvergiert $\sum a_n$ absolut.*

(ii) *Ist $\sqrt[n]{|a_n|} \geqslant 1$ für unendlich viele n (das tritt auf z.B. wenn $l > 1$), so ist $\sum a_n$ divergent.*

3.3.4. **Satz** (Quotientenkriterium)**.** *Es gelte $a_n \neq 0$ für fast alle $n \in \mathbb{N}$.*

(i) *Ist $l = \limsup \frac{a_{n+1}}{a_n} < 1$, so ist $\sum a_n$ absolut konvergent.*

(i) *Ist $\frac{a_{n+1}}{a_n} \geqslant 1$, für fast alle $n \in \mathbb{N}$, so ist $\sum a_n$ divergent.*

3.3.5. **Beispiel.** (i) Die Reihe

$$\sum_{n\geqslant 0} \frac{z^n}{n!} \tag{3.2}$$

heißt ***Exponentialreihe***. Es gilt $a_n = \frac{z^n}{n!}$. Für $z \neq 0$ ist $a_n \neq 0$ und $\frac{a_{n+1}}{a_n} = \frac{z}{n} \to 0$, $n \to \infty$ also konvergiert die Exponentialreihe. Sie konvergiert offensichtlich auch für $z = 0$. Die Summe der Exponentialreihe wird ***Exponentialfunktion*** und wird bezeichnet durch

$$\exp : \mathbb{C} \to \mathbb{C}, \quad \exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}. \tag{3.3}$$

Es gilt

$$\exp(0) = 1, \quad \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!} = e. \tag{3.4}$$

wobei die zweite Gleichcheit aus Beispiel 3.1.2 (iii) folgt. Für $x \in \mathbb{R}$ ist

$$\exp(x) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^n}{n!} \in \mathbb{R},$$

also definiert Einschränkung der Exponentialfunktion exp auf $\mathbb{R}$ eine Funktion $\exp : \mathbb{R} \longrightarrow \mathbb{R}$.

(ii) $\sum_{n\geqslant 1} nz^n$, $z\in\mathbb{C}$; $a_n = nz^n$, $\sqrt[n]{|a_n|} = \sqrt[n]{n|z|^n} = \sqrt[n]{n}|z| \longrightarrow |z|$, $n\to\infty$. Nach dem Wurzelkriterium folgt: Für $|z|<1$ ist die Reihe absolut konvergent; für $|z|\geqslant 1$ ist die Reihe divergent ($\sqrt[n]{n|z|^n} = \sqrt[n]{n}|z| \geqslant 1$).

(iii) $\sum_{n\geqslant 1} \frac{z^n}{n}$, $z\in\mathbb{C}$; $a_n = \frac{z^n}{n}$, $\lim|\frac{a_{n+1}}{a_n}| = \lim\frac{|z|^{n+1}}{n+1}\cdot\frac{n}{|z|^n} = |z|$. Für $|z|<1$ ist die Reihe absolut konvergent und für $|z|>1$ divergent. Ist $|z|=1$, so gilt: Für $z=1$ ist die Reihe divergent, für $z\neq 1$ konvergent.

(iv) Es reicht für absolute Konvergenz *nicht*, dass $\limsup\sqrt[n]{|a_n|}=1$ oder $\limsup|\frac{a_{n+1}}{a_n}|=1$ (oder $\sqrt[n]{|a_n|}<1$ oder $\frac{a_{n+1}}{a_n}<1$ für fast alle $n\in\mathbb{N}$) gilt. Die harmonische Reihe erfüllt all diese Bedingungen, ist aber divergent.

3.3.6. **Satz** (Cauchy-Produkt). *Seien $\sum_{k\geqslant 0} a_k$ und $\sum_{k\geqslant 0} b_k$ absolut konvergent. Setze*

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = a_0 b_n + a_1 b_{n-1} + \ldots + a_n b_0\,.$$

Dann ist $\sum_{n\geqslant 0} c_n$ konvergent, und

$$\sum_{n=0}^{\infty} c_n = \Big(\sum_{k=0}^{\infty} a_k\Big)\Big(\sum_{k=0}^{\infty} b_k\Big).$$

3.3.7. **Beispiele.** (i) Sei $|z|<1$. Dann gilt: $(\sum_{k=0}^{\infty} z^k)^2 = \frac{1}{(1-z)^2}$. Aus 3.3.6 folgt:

$$\Big(\sum_{k=0}^{\infty} z^k\Big)^2 = \Big(\sum_{k=0}^{\infty} z^k\Big)\Big(\sum_{l=0}^{\infty} z^l\Big) = \sum_{n=0}^{\infty}\big(z^0 z^n + \ldots + z^n z^0\big) = \sum_{n=0}^{\infty}(n+1)z^n = \sum_{k=1}^{\infty} k z^{k-1}.$$

Also $\displaystyle\sum_{k=1}^{\infty} k z^{k-1} = \frac{1}{(1-z)^2}$.

(ii) Seien $z,w\in\mathbb{C}$. Betrachte die Exponentialreihen $\sum_{k\geqslant 0} a_k = \sum_{k\geqslant 0}\frac{z^k}{k!}$ und $\sum_{k\geqslant 0} b_k = \sum_{k\geqslant 0}\frac{w^k}{k!}$. Dann gilt nach der Binomialformel

$$c_n = \sum_{k=0}^{n}\frac{z^k}{k!}\cdot\frac{w^{n-k}}{(n-k)!} = \frac{1}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!} z^k w^{n-k} = \frac{(z+w)^n}{n!}.$$

Das Cauchy-Produkt $\sum_{n\geqslant 0} c_n$ ist also wieder eine Exponentialreihe, nämlich $\sum_{n\geqslant 0}\frac{(z+w)^n}{n!}$. Nach dem Satz über Cauchy-Produkt 3.3.6 gilt

$$\boxed{\exp(z+w) = \exp(z)\exp(w),\quad z,w\in\mathbb{C}} \tag{3.5}$$

Die Gleichung (3.5) heißt die ***Funktionalgleichung der Exponentialfunktion***.

28.11.2011

Der Name zeigt, dass es um eine Gleichung handelt, deren Unbekannte eine Funktion $\varphi:\mathbb{C}\to\mathbb{C}$ ist, die die folgende Gleichung erfüllt

$$\varphi(z+w) = \varphi(z)\varphi(w),\quad z,w\in\mathbb{C} \tag{3.6}$$

Wir kennen schon eine Funktion die (3.6) für $z,w\in\mathbb{Q}$ erfüllt, nämlich die Exponentialfunktion $\mathbb{Q}\ni r\mapsto a^r$, wobei $a>0$. In der Tat, die Potenzgesetze lauten $a^q a^r = a^{q+r}$, siehe (1.7) und Aufgabe 1.8.12. Wir zeigen nun, dass die Funktionalgleichung (3.6) durch die Festsezung $\varphi(1)=a$ eindeutig bestimmt ist.

3.3.8. **Satz.** *Es gibt für $a>0$ genau eine Lösung $\varphi:\mathbb{Q}\to\mathbb{R}$ der Gleichung* (3.6) *mit $\varphi(1)=a$, nämlich $\varphi(r)=a^r$, $r\in\mathbb{Q}$.*

Die Funktion $\exp:\mathbb{Q}\to\mathbb{R}$ erfüllt die Gleichung (3.6) und $\exp(1)=e$. Nach dem Satz 3.3.8 gilt also

$$e^r = \exp(r),\quad r\in\mathbb{Q}. \tag{3.7}$$

Die Gleichung (3.7) und die Tatsache, dass die Funktion exp die Funktionalgleichung (3.5) für alle $z,w\in\mathbb{C}$ erfüllt, eröffnet die Möglichkeit, die Ausdehnung der Potenz e^z für alle $z\in\mathbb{C}$ durchzuführen.

3.3.9. **Definition.** Die ***komplexen Potenzen*** e^z mit $z\in\mathbb{C}$ sind definiert durch

$$e^z := \exp(z),\quad z\in\mathbb{C}. \tag{3.8}$$

Um Potenzen a^z mit $a>0$, $z\in\mathbb{C}$ zu definieren, müssen wir warten, bis das Logarithmus definiert wird. Als Vorgeschmack lautet die Definition $a^z := e^{z\log a}$ (siehe Definition 4.3.5).

3.4. **Potenzreihen.**

3.4.1. **Definition.** Sei $\mathcal{R}$ die Menge der Reihen mit komplexen Gliedern. Eine ***Potenzreihe*** ist eine Abbildung $P:\mathbb{C}\longrightarrow\mathcal{R}$ der Form $P(z)=\sum_{n\geqslant 0}a_nz^n$. Die Zahlen a_n heißen ***Koeffizienten*** der Reihe.

Gilt $a_n\in\mathbb{R}$ für alle n, so spricht man von einer reellen Potenzreihe $P:\mathbb{R}\longrightarrow\mathcal{R}$, $P(x)=\sum_{n\geqslant 0}a_nx^n$. Die Theorie der reellen Potenzreihen ergibt sich aus der Theorie der komplexen Potenzreihen durch Einschränkung der Variablen $z\in\mathbb{C}$ auf $\mathbb{R}$.

3.4.2. **Lemma** (Abelsches Konvergenzlemma). *Zur Potenzreihe $\sum_{n\geqslant 0}a_nz^n$ gebe es Zahlen $s,M>0$ mit $|a_n|s^n\leqslant M$ für alle $n\in\mathbb{N}_0$. Dann ist $P(z)=\sum_{n\geqslant 0}a_nz^n$ absolut konvergent in der Kreisscheibe $B_s(0)$.*

3.4.3. **Folgerung.** *Konvergiert $\sum_{n\geqslant 0}a_nz^n$ in $z_0\neq 0$, so konvergiert sie absolut in $B_{|z_0|}(0)$.*

3.4.4. **Satz.** *Sei $\sum_{n\geqslant 0}a_nz^n$ eine Potenzreihe. Sei*

$$R:=\sup\big\{t\in[0,\infty):(|a_n|t^n)\ \textit{beschränkt}\big\}\in[0,\infty].$$

Dann gilt:

(i) Für alle $z\in B_R(0)$ ist die Reihe absolut konvergent.

(ii) Für alle $z\in\mathbb{C}\setminus\overline{B_R(0)}$ ist die Reihe divergent.

Dabei ist $\overline{B_R(0)}:=\{z\in\mathbb{C}:|z|\leqslant R\}$.

3.4.5. **Definition.** R heißt ***Konvergenzradius*** und $B_R(0)$ heißt ***Konvergenzkreis(scheibe)*** der Potenzreihe. Wir verabreden, dass $B_R(0):=\mathbb{C}$, falls $R=\infty$.

3.4.6. **Satz.** *Sei R der Konvergenzradius von $P(z)=\sum_{n\geqslant 0}a_nz^n$. Dann gilt:*

(1) $R=\frac{1}{L}$ mit $L:=\limsup\limits_{n\to\infty}\sqrt[n]{|a_n|}\in[0,\infty]$ (Cauchy-Hadamard)

(2) $R=\frac{1}{q}$ mit $q:=\lim\limits_{n\to\infty}\left|\dfrac{a_{n+1}}{a_n}\right|\in[0,\infty]$, falls $\left(\left|\dfrac{a_{n+1}}{a_n}\right|\right)$ einen Grenzwert in $\overline{\mathbb{R}}$ hat.

Dabei wird vereinbart: $\frac{1}{0}:=\infty$, $\frac{1}{\infty}:=0$.

3.4.7. **Beispiel.** (1) Die ***geometrische Reihe*** $\sum_{k\geqslant 0}z^k$. Dann gilt $a_k=1$ für alle $k\in\mathbb{N}_0$; $|\frac{a_{n+1}}{a_n}|=\frac{1}{1}=1\longrightarrow 1$. Der Konvergenzradius ist $R=1$.

(2) $\sum_{n\geqslant 1}n^{-\alpha}z^n$, $\alpha\in\mathbb{Q}$. Dann $\sqrt[n]{\frac{1}{n^\alpha}}=(\frac{1}{\sqrt[n]{n}})^\alpha\longrightarrow 1^\alpha=1$. Der Konvergenzradius ist $R=1$. Verhalten auf dem Rand $\{z\in\mathbb{C}:|z|=1\}$ des Konvergenzkreiseses:

- (i) $\sum_{n\geqslant 0}z^n$ divergiert, da (z^n) keine Nullfolge ist.
- (ii) $\sum_{n\geqslant 1}n^{-\alpha}z^n$ divergiert für $\alpha<1$, da $(\frac{z^n}{n^\alpha})$ keine Nullfolge ist. (Für $\alpha<1$ gilt $\frac{1}{n^\alpha}>1$.)
- (iii) $\sum_{n\geqslant 1}n^{-\alpha}z^n$ ist absolut konvergent für $\alpha>1$, da $|\frac{z^n}{n^\alpha}|=\frac{1}{n^\alpha}$, $\sum_{n\geqslant 1}\frac{1}{n^\alpha}$ konvergent ist.
- (iv) $\sum_{n\geqslant 1}n^{-\alpha}z^n=\sum_{n\geqslant 1}n^{-\alpha}z^n$ für $\alpha=1$; die Reihe divergiert für $z=1$, sie konvergiert (aber nicht absolut) für $|z|=1,z\neq 1$.

(3) ***Exponentialreihe***: $\sum_{n\geqslant 0}\frac{z^n}{n!}$, $z\in\mathbb{C}$. Es ist $|\frac{a_{n+1}}{a_n}|=\frac{n!}{(n+1)!}=\frac{1}{n+1}\longrightarrow 0$. Es folgt $R=\infty$.

(4) ***Logarithmusreihe***:

$$\sum_{n\geqslant 1}\frac{(-1)^{n+1}}{n}z^n=z-\frac{z^2}{2}+\frac{z^3}{3}-\dots. \tag{3.9}$$

Dann gilt $\sqrt[n]{|a_n|}=\sqrt[n]{\frac{1}{n}}\longrightarrow 1$. Der Konvergenzradius ist $R=1$.

(5) Für $s\in\mathbb{C}$, $k\in\mathbb{N}_0$ definiere

$$\binom{s}{k}=\frac{1}{k!}\prod_{j=0}^{k-1}(s-j),\ \text{ also }\binom{s}{0}=1,\ \binom{s}{k}=\frac{s(s-1)\cdots(s-k+1)}{k!}\ \text{ für } k\in\mathbb{N}.$$

Die ***Binomialreihe zum Exponenten*** s ist definiert durch

$$B_s(z):=\sum_{k\geqslant 0}\binom{s}{k}z^k=1+sz+\binom{s}{1}z^2+\dots \tag{3.10}$$

• Falls $s\in\mathbb{N}_0$, so ist $\binom{s}{k}=0$ für $k>s$, also $\sum_{j=0}^{k}\binom{s}{j}z^j=\sum_{j=0}^{s}\binom{s}{j}z^j=(1+z)^s$ für alle $k>s$ (die reihe bricht ab). Die PR hat Konvergenzradius $R=\infty$; ihre Summe beträgt $(1+z)^s$.

• Falls $s\notin\mathbb{N}_0$: $\left|\binom{s}{n+1}/\binom{s}{n}\right|=|\frac{s-n}{n+1}|\longrightarrow 1$ für $n\longrightarrow\infty$ also $R=1$.

1.12.2

3.4.8. **Definition und Satz.** *Die* **Cosinus-Reihe** *und* **Sinus-Reihe**

$$\sum_{n\geqslant 0}\frac{(-1)^n}{(2n)!}z^{2n},\quad \sum_{n\geqslant 0}\frac{(-1)^n}{(2n+1)!}z^{2n+1}$$

haben Konvergenzradius $R=\infty$ *und definieren die* **Cosinus-** *und* **Sinus-Funktion**

$$\cos:\mathbb{C}\to\mathbb{C},\ \cos z=\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}z^{2n},\quad \sin:\mathbb{C}\to\mathbb{C},\ \sin z=\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}z^{2n+1} \tag{3.11}$$

3.4.9. **Satz** (Eulersche Formel)**.** *Für alle* $z\in\mathbb{C}$ *gilt*

$$\cos z+i\sin z=e^{iz} \tag{3.12}$$

Die Cosinus-Funktion ist *gerade* d. h. $\cos(-z)=\cos z$ und ie Sinus-Funktion ist *ungerade* d. h. $\sin(-z)=-\sin z$. So gilt $\cos z-i\sin z=e^{-iz}$. Durch Addition mit (3.12) und Subtraktion aus (3.12) folgen weiteren Eulerschen Formeln:

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

3.5. **Übungen.**

3.5.1. ***Aufgabe***. Berechne

$$\text{(a)}\ \sum_{k=0}^{\infty}\frac{(-1)^k}{3^k},\quad \text{(b)}\ \sum_{k=2}^{\infty}\frac{1}{4^{k-1}},\quad \text{(c)}\ \sum_{k=1}^{\infty}\frac{4}{5^k},\quad \text{(d)}\ \sum_{k=1}^{\infty}\frac{k^2+2k+5}{k!}.$$

Zeige

$$\text{(e)}\ \sum_{k=1}^{\infty}\frac{1}{4k^2-1}=\frac{1}{2},\qquad \text{(f)}\ \sum_{k=1}^{\infty}\frac{1}{k(k+1)(k+2)}=\frac{1}{4}.$$

3.5.2. ***Aufgabe***. Untersuche $\sum\limits_{n\geqslant 1}a_n$ auf Konvergenz und absolute Konvergenz, wobei a_n definiert sei als:

$$\text{(a)}\ \frac{1}{n^2}+\frac{(-1)^n}{n}\quad,\quad \text{(b)}\ \frac{4n^2+2n-3}{3n^4-n^3+7}\quad,\quad \text{(c)}\ \frac{n^n}{4^n\cdot n!}\quad,\quad \text{(d)}\ \frac{\sqrt{n+1}-\sqrt{n}}{n}\quad.$$

3.5.3. ***Aufgabe***. Untersuche folgende Reihen auf Konvergenz:

$$\text{(i)}\qquad \frac{1}{2}+\frac{1}{3}+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{2^3}+\frac{1}{3^3}+\ldots+\frac{1}{2^n}+\frac{1}{3^n}+\ldots$$

$$\text{(ii)}\qquad 1+a+ab+a^2b+a^2b^2+a^3b^2+\ldots+a^nb^n+a^{n+1}b^n+\ldots$$

3.5.4. ***Aufgabe***. Sei $x\in\mathbb{R}$, $x\geqslant 0$. Zeige, dass es eine eindeutig bestimmte Folge $(a_n)_{n\geqslant 0}$ gibt mit:

(i) $a_0\in\mathbb{N}_0$, $a_n\in\{0,1,2,\ldots,9\}$ für $n\in\mathbb{N}$;
(ii) für alle $n\in\mathbb{N}$ gilt $0\leqslant x-\sum\limits_{k=0}^{n}\frac{a_k}{10^k}<\frac{1}{10^n}$.

Tipp: Wähle induktiv $a_0=\lfloor x\rfloor$, $a_1=\lfloor 10(x-a_0)\rfloor$, $a_2=\lfloor 10^2(x-a_0-\frac{a_1}{10})\rfloor$, …, $a_n=\lfloor 10^n(x-a_0-\frac{a_1}{10}-\cdots\frac{a_{n-1}}{10^{n-1}})\rfloor$, …

Zeige, dass $x=\sum\limits_{k=0}^{\infty}\frac{a_k}{10^k}$. Man schreibt dann $x=a_0,a_1a_2a_3\ldots$.
Diese Darstellung heißt die *Dezimalbruchentwicklung* von x.

3.5.5. ***Aufgabe***. Seien (f_n) die *Fibonacci-Folge*, rekursiv definiert durch $f_1=f_2=1$ und $f_{n+1}=f_n+f_{n-1}$ für $n\geq 2$ und $g=\frac{1}{2}(1+\sqrt{5})$ die *Zahl des goldenen Schnittes*. Außerdem sei (x_n) rekursiv definiert durch $x_1=1$ und $x_{n+1}=1+\frac{1}{x_n}$ für $n\geq 1$. Sei g wie oben.

(a) Zeige induktiv $f_{n+1}-f_ng=(-1)^ng^{-n}$ und folgere $\frac{f_{n+1}}{f_n}\to g$.

(b) Zeige induktiv $x_n=\frac{f_{n+1}}{f_n}$ und folgere $x_n\to g$.

(c) Zeige $\sum_{k=1}^{\infty}\frac{1}{f_kf_{k+2}}=1$.

Tipp zu (c): $\frac{1}{f_kf_{k+1}}-\frac{1}{f_{k+1}f_{k+2}}=?$

3.5.6. ***Aufgabe***. Sei $z \in \mathbb{C}$. Zeige:

(i) Für jedes $\varepsilon > 0$ gibt es $N \in \mathbb{N}$, so dass $\sum_{k=N+1}^{\infty} \frac{|z|^k}{k!} < \frac{\varepsilon}{3}$.

Zudem gilt $\left| \sum_{k=N+1}^{n} \binom{n}{k} \frac{z^k}{n^k} \right| < \frac{\varepsilon}{3}$ für jedes $n > N$.

(ii) Für jedes $N \in \mathbb{N}$ gilt $\sum_{k=0}^{N} \left(\frac{z^k}{k!} - \binom{n}{k} \frac{z^k}{n^k} \right) \longrightarrow 0$ für $n \to \infty$.

Folgere: $\left(1 + \frac{z}{n}\right)^n \longrightarrow \exp(z)$ für $n \to \infty$.

3.5.7. ***Aufgabe***. Zu jedem $\alpha \in \mathbb{C}$ sei eine Folge (α_n) definiert durch

$$\alpha_0 = 1 \quad \text{und} \quad \alpha_n = \alpha(\alpha+1) \cdot \ldots \cdot (\alpha+n-1) \quad \text{für } n \in \mathbb{N}.$$

Seien $a, b, c \in \mathbb{C} \setminus \{0, -1, -2, \ldots\}$. Die *hypergeometrische Reihe* zu a, b, c ist definiert durch $F_{a,b,c}(z) := \sum_{k \geqslant 0} \frac{a_k b_k}{k!\, c_k} z^k = 1 + \frac{ab}{c} z + \frac{a(a+1)b(b+1)}{2!\, c(c+1)} z^2 + \ldots$

(a) Berechne den Konvergenzradius von $F_{a,b,c}$.

(b) Drücke die folgenden Potenzreihen in der Form $F_{a,b,c}(\pm z)$ oder $z \cdot F_{a,b,c}(\pm z)$ aus: (i) $\sum_{k \geqslant 0} z^k$,

(ii) $B_s(z) = \sum_{k \geqslant 0} \binom{s}{k} z^k$ mit $s \in \mathbb{C} \setminus \mathbb{N}_0$,

(iii) $L(z) = \sum_{k \geqslant 0} \frac{(-1)^k}{k+1} z^{k+1}$.

3.5.8. ***Aufgabe***. (i) Sei $a > 0$. Zeige, dass es eine einzige Funktion $\varphi : \mathbb{Q} \to \mathbb{R}$ gibt mit $\varphi(1) = a$ und $\varphi(r+s) = \varphi(r) \cdot \varphi(s)$ für alle $r, s \in \mathbb{Q}$.

(ii) Für $z \in \mathbb{C}$ mit $|z| < 1$ und $\alpha \in \mathbb{R}$ definiere die Binomialreihe (siehe)

$$B_\alpha(z) = \sum_{n=0}^{\infty} \binom{\alpha}{n} z^n .$$

Zeige mit Hilfe des Cauchy–Produkts, dass $B_{\alpha+\beta}(z) = B_\alpha(z) B_\beta(z)$ für alle $\alpha, \beta \in \mathbb{R}$. Hinweis: Zeige zunächst, dass für $\alpha, \beta \in \mathbb{R}$

$$\sum_{k=0}^{n} \binom{\alpha}{k} \binom{\beta}{n-k} = \binom{\alpha+\beta}{n}.$$

(iii) Leite für $\alpha \in \mathbb{Q}$ und $x \in \mathbb{R}$, $|x| < 1$, her:

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n .$$

3.5.9. ***Aufgabe***. Sei $(p_j)_{j \in \mathbb{N}}$ die Folge der Primzahlen $2, 3, 5, 7, 11, \ldots$, und sei $s \in \mathbb{Q}$ mit $s > 1$. Zeigen Sie:

(a) Mit $a_{\ell m} := \prod_{j=1}^{\ell} \sum_{k=0}^{m} p_j^{-ks}$ gilt $a_{\ell m} \leq \zeta(s)$ für alle $\ell, m \in \mathbb{N}$.

(b) Zu festem ℓ konvergiert $a_{\ell m}$ für $m \to \infty$ gegen einen Wert $a_\ell \leq \zeta(s)$, und für $\ell \to \infty$ konvergiert a_ℓ gegen einen Wert $a \leq \zeta(s)$. (Tipp: Monotonieprinzip)

(c) Zu jedem $\epsilon > 0$ gibt es $\ell, m \in \mathbb{N}$ mit $a_{\ell m} > \zeta(s) - \epsilon$, und es folgt $a \geq \zeta(s)$. Es gilt also die ***Eulersche Produktdarstellung*** der Riemannschen Zeta-Funktion:

$$\zeta(s) = \prod_{j=1}^{\infty} \frac{1}{1 - p_j^{-s}} := \lim_{\ell \to \infty} \prod_{j=1}^{\ell} \frac{1}{1 - p_j^{-s}}$$

3.5.10. ***Aufgabe***. Eine Potenzreihe $\sum_{n \geqslant 0} a_n z^n$ hat Konvergenzradius $R > 0$ genau dann, wenn es $q > 0$ mit $|a_n| \leqslant q^n$ für alle $n \in \mathbb{N}_0$ gibt.

3.5.11. ***Aufgabe***. Zeige die Additionstheoreme für alle $z, w \in \mathbb{C}$:

(a) $\cos(z+w) = \cos z \cos w - \sin z \sin w$, $\sin(z+w) = \cos z \sin w + \sin z \cos w$.
(Tipp: $\cos u + i \sin u = \exp(iu)$; betrachten Sie $u = z + w$ und $u = -(z+w)$.)

(b) $\cos^2 z + \sin^2 z = 1$, $\cos 2z = 2\cos^2 z - 1$, $\cos^2\left(\frac{z}{2}\right) = \frac{1+\cos z}{2}$, $\sin^2\left(\frac{z}{2}\right) = \frac{1-\cos z}{2}$.

(c) $\cos z + \cos w = 2 \cos \frac{z+w}{2} \cos \frac{z-w}{2}$, $\sin z + \sin w = 2 \sin \frac{z+w}{2} \cos \frac{z-w}{2}$.

3.5.12. ***Aufgabe***. Die ***Cosinus hyperbolicus-Reihe*** und ***Sinus hyperbolicus-Reihe***

$$\sum_{n\geqslant 0}\frac{1}{(2n)!}z^{2n}\,,\quad \sum_{n\geqslant 0}\frac{1}{(2n+1)!}z^{2n+1}$$

haben Konvergenzradius $R=\infty$ und definieren die ***Cosinus hyperbolicus-*** und ***Sinus hyperbolicus-Funktion***

$$\cosh z=\sum_{n=0}^{\infty}\frac{1}{(2n)!}z^{2n}\quad \sinh z=\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}z^{2n+1}.$$

Zeige für $z\in\mathbb{C}$, dass

$$\cosh z+\sinh z=e^{z}\,,$$

$$\sinh z=\frac{e^{z}-e^{-z}}{2}\,,\quad \cosh z=\frac{e^{z}+e^{-z}}{2}\,,$$

$$\cosh^2 z-\sinh^2 z=1\,,$$

$$\cosh z=\cos(iz)\,,\quad \sinh z=-i\sin(iz)\,.$$

Die Identität $\cosh^2 z-\sinh^2 z=1$ erklärt das Adjektiv „hyperbolisch“, wenn man sich an die reelle Hyperbelgleichung erinnert.
Zeige die Additionstheoreme für alle $z,w\in\mathbb{C}$:

$$\cosh(z+w)=\cosh z\cosh w+i\sinh z\sinh w\,,\quad \sinh(z+w)=\sinh z\cosh w+\cosh z\sinh w\,.$$

Zeige die Zerlegungen in Real- und Imaginärteil der trigonomotrischen Funktionen:

$$\cos(x+iy)=\cos x\cosh y-i\sin x\sinh y\,,\quad \sin(x+iy)=\sin x\cosh y+i\cos x\sinh y\,,\quad x,y\in\mathbb{R}\,.$$

3.6. **Notizen.** Wir sammeln in diesen Abschnitt Kommentare über das Produkt von Reihen, Umordnungen und summierbaren Familien.

Notiz zum Produkt von Reihen. Vorsicht: Aus der Konvergenz von $\sum_{n\geqslant 1}a_n$, $\sum_{n\geqslant 1}b_n$ folgt nicht die Konvergenz von $\sum_{n\geqslant 1}a_nb_n$! Beispiel: Die Reihe $\sum_{n\geqslant 1}\frac{(-1)^n}{\sqrt{n}}$ konvergiert nach Leibniz-Kriterium, aber

$$\sum_{n\geqslant 1}\frac{(-1)^n}{\sqrt{n}}\cdot\frac{(-1)^n}{\sqrt{n}}=\sum_{n\geqslant 1}\frac{1}{n}$$

divergiert.

3.6.1. **Satz** (Dirichlet-Kriterium). *Sei (a_n) eine monoton fallende Nullfolge, und die Folge der Partialsummen von $\sum_{n\geqslant 0}b_n$ sei beschränkt. Dann ist $\sum_{n\geqslant 0}a_nb_n$ konvergent.*

3.6.2. **Folgerung** (Abelsches Kriterium). *Sei (a_n) eine monoton fallende Nullfolge und $\sum_{n\geqslant 0}b_n$ sei konvergent. Dann ist $\sum_{n\geqslant 0}a_nb_n$ konvergent.*

Für $b_n=(-1)^n$ folgt das Leibniz-Kriterium aus dem Dirichlet-Kriterium. Der Beweis des Dirichlet-Kriteriums ist ganz analog zum Beweis des Leibniz-Kriteriums; dabei benutzt man die folgende Formel, genannt ***Abelsche partielle Summation***:

$$\sum_{k=1}^{n}a_kb_k=\sum_{k=1}^{n-1}(a_k-a_{k+1})s_k+a_ns_n\,. \tag{3.13}$$

3.6.3. ***Aufgabe***.

(i) Beweisen Sie die Abelsche partielle Summation (3.13).
(ii) Beweisen Sie das Dirichlet-Kriterium.
(iii) Beweisen Sie das Abelsche Kriterium.

3.6.4. **Lemma** (Kroneckers Lemma). *Seien (x_n) und (y_n) zwei Folgen reeller Zahlen, von denen die zweite monoton wachsend ist, nur Zahlen $y_n>0$ enthält und gegen $+\infty$ divergiert. Dann gilt:*

$$\sum_{i=1}^{\infty}\frac{x_i}{y_i}\ \textit{konvergent}\Longrightarrow \lim_{n\to\infty}\frac{1}{y_n}\sum_{i=1}^{n}x_i=0\,.$$

Notiz zur Umordnungen. Sei $\varphi:\mathbb{Z}(\geqslant q)\longrightarrow\mathbb{Z}(\geqslant q)$ bijektiv. Die Reihe $\sum_{n\geqslant q}a_{\varphi(n)}$ heißt ***Umordnung*** der Reihe $\sum_{n\geqslant q}a_n$. Nehmen wir an, dass $\sum_{n\geqslant q}a_n$ konvergiert. Fragen: Konvergiert $\sum_{n\geqslant q}a_{\varphi(n)}$? Gegen welchen Wert?

3.6.5. **Satz** (Umordnungssatz). *Ist $\sum_{n\geqslant q} a_n$ absolut konvergent, dann ist jede Umordnung $\sum_{n\geqslant q} a'_n$ absolut konvergent und $\sum_{n=q}^{\infty} a'_n = \sum_{n=q}^{\infty} a_n =: s$.*

3.6.6. **Satz** (Riemannscher Umordnungssatz). *Ist $\sum_{n\geqslant q} a_n$ konvergent, aber nicht absolut konvergent, und ist $a \in \overline{\mathbb{R}}$ vorgegeben, so gibt es eine Umordnung $\sum_{n\geqslant q} a'_n$ mit $\sum_{n=q}^{\infty} a'_n = a$.*

Beispiel: $\sum_{n\geqslant 1} \frac{(-1)^{n+1}}{n}$ ist konvergent, aber nicht absolut konvergent. Wir setzen $s = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ (eigentlich ist $s = \log 2$). Wir können so umordnen, dass die Umordnung $\sum_{n\geqslant 1} a'_n$ Summe $s/2$ hat.

Notiz zur summierbaren Familien. Sei I eine abzählbare Indexmenge und $(a_i)_{i\in I}$ eine Familie von komplexen Zahlen. Die Familie $(a_i)_{i\in I}$ heißt ***summierbar***, falls es eine bijektive Abbildung $\varphi : \mathbb{N} \longrightarrow I$ gibt, so dass $\sum_{n\geqslant 1} a_{\varphi(n)}$ absolut konvergiert. Nach 3.6.5 ist dann $\sum_{n\geqslant 1} a_{\varphi(n)}$ für jede Bijektion $\psi : \mathbb{N} \longrightarrow I$ absolut konvergent und $\sum_{n=1}^{\infty} a_{\psi(n)} = \sum_{n=1}^{\infty} a_{\varphi(n)} =: \sum_{i\in I} a_i$.

3.6.7. **Satz** (Großer Umordnungssatz). *Sei $(a_i)_I$ eine summierbare Familie und $I = \cup_{n\in\mathbb{N}} I_n$ eine Zerlegung von I. Dann sind auch $(a_i)_{i\in I_n}$ summierbar, $\sum_{n\geqslant 1}\sum_{i\in I_n} a_i$ ist absolut konvergent, und*

$$\sum_{i\in I} a_i = \sum_{n=1}^{\infty} \sum_{i\in I_n} a_i\,.$$

3.6.8. **Bemerkung.** $(a_i)_{i\in I}$ summierbar $\Leftrightarrow$ $\exists M > 0$ mit $\sum_{i\in J} |a_i| < M$ für alle endlichen Teilmengen $J \subset I$ $\Leftrightarrow$ $\exists$ Zerlegung $I = \cup_{n\geqslant 1} I_n$ mit $(a_i)_{i\in I_n}$ summierbar und $\sum_{n\geqslant 1}\sum_{i\in I_n} |a_i|$ konvergent.

3.6.9. **Satz** (Doppelreihensatz). (1) *Ist $(a_{i,j})_{(i,j)\in\mathbb{N}^2}$ summierbar, so ist für jedes $i \in \mathbb{N}$ die Reihe $\sum_{j\geqslant 1} a_{ij}$ absolut konvergent, und es gilt*

$$\sum_{(i,j)\in\mathbb{N}^2} a_{ij} = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij}\,. \tag{3.14}$$

(2) *Ist für jedes $i \in \mathbb{N}$ die Reihe $\sum_{j\geqslant 1} a_{ij}$ absolut konvergent und ist $\sum_{i\geqslant 1}\sum_{j=1}^{\infty} |a_{ij}|$ konvergent, so ist $(a_{ij})_{(i,j)\in\mathbb{N}^2}$ summierbar und es gilt* (3.14).

3.6.10. ***Aufgabe***. Betrachte die Familie $(a_{ij})_{(i,j)\in\mathbb{N}^2}$ definiert durch $a_{ij} = 1$ falls $i = j$, $a_{ij} = -1$ falls $i = j+1$ und $a_{ij} = 0$ sonst. Berechne $\sum_{i=1}^{\infty}(\sum_{j=1}^{\infty} a_{ij})$ und $\sum_{j=1}^{\infty}(\sum_{i=1}^{\infty} a_{ij})$ und zeige, dass (a_{ij}) nicht summierbar ist.

3.6.11. **Satz** (Multiplikationssatz). *Sind $\sum_{n\geqslant 1} a_n$ und $\sum_{n\geqslant 1} b_n$ absolut konvergent, so ist die Familie $(a_i b_j)_{(i,j)\in\mathbb{N}^2}$ summierbar und*

$$\sum_{(i,j)\in\mathbb{N}^2} a_i b_j = \Big(\sum_{i=1}^{\infty} a_i\Big)\Big(\sum_{j=1}^{\infty} b_j\Big).$$

Beispiel: Seien $z, w \in \mathbb{C}$, $|z| < 1$, $|w| < 1 \Rightarrow \sum_{k\geqslant 0} z^k$, $\sum_{l\geqslant 0} w^l$ absolut konvergent. Aus 3.6.11 folgt

$$\sum_{(k,l)\in\mathbb{N}_0^2} z^k w^l = \Big(\sum_{k=0}^{\infty} z^k\Big)\Big(\sum_{l=0}^{\infty} w^l\Big) = \frac{1}{(1-z)(1-w)}\,.$$

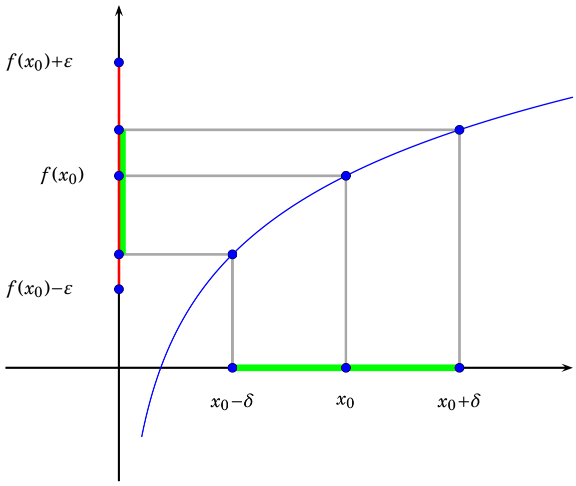
Der Multiplikationssatz gibt einen anderen Weg, der Satz über Cauchy-Produkt 3.3.6 zu beweisen.

4. Stetigkeit und Grenzwerte

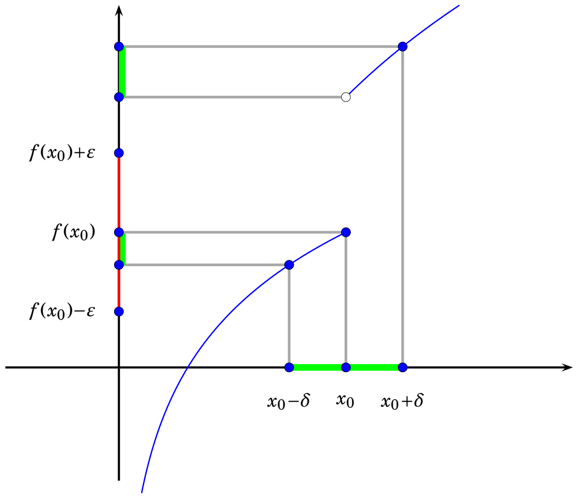
4.1. **Stetige Funktionen.** Wir betrachten Funktionen $f : D \longrightarrow \mathbb{C}$, wobei $D \subset \mathbb{C}$ beliebig ist. Wegen $\mathbb{R} \subset \mathbb{C}$ ist dabei auch $D \subset \mathbb{R}$ erlaubt. Die Funktion $f : D \longrightarrow \mathbb{C}$ heißt ***reell***, falls $f(x) \in \mathbb{R}$ für alle $x \in D$.

4.1.1. **Definition** (ε-δ-Definition der Stetigkeit)**.** Sei $f : D \longrightarrow \mathbb{C}$, $D \subset \mathbb{C}$ und $z_0 \in D$. Die Funktion f heißt ***stetig im Punkt*** z_0, falls für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon, z_0) > 0$ existiert, so dass für alle $z \in D$ mit $|z - z_0| < \delta$ gilt: $|f(z) - f(z_0)| < \varepsilon$.

Die Funktion f heißt ***stetig*** (auf D), falls f stetig in allen $z_0 \in D$ ist.



Stetig: zu jedem $\varepsilon > 0$, gibt es $\delta > 0$ mit $f\big((x_0-\delta, x_0-\delta)\big) \subset \big(f(x_0)-\varepsilon, f(x_0)-\varepsilon\big)$



Unstetig: es gibt ein $\varepsilon > 0$, so dass für alle $\delta > 0$ gilt $f\big((x_0-\delta, x_0-\delta)\big) \not\subset \big(f(x_0)-\varepsilon, f(x_0)-\varepsilon\big)$

Beispiele:

(1) Sei $f : D \longrightarrow \mathbb{C}$ eine konstante Abbildung, d.h. es gebe $c \in \mathbb{C}$ mit $f(z) = c$ für alle $z \in D$. Dann ist f stetig.

(2) $f : D \longrightarrow \mathbb{C}$, $f(z) = z$. Dann ist f stetig.

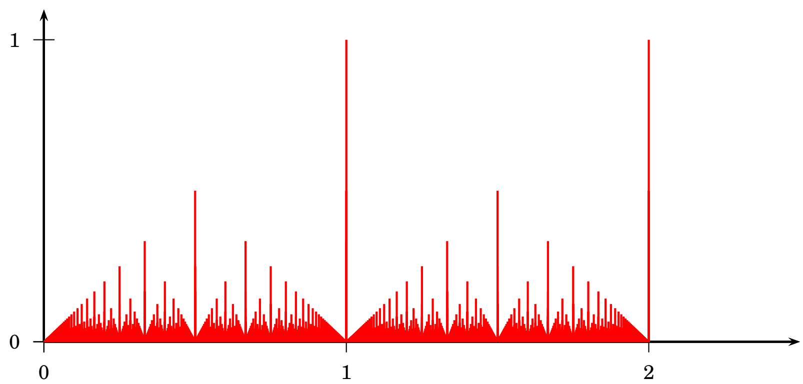
(3) $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = \begin{cases} 1, x \in \mathbb{Q}, \\ 0, x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$ f ist in keinem Punkt stetig; wir sagen: f ist unstetig.

Bemerkung: $f|_{\mathbb{Q}} : \mathbb{Q} \longrightarrow \mathbb{R}$, $x \mapsto 1$ (konstant); also ist $f|_{\mathbb{Q}}$ stetig!

(4) Die Funktion

$$f : \mathbb{R}_+ \longrightarrow \mathbb{R}, \quad f(x) = \begin{cases} \dfrac{1}{q}, x = \dfrac{p}{q}, (p,q) = 1, \\ 0, x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases} \tag{4.1}$$

heißt ***Dirichlet-Funktion***; sie ist in allen $x \in \mathbb{R} \setminus \mathbb{Q}$ stetig und unstetig in allen $x \in \mathbb{Q}$. Der Graph von f sieht wie ein Popcorn aus.



Weitere Beispiele liefern die Lipschitz-stetige Funktionen. Eine Funktion $f: D \longrightarrow \mathbb{C}$ heißt ***Lipschitz-stetig***, wenn es $L \in \mathbb{R}$ gibt, so dass für alle $z, w \in D$ gilt: $|f(z) - f(w)| \leqslant L|z-w|$. In diesem Fall heißt L eine ***Lipschitz-Konstante*** zu f.

4.1.2. **Satz.** *Sei $f: D \longrightarrow \mathbb{C}$ Lipschitz-stetig. Dann ist f stetig.*

Beispiele: (i) Eine lineare Funktion $f: \mathbb{C} \longrightarrow \mathbb{C}$ $f(z) = az+b$ ist Lipschitz-stetig mit Lipschitz-Konstante $|a|$.

(ii) $z \mapsto |z|$, $z \mapsto \overline{z}$, $z \mapsto \operatorname{Re} z$, $z \mapsto \operatorname{Im} z$ sind Lipschitz-stetig mit Lipschitz-Konstante 1 (z.B. $||z|-|w|| \leqslant |z-w|$).

4.1.3. **Satz.** *Sei $f: D \longrightarrow \mathbb{C}$ stetig in z_0. Sei $E \subset D$ mit $z_0 \in E$. Dann ist auch $f|_E : E \longrightarrow \mathbb{C}$ stetig in z_0.*

4.1.4. **Definition.** Sei $z_0 \in D \subset \mathbb{C}$. Eine Teilmenge $U \subset D$ heißt ***Umgebung von z_0 in D***, wenn es ein $\delta > 0$ mit $B_\delta(z_0) \cap D \subset U$ gibt.

Beispiele:

(i) $D = \mathbb{R}$, $B_\delta(z_0) \cap D = (z_0 - \delta, z_0 + \delta)$.

(ii) $D = [a,b] \subset \mathbb{R}$, $B_\delta(a) \cap D = [a, a+\delta) \cap [a,b]$.

(iii) Ist $U \subset D$ eine Umgebung von z_0 und $U \subset V \subset D$, so ist V eine Umgebung von z_0 in D. Sind U, V Umgebungen von z_0 in D, so ist $U \cap V$ eine Umgebung von z_0 in D.

4.1.5. **Satz.** *Sei $f: D \longrightarrow \mathbb{C}$ und $z_0 \in D$. Die folgenden Aussagen sind äquivalent:*

(i) *f ist stetig in z_0* (ε-δ-Definition der Stetigkeit).

(ii) *Zu jeder Umgebung V von $f(z_0)$ in $\mathbb{C}$ gibt es eine Umgebung U von z_0 in D mit $f(U) \subset V$* (Umgebungs-Definition der Stetigkeit).

(iii) *Für jede Folge (z_n) in D mit $z_n \to z_0$ gilt $f(z_n) \to f(z_0)$* (Folgenkriterium für Stetigkeit).

4.1.6. **Satz** (Rechenregeln). *Seien $f, g: D \longrightarrow \mathbb{C}$ stetig in $z_0 \in D$. Dann gilt:*

(i) *$f+g: D \longrightarrow \mathbb{C}$ und $fg: D \longrightarrow \mathbb{C}$ sind stetig in z_0.*

(ii) *Falls $g(z_0) \neq 0$, so ist $E = \{z \in D : g(z) \neq 0\}$ eine Umgebung von z_0 in D, und $f/g: E \longrightarrow \mathbb{C}$ ist stetig in z_0.*

4.1.7. **Folgerung.**

(i) **Polynome** $\mathbb{C} \ni z \mapsto a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 \in \mathbb{C}$ *sind stetig.*

(ii) **Rationale Funktionen** $D \ni z \mapsto \frac{P(z)}{Q(z)} \in \mathbb{C}$ *mit Polynomen P und Q sind stetig auf* $D = \{z \in \mathbb{C} : Q(z) \neq 0\}$. 5.12.2011

4.1.8. **Satz** (Kompositionsregel). *Seien $f: D \longrightarrow \mathbb{C}$, $g: E \longrightarrow \mathbb{C}$ mit $f(D) \subset E$. Ist f in $z_0 \in D$ stetig und g in $f(z_0)$ stetig, so ist $g \circ f: D \longrightarrow \mathbb{C}$ in z_0 stetig.*

4.1.9. **Folgerung.** *Ist $f: D \longrightarrow \mathbb{C}$ stetig, so sind $|f|$, $\overline{f}$, $\operatorname{Re} f$, $\operatorname{Im} f : D \longrightarrow \mathbb{C}$ stetig.*

4.1.10. **Satz** (Stetigkeit der Umkehrfunktion). *Sei $f: [a,b] \longrightarrow \mathbb{C}$ stetig und injektiv. Sei $E = \{f(x) : x \in [a,b]\}$ und sei $g: E \longrightarrow [a,b]$, $g := f^{-1}$. Dann ist g stetig.*

Beweis: Es gilt nach Definition der Umkehrfunktion $y = f(x) \iff x = g(y)$ für $x \in [a,b]$, $y \in E$. Sei $y_0 \in E$ und $y_n \to y_0$, $n \to \infty$, $y_n \in E$. Sei $x_n = g(y_n) \in [a,b]$, also $f(x_n) = y_n$. Nach dem Satz von Bolzano-Weierstass 2.3.4 existiert $L = \limsup x_n \in [a,b]$ und eine Teilfolge $x_{n_k} \to L$, $k \to \infty$. Da f stetig in L ist, gilt $y_{n_k} = f(x_{n_k}) \to f(L)$, $k \to \infty$. Wegen $y_{n_k} \to y_0$, $k \to \infty$ folgt, dass $y_0 = f(L)$ und deshalb $L = g(y_0)$. Analog zeigt man, dass für $\ell = \liminf x_n \in [a,b]$ gilt $\ell = g(y_0)$. Insgesamt $\liminf x_n = \limsup x_n = g(y_0)$ und somit $g(y_n) \to g(y_0)$, $n \to \infty$. □

Eine Version des Satzes 4.1.10 nur für die Stetigkeit in einem Punkt findet man in Übung 4.6.6.

4.1.11. **Definition.** Sei $I \subset \mathbb{R}$ ein Intervall. Eine Funktion $f: I \longrightarrow \mathbb{R}$ heißt ***monoton steigend*** (bzw. ***fallend***) $:\iff$ Für alle $x, y \in I$ mit $x \leqslant y$ gilt $f(x) \leqslant f(y)$ (bzw. $f(x) \geqslant f(y)$).

Eine Funktion $f: I \longrightarrow \mathbb{R}$ heißt ***streng monoton steigend*** (bzw. ***fallend***) $:\iff$ Für alle $x, y \in I$ mit $x < y$ gilt $f(x) < f(y)$ (bzw. $f(x) > f(y)$).

Bemerkung: Wenn f streng monoton ist, dann ist f injektiv.

Beispiel: Stetigkeit der Wurzelfunktion. Sei $k \in \mathbb{N}$. Seien $0 \leq a < b$. Dann ist $f: [\sqrt[k]{a}, \sqrt[k]{b}] \longrightarrow [a,b]$ mit $f(x) := x^k$ stetig monoton steigend. Wegen Satz 4.1.10 ist $g = f^{-1} = \sqrt[k]{\cdot} : [a,b] \longrightarrow [\sqrt[k]{a}, \sqrt[k]{b}]$ stetig; es folgt, dass die durch $h(x) := \sqrt[k]{x}$ definierte Funktion $h: [0,\infty) \longrightarrow \mathbb{R}$ stetig ist.

4.2. **Potenzreihen und Stetigkeit.**

4.2.1. **Definition.** Eine Folge (f_n) von Funktionen $f_n : D \longrightarrow \mathbb{C}$ ***konvergiert punktweise***, wenn für jedes $z \in D$ die Folge $(f_n(z))$ konvergiert. Ist dies der Fall, so heißt die durch $f(z) = \lim_{n\to\infty} f_n(z)$ definierte Funktion $f : D \longrightarrow \mathbb{C}$ die ***Grenzfunktion*** zu (f_n). Eine Reihe $\sum_{n\geqslant 0} f_n$ von Funktionen $f_n : D \longrightarrow \mathbb{C}$ heißt ***punktweise konvergent***, falls die Folge der Partialsummen (s_n), wobei $s_n = f_0 + f_1 + \ldots + f_n$, punktweise konvergiert. Für $D \ni z \mapsto \sum_{n=0}^{\infty} f_n(z) \in \mathbb{C}$ schreiben wir $\sum_{n=0}^{\infty} f_n$.

Beispiel: Seien $f_n : [0,1] \longrightarrow \mathbb{R}$, $f(x) = x^n$. Es gilt $f_n(1) = 1$ für alle n, also $\lim_{n\to\infty} f_n(1) = 1$. Für $x \in [0,1)$ gilt $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = 0$. Also konvergiert $(f_n)_{n\in\mathbb{N}}$ punktweise gegen

$$f : [0,1] \longrightarrow \mathbb{R}, \quad f(x) = \begin{cases} 0, x \in [0,1), \\ 1, x = 1. \end{cases}$$

Die Funktion f ist unstetig! Wir benötigen einen Konvergenzbegriff, der solche Phänomene ausschließt und garantiert, dass die Grenzfunktion stetig ist, wenn die Folgenglieder stetig sind. Der richtige Begriff für Funktionenfolgen ist die *gleichmäßige Konvergenz*. Wir sind in diesem Paragraph hauptsächlich mit Reihen beschäftigt, wir werden also einen für Reihen bequemeren Begriff, die *normale Konvergenz* einführen. Die gleichmäßige Konvergenz wird später diskutiert.

4.2.2. **Definition.** Für eine Funktion $f : D \longrightarrow \mathbb{C}$ wird die ***Supremumsnorm*** durch

$$\|f\|_D := \sup\big\{|f(z)| : z \in D\big\} \in [0,\infty].$$

eingeführt. Eine Funktion $f : D \longrightarrow \mathbb{C}$ heißt ***beschränkt***, falls $\|f\|_D < \infty$. Aquivalent dazu: das Bildmenge $f(D) \subset \mathbb{C}$ ist beschränkt (siehe Definition 2.4.4) oder $\{|f(z)| : z \in D\} \subset [0,\infty)$ ist beschränkt. Andernfalls heißt f unbeschränkt.

Die Menge $\mathscr{B}(D) = \{f : D \to \mathbb{C} : \|f\|_D < \infty\}$ aller beschränkten Funktionen ist ein $\mathbb{C}$-Vektrorraum; die Abbildung $f \mapsto \|f\|_D$ ist eine *Norm* auf $\mathscr{B}(D)$, genauer, sie erfüllt folgenden Eigenschaften:
1) $\|f\|_D \geqslant 0$; $\|f\|_D = 0 \Leftrightarrow f = 0$,
2) $\|cf\|_D = |c|\|f\|_D$ für alle $c \in \mathbb{C}$,
3) $\|f+g\|_D \leqslant \|f\|_D + \|g\|_D$.

4.2.3. **Definition.** Eine Reihe $\sum_{n\geqslant 1} f_n$ von Funktionen $f_n : D \longrightarrow \mathbb{C}$ heißt ***normal***[3] ***konvergent***, falls $\sum_{n\geqslant 1} \|f_n\|_D$ konvergiert.

4.2.4. **Satz.** *Ist $\sum_{n\geqslant 1} f_n$ normal konvergent, so ist $\sum_{n\geqslant 1} f_n$ punktweise konvergent,*

In der Tat, für jedes $z \in D$ ist die Reihe $\sum_{n\geqslant 1} \|f_n\|_D$ Majorante für $\sum_{n\geqslant 1} f_n(z)$.

4.2.5. **Satz.** *Sei $P(z) = \sum_{n\geqslant 0} a_n z^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Sei $0 < r < R$. Dann konvergiert $\sum_{n\geqslant 0} a_n z^n$ normal auf $B_r(0)$.*

Bemerkung: Die Bedingung $r < R$ ist wichtig!

8.12.2011

4.2.6. **Satz.** *Konvergiert $\sum_{n\geqslant 1} f_n$ normal gegen $f : D \longrightarrow \mathbb{C}$, und sind alle $f_n : D \longrightarrow \mathbb{C}$ stetig, so ist auch f stetig.*

4.2.7. **Satz.** *Sei $P(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann ist $P : B_R(0) \longrightarrow \mathbb{C}$ stetig. (Zur Erinnerung: Nach Vereinbarung ist $B_R(0) := \mathbb{C}$, falls $R = \infty$).*

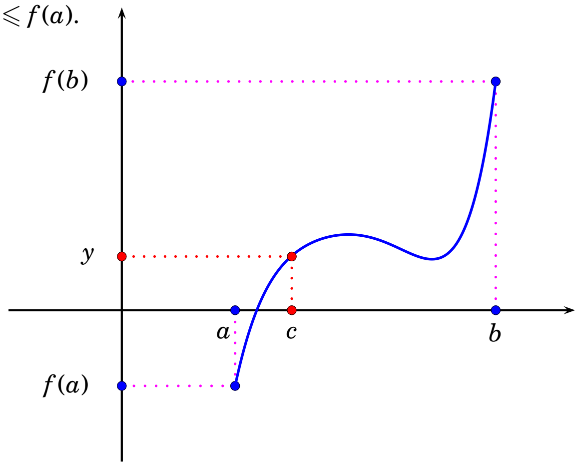
4.2.8. **Folgerung.** *Die Funktionen* exp, cos, sin : $\mathbb{C} \to \mathbb{C}$ *(definiert in* (3.4)*,* (3.11)*) und deren Einschränkungen* exp, cos, sin : $\mathbb{R} \to \mathbb{R}$ *sind stetig.*

[3]Das Wort „normal“ betont die Präsenz der Norm $\|\cdot\|_D$ in der Definition. Der Begriff der normalen Konvergenz wurde 1908 von René Baire eingeführt. Er notiert: „Un grand nombre des démonstrations soit dans la théorie des séries, soit plus loin dans la théorie des produits infinis, sont considérablement simplifiées quant on met en avance cette notion, beaucoup plus maniable que la propriété de convergence uniforme.“

4.3. **Der Zwischenwertsatz.** Der Zwischenwertsatz bildet die Grundlage vieler Existenzaussagen der Analysis.

4.3.1. **Satz** (Zwischenwertsatz). *Sei $f:[a,b]\longrightarrow\mathbb{R}$ stetig. Dann gibt es zu jedem y zwischen $f(a)$ und $f(b)$ ein $c\in[a,b]$ mit $f(c)=y$.*

Dabei bedeutet „y zwischen $f(a)$ und $f(b)$“: Falls $f(a)\leqslant f(b)$, so ist $f(a)\leqslant y\leqslant f(b)$, und falls $f(b)\leqslant f(a)$, so ist $f(b)\leqslant y\leqslant f(a)$.



Die Aussage des Zwischenwertsatzes ist falsch, wenn f nicht stetig oder nicht auf einem Intervall definiert ist.

Als erste Anwendung, zeigen wir noch einmal die Existenz der k-ten Wurzel (dies wurde bereits in Satz 1.5.8 direkt aus der Vollständigkeit von $\mathbb{R}$ hergeleitet). Sei $a\in\mathbb{R}$, $a>0$, und sei $k\in\mathbb{N}$. Betrachte die stetige Funktion $f:[0,\infty)\to\mathbb{R}$, $f(x)=x^k-a$. Es gilt $f(0)=-a<0$ und $f(1+a)=(1+a)^k>1+ka>0$ (Bernoullische Ungleichung). Der Zwischenwertsatz besagt, dass es ein $x\in[0,1+a]$ gibt, mit $f(x)=0$, d. h., $x^k=a$. Diese Lösung ist eindeutig, da f streng monoton wachsend ist.

4.3.2. **Folgerung.** *Sei $I\subset\mathbb{R}$ ein Intervall und sei $f:I\to\mathbb{R}$ eine stetige Funktion. Dann ist das Bild $f(I)$ auch ein Intervall.*

Eine weitere Anwendung des Zwischenwertsatzes ist die folgende Äquivalenz der Injektivität und Monotonie für stetige Funktionen.

4.3.3. **Satz.** *Sei $I\subset\mathbb{R}$ ein Intervall und sei $f:I\longrightarrow\mathbb{R}$ stetig. Dann sind äquivalent:*

(i) f ist injektiv,

(ii) f ist streng monoton.

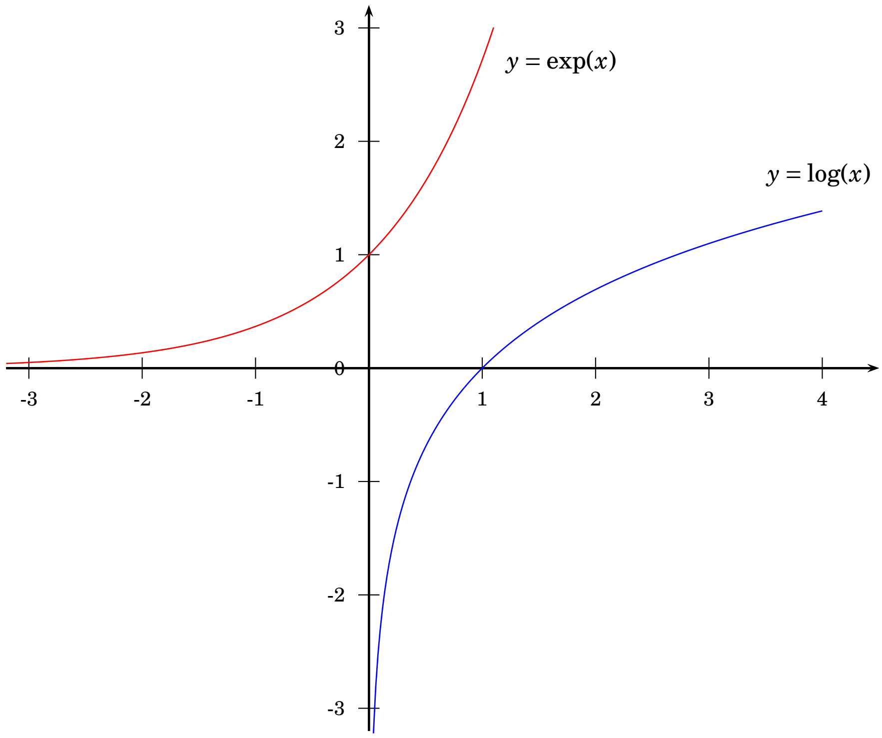
In diesem Fall ist $f^{-1}:f(I)\to I$ stetig und streng monoton.

Beweis: Wir nehmen an, dass f injektiv ist, aber weder streng monoton wachsend noch streng monoton fallend. Wenn f nicht streng monoton fallend ist, gibt es $x_1,x_2\in I$ mit $x_1<x_2$ und $f(x_1)\le f(x_2)$. Weil f injektiv ist, gilt dann auch rechts eine echte Ungleichung, also $x_1<x_2$ und $f(x_1)-f(x_2)<0$. Wenn f nicht streng monoton fallend ist, gibt es ebenso $y_1,y_2\in I$ mit $y_1<y_2$ und $f(y_1)?f(y_2)>0$. Betrachte die Funktion $g:[0,1]\to\mathbb{R}$ mit $g(t):=f((1-t)x_1+ty_1)-f((1-t)x_2+ty_2)$. Diese Funktion ist als Komposition stetiger Funktionen stetig auf dem Intervall $[0,1]$. Weiter ist $g(0)=f(x_1)-f(x_2)<0$, $g(1)=f(y_1)-f(y_1)>0$. Also gibt es nach dem Zwischenwertsatz ein $t\in(0,1)$ mit $g(t)=0$, also mit $f((1-t)x_1+ty_1)=f((1-t)y_2+ty_2)$. Aber nach den Rechenregeln für Ungleichungen gilt $(1-t)x_2+ty_1<(1-t)x_2+ty_2$. Das ist ein Widerspruch zur Injektivität von f.

Zur Stetigkeit von f^{-1}: Sei $y_0\in f(I)=J$. Nehmen wir an, dass y_0 im Innneren des Intervalls J liegt, d. h. es gibt $c,d\in J$ mit $y_0\in(c,d)\subset J$. Seien $a,b\in I$ mit $f(a)=c$, $f(b)=d$. Sei I_0 das kompakte Intervall mit Endpunkten a,b. Dann gilt $f(I_0)=[c,d]$ und $f|_{I_0}:I_0\to[c,d]$ is bijektiv und stetig. Der Satz 4.1.10 besagt, dass die Umkerfunktion $\left(f|_{I_0}\right)^{-1}=f^{-1}|_{[c,d]}$ ist stetig. Daraus folgt, dass f^{-1} stetig in y_0 ist, da $y_0\in(c,d)$. Falls y_0 ein Endpunkt von J ist, argumentiert man analog. □

Der Zwichenwersatz erlaubt eine vollständige Beschreibung der reellen Exponentialfunktion $\exp:\mathbb{R}\longrightarrow\mathbb{R}$ (liefert nämlich die Bildmenge) und dies führt wiederum zur Definition des Logarithmus. Diese Definition des Logarithmus als Umkehrung des Exponentialfunktion findet sich erstmals in Eulers Buch [8].

4.3.4. **Definition und Satz.** *(i)* $\exp : \mathbb{R} \longrightarrow \mathbb{R}_+$ *ist streng monoton wachsend, stetig und bijektiv.*
(ii) Sei $\log : \mathbb{R}_+ \longrightarrow \mathbb{R}$ *die Umkehrfunktion dazu, gennant natürlicher* ***Logarithmus****. Dann ist* $\log$ *streng monoton wachsend, stetig und bijektiv.*
(iii) $\log 1 = 0$, $\log(xy) = \log(x) + \log(y)$ *für alle* $x, y \in \mathbb{R}_+$.



4.3.5. **Definition.** Für $x > 0$ und $z \in \mathbb{C}$ definieren wir

$$x^z := \exp(z \cdot \log x) = e^{z \cdot \log x}. \tag{4.2}$$

Die Funktion $\mathbb{R}_+ \ni x \mapsto x^z \in \mathbb{C}$ (für festes $z \in \mathbb{C}$) heißt ***Potenzfunktion zum Exponenten*** z und die Funktion $\mathbb{C} \ni z \mapsto x^z \in \mathbb{C}$ (für festes $x > 0$) heißt ***Exponentialfunktion zur Basis*** x.

Für alle $x > 0$, $q \in \mathbb{Q}$ stimmt die obige Definition von x^q mit unserer früheren Definition 1.5.9 überein (siehe Übung 4.6.4). Als Komposition stetiger Funktionen sind sowohl die Potenzfunktionene als auch die Exponentialfunktionen stetig.

Es gelten die ***Potenzgesetze*** (vgl. (1.7))

$$(xy)^z = x^z y^z\,,\ x^{z+w} = x^z x^w\,,\ (x^a)^z = x^{az}\,,\ \text{für alle } x, y > 0,\ z, w \in \mathbb{C} \text{ und } a \in \mathbb{R} \tag{4.3}$$

12.12.2011

Außerdem gilt für $x > 1$ und $a < b$ gilt $x^a < x^b$; für $a > 0$ und $x < y$ gilt $x^a < y^a$.

4.4. Satz vom Maximum und Minimum.

4.4.1. **Definition.** Sei $f : D \to \mathbb{R}$. Die Funktion f heißt ***nach oben beschränkt*** (bzw. ***nach unten beschränkt***), wenn $f(D)$ nach oben beschränkt (bzw. nach unten beschränkt) ist.

Für eine beliebige Funktion $f : D \to \mathbb{R}$ setzen wir

$$\sup_D f := \sup f(D) = \sup\{f(z) : z \in D\} \in \overline{\mathbb{R}}, \quad \inf_D f := \sup f(D) = \inf\{f(z) : z \in D\} \in \overline{\mathbb{R}}$$

Dann gilt:

$$f \text{ ist nach oben beschränkt} \Leftrightarrow \sup_D f < \infty$$

$$f \text{ ist nach unten beschränkt} \Leftrightarrow \inf_D f > -\infty$$

4.4.2. **Definition.** Sei $f : D \to \mathbb{R}$. Wir sagen, dass f ihr ***Maximum*** (bzw. ***Minimum***) ***annimmt***, wenn $f(D)$ ein Maximum (bzw. Minimum) hat. Dann heißen

$$\max_D f := \max f(D) = \max\{f(z) : z \in D\}, \quad \min_D f := \min f(D) = \min\{f(z) : z \in D\}$$

Maximum (bzw. ***Minimum***) von f. Es gilt

$$f \text{ nimmt ihr Maximum an } \Leftrightarrow \exists x_1 \in D\ \forall x \in D\colon f(x) \leqslant f(x_1) \Leftrightarrow \exists x_1 \in D\colon f(x_1) = \sup_D f$$

$$f \text{ nimmt ihr Minimum an } \Leftrightarrow \exists x_2 \in D\ \forall x \in D\colon f(x) \geqslant f(x_2) \Leftrightarrow \exists x_2 \in D\colon f(x_2) = \inf_D f$$

Ein Punkt $x_1 \in D$ (bzw. $x_2 \in D$) wie oben heißt ***Maximalstelle*** (bzw. ***Minimalstelle***) von f.

4.4.3. **Satz** (Satz vom Maximum und Minimum (Weierstraß))**.** *Eine stetige Funktion $f : [a,b] \longrightarrow \mathbb{R}$ auf einem nichtleeren kompakten Intervall ist beschränkt und nimmt ihr Maximum und ihr Minimum an.*

4.4.4. **Bemerkung.** Die Kompaktheit des Intervalls ist wesentlich! Zum Beispiel ist die Funktion $f : (0,1] \longrightarrow \mathbb{R}$, $f(x) = \frac{1}{x}$ stetig, aber unbeschränkt. Die Funktion $g : (0,1) \longrightarrow \mathbb{R}$, $g(x) = x$ ist beschränkt, nimmt aber ihr Maximum und Minimum nicht an. Es gilt nämlich $\sup g(0,1) = 1$, $\inf g(0,1) = 0$, aber $0 < g(x) < 1$ für alle $x \in (0,1)$. Daher besitzt g kein Maximum und kein Minimum.

4.4.5. **Folgerung.** *Sei $f : [a,b] \longrightarrow \mathbb{R}$ eine stetige Funktion. Dann ist $f([a,b]) = [m,M]$, wobei $m = \min_{[a,b]} f$, $M = \max_{[a,b]} f$.*

4.5. **Fortsetzung stetiger Funktionen und Grenzwerte.**

4.5.1. **Definition.** Ein Punkt $z_0 \in \overline{\mathbb{C}}$ heißt ***Häufungspunkt*** einer Menge $D \subset \mathbb{C}$, wenn jede Umgebung von z_0 (in $\overline{\mathbb{C}}$) unendlich viele Punkte von D enthält. Analog heißt $x_0 \in \overline{\mathbb{R}}$ Häufungspunkt einer Menge $D \subset \mathbb{R}$, wenn jede Umgebung von x_0 (in $\overline{\mathbb{R}}$) unendlich viele Punkte von D enthält. Der Punkt $z_0 = \infty \in \overline{\mathbb{C}}$ (bzw. $x_0 = \pm\infty \in \overline{\mathbb{R}}$) heißt ***uneigentlicher Häufungspunkt*** von D.

4.5.2. **Bemerkung.** (i) Der Punkt z_0 kann Element von D sein oder auch nicht. Z.B. ist 0 ein HP von $[0,1]$ und von $(0,1]$.

(ii) Ein HW einer Folge $(x_n)_{n\in\mathbb{N}}$ ist nicht dasselbe wie ein HP der Menge $\{x_n : n \in \mathbb{N}\}$. Sei z.B. $z_n = 1$ für alle $n \in \mathbb{N}$. Dann ist $z_0 = 1$ ein HW von $(z_n)_{n\in\mathbb{N}}$, aber kein HP von $\{z_n : n \in \mathbb{N}\} = \{1\}$.

(iii) ∞ (bzw. $-\infty$) ist HP von $D \subset \mathbb{R}$ genau dann, wenn D nach oben (bzw. unten) unbeschränkt ist.

4.5.3. **Lemma.** *Die folgenden Aussagen sind äquivalent:*

(i) *z_0 ist HP von D.*
(ii) *Jede Umgebung von z_0 in $\overline{\mathbb{C}}$ enthält einen von z_0 verschiedenen Punkt aus D.*
(iii) *Es gibt eine Folge $(z_n)_{n\in\mathbb{N}}$ in D mit $z_n \to z_0$ für $n \to \infty$ und $z_n \neq z_0$ für alle $n \in \mathbb{N}$.*

Beweis: (i)⇒(ii) klar.

(ii)⇒(iii) Wähle $z_n \in B_{1/n}(z_0) \cap (D \setminus \{z_0\})$. Dann gilt $z_n \to z_0$ für $n \to \infty$ und $z_n \neq z_0$ für alle $n \in \mathbb{N}$.

(iii)⇒(i) Sei $(z_n)_{n\geqslant 1}$ eine Folge in $D \setminus \{z_0\}$ mit $z_n \to z_0$. Sei $\varepsilon > 0$ beliebig. Wähle $n_1 \in \mathbb{N}$ mit $z_{n_1} \in B_\varepsilon(z_0) \setminus \{z_0\}$. Sei $\varepsilon_1 := |z_{n_1} - z_0| > 0$. Wähle $n_2 \in \mathbb{N}$, $n_2 > n_1$ mit $z_{n_2} \in B_{\varepsilon_1}(z_0) \setminus \{z_0\}$. Dann ist $z_{n_1} \neq z_{n_2}$. Sei $\varepsilon_2 := |z_{n_2} - z_0| > 0$. Wähle $z_{n_3} \in B_{\varepsilon_2}(z_0) \setminus \{z_0\}$. Dann gilt $z_{n_1}, z_{n_2} \notin B_{\varepsilon_2}(z_0)$, also $z_{n_1} \neq z_{n_3}$, $z_{n_2} \neq z_{n_3}$. Durch Induktion konstruieren wir also eine Folge $(z_{n_k})_{k\geqslant 1}$ von paarweise verschiedenen Punkten in $B_\varepsilon(z_0) \cap D$.

4.5.4. **Definition.** Ein Punkt $z_0 \in D$ heißt ***isolierter Punkt*** von D, wenn es $\delta > 0$ mit $B_\delta(z_0) \cap D = \{z_0\}$ gibt.

4.5.5. **Bemerkung.** (i) $z_0 \in D$ ist entweder isoliert oder HP von D.

(ii) Ist $z_0 \in D$ isoliert, so ist jede Funktion $f : D \longrightarrow \mathbb{C}$ stetig in z_0.

4.5.6. **Definition.** (i) Sei $D \subset \mathbb{C}$, $z_0 \in \overline{\mathbb{C}}$ ein HP von D, $a \in \overline{\mathbb{C}}$. Wir sagen, dass $f : D \longrightarrow \mathbb{C}$ in z_0 den ***Grenzwert*** a hat, wenn es zu jeder Umgebung U von a eine Umgebung V von z_0 mit $f(V \cap (D \setminus \{z_0\})) \subset U$ gibt. Schreibweise: $\lim_{z\to z_0} f(z) = a$.

(ii) Sei $D \subset \mathbb{R}$, $x_0 \in \overline{\mathbb{R}}$ ein HP von D, $a \in \overline{\mathbb{R}}$. Wir sagen, dass $f : D \longrightarrow \mathbb{R}$ in x_0 den ***Grenzwert*** a hat, wenn es zu jeder Umgebung U von a eine Umgebung V von x_0 mit $f(V \cap (D \setminus \{x_0\})) \subset U$ gibt. Schreibweise: $\lim_{x\to x_0} f(x) = a$.

$\boxed{15.12.2011}$

4.5.7. **Bemerkung.** (i) Die Existenz und die Größe des Grenzwertes einer Funktion f in einem Punkt z_0 hat nichts damit zu tun, ob oder wie f in z_0 definiert ist. Wieso werden in der Definition nur Punkte $z \in V$, $z \neq z_0$ betrachtet? Es gibt zwei mögliche Gründe: Zunächst könnte $z_0 \notin D$ sein (z. B. $z_0 = \infty$); zweitens kann es für $z_0 \in D$ passieren, dass die Funktion „besser" in $V \setminus \{z_0\}$ ist als in V. Zum Beispiel ist die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x\sin(1/x)$ für $x \neq 0$ und $f(0) = 1$ nicht stetig in $x_0 = 0$, aber $\lim_{x\to 0} f(x) = 0$. Es wäre schade, diese schöne Eigenschaft von f nicht zu erkennen...

Wir können sagen, dass wir in der Definition der Stetigkeit das Verhalten einer Funktion in einer Umgebung eines Punktes z_0 im Definitionsbereich mit dem Wert der Funktion $f(z_0)$ vergleichen, während wir uns in der Definition des Grenzwertes nur für das Verhalten einer Funktion in Punkten $z \neq z_0$ interessieren.

Wenn wir akzeptieren, dass nur Punkte $z \in V$, $z \neq z_0$ betrachtet werden, dann ist es klar, dass die Definition nur für Häufungspunkte Sinn macht. Wäre z_0 in Definition 4.5.6 ein isolierter Punkt, so könnte man für jedes a und jedes U eine Umgebung V mit $V \cap (D \setminus \{z_0\}) = \emptyset$ wählen. Dann würde $f(V \cap (D \setminus \{z_0\})) = \emptyset \subset U$ und folglich $\lim_{z\to z_0} f(z) = a$ gelten. Weil aber a beliebig war, wäre $\lim_{z\to z_0} f(z)$ nicht wohldefiniert.

(ii) $a = \infty \in \overline{\mathbb{C}}$ und $a = \pm\infty \in \overline{\mathbb{R}}$ heißen ***uneigentliche Grenzwerte***.

Die Definition 4.5.6 ist einheitlich für eigentliche und uneigentliche Grenzwerte, man muss nur die jeweiligen Definitionen der Umgebungen beachten:

- Umgebung von $a \in \mathbb{R}$ in $\overline{\mathbb{R}}$: Obermenge eines Intervalles $(a-\delta, a+\delta)$
- Umgebung von $a = \infty$ (bzw. $-\infty$) in $\overline{\mathbb{R}}$: Obermenge eines Intervalles $(M, +\infty)$ (bzw. $(-\infty, M)$)
- Umgebung von $a \in \mathbb{C}$ in $\overline{\mathbb{C}}$: Obermenge einer Scheibe $B_\delta(a) \subset \mathbb{C}$
- Umgebung von $a = \infty$ in $\overline{\mathbb{C}}$: Obermenge des Äußeren $\{z \in \mathbb{C} : |z| > M\}$ einer Scheibe in $\mathbb{C}$

(iii) Ist $z_0 \in D$ ein HP von D, so ist $f : D \longrightarrow \mathbb{C}$ stetig in z_0 genau dann, wenn $\lim_{z\to z_0} f(z) = f(z_0)$.

(iv) Sei $z_0 \in \mathbb{C} \setminus D$ ein HP von D, und sei $a \in \mathbb{C}$. Sei $F : D \cup \{z_0\} \longrightarrow \mathbb{C}$ definiert durch

$$F(z) = \begin{cases} f(z), & z \in D, \\ a, & z = z_0. \end{cases}$$

Dann gilt:

$$F \text{ ist stetig in } z_0 \Leftrightarrow \lim_{z\to z_0} F(z) = F(z_0) \Leftrightarrow \lim_{z\to z_0} f(z) = a.$$

In diesem Fall heißt F die ***stetige Fortsetzung*** von f in z_0.

4.5.8. **Beispiele.** **(1)** Seien $D := \mathbb{C} \setminus \{0\}$, $f : D \longrightarrow \mathbb{C}$, $f(z) = \dfrac{\exp(z) - 1}{z}$; $z_0 = 0$ ist HP von D. Für $z \in D$ gilt

$$f(z) = \frac{1}{z}\Big(\sum_{k=0}^{\infty} \frac{z^k}{k!} - 1\Big) = \frac{1}{z}\sum_{k=1}^{\infty}\frac{z^k}{k!} = 1 + \frac{z}{2!} + \frac{z^2}{3!} + \ldots =: P(z).$$

$P : \mathbb{C} \longrightarrow \mathbb{C}$ ist definiert auch für $z = 0$ und ist die Summe einer Potenzreihe mit Konvergenzradius ∞, also

$$P(z) = \begin{cases} \dfrac{\exp(z)-1}{z}, & z \neq 0, \\ 1, & z = 0. \end{cases}$$

Aus Satz 4.2.7 folgt, dass P in 0 stetig ist und $\lim_{z\to 0} P(z) = P(0) = 1$ gilt. Also ist P die stetige Fortsetzung von f in 0. Wir erhalten

$$\lim_{z\to 0} \frac{\exp(z)-1}{z} = 1 \tag{4.4}$$

Analog ergibt sich

$$\lim_{z\to 0} \frac{\sin z}{z} = 1, \quad \lim_{z\to 0} \frac{1-\cos z}{z^2} = \frac{1}{2} \tag{4.5}$$

(2) Seien $D := \mathbb{R} \setminus \{0\}$, $z_0 = 0$, $f : D \longrightarrow \mathbb{R}$, $f(x) = x \sin\frac{1}{x}$. Dann gilt $\lim_{x\to 0} x \sin\frac{1}{x} = 0$. Also ist $F : \mathbb{R} \longrightarrow \mathbb{R}$ mit

$$F(x) = \begin{cases} x\sin\frac{1}{x}, & x \neq 0, \\ 0, & x = 0 \end{cases}$$

die stetige Fortsetzung von f in 0.

Analog gilt $\lim_{x\to 0} x^n \sin\frac{1}{x} = 0$ für $n \in \mathbb{N}$, aber $\lim_{x\to 0} \sin\frac{1}{x}$ existiert nicht.

(3) Sei $n \in \mathbb{N}$. Dann gilt:

$$\lim_{z\to\infty} z^n = \infty, \quad \lim_{x\to\infty} x^n = \infty, \quad \lim_{x\to-\infty} x^n = \begin{cases} -\infty, & n \quad \text{ungerade} \\ \infty, & n \quad \text{gerade} \end{cases}.$$

Sei nämlich $c \in \mathbb{R}$. Wähle $d = \max\{c, 1\}$. Für $|z| > d$ gilt dann $|z^n| \geqslant |z| > d \geqslant c$. Also $\lim_{z\to\infty} z^n = \infty$. Analog beweist man die anderen Aussagen.

(4) Für $n \in \mathbb{N}_0$ gilt

$$\lim_{x\to\infty} \frac{e^x}{x^n} = \infty \tag{4.6}$$

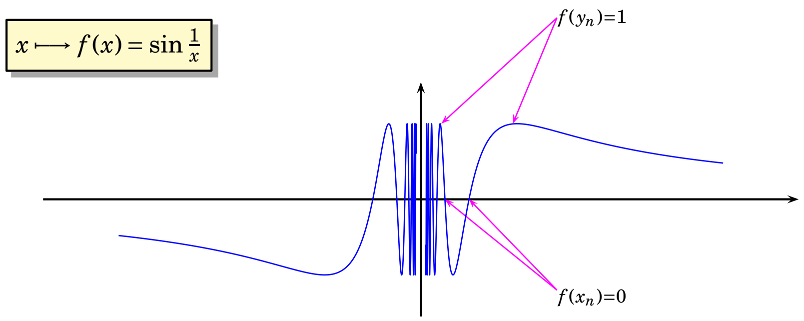
Dafür benutzen wir die folgende Abschätzung für $x > 0$:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} > \frac{x^{n+1}}{(n+1)!}, \quad \text{also} \quad \frac{e^x}{x^n} > \frac{x}{(n+1)!}$$

Zu $c \in \mathbb{R}$ wähle $d = \max\{0, c(n+1)!\}$. Für $x > d$ folgt dann $\dfrac{e^x}{x^n} > \dfrac{x}{(n+1)!} \geqslant \dfrac{c(n+1)!}{(n+1)!} = c$.

4.5.9. **Satz** (Folgenkriterium)**.** *Sei $f : D \longrightarrow \mathbb{C}$, sei $z_0 \in \overline{\mathbb{C}}$ ein HP von D, und sei $a \in \overline{\mathbb{C}}$. Es gilt $\lim_{z\to z_0} f(z) = a$ genau dann, wenn für alle gegen z_0 konvergenten Folgen $(z_n)_{n\in\mathbb{N}}$ in $D \setminus \{z_0\}$ gilt: $\lim_{n\to\infty} f(z_n) = a$. Analog für $D \subset \mathbb{R}$, $z_0 \in \overline{\mathbb{R}}$, $a \in \overline{\mathbb{R}}$.*

Mit dem Folgenkriterium sieht man leicht, dass $\lim_{x\to 0} \sin \frac{1}{x}$ nicht existiert:



Wir betrachten die Nullfolgen (x_n), $x_n = 1/n\pi$ und (y_n), $y_n = 1/(2n\pi + \pi/2)$. Dann gilt $f(x_n) \to 0$, $f(y_n) \to 1$ für $n \to \infty$. Wenn $\lim_{z\to z_0} f(z) = a$ existierte, dann müsste $a = 0$ und $a = 1$ sein, ein Widerspruch.

4.5.10. **Folgerung** (Rechenregeln für Grenzwerte)**.** *(i) Seien $f, g : D \longrightarrow \mathbb{C}$, und sei z_0 ein HP von D. Es gelte*

$$\lim_{z\to z_0} f(z) = a \in \overline{\mathbb{C}}, \qquad \lim_{z\to z_0} g(z) = b \in \overline{\mathbb{C}}.$$

Dann gilt:

$$\lim_{z\to z_0} (f+g)(z) = a + b, \ \lim_{z\to z_0} (f \cdot g)(z) = ab, \ \lim_{z\to z_0} \frac{f}{g}(z) = \frac{a}{b}$$

(letzteres falls $g(z) \neq 0$ in einer Umgebung von z_0), sofern die Ausdrücke auf der rechten Seiten definiert sind. Definiert sind hierbei

$$\infty + a = \infty, \ \frac{a}{\infty} = 0 \quad \text{für } a \in \mathbb{C},$$
$$\infty \cdot a = \infty, \ \frac{a}{0} = \infty \quad \text{für } a \in \mathbb{C} \setminus \{0\},$$
$$\infty \cdot \infty = \infty.$$

Nicht definierte (unbestimmte) Ausdrücke: $\infty \cdot 0$ und daraus abgeleitete Ausdrücke wie $\frac{0}{0}$, $\frac{\infty}{\infty}$ usw.

(ii) Analog für $f, g : D \longrightarrow \mathbb{R}$ und $a, b \in \overline{\mathbb{R}}$. Die definierten Ausdrücke sind in diesem Fall

$$\infty + a = \infty, \ -\infty + a = -\infty, \ \frac{a}{\infty} = \frac{a}{-\infty} = 0 \quad \text{für } a \in \mathbb{R},$$
$$\infty \cdot a = \infty \quad \text{für } a > 0, \ \infty \cdot a = -\infty \quad \text{für } a < 0,$$
$$\infty + \infty = \infty \cdot \infty = \infty,$$
$$-\infty + (-\infty) = (-\infty) \cdot \infty = -\infty.$$

Außerdem sind Ausdrücke definiert, welche aus den obigen durch Anwendung der Vorzeichenregeln oder des Kommutativgesetzes hervorgehen, zum Beispiel $(-\infty)(-\infty)=\infty$, $\infty-(-\infty)=\infty$. *Nicht definierte Ausdrücke:* $\infty-\infty$, $0\cdot\infty$ *und daraus abgeleitete Ausdrücke wie* $\frac{0}{0}$, $\frac{\infty}{\infty}$, $0\cdot(-\infty)$ *usw.*

(iii) (Vergleichsprinzip) Seien $f,g:D\longrightarrow\overline{\mathbb{R}}$, z_0 *ein HP von* D, $\lim_{z\to z_0}f(z)=a\in\overline{\mathbb{R}}$, $\lim_{z\to z_0}g(z)=b\in\overline{\mathbb{R}}$ *und* $f(z)\leqslant g(z)$ *für alle* $z\in D$ *mit* $z\neq z_0$. *Dann gilt* $a\leqslant b$.

(iv) (Substitutionsregel – erste Variante) Seien $f:D\longrightarrow\mathbb{C}$, $g:E\longrightarrow\mathbb{C}$ *mit* $f(D)\subset E$, $z_0\in\overline{\mathbb{C}}$ *ein HP von* D. *Falls* $\lim_{z\to z_0}f(z)=w_0\in E$ *und* g *stetig in* w_0 *ist, so gilt* $\lim_{z\to z_0}(g\circ f)(z)=g(w_0)$.

(v) (Substitutionsregel – zweite Variante). Seien f,g,z_0 *wie in (iv). Sei* $w_0\in\overline{\mathbb{C}}$ *ein HP von* E. *Falls* $f(z)\neq w_0$ *für* $z\neq z_0$, $\lim_{z\to z_0}f(z)=w_0$, $\lim_{w\to w_0}g(w)=a\in\overline{\mathbb{C}}$, *so ist* $\lim_{z\to z_0}(g\circ f)(z)=a$.

(v) (Grenzwerte der Umkehrfunktion). Seien I,J *Intervalle und* $f:I\to J$ *streng monoton und stetig. Sei* $x_0\in\overline{\mathbb{R}}$ *ein HP von* I *und* $y_0\in\overline{\mathbb{R}}$ *ein HP von* J. *Dann gilt* $\lim_{x\to x_0}f(x)=y_0$ *genau dann, wenn* $\lim_{y\to y_0}f^{-1}(y)=x_0$.

4.5.11. **Bemerkung.** (1) 4.5.10(iii) und (v) gelten auch dann, wenn die jeweiligen Voraussetzungen nur für $z\in D\cap U$, $z\neq z_0$ erfüllt sind, wobei U eine Umgebung von z_0 in $\overline{\mathbb{C}}$ ist.

(2) Die Substitutionsregel lässt sich durch das folgende Schema beschreiben:

$$f(z)\to w_0 \text{ für } z\to z_0 \text{ und } g(w)\to a \text{ für } w\to w_0 \overset{w=f(z)}{\Longrightarrow} g(f(z))\to a \text{ für } z\to z_0.$$

(3) Zur Grenzwerten der Umkehrfunktion (v): Wenn $x_0\in I$, so ist $y_0=f(x_0)$ und die Aussage folgt aus der Stetigkeit der Umkehrfunktion (Satz 4.1.10 oder Satz 4.3.3).

4.5.12. **Beispiel.** (1) Wir wollen $\lim_{x\to-\infty}e^x$ berechnen. Wir schreiben $e^x=e^{-(-x)}=(g\circ f)(x)$, wobei $g:\mathbb{R}_+\longrightarrow\mathbb{R}_+$, $g(y)=e^{-y}$, $f:(-\infty,\infty)\longrightarrow\mathbb{R}_+$, $f(x)=-x$. Dann ist $\lim_{y\to\infty}g(y)=0$, weil $\lim_{y\to\infty}e^y=\infty$ (nach (4.6)) und $\lim_{x\to-\infty}f(x)=\infty$. Folgerung 4.5.10(v) mit $z_0=-\infty$, $a=\infty$ impliziert $\lim_{x\to-\infty}(g\circ f)(x)=0$. Also

$$\lim_{x\to-\infty}e^x=0\,. \tag{4.7}$$

Die Rechnung auf einem Schlag:

$$\lim_{x\to-\infty}e^x=\lim_{x\to-\infty}\frac{1}{e^{-x}}\overset{y=-x}{=}\lim_{y\to\infty}\frac{1}{e^y}=0\,.$$

19.12.2011 (2) Wir wollen $\lim_{x\to0}\frac{\log(1+x)}{x}$ berechnen. Für die Substitution $y=f(x)=\log(x+1)$, wobei $f:(-1,\infty)\to\mathbb{R}$, gilt $\lim_{x\to0}f(x)=f(0)=0$, also $y\to0$ wenn $x\to0$. Ferner gilt $x=e^y-1$, also $\frac{\log(1+x)}{x}=\frac{y}{e^y-1}=g(y)$, wobei $g:\mathbb{R}\setminus\{0\}\to\mathbb{R}$ und $\lim_{y\to0}g(y)=1$ nach (4.4). Ferner ist $f(x)=\log(1+x)\neq0$ für $x\neq0$. Also ergibt sich

$$\lim_{x\to0}\frac{\log(1+x)}{x}\overset{y=\log(x+1)}{=}\lim_{y\to0}\frac{y}{e^y-1}=1\,.$$

(3) Durch Substitution $y=\log x$ (also $x=e^y$) ergibt sich

$$\lim_{x\to\infty}\frac{\log x}{x}=\lim_{y\to\infty}\frac{y}{e^y}=0\,.$$

(4) Für $x\to\infty$ gilt $\frac{\log x}{x}\to0$, also

$$x^{1/x}=\exp\log x^{1/x}=\exp(\tfrac{\log x}{x})\longrightarrow\exp(0)=1\,.$$

Daraus ergibt sich wieder der am Anfang mit Mühe gewonnene Grenzwert aus Satz 2.1.3: $\lim_{n\to\infty}\sqrt[n]{n}=1$.

4.5.13. **Bemerkung.** (Folgen-Grenzwerte und Funktionen-Grenzwerte) Man kann die Rechnung der Funktionen-Grenzwerte nutzen, um Folgen-Grenzwerte zu berechnen, wie in letztem Beispiel. Sei $f:[k,\infty)\to\mathbb{C}$, $k\in\mathbb{N}$, so dass $\lim_{x\to\infty}f(x)=\ell$. Dann gilt auch $\lim_{n\to\infty}f(n)=\ell$ (Folgenkriterium 4.5.9). Die Umkehrung ist falsch, z. B. gilt $\lim_{n\to\infty}\sin(n\pi)=0$, aber $\lim_{x\to\infty}\sin(\pi x)$ existiert nicht.

4.5.14. **Bemerkung.** (unbestimmte Ausdrücke) Die Rechenregeln geben nur dann Auskunft, wenn die Ausdrücke auf der rechten Seite bestimmt sind. Wie behandelt man unbestimmte Ausdrücke wie $\frac{0}{0}$, $\frac{\infty}{\infty}$, $\infty-\infty$, $0\cdot\infty$ usw.? Dazu ein paar Vorschläge:

- Zurückführen auf bekannte Grenzwerte durch Umformungen (z.B. Kürzen oder Erweitern von Brüchen) oder Substitutionen. Exponentialausdrücke behandelt man mittels der Definition $f^g:=e^{g\log f}$.

- Ist der unbestimmte Ausdruck von der Form f/g, und sind die auftretenden Funktionen durch Potenzreihen definiert, so schreibt man einige Potenzen der Reihe auf und behandelt den Quotienten f/g als eine rationale Funktion (siehe Beispiel 4.5.8 (1)).
- Die Regel von l'Hospital anwenden.

Die ersten zwei Methoden spenden mehr Einsicht; es empfiehlt deshalb, sie vorzuziehen.

4.5.15. **Definition** (Einseitige Grenzwerte)**.** Sei $D \subset \mathbb{R}$, $x_0 \in \mathbb{R}$.

(i) x_0 heißt rechtsseitiger (bzw. linksseitiger) HP von D, wenn x_0 HP von $(x_0,\infty)\cap D$ (bzw. $(-\infty,x_0)\cap D$) ist.

(ii) Sei x_0 rechtsseitiger (bzw. linksseitiger) HP von D. Die Funktion $f: D \longrightarrow \mathbb{R}$ hat in x_0 den rechtsseitigen (bzw. linksseitigen) Grenzwert $a \in \overline{\mathbb{R}}$, wenn $f|_{D\cap(x_0,+\infty)}$ (bzw. $f|_{(-\infty,x_0)\cap D}$) den Grenzwert $a \in \overline{\mathbb{R}}$ hat.

Notation:

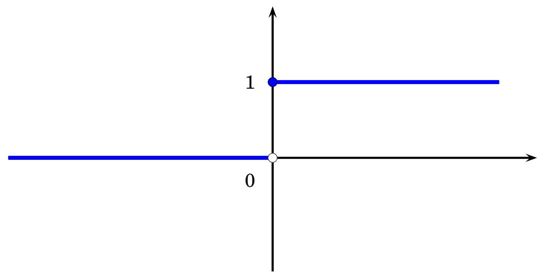
$$f(x_0+) := \lim_{x\searrow x_0} f(x) := \lim_{x\to x_0} f|_{D\cap(x_0,+\infty)} = a\,,$$
$$f(x_0-) := \lim_{x\nearrow x_0} f(x) := \lim_{x\to x_0} f|_{(-\infty,x_0)\cap D} = a\,.$$

4.5.16. **Beispiel.**

(1) Die ***Heaviside-Funktion*** ist definiert durch

$$f:\mathbb{R}\longrightarrow\mathbb{R},\quad f(x)=\begin{cases}1\,, x\geqslant 0\,,\\ 0\,, x<0\,.\end{cases} \tag{4.8}$$

Dann $f(0-)=0$, $f(0+)=1$.



(2) $f:\mathbb{R}\setminus\{0\}\longrightarrow\mathbb{R}$, $f(x)=\sin\frac{1}{x}$. Die Grenzwerte $f(0-)$, $f(0+)$ existieren nicht.

4.5.17. **Satz.**

(i) *Sei $f: D \longrightarrow \mathbb{R}$, $D\subset\mathbb{R}$, und sei x_0 ein links- und rechtsseitiger HP von D. Der Grenzwert $\lim_{x\to x_0} f(x)$ existiert genau dann, wenn die Grenzwerte $f(x_0-)$ und $f(x_0+)$ beide existieren und übereinstimmen. In diesem Fall gilt $\lim_{x\to x_0} f(x)=f(x_0-)=f(x_0+)$.*

(ii) *Sei zusätzlich $x_0\in D$. Die Funktion f ist stetig in x_0 genau dann, wenn die Grenzwerte $f(x_0-)$, $f(x_0+)$ in $\mathbb{R}$ existieren und $f(x_0)=f(x_0-)=f(x_0+)$ gilt.*

4.5.18. **Definition.** Seien f, x_0 wie in 4.5.17. Der Punkt x_0 heißt ***Unstetigkeitsstelle erster Art***, wenn $f(x_0-), f(x_0+)\in\mathbb{R}$ und $f(x_0-)\neq f(x_0+)$ oder $f(x_0-)=f(x_0+)\neq f(x_0)$. Der Punkt x_0 heißt ***Unstetigkeitsstelle zweiter Art***, falls x_0 eine Unstetigkeitsstelle ist, die nicht von erster Art ist.

4.6. Übungen.

4.6.1. ***Aufgabe***. Zeige mit Hilfe der ε-δ–Definition 4.1.1 der Stetigkeit die folgenden Behauptungen:

(a) Die Funktion $f:\mathbb{C}\to\mathbb{C}$, $f(z)=z^3$ ist stetig.

(b) Die Funktion $f:\mathbb{C}\to\mathbb{C}$, $f(z)=\begin{cases}\frac{\operatorname{Re}z}{|z|}\,, & z\neq 0\\ 0\,, & z=0\end{cases}$ ist nicht stetig in $z_0=0$.

(c) Die Funktion $f:\mathbb{C}\to\mathbb{C}$, $f(z)=\begin{cases}\frac{(\operatorname{Re}z)^2}{|z|}\,, & z\neq 0\\ 0\,, & z=0\end{cases}$ ist stetig in $z_0=0$.

4.6.2. ***Aufgabe***. Finde die Stetigkeits- und Unstetigskeitsstellen folgender Funktionen:

(a) $f:\mathbb{R}\to\mathbb{R}$, mit $f(x)=\begin{cases} x^2 & ,x\in\mathbb{Q}\\ 0 & ,x\in\mathbb{R}\setminus\mathbb{Q}\end{cases}$

(b) $g:\mathbb{R}\to\mathbb{R}$, mit $g(x)=\begin{cases} \dfrac{\cos x-1}{x} & ,x\neq 0\\ 0 & ,x=0\end{cases}$

(c) $\varphi:\mathbb{R}\to\mathbb{R}$, mit $\varphi(x)=\lfloor x+\frac{1}{2}\rfloor$

(d) $\psi:\mathbb{R}\to\mathbb{R}$, mit $\psi(x)=|\varphi(x)-x|$.

4.6.3. ***Aufgabe***.

(a) Zeigen Sie $\exp(z)\neq 0$ für alle $z\in\mathbb{C}$, $\exp(x)>0$ für alle $x\in\mathbb{R}$, $\exp(x)>1$ für alle $x\in\mathbb{R}_+$ und $\exp(x)<\exp(y)$ für alle $x,y\in\mathbb{R}$ mit $x<y$. (Tipp: $\exp(z+w)=\exp(z)\exp(w)$.)

(b) Zeigen Sie $|\exp(u)-1-u|\le|u|^2$ für alle $u\in\mathbb{C}$ mit $|u|<1$ und folgern Sie:
$|\exp(w)-\exp(z)-(w-z)\exp(z)|\le|w-z|^2\cdot|\exp(z)|$ für alle $z,w\in\mathbb{C}$ mit $|w-z|<1$.

(c) Beweisen Sie mit Hilfe von (b), dass $\exp:\mathbb{C}\to\mathbb{C}$ stetig ist.

4.6.4. ***Aufgabe***. Für $x>0$ und $z\in\mathbb{C}$ definieren wir $x^z:=\exp(z\cdot\log x)$ (Definition 4.3.5). Als Komposition stetiger Funktionen sind sowohl $\mathbb{R}_+\ni x\mapsto x^z\in\mathbb{C}$ (für festes $z\in\mathbb{C}$) als auch $\mathbb{C}\ni z\mapsto x^z\in\mathbb{C}$ (für festes $x>0$) stetig. Zeigen Sie:

(a) Für alle $x>0$, $q\in\mathbb{Q}$ stimmt die obige Definition von x^q mit unserer früheren Definition 1.5.9 überein.

(b) Für alle $x,y>0$, $z,w\in\mathbb{C}$ und $a\in\mathbb{R}$ gilt: $(xy)^z=x^zy^z$, $x^{z+w}=x^zx^w$, $(x^a)^z=x^{az}$; für $x>1$ und $a<b$ gilt $x^a<x^b$; für $a>0$ und $x<y$ gilt $x^a<y^a$.

4.6.5. ***Aufgabe***. Berechne die folgenden Grenzwerte:

(i) $\lim\limits_{x\to 0}\log x$, $\lim\limits_{x\to\infty}\log x$, $\lim\limits_{x\to 0}(1+x)^{\frac{1}{x}}$, $\lim\limits_{y\to\infty}(1+\frac{1}{y})^y$.

(ii) Sei $a>0$.
$\lim\limits_{x\to\infty}a^x$, $\lim\limits_{x\to-\infty}a^x$, $\lim\limits_{x\to 0}x^a$,
$\lim\limits_{x\to\infty}x^a$, $\lim\limits_{x\to 0}x^a\log x$, $\lim\limits_{x\to 0}\dfrac{\log x}{x^a}$.

(iii) $\lim\limits_{z\to\infty}P(z)$, wobei $P:\mathbb{C}\longrightarrow\mathbb{C}$, $P(z)=a_nz^n+\ldots+a_0$, $a_n\neq 0$.

4.6.6. ***Aufgabe***. Sei $f:[a,b]\longrightarrow\mathbb{R}$ streng monoton (steigend oder fallend). Sei $E=\{f(x):x\in[a,b]\}$. Dann ist $f:[a,b]\longrightarrow E$ bijektiv, und wir können $g:E\longrightarrow[a,b]$, $g:=f^{-1}$ definieren. Zeigen Sie: Falls f stetig in x_0 ist, dann ist g stetig in $y_0=f(x_0)$.

4.6.7. ***Aufgabe***. Eine Menge $\Omega\subset\mathbb{C}$ heißt *offen*, wenn es für jedes $z\in\Omega$ ein $\varepsilon>0$ mit $B_\varepsilon(z)\subset\Omega$ gibt. Eine Menge $A\subset\mathbb{C}$ heißt *abgeschlossen*, wenn ihr Komplement $\complement A:=\mathbb{C}\setminus A$ offen ist. Zeigen Sie:

(a) Die Vereinigung $\bigcup_{\lambda\in\Lambda}\Omega_\lambda$ einer beliebigen Familie $(\Omega_\lambda)_{\lambda\in\Lambda}$ offener Mengen $\Omega_\lambda\subset\mathbb{C}$ ist offen, und der Durchschnitt einer endlichen Familie offener Mengen ist offen. Was gilt für abgeschlossene Mengen?

(b) Eine Menge $A\subset\mathbb{C}$ ist genau dann abgeschlossen, wenn für jede konvergente Folge, deren Glieder sämtlich in A liegen, auch deren Grenzwert in A liegt.

4.6.8. ***Aufgabe***. Sei $D\subset\mathbb{C}$, und sei $\overline{D}$ die Vereinigung von D und der Menge aller Häufungspunkte von D. Beweisen Sie, dass $\overline{D}$ die kleinste abgeschlossene Menge ist, die D umfasst.
(Zu zeigen ist: $\overline{D}$ ist abgeschlossen, und jede abgeschlossene Menge $A\subset\mathbb{C}$, die D umfasst, umfasst auch $\overline{D}$.)

4.6.9. ***Aufgabe*** (lokaler Charakter der Stetigkeit).

(a) Sei $D\subset\mathbb{C}$. Eine Funktion $f:D\longrightarrow\mathbb{C}$ ist stetig genau dann, wenn jeder Punkt $z\in D$ eine Umgebung U_z von z in D besitzt, so dass $f|_{U_z}$ stetig ist.

(b) Seien $D_1,\ldots,D_m\subset\mathbb{C}$ und seien $f_k:D_k\to\mathbb{C}$ $(1\leqslant k\leqslant m)$ stetige Funktionen mit der Eigenschaft $f_i(x)=f_j(x)$ für alle $x\in D_i\cap D_j$. Man kann also $f:\bigcup_{k=1}^m D_k\to\mathbb{C}$ durch $f(x):=f_k(x)$ für $x\in D_k$ definieren. Wenn $D_1,\ldots,D_m$ offene Mengen sind, so ist f stetig.

(c) Man kann in (b) die Voraussetzung über die Offenheit der Mengen D_k nicht fortlassen.

4.6.10. ***Aufgabe***. Mit dieselben Notationen zeigen Sie: Wenn $D_1,\dots,D_m$ absgeschlossene Mengen sind, so ist f stetig. Kann man die Voraussetzung über die Abgeschlossenheit der Mengen D_k fortlassen?

4.6.11. ***Aufgabe***. Zeige die folgende Aussage: Sind $f,g:[a,b]\to\mathbb{R}$ beide stetig auf $[a,b]$ und gilt $f(x)=g(x)$ für alle $x\in\mathbb{Q}\cap[a,b]$ so gilt auch $f(x)=g(x)$ für alle $x\in[a,b]$.

4.6.12. ***Aufgabe*** (zur Definition 4.3.5 der Exponentialfunktion). Benutze Satz 3.3.8 und Übung 4.6.11 um zu zeigen: Es gibt für $a>0$ genau eine stetige Lösung $\varphi:\mathbb{R}\to\mathbb{R}$ der Funktionalgleichung der Exponentialfunktion (3.6)

$$\varphi(z+w)=\varphi(z)\varphi(w),\quad z,w\in\mathbb{C}$$

mit $\varphi(1)=a$, nämlich $\varphi(r)=a^r$, $r\in\mathbb{R}$.

4.6.13. ***Aufgabe***. Sei $\alpha\in\mathbb{R}$. Wir wissen, dass die Binomialreihe (3.10) für alle $x\in\mathbb{R}$ mit $|x|<1$ konvergent ist. Setze für $x\in\mathbb{R}$ mit $|x|<1$

$$B_\alpha(x)=\sum_{n=1}^{\infty}\binom{\alpha}{n}x^n=1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3+\cdots$$

(i) Sei $m>0$ und $\alpha\in[-m,m]$. Zeige, dass $\left|\binom{\alpha}{k}\right|\le\binom{m+k-1}{k}$ und die Reihe $\sum_{k\ge0}\binom{m+k-1}{k}|x|^k$ konvergent ist.

(ii) Zeige, dass die Funktionenreihe $\sum_{k\ge0}f_k$, $f_k:[-m,m]\to\mathbb{R}$, $f_k(\alpha)=\binom{\alpha}{k}x^k$, normal konvergent ist. Leite her, dass die Funktion $[-m,m]\ni\alpha\mapsto B_\alpha(x)$ stetig ist.

(iii) Beweise die ***Binomialentwicklung von Newton***:

$$(1+x)^\alpha=\sum_{k=0}^{\infty}\binom{\alpha}{k}x^k\,,\quad \alpha\in\mathbb{R}\,,\ x\in(-1,1)\,. \tag{4.9}$$

Hinweis: Benutze die Aufgaben 3.5.8 und 4.6.11.

4.6.14. ***Aufgabe*** (Stetigkeit der Zeta-Funktion im komplexen). Zeigen Sie:

(a) $|e^{it}|=1$ für alle $t\in\mathbb{R}$, und $|e^z|=e^{\operatorname{Re}z}$ für alle $z\in\mathbb{C}$.

(b) Sei $s\in\mathbb{C}$ mit $\operatorname{Re}s>1$. Dann konvergiert $\zeta(s):=\sum_{n=1}^{\infty}n^{-s}$ absolut.

(c) Für jedes $\varepsilon>0$ ist die Funktionenreihe $\sum_{n=1}^{\infty}f_n$ mit $f_n(s):=n^{-s}$ auf $D_\varepsilon:=\{s\in\mathbb{C}:\operatorname{Re}s\geqslant1+\varepsilon\}$ normal konvergent, und die Riemannsche Zeta-Funktion $\zeta:\{s\in\mathbb{C}:\operatorname{Re}s>1\}\to\mathbb{C}$ ist stetig.

4.6.15. ***Aufgabe***. Beweise den Zwischenwertsatz mit Hilfe einer Intervallschachtellung.

4.6.16. ***Aufgabe***. Es sei sei I ein nichttriviales kompaktes Intervall und $f:I\to I$ sei stetig. Zeige, das f einen Fixpunkt besitzt; d.h., es gibt ein $\xi\in I$ mit $f(\xi)=\xi$.

4.6.17. ***Aufgabe***. Sei $n\in\mathbb{N}$ ungerade, $P:\mathbb{R}\longrightarrow\mathbb{R}$, $P(x)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0$ ein reelles Polynom mit $a_n\ne0$. Dann hat P eine Nullstelle.

4.6.18. ***Aufgabe***. Zeigen Sie die folgenden Aussagen über die Einschränkungen der stetigen Funktionen sinh und cosh auf $\mathbb{R}$ bzw. $[0,\infty)$ (Tipp: "surjektiv" jeweils mit Zwischenwertsatz):

(a) $\sinh:\mathbb{R}\to\mathbb{R}$ ist stetig, streng monoton wachsend und bijektiv.

(b) $\cosh:[0,\infty)\to\mathbb{R}$ ist stetig, hat Bild $[1,\infty)$ und ist streng monoton wachsend. (Also ist $\cosh:[0,\infty)\to[1,\infty)$ bijektiv.)

4.6.19. ***Aufgabe***. (a) Seien $a_1,\dots,a_n\in\mathbb{R}$ mit $a_n>0$ und $a_0<0$. Zeige: Das Polynom $p(x):=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0$ besitzt eine positive Nullstelle.

(b) Zeige, daß die Gleichung $x^2+2=\exp(x)$ eine positive Lösung besitzt.

4.6.20. ***Aufgabe***. Betrachte die Funktion

$$f:[0,1]\to[0,1],\quad f(x)\begin{cases}x, & x \text{ rational},\\ 1-x, & x \text{ irrational},\end{cases}$$

und zeige: (a) f ist bijektiv, (b) f ist auf keinem Teilintervall von $[0,1]$ monoton, (c) f ist nur im Punkt 1/2 stetig.

4.6.21. ***Aufgabe***. Sei $f:(a,b)\to\mathbb{R}$ eine monotone Funktion. Zeige: (a) Für alle $x\in(a,b)$ existieren $f(x-)$, $f(x+)$, (b) f ist stetig an der Stelle $x\in(a,b)$ genau dann, wenn $f(x-)=f(x+)$, (c) Die Menge der Unstetigkeitsstellen von f ist höchstens abzählbar.

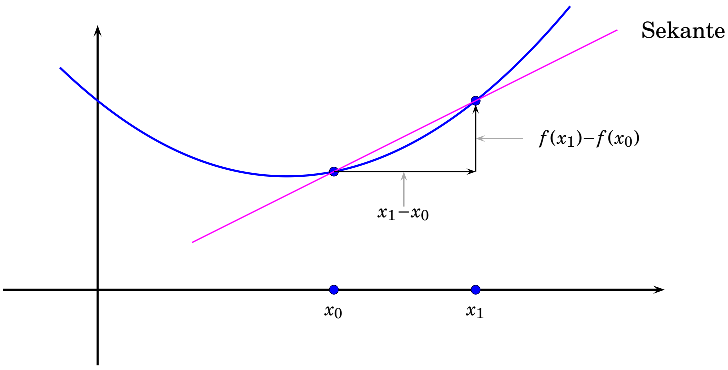
5. Differenzierbarkeit

5.1. **Definition und erste Eigenschaften.** Sei $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$, $f : I \longrightarrow \mathbb{C}$. Wir suchen die lineare (genauer: affine) Funktion

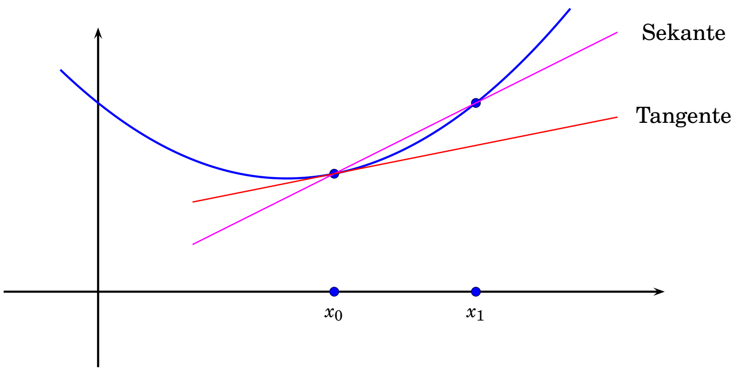
$$l : \mathbb{R} \longrightarrow \mathbb{C}, \quad l(x) = \lambda(x - x_0) + \mu,$$

die f in der Nähe von x_0 am besten approximiert. Wir ermitteln diese affine Funktion zunächst *geometrisch*. Wir betrachten die Gerade durch $(x_0, f(x_0))$ und $(x_1, f(x_1))$; sie heißt Sekante und ist der Graph der linearen (affinen) Abbildung

$$x \mapsto \frac{f(x_1) - f(x_0)}{x_1 - x_0} \cdot (x - x_0) + f(x_0).$$



Dabei ist $\frac{f(x_1)-f(x_0)}{x_1-x_0}$ die Steigung der Sekante, d.h. der Tangens des Winkels zwischen der x-Achse und der Sekante. Der Quotient $\frac{f(x_1)-f(x_0)}{x_1-x_0}$ heißt auch ein ***Differenzenquotient*** von f. Intuitiv ist der Graph der besten linearen Approximation die Grenzlage der Sekante, wenn sich $(x_1, f(x_1))$ auf $(x_0, f(x_0))$ hin bewegt. Diese Gerade, falls sie existiert, heißt ***Tangente*** an den Graphen von f im Punkte $(x_0, f(x_0))$.



Die Tangente erfüllt also zwei Bedingungen: $(x_0, f(x_0))$ liegt auf der Tangente (also $\mu = f(x_0)$), und ihre Steigung ist der Grenzwert des Differenzenquotienten $\frac{f(x_1)-f(x_0)}{x_1-x_0}$ für $x_1 \to x_0$. Daher die folgende Definition:

5.1.1. **Definition.** Sei $I \subset \mathbb{R}$ Intervall. Die Funktion $f : I \longrightarrow \mathbb{C}$ heißt ***differenzierbar im Punkt*** $x_0 \in I$, wenn $\lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0}$ in $\mathbb{C}$ existiert. Dann heißt dieser Limes ***Ableitung*** von f in x_0. Schreibweise:

$$f'(x_0) := \frac{df}{dx}(x_0) := \frac{df}{dx}\Big|_{x=x_0} := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Ist f in allen Punkten $x_0 \in I$ differenzierbar, so heißt f ***differenzierbar*** in I. Die Abbildung $f' : I \longrightarrow \mathbb{C}$, $x \mapsto f'(x)$ heißt die Ableitung von f.

Falls x_0 ein Randpunkt des Intervalls I ist, so ist der Grenzwert in der Definition als einseitiger Grenzwert zu verstehen (da die Punkte x immer in I liegen müssen).

Die affine Funktion

$$l:\mathbb{R}\longrightarrow\mathbb{C},\quad l(x)=f'(x_0)(x-x_0)+f(x_0) \tag{5.1}$$

heißt die lineare (genauer affine) Approximation von f bei x_0. Ihr Graph ist also die Tangente an den Graphen von f im Punkte $(x_0,f(x_0))$.

5.1.2. **Beispiele.** **(1)** Sei $f:I\longrightarrow\mathbb{C}$, $f(x)=c$ (konstant). Dann ist f differenzierbar und $f'(x)=0$ für alle $x\in I$. Kurz (aber mit einem Missbrauch der Notation):

$$\boxed{c'=0}$$

(2) Sei $n\in\mathbb{N}$, $f:\mathbb{R}\longrightarrow\mathbb{R}$, $f(x)=x^n$. Dann ist f differenzierbar und $f'(x)=nx^{n-1}$ für alle $x\in\mathbb{R}$.

$$\boxed{(x^n)'=nx^{n-1},\quad n\in\mathbb{N}}$$

(3) Sei $c\in\mathbb{C}$, $f:\mathbb{R}\longrightarrow\mathbb{C}$, $f(x)=\exp(cx)$. Dann ist f differenzierbar und $f'(x)=c\exp(cx)=cf(x)$ für alle $x\in\mathbb{R}$.

$$\boxed{(\exp(cx))'=c\exp(cx),\quad c\in\mathbb{C}} \tag{5.2}$$

Zu (3): Ist $c=0$ so ist f konstant und beide Seiten sind Null. Ist $c\neq0$, so gilt laut (3.5) und (4.4):

$$\frac{\exp(cx)-\exp(cx_0)}{x-x_0}=c\exp(cx_0)\frac{\exp(cx-cx_0)-1}{c(x-x_0)}\longrightarrow c\exp(cx_0),\quad \text{für } x\to x_0\,.$$

(4) Wenn das Argument einer Funktion f die Zeit bezeichnet, so schreibt man dafür t statt x und bezeichnet die Ableitung mit einem Punkt statt mit einem Strich: $\dot f:=df/dt$. Bewegt sich ein Punkt entlang einer Geraden (x-Achse), so kann die Bewegung dadurch beschrieben werden, dass der Ort x zur Zeit t als Funktion $x=x(t)$ vorgegeben wird. Der zugehörige Differenzenquotient $\frac{x(t)-x(t_0)}{t-t_0}$ ist die mittlere Geschwindigkeit zwischen den Zeiten t_0 und t, und

$$\boxed{v(t_0):=\dot x(t_0)=\lim_{t\to t_0}\frac{x(t)-x(t_0)}{t-t_0}}$$

ist die momentane Geschwindigkeit zur Zeit t_0. Sie beschreibt also die momentane Änderung des Ortes und wird meist nur ***Geschwindigkeit*** genannt. Die momentane Änderung der Geschwindigkeit ist die ***Beschleunigung***:

$$\boxed{\dot v(t_0)=\ddot x(t_0)} \tag{5.3}$$

Das ***Newtonsche Kraftgesetz*** „Kraft gleich Masse mal Beschleunigung“ führt zur ***Bewegungsgleichung***:

$$\boxed{m\ddot x(t)=K(x(t),\dot x(t),t)}$$

wobei m die Masse des Punktes und $K(x,\dot x,t)$ die Kraft ist.

5.1.3. **Satz.** *Sei $f:I\longrightarrow\mathbb{C}$ differenzierbar in $x_0\in I$. Dann ist f stetig in x_0.*

Die Umkehrung ist falsch. Die Funktion $f:\mathbb{R}\longrightarrow\mathbb{R}$, $f(x)=|x|$ liefert ein Gegenbeispiel; sie ist stetig in 0, aber

$$\lim_{x\nearrow0}\frac{f(x)-f(0)}{x-0}=\lim_{x\nearrow0}\frac{-x}{x}=-1$$
$$\lim_{x\searrow0}\frac{f(x)-f(0)}{x-0}=\lim_{x\searrow0}\frac{x}{x}=1\,.$$

Wegen Satz 4.5.17(i) existiert $\lim_{x\to0}\frac{f(x)-f(0)}{x-0}$ nicht. Man kann aber sagen, dass f einseitig differenzierbar ist, siehe dazu die Notizen §5.6.

Es gibt Funktionen, die auf ganz $\mathbb{R}$ stetig, aber nirgends differenzierbar sind! Siehe [13, § 9.11].

22.12.2011

5.2. **Ableitungsregeln.**

5.2.1. **Satz** (Ableitungsregeln). *Sind $f, g : I \longrightarrow \mathbb{C}$ in $x_0 \in I$ differenzierbar, so sind auch die Funktionen $f+g$, λf $(\lambda \in \mathbb{C})$, fg in x_0 differenzierbar; falls $g(x_0) \neq 0$, so ist auch $\dfrac{f}{g}$ in x_0 differenzierbar. Es gilt dann*

$$(f+g)'(x_0) = f'(x_0) + g'(x_0), \quad (\lambda f)'(x_0) = \lambda f'(x_0),$$
$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0) \quad \textit{(Produktregel)},$$
$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2} \quad \textit{(Quotientregel)}.$$

5.2.2. **Beispiel.** **(1)** Sei $c \in \mathbb{C}$ und f sei differenzierbar in x_0. Dann ist auch cf differenzierbar in x_0 und $(cf)'(x_0) = cf'(x_0)$ wegen der Produktregel und $c' = 0$.

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

$$\boxed{\sin' x = \cos x, \quad \cos' x = -\sin x}$$

Analog

$$\boxed{\sinh' x = \cosh x, \quad \cosh' x = \sinh x}$$

(3) $\tan x := \dfrac{\sin x}{\cos x}$ (wo $\cos x \neq 0$) $\rightsquigarrow$

$$\boxed{\tan' x = \frac{1}{\cos^2 x} = 1 + \tan^2 x}$$

Die Funktion tan erfüllt also die Differentialgleichung (DGL) $f' = 1 + f^2$.

5.2.3. **Satz** (Kettenregel). *Seien $I, J \subset \mathbb{R}$ Intervalle und $f : I \longrightarrow J$, $g : J \longrightarrow \mathbb{C}$. Ist f in $x_0 \in I$ und g in $y_0 := f(x_0)$ differenzierbar, so ist $g \circ f$ in x_0 differenzierbar, und*

$$(g \circ f)'(x_0) = g'(y_0) \cdot f(x_0) = g'(f(x_0)) \cdot f(x_0).$$

5.2.4. **Beispiel.** **(1)** Sei f sei differenzierbar in x_0. Dann ist auch e^f differenzierbar in x_0 und $(e^f)'(x_0) = e^{f(x_0)} f'(x_0)$, kurz

$$\boxed{(e^f)' = e^f f'}$$

wenn f differenzierbar ist.

(2) Ist f differenzierbar und $n \in \mathbb{N}$, so gilt

$$\boxed{(f^n)' = nf^{n-1}f'}$$

(3) Sei $a > 0$ fest, $f : \mathbb{R} \longrightarrow \mathbb{R}_+$, $f(x) = a^x := e^{x \log a}$ (Definition 4.3.5). Dann ist $f = g \circ h$, wobei $h : \mathbb{R} \longrightarrow \mathbb{R}$, $h(x) = x \log a$, $g : \mathbb{R} \longrightarrow \mathbb{R}_+$, $g(y) = e^y$. Es gilt $h'(x) = x' \log a + x(\log a)' = \log a$, $g'(y) = e^y$. Nach Kettenregel folgt $(a^x)' = e^{x \log a} \cdot (x \log a)' = a^x \log a$.

$$\boxed{(a^x)' = a^x \log a, \quad x \in \mathbb{R}, a > 0}$$

Insbesondere:

$$\lim_{x \to 0} \frac{a^x - 1}{x} = \log a.$$

5.2.5. **Satz** (Differenzierbarkeit der Umkehrfunktion). *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ stetig, streng monoton und differenzierbar in $x_0 \in I$. Sei $f'(x_0) \neq 0$. Dann ist die Umkehrfunktion $g : f(I) \longrightarrow \mathbb{R}$ differenzierbar in $y_0 := f(x_0)$, und es gilt*

$$g'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(g(y_0))} \quad \textit{oder } (f^{-1})'(y_0) = (f'(x_0))^{-1}. \tag{5.4}$$

5.2.6. **Beispiele.** **(1)** $\log : \mathbb{R}_+ \longrightarrow \mathbb{R}$ ist die Umkehrfunktion von $\exp : \mathbb{R} \longrightarrow \mathbb{R}_+$, also

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

$$\boxed{\log' x = \frac{1}{x}, \quad x > 0}$$

(2) Sei $z \in \mathbb{C}$ und $f : \mathbb{R}_+ \longrightarrow \mathbb{C}$, $f(x) = x^z = e^{z \log x}$. Dann ist f differenzierbar und $f'(z) = e^{z \log x} \cdot (z \log x)' = x^z \cdot z \cdot \frac{1}{x} = zx^{z-1}$.

$$\boxed{(x^z)' = zx^{z-1}, \quad x > 0, z \in \mathbb{C}}$$

5.3. **Mittelwertsätze.** Wir befassen uns nun mit dem vermutlich nützlichsten Satz der Analysis, dem Mittelwertsatz, und dessen Anwendungen.

Zur Erinnerung (Definition 4.1.4), eine Teilmenge $U \subset D \subset \mathbb{R}$ heißt Umgebung eines Punktes $x_0 \in D$ in D wenn es ein $\delta > 0$ mit $(x_0 - \delta, x_0 + \delta) \cap D \subset U$ gibt.

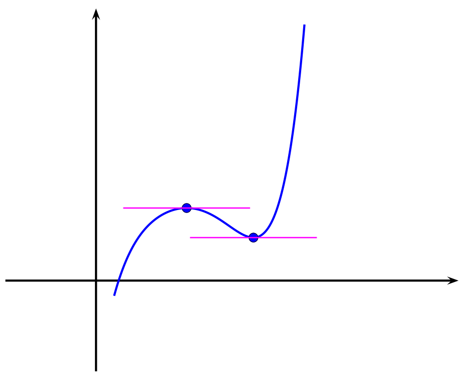
5.3.1. **Definition.** Sei $I \subset \mathbb{R}$ ein Intervall. Ein Punkt $x_0 \in I$ heißt ***innerer Punkt*** von I :⇔ es gibt eine Umgebung U von x_0 in $\mathbb{R}$ mit $U \subset I$ ⇔ es gibt $\delta > 0$ mit $(x_0 - \delta, x_0 + \delta) \subset I$ ⇔ $x_0 \notin \{\sup I, \inf I\}$. Die Menge $\mathring{I} := \{x \in I : x \text{ innerer Punkt von } I\} = (\sup I, \inf I)$ heißt ***das Innere*** von I.

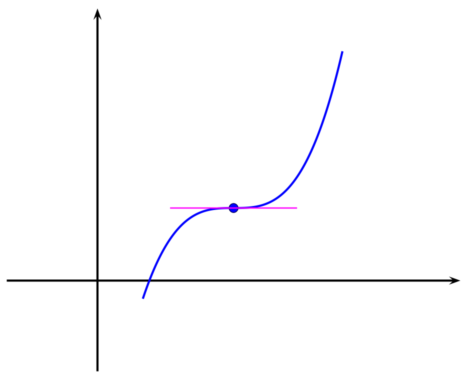
5.3.2. **Definition.** Sei $I \subset \mathbb{R}$ ein Intervall und $x_0 \in I$.
(1) Die Funktion $f : I \longrightarrow \mathbb{R}$ hat in x_0 ein ***lokales Maximum*** (bzw. ***Minimum***) :⇔ es gibt eine Umgebung U von x_0 in I mit $f(x) \leqslant f(x_0)$ (bzw. $f(x) \geqslant f(x_0)$) für alle $x \in U$. Ein Punkt heißt ***lokales Extremum***, falls er ein lokales Maximum oder ein lokales Minimum ist.
(2) Die Funktion $f : I \longrightarrow \mathbb{R}$ hat in x_0 ein ***globales Maximum*** (bzw. ***Minimum***) :⇔ es gilt $f(x) \leqslant f(x_0)$ (bzw. $f(x) \geqslant f(x_0)$) für alle $x \in I$. Ein Punkt heißt ***globales Extremum***, falls er ein globales Maximum oder ein globales Minimum ist.
(3) Ein innerer Punkt $x_0 \in I$ heißt ***kritischer Punkt*** für $f : I \longrightarrow \mathbb{R}$, falls f in x_0 differenzierbar ist und $f'(x_0) = 0$ gilt.
(4) Ein kritischer Punkt $x_0 \in I$ heißt ***Sattelpunkt*** von f, wenn es ein $\delta > 0$ gibt mit $(x_0 - \delta, x_0 + \delta) \subset I$ und $f(x) < f(x_0)$ für $x \in (x_0 - \delta, x_0)$ und $f(x) > f(x_0)$ für $x \in (x_0, x_0 + \delta)$ oder umgekehrt.

Um die x-Koordinate des globalen/lokalen Maximums bzw. Minimums vom dort angenommenen Funktionswert zu unterscheiden, spricht man auch von einer globalen/lokalen Maximumsstelle bzw. Minimumsstelle.

5.3.3. **Satz** (Notwendiges Extremwertkriterium; Fermat, um 1638)**.** *Sei $I \subset \mathbb{R}$ ein Intervall, $x_0 \in \mathring{I}$ ein innerer Punkt und $f : I \longrightarrow \mathbb{R}$ differenzierbar in x_0. Hat f in x_0 ein lokales Extremum, so gilt $f'(x_0) = 0$.*

5.3.4. **Bemerkung.** (i) Geometrische Interpretation: Die Tangente an den Graphen in einem inneren Extremum ist parallel zur x-Achse.


zwei lokale Extrema

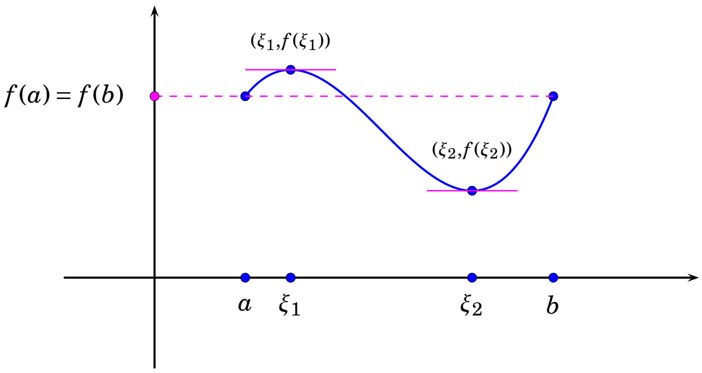

Sattelpunkt

(ii) Wenn x_0 kein innerer Punkt ist, so ist die Aussage i.A. falsch; Beispiel: $f : [0,1] \longrightarrow \mathbb{R}$, $f(x) = x$, $x_0 = 0$; dann ist x_0 ein Minimum, f differenzierbar in x_0, aber $f'(x_0) = 1$.

(iii) Die Umkehrung von Satz 5.3.3 ist falsch; Beispiel: $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = x^3$. Dann ist $x_0 = 0$ ein kritischer Punkt ($f'(0) = 3 \cdot 0^2 = 0$), aber kein lokales Extremum, sondern einen Sattelpunkt.

5.3.5. **Beispiel.** Als Anwendung des Satzes 5.3.3 betrachten wir das folgende Beispiel. Sei $f : \mathbb{R} \to \mathbb{R}$, $f(x) = (x^2 - 1)^2$. Wir suchen die Extremstellen von f. Dafür suchen wir die kritischen Punkte. Sie sind die Nullstellen von $f'(x)$, d. h. die Lösungen der Gleichung $4x(x^2 - 1) = 0$, also die Stellen $x_1 = -1$, $x_2 = 0$, $x_3 = 1$. Wegen $f(1) = f(-1) = 0$ und $f(x) \geq 0$ für alle $x \in \mathbb{R}$ sind x_1 und x_3 globale Minimumstellen von f. Die Einschränkung von f auf das Intervall $[-1,1]$ besitzt nach dem Satz von Weierstraß ein positives Maximum an einer Stelle $\xi \in (1,1)$, und nach Satz 5.3.3 gilt $f'(\xi) = 0$. Hieraus ergibt sich, dass $x = 0$ eine lokale Maximumstelle von f ist.

5.3.6. **Satz** (Rolle, um 1691)**.** *Sei $f : [a,b] \longrightarrow \mathbb{R}$ mit den Eigenschaften: (1) f stetig auf $[a,b]$, (2) f differenzierbar auf (a,b), (3) $f(a) = f(b)$. Dann gibt es $\xi \in (a,b)$ mit $f'(\xi) = 0$.*



Geometrische Interpretation: Es gibt innere Punkte $\xi \in (a,b)$, so dass die Tangente in $(\xi, f(\xi))$ parallel zur x-Achse ist.

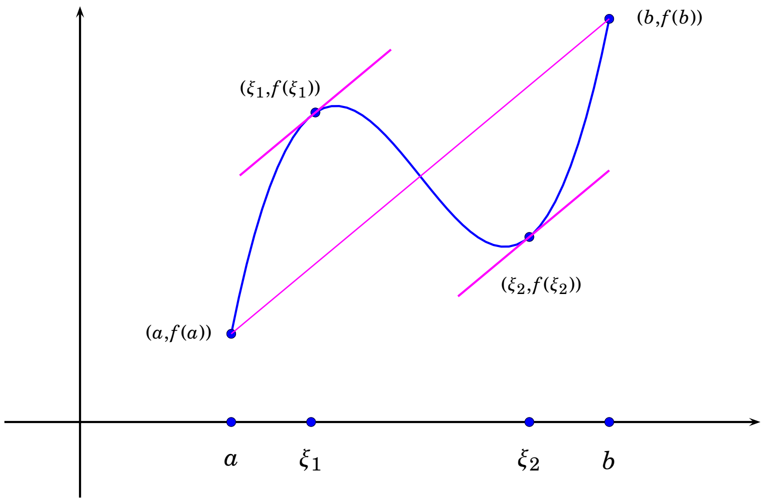
09.01.2012

5.3.7. **Satz** (Mittelwertsatz; Lagrange, um 1797). *Sei $f : [a,b] \longrightarrow \mathbb{R}$ stetig und auf (a,b) differenzierbar. Dann gibt es $\xi \in (a,b)$ mit $f(b) - f(a) = f'(\xi)(b-a)$.*

Geometrische Interpretation:

Die Gleichung der Sekante durch $(a, f(a))$, $(b, f(b))$ ist $y - f(a) = \frac{f(b)-f(a)}{b-a} \cdot (x-a)$, und die Gleichung der Tangente an Graph(f) in $(\xi, f(\xi))$ ist $y - f(\xi) = f'(\xi)(x-\xi)$. Die beiden Geraden haben dieselbe Steigung $f'(\xi) = \frac{f(b)-f(a)}{b-a}$, sind also parallel.

Also: Jede Sehnensteigung ist Tangentensteigung in einem geeigneten Zwischenpunkt.



Die Tangenten in den Punkten $(\xi_1, f(\xi_1))$, $(\xi_2, f(\xi_2))$ sind parallel zur Sekante durch $(a, f(a))$ und $(b, f(b))$

Wieso betrachtet man im Mittelwertsatz Funktionen, die nur auf (a,b) differenzierbar sind? Der Grund ist, dass wir Funktionen wie $f : [0,1] \to \mathbb{R}$, $f(x) = \sqrt{x}$ nicht ausschließen möchten (f ist in 0 nicht differenzierbar).

Für den Beweis der Regel von l'Hospital benötigen wir die folgende Verallgemeinerung des Mittelwertsatzes.

5.3.8. **Satz** (Verallgemeinerter Mittelwertsatz, Cauchy). *Seien $f, g : [a,b] \longrightarrow \mathbb{R}$ stetig und auf (a,b) differenzierbar. Dann gibt es $\xi \in (a,b)$ mit $(f(b) - f(a))g'(\xi) = (g(b) - g(a))f'(\xi)$.*

Wird zusätzlich $g'(x) \neq 0$ für alle $x \in (a,b)$ vorausgesetzt, so folgt $g(a) \neq g(b)$ und wir können den verallgemeinerter Mittelwertsatz in Bruchform schreiben,

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

5.4. **Anwendungen des Mittelwertsatzes.** Der Mittelwertsatz hat viele Anwendungen: Er ist die Brücke zwischen Ableitungsinformationen und Informationen über die Funktion selbst. Die wichtigsten sind: der Eindeutigkeitssatz der Differentialrechnung (Folgerung 5.4.1), das Monotoniekriterium (Folgerung 5.4.3), ein hinreichendes Extremwertkriterium (Folgerung 5.4.4), der Schrankensatz (Satz 5.4.6), ein Differenzierbarkeitssatz (Satz 5.4.7), die Regel von l'Hospital (Satz 5.4.8) und schließlich die Taylorformel (Sätze 6.5.4-6.5.6). Mit Hilfe dieser Mittel erläutern wir in Beispiel 5.4.9 das Studium des Funktionsverlaufs (Kurvendiskussion). Dies führt zu einem detaillierten Studium der trigonometrischen Funktionen. Zu den Anwendungen zählen auch die Konvexitätskriterien und die wichtigen Ungleichungen von Jensen, Hölder und Minkowski.

5.4.1. **Folgerung** (Eindeutigkeitssatz der Differentialrechnung).

(i) *Sei $I \subset \mathbb{R}$ ein Intervall, $f : I \longrightarrow \mathbb{R}$ stetig und auf $\mathring{I}$ differenzierbar. Falls $f'(x) = 0$ für alle $x \in \mathring{I}$, dann ist f konstant.*

(ii) *Seien $f, g : I \longrightarrow \mathbb{R}$ stetig und auf $\mathring{I}$ differenzierbar, und es gelte $f' = g'$ auf $\mathring{I}$. Dann gibt es $c \in \mathbb{R}$ mit $f(x) = g(x) + c$ für alle $x \in I$.*

5.4.2. **Bemerkung.** Es ist hierbei wesentlich, dass I ein Intervall ist: Man betrachte z.B. die Funktion $f : (-1,0) \cup (0,1) \longrightarrow \mathbb{R}$ mit $f(x) = 0$ für $x \in (-1,0)$ und $f(x) = 1$ für $x \in (0,1)$. Dann ist $f'(x) = 0$ für alle $x \in (-1,0) \cup (0,1)$, aber f ist nicht konstant.

5.4.3. **Folgerung** (Monotoniekriterium). *Sei I ein Intervall, $f : I \longrightarrow \mathbb{R}$ stetig und auf $\mathring{I}$ differenzierbar. Dann gilt:*

(1) $f'(x) > 0$ *für alle* $x \in \mathring{I} \Longrightarrow f$ *streng monoton wachsend,*
(2) $f'(x) \geqslant 0$ *für alle* $x \in \mathring{I} \Longleftrightarrow f$ *monoton wachsend,*
(3) $f'(x) < 0$ *für alle* $x \in \mathring{I} \Longrightarrow f$ *streng monoton fallend,*
(4) $f'(x) \leqslant 0$ *für alle* $x \in \mathring{I} \Longleftrightarrow f$ *monoton fallend.*

Beispiele. (1) Die Exponentialfunktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^x$ ist streng monoton wachsend, da $f'(x) = e^x > 0$.
(2) Das Logarithmus $f : (0,\infty) \to \mathbb{R}$, $f(x) = \log x$ ist auf $(0,\infty)$ streng monoton wachsend, da $f'(x) = \frac{1}{x} > 0$ für $x > 0$ ist.
(3) Die Funktion $f(x) = \frac{1}{x}$ ist auf jedem der Intervalle $(0,\infty)$, $(-\infty,0)$ streng monoton fallend, da da $f'(x) = -\frac{1}{x^2} < 0$ für $x \neq 0$ ist. Beachte aber, dass f nicht auf ganz $\mathbb{R} \setminus \{0\}$ monoton fallend ist.
(4) Vorsicht! Es gibt differenzierbare Funktionen, bei deren Ableitung sich das Vorzeichen in einer Umgebung eines Punktes x_0 unendlich oft ändert. Insbesondere ist die Funktion in keiner Umgebung des Punktes x_0 monoton. Zum Beispiel:

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = \begin{cases} x + x^2 \sin(\frac{1}{x}), & x \neq 0 \\ 0, & x = 0 \end{cases} \tag{5.5}$$

Für jedes $\varepsilon > 0$ gibt es $x_n = 2/(4n+1)\pi \in (0,\varepsilon)$ mit $f'(x_n) > 0$ und $y_m = 1/2m\pi \in (0,\varepsilon)$ mit $f'(y_m) < 0$. Da f' kein konstantes Vorzeichen in $(0,\varepsilon)$ hat, kann f dort nach Satz 5.4.3 (2), (4) nicht monoton sein. Obendrein gilt $f'(0) > 0$! (Allerdings ist f' in $x = 0$ nicht stetig.)

Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine stetige Funktion, die auf $\mathring{I}$ differenzierbar ist. Wir möchten die lokalen Extrema von f bestimmen. Die lokalen Extrema können (1) Endpunkte oder (2) innere Punkte von I sein.
(1) Lokale Extrema in Endpunkten. Mit Hilfe des Monotoniekriteriums ist es einfach zu bestimmen, ob in den Endpunkten lokale Extrema vorliegen. Sei $f : [a,b] \to \mathbb{R}$ oder $f : [a,b) \to \mathbb{R}$ stetig und auf (a,b) differenzierbar. Dann gilt:

- $f' \geqslant 0$ auf $(a, a+\varepsilon)$ für ein $\varepsilon < b-a \rightsquigarrow f$ hat in a ein lokales Minimum
- $f' \leqslant 0$ auf $(a, a+\varepsilon)$ für ein $\varepsilon < b-a \rightsquigarrow f$ hat in a ein lokales Maximum

Analog untersucht man, ob in rechten Endpunkt ein Extremum vorliegt.
(2) Innere lokale Extrema. Wir haben gesehen (Fermat-Kriterium 5.3.3), dass ein inneres lokales Extremum ein kritischer Punkt ist, d.h. $f'(x) = 0$ erfüllen muss; aber nicht jeder kritische Punkt ist ein lokales Extremum. Wir sagen, dass ein kritischer Punkt „extremumverdächtig" ist. Um zu entscheiden, ob der kritische Punkt tatsächlich ein lokales Extremum ist, müssen wir weitere Untersuchungen anstellen. Mit Hilfe des Monotoniekriteriums 5.4.3 erhalten wir ein hinreichendes Extremwertkriterium:

5.4.4. **Folgerung** (ersten hinreichendes Extremwertkriterium). *Sei $f:(a,b)\longrightarrow\mathbb{R}$ differenzierbar, $x_0\in(a,b)$. Falls $f'(x)\leqslant 0$ für alle $x\in(a,x_0)$ und $f'(x)\geqslant 0$ für alle $x\in(x_0,b)$, so hat f in x_0 ein Minimum. Falls $f'(x)\geqslant 0$ für alle $x\in(a,x_0)$ und $f'(x)\leqslant 0$ für alle $x\in(x_0,b)$, so hat f in x_0 ein Maximum.*

Unsere erste Regel zur Bestimmung der inneren lokalen Extrema lautet:

$\boxed{1}$ Bestimme die kritschen Punkte $x\in\mathring{I}$.

$\boxed{2}$ Für jeden kritischen Punkt x_0 bestimme das Vorzeichen von f' in einer linksseitigen Umgebung $(x_0-\varepsilon,x_0)$ und einer rechtsseitigen Umgebung $(x_0,x_0+\varepsilon)$ von x_0 (*falls* diese so existieren, dass f' dort konstantes Vorzeichen ungleich Null hat!).
Ändert sich das Vorzeichen von f' von Minus nach Plus $\rightsquigarrow$ x_0 lokales Minimum;
ändert sich das Vorzeichen von f' von Plus nach Minus $\rightsquigarrow$ x_0 lokales Maximum.

$\boxed{3}$ Ändert sich das Vorzeichen von f' nicht, so ist x_0 ein Sattelpunkt.

Beispiele. (1) Betrachten wir die Funktion $f:\mathbb{R}\to\mathbb{R}$, $f(x)=(x^2-1)^2$ aus Beispiel 5.3.5. Sie ist differenzierbar mit $f'(x)=4x(x^2-1)$. Die kritischen Punkte sind $x_1=-1$, $x_2=0$, $x_3=1$. Es gilt $f'<0$ auf die Intervallen $(-\infty,-1)$, $(0,1)$ und $f'>0$ auf die Intervallen $(-1,0)$, $(1,\infty)$. Somit sind -1 und 1 Minimumstellen und 0 Maximumstelle.
(2) Seien $a_1,\dots,a_n\in\mathbb{R}$; gesucht wird eine Zahl $x_0\in\mathbb{R}$, so dass $(x_0-a_1)^2+\ldots+(x_0-a_n)^2$ möglichst klein ist. Dazu betrachten wir $f:\mathbb{R}\longrightarrow\mathbb{R}$, $f(x)=(x-a_1)^2+\ldots+(x-a_n)^2$; die Funktion f ist differenzierbar, und $f'(x)=2nx-2(a_1+\ldots+a_n)$. Falls f in x_0 ein Minimum hat, so ist x_0 ein kritischer Punkt, d.h. $f'(x_0)=0$, folglich $x_0=\frac{1}{n}(a_1+\ldots+a_n)$. Also ist x_0 der einzige extremumverdächtige Punkt. Die Funktion f hat dort tatsächlich ein Minimum, denn $f'(x)<0$ für $x<x_0$ und $f'(x)>0$ für $x>x_0$.
(3) Emissionsmaximum eines strahlenden Körpers (beim Planckschen Strahlungsgesetz) und Wiensches Verschiebungsgesetz (siehe [13] §9.4).
(4) Für $f:\mathbb{R}\to\mathbb{R}$, $f(x)=x^3$ gilt $f'(0)=0$ und $f'(x)>0$ für $x\in(-\infty,0)\cup(0,\infty)$. Also ist $x=0$ ein Sattelpunkt.
(5) Vorsicht! Unter den Bedingungen von Satz 5.4.4 ist f auf einer geeigneten linksseitigen und einer geeigneten rechtsseitigen Umgebung von x_0 jeweils monoton. Es gibt aber differenzierbare Funktionen mit einem lokalen Extremum in x_0, die in keiner linksseitigen oder rechtsseitigen Umgebung von x_0 monoton sind. Zum Beispiel:

$$f:\mathbb{R}\to\mathbb{R},\quad f(x)=\begin{cases}2-x^2(2+\sin(\frac{1}{x})) & ,x\neq 0\\ 2 & ,x=0\end{cases}\tag{5.6}$$

ist differenzierbar, und

$$f':\mathbb{R}\to\mathbb{R},\quad f(x)=\begin{cases}-4x-2x\sin(\frac{1}{x})+\cos(\frac{1}{x}) & ,x\neq 0\\ 0 & ,x=0\end{cases}$$

Der Punkt $x_0=0$ ist ein lokales Maximum, da $f(0)=2\geqslant 2-x^2(2+\sin(\frac{1}{x}))$. Für jedes $\varepsilon>0$ gibt es $x_n=1/2n\pi\in(0,\varepsilon)$ und $y_m=2/(2m+1)\pi\in(0,\varepsilon)$. Dann gilt $f'(x_n)>0$ und $f'(y_m)<0$. Da f' kein konstantes Vorzeichen in $(0,\varepsilon)$ hat, kann sie dort nach dem Satz 5.4.3(2), (4) nicht monoton sein. Der Satz 5.4.4 ist also nicht immer anwendbar.

Ein weiteres hinreichendes Extremwertkriterium benutzt die zweite Ableitung. Wir sagen, dass eine Funktion $f:D\to\mathbb{C}$ ***zweimal differenzierbar*** in einem Punkt $x_0\in D$ ist, falls es $\delta>0$ gibt, so dass f differenzierbar auf $D\cap(x_0-\delta,x_0+\delta)$ ist und die Ableitung $f':D\cap(x_0-\delta,x_0+\delta)\to\mathbb{C}$ differenzierbar in x_0 ist. Wir bezeichnen $f''(x_0):=(f')'(x_0)$ die Ableitung von f' in x_0; sie wird zweite Ableitung von f in x_0 genannt.

5.4.5. **Folgerung** (zweiten hinreichendes Extremwertkriterium). *Sei $f:(a,b)\longrightarrow\mathbb{R}$ differenzierbar und zweimal differenzierbar in $x_0\in(a,b)$. Gilt $f'(x_0)=0$ und $f''(x_0)>0$, so ist x_0 eine lokale Minimumstelle. Gilt $f'(x_0)=0$ und $f''(x_0)>0$, so ist x_0 eine lokale Maximumstelle.*

Unsere zweite Regel zur Bestimmung der inneren lokalen Extrema lautet:

$\boxed{1}$ Bestimme die kritschen Punkte $x\in\mathring{I}$.

$\boxed{2}$ Für jeden kritischen Punkt x_0 berechne $f''(x_0)$ (*falls* f zweimal differenzierbar in x_0 ist).
Gilt $f''(x_0)>0$ $\rightsquigarrow$ x_0 lokales Minimum;
Gilt $f''(x_0)>0$ $\rightsquigarrow$ x_0 lokales Maximum.

$\boxed{3}$ Gilt $f''(x_0) = 0$, dann liefert das Kriterium 5.4.5 keine Auskunft und es kann passieren, dass x_0 eine Extremumstelle oder einen Sattelpunkt ist. Man kann höhere Ableitungen zur Untersuchung heranziehen, siehe Satz 5.4.5.

Beispiele. (1) Sei $f : \mathbb{R} \to \mathbb{R}$, $f(x) = (x^2-1)^2$ aus Beispiel 5.3.5. Sie ist zweimal differenzierbar mit $f'(x) = 4x(x^2-1)$, $f''(x) = 12x^2 - 4$. Die kritischen Punkte sind $x_1 = -1$, $x_2 = 0$, $x_3 = 1$. Es gilt $f''(-1) = f''(1) = 8 > 0$ und $f''(0) = -4 < 0$. Somit sind -1 und 1 Minimumstellen und 0 Maximumstelle.
(2) Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$ gilt $f'(0) = f''(0) = 0$ und $x = 0$ is ein Sattelpunkt. Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^4$ gilt $f'(0) = f''(0) = 0$ und $x = 0$ ist eine Minimumstelle.

Eine Funktion $f : I \to \mathbb{C}$ heißt ***stetig differenzierbar*** , falls f differenzierbar und $f' : I \to \mathbb{C}$ stetig ist. Vorsicht! Es gibt differenzierbare Funktionen, die nicht stetig differenzierbar sind, wie etwa (5.5) oder (5.6). In diesen Beispielen ist f' in $x = 0$ unstetig, da $\lim_{x\to 0} f'(x)$ nicht existiert.

5.4.6. **Satz** (Schrankensatz)**.** *Sei $f : [a,b] \longrightarrow \mathbb{R}$ stetig differenzierbar. Dann ist f Lipschitz-stetig mit Lipschitz-Konstante $L = \sup_{x\in[a,b]} |f'(x)| = \|f'\|_{[a,b]}$.*

5.4.7. **Satz** (Differenzierbarkeitssatz)**.** *Sei $I \subset \mathbb{R}$ ein Intervall und $x_0 \in I$. Sei $f : I \longrightarrow \mathbb{R}$ stetig und auf $I \setminus \{x_0\}$ differenzierbar. Es existiere $\lim_{x\to x_0} f'(x) =: \alpha \in \mathbb{R}$. Dann ist f in x_0 differenzierbar, und es gilt $f'(x_0) = \alpha$.*

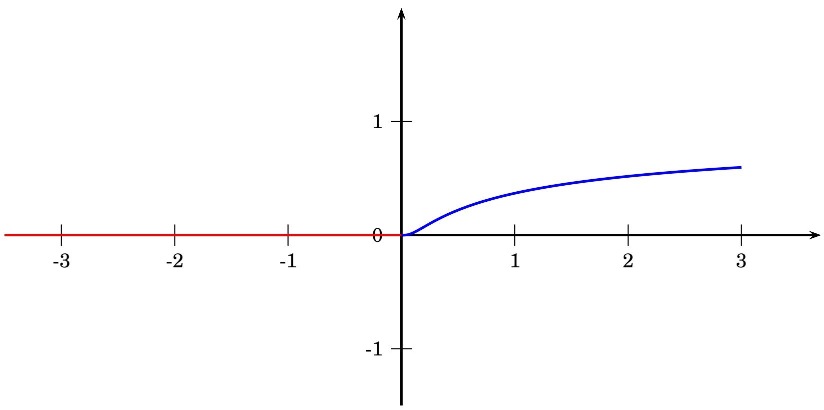
Als Beispiel betrachte die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$,

$$f(x) = \begin{cases} e^{-1/x}, & x > 0 \\ 0, & x \leqslant 0 \end{cases} \tag{5.7}$$

Die Funktion f ist stetig auf $\mathbb{R}$ und differenzierbar auf $\mathbb{R} \setminus \{0\}$, mit

$$f'(x) = \begin{cases} \frac{1}{x^2} e^{-1/x}, & x > 0 \\ 0, & x < 0 \end{cases}$$

Wegen $\lim_{x \nearrow 0} f'(x) = 0 = \lim_{x \searrow 0} f'(x)$ ist f auch in 0 differenzierbar, und es gilt $f'(0) = 0$.



5.4.8. **Satz** (Regel von l'Hospital)**.**
Sei $I \subset \mathbb{R}$, und sei $x_0 \in \overline{\mathbb{R}}$ ein Häufungspunkt von I. Seien $f, g : I \setminus \{x_0\} \longrightarrow \mathbb{R}$ differenzierbar. Es gelte :

(1) (a) $\lim\limits_{x\to x_0} f(x) = \lim\limits_{x\to x_0} g(x) = 0$
oder
(b) $\lim\limits_{x\to x_0} g(x) = \pm\infty$,

(2) $g'(x) \neq 0$ *für alle* $x \in I \setminus \{x_0\}$,

(3) $\lim\limits_{x\to x_0} \dfrac{f'(x)}{g'(x)} = \ell \in \overline{\mathbb{R}}$.

Dann gilt $g(x) \neq 0$ für $x \in I \setminus \{x_0\}$ nahe x_0, und $\lim\limits_{x\to x_0} \dfrac{f(x)}{g(x)} = \ell$.

Beispiele. (1) Sei $s>0$. Dann gilt:

$$\lim_{x\searrow 0} x^s\log x = \lim_{x\searrow 0}\frac{\log x}{x^{-s}} = \lim_{x\searrow 0}\frac{x^{-1}}{(-s)x^{-s-1}} \quad (\text{l'Hospital, Typ } \tfrac{\infty}{\infty},\ g'(x)\neq 0 \text{ für } x>0)$$
$$= \lim_{x\searrow 0}\frac{x^s}{(-s)} = 0\,.$$

(2)

$$\lim_{x\to 0}\frac{1-\cos\frac{x}{3}}{1-\cos x} \overset{(i)}{=} \lim_{x\to 0}\frac{\frac{1}{3}\sin\frac{x}{3}}{\sin x} \quad (\text{l'Hospital, Typ } \tfrac{0}{0},\ g'(x)\neq 0 \text{ für } x\in(-\pi,\pi)\setminus\{0\})$$
$$\overset{(ii)}{=} \lim_{x\to 0}\frac{\frac{1}{9}\cos\frac{x}{3}}{\cos x} \quad (\text{l'Hospital, Typ } \tfrac{0}{0},\ g'(x)\neq 0 \text{ für } x\in(-\tfrac{\pi}{2},\tfrac{\pi}{2}))$$
$$= \frac{\frac{1}{9}}{1} = \frac{1}{9}\,.$$

Wir haben hier die Regeln wiederholt angewendet, was natürlich nur dann erlaubt ist, wenn jeweils die Voraussetzungen des Satzes von l'Hospital erfüllt sind. Wir haben dies zunächst nur teilweise überprüft. Was in der Begründung der Zwischenschritte (i) und (ii) fehlt, ist die Existenz des Grenzwertes des Ableitungsquotienten auf der jeweiligen rechten Seite. Es muss ja immer erst der Ableitungsquotient konvergieren, bevor man sagen kann, dass der Funktionenquotient konvergiert. Daher ist in obiger Gleichungskette alles erst vom Ende her gerechtfertigt. Da der letzte Grenzwert existiert, tut es auch der vorletzte und dann schließlich der davorstehende.
Zum Vergleich die Berechnung des Grenzwertes in (2) mit Hilfe von Potenzreihen (Definition von cos):

$$\cos y = 1-\frac{y^2}{2!}+\sum_{k=2}^{\infty}\frac{(-1)^k y^{2k}}{(2k)!} = 1-\frac{y^2}{2}+o(y^2) \text{ für } y\to 0\,.$$

Also für $x\to 0$:

$$\frac{1-\cos\frac{x}{3}}{1-\cos x} = \frac{\frac{x^2}{2\cdot 9}+o(x^2)}{\frac{x^2}{2}+o(x^2)} = \frac{\frac{1}{18}+\frac{o(x^2)}{x^2}}{\frac{1}{2}+\frac{o(x^2)}{x^2}} \longrightarrow \frac{\frac{1}{18}}{\frac{1}{2}} = \frac{1}{9}\,.$$

(3) Durch n-malige Anwendung der Regel von l'Hospital erhält man erneut (4.6):

$$\lim_{x\to\infty}\frac{e^{cx}}{x^n} = \lim_{x\to\infty}\frac{ce^{cx}}{nx^{n-1}} = \ldots = \lim_{x\to\infty}\frac{c^n e^{cx}}{n!} = \infty, \quad c>0.$$

(4) Man betrachte die folgende Anwendung der Regel von l'Hospital:

$$\lim_{x\to 0}\frac{\sin x}{x} = \lim_{x\to 0}\frac{\cos x}{1} = \cos 0 = 1. \tag{5.8}$$

Dies ist *kein* Beweis von $\lim_{x\to 0}\frac{\sin x}{x}=1$, sondern nur eine Probe; in 5.8 haben wir nämlich schon benutzt, dass $\sin' 0 = 1$ gilt, und dies ist äquivalent zu $\lim_{x\to 0}\frac{\sin x}{x}=1$.
(5) Die Regel von l'Hospital bezieht sich auf unbestimmte Ausdrücke der Form $\frac{0}{0}$ und $\frac{\infty}{\infty}$. Andere wichtige unbestimmte Ausdrücke sind vom Typ $0\cdot\infty$ und $\infty-\infty$, die auf die Quotientenform zurückgeführt werden können: Bei Typ $0\cdot\infty$ schreibe z.B.: $f\cdot g = \frac{f}{1/g}$, bei Typ $\infty-\infty$ schreibe z.B.: $f-g=\frac{1/g-1/f}{1/(fg)}$.
Exponentialausdrücke der Form 1^∞, 0^0, ∞^0 behandelt man mittels der Definition $f^g := e^{g\log f}$.

5.4.9. **Beispiel** (Kurvendiskussion)**.** Sei $f: D\to\mathbb{R}$, wobei D eine Vereinigung von Intervallen sei. Die Kurvendiskussion der Kurve $y=f(x)$ besteht aus den folgenden Schritten:

$\boxed{1}$ Bestimme die Schnittpunkte mit den Achsen und das Vorzeichen der Funktion, d.h. die Mengen $\{x: f(x)>0\}$ und $\{x: f(x)<0\}$.

$\boxed{2}$ Studiere die Grenzwerte in Randpunkten (inklusive eventuell $\pm\infty$), Asymptoten, Stetigkeit der Funktion.

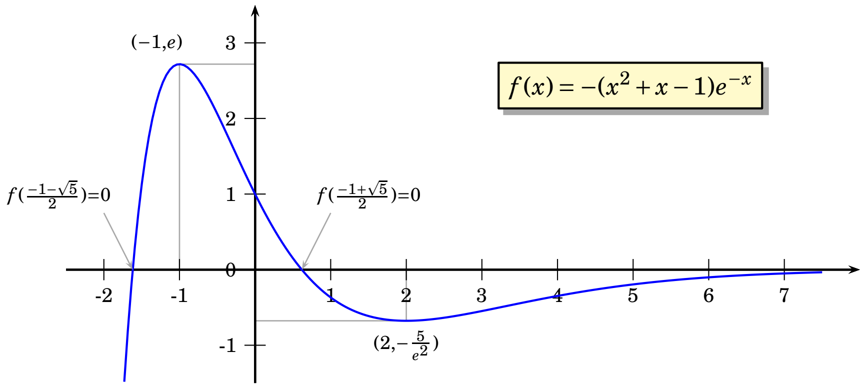
$\boxed{3}$ Studiere die Differenzierbarkeit, berechne f' und bestimme die Mengen $\{x: f'(x)=0\}$, $\{x: f'(x)>0\}$, $\{x: f'(x)<0\}$; bestimme die lokalen und globalen Extrema.

$\boxed{4}$ Erstelle eine Tabelle mit den wichtigsten Werten und mit den Monotoniebereichen der Funktion f. Schließlich skizziere den Graphen.

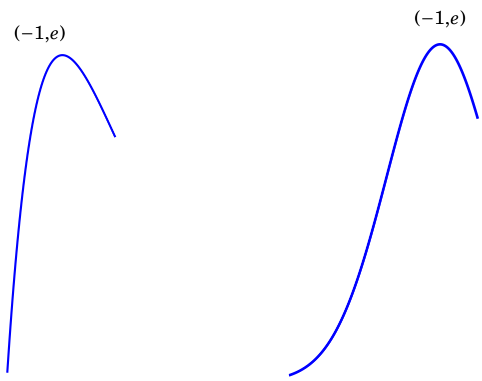
Wir erläutern das Vorgehen am Beispiel der Funktion $f:\mathbb{R}\longrightarrow\mathbb{R}$, $f(x)=-(x^2+x-1)e^{-x}$. Die Tabelle lautet:

x	$-\infty$		$\frac{-1-\sqrt{5}}{2}$		-1		0		$\frac{-1+\sqrt{5}}{2}$		2		∞
$f(x)$	$-\infty$	$\nearrow$	0	$\nearrow$	e	$\searrow$	1	$\searrow$	0	$\searrow$	$-\frac{5}{e^2}$	$\nearrow$	0
$f'(x)$		$+$		$+$	0	$-$		$-$		$-$	0	$+$	

Wir entnehmen der Tabelle, dass $f(x) \leqslant f(-1) = e$ für $x \in (-\infty, 2]$ und $f(x) \geqslant f(2) = -5/e^2$ für $x \in [-1, \infty)$. Die Funktion hat in $x = -1$ ein lokales Maximum und in $x = 2$ ein lokales Minimum. Da $\lim_{x\to-\infty} f(x) = -\infty$ gilt, besitzt f kein globales Minimum (Beweis?). Wegen $f(x) < 0 < e = f(-1)$ für $x \in [2, \infty)$ hat f in $x = -1$ ein globales Maximum.



Bei der vorigen Kurvendiskussion fehlt noch ein Argument. Wir müssen begründen, wieso der Graph wie in der linken Abbildung und nicht wie in der rechten Abbildung aussieht:



Dazu brauchen wir die zweite Ableitung (siehe Abschnitt 6.5)!

5.5. **Trigonometrische Funktionen.** Die trigonometrischen Funktionen *sin* und *cos* wurden in 3.4.8 eingeführt. Wir studieren sie nun mit Hilfe des Abschnitts 5.4.

5.5.1. **Lemma.** *Für* $x \in (0, 2]$ *gilt*

$$\sin x > x - \frac{x^3}{3!}, \quad \cos x < 1 - \frac{x^2}{2!} + \frac{x^4}{4!}.$$

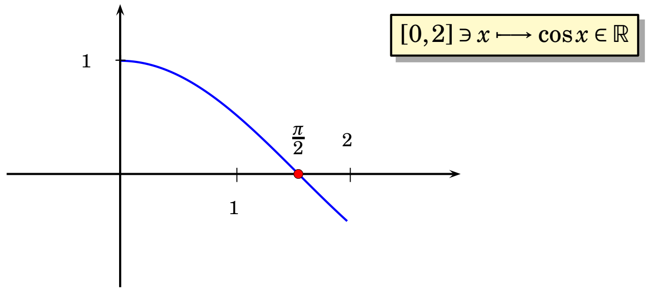
12.01.2012

5.5.2. **Lemma.** $\cos|_{[0,2]} : [0,2] \longrightarrow \mathbb{R}$ *hat genau eine Nullstelle* x_0 *; diese liegt in* $(1, 2)$.

5.5.3. **Definition.** Setze $\boxed{\boldsymbol{\pi} := 2x_0}$ (d.h. $\boxed{x_0 = \pi/2}$). Die Zahl π heißt ***Kreiszahl***.

5.5.4. **Bemerkung.**

(1) Wegen $\cos\frac{\pi}{2} = 0$ und $\cos^2 x + \sin^2 x = 1$ gilt $\sin\frac{\pi}{2} \in \{\pm 1\}$; wegen $\sin x > 0$ für $x \in (0, 2]$ folgt $\sin\frac{\pi}{2} = 1$.



(2) Durch die Additionssätze erhalten wir wegen (1):

$$\cos(z+\frac{\pi}{2}) = -\sin z\,, \quad \sin(z+\frac{\pi}{2}) = \cos z\,.$$

Daraus folgt:

$$\cos(z+\pi) = -\cos z, \qquad \sin(z+\pi) = -\sin z\,,$$
$$\cos(z+2\pi) = \cos z, \qquad \sin(z+2\pi) = \sin z\,,$$

und wir zeigen leicht, dass die Funktionen cos und sin beide 2π-periodisch sind.
Es gilt insbesondere:

x	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
$\cos x$	1	0	-1	0	1
$\sin x$	0	1	0	-1	0

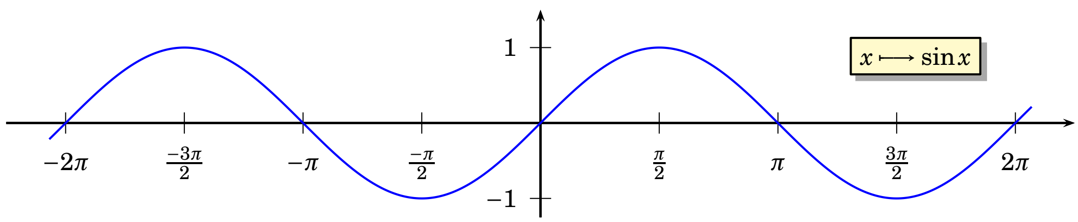
(3) $\cos x > 0$ für $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ und $\cos x < 0$ für $x \in (\frac{\pi}{2}, \frac{3\pi}{2})$.
(4) $\sin x > 0$ für $x \in (0,\pi)$ und $\sin x < 0$ für $x \in (\pi, 2\pi)$.
(5) $\sin x = 0 \Leftrightarrow x \in \{k\pi : k \in \mathbb{Z}\}$; $\cos x = 0 \Leftrightarrow x \in \{\frac{\pi}{2} + k\pi : k \in \mathbb{Z}\}$.
(6) $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \longrightarrow [-1,1]$ ist streng monoton wachsend, bijektiv und stetig; $\cos : [0,\pi] \longrightarrow [-1,1]$ ist streng monoton fallend, bijektiv und stetig. Deshalb sind diese Funktionen invertierbar.
(7) $\arcsin := \sin^{-1} : [-1,1] \longrightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ und $\arccos := \cos^{-1} : [-1,1] \longrightarrow [0,\pi]$ sind stetig und auf $(-1,1)$ differenzierbar. Es gilt $\arcsin' x = \frac{1}{\sqrt{1-x^2}}$ und $\arccos' x = -\frac{1}{\sqrt{1-x^2}}$.
(8) Wir definieren die Tangensfunktion $\tan : \mathbb{R} \setminus (\frac{\pi}{2} + \pi\mathbb{Z}) \longrightarrow \mathbb{R}$ durch $\tan x = \frac{\sin x}{\cos x}$ und die Cotangensfunktion $\cot : \mathbb{R} \setminus \pi\mathbb{Z} \longrightarrow \mathbb{R}$ durch $\cot = \frac{\cos x}{\sin x}$.
Die Funktion $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \longrightarrow \mathbb{R}$ ist bijektiv, streng monoton wachsend und differenzierbar. Die Umkehrfunktion $\arctan : \mathbb{R} \longrightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$ ist differenzierbar, und es gilt $\arctan' x = \frac{1}{\tan'(\arctan x)} = \frac{1}{1+\tan^2(\arctan x)} = \frac{1}{1+x^2}$. Also:

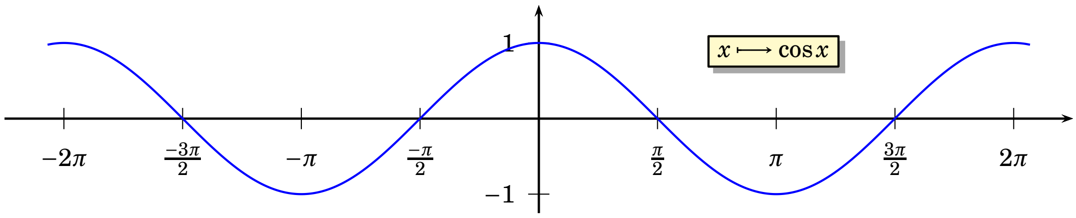
$$\boxed{\arctan' x = \frac{1}{1+x^2}}$$

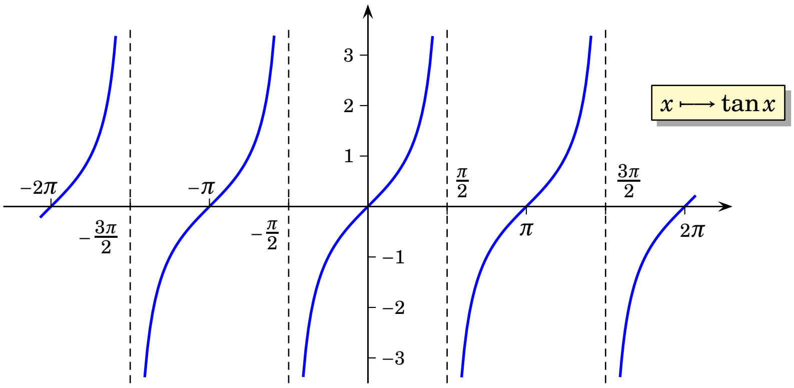
Die Funktion $\cot : (0,\pi) \longrightarrow \mathbb{R}$ ist bijektiv, streng monoton fallend und differenzierbar. Die Umkehrfunktion $\operatorname{arccot} : \mathbb{R} \longrightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$ ist differenzierbar, und es gilt $\operatorname{arccot}' x = \frac{1}{\cot'(\operatorname{arccot} x)} = -\frac{1}{1+\cot^2(\operatorname{arccot} x)} = -\frac{1}{1+x^2}$. Also:

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

Wir fassen die Eigenschaften der trigonometrischen Funktionen sin, cos und tan mit Hilfe ihrer Graphen zusammen:







5.6. **Notizen.** Wir sammeln in diesen Abschnitt Kommentare über einseitigen Ableitungen, umformulierung der Definition der Differenzierbarkeit, Landau-Symbole.

Einseitige Ableitungen. Sei $f : I \longrightarrow \mathbb{C}$ und $x_0 \in I$. Falls x_0 rechtsseitiger (bzw. linksseitiger) HP von I ist, sagen wir, dass f ***rechtsseitig*** (bzw. ***linksseitig***) ***differenzierbar in*** x_0 ist, falls $\lim_{x \searrow x_0} \frac{f(x)-f(x_0)}{x-x_0}$ (bzw. $\lim_{x \nearrow x_0} \frac{f(x)-f(x_0)}{x-x_0}$) in $\mathbb{C}$ existiert. In diesem Fall schreiben wir

$$f_r'(x_0) := \lim_{x \searrow x_0} \frac{f(x)-f(x_0)}{x-x_0} \quad \text{bzw.}$$
$$f_l'(x_0) := \lim_{x \nearrow x_0} \frac{f(x)-f(x_0)}{x-x_0}$$

und nennen dies die rechtsseitige (bzw. linksseitige) Ableitung von f in x_0. Es gilt nach Satz 4.5.17(i): f ist differenzierbar in x_0 genau dann, wenn f linksseitig und rechtsseitig differenzierbar in x_0 ist und $f_l'(x_0) = f_r'(x_0)$ gilt. In diesem Fall ist $f'(x_0) = f_l'(x_0) = f_r'(x_0)$.

Für $f(x) = |x|$ ist $f_l'(0) = -1$, $f_r'(0) = 1$. Für die Heaviside-Funktion (4.8) gilt $H_r'(0) = 0$, während die linksseitige Ableitung nicht existiert (bei entsprechender Erweiterung der Definition wäre sie ∞).

Aus physikalischen Gründen versucht man dennoch die Ableitung der Heaviside-Funktion zu betrachten. Die Physiker definieren sie als diejenige Funktion δ mit $\delta(x) = 0$ für $x \neq 0$, $\delta(0) = \infty$, deren Integral auf $\mathbb{R}$ eins beträgt. Dies ist die berühmte Dirac-δ-Funktion. Obwohl die obige naive Definition der üblichen Vorstellung von einer Funktion widerspricht, kann man sie im Rahmen der Distributionen-Theorie streng definieren. Sie spielt eine bedeutende Rolle auf dem Gebiet der partiellen Differentialgleichungen.

Umformulierung der Definition der Differenzierbarkeit, Landau-Symbole.

5.6.1. **Satz.** *Sei $f : I \longrightarrow \mathbb{C}$, $x_0 \in I$, I ein Intervall. Die folgenden Aussagen sind äquivalent:*

(i) *f ist differenzierbar in x_0.*
(ii) *Es gibt $\lambda \in \mathbb{C}$ und $\rho : I \longrightarrow \mathbb{C}$ mit $\rho(x_0) = 0$, $\lim_{x \to x_0} \frac{\rho(x)}{x-x_0} = 0$ und $f(x) = f(x_0) + \lambda(x-x_0) + \rho(x)$ für alle $x \in I$.*
(iii) *Es gibt eine in x_0 stetige Funktion $r : I \longrightarrow \mathbb{C}$ mit $f(x) = f(x_0) + r(x)(x-x_0)$ für alle $x \in I$.*

In diesem Falle gilt $f'(x_0) = \lambda = r(x_0)$.

5.6.2. **Definition** (Landau-Symbole)**.** Seien $D \subset \mathbb{C}$, $z_0 \in \overline{\mathbb{C}}$ ein HP von D, $f : D \longrightarrow \mathbb{C}$, $g : D \longrightarrow \mathbb{C}$.

(i) Wir schreiben $f(z)=o(g(z))$ für $z\to z_0$, $z\in D$ (gelesen: $f(z)$ gleich klein-o von $g(z)$), falls zu jedem $\varepsilon>0$ eine Umgebung U von z_0 existiert, so dass $|f(z)|\leqslant\varepsilon|g(z)|$ für alle $z\in D\cap U$. Ist $g(z)\neq 0$ für alle z in einer Umgebung U von z_0, so ist diese Bedingung äquivalent zu $\lim_{z\to z_0}\frac{f(z)}{g(z)}=0$.

(ii) Gilt zusätzlich $\lim_{z\to z_0} g(z)=0$, so sagen wir auch: „f strebt für $z\to z_0$ schneller gegen 0 als g“. Ferner schreibt man $f=h+o(g)$ für $f-h=o(g)$.

(iii) Wir schreiben $f(z)=O(g(z))$ für $z\in D$, wenn eine Konstante C existiert, so dass $|f(z)|\leqslant C|g(z)|$ für alle $z\in D$.

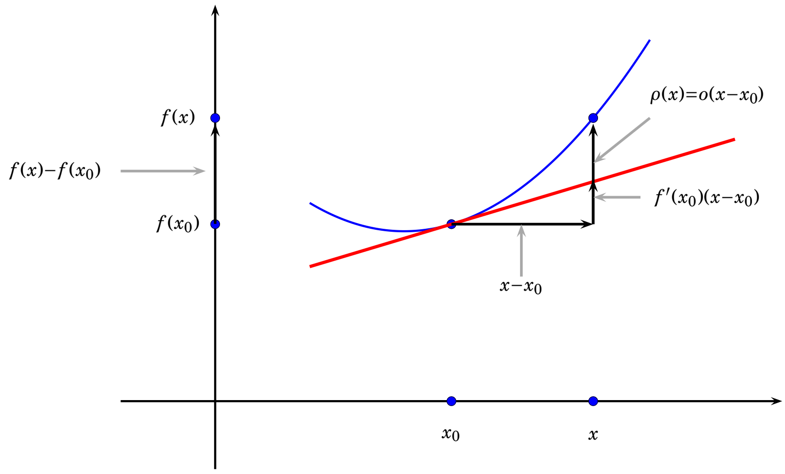
(iv) Wir schreiben $f(z)=O(g(z))$ für $z\to z_0$, $z\in D$, wenn eine Konstante C und eine Umgebung U von z_0 existieren, so dass $|f(z)|\leqslant C|g(z)|$ für alle $z\in D\cap U$.

5.6.3. **Bemerkung.**

(1) Mit dieser Symbolik bedeutet $\lim\limits_{x\to x_0}\frac{\rho(x)}{x-x_0}=0$ (siehe Satz 5.6.1(ii)), dass $\rho(x)=o(x-x_0)$ für $x\to x_0$. Man sagt auch, dass ρ „klein von 1. Ordnung“ in $(x-x_0)$ ist. Wir können also Satz 5.6.1 so umformulieren: f ist differenzierbar in x_0 genau dann, wenn es $\lambda\in\mathbb{C}$ gibt, so dass

$$f(x)=f(x_0)+\lambda(x-x_0)+o(x-x_0)\quad \text{für } x\to x_0.$$

(2) Satz 5.6.1(ii) zeigt, dass die Abbildung (5.1) die beste lineare Approximation von f auch im *analytischen* Sinne ist. Gesucht ist die affine Abbildung $l:\mathbb{R}\longrightarrow\mathbb{C}$, $l(x)=\lambda(x-x_0)+\mu$ mit $l(x_0)=f(x_0)$, d.h. $\mu=f(x_0)$, so dass $f(x)-l(x)$ schneller als jede lineare Funktion in der Variablen $h:=x-x_0$ gegen Null konvergiert. Das bedeutet gerade $f(x)-\lambda(x-x_0)-f(x_0)=o(x-x_0)$ für $x\to x_0$, und das ist äquivalent zu 5.6.1(ii).



Approximation des Funktionzuwachses $f(x)-f(x_0)$ durch den Wert $f'(x_0)(x-x_0)$

5.7. Übungen.

5.7.1. ***Aufgabe***. Zeigen Sie mit Hilfe von 5.6.1(ii): Ist $f:I\to\mathbb{R}$ in x_0 differenzierbar, so gilt

$$\begin{aligned}&f'(x_0)>0\Longrightarrow\exists\delta>0\;\forall h\in(0,\delta): f(x_0-h)<f(x_0)<f(x_0+h)\\&f'(x_0)<0\Longrightarrow\exists\delta>0\;\forall h\in(0,\delta): f(x_0-h)>f(x_0)>f(x_0+h)\end{aligned}\tag{5.9}$$

(Bemerkung: Die rechte Seite von (5.9) bedeutet nicht, dass f auf $(x_0-\delta,x_0+\delta)$ monoton ist.)

5.7.2. ***Aufgabe*** (Zwischenwertsatz für Ableitungen). Sei $I\subset\mathbb{R}$ ein Intervall und $f:I\to\mathbb{R}$ differenzierbar. Dann hat $f':I\to\mathbb{R}$ die Zwischenwerteigenschaft: Zu allen $a<b$ aus I und y zwischen $f'(a)$ und $f'(b)$ gibt es ein $c\in[a,b]$ mit $f'(c)=y$. (Tipp: Betrachten Sie $g(x)=f(x)-cx$ und zeigen Sie mit Hilfe von (5.9), dass für das Maximum von g in $[a,b]$ nur ein innerer Punkt $y\in(a,b)$ in Frage kommt. Wenden Sie dann Satz 5.3.3 auf g an. Man beachte, dass f' nicht unbedingt stetig ist; man darf also den Zwischenwertsatz 4.3.1 nicht anwenden.)

5.7.3. ***Aufgabe*** (Eine stetige nirgends differenzierbare Funktion). Für jedes $n\in\mathbb{N}$ definieren wir eine Funktion $f_n:[0,1]\to\mathbb{R}$ wie folgt: Ist $1\le m\le 2^n$ und $x\in\left[\frac{m-1}{2^n},\frac{m}{2^n}\right]$, so sei $f_n(x):=x-\frac{m-1}{2^n}$ für ungerades m und $f_n(x):=\frac{m}{2^n}-x$ für gerades m. Man überzeugt sich leicht, dass f_n wohldefiniert und stetig ist.

(a) Skizziere f_1, f_2, f_3, f_4 und $\sum_{k=1}^4 f_k$.
(b) Zeige, dass $f := \sum_{k=1}^\infty f_k : [0,1] \to \mathbb{R}$ existiert und stetig ist.
(c) Sei $g : I \to \mathbb{R}$ in $x_0 \in I$ differenzierbar. Seien $a_n \to x_0$ und $b_n \to x_0$ Folgen in I mit $a_n \le x_0 \le b_n$ und $a_n < b_n$ für alle n. Zeige: $d_n := \frac{g(b_n)-g(a_n)}{b_n-a_n} \to g'(x_0)$ für $n \to \infty$.

Tipp: Mit $g(x) = g(x_0) + r(x)(x - x_0)$ liegt d_n zwischen $r(a_n)$ und $r(b_n)$, warum?

(d) Sei nun $f : [0,1] \to \mathbb{R}$ wie in (b) und $x_0 \in [0,1]$. Zu jedem $n \in \mathbb{N}$ sei $[a_n, b_n]$ ein Intervall der Form $\left[\frac{m-1}{2^n}, \frac{m}{2^n}\right]$ (wie oben), das x_0 enthält. Zeige: $\frac{f(b_n)-f(a_n)}{b_n-a_n}$ ist für gerades n gerade und für ungerades n ungerade, und f ist in x_0 nicht differenzierbar.

5.7.4. ***Aufgabe***. Berechne jeweils die Ableitung der folgenden Funktionen:

$$(i)\ f_1 : \mathbb{R} \setminus \{\tfrac{-d}{c}\} \ni x \mapsto \frac{ax+b}{cx+d} \in \mathbb{C},\ \text{mit } a,b,c,d \in \mathbb{C}, c \neq 0$$
$$(ii)\ f_2 : \mathbb{R}_+ \ni x \mapsto x \log x \in \mathbb{R}\ ,$$
$$(iii)\ f_3 : \mathbb{R}_+ \ni x \mapsto x^x \in \mathbb{R}\ ,$$
$$(iv)\ f_4 : \mathbb{R} \ni x \mapsto \log\sqrt{1+\sin^2 x} \in \mathbb{R}\ ,$$
$$(v)\ f_5 := \tanh = \frac{\sinh}{\cosh} : \mathbb{R} \to \mathbb{R},$$

wobei $\cosh x = \frac{1}{2}(e^x + e^{-x})$ und $\sinh x = \frac{1}{2}(e^x - e^{-x})$.

Tipp: $\cosh^2(x) - \sinh^2(x) = 1$.

Sei $z \in \mathbb{C}$, und sei $f : \mathbb{R} \to \mathbb{C}$ definiert durch $f(x) := 0$ für $x \le 0$ und $f(x) := x^z$ für $x > 0$. Zeige: f ist genau dann im Punkt 0 differenzierbar, wenn $\operatorname{Re} z > 1$ gilt.

5.7.5. ***Aufgabe*** (Ableitungsregeln). Sei $f : \mathbb{R}_+^* \to \mathbb{R}$ mit $f(x) = \dfrac{\sin x}{\sqrt{x}}$. Zeigen Sie, daß

$$f''(x) + \frac{1}{x} f'(x) + \left(1 - \frac{1}{4x^2}\right) f(x) = 0\,.$$

5.7.6. ***Aufgabe***. (a) Zeigen Sie per Induktion: $\sin x = 2^n \sin\dfrac{x}{2^n} \cdot \displaystyle\prod_{k=1}^n \cos\frac{x}{2^k}$ für alle $x \in \mathbb{R}$, $n \in \mathbb{N}$.

(b)Schließen Sie aus (a): $\displaystyle\lim_{n\to\infty} \prod_{k=1}^n \cos\frac{x}{2^k} = \frac{\sin x}{x}$ für $x \in \mathbb{R} \setminus \{0\}$.

(c) Folgeren Sie mit $x = \frac{\pi}{2}$:

$$\sqrt{\tfrac{1}{2}} \cdot \sqrt{\tfrac{1}{2} + \tfrac{1}{2}\sqrt{\tfrac{1}{2}}} \cdot \sqrt{\tfrac{1}{2} + \tfrac{1}{2}\sqrt{\tfrac{1}{2} + \tfrac{1}{2}\sqrt{\tfrac{1}{2}}}} \cdot \ldots = \frac{2}{\pi} \qquad (5.10)$$

(***Vietasche Formel*** für π)

5.7.7. ***Aufgabe***. (a) Zeigen Sie mit Hilfe des Leibniz-Kriteriums 3.2.7, dass $\sin x > x - \dfrac{x^3}{3!}$ für $0 < x < \sqrt{6}$.

(b) Sei $(v_n)_{n\in\mathbb{N}}$ definiert durch $v_n = \displaystyle\prod_{k=1}^n \cos\frac{x}{2^{k+1}} = \frac{1}{2^n \sin\frac{\pi}{2^{n+1}}}$, $n \in \mathbb{N}$. Zeigen Sie, dass $(v_n)_{n\in\mathbb{N}}$ monoton fallend ist.

(c) Beweisen Sie, dass $0 < v_n - \dfrac{2}{\pi} < \dfrac{\sqrt{2}\,\pi^2}{48}\dfrac{1}{4^n} < \dfrac{3}{10}\dfrac{1}{4^n}$ (Tipp:(a)).

Dies bedeutet, dass $(v_n)_{n\in\mathbb{N}}$ sehr rasch gegen $\dfrac{2}{\pi}$ konvergiert; v_{21} liefert den ***approximativen Wert von*** π,

$$\boxed{\pi \approx 3{,}141592653589\ldots},$$

der bereits bis zur 12. Stelle hinter dem Komma korrekt ist.

5.7.8. ***Aufgabe***. Sei $P : \mathbb{R} \to \mathbb{R}$ ein Polynom. Zeigen Sie, dass die durch

$$f(x) = \begin{cases} P(\frac{1}{x})e^{-1/x}, x > 0 \\ 0, x \leqslant 0 \end{cases}$$

definierte Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ differenzierbar ist, und dass $f'(0) = 0$ gilt.

5.7.9. ***Aufgabe***. Sei $f : [-1,1] \longrightarrow \mathbb{R}$, $f(x) = x \arcsin x + \sqrt{1-x^2}$. Zeigen Sie, dass f differenzierbar in $x = -1$ und $x = 1$ ist, und dass $f'(1) = \frac{\pi}{2}$, $f'(-1) = -\frac{\pi}{2}$ gilt[4].

[4]Dies ist bemerkenswert, da arcsin und $x \to \sqrt{1-x^2}$ in 1 und -1 nicht differenzierbar sind.

6. Integralrechnung

Data aequatione quotcunque fluentes quantitates involvente, fluxiones invenire; et vice versa.

Brief von Newton an Leibniz

Die Integralrechnung behandelt zwei Problemstellungen:

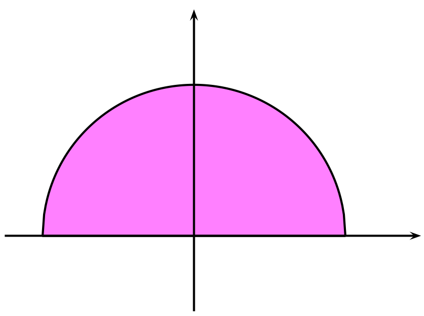
- Berechnung von Flächeninhalten.
- Inverse Operation zur Differentiation.

Wir gehen von der Problemstellung der Berechnung von Flächeninhalten aus. Zum Beispiel:

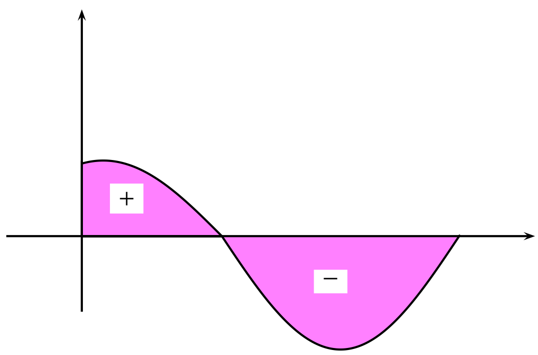
Problem: Berechne den Flächeninhalt einer Kreisscheibe vom Radius 1.

Es reicht, die Fläche einer halben Kreisscheibe zu berechnen. Betrachte dazu $f:[-1,1]\longrightarrow\mathbb{R}$, $f(x)=\sqrt{1-x^2}$. Die Halbkreisfläche ist der Flächeninhalt der Menge

$$\{(x,y)\in\mathbb{R}^2 : x\in[-1,1], 0\leqslant y\leqslant f(x)\}.$$



Halbkreisfläche als Ordinatenmenge von $x\mapsto\sqrt{1-x^2}$



Signierte Flächenbereiche einer Ordinatenmenge

Wir wollen das Integral so definieren, dass es unserer Vorstellung von einem Flächeninhalt zwischen dem Graphen und der x-Achse entspricht. Wenn die Funktion auch negative Werte annimmt, sollten die Flächenbereiche unterhalb der x-Achse negativ berechnet werden. So wird das Integral einer negativen Funktion negativ sein, was auch vernünftig ist (z.B. wegen der gewünschten Linearität des Integrals). Wir definieren also das Integral als den signierten Flächeninhalt zwischen dem Graphen und der x-Achse.

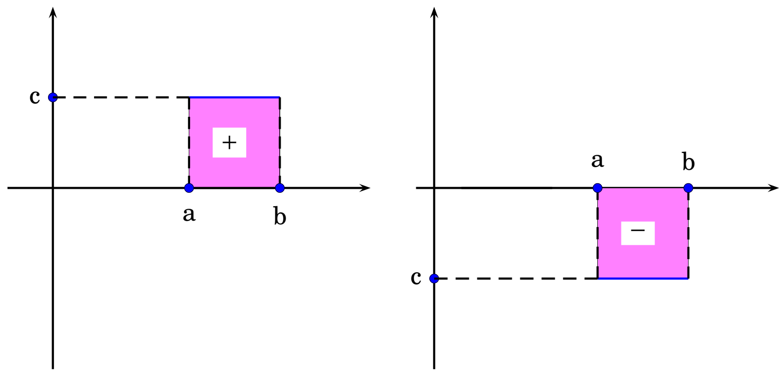
Sei $f:[a,b]\longrightarrow\mathbb{R}$. Die **Ordinatenmenge** von f ist

$$\mathcal{O}_f=\{(x,y)\in\mathbb{R}^2 : a\leqslant x\leqslant b\,,\, y\,\text{liegt zwischen 0 und } f(x)\}.$$

Wir definieren den signierten Flächeninhalt schrittweise:

- Für konstante Funktionen $f:[a,b]\longrightarrow\mathbb{R}$ mit $f(x)=c$ für alle $x\in[a,b]$.

$$\mathcal{O}_f=\begin{cases}[a,b]\times[0,c]\,,\ c\geqslant 0\,,\\ [a,b]\times[c,0]\,,\ c<0\,.\end{cases}$$

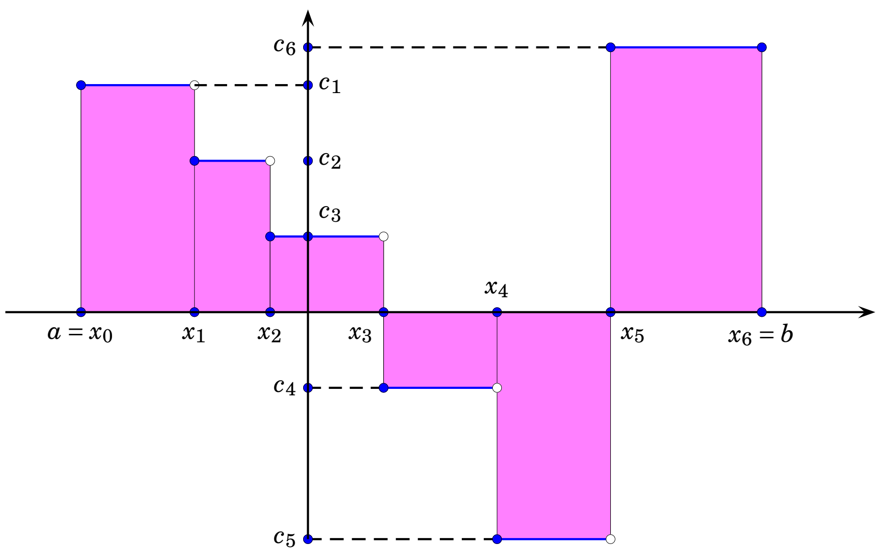


Der signierte Flächeninhalt von $\mathcal{O}_f$ ist dann $(b-a)\cdot c$. (Der echte Flächeninhalt ist $(b-a)\cdot|c|$, d.h. Höhe mal Breite.)

- Für stückweise konstante Funktionen. Diese Funktionen sind so definiert: Es gibt eine Zerlegung $a = x_0 < x_1 \ldots < x_{n-1} < x_n = b$ von $[a,b]$ so, dass für $k = 1,\ldots,n$ die Funktion f auf (x_{k-1},x_k) jeweils konstant mit dem Wert c_k ist. (Über die Funktionswerte $f(x_k)$ in den Teilpunkten wird nichts gefordert). Solche Funktionen heißen Treppenfunktionen (TF). Dann ist $\mathcal{O}_f$ eine disjunkte Vereinigung der Rechtecke

$$\begin{cases}(x_{k-1},x_k)\times[0,c],\ c\geqslant 0\\(x_{k-1},x_k)\times[c,0],\ c<0\end{cases}$$

und der Segmente $\{(x_k,y) : y \text{ liegt zwischen } 0 \text{ und } f(x_k)\}$.

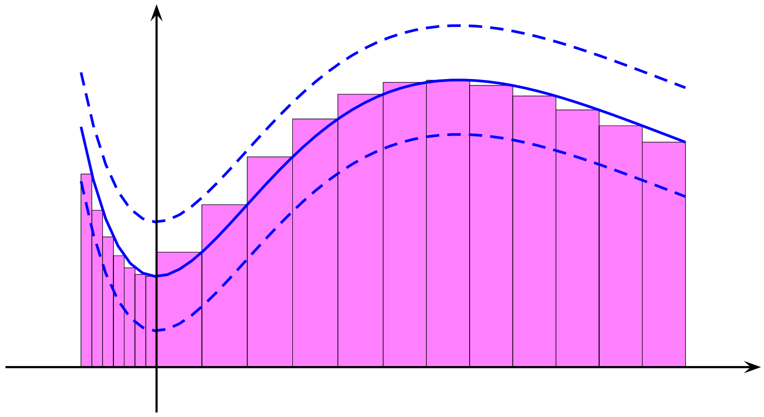


Da
- Segmente den Flächeninhalt Null haben sollen,
- der Flächeninhalt einer disjunkten Vereinigung gleich der Summe der Einzelflächeninhalte sein soll,

definieren wir den signierten Flächeninhalt von $\mathcal{O}_f$ als

$$\sum_{k=1}^{n} c_k(x_k - x_{k-1}).$$

- Für Funktionen, die durch Treppenfunktionen approximierbar sind. Für eine solche Funktion f gibt es zu jedem $\varepsilon > 0$ eine Treppenfunktion φ so, dass der Graph von φ zwischen den Graphen von $f-\varepsilon$ und $f+\varepsilon$ liegt. Solche Funktionen f heißen Regelfunktionen (RF). Der Flächeninhalt wird dann durch einen Grenzwertprozess für $\varepsilon \to 0$ definiert.



Programm:

- Definition des Integrals für Treppenfunktionen,
- Definition des Integrals für Regelfunktionen,
- Charakterisierung von Regelfunktionen,
- Hauptsatz der Differential- und Integralrechnung (Integration ist in gewissem Sinne die inverse Operation zur Differentiation),
- Integration rationaler Funktionen,
- uneigentliche Integrale.

.6.01.2012

6.1. **Das Integral von Treppenfunktionen.**

6.1.1. **Definition.** Eine Funktion $\varphi : [a,b] \longrightarrow \mathbb{C}$ heißt ***Treppenfunktion***, wenn es eine Zerlegung $Z = \{x_0,\dots,x_n\}$ mit $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ von $[a,b]$ gibt so, dass φ für jedes $k = 1,\dots,n$ auf (x_{k-1},x_k) jeweils konstant ist. Über die Funktionswerte $\varphi(x_k)$ in den Teilpunkten wird nichts gefordert. Notation: $\mathcal{T}([a,b]) := \{\varphi : [a,b] \longrightarrow \mathbb{C} : \varphi \text{ Treppenfunktion}\}$.

6.1.2. **Lemma.** *$\mathcal{T}([a,b])$ ist ein $\mathbb{C}$-Vektorraum. Für $\varphi,\psi \in \mathcal{T}([a,b])$ gehören auch $\varphi\psi$, $|\varphi|$ zu $\mathcal{T}([a,b])$. Sind φ,ψ reellwertig, so sind $\max\{\varphi,\psi\}$ und $\min\{\varphi,\psi\}$ Treppenfunktionen.*

Für $\varphi \in \mathcal{T}([a,b])$ und Z wie oben setzen wir:

$$I_Z(\varphi) := \sum_{k=1}^{n} c_k(x_k - x_{k-1})$$

6.1.3. **Lemma.** *Ist $\varphi \in \mathcal{T}([a,b])$, so ist $I_Z(\varphi)$ unabhängig von der Wahl von Z.*

6.1.4. **Definition.** Sei $\varphi \in \mathcal{T}([a,b])$, $Z = \{x_0,\dots,x_n\}$ eine Zerlegung von $[a,b]$ mit $\varphi(x) = c_k \in \mathbb{C}$ für alle $x \in (x_{k-1},x_k)$, $k = 1,\dots,n$. Dann heißt

$$\int_a^b \varphi(x)\,dx := I_Z(\varphi) := \sum_{k=1}^{n} c_k(x_k - x_{k-1})$$

das **Integral** von φ über $[a,b]$. Nach Lemma 6.1.3 ist das Integral wohldefiniert.

Notationen:

(i) Wir schreiben auch $\int_a^b \varphi\,dx$ statt $\int_a^b \varphi(x)\,dx$.

(ii) Für $f : [a,b] \longrightarrow \mathbb{C}$ setzen wir $\|f\| := \|f\|_{[a,b]} := \sup_{x\in[a,b]} |f(x)|$ (Supremumsnorm).

6.1.5. **Lemma.** *Seien $\varphi,\psi \in \mathcal{T}([a,b])$ und $\alpha,\beta \in \mathbb{C}$. Dann gilt:*

(i) $\displaystyle\int_a^b (\alpha\varphi + \beta\psi)\,dx = \alpha \int_a^b \varphi\,dx + \beta \int_a^b \psi\,dx$ *(Linearität)*

(ii) $\displaystyle\left|\int_a^b \varphi\,dx\right| \leqslant \int_a^b |\varphi|\,dx \leqslant (b-a)\|\varphi\|$ *(Beschränktheit)*

(iii) φ, ψ *reellwertig,* $\varphi \leqslant \psi \Rightarrow \int_a^b \varphi\, dx \leqslant \int_a^b \psi\, dx$ *(Monotonie)*

(iv) $c \in (a,b) \Rightarrow \int_a^b \varphi\, dx = \int_a^c \varphi\, dx + \int_c^b \varphi\, dx$ *(Intervall-Additivität)*

6.2. **Das Integral von Regelfunktionen.** Wie bestimmt man nun den Flächeninhalt bei „beliebigen“ Funktionen? Wir betrachten nur solche Funktionen, die sich durch Treppenfunktionen gleichmäßig approximieren lassen. (Den dabei auftretenden Begriff der gleichmäßigen Konvergenz von Funktionenfolgen werden wir im nächsten Kapitel ausführlicher untersuchen.)

6.2.1. **Definition.** Eine Funktion $f : [a,b] \longrightarrow \mathbb{C}$ heißt ***Regelfunktion***, wenn es zu jedem $\varepsilon > 0$ eine Treppenfunktion $\varphi \in \mathcal{T}([a,b])$ gibt mit $\|f - \varphi\| \leqslant \varepsilon$, d.h. $|f(x) - \varphi(x)| \leqslant \varepsilon$ für alle $x \in [a,b]$ (φ weicht nirgends mehr als ε von f ab). Dann heißt φ eine ε-approximierende TF zu f.

Sei $I \in \mathbb{R}$ ein beliebiges Intervall. Eine Funktion $f : I \longrightarrow \mathbb{C}$ heißt Regelfunktion, falls für jedes kompakte Intervall $[a,b] \subset I$ die Einschränkung $f|_{[a,b]}$ eine Regelfunktion ist.

Notation: $\mathcal{R}(I) = \{f : I \longrightarrow \mathbb{C} : f \text{ Regelfunktion}\}$.

Sei $f \in \mathcal{R}([a,b])$. Eine Folge $(\varphi_n)_{n\in\mathbb{N}}$ von TF heißt eine approximierende Folge von TF für f, falls $\|f - \varphi_n\| \to 0$ für $n \to \infty$. (Eine solche Folge existiert immer, z.B. φ_n TF mit $\|f - \varphi_n\| \leqslant \frac{1}{n}\cdot$ Man sagt auch, die Funktionenfolge (φ_n) konvergiert gleichmäßig gegen f).

Bemerkung:

(i) Jede Regelfunktion f auf einem kompakten Intervall $[a,b]$ ist beschränkt: Wähle eine TF $\varphi \in \mathcal{T}([a,b])$ mit $\|f - \varphi\| \leqslant 1$. Jede TF ist beschränkt, also $\|\varphi\| < \infty$. Nach der Dreiecksungleichung für die Supremumsnorm folgt $\|f\| = \|f - \varphi + \varphi\| \leqslant \|f - \varphi\| + \|\varphi\| \leqslant 1 + \|\varphi\| < \infty$.

(ii) Ist $(\varphi_n)_n$ eine approximierende Folge von TF für f, so gilt $\varphi_n(x) \to f(x)$ für alle $x \in [a,b]$ (d.h. $(\varphi_n)_n$ konvergiert punktweise gegen f) und $\|\varphi_n\| \to \|f\|$ für $n \to \infty$.

6.2.2. **Lemma.** *Die Menge* $\mathcal{R}([a,b])$ *ist ein* $\mathbb{C}$*-Vektorraum. Für* $f,g \in \mathcal{R}([a,b])$ *gehören auch* fg, $|f|$ *zu* $\mathcal{R}([a,b])$. *Sind* f,g *reellwertig, so sind auch* $\max\{f,g\}$ *und* $\min\{f,g\}$ *Regelfunktionen.*

6.2.3. **Beispiele.** Es ist klar, dass Treppenfunktionen auch Regelfunktionen sind d. h. $\mathcal{T}([a,b]) \subset \mathcal{R}([a,b])$ In §6.4 werden wir zeigen, dass auch stetige Funktionen und monotone Funktionen Regelfunktionen sind.

6.2.4. **Lemma.** *Sei* $f \in \mathcal{R}([a,b])$, *und sei* $(\varphi_n)_n$ *eine approximierende Folge von TF. Dann gilt:*

(a) $\lim\limits_{n\to\infty} \int_a^b \varphi_n\, dx$ *existiert,*

(b) $\lim\limits_{n\to\infty} \int_a^b \varphi_n\, dx$ *hängt nicht von der Wahl der Folge* $(\varphi_n)_n$ *ab.*

6.2.5. **Definition.** Sei $f \in \mathcal{R}([a,b])$, und sei $(\varphi_n)_n$ eine approximierende Folge von TF. Wir definieren ***das Integral*** von f über $[a,b]$ durch

$$\int_a^b f(x)\, dx := \int_a^b f\, dx := \lim_{n\to\infty} \int_a^b \varphi_n\, dx.$$

Nach Lemma 6.2.4 ist das Integral von f wohldefiniert. Weil dieses Integral für Regelfunktionen definiert ist, nennen wir es auch ***Regelintegral***.

Das von Leibniz eingeführte Integralzeichen $\int$ ist ein stilisiertes S und soll daran erinnern, dass das Integral (d.h. „das Ganze“) der Grenzwert einer Summe ist. Auf einem Notizzettel vom 29. Oktober 1675 schreibt Leibniz: Utile erit scribi $\int$ pro omn. ut $\int l$ pro omn. l, id est summa ipsorum l (es wird nützlich sein, $\int$ statt alle zu schreiben, wie $\int l$ anstatt alle l, das ist die Summe der Werte l).

6.2.6. **Beispiel.** $f : [0,b] \longrightarrow \mathbb{R}$, $f(x) = x^2$. Für $n \in \mathbb{N}$ sei

$$Z_n = \{0 < \tfrac{1}{n}b < \tfrac{2}{n}b < \ldots < \tfrac{n-1}{n}b < b\}.$$

Für $x \in \left[\frac{k-1}{n}b, \frac{k}{n}b\right)$ wähle $\varphi_n(x) = (\frac{k}{n}b)^2$, $\varphi_n(b) = b^2$. Dann ist

$$\|f - \varphi_n\| = \max\left\{\left|\left(\tfrac{k-1}{n}b\right)^2 - \left(\tfrac{k}{n}b\right)^2\right| : k = 1,\ldots,n\right\} = b^2 - (\tfrac{n-1}{n}b)^2$$

(weil $k^2-(k-1)^2 = 2k-1 \leqslant 2n-1 = n^2-(n-1)^2$). Also $\|f-\varphi_n\| \to 0$ für $n \to \infty$. Nach Definition folgt

$$\int_0^b x^2\,dx = \lim_{n\to\infty}\int_0^b \varphi_n\,dx = \lim_{n\to\infty}\sum_{k=1}^n \Big(\frac{k}{n}b\Big)^2 \cdot \frac{1}{n}b = b^3 \lim_{n\to\infty}\frac{1}{n^3}\sum_{k=1}^n k^2$$
$$= b^3 \lim_{n\to\infty}\frac{1}{n^3}\cdot\frac{n(n+1)(2n+1)}{6} = b^3\cdot\frac{2}{6} = \frac{b^3}{3}.$$

6.2.7. **Lemma.** *Seien $f,g \in \mathcal{R}([a,b])$ und $\alpha,\beta \in \mathbb{C}$. Dann gilt:*

(i) $\displaystyle\int_a^b (\alpha f+\beta g)\,dx = \alpha\int_a^b f\,dx + \beta\int_a^b g\,dx$ *(Linearität)*

(ii) $\displaystyle\Big|\int_a^b f\,dx\Big| \leqslant \int_a^b |f|\,dx \leqslant (b-a)\|f\|$ *(Beschränktheit)*

(iii) *f,g reellwertig, $f \leqslant g \Rightarrow \displaystyle\int_a^b f\,dx \leqslant \int_a^b g\,dx$ (Monotonie)*

(iv) *Sei $c \in (a,b)$. Dann sind $f|_{[a,c]}$ und $f|_{[c,b]}$ auch Regelfunktionen, und es gilt* $\displaystyle\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx$ *(Intervall-Additivität)*

6.2.8. **Notation.** Setze $\displaystyle\int_a^a f\,dx := 0$ und $\displaystyle\int_a^b f\,dx := -\int_b^a f\,dx$ falls $b<a$. Dann gilt

$$\Big|\int_a^b f\,dx\Big| \leqslant |b-a|\|f\|$$

20.01.2012

und 6.2.7 (iv) für alle $a,b,c \in \mathbb{R}$ (wenn f auf einem Intervall I mit $a,b,c \in I$ definiert ist).

6.3. **Der Hauptsatz der Differential- und Integralrechnung.** In einem Brief an Leibniz schreibt Newton „but because I cannot proceed with the explanation of it now, I have preferred to conceal it thus: 6accdae13eff7i3l9n4o4qrr4s8t12ux." Das Anagram bedeutet „Data aequatione quotcunque fluentes quantitates involvente, fluxiones invenire; et vice versa"; in unserer Sprache: „Gegeben eine Gleichung mit mehreren Funktionen, finde ihre Ableitungen und umgekehrt". Damit ist der Hauptsatz gemeint.

Wir beginnen mit einer physikalischen Motivation und zeigen, wie der Ort einer Bewegung als Integral dargestellt werden kann Wir betrachten einen Punkt, der sich entlang einer Geraden (x-Achse) bewegt (siehe auch Beispiel 5.1.2 (4)). Sei $x(t)$ der Ort und $v(t) := x'(t) =: \dot{x}(t)$ die Geschwindigkeit des Punktes zur Zeit t. Ist $v = v(t)$ konstant, so befindet sich der Punkt zum Zeitpunkt $t_1 > t_0$ am Ort $x(t_1) = x(t_0) + v\cdot(t_1-t_0)$ („Weg gleich Zeit mal Geschwindigkeit").

Nehmen wir nun an, dass a und b zwei Zeitpunkte sind, und dass die Geschwindigkeit $v:[a,b]\to\mathbb{R}$ zwar variabel, aber eine stetige Funktion ist. Sei $t \in [a,b]$ ein Zeitpunkt. Wir ermitteln den Ort $x(t)$, indem wir das Intervall $[a,t]$ in kleine Intervalle zerlegen, auf denen die Geschwindigkeit „fast konstant" ist. Dafür betrachten wir eine approximierende Folge (v_k) von TF mit $\|v-v_k\| \leqslant 1/k$. Sei $a = t_0 < t_1 < \ldots < t_n = t$ eine Zerlegung von $[a,t]$ mit $v_k = c_j$ auf (t_{j-1},t_j), $j = 1,\ldots,n$. Wenn wir uns vorstellen, dass sich der Punkt mit Geschwindigkeit v_k bewegt[5], so befindet er sich zum Zeitpunkt t am Ort

$$x_k(t) = x(a) + \sum_{j=1}^n c_j(x_j - x_{j-1}) = x(a) + \int_a^t v_k(s)\,ds. \tag{6.1}$$

Die Intuition besagt, dass $x_k(t)$ gegen $x(t)$ strebt, wenn k gegen Unendlich strebt. Andererseits gilt $\int_a^t v_k(s)\,ds \to \int_a^t v(s)\,ds$ für $k\to\infty$ nach der Definition des Integrals. Durch Grenzübergang in (6.1) erhalten wir die *Vermutung*

$$\boxed{x(t) = x(a) + \int_a^t \dot{x}(s)\,ds} \tag{6.2}$$

Leiten wir (6.2) ab, so finden wir $\dot{x}(t) = \dfrac{d}{dt}\displaystyle\int_a^t v(s)\,ds$, also

$$\boxed{v(t) = \frac{d}{dt}\int_a^t v(s)\,ds} \tag{6.3}$$

[5]Dies ist eine mathematische Idealisierung; physikalisch muss $t \mapsto x(t)$ zweimal differenzierbar sein, damit man von Beschleunigung sprechen kann. Die Geschwindigkeit soll also differenzierbar, insbesondere stetig sein.

Wegen (6.2) und (6.3) erweist sich die Integration als Umkehrung der Differentiation. Das ist im wesentlichen der Hauptsatz; wir möchten diesen streng beweisen. Die Formel (6.3) besagt, dass die Ableitung des Integrals als Funktion der oberen Grenze das Integrand selbst ist.

6.3.1. **Definition.** Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{C}$. Eine Funktion $F : I \to \mathbb{C}$ heißt ***Stammfunktion*** von f, falls F an jeder Stelle $x \in I$ differenzierbar ist mit $F'(x) = f(x)$.

Wir möchten zunächst die folgende Frage beantworten: Wie unterscheiden sich zwei Stammfunktionen? Es gilt $F_1' = F_2'$ auf I, und nach dem Eindeutigkeitssatz der Differentialrechnung 5.4.1 existiert $C \in \mathbb{C}$ mit $F_1 = F_2 + C$.

6.3.2. **Satz** (Hauptsatz der Differential- und Integralrechnung)**.** *Es sei $f : I \longrightarrow \mathbb{C}$ eine stetige Funktion auf einem Intervall I. Ein Punkt $a \in I$ sei fest gewählt, und für $x \in I$ setze man*

$$F(x) := \int_a^x f(t)\,dt.$$

Dann gilt:

(i) *F ist eine Stammfunktion zu f auf I;*
(ii) *Ist Φ eine beliebige Stammfunktion zu f auf I, so gilt für alle $x \in I$:*

$$\int_a^x f(t)\,dt = \Phi(x) - \Phi(a) =: \Phi\Big|_a^x$$

d.h. $\Phi = F + \Phi(a)$.

Der Hauptsatz hat zwei Aspekte:

- Auf der theoretischen Seite besagt der Hauptsatz, dass wir aus der lokalen Kenntnis der Änderung einer Funktion, d.h. ihrer Ableitung, die Funktion wieder zurückgewinnen können. Da die Naturgesetze als Differentialgleichungen formuliert sind, ist der Hauptsatz in gewissem Sinne die Grundlage der gesamten Naturwissenschaft.
- Auf der praktischen Seite besagt der Hauptsatz, dass wir aus der Kenntnis einer Stammfunktion das Intergral einer Funktion ohne Rückgriff auf Definition berechnen können. Zum Beispiel:

$$\int_0^b x^2\,dx = \frac{x^3}{3}\Big|_0^b = \frac{b^3}{3}, \qquad \int_0^b \frac{1}{1+x^2}\,dx = \arctan x\Big|_0^1 = \frac{\pi}{4}.$$

  Wenn möglich, werden wir ein bestimmtes Integral nach diesem einfachen Verfahren berechnen. Darum ist es wichtig, Integrationsregeln kennenzulernen (partielle Integration und Substitutionsregel), mit denen Stammfunktionen aufgefunden werden können.

6.3.3. **Definition.** Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{C}$ eine stetige Funktion. Das ***unbestimmte Integral*** $\int f(x)\,dx$ von f bezeichnet entweder die Menge aller Stammfunktionen oder (repräsentativ) eine bestimmte davon. Sei F_0 eine feste Stammfunktion; dann gilt

$$\int f(x)\,dx = \{F_0 + C : C \in \mathbb{C}\} = F_0 + \mathbb{C}.$$

Wir lassen üblicherweise die Mengenklammern fort und schreiben $\int f(x)\,dx = F_0 + C$.

Funktion	Stammfunktion
$f:\mathbb{R}_+ \longrightarrow \mathbb{C}$, $f(x)=x^\alpha$, $\alpha \in \mathbb{C} \setminus \{-1\}$	$\int x^\alpha\,dx = \dfrac{x^{\alpha+1}}{\alpha+1} + C$
$f: I \subset \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}$, $f(x)=\dfrac{1}{x}$	$\int \dfrac{1}{x}\,dx = \log\lvert x\rvert + C$
$f:\mathbb{R} \longrightarrow \mathbb{C}$, $f(x)=e^{cx}$ $c \in \mathbb{C} \setminus \{0\}$	$\int e^{cx} dx = \dfrac{1}{c} e^{cx} + C$
$f:\mathbb{R} \longrightarrow \mathbb{R}$, $f(x)=\sin x$	$\int \sin x\,dx = -\cos x + C$
$f:\mathbb{R} \longrightarrow \mathbb{R}$, $f(x)=\cos x$	$\int \cos x\,dx = \sin x + C$
$f:(-1,1) \longrightarrow \mathbb{R}$, $f(x)=\dfrac{1}{\sqrt{1-x^2}}$	$\int \dfrac{1}{\sqrt{1-x^2}}\,dx = \arcsin x + C$
$f:\mathbb{R} \longrightarrow \mathbb{R}$, $f(x)=\dfrac{1}{1+x^2}$	$\int \dfrac{1}{1+x^2}\,dx = \arctan x + C$

In der Definition des unbestimmtes Integrals $\int f(x)dx = F_0 + C$, kann die Konstante C in verschieden Formen geschrieben werden. Wesentlich ist, dass C die ganze Menge $\mathbb{C}$ läuft. Die Konstante C kann $\log C$, $C \in \mathbb{R}_+$, oder αC, mit $\alpha \neq 0$ fest $C \in \mathbb{C}$, oder $C_1 + C_2$, wobei $C_1, C_2 \in \mathbb{C}$. Wegen der Definition der Stammfunktion (und für bestimmte Integrale auch wegen des Hauptsatzes) kann jede formale Differentiationsregel in die Integralrechnung übertragen werden. Die Regeln $(F+G)' = F' + G'$, $(\lambda F)' = \lambda F'$ ergeben

$$\int (f+g)dx = \int f\,dx + \int g\,dx, \quad \int (\alpha f)dx = \alpha \int f\,dx \; (\alpha \neq 0)$$

Diese Gleichungen sind Gleichungen zwichen *Mengen* von Funktionen. Falls F, G, H sind Stammfunktionen von f, g, $f+g$, ist die Gleichung $F+G=H$ i. A. falsch. Geben Sie ein Gegenbeispiel! Eine Anwendung der obigen Regeln ergibt z. B.

$$\int (a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0)dx = a_n \frac{x^{n+1}}{n+1} + a_{n-1}\frac{x^n}{n} + \ldots + a_1 \frac{x^2}{2} + a_0 x + C.$$

Die Produktregel wird zur partiellen Integration.

6.3.4. **Satz** (Partielle Integration). *Seien $u, v : I \longrightarrow \mathbb{C}$ stetig differenzierbar. Dann ist auch uv stetig differenzierbar, und es gilt*

$$\int uv'\,dx = uv - \int u'v\,dx, \qquad \text{d.h.} \quad \int_a^b uv'\,dx = uv\Big|_a^b - \int_a^b u'v\,dx, \quad \text{für alle } a,b \in I.$$

Damit wird das Integral teilweise (partiell) gelöst. Man versucht also, den Integranden als $u'v$ zu schreiben, so dass das aus der partiellen Integration resultierende Integral leichter zu berechnen ist. Dabei bestimmen wir u aus u'; das ist auch eine Integration.

Beispiel: Berechne $\int \cos^n x\,dx$ für $n \geqslant 2$. Wir setzen $\cos^n x = \cos x \cos^{n-1} x = u'v$, wobei $u'(x) = \cos x$ und $v(x) = \cos^{n-1} x$. Dann gilt $u(x) = \sin x$ und $v'(x) = (n-1)\cos^{n-2} x(-\sin x)$. Nach Verwendung der Formel $\sin^2 x = 1 - \cos^2 x$ folgt:

$$\int \cos^n x\,dx = \frac{1}{n}\cos^{n-1} x \sin x + \frac{n-1}{n}\int \cos^{n-2} x\,dx.$$

und analog:

$$\int \sin^n x\,dx = -\frac{1}{n}\sin^{n-1}x\cos x + \frac{n-1}{n}\int \sin^{n-2}x\,dx.$$

Anwendung auf das Wallissche Produkt. Nach dem Hauptsatz gilt

$$I_n := \int_0^{\frac{\pi}{2}} \cos^n x\,dx = \frac{1}{n}\cos^{n-1}x\sin x\Big|_0^{\frac{\pi}{2}} + \frac{n-1}{n}\int_0^{\frac{\pi}{2}}\cos^{n-2}x\,dx = \frac{n-1}{n}I_{n-2}. \tag{6.4}$$

Daraus folgt

$$\begin{aligned} I_{2n} &= \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\cdots\frac{1}{2}\cdot I_0 = \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\cdots\frac{1}{2}\cdot\frac{\pi}{2}\\ I_{2n+1} &= \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\cdots\frac{2}{3}\cdot I_1 = \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\cdots\frac{2}{3}\cdot 1. \end{aligned} \tag{6.5}$$

Wir definieren die Wallissche Folge $(w_n)_{n\geqslant 1}$ durch

$$w_n := \frac{2\cdot 2}{1\cdot 3}\,\frac{4\cdot 4}{3\cdot 5}\cdots\frac{2n\cdot 2n}{(2n-1)(2n+1)}. \tag{6.6}$$

Dann ist

$$w_n = \frac{I_{2n+1}}{I_{2n}}\cdot\frac{\pi}{2}.$$

Aus $\frac{I_{2n+1}}{I_{2n}} \to 1$ folgt nun die Wallissche Formel

$$\boxed{\frac{\pi}{2} = \lim_{n\to\infty}\frac{2\cdot 2}{1\cdot 3}\,\frac{4\cdot 4}{3\cdot 5}\cdots\frac{2n\cdot 2n}{(2n-1)(2n+1)}} \quad \text{(Wallis 1655)}.$$

Die Kettenregel der Differentialrechnung liefert die Substitutionsregel der Integralrechnung.

6.3.5. **Satz** (Substitutionsregel)**.** *Seien I, J Intervalle in $\mathbb{R}$, sei $f : I \longrightarrow \mathbb{C}$ stetig und $\varphi : J \longrightarrow I$ stetig differenzierbar. Dann gilt:*

(i) *Die Funktion $(f\circ\varphi)\varphi'$ ist stetig, und für jede Stammfunktion $F : I \longrightarrow \mathbb{C}$ zu f gilt*

$$\int (f\circ\varphi)(t)\varphi'(t)\,dt = F\circ\varphi + C =: \int f(x)\,dx\Big|_{x=\varphi(t)}.$$

(ii) *Ist φ bijektiv mit $\varphi' \neq 0$ überall in J, und ist H eine Stammfunktion von $(f\circ\varphi)\varphi'$, so ist $H\circ\varphi^{-1} : I \longrightarrow \mathbb{C}$ eine Stammfunktion von f, und es gilt*

$$\int f\,dx = H\circ\varphi^{-1} + C =: \int (f\circ\varphi)(t)\varphi'(t)\,dt\Big|_{t=\varphi^{-1}(x)}.$$

(iii) *Für alle $\alpha, \beta \in J$ gilt*

$$\int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\,dx.$$

6.3.6. **Beispiele.** (1) Der Fall, dass das Integrand die besondere Form $(f\circ\varphi)(x)\varphi'(x)$ hat.

Aufgabe: Gesucht wird eine Stammfunktion von $g : (1,\infty) \longrightarrow \mathbb{R}$, $g(x) = \dfrac{1}{x\log x}$. Betrachte $\varphi : (1,\infty) \longrightarrow \mathbb{R}_+$, $\varphi(x) = \log x$ und $f : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$, $f(y) = \dfrac{1}{y}$. Aus $g(x) = \dfrac{(\log x)'}{\log x} = (f\circ\varphi)(x)\varphi'(x)$ folgt

$$\begin{aligned}\int \frac{1}{x\log x}\,dx = \int (f\circ\varphi)(x)\varphi'(x)\,dx = \int f(y)\,dy\Big|_{y=\varphi(x)} &= \int \frac{1}{y}\,dy\Big|_{y=\varphi(x)}\\ &= \log y\Big|_{y=\log x} + C = \log(\log x) + C.\end{aligned}$$

Es ist zweckmäßig, das Differential einer differenzierbaren Funktion φ einzuführen; es ist der formale Ausdruck $d\varphi = \varphi(x)\,dx$. Später werden wir das Differential als Differentialform definieren. Formal geht man so vor:

$\boxed{1}$ Setze $y = \varphi(x)$; dann $dy = \varphi'(x)\,dx$ und $f(\varphi(x))\varphi'(x)\,dx = f(y)\,dy$.

$\boxed{2}$ Evaluiere $F(y) = \int f(y)\,dy$.

$\boxed{3}$ Substituiere $y = \varphi(x)$ in $F(y)$ und erhalte $\int f(\varphi(x))\varphi'(x)\,dx = F(\varphi(x)) + C$.

(2) Das Integrand liegt oft nicht in der angenehmen Form $(f\circ\varphi)(x)\varphi'(x)$ vor. Man versucht dann, eine geeignete Substitution $x=\varphi(t)$ zu finden, durch die das Integral $\int f(x)dx$ in ein leichter zu berechnendes Integral $\int (f\circ\varphi)(x)\varphi'(x)$ übergeht.

Aufgabe: Wir suchen eine Stammfunktion von $f:(-1,1)\longrightarrow\mathbb{R}$, $f(x)=\sqrt{1-x^2}$. Betrachte $\varphi:(-\frac{\pi}{2},\frac{\pi}{2})\longrightarrow(-1,1)$, $\varphi(t)=\sin t$ (dann $\varphi'(t)=\cos t\neq 0$). Es gilt

$$\begin{aligned}\int (f\circ\varphi)(t)\varphi'(t)\,dt &= \int\sqrt{1-\sin^2 t}\cdot\cos t\,dt=\int|\cos t|\cos t\,dt=\int\cos^2 t\,dt\\ &=\frac{1}{2}(\sin t\cos t+t)+C,\ \text{daher}\\ \int f\,dx &= \frac{1}{2}(\sin t\cos t+t)\Big|_{t=\arcsin x}+C=\frac{1}{2}[x\cos(\arcsin x)+\arcsin x]+C\\ &=\frac{1}{2}\Big[x\sqrt{1-\sin^2(\arcsin x)}+\arcsin x\Big]+C=\frac{1}{2}\Big[x\sqrt{1-x^2}+\arcsin x\Big]+C.\end{aligned}$$

Formal geht man so vor:

$\boxed{1}$ Setze $x=\varphi(t)$; dann $dx=\varphi'(t)dt$ und $f(x)dx=f(\varphi(t))\varphi'(t)dt$.

$\boxed{2}$ Evaluiere $H(t)=\displaystyle\int f(\varphi(t))\varphi'(t)dt$.

$\boxed{3}$ Substituiere $t=\varphi^{-1}(x)$ in $H(t)$ und erhalte $\displaystyle\int f(x)dx=H(\varphi^{-1}(x))+C$.

Anwendung auf die Berechnung des Flächeninhalts des Kreises.

Sei $F:[-1,1]\to\mathbb{R}$, $F(x)=\frac{1}{2}[x\sqrt{1-x^2}+\arcsin x]$. Dies ist eine stetige Funktion auf $[-1,1]$ mit $F'=f$ auf $(-1,1)$. Der Differenzierbarkeitsatz 5.4.7 impliziert, dass F auch in $x=1$ und $x=-1$ differenzierbar ist und $F'(1)=f(1)$ und $F'(-1)=f(-1)$, d. h. F ist eine Stammfunktion von f auf $[-1,1]$. Daraus und aus dem Hauptsatz folgt

$$\begin{aligned}\int_{-1}^{1}\sqrt{1-x^2}\,dx &= \frac{1}{2}\Big(x\sqrt{1-x^2}+\arcsin x\Big)\Big|_{-1}^{1}=\frac{1}{2}\arcsin x\Big|_{-1}^{1}\\ &=\frac{1}{2}[\arcsin 1-\arcsin(-1)]=\frac{1}{2}\Big(\frac{\pi}{2}+\frac{\pi}{2}\Big)=\frac{\pi}{2}.\end{aligned}$$

Der Flächeninhalt des Halbkreises vom Radius 1 beträgt somit $\pi/2$. Der Flächeninhalt des Einheitskreises beträgt also π. Für einen Kreis vom Radius r betrachtet man den in der oberen Halbebene liegenden Teil des Kreises als Graph der Funktion $f:[-r,r]\to\mathbb{R}$, $f(x)=\sqrt{r^2-x^2}$. Mit der Substitution $x=r\sin t$ berechnet man wie oben, dass der Flächeninhalt πr^2 beträgt.

23.01.2012

6.4. **Charakterisierung und Eigenschaften der Regelfunktionen.** Wir geben nun eine schöne Charakterisierung der Regelfunktionen an, die viel einfacher zu überprüfen ist als die ursprüngliche Definition.

6.4.1. **Satz.** *Die Funktion $f:[a,b]\longrightarrow\mathbb{C}$ ist eine Regelfunktion genau dann, wenn f an jeder Stelle einseitige Grenzwerte besitzt, d.h. genau dann, wenn für alle $x\in[a,b)$ der Grenzwert $f(x+)=\lim_{t\searrow x}f(t)\in\mathbb{C}$ und für alle $x\in(a,b]$ der Grenzwert $f(x-)=\lim_{t\nearrow x}f(t)\in\mathbb{C}$ existiert.*

Diese Charakterisierung gilt auch für Regelfunktionen auf einem beliebigen Intervall $I\subset\mathbb{R}$.

6.4.2. **Folgerung.** *Ist $f:[a,b]\longrightarrow\mathbb{C}$ stetig, so ist f eine Regelfunktion. Ist $f:[a,b]\longrightarrow\mathbb{R}$ monoton, so ist f eine Regelfunktion.*

An einer Stetigkeitsstelle x einer Funktion f sind die einseitigen Grenzwerte gleich dem Wert $f(x)$ (wenn x ein Endpunkt des Intervalles ist, betrachten wir nur den einseitigen Grenzwert). Mit Blick auf Satz 6.4.1 fragen wir uns: Wie viele Unstetigkeitsstellen kann eine Regelfunktion haben? Die Antwort liefert der nächste Satz.

6.4.3. **Satz.** *Jede Regelfunktion $f:[a,b]\longrightarrow\mathbb{C}$ ist stetig mit Ausnahme höchstens abzählbar vieler Stellen, d.h. es gibt eine höchstens abzählbare Teilmenge $A\subset[a,b]$, so dass f an jeder Stelle $x\in[a,b]\setminus A$ stetig ist.*

6.4.4. **Satz** (Mittelwertsatz der Integralrechnung). *Es sei $f:[a,b]\longrightarrow\mathbb{R}$ eine stetige Funktion, und es sei $g:[a,b]\longrightarrow\mathbb{R}$ eine Regelfunktion mit $g\geqslant 0$ oder $g\leqslant 0$ auf $[a,b]$. Dann gibt es ein $\xi\in[a,b]$ mit*

$$\int_a^b f(x)g(x)\,dx=f(\xi)\cdot\int_a^b g(x)\,dx.$$

Im Spezialfall $g = 1$ folgt, dass es ein $\xi \in [a,b]$ gibt mit

$$\int_a^b f(x)\,dx = f(\xi)(b-a).$$

26.01.2012

6.5. Höhere Ableitungen. Die Taylorformel.

6.5.1. **Definition.** Sei I ein Intervall, $x_0 \in I$ und $f : I \to \mathbb{C}$. Sei $n \in \mathbb{N}_0$.

(i) Wir definieren den Begriff der n-maligen Differenzierbarkeit und die n-te Ableitung $f^{(n)}(x_0) = \frac{d^n f}{dx^n}(x_0)$ rekursiv, und zwar wie folgt:

(0) Jede beliebige Funktion f heißt 0-mal differenzierbar. Die 0-te Ableitung wird durch $f^{(0)} := f$ definiert.

(1) Für $n = 1$ sagen wir, dass f in $x_0 \in I$ 1-mal differenzierbar ist, falls f in x_0 differenzierbar ist. In diesem Fall ist die erste Ableitung von f in x_0 als $f^{(1)}(x_0) := f'(x_0)$ definiert.

(2) Sei $n \in \mathbb{N}$, $n \geqslant 2$; wir nehmen an, dass der Begriff der $(n-1)$-maligen Differenzierbarkeit und die $(n-1)$-te Ableitung schon definiert sind. Die Funktion f heißt ***n-mal differenzierbar*** in x_0, falls es $\varepsilon > 0$ gibt, so dass f in jedem Punkt von $(x_0-\varepsilon, x_0+\varepsilon) \cap I$ $(n-1)$-differenzierbar ist und die Abbildung $f^{(n-1)} : (x_0-\varepsilon, x_0+\varepsilon) \cap I \longrightarrow \mathbb{C}$, $x \mapsto f^{(n-1)}(x)$ in x_0 differenzierbar ist. Die ***n-te Ableitung*** von f in x_0 ist dann durch $f^{(n)}(x_0) := (f^{(n-1)})'(x_0)$ definiert.
Klassische Notationen: $f^{(1)} = f'$, $f^{(2)} = (f')' =: f''$, $f^{(3)} = (f'')' =: f'''$.

(ii) Wenn f n-mal differenzierbar in allen $x \in I$ ist, so sagen wir, dass f ***n-mal differenzierbar*** ist. Die Funktion $f^{(n)} : I \longrightarrow \mathbb{C}$ heißt die n-te Ableitung von f.
f heißt ***unendlich oft differenzierbar***, wenn f für jedes $n \in \mathbb{N}$ n-mal differenzierbar ist.

(iii) f heißt ***n-mal stetig differenzierbar***, wenn f n-mal differenzierbar und $f^{(n)} : I \longrightarrow \mathbb{C}$ stetig ist.

(iv) Wir führen für $n \in \mathbb{N}_0$ die folgenden Mengen ein:

$$\mathcal{C}^n(I) := \{f : I \longrightarrow \mathbb{C} : f \text{ ist } n\text{–mal stetig differenzierbar}\}$$
$$\mathcal{C}^\infty(I) := \{f : I \longrightarrow \mathbb{C} : f \text{ unendlich oft differenzierbar}\}.$$

Insbesondere ist $\mathcal{C}^0(I) = \{f : I \longrightarrow \mathbb{C} : f \text{ stetig}\}$, und es gilt

$$\mathcal{C}^\infty(I) = \cap_{n\in\mathbb{N}} \mathcal{C}^n(I).$$

6.5.2. **Beispiel.**

(1) Jedes Polynom $P : \mathbb{R} \longrightarrow \mathbb{R}$ gehört zu $\mathcal{C}^\infty(\mathbb{R})$, und $P^{(n)} = 0$ für $n > \mathrm{grad}(P)$.

(2) Sei $s \in \mathbb{R}$. Die Funktion $f : (0,\infty) \longrightarrow \mathbb{R}$, $x \mapsto x^s$ ist unendlich oft differenzierbar. Es gilt für alle $n \in \mathbb{N}$ und alle $x > 0$:

$$\Big(\frac{d}{dx}\Big)^n x^s = n!\binom{s}{n} x^{s-n}.$$

(3) Die Funktionen $\exp, \sin, \cos : \mathbb{R} \longrightarrow \mathbb{R}$ und $\log : (0,\infty) \longrightarrow \mathbb{R}$ sind unendlich oft differenzierbar.
Für $n \in \mathbb{N}$ und beliebiges $c \in \mathbb{R}$ ist

$$\Big(\frac{d}{dx}\Big)^n e^{cx} = c^n e^{cx} \quad \text{für alle } x \in \mathbb{R},$$

$$\Big(\frac{d}{dx}\Big)^n \sin x = \begin{cases} \sin x, & n = 4k \\ \cos x, & n = 4k+1 \\ -\sin x, & n = 4k+2 \\ -\cos x, & n = 4k+3 \end{cases} \qquad \text{für alle } x \in \mathbb{R},$$

$$\Big(\frac{d}{dx}\Big)^n \log x = (-1)^{n-1}(n-1)!\,\frac{1}{x^n} \quad \text{für alle } x > 0.$$

(4) Die Funktion (5.7) liegt in $\mathcal{C}^\infty(\mathbb{R})$. Man zeigt nämlich durch Induktion über $n \in \mathbb{N}$, dass f n-mal differenzierbar ist und dass

$$f^{(n)}(x) = \begin{cases} P_n(\frac{1}{x})e^{-1/x}, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

wobei P_n ein Polynom vom Grad $3n$ ist. Dabei benutzt man die Folgerung 5.4.7.

Sind $f,g:I\to\mathbb{C}$ n-mal differenzierbar, so sind auch $f+g$ und $f\cdot g$ n-mal differenzierbar, und es gilt:

$$(f+g)^{(n)}=f^{(n)}+g^{(n)},\quad (f\cdot g)^{(n)}=\sum_{k=0}^{n}\binom{n}{k}f^{(k)}g^{(n-k)}.$$

Die letzte Formel heißt ***Leibnizsche Produktregel***. Daraus folgt, dass für jedes Intervall $I\subset\mathbb{R}$ und alle $k\in\mathbb{N}\cup\{\infty\}$ die Menge $\mathcal{C}^k(I)$ eine $\mathbb{C}$-Algebra ist.

Darüber hinaus beweist man durch Induktion die folgenden Varianten der Sätze 5.2.3 und 5.2.5:

- Die Komposition n-mal (stetig) differenzierbarer Funktionen ist n-mal (stetig) differenzierbar.
- Ist $f:I\longrightarrow\mathbb{R}$ stetig, streng monoton und n-mal differenzierbar in $x_0\in I$ mit $f'(x_0)\neq 0$, so ist die Umkehrfunktion $g:f(I)\longrightarrow\mathbb{R}$ n-mal differenzierbar in $y_0:=f(x_0)$.

Aus Satz 5.6.1 wissen wir, dass eine in $x_0\in I$ differenzierbare Funktion $f:I\longrightarrow\mathbb{C}$ die Form $f(x)=f(x_0)+f'(x_0)(x-x_0)+\rho(x)$ hat, wobei $\rho:I\longrightarrow\mathbb{C}$ die Bedingung $\lim_{x\to x_0}\frac{\rho(x)}{x-x_0}=0$ erfüllt. Dies bedeutet, dass f in der Nähe von x_0 durch eine lineare Funktion (Polynom vom Grad Eins) approximiert werden kann. Inwieweit kann f durch Polynome höherer Ordnung approximiert werden?

Um die Form des approximierenden Polynoms zu erraten, betrachten wir zunächst den Fall, dass f selbst ein Polynom $f(x)=\sum_{k=0}^{n}a_k(x-x_0)^k$ vom Grad n ist. Wegen

$$\frac{d^j}{dx^j}(x-x_0)^k=\begin{cases}k(k-1)\cdots(k-j+1)(x-x_0)^{k-j}=\frac{k!}{(k-j)!}(x-x_0)^{k-j}, & j\leqslant k,\\ 0, & j>k\end{cases}$$

sind diese Ableitungen an der Stelle x_0 alle gleich Null, nur die k-te Ableitung hat den konstanten Wert $k!$. Die höheren Ableitungen ($j>k$) verschwinden wieder sämtlich. Daher folgt für alle $k\in\{0,\dots,n\}$:

$$f^{(k)}(x_0)=k!a_k\,.$$

Man kann die Koeffizienten von f also aus den höheren Ableitungen von f berechnen und erhält:

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k\,.$$

6.5.3. **Definition.** Sei $n\in\mathbb{N}$, und sei $f:I\longrightarrow\mathbb{C}$ n-mal differenzierbar. Zu $x_0\in I$ assoziieren wir das n-te ***Taylorpolynom*** zu f mit Entwicklungspunkt x_0:

$$T_{x_0,n}(x)=f(x_0)+\frac{1}{1!}f'(x_0)(x-x_0)+\ldots+\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k\,.$$

Das konstante Polynom $T_{x_0,0}(x)=f(x_0)$ heißt 0-tes Taylorpolynom.

Um festzustellen, wie gut das Taylorpolynom die Funktion in der Nähe des Punktes x_0 approximiert, müssen wir den Fehler oder das ***n-te Restglied***

$$R_{x_0,n}(x):=f(x)-T_{x_0,n}(x)$$

abschätzen. Die Taylorformel ist nichts anderes als die triviale Gleichheit $f(x)=T_{x_0,n}(x)+R_{x_0,n}(x)$ zusammen mit einer nützlichen Formel oder Abschätzung für das Restglied $R_{x_0,n}(x)$.

6.5.4. **Satz** (Taylorformel mit Integralrestglied)**.** *Es sei $I\subset\mathbb{R}$ ein Intervall, $n\in\mathbb{N}_0$, $f\in\mathscr{C}^{n+1}(I)$ und $x_0,x\in I$. Dann gilt*

$$(*)\qquad f(x)=T_{x_0,n}(x)+\int_{x_0}^{x}\frac{(x-t)^n}{n!}f^{(n+1)}(t)\,dt.$$

Diese Formel heißt ***Taylorformel mit Integralrestglied***.

6.5.5. **Satz** (Taylorformel)**.** *Es sei $I\subset\mathbb{R}$ ein Intervall, $n\in\mathbb{N}_0$, $f\in\mathscr{C}^{n+1}(I)$ und $x_0\in I$. Sei $p\in\mathbb{N}$ mit $1\leqslant p\leqslant n+1$. Dann gibt es zu jedem $x\in I\setminus\{x_0\}$ ein ξ zwischen x und x_0 mit*

$$R_{x_0,n}(x)=\frac{(x-\xi)^{n-p+1}(x-x_0)^p}{n!p}f^{(n+1)}(\xi)$$

(Schlömilch-Form des Restgliedes oder ***Restglied von Schlömilch****); es gilt also*:

$$(6.7)\qquad f(x)=T_{x_0,n}(x)+\frac{(x-\xi)^{n-p+1}(x-x_0)^p}{n!p}f^{(n+1)}(\xi)$$

(Taylorformel mit Schlömilch-Restglied).
Für $p=1$ ergibt sich die Cauchy-Form des Restgliedes oder ***Restglied von Cauchy*** :

$$(6.8)\qquad R_{x_0,n}(x)=\frac{(x-\xi)^n(x-x_0)}{n!}f^{(n+1)}(\xi).$$

Für $p = n+1$ ergibt sich die Lagrange-Form des Restgliedes oder ***Restglied von Lagrange*** :

$$R_{x_0,n}(x) = \frac{(x-x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi).$$

6.5.6. **Satz** (Qualitative Taylorformel oder Formel von Taylor-Young).
Es sei $I \subset \mathbb{R}$ ein Intervall, $n \in \mathbb{N}_0$, $f \in \mathscr{C}^n(I)$ und $x_0 \in I$. Dann ist $R_{x_0,n}(x) = o((x-x_0)^n)$ für $x \to x_0$, d.h.

$$f(x) = T_{x_0,n}(x) + o((x-x_0)^n) \quad \text{für } x \to x_0.$$

Mit Hilfe der Taylorformel können wir ein umfassendes Extremwertkriterium angeben:

6.5.7. **Satz** (dritten hinreichendes Extremwertkriterium). *Sei $f : I \longrightarrow \mathbb{R}$ n-mal differenzierbar in $x_0 \in I$. Es gelte $f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0$, aber $f^{(n)}(x_0) \neq 0$.*

(i) *Ist n ungerade, so hat f kein lokales Extremum in x_0.*

(ii) *Ist n gerade, dann gilt:* $\begin{cases} f^{(n)}(x_0) > 0 \Rightarrow x_0 \text{ ist lokales Minimum,} \\ f^{(n)}(x_0) < 0 \Rightarrow x_0 \text{ ist lokales Maximum.} \end{cases}$

Beispiel: Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = e^x + e^{-x} + 2\cos x$. Der Punkt $x_0 = 0$ ist ein kritischer Punkt für f, da $f'(x) = e^x - e^{-x} - 2\sin x$ in $x_0 = 0$ verschwindet. Ferner gilt:

$$f''(x) = e^x + e^{-x} - 2\cos x, \quad f''(0) = 0;$$
$$f'''(x) = e^x - e^{-x} + 2\sin x, \quad f'''(0) = 0;$$
$$f^{(4)}(x) = e^x + e^{-x} + 2\cos x, \quad f^{(4)}(0) = 4.$$

Da die erste Ableitung, welche nicht verschwindet, von gerader Ordnung ist, liegt ein Extremwert vor. Dieses ist ein Minimum wegen $f^{(4)}(0) > 0$. 30.01.2012

6.6. Uneigentliche Integrale.

6.6.1. **Definition.** Sei $-\infty \leqslant a < b \leqslant \infty$.
(i) Ist $a > -\infty$ und $f : [a,b) \longrightarrow \mathbb{C}$ Regelfunktion, so setzen wir

$$\int_a^b f(x)\,dx := \lim_{\beta \nearrow b} \int_a^\beta f(x)\,dx,$$

falls dieser Grenzwert existiert. In diesem Fall sagen wir, dass $\int_a^b f(x)\,dx$ ***konvergiert***; ein Integral dieser Form heißt ***uneigentliches Integral***.
(ii) Analog für $b < \infty$, $f : (a,b] \longrightarrow \mathbb{C}$ Regelfunktion:

$$\int_a^b f(x)\,dx := \lim_{\alpha \searrow a} \int_\alpha^b f(x)\,dx,$$

falls dieser Grenzwert existiert.
(iii) Ist $f : (a,b) \longrightarrow \mathbb{C}$ Regelfunktion, so wählen wir $c \in (a,b)$ und setzen

$$\int_a^b f(x)\,dx := \lim_{\alpha \searrow a} \int_\alpha^c f(x)\,dx + \lim_{\beta \nearrow b} \int_c^\beta f(x)\,dx,$$

falls diese beiden Grenzwerte existieren.
(iv) Konvergiert $\int_a^b |f(x)|\,dx$, so heißt $\int_a^b f(x)\,dx$ ***absolut konvergent***.

Beispiele:

(1) $\int_0^\infty e^{-x}\,dx = 1$, denn $\int_0^\beta e^{-x}\,dx = -e^{-x}\big|_0^\beta = 1 - e^{-\beta} \to 1$ für $\beta \to \infty$.

(2) $\int_0^\infty \sin x\,dx$ existiert nicht, denn $\int_0^\beta \sin x\,dx = -\cos x\big|_0^\beta = 1 - \cos\beta$ hat keinen Grenzwert für $\beta \to \infty$.

(3) $\int_1^\infty \frac{1}{x^s}\,dx$ konvergiert genau dann, wenn $\operatorname{Re} s > 1$. In der Tat:

$$\int_1^\infty \frac{1}{x^s}\,dx = \begin{cases} -\dfrac{1}{s-1} \cdot \dfrac{1}{x^s - 1}\Big|_1^\beta = \dfrac{1}{s-1} - \dfrac{\beta^{1-s}}{s-1}, s \neq 1, \\ \log x\big|_1^\beta = \log\beta, s = 1. \end{cases}$$

$\beta^{1-s} = \beta^{1-\mathrm{Re}\,s} \cdot e^{i(\log\beta)\mathrm{Im}\,s}$ konvergiert für $\beta \to \infty$ genau dann, wenn $\mathrm{Re}\,s > 1$. Es gilt also:

$$\boxed{\int_1^\infty \frac{1}{x^s}\,dx = \frac{1}{s-1} \quad \text{für } \mathrm{Re}\,s > 1}.$$

(4) $\displaystyle\int_0^1 \frac{1}{x^s}\,dx$ konvergiert genau dann, wenn $\mathrm{Re}\,s < 1$ und

$$\boxed{\int_0^1 \frac{1}{x^s}\,dx = \frac{1}{1-s} \quad \text{für } \mathrm{Re}\,s < 1}.$$

(5) $\displaystyle\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx = \pi$, denn $\displaystyle\int_0^\beta \frac{1}{1+x^2}\,dx = \arctan\beta \longrightarrow \frac{\pi}{2}$, $\beta \to \infty$ und $\displaystyle\int_\alpha^0 \frac{1}{1+x^2}\,dx = -\arctan\alpha \longrightarrow \frac{\pi}{2}$, $\alpha \to -\infty$.

Bemerkung:

(a) In (iii) ist die Existenz der Grenzwerte und der Wert der Summe unabhängig von der Wahl von c (Intevall-Additivität).

(b) Ist $f : [a,b] \longrightarrow \mathbb{C}$ eine Regelfunktion, so bedeutet $\displaystyle\int_a^b f(x)dx$ in der obigen Definition nichts Neues, denn z.B. $\lim_{\beta\to b} \displaystyle\int_a^b f(x)dx = \int_a^b f(x)dx$, weil $F : [a,b] \longrightarrow \mathbb{C}$, $F(x) = \displaystyle\int_a^x f(t)dt$ stetig ist.

6.6.2. **Satz** (Cauchy-Kriterium für Funktionengrenzwerte)**.** *Sei $f : D \longrightarrow \mathbb{C}$, und sei $z_0 \in \overline{\mathbb{C}}$ ein HP von D. Die Funktion f hat einen Grenzwert $\lim_{z\to z_0} f(z) \in \mathbb{C}$ genau dann, wenn es zu jedem $\varepsilon > 0$ eine Umgebung U von z_0 gibt, so dass für alle $z, w \in U \cap (D \setminus \{z_0\})$ gilt: $|f(z) - f(w)| < \varepsilon$.*

6.6.3. **Satz.** *Konvergiert $\displaystyle\int_a^b f(x)dx$ absolut, so konvergiert $\displaystyle\int_a^b f(x)dx$.*

Dies ist eine Anwendung des Cauchy-Kriteriums für Funktionsgrenzwerte wegen

$$\Big|\int_a^{\beta_1} f\,dx - \int_a^{\beta_2} f\,dx\Big| = \Big|\int_{\beta_1}^{\beta_2} f\,dx\Big| \leqslant \Big|\int_{\beta_1}^{\beta_2} |f|\,dx\Big| = \Big|\int_a^{\beta_1} |f|\,dx - \int_a^{\beta_2} |f|\,dx\Big|.$$

6.6.4. **Satz** (Majorantenkriterium für das uneigentliche Integral)**.** *Seien $f : [a,b) \longrightarrow \mathbb{C}$ und $g : [a,b) \longrightarrow \mathbb{R}$ Regelfunktionen mit $|f| \leqslant g$. Falls $\displaystyle\int_a^b g(x)dx$ konvergiert, so konvergiert $\displaystyle\int_a^b f(x)dx$ absolut.* (*Analog für $(a,b]$, (a,b)*).

Beispiel (6): Gaußsche Felherintegral (I). Wegen $e^{-x^2} \leqslant e^{-x}$ für $x \geqslant 1$ und Beispiel (1) implizert das Majorantenkriterium, dass das ***Gaußsche Fehlerintegral*** $\displaystyle\int_0^\infty e^{-x^2}\,dx$ konvergiert. Dieses Integral spielt eine bedeutende Rolle in der Wahrscheinlichkeitstheorie.

Beispiel (7): $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx$ konvergiert.

- $(0,1] \ni x \to \dfrac{\sin x}{x}$ ist stetig nach 0 fortsetzbar (mit Wert 1), also konvergiert $\displaystyle\int_0^1 \frac{\sin x}{x}\,dx$.
- $\displaystyle\int_1^\infty \frac{\sin x}{x}\,dx$ konvergiert: Es gilt

$$\int_1^\beta \frac{\sin x}{x}\,dx = \frac{-\cos x}{x}\Big|_1^\beta - \int_1^\beta \frac{\cos x}{x^2}\,dx = \frac{\cos\beta}{\beta} - \frac{\cos 1}{1} - \int_1^\beta \frac{\cos x}{x^2}\,dx.$$

Nun $\lim_{\beta\to\infty} \frac{\cos\beta}{\beta} = 0$ und $\displaystyle\int_1^\infty \frac{\cos x}{x^2}\,dx$ konvergiert, weil $\Big|\dfrac{\cos x}{x^2}\Big| \leqslant \dfrac{1}{x^2}$ und $\displaystyle\int_1^\infty \frac{1}{x^2}\,dx$ konvergiert.

02.02.2012 Bemerkung: $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx$ konvergiert nicht absolut!

6.6.5. **Satz** (Grenzwertkriterium)**.** *Seien $f : [a,b) \longrightarrow \mathbb{C}$ und $g : [a,b) \longrightarrow \mathbb{R}_+$ Regelfunktionen, und es existiere $\lim_{x\nearrow b} \frac{f(x)}{g(x)}$. Wenn $\displaystyle\int_a^b g(x)dx$ konvergiert, so konvergiert auch $\displaystyle\int_a^b f(x)dx$ absolut.*

Beispiel (8): Das Integral $\Gamma(s) = \displaystyle\int_0^\infty x^{s-1}e^{-x}\,dx$ konvergiert für $\mathrm{Re}\,s > 0$. Die Funktion

$$\boxed{\Gamma : \{s \in \mathbb{C} : \mathrm{Re}\,s > 0\} \to \mathbb{C}, \quad \Gamma(s) = \int_0^\infty x^{s-1}e^{-x}\,dx} \tag{6.9}$$

heißt ***Eulersche Gammafunktion***.

- $\int_0^1 x^{s-1}e^{-x}\,dx$ konvergiert: $|x^{s-1}e^{-x}| \leqslant x^{\mathrm{Re}\,s-1}$ und $\int_0^1 \frac{1}{x^{1-\mathrm{Re}\,s}}\,dx$ konvergiert für $\mathrm{Re}\,s > 0$.
- $\int_1^\infty x^{s-1}e^{-x}\,dx$ konvergiert: Wende das Grenzwertkriterium mit $g : [1,\infty) \to \mathbb{R}$, $g(x) := e^{-\frac{x}{2}}$ an; dann gilt $\frac{x^{s-1}e^{-x}}{e^{-x/2}} = x^{s-1}e^{-x/2} \longrightarrow 0$ für $x \to 0$. Außerdem $\int_1^\infty e^{-x/2}\,dx = -2e^{-x/2}\Big|_1^\infty = 2e^{-1/2}$.

6.6.6. **Satz.** *Die Γ-Funktion erfüllt die folgenden Eigenschaften:*

(a) $\Gamma(s+1) = s\Gamma(s)$,
(b) $\Gamma(1) = 1$,
(c) $\Gamma(n) = (n-1)!$.

6.6.7. **Satz** (Integralkriterium für Reihen-Konvergenz). *Sei $f : [m,\infty) \longrightarrow [0,\infty)$, $m \in \mathbb{N}$ eine monoton fallende Funktion. Dann gilt*

$$0 \leqslant \lim_{n\to\infty}\Big[\sum_{k=m}^{n} f(k) - \int_m^{n+1} f(x)dx\Big] \leqslant f(m).$$

Insbesondere konvergiert $\sum_{k\geqslant m} f(k)$ genau dann, wenn $\int_m^\infty f(x)dx$ konvergiert.

6.6.8. **Beispiel** (Abschätzung für die Zeta-Funktion). $f : [1,\infty) \longrightarrow \mathbb{R}$, $f(x) = \frac{1}{x^s}$, $s > 0$ ist monoton fallend. Nach dem Integralkriterium gilt:

$$\int_1^\infty \frac{1}{x^s}dx \quad \text{konvergiert genau dann, wenn} \quad \sum_{n\geqslant 1}\frac{1}{n^s} \quad \text{konvergiert}.$$

Wir wussten schon, dass beides genau dann zutrifft, wenn $s > 1$ ist (siehe Beispiel 3.2.6). Nun außerdem: Mit $\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}$ gilt $0 \leqslant \zeta(s) - \int_1^\infty \frac{1}{x^s}dx \leqslant f(1) = 1$. Daraus folgt

$$\zeta(s) \leqslant 1 + \frac{1}{s-1} = \frac{s}{s-1}.$$

Für $s = 1$ erhalten wir die Existenz des Limes

$$\lim_{n\to\infty}\Big(\sum_{k=1}^n \frac{1}{k} - \int_1^{n+1}\frac{1}{x}dx\Big) = \lim_{n\to\infty}\Big(\sum_{k=1}^n \frac{1}{k} - \log(n+1)\Big) = \lim_{n\to\infty}\Big(\sum_{k=1}^n \frac{1}{k} - \log n\Big) =: \gamma.$$

Dies bedeutet, dass die Partialsummen der harmonischen Reihe wie $\log n$ anwachsen. Die Zahl γ heißt ***Eulersche Konstante*** und beträgt approximativ $0{,}57721\ldots$. Es ist unbekannt, ob γ eine rationale oder eine irrationale Zahl ist.

6.7. Übungen.

6.7.1. ***Aufgabe***. Sei $b > 1$. Berechnen Sie $\int_1^b \frac{1}{x}dx$ aus der Definition des Integrals, *ohne* den Hauptsatz und Ihre Kenntnis einer Stammfunktion zu $f(x) = \frac{1}{x}$ vorauszusetzen. Benutzen Sie dazu die Unterteilungen $1 < b^{\frac{1}{n}} < b^{\frac{2}{n}} < \ldots < b^{\frac{n-1}{n}} < b$ des Intervalles $[1,b]$.

6.7.2. ***Aufgabe***. Zeigen Sie durch Betrachtung geeigneter Integrale $\int_a^b f(x)dx$:

(a) $\lim_{n\to\infty}\frac{1}{n^{k+1}}\left(1^k + 2^k + \ldots + n^k\right) = \frac{1}{k+1}$ für $k \in \mathbb{N}_0$
(b) $\lim_{n\to\infty}\left(\frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n}\right) = \log 2$

6.7.3. ***Aufgabe***. Sei $f \in \mathcal{R}([a,b])$

(i) Sei $\widetilde{f} : [a,b] \longrightarrow \mathbb{C}$ eine Funktion, die mit f außerhalb einer endlichen Menge von Punkten übereinstimmt. Zeigen Sie, dass $\widetilde{f} \in \mathcal{R}([a,b])$ gilt.
(ii) Sei $L \geqslant 0$, und sei $A \subset [a,b]$ eine endliche Menge so, dass $|f(x)| \leqslant L$ für alle $x \in [a,b] \smallsetminus A$. Zeigen Sie, dass $\Big|\int_a^b f\,dx\Big| \leqslant \int_a^b |f|\,dx \leqslant L(b-a)$ gilt.

6.7.4. ***Aufgabe***. Zeigen Sie: Falls es zu einer Regelfunktion $f : I \to \mathbb{R}$ eine Funktion $F : I \to \mathbb{R}$ mit $F' = f$ gibt, so ist f stetig (Tipp: Hauptsatz[6]).

[6]Ein anderer Beweis folgt aus den folgenden Aussagen: F' hat die Zwischenwerteigenschaft (Aufgabe 5.7.2), f hat nur Unstetigkeitsstellen erster Art (Satz 6.4.1), und eine Funktion mit der Zwischenwerteigenschaft kann keine Unstetigkeitsstellen erster Art haben (einfache Übung).

6.7.5. ***Aufgabe*** (***Riemannsches Lemma***).

(a) Zeigen Sie, dass $\lim\limits_{|y|\to\infty}\int_a^b \cos(yt)\,dt = \lim\limits_{|y|\to\infty}\int_a^b \sin(yt)\,dt = 0$ ist.

(b) Sei $f:[a,b]\to\mathbb{C}$ eine Treppenfunktion. Beweisen Sie

$$(*)\qquad \lim_{|y|\to\infty}\int_a^b f(t)\cos(yt)\,dt = \lim_{|y|\to\infty}\int_a^b f(t)\sin(yt)\,dt = 0\,.$$

(c) Zeigen Sie, dass $(*)$ auch für Regelfunktionen $f\in\mathcal{R}([a,b])$ gilt.

6.7.6. ***Aufgabe***. (a) Sei $\arcsin:[-1,1]\to\mathbb{R}$ die Umkehrfunktion zu $\sin|_{[-\frac{\pi}{2},\frac{\pi}{2}]}$. Finden Sie eine Stammfunktion zu arcsin (Tipp: $\arcsin x = (x)'\cdot\arcsin x$). Verifizieren Sie:
$\int_{\frac{1}{2}}^{1}\arcsin x\,dx = \frac{5\pi}{12}-\frac{\sqrt{3}}{2}$.

(b) Finden Sie eine Stammfunktion zu $\log^2:\mathbb{R}_+\to\mathbb{R}$ (Tipp: zweimal Tipp zu (a)). Verifizieren Sie: $\int_1^2\log^2(x)\,dx = 2\ln^2\left(\frac{e}{2}\right)$.

6.7.7. ***Aufgabe***. Beweisen Sie für $k,\ell\in\mathbb{Z}$ mit $|k|\neq|\ell|$:

$$\int_0^{2\pi}\sin(kx)\sin(\ell x)\,dx = 0\,,\quad \int_0^{2\pi}\cos(kx)\cos(\ell x)\,dx = 0\,,$$
$$\int_0^{2\pi}\sin(kx)\cos(\ell x)\,dx = 0\,.$$

6.7.8. ***Aufgabe***. Sei $f:[0,y]\to\mathbb{C}$ stetig. Beweisen Sie:

$$\int_0^y\Big(\int_0^{y_{n-1}}\Big(\dots\int_0^{y_2}\Big(\int_0^{y_1}f(x)\,dx\Big)dy_1\dots\Big)dy_{n-2}\Big)dy_{n-1} = \frac{1}{(n-1)!}\int_0^y (y-x)^{n-1}f(x)\,dx$$

für $n\in\mathbb{N}$. (Tipp: Induktion. Für $n=1$ ist die linke Seite als $\int_0^y f(x)\,dx$ zu verstehen. Fassen Sie im Induktionsschritt das *innerste* Integral als Funktion seiner oberen Grenze auf und wenden Sie die Induktionsvoraussetzung auf die restlichen Integrale an.)

6.7.9. ***Aufgabe***. (a) Seien $a,b\in\mathbb{R}\setminus\{0\}$. Bestimmen Sie $\int e^{ax}\cos bx\,dx$ auf zwei Weisen: (i) durch mehrmalige partielle Integration, (ii) durch Einsetzen der Definition: $\cos z = \frac{1}{2}(e^{iz}+e^{-iz})$.

(b) Bestimmen Sie $\int\cos(\log x)\,dx$ auf $\mathbb{R}_+$ auf zwei Weisen: (i) durch mehrmalige partielle Integration, (ii) mithilfe der Substitutionsregel und Teil (a).

6.7.10. ***Aufgabe***. (a) Bestimmen Sie auf $[-1,1]$ eine Stammfunktion zu $\sqrt{1-x^2}$ durch partielle Integration.

(b) Seien $a,b>0$. Berechnen Sie den Flächeninhalt der Ellipse, deren Rand durch die Gleichung $x^2/a^2+y^2/b^2=1$ gegeben ist. Betrachten Sie dazu z.B. den in der oberen Halbebene liegenden Teil der Randkurve als Graph einer Funktion.

6.7.11. ***Aufgabe***. Berechnen Sie mithilfe der Substitutionsregel die folgenden unbestimmten Integrale:

(a) $\displaystyle\int\frac{1}{x\log x}\,dx$ auf $(0,1)$ und $(1,\infty)$,

(b) $\displaystyle\int 9\sqrt{\frac{\operatorname{Arsinh}6x}{1+36x^2}}\,dx$,

(c) $\displaystyle\int\frac{1}{\sqrt{x}(1+x)}\,dx$ auf $\mathbb{R}_+$,

(d) $\displaystyle\int\frac{1}{a^2\cos^2x+b^2\sin^2x}\,dx$ mit $a,b\in\mathbb{R}\setminus\{0\}$ (Tipp: $y=\tan x$).

6.7.12. ***Aufgabe***. Sei $f\in\mathscr{C}^1([0,1])$.

(a) Beweisen Sie die Identität $f(x)=\int_0^1 f(u)\,du+\int_0^x u\,f'(u)\,du+\int_x^1(u-1)f'(u)\,du$, $x\in[0,1]$.

(b) Folgern Sie, dass für alle $x\in[0,1]$ gilt: $|f(x)|\leqslant\int_0^1\big(|f(u)|+|f'(u)|\big)\,du$.

6.7.13. ***Aufgabe***. (i) Zeigen Sie, dass die Dirichlet-Funktion (4.1) eine Regelfunktion ist.

(ii) Seien I, J Intervalle in $\mathbb{R}$; sei $f:I\to J$ eine Regelfunktion und $g:J\to\mathbb{C}$ stetig. Zeigen Sie, dass $g\circ f$ eine Regelfunktion ist.

6.7.14. ***Aufgabe***. Berechnen Sie $\displaystyle\int_0^1 f(x)\,dx$, wobei f die Dirichlet-Funktion (4.1) sei.

6.7.15. ***Aufgabe***. Es sei $f:[a,b]\longrightarrow\mathbb{R}$ eine Regelfunktion mit $f\geqslant 0$ und $\int_a^b f(x)dx=0$. Dann ist $f(x_0)=0$ an jeder Stetigkeitsstelle x_0 von f.

6.7.16. ***Aufgabe***. Zeigen Sie, dass $T_{x_0,n}$ das einzige Polynom P vom Grad höchstens n ist, das $f(x_0)=P(x_0)$, $f'(x_0)=P'(x_0),\ldots,f^{(n)}(x_0)=P^{(n)}(x_0)$ erfüllt.

6.7.17. ***Aufgabe***. Sei $I_n:=\int_0^{\pi/2}\cos^n t\,dt$ für $n\geqslant 0$, und sei $(w_n)_{n\geqslant 1}$ die Wallissche Folge (6.6). Zeigen Sie für alle $n\geqslant 2$:

$$\begin{aligned}
&I_{2n+1}^2=\frac{1}{2n+1}w_n\,,\quad I_{2n-2}^2=\frac{\pi^2}{4(2n-1)}\frac{1}{w_{n-1}}\\
&\int_0^1(1-x^2)^n\,dx=I_{2n+1}\\
&\int_0^\infty\frac{1}{(1+x^2)^n}\,dx=\int_0^{\pi/2}\sin^{2n-2}t\,dt=I_{2n-2}\,.
\end{aligned}$$

(Tip für das zweite Integral: Substitution $x=\cot t$ für ein eigentliches Integral und Grenzübergang.)

6.7.18. ***Aufgabe***. Sei $I:=\int_0^\infty e^{-x^2}\,dx$ das Gaußsche Fehlerintegral. Zeigen Sie für $n\geqslant 1$:

$$\int_0^\infty e^{-nx^2}\,dx=\frac{1}{\sqrt{n}}I\,,$$

$$\int_0^1(1-x^2)^n\,dx\leqslant\int_0^1 e^{-nx^2}\,dx\leqslant\int_0^\infty e^{-nx^2}\,dx\leqslant\int_0^\infty\frac{1}{(1+x^2)^n}\,dx\,,$$

Mit den Bezeichnungen der Aufgabe 6.7.17 zeige für alle $n\geqslant 2$:

$$\frac{n}{2n+1}w_n\leqslant I^2\leqslant\frac{n}{2n-1}\cdot\frac{1}{w_{n-1}}\cdot\frac{\pi^2}{4}\tag{6.10}$$

Folgere

$$\boxed{I=\int_0^\infty e^{-x^2}\,dx=\frac{\sqrt{\pi}}{2}}\tag{6.11}$$

6.7.19. ***Aufgabe*** (Geschwindigkeitsverteilung in einem idealen Gas). Zwischen dem quadratischen Mittelwert β aus den Geschwindigkeiten der Moleküle eines idealen Gases und ihrem linearen Mittelwert α besteht die Beziehung:

$$\beta=\sqrt{\frac{4\alpha^2}{\sqrt{\pi}}\int_0^\infty x^4e^{-x^2}\,dx}\,.$$

Zeigen Sie, dass $\beta=\alpha\sqrt{\frac{3}{2}}$.

6.7.20. ***Aufgabe***. (a) Zeigen Sie, dass $\int_0^\infty\frac{\sin x}{x}\,dx=\lim_{n\to\infty}\int_0^{\pi/2}\frac{\sin(2n+1)t}{t}\,dt$ ist.

(b) Zeigen Sie, dass die Funktion $f:[0,\pi/2]\to\mathbb{R}$ mit $f(t)=\frac{1}{t}-\frac{1}{\sin t}$ für $t\neq 0$ und $f(0)=0$ stetig ist.

(c) Wenden Sie das Riemannsche Lemma auf f aus (b) an und leiten Sie her, dass

$$\int_0^\infty\frac{\sin x}{x}\,dx=\lim_{n\to\infty}\int_0^{\pi/2}\frac{\sin(2n+1)t}{\sin t}\,dt\,.$$

(d) Zeigen Sie, dass für $n\in\mathbb{N}$ gilt: $\int_0^{\pi/2}\frac{\sin(2n+1)t}{\sin t}\,dt=\int_0^{\pi/2}\frac{\sin(2n-1)t}{\sin t}\,dt\,.$

(e) Folgern Sie, dass

$$\int_0^\infty\frac{\sin x}{x}\,dx=\frac{\pi}{2}\,.$$

6.7.21. ***Aufgabe***. Zeigen Sie mit Hilfe der Regel von l'Hospital:

$$\lim_{n\to\infty}n^s(\sqrt[n]{n}-1)=\begin{cases}0, & \text{falls } s<1,\\ \infty, & \text{falls } s\geqslant 1\,.\end{cases}$$

6.8. **Notizen: Der Riemannsche Integralbegriff.** Das Integral, das wir in 6.2 eingeführt haben, heißt ***Regelintegral***. Es ist einfacher als das gebräuchlichere ***Riemannsche Integral***. Das Regelintegral genügt aber für praktische Belange. Für wichtige theoretische Belange brauchen wir das kompliziertere Lebesgue-Integral (siehe Analysis III). Diese Integralbegriffe unterscheiden sich nur durch die Klassen der jeweiligen integrierbaren Funktion; auf dem Durchschnitt dieser Klassen stimmen die Integralbegriffe überein.

6.8.1. **Definition.** Sei $Z = \{a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b\}$ eine Zerlegung von $[a,b]$. Sei $\xi_k \in [x_{k-1}, x_k]$ für $k = 1,\ldots,n$.

(i) Für $f : [a,b] \longrightarrow \mathbb{C}$ heißt $\sum_{k=1}^n f(\xi_k)(x_k - x_{k-1})$ die ***Riemannsche Summe*** zu f bezüglich der Zerlegung Z und der Stützstellen $\xi_1,\ldots,\xi_k$.

(ii) Die Feinheit von Z ist definiert als $\max\{|x_k - x_{k-1}| : k = 1,\ldots,n\}$.

(iii) Eine Funktion $f : [a,b] \longrightarrow \mathbb{C}$ heißt ***Riemann-integrierbar***, wenn es eine Zahl $I \in \mathbb{C}$ mit der folgenden Eigenschaft gibt: Zu jedem $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für jede Zerlegung $Z = \{x_0,\ldots,x_n\}$ mit Feinheit $< \delta$ und für jede Wahl von Stützstellen $\xi_k \in [x_{k-1}, x_k]$ $(k = 1,\ldots,n)$ gilt:

$$\Big|\sum_{k=1}^n f(\xi_k)(x_k - x_{k-1}) - I\Big| < \varepsilon .$$

In diesem Fall heißt $I := R\text{-}\int_a^b f(x)\,dx$ das ***Riemann-Integral*** von f.

6.8.2. **Bemerkung.**

(i) Die Riemannschen Summen sind eigentlich Integrale von TF: Ist eine Zerlegung von $[a,b]$ wie oben gegeben, so setze $\varphi : [a,b] \longrightarrow \mathbb{C}$, $\varphi(x) = f(\xi_k)$ für $x \in (x_{k-1}, x_k)$ und definiere φ in den Teilpunkten x_k beliebig. Dann ist $\int_a^b \varphi\,dx = \sum_{k=1}^n f(\xi_k)(x_k - x_{k-1})$ eine Riemannsche Summe.

(ii) Ist $(Z_n)_{n\geqslant 1}$ eine Folge von Zerlegungen, deren Feinheit gegen Null strebt, und ist S_n eine Riemannsche Summe für f zur Zerlegung Z_n, so gilt $\lim_{n\to\infty} S_n = R\text{-}\int_a^b f\,dx$. Ist φ_n wie in (i) eine TF mit $\int_a^b \varphi_n\,dx = S_n$, so gilt $\lim_{n\to\infty} \int_a^b \varphi_n\,dx = R\text{-}\int_a^b f\,dx$.

(iii) In der Definition der Riemann-Integrierbarkeit verlangen wir nicht, dass die Folge (φ_n) die Funktion f geichmäßig approximiert, sondern dass für die φ_n aus (i) die Integrale $(\int_a^b \varphi_n\,dx)$ stets gegen denselben Wert konvergieren, sofern die Feinheit der Zerlegungen gegen Null strebt.

6.8.3. **Satz.** *Sei $f \in \mathcal{R}([a,b])$. Dann ist f auch Riemann-integrierbar, und*

$$R\text{-}\int_a^b f\,dx = \int_a^b f\,dx .$$

Insbesondere konvergieren die Riemannschen Summen zu einer Zerlegungsfolge mit gegen Null konvergierender Feinheit gegen das Regelintegral von f.

Als Anwendung von Satz 6.8.3 ergibt sich ein sehr kurzer Beweis der Hölderschen und Minkowskischen Ungleichungen für Integrale.

6.8.4. **Definition.** Sei $f \in \mathcal{R}([a,b])$ und $p \geq 1$. Die p-Norm von f ist definiert als

$$\|f\|_p := \Big(\int_a^b |f|^p\,dx\Big)^{1/p} .$$

6.8.5. **Satz** (Höldersche und Minkowskische Ungleichungen)**.** *Seien $f,g \in \mathcal{R}([a,b])$ und $p,q \in \mathbb{R}_+$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt:*

$$\int_a^b |fg|\,dx \leqslant \|f\|_p \|g\|_q , \tag{6.12}$$

$$\|f+g\|_p \leqslant \|f\|_p + \|g\|_p . \tag{6.13}$$

Für $p = q = 2$ heißt (6.12) *die Cauchy-Schwarz-Ungleichung für Integrale.*

6.8.6. **Bemerkung.** Die Menge der Riemann-integrierbaren Funktionen ist echt größer als die Menge der Regelfunktionen. Zum Beispiel ist die folgende Funktion Riemann-integrierbar, aber keine Regelfunktion: $f : [0,1] \to \mathbb{R}$ mit $f(x) = \cos(1/x)$ für $x \neq 0$ und $f(0) = 0$.

$$\boxed{\{\text{Regelfunktionen}\} \subsetneqq \{\text{Riemann-integriebare Funktionen}\}}$$

7. Gleichmässige Konvergenz

7.1. **Motivation und Definition.** Sei $D \subset \mathbb{C}$. In Definition 4.2.1 haben wir definiert: Eine Folge von Funktionen $(f_n)_{n\in\mathbb{N}}$, $f_n : D \longrightarrow \mathbb{C}$ ***konvergiert punktweise*** gegen $f : D \longrightarrow \mathbb{C}$, wenn $\lim_{n\to\infty} f_n(z) = f(z)$ für jedes $z \in D$ gilt.

Fragen:

(1) Wenn alle f_n stetig bzw. differenzierbar bzw. Regelfunktionen sind, gilt dies dann auch für die Grenzfunktion f ?

(2) Darf man Grenzwertprozesse vertauschen? Gelten Gleichungen wie:

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} (\lim_{n\to\infty} f_n(x)) = \lim_{n\to\infty} (\lim_{x\to x_0} f_n(x)) \tag{7.1}$$

$$\frac{d}{dx} f(x) = \frac{d}{dx}(\lim f_n(x)) = \lim_{n\to\infty} \frac{df_n}{dx} \tag{7.2}$$

$$\int_a^b f\,dx = \int_a^b \lim_{n\to\infty} f_n(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx? \tag{7.3}$$

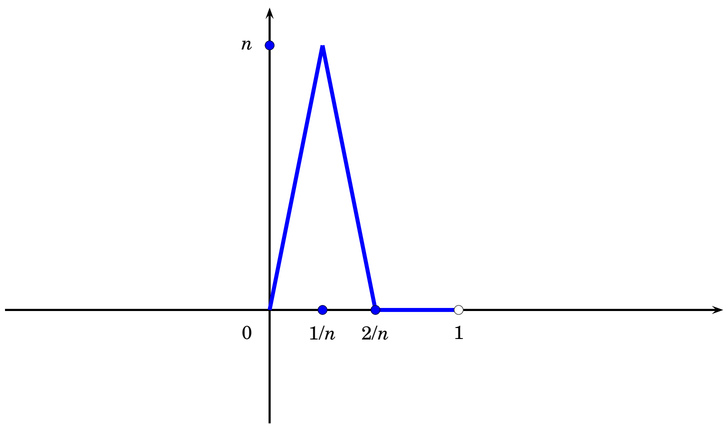
7.1.1. **Beispiele.** (1) Seien $f_n : [0,1] \longrightarrow \mathbb{R}$, $f_n(x) = x^n$. Dann gilt

$$\lim_{n\to\infty} f_n(x) = f(x) = \begin{cases} 0\,, x \in [0,1) \\ 1\,, x = 1\,. \end{cases}$$

Die Grenzfunktion ist nicht stetig! Die Gleichung (7.1) gilt für $x_0 = 1$ nicht.

(2) Seien $f_n : [0,1] \longrightarrow \mathbb{R}$, $f_n(x) = \frac{1}{n}\sqrt[n]{x + \frac{1}{n^n}}$. Für alle $x \in [0,1]$ gilt $0 \leqslant f_n(x) \leqslant f_n(1) < \frac{2}{n}$, also $\lim_{n\to\infty} f_n(x) = 0 =: f(x)$ (sogar $\|f_n - f\| \longrightarrow 0$). Es gilt $f_n'(x) = \frac{1}{n} \cdot \frac{1}{n}(x + \frac{1}{n^n})^{\frac{1}{n}-1}$. Daraus folgt $f_n'(0) = \frac{1}{n^2}\frac{1}{n^{1-n}} = \frac{n^{n-1}}{n^2} = n^{n-3} \longrightarrow \infty$. Also $f'(0) \neq \lim_{n\to\infty} f_n'(0)$ und (7.2) stimmt nicht.

(3) Seien $f_n : [0,1] \longrightarrow \mathbb{R}$,



$\lim_{n\to\infty} f_n(x) = 0 =: f(x)$, aber $\lim_{n\to\infty} \int_0^1 f_n\,dx = 1 \neq 0 = \int_0^1 f\,dx$ also (7.3) stimmt nicht..

7.1.2. **Definition.** Eine Folge $(f_n)_n$ von Funktionen $f_n : D \longrightarrow \mathbb{C}$ ***konvergiert gleichmäßig*** gegen $f : D \longrightarrow \mathbb{C}$, wenn $\|f_n - f\|_D \longrightarrow 0$ für $n \to \infty$, d.h.

$$\forall \varepsilon > 0 \quad \exists n_0 = n_0(\varepsilon) \quad \forall n \geq n_0(\varepsilon) \quad \forall x \in D : |f_n(x) - f(x)| < \varepsilon\,.$$

Eine Reihe $\sum_{n\geqslant 1} f_n$ konvergiert gleichmäßig, wenn die Folge der Partialsummen gleichmäßig konvergiert.

7.1.3. **Bemerkung.** (i) Vergleichen wir die Definitionen der gleichmäßigen und punktweisen Konvergenz. Es gilt $f_n \to f$ punktweise in D wenn

$$\forall x \in D \quad \forall \varepsilon > 0 \quad \exists n_0 = n_0(x,\varepsilon) \quad \forall n \geq n_0(x,\varepsilon) : |f_n(x) - f(x)| < \varepsilon\,.$$

Der Unterschied ist die Position des Quantors $\forall x \in D$; deshalb hängt der Index $n_0(x,\varepsilon)$ von x ab, währendessen $n_0(\varepsilon)$ in der Definition 7.1.2 hängt nur von ε ab, nicht von $x \in D$, d. h. $n_0(\varepsilon)$ und daher die Konvergenz ist "gleichmäßig in $x \in D$".

(ii) In obigen Beispielen gilt: (1) $\|f_n - f\| = 1$, (2) $\|f_n - f\| \to 0$, (3) $\|f_n - f\| = n$.
(iii) In der Definition der Regelfunktion konvergiert (φ_n) gleichmäßig gegen f.

Für Funktionenreihen haben wir in Definition 4.2.3 auch den Begriff *normale Konvergenz* eingeführt. Dies hat eine wichtige Rolle beim Beweis der Stetigkeit der Potenzreihen gespielt, darunter Exponentialfunktion (siehe Folgerung 4.2.8).

7.1.4. **Satz.**
(a) Konvergiert eine Folge oder Reihe gleichmäßig, so konvergiert sie auch punktweise.
(b) Konvergiert eine Reihe normal, so konvergiert sie auch gleichmäßig.

Es gilt also für eine Funktionenreihe $\sum f_n$:

$$\boxed{\text{Normale Konvergenz} \Rightarrow \text{Gleichmäßige Konvergenz} \Rightarrow \text{Punktweise Konvergenz}}$$

aber

$$\boxed{\text{Punktweise Konvergenz} \not\Rightarrow \text{Gleichmäßige Konvergenz} \not\Rightarrow \text{Normale Konvergenz}}$$

Siehe auch Beispiel 7.3.4.

7.2. **Vertauschungssätze.**

7.2.1. **Satz.** *Konvergiert eine Folge $(f_n)_n$ stetiger Funktionen $f_n : D \longrightarrow \mathbb{C}$ gleichmäßig gegen $f : D \longrightarrow \mathbb{C}$, so ist auch f stetig.*

Daraus folgt: Ist $\sum_{n\geqslant 0} f_n$ gleichmäßig konvergent und sind f_n stetig, so ist $f := \sum_{n=0}^{\infty} f_n$ stetig. Wir haben in Satz 4.2.6 dasselbe für normal konvergente Reihen bewiesen (dies war eine stärkere Voraussetzung). Wegen Satz 7.1.4, folgt Satz 4.2.6 aus Satz 7.2.1.

Satz 7.2.1 besagt auch, dass (7.1) gilt für jeder Häufungspunkt $x_0 \in D$ von D.

7.2.2. **Satz** (Vertauschung von Grenzwert und Integration)**.** *Konvergiert eine Folge $(f_n)_n$ von Regelfunktionen $f_n : [a,b] \longrightarrow \mathbb{R}$ gleichmäßig gegen $f : [a,b] \longrightarrow \mathbb{C}$, so ist auch f eine Regelfunktion, und es gilt*

$$\int_a^b f\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx\,.$$

Beispiel (2) in §7.1 zeigt, dass diese Aussage nicht stimmt, wenn die f_n nur punktweise gegen f konvergieren. Die Folge $(f_n)_n$ ist hier nicht gleichmäßig beschränkt. Der Satz über dominierte Konvergenz in der Lebesgue-Theorie des Integrals besagt, dass $\int_a^b f\,dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$ aber schon dann gilt, wenn $f_n \to f$ punktweise konvergiert und zusätzlich eine Lebesgue-integrierbare Funktion g (Regelfunktionen sind z.B. Lebesgue-integrierbar) existiert mit $|f_n| \leqslant g$ für alle n.

7.2.3. **Satz** (Vertauschung von Grenzwert und Differentiation)**.** *Sei $f_n : [a,b] \longrightarrow \mathbb{C}$, $n \in \mathbb{N}$ mit den Eigenschaften:*

(1) $\lim_{n\to\infty} f_n = f$ *punktweise,*
(2) f_n *stetig differenzierbar,*
(3) $(f_n')_n$ *konvergiert gleichmäßig auf* $[a,b]$.

Dann ist f auch differenzierbar, und es gilt $f' = \lim_{n\to\infty} f_n'$.

Die wesentliche Voraussetzung hier ist (3), wie das Beispiel 7.1.1 (2) zeigt: Dort konvergiert $f_n \to f$ gleichmäßig, jedoch nicht $f_n' \to f'$.

Wir können sogar eine ***stetige, nirgends differenzierbare Funktion*** als gleichmäßigen Limes von stetigen Funktionen konstruieren, siehe Übung 7.5.6. Für mehr dazu siehe [24, S. 353, 359].

7.3. **Potenzreihen und analytische Funktionen.** Sei $\sum_{n\geqslant 0} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$, wobei $a_n \in \mathbb{C}$ und $x \in \mathbb{R}$. Wir wissen, dass $\sum_{n\geqslant 0} a_n x^n$ konvergiert in $(-R,R)$ und divergiert auf $\mathbb{R} \setminus [-R,R]$. Wir wissen auch (Satz 4.2.5), dass die Potenzreihe konvergiert normal (also auch gleichmäßig wegen Satz 7.1.4 (b)) auf $[-r,r]$, für alle $0 < r < R$.

7.3.1. **Lemma.** *Hat $\sum_{n\geqslant 0} a_n x^n$ den Konvergenzradius R, so haben die durch gliedweise Differentiation bzw. Integration entstehenden Potenzreihen $\sum_{n\geqslant 0} n a_n x^{n-1}$ und $\sum_{n\geqslant 0} \frac{1}{n+1} a_n x^{n+1}$ den Konvergenzradius R.*

7.3.2. **Satz** (Vertauschbarkeit von Integration und Summation bei Potenzreihen). *Seien $\sum_{n\geqslant 0} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R>0$, und es seien $a,b\in(-R,R)$. Dann ist die Summe $P:(-R,R)\longrightarrow\mathbb{C}$, $P(x)=\sum_{n=0}^{\infty}a_nx^n$ stetig, und es gilt*

$$\int_a^b P(x)\,dx=\int_a^b\Big(\sum_{n=0}^{\infty}a_nx^n\Big)dx=\sum_{n=0}^{\infty}\int_a^b a_nx^n\,dx=\sum_{n=0}^{\infty}a_n\frac{b^{n+1}-a^{n+1}}{n+1},$$

d.h. die Potenzreihe kann gliedweise integriert werden.

Es gibt $0<r<R$, so dass $[a,b]\subset[-r,r]$, also konvergiert $\sum_{n\geqslant 0}a_nx^n$ gleichmäßig auf $[a,b]$. Der Satz 7.2.2 liefert dann die Behauptung.

7.3.3. **Beispiel** (Logarithmische Reihe). Wenn wir die geometrische Reihe

$$\frac{1}{1+t}=\frac{1}{1-(-t)}=\sum_{n=0}^{\infty}(-t)^n$$

über $[0,x]$ oder $[x,0]$ integrieren (für $x\in(-1,1)$), bekommen wir die Summe der Logarithmus-Reihe (siehe (3.9)):

$$\int_0^x\frac{1}{1+t}\,dt=\int_0^x\sum_{n=0}^{\infty}(-t)^n=\sum_{n=0}^{\infty}\int_0^x(-t)^n\,dt=\sum_{n=0}^{\infty}\frac{(-1)^n}{n+1}x^{n+1}$$

also

$$\log(1+x)=\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}x^n,\quad\text{für } x\in(-1,1).\tag{7.4}$$

Die Logarithmus-Reihe wurde von Nicolaus Mercator[7] 1668 entdeckt. Die Logarithmusreihe divergiert für $x>1$, obwohl die Logarithmusfunktion dort definiert ist. Für $x=1$ ist die Logarithmusreihe noch konvergent (alternierende harmonische Reihe); es ist aber keineswegs selbstverständlich, dass sie auch dort die Logarithmusfunktion darstellt. Dass dies doch der Fall ist, besagt die folgende Formel für die Summe der alternierenden harmonischen Reihe 12.04.2012

$$\log 2=\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}=1-\frac12+\frac13-\frac14+\frac15\mp\dots.\tag{7.5}$$

Zum Beweis beachten wir, dass für $x\in[0;1)$ die Logarithmusreihe die Hypothesen des Leibniz-Kriteriums 3.2.7 erfüllt (alternierende Reihe mit $\frac{x^n}{n}$ monoton fallende Nullfolge-Übung). Nach der Fehlerabschätzung des Leibniz-Kriteriums

$$\Big|\log(1+x)-\sum_{k=1}^{n}\frac{(-1)^{k-1}}{k}x^k\Big|\leqslant\frac{x^{n+1}}{n+1}\tag{7.6}$$

gilt. Wegen der Stetigkeit der angeschriebenen Funktionen im Punkt $x=1$ gilt diese Abschätzung auch noch in $x=1$:

$$\Big|\log 2-\sum_{k=1}^{n}\frac{(-1)^{k-1}}{k}\Big|\leqslant\frac{1}{n+1}.\tag{7.7}$$

Daraus folgt mit $n\to\infty$ die Behauptung.

7.3.4. **Beispiel** (Normale versus gleichmäßige Konvergenz). Die Umkehrung von Satz 7.1.4 (b) gilt im Allgemeinen nicht; die Logarithmus-Reihe $\sum_{n\geqslant1}\frac{(-1)^{n-1}x^n}{n}$ konvergiert nicht normal auf $[0,1]$, da

$$\sum_{n\geqslant1}\Big\|\frac{(-1)^{n-1}x^n}{n}\Big\|_{[0,1]}=\sum_{n\geqslant1}\frac1n,$$

konvergiert aber gleichmäßig, da

$$\Big\|\log(1+x)-\sum_{k=1}^{n}\frac{(-1)^{k-1}x^k}{k}\Big\|_{[0,1]}\leqslant\frac{1}{n+1}$$

nach der Restgliedabschätzung im Leibniz-Kriterium.

Der folgende, auf Abel zurückgehende Satz beantwortet die allgemeine Frage, ob eine im Randpunkt der Konvergenzbereich konvergente Potenzreihe, auch noch im Randpunkt dieselbe Funktion darstellt wie im Inneren des Konvergenzbereichs.

[7]Deutscher Mathematiker, Mitglied der Royal Society, entwarf die Springbrunnen von Versailles.

7.3.5. **Satz** (Abelscher Grenzwertsatz [1]). *Die Potenzreihe $\sum_{n\geqslant 0} a_n x^n$ habe den Konvergenzradius $R > 0$. Ist die Reihe auch für $x = R$ konvergent, so ist sie im Intervall $[0,R]$ gleichmäßig Konvergent. Die Summe $P : [0,R] \to \mathbb{C}$, $P(x) = \sum_{n=0}^{\infty} a_n x^n$ ist dann wegen Satz 7.2.1 im Punkt $x = R$ linksseitig stetig,*

$$\lim_{x\nearrow R} P(x) = P(R) = \sum_{n=0}^{\infty} a_n R^n .$$

Entsprechendes gilt, wenn die Reihe für $x = -R$ konvergiert (man betrachte dazu die Reihe für $P(-x)$).

7.3.6. **Satz** (Vertauschbarkeit von Differentiation und Summation bei Potenzreihen).
Sei $\sum_{n\geqslant 0} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann ist die Summe $P : (-R,R) \longrightarrow \mathbb{C}$, $P(x) = \sum_{n=0}^{\infty} a_n x^n$ differenzierbar, und es gilt $P'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$ für alle $x \in (-R,R)$, d.h. die Potenzreihe kann gliedweise differenziert werden.

Betrachte die gliedweise differenzierte Reihe $\sum_{n\geqslant 1} n a_n x^{n-1}$. Sie hat den Konvergenzradius R nach Lemma 7.3.1. Wenden wir den Satz 7.3.2 auf $\sum_{n\geqslant 1} n a_n x^{n-1}$ an, so erhalten wir

$$\int_0^x \Big(\sum_{n=1}^{\infty} n a_n t^{n-1}\Big) dt = \sum_{n=1}^{\infty} a_n x^n = P(x), \quad x \in (-R,R).$$

Der Hauptsatz liefert nun die Behauptung.

7.3.7. **Folgerung.** *Hat $\sum_{n\geqslant 0} a_n x^n$ den Konvergenzradius $R > 0$, so ist $P : (-R,R) \longrightarrow \mathbb{C}$, $P(x) = \sum_{n=0}^{\infty} a_n x^n$ unendlich oft differenzierbar in $(-R,R)$. Es gilt*

$$P^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1) a_n x^{n-k} = \sum_{n=k}^{\infty} k! \binom{n}{k} a_n x^{n-k}, \quad x \in (-R,R),\, k \in \mathbb{N}_0 \tag{7.8}$$

insbesondere

$$\frac{P^{(k)}(0)}{k!} = a_k \quad (\text{Taylorsche Koeffizientenformel}). \tag{7.9}$$

7.3.8. **Folgerung** (Identitätssatz für Potenzreihen). *Seien $\sum_{n\geqslant 0} a_n x^n$ und $\sum_{n\geqslant 0} b_n x^n$ Potenzreihen mit positiven Konvergenzradius. Es gelte $\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n$ für x in einer Umgebung von 0 in $\mathbb{R}$. Dann ist $a_n = b_n$ für alle n.*

Sei R_1 bzw. R_2 der Konvergenzradius der Reihe $\sum_{n\geqslant 0} a_n x^n$ bzw. $\sum_{n\geqslant 0} b_n x^n$ und sei $P_1 : (-R_1,R_1) \to \mathbb{C}$, $P_1(x) = \sum_{n=0}^{\infty} a_n x^n$ bzw. $P_2 : (-R_2,R_2) \to \mathbb{C}$, $P_2(x) = \sum_{n=0}^{\infty} b_n x^n$. Dann gilt $P_1^{(k)}(0) = P_2^{(k)}(0)$ für alle $k \in \mathbb{N}_0$ und nach (7.9) folgt $a_k = b_k$.

7.3.9. **Definition.** (i) Seien $I \subset \mathbb{R}$ ein Intervall, $f : I \longrightarrow \mathbb{C}$ unendlich oft differenzierbar und $x_0 \in I$. Die Potenzreihe

$$T_{x_0,\infty} = \sum_{k\geqslant 0} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

heißt ***Taylorreihe*** zu f in x_0. Beachte: Die n-te Partialsumme von $T_{x_0,\infty}$ ist das n-te Taylorpolynom $T_{x_0,n}$.
(ii) Wir sagen, dass die Funktion f auf $U = (x_0 - r, x_0 + r) \cap I$ durch ihre ***Taylorreihe darstellbar*** ist, wenn $T_{x_0,\infty}(x) = f(x)$ für alle $x \in U$ gilt. Wir sagen auch, dass f *besitzt in U eine Taylorentwicklung mit x_0 als Entwicklungspunkt*.
(iii) Wir sagen, dass f ***analytisch*** in x_0 ist, wenn es ein $r > 0$ gibt, so dass f auf $(x_0 - r, x_0 + r) \cap I$ durch ihre Taylorreihe darstellbar ist. Ist f analytisch in allen $x_0 \in I$, so heißt f ***analytisch***.

7.3.10. **Beispiele.** (1) Polynome sind durch ihre Taylorreihe in x_0 auf ganz $\mathbb{R}$ darstellbar, für alle $x_0 \in \mathbb{R}$.
(2) Die Exponentialfunktion ist durch ihre Taylorreihe in x_0 auf ganz $\mathbb{R}$ darstellbar, für alle $x_0 \in \mathbb{R}$. Insbesondere ist sie analytisch auf $\mathbb{R}$. Sei $x_0 \in \mathbb{R}$ fest. Dann gilt $f^{(n)}(x_0) = e^{x_0}$ für alle $n \geqslant 0$, und

$$T_{x_0,\infty}(x) = \sum_{k=0}^{\infty} \frac{e^{x_0}}{k!}(x-x_0)^k = e^{x_0} \sum_{k=0}^{\infty} \frac{(x-x_0)^k}{k!} = e^{x_0} \cdot e^{x-x_0} = e^x$$

für alle $x \in \mathbb{R}$.
(3) Sei $\sum_{n\geqslant 0} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann gilt nach (7.9)

$$P(x) = \sum_{n=0}^{\infty} a_n x^n = \sum_{k\geqslant 0} \frac{P^{(k)}(0)}{k!} x^k = T_{0,\infty}(x), \quad x \in (-R,R),$$

d. h. eine Potenzreihe ist die Taylorreihe ihrer Summe. Dies zeigt, dass die Summe einer Potenzreihe ist durch ihre Taylorreihe (in 0) auf dem Konvergenzbereich darstellbar und sie ist analytisch in $x = 0$. Wir können mehr zeigen: Die Summe einer Potenzreihe besitzt eine Taylorentwicklung mit Entwicklungspunkt x_0, für alle x_0 im Inneren des Konvergenzbereich. Sie ist also analytisch im Konvergenzbereich.

7.3.11. **Umentwicklungssatz.** *Die Potenzreihe* $\sum_{n\geqslant 0} a_n x^n$ *habe den Konvergenzradius* $R > 0$ *und bezeichne ihre Summe durch* $P : (-R,R) \longrightarrow \mathbb{C}$, $P(x) = \sum_{n=0}^{\infty} a_n x^n$. *Sei* $x_0 \in (-R,R)$. *Setze* $b_k = \frac{1}{k!}P^{(k)}(x_0) = \sum_{n=k}^{\infty} \binom{n}{k} a_n x_0^{n-k}$. *Dann hat* $\sum_{n\geqslant 0} b_n y^n$ *einen Konvergenzradius* $\geqslant R - |x_0|$ *und es gilt:*

$$P(x) = \sum_{n=0}^{\infty} b_n (x - x_0)^n\,, \quad \textit{für } |x - x_0| < R - |x_0|.$$

(4) Aus Beispiel 7.3.3 folgt: Der Logarithmus $\log : \mathbb{R}_+ \to \mathbb{R}$ ist durch seine Taylorreihe in $x_0 = 1$ auf $(-1,1]$ darstellbar.

7.3.12. **Bemerkung** (Glatte versus Analytische Funktionen)**.** Seien $I \subset \mathbb{R}$ ein Intervall, $f : I \longrightarrow \mathbb{C}$ unendlich oft differenzierbar und $x_0 \in I$.

1. Frage: Ist der Konvergenzradius von $T_{x_0,\infty}$ positiv? Die Antwort ist i.A. nein. Der folgende Satz zeigt, dass jede Potenzreihe (also auch jede Potenzreihe mit Konvergenzradius Null) als Taylor-Reihe realisiert werden kann.

7.3.13. **Satz** (Borel)**.** *Sei* $(a_n)_{n\geqslant 0}$ *eine beliebige Folge in* $\mathbb{C}$. *Dann gibt es* $f \in \mathcal{C}^{\infty}(I,\mathbb{C})$ *mit* $f^{(n)}(x_0) = a_n$ *für alle* $n \geqslant 0$.

Wir suchen eine beliebig oft differenzierbare Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ so, dass $f^{(n)}(0) = a_n$ für alle n. Wir wählen eine Funktion mit kompaktem Träger $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ so, dass $\phi = 1$ auf $[-1,1]$ und $\phi = 0$ auf $\mathbb{R}\setminus(-2,2)$ gilt. Wir setzen nun $f_n(x) = \dfrac{a_n}{n!} x^n \phi(r_n x)$, wobei die r_n positive Zahlen sind. Der Faktor $\phi(r_n x)$ produziert die Konvergenz. Die r_n können nämlich so (groß) gewählt werden, dass $|f_n^{(m)}| < 2^{-m}$ auf $\mathbb{R}$ für $m = 0,1,\dots,n-1$ gilt. (Beachte dazu: Was ist der Träger von $\phi(r_n x)$?) Die Folge $f(x) := \sum_{n=0}^{\infty} f_n(x)$ konvergiert und ist beliebig oft differenzierbar (gliedweise differenzieren). Wir können die Ableitungen $f_n^{(m)}(0)$ explizit berechnen (welche Werte hat $\phi(r_n x)$ in einer Umgebung von Null?). Wir sehen sofort, dass $f^{(m)}(0) = a_m$ ist.

2. Frage: Falls der Konvergenzradius von $T_{x_0,\infty}$ positiv ist, gibt es dann $r > 0$, so dass

$$T_{x_0,\infty}(x) = \sum_{k=0}^{\infty} \frac{f^k(x_0)}{k!}(x - x_0)^k = f(x)$$

für alle $x \in (x_0 - r, x_0 + r) \cap I$ ist? D. h. ist f durch ihre Taylorreihe darstellbar? Für ein $x \in I$ gilt offensichtlich $T_{x_0,\infty}(x) = f(x)$ genau dann, wenn $\lim_{n\to\infty} R_n(x) = 0$ ist. Die Antwort ist auch hier negativ. Ein Beispiel ist die Funktion (5.7). Für diese gilt $f^{(n)}(0) = 0$ für alle $n \geqslant 0$, ihre Taylor-Reihe ist also $T_{0,\infty}(x) = 0 \neq e^{-1/x} = f(x)$ für alle $x > 0$. In keiner Umgebung von Null stimmen diese Funktion und ihre Taylor-Reihe überein. Diese Funktion ist also glatt aber nicht analytisch.

16.04.2012

7.4. **Restgliedabschätzungen für die logarithmische und binomische Reihe.** Wir wollen nun den Fehler durch die Approximation der Partialsummen der logarithmischen und binomischen Reihen abschätzen.

7.4.1. **Satz.** *Für* $x \in (-1,1)$ *gilt*

$$\log(1+x) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} x^k + R_n\,, \quad |R_n| < \frac{|x|^{n+1}}{n+1} \cdot \frac{1}{1-|x|}\,. \tag{7.10}$$

In der Tat,

$$|R_n| = \Big|\sum_{k=n+1}^{\infty} \frac{(-1)^{k-1}}{k} x^k\Big| \leqslant \sum_{k=n+1}^{\infty} \frac{|x|^k}{k} < \frac{|x|^{n+1}}{n+1} \sum_{k=0}^{\infty} |x|^k = \frac{|x|^{n+1}}{n+1} \cdot \frac{1}{1-|x|}\,.$$

7.4.2. **Bemerkung.** Für $x \in (0,1)$ gilt die bessere Abschätzung $0 < R_n \leqslant \frac{x^{n+1}}{n+1}$ (siehe (7.6)), da $\frac{1}{1-x} > 1$.

7.4.3. **Folgerung.** *Für* $x \in (-1,1)$ *gilt*

$$\log\frac{1+x}{1-x} = 2\sum_{m=0}^{\infty} \frac{x^{2m+1}}{2m+1} = 2\sum_{m=0}^{n} \frac{x^{2m+1}}{2m+1} + R_n\,, \quad |R_n| < 2\frac{|x|^{2n+3}}{2n+3} \cdot \frac{1}{1-|x|^2}\,. \tag{7.11}$$

In der Tat,

$$|R_n| \leqslant 2 \sum_{m=n+1}^{\infty} \frac{|x|^{2m+1}}{2m+1} < \frac{|x|^{2n+3}}{2n+3} \sum_{k=0}^{\infty} |x|^{2k} = 2\frac{|x|^{2n+3}}{2n+3} \cdot \frac{1}{1-|x|^2}.$$

Für $x = \frac{1}{3}$ erhalten wir

$$\log 2 = 2 \sum_{m=0}^{\infty} \frac{1}{(2m+1)3^{2m+1}} = 2 \sum_{m=0}^{n} \frac{1}{(2m+1)3^{2m+1}} + R_n\,, \quad |R_n| < \frac{1}{3^{2n+1}} \cdot \frac{1}{4(2n+3)}.$$

Diese Reihe konvergiert viel schneller als die Reihe (7.5) und außerdem ist die Restgliedabschätzung viel besser als (7.7).

7.4.4. **Satz.** *Sei $s \in \mathbb{C}$. Für $x \in (-1,1)$ gilt*

$$(1+x)^s = \sum_{k=0}^{\infty} \binom{s}{k} x^k = \sum_{k=0}^{n} \binom{s}{k} x^k + R_n\,, \tag{7.12}$$

wobei

$$|R_n| \leqslant (n+1)\binom{s}{n+1}|x|^{n+1} \cdot \begin{cases} (1+|x|)^{\operatorname{Re} s-1}, & \operatorname{Re} s \geqslant 1, \\ (1-|x|)^{\operatorname{Re} s-1}, & \operatorname{Re} s \leqslant 1. \end{cases}$$

Diese Formel wurde 1669 von der 26järigen Newton entdeckt [17] (De analysis per aequationes numero terminorum infinitas, Band II, S. 206-247) für $s \in \mathbb{R}$. Abel [1] hat die Reihe für $s \in \mathbb{C}$ betrachtet. Aus (7.12) erhalten wir für $x \in (-1,1)$:

$$\begin{aligned}
\frac{1}{1+x} &= (1+x)^{-1} = 1 - x + x^2 - x^3 + \ldots + (-1)^n x^n + o(x^n), \\
\sqrt{1+x} &= (1+x)^{\frac{1}{2}} = 1 + \frac{1}{2}x - \frac{1}{2\cdot 4}x^2 + \frac{1\cdot 3}{2\cdot 4\cdot 6}x^3 + \ldots + (-1)^n \frac{1\cdot 3\cdots(2n-3)}{2\cdot 4\cdots(2n)}x^n + o(x^n), \\
\frac{1}{\sqrt{1+x}} &= (1+x)^{-\frac{1}{2}} = 1 + \frac{1}{2}x - \frac{1\cdot 3}{2\cdot 4}x^2 + \frac{1\cdot 3}{2\cdot 4\cdot 6}x^3 + \ldots + (-1)^n \frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots(2n)}x^n + o(x^n).
\end{aligned}$$

7.4.5. **Beispiel.** In der klassischen Mechanik ist die kinetische Energie eines Punktes der Masse m gegeben durch $T = \frac{1}{2}mv^2$, wobei v die Geschwindigkeit des Punktes ist. In der relativistischen Physik wird die (relativistische) Masse eines Punktes mit Geschwindigkeit v gegeben durch

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}, \quad m_0 = \text{Ruhemasse Masse}, \ c = \text{Lichtgeschwindigkeit}.$$

Die kinetische Energie wird als Differenz der Gesamtenergie und Ruheenergie gegeben

$$T = mc^2 - m_0c^2 = m_0c^2\left(\left(1-\frac{v^2}{c^2}\right)^{-\frac{1}{2}} - 1\right) = m_0c^2\left[\frac{1}{2}\left(\frac{v}{c}\right)^2 + \frac{3}{8}\left(\frac{v}{c}\right)^4 + \ldots\right]$$

also $T = \frac{1}{2}mv^2 + O\big((\frac{v}{c})^4\big)$. Für v sehr klein im Vergleich zu c, d. h. für $\frac{v}{c}$ sehr klein (Schreibweise $v \ll c$) kann man die höhere Ordnung Glieder $O\big((\frac{v}{c})^4\big)$ vernachlässigen und man erhält wieder die Formel der klassischen Mechanik $T = \frac{1}{2}mv^2$.

7.5. **Übungen.**

7.5.1. ***Aufgabe***. Beweise:

(a) $f_n : [0,1] \to \mathbb{R}$ mit $f_n(x) = x(1-x)^n$ konvergiert gleichmäßig gegen die Nullfunktion.

(b) $g_n := n \cdot f_n : [0,1] \to \mathbb{R}$ konvergiert punktweise, aber nicht gleichmäßig gegen die Nullfunktion. (Tipp: $g_n(\frac{1}{n}) = ?$)

Sei $h_n : \mathbb{R} \to \mathbb{R}$ definiert durch $h_n(x) := nxe^{-nx^2}$.

(c) Konvergieren die h_n punktweise? Falls ja, gegen welche Funktion?

(d) Auf welchen abgeschlossenen Intervallen $I \subset \mathbb{R}$ konvergieren die h_n gleichmäßig? (Tip: Funktionendiskussion)

7.5.2. ***Aufgabe***. Sei $D \subset \mathbb{C}$, und seien $f_n : D \to \mathbb{C}$, $g_n : D \to \mathbb{C}$ zwei Folgen von Funktionen, die gleichmäßig gegen $f : D \to \mathbb{C}$ bzw. $g : D \to \mathbb{C}$ konvergieren. Zeige:

(a) Die Funktionenfolge $f_n + g_n$ konvergiert gleichmäßig gegen $f + g$.

(b) Falls die f_n, g_n beschränkte Funktionen sind, so konvergiert die Funktionenfolge $f_n \cdot g_n$ gleichmäßig gegen $f \cdot g$. (Tip: Es gibt eine *gemeinsame* Schranke für alle $|f_n|$, warum?)

7.5.3. ***Aufgabe***. Zeige, daß für $x \in (-1,1)$ gilt:

(a) $\arctan x = x - \dfrac{x^3}{3} + \dfrac{x^5}{5} - \dfrac{x^7}{7} \pm \dots$

(b) $\arcsin x = x + \dfrac{1}{2}\cdot\dfrac{x^3}{3} + \dfrac{1\cdot 3}{2\cdot 4}\cdot\dfrac{x^5}{5} + \dfrac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\dfrac{x^7}{7} + \dots$

Anleitung: Entwickle jeweils zunächst die *Ableitung* der genannten Funktionen in eine Potenzreihe (man braucht dazu übrigens nicht auf die Definition der Taylorreihe zurückzugreifen).

7.5.4. ***Aufgabe***. Zeigen Sie:

(a) $\boxed{\dfrac{\pi}{4} = 1 - \dfrac{1}{3} + \dfrac{1}{5} - \dfrac{1}{7} \pm \dots}$ (Leibnizsche Reihe für π)

(b) $\boxed{\dfrac{\pi}{2} = 1 + \dfrac{1}{2}\cdot\dfrac{1}{3} + \dfrac{1\cdot 3}{2\cdot 4}\cdot\dfrac{1}{5} + \dfrac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\dfrac{1}{7} + \dots}$

Anleitung: (a) Das Leibniz-Kriterium liefert die Abschätzung $|\arctan x - \sum_{k=0}^{n} \frac{(-1)^k}{2k+1} x^{2k+1}| \leqslant \frac{|x^{2k+3}|}{2k+3}$ für alle $x \in (-1,1)$. (b) Sind $s_n(x)$ die Partialsummen der Reihe, so gilt $s_n(x) < \arcsin x < \arcsin 1$ für alle $x \in (0,1)$. Was passiert für $x \to 1$?

Mehr über die Berechnung von π in [13, § 8.11].

7.5.5. ***Aufgabe*** (Blatt 1,3). Sei $s \in \mathbb{R}$, und sei $f : (-1,\infty) \to \mathbb{R}$ durch $f(x) = (1+x)^s$ definiert.

(a) Zeigen Sie: Die Taylorreihe zu f in 0 ist die Binomialreihe $B_s(x) = \sum_{k=0}^{\infty} \binom{s}{k} x^k$.

(b) Zeigen Sie, dass f zumindest auf $\left(-\frac{1}{2}, 1\right)$ durch diese Taylorreihe dargestellt wird. (Tipp: Lagrange-Restglied; überlegen Sie, dass $\binom{s}{n}\lambda^n \to 0$ für $n \to \infty$ und $|\lambda| < 1$.)

Zusatz: Zeigen Sie, dass (b) auch auf $(-1,1)$ gilt. (Tipp: Cauchy-Restglied)

Es gilt also die folgende Verallgemeinerung der Binomialformel:

$$\boxed{(1+x)^s = \sum_{k=0}^{\infty} \binom{s}{k} x^k \quad \text{für } x \in (-1,1) \text{ und } s \in \mathbb{R}} \tag{7.13}$$

Diese grundlegende Formel wurde von Newton mit 24 Jahren entdeckt, und erst ein Jahrhundert später um 1774 von Euler bewiesen.

7.5.6. ***Aufgabe***. (a) Sei $f : I \to \mathbb{R}$ in $x_0 \in I$ differenzierbar. Seien $a_n \to x_0$ und $b_n \to x_0$ Folgen in I mit $a_n \leqslant x_0 \leqslant b_n$ und $a_n < b_n$ für alle n. Zeigen Sie: $d_n := \frac{f(b_n)-f(a_n)}{b_n - a_n} \to f'(x_0)$ für $n \to \infty$.
(Tipp: Mit $f(x) = f(x_0) + r(x)(x - x_0)$ liegt d_n zwischen $r(a_n)$ und $r(b_n)$, warum?)

(b) Für jedes $n \in \mathbb{N}$ definieren wir eine Funktion $f_n : [0,1] \to \mathbb{R}$ wie folgt: Ist $1 \leqslant m \leqslant 2^n$ und $x \in \left[\frac{m-1}{2^n}, \frac{m}{2^n}\right]$, so sei $f_n(x) := x - \frac{m-1}{2^n}$ für ungerades m und $f_n(x) := \frac{m}{2^n} - x$ für gerades m. Man überzeugt sich leicht, dass f_n wohldefiniert und stetig ist.

(i) Skizzieren Sie f_1, f_2, f_3, f_4 und $\sum_{k=1}^{4} f_k$.

(ii) Zeigen Sie, dass $f := \sum_{k=1}^{\infty} f_k : [0,1] \to \mathbb{R}$ existiert und stetig ist.

(iii) Zu jedem $n \in \mathbb{N}$ sei $[a_n, b_n]$ ein Intervall der Form $\left[\frac{m-1}{2^n}, \frac{m}{2^n}\right]$ (wie oben), das x_0 enthält. Zeigen Sie: $\dfrac{f(b_n) - f(a_n)}{b_n - a_n}$ ist für gerades n gerade und für ungerades n ungerade, und f ist in x_0 nicht differenzierbar.

Die Funktion f ist also stetig, aber nirgends differenzierbar.

8. Metrische und topologische Räume

Wir haben bisher die Konvergenz von Folgen in $\mathbb{R}$ und $\mathbb{C}$, die Stetigkeit und Differenzierbarkeit von Funktionen $f : I \longrightarrow \mathbb{C}$ ($I \subset \mathbb{R}$ Intervall) untersucht. Wir möchten nun die Analysis in Vektorräumen, insbesondere in $\mathbb{R}^n$ behandeln. Wir entwickeln dafür eine geometrische Sprache mit universellen Begriffen, die es uns ermöglicht:

- volle Allgemeinheit zu erreichen,
- einfache Beweise zu geben.

Die eingeführten Begriffe mögen sehr abstrakt erscheinen, sie haben aber eine sehr anschauliche Bedeutung, wenn man sie auf den zweidimensionalen Raum $\mathbb{R}^2$ oder den dreidimensionalen Raum $\mathbb{R}^3$ spezialisiert.

In einem metrischen Raum definieren wir die offenen Teilmengen; ihre Gesamtheit heißt Topologie. Begriffe wie Konvergenz oder Stetigkeit lassen sich allein mit dem Offenheitsbegriff ohne weiteren Rückgriff auf die Metrik definieren. Daher spielt die Metrik eigentlich nur eine Hilfsrolle, da die Topologie durch äquivalente Metriken erzeugt werden kann. Die Metrik oder die Norm sind dennoch sehr wichtige Mittel; je nach Problem kann man durch Wahl einer geeigneten Metrik oder Norm sehr rasch über die Konvergenz einer Folge, Stetigkeit einer Abbildung usw. entscheiden.

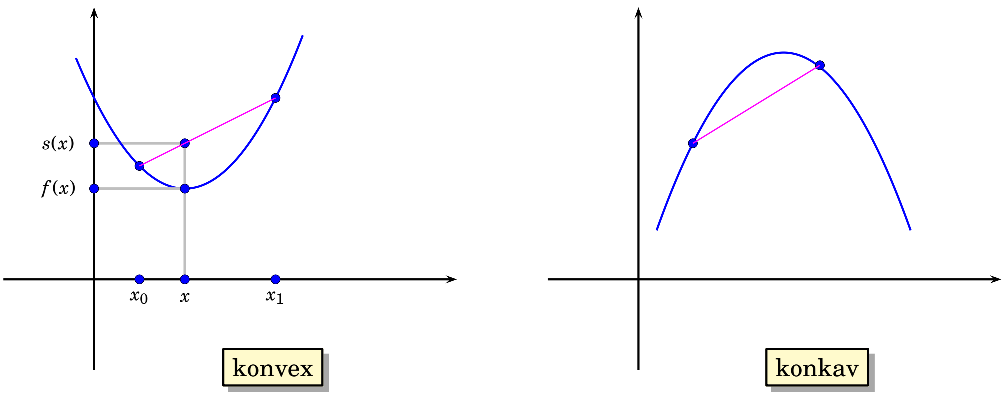
Die wichtigsten Begriffe dieses Kapitels sind die *Vollständigkeit*, die *Kompaktheit* und der *Zusammenhang*. Sie werden später sehr häufig benutzt. Es stellt sich heraus, dass die Kompaktheit und der Zusammenhang *topologische* Eigenschaften eines Raumes sind, d.h. sie hängen nur von der induzierten Topologie eines metrischen Raumes ab.

Einige Literaturhinweise für dieses Kapitel (abgesehen von den Analysisbüchern [13,24]) sind [5,11, 20]. Für Gegenbeispiele in der Topologie siehe [21].

8.1. Konvexität und wichtige Ungleichungen.

Die Konvexität ist eine bedeutende Eigenschaft der reellen Funktionen und liefert sehr nützliche Ungleichungen, die wir später brauchen.

Geometrische Definition: Sei $I \subset \mathbb{R}$ ein Intervall. Eine Funktion $f : I \longrightarrow \mathbb{R}$ heißt konvex (bzw. konkav), wenn jede Sekante durch zwei Punkte $(x_0, f(x_0))$, $(x_1, f(x_1))$ oberhalb (bzw. unterhalb) des Graphen liegt. Dies bedeutet, dass dass die Menge der Punkte oberhalb (bzw. unterhalb) des Graphen, eine konvexe Menge ist.



Die Gleichung der Sekante ist

$$y = s(x) := \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0) + f(x_0) = \frac{x_1 - x}{x_1 - x_0} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1).$$

Daher die folgende analytische Definition:

8.1.1. **Definition.** Eine Funktion $f : I \longrightarrow \mathbb{R}$ heißt ***konvex***, wenn für alle Tripel x_0, x, x_1 aus I mit $x_0 < x < x_1$ gilt:

$$f(x) \leqslant \frac{x_1 - x}{x_1 - x_0} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1). \tag{8.1}$$

f heißt ***streng konvex***, falls für alle solchen Tripel die Ungleichung (8.1) streng ist.
f heißt ***konkav***, wenn $-f$ konvex ist (d.h. (8.1) gilt mit $\geqslant$). f heißt ***streng konkav***, wenn $-f$ streng konvex ist (d.h. (8.1) gilt mit $>$).

8.1.2. **Bemerkung.** Seien $x_0 < x_1$. Die Abbildung[8] $\varphi : (0,1) \to (x_0, x_1)$, $\varphi(\lambda) = (1-\lambda)x_0 + \lambda x_1$ ist eine Bijektion mit Inverse $\varphi^{-1}(x) = \frac{x-x_0}{x_1-x_0}$, d.h. $\varphi(\lambda) = x \Leftrightarrow \lambda = \frac{x-x_0}{x_1-x_0}$.

Eine Funktion $f : I \longrightarrow \mathbb{R}$ ist konvex genau dann, wenn für alle $x_0, x_1 \in I$ mit $x_0 < x_1$ und alle $\lambda \in (0,1)$ gilt:

$$f((1-\lambda)x_0 + \lambda x_1) \leqslant (1-\lambda)f(x_0) + \lambda f(x_1).$$

8.1.3. **Lemma.** *Sei $f : I \longrightarrow \mathbb{R}$. Die folgenden Aussagen sind äquivalent:*

(1) *f ist konvex.*

(2) *Für alle $x_0 < x < x_1$ in I gilt:*

$$\frac{f(x)-f(x_0)}{x-x_0} \leqslant \frac{f(x_1)-f(x)}{x_1-x}.$$

(3) *Für alle $x_0 < x < x_1$ in I gilt:*

$$\frac{f(x)-f(x_0)}{x-x_0} \leqslant \frac{f(x_1)-f(x_0)}{x_1-x_0} \leqslant \frac{f(x_1)-f(x)}{x_1-x}.$$

8.1.4. **Satz** (Konvexitätskriterium I). *Sei $f : [a,b] \longrightarrow \mathbb{R}$ stetig und auf (a,b) differenzierbar. Dann gilt:*

(1) *f ist konvex $\Longleftrightarrow$ f' ist monoton wachsend auf (a,b).,*

(2) *f' ist streng monoton wachsend auf (a,b) $\Longleftrightarrow$ f ist streng konvex.*

8.1.5. **Satz** (Konvexitätskriterium II). *Sei $f : [a,b] \longrightarrow \mathbb{R}$ stetig und auf (a,b) zweimal differenzierbar. Dann gilt:*

(1) *f konvex $\Longleftrightarrow f'' \geqslant 0$,*

(2) *$f'' > 0 \Longrightarrow f$ streng konvex.*

8.1.6. **Definition.** Sei $f : I \longrightarrow \mathbb{R}$, und sei $x_0 \in I$ ein innerer Punkt. Die Funktion f hat in x_0 einen ***Wendepunkt***, wenn f stetig in x_0 ist und $(\alpha, \beta) \subset I$ existiert, so dass f konvex auf (α, x_0) und konkav auf (x_0, β) ist oder umgekehrt.

8.1.7. **Satz.** *Sei $f : (a,b) \longrightarrow \mathbb{R}$, und sei $x_0 \in (a,b)$.*

(i) *Ist f zweimal differenzierbar und existiert $(\alpha,\beta) \subset I$ mit $f'' \geqslant 0$ auf (α, x_0) und $f'' \leqslant 0$ auf (x_0, β) (oder umgekehrt), dann ist x_0 ein Wendepunkt.*

(ii) *Ist f 3-mal stetig differenzierbar und gilt $f''(x_0) = 0$ und $f'''(x_0) \neq 0$, dann ist x_0 ein Wendepunkt.*

8.1.8. **Satz (*Jensensche Ungleichung*).** *Eine Funktion $f : I \longrightarrow \mathbb{R}$ ist konvex genau dann, wenn für alle $n \in \mathbb{N}$, alle $x_1, x_2, \ldots, x_n \in I$ und alle $\lambda_1, \lambda_2, \ldots, \lambda_n \geqslant 0$ mit $\sum_{k=1}^n \lambda_k = 1$ gilt:*

$$f\Big(\sum_{k=1}^n \lambda_k x_k\Big) \leqslant \sum_{k=1}^n \lambda_k f(x_k).$$

8.1.9. **Bemerkung.** Ist f streng konvex, $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$ und $\sum_{k=1}^n \lambda_k = 1$, so gilt in der Ungleichung $f(\sum_{k=1}^n \lambda_k x_k) \leqslant \sum_{k=1}^n \lambda_k f(x_k)$ die Gleichheit genau dann, wenn $x_1 = \ldots = x_n$.

8.1.10. **Satz (*AGM-Ungleichung*).** *Seien $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$ mit $\sum_{k=1}^n \lambda_k = 1$. Dann gilt für alle $x_1, \ldots, x_n > 0$:*

$$x_1^{\lambda_1} \cdots x_n^{\lambda_n} \leqslant \lambda_1 x_1 + \ldots + \lambda_n x_n.$$

Gleichheit gilt genau dann, wenn $x_1 = \ldots = x_n$.

Für $\lambda_1 = \ldots = \lambda_n = \frac{1}{n}$ erhalten wir die klassische AGM-Ungleichung (1.8).

8.1.11. **Folgerung (*Youngsche Ungleichung*).** *Seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt für alle $x, y > 0$:*

$$x^{1/p} y^{1/q} \leqslant \frac{x}{p} + \frac{y}{q}. \tag{8.2}$$

Gleichheit gilt genau dann, wenn $x = y$.

Man kann die Youngsche Ungleichung auch folgendermaßen formulieren:

$$xy \leqslant \frac{x^p}{p} + \frac{y^q}{q}, \quad \text{für alle } x, y \geqslant 0 \tag{8.3}$$

mit Gleichheit, genau dann, wenn $x^p = y^q$.

[8]Wir nennen φ Parametrisierung.

8.1.12. **Definition.** Sei $z = (z_1, \dots, z_n) \in \mathbb{C}^n$, $p \geqslant 1$. Dann heißt

$$\|z\|_p := (|z_1|^p + \ldots + |z_n|^p)^{1/p}$$

die ***p*-Norm** von z. Für $p = 1$ also: $\|z\|_1 = \sum_{j=1}^n |z_j|$; für $p = 2$: $\|z\|_2 = (\sum_{j=1}^n |z_j|^2)^{1/2}$. Letzteres ist die ***euklidische Norm*** von z.

8.1.13. **Definition.** Sei V ein K-VR ($K = \mathbb{R}$ oder $\mathbb{C}$). Eine Abbildung $V \longrightarrow \mathbb{R}$, $v \mapsto \|v\|$ heißt ***Norm***, falls gilt:

(N_1) $\|v\| \geqslant 0$ für alle $v \in V$, und aus $\|v\| = 0$ folgt $v = 0$,
(N_2) $\|\lambda v\| = |\lambda| \|v\|$ für alle $\lambda \in K$, $v \in V$,
(N_3) $\|v + w\| \leqslant \|v\| + \|w\|$ für alle $v, w \in V$. (Dreiecksungleichung)

Unser Ziel ist es nun, zu zeigen, dass $\|\cdot\|_p$ eine Norm ist. Dazu brauchen wir die Höldersche Ungleichung.

8.1.14. **Satz (*Höldersche Ungleichung*).** *Seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt für alle $z, w \in \mathbb{C}^n$:*

$$\sum_{k=1}^n |z_k w_k| \leqslant \|z\|_p \|w\|_q .$$

Für $z, w \in \mathbb{C}^n$ setze $\langle z, w\rangle := \sum_{k=1}^n z_k \overline{w}_k$. Wir erhalten aus Satz 8.1.14 für $p = q = 2$ die ***Cauchy-Schwarz-Ungleichung***:

$$|\langle z, w\rangle| \leqslant \|z\|_2 \cdot \|w\|_2$$

mit Gleichheit genau dann, wenn z und w linear abhängig sind.

Es ist nämlich $|\langle z, w\rangle| \leqslant \sum_{k=1}^n |z_k \overline{w}_k| = \sum_{k=1}^n |z_k w_k| \leqslant \|z\|_2 \|w\|_2$. Die Behauptung über die Gleichheit folgt aus der Identität von Lagrange (1.12)[9].

8.1.15. **Satz (*Minkowski-Ungleichung*).** *Sei $p \geqslant 1$, und seien $z, w \in \mathbb{C}^n$. Dann gilt:*

$$\|z + w\|_p \leqslant \|z\|_p + \|w\|_p .$$

Insbesondere ist $z \mapsto \|z\|_p$ eine Norm.

8.2. **Metrische und normierte Räume.** Erinnern wir uns an die Definition der Konvergenz einer Folge $(z_n)_n$ in $\mathbb{C}$ gegen $z \in \mathbb{C}$: $\forall \varepsilon > 0\ \exists n_\varepsilon$ so, dass $\forall n \geqslant n_\varepsilon : |z_n - z| < \varepsilon$.

Im Wesentlichen benutzen wir hier nur den Abstand $|z_n - z|$ von z_n nach z. Wir definieren nun einen allgemeineren Abstandsbegriff, sogenannte Metriken. Deren Merkmale sind ganz analog zu den Merkmalen des Abstandes in $\mathbb{C}$.

8.2.1. **Definition.** Sei X eine nichtleere Menge. Eine ***Metrik*** auf X ist eine Abbildung $d : X \times X \longrightarrow \mathbb{R}$, für die gilt:

(M1) $d(x, y) \geqslant 0$ für alle $x, y \in X$; $d(x, y) = 0 \Leftrightarrow x = y$ (Positivität)
(M2) $d(x, y) = d(y, x)$ für alle $x, y \in X$ (Symmetrie)
(M3) $d(x, z) \leqslant d(x, y) + d(y, z)$ für alle $x, y, z \in X$ (Dreiecksungleichung)

Das Paar (X, d) heißt ein ***metrischer Raum***. Die Elemente $x \in X$ werden auch ***Punkte*** genannt. Die Zahl $d(x, y)$ heißt ***Abstand*** (bezgl. d) von x und y.

Die Dreiecksungleichung drückt aus, dass der „direkte Weg“ von x nach z kürzer ist als der „Umweg“ von x nach z über y.

Eine oft nützliche modifizierte Dreiecksungleichung ist: $|d(x, y) - d(y, z)| \leqslant d(x, z)$ für alle $x, y, z \in X$. Dies folgt sofort aus der Dreiecksungleichung.

8.2.2. **Beispiel.** **(1)** $X = \mathbb{R}$ oder $X = \mathbb{C}$, $d(x, y) = |x - y|$.

(2) $X = \mathbb{R}^n$ oder $X = \mathbb{C}^n$, $d_p(x, y) := \|x - y\|_p = (\sum |x_i - y_i|^p)^{1/p}$ für ein $p \geqslant 1$ (***ℓ^p-Metrik***). (Dreiecksungleichung: $\|x - z\|_p = \|x - y + y - z\|_p \leqslant \|x - y\|_p + \|y - z\|_p$ für $p \geqslant 1$ wegen der Minkowski-Ungleichung.) Für $p = 2$ heißt d_2 die ***euklidische Metrik*** auf $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$). Das ist das wichtigste Beispiel einer Metrik in Analysis II.

(3) $X = \mathbb{R}^n$, $\mathbb{C}^n$; $d_\infty(x, y) := \|x - y\|_\infty := \max\{|x_1 - y_1|, \dots, |x_n - y_n|\}$ (***ℓ^∞-Metrik***). Zur Rechtfertigung der

[9]In (1.12) wurde die Lagrange-Identität für reelle Zahlen formuliert, sie gilt aber offensichtlich auf für komplexe Zahlen.

Bezeichnung d_∞ beweisen Sie $\lim_{p\to\infty} d_p(x,y) = d_\infty(x,y)$.

(4) Für eine Menge $D \neq \emptyset$ betrachte $\mathcal{B}(D) := \{f : D \longrightarrow \mathbb{C} : \|f\|_D < \infty\}$, die Menge der beschränkten Funktionen auf D mit Werten in $\mathbb{C}$. Wir setzen $d_D(f,g) = \|f-g\|_D$; d_D heißt ***Supremumsmetrik***. Andere Bezeichnungen: $d_D = d_{sup} = d_\infty$.
Sei $D = [a,b] \in \mathbb{R}$ und sei $\mathcal{C}^0([a,b]) = \{f : [a,b] \to \mathbb{C} : f \text{ stetig}\}$. Jede stetige Funktion auf $[a,b]$ ist beschänkt, also $\mathcal{C}^0([a,b]) \subset \mathcal{B}([a,b])$. Dann ist auch $(\mathcal{C}^0([a,b]), d_{[a,b]})$ ein metrischer Raum.

(5) $X = \mathbb{R}^2$, $d(x,y) = \begin{cases} d_2(x,y), & \text{falls } x,y \text{ auf Gerade durch } 0 \\ d_2(x,0)+d_2(y,0), & \text{sonst} \end{cases}$
heißt ***Metrik der französischen Eisenbahn***.

(6) Für X beliebig heißt $d(x,y) = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases}$ die ***diskrete Metrik***.

(7) Eine große Klasse von Beispielen liefern die normierten Vektorräume (Definition 8.1.13). Sei $(V, \|\cdot\|)$ ein NVR. Dann heißt $d(x,y) := \|x-y\|$ die ***Norm-Metrik***. Normierte Vektorräume sind also spezielle metrische Räume. In den obigen Beispielen galt:

(i) d gehört zur Norm $|\cdot|$ auf $\mathbb{R}$ bzw. $\mathbb{C}$
(ii) d_p gehört zur Norm $\|\cdot\|_p$ auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$
(iii) d_∞ gehört zur Norm $\|\cdot\|_\infty$ auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$
(iv) d_D gehört zur Supremumsnorm auf $\mathcal{B}(D)$, $d_{[a,b]}$ gehört zur Supremumsnorm auf $\mathcal{C}^0([a,b])$.

Alle Begriffe, die wir für metrische Räume erklären, sind auch für normierte Vektorräume erklärt. Wenn wir in normierten Vektorräumen von Konvergenz, Stetigkeit, offenen Mengen etc. sprechen, beziehen wir uns immer auf die Norm-Metrik.

Nicht jede Metrik auf einem Vektorraum kommt von einer Norm. Die Metriken in (5),(6) rühren von keiner Norm her.

Wir geben noch zwei Beispiele von normierten Räumen.

(8) Sei $p \geqslant 1$; $l^p := \{(x_n)_{n\geqslant 0} : \sum_{n\geqslant 0} |x_n|^p \text{ konvergent}\}$ ist ein NVR mit der Norm:

$$\|(x_n)_n\|_p := \Big(\sum_{n=0}^{\infty} |x_n|^p\Big)^{1/p}.$$

Die Dreiecksungleichung folgt durch Grenzübergang in der Minkowski–Ungleichung. $\|\cdot\|_p$ heißt auch ***ℓ^p-Norm***. Die zugehörige Metrik wird mit d_p bezeichnet.

(9) Sei $X = \mathcal{C}^0([a,b])$ und $p \geqslant 1$. Dann heißt

$$\|f\|_p := \Big(\int_a^b |f(x)|^p\,dx\Big)^{1/p}$$

die ***L^p-Norm*** auf $\mathcal{C}^0([a,b])$.

(10) Sei (X,d) ein metrischer Raum und $Y \subset X$. Dann ist $d|_{Y\times Y} : Y \times Y \longrightarrow \mathbb{R}$ eine Metrik auf Y, genannt ***Spurmetrik*** oder induzierte Metrik auf Y.

Eine große Klasse von Beispielen von normierten Vektorräume liefern die Skalarprodukträume.

8.2.3. **Definition.** Sei V ein K-Vektorraum ($K = \mathbb{R}$ oder $\mathbb{C}$). Eine Abbildung $V \times V \longrightarrow K$, $(v,w) \mapsto \langle v,w\rangle$ heißt ***Skalarprodukt***, falls gilt:

(i) Die Abbildung $\langle\cdot,w\rangle : V \to K$ ist K-linear für alle $w \in V$,
(ii) $\langle v,w\rangle = \overline{\langle w,v\rangle}$ für alle $v,w \in V$,
(iii) $\langle v,v\rangle > 0$ für alle $v \in V \setminus \{0\}$.

8.2.4. **Beispiele.** Beispiele von Skalarprodukträumen:

- $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ mit $\langle x,y\rangle = \sum x_i y_i$, $(\mathbb{C}^n, \langle\cdot,\cdot\rangle)$ mit $\langle x,y\rangle = \sum x_i \overline{y}_i$,
- $(\ell^2, \langle\cdot,\cdot\rangle)$ mit $\langle (x_n)_n, (y_n)_n\rangle = \sum_{n=0}^{\infty} x_n \overline{y}_n$,
- $(\mathcal{C}^0[a,b], \langle\cdot,\cdot\rangle)$ mit $\langle f,g\rangle = \int_a^b f(x)\overline{g(x)}dx$.

8.2.5. **Satz.** *Sei $(V,\langle\cdot,\cdot\rangle)$ ein Skalarproduktraum. Sei $\|x\| := \sqrt{\langle x,x\rangle}$.*

(a) *Für alle $x,y \in V$ gilt $|\langle x,y\rangle| \leqslant \|x\|\cdot\|y\|$ (**Schwarzsche Ungleichung**).*
Die Gleichheit gilt genau dann, wenn x, y linear abhängig sind.
(b) *$\|\cdot\|$ ist eine Norm auf V.*

Beweis: Ist $y=0$ so ist die Aussage (a) klar. Sei $y\neq 0$ also $\|y\|\neq 0$. Dann gilt

$$\|x\|^2\cdot\|y\|^2-|\langle x,y\rangle|^2=\left\|x-\frac{\langle x,y\rangle}{\|y\|^2}y\right\|^2\geqslant 0\,.$$

8.2.6. **Definition.** Sei (X,d) ein metrischer Raum; eine Folge $(x_n)_{n\in\mathbb{N}}$ in X heißt ***konvergent***, falls es $x\in X$ gibt so, dass zu jedem $\varepsilon>0$ ein $n_0=n_0(\varepsilon)$ existiert mit $d(x_n,x)<\varepsilon$ für alle $n\geqslant n_0(\varepsilon)$. Dann heißt x ***Grenzwert*** der Folge, $\lim_{n\to\infty}x_n=x$. Eine Folge (x_n) heißt ***Cauchy-Folge***, falls gilt: $\forall\varepsilon>0\ \exists n_0=n_0(\varepsilon)$ so, dass $d(x_n,x_m)<\varepsilon$ für alle $m,n>n_0(\varepsilon)$.

Bemerkung:

(i) Der Grenzwert ist eindeutig bestimmt.

(ii) Jede konvergente Folge ist eine Cauchy-Folge.

8.2.7. **Beispiel.** **(i)** $(x_n)_n$ konvergiert in $(\mathbb{R},d_2)$ oder $(\mathbb{C},d_2)$ genau dann, wenn $(x_n)_n$ im Sinne der Definition 2.1.1, 2.4.5 konvergiert.

(ii) Sei $(x^{(k)})_k$ eine Folge in $\mathbb{R}^n$, $x^{(k)}=(x_1^{(k)},\dots,x_n^{(k)})\in\mathbb{R}^n$.

Behauptung: $(x^{(k)})_k$ ist konvergent in $(\mathbb{R}^n,d_2)$ gegen $x=(x_1\dots,x_n)$ genau dann, wenn für jedes $i=1,\dots,n$ die Folge $(x_i^{(k)})_k$ konvergent in $\mathbb{R}$ ist und $\lim_{k\to\infty}x_i^{(k)}=x_i$ in $\mathbb{R}$ gilt.

Also: Konvergenz im $(\mathbb{R}^n,d_2)$ bedeutet koordinatenweise Konvergenz.

Die Behauptung folgt aus der Abschätzung

$$0\leqslant|a_i|\leqslant\|a\|_2=\sqrt{|a_1|^2+\ldots+|a_n|^2}\leqslant|a_1|+\ldots+|a_n|,\ \text{d. h. } \|a\|_\infty\leqslant\|a\|_2\leqslant\|a\|_1\,. \tag{8.4}$$

Eine ähnliche Behauptung gilt für Folgen in $(\mathbb{C}^n,d_2)$.

Durch Benutzung von (8.4) beweist man analog: Ist $(x^{(k)})_k$ eine Cauchy-Folge in $(\mathbb{R}^n,d_2)$ genau dann, wenn für jedes $i=1,\dots,n$ die Folge $(x_i^{(k)})_k$ eine Cauchy-Folge in $\mathbb{R}$ ist.

(iii) Konvergenz in $(\mathcal{B}(D),\|\cdot\|_D)$ bedeutet gleichmäßige Konvergenz der Funktionen.

8.2.8. **Definition.** Ein metrischer Raum (X,d) heißt ***vollständig***, wenn jede Cauchy-Folge konvergiert. Ein normierter Raum $(V,\|\cdot\|)$ heißt ***Banachraum***, wenn V zusammen mit der Norm-Metrik vollständig ist. Ein Skalarproduktraum $(V,\langle\cdot,\cdot\rangle)$ heißt ***Hilbertraum***, wenn V mit der induzierten Norm ein Banachraum ist.

In einem vollständigen metrischen Raum gilt also das Cauchy-Kriterium. Der Unterschied zwischen der Anwendung dieses Kriteriums und der Verwendung der Definition der Konvergenz besteht darin, dass wir im Fall des Cauchy-Kriteriums den Wert des Grenzwertes nicht vorher zu kennen brauchen.

8.2.9. **Beispiele.** **(1)** $(\mathbb{R},d_2)$, $(\mathbb{C},d_2)$ sind vollständige Räume, $(\mathbb{Q},,d_2)$ aber ist nicht vollständig! (Betrachte die Folge (x_n) in $\mathbb{Q}$, gegeben durch $x_{n+1}=\frac{1}{2}(x_n+\frac{2}{x_n})$, wobei $x_0=2$. Dann $x_n\to\sqrt{2}$, $n\to\infty$ in $\mathbb{R}$, also ist (x_n) eine Cauchy-Folge in $\mathbb{R}$ also auch in $\mathbb{Q}$. Aber (x_n) konvergiert nicht in $\mathbb{Q}$, da $\sqrt{2}\not\in\mathbb{Q}$.)

(2) $(\mathbb{R}^n,d_2)$ ist vollständig (d.h. $(\mathbb{R}^n,\|\cdot\|_2)$ ist ein Banachraum, $(\mathbb{R}^n,\langle\cdot,\cdot\rangle)$ Hilbertraum).

(3) Sei $I\subset\mathbb{R}$ ein abgeschlossenes Intervall. Dann ist (I,d_2) vollständig.

(4) $(\mathcal{B}(D),d_D)$ ist vollständig, d. h. $(\mathcal{B}(D),\|\cdot\|_D)$ ist ein Banachraum.

$(\mathcal{C}^0([a,b]),d_{[a,b]})$ ist vollständig, d. h. $(\mathcal{C}^0([a,b]),\|\cdot\|_{[a,b]})$ ist ein Banachraum.

(5) ℓ^p ist ein Banachraum für alle $p\geqslant 1$, ℓ^2 ist ein Hilbertraum.

(6) $(\mathcal{C}^0([a,b]),\|\cdot\|_1)$ versehen mit der L^1-Norm ist kein Banachraum.

8.2.10. **Definition.** Eine Menge $A\subset(X,d)$ heißt ***beschränkt***, wenn es $x_0\in A$ und $M=M(x_0)>0$ gibt so, dass $d(x,x_0)\leqslant M$ für alle $x\in A$.

Bemerkung. **(1)** Ist $(V,\|\cdot\|)$ ein NVR, so ist $A\subset V$ beschränkt (bzgl. der Norm-Metrik) genau dann, wenn es $M>0$ gibt so, dass $\|x\|\leqslant M$ für alle $x\in A$.

(2) $A\subset(\mathbb{R},d_2)$ ist beschränkt genau dann, wenn A im üblichen Sinne beschränkt ist.

(3) Sei $\mathrm{pr}_i:\mathbb{R}^n\to\mathbb{R}$, $\mathrm{pr}_i(x)=x_i$ die Projektion auf die i-te Achse. $A\subset(\mathbb{R}^n,d_2)$ ist beschränkt genau dann, wenn für jedes $i=1,\dots,n$ die Menge $\mathrm{pr}_i(A)$ beschränkt in $\mathbb{R}$ ist.

8.2.11. **Satz** (Bolzano-Weierstraß in $\mathbb{R}^n$). *Jede beschränkte Folge in $(\mathbb{R}^n,d_2)$ besitzt eine konvergente Teilfolge.*

8.2.12. **Definition.** Zwei Normen $\|\cdot\|$, $\|\cdot\|_*$ auf einem Vektorraum V heißen ***äquivalent***, geschrieben $\|\cdot\| \sim \|\cdot\|_*$, wenn es zwei Konstanten $c_1, c_2 > 0$ gibt mit $c_1\|x\| \leqslant \|x\|_* \leqslant c_2\|x\|$ für alle $x \in V$. Dies ist eine Äquivalenzrelation.

Ist $\|\cdot\| \sim \|\cdot\|_*$, so gilt: $(x_n)_n$ ist konvergent gegen x (bzw. Cauchy-Folge) in $(V, \|\cdot\|)$ genau dann, wenn $(x_n)_n$ konvergent gegen x (bzw. Cauchy-Folge) in $(V, \|\cdot\|_*)$ ist.

8.2.13. **Satz.** *Sei V ein endlichdimensionaler K-Vektorraum ($K = \mathbb{R}, \mathbb{C}$). Dann gilt:*

(i) *Je zwei Normen auf V sind äquivalent.*
(ii) *$(V, \|\cdot\|)$ ist ein Banachraum für jede Norm auf V.*

Wieso interessieren wir uns für äquivalente oder vergleichbare Normen? Der Grund ist, dass wir die Konvergenz untersuchen möchten. Gilt die Ungleichung $\|\cdot\| \leqslant C\|\cdot\|_*$, so ist jede bezüglich $\|\cdot\|_*$ konvergente Folge auch konvergent bezüglich $\|\cdot\|$; gilt $\|\cdot\| \sim \|\cdot\|_*$, so stimmen die konvergenten Folgen in beiden Normen überein. Zum Beispiel bedeutet der Satz „$\int_a^b f_n\,dx \longrightarrow \int_a^b f\,dx$ falls $(f_n)_n$, $f_n \in \mathcal{C}([a,b])$ gleichmäßig auf $[a,b]$ gegen f konvergiert“ nichts anderes, als dass jede konvergente Folge in $(\mathcal{C}([a,b]), \|\cdot\|_{[a,b]})$ auch in $(\mathcal{C}([a,b]), \|\cdot\|_1)$ (wobei $\|f\|_1 = \int_a^b |f(x)|\,dx$) konvergiert. Dies folgt aus der Abschätzung $\|f\|_1 \leqslant (b-a)\|f\|_{[a,b]}$ für $f \in \mathcal{C}([a,b])$.

Übrigens sind die Normen $\|\cdot\|_{[a,b]}$ und $\|\cdot\|_1$ jedoch nicht äquivalent: Die Folge (f_n), $f_n(t) = t^n$ konvergiert in $\|\cdot\|_1$ gegen 0, aber konvergiert nicht in $\|\cdot\|_{[a,b]}$. Ein anderes Argument: $(\mathcal{C}([a,b]), \|\cdot\|_{[a,b]})$ ist vollständig, $(\mathcal{C}([a,b]), \|\cdot\|_1)$ jedoch nicht (siehe Aufgabe 8.8.5).

8.2.14. **Definition.** Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Ein Paar $(x_n)_{n\geqslant 0}$, $(s_n)_{n\geqslant 0}$ wird eine Reihe genannt, wenn $s_n = x_0 + \ldots + x_n$. Die Reihe wird durch $\sum_{n\geqslant 0} x_n$ bezeichnet. Wir sagen, dass die Reihe $\sum_{n\geqslant 0} x_n$ konvergiert, wenn $(s_n)_{n\geqslant 0}$ konvergiert. Dann heißt $s = \lim_{n\to\infty} s_n$ die Summe der Reihe. Wir schreiben $s = \sum_{n=0}^\infty x_n$. Eine Reihe $\sum_{n\geqslant 0} x_n$ heißt absolut konvergent, wenn $\sum_{n\geqslant 0} \|x_n\|$ in $\mathbb{R}$ konvergiert.

8.2.15. **Satz.** *Sei $(V, \|\cdot\|)$ ein Banachraum. Dann gilt:*

(i) *(Cauchy-Kriterium) $\sum_{n\geqslant 0} x_n$ konvergiert genau dann, wenn es für alle $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon)$ gibt so, dass für alle $m > n \geqslant n_0$ gilt: $\|\sum_{k=n+1}^m x_k\| < \varepsilon$.*
(ii) *Ist $\sum_{n\geqslant 0} x_n$ absolut konvergent, dann ist $\sum_{n\geqslant 0} x_n$ konvergent.*

8.2.16. **Definition.** Ein normierter K-Vektorraum $(V, \|\cdot\|)$ heißt normierte K-Algebra, wenn es eine assoziative Verknüpfung $V \times V \longrightarrow V$, $(x,y) \mapsto xy$ gibt so, dass $\|xy\| \leqslant \|x\| \cdot \|y\|$ für alle $x, y \in V$. Ist $(V, \|\cdot\|)$ zugleich ein Banachraum, so heißt $(V, \|\cdot\|)$ eine ***Banachalgebra***.

Beispiele: $V = M_{n\times n}(K)$ mit der Multiplikation der Matrizen und Norm $\|A\|_2 = (\sum_{i,j=1}^n |a_{ij}|^2)^{1/2}$ ist eine Banachalgebra. Auch $(\mathcal{B}(D), \|\cdot\|_D)$ und $(\mathcal{C}^0([a,b]), \|\cdot\|_{[a,b]})$ sind Banachalgebren.

8.2.17. **Satz.** *Sei $(V, \|\cdot\|)$ eine Banachalgebra mit Einselement e. Ist $\|x\| < 1$, so ist $e - x$ invertierbar, und es gilt $(e-x)^{-1} = \sum_{n=0}^\infty x^n = e + x + x^2 + \ldots$*

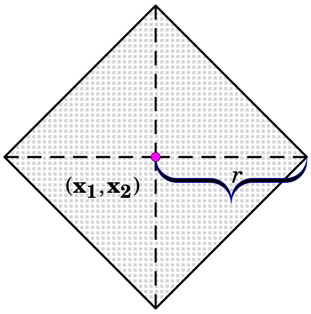
8.3. **Topologie eines metrischen Raumes. Topologische Räume.** Unser Ziel ist, bemerkenswerte Teilmengen eines metrischen Raumes zu beschreiben, die ein tieferes Studium der Funktionen ermöglichen.

8.3.1. **Definition.** Sei (X, d) ein metrischer Raum. Sei $a \in X$, $r > 0$. Dann heißt

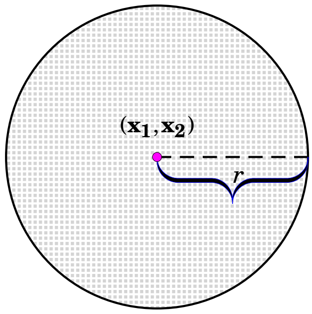
(i) $B_r(a) = \{x \in X : d(x,a) < r\}$ die offene Kugel mit dem Mittelpunkt a und dem Radius r,
(ii) $\overline{B}_r(a) = \{x \in X : d(x,a) \leqslant r\}$ die abgeschlossene Kugel mit dem Mittelpunkt a und dem Radius r,
(iii) $S_r(a) = \{x \in X : d(x,a) = r\}$ die Sphäre mit dem Mittelpunkt a und dem Radius r.

Beispiele:

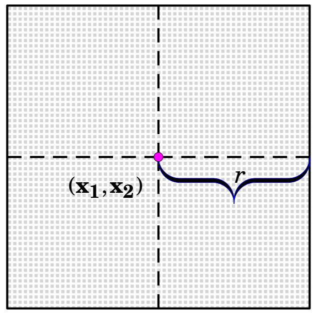
In $(\mathbb{R}, d_2)$: $B_r(a) = (a-r, a+r)$, $\overline{B}_r(a) = [a-r, a+r]$, $S_r(a) = \{a-r, a+r\}$.
In $(\mathbb{R}^2, d_2)$: $B_r(a) = \{(x_1,x_2) : (x_1-a_1)^2 + (x_2-a_2)^2 < r^2\}$.
In $(\mathbb{R}^2, d_1)$: $B_r(a) = \{(x_1,x_2) : |x_1-a_1| + |x_2-a_2| < r\}$.
In $(\mathbb{R}^2, d_\infty)$: $B_r(a) = \{(x_1,x_2) : \max\{|x_1-a_1|, |x_2-a_2|\} < r\}$.



$B_r(x)$ bzgl. d_1



$B_r(x)$ bzgl. d_2



$B_r(x)$ bzgl. d_∞

8.3.2. **Definition.** Sei (X,d) ein metrischer Raum. Eine Teilmenge $U \subset X$ heißt ***offen*** (bezüglich d) $:\Leftrightarrow$ für jedes $a \in U$ gibt es $r > 0$ mit $B_r(a) \subset U$. Eine Teilmenge $F \subset X$ heißt ***abgeschlossen*** $:\Leftrightarrow$ $X \setminus F$ ist offen.

Beispiele:

- Offene Intervalle in $\mathbb{R}$ sind offen in $(\mathbb{R}, d_2)$.
- Jede offene Menge in $\mathbb{R}$ ist eine disjunkte Vereinigung von höchstens abzählbar vielen offenen Intervalle.
- Abgeschlossene Intervalle in $\mathbb{R}$ sind abgeschlossen in $(\mathbb{R}, d_2)$.
- $B_r(a) \subset (X,d)$ ist offen: Für $b \in B_r(a)$ gilt $B_{r-d(a,b)}(b) \subset B_r(a)$.
- $\overline{B}_r(a) \subset (X,d)$ ist abgeschlossen.
- Jede einpunktige Teilmenge $\{a\} \subset X$ ist abgeschlossen.

8.3.3. **Satz.** *Sei (X,d) ein metrischer Raum. Dann gilt:*

(1) *X und die leere Menge $\emptyset$ sind offen.*
(2) *Die Vereinigung beliebig vieler offener Mengen ist offen.*
(3) *Der Durchschnitt endlich vieler offener Mengen ist offen.*
(1') *X und die leere Menge $\emptyset$ sind abgeschlossen.*
(2') *Die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.*
(3') *Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.*

8.3.4. **Bemerkung.** Die Aussagen (3) und (2') sind nicht wahr für beliebig viele Mengen: $\bigcap_{n\geqslant 1}\left(a-\frac{1}{n}, a+\frac{1}{n}\right) = \{a\}$ ist nicht offen, $\bigcup_{n\geqslant 2}\left[\frac{1}{n}, 1-\frac{1}{n}\right] = (0,1)$ ist nicht abgeschlossen in $\mathbb{R}$.

8.3.5. **Satz.** *Die Menge F ist abgeschlossen genau dann, wenn für jede Folge $(x_n)_n$ in F, die in X konvergiert, gilt:* $\lim_{n\to\infty} x_n \in F$.

Mittels Satz 8.3.5 kann man z. B. zeigen:

- Die Menge $\{x \in \mathbb{R}^n : Ax = b\}$ ist abgeschlossen in $\mathbb{R}^n$, wobei $A \in M_{m\times n}(R)$, $b \in \mathbb{R}^m$.
- Die Menge $\mathcal{C}^0([a,b])$ ist abgeschlossen in $(\mathcal{B}([a,b]), d_{[a,b]})$.

8.3.6. **Definition.** Sei X eine Menge. Eine ***Topologie*** auf X ist eine Teilmenge $\mathcal{O} \subset \mathcal{P}(X)$ von Teilmengen von X mit den Eigenschaften:

(1) $X, \emptyset \in \mathcal{O}$.
(2) Die Vereinigung beliebig vieler Mengen aus $\mathcal{O}$ gehört zu $\mathcal{O}$.
(3) Der Durchschnitt endlich vieler offener Mengen aus $\mathcal{O}$ gehört zu $\mathcal{O}$.

Das Paar $(X, \mathcal{O})$ heißt ***topologischer Raum***. Eine Teilmenge $U \subset X$ heißt ***offen***, falls $U \in \mathcal{O}$. Eine Teilmenge $F \subset X$ heißt ***abgeschlossen***, falls $X \setminus F$ offen ist.

8.3.7. **Beispiele.** **(1)** Ist (X,d) ein metrischer Raum, so ist

$$\mathcal{O} := \mathcal{O}(d) := \{U \subset X : U \text{ offen bezüglich } d\}$$

(siehe Def. 8.3.2) eine Topologie auf X, genannt die ***induzierte Topologie von*** d. Dies folgt aus Satz 8.3.3.

(2) Ein topologischer Raum $(X,\mathcal{O})$ heißt ***metrisierbar***, falls es eine Metrik gibt mit $\mathcal{O}(d)=\mathcal{O}$. Nicht alle topologischen Räume sind metrisierbar. Sei X eine Menge. Die Familie $\mathscr{O}=\{\emptyset,X\}$ bilden eine Topologie, genannt grobe Topologie. Die einzigen offenen Teilmengen sind $\emptyset$ und X, und die sind auch die einzigen abgeschlossenen Teilmengen. Nehmen wir an, X hat mindestens zwei Punkte. Dann is $(X,\mathscr{O})$ nicht metrisierbar. Wäre $(X,\mathscr{O})$ metrisierbar, so wäre $\{a\}$ abgeschlossen für $a\in X$, also $\{a\}=\emptyset$ oder $\{a\}=X$, Widerspruch.

Wir interessieren uns hier nicht unbedingt für allgemeine topologische Räume, sondern für die topologischen Begriffe und Eigenschaften von metrischen Räumen. Topologische Begriffe, d.h. solche, die ausschließlich von der Topologie des Raumes abhängen, sind flexibler als metrische Begriffe: Sie sind invariant unter sehr allgemeinen Transformationen, nämlich den sogenannten Homöomorphismen. Wichtige topologische Begriffe sind: Offene und abgeschlossene Menge (siehe oben), Umgebung, Inneres und Abschluss, Rand, stetige Abbildungen, Kompaktheit, Zusammenhang usw. Dagegen sind Kugel, Sphäre, beschränkte Menge metrische Begriffe, nicht topologische Begriffe.

(3) Seien $\|\cdot\|$ und $\|\cdot\|_*$ zwei Normen auf einen Vektorraum V. Es gilt $\|\cdot\|\sim\|\cdot\|_*$ genau dann, wenn die zugehörigen Metriken dieselbe Topologie induzieren.

(4) Sei V ein endlichdimensionaler Vektorraum. Dann sind alle Normen auf V äquivalent, und die zugehörigen Metriken induzieren dieselbe Topologie. Diese zu beliebige Normen gehörige Topologie heißt ***Standardtopologie*** auf V.

Ein endlichdimensionaler Vektorraum besitzt unendlich viele Normen, aber keine von diesen ist besonders ausgezeichnet; je nach Zweck ist mal die eine, mal eine andere Norm geeignet. Allerdings sind sie alle äquivalent: Die durch sie definierten Metriken liefern alle dieselbe Topologie (also dieselben offenen Mengen, dieselben konvergenten Folgen, dieselben stetigen Abbildungen).

8.3.8. **Definition.** Sei $(X,\mathcal{O})$ ein topologischer Raum, $a\in X$. Eine Teilmenge $U\subset X$ heißt ***Umgebung*** von a, wenn es eine offene Menge V gibt mit $a\in V\subset U$. Die Menge aller Umgebungen von a wird mit $\mathcal{U}_a$ bezeichnet.

Für einen metrischen Raum (X,d) gilt: $U\in\mathcal{U}_a \Leftrightarrow \exists\varepsilon>0: B_\varepsilon(a)\subset U$.

8.3.9. **Bemerkung.**

(i) Der Durchschnitt endlich vieler Umgebungen von a ist eine Umgebung von a.
(ii) Wenn $U\in\mathcal{U}_a$ und $U\subset V$, dann $V\in\mathcal{U}_a$.
(iii) U ist offen $\Leftrightarrow$ U ist Umgebung jedes seiner Punkte (d.h. $\forall x\in U\colon U\in\mathcal{U}_x$).

Es sei $(X,\mathcal{O})$ ein topologischer Raum, $\mathcal{U}_x$ die Menge aller Umgebungen von x. Eine Teilmenge $\mathcal{B}\subset\mathcal{O}$ heißt ***Basis der Topologie***, wenn zu jedem $U\in\mathcal{O}$ und zu jedem $x\in U$ ein $B\in\mathcal{B}$ existiert mit $x\in B\subset U$. Eine Teilmenge $\mathcal{V}_x\subset\mathcal{U}_x$ heißt ***Umgebungsbasis*** des Punktes x (auch: Fundamentalsystem der Umgebungen von x), wenn zu jedem $U\in\mathcal{U}_x$ ein $V\subset\mathcal{V}_x$ existiert mit $V\subset U$.

Wir sagen, dass $(X,\mathcal{O})$ das ***erste Abzählbarkeitsaxiom*** erfüllt, falls jeder Punkt eine abzählbare Umgebungsbasis hat. Wir sagen, dass $(X,\mathcal{O})$ das ***zweite Abzählbarkeitsaxiom*** erfüllt, falls die Topologie eine abzählbare Basis hat.

Jeder metrische Raum erfüllt das erste Abzählbarkeitsaxiom: $B_{1/n}(x)$ ist eine abzählbare Umgebungsbasis von x. Betrachte nun eine beliebige Menge X und setze $\mathcal{O}:=\{\emptyset\}\cup\{U\subset X : X\setminus U \text{ endlich}\}$. Ist X überabzählbar, so besitzt $x\in X$ keine abzählbare Umgebungsbasis: Angenommen, $\mathcal{V}\subset\mathcal{U}_x$ wäre eine abzählbare Umgebungsbasis. Dann gälte $\{x\}=\cap_{V\in\mathcal{V}}V$, $X\setminus\{x\}=\cup_{V\in\mathcal{V}}(X\setminus V)$, d. h. X wäre abzählbar; Widerspruch. Dies bedeutet, dass $(X,\mathcal{O})$ nicht metrisierbar ist, wenn X überabzählbar ist.

$\mathbb{R}^n$ mit der Standardtopologie erfüllt das zweite Abzählbarkeitsaxiom. Finden Sie eine abzählbare Basis!

8.3.10. **Definition.** Sei $(X,\mathcal{O})$ ein topologischer Raum und $A\subset X$. Ein Punkt $x\in X$ heißt:

(i) ***innerer Punkt*** von A, wenn A eine Umgebung von x ist.
(ii) ***Berührpunkt*** von A, wenn jede Umgebung von x (mindestens) einen Punkt aus A enthält.
(iii) ***Randpunkt*** von A, wenn x Berührpunkt von A und $X\setminus A$ ist.
(iv) ***Häufungspunkt*** von A, wenn jede Umgebung von x unendlich viele Punkte aus A enthält.
(v) ***isolierter Punkt*** von A, wenn es eine Umgebung U von x gibt mit $U\cap A=\{x\}$.

(vi) ***äußerer Punkt*** von A, wenn es eine Umgebung U von x gibt mit $U \cap A = \emptyset$.

8.3.11. **Bemerkung.** Sei (X,d) ein metrischer Raum und $a \in X$, $A \subset X$.
a ist Berührpunkt von A genau dann, wenn es eine Folge (x_n) in A gibt mit $\lim x_n = a$.
a ist Häufungspunkt von A genau dann, wenn es eine Folge (x_n) in A gibt mit $x_n \neq a$ und $\lim x_n = a$.

8.3.12. **Definition.** Sei $(X,\mathcal{O})$ ein topologischer Raum, $A \subset X$. Wir führen ein:

(i) $\mathring{A}$, die Menge aller inneren Punkte von A, genannt das ***Innere*** von A.
(ii) $\overline{A}$, die Menge aller Berührpunkte von A, genannt der ***Abschluss*** von A.
(iii) ∂A, die Menge aller Randpunkte, genannt der ***Rand*** von A.
(iv) A', die Menge der Häufungspunkte von A.

Als Beispiel betrachte $A = [0,1) \cup \{2\} \subset (\mathbb{R}, d_2)$. Dann ist $\mathring{A} = (0,1)$, $\overline{A} = [0,1] \cup \{2\}$, $\partial A = \{0,1,2\}$, $A' = [0,1]$. Die Menge der isolierten Punkten ist $\{2\}$, die Menge der äußeren Punkten ist $(-\infty,1)\cup(1,2)\cup(2,\infty)$. Sei nun $\mathbb{Q} \subset (\mathbb{R}, d_2)$. Da jedes offenes Intervall unendlich viele Punkte aus $\mathbb{Q}$ und aus $\mathbb{R} \setminus \mathbb{Q}$ enthält, gilt $\mathring{\mathbb{Q}} = \emptyset$, $\overline{\mathbb{Q}} = \mathbb{R}$, $\partial\mathbb{Q} = \mathbb{R}$, $\mathbb{Q}' = \mathbb{R}$, die Menge der isolierten Punkten und die Menge der äußeren Punkten sind leer.

8.3.13. **Satz.** *Sei $(X,\mathcal{O})$ ein topologischer Raum, $A \subset X$. Dann gilt:*

(1) $\mathring{A} = \cup\{U \text{ offen} : U \subset A\}$ *(d.h. $\mathring{A}$ ist die größte offene Menge, die in A enthalten ist). Insbesondere ist $\mathring{A}$ offen.*
(2) $\overline{A} = \cap\{F \text{ abgeschlossen} : F \supset A\}$ *(d.h. $\overline{A}$ ist die kleinste abgeschlossene Menge, die A enthält). Insbesondere ist $\overline{A}$ abgeschlossen.*
(3) $\partial A = \overline{A} \cap \overline{X \setminus A}$*; ∂A ist abgeschlossen.*
(4) $\mathring{A} = A \setminus \partial A$*,* $\overline{A} = A \cup \partial A = A \cup A' = \mathring{A} \sqcup \partial A$
(5) A *offen* $\Leftrightarrow A = \mathring{A} \Leftrightarrow A \cap \partial A = \emptyset \Leftrightarrow \partial A = \overline{A} \setminus A$
(6) A *abgeschlossen* $\Leftrightarrow A = \overline{A} \Leftrightarrow \partial A \subset A \Leftrightarrow A' \subset A$
(7) $\widehat{X \setminus \mathring{A}} = X \setminus \overline{A}$*,* $X = \mathring{A} \sqcup \partial A \sqcup \widehat{X \setminus \mathring{A}}$*.*

8.3.14. **Definition.** Ein topologischer Raum $(X,\mathcal{O})$ heißt ***Hausdorffraum***, wenn es zu jedem Paar $x,y \in X$ mit $x \neq y$ Umgebungen U von x und V von y gibt mit $U \cap V = \emptyset$.

Sei (X,d) ein metrischer Raum. Dann ist $(X,\mathcal{O}(d))$ ein Hausdorffraum: Für $x \neq y$ gilt

$$B\big(x,\tfrac{r}{2}\big) \cap B\big(y,\tfrac{r}{2}\big) = \emptyset, \quad \text{wobei } r = d(x,y) > 0\,.$$

In einem Hausdorffraum sind die einpunktigen Mengen $\{p\}$ mit $p \in X$ abgeschlossen: Zu jedem $x \in X \setminus \{p\}$ wähle eine offene Umgebung U_q von q mit $U_q \subset X \setminus \{p\}$. Dann gilt $X \setminus \{p\} = \bigcup_{q \in X \setminus \{p\}} U_q$ und $X \setminus \{p\}$ ist offen als Vereinigung von offenen Mengen.

8.3.15. **Definition.** Sei $(X,\mathcal{O})$ ein topologischer Raum. Eine Teilmenge $A \subset X$ heißt ***dicht in*** X falls $\overline{A} = X$.

Zum Beispiel, eine Teilmenge $A \subset \mathbb{R}$ ist dicht in $\mathbb{R}$, genau dann, wenn für alle Intervalle (a,b) in $\mathbb{R}$ mit $a < b$ gilt $A \cap (a,b) \neq \emptyset$. Daher sind $\mathbb{Q}$ und $\mathbb{R} \setminus \mathbb{Q}$ dicht in $\mathbb{R}$. Siehe dazu Satz 1.5.5; wir haben also in Definition 8.3.15 eine Verallgemeinerung der Dichtheit in $\mathbb{R}$. Ein Beispiel in $\mathbb{R}^n$: die Menge $\mathbb{Q}^n$ liegt dicht in $\mathbb{R}^n$.

8.3.16. **Definition.**

(i) Sei $(X,\mathcal{O})$ ein topologischer Raum, $A \subset X$. Dann ist $\mathcal{O}_A := \{U \cap A : U \in \mathcal{O}\}$ eine Topologie auf A, genannt ***induzierte Topologie*** oder ***Teilraumtopologie***.
(ii) Sind $(X_1,\mathcal{O}_1),\dots,(X_n,\mathcal{O}_n)$ topologische Räume, so ist

$$\mathcal{O} := \{U \subset X_1 \times \ldots \times X_n : \forall x \in U\, \exists U_i \in \mathcal{O}_i,\, i = 1,\dots,n : x \in U_1 \times \ldots \times U_n \subset U\}$$

eine Topologie auf $X_1 \times \ldots \times X_n$, genannt ***Produkttopologie*** von $\mathcal{O}_1,\dots,\mathcal{O}_n$.

8.3.17. **Bemerkung.** (i) Die Standardtopologie von $\mathbb{R}$ ist die Teilraumtopologie von $\mathbb{R} \subset \mathbb{C}$. Ist $U \subset \mathbb{R}$ offen, so gibt es zu jedem $x \in U$ ein $\varepsilon(x) > 0$ mit $(x - \varepsilon(x), x + \varepsilon(x)) \subset U$. Sei $B_{\varepsilon(x)}(x) = \{z \in \mathbb{C} : |z - x| < \varepsilon(x)\}$ und $V = \cup_{x \in U} B_{\varepsilon(x)}(x) \subset \mathbb{C}$; V ist offen in $\mathbb{C}$ als Vereinigung offener Mengen, und $U = V \cap \mathbb{R}$. Daraus folgt, dass U offen in der Teilraumtopologie ist. Sei nun $U \subset \mathbb{R}$ offen in der Teilraumtopologie. Laut Definition gibt es $V \subset \mathbb{C}$ offen mit $U = V \cap \mathbb{R}$. Sei $x \in U \subset V$. Dann gibt es $\varepsilon > 0$, so dass $B_\varepsilon(x) \subset V$. Daraus folgt schon $(x-\varepsilon, x+\varepsilon) = B_{\varepsilon(x)}(x) \cap \mathbb{R} \subset V \cap \mathbb{R} = U$, also U ist offen bzgl. der Standardtopologie von $\mathbb{R}$.

(ii) Sei (X,d) ein metrischer Raum, $A \subset X$. Dann ist die Teilraumtopologie $\mathcal{O}_A$ durch die Metrik d auf A induziert.

(iii) Sind $(X_1,d_1),\dots,(X_n,d_n)$ metrische Räume, so ist die Produkttopologie der induzierten Topologien $\mathcal{O}(d_1)$, ..., $\mathcal{O}(d_n)$ induziert durch die Metrik

$$d((x_1,\dots,x_n),(y_1,\dots,y_n)) = \max\{d_1(x_1,y_1),\dots,d_n(x_n,y_n)\}.$$

(iv) Die Standardtopologie auf $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ ist die Produkttopologie der Standardtopologien auf $\mathbb{R}$. Das folgt aus (ii) und die Tatsache, dass d_∞ auf $\mathbb{R}^n$ die Standardtopologie induziert.

(v) Die Begriffe der offenen und abgeschlossenen Mengen, des Inneren, des Abschlusses und des Randes einer Menge hängen von der Topologie des umgebenden Raumes ab, d.h. sind *extrinsische Begriffe*. Zum Beispiel ist $(0,1)$ keine abgeschlossene Menge in $\mathbb{R}$ bzgl. der Standardtopologie, ist aber abgeschlossen bzgl. der Teilraumtopologie von $(0,1)$.

Als weiteres Beispiel sei $A = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 < 1, z = 0\}$. Das ist eine Kreisscheibe in der Ebene $E = \{(x,y,z) \in \mathbb{R}^3 : z = 0\}$ von $\mathbb{R}^3$. Wir betrachten $\mathbb{R}^3$ versehen mit der Standardtopologie. Dann gilt $\mathring{A} = \emptyset$, $\partial A = \overline{A} = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 \leqslant 1, z = 0\}$. Betrachten wir nun A als Teilmenge von E, wobei E mit der Teilraumtopologie versehen sei. Dann ist $\mathring{A} = A$, $\overline{A} = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 \leqslant 1, z = 0\}$ und $\partial A = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 = 1, z = 0\}$.

8.4. **Stetige Abbildungen.**

8.4.1. **Definition.** Seien X,Y topologische Räume. Eine Abbildung $f : X \longrightarrow Y$ heißt ***stetig in*** $a \in X$, wenn es zu jeder Umgebung V von $f(a)$ eine Umgebung U von a gibt mit $f(U) \subset V$. Die Abbildung f heißt ***stetig***, wenn sie in allen $a \in X$ stetig ist. f heißt ***Homöomorphismus***, falls f bijektiv ist und sowohl f als auch f^{-1} stetig sind.

8.4.2. **Beispiel.**

(i) Konstante Abbildungen.

(ii) Identische Abbildung $\mathrm{Id} : X \to X$.

(iii) Eine Abbildung $f : (X,d_X) \to (Y,d_Y)$ zwischen zwei metrischen Räume heißt ***Lipschitz-stetig***, falls eine Konstante $L \geqslant 0$ existiert, so dass $d_Y(f(x_1),f(x_2)) \leqslant L d_X(x_1,x_2)$ für alle $x_1,x_2 \in X$. In diesem Fall heißt L eine ***Lipschitz-Konstante*** zu f. Die Abbildung f heißt ***Isometrie***, falls $d_Y(f(x_1),f(x_2)) = d_X(x_1,x_2)$ für alle $x_1,x_2 \in X$.

(iv) Ist $f : X \longrightarrow Y$ stetig, $A \subset X$ mit der Teilraumtopologie versehen, so ist auch $f|_A : A \longrightarrow Y$ stetig.

(v) Seien $X_1,\dots,X_n$ topologische Räume, $X_1 \times \ldots \times X_n$ versehen mit der Produkttopologie. Sei $\mathrm{pr}_i : X_1 \times \ldots \times X_n \longrightarrow X_i$, $\mathrm{pr}_i(x) = x_i$. Dann ist pr_i stetig: Ist V eine Umgebung von x_i in X_i, so ist $U = \mathrm{pr}^{-1}(V) = X_1 \times \ldots X_{i-1} \times V \times X_{i+1} \times \ldots \times X_n$ eine Umgebung von x, $\mathrm{pr}_i(U) = V$.

(vi) Sei $\mathbb{C}$ versehen mit der Standardtopologie (induziert z.B. durch die euklidische Metrik) und $D \subset \mathbb{C}$ versehen mit der Teilraumtopologie. Dann ist $f : D \to \mathbb{C}$ stetig im Sinne von Definition 8.4.1 genau dann, wenn f im Sinne von Definition 4.1.4 (Analysis I) stetig ist (siehe Satz 4.1.5).

(vii) Die Abbildung $\exp : \mathbb{R} \to \mathbb{R}_+$ ist ein Homöomorphismus (mit Inversem $\log : \mathbb{R}_+ \to \mathbb{R}$); Die Abbildung $\tan : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$ ist ein Homöomorphismus (mit Inversem $\arctan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$); im Allgemeinen zeigt Satz 4.1.10: Wenn $I \subset \mathbb{R}$ ein Intervall ist und $f : I \to \mathbb{R}$ eine stetige und streng monotone Funktion, dann ist $f(I)$ ein Intervall und $f : I \to f(I)$ ein Homöomorphismus.

(viii) Die Polarkoordinatenabbildung $P : \mathbb{R}_+ \times (0,2\pi) \longrightarrow \mathbb{R}^2 \setminus \{(x,y) : x \geqslant 0, y = 0\}$ ist ein Homöomorphismus mit Inversem $(x,y) \mapsto (\sqrt{x^2+y^2}, \arg(x+iy))$ (Satz 8.4.11).

8.4.3. **Satz.** *Seien $f,g : X \longrightarrow \mathbb{C}$ stetig in $a \in X$, und sei $\lambda \in \mathbb{C}$. Dann gilt:*

(1) *$f+g$, $f \cdot g$, $\lambda \cdot f$ sind stetig in a.*
(2) *f/g ist stetig in a, falls $g(a) \neq 0$.*

Daraus folgt, dass Polynome $P : \mathbb{R}^n \longrightarrow \mathbb{R}$, $P \in \mathbb{R}[x_1,\dots,x_n]$ stetig sind; rationale Funktionen $R = P/Q$, $R : \mathbb{R}^n \setminus \{x \in \mathbb{R}^n : Q(x) = 0\} \longrightarrow \mathbb{R}$ sind stetig.

8.4.4. **Satz.** *Sei $f : X \longrightarrow Y$ stetig in a, und sei $g : Y \longrightarrow Z$ stetig in $b = f(a)$. Dann ist $g \circ f$ stetig in a.*

Beweis: Sei W eine Umgebung von $(g\circ f)(a)=g(f(a))$. Da g stetig in $f(a)$ ist, existiert eine Umgebung V von $f(a)$, so dass $g(V)\subset W$. Da f stetig in a ist, existiert eine Umgebung U von a, so dass $f(U)\subset V$. Somit ist $(g\circ f)(U)=g(f(U))\subset g(V)\subset W$. Hier haben wir folgendes benutzt: Ist $g:Y\longrightarrow Z$ eine Abbildung und $A\subset B\subset Y$, so gilt $g(A)\subset g(B)$. $\square$

Beispiele:

(1) $f:X\longrightarrow\mathbb{R}^m$, $f=(f_1,\dots,f_m)$ ist stetig genau dann, wenn für jedes $i=1,\dots,m$ die Abbildung $f_i:X\longrightarrow\mathbb{R}$ stetig ist. Dies gilt allgemeiner: Sind $Y_1,\dots,Y_m$ topologische Räume, so ist $f=(f_1,\dots,f_m):X\to Y_1\times\dots\times Y_m$ stetig (bzgl. der Produkttopologie auf $Y_1\times\dots\times Y_m$) genau dann, wenn für jedes $i=1,\dots,m$ die Abbildung $f_i:X\to Y_i$ stetig ist.

(2) Die Funktion $f:\mathbb{R}^2\longrightarrow\mathbb{R}$,

$$f(x,y)=\begin{cases}\dfrac{x^3y}{x^2+y^2}\,, & \text{falls}\quad (x,y)\neq(0,0)\\ 0\,, & \text{falls}\quad (x,y)=(0,0)\end{cases}$$

ist stetig auf $\mathbb{R}^2$. In der Tat, f ist stetig auf $\mathbb{R}^2\setminus\{(0,0)\}$, da f dort eine rationale Funktion ist. Außerdem gilt die Abschätzung $|f(x,y)|=\frac{|x^3y|}{x^2+y^2}\leqslant\frac{|x^3y|}{2|xy|}=\frac12|x|^2$.
Sei $\varepsilon>0$. Ist $(x,y)\in B^\infty_\delta(0,0)=\{(x,y):\max\{|x|,|y|\}<\delta\}$ mit $\delta=\min\{1,\varepsilon\}$, so gilt $|f(x,y)-f(0,0)|\leqslant\frac12|x|^2<\varepsilon$. Daraus folgt, dass f stetig in $(0,0)$ ist.

(3) Sei $f:\mathbb{R}^2\longrightarrow\mathbb{R}$,

$$f(x,y)=\begin{cases}\dfrac{2xy}{x^2+y^2}\,, & \text{falls}\quad (x,y)\neq(0,0)\\ 0\,, & \text{falls}\quad (x,y)=(0,0).\end{cases}$$

f ist ***separat stetig*** d.h. $f(\cdot,y)$ ist stetig für alle y, $f(x,\cdot)$ ist stetig für alle x. Trotzdem ist f nicht stetig in $(0,0)$: Für $x\neq0$ gilt $f(x,x)=\frac{2x^2}{2x^2}=1$, also gibt es in jeder Umgebung von $(0,0)$ einen Punkt (x,x) mit $f(x,x)=1$ (wähle $x=\frac1n$ mit n genügend groß); keine Umgebung von $(0,0)$ wird also in die Umgebung $(-1,1)$ von $f(0)=0\in\mathbb{R}$ abgebildet.

Dieses Beispiel zeigt, warum separate Stetigkeit schwächer ist als Stetigkeit: Die Variable (x,y) muss sich der Stelle $(0,0)$ auf beliebige Weise nähern dürfen. Bei der partiellen Stetigkeit schränkt man sich aber auf achsenparallele Annäherung ein. Die Funktion f in dem Beispiel ist sogar auf allen Geraden durch den Nullpunkt jeweils konstant, und im Nullpunkt hat sie den Wert 0. Der kommt heraus, wenn man auf der x-Achse oder auf der y-Achse an den Nullpunkt heranläuft, aber eben nur dann.

8.4.5. **Satz.** *Sei $f:(X,d_1)\longrightarrow(Y,d_2)$ eine Abbildung, $a\in X$. Die folgenden Aussagen sind äquivalent:*

(1) *f ist stetig in a.*
(2) *(ε-δ-Kriterium) Für jedes $\varepsilon>0$ gibt es ein $\delta>0$ so, dass für jedes $x\in X$ mit $d(x,a)<\delta$ gilt: $d(f(x),f(a))<\varepsilon$.*
(3) *(Folgenkriterium) Für jede Folge $(x_n)_n$ mit $\lim\limits_{n\to\infty}x_n=a$ gilt $\lim\limits_{n\to\infty}f(x_n)=f(a)$.*

8.4.6. **Satz** (Globale Charakterisierung der Stetigkeit)**.** *Seien X,Y topologische Räume und $f:X\longrightarrow Y$. Folgende Aussagen sind äquivalent:*

(1) *f ist stetig.*
(2) *Für jedes offene $U\subset Y$ ist $f^{-1}(U)$ offen.*
(3) *Für jedes abgeschlossene $F\subset Y$ ist $f^{-1}(F)$ abgeschlossen.*

8.4.7. **Folgerung.** *Sei $f:X\longrightarrow\mathbb{R}$ stetig, und sei I ein offenes (bzw. abgeschlossenes) Intervall. Dann ist $f^{-1}(I)$ offen (bzw. abgeschlossen).*

8.4.8. **Beispiel.** Die ***n-Sphäre*** $S^n=\{x\in\mathbb{R}^{n+1}:x_1^2+\ldots+x_{n+1}^2=1\}$ ist abgeschlossen, da $S^n=f^{-1}(1)$, wobei f die stetige Funktion $f:\mathbb{R}^n\longrightarrow\mathbb{R}$, $f(x)=x_1^2+\ldots+x_{n+1}^2$ ist.

Wir beschreiben im folgenden zwei Beispielen von Abbildungen, die Polarkoordinaten in der Ebene und die linearen Abbildungen.

Wir schreiben $\mathbb{C}^*:=\mathbb{C}\setminus\{0\}$. Mit der Multiplikation von komplexen Zahlen ist $\mathbb{C}^*$ eine Gruppe $(\mathbb{C}^*,\cdot)$. Die Kreislinie $S^1:=\{z\in\mathbb{C}:|z|=1\}$ ist bezüglich der Multiplikation von komplexen Zahlen eine Untergruppe der Gruppe $(\mathbb{C}^*,\cdot)$.

8.4.9. **Satz** (Parametrisierung der Kreislinie). *(i) Die Abbildung* $p:\mathbb{R}\longrightarrow S^1$, $p(\varphi)=e^{i\varphi}=\cos\varphi+i\sin\varphi$ *ist ein Gruppenmorphismus der additiven Gruppe* $(\mathbb{R},+)$ *auf die multiplikative Gruppe* $(S^1,\cdot)$ *mit dem Kern* $2\pi\mathbb{Z}$. *Mit anderen Worten*

$$\forall z\in\mathbb{C},\,|z|=1\,\exists\varphi\in\mathbb{R}:e^{i\varphi}=z$$
$$e^{i\varphi_1}=e^{i\varphi_2}\Leftrightarrow\varphi_1-\varphi_2\in2\pi\mathbb{Z}$$

(ii) Sei $I\subset\mathbb{R}$ *ein halboffenes Intervall der Länge* 2π. *Dann ist* $p|_I:I\longrightarrow S^1$ *bijektiv und stetig. Die Umkehrfunktion von* $p:(-\pi,\pi]\longrightarrow S^1$ *bezeichnet man mit* $\arg:S^1\longrightarrow(-\pi,\pi]$. *Sie ist gegeben durch*

$$\arg:S^1\longrightarrow(-\pi,\pi],\,\arg z=\begin{cases}\arccos(\operatorname{Re}z), & \operatorname{Im}z\geqslant0,\\ -\arccos(\operatorname{Re}z), & \operatorname{Im}z<0.\end{cases}\tag{8.5}$$

$\arg:S^1\setminus\{-1\}\longrightarrow(-\pi,\pi)$ *ist stetig, aber* $\arg:S^1\longrightarrow(-\pi,\pi]$ *ist in* -1 *nicht stetig.*

Beweis: Für den Beweis der Stetigkeit von arg betrachte die offenen Mengen in S^1:

$$V_1=S^1\cap\{z:\operatorname{Im}z>0\},\quad V_2=S^1\cap\{z:\operatorname{Im}z<0\},\quad V_3=S^1\cap\{z:\operatorname{Re}z<0\}$$

und die stetigen Abbildungen

$$f_1:V_1\to\mathbb{R},\,f_1(z)=\arccos(\operatorname{Re}z),\quad f_2:V_2\to\mathbb{R},\,f_2(z)=-\arccos(\operatorname{Re}z),$$
$$f_3:V_3\to\mathbb{R},\,f_3(z)=\arcsin(\operatorname{Im}z).$$

(Verkettungen der stetigen Abbildungen $z\mapsto\operatorname{Re}z$ und arccos, bzw. $z\mapsto\operatorname{Im}z$ und arcsin). Es ist $S^1\setminus\{-1\}=V_1\cup V_2\cup V_3$ und

$$\arg z=\begin{cases}f_1(z), & z\in V_1,\\ f_2(z), & z\in V_2,\\ f_3(z), & z\in V_3.\end{cases}$$

also die zwei Definitionen in (8.5) "verkleben" sich stetigerweise in $z=1$. Formal, sei $z_0\in S^1\setminus\{-1\}$. Dann gibt es $k\in\{1,2,3\}$ mit $z_0\in V_k$. Da V_k offen ist, gilt

$$\lim_{\substack{z\to z_0\\ z\in S^1\setminus\{-1\}}}\arg(z)=\lim_{\substack{z\to z_0\\ z\in V_k}}f_k(z)=f_k(z_0)=\arg(z_0)$$

Andererseits gilt

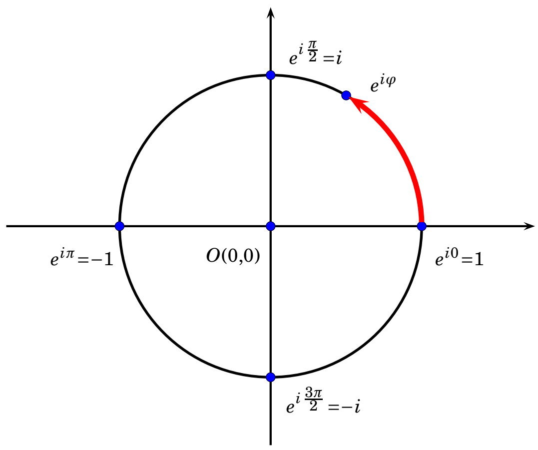
$$\lim_{\substack{z\to-1\\ \operatorname{Im}z<0}}\arg z=\lim_{x\to-1}(-\arccos x)=-\pi,\quad\lim_{\substack{z\to-1\\ \operatorname{Im}z>0}}\arg z=\lim_{x\to-1}\arccos x=\pi$$

also hat arg ein Sprung von 2π in $z=-1$. □

Die Aussage (i) kann man so umformulieren:

$$\forall(x,y)\in\mathbb{R}^2,\,x^2+y^2=1\,\exists\varphi\in\mathbb{R}:x=\cos\varphi,\,y=\sin\varphi$$
$$\cos\varphi_1=\cos\varphi_2,\,\sin\varphi_1=\sin\varphi_2\Leftrightarrow\varphi_1-\varphi_2\in2\pi\mathbb{Z}$$

Wir können uns $\varphi\mapsto p(\varphi)=e^{i\varphi}$ als Beschreibung der Bewegung eines Punktes vorstellen, der zum Zeitpunkt φ am Ort $e^{i\varphi}$ ist. Die Geschwindigkeit des Punktes ist $p'(\varphi)=ie^{i\varphi}$, ein Tangentialvektor an die beschriebene Kurve in $p(\varphi)$. Es gilt $|p'(\varphi)|=|ie^{i\varphi}|=1$ für alle φ.



8.4.10. **Definition.**

(a) Die Funktion $p : \mathbb{R} \longrightarrow S^1$, $p(\varphi) = e^{i\varphi}$ heißt die ***Standardparametrisierung der Kreislinie*** S^1.
(b) Für $z \in S^1$ heißt $\mathrm{Arg}(z) = \{\varphi \in \mathbb{R} : e^{i\varphi} = z\} = \arg z + 2\pi\mathbb{Z}$ die ***Menge der Argumente*** von z. Die Abbildung $\mathrm{Arg} : S^1 \longrightarrow \mathbb{R}/2\pi\mathbb{Z}$ ist ***mehrdeutig***, d.h. sie assoziiert zu $z \in S^1$ eine Menge von Werten in $\mathbb{R}$.
(c) Sei $I \subset \mathbb{R}$ ein halboffenes Intervall der Länge 2π. Die Umkehrung $(p|_I)^{-1} : S^1 \longrightarrow I$ heißt ein ***Zweig*** der Argument-Abbildung.
(d) Die Funktion $\arg = (p|_{(-\pi,\pi]})^{-1}$ heißt ***Hauptzweig*** von Arg.
(e) Wir erweitern nun den Definitionsbereich von Arg: Setze

$$\mathrm{Arg} : \mathbb{C}^* \longrightarrow \mathbb{R}/2\pi\mathbb{Z}, \mathrm{Arg}(z) := \mathrm{Arg}\big(\tfrac{z}{|z|}\big) = \{\varphi \in \mathbb{R} : e^{i\varphi} = \tfrac{z}{|z|}\}$$

Für $z \in \mathbb{C}^*$ heißt $\mathrm{Arg}(z)$ die Menge der Argumente von z.
(f) Die Funktion

$$\arg : \mathbb{C}^* \longrightarrow (-\pi,\pi], \quad \arg(z) := \arg\big(\frac{z}{|z|}\big) \tag{8.6}$$

heißt ***Hauptzweig*** von Arg.
(g) Für $z \in \mathbb{C}^*$ in der Form $\boxed{z = re^{i\varphi} = r(\cos\varphi + i\sin\varphi)}$ mit $r = |z|$ und $\varphi \in \mathrm{Arg}(z)$ (Polarkoordinatendarstellung von z (1.11)) heißen (r,φ) die ***Polarkoordinaten*** von z. Die Abbildung $\mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{C}^*$, $(r,\varphi) \mapsto re^{i\varphi}$ heißt ***Polarkoordinatenabbildung***.

Sei $\mathbb{R}_- = \{x \in \mathbb{R} : x < 0\}$ die negative reele Achse. Die Menge

$$\mathbb{C}_- := \mathbb{C}^* \setminus \mathbb{R}_- = \mathbb{C} \setminus \{z : \mathrm{Re}\, z \leqslant 0, \mathrm{Im}\, z = 0\}$$

heißt die ***geschlitzte Ebene***.

8.4.11. **Satz.** *Die Polarkoordinatenabbildung $P : \mathbb{R}_+ \times (-\pi,\pi] \longrightarrow \mathbb{C}^*$ ist bijektiv und stetig. Die Umkehrabbildung*

$$\mathbb{C}^* \longrightarrow \mathbb{R}_+ \times (-\pi,\pi], \quad z \longmapsto (|z|, \arg(z))$$

ist stetig auf $\mathbb{C}_-$ und unstetig in allen Punkten der negativen reellen Achse $\mathbb{R}_-$. Also ist

$$P : \mathbb{R}_+ \times (-\pi,\pi) \longrightarrow \mathbb{C}_-$$

ein Homöomorphismus.

Wir haben nun für $z \in \mathbb{C}^*$ zwei Koordinatensysteme: die kartesichen Koordinaten $(x,y) = (\mathrm{Re}\, z, \mathrm{Im}\, z)$ und die Polarkoordinaten $(r,\varphi) = (|z|, \arg(z))$. Für viele Probleme es ist vorteilhaft, die Polarkoordinaten zu benutzen.

Zum Beispiel ist die Multiplikation zweier in Polarkoordinaten $z = |z|e^{i\varphi}$, $w = |z|e^{i\psi}$ gegebener Zahlen besonders einfach: Es ist $zw = |z||w|e^{i(\varphi+\psi)}$. Daraus folgt die Formel von de Moivre (siehe (1.7)).

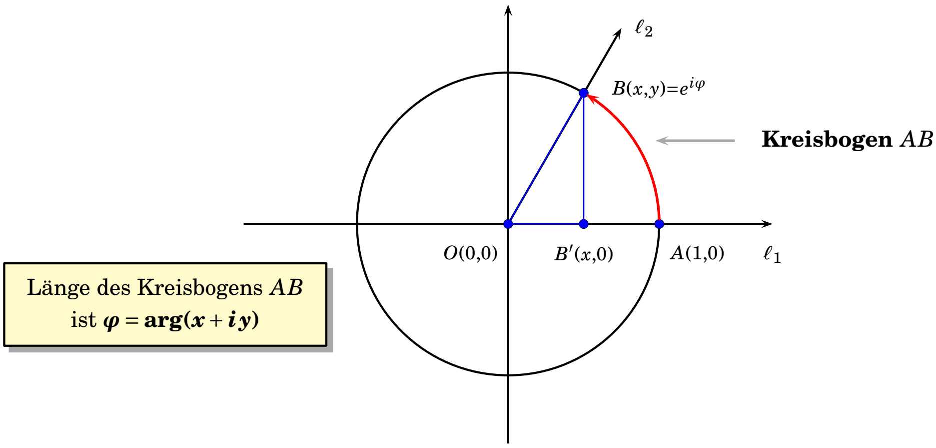
Geometrische Interpretation des Arguments: Definition des Winkelmaßes.

Seien l_1, l_2 zwei im Punkt O einer Ebene startende Strahlen. Wir legen die positive x-Achse auf l_1, so dass $O = (0,0)$. Sei $A = (1,0)$, $B = l_2 \cap S^1 = (x,y) \in \mathbb{R}^2 \cong \mathbb{C}$.

(Vorläufige) Definition des Winkelmaßes des Winkels $\sphericalangle l_1 l_2$ zwischen l_1 und l_2:

$$\sphericalangle l_1 l_2 := \sphericalangle AOB := \arg(b) = \varphi, \quad \text{wobei } b = x + iy.$$

Nach der Definition von *arg* gilt $\cos\varphi = x$ und $\sin\varphi = y$. Dies entspricht der üblichen geometrischen Definition von $\cos\varphi$ und $\sin\varphi$, wenn wir das rechtwinklige Dreieck $OB'B$ betrachten, wobei B' die Projektion von B auf die x-Achse ist.



Später beweisen wir, dass die Länge des Bogens AB genau $\varphi = \arg(b)$ beträgt (dafür brauchen wir das Kurvenintegral).

Dies können wir uns schon jetzt intuitiv so klarmachen: Wenn $\varphi \mapsto p(\varphi) = e^{i\varphi}$ die Bewegung eines Punktes beschreibt, so hat die Geschwindigkeit des Punktes den Betrag $|p'(\varphi)| = |ie^{i\varphi}| = 1$. Daher ist für $0 < \varphi < 2\pi$ die Länge des im Zeitintervall $[0,\varphi]$ durchlaufenen Weges gerade φ. Also ist φ die Bogenlänge des Kreisabschnitts von $1 = p(0)$ bis $p(\varphi)$ (für größere φ muss man die Bogenlänge mehrmals um den Kreis herummessen). Wegen $e^{i\frac{\pi}{2}} = i$ ist $\frac{\pi}{2}$ die Bogenlänge eines Viertels des Einheitskreises, also 2π die Bogenlänge des Einheitskreises. Dies entspricht der ***geometrischen Definition von*** π. Zur Kreiszahl π siehe auch [6, Kap. V].

8.4.12. **Satz** (n-te Einheitswurzeln). *Zu jeder natürlichen Zahl $n \in \mathbb{N}$ gibt es genau n verschiedene komplexe Nullstellen des Polynoms $z^n - 1$, nämlich*

$$\zeta_k = e^{i\frac{2\pi}{n}k} = (e^{i\frac{2\pi}{n}})^k, \quad \text{für } 0 \leqslant k < n.$$

Zu jedem $z_0 \in \mathbb{C} \setminus \{0\}$ gibt es genau n verschiedene komplexe Nullstellen des Polynoms $z^n - z_0$, nämlich $\zeta_k = \sqrt[n]{|z_0|}e^{i\frac{(\arg z_0 + 2\pi k)}{n}}$ für $0 \leqslant k < n$.

8.4.13. **Definition.** Seien X, Y topologische Räume, $A \subset X$ und a ein HP von A. Eine Funktion $f : A \longrightarrow Y$ hat in a den Grenzwert $b \in Y$, wenn es zu jeder Umgebung V von b eine Umgebung U von a gibt so, dass $f(A \cap U \setminus \{x_0\}) \subset V$. Schreibweise: $\lim_{x \to a} f(x) = b$.

8.4.14. **Bemerkung.** Sei $a \in A$ und a ein HP von A ist. Dann gilt: f ist stetig in a genau dann, wenn $\lim_{x \to a} f(x) = f(a)$.

8.4.15. **Satz.** *Seien X, Y metrische Räume, $A \subset X$, a ein HP von A, $b \in Y$ und $f : A \longrightarrow Y$. Die folgenden Aussagen sind äquivalent:*

(i) $\lim_{x \to a} f(x) = b$
(ii) *(ε-δ-Kriterium) Für alle $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle $x \in A$ mit $x \neq a$ und $d(x,a) < \delta$ gilt: $d(f(x), b)) < \varepsilon$*
(iii) *(Folgenkriterium) Für alle Folgen $(x_n)_n$ mit $x_n \in A \setminus \{a\}$ und $x_n \to a$ folgt $f(x_n) \to b$ für $n \to \infty$.*

8.4.16. **Satz.** *Sei $f : A \longrightarrow Y_1 \times \ldots \times Y_m$, $f = (f_1, \ldots, f_m)$, wobei $Y_1 \times \ldots \times Y_m$ ein Produkt von topologischen Räumen sei, versehen mit der Produkttopologie. Genau dann existiert der Grenzwert $\lim_{x \to a} f(x)$ und ist gleich $b = (b_1, \ldots, b_m) \in Y_1 \times \ldots \times Y_m$, wenn für alle $i = 1, \ldots, m$ der Grenzwert $\lim_{x \to a} f_i(x)$ existiert und gleich b_i ist.*

8.5. **Stetige lineare Abbildungen.**

8.5.1. **Satz.** *Seien V, W normierte Vektorräume, $T : V \to W$ eine lineare Abbildung. Dann sind äquivalent:*

(a) T ist stetig.
(b) T ist stetig in $0 \in V$.
(c) Es gibt $C \geqslant 0$, so dass für alle $x \in V$ gilt: $\|T(x)\| \leqslant C\|x\|$.

Wir setzen $\mathcal{L}(V, W) = \{T : V \to W : T \text{ linear und stetig}\}$.

8.5.2. **Satz.** *Ist V endlichdimensional, so ist jede lineare Abbildung $T : V \to W$ stetig, d.h. $\mathcal{L}(V,W) = \mathrm{Hom}(V,W)$. Insbesondere ist jeder Isomorphismus $T : V \to W$ zwischen endlichdimensionalen Vektorräumen stetig (bzgl. der Standardtopologien auf V und W).*

8.5.3. **Satz.** *Die Zahl*

$$\|T\| = \sup_{x \neq 0} \frac{\|T(x)\|}{\|x\|} = \sup_{\|x\| \leqslant 1} \|T(x)\| = \sup_{\|x\|=1} \|T(x)\|$$

definiert eine Norm auf $\mathcal{L}(V,W)$, genannt die ***Operatornorm****. Es gilt*

$$\|T\| = \inf\{C \geqslant 0 : \|T(x)\| \leqslant C\|x\| \textit{ für alle } x \in V\}$$

d.h. $\|T\|$ ist die kleinste nicht-negative reelle Zahl C mit $\|T(x)\| \leqslant C\|x\|$ für alle $x \in V$. Insbesondere gilt

$$\|T(x)\| \leqslant \|T\| \cdot \|x\|, \quad \textit{für alle } x \in V.$$

Wenn nichts anders gesagt, benutzen wir die Operatornorm auf $\mathcal{L}(V,W)$.

8.5.4. **Satz.** *Seien V, W, U NVR und $T \in \mathcal{L}(V,W)$, $S \in \mathcal{L}(W,U)$. Dann ist $S \circ T \in \mathcal{L}(V,U)$ und gilt*

$$\|S \circ T\| \leqslant \|S\| \, \|T\| \tag{8.7}$$

8.5.5. **Satz.** *Seien V ein NVR und W ein Banachraum. Dann ist $\mathcal{L}(V,W)$ versehen mit der Operatornorm auch ein Banachraum. Insbesondere ist $\mathcal{L}(V) := \mathcal{L}(V,V)$ eine Banachalgebra, falls V ein Banachraum ist.*

Aus Satz 8.2.17 erhalten wir:

8.5.6. **Satz.** *Sei V ein Banachraum und $T \in \mathcal{L}(V,W)$ mit $\|T\| \leqslant 1$. Dann ist $\mathrm{Id} - T$ invertierbar und*

$$(\mathrm{Id} - T)^{-1} = \sum_{n=0}^{\infty} T^n = \mathrm{Id} + T + T^2 + \dots$$

$$\|(\mathrm{Id} - T)^{-1}\| \leqslant \frac{1}{1 - \|T\|}, \quad \|(\mathrm{Id} - T)^{-1} - \mathrm{Id}\,\| \leqslant \frac{\|T\|}{1 - \|T\|}$$

$$\|(\mathrm{Id} - T)^{-1} - \mathrm{Id} - T\| \leqslant \frac{\|T\|^2}{1 - \|T\|}.$$

Die Reihe $\sum_{n=0}^{\infty} T^n$ heißt in diesem Zusammenhang auch ***Neumannsche Reihe***.

Sei nun V ein Banachraum. Nach einem wichtigen Satz von Banach gilt: ist T linear, stetig ($T \in \mathcal{L}(V)$) und invertierbar, so ist auch T^{-1} stetig d.h. $T^{-1} \in \mathcal{L}(V)$ (dies ist trivial, wenn V endlichdimensional ist). Wir bezeichnen mit $Gl(V) := \{T \in \mathcal{L}(V) : T \text{ invertierbar}\}$.

8.5.7. **Satz.** *Sei V ein Banachraum und $T \in Gl(V)$.*
(i) $Gl(V)$ offen in $\mathcal{L}(V)$. Genauer, sei $H \in \mathcal{L}(V)$ mit $\|H\| < 1/\|T^{-1}\|$. Dann ist $T+H$ invertierbar also besteht die offene Kugel in $\mathcal{L}(V)$ mit Mittelpunkt T und Radius $1/\|T^{-1}\|$ nur aus invertierbaren Elementen.
(ii) Die Abbildung $Gl(V) \to Gl(V)$, $T \mapsto T^{-1}$ ist stetig. Genauer, für $\|H\| \leqslant \delta/\|T^{-1}\|$, wobei $\delta < 1$, gilt

$$\|(T+H)^{-1} - T^{-1}\| \leqslant \frac{\|T^{-1}\|^2}{1-\delta} \cdot \|H\| \tag{8.8}$$

Beweis: Es gilt $T + H = (\mathrm{Id} + HT^{-1})T$. Wegen $\|H\|\|T^{-1}\| < 1$ ist $(\mathrm{Id} + HT^{-1})$ invertierbar und somit auch $T + H$, mit $(T+H)^{-1} = T^{-1}(\mathrm{Id} + HT^{-1})^{-1}$. Weiter

$$(T+H)^{-1} - T^{-1} = T^{-1}(\mathrm{Id} + HT^{-1})^{-1} - T^{-1} = T^{-1}\big((\mathrm{Id} + HT^{-1})^{-1} - \mathrm{Id}\big)$$

und nach Satz 8.5.6

$$\big\|(\mathrm{Id} + HT^{-1})^{-1} - \mathrm{Id}\,\big\| \leqslant \frac{\|HT^{-1}\|}{1 - \|HT^{-1}\|} \leqslant \|T^{-1}\| \cdot \frac{\|H\|}{1 - \|H\|\,\|T^{-1}\|}$$

Für $\|H\| \leqslant \delta/\|T^{-1}\|$ erhalten wir (8.8). □

8.5.8. **Satz.** *Seien $V_1, \dots, V_n, W$ normierte Vektorräume. Für eine multilineare Abbildung $T : V_1 \times \dots \times V_n \longrightarrow W$ sind äquivalent:*

(i) *T ist stetig,*
(ii) *T ist stetig in $(0, \dots, 0)$,*
(iii) $\exists C \geqslant 0\ \forall v_1 \in V_1, \dots, v_n \in V_n : \|T(v_1, \dots, v_n)\| \leqslant C\|v_1\| \cdot \|v_2\| \cdot \ldots \cdot \|v_n\|$.

Wir bezeichnen mit $\mathcal{L}(V_1,\dots,V_n;W)$ der Raum der stetigen multilinearen Abbildungen. Die durch $\|T\| :=$ $\sup\{\|T(v_1,\dots,v_n)\| : \|v_1\| \leqslant 1,\dots,\|v_n\| \leqslant 1\}$ *definierte Abbildung* $\|\cdot\| : \mathcal{L}(V_1,\dots,V_n;W) \to \mathbb{R}$ *ist eine Norm auf* $\mathcal{L}(V_1,\dots,V_n;W)$.

Sind $V_1,\dots,V_n$ endlichdimensional, so ist jede multilineare Abbildung $T : V_1 \times \ldots \times V_n \to W$ stetig.

8.6. **Kompaktheit.** Sei X eine Menge, $(V_i)_{i\in I}$ eine Familie von Teilmengen von X (d.h. eine Abbildung $I \longrightarrow \mathcal{P}(X)$, $i \mapsto V_i$). Wir setzen $\cup_{i\in I} V_i := \{x \in X : \exists\, i \in I, x \in V_i\}$. Ist $A \subset X$, so heißt $(V_i)_{i\in I}$ ***Überdeckung*** von A, falls $A \subset \cup_{i\in I} V_i$. Ist X ein topologischer Raum und sind die V_i offen, so heißt $(V_i)_{i\in I}$ eine ***offene Überdeckung*** von A.

8.6.1. **Definition.** Ein topologischer Raum X heißt ***kompakt***, wenn X die ***Heine-Borel-Über-deckungseigenschaft*** hat, d.h. wenn aus jeder offenen Überdeckung $(V_i)_{i\in I}$ von X endlich viele $i_1,\dots,i_k$ ausgewählt werden können so, dass schon $X = \cup_{r=1}^k V_{i_r}$. Die Familie $(V_{i_1},\dots,V_{i_k})$ heißt Teilüberdeckung, und die Heine-Borel-Überdeckungseigenschaft kann auch so formuliert werden: Jede offene Überdeckung besitzt eine endliche Teilüberdeckung[10].

Eine *Teilmenge* $K \subset X$ heißt *kompakt*, wenn sie zusammen mit der Teilraumtopologie ein kompakter topologischer Raum ist (dies bedeutet, dass aus jeder Familie $(V_i)_{i\in I}$ offener Mengen in X mit $K \subset \cup_{i\in I} V_i$ endlich viele $i_1,\dots,i_k$ ausgewählt werden können so, dass $K \subset V_{i_1} \cup \ldots \cup V_{i_k}$. Wieso ist das zur Definition äquivalent?).

Eine Teilmenge K heißt ***relativ kompakt***, falls $\overline{K}$ kompakt ist.

Da Kompaktheit einer Teilmenge K eine Eigenschaft des Raumes $(K, \mathcal{O}_K)$ ist, ist sie ein *intrinsischer Begriff*. Zum Beispiel, die Menge $\overline{A}$ aus 8.3.17 (v) ist kompakt sowohl als Teilmenge von $\mathbb{R}^3$ als auch als Teilmenge von E.

8.6.2. **Definition.** Sei X ein metrischer Raum. X heißt ***folgenkompakt***, wenn X die ***Bolzano-Weierstraß-Eigenschaft*** hat, d.h. jede Folge in X hat eine konvergente Teilfolge.

Eine *Teilmenge* $K \subset X$ heißt *folgenkompakt*, falls sie zusammen mit der Teilraumtopologie folgenkompakt ist (dies bedeutet, dass jede Folge in K eine konvergente Teilfolge hat, deren Grenzwert in K liegt).

Der metrische Raum X heißt ***total beschränkt*** oder ***präkompakt***, wenn es für jedes $\varepsilon > 0$ eine endliche Teilmenge $\{x_1,\dots,x_k\} \subset X$ gibt mit $X \subset \cup_{i=1}^k B_\varepsilon(x_i)$.

8.6.3. **Satz** (Bolzano-Weierstraß)**.** *Sei X ein metrischer Raum. Die folgenden Bedingungen sind äquivalent zueinander:*

(i) *X ist kompakt.*
(ii) *X ist folgenkompakt.*
(iii) *X ist vollständig und total beschränkt.*

Siehe dazu auch Aufgabe 8.8.15 (Lebesgue Lemma).

8.6.4. **Satz.** *Sei X ein metrischer Raum. Ist $K \subset X$ folgenkompakt, so ist K abgeschlossen und beschränkt.*

Die Umkehrung ist i.A. nicht wahr. Sei $(V, \langle\cdot,\cdot\rangle)$ ein Skalarproduktraum, $\dim_{\mathbb{C}} V = \infty$. Sei $(e'_k)_{k\in\mathbb{N}}$ eine Folge linear unabhängiger Vektoren. Durch das Schmidtsche Orthonormalisierungsverfahren können wir aus $(e'_k)_{k\in\mathbb{N}}$ ein ***Orthonormalsystem*** $(e_k)_{k\in\mathbb{N}}$ konstruieren, d.h. $\langle e_j, e_k\rangle = \delta_{jk}$ für alle $j,k \in \mathbb{N}$. Wegen $\|e_j - e_k\| = \sqrt{2}$ für alle $j \neq k$ kann $(e_k)_{k\in\mathbb{N}}$ keine Cauchy-Teilfolge enthalten, also erst recht keine konvergente Teilfolge. Trotzdem ist $(e_k)_{k\in\mathbb{N}}$ eine Folge in der beschränkten und abgeschlossenen Menge $\overline{B}_1(0) \subset V$.

8.6.5. **Satz** (Heine-Borel)**.** *Sei V ein endlichdimensionaler normierter Vektorraum. Sei $K \subset V$. Dann sind die folgenden Aussagen äquivalent zueinander:*

(i) *K ist kompakt.*
(ii) *K ist folgenkompakt.*
(iii) *K ist abgeschlossen und beschränkt.*

[10] „Teil“ heißt nicht, dass nur ein Teil von X überdeckt wird, sondern dass man nur eine Teilmenge der Indizes benutzt.

8.6.6. **Beispiel.** Wir haben gesehen in Beispiel 8.4.8, dass S^n abgeschlossen ist. Da sie auch beschränkt ist, folgern wir nach Heine-Borel, dass S^n kompakt ist.

8.6.7. **Satz.** *Seien X,Y topologische Räume, X kompakt und $f: X \longrightarrow Y$ stetig. Dann ist $f(X)$ kompakt.*

8.6.8. **Satz** (vom Maximum und Minimum)**.** *Sei X ein kompakter topologischer Raum, und sei $f: X \longrightarrow \mathbb{R}$ stetig. Dann ist f beschränkt und nimmt auf X ihr Maximum und Minimum an, d.h. $\exists \xi_1, \xi_2 \in X$ mit $f(\xi_1) = \inf f(X)$, $f(\xi_2) = \sup f(X)$.*

Satz 8.6.8 verallgemeinert Satz 4.4.3, weil Intervalle $[a,b] \subset \mathbb{R}$ mit $a,b \in \mathbb{R}$ kompakt sind. Für eine Anwendung auf die Abstand zwischen zwei Mengen siehe Aufgabe 8.8.17.

Zum Schluß beweisen wir den folgenden oft nützlichen Satz:

8.6.9. **Satz** (Lindelöf)**.** *Sei X ein topologischer Raum mit abzählbarer Basis. Jede offene Überdeckung von X enthält eine abzählbare Teilüberdeckung.*

Beweis: Sei $\mathcal{B}$ eine abzählbare Basis von X und $\mathcal{U}$ eine offene Überdeckung von X. Sei $\mathcal{B}' = \{B \in \mathcal{B} : \exists U \in \mathcal{U}, B \subset U\}$; da $\mathcal{B}' \subset \mathcal{B}$, ist $\mathcal{B}'$ abzählbar. Wähle für jedes $B \in \mathcal{B}'$ ein $U_B \in \mathcal{U}$ mit $B \subset U_B$. Sei $\mathcal{U}' = \{U_B : B \in \mathcal{B}'\}$. $\mathcal{U}'$ ist eine abzählbare Überdeckung von X: Sei $x \in X$. Dann gibt es $U_0 \in \mathcal{U}$ mit $x \in U_0$. Da $\mathcal{B}$ eine Basis ist, ist U_0 eine Vereinigung von Elementen aus $\mathcal{B}$, d.h. es gibt $B \in \mathcal{B}$ mit $x \in B \subset U_0$. Also gilt $B \in \mathcal{B}'$ und existiert $U_B \in \mathcal{U}'$ mit $x \in B \subset U_B$. □

8.7. **Zusammenhang.** Die Rolle der Intervalle in $\mathbb{R}$ wird in metrischen Räumen übernommen von den sogenannten zusammenhängenden Mengen, die wir jetzt kennenlernen.

8.7.1. **Definition.** Sei X ein topologischer Raum.

(i) X heißt ***zusammenhängend***, wenn es keine Zerlegung von X in zwei nichtleere disjunkte offene Teilmengen U,V gibt.

(ii) $A \subset X$ heißt ***zusammenhängend***, falls A versehen mit der Teilraumtopologie zusammenhängend ist.

X zusammenhängend $\Longleftrightarrow$ für alle offenen Teilmengen $U,V \subset X$ mit $U \cap V = \emptyset$ und $U \cup V = X$ gilt $U = \emptyset$ oder $V = \emptyset$ $\Longleftrightarrow$ $\emptyset$ und X sind die einzigen zugleich offen und abgeschlossen Teilmengen von X.

$A \subset X$ zusammenhängend $\Longleftrightarrow$ für alle offenen Teilmengen $U,V \subset X$ mit $A \subset U \cup V$ und $A \cap U \cap V = \emptyset$ gilt $A \cap U = \emptyset$ oder $A \cap V = \emptyset$.

Beispiele:

(i) Sei $f: \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}$, $f(x) = \frac{1}{x}$. Graph(f) ist nicht zusammenhängend.

(ii) $\mathbb{Q} \subset \mathbb{R}$ ist nicht zusammenhängend. Sogar: Für alle $x,y \in \mathbb{Q}$ mit $x \neq y$ gibt es $z \notin \mathbb{Q}$ zwischen x und y. Eine Teilmenge von $\mathbb{R}$ mit dieser Eigenschaft heißt Diskontinuum. Die Cantorsche Menge ist auch ein Diskontinuum.

Und nun ein positives Beispiel:

8.7.2. **Satz.** *$A \subset \mathbb{R}$ ist zusammenhängend genau dann, wenn A ein Intervall ist.*

8.7.3. **Satz.** *Seien X,Y topologische Räume, $f: X \longrightarrow Y$ stetig, X zusammenhängend. Dann ist $f(X)$ zusammenhängend.*

8.7.4. **Satz** (Verallgemeinerter Zwischenwertsatz)**.** *Sei $f: X \longrightarrow \mathbb{R}$ stetig, $A \subset X$ zusammenhängend. Dann gibt es zu allen $x,y \in A$ und jedem c zwischen $f(x)$ und $f(y)$ ein $z \in A$ mit $f(z) = c$.*

8.7.5. **Definition.** Sei X ein topologischer Raum.

(i) X heißt ***wegzusammenhängend***, wenn es zu allen $x,y \in X$ eine stetige Abbildung $\gamma: [a,b] \longrightarrow X$ mit $\gamma(a) = x$, $\gamma(b) = y$ gibt (solch ein γ heißt stetige Kurve).

(ii) $A \subset X$ heißt ***wegzusammenhängend***, wenn $(A, \mathcal{O}_A)$ wegzusammenhängend ist.

8.7.6. **Beispiel.** (i) Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Eine Teilmenge $A \subset V$ heißt ***konvex***, wenn für alle $x,y \in A$ die Verbindungsstrecke $[x,y] = \{(1-t)x + ty : t \in [0,1]\}$ in A enthalten ist. Zum Beispiel ist eine Kugel $B_r(x)$ in einem normierten Vektorraum konvex. Da $[0,1] \ni t \mapsto (1-t)x + ty \in V$ stetig ist, ist jede konvexe Menge auch wegzusammenhängend.

Hat man mehrere Punkte $x_0, x_1, \ldots, x_k$ in einem NVR V so bilden die einzelnen Strecken $[x_0,x_1]$, $[x_1,x_2]$,..., $[x_{k-1},x_k]$ zusammen einen ***Streckenzug***:

$$S(x_0,x_1,\ldots,x_k) = [x_0,x_1] \cup [x_1,x_2] \cup [x_{k-1},x_k].$$

(ii) Für $n \geqslant 2$ ist $\mathbb{R}^n \setminus \{0\}$ nicht konvex, aber offensichtlich wegzusammenhängend. Zu $x, y \in \mathbb{R}^n \setminus \{0\}$ gibt es nämlich $z \in \mathbb{R}^n \setminus \{0\}$, so dass $[x,z] \cup [z,y] \subset \mathbb{R}^n \setminus \{0\}$.

(iii) Für $n \geqslant 1$ ist die n-Sphäre $S^n = \{x \in \mathbb{R}^{n+1} : x_1^2 + \ldots + x_{n+1}^2 = 1\}$ wegzusammenhängend.

8.7.7. **Satz.** *Ist X wegzusammenhängend, dann ist X zusammenhängend.*

8.7.8. **Bemerkung.** Es gibt zusammenhängende Räume, die nicht wegzusammenhängend sind! Sei $A = \{(x, \sin\frac{1}{x}) : x > 0\}$, dann $\overline{A} = \{(x, \sin\frac{1}{x}) : x > 0\} \cup \{(0,y) : y \in [-1,1]\}$. A ist zusammenhängend, da $A = \gamma(\mathbb{R}_+)$, $\gamma(x) = (x, \sin\frac{1}{x})$, γ stetig; dann ist auch $\overline{A}$ zusammenhängend. Aber $\overline{A}$ ist nicht wegzusammenhängend! Es gibt keine stetige Kurve $c : [a,b] \to \overline{A}$ in $\overline{A}$ von $(0,0)$ nach $(x,y) \neq (0,0)$. Eine solche Kurve müsste zwischen den längs der x-Achse gegen 0 hin immer enger nebeneinander liegenden Extremstellen mit den Werten $1, -1$ hin- und herspringen! Dies würde der Stetigkeit von c widersprechen.

Sei $(V, \|\cdot\|)$ ein NVR. Eine offene und zusammenhängende Teilmenge $D \subset V$ heißt ***Gebiet***.

8.7.9. **Satz.** *Jedes Gebiet D in einem NRV ist wegzusammenhängend. Genauer: für jede $x, y \in D$ gibt es einen Streckenzug $[x_0,x_1] \cup [x_1,x_2] \cup \ldots \cup [x_{k-1},x_k] \subset D$ wobei $x_0 = x$ und $x_k = y$.*

8.7.10. **Satz.** *Für $n \geqslant 2$ ist $\mathbb{R}^n$ nicht homöomorph zu $\mathbb{R}$.*

8.7.11. **Bemerkung.** $(0,1) \simeq \mathbb{R}$ und $(0,1)^2 \simeq \mathbb{R}^2$ (wobei $X \simeq Y$ bedeutet: X homöomorph zu Y), also sind $(0,1)$ und $(0,1)^2$ nicht homöomorph zueinander. Dies ist kein Zufall:

8.7.12. **Satz** (Dimensionsinvarianz durch Homöomorphismen; Brouwer)**.** *Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ homöomorphe nicht-leere offene Mengen. Dann gilt $m = n$.*

Trotzdem gibt es surjektive stetige Abbildungen $c : [0,1] \longrightarrow [0,1]^2$, sogenannten ***Peanokurven*** oder flächenfüllende Kurven (siehe z.B. [19]).

8.7.13. **Definition** (Zusammenhangskomponente, Wegekomponente)**.** Sei X ein topologischer Raum. Die Vereinigung aller zusammenhängenden Teilmengen A von X, die $x \in X$ enthalten, heißt ***Zusammenhangskomponente*** $X(x)$ von x. Wir sprechen auch kurz von ***Komponenten***. Die Vereinigung aller wegzusammenhängenden Teilmengen A von X, die $x \in X$ enthalten, heißt ***Wegekomponente*** von x. Ein Raum X heißt ***lokal (weg-) zusammenhängend***, wenn zu jedem $x \in X$ und jeder Umgebung U von x eine (weg-) zusammenhängendende Umgebung $V \subset U$ existiert.

8.7.14. **Bemerkung.**

(1) Die Zusammenhangskomponenten von X sind zusammenhängende abgeschlossene Teilmengen von X (siehe die Aufgabe nach Satz 8.7.4).

(2) Ist $y \in X(x)$, so gilt $X(x) = X(y)$. Eine Komponente ist eine maximale zusammenhängende Teilmenge von X. Ein Raum ist disjunkte Vereinigung seiner Komponenten.

(3) Ein Punkt y ist in der Wegekomponente von x, genau dann, wenn x und y durch eine stetige Kurve in X verbindbar sind. Eine Wegekomponente ist eine maximale wegzusammenhängende Teilmenge von X. Ein Raum ist disjunkte Vereinigung seiner Wegekomponenten.

(4) Die offenen Mengen eines NVR sind lokal wegzusammenhängend (nach Satz 8.7.7 auch lokal zusammenhängend).

8.7.15. **Satz.**

(i) Die Komponenten eines lokal zusammenhängenden Raumes sind offen.

(ii) Die Wegekomponenten eines lokal wegzusammenhängenden Raumes sind offen und stimmen mit den Komponenten überein.

(iii) Die Aussage (ii) trifft insbesondere für offene Mengen in einem NVR. Eine solche Menge hat höchstens abzählbar viele Komponenten (= Wegwkomponenten) und diese sind offen. Die Komponenten einer offenen Menge in $\mathbb{R}$ sind offene Intervalle.

8.8. Übungen.

8.8.1. ***Aufgabe***. (X_1,d_1) und (X_2,d_2) seien metrische Räume. Für $x, y \in X_1 \times X_2$ sei $d(x,y) := \max\{d_1(x_1,y_1), d_2(x_2,y_2)\}$. Zeige, daß d eine Metrik auf $X_1 \times X_2$ ist.

8.8.2. ***Aufgabe***. Sei (X,d) ein metrischer Raum. Beweise:

(a) Für alle $x,y,z \in X$ gilt $|d(x,y) - d(y,z)| \leqslant d(x,z)$.

(b) Gilt $\lim\limits_{n\to\infty} x_n = x$ in (X,d), so gilt $\lim\limits_{n\to\infty} d(x_n,y) = d(x,y)$ für alle $y \in X$.

8.8.3. ***Aufgabe***. Sei $(V,\langle\cdot,\cdot\rangle)$ ein Skalarproduktraum. Zeigen Sie, dass für alle $x,y\in V$ die Polarisierungs-Beziehungen

$$\langle x,y\rangle = \frac{1}{4}(\|x+y\|^2-\|x-y\|^2),\quad \text{falls } V \text{ reell,}$$
$$\langle x,y\rangle = \frac{1}{4}(\|x+y\|^2-\|x-y\|^2+i\|x+iy\|^2-i\|x-iy\|^2),\quad \text{falls } V \text{ komplex,}$$

und die Parallelogramm-Formel

$$\|x+y\|^2+\|x-y\|^2=2(\|x\|^2+\|y\|^2).$$

gelten. Sei $(V,\|\cdot\|)$ ein normierter Raum. Die Norm $\|\cdot\|$ ist genau dann durch ein Skalarprodukt induziert, wenn die Parallelogramm-Formel gilt. Zeigen Sie, dass die Norm $\|\cdot\|_1$ auf $\mathbb{R}^n$ nicht von einem Skalarprodukt induziert wird.

8.8.4. ***Aufgabe***. Für $A\in M_{n\times n}(\mathbb{C})$, $A=(a_{ij})_{1\leqslant i,j\leqslant n}$, definiere $\|A\|_2 := \sqrt{\sum\limits_{i,j=1}^{n}|a_{ij}|^2}$.

(a) Zeige, daß $d(A,B):=\|A-B\|_2$ eine Metrik auf $M_{n\times n}(\mathbb{C})$ definiert.

(b) Sei $(A_k)_{k\in\mathbb{N}}$, $A_k=(a_{ij}^{(k)})_{1\leqslant i,j\leqslant n}$, eine Folge in $M_{n\times n}(\mathbb{C})$. Zeige, daß

$$\lim_{k\to\infty}A_k=A\quad\text{genau dann, wenn}\quad \lim_{k\to\infty}a_{ij}^{(k)}=a_{ij}\ \text{für alle } 1\leqslant i,j\leqslant n\ .$$

Folgere, daß $(M_{n\times n}(\mathbb{C}),d)$ ein vollständiger metrischer Raum ist.

(c) Beweise, daß $\|AB\|_2\leqslant\|A\|_2\cdot\|B\|_2$ für alle $A,B\in M_{n\times n}(\mathbb{C})$.

(d) Zeige, daß für jedes $A\in M_{n\times n}(\mathbb{C})$ die Reihe $\exp(A):=\sum\limits_{k\geqslant 0}\frac{1}{k!}A^k$ konvergiert.

(e) Berechne $\exp\begin{pmatrix}0 & -t\\ t & 0\end{pmatrix}$.

(f) Zeige: Wenn $AB=BA$, dann $\exp(A+B)=\exp(A)\exp(B)$.

8.8.5. ***Aufgabe***. Sei $D\subset\mathbb{C}$, und sei $V=\mathcal{C}_b^0(D,\mathbb{C})$ der $\mathbb{C}$-Vektorraum der beschränkten stetigen Funktionen $f:D\to\mathbb{C}$.

(a) Zeigen Sie, dass V mit der von der Supremumsnorm $\|\cdot\|_D$ induzierten Metrik ein vollständiger metrischer Raum ist.

(b) Sei $D=[a,b]\subset\mathbb{R}$ und $\|\cdot\|_1$ die durch $\|f\|_1:=\int_a^b|f(x)|\,dx$ definierte Norm auf V. Ist V mit der induzierten Metrik vollständig?

(Tipp: Betrachten Sie $f_n:[0,1]\to\mathbb{C}$, $f_n=0$ auf $[0,\frac12-\frac1n]$, $f_n=1$ auf $[\frac12+\frac1n,1]$ und linear auf $[\frac12-\frac1n,\frac12+\frac1n]$; zeigen Sie, dass $(f_n)_n$ eine Cauchy-Folge ist, die nicht konvergiert. Würde ein Grenzwert $f\in V$ existieren, so könnte man zeigen, dass $f=0$ auf $[0,\frac12)$, $f=1$ auf $(\frac12,1]$, so dass f nicht stetig wäre.)

8.8.6. ***Aufgabe***. Sei $p\geq 1$, und sei ℓ^p die Menge aller Folgen $(z_\nu)_{\nu\in\mathbb{N}}$ in $\mathbb{C}$ mit der Eigenschaft, dass $\sum_{\nu=1}^\infty|z_\nu|^p$ konvergiert. Zeigen Sie:

(a) Mit $z,w\in\ell^p$ und $\lambda\in\mathbb{C}$ liegen auch $z+w$ und λz in ℓ^p, wobei Addition und Multiplikation mit Skalaren gliedweise definiert sind. (Insbesondere ist ℓ^p ein $\mathbb{C}$-Vektorraum.)

(b) Durch $\ell^p\ni z=(z_\nu)_{\nu\in\mathbb{N}}\mapsto\|z\|_p:=\left(\sum_{\nu=1}^\infty|z_\nu|^p\right)^{1/p}$ ist eine Norm auf ℓ^p gegeben.

(c) $(\ell^p,\|.\|_p)$ ist ein Banachraum.

8.8.7. ***Aufgabe***. Wir betrachten die Normen $\|z\|_\infty=\max\{|z_1|,\dots,|z_n|\}$, $\|z\|_2=\sqrt{|z_1|^2+\ldots+|z_n|^2}$ und $\|z\|_1=|z_1|+\ldots+|z_n|$ auf $\mathbb{C}^n$. Finden Sie maximale Konstanten c_n,c_n' und minimale Konstanten C_n,C_n' mit

(a) $c_n\|\cdot\|_\infty\leqslant\|\cdot\|_2\leqslant C_n\|\cdot\|_\infty$, (b) $c_n'\|\cdot\|_1\leqslant\|\cdot\|_2\leqslant C_n'\|\cdot\|_1$.

Man kann zeigen, dass für $1\leqslant p\leqslant q$ und alle $x\in\mathbb{R}^n$ gilt $\|x\|_q\leqslant\|x\|_p\leqslant n^{\frac1p-\frac1q}\|x\|_q$.

8.8.8. ***Aufgabe***. Sei $\mathcal{C}^1([0,1])$ der Vektorraum der stetig differenzierbaren Funktionen auf $[0,1]$. Definiere $\|f\|_1=\int_0^1|f(t)|\,dt$, $\|f\|=\sup_{t\in[0,1]}|f(t)|$, $|||f|||=|f(0)|+\sup_{t\in[0,1]}|f'(t)|$.

(a) Zeigen Sie, dass $|||\cdot|||$ eine Norm ist.

(b) Zeigen Sie, dass eine bzgl. $\|\cdot\|$ konvergente Folge auch bzgl. $\|\cdot\|_1$ konvergiert und dass eine bzgl. $|||\cdot|||$ konvergente Folge auch bzgl. $\|\cdot\|$ konvergiert.

(c) Untersuchen Sie die Konvergenz der Folgen $f_n(t) = t^n$, $g_n(t) = \frac{1}{n}\sin nt$ bzgl. der drei Normen. Sind diese Normen äquivalent?

8.8.9. ***Aufgabe***. Sei (X,d) ein kompakter metrischer Raum. Zeigen Sie, dass (X,d) vollständig ist.

8.8.10. ***Aufgabe***. Sei (X,d) ein metrischer Raum, $a \in X$ und $r > 0$. Sei $\overline{B}_r(a) = \{x \in X : d(x,a) \leqslant r\}$ die abgeschlossene Kugel mit Zentrum a und Radius r und sei $\overline{B_r(a)}$ der Abschluss der offenen Kugel $B_r(a)$. Zeigen Sie:

(a) $\overline{B_r(a)} \subset \overline{B}_r(a)$.

(b) Ist $(X, \|\cdot\|)$ ein NVR, so gilt $\overline{B_r(a)} = \overline{B}_r(a)$.

Gilt es Gleichheit im Allgemeinen? (Hinweis: Betrachte einen metrischer Raum mit mehr als ein Punkt und diskreten Metrik.)

8.8.11. ***Aufgabe***. Sei X ein Hausdorffraum und $K \subset X$ kompakt. Zeigen Sie:

(a) K ist abgeschlossen.

(b) Ist F eine abgeschlossene Menge mit $K \cap F = \emptyset$, so existiert eine offene Menge U mit $K \subset U$ und $U \cap F = \emptyset$.

Gilt (a) im Allgemeinen? (Hinweis: Betrachte einen topologischer Raum mit der groben Topologie und mehr als zwei Punkten. Betrachte eine einpunktige Teilmenge.)

8.8.12. ***Aufgabe***. Sei K ein kompakter topologischer Raum, Y ein Hausdorffraum und $f : K \to Y$ bijektiv und stetig. Zeigen Sie, dass f ein Homöomorphismus ist. Was läßt sich aussagen, wenn K nicht kompakt vorausgesetzt wird?

8.8.13. ***Aufgabe***. Seien X, Y topologische Räume mit X kompakt. Zeigen Sie:

(a) Ist $y \in Y$ und W eine offene Menge in $X \times Y$ mit $X \times \{y\} \subset W$, so gibt es eine offene Umgebung $V \subset Y$ von y mit $X \times V \subset W$.

(b) Ist auch Y kompakt, so ist $X \times Y$ kompakt.

8.8.14. ***Aufgabe***. Zeigen Sie:

(a) Jede abgeschlossene Teilmenge A eines kompakten Raumes X ist kompakt.

(b) Ist $(K_n)_n$ eine absteigende Folge nicht-leeren kompakter Teilmengen eines Hausdorffraumes X (z. B. eines metrichen Raumes) d. h. gilt $K_{n+1} \subset K_n$ für alle $n \in \mathbb{N}$, so ist der Durchschnitt $\cap_{n\in\mathbb{N}} K_n$ nicht leer und kompakt (Cantorscher Durchschnittssatz).

8.8.15. ***Aufgabe***. Sei X ein metrischer Raum und $K \subset X$ folgenkompakt. Sei $(U_i)_{i\in I}$ eine Familie offener Mengen in X mit $K \subset \bigcup_{i\in I} U_i$. Zeigen Sie:

(a) Es existiert eine Zahl $r > 0$, so dass es zu jedem $x \in K$ ein $i \in I$ mit $B_r(x) \subset U_i$ gibt.

(b) Zu jedem $r > 0$ gibt es endlich viele $x_1, \dots, x_k$ in K mit $K \subset B_r(x_1) \cup \dots \cup B_r(x_k)$, d.h. K ist total beschränkt.

(Bemerkung: (a) und (b) ergeben offensichtlich einen Beweis dafür, dass folgenkompakte Teilmengen metrischer Räume kompakt sind, d.h. die Implikation (ii)⇔(i) in Satz 8.6.3. Aussage (a) heißt auch das ***Lebesgue-Lemma*** und r heißt ***Lebesgue Zahl*** der Überdeckung $(U_i)_{i\in I}$.)

8.8.16. ***Aufgabe***. Sei $(V, \|\cdot\|)$ ein normierter Vektorraum und $U \subsetneq V$ ein abgeschlossener Unterraum.

(i) Für $v \in V$ definieren wir den Abstand von v zu U durch $d(v,U) = \inf\{\|u - v\| \mid u \in U\}$. Zeigen Sie, dass $d(v,U) > 0$, falls $v \notin U$.

(ii) Sei $\delta > 0$. Dann existiert ein $v \in V$ mit $\|v\| = 1$ und $d(v,U) \geq 1 - \delta$.

(iii) Sei $W \subset V$ ein endlichdimensionaler Unterraum, d.h. $\dim W < \infty$. Zeigen Sie, dass W abgeschlossen ist.

(iv) (***Satz von Riesz***) Die Einheitskugel $B = \{v \in V \mid \|v\| \leqslant 1\}$ ist genau dann kompakt, wenn $\dim V < \infty$.

8.8.17. ***Aufgabe***. (Abstand zwischen zwei Mengen) Sei (X,d) ein metrischer Raum.

(1) Sei $A \subset X$. Zeigen Sie, dass $X \ni x \mapsto d(x,A) := \inf_{a\in A} d(x,a) \in \mathbb{R}$ Lipschitz-stetig ist.

(2) (a) Seien $A, B \subset X$ abgeschlossene Mengen mit $A \cap B = \emptyset$. Sei

$$f : X \to \mathbb{R} \quad f(x) = \frac{d(x,A)}{d(x,A) + d(x,B)}.$$

Zeigen Sie, dass f stetig ist, $0 \leqslant f \leqslant 1$, $A = f^{-1}(0)$, $B = f^{-1}(1)$.
(b) Zeigen Sie, dass zwei offene Mengen $U, V \subset X$ existieren, mit $A \subset U, B \subset V, U \cap V = \emptyset$.

(3) Für $A, B \subset X$ definiere $d(A,B) = \inf\{d(x,y) : x \in A, y \in B\}$.
(a) Sei $K \subset X$ kompakt, $F \subset X$ abgeschlossen, $K \cap F = \emptyset$. Zeigen Sie, dass es $x_0 \in K$ gibt mit $d(x_0, F) = d(K,F)$ und $d(K,F) > 0$.
(b) Seien $K, L \subset X$ kompakte Teilmengen. Zeigen Sie, dass es $x_1 \in K_1, x_2 \in K_2$ gibt mit $d(K_1, K_2) = d(x_1, x_2)$.

(4) Sei $(X,d) = (\mathbb{R}^n, d_2)$.
(a) Sei $F \subset X$ abgeschlossen und unbeschränkt und $f : F \to \mathbb{R}$ stetig mit $\lim\limits_{\|x\| \to \infty} f(x) = \infty$ (f heißt Ausschöpfungsfunktion). Zeigen Sie, dass $x \in F$ existiert mit $f(x) = \inf\limits_{y \in F} f(y)$.
(b) Sei $K \subset X$ kompakt, $F \subset X$ abgeschlossen. Zeigen Sie, dass $x \in K, y \in F$ existieren mit $d(K,F) = d(x,y)$. Bleibt die Aussage wahr, wenn X ein normierter Vektorraum von unendlicher Dimension ist?

8.8.18. ***Aufgabe***. Sei X ein topologischer Raum. Zeigen Sie:

(a) X ist zusammenhängend genau dann, wenn jede stetige Abbildung $f : X \longrightarrow \{0,1\}$ konstant ist (wobei $\{0,1\}$ mit der induzierten Topologie von $(\{0,1\}, d_2)$ versehen ist).

(b) Sei $A \subset X$ eine zusammenhängende Teilmenge. Falls $A \subset B \subset \overline{A}$, dann ist B zusammenhängend.

(c) Sei $(C_i)_{i \in I}$ eine Familie zusammenhängender Teilmengen von X, so dass $i_0 \in I$ existiert, mit $C_i \cap C_{i_0} \neq \emptyset$ für alle $i \in I$. Dann ist $\cup_{i \in I} C_i$ zusammenhängend.

(d) Sei $(C_i)_{i \in I}$ eine höchstens abzählbare Familie zusammenhängender Teilmengen von X (hier $I = \{1, \dots, n\}$ oder $I = \mathbb{N}$), so dass für alle $i \in I \setminus \{1\}$, $C_{i-1} \cap C_i \neq \emptyset$. Dann ist $\cup_{i \in I} C_i$ zusammenhängend.

(e) Seien $X_1, \dots, X_n$ topologische Räume. $X_1 \times \dots \times X_n$ ist zusammenhängend genau dann, wenn $X_1, \dots, X_n$ zusammenhängend ist.

8.8.19. ***Aufgabe***. Sei $S^n = \{x \in \mathbb{R}^{n+1} | x_1^2 + \cdots + x_{n+1}^2 = 1\}$ die n–Sphäre. Zeigen Sie, daß es zu jeder stetigen Funktion $f : S^n \longrightarrow \mathbb{R}$, $n \geqslant 1$, ein Paar antipodaler Punkte $x, -x \in S^n$ existiert, mit $f(x) = f(-x)$.
(Beispiel: Bei jeder stetigen Temperaturverteilung auf der Erdoberfläche gibt es antipodale Orte, in denen gleichzeitig dieselbe Temperatur herrscht.)

9. Differenzierbare Abbildungen

9.1. **Definition und einfache Regeln.** Wie für Funktionen einer Variablen definieren wir die Differenzierbarkeit im Allgemeinen als lineare Approximierbarkeit mit einem Fehler, der von höherer als erster Ordnung verschwindet.

9.1.1. **Definition.** Seien V,W endlichdimensionale reelle normierte Vektorräume. Sei $D \subset V$ offen, $a \in D$ und $f: D \longrightarrow W$ eine Abbildung. Wir sagen, dass f ***differenzierbar*** in a ist, falls es eine lineare Abbildung $T \in \mathcal{L}(V,W)$ gibt so, dass

$$\lim_{x\to a} \frac{\|f(x)-f(a)-T(x-a)\|}{\|x-a\|} = 0$$

d.h. $f(x) = f(a) + T(x-a) + \rho_a(x)$, mit $\lim_{x\to a} \dfrac{\rho_a(x)}{\|x-a\|} = 0$.

Die Abbildung T ist eindeutig bestimmt und heißt ***Differential*** von f in a, bezeichnet mit $T = df(a)$. Wir schreiben $T(v) = df(a)\cdot v = df(a)[v]$. Die Abbildung f heißt differenzierbar in D, falls f in allen $x \in D$ differenzierbar ist. Die Abbildung $df: D \longrightarrow \mathcal{L}(V,W)$, $x \mapsto df(x)$ heißt auch das Differential von f.

9.1.2. **Bemerkung.** (1) Sei $V = \mathbb{R}$, $D = I$ ein Intervall, $W = \mathbb{C} \cong \mathbb{R}^2$. In Satz 5.6.1 haben wir gesehen, dass $f: I \longrightarrow \mathbb{C}$ differenzierbar ist (im Sinne der Analysis I) genau dann, wenn es $\lambda \in \mathbb{C}$ und $\rho: I \longrightarrow \mathbb{C}$ gibt mit $\rho(a) = 0$ und $\lim_{x\to a} \frac{\rho(x)}{\|x-a\|} = 0$ so, dass $f(x) = f(a) + \lambda(x-a) + \rho(x)$; und dann ist $\lambda = f'(a)$. Jede lineare Abbildung $T \in \mathcal{L}(\mathbb{R},\mathbb{C})$ hat die Gestalt $T(t) = \lambda \cdot t$ für alle $t \in \mathbb{R}$, wobei $\lambda = T(1) \in \mathbb{C}$. Daraus folgt: Die Funktion f ist differenzierbar in a im Sinne der Analysis I genau dann, wenn es $T \in \mathcal{L}(\mathbb{R},\mathbb{C})$ und $\rho: I \longrightarrow \mathbb{C}$ gibt mit $\lim_{x\to a} \frac{\rho(x)}{\|x-a\|} = 0$ so, dass $f(x) = f(a) + T(x-a) + \rho(x)$, also genau dann, wenn f differenzierbar in a im Sinne der Definition 9.1.1 ist. Hierbei gilt $\boxed{df(a)\cdot v = f'(a)v}$.

(2) Wie in Analysis I ist $T = df(a)$ die beste lineare Approximation von f in der Nähe von a. Dies hat die folgende geometrische Deutung. Sei $D \subset \mathbb{R}^n$, $f: D \longrightarrow \mathbb{R}$. Dann ist $\mathrm{Graph}(f) = \{(x, x_{n+1}) \in \mathbb{R}^{n+1} : x_{n+1} = f(x)\}$ eine Hyperfläche in $\mathbb{R}^{n+1}$. Der Graph der Abbildung $x \mapsto f(a) + T((x-a)$ ist eine affine Hyperebene

$$\{(x, x_{n+1}) \in \mathbb{R}^{n+1} : x_{n+1} = f(a) + T(x-a)\},$$

genannt Tangentialebene an die Hyperfläche $\mathrm{Graph}(f)$ im Punkt $(a, f(a))$.

9.1.3. **Satz.** *Sei $D \subset V$ offen, $a \in D$ und $f: D \longrightarrow W$ differenzierbar in a. Dann gilt:*

(i) *Das Differential is gegeben für alle $v \in V$ durch*

$$df(a)\cdot v = \lim_{t\to 0} \frac{f(a+tv)-f(a)}{t}.$$

Insbesondere ist $df(a)$ eindeutig bestimmt.

(ii) *f ist stetig in a.*

9.1.4. **Beispiele.**

(1) Eine konstante Abbildung $f: D \longrightarrow W$ ist differenzierbar in D, und $df(x) = 0 \in \mathcal{L}(V,W)$ für alle $x \in D$.

(2) Eine lineare Abbildung $T \in \mathcal{L}(V,W)$ ist differenzierbar und $dT(x) = T \in \mathcal{L}(V,W)$ für alle $x \in V$. Daher ist dT eine konstante Abbildung mit Werten in $\mathcal{L}(V,W)$; ihr Wert ist aber die i.A. nicht konstante Abbildung T.

(3) Seien $V_1, \dots, V_n, W$ endlichdimensionale reelle normierte Vektorräume. Dann ist jede multilineare Abbildung $T: V_1 \times \ldots \times V_n \longrightarrow W$ stetig. Daraus folgt, dass T differenzierbar ist und

$$dT(x_1,\dots,x_n)[v_1,\dots,v_n] = T(v_1,x_2,\dots,x_n) + T(x_1,v_2,x_3,\dots,x_n) + \ldots + T(x_1,\dots,x_{n-1},v_n).$$

Als Beispiel betrachten wir $\det: M_{n\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$. Für $A \in M_{n\times n}(\mathbb{R})$ schreiben wir $A = (a_1,\dots,a_n)$ wobei $a_1,\dots,a_n \in \mathbb{R}^n$ die Spalten von A sind. Wir haben einen Isomorphismus $M_{n\times n}(\mathbb{R}) \longrightarrow \mathbb{R}^n \times \ldots \times \mathbb{R}^n$, $A \mapsto (a_1,\dots,a_n) \in \mathbb{R}^n \times \ldots \times \mathbb{R}^n$. Damit ist die Abbildung $\det: \mathbb{R}^n \times \ldots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ eine multilineare Abbildung. Sei $H \in M_{n\times n}(\mathbb{R})$. Dann gilt

$$d(\det)(A)\cdot H = \det(h_1,a_2,\dots,a_n) + \ldots + \det(a_1,\dots,a_{n-1},h_n) = \sum_{i,j=1}^{n} h_{ij}A_{ij},$$

wobei A_{ij} die Kofaktoren (algebraische Komplementen) von a_{ij} in A sind. Wenn $A^\# = (A_{ji})_{1\leqslant i,j\leqslant n}$, die Adjunktenmatrix Matrix ist, so folgt

$$d(\det)(A)\cdot H = \mathrm{Tr}(A^\# \cdot H).$$

Wir befassen uns nun mit ***Differentiationsregeln***. Die erste Regel besagt, dass wir die Frage der Differenzierbarkeit einer Abbildung mit Werten in einem Produkt auf die Differenzierbarkeit der einzelnen Komponenten der Abbildung reduzieren können.

9.1.5. **Satz** (Reduktionsregel). *Sei $D \subset V$ offen, und seien $f : D \longrightarrow W_1$, $g : D \longrightarrow W_2$ Abbildungen. Die Abbildung $(f,g) : D \longrightarrow W_1 \times W_2$ ist differenzierbar in a genau dann, wenn f und g differenzierbar in a sind. In diesem Fall gilt: $d(f,g)(a) = (df(a), dg(a))$.*

9.1.6. **Folgerung.** *Eine Abbildung $f : D \longrightarrow \mathbb{R}^m$, $f = (f_1,\dots,f_m)$ ist differenzierbar in a genau dann, wenn $f_1,\dots,f_m$ differenzierbar in a sind.*

Als nächstes zeigen wir, dass die Menge $E = \{f : D \to W : f \text{ differenzierbar in } a\}$ ein Unterraum des Vektorraums aller Abbildungen von D nach W ist und die Abbildung $f \to df(a)$ eine lineare Abbildung von E nach $\mathcal{L}(V,W)$ ist.

9.1.7. **Satz** (Linearität). *Seien $f,g : D \longrightarrow W$ differenzierbar in a, und seien $\alpha, \beta \in \mathbb{R}$. Dann ist die Abbildung $\alpha f + \beta g : D \longrightarrow W$ differenzierbar in a, und*

$$d(\alpha f + \beta g)(a) = \alpha df(a) + \beta dg(a) \in \mathcal{L}(V,W).$$

Die Kettenregel ist die meist benutzte Differentiationsregel. Sie schlägt die Brücke von der eindimensionalen zur mehrdimensionalen Analysis und gibt uns die Möglichkeit, zentrale Sätze wie den Mittelwertsatz und den Taylorschen Satz ohne große Mühe auf n Variablen zu verallgemeinern.

9.1.8. **Satz** (Kettenregel). *Seien V, W, Z endlichdimensionale normierte Räume, $D \subset V$, $G \subset W$ offen, $f : D \longrightarrow W$, $g : G \longrightarrow Z$ Abbildungen, $f(D) \subset G$. Ist f in $a \in D$ differenzierbar, g in $f(a) \in G$ differenzierbar, so ist $g \circ f : D \longrightarrow Z$ in a differenzierbar, und*

$$d(g \circ f)(a) = dg(f(a)) \circ df(a) \in \mathcal{L}(V,Z).$$

9.1.9. **Satz** (Leibniz-Regel). *Seien $f_1, f_2 : D \longrightarrow W_1, W_2$ differenzierbar in $a \in D$, und sei $\varphi : W_1 \times W_2 \longrightarrow Z$ bilinear. Dann ist $\varphi(f_1,f_2) : Z \longrightarrow D$ differenzierbar in a und*

$$d\varphi(f_1,f_2)(a) \cdot h = \varphi(df_1(a)\cdot h, f_2(a)) + \varphi(f_1(a), df_2(a)\cdot h)$$

Beweis: $\varphi(f_1,f_2)$ ist die zusammengesetzte Abbildung $D \xrightarrow{(f_1,f_2)} W_1 \times W_2 \xrightarrow{\varphi} Z$. Nach der Kettenregel folgt: $\varphi(f_1,f_2)$ ist differenzierbar in a, und

$$d\varphi(f_1,f_2)(a)\cdot h = d\varphi(f_1(a),f_2(a))\cdot(df_1(a)\cdot h, df_2(a)\cdot h) = \varphi(df_1(a)\cdot h, f_2(a)) + \varphi(f_1(a), df_2(a)\cdot h).$$

9.1.10. **Satz** (Differenzierbarkeit der Umkehrabbildung). *Es seien $D \subset V$ und $G \subset W$ offene Teilmengen und $f : D \longrightarrow G$ eine Abbildung mit folgenden Voraussetzungen:*

(i) *f ist bijektiv;*
(ii) *f ist in $a \in D$ differenzierbar und die Ableitung $df(a) : V \longrightarrow W$ ist bijektiv;*
(iii) *die Inverse $f^{-1} : G \longrightarrow D$ ist in $b := f(a)$ stetig.*

Dann ist f^{-1} in b differenzierbar, und

$$\boxed{d(f^{-1})(b) = (df(a))^{-1}}.$$

9.2. Richtungsableitungen und partielle Ableitungen.

9.2.1. **Definition.** Sei $D \subset V$ offen, $a \in D$ und $f : D \longrightarrow W$. Sei $v \in V$. Existiert der Limes

$$\partial_v f(a) = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t} \in W,$$

so heißt $\partial_v f(a)$ die ***Richtungsableitung*** von f in a in Richtung $v \in V$.

Sei nun $V = \mathbb{R}^n$ und $\{e_1,\dots,e_n\}$ die Standardbasis. f heißt ***partiell differenzierbar*** in a bzgl. der i-ten Koordinatenrichtung, falls die Richtungsableitung

$$\frac{\partial f}{\partial x_i}(a) := \partial_i f(a) := \partial_{e_i} f(a),$$

existiert und $\partial_i f(a)$ heißt die ***i-te partielle Ableitung*** von f in a.

Die Funktion f heißt *partiell differenzierbar*, falls $\partial_i f(x)$ für jedes $x \in D$ und jedes $1 \leqslant i \leqslant n$ existiert.

9.2.2. **Bemerkung.** Die Richtungsableitungen einer Funktion $f : D \longrightarrow W$ kann man als gewöhnliche Ableitungen von Funktionen einer Veränderlichen interpretieren. Sei $r > 0$ mit $B_r(a) \subset D$. Für $|t| < \varepsilon = \frac{r}{\|v\|}$ ist $a + tv \in B_r(a) \subset D$. Wir definieren die Kurve $\gamma : (-\varepsilon, \varepsilon) \longrightarrow D$, $\gamma(t) = a + tv$. Dann ist $\partial_v f(a) = (f \circ \gamma)'(0) \in W$. Wenn $V = \mathbb{R}^n$ und $v = e_i$, so ist $\gamma(t) = (a_1, \dots, a_i + t, \dots, a_n)$ und

$$\partial_i f(a) = \lim_{t \to 0} \frac{f(a_1, \dots, a_i + t, \dots, a_n) - f(a)}{t},$$

d.h. $\partial_i f(a)$ ist die Ableitung bzgl. der Variablen x_i, wenn man $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ festhält, d.h. als Konstanten auffasst. Deshalb gelten für die partiellen Ableitungen analoge Rechenregeln wie für die gewöhnlichen Ableitungen.

Der Fall $V = \mathbb{R}$ ist sehr einfach.

9.2.3. **Lemma.** *Sei $V = \mathbb{R}$. Dann ist $f : D \to W$ in a partiell differenzierbar genau dann, wenn f in a differenzierbar ist. Wir notieren dann*

$$\frac{\partial f}{\partial x}(a) := \frac{df}{dx}(a) := \dot{f}(a) := f'(a) \in W,$$

und es gilt $df(a) \cdot v = f'(a) \cdot v$ für alle $v \in \mathbb{R}$, wobei $f'(a) \cdot v$ das Produkt des Vektors $f'(a) \in W$ mit dem Skalar $v \in \mathbb{R}$ ist.

Sei $D \subset \mathbb{R}$ ein Intervall. Dann heißt eine Abbildung $f : D \to W$ eine ***Kurve*** in W. Ist f differenzierbar, so heißt der Vektor $f'(a) \in W$ ***Ableitung*** (und $\dot{f}(a)$ ***Geschwindigkeitsvektor***) zu f im Punkte a. Geometrisch ist $\dot{f}(a)$ der Tangentialvektor in $f(a)$ an das Bild $f(D)$ der Kurve f. Das Differential $df(a)$ ist einfach die Multiplikation mit $f'(a)$. Diese simple Beschreibung von $df(a)$ ist aber nur dann möglich, wenn $V = \mathbb{R}$ ist.

Ist $V = \mathbb{R}^n$ und $f : D \longrightarrow \mathbb{R}^m$, so schreiben wir $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} = (f_1, \dots, f_m)^T$ als Spaltenvektor.

Offensichtlich ist

$$\frac{\partial f}{\partial x_i}(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_i}(a) \\ \vdots \\ \frac{\partial f_m}{\partial x_i}(a) \end{pmatrix}.$$

Beispiel: $f : \mathbb{R}_+ \times \mathbb{R} \longrightarrow \mathbb{R}^2$, $f(r, \vartheta) = \begin{pmatrix} r\cos\vartheta \\ r\sin\vartheta \end{pmatrix}$; dann ist

$$\frac{\partial f}{\partial r}(r, \vartheta) = \begin{pmatrix} \cos\vartheta \\ \sin\vartheta \end{pmatrix}, \quad \frac{\partial f}{\partial \vartheta}(r, \vartheta) = \begin{pmatrix} -r\sin\vartheta \\ r\cos\vartheta \end{pmatrix}.$$

9.2.4. **Satz.** *Ist f in a differenzierbar, so existieren die Richtungsableitungen $\partial_v f(a)$ für alle $v \in V$, und es gilt*

$$\boxed{\partial_v f(a) = df(a) \cdot v}.$$

Ist $V = \mathbb{R}^n$, so gilt

$$\boxed{df(a) \cdot v = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot v_i}.$$

Ist $W = \mathbb{R}$, so hat $df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ die Darstellung $df(a) = \sum \frac{\partial f}{\partial x_i}(a) e_i^*$, wobei $\{e_1^*, \dots, e_n^*\}$ die duale Basis in $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ zur Standardbasis in $\mathbb{R}^n$ ist, d.h. $e_i^*(v) = v_i = \mathrm{pr}_i(v)$ für alle $v \in \mathbb{R}^n$. Wir betrachten nun pr_i als die „Funktion x_i“, $x \mapsto x_i(x) = x_i$; sie ist eine lineare Funktion, also $dx_i = \mathrm{pr}_i = e_i^*$. Wir haben daher die klassische Schreibweise

$$df(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) dx_i.$$

9.2.5. **Definition.** Seien $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, $D \subset \mathbb{R}^n$ offen, $a \in D$ und $f : D \longrightarrow \mathbb{R}^m$. Die ***Jacobi-Matrix*** $J_f(a) \in M_{m \times n}(\mathbb{R})$ von f in a ist die assoziierte Matrix zu $df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ bzgl. der kanonischen Basen B, C von $\mathbb{R}^n$, $\mathbb{R}^m$, also: $J_f(a) = M_C^B(df(a))$.

Aus der linearen Algebra wissen wir, dass die Spalten der Matrix $M_C^B(df(a))$ die Vektoren $df(a)\cdot e_i = \partial_{e_i} f(a) = \frac{\partial f}{\partial x_i}(a)$ sind. Also

$$J_f(a) = \Big(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)\Big) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

und

$$df(a)\cdot v = J_f(a)\cdot v = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Merkregel zur Bildung der Jacobi-Matrix: Komponenten untereinander, Ableitungen von links nach rechts. Beispiel: Für $f : \mathbb{R}_+ \times \mathbb{R} \longrightarrow \mathbb{R}^2$, $f(r,\vartheta) = \begin{pmatrix} r\cos\vartheta \\ r\sin\vartheta \end{pmatrix}$ gilt

$$J_f(r,\vartheta) = \begin{pmatrix} \cos\vartheta & -r\sin\vartheta \\ \sin\vartheta & r\cos\vartheta \end{pmatrix}.$$

Betrachten wir einige Spezialfälle:

- Falls $n = m = 1$, so ist $J_f(a) = f'(a) \in M_{1\times 1}(\mathbb{R}) = \mathbb{R}$ die Ableitung einer reellen Funktion von einer reellen Variablen.
- Falls $n = 1$, $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$, so ist $J_f(a) = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_m'(a) \end{pmatrix} = f'(a) \in M_{m\times 1}(\mathbb{R}) = \mathbb{R}^m$ (vgl. 9.2.3).
- Falls $m = 1$, so ist f eine skalare Funktion, und $J_f(a) \in M_{1\times n}(\mathbb{R}) = \mathbb{R}^n$ ist ein Zeilenvektor in $\mathbb{R}^n$.

9.2.6. **Definition.** Sei $f : D \to \mathbb{R}$ differenzierbar in $a \in D$. Dann heißt

$$\operatorname{grad} f(a) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix} = J_f(a)^T \in \mathbb{R}^n$$

der ***Gradient*** von f in a.

Für alle $v \in \mathbb{R}^n$ gilt nach Satz 9.2.4

$$df(a)\cdot v = (\partial_1 f(a), \dots, \partial_n f(a)) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \big\langle \operatorname{grad} f(a), v \big\rangle. \tag{9.1}$$

Der Gradient $\operatorname{grad} f(a)$ ist derjenige Vektor in $\mathbb{R}^n$, der bzgl. des euklidischen Skalarprodukts $\langle\cdot,\cdot\rangle$ dual zur Linearform $df(a) : \mathbb{R}^n \longrightarrow \mathbb{R}$ ist. Da $\operatorname{grad} f(a)$ hierdurch eindeutig bestimmt ist, kann (9.1) gut als basisfreie Definition des Gradienten dienen.

Die Gradientenabbildung (auch Gradient genannt) $D \to \mathbb{R}^n$, $x \mapsto \operatorname{grad} f(x)$ ist ein Spezialfall eines Vektorfeldes (im Sinne des Mathematikers), d.h. einer Abbildung $F : D \to \mathbb{R}^n$, oder Feldes (im Sinne des Physikers), d.h. jedem Punkt $x \in D$ wird der Vektor $\operatorname{grad} f(x)$ angeheftet.

Beispiel: Für $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $f(x) = \|x\|^2 = \langle x, x\rangle = x_1^2 + \ldots + x_n^2$ ist $\partial_j f(x) = 2x_j$ und $\operatorname{grad} f(x) = 2x$.

9.2.7. **Satz.** *Ist $f : D \longrightarrow \mathbb{R}$ in a differenzierbar, so gilt $\partial_v f(a) = \langle v, \operatorname{grad} f(a)\rangle$. Die Richtungsableitungen $\partial_v f(a)$ mit $v \in \mathbb{R}^n$, $\|v\| = 1$, nehmen genau die Werte zwischen $-\|\operatorname{grad} f(a)\|$ und $\|\operatorname{grad} f(a)\|$ an. Der größte Wert $\|\operatorname{grad} f(a)\|$ (im Falle $\operatorname{grad} f(a) \neq 0$) wird genau in der durch $\operatorname{grad} f(a)$ bestimmten Richtung $v = \operatorname{grad} f(a)/\|\operatorname{grad} f(a)\|$ angenommen.*

9.2.8. **Satz** (Umschreibung der Kettenregel). *Seien $D \subset \mathbb{R}^n$, $G \subset \mathbb{R}^m$ offen, $f : D \longrightarrow \mathbb{R}^m$, $g : G \longrightarrow \mathbb{R}^\ell$ Abbildungen, $f(D) \subset G$. Ist f in $a \in D$ differenzierbar und g in $f(a) \in G$ differenzierbar, so ist $g \circ f : D \longrightarrow \mathbb{R}^\ell$ in a differenzierbar, und es gilt*

$$\boxed{J_{g\circ f}(a) = J_g(f(a)) \cdot J_f(a)}\,. \tag{9.2}$$

Äquivalent dazu:

$$\boxed{\frac{\partial (g\circ f)_k}{\partial x_i}(a) = \sum_{j=1}^{m} \frac{\partial g_k}{\partial y_j}(f(a)) \cdot \frac{\partial f_j}{\partial x_i}(a)}$$

für alle $i \in \{1,\dots,n\}$ und $k \in \{1,\dots,\ell\}$.

Ist $I \subset \mathbb{R}$ ein Intervall, $\gamma : I \to D \subset \mathbb{R}^n$ eine differenzierbare Kurve und $f : D \to \mathbb{R}^m$ eine differenzierbare Abbildung, so folgt aus (9.2):

$$(f\circ\gamma)'(t_0) = \frac{d(f\circ\gamma)}{dt}(t_0) = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(\gamma(t_0)) \cdot \frac{d\gamma_j}{dt}(t_0), \quad \text{d.h. } (f\circ\gamma)'(t) = df(\gamma(t_0))\cdot\gamma'(t_0).$$

Insbesondere für $m = 1$ nach (9.1):

$$(f\circ\gamma)'(t_0) = \big\langle \operatorname{grad} f(\gamma(t_0)), \gamma'(t_0)\big\rangle .$$

9.2.9. **Satz.** *Ist $f : D \longrightarrow \mathbb{R}$ differenzierbar in D, so steht $\operatorname{grad} f$ senkrecht zu den Niveaumengen von f: Ist $N_c = \{x \in D : f(x) = c\} = f^{-1}(c)$ für $c \in \mathbb{R}$ und $\gamma : I \longrightarrow N_c$ ($I \subset \mathbb{R}$ Intervall) eine differenzierbare Kurve in N_c, so gilt $\operatorname{grad} f(\gamma(t)) \perp \dot\gamma(t)$.*

Somit gilt: Ist $\operatorname{grad} f(a) \neq 0$, so weist $\operatorname{grad} f(a)$ in Richtung des stärksten Anstiegs von f in a, und $\|\operatorname{grad} f(a)\|$ ist dieser stärkste Anstieg. Ist die Funktion f die Höhenfunktion in einer Landschaftskarte, so weist der Gradient dorthin, wo es von a aus am steilsten aufwärts geht. Diese Richtung steht senkrecht zu den Höhenlinien.

Bisher wurde vorausgesetzt, dass das Differential an einer Stelle existiert. Dann existieren dort die partiellen Ableitungen nach Satz 9.2.4. Umgekehrt erhebt sich die Frage, ob aus der Existenz der partiellen Ableitungen die Existenz des Differentials gefolgert werden kann. Die Umkehrung des Satzes 9.2.4 ist aber falsch. Sei $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definiert durch

$$f(x,y) = \begin{cases} \dfrac{xy^2}{x^2+y^4}\,, & \text{falls} \quad (x,y) \neq (0,0), \\ 0\,, & \text{falls} \quad (x,y) = (0,0). \end{cases}$$

Dann existieren alle Richtungsableitungen $\partial_v f(0,0)$, aber f ist nicht stetig in $(0,0)$, also erst recht nicht differenzierbar.

9.2.10. **Satz** (Hauptkriterium für Differenzierbarkeit). *Seien $D \subset \mathbb{R}^n$ offen, $a \in D$, $f : D \longrightarrow \mathbb{R}^m$ eine Abbildung. Falls in einer Umgebung U von a alle partiellen Ableitungen $\partial_1 f, \dots, \partial_n f$ existieren und in a stetig sind, so ist f in a differenzierbar.*

9.2.11. **Beispiel.** Sei $f : \mathbb{R}_+ \times \mathbb{R} \longrightarrow \mathbb{R}^2$, $f(r,\vartheta) = (r\cos\vartheta, r\sin\vartheta)$. Dann gilt

$$\frac{\partial f}{\partial r}(r,\vartheta) = \begin{pmatrix} \cos\vartheta \\ \sin\vartheta \end{pmatrix}, \quad \frac{\partial f}{\partial \vartheta}(r,\vartheta) = \begin{pmatrix} -r\sin\vartheta \\ r\cos\vartheta \end{pmatrix},$$

und $\dfrac{\partial f}{\partial r}, \dfrac{\partial f}{\partial \vartheta}$ sind stetig (sie haben stetige Komponenten). Daher ist f differenzierbar und

$$dP_2(r,\vartheta)\cdot(h,k) = J_{P_2}(r,\vartheta)\cdot\begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} \cos\vartheta & -r\sin\vartheta \\ \sin\vartheta & r\cos\vartheta \end{pmatrix}\begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} h\cos\vartheta - kr\sin\vartheta \\ h\sin\vartheta + kr\cos\vartheta \end{pmatrix}.$$

9.3. **Mittelwertsatz und Schrankensatz.** Sei $D \subset \mathbb{R}^n$ offen. Für *reellwertige* Abbildungen mehrerer Variablen $f : D \to \mathbb{R}$ können wir ein genaues Analogon des MWS für reellwertige Funktionen einer Variablen (Satz 5.3.7, Analysis I) beweisen:

9.3.1. **Satz.** *Sei $f : D \longrightarrow \mathbb{R}$ differenzierbar, und seien $a, b \in D$ so, dass die Verbindungsstrecke $[a,b]$ in D liegt. Dann gibt es $\xi \in [a,b]$ mit $f(b) - f(a) = df(\xi)\cdot(b-a)$.*

9.3.2. **Folgerung** (Konstanzkriterium). *Ist $D \subset \mathbb{R}^n$ offen und zusammenhängend, $f : D \longrightarrow \mathbb{R}^m$ differenzierbar und $df(x) = 0$ für alle $x \in D$, so ist f konstant.*

9.3.3. **Bemerkung.** Wenn D nicht zusammenhängend ist, wird die Folgerung falsch.

9.3.4. **Bemerkung.** Ist $f: D \longrightarrow \mathbb{R}^m$, $m > 1$, differenzierbar, so gilt das Analogon des MWS i.A. nicht. Beispiel: $f: \mathbb{R} \longrightarrow \mathbb{R}^2$, $f(t) = (\cos t, \sin t)$. Dann $f(2\pi) - f(0) = (0,0)$ aber $J_f(\xi) = \begin{pmatrix} -\sin\xi \\ \cos\xi \end{pmatrix}$ für alle $\xi \in [0,2\pi]$, also $df(\xi)(2\pi - 0) = 2\pi \cdot \begin{pmatrix} -\sin\xi \\ \cos\xi \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ für alle $\xi \in [0,2\pi]$ (sonst wäre $\cos^2\xi + \sin^2\xi = 0$). Ein Ersatz für den Mittelwertsatz bei vektoriellen Werten ist eine Abschätzung des Abstandes $\|f(b) - f(a)\|$ mit Hilfe der Norm des Differentials df. Dies kommt im Schrankensatz zum Ausdruck.

Dazu definieren wir zunächst das Integral vektorwertiger Kurven. Für eine stetige Kurve $\gamma: [a,b] \longrightarrow \mathbb{R}^m$ setzen wir

$$\int_a^b \gamma(t)\,dt := \Big(\int_a^b \gamma_1(t), \dots, \int_a^b \gamma_m(t)\,dt \Big).$$

Wir integrieren also eine stetige Kurve in $\mathbb{R}^m$ *komponentenweise*. Damit beweisen wir in Satz 9.3.5 eine Version des Fundamentalsatzes der Differential- und Integralrechnung: Sie leistet einen Aufbau der Differenz $f(b) - f(a)$ aus den Differentialen „dazwischen".

9.3.5. **Satz** (Integraldarstellung des Funktionszuwachses)**.** *Sei $f: D \longrightarrow \mathbb{R}^m$ stetig differenzierbar, $a, b \in D$ mit $[a,b] \subset D$. Dann gilt*

$$\begin{aligned} f(b) - f(a) &= \int_0^1 df((1-t)a + tb) \cdot (b-a)\,dt \\ &= \int_0^1 \sum_{k=1}^n \frac{\partial f}{\partial x_k}((1-t)a + tb) \cdot (b_k - a_k)\,dt. \end{aligned}$$

9.3.6. **Lemma** (Abschätzungsregel)**.**

$$\Big\| \int_a^b \gamma(t)\,dt \Big\| \leqslant \int_a^b \|\gamma(t)\|\,dt.$$

9.3.7. **Satz** (Schrankensatz)**.** *Sei $f: D \longrightarrow \mathbb{R}^m$ stetig differenzierbar, $a, b \in D$ mit $[a,b] \subset D$. Dann gilt*

$$\|f(b) - f(a)\| \leqslant \sup_{t \in [0,1]} \|df((1-t)a + tb)\|_{op} \, \|b - a\|.$$

Ist $K \subset D$ konvex und kompakt, so ist $f|_K$ Lipschitz-stetig mit der Lipschitz-Konstanten $\sup_K \|df(x)\|_{op}$. Dabei ist $\|df(x)\|_{op}$ die Operatornorm von $df(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ ist, definiert bzgl. der euklidischen Normen auf $\mathbb{R}^n$ und $\mathbb{R}^m$.

9.3.8. **Bemerkung.** Weil f stetig differenzierbar ist, ist $df: K \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ stetig und auch $\|df\|_{op}: K \longrightarrow \mathbb{R}$ stetig; da K kompakt ist, gilt $\sup_k \|df(x)\|_{op} = \max_K \|df\|_{op} < \infty$.

Wie berechnet man die Operatornorm explizit? Die lineare Abbildung $df(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ ist gegeben bezüglich der Standardbasen durch die Jacobi-Matrix $J = J_f(x) \in M_{m \times n}(\mathbb{R})$. Sei J^T die transponierte Matrix. Dann ist $J^T J$ eine symmetrische $n \times n$ Matrix, und $\|df(x)\|_{op}$ ist die Wurzel aus dem Maximum der Eigenwerte von $J^T J$.

Ist $f: \mathbb{R}^n \to \mathbb{R}$, so gilt $\|df(x)\|_{op} = \|\operatorname{grad} f(x)\| = \big(\sum_{i=1}^n |\partial_i f(x)|^2\big)^{1/2}$.

Eine Anwendung des Schrankensatzes ist der Satz über Vertauschung von Grenzwert und Differential. Eine Folge von Abbildungen $f_k: D \longrightarrow \mathbb{R}^m$ ***konvergiert lokal gleichmäßig*** gegen eine Abbildung $f: D \longrightarrow \mathbb{R}^m$, wenn es zu jedem Punkt $a \in D$ eine Umgebung U von a in D gibt, so dass $f_k|_U$ gleichmäßig gegen $f|_U$ konvergiert.

9.3.9. **Satz.** *Sei $f_k: D \longrightarrow \mathbb{R}^m$ eine Folge stetig differenzierbarer Abbildungen, und seien $f: D \longrightarrow \mathbb{R}^m$, $g: D \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Gilt $f_k \longrightarrow f$ und $df_k \longrightarrow g$ lokal gleichmäßig für $k \longrightarrow \infty$, so ist f stetig differenzierbar mit $df = g$.*

9.4. **Höhere Ableitungen und der Satz von Schwarz.**

9.4.1. **Definition.** Sei $f: D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ eine Abbildung, und seien $i, j \in \{1, \dots, n\}$. Die partielle Ableitung $\partial_j f$ existiere auf ganz U. Falls die i-te partielle Ableitung zu $\partial_j f$ in a existiert, so heißt $\partial_i(\partial_j f)(a)$ eine ***partielle Ableitung 2. Ordnung*** in a.

Schreibweisen:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) := \partial_{ij} f(a) := \partial_i(\partial_j f)(a), \quad \frac{\partial^2 f}{\partial x_i^2}(a) := \partial_i^2 f(a) := \partial_i(\partial_i f)(a).$$

Beispiel: Seien $D := \mathbb{R}_+ \times \mathbb{R} \subset \mathbb{R}^2$ und $f : D \longrightarrow \mathbb{R}$ definiert durch $f(x,y) := x^y$. Dann

$$\frac{\partial f}{\partial x}(x,y) = yx^{y-1}, \quad \frac{\partial^2 f}{\partial y \partial x}(x,y) = x^{y-1} + y\log x - x^{y-1},$$
$$\frac{\partial f}{\partial y}(x,y) = \log x \cdot x^y, \quad \frac{\partial^2 f}{\partial x \partial y}(x,y) = x^{y-1} + yx^{y-1}\log x,$$
$$\frac{\partial^2 f}{\partial x^2}(x,y) = y(y-1)x^{y-2}, \quad \frac{\partial f}{\partial y^2}(x,y) = \log x \cdot \log x \cdot x^y.$$

9.4.2. **Bemerkung.** Das Beispiel erweckt den Eindruck, als sei die Reihenfolge der partiellen Ableitungen für das Ergebnis nicht entscheidend, beispielsweise $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ oder $\partial_{ij} f = \partial_{ji} f$. Im Allgemeinen gilt aber $\partial_{ij} f \neq \partial_{ji} f$! Für ein Beispiel siehe Aufgabe 9.8.2.

9.4.3. **Satz** (Satz von Schwarz). *Sei $D \subset \mathbb{R}^n$ offen, und seien $i,j \in \{1,\dots,n\}$, $a \in D$. Die Abbildung $f : D \longrightarrow \mathbb{R}^m$ besitze auf D die partiellen Ableitungen $\partial_i f, \partial_j f$ und $\partial_{ij} f$. Außerdem sei $\partial_{ij} f$ in a stetig. Dann existiert auch $\partial_{ji} f(a)$, und es gilt $\partial_{ij} f(a) = \partial_{ji} f(a)$.*

Beweis: O.B.d.A. $n = 2$, $a = (0,0)$, $i = 1$, $j = 2$. Dann gilt:

$$\begin{aligned}
\partial_2(\partial_1 f)(0,0) &= \lim_{h\to 0} \frac{1}{h}\Big[(\partial_1 f)(0,h) - (\partial_1 f)(0,0)\Big] \\
&= \lim_{h\to 0} \frac{1}{h}\Big[\lim_{k\to 0} \frac{f(k,h)-f(0,h)}{k} - \lim_{k\to 0} \frac{f(k,0)-f(0,0)}{k}\Big] \\
&= \lim_{h\to 0}\lim_{k\to 0} \frac{1}{hk}\Big[f(k,h)-f(0,h)-f(k,0)+f(0,0)\Big]
\end{aligned}$$

Sei zunächst k und h fest. Nun verwenden wir den Mittelwertsatz für $g_k(y) := f(k,y) - f(0,y)$:

$$f(k,h)-f(0,h)-f(k,0)+f(0,0) = g_k(h) - g_k(0) = hg_k'(\eta)$$

für ein η zwischen 0 und h. Aber $g_k'(\eta) = \partial_2 f(k,\eta) - \partial_2 f(0,\eta)$. Setze $r(x) = \partial_2 f(x,\eta)$. Wir verwenden erneut den Mittelwertsatz: $r(k) - r(0) = kr'(\xi) = k\partial_1(\partial_2 f)(\xi,\eta)$ für ein ξ zwischen 0 und k. Also: Für alle k,h gibt es ξ,η (mit ξ zwischen 0 und k, sowie η zwischen 0 und h) mit

$$[f(k,h)-f(0,h)-f(k,0)+f(0,0)] = hk\partial_1(\partial_2 f)(\xi,\eta)$$

Da $\partial_1\partial_2 f$ stetig ist und $\xi,\eta \to 0$ für $k,h \to 0$ gilt, folgt:

$$\partial_2(\partial_1 f)(0,0) = \lim_{h\to 0}\lim_{k\to 0} \partial_1(\partial_2 f)(\xi,\eta) = \partial_1(\partial_2 f)(0,0). \quad \square$$

Die Aussage des Satzes von Schwarz trifft insbesondere für zweimal stetig differenzierbaren Funktionen zu (siehe Def. 9.4.5).

9.4.4. **Definition.** Für $i_1,\dots,i_k \in \{1,\dots,n\}$ definieren wir eine ***partielle Ableitung k-ter Ordnung*** in Richtungen $x_{i_1}, x_{i_2}, \dots, x_{i_k}$ durch $\partial_{i_1\cdots i_k} f = \partial_{i_1}(\cdots(\partial_{i_{k-1}}(\partial_{i_k} f))\cdots)$, falls f auf U in Richtung x_{i_k} differenzierbar ist, $\partial_{i_k} f$ auf U in Richtung $x_{i_{k-1}}$ differenzierbar ist etc.

Schreibweisen:

$$\frac{\partial^k f}{\partial x_{i_1}\dots\partial x_{i_k}} := \partial_{i_1\cdots i_k} f := \partial_{i_1}(\cdots(\partial_{i_{k-1}}(\partial_{i_k} f))\cdots),$$
$$\frac{\partial^\ell f}{\partial x_i^\ell} := \partial_i^\ell f := \underbrace{\partial_i\partial_i\dots\partial_i}_{\ell-mal} f, \quad \text{für } \ell \in \mathbb{N}_0 \text{ (wobei } \partial_i^0 f := f),$$
$$\frac{\partial^k f}{\partial x_1^{\ell_1}\dots\partial x_n^{\ell_n}} := \partial_1^{\ell_1}\cdots\partial_n^{\ell_n} f, \quad \text{für } \ell \in \mathbb{N}_0,\ \ell_1 + \dots + \ell_n = k.$$

9.4.5. **Definition.** Sei $D \subset \mathbb{R}^n$ offen und $k \in \mathbb{N}$.

(i) Eine Abbildung $f : D \longrightarrow \mathbb{R}^m$ heißt ***k-mal stetig differenzierbar***, wenn alle partiellen Ableitungen bis zur Ordnung k von f existieren und stetig sind. Wir bezeichnen mit $\mathcal{C}^k(D,\mathbb{R}^m)$ den Raum der k-mal stetig differenzierbaren Abbildungen von U nach $\mathbb{R}^m$.

(ii) Eine Abbildung $f : D \longrightarrow \mathbb{R}^m$ heißt ***unendlich oft differenzierbar*** oder ***glatt***, wenn für alle $k \in \mathbb{N}$ alle partiellen Ableitungen $\partial_{i_1\cdots i_k} f$ existieren und stetig sind. Wir bezeichnen mit $\mathcal{C}^\infty(D,\mathbb{R}^m)$ der Raum der unendlich oft differenzierbaren Abbildungen von U nach $\mathbb{R}^m$.

Offensichtlich gilt

$$\mathcal{C}^0(D,\mathbb{R}^m) \supset \mathcal{C}^1(D,\mathbb{R}^m) \supset \mathcal{C}^2(D,\mathbb{R}^m) \supset \ldots \supset \mathcal{C}^k(D,\mathbb{R}^m) \supset \mathcal{C}^{k+1}(D,\mathbb{R}^m) \supset \ldots$$

und $\mathcal{C}^\infty(D,\mathbb{R}^m) = \cap_{k\in\mathbb{N}} \mathcal{C}^k(D,\mathbb{R}^m)$.

9.4.6. **Bemerkung.** Mit dem Satz von Schwarz zeigt man induktiv: Ist $f \in \mathcal{C}^k(D,\mathbb{R}^m)$, so spielt die Reihenfolge der partiellen Ableitungen in $\partial_{i_1\cdots i_k} f$ keine Rolle: Für jede Permutation $(j_1,\ldots,j_k)$ von $(i_1,\ldots,i_k)$ gilt $\partial_{j_1\cdots j_k} f = \partial_{i_1\cdots i_k} f$ auf D.

Eine wichtige Rolle in der Analysis, partielle Differentialgleichungen, mathematische Physik usw. spielen die ***Differentialoperatoren***. Als Beispiel betrachten wir Der ***Laplace-Operator***

$$\Delta : C^2(U,\mathbb{R}) \to C^0(U,\mathbb{R}), \quad \Delta f := \partial_{11} f + \ldots + \partial_{nn} f. \tag{9.3}$$

wobei $U \subset \mathbb{R}^n$ offen ist.

9.4.7. **Definition.** Sei $f \in \mathcal{C}^2(D,\mathbb{R})$, $D \subset \mathbb{R}^n$ offen, $a \in D$. Dann heißt die Matrix

$$H_f(a) := \begin{pmatrix} \partial_{11} f(a) & \ldots & \partial_{1n} f(a) \\ \vdots & & \vdots \\ \partial_{n1} f(a) & \ldots & \partial_{nn} f(a) \end{pmatrix}$$

Hesse-Matrix zu f in a. Die Bilinearform

$$d^2 f(a) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, (u,v) \mapsto \partial_u(\partial_v f)(a).$$

heißt ***Hesse-Form*** oder ***Differential zweiter Ordnung*** zu f in a.

Wegen des Satzes von Schwarz ist $H_f(a)$ symmetrisch, d.h. $\partial_{ij} f(a) = \partial_{ji} f(a)$. Daher ist die Hesse-Matrix die Jacobi-Matrix des Gradienten: $H_f(a) = J_{\operatorname{grad} f}(a)$. Die Hesse-Form ist wohldefiniert, da $\partial_v f = v_1\partial_1 f + \ldots + v_n \partial_n f$ differenzierbar ist; daher ist $\partial_u(\partial_v f)(a)$ linear in u und v.

Es gilt

$$d^2 f(a)(u,v) = \sum_{i,j=1}^n u_i v_j d^2 f(a)(e_i,e_j) = \sum_{i,j=1}^n u_i v_j \partial_{ij} f(a) = u^T H_f(a) v.$$

Also hat die Bilinearform $d^2 f(a)$ bezüglich der Standardbasis $\{e_1,\ldots,e_n\}$ die Matrix $H_f(a)$. Insbesondere ist $d^2 f(a)$ symmetrisch. Wir werden sehen, dass die Hesse-Matrix bei der Extremwertbestimmung eine wesentliche Rolle spielt.

9.4.8. **Lemma.** *Sei $D \subset \mathbb{R}^n$ offen und $f \in \mathcal{C}^k(D,\mathbb{R}^m)$. Dann gilt für alle $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$:*

$$\partial_{v_1 v_2 \ldots v_k} f(a) := \partial_{v_1}(\ldots(\partial_{v_k} f)\ldots)(a) = \sum_{i_1,i_2,\ldots,i_k=1}^n v_{1i_1} v_{2i_2} \ldots v_{ki_k} \partial_{i_1 i_2 \ldots i_k} f(a).$$

9.4.9. **Definition.** Sei $f \in \mathcal{C}^k(D,\mathbb{R}^m)$ und $a \in D$. Die multilineare Abbildung

$$d^k f(a) : \mathbb{R}^n \times \ldots \times \mathbb{R}^n \longrightarrow \mathbb{R}^m,$$

$$(v_1, v_2, \ldots, v_k) \longmapsto \partial_{v_1 v_2 \ldots v_k} f(a) = \sum_{i_1,i_2,\ldots,i_k=1}^n v_{1i_1} v_{2i_2} \ldots v_{ki_k} \partial_{i_1 i_2 \ldots i_k} f(a)$$

heißt ***Differential k-ter Ordnung*** zu f in a. Nach Bemerkung 9.4.6 ist $d^k f(a)$ symmetrisch.

9.5. **Die Taylorformel.** Sei $D \subset \mathbb{R}^n$ offen, $f \in \mathcal{C}^k(D,\mathbb{R}^m)$, $a \in D$. Für $1 \leqslant r \leqslant k$ betrachte die multilineare Abbildung $d^r f(a)$. Für $v_1 = v_2 = \ldots = v_r = v$ setzen wir

$$d^r f(a) \cdot v^{(r)} := d^r f(a)(v, v, \ldots, v).$$

9.5.1. **Definition.** Sei $f \in \mathcal{C}^{p+1}(D,\mathbb{R})$ und $a \in D$. Das Polynom

$$T_{p,a} f(x) = f(a) + \frac{1}{1!} df(a)(x-a) + \frac{1}{2!} d^2 f(a)(x-a)^{(2)} + \ldots + \frac{1}{p!} d^p f(a)(x-a)^{(p)}$$

heißt ***p-tes Taylorpolynom*** zu f in $a \in D$.

Für $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ und $v \in \mathbb{R}^n$ schreibe

$$|\alpha| := \alpha_1 + \ldots + \alpha_n, \quad \alpha! := \alpha_1! \cdots \alpha_n!, \quad v^\alpha = v_1^{\alpha_1} \cdots v_n^{\alpha_n},$$

und setze:

$$\partial_i^\mu f := \frac{\partial^\mu}{\partial x_i^\mu}, \quad \partial^\alpha f := \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

9.5.2. **Lemma.** *Es gilt*

$$d^r f(a) \cdot v^{(r)} = \sum_{|\alpha|=r} \frac{r!}{\alpha!} \partial^\alpha f(a) v^\alpha.$$

9.5.3. **Satz** (Taylorformel). *Sei $f \in \mathcal{C}^{p+1}(D, \mathbb{R})$, und seien $a, x \in D$ mit $[a,x] \subset D$.*
(i) Taylorformel mit Lagrange-Restglied: Es gibt $\xi \in [a,x]$ mit

$$\begin{aligned} f(x) &= T_{p,a} f(x) + \frac{1}{(p+1)!} d^{p+1} f(\xi) \cdot (x-a)^{(p+1)} \\ &= \sum_{|\alpha| \leqslant p} \frac{1}{\alpha!} \partial^\alpha f(a)(x-a)^\alpha + \sum_{|\alpha|=p+1} \frac{1}{\alpha!} \partial^\alpha f(\xi)(x-a)^\alpha. \end{aligned}$$

(ii) Taylorformel mit Integralrestglied: Mit $x = a + h$ gilt:

$$\begin{aligned} f(x) &= T_{p,a} f(x) + \frac{1}{p!} \int_0^1 (1-t)^p d^{p+1} f(a+th) h^{(n+1)}\, dt \\ &= \sum_{|\alpha| \leqslant p} \frac{1}{\alpha!} \partial^\alpha f(a)(x-a)^\alpha + p \int_0^1 (1-t)^p \sum_{|\alpha|=p+1} \frac{1}{\alpha!} \partial^\alpha f(a+th) h^\alpha\, dt. \end{aligned}$$

(iii) Qualitative Taylorformel: Ist $f \in \mathcal{C}^p(D, \mathbb{R}^m)$ und $a \in D$ so gilt:

$$f(x) = T_{p,a} f(x) + o(\|x-a\|^p) \text{ für } x \to a; \text{ also } \lim_{x \to a} \frac{f(x) - T_{p,a} f(x)}{\|x-a\|^p} = 0.$$

9.6. Lokale Extrema.

9.6.1. **Definition.** Sei X ein topologischer Raum (z.B. $X \subset \mathbb{R}^n$ mit der induzierten Topologie).

(i) $f : X \longrightarrow \mathbb{R}$ hat in $a \in X$ ein ***Maximum*** (bzw. ***Minimum***), wenn $f(x) \leqslant f(a)$ für alle $x \in X$ (bzw. $f(x) \geqslant f(a)$ für alle $x \in X$). Man bezeichnet a auch als ***globales Maximum*** (bzw. ***globales Minimum***).
(ii) $f : X \longrightarrow \mathbb{R}$ hat in a ein ***lokales Maximum*** (bzw. ***lokales Minimum***), wenn es eine Umgebung U von a in X gibt mit $f(x) \leqslant f(a)$ (bzw. $f(x) \geqslant f(a)$) für alle $x \in U$.
(iii) $f : X \longrightarrow \mathbb{R}$ hat in a ein ***strenges Maximum*** (bzw. ***strenges Minimum***), wenn $f(x) < f(a)$ (bzw. $f(x) > f(a)$) für alle $x \in X \setminus \{a\}$.
(iv) $f : X \longrightarrow \mathbb{R}$ hat in a ein ***strenges lokales Maximum*** (bzw. ***strenges lokales Minimum***), wenn es eine Umgebung U von a gibt mit $f(x) < f(a)$ (bzw. $f(x) > f(a)$) für alle $x \in U \setminus \{a\}$.
(v) $f : X \longrightarrow \mathbb{R}$ hat in a ein ***lokales Extremum***, wenn f in a ein lokales Maximum oder ein lokales Minimum hat.

9.6.2. **Definition.** Seien $D \subset (V, \|\cdot\|)$ offen und $f : D \longrightarrow \mathbb{R}$ differenzierbar. Ein Punkt $a \in D$ mit $df(a) = 0 \in \mathcal{L}(V, \mathbb{R})$ heißt ***kritisch*** oder stationär.

9.6.3. **Satz** (Fermatsches Kriterium für lokale Extrema). *Sei $D \subset (V, \|\cdot\|)$ offen, $f : D \longrightarrow \mathbb{R}$ differenzierbar und a ein lokales Extremum. Dann ist a ein kritischer Punkt, d.h. $df(a) = 0$.*

Das Fermat-Kriterium ist notwendig, jedoch nicht hinreichend für das Vorhandensein eines Extremums (schon im Eindimensionalen bekannt). Ein 2-dimensionales Beispiel: $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $f(x,y) = xy$, $df(0,0) = 0$. In jeder Umgebung U von $(0,0)$ gibt es jedoch Punkte $u, v \in U$ mit $f(u) > 0 > f(v)$.

Der Linearen Algebra entnehmen wir folgende Definitionen. Eine symmetrische Bilinearform $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ heißt
(i) *positiv definit*, wenn $b(v,v) > 0$ für alle $v \neq 0$;
(ii) *negativ definit*, wenn $b(v,v) < 0$ für alle $v \neq 0$;
(iii) *indefinit*, wenn v, w mit $b(v,v) > 0$ und $b(w,w) < 0$ existieren.

9.6.4. **Satz** (Hinreichendes Kriterium für lokale Extrema)**.** *Sei $D \subset \mathbb{R}^n$ offen, $f \in \mathcal{C}^2(D,\mathbb{R})$ und $a \in D$ ein kritischer Punkt. Dann gilt:*

(i) *Ist die Hesse-Form $d^2f(a)$ **positiv definit**, so hat f in a ein strenges lokales Minimum.*

(ii) *Ist die Hesse-Form $d^2f(a)$ **negativ definit**, so hat f in a ein strenges lokales Maximum;*

(iii) *Ist die Hesse-Form $d^2f(a)$ **indefinit**, so hat f kein lokales Extremum in a.*

9.6.5. **Bemerkung.**

(i) $d^2f(a)$ ist positiv definit (bzw. negativ definit, indefinit) genau dann, wenn $H_f(a)$ positiv definit (bzw. negativ definit, indefinit) ist.

(ii) Sei $H \in M_{2\times 2}(\mathbb{R})$ symmetrisch, $H = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$.

Wenn $b_{11} > 0$ und $\det H > 0$, dann ist H positiv definit.
Wenn $b_{11} < 0$ und $\det H > 0$, dann ist H negativ definit.
Wenn $\det H < 0$, dann ist H indefinit.

(iii) Sei $H \in M_{n\times n}(\mathbb{R})$ symmetrisch. Dann gilt:

H ist positiv definit genau dann, wenn alle Eigenwerte positiv (>0) sind.
H ist negativ definit genau dann, wenn alle Eigenwerte negativ (<0) sind.
H ist indefinit genau dann, wenn Eigenwerte λ, μ von H mit $\lambda > 0$ und $\mu < 0$ existieren.

9.7. **Extremwertbestimmung.** Sei $f \in \mathcal{C}^2(D,\mathbb{R})$. Gesucht sind die lokalen Extrema von f.

$\boxed{1}$ Bestimme die kritischen Punkte, d.h. die Lösungen des Systems:

$$\frac{\partial f}{\partial x_1}(x) = \ldots = \frac{\partial f}{\partial x_n}(x) = 0.$$

$\boxed{2}$ Sei a ein kritischer Punkt. Berechne die Matrix $H_f(a)$ und deren Eigenwerte.

- Alle Eigenwerte sind positiv $\rightsquigarrow$ a ist ein strenges lokales Minimum.
- Alle Eigenwerte sind negativ $\rightsquigarrow$ a ist ein strenges lokales Maximum.
- Es gibt positive und negative Eigenwerte $\rightsquigarrow$ a ist kein lokales Extremum.
- 0 ist Eigenwert $\rightsquigarrow$ Wir können nicht entscheiden und müssen eventuell eine Taylorformel höherer Ordnung benutzen.

$\boxed{3}$ Dann betrachtet man ggf. andere kritische Punkte.

9.7.1. **Beispiel.** Sei $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $f(x,y) = x^3 + x^2 + y^2$; $\operatorname{grad} f(x,y) = (3x^2+2x, 2y) = (0,0)$ genau dann, wenn $(x,y) \in \{(0,0), (-\frac{2}{3}, 0)\}$. Die Hesse-Matrix ist: $H_f(x,y) = \begin{pmatrix} 6x+2 & 0 \\ 0 & 2 \end{pmatrix}$ mit $H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ positiv definit und $H_f(-\frac{2}{3}, 0) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$ indefinit. Daraus folgt: Der Punkt $(0,0)$ ist ein lokales Minimum, und $(-\frac{2}{3}, 0)$ ist kein lokales Extremum.

Punkte $(x,y) \in \mathbb{R}^2$ mit $df(x,y) = 0$ und $H_f(x,y)$ indefinit heißen ***Sattelpunkte***.

9.7.2. **Beispiel.** Betrachte die Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $f(x,y) = x^4 + y^3 - 4x^3 - 3y^2 + 3y$; $\operatorname{grad} f(x,y) = (4x^3 - 12x^2, 3y^2 - 6y + 3) = (0,0)$ genau dann, wenn $(x,y) \in \{(0,1), (3,1)\}$. Es ist

$$H_f(x,y) = \begin{pmatrix} 12x^2 + -24x & 0 \\ 0 & 6y-6 \end{pmatrix}, \quad H_f(0,1) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad H_f(0,1) = \begin{pmatrix} 36 & 0 \\ 0 & 0 \end{pmatrix}.$$

In beiden Fällen ist 0 ein Eigenwert der Hessematrix und das Kriterium 9.6.4 gibt keine Auskunft. Wir studieren daher das Vorzeichen der Differenzen $f(x,y) - f(0,1)$ und $f(x,y) - f(3,1)$. Für ein Polynom P gilt $P(x) = T_{k,a}P(x)$ für alle $k \geqslant \operatorname{grad} P$, d.h. wir können das Polynom nach der Potenzen von $x_1 - a_1, \ldots, x_n - a_n$ entwickeln. Also $f(x,y) - f(0,1) = x^4 - 4x^3 + (y-1)^2$ und dieser Ausdruck ist > 0 in allen Punktem $(0, 1+z)$ mit $z > 0$ und < 0 in allen Punktem $(0, 1+z)$ mit $z < 0$. Es folgt, dass $(0,1)$ kein Extremum ist. Es gilt $f(x,y) - f(3,1) = (x-3)^4 + 8(x-3)^3 + 18(x-3)^2 + (y-1)^2 = (x-3)^2(x^2+2x+3) + (y-1)^2 \geqslant 0$ also $(3,1)$ ist eine Minimumstelle.

9.8. **Übungen.**

9.8.1. ***Aufgabe***. (i) Es sei $f:\mathbb{R}^2\to\mathbb{R}$ definiert durch

$$f(x)=\begin{cases}\dfrac{x_1x_2^2}{x_1^2+x_2^2} & , \quad x\neq 0\\ 0 & , \quad x=0\end{cases}.$$

Zeige, daß f stetig ist und für jedes $x\in\mathbb{R}^2$ und jedes $y\in\mathbb{R}^2$

$$(*)\qquad \lim_{t\to 0}\frac{f(x+ty)-f(x)}{t}=g(x,y)$$

existiert, aber daß die Abbildung $y\mapsto g(0,y)$ nicht linear ist (so daß f im Punkt 0 nicht differenzierbar sein kann).

(ii) Es sei $f:\mathbb{R}^2\to\mathbb{R}$ definiert durch

$$f(x)=\begin{cases}\dfrac{x_1^3x_2}{x_1^4+x_2^2} & , \quad x\neq 0\\ 0 & , \quad x=0\end{cases}.$$

Zeige, daß der Limes $g(x,y)$ in $(*)$ für jedes $x\in\mathbb{R}^2$ und jedes $y\in\mathbb{R}^2$ existiert und daß $y\mapsto g(x,y)$ für jedes $x\in\mathbb{R}^2$ linear ist, aber f im Punkt 0 nicht differenzierbar ist.

9.8.2. ***Aufgabe***. Sei $f:\mathbb{R}^2\to\mathbb{R}$ definiert durch $f(0,0)=0$ und $f(x,y)=xy\dfrac{x^2-y^2}{x^2+y^2}$ für $(x,y)\neq(0,0)$. Zeigen Sie:

(a) $\partial_1 f$ und $\partial_2 f$ existieren in jedem Punkt des $\mathbb{R}^2$ (auch im Nullpunkt) und sind stetig; insbesondere ist f differenzierbar.

(b) $\partial_{12}f(0,0)$ und $\partial_{21}f(0,0)$ existieren, sind aber nicht gleich.

9.8.3. ***Aufgabe***. Wir betrachten nun $U:=\mathbb{R}^n\setminus\{0\}$ und die Funktion $r:U\ni x\mapsto\|x\|_2=\sqrt{\langle x,x\rangle}\in\mathbb{R}$.

(a) Berechnen Sie die Gradienten $\operatorname{grad}(r^{2-n})(x)$ und $\operatorname{grad}(\log r)(x)$.

(b) Zeigen Sie: $\Delta(r^{2-n})=0$; im Falle $n=2$ gilt $\Delta(\log r)=0$.

(c) Sei $f\in C^2(\mathbb{R}^2,\mathbb{R})$ und $P:\mathbb{R}_+\times\mathbb{R}\ni(r,\phi)\mapsto(r\cos\phi,r\sin\phi)\in\mathbb{R}^2$ die Polarkoordinatenabbildung. Zeigen Sie, daß mit $F:=f\circ P\in C^2(\mathbb{R}_+\times\mathbb{R},\mathbb{R})$ gilt:

$$\Big(\frac{\partial^2}{\partial r^2}F+\frac{1}{r^2}\cdot\frac{\partial^2}{\partial\phi^2}F+\frac{1}{r}\cdot\frac{\partial}{\partial r}F\Big)(r,\phi)=\big(\Delta f\big)(P(r,\phi)).$$

(d) Sei $k\in\mathbb{N}$. Wir betrachten die Abbildung $\mathbb{C}\ni z\mapsto z^k\in\mathbb{C}$ sowie die zugehörige Abbildung $f:\mathbb{R}^2\to\mathbb{R}^2$, die sich daraus nach der üblichen Identifikation von $\mathbb{C}$ mit $\mathbb{R}^2$ ergibt; diese ist ein $\mathbb{R}^2$-wertiges Polynom, insbesondere $f\in C^\infty(\mathbb{R}^2,\mathbb{R}^2)$. Sei P die Polarkoordinatenabbildung und sei $F:=f\circ P$. Berechnen Sie $F(r,\phi)$ und zeigen Sie $\Delta f_i=0$ für $i=1,2$.

9.8.4. ***Aufgabe*** (Lemma von Hadamard). Sei $f\in\mathcal{C}^\infty(\mathbb{R}^n)$. Zeigen Sie:

(a) Es gibt $g_i\in\mathcal{C}^\infty(\mathbb{R}^n)$ mit $f(x)=f(0)+\sum\limits_{i=1}^n x_ig_i(x)$, $x\in\mathbb{R}^n$.

(b) Ist $df(0)=0$, so existieren $h_{ij}\in\mathcal{C}^\infty(\mathbb{R}^n)$, mit $f(x)=f(0)+\sum\limits_{i,j=1}^n x_i\,x_j\,h_{ij}(x)$, $x\in\mathbb{R}^n$.

9.8.5. ***Aufgabe*** (Satz von Rolle in $\mathbb{R}^n$). Sei $f:\mathbb{R}^n\to\mathbb{R}$ differenzierbar, so dass $f|_{S^{n-1}}$ konstant ist (wobei $S^{n-1}=\{x\in\mathbb{R}^n:\|x\|=1\}$). Zeigen Sie, dass f einen kritischen Punkt in $B_1(0)=\{x\in\mathbb{R}^n:\|x\|<1\}$ hat.

9.8.6. ***Aufgabe***. (Homogene Funktionen) Die in $U=\mathbb{R}^m\smallsetminus\{0\}$ erklärte reellwertige Funktion f heißt homogen vom Grad $\alpha\in\mathbb{R}$, wenn $f(tx)=t^\alpha f(x)$ für alle $x\in U$ und $t>0$ gilt. Sei f außerdem unendlich oft differenzierbar in U.

(a) Zeigen Sie, daß ein kritischer Punkt von f auch eine Nullstelle ist, falls $\alpha\neq 0$.

b) Nehmen wir an, f sei erklärt und unendlich oft differenzierbar auf $\mathbb{R}^m$. Beweisen Sie, daß f ein Polynom ist.

9.8.7. ***Aufgabe***. Bestimmen Sie die Extrema der Funktion

$$f:\mathbb{R}^2 \longrightarrow \mathbb{R},\, f(x,y)=e^{xy}+x^2+\lambda y^2\,,\, \lambda>0$$

(Diskussion von Wirkung von λ auf die Extrema).

9.8.8. ***Aufgabe***. Für ein dem Einheitskreis einbeschriebenes Dreieck mit den Eckpunkten $a=(1,0)$, $b=(\cos x,\sin x)$ und $c=(\cos y,\sin y)$ ist der orientierte Flächeninhalt gegeben durch

$$f(x,y)=\frac{1}{2}\det\big(b-a\,,\,c-a\big)=\frac{1}{2}\det\begin{pmatrix}\cos x-1 & \cos y-1\\ \sin x & \sin y\end{pmatrix}.$$

Sei $D:=\{(x,y)\mid 0\le x\le y\le 2\pi\}$. Bestimme die lokalen Extrema von $f:D\to\mathbb{R}$ und interpretiere das Ergebnis geometrisch.

9.9. Notizen: Komplexe Differenzierbarkeit.

9.9.1. **Definition.** Sei $D\subset\mathbb{C}$ offen. Eine Funktion $f:D\longrightarrow\mathbb{C}$ heißt ***komplex differenzierbar*** in $z_0\in D$, wenn

$$f'(z_0):=\frac{df}{dz}(z_0):=\lim_{z\to z_0}\frac{f(z)-f(z_0)}{z-z_0}\quad \text{in } \mathbb{C} \text{ existiert.} \tag{9.4}$$

Diese Zahl heißt dann die (komplexe) ***Ableitung*** von f in z_0. Die Funktion $f:D\longrightarrow\mathbb{C}$ heißt ***holomorph*** in D, falls f komplex differenzierbar in allen Punkten $z\in D$ ist; f heißt ***holomorph in*** $z_0\in D$, wenn f holomorph in einer offenen Umgebung von z_0 ist.

Jedes Polynom $P:\mathbb{C}\longrightarrow\mathbb{C}$, $P(z)=a_nz^n+\ldots+a_0$ ist komplex differenzierbar, und

$$P'(z)=na_nz^{n-1}+\ldots+2a_2z+a_1.$$

Diese Definition erhalten wir durch Übertragung der Definition der Differenzierbarkeit in einer reellen Veränderlichen. Auf gleiche Weise wie in Satz 5.6.1 zeigt man:

9.9.2. **Lemma.** *Folgende Bedingungen sind äquivalent:*

(i) f ist komplex differenzierbar in z_0

(ii) Es gibt $\lambda\in\mathbb{C}$ mit $f(z)=f(z_0)+\lambda(z-z_0)+o|z-z_0|)$

(d.h. es gibt $\rho:D\longrightarrow\mathbb{C}$ mit $\lim_{z\to z_0}\frac{\rho(z)}{z-z_0}=0$ und $f(z)=f(z_0)+\lambda(z-z_0)+\rho(z)$).

(iii) Es gibt $L:\mathrm{Hom}_{\mathbb{C}}(\mathbb{C},\mathbb{C})$ mit $f(z)=f(z_0)+L(z-z_0)+o(|z-z_0|)$.

In diesem Falle gilt $f'(z_0)=\lambda$ und $L(v)=f'(z_0)\cdot v$ für $v\in\mathbb{C}$.

Dabei bezeichnet $\mathrm{Hom}_{\mathbb{C}}(\mathbb{C},\mathbb{C})$ der Raum der $\mathbb{C}$–linearen Abbildungen von $\mathbb{C}$ nach $\mathbb{C}$. Sei nun $f:D\longrightarrow\mathbb{C}$, $z=x+iy\mapsto f(z)=u(x,y)+iv(x,y)$, wobei $u=\mathrm{Re}f$ und $v=\mathrm{Im}f$. Wir identifizieren f mit einer Funktion $f:D\subset\mathbb{R}^2\to\mathbb{R}^2$, $(x,y)\mapsto((x,y),v(x,y))$:

$$\begin{array}{ccc}
\mathbb{C}\supset D & \xrightarrow{z\longmapsto f(z)} & \mathbb{C}\\
{\scriptstyle z\mapsto(x,y)}\big\downarrow & & \big\downarrow{\scriptstyle z\mapsto(x,y)}\\
\mathbb{R}^2\supset D & \xrightarrow[(x,y)\mapsto(u(x,y),v(x,y))]{} & \mathbb{R}^2
\end{array}$$

Laut Definition 9.1.1 ist die Funktion f reell-differenzierbar, wenn $L\in\mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^2,\mathbb{R}^2)=\mathrm{Hom}_{\mathbb{R}}(\mathbb{C},\mathbb{C})$ existiert mit $f(z)=f(z_0)+L(z-z_0)+o(|z-z_0|)$, $z\to z_0$. Die Abbildung L ist das Differential von f in z_0 und wird bezeichnet mit $L=df(z_0)$. Ist f reell-differenzierbar in z_0, so existieren die partiellen Ableitungen

$$\frac{\partial f}{\partial x}(z_0)=\frac{\partial u}{\partial x}(z_0)+i\frac{\partial v}{\partial x}(z_0),\quad \frac{\partial f}{\partial y}(z_0)=\frac{\partial u}{\partial y}(z_0)+i\frac{\partial v}{\partial y}(z_0),$$

und

$$df(z_0)=\frac{\partial f}{\partial x}(z_0)dx+\frac{\partial f}{\partial y}(z_0)dy.$$

Dabei sind

$$dx,dy:\mathbb{C}\to\mathbb{C},\qquad dx(z)=dx(x+iy)=x,\quad dy(z)=dy(x+iy)=y.$$

Wir benutzen oft die komplexen Differentiale

$$dz,d\overline{z}:\mathbb{C}\to\mathbb{C},\qquad dz(z)=z,\quad d\overline{z}(z)=\overline{z}.$$

also

$$dz = dx + i\,dy,$$
$$d\overline{z} = dx - i\,dy.$$

9.9.3. **Satz.** *Sei $D \subset \mathbb{C}$ offen, $f : D \longrightarrow \mathbb{C}$ und $z_0 \in D$. Die folgenden Aussagen sind äquivalent:*

(i) f ist komplex-differenzierbar in z_0,

(ii) f ist reell-differenzierbar in z_0 und $df(z_0)$ ist $\mathbb{C}$-linear,

(iii) f ist reell-differenzierbar in z_0 und erfüllt zusätzlich die ***Cauchy-Riemann-Gleichungen****:*

$$\frac{\partial f}{\partial x}(z_0) + i\frac{\partial f}{\partial y}(z_0) = 0 \tag{9.5}$$

d.h.

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{9.6}$$

In diesem Falle gilt

$$df(z_0)(v) = f'(z_0)\cdot v, \quad \textit{für } v \in \mathbb{C}. \tag{9.7}$$

Beweis: (i) $\Longleftrightarrow$ (ii) ist klar, da $\mathrm{Hom}_{\mathbb{C}}(\mathbb{C},\mathbb{C}) \subset \mathrm{Hom}_{\mathbb{R}}(\mathbb{C},\mathbb{C})$.
Zu (ii) $\Longleftrightarrow$ (iii): $df(z_0) \in \mathrm{Hom}_{\mathbb{C}}(\mathbb{C},\mathbb{C}) \Longleftrightarrow \exists\, \lambda \in \mathbb{C}$ mit $df(z_0)(z) = \lambda\cdot z \Longleftrightarrow \exists\, \lambda \in \mathbb{C}$ mit

$$df(z_0)\cdot 1 = \lambda, \quad df(z_0)\cdot i = i\lambda.$$

d.h. das folgende Diagramm kommutiert:

$$\begin{tikzcd}
\mathbb{C} \arrow[r, "z\mapsto\lambda\cdot z"] \arrow[d, "z\mapsto(x,y)"'] & \mathbb{C} \arrow[d, "z\mapsto(x,y)"] \\
\mathbb{R}^2 \arrow[r, "df(z_0)"'] & \mathbb{R}^2
\end{tikzcd}$$

Aber $\frac{\partial f}{\partial x}(z_0) = df(z_0)\cdot e_1$ und $\frac{\partial f}{\partial x}(z_0) = df(z_0)\cdot e_2$, wobei $e_1 = (1,0)$ und $e_2 = (0,1)$. Durch den Isomorphismus $\mathbb{R}^2 \cong \mathbb{C}$, $(x,y) \mapsto x + iy$ werden e_1 und e_2 auf 1 und i abgebildet. Es folgt

$$df(z_0)\cdot 1 = \frac{\partial f}{\partial x}(z_0), \quad df(z_0)\cdot i = \frac{\partial f}{\partial y}(z_0).$$

Also $df(z_0) \in \mathrm{Hom}_{\mathbb{C}}(\mathbb{C},\mathbb{C}) \Longleftrightarrow$ (9.5). Wenn wir Realteil und Imaginärteil von (9.5) betrachten, erhalten wir (9.6). □

Führen wir die folgenden partiellen differentiellen Operatoren ein:

$$\frac{\partial f}{\partial z} = \frac{1}{2}\Big(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\Big), \quad \frac{\partial f}{\partial \overline{z}} = \frac{1}{2}\Big(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\Big) \tag{9.8}$$

Damit gilt

$$df(z_0) = \frac{\partial f}{\partial z}(z_0)dz + \frac{\partial f}{\partial \overline{z}}(z_0)d\overline{z}.$$

Die Cauchy-Riemannschen Gleichungen (9.5) werden

$$\frac{\partial f}{\partial \overline{z}}(z_0) = 0. \tag{9.9}$$

Ist f komplex-diferrenzierbar, so gilt nach (9.5), (9.8)

$$\frac{\partial f}{\partial z}(z_0) = \frac{\partial f}{\partial x}(z_0) = df(z_0)\cdot 1 = f'(z_0). \tag{9.10}$$

Wenn wir z und $\overline{z}$ als Variablen betrachten und eine Funktion

$$f(x,y) = f\Big(\frac{z+\overline{z}}{2}, \frac{z-\overline{z}}{2i}\Big)$$

nach z und $\overline{z}$ mittels Kettenregel ableiten, erhalten wir die Formel (9.8). Dies bedeutet, dass wir $\dfrac{\partial f}{\partial z}$, $\dfrac{\partial f}{\partial \overline{z}}$ durch formelles Differenzieren nach den Variablen z und $\overline{z}$ erhalten.
Dies erleichtert viele Rechnungen. Z.B.

$$\frac{\partial}{\partial z}z^n = nz^{n-1}, \quad \frac{\partial}{\partial \overline{z}}z^n = 0.$$
$$\frac{\partial}{\partial z}|z|^2 = \frac{\partial}{\partial z}(z\overline{z}) = \overline{z}.$$

10. Umkehrsatz und Satz über implizite Funktionen

10.1. **Banachscher Fixpunktsatz.** Viele Gleichungen kann man in die Form $f(x) = x$ bringen. Beispiel: Für gegebenes $a > 0$ ist $x^2 = a$ ist äquivalent zu $x = f(x) := \frac{1}{2}(x + \frac{a}{x})$. Wir nennen einen Punkt ξ mit $f(\xi) = \xi$ einen ***Fixpunkt*** von f (ein Punkt, der unter der Abbildung f festbleibt). Zur Lösung der Gleichung $f(x) = x$ bilden wir, ausgehend von einem Wert $x_0 := a$, nacheinander

$$x_1 = f(x_0), x_2 = f(x_1), \dots, x_n = f(x_{n-1}), \dots$$

und sprechen von einem ***Iterationsverfahren*** oder ***Methode der sukzessiven Approximation*** zur Bestimmung des Fixpunktes. Im Falle der Bestimmung der Quadratwurzel setzen wir $x_0 := a > 0$, $x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n})$; das Verfahren konvergiert ziemlich rasch.

10.1.1. **Definition.** Sei (X,d) ein metrischer Raum. Eine Abbildung $f : X \longrightarrow X$ heißt ***Kontraktion***, falls es $q \in [0,1)$ gibt mit $d(f(x), f(y)) \leqslant q d(x,y)$ für alle $x, y \in X$, d.h. f ist Lipschitz-stetig mit Lipschitz-Konstante $q \in [0,1)$.

10.1.2. **Satz** (Banachscher Fixpunktsatz)**.** *Ist (X,d) ein vollständiger metrischer Raum, und ist $f : X \longrightarrow X$ eine Kontraktion, so hat f genau einen Fixpunkt $\xi \in X$. Für jedes $x \in X$ konvergiert die durch $x_0 := x$ und $x_n = f(x_{n-1})$ für $n \geqslant 1$ definierte Iterationsfolge $(x_n)_{n\geqslant 0}$ gegen ξ, und es gilt die Abschätzung*

$$d(x_n, \xi) \leqslant \frac{q^n}{1-q} d(x_1, x_0). \tag{10.1}$$

Wie überprüft man, dass eine Abbildung eine Kontraktion ist? Ein wichtiges Kriterium liefert den Schrankensatz 9.3.7: Ist $D \in \mathbb{R}^n$ konvex und $f : D \to D$ differenzierbar mit

$$q := \sup_D \|df(x)\|_{op} < 1,$$

so ist f eine Kontraktion, da für alle $x, y \in D$ gilt

$$\|f(y) - f(x)\| \leqslant \sup_{t\in[0,1]} \|df((1-t)x + ty)\|_{op} \, \|y - x\| \leqslant q \, \|y - x\| .$$

Wir bestimmen z.B. auf vier Nachkommastellen gerundet die reelle Lösung der Gleichung: $x^3 + 12x - 1 = 0$.

Es gibt eine eindeutige reelle Lösung im Intervall $X = [0,1]$. Durch Umformungen wird die Gleichung $x = \frac{1}{x^2+12}$; d.h. $f(x) = \frac{1}{x^2+12}$. Wir finden

$$q = \sup_{x\in[0,1]} |f'(x)| = \sup_{x\in[0,1]} \frac{2x}{(x^2+12)^2} = \frac{2}{169}.$$

Sei $x_0 = 0$, dann ist $x_1 = f(0) = \frac{1}{12}$ und $d(x_0, x_1) = \frac{1}{12}$. Aus der Formel (10.1) bestimmen wir n minimal, so dass

$$\frac{d(x_0,x_1)}{1-q} \cdot q^n < 10^{-4} \iff \frac{1}{12} \cdot \frac{169}{167} \cdot \Big(\frac{2}{169}\Big)^n < 10^{-4} \iff \Big(\frac{2}{169}\Big)^n < \frac{167 \cdot 12}{169 \cdot 10^4}.$$

Daraus folgt $n = 2$. Dann ist $\xi \cong x_2 = f(x_1) = \frac{1}{x_1^2+12} = \frac{144}{1729} \cong 0{,}08328$.

10.2. **Der Umkehrsatz.**
Die Diffeomorphismen spielen in der Analysis die Rolle der Koordinatenwechsel, also die Rolle der Isomorphismen in der Linearen Algebra. Sie vereinfachen oft die Berechnung von Integralen und das Lösen von Differentialgleichungen.

10.2.1. **Definition.** Seien $D \subset \mathbb{R}^n$, $G \subset \mathbb{R}^m$ offen, und sei $k \in \mathbb{N} \cup \{\infty\}$.

$f : D \longrightarrow G$ heißt ***$\mathcal{C}^k$-Diffeomorphismus***, wenn $f \in \mathcal{C}^k(D,G)$, f bijektiv und $f^{-1} \in \mathcal{C}^k(G,D)$ ist.

f heißt ***lokaler $\mathcal{C}^k$-Diffeomorphismus*** in $a \in D$, wenn offene Umgebungen U von a und V von $f(a)$ existieren so, dass $f : U \longrightarrow V$ ein $\mathcal{C}^k$-Diffeomorphismus ist.

f heißt ***lokaler $\mathcal{C}^k$-Diffeomorphismus***, wenn f ein lokaler $\mathcal{C}^k$-Diffeomorphismus in jedem $x \in D$ ist.

10.2.2. **Beispiele.**

(1) $\exp : \mathbb{R} \longrightarrow \mathbb{R}_+$ ist ein $\mathcal{C}^\infty$-Diffeomorphismus mit der Inversen $\log : \mathbb{R}_+ \longrightarrow \mathbb{R}$.

(2) Polarkoordinaten in $\mathbb{R}^2$. Die (Einschränkung der) Polarkoordinatenabbildung

$$P : \mathbb{R}_+ \times (-\pi, \pi) \longrightarrow \mathbb{R}^2 \setminus \{(x,y) : x \leqslant 0, y = 0\} =: \mathbb{R}^2_-, \quad P(r,\vartheta) = (r\cos\vartheta, r\sin\vartheta)$$

ist ein $\mathcal{C}^\infty$-Diffeomorphismus. Laut Satz 8.4.11 ist die Inverse von P gegeben durch

$$P^{-1}:\mathbb{R}^2_- \to \mathbb{R}_+\times(-\pi,\pi),\quad (x,y)\mapsto\left(\sqrt{x^2+y^2},\arg(x,y)\right)$$

wobei (siehe (8.5) und (8.6))

$$\arg:\mathbb{R}^2_-\to(-\pi,\pi),\quad \arg(x,y)=\begin{cases}\arccos\frac{x}{\sqrt{x^2+y^2}}, & x\geqslant 0,\\ -\arccos\frac{x}{\sqrt{x^2+y^2}}, & x<0.\end{cases}$$

Die Funktion $\mathbb{R}^2_-\to\mathbb{R}_+$, $(x,y)\mapsto\sqrt{x^2+y^2}$ ist glatt, als Zusammensetzung von $\mathbb{R}^2_-\to\mathbb{R}_+$, $(x,y)\mapsto x^2+y^2$ und $\mathbb{R}_+\to\mathbb{R}_+$, $t\mapsto\sqrt{t}$. Wir beweisen, dass $\arg:\mathbb{R}^2_-\to(-\pi,\pi)$ glatt ist, wie im Satz 8.4.9. Betrachte die offenen Mengen in $\mathbb{R}^2_-$:

$$U_1=\{(x,y):y>0\},\quad U_2=\{(x,y):y<0\},\quad U_3=\{(x,y):x>0\}$$

und die glatten Abbildungen

$$f_1:U_1\to\mathbb{R},\, f_1(x,y)=\operatorname{arccot}\frac{x}{y},\quad f_2:U_2\to\mathbb{R},\, f_2(x,y)=-\operatorname{arccot}\frac{x}{y},$$
$$f_3:U_3\to\mathbb{R},\, f_3(x,y)=\arctan\frac{y}{x}\cdot$$

Es ist $\mathbb{R}^2_-=U_1\cup U_2\cup U_3$ und

$$\arg(x,y)=\begin{cases}f_1(x,y), & z\in U_1,\\ f_2(x,y), & z\in U_2,\\ f_3(x,y), & z\in U_3.\end{cases}$$

woraus folgt, dass $\arg:\mathbb{R}^2_-\to(-\pi,\pi)$ glatt ist: gegeben $(x_0,y_0)\in\mathbb{R}^2_-$ beliebig, so gibt es $k\in\{1,2,3\}$ mit $(x_0,y_0)\in U_k$ also existieren $\partial^\alpha\arg(x_0,y_0)=\partial^\alpha f_k(x_0,y_0)$ und sind stetig für alle $\alpha\in\mathbb{N}_0^2$. Berechnen wir die Jacobi-Matrix er Abbildung P^{-1}. Zunächst

$$\frac{\partial}{\partial x}\arg(x,y)=\frac{\partial}{\partial x}\operatorname{arccot}\frac{x}{y}=-\frac{\partial}{\partial x}\arctan\frac{y}{x}=\frac{1}{1+\left(\frac{y}{x}\right)^2}\cdot\left(-\frac{y}{x^2}\right)=\frac{-y}{x^2+y^2}$$

und

$$\frac{\partial}{\partial y}\arg(x,y)=\frac{x}{x^2+y^2}$$

also für $x=r\cos\vartheta$, $y=r\sin\vartheta$,

$$J_{P^{-1}}(x,y)=\begin{pmatrix}\frac{x}{\sqrt{x^2+y^2}} & \frac{x}{\sqrt{x^2+y^2}}\\ \frac{y}{x^2+y^2} & \frac{x}{x^2+y^2}\end{pmatrix}=\begin{pmatrix}\cos\vartheta & \sin\vartheta\\ -\frac{1}{r}\sin\vartheta & \frac{1}{r}\cos\vartheta\end{pmatrix}.$$

Da

$$A=\begin{pmatrix}a & b\\ c & d\end{pmatrix},\det A\neq 0\rightsquigarrow A^{-1}=\frac{1}{\det A}\begin{pmatrix}d & -b\\ -c & a\end{pmatrix}$$

so folgt

$$\left(J_{P^{-1}}(x,y)\right)^{-1}=r\begin{pmatrix}\frac{1}{r}\cos\vartheta & -\sin\vartheta\\ \frac{1}{r}\sin\vartheta & \cos\vartheta\end{pmatrix}=\begin{pmatrix}\cos\vartheta & -r\sin\vartheta\\ \sin\vartheta & r\cos\vartheta\end{pmatrix}=J_P(r,\vartheta).$$

(3) $f:\mathbb{R}\longrightarrow\mathbb{R}$, $f(x)=x^3$ ist $\mathcal{C}^\infty$, bijektiv (sogar ein Homöomorphismus), aber kein Diffeomorphismus, weil $f^{-1}(y)=y^{1/3}$ nicht differenzierbar in $y=0$ ist.

(4) Die Polarkoodinatenabbildung $P:\mathbb{R}_+\times\mathbb{R}\longrightarrow\mathbb{R}^2\setminus\{(0,0)\}$ ist ein lokaler Diffeomorphismus, aber kein Diffeomorphismus, weil sie ist nicht injektiv ist.

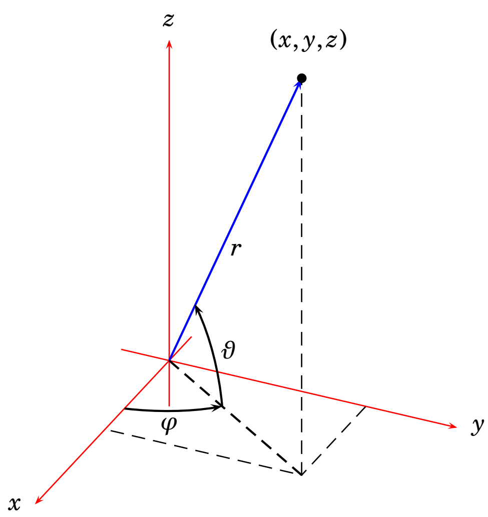
(5) Polarkoordinaten in $\mathbb{R}^3$ oder Kugelkoordinaten. Wir definieren

$$\begin{aligned}&P_3:\mathbb{R}_+\times(-\pi,\pi)\times(-\tfrac{\pi}{2},\tfrac{\pi}{2})\longrightarrow\mathbb{R}^3\\ &P_3(r,\varphi,\vartheta)=(P_2(r,\varphi)\cos\vartheta,r\sin\vartheta)=(r\cos\varphi\cos\vartheta,r\sin\varphi\cos\vartheta,r\sin\vartheta).\end{aligned}\tag{10.2}$$

Wir beweisen, dass P_3 ein Diffeomorphismus auf ihr Bild $\mathbb{R}^3 \setminus \{(x,y,z) : x \leqslant 0, y = 0\}$ ist. Wir schreiben $P_3 = \Psi_2 \circ \Psi_1$, wobei

$$\begin{aligned}
&\Psi_1 : \mathbb{R}_+ \times (-\pi,\pi) \times (-\tfrac{\pi}{2}, \tfrac{\pi}{2}) \longrightarrow \mathbb{R}_+ \times (-\pi,\pi) \times \mathbb{R},\\
&\Psi_1(r,\varphi,\vartheta) = (r\cos\vartheta, \varphi, r\sin\vartheta)\\
&\Psi_2 : \mathbb{R}_+ \times (-\pi,\pi) \times \mathbb{R} \longrightarrow \mathbb{R}^3 \setminus \{(x,0,z) : x \leqslant 0\},\\
&\Psi_2(\rho,\varphi,z) = \big(P_2(\rho,\varphi), z\big)
\end{aligned}$$

Da Ψ_1, Ψ_2 Diffeomorphismen sind, ist auch $\Psi_2 \circ \Psi_1$ ein Diffeomorphismus.



Berechnen wir $\det J_{P_3}(r,\varphi,\vartheta)$. Laut Kettenregel gilt $J_{P_3} = J_{\Psi_2}(\Psi_1) J_{\Psi_1}$ und $\det J_{\Psi_1} = r$, $\det J_{\Psi_2} = \rho$, also $\det J_{P_3} = r\cos\varphi \cdot r = r^2\cos\varphi$.

(6) Polarkoordinaten in $\mathbb{R}^n$. Wir definieren die Polarkoordinatenabbildung in $\mathbb{R}^n$ durch Rekursion:

$$\begin{aligned}
&P_n : \mathbb{R}_+ \times (-\pi,\pi) \times (-\tfrac{\pi}{2}, \tfrac{\pi}{2})^{n-2} \longrightarrow \mathbb{R}^n \setminus \{x \in \mathbb{R}^n : x_1 \leqslant 0, x_2 = 0\}\\
&P_n(r,\varphi,\vartheta_1,\dots,\vartheta_{n-2}) = (P_{n-1}(r,\varphi,\vartheta_1,\dots,\vartheta_{n-3})\cos\vartheta_{n-2}, r\sin\vartheta_{n-2})
\end{aligned} \tag{10.3}$$

Es ergibt sich $P_n(x) = y$ wobei

$$\begin{aligned}
y_1 &= r\cos\varphi\cos\vartheta_1\cdots\cos\vartheta_{n-2},\\
y_2 &= r\sin\varphi\cos\vartheta_1\cdots\cos\vartheta_{n-2},\\
y_3 &= r\sin\vartheta_1\cos\vartheta_2\cdots\cos\vartheta_{n-2},\\
&\vdots\\
y_{n-1} &= r\sin\vartheta_{n-3}\cos\vartheta_{n-2},\\
y_n &= r\sin\vartheta_{n-2}.
\end{aligned}$$

Behauptung:

P_n ist ein Diffeomorphismus und

$$\det J_{P_n}(r,\varphi,\vartheta_1,\dots,\vartheta_n) = r^{n-1}\cos\vartheta_1(\cos\vartheta_2)^2\cdots(\cos\vartheta_{n-2})^{n-2}.$$

Der Beweis erfolgt durch Induktion über n: Wir schreiben $P_n = \Psi_2 \circ \Psi_1$, wobei

$$\begin{aligned}
&\Psi_1 : \mathbb{R}_+ \times (-\pi,\pi) \times (-\tfrac{\pi}{2}, \tfrac{\pi}{2})^{n-2} \longrightarrow (0,\infty) \times (-\pi,\pi) \times (-\tfrac{\pi}{2}, \tfrac{\pi}{2})^{n-3} \times \mathbb{R},\\
&\Psi_1(r,\varphi,\vartheta_1,\dots,\vartheta_{n-2}) = (r\cos\vartheta_{n-2}, \varphi, \vartheta_1,\dots,\vartheta_{n-3}, r\sin\vartheta_{n-2}).\\
&\Psi_2 : \mathbb{R}_+ \times (-\pi,\pi) \times (-\tfrac{\pi}{2}, \tfrac{\pi}{2})^{n-3} \times \mathbb{R} \longrightarrow \mathbb{R}^n \setminus \{x : x_1 \leqslant 0, x_2 = 0\},\\
&\Psi_2(\rho,\varphi,\vartheta_1,\dots,\vartheta_{n-3},x_n) = (P_{n-1}(\rho,\varphi,\vartheta_1,\dots,\vartheta_{n-3}), x_n).
\end{aligned}$$

Ψ_1, Ψ_2 sind Diffeomorphismen also P_n ist ein auch Diffeomorphismus. Außerdem gilt $\det J_{\Psi_1} = r$, $\det J_{\Psi_2} = \det J_{P_{n-1}}$ und letzteres ist durch die Induktionsannahme bekannt. Folglich

$$\begin{aligned}\det J_{P_n} &= \det J_{\Psi_2}(\Psi_1)\cdot \det J_{\Psi_1} = \det J_{P_{n-1}}(\rho,\varphi,\vartheta_1,\dots,\vartheta_{n-3})\cdot r\\ &= r(r\cos\vartheta_{n-2})^{n-2}\cos\vartheta_1(\cos\vartheta_2)^2\cdots(\cos\vartheta_{n-3})^{n-3}\\ &= r^{n-1}\cos\vartheta_1(\cos\vartheta_2)^2\cdots(\cos\vartheta_{n-3})^{n-3}(\cos\vartheta_{n-2})^{n-2}.\end{aligned}$$

10.2.3. **Satz.** *Seien $D\subset\mathbb{R}^n$, $G\subset\mathbb{R}^m$ offen, $k\in\mathbb{N}\cup\{\infty\}$. Sei $f: D\longrightarrow G$ ein $\mathcal{C}^k$-Diffeomorphismus, $x\in D$, $y=f(x)$. Dann gilt $d(f^{-1})(y)=[df(x)]^{-1}$ und $J_{f^{-1}}(y)=[J_f(x)]^{-1}$. Insbesondere gilt $m=n$.*

10.2.4. **Satz.** *Sei $f: D\longrightarrow G$ ein Homöomorphismus mit $f\in\mathcal{C}^k(D,G)$. Die folgenden Aussagen sind äquivalent:*

(i) *f ist ein $\mathcal{C}^k$-Diffeomorphismus,*
(ii) *$df(x):\mathbb{R}^n\longrightarrow\mathbb{R}^n$ ist für jedes $x\in D$ ein Isomorphismus.*

10.2.5. **Satz** (Umkehrsatz)**.** *Seien $D,G\subset\mathbb{R}^n$ offen, $f\in\mathcal{C}^k(D,G)$, $a\in D$, $y_0=f(a)\in G$. Ist $df(a)$ ein Isomorphismus, so ist f ein lokaler $\mathcal{C}^k$-Diffeomorphismus in a.*

10.2.6. **Bemerkung.** $df(a)$ ist ein Isomorphismus genau dann, wenn $\det J_f(a)\neq 0$.

Der Satz besagt, dass die Gleichung $f(x)=y$ sich nahe a bzw. y_0 eindeutig lösen lässt, wenn die lineare Gleichung $df(a)\cdot v=w$ sich für jedes $w\in\mathbb{R}^n$ eindeutig lösen lässt, d.h. wenn die Zahl $\det J_f(a)\neq 0$ ist.

10.2.7. **Folgerung** (Diffeomorphiesatz)**.** *Seien $D,G\subset\mathbb{R}^n$ offen, und sei $f\in\mathcal{C}^k(D,G)$ bijektiv, $k\geqslant 1$. Falls $df(a)$ für jedes $a\in D$ invertierbar ist, so folgt $f^{-1}\in\mathcal{C}^k(G,D)$, d.h. f ist ein $\mathcal{C}^k$-Diffeomorphismus.*

10.2.8. **Beispiel.** Der Diffeomorphiesatz vereinfacht die Beweise, dass eine gegebene Abbildung ein Differomeorphismus ist; es genügt zu zeigen, dass die Abbildung bijektiv und von der Klasse $\mathcal{C}^k$ ist und zusätzlich, dass das Differential in jedem Punkt bijektiv ist. Man braucht nicht mehr die Inverse zu berechnen und nachzuweisen, dass sie $\mathcal{C}^k$ ist.

Betrachten wir nochmals die Abbildung $f:\mathbb{R}_+\times\mathbb{R}\longrightarrow\mathbb{R}^2$, $f(r,\varphi)=(r\cos\varphi, r\sin\varphi)$. Sie bildet $D=\mathbb{R}_+\times(\varphi_0-\pi,\varphi_0+\pi)$ bijektiv auf $G=\mathbb{R}^2\setminus\{te^{i\varphi_0}: t\leqslant 0\}$ ab. Außerdem gilt für alle $(r,\varphi)\in D$:

$$J_f(r,\varphi)=\begin{pmatrix}\cos\varphi & -r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{pmatrix},\quad \det J_f(r,\varphi)=r\neq 0.$$

Daraus folgt dass $df(r,\varphi)$ für jedes $(r,\varphi)\in\mathbb{R}_+\times\mathbb{R}$ invertierbar ist. Nach Folgerung 10.2.7 ist $f: D\to G$ ein $\mathcal{C}^\infty$-Diffeomorphismus.

Genauso zeigt man, dass die Kugelkoordinatenabbildung (10.2) ein $\mathcal{C}^\infty$-Diffeomorphismus auf ihr Bild ist (Übungsaufgabe).

10.2.9. **Bemerkung.** Ohne die Voraussetzung der Bijektivität ist der Diffeomorphiesatz falsch: Betrachte die Abbildung $f:\mathbb{R}^2\to\mathbb{R}^2$, $f(x,y)=(e^x\cos y, e^x\sin y)$, deren Jacobi-Determinante $\det J_f(x,y)=e^{2x}\neq 0$ ist. Die Abbildung ist aber nicht injektiv, denn $f(x,y+2k\pi)=f(x,y)$.

Man kann die Diffeomorphismen von $\mathbb{R}^n$ so charakterisieren:

10.2.10. **Satz** (Hadamard)**.** *Sei $f:\mathbb{R}^n\to\mathbb{R}^n$ stetig differenzierbar. Die folgenden Bedingungen sind äquivalent zueinander:*

(i) *f ist ein Diffeomorphismus von $\mathbb{R}^n$ auf $\mathbb{R}^n$.*

(ii) *$df(a)$ ist invertierbar für jedes $a\in\mathbb{R}^n$ und f ist eigentlich (d.h. $f^{-1}(K)$ ist kompakt für jede kompakte Teilmenge $K\subset\mathbb{R}^n$).*

10.3. **Der Satz über implizite Funktionen.**

Sei $f:\mathbb{R}^{k+m}\longrightarrow\mathbb{R}^m$ eine Abbildung, $z_0\in\mathrm{Im}(f)$, $(x_0,y_0)\in\mathbb{R}^k\times\mathbb{R}^m\longrightarrow\mathbb{R}^{k+m}$ liege in der Niveaumenge $f^{-1}(z_0)$. Wir möchten die Gleichung $f(x,y)=z_0$ nahe (x_0,y_0) durch eine Funktion $y=g(x)$ auflösen.

Beispiel: $f:\mathbb{R}^2\longrightarrow\mathbb{R}$, $(x,y)\mapsto y-x^2$, $z_0:=0$, $(x_0,y_0):=(0,0)$. Dann ist $y-x^2=0$ (nahe $(0,0)$) auflösbar durch $y=g(x):=x^2$. Sie ist aber nicht auflösbar in der Form $x=g(y)$: Es gibt für $y>0$ stets zwei Lösungen $x=\pm\sqrt{y}$ und für $y<0$ gar keine Lösung. Anders ausgedrückt: Die Menge $\{(x,y)\in\mathbb{R}^2: y=x^2\}$ ist ein Graph über der x-Achse, aber kein Graph über der y-Achse.

Wir führen nun die folgende Notation ein. Die Punkte in $\mathbb{R}^{k+m}$ schreiben wir als $(x,y)=(x_1,\dots,x_k,y_1,\dots,y_m)$. Ist $U\subset\mathbb{R}^{k+m}$ offen, $a\in U$ und $f\in\mathcal{C}^1(U,\mathbb{R}^m)$, also jeweils $df(a):\mathbb{R}^{k+m}\longrightarrow\mathbb{R}^m$ linear, so seien $d_xf(a)$ bzw. $d_yf(a)$ die Einschränkungen von $df(a)$ auf die Faktoren $\mathbb{R}^k$ bzw. $\mathbb{R}^m$ von $\mathbb{R}^k\times\mathbb{R}^m=\mathbb{R}^{k+m}$. Also:

$$d_xf(a):\mathbb{R}^k\longrightarrow\mathbb{R}^m,\quad d_xf(a)h=df_a(h,0),$$

$$d_yf(a):\mathbb{R}^m\longrightarrow\mathbb{R}^m,\quad d_yf(a)\widetilde{h}=df_a(0,\widetilde{h}),$$

$$df(a)(h,\widetilde{h})=d_xf(a)h+d_yf(a)\widetilde{h}.$$

Die Abbildgungen $d_xf(a)$ und $d_yf(a)$ sind linear. Die zugehörigen Matrizen bezüglich der Standardbasis sind

$$\frac{\partial f}{\partial x}(a):=\begin{pmatrix}\partial_1f_1(a) & \dots & \partial_kf_1(a)\\ \vdots & & \vdots\\ \partial_1f_m(a) & \dots & \partial_kf_m(a)\end{pmatrix}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_k}(a)\\ \vdots & & \vdots\\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_k}(a)\end{pmatrix}$$

zu $d_xf(a)$, bzw.

$$\frac{\partial f}{\partial y}(a):=\begin{pmatrix}\partial_{k+1}f_1(a) & \dots & \partial_{k+m}f_1(a)\\ \vdots & & \vdots\\ \partial_{k+1}f_m(a) & \dots & \partial_{k+m}f_m(a)\end{pmatrix}=\begin{pmatrix}\frac{\partial f_1}{\partial y_1}(a) & \dots & \frac{\partial f_1}{\partial y_m}(a)\\ \vdots & & \vdots\\ \frac{\partial f_m}{\partial y_1}(a) & \dots & \frac{\partial f_m}{\partial y_m}(a)\end{pmatrix}$$

zu $d_yf(a)$.

10.3.1. **Satz** (Satz über implizite Funktionen)**.** *Sei $U\subset\mathbb{R}^{k+m}$ offen, $f\in\mathcal{C}^\ell(U,\mathbb{R}^m)$, $(x_0,y_0)\in U$, $z_0=f(x_0,y_0)$. Sei $d_yf(x_0,y_0):\mathbb{R}^m\longrightarrow\mathbb{R}^m$ invertierbar. Dann gibt es eine offene Umgebung der Form $U'\times U''\subset U$ von (x_0,y_0) und eine Abbildung $g\in\mathcal{C}^\ell(U',U'')$ mit $g(x_0)=y_0$ so, dass für alle $(x,y)\in U'\times U''$ gilt: $f(x,y)=z_0$ genau dann, wenn $y=g(x)$.*

10.3.2. **Folgerung.** *Es gilt:*

$$\boxed{dg(x_0)=-(d_yf(x_0,y_0))^{-1}\circ d_xf(x_0,y_0)}\,.$$

Ist $f:\mathbb{R}^{k+1}\to\mathbb{R}$ (d.h. $m=1$), so gilt:

$$\boxed{\operatorname{grad}g(x_0)=-\frac{1}{\partial_{k+1}f(x_0,y_0)}(\partial_1f,\dots,\partial_kf)(x_0,y_0)}\,.$$

Beweis: Für alle $x\in U'$ gilt $F(x):=f(x,g(x))=z_0$, d.h. F ist konstant, also für alle $h\in\mathbb{R}^m$:

$$0=dF(x)\cdot h=df\big((x,g(x))\big)\cdot(h,dg(x)\cdot h)=d_xf(x,g(x))\cdot h+d_yf(x,g(x))\cdot(dg(x)\cdot h).$$

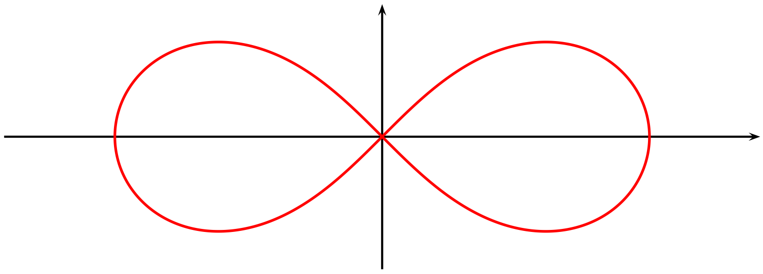
Daraus folgt: $dg(x_0)\cdot h=-(d_yf(x_0,y_0))^{-1}d_xf(x_0,y_0)\cdot h$. □

10.3.3. **Beispiel.**

(1) $f:\mathbb{R}^2\longrightarrow\mathbb{R}$, $f(x,y)=x^2+y^2$ (hier $k=1$, $m=1$). Dann ist $\frac{\partial f}{\partial y}(x,y)=2y$. Für alle $y_0\neq0$ und alle x_0 gilt also: $x^2+y^2=z_0=x_0^2+y_0^2$ kann lokal um (x_0,y_0) durch eine Gleichung $y=g(x)$ aufgelöst werden. (Klar auch ohne Satz über implizite Funktionen: $g(x):=\operatorname{sign}y_0\cdot\sqrt{z_0-x_0^2}$.) Analog für $x_0\neq0$.

(2) $f:\mathbb{R}^2\longrightarrow\mathbb{R}$, $f(x,y)=x^2-y^2$. Sei $z_0=0$, dann ist $f^{-1}(0)$ die Vereinigung der zwei Winkelhalbierenden $y=x$ und $y=-x$. Es ist $\operatorname{grad}f(x,y)=(2x,-2y)$. Daraus folgt: Für alle $(x_0,y_0)\neq(0,0)$ in $f^{-1}(0)$ ist $f(x,y)=0$ nahe (x_0,y_0) sowohl nach x als auch nach y auflösbar. Denn hier ist $2y_0\neq0\Leftrightarrow 2x_0\neq0$ wegen $x_0^2=y_0^2$. Im Nullpunkt ist die Gleichung weder nach x noch nach y auflösbar.

(3) Die ***Lemniskate*** von Bernoulli ist der Ort der Punkte $(x,y)\in\mathbb{R}^2$, deren Entfernungen von zwei festen Punkten $(-1,0)$ und $(1,0)$ das konstante Produkt 1 haben:



Die Lemniskate ist also die Menge der Punkte $(x,y) \in \mathbb{R}^2$, für die gilt:

$$|(x,y)-(1,0)|\cdot|(x,y)-(-1,0)| = 1 \iff$$

$$1 = ((x-1)^2+y^2)\cdot((x+1)^2+y^2) = (x^2+1+y^2)^2-4x^2 \iff$$

$$0 = (x^2+1+y^2)^2-4x^2-1 = (x^2+y^2)^2-2x^2+2y^2 =: f(x,y).$$

Der Gradient der Funktion f ist:

$$\operatorname{grad} f(x,y) = (4x(x^2+y^2)-4x, 4y(x^2+y^2)+4y) = 4(x(x^2+y^2-1), y(x^2+y^2+1)).$$

Ist $y_0 \neq 0$, so ist die Gleichung $f(x,y)=0$ nahe $(x_0,y_0) \in f^{-1}(0)$ auflösbar durch $y = g(x)$. Ist $x_0 \neq 0$ und $x_0^2+y_0^2 \neq 1$, so ist die Gleichung $f(x,y)=0$ auch nach x auflösbar nahe (x_0,y_0).

Wo gilt $x_0^2+y_0^2=1$? $f(x_0,y_0) = 1-2x_0^2+2y_0^2 = 3-4x_0^2 \Leftrightarrow x_0 = \pm\frac{\sqrt{3}}{2}$, $y_0 = \pm\frac{1}{2}$. In der Nähe der vier Punkte $(\pm\frac{\sqrt{3}}{2}, \pm\frac{1}{2})$ ist $f(x,y)=0$ jeweils nicht nach x auflösbar.

Nahe $(x_0,y_0) = (0,0)$ ist die Gleichung weder nach x noch nach y auflösbar.

(4) $f: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, $f(x,y,z) := (x^3+y^3+z^3-7, xy+yz+zx+2)$. Betrachte $f^{-1}(0,0)$ nahe $(2,-1,0)$.

$$\begin{pmatrix} \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \end{pmatrix}(x,y,z) = \begin{pmatrix} 3y^2 & 3z^2 \\ x+z & x+y \end{pmatrix}.$$

Es ist $\begin{pmatrix} \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \end{pmatrix}(2,-1,0) = \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}$. Diese Matrix ist invertierbar.

Es gibt also eine Umgebung U' von 2 in $\mathbb{R}$, eine Umgebung U'' von $(-1,0)$ in $\mathbb{R}^2$ und $g \in \mathcal{C}^\infty(U',U'')$ mit

$$f(x,y,z) = (0,0) \iff (y,z) = g(x) \text{ für alle } (x,y,z) \in U' \times U'',$$

und es gilt

$$g'(0) = -\begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} \frac{\partial f_1}{\partial x}(2,-1,0) \\ \frac{\partial f_2}{\partial x}(2,-1,0) \end{pmatrix} = -\frac{1}{3}\begin{pmatrix} 1 & 0 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 12 \\ -1 \end{pmatrix} = -\frac{1}{3} \cdot \begin{pmatrix} 12 \\ -27 \end{pmatrix} = \begin{pmatrix} -4 \\ 9 \end{pmatrix}.$$

(5) Sei $V = M_{n\times n}(\mathbb{C}) \cong \mathbb{R}^{2n^2}$, $f: V \times V \longrightarrow V$, $f(X,Y) := X - Y^2$.

Betrachte $f^{-1}(0)$ nahe (E_n, E_n). Es gilt

$$df(E_n,E_n)\cdot(0,H) = \frac{d}{dt}|_{t=0}(-(E_n+tH)^2) = -2H.$$

Daraus folgt: $d_Y f(0,0) = -2\,\mathrm{Id}$ ist invertierbar. Es gibt also offene Umgebungen U', U'' von E_n und $g \in \mathcal{C}^\infty(U',U'')$ mit:

$$X = Y^2 \iff Y = g(X)$$

für alle $(X,Y) \in U' \times U''$. Die Funktion g ist die „Wurzel“ von X und ist definiert nahe E_n. Außerdem gilt

$$dg(E_n) = -(d_Y f(E_n,E_n))^{-1} \circ d_X f(E_n,E_n) = -(-2\,\mathrm{Id})^{-1} \circ \mathrm{Id} = \frac{1}{2}\mathrm{Id},$$

folglich nach Taylor:

$$\sqrt{E_n + H} = E_n + \frac{1}{2}H + o(\|H\|).$$

Geometrische Interpretation:

• Linearer Fall. Nehmen wir an, dass $f : \mathbb{R}^{k+m} \longrightarrow \mathbb{R}^m$ linear ist. Sei $z_0 = 0$, $a \in f^{-1}(0)$. Dann ist $df(a) = f$ und $d_y f(a) = f|_{\{0\}\times\mathbb{R}^m}$, und $\frac{\partial f}{\partial y}(a)$ ist der linke $(m\times m)$-Block in der assoziierten Matrix $M(f)$ von f bzgl. der kanonischen Basen.

Die Hypothese, dass $df_y(a)$ invertierbar ist, impliziert $\operatorname{Rang} f = m$. Aus der Linearen Algebra wissen wir, dass $\operatorname{Ker} f$ ein k-dimensionaler Untervektorraum von $\mathbb{R}^{k+m}$ ist. Darüber hinaus kann das lineare Gleichungssystem $f(x,y) = 0$ nach y aufgelöst werden, d.h. es existiert $g : \mathbb{R}^k \longrightarrow \mathbb{R}^m$ mit $f(x,y) = 0 \Leftrightarrow y = g(x)$. Insbesondere ist $\operatorname{Ker} f = f^{-1}(0)$ ein Graph über $\mathbb{R}^k$.

Also steht die Aussage des Satzes über implizite Funktionen in Einklang mit diesen wohlbekannten Ergebnissen aus der Linearen Algebra.

• Allgemeiner Fall. Auch wenn f nicht linear ist, wird das lokale Verhalten von f nahe (x_0,y_0) durch die Eigenschaften des Differentials $df(x_0,y_0)$ bestimmt.

Ist $d_y f(x_0,y_0) : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ ein Isomorphismus, so ist $\operatorname{Ker} df(x_0,y_0)$ ein k-dimensionaler Untervektorraum von $\mathbb{R}^{k+m}$ und ein Graph über $\mathbb{R}^k$. Dann besagt der Satz über implizite Funktionen, dass auch $f^{-1}(z_0)$ lokal um (x_0,y_0) ein Graph über $\mathbb{R}^k$ ist. Wir werden später sehen, dass $f^{-1}(z_0)$ einen Tangentialraum in (x_0,y_0) (beste lineare Approximation) besitzt, nämlich $(x_0,y_0) + \operatorname{Ker} df(x_0,y_0)$.

10.4. **Extrema unter Nebenbedingungen.**

Wir betrachten nun das Problem, für eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ lokale Extrema der Einschränkung von f auf eine Menge $M \subset \mathbb{R}^n$ zu finden. Dabei sei M gegeben durch $M = \varphi^{-1}(0)$ mit einer Abbildung $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Gesucht sind Extrema von $f|_M$, d.h. Extrema von f unter der ***Nebenbedingung*** $\varphi(x_1,\dots,x_n) = 0$ für die Variablen $x_1,\dots x_n$.

Als Beispiel betrachten wir $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $f(x_1,x_2) = x_1x_2$ und $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $\varphi(x_1,x_2) = x_1^2 + x_2^2 - 1$. Gesucht sind die Extrema von f unter Nebenbedingung $\varphi = 0$, d.h. Extrema von $f|_{\varphi^{-1}(0)} = f|_{S^1}$. Für diese Funktion können wir die bekannten Kriterien für Extrema nicht direkt anwenden, da S^1 keine offene Menge in $\mathbb{R}^2$ ist. Zum Beispiel gilt $\operatorname{grad} f \neq 0$ überall auf S^1, obwohl $f|_{S^1}$ Extrema hat (Satz vom Maximum), weil S^1 kompakt ist.

Wie können wir das Fermat-Kriterium ersetzen? Ein Weg ist, die Nebenbedingung nach einer Variablen aufzulösen: $\varphi(x_1,x_2) = 0 \Leftrightarrow x_2 = \pm\sqrt{1-x_1^2}$ oder $x_1 = \pm\sqrt{1-x_2^2}$ und dann die Extrema der zugehörigen zusammengesetzten Funktionen, z.B. $(-1,1) \ni x_1 \mapsto f(x_1, \sqrt{1-x_1^2})$, zu untersuchen.

So haben wir das Problem auf die Bestimmung von Extrema von auf offenen Mengen definierten Funktionen zurückgeführt. Im Allgemeinen könnte man ebenso versuchen, das System $\varphi(x_1,\dots,x_n) = 0$ nach den Variablen $x_{n-m+1},\dots,x_n$ aufzulösen:

$$\varphi(x_1,\dots,x_n) = 0 \quad\Longleftrightarrow\quad x_{n-m+1} = \psi_1(x_1,\dots,x_{n-m}),\dots,x_n = \psi_m(x_1,\dots,x_{n-m})$$

wobei $\psi_1,\dots\psi_m$ implizite, durch das System $\varphi = 0$ definierte Funktionen sind. Somit reduziert sich das Problem eines Extremum mit Nebenbedingungen für f in einem Punkt $(x_1^\circ,\dots,x_n^\circ)$ auf das Problem eines gewöhnlichen Extremums für die zusammengesetzte Funktion

$$(x_1,\dots,x_{n-m}) \mapsto f(x_1,\dots,x_{n-m},\psi_1(x_1,\dots,x_{n-m}),\dots,\psi_m(x_1,\dots,x_{n-m}))$$

im Punkt $(x_1^\circ,\dots,x_{n-m}^\circ)$.

Diese Überlegungen deuten auch die Annahmen an, die die Nebenbedingungen erfüllen müssen, damit auf diese Weise die Extrema gefunden werden können: Laut dem Satz über implizite Funktionen sind die Nebenbedingungen nach m (geeigneten) Variablen dann auflösbar, wenn die Jacobi-Matrix J_φ den Rang m hat, d.h. wenn $\operatorname{grad}\varphi_1,\dots,\operatorname{grad}\varphi_m$ linear unabhängig sind.

Eine Auflösung der Gleichungsnebenbedingungen wie im Beispiel ist nicht immer möglich oder wird unüberschaubar. Wir geben nun einen Weg zur Bestimmung von Extrema von $f|_M$ an, welcher *keine* expliziten Ausdrücke für die impliziten Funktionen benötigt.

10.4.1. **Satz** (Multiplikatorregel von Lagrange). *Sei $U \subset \mathbb{R}^n$ offen, $f : U \longrightarrow \mathbb{R}$ differenzierbar, außerdem $\varphi \in \mathcal{C}^1(U,\mathbb{R}^m)$, $M := \varphi^{-1}(0) \subset U$. Sei $a \in M$ ein lokales Extremum von $f|_M$. Dann gilt: Falls $\operatorname{grad}\varphi_1(a),\dots,\operatorname{grad}\varphi_m(a)$ linear unabhängig sind, so folgt*

$$\operatorname{grad} f(a) \in span\{\operatorname{grad}\varphi_1(a),\dots,\operatorname{grad}\varphi_m(a)\},$$

d.h. $\exists\lambda_1^0,\dots,\lambda_m^0 \in \mathbb{R}$ („Lagrange–Multiplikatoren“) mit $\operatorname{grad} f(a) + \sum_{i=1}^m \lambda_i^0 \operatorname{grad}\varphi_i(a) = 0$.

Die Konklusion kann so umformuliert werden: Die Lagrange-Funktion $L: U\times\mathbb{R}^m\to\mathbb{R}$, $L(x,\lambda)=f(x)+\sum_{j=1}^m\lambda_j\varphi_j(x)$ *hat einen kritischen Punkt in* (a,λ^0), *d.h. es gilt* $\operatorname{grad}L(a,\lambda^0)=0$.

Beweis: Angenommen, die Vektoren $\{\operatorname{grad}f(a),\operatorname{grad}\varphi_1(a),\dots,\operatorname{grad}\varphi_m(a)\}$ wären linear unabhängig; insbesondere $m+1\leqslant n$. Wir betrachten dann in den $n+1$ Variablen $(x_1,\dots,x_n,y)=(x,y)$ das Gleichungssystem $f(x_1,\dots,x_n)-y=0$, $\varphi_1(x_1,\dots,x_n)=0,\cdots,\varphi_m(x_1,\dots,x_n)=0$ nahe $(a,f(a))$. Seine Jacobi-Matrix nach $x=(x_1,\dots,x_n)$ hat laut Annahme in $x=a$ den Zeilenrang $m+1$, also auch den Spaltenrang $m+1$. Es kann somit nahe $(a,f(a))\in\mathbb{R}^{n+1}$ nach geeigneten $m+1$ der Variablen $x_1,\dots,x_n$ aufgelöst werden. Daraus folgt, dass in jeder Umgebung von a sowohl $y>f(a)$ wie auch $y<f(a)$ möglich ist (unter Wahrung der Nebenbedingungen). Widerspruch!

Rezept. Zur Bestimmung der Kandidaten für Extremstellen von $f:U\longrightarrow\mathbb{R}$ unter den Nebenbedingungen $\varphi=0$ mit $\varphi:U\longrightarrow\mathbb{R}^m$, wobei $\operatorname{Rang}J_\varphi(x)=m$ für alle x mit $\varphi(x)=0$ sei, bildet man die Lagrange-Funktion $L:U\times\mathbb{R}^m\longrightarrow\mathbb{R}$, $L(x,\lambda)=f(x)+\sum_{j=1}^m\lambda_j\varphi_j(x)$ und sucht die Lösungen des Systems

$$\begin{aligned}&\operatorname{grad}L(x,\lambda)=0,\quad \text{d.h.}\\ &\frac{\partial L}{\partial x_i}(x,\lambda)=\frac{\partial f}{\partial x_i}+\sum_{j=1}^m\lambda_j\frac{\partial\varphi_j}{\partial x_i}=0\quad\text{für }1\leqslant i\leqslant n\quad\text{und}\\ &\frac{\partial L}{\partial\lambda_j}(x,\lambda)=\varphi_j(x)=0\quad\text{für }1\leqslant j\leqslant m\,.\end{aligned}$$

Das sind $n+m$ Gleichungen für die $n+m$ Variablen $x_1,\dots,x_n,\lambda_1,\dots,\lambda_m$. Nach dem Lösen kann man die λ_j wieder vergessen.

10.4.2. **Beispiel.** Sei $U:=\{(x_1,\dots,x_n)\in\mathbb{R}^n: x_j>0\,\forall j=1,\dots,n\}$ und $f:U\to\mathbb{R}$, $f(x)=x_1\cdots x_n$.

Wir suchen Extrema von f unter der Nebenbedingung $x_1+\ldots+x_n=1$, d.h. Extrema von $f|_{\varphi^{-1}(0)}$ mit $\varphi:U\to\mathbb{R}$, $\varphi(x)=x_1+\ldots+x_n-1$. Es ist $\operatorname{grad}\varphi(x)=(1,1,\dots,1)\neq0$ für jedes $x\in\varphi^{-1}(0)$, also ($m=1$) ist die Voraussetzung an die lineare Unabhängigkeit der Gradienten der φ_i erfüllt. Also: Ist $x\in U$ lokales Extremum von $f|_{\varphi^{-1}(0)}$, so gibt es $\lambda\in\mathbb{R}$ mit $\operatorname{grad}f(x)=\lambda\operatorname{grad}\varphi(x)=(\lambda,\dots,\lambda)$. Wegen $\operatorname{grad}f(x)=(x_2\cdots x_n,x_1x_3\cdots x_n,\dots,x_1\cdots x_{n-1})$ folgt $\frac{c}{x_1}=\lambda$, $\frac{c}{x_2}=\lambda,\dots,\frac{c}{x_n}=\lambda$, wobei $c=x_1\cdots x_n>0$. Daraus folgt $x_1=\ldots=x_n$, d.h. $x=(\frac1n,\dots,\frac1n)$; Wert von f dort: $\frac{1}{n^n}>0$.

Behauptung: $f|_{\varphi^{-1}(0)}$ hat in $(\frac1n,\dots,\frac1n)$ tatsächlich ein Extremum, und zwar ein globales Maximum.

Beweis:

$$\overline{\varphi^{-1}(0)}=\{x\in\mathbb{R}^n: x_1+\ldots+x_n=1, x_j\geqslant0\,\forall j=1,\dots,n\}$$

ist kompakt, also nimmt $f:x\mapsto x_1\cdots x_n$ auf $\overline{\varphi^{-1}(0)}$ ein Maximum an. In $\overline{\varphi^{-1}(0)}\smallsetminus U$ liegt es nicht (dort ist $f=0$), also liegt es in $\varphi^{-1}(0)$. Daraus folgt: Der oben gefundene Punkt ist notwendigerweise das Maximum. □

Anwendung: AGM-Ungleichung.
Sind $a_1,\dots,a_n>0$, so ist $b:=\left(\frac{a_1}{\sum_{j=1}^n a_j},\dots,\frac{a_n}{\sum_{j=1}^n a_j}\right)\in\varphi^{-1}(0)$, also $f(b)\leqslant f(\frac1n,\dots,\frac1n)=\frac{1}{n^n}$ und $f(b)=a_1\cdot\ldots\cdot a_n/(\sum_{j=1}^n a_j)^n$. Durch Wurzelziehen folgt:

$$\sqrt[n]{a_1\cdot\ldots\cdot a_n}\leqslant\frac{a_1+\ldots+a_n}{n}\,.$$

Die Gleichheit gilt genau dann, wenn $b=(\frac1n,\dots,\frac1n)$, also $a_1=\ldots=a_n$.

10.5. Übungen.

10.5.1. ***Aufgabe***. (Beispiele von Diffeomorphismen) Sei $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf $\mathbb{R}^n$.

(a) Zeigen Sie, daß die Abbildung $f:B_1(0)\longrightarrow\mathbb{R}^n$, $f(x):=\frac{x}{\sqrt{1-\langle x,x\rangle}}$ ein Diffeomorphismus ist, und berechnen Sie ihre Ableitung.

(b) Konstruieren Sie einen Diffeomorphismus der Kugel $B_1(0)$ auf den Würfel $(-1,1)^n$ in $\mathbb{R}^n$.

10.5.2. ***Aufgabe***. Sei $U=\{(r,\theta,\varphi)\in\mathbb{R}^3\mid r>0\}$ und

$$P:U\longrightarrow\mathbb{R}^3\quad,\quad P(r,\theta,\varphi)=(r\cos\theta\cos\varphi,r\sin\theta\cos\varphi,r\sin\varphi)\,.$$

(a) Zeige, daß P stetig differenzierbar ist, und bestimme die Menge der Punkte, in denen P lokal invertierbar ist.

(b) Bestimme die Ableitung einer lokalen Umkehrabbildung von P.

(c) Zeige, daß P auf $U_0=(0,\infty)\times\left(-\frac{\pi}{2},\frac{\pi}{2}\right)\times\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ invertierbar ist, und bestimme die Umkehrabbildung. Bilde die Ableitung der Umkehrabbildung und verifiziere das Ergebnis aus (b).

10.5.3. ***Aufgabe***. Sei $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$, so dass $\varphi = f - \mathrm{Id}_{\mathbb{R}^n}$ eine Kontraktion ist. Zeigen Sie, dass f ein $\mathcal{C}^1$-Diffeomorphismus ist. (Tipp: Diffeomorphiesatz; $d\varphi(x)$ ist eine Kontraktion also $df(x)$ ist injektiv; für f bijektiv: reduziere $f(x) = y$ zu einem Fixpunkt-Problem und wende den Banachschen Fixpunktsatz an.)

10.5.4. ***Aufgabe***. Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$,

$$f(x) = \begin{cases} x + 2x^2 \sin \frac{1}{x} & \text{für } x \neq 0, \\ 0 & \text{für } x = 0 \end{cases}$$

Zeigen Sie, daß f auf $\mathbb{R}$ differenzierbar ist mit $f'(0) = 1$, daß aber f nicht injektiv ist auf jedem offenen Intervall, das die Null enthält. Wiederlegt das den Satz über inverse Funktionen? Warum?

10.5.5. ***Aufgabe***. Seien $f_1 : U_1 \longrightarrow V_1$, $f_2 : U_2 \longrightarrow V_2$ zwei differenzierbare Funktionen, wobei $U_1, U_2 \subset \mathbb{R}^m$, $V_1, V_2 \subset \mathbb{R}^n$ offene Teilmengen sind. Wir definieren die Relation $\sim$ durch $f_1 \sim f_2$ genau dann, wenn es Diffeomorphismen $\varphi : U_1 \longrightarrow U_2$, $\psi : V_1 \longrightarrow V_2$ existieren so, daß $\psi f_1 \varphi^{-1} = f_2$. Zeigen Sie, daß $\sim$ eine Äquivalenzrelation ist.

10.5.6. ***Aufgabe***. (Das lokale Verhalten Funktionen einer Variablen)
Sei f eine reelle differenzierbare Funktion definiert in einer Umgebung von $a \in \mathbb{R}$, so daß $f^{(l)}(a) = 0$ für $l < k \in \mathbb{N}$.

(a) Zeigen Sie, daß es eine differenzierbare Funktion g in einer Umgebung von a existiert derart, daß $f(x) = (x-a)^k g(x)$ und $g(a) = \dfrac{1}{k!} f^{(k)}(a)$. (Anleitung: Benutzen Sie die Taylorformel mit Integralrestglied und die Leibnizregel)

(b) Ist $f^{(k)}(a) \neq 0$, so gibt es lokal um a eine Koordinatentransformation φ mit $\varphi(a) = 0$, $\varphi'(a) = 1$ und $f \circ \varphi^{-1}(y) = \dfrac{1}{k!} f^{(k)}(0) y^k$.

10.5.7. ***Aufgabe***. Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x,y) = xye^{-x-y}$.

(a) In der Umgebung welcher Punkte $(x_0, y_0) \in \mathbb{R}^2$ läßt sich die Bedingung $f(x,y) = f(x_0,y_0)$ gemäß dem Satz über implizite Funktionen durch eine C^1–Funktion $g : x \mapsto y(x)$ bzw. $\tilde{g} : y \mapsto x(y)$ auflösen?

(b) Berechne jeweils $dg(x_0)$ bzw. $d\tilde{g}(y_0)$.

Zeichne eine qualitative Skizze der Höhenlinien von f. Die Ergebnisse von (a) und (b) sind dabei hilfreich!

10.5.8. ***Aufgabe***. Bestimme Max. und Min. von $f : \mathbb{R}^3 \ni (x,y,z) \mapsto xy - z^4 - 2(x^2 + y^2 - z^2) \in \mathbb{R}$ auf dem Vollellipsoid $\{(x,y,z) \mid x^2 + y^2 + 2z^2 \leq 8\}$. (Tip: Diskutiere Inneres und Rand getrennt; benutze für den Rand die Methode der Lagrange-Multiplikatoren.)

10.5.9. ***Aufgabe***. Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch $f(x,y) := (2e^x - x^2, e^y + (x^2+1)y)$. Zeigen Sie:

(a) f ist bijektiv.

(b) f ist ein C^∞-Diffeomorphismus. Was ist $d(f^{-1})(2,1)$?

10.5.10. ***Aufgabe***. (a) Zeigen Sie, dass das Gleichungssystem

$$x^2 + uy + e^v = 0$$
$$2x + u^2 + uv = 5$$

für $x, y, u, v \in \mathbb{R}$ in einer Umgebung von $(x_0, y_0, u_0, v_0) := (2, 5, -1, 0)$ durch eine C^1-Abbildung $g : (x,y) \mapsto (u(x,y), v(x,y))$ aufgelöst werden kann.

(b) Berechnen Sie die Jacobi-Matrix $J_g(2,5)$.

10.5.11. ***Aufgabe***. Sei $V := M_{n \times n}(\mathbb{R}) \cong \mathbb{R}^{n^2}$. Es besitze $A_0 \in V$ einen reellen Eigenwert λ_0 mit algebraischer Vielfachheit 1 (d.h. λ_0 sei eine einfache reelle Nullstelle des charakteristischen Polynoms von A). Zeigen Sie: Es gibt eine Umgebung U von A_0, ein offenes Intervall I mit $\lambda_0 \in I$ und eine C^1-Funktion $\Lambda : U \to I$, so dass für alle $A \in U$, $\lambda \in I$ gilt: λ ist Eigenwert von A genau dann, wenn $\lambda = \Lambda(A)$ ist.

11. Untermannigfaltigkeiten des $\mathbb{R}^N$

Bisher haben wir differenzierbare Abbildungen behandelt, die auf offenen Teilmengen des $\mathbb{R}^N$ definiert sind. Für viele Bereiche der Mathematik und für viele Anwendungen z.B. in der mathematischen Physik reicht das nicht aus. In der Strömungsmechanik spielt der Fluss einer Strömung durch eine Kurve in $\mathbb{R}^2$ (bzw. eine Fläche in $\mathbb{R}^3$) eine bedeutende Rolle; der Nettofluss ist mathematisch ein Integral über die Kurve (bzw. Fläche) und lässt sich nach dem Satz von Gauß als ein Integral über einem Bereich in $\mathbb{R}^2$ (bzw. $\mathbb{R}^3$) ausdrücken.

Wie können wir eine Funktion auf einer gekrümmten Kurve oder Fläche zu einer „normalen" Funktion, definiert auf einer offenen Menge des euklidischen Raumes, reduzieren? Die Idee ist uns aus der Linearen Algebra bekannt und heißt „das Objekt parametrisieren": Wir wissen, wie man die Lösungsmenge eines linearen Systems parametrisieren kann und dass dadurch die Eigenschaften dieser Menge bekannt sind.

Die einfachsten Beispiele von Kurven oder gekrümmten Flächen sind Graphen:

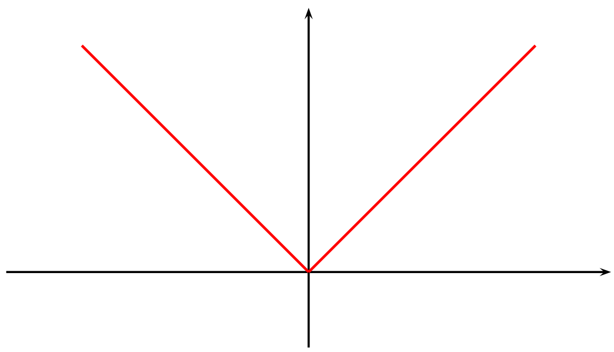
$$C := \{(t, f(t)) \in \mathbb{R}^2 : t \in I\}, \quad I \text{ Intervall in } \mathbb{R},\ f : I \to \mathbb{R} \text{ (Kurve)}$$
$$F := \{(u, v, f(u,v)) \in \mathbb{R}^3 : (u,v) \in D\}, \quad D \text{ offen in } \mathbb{R}^2,\ f : D \to \mathbb{R} \text{ (Fläche)}$$

Die Kurve C ist parametrisiert durch $t \in I$ und die Fläche F ist parametrisiert durch $(u,v) \in D$. Expliziter: Die Abbildung $\psi : D \to F$, $(u,v) \to (u,v,f(u,v))$ ist ein Homöomorphismus (die Inverse ist die Projektion $(u,v,z) \mapsto (u,v)$ eingeschränkt auf F). Eine beliebige Funktion $g : F \to \mathbb{R}$ ist nun eindeutig bestimmt durch die Funktion $\tilde{g} : D \to \mathbb{R}$, $\tilde{g}(u,v) := g(\psi(u,v)) = g(u,v,f(u,v))$; dabei ist $\tilde{g}$ definiert auf der offenen Menge $D \subset \mathbb{R}^2$!

Andere einfache Beispiele zeigen uns aber, dass sich nicht alle Flächen so parametrisieren lassen. Die Sphäre in $\mathbb{R}^3$ z.B. ist nicht homöomorph zu einer offenen Menge des $\mathbb{R}^2$. (Die Sphäre ist kompakt, eine offene Menge nicht!)

Mengen wie Kreise, Ellipsen, Sphären, Tori sind also *global* nicht durch n (in diesem Fall $n = 1$ bzw. $n = 2$) reelle Koordinaten zu beschreiben, sondern nur *lokal* in der Nähe jedes Punktes (für die Sphäre sieht man das leicht, indem man kleine Stücke auf geeignete Koordinatenebenen projiziert).

Solche Mengen nennen wir Untermannigfaltigkeiten, falls noch zusätzliche „gute" Eigenschaften erfüllt sind. Damit ist die Tatsache gemeint, dass diese Objekte „glatt" oder „ohne Ecken" sind, d.h. lokal durch lineare (affine) Räume approximierbar (deshalb werden sie in Analogie zu den Funktionen auch „differenzierbare Untermannigfaltigkeiten" genannt). Die Menge $\{(x,y) : y = |x|\}$ hat eine Ecke in $(0,0)$; die Parametrisierung $\mathbb{R} \ni t \mapsto (t, |t|)$ ist zwar ein Homöomorphismus, ist aber nicht differenzierbar.

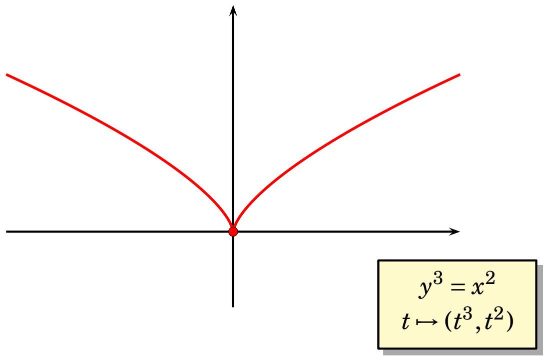


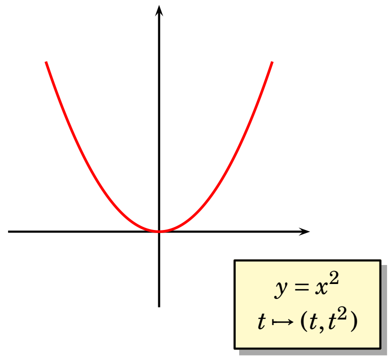
Betrachten wir also den Graph $G = \{(t, f(t)) : t \in I\}$ einer *differenzierbaren* Funktion $f : I \subset \mathbb{R} \to \mathbb{R}$. Wir wissen, dass G eine Tangente $T_{(t,f(t))}G$ in jedem Punkt $(t, f(t))$ besitzt, nämlich der Graph der affinen Abbildung $u \mapsto f(t) + f'(t)(u - t)$. Nun besitzt G die Parametrisierung $\psi : I \to \mathbb{R}^2$, $t \mapsto \psi(t) = (t, f(t))$ und die Tangente $T_{(t,f(t))}G$ die Parametrisierung $v \mapsto (t, f(t)) + (v, f'(t)v) = (t, f(t)) + \psi'(t)v$.

Dies bedeutet, dass die Ableitung der Parametrisierung ψ eine Parametrisierung der Tangente liefert. Wir werden diese Eigenschaft für alle Parametrisierungen verlangen. Insbesondere muss dann $\psi'(t) \neq 0$ (äquivalent dazu: $d\psi(t) \in \mathcal{L}(\mathbb{R}, \mathbb{R}^2)$ *injektiv*) sein, damit die Tangente hier ein affiner Raum der Dimension 1 wird.

Wir geben noch zwei konkrete Beispiele. Die Abbildung $\psi : t \mapsto (t, t^2)$ parametrisiert die (glatte) Parabel $y = x^2$; und $\psi'(t) = (1, 2t) \neq (0,0)$ liefert eine Parametrisierung der Tangente an die Parabel in (t, t^2).

Die Abbildung $\psi : t \mapsto (t^3, t^2)$ ist differenzierbar und parametrisiert die ***Neilsche Parabel*** $y^3 = x^2$; diese Menge hat eine Ecke in $(0,0)$, und tatsächlich gilt dort $\psi'(0) = (0,0)$.





11.1. **Immersionen, Einbettungen und Submersionen.** Die obige Diskussion motiviert das Studium von differenzierbaren Abbildungen, deren Differential injektiv ist.

Eine $\mathcal{C}^\infty$-Abbildung nennen wir ***glatt***. Einen $\mathcal{C}^\infty$-Diffeomorphismus nennen wir einfach einen Diffeomorphismus. In der Regel werden wir uns auf glatte Abbildungen beschränken. Sind alle Abbildungen in den folgenden Definitionen von der Klasse $\mathcal{C}^k$, wobei $1 \leqslant k \leqslant \infty$, so erhalten wir „$\mathcal{C}^k$-Immersionen", „$\mathcal{C}^k$-Submersionen", Untermannigfaltigkeiten „von der Klasse $\mathcal{C}^k$" usw.

11.1.1. **Definition.** Sei $D \subset \mathbb{R}^n$ offen. Eine glatte Abbildung $f : D \to \mathbb{R}^N$ heißt eine textbf*Immersion* in $a \in D$, wenn das Differential $df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^N)$ injektiv ist (d.h. die Jacobi-Matrix $J_f(a)$ hat Rang n). Ist f Immersion in allen Punkten von D, so heißt f eine *Immersion*.

Wenn $f : D \subset \mathbb{R}^n \to \mathbb{R}^N$ eine Immersion in $a \in D$ ist, muss notwendig $n \leqslant N$ gelten.

Immersionen sind unsere Kandidaten für Parametrisierungen einer Untermannigfaltigkeit. Trotzdem kann das Bild einer (sogar injektiven) Immersion „Spitzen" haben, siehe Beispiel 11.1.3 (iv) unten. Deshalb führen wir den folgenden Begriff ein.

11.1.2. **Definition.** Eine glatte Abbildung $f : D \to \mathbb{R}^N$ heißt ***Einbettung***, wenn gilt:

(i) f ist eine Immersion.
(ii) f ist injektiv.
(iii) $f : D \to f(D)$ ist ein Homöomorphismus.

Die Begingung (iii) sichert, dass $f(D)$ „glatt" ist. Da f schon stetig ist, reicht es für die Verifikation von (iii) zu zeigen, dass $f^{-1} : f(D) \to D$ stetig ist.

11.1.3. **Beispiel.**

(i) $f : \mathbb{R} \to \mathbb{R}^2$, $f(t) = (t, t^2)$ ist eine Immersion, deren Bild die Parabel $y = x^2$ ist.

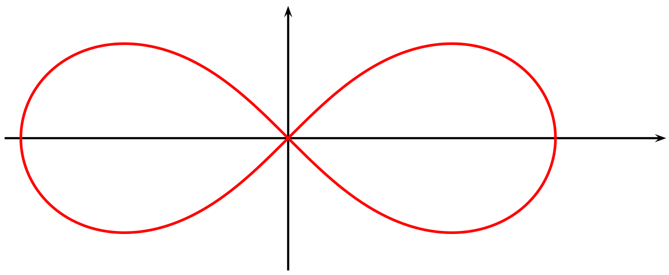
(ii) $f : \mathbb{R} \to \mathbb{R}^2$, $f(t) = (t^3, t^2)$ ist keine Immersion in $(0,0)$. Das Bild ist die Neilsche Parabel $\{(x, y) \in \mathbb{R}^2 : y^2 = x^3\}$.

(iii) $f : \mathbb{R} \to \mathbb{R}^2$, $f(t) = e^{i\varphi} = (\cos\varphi, \sin\varphi)$ ist eine Immersion, deren Bild der Einheitskreis ist. Sie ist nicht injektiv (sondern periodisch mit Periode 2π), aber für jedes offenes Intervall $I = (\varphi_0 - \pi, \varphi_0 + \pi)$ ist $\psi := f|_I : I \to \mathbb{R}^2$ injektiv und eine Einbettung. Das Bild von ψ ist $\operatorname{Im}(\psi) = f((\varphi_0 - \pi, \varphi_0 + \pi)) = S^1 \setminus \{e^{i(\varphi_0+\pi)}\}$. Die Umkehrfunktion ψ^{-1} ist ein Zweig der Argument-Abbildung (siehe Satz 8.4.9 und Definition 8.4.10 (d)). Wenn $\arg : S^1 \setminus \{1\} \to (0, 2\pi)$ den Hauptzweig bezeichnet, dann ist $\psi^{-1} : S^1 \setminus \{e^{i(\varphi_0+\pi)}\} \to (\varphi_0 - \pi, \varphi_0 + \pi)$ durch $z \mapsto \arg(ze^{-i(\varphi_0-\pi)}) + \varphi_0 - \pi$ gegeben.

(iv) $f : (-\pi/4, \pi/4) \cup (3\pi/4, 5\pi/4) \to \mathbb{R}^2$, $f(t) = (\cos t\sqrt{2|\cos(2t)|}, \sin t\sqrt{2|\cos(2t)|})$ ist eine Einbettung, deren Bild die Lemniskate von Bernoulli (Beispiel 10.3.3 (3)) ohne $(0,0)$ ist. Die kartesische Gleichung der Lemniskate ist

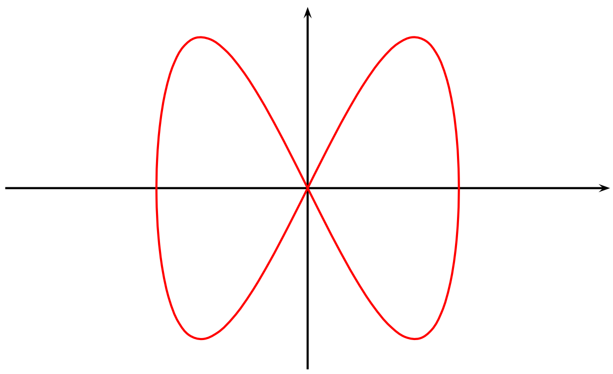
$$(x^2 + y^2)^2 - 2(x^2 - y^2) = 0. \tag{11.1}$$

Die parametrische Darstellung ergibt sich, wenn wir $(x,y) = (\rho\cos t, \rho\sin t)$ in (11.1) einsetzen und die Gleichung der Lemniskate in Polarkoordinaten erhalten, nämlich $\rho^2 = 2\cos(2t)$. $f|_{(-\pi/4,\pi/4)}$ parametrisiert die rechte Seite und $f|_{(3\pi/4,5\pi/4)}$ parametrisiert die linke Seite der Lemniskate.



Es ist klar, dass sich die Inverse f^{-1} im Punkte $(0,0)$ nicht stetig fortsetzen lässt.

(v) $f : (\frac{\pi}{2}, \frac{5\pi}{2}) \to \mathbb{R}^2$, $f(t) = (\cos t, \sin 2t)$ ist eine injektive Immersion, aber keine Einbettung; $\mathrm{Im}(f) = \{(x,y) \in \mathbb{R}^2 : y^2 - 4x^2 + 4x^4 = 0\}$ ist keine Untermannigfaltigkeit des $\mathbb{R}^2$.



(vi) Sei $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Dann ist $f : \mathbb{R} \to \mathbb{R}^4$, $t \mapsto (e^{it}, e^{i\alpha t})$ eine injektive Immersion, für die $f(\mathbb{R})$ dicht in $S^1 \times S^1$ ist.

(vii) Sphärische Koordinaten: Sei $r > 0$ fest. Dann ist

$$f : (0,2\pi) \times (-\pi/2, \pi/2) \ni (\phi, \theta) \mapsto (r\cos\phi\cos\theta, r\sin\phi\cos\theta, r\sin\theta) \in \mathbb{R}^3$$

eine Einbettung, deren Bild $S_r^2 \setminus \{(x,0,z) : x \geqslant 0\}$ ist, wobei S_r^2 die 2-Sphäre vom Radius r in $\mathbb{R}^3$ ist.

(viii) Ein Diffeomorphismus ist eine Einbettung.

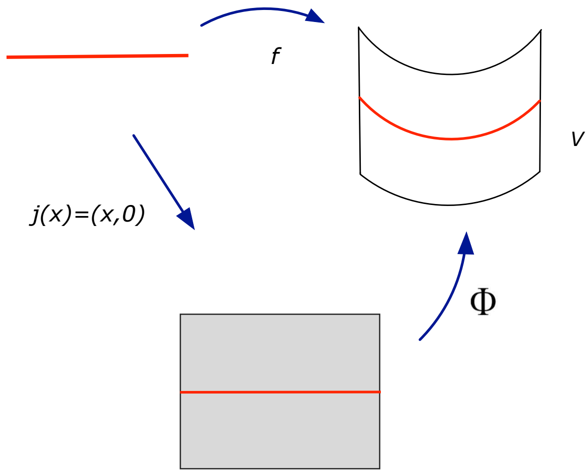
(ix) Für $n < N$, $x \in \mathbb{R}^n$ definieren wir $(x,0) \in \mathbb{R}^N$, wobei 0 der Nullvektor in $\mathbb{R}^{N-n}$ ist. Falls $n = N$ definieren wir $(x,0) \in \mathbb{R}^N = \mathbb{R}^n$ durch $(x,0) = x$. Sei $j : \mathbb{R}^n \longrightarrow \mathbb{R}^N$, $j(x) = (x,0)$. Die Abbildung j ist injektiv und linear. Daher $dj(x) = j$ für alle $x \in \mathbb{R}^n$, also ist j eine Immersion.

Wir beweisen nun, dass jede Immersion lokal (d.h. jeweils in einer kleinen Umgebung eines beliebigen gegebenen Punktes) bis auf einen Diffeomorphismus die Form (ix) hat.

Notation. Sei $W_\varepsilon^m(p) = \prod_{i=1}^m (p_i - \varepsilon, p_i + \varepsilon) \subset \mathbb{R}^m$ der Würfel mit Mittelpunkt $p \in \mathbb{R}^m$ und Kantenlänge 2ε.

11.1.4. **Satz** (lokale Struktur einer Immersion). *Sei $D \subset \mathbb{R}^n$ offen und $f : D \longrightarrow \mathbb{R}^N$ eine Immersion in $a \in D$. Dann gibt es $\varepsilon > 0$ mit $W_\varepsilon^n(a) \subset D$, eine offene Umgebung V von $f(a)$ und einen Diffeomorphismus $\Phi : W_\varepsilon^N(a,0) \longrightarrow V$ so, dass $f(W_\varepsilon^n(a)) \subset V$ und $f|_{W_\varepsilon^n(a)} = \Phi \circ j$, d.h. das folgende Diagramm ist kommutativ:*

$$\begin{tikzcd}
W_\varepsilon^n(a) \arrow[rr, "f"] \arrow[dr, "j"'] & & V \\
& W_\varepsilon^N(a,0) \arrow[ur, "\Phi"'] &
\end{tikzcd}$$



Beweis: Falls $n = N$, so ist f ein lokaler Diffeomorphismus in a (Umkehrsatz), und wir wählen $\varepsilon > 0$ so, dass $W^n_\varepsilon(a) \subset D$ und $f : W^n_\varepsilon(a) \longrightarrow V$ ein Diffeomorphismus ist. Da $j = \mathrm{Id}_{\mathbb{R}^n}$ ist, folgt die Behauptung mit $\Phi = f$.

Sei $n < N$. Wir betrachten die Abbildung $F : D \times \mathbb{R}^{N-n} \longrightarrow \mathbb{R}^N$, $F(x,z) = f(x)+T\cdot z$, wobei $T \in M_{N\times(N-n)}(\mathbb{R})$, $z \in \mathbb{R}^{N-n}$. Wir suchen T so, dass F ein lokaler Diffeomorphismus in a wird. Die Jacobi-Matrix von F in $(a,0)$ ist:

$$J_F(a,0) = \Big(J_f(a),\, T\Big) \in M_{N\times N}(\mathbb{R}).$$

Wegen $\operatorname{Rang} J_f(a) = n$ sind die Spalten $v_1,\dots,v_n$ von $J_f(a)$ linear unabhängig in $\mathbb{R}^N$. Wir wählen T so, dass die Spalten $v_{n+1},\dots,v_N$ von T eine Basisergänzung von $\{v_1,\dots,v_n\}$ zu einer Basis $\{v_1,\dots,v_n,v_{n+1},\dots,v_N\}$ von $\mathbb{R}^N$ sind. Dann ist $\operatorname{Rang} J_F(a) = N$ und $J_F(a)$ invertierbar. Nach dem Umkehrsatz gibt es $\varepsilon > 0$, so dass gilt: $W^n_\varepsilon(a) \subset D$, $F(W^N_\varepsilon(a,0)) =: V$ ist eine offene Umgebung von $f(a) = F(a,0)$ und $F : W^N_\varepsilon(a,0) \longrightarrow V$ ist ein Diffeomorphismus. Offensichtlich gilt $f(x) = F(x,0) = (F\circ j)(x)$ für $x \in W^n_\varepsilon(a)$. Die Behauptung folgt also mit $\Phi = F|_{W^N_\varepsilon(a,0)}$. □

11.1.5. **Definition.** Sei $D \subset \mathbb{R}^N$ offen. Eine glatte Abbildung $f : D \to \mathbb{R}^m$ heißt eine ***Submersion*** in $a \in D$, wenn das Differential $df(a) \in \mathcal{L}(\mathbb{R}^N,\mathbb{R}^m)$ surjektiv ist (d.h. wenn die Jacobi-Matrix $J_f(a)$ Rang m hat, d.h. wenn $\{\operatorname{grad} f_1(a),\dots,\operatorname{grad} f_m(a)\}$ eine linear unabhängige Menge ist). Wir sagen auch, dass die Komponenten $f_1,\dots,f_m$ von f unabhängige Funktionen in $a \in D$ sind. Ist f Submersion in allen Punkten von D, so heißt f eine *Submersion*. Ein Punkt $x \in D$ heißt ***regulärer Punkt*** von f, wenn f eine Submersion in x ist. Andernfalls heißt x ***kritischer Punkt***. Ein Punkt $c \in \mathbb{R}^m$ heißt ***regulärer Wert*** der Abbildung f, falls f eine Submersion in allen Punkten der Niveaumenge $f^{-1}(c)$ ist. Ansonsten heißt c ***kritischer Wert*** der Abbildung f.

Wenn $f : D \subset \mathbb{R}^N \to \mathbb{R}^m$ eine Submersion in $a \in D$ ist, muss also $N \geqslant m$ gelten. Im Falle $N < m$ sind alle Punkte $x \in D$ kritisch, und jedes $c \in \operatorname{Im}(f)$ ist ein kritischer Wert.
Generell ist jedes $y \in \mathbb{R}^m \setminus \operatorname{Im}(f)$ ein regulärer Wert, da in diesem Fall $f^{-1}(c) = \emptyset$ gilt.

11.1.6. **Beispiel.**

(i) $f : \mathbb{R}^{n+1}\setminus\{0\} \to \mathbb{R}$, $f(x) = x_1^2 + \ldots + x_{n+1}^2$ ist eine Submersion.
(ii) Sei $P : \mathbb{R}^N \to \mathbb{R}$ ein homogenes Polynom vom Grad $r \geqslant 1$. Dann ist jedes $c \in \mathbb{R}\setminus\{0\}$ ein regulärer Wert. Durch Ableiten der Identität $P(tx) = t^r P(x)$ nach t folgt $dP(tx)\cdot x = rt^{r-1}P(x)$ (Kettenregel); für $t = 1$ erhalten wir die ***Eulersche Identität***:

$$dP(x)\cdot x = rP(x) \text{ für alle } x \in \mathbb{R}^N.$$

Für $c \in \mathbb{R}\setminus\{0\}$ und jedes $x \in P^{-1}(c)$ folgt $dP(x) \neq 0$.
(iii) Falls $N \geqslant m$, so ist die Projektion $\pi : \mathbb{R}^N \to \mathbb{R}^m$, $\pi(x_1,\dots,x_m,x_{m+1},\dots,x_N) = (x_1,\dots,x_m)$ eine lineare surjektive Abbildung. Deshalb gilt $d\pi(x) = \pi$ für alle $x \in \mathbb{R}^N$ und π ist eine Submersion.

Wir beweisen nun, dass jede Submersion lokal bis auf einen Diffeomorphismus die Form (iii) hat.

11.1.7. **Satz** (lokale Struktur einer Submersion)**.** *Sei $D \subset \mathbb{R}^N$ offen und $f : D \longrightarrow \mathbb{R}^m$ eine Submersion in $a \in D$. Dann gibt es $\varepsilon > 0$, eine offene Umgebung V von a und einen Diffeomorphismus $\Phi : V \longrightarrow$*

$W^N_\varepsilon(f(a),0)$ so, dass $f(V) = W^m_\varepsilon(f(a))$ und $f|_V = \pi\circ\Phi$, d.h. das folgende Diagramm ist kommutativ:

$$\begin{tikzcd} V \arrow[rr, "f"] \arrow[dr, "\Phi"'] & & W^m_\varepsilon(f(a)) \\ & W^N_\varepsilon(f(a),0) \arrow[ur, "\pi"'] & \end{tikzcd}$$

Beweis: Falls $m = N$, so ist f ein lokaler Diffeomorphismus in a (Umkehrsatz), und wir wählen $\varepsilon > 0$, so dass $W^m_\varepsilon(a) \subset D$ und $f : W^m_\varepsilon(a) \longrightarrow V$ ein Diffeomorphismus ist. Da $j = \mathrm{Id}_{\mathbb{R}^m}$ ist, folgt die Behauptung mit $\Phi = f$.

Sei $m < N$. Wir betrachten die Abbildung $F : D \longrightarrow \mathbb{R}^m \times \mathbb{R}^{N-m}$, $F(x) = (f(x), T\cdot(x-a))$, wobei $T \in M_{m\times N}(\mathbb{R})$. Dann gilt $F(a) = (f(a),0)$. Wir suchen T so, dass F ein lokaler Diffeomorphismus in a wird. Die Jacobi-Matrix von F in a ist:

$$J_F(a) = \begin{pmatrix} J_f(a) \\ T \end{pmatrix} \in M_{N\times N}(\mathbb{R}).$$

Wegen $\operatorname{Rang} J_f(a) = m$ sind die Zeilen $v_1,\dots,v_m$ von $J_f(a)$ linear unabhängig in $\mathbb{R}^N$. Wir wählen T so, dass die Zeilen $v_{m+1},\dots,v_N$ von T eine Basisergänzung von $\{v_1,\dots,v_m\}$ zu einer Basis $\{v_1,\dots,v_m,v_{m+1},\dots,v_N\}$ von $\mathbb{R}^N$ sind. Dann ist $\operatorname{Rang} J_F(a) = N$ und $J_F(a)$ invertierbar. Nach dem Umkehrsatz gibt es $\varepsilon > 0$ und eine offene Umgebung V von a, so dass gilt: $F(V) = W^N_\varepsilon(f(a),0)$ und $\Phi := F|_V : V \longrightarrow W^N_\varepsilon(f(a),0)$ ist ein Diffeomorphismus. Offensichtlich gilt $f(x) = \pi\circ F(x) = \pi\circ\Phi(x)$ für alle $x \in V$. □

Wiederholung aus linearen Algebra: der Rang.

11.1.8. **Definition.** Sei K ein Körper, V, W endlich-dimensionalen K-Vektorrüme, $T : V \to W$ eine lineare Abbildung.

(i) Der **Rang** von T, bezeichnet $\operatorname{Rang} T$, ist definiert als Dimension des Bildes von T.
(ii) Ist $A \in M_{m\times n}(K)$, so definiert A eine lineare Abbildung $T_A : K^n \to K^m$. Wie setzen $\operatorname{Rang} A := \operatorname{Rang} T_A$.
(iii) Der **Spaltenrang** von A ist die Maximalzahl linear unabhängiger Spalten von A. Der **Zeilenrang** von A ist die Maximalzahl linear unabhängiger Zeilen von A.

11.1.9. **Proposition.** *Sei $A \in M_{m\times n}(K)$. Dann gilt:*

(i) $\operatorname{Rang} A$ = Zeilenrang A = Spaltenrang A.
(ii) *Der Rang von A ist das Maximum der Ränge aller quadratischen Untermatrizen von A.*
(iii) $\operatorname{Rang} A = r$ *bedeutet also:*
   - *Es gibt eine invertierbare $r\times r$-Untermatrix*
   - *Es gibt keine invertierbare $s\times s$-Untermatrix mit $s > r$*
(iv) *Ist A in Zeilenstufenform, so ist* $\operatorname{Rang} A$ *die Anzahl der Zeilenvektoren die ungleich Null sind.*
(v) *Ist $m = n$, so ist A genau dann invertierbar, wenn sie Maximalrang hat, d. h.* $\operatorname{Rang} A = n$.

11.2. **Definition und Charakterisierung der Untermannigfaltigkeiten.** Alle Teilmengen reeller Räume versehen wir mit der durch die euklidische Metrik induzierten Topologie. Ist $M \subset \mathbb{R}^N$, so ist $U \subset M$ offen in M genau dann, wenn es eine offene Menge $\tilde U \subset \mathbb{R}^N$ gibt mit $U = \tilde U \cap M$. Eine Umgebung von x in M ist dementsprechend eine in M offene Menge $U \subset M$ mit $x \in U$.
Sprechweise. Wir verwenden häufig die folgende Sprechweise: $M \subset \mathbb{R}^N$ bzw. $f : M \longrightarrow \mathbb{R}^m$ hat ***lokal in*** $a \in M$ eine Eigenschaft E, wenn a eine offene Umgebung $U \subset M$ besitzt, so dass U bzw. $f|_U$ die Eigenschaft E hat. Wir sagen auch: $M \subset \mathbb{R}^N$ bzw. $f : M \longrightarrow \mathbb{R}^m$ hat ***lokal*** eine Eigenschaft E, wenn jeder Punkt $x \in M$ eine offene Umgebung $U \subset M$ besitzt, so dass U bzw. $f|_U$ die Eigenschaft E hat.

11.2.1. **Definition.** Seien $n,N \in \mathbb{N}_0$. Eine nichtleere Teilmenge $M \subset \mathbb{R}^N$ heißt *n-**dimensionale Untermannigfaltigkeit*** des $\mathbb{R}^N$, falls gilt: Zu jedem Punkt $x \in M$ existieren

(i) eine offene Umgebung U von x in M,
(ii) eine Einbettung $\psi : D \to \mathbb{R}^N$ mit $D \subset \mathbb{R}^n$ offen, so dass $\psi(D) = U$.

Die Einbettung ψ heißt eine (***lokale***) ***Parametrisierung*** von M um den Punkt $x \in M$. Das Paar (U,φ), wobei $\varphi := \psi^{-1} : U \to D$, heißt ***Karte*** oder ***lokales Koordinatensystem*** von M um x mit ***Kartenbereich*** U.

Seien (U_1,φ_1) und (U_2,φ_2) zwei Karten der Untermannigfaltigkeit $M\subset\mathbb{R}^N$ um den Punkt $x\in M$. Die Abbildung $\varphi_2\circ\varphi_1^{-1}:\varphi_1(U_1\cap U_2)\subset\mathbb{R}^n\to\varphi_2(U_1\cap U_2)\subset\mathbb{R}^n$ heißt ***Koordinatentransformation*** oder ***Kartenübergang*** zwischen den Karten (U_1,φ_1) und (U_2,φ_2).

Eine Familie $\mathcal{A}=(U_i,\varphi_i)_{i\in I}$ heißt ***Atlas*** von M, falls die Karten ganz M überdecken, d.h. $M=\bigcup_{i\in I}U_i$.

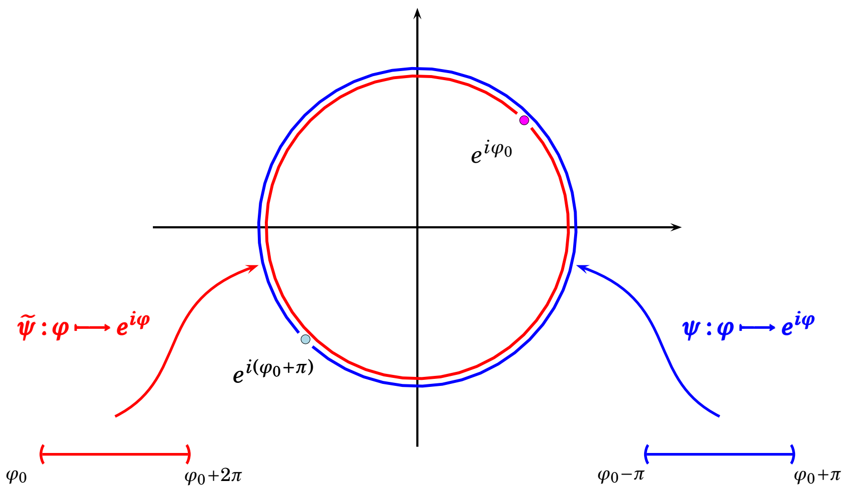
Ist M eine n-dimensionale Untermannigfaltigkeit, so heißt die Zahl n die ***Dimension*** von M. Wir schreiben oft M^n, um die Dimension anzudeuten. Die Zahl $N-n$ heißt die ***Kodimension*** von M^n in $\mathbb{R}^N$.

11.2.2. **Beispiel.** Um zu entscheiden, ob eine Teilmenge $M\subset\mathbb{R}^N$ eine Untermannigfaltigkeit ist, müssen wir einen Atlas für diese Menge angeben. Andererseits gibt es dann unendlich viele verschiedene Atlanten (Beispiel: Sphäre, siehe unten). Wegen $\mathbb{R}^0=\{0\}$ ist eine 0-dimensionale Untermannigfaltigkeit ein Punkt oder eine diskrete Menge von Punkten in $\mathbb{R}^N$.

(i) Jede ***offene Teilmenge*** $U\subset\mathbb{R}^N$ ist eine N-dimensionale Untermannigfaltigkeit des $\mathbb{R}^N$; eine Parametrisierung ist Id_U, und (U,Id_U) ist ein Atlas von U.

(ii) Der ***Kreis*** $S^1=\{(x,y)\in\mathbb{R}^2:x^2+y^2=1\}$ ist eine 1-dimensionale Untermannigfaltigkeit des $\mathbb{R}^2$. Als lokale Parametrisierung um einen Punkt $(x_0,y_0)=e^{i\varphi_0}$ kann man die Einbettung $\psi:(\varphi_0-\pi,\varphi_0+\pi)\to S^1$, $\psi(\varphi)=e^{i\varphi}$ nehmen. Die zugehörige Karte $\psi^{-1}:S^1\setminus\{e^{i(\varphi_0+\pi)}\}\to(\varphi_0-\pi,\varphi_0+\pi)$ ist ein Zweig der Argument-Abbildung und und heißt auch ***Winkelkoordinate***. Sie ist gegeben durch $z\mapsto\arg(ze^{-i(\varphi_0-\pi)})+\varphi_0-\pi$.

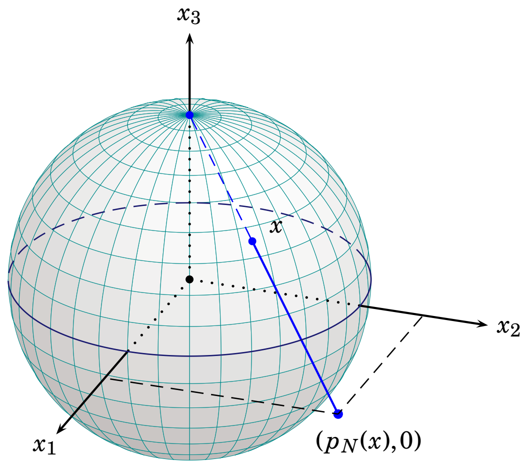
Da $\mathrm{Im}(\psi)=S^1\setminus\{e^{i(\pi+\varphi_0)}\}$, brauchen wir eine weitere Parametrisierung, um S^1 zu überdecken. Wir betrachten z.B. die Parametrisierung $\widetilde{\psi}:(\varphi_0,\varphi_0+2\pi)\to S^1$ mit $\mathrm{Im}(\widetilde{\psi})=S^1\setminus\{e^{i\varphi_0}\}$. Es gilt dann $S^1=\mathrm{Im}(\psi)\cup\mathrm{Im}(\widetilde{\psi})$, und $\{(\mathrm{Im}(\psi),\psi^{-1}),(\mathrm{Im}(\widetilde{\psi}),\widetilde{\psi}^{-1})\}$ ist ein Atlas von S^1. Der Kartenübergang ist $\widetilde{\psi}\circ\psi^{-1}:(\varphi_0-\pi,\varphi_0)\cup(\varphi_0,\varphi_0+\pi)\to(\varphi_0,\varphi_0+\pi)\cup(\varphi_0+\pi,\varphi_0+2\pi)$, $\varphi\mapsto\arg(e^{i(\varphi-\varphi_0)})+\varphi_0$, also $\varphi\mapsto\varphi+2\pi$ falls $\varphi\in(\varphi_0-\pi,\varphi_0)$ und $\varphi\mapsto\varphi$ falls $\varphi\in(\varphi_0,\varphi_0+\pi)$.

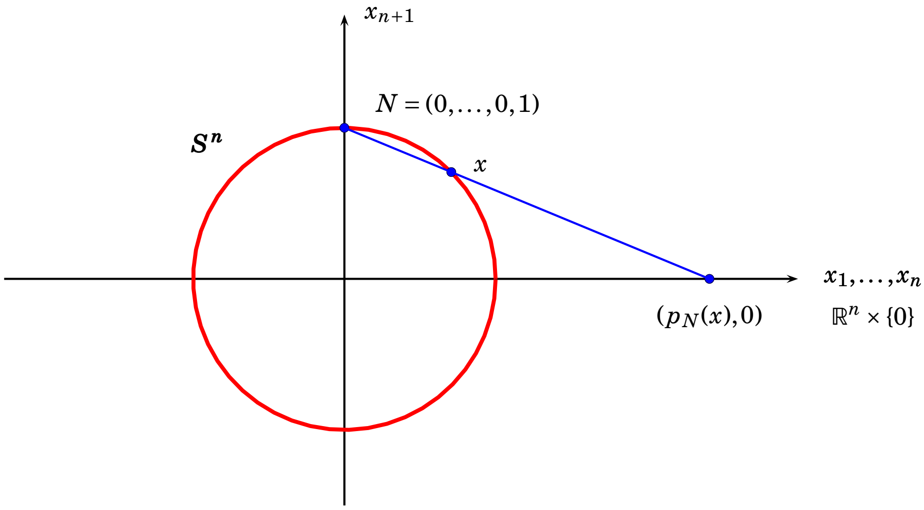


Beide Karten sind aus der Standardparametrisierung des Kreises (Def. 8.4.10) hergeleitet. Hier benutzen wir allerdings den Begriff Parametrisierung in einem engeren Sinn.

Wir nennen 1-dimensionale Untermannigfaltigkeiten des $\mathbb{R}^N$ ***reguläre*** oder ***glatte Kurven***.

(iii) Die n-***Sphäre*** $S^n_r=\{x\in\mathbb{R}^{n+1}:\|x\|_2=r\}$ vom Radius $r>0$ ist eine n-dimensionale Untermannigfaltigkeit des $\mathbb{R}^{n+1}$. Sie besitzt einen Atlas aus zwei Karten, nämlich den stereographischen Projektionen.





Sei $r=1$. Die stereographischen Projektionen aus dem Nordpol $N=e_n$ und Südpol $N=-e_n$ sind definiert durch

(11.2) $$p_N : S^n\setminus\{N\}\to\mathbb{R}^n,\quad p_N(q)=\frac{(q_1,\dots,q_n)}{1-q_{n+1}}$$

(11.3) $$p_S : S^n\setminus\{S\}\to\mathbb{R}^n,\quad p_N(q)=\frac{(q_1,\dots,q_n)}{1+q_{n+1}}$$

Die entsprechende Parametrisierungen sind gegeben durch

(11.4) $$p_N^{-1} : \mathbb{R}^n\to S^n\setminus\{N\},\quad p_N^{-1}(x)=\frac{(2x_1,\dots,2x_n,\|x\|^2-1)}{1+\|x\|^2}$$

(11.5) $$p_S^{-1} : \mathbb{R}^n\to S^n\setminus\{S\},\quad p_S^{-1}(x)=\frac{(2x_1,\dots,2x_n,1-\|x\|^2)}{1+\|x\|^2}$$

Es gibt viele anderen Karten auf S_r^n, z.B. die Koordinatenprojektionen der Halbsphären $S_r^n\cap\{x\in\mathbb{R}^{n+1}: x_i>0\}$ oder $S_r^n\cap\{x\in\mathbb{R}^{n+1}: x_i<0\}$ auf die Hyperebene $\{x\in\mathbb{R}^{n+1}: x_i=0\}\cong\mathbb{R}^n$.

Besonders nützliche Karten sind gegeben durch ***sphärische Koordinaten***. Die zugehörige Parametrisierung ist erhalten durch Einschränkung der Polarkoordinatenabbildung (10.3)auf $\{r\}\times(0,2\pi)\times(-\frac{\pi}{2},\frac{\pi}{2})^{n-2}$:

(11.6) $$\begin{aligned}&\psi_n : (0,2\pi)\times(-\tfrac{\pi}{2},\tfrac{\pi}{2})^{n-2}\longrightarrow S_r^{n-1}\setminus\{x\in\mathbb{R}^n : x_1\geqslant 0, x_2=0\}\\ &\psi_n(\varphi,\vartheta_1,\dots,\vartheta_{n-2})=P_n(r,\varphi,\vartheta_1,\dots,\vartheta_{n-2})\end{aligned}$$

insbesondere für $n=2$ gilt

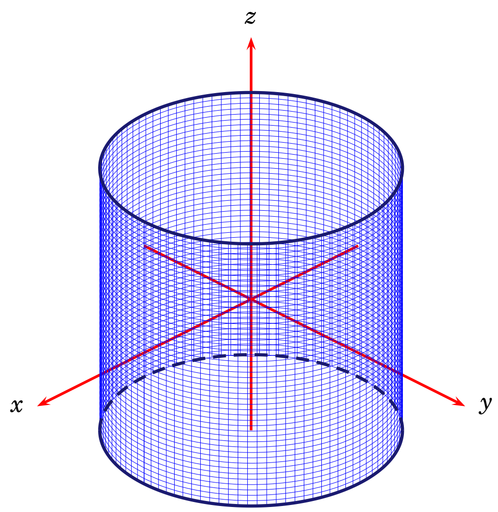
(11.7) $$\begin{aligned}&\psi_3 : (0,2\pi)\times(-\tfrac{\pi}{2},\tfrac{\pi}{2})\longrightarrow S_r^2\\ &\psi_3(\varphi,\vartheta)=(P_2(r,\varphi)\cos\vartheta, r\sin\vartheta)=(r\cos\varphi\cos\vartheta, r\sin\varphi\cos\vartheta, r\sin\vartheta).\end{aligned}$$

Wir nennen 2-dimensionale Untermannigfaltigkeiten des $\mathbb{R}^N$ ***reguläre*** oder ***glatte Flächen*** (Beispiele dafür sind die 2-Sphären $S_r^2 \subset \mathbb{R}^3$.).
Wir nennen $(N-1)$-dimensionale Untermannigfaltigkeiten des $\mathbb{R}^N$ ***Hyperflächen*** des $\mathbb{R}^N$. Die Sphäre S_r^n ist also eine Hyperfläche von $\mathbb{R}^{n+1}$.

(iv) Die Lemniskate ist keine Untermannigfaltigkeit, denn für den Schnittpunkt $(0,0)$ gibt es keine lokale Parametrisierung. Das ist anschaulich klar, da jede Umgebung von $(0,0)$ in M zwei sich kreuzende Linien enthält, also nicht homöomorph zu einer Teilmenge von $\mathbb{R}$ sein kann. (Dies kann man am besten mit dem Begriff der Zusammenhangskomponenten präzisieren: Jede Umgebung V von $(0,0)$ zerfällt bei Wegnehmen von $(0,0)$ in mindestens vier Teile (Zusammenhangskomponenten). Wenn $\psi : D \to \psi(D) = U$ eine Parametrisierung um $(0,0)$ wäre, so wäre ψ ein Homöomorphismus, daher müsste auch gelten: D zerfällt bei Wegnehmen eines Punktes in mindestens 4 Teile; da man o.B.d.A. die Menge D als zusammenhängend annehmen kann, also als offenes Intervall, zerfällt es aber in nur 2 Teile. Somit kann ψ nicht existieren.)

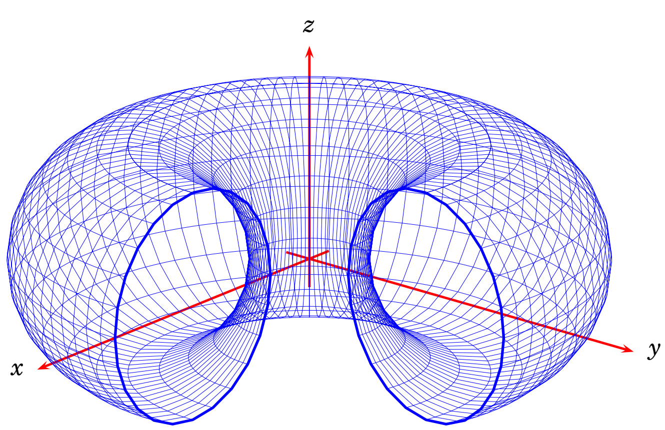
Andere Argumente erhält man mit Hilfe eines der Kriterien aus Satz 11.2.6.

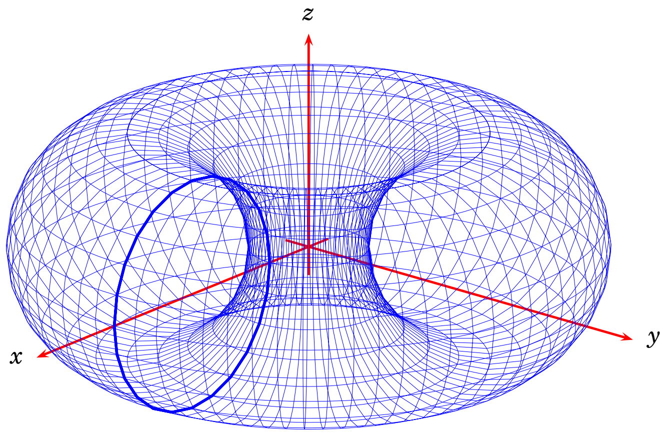
(v) Der gerade ***Zylinder*** entsteht durch Verschiebung eines Kreises parallel zu einer Geraden durch den Kreismittelpunkt, nämlich der Achse, die senkrecht zur Kreisebene steht.



Wenn diese Achse die z-Achse ist und der Kreis den Radius r hat, dann ist der Zylinder die Menge $M = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 = r^2\}$.

(vi) Wir betrachten den Rotationstorus im $\mathbb{R}^3$. Dies ist die auf folgende Weise definierte Menge T^2: Es sei ein Kreis in der (x,z)-Ebene um $(r_1, 0)$ mit Radius r_2 gegeben, wobei $0 < r_2 < r_1$. Sei T^2 die Punktmenge im $\mathbb{R}^3$, die bei Drehung dieses Kreises um die z-Achse entsteht. T^2 heißt ***Rotationstorus***.





Wir zeigen

$$T^2 = \{((r_1 + r_2 \cos u)\cos v, (r_1 + r_2 \cos u)\sin v, r_2 \sin u) : u, v \in \mathbb{R}\} :$$

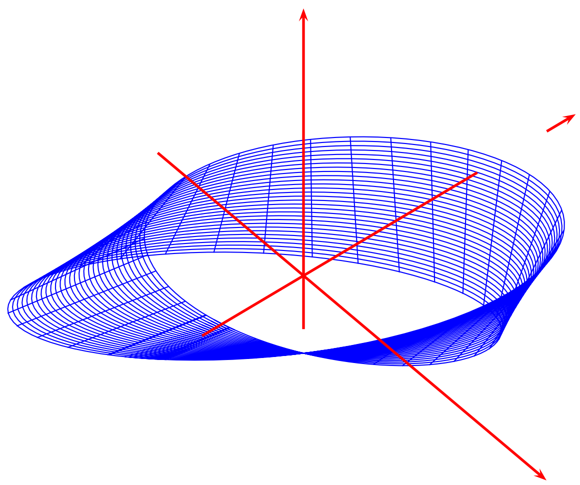
Der Kreis in der (x,z)-Ebene wird durch $\gamma(u) := (r_1 + r_2 \cos u, r_2 \sin u)$ parametrisiert. Bei Drehung um die z-Achse bleibt die z-Koordinate unverändert. Vom Nullpunkt verschiedene Punkte in der (x,y)-Ebene werden durch Polarkoordinaten der Form $pe^{iv} = p(\cos v + i \sin v) = (p\cos v, p \sin v)$ beschrieben, wobei $p \in \mathbb{R}_+$ den Abstand vom Nullpunkt und v den Winkel zur x-Achse beschreibt. Folglich gilt für die Koordinaten eines Punktes des Rotationstorus:

$$x = (r_1 + r_2 \cos u)\cos v\,, \quad y = (r_1 + r_2 \cos u)\sin v\,, \quad z = r_2 \sin u\,.$$

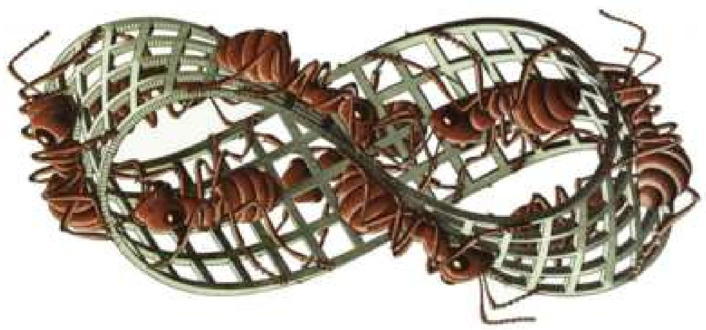
Dies ist eine parametrische Darstellung des Rotationstorus. Eine Darstellung als Lösungsmenge einer Gleichung ist

$$T^2 = \left\{(x,y,z) \in \mathbb{R}^3 : \left(\sqrt{x^2+y^2} - r_1\right)^2 + z^2 - r_2^2 = 0\right\}.$$

(vii) Das ***Möbiusband*** entsteht, wenn man einen langen rechteckigen Streifen Papier an beiden Enden zusammenklebt, ein Ende aber vor dem Zusammenkleben um 180° verdreht.



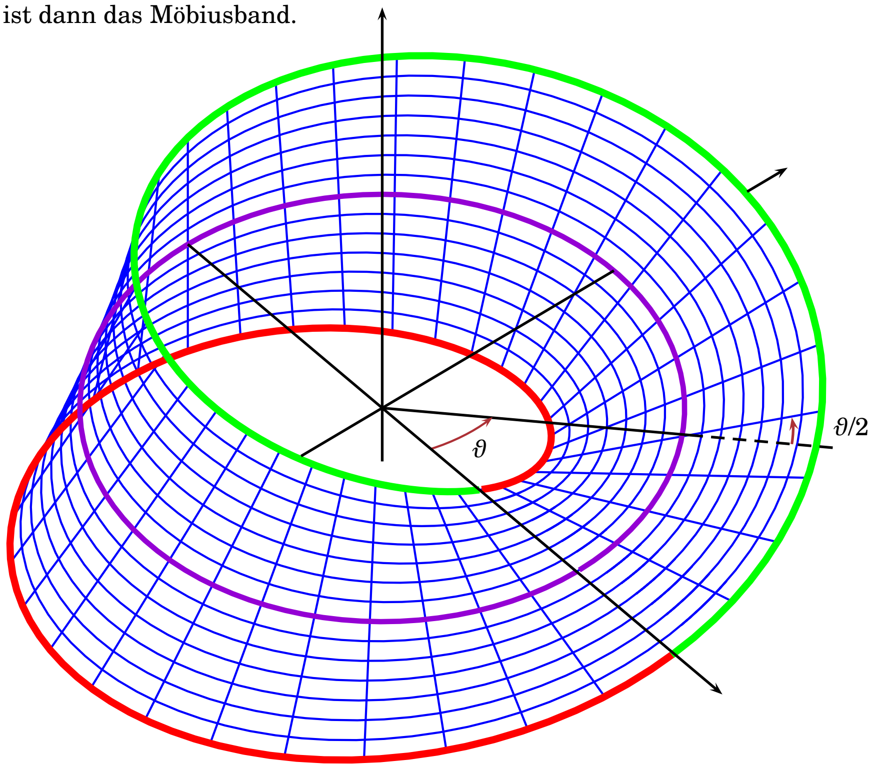
Wenn ein Käfer auf dieser Fläche entlangkriecht und immer in der Mitte des Streifens bleibt, so kommt er an seine Ausgangsstelle zurück, aber mit den Beinen nach oben. Anders als Sphäre oder Zylinder hat das Möbiusband nur eine Seite. Dies bedeutet, dass die Fläche *nicht orientierbar* ist.



Eine mögliche Realisierung des Möbiusbandes in $\mathbb{R}^3$ ergibt sich folgendermaßen. Betrachte das offene Segment $\overline{AB} \subset \mathbb{R}^3$, $A = (1,0,0)$, $B = (3,0,0)$. Wir bewegen das Segment in $\mathbb{R}^3$ so, dass zu dem Zeitpunkt $\vartheta \in \mathbb{R}$ das neue Segment $\overline{A_\vartheta B_\vartheta}$ die folgenden Eigenschaften hat:

- der Mittelpunkt von $\overline{A_\vartheta B_\vartheta}$ ist $(2\cos\vartheta, 2\sin\vartheta, 0)$,
- $\overline{A_\vartheta B_\vartheta}$ liegt in der von $(2\cos\vartheta, 2\sin\vartheta, 0)$, $(0,0,0)$ erzeugten Ebene, und der Winkel zwischen $\overline{A_\vartheta B_\vartheta}$ und dem Vektor $(2\cos\vartheta, 2\sin\vartheta, 0)$ ist $\frac{\vartheta}{2}$.

$M = \bigcup_{\vartheta\in\mathbb{R}} \overline{A_\vartheta B_\vartheta}$ ist dann das Möbiusband.



Es ist klar, dass

$$\begin{aligned}
A_\vartheta &= (2\cos\vartheta - \cos\tfrac{\vartheta}{2}\cos\vartheta,\ 2\sin\vartheta - \cos\tfrac{\vartheta}{2}\sin\vartheta,\ -\sin\tfrac{\vartheta}{2}),\\
B_\vartheta &= (2\cos\vartheta + \cos\tfrac{\vartheta}{2}\cos\vartheta,\ 2\sin\vartheta + \cos\tfrac{\vartheta}{2}\sin\vartheta,\ \sin\tfrac{\vartheta}{2}),\\
\overline{A_\vartheta B_\vartheta} &= \{(2\cos\vartheta + t\cos\tfrac{\vartheta}{2}\cos\vartheta,\ 2\sin\vartheta + t\cos\tfrac{\vartheta}{2}\sin\vartheta,\ t\sin\tfrac{\vartheta}{2}) =: f(t,\vartheta) : t\in(-1,1)\}.
\end{aligned}$$

Dann ist $M = \mathrm{Im}(f)$, wobei $f : (-1,1)\times\mathbb{R} \longrightarrow \mathbb{R}^3$. Nun ist M eine 2-dimensionale Untermannigfaltigkeit von $\mathbb{R}^3$:

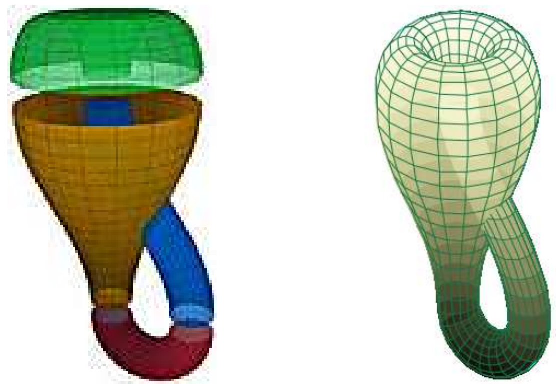
Sei $\psi = f|_{(-1,1)\times(0,2\pi)}$. Dann ist ψ eine Parametrisierung von $M \smallsetminus \overline{AB} = M \smallsetminus \overline{A_0B_0}$. Sei $G : \mathbb{R}^3 \smallsetminus \{(x,y,z) : y = 0, x \geqslant 0\} \longrightarrow \mathbb{R}^2$ mit $G(x,y,z) = \left(\arg(x,y), \dfrac{z}{2\sin(\arg(x,y))}\right)$.

Dann gilt $G\circ\psi = \mathrm{Id}_{(-1,1)\times(0,2\pi)}$. Dies beweist:

- ψ ist injektiv,
- ψ ist Immersion (Kettenregel: $dG(\psi(t,\vartheta))\cdot d\psi(t,\vartheta) = \mathrm{Id}_{\mathbb{R}^2}$),
- $\psi : (-1,1)\times(0,2\pi) \longrightarrow M \smallsetminus \overline{AB}$ ist Homöomorphismus (die Inverse ist $G|_{M\smallsetminus\overline{AB}}$).

Also ist ψ eine Einbettung und Parametrisierung von $\mathrm{Im}(\psi) = M \setminus \overline{AB}$. Genauso zeigt man, dass $\varphi = f|_{(-1,1)\times(-\pi,\pi)}$ eine Parametrisierung von $M \setminus \overline{A_\pi B_\pi}$ ist. Die Bilder der Parametrisierungen ψ und φ überdecken M. Daher ist M eine 2-dimensionale Untermannigfaltigkeit von $\mathbb{R}^3$.

(viii) Die ***Kleinsche Flasche*** entseht, indem man eine Zylinderoberfläche an ihren Enden verklebt, allerdings wird einer der Kreise am Ende mit Pfeilen im Uhrzeigersinn, einer mit Pfeilen gegen den Uhrzeigersinn versehen, und dann werden die Pfeile zur Deckung gebracht. Dieses Gebilde ist im $\mathbb{R}^3$ nicht realisierbar, d.h. es gibt keine Einbettung in $\mathbb{R}^3$. Eine gute Näherung erhalten wir, wenn wir ein Ende des Zylinders durch seine Oberfläche hindurchstecken und dann von innen mit dem anderen Ende verkleben.



Als nächstes wollen wir andere Möglichkeiten schildern, zu überprüfen, ob eine Menge M eine Untermanigfaltigkeit des $\mathbb{R}^N$ ist. Wir brauchen dazu die folgenden Begriffe.

11.2.3. **Definition.** Eine Teilmenge $G \subset \mathbb{R}^N$ heißt ***Graph der Dimension*** n, falls G bis auf eine Permutation der Variablen in $\mathbb{R}^N$ Graph einer glatten Abbildung $f : D \to \mathbb{R}^{N-n}$ mit $D \subset \mathbb{R}^n$ ist. Genauer muss gelten: Es existiert eine Permutation p von $\{1,\dots,N\}$, eine offene Menge $D \subset \mathbb{R}^n$ und eine glatte Abbildung $f : D \to \mathbb{R}^{N-n}$, so dass

$$G = \{(x_1,\dots,x_N) : (x_{p(1)},\dots,x_{p(n)}) \in D, (x_{p(n+1)},\dots,x_{p(N)}) = f(x_{p(1)},\dots,x_{p(n)})\}.$$

Beispiel: Der obere Einheitshalbkreis in $\mathbb{R}^2$ ist $G_o = \{(x,\sqrt{1-x^2}) : x \in (-1,1)\}$, also der Graph der Funktion $y = f(x) = \sqrt{1-x^2}$ auf $(-1,1)$. Der rechte Halbkreis ist $G_r = \{(\sqrt{1-y^2}, y) : y \in (-1,1)\}$, also der Graph der Funktion $x = f(y) = \sqrt{1-y^2}$ auf $(-1,1)$.

11.2.4. **Definition.** Eine Menge $M \subset \mathbb{R}^N$ heißt ***Lösungsmenge eines Systems von*** m ***unabhängigen Gleichungen***, falls gilt: Es gibt eine offene Umgebung D von M in $\mathbb{R}^N$, eine glatte Abbildung $F : D \to \mathbb{R}^m$ und einen regulären Wert $c \in \mathrm{Im}(F)$ von F, so dass $M = F^{-1}(c)$.

Beispiel: Die Sphäre $S^2 = f^{-1}(0)$, wobei $f(x,y,z) = x^2+y^2+z^2-1$. Es gilt $J_f(x,y,z) = (2x,2y,2z)$, und dies hat Rang 1, da $(x,y,z) \neq (0,0,0)$ für alle $(x,y,z) \in S^2$ gilt.

11.2.5. **Definition.** Eine Menge $M \subset \mathbb{R}^N$ ***ist zu*** $\mathbb{R}^n$ ***bügelbar***, falls gilt: Es existiert eine offene Umgebung D von M und ein Diffeomorphismus $\Phi : D \to D'$ von D auf eine offene Teilmenge $D' \subset \mathbb{R}^N$, so dass $\Phi(D \cap M) = \{y \in D' : y_{n+1} = \ldots = y_N = 0\} = D' \cap (\mathbb{R}^n \times \{0\})$. (Dabei bezeichnet 0 den Nullvektor im Teilraum R^{N-n}.)

11.2.6. **Satz** (Charakterisierung der Untermannigfaltigkeiten)**.** *Seien $n,N \in \mathbb{N}_0$ und $M \subset \mathbb{R}^N$. Die folgenden Bedingungen sind äquivalent:*

(i) *M ist eine n-dimensionale Untermannigfaltigkeit des $\mathbb{R}^N$;*
(ii) *M ist lokal zu $\mathbb{R}^n$ bügelbar;*
(iii) *M ist lokal Lösungsmenge eines Systems von $N-n$ unabhängigen Gleichungen;*
(iv) *M ist lokal Graph der Dimension n.*

Beweis: Wir zeigen zunächst die Implikation von **(i)** nach **(ii)**. Dazu benutzen wir den Satz über die Struktur einer lokalen Immersion. Sei also $p \in M$ und $\Psi : D \longrightarrow M$ eine Einbettung mit D offene Teilmenge des $\mathbb{R}^n$. Sei außerdem $a \in D$ mit $\Psi(a) = p$. Da $\Psi : D \longrightarrow \Psi(D)$ ein Homöomorphismus ist, so ist $\Psi(D)$ offen in M, also existiert eine offenen Menge $U \subset \mathbb{R}^n$ mit $U \cap M = \Psi(D)$. Da Ψ eine Immersion in a ist, folgt aus Satz 11.1.4, dass es ein $\varepsilon > 0$ gibt, mit $W^n_\varepsilon(a) \subset D$, eine offene Umgebung V von p in $\mathbb{R}^n$ und ein Diffeomorphismus $\Phi : W^N_\varepsilon(a,0) \longrightarrow V$ existieren, so dass $\Psi|_{W^n_\varepsilon(a)} = \Phi \circ j$ ist, wobei $j : W^n_\varepsilon(a) \longrightarrow W^N_\varepsilon(a,0)$ die Inklusion von $x \longmapsto (x,0)$ ist. Wir wählen V, so dass $V \subset U$. Dann gilt $M \cap V = f(W^n_\varepsilon(a)) \subset V$. Es folgt, dass $\Phi^{-1}(M \cap V) = \Phi^{-1}(W^n_\varepsilon(a)) = W^N_\varepsilon(a,0) \cap (\mathbb{R}^n \times 0)$, das heißt, Φ^{-1} bügelt $M \cap V$.

Nun zeigen wir die Implikation **(ii)** nach **(iii)**. Seien dazu $D, D' \subset \mathbb{R}^n$ offen und $\Phi : D \longrightarrow D'$ ein Diffeomorphismus mit $\Phi(D \cap M) = D' \cap \mathbb{R} \times 0$. Das heißt $D \cap M = \{p \in D : \Phi_{n+1}(p) = \ldots = \Phi_N(p) = 0\} = F^{-1}(0)$, wobei $F : D \longrightarrow \mathbb{R}^{N-n}$, $F = (\Phi_{n+1}, \ldots, \Phi_N)$. Nun ist 0 ein regulärer Wert für F und die Zeilen der Jacobimatrix J_F von F sind Zeilen der Jacobimatrix J_Φ von Φ. Da Φ ein Diffeomorphismus ist, sind die Zeilen von J_Φ linear unabhängig, also $\operatorname{Rang} J_\Phi = N - n$ in D.

Die Implikation von **(iii)** nach **(iv)** ist eine direkte Konsequenz aus dem Satz über implizierte Funktionen.

Also bleibt noch die Implikation **(iv)** nach **(i)**. Ist M lokal ein Graph, so können wir annehmen, dass $M \cap V = \{(x, f(x)) : x \in D\}$, mit V offene Teilmenge des $\mathbb{R}^N$ und D offene Teilmenge des $\mathbb{R}^n$ und $f : D \longrightarrow \mathbb{R}^{N-n}$ glatt ist. Dann ist $\Psi : D \longrightarrow \mathbb{R}^N$, $\Psi(x) = (x, f(x))$ eine Parametrisierung von $M \cap V$. □

Insbesondere folgt, dass M eine Untermannigfaltigkeit ist, wenn ganz M Lösungsmenge eines Systems von $N - n$ unabhängigen Gleichungen ist. Dies wollen wir als Satz ausdrücken:

11.2.7. **Satz** (Satz vom regulären Wert). *Seien $D \subset \mathbb{R}^N$ offen, $F : D \to \mathbb{R}^m$ eine glatte Abbildung und $c \in \operatorname{Im}(F)$ ein regulärer Wert von F. Dann ist $F^{-1}(c)$ eine $(N-m)$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^N$.*

11.2.8. **Beispiel.** Der Satz 11.2.7 ist sehr wichtig, denn Mannigfaltigkeiten sind oft als Lösungsmengen von Gleichungen gegeben.

(i) Wir betrachten nochmals die Sphäre vom Radius $r > 0$ in $\mathbb{R}^{n+1}$, $S^n_r = \{x \in \mathbb{R}^{n+1} : \|x\|_2 = r\}$. Offensichtlich gilt $S^n_r = F^{-1}(0)$ für die glatte Abbildung $F : \mathbb{R}^{n+1} \to \mathbb{R}$, $x \mapsto \|x\|^2 - r^2$. Für die Jacobimatrix von F in x gilt $J_F(x) = (2x_1, \ldots, 2x_{n+1}) = 2x$. Für $x \in S^n_r$ ist $\operatorname{Rang} J_F(x) = 1$. Somit ist S^n_r nach Satz 11.2.7 eine n-dimensionale Untermannigfaltigkeit des $\mathbb{R}^{n+1}$.

(ii) Sei allgemeiner $P : \mathbb{R}^N \to \mathbb{R}$ ein homogenes Polynom vom Grad $r \geqslant 1$. Für jedes $c \in \mathbb{R} \setminus \{0\}$ ist $P^{-1}(c)$ eine Hyperfläche.

(iii) Sei $A = (a_{ij}) \in M_{n\times n}(\mathbb{R})$ eine symmetrische Matrix und $c \in \mathbb{R} \setminus \{0\}$. Wir definieren

$$Q := \{x \in \mathbb{R}^n : \langle Ax, x\rangle = \sum a_{ij} x_i x_j = c\}, \text{ d.h. } Q = f^{-1}(c),$$

wobei $f(x) = \langle Ax, x\rangle$. Dann ist c ein regulärer Wert von f, da $J_f(x) = 2Ax \neq 0$ für $x \in Q$. Also ist Q eine $(n-1)$-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$. Eine in dieser Form gegebene Untermannigfaltigkeit des $\mathbb{R}^n$ heißt Quadrik.

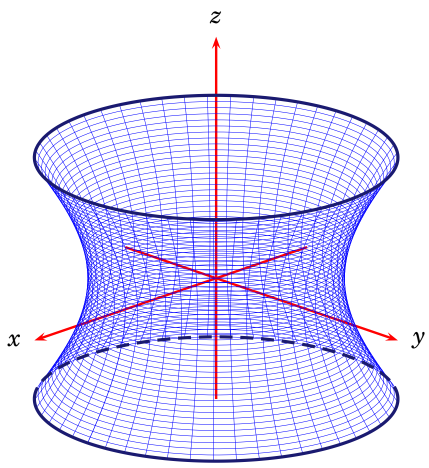
Beispiele in $\mathbb{R}^3$:

Für $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ und $c = 1$ ist Q ein Zylinder.

Für $A = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}$, $a_1, a_2, a_3 > 0$ und $c = 1$ ist Q ein Ellipsoid mit Halbachsen $\frac{1}{\sqrt{a_1}}, \frac{1}{\sqrt{a_2}}, \frac{1}{\sqrt{a_3}}$.

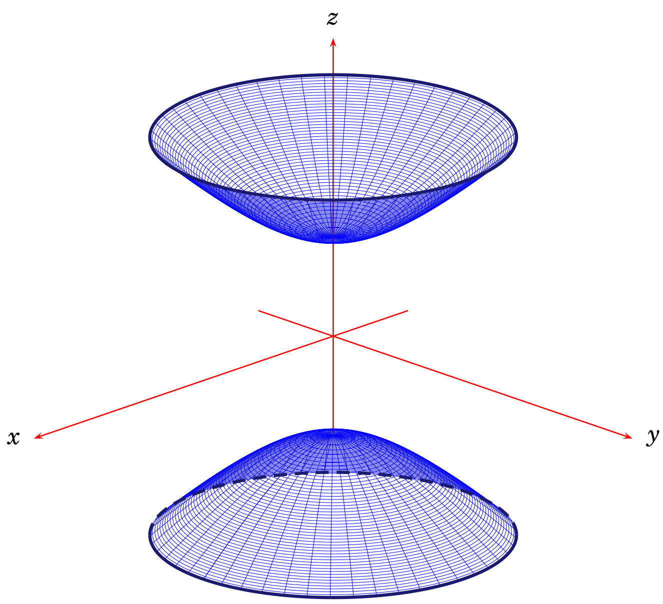
Für $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ und $c = 1$ ist Q ein einschaliges Hyperboloid:

$$Q = \{(x, y, z) : x^2 + y^2 - z^2 = 1\}.$$



Für $A = \left(\begin{smallmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$ und $c = 1$ ist Q ein zweischaliges Hyperboloid:

$$Q = \{(x, y, z) : -x^2 - y^2 + z^2 = 1\}.$$



(iv) Die orthogonale Gruppe $O(n) := \{A \in M_{n\times n}(\mathbb{R}) : A^T \cdot A = E_n\}$ ist eine $\binom{n}{2}$-dimensionale Untermannigfaltigkeit von $M_{n\times n}(\mathbb{R}) \cong \mathbb{R}^{n^2}$ (siehe [13, § 3.5]).

(v) Die Umkehrung des Satzes vom regulären Wert ist falsch, d.h. wenn $F^{-1}(c)$ eine Untermannigfaltigkeit ist, folgt i.A. nicht, dass c ein regulärer Wert wäre. Anders gesagt, es gibt kritische Werte c, so dass $F^{-1}(c)$ eine Untermannigfaltigkeit ist. Als Beispiel betrachte $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = x^2$. Dann ist 0 ein kritischer Wert, aber $F^{-1}(0)$ ist die y-Achse, also eine Untermannigfaltigkeit.

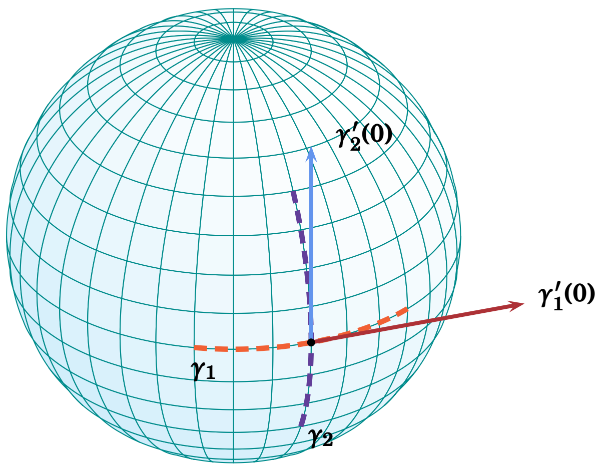
11.3. **Tangential- und Normalenräume an Untermannigfaltigkeiten.** Im folgenden sei $M \subset \mathbb{R}^N$ eine n–dimensionale Untermannigfaltigkeit.

11.3.1. **Definition.** Sei $x \in M$. Die Menge der Vektoren

$$\boxed{T_xM := \{v \in \mathbb{R}^N : \exists \varepsilon > 0\ \exists \text{ glatte Kurve } \gamma : (-\varepsilon, \varepsilon) \to M \text{ mit } \gamma(0) = x,\ \gamma'(0) = v\}}$$

heißt ***Tangentialraum*** von M im Punkt x. Die Elemente von T_xM heißen Tangentialvektoren an M im Punkt x.

Wir können anstelle von $t_0 = 0$ in der Definition eine beliebige Zahl (Zeit) $t_0 \in \mathbb{R}$ betrachten: $v = \gamma'(t_0)$, wobei $\gamma : (t_0 - \varepsilon, t_0 + \varepsilon) \to M$, $\gamma(t_0) = x$.



Tangentialvektoren sind genau die Tangenten (Geschwindigkeitsvektoren) an Kurven, die in der Untermannigfaltigkeit verlaufen. Man stellt sich einen Tangentialvektor als an x festgeklebt vor und damit den Tangentialraum als Unterraum eines Exemplars von $\mathbb{R}^N$, dessen Ursprung bei x sitzt.

11.3.2. **Satz.** (i) *Sei $\psi : D \subset \mathbb{R}^n \to U$ eine Parametrisierung von M um x mit $\psi(a) = x$ (wie in Def. 11.2.1). Dann gilt $T_xM = \operatorname{Im} d\psi(a)$. Insbesondere ist T_xM ein n–dimensionaler Untervektorraum des $\mathbb{R}^N$, und $\{\partial_1\psi(a), \dots, \partial_n\psi(a)\}$ ist eine Basis von T_xM.*
(ii) *Sei D eine offene Umgebung von x in $\mathbb{R}^N$, sei $F : D \to \mathbb{R}^{N-n}$ eine glatte Abbildung und $c \in \operatorname{Im}(F)$ ein regulärer Wert von F, so dass $M \cap D = F^{-1}(c)$. Dann gilt*

$$\boxed{T_xM = \ker dF(x) = \langle\{\operatorname{grad} F_1(x), \dots, \operatorname{grad} F_{N-n}(x)\}\rangle^{\perp}}.$$

11.3.3. **Beispiel.**

(i) Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und $x \in U$. Dann gilt $T_xU = \mathbb{R}^n$, da für jedes $v \in \mathbb{R}^n$ gilt: $v = \gamma'(0)$ mit $\gamma(t) := x + tv$.

(ii) Sei $S_r^n := \{x \in \mathbb{R}^{n+1} : \|x\|_2 = r\}$ eine n-dimensionale Sphäre und $x \in S_r^n$. Dann gilt $T_xS_r^n = \{v \in \mathbb{R}^{n+1} : \langle x, v\rangle = 0\}$.

(iii) Für eine Quadrik Q wie in Beispiel 11.2.8 und $x \in Q$ gilt $df(x)\cdot v = \langle Ax, v\rangle$ also $T_xQ = \ker df(x) = \{v \in \mathbb{R}^n : v \perp Ax\}$.

(iv) Sind M_1, M_2 zwei Untermannigfaltigkeiten und $M = M_1 \times M_2$, so gilt $T_xM = T_{x_1}M \times T_{x_2}M_2$ für alle $x = (x_1, x_2) \in M$.

11.3.4. **Definition.** Sei $M \subset \mathbb{R}^N$ eine n–dimensionale Untermannigfaltigkeit und $x \in M$. Der zum Tangentialraum T_xM orthogonale Vektorraum

$$\boxed{N_xM := \{w \in \mathbb{R}^N : w \perp T_xM\} = (T_xM)^{\perp}}$$

heißt ***Normalenraum*** an M im Punkt $x \in M$. Die Elemente von N_xM heißen Normalenvektoren an M im Punkt x.

11.3.5. **Beispiel.** Sei $M \subset \mathbb{R}^{n+1}$ eine Hyperfläche (d.h. $\dim M = n$). Aus Satz 11.3.2 folgt sofort:

(i) Ist $\psi : U \to \mathbb{R}^{n+1}$ eine lokale Parametrisierung um $x = \psi(u) \in M$, dann ist

$$N_xM = \mathbb{R}\underbrace{\partial_1\psi(u) \times \ldots \times \partial_n\psi(u)}_{\text{Kreuzprodukt}}.$$

(ii) Sei D eine offene Umgebung von x in $\mathbb{R}^N$, $F : D \to \mathbb{R}$ eine glatte Abbildung und $c \in \operatorname{Im}(F)$ ein regulärer Wert von F, so dass $M \cap D = F^{-1}(c)$. Dann gilt $N_xM = \mathbb{R}\operatorname{grad} F(x)$. Beispiel: $N_xS_r^n = \mathbb{R}x$.

(iii) Für eine Quadrik Q wie in Beispiel 11.2.8 und $x \in Q$ gilt $N_xQ = \mathbb{R}Ax$.

11.3.6. **Satz** (Umformulierung der Multiplikatorenregel von Lagrange). *Sei $M \subset \mathbb{R}^N$ eine Untermannigfaltigkeit, $U \subset \mathbb{R}^N$ offen und $f : U \to \mathbb{R}$ differenzierbar. Falls $f|_{U\cap M}$ in $a \in U \cap M$ ein lokales Extremum hat, so folgt* $\operatorname{grad} f(a) \in N_a M$.

11.4. **Glatte Abbildungen und ihr Differential.** Wir definieren nun den Begriff der differenzierbaren Abbildung zwischen Untermannigfaltigkeiten mit Hilfe von lokalen Parametrisierungen (oder Karten).

11.4.1. **Definition.** Seien $M_1^{n_1} \subset \mathbb{R}^{N_1}$ und $M_2^{n_2} \subset \mathbb{R}^{N_2}$ zwei Untermannigfaltigkeiten. Eine Abbildung $f : M_1 \to M_2$ heißt ***glatt***, falls für jede Karte (U,φ) von M_1 die Abbildung $f \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^{n_1} \to \mathbb{R}^{N_2}$ eine glatte Abbildung ist. Mit $\mathcal{C}^\infty(M)$ bezeichnen wir den Ring aller reellwertigen glatten Abbildungen auf M. Mit $\mathcal{C}^\infty(M_1, M_2)$ bezeichnen wir die Menge aller glatten Abbildungen zwischen den Untermannigfaltigkeiten M_1 und M_2.

11.4.2. **Beispiel.**
(i) Ist $M_1 = U \subset \mathbb{R}^{N_1}$ eine offene Teilmenge, dann stimmt der eben definierte Differenzierbarkeitsbegriff mit dem schon bekannten für Abbildungen zwischen reellen Vektorräumen überein: Betrachte die Karte $\varphi : U \to U$, $\varphi = \mathrm{Id}_U$.
(ii) Es genügt, die Differenzierbarkeit der Abbildungen $f \circ \varphi_i^{-1} : \varphi_i(U_i) \longrightarrow \mathbb{R}^{N_2}$ für einen Atlas $\mathcal{A} = \{(U_i, \varphi_i)\}_{i\in I}$ von M_1 zu überprüfen. Dazu brauchen wir das folgende wichtige Lemma:

11.4.3. **Lemma.** *Seien (U_1,φ_1) und (U_2,φ_2) zwei Karten der Untermannigfaltigkeit $M^n \subset \mathbb{R}^N$ um den Punkt $x \in M$. Dann ist der Kartenübergang $\varphi_2 \circ \varphi_1^{-1} : \varphi_1(U_1 \cap U_2) \subset \mathbb{R}^n \to \varphi_2(U_1 \cap U_2) \subset \mathbb{R}^n$ eine glatte Abbildung zwischen offenen Mengen des $\mathbb{R}^n$.*

Dies folgt aus der Tatsache, dass die Kartenabbildung φ_2 nach Satz 11.1.4 durch Einschränkung eines lokalen Diffeomorphismus im Einbettungsraum $\mathbb{R}^N$ entsteht (Φ^{-1} im Satz 11.1.4).

Dieses Lemma ist der Ausgangspunkt für die Definition der abstrakten differenzierbaren Mannigfaltigkeiten. Sei M ein topologischer Hausdorffraum. Ist $U \subset X$ eine offene Menge und $\varphi : U \to \varphi(U)$ ein Homöomorphismus von U auf eine offene Menge $\varphi(U) \subset \mathbb{R}^n$, dann heißt das Paar (U,φ) eine *n-dimensionale Karte* von M. Ein *Atlas* von M ist eine Familie $\mathcal{A} = \{(U_i,\varphi_i)\}_{i\in I}$ von n-dimensionalen Karten so, dass $M = \cup_{i\in I} U_i$ gilt und für alle $i,j \in I$ die Abbildung $\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \subset \mathbb{R}^n \to \varphi_i(U_i \cap U_j) \subset \mathbb{R}^n$ ein Diffeomorphismus ist. (Für $U_i \cap U_j = \emptyset$ ist diese Bedingung leer.) Sei $\mathcal{A}_{\max}$ ein maximaler Atlas (bzgl. der Inklusion von Atlanten, d.h. für jeden Atlas $\mathcal{B} \supset \mathcal{A}_{\max}$ folgt $\mathcal{B} = \mathcal{A}_{\max}$). Ist $\mathcal{A}$ ein Atlas von M, so existiert ein maximaler Atlas $\mathcal{A}_{\max} \supset \mathcal{A}$. Das Paar $(M, \mathcal{A}_{\max})$ heißt dann eine *n-dimensionale Mannigfaltigkeit*.

Mannigfaltigkeiten brauchen nicht in irgendeinem $\mathbb{R}^N$ zu liegen (man sagt auch: eingebettet zu sein), obwohl nach einem Satz von Whitney jede Mannigfaltigkeit mit abzählbarer Basis der Topologie in einen geeigneten $\mathbb{R}^N$ eingebettet werden kann. Der projektive Raum, der Quotient einer Mannigfaltigkeit bzgl. einer Gruppenoperation, die Verklebung zweier Mannigfaltigkeiten, das Kotangentialbündel sind Beispiele von Mannigfaltigkeiten, die a priori nicht eingebettet sind.

(iii) Sei $M \subset \mathbb{R}^N$ eine Untermannigfaltigkeit und (U,φ) eine Karte von M. Dann ist die Kartenabbildung $\varphi : U \longrightarrow V \subset \mathbb{R}^n$ glatt.
(iv) Ist $M \subset \mathbb{R}^N$ eine Untermannigfaltigkeit, so ist die Inklusionsabbildung $\iota : M \hookrightarrow \mathbb{R}^N$ glatt.
(v) Sind $f : M_1 \longrightarrow M_2$ und $g : M_2 \longrightarrow M_3$ glatt, so ist $g \circ f : M_1 \longrightarrow M_3$ glatt.
(vi) Seien $U \subset \mathbb{R}^N$ offen, $M \subset U$ eine Untermannigfaltigkeit und $f : U \to \mathbb{R}^m$ eine glatte Abbildung. Dann ist die Abbildung $f|_M : M \to \mathbb{R}^m$ ebenfalls glatt.
(vii) Es gilt:
$f \in \mathcal{C}^\infty(M) \iff$ Für jedes $x \in M$ existieren eine offene Umgebung U_x in $\mathbb{R}^N$ und $F_x \in \mathcal{C}^\infty(U_x)$, so dass $F_x|_{U_x\cap M} = f|_{U_x \cap M}$.
Dies folgt unmittelbar aus der Tatsache, dass M lokal bügelbar ist, und für $M = \mathbb{R}^n \times \{0\} \subset \mathbb{R}^N$ kann man die Fortsetzung $F : \mathbb{R}^N \to \mathbb{R}$ durch $F(x_1, \dots, x_N) = f(x_1, \dots, x_n)$ definieren. Mit Hilfe einer Zerlegung der Eins findet man eine offene Umgebung U von ganz M und $F \in \mathcal{C}^\infty(U)$, so dass $F|_M = f$ (siehe Aufgabe 11.10.3).

11.4.4. **Definition.** Sei $f : M_1^{n_1} \longrightarrow M_2^{n_2}$ eine Abbildung zwischen zwei Untermannigfaltigkeiten. Sei (U,φ) eine Karte um $x \in M_1$ und (V,ψ) eine Karte um $f(x) \in M_2$. Dann heißt die Abbildung

$$\psi \circ f \circ \varphi^{-1} : \varphi(U \cap f^{-1}(V)) \subset \mathbb{R}^{n_1} \longrightarrow \psi(V) \subset \mathbb{R}^{n_2}$$

die ***Kartendarstellung*** von f bzgl. der Karten (U,φ) und (V,ψ). Ist speziell $f : M_1 \longrightarrow \mathbb{R}^N$, so heißt $f \circ \varphi^{-1} : \varphi(U) \longrightarrow \mathbb{R}^N$ Kartendarstellung von f bzgl. der Karte (U,φ).

Es gilt: Die Abbildung F ist glatt genau dann, wenn alle ihre Kartendarstellungen $\mathcal{C}^\infty$-Abbildungen sind.

Sei (U,φ) eine Karte um $x \in M_1$, wobei $\varphi(U) \subset \mathbb{R}^n$. Wenn $x_1,\dots,x_n$ die Koordinaten von $\mathbb{R}^n$ bezeichnen, dann notieren wir die Komponenten von φ einfach als $\varphi_1 = x_1, \dots, \varphi_n = x_n$, und die Karte als $(U,\varphi = (x_1,\dots,x_n))$. Die Kartendarstellung $f \circ \varphi^{-1}$ wird auch $f \circ \varphi^{-1}(x_1,\dots,x_n) =: f(x_1,\dots,x_n)$ geschrieben.

Beispiel: Die Kartendarstellung von $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \sqrt{x^2+y^2}$ bzgl. der Polarkoordinaten (d.h. einer Karte (U,φ), wobei φ eine Inverse der Polarkoordinatenabbildung ist) ist $f(r,\varphi) := f(r\cos\varphi, r\sin\varphi) = r$.

Das Differential einer differenzierbaren Abbildung $f : M_1 \longrightarrow M_2$ definiert man in Analogie zu den Abbildungen zwischen reellen Räumen als Ableitung von f entlang von Kurven:

11.4.5. **Definition.** Sei $f : M_1 \longrightarrow M_2$ eine differenzierbare Abbildung zwischen Untermannigfaltigkeiten. Unter dem Differential der Abbildung f im Punkt $x \in M_1$ versteht man die Abbildung

$$\boxed{df(x) : T_xM_1 \longrightarrow T_{f(x)}M_2\,, \qquad \gamma'(0) \longmapsto (f\circ\gamma)'(0),}$$

wobei $\gamma : I \longrightarrow M_1$ eine glatte Kurve mit $\gamma(0) = x$ ist.

11.4.6. **Bemerkung.** (1) Die Definition von $df(x)$ ist korrekt, d.h. unabhängig von der Wahl von γ: Sei $v \in T_xM$, und sei γ eine beliebige glatte Kurve auf M_1 mit $\gamma(0) = x$ und $\gamma'(0) = v$. Wir wählen eine Karte (U,φ) um x. Dann folgt aus der Kettenregel für differenzierbare Abbildungen zwischen reellen Räumen, dass

$$(f\circ\gamma)'(0) = d(f\circ\varphi^{-1})(\varphi(x))\cdot(\varphi\circ\gamma)'(0).$$

Da die Kartenabbildung φ nach Satz 11.1.4 durch Einschränkung eines lokalen Diffeomorphismus Φ im Einbettungsraum $\mathbb{R}^N$ entsteht, folgt aus der Kettenregel weiterhin $(\varphi\circ\gamma)'(0) = (\Phi\circ\gamma)'(0) = d\Phi(x)\cdot\gamma'(0) = d\Phi(x)\cdot v$ und somit

$$(f\circ\gamma)'(0) = d(f\circ\varphi^{-1})(\varphi(x))\circ d\Phi(x)\cdot v\,.$$

Daher hängt $(f\circ\gamma)'(0)$ tatsächlich nicht von der Wahl der Kurve γ ab.

(2) Sind $M_1 \subset \mathbb{R}^{N_1}$ und $M_2 \subset \mathbb{R}^{N_2}$ offene Teilmengen, dann stimmt das soeben definierte Differential mit dem in Kapitel 9 definierten Differential überein. Für $v \in T_xM_1 = \mathbb{R}^{N_1}$ ist $v = \gamma'(0)$ mit $\gamma(t) := x + tv$. Laut Definition 11.4.5 gilt

$$df(x)(v) = \frac{d}{dt} f\,\underbrace{(x+tv)}_{\gamma(t)}|_{t=0}\,.$$

Letzteres ist aber die Richtungableitung $\partial_v f(x)$. Nach Satz 9.2.4 stimmt $df(x)$ daher mit dem üblichen Differential aus Definition 9.1.1 überein.

(vi) Viele Abbildungen sind in der Praxis als Einschränkungen gegeben. Seien $U \subset \mathbb{R}^N$ offen, $M \subset U$ und $F \in \mathcal{C}^\infty(U,\mathbb{R}^{N_2})$ mit $f = F|_M$. Dann gilt

$$\boxed{df(x) = dF(x)|_{T_xM}}\,.$$

Diese Formel erleichtert die Berechnung von $df(x)$, da $dF(x)$ die Multiplikation mit der Jacobi-Matrix $J_F(x)$ ist.

11.4.7. **Satz.** *Seien $f : M_1 \longrightarrow M_2$ und $g : M_2 \longrightarrow M_3$ differenzierbar, und sei $x \in M_1$.*

(1) *Das Differential $df(x) : T_xM_1 \longrightarrow T_{f(x)}M_2$ ist eine lineare Abbildung zwischen den Tangentialräumen.*
(2) *Es gilt die Kettenregel: $d(g\circ f)(x) = dg(f(x))\circ df(x)$.*

Andere Rechenregeln:

(1) Sind $f,g : M \to \mathbb{R}^m$ glatt, so gilt $d(f+g)(x) = d(f)(x) + d(g)(x)$ für alle $x \in M$.

(2) Ist $f=(f_1,f_2):M\to M_1\times M_2$ glatt, so ist $T_{f(x)}(M_1\times M_2)=T_{f_1(x)}M_1\times T_{f_2(x)}M_2$ (siehe Beispiel 11.3.3 (iv)) und entsprechend dieser Zerlegung gilt

$$df(x)=(df_1(x),df_2(x)):T_xM\to T_{f_1(x)}M_1\times T_{f_2(x)}M_2$$

für alle $x\in M$.

11.5. **Kanonische Basen.**

11.5.1. **Definition.** Sei $M^n\subset\mathbb{R}^N$ eine Untermannigfaltigkeit und $(U,\varphi=(x_1,\dots,x_n))$ eine Karte um $x\in M$. Sei e_i der i-te kanonische Basisvektor in $\mathbb{R}^n$. Für $1\leqslant i\leqslant n$ bezeichnen wir

$$\frac{\partial}{\partial x_i}(x):=\frac{\partial(\varphi^{-1})}{\partial x_i}\big(\varphi(x)\big)=d(\varphi^{-1})(\varphi(x))\cdot e_i\,.$$

Nach dem Satz 11.3.2 ist $\Big(\dfrac{\partial}{\partial x_1}(x),\dots,\dfrac{\partial}{\partial x_n}(x)\Big)$ eine Basis von T_xM, genannt die ***kanonische Basis in T_xM bezüglich der Karte*** (U,φ).

Der Vektor $\frac{\partial}{\partial x_i}(x)$ hat die folgende geometrische Interpretation: $\frac{\partial}{\partial x_i}(x)$ ist der Tangentialvektor im Punkt $x\in M$, der durch die Ableitung der i-ten Koordinatenlinie $\varphi^{-1}(\varphi(x)+te_i)$ durch x in $t=0$ definiert wird.

11.5.2. **Beispiel.**
Sei $M=\mathbb{R}^2$. Wir bestimmen die kanonischen Basen der durch die euklidischen und durch die Polarkoordinaten definierten Karten:
(i) Sei $\varphi:\mathbb{R}^2\longrightarrow\mathbb{R}^2$ die durch die euklidischen Koordinaten gegebene Karte $\varphi(x):=(x_1,x_2)$. Für diese Karte gilt offensichtlich

$$\frac{\partial}{\partial x_i}(x)=e_i\quad\text{für jeden Punkt } x=(x_1,x_2)\in M.$$

(ii) Die Polarkoordinaten auf $\mathbb{R}^2$ sind gegeben durch die Parametrisierung

$$\psi:\mathbb{R}_+\times(0,2\pi)\longrightarrow U:=\mathbb{R}^2\setminus\{(x,0):x\in\mathbb{R},x\geqslant 0\},\quad \psi(r,\vartheta):=(r\cos\vartheta,r\sin\vartheta).$$

Im Punkt $x=(x_1,x_2)=\psi(r,\vartheta)$ gilt dann für die durch die Parametrisierung ψ definierte Karte (U,ψ^{-1}):

$$\begin{aligned}\frac{\partial}{\partial r}(x)&=\frac{\partial\psi}{\partial r}(r,\vartheta)=(\cos\vartheta,\sin\vartheta)=\frac{1}{r}x\quad\text{und}\\ \frac{\partial}{\partial\vartheta}(x)&=\frac{\partial\psi}{\partial\vartheta}(r,\vartheta)=(-r\sin\vartheta,r\cos\vartheta)=(-x_2,x_1).\end{aligned}$$

Wir beschreiben nun die Basisdarstellung des Differentials einer differenzierbaren Abbildung zwischen Untermannigfaltigkeiten bezüglich kanonischer Basen:

11.5.3. **Satz** (Basisdarstellung des Differentials einer glatten Abbildung)**.**
Sei $F:M_1^n\longrightarrow M_2^m$ eine glatte Abbildung zwischen Untermannigfaltigkeiten und $x\in M_1$. Seien $(U,\varphi=(x_1,\dots,x_n))$ eine Karte um x und $(W,\psi=(y_1,\dots,y_m))$ eine Karte um $F(x)\in M_2$. Die Matrix der linearen Abbildung $dF(x):T_xM_1\longrightarrow T_{F(x)}M_2$ bezüglich der kanonischen Basen $\big(\frac{\partial}{\partial x_1}(x),\dots,\frac{\partial}{\partial x_n}(x)\big)$ und $\big(\frac{\partial}{\partial y_1}(F(x)),\dots,\frac{\partial}{\partial y_m}(F(x))\big)$ ist die Jacobi-Matrix der Kartendarstellung $\psi\circ F\circ\varphi^{-1}$ von F in $\varphi(x)$, d.h.

$$\Big(dF(x)\big(\tfrac{\partial}{\partial x_1}(x)\big),\dots,dF(x)\big(\tfrac{\partial}{\partial x_n}(x)\big)\Big)=\Big(\tfrac{\partial}{\partial y_1}(F(x)),\dots,\tfrac{\partial}{\partial y_m}(F(x))\Big)\cdot J_{\psi\circ F\circ\varphi^{-1}}(\varphi(x))$$

oder

$$dF(x)\Big(\frac{\partial}{\partial x_i}(x)\Big)=\sum_{j=1}^{m}\frac{\partial(\psi\circ F\circ\varphi^{-1})_j}{\partial x_i}(\varphi(x))\cdot\frac{\partial}{\partial y_j}(F(x)).$$

Dabei ist $\frac{\partial(\psi\circ F\circ\varphi^{-1})_j}{\partial x_i}(\varphi(x))$ die i-te partielle Ableitung der j-ten Komponente der Abbildung $\psi\circ F\circ\varphi^{-1}$.
Beweis: Wegen der Kettenregel gilt

$$d(\psi\circ F\circ\varphi^{-1})(\varphi(x))\circ d\varphi(x)=d(\psi\circ F\circ\varphi^{-1}\circ\varphi)(x)=d(\psi\circ F)(x)=d\psi(F(x))\circ dF(x)),$$

d.h. das folgende Diagramm ist kommutativ:

$$\begin{tikzcd} T_xM_1 \arrow[r, "dF(x)"] \arrow[d, "d\varphi(x)"'] & T_{F(x)}M_2 \arrow[d, "d\psi(F(x))"] \\ \mathbb{R}^n \arrow[r, "d(\psi\circ F\circ\varphi^{-1})(\varphi(x))"'] & \mathbb{R}^m \end{tikzcd}$$

Die vertikalen Abbildungen sind Isomorphismen und transportieren laut Definition 11.5.1 die kanonischen Basen $\left(\frac{\partial}{\partial x_1}(x),\dots,\frac{\partial}{\partial x_n}(x)\right)$ bzw. $\left(\frac{\partial}{\partial y_1}(F(x)),\dots,\frac{\partial}{\partial y_m}(F(x))\right)$ in die kanonischen Basen von $\mathbb{R}^n$ bzw. $\mathbb{R}^m$. Es folgt, dass die Matrix von $dF(x)$ bzgl. der kanonischen Basen der Karten (U,φ) und (W,ψ) gleich der Matrix von $d(\psi\circ F\circ\varphi^{-1})(\varphi(x))$ bzgl. der kanonischen Basen von $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ ist. Die letztere ist aber genau die Jacobi-Matrix $J_{\psi\circ F\circ\varphi^{-1}}(\varphi(x))$. □

Auf analoge Weise erhält man die Beziehung zwischen den kanonischen Basen zu zwei verschiedenen Karten um einen Punkt einer Untermannigfaltigkeit:

11.5.4. **Satz** (Transformationsformel für kanonische Basen). *Sei $M^n\subset\mathbb{R}^N$ eine Untermannigfaltigkeit und $x\in M$. Seien $(U,\varphi=(x_1,\dots,x_n))$ und $(V,\psi=(y_1,\dots,y_n))$ zwei Karten um x. Dann ist die Übergangsmatrix zwischen den kanonischen Basen bzgl. (U,φ) und (V,ψ) die Jacobi-Matrix des Kartenüberganges $\psi\circ\varphi^{-1}$, d.h.*

$$\left(\frac{\partial}{\partial x_1}(x),\dots,\frac{\partial}{\partial x_n}(x)\right)=\left(\frac{\partial}{\partial y_1}(x),\dots,\frac{\partial}{\partial y_m}(x)\right)\cdot J_{\psi\circ\varphi^{-1}}(\varphi(x))$$

oder

$$\frac{\partial}{\partial x_i}(x)=\sum_{j=1}^{n}\frac{\partial(\psi\circ\varphi^{-1})_j}{\partial x_i}(\varphi(x))\cdot\frac{\partial}{\partial y_j}(x).$$

Beweis: Wende Satz 11.5.3 auf $F=\mathrm{Id}_M$ an: Dann ist $dF(x)=\mathrm{Id}_{T_xM}$. □

Aus der Algebra wissen wir, dass man jedem Vektorraum sein algebraisches Dual zuordnen kann. Dies tun wir jetzt mit den Tangentialräumen an eine Untermannigfaltigkeit. Insbesondere wollen wir die dualen Basen zu den kanonischen Basen in den Tangentialräumen beschreiben.

11.5.5. **Definition.** Sei $M^n\subset\mathbb{R}^N$ eine Untermannigfaltigkeit, $x\in M$ und T_xM der Tangentialraum an M in x. Der Vektorraum $T^*_xM:=\{L:T_xM\longrightarrow\mathbb{R} : L\,\text{linear}\}$ heißt ***Kotangentialraum*** oder auch dualer Tangentialraum an M im Punkt x.

Sei $f:M\longrightarrow\mathbb{R}$ eine differenzierbare reellwertige Abbildung und $x\in M$. Das Differential $df(x):T_xM\longrightarrow T_{f(x)}\mathbb{R}=\mathbb{R}$ von f im Punkt x ist linear, also ist $df(x)\in T^*_xM$. Insbesondere sind die Koordinatenfunktionen $x_i:U\longrightarrow\mathbb{R}$ einer Karte $(U,\varphi=(x_1,\dots,x_n))$ glatte reellwertige Abbildungen. Also ist $(dx_i)_x:T_xM\longrightarrow\mathbb{R}$ linear, und somit gilt $(dx_1)_x,\dots,(dx_n)_x\in T^*_xM$.

11.5.6. **Satz.** *Sei $(U,\varphi=(x_1,\dots,x_n))$ eine Karte auf M um x mit kanonischer Basis $(\frac{\partial}{\partial x_1}(x),\dots,\frac{\partial}{\partial x_n}(x))$ in T_xM. Dann bilden die Differentiale $(dx_1(x),\dots,dx_n(x))$ die dazu duale Basis im Kotangentialraum T^*_xM, d.h.*

$$dx_i(x)\cdot\frac{\partial}{\partial x_j}(x)=\delta_{ij}\quad\forall i,j=1,\dots,n.$$

If $f:M\to\mathbb{R}$ glatt, so hat $d(x)$ die folgende Darstellung in fer Basis $(dx_1(x),\dots,dx_n(x))$:

$$df(x)=\sum_{i=1}^{n}\frac{\partial(f\circ\varphi^{-1})}{\partial x_i}(\varphi(x))\,dx_i(x).$$

Beweis: Einsetzen in der Definition ergibt

$$\begin{aligned}dx_i(x)\cdot\frac{\partial}{\partial x_j}(x)&=dx_i(x)\cdot\frac{d}{dt}(\varphi^{-1}(\varphi(x)+te_j))|_{t=0}=\frac{d}{dt}(\varphi_i\circ\varphi^{-1})(\varphi(x)+te_j)|_{t=0}\\&=\frac{d}{dt}(\varphi(x)+te_j)_i|_{t=0}=\delta_{ij}.\end{aligned}$$

Alternativ: Es gilt $d\varphi(x)\cdot\frac{\partial}{\partial x_j}(x)=e_j$ und $d\varphi(x)=(dx_1(x),\dots,dx_n(x))$, also $dx_i(x)\cdot\frac{\partial}{\partial x_j}(x)=\delta_{ij}$ für alle $i,j=1,\dots,n$. □

Aus Satz 11.5.4 und der Transformationsformel für duale Basen (siehe Lineare Algebra) erhalten wir die folgende Transformationsformel für unsere dualen Basen:

11.5.7. **Folgerung.** *Sind $(U,\varphi=(x_1,\dots,x_n))$ und $(W,\varphi=(y_1,\dots,y_n))$ zwei Karten von M um x, so gilt*

$$dy_j(x)=\sum_{i=1}^{n}\frac{\partial(\psi\circ\varphi^{-1})_j}{\partial x_i}(\varphi(x))\,dx_i(x).$$

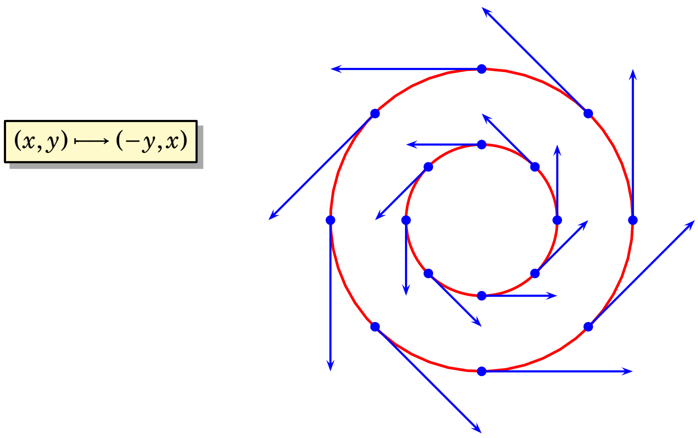
11.6. **Vektorfelder, Riemmansche Metrik, Gradient.**

11.6.1. **Definition.** Ein Vektorfeld auf einer Untermannigfaltigkeit $M^n \subset \mathbb{R}^N$ ist eine $\mathcal{C}^\infty$-Abbildung $X : M \longrightarrow \mathbb{R}^N$ so, dass $X(x) \in T_xM$ für alle $x \in M$ gilt. Die Menge aller Vektorfelder auf M wird mit $\mathfrak{X}(M)$ bezeichnet.

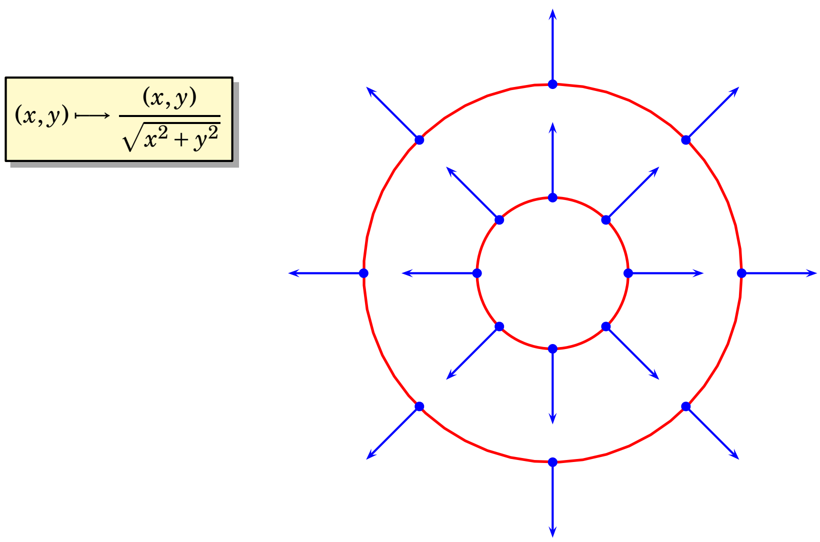
(Wir stellen uns ein Vektorfeld so vor, dass jedem Punkt ein Pfeil zugeordnet wird.)

11.6.2. **Beispiel.**

(i) Die durch $X(x,y) = (-y,x)$ definierte Abbildung $X : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ ist ein glattes Vektorfeld auf $\mathbb{R}^2$. Auch $X|_{S^1_r} : S^1_r \to \mathbb{R}^2$ ist ein glattes Vektorfeld auf S^1_r.



Die durch $X(x,y) = (x,y)/\|(x,y)\| = \dfrac{(x,y)}{\sqrt{x^2+y^2}}$ definierte Abbildung $X : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R}^2$ ist ein glattes Vektorfeld auf $\mathbb{R}^2 \setminus \{(0,0)\}$:



(ii) Die Abbildung $X : S^2 \longrightarrow \mathbb{R}^3$, definiert durch

$$X(x,y,z) := (-y,x,0),$$

ist ein Vektorfeld auf der Sphäre $S^2 = \{(x,y,z) \in \mathbb{R}^3 | x^2+y^2+z^2=1\} \subset \mathbb{R}^3$.

(iii) Sei $(U, \varphi = (x_1,\dots,x_n))$ eine Karte einer Untermannigfaltigkeit M, und bezeichne $\frac{\partial}{\partial x_i}(x) = \frac{\partial \varphi^{-1}}{\partial x_i}(\varphi(x))$ den im letzen Abschnitt definierten i-ten kanonischen Basisvektor bezüglich der Karte (U,φ) im Punkt $x \in U$. Da die Abbildungen $\frac{\partial \varphi^{-1}}{\partial x_i}$ und φ differenzierbar sind, ist die Abbildung

$$\frac{\partial}{\partial x_i} : U \subset M \longrightarrow \mathbb{R}^N,$$

die jedem Punkt $x \in U$ den kanonischen Basisvektor $\frac{\partial}{\partial x_i}(x)$ zuordnet, ein glattes Vektorfeld auf dem Kartenbereich $U \subset M$. Das n-Tupel der Abbildungen $(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n})$ heißt ***kanonisches Basisfeld bezüglich*** (U, φ).

Jedes Vektorfeld $X \in \mathfrak{X}(M)$ kann man über dem Kartenbereich U punktweise in der kanonischen Basis darstellen:

$$X(x) = \sum_{i=1}^{n} \xi_i(x) \frac{\partial}{\partial x_i}(x). \tag{11.8}$$

Die Komponenten in dieser Basisdarstellung definieren glatte Funktionen $\xi_i : U \longrightarrow \mathbb{R}$, $i = 1, \ldots, n$, auf dem Kartenbereich. Man nennt (11.8) die ***Basisdarstellung des Vektorfeldes*** X bezüglich der Karte (U, φ). Die Kartendarstellungen der Koordinatenfunktionen ξ_i, also

$$\xi_i \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^n \longrightarrow \mathbb{R},$$

heißen ***Komponenten von X bezüglich der Karte*** (U, φ). Ist z.B. $(U, \varphi = (r, \vartheta))$ die Inverse der Polarkoordinatenabbildung, so gilt $\frac{\partial}{\partial r}(x, y) = \frac{(x,y)}{\sqrt{x^2+y^2}}$ und $\frac{\partial}{\partial \vartheta}(x, y) = (-y, x)$.

11.6.3. **Definition.** Sei $M \subset \mathbb{R}^N$ eine Untermannigfaltigkeit und $g_x : T_xM \times T_xM \longrightarrow \mathbb{R}$ die Einschränkung des euklidischen Skalarproduktes, also $g_x(u, v) := \langle u, v \rangle$. Die Familie $g = \{g_x\}_{x \in M}$ dieser Skalarprodukte heißt ***induzierte Riemannsche Metrik*** auf M.

Sei $(U, \varphi = (x_1, \ldots, x_n))$ eine Karte von M um x. Wir betrachten die symmetrische positiv definite $(n \times n)$-Matrix $(g_{ij}(x))_{1 \leqslant i,j \leqslant n}$, wobei

$$\boxed{g_{ij}(x) := g_x\Big(\frac{\partial}{\partial x_i}(x), \frac{\partial}{\partial x_j}(x)\Big) = \Big\langle \frac{\partial}{\partial x_i}(x), \frac{\partial}{\partial x_j}(x) \Big\rangle.}$$

Die Funktionen $g_{ij} \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ heißen ***lokale Koeffizienten der Metrik*** g bezüglich der Karte (U, φ).

11.6.4. **Beispiel.**

(i) Die Matrix der induzierten Riemannschen Metrik in euklidischen Koordinaten: Sei $M = U$ eine offene Teilmenge des $\mathbb{R}^n$. Wir betrachten auf U die Karte, die durch die euklidischen Koordinaten gegeben wird: $\varphi(x) = (x_1, \ldots, x_n)$. Für die kanonische Basis gilt dann $\frac{\partial}{\partial x_i}(x) = e_i \in T_xM = \mathbb{R}^n$. Folglich ist

$$g_{ij}(x) = \langle e_i, e_j \rangle = \delta_{ij},$$

das heißt, in jedem Punkt x ist die Matrix der Metrik die Einheitsmatrix: $(g_{ij}(x)) = \mathrm{Id}$.

(ii) Die Matrix der induzierten Riemannschen Metrik in Polarkoordinaten: Wir betrachten Polarkoordinaten auf $\mathbb{R}^2$. Sei $\psi : \mathbb{R}_+ \times (0, 2\pi) \longrightarrow U = \mathbb{R}^2 \setminus \{(x, 0) : x \geqslant 0\} \subset \mathbb{R}^2$ die Parametrisierung durch Polarkoordinaten, $\psi(r, \vartheta) = (r\cos\vartheta, r\sin\vartheta)$. Für die Karte (U, ψ^{-1}) gilt dann in $x = \psi(r, \vartheta)$:

$$\frac{\partial}{\partial r}(x) = \frac{\partial \psi}{\partial r}(r, \vartheta) = (\cos\vartheta, \sin\vartheta) \quad \text{und} \quad \frac{\partial}{\partial \vartheta}(x) = \frac{\partial \psi}{\partial v}(r, \vartheta) = (-r\sin\vartheta, r\cos\vartheta).$$

Somit erhält man

$$\Big\langle \frac{\partial}{\partial r}(x), \frac{\partial}{\partial r}(x) \Big\rangle = 1, \quad \Big\langle \frac{\partial}{\partial r}(x), \frac{\partial}{\partial \vartheta}(x) \Big\rangle = 0 \quad \text{und} \quad \Big\langle \frac{\partial}{\partial \vartheta}(x), \frac{\partial}{\partial \vartheta}(x) \Big\rangle = r^2.$$

Die Matrix der induzierten Riemannschen Metrik in Polarkoordinaten hat folglich im Punkt $x = \psi(r, \vartheta)$ die Gestalt

$$(g_{ij}(x)) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}. \tag{11.9}$$

(iii) Die Matrix der induzierten Riemannschen Metrik auf der Sphäre S^2 in sphärischen Koordinaten (siehe (11.7)): Sei ψ die Parametrisierung durch sphärische Koordinaten,

$$\psi : (0, 2\pi) \times (-\frac{\pi}{2}, \frac{\pi}{2}) \longrightarrow S^2, \quad \psi(\varphi, \vartheta) = (\cos\varphi\cos\vartheta, \sin\varphi\cos\vartheta, \sin\vartheta). \tag{11.10}$$

Dann gilt in $x=\psi(\varphi,\vartheta)$ für die kanonischen Basisvektoren der durch diese Parametrisierung definierten Karte ψ^{-1}:

$$\begin{aligned}\frac{\partial}{\partial\varphi}(x)&=\frac{\partial\psi}{\partial\varphi}(\varphi,\vartheta)=(-\sin\varphi\cos\vartheta,\cos\varphi\cos\vartheta,0),\\ \frac{\partial}{\partial\vartheta}(x)&=\frac{\partial\psi}{\partial\vartheta}(\varphi,\vartheta)=(-\cos\varphi\sin\vartheta,-\sin\varphi\sin\vartheta,\cos\vartheta).\end{aligned}\tag{11.11}$$

Somit erhält man

$$\Big\langle\frac{\partial}{\partial\varphi}(x),\frac{\partial}{\partial\varphi}(x)\Big\rangle=\cos^2\vartheta,\quad \Big\langle\frac{\partial}{\partial\varphi}(x),\frac{\partial}{\partial\vartheta}(x)\Big\rangle=0\quad\text{und}\quad \Big\langle\frac{\partial}{\partial\vartheta}(x),\frac{\partial}{\partial\vartheta}(x)\Big\rangle=1.$$

Die Matrix der induzierten Metrik auf S^2 hat in sphärischen Koordinaten folglich im Punkt $x=\psi(\varphi,\vartheta)$ die Gestalt

$$(g_{ij}(x))=\begin{pmatrix}\cos^2\vartheta & 0\\ 0 & 1\end{pmatrix}.\tag{11.12}$$

Sei $f:M\longrightarrow\mathbb{R}$ eine glatte Funktion auf einer Untermannigfaltigkeit M. Dann ist das Differential $df(x):T_xM\longrightarrow\mathbb{R}$ eine lineare Abbildung auf dem euklidischen Vektorraum (T_xM,g_x). Wie man aus der Linearen Algebra weiß, entspricht dieser linearen Abbildung ein eindeutig bestimmter „dualer" Vektor in T_xM. Dies definiert den Gradienten von f:

11.6.5. **Definition.** Sei $M\subset\mathbb{R}^N$ eine Untermannigfaltigkeit, versehen mit der induzierten Riemannschen Metrik, und sei $f:M\longrightarrow\mathbb{R}$ eine glatte reellwertige Funktion auf M. Der ***Gradient*** von f ist das Vektorfeld

$$\operatorname{grad}f:M\longrightarrow\mathbb{R}^N$$

auf M, das jedem Punkt $x\in M$ den zur Linearform $df(x)$ dualen Vektor $\operatorname{grad}f(x)\in T_xM$ zuordnet, d.h.:

$$\boxed{\big\langle\operatorname{grad}f(x),v\big\rangle=df(x)\cdot v\quad\text{für alle } v\in T_xM.}$$

Wie im Fall einer Funktion in $\mathbb{R}^n$, weist $\operatorname{grad}f(x)$ in Richtung des stärksten Anstiegs der Funktion f auf M (falls $\operatorname{grad}f(x)\neq0$), siehe Bemerkung nach Satz 9.2.9. Der folgende Satz gibt die lokale Darstellung des Gradienten in der kanonischen Basis einer Karte an und zeigt insbesondere die Glattheit der Abbildung $\operatorname{grad}f$.

11.6.6. **Satz.** *Sei $f\in\mathcal{C}^\infty(M^n)$ und $(U,\varphi=(x_1,\dots,x_n))$ eine Karte von M. Sei weiterhin $f\circ\varphi^{-1}:\varphi(U)\longrightarrow\mathbb{R}$ die Kartendarstellung von f, und sei $(g^{ij}(x))$ die inverse Matrix zu $(g_{ij}(x))=\Big(g_x(\frac{\partial}{\partial x_i}(x),\frac{\partial}{\partial x_j}(x))\Big)$. Dann gilt über dem Kartenbereich U:*

$$\boxed{\operatorname{grad}f(x)=\sum_{i,j=1}^n g^{ij}(x)\frac{\partial(f\circ\varphi^{-1})}{\partial x_i}(\varphi(x))\cdot\frac{\partial}{\partial x_j}(x).}\tag{11.13}$$

11.6.7. **Beispiel.**

(i) Der Gradient einer Funktion in euklidischen Koordinaten: Sei U eine offene Teilmenge des $\mathbb{R}^n$ und $f:U\to\mathbb{R}$ eine glatte Abbildung. Wir bestimmen die Darstellung des Gradienten von f in der durch die euklidischen Koordinaten gegebenen Karte $\varphi(x)=(x_1,\dots,x_n)$. Die Matrix der induzierten Riemannschen Metrik bezüglich dieser Karte ist die Einheitsmatrix: $(g_{ij}(x))=\mathrm{Id}$, also auch $(g^{ij}(x))=\mathrm{Id}$, und $\frac{\partial}{\partial x_i}(x)=e_i$. Für den Gradienten gilt also in euklidischen Koordinaten:

$$\operatorname{grad}f(x)=\sum_{i=1}^n\frac{\partial f}{\partial x_i}(x)\frac{\partial}{\partial x_i}(x)=\sum_{i=1}^n\frac{\partial f}{\partial x_i}(x)e_i=\Big(\frac{\partial f}{\partial x_1}(x),\dots,\frac{\partial f}{\partial x_n}(x)\Big).$$

Dies stimmt mit unserer früheren Definition überein. Man kann der Gradient als ein vektorwertiger Differential auffassen:

$$\operatorname{grad}=\nabla=\Big(\frac{\partial}{\partial x_1},\dots,\frac{\partial}{\partial x_n}\Big).$$

Die Notation ∇ ist gerne benuzt in der Physik.

(ii) Der Gradient in Polarkoordinaten: Sei $\psi:\mathbb{R}_+\times(0,2\pi)\longrightarrow U=\mathbb{R}^2\setminus\{(x,0):x\geqslant0\}\subset\mathbb{R}^2$ die Parametrisierung durch Polarkoordinaten, $\psi(r,\vartheta)=(r\cos\vartheta,r\sin\vartheta)$. Sei $f:U\to\mathbb{R}$ glatt. Sei $x=\psi(r,\vartheta)$, und bezeichne $f(r,\vartheta):=(f\circ\psi)(r,\vartheta)=f(r\cos\vartheta,r\sin\vartheta)$ die Kartendarstellung von f. Dann gilt nach (11.9)

$$(g_{ij}(x))=\begin{pmatrix}1&0\\0&r^2\end{pmatrix},\qquad(g^{ij}(x))=\begin{pmatrix}1&0\\0&\frac{1}{r^2}\end{pmatrix}$$

und nach (11.13)

$$\operatorname{grad} f(x) = \frac{\partial f}{\partial r}(r,\vartheta)\frac{\partial}{\partial r}(x) + \frac{1}{r^2}\frac{\partial f}{\partial \vartheta}(r,\vartheta)\frac{\partial}{\partial \vartheta}(x).$$

(ii) Der Gradient einer Funktion $f : S^2 \longrightarrow \mathbb{R}$ in sphärischen Koordinaten: Es sei $\psi(\varphi,\vartheta) := (\cos\varphi\cos\vartheta, \sin\varphi\cos\vartheta, \sin\vartheta)$ die lokale Parametrisierung der Sphäre durch sphärische Koordinaten. Sei $x = \psi(\varphi,\vartheta)$, und bezeichne $f(\varphi,\vartheta) := (f\circ\psi)(\varphi,\vartheta) = f(\cos\varphi\cos\vartheta, \sin\varphi\cos\vartheta, \sin\vartheta)$ die Kartendarstellung von f in diesen Koordinaten. Für die Matrix der induzierten Riemannschen Metrik in der durch die sphärischen Koordinaten definierten Karte gilt (siehe (11.12))

$$(g_{ij}(x)) = \begin{pmatrix} \cos^2\vartheta & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad (g^{ij}(x)) = \begin{pmatrix} \dfrac{1}{\cos^2\vartheta} & 0 \\ 0 & 1 \end{pmatrix}.$$

Für den Gradienten gilt folglich

$$\operatorname{grad} f(x) = \frac{1}{\cos^2\vartheta}\frac{\partial f}{\partial \varphi}(\varphi,\vartheta)\cdot\frac{\partial}{\partial \varphi} + \frac{\partial f}{\partial \vartheta}(\varphi,\vartheta)\cdot\frac{\partial}{\partial \vartheta}(x).$$

Siehe auch Aufgabe 11.10.1.

11.7. **Lokale Diffeomorphismen. Fundamentalsatz der Algebra.**

11.7.1. **Definition.** Seien M_1, M_2 Untermannigfaltigkeiten. Eine Abbildung $f : M_1 \longrightarrow M_2$ heißt:

(1) ***Diffeomorphismus***, falls f bijektiv ist und f, f^{-1} glatt sind;

(2) ***lokaler Diffeomorphismus*** in $a \in M_1$, falls es offene Umgebungen U von a und V von $f(a)$ gibt so, dass $f : U \longrightarrow V$ ein Diffeomorphismus ist.

11.7.2. **Bemerkung.**

- Ist f ein Diffeomorphismus, so folgt, dass $df(a)$ invertierbar ist mit der Inversen $\big(df(a)\big)^{-1} = d(f^{-1})(b)$, wobei $f(a) = b$.
- Die Verkettung von Diffeomorphismen ist ein Diffeomorphismus.

11.7.3. **Satz.** *Sei $f : M_1 \longrightarrow M_2$ glatt, und $df(a) : T_aM_1 \longrightarrow T_{f(a)}M_2$ sei ein Isomorphismus. Dann ist f ein lokaler Diffeomorphismus in a.*

11.7.4. **Definition.** Sei $f : M_1 \longrightarrow M_2$ glatt.

$$\begin{aligned}
x \in M_1 \text{ heißt regulärer Punkt} &:\Longleftrightarrow df(x) : T_xM_1 \longrightarrow T_{f(x)}M_2 \text{ ist surjektiv},\\
x \in M_1 \text{ heißt kritischer Punkt} &:\Longleftrightarrow df(x) : T_xM_1 \longrightarrow T_{f(x)}M_2 \text{ ist nicht surjektiv},\\
y \in M_2 \text{ heißt regulärer Wert} &:\Longleftrightarrow \text{jedes } x \in f^{-1}(y) \text{ ist ein regulärer Punkt},\\
y \in M_2 \text{ heißt kritischer Wert} &:\Longleftrightarrow \text{mindestens ein } x \in f^{-1}(y) \text{ ist ein kritischer Punkt}.
\end{aligned}$$

11.7.5. **Bemerkung.** Ist $\dim M_1 = \dim M_2$, so ist $x \in M_1$ ein regulärer Punkt genau dann, wenn f ein lokaler Diffeomorphismus in x ist.

11.7.6. **Satz.** *Seien M_1^n, M_2^n zwei Untermannigfaltigkeiten von gleicher Dimension, und sei M_1 kompakt. Sei $f : M_1 \longrightarrow M_2$ glatt und $y \in M_2$ so, dass f für jedes $x \in f^{-1}(y)$ ein lokaler Diffeomorphismus in x ist. Dann gilt:*

(1) *$f^{-1}(y)$ ist endlich.*

(2) *Es gibt eine offene Umgebung V von y so, dass für alle $z \in V$ gilt: $|f^{-1}(z)| = |f^{-1}(y)|$.*

Beweis: 1. Fall: $f^{-1}(y) = \emptyset$, d.h. $y \in M_2 \setminus f(M_1)$, so reicht es (ii) zu beweisen: da M_1 kompakt ist, so ist $f(M_1)$ kompakt also abgeschlossen in M_2. Somit ist $V = M_2 \setminus f(M_1)$ offen und erfüllt (ii).

2.Fall: $f^{-1}(y) \neq \emptyset$. Nun ist $f^{-1}(y)$ kompakt, da $f^{-1}(y) \subset M_1$ abgeschlossen ist und M_1 kompakt. Sei $x \in f^{-1}(y)$. Dann existiert eine offene Umgebung $U_x \subset M_1$ von x, so dass $f|_{U_x} : U_x \to f(U_x)$ ein Diffeomorphismus ist. Insbesondere ist $f|_{U_x}$ injectiv, also $U_x \cap f^{-1}(y) = \{x\}$. Die offene Überdeckung $\{U_x\}_{x\in f^{-1}(y)}$ der kompakten Menge $f^{-1}(y)$ besitzt eine offene Teilüberdeckung $\{U_{x_i}\}_{1\leqslant i\leqslant p}$. Da jedes U_{x_i} genau einen Punkt aus $f^{-1}(y)$ enthält, ist $f^{-1}(y)$ endlich.

Wir können O.B.d.A. annehmen, dass $\{U_{x_i}\}_{1\le i\le p}$ paarweise disjunkt sind. Setze

$$V = f(U_{x_1})\cap\ldots\cap f(U_{x_p})\setminus f(\overbrace{M_1\setminus\underbrace{(U_{x_1}\cup\ldots\cup U_{x_p})}_{\text{offen}}}^{\text{kompakt}})$$

$M_1\setminus(U_{x_1}\cup\ldots\cup U_{x_p})$ ist abgeschlossen in kompaktem Raum M_1, also auch kompakt. Deshalb ist $f(M_1\setminus(U_{x_1}\cup\ldots\cup U_{x_p}))$ kompakt in M_2, also abgeschlossen. Folglich ist V eine offene Umgebung von y. Wegen der Konstruktion von V gilt für $z\in V$: $z\in f(U_{x_i})$, also $|f^{-1}(z)\cap U_{x_i}|=1$ und außerdem $f^{-1}(z)\subset U_{x_1}\cup\ldots\cup U_{x_p}$. Daraus folgt $|f^{-1}(z)|=|f^{-1}(y)|$. □

Wir beweisen nun den Fundamentalsatz der Algebra und folgen dabei dem Buch von Milnor [16]. Es gibt viele Beweise dieses Satzes. Der wohl einfachste Beweis wurde 1814 von Argand gegeben und benutzt nur den Satz über Maximum und Minimum für stetige Funktionen auf einer kompakten Menge sowie elementare Eigenschaften der komplexen Zahlen; siehe das lesenswerte Kapitel 4 von R. Remmert in [6]. Andere beliebte Beweise sind in der Funktionentheorie formuliert: Sie benutzen das Maximumprinzip oder den Offenheitssatz. Der Beweis von Milnor ist topologisch und hat den Vorteil, dass mit wenigen Mitteln viel über die Struktur des Polynoms als Abbildung gesagt werden kann.

11.7.7. **Satz** (Fundamentalsatz der Algebra). *Jedes nicht konstante komplexe Polynom hat im Körper $\mathbb{C}$ wenigstens eine Nullstelle.*

Beweis: Sei $P:\mathbb{C}\to\mathbb{C}, P(z)=a_nz^n+a_{n-1}z^{n-1}+\ldots+a_0, n\ge1, a_n\ne0$. Wir definieren eine Abbildung $\widetilde{P}:S^2\to S^2$ durch

$$\widetilde{P}(x)=\begin{cases}P_N{}^{-1}\circ P\circ P_N(x), & x\in S^2\setminus N\\ N, & x=N.\end{cases}$$

Wegen $\lim_{z\to\infty}P(z)=\infty$ ist $\widetilde{P}$ ist stetig und sogar C^∞. Ist $x\in S^2$ kritisch für $\widetilde{P}$ so ist $x=N$ oder ist $z=P_N(x)$ kritisch für P. Nun ist z kritisch für P genau dann, wenn $P'(z)=0$ (wegen (9.7)). Folglich hat $\widetilde{P}$ nur endlich viele kritische Werte. Sei F die Menge der kritischen Werte. Dann ist $S^2\setminus F$ zusammenhängend. Wir wenden den Satz 11.7.6 für $\widetilde{P}:S^2\to S^2$ und $y\in S^2\setminus F$ und erhalten, dass die Abbildung

$$S^2\setminus F\ni y\xrightarrow{g}|\widetilde{P}^{-1}(y)|\in\mathbb{N}$$

lokal konstant ist. Sie ist also stetig und ihr Bild ist eine zusammenhängende Teilmenge von $\mathbb{N}$, d.h. ihr Bild besteht aus einem Punkt. Mit anderen Worten ist g konstant. Wäre diese Konstante Null, dann hätte P nur kritische Werte, d.h. $P'(z)=0$ für alle $z\in\mathbb{C}$, was ein Widerspruch zu P nicht konstant ist. Also $|\widetilde{P}^{-1}(y)|\geqslant1$ für alle $y\in S^2\setminus F$. Daraus folgt, dass $\widetilde{P}$ surjektiv ist und erst recht auch $P:\mathbb{C}\to\mathbb{C}$ ist surjektiv. Insbesondere gilt $P^{-1}(0)\ne\emptyset$. □

In Satz 11.7.6 haben wir eigentlich mehr über die Struktur der Abbildung f gezeigt. Dies wollen wir nun in einer Definition festlegen.

11.7.8. **Definition.** Seien M_1,M_2 Untermannigfaltigkeiten. Eine Abbildung $f\in\mathcal{C}^\infty(M_1,M_2)$ heißt ***Überlagerung***, wenn gilt: (1) f ist surjektiv, und (2) für jedes $y\in M_2$ gibt es eine Umgebung V von y, so dass $f^{-1}(V)=\bigcup_{i\in I}U_i$ mit $U_i\subset M_1$ offen, $U_i\cap U_j=\emptyset$ für $i\ne j$, und $f|_{U_i}:U_i\longrightarrow V$ für jedes i ein Diffeomorphismus ist. Die Mengen U_i heißen Blätter der Überlagerung über V.

Die Indexmenge I kann von y abhängen. Die Kardinalzahl von I ist die Kardinalzahl von $f^{-1}(y)$: Es gilt $|I|=|f^{-1}(y)|$. Aus der Definition folgt $|f^{-1}(y')|=|f^{-1}(y)|$ für alle $y'\in V$; also ist die Funktion $Y\ni y\mapsto|f^{-1}(y)|\in\mathbb{N}_0$ lokal konstant. Wenn Y zusammenhängend ist, so ist diese Funktion konstant (siehe den Beweis des Satzes 11.7.7) und heißt Blätterzahl der Überlagerung. Ist die Blätterzahl gleich $k\in\mathbb{N}$, so sprechen wir von einer k-fachen Überlagerung.

Beispiele:

(i) $P:\mathbb{R}_+\times\mathbb{R}\longrightarrow\mathbb{R}^2\setminus\{0\}$, $P(r,\vartheta)=(r\cos\vartheta,r\sin\vartheta)$. Für $(x,y)=(re^{i\varphi_0})\in\mathbb{R}^2\setminus\{0\}$ setze $V=\mathbb{R}^2\setminus\{te^{i(\varphi_0+\pi)}:t\geqslant0\}$. Dann ist $f^{-1}(V)=\bigcup_{k\in\mathbb{Z}}\mathbb{R}_+\times(\varphi_0+k\pi,\varphi_0+(k+2)\pi)$.

(ii) $\psi:\mathbb{R}\longrightarrow S^1$, $\psi(t)=e^{it}$.

(iii) $\psi:\mathbb{R}^2\longrightarrow S^1\times S^1$, $\psi(s,t)=(e^{is},e^{it})$.

(iv) Unter der Annahme des Satzes 11.7.6 ist f eine Überlagerung. Aber nicht jeder lokale Diffeomorphismus ist eine Überlagerung: $\psi:(0,3\pi)\longrightarrow S^1$ mit $\psi(t)=e^{it}$ ist ein lokaler Diffeomorphismus, aber keine Überlagerung.

(v) Sei $P : \mathbb{C} \longrightarrow \mathbb{C}$ ein Polynom und $\widetilde{P} \in \mathcal{C}^\infty(S^2, S^2)$ seine Fortsetzung auf S^2 (die Riemannsche Zahlenkugel), dann existiert eine endliche Menge $F \subset S^2$, so dass $\widetilde{P} : S^2 \smallsetminus F \longrightarrow S^2 \smallsetminus \widetilde{P}(F)$ eine n-fache Überlagerung ist (wobei $n = \operatorname{grad} P$).

11.8. **Zerlegung der Eins.** Zerlegungen der Eins sind ein wichtiges Hilfsmittel in Analysis und Geometrie. Sie erlauben die „Lokalisierung" globaler Objekte auf einer Mannigfaltigkeit und umgekehrt das „Verschmelzen" lokaler Objekte zu globalen Objekten.

Motivation: Verkleben von Funktionen. Seien $a < b < c < d$ in $\mathbb{R}$ und $f : (a,c) \longrightarrow \mathbb{R}$, $g : (b,d) \longrightarrow \mathbb{R}$ glatt. Gibt es $h : (a,d) \longrightarrow \mathbb{R}$ glatt so, dass $h|_{(a,b)} = f$ und $h|_{(c,d)} = g$? Die Lösung benutzt eine Zerlegung der Eins. Wir konstruieren $\lambda, \eta \in \mathcal{C}^\infty(\mathbb{R})$ mit $\lambda = 1$ auf (a,b), $\lambda = 0$ auf (c,d), $\eta = 1$ auf (c,d) und $\eta = 0$ auf (a,b) und $\lambda + \eta = 1$. Dann erfüllt $h = \lambda f + \eta g$ die oben genannten Bedingungen.

11.8.1. **Lemma.** *Die Funktion* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$,

$$f(x) = \begin{cases} \exp\left(\dfrac{1}{\|x\|^2 - 1}\right), & \|x\| < 1, \\ 0, & \|x\| \geqslant 1 \end{cases}$$

ist $\mathcal{C}^\infty$.

Beweis: Die Funktion

$$g : \mathbb{R} \longrightarrow \mathbb{R}, \quad g(t) = \begin{cases} e^{-1/t}, & t > 0 \\ 0, & t \leqslant 0 \end{cases}$$

ist $\mathcal{C}^\infty$ (siehe (5.7) und Beispiel 6.5.2 (4)). Dann ist $f(x) = g(1 - \|x\|^2)$, also ist f ist eine Verkettung von $\mathcal{C}^\infty$-Abbildungen. □

11.8.2. **Definition.** Sei M ein topologischer Raum und $f : M \longrightarrow \mathbb{R}$. Der ***Träger*** von f ist die abgeschlossene Menge $\operatorname{supp} f := \overline{\{x \in M : f(x) \neq 0\}}$.
Sei M eine Untermannigfaltigkeit. Die Menge der glatten Funktionen auf M mit kompaktem Träger wird mit $\mathcal{C}_0^\infty(M)$ bezeichnet. $\mathcal{C}_0^\infty(M)$ ist ein Untervektorraum von $\mathcal{C}^\infty(M)$.

11.8.3. **Definition.** Sei M ein topologischer Raum. Eine Familie $(M_\alpha)_{\alpha \in A}$ von Teilmengen $M_\alpha \subset M$ heißt ***lokal endlich***, wenn zu jedem $x \in M$ eine Umgebung $V_x \subset M$ existiert so, dass $V_x \cap M_\alpha \neq \emptyset$ für nur endlich viele $\alpha \in A$. Eine Familie $(f_\alpha)_{\alpha \in A}$ von reellen Funktionen heißt ***lokal endlich***, wenn $(\operatorname{supp} f_\alpha)_{\alpha \in A}$ lokal endlich ist.

11.8.4. **Satz** (Zerlegung der Eins)**.**

Sei $M \subset \mathbb{R}^N$ eine Untermannigfaltigkeit und $M = \bigcup_{\alpha \in A} M_\alpha$ eine offene Überdeckung. Dann existiert eine Folge $\{\lambda_k\}_{k \in \mathbb{N}}$ von glatten Funktionen $\lambda_k \in \mathcal{C}^\infty(M)$, so dass gilt:

(1) $0 \leqslant \lambda_k \leqslant 1$ *für alle* $k \in \mathbb{N}$,
(2) $(\lambda_k)_{k \in \mathbb{N}}$ *ist lokal endlich,*
(3) *für jedes* $k \in \mathbb{N}$ *gilt:* $\operatorname{supp} \lambda_k$ *ist kompakt, und es gibt* $\alpha \in A$ *mit* $\operatorname{supp} \lambda_k \subset M_\alpha$,
(4) $\sum_{k \in \mathbb{N}} \lambda_k(x) = 1$ *für alle* $x \in M$.

Beweis: Für alle $\alpha \in A$ existiert eine offene Teilmenge $W_\alpha \subset \mathbb{R}^N$ mit $M_\alpha = M \cap W_\alpha$. Sei $W = \bigcup_{\alpha \in A} W_\alpha$. Dann ist W offen in $\mathbb{R}^N$. Wir konstruieren eine Folge $(B_k)_{k \in \mathbb{N}}$ von abgeschlossenen Kugeln so, dass:

(i) $\forall k\ \exists \alpha$: $B_k \subset W_\alpha$,
(ii) $(B_k)_{k \in \mathbb{N}}$ ist lokal endlich,
(iii) $W = \bigcup_{k \in \mathbb{N}} \mathring{B}_k$.

Zunächst konstruieren wir eine kompakte Ausschöpfung von W. Darunter versteht man eine Folge $K_1 \subset K_2 \subset \ldots \subset W$ kompakter Teilmengen mit $K_i \subset \mathring{K}_{i+1}$, $W = \bigcup_{i \in \mathbb{N}} K_i$. Setze z.B.

$$K_i = \{x \in \mathbb{R}^n : \|x\| \leqslant i, d(x, \mathbb{R}^N \smallsetminus W) \geqslant \tfrac{1}{i}\}.$$

Setze $K_0 = \emptyset$. Wir konstruieren nun induktiv eine Folge $(B_k)_{k \in \mathbb{N}}$, so dass (i) gilt und außerdem

$$(\star) \qquad K_1 \subset \bigcup_{1 \leqslant k \leqslant k_1} \mathring{B}_k \subset \mathring{K}_2$$

und

$$(\star\star) \qquad K_{i+1} \smallsetminus \mathring{K}_i \subset \bigcup_{k_i < k \leqslant k_{i+1}} \mathring{B}_k \subset \mathring{K}_{i+2} \smallsetminus K_{i-1}$$

für jedes $i \geqslant 1$. Für jedes $x \in K_1$ wähle $\alpha \in A$ mit $x \in W_\alpha$ und eine abgeschlossene Kugel B_x in $\mathbb{R}^N$ mit Mittelpunkt in x, so dass $B_x \subset W_\alpha \cap \mathring{K}_2$. Die Familie $(\mathring{B}_x)_{x \in K_1}$ ist eine offene Überdeckung von K_1. Wähle eine endliche Teilüberdeckung $\{B_1, \dots, B_{k_1}\}$; für diese gilt dann $(\star)$. Durch Induktion zeigen wir ähnlich $(\star\star)$. Daraus folgen (i)–(iii).

Sei $\eta_k \in \mathcal{C}^\infty(W)$ mit $\eta_k > 0$ auf B_k und $\eta_k = 0$ auf $\mathbb{R}^N \setminus B_k$ (Lemma 11.8.1). Weil $(B_k)_k$ lokal endlich ist, ist auch $(\eta_k)_k$ lokal endlich.
Setze $\eta = \sum_{k \in \mathbb{N}} \eta_k$. Dann ist $\eta \in \mathcal{C}^\infty(W)$ und $\eta > 0$ auf W. Sei

$$\lambda_k = \frac{\eta_k}{\eta}\Big|_M .$$

Dann erfüllt die Familie $(\lambda_k)_k$ die Behauptungen des Satzes. □

11.8.5. **Definition.** Die Familie $(\lambda_k)_{k \in \mathbb{N}}$ mit den Eigenschaften (1)–(4) aus Satz 11.8.4 heißt eine der Überdeckung $(M_\alpha)_{\alpha \in A}$ untergeordnete ***Zerlegung der Eins***.

In Satz 11.8.4 ist die Indexmenge der Familie $(\lambda_k)_{k \in \mathbb{N}}$ im Allgemeinen verschieden von der Indexmenge der Familie $(M_\alpha)_{\alpha \in A}$. Wir können eine Zerlegung der Eins mit derselben Indexmenge konstruieren, verlieren dabei aber die Eigenschaft, dass $\operatorname{supp} \lambda_k$ kompakt ist:

11.8.6. **Satz.** *Sei $(M_\alpha)_{\alpha \in A}$ eine offene Überdeckung einer Untermannigfaltigkeit M. Dann gibt es eine Familie $(\lambda_\alpha)_{\alpha \in A}$ von glatten Funktionen, so dass gilt:*

(1) *$(\lambda_\alpha)_{\alpha \in A}$ ist lokal endlich,*
(2) *$\operatorname{supp} \lambda_\alpha \subset M_\alpha$,*
(3) *$0 \leqslant \lambda_\alpha \leqslant 1$,*
(4) *$\sum_{\alpha \in A} \lambda_\alpha = 1$ auf M.*

11.8.7. **Definition.** Eine Familie $(\lambda_\alpha)_{\alpha \in A}$ mit den Eigenschaften (1)–(4) aus Satz 11.8.6 heißt eine der Überdeckung $(M_\alpha)_\alpha$ untergeordnete ***Zerlegung der Eins*** mit derselben Indexmenge.

11.8.8. **Folgerung** (Urysohn). *Sei $M \subset \mathbb{R}^N$ eine Untermannigfaltigkeit. Seien $A, B \subset M$ abgeschlossene Teilmengen mit $A \cap B = \emptyset$. Dann gibt es $f \in C^\infty(M)$, so dass $f|_A = 0$ und $f|_B = 1$.*

Wenn die Menge B kompakt ist, so kann man f mit kompaktem Träger finden (siehe Aufgabe 11.10.2). Die Zerlegung der Eins hat zahlreiche Anwendungen, z. B. im Beweis des Satzes von Stokes. (Siehe Aufgabe 11.10.3 für die Fortsetzung von Funktionen definiert auf einer Untermannigfaltigkeit zu einer Umgebung in $\mathbb{R}^N$).

11.9. **Glatt berandete Teilmengen einer Untermannigfaltigkeit.** Glatt berandete Teilmengen sind die Integrationsbereiche, die wir im Satz von Stokes zugrunde legen werden.

11.9.1. **Definition.** Sei $M \subset \mathbb{R}^N$ eine n-dimensionale Untermannigfaltigkeit und $D \subset M$. Wir bezeichnen mit ∂D den Rand von D bezüglich der Teilraumtopologie von M.
Ein Punkt $a \in \partial D$ heißt ***regulärer Punkt*** , falls es eine Umgebung U von a in M und eine glatte Funktion $q : U \longrightarrow \mathbb{R}$ gibt mit $dq(x) \neq 0$ für alle $x \in U$, so dass gilt: $U \cap D = \{x \in U : q(x) \leqslant 0\}$. Die Funktion q heißt dann eine ***lokal beschreibende Funktion*** für D um a. Falls U eine Umgebung von ∂D, so heisst q global beschreibende Funktion oder einfach beschreibende Funktion für D.
Ein Punkt $a \in \partial D$ heißt ***singulärer Punkt*** , falls a kein regulärer Punkt ist.
$D \subset M$ heißt ***glatt berandet*** , wenn jeder Randpunkt ein regulärer Punkt ist.
Gilt $\partial D = \emptyset$ (z.B. wenn $D = M$) so ist D glatt berandet.

Beispiele:
(i) $\mathbb{R}^n_- = \{x \in \mathbb{R}^n : x_n \leqslant 0\}$ ist glatt berandet (eine lokal beschreibende Funktion ist $q : \mathbb{R}^n \to \mathbb{R}$, $q(x) = x_n$) und $\partial D = \mathbb{R}^{n-1} \times \{0\}$.
(ii) $D = B_R(x_0) \subset \mathbb{R}^n$ ist glatt berandet. Sei $0 < \varepsilon < R$ und $U_\varepsilon = \{x \in \mathbb{R}^n : \|x - x_0\| > \varepsilon\}$. Eine (global) beschreibende Funktion ist $q : U_\varepsilon \longrightarrow \mathbb{R}$, $q(x) = \|x - x_0\|^2 - R^2$. Der Rand ist die Sphäre $\partial B_R(x_0) = S_R(x_0) = \{x : \|x - x_0\| = R\}$. Es gilt $dq(x) = 2\sum_{j=1}^n (x_j - x_{0j}) dx_j$ und $dq(x) \neq 0$ für $x \in U$, da $x_0 \notin U$.
(iii) $D = \{x \in \mathbb{R}^n : r \leq \|x\| \leq R\}$, wobei $0 < r < R$, ist glatt berandet. Der Rand hat zwei Komponenten $\partial D = S_r(x_0) \cup S_R(x_0)$. Sei $0 < \varepsilon < r$ und $U_\varepsilon = \{x \in \mathbb{R}^n : \|x - x_0\| > \varepsilon\}$. Eine lokal beschreibende Funktion für die Komponente $S_r(x_0)$ ist $q : U_\varepsilon \longrightarrow \mathbb{R}$, $q(x) = r^2 - \|x - x_0\|^2$ und eine lokal beschreibende Funktion für

die Komponente $S_R(x_0)$ ist $q : U_\varepsilon \longrightarrow \mathbb{R}$, $q(x) = \|x - x_0\|^2 - R^2$.

(iii) $D = S^n \cap \{x_{n+1} \leqslant 0\}$ ist glatt berandet. Eine beschreibende Funktion ist $q : S^n \setminus \{N, S\} \longrightarrow \mathbb{R}$, $q(x) = x_{n+1}$. Dabei sind N und S der Nord-und Südpol, wobei dq verschwindet. Der Rand ist $\partial D = S^{n-1} \times \{0\}$.

(iv) Sei $D = [a,b] \times [c,d] \subset \mathbb{R}^2$. Die Punkte $(a,c),(a,d),(b,c),(b,d)$ sind singuläre Punkte von ∂D, und alle anderen Punkten von ∂D sind reguläre Punkte.

(v) Eine Menge A mit der Eigenschaft, dass ∂A eine Untermannigfaltigkeit ist, ist nicht unbedingt eine glatt berandete Teilmenge. Beispiel: Für $A = \overline{B}_1(0) \cup \{(x,y) : x = 2\}$ gilt $\partial A = S^1 \cup \{(x,y) : x = 2\}$, die Definition regulärer Randpunkte ist aber für die Punkte $\{(x,y) : x = 2\}$ nicht erfüllt.

Die Definition besagt intuitiv, dass "das Innere $\mathring{D}$ liegt lokal genau auf einer Seite des Randes ∂D". Dies wird im folgenden Lemma präzisiert. Nämlich, $U \setminus \partial D$ ist die disjunkte Vereinigung von zwei offenen, nicht-leeren Mengen, $\mathring{D} \cap U$ und $(M \setminus \overline{D}) \cap U$. Die zwei Mengen kann man nach dem Vorzeichen von q unterscheiden: $\mathring{D} \cap U$ ist die Menge wo q negativ ist und $(M \setminus \overline{D}) \cap U$ die Menge wo q positiv ist.

11.9.2. **Lemma.** *Sei $a \in \partial D$ ein regulärer Punkt und $q : U \longrightarrow \mathbb{R}$ eine beschreibende Funktion. Dann gilt:*

$$\overline{D} \cap U = D \cap U\,, \quad \mathring{D} \cap U = \{x \in U : q(x) < 0\}\,, \quad \partial D \cap U = \{x \in U : q(x) = 0\}\,.$$

Ist D glatt berandet, so ist D abgeschlossen und $D \subset \overline{\mathring{D}}$ (= Abschluss von $\mathring{D}$ in M).

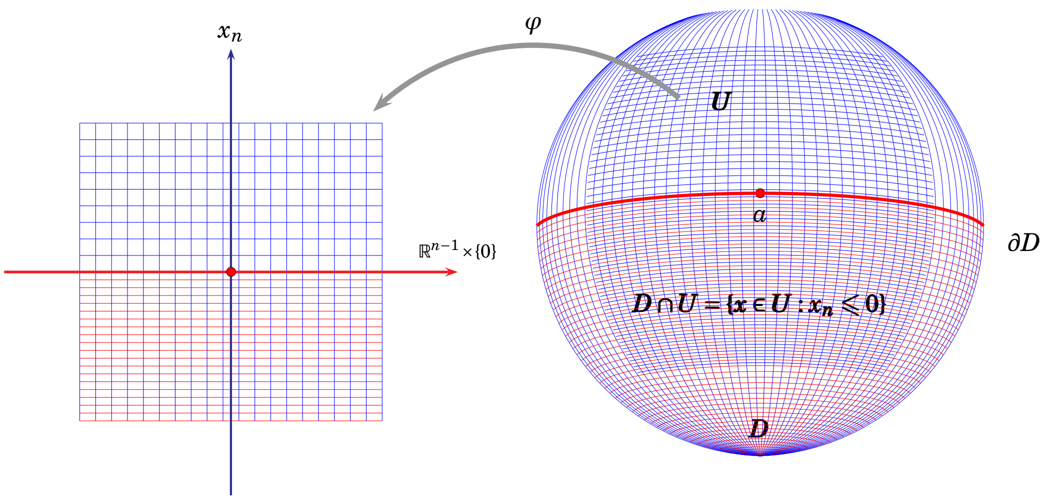
11.9.3. **Satz.** *Sei $M^n \subset \mathbb{R}^N$ eine Untermannigfaltigkeit, $D \subset M$ mit $\partial D \neq \emptyset$. Die folgenden Aussagen sind äquivalent:*

(1) *D ist glatt berandet.*
(2) *Für jedes $a \in \partial D$ existiert eine Karte $(U, \varphi = (x_1, \dots, x_n))$ auf M mit $D \cap U = \{x : x_n \leqslant 0\}$. (Dann heißt (U,φ) eine D-**angepasste Karte** oder D-**Karte**.)*
(3) *D ist abgeschlossen, $D \subset \overline{\mathring{D}}$ und ∂D ist eine $(n-1)$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^N$.*

Beweis: (1) $\Rightarrow$ (2): Sei $a \in \partial D$ und $q : U \longrightarrow \mathbb{R}$ eine beschreibende Funktion, wobei U eine Umgebung von a in M ist. O.B.d.A. können wir $U \subset \mathbb{R}^n$ annehmen (wir können U so verkleinern, dass $U \subset \widetilde{U}$, wobei $(\widetilde{U}, \widetilde{\varphi})$ eine Karte ist und $\widetilde{\varphi}(D) \subset \widetilde{\varphi}(U) \subset \mathbb{R}^n$ die beschreibende Funktion $q \circ \widetilde{\varphi}^{-1} : \widetilde{\varphi}(U) \longrightarrow \mathbb{R}$ um $\widetilde{\varphi}(a)$ hat). Wegen $dq(a) \neq 0$ ist q eine Submersion in a. Der Satz über die lokale Struktur einer Submersion besagt, dass es eine Umgebung $V \subset U$ von a und einen Diffeomorphismus $\varphi : V \longrightarrow W^n_\varepsilon(0)$ gibt so, dass das folgende Diagramm kommutativ ist:

$$\begin{tikzcd} V \arrow[rr, "q"] \arrow[dr, "\varphi"'] & & (-\varepsilon, \varepsilon) \\ & W^n_\varepsilon(0) \arrow[ur, "pr_n"'] & \end{tikzcd}$$

Folglich ist (V, φ) eine Karte um a mit $\varphi_n = q$, also $V \cap D = \{x \in V : q(x) \leqslant 0\} = \{x \in V : \varphi_n(x) \leqslant 0\}$.



D-Karte

(2) ⇒ (3): Sei (U,φ) eine D-Karte. Dann gilt $\partial D \cap U = \varphi^{-1}(\mathbb{R}^{n-1} \times \{0\} \cap \varphi(U))$. Sei $U' \subset \mathbb{R}^{n-1}$ das Bild der Projektion von $\varphi(U) \cap (\mathbb{R}^{n-1} \times \{0\})$ auf die ersten $n-1$ Koordinaten. Dann ist $U' \ni (x_1,\dots,x_{n-1}) \mapsto \varphi^{-1}(x_1,\dots,x_{n-1},0)$ eine Parametrisierung von $\partial D \cap U$, und somit ist $\partial D \cap U$ eine $(n-1)$-dimensionale Untermannigfaltigkeit.

(3) ⇒ (1): Sei $a \in \partial D$. Da das Problem lokal ist, können wir o.B.d.A. annehmen, dass $M = \mathbb{R}^n$ ist. Der Satz über die lokale Struktur einer Immersion impliziert, dass es $\varepsilon > 0$, eine offene Umgebung W von a und einen Diffeomorphismus $\Phi : W^n_\varepsilon(0) \longrightarrow W$ gibt mit $\Psi(W^{n-1}_\varepsilon(0) \times \{0\}) = W \cap \partial D$. Sei $W' = W \cap \mathring{D}$ und $W'' = W \setminus (W \cap D)$. Wegen $a \in D \subset \overline{\mathring{D}}$ und $a \notin \mathring{D}$ sind W' und W'' nicht leer. Außerdem sind W' und W'' offen, disjunkt, und $W \setminus \partial D = W' \cup W''$. Andererseits hat $W^n_\varepsilon(0) \setminus W^{n-1}_\varepsilon(0) \times \{0\}$ zwei Zusammenhangskomponenten, nämlich $W^n_\varepsilon(0) \cap \{x_n > 0\}$ und $W^n_\varepsilon(0) \cap \{x_n < 0\}$. Da Ψ ein Homöomorphismus ist, bildet Ψ Zusammenhangskomponenten auf Zusammenhangskomponenten ab. Daher muss nach eventueller Ersetzung des Basisvektors e_n durch $-e_n$ gelten:

$$\Psi(W^n_\varepsilon(0) \cap \{x_n < 0\}) = W \setminus \mathring{D} \quad \text{und} \quad \Psi(W^n_\varepsilon(0) \cap \{x_n > 0\}) = W \setminus (W \cap D).$$

Sei q die Submersion $\mathrm{pr}_n \circ \Psi^{-1} : W \longrightarrow \mathbb{R}$. Dann gilt $W \cap D = \{x \in W : q(x) \leqslant 0\}$. □

Sei D eine glatt berandete Teilmenge einer Untermannigfaltigkeit M^n. Sei $a \in \partial D$. Dann ist $T_a(\partial D)$ ein $(n-1)$-dimensionaler Untervektorraum von T_aM:

$$T_a(\partial D) = \{v \in \mathbb{R}^N : \exists \gamma : (-\varepsilon,\varepsilon) \longrightarrow \partial D \;\text{ mit } \gamma(0) = a \text{ und } \gamma'(0) = v\}.$$

Ist $q : U \longrightarrow \mathbb{R}$ eine lokal beschreibende Funktion um a, so gilt:

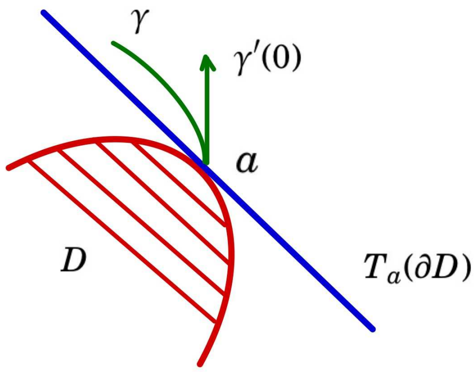
$$T_a(\partial D) = \operatorname{Ker} dq(a) = \{v \in T_aM : dq(a)\cdot v = 0\}.$$

Ist $(U, \varphi = (x_1,\dots,x_n))$ eine D-Karte, so gilt:

$$T_a(\partial D) = \left\langle \left\{ \frac{\partial}{\partial x_1}(a), \dots, \frac{\partial}{\partial x_{n-1}}(a) \right\} \right\rangle.$$

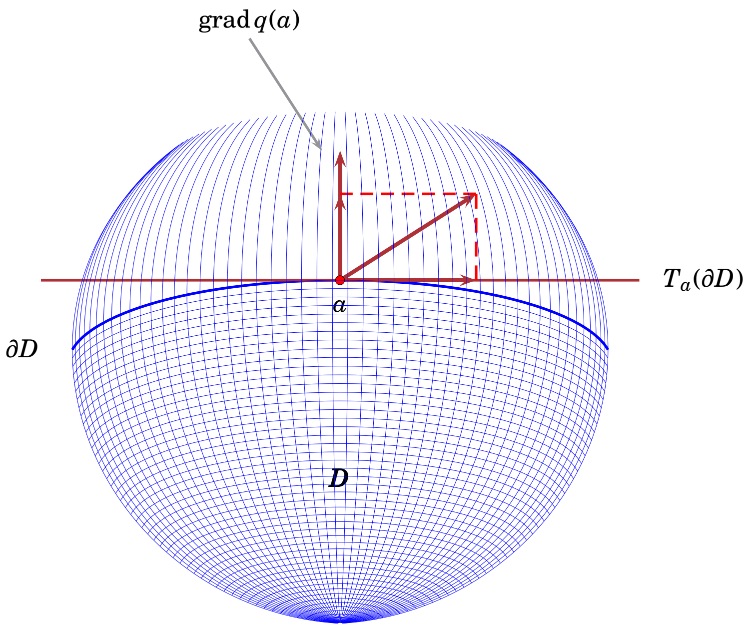
11.9.4. **Definition.** Sei $a \in \partial D$. Ein Vektor $v \in T_aM$ ist ein ***äußerer Vektor*** zu D in a, wenn $v \notin T_a(\partial D)$ ist und es eine glatte Kurve $\gamma : [0,\varepsilon) \longrightarrow M \setminus \mathring{D}$ mit $\gamma(0) = a$ und $v = \gamma'(0)$ gibt. Wir bezeichnen die Menge der äußeren Vektoren zu D in a mit $T^+_a D$.

Beispiel: Sei $M = \mathbb{R}^n$, $D = \mathbb{R}^n_-$, $a \in \partial\mathbb{R}^n_- = \mathbb{R}^{n-1} \times \{0\}$. Dann ist $T^+_a(\partial D) = \{v \in \mathbb{R}^n : v_n > 0\}$.



11.9.5. **Satz.** *Die folgenden Aussagen sind äquivalent zueinander:*

(i) $v \in T^+_a(D)$,
(ii) *für jede lokal beschreibende Funktion* $q : U \longrightarrow \mathbb{R}$ *um* a *gilt* $dq(a)\cdot v > 0$,
(iii) *für jede lokal beschreibende Funktion* $q : U \longrightarrow \mathbb{R}$ *um* a *gilt* $\langle v, \operatorname{grad} q(a)\rangle > 0$, *also* $v = v' + \lambda \operatorname{grad} q(a)$ *mit* $v' \in T_a(\partial D)$ *und einer Zahl* $\lambda > 0$,
(iv) *für jede* D*-Karte* $(U, \varphi = (x_1,\dots,x_n))$ *um* a *gilt* $v = \sum \lambda_i \frac{\partial}{\partial x_i}(a)$ *mit* $\lambda_n > 0$.



Zur Zerlegung
$v = v' + \lambda \operatorname{grad} q(a) \in T_a(\partial D) \oplus \mathbb{R} \operatorname{grad} q(a)$

11.9.6. **Satz.** *Es gibt eine eindeutig bestimmte glatte Abbildung $\nu : \partial D \longrightarrow \mathbb{R}^N$ so, dass für alle $x \in \partial D$ gilt:*

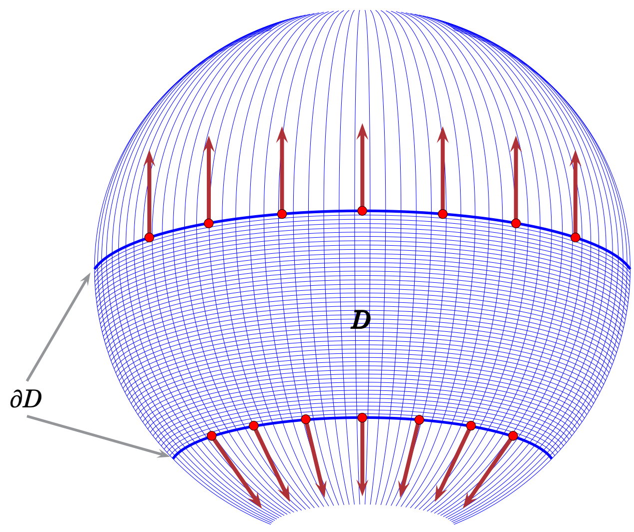
(1) *$\nu(x)$ ist ein äußerer Vektor,*
(2) *$\nu(x) \perp T_x(\partial D)$,*
(3) *$\|\nu(x)\| = 1$.*

Ist $q : U \longrightarrow \mathbb{R}$ eine lokal beschreibende Funktion um x, so gilt

$$\boxed{\nu(x) = \frac{\operatorname{grad} q(x)}{\|\operatorname{grad} q(x)\|}\,.}$$

11.9.7. **Definition.** Der Vektor $\nu(x)$ heißt ***äußerer Einheitsnormalenvektor*** zu D in $x \in \partial D$. Die Abbildung $\nu : \partial D \longrightarrow S^{N-1} \subset \mathbb{R}^N$ heißt ***äußeres Einheitsnormalenfeld*** an D oder ***Gauß-Abbildung***.

Als Beispiel betrachten wir $M = \mathbb{R}^n$, $D = B_R(0)$. Dann ist $\partial D = S_R(0) = \{x : \|x\| = R\}$ und $\nu(x) = \frac{x}{R}$ für alle $x \in S_R(0)$. Für $D = \{x : r \le \|x\| \le R\}$, wobei $0 < r < R$, so ist $\partial D = S_R(0) \cup S_r(0)$ und $\nu(x) = \frac{x}{R}$ für alle $x \in S_R(0)$, $\nu(x) = -\frac{x}{r}$ für alle $x \in S_r(0)$.



Äußeres Einheitsnormalenfeld

11.10. **Übungen.**

11.10.1. ***Aufgabe***. (a) Sei M eine Untermannigfaltigkeit des $\mathbb{R}^N$, und sei $F : \mathbb{R}^N \to \mathbb{R}$ glatt. Sei $f := F|_M : M \to \mathbb{R}$. Zeigen Sie, dass $\operatorname{grad} f(x)$ die orthogonale Projektion von $\operatorname{grad} F(x)$ auf T_xM ist.

(b) Sei $f : S^2 \to \mathbb{R}$ mit $f(x) = x_3$ die Höhenfunktion auf der Sphäre S^2. Berechnen Sie $\operatorname{grad} f$ (mit Hilfe von (a) oder (11.13)).

11.10.2. ***Aufgabe***. Sei M eine Untermannigfaltigkeit. Seien $A, B \subset M$, wobei A abgeschlossen, B kompakt und $A \cap B = \emptyset$ sei. Zeigen Sie, dass es eine Funktion $f \in C^\infty(M)$ mit kompaktem Träger gibt, so dass $f|_A = 0$ und $f|_B = 1$.

11.10.3. ***Aufgabe***. Sei $f \in C^\infty(M, \mathbb{R}^k)$. Zeigen Sie, dass es eine offene Menge $W \supset M$ in $\mathbb{R}^N$ und eine Funktion $F \in C^\infty(W, \mathbb{R}^k)$ gibt, so dass $F|_M = f$. Zeigen Sie außerdem: Falls M abgeschlossen ist, so kann $W = \mathbb{R}^N$ gewählt werden.

12. Differentialformen

12.1. Alternierende Multilinearformen.
Sei V ein $\mathbb{R}$-Vektorraum, $\dim V = n$.

12.1.1. Definition. Für $p \in \mathbb{N}$ setzen wir $V^p = \underbrace{V \times \ldots \times V}_{p-\text{mal}}$. Eine Funktion $f : V^p \longrightarrow \mathbb{R}$ heißt ***alternierende Multilinearform*** oder ***alternierende p-Form*** auf V, falls gilt:

(1) f ist multilinear, und

(2) $f(v_{\sigma(1)},\ldots,v_{\sigma(p)}) = \varepsilon(\sigma) f(v_1,\ldots,v_p)$ für alle $\sigma \in S_p$ und alle $(v_1,\ldots,v_p) \in V^p$ (wobei $\varepsilon(\sigma)$ das Vorzeichen der Permutation σ bezeichnet).

Eine alternierende 0-Form auf V ist einfach eine reelle Zahl. Für $p \in \mathbb{N}_0$ bezeichnen wir mit $\bigwedge^p V^*$ den Vektorraum der alternierenden p-Formen auf V.

Beispiele:

(i) $\bigwedge^0 V^* := \mathbb{R}$,

(ii) $\bigwedge^1 V^* = V^* = \operatorname{Hom}_{\mathbb{R}}(V,\mathbb{R})$.

(iii) Sei $(e_1,\ldots,e_n)$ eine Basis in V. Für ein n-Tupel $(v_1,\ldots,v_n)$ definiere die zugehörigen Koeffizienten v_i^j bezüglich der genannten Basis durch

$$v_i = \sum_{j=1}^n v_i^j e_i = (e_1 \ldots e_n) \begin{pmatrix} v_i^1 \\ \vdots \\ v_i^n \end{pmatrix}.$$

Dann ist

$$\omega : V^n \longrightarrow \mathbb{R}, \quad \omega(v_1,\ldots,v_n) = \det(v_i^j) \tag{12.1}$$

eine n-Form auf V, wobei (v_i^j) die Matrix mit Spalten die Vektoren $v_1,\ldots,v_n$ ist.

12.1.2. Definition (Dachprodukt)**.** Das ***Dachprodukt*** $f \wedge g \in \bigwedge^{p+q} V^*$ der Formen $f \in \bigwedge^p V^*$ und $g \in \bigwedge^q V^*$ ist definiert durch:

$$(f \wedge g)(v_1,\ldots,v_p,v_{p+1},\ldots,v_{p+q}) = \frac{1}{p!q!} \sum_{\sigma \in S_{p+q}} \varepsilon(\sigma) f(v_{\sigma(1)},\ldots,v_{\sigma(p)}) \cdot g(v_{\sigma(p+1)},\ldots,v_{\sigma(p+q)}).$$

12.1.3. Satz.

(i) *$\bigwedge V^* = \oplus_{p\geqslant 0} \bigwedge^p V^*$ (äußere oder Graßmann-Algebra) ist eine assoziative $\mathbb{R}$-Algebra mit Eins bezüglich $(+,\wedge,$Skalarmultiplikation$)$. $\bigwedge V^*$ ist antikommutativ, d.h. für $\omega \in \bigwedge^p V^*$ und $\eta \in \bigwedge^q V^*$ gilt $\eta \wedge \omega = (-1)^{pq} \omega \wedge \eta$.*

(ii) *Seien $f_1,\ldots,f_p \in \bigwedge^1 V^*$. Dann gilt:*

$$(f_1 \wedge \ldots \wedge f_p)(v_1,\ldots,v_p) = \det\big(f_i(v_j)\big) = \begin{vmatrix} f_1(v_1) & f_1(v_2) & \ldots & f_1(v_p) \\ f_2(v_1) & f_2(v_2) & \ldots & f_2(v_p) \\ \vdots & \vdots & \ldots & \vdots \\ f_p(v_1) & f_p(v_2) & \ldots & f_p(v_p) \end{vmatrix}. \tag{12.2}$$

Insbesondere: $f_{\sigma(1)} \wedge \ldots \wedge f_{\sigma(p)} = \varepsilon(\sigma) f_1 \wedge \ldots \wedge f_p$.

(iii) *Sei $(e_1,\ldots,e_n)$ eine Basis von V und $(e_1^*,\ldots,e_n^*)$ die dazu duale Basis von V^*, d.h. $e_i^*(e_j) = \delta_{ij}$. Dann ist $(e_{i_1}^* \wedge \ldots \wedge e_{i_p}^*)_{1\leqslant i_1<\ldots<i_p\leqslant n}$ eine Basis von $\bigwedge^p V^*$. Ein Element $\omega \in \bigwedge^p V^*$ hat die Darstellung*

$$\omega = \sum_{1\leqslant i_1<\ldots<i_p\leqslant n} \omega(e_{i_1},\ldots,e_{i_p}) e_{i_1}^* \wedge \ldots \wedge e_{i_p}^*. \tag{12.3}$$

Insbesondere gilt $\dim \bigwedge^p V^* = \begin{cases} \binom{n}{p}, & p \leqslant n, \\ 0, & p > n. \end{cases}$

Für $p = n$ erhalten wir, dass $\dim \bigwedge^n V^* = 1$ und eine Basis ist gegeben durch $e_1^* \wedge \ldots \wedge e_n^*$. Dies ist die n-Form (12.1).

Für einen Multiindex $I = (i_1,\ldots,i_p)$ aus $\{1,\ldots,n\}$ (d.h. mit $i_1,\ldots,i_p \in \{1,\ldots,n\}$) setzen wir $e_I^* = e_{i_1}^* \wedge \ldots \wedge e_{i_p}^*$. Ist der Multiindex leer, so vereinbaren wir $e_I^* = 1$. Die Länge des Multiindex I ist p; wir schreiben dafür $|I| = p$. Dann wird (12.3) so geschrieben

$$\omega = {\sum}'_{|I|=p} \omega_I e_I^*,$$

wobei $\omega_I = \omega(e_{i_1},\dots,e_{i_p})$, und das Symbol $\sum'_{|I|=p}$ bedeutet, dass nur über streng aufsteigende Multiindizes der Länge p summiert wird.

Der Fall der Formen vom maximaler Grad (d.h. $n = \dim V$) ist sehr wichtig und wir formulieren deshalb die folgende

12.1.4. **Folgerung.** *Es gilt* $\dim \bigwedge^n V^* = 1$. *Ist* $B = (b_1,\dots,b_n)$ *eine Basis von* V*, so ist* $(b_1^* \wedge \dots \wedge b_n^*)$ *eine Basis von* $\bigwedge^n V^*$ *und für* $\omega \in \bigwedge^n V^*$ *gilt*

$$\omega = \omega(b_1,\dots,b_n)\, b_1^* \wedge \dots \wedge b_n^* \tag{12.4}$$

Ist $C = (c_1,\dots,c_n)$ *eine weitere Basis von* V*, so ist*

$$c_1^* \wedge \dots \wedge c_n^* = \det M_C^B\, b_1^* \wedge \dots \wedge b_n^* \tag{12.5}$$

wobei M_C^B *die Basiswechselmatrix von der Basis* C *zur Basis* B *ist.*

Es sei erinnert, dass die Basiswechselmatrix von der Basis B zur Basis C ist definiert durch

$$M_C^B := \big(a_{ij}\big)_{1\leqslant i,j\leqslant n} := \big(c_j^*(b_i)\big)_{1\leqslant i,j\leqslant n} \in M_{n\times n}(\mathbb{R}), \quad \text{wobei } b_i = \sum_{j=1}^n a_{ji}\, c_j = \sum_{j=1}^n c_j^*(b_i)\, c_j\,. \tag{12.6}$$

d.h., M_C^B besitzt als Spalten die Koordinatenvektoren von $b_1,\dots,b_n$ bzgl. der Basis C. Fassen wir B und C als formale Zeilenvektoren so gilt

$$(b_1,\dots,b_n) = (c_1,\dots,c_n)\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad \text{kurz } B = CM_C^B \tag{12.7}$$

Beweis der Folgerung 12.1.4: (12.4) ist nichts anders als (12.3) für $p = n$. Wenden wir nun (12.4) für $\omega = c_1^* \wedge \dots \wedge c_n^*$, so erhalten wir $\omega = \alpha\, b_1^* \wedge \dots \wedge b_n^*$ wobei $\alpha = c_1^* \wedge \dots \wedge c_n^*(b_1,\dots,b_n) = \det\big(c_j^*(b_i)\big) = \det M_C^B$. □

12.1.5. **Definition.** Sei $v \in V$ und $\omega \in \bigwedge^p V^*$. Die Form $\iota_v\omega = v \lrcorner \omega \in \bigwedge^{p-1} V^*$ definiert für $w_1,\dots,w_{p-1} \in V$ durch

$$(\iota_v\omega)(w_1,\dots,w_{p-1}) := (v \lrcorner \omega)(w_1,\dots,w_{p-1}) := \omega(v,w_1,\dots,w_{p-1})$$

heißt ***inneres Produkt von*** ω ***mit dem Vektor*** v.

Zum Beispiel, für $v = \sum v_j e_j$ gilt

$$\begin{aligned}(v \lrcorner e_1^* \wedge \dots \wedge e_n^*)(w_1,\dots,w_{n-1}) &= (e_1^* \wedge \dots \wedge e_n^*)(v,w_1,\dots,w_{n-1}) = \\ &= \det(v,v_1,\dots,v_{n-1}) = \sum_{j=1}^n (-1)^{j+1} A_j v_j\end{aligned}$$

wobei $(-1)^{j+1}A_j$ die Kofaktoren (algebraische Komplemente) von v_j in der Matrix $A = (v,w_1,\dots,w_{n-1})$ sind (Eintwicklungssatz nach der ersten Spalte), siehe [12, S. 106-107]. Aber A_j ist die Determinante der Matrix, die aus A durch Streichen der ersten Spalte und j-ten Zeile entsteht, also die Determinante der Matrix, die aus $(w_1,\dots,w_{n-1})$ durch Streichen der j-ten Zeile entsteht. Es folgt nach (12.2) dass

$$A_j = e_1^* \wedge \dots \wedge \widehat{e_j^*} \wedge \dots \wedge e_n^*(w_1,\dots,w_{n-1})$$

wobei ein Hut auf e_j^* bedeutet, dass e_j^* weggelassen ist. Somit

$$v \lrcorner e_1^* \wedge \dots \wedge e_n^* = \sum_{j=1}^n (-1)^{j+1} v_j\, e_1^* \wedge \dots \wedge \widehat{e_j^*} \wedge \dots \wedge e_n^*\,.$$

12.1.6. **Definition** (Zurückholen von Formen)**.** Ist $A: V \longrightarrow W$ eine lineare Abbildung, so definieren wir $A^*: \bigwedge^p W^* \longrightarrow \bigwedge^p V^*$ durch $(A^*f)(v_1,\dots,v_p) = f(Av_1,\dots,Av_p)$. Offenbar ist tatsächlich $A^*f \in \bigwedge^p V^*$, und A^* ist eine lineare Abbildung. Für $p = 0$ ist $A^*f = f$, für $p = 1$ ist $A^*f = f \circ A$.

Ist $B: W \longrightarrow U$ linear, so gilt $(B\circ A)^* = A^* \circ B^*$:

$$\begin{tikzcd} V \arrow[rr, "A"] \arrow[dr, "B\circ A"'] & & W \arrow[dl, "B"] \\ & U & \end{tikzcd} \qquad \begin{tikzcd} \bigwedge^p U^* \arrow[rr, "B^*"] \arrow[dr, "(B\circ A)^* = A^*\circ B^*"'] & & \bigwedge^p W^* \arrow[dl, "A^*"] \\ & \bigwedge^p V^* & \end{tikzcd}\,.$$

Offensichtlich ist $\mathrm{Id}^* = \mathrm{Id}$, also ist A^* ein Isomorphismus, wenn A ein Isomorphismus ist.

12.1.7. **Satz.** *Sei $A : V \longrightarrow V$ linear und $\omega \in \bigwedge^n V^*$. Dann gilt:*

$$A^*\omega = (\det A)\cdot\omega .$$

12.2. **Orientierung. Volumenelement.**

12.2.1. **Definition.**
(i) Sei V ein $\mathbb{R}$-Vektorraum mit $\dim V = n$. Auf $\bigwedge^n V^* \setminus \{0\}$ führen wir eine Äquivalenzrelation ein: Es sei $\omega \sim \omega'$, wenn ein $\lambda > 0$ existiert mit $\omega = \lambda\omega'$. Wegen $\bigwedge^n V^* \cong \mathbb{R}$ gibt es genau zwei Äquivalenzklassen. Jede davon heißt eine ***Orientierung*** von V. Ein ***orientierter Vektorraum*** ist ein Vektorraum $(V, [\omega])$ zusammen mit einer fest gewählten Orientierung. Falls $V = \{0\}$, dann gilt nach Vereinbarung $\bigwedge^n V^* = \mathbb{R}$. Eine Orientierung auf $V = \{0\}$ ist einfach eine der Halbgeraden $\mathbb{R}_+$, $\mathbb{R}_-$, entspricht also der Wahl einer der Zahlen ± 1.
(ii) Eine Basis $B = (e_1, \dots, e_n)$ (die Anordnung der Elemente der Basis ist nun wichtig!) in einem orientierten Vektorraum $(V, [\omega])$ heißt ***positiv orientiert*** (bzw. ***negativ orientiert***), falls $\omega(e_1, \dots, e_n) > 0$ (bzw. < 0). Das Vorzeichen $\varepsilon(B)$ einer Basis $B = (e_1, \dots, e_n)$ ist definiert durch $\varepsilon(B) = \operatorname{sgn}\omega(e_1, \dots, e_n) \in \{\pm 1\}$. Zwei Basen B, B_1 heißen gleich orientiert, wenn $\varepsilon(B) = \varepsilon(B_1)$.

12.2.2. **Satz.** *Jede Basis $B = (e_1, \dots, e_n)$ definiert eine Orientierung von V so, dass B positiv orientiert ist, nämlich die Orientierung $[e_1^* \wedge \dots \wedge e_n^*]$. Zwei Basen B, B_1 sind gleich orientiert genau dann, wenn* $\det M^B_{B_1} > 0$.

12.2.3. **Bemerkung.** Wir arbeiten oft mit Orthonormalbasen. Ist $B = (b_1, \dots, b_n)$ eine beliebige positiv orientierte Basis, so erhält man durch das Gram-Schmidtsche Orthonormalisierungsverfahren eine positiv orientierte Orthonormalbasis $B' = (b'_1, \dots, b'_n)$. Das folgt aus der Tatsache, dass die Übergangsmatrix eine Dreiecksmatrix mit positiven Diagonaleinträge ist.

12.2.4. **Definition.** Sei $(V, [\omega], \langle\cdot,\cdot\rangle)$ ein orientierter Skalarproduktraum mit $\dim V = n$. Das ***Volumenelement*** ω_V ist die eindeutig bestimmte n-Form, die auf jeder positiv orientierten Orthonormalbasis den Wert 1 annimmt.

12.2.5. **Satz.** *Das Volumenelement ist wohldefiniert. Ist $B = (e_1, \dots, e_n)$ eine positiv orientierte Orthonormalbasis, so gilt $\omega_V = e_1^* \wedge \dots \wedge e_n^*$.*

Beweis: Ist $C = (c_1, \dots, e_n)$ eine positiv orientierte Orthonormalbasis, so gilt nach (12.4), (12.5) $e_1^* \wedge \dots \wedge e_n^*(c_1, \dots, c_n) = \det M^C_B$. Aber M^C_B ist eine orthogonale Matrix (als Basiswechselmatrix zwischen zwei orthonormalen Basen), hat also Determinante ± 1. Weil beide Basen B, C positiv orientiert sind, muss jedoch $\det M^C_B > 0$ sein, somit $\det M^C_B = 1$. Dies beweist die Existenz der Volumenform. Die Eindeutigkeit folgt leicht. □

12.2.6. **Definition.** Sei $V = \mathbb{R}^n$ und $B = (e_1, \dots, e_n)$ die Standardbasis. Die ***Standardorientierung*** des $\mathbb{R}^n$ ist gegeben durch $e_1^* \wedge \dots \wedge e_n^*$. Das Volumenelement von $\mathbb{R}^n$ ist dann $\omega_{\mathbb{R}^n}$ mit $\omega_{\mathbb{R}^n}(v_1, \dots, v_n) = \det(v_1, \dots, v_n)$.

Im $\mathbb{R}^3$ fixiert man die Orientierung durch die sogenannte Rechte-Hand-Regel. Danach ist ein eine Basis $\{u, v, w\}$ positiv orientiert wenn sich die Vektoren mit Daumen, Zeigefinger und Mittelfinger der rechten Hand in der angegebenen Reihenfolge zur Deckung bringen lassen. Andernfalls ist die Basis negativ orientiert.
Wir wollen nun die Beziehung zwischen dem Volumenelement und Volumen erklären. Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Skalarproduktraum, $\dim_{\mathbb{R}} V = n$, und $v_1, \dots, v_m \in V$, $m \leqslant n$. Die Determinante

$$G(v_1, \dots, v_m) := \det\big(\langle v_i, v_j\rangle\big) = \det\begin{pmatrix} \langle v_1,v_1\rangle & \langle v_1,v_2\rangle & \dots & \langle v_1,v_m\rangle \\ \langle v_2,v_1\rangle & \langle v_2,v_2\rangle & \dots & \langle v_2,v_m\rangle \\ \dots & \dots & \dots & \dots \\ \langle v_m,v_1\rangle & \langle v_m,v_2\rangle & \dots & \langle v_m,v_m\rangle \end{pmatrix} \tag{12.8}$$

heißt die ***Gramsche Determinante*** der Vektoren $v_1, \dots, v_m$.

12.2.7. **Lemma.** *Es gilt $G(v_1, \dots, v_m) \geqslant 0$ für alle $v_1, \dots, v_m \in V$ und $v_1, \dots, v_m$ sind genau dann linear unabhängig, wenn $G(v_1, \dots, v_k) > 0$.*

Beweis: Sei $B=(b_1,\dots,b_n)$ eine Orthonormalbasis von V und sei $v_i=\sum_{j=1}^n a_{ji}b_j$. Sei $A=(a_{ij})\in M_{n\times m}$ die Matrix mit Spalten die Koordinatenvektoren von $v_1,\dots,v_m$ bzgl. der Basis B. Dann gilt

$$\big(\langle v_i,v_j\rangle\big)=\Big(\sum_{k=1}^n a_{ki}a_{kj}\Big)=A^TA \tag{12.9}$$

also

$$G(v_1,\dots,v_m)=\det\big(\langle v_i,v_j\rangle\big)=\det(A^TA). \tag{12.10}$$

Es ist leicht zu sehen, dass für $A\in M_{n\times m}$ gilt: A^TA ist positiv semidefinit und positiv definit genau dann, wenn die Spalten von A linear unabhängig sind. Folglich, $\det(A^TA)\geqslant 0$ und $\det(A^TA)>0$ genau dann, wenn die Spalten von A linear unabhängig sind. □

Für $v_1,\dots,v_m\in V$, $m\leqslant n$ nennt man

$$P(v_1,\dots,v_m):=\Big\{\sum_{j=1}^m \alpha_jv_j\in V:0\leqslant\alpha_1,\dots,\alpha_m\leqslant 1\Big\}. \tag{12.11}$$

das von $v_1,\dots,v_m$ aufgespannte ***Parallelotop*** (oder *Parallelepiped*). Wenn die Vektoren $v_1,\dots,v_m$ linear unabhängig sind, nennt man (12.11) ein m-dimensionales Parallelotop.

Das ***m-dimensionale Volumen*** des Parallelotops (12.11) sollte folgenden geometrischen Vorstellungen erfüllen:

$$\mathrm{vol}_1 P(v_1):=\|v_1\| \tag{12.12}$$

(das 1-Volumen einer Strecke ist ihre Länge) und

$$\mathrm{vol}_m P(v_1,\dots,v_m,v_{m+1})=d(v_m,\mathrm{Span}(v_1,\dots,v_m))\,\mathrm{vol}_m P(v_1,\dots,v_m) \tag{12.13}$$

wobei $d(v_{m+1},\mathrm{Span}(v_1,\dots,v_m))$ die Abstand des Vektors v_{m+1} von dem durch $v_1,\dots,v_m$ erzeugten Unterraum $\mathrm{Span}(v_1,\dots,v_m)$ ist.

Nehmen wir an, dass $v_1,\dots,v_{m+1}$ linear unabhängig sind. Sei $L:=\mathrm{Span}(v_1,\dots,v_m)$. Wir bestimmen $d(v_{m+1},L)=\|v_{m+1}-v'_{m+1}\|$, wobei $v'_{m+1}\in L$ die orthogonale Projection von v_{m+1} auf L ist. Sei $v'_{m+1}=\sum_{j=1}^m\lambda_jv_j$. Weil $v'_{m+1}-v_{m+1}\perp L$, folgt $\langle\sum_{j=1}^m\lambda_jv_j-v_{m+1},v_k\rangle=0$, $k=1,\dots,m$ also

$$\sum_{j=1}^m\lambda_j\langle v_j,v_k\rangle=\langle v_{m+1},v_k\rangle,\quad k=1,\dots,m. \tag{12.14}$$

Die Determinante des Systems (12.14) ist $G(v_1,\dots,v_m)>0$, also hat das System eine eindeutig bestimmte Lösung $(\lambda_1,\dots,\lambda_m)$. Wegen des Satzes von Pythagoras gilt

$$d(v_{m+1},L)^2=\|v_{m+1}-v'_{m+1}\|^2=\|v_{m+1}\|^2-\|v'_{m+1}\|^2.$$

Außerdem

$$\|v'_{m+1}\|^2=\langle v'_{m+1},v'_{m+1}\rangle=\sum_j\lambda_j\Big(\sum_i\lambda_i\langle v_i,v_j\rangle\Big)=\sum_i\lambda_j\langle v_{m+1},v_j\rangle.$$

also

$$\sum_i\lambda_i\langle v_{m+1},v_i\rangle=\|v_{m+1}\|^2-d(v_{m+1},L)^2 \tag{12.15}$$

Wegen (12.14), (12.15) sind die Spalten der Matrix

$$B=\begin{pmatrix}\langle v_1,v_1\rangle & \langle v_1,v_2\rangle & \dots & \langle v_1,v_m\rangle & \langle v_1,v_{m+1}\rangle\\ \langle v_2,v_1\rangle & \langle v_2,v_2\rangle & \dots & \langle v_2,v_m\rangle & \langle v_1,v_{m+1}\rangle\\ \dots & \dots & \dots & \dots & \dots\\ \langle v_m,v_1\rangle & \langle v_m,v_2\rangle & \dots & \langle v_m,v_m\rangle & \langle v_m,v_{m+1}\rangle\\ \langle v_{m+1},v_1\rangle & \langle v_{m+1},v_2\rangle & \dots & \langle v_{m+1},v_m\rangle & \|v_{m+1}\|^2-d(v_{m+1},L)^2\end{pmatrix} \tag{12.16}$$

linear abhängig (nichtriviale Linearkombination mit Koeffizienten $(\lambda_1,\dots,\lambda_m,-1)$). Daher, durch Entwicklung nach der letzten Spalte

$$0=\det B=G(v_1,\dots,v_m,v_{m+1})-G(v_1,\dots,v_m)d(v_{m+1},L)^2$$

Die Gramsche Determinante erfüllt also die Bedingungen:

$$\sqrt{G(v_1)}=\|v_1\| \tag{12.17}$$

und

$$\sqrt{G(v_1,\dots,v_m,v_{m+1})}=\sqrt{G(v_1,\dots,v_m)}\,d(v_{m+1},L) \tag{12.18}$$

Ein Vergleich zwischen (12.12), (12.13) und (12.17)(12.18) führt zur folgenden

12.2.8. **Definition.** Das ***m-dimensionale Volumen*** des von $v_1,\dots,v_m$ aufgespannten Parallelotops ist durch

$$\operatorname{vol}_m P(v_1,\dots,v_m) := \sqrt{G(v_1,\dots,v_m)} \tag{12.19}$$

erklärt.

Wenn $v_1,\dots,v_m$ linear abhängig sind, ist die rechte Seite Null (Lemma 12.2.7), also wird auch $\operatorname{vol}_m P(v_1,\dots,v_m) = 0$ definiert, was auch die geometrische Vorstellung des m-dimensionalen Volumens entspricht ($v_1,\dots,v_m$ spannen ein Parallelotop vom Dimension $< m$ auf).

Aus (12.10) und (12.19) erhalten wir eine weitere nützliche Formel:

12.2.9. **Satz.** *(i) Ist $B = (b_1,\dots,b_n)$ eine Orthonormalbasis von V so gilt*

$$\operatorname{vol}_m P(v_1,\dots,v_m) = \sqrt{\det(A^TA)} \quad \textit{wobei } A = \big(b_j^*(v_i)\big) \in M_{n\times m}, \tag{12.20}$$

d.h. $A \in M_{n\times m}$ besitzt als Spalten die Koordinatenvektoren von $v_1,\dots,v_m$ bzgl. der Orthonormalbasis B.

(ii) Sei W ein Skalarproduktraum, $\dim W = m$ und sei $L : W \to V$ eine lineare Abbildung und sei A ihre Darstellungsmatrix bzgl. Orthonormalbasen von W und V. Dann gilt

$$\operatorname{vol}_m L(P) = \sqrt{\det(A^TA)}\operatorname{vol}_m P, \quad \textit{für jedes Parallelotop } P = P(w_1,\dots,w_m) \subset W. \tag{12.21}$$

12.2.10. **Definition.** Seien W, V Skalarprodukträume $L : W \to V$ eine lineare Abbildung und sei A ihre Darstellungsmatrix bzgl. Orthonormalbasen von W und V. Die Matrix A^TA heißt ***Maßtensor*** und $\det(A^TA)$ heißt die ***Gramsche Determinante*** der linearen Abbildung L.

Der Maßtensor ist von der Wahl der Basen abhängig, die Gramsche Determinante nicht.
Die Formel (12.21) besagt, dass die Wurzel der Gramsche Determinante die *Volumenverzerrung* durch die lineare Abbildung L angibt.

12.2.11. **Satz.** *Sei $(V,[\omega],\langle\cdot,\cdot\rangle)$ ein orientierter Skalarproduktraum. Seien $v_1,\dots,v_n \in V$. Dann gilt*

$$|\omega_V(v_1,\dots,v_n)| = \sqrt{G(v_1,\dots,v_n)} = \operatorname{vol}_n P(v_1,\dots,v_n). \tag{12.22}$$

Ist $(v_1,\dots,v_n)$ eine Basis von V, so gilt

$$\omega_V = \varepsilon(B)\sqrt{G(v_1,\dots,v_n)}\, v_1^* \wedge \dots \wedge v_n^*.$$

Beweis: Sei $(b_1,\dots,b_n)$ eine positiv orientierte Orthonormalbasis. Dann gilt nach Satz 12.2.5:

$$\omega_V = b_1^* \wedge \dots \wedge b_n^* \quad \text{und} \quad \omega_V(v_1,\dots,v_n) = \det\big(b_j^*(v_i)\big) = \det A. \tag{12.23}$$

(Beachte: $b_j^*(v_i)$ sind die Koordinaten von v_i bzgl. der Basis $(b_1,\dots,b_n)$). Nun ist $A \in M_{n\times n}$ quadratisch, also $\sqrt{\det(A^TA)} = |\det A|$. Aus (12.10) und (12.30) folgt (12.22). Falls $(v_1,\dots,v_n)$ eine Basis bilden, gilt $\omega_V = \omega_V(v_1,\dots,v_n)v_1^* \wedge \dots \wedge v_n^*$ und $\omega_V(v_1,\dots,v_n) = \varepsilon(v_1,\dots,v_n)G(v_1,\dots,v_n)^{1/2}$. □

12.2.12. **Definition.** Sei $(V,[\omega],\langle\cdot,\cdot\rangle)$ ein orientierter Skalarproduktraum mit $\dim V = n$ mit Volumenelement ω_V. Seien $v_1,\dots,v_{n-1} \in V$. Die Abbildung $V \to \mathbb{R}$, $w \mapsto \omega_V(w,v_1,\dots,v_{n-1})$ ist linear also existiert ein eindeutig bestimmter Vektor $v_1 \times \dots \times v_{n-1} \in V$, gennant das ***Kreuzprodukt*** der Vektoren $v_1,\dots,v_{n-1}$, mit der Eigenschaft

$$\omega_V(w,v_1,\dots,v_{n-1}) = \langle w, v_1 \times \dots \times v_{n-1}\rangle, \quad \text{für alle } w \in V.$$

Es gilt: $v_1,\dots,v_{n-1}$ sind linear abhängig, genau dann, wenn $v_1 \times \dots \times v_{n-1} = 0$.
Ist $B = (b_1,\dots,b_n)$ eine positiv orientierte Orthonormalbasis und $v_i = \sum v_{ij} b_j$, dann kann man formal schreiben

$$v_1 \times \dots \times v_{n-1} = \det\begin{pmatrix} b_1 & v_{11} & \dots & v_{n-11} \\ b_2 & v_{12} & \dots & v_{n-12} \\ \dots & \dots & \dots & \dots \\ b_n & v_{1n} & \dots & v_{n-1n} \end{pmatrix} \tag{12.24}$$

d.h. die Basisdarstellung von $v_1 \times \dots \times v_{n-1}$ bzgl. B erhält man durch Entwicklung der Determinante nach der ersten Spalte.

12.2.13. **Satz.** *Sei* $(V,[\omega],\langle\cdot,\cdot\rangle)$ *ein orientierter Skalarproduktraum mit* $\dim V = n$ *und Volumenelement* ω_V. *Seien* $v_1,\dots,v_{n-1}\in V$ *linear unabhängig. Dann gilt:*

(1) $v_1\times\cdots\times v_{n-1}$ *ist orthogonal auf* v_i, *für alle* $i=1,\dots,n-1$,

(2) $\|v_1\times\cdots\times v_{n-1}\| = \sqrt{G(v_1,\dots,v_{n-1})} = \mathrm{vol}_{n-1}P(v_1,\dots,v_{n-1})$,

(3) $(v_1\times\cdots\times v_{n-1},v_1,\dots,v_{n-1})$ *ist eine positiv orientierte Basis von* V,

(4) Für jede Permutation σ *von* $\{1,\dots,n-1\}$ *gilt*

$$v_{\sigma(1)}\times\cdots\times v_{\sigma(n-1)} = \varepsilon(\sigma)v_1\times\cdots\times v_{n-1}\,.$$

12.2.14. **Definition.** Sei $(V,[\omega],\langle\cdot,\cdot\rangle)$ ein orientierter Skalarproduktraum mit $\dim V = n$ und Volumenelement ω_V. Sei $W\subset V$ ein orientierter Unterraum von Kodimension 1. Der ***äußerer Einheitsnormalenvektor*** ist der eindeutig bestimmter Vektor $\boldsymbol{n}_W$ mit der Eigenschaften, dass

(1) $(\boldsymbol{n}_W,v_1,\dots,v_{n-1})$ ist eine positiv orientierte Basis in V für jede positiv orientierte Basis $(v_1,\dots,v_{n-1})$ in W,

(2) $\|\boldsymbol{n}_W\| = 1$.

12.2.15. **Satz.** *Seien* $W\subset V$ *wie in Definition 12.2.14 und sei* $B=(v_1,\dots,v_{n-1})$ *eine Basis in* W. *Dann gilt*

$$\boldsymbol{n}_W = \varepsilon(B)\frac{v_1\times\cdots\times v_{n-1}}{\|v_1\times\cdots\times v_{n-1}\|}\,. \tag{12.25}$$

Ist ω_W *das Volumenelement von* W *(versehen mit dem induzierten Skalarprodukt) so gilt*

$$\omega_W = \boldsymbol{n}_W \lrcorner\, \omega_V\,. \tag{12.26}$$

12.2.16. **Definition.** Ein Isomorphismus $T:(V,[\omega])\longrightarrow(W,[\eta])$ heißt ***orientierungserhaltend*** (bzw. ***orientierungsumkehrend***), falls $T^*\eta=\lambda\omega$ mit $\lambda>0$ (bzw. $\lambda<0$). Wir setzen

$$\varepsilon(T)=\begin{cases}1, & \text{falls } T \text{ orientierungserhaltend ist,}\\ -1, & \text{falls } T \text{ orientierungsumkehrend ist.}\end{cases}$$

12.3. Differentialformen auf Untermannigfaltigkeiten.

12.3.1. **Definition.** Sei $M^n\subset\mathbb{R}^N$ eine Untermannigfaltigkeit. Sei $p\in\mathbb{N}_0$. Unter einer ***Differentialform vom Grad*** p (kurz p-Form) auf M verstehen wir eine Abbildung $\omega: M\longrightarrow\bigcup_{x\in M}\bigwedge^p T_x^*M$ so, dass $\omega(x)\in\bigwedge^p T_x^*M$ für jedes $x\in M$.

Beispiele:

(1) Wegen $\bigwedge^0 T_x^*M=\mathbb{R}$ ist eine 0-Form auf M einfach eine Funktion $\omega: M\longrightarrow\mathbb{R}$.

(2) Ist $f\in\mathcal{C}^\infty(M)$, so ist $df(x)\in T_x^*M$, also ist die Abbildung $M\ni x\mapsto df(x)\in T_x^*M$ eine 1-Form auf M, genannt das Differential von f, bezeichnet mit df. Die Abbildungen $\mathbb{R}^n\longrightarrow\bigcup_{x\in\mathbb{R}^n}T_x^*\mathbb{R}^n$, $x\mapsto dx_i(x)=e_i^*=\mathrm{pr}_i$ sind 1-Formen auf $\mathbb{R}^n$.

(3) Sei $U\subset\mathbb{R}^n$ offen. Wegen $T_xU\cong\mathbb{R}^n$ für alle $x\in U$ gilt $\bigcup_{x\in M}\bigwedge^p T_x^*U\cong U\times\bigwedge^p(\mathbb{R}^n)^*$. Eine p-Form ist eine Abbildung $U\to U\times\bigwedge^p(\mathbb{R}^n)^*$, $x\mapsto(x,\omega(x))$, wobei $\omega(x)\in\bigwedge^p(\mathbb{R}^n)^*$. Wir identifiezieren deshalb eine p-Form auf U mit einer Abbildung $\omega: U\longrightarrow\bigwedge^p(\mathbb{R}^n)^*$.

(4) Für $p>n$ ist $\bigwedge^p T_x^*M=\{0\}$. Eine p-Form ist also für $p>n$ die Abbildung $x\mapsto 0\in\bigwedge^p T_x^*M$.

12.3.2. **Definition.** Sind ω,η zwei p-Formen auf M, dann ist $\omega+\eta$ die p-Form auf M mit $(\omega+\eta)(x):=\omega(x)+\eta(x)$. Ist ω eine p-Form und η eine q-Form auf M, dann ist $\omega\wedge\eta$ die $(p+q)$-Form auf M mit $(\omega\wedge\eta)(x)=\omega(x)\wedge\eta(x)$ für alle $x\in M$. Das Dachprodukt von Formen auf M ist bilinear, assoziativ und antikommutativ, d.h. $\omega\wedge\eta=(-1)^{pq}\eta\wedge\omega$.

Für $M=\mathbb{R}^n$ und ein p-Tupel $I=(i_1,\dots,i_p)$ aus $\{1,\dots,n\}$ setze $dx_I=dx_{i_1}\wedge\cdots\wedge dx_{i_p}$. Für $I=\emptyset$ setzen wir $dx_I=1$.

12.3.3. **Satz.** *Sei* $U\subset\mathbb{R}^n$ *offen. Jede p-Form auf* U *hat die Form*

$$\omega=\sum_{|I|=p}{}'\,\omega_I dx_I \quad \textit{mit}\quad \omega_I(x)=\omega(x)(e_{i_1},\dots,e_{i_p})\,.$$

12.3.4. **Folgerung.**

$$df=\sum_{i=1}^n\frac{\partial f}{\partial x_i}dx_i\,.$$

12.3.5. **Definition.** Sei $U \subset \mathbb{R}^n$ offen. Eine p-Form ω auf U heißt glatt (oder von der Klasse $\mathcal{C}^\infty$), wenn ω als Abbildung $\omega : U \longrightarrow \bigwedge^p(\mathbb{R}^n)^* \cong \mathbb{R}^{\binom{n}{p}}$ glatt ist, d.h. falls $\omega_I \in \mathcal{C}^\infty(U)$ für alle streng aufsteigenden Multiindizes I der Länge p. Sei $\Omega^p(U)$ der Vektorraum der glatten p-Formen auf U. Wir definieren das ***Differential*** (äußeres Differential, Cartan-Ableitung) $d : \Omega^p(U) \longrightarrow \Omega^{p+1}(U)$ durch

$$d\Big(\sum_{|I|=p}{}' \omega_I dx_I\Big) = \sum_{|I|=p}{}' d\omega_I \wedge dx_I = \sum_{|I|=p} \sum_{i=1}^n \frac{\partial \omega_I}{\partial x_i} dx_i \wedge dx_I .$$

Beispiele:

(1) Sei $\omega = Pdx + Qdy \in \Omega^1(U)$. Dann ist

$$d\omega = dP \wedge dx + dQ \wedge dy = \Big(\frac{\partial P}{\partial x}dx + \frac{\partial P}{\partial y}dy\Big) \wedge dx + \Big(\frac{\partial Q}{\partial x}dx + \frac{\partial Q}{\partial y}dy\Big) \wedge dy = \Big(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\Big) dx \wedge dy .$$

(2) Allgemeiner gilt:

$$d\Big(\sum_{j=1}^n f_j dx_j\Big) = \sum_{i<j} \Big(\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j}\Big) dx_i \wedge dx_j .$$

(3) Sei $X : U \longrightarrow \mathbb{R}^n$ ein Vektorfeld. Die Abbildung

$$i_X : \Omega^p(U) \longrightarrow \Omega^{p-1}(U),$$

definiert durch

$$(i_X\omega)(v_1,\dots,v_{p-1}) := (X \lrcorner \omega)(v_1,\dots,v_{p-1}) := \omega(X, v_1,\dots,v_{p-1})$$

heißt ***inneres Produkt von*** ω ***mit dem Vektorfeld*** X. Dann ist

$$\begin{aligned}
&(X \lrcorner \omega_{\mathbb{R}^n})(v_1,\dots,v_{n-1}) = (dx_1 \wedge \dots \wedge dx_n)(X, v_1,\dots,v_{n-1}) \\
&= \det(X, v_1,\dots,v_{n-1}) = \Big(\sum_{j=1}^n (-1)^{j+1} X_j\, dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n\Big)(v_1,\dots,v_{n-1}), \\
&d(X \lrcorner \omega_{\mathbb{R}^n}) = \sum_{j=1}^n (-1)^{j+1} dX_j \wedge dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n \\
&= \sum_{j=1}^n (-1)^{j+1} \sum_{k=1}^n \frac{\partial X_j}{\partial x_k} dx_k \wedge dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n \\
&= \sum_{j=1}^n (-1)^{j+1} \frac{\partial X_j}{\partial x_j} dx_j \wedge dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n \\
&= \Big(\sum_{j=1}^n \frac{\partial X_j}{\partial x_j}\Big) dx_1 \wedge \dots \wedge dx_n =: (\operatorname{div} X)\omega_{\mathbb{R}^n} .
\end{aligned}$$

Die Funktion $\operatorname{div} X := \sum_{j=1}^n \frac{\partial X_j}{\partial x_j}$ heißt die ***Divergenz*** des Vektorfeldes X.

12.3.6. **Satz.**

(a) *Für* $f \in \Omega^0(U)$ *ist* df *das übliche Differential der Funktion* f,

(b) $d(\omega + \eta) = d\omega + d\eta$ *für* $\omega, \eta \in \Omega^p(U)$,

(c) $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^p \omega \wedge d\eta$ *für* $\omega \in \Omega^p(U)$, $\eta \in \Omega^q(U)$ *(Leibniz-Regel)*,

(d) $d(d\omega) = 0$ *für alle* $\omega \in \Omega^p(U)$.

12.3.7. **Definition.** Eine p-Form $\omega \in \Omega^p(U)$ heißt ***geschlossen*** , falls $d\omega = 0$; sie heißt ***exakt***, falls es eine Form $\eta \in \Omega^{p-1}(U)$ mit $d\eta = \omega$ gibt. In diesem Falle heißt η ein Potential von ω; für $p = 1$ heißt η ***Stammfunktion*** von ω.

Aussage 12.3.6 (d) besagt, dass $d\omega = 0$ (ω ist geschlossen) eine notwendige Bedingung dafür darstellt, dass ω exakt ist. Diese Bedingung ist aber nicht hinreichend: Die ***Windungsform***

$$\omega = \frac{-y dx + x dy}{x^2 + y^2} \in \Omega^1(\mathbb{R}^2 \setminus \{0\}) \tag{12.27}$$

ist geschlossen, aber nicht exakt (siehe Übungsblatt).

12.3.8. **Definition.** Eine Menge $U \subset \mathbb{R}^n$ heißt ***sternförmig***, falls ein $x_0 \in U$ existiert so, dass für jedes $x \in U$ die Verbindungsstrecke $[x_0, x]$ in U liegt.

12.3.9. **Satz** (Poincaré-Lemma). *Sei $U \subset \mathbb{R}^n$ offen und sternförmig. Sei $p \geqslant 1$, und sei $\omega \in \Omega^p(U)$ mit $d\omega = 0$. Dann gibt es $\eta \in \Omega^{p-1}(U)$ mit $d\eta = \omega$.*

12.3.10. **Definition.** Seien M_1, M_2 Untermannigfaltigkeiten und $F \in \mathcal{C}^\infty(M_1, M_2)$. Sei ω eine p-Form auf M_2. Wir definieren die p-Form $F^*\omega$ auf M_1 folgendermaßen: $dF(x) : T_xM_1 \longrightarrow T_{F(x)}M_2$ ist linear und definiert eine Abbildung $(dF(x))^* : \bigwedge^p T^*_{F(x)}M_2 \longrightarrow \bigwedge^p T^*_x M_1$. Setze $(F^*\omega)(x) = (dF(x))^*\omega(F(x))$, d.h.

$$(F^*\omega)(x)\cdot(v_1,\ldots,v_p) = \omega(F(x))(dF(x)\cdot v_1,\ldots,dF(x)\cdot v_p).$$

Beispiele:

(i) Für $p = 0$ ist $\omega : M_2 \longrightarrow \mathbb{R}$ eine Funktion, und $F^*\omega = \omega \circ F$.

(ii) Seien $U, V \subset \mathbb{R}^n$ offen, $F \in \mathcal{C}^\infty(U, V)$ und $p = n$. Nach Satz 12.1.7 gilt $(F^*\omega)(x) = (\det dF(x))\cdot\omega(F(x))$. Für $\omega(y) = g(y)dy_1 \wedge \cdots \wedge dy_n$ gilt

$$(F^*\omega)(x) = (\det dF(x))\cdot(g\circ F)(x)\,dx_1 \wedge \cdots \wedge dx_n\,.$$

(iii) Sei $\omega \in \Omega^0(V) = \mathcal{C}^\infty(V)$. Dann gilt:

$$d(F^*\omega)(x) = d(\omega\circ F)(x) = d\omega(F(x))\circ dF(x) = (F^*(d\omega))(x),$$

also $d(F^*\omega) = F^*(d\omega)$.

12.3.11. **Satz.** *Seien $U, V \subset \mathbb{R}^n$ offen, $F \in \mathcal{C}^\infty(U, V)$. Dann gilt:*

(i) $F^*(\omega + \omega') = F^*\omega + F^*\omega'$ *für alle* $\omega, \omega' \in \Omega^p(V)$,

(ii) $F^*(\omega \wedge \eta) = F^*\omega \wedge F^*\eta$ *für alle* $\omega \in \Omega^p(V)$, $\eta \in \Omega^q(V)$,

(iii) $(G\circ F)^*\omega = F^*(G^*\omega)$ *für alle* $\omega \in \Omega^p(W)$, $G \in \mathcal{C}^\infty(V, W)$,

(iv) $d(F^*\omega) = F^*(d\omega)$ *für alle* $\omega \in \Omega^p(V)$,

(v) $F^*\big({\sum}'_{|I|=p}\omega_I dy_I\big) = {\sum}'_{|I|=p}{\sum}'_{|J|=p}(\omega_I\circ F)(x)\dfrac{\partial(f_{i_1}\cdots f_{i_p})}{\partial(x_{j_1}\cdots x_{j_p})}dx_J$

mit $\dfrac{\partial(f_{i_1}\cdots f_{i_p})}{\partial(x_{j_1}\cdots x_{j_p})} := \det\Big(\dfrac{\partial f_{i_\alpha}}{\partial x_{j_\beta}}\Big)_{1\leqslant\alpha,\beta\leqslant p}$.

Beispiele:

(i) Für die Windungsform (12.27) und die Polarkoordinatenabbildung P in $\mathbb{R}^2$ gilt $P^*\omega = d\vartheta$.

(ii) Sei die 2-Form ω auf $\mathbb{R}^3 \setminus \{0\}$ gegeben durch

$$\omega(x,y,z) = \frac{1}{(x^2+y^2+z^2)^{3/2}}(x\,dy\wedge dz + y\,dz\wedge dx + z\,dx\wedge dy),$$

und sei $P : \mathbb{R}_+ \times (0, 2\pi) \times (-\frac{\pi}{2}, \frac{\pi}{2}) \ni (r, \varphi, \vartheta) \mapsto (r\cos\varphi\cos\vartheta, r\sin\varphi\cos\vartheta, r\sin\vartheta) \in \mathbb{R}^3$ die Kugelkoordinatenabbildung. Dann ist $P^*\omega = \cos\vartheta\; d\varphi \wedge d\vartheta$.

12.3.12. **Definition.** Sei $M^n \subset \mathbb{R}^N$ eine Untermannigfaltigkeit. Eine p-Form $\omega : M \longrightarrow \bigcup_{x\in M}\bigwedge^p T^*_x M$ heißt ***glatt***, falls für jede Parametrisierung $\psi : D \to M$ von M (mit $D \subset \mathbb{R}^n$ offen) gilt: $\psi^*\omega \in \Omega^p(D)$. Der Vektorraum der glatten p-Formen auf M wird mit $\Omega^p(M)$ bezeichnet.

12.3.13. **Satz.** *Sei $M^n \subset \mathbb{R}^N$ eine Untermannigfaltigkeit und $\omega \in \Omega^p(M)$. Dann existiert eine eindeutig bestimmte $(p+1)$-Form $\eta \in \Omega^{p+1}(M)$, so dass $d(\psi^*\omega) = \psi^*\eta$ für alle Parametrisierungen $\psi : D \to M$ von M. Die Form η heißt das **Differential** von ω und wird mit $d\omega \in \Omega^{p+1}(M)$ bezeichnet.*

12.4. **Orientierbare Untermannigfaltigkeiten.** Sei $M^n \subset \mathbb{R}^N$ eine Untermannigfaltigkeit, $(U, \varphi = (x_1, \ldots, x_n))$ eine Karte. Dann ist $dx_1 \wedge \ldots \wedge dx_n = d\varphi_1 \wedge \ldots \wedge d\varphi_n \in \Omega^n(U)$ eine glatte n-Form auf U, und $(dx_1 \wedge \ldots \wedge dx_n)(\frac{\partial}{\partial x_1}(x), \ldots, \frac{\partial}{\partial x_n}(x)) = 1$, also $(dx_1 \wedge \ldots \wedge dx_n)(x) \in \bigwedge^n T^*_x M \setminus \{0\}$ für alle $x \in U$. Daher ist $dx_1 \wedge \ldots \wedge dx_n$ für jedes $x \in U$ eine Orientierung auf T_xM. Weil $dx_1 \wedge \ldots \wedge dx_n$ glatt ist, hängen diese Orientierungen gewissermaßen glatt von $x \in U$ ab.

12.4.1. **Definition.**

(a) Eine n-dimensionale Untermannigfaltigkeit $M^n \subset \mathbb{R}^N$ heißt ***orientierbar***, falls eine nirgends verschwindende n-Form $\omega \in \Omega^n(M)$ existiert, d.h. mit $\omega(x) \in \bigwedge^n T^*_x M \setminus \{0\}$ für alle $x \in M$.

(b) Sei M orientierbar, und seien $\omega, \omega' \in \Omega^n(M)$ zwei Formen wie in (i). Wegen $\bigwedge^n T^*_x M \cong \mathbb{R}$ und $\omega(x), \omega'(x) \in \bigwedge^n T^*_x M$ gibt es $\lambda(x) \neq 0$ mit $\omega(x) = \lambda(x)\omega'(x)$. Wir sagen, dass $\omega \sim \omega'$, falls $\lambda(x) > 0$ für alle $x \in M$ (d.h. wenn für jedes $x \in M$ die Formen $\omega(x)$ und $\omega'(x)$ dieselbe Orientierung auf T_xM induzieren). Dies ist eine Äquivalenzrelation. Eine Äquivalenzklasse $[\omega]$ heißt eine ***Orientierung*** auf M. Eine ***orientierte***

Untermannigfaltigkeit ist ein Paar $(M,[\omega])$ bestehend aus einer Untermannigfaltigkeit M und einer Orientierung $[\omega]$ auf M.

(c) Seien $(M,[\omega])$ eine orientierte Untermannigfaltigkeit, (U,φ) eine Karte von M, und sei $\lambda(x)=\omega(\frac{\partial}{\partial x_1},\dots,\frac{\partial}{\partial x_n})$ für $x\in U$, d.h. $\omega=\lambda\,dx_1\wedge\ldots\wedge dx_n$. Sei $\varepsilon_x(\varphi)=\operatorname{sgn}\lambda(x)\in\{\pm1\}$. Ist $x\mapsto\varepsilon_x(\varphi)$ konstant und gleich 1 (bzw. -1), so heißt (U,φ) ***positiv orientiert*** (bzw. ***negativ orientiert***).

(d) Seien $(M_1,[\omega_1])$ und $(M_2,[\omega_2])$ zwei orientierte Untermannigfaltigkeiten. Ein lokaler Diffeomorphismus $f:M_1\to M_2$ heißt ***orientierungserhaltend*** (bzw. ***orientierungsumkehrend***), falls $f^*(\omega_2)=\lambda\,\omega_1$, wobei $\lambda(x)>0$ (bzw. $\lambda(x)<0$) für alle $x\in M$.

Bemerkung:

(1) Ist M orientierbar und zusammenhängend, so gibt es genau zwei Orientierungen auf M.

(2) Es gilt $\varepsilon_x(\varphi)=\varepsilon(d\varphi(x))$, wobei $d\varphi(x):(T_xM,\omega_x)\to(\mathbb{R}^n,\omega_{\mathbb{R}^n})$ (siehe Definition 12.2.16).

(3) Die Abbildung $x\mapsto\varepsilon_x(\varphi)$ ist lokal konstant, weil λ stetig ist. Ist $x\mapsto\varepsilon_x(\varphi)$ konstant (z. B. wenn U zusammenhängend ist) so heisst diese Konstante das Vorzeichen der Karte φ, bezeichnet mit $\varepsilon(\varphi)$.

(4) Ist (U,φ) positiv orientiert, so ist $(\frac{\partial}{\partial x_1}(x),\dots,\frac{\partial}{\partial x_n}(x))$ für $x\in U$ eine positiv orientierte Basis von T_xM.

(5) Ist $f:(M_1,[\omega_1])\to(M_2,[\omega_2])$ ein lokaler Diffeomorphismus so kann man das Vorzeichen von f in einem Punkt als $\varepsilon_x(f):=\varepsilon(df(x))$ definieren, äquivalent $\varepsilon_x(f):=\operatorname{sgn}\lambda(x)$ wobei $f^*(\omega_2)=\lambda\,\omega_1$. Ist $x\mapsto\varepsilon_x(f)$ konstant (z. B. wenn M_1 zusammenhängend ist) so heisst diese Konstante $\varepsilon(f)$ das Vorzeichen von f. Offensichtlich ist $\varepsilon(f)=1$ falls f orientierungserhaltend ist bzw. $\varepsilon(f)=-1$ falls f orientierungsumkehrend ist.

12.4.2. **Beispiel.**

(1) Für $M=\mathbb{R}^n$ ist $\omega_{\mathbb{R}^n}(x)=dx_1(x)\wedge\ldots\wedge dx_n(x)$ die Standardorientierung auf jedem Tangentialraum $T_x\mathbb{R}^n=\mathbb{R}^n$. Die n-Form $\omega_{\mathbb{R}^n}=dx_1\wedge\ldots\wedge dx_n$ induziert also eine Orientierung, genannt die Standardorientierung auf $\mathbb{R}^n$.

(2) Existiert eine globale Karte $(U,\varphi=(\varphi_1,\dots,\varphi_n)=(x_1,\dots,x_n))$ mit $M=U$, so induziert $\omega=d\varphi_1\wedge\ldots\wedge d\varphi_n=dx_1\wedge\ldots\wedge dx_n$ eine Orientierung auf M.

(3) Jede 1-dimensionale Untermannigfaltigkeit ist orientierbar. Eigentlich ist jede 1-dimensionale Untermannigfaltigkeit diffeomorph zu S^1, falls sie kompakt ist, oder zu $(0,1)$ falls sie nicht-kompakt ist.

(4) Das Möbiusband ist nicht orientierbar.

(5) Ist $\varphi:U\to V$ ein Diffeomorphismus zwischen zwei offenen Teilmengen des $\mathbb{R}^n$, so gilt $\varepsilon_x(\varphi)=\operatorname{sgn}\det J_{\varphi^{-1}}(\varphi(x))=\operatorname{sgn}\det J_\varphi(x)$. Beispielsweise ist die Polarkoordinatenabbildung orientierungserhaltend, da $\det J_P(r,\vartheta)=r>0$ für alle $(r,\vartheta)\in\mathbb{R}_+\times\mathbb{R}$.

Sind $M_1\subset M_2$ zwei Untermannigfaltigkeiten, so bezeichnen wir durch $\iota:M_1\hookrightarrow M_2$ die Inklusionsabbildung, d.h. $\iota(x)=x$ für alle $x\in M_1$. Ist $\alpha\in\Omega^p(M_2)$, so ist die zurückgeholte Form $\iota^*\alpha\in\Omega^p(M_1)$ die Einschränkung von α auf Tangentialvektoren an M_1, d.h. $(\iota^*\alpha)(x)(v_1,\dots,v_p)=\alpha(x)(v_1,\dots,v_p)$ für alle $x\in M_1$ und alle $v_1,\dots,v_p\in T_xM_1\subset T_xM_2$.

12.4.3. **Satz.** *Sei $M^n\subset\mathbb{R}^N$ gegeben durch $M^n=\{x\in\mathbb{R}^N : f_1(x)=\ldots=f_{N-n}(x)=0\}$, wobei 0 ein regulärer Wert von $f=(f_1,\dots,f_{N-n}):\mathbb{R}^N\longrightarrow\mathbb{R}^{N-n}$ sei. Sei $\iota:M\hookrightarrow\mathbb{R}^N$ die Inklusion. Dann ist $\omega:=\iota^*(\operatorname{grad}f_{N-n}\lrcorner\ldots\lrcorner\operatorname{grad}f_1\lrcorner\,\omega_{\mathbb{R}^N})$ eine nirgends verschwindende n-Form auf M. Insbesondere ist M orientierbar.*

Beweis: Laut Definiton gilt $\omega(x)\in\bigwedge^n T_x^*M$ für alle $x\in M$, und

$$\begin{aligned}\omega(x)(v_1,\dots,v_n)&=\omega_{\mathbb{R}^N}(\operatorname{grad}f_1(x),\operatorname{grad}f_2(x),\dots,\operatorname{grad}f_{N-n}(x),v_1,\dots,v_n)\\&=\det(\operatorname{grad}f_1(x),\operatorname{grad}f_2(x),\dots,\operatorname{grad}f_{N-n}(x),v_1,\dots,v_n)\end{aligned}$$

für alle $v_1,\dots,v_n\in T_xM$. Da die Vektoren $\{\operatorname{grad}f_1(x),\operatorname{grad}f_2(x),\dots,\operatorname{grad}f_{N-n}(x)\}$ linear unabhängig und orthogonal zu T_xM sind, ist $\omega(x)(v_1,\dots,v_n)\neq0$ für jede Basis $(v_1,\dots,v_n)$ von T_xM. Außerdem ist ω glatt, da $\operatorname{grad}f_1,\operatorname{grad}f_2,\dots,\operatorname{grad}f_{N-n}$ und die Determinantenabbildung glatt sind. □

12.4.4. **Definition.** Seien $H^{n-1}\subset M^n$ Untermannigfaltigkeiten (H heißt Hyperfläche in M). Ein ***Einheitsnormalenfeld*** zu H bzgl. M ist eine glatte Abbildung $\boldsymbol{n}:H\to\mathbb{R}^N$, so dass für alle $p\in H$ gilt $\boldsymbol{n}(p)\in T_pM$, $\boldsymbol{n}(p)\perp T_pH$ und $\|\boldsymbol{n}(p)\|=1$.

12.4.5. **Satz.** *Seien $H^{n-1}\subset M^n$ Untermannigfaltigkeiten mit M^n orientierbar. Dann ist H orientierbar genau dann, wenn ein Einheitsnormalenfeld zu H bzgl. M existiert.*

12.4.6. **Folgerung.** *Sei M orientierbar und $D \subset M$ eine glatt berandete Teilmenge. Dann ist ∂D orientierbar. Ist $\boldsymbol{n}$ das äußere Einheitsnormalenfeld auf ∂D und $\omega \in \Omega^n(M)$ eine nirgends verschwindende n-Form, so ist $\iota^*(\boldsymbol{n} \lrcorner\, \omega) \in \Omega^{n-1}(\partial D)$ eine nirgends verschwindende $(n-1)$-Form auf ∂D.*

12.4.7. **Definition.** Ist $(M,[\omega])$ eine orientierte Untermannigfaltigkeit, $D \subset M$ eine glatt berandete Teilmenge und $\boldsymbol{n} : \partial D \longrightarrow S^{N-1} \subset \mathbb{R}^N$ das äußere Einheitsnormalenfeld, so heißt $[\iota^*(\boldsymbol{n} \lrcorner\, \omega)]$ die von $[\omega]$ induzierte ***Randorientierung*** von ∂D.

12.4.8. **Beispiele.**

(1) Seien $M = \mathbb{R}^n$, $D = \mathbb{R}^n_-$, $[\omega_{\mathbb{R}^n}] = [dx_1 \wedge \ldots \wedge dx_n]$ die Standardorientierung auf $\mathbb{R}^n$. Für $x \in \partial D = \mathbb{R}^{n-1} \times \{0\}$ gilt $\boldsymbol{n}_{\partial D}(x) = e_n$, und die Randorientierung ist $[e_n \lrcorner\, \omega_{\mathbb{R}^n}] = [(-1)^{n+1} dx_1 \wedge \ldots \wedge dx_{n-1}]$.

(2) Sei $D = \{x \in \mathbb{R}^n : f(x) \leqslant 0\}$, wobei $\operatorname{grad} f(x) \neq 0$ in einer Umgebung U von ∂D. Dann ist die Randorientierung induziert durch

$$\operatorname{grad} f \lrcorner\, \omega_{\mathbb{R}^n} = \sum_{j=1}^{n} (-1)^{j+1} \frac{\partial f}{\partial x_j}(x) dx_1 \wedge \ldots \wedge \widehat{dx_j} \wedge \ldots \wedge dx_n \,.$$

(3) Sei $D \subset M$ glatt berandet. Ist $v \in T_x(\partial D)$ ein äußerer Vektor, so definiert $v \lrcorner\, \omega_M$ die Randorientierung.

Sei nun $(U, \varphi = (x_1, \ldots, x_n))$ eine D-Karte mit konstanten Vorzeichen $\varepsilon(\varphi)$. Die Orientierung ist gegeben auf U durch $\varepsilon(\varphi) dx_1 \wedge \ldots \wedge dx_n$ und $\frac{\partial}{\partial x_n}$ ein äußerer Vektor auf $\partial D \cap U$. Die Randorientierung ist gegeben durch

$$\frac{\partial}{\partial x_n} \lrcorner\, \varepsilon(\varphi)(dx_1 \wedge \ldots \wedge dx_n) = (-1)^{n+1} \varepsilon(\varphi)(dx_1 \wedge \ldots \wedge dx_{n-1}).$$

Es gilt $D \cap U = \{x \in U : x_n \leqslant 0\}$, $\partial D \cap U = \{x \in U : x_n = 0\}$ und $\widetilde{\varphi} : \partial D \cap U \to \mathbb{R}^{n-1}$, $\widetilde{\varphi} = (x_1, \ldots, x_{n-1})$ ist eine Karte auf $\partial D \cap U$. Das Vorzeichen der Karte $(\partial D \cap U, \widetilde{\varphi})$ ist

$$\varepsilon(\widetilde{\varphi}) = (-1)^{n+1} \varepsilon(\varphi) \,.$$

Dabei betrachten wir Definitionsgemäß die Standardorientierung auf $\mathbb{R}^{n-1}$. Daraus folgt, dass wenn wir auf $\partial \mathbb{R}^n_-$ die Randorientierung betrachten, ist das Vorzeichen des Diffeomorphismus $\varphi|_{\partial D \cap U} : \partial D \cap U \to \partial \mathbb{R}^n_-$ auch $\varepsilon(\varphi)$.

(4) Sei $(U, \varphi = (x_1, \ldots x_{n-1}))$ eine Karte auf ∂D mit zusammenhängendem Kartengebiet U. Dann ist

$$\varepsilon(\varphi) = \varepsilon_x(\varphi) = \operatorname{sgn} \omega\Big(\boldsymbol{n}(x), \frac{\partial}{\partial x_1}(x), \ldots, \frac{\partial}{\partial x_{n-1}}(x)\Big) = \operatorname{sgn} \det\Big(v, \frac{\partial}{\partial x_1}(x), \ldots, \frac{\partial}{\partial x_{n-1}}(x)\Big),$$

wobei $v \in T_x(\partial D)$ ein beliebiger äußerer Vektor ist.

Wir führen nun eine kanonische positive n-Form auf einer orientierten Untermannigfaltigkeit der Dimension n ein.

12.4.9. **Satz.** *Sei $(M^n, [\omega])$ eine orientierte Untermannigfaltigkeit, vorgesehen mit der induzierten Riemannschen Metrik g. Definiere die n-From ω_M durch*

$$\omega_M(p) := dv_M(p) := \omega_{T_p M} \,, \quad p \in M \,,$$

*wobei $\omega_{T_p M}$ das Volumenelement des orientierten Skalarproduktraumes $(T_p M, [\omega(p)], g_p)$ ist Dann ist ω_M glatt und heißt die **Volumenform** der orientierten Untermannigfaltigkeit $(M, [\omega])$. Die Volumenform ist diejenige glatte n-Form auf M, welche auf jede posotiv orientierte Orthonormalbasis eines Tangentialraumes $T_p M$ den Wert $+1$ annimmt.*

12.4.10. **Satz** (globale Darstellung der Volumenform). *Sei $(M, [\omega])$ eine orientierte Untermannigfaltigkeit, ω_M die Volumenform, $D \subset M$ glatt berandet, $\boldsymbol{n}$ das äußere Einheitsnormalenfeld, und sei ∂D versehen mit der Randorientierung. Dann ist $\omega_{\partial D} = \iota^*(\boldsymbol{n} \lrcorner\, \omega_M)$.*

12.4.11. **Satz** (lokale Darstellung der Volumenform). *Sei $(M, [\omega])$ eine orientierte Untermannigfaltigkeit und $(U, \varphi = (\varphi_1, \ldots, \varphi_n))$ eine Karte von M mit zusammenhängendem Kartengebiet U. Dann gilt*

$$\boxed{\omega_M|_U = \varepsilon(\varphi) \sqrt{G(\tfrac{\partial}{\partial x_1}, \ldots, \tfrac{\partial}{\partial x_n})}\, d\varphi_1 \wedge \ldots \wedge d\varphi_n} \tag{12.28}$$

und

$$\boxed{(\varphi^{-1})^* \omega_M = \varepsilon(\varphi) \sqrt{G(\tfrac{\partial}{\partial x_1}, \ldots, \tfrac{\partial}{\partial x_n}) \circ \varphi^{-1}}\, dx_1 \wedge \ldots \wedge dx_n \,.} \tag{12.29}$$

Hierbei ist $\omega_M|_U$ die Einschränkung von ω_M auf U, also eine Form auf $U \subset M$ und $(\varphi^{-1})^*\omega_M$ ist ihr Urbild durch die Parametrisierung $\varphi^{-1} : \varphi(U) \to U$, also eine Form auf $\varphi(U) \subset \mathbb{R}^n$. Wenn wir wie üblich $\varphi = (\varphi_1, \dots, \varphi_n) = (x_1, \dots, x_n)$ bezeichnen, so schreiben wir auch

$$\boxed{\omega_M|_U = \varepsilon(\varphi)\sqrt{G(\tfrac{\partial}{\partial x_1}, \dots, \tfrac{\partial}{\partial x_n})}\, dx_1 \wedge \ldots \wedge dx_n}\,. \tag{12.30}$$

Beachte die starke formale Ähnlichkeit zwischen (12.30) und (12.29). Deswegen unterscheiden wir oft nicht zwischen $\omega_M|_U$ und $(\varphi^{-1})^*\omega_M$, wenn wir über lokale Darstellung der Volumenform sprechen.

12.4.12. **Beispiele.**

(1) Im Falle der Dimension 1 heißt die Volumenform ***Linienelement*** (oder Bogenelement) und wird in der klassischen Vektoranalysis mit ds bezeichnet. Berechnen wir die lokale Darstellung der Volumenform einer 1-dimensionalen orientierten Untermannigfaltigkeit C. Sei $\gamma : I \longrightarrow C$ eine positiv orientierte Parametrisierung und $\frac{\partial}{\partial t}|_{\gamma(t)} = \gamma'(t) \in T_{\gamma(t)}C$. Dann ist $g_{11}(\gamma(t)) = \langle \frac{\partial}{\partial t}, \frac{\partial}{\partial t} \rangle = \|\gamma(t)\|^2$. Daraus folgt

$$ds_C = \omega_C(\gamma(t)) = \|\gamma'(t)\| dt. \tag{12.31}$$

Betrachten wir $C = S^1 \subset \mathbb{R}^2$ versehen mit der Randorientierung $S^1 = \partial D$ wobei $D = B_1(0)$. Die Karte $(S^1 \setminus \{e^{i(\varphi_0+\pi)}\}, e^{i\vartheta} \mapsto \vartheta \in (\varphi_0 - \pi, \varphi_0 + \pi))$ ist positiv orientiert, da $\boldsymbol{n}(x,y) = \binom{x}{y}/\|\binom{x}{y}\| = \binom{x}{y}$ und $\frac{\partial}{\partial \vartheta} = \binom{-y}{x}$ also

$$\det\left(\boldsymbol{n}, \frac{\partial}{\partial \vartheta}\right)|_{(x,y)} = \det\begin{pmatrix} x & -y \\ y & x \end{pmatrix} = 1 > 0$$

auf S^1. Also ist

$$ds_{S^1} = \omega_{S^1} = \sqrt{G(\tfrac{\partial}{\partial \vartheta}, \tfrac{\partial}{\partial \vartheta})}\, d\vartheta = \sqrt{\|\tfrac{\partial}{\partial \vartheta}\|}\, d\vartheta = d\vartheta \quad \text{auf } S^1 \setminus \{e^{i(\varphi_0+\pi)}\}.$$

Eine andere Möglichkeit, diese lokale Darstellung zu berechnen, ist es, zunächst die Volumenform von S^1 nach der Formel

$$\omega_{S^1} = \iota^*((x,y) \lrcorner\, dx \wedge dy) = \iota^*(-y dx + x dy)$$

zu bestimmen. Für die zur obigen Karte gehörige Parametrisierung

$$\psi : (\varphi_0 - \pi, \varphi_0 + \pi) \ni \vartheta \mapsto (\cos\vartheta, \sin\vartheta) \in S^1$$

gilt:

$$\begin{aligned}(\psi^*\omega_{S^1})(\vartheta) &= \psi^*(-y dx + x dy)(\vartheta) = -\sin\vartheta\, d(\cos\vartheta) + \cos\vartheta\, d(\sin\vartheta) \\ &= -\sin\vartheta(-\sin\vartheta) d\vartheta + \cos^2\vartheta\, d\vartheta = d\vartheta.\end{aligned}$$

(2) Im Falle der Dimension 2 heißt die Volumenform ***Flächenelement*** und wird in der klassischen Vektoranalysis mit dA oder dF bezeichnet.

Wir berechnen die Volumenform einer 2-dimensionalen orientierten Untermannigfaltigkeit $M^2 \subset \mathbb{R}^N$: Sei $\psi : V \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ eine positiv orientierte Parametrisierung. Klassische Notationen in diesem Zusammenhang sind:

$$E = g_{11} = \left\langle \frac{\partial \psi}{\partial u_1}, \frac{\partial \psi}{\partial u_1} \right\rangle, F = g_{22} = \left\langle \frac{\partial \psi}{\partial u_2}, \frac{\partial \psi}{\partial u_2} \right\rangle, G = g_{12} = g_{21} = \left\langle \frac{\partial \psi}{\partial u_1}, \frac{\partial \psi}{\partial u_2} \right\rangle.$$

Dann gilt:

$$\det\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \det\begin{pmatrix} E & G \\ G & F \end{pmatrix} = EF - G^2 = \left\| \frac{\partial \psi}{\partial u_1} \times \frac{\partial \psi}{\partial u_2} \right\|^2.$$

Daraus folgt

$$dA_M = \omega_M = \left\| \frac{\partial \psi}{\partial u_1} \times \frac{\partial \psi}{\partial u_2} \right\| du_1 \wedge du_2 = \sqrt{EF - G^2}\, du_1 \wedge du_2 \quad \text{auf } U.$$

Seien $D = B_1(0) \subset \mathbb{R}^3$ und $M = \partial D = S^2$, versehen mit der Randorientierung. Dann gilt

$$\boldsymbol{n}(x,y,z) = \frac{(x,y,z)}{\|(x,y,z)\|} = (x,y,z),$$

und die Volumenform von S^2 ist

$$dA_{S^2} = \omega_{S^2} = \iota^*((x,y,z) \lrcorner\, dx \wedge dy \wedge dz) = \iota^*(x dy \wedge dz + y dz \wedge dx + z dx \wedge dy).$$

Betrachten wir die Parametrisierung durch sphärische Koordinaten (siehe (11.7), (11.10))

$$\psi : (0, 2\pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \longrightarrow S^2, \quad \psi(\varphi, \vartheta) = (\cos\varphi\cos\vartheta, \sin\varphi\cos\vartheta, \sin\vartheta).$$

Berechnen wir nach (12.24) das Kreuzprodukt

$$\frac{\partial\psi}{\partial\varphi}\times\frac{\partial\psi}{\partial\vartheta}=\det\begin{pmatrix} e_1 & -\sin\varphi\cos\vartheta & -\cos\varphi\cos\vartheta\\ e_2 & \cos\varphi\cos\vartheta & -\sin\varphi\sin\vartheta\\ e_3 & 0 & \cos\vartheta\end{pmatrix}=\cos\vartheta\begin{pmatrix}\cos\varphi\cos\vartheta\\ \sin\varphi\cos\vartheta\\ \sin\vartheta\end{pmatrix}=\cos\vartheta\cdot\boldsymbol{n}$$

Nach der Definition des Kreuzprodukts ist $(\cos\vartheta\cdot\boldsymbol{n},\frac{\partial\psi}{\partial\varphi},\frac{\partial\psi}{\partial\vartheta})$ positiv orientiert. Weil $\cos\vartheta>0$, ist auch $(\boldsymbol{n},\frac{\partial\psi}{\partial\varphi},\frac{\partial\psi}{\partial\vartheta})$ positiv orientiert. Die Parametrisierung ψ ist also orientierungserhaltend. Wegen $\|\frac{\partial\psi}{\partial\varphi}\times\frac{\partial\psi}{\partial\vartheta}\|=\cos\vartheta$ gilt

$$dA_{S^2}=\cos\vartheta\,d\varphi\wedge d\vartheta\quad\text{auf}\operatorname{Im}\psi.$$

So ist die zugehörige lokale Darstellung der Volumenform:

$$(\psi^*\omega_{S^2})(\varphi,\vartheta)=\cos\vartheta\,d\varphi\wedge d\vartheta.$$

12.5. **Divergenz, Rotation, Laplace-Operator.** Wir interpretieren zunächst den Gradienten mit Hilfe von Differentialformen. Sei $(V,\langle\cdot,\cdot\rangle)$ ein endlichdimensionaler Skalarproduktraum. Wir wissen aus der Linearen Algebra, dass $V\ni v\mapsto\omega^v:=\langle v,\cdot\rangle\in V^*$ ein Isomorphismus ist; ω^v heißt die duale Form zu v.

Sei nun M eine n-dimensionale Untermannigfaltigkeit des $\mathbb{R}^N$. Ein Vektorfeld $X\in\mathcal{X}(M)$ auf M definiert eine 1-Form $\omega^X\in\Omega^1(X)$ durch $\omega^X(x)=\omega^{X(x)}$, wenn wir diesen Isomorphismus $T_xM\to T_x^*M$ betrachten. Wir sehen leicht, dass ω^X eine glatte Form ist (wie im Satz 11.6.6). Nach Definition 11.6.5 ist das folgende Diagramm kommutativ:

$$\begin{array}{ccc}\mathcal{C}^\infty(M) & \xrightarrow{\text{grad}} & \mathcal{X}(M)\\ \Big\downarrow{\scriptstyle\mathrm{Id}} & & \Big\downarrow{\scriptstyle X\mapsto\omega^X}\\ \Omega^0(M) & \xrightarrow[d]{} & \Omega^1(M)\end{array}$$

Es gilt also $\omega^{\operatorname{grad}f}=df$ für alle $f\in\mathcal{C}^\infty(M)$.

12.5.1. **Definition.** Ein Vektorfeld $X\in\mathcal{X}(M)$ heißt ein ***Gradientenfeld*** (oder ***Potentialfeld***), wenn es $f\in\mathcal{C}^\infty(M)$ mit $X=\operatorname{grad}f$ gibt; in diesem Fall heißt f ein ***Potential*** zu X. In der Physik betrachtet man oft ein ***Kraftfeld*** in einer offenen Teilmenge des $\mathbb{R}^3$ (z.B. ein elektrisches Kraftfeld oder ein Gravitationsfeld). Wir werden Kraftfelder als Vektorfelder X auffassen. Ein Kraftfeld X heißt ***konservativ***, wenn es ein Potential besitzt (man schreibt dann $X=-\operatorname{grad}f$).

Das elektrostatische Feld $\boldsymbol{E}$ ist ein konservatives Feld, es kann über den Gradienten eines Potentials Φ ausgedrückt werden, mit $\boldsymbol{E}=-\operatorname{grad}\Phi$.

12.5.2. **Satz.** (i) *X ist ein Gradientenfeld genau dann, wenn ω^X eine exakte 1-Form ist.*

(ii) *Ist X ein Gradientenfeld, so ist notwendig ω^X geschlossen. Die Bedingung $d\omega^X=0$ heißt **Integrabilitätsbedingung** für X bzw. für ω_X.*

(iii) *Ist M eine offene sternförmige Teilmenge von $\mathbb{R}^N$, so ist die Integrabilitätsbedingung für $X\in\mathcal{X}(M)$ auch hinreichend dafür, dass X ein Gradientenfeld ist.*

Beweis: Zu (i): $X=\operatorname{grad}f\iff\omega^X=\omega^{\operatorname{grad}f}\iff\omega^X=df$.
Zu (ii): $X=\operatorname{grad}f\iff\omega^X=df\implies d\omega^X=d(df)=0$.
Zu (iii): Weil M offen und sternförmig ist, ist laut dem Poincaré-Lemma ω^X ist genau dann exakt, wenn ω^X geschlossen ist. □

12.5.3. **Bemerkung.** (i) Ist $U\subset\mathbb{R}^N$ eine offene Teilmenge und $X=\sum_{j=1}^N X_j\frac{\partial}{\partial x_j}$, so ist $\omega^X=\sum_{j=1}^N X_j dx_j$, und die Integrabilitätsbedingung lautet $\frac{\partial X_i}{\partial x_j}=\frac{\partial X_j}{\partial x_i}$ für alle $i,j\in\{1,\dots,N\}$.

(ii) Die Bedingung $d\omega^X=0$ heißt Integrabilitätsbedingung, weil man eine Lösung der Gleichung $df=\omega^X$ (d.h. ein Potential f für X, oder eine Stammfunktion für ω^X) durch Integration der 1-Form ω^X längs Kurven in U erhält (siehe Satz 12.6.4).

Wir betrachten nun zwei weitere Isomorphismen für eine n-dimensionale orientierte Untermannigfaltigkeit M mit Volumenform ω_M:

$$\mathcal{X}(M)\to\Omega^{n-1}(M),\quad X\mapsto X\lrcorner\,\omega_M$$
$$\mathcal{C}^\infty(M)\to\Omega^n(M),\quad f\mapsto f\omega_M$$

12.5.4. **Definition.** Sei M^n eine orientierte Untermannigfaltigkeit und X ein Vektorfeld auf M. Die ***Divergenz*** des Vektorfeldes ist die durch $d(X \lrcorner \omega_M) = \mathrm{div}(X)\omega_M$ definierte Funktion $\mathrm{div}(X) \in \mathcal{C}^\infty(M)$.

Laut Definition ist das folgende Diagramm kommutativ:

$$\begin{tikzcd} \mathcal{X}(M) \arrow[r, "\mathrm{div}"] \arrow[d, "X \mapsto X \lrcorner \omega_M"'] & \mathcal{C}^\infty(M) \arrow[d, "f \mapsto f\omega_M"] \\ \Omega^{n-1}(M) \arrow[r, "d"'] & \Omega^n(M) \end{tikzcd}$$

12.5.5. **Satz.** *Sei $(U,\varphi = (x_1,\dots,x_n))$ eine positiv orientierte Karte auf M, und sei $X = \sum_{j=1}^n X_j \frac{\partial}{\partial x_j}$ die lokale Darstellung von X bzgl. (U,φ). Sei $g = \det(g_{ij})$. Dann gilt*

$$\mathrm{div}(X) = \frac{1}{\sqrt{g}} \sum_{j=1}^n \frac{\partial}{\partial x_j}(\sqrt{g} X_j).$$

Beispiel: $M = \mathbb{R}^n$, $X = (X_1,\dots,X_n) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. Unter Benutzung der Karte $(U,\varphi) = (\mathbb{R}^n, \mathrm{Id})$ sieht man: $\mathrm{div} X = \sum \frac{\partial X_j}{\partial x_j}$. Mit der Physiker-Notation $\nabla = \left(\frac{\partial}{\partial x_1},\dots,\frac{\partial}{\partial x_n}\right)$ man kann schreiben $div = \nabla \cdot X = \sum \frac{\partial X_j}{\partial x_j}$ als formales Skalarprodukt.

12.5.6. **Definition.** Sei M^n eine orientierte Untermannigfaltigkeit, $f \in \mathcal{C}^\infty(M)$. Der ***Laplace-Operator*** $\Delta : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ ist definiert durch $\Delta f = \mathrm{div}(\mathrm{grad} f)$.

12.5.7. **Satz.** *Sei $(U,\varphi = (x_1,\dots,x_n))$ eine positiv orientierte Karte und $f \in \mathcal{C}^\infty(M)$. Dann gilt*

$$\Delta f = \frac{1}{\sqrt{g}} \sum_{i,j=1}^n \frac{\partial}{\partial x_j}\Big(\sqrt{g}\frac{\partial f}{\partial x_i} \cdot g^{ij}\Big), \quad \textit{wobei } (g^{ij}) = (g_{ij})^{-1}.$$

Beispiel: $M = \mathbb{R}^n$, $f \in \mathcal{C}^\infty(\mathbb{R}^n)$. Unter Benutzung der Karte $(U,\varphi) = (\mathbb{R}^n, \mathrm{Id})$ sieht man: $\Delta f = \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j^2}$ also $\Delta = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}\cdot$ Mit der Physiker-Notation kann man schreiben $\Delta = \nabla \cdot \nabla$ als formales Skalarprodukt.

Der Laplace-Operator wurde von Pierre-Simon Marquis de Laplace (1749–1827) eingeführt. Die Gleichung $\Delta u = 0$ heißt *Laplace-Gleichung*, ihre Lösungen *harmonische Funktionen*. Die Gleichung $\Delta u = f$ heißt *Poisson-Gleichung*. Der Poisson-Gleichung genügen das elektrostatische Potential und das Gravitationspotential. Dabei ist f proportional zur elektrischen Ladungsdichte beziehungsweise zur Massendichte.

Sei $\boldsymbol{E} = -\mathrm{grad}\,\Phi$ das elektrostatische Feld. Daraus folgt $\mathrm{div}\boldsymbol{E} = -\mathrm{div}\,\mathrm{grad}\,\Phi = \Delta\Phi$. Nach der ersten Maxwellgleichung gilt auch $\mathrm{div}\boldsymbol{E} = \frac{\rho}{\varepsilon}$, wobei ρ die Ladungsdichte und ε die Permittivität sind. Damit folgt die Poisson-Gleichung des elektrischen Feldes $\Delta\Phi = -\frac{\rho}{\varepsilon}$.

Der Laplace-Operator ist Bestandteil anderer wichtiger Differentialgleichungen der mathematischen Physik. Wir betrachten jetzt Funktionen von $n+1$ Variablen $(t,x_1,\dots,x_n) = (t,x)$. Die Schreibweise rührt davon her, dass in den Anwendungen t meist als Zeit interpretiert wird, während die $x_1,\dots,x_n$ räumliche Variable darstellen. Sei $u \in \mathcal{C}^2(\mathbb{R}_+ \times \mathbb{R}^n)$, $u = u(t,x_1,\dots,x_n)$. Die *Wellengleichung* bzw. die *Wärmeleitungsgleichung* sind die lineare partielle Differentialgleichungen zweiter Ordnung

$$\frac{\partial^2 u}{\partial t^2} - \Delta u = 0, \quad \text{bzw.} \quad \frac{\partial u}{\partial t} - \Delta u = 0.$$

12.5.8. **Definition.** und die Sei M^3 eine orientierte Untermannigfaltigkeit der Dimension 3. Sei $X \in \mathcal{X}(M)$ und ω^X die duale 1-Form ($\omega^X(x) \cdot u = \langle X(x), u\rangle$ für alle $u \in T_x(M)$). Dann ist $d\omega^X(x) \in \bigwedge^2 T_x^* M$. Wegen des Isomorphismus $T_x M \cong \bigwedge^2 T_x^* M$, $v \mapsto v \lrcorner \omega_M(x)$ gibt es ein Vektorfeld, bezeichnet mit $\mathrm{rot}(X)$, so dass $\mathrm{rot}(X) \lrcorner \omega_M = d\omega^X$. Ein Vektorfeld X heißt ***rotationsfrei***, falls $\mathrm{rot}(X) = 0$ (d.h. $d\omega^X = 0$). Ist X ein Kraftfeld, so sagen wir, das Kraftfeld mit $\mathrm{rot}(X) = 0$ sei ***wirbelfrei***.

Laut Definition ist das folgende Diagramm kommutativ:

$$\begin{tikzcd} \mathcal{X}(M^3) \arrow[r, "\mathrm{rot}"] \arrow[d, "X \mapsto \omega^X"'] & \mathcal{X}(M^3) \arrow[d, "X \mapsto X \lrcorner \omega_{M^3}"] \\ \Omega^1(M^3) \arrow[r, "d"'] & \Omega^2(M^3) \end{tikzcd}$$

Die Rotation ist nur für Vektorfelder auf 3-dimensionalen orientierten Untermannigfaltigkeiten definiert!

Beispiel: Sei $M\subset\mathbb{R}^3$, $(U,\varphi)=(\mathbb{R}^3,\mathrm{Id})$, $X=X_1\frac{\partial}{\partial x_1}+X_2\frac{\partial}{\partial x_2}+X_3\frac{\partial}{\partial x_3}=(X_1,X_2,X_3)\in\mathcal{C}^\infty(\mathbb{R}^3,\mathbb{R}^3)$. Dann ist

$$\begin{aligned}
&\omega^X=X_1\,dx_1+X_2\,dx_2+X_3\,dx_3\,,\\
&d\omega_X=\Big(\frac{\partial X_2}{\partial x_1}-\frac{\partial X_1}{\partial x_2}\Big)dx_1\wedge dx_2+\Big(\frac{\partial X_3}{\partial x_2}-\frac{\partial X_2}{\partial x_3}\Big)dx_2\wedge dx_3+\Big(\frac{\partial X_3}{\partial x_1}-\frac{\partial X_1}{\partial x_3}\Big)dx_1\wedge dx_3\,,\\
&v\lrcorner(dx_1\wedge dx_2\wedge dx_3)=v_1\,dx_2\wedge dx_3-v_2\,dx_1\wedge dx_3+v_3\,dx_1\wedge dx_2\,.
\end{aligned}$$

Daraus folgt

$$\mathrm{rot}(X)=\Big(\frac{\partial X_3}{\partial x_2}-\frac{\partial X_2}{\partial x_3},\frac{\partial X_1}{\partial x_3}-\frac{\partial X_3}{\partial x_1},\frac{\partial X_2}{\partial x_1}-\frac{\partial X_1}{\partial x_2}\Big).$$

12.5.9. **Satz.** *Sei U eine offene Teilmenge des $\mathbb{R}^3$. Das folgende Diagramm ist kommutativ:*

$$\begin{tikzcd}
\mathcal{C}^\infty(U) \arrow[r, "\mathrm{grad}"] \arrow[d, "\mathrm{Id}"] & \mathcal{X}(U) \arrow[r, "\mathrm{rot}"] \arrow[d, "X\mapsto\omega^X"] & \mathcal{X}(U) \arrow[r, "\mathrm{div}"] \arrow[d, "X\mapsto X\lrcorner\omega_{\mathbb{R}^3}"] & \mathcal{C}^\infty(U) \arrow[d, "f\mapsto f\omega_{\mathbb{R}^3}"] \\
\mathcal{C}^\infty(U) \arrow[r, "d"'] & \Omega^1(U) \arrow[r, "d"'] & \Omega^2(U) \arrow[r, "d"'] & \Omega^3(U)
\end{tikzcd}$$

Daraus kann man viele Eigenschaften von rot und div ablesen; wir haben $d\circ d=0$, und falls U sternförmig ist, so ist die Sequenz in der unteren Zeile des Diagramms exakt. Z.B.

$$\begin{aligned}
X=\mathrm{grad}\,f&\iff\omega^X=df\\
Y=\mathrm{rot}\,X&\iff d\omega^X=Y\rfloor\omega_{\mathbb{R}^3}\\
f=\mathrm{div}\,X&\iff d(X\rfloor\omega_{\mathbb{R}^3})=f\omega_{\mathbb{R}^3}\,.
\end{aligned}$$

Ein Vektorfeld X heißt ***Gradientenfeld (Potentialfeld)***, wenn es $f\in\mathcal{C}^\infty(U)$ existiert, mit $X=\mathrm{grad}\,f$; f heißt ***Potential*** (in der Physik schreibt man $X=-\mathrm{grad}\,f$).

12.5.10. **Folgerung.** *Sei U eine offene Teilmenge des $\mathbb{R}^3$.*

(i) *Ist $X\in\mathcal{X}(U)$ ein Gradientenfeld, so ist X rotationsfrei.*

(ii) *Falls U sternförmig ist, so gilt: $X\in\mathcal{X}(U)$ ist ein Gradientenfeld genau dann, wenn X rotationsfrei ist.*

In physikalischer Sprache gilt für ein Kraftfeld $X\in\mathcal{X}(U)$:

(i) Ist X konservativ, so ist X wirbelfrei.

(ii) Falls U sternformig ist, so gilt: X ist konservativ genau dann, wenn X wirbelfrei ist.

12.6. **Kurvenintegrale von 1-Formen. Bogenlänge.**

12.6.1. **Definition.** Eine stetige Abbildung $\gamma:[a,b]\longrightarrow\mathbb{R}^n$ heißt (parametrisierte) Kurve. Die Kurve γ heißt geschlossen, falls $\gamma(b)=\gamma(a)$. Ist γ glatt und $\gamma'(x)\neq0$ für alle $x\in[a,b]$, so heißt γ reguläre Kurve (d.h. γ ist eine Immersion). Eine Kurve $\gamma:[a,b]\longrightarrow\mathbb{R}^n$ heißt stückweise glatt, wenn eine Zerlegung $a<t_0<t_1<\ldots<t_k=b$ existiert, so dass $\gamma|_{[t_{j-1},t_j]}$ für jedes $j=1,\ldots,k$ glatt ist.

Sei X ein Kraftfeld in $\mathbb{R}^3$ und $\gamma:[a,b]\longrightarrow\mathbb{R}^3$ eine stückweise glatte glatte Kurve. Wir definieren die von X geleistete ***Arbeit***, die verrichtet wird, um das Teilchen längs γ von $\gamma(a)$ nach $\gamma(b)$ zu verschieben. Zunächst approximieren wir γ mit einen Streckenzug $[p_0,p_1,\ldots,p_k]$, wobei $p_j=\gamma(t_j)$ und $a=t_0<t_1<\ldots<t_k=b$ eine Zerlegung von $[a,b]$ ist. Das Skalarprodukt $\langle X(\gamma(t_j)),\gamma(t_j)-\gamma(t_{j-1})\rangle$ ist die Arbeit der (konstanten) Kraft $X(\gamma(t_j))$ bei der geradlinigen Verschiebung des Teilchens von $\gamma(t_{j-1})$ nach $\gamma(t_j)$ („Arbeit = Kraft × Weg“, oder genauer: Arbeit = Kraft × Verschiebung in Richtung der Kraft). Die Arbeit entlang $[p_0,\ldots,p_k]$ erklären wir als $\sum_{j=1}^k\langle X(p_j),p_j-p_{j-1}\rangle=\sum_{j=1}^k\langle X(\gamma(t_j)),\gamma(t_j)-\gamma(t_{j-1})\rangle$. Wenn die Norm der Zerlegung gegen Null strebt, konvergiert der Streckenzug in der Supremumsnorm gegen γ. Dann approximieren die Vektoren $\gamma(t_j)-\gamma(t_{j-1})$ die Vektoren $\gamma'(t_j)(t_j-t_{j-1})$, und die Summen $\sum\langle X(p_j),p_j-p_{j-1}\rangle$ approximieren $\sum\langle X(\gamma(t_j)),\gamma'(t_j)(t_j-t_{j-1})\rangle$, welches Riemannsche Summen der Funktion $t\mapsto\langle X(\gamma(t)),\gamma'(t)\rangle$ sind. Die Arbeit wird dann als $\int_a^b\langle X(\gamma(t)),\gamma'(t)\rangle dt$ definiert. Durch die gleiche Formel definieren wir allgemeiner die ***Zirkulation*** eines Vektorfeldes in $\mathbb{R}^n$ entlang einer Kurve γ.

Es ist vorteilhaft, mit Formen und nicht mit Vektorfeldern zu arbeiten. Dabei benutzt man die Korrespondenz $\mathcal{X}(\mathbb{R}^n)\cong\Omega^1(\mathbb{R}^n)$, $X=\sum X_i\frac{\partial}{\partial x_i}\mapsto\omega^X=\sum X_i dx_i$.

12.6.2. **Definition.** Seien $a < b$ in $\mathbb{R}$ und $\alpha \in \Omega^1([a,b])$, $\alpha = f\,dt$. Setze:

$$\int_a^b \alpha := \int_a^b f(t)dt,$$

wobei die rechte Seite das Regelintegral der Funktion f ist.
Sei $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ glatt, $\omega \in \Omega^1(\mathbb{R}^n)$. Das ***Integral von*** ω ***längs*** γ ist erklärt als

$$\int_\gamma \omega := \int_a^b \gamma^* \omega = \int_a^b \omega(\gamma(t))(\gamma'(t)) = \int_a^b \sum_{i=1}^n \omega_i(\gamma(t))\gamma_i'(t)dt.$$

Ist $\gamma : [a,b] \longrightarrow \mathbb{R}^k$ stückweise glatt wie in Definition 12.6.1, dann definieren wir

$$\int_\gamma \omega := \sum_{j=1}^k \int_{t_{j-1}}^{t_j} \gamma^* \omega.$$

Die Definition ist von der Wahl der Zerlegung unabhängig.

Für das Kurvenintegral gelten folgende **Rechenregeln**:
(1) Ist $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ stückweise glatt, und sind $\omega_1, \omega_2 \in \Omega^1(\mathbb{R}^n)$, $\alpha_1, \alpha_1 \in \mathbb{R}$, so gilt

$$\int_\gamma (\alpha_1\omega_1 + \alpha_2\omega_2) = \alpha_1 \int_\gamma \omega_1 + \alpha_2 \int_\gamma \omega_2.$$

(2) Sei $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ stückweise glatt und $c \in (a,b)$. Dann gilt

$$\int_\gamma \omega = \int_{\gamma|[a,c]} \omega + \int_{\gamma|[c,b]} \omega$$

(3) Sei $\omega \in \Omega^1([a,b])$, $\omega = f(s)ds$ und sei $\varphi : [\alpha,\beta] \longrightarrow [a,b]$ ein Diffeomorphismus. Dann gilt:

$$\int_a^b f(s)ds = \int_\alpha^\beta (f \circ \varphi)(t)\,|\varphi'(t)|\,dt, \quad \text{also} \quad \int_a^b \omega = \varepsilon(\varphi) \int_\alpha^\beta \varphi^* \omega,$$

wobei $\varepsilon(\varphi) = \operatorname{sgn} \varphi'$.
Ist $\gamma : [a,b] \longrightarrow U$ stückweise glatt und $\omega \in \Omega^1(U)$, so gilt

$$\int_{\gamma\circ\varphi} \omega = \varepsilon(\varphi) \int_\gamma \omega.$$

Die Kurve $\gamma \circ \varphi$ heißt eine ***Umparametrisierung*** von γ.
Insbesondere sei $\gamma^- : [a,b] \longrightarrow \mathbb{R}^n$, $\gamma^-(t) = \gamma(b+a-t)$ die umorientierte Kurve. Dann gilt

$$\int_{\gamma^-} \omega = -\int_\gamma \omega.$$

12.6.3. **Satz.** *Sei $U \subset \mathbb{R}^n$ offen, $\omega \in \Omega^1(U)$ eine exakte 1-Form und f eine Stammfunktion zu ω. Sei $\gamma : [a,b] \longrightarrow U$ glatte Kurve. Dann gilt*

$$\int_\gamma \omega = f(\gamma(b)) - f(\gamma(a)).$$

Ist γ geschlossen, so gilt $\int_\gamma \omega = 0$. Sind γ_1, γ_2 zwei Kurven mit gleichen Anfangspunkten und gleichen Endpunkten (d.h. $\gamma_1(a) = \gamma_2(a)$, $\gamma_1(b) = \gamma_2(b)$), so gilt $\int_{\gamma_1} \omega = \int_{\gamma_2} \omega$.

Wenn die letzte Eigenschaft zutrifft, sagen wir, dass das Kurvenintegral von ω in U ***wegunabhängig*** ist. Ist das der Fall, so setzen wir $\int_{x_0}^x \omega := \int_\gamma \omega$ für eine *beliebige* Kurve $\gamma : [a,b] \longrightarrow U$ mit $\gamma(a) = x_0$, $\gamma(b) = x$.

12.6.4. **Satz** (Hauptsatz über Kurvenintegrale)**.** *Seien $U \subset \mathbb{R}^n$ offen und zusammenhängend, und sei $\omega \in \Omega^1(U)$ eine 1-Form. Dann sind äquivalent*:

(i) *ω ist exakt.*
(ii) *$\int_\gamma \omega = 0$ für jede geschlossene stückweise glatte Kurve in U.*
(iii) *Das Kurvenintegral von ω in U ist wegunabhängig.*

Ist das der Fall, dann ist eine Stammfunktion von ω durch $f(x) = \int_{x_0}^x \omega$ gegeben, wobei $x_0 \in U$ fest gewählt ist.

Beispiel: Sei $U := \mathbb{R}^2 \setminus \{0\}$, und sei $\gamma : [0,2\pi] \longrightarrow U$ die geschlossene Kurve mit $\gamma(\vartheta) = (\cos\vartheta, \sin\vartheta)$. Sei $\omega = \dfrac{-y dx + x dy}{x^2+y^2} \in \Omega^1(U)$. Es gilt:

$$\int_\gamma \omega = \int_0^{2\pi} \gamma^*\omega = \int_0^{2\pi} dt = 2\pi.$$

Also ist ω nicht exakt, obwohl ω geschlossen ist. Man kann aber auf jeder sternförmigen Teilmenge von $\mathbb{R}^2 \setminus \{0\}$ eine Stammfunktion zu ω finden. Beispielsweise ist auf $U = \mathbb{R}^2 \setminus \{(x,y) : x \geqslant 0\}$ die Funktion $f = \arg(x,y)$ eine Stammfunktion zu ω.

12.6.5. **Definition.** Sei C eine 1-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$. Sei $\gamma : (a,b) \longrightarrow C$ eine Parametrisierung so, dass $C \setminus \gamma((a,b))$ endlich ist. Wir setzen

$$l(C) := \int_a^b \|\gamma'(t)\| \, dt \in \mathbb{R} \cup \{\infty\}.$$

$l(C)$ heißt die ***Länge*** von C.

12.6.6. **Bemerkung.** Laut Beispiel 12.4.2 ist C orientierbar. Wähle eine Orientierung und sei ds das entsprechende Linienelement auf C (siehe (12.31)). Dann gilt für eine positiv orientierte Parametrisierung γ:

$$l(C) = \int_a^b \gamma^*\omega_C = \int_C \omega_C = \int_C ds.$$

Dies zeigt, dass $l(C)$ von der Parametrisierung γ nicht abhängt.

13. MASS–UND INTEGRATIONSTHEORIE

Wir haben bisher das Regelintegral für reell- (oder komplex-)wertige Funktionenauf einem Intervall kennen gelernt (Kapitel 6). Das Regelintegral diente unter anderem zur Flächenberechnung. Möchte man auch Volumina berechnen, so scheint eine Erweiterung der Integration auf Funktionen von mehreren Variablen wünschenswert. Darüberhinaus, möchte man in der Stokastik Funktionen definiert auf allgemeine Räume (nicht unbedingt Teilmengen von $\mathbb{R}^n$), dir sog. Zufallsvariablen, integrieren. Wir müssen daher das Integralbegriff verallgemeinern. Gleichzeitig werden wir ein allgemeiner Integralbegriff sogar auf einem Intervall in $\mathbb{R}$ einführen.

Es gibt verschiedene Integralbegriffe: das Regelintegral, das Riemann-Integral, das Lebesgue-Integral. Für Treppenfunktionen, für stetige und monotone Funktionen, liefern diese Integrale denselben Wert. Sie unterscheiden sich aber hinsichtlich der jeweiligen Menge der "integrierbaren"Funktionen; diese Menge vergrößert sich bei den obigen drei Integralbegriffen in der angegebenenReihenfolge.

Es geht nicht nur darum, dass das Lebesgue-Integral algemeiner ist. Sie ist auch praktischer und leichter anzuwenden. In vielen Anwendungen der Analysis möchte man Grenzwertprozessein Funktionenräumen, zum Beispiel im Raum der integrierbaren Funktionen, durchführen. Ziel ist, dass unter möglichst allgemeinen Voraussetzungen gilt:

$$\lim_{n\to\infty}\int f_n = \int \lim_{n\to\infty} f_n\,. \tag{13.1}$$

Arbeiten wir mit dem Riemann-Integral, so gilt (13.1) wenn $f_n \to f$ gleichmäßig konvergiert. Das ist eine sehr starke Voraussetzung. Hingegen liefern die Konvergenzsätze der Lebesgue-Theorie die Gleichung (13.1) unter sehr milden Voraussetzungen (monotone oder dominierte Konvergenz). Deshalb hat die Lebesgue-Theorie die moderne Analysis, Funktionalanalysis (darunter die Quantenmechanik) und Wahrscheinlichkeitstheorie überhapt erst ermöglicht.

Sei $f:[a,b]\to\mathbb{R}$ eine stetige Funktion. Das Regelintegral von f können wir mittels Approximation mit Riemannschen-Summen berechnen (siehe § 6.8), die man durch Zerlegung des Definitionsberech $[a,b]$ erhält. Sei $Z=\{a=x_0<x_1<\ldots<x_{n-1}<x_n=b\}$ eine Zerlegung von $[a,b]$. Seien $\xi_k\in[x_{k-1},x_k]$ für $k=1,\ldots,n$ Stützpunkte. Wir definieren die Funktion $\varphi=\varphi_Z=\sum_{k=1}^n f(\xi_k)\mathbb{1}_{[x_{k-1},x_k)}$, wobei $\mathbb{1}_A$ die charakteristische Funktion der Menge A ist: $\mathbb{1}_A(x)=1$ falls $x\in A$ und $\mathbb{1}_A(x)=0$ falls $x\notin A$. Wir definieren das Integral von $\mathbb{1}_{[x_{k-1},x_k)}$ durch $\int_a^b \mathbb{1}_{[x_{k-1},x_k)}\,dx=x_k-x_{k-1}$ (die Länge von $[x_{k-1},x_k)$) und durch Linearität setzen wir $\int_a^b\varphi\,dx=\int_a^b\sum_{k=1}^n f(\xi_k)\mathbb{1}_{[x_{k-1},x_k)}=\int_a^b\sum_{k=1}^n f(\xi_k)(x_k-x_{k-1})$ (Riemannsche Summen).

Die charakteristische Funktion $\mathbb{1}_{\mathbb{Q}}$ ist auf $[0,1]$ nicht Riemann-integrierbar. Das liegt daran, dass man bei noch so feiner Unterteilung des Intervalls $[0,1]$ keinen Zwischenpunkt ξ_k aus irgendeinem Intervall $[x_k,x_{k+1}]$ als Argument für eine Näherung der Funktionswerte im gesamten Teilintervall nehmen kann. Eigentlich ist also der Riemannsche Integralbegriff auf stetige Funktionen zugeschnitten: ein Kriterium von Lebesgue besagt, dass f ist genau dann Riemann-integrierbar, wenn f beschränkt und die Menge der Unsetigkeitsstellen von f eine Lebesgue-Nullmenge ist.

Henri Lebesgue schlug ein Integrationsverfahren vor, bei dem von einer Unterteilung des Intervalls $[m,M]$ ausgegangen wird, in dem alle Funktionswerte liegen. Sei $Z=\{m=y_0<y_1<\ldots<y_{n-1}<y_n=M\}$ eine Zerlegung. Wir betrachten nun die Funktion $\varphi=\sum_{k=1}^n y_k\mathbb{1}_{f^{-1}[y_{k-1},y_k)}$. Wir möchten nun das Integral von $\mathbb{1}_{f^{-1}[y_{k-1},y_k)}$ definieren als $\int_a^b\mathbb{1}_{f^{-1}[y_{k-1},y_k)}\,dx=\lambda(f^{-1}[y_{k-1},y_k))$, wobei $\lambda(f^{-1}[y_{k-1},y_k))$ die „Länge" (das „Maß") der Menge $\lambda(f^{-1}[y_{k-1},y_k))$ bedeutet. Dann setzen wir $\int_a^b\varphi\,dx=\int_a^b\sum_{k=1}^n f(\xi_k)\mathbb{1}_{f^{-1}[y_{k-1},y_k)}=\int_a^b\sum_{k=1}^n y_k\lambda(f^{-1}[y_{k-1},y_k))$ (Lebesgue-Summen).

Lebesgue beschreibt sein Verfahren so: „Imaginez que je doive payer une certaine somme; je peux sortir les pièces de mon porte-monnaie comme elles viennent pour arriver à la somme indiquée, ou sortir toutes les pièces et les choisir selon leur valeur. La première méthode est l'intégrale de Riemann, la deuxième correspond à mon intégrale."

Man kann zeigen, dass jede beschränkte Funktion, für die die "Maße" $\lambda(f^{-1}[y_{k-1},y_k))$ erklärt sind, in dem Sinne integrierbar sind, dass die Lebesgue-Summen bei immer feiner werdender Unterteilung von $[m,M]$ gegen einen gemeinsamen Wert, das Lebesgue-Integral von f, konvergieren. Da jede Riemann-integrierbare Funktion auch in diesem Sinne integrierbar ist, ist dann der Integralbegriff auf eine weitaus größere Funktionenklasse verallgemeinert worden. Z.B. ist $\mathbb{1}_{\mathbb{Q}}$ ist auf $[0,1]$ Lebesgue-integrierbar mit Lebesgue-Integral Null.

Wir müssen nun die „Länge“ für $f^{-1}([y_{k-1}, y_k))$ definieren. Diese Menge ist i.A. kein Intervall aber $f^{-1}([y_{k-1}, y_k)) = f^{-1}(y_{k-1}) \cup f^{-1}((y_{k-1}, y_k))$. Weil f stetig ist, ist $f^{-1}((y_{k-1}, y_k))$ offen und $f^{-1}(y_{k-1})$ abgeschlossen. Für eine offene Menge $D \subset \mathbb{R}$ können wir die Länge so definieren: D ist eine höchstens abzählbare disjunkte Vereinigung von Intervalle $(I_k)_k$ und wir setzen $\lambda(D) := \sum_k \lambda(I_k)$. Die Länge einer abgeschlossenen Menge F definieren wir ähnlich: $F = \cap_k F_k$ wobei F_k endliche Vereinigungen von Intervallen sind und $F_k \supset F_{k+1}$. Wir setzen $\lambda(F) := \lim \lambda(F_k)$.

Wenn aber f nicht stetig ist, kann die Menge $f^{-1}([y_{k-1}, y_k))$ komplizierter sein. Deshalb werden wir als erstes eine große Klasse von Teilmengen in $\mathbb{R}$ studieren, die durch sukzessive abzählbare Vereinigungen und Komplemente aus der Menge der Intervalle erhalten werden kann. Diese Mengen heißen Borelmengen. Für die Integration brauchen wir noch kompliziertere Mengen, die Lebesgue-messbaren Mengen. Alle diese Mengen bilden so genannte σ-Algebren von meßbaren Mengen, die wir nun einführen. Danach werden wir einen entsprechen Begriff von Maß für diese Mengen definieren. Wir werden schließlich das Integral für *meßbaren Funktionen* definieren. Für eine Funktion $f : \mathbb{R} \to \mathbb{R}$ bedeutet das die die Eigenschaft, dass $f^{-1}([y_{k-1}, y_k))$ Lebesgue-messbar ist. Allerdings werden wir das Integral nicht mit Hilfe der Lebesgue-Summen einführen, sondern im Wesentlichen mit Hilfe der monotonen Folgen von einfachen Funktionen.

Henri Léon Lebesgue wurde am 28.6.1875 in Beauvais (Frankreich) geboren. Nach dem Studium an der École Normale Supérieure von 1894 bis 1897 unterrichtete er 1899-1902 als Gymnasiallehrer in Nancy. Dort gelang ihm auch die Entdeckung des nach ihm benannten Integrals (Sur une généralisation de l'intégrale définie, Comptes Rendus 1901). Er promovierte darüber 1902 (Intégrale, Longueur, Aire, Annali di Matematica) 1919 wurde er zum Professor an der Sorbonne ernannt und schließlich 1921 Professor am Collège de France, wo er bis zum Ende seines Lebens lehrte.

13.1. **σ–Algebren und Maße.** Sei X eine nicht–leere Menge, $\mathcal{P}(X)$ ihre Potenzmenge, d.h. die Menge aller Teilmengen von X. Eine Teilmenge von $\mathcal{P}(X)$ (eine Menge deren Elemente Teilmenge von X sind) wird Mengensystem genannt. Für $A, B \in \mathcal{P}(X)$ bezeichnen wir $\complement A := \{x \in X : x \notin A\}$ das Komplement von A und $A \setminus B := A \cap \complement B$. Die Gesetze von de Morgan lauten:

$$\complement(A \cap B) = \complement A \cup \complement B\,, \quad \complement(A \cup B) = \complement A \cap \complement B$$

Allgemeiner gilt für eine Familie $(A_i)_{i \in I}$ von Teilmengen von X

$$\complement \bigcap_{i \in I} A_i = \bigcup_{i \in I} \complement A_i\,, \quad \complement \bigcup_{i \in I} A_i = \bigcap_{i \in I} \complement A_i$$

Wir sagen, dass eine Teilmenge A die ***disjunkte Vereinigung*** der Familie $(A_i)_{i \in I}$ ist, geschrieben

$$A = \bigsqcup_{i \in I} A_i\,,$$

falls $A = \bigcup_{i \in I} A_i$ und $A_i \cap A_j = \emptyset$ für $i \neq j$. Wir sagen auch, dass $(A_i)_{i \in I}$ eine ***Zerlegung*** von A ist.

Sei $f : X \to Y$ eine Abbildung. Wir definieren das Bild von $A \subset X$ als

$$f(A) := \{y \in Y : \exists x \in A : y = f(x)\}$$

und das Urbild von $B \subset Y$ als

$$f^{-1}(B) := \{x \in X : \exists y \in B : y = f(x)\}.$$

Das Urbild ist verträglich mit der mengentheoretischen Operationen:

$$f^{-1}\big(\bigcap_{i \in I} A_i\big) = \bigcap_{i \in I} f^{-1}(A_i)\,, \quad f^{-1}\big(\bigcup_{i \in I} A_i\big) = \bigcup_{i \in I} f^{-1}(A_i)\,, \quad f^{-1}(\complement B) = \complement f^{-1}(B)\,.$$

Das Bild ist i.A. nur mit der Vereinigung verträglich:

$$f\big(\bigcup_{i \in I} A_i\big) = \bigcup_{i \in I} f(A_i)\,, \quad f\big(\bigcap_{i \in I} A_i\big) \subset \bigcap_{i \in I} f(A_i)$$

und die letzte Inklusion kann echt sein (Beispiel?).

13.1.1. **Definition.** Ein nichtleeres Mengensystem $\mathcal{A} \subset \mathcal{P}(X)$ heißt ***Algebra*** falls

(a) $X \in \mathcal{A}$,
(b) $A \in \mathcal{A} \Rightarrow \complement A = X \setminus A \in \mathcal{A}$,
(c) $A_1, \dots, A_n \in \mathcal{A} \Rightarrow A_1 \cup \ldots \cup A_n \in \mathcal{A}$.

Ein nichtleeres Mengensystem $\mathcal{A} \subset \mathcal{P}(X)$ heißt σ–***Algebra*** falls $\mathcal{A}$ eine Algebra ist und zusätzlich gilt:

(c)' $(A_n)_{n \in \mathbb{N}}$ Folge in $\mathcal{A} \Rightarrow \cup A_n \in \mathcal{A}$.

Ein Paar $(X,\mathcal{A})$ von einer nicht-leeren Menge X und einer σ–Algebra $\mathcal{A}\subset\mathcal{P}(X)$ heißt ***messbarer Raum***. Eine Teilmenge $A\in\mathcal{A}$ heißt ***$\mathcal{A}$-messbare Menge***.

Das „σ“ in der Bezeichnung σ-Algebra ist eine Abkürzung für „abzählbar“ und soll auf das Axiom (c)’ hinweisen, genauer darauf, dass die abzählbare Vereinigung von messbaren Mengen wieder messbar ist.

Eigenschaften der Algebren: Sei $\mathcal{A}\subset\mathcal{P}(X)$ eine Algebra. Dann gilt

(i) $\emptyset,\, X\in\mathcal{A}$,
(ii) $A,B\in\mathcal{A}\Rightarrow A\cup B,\, A\cap B,\, A\setminus B\in\mathcal{A}$,
(iii) $A_1,\dots,A_n\in\mathcal{A}\Rightarrow A_1\cap\ldots\cap A_n\in\mathcal{A}$.
Ist $\mathcal{A}$ eine σ–Algebra so gilt zusätzlich:
(iii)’ $(A_n)_n$ Folge in $\mathcal{A}\Rightarrow\cap_{n\in\mathbb{N}}A_n\in\mathcal{A}$.

Beispiele von Algebren:

(1) Betrachten wir die Menge $\mathcal{Q}_1$ aller Intervallen in $\mathbb{R}$. Dabei lassen wir leere, beschränkte, unbeschränkte, offene, abgeschlossene, halboffene Intervalle zu. Dann ist

$$\mathcal{R}_1:=\{\cup_{k=1}^n I_k : I_k\in\mathcal{Q}_1, n\in\mathbb{N}\}$$

eine Algebra, die Algebra der endlichen Vereinigungen von Intervalle.

13.1.2. ***Aufgabe***. Zeigen Sie:
(i) $\mathcal{R}_1$ ist eine Algebra.
(ii) Jedes $A\in\mathcal{R}_1$ besitzt eine Darstellung als disjunkte endliche Vereinigung von Intervalle.

(2) Eine menge $Q\subset\mathbb{R}^n$ heißt ***Quader*** falls $Q=I_1\times I_2\times\ldots\times I_n$, wobei $I_k\in\mathcal{Q}_1$, $1\leqslant k\leqslant n$, Intervalle in $\mathbb{R}$ sind. Wir vereinbaren dabei, dass $\emptyset\times A=A\times\emptyset=\emptyset$. Ein Quader heißt *entartet*, falls einer der Intervalle Einpunktig oder leer ist. Die Menge aller Quader wird mit $\mathcal{Q}_n$ bezeichnet. Das Mengensystem

$$\mathcal{R}_n:=\{\cup_{j=1}^k Q_j : Q_j\in\mathcal{Q}_n, k\in\mathbb{N}\}$$

eine Algebra, die Algebra der endlichen Vereinigungen von Quader. Eine Menge $R\in\mathcal{R}_n$ heißt ***Figur***. Es gilt:
(i) $\mathcal{R}_n$ ist eine Algebra.
(ii) Jedes $R\in\mathcal{R}_n$ besitzt eine Darstellung als disjunkte endliche Vereinigung von Quader.

Beispiele von σ–Algebren:

(1) $\{X,\emptyset\}$, $\mathcal{P}(X)$ sind σ–Algebren,
(2) $(\mathcal{A}_i)_{i\in I}$ familie von σ–Algebren $\Rightarrow\mathcal{A}=\cap_{i\in I}A_i$ σ–Algebra.
(3) Ist $Y\subset X$ und $\mathcal{A}\subset\mathcal{P}(X)$ eine σ–Algebra, so ist $\mathcal{A}_Y:=\{A\cap Y : A\in\mathcal{A}\}\subset\mathcal{P}(Y)$ eine σ–Algebra, genannt die Spur von $\mathcal{A}$ in Y.

13.1.3. **Satz.** *Sei $\mathcal{E}\subset\mathcal{P}(X)$ ein beliebiges nichtleeres Mengensystem. Dann ist*

$$\mathcal{A}_\sigma(\mathcal{E})=\cap\{\mathcal{A}:\mathcal{A}\ \sigma\text{–}Algebra\ \mathcal{E}\subset\mathcal{A}\}$$

eine σ–Algebra, die man die ***von $\mathcal{E}$ erzeugte σ–Algebra*** *nennt. Sie ist die kleinste σ–Algebra die $\mathcal{E}$ enthält, d.h. $\mathcal{E}\subset\mathcal{A}$ und ist $\mathcal{B}$ eine σ–Algebra mit $\mathcal{E}\subset\mathcal{B}$, dann ist $\sigma(\mathcal{E})\subset\mathcal{B}$.*

Diese Definition ist sehr einfach aber nicht konstruktiv; $\sigma(\mathcal{E})$ enthält viele Elemente und es ist schwer alle diese Elemente darzustellen oder durch eine konkrete Eigenschaft zu definieren. In der Praxis hat man zwei Typen von Aufgaben, die man mit Hilfe dieser Definition so lösst:

• Zu zeigen, dass eine gegebene Teilmenge $A\subset X$ zu $\sigma(\mathcal{E})$ gehört. Wir zeigen dann, dass $A=\cup_{n\in\mathbb{N}}A_n$ mit $A_n\in\mathcal{A}$ oder $A=\cup_{n\in\mathbb{N}}(\cap_{j\in\mathbb{N}}A_{nj})$ mit $A_{nj}\in\mathcal{A}$ oder $A=\cup_{n\in\mathbb{N}}(\cap_{j\in\mathbb{N}}(\cup_{k\in\mathbb{N}}A_{njk}))$ mit $A_{njk}\in\mathcal{A}$ usw. Normalerweise sind solchen Aufgaben leicht.

• Zu zeigen, dass alle $A\in\sigma(\mathcal{E})$ eine gegebene Eigenschaft (E) haben. Dann zeigt man:

(i) Alle $A\in\mathcal{E}$ haben die Eigenschaft (E),
(ii) $A\in\mathcal{P}(X)$ hat (E), dann $\complement A$ hat (E),
(iii) $A_1,\dots,A_n,\dots$ haben (E), dann $\cup_{n\in\mathbb{N}}A_n$ hat (E).

Es folgt, dass $\mathcal{A}:=\{A\subset X : A$ hat die Eigenschaft $(E)\}$ eine σ–Algebra ist und $\mathcal{E}\subset\mathcal{A}$, also $\mathcal{A}\subset\sigma(\mathcal{E})$.

13.1.4. **Definition.** Sei $(X,\mathcal{O})$ ein topologischer Raum. Wir bezeichnen mit $\mathcal{O}(X)$ das Mengensystem der Offenen Teilmengen. Die σ-Algebra

$$\mathcal{B}(X) := \sigma(\mathcal{O}(X))$$

heißt die ***Borelsche σ-Algebra*** des topologischen Raumes $(X,\mathcal{O})$. Eine Menge $A \in \mathcal{B}(X)$ heißt die ***Borelmenge***.

An vielen Stellen der Integrationstheorie genügt, mit Würfeln statt mit Quader zu argumentieren. Wir geben eine Zerlegung des $\mathbb{R}^n$ in beliebig feinen Netzen von Würfeln. Betrachte zunächst die Familie $\mathcal{I}(k) = \{[\frac{m}{2^k}, \frac{m+1}{2^k}) : m \in \mathbb{Z}\}$; dann gibt es die Zerlegung $\mathbb{R} = \cup_{m\in\mathbb{Z}}[\frac{m}{2^k}, \frac{m+1}{2^k})$ in halboffenen Intervalle der Länge $\frac{1}{2^k}$. Die Familie $\mathcal{W}(k) := \{W = I_1 \times \ldots \times I_n : I_1, \ldots, I_n \in \mathcal{I}(k)\}$ ist eine abzählbare Familie von Würfeln der Seitenlänge $\frac{1}{2^k}$, die eine Zerlegung des $\mathbb{R}^n$ liefert. Definiere $\mathcal{W}_n := \cup_{k\in\mathbb{N}_0}\mathcal{W}(k)$.

13.1.5. **Satz.** *Die folgende Familien erzeugen $\mathcal{B}(\mathbb{R}^n)$:*
(i) Die abgeschlossenen Mengen.
(ii) Die Figuren $\mathcal{R}_n$
(iii) Die Quader $\mathcal{Q}_n$ (und sogar die Quader mit rationalen Seiten und Zentrum)
(iv) Die halboffenen Würfeln $\mathcal{W}_n$.

Eine Beschreibung der Borelschen σ–Algebra. Wir führen die folgenden Klassen von Mengen:
• G_δ bezeichnet alle Durchschnitte von abzählbar vielen offenen Mengen; jede aggeschlossene Menge ist in G_δ. Die Familie G_δ ist stabil bei abzählbare Durchschnitte und endliche Vereinigungen (aber nicht abzählbare Vereinigungen).
• F_σ bezeichnet alle Durchschnitte von abzählbar vielen offenen Mengen; jede offene Menge ist in F_σ. Die Familie F_σ ist stabil bei abzählbare Vereinigungen und endliche Durchschnitte (aber nicht abzählbare Durchschnitte).
• $G_{\delta\sigma}$ bezeichnet alle Vereinigungen von abzählbar vielen G_δ Mengen.
• $F_{\sigma\delta}$ bezeichnet alle Durchschnitte von abzählbar vielen F_σ Mengen.
Wir konstruieren induktiv Familien $G_\delta, F_\sigma, G_{\delta\sigma}, F_{\sigma\delta}, G_{\delta\sigma\delta}, F_{\sigma\delta\sigma}$, ... Die Vereinigung dieser Klassen ist jedoch echt kleiner als die σ–Algebra der Borelmengen. Um eine vollständige Beschreibung zu erreichen, müssen wir die *transfinite Induktion* benutzen (siehe [7, I,(4.2)]). Glücklicherweise brauchen wir in der Praxis nicht die Beschreibung aller Borelmengen. Normalerweise genügt es „vernünftig definierte“ Teilmengen zu betrachten und man zeigt leicht, dass sie in der Klassen $G_\delta, F_\sigma, G_{\delta\sigma}, F_{\sigma\delta}$, ... liegen. (Sehr selten braucht man mehr als 3 Indizes.)

Wieviele Borelmengen gibt es, z.B. in $\mathbb{R}$? Man kann zeigen, dass die borelsche σ-Algebra gleichmächtig zur Menge der reellen Zahlen ist. Nach dem Satz 1.9 gilt $|\mathbb{R}| < |\mathcal{P}(\mathbb{R})|$, es gibt also nicht-Borelmengen in $\mathbb{R}$. Man kann explizit nicht–Borelmengen konstruieren! Für ein Beispiel von Lusin siehe [10].

Vereinbarung: Rechenregeln in $[0,\infty] = [0,\infty) \cup \{\infty\}$.

$$\begin{gathered} a + \infty = \infty + a = \infty \quad \text{für alle } a \in [0,\infty], \\ a \cdot \infty = \begin{cases} \infty, a > 0 \\ 0, a = 0 \end{cases} . \end{gathered} \tag{13.2}$$

Es folgt, dass Kommutativ-, Assoziativ- und Distributivgesetz für Addition und Multiplikation in $[0,\infty]$ erfüllt sind. Außerdem hat jede nicht-leere Teimenge $A \subset [0,\infty]$ ein Supremum in $[0,\infty]$ und jede monotone Folge in $[0,\infty]$ hat einen Grenzwert in $[0,\infty]$. Aus $0 \leqslant a_n \nearrow a \leqslant \infty$, $0 \leqslant b_n \nearrow b \leqslant \infty$ $(n \to \infty)$ folgt, dass für $n \to \infty$ auch $a_n + b_n \to a + b$ und $a_n \cdot b_n \to ab$.

13.1.6. **Definition.** Sei $\mathcal{A}$ eine Algebra. Ein ***Inhalt*** auf $\mathcal{A}$ ist eine Abbildung $\mu : \mathcal{A} \longrightarrow [0,\infty]$ so, dass

(1) $\mu(\emptyset) = 0$,
(2) μ endlich additiv, d.h. für alle paarweise disjunkten Mengen $A_1, \ldots, A_n \in \mathcal{A}$ gilt

$$\mu(A_1 \cup \ldots \cup A_n) = \mu(A_1) + \ldots + \mu(A_n).$$

Eine Abbildung $\mu : \mathcal{A} \longrightarrow [0,\infty]$ heißt ***σ–Inhalt*** oder ***Prämaß*** falls

(1) $\mu(\emptyset) = 0$,

(2) μ endlich σ-additiv, d.h. für jede Folge $(A_j)_{j\in\mathbb{N}}$ in $\mathcal{A}$, von paarweise disjunkten $\mathcal{A}$-messbaren Teilmengen mit $\cup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ gilt

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=1}^{\infty} \mu(A_n). \tag{13.3}$$

Ein Tripel $(X,\mathcal{A},\mu)$ aus einer nicht-leeren Menge X, einer Algebra $\mathcal{A}$ und einem Prämaß μ auf $\mathcal{A}$, heißt ***Prämaßraum***.
Ein Prämaß auf einer σ–Algebra heißt ***Maß***. Ein Tripel $(X,\mathcal{A},\mu)$ aus einer nicht leeren Menge, einer σ–Algebra und einem Maß heißt ***Maßraum***.
Sei $(X,\mathcal{A},\mu)$ ein Maßraum. Eine meßbare Menge $N \in \mathcal{A}$ mit $\mu(A) = 0$ heißt ***Nullmenge***. Für jedes $x \in X$ sei $E(x)$ eine Aussage. Man sagt $E(x)$ gilt für (μ-) fast alle $x \in X$ oder $A(x)$ gilt (μ-) ***fast überall*** (kurz ***μ–f.ü.***) auf X, falls eine Nullmenge $N \in \mathcal{A}$ existiert, so dass $E(x)$ für alle $x \in X\setminus N$ wahr ist.

Die unendliche Reihe in 13.3 ist folgendermaßen zu interpretieren: $\sum_{n=1}^{\infty}\mu(A_n) = \infty$ falls ein $\mu(A_j) = \infty$ ist, oder falls alle $\mu(A_j) < \infty$ sind, aber die Reihe divergent ist, und $\sum_{n=1}^{\infty}\mu(A_n) < \infty$ falls alle $\mu(A_j) < \infty$ sind, und die Reihe konvergent ist.

13.1.7. **Satz** (Eigenschaften von Inhalten)**.** *Seien $\mathcal{A}$ eine Algebra und $\mu : \mathcal{A} \longrightarrow [0,\infty]$ σ–Inhalt. Dann gilt:*

(i) $A,B \in \mathcal{A},\ A \subset B \Longrightarrow \mu(A) \leqslant \mu(B)$,
(ii) $A,B \in \mathcal{A},\ \mu(A\cup B) = \mu(A)+\mu(B)-\mu(A\cap B)$,
(iii) $\mu(A_1 \cup \ldots \cup A_k) \leqslant \mu(A_1) \cup \ldots \cup \mu(A_k)$.

13.1.8. **Satz** (Eigenschaften von Prämaße)**.** *Seien $\mathcal{A}$ eine Algebra und $\mu : \mathcal{A} \longrightarrow [0,\infty]$ ein Prämaß. Dann gilt:*

(i) $\mu(\cup_{i=1}^{\infty} A_i) \leqslant \sum_{i=1}^{\infty}\mu(A_i)$ (*σ–Subadditivität*),
(ii) $A_1 \subset A_2 \subset \ldots \subset A_n \subset A_{n+1} \subset \ldots \Longrightarrow \mu(\cup A_n) = \lim \mu(A_n)$ (*Stetigkeit von unten*),
(iii) $A_1 \supset A_2 \supset \ldots A_n \supset A_{n+1} \ldots$ *und* $\mu(A_1) < \infty \Longrightarrow \mu(\cap A_n) = \lim \mu(A_n)$ (*Stetigkeit von oben*).

13.1.9. **Folgerung.** *Eine Abzählbare Vereinigung von μ-Nullmengen ist eine μ-Nullmenge.*

In der Tat, aus $\mu(A_i) = 0$ für alle $i \in \mathbb{N}$ folgt $0 \leqslant \mu(\cup_{i=1}^{\infty} A_i) \leqslant \sum_{i=1}^{\infty}\mu(A_i) = 0$.

13.1.10. **Definition.** Ein Inhalt μ auf einer Algebra $\mathcal{A}$ über X heißt ***endllich***, falls $\mu(X) < \infty$, und ***σ-endllich***, falls eine monoton wachsende Folge von Algebraelemente $X_1 \subset X_2 \subset \ldots \subset X_n \subset X_{n+1} \subset \ldots$ existiert mit $X = \cup_{n\in\mathbb{N}} X_n$ und $\mu(X_n) < \infty$ für alle $n \in \mathbb{N}$.

Sei $(X,\mathcal{A},\mu)$ ein Maßraum. Das Maß heißt ***vollständig***, falls jede Teilmenge einer Nullmenge selbst meßbar ist, d.h. ist $A \in \mathcal{A}$ eine Mennge vom Maß $\mu(A) = 0$ und $B \subset A$, so gilt $B \in \mathcal{A}$ (und dann $\mu(B) = 0$, wegen $0 \leqslant \mu(B) \leqslant \mu(A) = 0$).

13.1.11. **Beispiel.**

(1) Das ***Zählmaß*** $\mu : \mathcal{P}(X) \to [0,\infty]$, $\mu(A) = |A|$ (Anzahl der Elemente von A), falls A endlich ist, und $\mu(A) = \infty$, falls unendlich ist. Das Zählmaß ist vollständig. Es ist endlich, falls X endlich ist, und σ-endlich, falls X abzählbar ist.

(2) Das ***Dirac-Maß***. Sei $a \in X$ fest und $\delta_a : \mathcal{P}(X) \to [0,\infty]$, $\delta_a(A) = 1$, falls $a \in A$ und $\delta_a(A) = 0$, falls $a \notin A$.

(3) Das ***elementar-geometrische Volumen***. Für ein Intervall $I \subset \mathbb{R}$ definieren wir die Länge von I durch $\lambda_1(I) := \sup I - \inf I \in [0,\infty]$. Für ein Quader $Q = I_1 \times \ldots \times I_n$ definieren wir sein elementar-geometrisches Volumen durch $\lambda_n(Q) := \lambda_1(I_1)\lambda_1(I_2)\ldots\lambda_1(I_n) \in [0,\infty]$ (das Produkt der Seitenlängen). Dabei beachten wir die Produktregeln in $[0,\infty]$. Ist $R \in \mathcal{R}_n$ eine Figur in $\mathbb{R}^n$, so gibt es eine Zelegung $R = Q_1 \sqcup \ldots \sqcup Q_k$ in Quader. Wir setzen $\lambda_n(R) := \lambda_n(Q_1)+\lambda_n(Q_2)+\ldots+\lambda_n(Q_k)$, genannt auch das elementar-geometrische Volumen von R.

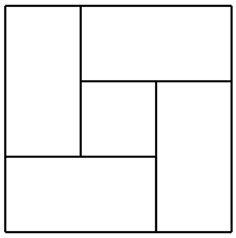
13.1.12. **Satz.** *Das elementar-geometrische Volumen $\lambda : \mathcal{R}_n \to [0,\infty]$ ist ein wohldefiniertes σ-endliches Prämaß, genannt Lebesguesches-Prämaß auf der Algebra der Figuren.*

Beweis: 1. Schritt. Seien $Q,K,L \in \mathcal{Q}_n$ mit $Q = Q_1 \times \ldots \times Q_n$, $K = K_1 \times \ldots \times K_n$, $L = L_1 \times \ldots \times L_n$, wobei Q_i, K_i, L_i Intervalle in $\mathbb{R}$ sind. Nehmen wie an, dass $Q = K \sqcup L$. Dann haben $\overline{K}$ und $\overline{L}$ eine gemeinsame

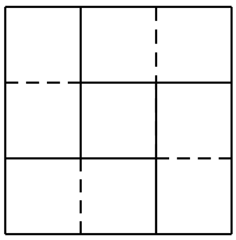
$(n-1)$-dimensionale Seite, d.h. es gibt $q \in \{1,\dots,n\}$, so dass $Q_i = K_i = L_i$ für alle $i \neq q$ und $Q_q = K_q \sqcup L_q$. Dann gilt:

$$\begin{aligned}\lambda_n(Q) &= \lambda_1(Q_1)\lambda_1(Q_2)\dots\lambda_1(Q_q)\dots\lambda_1(Q_n) = \lambda_1(Q_1)\lambda_1(Q_2)\dots\big(\lambda_1(K_q)+\lambda_1(L_q)\big)\dots\lambda_1(Q_n)\\ &= \lambda_1(K_1)\lambda_1(K_2)\dots\lambda_1(K_q)\dots\lambda_1(K_n)+\lambda_1(L_1)\lambda_1(L_2)\dots\lambda_1(L_q)\dots\lambda_1(L_n)\\ &= \lambda_n(K)+\lambda_n(L)\end{aligned}$$

2. Schritt. Seien $Q, Q_1, \dots, Q_k \in \mathcal{Q}_n$ mit $Q = Q_1 \sqcup \dots \sqcup Q_k$. Wir zeigen, dass $\lambda_n(Q) = \lambda_n(Q_1)+\lambda_n(Q_2)+\dots+\lambda_n(Q_k)$. Das Problem beim Beweis verdeutlicht die folgende Vereinigung von 5 Quader:



Keine zwei dieser Quader bilden vereinigt ein Quader. Deshalb kann man nicht einfach den 1. Schritt mehrfach anwenden. Die Lösung bietet eine Zerlegung in kleinere „atomare“ disjunkte Quader:



Daraus kann man erst vertikale „Spalten“ und daraus dann das ganze Quader so aufbauen, dass man jedesmal die Additivität anwenden kann. Das Maß des ganzen Quaders, aber auch das Maß jedes der obigen fünf Teilquader ist jeweils die Summe der Maße der beteiligten „Atome“.

3. Schritt: Wohldefinitheit. Seien $R \in \mathcal{R}_n$ und $R = Q_1 \sqcup \dots \sqcup Q_k = L_1 \sqcup \dots \sqcup L_p$ zwei Quader-Zerlegungen. Dann gilt $Q_j = (Q_j \cap L_1) \sqcup \dots \sqcup (Q_j \cap L_p)$ und nach 2. Schritt ist $\lambda_n(Q_j) = \sum_i \lambda_n(Q_j \cap L_i)$ für alle j. Analog gilt $\lambda_n(L_i) = \sum_j \lambda_n(Q_j \cap L_i)$ für alle i. Daraus folgt:

$$\sum_j \lambda_n(Q_j) = \sum_j \sum_i \lambda_n(Q_j \cap L_i) = \sum_i \sum_j \lambda_n(Q_j \cap L_i) = \sum_i \lambda_n(L_i).$$

3. Schritt: Additivität. Seien $R, S \in \mathcal{R}_n$ zwei disjunkte Figuren, $R = Q_1 \sqcup \dots \sqcup Q_k$, $S = L_1 \sqcup \dots \sqcup L_p$ Quader-Zerlegungen. Dann ist $R \cup S = Q_1 \sqcup \dots \sqcup Q_k \sqcup L_1 \sqcup \dots \sqcup L_p$ und die Definition liefert sofort $\lambda_n(R \cup S) = \sum_{i=1}^k \lambda_n(Q_i) + \sum_{j=1}^l \lambda_n(L_j) = \lambda_n(R)+\lambda_n(S)$ (wegen Wohldefinitheit) also λ_n ist endlich-additiv. Daraus folgt auch die Halbadditivität: $R \cup S = R \sqcup (S \setminus R)$ also $\lambda_n(R \cup S) = \lambda_n(R)+\lambda_n(S\setminus R) \leqslant \lambda_n(R)+\lambda_n(S)$ und die Monotonie: gilt $R \subset S$, so ist $S = R \sqcup (S \setminus R)$ also $\lambda_n(S) = \lambda_n(R)+\lambda_n(S \setminus R) \geqslant \lambda_n(R)$.

5. Schritt: σ-Additivität. Seien $Q, Q_k \in \mathcal{Q}_n$ ($k \in \mathbb{N}$) mit $Q = \sqcup_{k\in\mathbb{N}} Q_k$. Zu zeigen ist $\lambda_n(Q) = \sum_{k=1}^\infty \lambda_n(Q_k)$.

Zunächst zeigen wir $\lambda_n(Q) \geqslant \sum_{k=1}^\infty \lambda_n(Q_k)$. Es ist klar, dass $\lambda_n(Q) \geqslant \sum_{k=1}^N \lambda_n(Q_k)$ für alle N (Beweis?) und für $N \to \infty$ folgt die Behauptung.

Für die Ungleichung $\lambda_n(Q) \leqslant \sum_{k=1}^\infty \lambda_n(Q_k)$ wähle $\varepsilon > 0$ und $K \in \mathcal{Q}_n$ mit $\overline{K} \subset \mathring{Q}$ und $\lambda_n(K) > \lambda_n(Q) - \varepsilon$. Zu jedem $k \in \mathbb{N}$ wähle offene Quader $L_k \in \mathcal{Q}_n$ mit $Q_k \subset L_k$ und $\lambda_n(L_k) > \lambda_n(Q_k) - \frac{\varepsilon}{2^k}$. Dann gilt $K \subset \cup_{k\in\mathbb{N}} L_k$. K ist kompakt, also existiert $N \in \mathbb{N}$ mit $K \subset \cup_{k=1}^N L_k$ und deshalb

$$\lambda_n(Q) < \varepsilon + \lambda_n(K) \leqslant \varepsilon + \sum_{k=1}^N \lambda_n(L_k) \leqslant \varepsilon + \sum_{k=1}^N \big(\lambda_n(Q_k) + \frac{\varepsilon}{2^k}\big) \leqslant 2\varepsilon + \sum_{k=1}^\infty \lambda_n(Q_k)$$

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

Seien $R, S \in \mathcal{R}_n$ zwei disjunkte Figuren, $R = Q_1 \sqcup \dots \sqcup Q_k$, $S = L_1 \sqcup \dots \sqcup L_p$ Quader-Zerlegungen. Dann ist $R \cup S = Q_1 \sqcup \dots \sqcup Q_k \sqcup L_1 \sqcup \dots \sqcup L_p$ und die Definition liefert sofort $\lambda_n(R \cup S) = \sum_{i=1}^k \lambda_n(Q_i) + \sum_{j=1}^l \lambda_n(L_j) = \lambda_n(R)+\lambda_n(S)$ (wegen Wohldefinitheit) also λ_n ist endlich-additiv.

Seien nun $R_i \in \mathcal{R}_n$ paarweise disjunkt (oBdA $R_i \neq \emptyset$), $i \in \mathbb{N}$, mit $R = \cup_{i\in\mathbb{N}} R_i \in \mathcal{R}_n$. Wähle eine Zerlegung $R = B_1 \sqcup \ldots \sqcup B_m$ mit $B_1, \ldots, B_m \in \mathcal{Q}_n$ paarweise disjunkt. Zu jedem i wähle eine Zerlegung $R_i = Q_{i1} \sqcup \ldots \sqcup Q_{ik_i}$ mit $Q_{i1}, \ldots, Q_{ik_i} \in \mathcal{Q}_n$ paarweise disjunkt. Dann gilt

$$B_l = \bigsqcup_{i\in\mathbb{N}} (B_l \cap R_i) = \bigsqcup_{i\in\mathbb{N}, 1\leqslant j\leqslant k_i} (B_l \cap Q_{ij})$$

also

$$\begin{aligned}
\lambda_n(B_l) &= \sum_{i=1}^{\infty}\sum_{j=1}^{n_i} \lambda_n(B_l \cap Q_{ij}) \quad \text{(nach 1. Schritt)}\\
&= \sum_{i=1}^{\infty} \lambda_n(B_l \cap R_i) \quad \text{(nach der Def. des Volumen von Figuren)}
\end{aligned}$$

Es folgt

$$\begin{aligned}
\lambda_n(R) &= \sum_{l=1}^{m} \lambda_n(B_l) \quad \text{(nach der Def. des Volumen von Figuren)}\\
&= \sum_{l=1}^{m}\sum_{i=1}^{\infty} \lambda_n(B_l \cap R_i) \quad \text{(obige Gleichung)}\\
&= \sum_{i=1}^{\infty}\sum_{l=1}^{m} \lambda_n(B_l \cap R_i) \quad \text{(Doppelreihenesatz 3.6.9)}\\
&= \sum_{i=1}^{\infty} \lambda_n(R_i) \quad \text{(wegen endlicher Additivität)}
\end{aligned}$$

6. Schritt: σ-Endlichkeit. Schreibe $\mathbb{R}^n = \cup_k [-k,k]^n$, wobei $\lambda_n([-k,k]^n) = (2k)^n < \infty$. □

13.2. **Konstruktion von Maßräume nach Carathéodory.** Wir beschreiben ein Verfahren, das von Carathéodory stamm, mit dem man Maßräume konstruieren kann. Wir gehen von einem Prämßraum $(X, \mathcal{A}, \mu)$, wobei $\mathcal{A}$ eine Algebra ist und μ ein Prämaß. Wir wollen $(X, \mathcal{A}, \mu)$ zu einem Maßraum $(X, \widetilde{\mathcal{A}}, \widetilde{\mu})$ erweitern, wobei $\widetilde{\mathcal{A}}$ eine σ-Algebra ist und $\widetilde{\mu}$ ein Maß auf $\widetilde{\mathcal{A}}$ ist mit $\widetilde{\mu}|_{\mathcal{A}} = \mu$. Insbesondere wollen wir das Lebesgue-Prämaß zu einem Maß erweitern.

Zur Erläuterung betrachten wir folgende Situation. Ist $A \subset \mathbb{R}^2$ eine beschränkte Menge und $(Q_j)_{j\in\mathbb{N}}$ eine Folge offener Quader im $\mathbb{R}^2$ mit $A \subset \cup_{j\in\mathbb{N}} Q_j$, so stellt $\sum_{j=1}^{\infty} \lambda_2(Q_j)$ einen Näherungswert für den Inhalt von A dar, der den Inhalt umso besser approximiert je kleiner die Differenzmenge $\cup_{j\in\mathbb{N}} Q_j \setminus A$ ist. Deshalb kann man

$$\lambda_2^*(A) := \inf\Big\{ \sum_{j=1}^{\infty} \lambda_2(Q_j) : Q_j \in \mathcal{Q}_2, A \subset \cup_{j\in\mathbb{N}} Q_j \Big\} \tag{13.4}$$

als Approximation des Inhalts von A von *außen* verstehen. Ist D eine beschränkte Teilmenge des $\mathbb{R}^2$ mit $A \subset D$, etwa $D = [-R,R]^2$, mit R genügend groß, so kann entsprechend $\lambda_2^*(D \setminus A)$ als Approximation des Inhalts von $D \setminus A$ von außen betrachtet werden. Damit scheint es sinnvoll, die Zahl $\lambda_{2,D}^*(A) := \lambda_2^*(D) - \lambda_2^*(D \setminus A)$ als Approximation des Inhalts von A von *innen* zu betrachten. Eine besondere Rolle spielen die Teilmengen $A \subset \mathbb{R}^2$, für die bezüglich jeder beschränkten Obermenge D von A die Approximationen von außen und innen gleich sind, d.h. gilt: $\lambda_2^*(A) = \lambda_{2,D}^*(A) = \lambda_2^*(D) - \lambda_2^*(D \setminus A)$.

Wir fassen die Eigenschaften von λ_2^* zusammen wie folgt:

13.2.1. **Definition.** Eine Abbildung $\mu^* : \mathcal{P}(X) \to [0,\infty]$ heißt ***äußeres Maß*** auf X, wenn

(i) $\mu(\emptyset) = 0$,

(ii) μ^* monoton wachsend ist, d.h. wenn aus $A \subset B$ folgt $\mu^*(A) \leqslant \mu^*(B)$,und

(iii) μ^* σ-halbadditiv ist, d.h. wenn für jede Folge $(A_i)_{i\in\mathbb{N}}$ in $\mathcal{P}(X)$ gilt:

$$\mu^*\Big(\bigcup_{i=1}^{\infty} A_i\Big) \leqslant \sum_{i=1}^{\infty} \mu^*(A_i)$$

Bemerkung: Jedes Maß auf $\mathcal{P}(X)$ ist ein äußeres Maß, aber es existieren äußere Maße, die keine Maße sind.

13.2.2. **Definition** (Carathéodory). Sei μ^* ein äußeres Maß auf $\mathcal{P}(X)$. Eine Teilmenge $A \subset X$ heißt ***μ^*–meßbar*** , falls für alle Teilmengen $E \subset X$ gilt

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap (X \setminus A)) \tag{13.5}$$

d.h. A zerlegt jedes $E \subset X$ in 2 disjunkte Teilmengen, auf denen μ^* additiv ist. Da μ^* σ–halbadditiv ist, genügt es in (13.5) $\geqslant$ für alle $E \subset X$ mit $\mu(E) < \infty$ zu fordern. Die Menge aller μ^*–meßbaren Teilmengen von X bezeichnen wir mit $\mathcal{A}_{\mu^*}$.

Bemerkungen: Es gilt

- $\emptyset, X \in \mathcal{A}_{\mu^*}$ (nach Definition).
- Ist $A \subset X$ mit $\mu^*(A) = 0$ so ist A μ^*–meßbar.
- Sind $B \subset A \subset X$ und $\mu^*(A) = 0$ so ist B μ^*–meßbar.

13.2.3. **Satz** (Carathéodory)**.** *Sei* $\mu^* : \mathcal{P}(X) \longrightarrow [0,\infty]$ *ein äußeres Maß. Dann ist* $(X, \mathcal{A}_{\mu^*}, \mu^*|_{\mathcal{A}_{\mu^*}})$ *ein vollständiger Maßraum.*

13.2.4. **Satz.** *Sei* $\mu : \mathcal{R} \longrightarrow [0,\infty]$ *ein Inhalt auf einer Algebra* $\mathcal{R}$ *über* X*. Wir definieren* $\mu^* : \mathcal{P}(X) \longrightarrow [0,\infty]$ *durch*

$$\mu^*(E) := \inf\Big\{ \sum_{n=1}^{\infty} \mu(A_n) : \textit{wobei } E \subset \cup_{n=1}^{\infty} A_n, \textit{ und } A_n \in \mathcal{R} \textit{ für alle } n \in \mathbb{N} \Big\}.$$

Dann gilt:

(1) μ^* *ist ein äußeres Maß auf* X ***(das von*** μ ***erzeugte äußere Maß****)*.
(2) $\mathcal{R} \subset \mathcal{A}_{\mu^*}$.
(3) *Ist* μ *ein Prämass, so gilt* $\mu = \mu^*|_{\mathcal{R}}$.

Man kann also jeden Prämaßraum $(X, \mathcal{R}, \mu)$ zu einem vollständigen Maßraum $(X, \mathcal{A}_{\mu^*}, \mu^*)$ fortsetzen, wobei μ^* das von μ erzeugte Maß ist.

Sei $(X, \mathcal{R}, \mu)$ ein Prämaßraum. $\sigma(\mathcal{R})$ ist die kleinste σ–Algebra, die $\mathcal{R}$ enthält, also gilt $\sigma(\mathcal{R}) \subset \mathcal{A}_{\mu^*}$. $(X, \sigma(\mathcal{R}), \mu^*|_{\sigma(\mathcal{R})})$ ist die „kleinste" Fortsetzung des Prämaßraumes zu einem Maßraum und eventuell kleiner als der vollständige Maßraum $(X, \mathcal{A}_{\mu^*}, \mu^*|_{\mathcal{A}_{\mu^*}})$.

Wie verhalten sich diese Maßräume zueinander?

Dazu führen wir den Begriff der Vervollständigung eines Maßräumes ein und zeigen, dass im Falle σ–endlichen Inhalts die σ–Algebra $\mathcal{A}_{\mu^*}$ die Vervollständigung von $\sigma(\mathcal{R})$ ist.

Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Wir bezeichnen im folgenden mit

$$\mathcal{A}_0 := \{N \in \mathcal{A} \,|\, \mu(N) = 0\}$$

die Menge der Nullmengen von $(X, \mathcal{A}, \mu)$. Desweiteren definieren wir

$$\begin{aligned} \overline{\mu} : \mathcal{A}^{\mu} &\longrightarrow [0,\infty] \\ E = A \cup N &\mapsto \overline{\mu}(E) := \mu(A) \end{aligned}$$

($\overline{\mu}$ ist korrekt definiert.)

Bemerkung: Nach Definition ist $\mathcal{A} \subset \mathcal{A}^{\mu}$. Ist μ vollständig, so gilt $\mathcal{A} = \mathcal{A}^{\mu}$. Ist nämlich $E = A \cup N \in \mathcal{A}^{\mu}$, $A \in \mathcal{A}$, $N \subset N_0 \in \mathcal{A}_0$, so ist auf Grund der Vollständigkeit $N \in \mathcal{A}$ und somit $E \in \mathcal{A}$.

13.2.5. **Satz.** *Sei* $(X, \mathcal{A}, \mu)$ *ein Maßraum. Dann ist* $(X, \mathcal{A}^{\mu}, \overline{\mu})$ *ein vollständiger Maßraum* (***Vervollständigung von*** $(X, \mathcal{A}, \mu)$).

13.2.6. **Satz.** *Sei* $\mathcal{R}$ *eine Algebra auf* X *mit* σ*–endlichem Prämaß,* $\mu : \mathcal{R} \longrightarrow [0,\infty]$*, sei* μ^* *das von* μ *definierte äußere Maß und* $\widetilde{\mu} := \mu^*|_{\sigma(\mathcal{R})}$*. Dann gilt*

$$\sigma(\mathcal{R})^{\widetilde{\mu}} = \mathcal{A}_{\mu^*},$$

und $(X, \mathcal{A}_{\mu^*}, \mu^*|_{\mathcal{A}_{\mu^*}})$ *ist die Vervollständigung von* $(X, \sigma(\mathcal{R}), \widetilde{\mu})$.

Wir behandeln nun die Frage der Eindeutigkeit der Fortsetzung $\widetilde{\mu}$. Dazu zuerst zwei Gegenbeispiele:

(a) Sei S eine überabzählbare Menge, und $\mathcal{R} \subset \mathcal{P}(S)$ sei die Algebra erzeugt duch den endlichen Teilmengen von S sowie $\mu : \mathcal{R} \to [0,\infty]$ das triviale Prämaß $\mu = 0$. Da die von $\mathcal{A}$ erzeugte σ-Algebra die σ-Algebra $\{A \subset S :$ A oder $\complement A$ ist höchstens abzählbar$\}$ ist, ist für jedes $r \in [0,\infty]$

$$\mu_r(A) = \begin{cases} 0 & \text{falls } A \text{ abzählbar} \\ r & \text{falls } A \text{ überabzählbar} \end{cases}$$

eine Fortsetzung von μ zu einem Maß auf $\sigma(\mathcal{R})$.

(b) Betrachte $\mathbb{Q}$ und die σ-Algebra $\mathcal{P}(\mathbb{Q})$. Die Spur von $\mathcal{R}_1$ auf $\mathbb{Q}$ ist ein Erzeuger von $\mathcal{P}(\mathbb{Q})$, auf dem das zählende Maß μ und 2μ übereinstimmen.

Offenbar ist für die Nichteindeutigkeit im ersten Gegenbeispiel verantwortlich, dass die Mengen in $\mathcal{R}$ die Menge S „nicht erreichen“, im zweiten Gegenbeispiel ist es die „starke Unendlichkeit“ von μ auf $\mathcal{R}$. Wir betrachten jetzt Mengenfunktionen, die diese Defekte nicht aufweisen.

13.2.7. **Satz** (Eindeutigkeitssatz)**.** *Sei μ ein σ-endliches Prämaß auf einer Algebra $\mathcal{A}$. Dann existiert genau ein Maß $\widetilde{\mu}$ auf $\sigma(\mathcal{A})$, welches μ fortsetzt, d.h. für das $\widetilde{\mu}|_{\mathcal{A}} = \mu$ gilt. Dieses Maß $\widetilde{\mu}$ ist σ–endlich.*

Constantin Carathéodory wurde am 13.9.1873 in Berlin als Sohn einem griechischen Diplomaten im Dienste der Hohen Pforte, geboren. Carathéodory wuchs in Brüssel auf, wo sein Vater ab 1875 Botschafter war. 1891 legte er das belgische Abitur ab und trat als élève étranger in die École Militaire de Belgique in Brüssel ein. Als Ingenieur war Constantin Carathéodory von 1898 bis 1900 im englischen Dienst bei Staudammarbeiten am Nil beteiligt. Dann beschloss er, zur Überraschung seiner Familie, Mathematiker zu werden. Er studierte von 1900 bis 1902 bei Schwarz, Frobenius und Planck in Berlin und von 1902 bis 1904 bei Klein, Hilbert und Minkowski in Göttingen. 1913 wurde er Nachfolger von Klein in Göttingen und wechselte 1918 an die Universität in Berlin. Zusammen mit Albert Einstein wurde er 1919 in die Preußische Akademie der Wissenschaften aufgenommen. Bei der Aufnahme Carathéodorys hatte Max Planck die Laudatio gesprochen. 1920 ging er an die neu zu organiserende Ionische Universität im Smyrna (Izmir) aber seine Arbeit endete 1922 mit dem Einmarsch der Türken und Vertreibung der griechischen Bevölkerung. Carathéodory konnte noch rechtzeitig seine Familie auf der Insel Samos in Sicherheit bringen, um alleine nach Smyrna zurückzukehren. Dort organisiert er die Rettung kostbaren Schriftguts der Universität, das er auf Booten nach Griechenland transportieren lässt. Danach findet Carathéodory mit seiner Familie Zuflucht in Athen. 1924 folgte er einem Ruf nach München, wo er bis zu seinem Tod am 2.2.1950 blieb.

13.3. **Das Lebesgue-Maß in $\mathbb{R}^n$.** Betrachte den Prämaßraum $(\mathbb{R}^n, \mathcal{R}_n, \lambda_n)$, wobei $\mathcal{R}_n$ die Algebra der Figuren ist, λ_n das elementar geometrische Volumen der Figuren. Dies ist ein σ–endlicher Prämaßraum.

Sei $\lambda_n^* : \mathcal{P}(\mathbb{R}^n) \longrightarrow [0,\infty]$ das von λ_n erzeugte äußere Maß, d.h.

$$\lambda_n^*(A) = \inf\Big\{\sum_{k=1}^{\infty} \lambda(A_k) : A \subset \bigcup_{k=1}^{\infty} A_k, A_k \in \mathcal{R}_n\Big\}$$

13.3.1. **Definition.**

(i) λ_n^* heißt ***äußeres Lebesgue–Maß*** in $\mathbb{R}^n$.
(ii) $\mathcal{L}(\mathbb{R}^n) := \mathcal{A}_{\lambda_n^*} \subset \mathcal{P}(\mathbb{R}^n)$ heißt σ–Algebra der ***Lebesgue–meßbaren Mengen***.
(iii) $\lambda_n = \lambda_n^*|_{\mathcal{L}(\mathbb{R}^n)}$ heißt ***Lebesgue–Maß*** in $\mathbb{R}^n$.
(iv) $\widetilde{\lambda}_n = \lambda_n^*|_{\mathcal{B}(\mathbb{R}^n)}$ heißt ***Borel-Lebesgue–Maß*** in $\mathbb{R}^n$.

Wir wissen schon dass:
(1) $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{R}_n) \subset \mathcal{L}(\mathbb{R}^n)$.
(2) $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), \lambda_n)$ ist vollständig und ist eine Erweiterung von $(\mathbb{R}^n, \mathcal{R}_n, \lambda_n)$.
(3) $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), \lambda_n)$ ist die Vervollständigung von $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda_n)$. Insbesondere:

Eine Menge $A \subset \mathbb{R}^n$ ist genau dann Lebesgue-meßbar, wenn sie sich in der Form $A = B \cup N$ darstellen läßt, wobei $B \subset \mathbb{R}^n$ eine Borelmenge ist und $N \subset \mathbb{R}^n$ eine Lebesgue-Nullmenge ist.

Man kann zeigen, dass:
(4) $\mathcal{B}(\mathbb{R}^n) \neq \mathcal{L}(\mathbb{R}^n) \neq \mathcal{P}(\mathbb{R}^n)$.

13.3.2. **Satz** (Charakterisierung von Lebesgue–meßbaren Mengen)**.** *Sei $A \subset \mathbb{R}^n$, $A \neq \emptyset$. Betrachte die folgenden Bedingungen:*

(a) *Zu jedem $\varepsilon > 0$ existiert eine offene Menge $U \supset A$ mit $\lambda_n^*(U \setminus A) \leqslant \varepsilon$.*
(b) *Zu jedem $\varepsilon > 0$ existiert eine abgeschlossene Menge $F \subset A$ mit $\lambda_n^*(A \setminus F) \leqslant \varepsilon$.*

Ist $A \in \mathcal{L}(\mathbb{R}^n)$, so sind (a) und (b) erfüllt. Ist mindestens eine von (a) und (b) erfüllt, so ist $A \in \mathcal{L}(\mathcal{R}^n)$.

13.3.3. **Folgerung.**

(i) Eine Menge $A \subset \mathbb{R}^n$ ist genau dann Lebesgue–meßbar, wenn sie sich in der Form $A = (\cup_{k=1}^{\infty} F_k) \cup N$ darstellen läßt, wobei $F_k \subset \mathbb{R}^n$ abgeschlossene Mengen sind und $N \subset \mathbb{R}^n$ eine Lebesgue–Nullmenge ist.

(ii) Eine Menge $A \subset \mathbb{R}^n$ ist genau dann Lebesgue–meßbar, wenn es eine Darstellung der Form $A \sqcup N = \cap_{k=1}^{\infty} U_k$ gibt, wobei N eine Lebesguesche Nullmenge ist und $U_k \subset \mathbb{R}^n$ offene Menge sind.

(iii) Eine Teilmenge $N \subset \mathbb{R}^n$ ist genau dann eine Lebesguesche Nullmenge, wenn es zu jedem $\varepsilon > 0$ eine Folge von Würfeln $W_1, W_2, W_3, \dots$ im $\mathbb{R}^n$ gibt so, dass

$$N \subset \bigcup_{i=1}^{\infty} \mathring{W}_i \quad und \quad \sum_{i=1}^{\infty} \lambda_n(W_i) < \varepsilon\,.$$

Beispiele:

(1) Jede abzählbare Teilmenge $A \subset \mathbb{R}^n$ ist Lebesgue–Nullmenge.

(2) $F \in \mathcal{C}^1(D,G)$, $N \subset D$ Nullmenge $\Rightarrow F(D)$ Nullmenge.

(3) Jede Hyperebene ist eine Nullmenge.

(4) Die ***Cantorsche Menge*** ensteht aus dem Intervall $[0,1]$, indem man das mittlere Drittel $I_{11} := (\frac{1}{3}, \frac{2}{3})$ entfernt, aus jedem verbleibenden Intervall $[0, \frac{1}{3}]$, $[\frac{2}{3}, 1]$ das mittlere Drittel I_{21}, I_{22} entfernt, usw. Setze $C := [0,1] \setminus U$, wobei $U = I_{11} \cup (I_{21} \cup I_{22}) \cup (I_{31} \cup I_{32} \cup \ldots \cup I_{34}) \cup \ldots$. Die Cantorsche Menge C ist kompakt, also Borelmenge.

$$\lambda_1(U) = \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3^2} + \frac{1}{4} \cdot \frac{1}{3^3} + \ldots = \sum_{k=1}^{\infty} \frac{1}{2}\Big(\frac{2}{3}\Big)^k = \frac{1}{2} \cdot \frac{2}{3} \sum_{k=0}^{\infty} \Big(\frac{2}{3}\Big)^k = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{1}{1 - \frac{2}{3}} = 1\,.$$

Daraus folgt $\lambda_1(C) = 0$. Da λ_1 vollständig ist, so ist jede Teilmenge von C eine Lebesgue–Nullmenge. Außerdem kann man zeigen, dass $|C| = |\mathbb{R}|$. Dann $|\mathcal{L}(\mathbb{R})| \geqslant |\mathcal{P}(C)| > |\mathbb{R}|$. Mann kann aber beweisen, dass $|\mathcal{B}(\mathbb{R})| = |\mathbb{R}|$. Es gibt also mehr Lebesgue–meßbaren Mengen als Borelmengen, z.B. zwischen die Teilmengen der Cantorschen Menge C. Um eine nicht–Lebesgue-meßbaren Menge zu konstruieren, braucht man das Auswahlaxiom (siehe [7, III, §3]).

(5) Sei $E \in \mathcal{L}(\mathbb{R}^n)$. Setze $\mathcal{L}(E) = \{A \subset E : A \in \mathcal{L}(\mathbb{R}^n)\}$. $(E, \mathcal{L}(E), \lambda_n|_{\mathcal{L}(E)})$ ist dann ein vollständiger Maßraum. $\lambda_n|_{\mathcal{L}(E)}$ heißt das $\mathcal{L}$–Maß auf E.

(6) Wir wollen das Beispielreservoir für Maße auf $\mathcal{B}(\mathbb{R})$ noch ein wenig ausdehnen. Das Lebesgue-Maß ist dasjenige Maß, das von der üblichen Längenmessung abgeleitet ist. Was passiert bei einer gewichteten Längenmessung? Sei $F : \mathbb{R} \longrightarrow \mathbb{R}$ monoton wachsend, linkseitig–stetig. Setze $\mu_F([a,b)) = f(b) - f(a)$ für $a < b$ in $\mathbb{R}$. μ_F kann auf die Algebra der endlichen Vereinigungen von Intervalle fortgesetzt werden und wird ein Prämaß ([7, Satz 2.2,II.2]). Aus 13.2.3 folgt, dass es ein vollständiger Maßraum existiert $(\mathbb{R}, \mathcal{A}_{\mu_F^*}, \mu_F)$ so, dass $\mathcal{B}(\mathbb{R}) \subset \mathcal{A}_{\mu_F^*}$; μ_F heißt ***Lebesgue-Stieltjes–Maß***. Für den Beweis der σ-Additivität von μ_F benötigt man übrigens die linksseitige Stetigkeit von F. Umgekehrt kann man zeigen, dass ein Maß auf $\mathcal{B}(\mathbb{R})$, das auf kompakten Mengen endlich ist, ein Lebesgue-Stieltjes-Maß ist.

Wir stellen uns die Fragen, ob ein Inhalt oder Maß $\mu : \mathcal{P}(\mathbb{R}^n) \to [0,\infty]$, $\quad E \mapsto \mu(E)$ existiert mit folgenden Eigenschaften:

(1) Ist $E \subset \mathbb{R}^n$ *kongruent* zu $F \subset \mathbb{R}^n$ (E kann durch Translation, Rotation und Spiegelung in F überführt werden), dann gilt $\mu(E) = \mu(F)$.

(2) Der Einheitswürfel definiert durch $Q = \{x \in \mathbb{R}^n \mid 0 < x_i < 1,\ i = 1, \dots, n\}$ hat das Maß 1, also $\mu(Q) = 1$.

Hausdorff hat 1914 gezeigt:

In keiner Dimension $n \geqslant 1$ gibt es ein Inhalt μ mit den Eigenschaften (1) und (2). Fur $n = 1,2$ gibt es eine Lösung μ, aber μ ist in diesem Fall nicht eindeutig (Banach 1923). Die Nichtexistenz von μ für $n \geqslant 3$ folgt z.B. aus dem Banach-Tarski-Paradoxon (siehe [23]).

13.3.4. **Satz** (Banach-Tarski 1924). *Seien U, V beliebige nicht leere, offene Teilmengen des $\mathbb{R}^n$ für $n \geqslant 3$. Dann existiert ein $N \in \mathbb{N}$ und $E_1, \dots, E_N$ bzw. $F_1, \dots, F_N$, so dass*

(1) $U = \bigcup_{k=1}^{N} E_k, \quad E_i \cap E_k = \emptyset,\ i \neq k$

(2) $V = \bigcup_{k=1}^{N} F_k, \quad F_i \cap F_k = \emptyset,\ i \neq k$

(3) *E_k ist kongruent zu F_k für $k = 1, \dots, N$.*

Der Beweis dieses Satzes beruht auf dem Auswahlaxiom der Mengenlehre: Sei $\{E_\alpha \mid \alpha \in A\}$ eine beliebige Familie nichtleerer Mengen E_α. Dann gibt es eine Funktion $f : A \to \bigcup_{\alpha \in A} E_\alpha$ mit $f(\alpha) \in E_\alpha$.

Es gibt auch kein Maß $\mu : \mathcal{P}(\mathbb{R}^n) \to [0,\infty]$ mit den Eigenschaften (1) und (2) (Vitali, 1905). Das Maßproblem ist lösbar, wenn man statt beliebigen Mengen $E \subset \mathbb{R}^n$ nur *Lebesguesche* Mengen zulässt.

13.3.5. **Satz.** *Sei $F : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ eine $\mathcal{C}^1$–Abbildung. Dann gilt: Ist $A \subset U$ eine Lebesguesche Nullmenge, so ist auch $F(A)$ eine Lebesguesche Nullmenge.*

Beweis: Da $U \subset \mathbb{R}^n$ offen ist, ist U als disjunkte Vereinigung von abzählbar vielen halboffenen Quadern darstellbar: $U = \bigsqcup_{j=1}^{\infty} Q_j$ mit $\overline{Q}_j \subset U$. Dann ist

$$F(A) = F\big(A \cap (\cup_j Q_j)\big) = F\big(\cup_j (A \cap Q_j)\big) = \cup_j F(A \cap Q_j).$$

Folglich genügt es zu zeigen, dass $F(A \cap Q_j)$ eine Lebesguesche Nullmenge ist. Da A eine Lebesgue-Nullmenge ist, existieren nach dem Kriterium für Lebesgue-Nullmengen aus Folgerung 13.3.2 zu jedem $\varepsilon > 0$ abzählbar viele Würfel W_m, $m = 1,2,3,\dots$ so, dass

$$A \subset \bigcup_{m=1}^{\infty} \mathring{W}_m \quad \text{und} \quad \sum_{m=1}^{\infty} \operatorname{vol}(W_m) < \varepsilon .$$

Sei $2r_m$ die Seitenlänge von W_m und ξ_m das Zentrum von W_m. Für $x \in A \cap Q_j \cap \mathring{W}_m$ gilt $\|x - \xi_m\|_\infty := \max\{|x_i - \xi_{mi}| : i = 1,\dots,n\} < r_m$. Da F eine $\mathcal{C}^1$–Abbildung ist, ist $F|_{\overline{Q}_j}$ Lipschitzstetig. Somit existiert eine Konstante $M_j \in \mathbb{R}_+$ mit

$$\|F(x) - F(\xi_m)\|_\infty \leqslant M_j \|x - \xi_m\|_\infty < M_j r_m ,$$

d.h. $F(x)$ liegt im Inneren eines Würfels $\widetilde{W}_m$ mit dem Zentrum $F(\xi_m)$ und der Kantenlänge $2M_j r_m$. Dann gilt $F(A \cap Q_j) = \bigcup_{m=1}^{\infty} F(A \cap Q_j \cap \mathring{W}_m) \subset \bigcup_{m=1}^{\infty} \mathring{\widetilde{W}}_m$ und

$$\sum_{m=1}^{\infty} \lambda_n(\widetilde{W}_m) = \sum_{m=1}^{\infty} (2r_m M_j)^n \leqslant \sum_{m=1}^{\infty} \lambda_n(W_m) \cdot (M_j)^n \leqslant \varepsilon \cdot (M_j)^n .$$

Daher ist $F(A \cap Q_j)$ eine Lebesgue-Nullmenge. □

13.3.6. **Satz.** *Sei $G : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ eine abgeschlossene $\mathcal{C}^1$–Abbildung. Dann gilt: Ist $A \subset U$ Lebesgue–meßbar, so ist $G(A)$ ebenfalls Lebesgue–meßbar.*

Beweis: Nach 13.3.3 ist die Menge A genau dann Lebesgue–meßbar, wenn sie eine Darstellung der Form $A = \big(\bigcup_{k=1}^{\infty} F_k\big) \cup N$ besitzt, wobei $F_k \subset \mathbb{R}^n$ abgeschlossene Mengen sind und N eine Lebesguesche Nullmenge ist. Da die Abbildung F abgeschlossen ist, bildet sie die abgeschlossenen Mengen F_k in abgeschlossenen Mengen $G(F_k)$ ab. Nach 13.3.5 ist $G(N)$ eine Nullmenge. Nach 13.3.3 ist

$$G(A) = \Big(\bigcup_{k=1}^{\infty} G(F_k)\Big) \cup G(N)$$

ebenfalls eine Lebesgue–Menge. □

13.4. **Meßbare Funktionen.** Wie in der einleitende Diskussion über Lebesgue-Summen deutlich gemacht ist, sind die meßbaren Funktionen die einzigen die als Integranden vorkommen (siehe auch die Einleitung zum Kapitel III in [7]).

13.4.1. **Definition.** Seien $(X,\mathcal{A})$, $(Y,\mathcal{B})$ meßbaren Räume. Eine Abbildung $f : X \longrightarrow Y$ heißt ***meßbar*** (genauer $(\mathcal{A},\mathcal{B})$–***meßbar***), falls $f^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}$.

Man beachte die Ähnlichkeit zwischen dieser Definition und der Definition von stetigen Abbildungen zwischen topologischen Räumen.

13.4.2. **Beispiel.** (a) Konstante Abbildungen $(X,\mathcal{A}) \to (Y,\mathcal{B})$ sind stets messbar.
(b) Die charakteristische Funktion $\mathbb{1}_A : (X,\mathcal{A}) \to (\mathbb{R},\mathcal{B}(\mathbb{R}))$ einer Menge $A \subset X$ ist genau dann meßbar, wenn $A \in \mathcal{A}$ ist.

13.4.3. **Lemma.** *Ist $\mathcal{B} = \sigma(\mathcal{E})$ so gilt:*

$$f \textit{ ist } (\mathcal{A},\mathcal{B})\textit{–meßbar} \iff \forall B \in \mathcal{E} : f^{-1}(B) \in \mathcal{A} .$$

Beweis: Betrachte die σ–Algebra $\mathcal{C} = \{C \subset Y : f^{-1}(C) \in \mathcal{A}\}$; dann gilt $\mathcal{C} \supset \mathcal{E}$. Daraus folgt $\mathcal{C} \supset \sigma(\mathcal{E}) = \mathcal{B}$. □

Betrachte nun auf $\overline{\mathbb{R}}$ die folgende Topologie $\mathcal{O}(\overline{\mathbb{R}}) : A \in \mathcal{O}(\overline{\mathbb{R}}) :\Leftrightarrow A \cap \mathbb{R}$ offen in $\mathbb{R}$ und falls $+\infty \in A$ (bzw. $-\infty \in A$), existiert $a \in \mathbb{R}$ mit $(a,+\infty] \subset A$ (bzw. $[-\infty,a) \subset A$).

Sei $\mathcal{B}(\overline{\mathbb{R}})$ die σ–Algebra der Borelmengen in $\overline{\mathbb{R}}$:

$$\mathcal{B}(\overline{\mathbb{R}}) = \sigma(\mathcal{O}(\overline{\mathbb{R}})) = \{B \cup E : B \in \mathcal{B}(\mathbb{R}), E \subset \{-\infty,\infty\}\} .$$

Funktionen $f : X \to \overline{\mathbb{R}}$ heißen *numerische Funktionen*. Der Grund für die Einführung $\overline{\mathbb{R}}$-wertiger Funktionen liegt hauptsächlich darin, dass für jede Folge $f_k : X \to \overline{\mathbb{R}}$ die Funktion $\sup f_k$ stets definiert ist.

Außer den bereits getroffenen Vereinbarungen (13.2) benötigen wir noch:

$$a - \infty = -\infty \quad \text{für alle } a \in \mathbb{R} \text{ oder } a = -\infty,$$

(13.6)
$$a \cdot \infty = \begin{cases} -\infty, a < 0 \\ 0, a = 0 \end{cases}.$$

Der Ausdruck $\infty - \infty$ bleibt verboten; daher ist für $\overline{\mathbb{R}}$-wertige Funktionen $f - g$ nicht unbedingt überall definiert.

13.4.4. **Definition.** Sei $f : (X, \mathcal{A}) \longrightarrow (Y, \mathcal{B})$, wobei $Y = [0,\infty], \overline{\mathbb{R}}, \mathbb{C}$ und $\mathcal{B}$ die σ-Algebra der Borelmengen von Y ist. Dann heißt f ***$\mathcal{A}$-meßbar*** (oder einfach meßbar, wenn $\mathcal{A}$ klar vom Kontext ist). Falls X ein topologischer Raum ist und $\mathcal{A} = \mathcal{B}(X)$ die σ-Algebra der Borelmengen von X ist, so heißt f ***Borel-meßbar***. Falls $X \in \mathcal{L}(\mathbb{R}^n)$ und $\mathcal{A} = \mathcal{L}(X)$ (σ–Algebra der Lebesgue–meßbaren Teilmengen von X) sagen wir, dass f ***Lebesgue–meßbar*** oder $\mathcal{L}$–meßbar ist.

13.4.5. **Beispiel.**
(1) Seien X, Y topologische Räume und $f : X \longrightarrow Y$ stetig, dann ist f $(\mathcal{B}(X), \mathcal{B}(Y))$–meßbar. (Setze $\mathcal{E} = \mathcal{O}(Y)$ in Lemma 13.4.3; für $U \in \mathcal{O}(Y))$ gilt $f^{-1}(U) \in \mathcal{O}(X) \subset \mathcal{B}(X)$). Eine Abbildung $f : X \longrightarrow Y$ zwischen topologischen Räume heißt ***Borel-meßbar***, falls sie $(\mathcal{B}(X), \mathcal{B}(Y))$–meßbar ist.

Für stetiges $f : X \to Y$ ist also das Urbild einer Borelmenge wieder eine Borelmenge. Hingegen braucht das stetige Bild einer Borelmenge keine Borelmenge zu sein! Lebesgue hatte irrtümlich angenommen, dass das so ist; aber 1917 wurde ein Gegenbeispiel von Suslin konstruiert (siehe [2, S. 245]).

(2) Sei $X \in \mathcal{L}(\mathbb{R}^n)$, $f : X \longrightarrow \mathbb{R}$ fast überall stetig (d.h. existiert $N \subset X$ eine Lebesgue–Nullmenge so, dass $f : X \setminus N \longrightarrow \mathbb{R}$ stetig). Dann ist f meßbar.

(3) $f : (X, \mathcal{A}) \longrightarrow \mathbb{R}$ oder $[0,\infty]$ ist meßbar genau dann, wenn für jedes $b \in \mathbb{R}$ gilt

$$\{f < b\} := \{x \in X : f(x) < b\} = f^{-1}((-\infty, b)) \in \mathcal{A}.$$

(4) $f = (f_1, \ldots, f_n) : X \longrightarrow \mathbb{R}^n$ ist $(\mathcal{A}, \mathcal{B}(\mathbb{R}^n))$-meßbar genau dann, wenn $f_j : X \longrightarrow \mathbb{R}$ meßbar sind für alle $j = 1, \ldots, n$. Insbesondere ist $f : X \longrightarrow \mathbb{C}$ meßbar genau dann, wenn $\mathrm{Re}(f), \mathrm{Im}(f) : X \longrightarrow \mathbb{R}$ meßbar sind.

(5) Seien $f : (X, \mathcal{A}) \longrightarrow (Y, \mathcal{B})$, $g : (Y, \mathcal{B}) \longrightarrow (Z, \mathcal{C})$ meßbar, dann ist $g \circ f : (X, \mathcal{A}) \longrightarrow (Z, \mathcal{C})$ meßbar (weil $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$).

(6) Seien $f, g : X \longrightarrow [0,\infty]$ meßbar. Dann sind $f + g$, $f \cdot g$, $|f|$, $\min(f, g)$, $\max(f, g)$ meßbar. Gilt $g \neq 0$, so ist f/g meßbar. Ist der Wertebereich $\overline{\mathbb{R}}$, so gilt die Aussage noch, wenn die Operationen definiert sind (z.B. ist $f + g$ nicht immer definiert, wegen der verbotenen Operation $\infty - \infty$). Zum Beweis: Wir stellen $f + g$ z.B. als Zusammensetzung von $x \to (f(x), g(x))$ und $(y_1, y_2) \to y_1 + y_2$ dar.

(7) Seien $f_k : (X, \mathcal{A}) \longrightarrow \overline{\mathbb{R}}$ $\mathcal{A}$–meßbar. Dann sind $\sup f_k$, $\inf f_k$, $\limsup_{k\to\infty} f_k$ und $\liminf_{k\to\infty} f_k$ meßbar. Seien $f_k : X \longrightarrow \overline{\mathbb{R}}$ meßbar so, dass $\lim_{k\to\infty} f_k(x) = f(x)$ für alle $x \in X$. Dann ist f meßbar. In der Tat, sei $a \in \mathbb{R}$. Dann gilt $\{(\sup_{n\in\mathbb{N}} f_n) \le a\} = \bigcap_{n\in\mathbb{N}} \{f_n \le a\} \in \mathcal{B}(X)$ und $\{x \in X : \inf_{n\in\mathbb{N}} f_n(x) \ge a\} = \bigcap_{n\in\mathbb{N}} \{x : f_n(x) \ge a\} \in \mathcal{B}(X)$.
Mit **(3)** folgt dann, dass $\sup_{n\in\mathbb{N}} f_n$ und $\inf_{n\in\mathbb{N}} f_n$ messbar sind. Damit sind auch $\overline{\lim}_{n\to\infty} f_n = \inf_{j\in\mathbb{N}} \sup_{n\ge j} f_n$ und $\underline{\lim}_{n\to\infty} f_n = \sup_{j\in\mathbb{N}} \inf_{n\ge j} f_n$ messbar.
Wenn existent für alle $x \in X$, dann ist somit auch $\lim_{n\to\infty} f_n = \overline{\lim}_{n\to\infty} f_n$ messbar.

(8) Eine Funktion $f : X \to \overline{\mathbb{R}}$ heißt *einfach*, falls f messbar ist und nur endlich viele Werte annimmt (siehe Lemma 13.5.1). Es gilt:

(i) $f : X \to [0,\infty]$ ist meßbar genau dann, wenn es eine Folge $(s_n)_{n\in\mathbb{N}}$ einfacher Funktionen gibt, mit $0 \leqslant s_n \leqslant s_{n+1} \leqslant f$ für alle n und $\lim_{n\to\infty} s_n(x) = f(x)$ (kurz: $s_n \nearrow f$).

(ii) $f : X \to \overline{\mathbb{R}}$ ist meßbar genau dann, wenn es eine Folge $(s_n)_{n\in\mathbb{N}}$ einfacher Funktionen gibt, mit $\lim_{n\to\infty} s_n(x) = f(x)$.

(9) Sei $(X, \mathcal{A}, \mu)$ vollständig, $f_k : X \longrightarrow \overline{\mathbb{R}}$ meßbar für $k \in \mathbb{N}$. Sei $f : X \longrightarrow \overline{\mathbb{R}}$, so dass $\lim_{k\to\infty} f_k = f$ μ–fast überall (d.h. es gibt eine Nullmenge $N \subset X$ so, dass für jedes $x \in X \setminus N$ gilt $\lim_{k\to\infty} f_k(x) = f(x)$). Dann ist f meßbar.

(10) Sei $(X, \mathcal{A}, \mu)$ beliebig, $f_k : X \longrightarrow \overline{\mathbb{R}}$ meßbar für $k \in \mathbb{N}$. Sei $f : X \longrightarrow \overline{\mathbb{R}}$, so dass $\lim_{k\to\infty} f_k = f$ μ–fast überall. Sei $N \subset X$ eine Nullmenge so, dass für jedes $x \in X \setminus N$ gilt $\lim_{k\to\infty} f_k(x) = f(x)$. Definiere

$$\widetilde{f} : X \longrightarrow \overline{\mathbb{R}}, \quad \widetilde{f}(x) = \begin{cases} f(x), & x \notin N \\ 0, & x \in N \end{cases}$$

Dann ist $\widetilde{f}$ meßbar.

13.5. **Integration meßbarer Funktionen.** Wir erinnern, dass die charakteristische Funktion einer Teilmenge $A \subset X$ durch

$$\mathbb{1}_A : X \to \{0,1\}\,, \quad \mathbb{1}_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

definiert ist.

13.5.1. **Lemma.** *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $\varphi : X \to \mathbb{R}$. Dann sind äquivalent:*
(i) φ nimmt nur endlich viele verschiedene Werte an und für jedes $c \in \mathrm{Im}(\varphi)$ gilt $\varphi^{-1}(c) \in \mathcal{A}$,
(ii) φ besitzt eine Darstellung der Form $\varphi = \sum_{j=1}^m \alpha_j \mathbb{1}_{A_j}$, wobei $A_j \in \mathcal{A}$ und $\sqcup_{j=1}^m A_j = X$. Eine Funktion heißt ***einfach*** *falls eine der äquivalenten Bedingungen (i) oder (ii) erfüllt ist.*

Ist (i) erfüllt, so hat φ die Standardform

$$\varphi = \sum_{c \in \mathrm{Im}(\varphi)} c\, \mathbb{1}_{\varphi^{-1}(c)} = \sum_{j=1}^m \alpha_j \mathbb{1}_{A_j},$$

wobei $\mathrm{Im}(\varphi) = \{\alpha_1, \dots, \alpha_m\}$ und $A_j = \varphi^{-1}(\alpha_j)$. Eine Darstellung wie in (ii) ist <u>nicht</u> eindeutig! z.B.

$$\mathbb{1}_{[0,1]} = \mathbb{1}_{[0,\frac12)} + \mathbb{1}_{[\frac12,1]} = \mathbb{1}_{[0,\frac12)} + \mathbb{1}_{\{\frac12\}} + \mathbb{1}_{(\frac12,1]}$$

Eine einfache Funktion ist offensichtlich $\mathcal{A}$-meßbar.
Beispiele: Sei $[a,b] \subset \mathbb{R}$ ein Intervall. Dann ist eine Treppenfunktion (siehe Def. 6.1.1) eine einfache Funktion auf $([a,b], \mathcal{B}([a,b]), \lambda_1)$ oder $([a,b], \mathcal{L}([a,b]), \lambda_1)$. Aber auch nicht-Treppenfunktionen wie $\mathbb{1}_{\mathbb{Q}}$, $\mathbb{1}_{\mathbb{R}\setminus\mathbb{Q}}$ oder $\mathbb{1}_N$ (wobei N eine Lebesgue-Nullmenge ist) sind einfache Funktionen auf $([a,b], \mathcal{L}([a,b]), \lambda_1)$.

13.5.2. **Definition.** Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Wir bezeichnen die Menge der nicht-negativen einfachen Funktionen auf X mit $\mathcal{E}_+(X)$. Ist $\varphi \in \mathcal{E}_+(X)$, so heißt

$$\int_X \varphi\, d\mu := \sum_{c \in \mathrm{Im}(\varphi)} c\, \mu(\varphi^{-1}(c)), \tag{13.7}$$

das ***Integral*** von φ (bzgl. μ).

Beispiele: Ist $\varphi \in \mathcal{T}([a,b])$ eine nichtnegative Treppenfunktion, so gilt $\int_{[a,b]} \varphi\, d\lambda_1 = \int_a^b \varphi\, dx$, wobei die rechte Seite das Regelintegral von φ ist (Definition 6.1.4). Außerdem $\int_{[a,b]} \mathbb{1}_{\mathbb{Q}}\, d\lambda_1 = \lambda_1([a,b] \cap \mathbb{Q}) = 0$, $\int_{[a,b]} \mathbb{1}_{\mathbb{R}\setminus\mathbb{Q}}\, d\lambda_1 = \lambda_1([a,b] \setminus \mathbb{Q}) = b - a$.

13.5.3. **Lemma.** *Besitzt $\varphi \in \mathcal{E}_+(X)$ eine Darstellung $\varphi = \sum_{j=1}^m \alpha_j \mathbb{1}_{A_j}$, so gilt*

$$\int_X \varphi\, d\mu = \sum_{j=1}^m \alpha_j \mu(A_j), \tag{13.8}$$

Beweis: Sei $\varphi = \sum \alpha_j \mathbb{1}_{A_j} = \sum \beta_l \mathbb{1}_{B_l}$, wobei $X = \sqcup_j A_j = \sqcup_l B_l$. Z.z. $\sum_j \alpha_j \mu(A_j) = \sum_l \beta_l \mu(B_l)$. Dann ist $X = \sqcup_{j,l}(A_j \cap B_l)$, $A_j = \sqcup_l (A_j \cap B_l)$ und $B_l = \sqcup_j (A_j \cap B_l)$. Deshalb $\mu(A_j) = \sum_l \mu(A_j \cap B_l)$, $\mu(B_l) = \sum_j \mu(A_j \cap B_l)$. Außerdem gilt $\alpha_j = \beta_l$ falls $A_j \cap B_l \neq \emptyset$. Folglich

$$\begin{aligned}\sum_j \alpha_j \mu(A_j) &= \sum_{j,l} \alpha_j \mu(A_j \cap B_l) = \sum_{\substack{j,l \\ A_j \cap B_l \neq \emptyset}} \alpha_j \mu(A_j \cap B_l) = \sum_{\substack{j,l \\ A_j \cap B_l \neq \emptyset}} \beta_l \mu(A_j \cap B_l) \\ &= \sum_{j,l} \beta_l \mu(A_j \cap B_l) = \sum_l \beta_l \mu(B_l).\end{aligned}$$

$\square$

13.5.4. **Satz.** *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum, $\varphi, \psi \in \mathcal{E}_+(X)$. Dann gilt:*
(i) $\alpha \in [0, \infty) \Longrightarrow \alpha\varphi \in \mathcal{E}_+(X)$ und $\int_X \alpha\varphi\, d\mu = \alpha \int_X \varphi\, d\mu$.
(ii) $\varphi + \psi \in \mathcal{E}_+(X)$ und $\int_X (\varphi + \psi)\, d\mu = \int_X \varphi\, d\mu + \int_X \psi\, d\mu$.
(iii) $\varphi \leqslant \psi \Longrightarrow \int_X \varphi\, d\mu \leqslant \int_X \psi\, d\mu$
(iv) $A \in \mathcal{A} \Longrightarrow \mathbb{1}_A \in \mathcal{E}_+(X)$ und $\int_X \mathbb{1}_A\, d\mu = \mu(A)$.

Bemerkung: Ist X ein topologischer Raum und $\mathcal{A} = \mathcal{B}(X)$ so gilt $\int_X \varphi\, d\mu \leqslant \sup_X \varphi \cdot \mu(\mathrm{supp}\,\varphi)$; in der Tat, $\varphi \leqslant \sup_X \varphi \cdot \mathbb{1}_{\mathrm{supp}\,\varphi}$.

13.5.5. **Definition.** Sei $(X,\mathcal{A},\mu)$ ein Maßraum und $f : X \to [0,\infty]$ eine meßbare Funktion. Wir wissen nach Beispiel 13.4.5 (8), dass

$$f : X \longrightarrow [0,\infty] \text{ist meßbar} \iff \exists (\varphi_n)_{n\in\mathbb{N}} : \varphi_n \in \mathcal{E}_+ , \varphi_n \nearrow f . \tag{13.9}$$

Dann heißt

$$\int_X f\,d\mu := \lim \int_X \varphi_n\,d\mu = \sup \int_X \varphi_n\,d\mu, \tag{13.10}$$

für eine beliebige Folge $(\varphi_n)_{n\in\mathbb{N}}$ mit $\varphi_n \in \mathcal{E}_+, \varphi_n \nearrow f$. das ***Integral*** von φ (bzgl. μ). Ist $A \in \mathcal{A}$, so definieren wir das ***Integral von φ über*** A (bzgl. μ) durch

$$\int_A f\,d\mu := \int_X \mathbb{1}_A f\,d\mu. \tag{13.11}$$

Die Definition des Integrals ist von der Wahl der Folge $(\varphi_n)_{n\in\mathbb{N}}$ unabhängig. Wir können nämlich beweisen, dass für eine wachsende Folge $(\varphi_n)_{n\in\mathbb{N}}$ aus $\mathcal{E}_+$ und $\psi \in \mathcal{E}_+$ mit $\psi \leqslant \lim_{n\to\infty} \varphi_n$, so gilt $\int_X \psi\,d\mu \leqslant \lim_{n\to\infty} \int_X \varphi_n\,d\mu$. Sei nun $\varphi_n \nearrow f$, $\psi_n \nearrow f$ zwei Folgen wie in der Definition. Dann gilt $\psi_m \leqslant f = \lim_{n\to\infty} \varphi_n$ also $\int_X \psi_m\,d\mu \leqslant \lim_{n\to\infty} \int_X \varphi_n\,d\mu$ und folglich $\lim_{m\to\infty} \int_X \psi\,d\mu \leqslant \lim_{n\to\infty} \int_X \varphi_n\,d\mu$. Analog erhalten wir $\lim_{n\to\infty} \int_X \varphi_n\,d\mu \leqslant \lim_{m\to\infty} \int_X \psi\,d\mu$.

13.5.6. **Satz.** *Sei $(X,\mathcal{A},\mu)$ ein Maßraum, $f,g : X \to [0,\infty]$ seien meßbar.*
(i) $\alpha \in [0,\infty) \Longrightarrow \int_X \alpha f\,d\mu = \alpha \int_X f\,d\mu$.
(ii) $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.
(iii) $f \leqslant g \Longrightarrow \int_X f\,d\mu \leqslant \int_X g\,d\mu$

13.5.7. **Satz.** *Sei $(X,\mathcal{A},\mu)$ ein Maßraum, $f : X \to [0,\infty]$ sei messbar und $A,B,E \in \mathcal{A}$.*
(i) $\mu(E) = 0 \Longrightarrow \int_E f\,d\mu = 0$.
(ii) Ist $A \cap B$ eine Nullmenge, so gilt $\int_{A\cup B} f\,d\mu = \int_A f\,d\mu + \int_B f\,d\mu$.
(iii) $\int_E f\,d\mu = 0 \iff \mu\big(\{x \in X : f(x) \neq 0\}\big) = 0$, *d.h.* $f = 0$ *μ-f.ü.*

Der positive Teil f_+ und der negative Teil f_- einer Funktion $f : X \to \overline{\mathbb{R}}$ sind definiert durch

$$f_+(x) = \max\{f(x), 0\} \geqslant 0, \quad f_-(x) = \max\{-f(x), 0\} \geqslant 0$$

Ist f messbar, dann sind nach Beispiel 13.4.5 (6) auch f_+ und f_- messbar. Offensichtlich gilt

$$f = f_+ - f_-, \quad |f| = f_+ + f_-$$

und für $|f| < \infty$ folgt $f_\pm = \frac{1}{2}(|f| \pm f)$.

13.5.8. **Definition.** Sei $(X,\mathcal{A},\mu)$ ein Maßraum und $f : X \to \overline{\mathbb{R}}$ eine meßbare Funktion.
(i) Gilt $\displaystyle\int_X f_+\,d\mu < \infty$ oder $\displaystyle\int_X f_-\,d\mu < \infty$, dann heißt

$$\int_X f\,d\mu = \int_X f_+\,d\mu - \int_X f_-\,d\mu \in \overline{\mathbb{R}}$$

das ***Lebesgue–Integral*** von f über X (bzgl. μ) und wir sagen, dass $\displaystyle\int_X f\,d\mu$ existiert.
(ii) Sind beide Integrale $\displaystyle\int_X f_+\,d\mu$, $\displaystyle\int_X f_-\,d\mu$ endlich, so sagen wir, dass f ***Lebesgue–integierbar*** ist.
(iii) Die Menge der integrierbaren Funktionen $f : X \to \mathbb{R}$ wird mit $\mathcal{L}^1(X,\mu)$ bezeichnet.
(iv) Ist $A \in \mathcal{A}$, so definieren wir das ***Lebesgue–Integral von*** f ***über*** A (bzgl. μ) durch

$$\int_A f\,d\mu := \int_A f_+\,d\mu - \int_A f_-\,d\mu.$$

falls ein von $\int_A f_+\,d\mu$ und $\int_A f_-\,d\mu$ endlich ist. Wir sagen, dass f ***Lebesgue–integierbar über*** A ist, falls beide Integrale endlich sind. Eine äquivalente Definition erhalten wir, wenn wir die meßbare Funktion $f|_A : A \to \overline{\mathbb{R}}$ betrachten, und das Integral

$$\int_A f\,d\mu := \int_A f|_A\,d\mu_A.$$

als das Integral bzgl. der Maßraum $(A,\mathcal{A}^A,\mu_A)$ definieren, wobei $\mathcal{A}^A := \{B \in \mathcal{A} : B \subset A\}$ und $\mu_A := \mu|_{\mathcal{A}^A}$.

13.5.9. **Satz.** *Sei $(X,\mathcal{A},\mu)$ ein Maßraum und $f : X \to \overline{\mathbb{R}}$ eine meßbare Funktion. Die folgende Bedingungen sind äquivalent:*

(i) f ist integrierbar.

(ii) f_+ und f_- sind integrierbar.

(iii) Es gibt integrierbare Funktionen $\varphi \geqslant 0$, $\psi \geqslant 0$ mit $f = \varphi - \psi$.
In diesem Fall gilt $\int_X f\,d\mu = \int_X \varphi\,d\mu - \int_X \psi\,d\mu$.

(iv) Es gibt eine integrierbare Funktion $g \geqslant 0$ mit $|f| \leqslant g$.

(v) $|f|$ ist integrierbar.

13.5.10. **Satz.** *Sei $(X,\mathcal{A},\mu)$ ein Maßraum, $f,g : X \to \overline{\mathbb{R}}$ seien integrierbar.*

(i) $\alpha \in \mathbb{R} \Longrightarrow \alpha f$ ist integrierbar und $\int_X \alpha f\,d\mu = \alpha \int_X f\,d\mu$.

(ii) $f+g$ ist integrierbar, falls $f+g$ definiert ist und $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.

(iii) $\min\{f,g\}$, $\max\{f,g\}$ sind integrierbar.

(iv) $f \leqslant g \Longrightarrow \int_X f\,d\mu \leqslant \int_X g\,d\mu$.

(v) $\left|\int_X f\,d\mu\right| \leqslant \int_X |f|\,d\mu$.

13.5.11. **Folgerung.** *$\mathcal{L}^1(X,\mu)$ ist ein $\mathbb{R}$-Vektorraum und die Abbildung $\mathcal{L}^1(X,\mu) \to \mathbb{R}$, $f \mapsto \int_X f\,d\mu$ ist $\mathbb{R}$–linear.*

13.5.12. **Folgerung.** *Ist f integrierbar, so ist $N := \{x \in X : |f(x)| = \infty\}$ eine Nullmenge.*

13.5.13. **Folgerung.** *Sei $f : X \to \overline{\mathbb{R}}$ meßbar.*

(i) *Seien $A,B \in \mathcal{A}$, $A \subset B$, so dass f integrierbar auf B ist. Dann ist f integrierbar auf A.*

(ii) *Seien $A,B \in \mathcal{A}$, so dass f integrierbar auf A und B ist. Dann ist f integrierbar auf $A \cup B$ und*

$$\int_{A\cup B} f\,d\mu = \int_A f\,d\mu + \int_B f\,d\mu.$$

13.5.14. **Folgerung** (Modifikationssatz)**.**

(i) *Sei $f : X \to \overline{\mathbb{R}}$ integrierbar und $g : X \to \overline{\mathbb{R}}$ meßbar, so dass $f = g$ fast überall. Dann ist auch g integrierbar und gilt $\displaystyle\int_X f\,d\mu = \int_X g\,d\mu$.*

(ii) *Sei $N \subset X$ eine Nullmenge und $f : X \setminus N \to \overline{\mathbb{R}}$ integrierbar. Sei $\widetilde{f} : X \setminus N \to \overline{\mathbb{R}}$ eine meßbare Fortsetzung von f. Dann ist $\widetilde{f}$ integrierbar und gilt $\displaystyle\int_X \widetilde{f}\,d\mu = \int_{X\setminus N} f\,d\mu$. Wir setzen*

$$\int_X f\,d\mu := \int_X \widetilde{f}\,d\mu = \int_{X\setminus N} f\,d\mu \tag{13.12}$$

Wir sind oft in der Situation, dass eine Funktion f definiert und integriebar außerhalb einer Nullmenge ist. Der zweite Teil des Satzes besagt, dass wir das Integral von f auf dem ganzen Raum definieren dürfen, denn sie hängt nicht von der Auswahl von $\widetilde{f}$ und sie stimmt für auf ganz X definierte Funktionen mit der bisherigen Definition überein.

13.5.15. **Definition.** Sei $(X,\mathcal{A},\mu)$ ein Maßraum. Eine meßbare Funktion $f : X \to \mathbb{C}$ heißt ***Lebesgue–integierbar***, falls $\operatorname{Re} f, \operatorname{Im} f : X \to \mathbb{R}$ Lebesgue–integierbar sind. Das Integral von f wird definiert durch

$$\int_X f\,d\mu = \int_X \operatorname{Re} f\,d\mu + i\int_X \operatorname{Im} f\,d\mu \in \mathbb{C}.$$

Die Menge der integrierbaren Funktionen $f : X \to \mathbb{C}$ wird mit $\mathcal{L}^1_{\mathbb{C}}(X,\mu)$ bezeichnet.

13.5.16. **Satz.** *$\mathcal{L}^1_{\mathbb{C}}(X,\mu)$ ist ein $\mathbb{C}$-Vektorraum und die Abbildung $\mathcal{L}^1_{\mathbb{C}}(X,\mu) \to \mathbb{C}$, $f \mapsto \int_X f\,d\mu$ ist $\mathbb{C}$–linear. $f \in \mathcal{L}^1_{\mathbb{C}}(X,\mu)$ genau dann, wenn $|f| \in \mathcal{L}^1_{\mathbb{C}}(X,\mu)$ und dann gilt $\left|\int_X f\,d\mu\right| \leqslant \int_X |f|\,d\mu$.*

13.6. **Konvergenzsätze.**

13.6.1. **Satz** (Satz von Beppo Levi über monotone Konvergenz)**.** *Sei $f_k : X \longrightarrow [0,\infty]$ eine monoton wachsende Folge $\mathcal{A}$–meßbarer Funktionen: $f_1 \leqslant f_2 \leqslant \ldots \leqslant f_k \leqslant \ldots$.*

Sei $f : X \longrightarrow [0,\infty]$, $f(x) = \lim_{k\to\infty} f_k(x)$. Dann gilt:

$$\lim_{k\to\infty} \int_X f_k\,d\mu = \int_X f\,d\mu.$$

Bemerkung: Ohne Voraussetzung der Monotonie gilt der Satz nicht. Z.B.: Seien $f_n : \mathbb{R} \longrightarrow \mathbb{R}$, $f_n(x) = n\mathbb{1}_{[\frac{1}{n},\frac{2}{n}]}$, $f_n \to 0 =: f$ auf $\mathbb{R}$, aber $\int_{\mathbb{R}} f_n \, d\lambda_1 = 1 \not\to 0 = \int_{\mathbb{R}} f \, d\lambda_1$.

Gilt der Satz von Beppo Levi für fallende Folgen?

Beispiel: Sei $f_n : \mathbb{R} \longrightarrow \mathbb{R}$, $f_n = \mathbb{1}_{[n,\infty)}$. Es gilt $f_n \searrow 0 = f$, $\int f_n \, d\lambda_1 = \mu_1([n,\infty)) = \infty$ aber $\int f \, d\lambda_1 < 0$.

13.6.2. **Satz** (Satz von Beppo Levi für fallende Folgen). *Sei $f_k : X \longrightarrow [0,\infty]$ eine monoton fallende Folge $\mathcal{A}$–meßbarer Funktionen: $f_1 \geqslant f_2 \geqslant \ldots \geqslant f_k \geqslant \ldots$ mit $\int_X f_1 < \infty$.*

Sei $f : X \longrightarrow [0,\infty]$, $f(x) = \lim_{k\to\infty} f_k(x)$. Dann gilt:

$$\lim_{k\to\infty} \int_X f_k \, d\mu = \int_X f \, d\mu .$$

13.6.3. **Folgerung.** *Seien $f_n : X \longrightarrow [0,\infty]$ meßbar. Dann gilt*

$$\int_X \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int_X f_n \, d\mu .$$

13.6.4. **Satz** (Lemma von Fatou). *Seien $f_n : X \longrightarrow [0,\infty]$ meßbar. Dann gilt*

(i) $\int_X \liminf f_n \leqslant \liminf \int_X f_n$,

(ii) *Ist $\sup_{n\geqslant 1} f_n$ integrierbar, dann ist $\limsup \int f_n \leqslant \int \limsup f_n$.*

Bemerkung: für „<“ in (i) siehe Bemerkung nach Satz von Beppo Levi.

13.6.5. **Satz** (von Lebesgue über majorisierte Konvergenz). *Seien $f_n, f : X \longrightarrow \mathbb{C}$ ($n \in \mathbb{N}$) meßbar, so dass gilt:*

(i) $\lim_{n\to\infty} f_n = f$ *μ–fast überall,*

(ii) *es gibt eine integrierbare Funktion $g : X \longrightarrow [0,\infty]$, so dass für alle $n \in \mathbb{N}$ gilt $|f_n| \leqslant g$ μ–fast überall.*

Dann sind f_n, f integrierbar und es gilt

$$\lim_{n\to\infty} \int_X |f_n - f| \, d\mu = 0 , \quad \lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu .$$

Der Satz von Lebesgue über majorisierte Konvergenz ist der am häufigsten benutzte Konvergenzsatz (neben dem Satz von Beppo Levi). Wir verlangen lediglich *punktweise Konvergenz* von $(f_n)_n$. Der entsprechende Satz für Regelintegrale oder Riemann-Integrale setzt *gleichmäßige Konvergenz* auf kompakte Intervalle voraus. Die wesentliche Hypothese ist die Existenz der integrierbaren Majorante g.

Beispiel: Die Funktionenfolgen $(n^2\mathbb{1}_{[0,\frac{1}{n}]})_{n\in\mathbb{N}}$, $(\mathbb{1}_{[n,n+1]})_{n\in\mathbb{N}}$, $(\frac{1}{n}\mathbb{1}_{[0,n]})_{n\in\mathbb{N}}$ von λ_1– integrierbare Funktionen konvergiert überall gegen 0, die Grenzwertfunktion ist integrierbar, aber die Integralenfolge konvergiert nicht, weil die Funktionenfolge nicht dominiert ist. Z.B. sei g eine beliebige Funktion, die alle $f_n = \frac{1}{n}\mathbb{1}_{[0,n]}$ majoriert, d.h. $f_n \leqslant g$ (alle Funktionen sind positiv). Sei $x \in \mathbb{R}_+$ beliebig, $n \in \mathbb{Z}_+$ mit $n \leqslant x < n+1$. Da auch $g \geqslant f_{n+1}$ gilt, haben wir für alle $x \in \mathbb{R}$ die Abschätzung $g(x) \geqslant h(x) := \mathbb{1}_{[0,+\infty)} \frac{1}{[x]+1}$, wobei $[x]$ den ganzen Teil von x bezeichnet. Aber h ist nicht integrierbar: $\int h \, d\lambda_1 = \sum_{n=0}^{\infty} \frac{1}{n+1} = \infty$.

13.6.6. **Satz** (Ausschöpfungssatz). *Seien $(X,\mathcal{A},\mu)$ ein Maßraum, $(X_n)_{n\in\mathbb{N}}$ mit $X_n \in \mathcal{A}$, $X_n \subset X_{n+1}$ und $X = \cup_{n\in\mathbb{N}} X_n$. Sei $f : X \longrightarrow \overline{\mathbb{R}}$ meßbar. f ist integrierbar genau dann, wenn:*

(i) *f ist über jedem X_n integrierbar und*

(ii) *$(\int_{X_n} |f| \, d\mu)_{n\in\mathbb{N}}$ konvergiert in $\mathbb{R}$.*

In diesem Fall:

$$\lim_{n\to\infty} \int_{X_n} f \, d\mu = \int_X f \, d\mu .$$

13.6.7. **Satz.** *Sei $(f_k)_k$ eine Folge von meßbaren Funktionen $f_k : X \to \overline{\mathbb{R}}$ mit*

$$\sum_{k=1}^{\infty} \int |f_k| \, d\mu < +\infty .$$

Dann konvergiert die Reihe $\sum_{k=1}^{\infty} f_k$ fast überall und

$$\int \Big(\sum_{k=1}^{\infty} f_k \Big) d\mu = \sum_{k=1}^{\infty} \int f_k \, d\mu. \tag{13.13}$$

Beweis: Nach Folgerung 13.6.3 gilt $\int \sum_{k=1}^{\infty} |f_k| \, d\mu = \sum_{k=1}^{\infty} \int |f_k| \, d\mu < \infty$ also $\sum_{k=1}^{\infty} |f_k|$ ist integrierbar und damit endlich fast überall (Folgerung 13.5.12). Die Folge $(\sum_{k=1}^{n} f_k)_n$ hat die integrierbare Majorante $\sum_{k=1}^{\infty} |f_k|$ und $\lim_{n\to\infty} \sum_{k=1}^{n} f_k = \sum_{k=1}^{\infty} f_k$. Nach dem Satz von Lebesgue folgt (13.13). □

13.7. **Vergleich zwischen Lebesgue– und Regelintegral.**

13.7.1. **Satz.** *Sei $f \in \mathcal{R}([a,b])$ eine Regelfunktion. Dann ist f Lebesgue–integrierbar auf $[a,b]$ und*

$$\int_{[a,b]} f\,d\lambda_1 = \int_a^b f\,dx.$$

wobei die rechte Seite das Regelintegral von f ist.

13.7.2. **Bemerkung.**
(1) Der Satz gilt auch für Riemann–integriebare Funktionen.
(2) Die Dirichlet–Funktion $\mathbb{1}_{[0,1]\cap\mathbb{Q}}$ ist nirgends stetig (also keine Regelfunktion und nicht Riemann–integrierbar.) Sie ist aber Lebesgue–integrierbar, sogar einfache Funktion mit

$$\int \mathbb{1}_{[0,1]\cap\mathbb{Q}}\,d\lambda_1 = 0.$$

(3) Weil $\int_{[a,b]} f\,d\lambda_1 = \int_a^b f\,dx$ für $f \in \mathcal{R}([a,b])$ schreiben wir $\int_a^b f\,dx$ für das Lebesgue–Integral einer belibigen Funktion $f \in \mathcal{L}^1([a,b],\lambda_1)$.

13.7.3. **Satz** (uneigentliches Regelintegral und Lebesgue–Integral)**.** *Seien I ein beliebiges Intervall in $\mathbb{R}$ und $f : I \longrightarrow \mathbb{R}$ eine Regelfunktion. f ist Lebesgue–integrierbar genau dann, wenn $|f|$ uneigentlich regelintegrierbar ist. Dann stimmt das uneigentliche Regelintegral von f über I mit dem Lebesgue–Integral überein.*

13.7.4. **Bemerkung.** Sei $f : [0,\infty) \longrightarrow \mathbb{R}$ die stetige Funktion

$$f(x) = \begin{cases} \dfrac{\sin x}{x}, & x \neq 0 \\ 1, & x = 0 \end{cases}$$

Dann ist f uneigentlich regelintegrierbar auf $[0,\infty)$ aber nicht absolut uneigentlich regelintegrierbar. Nach dem Satz 13.7.3 kann f nicht Lebesgue–integrierbar sein. Wir haben also folgenden Verhältnisse:

f ist Lebesgue–integrierbar auf $I \iff |f|$ Lebesgue–integrierbar auf I.
f ist Lebesgue–integrierbar auf $I \iff |f|$ uneigentlich regelintegrierbar auf I.
$|f|$ uneigentlich regelintegrierbar $\Longrightarrow f$ uneigentlich regelintegrierbar auf I.
f uneigentlich regelintegrierbar $\not\Longrightarrow f$ uneigentlich regelintegrierbar auf I.

Vereinbarung: Sei $I \subset \mathbb{R}$ ein Intervall, $a = \inf I$, $b = \sup I$. I darf offen, halboffen oder unbeschränkt sein ($a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{+\infty\}$). Wir schreiben

$$\int_I f\,d\lambda_1 = \int_a^b f\,d\lambda_1 = \int_a^b f\,dx, \quad \text{für alle } f \in \mathcal{L}^1(I,\lambda_1).$$

Diese Schreibweise ist dadurch gerechtfertigt, dass $\{a\}$ und $\{b\}$ Nullmengen sind (falls $a,b \in \mathbb{R}$). Ist z.B. $f : [a,b] \longrightarrow \mathbb{R}$ Lebesgue–integrierbar auf $(a,b]$, dann ist f Lebesgue–integrierbar auf $[a,b]$ und $\int_{[a,b]} f\,d\lambda_1 = \int_{(a,b]} f\,d\lambda_1$.

13.7.5. **Satz.** *Sei $D \subset \mathbb{R}^n$ Lebesgue–meßbar. Ist $f : D \longrightarrow \mathbb{R}$ beschränkt und fast überall stetig, und gilt $\lambda_n(\operatorname{supp} f) < \infty$, so ist f Lebesgue–integrierbar und gilt*

$$\left|\int_D f\,d\lambda_n\right| \leqslant \int_D |f|\,d\lambda_n \leqslant \sup|f| \cdot \lambda_n(\operatorname{supp} f). \tag{13.14}$$

Insbesondere ist jede stetige Funktion $f : D \longrightarrow \mathbb{R}$ mit kompaktem Träger $\operatorname{supp} f$ *integrierbar.*

Beweis: Weil f fast überall stetig ist, ist f meßbar. Außerdem gilt $|f| \leqslant \sup|f| \cdot \mathbb{1}_{\operatorname{supp} f}$ und die letzte Funktion ist integrierbar: $\int_D \sup|f| \cdot \mathbb{1}_{\operatorname{supp} f}\,d\lambda_n = \sup|f| \cdot \lambda_n(\operatorname{supp} f) < \infty$. □

13.7.6. **Bemerkung.** Ein Satz von Lebesgue besagt:

$f : [a,b] \to \mathbb{R}$ *ist Riemann–integrierbar* $\iff$ *f beschränkt und fast überall stetig.*

Es gibt also die Inklusionen:

$$\mathcal{R}([a,b]) \subset \{f : [a,b] \to \mathbb{R} : f \text{ Riemann–integrierbar}\} \subset \mathcal{L}^1([a,b])$$

Alle diese Inklusionen sind echt. Siehe auch die Bemerkungen 6.8.6, 13.7.2.

13.8. Von einem Parameter abhängige Integrale.

13.8.1. **Satz** (stetige Abhängigkeit des Integrals von einem Parameter). *Seien T ein metrischer Raum, $t_0 \in T$, $(X, \mathcal{A}, \mu)$ ein Maßraum und $f : T \times X \to \mathbb{R}$ mit der Eigenschaften:*

(i) *Für alle $t \in T$ ist $f(t,\cdot) : X \to \mathbb{R}$ meßbar.*
(ii) *Es gibt eine Nullmenge $N = N(t_0) \subset X$, so dass für alle $x \in X \setminus N$ gilt: $f(\cdot, x)$ ist stetig in t_0.*
(iii) *Es gibt eine integrierbare Funktion $g : X \to [0,\infty]$, so dass für alle $t \in T$ eine Nullmenge $N(t) \subset X$ existiert mit $|f(t,x)| \leqslant g(x)$ für $x \in X \setminus N(t)$.*

Dann ist die Funktion $F : T \to \mathbb{R}$, $F(t) = \displaystyle\int_X f(t,x)d\mu(x)$, stetig im Punkte t_0.

13.8.2. **Bemerkung.**

(1) In vielen Anwendungen findet man eine Majorante g nur auf einer Umgebung U von t_0. Das reicht um die Stetigkeit in t_0 zu zeigen, da die Stetigkeit eine lokale Eigenschaft ist. Sei

$$f : [0,\infty) \times (0,\infty) \to \mathbb{R}, \quad f(x,t) = \frac{1}{t + x^2}\cdot \tag{13.15}$$

Für alle $t \in (0,\infty)$ ist $f(\cdot, t)$ auf $[0,\infty)$ integrierbar und wir betrachten das von dem Parameter t abhängigen Intergral

$$F : (0,\infty) \to \mathbb{R}, \quad F(t) = \int_0^\infty f(x,t)dx.$$

In der Tat, mit Hilfe der Substitution $y = \frac{x}{\sqrt{t}}$ sehen wir leicht, dass

$$F(t) = \frac{1}{\sqrt{t}} \int_0^\infty \frac{1}{1+y^2}dy = \frac{\pi}{2}\frac{1}{\sqrt{t}}\cdot$$

Wir müssen nun eine eine von t unabhängige Majorante g wie in (iii) finden. Wegen $x^2 \leqslant t + x^2$ für alle $t > 0$ die beste Majorante ist

$$\frac{1}{t+x^2} \leqslant \frac{1}{x^2} =: g(x),$$

sie ist leider auf $(0,\infty)$ nicht integrierbar (wegen des Verhaltens in der Umgebung von Null). Um die Stetigkeit von F zu zeigen, legen wir t_0 fest und betrachten die Umgebung $(t_0/2, \infty)$. Für alle $t \in (t_0/2,\infty)$ und $x \in [0,\infty)$ gilt

$$\frac{1}{t+x^2} \leqslant \frac{1}{(t_0/2)+x^2} = g_0(x)$$

und die Funktion g_0 ist eine von $t \in (t_0/2,\infty)$ unabhängige Majorante von $f(\cdot,t)$. Nach dem Satz ist F stetig in t_0. Da t_0 beliebig ist, ist F stetig.

(2) Wenn wir über eine kompakte Menge integrieren und f stetig ist, ist eine lokale Majorante automatisch vorhanden.

Seien $T \subset \mathbb{R}^m$ und $K \subset \mathbb{R}^n$ eine kompakte Menge. Sei $f : T \times K \to \mathbb{R}$ stetig. Dann ist die Funktion $F : T \to \mathbb{R}$, $F(t) = \displaystyle\int_K f(t,x)d\lambda_n(x)$, stetig.

Beweis: Sei U eine kompakte Umgebung von t_0 in T. Dann ist $U \times K$ kompakt und die stetige Funktion f ist beschränkt auf $U \times K$. Sei $M := \sup_{U\times K}|f|$. Die Funktion $g = M\mathbb{1}_K$ ist eine integrierbare Majorante für f auf $U \times K$. □

(3) Die Funktion $f : [0,\infty) \times [0,\infty) \to \mathbb{R}$, $f(t,x) = te^{-tx}$ ist stetig, also sind beide Voraussetzungen (i) und (ii) des Satzes für $t_0 = 0$ erfüllt. Für $t \in (0,\infty)$ ist die Funktion $f(t,\cdot)$ integrierbar, und

$$F(t) = \int_0^\infty te^{-tx}dx = -e^{-tx}\Big|_{x=0}^{\infty} = 0 - (-1) = 1.$$

Für $t = 0$ ist die Funktion $f(0,\cdot) = 0$ integrierbar mit $F(0) = \int_0^\infty f(0,x)dx = 0$. Die Funktion $F : [0,\infty) \to \mathbb{R}$, $F(t) = \int_0^\infty f(t,x)dx$ ist also nicht stetig in $t_0 = 0$. Man kann keine integrierbare Majorante g von $f(t,\cdot)$ für alle t in einer Umgebung von Null finden.

13.8.3. **Definition.** Sei $f \in \mathcal{L}^1_{\mathbb{C}}(\mathbb{R}^n, d\lambda_n)$. Dann heißt $\widehat{f} : \mathbb{R}^n \longrightarrow \mathbb{C}$,

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-i\langle \xi, x\rangle}d\lambda_n(x)$$

die ***Fourier–Transformierte*** von f.

13.8.4. **Folgerung** (Riemann-Lebesgue Lemma). *Die Fourier–Transformierte $\widehat{f}$ von f ist stetig auf $\mathbb{R}^n$ und gilt $\lim_{\|\xi\|\to\infty}\widehat{f}(\xi) = 0$.*

13.8.5. **Satz** (Differentiation unter Integralzeichen). *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \times X \longrightarrow \mathbb{R}$. Es gelte:*

(i) *$f(t,\cdot) \in \mathcal{L}^1(X,d\mu)$ für jedes $t \in I$ und*

(ii) *Es gibt eine Nullmenge $N \subset X$ und $g \in \mathcal{L}^1(X,d\mu)$ so, dass für jedes $t \in I$ und $x \in X \setminus N$ existiert $\frac{\partial f}{\partial t}(t,x)$ und $|\frac{\partial f}{\partial t}(t,x)| \leqslant g(x)$.*

Dann ist die Funktion

$$F : I \longrightarrow \mathbb{R}, \quad F(t) = \int_X f(t,x)\,d\mu(x) \tag{13.16}$$

differenzierbar und

$$F'(t) = \frac{d}{dt}\int_X f(t,x)\,d\mu(x) = \int_X \frac{\partial f}{\partial t}(t,x)\,d\mu(x). \tag{13.17}$$

Gilt zusätzlich in (ii), dass $\frac{\partial f}{\partial t}(\cdot,x)$ stetig für alle $x \in X \setminus N$ ist, so ist F der Klasse $\mathcal{C}^1$ auf I

Unter der Hypothesen des Satzes kann man also Differentiation und Integration vertauschen.

13.8.6. **Bemerkung.**

(1) Eine ähnliche Bemerkung wie in 13.8.2 gilt auch für die Differentiation unter Integral. Betrachten wir die Funktion (13.15). f ist differenzierbar und

$$\frac{\partial f}{\partial t}(x,t) = -\frac{1}{(t+x^2)^2}\,.$$

Legen wir t_0 fest und betrachten die Umgebung $(t_0/2,\infty)$ von t_0. Für alle $t \in (t_0/2,\infty)$ und $x \in [0,\infty)$ gilt

$$\Big|\frac{\partial f}{\partial t}(x,t)\Big| = \Big|\frac{1}{(t+x^2)^2}\Big| \leqslant \frac{1}{((t_0/2)+x^2)^2} \leqslant \frac{1}{(t_0/2)+x^2} =: g_0(x)$$

Die Funktion $g_0 : [0,\infty) \to \mathbb{R}$ ist Lebesgue-integierbar. Wir können also unter Integralzeichen differenzieren und erhalten

$$F'(t_0) = \int_0^\infty \frac{\partial f}{\partial t}(x,t_0)\,dx = -\int_0^\infty \frac{1}{(t_0+x^2)^2}\,dx\,.$$

(2) Wenn wir über eine kompakte Menge integrieren und f stetig differenzierbar ist, ist eine lokale Majorante automatisch vorhanden.
Seien $I \subset \mathbb{R}$ ein Intervall, $K \subset \mathbb{R}^n$ eine kompakte Menge und $f : I \times K \to \mathbb{R}$ stetig, so dass $\frac{\partial f}{\partial t}(t,x)$ existiert und ist stetig auf $I \times K$. Dann ist die Funktion F aus (13.16) *der Klasse $\mathcal{C}^1$ auf I und gilt* (13.16).

(3) Die Abbidung $f : \mathbb{R} \times [0,\infty) \longrightarrow \mathbb{R}$, $f(t,x) = t^2 e^{-x|t|}$ ist der Klasse $\mathcal{C}^1$ und $\frac{\partial f}{\partial t}(t,x) = (2-x|t|)te^{-x|t|}$ für alle $(t,x) \in \mathbb{R} \times [0,\infty)$. Für $t \neq 0$ gilt

$$F(t) = \int_0^\infty f(t,x)\,dx = t^2\int_0^\infty e^{-x|t|}\,dx = t^2\Big[-\frac{1}{|t|}e^{-x|t|}\Big]_{x=0}^\infty = |t|$$

und die Gleichung $F(t) = |t|$ gilt auch für $t = 0$. F ist also nicht differenzierbar auf $\mathbb{R}$. Die Ursache ist, dass es keine integrierbare Majorante $g : [0,\infty) \longrightarrow \mathbb{R}$, $|\frac{\partial f}{\partial t}(t,x)| \leqslant g(x)$ für alle t in einer Umgebung von Null gibt.

13.8.7. **Folgerung.** *Seien $f \in \mathcal{L}^1(\mathbb{R}^n,d\lambda_n)$ und $x_j f \in \mathcal{L}^1(\mathbb{R}^n,d\lambda_n)$ für jedes $1 \leqslant j \leqslant n$. Dann ist $\widehat{f} \in \mathcal{C}^1(\mathbb{R}^n)$ und $\dfrac{\partial \widehat{f}}{\partial x_j} = (-i)\widehat{x_j f}$.*

13.8.8. **Bemerkung.** Die Nullmengen spielen verschiedene Rollen in den Sätzen 13.8.1 und 13.8.5. In Satz 13.8.1 (ii) darf die Nullmenge $N(t_0)$ von t_0 abhängen, während in Satz 13.8.5 ist die Nullmenge N unabhängig von $t \in I$. Wir erklären das durch ein Beispiel. Sei $I = [0,1]$, $(X,\mathcal{A},\mu) = ([0,1],\mathcal{L}([0,1]),\lambda_1)$ und setze

$$f : [0,1] \times [0,1] \longrightarrow \mathbb{R}, \quad f(t,x) = \begin{cases} 1, & x < t \\ 0, & x \geqslant 0\,. \end{cases}$$

Offensichtlich $F(t) = \displaystyle\int_0^1 f(t,x)\,dx = t$. Die Voraussetzungen des Satzes 13.8.1 sind erfüllt:

* für fast alle x (eigentlich für $x \neq t_0$, d.h. $N(t_0) = \{t_0\}$) ist $t \longrightarrow f(t,x)$ stetig in t_0
* $|f(t,x)| \leqslant 1$ für alle $t \in [0,1]$.

Der Satz 13.8.1 liefert somit einen neuen Beweis, dass die Funktion $F(t) = t$ stetig ist!

Sei nun $t_0 \in [0,1]$ fest. Die partielle Ableitung $\frac{\partial f}{\partial t}(t_0,x)$ existiert für fast alle x (eigentlich für alle $x \neq t_0$) und sie ist Null. Die Voraussetzung des Satzes 13.8.5 verlangt aber mehr: wir müßten eine Nullmenge $N \subset [0,1]$ finden, so dass für alle $x \notin N$ existiert $\frac{\partial f}{\partial t}(t,x)$. Das ist klar unmöglich, da für alle x existiert $\frac{\partial f}{\partial t}(x,x)$ nicht. Würden wir unter Integralzeichen ableiten, so gälte:

$$1 = F'(t) = \int_0^1 \frac{\partial f}{\partial t}(t,x)\,dx = \int_0^1 0\,dx = 0\,, \quad \text{Widerspruch.}$$

13.8.9. ***Aufgabe***. Sei $f : \mathbb{R} \to \mathbb{R}$ der Klasse $\mathcal{C}^1$ und definiere

$$F : \mathbb{R}^2 \to \mathbb{R}, \quad F(x,y) = \begin{cases} \frac{f(x)-f(y)}{x-y} & x \neq y\,, \\ f'(x) & x = y\,. \end{cases}$$

(a) Zeige, dass F stetig ist.
(b) Ist $f \in \mathcal{C}^2$, so ist $F \in \mathcal{C}^1$.

13.9. **Produktmaße und der Satz von Fubini.**

13.9.1. **Definition.** Seien $(X,\mathcal{A},\mu)$, $(Y,\mathcal{B},\nu)$ Maßräume, $\mathcal{A} * \mathcal{B} := \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$. Die ***Produkt σ–Algebra*** $\mathcal{A} \otimes \mathcal{B}$ ist die σ–Algebra $\sigma(\mathcal{A} * \mathcal{B})$.

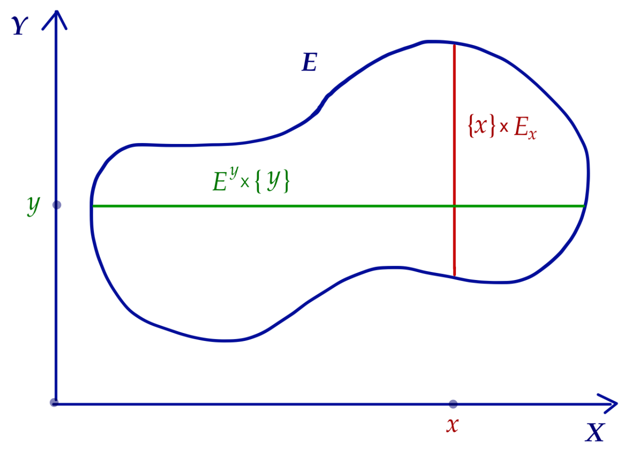
Das Symbol $\mathcal{A} \otimes \mathcal{B}$ sollte keine Angst machen. Es ist nicht möglich $\mathcal{A} \times \mathcal{B}$ zu nutzen, da $\mathcal{A} \times \mathcal{B}$ das Kartesische Produkt bezeichnet.

13.9.2. **Bemerkung.** Sei $\mathcal{A} = \sigma(\mathcal{M})$, $\mathcal{B} = \sigma(\mathcal{N})$ mit $\mathcal{M} \subset \mathcal{P}(X)$, $\mathcal{N} \subset \mathcal{P}(Y)$. Dann gilt

$$\mathcal{A} \otimes \mathcal{B} = \sigma(\{M \times N, X \times N, M \times Y : M \in \mathcal{M}, N \in \mathcal{N}\})$$

Daraus folgt, dass $\mathcal{B}(\mathbb{R}^p) \otimes \mathcal{B}(\mathbb{R}^q) = \mathcal{B}(\mathbb{R}^{p+q})$ (betrachte die Mengensysteme $\mathcal{M} = \mathcal{R}_p$ bzw. $\mathcal{N} = \mathcal{R}_q$ der Figuren in $\mathbb{R}^p$ bzw. $\mathbb{R}^q$).

13.9.3. **Lemma.** *Sei $M \in \mathcal{A} \otimes \mathcal{B}$. Dann liegen für alle $a \in X$, $b \in Y$ die Schnitte $M_a = \{y \in Y : (a,x) \in M\}$ bzw. $M^b = \{x \in X : (x,b) \in M\}$ in $\mathcal{A}$ bzw. $\mathcal{B}$. Für jede meßbare Abbildung $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}) \longrightarrow (Z, \mathcal{C})$ sind die Schnitte $f(a,\cdot) : (Y,\mathcal{B}) \longrightarrow (Z,\mathcal{C})$, $f(\cdot,b) : (X,\mathcal{A}) \longrightarrow (Z,\mathcal{C})$ meßbar.*



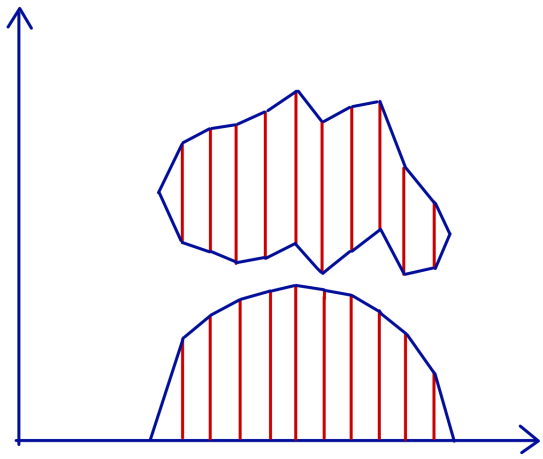
13.9.4. **Satz.** *Seien μ und ν σ–endlich. Dann gilt:*
(i) *Es gibt genau ein Maß $\mu \otimes \nu : \mathcal{A} \otimes \mathcal{B} \longrightarrow [0,\infty]$ mit*

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B) \quad A \in \mathcal{A}, B \in \mathcal{B}$$

*Das Maß $\mu \otimes \nu$ heißt **Produktmaß** von μ und ν.*
(ii) *Für jedes $M \in \mathcal{A} \otimes \mathcal{B}$ sind die Funktionen $X \ni x \longrightarrow \nu(M_x) \in [0,\infty]$ und $Y \ni y \longrightarrow \mu(M^y) \in [0,\infty]$ meßbar und gilt*

$$\boxed{(\mu \otimes \nu)(M) = \int_X \nu(M_x)\,d\mu(x) = \int_Y \mu(M^y)\,d\nu(y)}\,. \tag{13.18}$$

Das Produktmaß einer Menge $M \in \mathcal{A} \otimes \mathcal{B}$ berechnet man, indem man die Maße der Schnitte integriert. Die Gleichheit (13.18) heißt ***Prinzip von Cavalieri***.



Zur Illustrierung, im Bild oben betrachten wir eine Menge $M \subset \mathbb{R}^n \times \mathbb{R}$ und der Graph der Funktion $f : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto \lambda_1(M_x)$. Die vertikalen Schitte über $x \in \mathbb{R}^n$ von M und der Ordinatenmengen dieser Funktion haben dasselbe Mass per Konstruktion. Das Mass von M ist das Integral der Funkftion f und das ist zugleich das Mass der Ordinatenmenge von f. Das Prinzip von Cavalieri kann man auch so aussprechen: Seien $M, N \in \mathcal{A} \otimes \mathcal{B}$, so dass $\nu(M_x) = \nu(N_x)$ für alle $x \in X$, so ist $\mu \otimes \nu(M) = \mu \otimes \nu(N)$. Das wurde schon von Archimedes benutzt, um das Volumen der dreidimensionalen Kugel zu berechnen. Er betrachtete eine Halbkugel M von Radius R über die xy-Ebene und ein Kreiszylinder vom Radius R mit der Höhe R mit der z-Achse als Rotationsachse. Aus dem Zylinder entfernte er einen Kreiskegel mit der Spitze 0, der die Grundfläche des Zylinders zur Basis hat; das ergibt einen Restkörper N. Man sieht leicht, dass M_z und N_z dieselbe Fläche haben, somit haben M und N dasseble Volumen.

13.9.5. **Beispiel.**
(1) Es gilt $(\mathbb{R}^{p+q}, \mathcal{B}(\mathbb{R}^{p+q}), \lambda_{p+q}) = (\mathbb{R}^{p+q}, \mathcal{B}(\mathbb{R}^p) \otimes \mathcal{B}(\mathbb{R}^q), \lambda_p \otimes \lambda_q)$. Für die σ-Algebren der Lebesgue-meßbaren Mengen gilt $\mathcal{L}(\mathbb{R}^p) \otimes \mathcal{L}(\mathbb{R}^q) \subsetneq \mathcal{L}(\mathbb{R}^{p+q})$ aber $(\mathbb{R}^{p+q}, \mathcal{L}(\mathbb{R}^{p+q}), \lambda_{p+q})$ ist die Vervollständigung von $(\mathbb{R}^{p+q}, \mathcal{L}(\mathbb{R}^p) \otimes \mathcal{L}(\mathbb{R}^q), \lambda_p \otimes \lambda_q)$. Alle $M \in \mathcal{L}(\mathbb{R}^{p+q})$ haben eine Darstelung $M = N \cup B$, mit $B \in \mathcal{B}(\mathbb{R}^{p+q})$ und N eine Lebesgue–Nullmenge. Deshalb gilt

$$\lambda_{p+q}(M) = \int_{\mathbb{R}^p} \lambda_q(M_x)\, d\lambda_p(x) = \int_{\mathbb{R}^q} \lambda_p(M^y)\, d\lambda_q(y), \quad \text{für alle } M \in \mathcal{L}(\mathbb{R}^{p+q}).$$

(2) Seien $A \in \mathcal{L}(\mathbb{R}^n)$ und $f : A \longrightarrow [0,\infty]$ Lebesgue–meßbar. Wir definieren die Ordinatenmenge $U_f = \{(x,y) \in \mathbb{R}^{n+1} : x \in A, y \in [0, f(x)]\}$. Dann gilt

$$\boxed{\lambda_{n+1}(U_f) = \int_A f(x)\, d\lambda_n(x)}.$$

Das Integral entspricht unserer Vorstellung von einem Flächeninhalt zwischen dem Graphen und der x-Achse.
(3) Wir berechnen die Oberfläche der Kreisscheibe vom Radius R in $\mathbb{R}^2$. Für $f : [R, -R] \to \mathbb{R}$, $f(t) = \sqrt{R^2 - t^2}$ gilt $U_f = \overline{B}_R(0)$. Daraus folgt mit der Substitution $t = R\cos\varphi$:

$$\lambda_2(\overline{B}_R(0)) = 2\int_{-R}^{R} \sqrt{R^2 - t^2}\, dt = 2R^2 \int_{-\pi/2}^{\pi/2} \cos^2\varphi\, d\varphi$$

Aber

$$\int_{-\pi/2}^{\pi/2} \sin^2\varphi\, d\varphi + \int_{-\pi/2}^{\pi/2} \cos^2\varphi\, d\varphi = \pi$$

und

$$\int_{-\pi/2}^{\pi/2} \cos^2\varphi\, d\varphi = \int_{-\pi/2}^{\pi/2} \cos\varphi \sin'\varphi\, d\varphi = \cos\varphi\sin\varphi\Big|_{-\pi/2}^{\pi/2} + \int_{-\pi/2}^{\pi/2} \sin^2\varphi\, d\varphi = \int_{-\pi/2}^{\pi/2} \sin^2\varphi\, d\varphi.$$

also $\int_{-\pi/2}^{\pi/2} \cos^2 \varphi \, d\varphi = \frac{\pi}{2}$ und $\boxed{\lambda_2(\overline{B}_R(0)) = \pi R^2}$.

(4) Das Volumen der n-dimensionalen Kugel (I). Sei $n \in \mathbb{N}$ und $R > 0$. Setze $\omega_n(R) := \lambda_n(\bar{B}_R(0))$, $\omega_n(1) =: \omega_n$. Wir vereinbaren, dass $\omega_0 = 1$. Es ist klar, dass $\omega_1(R) := \lambda_1([-R,R]) = 2R$. Nach dem Prinzip von Cavalieri (13.18) gilt

$$\omega_n(R) = \int_{-R}^{R} \omega_{n-1}\left(\sqrt{R^2 - x^2}\right) dx$$

Behauptung:

$$\omega_n(R) = R^n \, \omega_n \, .$$

Beweis durch Induktion: $n = 1\checkmark$ (da $\omega_1 = 2$); $n - 1 \Rightarrow n$:

$$\begin{aligned}\omega_n(R) &= \int_{-R}^{R} \omega_{n-1}\left(\sqrt{R^2 - x^2}\right) dx = \omega_{n-1} \int_{-R}^{R} \left(R^2 - x^2\right)^{\frac{n-1}{2}} dx \\ &= \omega_{n-1} \int_{-1}^{1} \left(R^2 - R^2 t\right)^{\frac{n-1}{2}} R \, dt \\ &= R^n \omega_{n-1} \int_{-1}^{1} \left(1 - t^2\right)^{\frac{n-1}{2}} dt = R^n \omega_n\end{aligned}$$

Dies zeigt auch, dass

$$\omega_n = \omega_{n-1} \cdot I_n \, , \quad n \in \mathbb{N}_0$$

wobei

$$I_n := \int_{-1}^{1} \left(\sqrt{1 - t^2}\right)^{n-1} dt = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^n \phi \, d\phi, \quad n \in \mathbb{N}$$

(Substitution $[-\frac{\pi}{2}, \frac{\pi}{2}] \ni \phi \mapsto \sin\phi = t \in [-1,1]$). Es ist klar, dass

$$I_0 = \pi, \qquad I_1 = 2$$

und wie in (6.4) ergibt sich durch partielle Integration

(13.19) $$I_n = \int_0^{\frac{\pi}{2}} \cos^n x \, dx = \frac{1}{n} \cos^{n-1} x \sin x \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} + \frac{n-1}{n} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{n-2} x \, dx = \frac{n-1}{n} I_{n-2} \, .$$

also

(13.20) $$I_{2n} = \frac{2n-1}{2n} \cdot \frac{2n-3}{2n-2} \cdots \frac{1}{2} \cdot I_0 \, , \quad I_{2n+1} = \frac{2n}{2n+1} \cdot \frac{2n-2}{2n-1} \cdots \frac{2}{3} \cdot I_1 \, .$$

Folglich

(13.21) $$I_{n-1} I_n = \frac{1}{n} I_0 I_1 = \frac{2\pi}{n} \, .$$

Somit gilt

$$\omega_n = I_n \cdot \omega_{n-1} = I_{n-1} \cdot I_n \cdot \omega_{n-2} = \frac{2\pi}{n} \, \omega_{n-2}$$

und daher

$$\begin{aligned}\omega_{2m} &= \frac{2\pi}{2m} \omega_{m-2} = \frac{2\pi}{2m} \frac{2\pi}{2(m-1)} \omega_{2m-4} = \cdots = \frac{(2\pi)^{m-1}}{2m \cdot 2(m-1) \cdot 2(m-2) \cdot \cdots \cdot 4} \omega_2 \\ &= \frac{(2\pi)^{m-1}}{2^{m-1} \cdot m!} \omega_2 = \frac{\pi^m}{m!}\end{aligned}$$

und

$$\begin{aligned}\omega_{2m+1} &= \frac{2\pi}{2m-1} \omega_{2m-1} = \frac{2\pi}{2m+1} \cdot \frac{2\pi}{2m-1} \omega_{2m-3} \\ &= \frac{(2\pi)^m}{(2m+1)(2m-1)\ldots 3} \omega_1 = \frac{2^{m+1} \cdot \pi^m}{(2m+1)\ldots 3}\end{aligned}$$

Schließlich

(13.22) $$\boxed{\omega_{2m} = \frac{\pi^m}{m!} \, , \qquad \omega_{2m+1} = \frac{2^{m+1} \cdot \pi^m}{(2m+1)\ldots 3} \, , \quad m \in \mathbb{N}_0}$$

Eine überraschende Konsequenz hiervon ist $\omega_n \to 0$ für $n \to \infty$ (Beweis?).

Wir können die Formel für ω_{2m} und ω_{2m+1} einheitlich schreiben:

(13.23) $$\boxed{\omega_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)} \, , \quad n \in \mathbb{N}_0}$$

In der Tat, bezeichnen wir mit κ_n die rechte Seite. Wegen der Funktionalgleichung der Gamma-Funktion, $\Gamma(x+1) = x\Gamma(x)$ gilt

$$\frac{\kappa_n}{\kappa_{n-2}} = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} \cdot \frac{\Gamma\left(\frac{n-2}{2}+1\right)}{\pi^{\frac{n-2}{2}}} = \pi \cdot \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)} = \pi \cdot \frac{\Gamma\left(\frac{n}{2}\right)}{\frac{n}{2}\Gamma\left(\frac{n}{2}\right)} = \frac{2\pi}{n}$$

Dies bedeutet, dass die Folgen $(\omega_n)_{n\in\mathbb{N}_0}$ und $(\kappa_n)_{n\in\mathbb{N}_0}$ dieselbe Rekursionsformel genügen. Wir berechnen nun:

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty t^{-\frac{1}{2}} \cdot e^{-t} dt = \int_0^\infty \frac{1}{s} \cdot e^{-s^2} 2s\, ds = 2 \cdot \int_0^\infty e^{-s^2} ds = 2 \cdot \frac{\sqrt{\pi}}{2} = \sqrt{\pi} \tag{13.24}$$

wobei wir das Gaußsche Fehlerintegral (6.11) benutzt haben. (Wir werden das Gaußsche Fehlerintegral mit anderen Methoden in (13.26) und (13.31) berechnen.) Also

$$\kappa_1 = \frac{\pi^{\frac{1}{2}}}{\Gamma\left(\frac{1}{2}+1\right)} = \frac{\pi^{\frac{1}{2}}}{\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right)} = 2 = \omega_1, \quad \kappa_0 = \frac{1}{\Gamma(1)} = 1 = \omega_0$$

Es folgt, dass $\omega_n = \kappa_n$ für alle $n \in \mathbb{N}_0$.

13.9.6. **Satz** (Tonelli). *Seien $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ σ-endliche Maßräume und betrachte eine $\mathcal{A}\otimes\mathcal{B}$–meßbare Funktion $f : X \times Y \to [0,\infty]$. Dann sind die Funktionen*

$$X \ni x \longmapsto \int_Y f(x,y)\,d\nu(y), \quad Y \ni y \longmapsto \int_X f(x,y)\,d\mu(x)$$

messbar und gilt

$$\int_{X\times Y} f(x,y)\,d(\mu\otimes\nu) = \int_X \left(\int_Y f(x,y)\,d\nu(y)\right) d\mu(x) = \int_Y \left(\int_X f(x,y)\,d\mu(x)\right) d\nu(y).$$

13.9.7. **Satz** (Fubini). *Seien $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ σ-endliche Maßräume und betrachte eine $\mathcal{A}\otimes\mathcal{B}$–integrierbare Funktion $f : X \times Y \to \overline{\mathbb{R}}$. Dann gibt es Nullmengen $N \subset X$ und $M \subset Y$, so dass $f(x,\cdot) \in \mathcal{L}^1(Y, d\nu)$ für alle $x \in X \setminus N$ und $f(\cdot, y) \in \mathcal{L}^1(X, d\mu)$ für alle $y \in Y \setminus M$. Die Funktionen*

$$X\setminus N \ni x \longmapsto \int_Y f(x,y)\,d\nu(y), \quad Y\setminus M \ni y \longmapsto \int_X f(x,y)\,d\mu(x)$$

sind integrierbar auf $x \in X \setminus N$ bzw. $y \in Y \setminus M$ und gilt

$$\int_{X\times Y} f(x,y)\,d(\mu\otimes\nu) = \int_{X\setminus N} \left(\int_Y f(x,y)\,d\nu(y)\right) d\mu(x) = \int_{Y\setminus M} \left(\int_X f(x,y)\,d\mu(x)\right) d\nu(y).$$

Im Sinne der erweiterten Integraldefinition (13.12) können wir die Nullmengen M und N in der obigen Formel weglassen. Wir benutzen die beide Sätze oft in der folgenden Form:

13.9.8. **Folgerung** (Fubini-Tonelli). *Seien $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ σ-endliche Maßräume. Sei $f : X \times Y \to \overline{\mathbb{R}}$ eine $\mathcal{A}\otimes\mathcal{B}$–meßbare Funktion. Ist eines der Integrale*

$$\int_X \left(\int_Y |f(x,y)|\,d\nu(y)\right) d\mu(x), \quad \int_Y \left(\int_X |f(x,y)|\,d\mu(x)\right) d\nu(y)$$

endlich, so ist f $\mathcal{A}\otimes\mathcal{B}$–integrierbar und gilt

$$\int_{X\times Y} f(x,y)\,d(\mu\otimes\nu) = \int_X \left(\int_Y f(x,y)\,d\nu(y)\right) d\mu(x) = \int_Y \left(\int_X f(x,y)\,d\mu(x)\right) d\nu(y).$$

Wegen $(\mathbb{R}^{p+q}, \mathcal{B}(\mathbb{R}^{p+q}), \lambda_{p+q}) = (\mathbb{R}^{p+q}, \mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^q), \lambda_p\otimes\lambda_q)$ gelten diese Sätze zunächst für Borel-Lebesgue–meßbaren bzw. integrierbaren Funktionen. Weil $(\mathbb{R}^{p+q}, \mathcal{L}(\mathbb{R}^{p+q}), \lambda_{p+q})$ die Vervollständigung von $(\mathbb{R}^{p+q}, \mathcal{L}(\mathbb{R}^p)\otimes\mathcal{L}(\mathbb{R}^q), \lambda_p\otimes\lambda_q)$ ist, können wir die Sätze von Fubini und Tonelli auch für Lebesgue–meßbaren bzw. integrierbaren Funktionen ausdehnen. Dabei ersetzen wir $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ mit $(X, \mathcal{L}(X), \lambda_p)$, $(Y, \mathcal{L}(Y), \lambda_q)$, wobei $X \in \mathcal{L}(\mathbb{R}^p)$ sind, und $Y \in \mathcal{L}(\mathbb{R}^q)$ und $(X\times Y, \mathcal{A}\otimes\mathcal{B}, \mu\otimes\nu)$ mit $(X\times Y, \mathcal{L}(X\times Y), \lambda_{p+q})$. In diesem Sinne formulieren wir nur den Satz von Fubini-Tonelli:

13.9.9. **Folgerung** (Fubini-Tonelli für Lebesgue-integrierbare Funktionen). *Seien $X \in \mathcal{L}(\mathbb{R}^p)$ und $Y \in \mathcal{L}(\mathbb{R}^q)$ und sei $f : X\times Y \to \overline{\mathbb{R}}$ eine Lebesgue–meßbare Funktion. Ist eines der sukzessiven Integrale*

$$\int_X \left(\int_Y |f(x,y)|\,d\lambda_q(y)\right) d\lambda_p(x), \quad \int_Y \left(\int_X |f(x,y)|\,d\lambda_p(x)\right) d\lambda_q(y)$$

endlich, so ist f λ_{p+q}–integrierbar und gilt

$$\int_{X\times Y} f(x,y)\,d\lambda_{p+q} = \int_X \left(\int_Y f(x,y)\,d\lambda_q(y)\right) d\lambda_p(x) = \int_Y \left(\int_X f(x,y)\,d\lambda_p(x)\right) d\lambda_q(y). \tag{$*$}$$

Bemerkung. Für die Integrierbarkeit einer Funktion über einem Produktraum unterscheiden wir zwei Fälle:

(1) f ist nicht-negativ. Dann berchenen wir eines der iterierten Integrale

$$\int_X \left(\int_Y f(x,y)\, d\lambda_q(y) \right) d\lambda_p(x), \quad \int_Y \left(\int_X f(x,y)\, d\lambda_p(x) \right) d\lambda_q(y).$$

Ist eines endlich, so ist die Funktion integierbar und gilt $(*)$.

(2) f wechselt das Vorzeichen. Dann berechnen wir eines der iterierten Integrale des *Betrages* von f. Ist eines endlich, so ist die Funktion integierbar und gilt $(*)$. Es ist notwendig, die iterierten Integrale des Betrages von f zu betrachten, und nicht nur die von f (siehe Beispiel unten).

13.9.10. **Beispiele.**

(1) Sei $K = [a_1,b_1] \times [a_2,b_2] \times \cdots \times [a_n,b_n]$ mit $a_i \le b_i$ $(i = 1,\ldots,d)$. Es sei $f : K \to \mathbb{R}$ stetig. K ist kompakt, also gilt $K \in \mathcal{B}(\mathbb{R}^n)$. Nach (13.14) ist $f \in \mathscr{L}^1(K)$ und aus folgt

$$\int_K f(x_1,\ldots,x_n)\, d\lambda_n(x) = \int_{a_n}^{b_n} \left(\cdots \left(\int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1,\ldots,x_n)\, dx_1 \right) dx_2 \right) \cdots \right) dx_n$$

Die Reihenfolge der Integrationen darf beliebig vertauscht werden.

(2) Seien $a,b \in \mathbb{R}$ mit $a < b$ und $I := [a,b]$. Weiter seien $h_1, h_2 \in C(I)$ mit $h_1 \le h_2$ auf I und

$$A := \{(x,y) \in \mathbb{R}^2 : x \in I, h_1(x) \le y \le h_2(x)\}$$

Sei $f : A \to \mathbb{R}$ stetig. Da h_1 und h_2 stetig sind, ist A kompakt und somit gilt $A \in \mathcal{B}(\mathbb{R}^2)$. Aus Satz 13.7.5 folgt dann $f \in \mathcal{L}(A)$. Definiere

$$\tilde{f} : \mathbb{R}^2 \to \mathbb{R}, \quad \tilde{f}(x,y) = \begin{cases} f(x,y) & \text{, falls } (x,y) \in A \\ 0 & \text{, falls } (x,y) \notin A \end{cases}$$

Wir können der Satz 13.7.5 für $\tilde{f}$ anwenden und erhalten, dass $\tilde{f} \in \mathcal{L}(\mathbb{R}^2)$. Dann ist

$$\begin{aligned} \int_A f(x,y)\, d(x,y) &= \int_{\mathbb{R}^2} \tilde{f}(x,y)\, d(x,y) \\ &\overset{\text{Satz 13.9.7}}{=} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \tilde{f}(x,y)\, dy \right) dx \\ &= \int_a^b \left(\int_{h_1(x)}^{h_2(x)} f(x,y)\, dy \right) dx \end{aligned}$$

(3) Gaußsche Fehlerintegral (II). Betrachte die nicht-negative stetige (also Lebesgue-meßbare) Funktion

$$f : [0,\infty)^2 \to \mathbb{R}, \quad f(x,y) = y e^{-(1+x^2)y^2}.$$

Wir berechnen:

$$\int_0^\infty y e^{-(1+x^2)y^2} dy = -\frac{1}{2(1+x^2)} e^{-(1+x^2)y^2} \Big|_{y=0}^{\infty} = \frac{1}{2(1+x^2)}$$

(13.25)
$$\int_0^\infty \left(\int_0^\infty f(x,y) dy \right) dx = \frac{1}{2} \int_0^\infty \frac{1}{1+x^2} dx = \frac{1}{2} \arctan x \Big|_0^\infty = \frac{\pi}{4}$$

Weil $f \geqslant 0$ ist, können wir nach dem Satz 13.9.6 von Tonelli die Integrationsreihenfolge vertauschen:

$$\begin{aligned} \frac{\pi}{4} &= \int_0^\infty \left(\int_0^\infty f(x,y) dy \right) dx = \int_0^\infty \left(\int_0^\infty f(x,y) dx \right) dy \\ &= \int_0^\infty \left(\int_0^\infty y e^{-(1+x^2)y^2} dx \right) dy = \int_0^\infty \left(\int_0^\infty e^{-x^2y^2} dx \right) y e^{-y^2} dy \\ &= \int_0^\infty \left(\int_0^\infty e^{-t^2} dt \right) e^{-y^2} dy = \left(\int_0^\infty e^{-t^2} dt \right) \left(\int_0^\infty e^{-y^2} dy \right) \\ &= \left(\int_0^\infty e^{-y^2} dy \right)^2 \end{aligned}$$

(Substitution $xy = t$, $dx = \frac{dt}{y}$). Daraus folgt eine Auswertung des in der Wahrscheinlichkeitstheorie wichtigen und eindimensional nicht so leicht zu berechnenden ***Gaußschen Fehlerintegrals***:

(13.26)
$$\boxed{\int_0^\infty e^{-y^2} dy = \frac{\sqrt{\pi}}{2}}$$

Diese Methode stammt von Laplace (1778). Wir haben bereits in (6.11) das Integral mit einer komplizierteren Methode berechnet. Die schönste Herleitung werden wir mit Hilfe des Satzes von Tonelli und Transformationssatzes gewinnen (siehe (13.31)).

(4) Wir berechnen die Integrale

$$\int_0^1 \Big(\int_0^1 f(x,y)\,dx \Big)\,dy, \quad \int_0^1 \Big(\int_0^1 f(x,y)\,dy \Big)\,dx, \quad \int_{[0,1]^2} \big|f(x,y)\big|\,d\lambda_2$$

für die Funktion $f : [0,1]^2 \longrightarrow \mathbb{R}$:

$$f(x,y) = \begin{cases} \dfrac{x^2-y^2}{(x^2+y^2)^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0), \end{cases}$$

Für $y \neq 0$ gilt

$$\begin{aligned} \int_0^1 \frac{x^2-y^2}{\big(x^2+y^2\big)^2}\,dx &= 2\int_0^1 \frac{x^2}{\big(x^2+y^2\big)^2}\,dx - \int_0^1 \frac{dx}{x^2+y^2} \quad (x^2-y^2 = 2x^2-x^2-y^2) \\ &= \int_0^1 x\,\frac{2x\,dx}{\big(x^2+y^2\big)^2} - \int_0^1 \frac{dx}{x^2+y^2} \quad \text{(partielle Integration)} \\ &= -\frac{x}{x^2+y^2}\Big|_{x=0}^{x=1} + \int_0^1 \frac{dx}{x^2+y^2} - \int_0^1 \frac{dx}{x^2+y^2} \\ &= -\frac{x}{x^2+y^2}\Big|_{x=0}^{x=1} = -\frac{1}{1+y^2}. \end{aligned}$$

Deshalb gilt

$$\int_0^1 \Big(\int_0^1 f(x,y)\,dx \Big)\,dy = -2\int_0^1 \frac{dy}{1+y^2} = -2\arctan y\Big|_0^1 = -\arctan 1 = -\frac{\pi}{4}.$$

Ganz analog (vertausche x und y) ergibt sich

$$\int_0^1 \Big(\int_0^1 f(x,y)\,dy \Big)\,dx = \frac{\pi}{4}.$$

Sei $R_\varepsilon := \big\{(x,y) \in [0,1]^2 : \ \varepsilon \leqslant x^2+y^2 \leq 1\big\}$ für $\epsilon > 0$. Nach dem Satz von Beppo-Levi gilt:

$$\int_{[0,1]^2} \big|f(x,y)\big|\,d\lambda_2 = \lim_{\epsilon\to 0}\int_{R_0} \chi_{R_\epsilon}|f|\,d\lambda_2 = \lim_{\epsilon\to 0}\int_{R_\epsilon} |f|\,d\lambda_2. \tag{13.27}$$

Das Integral kann für jedes ε mit Hilfe von Polarkoordinaten berechnet werden:

$$\int_{R_\varepsilon} |f|\,d\lambda = \int_{R_\varepsilon} \frac{|\cos 2\varphi|}{r^2}\,r\,dr\,d\varphi = \int_0^{\pi/2} |\cos 2\varphi|\,d\varphi \cdot \int_\varepsilon^1 \frac{dr}{r} = 2\log r\Big|_\varepsilon^1 = -2\log\varepsilon,$$

also

$$\int_{[0,1]^2} \big|f(x,y)\big|\,d\lambda_2 = \infty.$$

Da f nicht intergrierbar ist, kann man den Satz von Fubini nicht anwenden, so dürfen die iterierten Integrale verschiedene Werte haben.

(5) Sei

$$f : [-1,1]^2 \longrightarrow \mathbb{R}, \quad f(x,y) = \begin{cases} \dfrac{xy}{(x^2+y^2)^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0) \end{cases}$$

Dann existieren die beiden iterierten Integrale von f und sind gleich, aber f ist auf $[-1,1]^2$ nicht integrierbar. In der Tat, für $x \in [-1,1]$ ist die Funktion $[-1,1] \ni y \mapsto f(x,y)$ stetig und ungerade, deshalb

$$\int_{-1}^1 f(x,y)dy = 0, \qquad \int_{-1}^1 \Big(\int_{-1}^1 f(x,y)dy \Big)dx = 0.$$

Die iterierten Intergale sind ebenfalls Null.

Nehmen wir an, dass f intergierbar auf $[-1,1]^2$ wäre. Dann müsste f auch auf $[0,1]^2$ intergierbar sein; nach dem Satz von Fubini ist auch die Funktion

$$g(x) := \int_0^1 f(x,y)dy$$

auf $[0,1]$ integrierbar. Wir berechnen jetzt g für $x \neq 0$:

$$\int_0^1 \frac{xy}{\left(x^2+y^2\right)^2}\,dy = \frac{x}{2}\int_0^1 \frac{d(x^2+y^2)}{\left(x^2+y^2\right)^2} = \frac{x}{2}\Big[-\frac{1}{x^2+y^2}\Big]\Big|_{y=0}^{y=1} = \frac{1}{2x} - \frac{x}{2x^2+2}.$$

Die Funktion $x/(2x^2+2)$ ist stetig auf $[0,1]$, und $1/(2x)$ ist nicht intergrierbar, d.h. g ist nicht intergierbar. Widerspruch.

13.10. **Der Transformationssatz.** Wir erinnern uns zunächst die Substitutionsregel in $\mathbb{R}$ (Satz 6.3.5). Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ stetig. Sei $[\alpha,\beta]$ ein weiteres Intervall, $\alpha < \beta$ und $\varphi \in \mathcal{C}^1([\alpha,\beta], I)$. Dann gilt

$$\int_\alpha^\beta f(\varphi(x))\varphi'(x)\,dx = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(y)\,dy\,.$$

Man beachte, dass diese Integrale eine Orientierung tragen, d.h. $\int_{\varphi(\alpha)}^{\varphi(\beta)} f(y)dy = -\int_{\varphi(\beta)}^{\varphi(\alpha)} f(y)dy$. Wir nehmen an, dass $\varphi'(x) \neq 0$ für alle $x \in (\alpha,\beta)$. Dann gibt es zwei Fälle:
Ist $\varphi'(x) > 0$ für alle $x \in (\alpha,\beta)$ d.h. φ ist monoton wachsend, so gilt

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(y)\,dy = \int_{\varphi([\alpha,\beta])} f\,d\lambda_1\,,$$

Ist $\varphi'(x) < 0$ für alle $x \in (\alpha,\beta)$ d.h. φ ist monoton fallend, so gilt

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(y)\,dy = -\int_{\varphi([\alpha,\beta])} f\,d\lambda_1\,.$$

Deshalb lautet die eindimensionale Substitutionsregel:

$$\int_{[\alpha,\beta]} (f\circ\varphi)\cdot|\varphi'|\,d\lambda_1 = \int_{\varphi([\alpha,\beta])} f\,d\lambda_1\,. \tag{13.28}$$

Das Ziel ist, die Transformationsformel der mehrdimensionalen Integralrechnung zu beweisen; sie lautet:

13.10.1. **Satz** (Transformationssatz). *Seien $U, V \subset \mathbb{R}^n$ offen und $\Phi : U \to V$ ein $\mathcal{C}^1$–Diffeomorphismus.*
(i) Ist $f : V \to [0,\infty)$ Lebesgue-meßbar, so ist $(f\circ\Phi)|\det J_\Phi| : U \to [0,\infty)$ Lebesgue-meßbar und gilt

$$\int_V f(y)\,d\lambda_n(y) = \int_U (f\circ\Phi)(x)\,|\det J_\Phi(x)|\,d\lambda_n(x)\,. \tag{13.29}$$

(ii) Eine messbare Funktion $f : V \to \mathbb{R}$ is genau dann integrierbar, wenn $(f\circ\Phi)|\det J_\Phi|$ integrierbar ist, und es gilt dann (13.29).

Beweis: Wir werden (i) beweisen; (ii) folgt aus (i), indem wir $f = f_+ - f_-$ zerlegen und die Aussage (i) für f_+, f_- anwenden. Den Beweis von (i) zerlegen wir in mehrere Teiletappen.
(1) Der Satz ist äquivalent zu einem Spezialfall.
Sei $A \subset U$ eine Lebesgue-meßbare Menge. Wir wissen nach dem Satz 13.3.6, dass $\Phi(A)$ Lebesgue-meßbar ist. Der Satz angewendet für $f = \mathbb{1}_A$ besagt:

$$\lambda_n(\Phi(A)) = \int_A |\det J_\Phi(x)|\,d\lambda_n(x)\,. \tag{13.30}$$

Umgekehrt liefert (13.30) die Transformationsformel (13.29):
(i) Ist f eine einfache Funktion, so folgt (13.29) aus (13.30) und die Linearität des Integrals.
(ii) Ist f eine meßbare nicht-negative Funktion, so wählen wir eine monotone Folge einfacher Funktionen $f_k \nearrow f$ und erhalten die Behauptung (13.29) aus (i) und dem Satz von Beppo-Levi.
Daher reicht es, (13.30) zu beweisen.
(2) Die Aussage ist lokal.
Weiterhin genügt es, folgende lokale Aussage zu beweisen: Jeder Punkt $p \in U$ hat eine offene Umgebung W so, dass (13.30) für die Transformation $\Phi|_W : W \longrightarrow \Phi(W)$ gilt. Ist dies gezeigt, so überdeckt man U durch abzählbar viele solche offenen Mengen W_j, $j \in \mathbb{N}$, z.B. durch Kugeln mit rationalem Durchmesser und Mittelpunkt. Dann zerlegt man A in disjunkte Teile $A = \cup_{k=1}^\infty A_k$ mit $A_k \subset W_{j(k)}$ und benutzt, dass beide Seiten von (13.30) σ–additiv sind.
(3) Die Formel (13.30) **gilt für** $n = 1$**.**
Dies ist für kompakte Intervalle der Fall (vgl. (13.28)). Daher übereinstimmen die Maße

$$\mu_1 : \mathcal{L}(U) \to [0,\infty], \quad \mu_1(A) = \lambda_1(\Phi(A)),$$

$$\mu_2 : \mathcal{L}(U) \to [0,\infty], \quad \mu_2(A) = \int_A |\Phi'(x)|\,d\lambda_1(x)\,.$$

auf allen kompakten Intervallen. Da Maße von unten stetig sind (Satz 13.1.8 (ii)), und $[a,b) = \cup_{k\in\mathbb{N}}[a,b+\frac{1}{n}]$ gilt, stimmen beide Maße auch auf halboffenen Intervallen überein. Außerdem sind μ_1 und μ_2 endlich auf kompakten Mengen und U läßt sich durch kompakten Mengen ausschöpfen. Folglich sind beide Maße σ-endlich. Nach dem Eindeutigkeitssatz 13.2.7 stimmen μ_1 und μ_2 auf $\mathcal{L}(U)$ überein.

(4) Die Formel (13.30) **gilt für eine Koordinatenpermutation.**

Vertauscht Φ die Koordinaten x_i und x_j, so ist Φ eine Spiegelung um die Hyperebene $x_i = x_j$ und $\det J_\Phi = \det\Phi = -1$. Die Formel (13.30) bedeutet also, dass $\lambda_n(A) = \lambda_n(\Phi(A))$. Dies beweisen wir auch Schrittweise. Die Aussage ist wahr für:

(1) Halboffenen Würfel A (offensichtlich).

(2) Offene Mengen A, denn eine offene Menge läßt sich als disjunkte abzählbare Vereinigung von halboffenen Würfel darstellen. Daraus folgt die Aussage auch für abgeschlossenen Mengen.

(3) Beliebeige Mengen $A \in \mathcal{L}(U)$. Das folgt aus (2) und Satz 13.3.2.

(5) Die Formel (13.30) **für Komposition zweier Diffeomorphismen.**

Gilt die Formel (13.30) für zwei Transformationen $\Phi : U \longrightarrow V$ und $\Psi : V \longrightarrow W$, so gilt sie auch für die Hintereinanderausführung $\Psi\circ\Phi : U \longrightarrow W$. Dies folgt aus der Produktregel für die Determinante

$$\det J_{\Psi\circ\Phi}(x) = \det J_\Psi(\Phi(x)) \cdot \det J_\Phi(x)$$

und der Äquivalenz von (13.30) und (13.29).

Aus **(4)** und **(5)** folgt, dass die Formel (13.30) insbesondere für eine beliebige Permutation von Variablen gilt. Wir können also oBdA Variablen im Bildbereich und im Urbildbereich vertauschen, ohne die Gültigkeit von (13.30) zu verletzen.

(6) Beweis der lokalen Aussage (2) durch Induktion über n.

Für $n = 1$ wurde die Behauptung bereits im Punkt **(4)** gezeigt. Wir setzen voraus, dass die Behauptung in Dimension $n-1$ gilt und schließen auf die Dimension n.

Sei $\Phi : U \longrightarrow \Phi(U) = V$ ein Diffeomorphismus und $p \in U$. Da $J_\Phi(p) \neq 0$, kann man nach Permutation von Koordinaten in U und V annehmen, dass $\frac{\partial\Phi_1}{\partial x_1}(p) \neq 0$. Wir zerlegen nun Φ lokal um p wie folgt in die Verknüpfung zweier Abbildungen:

Sei $\Psi(x_1,\dots,x_n) := (\Phi_1(x), x_2, \dots, x_n)$. Dann gilt

$$J_\Psi = \left(\begin{array}{c|ccc} \dfrac{\partial\Phi_1}{\partial x_1} & & * & \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{array}\right)$$

Also ist $\det J_\psi(p) \neq 0$. Folglich ist Ψ lokaler Diffeomorphismus um p (Umkehrsatz). Sei nun U so klein gewählt, dass $\Psi : U \longrightarrow W$ ein Diffeomorphismus von U auf eine offene Menge W ist.

Wir bezeichnen dann mit $\rho : W \longrightarrow V$ die Abbildung $\rho := \Phi\circ\Psi^{-1}$. Dann gilt offensichtlich $\rho(y_1, y_2, \dots, y_n) = (y_1, \rho_2(y), \dots, \rho_n(y))$. Wir haben also lokal um p den Diffeomorphismus Φ in die Verknüpfung zweier Diffeomorphismen Ψ und ρ zerlegt, die beide mindestens eine Koordinate festlassen:

$$\begin{tikzcd} U \arrow[rr, "\Phi"] \arrow[dr, "\Psi"'] & & V \\ & W \arrow[ur, "\rho"'] & \end{tikzcd}$$

Entsprechend Punkt **(5)** genügt es somit, die lokale Behauptung **(2)** für Abbildungen Φ zu beweisen, die die erste Koordinate festlassen. Sei also

$$\begin{aligned} &\Phi : (t,x) \in U \longrightarrow (t, \Phi_t(x)) \quad \text{mit} \\ &\Phi_t : U_t := \{x \in \mathbb{R}^{n-1} | (t,x) \in U\} \longrightarrow \mathbb{R}^{n-1}. \end{aligned}$$

Dann gilt

$$J_\Phi(t,x) = \left(\begin{array}{c|c} 1 & 0\dots 0 \\ * & J_{\Phi_t}(x) \end{array}\right)$$

d.h. $\det J_\Phi(t,x)) = \det J_{\Phi_t}(x)$. Wir wenden nun die Induktionsvoraussetzung, das Prinzip von Cavalieri und den Satz von Tonelli an, und erhalten

$$\begin{aligned}
\lambda_n(\Phi(A)) &\overset{\text{Cavalieri}}{=} \int_{\mathbb{R}} \lambda_{n-1}(\Phi(A)_t)\,d\lambda_1(t) = \int_{\mathbb{R}} \lambda_{n-1}(\Phi_t(A_t))\,d\lambda_1(t) \\
&= \int_{\mathbb{R}} \Big(\int_{A_t} |\det J_{\Phi_t}(x)|\,d\lambda_{n-1}(x) \Big)\,d\lambda_1(t) \\
&= \int_{\mathbb{R}} \Big(\int_{\mathbb{R}_{n-1}} \mathbb{1}_{A_t} |\det J_{\Phi_t}(x)|\,d\lambda_{n-1}(x) \Big)\,d\lambda_1(t) \\
&\overset{\text{Tonelli}}{=} \int_{\mathbb{R}^n} \mathbb{1}_A |\det J_\Phi|\,d\lambda_n \\
&= \int_A |\det J_\Phi|\,d\lambda_n\,.
\end{aligned}$$

□

13.10.2. **Folgerung** (Bewegungsinvarianz des Lebesgue-Maßes). *Sei $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ eine Euklidische Bewegung, d.h. $F(x) = L(x) + x_0$, wobei $L \in O(n)$, $x_0 \in \mathbb{R}^n$. Dann gilt: Ist $A \in \mathcal{L}(\mathbb{R}^n)$, so ist $F(A) \in \mathcal{L}(\mathbb{R}^n)$ und für die Maße gilt $\lambda_n(F(A)) = \lambda_n(A)$.*

Beispiele:

(1) Polarkoordinaten in der Ebene. Sei $U = (0,\infty) \times (0,2\pi)$, $V = \mathbb{R}^2 \setminus \{(x,0) : x \geqslant 0\}$ und sei $P : U \to V$, $(r,\theta) \to (r\cos\theta, r\sin\theta)$. Dann ist P ein Diffeomorphismus, und es gilt

$$J_P(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

Also $|\det J_P(r,\theta)| = r$. Da $\{(x,0) \mid x \geqslant 0\}$ eine Nullmenge ist, gilt für jede integrierbare Funktion $f : \mathbb{R}^2 \to \mathbb{C}$

$$\begin{aligned}
\int_{\mathbb{R}^2} f(x,y)\,d\lambda_2(x,y) &= \int_V f(x,y)\,d\lambda_2(x,y) \\
&\overset{(13.29)}{=} \int_U (f\circ P)(r,\theta)\, r\, d\lambda_2(r,\theta) \\
&\overset{\text{Fubini}}{=} \int_0^\infty \Big(\int_0^{2\pi} f(r,\theta)\, r\, d\theta \Big)\,dr
\end{aligned}$$

Gaußsche Fehlerintegral (III). Wir berechnen nochmal das Gaußsche Fehlerintegral (nach Gauß):

$$\begin{aligned}
\Big(\int_{\mathbb{R}} e^{-x^2}\,dx\Big)^2 &= \Big(\int_{\mathbb{R}} e^{-y^2}\,dy\Big)\cdot\Big(\int_{\mathbb{R}} e^{-x^2}\,dx\Big) = \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} e^{-(x^2+y^2)}\,dx \Big)\,dy \\
&\overset{\text{Tonelli}}{=} \int_{\mathbb{R}^2} e^{-(x^2+y^2)}\,d\lambda_2(x,y) = \int_0^\infty \Big(\int_0^{2\pi} e^{-r^2} r\,dt \Big)\,dr \\
&= 2\pi \int_0^\infty e^{-r^2} r\,dr = \lim_{R\to\infty} 2\pi \int_0^R e^{-r^2} r\,dr \qquad (u = r^2) \\
&= \lim_{R\to\infty} \pi \underbrace{\int_0^{R^2} e^{-u}\,du}_{\to 1} = \pi
\end{aligned}$$

also

$$\boxed{\int_{\mathbb{R}} e^{-x^2}\,dx = \sqrt{\pi}}. \tag{13.31}$$

(2) Zylinderkoordinaten in $\mathbb{R}^3$. Polarkoordinaten in der (x,y)–Ebene und z unverändert:

$$\begin{aligned}
&\psi : (0,\infty) \times (0,2\pi) \longrightarrow \mathbb{R}^3 \setminus \{(x,0,z) : x \geqslant 0, z \in \mathbb{R}\} \\
&\psi(r,\varphi,z) = (r\cos\varphi, r\sin\varphi, z) = (P_2(r,\varphi), z)\,.
\end{aligned}$$

Die Ebenen r = Konstante in (r,φ,z)–Raum werden in unendliche Zylinder in (x,y,z)–Raum abgebildet. Es ist leicht zu sehen, dass $\det J_\psi = r$.

(3) Polarkoordinaten in $\mathbb{R}^3$ oder Kugelkoordinaten. Wir haben in Beispiel 10.2.2 die Kugelkoordinaten definiert:

$$\begin{aligned}
&P_3 : \mathbb{R}_+ \times (0,2\pi) \times (-\tfrac{\pi}{2}, \tfrac{\pi}{2}) \longrightarrow \mathbb{R}^3 \\
&P_3(r,\varphi,\vartheta) = (P_2(r,\varphi)\cos\vartheta, r\sin\vartheta) = (r\cos\varphi\cos\vartheta, r\sin\varphi\cos\vartheta, r\sin\vartheta)\,.
\end{aligned}$$

Mit der dort eingeführten Notationen gilt $P_3 = \Psi_2 \circ \Psi_1$, also $J_{\Psi_2 \circ \Psi_1} = J_{\Psi_2}(\Psi_1) \cdot J_{\Psi_1}$. Aber $\det J_{\Psi_1}(r,\varphi,\vartheta) = r$ und $\det J_{\Psi_2}(\rho,\varphi,z) = \rho$ also

$$\det J_{P_3}(r,\varphi,\vartheta) = \det J_{\Psi_1}(r,\varphi,\vartheta) \cdot \det J_{\Psi_2}(r\cos\vartheta,\varphi,r\sin\vartheta) = r \cdot r\cos\vartheta = r^2\cos\vartheta\,.$$

Sei $I \subset [0,\infty)$ ein Intervall. Die *Kugelschale* K_I ist die Menge $K_I = \{x \in \mathbb{R}^n : \|x\|_2 \in I\}$. Z.B. $K_{[0,\infty)} = \mathbb{R}^n$, $K_{(0,\infty)} = \mathbb{R}^n \setminus \{0\}$, $K_{[0,R)} = B_R(0)$.

13.10.3. **Folgerung.** *Eine meßbare Funktion $f : K_I \longrightarrow \mathbb{C}$ ist über K_I genau dann integrierbar, wenn $(f \circ P_3)(r,\varphi,\vartheta)r^2\cos\vartheta$ auf $I \times (0,2\pi) \times (-\frac{\pi}{2},\frac{\pi}{2})$ integrierbar ist und dann gilt*

$$\int_{K_{a,b}} f\,d\lambda_n = \int_{I\times(0,2\pi)\times(-\frac{\pi}{2},\frac{\pi}{2})} f(r\cos\varphi\cos\vartheta, r\sin\varphi\cos\vartheta, r\sin\varphi)\,r^2\cos\vartheta\,d\lambda_3(r,\varphi,\vartheta)$$
$$= \int_I\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(r\cos\varphi\cos\vartheta, r\sin\varphi\cos\vartheta, r\sin\varphi)\,r^2\cos\vartheta\,d\vartheta\,d\varphi\,dr\,.$$

Beispiel: Das Volumen einer Kugelschale $K_I \subset \mathbb{R}^3$, mit $a = \inf I$, $b = \sup I$, $a,b \in \mathbb{R}$, ist

$$\lambda_3(K_I) = \int_a^b\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r^2\cos\vartheta\,d\vartheta d\varphi dr = \Big(\int_a^b r^2\,dr\Big)\Big(\int_0^{2\pi} d\varphi\Big)\Big(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos\vartheta\,d\vartheta\Big)$$
$$= \frac{b^3-a^3}{3}\cdot 2\pi\cdot(-\sin\vartheta)\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{4\pi}{3}(b^3-a^3).$$

(4) Polarkoordinaten in $\mathbb{R}^n$. Wir definieren die Polarkoordinatenabbildung in $\mathbb{R}^n$ durch Rekursion:

$$P_n : \mathbb{R}_+ \times (0,2\pi) \times (-\tfrac{\pi}{2},\tfrac{\pi}{2})^{n-2} \longrightarrow \mathbb{R}^n \setminus \{x \in \mathbb{R}^n : x_1 \geqslant 0, x_2 = 0\}$$
$$P_n(r,\varphi,\vartheta_1,\dots,\vartheta_{n-2}) = (P_{n-1}(r,\varphi,\vartheta_1,\dots,\vartheta_{n-3})\cos\vartheta_{n-2}, r\sin\vartheta_{n-2})$$

Es ergibt sich $P_n(x) = y$ wobei

$$\begin{aligned}
y_1 &= r\cos\varphi\cos\vartheta_1\cdots\cos\vartheta_{n-2}\,,\\
y_2 &= r\sin\varphi\cos\vartheta_1\cdots\cos\vartheta_{n-2}\,,\\
y_3 &= r\sin\vartheta_1\cos\vartheta_2\cdots\cos\vartheta_{n-2}\,,\\
&\vdots\\
y_{n-1} &= r\sin\vartheta_{n-3}\cos\vartheta_{n-2}\,,\\
y_n &= r\sin\vartheta_{n-2}\,.
\end{aligned}$$

In Beispiel 10.2.2 (6) haben wir gezeigt:
P_n ist ein Diffeomorphismus und

$$\det J_{P_n}(r,\varphi,\vartheta_1,\dots,\vartheta_n) = r^{n-1}\cos\vartheta_1(\cos\vartheta_2)^2\cdots(\cos\vartheta_{n-2})^{n-2}.$$

13.10.4. **Folgerung.** *Sei $f : \mathbb{R}^n \longrightarrow \mathbb{C}$, $f : K_{a,b} \longrightarrow \mathbb{C}$. f ist integrierbar genau dann, wenn $(f \circ P_n)|\det J_{P_n}|$ auf $\mathbb{R}_+ \times (0,2\pi) \times (-\frac{\pi}{2},\frac{\pi}{2})^{n-2}$ integrierbar ist und*

$$\int_{\mathbb{R}^n} = \int_0^\infty\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{-\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(P_n(r,\varphi))|\det J_{P_n}|\,d\vartheta_{n-2}\dots d\vartheta_1\,d\varphi\,dr\,.$$

Das Volumen der n-dimensionalen Kugel (II). Die Substitution $y = Rx$ zeigt sofort, dass $\omega_n(R) = R^n\omega_n$. Außerdem gilt

$$\begin{aligned}
\omega_n = \lambda_n(\overline{B}_1(0)) &= \int_{\mathbb{R}^n} \mathbb{1}_{\overline{B}_1(0)}\,d\lambda_n\\
&= \int_0^1\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{\frac{\pi}{2}}^{-\frac{\pi}{2}} r^{n-1}\cos\vartheta_1\cdots(\cos\vartheta_{n-2})^{n-2}\,d\vartheta_{n-2}\dots d\vartheta_1 dr\\
&= \Big(\int_0^1 r^{n-1}dr\Big)\Big(\int_0^{2\pi} d\varphi\Big)\Big(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos\vartheta_1\,d\vartheta_1\Big)\cdots\Big(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}(\cos\vartheta_{n-2})^{n-2}\,d\vartheta_{n-2}\Big)\\
&= \frac{1}{n}\cdot 2\pi\cdot I_1\cdots I_{n-2}\,,
\end{aligned}$$

wobei $I_n := \int_0^{\frac{\pi}{2}} \cos^n x\,dx$. Wir wissen nach (13.21), dass $I_{k-1}I_k = \frac{2\pi}{k}$, $k \geqslant 1$ und daraus erhalten wir wie im Beispiel 13.9.5:

$$\omega_n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}. \tag{13.32}$$

Sei $I \subset [0,\infty)$ ein Intervall und die *Kugelschale* $K_I = \{x \in \mathbb{R}^n : \|x\|_2 \in I\}$. Eine Funktion $f : K_I \longrightarrow \mathbb{R}$ heißt ***rotationssymmetrisch***, wenn es $g : I \longrightarrow \mathbb{R}$ existiert, mit $f(x) = g(\|x\|_2)$.

13.10.5. **Folgerung.** *Sei $g : I \longrightarrow \mathbb{R}$ eine meßbare Funktion und sei $f : K_I \longrightarrow \mathbb{R}$ die rotationssymmetrische Funktion definiert durch $f(x) = g(\|x\|_2)$.*
(i) Ist $g \geqslant 0$, so gilt

$$\int_{K_I} f\,d\lambda_n = n\,\omega_n \int_I g(r)r^{n-1}\,d\lambda_1(r). \tag{13.33}$$

(ii) f ist genau dann integrierbar, wenn $g(r)r^{n-1}$ integrierbar auf I ist, und dann gilt (13.33)*.*

Beweis:

$$\begin{aligned}\int_{K_I} f\,d\lambda_n &= \int_I \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \ldots \int_{-\pi/2}^{\pi/2} g(r)r^{n-1}\,d\varphi d\vartheta_1 \ldots d\vartheta_{n-2}\\ &= 2\pi I_1 \cdots I_{n-2} \int_I g(r)r^{n-1}\,dr = n\,\omega_n \int_I g(r)r^{n-1}\,d\lambda_1(r).\end{aligned}$$

□

Das Volumen der n-dimensionalen Kugel (III). Sei $f : \mathbb{R}^n \longrightarrow \mathbb{R}_+$, $f(x) = e^{-\|x\|^2}$. Dann gilt:

$$\begin{aligned}\int_{\mathbb{R}^n} e^{-\|x\|^2}\,dx &= n\,\omega_n \int_0^\infty e^{-r^2} r^{n-1}\,dr \quad \text{(nach (13.33))}\\ &= n\,\omega_n \int_0^\infty e^{-s} s^{\frac{n-1}{2}} \frac{1}{2s^{1/2}}\,ds \quad \text{(Substitution: } r^2 = s\,,\ 2rdr = ds))\\ &= \tfrac{n}{2}\,\omega_n \int_0^\infty e^{-s} s^{\frac{n}{2}-1}\,ds = \tfrac{n}{2}\Gamma(\tfrac{n}{2})\\ &= \omega_n \Gamma(\tfrac{n}{2}+1).\end{aligned}$$

Außerdem gilt

$$\int_{\mathbb{R}^n} e^{-\|x\|^2}\,dx \overset{\text{Tonelli}}{=} \Big(\int_{\mathbb{R}} e^{-x^2}\,dx\Big)^n = \pi^{n/2}.$$

also

$$\boxed{\omega_n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}}. \tag{13.34}$$

Dies ist die wohl eleganteste Herleitung der Formel für ω_n. Man kann
(1) Das Gaußsche Fehlerintegral wie in (13.31) berechnen.
(2) $\Gamma(\frac{1}{2})$ wie in (13.24) berechnen.
(3) (13.32) wie oben herleiten.
(4) $\Gamma(\frac{n}{2}+1)$ mit Hilfe der Funktionalgleichung der Gamma-Funktion berechnen und schließlich (13.22) erhalten.

13.11. **Die L^p-Räume.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum.

13.11.1. **Bemerkung.** Ist $g : \Omega \to [0,\infty]$ meßbar, so auch g^p für alle $p \in \mathbb{R}_+$ (mit $\infty^p := \infty$). In der Tat, ist (s_n) eine Folge einfacher Funktionen mit $s_n \nearrow g$ so sind s_n^p einfacher Funktionen und $s_n^p \nearrow g^p$.

13.11.2. **Definition.** Sei $f : \Omega \to \overline{\mathbb{R}}$ meßbar. Für $p \in \mathbb{R}_+$ setze

$$\|f\|_p := \Big(\int_\Omega |f|^p\,d\mu\Big)^{1/p} \in [0,\infty].$$

Für $p = \infty$ setze:

$$\|f\|_\infty := \operatorname{ess\,sup}|f| := \inf\Big\{\alpha \in [0,\infty] : |f| \leqslant \alpha \quad \mu-\text{f.ü.}\Big\} \in [0,\infty].$$

13.11.3. **Definition.**

(i) Für $0 < p \leqslant \infty$ sei

$$\mathscr{L}^p(\Omega, A, \mu) := \mathscr{L}^p(\mu) := \left\{ f : \Omega \to \mathbb{R} | f \text{ meßbar}, \|f\|_p < \infty \right\}.$$

Insbesondere

$$\mathscr{L}^1(\mu) = \{ f : \Omega \to \mathbb{R} : f \ \mu\text{-integrierbar} \}.$$

(ii) Sei $\mathscr{N}(\mu) := \left\{ f : \Omega \to \mathbb{R} : f \text{ meßbar}, f = 0 \ \mu\text{-f.ü.} \right\} \subset \mathscr{L}^p(\mu)$.

(iii) Betrachte auf $\mathscr{L}^p(\mu)$ die Äquivalenzrelation $f \sim g \Leftrightarrow f - g \in \mathscr{N}(\mu)$. Setze

$$L^p(\Omega, A, \mu) := L^p(\mu) := \mathscr{L}^p(\mu)/_\sim .$$

(iv) Für $f \in L^p(\mu)$ setze $\|f\|_p := \|\tilde{f}\|_p$, wo $\tilde{f}$ ein Repräsentant von f in $\mathscr{L}^p(\mu)$ ist. (Dann ist $\|f\|_p$ offenbar wohldefiniert!).

Wir wollen zeigen, dass für $1 \leqslant p \leqslant \infty$ gilt: $\mathscr{L}^p(\mu)$, $L^p(\mu)$ sind Vektorräume, $\|\cdot\|_p$ ist eine Norm auf $L^p(\mu)$, und $L^p(\mu)$ mit der zugehörigen Metrik ist vollständig. Dafür erst:

13.11.4. **Satz** (Höldersche Ungleichung). *Seien $1 \leqslant p, q \leqslant \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, $f, g : \Omega \to \overline{\mathbb{R}}$ meßbar. Dann gilt*

$$\|fg\|_1 \leqslant \|f\|_p \cdot \|g\|_q .$$

Beweis: Falls $p = \infty$: $|fg| = |f||g| \leqslant \|f\| \, \|g\|_\infty$ μ-f.ü., also $\int |fg| d\mu \leqslant \int |f| d\mu \cdot \|g\|_\infty$ (analog falls $q = \infty$).

Seien $1 < p$, $q < \infty$. Falls $\|f\|_p = 0$, so $|f|^p = 0$ μ-f.ü., also $f = 0$ μ-f.ü., also $fg = 0$ μ-f.ü. und die Behauptung folgt.

Seien $\|f\|_p, \|g\|_q > 0$. Falls $\|f\|_p = \infty$, so ist die Behauptung trivial.

Seien $0 < \|f\|_p, \|g\|_q < \infty$. Nach der Youngsche Ungleichung 8.2 gilt $\forall a, b \in [0, \infty]$: $a^{1/p} \cdot b^{1/q} \leqslant \frac{1}{p} a + \frac{1}{q} b$, also $\forall x \in \Omega$:

$$\frac{|f(x)|}{\|f\|_p} \cdot \frac{|g(x)|}{\|g\|_q} \leqslant \frac{1}{p} \frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q} \frac{|g(x)|^q}{\|g\|_q^q} \overset{\int_\Omega \dots d\mu}{\Longrightarrow}$$

$$\|fg\|_1 \leqslant \|f\|_p \cdot \|g\|_q \cdot \left(\frac{1}{p} \frac{\|f\|_p^p}{\|f\|_p^p} + \frac{1}{q} \frac{\|g\|_q^q}{\|g\|_q^q} \right) = \|f\|_p \cdot \|g\|_q . \quad \square$$

13.11.5. **Folgerung** (Cauchy-Schwarzsche Ungleichung). *Seien $f, g : \Omega \to \overline{\mathbb{R}}$ meßbar. Dann gilt*

$$\|fg\|_1 \leqslant \|f\|_2 \cdot \|g\|_2 ,$$

d.h.

$$\int |fg| d\mu \leqslant \left(\int |f|^2 d\mu \right)^{1/2} \cdot \left(\int |g|^2 d\mu \right)^{1/2} .$$

13.11.6. **Satz** (Minkowskische Ungleichung). *Seien $1 \leqslant p \leqslant \infty$, $f, g : \Omega \to \overline{\mathbb{R}}$ meßbar. Dann gilt*

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p .$$

Beweis:
- Falls $p = 1$: Klar wegen $|f + g| \leqslant |f| + |g|$.
- Falls $p = \infty$: zz $\operatorname{ess\,sup} |f + g| \leqslant \operatorname{ess\,sup} |f| + \operatorname{ess\,sup} |g|$:
  Linke Seite ist $\leqslant \alpha + \beta \forall \alpha \geq \operatorname{ess\,sup} |f|$, $\beta \geq \operatorname{ess\,sup} |g|$:
  Klar, da aus $|f| \leqslant \alpha \mu-$f.ü., $|g| \leqslant \beta \mu-$f.ü. folgt: $|f + g| \leqslant \alpha + \beta \mu-$f.ü..
- Sei $1 < p < \infty$. Sei $q := \frac{p}{p-1}$, also $\frac{1}{p} + \frac{1}{q} = 1$.
  Behauptung Inv., falls $\|f + g\|_p = 0$ oder $\|f\|_p = \infty$ oder $\|g\|_p = \infty$.
  Gelte also keine dieser Aussagen. Dann auch $\|f + g\|_p < \infty$, denn $|f + g|^p$ und:

$$\int_\Omega \underbrace{|f + g|^p d\mu}_{= \|f+g\|_p^p} \leqslant \int_\Omega |f||f + g|^{p-1} d\mu + \int_\Omega |g||f + g|^{p-1} d\mu$$

$$\overset{\text{Hölder}}{\leqslant} \left(\|f\|_p + \|g\|_p \right) \cdot \underbrace{\left\| |f + g^{p-1}| \right\|}_{= \|f+g\|_p^{\frac{p}{q}}} \cdot \|f + g\|_p^{\frac{-p}{q}}$$

$$\Rightarrow \|f + g\|_p^{p - \frac{p}{q} = 1} \leqslant \|f\|_p + \|g\|_p .$$

<u>Korollar:</u> $1 \leqslant p \leqslant \infty \Rightarrow$

- (i) $\mathscr{L}^P(\mu)$ ist ein $\mathbb{R}-VR$

$$\ulcorner$$

$$f \in \mathscr{L}^P(\mu), \alpha \in \mathbb{R} \Rightarrow \|\alpha f\|_p = |\alpha|\,\|f\|_p < \infty \quad \surd$$
$$f,g \in \mathscr{L}^P(\mu), \alpha \in \mathbb{R} \Rightarrow \|f+g\|_p \leqslant \|f\|_p + \|g\|_p < \infty \quad \surd.$$

$$\llcorner$$

- (ii) $L^P(\mu)$ ist ein $\mathbb{R}-VR$ (klar, da $\mathscr{N}(\mu)$ Unterraum, also $L^P(\mu)$ Quart.-Raum).
- (iii) $\|\cdot\|_p$ ist eine Norm auf $L^P(\mu)$.

$\ulcorner$

$\|f\|_p \geq 0$ klar, $\|f\|_p = 0 \quad \Rightarrow \int_\Omega |\widetilde{f}|^p d\mu = 0 \quad \Rightarrow |\widetilde{f}|^p = 0\mu-$f.ü. $\quad \Rightarrow \widetilde{f} = 0\mu-$f.ü. $\quad \Rightarrow f = 0 \quad \surd$.
$\|\alpha f\|_p = |\alpha|\,\|f\|_p$ klar (da $\alpha f :=$ Klasse von αf, wo $\widetilde{f}$ Rep. für f).

$$\|f+g\|_p = \underbrace{\|\widetilde{f}+\widetilde{g}\|_p}_{\text{Rep. für} f+g (f+g \text{ so def.})} \overset{\text{MINK}}{\leqslant} \|\widetilde{f}\|_p + \|\widetilde{g}\|_p = \|f\|_p + \|g\|_p.$$

$\llcorner$

- (iv) Mit $d_p(f,g) := \|f-g\|_p$ ist $L^P(\mu)$ ein metrischer Raum. $\quad (\surd)$.

13.11.7. **Bemerkung.** Hölder, Cauchy-Schwarz, Minkowski gelten auch auf $\left\{\text{meßbare Funktionen: } \Omega \to \mathbb{R}\right\} \big/ \mathscr{N}(\mu)$.

13.11.8. **Bemerkung.** Im folgenden betrachten wir nur die Funktionen-Klassen in $L^P(\mu)$, also “Funktionen“ auf Ω, d.h. wir meinen stets Repräsentanten (eindeutig bestimmt bis auf Unterschiede auf μ-Nullmengen).
Ist ungefährlich, da VR-Str. und $\|\cdot\|_p$ durch Repräsentanten definiert (und wohldefiniert).

13.11.9. **Satz** (Satz von Fischer-Riesz)**.** $1 \leqslant p \leqslant \infty \Rightarrow$
$\left(L^P(\mu), d_p\right)$ ist ein vollständig metrischer Raum, d.h.
$\left(L^P(\mu), \|\cdot\|_p\right)$ ist ein Banach-Raum.

Beweis zu Fischer-Riesz: Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\left(L^P(\mu), d_p\right)$.
Zu zeigen: $\exists f \in L^P(\mu)$ mit $\|f_n - f\|_p \to 0$ für $n \to \infty$.

- Falls $1 \leqslant p \leqslant \infty$:
  Wähle Teilfolge $(f_{n_k})_{n\in\mathbb{N}}$ mit $\|f_m - f_{n_k}\|_p < \frac{1}{2^k} \forall k \in \mathbb{N} \forall m \geq n_k$ (geht, da CF).
  Schreibe $g_k := f_{n_k} - f_{n_{k+n}}$. Dann $\forall n \in \mathbb{N}$:

$$\left\|\left(\sum_{k=1}^{n} |g_1|\right)\right\|_p \overset{\text{MINK}}{\leqslant} \sum_{k=1}^{n} \|g_n\|_p \leqslant \sum_{k=1}^{n} \frac{1}{2^k} < 1.$$

  Satz über non. Konv. $\left(\text{angewiesen auf Repräsentanten der } g_n \text{ und } \left(\sum_{k=1}^{n} |g_k|\right)^p \to \left(\sum_{k=1}^{\infty} |g_u|\right)^p\right)$:
  $\left|\left|\sum_{k=1}^{\infty} |g_u|\right|\right|_p \leqslant 1$, insbesondere gilt $\mu-$f.ü. $(\sum_{k=1}^{\infty} |g_k|)^p < \infty$, also $\sum_{k=1}^{\infty} |g_u| < \infty \mu-$f.ü..
  $\Rightarrow \sum_{k=1}^{\infty} g_k$ konv. $\mu-$f.ü., d.h. $\left(f_{n_1} - f_{n_k}\right)_{k\in\mathbb{N}}$ konv. $\mu-$f.ü., d.h. $\left(f_{n_k}\right)_{k\in\mathbb{N}}$ konv. $\mu-$f.ü..
  Sei $f : \Omega \to \mathbb{R}$ meßbar mit $f_{n_k} \to f\,\mu-$f.ü..
  <u>Behauptung:</u>
  $\|f_1 - f\|_p \to 0$ und (da insbesondere $\|f\|_p \leqslant \|f_n\|_p + \|f_n - f\|_p < \infty$):$f \in L^P$.
  <u>Beweis:</u>
  Sei $\varepsilon > 0$. Sei N so, daß $\|f_\ell - f_m\|_p < \varepsilon \forall \ell, m \geq N$.
  Dann $\forall m \geq N$:

$$\int_\Omega |f - f_m|^p d\mu = \int_\Omega \lim_{k\to\infty} |f_{n_k} - f|^p d\mu \overset{\text{FATOU}}{\leqslant} \liminf_{k\to\infty} \underbrace{\int_\Omega |f_{n_k} - f|^p d\mu}_{<\varepsilon^p \forall k \text{ mit } n_k \geq N} < \varepsilon^p.$$

  Also $\|f - f_m\|_p < \varepsilon^p$.
- Falls $p = \infty$: Nach Definition von $\|\cdot\|_\infty$ ist

$$C := \bigcup_{n\in\mathbb{N}} \{x \in \Omega | f_n(x)| > \|f_n\|_\infty\} \cup \bigcup_{m,n\in\mathbb{N}} \{x \in \Omega | f_m(x) - f_n(x)| > \|f_m - f_n\|_\infty\}$$

eine μ-Nullmenge. Außerhalb von C gilt: $f_n(x) \in \mathbb{R}$ und $|f_m(x) - f_n(x)| \leqslant \|f_m - f_n\|_\infty$, also $\Big(f_n(x)\Big)$ CF, also konv.
$\Rightarrow ?f : \Omega \to \mathbb{R}$ meßbar: $f_n \to f\,\mu-$f.ü., d.h. $\to f$ außerhalb von C.
Behauptung:
$\|f_n - f\|_\infty \to 0$ $\Big($ insbesondere $\underbrace{\|f\|_\infty}_{=\|f_n\|+f_n=\|f\|_\infty} < \infty$, d.h. $f \in L^\infty(\mu)\Big)$.
Beweis:
$\varepsilon > 0$ vorgegeben, N so, daß $\|f_n - f_m\|_\infty < \varepsilon \forall n,m \geq N$.
Dann $\forall n \geq N \forall x \notin C$:

$$|f_n(x) - f(x)| = \lim_m |f_n(x) - f_m(x)| \leqslant \lim_m \|f_n - f_m\|_\infty \leqslant \varepsilon.$$

Also $\|f_n - f\|_\infty \leqslant \varepsilon \forall n \ni N$.
$\llcorner$

13.11.10. **Bemerkung.** Für $f,g \in L^2(\mu)$ setze $\langle f,g\rangle := \int_\Omega f g d\mu$ ($|\ldots| < \infty$ wegen Hölder). Offenbar f,g $\mathbb{R}$-bilinear, symm., $\langle f,f\rangle \geq 0$, " = " $\Leftrightarrow f = 0$, also $\langle,\rangle$ ein SKP. Dann zugehörige Norm $\sqrt{\langle,\rangle_{L^2}} = \|\cdot\|_2$.

Korollar: $(L^2(\mu), \langle,\rangle)$ ist ein Hilbertraum, d.h. ein VR mit einem SKP, so daß der VR mit der zugehörigen Norm ein Banach-Raum ist.

§2.10 Konvergenz dem Maß nach

Sei (Ω, A, μ) Maßraum.

13.11.11. **Definition.** Seien $f, f_1, f_2, \ldots$ meßbare Funktionen: $\Omega \to \mathbb{R}$.
(f_n) konvergiert dem Maß nach oder μ-stochastisch gegen f, Bez.: $f_n \xrightarrow{\mu} f$, wenn gilt: $\forall \alpha > 0 \forall A \in \mathscr{A}$ mit $\mu(A) < \infty$:

$$\boxed{\mu\Big(\{|f_n - f| > \alpha\} \cap A\Big) \to 0 \text{ für } n \to \infty.}$$

(überfl. falls $\mu(\Omega) < \infty$).

13.11.12. **Satz.** *Sei μ σ-endlich, $f_n \xrightarrow{\mu} f$, $f_n \xrightarrow{\mu} g$. Dann $f = g\mu -$f.ü..*

Beweis: Ist $A \in \mathscr{A}$ mit $\mu(A) < \infty$, so

$$\mu\Big(\{f \neq g\} \cap A\Big) \leqslant \mu\Bigg(\bigcup_{k\in\mathbb{N}} \{|f - g| \geq \frac{1}{k}\} \cap A\Bigg) \leqslant \sum_{k=1}^{\infty} \underbrace{\mu\Big(\{|f - g| \geq \frac{1}{k}\} \cap A\Big)}_{\leqslant \mu(\{|f_n - f| \geq \frac{1}{2k}\} \cap A) + \mu(\{|f_n - g| \geq \frac{1}{2k}\} \cap A)} = \sum_{k=1}^{\infty} 0 = 0.$$

Wende dies an auf Folge A_1, A_2 mit $\bigcup, A_k = SL \Rightarrow \mu\{|f \neq g|\} = 0$.

13.11.13. **Satz.** *Falls $f_n \to f\,\mu -$f.ü. (reicht: $f_{n|A} \to f_{|A}\,\mu -$f.ü.$\forall A \in \mathscr{A}$ mit $\mu(A) < \infty$), so auch $f_n \xrightarrow{\mu} f$.*

Beweis: Sei $\mu(A) < \infty$ und $\alpha > 0$. Weiß:

$$\begin{aligned} 0 = \mu\Big(\{\lim_{n\to\infty} \sup |f_n - f| \geq \alpha\} \cap A\Big) &= \mu\Bigg(\Bigg\{\inf_{n\in\mathbb{N}} \Big(\sup_{m\geq n} |f_m - f|\Big) \geq \alpha\Bigg\} \cap A\Bigg) \\ = \mu\Big(\bigcap \{\sup |f_m - f| \geq \alpha\} \cap A\Big) &= \lim_{n\to\infty} \mu\Big(\{\sup |f_m - f| \geq \alpha\} \cap A\Big) \geq \lim_{n\to\infty} \mu\Big(\{|f_m - f| \geq \alpha\}\Big). \end{aligned}$$

13.11.14. **Satz.** $f_n \xrightarrow{\mu} f \Leftrightarrow \forall A \in \mathscr{A}$ *mit $\mu(A)$ endlich: Jede Teilfolge von (f_n) enthält eine Teilfolge, die auf A $\mu -$f.ü. gegen $f_{|A}$ konvergiert.*

Beweis: " $\Rightarrow$ ": Sei Teilfolge gegeben, ebenfalls $\xrightarrow{\mu} f \Rightarrow$ sei oBdA die ganze Folge. Zu $k \in \mathbb{N}$ wähle $n_k \in \mathbb{N}$ mit $\mu\Big(\{|f_m - f_{n_k}| \geq \frac{1}{k^2}\} \cap A\Big) \leqslant \frac{1}{k^2} \forall m \geq n_k$. (Geht, da $\ldots \leqslant \mu\Big(\{|f_m - f| \geq \frac{1}{2k^2}\} \cap A\Big) + \mu\Big(\{|f_{n_k} - f| \geq \frac{1}{2k^2}\} \cap A\Big)$ und Vorauss.). Wähle (n_k) außerdem strikt mon. steigend. Setze

$$C := \bigcap_{\ell\in\mathbb{N}} \bigcup_{k\geq\ell} \underbrace{\Bigg(\Bigg\{|f_{n_{k+1}} - f_{n_k}| \geq \frac{1}{k^2}\Bigg\} \cap A\Bigg)}_{\mu(\ldots) \leqslant \frac{1}{k^2}} \Rightarrow \mu(C) = \lim_{\ell\to\infty} \underbrace{\mu\Bigg(\bigcup_{k\geq\ell}\Bigg)}_{\leqslant \sum_{k\geq\ell} \frac{1}{k^2}} = 0.$$

Für jedes $w \in A \setminus C$ ist $|f_{n_{k+1}}(w) - f_{n_k}(w)| < \frac{1}{k^2}$ für fast alle k, also $\sum_k \Big(f_{n_{k+1}}(w) - f_{n_k}(w)\Big)$ (abs.) konvergent, also $\Big(f_{n_k}(w)\Big)$ konvergiert gegen $=: f^*(w)$. f meßbare Funktion auf $A \setminus C$, nach Forts. durch 0 oder auf A.

Nun $f_{n_{k|A}} \to f^* \mu-$f.ü. $\overset{\text{Satz 2}}{\Rightarrow} f_{n_{k|A}} \overset{\mu}{\to} f^* \underset{\mu(A)<\infty}{\overset{\text{Satz 1}}{\Rightarrow}} f_{|A} = f^y \mu-$f.ü. $\Rightarrow f_{n_{k|A}} \to f_{|A} \mu-$f.ü..

“ $\Leftarrow$ “: Sei $\mu(A) < \infty$ und $\alpha > 0$. Sei $\left(f_{n_k}\right)$ eine Teilfolge von (f_n).

Voraussetzung $\Rightarrow \left(f_{n_k}\right)$ hat Teilfolge, die $\mu-$f.ü., also auch μ-stoch., auf A gegen $f_{|A}$ kovergiert.
Also: $\mu\left(\{|f_{n_k} - f| \geq \alpha\} \cap A\right)$ hat Teilfolge $\to 0$.
Dies $\forall_{TF} \left(f_{n_k}\right) \Rightarrow \mu(\{|f_n - f| \geq \alpha\} \cap A) \to 0$.

Korollar: $f_n \overset{\mu}{\to} f$, $\gamma : \mathbb{R} \to \mathbb{R}$ stetig $\Rightarrow \gamma \circ f_n \mu \!\to\! \gamma \circ f$. [Klar mit Satz 3!]

13.11.15. **Satz.** *Sei* $1 \leqslant p < \infty$, $\|f_n - f\|_p \to 0 \Rightarrow f_n \overset{\mu}{\to} f$.

Beweis:

$$\mu(A) < \infty, \alpha > \sigma \Rightarrow \mu(\{|f_n - f| \geq \alpha\} \cap A) \leqslant \mu(|f_n - f| \geq \alpha) \overset{\text{Lemma s.u.}}{\leqslant} \frac{1}{\alpha^p} \int_\Omega |f_n - f|^p \, d\mu \to 0 \text{ für } n \to \infty.$$

Auch für $p = 0$ mit $\|f_{n_{|A}} - f_{|A}\|_1$ nach ÜA 22.

13.11.16. **Lemma** (Tschebyscheff-Markov-Ungleichung)**.** $f : \Omega \to \overline{\mathbb{R}}$ *meßbar* $(p = 2)$*,* $\alpha > 0$*,* $1 \leqslant p < \infty$

$$\Rightarrow \boxed{\mu(\{|f| \geq \alpha\}) \leqslant \frac{1}{\alpha^p} \int_\Omega |f|^p \, d\mu}.$$

Beweis:

$$\int_\Omega |f|^p \, d\mu \geq \int_{\{|f| \geq \alpha\}} |f|^p \, d\mu \geq \int_{\{|f| \geq \alpha\}} \alpha^p \, d\mu = \alpha^p \cdot \mu\left(\{|f| \geq \alpha\}\right).$$

14. Integration auf Untermannigfaltigkeiten

14.1. **Das Integral einer Differentialform.** In §12.6 haben wir das Integral einer 1-Form längs einer Kurve definiert. Analog führen wir das Integral einer n-Form auf einer n-dimensionalen Untermannigfaltigkeit.

14.1.1. **Definition.** Sei $A \in \mathcal{L}(\mathbb{R}^n)$. Für jede n–Form $\omega : A \to \Lambda^n(\mathbb{R}^n)^*$ gibt es eine eindeutig bestimmte $f : A \to \mathbb{R}$ mit $\omega = f\,dx_1 \wedge \ldots \wedge dx_n$. Die Form ω heißt *nicht-negativ*, geschrieben $\omega \geqslant 0$, falls $f \geqslant 0$. Die Form ω heißt *meßbar* (bzw. *integrierbar*), falls f meßbar (bzw. integrierbar) ist. Nehmen wir an, dass ω meßbar und $\omega \geqslant 0$ oder ω integrierbar ist, so definieren wir

$$\int_A \omega := \int_A f\,d\lambda_n\,.$$

14.1.2. **Satz** (Transformationssatz für n-Formen). *Seien $U,V \subset \mathbb{R}^n$ offen, $\Phi : U \longrightarrow V$ ein $\mathcal{C}^\infty$–Diffeomorphismus mit konstanten Vorzeichen $\varepsilon(\Phi)$. Sei ω eine meßbare n-Form auf V.*

(i) *Ist ω nicht-negativ, so ist $\varepsilon(\Phi)\Phi^*\omega$ nicht-negativ und meßbar und gilt*

$$\int_V \omega = \varepsilon(\Phi)\int_U \Phi^*\omega\,. \tag{14.1}$$

(ii) *ω is genau dann integrierbar, wenn $\Phi^*\omega$ integrierbar ist, und es gilt dann* (14.1).

Beweis: Sei $\omega = f\,dy_1 \wedge \ldots \wedge dy_n$. Dann gilt $\Phi^*\omega = (f\circ\Phi)\det J_\Phi\,dx_1 \wedge \ldots \wedge dx_n$ also

$$\begin{aligned}\int_V \omega &:= \int_V f\,d\lambda_n \overset{(13.29)}{=} \int_U (f\circ\Phi)|\det\Phi|\,d\lambda_n = \varepsilon(\Phi)\int_U (f\circ\Phi)\det J_\Phi\,d\lambda_n\\ &= \varepsilon(\Phi)\int_U \Phi^*\omega.\end{aligned}$$

□

14.1.3. **Definition.** Sei M eine n–dimensionale Untermannigfaltigkeit. Eine Teilmenge $A \subset M$ heißt Lebesgue–meßbar (bzw. Nullmenge), wenn für jede Karte (U,φ) von M die Menge $\varphi(A\cap U) \subset \mathbb{R}^n$ Lebesgue–meßbar (bzw. Nullmenge) ist.

Diese Begriffe sind wohldefiniert, da Lebesgue–meßbare Mengen und Nullmengen in $\mathbb{R}^n$ unter $\mathcal{C}^1$–Abbildungen in ebensolche übergehen und die Kartenübergänge zwischen zwei Karten der Klasse $\mathcal{C}^1$ sind.

14.1.4. **Satz.** *Sei $A \subset M$. Die folgenden Behauptungen sind äquivalent:*

(1) *A ist Lebesgue–meßbar (bzw. Nullmenge) in M,*

(2) *Für jedes $a \in A$, gibt es eine Karte $\varphi : U \longrightarrow V$ mit $a \in U$ derart, daß $\varphi(A\cap U)$ eine Lebesgue–meßbar (bzw. Nullmenge) in $\mathbb{R}^n$ ist.*

Beweis: Wir beweisen den Satz für Nullmengen. Der Fall der Lebesgue–meßbaren Mengen ist ganz analog.

(1) ⇒ (2) ist offensichtlich. Zu (2) ⇒ (1): Sei $\varphi_1 : U_1 \longrightarrow V_1$ eine Karte. Falls $A\cap U_1 = \emptyset$ ist $\varphi_1(A) = \emptyset$ eine Nullmenge. Falls $A\cap U_1 \neq \emptyset$ sei $a \in A\cap U_1$. Es gibt $\varphi : U \longrightarrow V$ mit $a \in U$ und $\varphi(A)$ Nullmenge in $\mathbb{R}^n$. Sei $W = U\cap U_1$. Die Abbildung $\Theta : \varphi(W) \longrightarrow \varphi_1(W)$, $\Theta = \varphi_1\circ\varphi^{-1}$ ist ein Diffeomorphismus. Der Satz 13.3.5 impliziert, daß $\Theta(\varphi(W\cap A)) = \varphi_1(A\cap W)$ eine Nullmenge ist. Weil der Punkt a willkürlich gewählt ist, finden wir eine Überdekung $\mathcal{U}$ von $A\cap U_1$ mit Kartengebiete U, so dass $\varphi_1(A\cap U)$ eine Nullmenge ist. Weil $A \subset \mathbb{R}^N$, besitzt A eine abzählbare Basis also nach dem Satz 8.6.9 von Lindelöf gibt es eine abzählbare Teilüberdeckung $A\cap U_1 = \cup_{k\geqslant 1}(A\cap W_k)$ mit $\varphi_1(A\cap W_k)$ Nullmenge. Dann ist auch $\varphi_1(A\cap U_1) = \cup_{k\geqslant 1}\varphi_1(A\cap W_k)$ eine Nullmenge. □

14.1.5. **Satz.** *Die Menge $\mathcal{L}(M)$ aller meßbaren Teilmenge von M bildet eine σ-Algebra, die die σ-Algebra der Borelmengen von M enthält. Die Vereinigung abzählbar vieler Nullmengen ist eine Nullmenge.*

14.1.6. **Definition.** Sei $M^n \subset \mathbb{R}^N$ eine orientierte n–dimensionale Untermanigfaltigkeit und ω_M die Volumenform von M. Für jede n–Form ω auf M gibt es eine eindeutig bestimmte $f : M \to \mathbb{R}$ mit $\omega = f\,\omega_M$. Die Form ω heißt *nicht-negativ*, geschrieben $\omega \geqslant 0$, falls $f \geqslant 0$. Setze $\omega_+ := f_+\,\omega_M$ und $\omega_- := f_-\,\omega_M$. Die Form ω heißt *meßbar*, falls f $(\mathcal{L}(M),\mathcal{B}(\mathbb{R}))$-meßbar ist.

14.1.7. **Bemerkung.** ω ist meßbar $\Longleftrightarrow$ Für jede Karte (U,φ) von M ist $(\varphi^{-1})^*\omega$ meßbar auf $\varphi(U) \subset \mathbb{R}^n$.

Wir definieren das Integral einer meßbaren nicht-negativen n-Form in drei Schritten. Zunächst auf kompakte Teilmengen enthalten in Kartengebiete, dann auf beliebige kompakte Teilmengen und schließlich auf der ganzen Untermannigfaltigkeit durch eine kompakte Ausschöpfung. Für den zweiten Schritt brauchen wir den folgenden Begriff.

14.1.8. **Definition.** Sei $K \subset M$ eine kompakte Teilmenge. Eine *fast disjunkte Zerlegung* von K ist eine Menge $\{K_1, \dots, K_m\}$ von kompakten Teilmengen mit

(1) $K = \cup_{i=1}^m K_i$
(2) $\forall i \neq j \quad K_i \cap K_j$ ist eine Nullmenge in M.

Beispiel. Sei W ein kompakter Würfel. Eine Unterteilung von W durch fortgesetzte Kanten–Halbierung ist eine fast–disjunkte Zerlegung.

14.1.9. **Satz.** *Jede kompakte Teilmenge $K \subset M$ hat eine fast disjunkte Zerlegung $\{K_1, \dots, K_m\}$ mit $K_i \subset U_i$, wobei (U_i, φ_i) sind Karten mit konstanten Vorzeichen.*

Beweis: Für $x \in K$ wähle eine Karte (U_x, φ_x) mit $x \in U_x$ und U_x zusammenhängend, also mit $\varepsilon(\varphi_x)$ konstant. Sei B_x eine kompakte Kugel mit $\varphi(x) \in B_x \subset \varphi_x(U_x)$. Dann ist $\{\varphi_x^{-1}(\mathring{B}_x)\}$ eine offene Überdeckung von K. Wähle eine endliche Teilüberdeckung $\{\varphi_i^{-1}(\mathring{B}_i)\}_{i=1}^m$.

Sei $K_1' = \varphi_1^{-1}(B_1)$, und $K_i' = \varphi_i^{-1}(B_l) \setminus \bigcup_{l<i} \varphi_l^{-1}(\mathring{B}_l)$ für $i = 2, \dots, m$. Für $i < j$ ist $K_i' \cap K_j' \subset \varphi_i^{-1}(\partial B_i)$ eine Nullmenge. Wir setzen $K_i = K \cap K_i'$. Dann ist $\{K_1, \dots, K_m\}$ eine fast disjunkte Zerlegung von K. □

14.1.10. **Definition.** Seien $M^n \subset \mathbb{R}^N$ eine orientierte n–dimensionale Untermannigfaltigkeit und ω eine meßbare nicht-negative n–Form auf M.

(i) Sei K eine kompakte Teilmenge von M und (U, φ) eine Karte mit konstanten Vorzeichen und $K \subset U$. Wir setzen

$$\int_K \omega := \varepsilon(\varphi) \int_{\varphi(K)} \left(\varphi^{-1}\right)^* \omega \tag{14.2}$$

Ist $\omega = f\omega_M$, so gilt nach (12.29)

$$\int_K \omega = \int_{\varphi(K)} (f \circ \varphi^{-1}) \sqrt{\det(g_{ij})}\, d\lambda_n .$$

(ii) Sei K eine beliebige kompakte Teilmenge von M und $\{K_1, \dots, K_m\}$ eine fast disjunkte Zerlegung von K. Wir setzen

$$\int_K \omega := \sum_{i=1}^m \int_{K_i} \omega = \sum_{i=1}^m \varepsilon(\varphi_i) \int_{\varphi_i(K_i)} (\varphi_i^{-1})^*(\omega)$$

(iii) Sei $\{K_i\}_{i \in \mathbb{N}}$ eine Ausschöpfung von M mit kompakten Teilmengen. Definiere

$$\int_M \omega := \lim_{i \to \infty} \int_{K_i} \omega$$

Ist $A \in \mathcal{L}(M)$ eine meßbare Teilmenge, so setzen wir

$$\int_A \omega := \int_M \mathbb{1}_A \omega$$

(iv) Das ***Volumen*** einer meßbaren Teilmenge-1n $A \in \mathcal{L}(M)$ wird definiert durch

$$\operatorname{vol}(A) := \operatorname{vol}_n(A) := \int_A \omega_M = \int_M \mathbb{1}_A \omega_M$$

Wir setzen ein Index in $\operatorname{vol}_n(A)$ weil z.B. $\operatorname{vol}_n(S^{n-1}) = 0$ (da S^{n-1} eine Nullmenge in $\mathbb{R}^n$ ist) aber $\operatorname{vol}_{n-1}(S^{n-1}) \neq 0$ (das ist das Öberflächeninhalt der Einheitssphäre).

14.1.11. **Satz.** *Das Integral einer nicht-negativen meßbaren n-Form ist wohldefiniert.*

Beweis: (i) Sei $\varphi_1 : U \longrightarrow V_1$ eine weitere Karte mit $\varepsilon(\varphi_1)$ konstant. Dann ist $\varphi \circ \varphi_1^{-1} : V_1 \longrightarrow V_1$ ein Diffeomorphismus und $\varepsilon(\varphi_1 \circ \varphi) = \varepsilon(\varphi_1) \cdot \varepsilon(\varphi)$. Nach dem Transformationssatz 14.1.2 gilt

$$\begin{aligned}\int_V (\varphi^{-1})^* \omega &= \varepsilon(\varphi \circ \varphi_1^{-1}) \int_{V_1} (\varphi \circ \varphi_1^{-1})^* (\varphi^{-1})^* \omega = \varepsilon(\varphi_1)\varepsilon(\varphi) \int_{V_1} (\varphi_1^{-1})^* \varphi^* (\varphi^{-1})^* \omega \\ &= \varepsilon(\varphi_1)\varepsilon(\varphi) \int_{V_1} (\varphi_1^{-1})^* \omega.\end{aligned}$$

weil $\varphi^*(\varphi^{-1})^* = (\varphi^{-1}\varphi)^* = \mathrm{Id}^* = \mathrm{Id}$. Also $\varepsilon(\varphi) \int_V (\varphi^{-1})^* \omega = \varepsilon(\varphi_1) \int_{V_1} (\varphi_1^{-1})^* \omega$.

(ii) Sei $\{L_1,\dots,L_q\}$ eine andere fast disjunkte Zerlegung von K mit $L_j \subset W_j$ wobei (W_j,ψ_j) Karten konstantem Vorzeichen sind. Dann ist $\{K_i\cap L_1,\dots,K_i\cap L_q\}$ eine fast disjunkte Zerlegung von K_i für jedes $i=1,\dots,m$ und daher $\{\varphi_i(K_i\cap L_1),\dots,\varphi_i(K_i\cap L_q)\}$ eine fast disjunkte Zerlegung von $\varphi_i(K_i)$. Dann gilt

$$\sum_{i=1}^m \varepsilon(\varphi_i)\int_{\varphi_i(K_i)}(\varphi_i^{-1})^*\omega = \sum_{i=1}^m\sum_{j=1}^q \varepsilon(\varphi_i)\int_{\varphi_i(K_i\cap L_j)}(\varphi_i^{-1})^*\omega$$
$$\overset{(i)}{=} \sum_{i=1}^m\sum_{j=1}^q \varepsilon(\psi_j)\int_{\psi_j(K_i\cap L_j)}(\psi_j^{-1})^*\omega = \sum_{j=1}^q \varepsilon(\psi_j)\int_{\psi_j(L_J)}(\psi_j^{-1})^*\omega\,.$$

(iii) Seien $K\subset L$ kompakte Teilmengen. Eine fast disjunkte Zerlegung von L induziert eine fast disjunkte Zerlegung von K. Daher folgt, dass $\int_K\omega \leqslant \int_L\omega$. Diese Bemerkung impliziert sofort, dass für zwei Ausschöpfungen $\{K_i\}_{i\in\mathbb{N}}$ und $\{L_j\}_{j\in\mathbb{N}}$ gilt

$$\lim_{i\to\infty}\int_{K_i}\omega \leqslant \lim_{j\to\infty}\int_{L_j}\omega$$

und auch die umgekehrte Ungleichung. □

14.1.12. **Definition.** Sei $M^n\subset\mathbb{R}^N$ eine orientierte n–dimensionale Untermanigfaltigkeit. Eine meßbare n-Form ω auf M heißt ***integrierbar*** auf M, falls

$$\int_M\omega_+<\infty,\quad \int_M\omega_-<\infty.$$

Wir setzen dann

$$\int_M\omega := \int_M\omega_+ - \int_M\omega_-\,.$$

Sei $A\in\mathcal{L}(M)$. Wir sagen, dass ω ist integrierbar auf A, falls $\mathbb{1}_A\,\omega$ inegrierbar auf M ist. Wir setzen

$$\int_A\omega := \int_M \mathbb{1}_A\,\omega\,.$$

Eine meßbare Funktion $f:M\to\mathbb{C}$ heißt ***integrierbar*** auf M, falls die n-Form $f\omega_M$ integrierbar ist, wobei ω_M die Volumenform von M ist. In der klassichen Vektoranalysis schreibt man $dS:=\omega_M$ und

$$\int_M f\,dS = \int_M f\omega_M\,.$$

Ist (U,φ) eine Karte, so ist

$$\int_U f\,dS = \int_{\varphi(U)}(f\circ\varphi^{-1})\sqrt{\det(g_{ij})}\,d\lambda_n\,.$$

Beispiele:

(i) Eine meßbare n–Form ist auf einer Nullmenge A integrierbar und $\int_A\omega=0$

(ii) Ist ω integrierbar auf $A\in\mathcal{L}(M)$ und $B\in\mathcal{L}(M)$, $B\subset A$, so ist ω integrierbar auf B.

(iii) Eine stetige n–Form ist auf einer kompakten Teilmenge integrierbar. Insbesondere ist das Volumen einer kompakten Teilmenge endlich.

(iv) Wir berechnen das Volumen $\kappa_{n-1}(R)$ der Sphäre $S_R^{n-1}=\{x\in\mathbb{R}^n:\|x\|_2=R\}$. Wir orientieren dabei die Späre als Rand der Kugel $B_R(0)$. Die Volumenform ist

$$\omega_{S_R^{n-1}} = \frac{1}{R}\sum_{j=1}^n(-1)^{j+1}x_j\,dx_1\wedge\ldots\wedge\widehat{dx}_j\wedge\ldots\wedge dx_n$$

Betrachte die Parametrisierung ist

$$\Psi:(0,2\pi)\times(-\tfrac{\pi}{2},\tfrac{\pi}{2})^{n-2}\longrightarrow S_R^{n-1}\setminus\{x\in\mathbb{R}^n:x_1\geqslant 0,x_2=0\}$$
$$\Psi(\varphi,\vartheta_1,\dots,\vartheta_{n-2}) = P_n(R,\varphi,\vartheta_1,\dots,\vartheta_{n-2})$$

Sie ist orientierungserhaltend, da

$$\Psi^*\omega_{S_R^{n-1}} = R^{n-1}(\cos\vartheta_1)(\cos\vartheta_2)^2\cdots(\cos\vartheta_{n-2})^{n-2}\,d\varphi\wedge d\vartheta_1\wedge\ldots\wedge d\vartheta_{n-2}\,,$$

und der Koeffizient von $d\varphi\wedge d\vartheta_1\wedge\ldots\wedge d\vartheta_{n-2}$ positiv ist. Außerdem ist

$$A=\{x\in S_R^{n-1}:x_1\geqslant 0,x_2=0\}$$

eine Nullmenge in S_R^{n-1} ist. Dann gilt

$$\operatorname{vol}(S_R^{n-1}) = \int\limits_{S_R^{n-1}}\omega_{S_R^{n-1}} = \int\limits_{S_R^{n-1}\setminus A}\omega_{S_R^{n-1}} = \int\limits_{(0,2\pi)\times\left(-\frac{\pi}{2},\frac{\pi}{2}\right)^{n-2}}\Psi^*\omega_{S_R^{n-1}}$$

also

$$\kappa_{n-1}(R) = \int_{(0,2\pi)\times\left(-\frac{\pi}{2},\frac{\pi}{2}\right)^{n-2}} R^{n-1}(\cos\vartheta_1)(\cos\vartheta_2)^2\cdots(\cos\vartheta_{n-2})^{n-2}\,d\varphi d\vartheta_1\dots d\vartheta_{n-2}$$

$$\kappa_{n-1}(R) = R^{n-1}\kappa_{n-1}, \tag{14.3}$$

wobei

$$\kappa_{n-1} := \mathrm{vol}(S^{n-1}) = n\,\omega_n, \quad \text{wobei } \omega_n = \lambda_n(\overline{B}_1(0)) \tag{14.4}$$

(v) Es gilt die folgende Form des Satzes von Fubini (zwiebelweise Integration): Ist $f:\mathbb{R}^n\to\mathbb{R}$ integrierbar, dann ist für fast alle $r\in\mathbb{R}_+$ die Funktion f über S_r^{n-1} integrierbar und es gilt:

$$\int_{\mathbb{R}^n} f\,d\lambda_n = \int_0^\infty\left(\int_{S_r^{n-1}} f\,dS\right)dr = \int_0^\infty\left(\int_{S^{n-1}} f(rx)\,dS\right)r^{n-1}dr$$

14.1.13. **Satz.**

(i) *Sind ω_1,ω_2 integrierbar auf $A\in\mathcal{L}(M)$ und $\lambda_1,\lambda_2\in\mathbb{R}$, so ist $\lambda_1\omega_1+\lambda_2\omega_2$ integrierbar auf $A\in\mathcal{L}(M)$ und gilt*

$$\int_A(\lambda_1\omega_1+\lambda_2\omega_2) = \lambda_1\int_A\omega_1 + \lambda_2\int_A\omega_2\,.$$

(ii) *Sei ω integrierbar oder meßbar und nicht–negativ. Sei $A\in\mathcal{L}(M)$ und $A=\bigsqcup_{i=1}^\infty A_i$ eine Zerlegung von A mit $A_i\in\mathcal{L}(M)$. Dann gilt*

$$\int_A\omega = \sum_{i=1}^\infty\int_{A_i}\omega\,.$$

(iii) *(Transformationssatz) Seien M, N orientiert und $f:M\to N$ ein Diffeomeorphismus mit konstantem Vorzeichen. Dann gilt: Eine n-Form $\omega\in\Omega^n(N)$ ist genau dann integrierbar, wenn $f^*\omega\in\Omega^n(M)$ integrierbar ist und dann*

$$\int_N\omega = \varepsilon(f)\int_M f^*\omega\,. \tag{14.5}$$

Wir geben nun eine Anwendung des Integrals.

14.1.14. **Satz** (Satz vom Igel, Hairy Ball Theorem)**.** *Sei n gerade, dann hat jedes Vektorfeld auf S^n eine Nullstelle.*

Beweis: Sei $X\in\mathfrak{X}(S^n)$ d.h. $X:S^n\longrightarrow\mathbb{R}^{n+1}$, $X(p)\perp p$ (weil $T_pS^n=\{v\in\mathbb{R}^{n+1}: v\perp p\}$). Nehmen wir an, $X(p)$ wäre verschieden von Null für alle $p\in S^n$. O.B.d.A. kann dann $\|X(p)\|=1$ gesetzt werden. Sei $\varepsilon>0$. Eine leichte Rechnung liefert $\|p+\varepsilon X(p)\|^2=1+\varepsilon^2$. Sei $\eta(\varepsilon):=(1+\varepsilon^2)^{1/2}$. Betrachte

$$f_\varepsilon: S^n\longrightarrow S^n_{\eta(\varepsilon)},\quad f_\varepsilon(p)=p+\varepsilon X(p)$$

Sei $\omega=\sum_{j=1}^{n+1}(-1)^{j+1}x^j dx^1\wedge\ldots\wedge\widehat{dx^j}\wedge\ldots\wedge dx^{n+1}\in\Omega^n(\mathbb{R}^{n+1})$. Dann ist $\omega_r:=\omega|_{S^n_r}$ positiv für die Orientierung von S_r^n als Rand von $B_r(0)$.

Wir berechnen zunächst $f_\varepsilon^*\omega=f_\varepsilon^*\omega_{\eta(\varepsilon)}$. Dafür erweitern wir f_ε auf einer Umgebung U von S^n (indem wir X erweitern zum Beispiel zu $\widetilde{X}(p)=X(p/|p|)$), berechnen $f_\varepsilon^*\omega$ in dieser Umgebung und dann schränken wir wieder auf S^n ein. Es ergibt sich

$$\begin{aligned} f_\varepsilon^*\omega &= \sum_{j=1}^{n+1}(-1)^{j+1}f_\varepsilon^j df_\varepsilon^1\wedge\ldots\wedge\widehat{df_\varepsilon^j}\wedge\ldots\wedge df_\varepsilon^{n+1}\\ &=\omega+\varepsilon\alpha_0+\ldots+\varepsilon^{n+1}\alpha_n = \omega+\varepsilon(\alpha_0+\ldots+\varepsilon^n\alpha_n),\end{aligned}$$

mit $\alpha_0,\dots,\alpha_n$ glatte n-Formen in einer Umgebung von S^n. Es folgt

$$f_\varepsilon^*\omega|_{S^n} = \omega+\varepsilon(\alpha_0+\ldots+\varepsilon^n\alpha_n)|_{S^n} = (1+(g_0+\ldots+\varepsilon^n g_n)\varepsilon)\omega,$$

wobei $g_0,\dots,g_n\in C^\infty(S^n)$.

Behauptung: Die Abbildung f_ε ist ein orientierungserhaltender Diffeomorphismus für ε genügend klein. Der *Beweis* der Behauptung erfolgt in drei Schritten.

Schritt 1: f_ε ist eine orientierungserhaltende Immersion für ε genügend klein. In der Tat, man kann ε_0 finden, so dass für $\varepsilon<\varepsilon_0$ gilt $1+\varepsilon(g_0+\ldots+\varepsilon^n g_n)>\frac{1}{2}$. Damit ist $f_\varepsilon^*\omega>0$ und daraus folgt leicht, dass f_ε eine orientierungserhaltende Immersion ist.

Schritt 2: f_ε ist injektiv für genügend kleine ε. Wäre dem nicht so, dann würden Folgen $\varepsilon_k\searrow 0$ und $p_k\neq q_k\in S^n$ existieren, so dass $f_{\varepsilon_k}(p_k)=p_k+\varepsilon_k X(p_k)=q_k+\varepsilon_k X(q_k)=f_\varepsilon(q_k)$. Dann wäre aber

$\frac{p_k-q_k}{|p_k-q_k|} = \varepsilon_k \frac{X(q_k)-X(q_k)}{|X(q_k)-X(q_k)|}$. Man sieht sofort, dass die linke Seite Norm 1 hat und die rechte Seite gegen Null konvergiert. Widerspruch.

Schritt 3: f_ε ist ein Diffeomorphismus für ε genügend klein. $\operatorname{Im} f_\varepsilon$ ist offen, weil f_ε lokaler Diffeomorphismus ist (da Immersion zwischen Mannigfaltigkeiten gleicher Dimension ist). $\operatorname{Im} f_\varepsilon$ ist weiter kompakt (f_ε ist stetig, S^n kompakt), also abgeschlossen. Außerdem ist das Bild von f_ε nicht leer. Daraus folgt, dass $\operatorname{Im} f_\varepsilon = S^n_{\eta(\varepsilon)}$ (weil $S^n_{\eta(\varepsilon)}$ zusammenhängend ist). Also ist f_ε bijektiv und somit f_ε Diffeomorphismus, da ein bijektiver lokaler Diffeomorphismus ein Diffeomorphismus ist. Die Schritte 1-3 beweisen die Behauptung.

Die Form $\frac{1}{r}\omega_r$ ist die Volumenform von S^n_r also $\operatorname{vol}(S^n_r) = \int_{S^n_r} \frac{1}{r}\omega_r$. Nun gilt nach (14.3): $\operatorname{vol}(S^n_r) = r^n \kappa_n$ mit $\kappa_n = \operatorname{vol}(S^n)$. Es folgt $\int_{S^n_r} \omega_r = r^{n+1}\kappa_n$. Mit Hilfe des Transformationssatzes 14.1.13(iii) erhalten wir:

$$\eta(\varepsilon)^{n+1}\kappa_n = \int\limits_{S^n_{\eta(\varepsilon)}} \omega = \int\limits_{S^n} f_\varepsilon^* \omega = \int\limits_{S^n} (\omega + \varepsilon\alpha_0 + \ldots + \varepsilon^{n+1}\alpha_n) = \kappa_n + \varepsilon \int\limits_{S^n} \alpha_0 + \ldots + \varepsilon^{n+1} \int\limits_{S^n} \alpha_n,$$

Man hat also insgesamt

$$(1+\varepsilon^2)^{\frac{n+1}{2}} \kappa_n = \kappa_n + \varepsilon \int\limits_{S^n} \alpha_0 + \ldots + \varepsilon^{n+1} \int\limits_{S^n} \alpha_n$$

Man sieht, dass auf der rechten Seite ein Polynom in ε steht. Auf der linken Seite steht nur dann ein Polynom in ε, wenn n ungerade ist. Für gerade n erhält man einen Widerspruch, es existiert also kein nullstellenfreies Vektorfeld. $\square$

14.2. **Der Satz von Stokes.**

14.2.1. **Satz** (Stokes). *Seien M eine n-dimensionale orientierte Untermannigfaltigkeit von $\mathbb{R}^N$ und $D \subset M$ eine glatt berandete Teilmenge. Sei ω eine $(n-1)$-Form der Klasse $\mathcal{C}^1$, so dass $D \cap \operatorname{supp}\omega$ kompakt ist. Dann gilt*

$$\int_{\partial D} \iota^*\omega = \int_D d\omega \tag{14.6}$$

wobei ∂D mit der Randorientierung versehen ist und $\iota : \partial D \to M$ die Inklusion ist.

Meistens wird kurz geschrieben

$$\int_{\partial D} \omega := \int_{\partial D} \iota^*\omega,$$

wobei $\iota : \partial D \to M$ die Inklusion ist, d.h. $\iota^*\omega$ ist die Einschänkung von ω auf ∂D. Der Beweis zeigt auch:

14.2.2. **Satz.** *Ist M n-dimensionale orientierte Untermannigfaltigkeit, so gilt für alle $(n-1)$-Formen ω der Klasse $\mathcal{C}^1$ mit kompaktem Träger auf M:*

$$\int_M d\omega = 0. \tag{14.7}$$

Insbesondere gilt (14.7) *für alle $(n-1)$-Formen ω der Klasse $\mathcal{C}^1$ mit auf einer* kompakten *Mannigfaltigkeit M.*

Wenn $\partial D = \emptyset$ vereinbaren wir, dass

$$\int_{\partial D} \omega = \int_{\emptyset} \omega := 0.$$

Wegen Satz 14.2.2 gilt dann mit dieser Vereinbarung den Satz 14.2.3 auch wenn $\partial D = \emptyset$.

Beweis: 1. Schritt. Wir reduzieren zunächst das Problem zum Fall, wenn den Träger von ω in einem Kartengebiet enthalten ist. Dies geschiet typischerweise durch Benutzung einer Zerlegung der Eins. Sei $K := D \cap \operatorname{supp}\omega$ und sei $(U_1, \varphi_1), \ldots, (U_m, \varphi_m)$ Karten auf M mit folgenden Eigenschaften:

- Alle (U_i, φ_i) haben konstanten Vorzeichen,
- $K \subset \cup_{i=1}^m U_i$,
- (U_i, φ_i) ist eine D-Karte, falls $U_i \cap \partial D \neq \emptyset$.

Dies erreichen wir, indem wir für jeden Punkt $x \in \partial D \cap K$ eine D-Karte (U_x, φ_x) und für jeden Punkt $x \in K \setminus \partial D$ eine Karte (U_x, φ_x) mit $U_x \cap \partial D = \emptyset$ und jeweils $x \in U_x$ wählen. Die U_x werden dabei zusammenhängend gewählt, also mit konstanten Vorzeichen. Wegen der Kompaktheit von K besitzt die offene Überdeckung $\{U_x\}_{x \in K}$ eine Teilüberdeckung.

Dann ist $\{U_1,\dots,U_m,X\setminus K\}$ eine offene Überdeckung von M. Sei $\{\lambda_1,\dots,\lambda_m,\lambda_{m+1}\}$ eine untergeordnete Zerlegung der Eins. Da $\lambda_{m+1}=0$ auf K gilt $\sum_{i=1}^m \lambda_i = 1$ auf K. Wir werden beweisen, dass

$$\int_{\partial D}\lambda_i\omega = \int_D d(\lambda_i\omega), \quad \text{für alle } i=1,\dots,m \tag{14.8}$$

Es gilt

$$\int_{\partial D}\omega = \int_{\partial D\cap K}\omega = \int_{\partial D\cap K}\Big(\sum_{i=1}^m \lambda_i\Big)\omega = \int_{\partial D}\Big(\sum_{i=1}^m \lambda_i\Big)\omega$$

und weiter erhalten wir

$$\int_{\partial D}\Big(\sum_{i=1}^m \lambda_i\Big)\omega = \sum_{i=1}^m\int_{\partial D}\lambda_i\omega \overset{(14.8)}{=} \sum_{i=1}^m\int_D d(\lambda_i\omega) = \int_D\sum_{i=1}^m d(\lambda_i\omega)$$

Nach der Leibniz-Regel gilt:

$$\sum_{i=1}^m d(\lambda_i\omega) = \sum_{i=1}^m\Big(d\lambda_i\wedge\omega + \lambda_i\,d\omega\Big) = d\Big(\sum_{i=1}^m\lambda_i\Big)\wedge\omega + \Big(\sum_{i=1}^m\lambda_i\Big)d\omega$$

Nun ist $\omega=0$ und $d\omega=0$ auf $M\setminus K$ und $\sum_{i=1}^m\lambda_i=1$, $d\Big(\sum_{i=1}^m\lambda_i\Big)=0$ auf K also

$$\begin{aligned}\int_D\sum_{i=1}^m d(\lambda_i\omega) &= \int_D d\Big(\sum_{i=1}^m\lambda_i\Big)\wedge\omega + \Big(\sum_{i=1}^m\lambda_i\Big)d\omega = \int_K d\Big(\sum_{i=1}^m\lambda_i\Big)\wedge\omega + \Big(\sum_{i=1}^m\lambda_i\Big)d\omega\\ &= \int_K d\omega = \int_D d\omega\end{aligned}$$

Diese Kette von Gleichungen zeigt, dass (14.6) aus (14.8) folgt. Wir beweisen nun (14.8).

2. Schritt. Den Beweis von (14.8) reduzieren wir zu einer Rechnung in $\mathbb{R}^n$. Der Einfachheitshalber schreiben wir $U=U_i$, $\lambda_i\omega=\eta$, $\varphi_i=\varphi$. Wegen der Definition 14.2 gilt

$$\int_D d\eta := \varepsilon(\varphi)\int_{\varphi(U\cap D)}(\varphi^{-1})^*(d\eta) = \varepsilon(\varphi)\int_{\varphi(U\cap D)} d(\varphi^{-1})^*\eta. \tag{14.9}$$

Ist $U\cap\partial D=\emptyset$, so gilt $\eta=0$ auf ∂D also

$$\int_{\partial D}\eta = 0. \tag{14.10}$$

Ist $U_i\cap\partial D\neq\emptyset$, so ist $(U,\varphi=(x_1,\dots,x_n))$ eine D-karte. Laut Beispiel 12.4.8 ist das Vorzeichen der Abbildung $\varphi|_{\partial D\cap U}=\varphi\circ\iota_{\partial D}:\partial D\cap U\to\partial\mathbb{R}^n_-$ auch $\varepsilon(\varphi)$. Es gilt $(\varphi|_{\partial D\cap U})^{-1}=\iota_{\partial D}^{-1}\circ\varphi^{-1}\circ\iota_{\partial\mathbb{R}^n_-}$ und deshalb $((\varphi|_{\partial D\cap U})^{-1})^* = \iota^*_{\partial\mathbb{R}^n_-}\circ(\varphi^{-1})^*\circ(\iota_{\partial D}^{-1})^*$. Mit (14.5) folgt

$$\int_{\partial D\cap U}\iota^*_{\partial D}\eta = \varepsilon(\varphi)\int_{\varphi(\partial D\cap U)}((\varphi|_{\partial D\cap U})^{-1})^*\iota^*_{\partial D}\eta = \varepsilon(\varphi)\int_{\partial\mathbb{R}^n_-}\iota^*_{\partial\mathbb{R}^n_-}(\varphi^{-1})^*\eta. \tag{14.11}$$

Die Gleichheiten (14.9) (14.10), (14.11) zeigen, dass um (14.8) zu beweisen reicht es, für die Form $\alpha := (\varphi^{-1})^*\eta$ folgendes zu zeigen:

$$\int_{\mathbb{R}^n} d\alpha = 0, \tag{14.12}$$

$$\int_{\mathbb{R}^n_-} d\alpha = \int_{\partial\mathbb{R}^n_-}\iota^*\alpha. \tag{14.13}$$

3. Schritt. Wir beweisen nun (14.12) und (14.13). Sei α eine $\mathcal{C}^1$ $(n-1)$-Form in $\mathbb{R}^n$, mit kompaktem Träger, $\operatorname{supp}\alpha\subset[-r,r]^n$, für $r>0$; α hat die Form

$$\alpha = \sum_{j=1}^n(-1)^{j+1}f_j\,dx_1\wedge\ldots\wedge\widehat{dx_j}\wedge\ldots\wedge dx_n$$

mit $f_j\in\mathcal{C}^1_0([-r,r]^n)$. Daher

$$d\alpha = \Big(\sum_{j=1}^n\frac{\partial f_j}{\partial x_j}\Big)dx_1\wedge\ldots\wedge dx_n\,,\qquad \int_{\mathbb{R}^n}d\alpha = \sum_{j=1}^n\int_{\mathbb{R}^n}\frac{\partial f_j}{\partial x_j}\,d\lambda_n$$

Sei $j\in\{1,\dots,n\}$ fest. Nach dem Satz von Fubini gilt

$$\int_{\mathbb{R}^n}\frac{\partial f_j}{\partial x_j}\,d\lambda_n = \int_{\mathbb{R}^{n-1}}\Big(\int_{\mathbb{R}}\frac{\partial f_j}{\partial x_j}\,dx_j\Big)d\lambda_{n-1}(x_1,\dots,\widehat{x_j},\dots,x_n) \tag{14.14}$$

Der Hauptsatz angewendet auf $t\mapsto f(x_1,\dots,t,\dots,x_n)$ liefert

$$\int_{\mathbb{R}}\frac{\partial f_j}{\partial x_j}\,dx_j = \int_{-r}^r\frac{\partial f_j}{\partial x_j}\,dx_i = f_j(x_1,\dots,r,\dots,x_n) - f_j(x_1,\dots,-r,\dots,x_n) = 0 \tag{14.15}$$

weil $f_j(x_1,\dots,r,\dots,x_n)=f_j(x_1,\dots,-r,\dots,x_n)=0$. Daher verschwinden alle Integrale aus (14.14) und

$$\int_{\mathbb{R}^n} d\alpha=\sum_{j=1}^n\int_{\mathbb{R}^n}\frac{\partial f_j}{\partial x_j}\,d\lambda_n=0$$

und somit gilt (14.12). Wir beweisen nun (14.13). Es ist

$$\int_{\mathbb{R}^n_-} d\alpha=\sum_{j=1}^n\int_{\mathbb{R}^n_-}\frac{\partial f_j}{\partial x_j}\,d\lambda_n \tag{14.16}$$

Sei $j\in\{1,\dots,n-1\}$ fest. Nach dem Satz von Fubini gilt

$$\int_{\mathbb{R}^n_-}\frac{\partial f_j}{\partial x_j}\,d\lambda_n=\int_{\mathbb{R}^{n-2}\times(-\infty,0]}\Big(\int_{\mathbb{R}}\frac{\partial f_j}{\partial x_j}\,dx_j\Big)\,d\lambda_{n-1}(x_1,\dots,\widehat{x_j},\dots,x_n)$$

und (14.15) liefert

$$\int_{\mathbb{R}^n_-}\frac{\partial f_j}{\partial x_j}\,d\lambda_n=0,\quad j\in\{1,\dots,n-1\}. \tag{14.17}$$

Für $j=n$ gilt nach dem Satz von Fubini

$$\int_{\mathbb{R}^n_-}\frac{\partial f_n}{\partial x_n}\,d\lambda_n=\int_{\mathbb{R}^{n-1}}\Big(\int_{-\infty}^0\frac{\partial f_n}{\partial x_n}\,dx_n\Big)\,d\lambda_{n-1}(x_1,\dots,x_{n-1})$$

und der Hauptsatz liefert

$$\begin{aligned}\int_{-\infty}^0\frac{\partial f_n}{\partial x_n}\,dx_n&=\int_{-r}^0\frac{\partial f_n}{\partial x_n}\,dx_n=f_n(x_1,\dots,x_{n-1},0)-f_n(x_1,\dots,x_{n-1},-r)\\&=f_n(x_1,\dots,x_{n-1},0)\end{aligned}$$

Folglich

$$\int_{\mathbb{R}^n_-}\frac{\partial f_n}{\partial x_n}\,d\lambda_n=\int_{\mathbb{R}^{n-1}}f_n(x_1,\dots,x_{n-1},0)\,d\lambda_{n-1}(x_1,\dots,x_{n-1}) \tag{14.18}$$

Nach Definition des Integrals

$$\int_{\partial\mathbb{R}^n_-}\iota^*\alpha=\varepsilon(\psi)\int_{\mathbb{R}^{n-1}}\psi^*\iota^*\alpha$$

wobei $\psi:\mathbb{R}^{n-1}\to\partial\mathbb{R}^n_-$ die Parametrisierung $(x_1,\dots,x_{n-1})\mapsto(x_1,\dots,x_{n-1},0)$ ist. Da $\frac{\partial}{\partial x_n}$ der äußere Normalenvektor zu $\mathbb{R}^n_-$ ist, ist das Vorzeichen dieser Parametrisierung das Vorzeichen der Basis $(\frac{\partial}{\partial x_n},\frac{\partial}{\partial x_1},\dots,\frac{\partial}{\partial x_{n-1}})=(e_n,e_1,\dots,e_{n-1})$, d.h. $(-1)^{n-1}$.

Es ist klar, dass $\iota\circ\psi=\psi$, $\psi^*\iota^*=(\iota\circ\psi)^*=\psi^*$ und die zurückgezogene Form $\psi^*\alpha$ auf $\mathbb{R}^{n-1}$ erhalten wir, indem wir „$x_n=0$ in der Ausdruck von α ersetzen“:

$$\psi^*\iota^*\alpha=\psi^*\alpha=(-1)^{n+1}f_n(x_1,\dots,x_{n-1},0)\,dx_1\wedge\ldots\wedge dx_{n-1}.$$

Also

$$\begin{aligned}\int_{\partial\mathbb{R}^n_-}\iota^*\alpha&=(-1)^{n-1}\int_{\mathbb{R}^{n-1}}(-1)^{n+1}f_n(x_1,\dots,x_{n-1},0)\,dx_1\wedge\ldots\wedge dx_{n-1}\\&=\int_{\mathbb{R}^{n-1}}f_n(x_1,\dots,x_{n-1},0)\,d\lambda_{n-1}(x_1,\dots,x_{n-1})\end{aligned} \tag{14.19}$$

Nach (14.16), (14.17), (14.18) und (14.19) erhalten wir schließlich

$$\begin{aligned}\int_{\mathbb{R}^n_-}d\alpha&=\sum_{j=1}^n\int_{\mathbb{R}^n_-}\frac{\partial f_j}{\partial x_j}\,d\lambda_n=\int_{\mathbb{R}^n_-}\frac{\partial f_n}{\partial x_n}\,d\lambda_n=\int_{\mathbb{R}^{n-1}}f_n(x_1,\dots,x_{n-1},0)\,d\lambda_{n-1}(x_1,\dots,x_{n-1})\\&=\int_{\partial\mathbb{R}^n_-}\iota^*\alpha\end{aligned}$$

d.h. (14.13). □

14.2.3. **Satz** (Satz von Gauss-Ostrogradski, Divergenzsatz)**.** *Seien M eine n-dimensionale orientierte Untermannigfaltigkeit von $\mathbb{R}^N$ und $D\subset M$ eine glatt berandete Teilmenge. Sei $X\in\mathscr{X}(M)$ ein Vektorfeld der Klasse $\mathcal{C}^1$, so dass $D\cap\operatorname{supp}X$ kompakt ist. Dann gilt*

$$\int_D\operatorname{div}(X)\omega_M=\int_{\partial D}\langle X,\nu\rangle\omega_{\partial D} \tag{14.20}$$

wobei ∂D mit der Randorientierung versehen ist, ν das äußere Einheitsnormalenfeld von D ist und $\omega_{\partial D}$ das Volumenform von ∂D ist.

15. Lösungen zu den Aufgaben

Aufgabe 1.8.4. Wir wenden den binomischer Lehrsatz:

$$(1+x)^n = \sum_{k=1}^{n} \binom{n}{k} x^k. \tag{15.1}$$

Zu a): Setze $x = 1$ in (15.1) $\rightsquigarrow$ Behauptung.

Alternativ: Die Anzahl der Teilmengen einer n-elementigen Menge ist 2^n. Diese Anzahl ist aber auch die Summe der Anzahlen der 0-elementigen Teilmengen (d. h. $\binom{n}{0}$), 1-elementigen Teilmengen (d. h. $\binom{n}{1}$), ..., k-elementigen Teilmengen (d. h. $\binom{n}{k}$), ..., n-elementigen Teilmengen (d. h. $\binom{n}{n}$).

b): Wir nutzen die Formel

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

(für $k = 0$ beide Glieder sind Null, weil $\binom{m}{l} = 0$ für $l < 0$). Dann gilt

$$\begin{aligned}
\sum_{k=0}^{n} k\binom{n}{k} &= \sum_{k=1}^{n} k\binom{n}{k} \\
&= \sum_{k=1}^{n} n\binom{n-1}{k-1} && \text{nach der obigen Formel} \\
&= n\sum_{j=0}^{n-1} \binom{n-1}{j} && \text{Indexverschiebung: } j := k-1 \\
&= n\cdot 2^{n-1} && \text{nach a)}
\end{aligned}$$

c): Setze $x = -1$ in (15.1) $\rightsquigarrow$ Behauptung.

d): Wir nutzen die Formel

$$\frac{1}{k+1}\binom{n}{k} = \frac{1}{n+1}\binom{n+1}{k+1}.$$

Damit folgt

$$\begin{aligned}
\sum_{k=0}^{n} (-1)^k \frac{1}{k+1}\binom{n}{k} &= \sum_{k=0}^{n} (-1)^k \frac{1}{n+1}\binom{n+1}{k+1} \\
&= -\frac{1}{n+1}\sum_{k=0}^{n} (-1)^{k+1}\binom{n+1}{k+1} \\
&= -\frac{1}{n+1}\sum_{j=1}^{n+1} (-1)^{j}\binom{n+1}{j} \quad \text{Indexverschiebung: } j := k+1 \\
&= -\frac{1}{n+1}\left[\sum_{j=0}^{n} (-1)^{j}\binom{n+1}{j} - \binom{n+1}{0}\right] \\
&= -\frac{1}{n+1}[0-1] \quad \text{nach c)} \\
&= \frac{1}{n+1}
\end{aligned}$$

Aufgabe 1.8.5. Wir betrachten zunächst den Spezialfall $x_1 \dots x_n = 1$. Dann lautet die Ungleichung $x_1 + \dots + x_n \geq n$. Für $n = 1$ lautet die Aussage $x_1 = 1 \Longrightarrow x_1 \geq 1$, sie ist also wahr. Induktionsschluss $n \rightsquigarrow n+1$. Seien $x_1, \dots, x_{n+1} > 0$ mit $x_1 \dots x_{n+1} = 1$. Sind alle $x_k = 1$ so ist $x_1 + \dots + x_{n+1} = n+1$ und der Beweis ist fertig. Ansonsten gibt es $k \in \{1, \dots, n+1\}$ mit $x_k \neq 1$; durch Umnummerierung oBdA $k = 1$, d.h. $x_1 \neq 1$. Ist nun $x_1 < 1$ so muss ein $\ell \in \{2, \dots, n+1\}$ existieren mit $x_\ell > 1$ (sonst gilt $x_\ell \leq 1$ für alle $\ell \in \{2, \dots, n+1\}$ und $x_1 \dots x_{n+1} \leq x_1 < 1$, Widerspruch). Wir können oBdA annehmen, dass $\ell = 2$, d.h. $x_2 > 1$. Sei $b = x_1 x_2$. Wegen $b x_3 \dots x_{n+1} = 1$ (n Faktoren) folgt aus der Induktionsvoraussetzung $x_1 x_2 + x_3 + \dots + x_{n+1} > n$. Nun gilt $(1-x_1)(1-x_2) < 0$ und das ist äquivalent zu $1 - x_1 - x_2 + x_1, x_2 < 0$ also $x_1 x_2 < x_1 + x_2 - 1$. Daraus folgt:

$$n \leq x_1 x_2 + x_3 + \dots + x_{n+1} < x_1 + x_2 - 1 + x_3 + \dots + x_{n+1}$$

also

$$n+1 < x_1 + \dots + x_{n+1}.$$

Ist $x_1 > 1$, so muss ein $\ell \in \{2,\dots,n+1\}$ existieren, mit $x_\ell < 1$ oBdA $x_2 < 1$. Dann gilt $(1-x_1)(1-x_2) < 0$ und der Beweis folgt wie oben. Allgemeiner Fall. Setze $\lambda = x_1 \dots x_n$. Gibt es $\ell \in \{1,\dots,n\}$ mit $x_\ell = 0$ so ist die Ungleichung trivial und die Gleichheit tritt ein, genau dann, wenn $x_1 = \dots = x_n = 0$. Sind alle $x_k > 0$, so ist $\lambda > 0$.

Setze $x_k' = \frac{x_k}{\sqrt[n]{\lambda}}$. Dann ist

$$x_1' \dots x_k' = \frac{x_1 \dots x_n}{\lambda} = 1.$$

Nach dem Spezialfall folgt $x_1' + \dots + x_n' \geq n$ also $x_1 + \dots + x_n \geq n\sqrt[n]{\lambda}$. Dabei gilt das Gleichheitszeichen genau dann, wenn alle $x_k' = 1$, d.h. $x_k = \sqrt[n]{\lambda}$.

Aufgabe 1.8.6. (i) Die AGM- Ungleichung liefert:

$$\sum_{k=1}^{n} a_k \geq n\sqrt[n]{a_1,\dots a_n}\,, \quad \sum_{k=1}^{n} \frac{1}{a_k} \geq n\sqrt[n]{\frac{1}{a_1 \dots a_n}}\,.$$

Durch Multiplikation folgt die Behauptung. Die Gleichheit tritt ein, genau dann, wenn Gleichheit in beiden obigen Ungleichungen eintritt, also genau dann, wenn $a_1 = \dots = a_n$.

(ii) Zur 1. Ungleichung. Wende die Bernoulli- Ungleichung an: $\left(1+\frac{a-1}{n}\right)^n > 1 + n \cdot \frac{a-1}{n} = a$ oder $p = 1$ in der 2. Ungleichung. Zur 2. Ungleichung. Die AGM-Ungleichung liefert:

$$\sqrt[n]{a^p} = \sqrt[n]{a \cdot \dots \cdot a \cdot 1 \cdot \dots \cdot 1} < \frac{1}{n}(p \cdot a + n - p) = 1 + \frac{p}{n}(a-1)\,.$$

wobei unter der n-ten Wurzel a p-mal und 1 $(n-p)$-mal vorkommt.

Aufgabe 1.8.8. (a) Fall 1: $k > m$. Dann $\binom{m}{k} = 0 < \binom{n}{k}$. Fall 2: $k \leq m$. Dann gilt $m < n, m-1 < n-1, \dots, m-k+1 < n-k+1$ und alle diese Zahlen sind positiv. Durch Multiplikation folgt $m(m-l)\dots(m-k+1) < n(n-1)\dots(n-k+1)$ und durch Division mit $k!$ folgt die Behauptung.

(b) Zur 1. Ungleichung. Klar falls $k > m$. Sei $k \leq m$, dann

$$\frac{1}{m^k}\binom{m}{k} = \frac{1}{k!}\cdot\frac{m}{m}\cdot\frac{m-1}{m}\cdots\frac{m-k+1}{m} < \frac{1}{k!}\cdot\frac{n}{n}\cdot\frac{n-1}{n}\cdots\frac{n-k+1}{n} = \frac{1}{n^k}\binom{n}{k}$$

denn $\frac{m-\ell}{m} < \frac{n-\ell}{n}$ für $\ell \in \{1,\dots,k-1\}$, denn $-\frac{\ell}{m} < -\frac{\ell}{n}$.

Zur 2. Ungleichung.

$$\frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k+1}{n} < 1, \quad \text{weil } \frac{n-1}{n} < 1, \dots, \frac{n-k+1}{n} < 1\,.$$

Zur 3. Ungleichung.

$$\frac{1}{k!} = \frac{1}{1}\cdot\frac{1}{2}\cdots\frac{1}{k} \leq 1 \cdot \underbrace{\frac{1}{2}\cdots\frac{1}{2}}_{(k-1)\text{–mal}} = \frac{1}{2^{k-1}}.$$

(c) Fall 1: $m = 1$ (dann $n \geq 2$). Dann $\left(1+\frac{1}{m}\right)^m = 2$, $\left(1+\frac{1}{n}\right)^n > 1 + n \cdot \frac{1}{n} = 2$, nach Bernoulli. Fall 2: $m, n \geq 2$. Dann

$$\left(1+\frac{1}{m}\right)^m = 1 + 1 + \sum_{k=2}^{m}\binom{m}{k}\frac{1}{m^k} \overset{(b)}{<} 1 + 1 + \sum_{k=2}^{m}\binom{n}{k}\frac{1}{n^k} < 1 + 1 + \sum_{k=2}^{n}\binom{n}{k}\frac{1}{n^k} = \left(1+\frac{1}{n}\right)^n.$$

Aufgabe 1.8.11. (a) $2+\frac{3}{2}$ ist eine obere Schranke von M_1, denn $2+\frac{3}{4}$, $2+\frac{3}{6}$, ... sind kleiner, übrige Elemente sogar negativ. Weil $2+\frac{3}{2} \in M_1$, ist $2+\frac{3}{2} = \max M_1$, also auch $2+\frac{3}{2} = \sup M_1$. $-(2+\frac{3}{1}) = -5$ ist untere Schranke von M_1, denn $-(2+\frac{3}{3})$, $-(2+\frac{3}{5})$, ... sind grösser, übrige Elemente von M_1 sogar positiv. Da $-5 \in M$, gilt $-5 = \min M_1, -5 = \sup M_1$.

(b) Seien $M_2' = \{-\frac{1}{3}, \frac{1}{9}, -\frac{1}{27}, \frac{1}{81}, \dots\}$, $M_2'' = \{5, \frac{5}{2}, \frac{5}{3}, \frac{5}{4}, \dots\}$. M enthält alle Summen von je einem Element von M_2' bzw. M_2''. Sowohl M_2' als auch M_2'' haben ein grösstes Element: $\frac{1}{5}$ bwz. 5. Also ist $5+\frac{1}{9}$ eine obere Schranke von M_2 und in M_2 enthalten, also $5+\frac{1}{9} = \max M_2$ und deshalb auch $5+\frac{1}{9} = \sup M_2$. Nun ist $-\frac{1}{3}$ eine untere Schranke von M_2, da $-\frac{1}{3} = \min M_2'$ und alle Elemente von M_2'' sind positiv. $-\frac{1}{3}$ ist auch die grösste untere Schranke. Denn wäre ein $s > -\frac{1}{3}$ untere Schranke, so für $n \in \mathbb{N}$ mit $\frac{5}{n} < s - (-\frac{1}{3})$ (existiert nach Eudoxos) wäre $-\frac{1}{3} + \frac{5}{n} < s$. Widerspruch zur Wahl von s als untere Schranke! Also ist $-\frac{1}{3} = \inf M_2$. Da $-\frac{1}{3} \notin M_2$ hat M_2 kein Minimum.

(c) Sei $f(x) = (x+a)(x+b)(x+c)$. Dann gilt $f(x) < 0$ für $x < -c$ und $-b < x < -a$ und $f(x) > 0$ für $-c < x < -b$ und $x > 0$. Also $M_3 = \{x \in R : -c < x < -b \text{ oder } x > a\}$, $\inf M_3 = -c \notin M_3$ und M_3 hat kein Minimum. Ausserdem hat M_3 keine obere Schranke, also kein Supremum und Maximum.

Aufgabe 1.8.14. (i) Induktion nach $n \in \mathbb{N}$. Für $n = 1$ gilt, nach der AGM- Ungleichung

$$b_1^2 = \left(\frac{a_0 + b_0}{2}\right)^2 > a_0 b_0 = b$$

(die Ungleichung ist strikt, da $a_0 \neq b_0$, wegen $a^2 \neq b$). Es folgt $b_1 > \frac{b}{b_1} = a_1$.
Induktionsschluss $n \rightsquigarrow n+1$: nach der AGM- Ungleichung gilt

$$b_{n+1}^2 = \left(\frac{a_n + b_n}{2}\right)^2 > a_n b_n = b$$

(die Ungleichung ist strikt, da $a_n \neq b_n$ nach der Induktionsannahme). Also $b_{n+1} > \frac{b}{b_{n+1}} = a_{n+1}$.
(ii) Sei $I_n = [a_n, b_n]$. Zu zeigen ist $a_n \leq a_{n+1} < b_{n+1} \leq b_{n+1}$, d.h. $I_{n+1} \subset I_n$. Zunächst gilt $b_{n+1} = \frac{a_n + b_n}{2} < b_n \Leftrightarrow a_n < b_n$ und dies ist wahr, nach (i). Ausserdem $a_{n+1} = \frac{b}{b_{n+1}} > \frac{b}{b_n} = a_n$. Berechne (wegen $a_n < a_{n+1}$):

$$|I_{n+1}| = b_{n+1} - a_{n+1} = \frac{a_n + b_n}{2} - a_{n+1} = \frac{a_n - 2a_{n+1} + b_n}{2} < \frac{b_n - a_n}{2} = \frac{1}{2}|I_n|$$

Durch Induktion folgt $|I_n| < \frac{1}{2^n}|I_1|$ also (I_n) ist eine Intervallschachtelung (da $\left(\frac{1}{2^n}\right)$ eine Nullfolge ist). Sei $\{x\} = \cap I_n$. Dann gilt $a_n \leq x \leq b_n$ für alle $n \in \mathbb{N}$. Ausserdem $a_n b_n = b$. Es folgt $x a_n \leq b \leq x b_n$, also $a_n \leq \frac{b}{x} \leq b_n$, äquivalent $\frac{b}{x} \in [a_n, b_n]$ für alle $n \in \mathbb{N}$, $\frac{b}{x} \in \cap I_n = \{x\}$. Somit gilt $\frac{b}{x} = x$, also $x^2 = b$.

Lösung mit dem Konvergenzbegriff: Es gilt $a_n \to x$, $b_n \to x$, $n \to \infty$. Wegen $x = \lim a_n = \lim \frac{b}{b_n} = \frac{b}{x}$ folgt $x^2 = b$, also $x = \sqrt{b}$.
(Beachte, dass $0 < a_1 \leq a_n < b_n$ für alle $n \in \mathbb{N}$, also $x > 0$).

Aufgabe 1.8.15.

(a) (i) Es gilt

$$\begin{aligned}\frac{3+4i}{2-i} &= \frac{(3+4i)(2+i)}{|2-i|^2} \\ &= \frac{2}{5} + i\frac{11}{5}\end{aligned}$$

und

$$\left|\frac{3+4i}{2-i}\right| = \sqrt{\frac{125}{25}} = \sqrt{5}.$$

(ii) Es gilt

$$(1+i)^2 = 2i \Rightarrow (1+i)^8 = (2i)^4 = 16.$$

(iii) Sei $n \in \mathbb{Z}$. Dann gilt

$$\left(\frac{1+i}{1-i}\right)^n = \left(\frac{(1+i)^2}{|1-i|^2}\right)^n = \left(\frac{2i}{2}\right)^n = i^n = \begin{cases} 1 & , n \in 4\mathbb{Z} \\ i & , n \in 4\mathbb{Z}+1 \\ -1 & , n \in 4\mathbb{Z}+2 \\ -i & , n \in 4\mathbb{Z}+3 \end{cases}$$

und $|i^n| = 1$, letzteres für alle n.

(b) Seien $z, w \in \mathbb{C}$.

(i) Eine direkte Rechnung gibt

$$|z+w|^2 + |z-w|^2 = (z+w)(\bar{z}+\bar{w}) + (z-w)(\bar{z}-\bar{w}) = |z|^2 + |w|^2.$$

(ii) Wir haben

$$\begin{aligned}|z+w|^2 &= |z|^2 + 2\operatorname{Re} z\bar{w} + |w|^2 \\ &\leqslant |z|^2 + 2|z\bar{w}| + |w|^2\end{aligned}$$

und

$$|z\bar{w}|^2 = \operatorname{Re} z\bar{w}^2 + \operatorname{Im} z\bar{w}^2.$$

Damit folgt

$$\begin{aligned} & \operatorname{Re} z\bar{w} = |z\bar{w}| \\ \Leftrightarrow \quad & \operatorname{Re} z\bar{w} \geq 0 \text{ und } \operatorname{Im} z\bar{w} = 0 \quad . \end{aligned}$$

Seien $z,w \neq 0$. Dann folgt damit

$$\begin{aligned} & \frac{z}{w} = \frac{z\bar{w}}{|w|^2} > 0 \\ \Leftrightarrow \quad & z\bar{w} > 0 \\ \Leftrightarrow \quad & \operatorname{Re} z\bar{w} \text{ und } \operatorname{Im} z\bar{w} = 0 \quad . \end{aligned}$$

Dies zeigt die Äquivalenz.

Aufgabe 1.8.16. (a) Der Tipp folgt leicht mit $g - h = 1$. Außerdem gilt

$$z^5 - 1 = (z-1)(z^4 + z^3 + z^2 + z + 1) = (z-1)(z^2 + gz + 1)(z^2 - hz + 1)$$

Die Gleichung $z^2 - hz + 1 = 0$ hat Lösungen $\frac{h}{2} \pm \sqrt{\frac{h^2}{4} - 1}$, wobei $\frac{h^2}{4} - 1 < 0$, also ist

$$\zeta = \frac{1}{2}(h + i\sqrt{4 - h^2})$$

eine Lösung von $z^2 - hz + 1 = 0$, daher auch von $z^5 - 1 = 0$.

(b) Wegen $\zeta^5 = 1$ gilt $|\zeta| = 1$ und $|\zeta^k| = |\zeta|^k = 1$ für alle $k \in \mathbb{N}$. Die angegebenen Punkte liegen also auf dem Einheitskreis. Sie sind paarweise verschieden: $\zeta \neq 1 \Rightarrow \zeta^2 \neq \zeta, \zeta^3 \neq \zeta^2, \zeta^4 \neq \zeta^3$. Wäre z. B. $\zeta^2 = 1 \quad (\Leftrightarrow \zeta^3 = \zeta \Leftrightarrow \zeta^4 = \zeta^2)$ so auch $\zeta^3 = \zeta^{5-2} = 1/\zeta^2 = 1$ also $\zeta = \zeta^3/\zeta^2 = 1$. Wäre $\zeta^3 = 1$, so $\zeta^2 = \zeta^5/\zeta^3 = 1$, Widerspruch.

Das von $1, \zeta, \ldots, \zeta^4$ gebildete Fünfeck ist regelmässig wegen $|\zeta^{k+1} - \zeta^k| = |\zeta - 1||\zeta|^k = |\zeta - 1|$ unabhängig von k. Es gilt:

$$\begin{aligned} \frac{|\zeta^2 - 1|}{|\zeta - 1|} &= |\zeta + 1| \\ &= \sqrt{(1+\zeta)(1+\bar{\zeta})} \\ &= \sqrt{1 + \zeta + \bar{\zeta} + 1} \\ &= \sqrt{2+h} = \sqrt{1+g} = \sqrt{g^2} \\ &= g. \end{aligned}$$

Aufgabe 2.5.2. (a) Wegen $\dfrac{1}{k(k+1)} = \dfrac{1}{k} - \dfrac{1}{k+1}$ hat man die Teleskopsumme

$$\sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1} \right) = 1 - \frac{1}{n+1} \to 1,\ n \to \infty.$$

(b) Für $n \geq k+1$ schreiben wir

$$\frac{n^k}{n!} = \frac{1}{(n-k-1)!} \cdot \frac{n^k}{(n-k) \cdots (n-1) \cdot n}.$$

Der erste Faktor ist nach oben beschränkt durch 1. In der letzten Bruch hat der Nenner $k+1$ Faktoren der Grossenordnung von n und wir werden damit der Zähler n^k kontrollieren. Wir betrachten n so dass $n - k > \frac{n}{2}$ d. h. $n > 2k \geq k+1$ und schätzen ab

$$\begin{aligned} 0 < \frac{n^k}{n!} &= \underbrace{\frac{1}{(n-k-1)!}}_{\leqslant 1} \cdot \frac{n^k}{\underbrace{(n-k)}_{>\frac{n}{2}} \cdots \underbrace{(n-1)}_{>\frac{n}{2}} \cdot \underbrace{n}_{>\frac{n}{2}}} \\ &< \frac{n^k}{\left(\frac{n}{2}\right)^{k+1}} = \frac{2^{k+1}}{n} \end{aligned}$$

und damit folgt die Behauptung.

(c) Für n gross genug ist $\frac{a}{n} < \frac{1}{2}$. Sei daher $k \in \mathbb{N}$, so dass $\frac{a}{k} < \frac{1}{2}$. Für $n > k$ gilt dann

$$\frac{a^n}{n!} = \frac{a^k}{k!} \cdot \prod_{j=k+1}^{n} \frac{a}{j} < \frac{a^k}{k!} \cdot \left(\frac{1}{2}\right)^{n-k} = \frac{(2a)^k}{k!} \cdot \frac{1}{2^n},$$

wobei $\frac{a}{j} < \frac{1}{2}$ wegen $j > k > 2a$. Da $\left(\frac{1}{2^n}\right)_{n \in \mathbb{N}_0}$ eine Nullfolge ist und $\frac{(2a)^k}{k!}$ eine Konstante, folgt die Behauptung.

(d) Man rechnet leicht nach, dass

$$\lim_{n\to\infty}\sqrt{n}\left(\sqrt{n+1}-\sqrt{n}\right)=\lim_{n\to\infty}\frac{\sqrt{n}(n+1-n)}{\sqrt{n+1}+\sqrt{n}}$$
$$=\lim_{n\to\infty}\frac{1}{\sqrt{1+\frac{1}{n}}+1}=\frac{1}{2}.$$

Aufgabe 2.5.3. (a) Beweis mit vollständiger Induktion über $n\in\mathbb{N}$:
(IA) $n=1$: : $x_1=1\in(0,2)$ wahr ✓
Induktionsschritt, $n\rightsquigarrow n+1$:
(IV): Sein $n\in\mathbb{N}$ und es gelte $x_n\in(0,2)$.
Induktionsbehauptung: dann gilt $x_{n+1}\in(0,2)$.
(IS): $x_n\in(0,2)\Rightarrow 2<x_n+2<4\Rightarrow 0<\sqrt{2}<\sqrt{x_n+2}<\sqrt{4}=2\Rightarrow 0<x_{n+1}<2$ ✓
Dabei wurde die Monotonie von $x\to\sqrt{x}$ benutzt.

(b) $x_n\in(0,2)\Rightarrow x_n$ liegt zwischen den Nullstellen von $x^2-x-2=0$ (d. h. zwischen -1 und 2) also $x_n^2-x_n-2<0\Rightarrow x_n^2<x_n+2=x_{n+1}^2$; da x_n, $x_{n+1}>0$ folgt $x_n<x_{n+1}$ für alle $n\in\mathbb{N}$.

(c) $(x_n)_{n\in\mathbb{N}}$ monoton (nach (b)) und beschränkt (nach (a)). Monotonieprinzip 2.1.5 $\rightsquigarrow (x_n)_{n\in\mathbb{N}}$ konvergent. Sei $x=\lim\limits_{n\to\infty}x_n$. Da $x_n>0$ impliziert das Vergleichsprinzip dass $x\geq 0$. Außerdem $x=\lim\limits_{n\to\infty}x_{n+1}$. Durch Limesübergang in $x_{n+1}^2=2+x_n$ folgt

$$x^2=\lim_{n\to\infty}x_{n+1}^2=\lim_{n\to\infty}(2+x_n)=2+\lim_{n\to\infty}x_n=2+x.$$

Schließlich $x^2=2+x$, also $x^2-x-2=0$, also $x=-1$ oder $x=2$. Da $x\geq 0$, folgt $x=2$.

Aufgabe 2.5.4. (a) Monotonie: Für alle $n\in\mathbb{N}$ gilt $x_n\leq x_{n+1}$:
(IA) $n=1$: $x_1=2<1+\sqrt{2}=x_2$.
(IV) Sei $n\in\mathbb{N}$ und gelte $x_n\leq x_{n+1}$ für dieses n.
(IS) Es gilt $x_{n+2}=1+\sqrt{x_{n+1}}\overset{\text{IV}}{\geq}1+\sqrt{x_n}=x_{n+1}$. Mit dem Induktionsprinzip folgt die Behauptung.
(b) Beschränktheit: Für alle $n\in\mathbb{N}$ gilt $2\leq x_n\leq 4$:
(IA) $n=1$: $2\leq x_1=2\leq 4$.
(IV) Sei $n\in\mathbb{N}$ und gelte $2\leq x_n\leq 4$ für dieses n.
(IS) Es ist $2<1+\sqrt{2}\overset{\text{IV}}{\leq}1+\sqrt{x_n}=x_{n+1}\overset{\text{IV}}{\leq}1+\sqrt{4}=3<4$. Nach dem Induktionsprinzip ist die Behauptung bewiesen.
(c) Nach (a) und (b) existiert wegen des Monotonieprinzips 2.1.5

$$x:=\lim_{n\to\infty}x_n=\lim_{n\to\infty}x_{n+1}=\lim_{n\to\infty}(1+\sqrt{x_n})\overset{\sqrt{}\ \text{stetig}}{\underset{\text{GWS}}{=}}1+\sqrt{x}.$$

Nun ist $x=1+\sqrt{x}\Leftrightarrow x-1=\sqrt{x}\Rightarrow x^2-2x+1=x\Leftrightarrow x^2-3x+1=0\Leftrightarrow x=\frac{3}{2}\pm\sqrt{\frac{9}{4}-1}=\frac{3}{2}\pm\sqrt{\frac{5}{4}}=\frac{1}{2}(3\pm\sqrt{5})$. Da $3-\sqrt{5}<4$ und damit $\frac{1}{2}(3-\sqrt{5})<2$ gilt, aber alle Folgenglieder ≥ 2 sind, muss $x=\frac{1}{2}(3+\sqrt{5})$ der Grenzwert der Folge sein.

Aufgabe 2.5.6. Man zeigt zunächst durch Induktion, dass $x_n>0$ für alle $n\in\mathbb{N}$. Für $n\in\mathbb{N}$ gilt nach der AGM-Ungleichung

$$x_{n+1}=\frac{1}{k}\left((k-1)x_n+\frac{a}{x_n^{k-1}}\right)\geqslant\sqrt[k]{x_n^{k-1}\cdot\frac{a}{x_n^{k-1}}}=\sqrt[k]{a}$$

also

$$x_n\geqslant\sqrt[k]{a} \tag{15.2}$$

für alle $n\geqslant 2$, daher ist die Folge nach untern beschränkt. Weiter kann die Rekursionsformel folgendermaßen geschrieben werden:

$$x_{n+1}=\frac{1}{k}\left((k-1)x_n+\frac{a}{x_n^{k-1}}\right)=x_n-\frac{x_n-a}{kx_n^{k-1}} \tag{15.3}$$

und daher $x_{n+1}\leqslant x_n$ für alle $n\geq 2$. Somit ist die Folge $(x_n)_{n\geq 2}$ monoton fallend. Nach dem Monotoniekriterium ist $(x_n)_{n\in\mathbb{N}}$ konvergent. Sei $x=\lim_{n\to\infty}x_n$. Nach dem Vergleichskriterium und nach (15.2) folgt $x\geqslant\sqrt[k]{a}>0$. Durch Limesübergang in der Rekursionsformel (15.3), so bekommt man

$$x=x-\frac{x^k-a}{kx^{k-1}}\Rightarrow x^k=a.$$

Für k ungerade hat $x^k = a$ die einzige Lösung $x = \sqrt[k]{a}$. Für k gerade hat $x^k = a$ zwei Lösungen, also $x \in \{-\sqrt[k]{a}, \sqrt[k]{a}\}$. Da $x > 0$ folgt aber, dass $x = \sqrt[k]{a}$. Der Grenzwert der Folge $(x_n)_n$ ist also $x = \sqrt[k]{a}$.

Aufgabe 2.5.7. (a) Da $\binom{n}{k}\frac{1}{n^k} \leqslant \frac{1}{k!}$ für alle $k \in \{0,1,\dots,n\}$ es gilt

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

(b) Wir rechnen

$$\begin{aligned}\left(1+\frac{1}{n}\right)^n &= \sum_{k=0}^{n}\binom{n}{k}\frac{1}{n^k} \overset{n>m}{>} \sum_{k=0}^{m}\binom{n}{k}\frac{1}{n^k}\\ &= 1+\frac{1}{1!}+\frac{1}{2!}\frac{n(n-1)}{n\cdot n}+\ldots+\frac{1}{m!}\frac{n(n-1)\cdots(n-m+1)}{n\cdots n}\\ &= 1+\frac{1}{1!}+\frac{1}{2!}\left(1-\frac{1}{n}\right)+\ldots+\frac{1}{m!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{m-1}{n}\right)\end{aligned}$$

für alle $m \in \mathbb{N}$ mit $2 \leqslant m < n$.
(c) Limesübergang auf der rechten Seite von (b) bezüglich n:

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \overset{(b)}{>} 1+\frac{1}{1!}+\frac{1}{2!}+\ldots+\frac{1}{m!} = x_m\,.$$

Limesübergang bezüglich m auf der rechten Seite gibt uns dann

$$e \geqslant \lim_{m\to\infty} x_m\,.$$

Betrachten wir den Limes in der Abschätzung in (a), so bekommen wir

$$e \leqslant \lim_{n\to\infty} x_n$$

also insgesamt $e = \lim_{n\to\infty} x_n$.
Aufgabe 2.5.8. (a) Für $m > n$ gilt

$$\begin{aligned}0 &< x_m - x_m\\ &= \frac{1}{(n+1)!}+\ldots+\frac{1}{m!} = \frac{1}{n!}\cdot\left(\frac{1}{n+1}+\frac{1}{(n+1)(n+2)}+\ldots+\frac{1}{(n+1)(n+2)\cdots m}\right)\\ &\leqslant \frac{1}{n!}\left(\frac{1}{n+1}+\frac{1}{(n+1)^2}+\ldots+\frac{1}{(n+1)^{m-n}}\right)\\ &\overset{\text{geom. Reihe}}{\leqslant} \frac{1}{n!}\cdot\frac{1}{n+1}\cdot\frac{1-\left(\frac{1}{n+1}\right)^{m-n}}{1-\frac{1}{n+1}}\\ &< \frac{1}{n!}\cdot\frac{1}{n+1}\cdot\frac{1}{\frac{n}{n+1}} = \frac{1}{n!n}\,.\end{aligned}$$

(b) Die Aussage folgt direkt aus $\lim_{m\to\infty} x_m = e$.
(c) Da $6!6 = 4320 > 10^3$, haben wir $e - x_6 < \frac{1}{6!6} < \frac{1}{10^3}$ und daher ist $x_6 = 2{,}718\ldots$.
(d) Angenommen $e = \frac{p}{q}$ mit $p, q \in \mathbb{N}$ teilerfremd. Da nach (c) $e \in (2,3)$, ist $q \geq 2$. Setze $n = q$, dann gilt nach (b):

$$0 < \frac{p}{q} - \left(1+\frac{1}{1!}+\ldots+\frac{1}{q!}\right) < \frac{1}{q!q}\,.$$

Multiplikation mit $q!$ gibt

$$0 < p(q-1)! - \left(q! + \frac{q!}{1!}+\ldots+\frac{q!}{q!}\right) < \frac{1}{q} < 1$$

und wir erhalten einen Widerspruch, da $p(q-1)! - \left(q! + \frac{q!}{1!}+\ldots+\frac{q!}{q!}\right) \in \mathbb{Z}$. Also ist e irrational.

Aufgabe 2.5.9. (a) Sei $x_n := \frac{1}{n}$, mit $n \in \mathbb{N}$. Da für alle $n \in \mathbb{N}$: $\frac{1}{n+1} < \frac{1}{n}$, ist die Folge streng monoton fallend. Dann gilt

$$\sup_{k \geqslant n} x_k = \sup\left\{\frac{1}{k} : k \geqslant n\right\} = \frac{1}{n},$$
$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup_{k\geqslant n} x_k = \lim_{n\to\infty} \frac{1}{n} = 0,$$
$$\inf_{k\geqslant n} x_k = \inf\left\{\frac{1}{k} : k \geqslant n\right\} = 0,$$
$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf_{k\geqslant n} x_k = \lim_{n\to\infty} 0 = 0.$$

(b) Sei $(x_n)_{n\in\mathbb{N}}$ eine beschänkte Folge in $\mathbb{R}$.

(i) Sei $C \in \mathbb{R}$ eine obere Schranke von $(x_n)_{n\in\mathbb{N}}$, d.h. $\forall n \in \mathbb{N} : x_n \leqslant C$. Betrachte die Folge $(c_n)_{n\in\mathbb{N}}$, wobei $c_n := \inf_{k\geqslant n} x_k = \inf A_n$, wobei $A_n := \{x_k : k \geqslant n\}$. Da $A_{n+1} \subset A_n$, gilt $c_{n+1} \geqslant c_n$, für alle $n \in \mathbb{N}$, das heißt, $(c_n)_{n\in\mathbb{N}}$ ist monoton wachsend. Außerdem ist $c_n \leqslant x_n \leqslant C$ für alle $n \in \mathbb{N}$. Das Monotonieprinzip zeigt, dass $(c_n)_{n\in\mathbb{N}}$ konvergent ist und

$$\lim_{n\to\infty} c_n = \sup_{n\geqslant 1} c_n = \sup_{n\geqslant 1} \inf_{k\geqslant n} x_k =: \liminf_{n\to\infty} x_n =: c.$$

Wir zeigen, dass c ein Häufungswert ist von $(x_n)_{n\in\mathbb{N}}$. Sei dafür $\varepsilon > 0$. Da $\lim_{n\to\infty} c_n = c$, gibt es ein $n(\varepsilon)$, so dass für alle $n \geqslant n(\varepsilon)$: $c_n \in (c-\varepsilon, c+\varepsilon)$. Insbesondere $\inf A_n = c_n < c+\varepsilon$. Nun ist $\inf A_n$ die größte untere Schranke von A_n, somit gibt es $k = k(n) \geqslant n$ mit $c - \varepsilon < c_n \leqslant x_k < c+\varepsilon$. Insgesamt

$$\exists n(\varepsilon) \forall n \geqslant n(\varepsilon) \exists k = k(n) \geqslant n : x_k \in (c-\varepsilon, c+\varepsilon).$$

Es folgt, dass die Menge aller $k \in \mathbb{N}$, für die $x_k \in (c-\varepsilon, c+\varepsilon)$ gilt, unendlich ist. (Denn wäre sie endlich, dann hätte sie ein Maximum m. Für $n > \max\{n(\varepsilon), m\}$ gibt es aber $k \geqslant n > m$ mit $x_k \in (c-\varepsilon, c+\varepsilon)$. Widerspruch.) Somit ist $c = \liminf_{n\to\infty} x_n$ ein Häufungswert von $(x_n)_{n\in\mathbb{N}}$.
Sei nun $c' < \liminf_{n\to\infty} x_n = c$ und sei $\varepsilon > 0$ mit $c' + \varepsilon < c$. Wegen $\lim_{n\to\infty} c_n = c$ existiert ein $n(\varepsilon)$, so dass für alle $n \geqslant n(\varepsilon)$ gilt $c_n > c - \varepsilon$. Da $c_n = \inf A_n$, gilt $x_n \geqslant c_n > c - \varepsilon$. Es gilt also $\{n \in \mathbb{N} : x_n < c' + \varepsilon\} \subset \{1, 2, \ldots, n(\varepsilon) - 1\}$. Folglich gibt es nur endlich viele $n \in \mathbb{N}$ mit $x_n \in (c' - \varepsilon, c' + \varepsilon)$ und somit ist c' kein Häufungswert von $(x_n)_{n\in\mathbb{N}}$.

(ii) ($\Rightarrow$) Sei $(x_n)_{n\in\mathbb{N}}$ konvergent mit $l = \lim_{n\to\infty} x_n$ und sei $\varepsilon > 0$. Es gibt $n(\varepsilon)$, so dass für alle $n \geqslant n(\varepsilon)$ gilt $x_n \in (l-\varepsilon, l+\varepsilon)$ also $A_n := \{x_k : k \geqslant n\} \subset (l-\varepsilon, l+\varepsilon)$. Folglich sind $\inf A_n, \sup A_n \in [l-\varepsilon, l+\varepsilon]$ und daher gilt

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup A_n \in [l-\varepsilon, l+\varepsilon], \text{ sowie}$$
$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf A_n \in [l-\varepsilon, l+\varepsilon].$$

Da $\varepsilon > 0$ beliebig ist, folgt $\liminf_{n\to\mathbb{N}} x_n = l = \limsup_{n\to\infty} x_n$.

($\Leftarrow$) Sei $l = \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$ und $\varepsilon > 0$. Dann existiert $n(\varepsilon)$, so dass für alle $n \geqslant n(\varepsilon)$ gilt $\sup A_n, \inf A_n \in (l-\varepsilon, l+\varepsilon)$. Wegen $\inf A_n \geqslant x_n \leqslant \sup A_n$ folgt auch $x_n \in (l-\varepsilon, l+\varepsilon)$. Somit ist $(x_n)_{n\in\mathbb{N}}$ konvergent und hat den Grenzwert l.

Aufgabe 2.5.10. Zu $(a_n)_n$: Wir haben $\lim_{n\to\infty} \frac{n+1}{n} = \lim_{n\to\infty} \frac{1}{1+\frac{1}{n}} = 1$ und

$$\begin{cases} \sin(\frac{n\pi}{2}) = 0, & n \in 4\mathbb{Z} \\ \sin(\frac{n\pi}{2}) = 1, & n \in 4\mathbb{Z}+1 \\ \sin(\frac{n\pi}{2}) = 0, & n \in 4\mathbb{Z}+2 \\ \sin(\frac{n\pi}{2}) = -1, & n \in 4\mathbb{Z}+3 \end{cases}$$

Die Teilfolgen $(a_{4n})_n$, $(a_{4n+1})_n$, $(a_{4n+2})_n$, $(a_{4n+3})_n$, sind also konvergent und ihre Grenzwerte sind 0, -1, 1. Somit sind 0, -1, 1 Häufungswerte von $(a_n)_n$. Sei a ein Häufungswert von $(a_n)_n$ und $(a_{n_k})_k$ eine Teilfolge, die gegen a konvergiert. Dann enthält die Menge der Indizes $\{n_k : k \in \mathbb{N}\}$ unendlich viele Indizes aus einer der Mengen $4\mathbb{Z}$, $4\mathbb{Z}+1$, $4\mathbb{Z}+1$, oder $4\mathbb{Z}+3$. Folglich besitzt $(a_{n_k})_k$ eine Teilfolge die gegen 0, -1 oder 1 konvergiert, also gilt $a \in \{0, -1, 1\}$. Die Menge der Häufungswerte von $(a_n)_n$ ist also $\{0, -1, 1\}$.
Zu $(b_n)_n$: Wir haben

$$\begin{cases} n \in 2\mathbb{Z}: & \cos(n\pi) = 1 \Rightarrow b_n = \left(1+\frac{1}{n}\right)^n \\ n \in 2\mathbb{Z}+1: & \cos(n\pi) = -1 \Rightarrow b_n = \left(1-\frac{1}{n}\right)^n \end{cases}$$

Nun gelten $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$ und

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = \lim_{n\to\infty}\left(\frac{n-1}{n}\right)^n = \lim_{n\to\infty}\frac{1}{\left(\frac{n}{n-1}\right)^n} = \lim_{n\to\infty}\frac{1}{\left(1+\frac{1}{n-1}\right)^n} = \frac{1}{e}\cdot$$

Die Teilfolgen $(a_{2n})_n$ und $(a_{2n+1})_n$, sind also konvergent und ihre Grenzwerte sind e und $\frac{1}{e}$. Wie oben zeigt man, dass e und $\frac{1}{e}$ die einzigen Häufungswerte sind.

Aufgabe 2.5.12. (i) **1. Fall:** Sei $(x_{n+1}-x_n)/(y_{n+1}-y_n)\to x\in\mathbb{R}$, $n\to\infty$. Für $\varepsilon>0$ beliebig, gibt es $n_0\in\mathbb{N}$, so dass für $n\geq n_0$

$$x-\varepsilon<\frac{x_{n+1}-x_n}{y_{n+1}-y_n}<x+\varepsilon \quad \text{und } y_n>0\,. \tag{15.4}$$

(y_n) monoton steigend $\rightsquigarrow y_{n+1}-y_n\geq 0$ also (15.4) impliziert

$$(x-\varepsilon)(y_{n+1}-y_n)<x_{n+1}-x_n<(x+\varepsilon)(y_{n+1}-y_n)$$

$$\Rightarrow \sum_{j=n_0}^{n}(x-\varepsilon)(y_{n+1}-y_n)\leq \sum_{j=n_0}^{n}x_{n+1}-x_n<\sum_{j=n_0}^{n}(x+\varepsilon)(y_{n+1}-y_n)$$

$$\Rightarrow (x-\varepsilon)(y_n-y_{n_0})<x_n-x_{n_0}<(x+\varepsilon)(y_n-y_{n_0}) \quad |:y_n$$

$$\Rightarrow (x-\varepsilon)\left(1-\frac{y_{n_0}}{y_n}\right)+\frac{x_{n_0}}{y_n}<\frac{x_n}{y_n}<(x+\varepsilon)\left(1-\frac{y_{n_0}}{y_n}\right)+\frac{x_{n_0}}{y_n}$$

n_0 fest und $y_n\to\infty\Rightarrow x_{n_0}/y_n$, $y_{n_0}/y_n\to 0$. Durch Limesübergang für Teilfolgen von (x_n/y_n) die gegen $\liminf(x_n/y_n)$, $\limsup(x_n/y_n)$ konvergieren,

$$x-\varepsilon\leq\liminf\frac{x_n}{y_n}\leq\limsup\frac{x_n}{y_n}\leq x+\varepsilon$$

ε beliebig $\Rightarrow \liminf(x_n/y_n)=\limsup(x_n/y_n)=x$, d. h. (x_n/y_n) konvergiert und $\lim(x_n/y_n)=x$.

2. Fall: $(x_{n+1}-x_n)/(y_{n+1}-y_n)\to\infty$, $n\to\infty$. Für $M>0$ beliebig, gibt es $n_0=n(M)\in\mathbb{N}$ sodass für $n\geq n_0$ gilt $x_{n+1}-y_n>M(y_{n+1}-y_n)$. Durch Summation von n_0 bis n wie oben,

$$x_n-x_{n_0}>M(y_n-y_{n_0}) \quad |:y_n>0$$

$$\Rightarrow \frac{x_n}{y_n}>\frac{x_{n_0}}{y_n}+M(1-\frac{y_{n_0}}{y_n})\Rightarrow \liminf\frac{x_n}{y_n}\geq M\Rightarrow \lim\frac{x_n}{y_n}=\infty.$$

3. Fall: $(x_{n+1}-x_n)/(y_{n+1}-y_n)\to\infty$, $n\to\infty$ behandelt man analog.

(ii) Für $x_n=a^n$,$y_n=n \rightsquigarrow$

$$\lim\frac{x_{n+1}-x_n}{y_{n+1}-y_n}=\lim a^n(a-1)=\begin{cases}0, & 0\leq a\leq 1\\ \infty, & a>1.\end{cases}$$

$$\Rightarrow \lim\frac{a^n}{n}=\begin{cases}0, & 0\leq a\leq 1\\ \infty, & a>1.\end{cases}$$

Für $x_n=1^p+2^p+\cdots+n^p$, $y_n=n^{p+1}$ gilt

$$\frac{x_{n+1}-x_n}{y_{n+1}-y_n}=\frac{(n+1)^p}{(n+1)^{p+1}-n^{p+1}}=\frac{n^p+\binom{p}{1}n^{p-1}+\cdots}{\binom{p+1}{1}n^p+\binom{p+1}{2}n^{p-1}+\cdots}\to\frac{1}{p+1}\Rightarrow\frac{x_n}{y_n}\to\frac{1}{p+1}.$$

Letztlich $\dfrac{(x_1+x_2+\cdots+x_{n+1})-(x_1+\cdots+x_n)}{(n+1)-n}=x_{n+1}$ also $\lim\dfrac{x_1+\cdots+x_n}{n}=\lim x_n$.

Aufgabe 2.5.13. (i) Wir zeigen $\limsup\sqrt[n]{a_n}\leq\limsup\frac{a_{n+1}}{a_n}=:a$. Falls $a=\infty$ O.K. Falls $a\in\mathbb{R}$ sei $\varepsilon>0$ vorgegeben. Es gibt höchstens endlich viele Glieder von (a_{n+1}/a_n) größer als $a+\varepsilon$, d. h. es gibt ein $n_0\in\mathbb{N}$ mit $\frac{a_{n+1}}{a_n}<a+\varepsilon$ für alle $n\geq n_0$.

$$\Rightarrow a_{n_0+1}<(a+\varepsilon)a_{n_0}, \quad a_{n_0}+2<(a+\varepsilon)a_{n_0+1}<(a+\varepsilon)^2a_{n_0}$$

Durch Induktion $a_{n_0+p}<(a+\varepsilon)^pa_{n_0}$ für $p\in\mathbb{N}$, oder $a_n\leq(a+\varepsilon)^{n-n_0}a_{n_0}$ für $n\geq n_0$. Es folgt:

$$\sqrt[n]{a_n}\leq\sqrt[n]{(a+\varepsilon)^{n-n_0}a_{n_0}}=\sqrt[n]{(a+\varepsilon)^{-n_0}a_{n_0}}(a+\varepsilon).$$

Nun gilt: $\limsup(x_ny_n)\leq\limsup x_n\limsup y_n$, also

$$\limsup\sqrt[n]{a_n}\leq\limsup\sqrt[n]{(a+\varepsilon)^{-n_0}a_{n_0}}(a+\varepsilon)=a+\varepsilon$$

weil $\sqrt[n]{(a+\varepsilon)^{-n_0}a_{n_0}}\to 1, n\to\infty$. Da $\varepsilon>0$ beliebig ist $\rightsquigarrow \limsup\sqrt[n]{a_n}\leqslant a$.

(ii) (a) Wähle $a_n=n!$.

(b) $a_n = \frac{n^n}{n!}$, $\frac{a_{n+1}}{a_n} = \frac{(n+1)^{n+1}}{(n+1)\cancel{!}} \cdot \frac{\cancel{n}!}{n^n} = \frac{(n+1)^n}{n^n} = \left(1+\frac{1}{n}\right)^n \to e,\ n \to \infty.$

(c) $\frac{\sqrt[n+1]{(n+1)!}}{\sqrt[n]{n!}} = \frac{\sqrt[n+1]{(n+1)!}}{n+1} \cdot \frac{n}{\sqrt[n]{n!}} \cdot \frac{n+1}{n} \longrightarrow \frac{1}{e} \cdot e \cdot 1 = 1,\ n \to \infty$ (nach (b)).

Aufgabe 3.5.1.

(a)

$$\sum_{k=0}^{n}\left(-\frac{1}{3}\right)^k = \frac{1-\left(-\frac{1}{3}\right)^{n+1}}{1-\left(-\frac{1}{3}\right)} \to \frac{1}{1-\left(-\frac{1}{3}\right)} = \frac{3}{4}, \quad n \to \infty$$

(b)

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

(c)

$$\sum_{k=1}^{n}\frac{4}{5^k} = 4\cdot\sum_{k=1}^{n}\frac{1}{5^k} = 4\cdot\left(\sum_{k=0}^{n}\left(\frac{1}{5}\right)^k - 1\right) \to 4\cdot\left(\frac{1}{1-\frac{1}{5}} - 1\right) = 1, \quad n \to \infty$$

(d)

$$\begin{aligned}
\sum_{k=1}^{n}\frac{k^2+2k+5}{k!} &= \sum_{k=1}^{n}\frac{k(k-1)+3k+5}{k!} \\
&= \sum_{k=2}^{n}\frac{k(k-1)}{k!} + 3\cdot\sum_{k=1}^{n}\frac{k}{k!} + 5\cdot\sum_{k=1}^{n}\frac{1}{k!} \\
&= \sum_{k=2}^{n}\frac{1}{(k-2)!} + 3\cdot\sum_{k=1}^{n}\frac{1}{(k-1)!} + 5\cdot\sum_{k=1}^{n}\frac{1}{k!} \\
&\to e + 3e + 5(e-1) = 9e - 5, \quad n \to \infty
\end{aligned}$$

da

$$\sum_{k=2}^{n}\frac{1}{(k-2)!} = \sum_{j=0}^{n-2}\frac{1}{j!}, \quad \sum_{k=1}^{n}\frac{1}{(k-1)!} = \sum_{j=0}^{n-1}\frac{1}{j!}, \quad \sum_{k=1}^{n}\frac{1}{k!} = \sum_{k=0}^{n}\frac{1}{k!} - 1.$$

(e)

$$\begin{aligned}
\sum_{k=1}^{\infty}\frac{1}{4k^2-1} &= \sum_{k=1}^{\infty}\frac{1}{(2k-1)(2k+1)} \\
&= \frac{1}{2}\cdot\sum_{k=1}^{\infty}\left(\frac{1}{2k-1} - \frac{1}{2k+1}\right) \\
&= \frac{1}{2}\cdot\left(\frac{1}{2\cdot 1-1} - \frac{1}{2n+1}\right) \\
&\to \frac{1}{2}\cdot\left(1-0\right) = \frac{1}{2}, \quad n \to \infty
\end{aligned}$$

(f)

$$\begin{aligned}
\sum_{k=1}^{\infty}\frac{1}{k(k+1)(k+2)} &= \frac{1}{2}\cdot\sum_{k=1}^{\infty}\frac{1}{k+1}\cdot\left(\frac{1}{k} - \frac{1}{k+2}\right) \\
&= \frac{1}{2}\cdot\left(\frac{1}{2\cdot 1} - \frac{1}{(n+1)(n+2)}\right) \\
&\to \frac{1}{2}\cdot\left(\frac{1}{2} - 0\right) = \frac{1}{4}, \quad n \to \infty
\end{aligned}$$

Aufgabe 3.5.2. (a) Die Reihen $\sum_{n\geqslant 1}\frac{1}{n^2}$ und $\sum_{n\geqslant 1}\frac{(-1)^n}{n}$ sind konvergent, daher auch die Reihe $\sum_{n\geqslant 1}\left(\frac{1}{n^2} + \frac{(-1)^n}{n}\right)$. Sie ist nicht absolut konvergent, da allein die Partialsummen für gerades n größer gleich

$$\frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{n} = \frac{1}{2}\left(1 + \frac{1}{2} + \ldots + \frac{1}{\frac{n}{2}}\right)$$

sind, was unbeschränkt ist.

(b) Es gilt für $n \geqslant 2$

$$\left|\frac{4n^2+2n-3}{3n^4-n^3+7}\right| \leqslant \frac{4n^2+2n+3}{3n^4-n^3-7} = \frac{1}{n^2}\cdot\frac{4+\frac{2}{n}+\frac{3}{n^2}}{3-\frac{1}{n}-\frac{7}{n^4}}.$$

Wegen $\frac{4+\frac{2}{n}+\frac{3}{n^2}}{3-\frac{1}{n}-\frac{7}{n^4}} \to \frac{4}{3}$, $n\to\infty$, es gibt $n_0\in\mathbb{N}$, so dass für alle $n\geqslant n_0$ gilt $\frac{4+\frac{2}{n}+\frac{3}{n^2}}{3-\frac{1}{n}-\frac{7}{n^4}} < 2$, also auch $|a_n| \leqslant \frac{2}{n^2}$. Das Majorantenkriterium impliziert, dass die Reihe Reihe absolut konvergent und damit auch konvergent ist.

(c) Es gilt

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)^{n+1}}{4^{n+1}(n+1)!}\cdot\frac{4^n n!}{n^n} = \frac{(n+1)^n}{4n^n} = \frac{1}{4}\left(1+\frac{1}{n}\right)^n \to \frac{e}{4} < 1$$

daher konvergiert die Reihe nach dem Quotientenkriterium absolut; insbesondere konvergiert die Reihe.

(d) Es gilt

$$0 < \frac{\sqrt{n+1}-\sqrt{n}}{n} = \frac{1}{n(\sqrt{n+1}+\sqrt{n})} < \frac{1}{n\sqrt{n}} = \frac{1}{n^{3/2}}$$

Nach dem Majorantenkriterium ist die Reihe konvergent und absolut konvergent.

Aufgabe 3.5.3. (i) Setze $x_{2n} := \frac{1}{3^n}$ für $n\in\mathbb{N}$ und $x_{2n+1} := \frac{1}{2^{n+1}}$ für $n\in\mathbb{N}_0$. Wir wenden das Wurzelkriterium an:

$$\sqrt[2n]{|x_{2n}|} = \sqrt[2n]{\frac{1}{3^n}} = \frac{1}{\sqrt{3}}, \qquad \sqrt[2n+1]{|x_{2n+1}|} = \left(\frac{1}{2}\right)^{\frac{n+1}{2n+1}}$$

also $\sqrt[2n]{|x_{2n}|} \to \frac{1}{\sqrt{3}}$ und $\sqrt[2n+1]{|x_{2n+1}|} \to \frac{1}{\sqrt{2}}$, $n\to\infty$. Die Folge $(\sqrt[n]{|x_n|})_{n\in\mathbb{N}}$ hat also genau zwei Häufungswert und $\limsup_{n\to\infty}\sqrt[n]{|x_n|} = \frac{1}{\sqrt{2}} < 1$. Nach dem Wurzelkriterium ist die Reihe absolut konvergent.

(ii) Setze $x_{2n} := a^n b^n$ für $n\in\mathbb{N}_0$ und $x_{2n+1} := a^{n+1}b^n$ für $n\in\mathbb{N}_0$. Wir wenden das Wurzelkriterium an:

$$\begin{aligned}\sqrt[2n]{|x_{2n}|} &= \sqrt[2n]{|a^n b^n|} = \sqrt{|ab|},\\ \sqrt[2n+1]{|x_{2n+1}|} &= |a^{n+1}b^n|^{\frac{1}{2n+1}} = |a|^{\frac{n+1}{2n+1}}|b|^{\frac{n}{2n+1}} \to \sqrt{|ab|}.\end{aligned}$$

Daher haben wir $\sqrt[n]{|x_n|} \to \sqrt{|ab|}$, $n\to\infty$. Ist nun $|ab|^{\frac{1}{2}} < 1$, was Äquivalent ist zu $|ab| < 1$, so ist die Reihe absolut konvergent. Ist $|ab| \geqslant 1$, so sind unendlich viele Folgeglieder im Betrag mindestens 1 und die Reihe ist nach dem Wurzelkriterium divergent.

(iii) Wir betrachten die Reihe $\sum_{n\geqslant 1}\frac{z^n n!}{n^n}$ und wenden das Quotientenkriterium mit $a_n := \frac{z^n n!}{n^n}$ an:

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{z(n+1)!n^n}{(n+1)^{n+1}n!}\right| = \frac{|z|}{\left(1+\frac{1}{n}\right)^n} \to \frac{|z|}{e}, n\to\infty.$$

Daraus folgt absolute Konvergenz für $|z|<e$ und Divergenz für $|z|>e$. Für $|z|=e$ ist wegen $(1+\frac{1}{n})^n < e$ auch $\left|\frac{a_{n+1}}{a_n}\right| > 1$ für alle $n\in\mathbb{N}$ und daher ist die Reihe dort divergent.

Aufgabe 3.5.4. Induktion nach $n\in\mathbb{N}_0$.

$n=0:$ Ist $x\in\mathbb{R}_{\geqslant 0}$, dann ist $a_0 := \lfloor x\rfloor \in\mathbb{N}_0$. Wegen $0\leqslant x-a_0<1$ ist $0\leqslant 10(x-a_0)<10 \Rightarrow a_1 := \lfloor 10(x-a_0)\rfloor \in \{0,1,2,\dots,9\}$ und $0\leqslant 10(x-a_0)-a_1 < 1 \Rightarrow 0\leqslant x-a_0-\frac{a_1}{10} < \frac{1}{10}$.

$n\to n+1:$ Nach IV ist $0\leqslant 10^n\left(x-\sum_{k=0}^n\frac{a_k}{10^k}\right) < 1 \Rightarrow 0\leqslant 10^{n+1}\left(x-\sum_{k=0}^n\frac{a_k}{10^k}\right) < 10$. Dann folgt mit $a_{n+1} := \lfloor 10^{n+1}\left(x-\sum_{k=0}^n\frac{a_k}{10^k}\right)\rfloor \in\{0,1,2,\dots,9\}$, dass

$$\begin{aligned} & 0\leqslant 10^{n+1}\left(x-\sum_{k=0}^n\frac{a_k}{10^k}\right)-a_{n+1} < 1\\ \Leftrightarrow\quad & 0\leqslant x-\sum_{k=0}^n\frac{a_k}{10^k}-\frac{a_{n+1}}{10^{n+1}} < \frac{1}{10^{n+1}}.\end{aligned}$$

Zur Eindeutigkeit:

Hat eine weitere Folge $(b_n)_{n\geqslant 0}$ die beiden Eigenschaften (a) und (b), dann folgt $b_0 = \lfloor x\rfloor = a_0$ und durch Induktion, falls $b_0=a_0,\dots,b_{n-1}=a_{n-1}$, folgt

$$\underbrace{0\leqslant 10^n x-\sum_{k=0}^{n-1}a_k 10^{n-k} < 1}_{=:y} \qquad\Rightarrow\qquad y-1<b_n\leqslant y.$$

Im Intervall $(y-1,y]$ befindet sich nur eine ganze Zahl, nämlich $\lfloor y\rfloor$. Daraus folgt $b_n=a_n$.

Aufgabe 3.5.5. Wir setzen $f_0 = 1$, so dass $f_{n+1} = f_n + f_{n-1}$ für alle $n \in \mathbb{N}$ (sonst Induktionanfang auch für $n = 2$ nötig).
(a) $n = 1$: $f_2 - f_1 g = 1 - g = -\frac{1}{g} = (-1)^1 g^{-1}$ (siehe Aufgabe 1.8.16 für die Eigenschaften der Zahl g insbesondere $1 - g = -\frac{1}{g}$).
$n-1 \rightsquigarrow n$: Es gilt

$$\begin{aligned}
f_{n+1} - f_n g &= f_{n-1} + f_n - f_n g \\
&= f_{n-1} + f_n(1-g) \\
&= f_{n-1} - \frac{f_n}{g} \\
&= -\frac{1}{g}(f_n - g f_{n-1}) \\
&\overset{IV}{=} (-g^{-1}) \cdot \left((-1)^{n-1} g^{-(n-1)}\right) \\
&= (-1)^n g^{-n}.
\end{aligned}$$

Daher

$$\left|\frac{f_{n+1}}{f_n} - g\right| = \frac{1}{f_n} |f_{n+1} - g f_n| = \frac{1}{f_n} g^{-n} \le g^{-n} \to 0, \quad n \to \infty,$$

wegen $g > 1$. Somit folgt $\frac{f_{n+1}}{f_n} \to g$, für $n \to \infty$.
(b) $n = 1$: ist klar. $n \rightsquigarrow n+1$: Es gilt

$$x_{n+1} = 1 + \frac{1}{x_n} \overset{IV}{=} 1 + \frac{f_n}{f_{n+1}} = 1 + \frac{f_{n+2} - f_{n+1}}{f_{n+1}} = \frac{f_{n+2}}{f_{n+1}}.$$

Damit folgt $x_n \to g$ für $n \to \infty$ nach (a).
(c) Es gilt

$$\begin{aligned}
\sum_{k=1}^{n} \frac{1}{f_k f_{k+2}} &= \sum_{k=1}^{n} \frac{f_{k+2} - f_k}{f_k f_{k+1} f_{k+2}} \\
&= \sum_{k=1}^{n} \left(\frac{1}{f_k f_{k+1}} - \frac{1}{f_{k+1} f_{k+2}}\right) \\
&= \frac{1}{f_1 f_2} - \frac{1}{f_{n+1} f_{n+2}} \\
&= 1 - \frac{1}{f_{n+1} f_{n+2}} \to 1, \quad n \to \infty.
\end{aligned}$$

Aufgabe 3.5.6. (a) Wegen $e^{|z|} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{|z|^k}{k!}$ gibt es ein $N \in \mathbb{N}$, so dass für alle $n > N$

$$\sum_{k=N+1}^{\infty} \frac{|z|^k}{k!} = e^{|z|} - \sum_{k=0}^{N} \frac{|z|^k}{k!} < \frac{\varepsilon}{3}.$$

Mit der Abschätzung aus 1.8.8 (b) folgt für alle $n \geqslant N+1$

$$\left|\sum_{k=N+1}^{n} \binom{n}{k} \frac{z^k}{n^k}\right| \overset{\Delta}{\leqslant} \sum_{k=N+1}^{n} \binom{n}{k} \frac{|z|^k}{n^k} \leqslant \sum_{k=N+1}^{n} \frac{|z|^k}{k!} < \frac{\varepsilon}{3}.$$

(b) Es gilt

$$\left|\sum_{k=0}^{N} \left(\frac{z^k}{k!} - \binom{n}{k} \frac{z^k}{n^k}\right)\right| \leqslant \sum_{k=0}^{N} \left|\frac{1}{k!} - \binom{n}{k} \frac{1}{n^k}\right| |z|^k = \sum_{k=0}^{N} \frac{|z|^k}{k!} \cdot \left(1 - \underbrace{\frac{n}{n}}_{=1} \cdot \underbrace{\frac{n-1}{n}}_{\to 1} \cdots \underbrace{\frac{n-k+1}{n}}_{\to 1}\right) \to 0.$$

für $n \to \infty$. Es folgt

$$\left|\left(1 + \frac{z}{n}\right)^n - e^z\right| = \left|\sum_{k=0}^{n} \binom{n}{k} \frac{z^k}{n^k} - \sum_{k=0}^{\infty} \frac{z^k}{k!}\right| \overset{\forall n>N}{\leqslant} \left|\sum_{k=N+1}^{n} \binom{n}{k} \frac{z^k}{n^k}\right| + \left|\sum_{k=0}^{N} \left(\binom{n}{k} \frac{z^k}{n^k} - \frac{z^k}{k!}\right)\right| + \left|\sum_{k=N+1}^{\infty} \frac{z^k}{k!}\right|.$$

Sei $\varepsilon > 0$. Nach (a) können wir ein $N_1 \in \mathbb{N}$ finden, so dass je der erste und dritte Summand $< \frac{\varepsilon}{3}$ sind für $n > N_1$. Nach der ersten Rechnung in (b) finden wir ein $N_2 \in \mathbb{N}$, so dass der zweite Summand $< \frac{\varepsilon}{3}$ ist für alle $n > N_2$. Mit $N := \max\{N_1, N_2\}$ folgt die Behauptung.

Aufgabe 3.5.7. (i) Wegen $1_n = n!$ haben wir

$$\sum_{k\geqslant 0} z^k = \sum_{k\geqslant 0}\frac{k!k!}{k!k!}z^k = \sum_{k\geqslant 0}\frac{1_k 1_k}{k! 1_k} z^k F_{1,1,1}(z).$$

(ii) Es gilt $\binom{s}{k} = \frac{1}{k!}s(s-1)\cdots(s-k+1) = (-1)^k\frac{(-s)_k}{k!}$ und daher

$$\sum_{k\geqslant 0}\binom{s}{k}z^k = \sum_{k\geqslant 0}\frac{(-1)^k(-s)_k}{k!}z^k = \sum_{k\geqslant 0}\frac{(-s)_k 1_k}{k! 1_k}(-z)^k = F_{-s,1,1}(-z).$$

(iii) Wegen $2_k = (k+1)!$ gilt

$$\sum_{k\geqslant 0}\frac{(-1)^k}{k+1}z^{k+1} = z\sum_{k\geqslant 0}\frac{1_k 1_k}{k! 2_k}(-z)^k = zF_{1,1,2}(-z).$$

Aufgabe 3.5.8. (i) Sei $\varphi : \mathbb{Q} \to \mathbb{R}$ mit $\varphi(1) = a > 0$ und $\varphi(r+s) = \varphi(r)\varphi(s)$ für alle $r, s \in \mathbb{Q}$. Wir untersuchen schrittweise unterschiedliche Eigenschaften der Funktion, die aus obigen Bedingungen folgen.

- Mit $r = 1$ und $s = 0$ gilt $a = \varphi(1+0) = \varphi(1)\varphi(0) = a\varphi(0)$ Da $a \neq 0$, folgt $\varphi(0) = 1$.
- Ist $s = -r$, dann $1 = \varphi(0) = \varphi(r-r) = \varphi(r)\varphi(-r)$, also $\varphi(r) \neq 0$ und $\varphi(-r) = (\varphi(r))^{-1}$ für alle $r \in \mathbb{Q}$.
- Mein zeigt durch Induktion, dass für alle $r \in \mathbb{Q}$ und alle $n \in \mathbb{N}$ gilt $\varphi(nr) = \varphi(r)^n$. Für $n \in \mathbb{Z}$ mit $n \leqslant 0$ gilt also $\varphi(nr) = (\varphi(-nr))^{-1} = ((\varphi(r))^{-n})^{-1} = (\varphi(r))^n$, also $\varphi(nr) = (\varphi(r))^n$ für alle $n \in \mathbb{Z}$ $(*)$.
- Sei $n \in \mathbb{N}$. Für $r = \frac{1}{n}$ in $(*)$ folgt dann
$$a = \varphi(1) = \varphi(n\cdot\frac{1}{n}) = (\varphi(\frac{1}{n}))^n \qquad \Rightarrow \qquad \varphi(\frac{1}{n}) = a^{\frac{1}{n}} = \sqrt[n]{a}.$$
- Für $m \in \mathbb{Z}$ haben wir daher
$$\varphi(\frac{m}{n}) = (\varphi(\frac{m}{n}))^m = (a^{\frac{1}{n}})^m = a^{\frac{m}{n}},$$
das heißt $\varphi(r) = a^r$ für alle $r \in \mathbb{Q}$.

Dabei haben wir dem Beweis aus der Vorlesung gefolgt, dass $\exp(r) = \exp^r$ für alle $r \in \mathbb{Q}$.

(ii) Sei $\alpha \in \mathbb{R}$ und $B_\alpha(z) = \sum_{n\geqslant 0}\binom{\alpha}{n}z^n$. Wir untersuchen zunächst das Konvergenzverhalten der Reihe für unterschiedliche α und unterscheiden nun die folgenden Fälle:

($\alpha \in \mathbb{N}_0$) Für $n \geqslant \alpha + 1$ ist $\binom{\alpha}{n} = 0$, daher ist $B_\alpha(z) = \sum_{n\geqslant 0}\binom{\alpha}{n}z^n$ $(= (1+z)^\alpha)$ ein Polynom in $z \in \mathbb{C}$. (Als solches hat es den Konvergenzradius ∞.) Ist β eine weitere natürliche Zahl (oder = 0), dann konvergiert $B_{\alpha+\beta}(z)$ ebenfalls absolut für alle $z \in \mathbb{C}$.

($\alpha \notin \mathbb{N}_0$) In diesem Fall ist $\binom{\alpha}{n} \neq 0$ für alle n und es gilt

$$\frac{\binom{\alpha}{n}}{\binom{\alpha}{n+1}} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}\cdot\frac{(n+1)!}{\alpha(\alpha-1)\cdots(\alpha-n+1)(\alpha-n)} = \frac{n+1}{\alpha-n}.$$

Der Konvergenzradius ergibt sich also aus der Formel

$$R = \lim_{n\to\infty}\frac{\left|\binom{\alpha}{n}\right|}{\left|\binom{\alpha}{n+1}\right|} = \lim_{n\to\infty}\frac{|n+1|}{|\alpha-n|} = 1.$$

Die Binomialreihe konvergiert also in $\{|z| < 1\}$ absolut für alle $\alpha \in \mathbb{R}$.

Für Produkte absolut konvergenter Reihen gilt der Satz über das Cauchy-Produkt, das heißt, für alle $\alpha, \beta \in \mathbb{R}$ und $z \in \{z \in \mathbb{C} : |z| < 1\}$ gilt

$$B_\alpha(z)B_\beta(z) = \sum_{n=0}^{\infty}\left(\sum_{0\leqslant k+l=n}\binom{\alpha}{k}\binom{\beta}{l}\right)z^n.$$

Wir zeigen nun für alle $\alpha, \beta \in \mathbb{R}$, dass

$$B_{\alpha+\beta}(z) = \sum_{n=0}^{\infty}\binom{\alpha+\beta}{n}z^n \stackrel{!}{=} \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\binom{\alpha}{k}\binom{\beta}{l}\right)z^n = B_\alpha(z)B_\beta(z)$$

für alle z mit $|z| < 1$. Die Behauptung ergibt sich also aus der Identität

$$P(\alpha,\beta,n) := \binom{\alpha+\beta}{n} - \sum_{k=0}^{n}\binom{\alpha}{k}\binom{\beta}{l} \stackrel{!}{=} 0.$$

Hierbei unterscheiden wir wieder mehrere Fälle:

1.Fall Sei $\alpha, \beta \in \mathbb{N}_0$. Es gilt $(1+z)^{\alpha+\beta} = (1+z)^\alpha(1+z)^\beta$ für alle $z \in \mathbb{C}$. Die Koeffizienten von z^n sind $\binom{\alpha+\beta}{n}$ auf der linken Seite und $\sum_{k=0}^{n}\binom{\alpha}{k}\binom{\beta}{l}$ auf der linken Seite. Durch Koeffizientenvergleich der Polynome sind diese beide Koeffizienten gleich. Die gewünschte Identität golgt also in diesem Fall.

2.Fall Sei $\beta \in \mathbb{N}_0$ fest. Dann ist $\alpha \mapsto P(\alpha,\beta,n)$ für alle $n \in \mathbb{N}_0$ ein Polynom in $\alpha \in \mathbb{R}$, das aufgrund des ersten Falles für alle $\alpha \in \mathbb{N}_0$ verschwindet, für jedes $n \in \mathbb{N}_0$. Ein nicht-triviales Polynom hat aber nur endlich viele Nullstellen, daher ist nur $P(\alpha,\beta,n) = 0$ für alle $\alpha \in \mathbb{R}$ möglich, für jedes $n \in \mathbb{N}_0$.

3. Fall Sei $\alpha \in \mathbb{R}$ fest. Dann ist $\beta \mapsto P(\alpha,\beta,n)$ für jedes $n \in \mathbb{N}_0$ ein Polynom in β und wie eben argumentiert man, dass $P(\alpha,\beta,n) = 0$ für alle $\beta \in \mathbb{N}_0$ und alle $n \in \mathbb{N}_0$ und folglich ist wieder nur $P(\alpha,\beta,n) = 0$ für alle $\beta \in \mathbb{R}$ möglich, für jedes $n \in \mathbb{N}_0$.

Insgesamt folgt für alle $n \in \mathbb{N}_0$ und alle $\alpha,\beta \in \mathbb{R}$, dass $P(\alpha,\beta,n) = 0$ und daher

$$B_{\alpha+\beta}(z) = \sum_{n=0}^{\infty} \binom{\alpha+\beta}{n} z^n = \sum_{n=0}^{\infty} \left(\sum_{0\leqslant k+l=n} \binom{\alpha}{k}\binom{\beta}{l} \right) z^n = B_\alpha(z) B_\beta(z).$$

(iii) Sei nun $x \in \mathbb{R}$ mit $|x| < 1$ und definiere $\varphi_x : \mathbb{R} \to \mathbb{R}$ durch $\varphi_x(\alpha) := B_\alpha(x)$. Es gilt $\varphi_x(1) = \sum_{n=0}^{1} \binom{1}{n} x^n = 1 + x > 0$ und nach (ii) gilt auch $\varphi_x(\alpha+\beta) = \varphi_x(\alpha)\varphi_x(\beta)$ für alle $\alpha,\beta \in \mathbb{R}$. Wendet man (i) an auf die Einschränkung $(\varphi_x)|_{\mathbb{Q}} : \mathbb{Q} \subset \mathbb{R} \to \mathbb{R}$ der Funktion φ_x auf die rationalen Zahlen, so folgt für alle $\alpha \in \mathbb{Q}$, dass

$$B_\alpha(x) = (\varphi_x)|_{\mathbb{Q}}(\alpha) = (1+x)^\alpha.$$

Aufgabe 3.5.11. Mit der Eulersche Formel (3.12), der Funktionalgleichung der Exponentialfunktion (3.5) und mit dem Tipp gilt

$$\begin{aligned}
&\cos(z+w) + i\sin(z+w) = \exp\,(i(z+w))\\
&= \exp(iz)\cdot\exp(iw) = (\cos z + i\sin z)\cdot(\cos w + i\sin w)\\
&= \cos z\cos w - \sin z\sin w + i(\cos z\sin w + \sin z\cos w)
\end{aligned}$$

Analog gilt

$$\begin{aligned}
&\cos(z+w) - i\sin(z+w) = \exp\,(-i(z+w))\\
&= \exp(-iz)\cdot\exp(-iw) = (\cos z - i\sin z)\cdot(\cos w - i\sin w)\\
&= \cos z\cos w - \sin z\sin w - i(\cos z\sin w + \sin z\cos w)
\end{aligned}$$

Hier wurde benutzt, dass $\cos(-z) = \cos z$, $\sin(-z) = -\sin z$. Addiere bzw. subtrahiere die Gleichungen $\rightsquigarrow$ 1. bzw. 2. Behauptung.

(b) Berechne

$$\cos^2 z + \sin^2 z = \frac{1}{4}(\exp(2iz) + 2 + \exp(-2iz)) + \frac{1}{4i^2}(\exp(2iz) - 2 + \exp(-2iz)) = \frac{1}{4}(2+2) = 1.$$

$$\cos 2z \overset{(a)}{=} \cos^2 z - \sin^2 z = 2\cos^2 z - (\cos^2 z + \sin^2 z) = 2\cos^2 z - 1$$

also

$$\cos z = 2\cos^2\frac{z}{2} - 1 \rightsquigarrow \cos^2\frac{z}{2} = \frac{1+\cos z}{2} \rightsquigarrow \sin^2\frac{z}{2} = 1 - \cos^2\frac{z}{2} = \frac{1-\cos z}{2}.$$

(c) Es gilt

$$\begin{aligned}
\cos z &\overset{(a)}{=} \cos\frac{z+w}{2}\cos\frac{z-w}{2} - \sin\frac{z+w}{2}\sin\frac{z-w}{2}\\
\cos w &\overset{(a)}{=} \cos\frac{z+w}{2}\cos\frac{-z+w}{2} - \sin\frac{z+w}{2}\sin\frac{-z+w}{2}\\
&= \cos\frac{z+w}{2}\cos\frac{z-w}{2} + \sin\frac{z+w}{2}\sin\frac{z-w}{2}
\end{aligned}$$

Addiere die Gleichungen $\rightsquigarrow$ 1. Behauptung. Außerdem

$$\begin{aligned}
\sin z &\overset{(a)}{=} \cos\frac{z+w}{2}\sin\frac{z-w}{2} + \sin\frac{z+w}{2}\cos\frac{z-w}{2}\\
\sin w &\overset{(a)}{=} \cos\frac{z+w}{2}\sin\frac{-z+w}{2} + \sin\frac{z+w}{2}\cos\frac{-z+w}{2}\\
&= -\cos\frac{z+w}{2}\sin\frac{z-w}{2} + \sin\frac{z+w}{2}\cos\frac{z-w}{2}
\end{aligned}$$

Addiere die Gleichungen $\rightsquigarrow$ 2. Behauptung.

Aufgabe 4.6.1. (a) Sei $z_0 \in \mathbb{C}$ und $\varepsilon > 0$. Für $z \in \mathbb{C}$ gilt

$$|z^3 - z_0^3| = |z - z_0|\,|z^2 + z\cdot z_0 + z_0^2| \le |z - z_0|\,(|z|^2 + |z|\cdot|z_0| + |z_0|^2) \le |z| - |z_0|(|z| + |z_0|)^2$$

und $|z| = |z-z_0+z_0| \leq |z-z_0|+|z_0|$ also $(|z|+|z_0|^2) \leq (|z-z_0|+2|z_0|)^2$. Für $\delta \leq 1$ gilt $|z-z_0|<\delta \Rightarrow (|z|-|z_0|) < 1 \Rightarrow (|z|+|z_0|^2) < (1+2|z_0|)^2$. Wähle $\delta = \delta(\varepsilon, z_0) := \min\left\{1, \frac{\varepsilon}{(2|z_0|+1)^2}\right\}$. Dann gilt

$$|z^3 - z_0^3| \leq |z-z_0|(1+2|z_0|)^2 < \frac{\varepsilon}{(2|z_0|+1)^2}(2|z_0|+1)^2 = \varepsilon$$

für alle $|z-z_0|<\delta$.

(b) Z. z. $\exists\ \varepsilon_0\ \forall \delta > 0\ \exists z = z(\delta)$mit $|z|<\delta$ und $|f(z)-f(0)| \geq \varepsilon_0$. Wähle $\varepsilon_0 = 1$. Sei $\delta > 0$ beliebig. Wähle $n \in N$ mit $\frac{1}{n} < \delta$ und setze $z(\delta) = \frac{i}{n}$. Dann gilt $|\frac{i}{n} - 0| < \delta$ aber $|f(z)-f(0)| = |f(z)| = \frac{1/n}{|1/n|} = 1 \geq 1$.

(c) Sei $\varepsilon > 0$. Wähle $\delta = \varepsilon$. Für $|z|<\delta$ gilt $|f(z)-f(0)| = |f(z)| = \frac{|\mathrm{Re}\, z|^2}{|z|} \leq \frac{|z|^2}{|z|} = |z| < \varepsilon$ (weil $|\mathrm{Re}\, z| \leq |z|$).

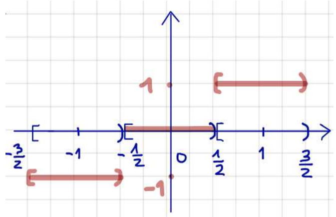
Aufgabe 4.6.2. (a) Die Funktion f ist stetig in 0 und unstetig auf $\mathbb{R}\setminus\{0\}$:
Wir weisen die Stetigkeit von f in 0 mit Hilfe der ε-δ-Definition nach. Sei dazu $\varepsilon > 0$. Wähle $\delta := \min\{\varepsilon, 1\}$. Für $|x| < \delta \leqslant 1$ und $x \in \mathbb{Q}$ gilt $|f(x)-f(0)| = |f(x)| = |x|^2 \leqslant |x| < \varepsilon$ (da $|x| < \delta \leqslant 1 \Rightarrow |x|^2 \leqslant |x|$ und auÃŸerdem $|x| < \delta \leqslant \varepsilon$). Für $|x| < \delta$ mit $c \in \mathbb{R}\setminus\mathbb{Q}$ gilt $|f(x)-f(0)| = |0-0| = 0 < \varepsilon$. Also ist f stetig in 0.
Sei nun $x \in \mathbb{R}\setminus\{0\}$. Ist $x \in \mathbb{Q}$, so ist $f(x) = x^2 \neq 0$. Wähle $(x_n)_n$ in $\mathbb{R}\setminus\mathbb{Q}$. mit $x_n \to x$. Dann gilt $f(x_n) = 0$ für alle $n \in \mathbb{N}$, also $f(x_n) \not\to f(x) = x^2 \neq 0$. Somit ist f nicht stetig in x. Ist $x \in \mathbb{R}\setminus\mathbb{Q}$ und $x \neq 0$, so ist $f(x) = 0$ und für eine Folge $(x_n)_n$ in $\mathbb{Q}$ mit $x_n \to x$ gilt $f(x_n) = x_n^2 \to x^2 \neq 0 = f(x)$. Daher ist f nicht stetig in x.

(b) Die Funktion g ist stetig auf $\mathbb{R}$. Für $x \neq 0$ gilt die Potenzreihendarstellung (Reihendarstellung des Cosinus):

$$g(x) = \sum_{n=1}^{\infty} (-1)^n \frac{x^{2n-1}}{(2n)!} = -\frac{x}{2!} + \frac{x^3}{4!} - \ldots =: P(x),$$

wobei die Potenzreihe auf $\mathbb{R}$ konvergiert. Sie ist insbesondere stetig auf $\mathbb{R}$ und $P(0) = 0 = g(0)$. Insgesamt gilt $g(x) = P(x)$ für alle $x \in \mathbb{R}$ und damit ist g stetig auf ganz $\mathbb{R}$.

(c)



Es gilt $\varphi(x) = n$ für $x \in [n-\frac{1}{2}, n+\frac{1}{2})$, φ ist also stetig auf $\mathbb{R}\setminus(\mathbb{Z}+\frac{1}{2})$ als eine auf $(n-\frac{1}{2}, n+\frac{1}{2})$ konstante Funktion. In $\mathbb{Z}+\frac{1}{2}$ ist sie unstetig: Für alle $k \geqslant 1$ gilt $\varphi(n-\frac{1}{2}-\frac{1}{k}) = n-1$, also

$$\lim_{k\to\infty} \varphi\left(n - \frac{1}{2} - \frac{1}{k}\right) = n-1 \neq n = \varphi\left(n - \frac{1}{2}\right).$$

(d) Für $x \in [n-\frac{1}{2}, n+\frac{1}{2})$ ist $\psi(x) = |n-x|$. Die Funktion ψ ist stetig auf $\mathbb{R}\setminus(\mathbb{Z}+\frac{1}{2})$, da $\psi|_{(n-\frac{1}{2},n+\frac{1}{2})}$ für alle $n \in \mathbb{Z}$ mit der Einschränkung der stetigen Funktion $x \mapsto |n-x|$ auf jedes Intervall $(n-\frac{1}{2}, n+\frac{1}{2})$ übereinstimmt. Stetigkeit von ψ gilt auch in $\mathbb{Z}+\frac{1}{2}$: Es gilt $\psi(n-\frac{1}{2}) = |n-(n-\frac{1}{2})| = \frac{1}{2}$. Für $x \in [n-\frac{1}{2}, n)$ gilt

$$\left|\psi(x) - \psi\left(n - \frac{1}{2}\right)\right| = \left||x-n| - \frac{1}{2}\right| = \left|n - x - \frac{1}{2}\right|.$$

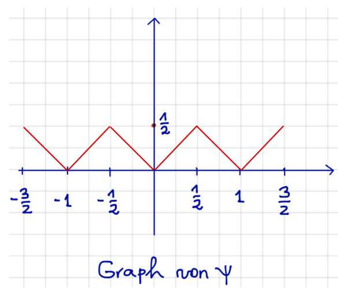
Für $x \in [n-1, n-\frac{1}{2})$ gilt

$$\left|\psi(x) - \psi\left(n - \frac{1}{2}\right)\right| = \left||x-(n-1)| - \frac{1}{2}\right| = \left|x - n + 1 - \frac{1}{2}\right| = \left|x - n + \frac{1}{2}\right| = \left|n - x - \frac{1}{2}\right|$$

Sei nun $\varepsilon > 0$. Wähle $\delta := \min\{\varepsilon, \frac{1}{2}\}$. Gilt $\left|x - \left(n - \frac{1}{2}\right)\right| < \delta$, so ist $x \in (n-1, n)$ also nach obiger Rechnung

$$\left|\psi(x) - \psi\left(n - \frac{1}{2}\right)\right| = \left|n - x - \frac{1}{2}\right| < \delta \leqslant \varepsilon.$$

Somit ist ψ stetig in $n - \frac{1}{2}$ für alle $n \in \mathbb{Z}$. Insgesamt ist ψ also stetig auf $\mathbb{R}\setminus\left(\mathbb{Z}+\frac{1}{2}\right) \cup \underbrace{\left(\mathbb{Z} - \frac{1}{2}\right)}_{=\mathbb{Z}+\frac{1}{2}} = \mathbb{R}$.



Alternative Lösungswege (für Stetigkeit in $\mathbb{Z}+\frac{1}{2}$):
Man kann auch einseitige Grenzwerte errechnen und vergleichen:

$$\lim_{x\nearrow n-\frac{1}{2}} \psi(x) = \lim_{x\nearrow\frac{1}{2}} |x-n+1| = \left|-\frac{1}{2}+1\right| = \frac{1}{2} = \psi(n-\frac{1}{2}),$$

$$\lim_{x\searrow n-\frac{1}{2}} \psi(x) = \lim_{x\searrow\frac{1}{2}} |x-n| = \left|-\frac{1}{2}\right| = \frac{1}{2} = \psi(n-\frac{1}{2}).$$

Da diese übereinstimmen, folgt, dass ψ in $n-\frac{1}{2}$ stetig ist, für alle $n \in \mathbb{Z}$. Oder:

$$\lim_{x\to n-\frac{1}{2}} |\psi(x)-\psi(n-\frac{1}{2})| = \lim_{x\to n-\frac{1}{2}} \left|n-x-\frac{1}{2}\right| = 0,$$

also $\lim_{x\to n-\frac{1}{2}} \psi(x) = \psi(n-\frac{1}{2})$.

Aufgabe 4.6.3. (a) Ist $z \in \mathbb{C}$, so gilt $\exp(z)\exp(-z) = \exp 0 = 1 \neq 0$ und daher $\exp(z) \neq 0$ für alle $z \in \mathbb{C}$. Für $x \in \mathbb{R}$ gilt daher $\exp(x) = \exp\left(\frac{x}{2}+\frac{x}{2}\right) = \exp\left(\frac{x}{2}\right)^2 > 0$ und falls $x > 0$, ist die Folge $\left(\sum_{k=0}^n \frac{x^k}{k!}\right)_{n\in\mathbb{N}}$ streng monoton steigend, also gilt $\exp(x) > 1$. Seien nun $x, y \in \mathbb{R}$ mit $x < y$. Dann gilt

$$\exp(y) = \exp(y-x+x) = \underbrace{\exp(y-x)}_{>1}\underbrace{\exp(x)}_{>0} > \exp(x).$$

(b) Sei $u \in \mathbb{C}$ mit $|u| < 1$. Dann gilt

$$\begin{aligned}|\exp(u)-1-u| &\leqslant \sum_{k=2}^{\infty} \frac{1}{k!}|u|^k \leqslant \sum_{k=2}^{\infty} \frac{1}{2^{k-1}}|u|^k \\ &= 2\sum_{k=2}^{\infty}\left(\frac{|u|}{2}\right)^k \\ &\overset{|u|<1}{=} 2\left(\frac{1}{1-\frac{|u|}{2}}-1-\frac{|u|}{2}\right) = \frac{|u|^2}{2-|u|} \\ &\overset{|u|<1}{<} |u|^2.\end{aligned}$$

Daraus folgt mit $u := w-z$, so dass $|w-z| < 1$:

$$\begin{aligned}&|\exp(w-z)-1-(w-z)| < |w-z|^2 \\ \Rightarrow\quad &|\exp(w)-\exp(z)-(w-z)\exp(z)| < |w-z|^2|\exp(z)|.\end{aligned}$$

(c) Falls $|w-z| < 1$ $(*)$, so folgt

$$\begin{aligned}|\exp(w)-\exp(z)| &\leqslant |\exp(w)-\exp(z)-(w-z)\exp(z)| + |(w-z)\exp(z)| \\ &\overset{(*)}{<} (|w-z|^2+|w-z|)|\exp(z)| \\ &\overset{(*)}{\leqslant} 2|w-z||\exp(z)|.\end{aligned}$$

Für $\varepsilon > 0$ gilt daher $|\exp(w)-\exp(z)| < \varepsilon$ für $|w-z| < \delta := \min\left\{1, \frac{\varepsilon}{2\exp(z)}\right\}$.

Aufgabe 4.6.4. (a) Die Funktionalgleichung der Exponentialfunktion, $\exp(z+w) = \exp(z)\cdot\exp(w)$, impliziert durch Induktion $\exp(nz) = \exp(z)^n$ für $n \in \mathbb{N}$. Daraus folgt mit $nz = w$: $\exp(w) = \exp(\frac{1}{n}w)^n$. Ist $w \in \mathbb{R}$, so $\exp(w) > 0$ und $\exp(\frac{1}{n}w)$ ist die n-te Wurzel von $\exp(w)$: $\exp(\frac{1}{n}w) = \sqrt[n]{\exp(w)}$. Also $\exp(\frac{m}{n}w) = \exp(\frac{1}{n}w)^m = \sqrt[n]{\exp(w)^m}$. Für $w = \log x \rightsquigarrow x^{\frac{m}{n}} = \sqrt[n]{\exp(\log x)}^m = (\sqrt[n]{x})^m$ und das ist genau die Definition 1.5.9.

(b) Es gilt

$$(xy)^z = \exp(z\log(xy)) = \exp(z\log x + z\log y) = \exp(z\log x)\cdot\exp(z\log y) = x^z\cdot y^z$$

also

$$x^{z+w} = \exp[(z+w)\log x] = \exp(z\log x + w\log x) = \exp(z\log x)\exp(w\log x) = x^z x^w$$

und

$$(x^a)^z = \exp(z\log(x^a)) \underset{(*)}{=} \exp(za\log x) = x^{az}$$

wobei $(*): x^a = \exp(a\log x)$ also $\log(x^a) = a\log x$. Außerdem

$$x > 1 \Rightarrow \log(x) > \log(1) = 0 \Rightarrow b\log(x) > a\log(x) \Rightarrow \exp(b\log x) > \exp(a\log x) \Rightarrow x^b > x^a$$

und

$$x < y \Rightarrow \log x < \log y \overset{a>0}{\Rightarrow} a\log x < a\log y \Rightarrow \exp(a\log x) < \exp(a\log y) \Rightarrow x^a < y^a.$$

Aufgabe 4.6.5.

- $\lim_{x\to 0}\log x$: Um den Limes zu erraten, betrachte eine partikuläre Folge $x_n = e^{-n} \to 0$, $n \to \infty$. Dann ist $\log x_n = -n \to -\infty$. Wir zeigen, dass $\lim_{x\to 0}\log x = -\infty$ mit der Umgebungs-Definition: Zu zeigen ist also:
$$\forall C \in \mathbb{R}\, \exists c > 0\, \forall x \in (0,c) : \log x \in (-\infty, C).$$
Wegen $\log x < C \Leftrightarrow x < e^C$ gilt also die Aussage mit $c := e^C$.
- $\lim_{x\to\infty}\log x$: Für $x_n = e^n \to \infty$, $n \to \infty$ gilt $\log x_n = n \to \infty$. Wir vermuten daher, dass $\lim_{x\to\infty}\log x = \infty$. Zu zeigen ist:
$$\forall C \in \mathbb{R}\, \exists c \in \mathbb{R}\, \forall x \in (c,\infty) : \log x \in (C,\infty).$$
Wegen $\log x > C \Leftrightarrow x > e^C$ gilt also die Aussage mit $c := e^C$.
- $\lim_{x\to 0}(1+x)^{\frac{1}{x}}$: Wir benutzen das Folgenkriterium. Sei (x_n) eine Folge reeller Zahlen mit $x_n \to 0$, $n \to \infty$. Wir schreiben
$$(1+x_n)^{\frac{1}{x_n}} = e^{\log(1+x_n)^{\frac{1}{x_n}}} = e^{\frac{\log 1+x_n}{x_n}}.$$
Sei $y_n := \log(1+x_n)$, also $x_n = e^{y_n} - 1$. Es gilt $y_n \to \log 1 = 0$ aufgrund der Stetigkeit der Logarithmusfunktion, also
$$\lim_{n\to\infty} \frac{\log(1+x_n)}{x_n} = \lim_{n\to\infty} \frac{y_n}{e^{y_n}-1} = 1,$$
da $\lim_{y\to 0}\frac{y}{e^y-1} = 1$. Wegen der Stetigkeit der Exponentialfunktion folgt
$$(1+x_n)^{\frac{1}{x_n}} = e^{\frac{\log 1+x_n}{x_n}} \to e^1 = e\,, n \to \infty.$$
Somit ist $\lim_{x\to 0}(1+x)^{\frac{1}{x}} = e$.
- $\lim_{y\to\infty}(1+\frac{1}{y})^y$: Sei (y_n) eine Folge reeller Zahlen mit $y_n \to \infty$, $n \to \infty$. Für fast alle $n \in \mathbb{N}$ gilt $y_n \neq 0$. Wir können daher ObdA annehmen, dass $y_n \neq 0$ für alle $n \in \mathbb{N}$. Sei $x_n := \frac{1}{y_n} \to 0$. Es gilt
$$\left(1+\frac{1}{y_n}\right)^{y_n} = (1+x_n)^{\frac{1}{x_n}} \to e$$
für $n \to \infty$ nach der vorigen Aufgabe. Somit gilt
$$\lim_{y\to\infty}\left(1+\tfrac{1}{y}\right)^y = e.$$
- $\lim_{x\to\infty} a^x$: Wir schreiben $a^x = e^{\log a^x} = e^{x\log a}$. Sei (x_n) eine Folge mit $x_n \to \infty$ für $n \to \infty$. Dann gilt
$$\lim_{n\to\infty} x_n \log a = \begin{cases} \infty & \text{, falls } \log a > 0 \Leftrightarrow a > 1 \\ 0 & \text{, falls } \log a = 0 \Leftrightarrow a = 1 \\ -\infty & \text{, falls } \log a < 0 \Leftrightarrow 0 < a < 1. \end{cases}$$
Wegen $\lim_{y\to\infty} e^y = \infty$ und $\lim_{y\to-\infty} e^y = 0$ folgt mit $y_n := x_n \log a$,
$$\lim_{n\to\infty} e^{x_n \log a} = \lim_{n\to\infty} e^{y_n} = \begin{cases} \infty & \text{, falls } a > 1 \\ 1 & \text{, falls } a = 1 \\ 0 & \text{, falls } 0 < a < 1. \end{cases}$$
Somit haben wir insgesamt
$$\lim_{x\to\infty} a^x = \begin{cases} \infty & \text{, falls } a > 1 \\ 1 & \text{, falls } a = 1 \\ 0 & \text{, falls } 0 < a < 1. \end{cases}$$
- $\lim_{x\to-\infty} a^x$: Sei (x_n) eine Folge mit $x_n \to -\infty$ für $n \to \infty$. Dann gilt $y_n := -x_n \to \infty$ für $n \to \infty$. Mit $\lim_{n\to\infty} a^{x_n} = \lim_{n\to\infty} a^{-y_n} = \lim_{y\to\infty}\left(\frac{1}{a}\right)^y$ und der vorigen Aufgabe folgt
$$\lim_{x\to-\infty} a^x = \begin{cases} 0 & \text{, falls } a > 1 \\ 1 & \text{, falls } a = 1 \\ \infty & \text{, falls } 0 < a < 1. \end{cases}$$

- $\lim_{x\to0} x^a$: Sei (x_n) eine Folge reeller Zahlen mit $x_n \to 0$ für $n\to\infty$. Da der Logarithmus nicht für negative reelle Zahlen definiert ist, dürfen wir $x_n > 0$ für alle n annehmen. Wir setzen $x_n^a = e^{\log x_n^a} = e^{a\log x_n} =: e^{y_n}$. Dann haben wir $\lim_{n\to\infty} y_n = -\infty$ und daher $\lim_{n\to\infty} x_n^a = \lim_{n\to\infty} e^{y_n} = 0$. Es folgt $\lim_{x\to0} x^a = \lim_{x\to0} e^{a\log x} = \lim_{y\to-\infty} e^y = 0$.
- $\lim_{x\to\infty} x^a$: Analog wie oben gilt $\lim_{x\to\infty} x^a = \lim_{x\to\infty} e^{a\log x} = \infty$.
- $\lim_{x\to0} x^a \log x$: Sei (x_n) eine Folge reeller Zahlen mit $x_n\to0$ für $n\to\infty$ und $x_n>0$ für alle n; setze $y_n := -\log x_n$, also $x_n = e^{-y_n}$ und $x_n^a \log x_n = -\frac{y_n}{e^{ay_n}}$. Wegen $y_n\to\infty$ für $n\to\infty$ gilt $-\frac{y_n}{e^{ay_n}}\to0$ für $n\to\infty$. Es folgt $\lim_{x\to0} x^a\log x = 0$.
- $\lim_{x\to0}\frac{\log x}{x^a}$: Mit $y:=\log x\to-\infty$ für $x\to0$ folgt wie oben $\lim_{x\to0}\frac{\log x}{x^a} = \lim_{y\to-\infty}\frac{y}{e^{ay}} = \infty$.
- $\lim_{z\to\infty}(a_nz^n+\ldots+a_0)$: Sei $P(z) = a_nz^n+\ldots+a_0$ mit $a_n\neq0$. Ist $n=0$, so gilt $P(z)=a_0$ für alle $z\in\mathbb{C}$, also $\lim_{z\to\infty}P(z)=a_0$. Für $n\geqslant1$ schreibe $P(z) = z^n\left(a_n+\frac{a_{n-1}}{z}+\ldots+\frac{a_0}{z^n}\right)$. Da $\lim_{z\to\infty}\left(a_n+\frac{a_{n-1}}{z}+\ldots+\frac{a_0}{z^n}\right) = a_n\neq0$ und $\lim_{z\to\infty}z^n=\infty$, folgt $\lim_{z\to\infty}P(z)=\infty$.

Aufgabe 4.6.7. (a) Sei $z\in\bigcup_{\lambda\in\Lambda}\Omega_\lambda$. Dann gibt es ein $\lambda_0\in\Lambda$ mit $z\in\Omega_{\lambda_0}$. Da Ω_{λ_0} offen ist, existiert ein $\varepsilon>0$, so dass $B_\varepsilon(z)\subset\Omega_{\lambda_0}\subset\bigcup_{\lambda\in\Lambda}\Omega_\lambda$.

Sei nun $(\Omega_k)_{k=1}^n$ eine endliche Familie offener Mengen. Sei $z\in\bigcap_{k=1}^n\Omega_k$. Da Ω_k offen, gibt es $\varepsilon_k>0$, so dass $B_{\varepsilon_k}(z)\subset\Omega_k$ für alle k. Setze $\varepsilon:=\min\{\varepsilon_1,\ldots,\varepsilon_n\}>0$. Es folgt $B_\varepsilon(z)\subset\bigcap_{k=1}^n\Omega_k$.

Es gelten duale Aussagen über abgeschlossene Mengen. Nämlich, der Dirchschnitt einer beliebigen Familie abgeschlossener Mengen ist abgeschlossen und die Vereinigung einer endlicher Familie abgeschlossener Mengen ist abgeschlossen. Dies sieht man, indem man in den obigen Argumenten $A_\lambda := \complement\Omega_\lambda$ und damit $\bigcap_{\lambda\in\Lambda}A_\lambda = \complement(\bigcup_{\lambda\in\Lambda}\Omega_\lambda)$ und $\bigcup_{k=1}^nA_k = \complement(\bigcap_{k=1}^n\Omega_k)$ betrachtet.

(b) $(\Rightarrow)$: Angenommen, es gibt $(a_n)_n$ in A, so dass $a_n\to a$, wobei $a\in\complement A =: \Omega$. Da Ω offen ist, existiert ein $\varepsilon>0$, so dass $B_\varepsilon(a)\subset\Omega$ und fast alle a_n liegen in $B_\varepsilon(a)$. Widerspruch zu $a_n\in A$.
$(\Leftarrow)$: Sei $a\in\Omega:=\complement A$. Angenommen, für alle $\varepsilon>0$ gilt $B_\varepsilon(a)\cap A\neq\emptyset$. Dann wähle $a_1\in B_1(a)\cap A, a_2\in B_{\frac12}(a)\cap A,\ldots,a_n\in B_{\frac1n}(a)\cap A$. Es folgt $a_n\to a$ und $a_n\in A$ für alle $n\in\mathbb{N}$ und wegen $a\notin A$ bekommen wir einen Widerspruch zur Voraussetzung.

Aufgabe 4.6.8. Abgeschlossenheit von $\overline{D}$: Sei $z\in\complement\overline{D}$, also $z\notin D$ und z ist kein Häufungspunkt von D. Dann gibt es ein $\varepsilon>0$, so dass $B_\varepsilon(z)$ enthält höchstens endlich viele Punkte von D. Ersetze ε durch das Minimum der Abstände von z zu diesen Punkten. Mit diesem neuen ε enthält $B_\varepsilon(z)$ keinen Punkt von D. Es folgt für alle $w\in B_{\varepsilon/2}(z)$: $B_{\varepsilon/2}(z)\subset B_\varepsilon(z)$ enthält keinen Punkt von D, also $w\notin D$ und w ist kein Häufungspunkt von $D\Rightarrow\complement\overline{D}$. Also gilt $B_{\varepsilon/2}(z)\subset\complement\overline{D}$, daher ist $\complement\overline{D}$ offen und $\overline{D}$ abgeschlossen.
Minimalität von $\overline{D}$: Sei A abgeschlossen mit $D\subset A$. Sei z ein Häufungspunkt von D und $(z_n)_n$ eine Folge in D mit $z_n\longrightarrow z$, für $n\longrightarrow\infty$. Wegen $z_n\in A$ für alle n und der Abgeschlossenheit von A muss dann auch $z\in A$ gelten. Also folgt $\overline{D}\subset A$. □

Aufgabe 4.6.9. (a) Sei $z_0\in D$ und $\varepsilon>0$ vorgegeben. Wähle eine Umgebung U von z_0, so dass $f|_U$ stetig ist. Dann gibt es $\delta>0$ mit $f(B_\delta(z_0)\cap U)\subset B_\varepsilon(f(z_0))$. Da U eine Umgebung von z_0 ist, gibt es $\delta'>0$ mit $B_{\delta'}(z_0)\cap D\subset U$ also auch ein $\delta'\in(0,\delta)$ mit dieser Eigenschaft. Dann $B_{\delta'}(z_0)\cap D = B_{\delta'}(z_0)\cap U\subset B_\delta(z_0)\cap U$, also $f(B_{\delta'}(z_0)\cap D)\subset f(B_\delta(z_0)\cap U)\subset B_\varepsilon(f(z_0))$.

(b) Für jedes $z\in D$ existiert j mit $z\in D_j$. Da D_j offen ist, so ist sie eine Umgebung von z. Nach Voraussetung ist $f|_{D_j}$ stetig. Die aussage folgt also aus (a).

(c) Nein. Gegenbeispiel: Seien $D_1 := \{z:\mathrm{Re}(z)>0\}$ und $f_1: D_1\to\mathbb{C}$, $f_1(z)=1$, sowie $D_2:=\{z:\mathrm{Re}(z)\le0\}$ und $f_2:D_2\to\mathbb{C}$, $f_2(z):=0$. Da $D_1\cap D_2=\emptyset$ gilt $f_1=f_2$ auf $D_1\cap D_2$. Sei f wie in der Aufgabenstellung. Dann ist D_1 offen aber D_2 abgeschlossen und f hat Unstetigkeitsstellen, beispielsweise in $z=0$, denn $f(\frac1n)=1\not\to0=f(0)$. □

Aufgabe 4.6.10. Sei $z_0\in\bigcup_{k=1}^mD_k =: D$ und $\varepsilon>0$ gegeben. Weil $f_{|D_k}$ stetig ist, gibt es zu jedem k mit $z_0\in D_k$ ein $\delta_k>0$, so dass $f(B_{\delta_k}(z_0)\cap D_k)\subset B_\varepsilon(f(z_0))$. Für diejenigen k mit $z_0\notin D_k$ gibt es, weil D_k abgeschlossen ist, ein $\delta_k>0$ mit $B_{\delta_k}(z_0)\cap D_k=\emptyset$, also ebenfalls $f(B_{\delta_k}(z_0)\cap D_k)\subset B_\varepsilon(f(z_0))$. Setze nun $\delta=\min\{\delta_1,\ldots,\delta_k\}$. Dann folgt

$$f(B_\delta(z_0)\cap D) = f\left(\bigcup_{k=1}^m(B_\delta(z_0)\cap D_k)\right)\subset f\left(\bigcup_{k=1}^m(B_{\delta_k}(z_0)\cap D_k)\right)\subset B_\varepsilon(f(z_0))$$

und damit die Behauptung. Falls nicht alle D_k abgeschlossen sind, funktioniert das Gegenbeispiel aus Aufgabe 4.6.9 (c).

Aufgabe 4.6.11. Sei $x\in[a,b]$ beliebig. Es gibt eine Folge (q_n) in $\mathbb{Q}\cap[a,b]$ mit $q_n\to x$, $n\to\infty$. Da f und g stetig (also auch folgenstetig) sind, gilt $f(x)=\lim_{n\to\infty}f(q_n)=\lim_{n\to\infty}g(q_n)=g(x)$.

Aufgabe 4.6.12. Benutze Satz 3.3.8 oder Übung 3.5.8 (i) und Übung 4.6.11.

Aufgabe 4.6.13. (a) Sei $m > 0$ und $\alpha \in [-m, m]$. Es gilt

$$\begin{aligned}\left|\binom{\alpha}{k}\right| &= \left|\frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}\right| \\ &\leqslant \frac{|\alpha|(|\alpha|+1)\cdots(|\alpha|+k-1)}{k!} \\ &\leqslant \frac{m(m+1)\cdots(m+k-1)}{k!} \\ &= \frac{(m+k-1)!}{(m-1)!k!} = \binom{m+k-1}{k}.\end{aligned}$$

Weil

$$\left|\frac{\binom{m+k}{k+1}}{\binom{m+k-1}{k}}\right| = \frac{k+m}{k+1} \to 1, \quad k \to \infty,$$

ist der Konvergenzradius der Potenzreihe $\sum_{k\geqslant 0}\binom{m+k-1}{k}|x|^k$ gleich 1.

(b) Sei $m > 0$. Es gilt

$$\sup_{\alpha\in[-m,m]} |f_k(\alpha)| \leqslant \binom{m+k-1}{k}|x|^k$$

und die Reihe aus (a) konvergent ist, ist die Reihe $\sum_{k\geqslant 0} f_k$ normal konvergent. Da $\alpha \mapsto f_k(\alpha)$ stetig ist, ist $\alpha \mapsto B_\alpha(x)\sum_{k\geqslant 0}\binom{\alpha}{k}x^k$ stetig.

(c) Betrachte die Funktionen $[-m,m] \ni \alpha \overset{f}{\mapsto} (1+x)^\alpha$ und $[-m,m] \ni \alpha \overset{g}{\mapsto} B_\alpha(x)$, für $x \in (-1,1)$ fest. Beide Funktionen sind stetig und nach Aufgabe 3.5.8 gilt $f(\alpha) = g(\alpha)$ für alle $\alpha \in [-m,m] \cap \mathbb{Q}$. Nach Aufgabe 4.6.11 folgt, dass $f(\alpha) = g(\alpha)$ für alle $\alpha \in [-m,m]$. Nun kann man m beliebig groß wählen, also haben wir $(1+x)^\alpha = B_\alpha(x)$ für alle $\alpha \in \mathbb{R}$.

Aufgabe 4.6.14. (a) Für $t \in \mathbb{R}$ gilt $|e^{it}| = |\exp(it)| = |\cos t + i\sin t| = \sqrt{\cos^2 t + \sin^2 t} = 1$.

(b) $|n^s| = n^{\operatorname{Re} s} > n^{s'}$ für $n > 1$, wenn $s' \in \mathbb{Q}$ mit $1 < s' < \operatorname{Re} s$. Wir wissen schon, dass die Reihe $\sum_{n\geqslant 1} n^{-s'}$ konvergiert. Sie ist auch eine Majorante für $\sum_{n\geqslant 1} n^{-s}$. Letztere ist also absolut konvergent.

(c) Es gilt

$$s \in D_\varepsilon \Longrightarrow |n^s| = n^{\operatorname{Re} s} \geqslant n^{1+\varepsilon} \Longrightarrow \|f_n\|_{D_\varepsilon} \leqslant \frac{1}{n^{1+\varepsilon}}$$

also $\sum_{n\geqslant 1}\|f_n\|_{D_\varepsilon}$ hat die konvergente Majorante $\sum_{n\geqslant 1}\frac{1}{n^{1+\varepsilon}}$. Somit ist $\sum_{n\geqslant 1}\|f_n\|_{D_\varepsilon}$ konvergent, d. h. $\sum_{n\geqslant 1} f_n|_{D_\varepsilon}$ konvergiert normal. Aus dem Satz 4.2.6 folgt, dass $\sum_{n=1}^\infty f_n|_{D_\varepsilon}$ ist stetig, und somit ist auch $\zeta|_{D_\varepsilon}$ stetig.

Für $s \in \mathbb{C}$ mit $\operatorname{Re} s > 1$ ist $B_{\frac{\operatorname{Re} s-1}{2}}(s) \subset D_{\frac{\operatorname{Re} s-1}{2}}$; folglich ist $\zeta|_{B_{\frac{\operatorname{Re} s-1}{2}}}$ stetig, also ζ ist stetig in s.

Aufgabe 4.6.15. Wir konstruieren eine Intervallschachtelung $I_n = [a_n, b_n]$ mit $f(a_n) \leq y \leq f(b_n)$ und $b_n - a_n = \frac{b-a}{2^n}$. Setze $I_0 = [a,b]$. Sei I_n gegeben, $c_n = \frac{a_n+b_n}{2}$. Definiere

$$I_{n+1} = \begin{cases} [a_n, c_n] & \text{falls } y \leq f(c_n) \\ [c_n, b_n] & \text{sonst} \end{cases}$$

Sei $x = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$. Dann $f(x) = \lim\limits_{n\to\infty} f(a_n) = \lim\limits_{n\to\infty} f(b_n) \Rightarrow f(x) \leq y$ und $f(x) \geq y$.

Aufgabe 4.6.18. Stetigkeit in (a) und (b) ist klar (da Einschränkungen stetiger Funktionen stetig sind- Satz 4.1.3).

(a) Wir zeigen, dass $\sinh : \mathbb{R} \to \mathbb{R}$ streng monoton wachsend ist: Da $\exp : \mathbb{R} \to \mathbb{R}_+$ streng monoton wachsend ist (Satz 4.3.4), gilt für $x < y$:

$$\sinh(x) = \frac{1}{2}\big(\exp(x) - \exp(-x)\big) < \frac{1}{2}\big(\exp(y) - \exp(-y)\big) = \sinh(y)$$

(es ist $\exp(x) < \exp(y)$, $\exp(-x) > \exp(-y)$). Insbesondere ist $\sinh : \mathbb{R} \to \mathbb{R}$ injektiv. Wir zeigen, dass $\sinh : \mathbb{R} \to \mathbb{R}$ surjektiv ist; sei $y \in \mathbb{R}$. Weil $\exp : \mathbb{R} \to \mathbb{R}_+$ surjektiv ist und streng monoton wachsend, gibt es $x_1 > 0$ mit $\frac{1}{2}(\exp(x_1) - 1) > |y|$ also auch

$$\sinh(x_1) = \frac{1}{2}\big(\exp(x_1) - \underbrace{\exp(-x_1)}_{<1}\big) > |y|$$

und $\sinh(-x_1) = -\sinh(x_1) < -|y|$.

Zwischenwertsatz 4.3.1 $\rightsquigarrow$ es gibt $x \in [-x_1, x_1]$ mit $\sinh(x) = y$.

(b) Für $x \geq 0$ gilt

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots \geqslant 1$$

also ist das Bild von $\cosh|_{[0,\infty)}$ in $[1,\infty]$ enthalten. Es gilt $\cosh 0 = 1$. Sei $y \in (1,\infty)$ und wähle $x_1 > 0$ mit $\exp(x_1) > 2y$. Dann auch $\cosh(x_1) = \frac{1}{2}(\exp(x_1) + \exp(-x_1)) > \frac{1}{2}\exp(x_1) > y$. Zwischenwertsatz 4.3.1 $\rightsquigarrow$ es gibt $x \in (0, x_1)$ mit $\cosh x = y \rightsquigarrow$ das Bild von $\cosh|_{(0,\infty)}$ ist $(1,\infty)$.

Wir zeigen, dass $\cosh : [0,\infty] \to [1,\infty]$ streng monoton wachsend ist. Vorüberlegungen: $1 \leq u < v \Rightarrow u + \frac{1}{u} < v + \frac{1}{v}$, denn $v + \frac{1}{v} - u - \frac{1}{u} = v - u - \frac{v-u}{uv} < v - u - \frac{v-u}{1} = 0$.

Also für $0 \leq x < y : 1 \leq \exp(x) < \exp(y)$ und

$$\cosh(x) = \frac{1}{2}\left(\exp(x) + \frac{1}{\exp(x)}\right) < \frac{1}{2}\left(\exp(y) + \frac{1}{\exp(y)}\right) = \cosh(y).$$

Andere Möglichkeit: Benutze $\sinh(x) = \frac{1}{2}(\exp(x) - \exp(-x)) > 0$ für $x > 0$ also $\sinh^2|_{[0,\infty]}$ streng monoton wachsend und $\cosh^2(x) = 1 + \sinh^2 x$.

Aufgabe 4.6.19. (a) Es gilt $p(0) = a_0 < 0$. Für $x \geq 1$ ist

$$p(x) \geqslant a_n x^n - |a_{n-1}|x^{n-1} - \ldots - |a_0| \geqslant a_n x^n - (|a_{n-1}| + \ldots + |a_0|)x^{n-1} > 0\,, \text{ falls } x > \frac{|a_{n-1}| + \ldots + |a_0|}{a_n}\,.$$

Insbesondere gibt es $x_1 > 0$ mit $p(x_1) > 0$. Da p stetig ist, folgt mit dem Zwischenwertsatz: Es existiert $x \in [0, x_1]$, $p(x) = 0$. Dieses ist > 0 wegen $p(0) < 0$.

(b) Sei $f : [0,\infty] \to \mathbb{R}$, $f(x) = \exp(x) - x^2 - 2$. Es gilt $f(0) = -1 < 0$. Für $x > 0$ ist $f(x) > \frac{x^3}{6} - x^2 - 2 > 0$ falls $x > \max\{1, \frac{1+2}{1/6}\}$. Insbesondere gibt es $x_1 > 0$ mit $f(x_1) > 0$. Da f stetig ist, folgt mit dem Zwischenwertsatz, dass es $x \in [0, x_1]$ existiert, mit $f(x) = 0$. Wegen $f(0) < 0$ ist $x > 0$.

Aufgabe 4.6.21. Sei o.B.d.A. f monoton wachsend. Falls f monoton fallend, betrachte man $g := -f$.

(a) Die Menge $\{f(y) \mid y \in (a,x)\}$ ist von $f(x)$ nach oben beschränkt. Damit existiert $l := \sup\{f(y) \mid y \in (a,x)\}$ und $l \leqslant f(x)$. Wir werden zeigen, daß $l = f(x-)$.

Sei (x_n) eine in (a,b) gegen $x \in (a,b)$ konvergente Folge für die gilt $\forall n \in \mathbb{N} : x_n < x$.

Behauptung: $\big(f(x_n)\big)_{n\in\mathbb{N}}$ konvergiert gegen l.

Beweis: Sei $\varepsilon > 0$ beliebig. Nach der Definition des Supremums gibt es ein $y \in (a,x)$ mit $l - \varepsilon < f(y)$. Da (x_n) konvergiert gegen x, gibt es ein $n(\varepsilon)$ ab dem für alle $n \geq n(\varepsilon)$ gilt $y < x_n < x$. Aufgrund der Monotonie von f gilt dann auch $f(y) \leq f(x_n)$, und da $l - \varepsilon < f(y)$ gilt auch $l - \varepsilon < f(x_n)$. Außerdem ist $f(x_n) \leq l$, damit natürlich $f(x_n) < l + \varepsilon$, also folgt:

$$-\varepsilon < f(x_n) - l < \varepsilon \iff |f(x_n) - l| < \varepsilon$$

Damit konvergiert $(f(x_n))$ gegen $l =: f(x-)$. Der Beweis für $f(x+) = \inf\{f(y) \mid y \in (x,b)\}$ verläuft analog.

(b) Richtung „$\Longrightarrow$": Sei f stetig an der Stelle $x \in (a,b)$. Für jede Forlge (x_n) in (a,b) mit Grenzwert x gilt daher, dass $(f(x_n))$ den Grenzwert $f(x)$ besitzt, so also auch die Spezialfälle davon mit $x_n < x$ bzw. $x_n > x$ für alle n, damit ist $f(x-) = f(x+) = f(x)$.

(b) Richtung „$\Longleftarrow$": Sei $f(x-) = f(x+)$. Nach 2a gilt:

$$f(x-) = \sup\{f(y) \mid y \in (a,x)\}$$

$$f(x+) = \inf\{f(y) \mid y \in (x,b)\}$$

Da f monoton wachsend, gilt damit $f(x-) \leq f(x) \leq f(x+)$, folglich $f(x) = f(x-) = f(x+)$.

Sei (x_n), $x_n \in (a,b)$, konvergent gegen x. Sei $N_< := \{n \in \mathbb{N} \mid x_n < x\}$, $N_> := \{n \in \mathbb{N} \mid x_n > x\}$ und $N_= := \{n \in \mathbb{N} \mid x_n = x\}$. Dann ist $\mathbb{N} = N_< \cup N_> \cup N_=$. Sei $\varepsilon > 0$. Dann existiert es $n_1(\varepsilon)$ derart, dass aus $n > n_1(\varepsilon)$ und $n \in \mathbb{N}_<$,

$$|f(x_n) - f(x)| = |f(x_n) - f(x-)| < \varepsilon\,.$$

folgt. (Die Menge $\{n > n_1(\varepsilon)\} \cap N_<$ könnte leer sein, wenn $N_<$ endlich ist.) Analog existiert es $n_2(\varepsilon)$ derart, daß aus $n > n_2(\varepsilon)$ und $n \in N_>$,

$$|f(x_n) - f(x)| = |f(x_n) - f(x+)| < \varepsilon\,.$$

folgt. (Die Menge $\{n > n_2(\varepsilon)\} \cap N_<$ könnte leer sein, wenn $N_>$ endlich ist.)

Wegen $\mathbb{N} = \mathbb{N}_< \cup \mathbb{N}_> \cup \mathbb{N}_=$ gilt nun $|f(x_n) - f(x)| < \varepsilon$ für $n > \max\{n_1(\varepsilon), n_2(\varepsilon)\}$.

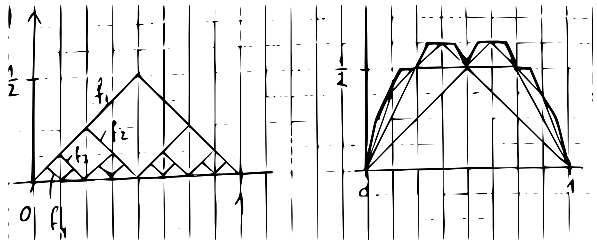
(c) Für jede Unstetigkeitsstelle x von f gilt $f(x-) \neq f(x+)$. Betrachte zu diesen x nun die Intervalle $I_x := \big(f(x-), f(x+)\big)$, welche damit nichtleer sind. Da f monoton, sind diese Intervalle I_x alle disjunkt: Seien $x_1 < x_2$ zwei verschiedene Unstetigkeitstellen und $y \in (x_1, x_2)$, es gilt wegen Monotonie:

$$f(x_1+) = \inf\big\{f(z) \mid z \in (x_1, b)\big\} \leq f(y) \leq \sup\big\{f(z) \mid z \in (a, x_2)\big\} \leq f(x_2-).$$

Damit sind alle $t_1 \in I_{x_1}$ kleiner als alle $t_2 \in I_{x_2}$.

Nach Satz 1.5.5 (Dichtheit von $\mathbb{Q}$ in $\mathbb{R}$) enthält nun aber jedes Intervall I_x eine rationale Zahl q_x, und da alle I_x disjunkt sind, sind alle q_x voneinander verschieden. Die Menge der rationalen Zahlen ist abzählbar, und da jeder Unstetigkeitsstelle eine andere rationale Zahl zugeordnet werden kann (d.h. die Abbildung ist injektiv), ist die Menge der Unstetigkeitsstellen demnach auch abzählbar. $\square$

Aufgabe 5.7.3. (a)



(b) Sei $D := [0,1]$. Für alle k hat man $\|f_k\|_D = \frac{1}{2^k}$. Da $\sum_{k\geqslant 1} \frac{1}{2^k}$ konvergiert, konvergiert nach dem Majorantenkriterium auch $\sum_{k\geqslant 1} \|f_k\|_D$ und nach Satz 4.2.4 konvergiert $\sum_{k\geqslant 1} f_k$ punktweise. Satz 4.2.6 ist ihre Summe ist stetig, da alle f_k stetig sind.

(c) Weil g in x_0 differenzierbar ist, gibt es laut Definition der Differenzierbarkeit eine Funktion $r : I \to \mathbb{R}$, die stetig ist in x_0 und $f(x) = f(x_0) + r(x)(x - x_0)$ für alle $x \in I$ erfüllt (siehe auch Satz 5.6.1). Außerdem gilt $r(x_0) = f(x_0)$. Nun ist

$$\begin{aligned} d_n &= \frac{r(b_n)(b_n - x_0) - r(a_n)(a_n - x_0)}{b_n - a_n} \\ &= \underbrace{\frac{b_n - x_0}{b_n - a_n}}_{\in[0,1]} \cdot r(b_n) + \underbrace{\frac{x_0 - a_n}{b_n - a_n}}_{\in[0,1]} \cdot r(a_n) \begin{cases} \leqslant \max\{r(a_n), r(b_n)\} \\ \geqslant \min\{r(a_n), r(b_n)\} \end{cases} \end{aligned}$$

wegen

$$\frac{b_n - x_0}{b_n - a_n} + \frac{x_0 - a_n}{b_n - a_n} = 1.$$

Damit folgt aber

$$|d_n - f'(x_0)| \leqslant \max\{\underbrace{|r(a_n) - f'(x_0)|}_{\longrightarrow 0}, \underbrace{|r(b_n) - f'(x_0)|}_{\longrightarrow 0}\} \longrightarrow 0,\, n \longrightarrow \infty.$$

(d) Setze

$$\frac{f(b_n) - f(a_n)}{b_n - a_n} = \sum_{k=1}^{\infty} \overbrace{\frac{f_k(b_n) - f_k(a_n)}{b_n - a_n}}^{=: d_{n,k}}.$$

Weil f_n - und erst recht alle f_k mit $k \leqslant n$ - auf $[a_n, b_n]$ linear sind, gilt $d_{n,k} \in \{\pm 1\}$ für alle k mit $1 \leqslant k \leqslant n$. Für $k > n$ ist $f_k(b_n) = f_k(a_n)$ (denn dann sind $a_n, b_n \in \frac{2\mathbb{Z}}{2^k}$), also $d_{n,k} = 0$. Daher gilt

$$\frac{f(b_n) - f(a_n)}{b_n - a_n} = \sum_{k=1}^{n} d_{n,k} = \underbrace{\pm 1 \pm 1 \pm \ldots \pm 1}_{n \text{ Summanden}}.$$

Dies ist gerade, falls n gerade ist und ungerade, falls n ungerade ist. Die Zahlen a_n und b_n erfüllen die Bedingungen aus (a). Wäre f also in x_0 differenzierbar, so müsste die Folge der $\frac{f(b_n)-f(a_n)}{b_n-a_n}$ also konvergieren. Widerspruch. $\square$

Aufgabe 5.7.4. (i) Nach der Quotientenregel gilt

$$f_1'(x) = \frac{a(cx+d) - c(ax+b)}{(cx+d)^2} = \frac{ad - bc}{(cx+d)^2}.$$

(ii) Nach der Produktregel gilt

$$f_2'(x) = 1 \cdot \log(x) + x \cdot \frac{1}{x} = 1 + \log(x)$$

(iii) $f_3(x) = x^x = e^{x\log(x)}$ und damit nach der Kettenregel

$$f_3'(x) = ex\log(x)\cdot(1+\log(x)) = x^x\cdot(1+\log(x))$$

(iv) Nach der Kettenregel gilt

$$f_4'(x) = \frac{1}{\sqrt{1+\sin^2(x)}}\cdot\frac{1}{2}(1+\sin^2(x))^{-\frac{1}{2}}\cdot 2\sin(x)\cos(x) = \frac{\sin(x)\cdot\cos(x)}{1+\sin^2(x)}$$

(v) Nach der Quotientenregel und dem Tipp.

$$f_5'(x) = \frac{\cosh(x)\cdot\cosh(x)-\sinh(x)\cdot\sinh(x)}{\cosh^2(x)} = \frac{1}{\cosh^2(x)}.$$

Zweiter Teil der Aufgabe: Sei $x\in\mathbb{R}\setminus\{0\}$. Dann gilt

$$\frac{f(x)-f(0)}{x-0} = \begin{cases} 0 & , x\leqslant 0 \\ \frac{x^z}{x} = x^{z-1} & , x>0. \end{cases}$$

Daher: f ist in 0 genau dann differenzierbar, wenn $x^{z-1}\longrightarrow 0$ für $x\longrightarrow 0^+$. Äquivalent dazu ist, dass $|x^{z-1}|\longrightarrow 0$ für $x\longrightarrow 0^+$. Es gilt $|x^{z-1}| = |e^{(z-1)\log(x)}| = e^{(\mathrm{Re}(z)-1)\log(x)}$ und für $x\longrightarrow 0^+$:

$$(\mathrm{Re}(z)-1)\log(x)\longrightarrow \begin{cases} -\infty & , \mathrm{Re}(z)>1 \\ 0 & , \mathrm{Re}(z)=1 \\ \infty & , \mathrm{Re}(z)<1 \end{cases}$$

also insgesamt, für $x\longrightarrow 0^+$,

$$e^{(\mathrm{Re}(z)-1)\log(x)}\longrightarrow \begin{cases} 0 & , \mathrm{Re}(z)>1 \\ 1 & , \mathrm{Re}(z)=1 \\ \infty & , \mathrm{Re}(z)<1 \end{cases}$$

und daraus folgt die Behauptung.

Aufgabe 5.7.5. Laut Beispiel 5.2.6 (2) gilt $(x^z)' = zx^{z-1}$ für alle $x\in\mathbb{R}_+^*$ und $z\in\mathbb{C}$. Screibe $f(x) = \frac{\sin x}{\sqrt{x}} = \sin x\cdot x^{-\frac{1}{2}}$. Da $\sin x$ differenzierbar ist und $\sqrt{x}$ differenzierbar auf $(0,\infty]$, ist f differenzierbar und die Ableitung ist (Produktregel):

$$f'(x) = -\frac{1}{2}x^{-\frac{3}{2}}\sin x + x^{-\frac{1}{2}}\cos x$$

Offenbar ist also auch f' differenzierbar, also ist die zweite Ableitung (Produktregel und Linearität):

$$\begin{aligned} f''(x) &= \frac{3}{4}x^{-\frac{5}{2}}\sin x - \frac{1}{2}x^{-\frac{3}{2}}\cos x - \frac{1}{2}x^{-\frac{3}{2}}\cos x - x^{-\frac{1}{2}}\sin x \\ &= \left(\frac{3}{4}x^{-\frac{5}{2}} - x^{-\frac{1}{2}}\right)\sin x - x^{-\frac{3}{2}}\cos x \end{aligned}$$

Setzt man dies nun in den linken Term der Behauptung ein, erhält man:

$$\begin{aligned} & f''(x) + \frac{1}{x}f'(x) + \left(1-\frac{1}{4x^2}\right)f(x) \\ =\ & \left(\frac{3}{4}x^{-\frac{5}{2}} - x^{-\frac{1}{2}}\right)\sin x - x^{-\frac{3}{2}}\cos \\ & + \frac{1}{x}\left(-\frac{1}{2}x^{-\frac{3}{2}}\sin x + x^{-\frac{1}{2}}\cos x\right) \\ & + \left(1-\frac{1}{4x^2}\right)\sin x\cdot x^{-\frac{1}{2}} \\ =\ & \left(\frac{3}{4}x^{-\frac{5}{2}} - x^{-\frac{1}{2}} - \frac{1}{2}x^{-\frac{5}{2}} + x^{-\frac{1}{2}} - \frac{1}{4}x^{-\frac{5}{2}}\right)\sin x \\ & + (-x^{-\frac{3}{2}} + x^{-\frac{3}{2}})\cos x \\ =\ & 0 \end{aligned}$$

Anhang A. Vorbemerkungen zur Aussagenlogik und Mengenlehre

Die großen Fortschritte, die in der Mathematik seit der Antike gemacht worden sind, sind zum wesentlichen Teil mit dadurch bedingt, dass es gelang, einen brauchbaren und leistungsfähigen Formalismus zu finden.

Hilbert&Ackermann

A.1. **Aussagenlogik.** In der Mathematik beschäftigen wir uns mit Aussagen über mathematische Objekte.

A.1.1. **Definition.** Eine ***Aussage*** ist ein sprachlich und grammatisch richtiger Ausdruck, von dem eindeutig feststeht, ob er wahr (richtig), bezeichnet w, oder falsch, bezeichnet f, ist. Wahre Aussagen nennen wir ***Sätze***. Wahr und falsch heißen auch ***Wahrheitswerte***.

A.1.2. **Beispiel.** A: „3 ist eine gerade Zahl.“ f
B: „Es gibt unendlich viele Primzahlen.“ w (Satz von Euklid)

Die Klasse aller Aussagen zerfällt also in zwei disjunkte Teilklassen, die Klasse der wahren und die Klasse der falschen Aussagen. Die Zweiwertigkeit besagt dabei lediglich, dass man von jeder Aussage *entscheiden* kann, ob sie wahr oder falsch ist. Aussagen, deren Wahrheitswert nicht bekannt ist, nennt man *Vermutungen*. So ist z.B. unbekannt, welchen Wahrheitswert die *Goldbachsche Vermutung* hat:
Jede gerade Zahl $\geqslant 4$ *lässt sich als Summe von zwei Primzahlen darstellen.*
Aussagenvariablen stehen für nicht weiter spezifizierte Aussagen, die wahr oder falsch sein können. Wir bezeichnen wahre Aussagen mit $\top$ und falsche Aussagen mit $\bot$.

A.1.3. **Definition** (logische Verknüpfungen, Formel)**.** Aus den Aussagenvariablen werden mit Hilfe von ***logischen Verknüpfungen*** kompliziertere Aussagen, genannt ***Formeln***, aufgebaut. Die logischen Verknüpfungen sind:

Verknüpfung	symbolisch	umgangssprachlich
Verneinung	$\neg$	non ; nicht
Konjunktion	$\wedge$	und
Disjunktion	$\vee$	oder
Implikation	$\Rightarrow$	wenn , dann
Äquivalenz	$\Leftrightarrow$	genau dann, wenn

Im folgenden bezeichnen wir Aussagenvariablen mit $A, B, \ldots$ und definieren die folgenden Formeln:

Formel	symbolisch	umgangssprachlich
Verneinung	$\neg A$	non (A); nicht A
Konjunktion	$A \wedge B$	A und B
Disjunktion	$A \vee B$	A oder B
Implikation	$A \Rightarrow B$	aus A folgt B; wenn A, dann B
Äquivalenz	$A \Leftrightarrow B$	A äquivalent mit B; A genau dann, wenn B

(i) Die Formel $\neg A$ bezeichnet die Negation (das Gegenteil) von A. Wenn A eine wahre Aussage ist, dann ist $\neg A$ eine falsche Aussage; ist A eine falsche Aussage, so ist $\neg A$ eine wahre Aussage.
(ii) Die Formel $A \wedge B$ bezeichnet die Konjunktion der beiden Aussagen A und B. Die Formel $A \wedge B$ ist dann und nur dann wahr, wenn A und B beide wahr sind.
(iii) Die Formel $A \vee B$ bezeichnet die Disjunktion von A und B. Die Formel $A \vee B$ ist dann und nur dann wahr, wenn mindestens eine der beiden Aussagen A und B wahr ist. Insbesondere ist $A \vee B$ wahr, wenn beide Aussagen A und B wahr sind. Man spricht auch vom inklusiven Oder im Gegensatz zum exklusiven Oder.

(iv) Die Formel $A \Rightarrow B$ bezeichnet die Implikation von A und B. Die Formel $A \Rightarrow B$ ist richtig, wenn A falsch (*ex falso quodlibet*) oder A und B richtig sind. Die Formel $A \Rightarrow B$ ist falsch, wenn A richtig und B falsch ist.

(v) Die Formel $A \Leftrightarrow B$ bezeichnet die Äquivalenz von A und B. Die Formel $A \Leftrightarrow B$ ist dann und nur dann richtig, wenn A und B denselben Wahrheitswert haben, d.h. wenn A und B beide richtig oder beide falsch sind.

A.1.4. **Bemerkung.** Eine Formel ist nicht einfach wahr oder falsch. Ihr Wahrheitswert hängt davon ab, ob die Aussagenvariablen, die in der Formel vorkommen, wahr oder falsch sind. Das ist ähnlich wie in der Algebra. Der Ausdruck $(x+y)z$ hat keinen Wert. Erst wenn wir z.B. x gleich 2, y gleich 3 und z gleich 4 setzen, wird der Ausdruck zu $(2+3)4$ und hat den Wert 20.

A.1.5. **Bemerkung.** Wir benutzen nur die Symbole „$\Rightarrow$“ und „$\Leftrightarrow$“. Für „$\wedge$“ und „$\vee$“ benutzen wir „und“, „oder“.

A.1.6. **Bemerkung** (Bedeutung der Implikation)**.** Nehmen wir an, dass der Professor zu den Studenten sagt: „Wenn Sie die Probleme korrekt lösen, bekommen Sie den Schein.“ Falls die Studenten zeitgerecht die Probleme korrekt lösen und den Schein bekommen, hat der Professor die Wahrheit gesagt. Dies ist der Fall „wahr $\Rightarrow$ wahr“. Under welchen Bedingungen würde man sagen, dass der Professor gelogen hat? Ganz klar: wenn die Studenten die Probleme korrekt lösen, aber keinen Schein bekommen. Das ist der Fall „wahr $\Rightarrow$ falsch“. Was kann man sagen, wenn die Studenten die Probleme nicht korrekt lösen? Entweder bekommen sie den Schein (falsch $\Rightarrow$ wahr) oder sie bekommen sie keinen Schein (falsch $\Rightarrow$ falsch). In keinem der beiden Fälle kann man sagen, dass der Professor gelogen hat.

Um Klammern zu sparen, vereinbaren wir folgende Konventionen:

- Das äußerste Paar von Klammern wird weggelassen.
- Die Bindungsstärke der logischen Verknüpfungen nimmt ab in der folgenden Reihenfolge:

$$\text{[bindet stark]} \quad \neg \quad \wedge \quad \vee \quad \Rightarrow \quad \Leftrightarrow \quad \text{[bindet schwach]}$$

- Eine logische Verknüpfung mit niedriger Bindungsstärke hat kleinere Präzedenz.
- Die Konjunktion ($\wedge$) und die Disjunktion ($\vee$) sind linksgeklammert, die Implikation ($\Rightarrow$) ist rechtsgeklammert.

Die nachstehenden Formeln werden daher wie folgt geklammert:

$$\neg A \wedge B \quad \text{ist} \quad (\neg A)\wedge B \quad \text{und nicht} \quad \neg(A \wedge B),$$
$$A \wedge B \Rightarrow C \vee D \quad \text{ist} \quad (A \wedge B) \Rightarrow (C \vee D).$$

A.1.7. **Definition.** Die Zuordnung von Wahrheitswerten wahr oder falsch zu Aussagenvariablen $A, B, \ldots$ heißt eine ***Belegung*** dieser Variablen. Die ***Wahrheitswerttabelle*** einer Formel gibt für alle Belegungen der Eingänge $A, B, \ldots$ die Belegung des Ausgangs an. Eine Wahrheitswerttabelle für eine Formel enthält für jede Aussagenvariable, die in der Formel vorkommt, eine Kolonne, und in den Kolonnen werden alle möglichen Kombinationen von Wahrheitswerten eingetragen.

Aus den Definitionen können wir die folgenden Wahrheitswerttabellen bilden:

A	$\neg A$
w	f
f	w

A	B	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
w	w	w	w	w	w
w	f	f	w	f	f
f	w	f	w	w	f
f	f	f	f	w	w

A.1.8. **Definition.** Eine Formel, deren Wahrheitswert für alle Belegungen der Aussagenvariablen immer falsch ist, heißt ***Kontradiktion***. Eine Formel, deren Wahrheitswert für alle Belegungen der Aussagenvariablen immer wahr ist, heißt ***Tautologie*** oder ***allgemeingültige Formel***. Zwei Formeln $\mathcal{A}$ und $\mathcal{B}$ heißen ***logisch äquivalent***, wenn die Formel $\mathcal{A} \Leftrightarrow \mathcal{B}$ eine Tautologie ist.

A.1.9. **Satz** (Aussagenlogische Gesetze)**.** *Die folgenden Formeln sind Tautologien:*

(A.1) $A \vee \top$

$A \vee \bot \Leftrightarrow A$

(A.2) $A \vee \neg A$ *(Gesetz vom ausgeschlossenen Dritten)*

$\neg(A \wedge \neg A)$ *(Gesetz vom ausgeschlossenen Widerspruch)*

$\neg(\neg A) \Leftrightarrow A$ *(Gesetz von der doppelten Verneinung)*

(A.3) $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$,

$(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$ *(Assoziativgesetze)*

(A.4) $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$,

$A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$ *(Distributivgesetze)*

(A.5) $\neg(A \wedge B) \Leftrightarrow (\neg A \vee \neg B)$,

$\neg(A \vee B) \Leftrightarrow (\neg A \wedge \neg B)$ *(De Morgansche Regeln)*

(A.6) $(A \Rightarrow B) \Leftrightarrow (\neg A \vee B)$

Abschlussregeln:

$A \Rightarrow (A \vee B)$, $A \Rightarrow (B \Rightarrow A)$, $(A \wedge B) \Rightarrow A$ *(Abschwächungsregeln)*

$((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C)$ *(Kettenschlussregel)*

(A.7) $((A \Rightarrow B) \wedge A) \Rightarrow B$ *(Abtrennungsregel, modus ponens)*

(A.8) $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$ *(Kontrapositionsgesetz)*

(A.9) $(A \Rightarrow B) \Leftrightarrow (A \wedge \neg B \Rightarrow \bot)$

$(A \Rightarrow B) \wedge \neg B \Rightarrow \neg A$ *(Gesetz zum modus tollens)*.

Ein Beispiel zu den de Morganschen Regeln: Die Negation von „Ich bin schön *und* ich bin reich“ ist „Ich bin nicht schön *oder* ich bin nicht reich“. Die Aussage $\neg B \Rightarrow \neg A$ heißt Kontraposition von $A \Rightarrow B$.

Beweis: Betrachten wir z.B. die Abtrennungsregel. Wir können sie beweisen, indem wir die Wahrheitswerttabelle erstellen.

A	B	$A \Rightarrow B$	$A \wedge (A \Rightarrow B)$	$(A \wedge (A \Rightarrow B)) \Rightarrow B$
w	w	w	w	w
w	f	f	f	w
f	w	w	f	w
f	f	w	f	w

Da in der Kolonne für die Formel (A.7) immer der Wahrheitswert w steht, ist die Formel eine Tautologie. Eine andere Möglichkeit ist, die Formel mit Hilfe der aussagenlogischen Gesetze zu einer Tautologie zu transformieren. Es ist

$(A \wedge (A \Rightarrow B)) \Rightarrow B$ äq. $(A \wedge (\neg A \vee B)) \Rightarrow B$ (nach (A.6))

$(A \wedge (\neg A \vee B)) \Rightarrow B$ äq. $((A \wedge \neg A) \vee (A \wedge B)) \Rightarrow B$ (nach (A.4))

$((A \wedge \neg A) \vee (A \wedge B)) \Rightarrow B$ äq. $(\bot \vee (A \wedge B)) \Rightarrow B$ (nach (A.2))

$(\bot \vee (A \wedge B)) \Rightarrow B$ äq. $(A \wedge B) \Rightarrow B$ (nach (A.1))

Weil $(A \wedge B) \Rightarrow B$ eine Tautologie (Abschwächungsregel) ist und (A.7) damit äquivalent ist, ist auch (A.7) eine Tautologie.

Wir erstellen die Wahrheitstabelle zum Kontrapositionsgesetz:

A	B	$A \Rightarrow B$	$\neg B$	$\neg A$	$\neg B \Rightarrow \neg A$	$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$
w	w	w	f	f	w	w
w	f	f	w	f	f	w
f	w	w	f	w	w	w
f	f	w	w	w	w	w

Da in der Kolonne für die Formel (A.8) immer der Wahrheitswert w steht, ist die Formel eine Tautologie. □

A.2. **Prädikatenlogik.**

A.2.1. **Definition.** Seien $X_1,\dots,X_k$ Mengen. Eine k-stellige ***Aussageform*** oder Prädikat mit freien Variablen aus $X_1,\dots,X_k$ ist ein sprachlicher Ausdruck $A(x_1,\dots,x_k)$, der endlich viele Variablen $x_1 \in X_1 \dots, x_k \in X_k$ enthält und zu einer Aussage wird, wenn alle Variablen mit Werten belegt werden. Eine Aussage ist per definitionem eine 0-stellige Aussageform.

A.2.2. **Beispiel.**

- $A(x)$: „Es gibt eine rationale Zahl x, für die gilt: $x^2 = 2$." (falsch)
- $B(x,y,z)$: „Für alle reellen Zahlen x, y und z gilt: $(x\cdot y)\cdot z = x\cdot(y\cdot z)$." (wahr)

Eine ***Quantifizierung*** von Variablen erzeugt aus einer Aussageform eine solche niedrigerer Stellenzahl.

Quantor	umgangssprachlich	symbolisch
Generalisierung	für alle x (für jedes x, für beliebiges x)	$\forall x : A(x,\dots)$
Partikularisierung	es gibt (mindestens) ein x (es existiert (mindestens) ein x)	$\exists x : A(x,\dots)$
verstärkte Partikularisierung	es gibt genau ein x (es existiert genau ein x; es gibt ein und nur ein x)	$\exists! x : A(x,\dots)$

A.2.3. **Beispiel.** Obige Aussagen können dann wie folgt geschrieben werden:

- $\exists x : x \in \mathbb{Q} \wedge x^2 = 2$ oder $\exists x \in \mathbb{Q} : x^2 = 2$.
- $\forall x\, \forall y\, \forall z \in \mathbb{R} : (x\cdot y)\cdot z = x\cdot(y\cdot z)$.

Andere Aussagen, die Quantifizierungen enthalten:
Gruppenaxiome:

$$\begin{gathered}\forall x\, \forall y\, \forall z : (x\cdot y)\cdot z = x\cdot(y\cdot z)\\ \exists e\, \forall x : e\cdot x = x\cdot e = x\\ \forall x\, \exists y : x\cdot y = y\cdot x = e\end{gathered}$$

A.2.4. **Satz** (Prädikatenlogische Gesetze.)**.**
Kommutativgesetze

$$\begin{gathered}\forall x\, \forall y : A(x,y,\dots) \Leftrightarrow \forall y\, \forall x : A(x,y,\dots)\\ \exists x\, \exists y : A(x,y,\dots) \Leftrightarrow \exists y\, \exists x : A(x,y,\dots)\\ \forall x : (A(x,\dots) \wedge B(x,\dots)) \Leftrightarrow (\forall x : A(x,\dots)) \wedge (\forall x : (B(x,\dots))\\ \exists x : (A(x,\dots) \vee B(x,\dots)) \Leftrightarrow (\exists x : A(x,\dots)) \vee (\exists x : (B(x,\dots))\end{gathered}$$

De Morgansche Regeln

$$\begin{gathered}\neg \forall x : A(x,\dots) \Leftrightarrow \exists x : \neg A(x,\dots)\\ \neg \exists x : A(x,\dots) \Leftrightarrow \forall x : \neg A(x,\dots)\end{gathered}$$

Abschwächungsregeln

$$\begin{gathered}\exists x\, \forall y : A(x,y,\dots) \Rightarrow \forall y\, \exists x : A(x,y,\dots)\\ \exists! x : A(x,\dots) \Rightarrow \exists x : A(x,\dots)\end{gathered}$$

A.2.5. **Bemerkung.** Um zu zeigen, dass $\forall x : A(x) \Rightarrow B(x)$ falsch ist, zeigen wir, dass die zugehörige Verneinung $\exists x : \neg\big(A(x) \Rightarrow B(x)\big)$ wahr ist. Wir sollten also ein x finden, so dass $A(x) \Rightarrow B(x)$ falsch ist, d.h. mit $A(x)$ wahr und $B(x)$ falsch. Solch ein x heißt ***Gegenbeispiel*** zu $\forall x : A(x) \Rightarrow B(x)$.

A.2.6. **Bemerkung.** Die Reihenfolge der Quantoren ist wesentlich, d.h.

$$\exists x\, \forall y : A(x,y,\dots) \Rightarrow \forall y\, \exists x : A(x,y,\dots) \tag{A.10}$$

ist *falsch*. Zum Beispiel für $A(x,y): y = x^2$ ist

(A.11) $$\forall y \geqslant 0 \, \exists x \in \mathbb{R} : y = x^2$$

wahr und besagt, dass alle nicht-negativen reellen Zahlen eine Quadratwurzel haben. Dagegen ist

(A.12) $$\exists x \in \mathbb{R} \, \forall y \geqslant 0 : y = x^2$$

falsch und besagt, dass es eine reelle Zahl x gibt derart, dass die Quadratwurzel aller nicht-negativen reellen Zahlen x ist! Ähnlich für $A(x,y): x < y$; $\forall y \exists x : A(x,y)$ ist wahr (z.B. $x = y - 1$) aber $\exists x \forall y : A(x,y)$ ist falsch und besagt, dass es eine kleinste reelle Zahl gibt.

Die Aussage $\forall y \exists x : A(x,y)$ bedeutet also, dass für alle y ein von y abhängiges x existiert. Dagegen ist x in $\exists x \forall y : A(x,y)$ fest und hängt von y nicht ab.

A.2.7. **Bemerkung.** Die Quantoren sind nicht einfache Abkürzungen, sondern „sprachliche Elemente der Logik", die syntaktisch richtig benutzt werden sollten. Eine Aussage über x steht *hinter* $\forall x$ bzw. $\exists x$. Nur bei Beachtung dieser Vorschrift kann man obige Regel systematisch zum Negieren von Aussagen benutzen.

A.3. **Beweistechnik.** In der Mathematik werden neue wahre Aussagen (d.h. Sätze) mittels des des logischen Schließens aus bereits als wahr bekannten Aussagen hergeleitet. Häufig erscheint ein Satz in der Form „Wenn A, dann B", d.h. $A \Rightarrow B$ oder $\forall x (A(x) \Rightarrow B(x))$. Dann heißt A (bzw. $A(x)$) die ***Voraussetzung***, und B (bzw. $B(x)$) die ***Behauptung*** des Satzes. Zum Beispiel:

(A.13) Ist $n \in \mathbb{N}$ ungerade, dann ist n^2 ungerade.

Der Satz hat der Form $A \Rightarrow B$, wobei A:"$n \in \mathbb{N}$ ungerade" und B:"n^2 ist ungerade". Wir nehmen an, dass die Voraussetzung A auch die Zusammenfassung (Konjunktion) aller bekannten wahren Aussagen (Sätze) enthält, auch wenn das explizit nicht geschrieben ist. Zum Beispiel gibt es Sätze, in denen nur die Konklusion B explizit vorkommt, etwa:

(A.14) Es gibt unendlich viele Primzahlen.

Dieser Satz gehört genau genommen auch zum Typ $A \Rightarrow B$; dafür bezeichnen wir mit A die Zusammenfassung aller bekannten wahren Aussagen und mit B die Behauptung des Satzes (hier B :"Es gibt unendlich viele Primzahlen".)

Die ***Umkehrung*** des Satzes $A \Rightarrow B$ bzw. $\forall x(A(x) \Rightarrow B(x))$ ist die Aussage $B \Rightarrow A$ oder $\forall x (B(x) \Rightarrow A(x))$. Zum Beispiel ist die Umkehrung des Satzes (A.13):

(A.15) Sei $n \in \mathbb{N}$. Ist n^2 ungerade, dann ist n ungerade.

Ein ***Beweis*** eines Satzes ist eine Herleitung der Konklusion aus aus der Voraussetzung und aus den schon bewiesenen Sätzen vermöge der Abschlussregeln. Ein Beweis kann entweder direkt oder indirekt geführt werden.

Ein ***direkter Beweis*** einer Aussage $A \Rightarrow B$ liegt vor, wenn es endlich vielen Zwischenbehauptungen $A_1, \dots A_n$ gibt, so dass die Implikationen $A \Rightarrow A_1$, $A_1 \Rightarrow A_2$, ... , $A_n \Rightarrow B$ wahr sind. Der Beweis hat die Form einer Schlusskette $A \Rightarrow A_1 \Rightarrow A_2 \Rightarrow \dots \Rightarrow B$.

A.3.1. **Beispiel. Beweis von** (A.13)**:** Definition: $n \in \mathbb{N}$ heißt ungerade genau dann, wenn es $k \in \mathbb{N}_0$ gibt mit $n = 2k+1$.

Ist $n \in \mathbb{N}$ ungerade, so gibt es nach Definition $k \in \mathbb{N}_0$ mit $n = 2k+1$. Dann ist $n^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2+2)+1$ auch von der Form $2m+1$ mit $m = 2k^2+2 \in \mathbb{N}_0$, also nach Definition ungerade. □

Oder formaler:

$$\begin{aligned} n \in \mathbb{N} \text{ ungerade} :\Leftrightarrow\ & \exists k \in \mathbb{N}_0 : n = 2k+1 \\ & \Rightarrow n^2 = (2k+1)^2 = 4k^2+4k+1 = 2(2k^2+2)+1 \\ & \Rightarrow \exists m \in \mathbb{N}_0 : n^2 = 2m+1 \\ & \Leftrightarrow: n^2 \text{ ungerade}. \end{aligned}$$

A.3.2. **Beispiel.** Dies ist ein komplizierteres Beispiel. Zu zeigen ist:

(A.16) Für alle $x, y, z \in \mathbb{R}$ mit $x < y$ und $y \leqslant z$ gilt $x < z$.

Beweis: Formal sieht der Satz so aus: $\forall x,y,z \in \mathbb{R} : (x < y) \wedge (y \leqslant z) \Rightarrow (x < z)$. Die Voraussetzung ist $(x,y,z \in \mathbb{R}) \wedge (x < y) \wedge (y \leqslant z)$ und die Behauptung $(x < z)$. Aus den Voraussetzungen folgt, dass die folgenden Aussagen wahr sind:

$$y \leqslant z :\Leftrightarrow (y < z \vee y = z) \quad \text{(nach Definition)}, \tag{A.17}$$

$$x < y \wedge y < z \Rightarrow x < z \quad \text{(Transitivität)}, \tag{A.18}$$

$$x < y \wedge y = z \Rightarrow x < z\,. \tag{A.19}$$

Nach (A.17) machen wir eine Fallunterscheidung.
Fall 1: Es ist $y < z$: Wir benutzen (A.18) und erhalten $x < z$.
Fall 2: Es ist $y = z$: Wir benutzen (A.19) und erhalten $x < z$.
Also ist in beiden Fällen $x < z$, und die Behauptung ist bewiesen. □

So schreibt man üblicherweise einen mathemathischen Beweis. Die Aussagenlogik taucht nicht explizit auf, wir haben sie jedoch stillschweigend benutzt. Nehmen wir als Besipiel den Fall 1. Für Aussagenvariablen A,B,C gilt die Tautologie

$$\big(A \wedge (B \wedge A \Rightarrow C)\big) \Rightarrow C \tag{A.20}$$

Sei nun A : „$y < z$“, B : „$x < y$“, C : „$x < z$“. Wir wissen nach (A.18), dass $B \wedge A \Rightarrow C$ wahr ist. Wenn A wahr ist, dann ist auch $(A \wedge (B \wedge A \Rightarrow C)$ wahr. Wegen (A.20) ist auch C wahr. Genauso behandelt man den Fall 2.

Wir haben hier (teilweise) die Struktur des Beweises mit Hilfe der Schlussregeln klar gemacht. In den meisten mathematischen Texten ist diese Struktur jedoch nicht explizit angegeben (das würde diese Texte mindestens dreimal so lang und unlesbar machen). Aber für Anfänger ist es sinnvoll, die logische Struktur mancher Beweise zu untersuchen.

A.3.3. *Aufgabe.* (i) Beweisen Sie (A.20). (ii) Geben Sie die logische Struktur des gesamten Beweises von Satz A.3.2 an.

Ein ***indirekter Beweis*** hat zwei Varianten: Beweis durch Kontraposition und Beweis durch Widerspruch (reductio ad absurdum).

Der ***Beweis durch Kontraposition*** (Beweis durch Umkehrschluss) beruht auf dem Kontrapositionsgesetz (A.8). Wenn ein Satz in der Form $A \Rightarrow B$ vorliegt, gewinnen wir einen Beweis des Satzes, indem wir einen direkten Beweis von $\neg B \Rightarrow \neg A$ geben.

A.3.4. **Beispiel. Beweis von** (A.15)**:** Zu zeigen: $A \Rightarrow B$, wobei A : „n^2 ist ungerade“, B : „n ist ungerade“. Dann ist $\neg B$: „n ist gerade“ und $\neg A$: „n^2 ist gerade“. Offensichtlich ist $\neg B \Rightarrow \neg A$ wahr. Also gilt auch $A \Rightarrow B$.

A.3.5. **Beispiel.** Zu beweisen:

A.3.6. **Satz.** *Wenn es in einer Schule* 733 *Schüler gibt, dann gibt es in einem Jahr mindestens* 3 *Schüler, die ihren Geburtstag am gleichen Tag feiern.*

Beweis: Wir beweisen den Satz durch Kontraposition. Wir bezeichnen
A: „In einer Schule gibt es 733 Schüler.“
B: „Es gibt in einem Jahr mindestens 3 Schüler, die ihren Geburtstag am gleichen Tag feiern.“
Zu zeigen: $A \Rightarrow B$. Wir beweisen die logisch äquivalente Aussage $\neg B \Rightarrow \neg A$, d.h. wir nehmen an, $\neg B$ wäre wahr, und leiten daraus einen Widerspruch zu A her.
$\neg B$: „Es gibt höchstens 2 Schüler, die ihren Geburtstag am gleichen Tag feiern.“
$\Rightarrow$ „Es gibt höchstens $2 \times 366 = 732$ Schüler in der Schule.“
$\Rightarrow$ „Es gibt nicht 733 Schüler in der Schule.“: $\neg A$ □

A.3.7. **Beispiel.** Zu beweisen:

A.3.8. **Satz.** *Sei (mindestens) eine von zwei ganzen Zahlen n und m nicht durch* 3 *teilbar. Dann ist auch die Summe oder die Differenz von n und m nicht durch* 3 *teilbar.*

Beweis: Wir beweisen den Satz durch Kontraposition. Um die Kontraposition der Aussage zu bilden, ist es hilfreich, die Implikation als logische Formel zu schreiben:

$$3 \nmid n \vee 3 \nmid m \Rightarrow 3 \nmid (m+n) \vee 3 \nmid (m-n)\,.$$

Unter Benutzung der de Morganschen Regel erhält man die zugehörige Kontraposition:

$$3|(m+n) \wedge 3|(m-n)) \Rightarrow 3|m \wedge 3|n\,.$$

Um diese Implikation für beliebige ganze Zahlen zu zeigen, genügt es den Fall zu betrachten, in dem die linke Seite wahr ist. (Andernfalls ist die Implikation sowieso wahr.) Gelte also $m+n=3k$ und $m-n=3k_0$ mit $k,k_0\in\mathbb{Z}$. Addition der beiden Gleichungen ergibt: $2m=3(k+k_0)$. Die rechte Seite davon muss also durch 2 teilbar sein. Da 2 und 3 teilerfremd sind, ist $k+k_0$ durch 2 teilbar, also $m=3(k+k_0)/2$. Damit ist also m durch 3 teilbar. Subtraktion der Gleichungen ergibt $2n=3(k-k_0)$. Wie oben kann man schließen, dass $n=3(k-k_0)/2$, also ist auch n durch 3 teilbar. □

Der ***Beweis durch Widerspruch*** beruht auf dem Gesetz (A.9). Wenn ein Satz in der Form $A\Rightarrow B$ vorliegt, gewinnen wir einen Beweis des Satzes, indem wir eine Implikation $A\wedge\neg B\Rightarrow\perp$ beweisen, wobei $\perp$ eine falsche Aussage ist. Wir nehmen also $\neg B$ mit als Voraussetzung auf und leiten mit Hilfe der Abschlussregeln daraus einen Widerspruch her, d.h. eine falsche Aussage, z.B. ($x\in M$ und $x\notin M$) oder ($a=b$ und $a\neq b$).

A.3.9. **Beispiel.** Wir beweisen:

A.3.10. **Satz.** *$\sqrt{2}$ ist irrational.*

Beweis: Die Konklusion ist B :“$\sqrt{2}$ ist irrational.“ Das Gegenteil ist $\neg B$:“$\sqrt{2}$ ist rational.“ Die folgende Schlusskette gilt:

$$\begin{aligned} x=\sqrt{2}\text{ ist rational} &:\Longleftrightarrow \exists p,q\in\mathbb{Z}, q\neq 0 : x=p/q\\ &\Longrightarrow \exists p,q\in\mathbb{Z}, q\neq 0 \text{ und } (p,q)=1 : x=p/q\\ &\Longrightarrow \exists p,q\in\mathbb{Z}, q\neq 0 \text{ und } (p,q)=1 : 2=x^2=p^2/q^2\\ &\Longrightarrow \exists p,q\in\mathbb{Z}, q\neq 0 \text{ und } (p,q)=1 : 2p^2=q^2\\ &\Longrightarrow \exists p,q\in\mathbb{Z}, q\neq 0 \text{ und } (p,q)=1 : 2|p \text{ und } 2|q\,. \end{aligned} \tag{A.21}$$

Die letzte Implikation zeigt man so:

$$\begin{aligned} &2p^2=q^2 \Longrightarrow 2|q^2 \Longrightarrow 2|q \Longrightarrow \exists k\in\mathbb{Z}: q=2k \Longrightarrow\\ &q^2=4k^2 \Longrightarrow 4k^2=q^2=2p^2 \Longrightarrow 2k^2=p^2 \Longrightarrow 2|p^2 \Longrightarrow 2|p\\ &\Longrightarrow 2|q \text{ und } 2|p\,. \end{aligned}$$

Die letzte Behauptung von (A.21) ($\exists p,q\in\mathbb{Z}, q\neq 0$ und $(p,q)=1 : 2|p$ und $2|q$) ist offensichtlich falsch. Wir haben also gezeigt, dass $\neg B\Longrightarrow\perp$ wahr ist. Das ist aber nur möglich, wenn $\neg B$ falsch ist (siehe Wahrheitswerttabelle der Implikation), d.h. wenn B wahr ist. □

In der Praxis wird man die Voraussetzung, Behauptung usw. nicht explizit angeben und bezeichnen. Ein normaler Beweis beginnt mit „Nehmen wir an, $x=\sqrt{2}$ wäre rational“ und fährt fort, bis man die falsche Aussage „$\exists p,q\in\mathbb{Z}, q\neq 0$ und $(p,q)=1 : 2|p$ und $2|q$“ hergeleitet hat. Dann schreibt man einfach „Widerspruch“, und der Beweis ist vollbracht!

A.3.11. **Bemerkung** (Was bedeuetet o.B.d.A.?)**.** Anstelle von

$$\begin{aligned} x=\sqrt{2}\text{ ist rational} &: \Longleftrightarrow \exists p,q\in\mathbb{Z}, q\neq 0 : x=p/q\\ &\Longrightarrow \exists p,q\in\mathbb{Z}, q\neq 0 \text{ und } (p,q)=1 : x=p/q \end{aligned}$$

schreibt man einfach „$x=\sqrt{2}$ ist rational $:\Longleftrightarrow \exists p,q\in\mathbb{Z}, q\neq 0 : x=p/q$. O.B.d.A. kann man annehmen, dass $(p,q)=1$“. Mit „***o.B.d.A.***“ ist gemeint, dass eine Einschränkung (z.B. des Wertebereichs einer Variablen) nur zur Vereinfachung der Beweisführung vorausgesetzt wird, ohne dass die Gültigkeit der im Anschluss getroffenen Aussagen in Bezug auf die Allgemeinheit darunter leidet. Man muss sich also klar machen, dass der allgemeine Fall aus dem schon bewiesenen Spezialfall folgt.

In unserem Beispiel muss man die Implikation beweisen:

$$\exists p,q\in\mathbb{Z}, q\neq 0 : x=p/q \Longrightarrow \exists p,q\in\mathbb{Z}, q\neq 0 \text{ und } (p,q)=1 : x=p/q \tag{A.22}$$

Beweis von (A.22): Wir wissen nach der Definition des g.g.T., dass $p=(p,q)p_1$, $q=(p,q)q_1$ wobei $p_1,q_1\in Z$, $q_1\neq 0$ und $(p_1,q_1)=1$. Dann gilt $r=\frac{p}{q}=\frac{(p,q)p_1}{(p,q)q_1}=\frac{p_1}{q_1}$. Mit anderen Worten, wir haben den Bruch $\frac{p}{q}$ vollständig gekürzt. Dann kann man p_1 und q_1 weiter mit p, q bezeichnen.

A.3.12. **Beispiel.** Definition: Eine Primzahl ist eine natürliche Zahl mit genau zwei natürlichen Zahlen als Teiler, nämlich 1 und sich selbst.

Primzahlen sind also $2,3,5,7,11,\dots$

Fundamentalsatz der Arithmetik: Jede natürliche Zahl lässt sich als Produkt von Primzahlen schreiben. Diese Produktdarstellung ist bis auf die Reihenfolge der Faktoren eindeutig.

Wir beweisen:

A.3.13. **Satz** (Euklid). *Es gibt unendlich viele Primzahlen $p \in \mathbb{N}$.*

Beweis: Bezeichnen wir die Menge der Primzahlen von $\mathbb{N}$ mit $P \subset \mathbb{N}$. Nehmen wir an, P wäre endlich, $P = \{p_1,\dots,p_k\}$. Sei $n := p_1 \dots p_k + 1 \in \mathbb{N}$; da $n > p_j$ für alle $j = 1,\dots,k$, ist $n \notin P$. Der Fundamentalsatz der Arithmetik besagt, dass es eine Primzahl $p \in P$ mit $p|n$ gibt. Weil p eine der Zahlen $p_1,\dots,p_k$ ist, folgt $p|p_1 \dots p_k$ und dann $p|n - p_1 \dots p_k = 1$. Wenn aber $p|1$ gilt, muss $p = 1$ sein. Widerspruch zur Definition der Primzahlen. Die Menge P muss also unendlich sein. □

Für weitere Beispiele und eine „Gebrauchsanleitung zur Formulierung mathematischer Gedanken mit vielen praktischen Tipps“ siehe das Buch [3]. Eine Geschichte des Beweisbegriffes findet man auch in "“The History and Concept of Mathematical Proof”".

A.4. **Mengenlehre.** Die Mengenlehre dient heute weitgehend als Grundlage der Mathematik, und viele fundamentale Begriffe und Methoden lassen sich auf die Mengenlehre zurückführen.

Der Begründer der Mengentheorie ist Georg Cantor (1845–1918). Er hat die Mengenvorstellung folgendermaßen ausgedrückt:

„Unter einer „Menge“ verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die „Elemente“ von M genannt werden) zu einem Ganzen“.

Es handelt sich um eine „naive“ Definition. Was bedeutet „Zusammenfassung“ oder „Objekt unserer Anschauung“? Diese Vorstellung führt zu Kontradiktionen, den sogenannten Russellschen Antinomien. Um diese Antinomien auszuschließen, wurden mehrere Wege vorgeschlagen (Russell, Zermelo–Fraenkel, Gödel–Bernays).

Zermelo (1871–1953) hat das erste brauchbare Axiomensystem der Mengenlehre formuliert. Ohne in Details zu gehen, beschreiben wir kurz sein Vorgehen. Wir setzen voraus, dass es zwei Typen von Objekten gibt, *Elemente* und *Mengen*. Diese Objekte können *Eigenschaften* haben oder zueinander in bestimmten *Beziehungen* (*Relationen*) stehen. Die Relation $x = y$ bedeutet, dass die mit x bzw. y bezeichneten Objekte identisch sind; die Negation dieser Relation wird durch $x \neq y$ ausgedrückt. Ist X die Menge, so bedeutet die Relation $x \in X$, dass x ein Element der Menge X ist oder zu X gehört; die Negation dieser Beziehung wird mit $x \notin X$ bezeichnet.

A.4.1. **Definition.** Sind X und Y Mengen, so bedeutet $X \subset Y$, dass jedes Element von X Element von Y ist. Ist $X \subset Y$, so sagt man, X sei eine ***Teilmenge*** von Y oder X sei in Y ***enthalten*** oder Y umfasse X; man schreibt auch $Y \supset X$. Die „$\subset$“–Beziehung heißt ***Inklusion***. Für die Negation von $X \subset Y$ schreibt man $X \not\subset Y$. Wenn X und Y Mengen sind, für die $X \subset Y$ und $X \neq Y$ gilt, so spricht man von einer ***echten*** Teilmenge (X ist ***echt enthalten*** in Y).

Offenbar gilt

$$X \subset X,$$
$$(X \subset Y \,\text{und}\, Y \subset Z) \Longrightarrow X \subset Z$$

A.4.2. **Extensionalitätsaxiom.** *Umfangsgleiche Mengen sind gleich; mit anderen Worten: Zwei Mengen sind genau dann gleich, wenn sie aus denselben Elementen bestehen.*

Daraus folgt

$$(X \subset Y \,\text{und}\, Y \subset X) \Leftrightarrow X = Y$$

Dementsprechend zerfallen fast alle Beweise von Gleichheiten zwischen zwei Mengen X und Y in zwei Teile; zuerst hat man $X \subset Y$ und dann $Y \subset X$ zu zeigen.

Wir geben nun ein wichtiges Mittel zum Bilden von neuen Mengen aus gegebenen Mengen an. Das nächste Axiom besagt, dass eine Aussage über Elemente einer gewissen Menge aus dieser eine Teilmenge aussondert, nämlich die Teilmenge derjenigen Elemente, für die die Aussage wahr ist.

A.4.3. **Aussonderungsaxiom.** *Zu jeder Menge X und jeder Aussageform $E(x)$ mit freien Variablen aus X existiert eine Menge Y, die genau aus den Elementen von X besteht, für die $E(x)$ wahr ist (zutrifft).*

Die Menge Y ist nach dem Extensionalitätsaxiom eindeutig bestimmt. Schreibweise:

$$Y = \{x \in X : E(x) \text{ trifft zu}\}.$$

A.4.4. **Definition.** Sind X, Y zwei Mengen und gilt $Y \subset X$, so ist die Menge $\{x \in X : x \notin Y\}$ eine Teilmenge von X, die sogenannte ***Differenz*** von X und Y oder das ***Komplement*** von Y bezüglich X, in Zeichen $X \setminus Y$ oder $\complement_X Y$ (oder auch $\complement Y$, wenn kein Missverständnis zu befürchten ist).

Sind zwei Mengen X,Y gegeben, so existiert eine Menge, die genau aus denjenigen Elementen besteht, die sowohl zu X als auch zu Y gehören, nämlich $\{x \in X : x \in Y\}$; sie wird ***Durchschnitt*** von X und Y genannt und mit $X \cap Y$ bezeichnet.

Ist X eine Menge und $x \in X$, so bedeutet $\{x\}$ die Menge, deren einziges Element x ist.

Um unserem Vorgehen Substanz zu geben, setzen wir voraus:

A.4.5. **Existenzaxiom.** *Es gibt eine Menge.*

Als logische Folgerung ergibt sich aus diesem Axiom, dass eine Menge ohne irgendein Element existiert. Hat man nämlich eine Menge X, so wende man das Aussonderungsaxiom mit der Aussage „$x \neq x$“. Das Ergebnis ist die Menge

$$\emptyset_X = \{x \in X : x \neq x\}, \tag{A.23}$$

gennant die leere Teilmenge von X. Es ist bemerkenswert, dass jede beliebige Eigenschaft auf Elemente von $\emptyset_X$ zutrifft.

A.4.6. *Aufgabe.* Betrachte die Aussage: “Alle Menschen über 200 Jahre sind Hochleistungssportler.“ Ist diese Aussage wahr oder falsch?

Ist E eine beliebige Eigenschaft, so gilt $x \in \emptyset_X \Rightarrow E(x)$ für jedes $x \in X$, da $x \in \emptyset_X$ eine falsche Aussage ist für jedes $x \in X$.

Sind X und Y Mengen, so zieht $x \in \emptyset_X$ stets $x \in \emptyset_Y$ nach sich; mit anderen Worten, es gilt $\emptyset_X \subset \emptyset_Y$, ebenso $\emptyset_Y \subset \emptyset_X$, also $\emptyset_X = \emptyset_Y$. Alle leeren Teilmengen sind somit einander gleich und werden daher einfach mit $\emptyset$ bezeichnet.

A.4.7. **Definition.** Die Menge $\emptyset$ heißt die ***leere Menge***. Zwei Mengen X und Y heißen ***disjunkt*** falls $X \cap Y = \emptyset$.

Das Paarmengenaxiom stellt sicher, dass jede Menge auch als Element einer Menge vorkommt und dass es zu je zwei Mengen stets eine dritte gibt, in der sie beide als Elemente enthalten sind. Das Axiom hat einen technischen Charakter.

A.4.8. **Paarmengenaxiom.** *Zu je zwei Mengen X, Y gibt es eine Menge Z, welche genau X und Y als Elemente besitzt.*

Wir schreiben $Z = \{X,Y\}$ und für $\{X,X\}$, die Einermenge von X, abkürzend $\{X\}$. Nach dem Extensionalitätsaxiom ist stets $\{X,Y\} = \{Y,X\}$; $\{X,Y\}$ hat also nicht die Eigenschaft eines geordneten Paares von X und Y. Eine Menge Z, deren Elemente Mengen sind, wird Mengensystem gennant.

A.4.9. **Vereinigungsmengenaxiom.** *Zu jedem Mengensystem Z gibt es die Menge Y der Elemente der Elemente von Z.*

A.4.10. **Definition.** Die Menge Y heißt ***Vereinigung*** von Z. Schreibweise:

$$Y = \bigcup_{z \in Z} z = \bigcup \{z : z \in Z\}.$$

Sind X und Y zwei Mengen und $Z = \{X,Y\}$ so wird $\cup\{z : z \in \{X,Y\}\}$ durch $X \cup Y$ bezeichnet. Folglich ist

$$X \cup Y = \{a : a \in X \quad \text{oder} \quad a \in Y\}.$$

Sei $x,y \in X$. Die Vereinigung $\{x\} \cup \{y\}$ wird mit $\{x,y\}$ bezeichnet; ähnlich schreibt man $\{x,y,z\}$ für $\{x\} \cup \{y\} \cup \{z\}$ usw.

A.4.11. **Potenzmengenaxiom.** *Zu jeder Menge X gibt es die Menge Y aller Teilmengen von X, die sogenannte **Potenzmenge** von X; wir bezeichnen diese mit $\mathcal{P}(X)$.*

Offenbar ist $\emptyset \in \mathcal{P}(X)$ sowie $X \in \mathcal{P}(X)$. Die Relationen $x \in X$ und $\{x\} \in \mathcal{P}(X)$ sind äquivalent, ebenso $Y \subset X$ und $Y \in \mathcal{P}(X)$.

Beispiel: Für $X = \{1,2,3\}$ ist $\mathcal{P}(X) = \big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, X\big\}$.

A.4.12. **Definition.** Je zwei Objekten x, y entspricht ein neues Objekt, das ***geordnete Paar*** (x,y); die Relation $(x,y) = (x',y')$ ist äquivalent der Relation ($x = x'$ und $y = y'$); insbesondere gilt $(x,y) = (y,x)$ genau dann, wenn $x = y$ ist. Das erste (bzw. zweite) Element eines geordneten Paares $z = (x,y)$ heißt die ***erste*** (bzw. ***zweite***) ***Projektion*** von z, in Zeichen $x = \mathrm{pr}_1 z$ (bzw. $y = \mathrm{pr}_2 z$).

Sind X, Y zwei (nicht notwendig verschiedene) Mengen, so gibt es eine (eindeutig bestimmte) Menge, deren Elemente genau die geordneten Paare (x,y), $x \in X$ und $y \in Y$ sind. Sie wird das ***kartesische Produkt*** (oder einfach ***Produkt***) ***von*** X ***und*** Y genannt und mit $X \times Y$ bezeichnet.

A.4.13. **Definition.** Seien X und Y Mengen. Eine ***Abbildung*** $f : X \to Y$ von der Menge X in die Menge Y ist eine Vorschrift, die jedem Element $x \in X$ genau ein Element $f(x) \in Y$ zuordnet. Wir schreiben auch $f : X \to Y$, $x \mapsto f(x)$. Der ***Graph*** der Funktion $f : X \to Y$ ist die folgende Teilmenge von $X \times Y$:

$$\mathrm{Graph}(f) = \{(x,y) \in X \times Y \,:\, y = f(x)\}. \tag{A.24}$$

Zwei Abbildungen $f, g : X \to Y$ sind gleich, wenn $f(x) = g(x)$ für alle $x \in X$, d.h. $\mathrm{Graph}(f) = \mathrm{Graph}(g)$.

Der Graph einer Abbildung charakterisiert die Abbildung vollständig. Eigentlich ist es korrekt zu denken, dass die Abbildung ein Graph ist! Eine formalere Definition einer Abbildung verläuft daher so:

A.4.14. **Definition.** Seien X, Y Mengen. Eine Abbildung von X in Y ist eine Teilmenge $G \subset X \times Y$, gennant funktionaler Graph, mit folgender Eigenschaft: Zu jedem $x \in X$ gibt es ein und nur ein $y \in Y$, so dass $(x,y) \in G$. Dieses y bezeichnet man dann als $f(x)$ und erhält eine „Abbildungsvorschrift“ $f : X \to Y$, $x \mapsto f(x)$.

A.4.15. **Beispiel.** (i) $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$. (ii) Für beliebige X, Y und $y \in Y$ hat man die konstante Abbildung vom Wert y, nämlich $f : X \to Y$, $x \mapsto y$. (iii) Für beliebiges X hat man die identische Abbildung $\mathrm{Id}_X : X \to X$, $x \mapsto x$.

A.4.16. **Definition.** Sei $f : X \to Y$ eine Abbildung.

(i) Die Menge X heißt der ***Definitionsbereich***, die Menge Y der ***Wertebereich*** von f.

(ii) Für $x \in X$ heißt $f(x)$ das Bild von x unter f oder, wenn f klar ist, einfach das Bild von x. Es heißt auch der Wert von f auf x oder an der Stelle x.

(iii) Die Menge $\{f(x) : x \in X\} \subset Y$ heißt das ***Bild*** oder die ***Bildmenge*** von f.

(iv) Für $A \subset X$ heißt $f(A) = \{f(x) : x \in A\}$ das Bild von A unter f.

(v) Für $B \subset Y$ heißt $f^{-1}(B) = \{x : f(x) \in B\} \subset X$ das ***Urbild*** von B unter f. (Es kann $f^{-1}(B) = \emptyset$ sein.)

A.4.17. **Definition** (Einschränkung)**.** Sei $f : X \longrightarrow Y$ eine Abbildung und $A \subset X$. Dann bezeichnen wir mit $f|_A : A \longrightarrow Y$ die Abbildung, die jedem $x \in A$ den Wert $f(x) \in Y$ zuordnet. Also $f|_A : A \longrightarrow Y$, $x \mapsto f(x)$. $f|_A$ heißt die ***Einschränkung*** (oder ***Restriktion***) von f auf A. Ist $A \neq X$, so gelten f und $f|_A$ also als verschiedene Abbildungen, obwohl sie mit jedem $x \in A$ „dasselbe machen“.

A.4.18. **Definition** (injektiv, surjektiv, bijektiv)**.** Sei $f : X \longrightarrow Y$ eine Abbildung.

(i) f heißt ***injektiv***, wenn eine der folgenden, zueinander äquivalenten Bedingungen erfüllt ist:
  a) $\forall x_1, x_2 \in X\, (x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2))$,
  b) $\forall x_1, x_2 \in X\, (f(x_1) = f(x_2) \Rightarrow x_1 = x_2)$.

(ii) f heißt ***surjektiv***, wenn $f(X) = Y$.

(iii) f heißt ***bijektiv***, wenn es injektiv und surjektiv ist, d.h. wenn jedes $y \in Y$ das Bild genau eines $x \in X$ ist.

A.4.19. **Definition** (Umkehrabbildung)**.** Sei $f : X \longrightarrow Y$ bijektiv. Zu jedem $y \in B$ gibt es also genau ein $x \in X$ mit $f(x) = y$. Wir bezeichnen dieses x mit $f^{-1}(y)$ und haben damit eine Abbildung $f^{-1} : Y \longrightarrow X$ definiert, welche ***Umkehrabbildung*** von f heißt. Die Umkehrabbildung existiert nur für bijektive Abbildungen $f : X \longrightarrow Y$.

LITERATUR

[1] N. H. Abel. Untersuchungen über die Reihe $1+\frac{m}{1}x+\frac{m(m-1)}{1\cdot2}x^2+\frac{m(m-1)(m-2)}{1\cdot2\cdot3}x^3+\ldots$. *Crelles Journal*, 1:311–339.

[2] E. Behrends. *Maß- und Integrationstheorie.* Springer, 1987.

[3] A. Beutelspacher. *"Das ist o. B. d. A. trivial!"*. Vieweg. Braunschweig. iv+96, 1997. Eine Gebrauchsanleitung zur Formulierung mathematischer Gedanken mit vielen praktischen Tips für Studierende der Mathematik und Informatik.

[4] P. Cohen. *Set Theory and the Continuum Hypothesis*. New York-Amsterdam: W.A. Benjamin, Inc. VI, 154 p., 1966.

[5] J. Dieudonné. *Grundzüge der modernen Analysis. Band 1. (Übers. aus dem Englischen: Ludwig Boll und Klaus Matthes. Übers. d. Berichtigungen u. Erg. aus d. Französ.: Ludwig Boll). 3., berichtigte u. erg. Aufl.* Logik und Grundlagen der Mathematik, Bd. 8. Braunschweig-Wiesbaden: Friedr. Vieweg & Sohn. 369 S., 1985.

[6] H.-D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher, K. Mainzer, J. Neukirch, A. Prestel, R. Remmert, and K. Lamotke. *Zahlen. 3., verb. Aufl.* Springer-Lehrbuch. Berlin: Springer-Verlag. xiv, 333 S., 1992.

[7] J. Elstrodt. *Maß- und Integrationstheorie*. Springer-Lehrbuch. Berlin: Springer. xvi, 434 S., 2007.

[8] L. Euler. *Foundations of differential calculus. Transl. from the Latin by John D. Blanton*. New York, NY: Springer. xiv, 194 p., 2000.

[9] P. R. Halmos. *Naive Mengenlehre. Aus dem Amerikanischen übersetzt von Manfred Armbrust und Fritz Ostermann. 4. Aufl.* Moderne Mathematik in elementarer Darstellung. Band 6. Göttingen: Vandenhoeck&Ruprecht. 132 S., 1976.

[10] A. S. Kechris. *Classical Descriptive Set Theory*. Springer-Verlag, Graduate texts in Math., vol. 156, 1995.

[11] J. L. Kelley. *General topology. 2nd ed.* Graduate Texts in Mathematics. 27. New York - Heidelberg - Berlin: Springer-Verlag. XIV, 298 p., 1975.

[12] M. Koecher. *Lineare Algebra und analytische Geometrie. 4., erg. u. aktual. Aufl.* Springer-Lehrbuch. Berlin: Springer. xiii, 291 S., 1997.

[13] K. Königsberger. *Analysis 1. 6., durchgesehene Aufl.* Springer-Lehrbuch. Berlin: Springer. xiii, 412 S., 2004.

[14] E. Landau. *Grundlagen der Analysis. (Das Rechnen mit ganzen, rationalen, irrationalen, komplexen Zahlen.) Ergänzung zu den Lehrbüchern der Differential- und Integralrechnung. 2. unveränd. Nachdruck*. Darmstadt: Wissenschaftliche Buchgesellschaft XIII, 134 S. , 1970.

[15] S. Lang. *Algebra. 3rd revised ed.* Graduate Texts in Mathematics. 211. New York, NY: Springer. xv, 914 p., 2002.

[16] J. M. Milnor. *Topology from the differentiable viewpoint. Based on notes by David W. Weaver. Revised 2nd ed.* Princeton Landmarks in Mathematics. Princeton, NJ: Princeton University Press. ix, 64 p., 1997.

[17] I. Newton. *Mathematical Works of Isaac Newton*.

[18] G. Pólya. *How to solve it. A new aspect of mathematical method. Expanded version of the 1988 edition, with a new foreword by John H. Conway.* Princeton Science Library. Princeton, NJ: Princeton University Press. xxviii, 253 p., 2004.

[19] H. Sagan. *Space-filling curves.* Universitext. New York: Springer-Verlag. xv, 193 S., 1994.

[20] H. Schubert. *Topologie. Ein Einführung. 4. Aufl.* Mathematische Leitfaden. Stuttgart: B. G. Teubner: 328 S. mit 23 Abb., 121 Aufg. und zahlr. Beisp., 1975.

[21] L. A. Steen and J. A. j. Seebach. *Counterexamples in topology. 2nd ed.* New York - Heidelberg - Berlin: Springer-Verlag. XI, 244 p. , 1978.

[22] B. L. van der Waerden. *Moderne Algebra*. Springer, Berlin, 1930/31, weitere Auflagen 1936, 1950, 1955, 1955, 1966 (ab 1955 unter dem Titel „Algebra“).

[23] S. Wagon. *The banach-Tarski Paradox*. Cambridge University Press, 1985.

[24] W. Walter. *Analysis 1. 7. Aufl.* Springer-Lehrbuch. Berlin: Springer. xiv, 398 S., 2004.

INDEX

MATHEMATISCHES INSTITUT, UNIVERSITÄT ZU KÖLN
Email address: gmarines@math.uni-koeln.de
URL: www.mi.uni-koeln.de/~gmarines