Analysis 1

SS 2006

WS 2006/07

Dr. R. Busam

Inhaltsverzeichnis

I. Zahlsysteme

„Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist *Menschenwerk*“

(Leopold Kronecker, 1886).

„Die Zahlen sind freie Schöpfungen des menschlichen Geistes, sie dienen als Mittel, um die Verschiedenheit der Dinge leichter und schärfer aufzufassen. Durch den rein logischen Aufbau der Zahlenwissenschaft und durch das in ihr gewonnene stetige Zahlen-Reich sind wir erst im Stand gesetzt, unsere Vorstellung von Raum und Zeit genau zu untersuchen, indem wir dieselben auf dieses in unserem Geiste geschaffene Zahlen-Reich beziehen.“

(Richard Dedekind (1888) in „Was sind und was sollen die Zahlen? “

Dieses erste Kapitel handelt von „Zahlsystemen“. Zahlen stellen eine wichtige Grundlage der gesamten Mathematik, speziell aber der Analysis dar. Man kann den von Dedekind und Kronecker ausgedrückten Standpunkt einnehmen und etwa die reellen Zahlen, ausgehend von den „natürlichen Zahlen“ über die „ganzen Zahlen“ und die „rationalen Zahlen“ konstruieren. Doch stellt sich gleich die Frage: „Was ist eine natürliche Zahl? “. Man kann – nach dem Vorbild von Peano und Dedekind (1888)– die natürlichen Zahlen mittels eines Axiomensystems (sog. Peano-Axiome) charakterisieren und aus diesen Grundannahmen alle Aussagen über das Rechnen mit natürlichen Zahlen, so wie man es von der Schule gewohnt ist, ableiten. Mit relativ einfachen algebraischen Mitteln, lassen sich aus den natürlichen Zahlen dann die *ganzen Zahlen* und die *rationalen Zahlen* konstruieren. Der Übergang von den *rationalen Zahlen* zu den *reellen Zahlen*, erfordert jedoch kompliziertere Begriffsbildungen aus der Analysis bzw. Algebra. Wir werden deshalb nicht so vorgehen, sondern den Standpunkt einnehmen, dass jede(r) intuitiv weiß, was reelle Zahlen sind.

„Was aber ist eine reelle Zahl? “

Wie die Diskussion in der Vorlesung über diese Frage ergeben hat, ist die Antwort hierauf gar nicht so einfach.

1 Axiomatische Einführung der reellen Zahlen

Um eine tragfähige Basis für den Aufbau der Analysis zu schaffen, fixieren wir den Begriff der reellen Zahl, indem wir *Axiome für die reellen Zahlen* angeben: Grundeigenschaften, die wir als wahr betrachten und nicht weiter hinterfragen, aus denen sich aber **alle** weiteren Aussagen über reelle Zahlen auf rein logischem Weg ableiten lassen. Diese Axiome sind

I. die Körperaxiome (1.1)

II. die Anordnungsaxiome (1.2)

III. (ein) Vollständigkeitsaxiom (1.3)

(GA): Wir nehmen also an, dass eine Menge $\mathbb{R}$ gibt – $\mathbb{R}$ heißt Menge der reellen Zahlen – , die den obigen Axiomen genügt. Was das im einzelnen bedeutet, wird im Folgenden erklärt. Statt „a ist eine reelle Zahl“, schreiben wir $a \in \mathbb{R}$ (Sprechweise: „a ist Element von $\mathbb{R}$“).

1.1 Die Körperaxiome

Hier werden solche Grundeigenschaften aufgelistet, die zur Begründung des „*Rechnens*“, d. h. der Addition, der Subtraktion, der Multiplikation und der Division, ausreichen. In den Axiomen kommen nur Addition und Multiplikation vor, Subtraktion und Division werden aus diesen abgeleitet.

1.1.1 Axiome der Addition

Wir nehmen folgendes an:

Je zwei reellen Zahlen a, b ist genau eine weitere reelle Zahl $a + b$ ihre *Summe*, zugeordnet und diese *Addition* genannte Zuordnung genügt den folgenden Bedingungen:

(A_1) Für alle $a, b \in \mathbb{R}$ gilt

$a + b = b + a$ (Kommutativität bezüglich der Addition)

(A_2) Für alle $a, b, c \in \mathbb{R}$ gilt

$(a + b) + c = a + (b + c)$ (Assoziativität bezüglich der Addition)

(A_3) Es gibt ein spezielles (eindeutig bestimmtes) Element $0 \in \mathbb{R}$, genannt Null(element), so dass für alle $a \in \mathbb{R}$ gilt

$a + 0 = a$ (Existenz eines neutralen Elements bezüglich der Addition).

(A_4) Zu jeder reellen Zahl a gibt es eine reelle Zahl $(-a)$ (genannt „Negatives von a") mit der Eigenschaft

$a + (-a) = 0$ (Existenz eines Negativen).

1.1.2 Axiome der Multiplikation

Wir nehmen weiter an:

Je zwei Zahlen $a, b \in \mathbb{R}$ ist genau eine weitere reelle Zahl $a \cdot b$, ihr *Produkt*, zugeordnet und diese *Multiplikation* genannte Zuordnung erfüllt folgende Bedingung

(M_1) Für alle $a, b \in \mathbb{R}$ gilt

$a \cdot b = b \cdot a$ (Kommutativität bzgl.. der Multiplikation)

(M_2) Für alle $a, b, c \in \mathbb{R}$ gilt

$(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (Assoziativität bzgl. der Multiplikation)

(M_3) Es gibt ein spezielles, (eindeutig bestimmtes,) von 0 verschiedenes Element $1 \in \mathbb{R}$, genannt Eins(element), so dass für alle $a \in \mathbb{R}$ gilt

$a \cdot 1 = a$ (Existenz eines neutralen Elements bzgl. der Multiplikation)

(M_4) Zu jedem $a \in \mathbb{R}, a \neq 0,$ existiert eine weitere reelle Zahl a^{-1} (genannt Inverses von a), so dass gilt

$a \cdot a^{-1} = 1$ (Existenz eines Inversen)

Ferner soll folgendes Axiom über die *Verbindung von Addition und Multiplikation* erfüllt sein:

(D) Für alle $a, b, c \in \mathbb{R}$ gilt

$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ (Distributivgesetz)

1.1.3 Bemerkung

Die Axiome besagen, dass die reellen Zahlen $\mathbb{R}$ mit den Verknüpfungen der Addition und Multiplikation einen *Körper* bilden,

wenn man darunter eine Menge K versteht, in welcher je zwei Elementen eine Summe $a+b$ und ein Produkt $a \cdot b$ so zugeordnet sind, dass für die Elemente von K die (A_1) bis (A_4), M_1 bis (M_4) und (D) entsprechenden Regeln gelten.

Wir werden noch zahlreiche Beispiele weiterer Körper kennen lernen. Der Körperbegriff wird auch in der Vorlesung über „Lineare Algebra 1“ ausführlich behandelt. Aus dem Schulunterricht geläufige Rechenregeln, wie

$$(-a)(-b) = ab, \quad \text{speziell} \quad (-1)(-1) = 1 \quad \text{oder}$$
$$(a-b)(a+b) = a^2 - b^2$$

lassen sich aus den obigen Axiomen einfach ableiten.

Zur Vereinfachung der Schreibweise treffen wir zunächst die folgenden *Vereinbarungen:*

(a) Statt $a \cdot b$ schreiben wir meistens einfach ab.

(b) Wir vereinbaren die *„Vorfahrtsregel“*

„Punktrechnung geht vor Strichrechnung“.

Dann schreibt sich z. B. das Distributivgesetz einfacher

$$a(b+c) = ab + ac$$

Wir ziehen einige Folgerungen aus den Axiomen, wobei wir zunächst nur die Axiome der Addition benutzen.

1.1.4 Folgerungen aus den Axiomen der Addition

(a) Die Summe der Zahlen $a, b, c \in \mathbb{R}$ ist unabhängig von Klammerung und Reihenfolge der Summanden. Man schreibt daher einfach $a + b + c$.

(b) Das Element Null in $(A3)$ ist eindeutig bestimmt.

(c) Das Negative zu a in $(A4)$ ist eindeutig bestimmt.

(d) Für je zwei Zahlen $a, b \in \mathbb{R}$ hat die Gleichung

$$a + x = b$$

genau eine Lösung x in $\mathbb{R}$ nämlich $x = b + (-a)$

(e) Für alle $a, b \in \mathbb{R}$ gilt (Vorzeichenregel)

$$-(-a) = a \quad \text{und} \quad -(a+b) = -a - b$$

Beweis:

(a) Wiederholte Anwendung von $(A1)$ und $(A2)$.

(b) Es gelte sowohl

$(*)\;\; a+0=a$ für alle $a \in \mathbb{R}$ und auch
$(**)\;\; a+0'=a$ für alle $a \in \mathbb{R}$.

Zu zeigen ist: $0=0'$.

Setzt man in $(*)\;\; a=0'$, so folgt

$$0'+0=0'$$

setzt man in $(**) a=0$ so folgt

$$0+0'=0\ .$$

Nach $(A1)$ ist aber $0+0'=0'+0$, daher

$$0=0'\ .$$

(c) Es sei sowohl

$$a+(-a)=0, \quad \text{als auch} \quad a+a'=0\ .$$

Zu zeigen ist: $-a=a'$. Nun ist

$$\begin{aligned} -a &= -a+0 \\ &= -a+(a+a') \\ &= -a+a'+a \\ &= a'+(-a)+a \\ &= a'+a+(-a) \\ &= a'+0 \\ &= a' \end{aligned}$$

(d) In der Tat ist $x:=b+(-a)$ eine Lösung der Gleichung

$$a+x=b$$

wie man durch Einsetzen bestätigt:

$$a+b+(-a)=a+(-a)+b=0+b=b+0=b\ .$$

Ist umgekehrt $x \in \mathbb{R}$ eine Lösung von

$a+x=b$, so folgt wie oben
$x=x+0=x+a+(-a)=a+x+(-a)=b+(-a)$.

Statt $b+(-a)$ schreiben wir einfach $b-a$ und nennen diese Zahl die *Differenz (von* b *und* a*)*.

Damit ist auch die *Subtraktion* von reellen Zahlen definiert. Man beachte, dass das Minus–Zeichen „–“ jetzt in zweierlei Bedeutung verwendet wird, einmal bedeutet $-a$ die zu a negative Zahl, in $b-a := (b+(-a))$ bedeutet das „–“ eine Rechenoperation, nämlich die Subtraktion.

(e) Nach Definition des Negativen ist

$$(-a)+(-(-a))=0\ .$$

Andererseits ist auch

$$(-a)+a=0$$

Wegen der Eindeutigkeit des Negativen ist daher

$$-(-a)=a\ .$$

Die zweite Behauptung beweist man etwa so: $-(a+b)$ und $(-a)+(-b)$ sind Lösungen der Gleichung $(a+b)+x=0$.

Wegen der eindeutigen Lösbarkeit dieser Gleichung muss

$$-(a+b)=(-a)+(-b)$$

gelten.

Völlig analoge Folgerungen kann man nun aus den Axiomen der Multiplikation ziehen, wir fassen uns kurz:

1.1.5 Folgerungen aus den Axiomen der Multiplikation

(a) Das Produkt der Zahlen $a,b,c \in \mathbb{R}$ ist unabhängig von der Klammerung und Reihenfolge der Faktoren, wir schreiben einfach

$$abc\ .$$

(b) Es gibt in $\mathbb{R}$ genau ein Einselement, *das* Einselement.

(c) Zu jedem $a \in \mathbb{R}, a \neq 0$, gibt es genau ein Inverses $a^{-1} \in \mathbb{R}$, das *Inverse* von a.

(d) Für je zwei reelle zahlen a, b mit $a \neq 0$ hat die Gleichung

$$ax=b$$

genau eine Lösung $x \in \mathbb{R}$, nämlich $x=a^{-1}b$. Statt $a^{-1}b$ schreibt man auch b/a oder $\frac{b}{a}$ und nennt eine solche Zahl *Quotient* (oder auch Bruch). Damit ist auch die Division durch reelle Zahlen $\neq 0$ definiert (als Umkehrung der Multiplikation).

(e) Für alle $a \in \mathbb{R}, a \neq 0$ gilt

$$(a^{-1})^{-1}=a \qquad \text{und}$$
$$(ab)^{-1}=b^{-1}a^{-1}=a^{-1}b^{-1}\ ,$$

falls $a \neq 0$ und $b \neq 0$ gilt. (Vgl. 1.1.6.(b))

Beweis: Man schließt wie im Beweis von 1.1.4, indem man $+$ durch $\cdot$ ersetzt.

Mit Hilfe des *Distributivgesetzes* ergeben sich weitere Folgerungen.

1.1.6 Folgerungen, in denen zusätzlich (D) benutzt wird

(a) Für alle $a \in \mathbb{R}$ gilt $a \cdot 0 = 0$.

(b) Für alle $a, b \in \mathbb{R}$ gilt: $ab=0$ genau dann, wenn $a=0$ oder $b=0$. (Nullteilersatz)

(c) Für alle $a, b \in \mathbb{R}$ gilt

$$(-a)b=a(-b)=-(ab)\ .$$

(d) Für alle $a, b \in \mathbb{R}$ gilt (die Vorzeichenregel)

$$(-a)(-b)=ab, \text{ speziell } (-1)(-1)=1\ .$$

e) Regeln der „Bruchrechnung“:

Für alle $a,b,c,d \in \mathbb{R},\ b \neq 0,\ d \neq 0$ gilt

(α) $\frac{a}{b} = \frac{c}{d}$ genau dann, wenn $ad = bc$. Speziell gilt für alle $x \neq 0 : \frac{ax}{bx} = \frac{a}{b}$.

(β) $\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}$.

(γ) $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$.

(δ) $\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc}$ (falls auch $c \neq 0$).

Beweis: (a) Aus $0 + 0 = 0$ folgt nach (D)

$$a \cdot 0 + a \cdot 0 = a(0 + 0) = a \cdot 0 \; .$$

Wegen der eindeutigen Lösbarkeit der Gleichung (bzgl. y).

$$a \cdot 0 + y = a \cdot 0$$

muss $a \cdot 0 = 0$ sein.

□

(b) Ist $b = 0$, dann gilt nach (a) $a \cdot 0 = 0$.

Genauso gilt natürlich $0 \cdot b = 0$. Sei umgekehrt $a \cdot b = 0$. Dann ist zu zeigen: $a = 0$ oder $b = 0$. Ist $b = 0$, dann braucht nichts bewiesen werden. Ist aber $b \neq 0$, dann existiert b^{-1}, und es folgt

$$0 = a \cdot b^{-1} = (a \cdot b)b^{-1} = a(b \cdot b^{-1}) = a \cdot 1 = a \; .$$

Beachte: Für die 2×2-Matrizen

$$A := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \text{ und } B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ gilt}$$

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \; (AB \text{ ist das Matrizenprodukt}) \; .$$

Frage: Was ist BA?

(c) Wir beweisen z. B. $a(-b) = -(ab)$. Es ist $b + (-b) = 0$ und daher. $ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0$.

$a(-b)$ löst also die Gleichung $ab + x = 0$. Andererseits ist nach Def. $ab + (-(ab)) = 0$, daher ist $a(-b) = -(ab) = -ab$.

(d) Es ist

$$\begin{aligned} (-a)(-b) &= -(a(-b)) && \text{nach (c)} \\ &= -(-(ab)) && \text{nach (c)} \\ &= ab && \text{nach 1.1.4(e)} \; . \end{aligned}$$

Speziell gilt also $(-1)(-1) = 1$.

Übung: Versuchen Sie die letzte Gleichung direkt zu beweisen, indem Sie von $1 + (-1) = 0$ ausgehen.

(e) Bei den Regeln der Bruchrechnung beweisen wir exemplarisch (α) und (β):

Für reelle Zahlen a, b, c, d mit $b \neq 0$ und $d \neq 0$ gilt:

$$\frac{a}{b} = \frac{c}{d} \qquad \text{genau dann, wenn gilt:} \qquad ad = bc \; .$$

Die Richtung $(\Rightarrow)$ in (α) : Wir multiplizieren die linke Gleichung mit bd und erhalten:

$$\begin{aligned} (ab^{-1})(bd) &= (cd^{-1})(bd) && \text{oder} \\ a(b^{-1}b)d &= c(d^{-1}d)b && \text{oder} \\ a1d &= c1b && \text{also} \\ ad &= cb = bc\,. \end{aligned}$$

Es folgt also $ad = bc$, die rechte Gleichung.

Die Richtung $(\Leftarrow)$ in (α) : Ist jetzt $ad = bc$, dann folgt durch Division dieser Gleichung durch $bd(\neq 0)$:

$$\begin{aligned} \frac{ad}{bd} &= \frac{bc}{bd} && \text{oder} \\ \frac{a}{b} &= \frac{c}{d}\,. \end{aligned}$$

Folgerung: Für jedes $x \in \mathbb{R}$ mit $x \neq 0$ ist deshalb auch:

$$\frac{a}{b} = \frac{ax}{bx} \qquad \text{(Erweitern und Kürzen)}$$

(β) : Es ist (– dabei wird die obige Folgerung verwendet –)

$$\frac{a}{b} + \frac{c}{d} = \frac{ad}{bd} + \frac{cb}{db} = \frac{ad}{bd} + \frac{bc}{bd} = \frac{ad+bc}{bd}\,.$$

(Beachte: Aus $b \neq 0$ und $d \neq 0$ folgt $bd \neq 0$.)

Übung: Geben Sie einen etwas anderen Beweis, indem Sie benutzen, dass $\frac{ad+bc}{bd}$ die eindeutig bestimmte Lösung der Gleichung $bd\,x = ad + bc$ ist.

1.1.7 Ein ausführliches Beispiel

Für alle $a, b \in \mathbb{R}$ gilt die *binomische Formel*

$$\boxed{(a+b)^2 = a^2 + 2ab + b^2\,.}$$

Beweis: Zunächst ist $2 := 1 + 1$ und für $x \in \mathbb{R}$ ist $x^2 := xx$.

$$\begin{aligned} (a+b)^2 &= (a+b)(a+b) && \text{nach Def.} \\ &= (a+b)a + (a+b)b && \text{nach } (D) \\ &= a(a+b) + b(a+b) && \text{nach } (A_1) \\ &= (aa+ab) + (ba+bb) && \text{nach } (D) \\ &= ((aa+ab)+ba) + bb && \text{nach } (A_2) \\ &= (aa+(ab+ab)) + bb && \text{nach } (A_2) \text{ und } (M_1) \\ &= (aa+(ab\cdot 1+ab\cdot 1)) + bb && \text{nach } (M_3) \\ &= (aa+ab(1+1)) + bb && \text{nach } (M_3) \text{ und } (D) \\ &= aa + (1+1)ab + bb && \text{nach } (M_1) \\ &= a^2 + 2ab + b^2 && \text{nach Def.} \end{aligned}$$

$\square$

Wenn bis jetzt auch nur Altbekanntes - wie z. B. die Vorzeichenregeln, die binomische Formel (im Spezialfall) o der die Regeln der Bruchrechnung herausgekommen sind, hat die axiomatische Methode den Vorteil, dass die aus den Körperaxiomen abgeleiteten Regeln für *alle Körper* gelten.

Davon gibt es - wie schon bemerkt - „sehr viele“, z.B. bilden die rationalen Zahlen oder die komplexen Zahlen Körper bzgl. der jeweils dort definierten Operationen Addition und Multiplikation. Einen „skurrilen“ Körper erhält man in folgendem

1.1.8 Beispiel: Körper mit 2 Elementen

Nach Def. muss ein Körper mindestens zwei Elemente enthalten (wegen $1 \neq 0$).

Sei $K = \{\bar{0}, \bar{1}\}$ eine Menge mit $\bar{0} \neq \bar{1}$.

Wir definieren eine Addition und Multiplikation durch folgende Tabellen:

$+$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{0}$

$\cdot$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$

Die Körperaxiome sind leicht nachzuprüfen, K ist tatsächlich ein Körper, in dem

$$\bar{1} + \bar{1} = \bar{0}$$

gilt. In der *Algebra* zeigt man, dass es zu jeder *Primzahl* p einen Körper mit p Elementen gibt, allgemeiner gibt es zu jeder Primzahlpotenz p^n, n eine natürliche Zahl, einen Körper mit p^n Elementen. Also gibt es z. B. Körper mit $4, 8, 9, 16, 65536$ oder auch 65537 Elementen, oder auch mit $2^{13\,466\,917} - 1$ Elementen. Der Körper $K = \{\bar{0}, \bar{1}\}$ ist der kleinste Körper, den man sich vorstellen kann. In der Literatur wird er auch häufig mit $\mathbb{F}_2$ oder $\mathbb{Z}/2\mathbb{Z}$ bezeichnet.

Dass auch in $\mathbb{R}$ die Gleichung

$$1 + 1 = 0$$

gilt, können wir also bisher nicht ausschließen. Aus den Anordnungsaxiomen wird sich jedoch ergeben, dass $\mathbb{R}$ die Gleichung

$$1 + 1 = 0$$

nicht erfüllt sein kann.

1.2 Die Anordnungsaxiome

Die Anordnungsaxiome, welchen wir uns nun zuwenden wollen, stellen die Grundlage für das Rechnen mit *Ungleichungen*, für *Abschätzungen* und auch für die *elementare Fehlerrechnung* dar. Ungleichungen sind in der Analysis (mindestens) so wichtig, wie Gleichungen in der Algebra.

Wir nehmen folgendes an:

In $\mathbb{R}$ ist eine Teilmenge P, genannt die *positiven Zahlen* - ausgezeichnet, so dass folgende Eigenschaften gelten:

(AnO$_1$) : Für jede reelle Zahl a gilt genau eine der folgenden Aussagen:

$a \in P, a = 0$ oder $(-a) \in P$

(AnO$_2$) : Für beliebige $a, b \in P$ gilt $a + b \in P$.

(AnO$_3$) : Für beliebige $a, b \in P$ gilt $ab \in P$.

Sprechweise: Statt „$a \in P$“ sagen wir auch „a ist positiv“ und verwenden dafür die Abkürzung $a > 0$. Ist $-a \in P$, dann heißt a negativ, Abkürzung $a < 0$.
Eine reelle Zahl ist also entweder positiv oder negativ oder gleich Null. Durch Auszeichnung der positiven reellen Zahlen kann man auch zwei reelle Zahlen der Größe nach vergleichen. Man definiert für $a, b \in \mathbb{R}$

$$a < b \quad :\Leftrightarrow \quad b > a \quad \Leftrightarrow \quad b - a \in P \quad .$$

$$a \leq b \quad :\Leftrightarrow \quad b \geq a \quad :\Leftrightarrow \quad [\, a < b \text{ oder } a = b] \quad .$$

Sprechweise: a kleiner b und b größer a, bzw. a kleinergleich b und b größergleich a.

Die Beziehungen

$$a < b \text{ bzw. } a \leq b$$

nennt man auch *„Ungleichungen“*, in manchen Zusammenhängen auch *„Abschätzungen“*

1.2.1 Definition (Angeordneter Körper)

Die reellen Zahlen $\mathbb{R}$ bilden einen *(an)geordneten Körper*, wenn man darunter einen Körper K versteht, in welchem ein Positivitätsbereich P so ausgezeichnet ist, dass die Grundeigenschaft (AnO_1), (AnO_2) und (AnO_3) gelten. Die rationalen Zahlen bilden auch einen angeordneten Körper.

1.2.2 Eigenschaften von “ $<$ “ und “ $\leq$ “

(a) *Trichotomiegesetz:* Für je zwei reelle Zahlen a, b gilt genau eine der Aussagen:

$$a < b \text{ oder } a = b \text{ oder } a > b$$

(b) *Transitivität:* Für beliebige $a, b, c \in \mathbb{R}$ gilt:

Aus $a < b$ und $b < c$ folgt $a < c$

(c) *Verträglichkeit mit der Addition:* Für beliebige $a, b, c, d \in \mathbb{R}$ gilt: Aus

$$a < b \quad \text{und} \quad c \leq d \quad \text{folgt}$$

$$a + c \quad < \quad b + d$$

„Gleichsinnige“ Ungleichungen darf man also addieren.

Insbesondere folgt aus $a < b$ für beliebiges $c \in \mathbb{R}$

$$a + c < b + c\ .$$

(d) *Verträglichkeit mit der Multiplikation* Für beliebige $a, b, c \in \mathbb{R}$ mit $c \neq 0$ gilt:

Aus $a < b$ und $c > 0$ folgt $ac < bc$.

Aus $a < b$ und $c < 0$ folgt $ac > bc$.

(d’) Aus $0 \leq a < b$ und $0 \leq c < d$ folgt $0 \leq ac < bd$.

Kurz: Ungleichungen zwischen nicht negativen Zahlen, darf man multiplizieren.

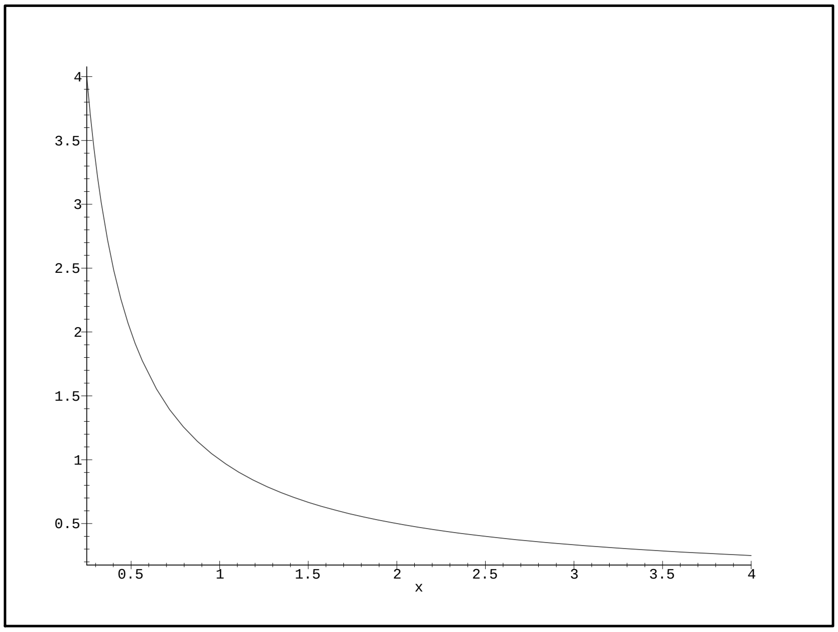


Abbildung 1: Graph der Funktion $f :]0,\infty[\to]0,\infty[,\ f(x) = 1/x$.

(e) *„Spiegelungseigenschaft"*: Für $a, b \in \mathbb{R}$ gilt:

$a < b$ genau dann, wenn $-a > -b$

(f) *Vergleichbarkeit:* Für $a, b \in \mathbb{R}$ gilt:

Aus $a \leq b$ und $b \leq a$ folgt $a = b$.

(g) Für jede $a \in \mathbb{R}$ mit $a \neq 0$ ist $a^2 > 0$, insbesondere ist $1 = 1^2 > 0$ und $-1 < 0$, also auch $-1 \neq x^2$ für alle $x \in \mathbb{R}$.

(h) Für $a \in \mathbb{R}$ gilt: Es ist $a > 0$ gleich bedeutend mit $a^{-1} > 0$.

(i) Für $a, b \in \mathbb{R}$ mit $0 < a < b$ folgt

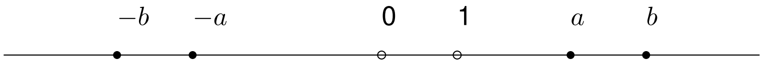
$$a^{-1} > b^{-1}$$

Hieraus folgt z.B. dass die Funktion aus der Abbildung 1

$$\begin{array}{rcl} f : P & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{x} \end{array}$$ streng monoton fällt.

Bemerkung:

Veranschaulicht man sich die Menge $\mathbb{R}$ der reellen Zahlen als die Punkte auf einer Geraden, auf der zwei verschiedene Punkte 0 und 1 als Wahl vom Ursprung (Nullpunkt) und Maßeinheit markiert sind und zwar so, dass 1 rechts von 0 liegt, dann liegen die positiven Zahlen rechts vom Nullpunkt, die negativen links davon. Von zwei Zahlen ist diejenige größer, die weiter rechts liegt. Addition einer Zahl bedeutet eine Verschiebung (Translation). Der Übergang von a zu $-a$ bedeutet eine *Spiegelung am Nullpunkt*, wobei die Rollen von rechts und links vertauscht werden.



Gelegentlich werden wir auch Ungleichungen verketten.

Statt $a < b$ und $b < c$, schreiben wir kurz
$a < b < c$, z. B.
$0 < 1 < 2$.

Dabei sollte man jedoch immer nur „gleichsinnige Ungleichungen“ verketten.

Beweis von 1.2.2

(a) Man muss nur die Definition einsetzen:

$$\begin{array}{lll} a < b & \text{bedeutet} & b - a \in P \\ a = b & \text{bedeutet} & b - a = 0 \\ a > b & \text{bedeutet} & a - b = -(b-a) \in P \end{array}$$

Nach (AnO_1) tritt für jede reelle Zahl x genau eine dieser Fälle ein.

(b) Aus $a < b$, d. h. $b - a \in P$ und
$b < c$, d. h. $c - b \in P$

folgt nach (AnO_2)
$(c - b) + (b - a) \in P$, das heißt aber
$(c - a) \in P$ und das bedeutet $a < c$.

(c) Wir behandeln zunächst den Fall $c < d$. Dann besagen die Voraussetzungen
$(b - a) \in P$ und $(d - c) \in P$,
dann gilt nach (AnO_2)

$$(b - a) + (d - c) \in P\ .$$

Es ist aber

$$(b - a) + (d - c) = b + d - (a + c), \text{ also}$$
$$a + c < b + d$$

Für $c = d$ schließt man analog: $a + c < b + c$ ist gleich bedeutend mit $(b + c) - (a + c) \in P$.

Es ist aber $(b + c) - (a + c) = (b - a) + (c - c) = (b - a) + 0 = b - a \in P$.

Aus $a < b$ folgt für beliebiges $c \in \mathbb{R}$ dann $a+c < b+c$. Das bedeutet eine Translationsinvarianz für Ungleichungen.

(d) Sei zunächst $c > 0$, dann ist zu zeigen, dass aus $a < b$ folgt: $ac < bc$ oder $bc - ac \in P$.

Nach Voraussetzung ist aber $a < b$, d.h. $b - a \in P$, und nach (AnO_2) folgt für $c > 0$ auch $(b - a)c \in P$.

$(b - a)c$ ist aber nach dem Distributivgesetz $bc - ac$, also gilt $bc - ac \in P$, d.h. $ac < bc$. Der Beweis im Fall $c < 0$ (dann ist $-c > 0$) verläuft völlig analog.

(e) Aus $a < b$ folgt durch Addition von $-b-a$ nach 1.2.1 (c) $-b-a+a < -b-a+b$, also $-b < -a$.

(f) Zu zeigen: Aus $a \leq b$ und $b \leq a$ folgt $a = b$. $a \leq b$ bedeutet: $a = b$ oder $a < b$. $b \leq a$ bedeutet: $b = a$ oder $b < a$. Wegen der Trichotomie können $a < b$ und $b < a$ nicht gleichzeitig gelten, also muss $a = b$ gelten.

(g) Ist $a > 0$, so ist auch $a^2 > 0$ nach (AnO_3). Ist $a < 0$, so ist $(-a) > 0$, also $(-a)(-a) > 0$.

Es ist aber

$$\begin{aligned} (-a)(-a) &= -(a(-a)) \\ &= -(-(aa)) = aa = a^2\ . \end{aligned}$$

Der endliche Körper $\mathbb{F}_2$ aus Beispiel 1.1.8 lässt sich nicht anordnen.

Wegen $\bar{1} \neq \bar{0}$ müsste die Alternative $\bar{0} < \bar{1}$ oder $\bar{1} < \bar{0}$ gelten.

Durch Addition von $\bar{1}$ zur linken Ungleichung folgt $\bar{1} = \bar{0}+\bar{1} < \bar{1}+\bar{1} = \bar{0}$, also ein Widerspruch, da nicht gleichzeitig $\bar{0} < \bar{1}$ und $\bar{1} < \bar{0}$ gelten kann.

Völlig analog schließt man, falls $\bar{1} < \bar{0}$ sein sollte. Auch der Körper $\mathbb{C}$ der komplexen Zahlen lässt sich nicht anordnen, weil es eine komplexe Zahl i gibt, für die $i^2 = -1$ gilt.

Erst jetzt sieht man, dass $\mathbb{R}$ mehr Zahlen als nur die Zahlen 0 und 1 enthält:

Definiert man $2 := 1 + 1,\ 3 := 2 + 1, \ldots$ so folgt jetzt $0 < 1,\ 1 < 1 + 1 = 2,\ 1 + 1 < 2 + 1 = 3$, also $2 < 3$ etc. $\mathbb{R}$ enthält also mindestens alle Zahlen, die man durch sukzessive Addition von 1 aus Null (bzw. 1) erhält.

(h) Ist $a > 0$, dann ist $a^{-1} \neq 0$, da $a \cdot a^{-1} = 1 \neq 0$ ist (Nullteilersatz).

Wäre $a^{-1} = \frac{1}{a} < 0$, dann folgt mit (d)

$$aa^{-1} = 1 < 0\ ,$$

das steht aber im Widerspruch zu 1.2.2(8). Also folgt aus $a > 0$ auch $a^{-1} = \frac{1}{a} > 0$. Insbesondere ist $0 < \frac{1}{2} < 1$. Die Umkehrung gilt natürlich auch:

Aus $a^{-1} > 0$ folgt $a > 0$.

Man verwendet den gleichen Schluss wie eben und beachtet $(a^{-1})^{-1} = a$.

(i) Man zeigt etwas allgemeiner:

Unter der Voraussetzung $ab > 0$ gilt:

$a < b$ genau dann, wenn $\frac{1}{b} < \frac{1}{a}$.

Es ist

$$\frac{1}{a} - \frac{1}{b} = \frac{b-a}{ab}\ .$$

„$\Leftarrow$“: Ist nun $a < b$, dann ist $b - a > 0$ und auch gilt $(ab)^{-1} > 0$, also ist die rechte Seite positiv, das heißt

$$\frac{1}{a} - \frac{1}{b} > 0 \ \text{ oder } \ \frac{1}{b} < \frac{1}{a}\ .$$

„$\Leftarrow$“: Ist die linke Seite positiv, dann muss wegen $ab > 0$ auch $b - a > 0$ gelten, d.h. es ist $a < b$.

1.2.3 Kleine Anwendungen

(a) Für alle $a, b \in \mathbb{R}$ mit $a < b$ gilt

$$a < \frac{a+b}{2} < b\ .$$

$A(a, b) := \dfrac{a+b}{2}$ heißt das *arithmetische Mittel* von a und b.

Zwischen je zwei verschiedenen reellen Zahlen gibt es stets weitere reelle Zahlen.

Der Beweis ist offensichtlich:

Aus $a < b$ folgt $a + a < a + b$ und $a + b < b + b$ also auch

$$\begin{aligned} 3a + a < a + b &\quad < b + b && \text{und damit} \\ a(1 + 1) < a + b &\quad < b(1 + 1) && \text{und mit der Definition } 2 := 1 + 1 > 0 \\ a < \frac{a+b}{2} &\quad < b\ . \end{aligned}$$

(b) Sind $a, b \in \mathbb{R}$, dann ist

$$0 \le (a-b)^2 = a^2 - 2ab + b^2 \,, \tag{1.1}$$ also

$$2ab \le a^2 + b^2 \,, \tag{1.2}$$ und damit

$$\boxed{ab \le \frac{a^2+b^2}{2}}$$

Addiert man in (1.2) auf beiden Seiten $2ab$, so folgt

$$4ab \le a^2 + 2ab + b^2 = (a+b)^2 \qquad \text{oder}$$

$$ab \le \left(\frac{a+b}{2}\right)^2 .$$

Sind $a, b \ge 0$, so kann man hierfür (unter der Verwendung der Existenz von Quadratwurzeln aus nicht negativen Zahlen) auch

$$\boxed{G(a,b) := \sqrt{ab} \le \frac{a+b}{2} =: A(a,b)}$$

schreiben. Das ist die Ungleichung zwischen dem **geometrischen** und **arithmetischen** Mittel.

(c) Für welche $x \in \mathbb{R}$ hat der Ausdruck

$$x^2 - x + \frac{1}{4}$$

seinen kleinsten Wert (Minimum), d. h. gibt es ein oder mehrere $x_0 \in \mathbb{R}$, so dass

$$x^2 - x + \frac{1}{4} \ge x_0^2 - x_0 + \frac{1}{4}$$

für alle $x \in \mathbb{R}$ gilt?

Es ist aber

$$x^2 - x + \frac{1}{4} = \left(x - \frac{1}{2}\right)^2 \ge 0 \,,$$

wobei das Gleichheitszeichen genau dann gilt, wenn $x = x_0 := \frac{1}{2}$.

Geometrische Interpretation ?

Man kann die Ungleichung auch in der Form schreiben:

$$x - x^2 \le \frac{1}{4} \,, \qquad \text{oder}$$

$$x(1-x) \le \frac{1}{4} \,, \qquad \text{für alle } x \in \mathbb{R} \,.$$

Die geometrische Interpretation dieser Ungleichung ist, dass für die Funktion

$$f : \mathbb{R} \to \mathbb{R} \,, \qquad x \mapsto x(1-x)$$

aus der Abbildung 2 die Stelle $x_0 := \frac{1}{2}$ die einzige Maximalstelle ist und $f\left(\frac{1}{2}\right) = \frac{1}{4}$ das Maximum von f auf $\mathbb{R}$ ist.

(d) *Fehlerabschätzungen für Näherungswerte*

Häufig verwendet man in physikalischen Anwendungen Näherungswerte, z.B. für $y := \dfrac{1}{\sqrt{1+x}}$

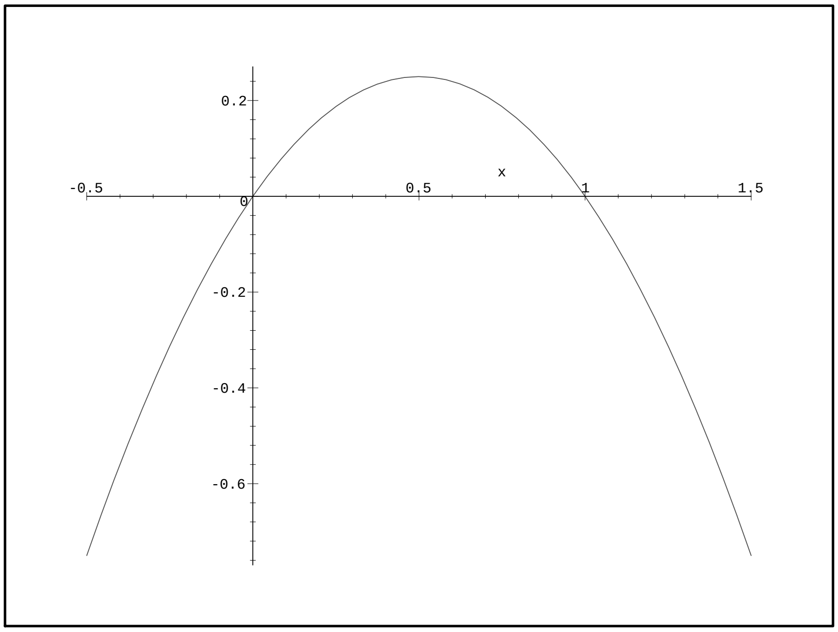


Abbildung 2: Graph der Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x(1-x)$.

den Näherungswert $y_0 := 1-\dfrac{1}{2}x+\dfrac{3}{8}x^2$ oder manchmal nur $1-\dfrac{1}{2}x$, und steht dann vor dem Problem, den relativen Fehler (bezogen auf den wahren Wert bzw. Näherungswert) abzuschätzen.

Wir betrachten ein einfacheres Beispiel:

Für $y := \dfrac{1}{1+x}$ (x klein) (aufgefasst als „wahrer Wert") wird häufig der Näherungswert

$$y_0 := 1 - x$$

verwendet.

Für den *Fehler*

$$
\begin{array}{ll}
y_0 - y = 1 - x - \dfrac{1}{1+x} = \dfrac{1-x^2}{1+x} - \dfrac{1}{1+x} = -\dfrac{x^2}{1+x} & \text{gilt} \\
y_0 - y \le 0 & \text{für alle } x \ge -1 \text{ und} \\
y_0 - y \ge -\dfrac{\frac{1}{100}}{1-\frac{1}{10}} = -\dfrac{1}{90} & \text{für alle } x \in \mathbb{R} \text{ mit } -\dfrac{1}{10} \le x \le \dfrac{1}{10}\,.
\end{array}
$$

Der *relative Fehler* (bezogen auf den wahren Wert)

$$
\begin{array}{rlll}
\dfrac{y_0 - y}{y} & = -x^2 \le 0 & \text{für alle } x \in \mathbb{R} & \text{und} \\
& \ge -\dfrac{1}{100} & \text{für alle } x \in \mathbb{R} & \text{mit } -\dfrac{1}{10} \le x \le \dfrac{1}{10}\,.
\end{array}
$$

Das folgende Lemma wird manchmal „**Fundamentallemma der Analysis**" genannt, wir werden es häufig anwenden.

1.2.4 (Fundamental-)Lemma

Ist $a \in \mathbb{R}$ und gilt für *jede* positive reelle Zahl ε die Ungleichung

$$(*) \qquad a \le \varepsilon \,,$$

dann gilt $a \le 0$.
Zusatz: Ist zusätzlich $a \ge 0$, dann muss $a = 0$ sein.

Beweis:(durch Widerspruch).

Wir nehmen an, dass $a > 0$, dann ist $\varepsilon := \dfrac{a}{2}$ ebenfalls positiv und eine zulässige Wahl. Da die Ungleichung $(*)$ für *alle* positiven $\varepsilon > 0$ gilt, muss sie speziell also auch $\varepsilon := \frac{a}{2}$ gelten und damit $2a \le a$ oder $a \le 0$ im Widerspruch zur Annahme $a > 0$.

□

Mit Hilfe der Anordnung der reellen Zahlen lassen sich wichtige Teilmengen von $\mathbb{R}$ definieren, die *Intervalle*.
Die meisten (reellen) Funktionen, die wir betrachten werden, haben Intervalle als Definitionsbereich, oder die Definitionsbereiche setzen sich aus Intervallen in einfacher Weise zusammen.
Angaben wie $3,14 < \pi < 3,15$ oder $1,414 < \sqrt{2} < 1,415$ bedeuten, dass π im Intervall zwischen $3,14$ und $3,15$ bzw. dass $\sqrt{2}$ im Intervall zwischen $1,414$ und $1,415$ liegt.

1.2.5 Definition (Intervalle)

Jede der in der folgenden Liste aufgeführte Teilmenge von $\mathbb{R}$ heißt *Intervall*.

1. *Abgeschlossene (sogar kompakte) Intervalle*
   (a) Für $a, b \in \mathbb{R}$ mit $a < b$ sei $[a, b] := \{x \in \mathbb{R}\,;\ a \le x \le b\}$.
   (b) Für $a = b$ sei $[a, a] := \{a\}$.
2. *Offene Intervalle*
   (a) $]a, b[= \{x \in \mathbb{R}\,;\ a < x < b\}\ (a < b)$
   (b) Im Fall $a = b$ ist $]a, a[:= \emptyset$,
   wir fassen die leere Menge also auch als Intervall auf.
3. *Halboffene Intervalle*
   (a) $[a, b[= \{x \in \mathbb{R}\,;\ a \le x < b\}$
   (b) $]a, b] = \{x \in \mathbb{R}\,;\ a < x \le b\}$
4. *Abgeschlossene Halbstrahlen*
   (a) $[a, \infty[= \{x \in \mathbb{R}\,;\ x \ge a\}$
   (b) $]-\infty, b] = \{x \in \mathbb{R}\,;\ x \le b\}$
5. *Offene Halbstrahlen*
   (a) $]a, \infty[= \{x \in \mathbb{R}\,;\ x > a\}$
   (b) $]-\infty, b[= \{x \in \mathbb{R}\,;\ x < b\}$

6. *Die Zahlengerade* $]-\infty, \infty[:= \mathbb{R}$.

   Mit der leeren Menge $\emptyset$, den einpunktigen Intervallen $\{a\}$ und ganz $\mathbb{R}$ gibt es also genau *11 Intervalltypen*.

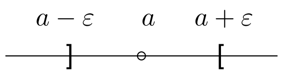
Die Bezeichnungen “abgeschlossenes“, “kompaktes“, “offenes“, ..., Intervall sind im Augenblick nur Namen. In den Fällen 1., 2. und 3. heißt $\ell := b - a$ die *Länge* des jeweiligen Intervalls.

Der Unterschied zwischen dem *abgeschlossenen* Intervall $[a, b]$ und dem *offenen* Intervall $]a, b[$ besteht offensichtlich darin, dass im ersten Fall die beiden „Randpunkte“ a und b zum Intervall gehören, im zweiten Fall aber nicht zum Intervall gerechnet werden.

Auf diese Begriffsbildungen kommen wir später ausführlich zurück. Für unsere späteren Kovergenzbetrachtungen wichtiges Intervall erhält man so:

Ist $a \in \mathbb{R}$ und $\varepsilon \in \mathbb{R},\ \varepsilon > 0$, dann heißt das *offene Intervall* $]a - \varepsilon,\ a + \varepsilon[$ auch *ε-Umgebung von a*.

Bezeichnung: $U_\varepsilon(a) :=]a - \varepsilon,\ a + \varepsilon[$ *Veranschaulichung:*



Mit Hilfe der Anordnung von $\mathbb{R}$ können wir einen wichtigen Begriff der Analysis einführen, den Betrag. Gleichzeitig erhalten wir ein Beispiel für eine wichtige Funktion (die stetig ist, aber im Nullpunkt nicht differenzierbar ist).

1.2.6 Definition (Betrag)

Ist $a \in \mathbb{R}$ so heißt die Zahl

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

der *Betrag* von a.

Ordnet man jeder reellen Zahl x ihren Betrag $|x|$ zu erhält man eine Funktion:

$$\begin{aligned} |\ | : \mathbb{R} &\to \mathbb{R}\,, \\ x &\mapsto |x|\,, \end{aligned}$$

deren Graph die bekannte Gestalt aus der Abbildung 3 hat. Der Betrag ignoriert das *Vorzeichen (Signum)* einer reellen Zahl, das man formal durch

$$\operatorname{sign}(a) = \begin{cases} 1, & \text{falls} \quad a > 0 \\ 0, & \text{falls} \quad a = 0 \\ -1 & \text{falls} \quad a < 0 \end{cases}$$

definieren kann.

Offenbar gilt

$$|a| = a \operatorname{sign}(a)$$

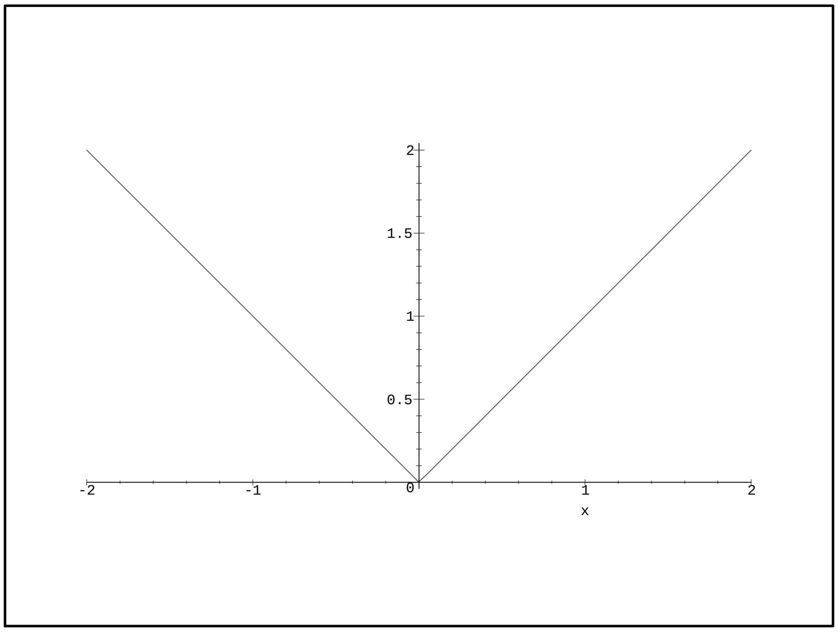


Abbildung 3: Graph der Betrag–Funktion $x \to |x|$.

1.2.7 Hilfssatz

Für $\varepsilon, x \in \mathbb{R}$, $\varepsilon \geq 0$ gilt

$$\boxed{|x| \leq \varepsilon \quad \text{ist gleichbedeutend mit} \quad -\varepsilon \leq x \leq \varepsilon}$$

insbesondere ist stets

$$-|x| \leq x \leq |x| \; .$$

Beweis: Die Richtung $(\Rightarrow)$:

Sei $|x| \leq \varepsilon$.

Ist $x \geq 0$, so ist $|x| = x$, also $-\varepsilon \leq 0 \leq x = |x| \leq \varepsilon$.

Ist $x \leq 0$, so ist $|x| = -x$, also $-\varepsilon \leq -|x| = x \leq 0 \leq \varepsilon$.

Die Richtung $(\Leftarrow)$:

Sei jetzt $-\varepsilon \leq x \leq \varepsilon$.

Ist $x \geq 0$, so ist $|x| = x \leq \varepsilon$.

Ist $x \leq 0$, so ist ($|x| = -x$ also) $-|x| = x \geq -\varepsilon$.

Durch Multiplikation mit (-1) folgt daher

$$|x| \leq \varepsilon \; .$$

Die Folgerung erhält man, indem man

$$\varepsilon = |x|$$

wählt. Sie folgt auch direkt aus der Definition von $|x|$.

Zusatz: Es gilt natürlich auch für $\varepsilon > 0$: $|x| < \varepsilon$ ist gleichbedeutend mit $-\varepsilon < x < \varepsilon$.

1.2.8 Satz: (Eigenschaften des Betrags)

Für $a, b \in \mathbb{R}$ gilt

(a) $|a| \geq 0$ und [$|a| = 0 \;\Leftrightarrow\; a = 0$] .

(b) $|a\,b| = |a|\,|b|$ (Multiplikativität), speziell $|-a| = |a|$,

(b') $\left|\dfrac{a}{b}\right| = \dfrac{|a|}{|b|}$, falls $b \neq 0$.

(c) $|a \pm b| \leq |a| + |b|$ (Dreiecksungleichung) ,

(d) $\Big|\, |a| - |b| \,\Big| \leq |a \pm b|$ (Dreiecksungleichung für Abschätzungen nach unten).

Beweis:

(a) folgt unmittelbar aus der Definition

(b) Wegen $|ab| \geq 0$ und $|a|\,|b| \geq 0$ gilt stets $|ab| = \pm ab = |a|\,|b|$.

Man kann den Beweis auch durch Fallunterscheidungen führen, man hat dann vier Fälle zu unterscheiden (siehe Vorlesung).

(b') folgt wegen $\dfrac{a}{b} \cdot b = a$ aus der Multiplikativität (b).

Man kann den Beweis auch durch Fallunterscheidungen führen, man hat dann vier Fälle zu unterscheiden (siehe Vorlesung).

(c) Aus

$-|a| \leq a \leq |a|$ und

$-|b| \leq b \leq |b|$ folgt (nach 1.2.2(c))

$-(|a| + |b| \leq a + b \leq |a| + |b|$ und damit nach Hilfssatz 1.2.7

$|a + b| \leq |a| + |b|$.

Ersetzt man b durch $-b$ in $|a + b| \leq |a| + |b|$, so erhält man

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

(d) Mit (c) folgt:

$|a| = |(a - b) + b| \leq |a - b| + |b|$, daher

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

Vertauscht man die Rollen von a und b, so folgt

$|b| - |a| \leq |b - a| = |a - b|$, daher

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

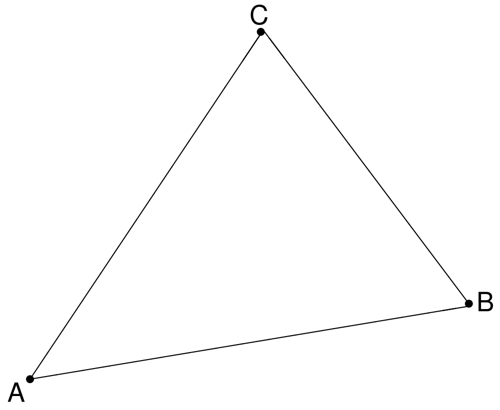
Ersetzt man b durch $-b$, so folgt auch $\Big||a| - |b|\Big| = \Big||a| - |-b|\Big| \leq |a - (-b)| = |a + b|$.

□

Bemerkungen:

- Definiert man als Abstand von $a, b \in \mathbb{R}$ die Zahl $d(a, b) := |a - b|$, dann besagt (d), dass der Abstand von a und b mindestens so groß ist, wie der Abstand ihrer Beträge.

- Auch in höheren Dimensionen, schon in $\mathbb{R}^2$, besagt die Dreiecksungleichung, dass zwei Seiten eines Dreiecks zusammen immer mindestens so lang sind wie die dritte Seite.



$$d(A,C) \leq d(A,B) + d(B,C)$$

  In der Dimension eins, degeneriert das Dreieck zu drei Punkten auf der Zahlengeraden die Aussage der Dreiecksungleichungen bleiben dennoch richtig, wir werden sie z. B. im nächsten Kapitel (Folgen und Reihen) dauernd verwenden.

- Ein Körper K zusammen mit einer Abbildung $|\ | : K \to \mathbb{R}$ mit den folgenden Eigenschaften:

  (a) $|a| \geq 0$ für alle $a \in K$ und $|a| = 0$ genau dann, wenn $a = 0$ ist,
  (b) $|a\, b| = |a|\, |b|$ für alle $a, b \in K$,
  (c) $|a + b| \leq |a| + |b|$ für alle $a, b \in K$,

  heißt bewerteter Körper.

  $\mathbb{R}$ ist also ein bewerteter Körper. Wir werden sehen, dass der Körper der komplexen Zahlen kein angeordneter Körper aber ein bewerteter Körper ist.

  Ferner gibt es für jede Primzahl p eine Bewertung $|\ |_p$ für den Körper der rationalen Zahlen $\mathbb{Q}$:

$$\left| p^u \frac{a}{b} \right|_p := p^{-u} \qquad a, b \in \mathbb{Z}, a \neq 0, b \neq 0\ , u \in \mathbb{Z}\ ,\ \ a,\, b \text{ nicht teilbar durch } p, |0|_p := 0\ .$$

Zeigen Sie als Übung:

(a) Für alle $a \in \mathbb{R}$ gilt $|a| = |-a|$, indem Sie die Definition des Betrags verwenden und nicht auf die Multiplikativität und $-a = (-1)a$ zurückgreifen.

(b) Für alle $a \in \mathbb{R}$ gilt $|a| \geq 0$ und $|a|^2 = a^2$. Zeigen Sie, dass diese Eigenschaften den Betrag charakterisieren, d.h. ist $x \in \mathbb{R}$, $x \geq 0$ und $x^2 = a^2$, dann ist $x = |a|$. Verwenden Sie diese Charakterisierung, um die Multiplikativität $|a\, b| = |a|\, |b|$ anders zu beweisen.

Für das in 1.2.3(d) behandelte Beispiel gilt also für den relativen Fehler $\left|\frac{y_0 - y}{y}\right| \leq \frac{1}{100}$ für alle $x \in \mathbb{R}$ mit $|x| \leq \frac{1}{10}$.

Als weiteres Beispiel behandeln wir die folgende

1.2.9 Aufgabe

Zu bestimmen ist die Menge

$$M := \left\{ x \in \mathbb{R} \,;\ x \neq 0,\ \left| \frac{5}{x} + x \right| < 6 \right\}$$

Ferner soll M möglichst einfach mit Hilfe von Intervallen dargestellt werden.

Lösung der Aufgabe:
Es gelten folgende Äquivalenzen für $x \in \mathbb{R}$, $x \neq 0$:

$$\begin{aligned}
x \in M \quad &\Leftrightarrow \quad \left| \frac{5}{x} + x \right| < 6 \\
&\Leftrightarrow \quad \left| \frac{5+x^2}{x} \right| < 6 \\
&\Leftrightarrow \quad \frac{|5+x^2|}{|x|} < 6 \\
&\Leftrightarrow \quad |5+x^2| < 6|x| \\
&\Leftrightarrow \quad 5+x^2 < 6|x| \\
&\Leftrightarrow \quad x^2 - 6|x| + 5 < 0 \\
&\Leftrightarrow \quad |x|^2 - 6|x| + 9 < 4 \\
&\Leftrightarrow \quad (|x|-3)^2 < 2^2 \\
&\Leftrightarrow \quad \big|\, |x| - 3 \,\big| < 2 \\
&\Leftrightarrow \quad -2 < |x| - 3 < 2 \\
&\Leftrightarrow \quad 1 < |x| < 5 \\
&\Leftrightarrow \quad x \in]-5,-1[\, \cup \,]1,5[\ .
\end{aligned}$$

M ist also die Vereinigung der (offenen) Intervalle $]-5,-1[$ und $]1,5[$.
Die geometrische Veranschaulichung ist in der Abbildung 4 festgehalten. Als Vorbereitung für das Vollständigkeitsaxiom führen wir noch die Begriffe *Maximum* und *Minimum* einer Teilmenge $M \subset \mathbb{R}$ ein.

1.2.10 Definition (Maximum, Minimum)

Ist $M \subset \mathbb{R}$ dann heißt eine reelle Zahl c *Maximum* von M – in Zeichen $c = \max M$ –, falls gilt:

(a) $c \in M$

(b) Fur alle $x \in M$ gilt $x \leq c$.

Völlig analog definiert man das Minimum von M, in Zeichen $\min M$.

Bemerkung: Wegen 1.2.2(f) hat jede Teilmenge $M \subset \mathbb{R}$ höchstens ein Maximum bzw. Minimum. Die Bezeichnungen $\max M$ und $\min M$ sind also sinnvoll.

Beispiele:

(a) Fur alle $a \in \mathbb{R}$ gilt $|a| = \max\{a, -a\}$

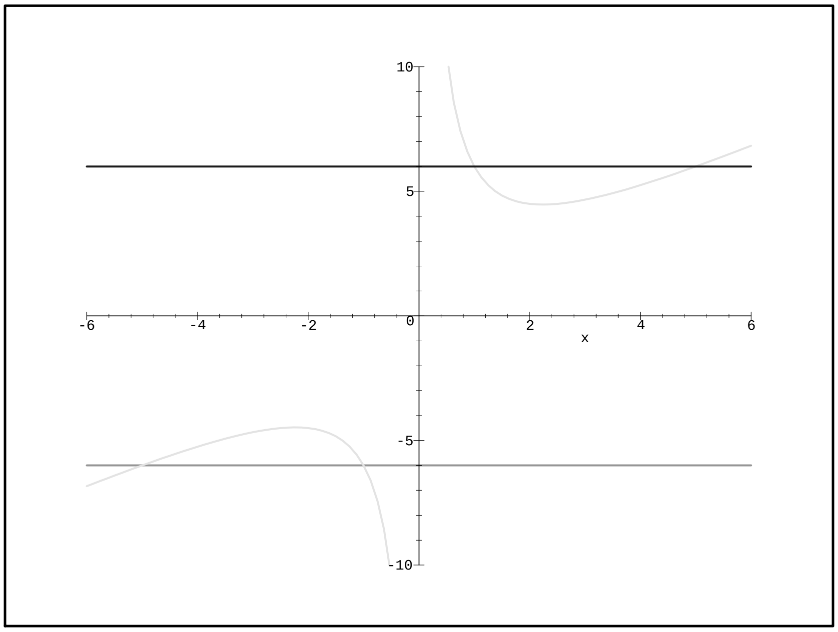


Abbildung 4: Graph der Funktionen $x \to -6$, $x \to \frac{5}{x} + x$, $x \to +6$.

(b) Fur alle $a, b \in \mathbb{R}$ mit $a < b$ gilt $\max[a,b] = \max]a,b] = \max]-\infty, b] = b$

(c) Fur alle $a, b \in \mathbb{R}$ mit $a < b$ gilt $\min[a,b] = \min[a,b[= \min[a,\infty[= a$

(d) Dagegen hat das offene Intervall $]a,b[$ kein Maximum, denn zu jedem $x \in]a,b[$ gibt es ein $y \in]a,b[$ mit $y > x$, etwa
$$y = \frac{x+b}{2} \ .$$

(e) Genauso sieht man, dass die Intervalle
$$[a,b[, \quad]-\infty,b[, \quad [a,\infty[, \quad]a,\infty[\quad \text{und} \quad]-\infty,\infty[= \mathbb{R}$$
kein Maximum, und die Intervalle
$$]a,b[, \quad]a,b], \quad]a,\infty[, \quad]-\infty,b], \quad]-\infty,b[\quad \text{und} \quad]-\infty,\infty[= \mathbb{R}$$
kein Minimum haben.

(f) Eine Teilmenge $M \subset \mathbb{R}$ hat genau dann ein Maximum (Minimum), falls die am „Nullpunkt gespiegelte Menge“
$$-M := \{x \in \mathbb{R}, -x \in M\}$$
ein Minimum (Maximum) hat. Es ist dann $\max M = -\min(-M)$ bzw. $\min M = -\max(-M)$.

Wir betrachten das spezielle Intervall
$$M = \{x \in \mathbb{R} \ ; \ x < 1\} =]-\infty, 1[$$
M hat nach Beispiel (a) kein Maximum, die Zahl 1 die als Kandidat für ein Maximum ins Auge springt, gehört aber nicht zur Menge M.

Jedoch haben alle Elemente $x \in M$ die Eigenschaft, dass $x < 1$ gilt, 1 ist also eine „obere Schranke für M“ im Sinne der folgenden Definition:

1.2.11 Definition (obere und untere Schranke)

Ist $M \subset \mathbb{R}$, so heißt eine Zahl $s \in \mathbb{R}$ *obere (bzw. untere) Schranke* von M, falls für alle $x \in M$ die Ungleichung

$$x \leq s \qquad (\text{ bzw. } x \geq s\ )$$

gilt.

Ist s eine obere (bzw. untere) Schranke von M, dann ist auch jede größere (bzw. kleinere) Zahl s' obere (bzw. untere) Schranke von M.

Besitzt M ein Maximum c, dann ist c obere Schranke von M und offensichtlich das Minimum in der Menge der oberen Schranken von M.

1.2.12 Definition (Beschränktheit)

Eine Menge $M \subset \mathbb{R}$ heißt *nach oben (unten) beschränkt*, falls M eine obere (untere) Schranke besitzt. M heißt *beschränkt*, wenn M sowohl nach oben als auch nach unten beschränkt ist.

Die Zahl 1 ist obere Schranke von

$$M = \{\ x \in \mathbb{R}\ ;\ x < 1\ \} =]-\infty, 1[\ .$$

Sie ist sogar eine besondere obere Schranke, nämlich die kleinste unter allen oberen Schranken (also wieder das Minimum in der Menge aller oberen Schranken von M).

Wir behaupten:

Ist s eine obere Schranke von M dann ist $s \geq 1$.

Beweis : Wir fuhren einen Widerspruchsbeweis, nehmen also an, es gebe eine obere Schranke s von M mit $s < 1$.

Wir betrachten dann die Zahl (das arithmetische Mittel von s und 1)

$$x := \frac{s+1}{2}\ ,$$

für sie gilt $s < x < 1$, also ist $x \in M$, aber $x > s$, das ist ein Widerspruch zur Voraussetzung, dass s obere Schranke von M ist.

Unser Ergebnis lautet:

$$M := \{\ x \in \mathbb{R}\ ;\ x < 1\ \}$$

ist nach oben beschränkt. M besitzt kein Maximum, aber die (nicht zur Menge gehörende Zahl 1 ist kleinste obere Schranke von M, d.h. das Minimum in der Menge der oberen Schranken von M.

Dass **jede** nicht leere nach oben beschränkte Menge reeller Zahlen eine kleinste obere Schranke hat, ist die letzte Grundvoraussetzung, die wir über die reellen Zahlen machen.

□

1.3 Vollständigkeitsaxiom

(V) Jede nicht leere nach oben beschränkte Menge M reeller Zahlen besitzt eine kleinste obere Schranke, d.h. es gibt $s_0 \in \mathbb{R}$ mit folgenden Eigenschaften:

- Für alle $x \in M$ ist $x \leq s_0$ (d.h. s_0 ist obere Schranke);
- Für *jede* obere Schranke s von M gilt $s_0 \leq s$;

Bemerkung: Ist $\bar{S}(M) = \{\ s \in \mathbb{R}\ ;\ s$ obere Schranke von $M\ \}$, so ist also

$$s_0 = \min\ \bar{S}(M)\ .$$

Als Minimum einer Menge ist s_0 eindeutig bestimmt (das folgt auch direkt aus der Definition).

Wir halten fest:

Bemerkung: Die kleinste obere Schranke einer nicht leeren nach oben beschränkten Menge $M \subset \mathbb{R}$ ist eindeutig bestimmt.

Wir bezeichnen diese reelle Zahl mit $\sup\ M$. Sie kann zur Menge M gehören oder auch nicht.

Bemerkung: Ist $M \subset \mathbb{R}$ nach oben beschränkt, dann ist die am *Nullpunkt gespiegelte Menge*

$$-M = \{\ x \in \mathbb{R}\ ;\ -x \in M\ \}$$

nach unten beschränkt. Es ergibt sich:

Folgerung: Ist M eine nicht leere nach unten beschränkte Menge reeller Zahlen, dann besitzt M eine eindeutig bestimmte *größte untere Schranke*, für die die Bezeichnung $\inf\ M$ (Infimum von M) verwendet wird.

Halten wir nochmals fest:

Bemerkung: Ist M eine Menge von reellen Zahlen, die ein Maximum bzw. Minimum besitzt, dann ist

$$\max\ M = \sup\ M \in M \qquad \text{bzw.} \qquad \min\ M = \inf\ M \in M\ .$$

Existieren umgekehrt $\sup M$ bzw. $\inf M$ und gilt $\sup M \in M$ bzw. $\inf M \in M$, dann ist $\sup M = \max M$ bzw. $\inf M = \min M$.

1.3.1 Satz (ε–Charakterisierung von sup bzw. inf)

Ist $s_0 \in \mathbb{R}$ eine obere (untere) Schranke der nicht leeren Teilmenge $M \subseteq \mathbb{R}$, dann ist $s_0 = \sup M$ (bzw. $s_0 = \inf M$) genau dann, wenn es zu jedem $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, ein $x \in M$ mit $s_0 - \varepsilon < x$ (bzw. $x < s_0 + \varepsilon$) gibt.

Beweis : Sei zunächst $s_0 = \sup M$ und wir nehmen an, es gäbe ein $\varepsilon > 0$ mit der Eigenschaft $x \leq s_0 - \varepsilon$ für alle $x \in M$.

Dann wäre auch $s_0 - \varepsilon$ eine obere Schranke von M und wegen $s_0 - \varepsilon < s_0$ eine kleinere obere Schranke als s_0. Widerspruch.

Zum Beweis der Umkehrung (– die ε–Bedingung sei als erfüllt –) betrachten wir neben s_0 eine weitere obere Schranke s von M. Wäre nun $s < s_0$, dann gibt es nach Voraussetzung (zu $\varepsilon := s_0 - s > 0$) ein $x \in M$ Mir $s < x$. Das widerspricht der Voraussetzung, dass$s$ obere Schranke von M ist. Daher ist $s_0 \leq s$ für jede obere Schranke s von M. s_0 ist also das Supremum von M.

□

Wir werden sehen, dass das Vollständigkeitsaxiom (V) die *rationalen Zahlen* von den reellen Zahlen unterscheidet. Die rationalen Zahlen, die Standardbezeichnung ist $\mathbb{Q}$, bilden ebenfalls einen angeordneten Körper, in welchem die Gleichung

$$r^2 = 2$$

jedoch keine Lösung hat, d.h. es gibt keine rationale r mit $r^2 = 2$. *Schlagwort*: $\sqrt{2}$ ist irrational.

Wir zeigen als Anwendung des Vollständigkeitsaxioms, dass es zu jeder nicht negativen reellen Zahl a eine eindeutig bestimmte nicht negative reelle Zahl x gibt, für die $x^2 = a$ gilt.

1.3.2 Satz (Existenzsatz für Quadratwurzeln)

Zu jeder reellen Zahl a mit $a \geq 0$ existiert eine eindeutig bestimmte nicht negative reelle Zahl x mit $x^2 = a$.

x heißt **die** Quadratwurzel aus a und wird mit $\sqrt{a}$ bezeichnet.

Beweis:
1. Fall: $a = 0$.
Obwohl man den Fall $a = 0$ im Folgenden mitbehandeln könnte, behandeln wir ihn vorab.
Die Gleichung $x^2 = 0$ hat nach der Nullteilerregel nur die Lösung $x = 0$. Wir setzen daher

$$\sqrt{0} = 0 \ .$$

2. Fall: $a > 0$.

(a) *Existenz*: Wir betrachten die folgende Menge

$$M := \{ \ y \in \mathbb{R} \ ; \ y \geq 0; \ y^2 \leq a \ \} \ .$$

Erinnerung: $y^2 := yy$.

Wegen $0^2 = 0 < a$ ist $0 \in M$, also $M \neq \emptyset$. Offensichtlich ist $a + 1$ eine obere Schranke von M, denn ist $y > a + 1$, so ist $y^2 > (a + 1)^2 = a^2 + 2a + 1 > a$.

Ist also $y \in M$, dann ist $y \leq a + 1$. M besitzt also nach dem Vollständigkeitsaxiom (V) ein Supremum: $x := \sup M$.

Behauptung: $x > 0$ und $x^2 = a$.

Wegen $0 \in M$ ist $0 \leq x$. Offensichtlich ist $y_0 := \min\{1, a\} \in M$, daher $0 < y_0 \leq x$, also $x > 0$.

Wir zeigen $x^2 = a$, indem wir die Annahmen $x^2 < a$ und $x^2 > a$ jeweils zu einem Widerspruch fuhren. Nach dem Trichotomiegesetz muss also $x^2 = a$ gelten.

Wäre $x^2 > a$, dann ist **auch**

$$s := x - \underbrace{\frac{x^2 - a}{2x}}_{>0} < x$$

eine obere Schranke von M mit $s < x$, denn es gilt:

$$\begin{aligned} s^2 &= \left(x - \frac{x^2 - a}{2x}\right)^2 \\ &= x^2 - 2x\frac{x^2 - a}{2x} + \underbrace{\left(\frac{x^2 - a}{2x}\right)^2}_{>0} \\ &> x^2 - 2x\frac{x^2 - a}{2x} \\ &= x^2 - (x^2 - a) \\ &= a\ . \end{aligned}$$

Es folgt dann für jedes $y \in M$:

$$y^2 \leq a < s^2$$

und hieraus (wegen $y \geq 0$ und $s > 0$)

$$y < s\ ,$$

das heißt, s wäre auch eine obere Schranke für M. Das widerspricht der Voraussetzung, dass x *kleinste* obere Schranke von M ist.

Wäre $x^2 < a$, dann setzen wir

$$\delta := \min\{\ 1\ ,\ \frac{a - x^2}{2x + 1}\ \}\ .$$

Dann ist offensichtlich $\delta > 0$ und

$$\begin{aligned} (x + \delta)^2 &= x^2 + 2x\delta + \delta^2 \\ &\leq x^2 + 2x\delta + \delta \qquad \text{(wegen } \delta \leq 1\text{)} \\ &= x^2 + (2x + 1)\delta \\ &\leq x^2 + (2x + 1)\frac{a - x^2}{2x + 1} \\ &= x^2 + (a - x^2) \\ &= a\ , \end{aligned}$$

also ist auch $x + \delta \in M$, und $x + \delta > x$.

Damit wäre x nicht obere Schranke von M. Also ist auch $x^2 < a$ unmöglich. Daher bleibt nur $x^2 = a$.

(b) *Eindeutigkeit*: Wäre für eine positive Zahl x_1 auch $x_1^2 = a$, so folgt:

$$(x - x_1)(x + x_1) = x^2 - x_1^2 = a - a = 0\ ,$$

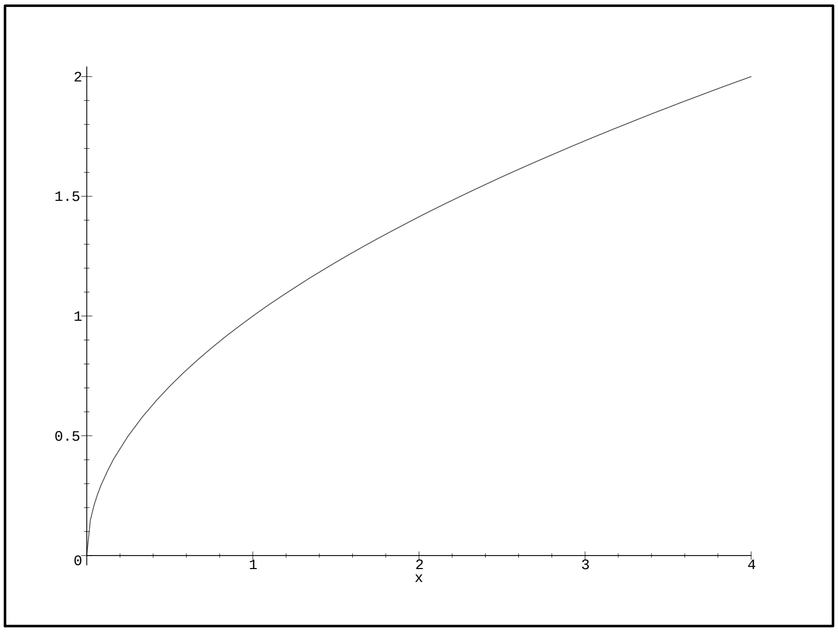


Abbildung 5: Graph der Funktion $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $f(x) = \sqrt{x}$.

Wegen $x + x_1 > x > 0$ folgt aus der Nullteilerregel

$$x - x_1 = 0, \qquad \text{also} \qquad x = x_1 \ .$$

Die Schreibweise $\sqrt{a}$ für die Zahl x ist also gerechtfertigt.

$\square$

Übung: Man zeige, dass für alle $a, b \in \mathbb{R}, a \geq 0, b \geq 0$ gilt:

$$\sqrt{a\,b} = \sqrt{a}\,\sqrt{b} \ .$$

Ferner beweise man die *Monotonie–Eigenschaft*: Aus $0 \leq a < b$ folgt $0 \leq \sqrt{a} < \sqrt{b}$.

Ordnet man jeder nicht negativen reellen Zahl ihre Quadratwurzel zu, so erhält man eine Funktion (Abbildung 5)

$$\sqrt{} : \mathbb{R}_+ := \{x \in \mathbb{R};\ x \geq 0\} \to \mathbb{R} \quad x \mapsto \sqrt{x} \ .$$

Man beachte: Es gilt $\sqrt{x^2} = |x|$ (und nicht etwa x) für alle $x \in \mathbb{R}$.

Bemerkung: $\sqrt{a}$ ist eine Lösung der Gleichung $x^2 = a$. Aber ihr Negatives $-\sqrt{a}$ ist auch Lösungen der Gleichung $x^2 = a$. Man könnte jeder nicht negativen reellen Zahl a auch die negative Zahl $-\sqrt{a}$ zuordnen und erhält dadurch ebenfalls eine Funktion (der sog. „Nebenzweig der Quadratwurzel“, Abbildung 6), die meistens in der reellen Analysis unterschlagen wird. $\sqrt{a}$ *und* $-\sqrt{a}$ sind also Lösungen der Gleichung $x^2 = a$, die verschieden sind, wenn $a > 0$ ist, und nur im Fall $a = 0$ zusammenfallen.

Bemerkung: Es gibt ein völlig analogen Existenzsatz für k–te Wurzeln, k eine natürliche Zahl $k \geq 2$:

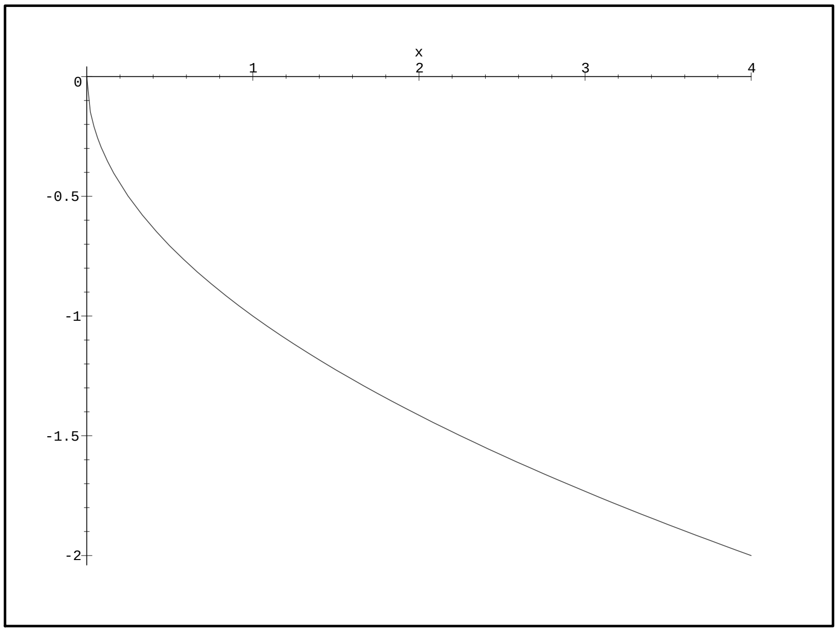


Abbildung 6: Graph der Funktion $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\leq 0}$, $f(x) = -\sqrt{x}$.

1.3.3 Satz (Existenzsatz für k–te Wurzeln)

Zu jeder nicht negativen reellen Zahl a und zu jeder natürlichen Zahl $k \geq 2$ gibt es genau eine nicht negative reelle Zahl x, für die $x^k = a$ gilt. x heißt **die** k–te Wurzel aus a und wird mit

$$\sqrt[k]{a} \qquad \text{oder} \qquad a^{1/k}$$

bezeichnet.

Man kann den Beweis in völliger Analogie zum Beweis des letzten Satzes fuhren, wir verzichten aber hier darauf, weil wir die Existenz der k–ten Wurzeln auf verschiedene andere Weisen beweisen werden.

Schlussbemerkung: Durch die Körperaxiome, die Anordnungsaxiome und den Vollständigkeitsaxiom (V) ist das System der reellen Zahlen eindeutig „bis auf Isometrie“ festgelegt. Ohne diesen Begriff zu präzisieren, sei angemerkt, dass wir z.B. später zeigen werden, dass sich jede reelle Zahl mit Hilfe der in den Axiomen auftretenden Zahlen 0 und 1 eindeutig als *Dualbruch* darstellen lässt.

2 Natürliche Zahlen (vollständige Induktion) ganze Zahlen, rationale Zahlen

In einem beliebigen Körper K gibt es zumindest die Elemente 0 und 1. Um weitere „Zahlen" zu definieren, liegt es nahe, einfach sukzessive die 1 zu addieren, also $2 := 1+1,\ 3 := 2 = 2+1,\ 4 := 3+1,$ usw. und die so erhaltenen Zahlen als *natürliche Zahlen* in K zu bezeichnen. Das Beispiel $\mathbb{F}_2$ des endlichen Körpers mit zwei Elementen zeigt jedoch, dass dieses Verfahren nicht unseren Vorstellungen entspricht, den in $\mathbb{F}$ ist $2 := 1 + 1 = 0$.

Auf Grund der Anordnungsaxiomen in $\mathbb{R}$ können jedoch in dem angeordneten Körper $\mathbb{R}$ (das Gleiche gilt für jeden angeordneten Körper K) solche Pathologien nicht auftreten.

Wir skizzieren im folgenden, wie man die natürlichen Zahlen (Standardbezeichnung ist $\mathbb{N}$) als Teilmenge von $\mathbb{R}$ definieren kann ($\mathbb{R}$ denken wir uns auf Grund der Grundannahme (GA) und der Körper- und Anordnungsaxiome gegeben. Als Nebenprodukt erhalten wir das Beweisprinzip der „vollständigen Induktion". Dieses ist ein beweisbarer Satz und kein Axiom (wie etwa im Buch von Forster oder bei Königsberger). Unsere Vorstellung von den natürlichen Zahlen ist, dass man einen *Anfang des Zählens* hat, nämlich 1, und wenn wir irgendeine noch so große natürliche Zahl, sagen wir N konstruiert haben, gibt es noch eine größere, z.B. $(N + 1)$. Die natürlichen Zahlen werden „größer und größer", die Menge der natürlichen Zahlen hat sicher kein Maximum. Aber sind die natürlichen Zahlen vielleicht doch durch eine reelle Zahl nach oben beschränkt?

Relativ klar ist auch, dass es unendlich viele natürliche Zahlen gibt. Aber was heißt das genau? Auch das werden wir (aber erst später) präzisieren.

2.1 Die natürlichen Zahlen

Die folgende Definition der natürlichen Zahlen mag auf den ersten Blick merkwürdig erscheinen, hat aber den Vorteil exakt zu sein, weil die Unbestimmtheit und das zeitliche Moment, die im Begriff der „sukzessiven Addition" enthalten sind, vermieden werden.

2.1.1 Definition (Zählmenge)

Eine Teilmenge $Z \subset \mathbb{R}$ heißt „*Zählmenge*" (induktiv oder Nachfolgemenge) falls gilt

(a) $1 \in Z$

(b) Für alle $x \in Z$ ist auch stets $(x + 1) \in Z$.

$\mathbb{R}$ selbst ist Zählmenge (eine ziemlich große), aber auch $\mathbb{R}_+ := \{\ x \in \mathbb{R}\ ;\ x \geq 0\ \}$.
$\mathbb{R} - \{2\} := \{\ x \in \mathbb{R}\ ;\ x \neq 2\ \}$ ist keine Zählmenge.

Wir wollen die natürlichen Zahlen als „kleinste" Zählmenge einführen.

2.1.2 Definition und Satz (natürliche Zahl)

Eine reelle Zahl n heißt *natürlich*, wenn n in *jeder* Zählmenge von $\mathbb{R}$ enthalten ist.

$\mathbb{N} := \{n \in \mathbb{R}\ ;\ n \text{ natürlich}\}$ wird als Menge der *natürlichen Zahlen* bezeichnet.

$\mathbb{N}$ ist selber Zählmenge und zwar die kleinste Zählmenge, weil $\mathbb{N}$ in jeder anderen Zählmenge enthalten ist: Mit Hilfe des Durchschnittes kann man auch schreiben

$$\mathbb{N} = \cap \{Z;\ Z \subset \mathbb{R} \text{ Zählmenge}\}.$$

Da $\mathbb{N}$ in jeder Zählmenge enthalten ist, ist nur noch nachzuweisen, dass $\mathbb{N}$ selber wieder eine Zählmenge ist. Dazu ist zu zeigen:

(a) $1 \in \mathbb{N}$, das ist aber klar, da $1 \in Z$ für jede Zählmenge Z gilt.

(b) Ist x eine beliebige reelle Zahl, die in $\mathbb{N}$ liegt, dann liegt x in jeder Zählmenge Z, dann ist auch $(x + 1) \in Z$. Da dies für jede Zählmenge gilt, ist auch $(x + 1) \in \mathbb{N}$.

Damit haben wir gezeigt, dass $\mathbb{N}$ die *kleinste Teilmenge* von $\mathbb{R}$ ist, die alle uns vom Zählen her bekannten „natürlichen" Zahlen $1, 2, 3, \ldots$ enthält. Die natürlichen Zahlen sind also genau das, was wir uns unter ihnen vorstellen. Als beweisbaren Satz erhalten wir nun den

2.1.3 Satz (Induktionssatz)

Ist $M \subset \mathbb{N}$ eine Teilmenge, für die gilt

(a) $1 \in M$

(b) Für alle $n \in M \Rightarrow n + 1 \in M$. Dann ist bereits $M = \mathbb{N}$.

Beweis:
Nach Voraussetzung ist $M \subset \mathbb{N}$. Da aber $\mathbb{N}$ in jeder Zählmenge Z enthalten ist, und M eine Zählmenge ist, muss also auch $\mathbb{N} \subset M$ gelten. Das bedeutet aber insgesamt $M = \mathbb{N}$.

Auf dem Induktionssatz beruht das *Beweisprinzip der Vollständigen Induktion*.

2.1.4 Satz (Beweisprinzip der Vollständigen Induktion)

Für jede natürliche Zahl n sei $A(n)$ eine Aussage (Behauptung) gegeben ($A(n)$ kann wahr oder falsch sein), und es gelte

(I A) $A(1)$ ist wahr. (Induktionsanfang oder Induktionsverankerung)

(I S) Für alle $\in \mathbb{N}:\ A(n) \Rightarrow A(n + 1)$ (Induktionsschritt)

$((A(n)$ nennt man auch Induktionsvoraussetzung und $A(n + 1)$ die Induktionsbehauptung).

Dann gilt $A(n)$ für alle natürlichen Zahlen n.

Zum *Beweis* braucht man nur zu beachten, dass die Menge

$M = \{n \in \mathbb{N}; A(n)$ ist wahr$\} \subset \mathbb{N}$ eine Zählmenge ist, also nach Satz 2.1.3. mit $\mathbb{N}$ identisch ist.

Auf Beispiele für das Beweisprinzip der vollständigen Induktion gehen wir in Abschnitt 2.4 ein. Es sei bemerkt, dass man aus unserer Definition der natürlichen Zahlen, die folgenden *Peano-Eigenschaften der natürlichen Zahlen* $\mathbb{N}$ ableiten kann

2.1.5 Satz (Peano-Eigenschaften der natürlichen Zahlen)

(P_1) : $1 \in \mathbb{N}$ (1 ist eine natürliche Zahl)

(P_2) : Jede natürliche Zahl n besitzt eine eindeutig bestimmte natürliche Zahl $n'(= n+1)$ als Nachfolger.

(P_3) : 1 ist nicht Nachfolger einer natürlichen Zahl $(n' \neq 1$ für alle $n \in \mathbb{N})$

(P_4) : Verschiedene natürliche Zahlen haben verschiedene Nachfolger:
Sind $n, m \in \mathbb{N}$ und ist $n \neq m$, dann ist auch $n' \neq m'$.

(P_5) : Ist $M \subset \mathbb{N}$ mit den Eigenschaften (a) $1 \in M$ (b). Für alle $n \in M \Rightarrow n+1 \in M$, dann ist $M = \mathbb{N}$.

Man kann umgekehrt diese Eigenschaften der natürlichen Zahlen als *Axiome* an den Anfang stellen, wie es R. Dedekind (1888) und G. Peano (1889) getan haben. Man kann dann aus den natürlichen Zahlen $\mathbb{N}$ die ganzen Zahlen $\mathbb{Z}$, die rationalen Zahlen $\mathbb{Q}$ und schließlich die reellen Zahlen $\mathbb{R}$ konstruieren. Eine klassische Darstellung findet man bei E. Landau: „Grundlagen der Analysis“, eine modernere in dem sehr empfehlenswerten Analysisbuch von Amann-Escher (vgl. Literaturverzeichnis). Wir stellen noch einige Eigenschaften von $\mathbb{N}$ zusammen.

2.1.6 Eigenschaften von $\mathbb{N}$

(a) Für jede natürliche Zahl n gilt $n \geq 1$;

(b) Summe und Produkt natürlicher Zahlen sind wieder natürliche Zahlen;

(c) Die Differenz $n - m$ zweier natürlicher Zahlen ist wieder eine natürliche Zahl falls $n > m$ gilt;

(d) Ist n eine beliebige natürliche Zahl, dann gibt es keine natürliche Zahl z mit $n < z < n+1$.

Zwischen n und $n+1$ existiert also keine weitere natürliche Zahl. Insbesondere folgt für $n \in \mathbb{N}$:

(e) Ist $\mathbb{A}_n := \{\, x \in \mathbb{N} \,;\, x \leq n \,\} = \{1, 2, \ldots, n\}$, dann ist

$$\mathbb{A}_{n+1} = \mathbb{A}_n \cup \{n+1\} = \{\, 1, \ldots, n, n+1 \,\};$$

(f) *Wohlordnungsssatz:* Jede nicht leere Teilmenge $B \subset \mathbb{N}$ besitzt ein Minimum;

Diese Eigenschaften lassen sich alle leicht mit vollständiger Induktion beweisen.

Wir beweisen nur den Wohlordnungsssatz (f), dessen Aussage offensichtlich ist, der aber dennoch bewiesen werden muss.

Beweis : Dazu nehmen wir an, es gebe eine nicht leere Teilmenge $B \subset \mathbb{N}$, die kein Minimum besitzt und betrachten die folgenden Menge von natürlichen Zahlen $M := \{n \in \mathbb{N};$ für alle $k \in B$ ist $n \leq k\}$, M besteht also aus allen natürlichen Zahlen, die untere Schranken von B sind. Dann ist $1 \in M$, denn jede natürliche Zahl k ist ≥ 1. Ist $n \in M$, dann enthält B nur Zahlen $\geq n$. Wäre jetzt n in B enthalten, so wäre n die kleinste Zahl in B, B hat aber nach Voraussetzung kein Minimum. Daher enthält B nur Zahlen $\geq n+1$, das bedeutet aber $n+1 \in M$.

Damit ist M eine induktive Teilmenge von $\mathbb{N}$, also $M = \mathbb{N}$ nach dem Induktionssatz. Dann muss B aber leer sein, denn wenn es $n \in B$ gibt, so ist $n+1 \notin M$, M ist aber gleich $\mathbb{N}$.

Daher ist doch $B = \emptyset$ im Widerspruch zur Voraussetzung $B \neq \emptyset$.

□

Zugegeben, die Beweisstruktur ist etwas kompliziert.

Bemerkung: Man kann zeigen, dass umgekehrt aus dem Wohlordnungssatz der Induktionssatz abgeleitet werden kann. Versuchen Sie dies zu beweisen !

Bemerkung: Die $\mathbb{A}_n$ sind Prototypen endlicher Mengen mit genau n Elementen. Wir nennen eine Menge *endlich*, wenn $M = \emptyset$ ist oder wenn man die Elemente von M mit den Zahlen $1, 2, \ldots, n$ so durchnummerieren kann, dass $M = \{ m_1, m_2, \ldots, m_n \}$ gilt, jedes $m \in M$ also eine Nummer erhält und verschiedene Elemente auch verschiedene Nummern erhalten:
Aus $j \neq k$ $(1 \leq j, k \leq n)$ folgt $m_j \neq m_k$.
Man kann leicht zeigen (Induktion nach n), dass dieses n eindeutig bestimmt ist und man nennt n die Elementanzahl von M.

Eine nicht endliche Menge nennen wir *unendlich*.

Aufgabe: Zeigen Sie: $\mathbb{N}$ ist unendlich.

Die Menge der natürlichen Zahlen besitzt kein Maximum, weil mit jeder natürlichen Zahl N auch $N + 1$ eine größere natürliche Zahl ist. Eine natürliche Zahl kann also nicht obere Schranke der Menge $\mathbb{N}$ sein. Möglich wäre jedoch, dass eine nicht zu $\mathbb{N}$ gehörende reelle Zahl s obere Schranke von $\mathbb{N}$ ist. Dass dies nicht der Fall ist, lässt sich mit Hilfe des Vollständigkeitsaxioms zeigen.

2.2 Die Archimedische Anordnung von $\mathbb{R}$

2.2.1 Satz (Archimedische Eigenschaft von $\mathbb{R}$)

Die Menge $\mathbb{N}$ der natürlichen Zahlen ist nicht nach oben beschränkt.

Man kann das auch so ausdrücken: Zu jeder reellen Zahl x gibt es eine natürliche Zahl n mit

$$n > x.$$

Beweis : Wenn $\mathbb{N}$ nach oben beschränkt wäre, müsste $\mathbb{N}$ nach dem Vollständigkeitsaxiom (V) eine kleinste obere Schranke s_o besitzen. Für jedes $n \in \mathbb{N}$ müsste also $n \leq s_0$ gelten. Andererseits muss es eine natürliche Zahl N mit $N > s_0 - 1$ geben, da sonst $s_0 - 1$ obere Schranke von $\mathbb{N}$ wäre, also s_0 nicht die kleinste obere Schranke von $\mathbb{N}$ sein könnte.

Aus $N > s_0 - 1$ folgt aber $N + 1 > s_0$, im Widerspruch zur Tatsache, dass s_0 obere Schranke von N ist. Daher kann nicht $n \leq s_0$ für alle $n \in \mathbb{N}$ gelten.

□

Aus diesem Satz ergibt sich die folgende Folgerung, die schon in der griechischen Mathematik (Eudoxos, 408?–355? v.Chr. – Archimedes 287?–212 v.Chr.) eine Rolle spielte.

2.2.2 Satz (Eudoxos - Archimedes)

Zu jeder reellen Zahl $a > 0$ und jedem $b \in \mathbb{R}$ gibt es eine natürliche Zahl n mit

$$na > b.$$

Beweis : Wenn es eine solche natürliche Zahl n nicht geben würde, müsste also für alle natürlichen Zahlen gelten

$$n \leq \frac{b}{a}.$$

$\frac{b}{a}$ wäre also eine obere Schranke für $\mathbb{N}$ im Widerspruch zu Satz 2.2.1.

$\square$

Bemerkung: Satz $2.2.1$ und Satz 2.2.1 (die übrigens äquivalente Aussagen beinhalten) bedeuten, dass der Körper der reellen Zahlen *archimedisch angeordnet* ist.

Bemerkung: Es gibt angeordnete Körper, in denen diese Aussage nicht gilt; es gibt also angeordnete Körper, die nicht archimedisch angeordnet sind. Wir werden später (nach Einführung der rationalen Zahlen) sehen, dass aus der Archimedischen Eigenschaft der reellen Zahlen folgt, dass zwischen je zwei reellen Zahlen stets eine (und damit auch unendlich viele) rationale Zahlen liegen, man sagt:

$$\boxed{\mathbb{Q} \text{ liegt dicht in } \mathbb{R}}$$

(Hinweis: Satz 2.3.5)

Als Folgerung wollen wir noch festhalten:

2.2.3 Folgerung

Ist ε eine beliebige positive reelle Zahl (also $\varepsilon > 0$), dann gibt es eine natürliche Zahl N mit

$$\frac{1}{N} < \varepsilon.$$

Zum Beweis wähle man in Satz $2.2.2$ $n = N,\ a = \varepsilon$ und $b = 1$.

Man kann auch direkt mit 2.2.1 schließen:

Zu der positiven reellen Zahl $\frac{1}{\varepsilon}$ gibt es eine natürliche Zahl N mit $N > \frac{1}{\varepsilon}$.

Diese Ungleichung ist aber äquivalent zu $\frac{1}{N} < \varepsilon$. Erst recht gilt dann für alle natürlichen Zahlen $n \geq N$, dass auch $\frac{1}{n} < \varepsilon$ gilt.

Diese Eigenschaft bedeutet, dass die Folge $(\frac{1}{n})$ eine Nullfolge ist.

2.3 Ganze Zahlen und rationale Zahlen

Die natürlichen Zahlen haben eine dürftige algebraische Struktur. Zwar ist mit je zwei natürlichen Zahlen m und n auch $m+n$ und $m \cdot n$ eine natürliche Zahl und für die natürliche Zahl 1 gilt $1 \cdot n = n$ für alle $n \in \mathbb{N}$, aber $\mathbb{N}$ besitzt kein neutrales Element bezüglich der Addition und für $n \in \mathbb{N}$ ist $(-n) \notin \mathbb{N}$. Daher erweitert man $\mathbb{N}$ zunächst durch Hinzunahme der Null zu

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

Jetzt ist zwar Null ein neutrales Element bezüglich der Addition in $\mathbb{N}_0$, aber die Gleichung $2+x=1$ besitzt keine Lösung $x \in \mathbb{N}_0$. Deshalb erweitert man $\mathbb{N}_0$ nochmals.

2.3.1 Definition (ganze Zahlen)

Eine reelle Zahl x heißt *ganz*, falls $x \in \mathbb{N}$ oder $x = 0$ oder $-x \in \mathbb{N}$ gilt.

$\mathbb{Z} := \{x \in \mathbb{R}; x \text{ ganz}\}$ heißt *Menge der ganzen Zahlen*. In aufzählender Schreibweise also

$$\mathbb{Z} = \{\, \dots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \dots \,\}$$

Die auf $\mathbb{R}$ erklärte Addition und Multiplikation vererben sich auf $\mathbb{Z}$.

$\mathbb{Z}$ ist bezüglich der Addition eine *abelsche Gruppe* (mit dem neutralen Element Null) und bezüglich der Addition und Multiplikation ein kommutativer *Ring* (mit den neutralen Elementen 0 und 1).

In einem kommutativen Ring mit Einselement gelten die Axiome (A_1) bis (A_4), $(M_1), (M_2), (M_3)$ und (D). Allerdings braucht (M_4) nicht zu gelten.

Ein kommutativer Ring unterscheidet sich von einem Körper dadurch, dass nicht alle von Null verschiedenen Elemente ein Inverses (bezüglich der Multiplikation) besitzen.

So sind z. B. 1 und -1 die einzigen Elemente in $\mathbb{Z}$, die bezüglich der Multiplikation ein Inverses in $\mathbb{Z}$ besitzen.

Diesem Mangel wird abgeholfen durch Einführung der *rationalen Zahlen*.

2.3.2 Definition (rationale Zahlen)

Die Menge

$$\mathbb{Q} := \left\{\, x \in \mathbb{R}\,;\ \text{Es gibt } m \in \mathbb{Z},\, n \in \mathbb{Z},\, n \neq 0, \text{ mit } x = \frac{m}{n} \,\right\}$$

heißt Menge der *rationalen Zahlen*.

Eine reelle Zahl, die nicht rational ist, heißt *irrational*.

Als Nenner in der Darstellung $x = \frac{m}{n}$ kann man statt der ganzen Zahl $n \neq 0$ stets eine natürliche Zahl wählen, weil man ein Vorzeichen im Nenner in den Zähler verlagern kann. Man beachte, dass die Darstellung einer rationalen Zahl in der Form $x = \frac{m}{n},\ m, n \in \mathbb{Z},\ n \neq 0$, nicht eindeutig ist. Nach den Regeln der Bruchrechnung aus 1.1.6 gilt $\frac{m}{n} = \frac{m'}{n'}$ genau dann, wenn $mn' = nm'$ $(m', n' \in \mathbb{Z}, n' \neq 0)$ ist.

2.3.3 Übungsaufgabe

Zeigen Sie, dass $\mathbb{Q}$ mit den von $\mathbb{R}$ geerbten Operationen Addition und Multiplikation ein angeordneter Körper ist, in dem auch die archimedische Eigenschaft gilt.

Wie wir schon früher bemerkt haben, gibt es keine rationale Zahl r mit $r^2 = 2$. Die Menge $M := \{y \in \mathbb{Q}; y \geq 0, y^2 \leq 2\}$ besitzt kein Supremum s_0 in $\mathbb{Q}$, weil für dieses $s_0^2 = 2$ gelten müsste. Vergleiche hierzu den Beweis von Satz 1.3.1.

Trotz dieser Lückenhaftigkeit der rationalen Zahlen werden wir gleich zeigen, dass $\mathbb{Q}$ in einem gewissen Sinne „dicht" in $\mathbb{R}$ ist. Dazu benötigen wir einen aus der archimedischen Eigenschaft von $\mathbb{R}$ folgenden Satz.

2.3.4 Satz

Zu jedem $x \in \mathbb{R}$ existieren eindeutig bestimmte ganze Zahlen p und q mit

(a) $p \leq x < p+1$ bzw.

(b) $q-1 < x \leq q$.

Bezeichnungen:

$p = [x] = \lfloor x \rfloor$ = floor x = größte ganze Zahl kleiner oder gleich x. $[x]$ heißt auch *Gauß-Klammer* von x.

$q = \lceil x \rceil$ = ceil x = kleinste ganz Zahl größer oder gleich x.

Beispiele:

$$\begin{aligned} [\pi] &= 3\ , & [-\pi] &= -4 \\ \lceil \pi \rceil &= 4\ , & \lceil -\pi \rceil &= -3\ , \end{aligned}$$

dabei sei $\pi = 3,141519265\ldots$.

Bemerkung:

Für $x \in \mathbb{R}$ gilt: $\lfloor x \rfloor = \lceil x \rceil \Leftrightarrow x \in \mathbb{Z}$.

Beweis : Wir beweisen nur (a).

Wir nehmen zunächst $x \geq 0$ an. Nach der archimedischen Eigenschaft gibt es ein $n \in \mathbb{N}$ mit $n > x$ und nach dem Wohlordnungssatz existiert dann

$$m := \min\{n \in \mathbb{N}; n > x\}.$$

Definiert man daher $p := m-1$, dann gilt

$$p \leq x < p+1.$$

Diese ganze Zahl p ist eindeutig bestimmt. Denn wenn für eine ganze Zahl $q \neq p$ auch

$$q \leq x < q+1$$

gilt, dann können wir ohne Einschränkung der Allgemeinheit $q < p$ annehmen.

Dann folgt aber

$$\begin{aligned} q+1 \leq p &\leq x < q+1, \text{ also} \\ q+1 &< q+1. \end{aligned}$$

Daher muss $p = q$ sein.

Dem Fall $x < 0$ führt man auf den eben behandelten Fall dadurch zurück, dass man ein geeignetes $k \in \mathbb{Z}$ zu x addiert, so dass

$$x + k \geq 0$$

gilt. Ein solches k existiert wieder nach der archimedischen Eigenschaft.

$\square$

2.3.5 Satz ($\mathbb{Q}$ liegt dicht in $\mathbb{R}$)

Sind $a, b \in \mathbb{R}$ und gilt $a < b$, dann gibt es im Intervall $]a, b[$ sowohl rationale als auch nicht rationale (= irrationale) Zahlen.
Hieraus folgt: Für jedes $x \in \mathbb{R}$ und jedes $\varepsilon > 0$ gibt es ein $r \in \mathbb{Q}$ mit $|r - x| < \varepsilon$.

Beweis : Wir wählen nach der Archimedischen Eigenschaft zunächst eine natürliche Zahl $n > \frac{2}{b-a}$ oder $\frac{1}{n} < \frac{b-a}{2}$ und suchen ein $m \in \mathbb{Z}$ mit

$$a < \frac{m}{n} < \frac{m+1}{n} < b.$$

Eine geeignete Wahl ist $m := [n \cdot a] + 1$, also $m \in \mathbb{Z}$ und $m - 1 = [n \cdot a] \leq n \cdot a < m$.

Es ist dann nämlich

$$a < \frac{m}{n} < \frac{m+1}{n} = \frac{m-1}{n} + \frac{2}{n} < a + (b - a) = b.$$

Also liegen die rationalen Zahlen

$$r_1 := \frac{m}{n} \qquad \text{und} \qquad r_2 = \frac{m+1}{n} \qquad \text{in} \qquad]a, b[\ ,$$

damit auch $r_3 := \frac{r_1 + r_2}{2}$.

Es ist damit offensichtlich, dass sogar „beliebig viele" rationale Zahlen in $]a, b[$ liegen.

Sind r_1 und r_2 die obigen rationalen Zahlen, dann ist für jedes $k \in \mathbb{N}$

$$x_k := r_1 + \frac{r_2 - r_1}{k\sqrt{2}}$$

eine irrationale Zahl mit $r_1 < x_k < r_2$ (weil sonst $\sqrt{2}$ rational wäre).

Zwischen je zwei rationalen Zahlen liegen also „unendlich viele" irrationale Zahlen, insbesondere liegen wegen $[r_1, r_2] \subset]a, b[$ unendlich viele irrationale Zahlen im Intervall $]a, b[$.

□

Bemerkung: Aus 2.3.4 folgt, dass sich jede reelle Zahl x eindeutig in der Form

$$x = [x] + \rho\ , \qquad x \in \mathbb{Z} \text{ und } \rho \in \mathbb{R}\ ,\ 0 \leq \rho < 1$$

darstellen lässt.

Hieraus ergibt sich ein einfacher Beweis eines wichtigen Satzes der elementaren Zahlentheorie:

2.3.6 Satz (Division mit Rest)

Zu einer ganzen Zahl n und einer positiven ganzen Zahl g gibt s eindeutig bestimmte ganze Zahlen q, r mit

$$n = q\,g + r\ , \qquad 0 \leq r < g\ .$$

r heißt der *reduzierte Rest modulo* g.
Zum Beweis braucht man nur die obige Bemerkung auf $:= \frac{n}{g}$ anzuwenden. Offensichtlich ist dann $q := \left[\frac{n}{g}\right]$ und $r := n - qg$.

2.4 Beispiele zum Beweisprinzip der vollständigen Induktion

Im Folgenden sind einige Beispiele für Beweise mit Hilfe des Beweisprinzips der „vollständigen Induktion“ zusammengestellt.

Man kann es anwenden, wenn eine Aussage $A(n)$ für alle natürlichen Zahlen n bewiesen werden soll. Induktionsbeweise spielen in allen Zweigen der Mathematik eine wichtige Rolle. Wir beschränken uns auf Beispiele aus der Analysis, Zahlentheorie und Geometrie. Wir werden dabei auch Varianten des Induktionsprinzips kennen lernen. Ferner kann man das Prinzip der vollständigen Induktion auch zur rekursiven Definition einsetzen.

(1) Summenformeln

Beweise von Summenformeln gelten allgemein als Inbegriff des Induktionsbeweises, dabei dienen sie nun zum Einüben des Induktionsschemas. Entdeckendes Lernen kann man dabei nicht praktizieren, zumal die zu beweisenden Formeln meist wie vom Heiligen Geist gegeben erscheinen. Der Weg zur Formel liefert meist schon den Beweis, das Induktionsschema ist dann meist umständlicher.

Wir geben dennoch einige typische Beispiele

1. Für alle $n \in \mathbb{N}$ gilt

$$A(n): \quad \boxed{1+2+3+4+\ldots n = \frac{n(n+1)}{2}}$$

Beweis : *Induktionsanfang*: $A(1)$ ist wahr:

$$A(1): \qquad 1 = \frac{1\cdot(1+1)}{2}\ .$$

Induktionsschritt: Wir zeigen:

Für alle $n \in \mathbb{N}$ gilt: $A(n) \Rightarrow A(n+1)$.

Wir betrachten

$$\begin{aligned}
\underbrace{1+2+\cdots+n}_{\substack{=n(n+1)/2\\ \text{Induktionsvoraussetzung !}}} \quad +(n+1) &= \frac{1}{2}n(n+1)+(n+1)\\
&= (n+1)\left[\frac{1}{2}n+1\right]\\
&= (n+1)\frac{n+2}{2}\\
&= \frac{(n+1)(\ (n+1)\ +1\ )}{2}\ .
\end{aligned}$$

Also ist $1+2+\cdots+n+(n+1) = \dfrac{(n+1)(n+2)}{2}$.

Wir haben aus $A(n)$ die Aussage $A(n+1)$ gefolgert.

Nach dem Induktionssatz gilt dann $A(n)$ für alle n.

□

Nach C. F. Gauß[1] (– der diese Summe als siebenjähriger Schüler im Spezialfall $n = 100$ nach anderen Quellen für $n = 60$ berechnet hat –) kann man so schließen:

Ist

$$\begin{aligned} s &:= 1 + 2 + \cdots + n\ , && \text{so ist auch} \\ s &= n + (n-1) + \cdots + 1\ , && \text{und durch Addition der Gleichungen folgt} \\ 2s &= \underbrace{(n+1) + (n+1) \cdots + \ldots (n+1)}_{n \text{ mal}} \\ &= n(n+1)\ , && \text{also } s = \frac{n(n+1)}{2}\ . \end{aligned}$$

2. Für alle $n \in \mathbb{N}$ gilt die Aussage $A(n)$: $1 + 3 + 5 + \cdots + (2n-1) = n^2$.

Beweis : *Induktionsanfang*: $A(1)$ ist wahr: $1 = 1^2$

Induktionsschritt: Für alle $n \in \mathbb{N}$ zeigen wir:

$$A(n) \Rightarrow A(n+1)\ .$$

Es ist

$$\underbrace{1 + 3 + 5 \cdots + (2n-1)}_{\substack{=n^2 \\ \text{nach Induktionsvoraussetzung !}}} + (2n+1) = n^2 + (2n+1) = (n+1)^2\ .$$

also gilt $A(n+1)$.

Damit gilt die Summenformel für alle $n \in \mathbb{N}$.

□

Wie man die letzten beiden Summenformeln völlig elementar mit Hilfe von *konfigurierten Zahlen* beweisen kann, wurde in der Vorlesung gezeigt.

3. Für alle $n \in \mathbb{N}$ gilt die Aussage

$$A(n)\ : \qquad 1^3 + 2^3 + \cdots + n^3 = \left(\frac{n(n+1)}{2}\right)^2 = (1 + 2 + \cdots + n)^2$$

Die rechte Gleichheit gilt wegen 1.

Induktionsanfang: $A(1)$: $1^3 = \left(\frac{1 \cdot 2}{2}\right)^2 = 1^2$ ist wahr.

Induktionsschritt: Wir zeigen $A(n) \Rightarrow A(n+1)$:

Betrachte

$$\begin{aligned} \underbrace{1^3 + 2^3 + \cdots + n^3}_{\substack{=n^2(n+1)^2/4 \\ \text{nach Induktionsvoraussetzung !}}} + (n+1)^3 &= \frac{1}{4}n^2(n+1)^2 + (n+1)^3 = (n+1)^2 \left[\frac{1}{4}n^2 + (n+1)\right] \\ &= (n+1)^2 \frac{n^2 + 4n + 4}{4} = (n+1)^2 \frac{(n+2)^2}{4} \\ &= \left(\frac{(n+1)(n+2)}{2}\right)^2\ . \end{aligned}$$

Ein Computeralgebrasystem berechnet solche Formeln schneller. Wir haben sie dennoch bewiesen, weil wir solche Formeln in der Integralrechnung benötigen.

Der kurze `maple`–Code für die Berechnung der Formel für $1^p + 2^p + \cdots + n^p$ als geschlossener Ausdruck in Abhängigkeit von n einschließlich Faktorisierung ist:

[1] Carl Friedrich Gauß: ∗ 30 April 1777 in Braunschweig,† 23 Februar 1855 in Göttingen

```
for p from 1 to 10 do
  print(p,   factor( sum('k'^p,'k'=1..n) ) );
od;
```

Die Ergebnisse sind dann:

$$
\begin{array}{ll}
1, & \frac{1}{2}\, n\,(n+1) \\[2mm]
2, & \frac{1}{6}\, n\,(n+1)\,(2\,n+1) \\[2mm]
3, & \frac{1}{4}\, n^2\,(n+1)^2 \\[2mm]
4, & \frac{1}{30}\, n\,(n+1)\,(2\,n+)\,(3\,n^2+3\,n-1) \\[2mm]
5, & \frac{1}{12}\, n^2\,(2\,n^2+\;n-1)\,(n+1)^2 \\[2mm]
6, & \frac{1}{42}\, n\,(n+1)\,(2\,n+)\,(3\,n^4+6\,n^3-3\,n+1) \\[2mm]
7, & \frac{1}{24}\, n^2\,(3\,n^4+\;n^3-n^2-4\,n+2)\,(n+1)^2 \\[2mm]
8, & \frac{1}{90}\, n\,(n+1)\,(2\,n+)\,(5\,n^6+15\,n^5+5\,n^4-15\,n^3-n^2+9\,n-3) \\[2mm]
9, & \frac{1}{20}\, n^2\,(n^2+n-)\,(2\,n^4+4\,n^3-n^2-3\,n+3)\,(n+1)^2 \\[2mm]
10, & \frac{1}{66}\, n\,(2\,n+1)\,(n+)\,(n^2+n-1)\,(3\,n^6+9\,n^5+2\,n^4-11\,n^3+3\,n^2+10\,n-5) \\[2mm]
11, & \frac{1}{24}\, n^2\,(2\,n^8+\;n^7+4\,n^6-16\,n^5-5\,n^4+26\,n^3-3\,n^2-20\,n+10)\,(n+1)^2
\end{array}
$$

Nur eine leichte LaTeX–Anpassung wurde vorgenommen. Der LaTeX–Code wurde von `maple` exportiert. Jacob Bernoulli (1654-1705) kannte schon die Struktur für die Summenformeln

$$\boxed{S_n^p := 1^p + 2 + p + \cdots + n^p} \qquad p, n \in \mathbb{N}\,.$$

Er zeigte nämlich, dass gilt:

$$\boxed{S_n^p = \frac{1}{p+1} n^{p+1} + \frac{1}{2} n^p + c_{p-1} n^{p-1} + c_{p-2} n^{p-2} + \cdots + c_1 n\,,}$$

wobei $c_{p-1}, c_{p-2}, \ldots, c_1$ rationale Zahlen sind, die eng mit den nach ihm benannten Bernoullischen Zahlen zusammenhängen.

4. *Geometrische Summenformel:*
   Für alle $n \in \mathbb{N}$ und alle $q \in \mathbb{R}$ gilt

$$A(n): \quad \boxed{(1-q)(1+q\ldots+q^n) = 1-q^{n+1}}$$

Beweis :

$A(1): (1-q)(1+q) = 1-q^2$.
Wir zeigen: Für alle $n \in \mathbb{N}$ gilt: $A(n) \Rightarrow A(n+1)$

Betrachte:

$$\begin{aligned}(1-q)(1+\cdots+q^n+q^{n+1}) &= (1-q)(1+\cdots+q^n) \quad + \quad (1-q)q^{n+1}\\ &= 1-q^{n+1} \quad + \quad q^{n+1}-q^{n+2}\\ &= 1-q^{n+2}\ ,\end{aligned}$$

d.h. aus $A(n)$ haben wir $A(n+1)$ gefolgert.

$\square$

Bemerkung: Hier bietet sich natürlich auch ein direkter Beweis an:

$$\begin{aligned}(1-q)(1+q+\cdots+q^n) = 1+q+q^2+\cdots+q^n&\\ -q-q^2-\cdots-q^n-q^{n+1} &= 1-q^{n+1}\ .\end{aligned}$$

Als letzte Summenformel beweisen wir die *binomische Formel*. Dazu sind einige Vorbereitungen erforderlich:

2.4.1 Definition (Binomialkoeffizienten)

Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ definiert man den *Binomialkoeffizienten* $\binom{\alpha}{k}$ durch

$$\binom{\alpha}{k} := \begin{cases} 1 & \text{für } k=0\ ,\\ \frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{1\cdot 2\cdot\ \ldots\ \cdot k} & \text{für } k\geq 1\ .\end{cases}$$

Dann gilt für alle $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}_0$

$(*)$
$$\boxed{\binom{\alpha}{k}+\binom{\alpha}{k+1}=\binom{\alpha+1}{k+1}}$$

Beweis : Für $k=0$ gilt $\binom{\alpha}{0}+\binom{\alpha}{1}=1+\alpha=\alpha+1=\binom{\alpha+1}{0+1}$.

Für $k \geq 1$ ist

$$\binom{\alpha}{k+1}=\frac{\alpha(\alpha-1)\ldots(\alpha-1-k+1)(\alpha-k)}{1\cdot 2\cdot\cdots\cdot k\cdot(k+1)}=\binom{\alpha}{k}\cdot\frac{\alpha-k}{k+1}\ ,$$

daher

$$\binom{\alpha}{k}+\binom{\alpha}{k+1}=\binom{\alpha}{k}\left(1+\frac{\alpha-k}{k+1}\right)=\binom{\alpha}{k}\frac{\alpha+1}{k+1}=\binom{\alpha+1}{k+1}\ .$$

$\square$

Im Folgenden sei $\alpha = n \in \mathbb{N}$.

5. *Binomischer Satz (oder binomische Formel):* Für $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt

$$\boxed{(a+b)^n=\binom{n}{0}a^n+\binom{n}{1}a^{n-1}b+\binom{n}{2}a^{n-2}b^2+\cdots+\binom{n}{n}b^n}$$

Beweis : Induktiv nach n.

$$n = 1 : \qquad (a+b)^1 = a + b\ .$$

Sei also $n \geq 1$ und $A(n)$ die Aussage

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \cdots + \binom{n}{n}b^n\ .$$

Multipliziert man diese Rechnung mit $(a+b)$, verwendet das Distributivgesetz und die Additionsformel $(*)$, dann folgt:

$$\begin{aligned}(a+b)^{n+1} &= (a+b)(a+b)^n = a(a+b)^n + b(a+b)^n\\ &= \binom{n}{0}a^{n+1} + \binom{n}{1}a^nb + \binom{n}{2}a^{n-1}b^2 + \cdots + \binom{n}{n-1}a^2b^{n-1} + \binom{n}{n}ab^n\\ &\qquad + \binom{n}{0}a^nb + \binom{n}{1}a^{n-1}b^2 + \cdots + \binom{n}{n-2}a^2b^{n-1} + \binom{n}{n-1}ab^n + \binom{n}{n}b^{n+1}\\ &= \underbrace{\binom{n+1}{0}}_{=1}a^{n+1} + \underbrace{\binom{n+1}{1}}_{=\binom{n}{0}+\binom{n}{1}}a^nb + \underbrace{\binom{n+1}{2}}_{=\binom{n}{1}+\binom{n}{2}}a^{n-1}b^2 + \cdots + \underbrace{\binom{n+1}{n}}_{=\binom{n}{n-1}+\binom{n}{n}}ab^n + \underbrace{\binom{n+1}{n+1}}_{=1}b^{n+1}\ ,\end{aligned}$$

und das ist die Aussage $A(n+1)$. Man vergleiche hierzu auch den in der Vorlesung etwas anders geführten Beweis.

$\square$

Auf Grund des Additionstheorems $(*)$ kann man die Binomialkoeffizienten $\binom{n}{k}$ mit Hilfe des sogenannten Pascalschen Dreiecks berechnen. Der „kurze“ `maple`–Code:

```
x := n-> seq( binomial(n,i) , i=0..n );
for n from 0 to 17 do print( x(n) ); od;
```

liefert dann sofort das Ergebnis aus der Abbildung 7 an der Seite 53.
Es hat die *Eigenschaften*:

1. Die erste und die letzte Zahl jeder Zeile ist 1.
2. Die $(k-1)$-te und die k-te Zahl in der n-ten Zeile haben als Summe die k-te Zahl in der $(n+1)$-ten Zeile.
3. Die Zahlen des Pascalschen Dreiecks sind „symmetrisch zur Höhe“, das bedeutet

$$\binom{n}{k} = \binom{n}{n-k}\ .$$

Die Eigenschaft 1. ist die Tatsache, dass $\binom{n}{0} = \binom{n}{n} = 1$ gilt.

2. ist gerade das Additionstheorem $(*)$. Um 3. einzusehen, schreiben wir die Binomialkoeffiziente $\binom{n}{k}$ ein bisschen um. Definiert man für $n \in \mathbb{N}$

$$n! = 1 \cdot 2 \cdot 3 \ldots n \ \text{ und } \ 0! = 1$$

(gesprochen n Fakultät), dann ist offensichtlich $n!$ das Produkt der ersten n natürlichen Zahlen

$$\begin{aligned}\binom{n}{k} &= \frac{n(n-1)\cdots(n-k+1)}{1\cdot 2\cdots k}\\ &= \frac{n(n-1)\cdots(n-k+1) \qquad (n-k)\cdots 1}{1\cdot 2\cdots k \qquad (n-k)\cdots 1}\\ &= \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!k!} = \frac{n!}{(n-k)!(n-(n-k))!}\\ &= \binom{n}{n-k}\ .\end{aligned}$$

```
1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
1 8 28 56 70 56 28 8 1
1 9 36 84 126 126 84 36 9 1
1 10 45 120 210 252 210 120 45 10 1
1 11 55 165 330 462 462 330 165 55 11 1
1 12 66 220 495 792 924 792 495 220 66 12 1
1 13 78 286 715 1287 1716 1716 1287 715 286 78 13 1
1 14 91 364 1001 2002 3003 3432 3003 2002 1001 364 91 14 1
1 15 105 455 1365 3003 5005 6435 6435 5005 3003 1365 455 105 15 1
1 16 120 560 1820 4368 8008 11440 12870 11440 8008 4368 1820 560 120 16 1
1 17 136 680 2380 6188 12376 19448 24310 24310 19448 2376 6188 2380 680 136 17 1
```

Abbildung 7: Das PASCALsche Dreieck

Kleine Tabelle für n!

$$\begin{aligned}
1! &= 1\\
2! &= 2\\
3! &= 6\\
4! &= 24\\
5! &= 120\\
6! &= 720\\
7! &= 5040\\
8! &= 40320\\
9! &= 362880\\
10! &= 3628800\\
11! &= 39916800\\
12! &= 479001600\\
13! &= 6227020800\\
14! &= 87178291200\\
15! &= 1307674368000\\
16! &= 20922789888000\\
17! &= 355687428096000\\
18! &= 6402373705728000\\
19! &= 121645100408832000\\
20! &= 2432902008176640000
\end{aligned}$$

Man sieht, dass $n!$ sehr schnell wächst. $100!$ hat z. B. schon 158 Dezimalstellen, es ist $100! = 9,3326\ldots \cdot 10^{157}$.
Die symmetrische Gruppe S_n hat die Ordnung $n!$. S_n ist die Gruppe der bijektiven Abbildungen von $\{1,2,\ldots,2\} \to \{1,2,\ldots,2\}$.

Bemerkung: Kombinatorische Interpretationen von $\binom{n}{k}$.
$\binom{n}{k}$ lässt sich interpretieren als die Anzahl der verschiedenen Teilmengen mit k Elementen in einer Menge mit n Elementen, z.B. der Menge $A_2 = \{\ 1,2,\ldots,n\ \}$.

Orientiert man sich an

1. $(a+b)^2 = a^2 + 2ab + b^2$ bzw.
2. $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$

und betrachtet dazu $(a+b)^n = \underbrace{(a+b)(a+b)\ldots(a+b)}_{n \text{ Faktoren}}$, dann wird offensichtlich, dass $(a+b)^n$ eine Summe von Zahlen der Gestalt $a^{n-k}\,b^k$ ergibt.

Jeder dieser Terme taucht genau so oft auf, wie man aus den n Faktoren k Faktoren auswählen kann und diese Anzahl ist $\binom{n}{k}$. Das ist auch ein Beweis des binomischen Satzes!

Folgerungen:

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

oder ohne Verwendung des Summenzeichens:

$$\boxed{\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n}$$

Kombinatorische Interpretation: Eine Menge mit n Elementen hat genau 2^n Teilmengen.
Durch Spezialisierung $a = 1$, $b = -1$ erhält man aus der binomischen Formel:

$$\boxed{\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots + (-1)^n \binom{n}{n} = 0}$$

2 Ungleichungen

Die folgende nach Jacob Bernoulli (1689) benannte, aber schon vorher bekannte Ungleichung werden wir häufig anwenden:

2.4.2 Satz (Bernoullische Ungleichung)

Ist $h \in \mathbb{R}$ und $h \geq -1$, dann gilt für alle $n \in \mathbb{N}$:

$$(1+h)^n \geq 1 + nh \; .$$

Zusatz: Für $h > -1, h \neq 0$ und $n \geq 2$ gilt die strikte Ungleichung:

$$(1+h)^n > 1 + nh \; .$$

Beweis : Induktion nach n
$n = 1$: $(1+h)^1 = 1 + 1 \cdot h$

Induktionsschritt: Es gelte für beliebiges $n \in \mathbb{N}$

$$(1+h)^n \geq 1 + nh.$$

Dann folgt durch Multiplikation mit der nicht negativen Zahl $1 + h$

$$(1+h)^{n+1} \geq (1+nh)(1+h) = 1 + (n+1)h + nh^2 \geq 1 + (n+1)h \; .$$

□

Folgerung: Ist $g \in \mathbb{R}$ und $g > 0$. Dann gilt

(a) ist $g > 1$, dann gibt es zu jedem $C \in \mathbb{R}$ ein $n \in \mathbb{N}$ mit $g^n > C$.

(b) Ist $0 < g < 1$, dann gibt es zu jedem $\varepsilon > 0$ ein $n \in \mathbb{N}$ mit $g^n < \varepsilon$

Beweis : (a) Sei $h := g-1$, dann ist nach Voraussetzung $h > 0$ und die Bernoullische Ungleichung liefert $g^n = (1+h)^n \geq 1 + nh$.

Nach der Archimedischen Eigenschaft von $\mathbb{N}$ gibt es ein $n \in \mathbb{N}$ mit $nh > C - 1$. Für dieses n ist dann $g^n > C$.

(b) Setzt man $g_1 = \dfrac{1}{g}$, dann ist $g_1 > 1$ und nach (a) gibt es zu $C := \dfrac{1}{\varepsilon}$ ein $n \in \mathbb{N}$ mit $g_1^n > \dfrac{1}{\varepsilon}$, hieraus folgt aber

$$g^n < \varepsilon \; .$$

□

Frage: Für welche $n \in \mathbb{N}$ gilt $2^n > n^2$?

Kleine Tabelle:

n	1	2	3	4	5	6	7	8
n^2	1	4	9	16	25	36	49	64
2^n	2	4	8	16	32	64	128	256

Vermutung: Für alle $n \geq 5$ gilt
$$2^n > n^2.$$

Wir verwenden zum Beweis eine *Variante des Induktionssatzes:*
Für jedes $n \in \mathbb{N}$ sei eine Behauptung $A(n)$ gegeben.
Gilt dann

(1) $A(n_0)$ für ein $n_0 \in \mathbb{N}$

(2) Für alle $n \geq n_0$ gilt die Implikation $A(n) \Rightarrow A(n+1)$,
dann ist $A(n)$ für alle $n \geq n_0$ wahr.
Zum Beweis braucht man die Menge $M = \{ \, n \in \mathbb{N} \, ; \, A(n_0 - 1 + n)$ ist wahr $\}$ zu betrachten. Dies ist eine Zählmenge (Beweis!) und in $\mathbb{N}$ enthalten, also gilt $M = \mathbb{N}$, das bedeutet, dass $A(n)$ für alle $n \geq n_0$ wahr ist.

Jetzt zu unserem Beispiel zurück:

Als *Induktionsverankerung* nehmen wir

$$n_0 = 5 \ : \qquad \text{Es gilt } 2^{n_0} = 2^5 = 32 > 25 = n_0^2 \, .$$

$A(n)$ sei die Aussage $2^n > n^2$.

Wir müssen zeigen für alle $n \geq 5$ gilt:

$$A(n) \Rightarrow A(n+1)$$

Durch Multiplikation der Ungleichung von $2^n > n^2$ mit 2 folgt

$$2^{n+1} > 2n^2 \, .$$

Wir müssen also nur noch die Ungleichung

$$2n^2 \geq (n+1)^2$$

beweisen. Dies geschieht durch Äquivalenzumformung:

$$\begin{aligned} 2n^2 \geq (n+1)^2 \qquad &\Leftrightarrow \qquad 2n^2 \geq n^2 + 2n + 1 \\ &\Leftrightarrow \qquad n^2 - 2n + 1 \geq 2 \\ &\Leftrightarrow \qquad (n-1)^2 \geq 2. \end{aligned}$$

Diese Ungleichung gilt somit für alle $n \geq 3$, ist also erst recht für alle $n \geq 5$ richtig.

Bemerkung: Die Induktionsverankerung ist wesentlicher Bestandteil des Beweisprinzips der vollständigen Induktion:

Die Implikation
$$A(n) \Rightarrow A(n+1)$$
kann richtig sein, aber die Aussage $A(n)$ gilt nicht für allen $n \in \mathbb{N}$, wenn es keine Induktionsverankerung gibt. Als Beispiel betrachten wir für $n \in \mathbb{N}$ die Aussage

$$A(n) : \qquad 1 + 2 + \cdots + n = \frac{n(n+1)}{2} + 2002 \, .$$

Für jedes $n \geq n_0 := 1$ gilt zwar

$$A(n) \Rightarrow A(n+1) \ :$$

Falls für ein $n \in \mathbb{N}$ die Gleichheit $1 + 2 + \cdots + n = \frac{n(n+1)}{2} + 2002$ wahr ist, dann folgt:

$$1 + 2 + \cdots + n + (n+1) = \frac{n(n+1)}{2} + 2002 + (n+1) = \frac{(n+1)(n+2)}{2} + 2002 \ .$$

Die Aussage $A(n)$ ist trotzdem falsch (siehe Beispiel weiter oben), der „Induktionsbeweis“ versagt, weil $A(1)$ nicht gilt.

3 Rekursive Definitionen, Summen, Produkte

In den neun Körperaxiomen (A_1) bis (A_4), (M_1) bis (M_4) und (D) tritt neben den Summen von zwei Elementen auch die Summe von drei Elementen auf:

$$(a+b)+c$$

nach (A_2) gilt

$$(a+b)+c = a+(b+c)$$

und wir hatten vereinbart, für diesen von der Klammerung (und Reihenfolge!) unabhängigen Ausdruck einfach

$$a+b+c$$

zu schreiben.

Hat man (für $n \in \mathbb{N}$) n reelle Zahlen $a_1, \ldots, a_n$ zu addieren, so liegt die folgende Vorgehensweise nahe:

Man definiert (so, wie man auf einem Taschenrechner eine echte Summe ausrechnet)

$$\begin{array}{rcl} s_1 & := & a_1 \\ s_2 & = & s_1 + a_2 = a_1 + a_2 \\ s_3 & = & s_2 + a_3 = (a_1 + a_2) + a_3 \\ \vdots & & \vdots \\ s_n & = & s_{n-1} + a_n \end{array}$$

Das ist eine rekursive Definition im Sinne von 3.1

Für $a \in \mathbb{R}$, $n \in \mathbb{N}$ haben wir die Definition

$$a^n = \underbrace{a \cdot a \cdot a \cdots \cdots a}_{n\text{-mal}} \quad (n \geq 2)$$

verwendet.

Die Pünktchenschreibweise ist aber unbefriedigend. Eine bessere Definition wäre gewesen:

$$(*) \begin{cases} a^1 = a \text{ und für } n \in \mathbb{N} \\ a^{n+1} = a^n \cdot a \\ \text{und ergänzend} \\ a^0 := 1 (\text{insbesodere } 0^0 = 1) \end{cases}$$

Die ist wieder eine rekursive Definition im Sinne von 3.1.

Um aus $(*)$ eine exakte Definition (Definition durch Induktion) zu machen, verwendet man das Induktionsprinzip, auch zur Definition von Objekten $D(n)$, $n \in \mathbb{N}$.

Wir wollen dies am Beispiel der rekursiven Definition von a^n erläutern.

Sei $a \in \mathbb{R}$ eine feste reelle Zahl. Für $n \in \mathbb{N}$ sei $E(n)$ die folgende Behauptung:
Es gibt zu jedem $k \in \mathbb{N}$ mit $k \leq n$ genau eine reelle Zahl a_k, so dass folgende beiden Eigenschaften erfüllt sind:
Es gilt $a_1 = a$, und für alle $k \in \mathbb{N}$ mit $k < n$ gilt

$$a_{k+1} = a_k \cdot a.$$

Wir wollen $E(n)$ für alle $n \in \mathbb{N}$ beweisen und benutzen dazu vollständige Induktion nach n.

$E(1)$: Man braucht noch $a_1 := a$ zu setzen und $E(1)$ lässt sich auch nur mit diesem a_1 erfüllen.

Induktionsschritt:
Setz man $E(n)$ als richtig voraus, nimmt also an, dass für alle $k \in \mathbb{N}$ mit $k \leq n$ Zahlen a_k mit den geforderten Eigenschaften existieren, dann kann man $E(n+1)$ wie folgt beweisen:

Für alle $k \in \mathbb{N}$ mit $k \leq n$ setze man $b_k := a_k$, und außerdem $b_{n+1} = a_n \cdot a$.
Dann ist offensichtlich $b_1 = a_1 = a$, und für $k < n$ ist auch $b_{k+1} = a_{k+1} = a_k \cdot a = b_k a$.
Für $k = n$ ist aber gerade $b_{n+1} = a_n \cdot a = b_n \cdot a$ definiert. Die Zahlen b_k erfüllen also die Existenzbedingungen von $E(n+1)$, sie sind aber auch eindeutig bestimmt.
Betrachtet man nämlich die b_k für *alle* $k \leq n$, so sind sie geeignete Kandidaten für die Eigenschaft $E(n)$, stimmen also mit den a_k überein; ferner muss notwendig

$$b_{n+1} = b_n \cdot a = a_n \cdot a$$

sein.
Wir haben aber gesagt: Für alle $n \in \mathbb{N}$

$$E(n) \Longrightarrow E(n+1);$$

damit gilt $E(n)$ für alle $n \in \mathbb{N}$.
Schrieben wir, um die formale Abhängigkeit von a_k von n zu verdeutlichen, statt a_k etwas genauer $a_{k,n}$, dann haben wir folgendes gezeigt:
Für alle natürlichen Zahlen k und n mit $k \leq n$ gibt es eindeutig bestimmte reelle Zahlen $a_{k,n}$, so dass für alle $n \in \mathbb{N}$ folgende Eigenschaften erfüllt sind:

(a) Es ist $a_{1,n} = a$;

(b) Für alle $k < n$ ist $a_{k+1,n} = a_{k,n} \cdot a$;

(c) Für alle $k < n$ ist $a_{k,n} = a_{k,n-1}$

Wir können jetzt definieren: Für $n \in \mathbb{N}$ sei

$$a^n := a_{n,n}.$$

Dann ist

$$\begin{aligned} a^1 &= a_{1,1} = a \text{ und} \\ a^{n+1} &= a_{n+1,n+1} = a_{n,n+1} \cdot a \\ &= a_{n,n} \cdot a = a^n \cdot a \end{aligned}$$

Dann ist a^n für alle $n \in \mathbb{N}$ eindeutig festgelegt.

Definiert man ergänzend $a^0 := 1$, dann gilt $a^{n+1} = a^n \cdot a$ auch noch im Fall $n = 0$ und $a^1 = a$ ist automatisch erfüllt.

Das Ganze mag reichlich kompliziert erscheinen (gegenüber der „naiven“ Definition von a^n), ist aber exakt.

Man erweitert die Definition dann für $a \neq 0$ auf negative Potenzen durch die Festsetzung ($n \in \mathbb{N}$)

$$a^{-n} := (a^{-1})^n.$$

Dann gelten die folgenden *Rechenregeln*:

$(R_1): \ a^n a^m = a^{n+m}$

$(R_2): \ (a^n)^m = a^{nm}$

$(R_3): \ a^n b^n = (ab)^n,$

dabei sind $n, m \in \mathbb{Z}$ beliebig und a, b sind reelle Zahlen, die ungleich 0 vorauszusetzen sind, wenn negative Potenzen auftreten. Man kann alle Regeln nicht für $n \in \mathbb{N}_0$ mit Induktion nach n beweisen (zu fest) und den Fall $n < 0$ auf den behandelten Fall zurückführen.

In völliger Analogie kann man die Summe $a_1 + a_2 + \ldots + a_n$ von n reellen Zahlen $a_1, \ldots, a_n$ definieren und mit vollständiger Induktion zeigen, dass diese Summe nicht von der Beklammerung oder der Reihenfolge abhängt.

Diese Beispiele sind Spezialfälle für den *Rekursionssatz*.

3.1 Rekursionssatz

(R.Dedekind, 1888)

Ist B eine Menge und für jedes $n \in \mathbb{N}$ sei F_n eine Vorschrift, die jedem $x \in B$ genau ein $F_n(x) \in B$ zuordnet. Ferner sei ein Element $b \in B$ vorgegeben. Dann gibt es genau eine Zuordnung

$$\begin{array}{ccc} \mathbb{N} & \to & B \\ n & \mapsto & f(n), \end{array}$$

die folgende Bedingungen erfüllt:

(a) $f(1) = b$

(b) $f(n+1) = F_n(f(n))$ für alle $n \in \mathbb{N}$.

Bemerkung: Der Satz gilt auch dann, wenn man zur Berechnung von $f(n+1)$ auch *alle* vorhergehende Werte $f(1), f(2), \ldots, f(n)$ benötigt.

Wir verzichten hier auf einen Beweis und verweisen auf die Literatur, insbesondere Barner-Flohr, Analysis I, Abschnitt 2.3 oder zu „Zahlen“(Herausgeber: Ebbinghaus, dort Teil A, Kap. 1, §2) oder Amann-Escher: Analysis I, Satz 5.11 (siehe Literaturverzeichnis).

Für die Definition von Summe und Produkt von n Zahlen, $n \in \mathbb{N}$, wählt man

$$F_n(x) = x + a_n \quad \text{bzw.} \quad F_n(x) = a_n \cdot x = x \cdot a_n$$

Die rekursive Definition

$$\begin{aligned} p(1) &= 1 \\ p(n+1) &= (n+1)p(n); \quad n \geq 1, \end{aligned}$$

liefert explizit

$$p(n) = n!,$$

also die Fakultät.
Ergänzend definiert man $p(0) = 1$.

Auch die aufwendige und unbefriedigende „Pünktchenschreibweise“ beim Beweis der Summenformeln, wie z.B.

$$1^2 + 2^2 + 3^2 + \ldots + 100^2$$

lässt sich nun völlig vermeiden, wenn man für die rekursiv definierte Summe von n reellen Zahlen $a_1, \ldots, a_n$ das *Summenzeichen* verwendet:

$$\sum_{j=1}^{n} a_j := a_1 + a_2 + \ldots + a_n$$

(In Worten: „Summe a_j, j von 1 bis n“).

Die rekursive Definition lautet also

$$\begin{aligned} \sum_{j=1}^{1} a_j &= a_1 \\ \sum_{j=1}^{n+1} a_j &:= \left(\sum_{j=1}^{n} a_j \right) + a_{n+1}. \end{aligned}$$

Mit Induktion zeigt man, dass $\sum\limits_{j=1}^{n} a_j$ von Klammerung und der Reihenfolge der Summanden unabhängig ist. Die Bestandteile des Summenzeichens haben im Einzelnen die folgende Bedeutung:

$\swarrow$ Obere Grenze des Laufindexes

Laufindex $\longrightarrow$ $\sum\limits_{j=1}^{n} a_j$ $\longleftarrow$ Summationsterm

$\nwarrow$ untere Grenze des Laufindexes

Der Laufindex j kann durch jedes andere Symbol ersetzt werden (das nicht schon eine andere Bedeutung hat), das muss dann allerdings an allen Stellen geschehen, wo j auftritt, z.B.

$$\sum_{j=1}^{n} a_j = \sum_{k=1}^{n} a_k = \sum_{l=1}^{n} a_l.$$

Folgende Manipulationen mit dem Summenzeichen sind erlaubt:

(a) *Beliebige Grenzen in* $\mathbb{Z}$: Sei $k, l \in \mathbb{Z}$, n ist für $l \geq k$

$$\sum_{j=k}^{l} a_j = a_k + a_{k+1} + \ldots + a_{k-1} + a_k.$$

Ist $k < l$, so setzt mann

$$\sum_{j=k}^{l} a_j = 0.$$

Die Indexmenge, über die summiert wird, ist leer, es wird also gar nicht summiert, das hat bezüglich der Addition die gleiche Wirkung, wie die Addition von Null (frei nach Forster: Analysis I).

(b) *Aufspaltung der Summe*: Ist $1 \leq m \leq n$, so ist

$$\sum_{j=1}^{n} a_j = \sum_{j=1}^{m} a_j + \sum_{j=m+1}^{n} a_j \quad \text{(Assoziativgesetz)}$$

(c) *Linearität der Summe*: Für jedes $\lambda, \mu \in \mathbb{R}$ gilt, falls $b_j,\ 1 \leq j \leq n$, weitere gegebene reelle Zahlen sind:

$$\sum_{j=1}^{n}(\lambda a_j + \mu b_j) = \lambda \sum_{j=1}^{n} a_j + \mu \sum_{j=1}^{n} b_j$$

(d) *Umnummerierung der Indizes*: Ist $m \leq n$, so gilt

$$\sum_{j=1}^{n} a_j = \sum_{j=1}^{n-m+1} a_{m-1+j}.$$

Versuchen Sie nicht, diese Regel klein zu machen.
Frage: Wenn die Terme $a_m, a_{m+1}, \ldots, a_n$ zu addieren sind, wieviele Summanden ergibt das insgesamt?

An einer 20m langen Grenze sollen Zaunpfähle im Abstand von je einem Meter aufgestellt werden. Wieviele Pfähle braucht man?
Ersetzt man $m-1$ durch ein beliebiges $k \in \mathbb{Z}$, so erhält man allgemeiner:

(e)

$$\sum_{j=m}^{n} a_j = \sum_{j=m-k}^{n-k} a_{j+k}$$

Die mit vollständiger Induktion bewiesene Summenformel, schreibt sich jetzt eleganter, z.B.

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

oder

$$\sum_{k=1}^{n}(2k-1) = n^2$$

oder der Binomische Satz:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k \quad (a, b \in \mathbb{R}, n \in \mathbb{N}_0).$$

Völlig analoge Eigenschaften, wie für die Summe von n reellen Zahlen, gelten für das *Produkt* von n reellen Zahlen $a_1, \ldots, a_n$. Man definiert:

$$\prod_{j=1}^{n} a_j \qquad \text{rekursiv durch}$$

$$\begin{aligned} \prod_{j=1}^{1} a_j &= a_1 \text{ und} \\ \prod_{j=1}^{n+1} a_j &= \left(\prod_{j=1}^{n} a_j\right) \cdot a_{n+1}. \end{aligned}$$

Und ergänzend für $m,\ l \in \mathbb{Z}$ mit $m \geq l$

$$\begin{aligned} \prod_{j=l}^{m} a_j &= a_l \cdot a_{l+1} \cdot \ \ldots \ \cdot a_m \text{ und} \\ \prod_{j=l}^{m} a_j &= 1, \text{ falls } m < l. \end{aligned}$$

„Wenn mit gar nichts zu multiplizieren ist, dann hat das bezüglich der Multiplikation die selbe Wirkung, wie, wenn man mit dem neutralen Element 1 der Multiplikation multipliziert.“

Wir erwähnen noch das *Allgemeine Kommutativ- und Distributivgesetz*:

Sind $a_1, \ldots, a_m$ und $b_1, \ldots, b_n$ vorgegebene reelle Zahlen, dann gilt

$$\begin{aligned} \left(\sum_{j=1}^{m} a_j\right)\left(\sum_{k=1}^{n} b_k\right) &= \sum_{j=1}^{m}\left(\sum_{k=1}^{n} a_j b_k\right) \\ &= \sum_{k=1}^{n}\left(\sum_{j=1}^{m} a_j b_k\right) \\ &= \sum_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}} a_j b_k \end{aligned}$$

Die ersten beiden Gleichheitszeichen beinhalten das allgemeine Kommutativgesetz, das dritte Gleichheitszeichen das allgemeine Distributivgesetz.

Man hat dabei jedes Glied der ersten Summe mit jedem Glied der zweiten Summe zu multiplizieren und muss die entstehende Produkte $a_j b_k$ aufsummieren, z.B. nach folgendem Schema:

$$\begin{array}{c|c} a_1b_1 + a_1b_2 + \ldots + a_1b_n & a_1 \sum\limits_{k=1}^{n} b_k \\ +a_2b_1 + a_2b_2 + \ldots + a_2b_n & +a_2 \sum\limits_{k=1}^{n} b_k \\ \vdots & \vdots \\ +a_mb_1 + a_mb_2 + \ldots + a_mb_n & +a_m \sum\limits_{k=1}^{n} b_k \\ \hline & = \left(\sum\limits_{j=1}^{m} a_j\right)\left(\sum\limits_{k=1}^{n} b_k\right) \end{array}$$

Nach dem Kommutativgesetz und dem Distributivgesetz ist das Resultat unabhängig von der Reihenfolge der Summanden und Faktoren und von Klammerungen. Man verwendet dafür die Schreibweise als Doppelsumme:

$$\sum_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}} a_j b_k .$$

4 Komplexe Zahlen

4.0 Motivation, historische Bemerkungen

Wir haben die reellen Zahlen als angeordneten Körper eingeführt, in dem zusätzlich noch das Vollständigkeitsaxiom (V) gilt.

Letzteres hat uns die Existenz von Quadratwurzeln aus nicht negativen reellen Zahlen beschert (allgemeiner die Existenz vom k-ten Wurzeln $(k \in \mathbb{N},\ k \geq 2)$ aus nicht negativen reellen Zahlen). Diese Eigenschaft unterscheidet den Körper $\mathbb{Q}$ der rationalen Zahlen von den reellen Zahlen, in $\mathbb{Q}$ ist z.B. die Gleichung $x^2 = 3$ nicht lösbar. Aber ganz einfache algebraische Gleichungen, wie z.B.

$$\begin{array}{lll} (*) & x^2+1=0 & \quad \textit{oder} \\ (**) & x^2-2x+3=0 & \end{array}$$

besitzen keine Lösungen $x \in \mathbb{R}$. Der Grund ist, dass das Quadrat einer reellen Zahl immer größer oder gleich Null ist. *Aufgabe:* Skizzieren Sie die Graphen der Funktionen

$$\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^2+1 \end{array} \qquad \text{bzw.} \qquad \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^2-2x+3\,. \end{array}$$

Wendet man jedoch z. B. auf $(**)$ formal die Lösungsformel für quadratische Gleichungen an, so erhält man

$$x_{1,2} = 1 \pm \sqrt{1-3} = 1 \pm \sqrt{-2}\ ,$$

also „Quadratwurzeln aus negativen Zahlen“, solche „Zahlen“ können aber keine reellen Zahlen sein. In der „Ars magna“(1545 in Nürnberg erscheinen) von G. Cardano (1501–1576) findet sich folgende Aufgabe

Zerlege die Zahl 10 so in zwei Summanden, dass ihr Produkt 40 ergibt .

Nehmen wir an, es gibt solche Zerlegungen und nennen wir die eine Zahl x, die andere y, so soll also gelten

$$10 = x + y \qquad \text{und} \qquad 40 = xy\ .$$

Einsetzen von $y = 10 - x$ in die zweite Gleichung ergibt die quadratische Gleichung

$$x^2 - 10x + 40 = 0\ ,$$

welche bei formaler Anwendung die Lösungsformel

$$x_{1,2} = 5 \pm \sqrt{25-40} = 5 \pm \sqrt{-15},$$

also wieder Wurzeln aus negativen Zahlen, ergibt.

Für eine Lösung der kubischen Gleichung

$$\boxed{x^3 - 15x - 4 = 0}$$

gab R. Bombelli (um 1560) den Wert

$$x_1 := \sqrt[3]{2+\sqrt{-121}} + \sqrt[3]{2-\sqrt{-121}}$$

an. Unter Beachtung von 8 Regeln mit dem Rechnen mit Wurzeln aus negativen Zahlen berechnete er

$$x_1 = (2+\sqrt{-1}) + (2-\sqrt{-1}) = 4.$$

Als weitere Lösungen gab er

$$x_2 = -2 + \sqrt{3} \qquad \text{und} \qquad x_3 = -2 - \sqrt{3}$$

an.

Aufgabe: Bestätigen Sie durch Einsetzen in die Gleichung, dass es sich wirklich um Lösungen der Gleichung handelt. Gibt es noch weitere Lösungen?

Für viele Mathematiker war es ein Mysterium, dass man über den Umweg über „imaginäre“ Zahlen (nur in der Einbildung, aber nicht wirklich existierende Zahlen) zu richtigen Ergebnissen kam. Vom ersten Auftreten bei den italienischen Mathematikern der Renaissance bis zur endgültigen Etablierung und Akzeptierung in der Mathematik hat es dann noch ca. 300 Jahre gedauert.

Wichtige Wegbereiter waren u. a. Leonard Euler, der 1748 die Formeln

$$\cos v = \frac{e^{+v\sqrt{-1}} \quad + \quad e^{-v\sqrt{-1}}}{2} \qquad \text{und}$$
$$\sin v = \frac{e^{+v\sqrt{-1}} \quad - \quad e^{-v\sqrt{-1}}}{2\sqrt{-1}}$$

aufstellte, also einen Zusammenhang zwischen den Winkelfunktionen nd der Exponentialfunktion. Von Euler wurde auch 1777 für $\sqrt{-1}$ die Abkürzung i eingeführt.

Der Fachausdruck „komplexe Zahl“ stammt von C. Fr. Gauß (1831), die heute am häufigsten verwendete Einführung der komplexen Zahlen über Paare reeller Zahlen geht auf Sir W. R. Hamilton (1837) zurück. Unser Ziel ist es, die reellen Zahlen in ein noch größeres Zahlsytem, das wieder ein Körper werden soll, so einzubetten, dass man durch Einschränkung der auf dem größeren Körper definierten Addition bzw. Multiplikation die bekannten Opertionen auf $\mathbb{R}$ erhält. Dazu machen wir folgende

Annahme:
Es gibt einen Körper K, der den Körper $\mathbb{R}$ als Unterkörper enthält, und der ein Element i enthält, für das $i^2 = -1$ gilt. Wie hat man in einem solchen Körper zu rechnen?

Als erstes halten wir fest, dass $i \notin \mathbb{R}$ gilt.

Aufgrund der Körperaxiome enthält K dann auch alle „Zahlen“ der Gestalt

$$z := a + bi \qquad \text{mit } a, b \in \mathbb{R}$$

und die Darstellung von $z \in K$ in dieser Form ist eindeutig, denn ist auch $z = c + di$ mit $c, d \in \mathbb{R}$, so folgt $a + bi = c + di$, also ist $a - c = (d - b)i$. Wäre jetzt $d \neq b$, dann wäre $i = \frac{a-c}{d-b} \in \mathbb{R}$, ein Widerspruch zu $i \notin \mathbb{R}$.

Also gilt $d = b$ und damit auch $a = c$. Betrachtet man jetzt die Menge

$$C := \{a + bi \in K \; ; \; a, b \in \mathbb{R} \} \, ,$$

so stellt man fest, dass bereits C ein Körper bezüglich der auf K erklärten Addition und Multiplikation ist.

Wir prüfen nicht alle Körperaxiome nach (das muss man auch gar nicht! Warum?), wir zeigen nun: Für $z, w \in C$ gilt auch $z \pm w \in C$ und $zw \in C$ und für $z \neq 0$ ist auch $z^{-1} \in C$.

Davon seien $z = a + bi$ und $w = c + di$ mit $a, b, c, d \in \mathbb{R}$.

Dann ist

$$\begin{array}{lrcl} (*) & z \pm w & = & (a + bi) \pm (c + di) = (a + c) \pm (b + d)i, \quad \text{also } z \pm w \in C \text{ und} \\ (**) & zw & = & ac + bic + adi + bdi^2 \quad (i^2 = -1) \\ & & = & (ac - bd) + (ad + bc)i, \text{ also auch } zw \in C \,. \end{array}$$

Ist $z = a + bi \neq 0$ (d.h. $a \neq b$ oder $b \neq 0$), dann existiert z^{-1} in K (K ist Körper) und man erhält

$$\frac{1}{z} = \frac{1}{a + bi} = \frac{a - bi}{(a + bi)(a - bi)} = \frac{a - bi}{a^2 + b^2} \, ,$$

also

$$(***) \quad \frac{1}{z} = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2} i$$

und somit $1/z \in C$.

Als Zusammenfassung erhalten wir: Wenn es überhaupt einen Körper gibt, der die reellen Zahlen $\mathbb{R}$ als Unterkörper enthält und in welchem es ein Element i mit $i^2 = -1$ gibt, dann gibt es auch einen „kleinsten" solcher Körper, und diese minimale Erweiterung ist „bis auf Isomorphie" eindeutig bestimmt: Die Operationen Addition und Multiplikation und die Existenz von Inversen von Elementen ungleich Null, sind durch die obigen Formeln $(*), (**), (***)$ mit den Operationen auf $\mathbb{R}$ verknüpft.

Offen geblieben in unserer heuristischen Betrachtung ist die Frage „Was ist i"? Aber die Darstellung $z = a + bi,\ a, b \in \mathbb{R}$, legt nahe, dass als wesentliches Bestimmungsstück der nun zu definierenden komplexen Zahlen die reellen Zahlen a und b anzusehen sind, also das Paar $(a, b) \in \mathbb{R} \times \mathbb{R}$. Das folgende Modell für die komplexen Zahlen geht – wie oben erwähnt – *Hamilton* (1831) zurück.

4.1 Komplexe Zahlen (Konstruktion von $\mathbb{C}$)

4.1.1 Definition und Satz

(a) Eine *komplexe Zahl* ist ein Element aus der Menge

$$\mathbb{C} := \mathbb{R}^2 := \{(a, b);\ a \in \mathbb{R},\ b \in \mathbb{R}\},$$

also ein (geordnetes) Zahlenpaar:

$$(a, b) = (c, d) \qquad \Leftrightarrow \qquad (\ a = c \ \text{ und } \ b = d\ )\ .$$

(b) Definiert man auf $\mathbb{C}$ eine *Addition* für $(a, b) \in \mathbb{C}$ und $(c, d) \in \mathbb{C}$ durch:

(A) : $(a, b) + (c, d) := (a + c, b + d)$ und eine *Multiplikation* durch

(M) : $(a, b) \cdot (c, d) = (ac - bd, ad + bc)$, so ist $(\mathbb{C}, +, \cdot)$ ein Körper.

Neutrale Elemente sind $\underline{0} = (0, 0)$ (bzgl. Addition) und $\underline{1} = (1, 0)$ (bzgl. Multiplikation).
Das *Negative* zu (a, b) ist $(-a, -b)$, das Inverse zu $(a, b) \neq (0, 0)$ ist

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

(c) *Einbettung von $\mathbb{R}$ in $\mathbb{C}$.*

$\mathbb{C}$ enthält als Teilmenge eine Kopie von $\mathbb{R}$. Definiert man nämlich

$$\mathbb{C}_{\mathbb{R}} := \{\ (a, 0)\ ;\ a \in \mathbb{R}\ \},$$

dann addieren und multiplizieren sich die Elemente aus $\mathbb{C}_{\mathbb{R}}$ wie die entsprechenden reellen Zahlen, das bedeutet, es gilt $(a, b, c, a' \in \mathbb{R})$

$$\begin{array}{rcl} (a, 0) + (c, 0) & = & (a + c, 0) \\ (a, 0) \circ (c, 0) & = & (ac, 0) \\ & \text{und} & \\ (a, 0) = (a', 0) & \Leftrightarrow & a = a'\,. \end{array}$$

Die Abbildung $j : \mathbb{R} \to \mathbb{C}_{\mathbb{R}},\ a \mapsto (a, 0)$ ist daher ein Körperisomorphismus. Wir identifizieren daher $\mathbb{R}$ mit seiner Kopie $\mathbb{C}_{\mathbb{R}} \subset \mathbb{C}$, schreiben also statt dem Paar $(a, 0)$ einfach a. Ferner ergibt die Multiplikation von $(a, b) \in \mathbb{C} = \mathbb{R}^2$ mit der Zahl $(r, 0) \in \mathbb{C}_{\mathbb{R}}$ das Resultat

$$(r, 0)(a, b) = (ra, rb),$$

das Resultat stimmt also überein mit dem „Produkt" $r \cdot (a, b) := (ra, rb)$, das man bei der Multiplikation mit Skalaren im $\mathbb{R}$-Vektorraum $\mathbb{R} \times \mathbb{R}$ betrachtet. Die Multiplikation mit Skalaren aus $\mathbb{R}$ ist also durch die Multiplikation im Körper $\mathbb{C}$ festgelegt.

(d) *Die imaginäre Einheit*

Es gilt

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

Definiert man also $i := (0, 1) \in \mathbb{C}$, dann ist $i^2 = -1$.

Wir nennen i die *imaginäre Einheit*.
i ist also eine Lösung der Gleichung $z^2 + 1 = 0$.
Diese Gleichung wird nur noch von $-i$ gelöst.

(e) *Standarddarstellung komplexer Zahlen*

Ist $z = (a, b) \in \mathbb{C}$, dann gilt:

$$z = (a, b) = (a, 0) + (0, b) = a + (b, 0)(0, 1) = a + b(0, 1) = a + bi\ .$$

In dieser eindeutigen Darstellung von z (– man vergleiche die Einführung –) heißt a **Realteil** von z und b **Imaginärteil** von z.

Bezeichnungen: $a = \text{Re}\ z,\ b = \text{Im}\ z$.

(f) Durch die Eigenschaften, dass $\mathbb{C}$ die reellen Zahlen als Unterkörper (bis auf Isomorphie) enthält, dass es in $\mathbb{C}$ ein Element i mit $i^2 = -1$ gibt und dass sich jedes $z \in \mathbb{C}$ in der Gestalt $z = a + bi$, $a, b \in \mathbb{R}$, ist $\mathbb{C}$ bis auf Isomorphie festgelegt.

Beweis : Wir prüfen nicht alle neun Körperaxiome nach. Wir weisen lediglich die Kommutativität der Multiplikation und das Assoziativgesetz bzgl. der Multiplikation (für Ungläubige) nach:

Ist $z = (a, b) \in \mathbb{C}$ und $z' = (a', b') \in \mathbb{C}$, dann ist $zz' = z'z$ zu eigen. Nun ist nach Definition:

$$\begin{aligned} zz' &= (a, b)(a', b') &&:= (aa' - bb', ab' + ba') &&\text{Definition des Produkts in } \mathbb{C} \\ z'z &= (a', b')(a, b) &&:= (a'a - b'b, a'b + b'a) &&\text{Definition des Produkts in } \mathbb{C} \end{aligned}$$

Unter Verwendung der Kommutativität des Produkts und der Summe in $\mathbb{R}$, sieht man die Gleichheit

$$\begin{aligned} (aa' - bb', ab' + ba') &= (a'a - b'b, a'b + b'a) \qquad \text{also} \\ zz' &= z'z \end{aligned}$$

Zur Überprüfung des Assoziativgesetzes setzen wir

$$z = (a, b) \in \mathbb{C}, \qquad z' = (a', b') \in \mathbb{C} \qquad \text{und} \qquad z'' = (a'', b'') \in \mathbb{C}\ .$$

Dann gilt:

$$
\begin{aligned}
(zz')z'' &= (\ (a,b)(a',b')\ )(a'',b'') \\
&= (\ aa' - bb'\ ,\ ab' + ba'\ )(a'',b'') \\
&= (\ (aa' - bb')a'' - (ab' + ba')b''\ ,\ (aa' - bb')b'' + (ab' + ba')a''\ ) \\
&= (\ (aa')a'' - (bb')a'' - (ab')b'' - (ba')b''\ ,\ (aa')b'' - (bb')b'' + (ab')a'' + (ba')a''\ )
\end{aligned}
$$

und weiter wegen der Assoziativität des Produkts in $\mathbb{R}$!

$$
= (\ a(a'a'') - b(b'a'') - a(b'b'') - b(a'b'')\ ,\ a(a'b'') - b(b'b'') + a(b'a'') + b(a'a'')\ )
$$

und weiter wegen der Kommutativität der Summe in $\mathbb{R}$!

$$
\begin{aligned}
&= (\ a(a'a'') - a(b'b'') - b(b'a'') - b(a'b'')\ ,\ a(a'b'') + a(b'a'') - b(b'b'') + b(a'a'')\ ) \\
&= (\ a(a'a'' - b'b'') - b(b'a'' + a'b'')\ ,\ a(a'b'' + b'a'') - b(b'b'' - a'a'')\ ) \\
&= (\ a(a'a'' - b'b'') - b(b'a'' + a'b'')\ ,\ a(b'a'' + a'b'') - b(b'b'' - a'a'')\ ) \\
&= (a,b)(\ a'a'' - b'b''\ ,\ b'a'' + a'b''\ ) \\
&= (a,b)(\ (a',b')(a'',b'')\ ) \\
&= z(z'z'')\ .
\end{aligned}
$$

□

4.1.2 Bemerkungen

(a) Da $\mathbb{C}$ ein Körper ist, gelten in $\mathbb{C}$ alle in $\mathbb{R}$ abgeleiteten Rechenregeln, bei denen nur die Körpereigenschaften benutzt waren.
Z.B. gilt für alle $n \in \mathbb{N}_0$ und alle $a, b \in \mathbb{C}$ die binomische Formel:

$$
(a+b)^n = a^n + \binom{n}{1} a^{n-1} b + \cdots + \binom{n}{n-1} a b^{n-1} + b^n
$$

und die geometrische Summenformel

$$
(1-q)(1+q+q^2+\cdots+q^n) = (1-q^{n+1}) \qquad q \in \mathbb{C}\ ,\ n \in \mathbb{N}_0\ .
$$

(b) Die Formel für die Multiplikation komplexer Zahlen braucht man sich nicht zu merken. Ist etwa $z = 2 - 6i$ und $w = 1 + i$, so rechnet man zw mit Hilfe des Distributivgesetzes und unter Verwendung von $i^2 = -1$ und erhält

$$
\begin{aligned}
zw &= (2-6i)(1+i) = 2\cdot 1 + 2\cdot i - 6i\cdot 1 - 6i\cdot i \\
&= 2 + 2i - 6i - 6i^2 = 2 + 2i - 6i - 6(-1) = 8 - 4i\ .
\end{aligned}
$$

Die Standarddarstellung für einen Quotienten wie z.B. $\frac{2+5i}{3-4i}$ berechnet man durch Erweiterung mit $3+4i$:

$$
\frac{2+5i}{3-4i} = \frac{(2+5i)(3+4i)}{(3-4i)(3+4i)} = \frac{6+15i+8i-20}{9+16} = \frac{-14+23i}{25} = -\frac{14}{25} + \frac{23}{25}i\ .
$$

Soll man die Standarddarstellung etwa von $z := \left(\frac{8-i}{5+i}\right)^4$ bestimmen, so berechnet man zunächst

$$
\frac{8-i}{5+i} = \frac{(8-i)(5-i)}{(5+i)(5-i)} = \frac{40-5i-8i-1}{25+1} = \frac{39}{26} - \frac{13}{26}i = \frac{3}{2} - \frac{1}{2}i = \frac{1}{2}(3-i)
$$

und verwendet dann die binomische Formel:

$$\begin{aligned} z &= \frac{1}{16}(3-i)^4 = \frac{1}{16}\left(3^4 - 4\cdot 3^3 i + 6\cdot 3^2 i^2 - 4\cdot 3 i^3 + i^4\right) \\ &= \frac{1}{16}\,(81 - 108i - 54 + 12i + 1) = \frac{28-96i}{16} = \frac{7}{4} - 6i\ . \end{aligned}$$

(c) Die Definition der Multiplikation komplexer Zahlen mag gekünstelt erscheinen, ist aber durch die Vorüberlegung in 3.0 motiviert.

Die komponentenweise Multiplikation $(a,b)*(c,d) := (ac,bd)$ als multiplikative Verknüpfung zu nehmen, scheitert deshalb, weil in jedem Körper K die Nullteilerregel gilt:
Für $z,w \in K$ gilt $zw = 0$ genau dann, wenn $z = 0$ oder $w = 0$ ist.

Bei komponentenweiser Multiplikation wäre aber z.B. $(1,0)*(0,1) = (1\cdot 0, 0\cdot 1) = (0,0)$.
Die komponentenweise Multiplikation der Paare ist aber immerhin noch kommutativ, assoziativ und distributiv bzgl. der Addition. Prüfen Sie dies als Übungsaufgabe nach.

4.1.3 Wichtige Bemerkung

$$\boxed{\mathbb{C} \text{ lässt sich nicht anordnen !}}$$

In $\mathbb{C}$ gibt es also keine Teilmenge P, so dass für P die Axiome (AO_1), (AO_2) und (AO_3) gelten. Denn in einem angeordneten Körper gilt für ein Element $z \neq 0$ stets $z^2 \in P$. Insbesondere ist $1 = 1^2 \in P$, also $-1 \notin P$. In $\mathbb{C}$ gilt aber $i^2 = -1$. Wenn sich $\mathbb{C}$ anordnen ließe, müsste einerseits (wegen $i \neq 0$) $i^2 \in P$ gelten, andererseits ist $i^2 = -1 \notin P$.

Wir können uns an dieser Stelle nur mit wenigen elementaren geometrischen und algebraischen Eigenschaften von $\mathbb{C}$ beschäftigen. Außer den in der Einführung genannten Gründen, den Körper der reellen Zahlen nochmals zu erweitern, gibt es viele zahlreiche weitere, z.B. viele innermathematische Gründe, aber auch etwa für die moderne Quantenmechanik (Schlagwort: Vertauschungsaxiom für Orts– und Impuls–Operator, Schrödinger–Gleichung, Hamilton–Operator) sind die komplexen Zahlen unverzichtbar.

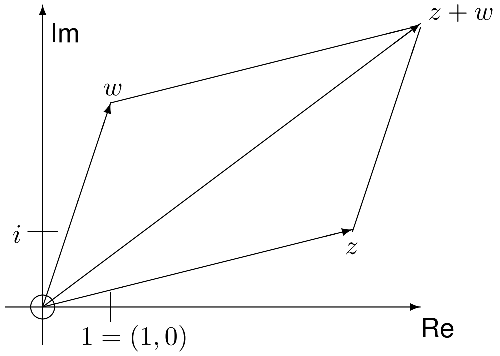
Analysis, die systematisch auf $\mathbb{C}$ aufbaut, heißt im deutschen Sprachgebrauch auch *Funktionentheorie*.

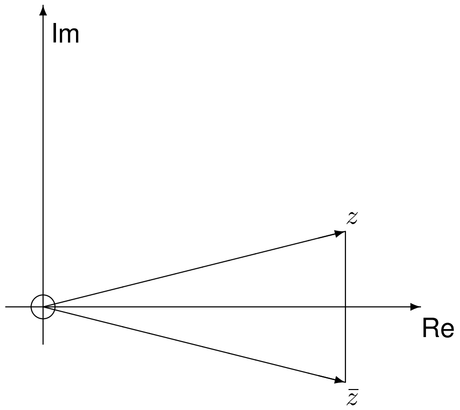
Wir werden uns hier nur mit Rudimenten der komplexen Analysis (=Funktionentheorie) beschäftigen. Im Folgenden stellen wir einige elementare Eigenschaften von $\mathbb{C}$ zusammen.

4.2 Elementare Eigenschaften von $\mathbb{C}$

Die komplexen Zahlen sind insbesondere ein $\mathbb{R}$-Vektorraum mit der speziellen Basis $1(=(1,0))$ und $i(=(0,1))$. Ein Vektorraum hat bekanntlich „viele“ Basen. Die Auswahl von 1 und i ist also in gewisser Weise willkürlich. Die Wahl von $1 = (1,0)$ als Basiselement ist aber dadurch gerechtfertigt, dass 1 neutrales Element bezüglich der Multiplikation in $\mathbb{C}$ ist und $i = (0,1)$ löst die Gleichung $z^2 + 1 = 0$. Aber auch $-i = (0,-1)$ löst diese Gleichung (sonst gibt es keine weiteren Lösungen). Dass man im Prinzip dieselbe Struktur erhält, wenn man in allen Rechnungen mit komplexen Zahlen i durch $-i$ ersetzt, wird sich im Folgenden ergeben.

Den $\mathbb{R}$-Vektorraum $\mathbb{C} = \mathbb{R}^2$ mit der ausgezeichneten Basis $1 = (1,0)$ und $i = (0,1)$ nennt man auch *Gaußsche Zahlenebene* und veranschaulicht sich komplexe Zahlen als Punkte oder Vektoren:





$\mathbb{R} \cong \mathbb{R} \times \{0\}$ nennt man die *reelle Achse* und $i\mathbb{R} = \{0\} \times \mathbb{R}$ die *imaginäre Achse*. Die *Summe* $z + w$ von zwei komplexen Zahlen z und w ist der vierte Eckpunkt des durch $0, z$ und w bestimmten Parallelogramms (falls diese Punkte nicht alle drei auf einer Geraden liegen). $z + w$ kann man auch als den Punkt beschreiben, der durch Translation um w aus z oder durch Translation um z aus w entsteht.
Eine geometrische Deutung für das Produkt zw zweier komplexer Zahlen geben wir weiter unten.
Die Gleichberechtigung von i und $-i$ begründet die Wichtigkeit, der Abbildung

$$^{-} : \mathbb{C} \to \mathbb{C}, z = x + yi \mapsto x - yi\ .$$

4.2.1 Definition (konjugiert komplexe Zahl)

Ist $z = x + yi,\ x, y \in \mathbb{R}$, dann heißt $\overline{z} = x - yi$ die zu z *konjugierte komplexe Zahl* und die Abbildung

$$^{-} : \mathbb{C} \to \mathbb{C}, z \mapsto \overline{z}$$

die *komplexe Konjugation*.
Geometrisch bedeutet der Übergang von z zu $\overline{z}$ eine Spiegelung an der reellen Achse.

4.2.2 Eigenschaften von $^{-} : \mathbb{C} \to \mathbb{C}, z \mapsto \overline{z}$

Für alle $z, w \in \mathbb{C}$ gilt:

(a) $\overline{z + w} = \overline{z} + \overline{w},\ \overline{zw} = \overline{z} \cdot \overline{w}$

(b) $\overline{\overline{z}} = z$

(c) $\text{Re}(z) = z + \overline{z}/2,\ \text{Im}(z) = z - \overline{z}/2i$

(e) $z \in \mathbb{R} \iff z = \overline{z},\ z \in i\mathbb{R} \iff z = -\overline{z}$ (im letzteren Fall nennt man z *rein imaginär*)

(f) $z\overline{z} = x^2 + y^2 \geq 0 \quad (z = x + yi, x, y \in \mathbb{R})$

(g) Ist $z = x + yi \neq 0$ (d. h. $x \neq 0$ oder $y \neq 0$), dann ist

$$z^{-1} = \frac{\overline{z}}{|z|^2} = \frac{x - yi}{x^2 + y^2} = \frac{x}{x^2 + y^2} + \frac{-y}{x^2 + y^2} i.$$

Die Eigenschaften lassen sich kurz zusammenfassen:
$- : \mathbb{C} \to \mathbb{C}$ ist ein *involutorischer Automorphismus* von $\mathbb{C}$ mit der Fixpunktmenge $\mathbb{R}$. Involutorisch bedeutet: Für alle $z \in \mathbb{C}$ gilt: $\overline{\overline{z}} = z$, insbesondere ist $^- : \mathbb{C} \to \mathbb{C}$ bijektiv. Nach (e) ist $\{z \in \mathbb{C}; z = \overline{z}\} = \mathbb{R}$, also ist $\mathbb{R}$ die genaue Fixpunktmenge von $^- : \mathbb{C} \to \mathbb{C}$.

Beweis : Die Rechenregeln ergeben sich alle unmittelbar aus der Definition. Wir beschränken uns auf den Nachweis der Formel $\overline{zw} = \overline{z}\,\overline{w}$.

Ist

$$\begin{aligned} z &= x + yi, \ x, y \in \mathbb{R} \\ w &= u + vi, \ u, v \in \mathbb{R}, \ \text{so ist} \\ \overline{z}\,\overline{w} &= (x - yi)(u - vi) = (xu - yv) - (xv + yu)i = \overline{zw} \end{aligned}$$

$\square$

4.2.3 Ein Beispiel zu §4.2.2(g)

$$\left(\frac{1+i}{\sqrt{2}}\right)^{-1} = \frac{1-i}{\sqrt{2}} \ ,$$

Wir knüpfen an (f) an:
Ist $z = x + yi$, $x, y \in \mathbb{R}$, dann ist $x^2 + y^2 \geq 0$ und $\sqrt{z\,\overline{z}} = \sqrt{x^2 + y^2}$, das ist nach dem Satz des Pythagoras der *Abstand* von z vom *Nullpunkt*. Diese reelle Zahl nennen wir den *Betrag von* z. Ist $z \in \mathbb{R}$, so stimmt die neue Definition mit der Betragsdefinition auf $\mathbb{R}$ überein.

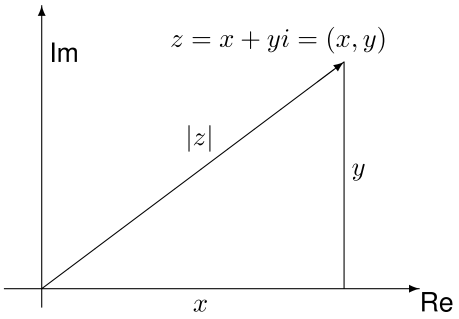
4.2.4 Definition und Satz (Betrag einer komplexen Zahl)

Ist $z = x + yi$ $(x, y \in \mathbb{R})$ eine komplexe Zahl, dann heißt die (nicht negative) reelle Zahl

$$|z| := \sqrt{z\,\overline{z}} = \sqrt{x^2 + y^2}$$

der *Betrag von* z. $|z|$ ist der euklidische Abstand des Punktes $z = x + yi = (x, \varphi)$ vom Nullpunkt.

Offensichtlich ist $|\overline{z}| = |z|$ für alle $z \in \mathbb{C}$.



Eigenschaften des Betrags: Für alle $z, w \in \mathbb{C}$ gilt

(B_1) $|z| \geq 0$; $|z| = 0 \Leftrightarrow z = 0$

(B_2) $|zw| = |z|\,|w|$ (Multiplikativität oder „Produktregel").Ist $w \neq 0$, dann ist $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$.

(B_3) $|\text{Re } z| \leq |z|$, $|\text{Im } z| \leq |z|$.

(B_4) Dreiecksungleichungen:

(a) $|z + w| \leq |z| + |w|$ (für Abschätzungen nach oben),
(b) $|z - w| \geq ||z| - |w||$ (für Abschätzungen nach unten).

Bemerkung: $\mathbb{C}$ lässt sich zwar nicht mehr anordnen, aber durch $|\ |$ ist eine Bewertung auf $\mathbb{C}$ definiert (vgl. ???)

Beweis : Wir beweisen exemplarisch (B_2) und (B_4)(a):

(B_2): Es gilt
$$|zw|^2 = (zw)(\overline{zw}) = (zw)(\overline{z}\,\overline{w}) = (z\overline{z})(w\overline{w}) = |z|^2|w|^2$$
und damit $|zw| = |z|\ |w|$.

Zu (B_4)(a): Es gilt

$$\begin{aligned} |z+w|^2 &= (z+w)(\overline{z+w}) = (z+w)(\overline{z}+\overline{w}) \\ &= z\overline{z} + +w\overline{z} + z\overline{w} + w\overline{w} \\ &= |z|^2 + 2\mathrm{Re}(w\overline{z}) + |w|^2 \\ &\leq |z|^2 + 2|w\overline{z}| + |w|^2 \\ &= |z|^2 + 2|w||\overline{z}| + |w|^2 \\ &= |z|^2 + 2|z||w| + |w|^2 \\ &= (|z| + |w|)^2 \end{aligned}$$

Hieraus folgt wegen der Monotonie von $\sqrt{}$
$$|z+w| \leq |z| + |w|\,.$$

Die Dreiecksungleichung für Abschätzungen nach unten beweist man wie im reellen Fall (vgl. ???).

$\square$

4.2.5 Bemerkungen

(a) Die Identität

$$\begin{aligned} |z+w|^2 &= |z|^2 + 2\ \mathrm{Re}\ (w\overline{z}) + |w|^2 \\ &= |z|^2 + 2\ \mathrm{Re}\ (z\overline{w}) + |w|^2 \end{aligned}$$

(Beachte: $\mathrm{Re}(w\overline{z}) = \mathrm{Re}(\overline{w\overline{z}}) = \mathrm{Re}\ (\overline{w}z) = \mathrm{Re}(z\overline{w})$) wollen wir aus einem unten erläuterten Grund *Cosinus-Satz* nennen.

(b) Die im Beweis verwendete Ungleichung

$$|\mathrm{Re}\ (w\overline{z}| = |\mathrm{Re}\ (z\overline{w})| \leq |z|\ |w|$$

ist die *Cauchy-Schwarzsche* Ungleichung (C.S.U.) in $\mathbb{C} = \mathbb{R}^2$. Schreibt man nämlich

$$\begin{aligned} z &= x + yi,\ x,\ y \in \mathbb{R}z \\ w &= u + vi,\ u, v \in \mathbb{R} \text{ dann ist} \end{aligned}$$

$$\mathrm{Re}(z\overline{w}) = \mathrm{Re}(\overline{z}w) = xu + yv$$

und das ist das (Standard)-*Skalarprodukt* der Vektoren $z = (x, y) \in \mathbb{R}^2$ und $w = (u, v) \in \mathbb{R}^2$.
In äquivalenter Form lautet die C.S.U.:

4.2.6 Satz (Cauchy-Schwarzsche Ungleichung)

Für alle $x, y, u, v \in \mathbb{R}$ gilt

$$|xu + yv|^2 \leq (x^2 + y^2)(u^2 + v^2)$$

oder für alle $z, w \in \mathbb{C}$ gilt

$$|\langle z, w\rangle|^2 \leq |z|^2 \; |w|^2 \, ,$$

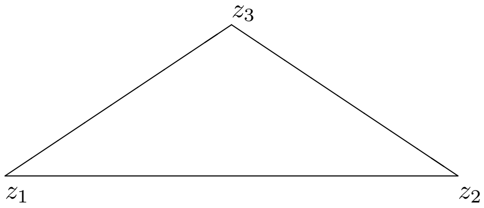
wenn wir für das Skalarprodukt von z und w die Bezeichnung $\langle z, w\rangle$ verwenden.

Bemerkung: Definiert man für beliebige $z, w \in \mathbb{C}$
den *Abstand von* z *und* w durch

$$d(z, w) := |z - w| \, ,$$

dann gilt für beliebige $z_1, z_2, z_3 \in \mathbb{C}$

$$(*) \quad \boxed{d(z_1, z_3) \leq d(z_1, z_2) + d(z_2, z_3)}$$



Der Name „Dreiecksungleichung“ wird jetzt geometrisch verständlich.

$(*)$ folgt aus der speziellen Dreiecksgleichung (Dreieck mit den Ecken $0, z, w)$). Wegen

$$z_1 - z_3 = (z_1 - z_2) + (z_2 - z_3) \quad \text{folgt} \quad |z_1 - z_3| \leq |z_1 - z_2| + |z_2 - z_3|$$

und das ist $(*)$.

Der Abstand d in $\mathbb{C}$ hat die *Eigenschaften*: ($z_1, z_2, z_3 \in \mathbb{C}$ beliebig)

(M_1) $d(z_1, z_2) = 0 \Leftrightarrow z_1 = z_2$

(M_2) $d(z_1, z_2) = d(z_2, z_1)$ (Symmetrie)

(M_3) $d(z_1, z_3) \leq d(z_1, z_2) + d(z_2, z_3)$ (Dreiecksungleichung)

$\mathbb{C}$ mit dieser Abstandsfunktion ist ein *metrischer Raum* (genauso wie $\mathbb{R}$ mit der Abstandsdefinition $d(a, b) = |a - b|; a, b \in \mathbb{R}$).

Mit Hilfe des Betrages bzw. des Abstandes in $\mathbb{C}$ erden wichtige Teilmengen von $\mathbb{C}$ definiert:

4.2.7 Definition (Einheitskreislinie)

$S^1 := \{z \in \mathbb{C}; |z| = 1\}$ heißt *Einheitskreislinie.*

Sie hat die bemerkenswerte Eigenschaft, dass sie bezüglich der *Multiplikation* eine *Gruppe* ist (Untergruppe von $\mathbb{C}^\bullet := \mathbb{C} - \{0\}$). Sind $z, w \in S^1$, so gilt auch $zw \in S^1$ und $\frac{z}{w} \in S^1$. Ferner ist $z^{-1} = \overline{z}$. Etwas allgemeiner ist

4.2.8 Definition (Kreisscheibe)

Für $z_o \in \mathbb{C}$ und $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, heißt

$$U_\varepsilon(z_0) := \{z \in \mathbb{C}; |z - z_0| < \varepsilon\}$$

die *offene Kreisscheibe mit Mittelpunkt* z_0 und Radius ε oder auch ε*-Umgebung von* z_o und

$$\overline{U_\varepsilon}(z_0) := \{z \in \mathbb{C}; |z - z_0| \leq \varepsilon\}$$

die *abgeschlossene Kreisscheibe* mit Mittelpunkt z_0 und Radius ε. Und

$$S_\varepsilon(z_0) := \{z \in \mathbb{C}; |z - z_0| = \varepsilon\}$$

Sphäre mit Mittelpunkt z_0 und Radius ε. Für $z_0 = 0$ und $\varepsilon = 1$, verwendet man die obige Bezeichnung

$$S^1 := S_1(0)$$

Wir wollen noch die Abbildung (Inversion)

$$\begin{array}{rccl} j: & \mathbb{C} - \{0\} & \to & \mathbb{C} - \{0\}\ , \\ & z & \mapsto & \dfrac{1}{z} \end{array}$$

etwas genauer betrachten. Die Formel

$$\frac{1}{z} = \frac{1}{|z|^2}\ \overline{z}$$

für das Inverse einer komplexen Zahl $z \neq 0$ zeigt, dass $\frac{1}{z}$ die „Richtung“ von $\overline{z}$ hat. Wie kann man $1/z$ geometrisch konstruieren?

4.2.9 Definition (Spiegelung am Einheitskreis)

Sind $z', z \in \mathbb{C}^\bullet = \mathbb{C} - \{0\}$, dann ließen z und z' *Spiegelpunkte* bezüglich der Einheitskreislinie S^1, wenn gilt

(a) $z' = az$ mit einem $a \in \mathbb{R}, a > 0$.

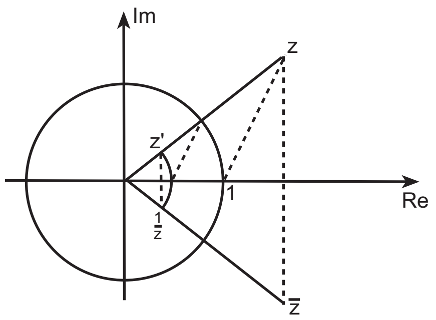
(b) $|z'|\,|z| = 1$.

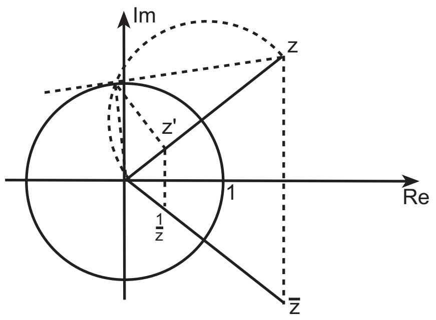
(a) bedeutet, dass z' und z auf dem selben von 0 ausgehenden Halbstrahl liegen.
Setzt man die Gleichung $z' = az$ in (b) ein, so folgt,

$$|z'|\,|z| = |az|\,|z| = |a|\,|z|^2 = az\overline{z} = 1, \text{ also}$$

$$\begin{aligned} a &= \frac{1}{z\overline{z}} \text{ und damit} \\ z' &= \frac{1}{z\overline{z}}\, z \end{aligned}$$

Wegen $\frac{1}{z} = \frac{1}{z\overline{z}}\,\overline{z} = \overline{z'}$ erhält man daher $1/z$, indem man den Punkt z' an der reellen Achse spiegelt.
Man erhält also $1/z$ durch zwei Spiegelungen: Spiegelung am Einheitskreis (ergibt z') und dann Spiegelung an der reellen Achse (ergibt $1/z$). Zum gleichen Resultat kommt man, indem man zuerst an der reellen Achse und dann am Einheitskreis spiegelt. Die Punkte z' sind bei der Spiegelung am Kreis Fixpunkte. Den Punkt z' kann man geometrisch mit verschiedenen Methoden konstruieren, in den folgenden Abbildungen sind zwei angedeutet.





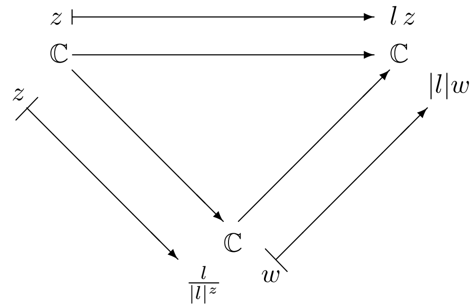
4.2.10 Geometrische Interpretation der Multiplikation

Die geometrische Interpretation der Multiplikation komplexer Zahlen erfordert etwas mehr Aufwand. Ist $\ell = a + bi \in \mathbb{C},\ a, b \in \mathbb{R}$ und $\ell \neq 0$, dann betrachten wir die $\mathbb{C}$-lineare Abbildung

$$\begin{aligned} \mu_\ell :\ \mathbb{C} &\to \mathbb{C}\,, \\ z &\mapsto \mu_\ell(z) = \ell z \end{aligned}$$

und zerlegen sie wie folgt:
Die Abbildung

$$\begin{aligned} \mathbb{C} &\to \mathbb{C} \\ w &\mapsto |\ell|\, w \end{aligned}$$



ist eine „Streckung“ mit dem Zentrum 0 und dem Streckungsfaktor $|\ell|$. Es ist also $\mu_\ell = \mu_{|\ell|} \circ \mu_d$. Ist $d := \frac{\ell}{|\ell|}$, dann ist $|d| = 1$ und

$$\begin{aligned} \mu_d :\ \mathbb{C} &\to \mathbb{C}\,, \\ z &\mapsto dz \end{aligned}$$

ist längentreu („abstandstreu“), denn

$$\begin{aligned} |\mu_d(z_1) - \mu_d(z_2)|\,| &= |dz_1 - dz_2| \\ &= |d(z_1 - z_2)| \\ &= |d|\,|z_1 - z_2| \\ &= |z_1 - z_2| \end{aligned}$$

Ferner ist μ_d $\mathbb{C}$-linear (insbesondere $\mathbb{R}$-linear) und bijektiv:
Die Abbildung

$$\begin{aligned} z &\mapsto dz \quad \text{wird durch} \\ w &\mapsto \frac{1}{d}\, w \quad \text{und umgekehrt}\,. \end{aligned}$$

Die Abbildung μ_d führt die Basisvektoren 1 und i in $d \cdot 1$ bzw. $d \cdot i$ über.

$$\begin{pmatrix} \mathrm{Re}\,(d \cdot 1) & \mathrm{Re}\,(d \cdot i) \\ \mathrm{Im}\,(d \cdot 1) & \mathrm{Im}\,(d \cdot i) \end{pmatrix} = \begin{pmatrix} \mathrm{Re}\,(d) & -\mathrm{Im}\,(d) \\ \mathrm{Im}\,(d) & \mathrm{Re}\,(d) \end{pmatrix}$$

Ihre Determinante ist $|d|^2 = 1$.

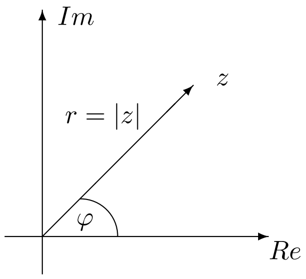
μ_d ist also eine längentreue, orientierungserhaltende lineare Abbildung von $\mathbb{C} = \mathbb{R}^2$, eine solche heißt „Drehung“ um den Nullpunkt.

Zusammengefasst:
Die Abbildung ($l \in \mathbb{C},\ l \neq 0$)

$$\begin{aligned} \mu_\ell : \mathbb{C} &\to \mathbb{C} \\ z &\mapsto \ell z \end{aligned}$$

ist eine *Drehstreckung*.

Übersichtlicher wird die geometrische Beschreibung der Multiplikation komplexer Zahlen durch Verwendung von *Polarkoordinaten*



4.2.11 Satz (Polarkoordinaten in $\mathbb{C} = \mathbb{R}^2$)

Jedes $z \in \mathbb{C}$ kann in der Form

$$z = r(\cos\varphi + i\sin\varphi)$$

mit $r, \varphi \in \mathbb{R},\ r \geq 0$, dargestellt werden. Dabei ist $r = |z|$ eindeutig bestimmt, φ für $z \neq 0$ eindeutig bis auf Addition ganzzahliger Vielfacher von 2π und φ ist beliebig für $z = 0$. φ ist der in Bogenmaß gemessene orientierte Winkel zwischen der positiven reellen Achse und dem Ortsvektor von z. r und φ nennt man *Polarkoordinaten* von z.

Zum Beweis dieser Darstellung benötigt man Eigenschaften der Funktionen $\cos$ und $\sin$, die wir später herleiten werden und die wir hier kurz zusammenstellen:

1. *Additionstheoreme:* Für beliebiges $\varphi, \psi \in \mathbb{R}$ ist

$$\boxed{\begin{aligned}\sin(\varphi+\psi) &= \sin\varphi\cos\psi + \cos\varphi\sin\psi\\ \cos(\varphi+\psi) &= \cos\varphi\cos\psi - \sin\varphi\sin\psi\end{aligned}}$$

2. Die Funktionen $\sin$ und $\cos$ sind *periodisch* mit der Periode 2π, d. h. für beliebiges $\varphi \in \mathbb{R}$ ist $\sin(\varphi + 2\pi) = \sin\varphi, \cos(\varphi + 2\pi) = \cos\varphi$

3. Zu jedem Punkt $(a, b) \in \mathbb{R}^2$ mit $a^2 + b^2 = 1$ (d. h. $a + bi \in S^1$) gibt es ein $\varphi \in \mathbb{R}$ mit $a = \cos\varphi$ und $b = \sin\varphi$

Wählt man z.B. φ im Intervall $[0, 2\pi[$ oder im Intervall $]-\pi, \pi]$, dann ist φ eindeutig bestimmt. Umgekehrt gilt für jedes $\varphi \in \mathbb{R}, \cos\varphi + i\sin\varphi \in S^1$. Generell ist φ die Länge des Bogens vom Punkt $(1, 0)$ zum Punkt $(a, b) \in S^1$. Es gilt für $\varphi, \varphi' \in \mathbb{R}$

$$(\cos\varphi, \sin\varphi) = (\cos\varphi', \sin\varphi') \Longleftrightarrow \varphi - \varphi' = 2\pi k, k \in \mathbb{Z}.$$

Soviel zu den Voraussetzungen.
Ist nun $z \in \mathbb{C},\ z \neq 0$, dann liegt $\frac{z}{|z|}$ auf der Kreislinie S^1, es gibt also ein $\varphi \in \mathbb{R}$ mit $\frac{z}{|z|} = \cos\varphi + i\sin\varphi$, daher ist

$$\boxed{z = |z|(\cos\varphi + i\sin\varphi) = |z|\mathrm{E}(\varphi)}$$

mit der Abkürzung $\mathrm{E}(\varphi) := \cos\varphi + i\sin\varphi$. Damit hat man die Existenz einer Polarkoordinatenstellung. Ist außerdem auch $z = r'(\cos\varphi' + \mathrm{ist}\,\varphi')$ mit $r' > 0$ und $\varphi' \in \mathbb{R}$, dann folgt

$$|z| = r'\sqrt{(\cos\varphi')^2 + (\sin\varphi')^2} = r',$$

also $r = r'$ und dann $\varphi - \varphi' = 2\pi k$ mit $k \in \mathbb{Z}$.

Jedes $\varphi \in \mathbb{R}$ mit $z = r\mathrm{E}(\varphi)$ heißt *ein Argument von* z. Eine komplexe Zahl hat also „viele“ Argumente. Wenn man Eindeutigkeit erreichen will, wählt man φ z.B. im Intervall $]-\pi, \pi]$ und spricht dann vom *Hauptwert des Arguments* von z und schreibt manchmal $\varphi =$Arg z.

Beispiele:

$$\mathrm{Arg(i)} = \frac{\pi}{2}; \mathrm{Arg}(-1) = \pi,$$
$$\mathrm{Arg(-i)} = \frac{-\pi}{2}; \mathrm{Arg}(1) = 0.$$

Sind nun $z, w \in \mathbb{C}$ und $z = r\mathrm{E}(\varphi)$ bzw. $w = \rho\mathrm{E}(\psi)$ Polarkoordinatendarstellungen von z bzw. w, dann ist

$$zw = r\rho(\mathrm{E}(\varphi)\,\mathrm{E}(\psi)) = \mathrm{r}\rho\mathrm{E}\,(\varphi + \psi)$$

denn die Additionstheoreme für $\cos$ und $\sin$ sind mit

$$\mathrm{E}(\varphi + \psi) = \mathrm{E}(\varphi)\mathrm{E}(\psi)$$

äquivalent.

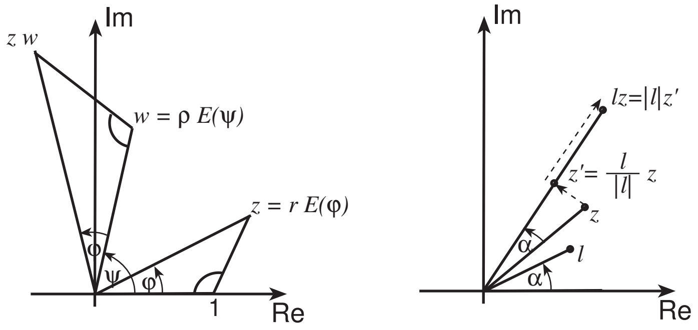
Daraus ergibt sich:

Komplexe Zahlen werden multipliziert, indem man ihre Beträge multipliziert und ihre Argumente addiert.

Speziell ergibt sich, wenn man etwa $w := \ell = \rho\mathrm{E}(\varphi) \in \mathbb{C}^\bullet$ fest wählt, dass die Abbildung

$$\begin{array}{rccl} \mu_\ell : & \mathbb{C} & \to & \mathbb{C}\,, \\ & z & \mapsto & \ell z \end{array}$$

eine Drehstreckung ist, eine Drehung um den Nullpunkt mit dem Drehwinkel φ und eine Streckung mit dem Streckungsfaktor $|l|$ und 0 als Streckungszentrum.



Auch die *Inversion* erhält nun eine einfache Interpretation: Ist $z = r\mathrm{E}(\varphi)$, $r > 0, \varphi \in \mathbb{R}$, dann gilt für den Spiegelpunkt z' bezüglich der Einheitskreislinie S^1

$$z' = \rho\mathrm{E}(\varphi) \;\text{ mit }\; r\rho = 1 \qquad \text{und} \qquad \overline{z'} = \frac{1}{z} = \rho\mathrm{E}(-\varphi)\;.$$

Wir haben die komplexen Zahlen eingeführt, um die Gleichung $z^2 + 1 = 0$ zu lösen. Überraschend und fundamental für die Anwendung der komplexen Zahlen ist die Tatsache, dass beliebige algebraische Gleichungen Lösungen in $\mathbb{C}$ besitzen. Das ist die Aussage des

Fundamentalsatzes der Algebra: Sind $a_0, \ldots, a_n \in \mathbb{C};\ a_n \neq 0, n \in \mathbb{N}$, beliebig vorgegebene Zahlen, dann besitzt die Gleichung

$$a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0 = 0$$

mindestens eine Lösung in $z \in \mathbb{C}$. Eine äquivalente Formulierung ist:
Jedes nicht konstante komplexe Polynom besitzt in $\mathbb{C}$ mindestens eine Nullstelle.

Obwohl der Satz „Fundamentalsatz der Algebra“ heißt, muss man bei den Beweisen auf Hilfsmittel der Analysis zurückgreifen. Sechs verschiedene Beweise finden sich in: Freitag/Busam: Funktionentheorie 1, Springer-Verlag, 3. Auflage, 2000. Es gibt ein Buch (Rosenberger/ ?) in welchem 200 Beweise aufgeführt sind!

Ein Spezialfall des Fundamentalsatzes ist der folgende Satz, der sich elementar beweisen lässt.

4.2.12 Existenzsatz für n-te Wurzeln

Sei $n \in \mathbb{N}$ und $w \in \mathbb{C}$. Ist $w = 0$, dann hat die Gleichung $z^n = w$ nur die Lösung $z = 0$. Ist $w \neq 0$ und $w = r\mathrm{E}(\varphi)$ eine Polarkoordinaten-Darstellung von w, dann ist $z_0 := \sqrt[n]{r}\ \mathrm{E}\left(\frac{\varphi}{n}\right)$ eine Lösung der Gleichung $z^n = w$, d.h. eine n-te Wurzel aus w. Ist

$$\zeta_n := E\left(\frac{2\pi}{n}\right) = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}\ ,$$

dann sind die n verschiedenen Zahlen

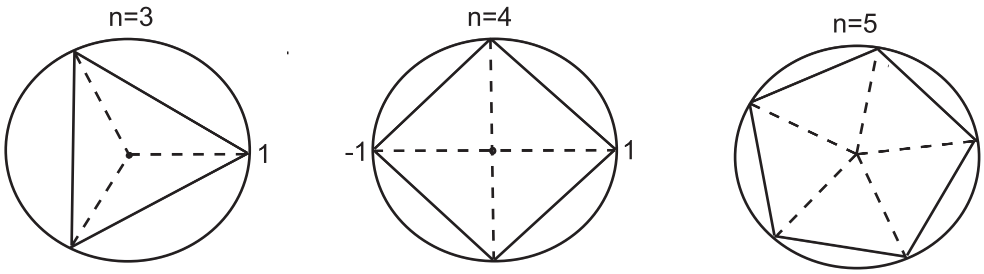
$$z_k = \zeta_n^k\, z_0 = \sqrt[n]{r}\ \mathrm{E}\left(\frac{\varphi + 2\pi k}{n}\right), \quad b = 0, 1 \ldots, n-1$$

sämtliche n-ten Wurzeln aus w. Sie bilden die Eckpunkte eines regelmäßigen n-Ecks, dessen Umkreis den Kreis um 0 mit dem Radius $\sqrt[n]{r}$ ist.

Insbesondere bilden im Fall $w = 1$ die Lösungen von $z^n = 1$, also die Zahlen

$$\zeta_{n,k} := \mathrm{E}\left(\frac{2\pi}{n}k\right), \quad k = 0, \ldots, n-1,$$

die sogenannten *n-ten Einheitswurzeln*, die Eckpunkte eines regelmäßigen n-Ecks (und einer Ecke in 1), dessen Umkreis die Einheitskreislinie S^1 ist.



Mit

$$\zeta_n := \mathrm{E}\left(\frac{2\pi}{n}k\right) \quad \text{gilt} \quad \zeta_{n,k} = \zeta_n^k;\ 0 \le k \le n-1$$

Wir erläutern zum Schluss noch, warum wir die Gleichung

$$|z + w|^2 = |z|^2 + 2\mathrm{Re}\,(w\overline{z}) + |w|^2$$

Cosinus-Satz genannt haben. Ist $w = \rho\mathrm{E}(\psi)$, $z = r\mathrm{E}(\varphi)$, dann ist $\overline{z} = r\mathrm{E}(-\varphi)$ und $w\overline{z} = |w|\,|z|\mathrm{E}(\psi - \varphi)$. Für das Skalarprodukt ergibt sich

$$< w, z >= \mathrm{Re}\,(w\overline{z}) = |w|\,|z|\cos\alpha,$$

wenn α der „Winkel zwischen z und w“ ist (vergl. die Bezeichnungen in der Abbildung). Wegen $\vartheta + \alpha = \pi$ gilt

$$\cos\alpha = -\cos\vartheta$$

und somit

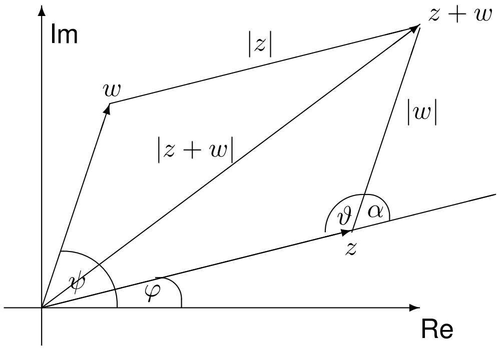
$$\boxed{|z + w|^2 = |z|^2 + |w|^2 - 2|z|\,|w|\cos\vartheta = |z|^2 + |w|^2 + 2|z|\,|w|\cos\alpha}$$

Im Cosinus-Satz ist im Spezialfall $\vartheta = \frac{\pi}{2}$ (also $\cos\vartheta = 0$) der Satz des Pythagoras enthalten.

Aber wie ist eigentlich ein Winkel zwischen zwei Vektoren zu definieren und zu messen?
Bevor man einen Winkel messen kann, muss man ihn erst mal definieren!

Im nächsten Paragrafen werden wir uns mit Skalarprodukten beschäftigen. Im Falle des Standardvektorraums $\mathbb{R}^n$ werden wir dann auch „Winkel zwischen Vektoren“ definieren können.

$$\alpha = \psi - \phi$$



5 Vektorräume mit Skalarprodukt, insbes. die Standardvektorräume $\mathbb{R}^n$ und $\mathbb{C}^n$

Im Vektorraum $\mathbb{C} = \mathbb{R}^2$ ließ sich neben der bekannten Addition von Summanden(Vektoren) eine Zusatzstruktur, nämlich eine Multiplikation so einführen, dass $\mathbb{C}$ mit der vorhandenen Addition und dieser Multiplikation ein Körper wird. (vgl. §4) Man kann nämlich fragen, ob man auch für $n > 2$ auf dem $\mathbb{R}^n$, auf dem eine Addition (siehe unten) erklärt ist, eine Multiplikation so definieren kann, dass wieder eine Körperstruktur herauskommt.
Dass man in $\mathbb{C} = \mathbb{R}^2$ noch eine Körperstruktur erklären kann, ist kein Zufall.

Es gilt der folgende Eindeutigkeitssatz:

Sei $\mathbb{K}$ ein Körper, der den Körper der reellen Zahlen als Unterkörper enthält und der ein endlich dimensionaler Vektorraum über $\mathbb{R}$ ist, $dim_{\mathbb{R}}\, K = n < \infty$, dann gilt:
entweder ist $n = 2$ und $\mathbb{K}$ ist als Körper isomorph zum Körper der komplexen Zahlen,
oder es ist $n = 1$ und $\mathbb{K}$ ist isomorph zum Körper $\mathbb{R}$ der reellen Zahlen. [2].

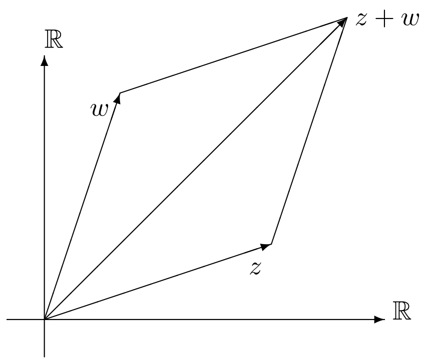
Im Fall $n = 3$ gibt es das auch für viele physikalische Anwendungen wichtige *Vektorprodukt* (Kreuzprodukt), das ist die Abbildung: für $x = (x_1, x_2, x_3),\ y = (y_1, y_2, y_3)$

$$\begin{array}{rcl} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x, y) & \mapsto & x \wedge y := x \times y := (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) \end{array}$$

Diese Abbildung ist bilinear, hat auch sonst weitere angenehme Eigenschaften, aber es gilt z.B. $y \times x = -(x \times y)$, insbesondere ist für alle $x \in \mathbb{R}^3 :\ x \times x = 0.$ [3]

Verzichtet man auf die Kommutativität der Multiplikation, so besitzt der $\mathbb{R}^4 = \mathbb{C} \times \mathbb{C}$ noch eine Struktur als *Schiefkörper*.

Der abstrakte Begriff des Vektorraums, wie es in der heutigen Mathematik, speziell der Analysis verwendet wird, ist aus elementargeometrischen und physikalischen Ansätzen aus der anschaulichen „Vektorrechnung mit Pfeilen“ entwickelt worden.
Eine geometrisch anschauliche Interpretation von Vektoren in $\mathbb{C} = \mathbb{R}^2$ ist im letzten Paragraphen verwendet worden. Dabei konnte man einerseits die Elemente $z \in \mathbb{C} = \mathbb{R}^2$ als „Punkte“ auffassen und aus Punkten bestehende Teilmengen von $\mathbb{C}$ betrachten, oder aber auch als Vektoren, genauer als Translationen. Betrachtet man die Addition zwischen komplexen Zahlen z und w, also die Summe $z + w$ und konstruiert diese Zahl mit Hilfe des Kräfteparallelogramms, so kann man auch sagen:
Man erhält $z + w$, indem man die Translation w auf den Punkt z oder die Translation z auf den Punkt w anwendet. Analog kann man die Addition in den Standardvektorräumen $\mathbb{R}^n$ bzw. $\mathbb{C}$ interpretieren.



Diese Räume tragen eine wichtige Zusatzstruktur, die sie eigentlich erst richtig interessant macht, weil man mit Hilfe dieser Zusatzstruktur, *Längen*(Normen) von *Vektoren* und Winkel zwischen Vektoren erklären kann und z.B. sagen kann, wann zwei Vektoren *senkrecht aufeinander* stehen.

Diese Zusatzstruktur erhält man mir Hilfe eines *Skalarprodukts*. Eine solche haben wir im Fall

[2]Für einen Beweis vergleiche man den Artikel von R.Remmert über komplexe Zahlen in „Zahlen“; Herausgeber: Ebbinghaus Kap.3; Springer Verlag 1988 (siehe Literaturverzeichnis)

[3]Näheres findet man etwa bei G.Fischer: Lineare Algebra, 11. Auflage.

$\mathbb{C} = \mathbb{R}^2$ schon betrachtet. Eine Verallgemeinerung auf $\mathbb{R}^n$ liegt auf der Hand.

Aber zunächst als Wiederholung aus der Linearen Algebra.

Ist $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, so besteht für $n \in \mathbb{N}$ der Standardvektorraum $\mathbb{K}^n$ aus allen n-Tupeln

$$x := (x_1, ..., x_n), \quad x_j \in \mathbb{K}, \quad 1 \leq j \leq n$$

x_j heißt dabei die *j-te Komponente* von x.
Man vereinbart die Gleichheit von n-Tupeln

$$(x_1, ..., x_n) = (y_1, ..., y_n) \iff x_j = y_j \quad \text{für } 1 \leq j \leq n$$

Der Punkt $\mathbf{0} := (0, ..., 0) \in \mathbb{K}^n$ heißt Nullpunkt oder Ursprung in $\mathbb{K}^n$.

Die *Addition* in $\mathbb{K}^n$ ist für $x = (x_1, ..., x_n)$ und $y = (y_1, ..., y_n)$ erklärt durch

$$x + y := (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) \qquad \text{(also komponentenweise)}$$

und die skalare Multiplikation für $\lambda \in \mathbb{K}$ durch

$$\lambda x := (\lambda x_1, ..., \lambda x_n)$$

Bezüglich der so definierten Addition und Multiplikation ist $\mathbb{K}^n$ ein Vektorraum der Dimension n, denn es ist etwa

$$\begin{array}{rcl} e_1 & = & (1, 0, 0, 0, ..., 0) \\ e_2 & = & (0, 1, 0, 0, ..., 0) \\ e_3 & = & (0, 0, 1, 0, ..., 0) \\ \vdots & & \vdots \\ e_n & = & (0, 0, 0, ..., 0, 1) \end{array}$$

eine Basis.

Neutrales Element bezüglich der Addition ist der *Nullvektor*: $\mathbf{0} = (0, ..., 0)$ und das Negative zu $x = (x_1, ..., x_n)$ ist $-x = (-x_1, ..., -x_n)$

Pflichtlektüre für alle (nicht nur für Physiker!)
Abschnitt 2.6 *„Was sind Vektoren? “ in K. Jänisch: Lineare Algebra; Springer-Verlag; 8. Afl. 2001 (dort siehe Seite 38 bis 50).*

In $\mathbb{C} = \mathbb{R}^2$ wird für $z = (x_1.y_1)$ und $w = (x_2, y_2)$ die reelle Zahl

$$\langle z, w \rangle = x_1 y_1 + x_2 y_2$$

das *Standardskalarprodukt* genannt.

Eine Verallgemeinerung auf beliebiges n liegt nahe.

5.1 Definition und Satz

Definiert man für $x = (x_1, ..., x_n) \in \mathbb{R}^n$ und $y = (y_1, ..., y_n) \in \mathbb{R}^n$

$$\langle x, y\rangle := x_1y_1 + x_2y_2 + \cdots + x_ny_n \ ,$$

dann heißt die so definierte reelle Zahl das *Standardskalarprodukt* von x und y.

Die nicht negative reelle Zahl

$$\|x\| := \|x\|_2 := \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

heißt die *euklidische Länge* oder *euklidische Norm* von x und die (ebenfalls) nicht negative Zahl

$$\|x - y\| := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

der *euklidische Abstand* von x und y.

Wir schreiben die wichtigsten *Eigenschaften* des Skalarprodukts, der Länge(Norm) und des Abstandes:

Das Standardskalarprodukt

$$\begin{aligned} \langle\ ,\ \rangle : \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\ (x, y) &\mapsto \langle x, y\rangle \end{aligned}$$

hat die folgenden Eigenschaften: für alle $x, y, z \in \mathbb{R}^n$, $\lambda, \mu \in \mathbb{R}$ gilt

(SP_1): $\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle$ (Linearität im ersten Bestandteil)
(SP_2): $\langle y, x\rangle = \langle x, y\rangle$ (Symmetrie)
(P): $\langle x, y\rangle \geq 0$, und $\langle x, x\rangle = 0 \iff x = 0$ (Positive Definitheit)

Ferner gilt für alle $x, y \in \mathbb{R}^n$ die Ungleichung

(C.S.U.): $|\langle x, y\rangle| \leq \|x\|\|y\|$ oder explizit

$$\left|\sum_{j=1}^{n} x_jy_j\right| \leq \sqrt{x_1^2 + \cdots + x_n^2} \cdot \sqrt{y_1^2 + \cdots + y_n^2}$$

Man nennt sie <u>**C**</u>auchy-<u>**S**</u>chwarzsche-<u>**U**</u>ngleichung (in $\mathbb{R}^n$).

Die *euklidische Norm* hat folgende Eigenschaften: für alle $x, y \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt

(N_1) : $\|x\| \geq 0, \ \ und \ \ \|x\| = 0 \iff x = 0$
(N_2) : $\|\lambda x\| = |\lambda|\ \|x\|$
(N_3) : $\|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung)

Der *euklidische Abstand* $d(x, y) = \|x - y\|$ hat folgende Eigenschaften: für alle $x, y, z \in \mathbb{R}^n$ gilt

(M_1) : $d(x, y) = 0 \iff x = y$
(M_2) : $d(x, y) = d(y, x)$ (Symmetrie)
(M_3) : $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung)

(dass $d(x, y) \geq 0$ ist, folgt aus (M_1) und (M_2))

Bemerkungen zum Beweis:
Die genannten Eigenschaften des Skalarproduktes, der euklidischen Norm und des euklidischen

Abstandes, bis auf die Cauchy-Schwarzsche Ungleichung, folgen alle unmittelbar aus der Definition, bei der (SP_1) benötigt man eine kleine Rechnung. Aus (SP_2) folgt auch die Linearität im zweiten Bestandteil des Skalarprodukts: für alle $x, y, z \in \mathbb{R}^n$ und alle $\lambda, \mu \in \mathbb{R}$ gilt:

$$\langle y, \lambda y + \mu z \rangle = \lambda \langle x, y \rangle + \mu \langle x, z \rangle$$

Man sagt: Das Standardskalarprodukt ist eine *positiv definite* (wegen (P)) *symmetrische Bilinearform* auf $\mathbb{R}^n$.

Auf den Beweis der Cauchy-Schwarz-schen Ungleichung gehen wir gleich im etwas allgemeineren Rahmen ein.

5.2 Definition (Vektorraum mit Skalarprodukt)

Sei V ein beliebiger $\mathbb{K}$-Vektorraum. Eine Abbildung

$$\begin{aligned} \langle\ ,\ \rangle : V \times V &\to \mathbb{K} \\ (v, w) &\mapsto \langle v, w \rangle \end{aligned}$$

heißt *Skalarprodukt auf V*, falls für alle $v, v', w, w' \in V$ und alle $\lambda \in \mathbb{K}$ gilt:

Im Fall $\mathbb{K} = \mathbb{R}$	Im Fall $\mathbb{K} = \mathbb{C}$
$(BF_1):\ \langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$ $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$	$(HF_1):\ \langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$ $\langle \lambda v, w \rangle = \bar{\lambda} \langle v, w \rangle$
$(BF_2):\ \langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$ $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$	$(HF_2):\ \langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$ $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$
$(BS):\ \langle v, w \rangle = \langle w, v \rangle$	$(HS):\ \langle v, w \rangle = \overline{\langle w, v \rangle}$

$$(P):\ \langle v, v \rangle\ \geq 0 \text{ und } \langle v, v \rangle = 0 \iff v = 0$$

5.3 Bemerkungen

(a) Im Fall $\mathbb{K} = \mathbb{R}$ ist ein Skalarprodukt eine positiv definite *Bilinearform*, im Fall $\mathbb{K} = \mathbb{C}$ eine positiv definite *hermitesche Form*.

(b) (BF_2) und (HF_2) sind eigentlich überflüssig. Diese beiden Eigenschaften folgen aus (BF_1) und (BS) bzw. (HF_1) und (HS). Da diese Eigenschaften aber dauernd gebraucht werden, haben wir sie unter die Grundeigenschaften aufgenommen.

(c) Ein Vektorraum mit Skalarprodukt heißt im Fall $\mathbb{K} = \mathbb{R}$ auch ein *euklidischer Vektorraum* und im Fall $\mathbb{R} = \mathbb{C}$ auch *unitärer Vektorraum*. Man beachte, dass im Fall $\mathbb{K} = \mathbb{C}$ wegen $\langle v, w \rangle = \overline{\langle w, v \rangle}$ die Zahl $\langle v, v \rangle$ eine nicht negative reelle Zahl ist. Man nennt auch hier:

$$\|v\| := \sqrt{\langle v, v \rangle}$$

die (aus dem Skalarprodukt abgeleitete) *Norm von* v und für $v, w \in V$ $\|v - w\|$ den *Abstand von* v *und* w.

$v, w \in V$ heißen *orthogonal* – in Zeichen $v \perp w$ –, falls $\langle v, w\rangle = 0$ gilt.

5.4 Beispiele

(a) Auf $V = \mathbb{R}^n$ ist $\langle x, y\rangle := x_1y_1 + \cdots + x_ny_n$ ein Skalarprodukt, das *Standardskalarprodukt*, dabei seien $x = (x_1, ..., x_n) \in \mathbb{R}^n,\ y = (y_1, ..., y_n) \in \mathbb{R}^n$ und es gilt $e_j \perp e_k$ für $k \neq j$.

Man sagt auch: Die Standardbasis $(e_1, e_2, ..., e_n)$ ist eine *Orthonormalbasis*.
Auf $V = \mathbb{R}^2$ ist aber auch z.B.

$$\langle x, y\rangle_S := (x_1, x_2) \cdot S \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

mit $S := \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ ein Skalarprodukt auf $\mathbb{R}^2$, expliziter

$$\langle x, y\rangle_S = 2x_1y_1 + x_1y_2 + x_2y_1 + 3x_2y_2$$

(b) Ist $V = \mathbb{C}^n$ und $z = (z_1, ..., z_n),\ w = (w_1, ..., w_n)$, dann ist

$$\langle z, w\rangle := \bar{z}_1w_1 + \bar{z}_2w_2 + \cdots + \bar{z}_nw_n$$

ein Skalarprodukt auf $\mathbb{C}^n$, man nennt es wieder das *Standardskalarprodukt*. Sind alle Komponenten reell, fällt es mit dem Standardskalarprodukt im $\mathbb{R}^n$ zusammen.

(c) Ist $M = [0, 1]$ und $V = \{f : [0, 1] \to \mathbb{K}\}$
der Vektorraum der stetigen, $\mathbb{K}$-wertigen Funktionen auf M, so erhält man für $f, g \in V$

1. Im Fall $\mathbb{K} = \mathbb{R}$ durch

$$\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$$

und

2. Im Fall $\mathbb{K} = \mathbb{C}$ durch

$$\langle f, g\rangle = \int_0^1 \overline{f(t)}g(t)\,dt$$

ein Skalarprodukt auf V.

(d) Definiert man auf $V = \mathbb{R}^4$ für $x = (t, x_1, x_2, x_3)$ und $y = (s, y_1, y_2, y_3)$

$$\langle x, y\rangle_L = st - x_1y_1 - x_2y_2 - x_3y_3$$

dann ist $\langle\ ,\ \rangle_L$ eine Bilinearform auf $\mathbb{R}$, die so genannte *Lorentz-Form*, sie spielt speziell in der Relativitätstheorie eine wichtige Rolle.

$(\mathbb{R}^4, \langle\ ,\ \rangle_L)$ heißt *Minkowski-Raum*. $\langle\ ,\ \rangle_L$ ist allerdings nicht positiv definit, denn z.B. für den Vektor $x = (1, -1, 0, 0) \neq 0$ gilt

$$\langle x, x\rangle_L = 1 \cdot 1 - 1 \cdot 1 - 0 \cdot 0 - 0 \cdot 0 = 0$$

und für $x = (1,1,1,1)$ gilt

$$\langle x, x\rangle_L = 1\cdot 1 - 1\cdot 1 - 1\cdot 1 - 1\cdot 1 = -2$$

Die Dreiecksungleichung für die entsprechende Normen bzw. Metriken spielen in den Anwendungen eine wichtige Rolle.
Wir beweisen deshalb die Cauchy-Schwarzsche Ungleichung für einen beliebigen $\mathbb{K}$-Vektorraum mit Skalarprodukt.

5.5 Satz (Cauchy-Schwarzsche Ungleichung)

Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$, dann gilt für alle $v, w \in V$ die *Cauchy-Schwarzsche Ungleichung*:

$$\boxed{\text{(C.S.U.)} : \quad |\langle v, w\rangle| \leq \|v\|\,\|w\|}$$

Beweis : Wir behandeln zunächst den Fall $\mathbb{K} = \mathbb{R}$.

Für beliebige $v, w \in V$ und beliebige $t \in \mathbb{R}$ gilt dann nach (P)

$$\begin{aligned} 0 \leq \langle v - tw, v - tw\rangle &= \langle v, v\rangle - 2t\langle v, w\rangle + t^2\langle w, w\rangle \\ &= \|v\|^2 - 2t\langle v, w\rangle + t^2\|w\|^2 \end{aligned}$$

Ist $w = 0$, so steht in der C.S.U. auf beiden Seiten Null. Wir können daher $w \neq 0$ annehmen, dann ist aber $\|w\| > 0$ und wir können für t den speziellen Wert $t = t_0 = \frac{\langle v,w\rangle}{\|w\|^2}$ einsetzen und erhalten

$$\begin{aligned} 0 &\leq \|v\|^2 - 2\frac{\langle v, w\rangle\langle v, w\rangle}{\|w\|^2} + \frac{\langle v, w\rangle^2}{\|w\|^2} && \text{oder} \\ 0 &\leq \|v\|^2\|w\|^2 - 2\langle v, w\rangle^2 + \langle v, w\rangle^2 && \text{oder} \\ \langle v, w\rangle^2 &\leq \|v\|^2\|w\|^2 && \text{oder (durch Wurzelziehen)} \\ |\langle v, w\rangle| &\leq \|v\|\,\|w\|. \end{aligned}$$

□

5.6 Bemerkung

Für $t \in \mathbb{R}$ ist durch

$$P(t) = At^2 + 2Bt + C$$

mit $A = \|w\|^2,\ B = -\langle v, w\rangle$ und $C = \|v\|^2$ eine nach oben geöffnete Parabel definiert, deren Graph oberhalb der x-Achse verläuft und die ihr Minimum im Punkt $t_0 = -\frac{B}{A}$ annimmt.
Dort ist der Funktionswert $\frac{AC-B^2}{A^2}$ und dieser ist im Fall $A > 0$ genau dann größer oder gleich Null, wenn $B^2 \leq AC$ gilt. Denn im Fall $A \neq 0$ kann man $P(t)$ in folgender Form schreiben:

$$P(t) = A\left(t + \frac{B}{A}\right)^2 + \frac{AC - B^2}{A^2}$$

Aus dieser Darstellung kann man die Behauptung ablesen.
$B^2 \leq AC$ ist gerade die Cauchy-Schwarzsche Ungleichung.

Der *Beweis* im komplexen Fall ist eine kleine Modifikation des Beweises im Fall $\mathbb{K} = \mathbb{R}$, besitzt aber nicht mehr die geometrische Interpretation. Man benutzt genauso die positive Definitheit und die Eigenschaften des Skalarproduktes:

Für alle $t \in \mathbb{C}$ gilt:

$$0 \leq \quad \langle v - tw, v - tw\rangle = \|v\|^2 - t\langle v, w\rangle - \bar{t}\,\overline{\langle v, w\rangle} + \bar{t}\langle v, w\rangle$$

Ist wieder $w = 0$, so gilt in der C.S.U. das Gleichheitszeichen.
Ist $w \neq 0$, so wähle man

$$t = t_0 := \frac{\overline{\langle v, w\rangle}}{\|w\|^2}$$

so erhält man durch Multiplikation mit $\|w\|^2$

$$\begin{aligned} 0 &\leq \|v\|^2\|w\|^2 - \langle v, w\rangle\overline{\langle v, w\rangle} = \|v\|^2\|w\|^2 - |\langle v, w\rangle|^2 \quad \text{oder} \\ |\langle v, w\rangle|^2 &\leq \|v\|^2\|w\|^2 \end{aligned}$$

□

Zusatz: In der C.S.U. gilt genau dann das Gleichheitszeichen, wenn (v, w) ein linear abhängiges System ist.

5.7 Bemerkungen

(a) Man beachte, dass beim Beweis die C.S.U. die Dimension von V keine Rolle gespielt hat.

(b) Für $V = \mathbb{R}^n$ bzw. $V = \mathbb{C}^n$ mit dem Standardskalarprodukt lautet die C.S.U. explizit

$$\left|\sum_{j=1}^{n} x_j y_j\right| \leq \sqrt{\sum_{j=1}^{n} x_j^2}\sqrt{\sum_{j=1}^{n} y_j^2} \tag{5.1}$$

bzw.

$$\left|\sum_{j=1}^{n} \bar{z}_j w_j\right| \leq \sqrt{\sum_{j=1}^{n} |z_j|^2}\sqrt{\sum_{j=1}^{n} |w_j|^2} \tag{5.2}$$

(5.1) ist dabei wieder ein Spezialfall von (5.2).

Diese Ungleichungen für endliche Summen kann man auch durch vollständige Induktion beweisen. Das ist aber recht mühselig.
(5.1) ist äquivalent zur Gleichung

$$\sum_{j=1}^{n} |x_j y_j| \leq \sqrt{\sum_{j=1}^{n} x_j^2}\sqrt{\sum_{j=1}^{n} y_j^2}$$

und (5.2) ist äquivalent zur Gleichung

$$\sum_{j=1}^{n} |\bar{z}_j w_j| \leq \sqrt{\sum_{j=1}^{n} |z_j|^2} \sqrt{\sum_{j=1}^{n} |w_j|^2} \ ,$$

dabei kann man $\bar{z}_j$ auch durch z_j ersetzen (wegen $|\bar{z}_j| = |z_j|$)
Diese Ungleichungen werden deshalb häufig auch als Cauchy-Schwarzsche Ungleichungen bezeichnet.

(c) Im Beispiel lautet die C.S.U.

$$\left| \int_0^1 f(t)g(t) \right| \quad \leq \quad \left(\int_0^1 f(t)^2 \, dt \right)^{\frac{1}{2}} \left(\int_0^1 g(t)^2 \, dt \right)^{\frac{1}{2}}$$

bzw.

$$\left| \int_0^1 \bar{f}(t)g(t) \, dt \right| \quad \leq \quad \left(\int_0^1 |f(t)|^2 \, dt \right)^{\frac{1}{2}} \left(\int_0^1 |g(t)|^2 \, dt \right)^{\frac{1}{2}} \ .$$

Dass die euklidische Norm in $\mathbb{R}^n$ (ebenso in $\mathbb{C}^n$) und die euklidische Metrik, die in $\S$5.1 aufgezählten Eigenschaften haben, ist schon offensichtlich bis auf die Dreiecksungleichung:
Diese ergibt sich (in einem beliebigen Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$) so:
Man berechnet für $v, w \in V$ das Skalarprodukt $v + w$ mit sich selbst

$$\begin{aligned}
\|v+w\|^2 = \langle v+w, v+w \rangle \quad &= \quad \langle v,v \rangle + \langle w,v \rangle + \langle v,w \rangle + \langle w,w \rangle \\
&= \quad \|v\|^2 + \overline{\langle v,w \rangle} + \langle v,w \rangle, \|w\|^2 \\
&= \quad \|v\|^2 + 2Re\langle v,w \rangle + \|w\|^2 \\
&\leq \quad \|v\|^2 + 2|\langle v,w \rangle| + \|w\|^2 \text{ (da } |Re(z)| \leq |z| \text{ für alle } z \in \mathbb{C}) \\
&\leq \quad \|v\|^2 + 2\|v\| \, \|w\| + \|w\|^2 \ \text{ (nach C.S.U.)} \\
&= \quad (\|v\| + \|w\|)^2
\end{aligned}$$

Wegen der Monotonie der Quadratwurzel folgt die Dreiecksungleichung

$$\|v + w\| \leq \|v\| + \|w\| \ .$$

Die entsprechende Dreiecksungleichung für die induzierte Metrik

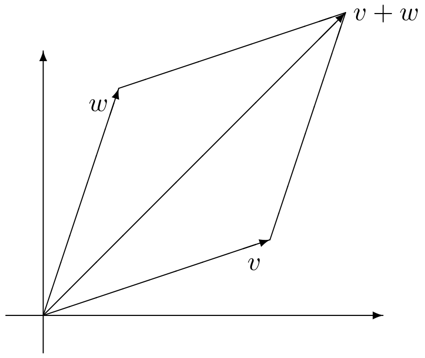
$$d(v, w) = \|v - w\|$$

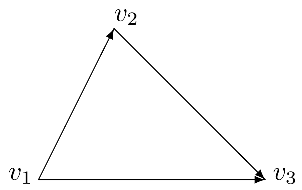
folgt aus der Dreiecksungleichung für die Norm.
Setzt man $v := v_2 - v_1;\ w := v_3 - v_1$, dann ist

$$v + w = v_3 - v_1$$

und es folgt

$$\begin{aligned}
d(v_1, v_3) = \|v_1 - v_3\| = \|v_3 - v_1\| = \|v + w\| \quad &\leq \quad \|v\| + \|w\| \\
&= \quad \|v_2 - v_1\| + \|v_1 - v_2\| \\
&= \quad \|v_1 - v_2\| + \|v_2 - v_3\| \\
&= \quad d(v_1, v_2) + d(v_2, v_3)
\end{aligned}$$





Jeder Vektorraum V mit Skalarprodukt ist ein *normierter Raum* im Sinne der folgenden Definition.

5.8 Definition (Norm, normierter Raum)

Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$\begin{array}{rcl} \| \ \| : V & \to & \mathbb{R} \\ v & \mapsto & \|v\| \end{array}$$

heißt *Norm* auf V, falls für alle $v, w \in V$ und alle $\lambda \in \mathbb{K}$ gilt

$(N_1): \ \|v\| = 0 \iff v = 0$
$(N_2): \ \|\lambda v\| = |\lambda| \ \|v\|$
$(N_3): \ \|v + w\| \leq \|v\| + \|w\|$ (Dreiecksungleichung)

Die (nicht negative!) reelle Zahl $\|v\|$ heißt *Norm* (auch Betrag, Länge) des Vektors v. Ein Vektor v mit $\|v\| = 1$ heißt *Einheitsvektor*.

Ein Vektorraum V zusammen mit einer Norm $\| \ \|$ auf V heißt *normierter Raum*.

Beispiele:

(a) $V = \mathbb{R}$; $\|x\| = |x| =$ Betrag von x; $(x \in \mathbb{R})$

(b) $V = \mathbb{C} = \mathbb{R}^2$; $\|z\| = |z| =$ Betrag von z; $(z \in \mathbb{C})$

(c) $V = \mathbb{K}^n$; $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$
$\|z\| = \|z\|_2 = \sqrt{|z_1|^2 + \cdots + |z_n|^2} = \sqrt{\langle z, z \rangle}$, dabei sei $\langle z, z \rangle$ das Standardskalarprodukt.

(d) Ist V ein beliebiger $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle \ , \ \rangle$, dann ist Norm der bewiesene Satz (diese heißt auch manchmal *Hilbert-Norm*).
Auch der Begriff des Abstandes lässt sich axiomatisieren.

5.9 Definition (Metrik, metrischer Raum)

Ist X eine nicht leere Menge. Eine Abbildung

$$\begin{array}{rcl} d : X \times X & \to & \mathbb{R} \\ (x, y) & \mapsto & d(x, y) \end{array}$$

heißt *Metrik auf* X, falls für alle $x, y, z \in X$ gilt:

$(M_1): \ d(x,y) = 0 \iff x = y$
$(M_2): \ d(x,y) = d(y,x)$ (Symmetrie)
$(M_3): \ d(x,z) \leq d(x,y) + d(y,z)$ (Dreiecksungleichung)

Die (nicht negative!) reelle Zahl $d(x,y)$ heiße *Abstand* oder *Distanz* zwischen x und y.
X zusammen mit einer Metrik d auf X heißt *metrischer Raum*.

5.10 Beispiele

(a) Ist V ein normierter $\mathbb{K}$-Vektorraum, dann ist für alle $v, w \in V$ durch

$$d(v,w) := \|v - w\|$$

eine Metrik auf V definiert.
Speziell sind $V = \mathbb{R},\ V = \mathbb{C} = \mathbb{R}^2$ und allgemeiner $V = \mathbb{R}^n$ bzw. $V = \mathbb{C}^n$ metrische Räume bezüglich der mit der euklidischen Norm definierten Metrik.

(b) Nicht jede Metrik stammt von einer Norm, z.B. gilt dies für das folgende Beispiel:
Ist $V = \mathbb{R}$ und

$$\delta(x,y) = \begin{cases} 0, & \text{für } x = y \\ 1, & \text{für } x \neq y \end{cases}$$

dann ist δ eine Metrik auf X. Sie heißt *diskrete Metrik* auf $\mathbb{R}$.

(c) Auch stammt nicht jede Norm von einem Skalarprodukt, denn wenn eine Norm von einem Skalarprodukt abgeleitet ist, dann muss für alle $v, w \in V$ gelten

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2)$$

(Parallelogrammidentität) und im Fall $\mathbb{K} = \mathbb{R}$

$$\|v + w\|^2 - \|v - w\|^2 = \frac{1}{4}\langle v, w\rangle$$

(sog. *Polarisationsgleichung*)
Geben Sie für den Fall $V = \mathbb{R}^2$ geometrische Interpretation von dieser Gleichung.
Eine wichtige Norm auf $\mathbb{R}^2$ (auch auf $\mathbb{C}$) ist die *Maximumsnorm*.

5.11 Definition (Maximumsnorm)

Für $x = (x_1, ..., x_n) \in \mathbb{K}^n$ heißt

$$\|x\|_\infty = max\{|x_1|, |x_2|, ..., |x_n|\}$$

die *Maximumsnorm* von x.

Dass es sich tatsächlich um eine Norm handelt ist leicht nachzurechnen. Die Maximumsnorm stammt aber nicht von einem Skalarprodukt auf $\mathbb{K}^n$, denn sie erfüllt *nicht* die Parallelogramm-Identität.

Vergleicht man $\|x\|_\infty$ mit der euklidischen Norm $\|x\|_2$ auf $\mathbb{K}^n$, so gelte offensichtlich die Ungleichung

$$|x_i| \leq max\{|x_1|, ..., |x_n|\} = \|x\|_\infty \ \leq \sqrt{|x_1|^2 + \cdots + |x_n|^2} = \|x\|^2 \leq \sqrt{n}\, \|x\|_\infty$$

5.12 Definition (Eins-Norm)

Eine weitere Norm auf $\mathbb{K}^n$ ist für $x = (x_1, ..., x_n) \in \mathbb{K}^n$ die Definition

$$\|x\|_1 := \sum_{i=1}^{n} |x_i| \qquad \text{(sog. } \textit{Eins-Form}\text{)}$$

Hier gilt

$$\|x\|_2 \leq \|x\|_1 \qquad \text{und} \qquad \|x\|_1 = \sum_{i=1}^{n} 1 \cdot |x_i| \leq \left(\sum_{i=1}^{n} 1^2\right)^{\frac{1}{2}} \cdot \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} = \sqrt{n}\|x\|_2$$

(letztes nach der Cauchy-Schwarzschen Ungleichung)

Fassen wir zusammen:

5.13 Satz

Für die drei Normen $\| \ \|_1, \ \| \ \|_2$ und $\| \ \|_\infty$ auf $\mathbb{K}^n$ gelten die Ungleichungen

$$\boxed{\begin{array}{c} \|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\, \|x\|_\infty \\ \text{und} \\ \frac{1}{\sqrt{n}}\|x\|_1 \leq \|x\|_2 \leq \|x\|_1 \end{array}}$$

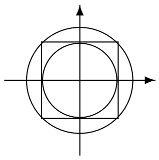
Analoge Ungleichungen gelten für die an diesen Normen abgeleiteten Metriken. Für die Konvergenztheorie in $\mathbb{R}^n$ hat das zu Folge, dass man jeweils den gleichen Konvergenzbegriff erhält. Definiert man nämlich für $a \in \mathbb{K}^n$ und $\varepsilon > 0$, die sog. $\mathbb{K}$-*Kugel* bezüglich der entsprechenden Metriken d_2, d_∞, d_1

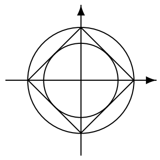
$$\begin{aligned} U_\varepsilon^{d_2}(a) &:= \{x \in \mathbb{K}^n;\ d_2(x,a) < \varepsilon\} \ , \\ U_\varepsilon^{d_\infty}(a) &:= \{x \in \mathbb{K}^n;\ d_\infty(x,a) < \varepsilon\} \ , \\ U_\varepsilon^{d_1}(a) &:= \{x \in \mathbb{K}^n;\ d_1(x,a) < \varepsilon\} \ , \end{aligned}$$

dann gilt

$$\begin{aligned} U_\varepsilon^{d_2}(a) &\subset U_\varepsilon^{d_\infty}(a) \subset \sqrt{n}\, U_\varepsilon^{d_2}(a) \ , \\ U_\varepsilon^{d_1}(a) &\subset U_\varepsilon^{d_2}(a) \subset \sqrt{n}\, U_\varepsilon^{d_1}(a) \ . \end{aligned}$$

Für den Fall $n = 2, \ a = 0$ und $\varepsilon = 1$ sind die Inklusionen in den folgenden Skizzen veranschaulicht.





Hierauf kommen wir im nächsten Kapitel ($\mathbb{C} = \mathbb{R}^2$) und in der Vorlesung Analysis 2 zurück.

Wir geben zum Schluss für den Fall $V = \mathbb{R}^n$ mit dem Standardskalarprodukt noch eine *geometrische Interpretation des Skalarprodukts* und der *Cauchy-Schwarzschen Ungleichung*.
Wir gehen aus von folgendem *Problem*:

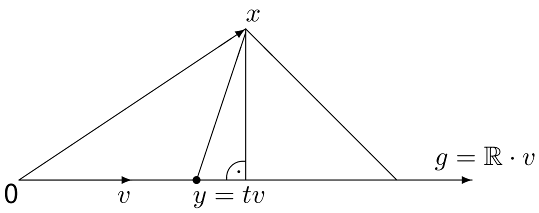
Gegeben seien $x, v \in \mathbb{R}^n$ und v sei ein Einheitsvektor, d.h. $\|v\| = \|v\|_2 = 1$.
Wir suchen auf der Geraden

$$g := \mathbb{R}v := \{tv, t \in \mathbb{R}\}$$

einen Punkt, der minimalen Anstand von x hat, mit anderen Worten: wir suchen das Minimum von

$$\|x - tv\| \text{ für } t \in \mathbb{R},$$

falls es ein solches gibt.
Wir nehmen zunächst $x \notin g$ an. Wir vermuten, dass die „orthogonale Projektion" von x auf g gerade dieses Minimum ergibt.



Dazu betrachten wir das Quadrat des Abstandes eines beliebigen Punktes $tv \in g$ von x:

$$\begin{aligned}
\|x - tv\|^2 &= \langle x - tv, x - tv\rangle = \\
&= \langle x, x\rangle - t\langle v, x\rangle - t\langle x, v\rangle + t^2\langle v, v\rangle = \\
&= \|x\|^2 - 2t\langle x, v\rangle + t^2 \quad (\langle v, v\rangle = 1) \\
&= \|x\|^2 - \langle x, v\rangle^2 + (\langle x, v\rangle - t)^2
\end{aligned}$$

Da $(\langle x, v\rangle - t)^2$ als Quadrat immer größer oder gleich Null ist und gleich Null genau dann, wenn $t = \langle x, v\rangle$ gilt, sieht man, dass tatsächlich ein Minimum existiert und dass für die reelle $t = \langle x, v\rangle$ erreicht wird.

Damit haben wir folgendes:

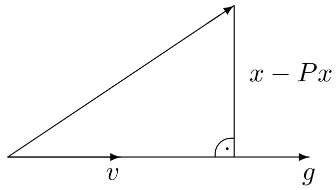
Satz: Seien $x, v \in \mathbb{R}^n, \|v\| = 1, g := \mathbb{R}v$, dann gibt es einen eindeutig bestimmten Punkt $y \in g$, der von x minimalen Abstand hat. Dieser Punkt ist gegeben durch $y := Px := \langle x, v\rangle v$.

Für das Abstandsquadrat gilt:

(a) $\|x - Px\|^2 = \|x\|^2 - \langle x, v\rangle^2$ (Pythagoras)
Wegen $\|x - Px\|^2 > 0$ folgt

(b) $\langle x, v\rangle^2 \leq \|x\|^2$.
Für alle $t \in \mathbb{R}$ gilt

(c) $\langle x - Px, tv\rangle$=0,
d.h. der Richtungsvektor $x - Px$ ist orthogonal zu g, d.h. zu allen Vektoren aus g.
Px ist also der Fusspunkt des Lotes von x auf g, deshalb heißt Px auch die *orthogonale Projektion* von x auf g.



Die Relation (c) lässt sich leicht nachprüfen (beachte dabei $\langle v,v\rangle = 1$):

$$\begin{aligned}\langle v - Px, tv\rangle &= \langle x - \langle x,v\rangle, tv\rangle \\ &= t\langle x,v\rangle - \langle\langle x,v\rangle v, tv\rangle \\ &= t\langle x,v\rangle - t\langle x,v\rangle\langle v,v\rangle \quad (\langle v,v\rangle = 1) \\ &= t(\langle x,v\rangle - \langle x,v\rangle) = 0\end{aligned}$$

Aus (1) erhält man ganz einfach die *Cauchy-Schwarzsche Ungleichung in* $\mathbb{R}^n$:
Für alle $x, y \in \mathbb{R}^n$ gilt

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

wobei das Gleichheitszeichen genau dann gilt, wenn das System (x, y) linear abhängig ist.

Beweis : Sei $y = 0$, dann ist

$$\langle x,y\rangle = \langle x,0\rangle = 0 = \|x\|\,\|y\|\ ,$$

und x, y sind linear abhängig. (Weil einer der Vektoren des Systems ein Nullvektor ist)

Ist $y \neq 0$, so können wir den Einheitsvektor $v = \frac{y}{\|y\|}$ betrachten. Es ist also $y = \|y\| \cdot v$ mit $\|v\| = 1$. Nach (b) folgt daher

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

Das Gleichheitszeichen gilt genau dann, wenn $|\langle x,v\rangle| = \|x\|$ und das bedeutet nach (a)

$$x = \langle x,v\rangle\,v = \frac{\langle x,v\rangle}{\|y\|}\,y = \lambda y$$

mit $\lambda := \frac{\langle x,v\rangle}{\|y\|} \in \mathbb{R}$, das bedeutet, dass x und y linear abhängig sind.

$\square$

Für Vektoren $x, y \in \mathbb{R}^n$, $x \neq 0$, $y \neq 0$ kann man die C.S.U.

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

in der Form

$$-1 \leq \frac{\langle x,y\rangle}{\|x\|\,\|y\|} \leq 1$$

schreiben.

Ein Blick auf den Graphen des *Cosinus* (Abb. 8) zeigt, dass die reelle Zahl $\frac{\langle x,y\rangle}{\|x\|\,\|y\|}$ im Wertebereich von $\cos$ liegt.

Da $\cos$ im Intervall $[0,\pi]$ streng monoton fällt (das werden wir später beweisen), gibt es genau ein $\varphi \in [0,\pi]$ mit

$$\cos\varphi = \frac{\langle x,y\rangle}{\|x\|\,\|y\|}.$$

φ heißt „*Winkel zwischen* x *und* y“. Schreibweise: $\varphi = \measuredangle(a,b)$.

Die Definition lautet also:

$$\varphi = \measuredangle(x,y) \iff 0 \leq \varphi \leq \pi \qquad \text{und} \qquad \cos\varphi = \frac{\langle x,y\rangle}{\|x\|\,\|y\|}$$

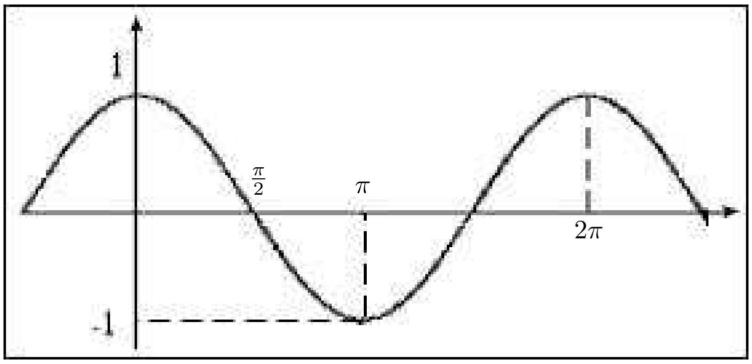


Abbildung 8: Graph der Funktion $f(x) = \cos x$.

oder für Formelgläubige (falls man den $\arccos$ kennt)

$$\boxed{\varphi = \arccos \frac{\langle x,y\rangle}{\|x\|\,\|y\|}}$$

Wegen

$$\frac{\langle x,y\rangle}{\|x\|\,\|y\|} = \frac{\langle \frac{1}{\|x\|}x, y\rangle}{\|y\|} = \frac{\langle y, \frac{x}{\|x\|}\rangle}{\|y\|} = \frac{\langle x, \frac{y}{\|y\|}\rangle}{\|x\|}$$

wird $\cos\measuredangle(x,y)$ auch als *Richtungscosinus von y in Richtung x* bzw. als *Richtungscosinus von x in Richtung y* interpretiert.

Nach der obigen Formel ist aber

$$\boxed{\langle x,y\rangle = \|x\|\,\|y\|\,\cos\measuredangle(x,y)} \quad (*)$$

Diese Formel gilt auch noch, wenn $x = 0$ oder $y = 0$ ist.
Die Identität

$$\|x+y\|^2 = \|x\|^2 + 2\|x\|\,\|y\|\cos\varphi + \|y\|^2$$

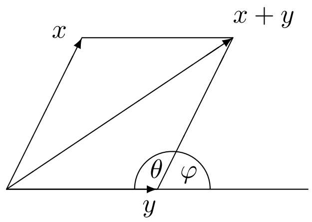
schreibt sich jetzt mit $\measuredangle(x,y) =: \varphi$ so:

$$\boxed{\|x+y\|^2 = \|x\|^2 + 2\|x\|\,\|y\|\cos\varphi + \|y\|^2}$$

Man kann diese Formel als *Cosinus-Satz* interpretieren. Wenn man $\theta + \varphi = \pi$ beachtet, erhält man die klassische Form

$$\boxed{\|x+y\|^2 = \|x\|^2 - 2\|x\|\,\|y\|\cos\theta + \|y\|^2}$$

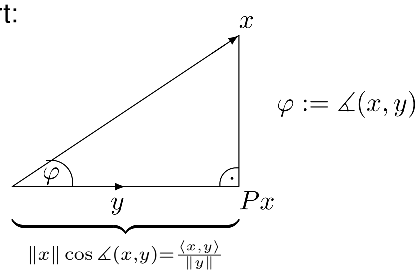
Dies Aussage des Satzes von Pythagoras gilt genau dann, wenn $x\perp y$, d.h. $\theta = \varphi = \frac{\pi}{2}$ gilt.



In physikalischen Anwendungen (z.B. Arbeit = Kraft · Weg) wird das Skalarprodukt zwischen zwei Vektoren x und y manchmal durch die Gleichung $(*)$ definiert:

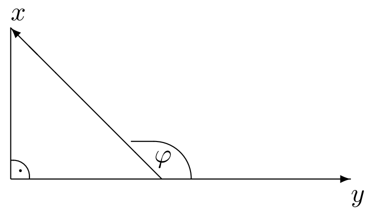
$$\langle x,y\rangle = \|x\|\,\|y\|\cos\measuredangle(x,y)\ ,$$

wobei $\|x\|\cos\measuredangle(x,y)$ die Länge der orthogonalen Projektion Px von x auf die Gerade $g = \mathbb{R}y$ ist.



Allerdings muss man das Vorzeichen beachten

$$\cos\measuredangle(x,y) = \begin{cases} > 0, \text{ falls } 0 \leq \measuredangle(x,y) < \frac{\pi}{2} \\ < 0, \text{ falls } \frac{\pi}{2} < \measuredangle(x,y) \leq \pi \end{cases}$$



Dass in $\mathbb{R}^2$ die obige Winkeldefinition mit der „üblichen“ übereinstimmt, kann man mittels Polarkoordinaten zeigen. Wir stellen zum Abschluss (ohne Beweis) einige Eigenschaften von $\measuredangle(x,y)$ zusammen.

Eigenschaften von $\measuredangle(x,y)$:

(a) $\measuredangle(x,y) = \measuredangle(y,x)$

(b) $\measuredangle(x,-y) = \pi - \measuredangle(x,y)$

(c) $\measuredangle(rx,sy) = \begin{cases} \measuredangle(x,y), \text{ falls } rs > 0 \\ \pi - \measuredangle(x,y), \text{ falls } rs < 0 \end{cases}$

(d) $\measuredangle(x,y) = \begin{cases} 0 \\ \pi \end{cases} \iff$ es gibt ein $r \in \mathbb{R}$ mit $x = ry$

(e) $\measuredangle(x,y) = \frac{\pi}{2} \iff \langle x,y\rangle = 0 \iff x \perp y$

(f) In Fall $n = 3$, also in $\mathbb{R}^3$ gelten folgende Zusammenhänge zwischen Skalarprodukt, Vektorprodukt und $\measuredangle(x,y)$:

Für alle $x,y,z \in \mathbb{R}^3$ gilt:

1. $x \times (y \times z) = \langle x,z\rangle y - \langle x,y\rangle z$ bzw. $(x \times y) \times z = \langle x,z\rangle y - \langle y,z\rangle x$ (Graßmann-Identitäten)
2. $\|x \times y\|^2 = \|x\|^2\|y\|^2 - \langle x,y\rangle^2$ (spezielle Lagrangesche Identität)

($(x \times y)$ ist ein Vektor, der auf der von x und y aufgespannten Ebene (x und y seien dabei linear unabhängig) senkrecht steht und dessen Länge (Norm) den Flächeninhalt des von x und y aufgespannten Parallelogramms ist.)

Ferner gilt die sog. *Jacobi-Identität*:

$$(a \times b) \times c + (b \times c) \times a + (c \times a) \times b = 0$$

6 Einige nützliche Ungleichungen

Wir stellen im Folgenden einige *nützliche Ungleichungen* zusammen, die wir zum *Basiswissen* rechnen und die in jedem Zusammenhang richtig erkannt werden sollten.

Für $x \in \mathbb{R}$ gilt stets

6.1 $x^2 \geq 0$, und $x^2 = 0$ ist mit $x = 0$ äquivalent.

Aus 6.1 folgt schon eine Reihe von Ungleichungen, die lediglich Äquivalenzumformungen von 6.1 sind.

Für alle $a, b \in \mathbb{R}$ gilt:

6.2 $a^2 + b^2 \geq 2ab$ bzw.

6.3 $(\frac{a+b}{2})^2 \geq ab$ (binomische Ungleichung)

Beweis : Setzt man $x := a - b$, dann gilt $x^2 \geq 0$ nach 6.1 und

$$\begin{aligned} 0 \leq x^2 = (a-b)^2 = a^2 - 2ab + b^2 & \qquad \Longleftrightarrow \\ 2ab \leq a^2 + b^2 & \qquad \Longleftrightarrow \\ 2ab + 2ab \leq a^2 + b^2 + 2ab & \qquad \Longleftrightarrow \\ 4ab \leq (a+b)^2 & \qquad \Longleftrightarrow \\ ab \leq \left(\frac{a+b}{2}\right)^2 . & \end{aligned}$$

□

Sind $a, b \in \mathbb{C}$, so sind $|a|,\ |b| \in \mathbb{R}$ und es gilt für beliebige komplexe Zahlen $a,\ b$

6.4 $|a|\,|b| = |ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$.

Das ist ein Spezialfall der sog. *Young-schen Ungleichung*

$$\boxed{|ab| \leq \tfrac{1}{p}|a|^p + \tfrac{1}{q}|b|^q}$$

$a, b \in \mathbb{C};\ p, q \in \mathbb{R}$ mit $p, q > 1$ und $\frac{1}{p} + \frac{1}{q} = 1$.

Sind $a,\ b \in \mathbb{R}$ und ist $a > 0$ und $b > 0$, dann folgt aus 6.2 und 6.3

6.5 $\frac{a+b}{2} \geq \sqrt{ab}$ bzw.

6.6 $\sqrt{ab} \geq \frac{2ab}{a+b}$ bzw.

6.7 $\frac{a+b}{2} \leq \sqrt{\frac{a^2+b^2}{2}}$

Nennt man (vergleiche Übungsblatt 1):

$$\left.\begin{array}{ll} A(a,b) := \frac{a+b}{2} & \text{das arithmetische} \\ G(a,b) := \sqrt{ab} & \text{das geometrische} \\ H(a,b) = \frac{2ab}{a+b} & \text{das harmonische} \\ Q(a,b) = \sqrt{\frac{a^2+b^2}{2}} & \text{das quadratische} \end{array}\right\} \text{Mittel von } a \text{ und } b,$$

so gilt also die *Ungleichungskette*

$$\boxed{min\{a,b\} \leq H(a,b) \leq G(a,b) \leq A(a,b) \leq Q(a,b) \leq max\{a,b\}}$$

Man beachte, jede der Ungleichungen 6.2, 6.3, 6.4, 6.5, 6.6, 6.7 ist im Fall $a \neq 0$, $b \neq 0$ mit

$$\begin{array}{rcll} \frac{|a|}{|b|} + \frac{|b|}{|a|} & \geq & 2 & \text{und damit mit} \\ x + \frac{1}{x} & \geq & 2 & \text{für } x > 0 \text{ äquivalent.} \end{array}$$

Diese Ungleichungen lassen sich durch Induktion nach n sofort auf n reelle Zahlen verallgemeinern.
Definiert man für $a_1, \ldots, a_n \in \mathbb{R}$

$$\begin{array}{rll} A(a_1,\ldots,a_n) := & \frac{a_1 + a_2 + \ldots + a_n}{n} & \text{(arithmetisches Mittel)} \\ Q(a_1,\ldots,a_n) := & \sqrt{\frac{a_1^2 + \ldots + a_n^2}{n}} & \text{(quadratisches Mittel)} \\ G(a_1,\ldots,a_n) := & \sqrt[n]{a_1 a_2 \ldots a_n} & \text{(geometrisches Mittel)} \\ & & \text{(hier wird } a_1 \geq 0, \ldots, a_n \geq 0 \text{ vorausgesetzt)} \\ & & \text{und} \\ H(a_1,\ldots,a_n) := & \frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \ldots + \frac{1}{a_n}} = & \\ & = \frac{n}{A(\frac{1}{a_1},\ldots,\frac{1}{a_n})} & \text{(harmonisches Mittel)} \end{array}$$

wobei man hier $a_1 \neq 0, \ldots, a_n \neq 0$ voraussetzen muss, dann gilt die *Ungleichungskette* (für positive $a_1, \ldots, a_n$)

6.8
$$\begin{aligned} min\{a_j;\ 1 \leq j \leq n\} &\leq H(a_1,\ldots,a_n) \leq G(a_1,\ldots,a_n) \leq \\ &\leq A(a_1,\ldots,a_n) \leq Q(a_1,\ldots,a_n) \leq max\{a_j;\ 1 \leq j \leq n\} \end{aligned}$$

Wir skizzieren, wie man etwa

$$(*) \qquad G(a_1,\ldots,a_n) \leq A(a_1,\ldots,a_n)$$

beweisen kann.
OBdA können wir $a_1 > 0, \ldots, a_n > 0$ annehmen. Die Fälle $n = 1$ und $n = 2$ sind klar.

Also können wir mit Induktion schließen:

Die Ungleichung $(*)$ gelte für je n positive reelle Zahlen. Sind nun $n+1$ positive reelle Zahlen $a_1, \ldots, a_n, a_{n+1}$ vorgegeben, so können wir durch Umnummerieren $a_{n+1} \geq a_1, \ldots, a_n$ erreichen.

Dann ist offensichtlich

$$\begin{aligned}\alpha := \frac{a_1+\ldots+a_n}{n} &\le \frac{a_{n+1}}{n}+\ldots+\frac{a_{n+1}}{n}\\ &= \left(\frac{1}{n}+\ldots+\frac{1}{n}\right)a_{n+1}\\ &= n\frac{1}{n}a_{n+1} = a_{n+1}\ ,\end{aligned}$$

also: $h := \dfrac{a_{n+1}-\alpha}{(n+1)\alpha} \ge 0$.

Mit der Bernoullischen Ungleichung 6.9 folgt jetzt

$$\left(\frac{a_1+a_2+\ldots+a_n}{(n+1)\alpha}\right)^{n+1} = (1+h)^{n+1} \ge 1+(n+1)h = \frac{a_{n+1}}{\alpha}\ .$$

Mit der Induktionsvoraussetzung (I.V.) folgt hieraus

$$\left(\frac{a_1+a_2+\ldots+a_{n+1}}{n+1}\right)^{n+1} \ge \alpha^{n+1}\,\frac{a_{n+1}}{\alpha} = \alpha^n\,a_{n+1} \overset{\text{(I.V.)}}{\ge} a_1\cdot\ldots\cdot a_n\cdot a_{n+1}\ .$$

□

In §2.4.2 haben wir die Bernoullische Ungleichung bewiesen:

6.9 *Bernoullische Unlgeichung*

Für alle $h\in\mathbb{R}$ mit $h\ge -1$ und alle $n\in\mathbb{N}$ gilt

$$(1+h)^n \ge 1+nh\ .$$

Zusätze:
Die Bernoullische Ungleichung gilt sogar für alle $h\in\mathbb{R}$ mit $h\ge -2$ und alle $n\in\mathbb{N}$. Ist $h>-1$, $h\neq 0$, dann gilt für $n\ge 2$ die strikte Ungleichung $(1+h)^n > 1+nh$.

Der Beweis der Zusätze sei als Übungsaufgabe gestellt.

6.10 *Abschätzung für die endliche geometrische Reihe*

Für $q\in\mathbb{R}$ mit $0<q<1$ und alle $n\in\mathbb{N}$ gilt

$$\boxed{1+q+q^2+\ldots+q^n < \tfrac{1}{1-q}}$$

nach der Summenformel $\S 2.4(1)$ gilt für $q\neq 1$

$$1+q+q^2+\ldots+q^n = \frac{1-q^{n+1}}{1-q} < 1-q\ ,$$

denn $1-q^{n+1}<1 \iff 0<q^{n+1}$ (beachte $q>0$).

Für einen beliebigen $\mathbb{K}$-Vektorraum V mit Skalarprodukt $\langle\ ,\ \rangle$ haben wir die *Cauchy-Schwarzsche-Ungleichung* bewiesen und für den Fall des Standardvektorraumes $\mathbb{R}^n$ einen einfachen geometrischen Beweis gegeben.

Wir formulieren die C.S.U. nochmal für den Standardvektorraum $\mathbb{K}^n$ mit dem Standardskalarprodukt

$$\langle z,w\rangle := \sum_{j=1}^{n}\bar z_j w_j\ ,$$

dabei sei

$$z := (z_1, \ldots, z_n) \in \mathbb{K}^n , \qquad w := (w_1, \ldots, w_n) \in \mathbb{K}^n$$

Die C.S.U. lautet:

6.11 $\left|\sum\limits_{j=1}^{n} \bar{z}_j w_j\right| \leq \sqrt{\sum\limits_{j=1}^{n} |z_j|^2}\sqrt{\sum\limits_{j=1}^{n} |w_j|^2}$,

Da sie für alle $z_1, \ldots, z_n \in \mathbb{C}$ gilt, ist 6.11 äquivalent mit

6.12 $\left|\sum\limits_{j=1}^{n} z_j w_j\right| \leq \sqrt{\sum\limits_{j=1}^{n} |z_j|^2}\sqrt{\sum\limits_{j=1}^{n} |w_j|^2}$.

Sind die z_j bzw. w_j alle reell, so kann man auf der rechten Seite die Beträge weglassen. Manchmal nennt man auch die Ungleichung

6.13 $\sum\limits_{j=1}^{n} |z_j w_j| \leq \sqrt{\sum\limits_{j=1}^{n} |z_j|^2}\sqrt{\sum\limits_{j=1}^{n} |w_j|^2}$ Cauchy-Schwarzsche-Ungleichung.

6.12 folgt nach der Dreiecksungleichung

$$\left|\sum_{j=1}^{n} z_j w_j\right| \leq \sum_{j=1}^{n} |z_j w_j| = \sum_{j=1}^{n} |z_j|\,|w_j| \leq \sqrt{\sum_{j=1}^{n} |z_j|^2} \cdot \sqrt{\sum_{j=1}^{n} |w_j|^2}$$

sofort aus 6.13.

Es gilt aber auch die Umkehrung:

Gilt nämlich 6.11 für alle z_j bzw. $w_j \in \mathbb{K}$, so gilt die Ungleichung auch, wenn man z_j bzw. w_j durch $|z_j|$ bzw. $|w_j|$ ersetzt. Das ist aber gerade die Ungleichung 6.13. (Den äußeren Betrag kann man dann weglassen.)

Wir geben hier für 6.13 und damit für 6.12 einen weiteren Beweis, der sich auf die elementare Ungleichung 6.4

$$|ab| = |a|\,|b| \leq \frac{1}{2}\left(|a|^2 + |b|^2\right)$$

für beliebige komplexe $a,\ b$ stützt.

Sei $A := \sqrt{\sum\limits_{j=1}^{n} |z_j|^2}$, $B := \sqrt{\sum\limits_{j=1}^{n} |w_j|^2}$.

Da im Fall $A = B = 0$ nichts zu beweisen ist, können wir $A > 0$ und $B > 0$ annehmen. Mit $a_j := \frac{|z_j|}{A}$, $b_j := \frac{|w_j|}{B}$, $1 \leq j \leq n$ geht die Ungleichung 6.13 in die äquivalente Ungleichung

$$\sum_{j=1}^{n} a_j b_j \leq 1$$

über. Deren Richtigkeit ist aber mit 6.4 evident:

Es ist nämlich

$$\sum_{j=1}^{n} a_j b_j = \sum_{j=1}^{n} \frac{|z_j|}{A} \cdot \frac{|w_j|}{B} \leq \frac{1}{2}\sum_{j=1}^{n}\left(\frac{|z_j|^2}{A^2} + \frac{|w_j|^2}{B^2}\right) = \frac{1}{2}(1+1) = 1 \ .$$

Wir weisen aber ausdrücklich darauf hin, dass wir die C.S.U. für *jeden* $\mathbb{K}$-Vektorraum mit Skalarprodukt bewiesen haben. Insbesondere gilt sie auch für den Vektorraum V der stetigen komplexwertigen Funktionen auf einem Intervall $[a, b] \subset \mathbb{R}$ mit dem Skalarprodukt

$$< f, g > := \int_a^b \bar{f}(t) g(t)\, dt$$

bzw. für den *Hilbertschen Folgenraum* $l^2 = l^2(\mathbb{C})$ der quadratsummierbaren Folgen (vergleiche Aufgabenblatt 9)

Es sei auch nochmal ausdrücklich bemerkt, dass aus der C.S.U. für einen Vektorraum mit Skalarprodukt die Dreiecksungleichung für die entsprechende Norm folgt:
Ist $\|x\| = \sqrt{< x, x >}$, dann gilt für alle $x, y \in V$

$$\|x + y\| \leq \|x\| + \|y\| \, .$$

Dies ist ein Spezialfall der sog. *Minkowskischen Ungleichung* (H.Minkowski, 1864-1909)
Im Fall $V = \mathbb{K}^n$ lautet diese für p-Norm $(p \in (\mathbb{N} \cup \{\infty\}))$

6.14 $\|x + y\|_p \leq \|x\|_p + \|y\|_p$

Dabei ist $\|x\|_p = (|x_1|^p + \ldots + |x_n|^p)^{\frac{1}{p}}$ für $p \in \mathbb{N}$, $\|x\|_\infty = max\{|x_1, \ldots, |x_n|\}$.

Die Minkowski Ungleichung steht in einem Zusammenhang mit der *Hölderschen Ungleichung* (Otto Hölder, 1859-1937), deren einfachste Form so lautet:
Ist $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$ und sind $(a_1, \ldots, a_n) \in \mathbb{K}^n$, $(b_1, \ldots, b_n) \in \mathbb{K}^n$, so gilt

6.15 $\sum_{j=1}^{n} |a_j b_j| \leq \left(\sum_{j=1}^{n} |a_j|^p \right)^{\frac{1}{p}} \cdot \left(\sum_{j=1}^{n} |b_j|^q \right)^{\frac{1}{q}}$.

Der Spezialfall $p = q = 2$ ergibt wieder die klassische C.S.U. .

Die Höldersche Ungleichung gilt jedoch auch für *Integrale* und *Reihen*.

Später werden wir sehen, dass sich all diese Ungleichungen aus *Konvexitätseigenschaften* geeigneter Funktionen relativ einfach ergeben.

Zum Abschluss sei bemerkt, dass sich die *Dreiecksungleichung*

6.16 $|z + w| \leq |z| + |w|$ $\qquad (z, w \in \mathbb{C}$ oder $z, w \in \mathbb{R})$

sofort (via Induktion) auf n Summanden überträgt.

Für beliebige $z_1, \ldots, z_n \in \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) gilt

6.17 $|z_1 + z_2 + \ldots + z_n| \leq |z_1| + |z_2| + \ldots + |z_n|$.

II. Folgen und Reihen

Wie in der Einführung in diese Vorlesung schon verdeutlicht wurde, steht im Mittelpunkt der Analysis die Untersuchung von *Grenzwerten*. Viele wichtige mathematische und physikalische Begriffe lassen sich (nur) durch Grenzwerte definieren, z.B. die Geschwindigkeit, Beschleunigung, Arbeit, Energie, Leistung, Wirkung, Volumen und Oberfläche, Länge und Krümmung von Kurven, die Krümmung von Flächen. Grundidee ist dabei, dass man versucht ein unbekanntes oder schwierig zu erfassendes (mathematisches oder physikalisches) Objekt durch bekannte oder leichter zugängliche Objekte zu approximieren. Viele Größen (wie z.B. die Kreiszahl π) lassen sich nicht durch einen in endlich vielen Schritten exakt berechenbaren Ausdruck darstellen, sondern nur mit „beliebiger Genauigkeit" approximieren.

Was das genau bedeutet, wollen wir zunächst am Beispiel reeller und komplexer Zahlenfolgen verdeutlichen. Später werden wir so wichtige Funktionen, wie z.B. die Exponentialfunktion oder die Winkelfunktionen, ebenfalls mittels Grenzprozessen einführen.

Grenzprozesse (der verschiedensten Arten) sind ein wesentliches Kennzeichen der Analysis.

Quadratwurzeln aus nicht negativen Zahlen existieren zwar in $\mathbb{R}$, allgemeiner k-te Wurzeln ($k \in \mathbb{N},\ k \geq 2$), aber wir kennen bis jetzt kein systematisches Verfahren, das den Wurzelexponenten k und ein beliebiges $a \in \mathbb{R}, a \geq 0$ als Input akzeptiert und Zahlendaten vorgeschriebener Genauigkeit für $\sqrt[k]{a}$ produziert.

Frage: Wie berechnet ein Taschenrechner $\sin 23,5°$?

Wir entwickeln daher ein allgemeines Konstruktionswerkzeug für Zahlen und Funktionen, und zwar in zweierlei Hinsicht:

- erstens geht es darum, neue Objekte begrifflich zu konzipieren und formelmäßig darzustellen, wie z.B. die Eulersche Zahl
$$e = \left(\lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!} \right)$$
oder die Kreiszahl π.
- zweitens sollte man in die Lage versetzt werden, diese Objekte mit endlich vielen Schritten mit beliebiger Genauigkeit (nummerisch) zu berechnen.

Ein derartiger Konstruktionswerkzeug ist der *Folgenbegriff* und im unmittelbaren Zusammenhang damit steht der Begriff der „*Reihe*".

7 Folgen (Begriff der Konvergenz, erste Beispiele)

7.1 Definition

Ist X eine beliebige (nicht leere) Menge, so versteht man unter einer *Folge von Elementen aus* X eine Zuordnung (Abbildung), die jeder natürlichen Zahl $n \in \mathbb{N}$ ein eindeutig bestimmtes Element $a_n \in X$ zuordnet.
Man schreibt hierfür: $(a_n)_{n\in\mathbb{N}}$ oder suggestiver $(a_1, a_2, a_3, a_4, \ldots)$ oder kurz (a_n)

a_n heißt das n-te Glied der Folge, n der Folgenindex, er kann durch jedes andere Symbol ersetzt werden. Statt die natürlichen Zahlen als Indexbereich zu verwenden, kann man etwas allgemeine Menge vom Typ $\{n \in \mathbb{Z};\ n \geq k\}$ mit einer Zahl $k \in \mathbb{Z}$ betrachten, so erhält man Folgen

$$(a_n)_{n\geq k} \quad \text{oder} \quad (a_k, a_{k+1}, a_{k+2}, \ldots)$$

Häufig ist $k = 0$ oder $k = 1$.

Wir betrachten zunächst nur *reelle* oder *komplexe* Zahlenfolgen, d.h. wir wählen $X = \mathbb{R}$ oder $X = \mathbb{C}$; hierbei verwenden wir wieder die gemeinsame Bezeichnung $\mathbb{K}$.

7.2 Beispiele

(1) Ist $c \in \mathbb{K}$ ein festes Element und setzt man $a_n := c$ für alle $n \in \mathbb{N}$, so erhält man die *konstante Folge*

$$(a_n) = (c, c, c, \ldots)$$

(2) Setzt man $a_n := n$ für alle $n \in \mathbb{N}$, so erhält man

$$(a_n) = (1, 2, 3, 4, 5, 6, \ldots),$$

also die *Folge der natürlichen Zahlen*.

(3) Setz man $b_n := n^2$ für alle $n \in \mathbb{N}$, so erhält man die Folge der *Quadratzahlen*

$$(b_n) = (1, 4, 9, 16, 25, \ldots)\ .$$

Offensichtlich gilt $b_n = a_{n^2}$.

(4) Setzt man $a_n := p_n$, wobei p_n die n-te Primzahl in der natürlichen Reihenfolge ist, so erhält man die Folge der Primzahlen

$$(p_n) = (2, 3, 5, 7, 11, 13, 17, 19, \ldots)$$

(5) Setzt man $a_n := \frac{1}{n}$ für alle $n \in \mathbb{N}$, so erhält man die *Folge der Stammbrüche*

$$(a_n) = \left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots\right)$$

(6) Setzt man $a_n := (-1)^{n+1}$ für $n \in \mathbb{N}$, so erhält man die *alternierende Folge*

$$(a_n) = (1, -1, 1, -1, 1, -1, \ldots)$$

Man beachte, dass hier

$$\{a_n;\ n \in \mathbb{N}\} = \{1, -1\}$$

gilt. Man muss also zwischen einer Folge und ihrer *Wertemenge* $W(a_n) := \{a_n;\ n \in \mathbb{N}\}$ unterscheiden.

(7) Für $a_n := \frac{n}{n+1}$ für $n \in \mathbb{N}_0$, erhält man

$$(a_n) = \left(0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \dots\right)$$

(8) $a_n := \frac{n}{2^n}$ für $n \in \mathbb{N}_0$ liefert

$$(a_n) = \left(0, \frac{1}{2}, \frac{1}{2}, \frac{3}{8}, \frac{1}{4}, \frac{5}{32}, \frac{3}{32}, \dots\right)$$

(9) Häufig sind Folgen *rekursiv* definiert, z.B. erhält man für $f_1 := f_2 := 1$ und

$$f_n = f_{n-1} + f_{n-2}$$

für $n \geq 3$ eine berühmte *Fibonacci-Folge*

$$(1, 1, 3, 5, 8, 13, 21, 34, \dots)$$

(10) Für $q \in \mathbb{C}$ und $a_n := q^n$ für $n \in \mathbb{N}$ erhält man die *geometrische Folge*

$$(q^n)_{n \geq 0} = (1, q, q^2, q^3, \dots)$$

(11) $a_n := (1 + \frac{1}{n})^n,\ n \in \mathbb{N}$, liefert

$$(a_n) = \left(2, \frac{9}{4}, \frac{64}{27}, \frac{625}{265}, \dots\right) .$$

(12) $a_n := \frac{n!}{20^n}$ für $n \in \mathbb{N}$ liefert

$$(a_n) = \left(\frac{1}{20}, \frac{1}{200}, \frac{3}{4000}, \frac{3}{20000}, \frac{3}{80000}, \frac{9}{800000}, \frac{63}{16 \cdot 10^6}, \dots\right)$$

Hier geben die ersten Folgenglieder einen völlig falschen Eindruck von der „wahren Größe“ der späteren Folgenglieder, hier ist z.B.

$$a_{60} \approx 7217,3 \text{ und } a_{100} \approx 0,375 \cdot 10^{1267}$$

(13) $a_n := i^n,\ n \in \mathbb{N}_0$, liefert

$$(a_n) = (1, i, -1, -i, 1, i, -1, -i, \dots)$$

Hier gilt also $a_{n+4} = a_n$ für alle $n \in \mathbb{N}_0$.

(14) Die Folge $a_n := (1 - \frac{1}{n})(1 + i)$ für $n \in \mathbb{N}$ liefert

$$(a_n) = \left(0, \frac{1+i}{2}, \frac{2}{3}(1+i), \frac{3}{4}(1+i), \dots\right) .$$

Bei wachsenden n verhalten sich die betrachteten Folgen sehr unterschiedlich.
Während die Folge $(a_n) = (n)$ der natürlichen Zahlen „größer und größer“ wird, kommen die Glieder a_n der Folge $a_n = \frac{1}{n}$ der Zahl Null „beliebig nahe“.
Die in der Vorlesung behandelten Beispiele zur Motivation des Konvergenzbegriffes:

- Das *Babylonische Wurzelziehen* (für die k-te Wurzel, $k \geq 2$)
- *„Multiplizieren statt Dividieren"* (Lösung der Gleichung $ax = 1$ für $a \neq 0$)
- *Kreismessung nach Archimedes* (siehe Übungsaufgabe 5 auf dem Blatt 7)
- die Auswirkungen einer Investition von K Euro auf das Volkseinkommen

werden im laufenden Text als Beispiele vorkommen.

Um die obigen wagen Begriffe (wie z.B. „beliebig nahe") zu präzisieren führen wir den Begriff der ε *-Umgebung* ein.

7.3 Definition (ε -Umgebung)

Sei $\varepsilon > 0$ eine beliebige (positive) reelle Zahl und $a \in \mathbb{K}$. Dann heißt (vergl. $\S 5.13$)

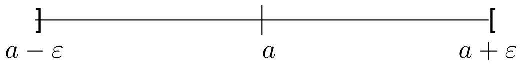
$$U_\varepsilon(a) := \{z \in \mathbb{K};\ |z - a| < \varepsilon\}$$

ε *-Umgebung von a.*

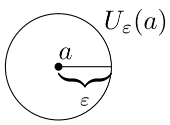
Im Fall $a \in \mathbb{R}$ ist

$$U_\varepsilon(a) = \{z \in \mathbb{R};\ |z - a| < \varepsilon\} =]a - \varepsilon, a + \varepsilon[\ ,$$

also das *offene Intervall* mit Mittelpunkt a und Randpunkten $a - \varepsilon$ bzw. $a + \varepsilon$:



Im Fall $a \in \mathbb{C}$ ist $U_\varepsilon(a)$ die *offene Kreisscheibe* mit Mittelpunkt a und Radius ε:



7.4 Definition (Konvergenz von Folgen)

Eine Folge (a_n) mit $a_n \in \mathbb{K}$ heißt *konvergent* (in $\mathbb{K}$), wenn es ein $a \in \mathbb{K}$ mit folgender Eigenschaft gibt:
Zu jeder ε*-Umgebung* $U_\varepsilon(a)$ gibt es eine natürliche Zahl $N \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt:

$$(*) \qquad a_n \in U_\varepsilon(a) \qquad (\Longleftrightarrow |a_n - a| < \varepsilon)\ .$$

Die Zahl a heißt *Grenzwert(Limes)* der Folge (a_n) und man schreibt

$$a = \lim_{n\to\infty} a_n$$

oder

$$a_n \to a \text{ für } n \to \infty$$

oder

$$a_n \overset{n\to\infty}{\longrightarrow} a.$$

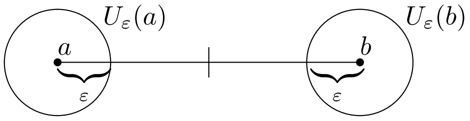
Diese Schreibweise ist jedoch erst gerechtfertigt, wenn man sich überlegt hat, dass eine konvergente Folge nicht zwei *verschiedene Grenzwerte* haben kann: Das liegt an der sogenannten

7.5 Hausdorffsche Trennungseigenschaft von $\mathbb{R}$ bzw $\mathbb{C}$

Sind $a, b \in \mathbb{K},\ a \neq b$, dann gibt es stets ε-Umgebungen $U_\varepsilon(a)$ und $U_\varepsilon(b)$ mit

$$U_\varepsilon(a) \cap U_\varepsilon(b) = \emptyset\ .$$

Kurz: Verschiedene Punkte lassen sich durch ε-Umgebungen trennen:



Wegen $a \neq b$ ist $\varepsilon := \frac{|a-b|}{3} > 0$, und es gilt

$$U_\varepsilon(a) \cap U_\varepsilon(b) = \emptyset\ .$$

Wenn (a_n) gegen c konvergiert, dann gibt es ein $N \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt

$$a_n \in U_\varepsilon(a)$$

In $U_\varepsilon(b)$ können dann höchstens die Glieder $a_1, \ldots, a_{N-1}$ liegen, d.h. eine konvergente Folge kann nicht zwei verschiedene Grenzwerte haben.

7.6 Bemerkungen

(1) Eine mit dem Grenzwert Null konvergente Folge heißt *Nullfolge*.

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

(3) Die geometrische Bedingung $(*)$ aus Def. 7.4 ist äquivalent zu $|a_n - a| < \varepsilon$, d.h. (a_n) konvergiert genau dann gegen a, wenn die Folge $(a_n - a)$ eine Nullfolge ist; letztes ist jedoch auch mit „$(|a_n - a|)$ ist Nullfolge“ äquivalent. Das bedeutet geometrisch, dass die *Folge der Abstände* eine Nullfolge ist.

Man kann die Konvergenzbedingung noch schlagkräftiger formulieren, wenn man sich der folgenden Sprechweise bedient:

7.7 Sprechweise (fast alle, schließlich alle)

Ist $E(n),\ n \in \mathbb{N}$ eine Eigenschaft, die eine natürliche Zahl haben kann oder nicht, so sagt man: $E(n)$ gilt für *fast alle* $n \in \mathbb{N}$ oder *schließlich alle* $n \in \mathbb{N}$, wenn es ein $N \in \mathbb{N}$ gibt, so dass für alle Indizes $n \in \mathbb{N}$ mit $n \geq N$ die Eigenschaft $E(n)$ richtig ist.

Mit dieser Sprechweise kann man die Konvergenzdefinition dann so fassen:

Eine Folge (a_n), $a_n \in \mathbb{K}$ konvergiert genau dann gegen $a \in \mathbb{K}$, wenn in jeder ε-Umgebung $U_\varepsilon(a)$ fast alle Glieder der Folge liegen.

Aufgabe: Formulieren Sie möglichst einfach die Aussage: Die Folge (a_n) konvergiert *nicht* gegen $a \in \mathbb{K}$.

7.8 Beispiele konvergenter und divergenter Folgen

(1) Die konstante Folge $(a_n) = (c, c, c, \ldots)$, $c \in \mathbb{K}$, ist konvergent mit dem Grenzwert c, denn es ist für alle $n \in \mathbb{N}$

$$a_n - c = 0;\ \text{damit auch für jedes } \varepsilon > 0$$
$$a_n \in U_\varepsilon(c),\ \text{sogar für } \textit{alle}\ n \in \mathbb{N}\ .$$

(2) Obwohl wir später einfacher beweisen können, dass die Folge $(a_n) = (n)$ der natürlichen Zahlen, nicht konvergiert, geben wir hier einen anderen Beweis:
Wir nehmen an, $(a_n) = (n)$ sei konvergent und $a = \lim\limits_{n\to\infty} a_n$.
Wir wählen $\varepsilon = 1$, dann gibt es $N \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$, $n \geq N$, gilt $a_n \in U_\varepsilon(a)$, d.h.

$$(*) \qquad |a_n - a| = |n - a| < 1\ .$$

Die Ungleichung $(*)$ gilt für alle $n \geq N$, also z.B. für $n := N$ und $m := N + 2$. Wir erhalten also

$$|a_n - a| = |N - a| < 1 \qquad \text{und}$$
$$|a_m - a| = |a_{N+2} - a| = |N + 2 - a| < 1\ .$$

Mit der Dreiecksungleichung folgt dann aber

$$2 = |2| = |(N + 2 - a) - (N - a)| \leq |N + 2 - a| + |N - a| < 1 + 1 = 2\ ,$$

also ein Widerspruch. □

(3) Dass $(a_n) = \left(\frac{1}{n}\right)$ eine *Nullfolge* ist, ist mit der *Archimedischen Eigenschaft* von $\mathbb{R}$ äquivalent:
Ist nämlich $\varepsilon > 0$ beliebig vorgegeben, dann gibt es wegen der Archimedischen Eigenschaft von $\mathbb{R}$ ein $N \in \mathbb{N}$ mit $N > \frac{1}{\varepsilon}$.

Diese Bedingung ist äquivalent zu $\frac{1}{N} < \varepsilon$. Ist nun $n \in \mathbb{N}$, $n \geq N$, so gilt

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

d.h. $(a_n) = (\frac{1}{n})$ ist eine Nullfolge.

(4) Betrachtung der ersten Folgenglieder $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots$ lassen die Vermutung aufkommen, dass die Folge

$$(a_n) = \left(\frac{n}{n+1}\right) = \left(\frac{n+1-1}{n+1}\right) = \left(1 - \frac{1}{n+1}\right)$$

den Grenzwert 1 hat, denn es ist

$$|a_n - 1| = |1 - a_n| = \left|1 - \frac{n}{n+1}\right| = \left|\frac{n+1-n}{n+1}\right| = \left|\frac{1}{n+1}\right| = \frac{1}{n+1} \; .$$

Ist nun $\varepsilon > 0$ beliebig vorgegeben, so bestimmen wir wieder auf Grund der Archimedischen Eigenschaft von $\mathbb{R}$ eine natürliche Zahl N mit $\frac{1}{N} < \varepsilon$.
Für alle $n \in \mathbb{N}$, $n \geq N$ ist dann

$$|a_n - 1| = \frac{1}{n+1} < \frac{1}{n} \leq \frac{1}{N} < \varepsilon \; ,$$

also $a_n \in U_\varepsilon(1)$ für alle $n \geq N$, d.h. $\lim\limits_{n\to\infty} a_n = 1$.

(5) Dass die Folge $(a_n) = ((-1)^{n+1}) = (1, -1, 1, -1, \ldots)$ divergiert, zeigt man nach dem gleichen Muster, wie beim Beispiel(2):
Wir nehmen an, (a_n) sei konvergent und $a := \lim\limits_{n\to\infty} a_n$. Zu $\varepsilon := 1$ gibt es ein $N \in \mathbb{N}$, so dass $|a_n - a| < 1$ gilt für alle $n \geq N$.
Für alle $n \geq N$ folgt dann mit der Dreiecksungleichung

$$2 = |a_{n+1} - a_n| = |(a_{n+1} - a) + (a - a_n)| \leq |a_{n+1} - a| + |a_n - a| < 1 + 1 = 2 \; ,$$

also erhalten wir den Widerspruch $2 < 2$.

(6) Welchen falschen Eindruck die ersten Glieder vermitteln können, zeigt der Beispiel

$$(a_n) = \left(\frac{n!}{20n}\right) \; .$$

Hier gilt

$$(a_n) = (0,05;\; 0,005;\; 7,5 \cdot 10^{-4};\; 1,5 \cdot 10^{-4}; \ldots)$$

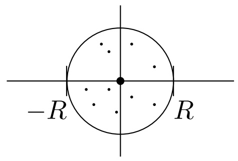
Es ist $a_{20} \approx 2,32 \cdot 10^{-8}$ und die Vermutung, dass (a_n) eine Nullfolge ist, liegt auf der Hand. Berechnet man jedoch a_{60} oder a_{100}, so ergibt sich

$$a_{60} \approx 7217,3; \quad a_{100} \approx 0,375 \cdot 10^{1267} \; .$$

Die Folge $(a_n) = \left(\frac{n!}{n^n}\right)$ ist nicht (nach oben) beschränkt, und damit divergent, denn:

(7) Eine konvergente Folge (a_n) ist *beschränkt*, dabei heißt eine Folge (a_n), $a_n \in \mathbb{K}$, *beschränkt*, wenn es ein $R > 0$ gibt, so dass für alle $n \in \mathbb{N}$ gilt $|a_n| \leq R$.

Alle (a_n) liegen also im komplexen Fall in einer abgeschlossenen Kreisscheibe vom Radius R um das Nullpunkt, im reellen Fall im abgeschlossenen Intervall $[-R, R]$.



Ist nämlich $a = \lim_{n\to\infty} a_n$, so gibt es zu $\varepsilon = 1$ ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $|a_n - a| < 1$. Also folgt
$$|a_n| \leq |a| + |a_n - a| \leq |a| + 1$$
für alle $n \geq 1$. Setzt man $R := max\{|a_1|, |a_2|, \ldots, |a_{N-1}|, |a| + 1\}$, so ist $|a_n| \leq R$ für alle $n \in \mathbb{N}$.

Mit Hilfe dieses Kriteriums erhalten wir einen viel einfacheren Beweis, dass die Folge
$$(a_n) = (n)$$
der natürlichen Zahlen nicht konvergent ist, denn sie ist nicht beschränkt.
Die Folge (f_n) der Fibonacci-Zahlen
$$(1, 1, 2, 3, 5, 8, 17, 21, 34, 55, \ldots)$$
kann damit auch nicht konvergent sein, denn es gilt $f_{n+1} \geq n$ für alle $n \in \mathbb{N}$, die Folge (f_n) ist also nicht (nach oben) beschränkt.

Wir werden in Beispiel (13) sehen, dass sogar für jedes $z \in \mathbb{C}$ die Folge $\left(\frac{z^n}{n!}\right)$ eine Nullfolge ist, insbesondere ist $(\frac{20^n}{n!})$ ein Nullfolge:
Das ist äquivalent mit der Eigenschaft, dass es zu jedem $C > 0$ ein $n \in \mathbb{N}$ gibt, so dass für alle $n \geq N$ gilt
$$\frac{n!}{20^n} > C \ .$$

Eine Fülle von weiteren Beispielen für Nullfolgen erhält man mit Hilfe des folgenden Vergleichssatzes.

(7) Ist (r_n) eine reelle Nullfolge und (a_n) eine beliebige Folge mit $a_n \in \mathbb{K}$. Gibt es dann eine positive reelle Zahl C, so dass für fast alle n gilt $|a_n| \leq Cr_n$, dann ist (a_n) ebenfalls eine Nullfolge.

Der Beweis ist evident: Da (r_n) eine (reelle) Nullfolge ist, gibt es zu jedem $\varepsilon > 0$ ein $N_1 \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geq N_1$ gilt
$$r_n = |r_n| < \frac{\varepsilon}{C} \ .$$
(Beachte: $\frac{\varepsilon}{C}$ ist ebenfalls eine positive reelle Zahl).

Außerdem gilt für fast alle n, etwa für alle $n \geq N_1$, die Ungleichung $|a_n| \leq Cr_n$.
Für alle $n \geq N := max\{N_1, N_2\}$ folgt daher
$$|a_n| \leq Cr_n < C \cdot \frac{\varepsilon}{C} = \varepsilon \ .$$
Hieraus ergibt sich sofort, dass für jedes $k \in \mathbb{N}$ die Folge $\left(\frac{1}{n^k}\right)$ eine Nullfolge ist.

(8) Wir behaupten, dass auch die Folge
$$(a_n) = \left(\frac{1}{1 + \sqrt{n}}\right)$$
eine Nullfolge ist.

Wir gehen dabei heuristisch vor und versuchen zu vorgegebenem $\varepsilon > 0$, ein $N \in \mathbb{N}$ so zu bestimmen, dass für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt
$$|a_n - 0| = |a_n| < \varepsilon \ .$$
Wir gehen dabei versuchsweise von der Ungleichung, die wir beweisen sollen, aus:

(1) $\frac{1}{1+\sqrt{n}} < \varepsilon$.
Hieraus folgt $1 + \sqrt{n} > \frac{1}{\varepsilon}$, oder $\sqrt{n} > \frac{1}{2} - 1$, oder also

(2) $n > (\frac{1}{\varepsilon} - 1)^2$.

Damit scheint das Problem gelöst: Denn wählt man eine natürliche Zahl $N > (\frac{1}{\varepsilon} - 1)^2$, so gilt für alle $n \geq N$ die gewünschte Ungleichung

$$|a_n - 0| < \varepsilon \ .$$

Aber: Wir haben lediglich gesagt, dass aus (1) die Ungleichung (2) folgt.

Der eigentliche Konvergenzbeweis verläuft nun so:

Zum beliebig vorgegebenem $\varepsilon > 0$, wähle man ein $N \in \mathbb{N}$, so dass

$$N > \left(\frac{1}{\varepsilon} - 1\right)^2$$

gilt.

Für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt dann (Monotonie der Wurzel)

$$\sqrt{n} \geq \sqrt{N} > \left|\frac{1}{\varepsilon} - 1\right| \geq \frac{1}{\varepsilon} - 1 \ ,$$

also

$$1 + \sqrt{n} \geq \frac{1}{\varepsilon} \qquad \text{oder} \qquad \frac{1}{1+\sqrt{n}} < \varepsilon \ .$$

(9) Für die Folge

$$(a_n) = \left(\frac{2n^2}{n^2 + 2n + 3}\right)$$

legt die Umformung

$$\frac{2n^2}{n^2 + 2n + 3} = \frac{2}{1 + \frac{2}{n} + \frac{3}{n^2}}$$

die Vermutung nahe, dass $\lim\limits_{n\to\infty} a_n = 2$ gilt.

Wir betrachten zum Beweis

$$\begin{aligned}
|a_n - 2| &= \left|\frac{2n^2}{n^2 + 2n + 3} - 2\right| = \left|\frac{2n^2 - 2n^2 - 4n - 6}{n^2 + 2n + 3}\right| = \\
&= \left|\frac{-4n - 6}{n^2 + 2n + 3}\right| = \frac{4n + 6}{n^2 + 2n + 3} = \\
&= \frac{4n}{n^2 + 2n + 3} + \frac{6}{n^2 + 2n + 3} \leq \frac{4n}{n^2} + \frac{6}{2n} = \frac{7}{n} \ .
\end{aligned}$$

Jetzt wende man den Vergleichssatz mit $C = 7$ und $r_n = \frac{1}{n}$ an.

(10) Für jedes reelle $a > 0$ gilt

$$\boxed{\lim_{n\to\infty} \sqrt[n]{a} = 1}$$

Bemerkung: Die Existenz n-ter Wurzeln aus nicht negativen reellen Zahlen wird hier vorausgesetzt. Vergleiche hierfür den im Anschluss an §1.3 formulierten Existenzsatz.

Wir benutzen: Zu jedem $n \in \mathbb{N}$, $n \geq 2$, und jedem $a \in \mathbb{R}_+$ gibt es genau ein $x \in \mathbb{R}_+$ mit $x^n = a$.
Bez.: $x = \sqrt[n]{a}$.

Weitere Beweise für die Existenz n-ter Wurzel werden wir in dem nächsten Paragraphen über konvergente Folgen bzw. im nächsten Kapitel (Zwischenwertsatz) kennen lernen.

Im Fall $a = 1$ haben wir die konstante Folge $(1, 1, 1, \ldots)$.
Sei daher zunächst $a > 1$.

Dann ist

$$\delta_n := \sqrt[n]{a} - 1 > 0 \text{ und } a = (1 + \delta_n)^n \ .$$

Nach dem binomischen Satz oder der Bernoullischen Ungleichung ist $(1 + \delta_n)^n \geq 1 + n\delta_n$, daher

$$0 \leq \delta_n \leq \frac{a-1}{n} = \frac{C}{n} \qquad (\text{mit } C = a - 1) \ .$$

Also folgt

$$|\sqrt[n]{a} - 1| = |\delta_n| = \delta_n \leq \frac{C}{n} \ .$$

Damit ist $(\sqrt[n]{a} - 1)$ eine Nullfolge, d.h. $\lim\limits_{n\to\infty} \sqrt[n]{a} = 1$.

Mit den Rechenregeln aus dem nächsten Paragraphen führt man den Fall $0 < a < 1$ auf den oben behandelten zurück.

(11) Wie „nivellierend“ die n-te Wurzel ist, zeigt sich am nächsten Beispiel:

$$\boxed{\lim_{n\to\infty} \sqrt[n]{n} = 1}$$

Auf die Vermutung, dass diese Folge den Grenzwert 1 hat, kann man mit Hilfe eines einfachen Taschenrechners kommen. Stellt man für die x-Taste $x = n$, die y-Taste $y = \frac{1}{n}$ ein und betätigt dann die x^y-Taste, so erhält man etwa folgende Werte:

$$\begin{array}{ll} n = 10\,000 & a_n = 1,0009215 \\ n = 100\,000 & a_n = 1,0001151 \\ n = 10^6 & a_n = 1,0000138 \\ n = 10^7 & a_n = 1,0000016 \end{array}$$

Die Folge konvergiert „langsam“ gegen 1, d.h. man braucht sehr große Indizes, um der 1 beliebig nahe zu kommen.
Zum Beweis setzen wir wieder

$$(*) \qquad \sqrt[n]{n} = 1 + \delta_n$$

mit $\delta_n > 0$ und haben $\lim\limits_{n\to\infty} \delta_n = 0$ zu zeigen.

Aus $(*)$ folgt nach dem binomischen Satz

$$n = (1 + \delta_n)^n \geq 1 + \binom{n}{2} \delta_n^2 \qquad \text{oder} \qquad \binom{n}{2} \delta_n^2 \leq n - 1 \ .$$

Nun ist aber

$$\binom{n}{2} = \frac{n(n-1)}{2}$$

und daher ist für $n \geq 2$

$$\delta_n^2 \leq \frac{2}{n} \qquad \text{oder} \qquad \delta_n \leq \sqrt{\frac{2}{n}} \ .$$

Wählt man jetzt zum vorgegebenem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit $N > \frac{2}{\varepsilon^2}$, so ist diese Bedingung äquivalent mit $\frac{2}{N} < \varepsilon^2$. Daher folgt für alle $n \in \mathbb{N}$ mit $n \geq N$

$$|\sqrt[n]{n} - 1| = |\delta_n| = \delta_n \leq \sqrt{\frac{2}{n}} \leq \sqrt{\frac{2}{N}} < \varepsilon \; .$$

(12) Auf den ersten Blick überraschend ist es, dass $(\sqrt{n+1} - \sqrt{n})$ eine Nullfolge ist.
Nun gilt aber für beliebige $a, b \in \mathbb{R}$ die Formel

$$(a-b)(a+b) = a^2 + b^2 \; .$$

Wendet man diese hier an mit

$$a = \sqrt{n+1}, \qquad b = \sqrt{n} \; ,$$

so erhält man

$$a_n = \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \frac{n+1-n}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n+1} + \sqrt{n}} \; .$$

Also ist $|a_n - 0| = |a_n| < \frac{1}{\sqrt{n}}$.

Der Trick $a - b$ in der Form $a - b = \frac{a^2 - b^2}{a+b}$ mit $a \neq -b$ zu schreiben, wird uns noch mehrmals begegnen.

(13) Für alle $z \in \mathbb{C}$ ist $\left(\frac{z^n}{n!}\right)$ eine Nullfolge. Für $z = 0$ ist dies ohnehin klar.

Setze $a_n := \frac{|z|^n}{n!}$, dann ist

$$a_n = \frac{|z|}{n} a_{n-1} \leq \frac{1}{2} a_{n-1} \; ,$$

falls $n \geq N$ und $N > 2|z|$

$$\Longrightarrow a_n \leq C \left(\frac{1}{2}\right)^n \qquad \text{mit} \qquad C := \frac{(2|z|)^N}{N!}$$

Nach dem Vergleichssatz ist damit $\left(\frac{z^n}{n!}\right)$ für alle $z \in \mathbb{C}$ eine Nullfolge.

(14) *Geometrische Folge*

Das Grenzwertverhalten der geometrische Folge (q^n) hängt von q ab. Für $|q| > 1$ ist (q^n) nicht beschränkt (nach ???), also nicht konvergent. Für $q = 1$ ist (q^n) die konstante Folge $(1, 1, 1, \ldots)$, also konvergent. Für $|q| = 1$, aber $q \neq 1$, ist die Folge (q^n) nicht konvergent, wie wir im nächsten Paragraphen sehen werden.

Es gilt jedoch: für $|q| < 1$ ist (q^n) eine Nullfolge. Für $q = 0$ ist dies klar.
Sei also $0 < |q| < 1$, dann betrachten wir $\frac{1}{|q|} = 1 + h$, dabei ist dann $h > 0$ und die Bernoullische Ungleichung oder der Binomische Satz liefert

$$\left(\frac{1}{|q|}\right)^n = (1+h)^n \geq 1 + nh > nh \; ,$$

und damit

$$|q^n - 0| = |q^n| = |q|^n = \frac{1}{(1+h)^n} < \frac{1}{h}\frac{1}{n} = C \cdot \frac{1}{n} \qquad \text{mit } C = \frac{1}{h} \; .$$

(15) Bei der Untersuchung, wie sich eine Investition von K Euro auf das Volkseinkommen auswirkt, waren wir nach n Konsumschritten auf die Erhöhung um

$$a_n = K + qK + q^2K + \ldots + q^nK = \frac{K}{1-q} - K\frac{q^{n+1}}{1-q}$$

Euro gekommen, dabei ist $q (0 < q < 1)$ die *Grenzneigung zum Verbrauch*. Die damalige Vermutung, dass a_n für große n durch $\frac{K}{1-q}$ „hinreichend gut" approximiert wird, können wir jetzt bestätigen, denn wegen $\lim\limits_{n\to\infty} q^{n+1} = 0$, gilt tatsächlich $\lim\limits_{n\to\infty} a_n = \frac{K}{1-q}$. Eine Investition von $K = 10^9$ Euro, ergibt also eine Erhöhung des Volkseinkommen auf $3 \cdot 10^9$ Euro.

In diesem Beispiel hat uns der *Grenzwert* als Näherungswert für alle a_n mit hinreichend großem Folgenindex gedient. Meist ist es aber umgekehrt:
Wir sind nicht so sehr an der Folge (a_n) interessiert, sondern an ihrem Grenzwert. Dabei werden die Folgenglieder mit hinreichend großem Index als Approximationen für den nicht der Größe nach bekannten Grenzwert betrachtet.

(16) Für jedes feste $q \in \mathbb{C}$ mit $|q| < 1$ und jedes feste $p \in \mathbb{N}$ gilt

$$\lim_{n\to\infty} n^p q^n = 0 \ .$$

Man sagt: „(q^n) geht schneller gegen Null als (n^p) gegen Unendlich."

Hier sind offensichtlich zwei gegensätzliche Kräfte am Werk:
Die Folge (n^p) wächst über jede Grenze, d.h. zu jedem $C > 0$ gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $n^p > C$, während (q^n) für $|q| < 1$ eine Nullfolge ist. Die Nullfolge bleibt bei diesem Kräftemessen Sieger.

Zum Beweis können wir $0 < |q| < 1$ annehmen. Dann ist wieder

$$\frac{1}{|q|} = 1 + h \qquad \text{mit } h > 0 \ .$$

Für alle $n \geq p + 1$ ist dann

$$\begin{aligned}
|n^p q^n - 0| = n^p |q|^n &= \frac{n^p}{(1+h)^n} = \frac{n^p}{1 + nh + \ldots + \binom{n}{p+1} h^{p+1} + \ldots + \binom{n}{n} h^n} \\
&\leq \frac{n^p}{\binom{n}{p+1} h^{p+1}} = \frac{(p+1)!}{n\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{p}{n}\right)h^{p+1}} \\
&\leq \frac{(p+1)!}{n\left(1-\frac{1}{p+1}\right)\left(1-\frac{2}{p+1}\right)\cdots\left(1-\frac{p}{p+1}\right)h^{p+1}} \\
&= \frac{C}{n} \ ,
\end{aligned}$$

wobei $C := \frac{(p+1)!}{\left(1-\frac{1}{p+1}\right)\left(1-\frac{2}{p+1}\right)\cdots\left(1-\frac{p}{p+1}\right)h^{p+1}}$ von n unabhängig ist.

Nach dem Vergleichssatz ist also $(n^p q^n)$ eine Nullfolge.

(17) Die Folge $(H_n)_{n\geq 1}$(sog. „Harmonische Reihe") mit

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

ist divergent, denn sie ist nicht (nach oben) beschränkt. Betrachtet man nämlich die Partialsumme

$$H_{2^k} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^{k-1}+1} + \ldots + \frac{1}{2^k}\right),$$

so ist offensichtlich

$$H_{2^k} \geq 1 + \frac{1}{2} + \frac{2}{2^2} + \frac{2^2}{2^3} + \ldots + \frac{2^{k-1}}{2^k} > 1 + \frac{k}{2}\ ,$$

also sind die Partialsummen nicht beschränkt, denn ist $n \geq 2^k$, so ist $H_n \geq H_{2^k}$.

Die Folge (H_n) wächst sehr langsam. Einige Werte seien (mit Hilfe von Maple) angegeben:

n	$(H_n) = \sum\limits_{k=1}^{n} \frac{1}{k}$	$\ln n + \gamma$
10	2.928968253968253968253968253968253968253968	2.8798007578955785446245035447667666
10^2	5.1873775176396202608051176756582531579	5.1823858508896242286424949994511308
10^3	7.4854708605503449126565182043339001765	7.4849709438836699126604864541354950
10^4	9.7876060360443822641784779048516053348	9.7875560368777155966784779088198592
10^5	12.090146129863427947363219363504219500	12.090141129871761280696469363504223
10^6	14.392726722865723631381127493188587676	14.392726222865806964714460818188587
10^7	16.695311365859851815399118939540451884	16.695311315859852648732452272872951
10^8	18.997896413853898324417110394223982841	18.997896408853898332750443727557316
10^9	21.300481502347944016685101848908346966	21.300481501847944016768435182241680
10^{10}	23.603066594891989700785593303592711173	23.603066594841989700786426636926044
10^{11}	25.905651687841035384804409758277075381	25.905651687836035384804418091610408
10^{12}	28.208236780830581068822409462961439588	28.208236780830081068822409546294772
10^{13}	30.510821873824176752840401000145803796	30.510821873824126752840401000979137
10^{14}	32.813406966818177436858392455655168004	32.813406966818172436858392455663501
10^{15}	35.115992059812218620876383910347782211	35.115992059812218120876383910347865
10^{16}	37.418577152806263854894375365032228919	37.418577152806263804894375365032229
10^{17}	39.721162245800309493912366819716593951	39.721162245800309488912366819716593
10^{18}	42.023747338794355173430358274400958167	42.023747338794355172930358274400958
10^{19}	44.326332431788400856998349729085322375	44.326332431788400856948349729085322
10^{20}	46.628917524782446540971341183769686583	46.628917524782446540966341183769686

$H_{250 \cdot 10^6}$ ist noch kleiner als 20, jedoch $H_{500 \cdot 10^6} > 20$. Um über Hundert zu kommen, braucht man schon etwa $15 \cdot 10^{42}$ Summanden.

Ein *Divergenzkriterium*, mit welchem man häufig leicht über die Divergenz einer Folge entscheiden kann, beruht auf folgender Tatsache:

(18) Ist (a_n) eine konvergente Folge mit dem Grenzwert a, dann konvergiert auch jede *Teilfolge* und jede *Umordnung* gegen a. Dabei heißt eine Folge (b_k) *Umordnung* der Folge (a_n), wenn es eine Folge (n_k) von natürlichen Zahlen gibt, in der jede natürliche Zahl genau einmal vorkommt und für die $b_k = a_{n_k}$ gilt.

Eine Folge (b_k) heißt Teilfolge vom (a_n), falls es eine Folge (n_k) natürlicher Zahlen gibt mit $n_1 < n_2 < n_3 < \ldots < n_k < n_{k+1} < \ldots$ und falls $b_k = a_{n_k}$ gilt.

So ist z.B.

$$\left(1, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \frac{1}{9}, \ldots\right)$$

eine Teilfolge von $(a_n) = \left(\frac{1}{n}\right)$, dabei ist $n_k = 2k - 1,\ k \geq 1$ und

$$\left(\frac{1}{3}, \frac{1}{2}, 1, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{7}, \frac{1}{8}, \frac{1}{9}, \ldots\right)$$

ist eine Umordnung von $\left(\frac{1}{n}\right)$.

Ferner konvergiert jede Folge, die aus (a_n) durch *Abänderung endlich vieler Glieder* entsteht, ebenfalls gegen a.

Ist nun (a_n) konvergent mit dem Grenzwert a, dann gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt $a_n \in U_\varepsilon(a)$. Höchstens die Folgenglieder a_n mit $n \in \{1, 2, \ldots, N-1\}$ liegen nicht in $U_\varepsilon(a)$. Für $k \neq k'$ ist aber (in beiden Fällen) $n_k \neq n_{k'}$, daher gibt es höchstens $N-1$ Zahlen n_k mit $n_k \in \{1, \ldots, N-1\}$, also für fast alle k gilt $n_k \notin \{1, \ldots, N-1\}$.
Daher liegt für fast alle k das Folgenglied a_{n_k} in $U_\varepsilon(a)$, d.h. $\lim\limits_{k\to\infty} a_{n_k} = a$.

Ist schließlich (b_k) eine Folge, die aus (a_n) durch Abänderung endlich vieler Glieder entsteht, dann gibt ab einem $N_1 \in \mathbb{N}$ $b_n = a_n$ für alle $n \geq N_1$.

Also ist $|b_n - a| = |a_n - a| < \varepsilon$ für alle $n \geq M := max\{N_1, N\}$.

Wir schließen diesen Paragraphen mit einer Feststellung, die sich auch unmittelbar aus der Tatsache ergibt, dass $\mathbb{Q}$ *dicht* in $\mathbb{R}$ ist (sogar mit dieser Tatsache äquivalent ist).

(19) Jede reelle Zahl x ist Grenzwert einer Folge von rationalen Zahlen, d.h. zu jedem $x \in \mathbb{R}$ gibt es eine Folge $(r_n),\ r \in \mathbb{Q}$, mit $\lim\limits_{n\to\infty} r_n = x$.

Beweis : Man kann $r_n = \frac{[nx]}{n},\ n \in \mathbb{N}$, nehmen. Man kann auch so schließen: Aus Satz 2.3.5 folgt, dass es zu jeder natürlichen Zahl n eine rationale Zahl r_n gibt mit

$$|r_n - x| < \frac{1}{n}\ .$$

Nach dem Vergleichssatz ist also

$$\lim_{n\to\infty} r_n = x\ .$$

□

Zusatz: Man kann sogar erreichen, dass $r_n \leq r_{n+1}$ bzw. $r_{n+1} \leq r_n$ für alle $n \in \mathbb{N}$ gilt.

8 Rechenregel für konvergente Folgen

Die Beispiele zeigen, dass es für eine vorgelegte Folge mühselig sein kann nachzuweisen, dass sie konvergiert, insbesondere, wenn sie durch einen komplizierten Rechenausdruck gegeben ist, oder etwa rekursiv definiert ist. Die folgenden *Rechenregeln* erleichtern die Bestimmung von Grenzwerten, sie stellen Beziehungen her zwischen den *algebraischen Operationen* in $\mathbb{R}$ bzw. $\mathbb{C}$ und der *Grenzwertbildung* bzw. zwischen der *Anordnung* in $\mathbb{R}$ und der *Grenzwertbildung*.
Betrachten wir als Beispiel die reellen Zahlenfolgen

$$(a_n) = \left(1 - \frac{1}{n}\right), \qquad (b_n) = \left(1 + \frac{1}{n}\right), \qquad n \in \mathbb{N} ,$$

dann gilt für *alle* $n \in \mathbb{N}$

$$a_n < b_n \ ,$$

aber

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 1 \ .$$

Es gilt jedoch wenigstens

8.1 Satz (Monotonie des Grenzwertes)

Sind (a_n) und (b_n) konvergente reelle Zahlenfolgen mit $a := \lim\limits_{n\to\infty} a_n$ und $b := \lim\limits_{n\to\infty} b_n$ und gilt für fast alle n

$$a_n \leq b_n \ ,$$

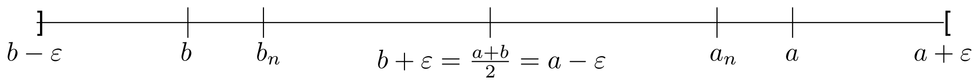
dann gilt auch

$$a = \lim a_n \leq \lim b_n = b \ .$$

1. Beweis durch Widerspruch.

   Wir nehmen $a > b$ an und wählen

   $$\varepsilon := \frac{a-b}{2} > 0 \ ,$$

   Dann gilt für fast alle n

   $$a_n \in U_\varepsilon(a) \text{ und } b_n \in U_\varepsilon(b) \ ,$$

   daher ist für fast alle n

   $$b_n < b + \frac{a-b}{2} = \frac{a+b}{2} = a - \frac{a-b}{2} < a_n$$

   ein Widerspruch zur Voraussetzung

   $$a_n \leq b_n, \text{ für fast alle } n \ .$$

   $\square$

2. Beweis mit dem Fundamentallemma der Analysis, das ja (vgl. §1.2.4) besagt: Ist $x \in \mathbb{R}$ und $x \leq \varepsilon$ für *jedes* $\varepsilon > 0$, dann ist $x \leq 0$.

Nach Voraussetzung gibt es ein $n \in \mathbb{N}$, so dass gleichzeitig gilt (für jedes $\varepsilon > 0$)

$$a - \frac{\varepsilon}{2} < a_n, \quad a_n \leq b_n, \quad b_n < b + \frac{\varepsilon}{2}\ ,$$

daher gilt auch

$$a - \frac{\varepsilon}{2} < b + \frac{\varepsilon}{2} \qquad \text{oder} \qquad a - b < \varepsilon$$

für *jedes* $\varepsilon > 0$, nach dem Fundamentallemma ist daher $a \leq b$. □

8.2 Folgerung

Gilt für eine mit dem Grenzwert a konvergente Folge (a_n) reeller Zahlen die Ungleichung ($\alpha, \beta \in \mathbb{R}$)

$$\alpha \leq a_n \leq \beta$$

für fast alle n, dann gilt auch

$$\alpha \leq a \leq \beta\ .$$

8.3 Satz (Sandwich-Lemma)

Sind (a_n) und (b_n) konvergente reelle Zahlenfolgen mit $a = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$, und ist (x_n) eine reelle Zahlenfolge mit $a_n \leq x_n \leq b_n$, für fast alle n, dann ist auch (x_n) konvergent, und es gilt $\lim\limits_{n\to\infty} x_n = a$.

Beweis : Ist $\varepsilon > 0$ beliebig vorgegeben, dann liegen in $U_\varepsilon(a)$ fast alle Folgenglieder (a_n) und (b_n), wegen

$$a_n \leq x_n \leq b_n$$

liegen auch fast alle x_n in $U_\varepsilon(a)$, d.h. $\lim\limits_{n\to\infty} x_n = a$.

□

8.4 Beispiel

Sind $a_1, \ldots, a_k,\ (k \in \mathbb{N}$ fest), nicht negative reelle Zahlen, dann ist die Folge (x_n) mit

$$x_n = \sqrt[n]{a_1^n + a_2^n + \ldots + a_k^n}$$

konvergent, und es gilt

$$\boxed{\lim_{n\to\infty} x_n = max\{a_1, \ldots, a_k\}}$$

Beweis : Ist $m := max\{a_1, \ldots, a_k\}$, dann gilt

$$m = \sqrt[n]{m^n} \leq \sqrt[n]{a_1^n + \ldots + a_k^n} \leq \sqrt[n]{km^n} = \sqrt[n]{k}m \, .$$

Wegen $\lim\limits_{n\to\infty} \sqrt[n]{k} = 1$ (beachte k ist fest!) folgt nach dem Sandwich-Lemma

$$m \leq \lim_{n\to\infty} \sqrt[n]{a_1^n + \ldots + a_k^n} \leq m \, , \qquad \text{also} \qquad \lim_{n\to\infty} \sqrt[n]{a_1^n + \ldots + a_k^n} = m \, .$$

□

Betrachtet man die Folge

$$(c_n) = \left(\frac{n+1}{n^2}\right) = \left(\frac{1}{n} + \frac{1}{n^2}\right) = (a_n + b_n) \, ,$$

mit $a_n = \frac{1}{n}$, $b_n = \frac{1}{n^2}$, dann lässt es nahe zu vermuten, dass (c_n) eine Nullfolge ist, da (a_n) und (b_n) Nullfolgen sind.

Die folgenden Permanenzeigenschaften von $\lim$ vereinfachen die Bestimmung von Grenzwerten wesentlich. Wir betrachten dabei i.A. Folgen in $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$).

8.5 Satz (Permanenzeigenschaften von lim)

Die Zahlenfolgen (a_n) und (b_n), $a_n, b_n \in \mathbb{K}$, seien konvergent und $a := \lim\limits_{n\to\infty} a_n$ und $b := \lim\limits_{n\to\infty} b_n$. Dann gilt:

(1) Die (Summen-)Folge $(a_n + b_n)$ ist konvergent, und es gilt

$$\lim_{n\to\infty} (a_n + b_n) = a + b = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n \, .$$

(2) Die (Produkt-)Folge $(a_n b_n)$ ist konvergent, und es gilt

$$\lim_{n\to\infty} (a_n b_n) = ab = \lim_{n\to\infty} a_n \lim_{n\to\infty} b_n \, .$$

Speziell ist für $\lambda \in \mathbb{K}$ die Folge (λa_n) konvergent mit

$$\lim_{n\to\infty} (\lambda a_n) = \lambda a = \lambda \lim_{n\to\infty} a_n \, .$$

(3) Ist $b \neq 0$, dann gibt es ein $n_1 \in \mathbb{N}$, so dass $b_n \neq 0$ für alle $n \geq n_1$ gilt:
Die (Quotienten-)Folge $\left(\frac{a_n}{b_n}\right)_{n\geq n_1}$ ist konvergent, und es gilt

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b} = \frac{\lim\limits_{n\to\infty} a_n}{\lim\limits_{n\to\infty} b_n} \, .$$

(4) Die (reelle) Zahlenfolge $(|a_n|)$ ist konvergent, und es gilt

$$\lim_{n\to\infty} |a_n| = |a| = |\lim_{n\to\infty} a_n| \, .$$

(5) Die Folge $(\bar{a}_n)$ ist konvergent, und es gilt

$$\lim_{n\to\infty} \bar{a}_n = \bar{a} = \overline{\lim_{n\to\infty} a_n} \; .$$

(6) Für $a_n \in \mathbb{C}$ konvergieren die Folgen $(Re(a_n))$ und $(Im(a_n))$, und es gilt

$$\lim_{n\to\infty} Re(a_n) = Re(a) = Re(\lim_{n\to\infty} a_n) \quad \text{und}$$
$$\lim_{n\to\infty} Im(a_n) = Im(a) = Im(\lim_{n\to\infty} a_n) \; .$$

Umgekehrt folgt aus der Konvergenz von $(Re(a_n))$ und $(Im(a_n))$ die Konvergenz von (a_n), wobei

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} Re(a_n) + i \lim_{n\to\infty} Im(a_n) \quad \text{gilt.}$$

Beim Beweis werden folgende Prinzipien verwendet:

(a) Ist (a_n) konvergent mit dem Grenzwert a und ist (b_n) konvergent mit dem Grenzwert b, dann gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ (gleichzeitig) gilt

$$a_n \in U_\varepsilon(a) \text{ und } b_n \in U_\varepsilon(b) \; .$$

(b) Ist (x_n) eine Nullfolge und (y_n) eine beschränkte Folge, dann ist auch die Produktfolge $(x_n y_n)$ eine Nullfolge.

Beweis zu (1): Die Summenregel ist klar:
Zu jedem $\varepsilon > 0$ gibt es (nach (a)) ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt

$$|a_n - a| < \frac{\varepsilon}{2} \text{ und } |b_n - b| < \frac{\varepsilon}{2} \; .$$

Nach der Dreiecksungleichung erhält man

$$|a_n + b_n - (a+b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} < \varepsilon$$

für alle $n \geq N$.

□

Beweis zu (2): Der Beweis für das Produkt ist etwas trickreicher. Man benutzt die Zerlegung

$$a_n b_n - ab = a_n b_n - ab_n + ab_n - ab = (a_n - a)b_n + a(b_n - b)$$

und die Voraussetzung, dass a_n gegen a konvergiert, also $(a_n - a)$ eine Nullfolge ist, dass (b_n) gegen b konvergiert, also $(b_n - b)$ eine Nullfolge ist und dass (b_n) als konvergente Folge beschränkt ist und verwendet dann (b).

Der geneigte Leser (die geneigte Leserin) möge die Einzelheiten ausführen.
Den Spezialfall erhält man, in dem man für (b_n) die konstante Folge $(b_n) = (\lambda)$ wählt. Dieser Spezialfall lässt sich natürlich auch einfach direkt beweisen.

□

Beweis zu (3): Ist $b \neq 0$, dann ist $\varepsilon := \frac{|b|}{2} > 0$ und in $U_\varepsilon(b)$ liegen fast alle Glieder der Folge (b_n), d.h. es gibt ein $n_1 \in \mathbb{N}$, so dass für alle $n \geq n_1$ gilt

$$|b_n - b| < \frac{|b|}{2} \ .$$

Aus der Dreiecksungleichung für Abschätzungen nach unter folgt daher

$$|b| - |b_n| \leq |b_n - b| < \frac{|b|}{2} \quad \text{und damit} \quad |b_n| \geq \frac{|b|}{2} > 0$$

für alle $n \geq n_1$.
Nun folgt nach Standardschlüssen die Konvergenz von $(\frac{1}{b_n})_{n \geq n_1}$ gegen $\frac{1}{b}$ aus

$$\left| \frac{1}{b_n} - \frac{1}{b} \right| = \frac{|b_n - b|}{|b_n|\,|b|} \leq \frac{2}{|b|^2} |b_n - b|$$

und die Konvergenz von $(\frac{a_n}{b_n})_{n \geq n_1}$ folgt wegen $\frac{a_n}{b_n} = a_n \cdot \frac{1}{b_n}$ aus (2) und dem ersten Teil von (3).

□

Beweis zu (4) folgt unmittelbar aus $||a_n| - |a|| \leq |a_n - a|$.

Beweis zu (5): Wegen $|z| = |\bar{z}|$ für $z \in \mathbb{C}$ und $\overline{a_n - a} = \bar{a}_n - \bar{a}$ ist (5) ebenfalls klar:

$$|\bar{a}_n - \bar{a}| = |\overline{a_n - a}| = |a_n - a| \ .$$

□

Beweis zu (6): Ist die komplexe Zahlenfolge (a_n) konvergent mit dem Grenzwert a, dann konvergiert nach (5) die Folge $(\bar{a}_n)$ und zwar mit dem Grenzwert $\bar{a}$, daher konvergiert auch die Folge

$$\left(\frac{1}{2}(a_n + \bar{a}_n) \right) = (Re(a_n))$$

und zwar gegen $\frac{1}{2}(a + \bar{a}) = Re(a)$.

Ebenso kann man für den Imaginärteil schließen:

Die Konvergenz von $(Re(a_n))$ gegen $Re(a)$,
von $(Im(a_n))$ gegen $Im(a)$

ergibt sich auch aus den elementaren Abschätzungen

$$|Re(a_n - a)| \leq |a_n - a| \qquad \text{bzw.} \qquad |Im(a_n - a)| \leq |a_n - a| \ .$$

Die Umkehrung ergibt sich etwa aus der Ungleichung

$$|z - a| \leq |Re(z - a)| + |Im(z - a)| \ .$$

□

8.6 Bemerkungen und Beispiele

(a) Für die Folge (a_n) mit

$$a_n = \frac{18n^3 + 2n^2 - 4711}{3n^3 - 17n^2 - 11n - 19}$$

ergibt sich durch Kombination der Rechenregeln

$$(a_n) = \left(\frac{18 + 2 \cdot \frac{1}{n} - \frac{4711}{n^3}}{3 - 17 \cdot \frac{1}{n} - 11 \cdot \frac{1}{n^2} - \frac{19}{n^3}} \right) ,$$

also

$$\lim_{n\to\infty} a_n = \frac{18}{3} = 6 .$$

(b) Was würden Sie als Tutor einer Übungsgruppe zur Analysis 1 zu folgendem Argument eines Übungsgruppenteilnehmers sagen?
Die Folge (a_n) mit

$$(a_n) = \left(\frac{1+2+3+\ldots+n}{n^2} \right)$$

ist eine Nullfolge, denn es gilt

$$(a_n) = \left(\frac{1}{n^2} + \frac{2}{n^2} + \frac{3}{n^2} + \ldots + \frac{1}{n} \right)$$

und jeder Summand ist eine Nullfolge.

Beachte: Die Summandenregel (2) lässt sich sofort auf eine *endliche* Anzahl von Summanden ausdehnen, in diesem Beispiel wächst aber die Anzahl der Summanden mit n gegen Unendlich.

Korrekt kann man so schließen: Nach der Gaußschen Formel ist

$$1+2+3+\ldots+n = \frac{n(n+1)}{2}, \text{ also}$$

$$(a_n) = \left(\frac{1+2+\ldots+n}{n^2} \right) = \left(\frac{\frac{n(n+1)}{2}}{n^2} \right) = \left(\frac{n+1}{2n} \right) = \left(\frac{1}{2} \left(1 + \frac{1}{n} \right) \right) ,$$

daher $\lim\limits_{n\to\infty} a_n = \frac{1}{2}$.

(c) Der Fall $0 < a < 1$ in Beispiel (10) aus §7.8 lässt sich ganz einfach behandeln. Ist nämlich $0 < a < 1$, dann ist $a^{-1} > 1$ und nach $2.5(3)$ ist

$$\lim_{n\to\infty} \sqrt[n]{a} = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{a^{-1}}} = \frac{1}{1} = 1 .$$

(d) Aus $\lim\limits_{n\to\infty} a_n = a$, folgt durch mehrfache Anwendung von 2.5(2)

$$\lim_{n\to\infty} a_n^k = a^k \text{ für } k \in \mathbb{N} .$$

Ist daher $P : \mathbb{K} \to \mathbb{K}$ definiert durch

$$P(x) := a_0 + a_1 x + \ldots + a_m x^m \qquad (a_j \in \mathbb{K}) ,$$

(P ist ein Polynom) dann folgt (mehrfache Anwendung 2.5(1) und 2.5(2))

$$\lim_{n\to\infty} P(a_n) = P(a) \ .$$

Das bedeutet, dass *Polynome* (in ihrem Definitionsbereich $\mathbb{R}$ oder $\mathbb{C}$) *stetige Funktionen* sind (vgl. Kap.III).

(e) Ist $s = s(\mathbb{K}) = Abb(\mathbb{N};\mathbb{K})$ die Menge aller Folgen in $\mathbb{K}$ und $c = c(\mathbb{K})$ die Menge aller *konvergenter Folgen* in $\mathbb{K}$, dann kann man die Eigenschaften von Satz 2.5 auch so zusammenfassen:

$c(\mathbb{K})$ ist ein Untervektorraum von $s = s(\mathbb{K})$ und Abbildung

$$\begin{aligned} \lim: \ c(\mathbb{K}) &\to \mathbb{K} \ , \\ (a_n) &\mapsto \lim_{n\to\infty} a_n \ . \end{aligned}$$

ist $\mathbb{K}$-linear.

Was ist der Kern von lim?
Antwort: $Kern(\lim) = c_0(\mathbb{K}) := \{(a_n) \in s;\ (a_n)$ ist Nullfolge$\}$. Der Kern besteht also gerade aus allen Nullfolgen. $s(\mathbb{K})$ ist mit der Multiplikation

$$(a_n)\cdot(b_n) = (a_n b_n)$$

sogar ein *Ring*, also insgesamt eine $\mathbb{K}$-*Algebra*.

Allerdings hat $s(\mathbb{K})$ Nullteiler, wie etwa das einfache Beispiel

$$(a_n) = (1,0,0,0,\ldots) \ , \quad (b_n) = (0,1,0,0,\ldots)$$

zeigt.

(f) Um die Konvergenz einer komplexen Zahlenfolge zu untersuchen, braucht man sie nicht unbedingt in Real- und Imaginärteil zu zerlegen. Die Folge $\left(\frac{i^n}{1+in}\right)$ ist eine Nullfolge in $\mathbb{C}$. Eine elementare Umformung ergibt

$$\frac{i^n}{1+in} = \frac{1}{n}\frac{i^n}{i+\frac{1}{n}} \ .$$

Die Abschätzung (für $n \geq 2$)

$$\left|i+\frac{1}{n}\right| \geq |i| - \frac{1}{n} = 1 - \frac{1}{n} \geq 1 - \frac{1}{2} = \frac{1}{2}$$

zeigt, dass $\left(\left|\frac{i^n}{i+\frac{1}{n}}\right|\right)$ beschränkt ist:

$$\left|\frac{i^n}{i+\frac{1}{n}}\right| = \frac{|i^n|}{|i+\frac{1}{n}|} \leq \frac{1}{1-\frac{1}{n}} \leq \frac{1}{\frac{1}{2}} = 2$$

für $n \geq 2$, also ist $\left(\frac{i^n}{1+in}\right)$ eine Nullfolge. □

(g) So suggestiv die Schreibweisen, wie z.B.

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n \ ,$$

auch sein mögen, so muss man doch die typische Struktur beachten.
Eigentlich muss man sie von recht nach links lesen: Wenn $\lim\limits_{n\to\infty} a_n$ und $\lim\limits_{n\to\infty} b_n$ existieren, dann

existiert auch $\lim_{n\to\infty}(a_n + b_n)$, und es gilt $\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$.

Aus der Existenz von $\lim_{n\to\infty}(a_n + b_n)$ kann man u.a. nicht schließen, dass die Summenformel gilt, denn die Folgen (a_n) und (b_n) brauchen gar nicht zu konvergieren, wenn die Summenfolge $(a_n + b_n)$ konvergiert. Ein einfaches Beispiel ist

$$(a_n) = (1, -1, 1, -1, \ldots) \qquad \text{und} \qquad (b_n) = (-1, 1, -1, 1, \ldots)\ .$$

Beide Folgen sind divergent, die Summenfolge $(a_n + b_n)$ ist die Konstante Folge $(0, 0, 0, \ldots)$, also konvergent.

9 Konvergenzkriterien

Wir beschäftigen uns in diesem Anschnitt mit Kriterien für die Konvergenz von Folgen, die *ohne Kenntnis des Grenzwertes* erlauben, auf die Konvergenz bzw. Divergenz einer Folge zu schließen. Wir beschäftigen uns zunächst mit *reellen Folgen*. Die entsprechenden Kriterien beruhen auf dem Vollständigkeitsaxiom für $\mathbb{R}$.

Zunächst erinnern wir an ein

9.1 Divergenzkriterium

Ist eine Folge (a_n) reeller oder komplexer Zahlen nicht beschränkt, so ist sie nicht konvergent. So ist also z.B. die Folge (n) der natürlichen Zahlen oder die Folge (f_n) der Fibonacci-Zahlen (vgl. §7.2(9)) nicht konvergent.

9.2 Definition (Monotonie)

Eine *reelle* Zahlenfolge (a_n) heißt

$$\left.\begin{array}{r}\text{monoton wachsend}\\ \text{monoton fallend}\\ \text{streng monoton wachsend}\\ \text{streng monoton fallend}\end{array}\right\}\ \text{falls für alle } n \text{ gilt}\ \left\{\begin{array}{l}a_n \leq a_{n+1}\\ a_n \geq a_{n+1}\\ a_n < a_{n+1}\\ a_n > a_{n+1}\end{array}\right.$$

Für *monotone* Folgen gilt der folgende fundamentale Konvergenzsatz:

9.3 Monotonieprinzip

Eine *monotone* reelle Folge konvergiert genau dann, wenn sie beschränkt ist. Sie konvergiert gegen das Supremum ihrer Wertemenge, wenn sie nach oben beschränkt und gegen das Infimum ihrer Wertemenge, wenn sie nach unter beschränkt ist.

Beweis : Sei (a_k) die gegebene Folge und etwa (a_k) monoton wachsend und $M = \{a_k;\ k \in \mathbb{N}\}$ ihre Wertemenge, dann ist $M \neq 0$.

Da (a_k) nach oben beschränkt ist, existiert $a = sup\, M$ nach dem Vollständigkeitsaxiom. Ist nun $\varepsilon > 0$ beliebig vorgegeben, dann ist $a - \varepsilon$ keine obere Schranke von M mehr, es gibt daher ein $N \in \mathbb{N}$, so dass $a - \varepsilon < a_N$ gilt.
Wegen der Monotonie der (a_k) gilt dann für alle $n \in \mathbb{N}$ mit $n \geq N$ $a_N \leq a_n$, also auch

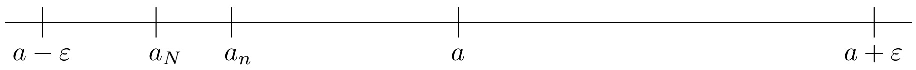
$$a - \varepsilon < a_n \leq a < a + \varepsilon\ .$$

Daher gilt für jede ε-Umgebung $U_\varepsilon(a)$

$$a_n \in U_\varepsilon(a)$$

für alle $n \geq N$, d.h.

$$a = \lim_{n\to\infty} a_n \ .$$



□

Für monoton fallende, nach unter beschränkte Folgen, schließt man völlig analog oder verwendet das *Spiegelungsprinzip* (vgl. §1.2.2).

Wenn wir Beispiele betrachten, formulieren wir ein eng mit dem Monotonieprinzip verwandtes Prinzip, das *Intervallschachtelungsprinzip*:

9.4 Definition und Satz

Gegeben sei eine Folge von Intervallen $([a_n, b_n])_{n\in\mathbb{N}}$, $a_n \leq b_n$ für alle $n \in \mathbb{N}$ mit den Eigenschaften:

(a) $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ für alle $n \in \mathbb{N}$;

(b) $\lim\limits_{n\to\infty} (b_n - a_n) = 0$.

Eine solche Folge nennen wir eine *Intervallschachtelung*.
Dann existiert genau eine reelle Zahl x, die in allen Intervallen liegt: $x \in [a_n, b_n]$ für alle $n \in \mathbb{N}$, m.a.W.:

$$\bigcap_{n\geq 1} [a_n, b_n] = \{x\} \ .$$

Man sagt: Die Intervallschachtelung erfasst die Zahl x.

Ferner gilt

$$x = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n \ .$$

Der *Beweis* versteht sich fast von selbst. Aus der Voraussetzung $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ für alle $n \in \mathbb{N}$ ergibt sich, dass die Folge (a_n) monoton wächst und die Folge (b_n) monoton fällt.

Andererseits sind beide Folgen (a_n) und (b_n) beschränkt, da alle Folgenglieder im Intervall $[a_1, b_1]$ liegen. Also existieren nach dem *Monotonieprinzip*

$$a := \lim_{n\to\infty} a_n \qquad \text{und} \qquad b := \lim_{n\to\infty} b_n \ .$$

Nach Voraussetzung ist die Folge der Intervallängen $(b_n - a_n)$ eine Nullfolge:

$$0 = \lim_{n\to\infty} (a_n - b_n) \ .$$

Andererseits ist nach den *Rechenregeln für konvergente Folgen*

$$\lim_{n\to\infty} (b_n - a_n) = \lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = b - a \ ,$$

also $a = b$.

Für alle $n \in \mathbb{N}$ gilt

$$a_n \leq sup\{a_k, k \in \mathbb{N}\} = a = b = sup\{b_k; k \in \mathbb{N}\} \leq b_k \ ,$$

also ist für alle $n \in \mathbb{N}$

$$a \in [a_n, b_n] \ .$$

Ist $c \in \mathbb{R}$ eine weitere Zahl mit $a_n \leq c \leq b_n$, für alle $n \in \mathbb{N}$, so folgt nach dem *Sandwich-Lemma* $a \leq c \leq a$, also $c = a$.

9.5 Beispiele und Bemerkungen zum Monotonie- und zum Intervallschachtelungsprinzip

(1) „*Multiplikation statt Division*"

Ist $a > 0$, so konvergiert die rekursive Folge (x_n) mit beliebigem Startwert

$$x_0 \in \left]0,\ \frac{1}{a}\right[\qquad \text{und} \qquad x_{n+1} := 2x_n - ax_n^2$$

für $n \geq 1$ monoton wachsend gegen $\frac{1}{a}$. Die Konvergenz ist sogar quadratisch.

Auf diese Folge kommt man, wenn man die Gleichung $(a > 0)$

$$(*) \qquad ax = 1$$

zur Bestimmung von a^{-1} äquivalent umschreibt:

$$(**) \qquad x = 2x - ax^2 \ .$$

x ist die von Null verschiedene Lösung der Gleichung $(**)$.
Setzt man $p(x) := 2x - ax^2 = x(2 - ax)$, so haben wir die *Fixpunktgleichung* $x = p(x)$ zu lösen.

Durch die Abbildung 9 wird die rekursive Definition obiger Folge motiviert. Man wählt einen beliebigen Grenzwert $x_0 \in]0, \frac{1}{a}[$. Mit Induktion zeigt man leicht, dass (x_n) (streng) monoton wächst und $x_n \leq \frac{1}{a}$ für alle $n \in \mathbb{N}$ gilt.

(x_n) ist also nach dem Monotonieprinzip konvergent und der Grenzwert ist die positive Lösung der Gleichung $(**)$, also $\frac{1}{a}$.
Aus der Beziehung

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

ist ersichtlich, dass die Folge (x_n) *quadratisch* gegen $\frac{1}{a}$ konvergiert.

Für $a = 2,\ x_0 = 0,25$ liefert ein primitiver Taschenrechner bereits

$$\begin{array}{lll} x_1 = 0,375, & x_2 = 0,46875, & x_3 = 0,498046873 \ , \\ x_4 = 0,499992371, & x_5 = 0,5, & x_6 = 0,5 \ , \end{array}$$

also auch $x_n = 0,5$ für alle $n > 6$.

Abbildung 9: ???

Zusatz: Man kann den Grenzwert x_0 sogar im größeren Intervall $]0, \frac{2}{a}[$ wählen, dann konvergiert die Folge (x_n) ab $n = 1$ monoton wachsend gegen $\frac{1}{a}$.

(2) Wir wissen, dass die rationalen Zahlen $\mathbb{Q}$ dicht in $\mathbb{R}$ liegen (vgl. §2.3.5), was man auch in der Sprache der Folgen so ausdrücken kann:

Jede reelle Zahl ist der Grenzwert einer Folge von rationalen Zahlen (vgl. §7.8(19)) Durch eine kleine Variation des Beweisprinzips von 2.3.5 erhält man zu einer beliebigen reellen Zahl x und beliebigem $n \in \mathbb{N}$ rationale Zahlen p_n und q_n, so dass die folgende Ungleichungskette erfüllt ist

$$x - \frac{1}{n} < p_n < x - \frac{1}{n+1} < x < x + \frac{1}{n+1} < q_n < x + \frac{1}{n} \ .$$

Offensichtlich ist $p_n < x < q_n$, (x_n) monoton wachsend und (q_n) monoton fallend, außerdem ist

$$q_n - p_n < \frac{2}{n} \ ,$$

also ist $([p_n, q_n])_{n\in\mathbb{N}}$ eine Intervallschachtelung mit

$$x = \lim_{n\to\infty} p_n = \lim_{n\to\infty} q_n \ .$$

(3) *Die Eulersche Zahl e*

Wir betrachten die Folgen (E_n) und $(\tilde{E}_n)$ mit

$$E_n = \sum_{k=0}^{n} \frac{1}{k!}, \ \ n \in \mathbb{N}_0 \ , \qquad \text{und} \qquad \tilde{E}_n = E_n + \frac{1}{n \cdot n!}, \ \ n \in \mathbb{N} \ ,$$

und behaupten, dass durch $([E_n, E_n^*])_{n\in\mathbb{N}}$ eine Intervallschachtelung definiert wird. Die Folge (E_n) ist offensichtlich (streng) monoton wachsend:

$$E_{n+1} = E_n + \frac{1}{(n+1)!}, \ \text{ also } \ E_{n+1} > E_n \ .$$

Aus

$$E_{n+1} - E_n = \frac{1}{(n+1)!} < \frac{1}{n \cdot n!} - \frac{1}{(n+1)(n+1)!}$$

folgt

$$\tilde{E}_{n+1} = E_{n+1} + \frac{1}{(n+1)(n+1)!} < E_n + \frac{1}{n \cdot n!} = \tilde{E}_n \ ,$$

d.h. die Folge $(\tilde{E}_n)$ ist streng monoton fallend und es gelte die Ungleichung

$$2 = E_1 < 2,5 = E_2 < \ldots < E_n < \tilde{E}_n < \ldots < 2,75 = \tilde{E}_2 < 3 = \tilde{E}_1 \ .$$

Ferner ist

$$\tilde{E}_n - E_n = \frac{1}{n \cdot n!} \ .$$

Nach dem Intervallschachtelungsprinzip gibt es also genau eine reelle Zahl –wir nennen sie nach dem Vorbild von L.Euler (1731) die *Eulersche Zahl* e– mit $e \in [E_n, \tilde{E}_n]$ für alle n.

e ist die Basis für den natürlichen Logarithmus.

Die Eulersche Zahl e lässt sich auch durch eine andere Intervallschachtelung erfassen.
Dazu betrachten wir die beiden Folgen (e_n) und $(\tilde{e}_n)$ mit

$$e_n = \left(1 + \frac{1}{n}\right)^n , \ n \in \mathbb{N} \qquad \text{und} \qquad \tilde{e}_n = \left(1 + \frac{1}{n}\right)^{n+1} , \ n \in \mathbb{N}$$

und behaupten, dass $([e_n, \tilde{e}_n])_{n \in \mathbb{N}}$ ebenfalls eine Intervallschachtelung definiert, die dann eine reelle Zahl e' erfasst.

Wir werden zeigen, dass $e' = e$ gilt.

Die Folge (e_n) taucht bei *Wachstums-* und *Zerfallsprozessen* (z.B. stetige Verzinsung) auf. Das Monotonieverhalten und die Beschränktheit der beiden Folgen ist nicht offensichtlich.
Wir betrachten zunächst

$$\tilde{e}_n = e_n \left(1 + \frac{1}{n}\right) > e_n \ \ (n \in \mathbb{N}) \ .$$

Wenn wir wüssten, dass beide Folgen konvergent sind, folgte hieraus also

$$\lim_{n \to \infty} \tilde{e}_n = \lim_{n \to \infty} e_n \ .$$

Wir zeigen, dass (e_n) (streng) monoton wächst, $(\tilde{e}_n)$ (streng) monoton fällt.

Für $n \geq 2$ ist nämlich

$$\begin{aligned}
\frac{e_n}{e_{n-1}} &= \frac{\left(1+\frac{1}{n}\right)^n}{\left(1+\frac{1}{1+\frac{1}{n-1}}\right)} \\
&= \left(\frac{n+1}{n}\right)^n \cdot \left(\frac{n-1}{n}\right)^{n-1} \\
&= \left(\frac{n+1}{n}\right)^n \cdot \left(\frac{n-1}{n}\right)^{n} \cdot \frac{n}{n-1} \\
&= \left(\frac{n^2-1}{n^2}\right)^n \cdot \frac{n}{n-1} \\
&= \left(1-\frac{1}{n^2}\right)^n \cdot \frac{n}{n-1} \\
&> \left(1-\frac{n}{n^2}\right) \cdot \frac{n}{n-1} \qquad \text{(nach Bernoulli)} \\
&= \frac{n^2-n}{n^2} \cdot \frac{n}{n-1} = 1\ ,
\end{aligned}$$

also $e_n > e_{n-1}$ für alle $n \geq 2$ oder $e_{n+1} > e_n$ für alle $n \in \mathbb{N}$.

Andererseits ist

$$\begin{aligned}
\frac{\tilde{e}_{n-1}}{\tilde{e}_n} &= \frac{\left(1+\frac{1}{n-1}\right)^n}{\left(1+\frac{1}{n}\right)^{n+1}} \\
&= \left(\frac{n}{n-1}\right)^n \cdot \left(\frac{n}{n+1}\right)^{n+1} \\
&= \left(\frac{n}{n-1}\right)^n \cdot \left(\frac{n}{n+1}\right)^{n} \cdot \frac{n}{n+1} \\
&= \left(\frac{n^2}{n^2-1}\right)^n \cdot \frac{n}{n+1} \\
&= \left(1+\frac{1}{n^2-1}\right)^n \cdot \frac{n}{n+1} \\
&> \left(1+\frac{n}{n^2-1}\right) \cdot \frac{n}{n+1} \qquad \text{(nach Bernoulli)} \\
&> \left(1+\frac{1}{n}\right) \cdot \frac{n}{n+1} = 1\ , \qquad \left(\text{beachte } \frac{n}{n^2-1} > \frac{1}{n}\right)
\end{aligned}$$

also ist $\tilde{e}_{n-1} > \tilde{e}_n$ für alle $n \geq 2$ oder $\tilde{e}_{n+1} < \tilde{e}_n$ für alle $n \in \mathbb{N}$, d.h. die Folge $(\tilde{e}_n)$ fällt streng monoton.

Es gilt also insgesamt

$$2 = e_1 < 2{,}25 = e_2 < \ldots < e_n < \tilde{e}_n < \ldots < 3{,}375 = \tilde{e}_2 < 4 = \tilde{e}_1\ .$$

Ferner ist für alle $n \in \mathbb{N}$

$$\tilde{e}_n - e_n = \frac{1}{n} e \leq \frac{1}{n} 4\ ,$$

also gilt $\lim\limits_{n\to\infty} (\tilde{e}_n - e_n) = 0$, die Voraussetzungen des Intervallschachtelungsprinzips sind also erfüllt. Es gibt also eine eindeutig bestimmte reelle Zahle e' mit

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

Wir zeigen nun, dass $e = e'$ gelten muss. Dazu zeigen wir zunächst, dass für alle $n \in \mathbb{N}$ $e_n \le E_n$ gilt.

Nach dem binomischen Satz ist nämlich

$$\begin{aligned}
e_n &= \left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \frac{1}{n^k} \\
&= 1 + \sum_{k=1}^{n} \frac{n}{n} \frac{n-1}{n} \cdots \frac{n-k+1}{n} \frac{1}{k!} \\
&= 1 + \sum_{k=1}^{n} \frac{n}{n} \left(1-\frac{1}{n}\right) \cdot (1-\frac{k+1}{n}) \frac{1}{k!} \\
&\le 1 + \sum_{k=1}^{n} \frac{1}{k!} = \sum_{k=0}^{n} \frac{1}{k!} = E_n \; .
\end{aligned}$$

Durch Grenzübergang $n \to \infty$ folgt hieraus

$$e' = \lim_{n\to\infty} e_n \le \lim_{n\to\infty} E_n = e \; ,$$

also $e' \le e$.

Wir zeigen nun, dass auch $e' \ge e$, insgesamt also $e = e'$ gilt.

Dazu sei $m \in \mathbb{N}$ zunächst fest gewählt und $n \in \mathbb{N}$, $n \ge m$. Dann folgt mit einer völlig analogen Schlussweise

$$\begin{aligned}
e_n &= \left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \\
&= 1 + \sum_{k=1}^{n} \frac{n}{n} \frac{n-1}{n} \cdots \frac{n-k+1}{n} \frac{1}{k!} \\
&\ge 1 + \sum_{k=1}^{m} \frac{n}{n} \left(1-\frac{1}{n}\right) \cdots \left(1-\frac{k+1}{n}\right) \frac{1}{k!}
\end{aligned}$$

und hieraus nach den Rechenregeln für konvergente Folgen (beachte m ist fest)

$$e' = \lim_{n\to\infty} e_n \ge \sum_{k=0}^{m} \frac{1}{k!} = E_m \; .$$

Weil diese Ungleichung für *alle* $m \in \mathbb{N}$ gilt, folgt nun wegen der Monotonie des Limes auch

$$e' \le e \; .$$

□

Bemerkung:
Die Folge (E_n) konvergiert wesentlich schneller gegen e wie die Folge (e_n). Für die Intervalle $[\tilde{e}_n, e_n]$ und $[\tilde{E}_n, E_n]$ gilt für alle $n \in \mathbb{N}$

$$[\tilde{E}_n, E_n] \subset [\tilde{e}_n, e_n] \qquad \text{und} \qquad \tilde{E}_n - E_n = \frac{1}{n \cdot n!} \; ,$$

während

$$\tilde{e}_n - e_n = \frac{1}{n} e_n \le \frac{4}{n}$$

gilt.

Aus

$$E_n \leq e_n \leq E_n + \frac{1}{n \cdot n!}$$

und der strengen Monotonie von E_n und $\tilde{E}_n$ erhält man sogar die Abschätzung

$$(*) \qquad \boxed{0 < e - E - n < \frac{1}{n \cdot n!}, \quad n \in \mathbb{N}}$$

Die folgende Tabelle verdeutlicht den Sachverhalt.

n	e_n	E_n	$\tilde{E}_n$	$\tilde{e}_n$
1	2	2.7182818284590452353	3	4
2	2.250000000	2.5000000000000000000	2.7500000000000000000	3.375000000
3	2.370370370	2.6666666666666666666	2.7222222222222222222	3.160493827
4	2.441406250	2.7083333333333333333	2.7187500000000000000	3.051757813
5	2.488320000	2.7166666666666666666	2.7183333333333333333	2.985984000
6	2.521626372	2.7180555555555555555	2.7182870370370370370	2.941897434
7	2.546499697	2.7182539682539682539	2.7182823129251700680	2.910285368
8	2.565784514	2.7182787698412698412	2.7182818700396825396	2.886507578
9	2.581174792	2.7182815255731922398	2.7182818317656280619	2.867971991
10	2.593742460	2.7182818011463844797	2.7182818287037037037	2.853116706
11	2.604199012	2.7182818261984928651	2.7182818284759572638	2.840944377
12	2.613035290	2.7182818282861685639	2.7182818284601415388	2.830788231
13	2.620600888	2.7182818284467590023	2.7182818284591121129	2.822185572
14	2.627151556	2.7182818284582297479	2.7182818284590490868	2.814805239
15	2.632878718	2.7182818284589944642	2.7182818284590454453	2.808403966
16	2.637928497	2.7182818284590422590	2.7182818284590452462	2.802799028
17	2.642414375	2.7182818284590450705	2.7182818284590452358	2.797850515
18	2.646425821	2.7182818284590452267	2.7182818284590452353	2.793449478
19	2.650034327	2.7182818284590452349	2.7182818284590452353	2.789509818
20	2.653297705	2.7182818284590452353	2.7182818284590452353	2.785962590
21	2.656263214	2.7182818284590452353	2.7182818284590452353	2.782751938
22	2.658969859	2.7182818284590452353	2.7182818284590452353	2.779832125
23	2.661450119	2.7182818284590452353	2.7182818284590452353	2.777165341
24	2.663731258	2.7182818284590452353	2.7182818284590452353	2.774720060
25	2.665836331	2.7182818284590452353	2.7182818284590452353	2.772469785
26	2.667784967	2.7182818284590452353	2.7182818284590452353	2.770392081
27	2.669593978	2.7182818284590452353	2.7182818284590452353	2.768467829
28	2.671277853	2.7182818284590452353	2.7182818284590452353	2.766680634
29	2.672849144	2.7182818284590452353	2.7182818284590452353	2.765016356
30	2.674318776	2.7182818284590452353	2.7182818284590452353	2.763462735
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
∞	2.718281828	2.718281828	2.718281828	2.718281828

Aus $(*)$ ergibt sich ein einfacher Beweis für die *Irrationalität* von e.

9.6 Satz: Die Eulersche Zahl e ist irrational.

Wegen $2 < e < 3$ ist e sicher keine ganze Zahl. Wäre nun e rational, so müsste sich e in der Form $e = \frac{p}{q}$ mit $p, q \in \mathbb{N}$, $q \geq 2$, schreiben lassen.
Die Formel $(*)$ liefert für $n = q$

$$0 < e - E_q < \frac{1}{q \cdot q!} \; .$$

Multipliziert man diese Ungleichung mit $q!$, so folgt

$$0 < e \cdot q! - q! E_q < \frac{1}{q} \leq \frac{1}{2} < 1$$

Dann ist $g := q!E_q \in \mathbb{Z}$ und $eq! = p(q-1)!$ ist ebenfalls eine ganze Zahl, also ist ihre Differenz $d := eq! - q$ eine ganze Zahl, das steht aber im Widerspruch zu $0 < d < 1$. □

Die Eulersche Zahl e ist sogar *transzendent* (vgl. §10.5.17). Dies wurde zuerst von C.Hermite (1873) bewiesen.
Mit der Eulerschen Zahl e erhalten wir auch bessere Abschätzung für $n!$

9.7 Satz (Abschätzung für $n!$)

Für alle $n \in \mathbb{N}$ gilt

$$\boxed{e\left(\frac{n}{e}\right)^n \leq n! \leq en\left(\frac{n}{e}\right)^n}$$

Beweis : Für $n = 1$ ist die Ungleichung offensichtlich. Ist $n \geq 2$ so multipliziert man die Ungleichungen $e_k \leq e \leq \tilde{e}_k$ für $k = 1, 2, \ldots, n-1$ miteinander, wobei fest „Teleskopprodukte" entstehen, d.h. viele Faktoren kürzen sich raus:

$$\begin{aligned} e_1 \cdot e_2 \cdot e_3 \cdot \ldots \cdot e_{n-1} &= \left(\frac{2}{1}\right)^1 \cdot \left(\frac{3}{1}\right)^2 \cdot \ldots \cdot \left(\frac{n-1}{n-2}\right)^{n-2} \cdot \left(\frac{n}{n-1}\right)^{n-1} \\ &= \frac{n^{n-1}}{1 \cdot 2 \cdot (n-2)(n-1)} \\ &= \frac{n^{n-1}}{(n-1)!} \; . \end{aligned}$$

Analog erhält man

$$\tilde{e}_1 \cdot \tilde{e}_2 \cdot \tilde{e}_3 \cdot \cdots \cdot \tilde{e}_{n-1} = \frac{n^n}{(n-1)!}$$

und damit

$$\frac{n^{n-1}}{(n-1)!} \leq e^{n-1} \leq \frac{n^n}{(n-1)!} \; ,$$

woraus die Behauptung unmittelbar folgt. □

Bemerkung: Die Abschätzung für $n!$ werden wir später nochmals stark verbessern (Stichwort: *Stirlingsche Formel*)

Weitere Beispiele:
Als weiteres Beispiel betrachten wir das mehrfach in der Vorlesung vorgestellte „*Babylonische Wurzelziehen*“ von einem systematischen Standpunkt. Wir erhalten gleichzeitig einen neuen *Existenzbeweis für Quadratwurzeln* aus positiven reellen Zahlen.

Sei also $a > 0$ eine reelle Zahl, deren Quadratwurzel bestimmt werden soll. Wenn $x > 0$ Quadratwurzel von a sein soll, also der Gleichung $x^2 = a$ genügt, gilt $x = \frac{a}{x}$, anderfalls ist $x \neq \frac{a}{x}$. Dann ist das arithmetische Mittel $\frac{1}{2}\left(x + \frac{a}{x}\right)$ vielleicht ein besserer Näherungswert für die Quadratwurzel und durch Iteration erhält man tatsächlich eine Folge, die gegen die Wurzel aus a konvergiert.

9.8 Satz

Seien $a > 0$ und $x_0 > 0$ reelle Zahlen. Die Folge $(x_n)_{n\geq 0}$ sei rekursiv definiert durch

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) .$$

Dann gilt $x_n > 0$ für alle $n \in \mathbb{N}_0$ und die Folge (x_n) konvergiert monoton fallend (ab $n \geq 1$) gegen die eindeutig bestimmt positive Lösung der Gleichung $x^2 = a$.

Die Folge (y_n) mit $y_n = \frac{a}{x_n}$ konvergiert monoton wachsend ($n \geq 1$) gegen die mit $\sqrt{a}$ berechnete positive Lösung der Gleichung $x^2 = a$.

$([y_n, x_n])_{n\geq 1}$ ist eine Intervallschachtelung mit $y_n \leq \sqrt{a} \leq x_n (n \in \mathbb{N})$.

Beweis in mehreren Schritten

(1) Für alle n gilt $x_n > 0$.
Induktion nach n:
Induktionsanfang: $n = 0;\ x_0 \geq a > 0$.
Induktionsschritt: Sei $x_n > 0$ für $n \in \mathbb{N}$ schon gezeigt. Dann folgt auch

$$x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n}) > \frac{1}{2}(0+0) = 0 .$$

Also ist $y_n := \frac{a}{x_n}$ eine zulässige Bildung.

(2) Für alle $n \in \mathbb{N}$ gilt $x_n^2 \geq c$.
Es ist nämlich für $n > 1$

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

nach der AGM-Ungleichung.

(3) (x_n) ist monoton fallend: $x_{n+1} \leq x_n$ für alle $n \in \mathbb{N}$.

Es ist nämlich

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

(4) Mit $y_n = \frac{a}{x_n}$ gilt $y_n^2 \leq a$ für alle $n \in \mathbb{N}$. Denn aus $\frac{1}{x_n^2} \leq \frac{1}{a}$ folgt durch Multiplikation mit a^2:

$$y_n^2 = \frac{a^2}{x_n^2} \leq \frac{1}{a}a^2 = a \ .$$

(5) (y_n) ist monoton wachsend: $y_n \leq y_{n+1}$ für alle n, denn aus $x_{n+1} \leq x_n$ folgt $\frac{1}{x_n} \leq \frac{1}{x_{n+1}}$ und durch Multiplikation mit a

$$y_n = \frac{a}{x_n} \leq \frac{a}{x_{n+1}} = y_{n+1} \ .$$

(6) $y_n \leq x_n$ für alle $n \in \mathbb{N}$, denn andernfalls wäre $y_n > x_n > 0$ und damit $y_n^2 > x_n^2 \geq a$ mit Widerspruch zu (4).

(7) Wegen (3) und $y_1 \leq x_n$ für alle n ist (x_n) nach dem Monotonieprinzip konvergent.
Falls $x := \lim\limits_{n\to\infty} x_n$ gilt, dann auch $x \geq y_1 > 0$.
Wegen $y_n \leq x_n \leq x_1$ und (5) ist (y_n) ebenfalls konvergent und für $y := \lim\limits_{n\to\infty} \frac{a}{x_n}$ gilt

$$y = \frac{a}{x} \ .$$

Aus $x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$ folgt nach den Rechenregeln für Grenzwerte

$$\begin{aligned} \lim_{n\to\infty} x_{n+1} &= \lim_{n\to\infty} \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \\ &= \frac{1}{2}\left(\lim_{n\to\infty} x_n + \lim_{n\to\infty} \frac{a}{x_n}\right) \\ &= \frac{1}{2}\left(x + \frac{a}{x}\right) \end{aligned}$$

und wegen $x = \lim\limits_{n\to\infty} x_{n+1}$ genügt x der Gleichung

$$x = \frac{1}{2}\left(x + \frac{a}{x}\right) \ ,$$

also der Gleichung $x^2 = a$.

Damit ist gezeigt, dass die Folge (x_n) (und auch die Folge (y_n)) gegen eine Lösung der Gleichung $x^2 = a$ konvergiert.

Wegen $a \geq y_1 > 0$ ist es die positive Lösung der Gleichung.

Bezeichnung (wie früher): $x = \sqrt{a}$.

Damit haben wir einen weiteren Existenzbeweis für Wurzeln aus positiven reellen Zahlen und ein effektives Verfahren, solche Wurzeln mit *beliebiger Genauigkeit* zu berechnen.
Wegen

$$y_n = \frac{a}{x_n} \leq \sqrt{a} \leq x_n$$

hat man bei jedem Iterationsschritt eine Fehlerabschätzung für die gesuchte Quadratwurzel.

(8) Das $[y_n, x_n]$ eine Intervallschachtelung ist, sei als Aufgabe gestellt.
Die folgende Tabelle lässt erkennen, dass die Folge sehr schnell konvergiert, selbst bei exotischen Startwerten.

$a = 2$, $x_0 = 1$, $\sqrt{2} = 1.41421356237309504880168872 4$

n	x_n	$\frac{a}{x_n}$	$x_n - \sqrt{2}$
0	1	2.0000000000000000000000000000	0.414213
1	1.5	1.3333333333333333333333333333	0.085786
2	1.4166666666666666666666666666	1.4117647058823529411764705882 3	0.002453
3	1.4142156862745098039215686274	1.4142114384748700173310225303 2	$0.2 \cdot 10^{-5}$
4	1.4142135623746899106262955788	1.4142135623715001869770836681 1	$0.1 \cdot 10^{-11}$
5	1.4142135623730950488016896235	1.4142135623730950488016878249 1	$0.8 \cdot 10^{-24}$
6	1.4142135623730950488016887242	1.4142135623730950488016887242 0	$0.2 \cdot 10^{-48}$
7	1.4142135623730950488016887242	1.4142135623730950488016887242 0	$0.2 \cdot 10^{-97}$
8	1.4142135623730950488016887242	1.4142135623730950488016887242 0	0
9	1.4142135623730950488016887242	1.4142135623730950488016887242 0	0
10	1.4142135623730950488016887242	1.4142135623730950488016887242 0	0

$a = 2$, $x_0 = 0{,}001$

n	x_n	$\frac{a}{x_n}$	$x_n - \sqrt{2}$
0	0,001	2000.0000000000000000000000000	-1,4132135
1	1000,0005	0,0019999990000004999997500000	998,590786
2	500,0012499995000002499998750	0,0039999900000289999155002 46	498,587036
3	250,0026249947500146249576876	0,0079999160010499867431675 14	248,588411
4	125,0053124553755323058504275	0,0159993200346102181858296 91	123,591098
5	62,5106558877050712620181286 3	0,0319945451155211867197745 87	61,0964423
6	31,2713252164102962243689516 0	0,0639563557399370224603245 77	29,8571116
7	15,6676407860751166234146380 9	0,1276516373656928720977212 79	14,2534272
8	7,8976462117204047477561796 86	0,2532400092868080862919733 56	6,48343264
9	4,0754431105036064170240765 21	0,4907441830915063565389529 11	2,66122954
10	2,2830936467975563867815147 16	0,8760043648692880309255676 37	0,86888008
11	1,5795490058334222088535411 77	1,2661842036010360844862566 32	0,16533544
12	1,4228666047172291466698989 04	1,4056131427706579899900484 21	0,00865308
13	1,4142398737439435683299736 63	1,4141872514917591280771807 54	0,00002631
14	1,4142135626178513482035772 08	1,4142135621283387494421599 28	$2{,}3 \cdot 10^{-10}$
15	1,4142135623730950488228685 68	1,4142135623730950488228685 68	$2{,}1 \cdot 10^{-20}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Man spricht von *quadratischer Konvergenz*, weil

$$
\begin{aligned}
|x_{n+1} - \sqrt{a}| &= x_{n+1} - \sqrt{a} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) - \sqrt{a} \\
&= \frac{1}{2}\frac{(x_n - \sqrt{a})^2}{x_n} \\
&\leq \frac{|x_n - \sqrt{a}|^2}{2\sqrt{a}} = C|x_n - \sqrt{a}|^2
\end{aligned}
$$

(mit $C := \frac{1}{2\sqrt{a}}$) gilt.

Hat man also z.B. $\sqrt{2}$ beim n-ten Iterationsschritt mit einem Fehler $< 10^{-n}$ approximiert, dann gilt beim nächsten Schritt

$$|x_{n+1} - \sqrt{2}| < \frac{10^{-2n}}{2\sqrt{2}} < 10^{-2n}.$$

Da –wie mehrfach plausibel gemacht– das Babylonische Wurzelziehen ein Spezialfall des Newton-Verfahren ist -und dieses, wenn es konvergiert, i.A. ziemlich blitzartig konvergiert, ist die quadratische Konvergenz nicht sehr überraschend.

Das Verfahren hat außerdem den Vorteil *selbstkorrigierend* zu sein:
Da *jede* positive Zahl als Startwert verwendet werden kann, können Rechenfehler, insbesondere Rundungsfehler, den Ablauf des Algorithmus nicht gänzlich verfälschen, sondern allerfalls verzögern. Wollen wir z.B. $\sqrt{2}$ auf 100 Nachkommastellen genau berechnen, bräuchten wir im obigen Beispiel die Rechnung nicht von Anfang an mit 100-stelliger Genauigkeit durchzuführen, sondern etwa mit einem Näherungswert mit 35 Nachkommastellen bestimmen und erhielten nach zwei weiteren Iterationsschritten das gewünschte Ergebnis.
Sind außerdem a und der Startwert x_0 *rational*, so sind alle x_n rational.

Alles in Allem: Es handelt sich hier um den *Idealfall* eines schnell konvergierenden und stabilen *nummerischen Verfahrens*.
Leider sind die Verhältnisse bei anderen nummerischen Verfahren nicht immer so ideal.

9.9 Existenz k-ter Wurzeln ($k \in \mathbb{N},\ k \geq 2$)

Eine Verallgemeinerung auf k-te Wurzeln ($k \in \mathbb{N},\ k \geq 2$) liegt nahe.

Sind $a > 0$ und $x_0 > 0$ beliebige reelle Zahlen. Definiert man die Folge $(x_n)_{n \geq 0}$ rekursiv durch

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

dann konvergiert die Folge (x_n) gegen die eindeutig bestimmte positive Lösung der Gleichung

$$x^k = a.$$

Wie man auf diese rekursive Folge kommt, wurde in der Vorlesung begründet.

Damit hat man einen weiteren *Existenzbeweis* für k-te Wurzeln. Der Spezialfall $k = 2$ ist hierin enthalten.

Zum Beweis vergleiche man die Musterlösung zu Aufgabe 4 von Blatt 2.

9.10 Kreismessung nach Archimedes

Die Berechnung des Flächeninhalts der Einheitsscheibe $\bar{\mathbb{E}} := \{(x, y) \in \mathbb{R};\ x^2 + y^2 \leq 1\}$ nach dem Vorbild von Archimedes sei auf Übungsblatt 7 Aufgabe 5 und die entsprechende Musterlösung verwiesen.

9.11 Dezimalbruchdarstellung einer reellen Zahl

Als Vorbereitung für die *Theorie der Reihen* betrachten wir ein letztes Beispiel für eine Intervallschachtelung.

Ist $x \in \mathbb{R}$, $x \geq 0$, dann sei

$$\begin{aligned}
Z &:= \{0, 1, \ldots, 9\} && \text{und}\\
x_0 &:= [x] = max\{k \in \mathbb{R};\ k \leq x\} && \text{und}\\
x_1 &:= max\{k \in \mathbb{Z};\ x_0 + \frac{k}{10} \leq x\}\\
&\vdots\\
x_n &:= max\{k \in \mathbb{Z};\ x_0 + \frac{x_1}{10} + \frac{k}{10^2} \leq x\}\\
&\vdots
\end{aligned}$$

Ist x_n schon definiert, so sei

$$x_{n+1} := max\left\{k \in \mathbb{Z};\ x_0 + \frac{x_1}{10} + \cdots + \frac{x_n}{10^n} + \frac{k}{10^{n+1}} \leq x\right\}.$$

Man erhält zwei Folgen (a_n) und (b_n)

$$\begin{aligned}
a_n &:= x_0 + \frac{x_1}{10} + \cdots + \frac{x_n}{10^n} \quad \text{und}\\
b_n &:= a_b + \frac{1}{10^n}
\end{aligned}$$

mit $a_n \leq x < b_n$ für alle $n \in \mathbb{N}$.
Offensichtlich ist (a_n) monoton wachsend und (b_n) monoton fallend. Wegen

$$b_n - a_n = \frac{1}{10^n}$$

ist $([a_n, b_n])$ eine Intervallschachtelung mit $x = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$. Man schreibt auch

$$x = x_0 + \lim_{n\to\infty} \sum_{k=1}^{n} \frac{x_k}{10^k} = x_0, x_1 x_2 x_3 x_4 \ldots$$

Die so gewonnene Darstellung von x nennt man die *Dezimalbruchdarstellung von x*. Wir kommen hierauf im Abschnitt über Reihen ausführlich zurück.

Monotone und beschränkte Folgen haben also ein gutes Konvergenzverhalten.
Um unsere Sätze über monotone Folgen auf allgemeine Folgen anwenden zu können, zeigen wir einen bemerkenswerten Satz, dass nämlich *jede* reelle Zahlenfolge eine *monotone Teilfolge* besitzt.
Betrachtet man z.B. die Folge

$$\begin{aligned}
(a_n) &= \left(\frac{(-1)^n}{n}\right),\ n \in \mathbb{N} \text{ und die Teilfolgen}\\
(a_{2k}) &= (a_2, a_4, a_6, \ldots) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \ldots\right),\\
(a_{2k-1}) &= (a_1, a_3, a_5, \ldots) = \left(-1, -\frac{1}{3}, -\frac{1}{5}, \ldots\right),\\
(a_{2^k}) &= (a_2, a_4, a_8, \ldots) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots\right),
\end{aligned}$$

so sind diese alle monoton, was auf die ursprüngliche Folge nicht zutrifft. Mit Hilfe der folgenden Begriffsbildung lässt sich das folgende etwas überraschende Lemma 9.13 beweisen.

9.12 Definition

Eine Zahl $n \in \mathbb{N}$ heißt *Gipfelindex* der reellen Zahlenfolge (a_n), falls für alle $m > n$ gilt $a_n > a_m$.

9.13 Lemma

Jede reelle Zahlenfolge (a_n) besitzt eine monotone Teilfolge.

Beweis : Zwei Fälle sind denkbar:

(a) Es gibt eine Folge
$$n_1 < n_2 < \ldots < n_k < n_{k+1} < \ldots$$
von Gipfelindizes, dann ist
$$a_{n_1} > a_{n_2} > a_{n_3} > \ldots > a_{n_k} > a_{n_{k+1}} > \ldots\ ,$$
also (a_{n_k}) eine (sogar streng) monoton fallende Teilfolge von (a_n).

(b) Die Menge der Gipfelindizes ist endlich: Seien etwa
$$n'_1 < n'_2 < \ldots < n'_r \qquad (r \in \mathbb{N})$$
alle Gipfelindizes.
Man setze $n_1 := n'_r + 1$. Da n_1 kein Gipfelindex ist, gibt es ein $n_2 \in \mathbb{N}$ mit $n_2 > n_1$ und $a_{n_1} \leq a_{n_2}$. Da auch n_2 kein Gipfelindex ist, gibt es $n_3 > n_2$ mit $a_{n_2} \leq a_{n_3}$ etc.
Man erhält eine monoton wachsende Teilfolge (a_{n_k}) von (a_n).

$\square$

Mit Hilfe des Lemmas ergibt sich nun ganz einfach das Auswahlprizip von *Bolzano-Weierstraß*.

9.14 Theorem (Bolzano-Weierstraß)

(B.Bolzano(1781-1841), K.Weierstraß (1815-1897))

Jede beschränkte reelle Zahlenfolge besitzt eine konvergente Teilfolge.

Beweis : *Jede* reelle Zahlenfolge besitzt nach dem Lemma eine monotone Teilfolge. Ist die gegebene Folge beschränkt, dann ist auch jede Teilfolge beschränkt, insbesondere also die nach dem Lemma existierende monotone Teilfolge. Diese ist dann nach dem Monotonieprinzip konvergent. $\square$

Bemerkung:
Ist $h \in \mathbb{R}$ bzw. $h \in \mathbb{C}$ der Grenzwert einer konvergenten Teilfolge einer reellen bzw. komplexen Zahlenfolge (a_n), dann nennt man h auch *Häufungswert* der Folge (a_n).

Man kann den Satz von Bolzano-Weierstraß dann auch so aussprechen:

Satz: Jede *beschränkte* reelle oder komplexe Zahlenfolge besitzt einen Häufungswert.

Die Übertragung auf den komplexen Fall erhält man dabei sofort durch Anwendung des reellen Satzes auf Realteil und Imaginärteil der komplexen Folge.

Wir beweisen als nächstes eine bemerkenswerte *Verdichtungseigenschaft* konvergenter reeller oder komplexer Zahlenfolgen. Sei dazu (a_n) eine konvergente (reelle oder komplexe) Zahlenfolge mit dem Grenzwert a ($a \in \mathbb{R}$ oder $a \in \mathbb{C}$). Dann gibt es zu jedem $\varepsilon > 0$ einen Index $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt

$$|a_n - a| < \frac{\varepsilon}{2}\ .$$

Ist dann auch $m \in \mathbb{N}$ und $m \geq N$, dann ist ebenso

$$|a_m - a| < \frac{\varepsilon}{2}\ .$$

Deshalb folgt für alle $n, m \geq N$

$$\begin{aligned} |a_m - a_n| &= |a_m - a + a - a_n| \\ &= |a_m - a| + |a - a_n| \\ &= |a_m - a| + |a_n - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\ . \end{aligned}$$

Etwas locker formuliert bedeutet das:
Bei einer konvergenten Folge (a_n) liegen die Folgenglieder mit hinreichend großem Index beliebig dicht beieinander. Zur Präzisierung führt man den folgenden Begriff ein.

9.15 Definition (Cauchy-Folge)

(A.L.Cauchy (1789-1857))

Eine reelle ode komplexe Zahlenfolge heißt *Cauchy-Folge* (oder Fundamental-Folge), wenn sie die folgende Bedingung erfüllt:

(CF): Zu jedem $\varepsilon > 0$ gibt es einen Index $N \in \mathbb{N}$, so dass für alle $m, n \in \mathbb{N}$ mit $m, n \geq N$ gilt:

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

Unsere obige Überlegung hat gezeigt:

Satz: Jede *konvergente* reelle oder komplexe Zahlenfolge ist eine Cauchy-Folge.

Von großer Bedeutung ist, dass hiervon auch die Umkehrung, also das folgende Theorem gilt:

9.16 Theorem (Cauchysches Konvergenzprinzip)

Eine reelle oder komplexe Zahlenfolge ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.

Beweis : Da eine komplexe Zahlenfolge (z_n) offensichtlich genau dann eine Cauchy-Folge ist, wenn die Folgen $(Re\ z_n)$ und $(Im\ z_n)$ Cauchy-Folgen in $\mathbb{R}$ sind, können wir uns auf reelle Zahlenfolgen beschränken und haben also nur noch zu zeigen:
Wenn (a_n) eine Cauchy-Folge reeller Zahlen ist, dann ist (a_n) konvergent.

Wir müssen also aus der Bedingung (CF) uns irgendwie einen Kandidaten für einen möglichen Grenzwert der Folge (a_n) verschaffen. Wir tun dies in drei Schritten:

1.Schritt: Jede Cauchy-Folge ist beschränkt.

2.Schritt: Jede beschränkte Folge besitzt eine konvergente Teilfolge (das ist die Aussage des Satzes von Bolzano-Weierstraß)

3.Schritt: Ist a der Grenzwert der konvergenten Teilfolge, dann konvergiert auch die Ausgangsfolge gegen a.

Zum 1.Schritt: Sei (a_n) die gegebene Cauchy-Folge. Dann erfüllt (a_n) also die Bedingung

(CF) : Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass für alle $m, n \in \mathbb{N}$ mit $m, n \geq N$ gilt:

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

Wählt man speziell $\varepsilon = 1$ und $n = N$, dann gilt $|a_m - a_N| < 1$ für alle $m \geq N$. Daher

$$|a_m| = |a_m - a_N + a_N| \leq |a_m - a_N| + |a_N| < 1 + |a_N|$$

für alle $m \geq N$.

Fast alle a_m liegen also in $U_{1+|a_N|}(0)$. Ist

$$M := max\{|a_1|, |a_2|, \dots, |a_{N-1}|\} \quad \text{und} R = max\{M, 1 + |a_N|\}\ ,$$

dann gilt offenbar $|a_n| \leq R$ für *alle* $n \in \mathbb{N}$.

Bei dieser Schlussweise war es übrigens unwesentlich, ob (a_n) eine reelle oder komplexe Folge ist.

Zum 2.Schritt: Nach Bolzano-Weierstraß hat jede beschränkte (reelle oder komplexe) Zahlenfolge eine konvergente Teilfolge. Wir können uns –wie oben bemerkt– auf den reelle Fall beschränken, das ist aber unwesentlich.

Sei also $(b_k) = (a_{n_k})$ die konvergente Teilfolge von (a_n) und $a := \lim\limits_{n\to\infty} b_k$.

Zum 3.Schritt: $a = \lim_{n\to\infty} a_n$. Dann müssen wir ausnutzen, dass (b_k) gegen a konvergiert und (a_n) die Bedingung (CF) erfüllt. Sei $\varepsilon > 0$ beliebig vorgegeben.
Wegen (CF) gibt es ein $N \in \mathbb{N}$, so dass für alle $n, m \in \mathbb{N}$ mit $n, m \geq N$ gilt $|a_n - a_m| < \varepsilon$.
Wegen $a = \lim_{k\to\infty} b_k = \lim_{k\to\infty} a_{n_k}$ gibt es ein $k \in \mathbb{N}$, so dass für alle $b \in \mathbb{N}$ mit $b \geq k$ gilt

$$|b_k - a| = |a_{n_k} - a| < \frac{\varepsilon}{2}\ .$$

Ist nun $n \geq M := \max\{N, k\}$, so ist (beachte $n_M \geq M$)

$$\begin{aligned} |a_n - a| &= |a_n - a_{n_M} + a_{n_M} - a| \\ &\leq |a_n - a_{n_M}| + |a_{n_M} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\ , \end{aligned}$$

also konvergiert (a_n) gegen a.

□

9.17 Beispiele zum Cauchy-Kriterium

Das Cauchy-Kriterium erlaubt es Aussagen über die Konvergenz bzw. Divergenz einer Folge zu machen, ohne dass man den Grenzwert kennt.

Wir benutzen es als Divergenzkriterium:
Gibt es ein $\varepsilon_0 > 0$, so dass es für alle $N \in \mathbb{N}$ stets natürliche Zahlen $m, n \geq N$ gibt, für welche $|a_m - a_n| \geq \varepsilon_0$ gilt, dann ist (a_n) nicht konvergent, denn (CF) ist für ε_0 nicht erfüllt.

(1) Die Folge (H_n) mit

$$H_n = 1 + \frac{1}{2} + \ldots + \frac{1}{2} = \sum_{k=1}^{n} \frac{1}{k}$$

ist nicht konvergent, denn wählt man $\varepsilon_0 = \frac{1}{2},\ N \in \mathbb{N}$ beliebig, $n = N$ und $m = 2N$, dann ist

$$H_m - H_n = \frac{1}{N+1} + \frac{1}{N+2} + \ldots + \frac{1}{2N} > N \cdot \frac{1}{N} = \frac{1}{2}\ .$$

Die Folge (H_n) ist also keine Cauchy-Folge, also nicht konvergent. Dies hatten wir auch schon aus der Unbeschränktheit von (H_n) gefolgert:
Für $n \geq 2^k$ ist (vgl. §7.8(17))

$$H_n \geq H_{2^k} \geq 1 + \frac{k}{2} > \frac{k}{2}\ .$$

(2) Bei der Folge (s_n) mit

$$s_n = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{(-1)^{n-1}}{n}\ ,$$

berechnet man zunächst einige Glieder:

$$\begin{array}{lll} s_1 & = 1 & \\ s_2 & = 1 - \frac{1}{2} & = 0,5 \\ s_3 & = 1 - \frac{1}{2} + \frac{1}{3} & = 0,833 \\ s_4 & = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} & = 0,583\ldots \\ s_5 & = \ldots & = 0,783\ldots \\ s_6 & = \ldots & = 0,616\ldots \\ s_7 & = \ldots & = 0,759\ldots \\ s_8 & = \ldots & = 0,634\ldots \\ s_9 & = \ldots & = 0,745\ldots \end{array}$$

Das legt die Vermutung nahe, dass die Teilfolge (s_{2k+1}) monoton fällt und die Teilfolge (s_k) monoton wächst und dass sie möglicherweise gegen den gleichen Grenzwert, der ungefähr bei 0,7 liegt, konvergieren.

Da wir aber keine Vermutung über den potentiellen Grenzwert haben, verwenden wir das Cauchy-Kriterium um die Konvergenz von (s_n) zu zeigen.
Wir können im Cauchy-Kriterium oBdA $m > n$, also $m = n + k$ mit $k \in \mathbb{N}$ voraussetzen. Dann ist

$$\begin{aligned} s_m - s_n = s_{n+k} - s_n &= \frac{(-1)^n}{n+1} + \frac{(-1)^{n+1}}{n+2} + \ldots + \frac{(-1)^{n+k-1}}{n+k} \\ &= (-1)^n \left(\frac{1}{n+1} - \frac{1}{n+1} + \frac{1}{n+3} \mp \ldots + \frac{(-1)^{k-1}}{n+k} \right) \end{aligned}$$

Wir betrachten den Klammerausdruck: Sei

$$a := \frac{1}{n+1} - \frac{1}{n+1} + \frac{1}{n+3} \mp \ldots + \frac{(-1)^{k-1}}{n+k} \ .$$

Wegen $\frac{1}{k} - \frac{1}{k+1} > 0$ ergibt sich durch Zusammenfassen zweier aufeinander folgenden Summanden

$$a = \left(\frac{1}{n+1} - \frac{1}{n+2} \right) + \left(\frac{1}{n+3} - \frac{1}{n+4} \right) + \ldots + \begin{cases} \frac{1}{n+k-1} - \frac{1}{n+k} \ , & \text{falls } k \text{ gerade,} \\ \frac{1}{n+k} \ , & \text{falls } k \text{ ungerade,} \end{cases}$$

dass $a > 0$ ist.

Lässt man in a den ersten Summanden unverändert und fasst dann je zwei aufeinander folgende Summanden zusammen, so ergibt sich

$$a = \frac{1}{n+1} - \left(\frac{1}{n+2} - \frac{1}{n+3} \right) - \left(\frac{1}{n+4} - \frac{1}{n+5} \right) - \ldots - \begin{cases} \frac{1}{n+k} \ , & \text{falls } k \text{ gerade,} \\ \frac{1}{n+k-1} - \frac{1}{n+k} \ , & \text{falls } k \text{ ungerade.} \end{cases}$$

Wegen $\frac{1}{k} - \frac{1}{k+1} > 0$ folgt nun $a < \frac{1}{n+1}$ und damit folgt

$$|s_{n+k} - s_n| = |a| < \frac{1}{n+1} \ .$$

Jetzt läuft der eigentliche Konvergenzbeweis routinemäßig:
Zu vorgegebenem $\varepsilon > 0$, gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $\frac{1}{n} < \varepsilon$.

Für alle $n \geq N$ ist dann

$$|s_{n+k} - s_n| < \frac{1}{n+1} < \frac{1}{n} < \varepsilon \ .$$

(s_n) ist also eine Cauchy-Folge, nach dem Cauchyschen Konvergenzkriterium also konvergent. Wie wir später sehen werden ist der Grenzwert der natürliche Logarithmus von 2:

$$\ln 2 = 0,6931471805599\ldots$$

(3) Setzt man im Cauchy-Kriterium $m = n + 1$, so erhält man, dass die Folge $(|a_{n+1} - a_n|)$ eine Nullfolge sein muss. Hieraus folgt aber noch nicht die Konvergenz von (a_n). Im Beispiel (1) ist

$$|H_{n+1} - H_n| = \frac{1}{n+1} ,$$

also ist $(H_{n+1} - H_n)$ eine Nullfolge, $H(n)$ ist aber divergent.

Wenn jedoch die Folge $|a_{n+1} - a_n|$ „schnell" gegen Null konvergiert, folgt auch die Konvergenz von (a_n):

Für eine reelle oder komplexe Zahlenfolge (a_n) gelte folgende Bedingung:

(B): Es gibt ein $q \in]0, 1[$ und ein $C > 0$ und ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt

$$|a_{n+1} - a_n| \leq Cq^n ,$$

dann ist (a_n) konvergent.

Man zeigt, dass (a_n) eine Cauchy-Folge ist: Für $m > n \geq N$ gilt

$$\begin{aligned} |a_m - a_n| &= \left|\sum_{k=n}^{m-1}(a_{k+1} - a_k)\right| \leq \sum_{k=n}^{m-1} |(a_{k+1} - a_k)| \\ &\leq C\sum_{k=n}^{m-1} q^k = Cq^n \sum_{k=0}^{m-1-n} q^k \leq \frac{C}{1-q} q^n . \end{aligned}$$

Da (q^k) eine Nullfolge ist, gibt es zu beliebig vorgegebenem $\varepsilon > 0$ ein $n \geq N$ mit

$$\frac{C}{1-q}\, q^n < \varepsilon ,$$

also mit $|a_m - a_n| < \varepsilon$, für alle $m > n \geq N$.

9.18 Zur Bedeutung des Cauchy-Kriteriums

Seine Bedeutung liegt zunächst darin, dass es ein notwendiges und hinreichendes Kriterium für die Konvergenz einer Folge darstellt, in welchem der hypotetische Grenzwert der Folge nicht vorkommt. Wir haben das Cauchy-Kriterium aus dem Satz von Bolzano-Weierstraß, diesen wiederum aus dem Monotonie-Prinzip gefolgert. Das Monotonieprinzip war eine unmittelbare Konsequenz aus dem Vollständigkeitsaxiom (Supremumsaxiom). Lehrbuchautoren wie O.Forster verwenden das Cauchy-Kriterium (und das Archimedische Axiom) als Vollständigkeitsaxiom. Letztlich sind diese Axiome alle äquivalent. Dass das Cauchy-Kriterium besonders anschaulich ist, kann man wohl nicht behaupten. Dennoch ist es in gewisser Weise fundamental, wenn man allgemeinere Räume, etwa metrische Räume, betrachtet und dort den Begriff der Vollständigkeit definieren will. Eine Folge (a_n) in einem metrischen Raum (X, d) heißt *konvergent*, wenn es ein $a \in X$ mit folgenden Eigenschaften gibt:
In jeder ε-Umgebung $U_\varepsilon(a)$ liegen fast alle Folgenglieder, d.h. zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$, $n \geq N$, gilt $d(x_n, a) < \varepsilon$.

Das bedeutet nichts anderes als dass die Folge $(d(x_n, a))$ der Abstände eine Nullfolge in $\mathbb{R}$ ist. Der Begriff der Cauchy-Folge lässt sich in jedem metrischen Raum prägen:

Eine Folge (a_n) von Elementen eines metrischen Raums (X, d) heißt *Cauchy-Folge*, wenn sie folgende Bedingung erfüllt:

(CF): Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass für alle $m, n \in \mathbb{N}$, mit $m, n \geq N$ gilt:

$$d(a_m, a_n) < \varepsilon$$

Jede konvergente Folge in X ist Cauchy-Folge. Dies zeigt man wie in $\mathbb{R}$ oder $\mathbb{C}$.

(X, d) heißt *vollständiger metrischer Raum*, falls jede Cauchy-Folge in X einen Grenzwert in X hat.

Unsere Überlegungen haben gezeigt:

$\boxed{\mathbb{R} \text{ und } \mathbb{C} \text{ sind bezüglich der aus dem Betrag abgeleiteten Metrik vollständige metrische Räume.}}$

Ist X ein *normierter Raum* (vgl. §5.8) und ist X bzgl. der aus der Norm abgeleiteten Metrik ein vollständiger metrischer Raum, dann heißt X ein *Banach-Raum*.

Ist die Norm aus einem *Skalarprodukt* auf den $\mathbb{K}$-Vektorraum X abgeleitet und X vollständig bezüglich der aus der Norm abgeleiteten Metrik, dann heißt X ein *Hilbert-Raum*.

$\boxed{\mathbb{R},\ \mathbb{C},\ \mathbb{R}^n,\ \mathbb{C}^n \text{ sind Hilbert-Räume, speziell also auch Banach-Räume.}}$

$\mathbb{Q}$ ist kein vollständiger metrische Raum bezüglich der von $\mathbb{R}$ geerbten Metrik, denn es gibt jede Menge Cauchy-Folgen in $\mathbb{Q}$, die nicht gegen Elemente aus $\mathbb{Q}$ konvergieren:

So gilt z.B. für die durch $x_0 := 1$ und $x_{n+1} := \frac{1}{2}\left(x_n + \frac{2}{x_n}\right)$ für $n \geq 1$ definierte Folge (x_n), $x_n \in \mathbb{Q}$ für alle $n \in \mathbb{N}$: (x_n) konvergiert zwar in $\mathbb{R}$ mit $\lim\limits_{n\to\infty} x_n = \sqrt{2}$, aber $\sqrt{2} \notin \mathbb{Q}$.

Auch die Folgen (e_n) und (E_n) mit

$$e_n = \left(1 + \frac{1}{n}\right)^n \quad \text{bzw.} \quad E_n = \sum_{k=0}^{n} \frac{1}{k!}$$

sind Cauchy-Folgen rationaler Zahlen, die aber in $\mathbb{Q}$ nicht konvergieren.

Bei einem konstruktiven Aufbau des Zahlensystems, ausgehend von den natürliche Zahlen $\mathbb{N}$, den ganzen Zahlen, den rationalen Zahlen, kann man reelle Zahlen als Äquivalenzklassen von Cauchy-Folgen rationaler Zahlen einführen.
Die rationalen Zahlenfolgen (e_n) und (E_n) ergeben dabei dieselbe reelle Zahl, nämlich die Eulersche Zahl e.
Zwei Cauchy-Folgen (a_n) und (b_n) rationaler Zahlen heißen dabei äquivalent, wenn $(a_n - b_n)$ eine (rationale) Nullfolge ist.
Auf der Menge der Folgen rationaler Zahlen wird hierdurch eine Äquivalenzrelation definiert. Auf der Menge der Äquivalenzklassen kann man dann eine Addition und Multiplikation so erklären, dass eine Körperstruktur entsteht. Algebraisch liegt dies daran, dass die Menge aller Nullfolgen von rationalen Zahlen ein *Ideal* ist, sogar ein *maximales* und der Restklassen-Ring R/m eines Ringes R nach einem maximalen Ideal m ist ein Körper (vgl. z.B. Fischer-Sacher: Algebra; Teubner Studienbücher, Satz 3.2.2)

Eine solche auf G.Cantor(1845-1918) zurückgehende Konstruktion von $\mathbb{R}$ findet man z.B. in W.Kaballo: Einführung in die Analysis 1; 3.Auflage, dort siehe §15 oder in Amann-Escher: Analysis 1, Seite 180ff.

9.19 Häufungswerte, $\underline{lim}$, $\overline{lim}$, uneigentliche Konvergenz

Die konvergenten reellen oder komplexen Zahlenfolgen bilden eine außerordentlich wichtige Klasse in der Gesamtheit aller reellen bzw. komplexen Zahlenfolgen (sie bilden jeweils $\mathbb{K}$-Vektorräume).

Dies zeigt sich insbesondere in den einfachen Rechenregeln, die wir für konvergente Folgen ableiten konnten.

Einfache Folgen, wie z.B. $((-1)^n)$, sind jedoch nicht konvergent. Die Folge $((-1)^n)$ besitzt jedoch die beiden konvergenten Teilfolgen $((-1)^{2k})$ und $((-1)^{2k-1})$, $k \in \mathbb{N}$. Nach dem Satz von Bolzano-Weierstraß (vgl. §9.14) besitzt *jede* beschränkte reelle oder komplexe Zahlenfolge eine konvergente Teilfolge.

Besitzt eine reelle oder komplexe Zahlenfolge (a_n) eine konvergente Teilfolge (a_{n_k}) und ist $a = \lim\limits_{n\to\infty} a_{n_k}$ ihrer Grenzwert, so hatten wir diesen in §9.14 auch einen *Häufungswert* der Folge (a_n) genannt. Der Satz von Bolzano-Weierstraß lässt sich also umformulieren in

Satz (Bolzano-Weierstraß):
Jede beschränkte reelle oder komplexe Zahlenfolge besitzt einen Häufungswert.

9.19.1 Beispiele

(a) Die Folge $((-1)^n)$ hat die Häufungswerte 1 und -1 (und sonst keine weiteren, Beweis?).

(b) Die Folge $\left((-1)^n \frac{n}{n+1}\right)$ besitzt die Häufungswerte 1 und -1 und sonst keine weiteren.

(c) Die nicht beschränkte Folge $(n(1 + (-1)^n))$ besitzt den Häufungswert Null.

(d) Die Folge $(i^n)_{n\in\mathbb{N}}$ hat die Häufungswerte $1, i, -1, -i$ und sonst keine weiteren.

(e) Eine konvergente reelle oder komplexe Zahlenfolge (a_n) besitzt genau einen Häufungspunkt, nämlich ihren Grenzwert.

Das ist offensichtlich, denn ist $a = \lim\limits_{n\to\infty} a_n$, dann ist a Häufungswert von (a_n), denn durch $n_k = k$ wird eine Teilfolge mit $\lim\limits_{k\to\infty} a_{n_k} = a$ definiert.

Ist umgekehrt $h \in \mathbb{K}$ Häufungswert von (a_n), gibt es also eine Teilfolge (a_{n_k}) mit $h = \lim\limits_{k\to\infty} a_{n_k}$. Aber *jede* Teilfolge von (a_n) konvergiert gegen a, also ist $h = a$.

Umgekehrt gilt:

(f) Ist (a_n) beschränkt und besitzt (a_n) genau einen Häufungswert a, dann ist (a_n) konvergent und $a = \lim\limits_{n\to\infty} a_n$.

Beweis als Übungsaufgabe.

Unter Verwendung der Begriffe der *endlichen* bzw. *unendlichen* Menge (vgl. §2.1.6) lässt sich der Begriff des Häufungswertes einer Zahlenfolge wie folgt charakterisieren.

9.19.2 Satz (Charakterisierung von Häufungswerten)

Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.
Für eine Folge (a_n), $a_n \in \mathbb{K}$, und einen Punkt $a \in \mathbb{K}$ sind folgende Aussagen äquivalent:

(1) Zu jedem $\varepsilon > 0$ und jedem $N \in \mathbb{N}$ gibt es ein $n \in \mathbb{N}$ mit $n \geq N$ und $|a_n - a| < \varepsilon$.
Beachte: zu jedem (noch so großen) N gibt es noch größere Indizes n mit $|a_n - a| < \varepsilon$.

(2) Für jedes $\varepsilon > 0$ ist die Menge $\{n \in \mathbb{N};\ |a_n - a| < \varepsilon\}$ unendlich.

(3) Es gibt eine Teilfolge (a_{n_k}) von (a_n), die gegen a konvergiert (d.h. a ist Häufungswert von (a_n)).

Beweis : Wir zeigen

$$(1) \Longrightarrow (3) \Longrightarrow (2) \Longrightarrow (1)$$

und damit die Äquivalenz aller der Aussagen.

Es gelte (1). Zu $\varepsilon = 1$ und $N = 1$ gibt es dann eine natürliche Zahl n_1 mit $|a_{n_1} - a| < 1$. Nun setzte man $\varepsilon := \frac{1}{2}$ und $N := n_1 + 1$. Hierzu gibt es eine natürliche Zahl n_2 mit $n_2 > n_1$ und $|a_{n_2} - a| < \frac{1}{2}$. Durch Rekursion erhält man eine Folge (n_k) natürlicher Zahlen mit

$$n_k < n_{k+1} \quad \text{und} \quad |a_{n_k} - a| < \frac{1}{k} \qquad \text{für alle } k \in \mathbb{N}.$$

Offensichtlich ist (a_{n_k}) eine Teilfolge von (a_n) mit $\lim\limits_{k\to\infty} a_{n_k} = a$.
a ist also Häufungswert von (a_n).

$(3) \Longrightarrow (2)$
Ist $(b_k) = (a_{n_k})$ irgendeine gegen a konvergente Teilfolge von (a_n).
Fast alle (b_k) liegen dann in $U_\varepsilon(a)$ ($\varepsilon > 0$ sei vorgegeben), d.h. es gibt ein K, so dass für alle $k \geq K$ gilt

$$|b_k - a| = |a_{n_k} - a| < \varepsilon\ .$$

Das bedeutet, dass alle Indizes n_k mit $k \geq K$ in der Menge $\{m \in \mathbb{N};\ |a_m - a| < \varepsilon\}$ liegen. Weil aber $n_k \geq k$ für alle k gilt, kann die Menge dieser n_k nicht noch nach oben beschränkt sein, insbesondere kann sie nicht endlich sein, diese Menge ist also unendlich.

$(2) \Longrightarrow (1)$
Sind $\varepsilon > 0$ und $n \in \mathbb{N}$ beliebig vorgegeben, so muss es noch Indizes $n \in \mathbb{N}$ geben, für die $|a_n - a| < \varepsilon$ und n *nicht* $< N$ ist, denn sonst wäre die Menge $\{n \in \mathbb{N}; |a_n - a| < \varepsilon\}$ endlich. □

Als nützliche *Übung* für den Umgang mit Häufungswerten zeige man:

Aufgabe: Ist $x \in \mathbb{R}$ eine beliebig vorgegebene Zahl und die Folge $(a_n(x))$ sei definiert durch

$$a_n(x) = nx - [nx] \qquad (x \in \mathbb{R}, n \in \mathbb{N}_0)\ .$$

Ist x rational, dann hat die Folge unendlich viele Häufungswerte.
Ist jedoch x irrational (z.B. $x = \sqrt{2}$), so ist jede reelle Zahl $a \in [0, 1]$ Häufungswert der Folge $(a_n(x))$.

Ist nun (a_n) eine *beschränkte reelle* Zahlenfolge, so wollen wir zeigen, dass es Häufungswerte a^* und a_* von (a_n) gibt, so dass für *jeden* Häufungswert h von (a_n) gilt

$$a_* \leq h \leq a^* \ .$$

a^* heißt der *limes superior* von (a_n) und a_* heißt der *limes inferior* von (a_n).

Wir geben hierfür zwei Beweise, einen ausführlichen und für den zweiten lediglich eine Beweisskizze.

Ist $h \in \mathbb{R}$ ein Häufungswert der beschränkten Folge (a_n) und gilt etwa $\alpha \leq a_n \leq \beta$ für alle n, so gilt auch $\alpha \leq a_{n_k} \leq \beta$ für jede Teilfolge a_{n_k} von (a_n) und damit nach §8.2 auch

$$\alpha \leq h = \lim_{k\to\infty} a_{n_k} \leq \beta \ .$$

Die Menge H der Häufungswerte von (a_n) ist also ebenfalls nach oben (durch β) und nach unten (durch α) beschränkt. Also existieren $a_* = \inf H$ und $a^* = \sup H$.

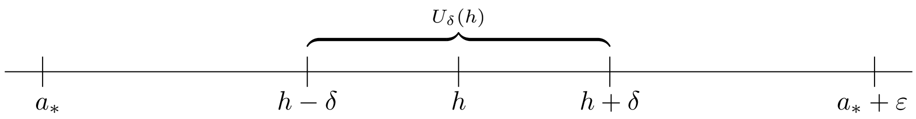
Wir zeigen: $a_* \in H$ und $a^* \in H$, also gilt $a_* = \min H$ und $a^* = \max H$.
Wir beschränken uns auf den Nachweis von $a_* = \min H$.

Nach der ε-Charakterisierung von inf gibt es zu jedem $\varepsilon > 0$ ein $h \in H$ mit

$$a_* \leq h < a_* + \varepsilon \ .$$

Ist nun $a_* = h$, dann ist $a_* = h = \min H \in H$. Ist aber $a_* < h$, dann gibt es ein $\delta > 0$ mit

$$a_* < h - \delta < h + \delta < a_* + \varepsilon \ .$$



Da h Häufungswert von (a_n) ist, gibt es unendlich viele $n \in \mathbb{N}$ mit $a_n \in U_\delta(h)$, wegen $U_\delta(h) \subset U_\varepsilon(a_*)$ liegen diese a_n auch in $U_\varepsilon(a_*)$, d.h. es gibt unendlich viele $n \in \mathbb{N}$ mit $a_n \in U_\varepsilon(a_*)$, d.h. a_* ist Häufungswert von (a_n).

Für a^* schließt man völlig analog.

Wir haben also Folgendes bewiesen:

9.19.3 Satz (Existenz vom $\underline{lim}$ und $\overline{lim}$)

Jede beschränkte reelle Zahlenfolge (a_n) besitzt einen größten Häufungswert (limes superior) und einen kleinsten Häufungswert (limes inferior).

Bezeichnungen: $\overline{\lim\limits_{n\to\infty}}\, a_n = \lim\limits_{n\to\infty} \sup a_n$ bzw. $\underline{\lim\limits_{n\to\infty}}\, a_n = \lim\limits_{n\to\infty} \inf a_n$.

Einen zweiten Beweis über die Existenz von $\overline{\lim}$ und $\underline{\lim}$ einer beschränkten reellen Zahlenfolge

$$(a_n), \ (a_n \in [\alpha, \beta], \ n \in \mathbb{N})$$

erhält man durch folgende Überlegung:

Für $n \in \mathbb{N}$ betrachten wir die Intervalle

$$K_n := [\inf\{a_k; k \geq n\}, \sup\{a_k; k \geq n\}] \ .$$

Dann gilt offenbar

$$[\alpha, \beta] \supset K_1 \supset K_2 \supset \ldots \supset K_n \supset K_{n+1} \supset \ldots$$

Die Folge (x_n) der linken Intervallenden

$$x_n = \inf\{a_k; k \geq n\}$$

ist monoton wachsend (warum?) und nach oben beschränkt (etwa durch β), die Folge (y_n) der rechten Intervallenden

$$y_n = \sup\{a_k; k \geq n\}$$

ist monoton fallend (warum?) und nach unten beschränkt (etwa durch β), also sind beide Folgen konvergent.

Sei $x := \lim\limits_{n\to\infty} x_n$ und $y := \lim\limits_{n\to\infty} y_n$. Da eine monoton wachsende, nach oben beschränkte Folge gegen das Supremum ihrer Wertemenge konvergiert, ist also

$$x := \lim_{n\to\infty} x_n = \lim_{n\to\infty} \{\inf\{a_n, a_{n+1}, \ldots\}\} = \sup\{x_n; n \in \mathbb{N}\}$$

und analog

$$y := \lim_{n\to\infty} y_n = \lim_{n\to\infty} \{\sup\{a_n, a_{n+1}, \ldots\}\} = \inf\{y_n; n \in \mathbb{N}\} \ .$$

Wegen $x_n \leq y_n$ für alle n ist auch $x \leq y$.

9.19.4 Satz

Es gilt

$$x = a_* = \min H = \varliminf_{n\to\infty} a_n \quad \text{und} \quad y = a^* = \max H = \varlimsup_{n\to\infty} a_n \ .$$

Wir zeigen $y = a^*$, indem wir $y \in H$ und $h \leq y$ für alle $y \in H$ nachweisen.

Wir verwenden die Charakterisierung von Häufungswerten aus Satz 19.9.2 und zeigen: Zu jedem $\varepsilon > 0$ und jedem $N \in \mathbb{N}$ gibt es ein $n \in \mathbb{N}$ mit $n \geq N$ und $|a_n - y| < \varepsilon$. Wegen $\lim\limits_{n\to\infty} y_n = y$, gibt es zunächst zu beliebig vorgegebenem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $m \in \mathbb{N}$ mit $m \geq N$ gilt

$$|y_m - y| < \frac{\varepsilon}{2} \ .$$

Nach Definition vom y_m gibt es dann ein $n \geq m$, so dass

$$|a_n - y_m| < \frac{\varepsilon}{2} \ .$$

Für $n \geq N$ ist aber dann

$$|a_n - y| \leq |a_n - y_m| + |y_m - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \ .$$

y ist also Häufungswert von (a_n).

Ist andererseit h ein Häufungswert von (a_n), d.h. $h \in H = H(a_n)$, dann gibt es eine Teilfolge (a_{n_k}) von (a_n) mit $\lim\limits_{k\to\infty} a_{n_k} = h$.

Nach der Definition ist aber $y_{n_k} \geq a_{n_k}$. Hieraus folgt

$$y = \lim_{n\to\infty} y_n = \lim_{k\to\infty} y_{n_k} \geq \lim_{k\to\infty} a_{n_k} = h \ .$$

Damit ist y der größte Häufungswert von (a_n), also $y = a^* = \max H = \overline{\lim_{n\to\infty}}(a_n)$.
Für $x = \lim_{n\to\infty} x_n$ schließt man völlig analog.

Bemerkung:
Für *konvergente* reelle Folgen sind die Begriffe $\overline{\lim}$ und $\underline{\lim}$ eigentlich überflüssig, weil in diesem Fall der Grenzwert der einzige Häufungswert ist und $\overline{\lim}$ und $\underline{\lim}$ mit dem Grenzwert übereinstimmt. Für *divergente* reelle Zahlenfolgen messen $\overline{\lim}$ und $\underline{\lim}$ einen wichtigen Aspekt der Nichtkonvergenz.

Als Übung im Umgang mit $\overline{\lim}$ und $\underline{\lim}$ zeige man:
Aufgabe: Eine Folge (a_n) reeller Zahlen ist ganau dann konvergent gegen $a \in \mathbb{R}$, wenn

$$\overline{\lim_{n\to\infty}}\, a_n = \underline{\lim_{n\to\infty}}\, a_n = a$$

gilt.

Häufig nützlich sind die folgenden ε-Charakterisierungen vom $\overline{\lim}$ und $\underline{\lim}$.

9.19.5 Satz (ε-Charakterisierung von $\overline{lim}$ und $\underline{lim}$)

Ist (a_n) eine beschränkte reelle Zahlenfolge und sind $a_*, a^* \in \mathbb{R}$. Dann ist

(a) $a^* = \overline{\lim_{n\to\infty}}\, a_n \Longleftrightarrow$ wenn für jedes $\varepsilon > 0$ gilt

- (α) Für fast alle Indizes $n \in \mathbb{N}$ ist $a_n < a^* + \varepsilon$.
- (β) Es gibt unendlich viele Indizes $m \in \mathbb{N}$ mit $a_m > a^* - \varepsilon$.

(b) $a_* = \underline{\lim_{n\to\infty}}\, a_n$ genau dann, wenn für jedes $\varepsilon > 0$ gilt:

- (α) Für fast alle Indizes n ist $a_* - \varepsilon < a_n$.
- (β) Es gibt unendlich viele $m \in \mathbb{N}$ mit $a_m < a_* + \varepsilon$.

Beweis : Wir zeigen exemplarisch (a).
Sei dazu $A_n := \{a_k; k \geq n\}$ und $y_n = \sup A_n,\ n \in \mathbb{N}$. Sei $a^* = \lim_{n\to\infty} y_n = \overline{\lim_{n\to\infty}}\, a_n$ und $\varepsilon > 0$ beliebig vorgegeben.

Da die Folge (y_n) monoton fällt, gilt $y_n \geq a^*$ für alle n.
Wegen $y_n = \sup A_n$, gibt es dann nach der ε-Charakterisierung von sup ein a_n mit

$$a_n > y_n - \varepsilon \geq a^* - \varepsilon \ .$$

Hieraus folgt (α).
Andererseits gibt es wegen $\lim_{n\to\infty} y_n = a^*$ ein $N \in \mathbb{N}$, so dass $y_n < a^* + \varepsilon$ für alle $n \geq N$. Hieraus folgt aber auch $a_n < a^* + \varepsilon$ für alle $n \geq N$. Das ist die Bedingung (β).

Seien jetzt umgekehrt die Bedingungen (α) und (β) erfüllt.

Zu vorgegebenem $\varepsilon > 0$ gibt es dann ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $a_n < a^* + \varepsilon$.
Hieraus folgt für alle $n \geq N \;\; y_n \leq a^* + \varepsilon$.
Andererseits folgt aus (β) $y_n > a^* - \varepsilon$ für alle $n \in \mathbb{N}$, also insgesamt

$$|y_n - a^*| \leq \varepsilon$$

für alle $n \geq N$, d.h. $\lim\limits_{n\to\infty} y_n = a^*$. □

9.19.6 Beispiele

(a) Für die Folge $(a_n) = \left((-1)^n \left(1 + \frac{1}{n}\right)\right)$ gilt

$$y_n := \sup\{a_k;\ k \geq n\} = \begin{cases} 1 + \frac{1}{n}\ , & \text{falls } n \text{ gerade,} \\ 1 + \frac{1}{n+1}\ , & \text{falls } n \text{ ungerade,} \end{cases}$$

also ist $\overline{\lim\limits_{n\to\infty}}\, a_n = \lim\limits_{n\to\infty} y_n = 1$.

Wegen

$$x_n := \inf\{a_k;\ k \geq n\} = \begin{cases} -\left(1 + \frac{1}{n}\right)\ , & \text{falls } n \text{ ungerade,} \\ -\left(1 + \frac{1}{n+1}\right)\ , & \text{falls } n \text{ gerade,} \end{cases}$$

gilt $\underline{\lim}_{n\to\infty} = \lim\limits_{n\to\infty} x_n = -1$.

(b) Für die Folge $(a_n) = (n)$ der natürlichen Zahlen ist nicht nach oben beschränkt. Es ist jedoch sinnvoll, in diesem Fall $\sup\{a_k;\ k \geq n\} = \infty$ zu setzen und auch $\overline{\lim\limits_{n\to\infty}}\, a_n = \infty$ zu definieren.

Wegen $\inf\{a_k,\ k \geq n\} = n$, ist auch hier die Definition $\underline{\lim\limits_{n\to\infty}}\, a_n = \infty$ sinnvoll.

Allgemein setzt man für eine nicht nach oben (bzw. unten) beschränkte Folge (a_n) reeller Zahlen

$$\overline{\lim_{n\to\infty}}\, a_n = \infty \quad \text{bzw.} \quad \underline{\lim_{n\to\infty}}\, a_n = -\infty\ .$$

Damit haben wir keine reellen Zahlen „∞“ bzw „$-\infty$“ eingeführt, sondern lediglich eine bequeme Sprechweise (façon de parler nach C.F.Gauß). Wir schreiben jedoch, dass für jede reelle Zahl x gelten soll

$$-\infty < x < \infty \quad \text{und} \quad -\infty < \infty\ .$$

Statt „∞“ findet man auch die Schreibweise „$+\infty$“.

Diese Betrachtungen hängen eng mit einer bestimmten *Sorte von Divergenz* reeller Zahlenfolgen zusammen.

9.19.7 Sprechweise

Eine reelle Zahlenfolge (a_n) wächst über jede Grenze (oder ist bestimmt divergent gegen ∞), wenn es zu jedem $C > 0$ einen Index $N \in \mathbb{N}$ gibt, so dass für alle $n \geq N$ gilt $a_n > C$.

Die reelle Folge (a_n) fällt unter jede Grenze (divergiert bestimmt gegen $-\infty$), falls die Folge $(-a_n)$ über jede Grenze wächst.

Schreibweise: Wächst (bzw. fällt) (a_n) über(unter) jede Grenze, so schreibt man

$$\lim_{n\to\infty} a_n = \infty \quad \text{bzw.} \quad \lim_{n\to\infty} a_n = -\infty \, .$$

Statt *bestimmt divergent* verwendet man auch den Ausdruck *uneigentlich konvergent*.

9.19.8 Beispiele und Bemerkungen

(a) Die Folge $(a_n) = (n)$ der natürlichen Zahlen wächst über jede Grenze.

(b) Die Folge $(a_n) = (-3^n)$ fällt unter jede Grenze.

(c) Die Folge $(a_n) = ((-1)^n n)$ divergiert, sie divergiert aber weder bestimmt gegen ∞ noch gegen $-\infty$. Jedoch wächst die Folge $(|a_n|)$ über jede Grenze.

(d) Eine Folge, die über jede Grenze wächst, kann nicht nach oben beschränkt sein. Die Umkehrung hiervon gilt i.A. nicht, wie Beispiel (c) zeigt.

(e) ∞ und $-\infty$ sind keine reellen Zahlen, vielmehr ist ihre Bedeutung durch die Definition der bestimmten Divergenz genau festgelegt.

Auch bei der Übertragung der *Rechenregeln für konvergente reelle Folgen* auf uneigentlich konvergente Folgen ist Vorsicht geboten. Wenn man mit ∞ und $-\infty$ wie mit reellen Zahlen rechnen könnte, so ergeben sich leicht Widersprüche:

Ist etwa $a_n = n,\ b_n = 1$ und $c_n = a_n + 1 = n + 1$. Dann ist

$$\lim_{n\to\infty} a - n = \infty, \quad \lim_{n\to\infty} b_n = 1, \quad \lim_{n\to\infty} c_n = \infty \, .$$

Ließen sich die Rechenregeln übertragen, so müßte gelten $\infty + 1 = \infty$. Nach den Körperaxiomen besitzt aber die Gleichung $a + x = a$ die eindeutige Lösung $x = 0$. Man erhält so den Widerspruch $1 = 0$.

Auf der Menge $\tilde{\mathbb{R}} := \mathbb{R} \cup \{\infty, -\infty\}$ lässt sich also keine Körperstruktur erklären.

Nützlich ist der folgende Zusammenhang zwischen der bestimmten Divergenz einer Folge und dem Begriff der Nullfolge:

9.19.9 Satz

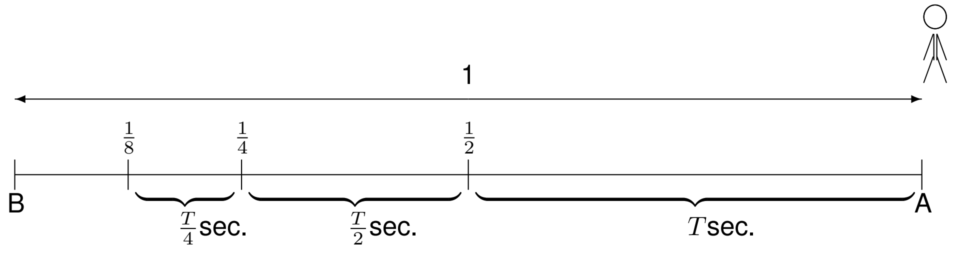
(a) Die reelle Folge (a_n) sei bestimmt divergent mit $\lim\limits_{n\to\infty} a_n = \infty$ oder $\lim\limits_{n\to\infty} a_n = -\infty$. Dann gibt es ein $N \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt $a_n \neq 0$ und für die Folge $\left(\frac{1}{a_n}\right)_{n \geq \infty}$ gilt $\lim\limits_{n\to\infty} a_n = 0$.

(b) Ist umgekehrt (a_n) eine Nullfolge mit $a_n > 0$ für (fast) alle n (bzw. < 0 für (fast) alle n), dann divergiert die Folge $\left(\frac{1}{a_n}\right)$ bestimmt gegen ∞ (bzw. $-\infty$).

Der Beweis dürfte der geneigten Leserin / dem geneigten Leser keine Schwierigkeiten bereiten. Wegen $\lim\limits_{n\to\infty} \frac{n^2}{2^n} = 0$ ist z.B. $\lim\limits_{n\to\infty} \frac{2^n}{n^2} = \infty$.

10 Reihen

Einführung: Wir betrachten einen Sprinter, der vom Punkt A zum Punkt B mit konstanter Geschwindigkeit läuft:



Wir nehmen an, dass die Länge der Strecke $\overline{AB}$ (gemessen in einer geeigneten Längeneinheit) 1 ist. Braucht er für die erste Hälfte der Strecke T sec., dann braucht er erfahrungsgemäß für die Gesamtstrecke $2T$ sec.

Man kann das Problem aber auch so betrachten:
Braucht er für die halbe Strecke T sec., so braucht er für das nächste Viertel $\frac{T}{2}$ sec,
für das nächste Achtel $\frac{T}{4}$ sec.,

$\vdots$

für die k-te Teilstrecke der Länge $\frac{1}{2^k}$ braucht er $\frac{T}{2^{k-1}}$ sec.

Die Frage, wie groß die Summe all dieser Zeiten (Gesamtzeit) ist, scheint geklärt, nämlich $2T$. Der Versuch die Teilzeiten aufzuaddieren führt auf das Problem, die unendlich vielen Zeitspannen

$$T, \frac{T}{2}, \frac{T}{4}, \frac{T}{8}, \dots, \frac{T}{2^{k-1}}, \dots$$

aufzuaddieren.

Nun ist in jedem Körper, die Addition von endlich vielen Zahlen $a_1, \dots, a_n$ erklärt und die Summe

$$s := a_1 + a_2 + \dots + a_n$$

ist unabhängig von der Reihenfolge und möglichen Klammerungen.
Für „unendlich viele“ Zahlen haben wir eine solche Summe aber bis jetzt nicht erklärt.

Die obige Aufgabe lässt sich aber lösen, wenn wir sie so interpretieren, dass lediglich die Folge (t_n) der „Teilzeiten“(Teilsummen)

$$t_1 := T, \quad t_2 := T + \frac{T}{2}, \quad t_3 = T + \frac{T}{2} + \frac{T}{4}, \quad \dots$$

also die Folge (t_n) mit $t_n = \sum\limits_{k=1}^{n} \frac{T}{2^{k-1}}$ auf Konvergenz zu untersuchen ist. Tatsächlich gilt nach der *geometrischen Summenformel* (vgl. §2.4(1))

$$\lim_{n\to\infty} t_n = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{T}{2^{k-1}} = T \cdot \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1-\left(\frac{1}{2}\right)^n}{1-\frac{1}{2}} = 2T\ .$$

Der griechische Philosoph Zenon von Elea (495-435 v.Chr.) argumentierte wie folgt:
„Der Läufer erreicht nie das Ziel, weil unendlich viele solche Zeitspannen keine endliche Gesamtheit ergeben können.“

Ein ähnliches von Zenon betrachtetes Beispiel ist „Achilles und die Schildkröte“. Der berühmte Held Achill veranstaltet einen Wettlauf mit einer (ziemlich schnellen) Schildkröte. Achill kann zehnmal schneller als die Schildkröte. Als fairer Sportsman gibt er der Schildkröte einen Vorsprung von 10 Ellen (1 Elle = geeignete Längeneinheit). Die Schildkröte und Achill starten zur gleichen Zeit.
Hat Achill die ersten 10 Ellen durcheilt, so ist die Schildkröte um eine Elle vorangekommen. Hat Achill diese Elle zurückgelegt, beträgt der Vorsprung der Schildkröte immer noch $\frac{1}{10}$ Ellen. Bringt Achill diese Strecke hinter sich, beträgt der Vorsprung der Schildkröte noch $\frac{1}{100}$ Ellen etc.
Der Vorsprung der Schildkröte wir zwar immer kleiner, aber er wird nie Null. Deshalb, so argumentiert Zenon, kann Achill die Schildkröte „niemals“ einholen.

Die Aufklärung des Paradoxon ergibt sich, wenn man die doppelte Bedeutung der Wortes „niemals“ als „nicht in endlich vielen der (immer kleiner werdenden) Schritte“ bedeutet. Erst nach unendlich vielen Schritten wir Achill die Schildkröte erreichen. Der Treffpunkt ist der Grenzwert einer Reihe. Das Paradoxon entsteht, weil das Wort „niemals“ zugleich als „nicht in endlicher Zeit“ verstanden wird. Das ist jedoch ein Fehlschluss. Denn auch die für die einzelnen Wegabschnitte nötigen Zeitabschnitte werden kleiner und die Folge der Zeitabschnitte hat einen endlichen Grenzwert.

Aufgabe: Wenn Achill in einer Zeiteinheit 10 Ellen läuft, wann holt er dann die Schildkröte ein?

Folgen (s_n), bei denen die s_n solche endliche Summen der Gestalt $s_n = a_0 + a_1 + \ldots + a_n$ sind, wobei (a_k) eine weitere (reelle oder komplexe) Zahlenfolge ist, sind uns schon mehrmals begegnet:

(a) Bei der Frage der Erhöhung des Volkseinkommen bei einer Investition von K Euro:
Im Wesentlichen kamen wir auf die geometrische Folge (s_n) mit

$$s_n = 1 + q + q^2 + \ldots + q^n \ ;$$

(b) Bei der Definition der Eulerschen Zahle e:

$$e = \lim_{n\to\infty} E_n,\ E_n = \sum_{k=0}^{n} \frac{1}{k!} \ ;$$

(c) Allgemeiner bei der Definition der exp-Funktion:

$$exp : \mathbb{C} \to \mathbb{C} \quad \text{mit} \quad exp(z) = \lim_{n\to\infty} E_n(z),\ E_n(z) = \sum_{k=0}^{n} \frac{z^k}{k!} \ ;$$

(d) Die Dezimaldarstellung reeller Zahlen

$$x = x_0 + \lim_{n\to\infty} \sum_{k=1}^{n} \frac{z_k}{10^k} \qquad \text{(vgl. §9.11).}$$

Reihen, mit denen wir uns nun zu beschäftigen haben, sind Folgen spezieller Bauart. Man untersucht eine Folge (s_n), in dem man sich anschaut, wie sie sich von Glied zu Glied ändert, betrachtet also die Differenzen:

$$\begin{array}{ll}
a_0 = s_0 & \Longrightarrow\ s_0 = a_0 \\
a_1 = s_1 - s_0 & \Longrightarrow\ s_1 = a_0 + a_1 \\
a_2 = s_2 - s_1 & \Longrightarrow\ s_2 = a_0 + a_1 + a_2 \\
a_3 = s_3 - s_2 & \Longrightarrow\ s_3 = a_0 + a_1 + a_2 + a_3 \\
\vdots & \vdots \\
a_n = s_n - s_{n-1} & \Longrightarrow\ s_n = a_0 + a_1 + a_2 + \ldots + a_n .
\end{array}$$

Diese Vorüberlegungen werden uns zeigen, dass Reihen und Folgen im Prinzip äquivalente Begriffe sind:
Jede Folge lässt sich als Reihe auffassen und jede Reihe ist eine Folge spezieller Bauart. Wir werden das noch präzisieren.

Reihen (insbesondere Potenzreihen und Fourier-Reihen) sind eines der wichtigsten Mittel zur Konstruktion und Darstellung von Funktionen. So werden wir z.B. alle nicht *elementaren* Funktionen (exp, cos, sin, etc.) über Reihen einführen.

10.1 Der Begriff der Reihe, erste Beispiele

10.1.1 Definition

Ist $(a_k)_{k\geq 0}$ eine Folge reeller oder komplexer Zahlen, dann heißt die Folge

$$(s_n) \quad \text{mit} \quad s_n := \sum_{k=0}^{n} a_k$$

die der Folge (a_k) *zugeordnete Reihe*.

Sie wird mit dem Symbol $\sum(a_k)$ oder $\sum a_k$ bezeichnet.

(a_k) heißt das k-te Reihenglied (besser sollte man (a_k) die *Summenfolge* und a_k den k-ten Summanden nennen).

$s_n = \sum\limits_{k=1}^{n} a_k$ heißt die *n-te Partialsummen* der Reihe $\sum a_k$.

k heißt *Summationsindex*. Wie bei endlichen Summen kann er durch jeden anderen (nicht schon mit einer festen Bedeutung belegten) Buchstaben ersetzt werden.
Ist die Folge (s_n) der Partialsummen *konvergent* und der Grenzwert s von (s_n) heißt *Summe* oder *Wert* der Reihe:

$$s = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=0}^{n} a_k =: \sum_{k=0}^{\infty} a_k \ .$$

Eine nicht konvergente Reihe heißt *divergent*.

10.1.2 Bemerkungen

(a) Man beachte, dass in der Literatur das Symbol $\sum a_k$ (gesprochen „Reihe a_k“) meist mit zweierlei Bedeutung verwendet wird:

(1) $\sum(a_k)$ oder $\sum a_k$ (alleinstehend) bedeutet nichts anderes als die Folge der Partialsummen:

$$(s_n), \quad s_n = a_0 + a_1 + a_2 + \ldots + a_n \ ;$$

(2) Die Gleichung $\sum\limits_{k=0}^{\infty} a_k = s$ bedeutet, dass die Reihe $\sum a_k$, also die Folge der Partialsummen, konvergiert und den Grenzwert s hat.

Welche der beiden Bedeutungen gemeint ist, ist meist aus dem Zusammenhang ersichtlich.

Wir wollen zunächst versuchen, die beiden Bedeutungen auseinander zu halten, denn sie bedeuten ja völlig verschiedene Dinge:

(1) einmal eine Folge (die Folge der Partialsummen) und

(2) im Fall der Konvergenz dieser Folge, deren Grenzwert, also eine Zahl.

(b) Auch Folgen allgemeineren Typs, etwa $(a_k)_{k\geq k_0},\ k_0 \in \mathbb{Z}$, kann man eine Reihe zuordnen:

$$(s_n) \quad \text{mit} \quad s_n = a_{k_0} + a_{k_0+1} + \ldots + a_n \ .$$

Offensichtlich ist $s_n = \sum\limits_{k=0}^{n} b_k$ mit $b_k := a_{k_0+k}$.
In vielen Fällen ist $k_0 = 0$ oder $k_0 = 1$.

10.1.3 Beispiele

(a) *Die geometrische Reihe*:
Für $z \in \mathbb{C}$ ist die Reihe $\sum(z^k)_{k\geq 0}$ genau dann konvergent, wenn $|z| < 1$ ist und dann gilt

$$\boxed{\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}}$$

Dass die Reihe $\sum\left(z^k\right)_{k\geq 0}$ für $|z| < 1$ konvergiert und die angegebene Summe hat ist klar wegen

$$\sum_{k=0}^{\infty} z^k = \frac{1-z^{n+1}}{1-z} \qquad \text{für } z \neq 1$$

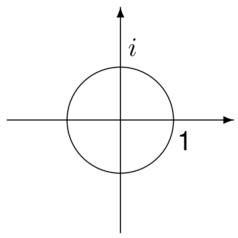
und der Tatsache, dass (z^n) für $|z| < 1$ eine Nullfolge ist.
Bemerkenswert ist, dass die Menge

$$M = \left\{z \in \mathbb{C}\ ;\ \sum(z^n) \text{ ist konvergent}\right\}$$

eine Kreisscheibe ist, nämlich die Einheitskreisscheibe

$$\mathbb{E} := \{z \in \mathbb{C};\ |z| < 1\}\ .$$

(siehe Abb.)
Wir kommen auf diese Phänomene im Zusammenhang mit *Potenzreihen* ausführlich zurück.



Für $|z| \geq 1$ ist die geometrische Reihe divergent, denn dann ist (z^k) oder äquivalent $(|z|^k)$ keine Nullfolge, die Summenfolge (a_k) einer konvergenten Reihe muss aber eine Nullfolge sein:

(b) *Eine notwendige Bedingung für die Konvergenz einer Reihe:*
Ist $\sum(a_k)$ konvergent, so ist (a_k) eine Nullfolge oder äquivalent:
Ist (a_k) keine Nullfolge, so ist $\sum a_k$ nicht konvergent.

Denn ist $(s_n) = \left(\sum\limits_{k=0}^{n} a_k\right)$ konvergent und $s = \lim\limits_{n\to\infty} s_n = \sum\limits_{k=0}^{\infty} a_k$, so folgt aus $a_n = s_n - s_{n-1}$

$$\lim_{n\to\infty} a_n = s - s = 0\ .$$

Dieses leicht nachprüfbare notwendige Kriterium für die Konvergenz einer Reihe ist aber leider nicht hinreichend, wie das folgende Beispiel zeigt.

(c) *Die harmonische Reihe*:

$$(H_n) \quad \text{mit} \quad H_n := \sum_{k=1}^{n} \frac{1}{k}$$

ist divergent.

Dies haben wir auf zwei verschiedene Weisen schon nachgewiesen:
Einmal, indem wir gezeigt haben, dass die Teilfolge H_{2^k} wegen $H_{2^k} > 1 + \frac{k}{2}$ nicht beschränkt ist und weiter mit Hilfe des Cauchy-Kriteriums (vgl. §9.16)

Die Teilsummen der harmonischen Reihe (H_n) wachsen also über jede Grenze, das Wachstum ist aber im „Schneckentempo“, wie die folgende Tabelle zeigt:

n	$(H_n) = \sum\limits_{k=1}^{n} \frac{1}{k}$	$\ln n + \gamma$
10	2.9289682539682539682539682539682539682	2.8798007578955785446245035447667666
10^2	5.1873775176396202608051176756582531579	5.1823858508896242286424949994511308
10^3	7.4854708605503449126565182043339001765	7.4849709438836699126604864541354950
10^4	9.7876060360443822641784779048516053348	9.7875560368777155966784779088198592
10^5	12.090146129863427947363219363504219500	12.090141129871761280696469363504223
10^6	14.392726722865723631381127493188587676	14.392726222865806964714460818188587
10^7	16.695311365859851815399118939540451884	16.695311315859852648732452272872951
10^8	18.997896413853898324417110394223982841	18.997896408853898332750443727557316
10^9	21.300481502347944016685101848908346966	21.300481501847944016768435182241680
10^{10}	23.603066594891989700785593303592711173	23.603066594841989700786426636926044
10^{11}	25.905651687841035384804409758277075381	25.905651687836035384804418091610408
10^{12}	28.208236780830581068822409462961439588	28.208236780830081068822409546294772
10^{13}	30.510821873824176752840401000145803796	30.510821873824126752840401000979137
10^{14}	32.813406966818177436858392455655168004	32.813406966818172436858392455663501
10^{15}	35.115992059812218620876383910347782211	35.115992059812218120876383910347865
10^{16}	37.418577152806263854894375365032228919	37.418577152806263804894375365032229
10^{17}	39.721162245800309493912366819716593951	39.721162245800309488912366819716593
10^{18}	42.023747338794355173430358274400958167	42.023747338794355172930358274400958
10^{19}	44.326332431788400856998349729085322375	44.326332431788400856948349729085322
10^{20}	46.628917524782446540971341183769686583	46.628917524782446540966341183769686

Mit Hilfe der Integralrechnung werden wir zeigen, dass H_n ungefähr wie $\log n$ wächst.
Die obige Zahl γ ist die Euler-Mascheronische-Konstante:

$$\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log n \right) = 0,57721566490...$$

(d) Die Reihe $\sum \left(\dfrac{1}{k(k+1)} \right)_{k \geq 1}$ ist konvergent und es gilt

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \ldots = 1 \; .$$

Hier gelingt es für die n-te Partialsumme einen expliziten Ausdruck anzugeben:
Wegen $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$, ist nämlich

$$s_n = \sum_{k=1}^{n} \frac{1}{k(k+1)} = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \ldots + \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{n+1}$$

und damit

$$\boxed{\lim_{n \to \infty} s_n = \sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1}$$

Der gleiche Trick funktioniert immer, wenn es gelingt, das Reihenglied a_k in der Form

$$a_k = x_k - x_{k-1} \quad \text{oder} \quad a_k = x_k - x_{k+1}$$

darzustellen.
Eine solche Reihe nennt man *teleskopisch*, weil z.B. im Fall $a_k = x_k - x_{k+1}$ für die n-te Partialsumme gilt

$$s_n = (x_1 - x_2) + (x_2 - x_3) + (x_3 - x_4) + \ldots + (x_n - x_{n+1}) = x_1 - x_{n+1}$$

(es bleiben nur x_1 und x_{n+1} übrig, die Glieder dazwischen heben sich weg).

(e) Eine teleskopische Reihe $\sum a_k$ mit $a_k = x_k - x_{k-1}$ bzw. $a_k = x_k - x_{k+1}$ ist genau dann konvergent, wenn $\lim\limits_{n \to \infty} x_n$ existiert und dann gilt

$$\sum_{k=0}^{\infty} a_k = \lim_{n \to \infty} x_{n+1} - x_0 \quad \text{bzw.} \quad \sum_{k=0}^{\infty} a_k = x_0 - \lim_{n \to \infty} x_{n+1} \; .$$

Dazu ein weiteres Beispiel:

$$\begin{aligned} \text{Wegen} \qquad \frac{1}{4k^2 - 1} &= \frac{1}{2}\left(\frac{1}{2k-1} - \frac{1}{2k+1}\right) \qquad && \text{ist} \\ \sum_{k=1}^{n} \frac{1}{4k^2 - 1} &= \frac{1}{2}\left(1 - \frac{1}{2n+1}\right) && \text{also} \\ \sum_{k=1}^{\infty} \frac{1}{4k^2 - 1} &= \frac{1}{2} \end{aligned}$$

(f) Die Reihe $\sum \left(\dfrac{1}{k^2} \right)_{k \geq 1}$ ist konvergent.

Betrachtet man nämlich die n-te Partialsumme $s_n = \sum\limits_{k=1}^{n} \frac{1}{k^2}$, so ist (s_n) streng monoton wachsend.

Wenn wir zeigen können, dass (s_n) nach oben beschränkt ist, ist (s_n) also nach dem Monotonieprinzip konvergent.
Für $k \geq 2$ ist aber $\frac{1}{k^2} \leq \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$ und damit

$$\begin{aligned} s_n = \sum_{k=1}^{n} \frac{1}{k^2} &= 1 + \sum_{k=2}^{n} \frac{1}{k^2} \leq 1 + \sum_{k=2}^{n} \left(\frac{1}{k-1} - \frac{1}{k} \right) \\ &\leq 1 + \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \ldots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 2 - \frac{1}{n} < 3. \end{aligned}$$

Später werden wir sehen, dass für den Grenzwert gilt:

$$s = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k^2} = \frac{\pi^2}{6} \ .$$

Dieses Beispiel zeigt, dass es relativ einfach ist, die Konvergenz der Reihe zu zeigen, aber schwierig sein kann, die Summe der Reihe explizit anzugeben.

Dass Reihen und Folgen im Prinzip das Gleiche sind, wird nochmals festgehalten in folgenden Bemerkungen.

10.1.4 Bemerkungen

Sei $s(\mathbb{K}) = \mathit{Abb}(\mathbb{N}_0, \mathbb{K})$ der $\mathbb{K}$-Vektorraum aller Folgen ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$), dann ist die Abbildung

$$\begin{aligned} \sum \ : \ s(\mathbb{K}) &\to s(\mathbb{K}) \\ (a_k) &\mapsto (s_n) = \left(\sum_{k=0}^{n} a_k \right) \end{aligned}$$

bijektiv.
Gilt nämlich für zwei Folgen (a_k) und (b_k)

$$\sum(a_k) = \sum(b_k) \ ,$$

so bedeutet das, dass für alle n die Partialsummen gleich sind

$$\begin{aligned} a_0 &= b_0 \\ a_0 + a_1 &= b_0 + b_1 \\ a_0 + a_1 + a_2 &= b_0 + b_1 + b_2 \\ &\vdots \\ a_0 + a_1 + \ldots + a_n &= b_0 + b_1 + \ldots + b_n \end{aligned}$$

Hieraus folgt aber $a_0 = b_0, \ a_1 = b_1, \ldots, a_n = b_n$ für alle $n \in \mathbb{N}$.

Die Surjektivität der Abbildung $\sum$ haben wir schon beim Teleskoptrick benutzt.
Ist $(s_0, s_1, s_2, \ldots)$ irgendeine vorgegebene Folge, so definiert man

$$a_0 := s_0; \ a_1 := s_1 - s_0; \ \ldots; \ a_k := s_k - s_{k-1} \qquad (k \geq 1) \ ,$$

dann ist $\sum(a_k) = (s_n)$, $\sum$ ist also surjektiv.

Weil Reihen nur eine ganz besondere Sorte von Folgen sind, gelten folgenden

10.1.5 Rechenregeln für Reihen

Sind (a_k) und (b_k) Folgen in $\mathbb{K}$ und sind die Reihen $\sum a_k$ und $\sum b_k$ konvergent, dann konvergiert für beliebige $\lambda, \mu \in \mathbb{K}$ auch die Reihe $\sum(\lambda a_k + \mu b_k)$ und es gilt

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

Der Beweis folgt sofort durch Übergang zu den Partialsummen aus den Rechenregeln für konvergente Folgen. Man beachte jedoch, dass aus der Konvergenz von $\sum a_k$ und $\sum b_k$ *nicht* die Konvergenz von $\sum a_k b_k$ zu folgen braucht. Nimmt man z.B. $a_0 = 0$ und $a_k = \frac{(-1)^k}{\sqrt{k}}$, $k \geq 1$, dann ist $\sum a_k$ konvergent (wie wir gleich sehen werden), aber

$$\sum a_k^2 = \sum \frac{1}{k}$$

ist die harmonische Reihe, also divergent.

10.2 Konvergenzkriterien für Reihen

Wir wollen uns nun mit *Konvergenzkriterien für Reihen* beschäftigen, die es gestatten aus *Eigenschaften der Reihenglieder* a_k auf die Konvergenz von $\sum a_k$ zu schließen. Wir beginnen mit zwei einfachen Konvergenzkriterien für Reihen mit *reellen* Summanden:

10.2.1 Satz

Ist (a_k) eine Folge nicht negativer reeller Zahlen, dann ist die Reihe $\sum a_k$ genau dann konvergent, wenn die Folge ihrer Partialsummen nach oben beschränkt ist.

Beweis : Wegen der Voraussetzung $a_k \geq 0$ ist die Folge der Partialsummen (s_n), $s_n = \sum\limits_{k=0}^{n} a_k$, monoton wachsend, nach dem Monotonieprinzip also genau dann konvergent, wenn sie nach oben beschränkt ist.

Dieses Prinzip haben wir im Beispiel 10.1.3(f) schon benutzt.

□

Ein weiteres Kriterium betrifft reelle Reihen, bei denen die Reihenglieder abwechselnde Vorzeichen haben, wie z.B. die Reihe $\sum \left(\frac{(-1)^k}{k+1}\right)_{k\geq 0}$.

10.2.2 Leibniz-Kriterium für alternierende Reihen (G.W.Leibniz, 1705)

Ist (a_k) eine monoton fallende Nullfolge reeller Zahlen, dann ist die *alternierende Reihe*

$$\sum \left((-1)^k a_k\right)_{k\geq 0}$$

konvergent.

Für die Partialsummen gilt

$$s_{2n+1} \leq \sum_{k=0}^{\infty}(-1)^k a_k \leq s_{2n} \qquad \text{für alle } n\ ,$$

insbesondere approximiert die Partialsumme

$$s_m = \sum_{k=0}^{m}(-1)^k a_k,\ m \in \mathbb{N}_0\ ,$$

den Reihenwert $s := \sum\limits_{k=0}^{\infty}(-1)^k a_k$ bis auf einen Fehler, der höchstens so groß ist, wie der erste nicht berücksichtigte Summand (die Summanden sind nicht negativ!).

$$(*) \qquad \left| s - \sum_{k=0}^{m}(-1)^k a_k \right| \leq a_{m+1}$$

Beweis : Da die Folge (a_k) eine monoton fallende Nullfolge ist, ist $a_k \geq \inf\{a_n,\ n \in \mathbb{N}\} = 0$ für alle $k \in \mathbb{N}$.

Wir zeigen, dass $([s_{2n+1}, s_{2n}])_{n\in\mathbb{N}_0}$ eine Intervallschachtelung ist, die s erfasst.

Es ist

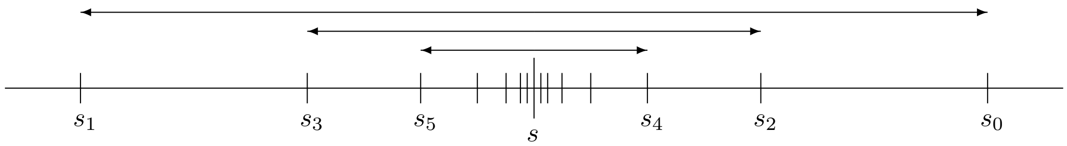
$$\begin{aligned}
s_{2n+1} &= s_{2n} + (-1)^{2n+1}a_{2n+1} = s_{2n} - \underbrace{a_{2n+1}}_{\geq 0} \leq s_{2n},\\
s_{2n+3} &= s_{2n+1} + (\underbrace{a_{2n+2} - a_{2n+3}}_{\geq 0}) \geq s_{2n+1}\\
s_{2n+2} &= s_{2n} - (\underbrace{a_{2n+1} - a_{2n+2}}_{\geq 0}) \leq s_{2n}
\end{aligned}$$

und schließlich

$$s_{2n} - s_{2n+1} = a_{2n+1},\ n \in \mathbb{N}_0\ ,$$

d.h. $([s_{2n+1}, s_{2n}])_{n\in\mathbb{N}_0}$ ist tatsächlich eine Intervallschachtelung:

$$[s_1, s_0] \supset [s_3, s_2] \supset [s_5, s_4] \supset \ldots$$



Wegen $s \in [s_{2n+1}, s_{2n}]$ für alle n und $|s_{m+1} - s_m| = a_{m+1}$ folgt auch die Fehlerabschätzung $(*)$

$$\left| s - \sum_{k=0}^{m} (-1)^k a_k \right| \leq a_{m+1} \; .$$

□

10.2.3 Beispiele zum Leibniz-Kriterium

Nach dem Leibniz-Kriterium sind die Reihen

$$\sum \left(\frac{(-1)^k}{k+1} \right)_{k \geq 0} , \quad \sum \left(\frac{(-1)^k}{2k+1} \right)_{k \geq 0} \quad \text{und} \quad \sum \left(\frac{(-1)^k}{k!} \right)_{k \geq 0}$$

konvergent.
Wie wir später zeigen werden gilt

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots = \log 2, \\ \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} &= 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots = \frac{\pi}{4}, \\ \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} &= \frac{1}{e}. \end{aligned}$$

Die Konvergenz der ersten beiden Reihen gegen die angegebenen Grenzwerte ist nicht besser als die, die Abschätzung $(*)$ aus dem Leibniz-Kriterium angibt. Zur numerischen Berechnung von $\frac{\pi}{4}$ bzw. $\log 2$ sind die Reihen also nicht sonderlich geeignet.

Dagegen konvergiert die Reihe $\sum \frac{(-1)^k}{k!}$ sehr schnell gegen $\frac{1}{e}$ und kann zur nummerischen Berechnung von $\frac{1}{e}$ verwendet werden.

Da eine Reihe $\sum (a_k)$ nichts anderes als die Folge (s_n) ihrer Partialsummen ist, überträgt sich das Cauchy-Kriterium für die Konvergenz von Folgen unmittelbar auf Reihen:

10.2.4 Cauchy'sches Konvergenzkriterium für Reihen

Die Reihe $\sum a_k$ ($a_k \in \mathbb{K}$) konvergiert genau dann, wenn es zu jedem (noch so kleinen) $\varepsilon > 0$ einen Index $N \in \mathbb{N}_0$ gibt, so dass für alle $m > n \geq N$ gilt

$$\left| \sum_{k=n+1}^{m} a_k \right| < \varepsilon \; .$$

Der Beweis ergibt sich sofort durch Anwendung des Cauchy-Kriteriums für Folgen auf die Partialsummenfolge (s_n).
Es ist nämlich (für $m > n$)

$$(*) \qquad s_m - s_n = \sum_{k=n+1}^{m} a_k$$

Bemerkung: Hieraus ergibt sich nochmals, dass die Folge (a_k) der Summanden eine Nullfolge sein muss, wenn die Reihe $\sum a_k$ konvergiert.

Ferner ergibt sich, dass auch die Folge der Reihenreste (r_n)

$$r_n := \sum_{k=n+1}^{\infty} a_k$$

für eine konvergente Reihe $\sum a_k$ ebenfalls eine Nullfolge ist, denn aus $(*)$ erhält man für $m \to \infty$ (bei festem n)

$$s := \lim_{m\to\infty} s_m = s_n + \lim_{m\to\infty} \sum_{k=n+1}^{m} a_k = s_n + r_n$$

und da $(s - s_n)$ eine Nullfolge ist, muss auch (s_n) eine Nullfolge sein. Dies folgt übrigens auch sofort aus der Definition der Konvergenz einer Reihe.

Zu jeder Reihe $\sum a_k$ ($a_k \in \mathbb{K}$) kann man die Reihe $\sum |a_k|$ betrachten. Da $|a_k|$ stets eine nicht negative reelle Zahl ist, ist die „Absolutreihe“ $\sum |a_k|$ genau dann konvergent (vgl. §10.2.1), wenn die Folge $(\sigma_n) = \left(\sum\limits_{k=0}^{n} |a_k|\right)$ ihrer Partialsummen nach oben beschränkt ist.

In welchem Verhältnis stehen nun die Konvergenz von $\sum a_k$ und $\sum |a_k|$?
Bezeichnet man mit

$$s_n := \sum_{k=0}^{n} a_k \qquad \text{bzw.} \qquad \sigma_n := \sum_{k=0}^{n} |a_k|$$

die entstehende Partialsummen und nehmen wir an, dass (σ_n) konvergiert, dann gibt es nach dem Cauchy-Kriterium zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}_0$, so dass für alle $m, n \in \mathbb{N}_0$ mit $m > n \geq N$ gilt

$$|\sigma_m - \sigma_n| = \left|\sum_{k=n+1}^{m} |a_k|\right| = \sum_{k=n+1}^{m} |a_k| < \varepsilon\ .$$

Nach der Dreiecksungleichung ist aber

$$|s_m - s_n| = \left|\sum_{k=n}^{m} a_k\right| \leq \sum_{k=n}^{m} |a_k| < \varepsilon$$

für alle $m, n \in \mathbb{N}_0$ mit $m > n \geq N$, d.h. nach dem Cauchy-Kriterium ist auch die Reihe $\sum a_k$ konvergent.

Wir haben also gezeigt:

$$\boxed{\text{Aus der Konvergenz von } \sum |a_k| \text{ folgt die Konvergenz von } \sum a_k.}$$

Man führt an dieser Stelle die folgende Sprechweise ein:

10.2.5 Definition

Eine Reihe $\sum a_k$ ($a_k \in \mathbb{K}$) heißt *absolut konvergent*, wenn die Reihe $\sum |a_k|$ konvergiert. (Das ist in jedem Fall ($a_k \in \mathbb{R}$ oder $a_k \in \mathbb{C}$) eine Reihe mit nicht negativen reellen Summanden.)

Wir haben also gezeigt:

10.2.6 Satz

Eine absolut konvergent Reihe $\sum a_k$ ist konvergent (im gewöhnlichen Sinne), und es gilt die *Dreiecksungleichung für Reihen*

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

Beweis : Die zusätzliche Behauptung über die Dreiecksungleichung ergibt sich aus

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

durch Grenzübergang $n \to \infty$.

□

10.2.7 Bemerkungen

(a) Wie das einfache Beispiel der nach dem Leibniz-Kriterium konvergente alternierende Reihe

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots$$

zeigt, gilt die Umkehrung vom Satz i.A. nicht, denn die „Absolutreihe" ist die harmonische Reihe.

(b) Für Reihen $\sum a_k$ mit $a_k \in \mathbb{R}_+$stimmen die Begriffe *Konvergenz* und *absolute Konvergenz* überein.

(c) Die *geometrische Reihe* $\sum (z^k)_{k_0}$ ist für $|z| < 1$ absolut konvergent. Sie ist ein Beispiel für eine *Potenzreihe* $\sum a_k z^k$, $a_k \in \mathbb{K}$.
Von solchen Reihen werden wir zeigen (vgl. ???):
Wenn sie konvergieren, dann konvergieren sie sogar absolut (in geeigneten Kreisscheiben).

(d) Die grundsätzliche Bedeutung des Begriffes *absolute Konvergenz* einer Reihe wird nun auch im Zusammenhang mit *Umordnungssätzen für Reihen* deutlich werden.

(e) Eine konvergente, aber nicht absolut konvergente Reihe nennt man auch *bedingt konvergent*. Die obige alternierende harmonische Reihe ist hierfür ein Beispiel.

(f) Eine Folge (a_k) komplexer Zahlen heißt *quadratsummierbar*, wenn $\sum(|a_k|^2)$ konvergiert.
In Übungsaufgabe 4 von Blatt 9 haben wir gezeigt, dass für quadratsummierbare Folgen (a_k) und (b_k) auch die Folgen $(a_k b_k)$ und $(a_k + b_k)$ quadratsummierbar sind. Hieraus folgt, dass die Menge $l^2(\mathbb{N}_0)$ der quadratsummierbaren Folgen eines $\mathbb{C}$-Vektorraums ist, auf welchem durch

$$\langle (a_k), (b_k) \rangle := \sum_{k=0}^{\infty} \bar{a}_k b_k$$

ein Skalarprodukt definiert wird, für welchen die C.S.U. gilt:

$$\left|\sum_{k=0}^{\infty} \bar{a}_k b_k\right| \leq \sum_{k=0}^{\infty} |a_k| \cdot |b_k| \leq \sqrt{\sum_{k=0}^{\infty} |a_k|^2} \cdot \sqrt{\sum_{k=0}^{\infty} |b_k|^2}$$

$l^2(\mathbb{C})$ heißt Hilbertscher Folgenraum. Er ist vollständig, d.h. ein Banach-Raum.

Ein weiteres einfache Konvergenzkriterium beinhaltet

10.2.8 Satz

Sind (a_k) und (b_k) Folgen in $\mathbb{K}$ und konvergiert die Reihe $\sum a_k$ absolut und ist die Folge (b_k) beschränkt, dann konvergiert auch die Reihe $\sum a_k b_k$ absolut.

Beweis : Ist etwa $|b_k| \leq S$ für alle $k \in \mathbb{N}_0$, dann ist

$$\sum_{k=0}^{n} |a_k b_k| \leq \sum_{k=0}^{n} |a_k| \cdot S = S \cdot \sum_{k=0}^{n} |a_k| \leq S \cdot \underbrace{\sum_{k=0}^{\infty} |a_k|}_{\sigma} = S \cdot \sigma = C \ ,$$

somit ist die Folge $\left(\sum\limits_{k=0}^{n} |a_k b_k|\right)$ der Partialsummen von $(\sum |a_k b_k|)$ nach oben beschränkt und damit konvergent.

□

Sehr häufig beweist man die Konvergenz oder Divergenz einer vorgegebene Reihe durch *Vergleich mit Reihen*, deren Konvergenz- bzw. Divergenzverhalten bekannt ist. Eine wichtige Vergleichsreihe ist *die geometrische Reihe* $\sum q^k$, $q \in \mathbb{C}$.

10.2.9 Theorem (Majoranten-Kriterium)

Es sei $\sum a_k$ eine Reihe mit $a_k \in \mathbb{K}$ und (b_k) eine Folge nicht negativer reeller Zahlen, für welche die Reihe $\sum b_k$ konvergiert und es gelte $|a_k| \leq b_k$ für (fast alle) $k \in \mathbb{N}_0$, dann ist auch die Reihe $\sum a_k$ konvergent (sogar absolut), und es gilt

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

Beweis : Nach Voraussetzung und nach der Dreiecksungleichung ist

$$\left|\sum_{k=n+1}^{m} a_k\right| \leq \sum_{k=n+1}^{m} |a_k| \leq \sum_{k=n+1}^{m} |b_k| \ ,$$

woraus nach dem Cauchy-Kriterium alles folgt.

□

Bemerkung: Für die Konvergenzaussage in §10.2.9 genügt es natürlich $|a_k| \leq b_k$ für *fast alle* $k \in \mathbb{N}_0$ zu fordern.

Man nennt die Reihe $\sum b_k$ eine *konvergente Majorante* für die Reihe $\sum a_k$.

Umgekehrt kann man durch Vergleich mit einer bekannten divergenten Reihe gelegentlich auf die *Divergenz* einer vorgegebenen Reihe schließen: Ist etwa $0 \leq d_k \leq c_k$ für fast alle k und divergiert die Reihe $\sum d_k$, dann divergiert auch die Reihe $\sum c_k$.
Man nennt in diesem Fall die Reihe $\sum d_k$ eine *divergente Minorante* für die Reihe $\sum c_k$.

10.2.10 Beispiele

Vom Majorantenkriterium haben wir schon mehrfach Gebrauch gemacht.

(a) Die Reihe $\sum \left(\frac{1}{k^p}\right)_{k\geq 1}$, $p \in \mathbb{N}$, ist konvergent für $p \geq 2$, denn für $p \geq 2$ ist $\frac{1}{k^p} \leq \frac{1}{k^2}$ und $\sum \left(\frac{1}{k^2}\right)$ ist konvergent. Für $p \leq 1$ ist $\sum \frac{1}{k^p}$ sicher divergent, da die harmonische Reihe $(H_n) = \left(\sum\limits_{k=1}^{n} \frac{1}{k}\right)$ eine divergente Minorante ist.

Wie steht es mit der allgemeinen harmonischen Reihe $\sum \left(\frac{1}{k^s}\right)_{k\geq 1}$, $s \in \mathbb{Q}$?
Insbesondere: Was gilt für s mit $1 < s < 2$?

Die Antwort lautet (Verallgemeinerung von (a)):

(b) $\sum \left(\frac{1}{k^s}\right)_{k\geq 1}$, $s \in \mathbb{Q}$, ist genau dann konvergent, wenn $s > 1$ gilt. Wählt man nämlich im Fall $s > 1$ zu $n \in \mathbb{N}$ ein $m \in \mathbb{N}$, so dass $n \leq 2^m - 1$ gilt, so wird

$$\begin{aligned}
s_n = \sum_{k=1}^{n} \frac{1}{k^s} \quad &\leq \quad \sum_{k=1}^{2^m-1} \frac{1}{k^s} \\
&= \quad 1 + \left(\frac{1}{2^s} + \frac{1}{3^s}\right) + \ldots + \left(\frac{1}{(2^{m-1})^s} + \ldots + \frac{1}{(2^m-1)^s}\right) \\
&\leq \quad 1 + \frac{2}{2^s} + \frac{4}{4^s} + \ldots + \frac{2^{m-1}}{(2^{m-1})^s} \\
&= \quad \frac{1-2^{(1-s)m}}{1-2^{1-s}} < \frac{1}{1-2^{1-s}}
\end{aligned}$$

Die Folge (s_n) der Partialsummen ist also nach oben beschränkt und damit konvergent. Ist $s \leq 1$, so ergibt sich wegen $\frac{1}{k^s} \geq \frac{1}{k}$

$$s_n := 1 + \frac{1}{2^s} + \frac{1}{3^s} + \ldots + \frac{1}{n^s} \geq 1 + \frac{1}{2} + \ldots + \frac{1}{n} = H_n \;.$$

Die harmonische Reihe $(H_n) = \left(\sum\limits_{k=1}^{n} \frac{1}{k}\right)$ ist also eine divergente Minorante.

Die sogar für komplexes s mit $Re(s) > 1$ durch

$$\zeta(s) := \sum_{k=1}^{\infty} \frac{1}{k^s}$$

definierte Funktion heißt *Riemannsche Zetafunktion*. Sie spielt für die *Verteilung der Primzahlen* eine *herausragende Rolle*.
Wie wir später zeigen werden, gilt etwa

$$\zeta(2) = \frac{\pi^2}{4}, \qquad \zeta(4) = \frac{\pi^4}{90}, \qquad \zeta(6) = \frac{\pi^6}{945}$$

(L. Euler, 1734).

(c) Ein Trivialbeispiel für Divergenz: (es ist im Beispiel (a) für $s = \frac{1}{2}$ enthalten)
Ist $d_k = \frac{1}{k};\ k \in \mathbb{N}$, und $c_k = \frac{1}{\sqrt{k}}$, dann gilt $d_k \leq c_k$ für alle $k \in \mathbb{N}$ und aus der Divergenz der harmonischen Reihe $\sum \frac{1}{k}$ folgt die Divergenz von $\sum \frac{1}{\sqrt{k}}$.

(d) $\sum \left(\frac{k!}{k^k}\right)_{k\geq 1}$ ist konvergent, dann es ist für $k \geq 2$

$$\frac{k!}{k^k} = \frac{1 \cdot 2 \cdot 3 \cdot \ldots \cdot k}{k \cdot k \cdot \ldots \cdot k} \leq \frac{2}{k^2}$$

und mit $b_k = \frac{2}{k^2}$ ist $\sum b_k$ eine konvergente Majorante.

Häufig kann man die *geometrische Reihe* als konvergente Majorante wählen.

10.2.11 Satz (Quotientenkriterium: d'Alembert, 1768)

Ist $\sum a_k$ eine gegebene Reihe mit $a_k \in \mathbb{K}$. Es gelte $a_k \neq 0$ für fast alle k und es gebe eine reelle Zahl q mit $0 < q < 1$ und $\left|\frac{a_{k+1}}{a_k}\right| \leq q$ für fast alle $k \in \mathbb{N}$.

Dann ist die Reihe $\sum a_k$ (sogar absolut) konvergent. Gilt jedoch $\left|\frac{a_{k+1}}{a_k}\right| \geq 1$ für fast alle k, so ist die Reihe $\sum a_k$ divergent.

Zusatz: (Limesform des Quotientenkriteriums):
Existiert $\alpha := \lim\limits_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right|$ und gilt $\alpha < 1$, dann ist $\sum a_k$ (absolut) konvergent.

Beweis : Für alle $k \geq N$ gelte $a_k \neq 0$ und $|a_{k+1}| \leq q|a_k|$. Insbesondere ist $|a_{N+1}| \leq q|a_N|$.
Dann ist auch

$$|a_{k+2}| \leq q|a_{k+1}| \leq q^2|a_k|$$

und allgemeiner für $m > N$

$$|a_m| = \left|\frac{a_m}{a_{m-1}}\right| \cdot \left|\frac{a_{m-1}}{a_{m-2}}\right| \cdots \left|\frac{a_{N+1}}{a_N}\right| |a_N| \leq q|a_{m-1}|q|a_{m-2}| \ldots q|a_N| = q^{m-N}|a_N| \,.$$

Damit gilt für $m \geq N$

$$|a_m| \leq \frac{|a_N|}{q^N} \cdot q^m = C \cdot q^m \qquad \text{mit } C := \frac{|a_N|}{q^N}$$

und damit ist die Reihe $C \cdot \sum (q^m)_{m\geq N}$ eine konvergente Majorante.
Gilt dagegen $\left|\frac{a_{k+1}}{a_k}\right| \geq 1$ für fast alle k, so kann (a_k) keine Nullfolge sein, $\sum (a_k)$ ist also divergent.

$\square$

Zur Limesform:
Existiert $\alpha := \lim\limits_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right|$ und ist $\alpha < 1$, so ist $\sum a_k$ (sogar absolut) konvergent.

Dies führt man sofort auf den ersten Teil zurück: Es ist $0 < \alpha < 1$ und definiert man $q := \frac{1+\alpha}{2}$, dann gilt

$$0 < \alpha < \frac{1+\alpha}{2} = q < 1$$

und für fast alle k gilt dann

$$\left|\frac{a_{k+1}}{a_k}\right| < q \ ,$$

denn für fast alle k gilt $\frac{a_{k+1}}{a_k} \in U_\varepsilon(\alpha)$, mit $\varepsilon = \frac{1-\alpha}{2} > 0$.

Gilt jedoch $\lim\limits_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = 1$, so liefert die Limesform keine Aussage. Es kann dann sowohl Divergenz als auch Konvergenz vorliegen, wie wir gleich sehen werden.

10.2.12 Beispiele

(a) Für alle $z \in \mathbb{K}$ konvergiert die Reihe $\sum \left(\frac{z^k}{k!}\right)_{k\geq 0}$ absolut.
Für $z = 0$ ist das klar. Für $z \neq 0$ und $a_k = \frac{z^k}{k!}$ erhält man

$$\left|\frac{a_{k+1}}{a_k}\right| = \left|\frac{z^{k+1}\cdot k!}{(k+1)!\cdot z^k}\right| = \frac{|z|}{k+1} \leq \frac{1}{2} \qquad \text{für } \ k \geq 2|z| - 1 \ .$$

Wählt man daher eine natürliche Zahl N mit $N \geq 2|z| - 1$, so gilt für alle $k \geq N$

$$\left|\frac{a_{k+1}}{a_k}\right| = \frac{|z|}{k+1} \leq q := \frac{1}{2} \ ,$$

die Voraussetzung des Quotientenkriteriums wird also erfüllt. Die „Teilreihen“

$$\sum \left(\frac{z^{2k}}{(2k)!}\right)_{k\geq 0} \quad \text{bzw.} \quad \sum \left(\frac{z^{2k+1}}{(2k+1)!}\right)_{k\geq 0}$$

sind ebenfalls für alle $z \in \mathbb{K}$ absolut konvergent.
Durch diese Reihe werden *wichtige Funktionen* definiert:
Im Fall $\mathbb{K} = \mathbb{C}$ durch

$$\begin{aligned} exp_{\mathbb{C}} : \mathbb{C} &\to \mathbb{C} \\ z &\mapsto \sum_{k=0}^{\infty} \frac{z^k}{k!} \end{aligned}$$

die *komplexe Exponentialfunktion*, durch

$$\begin{aligned} exp_{\mathbb{R}} : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \sum_{k=0}^{\infty} \frac{x^k}{k!} \end{aligned}$$

die *reelle Exponentialfunktion*.
Offensichtlich gilt für $x \in \mathbb{R}$

$$exp_{\mathbb{C}}(x) = exp_{\mathbb{R}}(x) \ , \text{ oder } \ exp_{\mathbb{C}}|\mathbb{R} = exp_{\mathbb{R}} \ .$$

Wir lassen also den Index $\mathbb{C}$ bzw. $\mathbb{R}$ in Zukunft weg.

Trotz der gleichen Definition und vieler gemeinsamer Eigenschaften (Stetigkeit, Differenzierbarkeit) unterscheiden sich die komplexe und reelle Exponentialfunktion wesentlich:
Die reelle Exponentialfunktion ist streng monoton wachsend (und damit injektiv), die komplexe Exponentialfunktion besitzt die *rein imaginäre* –und damit im Reellen nicht sichtbare– Periode $2\pi i$, d.h. es gilt für alle $z \in \mathbb{C}$

$$\boxed{exp(z + 2\pi i) = exp(z)}$$

Hierauf kommen wir ausführlich zurück: Durch $\sum \frac{z^{2k}}{(2k)!}$ und $\sum \frac{z^{2k+1}}{(2k+1)!}$ werden *cosh* und *sinh*, sogar für komplexe, insbesondere also für reelle Argumente, definiert.

(b) Die Reihe $\sum \left(\frac{k!}{k^k} z^k\right)_{k\geq 1}$ ist absolut konvergent für $|z| < e$. Für $z = 0$ ist die Konvergenz klar. Sei also $z \neq 0$ und $a_k := \frac{k!}{k^k} z^k$, dann ist

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

$\alpha := \lim\limits_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right|$ existiert und es gilt $\alpha = \frac{|z|}{e}$, also ist die Reihe konvergent für alle $z \in \mathbb{K}$ mit $|z| < e$.

Komplizierter in der Anwendung ist das weitere hinreichende Kriterium für die absolute Konvergenz einer Reihe, das *Wurzelkriterium*:

10.2.13 Satz (Wurzelkriterium: Cauchy, 1821)

Gibt es eine reelle Zahl q mit $0 < q < 1$, so dass für fast alle Glieder der Folge (a_k) gilt $\sqrt[k]{|a_k|} \leq q$, dann ist die Reihe $\sum a_k$ (sogar absolut) konvergent.

Gilt dagegen für fast alle k $\sqrt[k]{|a_k|} \geq 1$ oder gilt auch nur für unendlich viele k $\sqrt[k]{|a_k|} \geq 1$, dann ist $\sum a_k$ divergent.

Zusatz (Limesform):
Existiert $\alpha := \lim\limits_{k\to\infty} \sqrt[k]{|a_k|}$ und gilt $\alpha < 1$, dann ist $\sum a_k$ (sogar absolut) konvergent, im Fall $\alpha > 1$ ist $\sum a_k$ divergent, im Fall $\alpha = 1$ ist keine Entscheidung möglich.

Beweis : Der Beweis folgt hier direkt aus dem Vergleich mit der geometrischen Reihe:
Gilt etwa $\sqrt[k]{|a_k|} \leq q$ für alle $k \geq N$, so folgt für alle $k \geq N$ $|a_k| \leq q^k$.

Gilt jedoch für fast alle k, dass $\sqrt[k]{|a_k|} \geq 1$ ist, dann ist (a_k) keine Nullfolge.

□

Zur Limesform:
Existiert $\alpha := \lim\limits_{k\to\infty} \sqrt[k]{|a_k|}$ und ist $\alpha < 1$, so ist für fast alle k

$$\sqrt[k]{|a_k|} \leq q := \alpha + \frac{1-\alpha}{2} = \frac{1+\alpha}{2} < 1$$

und man ist in den Voraussetzungen des ersten Teiles des Satzes.

Ist $\alpha > 1$, so ist (a_k) keine Nullfolge.

Im Fall $\alpha = 1$ kann Konvergenz oder Divergenz vorliegen, wie die folgenden Beispiele zeigen.

10.2.14 Beispiele

(a) $\sum\left(\left(\frac{k}{2k+1}\right)^k\right)_{k\geq 1}$ ist konvergent, denn $\lim\limits_{k\to\infty}\sqrt[k]{|a_k|} = \lim\limits_{k\to\infty}\frac{k}{k+1} = \frac{1}{2}$.

(b) $\sum\left(\frac{(k+1)^{k^2}}{k^{k^2}2^k}\right)_{k\geq 1}$ ist divergent, denn $\lim\limits_{k\to\infty}\sqrt[k]{|a_k|} = \lim\limits_{k\to\infty}\frac{1}{2}\left(1+\frac{1}{k}\right)^k = \frac{1}{2}e > 1$.

(c) Für alle $x \geq 0$ ist $\sum\left(\frac{x^k}{k^k}\right)_{k\geq 1}$ konvergent, denn es ist $\sqrt[k]{\frac{x^k}{k^k}} = \frac{x}{k}$ und $\lim\limits_{k\to\infty}\frac{x}{k} = 0$, also sicher $\frac{x}{k} \leq \frac{1}{2}$ für fast alle k.

(d) Für die Reihe $\sum\frac{1}{k^2}$ und $\sum\frac{1}{k}$ liefert sowohl das Quotientenkriterium als auch das Wurzelkriterium, dass die Quotienten

$$\frac{\frac{1}{(k+1)^2}}{\frac{1}{k^2}} = \frac{k^2}{(k+1)^2} \quad \text{bzw.} \quad \frac{\frac{1}{k+1}}{\frac{1}{k}} = \frac{k}{k+1} = 1 - \frac{1}{k+1} < 1$$

kleiner 1 sind, der limes ist aber jeweils 1, es gibt aber keinen *echten Bruch* q $(0 < q < 1)$ mit $\left|\frac{a_{k+1}}{a_k}\right| \leq q$ für fast alle k.

Die erste Reihe ist konvergent, die zweite ist divergent.

(e) Obwohl das Quotientenkriterium und das Wurzelkriterium beide auf die Konvergenz der geometrischen Reihe beruhen, liefert das Wurzelkriterium in manchen Fällen noch eine Entscheidung, wenn das Quotientenkriterium versagt:

Setzt man $a_k = (\frac{1}{2})^{k+(-1)^k}$ für $k \in \mathbb{N}_0$ und betrachte die Reihe

$$\sum a_k = \frac{1}{2} + \frac{1}{1} + \frac{1}{8} + \frac{1}{4} + \frac{1}{32} + \frac{1}{16} + \dots$$

so ist

$$\frac{a_{k+1}}{a_k} = \begin{cases} 2, & \text{falls } k \text{ gerade,} \\ \frac{1}{8}, & \text{falls } k \text{ ungerade,} \end{cases}$$

das Quotientenkriterium ist also nicht anwendbar. Beachtet man jedoch

$$a_k = \frac{1}{2^{k+(-1)^k}} \leq 2\left(\frac{1}{2}\right)^k \quad \text{für } \; k \geq 0 \; ,$$

so sieht man, dass das Wurzelkriterium anwendbar ist.

Wir beschließen die Reihe der Konvergenzkriterien mit zwei nützlichen Kriterien, die nach Abel (N.H.Abel, 1801-1829) und Dirichlet (P.G.L. Dirichlet, 1805-1859) benannt sind und auf dem Prinzip der *Abelschen partiellen Summation* beruhen.

10.2.15 Satz (Abelsche partielle Summation: N.H.Abel, 1826)

Seien (a_k) und (b_k) zwei Folgen in $\mathbb{K}$.
Setzt man $A_m := \sum\limits_{k=0}^{m} a_k$, $m \in \mathbb{N}_0$, so ist für alle $n \in \mathbb{N}_0$ (mit $A_{-1} := 0$)

$$\boxed{\sum_{k=0}^{n} a_k b_k = \sum_{k=0}^{n-1} A_k(b_k - b_{k+1}) + A_n b_n}$$

Beweis : Mit der Fortsetzung $A_{-1} = 0$ gilt

$$\begin{aligned}
\sum_{k=0}^{n} a_k b_k &= \sum_{k=0}^{n}(A_k - A_{k-1})b_k = \sum_{k=0}^{n} A_k b_k - \sum_{k=0}^{n} A_{k-1} b_k \\
&= \sum_{k=0}^{n} A_k b_k - \sum_{k=0}^{n-1} A_k b_{k+1} \\
&= \sum_{k=0}^{n-1} A_k(b_k - b_{k+1}) + A_n b_n \ .
\end{aligned}$$

□

Abelsche partielle Summation kann man häufig dann anwenden, wenn das allgemeine Reihenglied c_k die Gestalt $c_k = a_k b_k$ mit geeigneten Zahlen a_k bzw. b_k hat.

10.2.16 Satz (Abelsches Konvergenzkriterium)

Ist die Reihe $\sum a_k$ konvergent, $a_k \in \mathbb{C}$ und ist (b_k) eine monotone und beschränkte Folge *reeller* Zahlen, dann ist auch die Reihe $\sum a_k b_k$ konvergent.

Beweis : Nach §10.2.15 ist

$$\sum_{k=0}^{n} a_k b_k = \sum_{k=0}^{n-1} A_k(b_k - b_{k+1}) + A_n b_n \ .$$

Da die Folge $(A_n) = \left(\sum\limits_{k=0}^{n} a_k\right)$ der Partialsummen von $\sum a_k$ konvergiert und $\sum(b_k - b_{k+1})$ wegen der Monotonie der Folge (b_k) sogar absolut konvergiert, konvergiert nach §10.2.8 auch die Reihe

$$\sum A_k(b_k - b_{k+1}) \ .$$

Wegen der Konvergenz der Folge $(A_n b_n)$ ist die Folge $\left(\sum\limits_{k=0}^{n} a_k b_k\right)$ dann ebenfalls konvergent.

□

10.2.17 Satz (Konvergenzkriterium von Dirichlet)

Ist (a_k) eine Folge komplexer Zahlen für welche die Folge $(A_n) = \left(\sum\limits_{k=0}^{n} a_k\right)$ der Partialsummen beschränkt ist und ist (b_k) eine monotone Nullfolge reeller Zahlen, dann ist die Reihe $\sum a_k b_k$ konvergent.

Beweis : Nach der Formel für die abelsche partielle Summation ist für $n \in \mathbb{N}_0$

$$\sum_{k=0}^{n} a_k b_k = \sum_{k=0}^{n-1} A_k(b_k - b_{k+1}) + A_n b_n \ .$$

Die Folge $(A_n b_n)$ konvergiert gegen Null, und die Reihe $\sum A_k(b_k - b_{k+1})$ konvergiert (wieder nach $\S$10.2.8). Daher ist $\sum a_k b_k$ konvergent.

□

Beispiel:
Mit Hilfe des Dirichlet-Kriteriums erhält man einen weiteren Beweis für das Leibnizkriterium (vgl. $\S$10.2.2)
Ist $a_k = (-1)^k$, dann bildet $\sum a_k = \sum (-1)^k$ beschränkte Partialsummen. Ist dann (b_k) eine monotone Nullfolge, so ist $\sum (-1)^k b_k$ konvergent. (Beachte: Die Rolle der a_k in $\S$10.2.2 übernehmen jetzt die b_k).

Hauptanwendung des Dirichlet-Kriterium ist eine Anwendung auf sog. Dirichlet-Reihe; das sind Reihen vom Typ

$$\sum \frac{a_k}{k^s}, \quad a_k \in \mathbb{C},\ s \in \mathbb{C},$$

z.B. ist $\sum \frac{1}{k^s}$, $s \in \mathbb{C},\ Re(s) > 1$, (hier ist $a_k = 1$ für alle k) eine berühmte Dirichlet-Reihe. Durch sie wird die *Riemannsche Zetafunktion* (für $Re(s) > 1$) definiert.

10.2.18 Bemerkung

Eine wichtige Technik zum Untersuchen der Konvergenz bzw. Divergenz einer Reihe ist der Vergleich mit Integralen.

So kann man etwa die Partialsumme $\sum\limits_{k=1}^{n} \frac{1}{k^s}$ der Reihe $\sum \left(\frac{1}{k^s}\right)$ mit dem Integral $\int\limits_1^n \frac{1}{x^s}\ dx$ vergleichen. Wir kommen hierauf später (im Zusammenhang mit der Integralrechnung) zurück.

Es gibt zahlreiche weitere hinreichende Konvergenzkriterien für Reihen, wie z.B. das *Verdichtungskriterium* (vgl. Übungsblatt 10 Aufgabe 67). Wenn man gar nicht mehr weiterkommt, konsultiere man das 1921 in erster Auflage und 1996 in 6. Auflage erschienene meisterhafte Werk von *Konrad Knopp*: Theorie und Anwendung der unendlichen Reihen, Springer-Verlag, HD.

10.3 Exkurs: Die g-al Darstellung reeller Zahlen

In diesem Exkurs beschäftigen wir uns mit der Darstellung reeller Zahlen z.B. als *Dezimalzahlen*. Die Grundzahl 10 ist aber mathematisch nicht ausgezeichnet, wir wählen daher eine beliebige natürliche Zahl $g \in \mathbb{N},\ g \geq 2$, als Grundzahl. Bei der Darstellung reeller Zahlen bzw. ihrer rationalen Approximationen im Computer sind die Grundzahlen $g = 2,\ g = 8$ bzw. $g = 16$ von Bedeutung.

Für die *Darstellung reeller Zahlen im Computer* vergleiche man die Ausführungen von O.Forster in Analysis 1, (6.Aufl., S.49-51).

Zunächst einige Bemerkungen über natürliche Zahlen und ihre Darstellungen in verschiedenen Positionssystemen (Stellenwertsystemen).

Bei der Beschäftigung mit Zahlen müsste man eigentlich unterscheiden zwischen dem „Zahlwort", z.B. „Neun" und dem zugehörigen „Zahlzeichen", z.B. 9 oder IX (im römischen Zahlensystem). $\sqrt{2}$ ist die Bezeichnung für die eindeutig bestimmte positive reelle Zahl x, für die $x^2 = 2$ gilt.

Im täglichen Leben verwendet man das Dezimalsystem, ein Stellenwertsystem mit der Grundzahl $g = 10$. Man kommt dabei mit 10 Zahlzeichen (Ziffern) aus: $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$.

Die Dezimalzahl 444 ist eine Abkürzung für $4 \cdot 10^2 + 4 \cdot 10^1 + 4 \cdot 10^0$.
Die Ziffer 4 hat also im Zahlzeichen 444 einen unterschiedlichen Stellenwert.
Wie aus der 6.Klasse (Lehrplan Ba-Wü) bekannt, kann man natürliche Zahlen z.B. auch mit Hilfe von Zweierpotenzen schreiben (Grundzahl $g = 2$). So ist z.B.

$$25 = 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = (11001)_2 \ .$$

Es gilt aber auch etwa

$$25 = 2 \cdot 3^2 + 2 \cdot 3^1 + 1 \cdot 3^0 = (221)_3 \ .$$

Zum Verdeutlichen um welche Grundzahl es sich bei der Darstellung handelt, schreibt man die Grundzahl als Index an die Ziffernfolge.
Bei der Grundzahl $g = 2$ braucht man nur die Ziffern 0 und 1 (häufig wird 0 und L verwendet). Ist die Grundzahl $g \leq 10$, so benötigt man nur die Ziffern $0, 1, \ldots, g - 1$. Ist die Grundzahl g jedoch > 10, so müssen, um Zweideutigkeit zu vermeiden, neue Ziffern eingeführt werden.
Für die Grundzahl $g = 12$ (Duodezimalsystem) verwendet man z.B. für die Dezimalzahl 10 die Bezeichnung A und für die Dezimal Zahl 11 die Bezeichnung B. Dann ist z.B.

$$275 = 1 \cdot 12^2 + 10 \cdot 12 + 11 \cdot 12^0 = (1AB)_{12} \ .$$

Würde man für 10 und 11 im Dezimalsystem keine neuen Ziffern einführen so könnte das Zahlwort $(101)_{12}$

$$\text{entweder } \ 1 \cdot 12^2 + 0 \cdot 12^1 + 1 \cdot 12^0 = 145$$
$$\text{oder } \ 10 \cdot 12^1 + 1 \cdot 12^0 = 121$$

bedeuten.
Die Grundzahl 10 ist höchstens aufgrund historischer, kultureller oder praktischer Gründe ausgezeichnet (jeder Mensch hat nun mal 10 Finger, die man in allen Kulturen und schon immer zum Zählen zu Hilfe genommen hat). Von mathematischem Gesichtspunkt ist die Grundzahl $g = 10$ aber nicht die beste und zweckmäßigste.
Mit vollständiger Induktion lässt sich leicht folgender Satz beweisen:

10.3.1 Satz (g-al-Darstellung natürlicher Zahlen)

Jedes $a \in \mathbb{N}$ lässt sich bei gegebenem $g \in \mathbb{N}$, $g \geq 2$, auf genau eine Weise in der Form

$$a = x_n g^n + x_{n-1} g^{n-1} + \ldots + x_1 g + x_0 g^0$$

mit $n \in \mathbb{N}_0$, $x_n > 0$ und $x_j \in \{0, 1, \ldots, g - 1\}$ für $0 \leq j \leq n$ darstellen.

Kurzschreibweise: $a = (x_n x_{n-1} \dots x_0)_g$

10.3.2 Beispiele

(a) $24 = 2 \cdot 3^2 + 2 \cdot 3^1 + 0 \cdot 3^0 = (220)_3$

(b) $36 = 1 \cdot 5^2 + 2 \cdot 5^1 + 1 \cdot 5^0 = (121)_5$

(c) $45054 = (AFFE)_{16}$
Wenn man im Hexadezimalsystem ($g = 16$) für die Dezimalzahlen

10	11	12	13	14	15	die Bezeichnungen
A	B	C	D	E	F	

verwendet.
Steht an den Ziffernfolge kein Index, so ist in der Regel das Zahlwort mit Hilfe der Basis $g = 10$ geschrieben.

10.3.3 Bemerkung

g-al-Ziffern erhält man einfach durch fortgesetzte Division mit Rest durch die Grundzahl g:

$$\begin{aligned} 24 &= 2 \cdot 12 + 0 \\ 12 &= 2 \cdot 6 + 0 \\ 6 &= 2 \cdot 3 + 0 \\ 3 &= 2 \cdot 1 + 1 \\ 1 &= 2 \cdot 0 + 1 \end{aligned}$$

und damit

$$24 = (11000)_2\ .$$

Man beachte die umgekehrten Reihenfolgen in der Ziffernfolge.

$$\begin{aligned} 45054 &= 16 \cdot 2815 + 14 \\ 2815 &= 16 \cdot 175 + 15 \\ 175 &= 16 \cdot 10 + 15 \\ 10 &= 16 \cdot 0 + 10, \text{ also} \\ (45054)_{10} &= (AFFE)_{16}. \end{aligned}$$

Die Darstellung natürlicher Zahlen mit einer beliebigen Grundzahl $g \in \mathbb{N},\ g \geq 2$, wollen wir nun auf beliebige reelle Zahlen verallgemeinern.

Dabei werden wir unsere Kenntnisse über *Reihen* entscheidend benutzen. Die auftretende konvergente Majorante ist dabei die geometrische Reihe. Bekanntlich ist

$$0,4999\ldots = 0,4\bar{9}$$

eine Abkürzung für die Reihe

$$\begin{aligned}\frac{4}{10}+\frac{9}{10^2}+\frac{9}{10^3}+\frac{9}{10^4}+\cdots &= \frac{4}{10}+\frac{9}{10^2}\left(1+\frac{1}{10}+\frac{1}{100}+\cdots\right)= \\ &= \frac{4}{10}+\frac{9}{10^2}\cdot\frac{1}{1-\frac{1}{10}} \\ &= \frac{4}{10}+\frac{9}{10^2}\cdot\frac{10}{9} \\ &= \frac{4}{10}+\frac{1}{10}=\frac{5}{10}=\frac{1}{2}\end{aligned}$$

Durch die Reihe wird also die rationale Zahl $\frac{1}{2}$ dargestellt. Andererseits ist aber auch

$$\frac{1}{2}=\frac{5}{10}=0,5=0,5000\ldots\ .$$

Man hat also für die rationale Zahl $\frac{1}{2}$ zwei Darstellungen:

$$\frac{1}{2}=0,5=0,500\ldots=0,4999\ldots=0,4\bar{9}\ .$$

Um Eindeutigkeit zu erreichen, werden wir uns im Folgenden für die Darstellung entscheiden (Grundzahl $g=10$), in der „Neunerperioden“ ausgeschlossen sind (analog für andere Grundzahlen).

Ist nun x eine beliebige reelle Zahl ≥ 0, dann gilt

$$x=[x]+r,\quad \text{wobei}\quad 0\leq r<1\quad \text{und}\quad [x]\in\mathbb{N}_0$$

gilt.
Für den „Rest“ r geben wir im Folgenden eine Intervallschachtelung an, die r erfasst.

10.3.4 Theorem (g-al-Darstellung)

Es sei $r\in\mathbb{R},\ 0\leq r<1$, und $g\in\mathbb{N},\ g\geq 2$, dann gilt

(a) Es gibt genau eine Folge (z_j) von „Ziffern“

$$z_j\in Z_g:=\{0,1,\ldots,g-1\}\ ,$$

so dass für jedes $n\in\mathbb{N}$ gilt

$$(*)\qquad \sum_{j=1}^{n}\frac{z_j}{g^j}\leq r<\sum_{j=1}^{n}\frac{z_j}{g^j}+\frac{1}{g^n}\ .$$

Insbesondere ist

$$r=\lim_{n\to\infty}\sum_{j=1}^{n}\frac{z_j}{g^j}=\sum_{j=1}^{\infty}\frac{z_j}{g^j}\ .$$

(b) Für jede Folge (z_j) mit der Eigenschaft $(*)$ gilt *nicht* für fast alle $j\in\mathbb{N}$

$$z_j=g-1\ .$$

(c) Für jede Folge (z_j) mit $z_j\in Z_g$ für alle j und für die nicht $z_j=g-1$ für fast alle j gilt, gibt es genau eine reelle Zahl r mit $(*)$ und $0\leq r<1$.

Vor dem Beweis überlegen wir uns, wie man im Fall der Grundzahl $g = 10$ die Ziffern der Zahl

$$r := 0,1234567\ldots$$

durch r ausdrücken kann.

Es ist $10r = 1,23\ldots$,	also $[10r] = 1$.	
Es ist $100r = 12,3\ldots$,	also $[100r] = 12$	und damit $[100r] - 10[10r] = 2$,
Es ist $1000r = 123,4\ldots$,	also $[1000r] = 123$	und damit $[1000r] - 10[100r] = 3$,

entsprechend ist $4 = [10^4 r] - 10[10^3 r]$, etc.

Beweis des Satzes:

(a) Der folgende Algorithmus (g-al-Algorithmus) ist eine Verallgemeinerung der Division mit Rest. Wir definieren

$$z_0 := 0,\ r_0 := r \ \text{ und } \ z_j := [r_{j-1}g],\ r_j := r_{j-1}g - z_j$$

für $j \geq 1$ und für $n \in \mathbb{N}$

$$a_n := \sum_{j=1}^{n} \frac{a_j}{g^j} \ \text{ und } \ b_n := a_n + \frac{1}{g^n} \ .$$

Mittels Induktion folgt dann

$(**)$
$$z_n = [g^n r] - g[g^{n-1} r]$$

und

$(***)$
$$[g^n r] = g^n a_n$$

für jedes $n \in \mathbb{N}$ und damit

$$g^n a_n = [g^n r] \leq g^n r < [g^n r] + 1 = g^n a_n + 1 = g^n b_n$$

und damit für jedes $n \in \mathbb{N}$

$$a_n \leq r < b_n \ .$$

Das ist die Behauptung $(*)$ von (a).

Man beachte dabei, dass für jedes reelle y nach Definition der Gauß-Klammer gilt

$$[y] \leq y < [y] + 1 \ .$$

Zur Eindeutigkeit der Folge (z_n):
Ist (z_n) eine Folge, für die $(*)$ gilt, so folgt mittels $(***)$

$$z_n = g^n(a_n - a_{n-1}) = g^n a_n - g(g^{n-1}a_{n-1}) = [g^n r] - g[g^{n-1} r] \ ,$$

z_n ist also gleich der durch $(**)$ eindeutig bestimmten Zahl.

Zu zeigen bleibt: $z_n \in Z_g$.
Für $n = 1$ gilt

$$0 \leq [gr] = z_1 < gr < g, \ \text{ also } \ 0 \leq z_1 < g \ .$$

Für $n > 1$ benutzen wir $r < b_{n-1}$ und erhalten mit $(**)$ und $(***)$

$$0 \leq [g^n r] < g^n r = g(g^{n-1} r) < g(g^{n-1} b_{n-1}) = g(g^{n-1} a_{n-1} + 1) = g[g^{n-1} r] + g \ .$$

Hieraus folgt aber

$$0 \leq [g^n r] - g[g^{n-1} r] < g,$$

also $0 \leq z_n < g$ für alle $n \in \mathbb{N}$.
Als Differenz ganzer Zahlen ist z_n eine ganze Zahl und wegen $0 \leq z_n < g$ gilt daher $z_n \in Z_g$.

(b) ist eine unmittelbare Folgerung aus $(*)$. Gäbe es ein $N \in \mathbb{N}$, so dass für alle $n > N$ gilt $z_n = g - 1$, dann gilt für $n > N \quad b_n = b_N$, also $a_n \leq x < b_N$.

Wählt man n so groß, dass $n > N$ und $\frac{1}{g^n} < b_N - r$ gilt, so folgt

$$r < b_N - \frac{1}{g^n} = b_n - \frac{1}{g^n} = a_n \ ,$$

also erhalten wir den Widerspruch $r < a_n$.

(c) Ist (z_j) eine Folge mit $z_j \in Z_g$ und gilt nicht für fast alle j $z_j = g - 1$, dann gilt für

$$a_n := \sum_{j=1}^{n} \frac{z_j}{g^j} \quad \text{und} \quad b_n := a_n + \frac{1}{g^n}$$

$a_n \leq a_{n+1} < b_{n+1}$ für alle $n \in \mathbb{N}$.
Ferner ist

$$b_{n+1} = \sum_{j=1}^{n+1} \frac{a_j}{g^j} + \frac{1}{g^{n+1}} = \sum_{j=1}^{n} \frac{a_j}{g^j} + \frac{(a_{n+1}+1)}{g^{n+1}} \leq \sum_{j=1}^{n} \frac{a_j}{g^j} + g \cdot \frac{1}{g^{n+1}} = b_n.$$

Also gilt $a_n \leq a_{n+1} < b_{n+1} \leq b_n$ für alle $n \in \mathbb{N}$ und wegen $b_n - a_n = \frac{1}{g^n}$ gilt $\lim\limits_{n\to\infty}(b_n - a_n) = 0$.

$([a_n, b_n])_{n\in\mathbb{N}}$ ist also eine Intervallschachtelung, die genau eine reelle Zahl r erfasst: $a_n \leq r \leq b_n$ für alle $n \in \mathbb{N}$, also

$$a_1 \leq a_n \leq a_{n+1} \leq r \leq b_{n+1} \leq b_n \leq b_1 \ .$$

Es gilt $0 \leq \frac{z_j}{g} = a_1$ und $b_1 = \frac{z_1}{g} + \frac{1}{g} \leq 1$. Da nicht $z_n = g - 1$ für fast alle n gilt, gibt es ein $n \in \mathbb{N}$ mit $z_{n+1} < g$. Damit gilt sogar $b_{n+1} < b_n$ und daraus folgt

$$0 \leq r < 1 \ .$$

□

10.3.5 Bemerkungen

(a) Die Darstellung $r = \sum\limits_{j=1}^{\infty} \frac{z_j}{g^j}$ heißt *g-al-Darstellung* von r.

Ist $z_n = 0$ für alle $n > n_0$ und $z_{n_0} \neq 0$, dann nennt man r auch eine endliche, n_0-stellige g-al-Zahl, andernfalls heißt r eine unendliche g-al-Zahl. Im ersten Fall sagt man auch, dass r eine (mit Null) *abbrechende* g-al-Darstellung besitzt, im zweiten Fall spricht man auch von einer nicht abbrechenden Darstellung.
Gibt es $l \in \mathbb{N}$ und $p \in \mathbb{N}$, so dass $z_{j+p} = z_j$ für alle $j \geq l$ gilt, dann nennt man die g-al-Darstellung *periodisch*.

(b) Ist $r = \frac{a}{b}$, $a, b \in \mathbb{N}$, $a < b$, eine rationale Zahl, so gehören die Reste r_j zu den b verschiedenen Brüchen $\frac{0}{b}, \frac{1}{b}, \cdots, \frac{b-1}{b}$ und müssen sich also wiederholen. Wiederholen sich aber die Reste, dann wiederholen sich auch die Ziffern, also ist die g-al-Entwicklung einer rationalen Zahl *periodisch*.

Als Übungsaufgabe überlege man sich, dass umgekehrt jede Zahl $r \in [0,1[$ mit einer periodischen g-al-Entwicklung rational ist.

(c) Ist $x \in \mathbb{R}$, $x \geq 0$ und $x = [x] + r$ mit $0 \leq r < 1$ und $r = 0, z_1 z_2 z_3 \ldots = \frac{z_1}{g} + \frac{z_2}{g^2} + \ldots$ die g-al-Darstellung von r und

$$[x] = x_n g^n + x_{n-1} g^{n-1} + \ldots + x_0 = (x_n x_{n-1} \ldots x_0)_g \ ,$$

so erhält man für x die g-al-Darstellung

$$x = x_n g^n + x_{n-1} g^{n-1} + \ldots + x_0 + \frac{z_1}{g} + \frac{z_2}{g^2} + \ldots$$
$$x = (x_n x_{n-1} \ldots x_0, z_1 z_2 z_3 \ldots)_g \ .$$

Ist $x < 0$, so entwickle man $y := -x$.

Jede reelle Zahl besitzt also eine (eindeutige) g-al-Darstellung und umgekehrt stellt ($n \in \mathbb{N}_0,\ g \in \mathbb{N},\ g \geq 2$)

$$x_n g^n + \ldots + x_0 + \frac{z_1}{g} + \frac{z_2}{g^2} + \frac{z_3}{g^3} + \ldots$$

mit $x_j, z_j \in \mathbb{Z}_g$, wobei *nicht* $z_n = g - 1$ für fast alle n gilt, eine (eindeutig bestimmte) reelle Zahl dar.

Man hätte daher von vornherein die Menge der reellen Zahlen etwa als die *Menge aller Dezimalzahlen* definieren können. Dabei muss man allerdings mit dem Reihenbegriff vertraut sein. Dabei stößt man aber gleich auf Probleme: Ist z.B. $x := 3,14159265358\ldots$ und $y := 2,71828182845\ldots$
Was ist dann $x + y$ oder gar $x \cdot y$?
Genauer: Wie erhält man die g-al Ziffern von $x + y$ bzw $x \cdot y$ aus denen von x und y?

Was mit dem Satz über die g-al-Entwicklung einer reellen Zahl neu bewiesen wurde, ist die Tatsache, dass $\mathbb{Q}$ *dicht in* $\mathbb{R}$ ist, d.h. jede reelle Zahl ist Grenzwert einer Folge von rationalen Zahlen, denn die a_n aus Satz ??? sind alle rational und es gilt $r = \lim\limits_{n\to\infty} a_n$.

Jede reelle Zahl lässt sich also mit beliebiger Genauigkeit durch rationale Zahlen approximieren. Es gibt jedoch auch andere (bessere) Approximationen reeller Zahlen durch rationale Zahlen, nämlich durch sog. *Kettenbrüche*.
Diese Approximationen beruhen auf der *Kettenbruchentwicklung* einer reellen Zahl, das ist ein Algorithmus, der den gewöhnlichen euklidischen Algorithmus verallgemeinert.
Es sind z.B. $\frac{22}{7}$ oder $\frac{333}{106}$ oder $\frac{355}{113}$ (Näherungs-)Kettenbrüche für die Kreiszahl π.

Für Einzelheiten vergleiche man etwa P.Bundschuh: „Einführung in die Zahlentheorie“, Springer-Verlag, 4.Auflage 1998 Kap.V, §3.

10.3.6 Beispiele

(1) $g = 10;\ r = \frac{1}{7}$.

Der g-al-Algorithmus liefert

$$\begin{array}{lll}
 & z_0 = 0; & r_0 = r = \frac{1}{7} \\
 & z_1 = [10r_0] = \left[\frac{10}{7}\right] = 1; & r_1 = 10r_0 - z_1 = \frac{10}{7} - \frac{7}{7} = \frac{3}{7} \\
 & z_2 = [10r_1] = \left[\frac{30}{7}\right] = 4; & r_2 = 10r_1 - z_2 = \frac{30}{7} - \frac{28}{7} = \frac{2}{7} \\
 & z_3 = [10r_2] = \left[\frac{20}{7}\right] = 2; & r_3 = 10r_2 - z_3 = \frac{20}{7} - \frac{14}{7} = \frac{6}{7} \\
 & z_4 = [10r_3] = \left[\frac{60}{7}\right] = 8; & r_4 = 10r_3 - z_4 = \frac{60}{7} - \frac{56}{7} = \frac{4}{7} \\
 & z_5 = [10r_4] = \left[\frac{40}{7}\right] = 5; & r_5 = 10r_4 - z_5 = \frac{40}{7} - \frac{35}{7} = \frac{5}{7} \\
 & z_6 = [10r_5] = \left[\frac{50}{7}\right] = 7; & r_6 = 10r_5 - z_6 = \frac{50}{7} - \frac{49}{7} = \frac{1}{7} \\
\text{Damit ist } r_6 = r_0 \text{ und} & z_7 = z_1 = 1 & \\
 & z_8 = z_2 = 4 & \\
 & z_9 = z_3 = 2 & \\
 & z_{10} = z_4 = 8 & \\
 & z_{11} = z_5 = 5 & \\
 & z_{12} = z_6 = 7 &
\end{array}$$

etc, also $\frac{1}{7} = 0,142857142857\ldots = 0,\overline{142857}$.

Die Überstreichung bedeutet, dass sich der Zifferblock 142857 periodisch wiederholt.
Die *Periodenlänge* beträgt hier 6.

Mit elementarer Zahlentheorie (vergleiche etwa das oben zitierte Buch von P.Bundschuh) kann man das theoretisch untermauern.

(2) Ist $g = 2$ und etwa $r = 0.010101\ldots = 0,\overline{01}$, so ist die rechte Seite eine Abkürzung für die Reihe

$$\begin{aligned}
\frac{0}{2} + \frac{1}{2^2} + \frac{0}{2^3} + \frac{1}{2^4} + \frac{0}{2^5} + \frac{1}{2^6} + \cdots \quad &= \quad \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \cdots \\
&= \quad \frac{1}{2^2}\left(1 + \frac{1}{2^2} + \frac{1}{2^4} + \cdots\right) \\
&= \quad \frac{1}{2^2} \cdot \frac{1}{1 - \frac{1}{4}} = \frac{1}{2^2} \cdot \frac{1}{\frac{3}{4}} = \frac{1}{2^2} \cdot \frac{4}{3} = \frac{1}{3}
\end{aligned}$$

Die periodische Dual-Zahl $0,\overline{01}$ stellt also die rationale Zahl $r = \frac{1}{3}$ dar.

(3) Ob eine rationale Zahl eine abbrechende oder einer nicht abbrechende g-al-Darstellung hat, hängt i.A. von der Grundzahl g ab:
So besitzt z.B. die rationale Zahl $x = \frac{5}{g}$(im Dezimalsystem) im System mit der Grundzahl $g = 3$ die abbrechende Entwicklung(Periode Null)

$$x = 0,12 = \frac{1}{3} + \frac{2}{3^2}\ ,$$

während im Dezimalsystem ($g = 10$) bekanntlich

$$x = 0,555\ldots = 0,\bar{5}$$

gilt.

Rationale Zahlen sind –wie wir bereit festgestellt haben– aber für *jede* Grundzahl $g \geq 2$ stets durch periodische Entwicklungen gekennzeichnet (abbrechende Entwicklung bedeutet: Periode Null).

10.3.7 Aufgabe

Begründen Sie folgende Regeln für die Umwandlung eines periodischen Dezimalbruchs in eine rationale Zahl, aus ???: HANDBUCH ??? (FALKENVERLAG, 1997).

1.Fall: Die Periode umfasst alle Ziffern nach dem Komma.

Man schreibt alle Ziffern nach dem Komma in den Zähler. In den Nenner schreibt man so viele Neunen, wie die Periode Ziffern umfasst.

Beispiele:

(1) $x = 0,\overline{142857} = \dfrac{142857}{999999} = \dfrac{47619}{333333} = \dfrac{5291}{37037} = \dfrac{402}{2849} = \dfrac{1}{7}.$

(2) $x = 0,\overline{425871} = \dfrac{428571}{999999} = \dfrac{3}{7}.$

Diese Beispiele lassen vermuten, dass die Periodenlänge (hier 6) nur vom Nenner, nicht aber vom Zähler abhängt. Tatsächlich haben

$$\frac{1}{7}, \frac{2}{7}, \frac{3}{7}, \frac{4}{7}, \frac{5}{7} \text{ und } \frac{6}{7}$$

alle die Periodenlänge 6 und man erhält die Entwicklung von $\frac{2}{7}, \frac{3}{7}, \frac{4}{7}, \frac{5}{7}, \frac{6}{7}$ aus der von $\frac{1}{7} = 0,\overline{142857}$ durch zyklische Vertauschung der Ziffernfolge, etwa

$$\begin{aligned} \frac{3}{7} &= 0,\overline{428571} \\ \frac{2}{7} &= 0,\overline{285714} \\ \frac{6}{7} &= 0,\overline{857142} \\ \frac{4}{7} &= 0,\overline{571428} \\ \frac{5}{7} &= 0,\overline{714285} \end{aligned}$$

Können Sie das begründen?

2.Fall: Zwischen dem Komma und der Periode stehen n Nullen.

Man geht vor, wie im ersten Fall, hängt aber im Nenner an die Neuner n Nullen an.

Beispiel: $x = 0,000\overline{56} = \dfrac{56}{99000} = \dfrac{7}{12375}.$

3.Fall: Zwischen dem Komma und der Periode stehen Ziffern, die nicht alle Null sind.

In diesem Fall zerlegt man die Dezimalzahl in eine Summe aus einer endlichen Dezimalzahl und einer dem 2.Fall entsprechenden periodischen Dezimalzahl und addiert die beiden sich hieraus ergebenden Brüche.

Beispiel: $x = 0,123\overline{56} = 0,123 + 0,00\overline{56} = \dfrac{123}{1000} + \dfrac{56}{99000} = \dfrac{12233}{99900}$.

10.4 Umordnungssätze für Reihen, Reihenprodukte

Die Hauptregeln für das Rechnen mit *endlichen Summen* sind

- das Assoziativgesetz
- das Distributivgesetz
- das Kommutativgesetz.

Wie übertragen sich diese Regeln auf das Rechnen mit (unendlichen) Reihen?

Zunächst gilt (wie immer sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$)

10.4.1 Satz

Ist $\sum(a_k),\ a_k \in \mathbb{K}$, eine konvergente Reihe und fasst man ihre Summanden durch Klammern zu neuen Summanden zusammen, d.h. setzt man für $0 \le k_1 < k_2 < \ldots < k_\nu < \ldots$

$$A_1 := a_0 + \ldots + a_{k_1},\quad A_2 := a_{k_1+1} + \ldots + a_{k_2},\quad \ldots$$

dann ist auch die Reihe $\sum(A_\nu)_{\nu\ge 1}$ konvergent, und es gilt

$$\sum_{\nu=1}^{\infty} A_\nu = \sum_{k=0}^{\infty} a_k\ .$$

Zusatz: Schon vorhandene Klammern in einer konvergenten Reihe darf man jedoch dann und nur dann weglassen, wenn die so entstehende (unbeklammerte) Reihe wieder konvergiert.

Beweis : $(S_m) := \left(\sum\limits_{\nu=1}^{m} A_\nu\right)$ ist eine Teilfolge von $(s_n) := \left(\sum\limits_{k=0}^{n} a_k\right)$.

Lässt man jedoch in der konvergenten Reihe

$$(1-1) + (1-1) + (1-1) + \ldots\ (=0)$$

die Klammern weg, erhält man die divergente Reihe

$$1 - 1 + 1 - 1 + 1 - 1 + 1 - 1 + \ldots\quad .$$

$\square$

Wir wissen (vgl. §7.8(18)):
Ist $(a_k)_{k\ge 0}$ eine reelle oder komplexe Zahlenfolge und $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ eine Abbildung, bei der jede natürliche Zahl genau einmal als Bild auftritt(Permutation, Bijektion), dann heißt die Folge

$$(b_k) = \left(a_{\varphi(k)}\right)$$

eine *Umordnung von* (a_k) und ist (a_k) konvergent mit dem Grenzwert a, dann ist auch (b_k) konvergent mit dem selben Grenzwert:

$$\lim_{k\to\infty} b_k = \lim_{k\to\infty} a_{\varphi(k)} = a \ .$$

Wir werden im Folgenden sehen, dass sich eine Permutation $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ sehr wohl auf das Konvergenzverhalten und den Wert (Summe) einer Reihe auswirken kann. Dazu betrachten wir zunächst das folgende *Beispiel*:
Wir wissen, dass die alternierende harmonische Reihe

$$\sum \left(\frac{(-1)^k}{k+1}\right)_{k\geq 0} = \left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \ldots\right)$$

nach dem Leibniz-Kriterium konvergiert und dass für ihre Summe S gilt

$$\frac{1}{2} < S < \frac{5}{6} \ .$$

Betrachtet man die Abbildung (Permutation) $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$

$$\left.\begin{array}{l} \varphi(3k) = 4k \\ \varphi(3k+1) = 4k+2 \\ \varphi(3k+2) = 2k+1 \end{array}\right\} \text{ und setzt } \ b_k := a_{\varphi(k)} \ ,$$

dann nennen wir die Reihe $\sum(b_k) = \sum(a_{\varphi(k)})$ eine *Umordnung* der Reihe $\sum(a_k)$. Es ist also in suggestiver Schreibweise

$$\begin{aligned} \sum b_k \quad &= \quad a_0 + a_2 + a_1 + \\ & \quad\quad a_4 + a_6 + a_3 + \\ & \quad\quad \vdots \\ & \quad\quad a_{4k} + a_{4k+2} + a_{2k+1} + \\ & \quad\quad \vdots \\ &= \quad 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \ldots + \frac{1}{4k+1} + \frac{1}{4k+3} - \frac{1}{2k+2} + \ldots \end{aligned}$$

Diese Reihe ist also eine Umordnung der ursprünglichen, sie ist wieder konvergent (Beweis?) und für ihre Summe $\tilde{S}$ gilt

$$\tilde{S} = \frac{3}{2} S \ .$$

Man vergleiche hierzu die Musterlösung von Aufgabe 68.
Die umgeordnete Reihe ist zwar wieder konvergent, aber die Summe hat sich geändert.
Unser Ziel ist es zu zeigen, dass genau die *absolut konvergenten Reihen* diejenigen Reihen sind, deren Konvergenzverhalten sich bei einer Umordnung nicht ändert (genauer: *jede* Umordnung ist ebenfalls konvergent und hat dieselbe Summe wie die Ausgangsreihe).

10.4.2 Definition

Ist $\sum(a_k), \ a_k \in \mathbb{K}$, eine Reihe und $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ bijektiv (Permutation, Bijektion), dann heißt die Reihe $\sum(a_{\varphi(k)})$ eine *Umordnung der Reihe* $\sum(a_k)$.

10.4.3 Satz (Kleiner Umordnungssatz)

Ist $\sum(a_k)$ absolut konvergent, $a_k \in \mathbb{K}$, dann ist auch jede Umordnung $\sum(a_{\varphi(k)})$ (absolut) konvergent, und es gilt

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} a_{\varphi(k)} \ .$$

Kurz: Absolut konvergente Reihen darf man beliebig umordnen ohne dass sich das Konvergenzverhalten oder die Summe ändern.

Beweis : Wir zeigen zunächst die absolute Konvergenz (und damit insbesondere die Konvergenz) einer beliebigen Umordnung $\sum a_{\varphi(k)}$ von $\sum a_k$.
Wir benutzen das Kriterium 10.4.1.
Nach Voraussetzung gibt es ein $C > 0$, so dass

$$\sigma_n := \sum_{k=0}^{n} |a_k| \leq C \ \text{ für alle } n \in \mathbb{N}_0 \text{ gilt.}$$

Hieraus folgt unmittelbar

$$|a_{\varphi(1)}| + |a_{\varphi(2)}| + \ldots + |a_{\varphi(n)}| \leq C$$

für alle $n \in \mathbb{N}_0$, denn diese Summe wird ja nach oben abgeschätzt durch

$$\sigma_N := |a_1| + |a_2| + \ldots + |a_N| \ \text{ mit } \ N := \max\{\varphi(1), \ldots, \varphi(n)\}$$

Wir wissen also, dass auch jede Umordnung wieder konvergiert (sogar absolut).
Ist nun $s := \sum\limits_{k=0}^{\infty} a_k$, dann haben wir noch

$$s = \sum_{k=0}^{\infty} a_{\varphi(k)}$$

zu zeigen.
Dazu sei $\varepsilon > 0$ beliebig vorgegeben. Wegen der absoluten Konvergenz von $\sum(a_k)$, d.h. wegen der Konvergenz von $\sum(|a_k|)$ gibt es ein $N \in \mathbb{N}_0$, so dass für alle $n \geq N$ gilt

$$\left| \sum_{k=0}^{\infty} |a_k| - \sum_{k=0}^{n} |a_k| \right| < \frac{\varepsilon}{2} \ ,$$

d.h. insbesondere gilt

$$\sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{2} \ \text{ und auch}$$

$$\left| s - \sum_{k=0}^{N} a_k \right| = \left| \sum_{k=N+1}^{\infty} a_k \right| \leq \sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{2} \ .$$

Wir wählen nun eine natürliche Zahl M so groß, dass die Zahlen $0, 1, \ldots, N$ unter den Zahlen $\varphi(0), \varphi(1), \ldots, \varphi(M)$ vorkommen (φ ist bijektiv!), also

$$\{0, 1, \ldots, N\} \subset \{\varphi(0), \varphi(1), \ldots, \varphi(M)\} \ .$$

Dann ist für $m \in \mathbb{N}$ mit $m \geq N$:

$$\left| \sum_{k=0}^{m} a_{\varphi(k)} - s \right| = \left| \sum_{k=0}^{m} a_{\varphi(k)} - \sum_{k=0}^{N} a_k + \sum_{k=0}^{N} a_k - s \right| \leq \sum_{k=N+1}^{\infty} |a_k| + \left| \sum_{k=0}^{N} a_k - s \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

$\square$

Wir haben also gesehen:
Eine absolut konvergente Reihe $\sum a_k$ ist gegenüber einer Umordnung ihrer Summanden unempfindlich. Ganz anders ist die Situation, wenn $\sum a_k$ zwar konvergiert, aber $\sum |a_k|$ nicht konvergiert. Eine solche Reihe nennt man gelegentlich auch *bedingt konvergent*.

Wir führen zunächst folgende Bezeichnungen ein:
Für $a_k \in \mathbb{R},\ k \in \mathbb{N}_0$, sei

$$a_k^+ = \begin{cases} a_k & \text{, falls } a_k \geq 0 \\ 0 & \text{, falls } a_k < 0 \end{cases} \quad \text{und} \quad a_k^- = \begin{cases} 0 & \text{, falls } a_k \geq 0 \\ -a_k & \text{, falls } a_k < 0, \end{cases}$$

dann gilt offenbar ($k \in \mathbb{N}_0$)

$$(*) \qquad a_k = a_k^+ - a_k^-$$

und

$$(**) \qquad |a_k| = a_k^+ + a_k^-$$

Bemerkungen:

(a) Eine Reihe $\sum a_k,\ a_k \in \mathbb{R}$, ist genau dann absolut konvergent, wenn $\sum a_k^+$ und $\sum a_k^-$ konvergieren.

(b) Ist $\sum a_k$ konvergent ($a_k \in \mathbb{R}$), aber nicht absolut konvergent (also nur bedingt konvergent), dann sind $\sum a_k^+$ und $\sum a_k^-$ beide divergent.

Beweis :

(a) folgt sofort aus ($**$), da nur nicht negative Summanden auftreten.

(b) Wäre eine der Reihen $\sum a_k^+;\ \sum a_k^-$ konvergent, dann folgt aus $a_k^+ = a_k + a_k^-$ bzw. $a_k^- = a_k^+ - a_k$, dass auch die andere konvergiert. Wegen $|a_k| = a_k^+ + a_k^-$, wäre dann auch $\sum |a_k|$ konvergent im Widerspruch zur Voraussetzung.

$\square$

Der folgende erstaunliche Satz geht auf P.G.L.Dirichlet (1837) zurück.

10.4.4 Satz

Ist $\sum a_k,\ a_k \in \mathbb{C}$, eine bedingt konvergente Reihe, dann gibt es eine Umordnung $\sum a_{\varphi(k)}$, die divergiert.

Beweis :

(a) Sei zunächst $a_k \in \mathbb{R},\ k \in \mathbb{N}_0$.
Dann sind die Reihen $\sum a_k^+$ und $\sum a_k^-$ beide divergent und da $a_k^+ \geq 0$ und $a_k^- \geq 0$ gilt, sind ihre Teilsummen insbesondere nach oben nicht beschränkt.
Zur Vereinfachung setzen wir $p_k := a_k^+$ und $q_k := a_k^-$. Nun bestimmt man induktiv (rekursiv) eine Folge $(\varphi(k))$ natürlicher Zahlen mit $\varphi(k) < \varphi(k+1)$, so dass für alle $k \in \mathbb{N}_0$ gilt

$$p_0 + p_1 + \ldots + p_{\varphi(k)} > k + q_0 + q_1 + \ldots + q_k$$

und betrachtet die Reihe

$$(*) \qquad p_0 + p_1 + \ldots + p_{\varphi(0)} - q_0 + p_{\varphi(0)+1} + \ldots + p_{\varphi(1)} - q_1 + p_{\varphi(1)+1} + \ldots ,$$

dann ist diese eine Umordnung der ursprünglichen Reihe, die offensichtlich divergiert, da ihre Partialsummen nicht beschränkt sind, denn die Partialsumme der Reihe $(*)$, deren letzter Summand q_k ist.

(b) Sei jetzt $a_k \in \mathbb{C}$. Wegen

$$|a_k| \leq |Re\, a_k| + |Im\, a_k|$$

können $\sum |Re\, a_k|$ und $\sum |Im\, a_k|$ nicht beide konvergieren.
Ist etwa $\sum |Re\, a_k|$ divergent, dann gibt es wegen der Konvergenz von $\sum Re\, a_k$ nach (a) eine Umordnung $\sum Re\, a_{\varphi(k)}$, die divergiert.
Unabhängig vom Konvergenzverhalten von $\sum Im\, a_{\varphi(k)}$, ist dann $\sum a_{\varphi(k)}$ divergent.

□

Beispiel:
Wir erläutern den Satz am Beispiel der alternierenden harmonischen Reihe:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \ldots$$

Hier ist also

$$(p_0, p_1, p_2, \ldots) = \left(1, \frac{1}{3}, \frac{1}{5}, \ldots\right)$$

$$(q_0, q_1, a_2, \ldots) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \ldots\right)$$

Hier nimmt man zunächst so viele positive Glieder in ihrer natürlichen Anordnung

$$p_k = \frac{1}{2k-1},\ k \in \mathbb{N}_0$$

so, dass ihre Summe die Zahl 1 zum ersten Mal überschreitet, danach subtrahiert man so viele Glieder

$$q_k = \frac{1}{2k}$$

(d.h. man addiert Glieder -$\frac{1}{2k}$), bis die Summe dieser Glieder und der bislang aufgelisteten Glieder, die 1 zum ersten Mal unterschreiten. Im nächsten Schritt nimmt man von der nächsten positiven Glieder so viele, dass 2 überschritten wird, dann wieder negative bis die 2 zum ersten Mal unterschritten wird. Entsprechend geht man für jede der Zahlen 3,4,5,... vor. Die ersten Schritte lassen sich folgendermaßen darstellen

$$\underbrace{\underbrace{\underbrace{\underbrace{1 + \frac{1}{3}}_{>1} - \frac{1}{2}}_{<1} + \frac{1}{5} + \frac{1}{7} + \frac{1}{9} + \ldots + \frac{1}{41}}_{>2} - \frac{1}{4}}_{<2} + \frac{1}{43} + \frac{1}{45} + \ldots$$

Berechnet man mit s_n die Partialsummen dieser Reihe und ist $C > 0$ beliebig vorgegeben, so wählt man zunächst eine natürliche Zahl $m > C$ und ist dann s_N diejenige Partialsumme, welche die natürliche Zahl $m+1$ zum ersten Mal überschreitet, so gilt nach Konstruktion (insbesondere $|-\frac{1}{2k}| < 1,\ k \in \mathbb{N}$)

$$s_n \geq m > C$$

für alle $n \geq N$. Die Folge (s_n) ist also nicht beschränkt und damit nicht konvergent.
Man beachte, dass die negative Summanden

$$-\frac{1}{2}, -\frac{1}{4}, -\frac{1}{6}, \dots$$

der Ausgangsreihe zwar alle auftauchen, aber „immer später".

Nennt man eine konvergente Reihe $\sum a_k,\ a_k \in \mathbb{K}$, *unbedingt konvergent*, wenn auch *jede Umordnung* konvergiert und zwar mit derselben Summe wie die Ausgangsreihe, so kann man jetzt sagen:

10.4.5 Satz (Dirichlet, 1837)

Eine Reihe $\sum a_k,\ a_k \in \mathbb{K}$, ist genau dann absolut konvergent, wenn sie unbedingt konvergent ist.

Beweis :

„$\Longrightarrow$"
Nach dem kleinen Umordnungssatz konvergiert für eine absolut konvergente Reihe jede Umordnung mit derselben Summe wie die Ausgangsreihe. Eine absolut konvergente Reihe ist also unbedingt konvergent.

„$\Longleftarrow$"
Sei also umgekehrt $\sum |a_k|$ nicht konvergent, s.h. wäre $\sum a_k$ nicht absolut konvergent, dann gibt es nach Satz 10.4.4 eine Umordnung, die divergiert, das ist ein Widerspruch zur Voraussetzung, dass *jede* Umordnung konvergiert.

$\square$

Übungsaufgabe:
Man zeige: Eine Reihe $\sum a_k$ ist bereits dann unbedingt konvergent, wenn jede ihre Umordnungen konvergiert (über die Summen dieser Umordnungen braucht man nichts vorauszusetzen).

Für *reelle* Reihen $\sum a_k,\ a_k \in \mathbb{R}$, die nun bedingt konvergieren, hat B.Riemann in seiner berühmten Göttinger Habilitationsschrift (1854, veröffentlicht erst 1866/68) bewiesen, dass man sogar durch geeignete Umordnung der Summanden eine mit einer *beliebigen Summe* s konvergente Reihe erhalten kann.

10.4.6 Satz (Riemannscher Umordnungssatz)

Ist $\sum a_k$ eine bedingt konvergente reelle Reihe (d.h. $a_k \in \mathbb{R}$) und ist $s \in \mathbb{R}$ beliebig, dann gibt es eine Umordnung $\sum a_{\varphi(k)}$ mit

$$\sum_{k=0}^{\infty} a_{\varphi(k)} = s\ .$$

Der Beweis ist eine Modifikation des Beweises des Satzes von Dirichlet. Einen Beweis findet man z.B. in *Mangoldt/Knopp: „Einführung in die höhere Mathematik", Band 2, 80, S.227* (siehe Literaturverzeichnis). Weitere Verallgemeinerungen findet man in „Reihen-Knopp".

Für eine konvergente Reihe $\sum a_k,\ a_k \in \mathbb{C}$, die nun bedingt konvergiert, gibt es die Möglichkeit, dass
$\sum Re(a_k)$ konvergiert und $\sum Im(a_k)$ bedingt konvergiert oder, dass
$\sum Re(a_k)$ bedingt konvergiert und $\sum Im(a_k)$ konvergiert oder, dass
$\sum Re(a_k)$ und $\sum Im(a_k)$ bedingt konvergieren.

In den ersten beiden Fällen erhält man als mögliche Summen der Umordnungen jeweils eine Gerade vom Typ

$$s + \mathbb{R}i := \{s + ri;\ r \in \mathbb{R}\}; \quad \text{bzw.} \quad \mathbb{R} + ti := \{r + ti;\ r \in \mathbb{R}\}\ \ s, t \in \mathbb{R}\ ,$$

oder ganz $\mathbb{C}$ (im dritten Fall).

Aufgabe:
Geben Sie eine (bedingt konvergent) Reihe $\sum a_k,\ a_k \in \mathbb{C}$ an, so dass für die Menge S der möglichen Summen der jeweiligen Umordnung gilt

$$S = \mathbb{R} + i\ .$$

Wir wollen uns noch abschließend mit der *Multiplikation von Reihen* beschäftigen.
Hat man zwei endliche Summen etwa

$$s_n := a_0 + a_1 + \ldots + a_n;\ a_j \in \mathbb{K},\ n \in \mathbb{N}_0\ ,$$

$$t_m := b_0 + b_1 + \ldots + b_m,\ b_j \in \mathbb{K},\ m \in \mathbb{N}_0$$

zu multiplizieren, so hat man jeden Summanden der ersten Summe mit jedem Summanden der zweiten Summe zu multiplizieren und die entsprechenden Produkte aufzuaddiren:

$$\begin{aligned} s_n \cdot t_m = a_0 t_m + a_1 t_m + \ldots + a_n t_m &= a_0 b_0 + a_0 b_1 + \ldots + a_0 b_m \\ &+ a_1 b_0 + a_1 b_1 + \ldots + a_1 b_m \\ &+ \qquad \vdots \\ &+ a_n b_0 + a_n b_1 + \ldots + a_n b_m \end{aligned}$$

Für das Resultat schreibt man ja auch

$$S := s_n t_m = \sum_{\substack{0 \le k \le n \\ 0 \le j \le m}} a_k b_j$$

Hier ist ein allgemeines Kommutativ- und Distributivgesetz enthalten:
Ob man etwa in der $n \times m$-Matrix

$$(*) \qquad \begin{pmatrix} a_0 b_0 & a_0 b_1 & \cdots & a_0 b_m \\ a_1 b_0 & a_1 b_1 & \cdots & a_1 b_m \\ \vdots & & & \vdots \\ a_n b_0 & a_n b_1 & \cdots & a_n b_m \end{pmatrix}$$

zuerst die in den Zeilen stehenden Elemente aufsummiert, also

$$\begin{aligned} Z_0 &= a_0 b_0 + a_0 b_1 + \cdots + a_0 b_m \\ Z_1 &= a_1 b_0 + a_1 b_1 + \cdots + a_1 b_m \\ \vdots & \qquad \vdots \\ Z_n &= a_n b_0 + a_n b_1 + \cdots + a_n b_m \end{aligned}$$

bildet, oder ab man die in den Spalten stehende Elemente aufsummiert, also

$$\begin{aligned} S_0 &= a_0 b_0 + a_1 b_0 + \cdots + a_n b_0 \\ S_1 &= a_0 b_1 + a_1 b_1 + \cdots + a_n b_1 \\ \vdots & \qquad \vdots \\ S_n &= a_0 b_m + a_1 b_m + \cdots + a_n b_m \end{aligned}$$

bildet, so ist jeweils

$$S = \sum_{k=0}^{\infty} Z_k = \sum_{j=0}^{\infty} S_j \ .$$

Etwas allgemeiner gilt:
Ordnet man alle möglichen Produkte $a_k b_j$ in *beliebiger Weise* zu einer endlichen Folge $p_0, \dots, p_s$ an, dann gilt

$$S = p_0 + p_1 + \ldots + p_s \ .$$

Es liegt nun nahe das Verfahren auf die Multiplikation zweier konvergenter Reihen $\sum a_k$ und $\sum b_k$ zu übertragen.
Man bildet alle möglichen Produkte

$$a_k b_l, \ k \in \mathbb{N}_0, l \in \mathbb{N}_0$$

und ordnet sie etwa in dem (doppelt unendlichen) Schema ($\infty \times \infty$ - Matrix)

$$(**) \qquad \begin{pmatrix} a_0b_0 & a_0b_1 & a_0b_2 & \cdots \\ a_1b_0 & a_1b_1 & a_1b_2 & \cdots \\ a_2b_0 & a_2b_1 & a_2b_2 & \cdots \\ \vdots & \vdots & \vdots & \ldots \end{pmatrix}$$

an und schreibt die Produkte in $(**)$ als einfache Folge

$$p_0, p_1, p_2, \ldots, p_n, \cdots \ ,$$

d.h. in der Folge (p_n) kommt jedes Produkt $a_k b_j, \ k, l \in \mathbb{N}_0$ genau einmal vor (wegen der Abzählbarkeit von $\mathbb{N}_0 \times \mathbb{N}_0$ geht das immer, siehe §10.5).

Dann ist aber die Frage:
Gilt dann immer $\sum\limits_{k=0}^{\infty} a_k \cdot \sum\limits_{l=0}^{\infty} b_l = \sum\limits_{n=0}^{\infty} p_n$ (vorausgesetzt, $\sum p_n$ ist überhaupt konvergent)?

Zunächst eine Sprechweise:
Sind $\sum a_k$ und $\sum b_l$ zwei (konvergente) Reihen, $a_k, b_l \in \mathbb{K}$, und ist (p_n) eine Anordnung der möglichen Produkte $a_k b_l, \ k, l \in \mathbb{N}_0$ als Folge, dann heißt die Reihe $\sum (p_n)_{n \geq 0}$ eine *Produktreihe* von $\sum a_k$ und $\sum b_k$.

Für spätere Anwendungen (Potenzreihen) erweist sich folgende Anordnung der Produkte $a_k b_l$ nach „Schrägzeilen" (Diagonalen)

$$\begin{pmatrix} a_0b_0 & a_0b_1 & a_0b_2 & a_0b_3 & \cdots & a_0b_n & \cdots \\ a_1b_0 & a_1b_1 & a_1b_2 & \cdots & & & \\ a_2b_0 & a_2b_1 & \cdots & & & & \\ a_3b_0 & \cdots & & & & & \\ \vdots & & & & & & \\ \vdots & & & & & & \\ a_nb_0 & \cdots & & & & & \\ \vdots & & & & & & \end{pmatrix}$$

also besonders zweckmäßig. Hier ist also

$$p_0 = a_0b_0, \quad p_1 = a_0b_1, \quad p_2 = a_0b_2, \quad p_3 = a_1b_1, \quad p_4 = a_2b_0, \quad p_5 = a_0b_3, \quad \cdots$$

Wir setzten nun $c_0 := p_0 = a_0b_0$ und bezeichnen die Summe der durch Pfeile auf der „n-ten Diagonale“ miteinander verbundenen Produkte mit c_n, also für $n \in \mathbb{N}_0$:

$$\begin{aligned}
c_0 &= a_0b_0 \\
c_1 &= a_0b_1 + a_1b_0 \\
c_2 &= a_0b_2 + a_1b_1 + a_2b_0 \\
\vdots &\quad \vdots \\
c_n &= a_0b_n + a_1b_{n-1} + \ldots + a_{n-1}b_1 + a_nb_0
\end{aligned}$$

Ein „Summenblock“ c_n besteht also genau aus den Produkten a_kb_l, für welche $k + l = n$ gilt, die also die gleiche Indexsumme haben.
Die so gebildete Reihe $\sum c_n$ hat genau alle Produkte $a_kb_l,\ k, l \in \mathbb{N}_0$, als Summanden, denn sie entsteht ja aus $\sum p_\nu$, indem man Summanden mit gleicher Indexsumme zusammenfasst.
Die so gebildete Produktreihe $\sum c_n$ mit $c_n = \sum\limits_{k=0}^{\infty} a_kb_{n-k}$ aus $\sum a_k$ und $\sum b_l$ hat einen eigenen Namen, man nennt sie das *Cauchy-Produkt*.
Cauchy (1821) hat schon folgendes Beispiel angegeben:
Nimmt man

$$a_k = b_k = \frac{(-1)^k}{\sqrt{k+1}}, \quad k \in \mathbb{N}_0 \ ,$$

dann sind $\sum a_k$ und $\sum b_k$ beide bedingt konvergent, aber es ist

$$c_n = (-1)^n \left(\frac{1}{\sqrt{1}\sqrt{n+1}} + \frac{1}{\sqrt{2}}\frac{1}{\sqrt{n}} + \ldots + \frac{1}{\sqrt{n+1}}\frac{1}{\sqrt{1}} \right) \ .$$

Ersetzt man in der Klammer jeden Radikanden durch größten von ihnen, also $n + 1$, so ergibt sich

$$|c_n| \geq \frac{n+1}{\sqrt{(n+1)^2}} = 1 \ .$$

Daher kann die Cauchy-Produkt-Reihe $\sum c_n$ von $\sum a_k$ und $\sum b_l$ nicht konvergieren. Wenn jedoch die Ausgangsreihen $\sum a_k$ und $\sum b_l$ beide *absolut* konvergieren, dann konvergiert jede Produktreihe $\sum p_n$ auch absolut, und es gilt

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

Insbesondere gilt auch

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

Dies ist der Inhalt des folgenden Satzes.

10.4.7 Satz (Produktreihensatz)

Sind die beiden Reihen $\sum a_k$ und $\sum b_l,\ a_k, b_l \in \mathbb{K}$, absolut konvergent, dann ist auch *jede* Produktreihe $\sum p_n$ absolut konvergent, und es gilt

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

Insbesondere ist die Cauchy-Produktreihe $\sum c_n$ mit $c_n = \sum\limits_{k=0}^{\infty} a_k b_{n-l}$ absolut konvergent, und es gilt

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

Beweis : Sei $\sum p_n$ irgendeine Produktreihe von $\sum a_k$ und $\sum b_l$ und $\sigma_m = \sum\limits_{n=0}^{m} p_n$ eine Partialsumme.

Ist dann q der größte Index der Zahlen a_k und b_l aus denen die Partialsumme σ_m gebildet werde, so ist

$$\begin{aligned} |\sigma_m| \leq \sum_{n=0}^{m} |p_n| &\leq |a_0 b_0| + \ldots + |a_0 b_q| \\ &+ |a_1 b_0| + \ldots + |a_1 b_q| \\ &\quad\vdots \\ &+ |a_q b_0| + \ldots + |a_q b_q| \\ &= \sum_{k=0}^{q} |a_k| \cdot \sum_{l=0}^{q} |b_l| \ . \end{aligned}$$

Wegen der vorausgesetzten Konvergenz von $\sum |a_k|$ und $\sum |b_l|$ folgt nach dem Beschränktheitskriterium, die absolute Konvergenz von $\sum p_n$,denn für alle m gilt

$$\sum_{n=0}^{m} |p_n| \leq \sum_{k=0}^{m} |a_k| \cdot \sum_{l=0}^{m} |b_l| \ .$$

Nach dem kleinen Umordnungssatz haben dann *alle* Produktreihen die Summe $s := \sum\limits_{n=0}^{\infty} p_n$.

Um s zu ermitteln bilden wir eine spezielle Produktreihe $\sum q_n$, indem wir die Produkte gemäß dem folgenden quadratischen Schema anordnen und zusammenfassen:

$$\begin{array}{ccccccc} a_0 b_0 & & a_0 b_1 & & a_0 b_2 & \cdots \\ & & \downarrow & & \downarrow & \cdots \\ a_1 b_0 & \leftarrow & a_1 b_1 & & a_1 b_2 & \cdots \\ & & & & \downarrow & \cdots \\ a_2 b_0 & \leftarrow & a_2 b_1 & \leftarrow & a_2 b_2 & \cdots \\ \vdots & & \vdots & & \vdots & \end{array}$$

Dann konvergiert die Partialsumme

$$q_0 + q_1 + \ldots + q_{(m+1)^2-1} = (a_0 a_1 + \ldots + a_m)(b_0 + b_1 + \ldots + b_m)$$

einerseits gegen

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

nach dem oben Bewiesenen aber ebenfalls gegen s, so dass in der Tat gilt

$$s = \sum_{k=0}^{\infty} a_k \cdot \sum_{l=0}^{\infty} b_l \ .$$

Speziell gilt für die Cauchy-Produkt-Reihe

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

(dabei ist $c_n = \sum\limits_{k=0}^{m} a_k b_{n-k}$).

□

10.4.8 Beispiele und Bemerkungen

(1) Die Cauchy-Produkt-Reihe $\sum c_n$ konvergiert bereits dann schon gegen $\sum\limits_{k=0}^{\infty} a_k \cdot \sum\limits_{l=0}^{\infty} b_l$ (allerdings i.A. nur bedingt), wenn nur *eine* der beiden Reihen $\sum a_k$ oder $\sum b_l$ absolut und die andere nur bedingt konvergiert. (Satz von F.Mertens(1875), siehe Reihen-Knopp S.330/331)

(2) Ist das Cauchy-Produkt $\sum c_n$ der beiden konvergenten Reihen $\sum a_k$ und $\sum b_l$ selber konvergent, so gilt auch

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

(Satz von Abel, 1826)

Hierfür werden wir später mit Hilfe des *Abelschen Grenzwertsatzes* einen einfachen und durchsichtigen Beweis geben.

(3) Wir wissen, dass die geometrische Reihe $\sum z^k$ für $z \in \mathbb{C}$ genau für $|z| < 1$ konvergiert, und dass dann gilt

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

Bildet man das Cauchy-Produkt der Reihe mit sich selbst, so erhält man

$$\boxed{\frac{1}{(1-z)^2} = \sum_{k=0}^{\infty} z^k \cdots \sum_{l=0}^{\infty} z^l = \sum_{n=0}^{\infty} (n+1) z^n, \qquad |z| < 1}$$

und etwas allgemeiner für alle $k \in \mathbb{N}_0$ und alle $z \in \mathbb{C}$, mit $|z| < 1$

$$\boxed{\frac{1}{(1-z)^{k+1}} = \sum_{n=0}^{\infty} \binom{n+k}{k} z^n = \sum_{n=0}^{\infty} \binom{n+k}{k} z^n}$$

(4) *Funktionalgleichung von exp.*

Wir wissen, dass die Exponentialreihe $\sum \frac{z^k}{k!}$ für alle $z \in \mathbb{C}$ absolut konvergiert. Nach Definition ist für $z \in \mathbb{C}$

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

Ist nun $w \in \mathbb{C}$ und

$$exp(w) = \sum_{l=0}^{\infty} \frac{w^l}{l!} \ ,$$

dann ist das n-te Glied ihres Cauchy-Produktes

$$\begin{aligned} c_n &:= \frac{z^0 w^n}{0!n!} + \frac{z^1 w^{n-1}}{1!(n-1)!} + \ldots + \frac{z^n w^0}{n!0!} \\ &= \frac{1}{n!}\left\{\binom{n}{0}w^n + \binom{n}{1}zw^{n-1} + \ldots + \binom{n}{n}z^n\right\} \\ &= \frac{(z+1)^n}{n!} \qquad \text{(nach der Binomischen Formel).} \end{aligned}$$

Also gilt für alle $z, w \in \mathbb{C}$

$$\sum_{k=0}^{\infty} \frac{z^k}{k!} \cdot \sum_{l=0}^{\infty} \frac{w^l}{l!} = \sum_{n=0}^{\infty} \frac{(z+w)^2}{n!} ,$$

oder

$$\boxed{exp(z) \cdot exp(w) = exp(z+w), \qquad z, w \in \mathbb{C}}$$

Das ist die *Funktionalgleichung* (Additionstheorem) der exp-Funktion. Sie gilt für komplexe wie für reelle Argumente. Sie hat wichtige Konsequenzen, z.B. folgt aus ihr, dass für alle $z \in \mathbb{C}$ gilt

$$exp(z) \neq 0 \quad \text{und} \quad (exp(z))^{-1} = exp(-z) .$$

Wir werden uns mit der exp-Funktion, insbesondere der mit reellen Argumenten, demnächst ausführlich beschäftigen.

10.5 Abbildungen, Abzählbarkeit

In allen Teildisziplinen der Mathematik und ihren Anwendungen, sogar im täglichen Leben, ist der Begriff der *Abbildung* oder *Funktion* von zentraler Bedeutung. Wir haben ihn schon mehrfach –mehr oder weniger explizit– verwendet und gehen davon aus, dass er der geneigten Leserin / dem geneigten Leser hier nicht zum ersten Mal begegnet.

Wir stellen die wichtigen Grundtatsachen zusammen. Der Abbildungsbegriff hat im Laufe seiner historischen Entwicklung (Fermat, Descartes, Leibniz, Euler, Dirichlet) viele Abänderungen erfahren, eine exakte Definition führt ihn auf Grundbegriffe der Mengenlehre zurück.

Wir wollen den Abbildungsbegriff im Sinne der folgenden Arbeitsdefinition verwenden:

10.5.1 Arbeitsdefinition

Seien X und Y beliebige Mengen. Eine Abbildung (oder Funktion) f von X *nach* Y (oder *in* Y) ist eine Vorschrift, die jedem Element $x \in X$ genau ein Element $y \in Y$ zuordnet.

Schreibweise: $f : X \to Y$
Sprechweise: f ist Abbildung von X nach Y

Das dem Element $x \in X$ in eindeutiger Weise durch f zugeordnete Element y wird auch mit $f(x)$ bezeichnet, $f(x)$ heißt auch das *Bild* von x unter f oder der Wert von f an der Stelle x.
Um Verwechslungen zu vermeiden, wird die Zuordnung von x zu $f(x)$ mit einem gesonderten Pfeil bezeichnet:

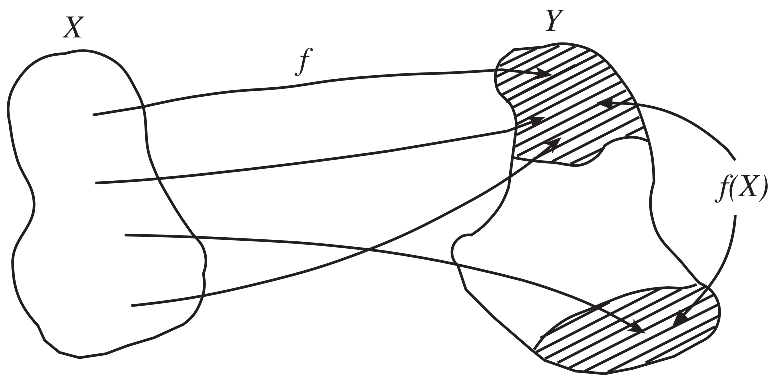
$$x \mapsto f(x) .$$

Die Menge X heißt der *Definitionsbereich*(Definitionsmenge) von f, die Menge Y der *Zielbereich* (Zielmenge) von f und

$$W_f := \{y \in Y;\ \text{es gibt ein } x \in X \text{ mit } y = f(x)\}$$

der (exakte) Wertebereich von f oder die Bildmenge von f oder kurz das *Bild* von f, andere Bezeichnung: $f(X)$.
Das Element $x \in X$ nennt man auch *Argument von* f.

Man beachte, dass die Fälle $X = \emptyset$ oder $Y = \emptyset$ nicht ausgeschlossen sind. Ist $X = \emptyset$, dann gibt es genau eine Abbildung von X nach Y, nämlich die leere Abbildung $\emptyset : \emptyset \to Y$.
Ist aber $Y = \emptyset$ und $X \neq \emptyset$, dann kann es offensichtlich überhaupt keine Abbildung von X nach Y geben.
Für die Menge aller Abbildungen $f : X \to Y$ verwendet man häufig die Bezeichnung $Abb(X, Y)$.



Ist $f : X \to Y$ eine Abbildung, so heißt

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

der *Graph* von f. Insbesondere Im Fall $X \subset \mathbb{R}$ und $Y = \mathbb{R}$ veranschaulicht man sich Funktionen (Abbildungen) $f : X \to Y$ mit Hilfe des Graphen. Beispiele hierzu haben wir schon genügend kennen gelernt. Auch im Fall $X \subset \mathbb{R} \times \mathbb{R}$ und $Y = \mathbb{R}$ kann man sich eine Abbildung $f : X \to Y$ häufig durch ihren Graphen veranschaulichen.

Was schreibt man hin, wenn man eine Abbildung anzugeben hat?

Dazu einige Formulierungen zur Auswahl (frei nach JÄNISCH: LINEARE ALGEBRA):
Als Beispiele betrachten wir die Abbildung f von $X = \mathbb{Z}$ nach $Y = \mathbb{N}_0$, die jeder ganzen Zahl x ihr Quadrat x^2 zuordnet. Dann kann man schreiben:

Sei $f : \mathbb{Z} \to \mathbb{N}_0$ die durch $f(x) = x^2$ für alle $x \in \mathbb{Z}$ gegebene Abbildung oder kürzer:
Sei

$$\begin{aligned} f : \mathbb{Z} &\to \mathbb{N}_0, \ \text{die durch} \\ x &\mapsto x^2 \end{aligned}$$

gegebene Abbildung oder noch kürzer:
Betrachte

$$\begin{aligned} f :\ \mathbb{Z} &\to \mathbb{N}_0\ , \\ x &\mapsto x^2\ . \end{aligned}$$

Manchmal ist es gar nicht nötig, der Abbildung einen eigenen Namen zu geben, in unserem Beispiel könnte man einfach

$$\begin{array}{rcl} \mathbb{Z} & \to & \mathbb{N}_0 \, , \\ x & \mapsto & x^2 \end{array}$$

schreiben, was sicherlich praktisch und suggestiv ist.

Die Aufgabe von Definitionsbereich, hier $X = \mathbb{Z}$, und Zielbereich, hier $Y = \mathbb{N}_0$, kann man sich aber nicht ersparen, und es ist auch nicht zulässig, unsere Abbildung einfach x^2 zu nennen. x^2 ist der Wert unserer Abbildung an der Stelle x oder das Bild von x unter der gegebenen Abbildung, aber nicht die Abbildung selber. Sprechweisen, wie „die Funktion x^2" sind in der Literatur jedoch üblich. aber nicht gut.
Für den Begriff „Abbildung" verwendet man als Synonym häufig auch die Bezeichnung „Operator", „Transformation" oder „Funktion", insbesondere, wenn $Y = \mathbb{R}$ oder $Y = \mathbb{C}$ gilt. Man spricht dann auch von einer „reell-" bzw. „komplexwertigen" Funktion.
Der Begriff „Abbildung" stammt eigentlich aus der Geometrie, wo man spezielle Abbildungen wie *Projektionen*, *Drehungen* oder *Translationen* betrachtet.

10.5.2 Einfache Beispiele und Bemerkungen:

(1) Wir haben §10.5.1 „Arbeitsdefinition" genannt, weil die „Vorschrift" und „Zuordnung" undefiniert sind. Glaubt man an die Mengenlehre und ihre Axiome, so kann man den Begriff „Abbildung" exakt definieren:
Seien X und Y beliebige Mengen. Eine Teilmenge $f \subset X \times Y$ heißt *Abbildung* von X nach Y, wenn es zu jedem $x \in X$ genau ein $y \in Y$ mit $(x, y) \in f$ gibt. Dieses y wird mit $f(x)$ bezeichnet und man schreibt wieder: $f : X \to Y$. Bei dieser Definition wird also eine Funktion mit ihrem Graphen identifiziert.

(2) Abbildungen $f : X \to Y$ werden häufig dadurch definiert, dass man ein „Verfahren"(Rechen-Ausdruck, Algorithmus) angibt, mit dessen Hilfe man $f(x)$ für jedes $x \in X$ aus x berechnen kann (vgl. etwa $f : \mathbb{Z} \to \mathbb{N}_0,\ x \mapsto x^2$).

(3) Die Abbildung (Funktion) $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} x, & \text{falls } x \geq 0, \\ -x, & \text{falls } x < 0, \end{cases}$$

ist eine andere Schreibweise für die „Betragsfunktion"

$$\begin{array}{rrcl} |\,| : & \mathbb{R} & \to & \mathbb{R} \, , \\ & x & \mapsto & |x| \end{array}$$

(4) Die Abbildung

$$\begin{array}{rrcl} id_X : & X & \to & X \, , \\ & x & \mapsto & x \end{array}$$

heißt *Identität* auf X (oder identische Abbildung auf X)

(5) Sind X und Y nicht leer und ist $b \in Y$ fest gewählt, dann ist

$$\begin{array}{ccc} X & \to & Y \\ x & \mapsto & b \end{array}$$

eine *konstante Abbildung* (mit dem Wert b).

(6) Ist X eine *endliche* Menge (vgl. §2.1.6), dann kann man eine Abbildung $f : X \to Y$ dadurch definieren, dass man für jedes $x \in X$ das Bild $f(x) \in Y$ explizit mit einer Wertetafel angibt: $X = \{1, 2, 3\},\ Y = \{1, 2\}$, etwa

$$f(1) = 2, \quad f(2) = 1, \quad f(3) = 2,$$

oder

$$\begin{array}{lcrccc} X = \{1,2,3\} & \qquad & x: & 1 & 2 & 3 \\ & & & \downarrow & \downarrow & \downarrow \\ Y = \{1,2\} & & f(x): & 2 & 1 & 2 \end{array}$$

oder kurz:

x	1	2	3
$f(x)$	2	1	2

(7) Abbildungen spielen auch bei Anwendungen der Mathematik eine zentrale Rolle:

(a) Der Weg s, den ein Körper beim freien Fall zurücklegt, hängt von der Fallzeit t ab:

$$s(t) = \frac{g}{2} t^2$$

(g = Erdbeschleunigung).

(b) Die Schwingungsdauer T eines mathematischen Pendels wird bei kleinen Auslenkungen von der Pendellänge l bestimmt:

$$T(l) = 2\pi \sqrt{\frac{l}{g}}\ .$$

(c) das Volumen V eines idealen Gases hängt vom Druck p und der Temperatur T ab:

$$V = V(p, T) = c \frac{T}{p}$$

(c eine Konstante).

(d) Die Anziehungskraft F zweier Massen m und M hängt von m und M und ihrem Abstand r ab.

$$F = \gamma \frac{mM}{r^2}$$

(γ Gravitationskonstante).

Die Beispiele –auch aus anderen Naturwissenschaften– ließen sich beliebig fortsetzen.

Definitionsbereiche für die genannten Abbildungen (Funktionen) findet man in Physikbüchern selten oder gar nicht.
Sie sollten dann immer davon ausgehen, dass die Autorin / der Autor einen der folgenden Standpunkte einnimmt.

(a) Der Definitionsbereich ergibt sich aus dem Zusammenhang, oder
(b) Die Abbildung(Funktion) ist überall dort definiert, wo der angegebene Funktionsausdruck (Funktionsterm) $f(x)$ sinnvoll gelesen werden kann, z.B.

$$f(x) = \frac{1}{x} \text{ auf } \mathbb{R} - \{0\} .$$

(c) Der Definitionsbereich spielt für das, was über die Abbildung f zu sagen ist, keine Rolle.

Halten wir deshalb nochmals fest:

Zur Definition einer Funktion(Abbildung) gehören ein Definitionsbereich, ein Zielbereich und eine wie auch immer gegebene Zuordnungsvorschrift.

(8) Gleichheit bei Abbildungen:
Zwei Abbildungen $f : X \to Y$ und $g : A \to B$ heißen gleich -in Zeichen $f = g$- wenn $X = A,\ Y = B$ und $f(x) = g(x)$ für alle $x \in X$ gilt.

Es sind z.B. die Abbildungen

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2 \end{aligned} \qquad \text{und} \qquad \begin{aligned} g : \mathbb{R} &\to \mathbb{R}_+ \\ x &\mapsto x^2 \end{aligned}$$

im Sinne unserer Definition verschieden, obwohl sie durch die gleiche Abbildungsvorschrift gegeben sind.
Sie haben auch unterschiedliche Eigenschaften, wie wir gleich sehen werden.

Ob zwei Abbildungen gleich sind, ist manchmal nicht einfach zu entscheiden:
Sei z.B.

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} , \\ x &\mapsto \frac{1}{2}x^2 + x - 2 , \end{aligned} \qquad \begin{aligned} g : \mathbb{R} &\to \mathbb{R} , \\ x &\mapsto \frac{1}{2}(x + 1 + \sqrt{x})(x + 1 - \sqrt{x}) , \end{aligned}$$

dann sind Definitions- und Zielbereich von vornherein gleich, und es gilt $f(x) = g(x)$ für alle $x \in \mathbb{R}$, wie man durch Multiplizieren der rechten Seite im Funktionsausdruck für g sofort sieht.
Komplizierter ist es im folgenden Beispiel:

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} , \\ x &\mapsto \sin 2x , \end{aligned} \qquad \begin{aligned} g : \mathbb{R} &\to \mathbb{R} , \\ x &\mapsto 2 \sin x \cos x . \end{aligned}$$

Aufgrund des *Additionstheorems* für sin gilt aber $\sin 2x = 2 \sin x \cos x$ für alle $x \in \mathbb{R}$, also $f = g$.
Um die Notation jedoch nicht unnötig zu komplizieren, unterscheidet man aber häufig nicht zwischen der Abbildung $f : X \to Y$ und der Abbildung $g : X \to f(X)$, die durch $g(x) := f(x)$ für $x \in X$ definiert ist.

Das hat gute Gründe, wie wir im Zusammen hang mit der Umkehrabbildung sehen werden.

(9) Einschränkung einer Abbildung:
Ist $f : X \to Y$ eine Abbildung und $X_0 \subset X$ eine Teilmenge, so heißt die Abbildung

$$\begin{array}{rcl} f_0 : \ X_0 & \to & Y \, , \\ x & \mapsto & f_0(x) := f(x) \end{array}$$

die *Einschränkung von f auf X_0* und f eine Fortsetzung von f_0. Häufig wird auch hier nicht zwischen f und f_0 unterschieden.

10.5.3 Definition (injektiv, surjektiv, bijektiv)

Eine Abbildung $f : X \to Y$ heißt

(a) *injektiv*, falls verschiedene Punkte aus X stets verschiedene Bilder in y haben:
Für alle $x, x' \in X$ mit $x \neq x'$ gilt $f(x) \neq f(x')$;

(b) *surjektiv*, falls $f(X) = Y$ gilt, falls es also zu jedem $y \in Y$ (mindestens) ein $x \in X$ mit $f(x) = y$ gibt.
Eine surjektive Abbildung $f : X \to Y$ nennt man auch eine Abbildung von X *auf* Y (im Unterschied zum allgemeinen Fall (Abbildung nach oder in Y));

(c) *bijektiv*, falls f injektiv und surjektiv ist. f nennt man in diesem Fall auch eine Bijektion.
Die Abbildung

$$\begin{array}{rcl} f : \ \mathbb{R} & \to & \mathbb{R} \, , \\ x & \mapsto & x^2 \end{array}$$

ist weder injektiv, noch surjektiv, also bestimmt nicht bijektiv. Betrachtet man jedoch die Abbildung

$$\begin{array}{rcl} \tilde{f} : \ \mathbb{R}_+ & \to & \mathbb{R}_+ \, , \\ x & \mapsto & x^2 \end{array}$$

dann ist $\tilde{f}$ injektiv und surjektiv, also bijektiv. Man beachte, dass f und $\tilde{f}$ durch die gleiche Zuordnungsvorschrift definiert sind.

Aus gegebenen Abbildungen kann man neue Abbildungen konstruieren:
Sei z.B. $f : X \to Y$ und $g : X \to Z$ Abbildungen und $f(X) = Y$. Dann kann man auf das Bild $f(x)$ jedes Element $x \in X$ der Abbildung g anwenden und erhält ein Element $g(f(x)) \in Z$. Durch diese Hintereinanderausführung (erst f, dann g) haben wir eine Abbildung von X nach Z konstruiert:

$$\begin{array}{rcl} X & \to & Z \\ x & \mapsto & g(f(x)). \end{array}$$

Wir bezeichnen diese Abbildung mit $g \circ f$ und nennen sie die *Komposition* von f mit g.

Sprechweise: „g nach f“.

$g \circ f$ kann man auch dann bilden, wenn nur $f(X) \subset Y$ gilt.
Man hat dann eigentlich g durch $g_0 = g|f(x)$ zu ersetzen, aber da g_0 und g auf der Teilmenge dieselben Werte haben, erhält man

$$g_0(f(x)) = g(f(x)) \; .$$

10.5.4 Definition

Seien $f : X \to Y$ und $f : Y \to Z$ Abbildungen.

Die Abbildung $h : X \to Z$, die für $x \in X$ durch $h(x) = g(f(x))$ definiert ist, heißt die *Komposition* (Zusammensetzung von f mit g) und man schreibt

$$h = g \circ f \text{ (gelesen } g \text{ nach } f\text{)} .$$

Veranschaulichung im Diagramm:

$$\begin{tikzcd} X \arrow[rr, "h := g \circ f"] \arrow[dr, "f"'] & & Z \\ & Y \arrow[ur, "g"'] & \end{tikzcd} \qquad \text{oder} \qquad \begin{tikzcd} X \arrow[r, "f"] \arrow[rr, bend right, "h := g \circ f"'] & Y \arrow[r, "g"] & Z \end{tikzcd}$$

Beachte: Wenn $g \circ f$ definiert ist, braucht $f \circ g$ nicht definiert zu sein und selbst wenn die Bildung $f \circ g$ möglich ist, gilt i.A. $f \circ g \neq g \circ f$.

Man darf $g \circ f$ auch nicht mir dem Produkt $g \cdot f$ verwechseln, das im Fall reellwertiger Funktionen f und g durch $(g \cdot f)(x) = g(x)f(x)$ definiert ist.

So ist im Fall $X = Y = Z$ und $f(x) = x^2$ und $g(x) = \sin x$ etwa $h(x) = g(f(X)) = \sin(x^2)$, aber das Produkt ist $g(x) \cdot f(x) = \sin(x) \cdot (x^2)$.

Hat man drei Abbildungen $f : X \to Y$, $g : Y \to Z$ und $h : Z \to T$, dann kann man die Zusammensetzung

$$g \circ f : X \to Z, \quad h \circ g : Z \to T, \quad h \circ (g \circ f) : X \to T \quad \text{und} \quad (h \circ g) \circ f : X \to T$$

bilden.

Die Frage ist: Gilt $h \circ (g \circ f) = (h \circ g) \circ f$?

Diagramm:

$$\begin{tikzcd} & X \arrow[dl, "f"'] \arrow[dr, "g \circ f"] \arrow[dd, bend right=80, "(h \circ g) \circ f"'] \arrow[dd, bend left=80, "h \circ (g \circ f)"] & \\ Y \arrow[rr, "g"] \arrow[dr, "h \circ g"'] & & Z \arrow[dl, "h"] \\ & T & \end{tikzcd}$$

Man rechnet für $x \in X$

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

Es ist also in der Tat

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

Man sagt: Die Zusammensetzung von Abbildungen ist assoziativ und lässt die Klammerung weg, schreibt also einfach $h \circ g \circ f$.

Insbesondere kann man eine Abbildung $f : X \to X$ mehrfach mit sich selbst zusammensetzen, also

$$\underbrace{f \circ f \circ f \circ \ldots \circ f}_{n \text{ mal}} =: f^{[n]},\ n \in \mathbb{N},$$

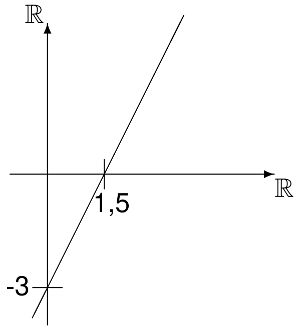
bilden.
Im Fall $X = \mathbb{R}$ darf man $f^{[n]}(x)$ nicht mit der Potenz

$$f^n(x) = \underbrace{f(x) \cdot \ldots \cdot f(x)}_{n \text{ mal}},\ n \in \mathbb{N},$$

verwechseln. Um den folgenden Begriff zu motivieren betrachten wir das folgende einfache

Beispiel: Die (linear-affine) Funktion

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}\,, \\ x &\mapsto 2x - 3 \end{aligned}$$



(siehe Abb.) ist injektiv, wie man sofort nachrechnet. Sie ist auch surjektiv. Denn ist $y \in \mathbb{R}$ beliebig vorgegeben, dann gibt es immer ein (sogar eindeutig bestimmtes) $x \in \mathbb{R}$ mit $f(x) = y$.

Wir können dieses x aus der Gleichung

$$2x - 3 = y$$

einfach ausrechnen und erhalten

$$x = \frac{y+3}{2} \ .$$

Für dieses x gilt tatsächlich

$$f(x) = 2x - 3 = 2 \cdot \frac{x+3}{2} - 3 = y \ .$$

Die Funktion

$$\begin{aligned} g : \mathbb{R} &\to \mathbb{R}\,, \\ y &\mapsto \frac{y+3}{2} \end{aligned}$$

ist wieder bijektiv und hat die Eigenschaft $g \circ f = id_{\mathbb{R}}$ und $f \circ g = id_{\mathbb{R}}$. g und f sind Umkehrfunktionen(Umkehrabbildungen) voneinander im Sinne der folgenden allgemeinen Definition.

10.5.5 Definition (Umkehrabbildung)

Seien X und Y nicht leere Mengen und $f : X \to Y$ eine *bijektive* Abbildung. Dann gibt es eine eindeutig bestimmte Abbildung $g : Y \to X$ mit

$$g \circ f = id_X \quad \text{und} \quad f \circ g = id_Y \ .$$

g ist wieder bijektiv und heißt die *Umkehrabbildung* von f. Sie wird häufig mit f^{-1} (nicht ganz ungefährlich!) oder auch mit $f^{\#}$ bezeichnet.

Beweis :

(1) Zu jedem $y \in Y$ gibt es wegen der Bijektivität von f genau eine $x \in X$ mit $f(x) = y$. Ordnet man jedem $y \in Y$ diese x mit $f(x) = y$ zu, so erhält man eine Abbildung $g : Y \to X$ mit $y = f(g(y)) = f \circ g(y)$ für alle $y \in Y$ und $x = g(f(x)) = g \circ f(x)$ für alle $x \in X$, also

$$f \circ g = id_Y \text{ und } g \circ f = id_X .$$

Wegen der ersten Gleichung ist g injektiv, dann ist $g(y) = g(y'),\ y, y' \in Y$, so folgt

$$y = f(g(y)) = f(g(y')) = y' .$$

Wegen der zweiten Gleichung ist g surjektiv. Denn ist $x \in X$ ein beliebig vorgegebenes Element, dann gilt für das Element $y := f(x) \in Y$

$$g(y) = g(f(x)) = (g \circ f)(x) = x .$$

(2) Zur Eindeutigkeit:
Sei auch $g^* : Y \to X$ eine Abbildung mit $g^* \circ f = id_X$ und $f \circ g^* = id_Y$. Dann gilt

$$\begin{aligned} g^* = g^* \circ id_Y &= g^* \circ (f \circ g) \\ &= (g^* \circ f) \circ g \\ &= id_X \circ g = g. \end{aligned}$$

□

Als Übungsaufgabe zeige man:

10.5.6 Übungsaufgabe

(a) Seien X und Y nicht leere Mengen und $f : X \to Y$ eine Abbildung, für die gilt:
Es gibt eine Abbildung $g : Y \to X$ mit $g \circ f = id_X$ und $f \circ g = id_Y$, dann ist f bijektiv und es gilt $g = f^{\#}$.

(b) Ist $f : X \to Y$ bijektiv und $f^{\#} : Y \to X$ die Umkehrabbildung, dann ist $f^{\#}$ wieder bijektiv und $(f^{\#})^{\#} = f$.

Wir werden nun den Begriff der *bijektiven Abbildung* benutzen, um Mengen der *Größe* nach zu vergleichen. Eine grobe Einteilung der Mengen ist eine Einteilung in

- *endliche Mengen*
- *unendliche Mengen.*

Dabei werden wir sehen, dass es bei den unendlichen Mengen auch noch Abstufungen in der Größenordnung gibt. Es gibt z.B. abzählbar-unendliche Mengen, typisches Beispiel ist die Menge $\mathbb{N}$ der natürlichen Zahlen, und *überabzählbare Mengen*, typisches Beispiel ist die Menge $\mathbb{R}$ aller reellen Zahlen.

Ohne zu zählen, ja ohne überhaupt zählen zu können, kann man feststellen, ob in einer Straßenbahn ebenso viele Sitzplätze wie Fahrgäste vorhanden sind:
Das ist genau dann der Fall, wenn jeder Sitzplatz, mit einem Fahrgast besetzt ist und kein Fahrgast steht.

Etwas vornehmer ausgedrückt:
Wenn es eine bijektive Abbildung der Menge S aller Sitzplätze auf die Menge F aller Fahrgäste gibt. Analog kann man (ohne zu zählen) feststellen, ob es mehr (oder weniger) Sitzplätze als Fahrgäste gibt.

Aufgabe: Denken Sie sich in die Prähistorie zurückgesetzt (Jäger und Sammler!). Sie sitzen vor ihrer Höhle und haben zwei Nusshaufen vor sich und sollen feststellen, welche Nusshaufen mehr Nüsse enthält. Wie würden Sie das (ohne zu zählen) machen?

10.5.7 Definition (Gleichmächtigkeit, G.Cantor, 1878)

Zwei Mengen X und Y heißen *gleichmächtig* (Symbol: $X \sim Y$), wenn es eine bijektive Abbildung $f : X \to Y$ gibt.

Statt der Sprechweise: X und Y sind gleichmächtig (haben dieselbe Mächtigkeit) und der symbolischen Schreibweise $X \sim Y$, sagt man auch: X und Y haben die gleiche *Kardinalzahl*.

Schreibweise: $\#X = \#Y$.

10.5.8 Satz (Eigenschaften von „$\sim$“)

Seien X, Y, Z Mengen. Dann gilt

(a) $X \sim X$ (Reflexivität)

(b) $X \sim Y \Longrightarrow Y \sim X$ (Symmetrie)

(c) $X \sim Y$ und $Y \sim Z \Longrightarrow X \sim Z$ (Transitivität)

In einem festen System von Mengen ist also „$\sim$“ eine Äquivalenzrelation und äquivalente Mengen haben gleiche Mächtigkeit (Kardinalzahl).

Der **Beweis** ist offensichtlich, da $f = id_X$ bijektiv ist (a) und mit f auch die Umkehrabbildung $g = f^{\#}$ bijektiv ist (b).
Sind $f : X \to Y$ und $g : Y \to Z$ bijektiv, dann ist auch die Zusammensetzung

$$g \circ f : X \to Z$$

bijektiv, also gilt auch (c). $\square$

10.5.9 Definition

Eine Menge X heißt

(a) *endlich*, falls $X \neq \emptyset$ oder falls es ein $n \in \mathbb{N}$ gibt mit $\mathbb{A}_n \sim X$.
Dabei ist $\mathbb{A}_n := \{1, 2, \ldots, n\},\ n \in \mathbb{N}$.

(b) *unendlich*, falls sie nicht endlich ist.

Wir haben schon früher (vgl. §2.1.6) betont, dass wir die Mengen $\mathbb{A}_N := \{1, \ldots, n\},\ n \in \mathbb{N}$, als Prototypen von endlichen Mengen mit n Elementen betrachten. Hier ist $\#\mathbb{A}_n = n$, es sei $\mathbb{A}_0 := \emptyset$ und $\#\mathbb{A}_0 := 0$. Ist X endlich, dann ist die Kardinalzahl (Mächtigkeit) $\#X$ eindeutig definiert, denn es gilt

10.5.10 Bemerkung

Gibt es eine bijektive Abbildung $f : \mathbb{A}_n \to X$ und eine bijektive Abbildung $g : \mathbb{A}_m \to X$, dann ist $n = m$ $(n, m \in \mathbb{N})$.

Zum **Beweis** hat man lediglich zu zeigen:
Ist $\varphi : \mathbb{A}_n \to \mathbb{A}_m$ bijektiv, dann ist $n = m$.

Hier bietet sich ein Induktionsbeweis, etwa nach n, an.
Sei also für beliebiges $n \in \mathbb{N}$ $A(n)$ die Eigenschaft: Für $n \in \mathbb{N}$ und beliebiges $m \in \mathbb{N}$ gilt: Gibt es eine bijektive Abbildung $\varphi : \mathbb{A}_n \to \mathbb{A}_m$, dann ist $n = m$.

A(1): Ist $n = 1$, dann ist notwendig auch $m = 1$. Für den Induktionsschritt unterscheiden wir zwei Fälle:

1.Fall: Es ist $\varphi(n+1) = m$. Dann können wir eine Abbildung $g : \mathbb{A}_n \to \mathbb{A}_{m-1}$ dadurch definieren, dass wir $g(k) := f(k)$ für alle $k \leq n$ setzen, denn für kein $k \leq n$ kann $f(k)$ mit m übereinstimmen, da f bijektiv ist. g ist dann auch bijektiv und nach der Induktionsvoraussetzung ist $n = m - 1$, also auch $n + 1 = m$.

2.Fall: $\varphi(n+1) < m$.
Dann gibt es ein $k < n + 1$ mit $\varphi(k) = m$.

$$\tau:\quad \begin{array}{l} \{1, 2, \ldots, k, ..., n+1\} \\ \;\downarrow\downarrow \qquad\quad \times \\ \{1, 2, \ldots, k, ..., n+1\} \end{array}$$

Betrachte die Abbildung $\tau : \mathbb{A}_{n+1} \to \mathbb{A}_{n+1}$ mit $\tau(k) = n + 1,\ \tau(n+1) = k$ und $\tau(j) = j$ für $j \leq n, j \neq k$.
τ vertauscht also die Zahlen $n + 1$ und k und lässt alle übrige Zahlen fest (Transposition). τ ist bijektiv ($\tau \circ \tau = id_{\mathbb{A}_{n+1}}$) und damit ist auch

$$g := \varphi \circ \tau : \mathbb{A}_{n+1} \to \mathbb{A}_m$$

bijektiv und es gilt

$$g(n+1) = \varphi(\tau(n+1)) = \varphi(k) = m\ .$$

Nach dem ersten Fall folgt $n + 1 = m$. $\square$

Als **Übung** zeigen man:

(a) Ist Y Teilmenge der endlichen Menge X, dann ist Y ebenfalls endlich und $\#Y \leq \#X$.

(b) Sind X und Y disjunkte endlichen Mengen, d.h. es gilt $X \cap Y = \emptyset$, dann ist auch $X \cup Y$ endlich (das gilt auch ohne die Disjunktheit), und es ist

$$\#(X \cup Y) = \#X + \#Y\ .$$

(c) Sind X und Y endliche Mengen, dann ist auch das kartesische Produkt $X \times Y$ endlich, und es gilt

$$\#(X \times Y) = \#X \cdot \#Y\ .$$

Offensichtlich ist die Menge $\mathbb{N}$ der natürlichen Zahlen eine unendliche Menge. Denn wäre $\mathbb{N}$ endlich, so gäbe es ein $n \in \mathbb{N}$ mit $\mathbb{A}_n \sim \mathbb{N}$.

Wegen $\mathbb{A}_{n+1} \subset \mathbb{N}$ müsste dann $n+1 \leq n$ sein. Offensichtlich ist dann auch $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ eine unendliche Menge und die Abbildung

$$\begin{array}{rcl} f: \mathbb{N}_0 & \to & \mathbb{N}\,, \\ n & \mapsto & n+1 \end{array}$$

ist bijektiv, also ist im Sinne unserer Definition $\#\mathbb{N}_0 = \#\mathbb{N}$.

Noch erstaunlicher ist, dass im Sinne der Kardinalzahlen nicht *mehr* natürliche Zahlen als gerade natürliche Zahlen gibt, denn die Abbildung

$$\begin{array}{rcl} f: \mathbb{N}_0 & \to & 2\mathbb{N}_0\,, \\ n & \mapsto & 2n \end{array}$$

ist bijektiv, also ist wieder $\#\mathbb{N}_0 = \#(2\mathbb{N}_0)$.

Im Bereich der unendlichen Mengen kann also –im Gegensatz zu endlichen Mengen– Teil aus einer Menge zu einer echten Teilmenge äquivalent sein.

10.5.11 Definition (Abzählbarkeit)

Eine Menge X heißt

(a) *abzählbar unendlich*, wenn $\mathbb{N} \sim X$ gilt,

(b) *abzählbar*, falls X endlich oder abzählbar unendlich ist,

(c) *überabzählbar*, falls X nicht abzählbar ist.

Man beachte, dass die Terminologie in der Literatur nicht einheitlich ist.

Eine Menge X, die endlich oder abzählbar unendlich ist, nennt man häufig auch eine „höchstens abzählbare“ Menge. Manche Autoren verwenden den Begriff „abzählbar“ auch im Sinne von „abzählbar unendlich“.

10.5.12 Beispiele und Bemerkungen

(1) Wir wissen schon:

$$\mathbb{N}_0 \sim \mathbb{N} \quad \text{oder} \quad \#\mathbb{N}_0 = \#\mathbb{N}$$

(2) Erstaunlicherweise gibt es aber auch „nicht mehr“ ganze Zahlen als natürliche Zahlen.
Dass eine Menge (abzählbar) unendlich ist, behandelt nicht anderes, als dass man ihre Elemente als Folge schreiben kann und die Folgenglieder paarweise verschieden sind.
Die folgende Tabelle liefert eine *Nummerierung* der ganzen Zahlen durch die natürliche Zahlen

$\mathbb{N}_0$	0	1	2	3	4	$\dots$	$2k-1$	$2k$	$\dots$
$\mathbb{Z}$	0	1	-1	2	-2	$\dots$	k	$-k$	$\dots$

Etwas formaler: Die Abbildung $f : \mathbb{N}_0 \to \mathbb{Z}$ mit

$$f(n) := \begin{cases} 0, & \text{falls } n = 0 \\ k, & \text{falls } n = 2k-1 \\ -k, & \text{falls } n = 2k \end{cases}$$

für $k \in \mathbb{N}$ ist bijektiv.

(3) Auch die Menge $\mathbb{Q}$ aller rationaler Zahlen ist abzählbar (unendlich), also gilt auch $\#\mathbb{Q} = \#\mathbb{N}$. Es gibt also nicht „mehr“ (im Sinne der Kardinalzahlen) rationaler Zahlen als natürlicher Zahlen.

Zum **Beweis** ordnet man zunächst die *positiven* rationalen Zahlen im folgenden Schema an:

$$
\begin{array}{cccccccccc}
\frac{1}{1} & \rightarrow & \frac{2}{1} & & \frac{3}{1} & \rightarrow & \frac{4}{1} & & \frac{5}{1} & \rightarrow \cdots \\
 & \swarrow & & \nearrow & & \swarrow & & \nearrow & & \\
\frac{1}{2} & & \frac{2}{2} & & \frac{3}{2} & & \frac{4}{2} & \cdots & & \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & & & \\
\frac{1}{3} & & \frac{2}{3} & & \frac{3}{3} & \cdots & & & & \\
 & \swarrow & & \nearrow & & & & & & \\
\frac{1}{4} & & \frac{2}{4} & \cdots & & & & & & \\
\downarrow & \nearrow & & & & & & & & \\
\frac{1}{5} & \cdots & & & & & & & & \\
\vdots & & & & & & & & &
\end{array}
$$

und nummeriert sie in Richtung der angegebenen Pfeile, wobei man die ungekürzten Brüche (wie z.B. $\frac{1}{2}$ oder $\frac{2}{4}$ oder $\frac{4}{2}$ oder $\frac{3}{3}$ etc.) überspringt, um „Mehrfachaufzählungen“ zu vermeiden. In der Weise erhält man eine Abzählung von $\mathbb{Q}_+^\bullet$. $\mathbb{Q}_+^\bullet$ erscheint als

$$\mathbb{Q}_+^\bullet = (r_1, r_2, r_3, \ldots) = \left(1, 2, \frac{1}{2}, \frac{1}{3}, 3, 4, \frac{3}{2}, \frac{2}{3}, \frac{1}{4}, \frac{1}{5}, \ldots\right) .$$

Dann ist aber

$$(0, r_1, -r_1, r_2, -r_2, r_3, \ldots)$$

eine Anordnung von $\mathbb{Q}$ als Folge, damit ist $\mathbb{Q}$ abzählbar (unendlich).

(4) *Verallgemeinerung*:
Die Vereinigung höchstens abzählbar vieler höchstens abzählbaren Mengen ist wieder höchstens abzählbar.

Beweis: Man muss den Beweis von (3) ein klein wenig modifizieren, da wir zugelassen haben, dass die beteiligten Mengen auch endlich sein können.

(5) $\mathbb{N}_0 \times \mathbb{N}_0$ ist abzählbar (unendlich). Dies hatten wir im Zusammenhang mit Produktreihen benutzt, deshalb geben wir hier einen Beweis:

Man benutzt (ähnlich wie bei (4)) das erste (oder Cauchy'sche) Diagonalverfahren zur Abzählung der Paare $(m, n) \in \mathbb{N}_0 \times \mathbb{N}_0$ in dem doppelt unendlichen Schema

$$
\begin{array}{ccccccccccc}
(0,0) & & (0,1) & & (0,2) & & (0,3) & & \cdots & (0,n) & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & & \swarrow & & \\
(1,0) & & (1,1) & & (1,2) & & \cdots & & & & \\
 & \swarrow & & \swarrow & & & & & & & \\
(2,0) & & (2,1) & & \cdots & & & & & & \\
 & \swarrow & & & & & & & & & \\
(3,0) & & \cdots & & & & & & & & \\
\vdots & & & & & & & & & & \\
 & \swarrow & & & & & & & & & \\
(n,0) & & \cdots & & & & & & & & \\
\vdots & & & & & & & & & &
\end{array}
$$

$$0 \mapsto (0,0),\ 1 \mapsto (0,1),\ 2 \mapsto (1,0),\ 3 \mapsto (0,2),\ 4 \mapsto (1,1),\ 5 \mapsto (2,0),\ 6 \mapsto (0,3),\ \ldots$$

Einfacher ist es eine bijektive Abbildung $\varphi : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$, die so genannte „Cantor'sche Paarungsfunktion" anzugeben. Diese ist definiert durch

$$\varphi(m,n) = \frac{1}{2}(m+n)(m+n+1) + m \ ,$$

also z.B. $\varphi(3,2) = \frac{30}{2} + 3 = 18$.
Weisen Sie als Übungsaufgabe nach, dass die Abbildung φ bijektiv ist.
Eine andere Bijektion ist

$$\begin{aligned} \psi :\ \mathbb{N}_0 \times \mathbb{N}_0 &\to \mathbb{N}_0 \ , \\ (m,n) &\mapsto 2^m(2n+1) - 1 \ , \end{aligned}$$

dabei gibt m an, wie oft $\psi(m,n)+1$ den Faktor 2 enthält und $2n+1$ ist der ungerade Restfaktor.

(6) **Bemerkung**:
Zur Abkürzung wird für $\#\mathbb{N}_0$ der hebräische Buchstabe $\aleph_0$ (Aleph Null) verwendet. Die merkwürdige und befremdliche Arithmetik im Bereich der Kardinalzahlen, die z.B. in

$$\begin{aligned} \aleph_0 + 1 &= \aleph_0 && \text{oder auch} \\ \aleph_0 + 1000 &= \aleph_0 && \text{oder auch} \\ \aleph_0 + \aleph_0 &= \aleph_0 \end{aligned}$$

zum Ausdruck kommt, wurde in der Vorlesung im Rahmen der „mathematischen Folklore" am Beispiel von *Hilberts Hotel* illustriert. (Es kommt 1 weiterer Gast, es kommen nochmals 1000 Gäste, es kommen nochmals abzählbar unendlich viele Gäste, es kommen abzählbar viele Busse mit jeweils abzählbar vielen Gästen.)

Wenn also abzählbar unendliche Vereinigungen von abzählbar unendlichen Mengen wieder abzählbar sind, ist dann überhaupt zu erwarten, dass es überabzählbare Mengen gibt? Ihre Existenz zeigt der folgende Satz.

10.5.13 Satz

Ist $A := \{0,1\}$ und $F := Abb(\mathbb{N}_0, A)$ die Menge aller 0-1-Folgen, dann ist F überabzählbar.

Beweis : F besteht aus Folgen, als deren Glieder nur die Elemente 0 oder 1 auftreten, z.B. ist

$$(1,0,0,0,\ldots) \in F \ \text{ oder } \ (0,1,0,0,\ldots) \in F \ \text{ oder } \ (1,1,0,0,0,\ldots) \in F \ .$$

Wäre F nun abzählbar, dann müsste sich F als Folge (von Folgen) schreiben lassen, etwa

$$F = (f_0, f_1, f_2, \ldots) \ ,$$

wobei jedes f_j eine 0-1-Folge ist, etwa ($j \in \mathbb{N}_0$)

$$\begin{aligned} f_0 &= (f_{00}, f_{01}, f_{02}, \ldots) && \text{mit } f_{oj} \in \{0,1\} \\ f_1 &= (f_{10}, f_{11}, f_{12}, \ldots) && \text{mit } f_{1j} \in \{0,1\} \\ f_2 &= (f_{20}, f_{21}, f_{22}, \ldots) && \text{mit } f_{2j} \in \{0,1\} \\ &\vdots \\ f_n &= (f_{n0}, f_{n1}, f_{n2}, \ldots) && \text{mit } f_{nj} \in \{0,1\} \\ &\vdots \end{aligned}$$

Wir konstruieren nun ein Element $g \in F$, d.h. eine 0-1-Folge, die in obiger Auflistung *nicht* vorkommt. Wir setzen nämlich für $n \in \mathbb{N}_0$

$$g_n = \begin{cases} 1, & \text{falls } f_{nn} = 0, \\ 0, & \text{falls } f_{nn} = 1, \end{cases}$$

und betrachten $g = (g_0, g_1, \ldots, g_n, \ldots)$, dann ist offensichtlich $g \in F$, aber g kommt in obiger Aufzählung nicht vor (weil $g_n \neq f_{nn}$ für alle $n \in \mathbb{N}_0$ gilt).
Dieser Widerspruch zeigt, dass F in der Tat überabzählbar ist.

□

Einer Folge $f = (f_0, f_1, \ldots) \in F$ kann man den Dezimalbruch

$$x_f := \frac{f_1}{10^1} + \frac{f_2}{10^2} + \ldots + \frac{f_n}{10^n} + \ldots \in [0,1[$$

zuordnen.
Da die Dezimalbruchdarstellung einer reeller Zahl $r \in [0,1[$ eindeutig ist, ist diese Abbildung injektiv. Daher ist das Bild $\varphi(F) \subset [0,1[$ eine überabzählbare Teilmenge, also ist $[0,1[$ selber überabzählbar.
Offensichtlich ist dann auch $]0,1[$ überabzählbar und da durch

$$\begin{array}{rcl} \varphi :]0,1[& \rightarrow & \mathbb{R} \\ x & \mapsto & ??? \end{array}$$

eine bijektive Abbildung definiert wird, haben wir:

10.5.14 Satz

Das Intervall $]0,1[$ und die Menge $\mathbb{R}$ der reellen Zahlen sind überabzählbar.

10.5.15 Aufgaben

(a) Zu je zwei offenen Intervallen

$$]a,b[\ (a,b \in \mathbb{R},\ a < b) \qquad \text{und} \qquad]c,d[\ (c,d \in \mathbb{R},\ c < d)$$

gebe man eine Bijektion $f :]a,b[\rightarrow]c,d[$ an.

Kann man die Bijektion auf die abgeschlossene Intervalle festsetzen?

(b) Man zeige: Alle echten Intervalle (d.h. Intervalle, die mindestens 2 Punkte enthalten) sind überabzählbar und haben die gleiche Mächtigkeit wie $\mathbb{R}$.

10.5.16 Folgerung

Die Menge $\mathbb{R} - \mathbb{Q}$ der irrationalen Zahlen ist überabzählbar.

Beweis: Denn wenn $\mathbb{R} - \mathbb{Q}$ abzählbar wäre, dann wäre auch $(\mathbb{R} - \mathbb{Q}) \cup \mathbb{Q} = \mathbb{R}$ abzählbar.

10.5.17 Definition (algebraische Zahl)

Eine komplexe Zahl α heißt *algebraisch*, wenn es ein $n \in \mathbb{N}_0$ und Zahlen $a_0, \ldots, a_n \in \mathbb{Z}$ gibt, die nicht alle gleich Null sind, so dass

$$a_n\alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_0 = 0$$

gilt.
M.a.W.: α ist genau dann algebraisch, wenn es ein Polynom $P \in \mathbb{Z}[X],\ P \not\equiv 0$ gibt mit $P(\alpha) = 0$.

$$\mathbb{A} := \{\alpha \in \mathbb{C};\ \alpha \text{ algebraisch}\}$$

ist ein Körper (Unterkörper von $\mathbb{C}$; Der Beweis ist nicht ganz elementar!) und heißt Körper der *algebraischen Zahlen*.

Eine komplexe Zahl, die nicht algebraisch ist, heißt *transzendent*.

10.5.18 Aufgabe

Man erhält die selbe Menge $\mathbb{A}$, wenn man Polynome $Q \in \mathbb{Q}[X],\ Q \not\equiv 0$, zur Definition verwendet.

10.5.19 Beispiele

(a) Die imaginäre Einheit i ist algebraisch, da für $P = X^2 + 1 \in \mathbb{Z}[X]$ gilt

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

(b) Die achte Einheitswurzel

$$\zeta_8 := exp\left(\frac{2\pi i}{8}\right)$$

ist algebraisch, da für $P = x^8 - 1$ gilt $P(\zeta_8) = 0$.

(c) Jede rationale Zahl ist algebraisch, denn ist $r \in \mathbb{Q}$, dann ist r Nullstelle des Polynoms $P = x - r$.

(d) Die nicht rationale Zahl (Goldener Schnitt!)

$$g := \frac{1+\sqrt{5}}{2} = 1,618033988\ldots$$

ist algebraisch, denn sie ist Nullstelle des Polynoms

$$P = X^2 - X - 1 \text{ (nachprüfen!)}$$

(e) $\alpha; = \sqrt{2} + \sqrt{3}$ ist algebraisch, denn α ist eine Nullstelle des Polynoms

$$P = X^4 - 10X^2 + 1 \text{ (nachrechnen!)}$$

Gibt es überhaupt komplexe Zahlen die nicht algebraisch sind, d.h. gibt es transzendente Zahlen?
Ohne dass man jemals eine transzendente Zahl gesehen hat, ergibt sich ihre Existenz am

10.5.20 Satz

Die Menge $\mathbb{A}$ aller algebraischen Zahlen ist abzählbar(unendlich)

Beweis : Für $n \in \mathbb{N}_0$ sei

$$A_n := \{\alpha \in \mathbb{A};\ \alpha \text{ ist Nullstelle eins nicht-trivialen Polynoms } P \in \mathbb{Z}[X] \text{ vom Grad } \leq n, \text{ für dessen Koeffizienten gilt } |a_n| + |a_{n-1}| + \ldots + |a_0| \leq n.\}$$

Nach dem Fundamentalsatz der Algebra hat ein Polynom vom Grad n höchstens n Nullstellen. A_n ist daher abzählbar, sogar endlich. Offensichtlich gilt

$$\mathbb{A} = \bigcup_{n \in \mathbb{N}_0} A_n \,.$$

Als abzählbare Vereinigung endlicher Mengen ist $\mathbb{A}$ daher abzählbar.

□

Der folgende Satz zeigt, dass die „Masse“ der komplexen Zahlen transzendent ist.

Satz: Die Menge $\mathbb{T} = \{z \in \mathbb{C};\ z \text{ transzentent } \}$ ist überabzählbar.

Beweis: Es gilt $\mathbb{C} = \mathbb{A} \cup \mathbb{T}$. Da $\mathbb{A}$ abzählbar, ist die Menge $\mathbb{T}$ der transzendenten Zahlen überabzählbar (sonst wäre $\mathbb{C}$ abzählbar!).

Bemerkung:
Der Nachweis, ob eine konkrete reelle oder komplexe Zahl transzendent ist, ist im allgemeinen ein sehr schwieriges Problem.

Im Jahr 1851 konnte J.Liouoille zeigen, dass die (irrationale!) Zahl

$$0,110001000000000000000000010\ldots$$

(nach welchem Muster ist die Zahl gebildet?) transzendent ist. C.Hermite gelang 1872 der Nachweis, dass die Eulersche Zahl $e := \sum\limits_{k=0}^{\infty} \frac{1}{k!}$ transzendent ist und C.L.F.Lindenmann konnte 1882 zeigen, das die Kreiszahl π transzendent ist. Im Jahr 1934 haben C.L.Siegel und A.O.Gelfond unabhängig voneinander gezeigt:

10.5.21 Satz (Siegel-Gelfond)

Ist a algebraisch, $a \neq 0,\ a \neq 1$, und ist b irrational und algebraisch, dann ist a^b transzendent.
Beispielsweise ist also $2^{\sqrt{2}}$ eine transzendente Zahl ($2^{\sqrt{2}} := exp(\sqrt{2}\log 2)$).

Den Siegel'schen Beweis findet man in C.L.Siegel „Transzendente Zahlen“, BI-Taschenbücher, Nr.137, Bibliograph. Institut, Mannheim 1967.

Nach dieser kurzen Lektüre werden Sie feststellen, dass explizite Transzendenzbeweise ein hartes Geschäft sind.

Da $\mathbb{N}$ bzw $\mathbb{N}_0$ als Prototypen einer unendlichen Menge abzählbar ist, die Menge $\mathbb{R}$ der reellen Zahlen aber überabzählbar ist, ist man geneigt zu sagen, dass es „mehr“ reelle Zahlen als natürliche oder auch rationale Zahlen gibt. das „mehr“ kann man mit folgenden Definition präzisieren:

10.5.22 Definition (strikt kleinere Mächtigkeit)

Seinen X und Y Mengen. Man sagt X ist von strikt kleineren Mächtigkeit als Y, in Zeichen $\#X < \#Y$, falls es eine injektive Abbildung $\varphi : X \to Y$ aber keine bijektive Abbildung von X auf Y gibt.

Beispiel: $\#\mathbb{N} = \#\mathbb{Q} < \#\mathbb{R}$.

Frage: Kann man zu jeder Menge X eine Menge Y mit $\#X < \#Y$ konstruieren?

Ist X eine endliche Menge, etwa

$$X = \mathbb{A}_n := \{1, 2, \ldots, n\},\ n \in \mathbb{N}\ ,$$

und ist Y die Potenzmenge von X, also die Menge aller Teilmengen, dann hat Y 2^n Elemente, die Elementenanzahl wächst also exponentiell.

G.Cantor hat nun gezeigt, dass für *jede* Menge X die Potenzmenge $P(X) := \{A;\ A \subset X\}$ eine echt größere Mächtigkeit als X hat.

10.5.23 Satz (von Cantor über die Mächtigkeit der Potenzmenge)

Ist X eine beliebige Menge. $P(X) = \{A,\ A \subset X\}$ ihre Potenzmenge, dann gilt $\#X < \#P(X)$.

Beweis : OBdA können wir $X \neq \emptyset$ voraussetzen. Wir betrachten die Menge F der auf X definierten Funktionen, die nur die Werte 0 und 1 annehmen.

Ordnet man jedem $x \in M$ die Funktion $f_x : M \to \{0, 1\}$ mit

$$f_x(y) = \begin{cases} 1, & \text{falls } y = x, \\ 0, & \text{falls } y \neq x, \end{cases}$$

zu, dann ist $X \sim \{f_x;\ x \in X\} \subset F$. X ist jedoch nicht zur ganzen Menge F äquivalent. Denn sonst gäbe es eine bijektive Abbildung $x \mapsto f^{(x)}$ von X auf F. Die Funktion $f : X \to \{0, 1\}$ mit

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

gehört zu F, stimmt aber mit keinem $f^{(x)}$ überein.

X und F sind also nicht gleich mächtig und F hat eine größere Mächtigkeit als X. Definiert man nun als *charakteristische* Funktion einer Teilmenge $A \subset X$ die Funktion

$$\chi_A(x) = \begin{cases} 1, & \text{falls } x \in A, \\ 0, & \text{falls } x \notin A, \end{cases}$$

dann liefert offenbar die Abbildung $A \mapsto \chi_A$ eine Bijektion von

$$P(X) \to \{\chi_A;\ A \subset X\} = F\ .$$

$\square$

Bemerkung:
Die Konstruktion kann man natürlich wiederholen und etwa $P(P(X))$ bilden etc. Sie zeigt, dass im Bereich der unendlichen Mengen noch unendlich viele Abstufungen in der Mächtigkeit dieser Mengen gibt.

D.Hilbert hat 1900 auf dem Internationalen Mathematikkongress in Paris die Frage gestellt:
Um wieviel größer als $\mathbb{N}$ ist $\mathbb{R}$?

Die so genannte *Kontinuumshypothese (CH)* besagt, dass die Kluft zwischen $\mathbb{N}$ und $\mathbb{R}$ minimal ist. Sie besagt nämlich:

Kontinuumshypothese (CH; G.Cantor, 1878): Es gibt keine Menge X mit $\#\mathbb{N} < \#X < \#\mathbb{R}$.

Man kann die Cantor'sche Kontinuumshypothese auch so ausdrücken: Jede Teilmenge $M \subset \mathbb{R}$ ist abzählbar oder gleichmächtig zu $\mathbb{R}$.

Die Antwort auf die Frage, ob CH richtig oder falsch ist, ist beunruhigend, denn es gilt

Fundamentalsatz der Mengenlehre:
Die Kontinuumshypothese ist weder beweisbar noch widerlegbar (im Rahmen des der Mengenlehre üblicherweise zu Grunde gelegten Axiomensystems).

Der Beweis beruht auf zwei Säulen:
Bereits 1938 hat Kurt Gödel gezeigt, das $\neg$(CH) nicht beweisbar ist, d.h. CH ist nicht widerlegbar, und 1963 hat Paul Cohen gezeigt, dass CH selber nicht beweisbar ist.

III. Stetigkeit, Grenzwerte bei Funktionen

„Natura non facit saltus“ (Die Natur macht keine Sprünge), dieser Anspruch von Raoul Fournier (1627) galt lange bei der mathematischen Behandlung von Naturvorgängen. Bei der Anwendung des Funktionsbegriffs auf Probleme der Naturwissenschaften z.B. auf Bewegungsvorgänge in der Physik, aber natürlich auch bei innermathematischen Anwendungen des Funktionsbegriffs, kann man feststellen, dass die betrachteten Funktionen (Abbildungen), häufig die folgende Eigenschaft haben:
Ändert man das Argument x einer Funktion f nur wenig, dann ändern sich auch die Funktionswerte nur wenig und zwar ändert sich $f(x)$ *beliebig wenig*, wenn sich x *genügend wenig* ändert. Dies kann man in der ε - δ - Sprache genau ausdrücken, wir kommen darauf zurück, wir geben zurecht für die Stetigkeit eine andere -gleichwertige- Definition, die auf dem Folgenbegriff basiert. Der Begriff des *Grenzwertes einer Funktion* führen wir auf den Stetigkeitsbegriff zurück.

11 Der Begriff der Stetigkeit

Wir betrachten im Folgenden Funktionen, deren Definitionsbereich D i.A. eine nicht leere Teilmenge von $\mathbb{C}$ und deren Werte in $\mathbb{K}$ liegen, dabei ist $\mathbb{K}$ wie bisher immer entweder $\mathbb{R}$ oder $\mathbb{C}$. Verallgemeinerungen, etwa auf Abbildungen

$$f : D \to \mathbb{R}^m;\ D \subset \mathbb{R}^n,\ D \neq \emptyset,$$

liegen auf der Hand. Systematisch kommen wir hierauf aber erst später zurück.

11.1 Definition (Folgenstetigkeit)

Sei $D \subset \mathbb{C},\ D \neq \emptyset$ und $f : D \to \mathbb{K}$ eine Funktion.
f heißt stetig an der Stelle $a \in D$ oder im Punkt $a \in D$, wenn für *jede* Folge (x_n) mit $x_n \in D$ und $\lim\limits_{n\to\infty} x_n = a$, die Bildfolge $(f(x_n))$ gegen $f(a)$ konvergiert:

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

f heißt stetig schlechthin oder stetig auf D, wenn f an *jeder* Stelle $a \in D$ stetig ist.
Ist f nicht stetig an der Stelle $a \in D$, dann heißt f *unstetig* in a.

Bemerkungen:

(a) Wegen $a = \lim\limits_{n\to\infty} x_n$, kann man die die Stetigkeit charakterisierende Bedingung auch als

$$\boxed{\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right)}$$

schreiben: Grenzwertbildung und Funktionauswertung dürfen also vertauscht werden.

(b) Die hier definierte Stetigkeit ist eigentlich die *Folgenstetigkeit*, sie ist jedoch mit der ε - δ - Stetigkeit äquivalent (vgl. §11.2).

Aus den Rechenregeln für konvergente Folgen (vgl. §8) folgt sofort:

11.1.1 Satz

Die reelle oder komplexe Betragsfunktion, alle Polynome und rationale Funktionen sind in ihren jeweiligen Definitionsbereichen stetig.

Wir wiederholen den **Beweis** etwa für Polynome: Sei

$$\begin{array}{rcl} P: \mathbb{C} & \to & \mathbb{C}\,, \\ z & \mapsto & a_m z^m + a_{m-1} z^{m-1} + \ldots + a_o \end{array}$$

($a_j \in \mathbb{C}$) ein Polynom; $m \in \mathbb{N}_0,\ a \in \mathbb{C}$ und $(z_n), z_n \in \mathbb{C}$, eine Folge mit $\lim\limits_{n\to\infty} z_n = a$.

Dann gilt auch $\lim\limits_{n\to\infty} z_n^2 = a^2$ und allgemeiner für jedes $m \in \mathbb{N}_0$

$$\lim_{n\to\infty} z_n^m = a^m$$

und damit auch

$$\begin{array}{rcl} \lim\limits_{n\to\infty} P(z_n) & = & \lim\limits_{n\to\infty} (a_m z_n^m + a_{m-1} z_n^{m-1} + \ldots + a_0) \\ & = & \lim\limits_{n\to\infty} a_n z_n^m + \lim\limits_{n\to\infty} a_{m-1} z_n^{m-1} + \ldots + a_0 \\ & = & a_m a^m + a_{m-1} a^{m-1} + \ldots + a_0 \\ & = & P(a)\ . \end{array}$$

Wir geben für die Stetigkeit der m-ten Potenz $(m \in \mathbb{N})$

$$\begin{array}{rcl} p_m: \mathbb{C} & \to & \mathbb{C}\,, \\ z & \mapsto & z^m \end{array}$$

einen direkten **Beweis**:

Sei $a \in \mathbb{C}$ und (z_n) eine Folge mit $\lim\limits_{n\to\infty} z_n = a$.
Als konvergente Folge ist die Folge (z_n) beschränkt, es gilt etwa $|z_n| \leq C$ für alle $n \in \mathbb{N}$. Dann gilt auch $|a| \leq C$. Es gilt dann

$$\begin{array}{rcl} |p_m(z_n) - p_m(a)| & = & |z_n^m - a^m| \\ & = & |(z_n - a) \sum\limits_{j=1}^{m} z_n^{j-1} a^{m-j}| \\ & \leq & |z_n - a| \sum\limits_{j=1}^{m} C^{j-1} C^{m-j} \\ & = & K|z_n - a| \ \text{ mit } K := \sum\limits_{j=1}^{m} C^{j-1} C^{m-j}\ . \end{array}$$

Weil die rechte Seite gegen Null konvergiert, konvergiert auch die linke Seite gegen Null, d.h. es gilt

$$\lim_{n\to\infty} p_m(z_n) = p_m(a) = a^m\ .$$

11.1.2 Weitere Beispiele und Bemerkungen

(a) Die Dirichlet-Funktion $\delta : [0,1] \to \mathbb{R}$ mit

$$\delta(x) = \begin{cases} 1, & \text{falls } x \text{ rational,} \\ 0, & \text{falls } x \text{ irrational,} \end{cases}$$

ist an keiner Stelle $a \in [0,1]$ stetig. Ist nämlich $a \in [0,1]$ rational, so wählen wir eine Folge (x_n) $\lim\limits_{n\to\infty} x_n = a$, für welche die x_n irrational sind. Dann gilt $f(x_n) = 0$, also auch

$$\lim_{n\to\infty} f(x_n) = 0 \neq 1 = f(a).$$

Ist $a \in [0,1]$ irrational, so wählen wir eine Folge (x_n) mit $\lim\limits_{n\to\infty} x_n = a$, für welche die x_n rational sind. Dann gilt $f(x_n) = 1$, also auch

$$\lim_{n\to\infty} f(x_n) = 1 \neq 0 = f(a).$$

(b) Die Riemann-Funktion $f : [0,1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} \frac{1}{q_x}, & \text{falls } x \text{ rational und } x = \frac{p_x}{q_x} \text{ (Darstellung gekürzt:} (p_x, q_x) \in \mathbb{N}_0 \times \mathbb{N}), \\ 0, & \text{sonst,} \end{cases}$$

ist unstetig in allen rationalen Punkten aus $[0,1]$ und stetig in allen irrationalen Punkten aus $[0,1]$.

Der Beweis der Unstetigkeit in den rationalen Punkten verläuft nach dem gleichen Muster, wie bei Beispiel (a):
Ist $a \in [0,1] \cap \mathbb{Q}$, dann wählt man eine Folge (x_n) mit $x_n \in [0,1]$, x_n irrational und $\lim\limits_{n\to\infty} x_n = a$. Dann ist $f(x_n) = 0$, also auch $\lim\limits_{n\to\infty} f(x_n) = 0$.

Der Beweis, dass f in allen irrationalen Punkten $a \in [0,1]$ stetig ist, sei als Übungsaufgabe gestellt (vgl. auch Übungsaufgabe ???).
Versuchen Sie, den Graphen von f für wenigstens für einige Argumente $x \in [0,1]$ zu skizzieren.

(c) Die Größte-Ganze Funktion

$$\begin{aligned} [\ ] : \mathbb{R} &\to \mathbb{R}\,, \\ x &\mapsto [x] = \max\{k \in \mathbb{Z};\ k \leq x\} \end{aligned}$$

ist unstetig für alle $k \in \mathbb{Z}$, stetig für alle $k \in \mathbb{R} - \mathbb{Z}$ (siehe Abb.10).

Beweis :

(1) Wir zeigen, dass [ ] an jeder Stelle $k \in \mathbb{Z}$ unstetig ist. Dazu wählen wir die Folge (x_n) mit $x_n := k - \frac{1}{n}$ für alle $n \in \mathbb{N}$. Dann gilt $\lim\limits_{n\to\infty} x_n = k$, aber mit $f :=$ [ ] ist

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f\left(k - \frac{1}{n}\right) = \lim_{n\to\infty} (k-1) = k - 1 \neq k = f(k)$$

Die Bildfolge konvergiert also nicht gegen den Funktionswert $f(k)$.

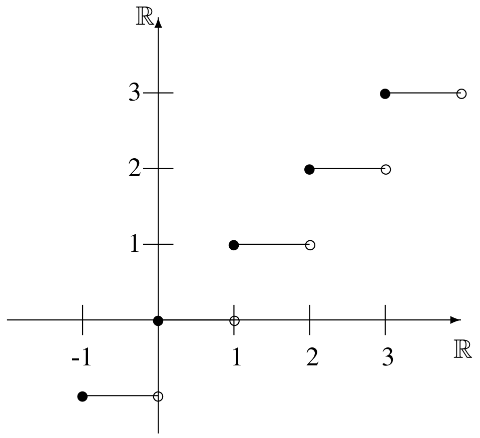


Abbildung 10: Skizze des Graphen von [ ]

(2) Wir zeigen, dass $f := [\ \ ]$ an jeder Stelle $a \in \mathbb{R} - \mathbb{Z}$ stetig ist.
Nach der Definition von $[\ \ ]$ gibt es genau eine ganze Zahl k mit

$$k < a < k+1 \qquad \text{und} \qquad f(a) = [a] = k\ .$$

Ist nun (x_n) eine beliebige reelle Folge mit $\lim\limits_{n\to\infty} x_n = a$.

Zu $\delta := \min\{a-k, k+1-a\} > 0$ gibt es dann ein $N = N(\delta) \in \mathbb{N}$, so dass $|x_n - a| < \delta$ für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt.
Nun gilt

$$|x_n - a| < \delta \iff a - \delta < x_n < a + \delta$$

und hieraus folgt wegen der Wahl von δ

$$k < x_n < k+1$$

für alle $n \geq N$. Dann gilt aber für alle $n \geq N$

$$f(x_n) = [x_n] = k,$$

also gilt

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

□

Übungsaufgabe:
Geben Sie für $k \in \mathbb{Z}$ eine Folge $(x_n),\ x_n \in \mathbb{R}$, für welche $\lim\limits_{n\to\infty} x_n = k \in \mathbb{Z}$ gilt, für welche aber die Bildfolge $f(x_n) = [x_n]$ divergiert.

(d) Die Funktionen

$$\begin{aligned} {}^{-} : \mathbb{C} &\to \mathbb{C}\,, \\ z &\mapsto \overline{z}\,, \end{aligned} \qquad\qquad \begin{aligned} \mathrm{Re} : \mathbb{C} &\to \mathbb{C}\,, \\ z &\mapsto \mathrm{Re}\, z\,, \end{aligned}$$

$$\begin{aligned} \mathrm{Im} : \mathbb{C} &\to \mathbb{C}\,, \\ z &\mapsto \mathrm{Im}\, z\,, \end{aligned} \qquad\qquad \begin{aligned} |\ | : \mathbb{C} &\to \mathbb{R}\,, \\ z &\mapsto |z|\,, \end{aligned}$$

sind stetig. Denn wir wissen: Konvergiert die komplexe Zahlenfolge (z_n) gegen $z \in \mathbb{C}$, dann konvergiert die Folge $(\bar{z}_n)$ gegen $\bar{z}$ und die Folge Re(z_n)) konvergiert gegen Re(z) bzw. Im(z), und die Folge $(|z_n|)$ konvergiert gegen $|z|$.

11.2 ε - δ - Charakterisierung der Stetigkeit

Einleitend haben wir schon bemerkt, dass Stetigkeit von f an der Stelle x folgendes bedeutet: $f(x)$ ändert sich beliebig wenig, wenn sich x nur hinreichend wenig ändert.

Wir beweisen gleich den folgenden

11.2.1 Theorem (Äquivalenzsatz für Stetigkeit)

Ist $f : D \to \mathbb{K}$ eine Funktion, $a \in D$. Dann sind folgende Aussagen äquivalent:

(a) Für jede Folge (x_n), $x_n \in D$, mit $\lim\limits_{n\to\infty} x_n = a$ konvergiert die Bildfolge $(f(x_n))$ gegen $f(a)$:

$$\lim_{n\to\infty} f(x_n) = f(a) \ ,$$

d.h. f ist (folgen-)stetig in a.

(b) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in D$ mit $|x - a| < \delta$ gilt $|f(x) - f(a)| < \varepsilon$. ($\varepsilon$ - δ - Stetigkeit von f in a).

(c) Zu jeder ε - Umgebung $U_\varepsilon(f(a))$ gibt es eine δ - Umgebung $U_\delta(a)$ mit

$$f(D \cap U_\delta(a)) \subset U_\varepsilon(f(a)) \ .$$

(Umgebungsdefinition der Stetigkeit in a.)

Beweis : Wir zeigen: $(a) \Longrightarrow (b),\ (b) \Longrightarrow (a)$ und $(b) \Longleftrightarrow (c)$.

Wir zeigen zunächst $(b) \Longrightarrow (a)$:
Die ε - δ - Bedingung sei also erfüllt und es sei (x_n) eine Folge mit $x_n \in D$ und $\lim\limits_{n\to\infty} x_n = a$. Zu zeigen ist: $\lim\limits_{n\to\infty} f(x_n) = f(a)$. Ist $\varepsilon > 0$ beliebig vorgegeben, dann wähle man dazu ein $\delta > 0$ nach Definition der ε - δ - Stetigkeit. Weil (x_n) gegen a konvergiert, liegen in $U_\delta(a)$ fest alle Folgenglieder, d.h. es gibt ein $N \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geq \mathbb{N}$ gilt

$$x_n \in U_\delta(a) \iff |x_n - a| < \delta \text{ für } n \geq N \ .$$

Nach Voraussetzung folgt dann $f(x_n) \subset U_\varepsilon(f(a))$ oder äquivalent $|f(x_n) - f(a)| < \varepsilon$ für alle $n \geq N$, d.h. $\lim\limits_{n\to\infty} f(x_n) = f(a)$.

$(a) \Longrightarrow (b)$
Sei also jetzt f folgenstetig in a.

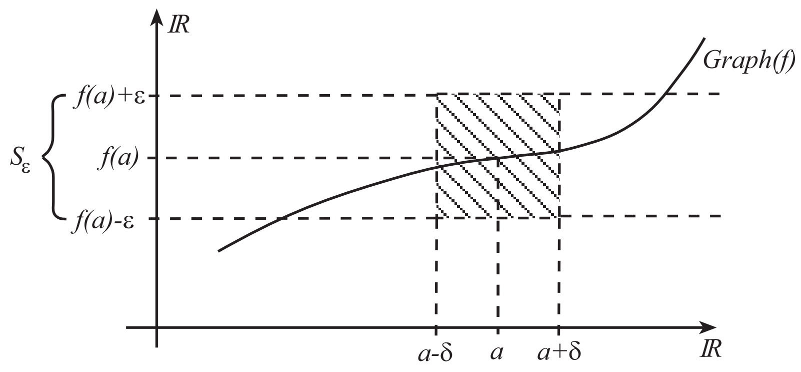


Abbildung 11: ε - δ - Stetigkeit von f in a

Wir führen ein Widerspruchsbeweis und nehmen an, dass die ε - δ - Bedingung i.A. *nicht* erfüllt ist. Dann gilt: Es gibt ein (Ausnahme) $\varepsilon_0 > 0$, so dass für *alle* $\delta > 0$ ein $x \in D$ existiert, für das zwar $|x - a| < \delta$, aber $|f(x) - f(a)| \geq \varepsilon_0$ gilt.
Wenn dies für *alle* $\delta > 0$ gilt, dann gilt dieses insbesondere für $\delta = \frac{1}{n}$, $n \in \mathbb{N}$ beliebig. Es gibt also ein $\varepsilon_0 > 0$ und zu jedem n ein $x_n \in D$ mit

$$|x_n - a| < \frac{1}{n} \qquad \text{und} \qquad |f(x_n) - f(a)| \geq \varepsilon_0$$

für alle $n \in \mathbb{N}$.

Die Folge (x_n) konvergiert offensichtlich gegen a (($\frac{1}{n}$) ist eine Nullfolge), aber die Bildfolge konvergiert nicht gegen $f(a)$. Das ist ein Widerspruch zur vorausgesetzten Folgenbedingung. Deshalb war unsere Annahme falsch, d.h. die ε - δ - Bedingung ist doch erfüllt, wenn f in a folgenstetig ist.

$(b) \Longleftrightarrow (c)$ ist offensichtlich aufgrund der Definition des Begriffs ε - Umgebung bzw. δ - Umgebung.
□

11.2.2 Bemerkungen und Beispiele

(a) Ist $D \subset \mathbb{R}$ und $f : D \to \mathbb{R}$ eine reellwertige Funktion und veranschaulicht man sich f mit Hilfe von $Graph(f)$, dann bedeutet die Stetigkeit von f in a folgendes:

Für beliebiges $\varepsilon > 0$ sei

$$S_\varepsilon := \{(x, y) \in \mathbb{R} \times \mathbb{R};\ f(a) - \varepsilon < f(a) + \varepsilon\}$$

der ε - Streifen (der Breite 2ε), dann gibt es dazu ein $\delta > 0$, so dass für alle $x \in D \cap U_\delta(a)$ gilt $f(x) \in S_\varepsilon$, d.h. man kann den Graphen von f über $D \cap U_\delta(a)$ in einem Kasten von beliebig kleiner Höhe einsperren (siehe Abb.11).

Das δ wird in der Regel von ε und der Stelle a abhängig, wählt man ein beliebiges ε, wird man in der Regel auch δ verkleinern müssen.

Zur Veranschaulichung von Funktionen $f : D \to \mathbb{C}$ $(D \subset \mathbb{C})$ kann man sich zwei Exemplare von $\mathbb{C}$ vorstellen, in der einen Ebene denkt man sich die Punkte von D aufgetragen (wobei

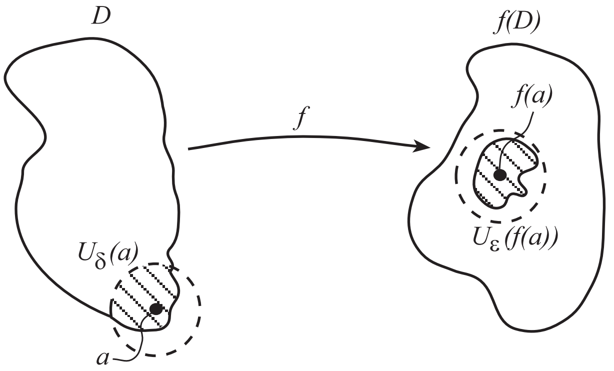


Abbildung 12: Umgebungsdefinition der Stetigkeit in a

$D = \mathbb{C}$ sein kann), in der anderen Ebene die Bildpunkte $f(z)$, $z \in D$. (siehe Abb.12)

Zu einer beliebig vorgegebenen ε - Umgebung $U_\varepsilon(a)$ mit $f(D \cap U_\delta(a)) \subset U_\varepsilon(f(a))$

(b) Die Quadratwurzelfunktion

$$\begin{array}{rccl} \sqrt{} : & \mathbb{R}_+ & \rightarrow & \mathbb{R}\,, \\ & x & \mapsto & \sqrt{x} \end{array}$$

ist in jedem Punkt $a \in \mathbb{R}_+$ stetig.

Beweis : Ist $a = 0$ und $\varepsilon > 0$ beliebig vorgegeben, so kann man $\delta := \varepsilon^2$ wählen, denn wegen der Monotonie der Wurzelfunktion gilt dann

$$\sqrt{x} - \sqrt{0} = \sqrt{x} < \varepsilon$$

für alle x mit $0 \leq x < \varepsilon^2 = \delta$. Ist $a > 0$, so formen wir zunächst um:

$$\sqrt{x} - \sqrt{a} = \frac{x-a}{\sqrt{x} + \sqrt{a}}.$$

Hieraus erhält man, dass man $\delta := \sqrt{a} \cdot \varepsilon$ wählen kann, denn wenn $|x - a| < \delta$ und $x \in \mathbb{R}_+$ gilt, so ist

$$|\sqrt{x} - \sqrt{a}| < \frac{\sqrt{a}\varepsilon}{\sqrt{x} + \sqrt{a}} < \varepsilon\ .$$

$\square$

(c) Allgemeiner gilt: Für jedes $k \in \mathbb{N}$, $k \geq 2$, ist die k-te Wurzel $\sqrt[k]{} :\ \mathbb{R}_+ \rightarrow \mathbb{R}$ stetig.

Beweis : Wir verwenden die Ungleichung

$$(*) \qquad |\sqrt[n]{a} - \sqrt[k]{b}| \leq \sqrt[k]{|a-b|}$$

für $a \geq 0,\ b \geq 0$.

Beweisen Sie die Ungleichung $(*)$ mit Hilfe des binomischen Satzes, indem Sie

$$a \leq (\sqrt[k]{a-b} + \sqrt[k]{b})^k$$

für $a \geq b \geq 0$ zeigen.

Seien nun $x, a \in \mathbb{R}_+$ und $\varepsilon > 0$ beliebig vorgegeben. Man kann $\delta = \varepsilon^k$ wählen, denn es gilt dann für $|x-a| < \delta$

$$|\sqrt[k]{x} - \sqrt[k]{a}| \leq \sqrt[k]{|x-a|} < \sqrt[k]{\delta} = \sqrt[k]{\varepsilon^k} = \varepsilon.$$

□

(d) Wir beweisen nochmal die Stetigkeit von

$$\begin{array}{rrcl} [\ ]: & \mathbb{R} & \to & \mathbb{R}\,, \\ & x & \mapsto & [x] \end{array}$$

für nicht ganzzahlige Argumente $a \in \mathbb{R}$.

Es gibt genau eine ganze Zahl k mit $k < a < k+1$.
Man wähle

$$\delta := \min\{a-k, k+1-a\} \quad \text{(Zeichnung!)},$$

dann folgt aus $|x-a| < \delta$

$$[x] - [a] = 0 < \varepsilon$$

für jedes positive ε.
Zu jedem $\varepsilon > 0$ kann man also stets dasselbe δ wählen.

(e) Das zur vorgegebenen Toleranzschwelle ε zu bestimmende δ wird i.A. von ε und der Stelle a abhängen. Als einfaches Beispiel betrachten wir

$$\begin{array}{rrcl} f: & \mathbb{C} & \to & \mathbb{C}\,, \\ & z & \mapsto & z^2\,. \end{array}$$

Wir wissen, dass f stetig ist (nach dem Folgenkriterium), geben aber nochmals zur Einüben der Technik einen ε - δ - Beweis.

Für $z, a \in \mathbb{C}$ versuchen wir unter der Voraussetzung $|z-a| < \delta$

$$|f(z) - f(a)| = |z^2 - a^2|$$

so durch einen von z unabhängigen Ausdruck nach oben abzuschätzen, dass wir aus dem größeren Ausdruck die Zahl δ mit der gewünschten Genauigkeit berechnen können. Es gilt

$$\begin{aligned} |z^2 - a^2| = |(z-a)(z+a)| &= |z+a||z-a| \\ &\leq (|z| + |a|)|z-a| \\ &< (|z| + |a|)\delta\,. \end{aligned}$$

Aus $|z-a| < \delta$ folgt $|z| - |a| \leq |z-a| < \delta$ also $|z| < |a| + \delta$ und $|z| + |a| < 2|a| + \delta$.

Wir wollen $\delta > 0$ so bestimmen, dass $(2|a| + \delta)\delta \leq \varepsilon$ wird.

Da δ lediglich der Einschränkung $\delta > 0$ unterworfen ist, sonst aber beliebig ist, betrachten wir

nur Zahlen δ mit $0 < \delta \leq 1$.
Dann gilt

$$|z^2 - a^2| < (2|a| + \delta)\delta \leq (2|a| + 1)\delta.$$

Aus der Bedingung $(2|a| + 1)\delta \leq \varepsilon$ erhält man

$$\delta \leq \frac{\varepsilon}{2|a| + 1} .$$

Wählt man daher von vornherein

$$0 < \delta \leq \min\left\{1,\ \frac{\varepsilon}{2|a| + 1}\right\} ,$$

dann folgt aus

$$|f(z) - f(a)| < (2|a| + \delta)\delta \leq (2|a| + 1)\delta$$

für alle z mit $|z - a| < \delta$

$$|f(z) - f(a)| < (2|a| + 1) \cdot \frac{\varepsilon}{2|a| + 1} = \varepsilon .$$

□

Übungsaufgabe:
Versuchen Sie nach dem gleichen Schema einen ε - δ - Beweis für die Stetigkeit von

$$f : \begin{array}{ccc} \mathbb{C} & \to & \mathbb{C}\,, \\ z & \mapsto & z^3 \end{array}$$

zu geben.

(f) Funktionen $f : D \to \mathbb{K}$, für die es eine Konstante $L \geq 0$ gibt, so dass für alle $x, y \in D$ gilt

$$|f(x) - f(y)| \leq L|x - y|,$$

sind stetig auf D.

Solche Funktionen nennt man auch *Lipschitz-stetig* (Rudolf Lipschitz, 1832-1903). Im Fall $L < 1$ spricht man auch von einer *Kontraktion*.

Für den Stetigkeitsbeweis setze man $\delta = \frac{\varepsilon}{L}$ falls $L \neq 0$ und $\delta = 1$ falls $L = 0$.
Das zur Toleranzschwelle $\varepsilon > 0$ zu bestimmende δ ist also unabhängig von der betrachteten Stelle.

Beispiele für Lipschitz-stetige Funktionen auf $\mathbb{C}$ sind:

(α) die linear-affine Funktion

$$\begin{array}{ccc} \mathbb{C} & \to & \mathbb{C} \quad (a, b \in \mathbb{C}) \\ z & \mapsto & az + b \end{array}$$

mit $L = |a|$.
Insbesondere sind alle konstante Funktionen auf $\mathbb{C}$ und $id_{\mathbb{C}}$ Lipschitz-stetig;

(β) die Funktionen

$$\begin{aligned} {}^{-} : \mathbb{C} &\to \mathbb{C}\,, \\ z &\mapsto \overline{z}\,, \end{aligned} \qquad \begin{aligned} \mathrm{Re} : \mathbb{C} &\to \mathbb{C}\,, \\ z &\mapsto \mathrm{Re}\, z\,, \end{aligned}$$

$$\begin{aligned} \mathrm{Im} : \mathbb{C} &\to \mathbb{C}\,, \\ z &\mapsto \mathrm{Im}\, z\,, \end{aligned} \qquad \begin{aligned} |\;| : \mathbb{C} &\to \mathbb{R}\,, \\ z &\mapsto |z|\,, \end{aligned}$$

mit der Lipschitz-Konstanten $L = 1$;

(g) Ist $D = [1, 2] \cup 3$ und $f : D \to \mathbb{C}$ irgendeine Funktion, dann ist f stetig im Punkt $a = 3$. Zu beliebig vorgegebenem $\varepsilon > 0$ wähle wir nämlich etwa $\delta = \frac{1}{2}$, dann folgt aus $|x - 3| < \frac{1}{2}$ und $x \in D$ notwendig $x = 3$, also ist

$$|f(x) - f(a)| = 0 < \varepsilon$$

für jedes $\varepsilon > 0$.

Mit dem Folgenkriterium kann man so argumentieren:
Ist (x_n) eine Folge mit $x_n \in D$ und $\lim\limits_{n\to\infty} x_n = 3$, dann liegen fast alle Folgenglieder in $U_{\frac{1}{2}}(3)$, das heißt fast alle Folgenglieder x_n sind gleich 3. Dann ist auch $f(x_n) = f(3)$ für fast alle n, d.h. $\lim\limits_{n\to\infty} f(x_n) = f(3)$.

Verallgemeinerung:
Ist $a \in D$ ein *isolierter Punkt* von D, d.h. gibt es eine r-Umgebung $U_r(a)$ ($r > 0$) mit $D \cap U_r(a) = \{a\}$, dann ist *jede* Funktion $f : D \to \mathbb{K}$ stetig in a.

11.2.3 Bemerkung

Stetigkeit in einem Punkt $a \in D$ is eine lokale Eigenschaft, d.h. ist $f : D \to \mathbb{K}$ eine Funktion und $a \in D$, dann ist f genau dann stetig in a, wenn es eine r-Umgebung von a gibt, so dass die Einschränkung

$$f|D \cap U_r(a)$$

stetig in a ist.

Ist $M := D \cap U_r(a)$, dann ist $f|M$ natürlich stetig in a, wenn f stetig in a ist, denn jede Folge (x_n), $x_n \in M$, die gegen a konvergiert, konvergiert auch in D gegen a.

12 Rechenregeln für stetige Funktionen

Wie wiederholen nochmals (vergl. §11.1.1)

12.1 Satz

Sind $f, g : D \to \mathbb{K}$ stetig in $a \in D$, dann sind auch $f + g$ und fg stetig in a. Insbesondere ist cf stetig in a für jedes $c \in \mathbb{K}$. Ist außerdem $g(a) \neq 0$, dann ist

$$\frac{f}{g} : D' := \{x \in D;\ g(x) \neq 0\} \to \mathbb{K}$$

stetig in a.

Beweis Ist (x_n) eine Folge mit $x \in D$ und $\lim\limits_{n\to\infty} x_n = a$. Dann gilt $\lim\limits_{n\to\infty} f(x_n) = f(a)$ und $\lim\limits_{n\to\infty} g(x_n) = g(a)$ wegen der Stetigkeit von f bzw. g in a.
Nach den Rechenregeln für Folgen (vergl. §8) folgt

$$\lim_{n\to\infty}(f+g)(x_n) = \lim_{n\to\infty} f(x_n) + \lim_{n\to\infty} g(x_n) = f(a) + g(a) = (f+g)(a).$$

Ist (x_n) eine Folge mit $x_n \in D'$ und $\lim\limits_{n\to\infty} x_n = a$, dann gilt

$$\lim_{n\to\infty} g(x_n) = g(a)(\neq 0).$$

Wegen $x_n \in D'$ ist $g(x_n) \neq 0$, daher folgt

$$\lim_{n\to\infty} \frac{1}{g}(x_n) = \lim_{n\to\infty} \frac{1}{g(x_n)} = \frac{1}{g(a)} = \frac{1}{g}(a).$$

□

Die allgemeine Behauptung folgt jetzt aus

$$\frac{f}{g} = f \cdot \frac{1}{g}.$$

Aus Satz 12.1 folgt, dass die stetige Funktion $f : D \to \mathbb{K}$ eines $\mathbb{K}$-Vektorraums, sogar eine $\mathbb{K}$-Algebra bildet.
Aus dem Satz ergibt sich nochmals die Stetigkeit von Polynomen in $\mathbb{R}$ bzw. $\mathbb{C}$ und die Stetigkeit rationaler Funktionen auf ihren jeweiligen Definitionsbereichen. Ist nämlich $R = \frac{f}{g}$, $f,\ g$ Polynome, so ist

$$S = \{z \in \mathbb{C};\ g(z) = 0\}$$

eine endliche Menge (nach dem Fundamentalsatz der Algebra) und

$$D = Def(R) = \mathbb{C} - S = \{z \in \mathbb{C};\ g(z) \neq 0\}$$

So ist z.B. die rationale Funktion

$$\begin{aligned} \sigma : \mathbb{R} &\to \mathbb{C} \\ x &\mapsto \frac{2x + i(x^2-1)}{x^2+1} \end{aligned}$$

stetig auf $\mathbb{R}$.
Stetigkeit ist auch mit dem Zusammensetzen von Funktionen verträglich.

12.2 Satz

Sind D und $E \subset \mathbb{K}$ und $f : D \to E$ und $g : E \to \mathbb{K}$ Funktionen und ist f stetig in $a \in D$ und g stetig in $b := f(a)$, dann ist die Zusammensetzung $g \circ f : D \to \mathbb{K}$ stetig in a.

Schlagwort: Die Zusammensetzung stetiger Funktionen ist stetig.

Beweis: Ist (x_n) eine Folge mit $x_n \in D$ und $\lim\limits_{n\to\infty} x_n = a \in D$ und $y_n := f(x_n)$.
Dann ist (y_n) eine Folge in E mit

$$b = f(a) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} y_n.$$

Wegen der Stetigkeit von g in $b = f(a)$, gilt

$$g(f(a)) = g(b) = \lim_{n\to\infty} g(y_n) = \lim_{n\to\infty} g(f(x_n)) = \lim_{n\to\infty} (g \circ f)(x_n).$$

□

Geben Sie einen weiteren Beweis für Satz 12.2 mit Hilfe der $\varepsilon - \delta-$Stetigkeitsdefinition.

12.3 Folgerung

Insbesondere sind mit f auch $\bar{f},\ |f|,\ Re\, f$ und $Im\, f$ stetig. Als Zusammensetzung stetiger Funktionen ist z.B. $|f| = |\ | \circ f$ stetig.

12.4 Bemerkung

Sind $f : \mathbb{R} \to \mathbb{R}$ und $g : \mathbb{R} \to \mathbb{R}$ stetige Funktionen und gilt

$$f|\mathbb{Q} = g|\mathbb{Q},$$

dann gilt $f(x) = g(x)$ für alle $x \in \mathbb{R}$.
Die Werte einer stetigen Funktion auf $\mathbb{R}$ sind also durch ihre Werte auf $\mathbb{Q}$ schon eindeutig bestimmt. Hieraus folgt z.B., dass die Menge $C(\mathbb{R}, \mathbb{R}) := \{f \in Abb(\mathbb{R}, \mathbb{R});\ f \text{ stetig}\}$ der stetigen Funktionen auf $\mathbb{R}$, die gleiche Mächtigkeit wie $\mathbb{R}$ hat:

$$\#\mathbb{R} = \#C(\mathbb{R}, \mathbb{R}).$$

Beweis der Bemerkung: Zu zeigen ist: Es gilt $f(x) = g(x)$ für alle $x \in \mathbb{R}$, wenn $f|\mathbb{Q} = g|\mathbb{Q}$ gilt.
Nun gibt es aber zu jedem $x \in \mathbb{R}$ eine Folge (r_n) von rationaler Zahlen mit $\lim\limits_{n\to\infty} r_n = x$. Aus $f(r_n) = g(r_n)$ und der Stetigkeit von f folgt dann

$$f(x) = \lim_{n\to\infty} f(r_n) = \lim_{n\to\infty} g(r_n) = g(x).$$

□

13 Abbildungseigenschaften stetiger Funktionen

Unsere alltägliche Stetigkeitsvorstellung unterstellt, dass „stetige“ Veränderungen oder Abläufe (etwa Bewegungsvorgänge) keinen abrupten, jähen Schwankungen unterworfen sind. Die $\varepsilon-\delta-$Definition der Stetigkeit bringt diese Vorstellung exakt zum Ausdruck. Die folgenden Sätze sind ebenfalls Präzisierungen des oben beschriebenen Sachverhalts, insbesondere werden wir sehen, dass -obwohl Stetigkeit eine *lokale* Eigenschaft ist- sie häufig auch Rückschlüsse auf den *globalen* Verlauf der Funktion (z.B. Existenz von Maxima und Minima) gestattet, insbesondere wenn man über den Definitionsbereich D zusätzliche Forderungen stellt (D z.B. ein Intervall oder D kompakt).

13.1 Eine einfache Abbildungseigenschaften stetiger Funktionen

Wir beginnen mit einem einfachen Satz, der etwa für die Integralrechnung nützlich sein wird.

13.1.1 Satz

Ist $f : D \to \mathbb{K}$ stetig in $a \in D$ und gilt $f(a) \neq 0$, dann gibt es eine $\delta-$Umgebung von a, so dass $f(x) \neq 0$ für alle $x \in U_\delta(a) \cap D$ gilt.

Beweis : Zu $\varepsilon := \frac{1}{2}|f(x)| > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in U_\delta(a) \cap D$ gilt

$$|f(x) - f(a)| < \varepsilon.$$

Aus

$$|f(x)| \geq |f(a)| - |f(a) - f(x)| \geq |f(a)| - \varepsilon = \frac{|f(a)|}{2} > 0$$

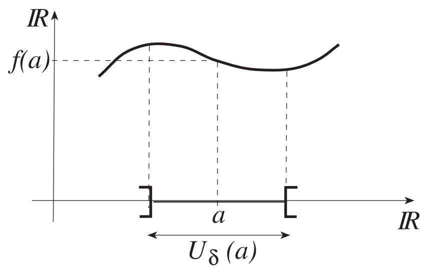
folgt die Behauptung unmittelbar.

$\square$

Bemerkung:
Ist f insbesondere reellwertig und etwa $f(a) > 0$, dann gibt es eine $\delta-$Umgebung $U_\delta(a)$, so dass für alle $x \in U_\delta(a) \cap D$ gilt

$$f(x) > 0.$$

Wenn also eine stetige reellwertige Funktion an einer Stelle a ihres Definitionsbereichs einen positiven Wert annimmt, dann gibt es eine Umgebung $U_\delta(a)$, so dass für alle $x \in U_\delta(a)$, soweit sie zum Definitionsbereich von f gehören, die entsprechende Werte auch positiv sind:



13.2 Stetige reellwertige Funktionen auf Intervallen: Der Nullstellensatz von Bolzano und der Zwischenwertsatz

Besonders angenehme Abbildungseigenschaften haben stetige reellwertige Funktionen mit einem *Intervall* als Definitionsbereich.

13.2.1 Satz (Nullstellensatz von Bolzano, B.Bolzano (1817))

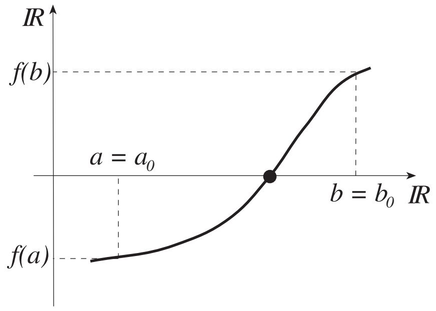
Ist $D \subset \mathbb{R}$ ein (echtes) Intervall (d.h. $\#D \geq 2$) $f : D \to \mathbb{R}$ eine stetige Funktion und sind $a, b \in D$, $a < b$ und gilt

$$\begin{array}{lll} f(a) < 0 & \text{und} & f(b) > 0 \text{ bzw.} \\ f(a) > 0 & \text{und} & f(b) < 0, \end{array}$$

dann existiert ein $p \in]a, b[$ mit $f(p) = 0$.

13.2.2 Bemerkung

Die Aussage des Satzes ist anschaulich klar, vgl. dazu die folgende Abbildung



Hier kommt nochmals der wesentliche Unterschied zwischen $\mathbb{Q}$ und $\mathbb{R}$ zum Ausdruck. Die Aussage wird nämlich falsch, wenn man nur im Körper $\mathbb{Q}$ der rationalen Zahlen arbeitet. Ist etwa $M := \{x \in \mathbb{Q};\ 1 \leq x \leq 2\}$ und

$$\begin{array}{rcl} f : M & \to & \mathbb{R} \\ x & \to & x^2 - 2, \end{array}$$

die stetige Funktion, dann ist $f(1) = -1 < 0$ und $f(2) = 2 > 0$, aber bekanntlich (vgl. ???) gibt es keine *rationale* Zahl p mit $f(p) = 2$, d.h. $p^2 = 2$. (Schlagwort: $\sqrt{2}$ ist irrational)

Beweis von 13.2.1: Wir geben einen konstruktiven Beweis mit Hilfe der Intervallhalbierungsmethode. Wir konstruieren induktiv eine Intervallschachtelung $[a_n, b_n],\ n \in \mathbb{N}_0$, in $[a, b]$ mit folgenden Eigenschaften:

(1) $[a_n, b_n] \subset [a_{n-1}, b_{n-1}]$ für $n \in \mathbb{N}$
(2) $b_n - a_n = \frac{b-a}{2^n}$ für $n \in \mathbb{N}_0$
(3) $f(a_n) \leq 0$ und $f(b_n) \geq 0$ für $n \in \mathbb{N}_0$.

Induktionsanfang: Wir setzen $[a_0, b_0] := [a, b]$;
Induktionsschritt: Wenn das Intervall $[a_n, b_n]$ (mit den Eigenschaften (1), (2), (3)) bereits definiert ist, so setzen wir

$$m_n := \frac{a_n + b_n}{2} \qquad \text{(Intervallmitte)}$$

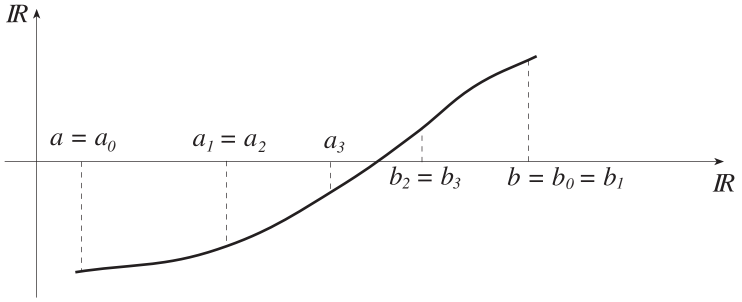
Dann können im Prinzip 3 Fälle auftreten:

1. Fall: $f(m_n) = 0$ und wir setzen $p = m_n$. Damit haben wir eine Nullstelle gefunden.

2. Fall: $f(m_n) > 0$ und wir setzen $[a_{n+1}, b_{n+1}] = [a_n, m_n]$

3. Fall: $f(m_n) < 0$ und wir setzen $[a_{n+1}, b_{n+1}] = [m_n, b_n]$

Obwohl im ersten Fall unser Verfahren schon beendet ist, fassen wir den ersten und den zweiten Fall zusammen:
Falls $f(m_n) \geq 0$ gilt, setzen wir $[a_{n+1}, b_{n+1}] = [a_n, m_n]$.



Für das neue Intervall $[a_{n+1}, b_{n+1}]$ sind wieder die Eigenschaften (1) und (3) erfüllt und $[a_n, b_n]$ ist eine Intervallschachtelung. Nach dem Intervallschachtelungsprinzip gibt es eine Zahl p mit $p \in [a_n, b_n]$ für alle $n \in \mathbb{N}_0$ und $p = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$.

Wegen der Stetigkeit von f in p gilt dann auch

$$\lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n) = f(p).$$

Wegen der Monotonie des Grenzwertes folgt aus (3)

$$f(p) = \lim_{n\to\infty} f(a_n) \leq 0 \text{ und } f(p) = \lim_{n\to\infty} f(b_n) \geq 0,$$

d.h. es gilt $f(p) = 0$.

□

13.2.3 Beispiele

(1) **Existenz k-ter Wurzeln**
Für jede $k \in \mathbb{N}$ und jede $\alpha \in \mathbb{R}$, $\alpha > 0$, gibt es ein (eindeutig bestimmtes) $x \in \mathbb{R}$ mit $x^k = \alpha$.

Zum **Beweis** betrachten wir die stetige Funktion

$$\begin{array}{rcl} p : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^k - \alpha. \end{array}$$

Es gilt

$$\begin{array}{rcll} p(0) & = & -\alpha > 0 \text{ und} & \\ p(\alpha+1) & = & (\alpha+1)^k - \alpha & \\ & \geq & (1+\alpha k) - \alpha & \text{(nach der Bernoullischen Ungleichung)} \\ & = & 1 + \alpha(k-1) \geq 1 > 0. & \end{array}$$

Es ist also $p(0) < 0 < p(\alpha+1)$ und da p als Polynom auf ganz $\mathbb{R}$ stetig ist, gibt es nach dem Nullstellensatz ein $x \in \mathbb{R}$ mit $0 < x < \alpha + 1$ mit $p(x) = 0$, d.h. $x^n = \alpha$.

(2) Ist

$$\begin{aligned} P : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto x + +a_{k-1}x^{k-1} + \ldots + a_0, \\ &\quad (k \in \mathbb{N},\ k \text{ ungerade und } a_j \in \mathbb{R} \text{ für } j = 0, \ldots, k), \end{aligned}$$

also ein Polynom ungeraden Grades, dann hat p mindestens eine reelle Nullstelle, d.h. es gibt (mindestens) ein $p \in \mathbb{R}$ mit $P(p) = 0$.

Die Idee des **Beweises** besteht darin zu zeigen, dass $P(x)$ für genügend große positive x positiv und für genügend kleine negative x negativ ist.
Wir schreiben für $x \neq 0$

$$P(x) = x^k \underbrace{\left(1 + \frac{a_{k-1}}{x} + \ldots + \frac{a_0}{x^k}\right)}_{=g(x)} = x^k g(x)$$

Für alle $x \in \mathbb{R}$ mit $|x| \geq c := \max\{1, 2k|a_{k-1}|, \ldots, 2k|a_0|\}$ gilt dann wegen $|x^k| = |x|^k \geq |x|$

$$\begin{aligned} |g(x) - 1| &= \left|\frac{a_{k-1}}{x} + \frac{a_{k-2}}{x^2} + \ldots + \frac{a_0}{x^k}\right| \\ &\leq \frac{|a_{k-1}|}{|x|} + \frac{|a_{k-2}|}{|x^2|} + \ldots + \frac{|a_0|}{|x^k|} \\ &\leq \frac{1}{2k} + \frac{1}{2k} + \ldots + \frac{1}{2k} = \frac{k}{2k} = \frac{1}{2} \end{aligned}$$

M.a.W., es gilt für die betrachteten x

$$-\frac{1}{2} \leq g(x) - 1 \leq \frac{1}{2}$$

und daher insbesondere

$$\frac{1}{2} \leq g(x)$$

für die betrachteten x.
Wählen wir also ein $x_1 > 0$ mit $x_1 = |x_1| \geq c$, dann ist

$$\frac{x_1^k}{2} \leq x_1^k g(x_1) = P(x_1),$$

also $P(x_1) > 0$.
Wählen wir $x_2 = -x_1$, dann ist $x_2 < 0$ und da k ungerade ist gilt $x_2^k < 0$ und

$$\frac{x_2^k}{2} \geq x_2^k g(x_2) = P(x_2),$$

also $P(x_2) < 0$.
Nach dem Nullstellensatz von Bolzano gibt es also ein p mit $x_2 < p < x_1$ und $P(p) = 0$. In beiden Beispielen handelt es sich um Spezialfälle des *Fundamentalsatzes der Algebra*.

Ein Polynom ungeraden Grades kann mehrere reelle Nullstellen haben, z.B. das Polynom $x \mapsto P(x) = x^5 - 4x + 2$ (siehe Abb. 13).

Wie wir wissen, braucht ein Polynom *geraden* Grades mit reellen Koeffizienten keine reelle Nullstellen zu besitzen.
Das einfache Beispiel ist das quadratische Polynom

$$\begin{aligned} P : \mathbb{R} &\to \mathbb{R}\,, \\ x &\mapsto x^2 + 1 \end{aligned}$$

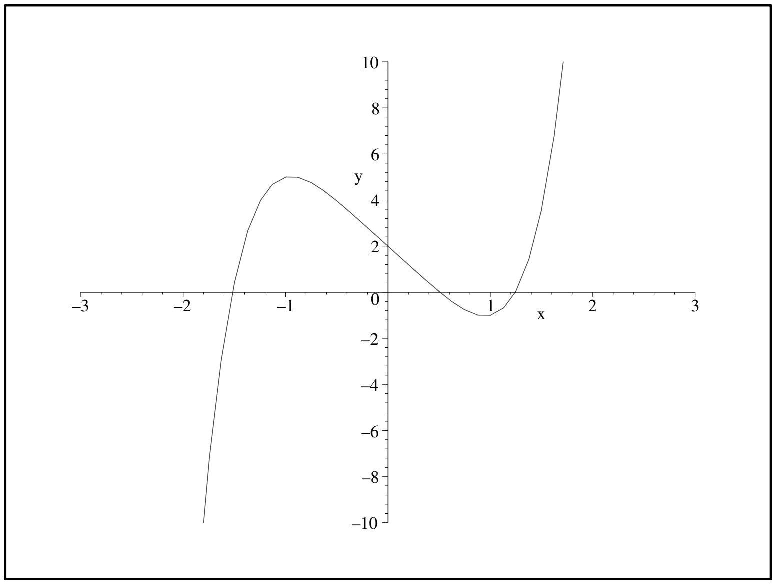


Abbildung 13: Graph des Polynoms $x^5 - 4x + 2$.

13.2.4 Korollar (Zwischenwertsatz)

Ist $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$ stetig und sind $a, b \in D$, $a < b$ und ist c irgendeine reelle Zahl mit

$$f(a) \leq c \leq f(b) \text{ oder } f(b) \leq c \leq f(a),$$

dann gibt es stets (mindestens) ein $p \in [a, b]$ mit $f(p) = c$.

Zum **Beweis** sei oBdA $f(a) < c < f(b)$.
Wir betrachten die stetige Funktion

$$\begin{aligned} g : [a, b] &\to \mathbb{R}\,, \\ x &\mapsto f(x) - c. \end{aligned}$$

Für sie gilt

$$g(a) = f(a) - c < 0 \text{ und } g(b) = f(b) - c > 0.$$

Nach dem Nullstellensatz von Bolzano gibt es ein $p \in [a, b]$ mit $g(p) = 0$, d.h. mit $f(p) = c$.

Natürlich folgt aus dem Zwischenwertsatz (ZWS) der Nullstellensatz von Bolzano als Spezialfall. In Wirklichkeit sind sie aber äquivalent.

Eine weitere äquivalente Formulierung für den Zwischenwertsatz ist

13.2.5 Korollar

Ist $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$ eine stetige Funktion, dann ist das Bild $f(D)$ wieder ein Intervall.

Kurz: Das stetige Bild eines Intervalls ist ein Intervall

Zum **Beweis** beachte man:
Ist $D = \emptyset$, dann ist $f(D) = f(\emptyset) = \emptyset$ und $\emptyset$ ist per Definition auch ein Intervall. Besteht D nur aus einem Element, etwa $D = \{a\}$, dann ist $f(\{a\}) =: \{f(a)\} = [f(a), f(a)]$, also ein Intervall.

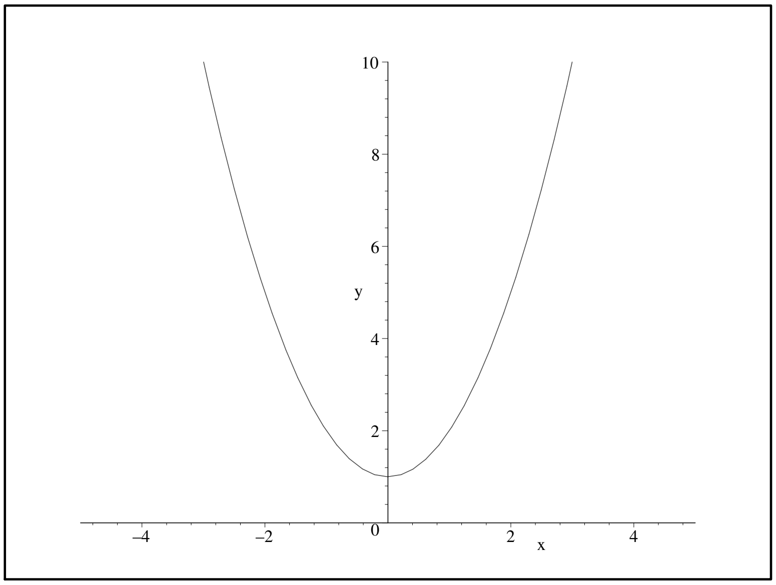


Abbildung 14: Graph des Polynoms $x^2 + 1$.

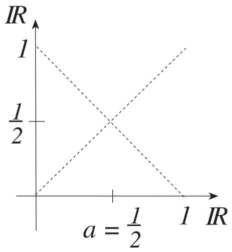
Ist D ein echtes Intervall ($\#D \geq 2$), dann kann man den ZWS anwenden:
Mit je zwei Funktionswerten, die von f angenommen werden, wird auch jeder „Wert dazwischen“ angenommen, diese Eigenschaft charakterisiert aber gerade die Intervalle (vgl. Übungsaufgabe ??? auf Blatt ???).

13.2.6 Bemerkung

Man könnte auf die Idee kommen, dass der ZWS für stetige Funktionen charakteristisch ist, d.h. wenn eine Funktion $f : [a, b] \to \mathbb{R}$ jeden Wert zwischen $f(a)$ und $f(b)$ annimmt, dass sie dann auch stetig ist.
Die Funktion $f : [0, 1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} x, & \text{falls } x \text{ rational,} \\ 1 - x, & \text{falls } x \text{ irrational,} \end{cases}$$

nimmt *jeden* Wert zwischen $f(0) = 0$ und $f(1) = 1$ an, ist aber nur im Punkt $a = \frac{1}{2}$ stetig (vgl. Beispiel ???)



Wir werden jedoch bald den erstaunlichen Satz beweisen:
Ist $f : D \to \mathbb{R}$ *monoton* und ist $f(D)$ ein Intervall, dann ist f stetig. ($D \neq \emptyset,\ f$ beliebige Funktion)

Eine einfache Folgerung aus dem ZWS ist auch der folgende Fixpunktsatz:

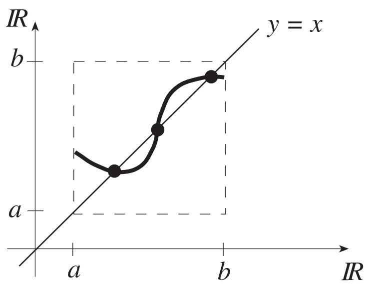


Abbildung 15: Geometrische Veranschaulichung des Fixpunktsatzes

13.2.7 Satz (Fixpunktsatz)

Sind $a, b \in \mathbb{R},\ a < b$ und ist $f : [a, b] \to \mathbb{R}$ stetig und gilt $f([a, b]) \subset [a, b]$, dann gibt es (mindestens) ein $\xi \in [a, b]$ mit $f(\xi) = \xi$, d.h. f hat in $[a, b]$ (mindestens) einen Fixpunkt. (siehe Abb. 15)

Beweis : Wir betrachten die stetige Funktion

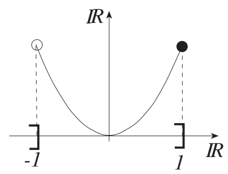
$$\begin{aligned} g:\ [a,b] &\to \mathbb{R}\,, \\ x &\mapsto f(x) - x. \end{aligned}$$

Wegen $f(a) \in [a, b]$ und $f(b) \in [a, b]$ gilt $g(a) \geq 0$ und $g(b) \leq 0$. g hat also in $[a, b]$ (mindestens) eine Nullstelle, diese ist ein Fixpunkt von f.

□

Dieser Fixpunktsatz ist ein Spezialfall des sog. *Brower'schen Fixpunktsatzes* (siehe z.B. Königsberger, Analysis 2). *Fixpunktsätze* sind in der Analysis ein starkes Hilfsmittel zum Beweis von Existenzaussagen. Man vergleiche hierzu auch Übungsaufgabe ??? von Blatt ???.
Der ZWS ist äquivalent zur Aussage, dass das stetige Bild eines Intervalls wieder ein Intervall ist. Der Typ des Intervalls kann sich dabei aber ändern:
Betrachtet man z.B.

$$\begin{aligned} f :]-1,1] &\to \mathbb{R}\,, \\ x &\mapsto x^2, \end{aligned}$$



dann ist $f(]-1,1]) = [0,1]$ oder

$$\begin{aligned} \sin :]0,2\pi[&\to \mathbb{R}\,, \\ x &\mapsto \sin x, \end{aligned}$$

dann ist $\sin(]0, 2\pi[) = [-1, 1]$, das *offene* Intervall wird also auf das *abgeschlossene* (sogar kompakte) Intervall $[-1, 1]$ abgebildet. Diese Eigenschaft des (reellen) Sinus werden wir in §13.3 beweisen.
Wir werden jedoch bald sehen:
Ist $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall und $f : [a, b] \to \mathbb{R}$ stetig, dann ist das Bildintervall $f([a, b])$ wieder ein kompaktes Intervall.
Wir zeigen zunächst, dass abgeschlossene bzw. kompakte Intervalle, also Intervalle vom Typ

$[a,b],\ a,b \in \mathbb{R},\ a \leq b$, im Sinne einer allgemeinen Kompaktheitsdefinition, tatsächlich abgeschlossen bzw. kompakt sind. „Kompakt" war bis jetzt nur ein Name für das Intervall $[a,b]$.

13.3 Stetige Funktionen auf kompakten Mengen, der Satz vom Maximum und Minimum

13.3.1 Definition

Eine Teilmenge $A \subset \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) heißt *abgeschlossen*, wenn der Grenzwert jeder konvergenten Folge $(a_n),\ a_n \in A$, wieder in A liegt. Die leere Menge $\emptyset$ zählen wir ebenfalls zu den abgeschlossenen Mengen.

13.3.2 Beispiele

Ist $A := [a,b] \subset \mathbb{R}$ ein abgeschlossenes Intervall ($a,b \in \mathbb{R},\ a \leq b$) und ist (a_n) eine konvergente Folge mit $a_n \in A$ und $\lim\limits_{n\to\infty} a_n = p (\in \mathbb{R})$, dann gilt wegen $a \leq a_n \leq b$ auch $a \leq \lim\limits_{n\to\infty} a_n = p \leq b$, also $p \in [a,b]$, d.h. ein abgeschlossenes Intervall ist abgeschlossen im Sinne der Definition.

Ein *offenes* Intervall $]a,b[$ ($a,b \in \mathbb{R},\ a < b$) hat diese Eigenschaft nicht:
Für die Folge (a_n) mit $a_n = a + \frac{b-a}{2^n},\ n \in \mathbb{N}$, gilt $a_n \in]a,b[,\ (a_n)$ ist konvergent, aber es gilt $\lim\limits_{n\to\infty} a_n = a \notin]a,b[$.

Die folgenden beiden Hilfssätze zeigen, wie man abgeschlossene Teilmengen in $\mathbb{R}$ oder $\mathbb{C}$ konstruieren kann.

13.3.3 Hilfssatz

Sind $f_j : \mathbb{C} \to \mathbb{R}$ stetige Funktionen ($1 \leq j \leq r,\ r \in \mathbb{N}$) und sind $a_1, \ldots, a_r \in \mathbb{R}$ reelle Konstanten, dann ist die Menge $A := \{z \in \mathbb{C};\ f_1(z) \leq a_1, \ldots, f_r(z) \leq a_r\}$ abgeschlossen.

Zum **Beweis** sei (z_n) eine konvergente Folge von Elementen $z_n \in A$ und $p := \lim\limits_{n\to\infty} z_n$. Wegen $z_n \in A$ ist $f_j(z_n) \leq a_j$ für alle $n \in \mathbb{N}$ und alle $j \in \{1,2,\ldots,r\}$. Wegen der Stetigkeit der f_j gilt dann auch

$$f_j(p) = \lim_{n\to\infty} f_j(z_n) \leq a_j \quad \text{für} \quad j \in \{1,2,\ldots,r\}$$

d.h. $p \in A$, A also abgeschlossen.

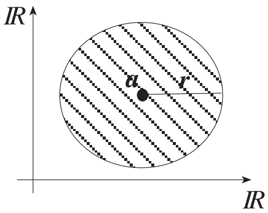
13.3.4 Beispiele

Abgeschlossene Mengen, die man so konstruieren kann sind:

(a) Die abgeschlossene Kreisscheibe

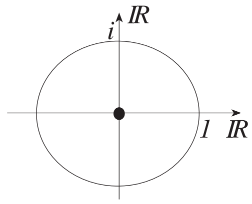
$$\overline{U_r}(a) := \{z \in \mathbb{C};\ |z-a| \leq r\} \qquad (a \in \mathbb{C},\ r \in \mathbb{R},\ r \geq 0)$$

Beachte: $\overline{U_0}(a) = \{a\}$



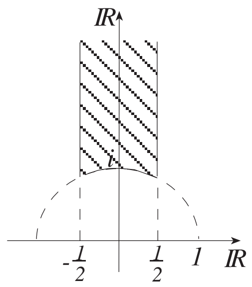
(b) Die 1-Sphäre

$$S^1 := \{z \in \mathbb{C};\ |z| = 1\}$$



(c) Die sog. „Modulfigur"

$$\mathcal{F} := \left\{z \in \mathbb{C};\ |z| \geq 1;\ |Re\, z| \leq \frac{1}{2};\ Im\, z > 0\right\}$$



13.3.5 Hilfssatz

Sind $A_1, \ldots, A_r \subset \mathbb{K}$ abgeschlossen, dann ist auch $A := A_1 \cup A_2 \cup \ldots \cup A_r$ abgeschlossen, d.h. die *Vereinigung endlich vieler abgeschlossenen Mengen ist wieder abgeschlossen*.
Ist $(A_\lambda)_{\lambda \in \Lambda}$ ein System abgeschlossener Menge, dann ist auch $\bigcap\limits_{\lambda \in \Lambda} A_\lambda$ abgeschlossen, d.h. *der Durchschnitt beliebig vieler abgeschlossenen Mengen ist wieder abgeschlossen*.
Der zweite Teil ist auf Grund der Definition evident, der erste Teil sei als Übungsaufgabe gestellt.
Man beachte jedoch, dass die Vereinigung beliebig vieler abgeschlossenen Mengen nicht wieder abgeschlossen zu sein braucht, wie etwa das folgende Beispiel.

13.3.6 Beispiel

$A_n = \{\frac{1}{n}\},\ n \in \mathbb{N}$, zeigt: Jede A_n ist abgeschlossen, aber

$$X := \bigcup_{n=1}^{\infty} A_n = \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\right\}$$

ist nicht abgeschlossen, da die Folge $(x_n) = \left(\frac{1}{n}\right)$, also $x_n \in X$, zwar konvergiert in X, ihr Grenzwert 0 aber nicht in X liegt.

13.3.7 Definition (kompakt)

Eine Teilmenge $K \subset \mathbb{K}$ heißt *kompakt*, wenn sie abgeschlossen und beschränkt ist. (Beschränkt heißt: Es gibt ein $R \in \mathbb{R}_+$, so dass für alle $z \in K$ gilt $|z| \leq R$).
Die leere Menge $\emptyset$ rechnen wir ebenfalls zu den kompakten Mengen.

13.3.8 Beispiele

(a) Für jedes $a \in \mathbb{C}$ und jedes $r \in \mathbb{R}_+$ sind die abgeschlossenen Kreisscheiben

$$\overline{U_r}(a) = \{z \in \mathbb{C};\ |z-a| \leq r\}$$

kompakt, ebenso die Intervalle $[a-r, a+r] \subset \mathbb{R}$.

(b) Die 1-Sphäre $S^1 = \{z \in \mathbb{C};\ |z| = 1\}$ ist kompakt.

(c) Die Modulfigur $\mathcal{F}$ aus §13.3.4(c) ist nicht kompakt.

Die beiden Hilfssätze über abgeschlossene Mengen ergeben analoge Sätze für kompakte Mengen. Es gilt etwa

13.3.9 Hilfssatz

(a) Die Vereinigung *endlich* vieler kompakten Mengen ist wieder kompakt.

(b) Der Durchschnitt beliebig vieler kompakten Mengen ist wieder kompakt.

(c) Ist $K \subset \mathbb{K}$ kompakt und $A \subset \mathbb{K}$ abgeschlossen, so ist der Durchschnitt $A \cap K$ wieder kompakt.

Kompakte Mengen können jedoch eine sehr komplizierte Struktur haben. Man vergleiche hierzu etwa die Konstruktion des *Cantorschen Diskontinuums* (Cantorsche „Wischmenge") bei Königsberger: Analysis 1 (Abschnitt 7.5). Eine brauchbare Charakterisierung kompakter Mengen erhält man mit Hilfe des Satzes von Bolzano-Weierstrass.

13.3.10 Lemma (Bolzano-Weierstrass-Charakterisierung von „kompakt")

Eine Teilmenge $K \subset \mathbb{K}$ ist genau dann kompakt, wenn jede Folge von Elementen aus K eine Teilfolge besitzt, die gegen ein Element aus K konvergiert.

Beweis : Sei K kompakt, also abgeschlossen und beschränkt. Dann ist jede Folge $(z_n),\ z_n \in K$, beschränkt und besitzt nach dem Satz von Bolzano-Weierstrass für Folgen eine konvergente Teilfolge. Der Grenzwert dieser Folge liegt aber in K, da K abgeschlossen ist.

Umgekehrt besitzt *jede* Folge aus K eine Teilfolge, deren Grenzwert in K liegt.

Wir zeigen zunächst, dass K abgeschlossen ist. Denn ist (z_n) irgendeine konvergente Folge mit $z_n \in K$ und $\xi := \lim\limits_{n\to\infty} z_n$, dann gilt auch $\xi \in K$, denn nach Voraussetzung hat (z_n) eine konvergente Teilfolge (z_{n_k}) mit $\eta := \lim\limits_{n\to\infty} z_{n_k} \in K$.
da aber jede Teilfolge einer konvergenten Folge den selben Grenzwert wie die Ausgangsfolge hat, gilt also $\xi = \eta \in K$. K ist also abgeschlossen.

Wir zeigen, dass K auch beschränkt ist. Wäre dies nämlich nicht der Fall, dann gibt es zu jedem $n \in \mathbb{N}$ ein $z_n \in K$ mit $|z_n| > n$. Eine solche Folge (z_n) kann aber keine konvergente Teilfolge besitzen.

□

Wir betrachtet nur stetige Funktionen f mit *kompaktem* Definitionsbereich X, dabei sei X zunächst eine beliebige nicht leere kompakte Teilmenge $X \subset \mathbb{K}$.

13.3.11 Fundamentallemma

Ist $K \subset \mathbb{K}$ und kompakt und $f : K \to \mathbb{K}$ stetig, dann ist auch das Bild $f(K)$ kompakt.

Kurz: Stetige Bilder kompakter Mengen sind kompakt

Diesen Argument werden wir häufig anwenden.

Beweis : Wir zeigen, dass mit K auch $Y := f(K)$ die Bolzano-Weierstrass-Eigenschaft für „kompakt“ hat.

Sei dazu $(f(z_n))$ eine Folge in $f(K),\ z_n \in K$. Da K kompakt ist besitzt (z_n) eine Teilfolge (z_{n_k}), die gegen ein Element $\xi \in K$ konvergiert. Wegen der Stetigkeit von f konvergiert dann die Bildfolge $(f(z_{n_k}))$ gegen $f(\xi) \in f(K)$. Nach §13.3.10 ist daher $f(K)$ kompakt.

□

Eine besonders wichtige Anwendung des Fundamentallemmas ist der Satz vom Maximum und Minimum.

13.3.12 Theorem (K.Weierstraß, 1861)

Ist $K \subset \mathbb{K}$ kompakt ($\neq \emptyset$) und $f : K \to \mathbb{R}$ eine stetige Funktion, dann gibt es Elemente $x_{min} \in K$ und $x_{max} \in K$, so dass für *alle* $x \in K$ gilt

$$\boxed{f(x_{min}) \leq f(x) \leq f(x_{max})}$$

M.a.W: Eine stetige reellwertige Funktion auf einer nicht leeren kompakten Menge K nimmt dort ihr globales Maximum und ihr globales Minimum an.

Da $f(K)$ eine kompakte Teilmenge von $\mathbb{R}$ ist, ist $f(K)$ beschränkt (und $\neq \emptyset$), also existieren

$$m := \inf\{f(x);\ x \in K\} \text{ und } M := \sup\{f(x);\ x \in K\}.$$

Weil es Folgen (x_n) mit $x_n \in K$ bzw. (y_n) mit $y_n \in K$ und $m = \lim\limits_{n\to\infty} f(y_n)$ bzw. $M = \lim\limits_{n\to\infty} f(x_n)$ gibt und (y_n) bzw. (x_n) gegen Elemente aus K konvergieren, etwa $x_{min} := \lim\limits_{n\to\infty} y_n \in K$ und

$x_{max} = \lim\limits_{n\to\infty} x_n \in K$, gilt also $m = f(x_{min})$ und $M = f(x_{max})$, d.h. $m \in f(K)$ und $M \in f(K)$ und damit

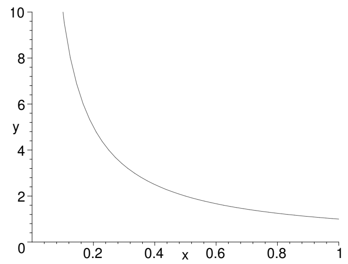
$$f(x_{min}) \le f(x) \le f(x_{max}) \ \text{ für alle } \ x \in K.$$

13.3.13 Bemerkung

Die Kompaktheit ist in §13.3.12 wesentlich, so ist Funktion

$$\begin{aligned} j :]0,1] &\to \mathbb{R}\,, \\ x &\mapsto \frac{1}{x} \end{aligned}$$

stetig auf $]0,1]$, aber nicht beschränkt.



j besitzt zwar ein Minimum (mit dem Wert 1) an der Stelle $x = 1$, aber kein Maximum.

Hätten wir die gleiche Funktion auf dem *offenen* Intervall $]0,1[$ betrachtet, dann besitzt sie dort auch kein Minimum.

13.3.14 Korollar

Ist $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall ($a, b \in \mathbb{R},\ a \le b$) und $f : [a,b] \to \mathbb{R}$ stetig, dann ist auch das Bild $f([a,b])$ ein kompaktes Intervall.

Beweis : Wir wissen schon, dass $f([a,b])$ wieder ein Intervall ist. Da aber $K := [a,b]$ kompakt ist, ist auch $f(K)$ kompakt und

$$m = \min\{f(x);\ x \in [a,b]\} \subset f([a,b])$$

und

$$M = \max\{f(x);\ x \in [a,b]\} \subset f([a,b]),$$

daher gilt $f([a,b]) = [m, M]$.

□

Für komplexwertige Funktionen mit nicht leerem kompakten Definitionsbereich gilt

13.3.15 Korollar

Ist $K \subset \mathbb{K}$ kompakt ($\neq \emptyset$) und $f : K \to \mathbb{C}$ stetig, dann gibt es Elemente x_m und $x_M \in K$ mit

$$|f(x_m)| \le |f(x)| \le |f(x_M)|$$

für alle $x \in K$.

Zum Beweis braucht man nur die auf stetige Funktion $|f|$ den Satz 13.3.12 anzuwenden.

13.3.16 Eine kleine Anwendung

Ist $K \subset \mathbb{K}$ kompakt ($\neq \emptyset$), dann gibt es zu jedem Punkt $p \notin K$ eine Punkt $k \in K$, so dass für *alle* $z \in K$ die Ungleichung

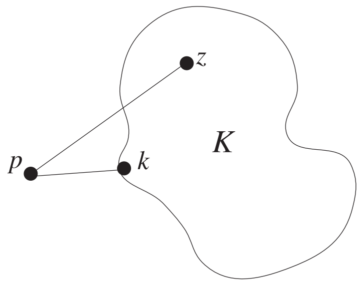
$$|k - p| \leq |z - p|$$

gilt.
Das bedeutet, es gibt einen Punkt $k \in K$, der von p minimalen Abstand hat.

Zum Beweis braucht man nur zu beachte, dass die stetige

$$\begin{array}{rcl} Funktion: \ K & \rightarrow & \mathbb{R}\,, \\ z & \mapsto & |z - p| \end{array}$$

auf K ein absolutes Minimum besitzt.



Zum Abschluss erwähnen wir eine weitere Eigenschaft von stetigen Funktionen mit kompaktem Definitionsbereich: Ist $D \subset \mathbb{K}$ kompakt ($\neq \emptyset$) und $f : K \rightarrow \mathbb{K}$ stetig, dann gilt: Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für *alle* $x, y \in D$ mit $|x - y| < \delta$ gilt

$$|f(x) - f(y)| < \varepsilon.$$

Diese Eigenschaft, die stetige Funktionen mit kompaktem Definitionsbereich haben, ist die sog. *gleichmäßige Stetigkeit*. Eine gleichmäßig stetige Funktion ist natürlich stetig. Im allgemeinen hängt das zu vorgegebenem $\varepsilon > 0$ zu bestimmendes δ von ε (was selbstverständlich ist) und von der betrachteten Stelle $a \in D$ ab, wie unsere konkreten Beispiele zeigen.

Manchmal kann man jedoch ein *universelles* δ finden, das für alle Punkte des Definitionsbereichs die Stetigkeitsbedingung erfüllt. Wir kommen hierauf später (bei der Integralrechnung) ausführlich zurück.

Zeigen Sie als Vorbereitung:

(a)

$$\begin{array}{rcl} f: \ [0,1] & \rightarrow & \mathbb{R}\,, \\ x & \mapsto & x^2 \end{array}$$

ist gleichmäßig stetig auf $[0,1]$

(b)

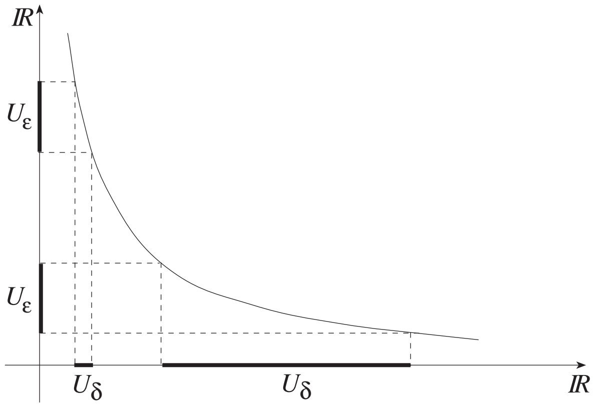
$$\begin{array}{rcl} j:]0,1] & \rightarrow & \mathbb{R}\,, \\ x & \mapsto & \dfrac{1}{x} \end{array}$$

ist stetig, aber nicht gleichmäßig stetig.

Ist $a \in]0,1]$ und $\varepsilon > 0$ beliebig vorgegeben, so kann man etwa $\delta = \min\left\{\frac{a}{2}, \frac{a^2\varepsilon}{2}\right\}$ wählen. Dann gilt für alle $x \in]0,1]$ mit $|x-a| < \delta$

$$\begin{aligned} |j(x) - j(a)| &= \left|\frac{1}{x} - \frac{1}{a}\right| = \left|\frac{x-a}{xa}\right| \\ &\leq \frac{2|x-a|}{a^2} < \frac{2\delta}{a^2} \leq \varepsilon. \end{aligned}$$

Bei jedem $\varepsilon > 0$ muss man also das zugehörige δ immer kleiner wählen je näher a bei Null liegt:



14 Stetige Fortsetzbarkeit. Grenzwerte bei Funktionen

14.1 Stetige Fortsetzbarkeit. Grenzwerte (bei Funktionen)

In 11.1 hatten wir die Stetigkeit einer Funktion $f : D \to \mathbb{K}$ im Punkt $a \in D$ folgendermaßen definiert:
f heißt *stetig* im Punkt $a \in D$, wenn für *jede* Folge $(x_n),\ x_n \in D$ für alle $n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} x_n = a$ die Bildfolge $(f(x_n))$ gegen $f(a)$ konvergiert:

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

Bemerkung: Es hätte genügt zu fordern, dass für *jede* Folge $(x_n),\ x_n \in D$ für alle n und $\lim\limits_{n\to\infty} x_n = a$ die Bildfolge $(f(x_n))$ konvergiert (der Grenzwert all dieser Folgen ist dann notwendig $f(a)$). Es genügt sogar, dass man sich auf solche Folge $(x_n),\ x_n \in D$, beschränkt, die zwar gegen a konvergieren, in denen aber a als Folgeglied gar nicht auftritt (Beweis als Übungsaufgabe).

In diesem Abschnitt wollen wir uns hauptsächlich mit der Frage befassen, ob die gegebene Funktion f die Eigenschaft hat, dass für alle Folgen $(x_n),\ x_n \in D,\ x_n \neq a$ für alle n, die Bildfolgen $(f(x_n))$ konvergieren. Dabei kann a ein Punkt von D sein, a muss aber nicht zu D gehören, wir setzen aber voraus, dass es mindestens eine Folge $(x_n),\ x_n \in D,\ x_n \neq a$ für alle n und $\lim\limits_{n\to\infty} x_n = a$ gibt. Anschaulich bedeutet dies, dass man sich der Stelle a mit Punkten aus $D - \{a\}$ ***beliebig*** nähern kann. Im Folgenden sei oder $D \subset \mathbb{K}$ eine nichtleere Teilmenge ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) und $f : D \to \mathbb{K}$ eine Funktion und $a \in \mathbb{K}$ ein Punkt der folgende Eigenschaft besitzt:

14.1.1 (Häufungspunkt):

$a \in \mathbb{K}$ heißt *Häufungspunkt* von $D(\subset \mathbb{K})$, wenn es (mindestens) eine Folge (x_n) gibt mit $x_n \in D$ und $x_n \neq a$ für alle $n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} x_n = a$.

14.1.2 Bemerkungen:

1. Ein Häufungspunkt einer Menge D braucht kein Element von D zu sein. Ist z. B. $D :=]0,1]$, so ist 0 zwar ein Häufungspunkt von D, gehört aber nicht zu D (die Folge $(x_n) = \left(\frac{1}{n}\right)$ hat offensichtlich die in der Definition geforderte Eigenschaft).
2. Ist $D := [0,1]$, dann ist jeder Punkt von D auch Häufungspunkt von D.
3. Ist D endlich, so ist die Menge der Häufungspunkte von D die leere Menge.
4. Die Menge der Häufungspunkte von $\mathbb{Q}$ ist die Menge $\mathbb{R}$ der reellen Zahlen.
5. Die Menge $\left\{\frac{1}{n};\ n \in \mathbb{N}\right\}$ hat als einzigen Häufungspunkt in $\mathbb{R}$ den Nullpunkt.

   Beachte: Betrachtet man die Menge $M := \left\{\frac{1}{n};\ n \in \mathbb{N}\right\}$ als Teilmenge von $\mathbb{R} - \{0\}$, dann hat M keinen Häufungspunkt in $\mathbb{R} - \{0\}$.

3. ergibt sich z. B. aus

14.1.3 (ε-Charakterisierung von Häufungspunkten)

$a \in \mathbb{K}$ ist genau dann Häufungspunkt von $D \subset \mathbb{K}$, wenn in *jeder* ε-Umgebung $U_a(a) := \{z \in \mathbb{K};\ |z-a| < \varepsilon\}$ mindestens ein von a verschiedener Punkt $z \in D$ liegt.

Beweis als Übungsaufgabe.

Führt man den Begriff der *punktierten ε-Umgebung* ein: $\dot{U}_\varepsilon(a) := \{z \in \mathbb{K};\ 0 < |z-a| < \varepsilon\}$, dann ist $a \in \mathbb{K}$ genau dann Häufungspunkt von D, wenn $\dot{U}_\varepsilon(a) \cap D \neq \emptyset$ für jedes $\varepsilon > 0$ gilt. In jeder ε-Umgebung von a liegen sogar unendlich viele von a verschiedene Punkte von D.

14.1.4 Beachte:

(a) Ist $a \in \mathbb{K}$ Häufungspunkt von D, dann ist a auch Häufungspunkt jeder Obermenge $\tilde{D}$ von D.

(b) Man verwechsle nicht den Begriff der Häufungspunkte einer Menge mit dem Begriff des *Hauptwertes* einer Folge p.t.o. Ist z.B. (x_n) eine konstante Folge (reelle neue komplexe Zahl, also $x_n = c$ mit einem geeigneten $c \in \mathbb{K}$, dann ist c zwar Häufungswert von (x_n), aber die Menge $\{x_n, n \in \mathbb{N}\}$ der Folgenglieder besteht nun aus der einelementigen Menge $\{c\}$, hat also keinen Häufungspunkt.

Die eingangs gestellte Frage über die Konvergenz der Bildfolgen hängt eng zusammen mit folgendem

14.1.5 Fortsetzungsproblem

Gegeben eine Funktion $f : D \to \mathbb{K}$ und ein Punkt $a \in \mathbb{K}$, der nicht zu D gehören muss, aber gehören darf. Frage: Gibt es eine Funktion

$$\tilde{f} : D \cup \{a\} \to \mathbb{K},$$

die in a stetig ist und deren Einschränkung auf $D - \{a\}$ mit f übereinstimmt: $\tilde{f}|\, D - \{a\} = f$? Im Existenzfall nennen wir eine solche Funktion eine *stetige Fortsetzung von* f. Selbst wenn der Definitionsbereich D von f "einfach" ist, etwa ein offenes Intervall oder eine punktierte r-Umgebung, und f stetig auf ganz D ist, so braucht es keine stetige Fortsetzung zu geben:

14.1.6 Beispiele:

1.
$$\begin{aligned} f :]0,1[&\to \mathbb{R},\ a = 0. \\ x &\mapsto \frac{1}{x}; \end{aligned}$$

f lässt sich nicht stetig nach Null fortsetzen, denn für die Folge $(x_n) = \left(\frac{1}{n}\right)$ ist die Bildfolge $(f(x_n)) = (n)$ divergent.

2.
$$\begin{aligned} f : \dot{U}_1(0) &\to \mathbb{R},\ a = 0. \\ x &\mapsto \frac{x^2}{|x|}. \end{aligned}$$

Für $x \neq 0$ gilt hier $\frac{x^2}{|x|} = |x|$ und

$$\begin{aligned} \tilde{f} : U_1(0) &\to \mathbb{R} \\ x &\mapsto |x| \end{aligned}$$

ist eine stetige Fortsetzung von f mit $\tilde{f}(0) = 0$.

3.
$$\begin{aligned} f : \mathbb{C} - \{0\} &\to \mathbb{C};\ a = 0; \\ z &\mapsto \frac{\exp(z) - 1}{z}. \end{aligned}$$

Hier ist

$$\tilde{f} : \mathbb{C} \to \mathbb{C} \ \text{ mit}$$
$$\tilde{f}(z) = \begin{cases} \frac{\exp(z)-1}{z}, & z = 0 \\ 1, & z = 0 \end{cases}$$

eine stetige Fortsetzung von f (wie wir sehen werden).

Das Fortsetzungsproblem ist nur sinnvoll, wenn a *Häufungspunkt* von D ist, denn wenn a kein Häufungspunkt von D ist, dann wird $\tilde{f}$ bei *jeder* Festsetzung des Funktionswertes $\tilde{f}(c)$ stetig bei a.
Sei z.B.

$$D := \quad]-\infty, 0] \cup [2, \infty[$$
$$f(x) = \quad \begin{cases} 0, \text{ für } x \leq 0 \\ 1, \text{ für } x \geq 2 \end{cases}$$

und z. B.

$$\tilde{f} : D \cup \{1\} \to \mathbb{R} \text{ mit}$$
$$\tilde{f}|D = f \text{ und}$$
$$\tilde{f}(1) = c$$

dann ist $\tilde{f}$ stetig in 1 für jede Wahl $c \in \mathbb{R}$. (1 ist isolierter Punkt des Definitionsbereichs von $\tilde{f}$). Ist jedoch $a \in \mathbb{K}$ ein Häufungspunkt von D, dann besitzt f höchstens eine in a stetige Fortsetzung $\tilde{f}$ (Beweis als Übungsaufgabe). Wenn f eine in a stetige Fortsetzung besitzt, so sagt man auch, dass f bei *Annäherung an* a *einen Grenzwert* besitzt. Genauer definieren wir:

14.1.7 Definition:

Die Funktion

$$f : D \to \mathbb{K}$$

besitzt bei Annäherung an den Häufungspunkt a von D den *Grenzwert* $\ell(\in \mathbb{K})$, wenn die Funktion

$$\tilde{f} : D \cup \{a\} \rightharpoonup \mathbb{K} \text{ mit}$$

$$(*)\ \tilde{f}(z) = \begin{cases} f(z); & z \in D,\ z \neq a \\ \ell, & z = a \end{cases}$$

im Punkt a stetig ist.

Sprechweise: Man sagt auch f konvergiert bei Annäherung an a gegen ℓ und verwendet die

Schreibweise: $\lim\limits_{z \to a} f(z) = \ell$ oder ´kurz $f(z) \to \ell$ für $z \to a$ (da ℓ eindeutig bestimmt ist (Begründung?), sind diese Sprech-und Schreibweisen gerechtfertigt. Der Wert $\tilde{f}(a)$ der in a stetigen Fortsetzung $\tilde{f}$ von f ist also der Grenzwert von f bei der Annäherung an a. Ist $a \in D$, dann ist $\tilde{f} = f$ und $\ell = f(a)$ und die Stetigkeit f in a ist äquivalent mit:

$$(**) \qquad \boxed{\lim_{z \to a} f(z) = f(a)}$$

Der Grenzwert von f in a stimmt also im Fall der Stetigkeit mit dem Funktionswert $f(a)$ überein und umgekehrt:
Falls der Grenzwert bei a mit dem Funktionswert übereinstimmt, ist f stetig in a.

14.1.8 Bemerkung:

1. In der Schulmathematik wird $(**)$ meist als Definition für die Stetigkeit verwendet. Der Grenzwertbegriff ist jedoch komplizierter als der Stetigkeitsbegriff, von Häufungspunkten ist in der Schulmathematik kaum die Rede, da die Definitionsbereich meist echte Intervalle sind und für diese jeder Punkt auch Häufungspunkt ist.

2. Man beachte, dass es bei der Definition des Grenzwertes keine Rolle spielt, ob die Funktion f in a definiert ist und welchen Wert sie dort hat. Die $\varepsilon - \delta$-Definition der Stetigkeit von $\tilde{f}$ übersetzt sich für den Grenzwert von f in die folgende

14.1.9 Satz ($\varepsilon - \delta$-Charakterisierung des Grenzwertes)

$f : D \to \mathbb{K}$ hat im Häufungspunkt a von D $(a \in \mathbb{K})$ genau dann den Grenzwert $\ell \in \mathbb{K}$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $z \in D$ mit $0 < |z - a| < \delta$ gilt

$$|f(z) - \ell| < \varepsilon\,.$$

Beweis: Die Existenz des Grenzwertes ℓ ist äquivalent mit der Stetigkeit von $\tilde{f}$ in a (ist $a \in D$ so ist $\tilde{f} = f$). Die Stetigkeit von $\tilde{f}$ besagt: Zu jedem $\varepsilon \to 0$ gibt es ein $\delta > 0$, so dass für alle $z \in D \cup \{a\}$ mit $|z - a| < \delta$ gilt

$$|\tilde{f}(z) - \ell| < \varepsilon\,.$$

Diese Bedingung ist jedoch zu

$$z \in D,\ z \neq a,\ |z - a| < \delta \Rightarrow |f(z) - \ell| < \varepsilon$$

äquivalent wegen

$$\tilde{f}(z) = \begin{cases} f(z), z \in D, z & \neq a \\ \ell,\ \text{für}\ z & = a\,. \end{cases}$$

Man hätte die Grenzwertdefinition auch über Folgen einführen können:

14.1.10 Satz (Folgenkriterium für Grenzwerte):

Die Funktion $f : D \to \mathbb{K}$ hat im Häufungspunkt a von $D(a \in \mathbb{K})$ genau dann den Grenzwert $\ell \in \mathbb{K}$, wenn für jede Folge $(z_n),\ z_n \in D - \{a\}$ für alle n und $\lim\limits_{n\to\infty} z_n = a$ gilt

$$\boxed{\lim_{n\to\infty} f(z_n) = \ell}$$

Beweis: Mit der durch $(*)$ erklärten stetigen Fortsetzung $\tilde{f}$ von f bestehen folgende Äquivalenzen $\lim\limits_{z\to a} f(z) = \ell \Leftrightarrow \tilde{f}$ ist stetig in a $\Leftrightarrow$ die Folgenbedingung $(***)$ ist erfüllt (beachte für $z_n \in D - \{a\}$, ist $\tilde{f}(z_n) = f(z_n)$).

Aus den Rechenregeln für Folgen bzw. den Permanenzeigenschaften stetiger Funktionen ergeben sich mittelbar folgende *Rechenregeln* für *Grenzwerte* von Funktionen.

Regel 0 : $\lim\limits_{z\to a} f(z) = 0 \Leftrightarrow \lim\limits_{z\to a} |f(z)| = 0$

Regel 1 : Haben die Funktionen $f : D \to \mathbb{K}$ bzw. $g : D \to \mathbb{K}$ im Häufungspunkt a ihres Definitionsbereichs die Grenzwerte ℓ bzw. ℓ', dann haben auch die Funktionen $f{\pm}g : D \to K,\ f{\cdot}g : D \to \mathbb{K}$ und falls $\lim\limits_{z\to a} g(z) \neq 0$, auch $\frac{f}{g}$ und f (definiert in eine geeignete punktierte Umgebung von 0) und es gibt

(a) $\lim\limits_{z\to a}(f(z) \pm g(z)) = \ell \pm \ell' = \lim\limits_{z\to a} f(z) \pm \lim\limits_{z\to a}(z)$.

(b) $\lim\limits_{z\to a}(f(z)g(z)) = \ell \cdot \ell' = \lim\limits_{z\to a} f(z) \cdot \lim\limits_{z\to a} g\, z$.

(c) $\lim\limits_{z\to a} \frac{f(z)}{f(z)} = \frac{\ell}{\ell'} = \frac{\lim_{z\to a} f(z)}{\lim_{z\to a} f(z)}$.

Regel 2 : Gegeben $D \xrightarrow{f} \mathbb{E} \xrightarrow{g} > \mathbb{K}$. Es gelte $\lim\limits_{z\to a} = \ell \in E$ und g sei stetig in ℓ. Dann gilt $\lim\limits_{z\to a}(f(z)) = g(\ell)$.

Beispiel: Ist $\lim\limits_{z\to a} f(z) = \ell$ so folgt $\lim\limits_{z\to a} |f(z)| = |\ell|$

Regel 3 : Sind f und g reellwertige und besitzen f und g Grenzwerte in a, dann folgt aus $f(x) \leq f(x)$ für alle x aus einer punktierten Umgebung von a auch $\lim\limits_{x\to a} f(x) \leq \lim\limits_{x\to a} f(x)$.

Regel 4 : (Analogon zum Sandwich-Theorem:) Sind f und g reellwertig und gilt

$$\lim_{x\to a} f(x) = \lim_{x\to a} f(x) =: \ell$$

und ist h eine Funktion mit $f(x) \leq h(x) \leq g(x)$ für alle x aus einer punktierten Umgebung von a, dann existiert auch $\lim\limits_{x\to a} h(x)$ und es gilt $\lim\limits_{x\to a} h(x) = \ell$.

Wenn man nachweisen muss, dass $f \to a$ einen Grenzwert besitzt, ohne dass man diesen kennt, kann man benutzen:

14.1.11 Cauchy-Kriterium für Grenzwerte

f hat im Häufungspunkt a ihres Definitionsbereiches D genau dann einen Grenzwert, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass gilt:

CB: Für alle $x, y \in D,\ x \neq a,\ y \neq a$ mit $|x-a| < \delta,\ |y-a| < \delta$ gilt

$$|f(x) - f(y)| < \varepsilon\,.$$

Beweis: Wenn f in a einen Grenzwert ℓ besitzt, so folgt aus der $\varepsilon - \delta$-Charakterisierung leicht die angegebene Cauchy-Bedingung.

Sei nun umgekehrt die im Kriterium formulierte Cauchy-Bedingung erfüllt. Zu jeder Folge $(x_n),\ x_n \in D,\ x_n \neq a$ und $\lim\limits_{n\to\infty} x\,n = a$ ist die Bildfolge $f(x_n)$ eine Cauchy-Folge in $\mathbb{K}$, also konvergent. Denn zu beliebig vorgegebenen $\varepsilon > 0$ gibt es ein $\delta > 0$ und eine $N \in \mathbb{N}$, so dass für alle $n, m \geq N$ gilt $x_n \in \dot{U}_\delta(a),$ d.h. $0 < |x_n - a| < \delta$ bzw. $x_m \in \dot{U}_\delta(a)$ d.h. $0 < |x_m - a| < \delta$, also gilt auch

$$|f(x_n) - f(x_m)| < \varepsilon\,.$$

Je zwei solcher Cauchy-Folgen müssen aber denselben Grenzwert haben, denn ist (y_n) eine weitere Folge mit $y_n \in D - \{a\}$ und $\lim\limits_{n\to\infty} y_n = a$, dann gilt für die "Mischungsfolge" (auch Reißverschlussfolge genannt) $(z_n) := (x_1, y_1, x_2, y_2, x_3, y_3, \dots)$ ebenfalls $z_n \in D - \{a\}$ und $\lim\limits_{n\to\infty} z_n = a$ und $(f(z_n))$ ist als Cauchy-Folge konvergent. $(f(x_n))$ und $(f(y_n))$ sind aber Teilfolgen von $(f(z_n))$, haben also denselben Grenzwert.

14.1.12 Beispiele

1. $$\lim_{x\to\infty} \frac{x^2}{|x|} = 0\,,$$

   dem

   $$\tilde{f}(x) = \begin{cases} \frac{x^2}{|x|}, & \text{falls } x \neq 0 \\ \\ |x|, & \text{falls } x = 0 \end{cases}$$

   ist eine stetige Fortsetzung von

   $$\begin{array}{rcl} f : \mathbb{R} - \{0\} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{x^2}{|x|} \end{array}$$

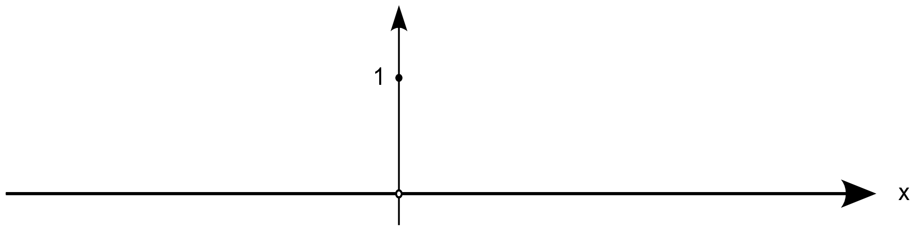
   mit $\tilde{f}(0) = 0\,.$

2. $\lim\limits_{x\to 1} \frac{\sqrt{x}-1}{x-1} = \frac{1}{2}, \quad (D = \mathbb{R}_+ - \{1\})$, denn $x \mapsto \frac{1}{\sqrt{x}+1}$ liefert die stetige Fortsetzung

3. $\lim\limits_{z\to a} \frac{z^k-a^k}{z-a} = ka^{k-1}$

   Beweis: Für $z \in \mathbb{C} - \{a\}$ sei
   $$g(z) = \frac{z^k - a^k}{z - a}\,.$$
   Nun gilt aber die Faktorisierung $z^k - a^k = (z-a)(z^{k-1} + z^{k-2}a + \cdots + z \cdot a^{k-2} + a^{k-1})$ und das Polynom
   $$\begin{array}{rcl} \tilde{f} : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & z^{k-1} + z^{k-1}a + \cdots + a^{k-1} \end{array}$$
   ist insbesondere stetig in a und hat auch den Funktionswert
   $$\tilde{f}(a) = ka^{k-1}\,.$$
   Daher ist
   $$\boxed{\lim_{z\to a} g(z) = \lim_{z\to a} \frac{z^k - a^k}{z - a} = ka^{k-1}\,.}$$

4. Die Funktion
   $$f : \mathbb{R} \to \mathbb{R} \text{ mit}$$
   $$f(x) = \begin{cases} 0, & \text{falls} \quad x \neq 0 \\ 1, & \text{falls} \quad x = 0\,. \end{cases}$$



   f ist unstetig an der Stelle 0 für jede Nullfolge (x_n) aus $\mathbb{R} = -\{0\}$ konvergiert aber die Bildfolge $(f(x_n))$ gegen Null. Durch $\tilde{f}(x) = 0$ für *alle* $x \in \mathbb{R}$ erhält man die stetige Fortsetzung von f.

   Hier (und im vorigen Beispiel) haben wir jeweils eine Ableitung ausgerechnet, nämlich die Ableitung der Potenzfunktion
   $$\begin{array}{rcl} p : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & z^k \end{array}$$
   an der Stelle $a \in \mathbb{C} : p'(a) = ka^{k-1}$. Eine wichtige Anwendung findet die Theorie der Grenzwerte in der Differentialrechnung. Auch das nächste Beispiel fällt in diese Kategorie.

5. $\lim\limits_{z\to 0} \dfrac{\exp(z) - 1}{z} = 1$ und

6. $\lim\limits_{z\to a} \dfrac{\exp(z)-\exp(a)}{z-a} = \exp(a)\,.$

Aus der Darstellung (vergl. ...) $\exp(z) = \sum\limits_{k=0}^{n} \frac{z^k}{k!} + R_{n+1}(z)$ mit $|R_{n+1}(z)| \leq 2\dfrac{|z|^{n+1}}{(n+1)!}$ für $|z| \leq 1$ folgt für $n = 1$:

$\exp(z) = 1 + z + R_2(z)$ mit $|R_2(z) \leq 2\frac{|z|^2}{2!} = |z|^2$, und daher $\left|\dfrac{\exp(z)-1}{z} - 1\right| \leq |z|$ für $0 < |z| \leq 1$ und damit ergibt sich

$$\boxed{\lim_{z\to 0} \frac{\exp(z)-1}{z} = 1}$$

Bemerkung: Mit Hilfe dieses Grenzwertes 5. (oder auch direkt) kann man leicht

$$\boxed{\lim_{z\to 0} \frac{\sin z}{z} = 1 \quad (=\cos(0)\,)}$$

zeigen (vergl. auch Abschnitt 16).

$\sin : \mathbb{C} \to \mathbb{C}$ ist dabei durch die für alle $z \in \mathbb{C}$ absolut konvergente Potenzreihe $\sum\limits_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!}$ definiert.

Die zweite Grenzwertaussage wird mittels der Funktionalgleichung der Exponentialfunktion auf die erste zurückgeführt:

Für $z \neq a$ ist (Funktionalgleichung) $\dfrac{\exp(z)-\exp(a)}{z-a} = \exp(a) \cdot \dfrac{\exp(z-a)-1}{z-a}$ und damit

$$\begin{aligned}\lim_{z\to a} \frac{\exp(z)-\exp(a)}{z-a} &= \exp(a) \lim_{z\to a} \frac{\exp(z-a)-1}{z-a} \\ &= \exp(a) \cdot 1 \ \text{(nach 5.)} \\ &= \exp(a)\,.\end{aligned}$$

Da der linke Grenzwert dem Wert der Ableitung der Exponentialfunktion an der Stelle a liefert, haben wir gezeigt: $\exp : \mathbb{C} \to \mathbb{C}$ ist für alle $a \in \mathbb{C}$ (komplex) differenzierbar, und es gilt $\exp'(a) = \exp(a)$ für alle $a \in \mathbb{C}$. Speziell ist $\exp'(0) = \exp(0) = 1$.

Mit 5. zeigt man leicht $(x \in \mathbb{R})$

7. $\lim\limits_{x\to 0} \dfrac{\log(1+x)}{x} = 1$ (dabei ist $\log$ der natürliche Logarithmus). Dieser Grenzwert lässt sich ebenfalls als Ableitung interpretieren.

14.2 Einseitige Grenzwerte

In den bisherigen Überlegungen spielte es i. a. keine Rolle, ob D eine Teilmenge von $\mathbb{R}$ oder $\mathbb{C}$ ist. Jetzt wollen wir ausnutzen, dass $\mathbb{R}$ ein *angeordneter Körper* ist, wir betrachten daher ab jetzt ausschließlich Funktionen $f : D \to \mathbb{K}$ mit einem Definitionsbereich $D \subset \mathbb{R}$ (die Werte von f dürfen zunächst auch komplex sein). Spezielle Grenzwerte sind die sogenannten "einseitigen" Grenzwerte.

14.2.1 Definition:

Sei $a \in \mathbb{R}$ Häufungspunkt von $D\cap]-\infty, a[$ bzw. $D\cap]a, \infty[$.

Man sagt: f hat bei Annäherung an a (oder in a) den linksseitigen bzw. rechtsseitigen Grenzwert $\ell \in \mathbb{C}$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass gilt

$$|f(x) - \ell| < \varepsilon \begin{cases} \text{für alle } x \in D\cap]x_0 - \delta, a[\text{ bzw.} \\ \text{für alle } x \in D\cap]a, x_0 + \delta[\end{cases}$$

Schreibweisen:

$$\ell = \lim_{\substack{x \to x_0 \\ x < x_0}} f(x) = f(x_0-) \quad \text{(linksseitiger Grenzwert)}$$

$$\ell = \lim_{\substack{x \to x_0 \\ x > x_0}} f(x) = f(x_0+) \quad \text{(rechtsseitiger Grenzwert)}$$

14.2.2 Beispiele:

1. $$\lim_{\substack{x \to 0 \\ x > 0}} \frac{x}{|x|} = 1, \quad \lim_{\substack{x \to 0 \\ x < 0}} \frac{x}{|x|} = -1$$

2. Die Gauß-Klammer

$$\begin{array}{rcl} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & [x] = \max\{k \in \mathbb{Z},\ k \leq x\} \end{array}$$

besitzt an jeder Stelle $k \in \mathbb{Z}$ den linksseitigen Grenzwert $k-1$ und den rechtsseitigen Grenzwert k. Die einseitigen Grenzwerte kann man auch als gewöhnliche Grenzwerte auffassen, man beachte f nun auf $M := D\cap]-\infty, a[$ bzw. $M := D\cap]a, \infty[$ einzuschränken.

Für die einseitigen Grenzwerte gibt es auch eine offensichtliche Folgencharakterisierung, (versuchen Sie diese zu formulieren).

14.2.3 Satz:

Hat eine Funktion $f : D \to \mathbb{K}$ $(D \subset \mathbb{R})$ in einem Punkt $a \in \mathbb{R}$ sowohl einen linksseitigen Grenzwert als auch einen rechtsseitigen, und sind diese beiden Grenzwerte gleich (etwa gleich ℓ), dann besitzt f in a auch einen Grenzwert und dieser stimmt mit den einseitigen Grenzwerten überein.

Beweis: Wir benutzen eine $\varepsilon - \delta$-Charakterisierung für die fraglichen Grenzwerte. Dazu sei $\varepsilon > 0$ beliebig vorgegeben. Dann gibt es Zahlen $\delta_1 > 0$ und $\delta_2 > 0$, so dass für alle $x \in U_{\delta_1}(a)\cap]-\infty, a[$ bzw. $x \in U_{\delta_2}(a)\cap]a, \infty[$ gilt $|f(x) - \ell| < \varepsilon$ bzw. $|f(x) - \ell| < \varepsilon$.
Setzt man $\delta := \min\{\delta_1, \delta_2\}$, dann gilt für alle $x \in U_\delta(a) \cap D$ $|f(x) - \ell| < \varepsilon$, d.h. f besitzt ebenfalls den Grenzwert ℓ.

Der folgende Satz zeigt, dass *monotone* Funktionen auf Intervallen immer einseitige Grenzwerte besitzen (möglicherweise uneigentliche).

14.2.4 Satz:

Sei $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$ eine monotone Funktion.

(a) f besitzt in jedem *inneren Punkt* von D einen links- und rechtsseitigen Grenzwert.

(b) Hat D einen linken Randpunkt, der zu D gehört, so besitzt f in a einen rechtsseitigen Grenzwert.

(c) Hat D einen linken Randpunkt, der nicht zu D gehört, so hat f in a einen rechtsseitigen Grenzwert, falls f auf $U_\delta(a) \cap D$ für ein geeignetes $\delta > 0$ beschränkt ist.

(d) Analoge Aussagen zu (b) und (c) gelten für einen rechten Randpunkt von D.

14.2.5 Beweis:

Wir beweisen nun die Aussage (a) und nehmen OBdA an, dass f monoton wächst (sonst betrachte man $-f$). Ist $x_0 \in D$ ein innerer Punkt von D, dann sind die Mengen

$$\begin{aligned} U &:= \{f(x);\quad x \in D;\quad x < x_0\} \text{ und} \\ O &:= \{f(x);\quad x \in D;\quad x > x_0\} \text{ nicht leer.} \end{aligned}$$

$f(t)$ ist eine obere Schranke für U und eine untere Schranke für O, daher existieren $s := \sup U$ und $t := \inf 0$.
Wir behaupten

$$\lim_{\substack{x \to x_0 \\ x < x_0}} f(x) = s \text{ und } \lim_{\substack{x \to x_0 \\ x > x_0}} f(x) = t\,.$$

Dazu verwenden wir die ε-Charakterisierung von $\sup$ bzw. $\inf$. Zu beliebig vorgegebenem $\varepsilon > 0$ gibt es ein $\xi \in D$ mit $\xi < x_0$ und $s - \varepsilon < f(\xi) \le s$. Nach Definition von s und wegen der Monotonie von f gilt dann $s - \varepsilon < f(x) \le s (\le s + \varepsilon)$ für alle $x \in]\xi, x_0[=]-\infty, x_0[\cap D$. Das bedeutet

$$\lim_{\substack{x \to x_0 \\ x < x_0}} f(x) = s\,.$$

Völlig analog zeigt man

$$\lim_{\substack{x \to x_0 \\ x > x_0}} f(x) = t\,.$$

Aus dem bewiesenen Satz kann man interessante Folgerungen ziehen:
Eine auf einem Intervall D monotone Funktion f ist genau dann in einem inneren Punkt unstetig, wenn

$$f(x_0-) = \lim_{\substack{x \to x_0 \\ x < x_0}} f(x) \ne \lim_{\substack{x \to x_0 \\ x > x_0}} f(x) = f(x_0+)$$

gilt. Eine solche Unstetigkeitsstelle x_0 von f (man nennt sie auch Unstetigkeitsstelle 1. Art), können wir eine rationale Zahl $r(x_0)$ zuordnen, die echt zwischen $f(x_0)-$ und $f(x_{0+})$ liegt. Für eine monoton wachsende Funktion f gilt daher

$$f(x_0-) < r(x_0) < f(x_0+)$$

und entsprechend für eine monoton fallende Funktion

$$f(x_0-) > r(x_0) > f(x_0+)\,.$$

Ist M die Menge der im Inneren D gelegenen Unstetigkeitsstellen von f, so wird durch

$$\begin{aligned} r : M &\to \mathbb{Q} \\ x &\mapsto r(x) \end{aligned}$$

eine Funktion definiert, deren Definitionsbereich M ist und deren Werte in $\mathbb{Q}$ liegen. Diese Funktion ist injektiv, denn sind $x_1, x_2 \in M$ und ist $x_1 < x_2$, so folgt

$$f(x_1-) < r(x_1) < f(x_1+) \le f(x_2-) < r(x_2) < f(x_2+)$$

falls f monoton wächst (analog falls f monoton fällt). Als Teilmenge der abzählbaren Menge $\mathbb{Q}$ ist der Bildbereich von r abzählbar und da r injektiv ist, gilt dies für M. Damit haben wir das bemerkenswerte Ergebnis erhalten:

Die Menge der Unstetigkeitsstellen einer auf einem Intervall monotonen Funktion ist abzählbar.

(denn zu M können höchstens noch die beiden Randpunkte von D hinzukommen, falls f dort unstetig sein sollte).

Beispiel: Schränkt man die Funktion aus Übungsaufgabe 47 (SS 2006) auf das Intervall $[0, 1]$ ein; so erhält man eine monotone Funktion mit einer abzählbar unendlichen Menge von Unstetigkeiten.
Die Unstetigkeiten 1. Art nennt man auch *Sprungstellen*.
Wenn mindestens eine der beiden einseitigen Grenzwerte $f(x_0-)$ bzw. $f(x_0+)$ nicht existiert, spricht man auch von Unstetigkeitsstellen *zweiter Art*, so hat etwa die Dirichlet-Funktion

$$\delta : \mathbb{R} \to \mathbb{R},\ \delta(x) = \begin{cases} 1, & \text{falls } x \in \mathbb{Q} \\ 0, & \text{falls } x \in \mathbb{R} - \mathbb{Q} \end{cases}$$

nur Unstetigkeitsstellen zweiter Art.

14.3 Weitere Schreib- und Sprechweisen bei Grenzwerten

14.3.1 Die Schreibweise

$$\lim_{x\to\infty} f(x) = \ell \text{ oder } f(x) \to \ell \ \text{ für } \ x \to \infty)$$

bedeutet:

Der Definitionsbereich D von f ist nicht nach oben beschränkt und zu jedem $\varepsilon > 0$ gilt ein $R \subset \mathbb{R}$, so dass für alle $x \in D$ mit $x > R$ gilt

$$|f(x) - \ell| < \varepsilon\,.$$

Analoge Definition für $\lim\limits_{x\to-\infty} f(x)$, falls D nicht nach unten beschränkt ist. Geometrische Interpretation für reellwertige f:

Für alle $x > R$ verläuft der Graph von f ganz im ε-Streifen um ℓ, d.h. in

$$S_\varepsilon(\ell) := \{(x, y) \in \mathbb{R} \times \mathbb{R};\ \ell - \varepsilon < y < \ell + \varepsilon\}$$

14.3.2 Bemerkung: (Bei Königsberger 'Reduktionslemma')

Es gilt: $\lim\limits_{x\to\infty} f(x)$ existiert genau dann, wenn die Funktion φ mit $\varphi(t) := f(\frac{1}{t})$ falls $\frac{1}{t} \in D$ bei 0 einen rechtsseitigen Grenzwert besitzt und es gilt dann

$$\lim_{\substack{t\to 0\\ t>0}} \varphi(t) = \lim_{x\to\infty} f(x)$$

Der Beweis ist klar, denn die Bedingung $|\varphi(t) - \ell| < \varepsilon$ für $0 < t < \delta$ ist äquivalent mit $|f(x) - \ell| < \varepsilon$ für $\ x > \frac{1}{\delta} > 0\,.$

14.3.3 Beispiele:

$\lim\limits_{x\to\infty} \frac{1}{x} = 0, \quad \lim\limits_{x\to-\infty} \frac{1}{x} = 0 \qquad \lim\limits_{x\to\infty} \frac{x^2-2006}{x^2+2006} = 1, \quad \lim\limits_{x\to-\infty} \exp(x) = 0 \qquad \lim\limits_{x\to\infty} \left(\sqrt{x+1} - \sqrt{x}\right) = 0$

Die Beweise sind einfach und seien dem geneigten Leser überlassen.

Sei

$$\begin{array}{rccl} p: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & a_n x^n + \cdots + a_0,\ a_n \neq 0 \end{array}$$

ein Polynom, dann gilt

$$\lim_{x\to\infty} \frac{P(x)}{a_n x^n} = \lim_{\substack{t\to 0\\ t>0}} \frac{a_n + a_{n-1}t + \cdots + a_0 t^n}{a_n} = 1$$

Man sagt: Das Polynom $P(x)$ ist asymptotisch gleich $a_n x^n$.

Dabei nennt man zwei Funktionen (mit nicht nach oben beschränktem Definitionsbereich) $f, g : D \to \mathbb{R}$ asymptotisch gleich für $x \to \infty$, falls $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} = 1$ gilt.

Hierfür verwendet man häufig das Zeichen

$$f(x) \simeq g(x) \text{ für } x \to \infty\,.$$

14.3.4 Uneigentliche Grenzwerte

Jetzt seien die Werte der betrachteten Funktionen ebenfalls reell. Wir betrachten als einführende Beispiele die Funktionen

1. $\begin{array}{rcl} f : \mathbb{R} - \{0\} & \to & \mathbb{R} \\ x & \mapsto & \frac{1}{x} \end{array}$ und die Funktionen

2. $\begin{array}{rcl} g : \mathbb{R} - \{1\} & \to & \mathbb{R} \\ x & \mapsto & \frac{1}{(x-1)^2} \end{array}$.

3. $\begin{array}{rcl} h : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^k,\ k \in \mathbb{N} \end{array}$.

   (a) Für jede Folge (x_n) mit $x_n > 0$ für alle n und $\lim\limits_{n\to\infty} x_n = 0$ wächst die Bildfolge $(f(x_n)) = \left(\frac{1}{x_n}\right)$ über alle Grenzen. ($(f(x_n))$ ist divergent mit dem uneigentlichen Grenzwert $+\infty$).

   (b) Für jede Folge (x_n) mit $x_n < 0$ für alle n und $\lim\limits_{n\to\infty} x_n = 0$ ist die Bildfolge $(f(x_n)) = (\frac{1}{x_n})$ bestimmt divergent mit dem uneigentlichen Grenzwert $-\infty$.

   (c) Für die Nullfolge $(x_n) = \left(\frac{(-1)^n}{n}\right)$ ist die Bildfolge $f(x_n) = \frac{1}{x_n}$ divergent.

Im Beispiel 2. gilt für jede Folge (x_n) mit $\lim\limits_{n\to\infty} x_n = 1$ und $x_n \neq 1$ für alle $n \in \mathbb{N}$:

$$\lim_{n\to\infty} f(x_n) = \infty\,.$$

Im Beispiel 3. gilt für jede bestimmt gegen $+\infty$ divergente Folge $(x_n),\ x_n \in \mathbb{R}$ für alle $n \in \mathbb{N}$

$$\lim_{n\to\infty} x_n^k = +\infty\,.$$

Folgende Schreib- und Sprechweisen liegen nahe:

1. $\lim\limits_{x\to a} f(x) = +\infty$ (oder $f(x) \to \infty$ für $x \to a$) bedeutet:

   a ist Häufungspunkt des Definitionsbereichs von f und für jede gegen a konvergierende Folge $(x_n),\ x_n \in D - \{a\}$, ist die Bildfolge $(f(x_n))$ bestimmt divergent mit dem Grenzwert $+\infty$.

2. Die Interpretation von
   $\lim\limits_{x\to a} f(x) = -\infty$
   ist völlig analog.

3. Die Interpretation von
   $\lim\limits_{\substack{x\to a\\ x>a}} f(x) = +\infty$ bzw. $\lim\limits_{\substack{x\to a\\ x<a}} f(x) = +\infty$
   bzw. von
   $\lim\limits_{\substack{x\to a\\ x>a}} f(x) = \infty$ und $\lim\limits_{\substack{x\to a\\ x<a}} f(x) = -\infty$
   seien ebenfalls dem Leser/ der Leserin überlassen.

Schließlich bedeutet

$\lim\limits_{x\to\infty} f(x) = +\infty$ die Aussage: Der Definitionsbereich von f ist nicht nach oben beschränkt und für jede bestimmt gegen $+\infty$ divergente Folge $(x_n),\ x_n \in D$, ist die Bildfolge $(f(x))$ ebenfalls bestimmt divergent mit dem uneigentlichen Grenzwert $+\infty$. Weitere Fälle sind möglich.

In den einführenden Beispielen gilt etwa

$\lim\limits_{\substack{x\to 0\\ x>0}} f(x) = +\infty,\ \lim\limits_{\substack{x\to 0\\ x<0}} f(x) = -\infty$
$\lim\limits_{x\to 1} g(x) = \infty$ und
$\lim\limits_{x\to+\infty} h(x) = +\infty$.

Häufig nützlich ist die folgende

Bemerkung: Ist $f : D \to \mathbb{R}$ auf dem nach oben nicht beschränkten Definitionsbereich D monoton wachsend und nach oben nicht beschränkt, dann ist

$$\lim_{x\to\infty} f(x) = +\infty\,.$$

Beweis als Übungsaufgabe.

Nach der Bemerkung gilt etwa

$$\lim_{x\to\infty} \exp(x) = +\infty,\quad \lim_{x\to\infty} \log(x) = \infty,\quad \lim_{x\to\infty} \sqrt{x} = \infty,\quad \lim_{x\to\infty} x^k = +\infty,\quad \text{für jedes } k \in \mathbb{N}\,.$$

Wir beschließen diesen Abschnitt mit folgendem Satz über das Wachstum der reellen Experimentalfunktion im Vergleich zu allgemeinen Exponentialfunktion $p_k(x) = x^k \ (k \subset \mathbb{N})$.

14.3.5 Satz vom Wachstum (der reellen Exponentialfunktion):

Es gilt für jedes $k \in \mathbb{N}$

$$\boxed{\lim_{x\to\infty} \frac{x^k}{\exp(x)} = 0 \text{ oder äquivalent } \lim_{x\to\infty} \frac{\exp(x)}{x^k} = \infty\,.}$$

Man sagt: Die reelle Exponentialfunktion wächst beim Grenzübergang $x \to +\infty$ schneller gegen $+\infty$ als jede positive Potenz.

Beweis des Satzes: Aus

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!} + \ldots$$

folgt für $x > 0$

$$\exp(x) > \frac{x^{n+1}}{(n+1)!} \quad \text{und damit} \quad \frac{\exp(x)}{x^n} > \frac{x}{(n+1)!} \ ,$$

woraus die Behauptung unmittelbar folgt. Das schnelle (starke) Wachstum der Exponentialfunktion überträgt sich in ein (extrem) langsames Wachstum der Umkehrfunktion $\log$.

14.3.6 Satz vom Wachstum (des natürlichen Logarithmus):

$$\begin{array}{rcl} \log : \mathbb{R}_+^\bullet & \to & \mathbb{R} \\ x & \mapsto & \log x \end{array}$$

wächst für $x \to \infty$ *schwächer* als jede n-te Wurzel, d. h. für jedes $n \in \mathbb{N}$ gilt

$$\boxed{\lim_{x\to\infty} \frac{\log x}{\sqrt[n]{x}} = 0 \, .}$$

Durch die Ersetzung $x := e^{ny}$ reduziert man die Aussage auf den entsprechenden Satz über das Wachstum von $\exp$.

Wie schon bemerkt, baut die Differentialrechnung auf dem Grenzwertbegriff auf. Die Differentialrechnung liefert auch brauchbare Mittel, u. U. auch komplizierte Grenzwerte zu berechnen, wie wir später noch sehen werden.

15 Funktionenfolgen, Funktionenreihen, gleichmäßige Kovergenz, Potenzreihen

Beispiele für Funktionenfolgen und Funktionenreihen sind uns in den vorangehenden Abschnitten schon häufig begegnet. So hatten wir z.B. für $z \in \mathbb{C}$ die komplexe Exponentialfunktion durch

$$\exp(z) := \lim_{n\to\infty} E_n(z) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{z^k}{k!} =: \sum_{k=0}^{\infty} \frac{z^k}{k!} \; (z \in \mathbb{C})$$

definiert.
Für festes $z \in \mathbb{C}$ konvergiert also die komplexe Zahlenfolge $(E_n(z))$ gegen $\exp(z)$. Für festes $z \in \mathbb{C}$ konvergiert auch die komplexe Zahlenfolge $f_n)(z) = \big((1+\frac{z}{n})^n\big)$ gegen $\exp(z)$. Für festes $n \in \mathbb{N}_0$, bzw. $n \in \mathbb{N} \mapsto E_n(z)$ bzw. $z \mapsto f_n(z)$ Funktionen auf $\mathbb{C}$. Ist D irgend eine nichtleere Masse (konkret etwa eine nichtleere Teilmenge von $\mathbb{R}$ oder $\mathbb{C}\backslash\mathbb{K} = \mathbb{C}$) und ist für jedes $n \in \mathbb{N}$ (bzw. $\mathbb{N}_0$) eine Funktion $f_n : D \to \mathbb{K}$ $(\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C})$ gegeben, so spricht man von einer auf D erklärten *Funktionenfolge*. Eine auf D erklärte Funktionenfolge kann man als Element der Menge Abb $(\mathbb{N}, \mathbb{K})$ bzw. Abb $(\mathbb{N}_{\nu}, \mathbb{K})$ auffassen.
Kurzschreibweise $n \mapsto f_n$ oder (f_n).
Wählt man $z \in D$ fest, dann ist $(f_n(z))$ eine Zahlenfolge in $\mathbb{K}(\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C})$, die aus (f_n) durch Einsetzen von z entsteht.
Bei der Folge $(f_n(z)$ kann man nun fragen, ob sie konvergiert oder nicht. Wenn $(f_n(z)$ für alle $z \in D$ konvergiert, dann wird durch $f(x) = \lim\limits_{n\to\infty} f_n(z)$ eine Funktion $f : D \to \mathbb{K}$ definiert. Die Funktion f ist dann eindeutig bestimmt und man sagt, dass (f_n) *punktweise* gegen f konvergiert.
So konvergiert etwa die Folge (E_n) mit $E_n(z) = \sum\limits_{k=0}^{n} \frac{z^k}{k!}$ punktweise gegen die Exponentialfunktion.
Eine naheliegende Frage ist nun, ob sich fundamentale Eigenschaften der f_n wie

- die Stetigkeit
- die Differenzierbarkeit
- die Integrierbarkeit

auf die Grenzfunktion übertragen.
Wir werden sehen, dass diese "guten"Eigenschaften der f_n sich i.a. nicht auf die Grenzwertfunktion f übertragen. Wir benötigen einen stärkeren Konvergenzbegriff, den Begriff der *gleichmäßigen Konvergenz*, der einfache Aussagen über die *Vertauschbarkeit von Grenzprozessen* ermöglicht.
Unser Aufbau der Integralrechnung wird sich wesentlich auf dem Begriff der gleichmäßigen Konvergenz stützen. Wir werden sehen, dass sich der Begriff der gleichmäßigen Konvergenz von Funktionenfolgen mit Hilfe einer *Norm*, der sogenannten *Supremumsnorm* übersichtlich behandeln lässt. Diese ist für *beschränkte* Funktionen $f : D \to \mathbb{K}$ durch

$$\|f\| := \sup\{|f(z)|;\ z \in D\}$$

definiert und hat ähnliche Eigenschaften wieder (Absolut-) Betrag einer reellen oder komplexen Zahl.

15.1 Punktweise Konvergenz von Funktionenfolgen

Wir betrachten als einführendes Beispiel die Funktionenfolge (f_n) mit $D = [0, 1]$ und $f_n(x) = x^n$. Für jedes x mit $0 \le x < 1$ ist bekanntlich (x^n) eine Nullfolge und damit gilt $\lim\limits_{n\to\infty} f_n(x) = 0$ für $0 \le x < 1$. Für $x = 1$ ist jedoch $f_n(x) = 1$ und damit auch $\lim\limits_{n\to\infty} f_n(x) = 1$.
Also konvergiert die Folge $(f_n(x))$ gegen 0 falls $0 < x < 1$ und gegen 1 falls $x = 1$, h.h. die Grenzwertfunktion f ist gegeben durch $f(x) = \begin{cases} 0, & \text{falls } 0 \le x < 1 \\ 1, & \text{falls } x = 1\,. \end{cases}$

Obwohl alle (f_n) stetig sind (Polynome!), ist die Grenzfunktion unstetig an der Stelle 1. Man könnte an dieser Stelle schon den Grund für dieses Phänomen analysieren, wir kommen später darauf zurück. Zunächst geben wir eine exakte Definition für die punktweise Konvergenz einer Funktionenfolge.

Definition: Eine Folge (f_n) von Funktionen $f_n : D \to \mathbb{K}$ konvergiert punktweise auf D, wenn r eine Funktion $f : D \to \mathbb{K}$, so dass für jede $z \in D$ die Zahlenfolge $(f_n(z))$ gegen $f(z)$ konvergiert. f heißt die Grenzfunktion der $f_n,\ f$ ist dan eindeutig bestimmt (warum?).
Wenn man keinen Kandidaten für die Grenzfunktion f findet oder raten kann, hilft das folgende Cauchy-Kriterium weiter, in welchen die Grenzfunktion nicht vorkommt.

Satz: Notwendig und hinreichend dafür, dass die Funktionenfolge (f_n) punktweise (gegen f konvergiert ist die folgende Cauchy-Bedingung: Zu jedem $\varepsilon > 0$ und jedem $z \in D$ gibt ρ ein $\mathbb{N} \in \mathbb{N}$, so

dass für alle $m, n \geq N$ gibt

$$(*) \qquad\qquad |f_n(z) - f_m(z)| < \varepsilon\,.$$

Wenn (f_n) punktweise gegen f konvergiert, so gibt es für beliebiges $\varepsilon >=$ und beliebiges, aber festes $z \in D$ ein $\mathbb{N} \in \mathbb{N}$, so dass für alle $n \geq \mathbb{N}$ gilt

$$|f_m(z) - f(z)| < \frac{\varepsilon}{2}\,.$$

Ist auch $m \geq \mathbb{N}$ so gilt entsprechen

16 Die elementaren Funktionen

IV. Grundlagen der Integral- und Differentialrechnung

18 Grundlagen der Integralrechnung

Sowohl die *Integralrechnung* als auch die *Differentialrechnung* gehören zum Kernbestand der Analysis. Beider gehen ursprünglich von geometrischen Fragestellungen aus:
Bei der *Integralrechnung* will man *Flächeninhalte* oder allgemeiner *Volumina* von geometrischen Objekten bestimmen, aber auch *Längen von Kurven*, Dichten, Schwerpunkte, Mittelwerte oder auch Arbeit (Energie) in einem nicht konstanten Kraftfeld berechnen. Auch die Frage, wie man aus einer *Änderungsrate einer mathematischen oder physikalischen Größe* auf diese Größe selber zurückschließen kann, lässt sich mit Hilfe der Integralrechnung beantworten. Ein Fahrtenschreiber z.B. zeichnet die Geschwindigkeit eines Fahrzeugs auf, mit Hilfe der Integralrechnung lässt sich die Fahrtstrecke rekonstruieren.
Das Prinzip der *Trägheitsnavigation* beruht auf dem Messen der Beschleunigung, der daraus rekonstruierten Geschwindigkeit und dem wiederum hieraus rekonstruieren zurückgelegten Weg.
Geometrische Aspekte bei der Differentialrechnung sind etwa das *Tangentenproblem* (Tangente an eine Kurve), physikalische etwa die *momentane Änderung* einer Größe etwa die Momentangeschwindigkeit oder die Momentanbeschleunigung. Mathematisch handelt es sich hierbei jeweils um das gleiche Problem. Die Differentialrechnung ist auch an sich starkes Hilfsmittel zur Untersuchung qualitativer Eigenschaften von Funktionen (Abbildungen).
Für Funktionen auf Intervallen werden wir sehen, dass die Differential- und Integralrechnung, die zunächst nichts untereinander zu tun zu haben scheinen, über den sogenannten Hauptsatz der *Differential- und Integralrechnung* untereinander schon eng verbandelt sind.
Bei Archimedes (†212 v.Chr.) findet man schon als Vorstufe der Integration die Bestimmung von Flächen- und Rauminhalten, den Zusammenhang zwischen Differential- und Integralrechnung haben Leibniz und Newton um ca. 1670 entdeckt.

Eine systematische Fassung des Flächen- und des Volumenbegriffs für allgemeine geometrische Objekte beginnt mit J.Kepler (1615, Volumenbestimmung von Weinfässern), wichtige Beiträge stammen von dem Galilei-Schüler B.Cavalieri (1635, sog. Cavalierisches Prinzip), dann Newton's Lehrer I.Barrow und viele anderen. Strenge systematische Behandlungen solcher Probleme stehen im engen Zusammenhang mit der Entwicklung der Maß- und Integrationstheorie am Anfang des 20. Jahrhunderts.

Es gibt verschiedene Integralbegriffe und die Mengen, denen man in vernünftiger Weise mit Hilfe des Integrals eine Zahl zuordnen kann, die man als Volumen der Menge ansprechen kann, hängen vom verwendeten Integralbegriff ab. Wir beschränken uns hier auf das sog. *Regel-Integral* (auch Cauchy-Integral genannt, Cauchy 1823).
Das sog. *Riemann-Integral* (B.Riemann) streifen wir nur am Rande.
In *mehreren Veränderlichen* betrachten wir später das von H.Lebesgue 1902 eingeführte *Lebesgue -Integral*, das allen anderen Integralbegriffen weit überlegen ist.
Unser Ziel ist zunächst für eine einfache Klasse von Funktionen -die *Treppenfunktionen*- ein Integral zu definieren und dieses dann auf eine größere Klasse von Funktionen -die *Regelfunktionen*- zu erweitern.

18.1 Das Integral für Treppenfunktionen

Im Folgenden legen wir bis auf Widerruf ein kompaktes Intervall

$$M := [a, b] := \{x \in \mathbb{R};\ a \le x \le b\} \qquad (a, b \in \mathbb{R},\ a < b)$$

zu Grunde.

Schränkt man die Größte-Ganze-Funktion

$$\begin{aligned} [\,] : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto [x] = \max\{k \in \mathbb{Z};\ k \leq x\} \end{aligned}$$

auf ein kompaktes Intervall, etwa $[-1, 4]$ ein, erhält man ein typisches Beispiel für eine Treppenfunktion im Sinne der folgenden allgemeinen Definition (vgl. Abb 10).

18.1.1 Definition (Treppenfunktion, Zerlegung)

Eine Funktion

$$t : \ M \rightarrow \mathbb{R}$$

heißt *Treppenfunktion*, wenn es ein $r \in \mathbb{N}$ und reelle Zahlen $x_0, x_1, \ldots, x_r$ mit

$$a = x_0 < x_1 < x_2 < \ldots < x_r = b$$

gibt, so dass die Einschränkungen

$$f|]x_{k-1}, x_k[\ \text{ für } \ k = 1, \ldots, r$$

konstant, etwa $= c_k \in \mathbb{R}$, sind.
Man sagt: Die endliche Menge

$$\{a = x_0, x_1, \ldots, x_r = b\}$$

bildet eine *Zerlegung* Z des Intervalls $[a, b]$ und schreibt suggestiv

$$Z := \{a = x_0 < x_1 < \ldots < x_r = b\}.$$

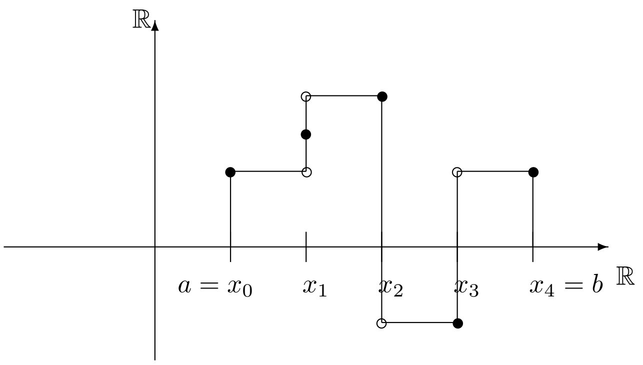
Die Punkte x_k $(0 \leq k \leq r)$ nennt man auch *Teilpunkte* (oder *Stützstellen*) der Zerlegung Z.
Man nennt dann t auch eine Treppenfunktion zur Zerlegung Z.

18.1.2 Bemerkungen

(a) Da t eine Funktion ist, hat t natürlich auch eindeutig bestimmte Werte in den Teilpunkten

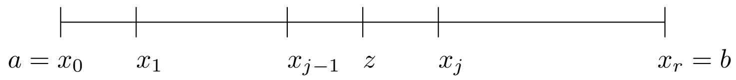
$$x_k \ (0 \leq k \leq r).$$

Diese sind für unsere Integraldefinition allerdings unerheblich.



(b) Bei gegebener Treppenfunktion $t: \ [a,b] \to \mathbb{R}$ ist die Zerlegung Z nicht eindeutig bestimmt, so kann man beispielsweise zu einer gegebenen Zerlegung Z stets Teilpunkte hinzufügen. Ist etwa z ein Punkt mit $x_{j-1} < z < x_j$ für ein geeignetes $j \in \{1, 2, \ldots, r\}$, dann erhält man aus Z durch das Hinzufügen von z die Zerlegung

$$Z' := \{a = x_0 < x_1 < \ldots < x_{j-1} < z < x_j < \ldots < x_r = b\}$$



(c) Wir nennen eine Zerlegung

$$Z^* = \{a = x_0^* < x_1^* < \ldots < x_s^* = b\} \qquad (s \in \mathbb{N})$$

feiner als die Zerlegung

$$Z = \{a = x_0 < x_1 < \ldots < x_{j-1} < z < x_j < \ldots < x_r = b\} \qquad (r \in \mathbb{N})$$

wenn $Z^* \supset Z$ gilt, wenn also jedes x_j $(0 \leq j \leq r)$ unter den x_l^* $(0 \leq l \leq s)$ vorkommt.

(d) Zu je zwei Zerlegungen Z und Z' von $[a,b]$ gibt es stets eine Zerlegung Z^*, die feiner ist als Z und auch feiner als Z'. Offensichtlich ist die Zerlegung $Z^* := Z \vee Z'$, die aus Z durch Hinzunahme der (nicht in Z gelegenen) Punkten von Z' entsteht, eine *gemeinsame Verfeinerung* von Z und Z'.

18.1.3 Definition

Ist $t: \ M \to \mathbb{R}$ eine Treppenfunktion, die bezüglich der Zerlegung

$$Z := \{a = x_0 < x_1 < \ldots < x_{j-1} < z < x_j < \ldots < x_r = b\}$$

definiert ist und gilt etwa $f| \,]x_{k-1}, x_k[= c_k$ für $k = 1, \ldots, n$, dann heißt die reelle Zahl

$$I_Z(t) := \sum_{k=1}^{r} c_k(x_k - x_{k-1})$$

das *Integral von t bezüglich der Zerlegung Z*.

Wir werden gleich sehen, dass dieses Integral in der Wirklichkeit nicht von der Zerlegung Z abhängt, mit deren Hilfe t erklärt ist.

18.1.4 Satz

Ist $t: \ M \to \mathbb{R}$ eine Treppenfunktion bezüglich der Zerlegung Z und Z', dann gilt

$$I_Z(t) = I_{Z'}(t).$$

Beweis : Wir betrachten zunächst die Zerlegung Z_1, die aus Z durch Hinzufügen eines weiteren Teilpunktes z, etwa Intervall $]x_{j-1}, x_j[$, entsteht. z sei also ein neuer Teilpunkt mit $x_{j-1} < z < x_j$ für ein $j \in \{1, 2, \dots, n\}$.
Dann ist

$$I_{Z_1}(t) = \sum_{k=1}^{j-1} c_k(x_k - x_{k-1}) + c_j(z - x_{j-1}) + c_j(x_j - z) + \sum_{k=j+1}^{r} c_k(x_k - x_{k-1})$$

Nun ist aber

$$c_j(z - x_{j-1}) + c_j(x_j - z) = c_j z - c_j x_{j-1} + c_j x_j - c_j z = c_j(x_j - x_{j-1}).$$

Daher ist

$$\begin{aligned} I_{Z_1}(t) &= \sum_{k=1}^{j-1} c_k(x_k - x_{k-1}) + c_j(x_j - x_{j-1}) + \sum_{k=j+1}^{r} c_k(x_k - x_{k-1}) \\ &= \sum_{k=1}^{r} c_k(x_k - x_{k-1}) = I_Z(t). \end{aligned}$$

Beim Hinzufügen *eines* weiteren Teilpunktes ändert sich also $I_Z(t)$ nicht. Durch Induktion folgt etwa, dass sich $I_Z(t)$ durch Hinzufügen endlich vieler weiteren Teilpunkte nicht ändert.

Wir betrachten nur zwei Zerlegungen Z und Z' die *gemeinsame Verfeinerung* $Z^* := Z \vee Z'$. Sie entsteht aus Z bzw. Z' durch Hinzunahme endlich vieler weiteren Teilpunkte, daher ist

$$I_Z(t) = I_{Z \vee Z'}(t) = I_{Z'}(t),$$

d.h. also $I_Z(t) = I_{Z'}(t)$.

□

18.1.5 Satz und Definition

Ist $t: \ [a, b] \to \mathbb{R}$ eine Treppenfunktion und

$$Z = \{a = x_0 < x_1 < \ldots < x_{j-1} < z < x_j < \ldots < x_r = b\}$$

irgendeine Zerlegung von $[a, b]$ mit $t| \ ]x_{k-1}, x_k[= c_k \in \mathbb{R}$ für $k = 1, \dots, n$, dann heißt die von der Zerlegung unabhängige Zahl

$$I(t) := I_Z(t) := \sum_{k=1}^{r} c_k(x_k - x_{k-1})$$

das *Integral der Treppenfunktion* über das Intervall $M = [a, b]$.

Schreibweise: $I(t) := \int\limits_a^b t = \int\limits_a^b t(x)dx.$

x heißt auch *Integrationsvariable*. Sie ist eine sogenannte freie Variable, kann also mit jedem anderen sinnvollen Buchstaben bezeichnet werden.

$\boxed{\text{Das Integral einer Treppenfunktion ist eine endliche Summe von Produkten.}}$

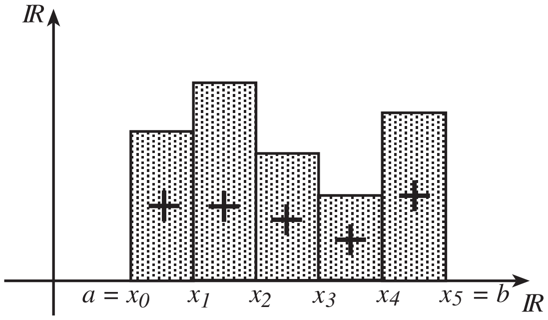
18.1.6 Geometrische Interpretation

Nimmt die Treppenfunktion t nur nicht negative Werte an, dann kann man

$$I(t) = \sum_{k=1}^{r} c_k(x_k - x_{k-1})$$

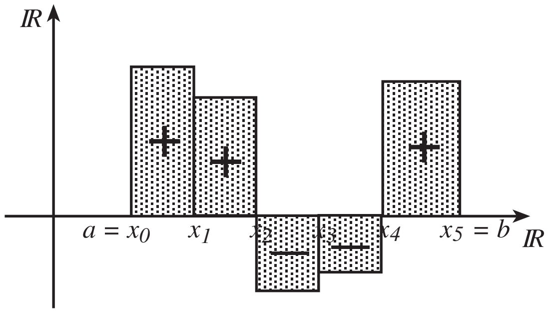
als Flächeninhalt der Punktmenge (Ordinatenmenge)

$$O(t) := \{(x, y) \in \mathbb{R} \times \mathbb{R};\ x \in [a, b];\ 0 \le y \le t(x)\}$$



deuten: $O(t)$ ist eine Vereinigung von Rechtecken und $I(t)$ liefert den elementargeometrischen Flächeninhalt dieser Rechtecke.

Nimmt t auch negative Werte an, dann ist $I(t)$ ein Maß für den *orientierten* Flächeninhalt (d.h. Flächeninhalt mit Vorzeichen) der entsprechenden Rechtecke.



Bezeichnung: Sei $T(M) = \{t : M \to \mathbb{R};\ t\ \text{Treppenfunktion}\}$ die Menge aller Treppenfunktionen auf $M = [a, b]$, dann gilt

18.1.7 Satz (Eigenschaften von $T(M)$ und $I : M \to \mathbb{R}$)

(a) $T(M)$ ist ein Unterraum von $B(M) := \{f : M \to \mathbb{R};\ f\ \text{beschränkt}\}$ und damit selbst ein $\mathbb{R}$-Vektorraum und die Abbildung

$$\begin{array}{rcl} I:\ T(M) & \to & \mathbb{R} \\ t & \mapsto & I(t) \end{array}$$

ist linear, d.h. es gilt

(α) $I(t_1 + t_2) = I(t_1) + I(t_2)$ für alle $t_1, t_2 \in T(M)$

(β) $I(ct) = cI(t)$ für alle $c \in \mathbb{R}$ und alle $t \in T(M)$.

Das Integral für Treppenfunktionen ist also ein *lineares Funktional* auf dem Vektorraum $T(M)$.

(b) Für alle $t \in T(M)$ gilt

$$|I(t)| \leq \underbrace{\sup\{|t(x)|;\ x \in M\}}_{\|t\|_\infty}(b-a) = \|t\|_\infty(b-a) \qquad \text{(Standardabschätzung)}$$

(c) Ist $t(x) \geq 0$ für alle $x \in M$ (kurz $t \geq 0$), dann gilt auch $I(t) \geq 0$.
$I : T(M) \to \mathbb{R}$ ist ein nicht negatives lineares Funktional.

(c') Gilt für $t_1, t_2 \in T(M)$, $t_1(x) \leq t_2(x)$ für alle $x \in M$ (kurz: $t_1 \leq t_2$), dann gilt auch

$$I(t_1) \leq I(t_2).$$

Man sagt: I ist ein monotones Funktional.

(d) Mit $t \in T(M)$ gilt auch $|t| \in T(M)$, und es ist

$$|I(t)| \leq I(|t|).$$

Beweis (a): Seien $t_1, t_2 \in T(M)$. Wir können oBdA annehmen, dass t_1 und t_2 bezüglich der gleichen Zerlegung Z von $[a, b]$ definiert sind, dann ist aber klar, dass $t_1 + t_2$ wieder Treppenfunktion ist, die Funktionswerte in den Teilintervallen $]x_{k-1}, x_k[$ sind einfach zu addieren. Sei etwa $t_1|\,]x_{k-1}, x_k[)c_k$ und $t_2|\,]x_{k-1}, x_k[= d_k$, dann gilt

$$\begin{aligned}
I(t_1 + t_2) &= I_Z(t_1 + t_2) \\
&= \sum_{k=1}^{r}(c_k + d_k)(x_k - x_{k-1}) \\
&= \sum_{k=1}^{r} c_k(x_k - x_{k-1}) + \sum_{k=1}^{r} d_k(x_k - x_{k-1}) \\
&= I_Z(t_1) + I_Z(t_2) = I(t_1) + I(t_2).
\end{aligned}$$

□

Dass mit $t \in T(M)$ auch $ct \in T(M)$ ist evident, ebenso die Regel $I(ct) = cI(t)$

Beweis (b): Es gilt

$$\begin{aligned}
|I(t)| &= \left|\sum_{k=1}^{r} c_k(x_k - x_{k-1})\right| \\
&\leq \sum_{k=1}^{r} |c_k|(x_k - x_{k-1}) \qquad (Dreiecksungleichung) \\
&\leq \sum_{k=1}^{r} \underbrace{\sup\{|t(x)|;\ x \in M\}}_{\|t\|_\infty}(x_k - x_{k-1}) \\
&= \|t\|_\infty(b-a),
\end{aligned}$$

da $\sum\limits_{k=1}^{r}(x_k - x_{k-1})$ als endliche teleskopische Summe den Wert $b - a$ hat:

$$\begin{aligned}
\sum_{k=1}^{r}(x_k - x_{k-1}) &= (x_1 - x_0) + (x_2 - x_1) + \ldots + (x_r - x_{r-1}) \\
&= -x_0 + x_r = x_r - x_0 = b - a
\end{aligned}$$

□

Beweis (c): Ist $t(x) \geq 0$, dann gilt speziell $c_k \geq 0$ für alle $k = 1, 2, \ldots, n$ und es folgt

$$I(t) = \sum_{k=1}^{r} c_k \underbrace{(x_k - x_{k-1})}_{>0} \geq 0.$$

□

Beweis (c'): $t_1 \leq t_2 \iff t_2 - t_1 \geq 0$. Aus der Linearität von I und (c) folgt dann

$$0 \leq I(t_2 - t_1) = I(t_2) - I(t_1) \text{ oder } I(t_1) \leq I(t_2).$$

□

Beweis (d): Aus $t \leq |t|$ folgt nach (c') daher $I(t) \leq I(|t|)$. Weil auch $-t \leq |t|$ gilt, folgt mit (a) und (c')

$$-I(t) = I(-t) \leq I(|t|),$$

also insgesamt

$$|I(t)| \leq I(|t|).$$

□

Wir betrachten ein Beispiel:

18.1.8 Beispiel

Wir betrachten die Treppenfunktion

$$t : [0, 1] \to \mathbb{R} \text{ mit}$$

$$t(x) := \frac{[nx]}{n},$$

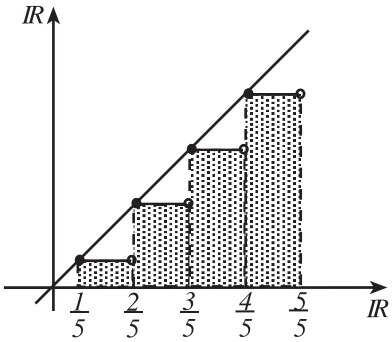
dabei ist $n \in \mathbb{N}$ fest.
Eine zu t passende Zerlegung von $[0, 1]$ ist

$$z := 0 = x_0 < \frac{1}{n} < \frac{2}{n} < \ldots < \frac{k}{n} < \ldots < \frac{n-1}{n} < \frac{n}{n} = 1 = x_n.$$

Es ist also $x_k = \frac{k}{n}$ und $x_k - x_{k-1} = \frac{1}{n}$ und $c_k := t|\,]x_k - x_{k-1}[= \frac{k-1}{n}$ für $k = 1, \ldots, n$.
Das Integral $I(t)$ ist leicht zu berechnen:

$$\begin{aligned} I(t) = \sum_{k=1}^{r} c_k (x_k - x_{k-1}) &= \sum_{k=1}^{r} \frac{k-1}{n} \cdot \frac{1}{n} = \frac{1}{n^2} \sum_{k=1}^{r} (k-1) \\ &= \frac{1}{n^2} \frac{n(n-1)}{2} \quad \text{(nach Gauß)} \\ &= \frac{1}{2}\left(1 - \frac{1}{n}\right). \end{aligned}$$

Veranschaulichung für $n = 5$



Lässt man n in alle natürlichen Zahlen durchlaufen und schreibt $t_n(x) = \frac{[nx]}{n}$, dann ist (t_n) eine Folge von Treppenfunktionen, die auf $[0,1]$ gleichmäßig gegen f mit $f(x) = x$ für $x \in [0,1]$ konvergiert.
Beachte dabei auch $\lim\limits_{n\to\infty} \frac{1}{2}\left(1 - \frac{1}{n}\right) = \frac{1}{2}$.

Wir kommen in Abschnitt 17.2 hierauf zurück.

Man kommt zu einer kleinen Erweiterung der Integralfunktion, in der man

$$\int\limits_a^a t = 0$$

definiert und beachtet, dass für eine Treppenfunktion $t : [0,1] \to \mathbb{R}$ für jedes c mit $a < c < b$

$$t|\,[a,c] \text{ und } t|\,[c,b]$$

wieder Treppenfunktionen sind, die wir der Einfachheit halber auch wieder mit t bezeichnen.

18.1.9 Satz

Für jedes $t \in T(M)$ und für jedes $c \in [a,b]$ gilt

$$\boxed{\int\limits_a^c t(x)dx + \int\limits_c^b t(x)dx = \int\limits_a^b t(x)dx}$$

Diese Formel bringt eine *Intervalladditivität* des Integrals I zum Ausdruck.

Zum **Beweis** braucht man nur c unter die Teilpunkte der Zerlegung aufzunehmen, mit deren Hilfe man $I(t)$ berechnet.

18.1.10 Bemerkung

Zum Abschluss sei bemerkt, dass mit $t_1, t_2 \in T(M)$ auch $t_1 t_2 \in T(M)$ gilt, das Produkt von zwei Treppenfunktionen ist also wieder eine Treppenfunktion, aber wie einfache Beispiele zeigen (vgl. Übungsblatt 1, Aufgabe 4) gilt i.A. nicht $I(t_1 t_2) = I(t_1)I(t_2)$.

Es gilt jedoch die *Cauchy-Schwarzsche Ungleichung* in der Form

$$\boxed{(I(t_1 t_2))^2 \leq I(t_1^2)I(t_2^2)}$$

Vergleiche hierzu auch die Musterlösung von Aufgabe 4 von Blatt 1.

18.2 Das Integral für Regelfunktionen

Ziel diese Abschnitts ist es, die Integraldefinition, die sich bis jetzt nur auf Treppenfunktionen erstreckt hat, auf eine größere Klasse von Funktionen, die sog. *Regelfunktionen* so auszudehnen, dass die Eigenschaften des Integrals, d.h. der Abbildung

$$\begin{aligned} I:\ T(M) &\to \mathbb{R} \qquad (M := [a,b]) \\ t &\mapsto I(t) = \int_a^b t(x)dx \end{aligned}$$

erhalten bleiben (I ist ein nicht negatives, beschränktes lineares Funktional). Wir werden sehen, dass stetige bzw. monotone Funktionen auf M Regelfunktionen sind. Da viele der uns geläufigen Funktionen (Polynome höheren Grades, rationale Funktionen, sin, cos, exp, log) keine Treppenfunktionen sind, werden wir versuchen, sie in geeigneter Weise durch Treppenfunktionen zu approximieren.

18.2.1 Definition

Sei $M = [a,b]$ $(a, b \in \mathbb{R},\ a < b)$ ein kompaktes Intervall in $\mathbb{R}$.
Eine Funktion $f : M \to \mathbb{R}$ heißt *Regelfunktion*,wenn es eine Folge (t_n) von Treppenfunktionen $t_n \in T(M)$ gibt, die *gleichmäßig* gegen f konvergiert.

Beachte:
Gleichmäßige Konvergenz ist die Konvergenz bezüglich der Supremumsnorm:
Das die Folge (t_n) gleichmäßig gegen f konvergiert bedeutet also, dass die (reelle) Zahlenfolge $(\|t_n - f\|)$ der Normen eine Nullfolge ist.
Die Supremumsnorm einer beschränkten Funktion $f : M \to \mathbb{R}$ ist dabei die nicht negative reelle Zahl

$$\|f\|_\infty := \|f\|_M = \sup\{|f(x)|;\ x \in M\}$$

Da der Definitionsbereich M zunächst festgehalten wird, lassen wir den Index M auch der Einfachheit halber weg und schreiben nur $\|f\|$.
Die *Supremumsnorm* ist eine *Norm* auf dem Vektorraum $B(M) = \{f : M \to \mathbb{R};\ f$ beschränkt$\}$ der beschränkten Funktionen auf M im Sinne von Def. ???.
Für die gleichmäßige Konvergenz der Folge (t_n) gegen f verwenden wir die Kurzschreibweise

$$t_n \rightrightarrows f.$$

Gilt nun $t_n \rightrightarrows f$, d.h. konvergiert (t_n) gleichmäßig gegen f, dann liegt es nahe für die Funktion f ein Integral durch

$$I(f) := \lim_{n\to\infty} I(t_n)$$

zu erklären.
Damit dies sinnvoll ist, ist zweierlei zu zeigen:

(a) Die Folge $(I(t_n))$ der Integrale der Treppenfunktionen t_n ist überhaupt konvergent.

(b) Ist $(\tilde{t}_n)$ eine weitere Folge von Regelfunktionen $\tilde{t}_n \in T(M)$ mit $\tilde{t}_n \rightrightarrows f$, dann gilt

$$\lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} I(\tilde{t}_n).$$

Dass beide Aussagen richtig sind, ist die Aussage von

18.2.2 Lemma

Sei $f : M \to \mathbb{R}$ eine Regelfunktion.

(a) Ist $t_n \in T(M)$ für $n \in \mathbb{N}$ und gilt $t_n \rightrightarrows f$, dann ist die Folge $(I(t_n))$ der Integrale eine konvergente reelle Zahlenfolge.

(b) Gilt auch $\tilde{t}_n \rightrightarrows f$ mit $\tilde{t}_n \in T(M)$, dann ist

$$\lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} I(\tilde{t}_n)$$

Beweis (a): Da wir den potenziellen Grenzwert nicht kennen, verwenden wir das Cauchy-Kriterium, um die Konvergenz der Folge $(I(t_n))$ zu zeigen.
Da (t_n) gleichmäßig gegen f konvergiert, gibt es zu jedem $\varepsilon > 0$ einen Index $N \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt

$$\|t_n - f\| < \frac{\varepsilon}{2(b-a)}$$

Für alle $m \in \mathbb{N}$ mit $m \geq N$ gilt dann auch

$$\|t_m - f\| < \frac{\varepsilon}{2(b-a)}$$

Daher ist

$$\begin{aligned}
\|t_m - t_n\| &= \|t_m - f + f - t_n\| \\
&\leq \|t_m - f\| + \|f - t_n\| \\
&= \|t_m - f\| + \|t_n - f\| < \frac{\varepsilon}{2(b-a)} + \frac{\varepsilon}{2(b-a)} = \frac{\varepsilon}{b-a}
\end{aligned}$$

für alle $m, n \geq N$.
Nach dieser Vorbereitung betrachten wir die Folge $(I(t_n))$ der Integrale und zeigen, dass sie eine Cauchy-Folge reeller Zahlen, also konvergent ist.
Dazu nutzen wir die Linearität des Integrals für Treppenfunktionen und die Standardabschätzung aus:

$$\begin{aligned}
|I(t_m) - I(t_n)| \overset{\text{Linearität}}{=} |I(t_m - t_n)| &\leq (b-a)\|t_m - t_n\| \\
&< (b-a)\frac{\varepsilon}{b-a} = \varepsilon
\end{aligned}$$

für alle m, n mit $m, n \geq N$. Die Folge $(I(t_n))$ ist daher als Cauchy-Folge reeller Zahlen konvergent.

□

Beweis (b): Für (b) geben wir zwei Beweise. Da wir nach (a) schon wissen, dass die Folgen $(I(t_n))$ und $(I(\tilde{t}_n))$ konvergent sind, genügt es zu zeigen, dass die Folge der Differenzen eine Nullfolge ist:
Aus

$$\|t_n - \tilde{t}_n\| \leq \|t_n - f\| + \|\tilde{t}_n - f\|$$

ist ersichtlich, dass die Folge $(\|t_n - \tilde{t}_n\|)$ eine Nullfolge ist.
Nun gilt

$$|I(t_n) - I(\tilde{t}_n)| \overset{\text{Linearität}}{=} |I(t_n - \tilde{t}_n)| \overset{\text{Standardabschätzung}}{\leq} (b-a)\|t_n - \tilde{t}_n\|$$

Weil die rechte Seite eine Nullfolge ist, ist auch die linke Seite eine Nullfolge und nach den Rechenregeln für konvergente reelle Zahlenfolgen folgt

$$\lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} I(\tilde{t}_n)$$

□

Ein zweiter Beweis beruht auf dem Prinzip der *Folgenmischung*, das wir schon mehrfach verwendet haben:
Hier bedeutet es:
falls $t_n \rightrightarrows f$ und $\tilde{t}_n \rightrightarrows f$, dann gilt auch, dass die durch Mischung entstehende „Zickzackfolge" („Reißverschlussfolge")

$$(t_n^*) = (t_1, \tilde{t}_1, t_2, \tilde{t}_2, t_3, \tilde{t}_3, \ldots)$$

gleichmäßig gegen f konvergiert.
Zum **Beweis** beachte man

$$t_n^* = \begin{cases} t_{\frac{n+1}{2}}, & \text{falls } n = 2m-1,\ m \in \mathbb{N}; \\ \tilde{t}_{\frac{n}{2}}, & \text{falls } n = 2m,\ m \in \mathbb{N}. \end{cases}$$

Nach (a) folgt dann aber, dass die Folge $I(t_n^*))$ der Integrale konvergiert, sei etwa $A := \lim\limits_{n\to\infty} I(t_n^*)$.
Da aber $(I(t_n))$ und $(I(\tilde{t}_n))$ beides Teilfolgen von $(I(t_n^*))$ sind, konvergieren $(I(t_n))$ und $(I(\tilde{t}_n))$ auch beide gegen A, also ist

$$\lim_{n\to\infty} I(t_n) = A = \lim_{n\to\infty} I(\tilde{t}_n).$$

□

Aufgrund des Lemmas ist die folgende Definition sinnvoll.

18.2.3 Definition und Satz (Integral einer Regelfunktion)

Ist $f : M \to \mathbb{R}$ eine Regelfunktion und ist (t_n) irgend eine Folge von Treppenfunktionen t_n auf M, die gleichmäßig gegen f konvergiert: $t_n \rightrightarrows f$, dann existiert der Grenzwert

$$I(f) := \lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} \int_a^b t_n(x)dx.$$

Es ist unabhängig von der Approximationsfolge (t_n) und heißt das *Integral von* f über $[a, b]$.

Bezeichnung: $I(f) = \int\limits_a^b f = \int\limits_a^b f(x)dx.$

x nennt man *Integrationsvariable* und die Funktion f auch *Integrand*.

18.2.4 Bemerkungen

(a) Falls f noch von zusätzlichen Parametern abhängt, z.B. $f(x) = x^2y$, sollte man die ausführliche Schreibweise verwenden, auf die Bezeichnung der Integrationsvariablen kommt es dabei aber nicht an:

$$\int_a^b f(x)dx = \int_a^b f(u)du = \int_a^b = f(t)dt = \cdots .$$

Wie wir bald sehen werden, gilt z.B.

$$\int_0^1 x^2ydx = \frac{y}{3}, \text{ aber } \int_0^1 x^2ydy = \frac{x^2}{2}.$$

(b) Wir haben für das Integral einer Regelfunktion keine neue Bezeichnung eingeführt, das macht Sinn, da jede Treppenfunktion $f = t$ eine Regelfunktion ist, als approximierende Folge (t_n) kann man $t_n = t$ für alle $n \in \mathbb{N}$ verwenden:

$$I(f) = I(t) = I(t_n).$$

18.2.5 Beispiel

Ein erstes Beispiel:
Um $\int\limits_0^1 x\ dx$ zu berechnen, betrachten wir die Folge (t_n) von Treppenfunktionen mit $t_n(x) := \frac{[nx]}{n}$, von der wir wissen, dass sie gleichmäßig gegen f mit $f(x) = x$ (sogar auf ganz $\mathbb{R}$) konvergiert.
Für jedes $n \in \mathbb{N}$ ist die Zerlegung $Z^{(n)} := \{0 = x_0 < x_1 < \ldots < x_{k-1} < x_k < \ldots < x_n = 1\}$ mit

$$x_k := \frac{k}{n} \text{ für } k = 0, \ldots, n$$

eine passende Zerlegung zu t_n und $I(t_n) = \frac{1}{2}\left(1 - \frac{1}{n}\right)$ nach ???.
Da $t_n \rightrightarrows id_{[0,1]}$ gilt, ist die Folge $(I(t_n))$ konvergent (das folgt hier auch aus den Rechenregeln für konvergente Folgen), und es gilt

$$\int\limits_0^1 x\ dx = \lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} \frac{1}{2}\left(1 - \frac{1}{n}\right) = \frac{1}{2}.$$

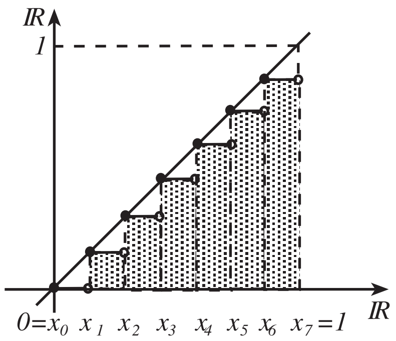


Abbildung 16: Visualisierung von t_7

Wir beschäftigen uns im Folgenden mit *Permanenzeigenschaften des Regelintegrals* und der Frage, wie umfangreich die Klasse der Regelfunktionen ist, einigen speziellen Integralberechnungen und einer Charakterisierung von Regelfunktionen durch innere Eigenschaften.
Da eine Treppenfunktion $t : [a, b] \to \mathbb{R}$ nur endlich viele Werte annimmt, ist jede Treppenfunktion $t : [a, b] \to \mathbb{R}$ beschränkt. Diese Eigenschaft überträgt sich auf Regelfunktionen:

Bezeichnung: $R(M) := \{f : M \to \mathbb{R};\ f \text{ Regelfukntion}\}$

Dann gilt:

18.2.6 Bemerkung

Es gilt $R(M) \subset B(M)$, d.h. jede Regelfunktion ist beschränkt.

Beweis : Gilt $t_n \rightrightarrows f$ mit $t_n \in T(M)$ und beachtet man

$$f = f - t_n + t_n \text{ und}$$

$$(*) \qquad \|f\| \leq \|f - t_n\| + \|t_n\|$$

und die Tatsache, dass Treppenfunktionen beschränkt sind, folgt die Behauptung unmittelbar aus $(*)$.

□

18.2.7 Satz (Eigenschaften des Regelintegrals)

Sind $f, g \in R(M)$ (also Regelfunktionen), dann gilt:

(a) $\alpha) f + g \in R(M)$ und $\beta) cf \in R(M)$ für beliebiges $c \in \mathbb{R}$, das bedeutet. $R(M)$ ist eine $\mathbb{R}$-Vektorraum (Untervektorraum von $B(M)$). Ferner ist

$$I(f+g) = I(f) + I(g) \text{ und } I(cf) = cI(f)$$

($I : R(M) \to \mathbb{R}$ ist also ein lineares Funktional)

Zusatz: Es ist auch $fg \in R(M)$, $R(M)$ ist also sogar eine *Funktionenalgebra*.

(b) $|I(f)| \leq (b-a)\|f\|$ (Standartabschätzung)
(I ist also ein beschränktes lineares Funktional)

(c) Aus $f \geq 0$ folgt $I(f) \geq 0$
(I ist ein nicht negatives lineares Funktional).

(c') Aus $f \leq g$ (d.h. $f(x) \leq g(x)$ für alle $x \in M$) folgt

$$I(f) \leq I(g)$$

(I ist also ein monotones lineares Funktional).

(d) Aus $m_1 \leq f(x) \leq m_2$ für alle $x \in M$ ($m_1, m_2 \in \mathbb{R}$ fest) folgt

$$m_1(b-a) \leq I(f) \leq m_2(b-a)$$

(e) Mit $f \in R(M)$ ist auch $|f| \in R(M)$, und es gilt

$$|I(f)| \leq I(|f|).$$

Beweis (a): Aus $t_n \rightrightarrows f,\ \tilde{t}_n \rightrightarrows g\ (t_n, \tilde{t}_n \in T(M))$ folgt

$$t_n + \tilde{t}_n \rightrightarrows f + g$$

und damit

$$\begin{aligned} I(f+g) = \lim_{n\to\infty} I(t_n + \tilde{t}_n) &= \lim_{n\to\infty} (I(t_n) + I(\tilde{t}_n)) \\ &= \lim_{n\to\infty} I(t_n) + \lim_{n\to\infty} I(\tilde{t}_n) \\ &= I(f) + I(g). \end{aligned}$$

Wegen $ct_n \rightrightarrows cf$ folgt

$$I(cf) = \lim_{n\to\infty} I(ct_n) = \lim_{n\to\infty} cI(t_n) = c \lim_{n\to\infty} I(t_n) = cI(f).$$

□

Will man den Zusatz beweisen, dann muss man zeigen:
Aus $t_n \rightrightarrows f$ und $\tilde{t}_n \rightrightarrows g$ folgt $t\tilde{t}_n \rightrightarrows fg$. Dazu verwendet man Standardschlüsse:

$$\|fg - t_n\tilde{t}_n\| = \|(f - t_n)g + t_n(g - \tilde{t}_n)\| \leq \|g\|\,\|f - t_n\| + \|t_n\|\,\|g - \tilde{t}_n\|.$$

Mit der Beschränktheit von $\|g\|$ und $\|t_n\|$ folgt die Behauptung.

Beweis (b): Wegen $|\,\|t_n\| - \|f\|\,| \leq \|t_n - f\|$ folgt aus $t_n \rightrightarrows f$ auch $\lim\limits_{n\to\infty} \|t_n\| = \|f\|$, daher

$$\begin{aligned} |I(f)| = |\lim_{n\to\infty} I(t_n)| &= \lim_{n\to\infty} |I(t_n)| \\ &\leq \lim_{n\to\infty} \|t_n\|(b-a) = \|f\|(b-a). \end{aligned}$$

□

Beweis (c): Aus $t_n \rightrightarrows f$ folgt auch $|t_n| \rightrightarrows |f| = f$
(beachte für alle $c, d \in \mathbb{R}$ gilt

$$|\,|c| - |d|\,| \leq |c - d| \text{ und damit (für } \textit{alle } x \in [a,b])$$

$$|\,|t_n(x)| - |f(x)|\,| \leq |t_n(x) - |f(x)|\,| \leq |t_n(x) - f(x)| \leq \|t_n - f\|)$$

Damit gilt auch $|t_n| \rightrightarrows |f|$ und somit $|f| \in R(M)$ und

$$I(f) = I(|f|) = \lim_{n\to\infty} I(|t_n|) \geq 0,$$

da $I(|t_n|) \geq 0$ für alle $n \in \mathbb{N}$ gilt.

□

Beweis (c'): (c') folgt aus (c) wegen der Linearität von I:

$$f \leq g \iff g - f \geq 0,$$

daher

$$I(g) - I(f) = I(g - f) \geq 0,$$

also $I(f) \leq I(g)$.

□

Beweis (d): Wegen $\int\limits_a^b 1 \cdot dx = b - a$, folgt (d) sofort durch mehrfache Anwendung von (c').
Aus $m_1 \cdot 1 \leq f(x)$ folgt $m_1(b-a) \leq I(f)$ und aus $f(x) \leq m_2 \cdot 1$ folgt $I(f) \leq m_1(b-a)$, also

$$m_1(b-a) \leq I(f) \leq m_2(b-a).$$

□

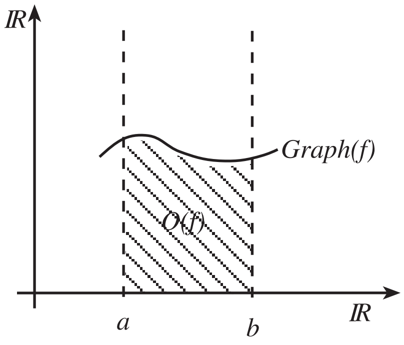
Beweis (e): Aus $t_n \rightrightarrows f$ folgt auch $|t_n| \rightrightarrows |f|$ wie unter (b) schon gezeigt, daher ist

$$\begin{aligned} |I(f)| = |\lim_{n\to\infty} I(t_n)| &= \lim_{n\to\infty} |I(t_n)| \\ &\leq \lim_{n\to\infty} I(|t_n|) = I(|f|). \end{aligned}$$

□

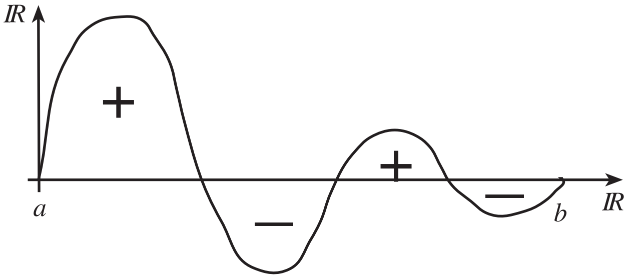
18.2.8 Geometrische Interpretation

Wie bei Treppenfunktionen (vgl. §17.1.6) hat man die folgende geometrische Inteerpretation:
Sei $f : [a,b] \to \mathbb{R}$ eine Regelfunktion, etwa f stetig und $f(X) \geq 0$ für alle $x \in [a,b]$ und $O(f) := \{(x,y) \in \mathbb{R} \times \mathbb{R};\ x \in (a,b);\ o \leq y \leq f(x)\} \subset \mathbb{R} \times \mathbb{R}$ ihre Ordinatenmenge. Dann ist $I(f) \geq 0$ und $I(f) = \int\limits_a^b f(x)dx$ kann man dann als Maß für den Flächeninhalt von $O(f)$ *definieren*.



(ein Flächeninhalt für solche (spezielle) Teilmengen von $\mathbb{R}^2$ ist bisjetzt nihct definiert!).

Nimmt f auch negative Werte an, so liefert $I(f) = \int\limits_a^b f$ ein Maß für den „orientierten Flächeninhalt" (d.h. Flächeninhalt mitt Vorzeichen) der Menge, deri zwischen dem Graphen von f und der x-Achse liegt:



Indem wir gleichmäßige Limites von Folgen von Treppenfunktionen betrachtet haben, ist es und gelungen, den Raum der Treppenfunktionen zu erweitern zum Raum der Regelfunktionen. Man könnte versuchen, diesen Prozess nochmals zu wiederholen, d.h. gleichmäßige Limites von Folgen von Regelfunktionen zu betrachten und man könnte erwarten, nochmals eine größere Klasse von Funktionen zu erhalten. Der folgende Satz besagt aber, dass der Raum der Regelfunktionen *stabil* gegenüber der Bildung von gleichmäßigen Limites von Regelfunktionen ist (hieraus folgt auch, dass $R(M)$ ein Banach-Raum beütglich der Supremumsnorm ist) und beinhaltet gleichzeitig eine wichtige *Vertauschungseigenschaft des Regelintegrals*.

18.2.9 Satz (Stabilitätssatz, Vertauschungssatz)

(a) Seien $f_n \in R(M)$ für $n \in \mathbb{N}$ und es gelte $f_n \rightrightarrows f : [a,b] \to \mathbb{R}$, dann gilt $f \in R(M)$ und es ist $I(f) = \lim\limits_{n\to\infty} I(f_n)$, oder vielleicht suggestiver

$$\int\limits_a^b f(x)dx = \int\limits_a^b \lim_{n\to\infty} f_n(x)dx = \lim_{n\to\infty} \int\limits_a^b f(x)dx.$$

D.h. der gleichmäßige limes einer Folge von Regelfunktionen ist wieder eine Regelfunktion und man darf zur Berechnung Integration und Grenzwertbildung vertauschen.

Da die Integration ja auch durch ein Grenzprozess definiert ist, handelt es sich dabei um die Vertauschung zweier Grenzprozesse.

(b) Ist (f_k) eine Folge von Regelfunktionen $f_k \in R(M)$, so dass die Funktionsreihe $\sum\limits_{k=0}^{\infty} f_k$ (d.h. die Folge $F_n := \sum\limits_{k=0}^{n} f_k$ der Partialsummen) gleichmäßig auf M gegen $F : [a,b] \to \mathbb{R}$ konvergiert, dann ist F auch eine Regelfunktion und man „darf Summation und Integration vertauschen".

$$\int_a^b F(x)dx = \int_a^b \left(\sum_{k=0}^{\infty} f_k(x) \right) dx = \sum_{k=0}^{\infty} \int_a^b f_k(x)dx.$$

Beweis (a): Wegen $f_n \in R(M)$ für $n \in \mathbb{N}$, kann man für jedes n ein $t_n \in T(M)$ mit $\|t_n - f_n\| < \frac{1}{n}$ wählen. Aus

$$\begin{aligned} \|t_n - f\| &= \|t_n - f_n + f_n - f\| \\ &\leq \|t_n - f_n\| + \|f_n - f\| \\ &\leq \frac{1}{n} + \|f_n - f\| \end{aligned}$$

und $f_n \rightrightarrows f$ folgt daher $\lim\limits_{n\to\infty} \|t_n - f\| = 0$, also $t_n \rightrightarrows f$, also $f \in R(M)$.
Nach uns jetzt schon vertrauter Schlussweise folgt nun

$$|I(f_n) - I(f)| = |I(f_n - f)| \leq (b-a)\|f_n - f\|, \text{ also}$$

$$\lim_{n\to\infty} I(f_n) = I(f) \text{ wegen } \lim_{n\to\infty} \|f_n - f\| = 0$$

□

Beweis (b): Man braucht (a) nur auf die Partialsummen $F_n := \sum\limits_{k=0}^{n} f_k$ anzuwenden und die Vertauschbarkeit der Indizes mit endlichen Summen beachten, via Induktion folgt aus der Additivität:

$$\int_a^b F_n(x)dx = \int_a^b \left(\sum_{k=0}^{n} f_k(x)dx \right) = \sum_{k=0}^{n} \int_a^b f_k(x)dx.$$

Der Grenzübergang $n \to \infty$ liefert dann nach (a)

$$\begin{aligned} \int_a^b \left(\sum_{k=0}^{\infty} f_k(x) \right) dx &= \int_a^b F(x)dx \\ &= \lim_{n\to\infty} \int_a^b F_n(x)dx \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} \int_a^b f_k(x)dx = \sum_{k=0}^{\infty} \int_a^b f_k(x)dx. \end{aligned}$$

□

Gegenbeispiel:
Bei nur punktweisen Konvergenz ist der Vertauschungssatz i.A. falsch:
Betrachte $t_n : [0,1] \to \mathbb{R}$ mit

$$t_n(x) = \begin{cases} 0, & \text{für } x = 0; \\ n, & \text{für } 0 < x < \frac{1}{n}; \\ 0, & \text{für } \frac{1}{n} < x \leq 1. \end{cases}$$

Die Folge (t_n) konvergiert punktweise gegen die Nullfunktion, es gibt aber $I(t_n) = \int\limits_0^1 t_n(x)dx = 1$ und damit auch $\lim\limits_{n\to\infty} I(t_n) = 1 \neq 0 = \int\limits_0^1 \lim\limits_{n\to\infty} t_n(x)dx$.

Wie umfangreich die Klasse der Regelfunktionen ist zeigt

18.2.10 Theorem

Jede stetige Funktion bzw. jede monotone Funktion $f : M \to \mathbb{R}$ ist eine Regelfunktion.

Der Beweis für *stetige* Funktionen beruht auf der Tatsache, dass eine stetige Funktion $f : K \to \mathbb{K}$, wobei $K \subset \mathbb{K}$ kompakt ist, auf K *gleichmäßig stetig* ist.

18.2.11 Definition

Eine Funktion $f : D \to \mathbb{K}$ $(D \subset K,\ D \neq \emptyset)$ heißt *gleichmäßig stetig auf* D, wenn folgendes gilt:
Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für *alle* $x, y \in D$ mit $|x - y| < \delta$ gilt $|f(x) - f(y)| < \varepsilon$.

Bemerkung:
Wählt man einen festen Punkt $a \in D$, dann sieht man sofort, dass eine gleichmäßige stetige Funktion in jedem Punkt $a \in D$ stetig ist. Der Unterschied zur gewöhnlichen Stetigkeit besteht darin, dass das zu jedem $\varepsilon > 0$ bei der gewöhnlichen Stetigkeitsdefinition existierende $\delta > 0$ neben seiner Abhängigkeit von ε i.A. von der betrachtenden Stelle a abhängt (man vergleiche dazu die Beispiele zur $\varepsilon - \delta-$Definition der Stetigkeit in §12).
Manchmal lässt sich jedoch ein universelles δ wählen, das nur von ε und nicht von der betrachtenden Stelle $a \in D$ abhängt, z.B. bei

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x \end{aligned} \qquad \text{oder} \qquad \begin{aligned} g : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto |x| \end{aligned}$$

Hier ist die Wahl $\delta = \varepsilon$ möglich. Ist jedoch der Definitionsbereich einer stetigen Funktion kompakt (vgl. ???), dann ist f gleichmäßig stetig.

18.2.12 Satz

Ist $K \subset \mathbb{K}$ kompakt($\neq \emptyset$) und $f : K \to \mathbb{K}$ stetig auf K.
Da ein Intervall $[a,b] \to \mathbb{R}$ kompakt ist, ist jede stetige Funktion $f : [a,b] \to \mathbb{R}$ gleichmäßig stetig auf $[a,b]$.

Beweis : Wir führen den Beweis für eine beliebige kompakte Teilmenge $K \subset \mathbb{R}$ oder $K \subset \mathbb{C}$, da beweistechnisch kein Unterschied besteht.
Wir führen einen indirekten Beweis, nehmen also an, dass f stetig auf K aber nicht gleichmäßig

stetig ist. Dann gibt es eine Ausnahme ε, nennen wir es ε_0, so dass kein geeignetes δ mit der genannten Eigenschaft existiert, das heißt für *jedes* $\delta > 0$ gibt es Paare $(x_\delta, y_\delta) \in K \times K$ für die zwar $|x_\delta - y_\delta| < \delta$ aber *nicht* $|f(x_\delta) - f(y_\delta)| < \varepsilon_0$ gilt, das heißt, es gilt $|f(x_\delta) - f(y_\delta)| \geq \varepsilon_0$. Wir setzen $\delta = 1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$ und bezeichnen die entsprechenden Punkte $x_{\delta_n} \in K$ und $y_{\delta_n} \in K$ einfach mit x_n und y_n. Für diese Punkte gilt also

$$|x_n - y_n| < \frac{1}{n}, \text{ aber } |f(x_n) - f(y_n)| \geq \varepsilon_0.$$

Da K kompakt ist, also insbesondere beschränkt ist, ist die Folge (x_n) beschränkt, besitzt also nach Bolzano-Weierstrass eine konvergente Teilfolge (x_{n_k}), die wegen der Kompaktheit (an der Stelle geht die Abgeschlossenheit von K ein) gegen einen Punkt $\xi \in K$ konvergiert:

$$\xi := \lim_{k\to\infty} x_{n_k}.$$

Wegen $|x_n - y_n| < \frac{1}{n}$ gilt dann auch $\lim\limits_{k\to\infty} y_{n_k} = \xi$.
Wegen der Stetigkeit von f in ξ folgt daher

$$\lim_{k\to\infty} f(x_{n_k}) = f(\xi) = \lim_{k\to\infty} f(y_{n_k}),$$

oder auch $\lim\limits_{k\to\infty} |f(x_{n_k}) - f(y_{n_k})| = 0$ im Widerspruch zu $|f(x_n) - f(y_n)| \geq \varepsilon_0$ für alle $n \in \mathbb{N}$.

Dieser Widerspruch zeigt, dass unsere Annahme falsch war, f also doch gleichmäßig stetig auf K ist.

□

Wir kommen nun zum Beweis von Theorem 17.2.10 und zeigen zunächst, dass jede stetige Funktion

$$f : [a, b] \to \mathbb{R}$$

eine Treppenfunktion ist.
Die Idee ist, das Intervall $[a, b]$ hinreichend fein zu unterteilen und eine geeignete Folge von Treppenfunktionen (t_n) mit $t_n \rightrightarrows f$ zu konstruieren. Dazu wählen wir eine Folge $(Z^{(n)})$ von äquidistanten Zerlegungen

$$Z^{(n)} = \{a = x_0^{(n)} < x_1^{(n)} < \ldots < x_{k-1}^{(n)} < x_k^{(n)} < \ldots < x_n^{(n)}\}$$

mit $x_k^{(n)} = a + k \cdot \frac{b-a}{n},\ k = 0, 1, \ldots, n$ und definieren $t_n : [a, b] \to \mathbb{R}$ durch

$$t_n(x) = \begin{cases} f(x_{k-1}), & \text{für } x \in \left[x_{k-1}^{(n)}, x_k^{(n)}\right[,\ 1 \leq k \leq n \\ f(b), & \text{für } x = x_n^{(n)} = b \end{cases}$$

Für diese speziell ausgewählte Folge (t_n) von Treppenfunktionen zeigen wir, dass $t_n \rightrightarrows f$ gilt.
Dazu benutzen wir die *gleichmäßige Stetigkeit* von f auf $[a, b]$:
Zu jedem $\varepsilon > 0$ gibt es daher ein Universal-$\delta > 0$, so dass für alle $x, y \in [a, b]$ mit $|x - y| < \delta$ gilt $|f(x) - f(y)| < \varepsilon$.
Zunächst wählen wir eine natürliche Zahl N so groß, dass $\frac{b-a}{N} < \delta$ gilt.
Für $n \geq N$ betrachten wir $|f(x) - t_n(x)|$ für $x \in [a, b]$. Jedes $x \in [a, b],\ x \neq b$, liegt in genau einem der Intervalle $\left[x_{k-1}^{(n)},\ x_k^{(n)}\right[$, $k = 1, \ldots, n$. Sei also $x \in \left[x_{k-1}^{(n)},\ x_k^{(n)}\right[$, dann ist nach Definition von t_n

$$f(x) - t_n(x) = f(x) - f\left(x_{k-1}^{(n)}\right)$$

und weiter

$$\left|x - x_{k-1}^{(n)}\right| \leq \left|x_k^{(n)} - x_{k-1}^{(n)}\right| = x_k^{(n)} - x_{k-1}^{(n)} \leq \frac{b-a}{n} < \frac{b-a}{N} < \delta$$

Aus $\left|x - x_{k-1}^{(n)}\right| < \delta$ folgt dann wegen der gleichmäßigen Stetigkeit von f

$$|f(x) - t_n(x)| = \left|f(x) - f(x_{k-1}^{(n)})\right| < \varepsilon$$

Da für $x = b$ nach der Definition

$$|f(b) - t_n(b)| = |f(b) - f(b)| = 0$$

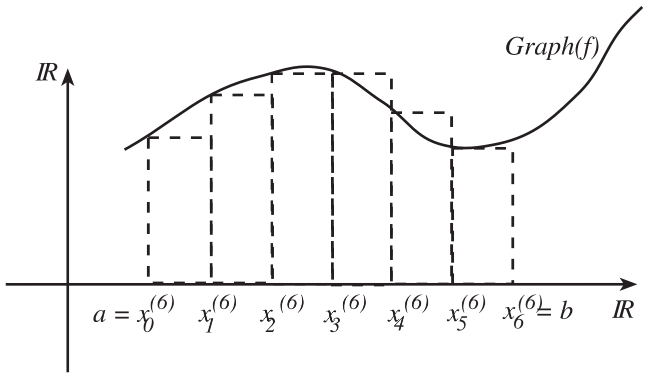
gilt, ist also für alle $n \geq N$ und alle $x \in [a, b]$

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

und damit

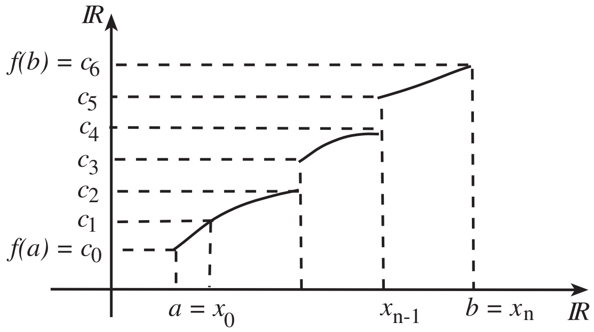
$$\|f - t_n\| \leq \varepsilon$$

für alle $n \geq N$, d.h. $t_n \rightrightarrows f$.
Damit haben wir also nach einem ganz speziellen Verfahren (äquidistante Festlegung des Funktionswertes) eine Folge (t_n) von Treppenfunktionen $t_n : [a, b] \to \mathbb{R}$ konstruiert, die gleichmäßig gegen die gegebene stetige Funktion f konvergiert. Damit ist f also eine Regelfunktion (Die Abbildung veranschaulicht den Fall $n = 6$.):



Wir kehren zum Beweis vom Theorem 17.2.10 zurück und zeigen jetzt, dass jede *monotone* Funktion $f : [a, b] \to \mathbb{R}$ eine Regelfunktion ist.

OBdA sei f monoton wachsend (sonst gehe man von f zu $-f$ über)



Die Bildmenge $f([a, b])$ ist dann im Intervall $[f(a), f(b)]$ enthalten. Wir unterteilen $[f(a), f(b)]$ äquidistant, etwa

$$f(a) = c_0 < c_1 < c_2 < \ldots < c_{\nu-1} < c_\nu < \ldots < c_n = f(b)$$

mit $c_\nu := f(a) + \frac{\nu(f(b)-f(a))}{n}$ für $0 \leq \nu \leq n$. Außerdem setzen wir

$$B_\nu = [c_{\nu-1}, c_\nu] \text{ für } 1 \leq \nu \leq n \text{ und}$$

$$A_\nu = f^{-1}(B_\nu) := \{x \in [a, b];\ c_{\nu-1} \leq f(x) \leq c_\nu\}.$$

Aus der Monotonie von f folgt, dass A_ν ein (möglicherweise leeres) Intervall ist. Wir interessieren uns nur für die „echten“ Intervalle, also diejenigen, die mehr als einen Punkt enthalten. Wir nummerieren die Randpunkte dieser Intervalle durch und erhalten in

$$Z := \{a = x_0 < x_1 < \ldots < x_{j-1} < x_j < \ldots < x_n := b\}$$

eine Zerlegung von $[a, b]$.
Wir wählen nun noch irgendeinen Zwischenpunkt $\xi_j \in]x_{j-1}, x_j[$ mit $1 \leq j \leq k$. (z.B. kann man $\xi_j = \frac{x_{j-1}+x_j}{2}$, also der Mittelpunkt des Intervalls wählen) und definieren nun

$$t_n : [a, b] \to \mathbb{R}$$

durch

$$t_n(x) = \begin{cases} f(\xi_j), & \text{für } x \in]x_{j-1}, x_j[\ (1 \leq j \leq k); \\ f(x_j), & \text{für } x = x_j, \ (j = 0, \ldots, k). \end{cases}$$

Dann ist t_n eine Treppenfunktion bezüglich der Zerlegung Z.
Wir schätzen jetzt $f(x) - t_n(x)$ ab. Nach Konstruktion ist das Bild irgendeines der Intervalle $]x_{j-1}, x_j[$ enthalten in einem Intervall der Länge $\frac{f(b)-f(a)}{n}$. Hieraus und aus $t_n(x_j) = f(x_j)$ folgt dann für alle $x \in [a, b]$

$$|f(x) - t_n(x)| \leq \frac{f(b) - f(a)}{n},$$

und damit

$$\lim_{n\to\infty} \|f - t_n\| = 0.$$

Damit ist gezeigt, dass die Folge (t_n) gleichmäßig gegen f konvergiert. Damit ist f eine Regelfunktion.

□

Mit der stetigen Funktion und der monotonen Funktion haben wir einen großen Vorrat von Regelfunktionen.
Da die Funktionen, für die wir ein Integral erklärt haben genau die Regelfunktionen $f : [a, b] \to \mathbb{R}$ sind, nennen wir die Regelfunktionen auch die *integrierbare Funktionen*:

$$\boxed{\text{Integrierbare Funktion = Regelfunktion}}$$

Bei Zugrundelegung anderer Integralbegriffe erhält man andere Funktionsklassen.

$$\text{Riemann-Integral} \Longleftrightarrow \text{Riemann-integrierbare Funktion}$$

$$\text{Lebesque-Integral} \Longleftrightarrow \text{Lebesque-integrierbare Funktion}$$

Obwohl wir häufig mit äquidistanten Zerlegungen gearbeitet haben, erhält man eine größere Flexibilität bei der konkreten Integralberechnung, wenn man auch andere Zerlegungen (oder Folgen von Zerlegungen) betrachtet, die nur die Eigenschaft haben müssen, dass sie „fein genug“ sind. Diesen Begriff werden wir gleich präzisieren.
Außerdem wollen wir auch eine größere Flexibilität bei der Auswahl der Zwischenpunkte ξ_j im j-ten Teilungintervall. Wir führen dazu die folgende Sprechweise ein.

18.2.13 Definition

Ist $f : [a, b] \to \mathbb{R}$ eine Regelfunktion

$$Z^{(n)} := \{a = x_0^{(n)} < x_1^{(n)} < \ldots < x_{k-1}^{(n)} < x_k^{(n)} < \ldots < x_{r_n}^{(n)} = b\}$$

eine (zunächst feste Zerlegung von $[a,b]$ ($n \in \mathbb{N}$ fest). Wählt man für jedes k mit $1 \leq k \leq r_n$ einen Punkt $\xi_k^{(n)} \in [x_{k-1}, x_k],\ 1 \leq k \leq r_n$, dann heißt der Vektor

$$\xi_k^{(n)} = (\xi_1^{(n)}, \dots, \xi_{r_n}^{(n)})$$

ein „Zwischenvektor“ und die Summe

$$S(f; Z^{(n)}, \xi^{(n)}) = \sum_{k=0}^{r_n} f(\xi_k^{(n)})(x_k^{(n)} - x_k^{(n-1)})$$

eine *Riemannsche Summe* für f zur Zerlegung $Z^{(n)}$ und zum Zwischenvektor $\xi^{(n)}$.
Die maximale Länge der Teilintervalle

$$\eta(Z^{(n)}) = \max\{x_k^{(n)} - x_{k-1}^{(n)};\ 1 \leq k \leq r_n\}$$

heißt die *Feinheit*(Feinheitsmaß, Feinheitsgrad) der Zerlegung $Z^{(n)}$.
Beachte: $S(f, Z^{(n)}, \xi^{(n)})$ ist das Interval einer geeigneten Treppenfunktion (welcher?) und jedes Integral einer Treppenfunktion ist eine spezielle Riemannsche Summe.

18.2.14 Satz (Berechnung von Integralen mit der Riemannschen Summe)

Ist $f : [a,b] \to \mathbb{R}$ eine Regelfunktion und $Z^{(n)} := \{a = x_0^{(n)} < x_1^{(n)} < \ldots < x_k^{(n)} < \ldots < x_{r_n}^{(n)} = b\}$ irgendeine Folge von Zerlegungen mit $\lim\limits_{n\to\infty} \eta(Z^{(n)}) = 0$ („Zerlegungsnullfolge“), dann konvergiert für *jede* Folge von Zwischenvektoren $(\xi^{(n)})$, $\xi^{(n)} = (\xi_1^{(n)}, \dots, \xi_{r_n}^{(n)})$ mit $\xi_k^{(n)} \in [x_{k-1}^{(n)}, x_k^{(n)}];\ 1 \leq k \leq r_n$ die Folge der Riemannschen Summen $\big(S(f, Z^{(n)}, (\xi^{(n)})\big)_{n\geq 1}$ gegen das Integral von f:

$$\lim_{n\to\infty} S(f, Z^{(n)}, (\xi^{(n)})) = \lim_{n\to\infty} \sum_{k=0}^{r_n} f(\xi_k^{(n)})(x_k^{(n)} - x_{k-1}^{(n)}) = I(f) = \int\limits_a^b f.$$

Beweisskizze:
Wir beweisen die Behauptung nur für *stetige* Funktionen $f : [a,b] \to \mathbb{R}$.
Für eine *beliebige* Regelfunktion vergleiche man die Musterlösung zu Aufgabe ?? von Blatt 2.
Man imitiere den Beweis vom Theorem und wähle eine Folge von $\eta(Z^{(n)})$ der Zerlegungen, so dass $\eta(Z^{(n)}) < \delta$ für $n \geq N$ gilt (dieses δ ist das Universaldelta, was es auf Grund der gleichmäßigen Stetigkeit von f zu einem vorgegebenem $\varepsilon > 0$ gibt).
Definiert man dann t_n $(n \geq N)$ in Analogie zum Beweis von 17.2.10 für stetige Funktionen, wobei man die Zwischenpunkte $\xi_k^{(n)} \in [x_{k-1}^{(n)}, x_k^{(n)}]$ beliebig wählen kann, dann gilt $t_n \rightrightarrows f$ und

$$S(f, Z^{(n)}, (\xi^{(n)})) = I(t_n)$$

und damit

$$\lim_{n\to\infty} S(f, Z^{(n)}, (\xi^{(n)})) = \lim_{n\to\infty} I(t_n) = I(f).$$

18.2.15 Beispiele und Bemerkungen

(a) Häufig verwendet man eine Folge $Z^{(n)}$ von äquidistanten Zerlegungen von $[a,b]$, hier gilt also $r_n = n$ für alle $n \in \mathbb{N}$ und $x_k = a + k\frac{b-a}{n},\ h = 0, \dots, n.$

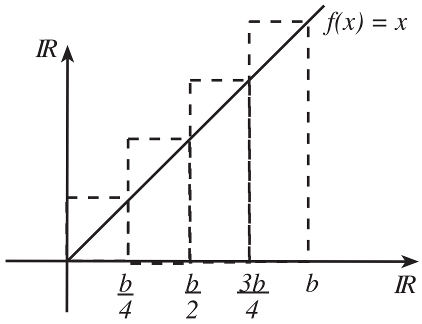
Als erstes Beispiel hatten wir in §17.2.5 das Integral $\int\limits_0^1 x\, dx$ berechnet und den Wert $\frac{1}{2}$ erhalten (was zu erwarten war). Wir wollen etwas allgemeiner das Integral $\int\limits_0^b x\, dx$ berechnen und

verwenden dazu Riemansche Summen.
Wir wählen für jedes $n \in \mathbb{N}$ äquidistante Zerlegungen

$$Z^{(n)} := \left\{0, \frac{b}{n}, \frac{2b}{n}, \ldots, b\right\}$$

sowie die Zwischenpunkte (Stützpunkte): $\xi_k^{(n)} = x_k^{(n)} = k\frac{b}{n}$ für $b = 1, \ldots, n$. (als Zwischenpunkte haben wir also die rechten Randpunkte des Intervall $[x_{k-1}^{(n)}, x_k^{(n)}]$ gewählt). Es gilt also

$$S_n := S(f;\ Z^{(n)}, \xi^{(n)}) = \sum_{k=1}^{n} \left(\frac{kb}{n}\right) \frac{b}{n} = \frac{b^2}{n^2} \sum_{k=1}^{n} k = \frac{b^2}{n^2} \frac{n(n+1)}{2}$$



Also ist (beachte $\lim\limits_{n\to\infty} \eta(Z^{(n)}) = \lim\limits_{n\to\infty} \frac{b-a}{n} = 0$)

$$\lim_{n\to\infty} S_n = \int_a^b f(x)dx = \int_0^b x\ dx = \frac{b^2}{2}.$$

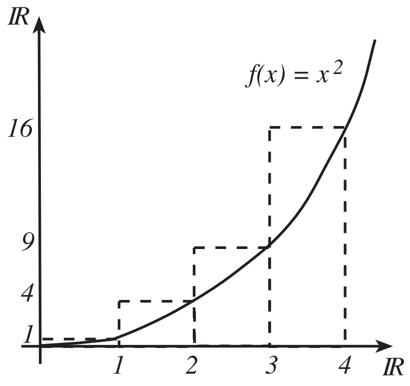
$\square$

(b) Mit der Summenformel (vgl. §2.4(1))

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

findet man völlig analog

$$\int_0^b x^2 dx = \frac{b^3}{3}.$$



(c) Wir wollen allgemeiner das Integral $\int\limits_a^b p_\alpha(x)dx$ mit $p_\alpha(x) = x^\alpha,\ \alpha \in \mathbb{R},\ \alpha \neq -1$ und $0 < a < b$ berechnen.
Wir wählen dazu schon von Fermat benutzte „geometrische Zerlegung " von $[a, b]$.
Für $n \in \mathbb{N}$ sei

$$Z_q^{(n)} := \{a = x_0^{(n)} < x_1^{(n)} = aq < x_2^{(n)} = aq^2 < \ldots < x_k^{(n)} = aq^k < \ldots < x_n^{(n)} = aq^n = b\}$$

mit $q := \sqrt[n]{\frac{b}{a}}$.

$\xi_k^{(n)} = x_k^{(n)} = aq^{k-1}$ für $k = 1, \ldots, n$.
Es ist $\eta(Z_q^{(n)}) = \max\{qq^{k-1}(q-1);\ k = 1, \ldots, n\} \leq b(q-1)$ und wegen $\lim\limits_{n\to\infty} \sqrt[n]{\frac{b}{a}} = 1$ gilt also $\lim\limits_{n\to\infty} \eta(Z_g^{(n)}) = 0,\ (Z_g^{(n)})$ ist also eine Zerlegungsnullfolge.
Es ist dann

$$\begin{aligned}
S_n := S(p_\alpha, Z_g^{(n)}, \eta^{(n)}) &= \sum_{k=1}^{n} p_\alpha(\eta_k^{(n)})(x_k^{(n)} - x_{k-1}^{(n)}) \\
&= a \cdot \sum_{k=1}^{n} (aq^{k-1})\alpha(q^k - q^{k-1}) \\
&= a^{\alpha+1}(q-1) \sum_{k=1} q^{(k-1)(\alpha+1)} \\
&= a^{\alpha+1} \frac{q^{(\alpha+1)n} - 1}{q^{n+1} - 1} \quad \text{(Summenformel für die endliche geometrische Reihe)} \\
&= (b^{\alpha+1} - a^{\alpha+1}) \frac{q-1}{q^{n+1} - 1}.
\end{aligned}$$

Nun gilt aber für $q := q(n) := \sqrt[n]{\frac{b}{a}} > 1$

$$\lim_{n\to\infty} q(n) = 1$$

und mit der Substitution $q := \exp(t)$ erhält man mit $\beta = \alpha + 1$

$$\begin{aligned}
\lim_{q \to 1} \frac{q^\beta - 1}{q - 1} &= \lim_{t \to 0} \frac{\exp(\beta t) - 1}{\exp t(t) - 1} \\
&= \lim_{t \in \infty} \beta \frac{\exp(\beta t) - 1}{\beta t} \cdot \frac{t}{\exp(t) - 1} \\
&= \beta \lim_{t \to 0} \frac{\exp(\beta t) - 1}{\beta t} \lim_{t \to 0} \left(\frac{\exp(t) - 1}{t} \right)^{-1} \\
&= \beta \cdot 1 \cdot 1 = \beta = \alpha + 1.
\end{aligned}$$

(Hierbei wurde $\lim\limits_{y \to 0} \frac{\exp(y)-1}{y} = 1$ verwendet).

Daher ist $\lim\limits_{n \to \infty} S_n = \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha+1}$, also

$$\boxed{\int_a^b x^\alpha dx = \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha+1}}.$$

Mit dem Hauptsatz der Differenzial-und Integralrechnung (vgl. §20) ist ist dieses doch etwas mühselig gewonnene Resultat trivial.

(d) Als weiteres Beispiel wollen wir das Integral $\int\limits_a^b \frac{1}{x}\, dx$ für $0 < a < b$ berechnen.

Wir wählen wieder für jedes $n \in \mathbb{N}$ die geometrische Zerlegung

$$Z_g^{(n)} := \{a = x_0^{(n)} < x_1^{(n)} = aq < \ldots < a_k^{(n)} = aq^k < \ldots < x_n^{(n)} = aq^n = b\}$$

mit $q := \sqrt[n]{\frac{b}{a}}$.

Wir wählen $\eta_k^{(n)} = x_{k-1}^{(n)} = aq^{k-1}$ für $k = 1, \ldots, n$ und erhalten mit $J(x) := \frac{1}{x}$

$$\begin{aligned}
S_n := S(J; Z_g^{(n)}, \xi^{(n)}) &= \sum_{k=1}^n J(\xi_k^{(n)})(x_k^{(n)} - x_{k-1}^{(n)}) \\
&= \sum_{k=1}^n \frac{1}{aq^{k-1}}(aq^k - aq^{k-1}) \\
&= \sum_{k=1}^n (q - 1) \\
&= n \cdot (q - 1) = n \left(\sqrt[n]{\frac{b}{a}} - 1 \right).
\end{aligned}$$

Nach Übungsaufgabe ??? von Blatt ??? gilt aber

$$\lim_{n \to \infty} n \left(\sqrt[n]{\frac{b}{a}} - 1 \right) = \log \frac{b}{a}.$$

Dies kann man auch direkt sehen, denn es ist

$$\begin{aligned}
\lim_{n \to \infty} n \left(\sqrt[n]{\frac{b}{a}} - 1 \right) &= \lim_{n \to \infty} \frac{\exp\left(\frac{1}{n} \log \frac{b}{a} - 1\right)}{\log \frac{b}{a}} \log \frac{b}{a} \\
&= \log \frac{b}{a} \lim_{t \to 0} \frac{\exp(t) - 1}{t} = \log \frac{b}{a} \quad \text{(mit } t := \tfrac{\log \frac{b}{a}}{n}\text{)}
\end{aligned}$$

Daher ist

$$\lim_{n \to \infty} S_n = \int_a^b J(x) dx = \int_a^b \frac{1}{x} dx = \log \frac{b}{a} = \log b - \log a$$

(Letzteres nach der Funktionalgleichung des Logarithmus).
Auch hier fällt uns das Ergebnis mit dem Hauptsatz in den Schoss, denn es ist $\log'(x) = \frac{1}{x} = J(x)$, der Logarithmus ist also eine Stammfunktion von J.

(e) Mit Hilfe einer Folge von äquidistanten Zerlegungen von $[a,b]$ lässt sich leicht $\int\limits_a^b \exp(x)dx = \exp(b) - \exp(a)$ zeigen.
Wir wollen das Integral mit Hilfe des Vertauschungssatzes (vgl. §17.2.9), aber einfacher berechnen:
Da die Exponentialreihe das Konvergenzradius ∞ hat, gilt für alle $a, b \in \mathbb{R}$ mit $a < b$

$$\begin{aligned}
\int_a^b \exp(x)dx &= \int_a^b \sum_{k=0}^{\infty} \frac{x^k}{k!}dx \\
&= \sum_{k=0}^{\infty} \int_a^b x^k dx \\
&= \sum_{k=0}^{\infty} \frac{b^{k+1} - a^{k+1}}{(k+1)!} \\
&= 1 + \sum_{k=0}^{\infty} \frac{b^{k+1}}{(k+1)!} - \left(1 + \sum_{k=0}^{\infty} \frac{a^{k+1}}{(k+1)!}\right) \\
&= \exp(b) - \exp(a).
\end{aligned}$$

Genauso zeigt man etwa

$$\int_b^a \cos x \; dx = \sin b - \sin a$$

Wir wollen uns noch mit der Frage beschäftigen, welche Funktionen $f : [a,b] \to \mathbb{R}$ *keine* Regelfunktion sind.
Dazu stellen wir fest:

18.2.16 Bemerkung

Für jede Treppenfunktion $t : M := [a,b] \to \mathbb{R}$ existieren für jeden Punkt x_0 mit $a < x_0 < b$ die einseitigen Grenzwerte

$$f(x_{0_+}) := \lim_{\substack{x \to x_0 \\ x \in M \\ x > x_0}} f(x) \quad \text{(rechtseitige Grenzwerte)}$$

und

$$f(x_{0_-}) = \lim_{\substack{x \to x_0 \\ x \in M \\ x < x_0}} f(x)$$

und für die Endpunkte a und b existieren

$$f(a_+) = \lim_{\substack{x \to a \\ x > a}} f(x) \text{ und } f(b_-) = \lim_{\substack{x \to b \\ x < b}} f(x).$$

Diese Eigenschaft überträgt sich sofort auch auf einen gleichmäßigen limes vom Treppenfunktionen (Skizzieren Sie einen Beweis!), also Regelfunktionen. Damit haben wir:

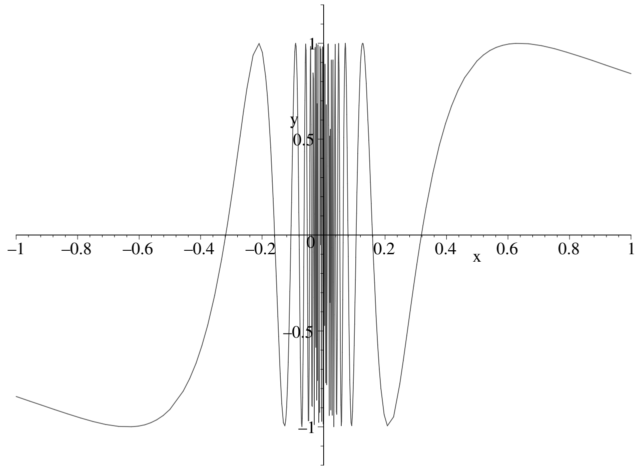


Abbildung 17: Graph der Funktion $\sin \frac{1}{x}$

18.2.17 Satz

Für jede Regelfunktion $f : M \to \mathbb{R}$ existieren für alle $x_0 \in]a, b[$ Grenzwerte $f(x_{0_+})$, $f(x_{0_-})$ und für a bzw. b existieren $f(a_+)$ bzw. $f(b_-)$.

Im Umkehrschluss bedeutet dies:
Wenn für eine Funktion $f : [a, b] \to \mathbb{R}$ einer dieser Grenzwerte *nicht existiert*, dann ist f keine Regelfunktion. So ist also zum Beispiel die Dirichlet- Funktion $\delta : [0, 1] \to \mathbb{R}$

$$\delta(x) = \begin{cases} 1, & \text{falls } x \in \mathbb{Q} \\ 0, & \text{sonst} \end{cases}$$

keine Regelfunktion, weil in keinem Punkt $x_0 \in [0, 1]$ einer der möglichen Grenzwerte existiert. Mit dem obigen Kriterium erhält man auch, dass die oszilierende Funktionen $f : [0, 2\pi] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} \sin \frac{1}{x}, & \text{falls } x \neq 0 \\ 0, & \text{falls } x = 0 \end{cases}$$

keine Regel Funktion ist. (siehe Abb. 17)
Der Grenzwert $f(0_+)$ existiert nicht, denn für die beiden Folgen (x_n) und (y_n) mit $x_n = \frac{2}{(4n+1)\pi}$ bzw. $y_n = \frac{2}{(4n-1)\pi}$ gilt $x_n > 0$ und $y_n > 0$ und $\lim\limits_{n\to\infty} f(x_n) = 1 \neq -1 = \lim\limits_{n\to\infty} f(y_n)$ (beachte: es gilt zum Beispiel schon $f(x_n) = 1$).
Das f keine Regelfunktion sein kann, kann man auch so einsehen:
Ist $Z = \{0 = x_0 < x_1 < \ldots < x_n = 2\pi\}$ eine Zerlegung von $[0, 2\pi]$ und t eine Treppenfunktion zu dieser Zerlegung, dann ist t auf dem Interval $]0, x_1[$ konstant. Zu diesem Interval gibt es sowohl Punkte x mit $f(x) = 1$, als auch solche mit $f(x) = -1$. (Man braucht bei den Folge (x_n) und (y_n) n nur hinreichend groß zu wählen.) Daher ist

$$\|f - t\| = \sup\{|f(x) - t(x)|;\ x \in [0, 2\pi]\} \geq 1$$

für jede Treppenfunktion $t : [0, 2\pi] \to \mathbb{R}$, es kann daher auch keine Folge $(t_n), t_n \in T([0, 2\pi])$ mit $t_n \rightrightarrows f$ geben.
f ist auch aus diesem Grund keine Regelfunktion (jedoch Riemann-integrierbar).

Nach Satz ??? gilt auch die Umkehrung, das heißt es gilt:

18.2.18 Charakterisierung von Regelfunktionen durch innere Eigenschaft

Eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ ist genau dann eine Regelfunktion, wenn in jedem Punkt $x_0 \in [a, b]$ die dort möglichen einseitigen Grenzwerte existieren, das heißt es existieren $f(x_{0_+})$ für $a \leq x_0 < b$ und $f(x_{0_-})$ für $a < x_0 \leq b$.

Beweis : Eine Regelfunktion $f : [a, b] \to \mathbb{R}$ ist beschränkt und es existieren die möglichen einseitigen Grenzwerte.
Für die Umkehrung nehmen wir an, dass für $x \in [a, b[$ bzw. $x \in]a, b]$ die Grenzwerte $f(x_+)$ bzw. $f(x_-)$ existieren.
Zu vorgegebenem $\varepsilon > 0$ bilden wir rekursiv

$$x_0 = a, x_1, x_2, \ldots$$

als

$$x_j := sup\{x \in]x, b[;\ |f(x) - f(x_{j-1\,+}| < \varepsilon\}$$

für $j \geq 1$.

Dann gilt offensichtlich $a = x_0 < x_1 < x_2 < \ldots$. Dieser Prozess muss aber nach endlich vielen Stellen abbrechen, und es gilt dann $x_m = b$ mit einem geeigneten $m \in \mathbb{N}$. Würde der Prozess nämlich nicht abbrechen, dann gebe es einen Punkt $x_* \in]a, b]$, gegen den die Folge (x_j) streng monoton konvergiert. Dann kann aber der einseitige Grenzwert $f(x_{*_-})$ nicht existieren, weil in jedem Interval $]x_* - \delta, x_*[$ mit $\delta > 0$ gelten würde

$$sup\{|f(x) - f(y)|;\ x, y \in]x_* - \delta, x_*[\} \geq \varepsilon.$$

Also gibt es ein $m \in \mathbb{N}$ mit $x_m = b$ und mit Hilfe der Zerlegung

$$Z = \{a = x_0 < x_1 < \ldots < x_m = b\}$$

definieren wir eine Treppenfunktion $t : M = [a, b] \to \mathbb{R}$ durch

$$t(x) := \begin{cases} f(x_{j-1_+}), & \text{für } x_{k-1} \leq x < x_k;\ 1 \leq k \leq m \\ f(b) & \text{für } x = b. \end{cases}$$

Dann folgt $|f(x) - t(x)| < \varepsilon$ für alle $x \in [a, b]$ und damit $\|f - t\| \leq \varepsilon$.
Setzt man nun die Reihe nach $\varepsilon = 1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots$ein, so erhält man eine Folgen (t_n) von Treppenfunktionen $t_n : [a, b] \to \mathbb{R}$ mit $t_n \rightrightarrows f$, also ist f eine Regelfunktion.
Dieser Beweis folgt im Wesentlichen dem von *S.Hildebrandt in S.Hildebrandt: Analysis 1, Springer-Verlag 2002, Seite 330.*

□

Einen weiteren Beweis für die schwierige Beweisrichtung findet man zum Beispiel bei: *K.Königsberger: Analysis 1, Springer-Verlag, 5.Auflage, 2000, S.194.*
Einen weiteren (relativ einfachen) Beweis findet man bei: *Barmer-Floh: Analysis 1*;
Dort wird die „Überdeckungskompaktheit“ des kompakten Intervals benutzt.

18.2.19 Bemerkungen und Beispiele

(a) Aus Satz 17.2.18 folgt nochmals, dass jede *stetige* Funktion $f : [a, b] \to \mathbb{R}$ bzw. jede *monotone* Funktion $f : [a, b] \to \mathbb{R}$ eine Regelfunktion ist. Insbesondere ist auch jede *stückweise stetige* Funktion $f : [a, b] \to \mathbb{R}$ eine Regelfunktion. Dabei heißt f stückweise stetig, wenn es eine Zerlegung

$$Z := \{a = x_0 < x_1 < \ldots < x_{k-1} < x_k < \ldots < x_n = b\}$$

gibt, so dass $f_k := f|\,]x_{k-1}, x_k[$ stetig ist für $1 \leq k \leq n$ und die f_k zu stetigen Funktionen auf $[x_{k-1}, x_k]$ fortgesetzt werden kann.

(b) Nach Satz 17.2.18 ist die Riemann-Funktion (vgl. §11.1.2(b)) $f : [0,1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 0, & \text{falls } x \text{ irrational} \\ \frac{1}{q_x}, & \text{falls } x = \frac{p_x}{q_x} \text{ mit } (p_x, q_x) \in \mathbb{N}_0 \times \mathbb{N}, q_x \text{ minimal} \end{cases}$$

eine Regelfunktion, weil für *jeden* Punkt $x_0 \in [0,1]$ gilt $\lim\limits_{x \to x_0} f(x) = 0$. (Vgl. dazu auch die Musterlösung zur Aufgabe ??? von Blatt ???).

Man kann auch direkt eine ε-Approximation durch eine Treppenfunktion für f angegeben:
In der Übungsaufgabe wurde gezeigt, dass es zu vorgegebene $\varepsilon > 0$ nur endlich viele rationale Zahlen $r_1, \dots, r_N \in [0,1]$ gibt mit $f(r_j) > \varepsilon$. Definiert man t_ε als die Summe der charakteristischen Funktion der endlich vieler Punkte, das heißt $t_\varepsilon := \sum\limits_{k=1}^{N} \chi_{\{x_k\}}$, dann ist $t_\varepsilon \in T([0,1])$, $I(t_\varepsilon) = 0$ und $\|f - t_\varepsilon\| \leq \varepsilon$.

Das zeigt, dass f ein Regelfunktion ist und dass gleichzeitig $I(f) = \int\limits_0^1 f(x)dx = 0$ gilt.

18.2.20 Definition

Eine Unstetigkeitsstelle x_0 einer beschränkten Funktion $f : [a,b] \to \mathbb{R}$ heiße *Sprungstelle*, wenn $f(x_{0_+})$ im Fall $a \leq x_0 < b$ und $f(x_{0_-})$ im Falle $a < x_0 \leq b$ existieren.
Eine Umformulierung von Satz ist dann:

18.2.21 Satz

Für eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ gilt:
f ist genau dann eine Regelfunktion, wenn f (höchstens) abzählbar viele Sprungstellen als Unstetigkeitsstellen hat. ((Höchstens) Abzählbar soll heißen: endlich oder abzählbar unendlich.)
Ist f eine Regelfunktion, dann hat nach Satz 17.2.18 f höchstens Sprungstellen als Unstetigkeitsstellen. (für eine stetige Funktion ist diese Menge leer).
Ferner kann man sich leicht überlegen, dass es für jedes $n \in \mathbb{N}$ höchste endlich viele Sprungstellen $x \in M$ gibt, für welche eine der Zahlen

$$|f(x_{0_+}) - f(x_{0_-}),\ |f(x_{0_+}) - f(x_0|,\ |f(x_{0_-}) - f(x_0)|$$

nicht kleiner als $\frac{1}{n}$ ist. Daher gibt es (höchstens) abzählbar viele Sprungstellen von f. Hat umgekehrt f nur Sprungstellen als Unstetigkeitsstellen, so existieren die möglichen einseitigen Grenzwerte für alle $x_0 \in [a,b]$, f ist also nach 17.2.18 eine Regelfunktion.
In 17.2.18 haben wir eine Intervalladditivität für das Integral für Treppenfunktionen bewiesen. Diese überträgt sich sofort auf das Integral für Regelfunktion. Dazu stellen wir fest: Ist $f : [a,b] \to \mathbb{R}$ eine Regelfunktion und $c \in [a,b]$ mit $a < c < b$, dann ist $f_1 := f|\ [a,b]$ und $f_2 := f|\ [a,b]$ ebenfalls eine Regelfunktion.
Sind umgekehrt $f_1 : [a,c] \to \mathbb{R}$ und $f_2 : [a,b] \to \mathbb{R}$ Regelfunktion, dann ist auch $f : [a,b] \to \mathbb{R}$ mit

$$f(x) := \begin{cases} f_1(x), & \text{für } x \in [a,c]; \\ f_2(x), & \text{für } x \in]c,b]. \end{cases}$$

eine Regelfunktion. Hieraus ergibt sich dann die Intervalladditivität auch für Regelfunktion.

18.2.22 Satz: Intervalladditivität des Regelintegrals

Ist $f : [a, b] \to \mathbb{R}$ eine Regelfunktion und gilt $a < c < b$, dann ist

$$\boxed{\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^d f(x)dx}$$

Beweis : Ist (t_n) eine Folge von Treppenfunktionen $t_n : [a, b] \to \mathbb{R}$ mit $t_n \rightrightarrows f$, dann gilt auch

$$t_n \| \; [a, b] \rightrightarrows f| \; [a, c] \; \text{ und } \; t_n| \; [c, b] \rightrightarrows f| \; [c, b].$$

Daher ist (weil für Treppenfunktionen die Intervalladditivität schon gezeigt ist)

$$\begin{aligned} \int_a^b f = \lim_{n\to\infty} \int_a^b t_n &= \lim_{n\to\infty} \left(\int_a^c t_n + \int_c^b t_n \right) \\ &= \lim_{n\to\infty} \int_a^c t_n + \lim_{n\to\infty} \int_c^b t_n \\ &= \int_a^c f + \int_c^b f. \end{aligned}$$

□

Um das Integral etwas flexibler handhaben zu können, wollen wir uns von der Generalvoraussetzung $a < b$ für das Integrationsintervall $[a, b]$ befreien und definieren $\int_a^a f = 0$ für jede Regelfunktion f und falls $b < a$ gilt

$$\int_a^b f = -\int_b^a f.$$

Mit diesen Festlegungen gilt dann

18.2.23 Satz

Sind $a, b, c \in \mathbb{R}$ und ist $f : [\min\{a, b, c\}, \max\{a, b, c\}] \to \mathbb{R}$ eine Regelfunktion, dann gilt bei beliebiger Lage von a, b, c

$$\boxed{\int_a^b f + \int_b^c f + \int_c^a f = 0}$$

Der einfache Beweis durch Fallunterscheidung ($3! = 6$ Fälle) sei den geneigten Lesern überlassen.

Zum Abschluss gehen wir auf zwei Sätze ein, die uns von Nutzen sein werden:

- Mittelwertsatz der Integralrechnung und
- ein „Verschwindungslemma“.

18.2.24 Theorem (Mittelwertsatz der Integralrechnung)

Sei $a, b \in \mathbb{R}$, $a < b$ und $f : [a,b] \to \mathbb{R}$ stetig und $p : [a,b] \to \mathbb{R}$ eine Regelfunktion mit $p \geq 0$ (das heißt $p(x) \geq 0$ für alle $x \in [a,b]$), dann gibt es ein $\xi \in [a,b]$ mit

$(*)$ $\int_a^b f(x)p(x)dx = f(\xi)\int_a^b p(x)dx$ (verallgemeinerter Mittelwertsatz).

Im Spezialfall $p(x) = 1$ für alle $x \in [a,b]$ folgt:

$(**)$ $\int_a^b f(x)dx = f(\xi)(b-a)$ mit $\xi \in [a,b]$ (spezieller Mittelwertsatz)

$\mu := \frac{\int_a^b f(x)dx}{b-a}$ nennt man auch *Integral-Mittelwert* (von f im Intervall $[a,b]$). Der spezielle MWS besagt also, dass der Integralmittelwert ein Funktionswert ist.

Beweis : Sei $m_1 = \min\{f(x);\ x \in [a,b]\}$ und $m_2 = \max\{f(x);\ x \in [a,b]\}$.
Nach ??? gilt $f \cdot p \in R([a,b])$ und aus $m_1 p(x) \leq f(x)p(x) \leq m_2 p(x)$ folgt nach ???

$$m_1 \int_a^b p(x)dx \leq \int_a^b f(x)p(x)dx \leq m_2 \int_a^b p(x)dx.$$

Ist nun $\int_a^b p(x)dx = 0$, dann ist auch

$$\int_a^b f(x)p(x)dx = 0$$

und die Gleichung $(*)$ ist für alle $\xi \in [a,b]$ richtig. Gilt jedoch $\int_a^b p(x)dx \neq 0$, so folgt

$$m_1 \leq \frac{\int_a^b f(x)p(x)dx}{\int_a^b p(x)dx} \leq m_2$$

Der Quotient in der Mitte liegt also im Intervall $[m_1, m_2]$, ist also ein Funktionswert von f, daher gibt es nach dem Zwischenwert Satz ein $\xi \in [a,b]$ mit

$$f(\xi) = \frac{\int_a^b f(x)p(x)dx}{\int_a^b p(x)dx},$$

d.h. es gilt

$$\int_a^b f(x)p(x)dx = f(\xi)\int_a^b p(x)dx.$$

$\square$

18.2.25 Beispiele und Bemerkungen

(a) Für Treppenfunktionen ist der MWS nicht richtig, denn für die Treppenfunktionen

$$t : [0,2] \to \mathbb{R} \text{ mit } t(x) = \begin{cases} 1, & \text{falls } 0 \leq x \leq 1; \\ 0, & \text{falls } 1 < x \leq 2. \end{cases}$$

gilt $\frac{1}{2}\int\limits_0^2 t(x)dx = \frac{1}{2} \neq f(\xi)$ für alle $\xi \in [a,b]$.

(b) Man braucht denn Zwischenwertsatz nur im Fall

$$m_1 < \frac{\int\limits_a^b f(x)p(x)dx}{\int\limits_a^b p(x)dx} < m_2,$$

dann kann man ξ sogar im $]a,b[$ wählen. Gilt etwa $m_1 = \frac{\int_a^b f(x)p(x)dx}{\int_a^b p(x)dx}$, dann gibt es ein $x_1 \in [a,b]$ mit $m_1 = f(x_1)$ und $(*)$ ist mit $\xi = x_1$ erfüllt.

Analog schließt man, falls $\frac{\int_a^b f(x)p(x)dx}{\int_a^b p(x)dx} = m_2$ gilt.

(c) Der verallgemeinern Mittelwertsatz macht die Aussage über *gewichtete Mittelwerte* einer stetigen Funktionen. Es gilt auch im Fall $p \leq 0$ (das heißt $p(x) \leq 0$ für alle $x \in [a,b]$), jedoch *nicht* für Regelfunktionen mit Vorzeichenwechsel in $[a,b]$.

Gegenbeispiel: $p(x) = x - \frac{1}{2}$ mit $\int\limits_0^1 p(x)dx = \frac{1}{2} - \frac{1}{2} = 0$ und $f(x) = x$ für $x \in [0,1]$.

Hier ist

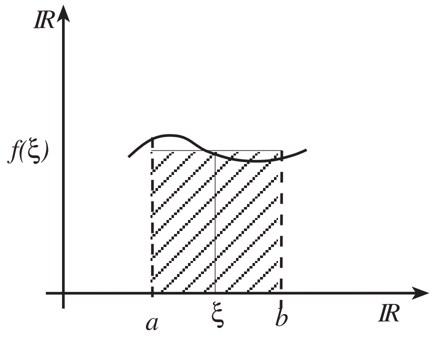
$$\int\limits_0^1 f(x)p(x)dx = \int\limits_0^1 x^2dx - \frac{1}{2}\int\limits_0^1 x\ dx = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \neq 0,$$

während $f(\xi)\int\limits_0^1 p(x)dx = f(\xi)\cdot 0 = 0$ für alle $\xi \in [0,1]$ gilt.

(d) Der spezieller Mittelwertsatz bedeutet geometrische interpretiert, wegen

$$I(f) = \int\limits_a^b f(x)dx = f(\xi)(b-a),$$

dass das Integral gleicht dem Flächeninhalt eines Rechtecks ist, dessen eine Seitenlänge $(b-a)$ und dessen anderer Seitenlänge ein Funktionswert ist:



Die Mittelwertsätze beruhen auf der Monotonie des Integrals. Die folgende Aussage, die wir häufig benutzen werden, beruht auch wesentlich auf der Monotonie des Integral:

18.2.26 Lemma („Verschwindungslemma“)

Ist $f : [a,b] \to \mathbb{R}$ eine Regelfunktion mit $f \geq 0$ und $\int\limits_a^b f(x)dx = 0$, dann ist $f(x_0) = 0$ an jeder Stetigkeitsstelle $x_0 \in [a,b]$. Insbesondere ist $f(x) = 0$ für fast alle $x \in [a,b]$, d.h. die Menge der $x \in [a,b]$ mit $f(x) \neq 0$ ist höchstens abzählbar.

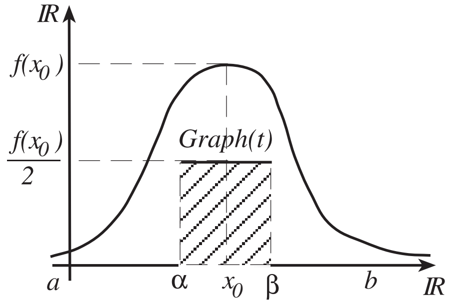
Beweis : Wir nehmen an $x_0 \in [a,b]$ sei eine Stetigkeitsstelle mit $f(x_0) > 0$ und $[\alpha,\beta] \subset [a,b]$ ein Intervall mit $x_0 \in [\alpha,\beta]$ und $f(x) > \frac{1}{2}f(x_0)$ für alle $x \in [\alpha,\beta]$ (ein solches Intervall existiert wegen der Stetigkeit von f in x_0). Dann gilt für die durch

$$t(x) = \begin{cases} \frac{1}{2}f(x_0), & \text{für } x \in [\alpha,\beta]; \\ 0, & \text{für } x \in [a,b] - [\alpha,\beta]; \end{cases}$$

definierte Treppenfunktion

$$\int\limits_a^b f(x)dx \geq \int\limits_a^b t(x)dx = (\beta-\alpha)\frac{f(x_0)}{2} > 0$$

im Widerspruch zur Voraussetzung $\int\limits_a^b f(x)dx = 0$.



Also gilt für jede stetig Stetigkeitsstelle x_0 unter den genannten Voraussetzungen $f(x_0) = 0$. Der Zusatz, dass für fast alle $x \in [a,b]$ auch $f(x) = 0$ gelten muss, folgt aus der Eigenschaft, dass eine Regelfunktion höhchstens abzählbar viele Unstetigkeitsstellen hat (vgl. 17.2.21). Ist f also stetig auf $[a,b]$, $f \geq 0$ und $I(f) = 0$, dann muss $f(x) = 0$ für alle $x \in [a,b]$ gelten, denn hier ist $S_f := \{x \in [a,b];\ x \text{ Sprungstelle von } f\} = \emptyset$.

$\square$

19 Grundlagen der Differentialrechnung

Der Ausgangspunkt für die Differentialrechnung war bei I.Newton (1643-1727) das Problem der Bestimmung von *Momentangeschwindigkeiten* bzw. *Momentanbeschleunigungen* für ungleichförmigen Bewegungen, bei G.W.Leibniz (1646-1716) stand das Problem der Bestimmung von *Tangenten* an zunächst ebenen Kurven (das heißt Kurven in $\mathbb{R}^2$) im Mittelpunkt des Interesses. Wer untersuchen dieses Problem zunächst für solche Kurven in $\mathbb{R}^2$, die als Graphen von Funktionen auftreten. Wer werde sehen, dass beide Probleme den selben mathematischen Kern haben und darauf hinaus laufen, gegebene Funktionen (hier die differenzierbaren) durch eine einfach Klasse von Funktionen (wenigstens lokal, d.h. in der Umgebung eines festen Punktes) „gut" zu approximieren. Im ersten Abschnitt beschäftigen wir uns mit dem *Begriff der Ableitung*, dann mit *Ableitungsregeln* (z.B. *Quotientenregel*, *Kettenregel*, *Differenzierbarkeit der Umkehrfunktion*). Dann werden wir sehen, wie die Differenzialrechnung geeignete Mittel bereitstellt, um zum Beispiel *lokale Extrema* differenzierbaren Funktionen zu ermitteln. Von großen theoretischen Interesse ist der *Mittelwertsatz der Differenzialrechnung* (abgekürzt MWSD). Wichtig sind die Anwendungen des Mittelwertsatzes: *Monotoniekriterium*, *Konvexität*, *Schrankensatz*, *Charakterisierung konstanter Funktionen* etc.
Wir werden dann auch sehen (siehe den Abschnitt über den Hauptsatz der Differenzial-und Integral Rechnung), dass Differentiation und Integration *Umkehroperationen* voneinander sind.

19.1 Der Begriff der Ableitung, erste Beispiele differenzierbaren Funktionen

19.1.1 Definition: Begriff der Differenzierbarkeit

Sei (zunächst) $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$ eine Funktion.
f heißt *differenzierbar* (ableitbar) im Punkt $a \in D$, falls der Grenzwert

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

existiert, der dann eindeutig bestimmt ist. $f'(a)$ heißt dann die *Ableitung* oder *Differentialquotient* von f in a.

Schreibweise: $f'(a) = \mathsf{D}f(a) = \frac{df}{dx}(a) = \frac{d}{dx}f(a) = \dot{f}(a)$ oder ähnlich.

Die Funktion $\Delta f(a, \cdot) : D - \{a\} \to \mathbb{R}$ mit $\Delta f(a;x) = \frac{f(x)-f(a)}{x-a}$ heißt *Differenzenquotient* von f in a.
Differenzierbarkeit von f in a bedeutet also, dass

$$\lim_{x \to a} \Delta f(a;x) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

existiert.
f heißt *differenzierbar* in D, wenn f in jedem Punkt $x \in D$ differenzierbar ist.
Die Funktion

$$\begin{array}{rccc} f' : & D & \to & \mathbb{R} \\ & x & \mapsto & f'(x) \end{array}$$

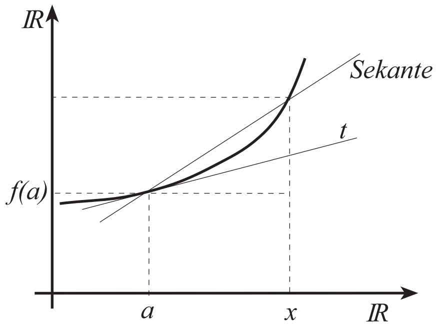
heißt dann Ableitungsfunktion von f.
Schreibweisen: $f' = \mathsf{D}f = \frac{d}{dx}f = \dot{f}$ oder ähnlich.

19.1.2 Geometrische Interpretation

Ist $f : D \to \mathbb{R}$ in $a \in D$ differenzierbar, so kann man den Differenzenquotient

$$\Delta f(a;x) = \frac{f(x) - f(a)}{x - a} \quad (x \neq a)$$

interpretieren als Steigerung der Geraden durch die Punkte $(a, f(a))$ und $(x, f(x))$ („Sekante")



Ist f differenzierbaren in a, so konvergiert $\Delta f(a;x)$ für $x \to a$ gegen $f'(a)$. Die durch die Gleichung $y = f(a) + f'(a)(x - a)$ für $x \in \mathbb{R}$ definierte Gerade t heißt *Tangente an den Graphen* von f im Punkt $(a, f(a))$.

19.1.3 Physikalische Interpretation

Ist $s(t)$ der Ort eines sich auf einer Geraden bewegten Massenpunktes zur Zeit („tempus") $t \in \mathbb{R}$, dann ist (für $t \neq t_0$) $\Delta s(t;t_0) = \frac{s(t)-s(t_0)}{t-t_0}$ die *mittlere Geschwindigkeit* (Durchschnittsgeschwindigkeit) des Massenpunktes im Zeit Intervall $[t, t_0]$. Existiert die Ableitung von s in t_0, existiert also

$$\dot{s}(t) := \lim_{t \to t_0} \frac{s(t) - s(t_0)}{t - t_0},$$

so kann man $\dot{s}(t_0)$ als momentane Geschwindigkeit des Massenpunktes zum Zeitpunkt t_0 interpretieren. Auch wenn das Weg-Zeit-Gesetz $t \to s(t)$ nicht durch eine Gerade beschrieben wird, kann man $\dot{s}(t_0)$ als Momentangeschwindigkeit interpretieren. In der Physik wird die schon auf Newton zurückgehende Notation „$\dot{} = \frac{d}{dt}$" für Ableitung nach der Variablen t (Zeit) verwendet.

19.1.4 Ableitung komplexwertiger Funktionen

Für Funktionen $f : D \to \mathbb{C}$ kann man denn Begriff der Differenzierbarkeit durch den limes des Differenzenquotienten wie in 18.1.1 erklären ($\mathbb{C}$ ist ja ein Körper!). Es gilt: Ist $f = u + iv$, $u = Re\ f$, $v = Im\ f$, dann ist f genau dann in $a \in D$ differenzierbar, wenn die reellwertigen Funktionen u und v in a differenzierbar sind, und es gilt dann

$$f'(a) = u'(a) + iv'(a).$$

19.1.5 Erste Beispiele

(a) Jede konstanter Funktionen $f : D \to \mathbb{C}$ ist differenzierbar, und es gilt $f'(x) = 0$ für alle $z \in D$. Denn ist etwa $f(x) = c$, $c \in \mathbb{C}$ ($c \in \mathbb{R}$ ist natürlich zulässig), dann ist schon ($x \neq c$) $\Delta f(a;x) = \frac{0-0}{x-a} = 0$, also auch

$$f'(a) = \lim_{x \to a} \Delta f(a;x) = 0.$$

(b)

$$\begin{array}{rcl} p_n : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^n \quad (n \in \mathbb{N}) \end{array}$$

ist differenzierbar und es gilt

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

Für $x \neq a$ ist

$$\Delta p_n(a;x) = \frac{x^n - a^n}{x-a} = x^{n-1} + x^{n-2}a + \ldots + a^{n-1}$$

und damit $\lim\limits_{x \to a} \Delta p_n(a,x) = na^{n-1}$.

(c) Für $\alpha \in \mathbb{C}$, ist

$$\begin{array}{rcl} f : \mathbb{R} & \to & \mathbb{C} \\ x & \mapsto & \exp(\alpha x) \end{array}$$

differenzierbar, und es gilt für alle $x \in \mathbb{R}$

$$f'(x) = \alpha \exp(\alpha x) = \alpha f(x).$$

Es ist nämlich für $x \neq a$

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

und daher

$$\lim_{x \to a} \frac{\exp(\alpha x) - \exp(\alpha a)}{x-a} = \alpha \exp(\alpha a).$$

Insbesondere ist für $\alpha = 1$

$$exp'(x) = exp(x)$$

für alle $x \in \mathbb{R}$ und für $\alpha = i$ gilt mit $f(x) = \exp(ix)$

$$f'(x) = i \exp(ix) = if(x).$$

(d) Die Betragsfunktion $f = abs = |\ | : \mathbb{R} \to \mathbb{R}$ ist in $a = 0$ *nicht* differenzierbar, denn es ist für $x \neq 0$

$$\Delta f(0,x) = \frac{|x| - 0}{x - 0} = \begin{cases} 1, & \text{falls } x > 0; \\ -1, & \text{falls } x < 0; \end{cases}$$

und dieser Differenzenquotient hat keinen Grenzwert für $x \to 0$ (für die Folge $x_n = \frac{1}{n}$ gilt $\lim\limits_{n \to \infty} \Delta f(0;x_n) = 1 \neq -1 = \lim\limits_{n \to \infty} \Delta f(0,y_n)$, $y_n = -\frac{1}{n}$).
Betrachtet man jedoch die Einschränkung g von $f = abs$ auf $\mathbb{R}_+ = \{x \in \mathbb{R}, x \geq 0\}$, dann existiert

$$g'(0) = \lim_{x \to 0} \frac{g(x) - g(0)}{x - 0} = \lim_{x \to 0} \frac{x}{x} = 1.$$

Man definierte $f_+'(0) := g'(0)$ und nennt $f_+'(0)$ die rechtsseitige Ableitung von f im Nullpunkt. Entsprechend ist die linksseitige Ableitung von f im Nullpunkt -1.

19.1.6 Feststellung

Dass der Grenzwert $f'(a) := \lim\limits_{x\to a} \frac{f(x)-f(a)}{x-a}$ in 18.1.1 existiert ist nach der Definition des Grenzwertes äquivalent mit der Fortsetzbarkeit des Differentialquotienten $\Delta f(a;\ ) : D \to \{a\} \to \mathbb{R}$ einer auf ganz D definierte Funktion, die in a stetig ist. Nennen wir die in a stetige Fortsetzung der Einfachheit halber φ (sie hängt von a und f ab), damit ist also $\varphi : D \to \mathbb{R}$ mit

$$\varphi(x) = \begin{cases} \frac{f(x)-f(a)}{x-a}, & \text{für } x \neq a \\ f'(a), & \text{für } x = a \end{cases}$$

stetig in a.
Für $x \neq a$ kann man $f(x) - f(a) = (x-a)\varphi(x)$ schreiben. Diese Gleichung gilt auch für $x = a$. Damit haben wir:
$f : D \to \mathbb{R}$ ist genau dann differenzierbar eben $a \in D$, wenn es eine Funktion $\varphi : D \to \mathbb{R}$ gibt, die in a stetig ist und für die

$$(*) \qquad f(x) = f(a) + (x-a)\varphi(x)$$

für alle $x \in D$ und $\varphi(a) = f'(a)$ gilt. Setzt man $\varrho(x) := \varphi(x) - f'(x)$ für $x \in D$ und setzt diese in $(*)$ ein, so erhält man

$$(**) \qquad f(x) = f(a) + (x-a)f'(a) + (x-a)\varrho(x),$$

dabei ist $g : D \to \mathbb{R}$ stetig in a und $g(a) = 0$. Definiert man $r(x) := (x-a)\varrho(x)$ für $x \in D$, dann ist $(**)$ äquivalent zu

$$(***) \qquad f(x) = f(a) + (x-a)f'(a) + r(x)$$

mit

$$\lim_{x\to a} \frac{r(x)}{x-a} = \lim_{x\to a} g(x) = g(a) = 0.$$

Damit haben wir die meisten Teile des folgenden Äquivalenssatzes für Differenzierbarkeit:

19.1.7 Theorem (Äquivalentssatz für Differenzierbarkeit)

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f : D \to \mathbb{R}$ eine Funktion, $a \in D$. Dann sind folgende Aussagen äquivalent:

(1) f ist in a differenzierbar

(2) Es gibt eine in a stetige Funktionen $\varphi : D \to \mathbb{R}$ mit $f(x) = f(a) + (x-a)\varphi(x)$ für alle $x \in D$

(3) Es gibt eine Zahl $l \in \mathbb{R}$ und eine in a stetige Funktionen $g : D \to \mathbb{R}$ mit $g(a) = 0$ und

$$f(x) = f(a) + l(x-a) + (x-a)g(x)$$

(4) Es gibt eine Zahl $l \in \mathbb{R}$, so dass die durch die Gleichung

$$f(x) = f(a) + l(x-a) + r(x)$$

definierte Funktion („Rest“) $r : D \to \mathbb{R}$ die Bedingung $\lim\limits_{x\to a} \frac{r(x)}{x-a} = 0$ erfüllt.

(5) Es gibt eine lineare Abbildung $L : \mathbb{R} \to \mathbb{R}$ mit $\lim\limits_{x\to a} \frac{f(x)-f(a)-L(x-a)}{x-a} = 0$

Falls die fünf Eigenschaften zu treffen, so sind die Funktionen φ, g, r und die linear Abbildung L und die Konstante $l \in \mathbb{R}$ eindeutig bestimmt, und es gilt $l = \varphi(a) = f'(a) = L(1)$.
Zum **Beweis** zeigen wir noch $(4) \Longrightarrow (5)$ und $(5) \Longrightarrow (1)$, da wir in der Vorüberlegung schon die paarweise Äquivalenz von (1) bis (4) gezeigt haben.

Beachte: Jede lineare Abbildung $L:\mathbb{R}\to\mathbb{R}$ hat die Gestalt $h\mapsto lh$ mit $l:=L(1)\in\mathbb{R}$.
Ist nun f in a differenzierbar, so gilt etwa (4) und wenn man $L(h):=lh:=f'(a)h$ definiert hat man eine lineare Abbildung $L:\mathbb{R}\to\mathbb{R}$ gefunden, die

$$\lim_{x\to a}\frac{f(x)-f(a)-L(x-a)}{x-a}=0$$

erfüllt.
Ist umgekehrt $L:\mathbb{R}\to\mathbb{R}$ eine lineare Abbildung, wie die die Gleichung in (5) erfüllt und gilt etwa $L(h)=lh$ mit einem $l\in\mathbb{R}$, so folgt

$$0=\lim_{x\to a}\frac{f(x)-f(a)-l(x-a)}{x-a}=\lim_{x\to a}\frac{f(x)-f(a)}{x-a}-l,$$

das bedeutet, dass f in a differenzierbar ist und $l=f'(a)$ gelten muss.
Da Äquivalenz der Eigenschaft (4) oder (5) mit der Differenzierbarkeit stellt den Aspekt der „linearen Approximierbarkeit“ in den Vordergrund. Die lineare Abbildung $L:\mathbb{R}\to\mathbb{R}$ mit $h\mapsto f'(a)h$ nennt man auch das Differenzial (oder Linearisierung) von f im Punkt a und bezeichnet es mit d_af oder $df(a)$. Nach Definition gilt also

$$d_af(h)=df(a)(h)=f'(a)h.$$

Die linear-affine Funktion (Tangente)

$$\begin{aligned} t:\mathbb{R} &\to \mathbb{R}\\ x &\mapsto f(a)+L(x-a)=f(a)+f'(a)(x-a)\end{aligned}$$

heißt auch *lineare Approximation* von f in a.

19.1.8 Feststellung

Ist f differenzierbar in $a\in D$, dann ist f auch stetig in a.

Das folgt unmittelbar aus 18.1.7(2) oder auch einfach aus

$$f(x)-f(a)=\frac{f(x)-f(a)}{x-a}(x-a)\quad\text{(für } x\neq a),$$

da die rechte Seite für $x\to a$ dem Grenzwert Null hat. Dass die Umkehrung hier von im Allgemeinen nicht gilt, seit das einfache Beispiel der Betragsfunktion im Punkt $a=0$.
Das folgende Beispiel demonstriert aber, dass aus der Differenzierbarkeit von f in a nichts über die Stetigkeit in einer Umgebung von a zu folgen braucht: die Funktion

$$f:\mathbb{R}\to\mathbb{R}\ \text{ mit }\ f(x)=\begin{cases} x, & \text{falls } x\in\mathbb{Q},\\ x+x^2, & \text{falls } x\notin\mathbb{Q},\end{cases}$$

ist an der Stelle $a=0$ differenzierbar und hat dort die Ableitung 1. Für alle Punkte $a\neq 0$ ist f nicht stetig und damit auch erst recht nicht differenzierbar. Für $a=0$ und $x\neq 0$ ist der Differenzenquotient

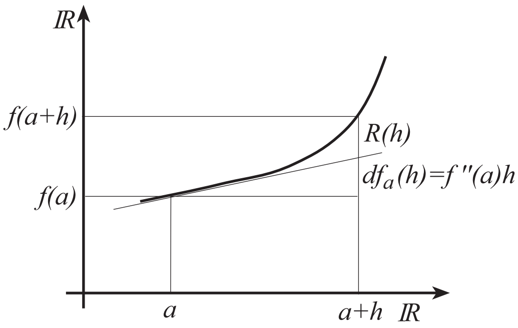
$$\frac{f(x)-f(0)}{x-0}=\begin{cases} 1, & \text{falls } x\in\mathbb{Q};\\ 1+x, & \text{falls } x\notin\mathbb{Q}.\end{cases}$$

Die Funktion φ mit

$$\varphi(x)=\begin{cases} 1, & \text{falls } x\in\mathbb{Q}\\ 1+x, & \text{falls } x\notin\mathbb{Q}\end{cases}\qquad (x\neq 0)$$

ist aber durch $\varphi(0):=1$ stetig nach Null fortsetzbar, also ist $f'(0)=1$.

19.1.9 Geometrische Interpretation des Differenzials



setzt man $x = a + h$, dann wird

$$f(a+h) - f(a) = L(h) + R(h).$$

Die Änderungsrate der Funktion beim Übergang von Argument a Argument $a + h$ wird also beschrieben durch eine lineare Funktion L und einen Rest R, der mit $h \to 0$ schneller gegen Null geht als h selber:

$$\lim_{h \to 0} \frac{R(h)}{h} = 0.$$

Ist $f : D \to \mathbb{R}$ in jedem Punkt $a \in D$ differenzierbar, dann besitzt also f in jedem $a \in D$ ein Differenzial:

$$\begin{aligned} d_a f : \mathbb{R} &\to \mathbb{R} \\ h &\mapsto d_a f(h) := f(a)h. \end{aligned}$$

Beachtet man speziell die Identität eingeschränkt auf D, dann gilt also $id(x) = x$ für alle $x \in D$ und $id'(x) = 1$ für alle $x \in D$. Das Differenzial der Identität

$$\begin{aligned} d\, id_a : \mathbb{R} &\to \mathbb{R} \\ h &\mapsto d\, id_a(h) = 1 \cdot h = h, a \in D, \end{aligned}$$

ist also von a unabhängig und wirkt als Identität.
Wegen $id(x) = x$ schreibt man für denn Differential der Identität einfach dx.
Wohlgemerkt dx ist die $\mathbb{R}$-lineare Abbildung

$$\begin{aligned} dx : \mathbb{R} &\to \mathbb{R} \\ h &\mapsto dx(h) = 1 \cdot h = h. \end{aligned}$$

Wie jede linear Abbildung $L : \mathbb{R} \to \mathbb{R}$ ein Vielfaches der identische Abbildung ist $(L = L(1) \cdot id_{\mathbb{R}})$, ist auch das Differential einer differenzierbaren Funktionen ein skalares vielfaches von dx:

$$\boxed{df = f' dx}$$

Das bedeutet, dass für alle $a \in D$ und alle $h \in \mathbb{R}$ gilt

$$d_a f(h) = f'(a) dx(h) = f'(a)h.$$

Betrachtet man die konstante Abbildung $h \mapsto f'(a)$, dann kann sie als Quotient der Differentiale (linear Funktionen) df_a und dx auffassen, denn für $h (\neq 0)$ gilt

$$\frac{d_a f}{dx}(h) = \frac{d_a f(h)}{dx(h)} = \frac{f'(a)h}{h} = f'(a).$$

Im eindimensionalen $\mathbb{R}$-Vektorraum $\mathbb{R}$ unterscheidet man in der Regel nicht zwischen linear Abbildungen $L : \mathbb{R} \to \mathbb{R}$ und Zahlen: jede linear Abbildung $L : \mathbb{R} \to \mathbb{R}$ hat die Gestalt $L(h) = L(1)h = lh$

mit einer eindeutig bestimmten Zahl $l(= L(1)) \in \mathbb{R}$. Umgekehrt bestimmt jede reelle Zahlen l eine linear Abfindung $L : \mathbb{R} \to \mathbb{R}$, nämlich die Linksmultiplikation mit l:

$$\begin{array}{rcl} L : \mathbb{R} & \to & \mathbb{R} \\ h & \mapsto & lh. \end{array}$$

Algebraisch bedeutet das:

$$\boxed{Hom(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}}$$

Erst in der Differenzialrechnung mehrerer Veränderlichen, wird die Bedeutung der Differenziale voll zum Tragen kommen.

19.1.10 Definition (höhere Ableitungen)

Ist $f : D \to \mathbb{R}$ differenzierbar ($D \subset \mathbb{R}$ echtes Intervall) und $f' : D \to \mathbb{R}$ die Ableitung von f. Ist f' wiederum differenzierbar, dann heißt f zweimal differenzierbar und man setzt $f'' := (f')'$ und nennt f'' die *zweite* Ableitung von f.
Allgemein definiert man rekursiv höhere Ableitungen von f:
$f : D \to \mathbb{R}$ heißt k-mal differenzierbar ($k \in \mathbb{N}$), falls $f^{(k-1)} : D \to \mathbb{R}$ differenzierbar ist und man setzt dann

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

Zur Ergänzung setzt man noch $f^{(0)} := f$. *Jede* Funktion ist als Null-mal differenzierbar.

f heißt im Punkt $a \in D$ k-mal differenzierbar, $k \in \mathbb{N}$, $k \geq 1$, falls es ein $r > 0$ gibt, so dass f in $U := U_r(a) \cap D$ $(k-1)$-mal differenzierbar ist und $f^{(k-1)} : U \to \mathbb{R}$ in a differenzierbar ist. Eine Funktion $f : D \to \mathbb{R}$ ist also genau dann k-mal differenzierbar, wenn sie jedem Punkt $a \in D$ k-mal differenzierbar ist.
Besitzt $f : D \to \mathbb{R}$ eine k-te Ableitung $f^{(k)} : D \to \mathbb{R}$, dann sind alle

$$f^{(0)} = f, f' = f =^{(1)}, \ldots, f^{(k-1)}$$

stetig. Für $f^{(k)}$ braucht dies jedoch nicht zu zutreffen. Ist jedoch auch $f^{(k)}$ stetig, dann heißt f *k-mal stetig differenzierbar*.

Standardbezeichnung: $C^k(D) := \{f : D \to \mathbb{R};\ fk\text{-mal stetig diffferenzierbar}\}$; Alle $C^k(D)$ sind $\mathbb{R}$-Vektorräume, und es gelten die Inklusionen

$$C(D) := C^0(D) \supset C^1(D) \supset \ldots \supset C^k(D) \supset \ldots, \qquad k \in \mathbb{N}_0$$

wobei die Inklusionen echt sind. Existiert $f^{(k)}$ für *jedes* $k \in \mathbb{N}_0$, so heißt f beliebig oft (stetig) differenzierbar und man setzt $C^\infty(D) := \bigcap\limits_{k=0}^{\infty} C^k(D)$. $C^\infty(D)$ ist ebenfalls ein $\mathbb{R}$-Vektorraum. Es gilt etwa $\exp \in C^\infty(\mathbb{R})$.

Vom physikalischen Standpunkt besonders wichtig ist wegen der Newtonschen Bewegungsgesetze die zweite Ableitung.

19.1.11 Bemerkung

Die Definition der k-maligen Differenzierbarkeit in einem Punkt scheint gekünstelt, wir haben sie deshalb gewählt, weil wir auch Funktionen und ihre Ableitungen betrachten wollen, wenn die Definitions Bereiche keinen Intervalle sind. Die in der Praxis vorkommenden Definitionsbereiche sind aber meist endliche oder abzählbar unendliche Vereinigungen von (echte) Intervallen: $D = \bigcup\limits_{j \in J} D_j$, J endliche oder abzählbar unendliche Indexmenge.

So ist zum Beispiel für

$$\begin{aligned} j : D := \mathbb{R} - \{0\} &\to \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned}$$

$D = D_1 \cup D_2$ mit $D_1 := \{x \in \mathbb{R};\ x > 0\}$ und $D_2 := \{x \in \mathbb{R};\ x < 0\}$ und für den Definitionsbereich des Tangens gilt

$$D = Def(\tan) = \bigcup_{j \in \mathbb{Z}} \left] \frac{\pi}{2}(2j-1);\ \frac{\pi}{2}(2j+1) \right[= \left\{ x \in \mathbb{R};\ x \neq \frac{\pi}{2}(2k+1),\ k \in \mathbb{R} \right\}.$$

Eine Funktion $f : D = \bigcup\limits_{j \in \mathbb{Z}} D_j \to \mathbb{R}$ wollen wir differenzierbar nennen, wenn $f|\ D_j$ differenzierbar ist ($j \in \mathbb{Z}$). Wie wir im nächsten Abschnitt sehen werden, sind zum Beispiel *rationale Funktionen* in ihrem Definitionsbereich (der die Vereinigung von endlich viele Intervallen ist) differenzierbar.
Allgemein benötigt man zur Definition der Ableitung einer Funktion f im Punkt $a \in D$ ($D \neq \emptyset$ beliebig) lediglich die Eigenschaft, dass „hinreichend nahe bei a“ Punkte $x \in D$ geben muss, die von a verschieden sind. Das ist sicherlich dann gegeben, wenn $a \in D$ ein Häufungspunkt von D ist. Ist D ein (echtes) Intervall, dann ist jeder Punkt $a \in D$ auch Häufungspunkt von D, die obige Bedingung also automatisch erfüllt.
Die Differenzierbarkeit von f in einem Punkt $a \in D$ ist eine *lokale Eigenschaft*, das heißt f ist genau dann differenzierbar in a, wenn es ein $r > 0$ gibt, so dass $f|\ D \cap U_r(a)$ in a differenzierbar ist. Man beweise diese Feststellung.

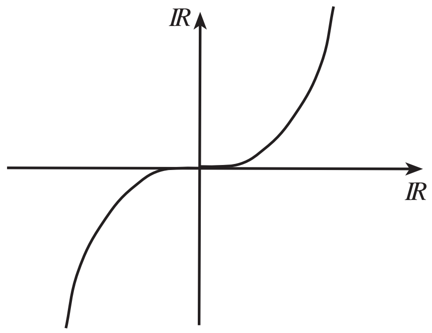
19.1.12 Abschreckende Beispiele

(a) Die Ableitung einer differenzierbaren Funktion $f : D \to \mathbb{R}$ muss nicht notwendig selbst wieder differenzierbar sein. Ein einfaches Beispiel ist

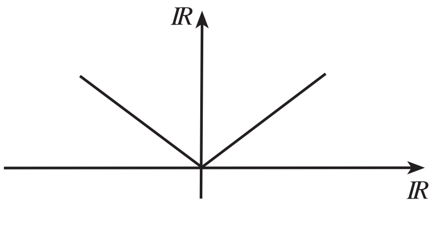
$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \frac{1}{2}|x|x, \end{aligned}$$

es ist also

$$f(x) = \begin{cases} \frac{1}{2}x^2, & \text{falls } x \geq 0 \\ -\frac{1}{2}x^2, & \text{falls } x \leq 0. \end{cases}$$



Graph(f)



Graph(f ')

Offensichtlich ist $f'(x) = |x|$, f' ist aber in Null nicht differenzierbar.
Nach dem gleichen Prinzip kann man Funktionen konstruieren, die in einem Punkt k-mal, aber nicht $(k+1)$-mal differenzierbar sind.

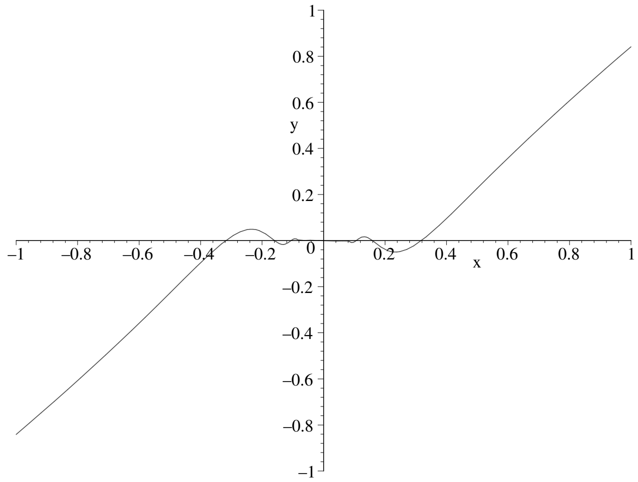


Abbildung 18: Graph von $f(x) = x^2 \sin \frac{1}{x}$

(b) Die Ableitung einer differenzierbaren Funktion $f : D \to \mathbb{R}$ braucht auch nicht stetig zu sein, wie das Beispiel

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \begin{cases} x^2 \sin \frac{1}{x}, & \text{falls } x \neq 0 \\ 0, & \text{falls } x = 0 \end{cases} \end{aligned}$$

zeigt. (siehe Abb. 18)

Hier ist (wie man sich leicht überzeugt)

$$f'(x) = \begin{cases} 2x \sin(\frac{1}{x}), & \text{falls } x \neq 0 \\ 0, & \text{falls } x = 0 \end{cases}$$

aber f' ist in $a = 0$ stetig.

(c) Es gibt Beispiele für stetige Funktionen $f : \mathbb{R} \to \mathbb{R}$, die in *keinem* Punkt differenzierbar sind. ein solches Beispiel ist etwa $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \sum \ldots \sum\limits_{n=0}^{\infty} \frac{\{10^n x\}}{10^n}$, wo $\{y\}$ der kleinste Abstand von y zur nächsten ganzen Zahl ist.

19.2 Die Technik des Differenzierens

Um mit differenzierbaren Funktionen erfolgreich arbeiten zu können, stellen wir in diesem Abschnitt mit den *Ableitungsregeln* das entsprechende Handwerkszeug zur Verfügung. Wir teilen die Regel in drei Gruppen ein:

I. die algebraischen Regeln (18.2.1)

II. die Kettenregel (18.2.3)

III. der Satz über die Differentiation der Umkehrfunktion (18.2.5)

Die Kombination dieser Regel wird es uns ermöglichen, die Ableitungen auch komplizierter Funktionen zu bestimmen. Die Definitionsbereiche der betrachteten Funktionen sind in der Regel echte Intervalle.

19.2.1 Satz (algebraische Regeln)

Sei $D \subset \mathbb{R}$ ein echtes Intervall, $f, g : D \to \mathbb{R}$ Funktionen, die in $a \in D$ differenzierbar sind. So gilt dies auch für jede ihrer Linearkombinationen $\lambda f + \mu g$ $(\lambda, \mu \in \mathbb{R})$, ihr Produkt fg und falls $g(a) \neq 0$ gilt auch für den Quotienten $\frac{f}{g}$ und zwar gilt dann

(a) $(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$ (Linearität der Summenregel)
(b) $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$ (Produktregel)
(c) $\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}$ (Quotientenregel)

Zusatz:
Sind f und g in a k-mal differenzierbar ($n \in \mathbb{N}$), dann sind auch $\lambda f + \mu g$ $(\lambda, \mu \in \mathbb{R})$ und fg in a k-mal differenzierbar, und es gilt

$(d)(\lambda f + \mu g)^{(n)} = \lambda f^{(n)}(a) + \mu g^{(n)}(a)$ (Linearität)
$(e)(fg)^{(n)}(a) = \sum\limits_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} f^{(n-k)}(a) g^{(n)}(a) = \sum\limits_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} f^{(k)}(a) g^{(n-k)}(a)$ (Leibniz'sche Regel)

speziell also für $n = 2$ und $n = 3$

$$(fg)''(a) = f''g + f'g' + f'g' + fg'' = f''g + 2f'g' + fg''$$
$$(fg)'''(a) = f'''g + 3f''g' + 3f'g'' + g'''.$$

Zum Beweis von (a), (b), (c) benutzt man am besten die Äquivalenz von (1) und (2) aus dem Äquivalenzsatz:
Wegen der Differenzierbarkeit von f in a gibt es eine im a stetige Funktionen $\varphi : D \to \mathbb{R}$ mit

$$(*) \qquad f(x) = f(a) + (x-a)\varphi(x) \text{ und } \varphi(a) = f'(a).$$

Analog gibt es wegen der Differenzierbarkeit von f in a eine in a stetige Funktionen $\psi : D \to \mathbb{R}$ mit

$$(**) \qquad f(x) = g(a) + (x-a)\psi(x) \text{ und } \psi(a) = g'(a).$$

Dann gilt aber

$$\begin{aligned}
(\lambda f + \mu g)(x) &= \lambda f(x) + \mu g(x) \\
&= \lambda f(a) + (x-a)\lambda\varphi(x) + \mu g(a) + (x-a)\mu\psi(x) \\
&= \lambda f(a) + \mu g(a) + (x-a)(\underbrace{\lambda\varphi(x) + \mu\psi(x)}_{=\rho(3x)}) \\
&= (\lambda f + \mu g)(a) + (x-a)\rho(3x),
\end{aligned}$$

da g in a differenzierbar ist, und $\rho(a) = \lambda\varphi(a) + \mu\psi(a) = \lambda f'(a) + \mu g'(a)$ gilt, folgt also $(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$. Damit ist (a) bewiesen.
Aus $(*)$ und $(**)$ folgt für das Produkt

$$(fg)(x) = f(a)g(a) + (\varphi(x)g(a) + f(a)\psi(x) + \varphi(x)\psi(x)(x-a))(x-a)$$

Da als Faktor von $(x-a)$ stehende Funktion ist also Zusammensetzung in a stetiger Funktionen stetig in a und ihr Wert in a ist

$$\varphi(a)g(a) + f(a)\psi(a) = f'(a)g(a) + f(a)g'(a),$$

das beweist (b).
Für die Quotientenregel (c) beweisen wir zunächst den Spezialfall

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

der allgemeine Fall folgt dann wegen $\frac{f}{g} = f \cdot \frac{1}{g}$ aus der Produktregel.
Wegen der Stetigkeit von g gibt es eine Umgebung $U_r(a)$ mit $g(x) \neq 0$ für alle $x \in D \cap U_r(a)$. Für dieses x gilt dann

$$\frac{1}{g(x)} - \frac{1}{g(a)} = -\frac{g(x) - g(a)}{g(a)g(x)} = \left(-\frac{\psi(x)}{g(a)g(x)}\right)(x - a).$$

Dies als Faktor $(x - a)$ stehende Funktion ist dann (wegen der Stetigkeit von f und ψ in a) auch stetig in a, und ihr Wert in a ist

$$-\frac{g'(a)}{(g(a))^2}.$$

Das beweist den Spezialfall der Quotientenregel. Der allgemeine Fall ergibt sich -wie schon bemerkt- durch Anwendung der Produktregel an

$$\frac{f}{g} = f \cdot \frac{1}{g}.$$

Die Regel (d) ist evident.
Die allgemeine Leibniz'sche Regel (e) für die n-te Ableitung eines Produkts ergibt sich durch Induktion nach n in völliger Analogie zum Beweis des Binomischen Satzes (vgl. §2.4.1).

□

19.2.2 Beispiele und Bemerkungen

(a) Für $n \in \mathbb{N}$ ergibt sich nochmals als Ableitung für die Potenzfunktion

$$\begin{aligned} p_n : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto x^n \end{aligned}$$

$p'_n(x) = nx^{n-1} = np_{n-1}(x)$ aus $p_{n+1} = p_1 p_n$ durch vollständige Induktion nach n unter Verwendung der Produktregel.

(b) Für

$$\begin{aligned} f : \mathbb{R}_+^* &\rightarrow \mathbb{R} \\ x &\mapsto \frac{1}{x^n} \quad (n \in \mathbb{N}) \end{aligned}$$

liefert die Quotientenregel

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

Diese Regel gilt auch für $x < 0$.
Zusammengefasst: Es gilt für alle $n \in \mathbb{Z}$

$$\mathrm{D}x^n = \frac{d}{dx}(x^n) = nx^{n-1}$$

und alle $x \in \mathbb{R}$ mit $x \geq 0$ bzw. alle $x \in \mathbb{R}$ mit $x < 0$

(c)

$$\begin{aligned} tan : \left]-\frac{\pi}{2}; \frac{\pi}{2}\right[&\rightarrow \mathbb{R} \\ x &\mapsto \tan x = \frac{\sin x}{cosx} \end{aligned}$$

ist differenzierbar, und es gilt nach der Quotientenregel

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

In jedem Intervall $(k \in \mathbb{Z})\ \left](2k-1)\frac{\pi}{2}, (2k+1)\frac{\pi}{2}\right[$ hat den Tangens diese Ableitung.

(d) Aus 18.2.2 folgt insbesondere, dass jedes Polynom

$$\begin{aligned} p : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto a_0 + a_1 x + \ldots + a_n x^n,\ a_j \in \mathbb{R}, 0 \leq j \leq n \end{aligned}$$

auf ganz $\mathbb{R}$ differenzierbar ist, und dass gilt

$$p'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1}.$$

Beim (einmaligen) Differenzieren erniedrigt sich also der Grad eines Polynoms um 1. Es ist $a_0 = p(0)$ und $a_1 = p'(0)$ und allgemein $a_j := \frac{p^{(j)}(0)}{j!},\ 0 \leq j \leq n$.

Die Koeffizienten a_j sind also durch p eindeutig bestimmt:

$$\boxed{a_j = \frac{p^{(j)}(0)}{j!}} \qquad 0 \leq j \leq n.$$

Offensichtlich gilt $p^{(n+1)}(x) = 0$ für alle $x \in \mathbb{R}$ und damit auch $p^{(k)}(x) = 0$ für alle $k \geq n+1$. Offensichtlich gilt damit für jedes Polynom $p \in C^\infty(\mathbb{R})$.

Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ kann damit kein Polynom sein, für sie gilt $\exp \in C^\infty(\mathbb{R})$, aber $\exp^{(k)}(x) \neq 0$ für alle $k \in \mathbb{N}_0$ und alle $x \in \mathbb{R}$.

(e) Gilt $f, g \in C^{(n)}(D)$, dann ist nach (c) auch $fg \in C^{(n)}(D)$, $C^{(n)}(D)$ ist daher sogar eine Funktionenalgebra.

(f) Eine rationale Funktion $f = \frac{P}{Q}$, P, Q Polynome (oBdA Grad Q minimal) ist in ihrem natürlichen Definitionsbereich $D = \{x \in \mathbb{R};\ Q(x) \neq 0\}$ nach unserer Vereinbarung differenzierbar. D ist die Vereinigung von endlich vielen Intervallen, welche durch die endlich vielen Nullstellen von Q bestimmt sind.

(g) Beispiel:

$$f(x) = \frac{1}{1-x^2} = \frac{1}{Q(x)}$$

Die Nullstellen des Nenners Q sind 1 und -1. Der Definitionsbereich von f ist daher

$$D :=]-\infty, -1[\ \cup\ ]-1, 1[\ \cup\ ]1, \infty[,$$

also eine Vereinigung von Intervallen. In jedem Teilintervall ist f differenzierbar und hat dort die Ableitung

$$f'(x) = \frac{2x}{(1-x)^2},$$

$f'(x)$ ist also wieder eine rationale Funktion. (siehe Abb. 19)

(h) Sind f und g in allen Punkten $x \in D$ differenzierbar, so besagen die Regeln aus 18.2.2 insbesondere, dass das Differentialquotient

$$\begin{aligned} \mathsf{D} : V &\to \mathit{Abb}(D, \mathbb{R}) \\ f &\mapsto f', \end{aligned}$$

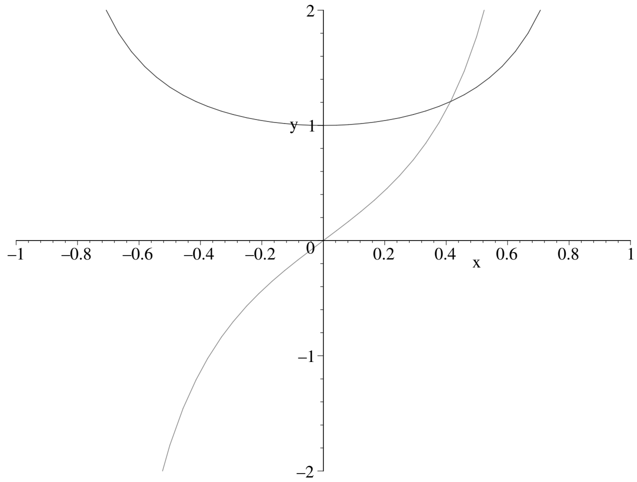


Abbildung 19: Graphen von $f(x) = \frac{1}{1-x^2}$ und $f'(x)$

der also jeder differenzierbaren Funktion auf D ihre Ableitung D$f = f'$ zuordnet, ein linearer Operator (eine lineare Abbildung) ist. V ist dabei der $\mathbb{R}$-Vektorraum der differenzierbaren Funktionen auf D.

Beachte: Es gilt auch $fg \in V$, falls $f, g \in V$ gilt.

Ist etwa $f(x) = x^2 \exp(x)$, dann ist

$$f^{(2002)}(x) = x^2 \exp(x) + 2002 \cdot 2x \exp(x) + \binom{2002}{2} \cdot 2 \exp(x) = (x^2 + 4004x + 2002) \exp(x)$$

Wenn man die Ableitung von

$$\begin{array}{rcl} h : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & (x^2+1)^{1001} \end{array}$$

ausrechnen sollte, könnte man $(x^2 + 1)^{1001}$ nach der Binomischen Formel berechnen, man erhält ein Polynom der Gestalt

$$x \mapsto x^{2002} + 1001x^{2000} + \ldots + 1$$

und kann dieses differenzieren.
Einfacher geht es mit der Kettenregel, wenn man beachtet, dass sich h aus zwei Funktionen zusammensetzt

$$x \mapsto x^2 + 1 := y \ \text{ und } \ y \mapsto y^{1001}$$

19.2.3 Satz (Kettenregel)

Sind $D, D^* \subset \mathbb{R}$ echte Intervalle und $f : D \to \mathbb{R}$ und $g : D^* \to \mathbb{R}$ Funktionen und es gelte $f(D) \subset D^*$.
Dann gilt: Ist f differenzierbar in $a \in D$, g ist differenzierbar in $b = f(a) \in D^*$, dann ist auch die Zusammensetzung

$$g \circ f : D \to \mathbb{R}$$

differenzierbar in a und es gilt

$$\boxed{(g \circ f)'(a) = g'(f(a)) \cdot f'(a)}$$

Zum Beweis sei wieder

$$f(x) = f(a) + (x - a)\varphi(x), \ \varphi \text{ stetig in } a \text{ und } \varphi(a) = f'(a)$$

bzw.

$$g(y) = g(b) + (y-b)\psi(y),\ \psi \text{ stetig in } b = f(a) \text{ und } \psi(b) = g'(b).$$

Für die Zusammensetzung $g \circ f$ ergibt sich

$$(g \circ f)(x) = g(f(x)) = g(b) + \psi(b + \varphi(x)(x-a))\varphi(x)(x-a).$$

Da $x \mapsto \psi(b+\varphi(x)(x-a))\varphi(x)$ an der Stelle a stetig ist und dort den Wert $\psi(b)\varphi(a) = g'(f(a))f'(a)$ hat, ergibt sich die Behauptung.

19.2.4 Bemerkungen und Beispiele

(a) Der naheliegende Beweisansatz mit Hilfe des Differenzenquotienten ($x \neq a$)

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

erfordert, dass wenigstens in einer Umgebung von a $f(x) \neq f(a)$ gelten muss.

In jeder Umgebung von a kann es jedoch Punkte $x \neq a$, für die $f(x) - f(a) = 0$ gilt. Mit etwas Mühe (man unterscheidet die Fälle: $f'(a) \neq 0$ und $f'(a) = 0$) kann man aber auch den obigen Beweisansatz retten, einfacher geht es mit dem von uns verwendeten Ansatz.

(b) Unser Eingangsbeispiel

$$\begin{array}{rcl} h : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & (x^2+1)^{1001} \end{array}$$

wird jetzt erledigt durch die Beobachtung, dass $h = g \circ f$ mit $f(x) = x^2 + 1$ und $g(y) = y^{1001}$ ist. Es ist also (für alle $x \in \mathbb{R}$)

$$h'(x) = (g \circ f)'(x) = g'(f(x)) \cdot f'(x) = 1001y^{1000} \cdot 2x = 2002x(x^2+1)^{1000}.$$

(c) Die allgemeinen Exponentialfunktionen

$$\begin{array}{rcl} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \exp_a(x) = \exp(x \log a) =: a^x \quad (a \in \mathbb{R}_+^*) \end{array}$$

sind differenzierbar, und es gilt

$$\exp_a'(x) = \log a \exp(x \log a) = \log a \exp_a(x)$$

Als letzte Regel für die Technik des Differenzierens wollen wir eine Regel ableiten, die es gestattet, aus Eigenschaften einer differenzierbaren Funktion f, die umkehrbar ist, z.B. weil sie streng monoton ist, auf die Differenzierbarkeit der Umkehrfunktion zu schließen und einen Zusammenhang zwischen der Ableitung des Umkehrfunktion und der Funktion abzuleiten. Wie dieser Zusammenhang aussieht, ist leicht zu erraten.
Wir nehmen an, dass g die Funktion f umkehrt, also $g(f(y)) = y$ für alle $y \in D = Def(f)$ und $f(g(x)) = x$ für alle $x \in Def(g)$ gilt. Nehmen wir weiter an, dass f in $b \in D$ und g in $a := f(b) \in Def(g)$ differenzierbar ist. Dann können wir die Kettenregel verwenden und erhalten aus $(f \circ g)(x) = x$ für alle $y \in D$

$$(f \circ g)'(a) = f'(g(a))g'(a) = 1.$$

Also eine notwendige Bedingung für die Differenzierbarkeit von g in $f(b)$, dass $f'(b) \neq 0$ sein muss.

Wir wollen jedoch die Differenzierbarkeit von f aus Eigenschaften von f schließen, die Bedingung $f'(b) \neq 0$ ist aber ein wichtiger Hinweis.
Wir formulieren den Satz über die Differenzierbarkeit der Umkehrfunktion gleich so, wie wir in der meisten Fällen anwenden werden.
Man beachte in diesem Zusammenhang auch, dass für stetige Funktionen f auf (echten) Intervallen die *strenge Monotonie* (d.h. f ist streng monoton wachsend oder streng monoton fallend) und die Injektivität äquivalente Begriffe sind (vgl. ???). Erinnern Sie sich noch an den Beweis? Die eine Richtung ist offensichtlich. Welche?

19.2.5 Theorem (Differenzierbarkeit der Umkehrfunktion)

Sei $D \subset \mathbb{R}$ ein echtes Intervall: $f : D \to \mathbb{R}$ differenzierbar, streng monoton und $f'(y) \neq 0$ für alle $y \in D$.
Dann besitzt $f : D \to f(D) =: D^*$ (D^* ist wieder Intervall) eine differenzierbare Umkehrfunktion $g : D^* \to D \subset \mathbb{R}$,

$$g(x) = y \iff x = f(y),$$

und es gilt für alle $x \in D^*$

$$g'(x) = \frac{1}{f'(y)}.$$

Zum **Beweis** sei f differenzierbar in $b \in D$ und es sei $a := f(b)$ ($\iff g(a) = b$).

Die Existenz der Umkehrfunktion ist nach der Voraussetzung klar, nach Satz ???. Da f insbesondere stetig ist, ist $D^* = f(D)$ auch wieder ein Intervall.
Die Umkehrfunktion g ist selbst wieder streng monoton (im selben Sinne wie f) und auch stetig (da $g(D^*) = D$ ein Intervall ist). Wegen der Differenzierbarkeit von f in b gibt es eine in b stetige Funktion $\varphi : D \to \mathbb{R}$ mit

$$f(y) - f(b) = (y - b)\varphi(y) \text{ und } \varphi(b) = f'(b).$$

Nach Voraussetzung ist $\varphi(b) = f'(b) \neq 0$, daher ist wegen der strengen Monotonie von f sogar $\varphi(y) \neq 0$ für alle $y \in D$.
Daher folgt

$$y - b = \frac{1}{\varphi(y)}\,(f(y) - f(b))$$

oder

$$g(x) - g(a) = \frac{1}{\varphi(g(x))}(x - a)$$

Da aber φ und g stetig in a sind, ist auch $\varphi \circ g$ stetig in a und damit auch $\frac{1}{\varphi \circ g}$, daher ist g differenzierbar in a und

$$g'(a) = \frac{1}{\varphi(g(a))} = \frac{1}{f'(g(a))} = \frac{1}{f'(b)}.$$

□

19.2.6 Beispiele und Bemerkungen

(a) Die Voraussetzung, dass $f'(x) \neq 0$ ist wesentlich, so ist etwa die Potenzfunktion

$$\begin{array}{rcl} p_3 : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 \end{array}$$

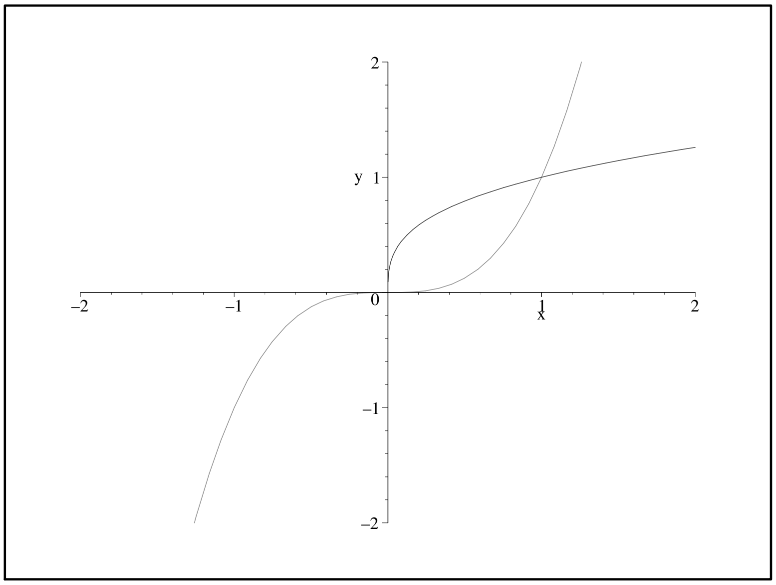


Abbildung 20: Graphen der Funktionen x^3 und $\sqrt[3]{x}$

streng monoton wachsend, differenzierbar auf ganz $\mathbb{R}$, und es gilt $P_3'(x) = 3x^2$ und $p_3(x) = 0 \iff x = 0$.
Die Umkehrfunktion $\sqrt[3]{x}$, ist aber in $x = 0$ nicht differenzierbar, jedoch für alle $x > 0$ oder für alle $x < 0$. (siehe Abb. 20)

(b) Wegen $\exp'(y) = \exp(y) \neq 0$ für alle $y \in \mathbb{R}$ ist $\log : \mathbb{R}_+^* \to \mathbb{R}$ differenzierbar, und es gilt

$$\log'(x) = \frac{1}{\exp'(\log x)} = \frac{1}{\exp(\log x)} = \frac{1}{(\exp \circ \log)(x)} = \frac{1}{x}$$

für alle $x > 0$.

(c) Für alle $\alpha > 0$ sind die Potenzfunktionen

$$\begin{aligned} p_\alpha : \mathbb{R}_+^* &\to \mathbb{R} \\ x &\mapsto x^\alpha := \exp(\alpha \log x) \end{aligned}$$

differenzierbar, und es gilt für alle $x \in \mathbb{R}_+^*$

$$p_\alpha'(x) = \alpha x^{\alpha-1} = \alpha p_{\alpha-1}(x).$$

Der Beweis ist klar nach der Kettenregel und (b)

$$p_\alpha'(x) = \exp(\alpha \log x) \cdot \alpha \cdot \frac{1}{x},$$

also

$$p_\alpha'(x) = \alpha x^{\alpha-1} = \alpha p_{\alpha-1}(x) \quad (\alpha \in \mathbb{R},\ x > 0$$

(d) Ist $f : D \to \mathbb{R}$ eine *nullstellfreie* differenzierbare Funktion, dann ist nach dem Zwischenwertsatz $f(x) > 0$ für alle $x \in D$ oder $f(x) > 0$ für alle $x \in D$. In jedem Fall ist

$$(\log \circ |f|)' = \frac{f'}{f}.$$

Man nennt allgemein $\frac{f'}{f}$ *logarithmische Ableitung* von f' oder $|f|$.

19.3

19.4

19.5

20 Lokale Extrema, Mittelwertsätze, Anwendungen

Eine der wesentlichen Aufgaben der Analysis besteht im Studium des Änderungsverhaltens von Funktionen in Abhängigkeit von Änderungen der betreffenden Argumente. Die Ableitungen im geometrischen Bild, die Steigung der Tangente an den Graphen der Funktion, geben diese Änderung „lokal" an. Es zeigt sich nun, dass man aus der Kenntnis des *lokalen* Veränderungsverhaltens einer differenzierbaren Funktion auf das *globale* Verhalten schließen kann. Viele Eigenschaften einer differenzierbaren Funktion spiegeln sich in Eigenschaften ihrer Ableitungen f' bzw. der höheren Ableitungen. So können etwa *lokale Extrema*, das *Monotonieverhalten*, die *Konvexität* bzw. *Konkavität* einer Funktion mit Hilfe der Ableitungen untersuchen werden. Der zentrale Satz ist der (erste) *Mittelwertsatz der Differentialrechnung* (1.MWSD). Wir folgern den klassischen Beweisschema und folgern ihn aus dem Satz *von Rolle*, diesen wiederum aus einem schon *Fermat* geläufigen notwendigen Kriterium für das Vorliegen eines Extremum für differenzierbare Funktionen in offenen Intervallen. In Wirklichkeit sind alle die Sätze äquivalent (in Wahrheit sogar äquivalent zum Vollständigkeitsaxiom für $\mathbb{R}$) und auch äquivalent zum zweiten MWSD, mit dessen Hilfe wir auch die sog. *Bernoulli-l'Hospital*'schen Regeln ableiten werden. Als Anwendung werden wir auch einige *Extremalprobleme* (aus verschiedenen Anwendungsbereichen) behandeln und weitere *nützliche Ungleichungen* ableiten.

20.1 Lokale Extrema, der MWSD

Wir erinnern an folgende Definition

20.1.1 Definition (lokale und globale Extrema)

Ist $D \neq \emptyset$ und $f : D \to \mathbb{R}$ eine Funktion. Man sagt: f hat in $x_0 \in D$

(a) ein *globales Maximum* (bzw. globales Minimum), wenn für *alle* $x \in D$ gilt

$$f(x) \leq f(x_0) \text{ bzw. } f(x) \geq f(x_0)$$

(b) ein *lokales Maximum* (bzw. lokales Minimum), wenn es ein $r > 0$ gibt, so dass für alle $x \in U_r(x_0) \cap D$ gilt

$$f(x) \leq f(x_0)(\text{ bzw. } f(x) \geq f(x_0))$$

(c) In allen Fällen nennt man x_0 eine lokale bzw. globale Extremalstelle von f und der Funktionswert $f(x_0)$ das globale bzw. lokales Extremum von f in x_0.

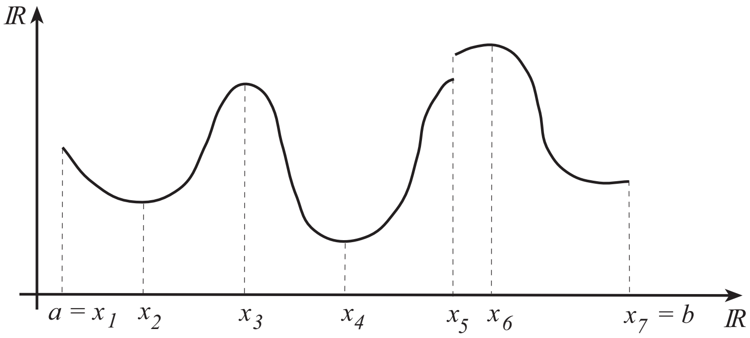
(d) Gilt an den entsprechenden Ungleichungen das Gleichheitszeichen jeweils nur für $x = x_0$, so spricht man von einem *strikten* lokalen oder globalen Maximum bzw. Minimum an der Stelle x_0.

(e) Satt globales Maximum bzw. Minimum sagt man gelegentlich auch *absolutes* Maximum bzw. minimum und statt lokales Maximum bzw. Minimum auch *relatives* Maximum bzw, Minimum.

20.1.2 Bemerkungen

(a) Eine auf einem kompakten Intervall $[a, b]$ stetige Funktionen, hat nach dem Satz von Weierstrass dort ein globales Maximum und ein globales Minimum. Dieser Satz ist jedoch ein reiner Existenzsatz und die kaum einen Hinweis, wie man die jeweilige Extrema Stelle ermittelt.

(b) Eine Funktion kann viele lokale Extrema haben, ein globales Maximum (bzw. Minimum) ist natürlich auch ein lokales Maximum (bzw. Minimum), aber im Allgemeinen nicht umgekehrt. Auch er kann ein globales Maximum (bzw. Minimum) an verschiedenen Stellen des Definition Bereiches angenommen werden. Der geneigte Leser macht sich klar, an welcher Stelle die Funktion f, deren Graf in der folgenden Abbildung skizziert ist, jeweils lokale bzw. globalen Maxima besitzt.



20.1.3 Fermat, $\approx$ 1638

Ist $D \subset \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion, $a \in D$ und es gebe ein $r > 0$ mit $U_r(a) \subset D$. Hat dann f in a ein lokales Extremum und ist f in a differenzierbar, dann gilt

$$\boxed{f'(a) = 0}$$

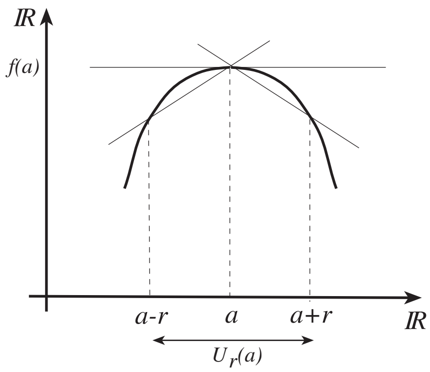
Beweis : Wir nehmen oBdA an, das f an der Stelle a ein lokales Maximum hat (im Falle eines Minimums ersetzt man f durch $-f$). Wir greifen auf die Definition der Differenzierbarkeit als limes des Differenzenquotient zurück.

Für alle $x \in U_r(a)$ mit $x > a$ gilt wegen $f(x) \leq f(a)$

$$\frac{f(x) - f(a)}{x - a} \leq 0$$

also ist auch

$$f'(a) = \lim_{\substack{x \to a \\ x > a}} \frac{f(x) - f(a)}{x - a} \leq 0.$$



Für alle $x \in U_r(a)$ mit $x < a$ gilt auch wegen $f(x) \leq f(a)$

$$\frac{f(x) - f(a)}{x - a} \geq 0$$

und damit auch

$$f'(a) = \lim_{\substack{x \to a \\ x < a}} \frac{f(x) - f(a)}{x - a} \geq 0,$$

also insgesamt $f'(a) = 0$.

□

20.1.4 Bemerkungen

(a) Geometrisch besagt der Satz von Fermat:

Wenn f in einem *inneren Punkt* $a \in D$ (das heißt es gibt ein $r > 0$ mit $U_r(a) \subset D$) ein lokales Extremum besitzt und wenn f in a differenzierbar ist, besitzt der Graph von f im Punkt $(a, f(a)) \in \mathbb{R}^2$ eine waagerechte Tangente.

(b) Die Aussage des Fermatschen Lemmas wird für Randpunkte im Allgemein falsch.

Ist etwa $D := [0, 1]$ und

$$\begin{aligned} f : D &\to \mathbb{R} \\ x &\mapsto x \end{aligned}$$

die Einschränkung von $id_{\mathbb{R}}$ auf $[0, 1]$, dann hat f an der Stelle 0 ein (sogar globales) Minimum und an der Stelle 1 ein (sogar globales) Maximum, es ist aber $f'(x) = 1 \neq 0$ für alle $x \in [0, 1]$.

(c) Wie das einfache Beispiel des Polynoms

$$\begin{aligned} p_3 : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^3 \end{aligned}$$

und $a = 0$ zeigt, bedeutet das Verschwinden der Ableitung in einem (inneren) Punkt keineswegs, dass an dieser Stelle ein lokales Extrema vorliegen muss.

> Das Verschwinden der Ableitung in einen inneren Punkt (Differenzierbarkeit natürlich vorausgesetzt), ist also lediglich eine *notwendige Bedingung* für das Vorliegen eines lokalen Extremum.

(d) Ist $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$ eine Funktion, die auf Extremalstellen zu untersuchen ist, dann hat man folgende Kandidaten ins Auge zu fassen:

1. Die Randpunkte vor D (also 2 Punkte);
2. Die Punkte von D, in denen f nicht differenzierbar ist.
3. Die inneren Punkte $x_0 \in D$, in denen f differenzierbar ist und für die $f'(x_0) = 0$ gilt. Solche Punkte nennt man auch *kritische Punkte*.

> In keinem dieser Punkte braucht aber eine Extremalstelle vorzuliegen.

ABBILDUNG

Wir werden aber bald auch *hinreichende Kriterien* dafür ableiten. Obwohl der Mittelwertsatz der Differenzialrechnung eine sehr anschauliche geometrische Interpretation hat, muss er bewiesen werden. Wir verweisen zunächst einen Spezialfall, aus dem der MWSD dann einfach abzuleiten ist.

20.1.5 Satz (von Rolle, Michael Rolle (1652-1719))

Sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall ($a < b$) und $f : [a,b] \to \mathbb{R}$ eine stetige Funktionen, die im offenen Intervall $]a,b[$ differenzierbar ist und für welche $f(a) = f(b)$ gilt. Dann gibt es einen Punkt $\xi \in]a,b[$ mit $f'(\xi) = 0$.
Insbesondere liegt zwischen zwei Nullstellen von f $(f(a) = f(b) = 0)$ stets eine Nullstellen der Ableitung.

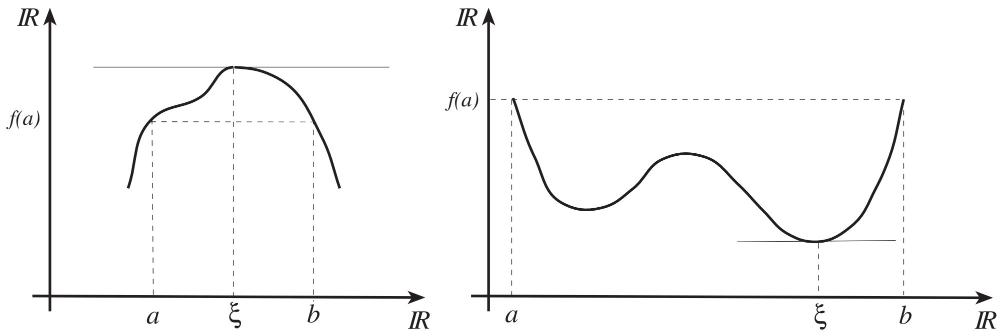
Beweis : Ist f konstant auf $[a,b]$, dann gilt $f'(x) = 0$ für alle $x \in [a,b]$, insbesondere gilt etwa für $\xi := \frac{a+b}{2} \in]a,b[$ auch $f'(\xi) = 0$.

Ist f aber nicht konstant, dann hat f als stetige Funktionen auf der kompakte Intervall $[a,b]$ ein globales Maximum und ein globales Minimum. Diese können wegen $f(a) = f(b)$ nicht bei der in den Endpunkten des Intervall vorliegen. Wenn ein globales Minimum bei a vorliegt und damit auch bei b, dann liegt ein existierendes globales Maximum in einem Punkt $\xi \in]a,b[$. Wegen $]a,b[= U_r(\xi)$ mit $\xi := \frac{a+b}{2}$ und $r := \frac{b-a}{2}$ sind Voraussetzungen des Fermatschen Lemmas erfüllt. Es gilt also $f'(\xi) = 0$.

Im Fall, dass in a (und damit auch in b) ein globales Maximum vorliegt, schließt man völlig analog: Ein sicher vorhandenes globales Minimum muss dann wieder in $[a,b[$ liegen etc.

□

20.1.6 Geometrische Veranschaulichung



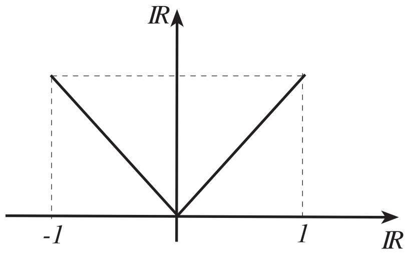
20.1.7 Bemerkungen zu den Voraussetzungen des Satzes von Rolle

Die Voraussetzungen des Satzes von Rolle sind im folgenden Sinne „scharf“:
Verzichtet man auf eine der Voraussetzungen, dann wird der Satz i.A. falsch:

(a) Verzichtet man auf die Voraussetzung der Differenzierbarkeit auch nur in einem Punkt $x_0 \in]a,b[$, dann wird der Satz i.A. falsch.
Ist z.B. $[a,b] = [-1,1]$ und

$$\begin{aligned} f : [-1,1] &\to \mathbb{R} \\ x &\mapsto |x| \end{aligned}$$

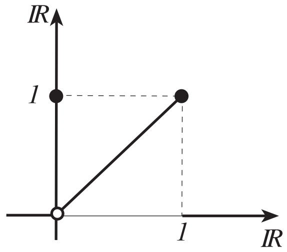
die Betragsfunktion, dann gilt $f(-a) = f(a) = 1$ und f ist stetig in $[-1,1]$, aber f ist nicht differenzierbar in 0, es gibt bei $\xi \in]-1,1[$ mit $f'(\xi) = 0$.



(b) $f : [0,1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 1, & \text{für } x = 0; \\ x, & \text{für } 0 \le x \le 1 \end{cases}$$

erfüllt $f(0) = f(1)(= 1)$, ist differenzierbar in $]0,1[$ mit Ableitung $f'(x) = 1$ für alle $x \in]0,1[$, aber es gibt kein $\xi \in]0,1[$ mit $f'(\xi) = 0$. Hier ist f nicht stetig in Null.



Der Mittelwertsatz ist nun eine einfache Konsequenz aus dem Satz von Rolle, er hat zunächst die gleichen Voraussetzungen, man verzichtet jedoch auf die Bedingung

$$f(a) = f(b).$$

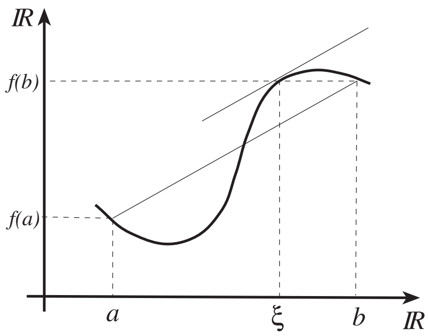
20.1.8 Theorem (MWSD, Lagrange (1797) (J.L.Lagrange, 1736-1813))

Ist $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall ($a < b$) und $f : [a,b] \to \mathbb{R}$ eine stetige Funktion, die im offenen Intervall $]a,b[$ differenzierbar ist. Dann gibt es ein $\xi \in]a,b[$ mit

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

20.1.9 Geometrische Veranschaulichung

Auf dem Graphen von f gibt es mindestens einen Punkt $(\xi, f(\xi))$, in welchem die Tangente parallel zur Sekanten durch die Punkte $(a, f(a))$ und $(b, f(b))$ ist.



Zum **Beweis** betrachtet man die Hilfsfunktion $h : [a,b] \to \mathbb{R}$ mit

$$h(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

(Geometrisch bedeutet dies eine Scherung des Graphen von f.)

Dann ist auch h stetig in $[a,b]$ und differenzierbar in $]a,b[$, und es gilt

$$h(a) = f(a) \text{ und } h(b) = f(b) - \frac{f(b)-f(a)}{b-a}(b-a) = f(a),$$

also auch $h(a) = h(b)$.
damit erfüllt h die Voraussetzung des Satzes von Rolle, es gibt also wegen $h'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$ ein $\xi \in]a,b[$ mit

$$0 = h'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{b-a},$$

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

20.1.10 Bemerkungen

(a) Der „Nachteil" des MWSD liegt in der Tatsache, dass die Zwischenstelle $\xi \in]a,b[$ nicht näher spezifiziert wird, wie einfache Beispiele zeigen, kann es mehrere ξ mit der genannten Eigenschaft geben.

(b) Häufig verwendet man den MWSD in der folgenden Form:
Ist f stetig auf dem Intervall $[a, a+h]$ $(h > 0)$ bzw.$[a+h, a]$ (falls $h < 0$) und differenzierbar in $]a, a+h[$ bzw. $]a+h, a[$, dann gibt es ein ϑ mit $0 < \vartheta < 1$ und

$$f(a+h) = f(a) + f'(a+\vartheta h)h.$$

(c) Eine kleine Anwendung:
Der MWSD eignet sich zur näherungsweisen Berechnung von Funktionswerten.
Um etwa $\sqrt{2}$ näherungsweise zu berechnen setzen wir $2 = a + h$ mit $a = 1,96$ und $h = 0,04$ und erhalten wegen

$$\sqrt{a+h} = \sqrt{a} + \frac{1}{2\sqrt{a+\vartheta h}}$$

mit $0 < \vartheta < 1$

$$\sqrt{2} = \sqrt{1,96 + 0,04} = \sqrt{1,96} + 0,04 \cdot \frac{1}{2\sqrt{3}}$$

mit $1,96 < \xi < 2\ (< 2,0164)$.
Hieraus folgt mit $1,4 < \sqrt{\xi} < 1,42$ bereits $1,41414 < \sqrt{2} < 1,41429$.
Gar nicht mal so schlecht!

In der anspruchsvolleren Schulbuchliteratur zu Analysis wird die folgende Variante der MWSD diesem meist vorgezogen.

20.1.11 Schrankensatz

Ist $f : [a,b] \to \mathbb{R}$ eine stetige und in $]a,b[$ differenzierbare Funktion.
Für die Ableitung gelte $m \leq f'(\xi) \leq M$ für alle $\xi \in]a,b[$ mit festen Konstanten $m, M \in \mathbb{R}$. Dann gilt für alle $x, y \in [a,b]$ die Abschätzung

$$\boxed{m(y-x) \leq f(y) - f(x) \leq M(y-x)}$$

Der Beweis folgt unmittelbar aus

$$f(y) - f(x) = f'(\xi)(y-x)$$

20.1.12 Folgerung

Eine differenzierbare Funktion f auf einem echten Intervall $D \subset \mathbb{R}$ mit beschränkter Ableitung ist dort Lipschitz-stetig. Genauer gilt:
Ist $f : D \to \mathbb{R}$ differenzierbar und gilt $|f'(\xi)| \leq L$ für alle $x \in D$, dann gilt für alle $x, y \in D$

$$(*) \qquad |f(y) - f(x)| \leq L|y - x|$$

Der Beweis folgt unmittelbar aus dem Schrankensatz oder direkt aus dem MWSD. Zum Beispiel ist also $\sin : \mathbb{R} \to \mathbb{R}$ Lipschitz-stetig (und damit gleichmäßig stetig auf ganz $\mathbb{R}$), dann gilt wegen $\sin'(\xi) = \cos(\xi)$ und $|\cos\xi| \leq 1$ gilt für alle $x, y \in \mathbb{R}$

$$|\sin y - \sin x| \leq 1 \cdot |y - x|$$

mit der Lipschitz-Konstante $L = 1$.
Die Ungleichung $(*)$ gilt auch für komplexwertige (differenzierbare) Funktionen $f : D \to \mathbb{C}$ ($D \subset \mathbb{R}$ Intervall) obwohl für solche Funktionen der Satz von Rolle bzw. der MWSD i.A. nicht gilt. Weitere Folgerungen aus dem MWSD bzw. dem Schrankensatz sind:

20.1.13 Satz (Charakterisierung konstanter Funktionen und von Stammfunktionen)

(a) Eine differenzierbare Funktion $f : D \to \mathbb{R}$ ($D \subset \mathbb{R}$ echtes Intervall) ist genau dann konstant, wenn $f'(x) = 0$ für alle inneren Punkte $x \in D$ gilt.

(b) Ist $F : D \to \mathbb{R}$ *eine* Stammfunktion von $f : D \to \mathbb{R}$, d.h. ist F differenzierbar und gilt $F'(x) = f(x)$ für alle $x \in D$, dann ist $\{F + c;\ c \in \mathbb{R}\}$ die Menge *aller* Stammfunktionen von f auf dem Intervall D.

Die eine Richtung ist evident: Ist $f(x) = const.$ für alle $x \in D$, dann ist $f'(x) = 0$ für alle $x \in D$, insbesondere für alle inneren Punkte.
Die Umkehrung folgt aus dem Schrankensatz mit $m = M = 0$ oder direkt aus dem MWSD (Ist $x_0 \in D$ ein fester Punkt, dann hat man $f(x) = f(x_0)$ zu zeigen; Man kann auf $[x_0, x]$, $[\bar{x}_0, x]$ bzw. $[\bar{x}, x_0]$ den MWSD anwenden).
Zum Beweis von (b) bemerken wir, dass natürlich $G = F + c$, c konstant, ebenfalls eine Stammfunktion ist. Die Umkehrung folgt aus (a), den G und F dieselbe Ableitung haben, ist $(G - F)' = 0$ gilt und damit $G = F + c$ gelten muss.

Die Funktionen

$$\begin{aligned} f_c : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \exp(cx) \quad (c \in \mathbb{R}) \end{aligned}$$

erfüllen die Gleichung $f_c' = cf_c$. Hiervon gilt eine gewisse Umkehrung.

20.1.14 Folgerung (Charakterisierung von exp durch eine Differentialgleichung)

Ist $c \in \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion mit $f'(x) = cf(x)$ für alle $x \in \mathbb{R}$, dann gilt für alle $x \in \mathbb{R}$

$$f(x) = A\exp(cx) \ \text{ mit } \ A := f(0).$$

Insbesondere ist $\exp : \mathbb{R} \to \mathbb{R}$ die eindeutig bestimmte Lösung des Anfangswertproblems $f' = f$ und $f(0) = 1$.

20.1.15 Monotoniekriterien

Ist $D \subset \mathbb{R}$ ein (echtes) Intervall und $f : D \to \mathbb{R}$ eine differenzierbare Funktion. Dann gilt

(a) f ist *genau dann* monoton wachsend (bzw. monoton fallend) auf D, wenn $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$) für alle inneren Punkte $x \in D$ gilt.

(b) Wenn $f'(x) > 0$ (bzw. $f'(x) < 0$) ist für alle inneren Punkte $x \in D$, dann ist f sogar streng monoton wachsend (bzw. fallend) auf D.

Der **Beweis** ergibt sich in der einen Richtung („$\Longleftarrow$") sofort aus dem MWSD: Sind $x, y \in D$ und gilt $x < y$, dann gibt es ein $\xi \in]x, y[$ (ξ ist dann innerer Punkt von D) mit $f(y) - f(x) = f'(\xi)(y - x) \geq 0$, also ist

$$f(x) \leq f(y).$$

$\Longrightarrow$ Ist $x \in M$ fest, dann ist für alle $y \in D,\ y \neq x$ stets der Differenzenquotient $\frac{f(y)-f(x)}{y-x} \geq 0$, also auch $f'(x) \geq 0$.
Ist sogar $f'(x) > 0$ für alle inneren Punkte $x \in D$, dann gilt auch $f'(\xi) > 0$, also folgt $f(y) - f(x) = f'(\xi)(y - x) > 0$.

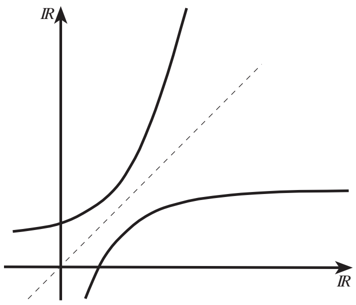
20.1.16 Bemerkungen und Beispiele

(a) Die Potenzfunktion

$$\begin{aligned} p_3 : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^3 \end{aligned}$$

ist streng monoton wachsend, obwohl $p_3'(0) = 0$ gilt.
Die Positivität der Ableitung ist also lediglich ein *hinreichendes* Kriterium für die strenge Monotonie.

(b) Wegen $\exp'(x) = \exp(x) > 0$ für alle $x \in \mathbb{R}$ ist $\exp : \mathbb{R} \to \mathbb{R}$ streng monoton wachsend; Wegen $\log'(x) = \frac{1}{x} > 0$ für $x > 0$ ist auch $\log : \mathbb{R}_+ \to \mathbb{R}$ streng monoton wachsend.



(c) sin ist streng monoton wachsend in $[-\frac{\pi}{2}, \frac{\pi}{2}]$;
cos ist streng monoton fallend in $[0, \pi]$;
tan ist streng monoton wachsend in $]-\frac{\pi}{2}, \frac{\pi}{2}[$;
cot ist streng monoton fallend in $]0, \pi[$.

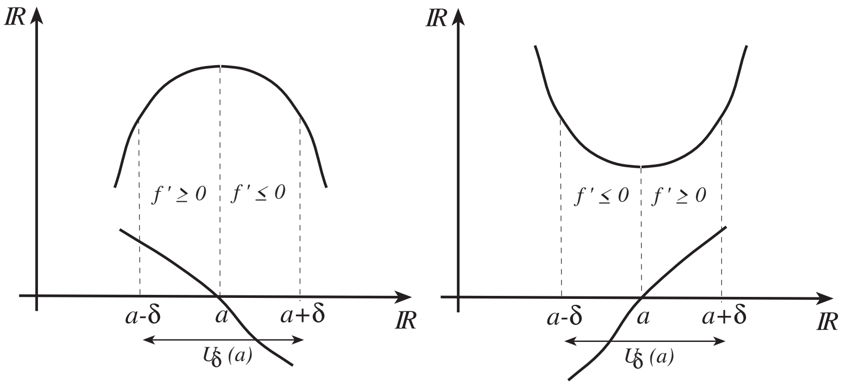
Dort haben diese (auf diese Intervalle eingeschränkte) Funktionen Umkehrfunktionen, die im gleichen Sinne wie die Funktionen monoton sind (vgl. ???)

Eine weitere Anwendung des Monotoniekriteriums sind *hinreichende Bedingungen* für Extremwerte. Wir wissen, dass das Verschwinden der Ableitung in einem inneren Punkt a des Definitionsbereichs für eine in a differenzierbare Funktion ein *notwendiges* Kriterium für ein lokales Extremum ist, wir wissen auch - das obige Beispiel von $p_3 : \mathbb{R} \to \mathbb{R},\ x \mapsto x^3$ und $a = 0$, zeigt, dass i.A. es nicht hinreichend ist.

20.1.17 Satz (hinreichende Kriterien für lokale Extrema)

Ist $D = U_\delta(a),\ \delta > 0$, eine δ-Umgebung von $a \in \mathbb{R}$ und $f : D \to \mathbb{R}$ differenzierbar und $f'(a) = 0$. Dann gilt

(a) Wechselt f' beim Durchgang durch a das Vorzeichen, das heißt gilt für alle $x \in U_\delta(a)$ $(x-a)f'(x) \geq 0$ (bzw. ≤ 0), dann hat f in a ein lokales Minimum (bzw. Maximum)



a ist die einzige Minimal bzw. Maximalstelle von f in $U_\delta(a)$, wenn a die einzige Nullstelle von f' in $U_\delta(a)$ ist.

(b) Ist f in a zweimal differenzierbar und gilt $f''(a) > 0$ (bzw. $f''(a) < 0$), dann hat f in a ein (sogar striktes) lokales Minimum (bzw. Maximum).

Beweis von (a): Der Beweis ergibt such unmittelbar aus dem Monotoniekriterium:
Gilt etwa $(x-a)f'(x) \leq 0$, dann ist f in $]a-\delta, a]$ monoton fallend und in $[a, a+\delta]$ monoton wachsend. f muss also in a ein lokales Minimum haben. Im Fall $(x-a)f'(x) \geq 0$ ist f in $]a-\delta, a]$ monoton fallend und in $[a, a+\delta]$ monoton wachsend, bei a liegt also ein lokales Maximum.

□

Beweis von (b): Wir behandeln exemplarisch den Fall $f''(a) > 0$ (man gehe nicht von f zu $-f$ über).
Wegen

$$f''(a) = \lim_{\substack{x \to a \\ x \neq a}} \frac{f'(x) - f'(a)}{x - a} > 0$$

gibt es ein δ^* $(0 < \delta' \leq \delta)$, so dass $\frac{f'(x)-f'(a)}{x-a} > 0$ gilt für alle $x \in U_\delta(a)$ mit $0 < |x-a| < \delta^*$.
Wegen $f'(a) = 0$ folgt hieraus $f'(x) < 0$ für $a - \delta^* < x < a$ und $f'(x) > 0$ für $a < x < a + \delta^*$; f ist also in $]a-\delta^*, a[$ streng monoton fallend und in $]a, a+\delta^*[$ streng monoton wachsend. f besitzt also in a ein striktes lokales Minimum.

□

Bemerkung zum letzten Beweis:
Hätten wir vorausgesetzt, dass sogar $f \in \mathcal{C}^2(D)$ gilt, dann hätten man aus der Stetigkeit von f'

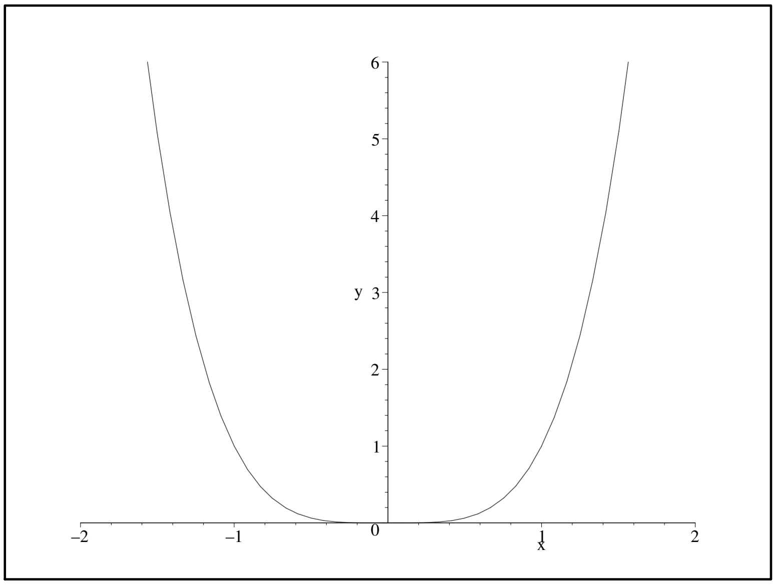


Abbildung 21: Graph der Funktion x^4

auch direkt auf die Existenz eines weiteren δ^* schließen können.

20.1.18 Beispiele und Bemerkungen

(a) Das einfache Beispiel

$$\begin{array}{rcl} p_4 : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^4 \end{array}$$

mit $a = 0$ zeigt, dass das Kriterium (b) nicht notwendig für das Vorliegen eines lokalen Extremum ist.
Es gilt $p_4'(0) = 0$ und p_4 hat in $x = 0$ ein lokales Minimum. (siehe Abb. 21)

Zur Entscheidung kann man in einem solchen Fall die höheren Ableitungen herausziehen: Offensichtlich gilt $p_4'(x) = 4x^3,\ p_4''(x) = 12x^2,\ p_4''' = 24x$ und $p_4^{(4)} = 24$, insbesondere $p_4^{(4)} = 24 > 0$.

(b) Ob ein Extremum vorliegt, kann man häufig mit Hilfe der Taylorschen Formel (vgl. ???) entscheiden. Im konkreten Fall ist offensichtlich, dass an der Stelle $a < 0$ sogar ein globales Minimum von p_4 vorliegt.

Das aus der Schulmathematik bekannte Verfahren zur Extremwertbestimmung findet also durch Satz 19.1.17 eine Rechtfertigung. Ob eine gegebene Funktion an einer Stelle $a \in D$ ein lokales Extremum hat, ist häufig ein schwer zu entscheidendes Problem. Hat man nur notwendige Kriterien zur Verfügung, so ergibt sich häufig aus der Fragestellung selbst oder auf Grund der Herkunft aus einem naturwissenschaftlichen oder geometrischen Problem oder einem Problem aus einem anderen Anwendungsgebiet die Möglichkeit zu entscheiden, welche nach den notwendigen Bedingungen in Frage kommende Stellen die gesuchten Extremalstellen sind. Eine solche Überlegung ist häufig aufschlussreicher, als die formale Anwendung hinreichender Kriterien. Dennoch geben wir ein paar Beispiele für diese, um die Bandbreite der Anwendungen auszudehnen.

(c) Wir bemerken noch eine interessante Eigenschaft der Ableitung einer differenzierbaren Funktion f auf einem Intervall (in der Literatur häufig als **Satz von Darboux** zitiert:
Ist $f[a,b] \to \mathbb{R}$ differenzierbar und ist c eine Zahl mit

$$f'(a) < c < f'(b) \text{ oder } f'(b) < c < f'(a),$$

dann gibt es ein $\xi \in]a,b[$ mit $f'(\xi) = c$.

Zum **Beweis** betrachte man

$$\begin{aligned} h : [a,b] &\to \mathbb{R} \\ x &\mapsto f(x) - cx \end{aligned}$$

Die Ableitung einer differenzierbaren Funktion hat also die „Zwischenwerteigenschaft“. Da es differenzierbare, aber nicht stetig differenzierbare Funktionen gibt (vgl. ???), zeigt dies nochmals, dass die Zwischenwerteigenschaft die stetigen Funktionen nicht charakterisiert. Außerdem zeigt der Satz von Darboux, dass eine Funktion $f : D \to \mathbb{R}$ höchstens dann eine Stammfunktion F haben kann, wenn $f(= F')$ die Zwischenwerteigenschaften hat.

20.1.19 Beispiele für Extremwertaufgaben (wird nachgetragen)

20.1.20 Konvexe Funktionen

Ist eine Funktion $f : D \to \mathbb{R}$ zweimal differenzierbar, dann gibt die zweite Ableitung die Änderung der Tangenrichtungen an, beschreibt also wie sich der Graph von f „krümmt“.
Eine präzise Definition des Begriffs „Krümmung“ geben wir erst später.
Völlig ohne Differezierbarkeitsvoraussetzungen kann man den Begriff der *konvexen Funktionen* einführen, ist f allerdings zusätzlich zweimal differenzierbar, so ergibt sich ein einfaches Konvexitätskriterium über die zweite Ableitung ($f''(x) \geq 0$ für alle $x \in D$).
Zentrale Ungleichungen der Analysis beruhen auf der Konvexität einfacher Funktionen.

20.1.21 Definition (konvexe und konkave Funktion)

Ist $D \subset \mathbb{R}$ ein echtes Intervall:
Eine Funktion $f : D \to \mathbb{R}$ heißt *konvex*, wenn für je zwei Punkte $x_1, x_2 \in D$ und alle $\lambda \in [0,1]$ gilt

$$\boxed{f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2)}$$

f heißt *konkav* $\Longleftrightarrow -f$ konvex ist.

Die Fälle $\lambda = 0$ oder $\lambda = 1$ können wir dabei unberücksichtigt lassen, da die obige Ungleichung für diese Werte automatisch erfüllt ist. Auch der Fall $x_1 = x_2$ ist ohne Bedeutung, so dass wir oBdA immer $x_1 < x_2$ und $\lambda \in]0,1[$ annehmen können und sollen.

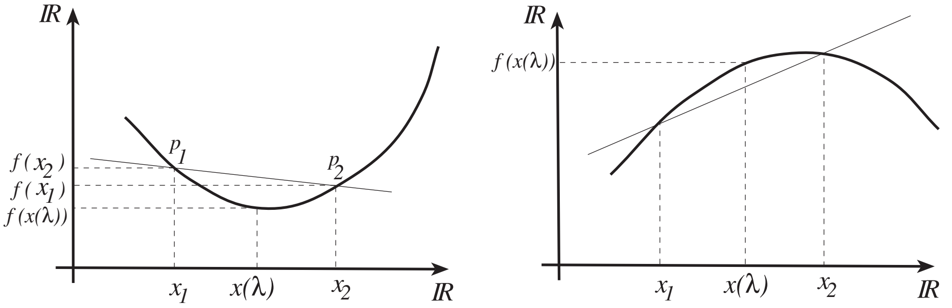
20.1.22 Geometrische Interpretation der Konvexität

Ist $x_1 < x_2$, dann durchläuft der Punkt $x(\lambda) = (1-\lambda)x_1 + \lambda x_2 = x_1 + \lambda(x_2 - x_1)$ gerade das Intervall $]x_1, x_2[$, wenn λ das Intervall $]0,1[$ durchläuft. Durchläuft λ das Intervall $]0,1[$, so durchläuft wegen

$$x(\lambda) = (1-\lambda)f(x_1) + \lambda f(x_2) = f(x_1) + \lambda(f(x_2) - f(x_1)) = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x(\lambda) - x_1)$$

der Punkt $(x(\lambda), f(x))$ gerade das Geradenstück s zwischen den Punkten $P_1 = (x_1, f(x_1))$ und $P_2 = (x_2, f(x_2))$ auf dem Graphen von f. s ist ein Stück der Sekanten durch die Punkte P_1 und P_2.

Konvexität von f bedeutet also, dass der Graph von f zwischen zwei Punkten x_1 und x_2 immer unterhalb der Sehne (Sekante) s verläuft. Konkavität bedeutet, dass s immer *unterhalb* der Graphen von f liegt.



Man könnte etwas unpräzise, aber anschaulich sagen:
f konvex bedeutet, dass der Graph von f mit wachsenden Argumenten eine „Linkskurve“ beschreibt (stellen Sie sich vor, der Graph von f ist in einer Ebene irgendwie markiert und Sie fahren auf der Markierung mit dem Fahrrad entlang).

Übungsaufgabe:
Ist $f : D \to \mathbb{R}$ konvex und besitzt f in $x_0 \in D$ ein *lokales* Minimum, so ist dieses auch *globales* Minimum für f.

Konvexität bzw. Konkavität lässt sich bei *differenzierbaren* Funktionen mit Hilfe der zweiten Ableitung charakterisieren:

20.1.23 Satz

Ist $f : D \to \mathbb{R}$ zweimal differenzierbar, dann ist f genau dann konvex (bzw. konkav), wenn $f''(x) \geq 0$ (bzw. $f''(x) \leq 0$) für alle inneren Punkte $x \in D$ gilt.

Beweis : „$\Longleftarrow$“

Sei $f''(x) \geq 0$ für alle inneren Punkt $x \in D$. Dann ist nach 19.1.15 $f| \ : D \to \mathbb{R}$ monoton wachsend.
Seien $x_1, x_2 \in D$ und oBdA $x_1 < x_2$ und $\lambda \in]0,1[$ und $x = x(\lambda) = (1-\lambda)x_1 + \lambda x_2$. Es ist dann $x_1 < x < x_2$. Wir wenden den MWSD auf die Intervalle $[x_1, x]$ und $[x, x_2]$ an:
Es existieren daher $\xi_1 \in]x_1, x[$ und $\xi_2 \in]x, x_2[$ mit

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) \leq f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}$$

(die mittlere Ungleichung gilt, da $f|$ monoton wächst). Da nun

$$x - x_1 = \lambda(x_2 - x_1) \ \text{ und } \ x_2 - x = (1-\lambda)(x_2 - x_1)$$

gilt, folgt

$$\frac{f(x)-f(x_1)}{\lambda} \leq \frac{f(x_2)-f(x)}{1-\lambda}$$

oder auch

$$f(x) = f(x(\lambda)) \leq (1-\lambda)f(x_1) + \lambda f(x_2),$$

d.h.: f ist konvex.

„$\Longrightarrow$"

Sei umgekehrt $f : D \to \mathbb{R}$ konvex.

Wir führen einen indirekten Beweis und nehmen an, dass *nicht* $f''(x) \geq 0$ für alle inneren Punkte $x \in D$ gilt. Dann gibt es einen inneren Punkt $x_0 \in D$, für den $f''(x_0) \geq 0$ gilt. Sei $l := f'(x_0)$ und $\varphi : D \to \mathbb{R}$ definiert durch $x \mapsto f(x) - l(x - x_0)$.

Dann ist φ zweimal differenzierbar und $\varphi(x_0) = 0$ und $\varphi''(x_0) = f''(x_0) < 0$. φ hat also nach 19.1.17(b) in x_0 ein striktes lokales Maximum. Es gibt also ein $h > 0$, so dass $[x_0-h, x_0+h] \subset D$ und $\varphi(x_0 - h) < \varphi(x_0)$ und $\varphi(x_0 + h) < \varphi(x_0)$ gilt. Hieraus folgt aber

$$f(x_0) = \varphi(x_0) > \frac{1}{2}(\varphi(x_0 - h) + \varphi(x_0 + h)) = \frac{1}{2}(f(x_0 - h) + f(x_0 + h))$$

Setzt man nun

$$x_1 := x_0 - h,\ x_2 := x_0 + h, \lambda := \frac{1}{2},$$

so ist $x_{=}(1-\lambda)x_1 + \lambda x_2$, also

$$f((1-\lambda)x_1 + \lambda x_2) > (1-\lambda)f(x_1) + \lambda f(x_2).$$

Diese Ungleichung steht aber im Widerspruch zur vorausgesetzten Konvexität von f.

Also war unsere Annahme falsch. Wenn f konvex ist, gilt daher $f''(x)$ für alle inneren Punkte $x \in D$.

□

Eine einfache, aber wichtige Anwendung der Konvexität von exp ist die folgende Ungleichung.

20.1.24 Satz (Youngsche Ungleichung)

Seien $p, q \in]1, \infty[$ mit $\frac{1}{p} + \frac{1}{q} = 1$.

Dann gilt für alle $a, b \in \mathbb{R}_+$ die Unlgeichung

$$\boxed{a \cdot b \leq \frac{a^p}{p} + \frac{b^q}{q}}$$

Der Beweis ergibt sich einfach für $a, b > 0$. (für $ab = 0$ ist die Ungleichung richtig) Mit $x := \log a$ und $y := \log b$ folgt wegen der Konvexität von exp

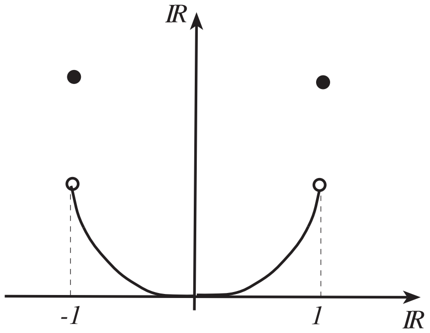
$$\begin{aligned} ab = \exp(x)\exp(y) &= \exp(x+y) \\ &= \exp(\frac{1}{p}px + \frac{1}{q}qy) \\ &\leq \frac{1}{p}\exp(px) + \frac{1}{q}(\exp qy) \\ &= \frac{ap}{p} + \frac{bq}{q} \end{aligned}$$

(man verwende die Konvexitätsbedingung mit $1 - \lambda = \frac{1}{p}$ und $\lambda = \frac{1}{q}$).
Die Youngsche Ungleichung ist die *Quelle für viele wichtige Ungleichungen* etwa die *Höldersche Ungleichung* für die p-Normen im $\mathbb{R}^n$ oder $\mathbb{C}^n$ als Spezialfall die CSU für das Standardskalarprodukt im $\mathbb{K}^n$ die Minkowskische Ungleichung in $\mathbb{K}^n$ und auch die entsprechende Ungleichungen für Integrale.

Zum Schluss sei bemerkt, dass die Konvexität eine Funktion eine sehr starke Eigenschaft ist. Ist $f : D \to \mathbb{R}$ konvex, dann ist f in allen inneren Punkten stetig. (man kann sogar zeigen, dass für alle inneren Punkte $x \in D$ f links- und rechtsseitig differenzierbar und, dass $f'_-(x) \leq f'_+(x)$ gilt.
In Randpunkten gilt dies i.A. nicht, wie das einfache Beispiel $f : [-1, 1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 2, & \text{falls } x = -1 \text{ oder } x = 1; \\ x^2, & \text{falls } x \in]-1, 1[\end{cases}$$

zeigt.



20.1.25 Beispiel (Wendepunkte)

Sei $f : D \to \mathbb{R}$ eine Funktion (auf einem echten Intervall D). Dann heißt $a \in D$ Wendepunkt von f, wenn a innerer Punkt von D ist und ein $\delta > 0$ existiert, so dass f auf $[a - \delta, a]$ konvex und auf $[a, a + \delta]$ konkav ist oder umgekehrt. Ist f in a differenzierbar, dann heißt die Tangente $x \mapsto f(a) + f'(a)(x - a)$ an den Graphen von f im Punkt $(a, f(a))$ *Wendepunkt* von f in a. Sie durchsetzt den Graphen im Punkt $(a, f(a))$.

ZEICHNUNG

Ist a innerer Punkt und in einer Umgebung von a differenzierbar und im Punkt a zweimal differenzierbar, dann ist $f''(a) = 0$ eine notwendige Bedingung für das Vorliegen eines Wendepunkts von f in a, da f' im Wendepunkt dann ein lokales Extremum hat. Ein hinreichendes Kriterium enthält

20.1.26 Satz

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $a \in D$ ein innerer Punkt. Ferner sei $n \geq 3$ ungerade. Ist dann $f^{(n-1)}$ in a differenzierbar und gilt

$$f^{(3)}(a) = \ldots = f^{(n-1)} = 0,\ f^{(n)}(a) \neq 0,$$

so hat f in a einen Wendepunkt.

Beweis : f'' wechselt beim Durchgang durch a das Vorzeichen.

□

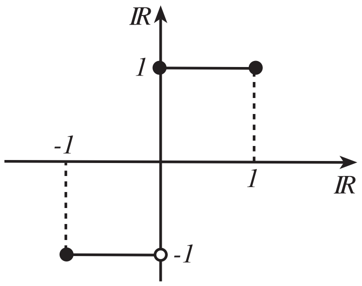
21 Zusammenhang zwischen Integral- und Differenzialrechnung: Der Hauptsatz und seine Anwendungen

In diesem Abschnitt zeigen wir, dass Integralrechnung und Differentialrechnung, die aus völlig verschiedenen Motiven entwickelt wurden (etwa Flächeninhaltsproblem bzw. Tangentenproblem) über den sog. Hauptsatz der Differential- und Integralrechnung eng miteinander verquickt sind. Wir werden auch sehen, wie man das Umkehrproblem der Differentialrechnung, d.h. zu einer vorgegebenen Funktion $f : D \to \mathbb{R}$(oder $\mathbb{C}$) eine *Stammfunktion* zu finden, d.h. eine differenzierbare Funktion $F : D \to \mathbb{R}$(oder $\mathbb{C}$) mit $F' = f$, zumindest für *stetige* F auf Intervallen lösen kann. Der Hauptsatz wird uns auch schlagkräftige *Methoden zur Integralberechnung* zur Verfügung stellen.
Dass Integration und Differentiation Umkehroperationen von einander sind, kommt insbesondere in der *algebraischen Formulierung* des Hauptsatzes zum Ausdruck. Schließlich werden wir und mit einigen Anwendungen des Hauptsatzes beschäftigen.

21.1 Der Hauptsatz der Differential- und Integralrechnung

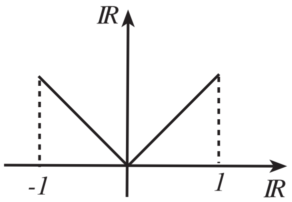
Ist $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$(oder $\mathbb{C}$) eine Funktion, dann ist bis jetzt nicht klar, welche Eigenschaften f erfüllen muss, damit es eine differenzierbare Funktion $F : D \to \mathbb{R}$(oder $\mathbb{C}$) mit $F' = f$ gibt (Problem der Existenz einer Stammfunktion). Das „abschreckende Beispiel" aus ??? zeigt, dass auch unstetige Funktionen Stammfunktionen haben können. Natürlich kann man in konkreten Fällen mit den Rechenregeln für Ableitungen sofort bestätigen, dass eine Funktion F Stammfunktion einer Funktion f ist.
Wählt man z.B. $D = \mathbb{R}$ und $f := \cos$ und $F := \sin$, dann ist $F' = \sin' = \cos$, also ist der Sinus (und zwar auf ganz $\mathbb{R}$) eine Stammfunktion von $\cos$. Sicherlich gibt es nicht für jede Funktion $f : D \to \mathbb{R}$ eine Stammfunktion, denn nach dem Satz von Darboux (vergl. ???), muss f jedes Teilintervall $D_0 \subset D$ wieder auf ein Intervall abbilden. So kann etwa die Treppenfunktion $t : [-1, 1] \to \mathbb{R}$ mit

$$t(x) = \begin{cases} 1, & \text{für } 0 \leq x \leq 1; \\ -1, & \text{für } -1 \leq x < 0 \end{cases}$$



keine Stammfunktion haben (obwohl t als Treppenfunktion und damit als Regelfunktion) natürlich integrierbar ist (das Integral ist Null).
Ein Kandidat für eine Stammfunktion wäre

$$\begin{aligned} F : [-1, 1] &\to \mathbb{R} \\ x &\mapsto |x|, \end{aligned}$$



also die Einschränkung der Betrags auf $[-1,1]$, aber F ist in 0 nicht differenzierbar.
Fasst man den Stammfunktionsbegriff allgemeiner (vergl. Königsberger, Analysis 1, 5.Auflage., Seite ???), lässt also abzählbare Ausnahmemengen $A \subset D$ zu, in denen F nicht differenzierbar zu ein braucht und fordert von einer Stammfunktion $F : D \to \mathbb{R}$ die Eigenschaften

(a) F ist stetig;
(b) In $D - A$ ist F differenzierbar und für $x \in D - A$ gilt $F'(x) = f(x)$;

dann wäre F mit $F(x) = |x|$ für $-1 \leq x \leq 1$ eine Stammfunktion von f auf $[-1,1]$. Nach dem Zwischenwertsatz haben aber stetige reellwertige Funktionen auf Intervallen die Zwischenwerteingenschaft („das stetige Bild eines Intervalls ist eine Intervall“), wenn wir daher f als stetig voraussetzen, widerspricht der Satz von Dauboux auf jeden Fall nicht der möglichen Existenz einer Stammfunktion. Der Hauptsatz besagt u.a., dass die Stetigkeit auch hinreichend für Existenz einer Stammfunktion ist und dass umgekehrt einer Regelfunktion f, die auf einem kompakten Intervall $[a,b]$ eine Stammfunktion besitzt, auch stetig ist. Wir wollen folgende Vereinbarungen treffen:

Ist $D \subset \mathbb{R}$ ein echtes Intervall, und $f : D \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$) eine Funktion, dann heißt f Regelfunktion, wenn die Einschränkung von f auf *jedes* kompakte Teilintervall $[a,b] \subset D$ eine Regelfunktion (im bisherigen Sinne) ist.

21.1.1 Theorem (Hauptsatz der Differential- und Integralrechnung, 1. Version)

(a) Ist $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$(oder $\mathbb{C}$) eine Regelfunktion, dann besitzt f genau dann eine Stammfunktion, wenn f stetig ist.
Ist f stetig, $a \in D$, so ist *eine* Stammfunktion $F : D \to \mathbb{R}$ gegeben durch die Integralfunktion $x \mapsto \int\limits_a^x f(t)dt$.

Zusatz:
$\{F + c;\ c \in \mathbb{R}\}$ ist dann die Menge *aller* Stammfunktionen von f.

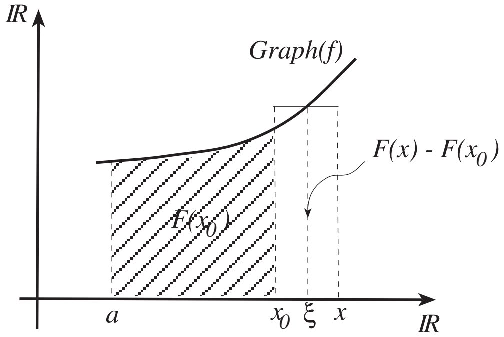
(b) Ist G eine *beliebige* Stammfunktion von $f \in C([a,b])$, dann gilt

$$\int\limits_a^b f(t)dt = G(b) - G(a)$$

Beweis: (a) Sei zunächst f stetig in $x_0 \in D$ und $F(x) := \int\limits_a^x f(t)dt$.
Wir zeigen, dass F in x_0 differenzierbar ist und $F'(x) = f(x_0)$ gilt. Wenn f in jedem Punkt stetig ist, ist also gezeigt, dass f eine Stammfunktion auf D besitzt, nämlich F.
Wir geben zwei Beweisvarianten:

Bei der ersten wird der 1. Mittelwertsatz der Integralrechnung (vgl. ???) verwendet:
Nach diesem gibt es ein ξ in $[b,x]$ bzw. $[x,b]$ (m.a.W. $\xi = x_0 + \vartheta(x - x_0),\ 0 \leq \vartheta \leq 1$)

$$F(x) - F(x_0) = \int\limits_a^x f(t)dt - \int\limits_a^{x_0} f(t)dt = \int\limits_{x_0}^x f(t)dt = f(\xi)(x - x_0)$$



Für $x \neq x_0$ folgt

$$\frac{F(x) - F(x_0)}{x - x_0} = f(\xi).$$

Für $x \to x_0$ folgt dabei wegen der Stetigkeit von f in x_0

$$F'(x_0) = \lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0)$$

Wegen der Wichtigkeit des Satzes geben wir einen zweiten Beweis ($\varepsilon - \delta-$Beweis):
Dabei schreiben wir die Differenz $F(x) - F(x_0)$ in der Form

$$\begin{aligned}\int_{x_0}^{x} f(t)dt &= \int_{x_0}^{x} f(x_0)dt + \int_{x_0}^{x} (f(t) - f(x_0))dt \\ &\qquad \uparrow \text{ man beachte: der Integrand } x_0 \text{ ist konstant} \\ &= (x - x_0)f(x_0) + (x - x_0)\varrho(x)\end{aligned}$$

mit

$$\varrho(x) := \begin{cases} \frac{\int_{x_0}^{x} (f(t) - f(x_0))dt}{x - x_0}, & \text{falls } x \neq x_0; \\ 0, & \text{falls } x = x_0. \end{cases}$$

Damit haben wir also

$$F(x) - F(x_0) = (x - x_0)f(x_0) + (x - x_0)\varrho(x).$$

Nach dem Äquivalenzsatz über Differenzierbarkeit müssen wir lediglich noch nachweisen, dass ϱ an der Stelle x_0 stetig ist (denn dann ist $F'(x_0) = f(x_0)$).
Sei $\varepsilon > 0$ beliebig vorgegeben. Wegen der Stetigkeit von f in x_0 gibt es dazu ein $\delta > 0$, mit

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

für alle $t \in D$ mit $|t - x_0| < \delta$.
Dann gilt zunächst für $x > x_0$

$$\left| \int_{x_0}^{x} (f(t) - f(x_0))dt \right| \le \int_{x_0}^{x} |f(t) - f(x_0)|dt \le \int_{x_0}^{x} \varepsilon \, dt = \varepsilon |x - x_0|$$

Diese Abschätzung funktioniert genauso, wenn $x < x_0$ sein sollte.
Insgesamt folgt also $|\varrho(x)| \le \varepsilon$ für alle $x \in D$ mit $|x - x_0| \le \delta$. ϱ ist also stetig in x_0, damit gilt $F'(x_0) = f(x_0)$.

Damit haben wir also gezeigt:

$\boxed{\text{Jede stetige Funktion auf einem Intervall besitzt eine Stammfunktion.}}$

Jetzt sei umgekehrt $f : D \to \mathbb{R}$ eine Regelfunktion, die eine Stammfunktion F besitzt. Wir zeigen die Stetigkeit von f in $x_0 \in D$ in folgender Weise:
Da f eine Regelfunktion ist, existieren in jedem inneren Punkt $x_0 \in D$ die links- und rechtsseitigen Grenzwerte $f(x_{0_-})$ und $f(x_{0_+})$ (in den Randpunkten nur die einseitigen).
Sei (x_n) eine Folge mit $x_n \in D$ und $x_n > x_0$ und $\lim\limits_{n\to\infty} x_n = x_0$. Nach dem MWSD gibt es dann ein $\xi \in]x_0, x_n[$ mit

$$\frac{F(x) - F(x_0)}{x - x_0} = F'(\xi_n) = f(\xi_n),\ x_0 < \xi_n < x_n.$$

Die linke Seite konvergiert nach Voraussetzung gegen $F'(x_0) = f(x_0)$, die rechte Seite aber gegen $f(x_{0_+})$, also folgt

$$f(x_0) = f(x_{0_+}).$$

Analog zeigt man $f(x_0) = f(x_{0_-})$ (falls x_0 nicht linker Randpunkt von D ist.
Insgesamt gilt also: f ist stetig in x_0. Da wir schon wissen (vgl. ???), dass wenn eine Funktion $f : D \to \mathbb{K}$ eine Stammfunktion F besitzt, eine weitere Funktion $G : D \to \mathbb{K}$ *genau dann* auch Stammfunktion von f ist, wenn $G = F + c$ mit irgendeiner Konstanten Funktion gilt, ist die Aussage über die Menge aller Stammfunktionen klar.

(b) Nach (a) ist für stetige $f : [a, b] \to \mathbb{K}$ die Integralfunktion

$$\begin{aligned} F : [a,b] &\to \mathbb{K} \\ x &\mapsto \int_a^x f(t)dt \end{aligned}$$

eine Stammfunktion von f mit $F(a) = \int_a^a f = 0$. Für eine beliebige Stammfunktion G von f gilt nach (a) dann $G = F + c$ mit einer Konstanten c und damit

$$\begin{aligned} G(b) - G(a) &= (F(b) + c) - (F(a) + c) \\ &= F(b) - F(a) \\ &= F(b) = \int_a^b f. \end{aligned}$$

$\square$

21.1.2 Bemerkungen

(a) Der Hauptsatz garantiert die *Existenz* von Stammfunktionen zu allen stetigen Funktionen auf Intervallen. Darin liegt seine große theoretische Bedeutung. Für konkrete Funktionen kann es jedoch *schwierig* oder *sogar unmöglich* sein, Stammfunktionen mit Hilfe bereits bekannter Funktionen auszudrücken.

(b) Diese Phänomen kann man schon an folgenden Beispielen sehen:
Betrachtet man

$$\begin{aligned} j :]0, \infty[&\to \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned}$$

dann ist der Logarithmus

$$L := \log :]0, \infty[\to \mathbb{R}$$

eine Stammfunktion von j:

$$L'(x) = \frac{1}{x} = j(x)$$

für alle $x \in]0, \infty[$. Man beachte hier zum einen: j ist eine rationale Funktion, L aber nicht (warum?). Der Übergang zu Stammfunktionen kann also durchaus aus einer Klasse von Funktionen herausführen. Zum anderen: Wenn wir den Logarithmus nicht schon gekannt hätten, der ja als Umkehrung von exp definiert wurde, hätten wir keine Stammfunktion explizit angeben können.
Wie in Aufgabe ??? von Blatt ??? zu sehen ist, könnte man sogar umgekehrt $L := \log :]0, \infty[\to \mathbb{R}$ durch $x \mapsto L(x) = \int\limits_1^x \frac{1}{t} dt$ definieren, Eigenschaften von L beweisen (Differenzierbarkeit, strenge Monotonie, Funktionalgleichung) und dann exp als Umkehrfunktion von L definieren.
Diesen Weg in der Schulmathematik einzuschlagen, wurde von Felix Klein angeregt. Nach den gängigen Lehrplänen ist dies aber kaum möglich, da die Exponentialfunktion (Wachstum und Zerfallsprozesse) sogar auch in anderen Fächern (Physik, Biologie) benötigt wird, zu einem Zeitpunkt also, wo der Integralbegriff noch nicht zur Verfügung steht.
Nach den Rechenregeln für Ableitungen hat arcsin in $]-1, 1[$ die Ableitung

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

Auch hier könnte man umgekehrt vorgehen und für $x \in]-1, 1[$

$$a(x) := arcsin(x) := \int\limits_0^x \frac{1}{\sqrt{1-t^2}} dt$$

definieren und dann den Sinus als Umkehrfunktion. Die Kreiszahl π wird in diesem Fall als $\frac{\pi}{2} := \lim\limits_{\substack{x \to 1 \\ x < 0}} a(x)$ definiert.

Dieses Vorgehen hat Vorteile: Man hat gleich die richtige Interpretation von x als Bogenlänge. Der entscheidende Nachteil ist, dass man zur Definition das Integral braucht, das man auch noch in die Punkte $x = -1$ und $x = 1$ fortsetzen muss. Außerdem ist die Periodizität von sin nicht offensichtlich: Man muss sin zu einer periodischen Funktion auf ganz $\mathbb{R}$ fortsetzen.

(c) Die große *praktische Bedeutung* des Hauptsatzes liegt darin, dass es eine äußerst elegante und bequeme Methode darstellt, Integrale zu berechnen (deren Existenz ohnehin gesichert ist), wenn man eine explizite Stammfunktion kennt: es ist ja

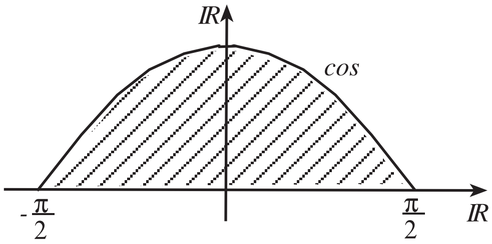
$$(*) \qquad \int\limits_a^b f(x)dx = G(b) - G(a)$$

mit irgendeiner Stammfunktion G von f.
Man führt für die rechte Seite von (x) auch die Notation

$$G(x)|_a^b \quad \text{oder} \quad [G(x)]_a^b$$

ein. So ist z.B.

$$\int\limits_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(x)dx = \sin(x)|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \sin(\pi) - \sin(-\frac{\pi}{2}) = -1 - (-1) = 2.$$



Die in ??? mit etwas Mühe berechneten Integrale fallen uns nun in den Schoss, etwa

$$\int_a^b \exp(x)dx = \exp(b) - \exp(a)$$

oder

$$\int_a^b x^\alpha \, dx = \left.\frac{x^\alpha + 1}{\alpha + 1}\right|_a^b = \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha + 1} \quad (\alpha \neq 1, \ a, b > 0)$$

(d) Für *eine* Stammfunktion F von f wird in der Literatur häufig auch das Symbol

$$\int f(x)dx$$

verwendet und „unbestimmtes Integral“ von f genannt (im Gegensatz zum „bestimmten“ Integral $\int_a^b f(x)dx$ mit festen Grenzen).

Man beachte aber, dass in der Literatur unter diesem Symbol auch häufig die *Menge aller Stammfunktionen* von f verstanden wird.

Aus

$$\int x^3 dx = \frac{1}{4}x^4 \quad \text{und} \quad \int x^3 dx = \frac{1}{4}x^4 + 17$$

kann man jedoch nicht auf $\frac{x^4}{4} = \frac{x^4}{4} + 17$ für alle $x \in \mathbb{R}$ schließen. Also ist

$$\int f(x)dx = F(x)$$

zu lesen: F ist eine Stammfunktion von f. Der Nachteil dieser Schreibweise ist auch, dass kein Definitionsbereich angegeben ist. Aus $\int f(x)dx)F(x)$ und $\int f(x) = G(x)$ darf man lediglich $G = F + c$ schließen.

Aus jedem aus der Differentialrechnung gewonnenen Resultat, ergibt sich umgekehrt eine Aussage über Integrale (sofern die Ableitungen stetig sind). Bevor wir eine Liste von *Grundintegralen* zusammenstellen, seien noch *weitere Versionen des Hauptsatzes* aufgeführt:

Ist $f : D \to \mathbb{R}$ stetig differenzierbar, dann liefert der Hauptsatz für $x \in D$ und $x_0 \in D$

$$f(x) = f(x_0) + \int_{x_0}^x f'(t)dt,$$

speziell für stetig differenzierbares $f : [a, b] \to \mathbb{R}$

$$f(b) = f(a) + \int_a^b f'(t)dt.$$

Man gewinnt also die Funktion f (bis auf eine Konstante) aus ihrer Ableitung („Änderungsrate“) zurück.

Diese Version des Hauptsatzes wird von vielen Autoren als das wichtigere Resultat betrachtet (wegen weittragenden Anwendungen in den Naturwissenschaften).

21.1.3 Theorem (2.Version des Hauptsatzes)

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f \in C^1(D)$, dann gilt für alle $x, x_0 \in D$

$$\boxed{f(x) = f(x_0) + \int\limits_{x_0}^{x} f'(t)dt}$$

speziell für $[a, b] \subset D$ ist

$$\boxed{f(b) = f(a) + \int\limits_{a}^{b} f'(t)dt}$$

Wir geben noch eine weitere Formulierung des Hauptsatzes in der Sprache der Linearen Algebra, indem wir Ableitung und Integration als lineare Operatoren (lineare Abbildungen) auffassen. Diese Version zeigt überdeutlich, dass Integration und Differentiation Umkehroperationen voneinander sind.
Für ein kompaktes Intervall $[a, b] \subset \mathbb{R}$ ($a < b$) betrachten wir die $\mathbb{R}$-Vektorräume

$$V = \{f : [a, b] \to \mathbb{R},\ f \text{ stetig } \} =: C([a, b])$$

und den Untervektorraum

$$W = \{f : [a, b] \to \mathbb{R},\ f\text{stetig differenzierbar}\} =: C^1([a, b]).$$

Man definiert den Ableitungsoperator $\mathsf{D} : W \to V$ durch $f \mapsto \mathsf{D}f$ mit $(\mathsf{D}f)(x) = f'(x)$.
Nach den Rechenregeln ??? ist D ein linearer Operator (eine lineare Abbildung). Weiter definiert man

$$\begin{aligned} I : V &\to W \\ f &\mapsto If \ \text{ mit } \ (If)(x) = \int\limits_a^x f(t)dt. \end{aligned}$$

Nach der ersten Version des Hauptsatzes liegt $F := If$ in W, ferner ist nach Grundeigenschaften des Integrals I auch eine lineare Abbildung (Integrationsoperator) und es gilt wegen $\mathsf{D}F = \mathsf{D}(If) = f$.

$$(*) \qquad \mathsf{D} \circ I = id_V .$$

$$V \underset{\mathsf{D}}{\overset{I}{\rightleftarrows}} W$$

Startet man umgekehrt mit einem $f \in W$, dann erhält man nach der 2. Version des Hauptsatzes

$$((I \circ \mathsf{D})f)(x) = (I(\mathsf{D}f))(x) = \int\limits_a^x (\mathsf{D}f)(t)dt = \int\limits_a^x f'(t) = f(x) - f(a).$$

Man erhält also f bis auf eine Konstante zurück. Will man den Sachverhalt, dass Integration und Differentiation wirklich Umkehroperationen voneinander sind (also Bijektionen), dann muss man den Unterraum $W_0 := \{f \in W;\ f(a) = 0\}$ von W einführen. Für $f \in W_0$ gilt dann auch

$$(**) \qquad I \circ \mathsf{D} = id_{W_0}$$

Da $\mathsf{D} \circ I = id_V$ immer gilt, haben wir die folgende algebraische Formulierung des Hauptsatzes mit Hilfe linearer Operatoren:

21.1.4 Theorem (algebraische Form des Hauptsatzes)

Die Abbildung

$$\begin{aligned} I; V &\to W_0 \\ f &\mapsto If \end{aligned}$$

ist ein Isomorphismus mit der Umkehrabbildung

$$\begin{aligned} \mathsf{D} : W_0 &\to V \\ f &\mapsto \mathsf{D}f. \end{aligned}$$

21.1.5 Bemerkung

Der Vektorraum $V = C([a, b])$ ist isomorph zum echten Unterraum W_0. Aus der Linearen Algebra sollte man wissen, dass ein *endlich dimensionaler* Vektorraum *nie* zu einem echten Unterraum isomorph sein kann.

Wir stellen eine (sehr kleine) Liste von Stammfunktionen stetiger Funktionen zusammen, links steht eine stetige Funktion f, in der Mitte eine Stammfunktion von f, in der dritten Spalte stehen Angaben zum Gültigkeitsbereich:
Diese Liste sollte man beherrschen ohne ein Tafelwerk oder ein CAS zur Hilfe zu nehmen. In klassischen Tafelwerken findet man Tausende von Stammfunktionen.

21.1.6 Eine kleine Liste von Stammfunktionen (Grundintegralen)

$f(x)$	$F(x) = \int f(x)dx$	Gültigkeitsbereich, Bemerkungen
$x^k,\ k \in \mathbb{Z},\ k \neq -1$	$\frac{x^{k+1}}{k+1}$	$x \in \mathbb{R}$ für $k \geq 0,\ x \neq 0$ für $k \leq -2$
$x^\alpha,\ \alpha \in \mathbb{R},\ \alpha \neq -1$	$\frac{x^{\alpha+1}}{\alpha+1}$	$x > 0$
$\frac{1}{x}$	$\log \lvert x \rvert$	$x > 0$ oder $x < 0$
$\sin(x)$	$-\cos(x)$	$x \in \mathbb{R}$
$\cos(x)$	$\sin(x)$	$x \in \mathbb{R}$
$\tan(x)$	$-\log \lvert \cos(x) \rvert$	$x \neq (k + \frac{1}{2})\pi,\ k \in \mathbb{Z}$
$\cot(x)$	$\log \lvert \sin(x) \rvert$	$x \neq k\pi,\ k \in \mathbb{Z}$
$\frac{1}{\cos^2(x)} = 1 + \tan^2(x)$	$\tan(x)$	$x \neq (k + \frac{1}{2})\pi, k \in \mathbb{Z}$
$\frac{1}{\sin^2(x)}$	$-\cot(x)$	$x \neq k\pi, k \in \mathbb{Z}$
$a^x\ (a \neq 1)$	$\frac{a^x}{\log a}$	$x \in \mathbb{R}$
$\exp(ax),\ a \in \mathbb{C},\ a \neq 0$	$\frac{1}{a}\exp(ax)$	$x \in \mathbb{R}$
$\exp(ix)$	$-i\exp(ix)$	$x \in \mathbb{R}$
$\sinh(x)$	$\cosh(x)$	$x \in \mathbb{R}$
$\cosh(x)$	$\sinh(x)$	$x \in \mathbb{R}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$	$\lvert x \rvert < 1$
$\frac{1}{\sqrt{1+x^2}}$	Arsinh(x)	$x \in \mathbb{R}$
$\frac{1}{\sqrt{x^2-1}}$	$\log \lvert x + \sqrt{x^2-1} \rvert$	$\lvert x \rvert > 1$
$\frac{1}{1+x^2}$	$\arctan(x)$	$x \in \mathbb{R}$
$\frac{f'(x)}{f(x)}$	$\log \lvert f(x) \rvert$	Intervalle, in denen $f(x) \neq 0$ ist.
$a_0 + a_1x + \ldots + a_nx^n$ $(n \in \mathbb{N}_0, a_j \in \mathbb{R})$	$a_0x + \frac{a_1x^2}{2} + \ldots + \frac{a_n}{n+1}x^{n+1}$	$x \in \mathbb{R}$
$\sum\limits_{k=0}^{\infty} a_kx^k\ (a_k \in \mathbb{R})$	$\sum\limits_{k=0}^{\infty} \frac{a_k}{k+1}x^{k+1}$	x im Konvergenzintervall der Potenzreihe

21.1.7 Bemerkung

Verschafft man sich Stammfunktionen mit Hilfe eines Computeralgebrasystems, etwa Maple oder Mathematica, so sollte man zur Sicherheit immer den Differentiations-Operator D anwenden, um zu testen, ob tatsächlich die Ausgangsfunktion herauskommt.
Mit Hilfe des Hauptsatzes lassen sich Aussagen der Differentialrechnung in Aussagen der Integralrechnung transformieren. Hierauf beruhen die folgenden *Integrationstechniken*.

21.2 Integrationstechniken, erste Anwendungen des Hauptsatzes

„Differentiation ist Technik, Integration ist Kunst" (N..N.)

Auch im Zeitalter der Coputeralgebrasysteme, in dem einem etwa für die Funktion f mit

$$f(x) = \frac{x^5 - x^4 + 1}{x^2\sqrt{x^2-1}} \qquad (x \neq 0,\ x \neq \pm 1)$$

in Bruchteilen von Sekunden eine Stammfunktion angeboten wird, etwa F mit

$$F(x) = x^2\sqrt{x^2-1} - \frac{2}{3}(x^2-1)^{3/2} - \frac{1}{2}x\sqrt{x^2-1} + \frac{1}{2}\log(x+\sqrt{x^2-1}) + \frac{\sqrt{x^2-1}}{x},$$

sollte jede(r) einige *Grundtechniken zur Integralrechnung* beherrschen. Zwei dieser Techniken beruhen auf der Umkehrung der *Produktregel* bzw. der *Kettenregel* der Differentiation. Man kann sie für das unbestimmte oder das bestimmte Integral formulieren. Diese Regeln ermöglichen es in Fällen, in denen eine Stammfunktion nicht unmittelbar erkennbar ist, den Integranden in geeigneter Form in der Hoffnung umzuformen, doch eine Stammfunktion zu finden oder ein bestimmtes Integral explizit zu berechnen.

21.2.1 Satz (Partielle Integration oder Produktintegration)

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f : D \to \mathbb{R}$(oder $\mathbb{C}$) stetig und F eine Stammfunktion von f(auf D) und $g : D \to \mathbb{R}$ (oder $\mathbb{C}$) sei stetig differenzierbar.
Dann gilt

$$\int f(x)g(x)dx = F(x)g(x) - \int F(x)g'(x)dx.$$

Ist $D = [a, b]$, dann gilt

$$\int_a^b f(x)g(x)dx = F(x)g(x)\Big|_a^b - \int_a^b F(x)g'(x)dx.$$

21.2.2 Korollar

Sind $u, v : D \to \mathbb{R}$(oder $\mathbb{C}$) stetig differenzierbar, dann gilt

$$\int u'(x)v(x)dx = u(x)v(x) - \int u(x)v'(x)dx$$

Kürzer: $\int vdu = uv - \int udv$ und für $D = [a, b]$ gilt

$$\boxed{\int_a^b u'(x)v(x)dx = u(x)v(x)\Big|_a^b - \int_a^b u(x)v'(x)dx}$$

Der Beweis von 22.2.1 ergibt sich unmittelbar aus der Produktregel:

$$(Fg)' = F'g + Fg' = fg + Fg',$$

also ist Fg eine Stammfunktion von $fg+Fg'$ und daher (die Bildung von Stammfunktionen ist eine lineare Operation)

$$\int f(x)g(x)dx = F(x)g(x) - \int F(x)g'(x)dx.$$

Das Korollar ergibt sich mit $u' = f,\ F = u$ und $g = v$ oder auch direkt aus $(uv)' = u'v + uv'$.

21.2.3 Beispiel

Zu berechnen sei $\int x\exp(x)dx$.

1.Ansatz: $u'(x) = x,\ v(x) = \exp(x)$.
Dann ist $x \to \frac{1}{2}x^2$ eine Stammfunktion von u' und $v'(x)\exp(x)$. Folglich ist

$$\int x\exp(x)dx = \frac{1}{2}x^2\exp(x) - \int \frac{x^2}{2}\exp(x)dx.$$

Das letzte Integral ist aber komplizierter als das Ausgangsintegral, der Ansatz hat keine Vereinfachung gebracht.

2.Ansatz: $u'(x) = \exp(x),\ v(x) = x$, also $u(x) = \exp(x)$ und $v'(x) = 1$.
Man erhält so

$$\int x\exp(x) = x\exp(x) - \int 1\cdot\exp(x)dx = x\exp(x) - \exp(x) = (x-1)\exp(x),$$

also ist $x \mapsto (x-1)\exp(x)$ eine Stammfunktion von $x \mapsto x\exp(x)$ (auf ganz $\mathbb{R}$).

21.2.4 Bemerkungen und Beispiele

(a) Strategie bei der Anwendung der Methode der partiellen Integration sollte es sein, auf der rechten Seite der Formel im nicht ausintegrierten Bestandteil ein Integral zu erhalten, das einfacher ist als das Ausgangsintegral. Das erfordert ein gewisses Geschick und Übung. Ferner ist es zweckmäßig im Ausgangsintegral, den Faktor der im Laufe der Rechnung differenziert wird mit einem nach unten weisenden Pfeil („↓“) und den Faktor, der integriert wird, mit einem nach oben weisenden Pfeil („↑“) zu bezeichnen. Nach Möglichkeit sollte man den Faktor differenzieren, der bei der Differetiation einfacher wird und den Faktor integrieren, der sich bei der Integration wenigstens nicht allzu verkompliziert.

(b) Um eine Stammfunktion von $x \mapsto \log x$ $(x > 0)$ also $\int \log x$, zu berechnen, beachte man $1\cdot\log x = \log x$. So erhält man

$$\begin{aligned}\int \log x dx = \int \underset{\uparrow}{1}\underset{\downarrow}{\log x} dx &= x\log x - \int x\frac{1}{x}dx\\ &= x\log x - \int 1\ dx\\ &= x\log x - x\\ &= x(\log x - 1).\end{aligned}$$

(c) Bei Integration vom Typ

$$\int x^n\exp(\lambda x)dx\ (\lambda\in\mathbb{C})\ \text{ oder } \int x^n\sin x dx \text{ oder } \int x^n\cos x dx, \qquad n\in\mathbb{N}_0,$$

ergeben sich so Rekursionsformeln. Ist z.B.

$$I_n := \int \underset{\downarrow}{x^n}\underset{\uparrow}{\exp}(x)dx = x^n\exp(x) - n\int x^{n-1}\exp(x)dx = x^n\exp(x) - nI_{n-1}.$$

Diese Rekursionsformeln führen I_n schließlich auf $I_0 = \int\exp(x)dx = \exp(x)$ zurück.

(d) Für $m, n\in\mathbb{N}_0$ sei $I_{n,m} = \int\limits_0^1 x^n(1-x)^m dx$ zu berechnen.

Partielle Integration $I_{n,m} = \int\limits_0^1 \underset{\uparrow}{x^n}(1-\underset{\downarrow}{x})^m dx$ ergibt die Rekursionsformel

$$I_{n,m} = \frac{m}{n+1}I_{n+1,m-1}.$$

Mit $I_{n+m,0} = \frac{1}{m+n+1}$ folgt schließlich

$$I_{n,m} = \frac{n!m!}{(n+m+1)!}:$$

damit kann man auch das Integral $I := \int\limits_{-1}^{1} (1+x)^n(1-x)^m dx$ leicht berechnen.

Setzt man $x := 2t - 1$, so erhält man mit der Substitutionsformel (vgl. 22.2.1), die wir hier vorwegnehmen,

$$I = 2^{n+m+1} \int\limits_0^1 t^n(1-t^m)dt = 2^{n+m+1} I_{n,m}.$$

(e) Um $\int \arctan x dx$ zu berechnen, beachten wir wieder $1 \cdot \arctan x = \arctan x$ und erhalten

$$\begin{aligned}\int \arctan x dx = \int \underset{\uparrow}{1} \cdot \underset{\downarrow}{\arctan} x dx &= x \arctan x - \int \frac{x}{1+x^2} dx \\ &= x \arctan x - \frac{1}{2} \int \frac{2x}{1+x^2} dx \\ &= x \arctan x - \frac{1}{2} \log(1+x^2).\end{aligned}$$

Man beachte, dass im letzten Integral $\int \frac{2x}{1+x^2} dx$ im Zähler die Ableitung des Nenners steht, ein solches Integral ist ein Grundintegral (vgl. 20.1.6).

(f) Für $x \in [-1,1]$ gilt

$$\int \sqrt{1-x^2} dx = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x).$$

Zunächst sei $x \in]-1,1[$, dann erhält man

$$\begin{aligned}I := \int \sqrt{1-x^2} dx = \int \underset{\uparrow}{1} \cdot \underset{\downarrow}{\sqrt{1-x^2}} dx &= x\sqrt{1-x^2} - \int \frac{x(-2x)}{2\sqrt{1-x^2}} dx \\ &= x\sqrt{1-x^2} + \int \frac{x^2}{\sqrt{1-x^2}} dx \\ &= x\sqrt{1-x^2} + \int \frac{1}{\sqrt{x^2-1}} dx - \int \frac{1-x^2}{\sqrt{1-x^2}} dx \\ &= x\sqrt{1-x^2} + \arcsin x - \int \sqrt{1-x^2} dx \\ &= x\sqrt{1-x^2} + \arcsin x - I,\end{aligned}$$

also

$$2I = x\sqrt{1-x^2} + \arcsin x \text{ oder } I = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x).$$

Da die Ableitungsregel $\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$ nur für $x \in]-1,1[$ gilt, gilt die abgeleitete Formel zunächst nur für die x im offenen Intervall $]-1,1[$.

Dass sie auch für die Randpunkte, also für alle $x \in [-1,1]$ gilt, folgt aus dem folgenden allgemeinen bemerkenswerten

Hilfssatz:

Ist $D \subset \mathbb{R}$ ein echtes Intervall und sind f und G stetige Funktionen auf D und ist $a \in D$ ein fester Punkt. Für alle $x \in D$, $x \neq a$, sei G differenzierbar bei x und $G'(x) = f(x)$. Dann ist G auch in a differenzierbar und es gilt $G'(a) = f(a)$.

Zum Beweis beachten wir, dass f nach dem Hauptsatz eine Stammfunktion $F : D \to \mathbb{R}$

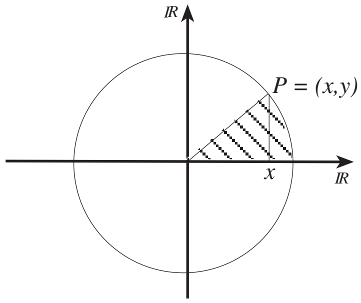
besitzt. Die Menge $D - \{a\}$ besteht entweder aus zwei getrennten Intervallen oder ist (im Fall, dass a ein Randpunkt ist) ein Teilintervall von D. Auf diesem Intervall ist die Differenz $G - F$ jeweils konstant, aber weil beide Funktionen auf D stetig sind, ist $G - F$ insgesamt konstant, etwa $G - F = c$ oder $G = F + c$. Da F in a differenzierbar ist und $F'(a) = f(a)$ gilt, ist auch G in a differenzierbar und es gilt $G'(a) = f(a)$.

Wendet man diesen Hilfssatz auf $f(x) = \sqrt{1-x^2}$ und $G(x) = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)$ an (einmal für $a = 1$ und einmal für $a = -1$), erhält man für alle $x \in [-1, 1]$

$$\boxed{\int \sqrt{1-x^2} = \tfrac{1}{2}(x\sqrt{1-x^2} + \arcsin x)}$$

□

(g) **Anwendung: Die Fläche des Einheitskreises**
Als Anwendung von (f) erhalten wir, dass der Flächeninhalt des Einheitskreises π ist. Wir betrachten dazu zunächst den straffierten Sektor in der und nennen seinen Flächeninhalt $\frac{1}{2}F$.



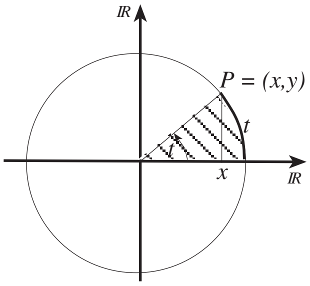
(Es setzt sich aus dem Dreieck OPx und dem durch $\int\limits_x^1 \sqrt{1-t^2}dt$ gegebenen Flächeninhalt zwischen dem Graphen von $x \mapsto \sqrt{1-x^2}$ und der reellen Achse zwischen Punkten x und 1 zusammen).
Wegen der Kreisgleichung $x^2 + y^2 = 1$ gilt

$$\begin{aligned} F = xy + 2\int\limits_x^1 \sqrt{1-t^2}dt &= x\sqrt{1-x^2} + \left(t\sqrt{1-t^2} + \arcsin t\right)\Big|_x^1 \\ &= \arcsin 1 - \arccos x \\ &= \frac{\pi}{2} - \arcsin x = \arccos x. \end{aligned}$$

Speziell für $x = -1$ erhält man wegen $\arccos(-1) = \pi$ den Flächeninhalt des Einheitskreises (besser der Einheitskreisscheibe).

$$\boxed{\text{Der Flächeninhalt der Einheitskreisscheibe ist also } \pi.}$$

Man beachte, dass wir $\frac{\pi}{2}$ als kleinste positive Nullstelle den $\cos : \mathbb{R} \to \mathbb{R}$ eingeführt haben. π hat also auch eine klassische Bedeutung. Wie wir wissen (und auch nochmals sehen werden) ist 2π die *Länge* der Einheitskreislinie.
Gehört zum Punkt $P \in S^1 := \{(x,y) \in \mathbb{R}^2;\ x^2 + y^2 = 1\}$ der im Bogenmaß gemessene Winkel t, so ist also $\cos t = x$ oder $t = \arccos x$. Der im Bogenmaß gemessene Winkel ist also gleich dem doppelten Flächeninhalt des straffierten Sektors oder:
Der straffierte Sektor hat den Flächeninhalt $\frac{1}{2}t$.



Die die Einheitskreislinie die Länge 2π hat, ergibt sich nochmals, das π der Flächeninhalt der Einheitskreisscheibe ist.

Eine schöne Anwendung der Produktintegration ist der sog. *Wallis'sche Produktformel* für $\frac{\pi}{2}$.

21.2.5 Die Wallische Produktformel für $\frac{\pi}{2}$ (John Wallis, Arithmetica Infiniforum, 1666)

Ist

$$w_n = \frac{2\cdot 2}{1\cdot 3}\cdot\ldots\cdot\frac{2n\cdot 2n}{(2n-1)(2n+1)} = \prod_{k=1}^{n}\frac{(2k)^2}{(2k-1)(2k+1)} = \prod_{k=1}^{n}\frac{4k^2}{4k^2-1}$$

für $n\in\mathbb{N}$, dann ist die Folge (w_n) konvergent und es gilt

$$\boxed{\lim_{n\to\infty} w_n = \lim_{n\to\infty}\prod_{k=1}^{n}\tfrac{4k^2}{4k^2-1} = \tfrac{\pi}{2}}$$

Zum **Beweis** betrachten wir die Folge (I_k), $k\in\mathbb{N}_0$ mit

$$I_k := \int_0^{\frac{\pi}{2}} \sin^k x dx$$

Dann ist $I_0 = \frac{\pi}{2}$ und $I_1 = 1$ und für $k\geq 2$ erhält man

$$\begin{aligned}
I_k &= \int_0^{\frac{\pi}{2}} \underset{\uparrow}{\sin}(x)\underset{\downarrow}{\sin}^{k-1}(x)dx \\
&= -\cos x\sin^{k-1}(x)\Big|_0^{\frac{\pi}{2}} + (k-1)\int_0^{\frac{\pi}{2}}\cos^2(x)\sin^{k-2}(x)dx \\
&= (k-1)\int_0^{\frac{\pi}{2}}(1-\sin^2(x))\sin^{k-2}(x)dx,
\end{aligned}$$

also

$$I_k = (k-1)I_{k-2} - (k-1)I_k \text{ oder } I_k = \frac{k-1}{k}I_{k-2} \text{ für } k\geq 2.$$

Ist k gerade, also $k=2n,\ n\in\mathbb{N}_0$, so endet der iterative Abstieg bei 0 und man erhält wegen $I_0=\frac{\pi}{2}$

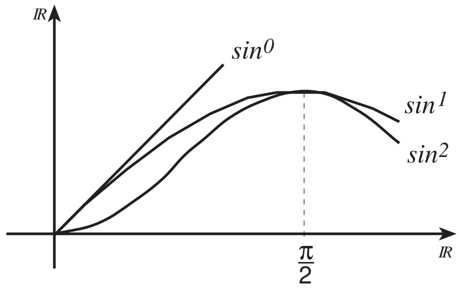
$$I_{2n} = \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\cdot\ldots\cdot\frac{3}{4}\cdot\frac{1}{2}\cdot I_0.$$

Ist jedoch k ungerade; $k=2n+1,\ n\in\mathbb{N}_0$, so ergibt sich analog wegen $I_1=1$

$$I_{2n+1} = \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\cdot\ldots\cdot\frac{4}{5}\cdot\frac{2}{3}\cdot 1.$$

Weil für $x \in [0, \frac{\pi}{2}]$ gilt $\sin^{2n+2} x \leq \sin^{2n+1} x \leq \sin^{2n}$ folgt

$$(*) \qquad I_{2n+2} \leq I_{2n+1} \leq I_{2n}.$$



Nach Definition ist $w_n = \frac{\pi}{2} \frac{I_{2n+1}}{I_{2n}}$. Es bleibt also $\lim\limits_{n\to\infty} \frac{I_{2n+1}}{I_{2n}} = 1$ zu zeigen.
Aus $(*)$ folgt aber

$$(**) \qquad \frac{I_{2n+2}}{I_{2n}} \leq \frac{I_{2n+1}}{I_{2n}} \leq 1$$

und wegen

$$\lim_{n\to\infty} \frac{I_{2n+2}}{I_{2n}} = \lim_{n\to\infty} \frac{2n+1}{2n+2} = \lim_{n\to\infty} \left(1 - \frac{1}{2n+2}\right) = 1$$

folgt nach dem Sandwich-Theorem auch $\lim\limits_{n\to\infty} \frac{I_{2n+2}}{I_{2n}} = 1$, also ist

$$\boxed{\lim_{n\to\infty} w_n = \lim_{n\to\infty} \prod_{k=1}^{n} \frac{4k^2}{4k^2-1} = \frac{\pi}{2}}$$

$\square$

Für den Limes in der Mitte verwendet man auch die Bezeichnung $\prod\limits_{k=1}^{\infty} \frac{4k^2}{4k^2-1}$ und nennt einen solchen Ausdruck auch ein *unendliches Produkt*.
Wegen der Sonderrolle der Null bezüglich der Multiplikation ist die Konvergenztheorie für unendliche Produkte schwieriger als für Reihen. Wir gehen hier nicht darauf ein.

21.2.6 Bemerkung

Die Wallis'sche-Produktformel für $\frac{\pi}{2}$ wird uns nützlich sein im Zusammenhang mit der Theorie der Gamma-Funktion (sie erlaubt eine einfache Bestimmung des Funktionswertes $\Gamma(\frac{1}{2})(= \sqrt{\pi})$ und im Zusammenhang mit der Stirlingschen Formel (Wachstum von $n!$).
Man vergleiche hierzu den Abschnitt 22. Zur numerischen Berechnung von π ist die Wallis'sche Produktformel nicht besonders gut geeignet, da die Folge (w_n) sehr langsam konvergiert:

n	w_n			
1	w_1	$=$	$\frac{4}{3}$	$= 1,333333\ldots$
2	w_2	$=$	$\frac{64}{45}$	$= 1,422222\ldots$
3	w_3	$=$	$\frac{256}{175}$	$= 1,46286\ldots$
4	w_4	$=$	$\frac{16834}{11025}$	$= 1,48608\ldots$
5	w_5	$=$		$= 1,50108\ldots$
10	w_{10}	$=$		$= 1,53385\ldots$
100	w_{100}	$=$		$= 1,56689\ldots$
1000	w_{1000}	$=$		$= 1,57040\ldots$

Der „exakte" Wert von $\frac{\pi}{2}$ ist $\frac{\pi}{2} = 1,5707963267\ldots$.
Immerhin stimmt also w_{1000} in den ersten drei Nachkommastellen mit dem exakten Wert überein.

Die heutigen schnellen Algorithmen zur Berechnung von π liefern auch die Möglichkeit, einzelne (Dezimal- bzw. Hexadezimal-)Stellen „irgendwo“ in π zu berechnen. Eine solche Formel ist etwa

$$\boxed{\pi = \sum_{k=0}^{\infty} \frac{1}{16^k}\left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6}\right)}$$

Diese sog. BBP-Reihe wurde mit dem Computer entdeckt, man kann sie mit Hilfe des Vertauschungssatzes für Integration und Summation relativ einfach beweis. Vergleiche hierzu:

Davis H.Bailey, Jonathan M.Borwein, Peter B.Borwein und Simon Plouffe:
The Quest for Pi, The mathematical Intelligencer,
Vol.19(1997), No.1, 50-56.

oder auch

Hort S. Holdgrün: Analysis, Band 1, Leins-Verlag, 1998,
dort Abschnitt 38.20 und 38.21

Ist s_n die n-te Partialsumme der obigen Reihe, so gilt $|s_n - \pi| < \frac{1}{30(n+1)16^n}$. Berechnet man 25 Summanden der Partialsumme s_{24}, so erhält man

$$s_{24} = 3,14\ 15\ 92\ 653\ 589\ 793\ 238\ 462\ 643\ 383\ 279\ 50\ldots$$

Der „exakte“ Wert von π ist

$$\pi = 3,14\ 15\ 92\ 653\ 589\ 793\ 238\ 462\ 643\ 383\ 279\ 502\ 884\ldots,$$

die ersten 32 Nachkommastellen von s_{24} sind also korrekt.
Die Darstellung von π mit der BBP-Reihe erlaubt es auch, eine beliebige Ziffer in der Hexa-Dezimal-Darstellung von π zu berechnen, ohne dass man die vorausgehende Ziffern zu kennen braucht. dabei genügt etwas Ausdauer und ein einfacher Taschenrechner! Versuchen Sie es!
Nebenbei: Die 20 billionste Hexadezimalstelle von π ist „A“, die 10 milliardste Hexadezimalstelle ist eine „9“.

Die Substitutionsmethode erhält man durch Umkehrung der Kettenregel der Differentiation. Dazu seien $M, N \subset \mathbb{R}$ echte Intervalle, $\varphi : M \to \mathbb{R}$ stetige differenzierbare Funktion mit $\varphi(M) \subset N$ und $F : N \to \mathbb{R}$ stetig differenzierbar. Dazu ist auch die Zusammensetzung $G = F \circ \varphi : M \to \mathbb{R}$ stetig differenzierbar und für $t \in M$ gilt

$$G'(t) = F'(\varphi(t))\varphi'(t).$$

Übergang zu Stammfunktionen auf beiden Seiten liefert

$$\int F'\varphi(t)\varphi'(t)dt = G(t) = F(\varphi(t))$$

Setzt man $x = \varphi(t)$ und $F'(x) = f(x)$, so erhält man

$$\int f(\varphi(t))\varphi'(t)dt = \int f(x)dx|_{x=\varphi(t)},$$

Wobei die Schreibweise rechts bedeuten soll, dass man in der Stammfunktion $F(x)$ und $f(x)$ für x die Funktion $\varphi(t)$ einzusetzen hat.

21.2.7 Satz (Substitutionsregel)

(a) Sind $M, N \subset \mathbb{R}$ echte Intervalle, $\varphi : M \to \mathbb{R}$ stetig differenzierbar mit $\varphi(M) \subset N$ und $f : N \to \mathbb{R}$ stetig, so gilt mit $x = \varphi(t)$

$$\int f(\varphi(t))\varphi'(t)dt = \int f(x)dx\big|_{x=\varphi(t)}$$

(b) Ist $M = [a, b]$, so gilt speziell

$$\int\limits_a^b f(\varphi(t))\varphi'(t)dt = \int\limits_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

Dabei ergibt sich (b) so: Ist F eine Stammfunktion von f, dann ist $F \circ \varphi$ eine Stammfunktion von $(f \circ \varphi)\varphi'$, also

$$\int\limits_a^b f(\varphi(t))\varphi'(t)dt = F(\varphi(t))\big|_a^b = F(\varphi(b)) - F(\varphi(a)) = \int\limits_{\varphi(a)}^{\varphi(b)} f(x)dx$$

21.2.8 Bemerkung

Unter Verwendung der Leibniz'scher Schreibweise (man kann diese durch das Rechnen mit Differentialen rechtfertigen) $dx = d\varphi(t) = \varphi'(t)dt$ lautet die Substitutionsregel einfach

$$\int f(\varphi(t))\varphi'(t)dt = \int f(\varphi(t))d\varphi(t) = \int f(x)dx.$$

Analog für das bestimmte Integral

$$\int\limits_a^b f(\varphi(t))\varphi'(t)dt = \int\limits_a^b f(\varphi(t))d\varphi(t) = \int\limits_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

Man hat einfach x durch $\varphi(t)$ zu ersetzen: Läuft t von a nach b, so läuft $x = \varphi(t)$ von $\varphi(a)$ nach $\varphi(b)$. dabei hat man die Konvention $\int\limits_c^d f = -\int\limits_d^c f$ zu beachten.

21.2.9 Beispiele und Bemerkungen

Die Substitutionsregel lässt sich für unbestimmte und bestimmte Integrale in zwei Richtungen anwenden: Von links nach rechts und von rechts nach links. Wir behandeln jeweils typische Beispiele.

(a) Zu berechnen sei $\int e^{t^2} \cdot tdt$.
Setzt man versuchsweise $x = \varphi(t) = t^2$, so folgt $\varphi'(t) = 2t$, also mit $f(x) = e^x$

$$2\int e^{t^2} \cdot tdt = \int e^x dx \quad (x = t^2),$$

folglich

$$\int e^{t^2} \cdot tdt = \frac{1}{2}\int e^x = \frac{e^x}{2} = \frac{e^{t^2}}{2}.$$

Mit einem scharfer Blick hätte man auch sofort erkennen können, dass $\frac{1}{2}\int e^{t^2} \cdot 2tdt = \frac{1}{2}\int e^x dx = \frac{e^x}{2} = \frac{e^{t^2}}{2}$ ist.

(b) Zu berechnen sei $\int \left(\sin^3 t + e^{\sin t} \cos t dt\right)$.
Man sieht direkt, dass G mit $G(t) = \frac{1}{4}\sin^4 t + e^{\sin t}$ eine Stammfunktion ist, also ist

$$\int (\sin^3 t + e^{\sin t}) \cos t dt = \frac{1}{4}\sin^4 t + e^{\sin t}$$

(c) $I := \int\limits_0^{\frac{\pi}{2}} \sin^3 t \cdot \cos t dt$ ist zu berechnen mit $f(x) = x^3$ und $x = \varphi(t) = \sin t$ erhält man

$$I = \int\limits_0^{\frac{\pi}{2}} \sin^3 t \cos t dt = \int\limits_0^{\frac{\pi}{2}} f(\varphi(t))\varphi'(t) dt = \int\limits_0^1 x^3 dx = \left.\frac{x^4}{4}\right|_0^1 = \frac{1}{4}.$$

(d) Das Integral $I = \int (\cos t + \cos^3 t) dt$ hat zunächst nicht die benötigte Form, um 20.2.7(a) von links nach rechts anwenden zu können. Beachtet man aber

$$I = \int (1 + \cos^2 t) \cos t dt = \int (2 - \sin^2 t) \cos t dt$$

und setzt jetzt $\sin t := \varphi(t) = x$, $dx = \cos t dt$, so erhält man

$$I = \int \left. (2 - x^2) dx\right|_{x=\sin t} = \left.\left[2x - \frac{x^3}{3}\right]\right|_{x=\sin t} = (2\sin t - \frac{1}{3}\sin^3 t)$$

Für das bestimmte Integral $I_0 = \int\limits_{\frac{\pi}{6}}^{\pi} (\cos t + \cos^3 t) dt$ erhält man so z.B.

$$I_0 = \int\limits_{\frac{1}{2}}^{0} (2 - x^2) dx = \left.\left(2x - \frac{x^3}{3}\right)\right|_{\frac{1}{2}}^{0} = -\left(1 - \frac{1}{24}\right) = -\frac{23}{24}.$$

(e) Ist $f = F'$ und $\alpha \neq 0$, so gilt

$$\int f(\alpha t + \beta) dt = \frac{1}{\alpha} \int f(\alpha t + \beta)\alpha dt = \frac{1}{\alpha} \int \left. f(x) dx\right|_{x=\alpha t+\beta} = \frac{1}{\alpha} F(\alpha t + \beta).$$

Hiermit ergibt sich z.B. (für $a \neq 0$)

$$\int \frac{1}{t^2 + a^2} dt = \frac{1}{a^2} \int \frac{1}{\left(\frac{t}{a}\right)^2 + 1} dt = \frac{1}{a^2}(a \arctan \frac{x}{a}) = \frac{1}{a} \arctan \frac{x}{a}$$

und für bestimmte Integrale als Spezialfälle

$(*)$ $\int\limits_a^b f(t + a) dt = \int\limits_{a+c}^{b+c} f(x) dx$ und

$(**)$ $\int\limits_a^b f(\alpha t) dt = \frac{1}{\alpha} \int\limits_{\alpha a}^{\alpha b} f(x) dx \quad (\alpha \neq 0)$

(f) Ist $\varphi : M \to \mathbb{R}$ stetig differenzierbar und $\varphi'(t) \neq 0$ für alle $t \in [a, b]$, dann ist

$$\int \frac{\varphi'(t)}{\varphi(t)} dt = \log|\varphi(t)| \quad (f(x) = \frac{1}{x};\ x = \varphi(t))$$

Für $a, b \in M$ ist also

$$\int\limits_a^b \frac{\varphi'(t)}{\varphi(t)} dt = \log|\varphi(t)|\Big|_a^b$$

(siehe auch die Tabelle der Grundintegrale)

(g) Für $a, b \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ gilt

$$\int_a^b \tan t dt = \int_a^b \frac{\sin t}{\cos t} dt = -\log \cos t\big|_a^b$$

(h) Zu berechnen sei $\int_1^4 (1+x)\sqrt{x}dx$.

Hier wird man versuchen die Substitutionsformel von rechts nach links anzuwenden. Um die Wurzel zu beseitigen, setzen wir $x = \varphi(t) = t^2$, $dx = \varphi'(t)dt = 2tdt$ und rechnen die Grenzen um: $x_1 = 1$ wird für $t_1 = 1$ und $x_2 = 4$ wird für $t_2 = 2$ angenommen. Damit erhalten wir

$$\begin{aligned}\int_1^4 (1+x)\sqrt{x}dx = \int_1^2 (1+t^2)t \cdot 2tdt &= 2\int_1^2 (t^2+t^4)dt \\ &= 2\left(\frac{1}{3}t^3 + \frac{1}{}5t^5\right)\Big|_1^2 \\ &= \frac{14}{3} + \frac{62}{5} = \frac{256}{15}.\end{aligned}$$

(i) Zu berechnen sei $I = \int_1^2 x\sqrt{5x-1}dx$.

Wir wollen φ so bestimmen, dass $\sqrt{5\varphi(t)-1} = t$ wird. Dazu lösen wir $\sqrt{5x-1} = t$ nach x auf und erhalten

$$x = \frac{1}{5}(t^2+1),\ dx = \frac{2}{5}tdt$$

und erhalten mit $t_1 = \sqrt{5x_1 - 1} = 2$ und $t_2 = \sqrt{5x_2 - 1} = 3$

$$\begin{aligned}\int_1^2 x\sqrt{5x-1} &= \int_2^3 \frac{1}{5}(t^2+1)t \cdot \frac{2}{5}tdt \\ &= \frac{2}{25}\int_2^3 (t^4+t^2)dt \\ &= \frac{2}{25}\left(\frac{t^5}{5} + \frac{t^3}{3}\right)\Big|_2^3.\end{aligned}$$

Durch Einsetzen der Grenzen erhält man den gesuchten Integralwert.
Hier haben wir eine geeignete Subsitutionsfunktion gefunden, indem wir die Gleichung $\sqrt{5x-1} = t$ nach x aufgelöst haben, d.h. wir haben die Umkehrfunktion von $x \mapsto \sqrt{5x-1}$ bestimmt oder $\sqrt{5x-1} = \psi(t)$ gesetzt, wobei ψ die Umkehrfunktion von φ ist.
Das hat funktioniert, weil

$$\begin{aligned}\varphi : [2,3] &\to [1,2] \\ t &\mapsto \frac{1}{5}(t^2+1)\end{aligned}$$

bijektiv ist mit der Umkehrfunktion

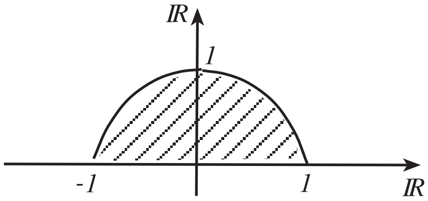
$$\begin{aligned}\psi : [1,2] &\to [2,3] \\ x &\mapsto \sqrt{5x-1}.\end{aligned}$$

Ist $\varphi : M \to N$ bijektiv mit der Umkehrabbildung $\psi : N \to M$, dann kann man die Substitutionsformel auch in der Form $(x_1, x_2 \in N)$

$$\int_{\psi(x_1)}^{\psi(x_2)} f(\varphi(t))\varphi'(t)dt = \int_{x_1}^{x_2} f(x)dx$$

schreiben.

(j) Zu berechnen sei $\int\limits_{-1}^{1} \sqrt{1-x^2}dx$.



In der geometrischen Interpretation des Integrals erhält man also den halben Flächeninhalt der Einheitskreises. Unsere Berechnung sollte also den Wert $\frac{\pi}{2}$ ergeben.
Hier liegt die Substitution $x = \varphi(t) = \cos t$ nahe:
Wir integrieren unbestimmt und erhalten

$$\int \sqrt{1-x^2}dx = -\int \sqrt{1-\cos^2 t}\sin t dt = -\int \sin^2 t dt.$$

$\int \sin^2 t$ kann man mit partieller Integration berechnen oder unter Beachtung von $\sin^2 t = \frac{1}{2}(1-\cos(2t))$. Das folgt aus den Additionstheoremen

$$\cos(2x) = \cos^2 t - \sin^2 t \text{ und } 1 = \cos^2 t + \sin^2 t.$$

Also ergibt sich

$$\begin{aligned}\int \sqrt{1-x^2}dx &= -\int \sin^2 t dt \\ &= -\frac{1}{2}(t - \sin t \cos t) \\ &= -\frac{1}{2}\arccos x + \frac{x}{2}\sqrt{1-x^2}.\end{aligned}$$

Für das bestimmte Integral ergibt sich daher

$$\begin{aligned}\int\limits_{-1}^{1} \sqrt{1-x^2}dx &= \left(-\frac{1}{2}\arccos x + \frac{x}{2}\sqrt{1-x^2}\right)\Big|_{-1}^{1} \\ &= -\frac{1}{2}\arccos 1 + \frac{1}{2}\arccos(-1) \\ &= 0 + \frac{1}{2}\pi = \frac{\pi}{2}\end{aligned}$$

in Übereinstimmung mit unseren früheren Resultaten.

Die Beispiele zeigen, dass ein wesentlicher Teil der „Kunst der Integration" in der Auffindung einer geeigneten Substitution besteht. Durch raffinierte Substitutionen, die wie ein Zaubertrick wirken, lassen sich manchmal sehr komplizierte Integrale so vereinfachen, dass man Stammfunktion angeben oder die Integrale berechnen kann.

Um z.B. das Integral $\int \frac{x^7}{x^4+2}dx$ zu berechnen, kann man durch $u := x^4$ eine neue Variable einführen, und man erhält sofort

$$\int \frac{x^7}{x^4+2}dx = \frac{1}{4}\int \frac{u}{u+2}du = \frac{1}{4}\int \frac{u+2-2}{u+2}du = \frac{1}{4}u - \frac{1}{2}\log(u+2) = \frac{1}{4}x^4 - \frac{1}{2}\log(x^4+2).$$

Das gleiche Resultat, allerdings mit erheblich mehr Aufwand, hätte man mit der *Methode der Partialbruchzerlegung* einer rationalen Funktion erhalten können.
Wie schon festgestellt, garantiert der Hauptsatz zwar, das jede stetige Funktion f auf einem Intervall D eine Stammfunktion besitzt. Doch ist damit noch lange nicht gesagt, dass eine Stammfunktion zu Klasse der *elementaren Funktionen* gehören muss. Unter der *elementaren Funktion* versteht man dabei die Gesamtheit aller Funktionen, die sich aus Polynomen, der Exponentialfunktionen, dem Sinus und allen denjenigen Abbildungen zusammensetzt, die sich hieraus mittels der vier „Grundrechenarten“ (Addition, Subtraktion, Multiplikation und Division), sowie den Operationen „Zusammensetzen“ und Bildung der Umkehrfunktionen in endlich vielen Schritten gewonnen werden können.
Aus unseren früheren Überlegungen folgt:

> Die Ableitung einer elementaren Funktion ist wieder eine elementare Funktion.

Bei Stammfunktionen elementarer Funktionen braucht dies nicht der Fall zu sein.
Systematisch hat sich wohl zuerst J.Liouville (1833) damit beschäftigt. Er hat für eine große Klasse stetiger Funktionen nachgewiesen, dass sie nicht elementar invertierbar sind. In jüngster Zeit hat man sich mit dem Aufkommen der Coputeralgebrasystemenen wieder mit dieser Frage beschäftigt und dabei große Fortschritte erzielt (vgl. z.B. R.H.Risch: The problem of integration in finite terms, Transac.Amer.Math.Soc.139 (1969)167-183).
Durch Integration lassen sich aus gegebenen Funktionen neue Funktionen gewinnen, manche haben wegen ihrer Wichtigkeit einen eigenen Namen:
Der *Integralsinus* $Si(x)$ ist definiert durch

$$Si(x) = \int_0^x \frac{\sin t}{t} dt = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)(2k+1)!}, \quad x \in \mathbb{R}.$$

Die Fehlerfunktion (Errorfunction) ist definiert durch

$$Err(f) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)k!}$$

Durch die angegebenen Reihenentwicklungen rechts lassen sich jedoch die Funktionswerte beliebig genau berechnen.
Der *Integralalgorithmus* definiert durch

$$Li(x) = \int_2^x \frac{1}{\log t} dt$$

Dazu gehören auch die sog. *elliptischen Integrale*, wie

$$\int \frac{1}{\sqrt{(1-t^2)(1-k^2t^2)}} dt \quad \text{und} \quad \int \sqrt{\frac{1-k^2t^2}{1-t^2}} dt \quad (0 < k^2 \neq 1)$$

(sog. Legendiesche Normalintegrale) oder die Weierstrass'schen Normalintegrale

$$\int \frac{1}{\sqrt{4t^3 - g_2 t - g_3}} dt \quad \text{oder} \quad \int \frac{t}{\sqrt{4t^3 - g_2 t - g_3}} dt,$$

wobei die Konstanten g_2, g_3 die Bedingung $\Delta(g_2, g_3) := g_2^3 - 27g_3^2 \neq 0$ erfüllen sollen. Das bedeutet, dass das kubische Polynom unter der Wurzel keine mehrfachen Nullstellen hat.
Historisch am Anfang dieser Integralbehandlungen (G.C.Fagnano, 1718) steht das elliptische integral

$$E(x) = \int_0^x \frac{1}{\sqrt{1-t^4}} dt, \quad 0 \leq x < 1.$$

Die Umkehrfunktion von E, nennen wir sie G, besitzt eine Fortsetzung als (meromorphe) Funktion nach $\mathbb{C}$, die Fortsetzung $\tilde{G}$ ist *doppelt-periodisch*, d.h. es gibt zwei über $\mathbb{R}$-linear unabhängige komplexe Zahlen w_1, w_2 (Perioden), mit $\tilde{G}(z + w_1) = \tilde{G}(z + w_2) = \tilde{G}(z)$. Da unter den elementaren Funktionen keine doppelt-periodischen vorkommen, kann $\tilde{G}$ nicht zu den elementaren Funktionen zählen.

Im folgenden Abschnitt werden wir jedoch zeigen, dass *rationale Funktionen elementar* invertierbar sind. Diese Aussage werden wir noch wesentlich präzisieren können.

21.3 Partialbruchzerlegung und die Integration der rationalen Funktionen

Dieser Abschnitt wird nachgetragen.

Als weitere Anwendung des Hauptsatzes beschäftigen wir uns mit der Frage der *Vertauschbarkeit von Differentiation mit Grenzprozessen.*

21.4 Verträglichkeit der Differentiation mit Grenzprozessen

Wir gehen von folgender Fragestellung aus: Ist $D \subset \mathbb{R}$ ein echtes Intervall und (f_n) eine Folge von differenzierbaren Funktionen $f_n : D \to \mathbb{R}$, die gegen eine Funktion $f : D \to \mathbb{R}$ konvergiert (etwa punktweise oder gleichmäßig).
Die Frage, die sich automatisch stellt: Ist dann f auch differenzierbar und gilt etwa $f' = \lim\limits_{n\to\infty} f'_n$?

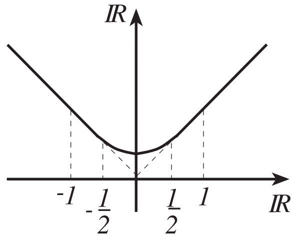
(1) Die schlechten Erfahrungen, die wir bei punktweisen Konvergenz gemacht haben, lassen vermuten, dass f i.A. nicht mehr differenzierbar ist. Das einfache Beispiel $f_n : [0,1] \to \mathbb{R}$ mit $f_(x) = x^n$ zeigt, dass die Grenzfunktion nicht einmal stetig, geschweige denn differenzierbar ist.

(2) Auch die gleichmäßige Konvergenz garantiert nicht die Differenzierbarkeit der Grenzfunktion: Dazu geben wir eine Folge (f_n) von Funktionen $f_n : [-1,1] \to \mathbb{R}$ an, die gleichmäßig gegen

$$\begin{aligned} abs : [-1,1] &\to \mathbb{R} \\ x &\mapsto abs(x) = |x| \end{aligned}$$

konvergiert. Dazu definieren wir

$$f_n(x) = \begin{cases} \frac{n}{2}x^2 + \frac{1}{2n}, & \text{für } |x| \leq \frac{1}{2} \\ |x|, & \text{für } |x| \geq \frac{1}{2} \end{cases}$$



Kritische sind die Punkte $-\frac{1}{n}$ und $\frac{1}{n}$. Dort haben beide Funktionen rechts in der Definition die Steigung -1 bzw. 1. Das Minimum von f_n liegt bei 0 und hat den Wert $\frac{1}{2}n$, daher ist

$$\|f_n - abs\| \leq \frac{1}{2n}.$$

f_n konvergiert also gleichmäßig gegen abs, alle f_n sind differenzierbar, die Grenzfunktion abs ist aber nicht differenzierbar (in Null).
Nach dem Weierstrass'schen Approximationssatz (siehe Kap VI) ist jede stetige Funktion $f : [a,b] \to \mathbb{R}$ gleichmäßige Limes einer Folge von Polynomfunktionen p_n. Jedes $p_n \in C^\infty([a,b])$, eine beliebige stetige Funktion braucht aber nicht differenzierbar zu sein.

(3) Selbst wenn die Grenzfunktion f eine gleichmäßig konvergente Folge differenzierbarer Funktionen f_n differenzierbar ist, braucht nicht $f' = \lim\limits_{n\to\infty} f_n'$ zu gelten.
Sei dazu für $n \in \mathbb{N}$

$$\begin{array}{rcl} f_n : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{\sin nx}{n}, \end{array}$$

(f_n) ist wegen $|\sin x| \le 1$ gleichmäßig konvergent mit dem Nullfunktion $f = 0$ als Grenzfunktion.
Es ist aber $f_n'(x) = \cos nx,\ (f_n'(x))$ konvergiert z.B. für $x = 0,\ f_n'(0) = \cos 0 = 1$, aber $\lim\limits_{n\to\infty} f_n'(0) = 1 \ne 0 = f'(0)$.

(4) Analoge Beispiele lassen sich bei Reihen differenzierbarer Funktionen finden:
Die Reihe $\sum\limits_{n=1}^{\infty} \frac{\sin(nx)}{n^2}$ konvergiert gleichmäßig auf $\mathbb{R}$, die Ableitung $\sum\limits_{n=1}^{\infty} \frac{\cos nx}{n}$ konvergiert aber z.B. nicht für $x = 0$.

Ein Vertauschungssatz für die Differentiation und Grenzwertbildung muss also etwa komplizierter aussehen.

21.4.1 Satz (Vertauschbarkeit von Differentiation mit Grenzprozessen)

Ist $D \subset \mathbb{R}$ ein echtes Intervall und gegeben sei eine Folge (f_n) von differenzierbaren Funktionen $f_n : D \to \mathbb{R}$.
Wir machen folgende Voraussetzungen:

(a) Die Folge (f_n) konvergiert in mindestens einem Punkt $x_0 \in D$.
(b) Die Funktionen $f_n',\ n \in \mathbb{N}$, sind stetig (m.a.W. $f_n \in C^1(D)$)
(c) Die Folge (f_n') konvergiere gleichmäßig. (es reicht: gleichmäßig auf jedem kompakten Teilintervall $[\alpha, \beta] \subset D$.

Dann konvergiert die Folge (f_n) gegen eine differenzierbare Funktion $f : D \to \mathbb{R}$ und es gilt $f = \lim\limits_{n\to\infty} f_n'$ oder suggestiver

$$\boxed{(\lim_{n\to\infty} f_n)' = \lim_{n\to\infty} f_n'}$$

Zusatz: Ist $D = [a,b]$ eine kompaktes Intervall, so konvergiert die Folge (f_n) gleichmäßig gegen f. Die Grenzfunktion der Folge (f_n') sei g. Als gleichmäßiges Limes stetiger Funktionen ist g stetig (vgl. ???). Auf Grund des Hauptsatzes gilt

$$f_n(x) = f_n(x_0) + \int\limits_{x_0}^{x} f_n'(t)dt \ \text{ für } \ x \in D \ \text{ und } \ n \in \mathbb{N}.$$

Nach dem Stabilitätssatz (vgl. ???) gilt

$$\lim_{n\to\infty} \int\limits_{x_0}^{x} f_n'(t)dt = \int\limits_{x_0}^{x} g(t)dt \ \text{ für alle } \ x \in D$$

und damit existiert

$$f(x) := \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} f_n(x_0) + \int\limits_{x_0}^{x} g(t)dt = f(x_0) + \int\limits_{x_0}^{x} f(t)dt \quad \text{für alle} \quad x \in D$$

Der Hauptsatz liefert nun für alle $x \in D$

$$f'(x) = g(x) = \lim_{n\to\infty} f_n'(x).$$

Beweis des Zusatzes:
Ist $D = [a, b]$ ein kompaktes Intervall, so folgt aus der Ungleichung

$$\begin{aligned} |f_n(x) - f(x)| &\leq |f_n(x_0) - f(x_0)| + \left| \int\limits_{x_0}^{x} f_n'(t)dt - \int\limits_{x_0}^{x} g(t)dt \right| \\ &\leq |f_n(x_0) - f(x_0)| + |x - x_0| \|f_n' - g\| \\ &\leq |f_n(x_0) - f(x_0)| + (b - a)\|f_n' - g\| \end{aligned}$$

für alle $x \in [a, b]$ die gleichmäßige Konvergenz von (f_n) gegen f.
Wendet man den Satz auf die Partialsummen einer Funktionenreihe an, so erhält man

21.4.2 Korollar

Ist $D \subset \mathbb{R}$ ein echtes Intervall, (f_n) eine Folge von stetig differenzierbaren Funktionen mit folgenden Eigenschaften:
Es gibt einen Punkt $x_0 \in D$, so dass die Folge $(F_n) = \left(\sum\limits_{k=0}^{n} f_k\right)$ der Partialsummen in x_0 konvergiert.
Die Folge $(F_n') = \left(\sum\limits_{k=0}^{n} f_k'\right)$ konvergiere gleichmäßig auf D(es genügt gleichmäßig auf jedem kompakten Teilintervall). Dann konvergiert die Folge (F_n) gegen eine differenzierbare Funktion $F : D \to \mathbb{R}$ und es gilt

$$F' = \left(\sum_{k=0}^{\infty} f_k\right)' = \sum_{k=0}^{\infty} f_k'.$$

Ist D kompakt, dann konvergiert (F_n) gleichmäßig gegen F. Die wichtigste Anwendung dieses Satzes betrifft Potenzreihen:

21.4.3 Korollar

Ist $\sum\limits_{k=0}^{\infty} a_k(x - a)^k$ eine Potenzreihe mit Konvergenzradius $r > 0$, $D := U_r(a)$ und ist

$$\begin{aligned} p : D &\to \mathbb{R} \\ x &\mapsto \sum_{k=0}^{\infty} a_k(x - a)^k \end{aligned}$$

die durch die Potenzreihe dargestellte Funktion, dann gilt

$$p'(x) = \sum_{k=1}^{\infty} k a_k(x - a)^{k-1} = \sum_{k=0}^{\infty} (k + 1)a_{k+1}(x - a)^k.$$

Die Ableitung einer Potenzreihe erhält man also (wie bei Polynomen) durch gliedweise Differentiation der Reihe.

Zusatz: $q : D \to \mathbb{R}$ mit $q(x) := \sum_{k=0}^{\infty} \frac{a_k}{k+1}(x-a)^{k+1}$ ist eine Stammfunktion von p.

Beweis: Wir wissen, dass die formal differenzierbare Reihe $\sum_{k=1}^{\infty} ka_k(x-a)^{k-1}$ den gleichen Konvergenzradius wie die Ausgangsreihe hat (das folgt auch einfach aus der Cauchy-Hadamardschen Formel für den Konvergenzradius). Daher erfüllt p die Voraussetzungen des Korollars, weil jedes $x \in U_r(a)$ in einem abgeschlossenen Teilintervall $[a-\varrho, a+\varrho] \in U_r(a)$, in welchem die formal abgeleitete Reihe gleichmäßig konvergiert ($0 < \varrho < r$ geeignet).
Durch mehrfache Anwendung folgt $p \in C^\infty(D)$ und für $k \in \mathbb{N}_0$ gilt

$$\boxed{a_k = \frac{f^{(k)}(a)}{k!}}$$

Hieraus kommen wir im nächsten Abschnitt zurück.

Da auch $\sum \frac{a_k}{k+1}(x-a)^{k+1}$ den Konvergenzradius r hat, gilt auch der Zusatz.

Durch Anwendung des Korollars erhält man völlig neue Beweise für die Ableitungen von exp, cos, sin, cosh, sinh, etc.

21.4.4 Beispiele und Bemerkungen

(a) Wegen $\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^k}{k!} + \ldots$ folgt

$$\begin{aligned} \exp'(x) &= 1 + \frac{2}{2!}x + \ldots + \frac{k+1}{(k+1)!}x^{k-1} + \ldots \\ &= 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^k}{k!} + \ldots \\ &= \exp(x) \end{aligned}$$

Analog folgt für $x \in \mathbb{R}$

$$\sin' x = \cos x,\ \cos' x = -\sin x,\ cosh' x = \sinh x \text{ und } sinh' x = \cosh x.$$

(b) Beim Beweis der Ableitungsformel für exp wurde die *Funktionalgleichung* $(x, y \in \mathbb{R})$ $\exp(x+y) = \exp(x)\exp(y)$ wesentlich benutzt.
Weiß man jetzt über die gliedweise Differentiation der Reihe, dass $\exp' = \exp$ gilt, so kann man die Funktionalgleichung (das Additionstheorem) so beweisen:
Bei festem, aber beliebigen $y \in \mathbb{R}$, betrachte man

$$\begin{aligned} g : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \exp(x+y)\exp(-x) \end{aligned}$$

Nach der Produktregel ist

$$g'(x) = \exp(x+y)\exp(y) - \exp(x+y)\exp(-x) = (\exp(x+y) - \exp(x+y))\exp(-x) = 0.$$

Da $\mathbb{R}$ ein Intervall ist, ist also $g(x) = g(0) = \exp(y)$, also

$$\boxed{\exp(x+y) = \exp(x)\exp(y)}$$

Damit ist die Funktionalgleichung (das Additionstheorem) von exp mit Hilfe der Differentialgleichung von exp bewiesen.
Man beachte, dass wir in einen circuluc vitiosus geraten wären, wenn wir $\exp' = \exp$ nicht mit Hilfe der gliedweisen Differentiation der Reihe bewiesen hätten.

21.5 Taylor'sche Formel, Taylor-Reihen, Abelscher Grenzwertsatz

Wir haben die Funktionen exp, sin, cos, sinh, cosh etc durch Potenzreihen eingeführt (sogar für komplexe Argumente). Potenzreihen sind offensichtliche Verallgemeinerungen von Polynomen,, diese wiederum sind sehr gut handhabbare Funktionen mit angenehmen Eigenschaften (stetig, beliebig oft differenzierbar). In diesem Abschnitt beschäftigen wir uns systematisch mit der Entwicklung von Funktionen in Potenzreihen. Wir werden sehen: Ist eine Funktion überhaupt in eine Potenzreihe entwickelbar, dann ist diese Potenzreihe notwendig die Taylor-Reihe, zunächst zur Erinnerung.
Ist $f : D \to \mathbb{R}$ ($D \subset \mathbb{R}$ ein echtes Intervall) $a \in D$ und f in a differenzierbar, dann approximiert die Tangente

$$\begin{aligned} T : D &\to \mathbb{R} \\ x &\mapsto f(a) + f'(a)(x-a) \end{aligned}$$

an den Graphen von f die Funktion f in einer Umgebung von a so gut, dass sogar

$$(*) \qquad \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - T(x)}{x-a} = 0$$

gilt.
Ferner ist $f(a) = T(a)$ und $f'(a) = T'(a)$. $T_1 := T$ ist ein Polynom von Grad ≤ 1. Wir wollen in Verallgemeinerung dieser Situation von folgender *Problemstellung* ausgehen:
Gegeben sei eine echtes Intervall $D \subset \mathbb{R}$, eine hinreichend oft differenzierbare Funktion, sagen wir n-mal differenzierbare Funktion ($n \in \mathbb{N}$). Gesucht ist ein Polynom $T_n : \mathbb{R} \to \mathbb{R}$ vom Grad $\leq n$ mit der folgenden Eigenschaft:

$$\begin{aligned} (**) \qquad & T_n(a) = f(a) = f^{(0)}(a), \\ & T_n'(a) = f'(a), \ldots, T^{(n)}(a) = f^{(n)}(a). \end{aligned}$$

Wir machen für das Polynom T_n versuchsweise den Ansatz

$$T_n(x) = \sum_{k=0}^{n} a_k (x-a)^k \qquad (a_k \in \mathbb{R};\ 0 \leq k \leq n)$$

weil wir f in einer Umgebung von a approximieren wollen.
Wegen $T_n^{(k)}(a) = k! a_k$ für $0 \leq k \leq n$, ergibt sich, dass die a_k durch f eindeutig bestimmt sind: $a_k = \frac{T_n^{(k)}(a)}{k!} = \frac{f^{(k)}(a)}{k!}$, $k = 0, \ldots, n$.

21.5.1 Satz und Definition

Es gibt genau ein Polynom $T_n : \mathbb{R} \to \mathbb{R}$ vom Grad $\leq n$, das $(**)$ erfüllt ist, nämlich

$$\begin{aligned} T : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto T_n(x) := T_{n,a}(f)(x) := f(a) + f'(a)(x-a) + \ldots + \frac{f^{(k)}(a)}{k!}(x-a)^k = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \end{aligned}$$

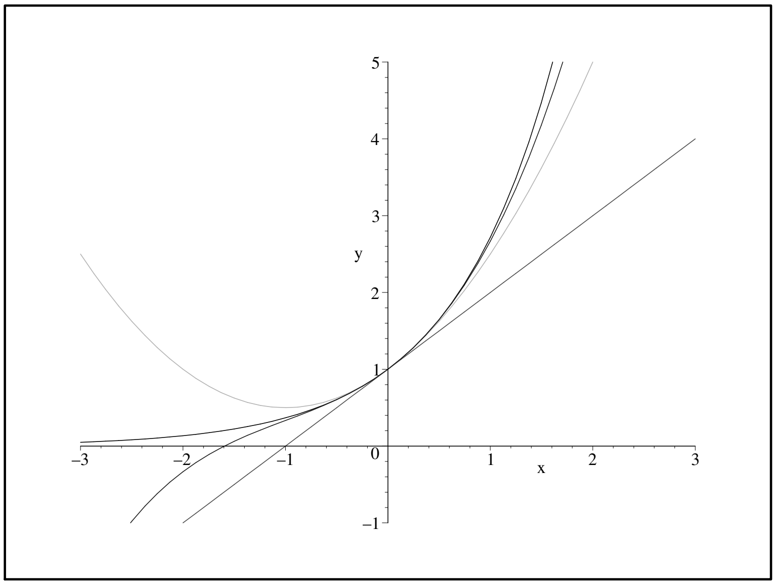


Abbildung 22: Schmiegparabeln der Grade 1, 2, 3 der Exponentialfunktion am Punkt 0

T_n heißt n*-tes Taylor-Polynom* von f zum Punkt a. (B. Taylor 1685-1721, Schüler von I.Newton)
$a_b := \frac{f^{(k)}(a)}{k!}$ heißt der k-te Taylor-Koeffizient von f im Punkt a.
Wir werden sehen, dass für den Fall, dass f n-mal stetig differenzierbar ist, in Verallgemeinerung von $(*)$ gilt

$$\lim_{x \to a} \frac{f(x) - T_{n,a}(a)}{(x-a)^n} = 0$$

Das Taylor-Polynom T_1 ist die Tangente an den Graphen von f, das Taylor-Polynom T_2 eine Parabel, falls $f''(a) \neq 0$. Den Graphen von $T_{n,a}(f)$ nennt man auch Schmiegeparabel n-ten Grades von f in a. (siehe Abb.22, Abb. 23 und Abb. 24)
Um die Güte der Approximation von f durch $T_{n,a}(f)$ zu messen, bezeichnen wir die Differenz (Abweichung) $f - T_{n,a}(f)$ mit R_{n+1}:

$$R_{n+1}(x) = f(x) - T_{n,a}(f)(x).$$

Ob und in wieweit Taylor-Polynome als brauchbare Näherungen für f (wenigstens „nahe bei a“) zu betrachten sind, kann erst durch Analyse dieses Fehlers beantwortet werden. Es gilt nun

21.5.2 Satz (Taylor'sche Formel mit Integralrestglied)

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f \in C^{n+1}(D)$ (f also $n+1$-mal stetig differenzierbar), $a \in D$.
Dann gilt für alle $x \in D$

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!} + R_{n+1}(x)$$

mit

$$R_{n+1}(x) = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t) dt,$$

also $f(x) = T_{n,a}(f)(x) + R_{n+1}(x)$.

Bemerkung: Der Rest R_{n+1} hängt natürlich auch (wie das Taylor-Polynom) von f und der Stelle a.

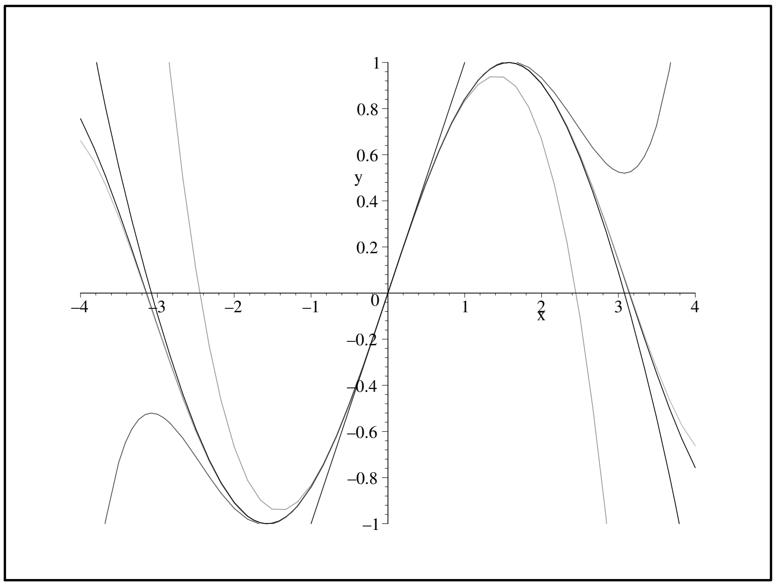


Abbildung 23: Schmiegparabeln der Grade 1, 3, 5, 7, 9 des Sinus am Punkt 0

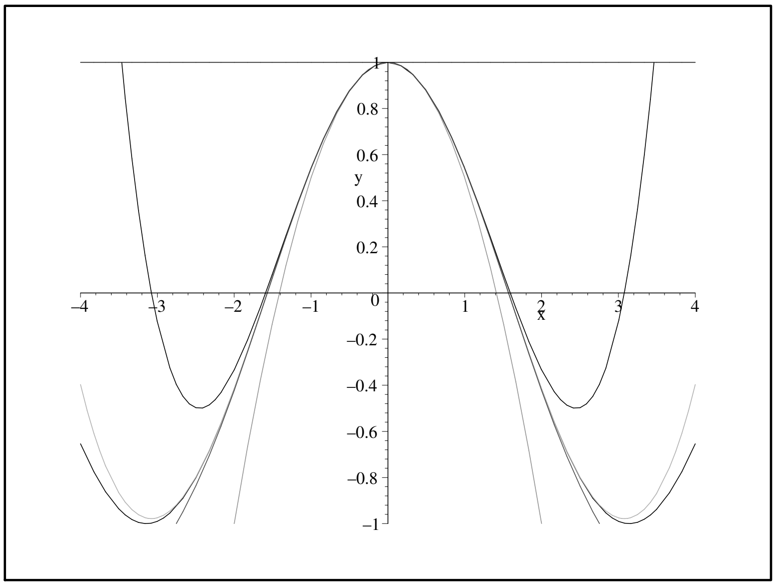


Abbildung 24: Schmiegparabeln der Grade 0, 2, 4, 6, 8 des Cosinus am Punkt 0

Diese Abhängigkeit bringen wir jedoch in der Notation (im Gegensatz zum Taylor-Polynom) nicht extra zum Ausdruck.

Beweis von 20.5.2: Im Fall $n = 0$ lautet die Aussage $f(x) = f(a) + \int_a^x f'(t)dt$, das ist genau die Aussage des Hauptsatzes (Version 2 für stetig differenzierbare Funktionen f). Es bietet sich also ein Induktionsbeweis mit der Induktionsverankerung $n = 0$ an.
Schluss von $n-1$ auf n:
Wir nehmen an, dass die Formel für $n-1$ gilt:

$$(*) \qquad f(x) = T_{n-1,a}(f)(x) + \underbrace{\frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f^{(n)}(t)dt}_{R_n(x)}$$

Mit partieller Integration folgt

$$R_n(x) = -\left.\frac{(x-t)^n}{n!} f^{(n)}(t)\right|_a^x + \underbrace{\frac{1}{n!}\int_a^x (x-a)^n f^{(n-1)}(t)dt}_{R_{n+1}(x)} = \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n+1}(x).$$

Setzt man diesen Wert für $R_n(x)$ in die Darstellung $(*)$ ein, so folgt

$$f(x) = T_{n-1,a}(f)(x) + \frac{f^n(a)}{n!}(x-a)^n + R_{n+1}(x) = T_{n,a}(f)(x) + R_{n+1}(x)$$

und das ist die Aussage für n.

□

Obwohl man an der Integralform des Restgliedes schon einiges ablesen kann, ist die folgende Form des Restgliedes für Anwendungen häufig günstiger.

21.5.3 Satz (Lagrange'sche Form des Restgliedes)

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f \in C^{n+1}(D)$, $a, x \in D$. Dann gibt es ein ξ mit $\xi \in [a, x]$ oder $\xi \in [x, a]$ mit

$$f(x) = T_{n,a}(f)(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

Nach dem Mittelwertsatz der Integralrechnung (vgl. ???) existiert ein ξ, so dass gilt

$$R_{n+1}(x) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)dt = f^{(n+1)}(\xi)\int_a^x \frac{(x-t)^n}{n!}dt = -f^{(n+1)}(\xi)\left[\frac{(x-t)^{n+1}}{(n+1)!}\right]_a^x = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

Man kann die Voraussetzung noch abschwächen.

21.5.4 Satz

Sei $f \in C^n(D)$ und $f^{(n)} : D \to \mathbb{R}$ noch differenzierbar in alle inneren Punkten von D. Dann gibt es zu $a \in D$ und $x \in D$ ein ϑ mit $0 < \vartheta < 1$, so dass gilt

$$f(x) = T_{n,a}(f)(x) + \frac{f^{(n+1)}(a + \vartheta(x-a))}{(n+1)!}(x-a)^{n+1}$$

Beweis : Man betrachte für festes x

$$g(u) = f(x) - f(u) - f'(u)(x-u) - \ldots - \frac{f^{(n)}(u)}{n!}(x-u)^n - \sigma \frac{1}{(n+1)!}(x-n)^{n+1}$$

wobei die Konstante $\sigma \in [a,x]$ bzw $[x,a]$ gewählt sei und so bestimmt wird, dass $g(a) = 0$ gilt. Da außerdem $g(x) = 0$, gibt es nach dem Satz von Rolle ein $\vartheta \in]0,1[$ mit $g'(a_0 + \vartheta(x-a)) = 0$. Andererseits folgt

$$g'(u) = -\frac{1}{n!} f^{(n+1)}(u)(x-u)^n + \frac{\sigma}{n!}(x-u)^n$$

und damit folgt $\sigma = f^{(n+1)}(a_0 + \vartheta(x-a))$.

□

Man beachte, dass dieser Satz im Fall $n = 0$ gerade der MWSD ist.

21.5.5 Folgerung

Ist $f \in C^n(D)$ und $f^{(n+1)} = 0$, dann ist f ein Polynom vom Grad $\leq n$.
Denn das Restglied im letzten Satz ist dann immer Null und f reduziert sich auf das Taylor-Polynom von Grad $\leq n$.

Alle Polynome vom Grad $\leq n$ sind also Lösungen der Differentialgleichung $f^{(n+1)} = 0$ (sogar auf ganz $\mathbb{R}$)

Die verschiedenen Darstellungen des Restgliedes werden wir für Abschätzungen der Größe des Fehlers als auch (speziell die Lagrange Form) zur Bestimmung des Vorzeichens des Fehlers benutzen.

21.5.6 Beispiele und Bemerkungen

(a) Es gilt $\sin, \cos \in C^\infty(D)$ und für $a = 0$ ist

$$\sin x = x - \frac{x^3}{3!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + R_{2n+3}(x)$$

mit

$$R_{2n+3}(x) = \frac{\sin^{2n+3}(\xi)}{(2n+3)!} x^{2n+3} = (-1)^{n+1} \frac{\cos\xi}{(n+1)!} x^{2n+3}$$

mit ξ zwischen 0 und x, und damit gilt für *alle* $x \in \mathbb{R}$

$$\left| \sin x - \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \right| \leq \frac{|x|^{2n+3}}{(2n+3)!}$$

Als einfaches Beispiel berechnen wir den relativen Fehler $r(x)$, wenn man im Intervall $]0, \frac{\pi}{4}]$ den Sinus durch des Taylor-Polynom $P(x) := T_3(x) := x - \frac{x^3}{3!}$ approximiert. Es ist (beachte $\sin x > 0$)

$$0 < r(x) = \frac{\sin x - P(x)}{\sin x} \leq \frac{x^5}{5! \sin x}.$$

Für $0 < x \leq \frac{\pi}{4}$ ist aber

$$\sin x \geq x - \frac{x^3}{3!} = x(1 - \frac{x^2}{3!}) \geq x(1 - \frac{\pi^2}{4^2 \cdot 3!}),$$

also

$$0 < r(x) \le \frac{x^4}{5!(1-\frac{\pi^2}{4^2\cdot 3!})} \le \frac{\pi^4}{4^4\cdot 5!}\frac{1}{(1-\frac{\pi^2}{4^2\cdot 3!})} < 0,0036 \text{ (Taschenrechner oder Maple)}$$

Also ist der relative Fehler kleiner als $0,36\%$.

Analog gilt für *alle* $x \in \mathbb{R}$:

$$\left|\cos x - \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!}\right| \le \frac{|x|^{2n+2}}{(2n+2)!}$$

(b) Für $\exp \in C^\infty(\mathbb{R})$ und $a = 0$ folgt

$$e^x := \exp(x) = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!}e^{\vartheta x}$$

mit $0 < \vartheta < 1$.
Will man etwa $\exp$ im Intervall $[-1, 1]$ durch ein Polynom bis auf 2 Stellen nach dem Komma genau approximieren, so leistet das Taylor-Polynom T_5 das Gewünschte:
So ist nämlich

$$P(x) := T_{5,0}(\exp)(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120}$$

und damit

$$|e^x - P(x)| = \frac{x^6}{720}e^{\vartheta x} < \frac{e}{720} < 0,0038.$$

Aus der obigen Darstellung von $\exp x$ ergibt sich für $x = 1$ auch die Abschätzung

$$0 < e - E_n < \frac{3}{(n+1)!}$$

für alle $n \in \mathbb{N}_0$, dabei ist $E_n = \sum\limits_{k=0}^{n} \frac{1}{k!}$.
Aus dieser Abschätzung ergibt sich (vgl. ???) leicht die *Irrationalität* von e.

Die hinreichende Kriterien für lokale Extrema oder Wendepunkte (vgl. ???) versagen schon in einfachen Beispielen: $P_4(x) = x^4$ oder $P_5(x) = x^5$

Mit Hilfe des Lagrange'schen Restgliedes erhält man folgendes hinreichende Kriterium für lokale Extrema.

21.5.7 Satz (hinreichendes Kriterium für lokale Extrema)

Sei $D \subset \mathbb{R}$ ein echtes Intervall, $f \in C^{n+1}(D)$. In einem Punkt $a \in D$ gelte $f'(a) = f''(a) = \ldots = f^{(n)}(a) = 0$, jedoch $f^{(n+1)}(a) \neq 0$. Dann hat f in a

(a) ein *strenges (striktes) lokales Minimum*, falls n ungerade und $f^{(n+1)}(a) < 0$ gilt;
(b) ein strenges *lokales Maximum*, falls n un gerade und $f^{(n+1)}(a) < 0$,
(c) kein Extremum, falls n gerade ist.

Man hat die Differenz $f(x) - f(a)$ zu betrachten. Zunächst besagt die Voraussetzung

$$f(x) - f(a) = \frac{f^{(n+1)(\xi)}}{(n+1)!}(x-a)^{n+1} = R_{n+1}(x)$$

Wir betrachten den Fall $f^{n+1}(a) > 0$ und wählen ein möglicherweise kleines Intervall $U_\delta(a)$, so dass für alle $x \in U_\delta(a) \cap D$ gilt $f^{(n+1)}(x) > 0$, insbesondere ist $f^{(n+1)}(\xi) > 0$.

Ist nun n ungerade, also $n+1=2k$ gerade, dann gilt für $x \in U_\delta(a) \cap D$ immer $(x-a)^{2k \geq 0}$, wobei das Gleichheitszeichen nur für $x=a$ gilt. Es ist also $R_{n+1}(x) > 0$ für alle $x \in U_\delta(a);\ x \neq a$. An der Stelle a liegt also ein strenges lokales Minimum vor. Die Aussage für den Maximum erhält man durch Übergang zu $-f$.
Ist jedoch n gerade, also $n+1$ ungerade, so ist $R_{n+1}(x) > 0$ für $x > a$ und $R_{n+1}(x) < 0$ für $x < a$; In a kann in diesem fall weder ein Maximum noch ein Minimum vorliegen.

□

Wir beweisen nun die angekündigte Verallgemeinerung der Formel $(*)$ aus der Einführung zu diesem Abschnitt:

21.5.8 Satz (qualitative Form der Taylorschen Formel)

Ist $D \subset \mathbb{R}$ ein echtes Intervall und $f \in C^n(D)$, $n \in \mathbb{N}$, $a \in D$. Dann gilt für alle $x \in D$

$$(*) \qquad f(x) = T_{n,a}(f)(x) + (x-a)^n \varrho(x),$$

dabei ist $\varrho : D \to \mathbb{R}$ eine in a stetige Funktion mit $\varrho(a) = 0$.

Man sagt hierfür auch: f und $T_{n,a}f$ stimmen in a in n-ten Ordnung überein.
Zum Beweis gehen wir von (Lagrangesches Restglied)

$$\begin{aligned} f(x) - T_{n-1,a}(f)(x) &= f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k \\ &= \frac{f^{(n)}(\xi)}{n!}(x-a)^n \end{aligned}$$

aus.
Man beachte: wir haben nur $f \in C^n(D)$ vorausgesetzt.
Dann folgt

$$f(x) - T_{n-1,a}(f)(x) = \frac{f^{(n)}(a)}{n!}(x-a) + \underbrace{\frac{f^{(n)}(\xi) - f^{(n)}(a)}{n!}}_{\varrho(x)}(x-a)^n$$

Da ξ zwischen a und x liegt, also die Gestalt $\xi = \xi(x) = a + \vartheta(x-a)$ hat mit $0 \leq \vartheta \leq 1$, folgt wegen der Stetigkeit von $f^{(n)}$

$$\lim_{\substack{x \to a \\ x \neq a}} \varrho(x) = \lim_{x \to a} \frac{f^{(n)}(\xi(n)) - f^{(n)}(a)}{n!} = 0,$$

d.h. ϱ ist stetig in a und $\varrho(a) = 0$.

21.5.9 Beispiele und Bemerkungen

(a) Die Gleichung $(*)$ in 20.5.(8) ist äquivalent mit

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - T_{n,a}(f)(x)}{(x-a)^n} = 0$$

Im Fall $n=1$ erhält man also einfach die Charakterisierung der Differezierbarkeit durch die lineare Approximierbarkeit. Das Taylor-Polynom $T_{n,a}(f)$ ist durch $(*)$ charakterisiert, denn es gilt:

Ist $p_n : \mathbb{R} \to \mathbb{R}$ ein Polynom vom Grad $\leq n$, mit $\lim\limits_{\substack{x \to a \\ x \neq a}} \frac{f(x)-p_n(a)}{(x-a)^n} = 0$, dann ist $p_n = T_{n,a}(f)$, denn aus der Voraussetzung folgt dann auch

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{T_{n,a}(f)(x) - p_n(x)}{(x-a)^n} = 0$$

das gehr aber nur wenn $T_{n,a}(f) - p_n$ das Nullpolynom ist.

(b) Häufig verwendete *Näherungsformeln* beruhen auf der qualitativen Form des Taylorschen Satzes.

(α) $\sqrt{1+x} \approx 1 + \frac{x}{2} - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4$
(β) $\frac{1}{\sqrt{1+x}} \approx 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{6}{16}x^8$
(γ) $\frac{1}{1-x} \approx 1 + x$ ($|x|$ immer hinreichend klein).

Häufig benügt man sich z.B. im Fall (α) schon mit $\sqrt{1+x} = 1 + \frac{x}{2} + \varrho(x)x \quad (\lim\limits_{x\to 0} \varrho(x) = 0)$ (hier wird

$$\begin{aligned} f :]-1,1[&\to \mathbb{R} \\ x &\mapsto \sqrt{1+x} \end{aligned}$$

betrachtet, $a = 0$)

Wir benutzen die obige Formel, um ganz einfach

$$\boxed{\lim_{n\to\infty} (\sqrt{n+\sqrt{n}} - \sqrt{n}) = \tfrac{1}{2}}$$

zu zeigen.
Für $n > 1$ ist nämlich

$$\begin{aligned} \sqrt{n+\sqrt{n}} = \sqrt{n(1+\frac{1}{\sqrt{n}}} &= \sqrt{n}\sqrt{1+\frac{1}{\sqrt{n}}} \\ &= \sqrt{n}(1 + \frac{1}{2\sqrt{n}} + \varrho(\frac{1}{\sqrt{n}})\frac{1}{\sqrt{n}}) \\ &= \sqrt{n} + \frac{1}{2} + \varrho(\frac{1}{\sqrt{n}}). \end{aligned}$$

Hieraus folgt aber wegen $\lim\limits_{n\to\infty} \varrho(\frac{1}{\sqrt{n}} = \varrho(0) = 0$

$$\lim_{n\to\infty} (\sqrt{n+\sqrt{n}} - \sqrt{n}) = \frac{1}{2}.$$

Ist nun $D \subset \mathbb{R}$ ein echtes Intervall und $f \in C^\infty(D)$, $a \in D$, dann ist ja insbesondere $f \in C^n(D)$ für jedes $n \in \mathbb{N}_0$ und man kann für jedes n das n-te Taylor-Polynom $T_{n,a}(f)(x) = \sum\limits_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$ betrachten und fragen: Existiert $\lim\limits_{n\to\infty} T_{n,a}(f)(x) = \lim\limits_{n\to\infty} \sum\limits_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$?
Das führt zu folgender Definition:

21.5.10 Definition (Taylor-Reihe von $f \in C^\infty(D)$

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f \in C^\infty(D)$, $a \in D$. Dann heißt die Potenzreihe $T_{\infty,a}(f)(x) := \sum\limits_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$ die Taylor-Reihe von f im Punkt a (oder zum Entwicklungspunkt a).
Dabei stellt sich die Frage:
Konvergiert die Taylor-Reihe -wenn sie denn konvergiert- immer gegen f ?

21.5.11 Bemerkungen und Beispiele

(a) Auf Grund des Taylor'schen Formel mit Restglied konvergiert die Taylorreihe $T_{\infty,a}(f)(x)$ für $x \in D$ genau dann gegen $f(x)$, wenn $\lim\limits_{n\to\infty} R_{n+1}(x) = 0$. Oder äquivalent: Die Folge $(T_{n,a}(f))$ der Taylor-Polynome konvergiert genau dann und $T_{\infty,a}f$ stellt dann die Funktion f ander Stelle x dar, wenn die Folge $(R_{n+1}(x)$ der Reste eine Nullfolge ist.

(b) Weiss man, dass f irgendwie um a in eine Potenzreihe entwickelt werden kann, gibt es also eine Potenzreihe $\sum a_k(x-a)^k$ mit positivem konvergenzradius $r > 0$, so dass für alle $x \in U_r(a) \cap D$ gilt

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

dann folgt aus dem Satz über die gliedweise Differenzierbarkeit einer Potenzreihe für $k \in \mathbb{N}_0$,

$$\boxed{a_k = \tfrac{f^{(k)}(a)}{k!}}$$

d.h. die Potenzreihe stimmt mit der Taylor-Reihe überein. Es gillt also $T_{\infty,a}(f)(x) = f(x)$ und unsere obige Folge ist positiv zu beantworten.

(c) Eine Potenzreihe kann aber auch nur für den Entwicklungsopunkt a konvergieren (Konvergenzradius Null). Es kann tatsächlich vorkommen, dass die Taylorreihe einer Funktion $f \in C^\infty(D)$ nur für den Entwicklungspunkt a konvergiert. Ein solches Beispiel wird z.B. durch die Funktion $f \in C^\infty(\mathbb{R})$ mit $(x \in \mathbb{R}, a = 0)$

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{2^k} \cos(kx)$$

geliefert. (vgl. z.B. Barner-Flohr: Analysis 1, 9.5)

(d) Für die Funktion $f \subset C^\infty(\mathbb{R})$ mit

$$f(x) = \begin{cases} \exp(-\frac{1}{x^2}, & \text{für } x \neq 0 \\ 0, & \text{für } x = 0 \end{cases}$$

gilt (vgl. Übungsaufgabe ??? vom Blatt ???) $f \in C^\infty(\mathbb{R})$ und $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}_0$.

$T_{\infty,0}f(x)$ konvergiert also für alle $x \in \mathbb{R}$, aber nur für $x = 0$ wird die Funktion dargestellt. Obwohl hier der Konvergenzradius der Taylor-Reihe positiv ist, stellt sie nur für den Entwicklungspunkt die Funktion dar.
Aus der Bemerkung (b) folgt der bemerkenswerte

21.5.12 Eindeutigkeitssatz (Identitätssatz) für Potenzreihen

Konvergieren die Potenzreihen

$$\sum_{k=0}^{\infty} a_k(x-a)^k \; und \; \sum b_k(x-a)^k$$

in $U_R(a)$ $(R > 0)$ und stellen sie dort dieselbe Funktion dar, dann gilt $a_k = b_k$ für alle $k \in \mathbb{N}_0$.

Beweis : Die dargestellte Funktion sei etwa f. Dann gilt nach (b) einerseits $\frac{f^{(k)}}{k!} = a_k$, aber auch $\frac{f^{(k)}(a)}{k!} = b_k$ für alle $k \in \mathbb{N}_0$, d.h. $a_k = b_k$.

□

Die beiden wichtigsten Interpretationen des Satzes sind:

(1) Wenn es überhaupt möglich ist eine Funktion in $U_R(a) =]a-R, a+R[$ als Potenzreihe mit Entwicklungspunkt a darzustellen, dann nur als Taylor-Reihe:

$$f(x) = T_{\infty,a}(f)(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

(2) Wird eine Funktion f auf zwei verschieden Weisen als Potenzreihe mit Entwicklungspunkt a dargestellt, dann sind die Koeffizienten entsprechender $(x-a)^n$-Potenzen gleich.
In dieser Form nennt man Satz auch *Satz von Koeffizientenvergleich*.
Mit ihm gelangt man häufig zu nicht-trivialen Beziehungen.
Der Satz enthält auch als Spezialfall den Identitätssatz für Polynome.

21.5.13 Satz (Hinreichnende Bedingung für die Darstellbarkeit einer C^∞-Funktion durch ihre Taylor-Reihe)

Ist $f \in C^\infty(D)$, $a \in D$ und gibt es Konstanten A und B, so dass für alle $x \in D$ und alle $n \in \mathbb{N}_0$ die Abschätzung $|f^{(n)}(x)| \le AB^n$ gilt, dann gilt

$$f(x) = T_{\infty,a}(f)(x)$$

Unter dieser Voraussetzung ist nämlich

$$|R_{n+1}(x)| \le \frac{AM^{n+1}|x-a|^{n+1}}{(n+1)!} = A\frac{(B|x-a|)^{n+1}}{(n+1)!}$$

und die Folge rechts ist eine Nullfolge.

Zusatz: Gilt sogar $|f^{(n)}(x)| \le M$ für alle $x \in D$ und alle $n \in \mathbb{N}_0$, dann gilt $\lim\limits_{n\to\infty} R_{n+1}(x) = 0$.

21.5.14 Weitere Beispiele für Taylor-Reihen

(a) Die C^∞-Funktionen sin, cos, sinh, cosh, exp wurden durch ihre Taylor-Reihe zum Entwicklungspunkt 0 eingeführt.
Es gilt aber für *alle* $a \in \mathbb{R}$, z.B. $\sin x = T_{\infty,a}(\sin)(x)$ bzw. $\cos x = T_{\infty,a}(\cos)(x)$ bzw.

$$\begin{aligned} \exp(x) = T_{\infty,a}(\exp)(x) &= \sum_{k=0}^{\infty} \frac{\exp^{(k)}(a)}{k!}(x-a)^k \\ &= \sum_{k=0}^{\infty} \frac{\exp(a)}{k!}(x-a)^k. \end{aligned}$$

sin, cos, exp werden also für alle $a \in \mathbb{R}$ durch ihre Taylor-Reihe dargestellt und zwar auf ganz $\mathbb{R}$.
Denn es gilt $|\sin^{(k)}(x)| \le 1$ und $|\cos^{(k)}(x)| \le 1$ für alle $x \in \mathbb{R}$ und alle $k \in \mathbb{N}_0$, daher folgt aus dem Zusatz von 20.5.13 die Behauptung.
Für die Exponentialfunktion kann man analog schließen:
Man nimmt die Lagrange'sche Form des Restgliedes $R_{n+1}(x) = \frac{\exp^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$, wobei $\xi \in [a,x]$ oder $\xi \in [x,a]$ gilt. Ist dann $k = \max\{a,x\}$, so ist

$$|\exp^{(k)}(\xi)| = |\exp(\xi)| \le \exp(b) =: M,$$

also ist,

$$|R_{n+1}(x)| \leq \frac{M|x-a|^{n+1}}{(n+1)!}$$

und damit $\lim\limits_{n\to\infty} R_{n+1}(x) = 0$ für alle $x \in \mathbb{R}$.
Im Fall der Exponentialfunktion hätte man auch mittels der Funktionalgleichung auf $\exp(x) = T_{\infty,a}(\exp)(x)$ für alle $a \in \mathbb{R}$ und alle $x \in \mathbb{R}$ schließen können:
Es ist ja

$$\begin{aligned}
\exp(x) &= \exp(a)\exp(x-a) \\
&= \exp(a)\sum_{k=0}^{\infty}\frac{(x-a)^k}{k!} \qquad \text{(nach Def. von exp)} \\
&= \sum_{k=0}^{\infty}\exp(a)\frac{(x-a)^k}{k!} \\
&= \sum_{k=0}^{\infty}\frac{\exp^{(k)}(a)}{k!}(x-a)^k = T_{\infty,a}(\exp)(x).
\end{aligned}$$

(b) In Spezialfällen kann man ohne Benutzung der Darstellungen in ??? das Restglied erhalten: Sei dazu

$$\begin{aligned}
f:\ \mathbb{R}-\{1\} &\to \mathbb{R} \\
x &\mapsto \frac{1}{1-x}
\end{aligned}$$

Nach der Summenformel für die geometrische Reihe ist für $|x| < 1$

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

Daher $P_n(x) := T_{n,0}(f)(x) = \sum\limits_{k=0}^{n} x^k$.
Statt aber das Restglied aus der Potenzreihendarstellung abzulesen, was dieses nur für $|x| < 1$ liefern würde, beachten wir, dass für $x \neq 1$ gilt

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

Es ist also für alle $x \neq 1$ (d.h. in $]-\infty, 1[$ bzw. $]1, \infty[$)

$$R_{n+1}(x) = \frac{x^{n+1}}{1-x}.$$

(c) Für $x > -1$ sei $f(x) = \log(1+x)$ (beachte $f'(x) = \frac{1}{1+x}$).
Wir wollen die Taylor-Reihe von f zum Entwicklungspunkt $a = 0$ aufstellen und nach Möglichkeit nur wenige Ableitungen ausrechnen.
Dazu beachten wir, dass nach dem Hauptsatz gilt

$$f(x) = \int_0^x \frac{1}{1+t}\,dt.$$

Für $|x| < 1$ gilt daher

$$f(x) = \log(1+t)|_0^x = \int_0^x \frac{1}{1+t}\,dt = \int_0^x \left(\sum_{k=0}^{\infty}(-1)^k t^k\right) dt.$$

Nun konvergiert $\sum\limits_{k=0}^{\infty}(-1)^k t^k$ gleichmäßig auf $[-|x|, |x|]$. Mit ??? folgt daher

$$f(x) = \log(1+x) = \sum_{k=0}^{\infty}(-1)^k \int\limits_0^x t^k dt = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1}$$

Damit haben wir: Für $|x| < 1$ gilt

$$\log(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k.$$

Da die Reihe rechts auch noch (nach dem Leibniz-Kriterium) für $x = 1$ konvergiert, ist zu vermuten, dass die Darstellung auch noch für $x = 1$ gilt, dass also

$$\boxed{\log 2 = \sum_{k=1}^{\infty} \tfrac{(-1)^{k-1}}{k} = 1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \tfrac{1}{5} \mp \ldots}$$

gilt.
Dazu benutzen wir die Taylorsche Formel mit dem Lagrange'schen Restglied ($0 \le \vartheta \le 1$)

$$\log(1+x) = \sum_{k=1}^{n}(-1)^{k-1}\frac{x^k}{k} + \underbrace{\frac{(-1)^n}{n+1}\frac{1}{(1+\vartheta x)^{n+1}}x^{n+1}}_{R_{n+1}(x)}$$

Für $0 \le x \le 1$ ist dann $1 + \vartheta x \le 1$, also $0 \le \frac{x}{1+\vartheta x} \le 1$ und damit $|R_{n+1}(x)| \le \frac{1}{n+1}$, also $\lim\limits_{n\to\infty} R_{n+1}(x) = 0$, speziell ist $\lim\limits_{n\to\infty} R_{n+1}(1) = 1$.

Fassen wir zusammen:

21.5.15 Satz

Für alle $x \in]-1,1]$ gilt

$$\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k,$$

Speziell gilt:

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \mp \ldots \approx 0$$

21.5.16 Bemerkung

Da $\lim\limits_{\substack{x \to 1 \\ x > -1}} \log(1+x) = -\infty$ gilt und die rechte Seite für $x \to -1$ $(x > -1)$ ebenfalls diesen uneigentlichen Grenzwert hat, gilt die Darstellung bei großzügiger Interpretation auch noch für $x = -1$.

Auf die Konvergenz der Taylor-Reihe im Punkt $x = 1$ hätte man auch mit Hilfe des folgenden sehr nützlichen Satzes schließen können.

21.5.17 Satz (Abelscher Grenzwertsatz, N.H.Abel, 1826)

Die Potenzreihe $\sum a_k x^k$ besitze den endlichen und positiven Konvergenzradius r. Die Reihe $\sum a_k x^k$ sei noch für $x = r$ konvergent. Dann ist die durch die Potenzreihe dargestellte Funktion $f :]-r, r] \to \mathbb{R}$ in r noch (rechtsseitig) stetig, d.h. es gilt

$$\lim_{\substack{x \to r \\ x < r}} f(x) = \sum_{k=0}^{\infty} a_k r^k =: f(r).$$

Beweis : Wir können oBdA $r = 1$ annehmen, denn hat die Potenzreihe $\sum a_n x^n$ den Konvergenzradius $r > 0$, so hat die Reihe $\sum b_k x^k$ mit $b_k := a_k r^k$ offensichtlich den Konvergenzradius 1 und $\sum b_k x^k$ ist dann und nur dann für $x = 1$ konvergent, wenn $\sum a_k x^k$ für $x = r$ konvergent ist.
Wir setzen $s_{-1} := 0$ und $s_n := a_0 + a_1 + \ldots + a_n$ und erhalten für $|x| < 1$

$$\sum_{k=0}^{n} a_k x^k = \sum_{k=0}^{n} (s_k - s_{k-1}) x^n = (1-x) \sum_{k=0}^{n-1} s_k x^k + s_n x^n$$

und damit $(n \to \infty)$ für $|x| < 1$

$$f(x) = (1-x) \sum_{k=0}^{\infty} s_k x^k.$$

Ist

$$s = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^{\infty} a_k = \sum_{k=0}^{\infty} a_k,$$

dann folgt wegen $(1-x) \sum\limits_{k=0}^{\infty} x^k = 1$ $(|x| < 1)$ (Summenformel für die geometrische Reihe)

$$(*) \qquad f(x) - s = (1-x) \sum_{k=0}^{\infty} (s_k - s) x^k.$$

Wegen $s = \lim\limits_{n\to\infty} s_n$ gibt es zu beliebig vorgegebenem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $|s_n - s| < \frac{\varepsilon}{2}$.
Für $0 \leq x < 1$ folgt dann aus $(*)$

$$\begin{aligned} |f(x) - s| &\leq (1-x)\sum_{k=0}^{\infty} |s_n - s|x^k \\ &\leq (1-x)\sum_{k=0}^{N-1} |s_k - s| + (1-x)\frac{\varepsilon}{2}\sum_{k=N}^{\infty} x^k \\ &\leq (1-x)\sum_{k=0}^{N-1} |s_k - s| + \frac{\varepsilon}{2}. \end{aligned}$$

Wählen wir nun ein $\delta \in]0,1[$, so dass für alle $x \in]1-\delta, 1[$ gilt

$$(1-x)\sum_{k=0}^{N-1} |s_k - s| < \frac{\varepsilon}{2},$$

so folgt für diese x

$$|f(x) - s| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

d.h. es ist

$$\lim_{\substack{x \to 1 \\ x < 1}} f(x) = s = \sum_{k=0}^{\infty} a_k.$$

□

Um die Tragweite des Abelschen Grenzwertsatzes zu demonstrieren, beweisen wir an dieser Stelle einen Satz über das Cauchy-Produkt *nicht notwendig absolut konvergenter* Reihen.

21.5.18 Satz (N.H. Abel, 1826)

Sind die drei (reellen) Reihen $\sum a_k,\ \sum b_l$ und $\sum c_n$ mit $c_n := \sum\limits_{k=0}^{n} a_k b_{n-k}$ konvergent und sind $A,\ B$ bzw. C ihre Summen, dann gilt $C = A \cdot B$ oder explizit

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

Man beachte, dass hier keine absolute Konvergenz vorausgesetzt wird.
Für $x \in [0,1[$ sei

$$A(x) = \sum_{k=0}^{\infty} a_k x^k,\ B(x) = \sum_{l=0}^{\infty} b_l x^l \text{ und } C(x) = \sum_{n=0}^{\infty} c_n x^n$$

und es gilt $A(x)B(x) = C(x)$ für $0 \leq x \leq 1$. Wegen der vorausgesetzten Konvergenz von $\sum a_k,\ \sum b_l$ und $\sum c_n$ besitzen die Funktionen nach dem Abelschen Grenzwertsatz einen Grenzwert für $x \to 1$ $(x < 1)$.
Aus $A(x)B(x) = C(x)$ folgt daher wegen $\lim\limits_{\substack{x \to 1 \\ x < 1}} A(x) = A,\ \lim\limits_{\substack{x \to 1 \\ x < 1}} B(x) = B$ und $\lim\limits_{\substack{x \to 1 \\ x < 1}} C(x) = C$, auch

$$A \cdot B = C.$$

(d) Als weiteres Beispiel behandeln wir die Taylor-Reihe von arctan und gehen ähnlich wie bei Beispiel (c) vor.

Wir nehmen das Ergebnis vorweg:

21.5.19 Satz

Für alle $x \in \mathbb{R}$ mit $|x| \leq 1$ gilt

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \pm \ldots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}$$

Speziell gilt

$$\arctan 1 = \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \pm \ldots$$

Denn für $|x| < 1$ gilt

$$\arctan x = \int_0^x \frac{1}{1+t^2} dt = \int_0^x \sum_{k=0}^{\infty} (-1)^k t^{2k} dt = \sum_{k=0}^{\infty} (-1)^k \int_0^x t^{2k} dt = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}.$$

(da $\sum (-1)^k t^{2k}$ auf $[0, |x|]$ gleichmäßig konvergiert)

Für $x = 1$ folgt die Gültigkeit der Formel aus dem Abelschen Grenzwertsatz, da nach dem Leibniz-Kriterium die Reihe $\sum\limits_{k=0}^{\infty} \frac{(-1)^k}{k+1}$ konvergiert.
Da der Abelsche Grenzwertsatz aber auch für den Fall gilt, dass die Potenzreihe $\sum a_k x^k$ noch für $x = -r$ konvergiert, gilt die Darstellung für alle $x \in \mathbb{R}$ mit $|x| \leq 1$.

Bemerkung: Zur praktischen Berechnung von π ist die Formel

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{1} \pm \ldots$$

nicht sonderlich gut geeignet, da die Reihe zu langsam konvergiert.

Der BBP-Algorithmus (vgl. ???) ist da wesentlich effizienter.
Im Vor-Computer-Zeitalter hat man mit der Darstellung (sog. Machinsche Formel)

$$\frac{\pi}{4} = 4 \arctan \frac{1}{5} - \arctan \frac{1}{239}$$

gearbeitet, die schneller konvergiert.

(e) **Binomische Reihe**
Eine wichtige Reihe, die als Spezialfall sowohl die binomische Formel als auch die geometrische Reihe umfasst, ist die *binomische Reihe*, dahinter versteht man die Reihe

$$b_\alpha(x) := \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \qquad \alpha \in \mathbb{R}.$$

$\binom{\alpha}{k}$ ist dabei der schon in ??? definierte Binomialkoeffizient

$$\binom{\alpha}{0} := 1; \quad \binom{\alpha}{k} = \frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{1 \cdot 2 \cdot 3 \cdot \ldots \cdot k} \quad (k \geq 1).$$

Wir werden sehen, dass durch diese Reihe die Potenzfunktion $x \mapsto (1+x)^\alpha = \exp(\alpha \log(1+x))$ wenigstens in $]-1,1[$ dargestellt wird. Berechnet man nämlich die Taylor-Koeffizienten von $f(x) = (1+x)^\alpha$ zum Entwicklungspunkt $a = 0$, so erhält man

$$f^{(k)}(x) = \alpha(\alpha-1)\dots(\alpha-k+1)(1+x)^{\alpha-k} = k! \left(\begin{array}{c} \alpha \\ k \end{array} \right) (1+x)^{\alpha-k}$$

und damit $\frac{f^{(k)}(0)}{k!} = \left(\begin{array}{c} \alpha \\ k \end{array} \right)$ und daher

$$T_{\infty,0}(f)(x) = \sum_{k=0}^{\infty} \left(\begin{array}{c} \alpha \\ k \end{array} \right) x^k = b_\alpha(x).$$

Die Frage ist zunächst, für welche $x \in \mathbb{R}$ die Binomialreihe $b_\alpha(x)$ konvergiert.
Um triviale Fälle auszuschließen, nehmen wir $\alpha \notin \mathbb{N}_0$ und $x \neq 0$ an.
Wir wenden das Quotientenkriterium mit $a_n := \left(\begin{array}{c} \alpha \\ n \end{array} \right) x^n$ an:

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \left| \frac{\left(\begin{array}{c} \alpha \\ n+1 \end{array} \right) x^{n+1}}{\left(\begin{array}{c} \alpha \\ n \end{array} \right) x^n} \right| = |x| \underbrace{\lim_{n\to\infty} \left| \frac{\alpha - n}{n+1} \right|}_{=1} = |x|.$$

Ist also $|x| < 1$, so ist die Binomialreihe nach der Limesform des Quotientenkriteriums konvergent.

Um zu zeigen, dass für $|x| < 1$ gilt

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \left(\begin{array}{c} \alpha \\ k \end{array} \right) x^k := b_\alpha(x)$$

oder m.a.W. (mit $f(x) = (1+x)^\alpha$)

$$f(x) = T_{\infty,0}(f)(x) = b_\alpha(x),$$

könnte man zeigen, dass das entsprechende Restglied $R_{n+1}(x)$ für $|x| < 1$ gegen Null konvergiert. Das ist im Prinzip möglich, wir gehen aber anders vor und benutzen einen Differetialgleichungstrick.
Aus

$$b_\alpha(x) = \sum_{k=0}^{\infty} \left(\begin{array}{c} \alpha \\ k \end{array} \right) x^k \quad (|x| < 1)$$

folgt

$$b'_\alpha(x) = \sum_{k=1}^{\infty} k \left(\begin{array}{c} \alpha \\ k \end{array} \right) x^{k-1} = \sum_{k=0}^{\infty} (k+1) \left(\begin{array}{c} \alpha \\ k+1 \end{array} \right) x^{k+1}.$$

Da aber

$$(k+1) \left(\begin{array}{c} \alpha \\ k+1 \end{array} \right) = (k+1) \frac{\alpha(\alpha-1)\dots(\alpha-k+1)(\alpha-k)}{1 \cdot 2 \cdot \ldots \cdot k \cdot (k+1)} = \alpha \left(\begin{array}{c} \alpha - 1 \\ k \end{array} \right)$$

gilt, ergibt sich weiter

$$b'_\alpha(x) = \alpha b_{\alpha-1}(x).$$

Nun ist aber

$$\begin{aligned}(1+x)b_{\alpha-1}(x) &= (1+x)\sum_{k=0}^{\infty}\binom{\alpha-1}{k}x^k \\ &= 1+\sum_{k=1}^{\infty}\underbrace{\left\{\binom{\alpha-1}{k}+\binom{\alpha-1}{k-1}\right\}}_{=\binom{\alpha}{k}}x^k \\ &= \sum_{k=0}^{\infty}\binom{\alpha}{k}x^k.\end{aligned}$$

Also gilt für alle $x \in \mathbb{R}$ mit $|x| < 1$ die (Differential-)Gleichung

$$(*) \qquad (1+x)b'_\alpha(x) = \alpha b_\alpha(x).$$

Da für $|x| < 1$ aber $(1+x)^\alpha \neq 0$ gilt, so folgt aus $(*)$, dass $x \mapsto b_\alpha(x)(1+x)^{-\alpha}$ in $]-1,1[$ die Ableitung Null hat, dass also $x \mapsto b_\alpha(x)(1+x)^{-\alpha}$ dort konstant ist.
Für $x = 0$ erhält man, dass die Konstante gleich 1 sein muss. Daher gilt

$$(1+x)^\alpha = b_\alpha(x)$$

für alle $x \in]-1,1[$ und beliebige $\alpha \in \mathbb{R}$.
Ohne Beweis sei mitgeteilt, dass die Formel auch noch im Fall $x = -1$ gilt, falls $\alpha > 0$ ist und im Fall $x = 1$, falls $\alpha > -1$. Zu allen anderen Fällen divergiert die Binomialreihe $b_\alpha(x)$ (man beachte $\alpha \notin \mathbb{N}_0$ für $\alpha \in \mathbb{N}_0$ „erstarrt" die Binomialreihe zur Binomischen Formel).

Fassen wir zusammen:

21.5.20 Satz

Für $\alpha \in \mathbb{R} - \mathbb{N}_0$ gilt

$$(1+x)^\alpha = \sum_{k=0}^{\infty}\binom{\alpha}{k}x^k, \text{ falls } \begin{cases} |x| < 1, & \\ x = -1, & \text{und } \alpha > 0 \\ x = 1, & \text{und } \alpha > -1. \end{cases}$$

Für alle anderen $x \in \mathbb{R}$ divergiert die Binomialreihe.

21.5.21 Beispiele und Bemerkungen zur Binomialreihe

(a) Für $\alpha = -1$ mit $\binom{-1}{k} = (-1)^k$, daher ergibt sich für $\alpha = -1$ aus der binomischen Reihe für $|x| < 1$

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 \pm \ldots = \sum_{k=0}^{\infty}(-1)^k x^k.$$

Ersetzt man x durch $-x$, so ergibt sich

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots = \sum_{k=0}^{\infty} x^k.$$

Die geometrische Reihe ist also ein Spezialfall der Binomischen Reihe.
Die häufig verwendeten Näherungsformel

$$\frac{1}{1-x} \approx 1+x \text{ bzw. } \frac{1}{1+x} \approx 1-x$$

(für kleine $|x|$) sind nun evident.
Wichtige Spezialfälle sind $\alpha = \frac{1}{2}$ und $\alpha = -\frac{1}{2}$. Berechnet man im Fall $\alpha = \frac{1}{2}$ die ersten Binomialkoeffizienten, so erhält man

$$\binom{\frac{1}{2}}{0} = 1, \qquad \binom{\frac{1}{2}}{1} = \tfrac{1}{2}, \qquad \binom{\frac{1}{2}}{2} = \tfrac{\frac{1}{2}(-\frac{1}{2})}{1\cdot 2} = -\tfrac{1}{8},$$

$$\binom{\frac{1}{2}}{3} = \tfrac{(\frac{1}{2})(-\frac{1}{2})(-\frac{3}{2})}{1\cdot 2\cdot 3} = \tfrac{1}{16}, \quad \binom{\frac{1}{2}}{4} = \binom{\frac{1}{2}}{3}\tfrac{-\frac{5}{2}}{4} = -\tfrac{5}{128}$$

und damit für $|x| < 1$

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 - \frac{5}{128}x^4 + \mathcal{O}(x^6),$$

dabei steht $\mathcal{O}(x^6)$ für eine Potenzreihe, die mit dem Glied cx^6 $(c \in \mathbb{R}_+^*)$ beginnt.

Wie schon früher im Fall der linearen Approximation $\sqrt{1+x} \approx 1 + \frac{x}{2}$ ausgeführt, kann man diese Formel benutzen um Wurzeln näherungsweise zu berechnen und den dabei gemachten Fehler abzuschätzen. Um etwa $\sqrt{10}$ näherungsweise zu berechnen, könnte man in erster Näherung $\sqrt{10} = \sqrt{1+9} \approx 1+\frac{9}{2} = 5,5$ rechnen, der Näherungswert ist aber viel zu schlecht. Um die Formel für $|x| < 1$ zu benutzen, schreiben wir

$$\sqrt{10} = \sqrt{9\cdot\frac{10}{9}} = 3\cdot\sqrt{1+\frac{1}{9}} = 3\left(1+\frac{1}{2\cdot 9} - \frac{1}{8\cdot 9^2} + \dots\right) = 3 + \frac{1}{6} - \frac{1}{8\cdot 27} + \dots \approx 3,162.$$

Der „exakte“ Wert ist

$$\sqrt{10} = 3,1622776602\dots$$

Für $\alpha = -\frac{1}{2}$ erhält man

$$\binom{-\frac{1}{2}}{0} = 1 \text{ und } \binom{-\frac{1}{2}}{k} = (-1)^k \frac{1\cdot 3\cdot 5\cdot\ldots(2k-1)}{2\cdot 4\cdot\ldots\cdot 2k}$$

für $k \geq 1$, also z.B.

$$\binom{-\frac{1}{2}}{1} = -\frac{1}{2};, \binom{-\frac{1}{2}}{2} = \frac{3}{8}, \binom{-\frac{1}{2}}{3} = -\frac{5}{16}$$

und daher (für $|x| < 1$)

$$\boxed{\tfrac{1}{\sqrt{1+x}} = 1 - \tfrac{1}{2}x + \tfrac{3}{8}x^2 - \tfrac{5}{16}x^3 + \tfrac{35}{128}x^4 \mp \dots}$$

Anwendung: Nach A.Einstein beträgt die Energie einer „relativistischen“ Teilchens der Masse $m = \frac{m_0}{\sqrt{1-(\frac{v}{c})^2}}$ (m_0 Ruhemasse, v die Geschwindigkeit des Teilchens, c Lichtgeschwindigkeit)

$$E = mc^2.$$

Die kinetische Energie ist definiert durch

$$E_{kin} = mc^2 - m_0c^2.$$

Ist $v < c$, so kann man zur Berechnung von E_{kin} die binomische Reihe (mit $x = -(\frac{v}{c})^2$) verwenden:

$$\begin{aligned} E_{kin} &= mc^2 - m_0c^2 \\ &= m_0c^2\left(\frac{1}{\sqrt{1-(\frac{v}{c})^2}} - 1\right) \\ &= m_0c^2\left(\frac{1}{2}(\frac{v}{c})^2 + \frac{3}{8}(\frac{v}{c})^4 + \frac{35}{128}(\frac{v}{c})^8 + \ldots\right) \\ &= \frac{1}{2}m_0v^2 + \frac{3}{8}m_0v^2(\frac{v}{c})^2 + \text{ Glieder höherer Ordnung} \end{aligned}$$

Der Term $\frac{1}{2}m_0v^2$ repräsentiert die kinetische Energie im klassischen Fall (v sehr klein gegenüber c), der sog. „3/8-Term“ $\frac{3}{8}m_0v^2(\frac{v}{c})^2$ ist das Glied niedrigster Ordnung der Abweichung zwischen dem relativistischen und nicht-relativistischen Fall.

Die Beispiele haben gezeigt, dass zum Aufstellen der Taylor-Entwicklung einer Funktion die Benutzung der Taylorschen Formel mit Restglied und dem Nachweis $\lim\limits_{n\to\infty} R_{n+1}(x) = 0$ häufig schwerfällig ist. Es ist häufig günstiger, bekannte Reihen zu differenzieren oder zu integrieren (vgl. die Beispiele ???).
Ferner kann man versuchen, die gegebene Funktion als Summe und/oder Produkt von Funktionen mit bekannten Reihenentwicklung darzustellen, z.B.

$$\begin{aligned} \cosh x &= \frac{1}{2}(\exp(x) + \exp(-x)) \\ &= \frac{1}{2}\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \ldots\right) \\ &= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6} + \ldots = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \quad (x \in \mathbb{R}) \end{aligned}$$

$$\begin{aligned} \frac{\cos x}{1-x} &= \cos x \frac{1}{1-x} \\ &= \left(\sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!}\right) \cdot \sum_{l=0}^{\infty} x^l \\ &= 1 + x + \left(1 - \frac{1}{2!}\right)x^2 + \left(1 - \frac{1}{2!}\right)x^3 + \left(1 - \frac{1}{2!} + \frac{1}{4!}\right)x^4 + \ldots, \quad |x| < 1 \end{aligned}$$

In der folgenden Tabelle sind einige häufig benutzte Taylor-Entwicklungen zusammengestellt.

Funktion	Taylor-Reihe um $a = 0$	Gültigkeitsbereich
$\frac{1}{1-x}$	$\sum\limits_{k=0}^{\infty} x^k$ (geom. Reihe)	$\lvert x\rvert < 1$
$\exp x$	$\sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$	$x \in \mathbb{R}$
$\sin x$	$\sum\limits_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots$	$x \in \mathbb{R}$
$\cos x$	$\sum\limits_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} - \frac{x^4}{4!} + \ldots$	$x \in \mathbb{R}$
$\sinh x$	$\sum\limits_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots$	$x \in \mathbb{R}$
$\cosh x$	$\sum\limits_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots$	$x \in \mathbb{R}$
$(1+x)^\alpha$	$\sum\limits_{k=0}^{\infty} \binom{\alpha}{k} x^k$	$\lvert x\rvert < 1$ $(\alpha \in \mathbb{R} - \mathbb{N}_0)$
$(1+x)^{\frac{1}{2}}$	$1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{10}x^3 - \frac{5}{128}x^4 \pm \ldots$	$\lvert x\rvert < 1$
$(1+x)^{-\frac{1}{2}}$	$1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + \frac{35}{128}x^4 \mp$	$\lvert x\rvert < 1$
$\log(1+x)$	$\sum\limits_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \pm \ldots$	$-1 \leq x \leq 1$
$\arctan x$	$\sum\limits_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \pm \ldots$	$-1 \leq x \leq 1$
$\arcsin x$	$\sum\limits_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} \frac{x^{2k+1}}{2k+1} = x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \ldots$	$-1 \leq x \leq 1$
	beachte: $(-1)^k \binom{-\frac{1}{2}}{k} = \frac{1\cdot 3\cdot 5\cdot \ldots \cdot (2k-1)}{2\cdot 4\cdot \ldots \cdot 2k}$ für $k \geq 1$	

Die Reihenentwicklung von arcsin erhält man aus der Reihenentwicklung der Ableitung $\arcsin' x = \frac{1}{\sqrt{1-x^2}}$ $(\lvert x\rvert < 1)$ und durch Anwendung des Abelschen Grenzwertsatzes durch gliedweise Integration.

21.6 Iterationsverfahren zur Berechnung von Fixpunkten und Nullstellen (Newton-Verfahren)

21.6.1 Einführung:

Das Lösen von Gleichungen der verschiedensten Art ist eines der ältesten Probleme der Mathematik. Bereits die Babylonier haben Verfahren zur Lösung der Gleichung

$$(1) \qquad x^2 = a \qquad (a \in \mathbb{R},\ a > 0)$$

entwickelt (vergl.).

Die Gleichung (1) ist äquivalent mit der *Fixpunktgleichung*

$$(2) \qquad \frac{1}{2}\left(x + \frac{a}{x}\right) = x\,.$$

Wir hatten gezeigt: Wählt man einen beliebigen Startwert $x_0 > a+1$ und definiert rekursiv (iterativ)

$$x_{n+1} := \frac{1}{2}\left(x_n + \frac{a}{x_n}\right);\ n \in \mathbb{N}_0,$$

dann konvergiert die Folge (x_n) (sogar monoton fallend) gegen die eindeutig bestimmte Lösung ξ (Fixpunkt) der Gleichung (2), d.h. es gilt $\xi = \sqrt{a}$.

Auch beim Problem ('*Multiplizieren statt dividieren*') (vergl.) waren wir auf eine Fixpunktgleichung gestoßen. Statt das Problem zu einer positiven reellen Zahl a die inverse Zahl (bezüglich der Multiplikation) zu bestimmen, also die Gleichung

$$ax = 1 \tag{3}$$

zu lösen, haben wir ein äquivalentes Problem gelöst: Wir haben die von Null verschiedene Lösung der Gleichung

$$x = 2x - ax^2 \tag{4}$$

bestimmt.

Setzt man $g(x) := 2x - ax^2$, so hat man wieder das Fixpunktproblem $g(x) = x$ zu lösen.
Wir hatten einen beliebigen Startwert x_0 mit $0 < x_0 < \frac{1}{a}$ gewählt und gesehen, dass die durch

$x_{n+1} = g(x_n)$ definierte rekursive Folge "sehr schnell" gegen die eindeutig bestimmte von Null verschiedene Lösung von (4), also gegen $\xi := \frac{1}{a}$ konvergiert.

Der Zwischenwertsatz (vergl.) liefert die *Existenz von Fixpunkten*: Ist $f : [a, b] \to \mathbb{R}$ eine stetige Abbildung und gilt $f([a, b]) \subset [a, b]$, dann gilt es (mindestens) ein $\xi \in [a, b]$ mit $f(\xi) = \xi$.

Sind $f, g : D \to \mathbb{R}$ etwa stetige Funktionen gemeinsamen Definitionsbereich, so kann man sich für die Menge

$$\{x \in D;\ f(x) = g(x)\} \tag{5}$$

also die Menge der Schnittpunkte der Graphen von f und g interessieren.

Diese geometrische Interpretation ist nützlich, weil eine Skizze der Graphen (die man selbst anfertigt oder die man von einem GTR oder CAS geliefert bekommt) schon häufig eine gute Information über die Anzahl und die (ungefähre) Lage des Schnittpunktes gibt.

Beispiel $f(x) = \cos x,\ g(x) = x$.

Hier sind wir wieder auf ein Fixpunktproblem gestoßen: In der graphischen Interpretation geht es um das Auffinden der Schnittpunkte der Graphen von f mit der Winkelhalbierenden des 1. und 3. Quadranten in einem cartesischen Koordinatensystem.

Wählt man in (5) für f die Nullfunktion, d.h. $f(x) = 0$ für alle $x \in D$, dann hat man alle $x \in D$ mit

$$(6) \qquad g(x) = 0$$

zu bestimmen. Geometrisch sind also die Schnittpunkte des Graphen von g mit der reellen Achse zu bestimmen.

Dieses Problem zu lösen ist für eine linear-affine Funktion trivial, für nicht-lineare Funktionen kann es recht kompliziert sein.

Das Nullstellenproblem und das Fixpunktproblem sind im Prinzip nicht verschieden; man kann sie in verschiedener Weise einander überführen.

Aus $g(x) = 0$ erhält man durch Addition von x das Fixpunktproblem

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

Umgekehrt erhält man aus dem Fixpunktproblem $f(x) = x$ durch Subtraktion von x das Nullstellenproblem

$$g(x) := f(x) - x = 0\,.$$

Aus dem Fixpunktproblem

$$f(x) := \exp(-x) = x$$

erhält man durch Anwendung der Umkehrfunktion $\log$ von $\exp$ das Nullstellenproblem

$$g(x) := x + \log x = 0\,.$$

Welche Übergänge gewählt werden und ob das gestellte Problem in der Fixpunktform oder in der Nullstellenform gelöst wird, hängt von den jeweiligen Bedingungen ab. Für beide Probleme gibt es spezielle unter etwas verschiedenen Voraussetzungen arbeitende Lösungstechniken, von denen wir zwei ausführlich behandeln:

- den *Fixpunktsatz für kontrahierende Selbstabbildungen* einer nichtleeren abgeschlossene Teilmenge $A \subset \mathbb{C}$ (es ist ein Spezialfall des allgemeinen *Banachschen Fixpunktsatzes* (vergl.).
- das *Newton-Verfahren* zur Nullstellenbestimmung (das auch im Komplexen funktioniert, wir beschränken uns aber hier auf den reellen Fall).

Zur Einstimmung auf mögliche Verallgemeinerungen versuche der geneigte Leser (die geneigte Leserin) ein $x := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4$ zu bestimmen, welches die vier Gleichungen

$$\begin{aligned}
x_1^3 - x_2^3 - 3x_1x_2x_4 - 8 &= 0 \\
\sqrt{25 - x_1^2} + 8x_3 + 4 &= 0 \\
x_1 + x_2 + x_3 + x_4 - 5 &= 0 \\
2x_1\,x_2\,x_3 - x_4 + 8 &= 0
\end{aligned}$$

offensichtlich ist das Gleichungssystem für $x_1 = 3$, $x_2 = 1$, $x_3 = -1$ und $x_4 = 2$ erfüllt. Gibt es noch andere Lösungen?

Für m *lineare* Gleichungen in n Unbekannten liefert die Lineare Algebra Lösungsverfahren. Der Schwerpunkt unserer Überlegungen liegt bei der *numerischen* Bestimmung der Lösungen und *Fehlerabschätzungen*.

21.6.2 Theorem: Fixpunktsatz für kontrahierende Abbildungen, Kontraktionssatz (Spezialfall des allgemeinen Banachschen Fixpunktsatzes)

Es sei A eine nichtleere *abgeschlossene* Teilmenge von $\mathbb{C}$ und
$f : A \to \mathbb{C}$ eine Funktion mit folgenden Eigenschaften:

(a) Für alle $x \in A$ ist $f(x) \in A$ (d.h. $f(A) \subset A$, m. a. W. f ist eine Selbstabbildung von A).

(b) f ist eine Kontraktion (man sagt auch: f sei kontrahierend), d. h. es gibt ein $q \in \mathbb{R}$ mit $0 \le q < 1$, so dass für alle $x, y \in A$ gilt

$$\boxed{(*)\ \ |f(x) - f(y)| \ \le q|x - y|}\,.$$

Dann gilt

1. f besitzt in A genau einen Fixpunkt, d.h. es gibt genau ein $\xi \in A$ mit $f(\xi) = \xi$.

   Für *jeden* Startwert $x_0 \in A$ konvergiert dadurch

   $$(**)\ \ x_{n+1} := f(x_n)$$

   rekursiv definierte Folge (x_n) gegen den Fixpunkt ξ:

   $$\lim_{n\to\infty} x_n = \xi = f(\xi)\,.$$

2. Es gelten die *Fehlerabschätzungen*

   $$(***)\ \ |x_n - \xi| \le \frac{q}{1-q}|x_n - x_{n-1}| \le \frac{q^n}{1-q}|x_1 - x_0|\,(n \in \mathbb{N})\,.$$

Beweis: Die Kontraktionseigenschaft $(*)$ impliziert zunächst $(k \in \mathbb{N}_0)$.

$$\begin{aligned}(****) : |x_{n+1} - x_n| = |f(x_n) - f(x_{n-1}| &\le q|x_n - x_{n-1}| \\ &\le q^2|x_{n-1} - x_{n-2}|^2 \\ &\ \vdots \\ &\le q^n|x_1 - x_0|\end{aligned}$$

Für fixiertes n und $m > n$ folgt daher

$$\begin{aligned}|x_m - x_n| &= |x_m - x_{m-1} + x_{m-1} - x_{m-2} + \cdots + x_{n+1} - x_n| \\ &\le |x_m - x_{m-1} + |x_{m-1} - x_{m-2}| + \cdots + |x_{n+1} - x_n| \\ &\le q^{m-1}|x_1 - x_0| + q^{m-2}|x_1 - x_0| + \cdots + q^n|x_1 - x_0| \\ &= q^n(q^{m-n-1} + \cdots + q^2 + q + 1)\,|\,x_1 - x_0|\end{aligned}$$

und wegen

$$1 + q + q^2 + \cdots + q^{m-n-1} = \frac{1 - q^{m-n}}{1-q} \le \frac{1}{1-q}$$

folgt

$$(*1) :\ \ |x_m - x_n| \le \frac{q^n}{1-q}|x_1 - x_0|\ \ (m > n)\,.$$

Da die Folge (q^n) aber wegen $0 \le q < 1$ eine Nullfolge ist, ist die Folge (x_n) eine Cauchy-Folge, also konvergent. Für ihren Grenzwert $\xi := \lim\limits_{n\to\infty} x_n$ gilt:

1. $\xi \in A$, da A abgeschlossen ist (hier geht die Folgen-Abgeschlossenheit von A ein!)

2. ξ erfüllt die Gleichung $f(\xi) = \xi$, d. h. ξ ist Fixpunkt, denn aus der Rekursionsgleichung
$$x_{n+1} = f(x_n)$$
folgt für $n \to \infty$ wegen der (Lipschitz-)Stetigkeit von f auch
$\xi = \lim\limits_{n\to\infty} x_{n+1} = \lim\limits_{n\to\infty} f(x_n) = f(\lim\limits_{n\to\infty} x_n) = f(\xi)$, d. h. ξ ist Fixpunkt von f.

3. ξ ist der einzige Fixpunkt von f in A. Denn wäre $\eta \in A$ ebenfalls Fixpunkt von f und $\eta \neq \xi$, dann erhält man
$$|\xi - \eta| = |f(\xi) - f(\eta)| \leq q|\xi - \eta|$$
und damit $q \geq 1$ im Widerspruch zur Voraussetzung $0 \leq q < 1$.

Die Abschätzung
$$(**2): \quad |x_n - \xi| \leq \frac{q^n}{1-q}|x_1 - x_0|$$
erhält man aus $(*1)$ durch Grenzübergang $m \to \infty$.

Wir zeigen noch
$$(**1): \quad |x_n - \xi| \leq \frac{q}{1-q}|x_n - x_{n-1}| \qquad (n \in \mathbb{N}).$$
Aus dieser Abschätzung ergibt sich mittels $(****)$
$$\begin{aligned} |x_n - \xi| \leq \frac{q}{1-q}|x_n - x_{n-1}| &\leq \frac{q}{1-q}q^{n-1}|x_n - x_0| \\ &= \frac{q^n}{1-q}|x_1 - x_0|, \end{aligned}$$
also die Abschätzung $(**2)$. Zum Beweis von $(**1)$:
$$\begin{aligned} \text{Aus } |x_n - \xi| = |f(x_{n-1}) - f(\xi)| &\leq q|x_{n-1} - \xi| \\ &= q|x_{n-1} - x_n + x_n - \xi| \\ &\leq q(|x_{n-1} - x_n| + |x_n - \xi|) \end{aligned}$$
folgt
$$\begin{aligned} (1-q)|x_n - \xi| &\leq q|x_n - x_{n-1}| \text{ oder} \\ |x_n - \xi| &\leq \frac{q}{1-q}|x_n - x_{n-1}| \text{ und damit} \\ |x_n - \xi| &\leq \frac{q}{1-q}q^{n-1}|x_1 - x_0| = \frac{q^n}{1-q}|x_1 - x_0|. \end{aligned}$$
Aus der Abschätzung $(**2)$ erhält man umgekehrt auch die Abschätzung $(**1)$:

Für $n = 0$ in $(**2)$ erhält man
$$|x_0 - \xi| \leq \frac{1}{1-q}|x_1 - x_0|.$$
Beachtet man, dass der Startwert $x := x_0 \in A$ beliebig gewählt war, so folgt wegen $f(x) = x_1$ aus $(**2)$
$|x - \xi| \leq \frac{1}{1-q}|f(x) - x|$ für alle $x \in A$, speziell für $x = x_n$ also $|x_n - \xi| \leq \frac{1}{1-q}|(f(x_n) - x_n| = \frac{1}{1-q}|x_{n+1} - x_n| \leq \frac{q}{1-q}|x_n - x_{n-1}|$.

21.6.3 Bemerkungen

a) Da man jeden gewonnenen Wert $x_n \in A$ als neuen Startwert wählen kann, ist das Verfahren relativ unempfindlich gegen Rundungsfehler und Fehlerfortpflanzung.

b) Das Theorem gilt speziell für eine kontrahierende Selbstabbildungen eines kompakten (und damit speziell abgeschlossenen) Intervalls $A := [a, b] \subset \mathbb{R}$.

c) Ist $f : [ab] \to [a, b]$ stetig differenzierbar, so ist f sicher dann eine Kontraktion, wenn

$q := \max\{|f'(x)|;\ x \in [a, b]\} < 1$ gilt.

Dies ergibt sich sofort aus dem Mittelwertsatz der Differentialrechnung $(MWSD)$ oder dem Schrankensatz. Wegen $\xi - x_{n+1} = f(\xi) - f(x_n) = f'(\xi_n)(\xi - x_n)$ mit $\xi_n \in [\xi, x_n]$ oder $\xi_n \in [x_n, \xi]$ ist die Konvergenz linear und für große n ist die Konvergenzrate jeweils etwa $|f'(\xi)|$.

d) Die Bedingung, dass f eine Kontraktion ist, ist lediglich eine *hinreichende Bedingung* für die Existenz eines Fixpunkts, sie ist nicht notwendig, wie das Beispiel

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto -2x + 3 \end{aligned}$$

zeigt. f hat den Fixpunkt 1, ist aber keine Kontraktion! Warum nicht?

e) Im Fall $f; [a, b] \to [a, b]$ ergibt der Fixpunkt ξ den Schnittpunkt (ξ, ξ) des Graphen von f mit der Diagonalen $\Delta := \{(x, x),\ x \in [a, b]\}$ des Quadrats $[a, b] \times [a, b]$.

Der Ablauf der Iteration $x_{n+1} = f(x_n =)$ stellt sich geometrisch übersichtlich dar durch den Streckenzug, der bei (x_0, x_0) beginnt und der Reihe nach die Punkte $(x_n, x_n), (x_n, x_{n+1})$, (x_{n+1}, x_{n+1}) etc. verbindet:

Beispiel: Monotone Konvergenz. Beispiel: Alternierende Konvergenz

Auch die Divergenz kann man sich graphisch veranschaulichen

Abstoßende Fixpunkte

Ist $|f'(\xi)| > 1$ so kann zwar f einen Fixpunkt ξ haben, aber auch wenn man den Startwert noch so nahe bei ξ wählt ($\xi_0 \neq \xi$), divergiert die Iterationsfolge (x_n), der Fixpunkt heißt dann ein *abstoßender Fixpunkt*'.

f) Aussagen über die Existenz von Fixpunkten und Verfahren zur deren Ermittlung starke Hilfsmittel der Analysis und der numerischen Mathematik dar. Fixpunktsätze werden in den verschiedensten Zweigen der Mathematik und ihren Anwendungen verwendet, speziell in der Analysis, etwa für Existenzbeweise für Differentialgleichungen oder beim Satz über die Existenz einer lokalen Umkehrabbildung einer differenzierbare Funktion (von mehreren Veränderlichen). Wir kommen hierauf bald zu sprechen.

g) Gilt etwa $f \in C^p([a,b])$ und ist $f'(\xi) = f''(\xi) = \cdots = f^{(p-1)}(\xi) = 0$, aber $f^{(p)}(\xi) \neq 0$, so liefert die Taylor-Formel für den Entwicklungspunkt ξ

$$(!)\;\; \xi - x_{n+1} = f(\xi) - f(x_n) = \pm \frac{f^{(p)}(z_n)}{p!}(\xi - x_n)^p \;\text{ mit } z_n \in \left\{ \begin{array}{c} [x, x_n] \\ \text{oder} \\ [x_n, \xi] \end{array} \right.$$

und man hat sogar Konvergenz der Ordnung p, d. h. es gilt für alle $n \in \mathbb{N}_0$

$|x_{n+1} - \xi| \leq C|x_n - \xi|^p$, mit einer geeigneten Konstanten $C > 0$.

21.6.4 Ein Beispiel: Näherungsweise Berechnung von $\frac{\pi}{2}$

Wir definieren $f(x) := x + \cos x$. Dann ist $f(x) = x$ äquivalent mit $\cos x = 0$ und $\pi/2$ ist die einzige Lösung dieser Gleichung im Intervall $]0, \pi[$.

Wegen $f'(x) = 1 - \sin x$ hat man für $A := \left[\frac{\pi}{2} - r,\, \frac{\pi}{2} + r\right]$ mit $0 < r < \frac{\pi}{2},\; q := \max\{|f'(x)|;\; x \in A\} < 1$. Für $x \in A$ folgt dann auch

$$\left|f(x) - \frac{\pi}{2}\right| = \left|f(x) - f\left(\frac{\pi}{2}\right)\right| \leq q\left|x - \frac{\pi}{2}\right| \leq r\,,$$

d. h. es ist $f(A) \subset A$ und f ist eine Kontraktion.

Die Fixpunktfolge $x_{n+1} = f(x_n)$ konvergiert also für jeden Startwert $x_0 \in]0, \pi[$ und wegen $f'\left(\frac{\pi}{2}\right) = f''\left(\frac{\pi}{2}\right) = 0$ liegt *kubische Konvergenz* vor.

Wegen $|f'''(z)| \leq 1$ liefert (!) für den Fehler $\varepsilon_n := |x_n - \frac{\pi}{2}|$ die Abschätzung

$\varepsilon_{n+1} \leq \frac{1}{6}\,\varepsilon_n^3 \leq \frac{1}{6^4}\varepsilon_{n-1}^9 \leq \cdots \leq \frac{1}{6^{\alpha_n}}\,\varepsilon_0^{3^{(n+1)}}$ mit $\alpha_n := \sum\limits_{k=0}^{n} 3^k = \frac{1}{2}(3^{n+1} - 1)$.

Aus elementar-geometrischen Gründen oder nach Definition von $\frac{\pi}{2}$ als kleinste positive Nullstelle von $\cos$ gilt $1.4 < \sqrt{2} < \frac{\pi}{2} < 1.6 < 2$. Wählt man als Startwert $x_0 := 2$, dann ist $\varepsilon_0 \leq 0.6$..

Die folgende Tabelle für die Werte x_n, der Fehler $\varepsilon_n := |x_n - \frac{\pi}{2}|$ und die Fehlerschranken $A_n := \frac{1}{6^{\alpha_{n-1}}}(0,6)^{(3^n)}$ zeigen die extrem schnelle Konvergenz von x_n (kubische Konvergenz), jedoch muss stets $\cos x_n$ mit der gleichen Genauigkeit wie x_n berechnet werden:

n	x_n	ε_n	A_n
0	2	$0,43$	$0,6$
1	$1,58385$	$0,0013$	$0,0036$
2	$1,57079669778$	$3,71 \cdot 10^{-7}$	$7,8 \cdot 10^{-6}$
3	$1,57079632679489661923983$	$8,51 \cdot 10^{-21}$	$7,8 \cdot 10^{-77}$

vergleiche auch W. Kaballo: Einführung in die Analysis I, Spektrum, Akademischer Verlag, 2. Auflage S. 271.

Wir haben in der Einführung (20.6.1) gesehen, dass das *Fixpunktproblem* und das *Nullstellenproblem* äquivalente Probleme sind. Beim Fixpunktsatz (Kontraktionssatz) 20.6.2, haben wir keine Differenzierbarkeitsvoraussetzungen für die Funktion f benutzt. Wir haben jedoch gesehen (vergl. 20.6.3 Bem. g)), dass für den Fall, dass $f \in C^p([a, b])$ gilt und der Fixpunktξ eine Nullstelle der Ordnung p ist, diese Iterationsfolge (x_n) sogar mit der Ordnung p gegen den Fixpunkt konvergiert.

Beim *Newton-Verfahren* setzen wir voraus, dass $f : [a, b] \to \mathbb{R}$ *zweimal stetig differenzierbar* ist. Die Idee des klassischen Newton-Verfahrens (auch Newton-Raphson-Verfahren genannt)zur Nullstellenbestimmung lässt sich einfach geometrisch motivieren.

Dazu nehmen wir einmal an, der Graph von f habe die folgende geometrische Gestalt

und x_k sei bereits eine Näherung für die gesuchte Lösung ξ der Gleichung $f(x) = 0$. Betrachtet man jetzt die Tangente T an den Graphen von f im Punkt $(x_k, f(x_k))$, so hat diese die Gleichung

$$T(x) = f(x_k) + f'(x_k)(x - x_k)$$

Ihr Schnittpunkt x_{a+1} mit der x-Achse ist gegeben durch

$$0 = f(x_k) + f'(x_a)(x_{k+1} - x_k) \quad \text{oder}$$

$$\boxed{x_{k+1} = x_k - \tfrac{f(x_k)}{f'(x_k)}\,.}$$

"Offensichtlich" ist x_{k+1} eine bessere Näherung für ξ als x_k (es muss natürlich $f'(x_k) \neq 0$ sein). Jetzt wird man das Verfahren mit dem neuen Startwert x_{k+1} fortsetzen.

Wir werden sehen, dass unter geeigneten Voraussetzungen über f die *Newton-Folge*

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \qquad (k \in \mathbb{N}_0)\,,$$

– wobei noch ein geeigneter Startwert x_0 zu wählen ist – tatsächlich gegen die Nullstelle ξ konvergiert. Hier kommt die Grundidee der Differenzierbarkeit

– Differenzierbarkeit bedeutet in erster Näherung lineare Approximierbarkeit zum Tragen: Der Graph von f wird ersetzt durch die Tangente, also eine linear-affine Funktion, deren Schnittpunkt mit der x-Achse sich elementar berechnen lässt.

Man kann den Satz über das Newton-Verfahren unter verschiedenen Voraussetzungen formulieren. Wir wählen eine Formulierung, in welcher die Existenz einer (einfachen) Nullstelle der betreffenden Funktion vorausgesetzt wird (vergl. hierzu die Bemerkungen im Anschluss an das Theorem).

21.6.5 Theorem (Newton-Verfahren zur Nullstellenbestimmung)

Sei $[a, b] \in \mathbb{R}(a < b)$ ein kompaktes Intervall und $f : [a, b] \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion (also $f \in C^2/[a, b]$), die im offenen Intervall $]a, b[$ eine Nullstelle hat: $f(\xi) = 0$ für $\xi \in]a, b[$.

Ferner sei

$$m := \min\{|f'(x)|,\, x \in [a, b] > 0\} \;\text{ und }\; M := \max\{|f''(x)|;\; x \in [a, b]\}$$

Außerdem sei $\delta > 0$ so gewählt, dass

$$q := \frac{M}{2m}\delta < 1 \;\text{ und } K := [\xi - \delta, \xi + \delta] \subset [a, b] \text{ gilt}\,.$$

a) Für jeden Startwert $x_0 \in K$ liegen die Newton-Iterierten x_n in K $(x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},\;\; n \in \mathbb{N}_0$ und die Folge $(x_n) \geq 0$ der Newton-Iterierten konvergiert gegen die Nullstelle ξ.

   Für $n \in \mathbb{N}_0$ gilt

$$(*) \qquad x_{n+1} - \xi = \tfrac{1}{f'(xn)} \int_{x_n}^{\xi} f''(t)(\xi - t)\, dt \text{ und}$$

$$(**) \quad |x_{n+1} - \xi| \leq \tfrac{M}{2m} |x_n - \xi|^2$$

b) Die Folge (x_n) konvergiert gegen ξ (wegen $(**)$ sogar quadratisch). Genauer gelten die Fehlerabschätzungen

$$(3) \quad |x-\xi| \leq \tfrac{2m}{M} q^{(2^n}, \; n \in \mathbb{N} \text{ und}$$

$$(4) \quad |x_n-\xi| \leq \tfrac{1}{m}|f(x_n)| \leq \tfrac{M}{2m}|x_n-x_{n-1}|^2, \; n \in \mathbb{N}$$

Beweis: Wesentliches Hilfsmittel beim Beweis wird die Taylorsche Formel mit Restglied sein.

(a) Wir beweisen die Aussage $x_n \in K$ für alle $n \in \mathbb{N}_0$ mit Induktion nach n.

$x_0 \in K$ gilt nach Voraussetzung. Es sei bereits $x_k \in K$ für $k = 0, \dots, n$ gezeigt. Zu zeigen ist dann $x_{n+1} \in K$.

Wir benutzen die Taylorsche Formel mit Integralrestglied (vergl.) mit $x = \xi_n$ und $a = x_n$ und erhalten

$$0 = f(\xi) = f(x_n) + f'(x_n)(\xi - x_n) + \int_{x_n}^{\xi} f''(t)(\xi - t)\,dt$$

Subtrahiert man hiervon die Tangentengleichung $0 = f(x_n) + f'(x_n)(x_{n+1} - x_n)$, so erhält man

$$0 = f'(x_n)(\xi - x_{n+1}) + \int_{x_n}^{\xi} f''(t)(\xi - t)\,dt \;\text{ oder}$$

$$x_{n+1} - \xi = \frac{1}{f'(x_n)} \int_{x_n}^{\xi} f''(t)(\xi - t)\,dt\,.$$

Das ist die gerade die Gleichung (1). Aus den Voraussetzungen ergibt sich mit der Standardintegralabschätzung hieraus auch $|x_{n+1} - \xi| \leq \frac{M}{2m}|x_n - \xi|^2$, also (2) und damit auch weiter

$$|x_{n+1} - \xi| \leq \frac{M}{2m}|x_n - \xi|^2 \leq q|x_n - \xi| < |x_n - \xi| \leq \delta$$

und damit $x_n \in K$ für alle $n \in \mathbb{N}_0$.

(b) Für $n = 0$ ist (3) richtig. (3) gelte für ein beliebiges $n \in \mathbb{N}$ dann folgt mit (2)

$$|x_{n+1} - \xi| \leq \frac{M}{2m}|x_n - \xi|^2 \leq \frac{M}{2m}\left(\frac{2m}{M}\right)^2 q^{2\cdot 2^n} = \frac{2m}{M} q^{(2^{n+1})}$$

und damit (3) auch für $n + 1$.

Zum Nachweis von (4) verwenden wir zunächst den Mittelwertsatz der Differentialrechnung (MWSD) $f(x_n) - f(\xi) = f'(\xi_n)(x_n - \xi)$ mit einem geeigneten Zwischenwert $\xi_n \in \left\{ \begin{array}{l} [x_n, \xi] \\ [\xi, x_n] \end{array} \right.$

Damit folgt wegen $f(\xi) = 0$

$$|f(x_n)| = f'(\xi_n)|\, x_n - \xi| \geq m|x_n - \xi| \text{ oder}$$

(4a) : $|x_n - \xi| \leq \frac{|f'(x_n)|}{m}$.

Eine weitere Anwendung der Taylorschen Formel $(x = x_n;\ a = x_{n-1})$ mit Lagrange-Restglied liefert

$$\begin{aligned} f(x_n) &= \underbrace{f(x_{n-1}) + f'(x_{n-1})(x_n - x_{n-1})}_{=0} + \tfrac{f''(z_n)}{2!}(x_n - x_{n+1})^2 \\ &= \tfrac{f''(z_n)}{2!}(x_n - x_{n-1})^2 \end{aligned}$$

dabei ist z_n eine geeignete Zwischenstelle zwischen x_{n-1} und x_n.

Mit (4a) folgt dann

$$|x_n - \xi| \leq \frac{|f(x_n)|}{m} = \frac{|f''(z_n)|}{2m} \leq \frac{M}{2m}|x_n - x_{n-1}|^2$$

Das ist die (a-posteriori-) Abschätzung (4).

21.6.6 Bemerkungen:

a) In der Situation von Theorem 20.6.3 hat f wegen $f'(x) \neq 0$ für alle $x \in [a,b]$ und des Satzes von Rolle höchstens eine Nullstele in $]a,b[$, wegen der vorausgesetzten Existenz einer Nullstelle, hat f also genau eine Nullstelle in $]a,b[$,. Die Existenz einer solchen Nullstelle kann man häufig mit Hilfe des *Zwischenwertsatzes* zeigen.

b) Ist diese garantiert und ist $f'(\xi) \neq 0$ (d.h. ist ξ eine einfache Nullstelle), so existiert stets eine möglicherweise sehr kleine abgeschlossene Umgebung $K := K_\delta := [\xi - \delta,\, \xi + \delta]$ für welche die Voraussetzungen des Theorems erfüllt sind. Das Problem beim Newton-Verfahren besteht also darin, einen im "Einzugsbereich" der Nullstelle ξ gelegenen geeigneten Startwert x_0 zu finden.

Man vergleiche hierzu auch die Übungsaufgabe über "anziehende" und "abstoßende" Fixpunkte. Wir hätten das Newton-Verfahren auch als Fixpunktiteration für $N(x) = x$ auffassen können, dabei sei

$$N_f(x) := N(x) = x - \frac{f(x)}{f'(x)}$$

die zu f gehörige *Newton-Funktion*, wobei natürlich der Ausdruck $\frac{f(x)}{f'(x)}$ einen Sinn haben muss.

Hat man einen Startwert x_0 im Einzugsbereich der quadratischen Konvergenz gefunden, so konvergiert das Newton-Verfahren extrem schnell gegen die Nullstelle ξ:

Im Fall $q \leq \frac{1}{2}$ gilt nach 10 Iterationsschnitten – wenn man $2^{10} = 1024 > 10^3$ beachtet – bereits

$$|x_{10} - \xi| \leq \frac{2m}{M} q^{1000} \approx \frac{2m}{M} 10^{-300}\,.$$

c) Auch im Fall $f'(\xi) = 0$ ist das Newton-Verfahren nicht unbedingt zum Scheitern verurteilt, denn es könnte ja auch $f(\xi) = 0$ sein und der Grenzwert $\lim\limits_{s \to x} \frac{f(x)}{f'(x)}$ existieren

Übung: Man untersuche hierzu $a - B$

$$\begin{aligned} f : \mathbb{R} &\;\to\; \mathbb{R} \\ x &\;\mapsto\; x^p \;\; (p \in \mathbb{N},\, p \geq 2) \;\text{ mit }\; \xi = 0 \;\text{ und} \end{aligned}$$

beliebigem Startwert $x_0 \rightharpoondown 0$.

d) Ist $f \in C^{p+1}([a,b])$ und $\xi \in]a,b[$ eine p-fache Nullstelle von f, d. h. es gelte

$$f(\xi) = f'(\xi) = \cdots = f^{(p-1)}(\xi) = 0, \text{ aber } f^{(p)}(\xi) \neq 0,$$

dann konvergiert das Newton-Verfahren noch linear mit Konvergenzrate $1 - \frac{1}{p}$, d. h. für hinreichend große n gilt $x_{n+1} - \xi \sim (1 - \frac{1}{p})(x_n - \xi)\,.$

Modifiziert man in diesem Fall die Newton-Folge, indem man $x_{n+1} := x_n - p\frac{f(x_n)}{f'(x_n)}$ für $n \in \mathbb{N}_0$ setzt, so hat man für geeignete Startwerte x_0 wieder quadratische Konvergenz.

e) Wegen $m := \min\{|f'(x)|,\ x \in [a,b]\} > 0$ hat f' konstante Vorzeichen im Intervall $[a,b)$. Ändert auch f'' sein Vorzeichen nicht auf $[a,b]$, so ist f konvex oder konkav auf $[a,b]$ und hat $\frac{f''}{f'}$ das gleiche Vorzeichen wie $x_0 - \xi$, so impliziert die Bedingung "$\frac{M}{2m}\,\delta < 1$" wegen (1) in 20.6.5 ($*$) bereits $x_n \in \begin{cases} [x_0,\, \xi] \\ \text{oder} \\ [\xi,\, x_0] \end{cases}$ für alle n.

Dann konvergiert (x_n) *monoton* gegen ξ und die Bedingung $[\xi - \delta,\, \xi + \delta] \subset [a,b]$ ist überflüssig.

21.6.7 Beispiel:

Newton-Verfahren zur Berechnung der p-ten Wurzel aus einer positiven reellen Zahl.

Sei $a > 0$. Die p-te Wurzel aus $a(p \in \mathbb{N}, p \geq 2)$ ist (eindeutig bestimmte) Nullstelle $\xi > 0$ der Funktion $f : \mathbb{R}_+^\bullet \to \mathbb{R};\ x \mapsto x^p - a$. Das Newton-Verfahren liefert hier

$$x_{n+1} = x_n - \frac{x_n^p - a}{n x_n^{p-1}} = \frac{1}{n}\left\{\left((p-1)x_n + \frac{a}{x_n^{p-1}}\right)\right\}.$$

Für $n \to \infty$ konvergiert (x_n) nach Theorem gegen $\xi = \sqrt[p]{a}$, wenn nun der Startwert x_0 nahe genug bei ξ gewählt wird. Hier ist aber wegen der Konvexität und Monotonie von f die Konvergenz für *jeden* Startwert $x_0 > 0$ gesichert.

Die Folge $(x_n)_{n\in\mathbb{N}}$ ist in diesem Fall monoton fallend und konvergiert notwendig gegen $\sqrt[p]{a}$. Für hinreichend große n liegt dann x_n im Einzugsbereich der Nullstelle ξ und die Fehlerabschätzungen gelten mit diesem x_n als neuen Startwert.

Aus $x_0 > 0$ folgt $\left\{\begin{array}{l} x_n > \sqrt[p]{a},\ n \in \mathbb{N} \\ \sqrt[p]{a} < x_{n-1} < x_n \end{array}\right.$.

Den Spezialfall $p = 2$ haben wir früher (vergleiche) bereits ausführlich betrachtet. Wir wollen für diesen Fall den Einzugsbereich der quadratischen Konvergenz bestimmen. Es gilt (vergl.).

$$\begin{aligned}(x_{n+1} - \sqrt{a}) = \frac{1}{2}\left(x_n + \frac{a}{\sqrt{n}}\right) - \sqrt{a} \quad &= \quad \frac{1}{2x_n}\left(x_n^2 + a - 2x_n\sqrt{a}\right) \\ &= \quad \frac{1}{2x_n}\left(x_n - \sqrt{a}\right)^2,\end{aligned}$$

also für $n \geq 1$ (wegen $x_n > \sqrt{a}$)

$$\left|x_{n+1} - \sqrt{a}\right| \leq \frac{1}{2\sqrt{a}}\left|x_n - \sqrt{a}\right|^2.$$

Quadratische Konvergenz erhält man also für Startwerte x_0 mit der Eigenschaft

$$\frac{1}{2\sqrt{a}}\left|x_0 - \sqrt{a}\right| < 1 \ \text{ bzw. } \ \left|x_0 - \sqrt{a}\right| < 2\sqrt{a}.$$

Da die Folge $\left(\frac{a}{x_n}\right)$ monoton wachsend gegen $\sqrt{a}$ konvergiert, erhält man aus der Beziehung

$$\frac{a}{x_n} \leq \sqrt{a} \leq x_n$$

ein Abbruchkriterium: Man stoppt die Berechnung falls bei vorgegebenem ϵ

$$0 \leq x_n - \frac{a}{x_n} \leq \varphi$$

gilt.

Man vergleiche hierzu nochmals die Tabellen zur Berechnung von $\sqrt{2}$ in Abschnitt ? .

Auch bei absurder Wahl des Startwerts x_0 erhält man schnelle Konvergenz. Neben der schnellen Konvergenz ist das Newton-Verfahren zur Wurzelberechnung auch noch selbstkorrigierend, wenn man $\sqrt{2}$ etwa auf 100 Dezimalstellen genau berechnen, so müssten wir nicht die Rechnung von Anfang an mit 100-stelliger Genauigkeit durchführen. Wählt man den Startwert $x_0 = 2$, so liefert das Newton-Verfahren schon nach 6 Iterationsschritten 36 exakte Dezimalstellen. Nimmt man x_6 als neuen Startwert, erhält man nach zwei weiteren Iterationsschritten 100 exakte Dezimalstellen. Die Verhältnisse sind jedoch nicht immer so ideal wie im letzten Beispiel.

21.6.8 Ärger mit dem Newton-Verfahren

a) Betrachtet man das Beispiel

$$\begin{aligned} f:\mathbb{R} &\to \mathbb{R} \\ x &\mapsto \arctan x\,. \end{aligned}$$

mit der einzigen Nullstelle $\xi = 0$, so divergiert das Newton-Verfahren für jeden Startwert x_0 mit $|x_0| \geq 2$. Vergleiche Übungsaufgabe ? Blatt ?. Wählt man jedoch einen Startwert x_0 mit $|x_0| < 2$, so konvergiert das Newton-Verfahren gegen die Nullstelle $\xi = 0$ von $\arctan$.

b) **Zykel im Newton-Verfahren**

Betrachtet man

$$\begin{aligned} f:\mathbb{R} &\to \mathbb{R} \ \text{ mit} \\ x &\mapsto \operatorname{sign}(x)\sqrt{x}\,. \end{aligned}$$

Dann gilt für die entsprechende Newton-Funktion $N_f(x) = N(x) = x - \sqrt{x}\,|(1/(2\sqrt{x})| = -x$ für $x > 0$ und $N(-x) = x$ für $x < 0$.

Für *jeden* Startwert tritt das Newton-Verfahren auf der Stelle.

Beachte $|N'(x)| = 1$ für alle $x \in \mathbb{R}$.

21.7 Uneigentliche Integrale, Γ-Funktion, Stirlingsche Formel (wird nachgetragen)

21.8 Fourier-Reihen: Eine Einführung

Diese Einführung enthält:

0. Einführung, Motivation
1. Grundbegriffe: Trigonometrische Polynome und trigonometrische Reihen, der Begriff der *Fourier*-Reihe, erste Beispiele
2. Eine Minimaleigenschaft der Fourier-Polynome, Konvergenz im quadratischen Mittel, Besselsche Ungleichung.
3. Die Sätze von *Dirichlet* und *Féjer*, Beispiele und Anwendungen: Partialbruchzerlegung des Cotangens und Produktentwicklung von $\sin \pi x$.
4. Die Vollständigkeit des Systems $(e_k)_{k\in\mathbb{Z}}$ im Vektorraum V.
5. Eine Anwendung: Die isoperimetrische Ungleichung.

0. **Einführung, Motivation**
Eine der Grundaufgaben der Analysis ist es, ”relativ komplizierte” Funktionen mit Hilfe ”relativ einfacher” Funktionen (z. B. durch Polynome) darzustellen oder wenigstens zu approximieren. Sind z. B. von einer beliebigen Funktion: $f : \mathbb{R} \to \mathbb{R}$, von der wir allerdings voraussetzen, dass sie n-mal stetig differenzierbar ist, der Funktionswert $f(a)$ und die Werte $f^{(k)}(a)$, $k = 1, \ldots, n$ für eine Stelle $a \in \mathbb{R}$ bekannt, dann approximiert das n-te Taylor-Polynom
$$T_{n,a}(f)(x) := f(a) + \frac{f'(a)}{1}(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$
von f die Funktion f, in einer Umgebung de Stelle a in dem Sinne gut, dass für $0 \le k \le n$ die Ableitungen $f^{(k)}(a)$ und $T_{n,a}^{(k)}(f)(a)$ übereinstimmen und für den Fehler $R_{n,a}(f)(x) := f(x) - T_{n,a}(f)(x)$ gilt
$$R_{n,a}(f)(x) = (x-a)^n r(x)$$
mit einer stetigen Funktion r, für die $r(a) = 0$ gilt. Verallgemeinerungen von Polynomen sind Potenzreihen, jedoch ist die Klasse der Funktionen, die sich wenigstens lokal durch Potenzreihen darstellen lassen (die sog. (reell-) analytischen Funktionen) verhältnismäßig eng. Weil Potenzreihen stetig, differenzierbar, ja sogar beliebig oft differenzierbar sin, haben die dargestellten Funktionen die entsprechenden Eigenschaften.

Umso erstaunlicher ist es, dass mit Hilfe von Linearkombinationen aus
$$1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx, \ldots$$
relativ komplizierte ”willkürliche” Funktionen dargestellt werden können. Diese Entdeckung hat im 18. und 19. Jahrhundert wesentlich zur Begriffsbildung und zur Klärung von Grundlagen der Analysis beigetragen.

Eine Grundeigenschaft der Funktionen $\cos$ und $\sin$ ist ihre 2π-Periodizität: Es gilt für alle $x \in \mathbb{R}$
$$\cos(x + 2\pi) = \cos x \ \text{ und } \ \sin(x + 2\pi) = \sin x\,.$$
Neben diesen ”Grundschwingungen” haben aber auch die ”Oberschwingungen”
$$\cos(kx) \ \text{ bzw. } \ \sin(kx)\ (k \in \mathbb{N},\, k \ge 2)$$

die Periode 2π (und damit auch die Perioden $2\pi\ell,\ \ell \in \mathbb{Z}$). Idee ist es nun, möglichst allgemeine 2π-periodische Funktionen als Überlagerungen dieser harmonischen Schwingungen zu schreiben (also die Frequenzen zu erhöhen), d. h. endliche Summen

$$\frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \qquad \text{bzw. Reihen}$$
$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \quad a_k, b_k \in \mathbb{C}, x \in \mathbb{R}$$

zu betrachten. Warum man das das konstante Glied in der Form $\frac{a_0}{2}$ schreibt, wird sich bald zeigen.

Periodische Vorgänge spielen auch bei zahlreichen Anwendungen eine wichtige Rolle (Astronomie, Physik etc.).

21.8.1 Grundbegriffe:

Der Begriff der periodischen Funktion, trigonometrische Polynome und Reihen, der Begriff der Fourier-Reihe, erste Beispiele.

21.8.2 Definition:

Eine Funktion $f : \mathbb{R} \to \mathbb{C}$ heißt *periodisch* mit der Periode $p (\in \mathbb{R})$, falls für alle $x \in \mathbb{R}$ gilt

$$f(x + p) = f(x) .$$

Da die Zahl 0 Periode für jede Funktion ist, sind natürlich nur Perioden $p \neq 0$ von echtem Interesse. Die Menge Per(f) der Perioden einer Funktion $f : \mathbb{R} \to \mathbb{C}$ bildet offensichtlich einen $\mathbb{Z}$-Modul, d. h. sind p_1 und p_2 Perioden von f, dann ist auch $m_1 p_1 + m_2 p_2$ für beliebiges $m_1, m_2 \in \mathbb{Z}$ eine Periode von f.

Aufgabe: Zeigen Sie: Gibt es unter den positiven Perioden eine Funktion $f : \mathbb{R} \to \mathbb{C}$ eine kleinste, etwa p, so ist *jede* Periode ω von f ein ganzzahliges Vielfaches von $p : \omega = \ell p,\ \ell \in \mathbb{Z}$.

Aufgrund der folgenden Bemerkung kann man sich bei der Untersuchung periodischer Funktionen immer auf 2π-periodische Funktionen beschränken:

21.8.3 Bemerkung:

Hat die Funktion $f : \mathbb{R} \to \mathbb{C}$ die Periode $p \neq 0$ dann hat die Funktion $g : \mathbb{R} \to \mathbb{C}$ mit $g(x) = f\left(\frac{p}{2\pi} x\right)$ die Periode 2π. Hat umgekehrt $g : \mathbb{R} \to \mathbb{C}$ die Periode 2π, dann hat $f : \mathbb{R} \to \mathbb{C}$ mit $f(y) = g\left(\frac{2\pi}{p} y\right)$ die Periode p.

Wir werden uns daher bei unseren Untersuchungen auf 2π-periodischen Funktionen beschränken.

Für die Kombinationen $\cos kx + i \sin kx = \exp(ikx) (k \in \mathbb{Z},\ x \in \mathbb{R})$ führen wir eine spezielle Bezeichnung ein: $e_k : \mathbb{R} \to \mathbb{C}$ sei definiert durch $e_k(x) := \exp(ikx)$. Man beachte, dass die 'Basisfunktionen' $e_k : \mathbb{R} \to \mathbb{C}$ 2π-periodisch sind, also alle Werte, die sie überhaupt annehmen, schon etwa im Intervall $[0, 2\pi]$, sogar in $[0, 2\pi[$ annehmen, und dass das Bild von $[0, 2\pi]$ unter der Abbildung e_k im Fall $k \neq 0$ die Einheitskreislinie $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ ist. Insbesondere bildet die Funktion: cis$:= e_1$ das Intervall $[0, 2\pi[$ bijektiv auf die Einheitskreislinie S^1 ab.

Man beachte auch die *Rechenregeln*

$$\boxed{\begin{array}{rcll} e_k \cdot e_\ell &=& e_{k+\ell} & (k, \ell \in \mathbb{Z}) \\ (e_k)^r &=& e_{rk} & (k, r \in \mathbb{Z})\,. \end{array}}$$

Eine Funktion der Gestalt $(n \in \mathbb{N}_0)$ $T_n : \mathbb{R} \to \mathbb{C}$

$$x \mapsto T_n(x) := \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos \ell x + b_k \sin kx) \qquad (*)$$

$$(a_k, b_k \in \mathbb{C} \leq k \leq n,\ x \in \mathbb{R})$$

nennen wir ein *trigonometrisches Polynom vom Grad höchstens* n. Aufgrund der Eulerschen Formeln

$$\begin{aligned} \cos kx &= \frac{\exp(ikx) + \exp(-ikx)}{2} = \frac{e_k(x) + e_k(-x)}{2} \\ \sin kx &= \frac{\exp(ikx) + \exp(-ikx)}{2i} = \frac{e_k(x) - e_k(-x)}{2i} \end{aligned}$$

lässt sich $T_n(x)$ auch in der Form

$$T_n(x) = \sum_{k=-n}^{n} c_k e_k(x),\ c_k \in \mathbb{C} \qquad (**)$$

ausdrücken, wobei zwischen den Koeffizienten a_k, b_k und c_k der folgende Zusammenhang besteht:

$$\boxed{c_k = \left\{ \begin{array}{ll} \frac{1}{2}(a_k - ib_k), & k \geq 1 \\ \frac{1}{2}a_0, & k = 0 \\ \frac{1}{2}(a_{-k} + ib_{-k}, & k < 0 \end{array} \right. .}$$

Umgekehrt lassen sich die a_k und b_k durch die c_k ausdrücken:

$$\boxed{\begin{array}{ll} a_k = c_k + c_{-k}, & k \geq 0 \\ b_k = i(c_k - c_{-k}), & k \geq 1 \end{array}}$$

Dass a_k und b_k *reell* sind, ist mit $\overline{c_k} = c_{-k}$ äquivalent.

Frage: Sind die Koeffizienten c_k in der Darstellung für die a_k und b_k durch T_n eindeutig bestimmt? Die Antwort ist "ja" und der Grund liegt in den *Orthogonalitätsrelationen*, welche das System $(e_k)_{k\in\mathbb{Z}}$ erfüllt.

Orthogonalitätsrelationen

Für $k, \ell \in \mathbb{Z}$ gilt

$$\frac{1}{2\pi} \int_0^{2\pi} \exp(ikx) \exp(-i\ell x) dx = \delta_{k\ell} = \left\{ \begin{array}{l} 1, k = \ell \\ 0, k \neq \ell \end{array} \right.$$

Beweis: Nach dem Additionstheorem von $\exp$ gilt

$$\frac{1}{2\pi} \int_0^{2\pi} \exp(ikx) \exp(-i\ell x) dx = \frac{1}{2\pi} \int_0^{2\pi} \exp(i(k - \ell)x) dx$$

Für $k = \ell$ ergibt sich $\frac{1}{2\pi} \cdot 2\pi = 1$.

Für $k \neq \ell$ ist

$$\frac{1}{2\pi}\int\limits_0^{2\pi} \exp(i(k-\ell)x)dx = \frac{1}{2\pi}\,\frac{\exp i(k-\ell)x}{k-\ell}\,\Big|_0^{2\pi} = 0\,,$$

da $e_{k-\ell}(x)$ die Periode 2π hat.

Bemerkung: Man spricht von Orthogonalitätsrelationen, weil die Funktionen im Sinne des Skalarprodukts, das wir im Abschnitt 2 definieren, tatsächlich orthogonal zueinander sind.

Aufgabe: Beweisen Sie Orthogonalitätsrelationen für die Funktionen $\cos kx,\ \sin kx$.

Multipliziert man dabei in der Darstellung

$$T_n(x) = \sum_{k=-n}^{n} c_k e_k(x) \ \text{ mit } \ e_\ell(-x)\ (= e_{-\ell}(x))$$

und integriert, dann folgt

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} T_n(x)e_\ell(-x)dx &= \frac{1}{2\pi}\int_0^{2\pi}\sum_{k=-n}^{n} c_k e_k(x)e_\ell(-x)dx \\
&= \sum_{k=-n}^{n} c_k \frac{1}{2\pi}\int\limits_0^{2\pi} e_k(x)e_\ell(-x)dx = \\
&= \sum_{k=-n}^{n} c_k \frac{1}{2\pi}\int\limits_0^{2\pi} e_{k-\ell}(x)dx = \sum_{k=-n}^{n} c_k\delta_{k\ell} = c_\ell\,,
\end{aligned}$$

also sind die Koeffizienten c_k (wir nennen den Index wieder k) in der Darstellung durch T_n eindeutig bestimmt:

$$\boxed{c_k = \frac{1}{2\pi}\int\limits_0^{2\pi} T_n(x)\exp(-ikx)dx}$$

Betrachtet man in dieser Gleichung den Grenzübergang $n \to \infty$, dann erhält man eine Reihe der Gestalt

$$\sum_{k=-\infty}^{\infty} c_k e_k(x) := \lim_{n\to\infty}\sum_{k=-\infty}^{n} c_k e_k(x)$$

Dabei sind die $c_k \in \mathbb{C}$ beliebig vorgegeben. Eine Reihe vom Typ heißt *trigonometrische* Reihe. Sie heißt konvergent, wenn (der symmetrisch gebildete) Grenzwert rechts existiert.

Beachte: Man verwendet das Symbol $\sum\limits_{k=-\infty} \ldots$ in zweierlei Bedeutung: Einmal ist es ein Synonym für die Folge $\left(\sum\limits_{k=-n}^{n} c_k e_k(x)\right)$ der Partialsummen, zum anderen (im Fall der Konvergenz) bedeutet es den Grenzwert der Folge der Partialsummen (bei punktweiser bzw. eventuell gleichmäßiger Konvergenz).

Wenn z. B. die Reihe $\sum\limits_{k=-\infty} \lvert c_k\rvert$ konvergiert, dann ist die trigonometrische Reihe $\sum\limits_{k=-\infty} c_k e_k(x)$ auf $\mathbb{R}$ gleichmäßig konvergent; durch $f:\mathbb{R}\to\mathbb{C},\ x\mapsto \sum\limits_{k=-\infty} c_k e_k(x)$ wird also eine *stetige* und $2\pi-$ *periodische* Funktion definiert

Es stellen sich nun fast von selbst die folgenden Fragen:

(A) Welche 2π-periodischen Funktionen: $f : \mathbb{R} \to \mathbb{C}$ lassen sich durch trigonometrische Polynome approximieren oder sogar durch trigonometrische Reihen darstellen?

(B) Wie lassen sich die auftretenden Koeffizienten c_k berechnen? Sind sie etwa wie im Fall der trigonometrischen Polynome durch f eindeutig bestimmt? Was wir feststellen können, ist zunächst folgendes:

Feststellung: Ist die Funktion $f : \mathbb{R} \to \mathbb{C}$ lokal integrierbar (es reicht, dass $f|[0, 2\pi]$ eine Regelfunktion ist) und es gelte

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e_k(x),$$

wobei die Reihe rechts auf $\mathbb{R}$ *gleichmäßig* konvergiere, dann gilt notwendig für alle $k \in \mathbb{Z}$:

$$\boxed{c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) \exp(-ikx) dx}$$

Beweis: Wir betrachten ein festes $\ell \in \mathbb{Z}$, multiplizieren mit $e_\ell(-x) = e_{-\ell}(x)$:

$$f(x)e_{-\ell}(x) = \sum_{k=-\infty}^{\infty} c_k e_{k-\ell}(x) \,.$$

Wegen $|e_{k-\ell}(x)| = 1$ für alle $x \in \mathbb{R}$ ist auch die Reihe rechts auf $\mathbb{R}$ gleichmäßig konvergent, so können wir daher gliedweise über $[0, 2\pi]$ integrieren. Das ergibt

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} f(x) \exp(-ilx) dx &= \sum_{k=-\infty}^{\infty} c_k \frac{1}{2\pi} \int_0^{2\pi} e_{k-\ell}(x) dx \\ &= \sum_{k=-\infty}^{\infty} c_k \delta_{k\ell} = c_\ell \end{aligned}$$

Fazit: Die Koeffizienten c_k lassen sich in einfacher Weise gewinnen, nämlich durch Integration der darzustellenden Funktion f, multipliziert mit $e_k(-x) = e_{-k}(x) = e_k(x)$, über das Periodenintervall $[0, 2\pi]$. Diese werden unabhängig voneinander für jedes k berechnet; der Aufwand ist für jedes k gleich groß.

Wir drehen nun gewissermaßen den Spieß um, gehen nicht von trigonometrischen Reihen aus, sondern von Funktionen $f : \mathbb{R} \to \mathbb{C}$.

Wir treffen die Vereinbarungen:

Wir bezeichnen mit V den $\mathbb{C}$-Vektorraum der 2π-periodischen Funktion $f : \mathbb{R} \to \mathbb{C}$, deren Beschränkung auf das kompakte Intervall $[0, 2\pi]$ eine Regelfunktion ist. Für $f \in V$ heißt

$$\boxed{\begin{gathered}\hat{f}(k) := \frac{1}{2\pi} \int_0^{2\pi} f(x) \exp(-ikx) dx,\ k \in \mathbb{Z}, \\ \text{EULER-FOURIER}\textit{sche Formel}\end{gathered}}$$

der k-te *Fourier-Koeffizient* von f und die Reihe

$$S_\infty(f) := \lim_{n\to\infty} \sum_{k=-n}^{n} \hat{f}(k) e_k \text{ *die Fourier-Reihe von* } f \,.$$

Die Partialsumme $S_n(f) := \sum\limits_{k=-n}^{n} \hat{f}(k)e_k (n \in \mathbb{N}_0)$ nennen wir auch das n-te *Fourier-Polynom* von f. Diese kann punktweise oder gleichmäßig konvergieren. Werten wir die Partialsumme bzw. die Reihe an eine Stelle $x \in \mathbb{R}$ aus, so schreiben wir auch

$$S_n(f)(x) \text{ bzw. } S_\infty(f)(x).$$

Jeder Funktion $f \in V$ kann man also die Fourier-Reihe $S_\infty(f)$ zuordnen:

$$f \to S_\infty(f).$$

Fragen: Gegeben sei eine Funktion $f \in V$.

(A) Für welche $x \in \mathbb{R}$ konvergiert die Fourier-Reihe von f punktweise?
Für welche Funktion konvergiert sie sogar gleichmäßig?

(B) Gilt im Falle der Konvergenz der Fourier-Reihe notwendig

$$f(x) = S_\infty(f)(x)\ ?$$

Das heißt, stellt die Fourier-Reihe die Funktion dar?

(C) Gibt es Bedingungen dafür, dass eine Funktion $f \in V$ durch ihre Fourier-Reihe dargestellt wird?

(D) Ist die Fourier-Reihe einer *stetigen* Funktion $f \in V$ in jedem Punkt $x \in \mathbb{R}$ konvergent?

Letzteres ist nicht der Fall. Ein Gegenbeispiel findet man etwa in Zygmund: *Trigonometrical Series*, Warschau 1935 auf Seite 298 oder auch *Königsberger Analysis 1*, 16.4.

(E) Wenn aber die Fourier-Reihe einer stetigen V Funktion $f \in V$ doch an einer Stelle $x \in \mathbb{R}$ konvergiert, gilt dann $S_\infty(f)(x) = f(x)$?

Wir werden wenigstens Teilantworten auf die gestellten Fragen geben. Eines der Ergebnisse wird sein, wenn die Fourier-Reihe einer Funktion $f \in V$ in einer Stelle $x \in \mathbb{R}$ konvergiert, dann konvergiert sie gegen $\tilde{f}(x) := \frac{f(x_+)+f(x_-}{2}$, d. h. gegen das arithmetische Mittel aus dem rechts- bzw. linksseitigen Grenzwert an der Stelle x. Die Grenzwerte

$$\begin{aligned} f(x_+) &:= \lim_{\substack{h\to 0\\ h<0}} f(x+h) - f(x) \text{ und} \\ f(x_-) &:= \lim_{\substack{h\to 0\\ h<0}} f(x-h) - f(x) \end{aligned}$$

existieren stets, da f lokal integrierbar auf $\mathbb{R} \to \mathbb{C}$ 2π-periodisch, $h|[0, 2\pi]$ Regelfunktion, dann ist h lokal integrierbar, also über jedes kompakte Teilintervall von $\mathbb{R}$ integrierbar, speziell gilt für beliebige $a, a' \in \mathbb{R}$

$$\int\limits_0^{2\pi} h(t)dt = \int\limits_a^{a+2\pi} h(t)dt = \int\limits_0^{2\pi} h(a'+t)dt\,,$$

speziell für $a = -\pi$ also etwa

$$\int\limits_0^{2\pi} h(t)dt = \int\limits_{-\pi}^{\pi} h(t)dt = \int\limits_{-\pi}^{\pi} h(a'+t)dt\,.$$

Die Fourierkoeffizienten $\hat{f}(k)$ einer Funktion $f \in V$ lassen sich deshalb auch in der Form

$$\boxed{\hat{f}(k) := \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} f(t)\exp(-ikx)dx,\ k \in \mathbb{Z}}$$

schreiben. Das ergibt für

$$a_k := c_k + c_{-k} = \frac{1}{\pi} \int\limits_{-\pi}^{\pi} f(t) \cos(kt) dt;\ k \geq 0$$

bzw.

$$b_k := i(c_k - c_{-k}) = \frac{1}{\pi} \int\limits_{-\pi}^{\pi} f(t) \sin(kt) dt;\ k \geq 1\,.$$

Ist $f \in V$ eine *gerade* Funktion, d. h. $f(-x) = f(x)$ für alle $x \in \mathbb{R}$, so sind alle $b_k = 0 (k \geq 1)$ und für die a_k gilt $a_k = \frac{2}{\pi} \int\limits_0^{\pi} f(t) \cos(kt) dt,\ k \geq 0$.

Ist $f \in V$ ungerade, d. h. $f(-x) = -f(x)$ für alle $x \in \mathbb{R}$, dann sind alle $a_k = 0\ (k \geq 0)$ und es gilt

$$b_k = \frac{2}{\pi} \int\limits_0^{\pi} f(t) \sin(kt) dt,\ k \geq 1\,.$$

Die Fourier-Reihe reduziert sich also auf eine reine $\cos$-Reihe bzw. reine $\sin$-Reihe.

Bemerkung: Für theoretische Überlegungen ist die Verwendung der komplexen Fourier-Koeffizienten $\hat{f}(k)$ meist zweckmäßiger, für konkrete Rechnungen, insbesondere, falls die Funktion f gerade oder ungerade ist, die Verwendung der a_k bzw. b_k, das Rechnen im 'Komplexen' ist jedoch häufig einfacher.

Die Berechnung der Fourier-Koeffizienten für die Aufstellung der Fourier-Reihe einer Funktion $f \in V$ geschieht unter Verwendung der bekannten Integrationstechniken (z. B. partielle Integration, Substitutionsmethode, Integration der rationalen Funktionen). Wir begnügen uns zunächst mit 2 Beispielen:

(1) Die stückweise lineare 2π-periodische Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(t) = \begin{cases} \frac{1}{2}(\pi - t) & \text{für} & (0 < t < 2\pi \\ 0, & \text{für} & t = 0 \\ f(t + 2\pi), & \text{für alle} & t \in \mathbb{R} \end{cases}$$

besitzt Sprungstellen in den Punkten $t = 2\pi k,\ k \in \mathbb{Z}$ und ist im übrigen ungerade.
Sägezahn:

Nach der Bemerkung gilt $a_k = 0$ für $k \geq 0$ und für die b_k gilt

$$\begin{aligned}
b_k &= \frac{2}{\pi}\int_0^{\pi} \frac{1}{2}(\pi - t)\sin kt\,dt \\
&= \qquad\quad \downarrow \quad \uparrow \\
&= \frac{1}{\pi}\left(\left(-(\pi - t)\frac{\cos kt}{k}\right)\Big|_0^{\pi} - \int_0^{\pi} \frac{\cos kt}{k}\,dt\right. \\
&= \frac{1}{k} - \frac{1}{\pi k^2}\sin kt\Big|_0^{\pi} = \frac{1}{k}
\end{aligned}$$

Also ist die Fourier-Reihe von f gegeben durch

$$S_\infty(f)(t) = \sum_{k=1}^{\infty} \frac{\sin kt}{k}\,.$$

Man kann zeigen, dass für $0 < t < 2\pi$ gilt

$$\sum_{k=1}^{\infty} \frac{\sin kt}{k} = \frac{\pi - t}{2}$$

Für $t = 0$ gilt offensichtlich $S_\infty(f)(0) = 0$, ebenso $S_\infty(f)(2\pi) = 0$. Daher gilt für alle $t \in \mathbb{R}$

$$S_\infty(f)(t) = \sum_{k=1}^{\infty} \frac{\sin kt}{k}\,.$$

Die Fourier-Reihe von f konvergiert für alle $t \in \mathbb{R}$ und stellt die Funktion f dar. Man beachte dabei, dass der Funktionswert 0 bei den Sprungstellen $2\pi k,\ k \in \mathbb{Z}$ das arithmetische Mittel aus dem rechts- und linksseitigen Grenzwert ist:

$$\frac{\lim\limits_{\substack{h\to 0\\ h>0}} f(2\pi k + h) + \lim\limits_{\substack{h\to 0\\ h>0}} f(2\pi k - h)}{2} = \frac{\frac{\pi}{2} + \left(\frac{-\pi}{2}\right)}{2} = 0$$

In einer Umgebung einer der Sprungstellen ist die Konvergenz nicht sonderlich gut. Sprungstellen führen zu Fourier-Koeffizienten der Größenordnung $\frac{1}{k}(k \to \infty)$ und beeinträchtigen damit die Konvergenzqualität für alle t.

An jeder Sprungstelle tritt das "Gibbs-sche Phänomen" auf.

In jedem kompakten Teilintervall $[a, \beta] \subset]0, 2\pi[$ konvergiert die Fourier-Reihe sogar gleichmäßig.

(2) Wir setzen die Einschränkung der Betragsfunktion

$$f : [-\pi, \pi] \to \mathbb{R}, x \mapsto |x|$$

auf ganz $\mathbb{R}$ mit der Periode 2π periodisch fort.

Vorspann:

Behauptung: $\sum\limits_{k=1} \frac{1}{k^2} \leq 2$

Beweis: $s_n := \sum\limits_{k=1}^{n} \frac{1}{k^2}$

$$\begin{aligned}
S_n &= 1 + \frac{1}{2\cdot 2} + \frac{1}{3\cdot 3} + \frac{1}{4\cdot 4} + \cdots + \frac{1}{n\cdot n} \\
&\leq 1 + \frac{1}{2\cdot 1} + \frac{1}{3\cdot 2} + \frac{1}{4\cdot 3} + \cdots + \frac{1}{n\cdot(n-1)} \\
&= 1 + \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\
\Rightarrow S_n &\leq 1 + 1 - \left(\frac{1}{2} - \frac{1}{2}\right) - \left(\frac{1}{3} - \frac{1}{3}\right) - \left(\frac{1}{4} - \frac{1}{4}\right) - \cdots - \left(\frac{1}{n-1} - \frac{1}{n-1}\right) - \frac{1}{n} \\
&= 1 + 1 - \frac{1}{n} \leq 2
\end{aligned}$$

□

f ist offensichtlich gerade, d. h. alle b_k verschwinden und für die a_k gilt

$$\begin{aligned}
a_0 &= \frac{2}{\pi}\int_0^\pi t dt = \frac{2}{\pi}\frac{t^2}{2}\Big|_0^\pi = \pi \text{ und für } k \geq 1 \\
a_k &= \frac{2}{\pi}\int_0^\pi \underset{\downarrow\atop u}{t} \cos \underset{\uparrow\atop dv}{k} \\
&= \frac{2}{\pi}\frac{t \sin kt}{k}\Big|_0^\pi - \frac{2}{\pi}\int_0^\pi \frac{\sin kt}{k}\, dt \\
&= \frac{2}{\pi k^2}(\cos kt)\Big|_0^\pi = \frac{2}{\pi k^2}\left((-1) - 1\right) = \begin{cases} 0, & \text{falls } k \text{ gerade} \\ \frac{-4}{\pi k^2}, & \text{falls } k \text{ ungerade} \end{cases}
\end{aligned}$$

also lautet die Fourier-Reihe von f

$$S_\infty(f) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{\cos(2k-1)x}{(2k-1)^2}.$$

Die Fourier-Reihe $S_\infty(f)$ von $f = |\cdot|$ konvergiert offensichtlich gleichmäßig auf $\mathbb{R}$, stellt also eine stetige Funktion dar, gilt aber

$$\begin{aligned}
S_\infty(f)(x) &= |f| \text{ d. h. also} \\
\frac{\pi}{2} - \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{\cos(2k-1)x}{(2k-1)^2} &= |x| \text{ für alle } x \in \mathbb{R}\,?
\end{aligned}$$

Das ist – wie wir bald zeigen werden – tatsächlich der Fall. Durch eine einfache Umformung erhält man

$$\cos x + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \cdots = \frac{\pi^2}{8} - \frac{\pi}{4}|x|.$$

Setzt man hier $x = 0$, so folgt die bemerkenswerte Gleichung

$$\boxed{\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2}\cdots = \sum_{k=1}^\infty \frac{1}{(2k-1)^2}}$$

Damit kann man auch den Wert der Riemannschen ζ-Funktion $\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}$ an der Stelle $s = 2$ berechnen.

Setzt man nämlich

$$S := \sum_{k=1}^{\infty} \frac{1}{k^2} (= \zeta(2))$$

und schreibt (man rechtfertige die Zerlegung)

$$S = \sum_{\substack{k=1 \\ k \text{ungerade}}}^{\infty} \frac{1}{k^2} + \sum_{\substack{k=1 \\ k \text{ gerade}}}^{\infty} \frac{1}{k^2}$$

so ergibt sich

$$S = \frac{\pi^2}{8} + \frac{1}{2^2} S, \text{ also}$$

$$\boxed{\zeta(2) = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6} \quad (\text{EULER}, 1776)}$$

(3) Die Funktion $f; \mathbb{R} \to \mathbb{C},\ t \mapsto \frac{4-2\cos t}{5-4\cos t}$ ist 2π-periodisch und stetig (Plotten Sie den Graphen!). Zeigen Sie, dass die Fourier-Reihen von f gegeben ist als $S_\infty(f)(x) = \sum_{k=0}^{\infty} \frac{1}{2^k} \cos(kx)$ und zeigen Sie direkt, (mit Hilfe der Summenformel für die geometrische Reihe) dass

$$S_\infty(f)(x) = \sum_{k=0}^{\infty} \frac{1}{2^k} \cos kx = f(x) \text{ für alle } x \in \mathbb{R} \text{ gilt}.$$

Schlussbemerkungen

Da eine 2π-periodische Funktion jeden Wert, den sie annimmt, schon in einem Intervall der Länge 2π annimmt, ist es nahe liegend, Punkte $x, x' \in \mathbb{R}$, deren Differenz $x - x'$ ein ganzzahliges Vielfaches von 2π ist, unter dem Gesichtspunkt der 2π-periodischen Funktion nicht zu unterscheiden, es gilt ja $f(x) = f(x')$, falls $x - x' \in 2\pi\mathbb{Z}$. Wir wollen daher zwei Punkte x und $x' \in \mathbb{R}$ äquivalent (mod2π) nennen, wenn für ihre Differenz $x - x' \in 2\pi\mathbb{Z}$ gilt. Das definiert eine Äquivalenzrelation auf $\mathbb{R}$, die zu x gehörige Äquivalenzklasse sei $[x]$, also $[x] = \{y \in \mathbb{R},\ y \sim x\}$. Eine 2π-periodische Funktion lässt sich dann auch als Funktion auf der Menge $\{[x]\};\ x \in \mathbb{R}\}$ der Äquivalenzklassen auffassen. Diese Menge der Äquivalenzklasse (es ist eine Menge von Mengen!) bezeichnet man hier auch mit $\mathbb{R}/2\pi\mathbb{Z}$. Die Abbildung $p : \mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z},\ x \mapsto [x]$ nennt man auch die *kanonische Projektion* und man hat eine Zerlegung

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{f} & \mathbb{C} \\ p\downarrow & \nearrow \tilde{f} & \\ \mathbb{R}/2\pi\mathbb{Z} & & \end{array}$$

der Abbildung f in die kanonische Projektion p und die Abbildung $\tilde{f} : \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{C}$ eine Funktion, dann wird durch $f(x) := \tilde{f}(x)$.

Ist umgekehrt $\tilde{f} : \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{C}$ eine Funktion, dann wird durch $f(x) := \tilde{f}([x])$ eine 2π-periodische Funktion auf $\mathbb{R}$ definiert.

Wir wissen ferner, dass cis$\tilde{}$: $\mathbb{R}/2\pi\mathbb{Z} \to S^1$, $[t] \mapsto_1 (t) = \exp(it)$ bijektiv ist, daher können wir eine 2π-periodische Funktion $f : \mathbb{R} \to \mathbb{C}$ bzw. eine Funktion $\tilde{f} : \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{C}$ auch als

Funktion f^0 auf der Einheitskreislinie auffassen. $f^0 : S^1 \to \mathbb{C},\ exp\,(it) = e_1(t) \mapsto \tilde{f}([t]) = f(t), (t \in \mathbb{R})$. Jede 2π-periodische Funktion $f : \mathbb{R} \to \mathbb{C}$ und umgekehrt.

$$\begin{array}{ccc}
\mathbb{R} & \xrightarrow{f} & \mathbb{C} \\
p \downarrow & \tilde{f} \nearrow & \uparrow f^0 \\
\mathbb{R}/2\pi\mathbb{Z} & & S^1 \\
\multicolumn{3}{c}{[t] \overset{\cong}{\mapsto} \exp\,(it) = e_1(t)}
\end{array}$$

21.9 Eulersche Summenformel und Bernoullische Polynome (wird nachgetragen)

21.10 Interpolation und elementare Aspekte der numerischen Integration (wird nachgetragen)

21.11 Einige elementar integrierbare Differentialgleichungen (wird nachgetragen)

V: Weiterer Ausbau der Integral- und Differentialrechnung
ist noch in Bearbeitung und enthält u.a.:

- Uneigentliche Integrale, $\Gamma-$Funktion, Stirlingsche Formel
- Eulersche Summenformel und Bernoullische Polynome
- Interpolation und elementare Aspekte der numerischen Integration
- Einige elementar integrierbare Differentialgleichungen
- Einige geometrische Anwendungen: insb. Kurven und ihre Längen

(wird nachgetragen)

VI. Metrische Räume und ihre Topologie

Mit diesem Kapitel beginnen wir mit dem systematischen Studium von Funktionen mehrerer Veränderlicher. Wir wollen Begriffe wie etwa „Konvergenz“, „Stetigkeit“, „Kompaktheit“ auf allgemeinere Situationen übertragen. In den Anwendungen der Mathematik (speziell in der Physik, aber nicht nur dort) hat man es häufig mit Phänomenen zu tun, die nicht nur von einem Parameter abhängen, sondern von mehreren. Dazu einige Beispiele:

(1) Die Bewegung eines Teilchens in $\mathbb{R}^3$ kann man als (stetige) Abbildung eines Intervalls $[a, b] \subset \mathbb{R}$ in den $\mathbb{R}^3$ auffassen:

$$\begin{aligned} \alpha: \ [a,b] &\to \mathbb{R}^3 , \\ t &\mapsto (\alpha_1(t), \alpha_2(t), \alpha_3(t)) \end{aligned}$$

(t etwa ein Zeitparameter)

(2) Betrachtet man etwas allgemeiner die Bewegung von m Massenpunkten in $\mathbb{R}^3$, dann wird die Bewegung beschrieben durch m Abbildungen ($1 \le j \le m$)

$$\begin{aligned} \alpha^{(j)}: \ [a,b] &\to \mathbb{R}^3 , \\ t &\mapsto \left(\alpha_1^{(j)}(t), \alpha_2^{(j)}(t), \alpha_3^{(j)}(t)\right) , \end{aligned}$$

also insgesamt durch eine Abbildung

$$\begin{aligned} \alpha: \ [a,b] &\to (\mathbb{R}^3)^n = \mathbb{R}^{3n} , \\ t &\mapsto \left(\alpha_1^{(1)}(t), \alpha_2^{(1)}(t), \alpha_3^{(1)}(t), \alpha_1^{(2)}(t), \alpha_2^{(2)}(t), \alpha_3^{(2)}(t), \ldots, \alpha_1^{(m)}(t), \alpha_2^{(m)}(t), \alpha_3^{(m)}(t)\right) \end{aligned}$$

(3) Befindet sich im Nullpunkt $0 = (0,0,0) \in \mathbb{R}^3$ eine Masse M, so übt sie nach dem Gravitationsgesetz auf eine im Punkt $x = (x_1, x_2, x_3) \in \mathbb{R}^3 - \{0\}$ befindliche Masse m eine Kraft aus:

$$F(x) = \frac{C}{r^2} \cdot \frac{x}{r} \qquad r := r(x) := \sqrt{x_1^2 + x_2^2 + x_3^2}$$

(dabei ist (auch im Folgenden) $\frac{x}{r}$ als $\frac{1}{r} \cdot x$ aufzufassen und C ist eine geeignete Konstante).

Durch

$$\begin{aligned} F: \ \mathbb{R}^3 - \{0\} &\to \mathbb{R}^3 , \\ (x_1, x_2, x_3) &\mapsto F(x) \end{aligned}$$

wird ein *Vektorfeld* (Kraftfeld) definiert.

(4) Nach dem Boyle-Mariotteschen Gesetzt hängt das Volumen eines idealen Gases vom Druck p und der Temperatur T ab:

$$V = V(p, T) := c\frac{T}{p}$$

mit einer Konstanten c.

(5) Für den Drehimpuls L (im Bezug auf $0 = (0,0,0)$) eines Teilchens mit der Masse m, das sich in $\mathbb{R}^3$ bewegt, gilt

$$L(t) = x(t) \times p(t),$$

dabei bezeichnet $\times$ das Vektorprodukt in $\mathbb{R}^3$, $x(t)$ den Ort und $p(t)$ den Impuls zur Zeit $t \in [a,b]$.

Diese Liste von Beispielen lässt sich beliebig fortsetzen. Der Schritt von der eindimensionalen Analysis zur mehrdimensionalen Analysis bereitet häufig Schwierigkeiten u.a. weil

- die Theorie nicht mehr vertraut ist,
- die Rechenverfahren komplizierter werden,
- die begrifflichen Anstrengungen erhöht werden müssen.

In der Analysis 1 haben wir einen kleinen Schritt in diese Richtung beim Übergang von $\mathbb{R}$ zu $\mathbb{R}^2 = \mathbb{C}$ bereits getan. Wir werden versuchen, einen Mittelweg zwischen geometrischer Anschaulichkeit und Abstraktion zu finden. Wir entwickeln die Begriffe im Kontext der „metrischen Räumen“, der allgemeine Begriff des „topologischen Raumes“ wird lediglich gestreift. Unsere Anschauung ist häufig nur ein unzureichendes Hilfsmittel zum Verständnis komplizierter Phänomene. Der Einsatz eines Computeralgebrasystems (CAS) zur Visualisierung vieler Phänomene ist dabei ausgesprochen nützlich.

22 Der Begriff des metrischen Raums, Beispiele, topologische Grundbegriffe

Wir wiederholen zunächst den Begriff des *metrischen Raumes* (vgl. Def 5.9), betrachten weitere Beispiele und führen *topologische Grundbegriffe* (Umgebung, offen, abgeschlossen, Randpunkt, Häufungspunkt, innerer Punk etc.) ein.

22.1 Definition (Metrik, metrischer Raum)

Sei $X \neq \emptyset$. Eine Abbildung $d : X \times X \to \mathbb{R}$ heißt *Metrik auf* X, wenn für alle $x, y, z \in X$ gilt

$$\begin{array}{llcll} (M_1): & d(x,y) & = & 0 \iff x = y & \\ (M_2): & d(x,y) & = & d(y,x) & \text{(Symmetrie)} \\ (M_3): & d(x,z) & \leq & d(x,y) + d(y,z) & \text{(„Dreiecksungleichung“)} \end{array}$$

Das Paar (X,d) heißt *metrischer Raum*.
Die Elemente $x \in X$ nennen wir auch *Punkte* des metrischen Raums. $d(x,y)$ nennen wir auch den *Abstand* von x und y.

22.1.1 Bemerkungen und Beispiele

(a) Häufig schreiben wir nur X, wenn klar ist, welche Metrik betrachtet wird. Da es aber auf einer Menge X verschiedene Metriken geben kann, wie die folgende Beispiele zeigen, ist manchmal die präzise Schreibweise (X, d) nötig.

(b) Wie früher (vgl. §5.9) gezeigt gilt $d(x, y) \geq 0$ für alle $x, y \in X$.

(c) Gilt nur $(M_1') : \; x = y \Longrightarrow d(x, y) = 0$ und (M_2) und (M_3) so spricht man auch von einer *Pseudo-Metrik* auf X.

(d) Für alle $x, y, z \in X$ gilt

$$|d(x, z) - d(x, y)| \leq d(y, z)$$

(Jede Dreiecksseite ist mindestens so groß wie die Differenz der beiden anderen.)

Für $n \in \mathbb{N}$ und $X = \mathbb{K}^n$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) und $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ bzw. $y = (y_1, \ldots, y_n) \in \mathbb{K}^n$ wird durch

$$d_2(x, y) := \sqrt{\sum_{\nu=1}^{n} |x_\nu - y_\nu|^2}$$

die *euklidische Metrik* auf $\mathbb{K}^n$ definiert.
Die euklidische Metrik auf $\mathbb{K}^n$ nennen wir auch die *natürliche Metrik* auf $\mathbb{K}^n$. Sie entspricht unserem geometrischen Abstandsbegriff.

Im Fall $n = 1$ stimmt sie mit dem Betrag auf $\mathbb{R}$ bzw. $\mathbb{C} = \mathbb{R}^2$ überein.

Andere Metriken auf $\mathbb{K}^n$ sind

$$d_1(x, y) := \sum_{\nu=1}^{n} |x_\nu - y_\nu| \qquad \text{(„Summenmetrik“)}$$

und

$$d_\infty(x, y) := \max_{1 \leq \nu \leq n} \{|x_\nu - y_\nu|\} \qquad \text{(„Maximumsmetrik“)}$$

Für jedes $p \in \mathbb{N}$ liefert

$$d_p(x, y) := \left(\sum_{\nu=1}^{n} |x_\nu - y_\nu|^p \right)^{\frac{1}{p}}$$

eine Metrik auf $\mathbb{K}^n$, die für $p = 2$ und $p = 1$ bzw. für $p \to \infty$ die Metriken $d_2, \; d_1$ bzw. d_∞ liefern.

(e) Ist X eine Menge mit mindestens zwei Elementen, dann wird durch

$$d(x, y) = \begin{cases} 1, & \text{falls } x \neq y; \\ 0, & \text{falls } x = y \end{cases}$$

eine Metrik auf X definiert, die sog. *diskrete Metrik*.
Eine Menge X versehen mit der diskreten Metrik hat ungewohnte Eigenschaften. Man vergleiche hierzu das Aufgabenblatt ??? zur Analysis 2.

(f) Die Metriken aus Beispiel (d) erhält man aus *Normen*. Dazu sei an die Definition 5.8 (Norm und normierter Raum) erinnert:

Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$\begin{aligned} \| \; \| : V &\to \mathbb{R} \\ v &\mapsto \|v\| \end{aligned}$$

heißt *Norm* auf V, falls für alle $v, w \in V$ und alle $\lambda \in \mathbb{K}$ gilt

$(N_1):\ \|v\| = 0 \iff v = 0$
$(N_2):\ \|\lambda v\| = |\lambda|\,\|v\|$
$(N_3):\ \|v + w\| \leq \|v\| + \|w\|$ (Dreiecksungleichung)

Die (nicht negative!) reelle Zahl $\|v\|$ heißt *Norm* (auch Betrag, Länge) des Vektors v. Ein Vektor v mit $\|v\| = 1$ heißt *Einheitsvektor*.
Ein Vektorraum V zusammen mit einer Norm $\| \; \|$ auf V heißt *normierter Raum*.
Hierfür schreibt man auch $(V, \| \; \|)$, insbesondere wenn man die gerade betrachtete Norm besonders betonen will (das ist manchmal nötig, da es auf einem Vektorraum V viele Normen geben kann).

Wir wissen auch, dass jeder normierte Raum $(V, \| \; \|)$ durch die Definition $(v, w \in V)$

$$d(v, w) := \|v - w\|$$

zu einem metrischen Raum wird (vgl. §5.10). Insbesondere erhält man die Metriken d_1, d_2, d_∞ und allgemeine d_p aus (d) aus den $p-$Normen

$$\|x\|_p = \left(\sum_{\nu=1}^{n} |x_\nu|^p \right)^{\frac{1}{p}}.$$

Aus Normen abgeleitete Metriken haben den Vorteil, dass sie *translationsinvariant* sind.
Für alle $u, v, w \in V$ gilt:

$$d(v + u, w + u) = \|(v + u) - (w + u)\| = \|v - w\| = d(v, w)$$

Weitere wichtige Beispiele:

(g) Der Hilbert'sche Folgenraum l^2:

$$l^2 := l^2(\mathbb{C}) = \left\{ (x_\nu)_{\nu \geq 1},\ x_\nu \in \mathbb{C};\ \sum_{\nu=1}^{\infty} |x_\nu|^2 < \infty \right\}$$

mit der l^2-Norm

$$\|(x_\nu)\|_2 = \sqrt{\sum_{\nu=1}^{\infty} |x_\nu|^2}.$$

Der Raum l^2 ist ein wichtiges Beispiel, wo die Norm aus einem Skalarprodukt abgeleitet ist. Wir kommen auf dieses Beispiel ausführlich zurück.

(h) Ist M irgendeine nicht leere Menge und

$$B(M) = \{f : M \to \mathbb{C},\ |f(x)| \leq C \text{ für alle } x \in M,\ C \in \mathbb{R}_+ \text{ geeignet}\}$$

die Menge aller *beschränkten* Funktionen auf M, so wird $B(M)$ zu einem normierten Raum, wenn man

$$\|f\| := \|f\|_M = \|f\|_\infty := \sup\{|f(x)|;\ x \in M\}$$

definiert. (Supremums-Norm)
Die zugehörige Metrik liefert als Konvergenzbegriff den Begriff der *gleichmäßigen Konvergenz*.

(i) Ist (X,d) ein metrischer Raum und $X_0 \subset X$ eine nicht leere Teilmenge, so liefert die Einschränkung der Metrik $d : X \times X \to \mathbb{R}$ auf die Teilmenge X_0 eine Metrik $d_0 : X_0 \times X_0 \to \mathbb{R}$:

$$d_0(x,y) := d(x,y) \text{ für } x,y \in X_0.$$

d_0 heißt die auf X_0 (von der Metrik d auf X) *induzierte* Metrik.

Ist X_0 eine Teilmenge des metrischen Raums X, so werden wir in der Regel X_0 mit der induzierten Metrik versehen. Ist z.B. $X = \mathbb{R}$ mit der üblichen Metrik und $X_0 = \mathbb{Z}$, dann ist die auf X_0 induzierte Metrik die *diskrete* Metrik (vgl. (e)).

Wir kommen auf die induzierte Metrik häufig zurück.

(j) Die Metriken $d_1,\ d_2$ und d_∞ auf $\mathbb{K}^n$ besitzen Analoga für Produkte von metrischen Räumen:

Sind $(X_1,d_1),(X_2,d_2),\ldots,(X_n,d_n)$ metrischen Räume und ist

$$X := X_1 \times X_2 \times \ldots \times X_n,$$

dann werden durch

$$\begin{aligned} d_1(x,y) &:= \sum_{j=1}^{n} d_j(x_j,y_j) \text{ bzw.} \\ d_2(x,y) &:= \sqrt{\sum_{j=1}^{n} d_j(x_j,y_j)^2} \text{ bzw.} \\ d_\infty(x,y) &:= \max_{1\le j\le n} \{d_j(x_j,y_j)\} \end{aligned}$$

jeweils Metriken definiert ($x = (x_1,\ldots,x_n) \in X,\ y = (y_1,\ldots,y_n) \in X$).

Das nachzuweisen, sei als Übungsaufgabe gestellt.

(k) Auf einem *unendlichen* Produkt $X := \prod_{n=0}^{\infty} X_n$ von metrischen Räumen (X_n,d_n) erhält man durch

$$d(x,y) = \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} \frac{d_n(x_n,y_n)}{1+d_n(x_n,y_n)} \quad (x = (x_n) \in X,\ y = (y_n) \in X)$$

eine Metrik.

Warum kann man das Prinzip aus (i) nicht direkt übertragen?

(l) Wir schließen die Liste der Beispiele mit einem unserer geometrischen Anschauung ferner liegenden Beispiel, das jedoch für die Zahlentheorie von grundlegender Bedeutung ist.

Sei $X = \mathbb{Z},\ p \in \mathbb{N}$ eine feste Primzahl. Für $x \in \mathbb{Z},\ x \neq 0$, ist $|x| \in \mathbb{N}$ und nach dem sog. „Hauptsatz der elementaren Zahlentheorie“ besitzt $n := |x|$ eine Zerlegung in Primzahlpotenzen

$$\boxed{n = \prod_{p\in\mathbb{P}} p^{\nu_p(|x|)}}$$

dabei sei $\mathbb{P} := \{2,3,5,7,\ldots\}$ die Menge aller Primzahlen.

Das obige Produkt ist nur formal ein unendliches Produkt, denn für die Primzahlen p, die nicht in der Primzahlzerlegung von n vorkommen, setzt man einfach $\nu_p(n) = 0$.

Die durch

$$d_p(x,x) = 0 \text{ und } d_p(x,y) = p^{-\nu_p(|x-y|)}$$

für $x \neq y$ definierte Metrik auf $\mathbb{Z}$ heißt *p-adische Metrik* auf $\mathbb{Z}$. Sie ist eine sog. *Ultra-Metrik*, d.h. es gilt

$$d_p(x,z) \leq \max\{d_p(x,y), d_p(x,z)\} \text{ für } x,y,z \in \mathbb{Z}.$$

Diese Metrik lässt sich in naheliegende Weise auf den Körper $\mathbb{Q}$ fortsetzen.

Die sog. „Komplettierung“ von $\mathbb{Q}$ bezüglich dieser Metrik ergibt den Körper $\mathbb{Q}_p$ der $p-$adischen Zahlen. Für jede Primzahl p erhält man so Körper $\mathbb{Q}_p$, welche den Körper $\mathbb{Q}$ der rationalen Zahlen als Unterkörper enthalten. Interessenten an dieser Tatsache seien verwiesen etwa auf:
G.Frei: Elementare Zahlentheorie, Vieweg Studium, Kap III §4.

Wir fahren nun mit der allgemeinen Theorie fort und kommen zu einem fundamentalen Begriff, nämlich dem der *Kugel* in einem metrischen Raum.

22.2 Definition (offene und abgeschlossene Kugel, $r-$Umgebung, Sphäre)

Sei (X,d) ein metrischer Raum, $a \in X$ und $r \in \mathbb{R},\ r \geq 0$, dann heißt

(a) $U_r(a) := \{x \in X;\ d(x,a) < r\}$ die *offene Kugel* mit *Mittelpunkt* a und *Radius* r. Statt *offene Kugel* sagt man auch kurz *$r-$Umgebung* von a.

(b) $\overline{U_r}(a) := \{x \in X;\ d(x,a) \leq r\}$ die *abgeschlossene Kugel* mit Mittelpunkt a und Radius r.

(c) $S_r(a) := \{x \in X;\ d(x,a) = r\}$ die *Sphäre* mit *Mittelpunkt* a und Radius r.

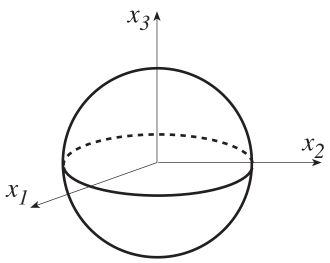
22.2.1 Beispiele und Bemerkungen

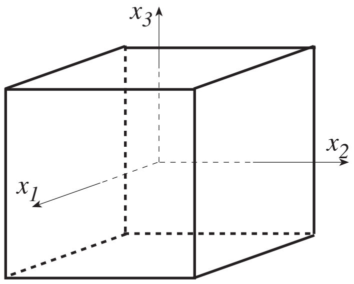
(a) Ist $X = \mathbb{R}^3$ versehen mit der euklidischen Metrik, $a \in \mathbb{R}^3$ und $r > 0$, dann hat $U_r(a)$ tatsächlich die geometrische Gestalt einer Kugel ($\overline{U_r}(a)$ ist eine „Vollkugel“), weshalb man auch im allgemeinen Fall von einer „Kugel“ spricht.

Wählt man in $\mathbb{R}^3$ etwa die Maximumsmetrik, so ist die entsprechende „Kugel“ ein „Würfel“, nämlich die Menge

$$\{x \in \mathbb{R}^3;\ \max\{|x_1 - a_1|, |x_2 - a_2|, |x_3 - a_3| < r\},$$

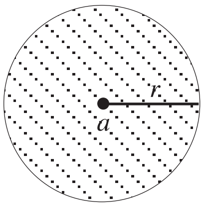
dabei $x = (x_1, x_2, x_3),\ a = (a_1, a_2, a_3) \in \mathbb{R}^3$.

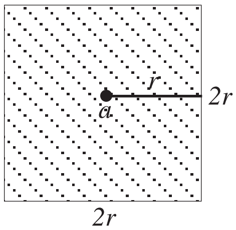




$$= [a_1 - r, a_1 + r[\times]a_2 - r, a_2 + r[\times]a_3 - r, a_3 + r[$$

Im Fall $n = 2$ erhält man bezüglich der euklidischen Metrik eine Kreisscheibe bzw. ein Quadrat bei der Maximumsmetrik.

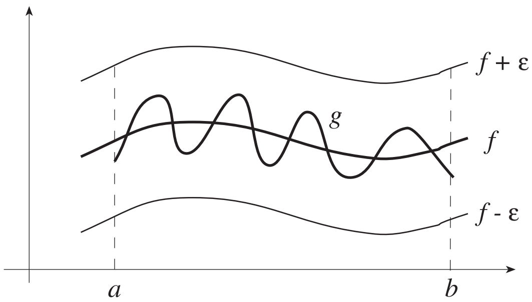




Im Fall $n = 1$ fallen die Metriken zusammen und es ist $U_r(a) =]a - r, a + r[$ ein „offenes Intervall".

Ist $X = \mathcal{C}([a, b]) = \{f : [a, b] \to \mathbb{R},\ f$ stetig$\}$ und $d(f, g) = \|f - g\|_\infty$ die Supremums-Metrik auf X, dann ist für $f \in X$

$$U_r(f) = \{g \in X;\ \|g - f\|_\infty < \varepsilon\}$$



die Menge aller Funktionen $g \in X$, die im „ε-Schlauch" um f liegen.

Versieht man allgemein ein Produkt von metrischen Räumen

$$(X_1, d_1), (X_2, d_2), \dots, (X_n, d_n)$$

mit der Maximumsmetrik d_∞:

$$d_\infty(x, y) = \max_{1 \le j \le n} \{d_j(x_j, y_j)\}, \text{ dabei seien}$$

$$x = (x_1, \dots, x_n) \in X_1 \times \dots \times X_n$$
$$y = (y_1, \dots, y_n) \in X_1 \times \dots \times X_n,$$

dann zerfällt die $r-$Kugel bezüglich d_∞ in ein Produkt von $r-$Kugeln:

$$\begin{aligned} U_r(a) &:= \{x \in X;\ d_\infty(x, a) < r\} \\ &= \{x_1 \in X;\ d_1(x_1, a_1) < r\} \times \dots \times \{x_n \in X_n;\ d_n(x_n, a_n) < r\} \\ &= U_r(a_1) \times U_r(a_2) \times \dots \times U_r(a_n). \end{aligned}$$

(b) Man beachte: Im Fall $r = 0$ ist

$$U_0(a) = \emptyset,\ \overline{U_0}(a) = \{a\} \text{ und } S_0(a) = \{a\}$$

(c) Sind $a, b \in \mathbb{R},\ a \leq b$, dann ist das „offene Intervall"

$$]a, b[= U_r(x_0) \ \text{ mit } \ r := \frac{b-a}{2} \ \text{ und } \ x_0 = \frac{a+b}{2}.$$

Wir werden jetzt den Begriff „offen" spezifizieren und dann sehen, dass die bisher als „offen" bezeichneten Mengen tatsächlich offen im Sinne dieser Definition sind.

(d) Statt der Bezeichnung $U_r(a)$ bzw. $\overline{U_r}(a)$ für Kugeln findet man in der Literatur auch häufig $B_r(a)$ bzw. $B(a, r)$ und $\overline{B_r}(a)$ bzw. $\bar{B}(a, r)$ (von Englisch „Ball" oder französisch „Boule")

(e) Welche „merkwürdigen", von der geometrischen Vorstellung nicht zu erwartenden Eigenschaften Kugeln bzw. Sphären in einem metrischen Raum haben können, zeigt das Beispiel eines diskreten metrischen Raumes (X, d) (vgl. 21.1(e)):

Hier gilt

(α) für die offenen Kugeln und $x \in X$:
Ist $0 < r \leq 1$, so ist $U_r(x) = \{x\}$.
Ist $r > 1$, so ist $U_r(x) = X$.

(β) für die abgeschlossenen Kugeln:
Ist $0 < r < 1$, so ist $\overline{U_r}(x) = \{x\}$.
Ist $r \geq 1$, so ist $\overline{U_r}(x) = X$.

(γ) für die Sphären:
Ist $0 < r < 1$, so ist $S_r(x) = \emptyset$.
Ist $r = 1$, so ist $S_r(x) = X - \{x\}$.
Ist $r > 1$, so ist $S_r(x) = \emptyset$.

Für alle $y \in U_r(x)$ gilt $U_r(y) = U_r(x)$.

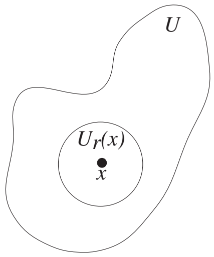
Ein Beweis findet man in der Lösung der Übungsaufgabe ??? aus Analysis 2. Es ist im Übrigen elementar.

Zuvor führen wir folgende Sprechweise ein:

22.3 Definition (Umgebung, offen, abgeschlossen)

Sei (X, d) ein metrischer Raum.

Eine Teilmenge $U \subset X$ heißt



(a) Umgebung von $x \in X$, wenn es ein $r \in \mathbb{R},\ r > 0$ gibt mit

$$U_r(x) \subset U$$

Beachte: $x \in U_r(x) \subset U \subset X$.

(b) offen (in X), wenn U Umgebung jedes Punktes $x \in U$ ist, d.h. wenn es zu jedem $x \in U$ ein $\varepsilon_x > 0$ gibt mit

$$U_{\varepsilon_x}(x) \subset U.$$

(c) Eine Teilmenge $A \subset X$ heißt *abgeschlossen* (in X), wenn ihr Komplement $X - A$ offen in X ist.

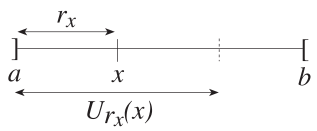
22.3.1 Beispiele und Bemerkungen

(a) Sind $a, b \in \mathbb{R},\ a < b$, dann ist das „offene Intervall"

$$]a, b[= \{x \in \mathbb{R};\ a < x < b\}$$

offen, denn ist $x \in]a, b[$, so gilt

$$U_{r_x}(x) \subset]a, b[\text{ mit } r_x := \min\{|a - x|, |b - x|\}$$



Auch die „uneigentlichen Intervalle"

$$]a, \infty[\text{ bzw. }] - \infty, a[\quad (a \in \mathbb{R})$$

sind offen, ebenso das Intervall $] - \infty, \infty[= \mathbb{R}$.

Das Intervall $[a, b]$ $(a \leq b)$ ist nicht offen, auch nicht das Intervall $[a, b[$, denn für kein $r > 0$ liegt $U_r(a)$ ganz in $[a, b]$ oder $[a, b[$.

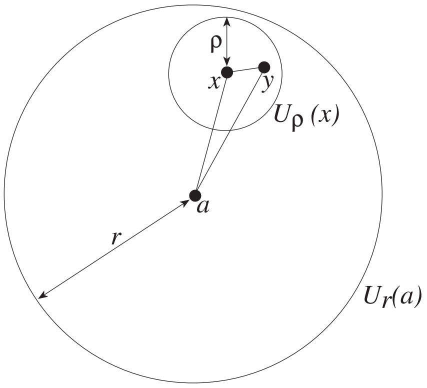
Das „abgeschlossene Intervall" $[a, b]$ ist jedoch abgeschlossen im Sinne unserer Definition, wie man direkt zeigen kann oder aus dem nächsten Beispiel folgern.

(b) Ist (X, d) ein metrischer Raum, $a \in X,\ r \geq 0$, dann ist die offene Kugel $U_r(a)$ offen und die abgeschlossene Kugel $\overline{U_r}(a)$ abgeschlossen.

Für $r = 0$ ist $U_0(a) = \emptyset$ und die leere Menge nach Konvention offen.

Sei $r > 0$ und ist $x \in U_r(a)$ und $\rho := r - d(x, a) > 0$, dann ist $U_\rho(x) \subset U_r(a)$, denn für $y \in U_\rho(x)$ gilt nach der Dreiecksungleichung

$$\begin{aligned} d(y, a) &\leq d(y, x) + d(x, a) \\ &< \rho + d(x, a) \\ &= r - d(x, a) + d(x, a) \\ &= r. \end{aligned}$$



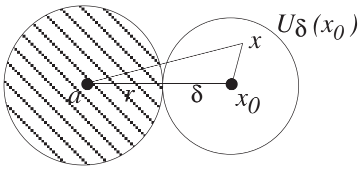
Sei nun $r \geq 0$ und $M := \overline{U_r}(a)$ die abgeschlossene Kugel. Wir haben zu zeigen, dass $U := X - M$ offen ist. Sei dann $x_0 \in U$, dann ist

$$d(a, x_0) := r + \delta \text{ mit } \delta > 0.$$

Für $x \in U_\delta(x_0)$ ist dann

$$d(a, x) \geq d(a, x_0) - d(x_0, x) > r + \delta - \delta = r,$$

d.h. auch x liegt in U, d.h. das Komplement von $M = \overline{U_r}(a)$ ist offen, d.h. $\overline{U_r}(a)$ ist abgeschlossen.



Beachte im Fall $r = 0$ ist

$$U_0(a) = \{a\}.$$

Im Fall reeller Zahlen a, b mit $a \leq b$ ist das „abgeschlossene Intervall"

$$[a, b] = \{x \in \mathbb{R},\ a \leq x \leq b\}$$

ist abgeschlossen im Sinne unserer neuen Definition, denn es ist

$$[a, b] = \overline{U_r}(x_0) \text{ mit } x_0 = \frac{a+b}{2} \text{ und } r = \frac{b-a}{2}.$$

Auch die Intervalle $[a, \infty[$ und $] - \infty, a]$ sind abgeschlossen.

(c) Die Beispiele $[a < b]$ $[a, b[$ bzw. $]a, b]$ zeigen, dass eine Teilmenge von $\mathbb{R}$ (allgemeiner ein metrischer Raum) weder offen noch abgeschlossen zu sein braucht. Auch bedeutet die Tatsache, dass eine Teilmenge $M \subset X$ *nicht offen* ist, *nicht*, dass sie automatisch abgeschlossen ist. Die Begriffe „offen" und abgeschlossen sind über die Komplementbildung miteinander verbunden.

In $\mathbb{R}$ sind die leere Menge $\emptyset$ und der ganze Raum jedoch die einzigen Teilmengen, die zugleich abgeschlossen *und* offen (=„abgeschloffen") sind.

In einem Raum mit der diskreten Metrik ist jedoch jede offene Kugel auch abgeschlossen und jede abgeschlossene Kugel auch offen. Das ist aber nicht der „Normalfall".

Auf diese Phänomene kommen wir bei der Behandlung der „Zusammenhangsbegriffe" zurück.

Die offenen Mengen in einem metrischen Raum haben bemerkenswerte Eigenschaften, die auf elementaren Eigenschaften von offenen Kugeln beruhen. Bevor wir diese allgemeinen Eigenschaften formulieren halten wir eine wichtige Eigenschaft offener Kugeln bzw. Umgebungen fest:

22.4 Feststellung

Ist (X, d) ein metrischer Raum und sind $x, y \in X$ zwei verschiedene Punkte. Dann gibt es stets Umgebungen $U(x)$ von x und $U(y)$ von y mit

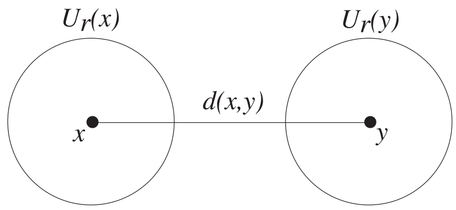
$$U(x) \cap U(y) = \emptyset.$$

Kurz: Verschiedene Punkte lassen sich stets durch Umgebungen (sogar durch $r-$Umgebungen) trennen.

> Auf dieser sog. *Hausdorff-Eigenschaft* der Umgebungen in einem metrischen Raum beruht die Tatsache, dass eine konvergente Folge einen eindeutig bestimmten Grenzwert hat.

Der **Beweis** der Feststellung ist elementar:
Sind $x, y \in X$ zwei verschiedene Punkte, dann ist $d(x, y) > 0$ und setzt man etwa $r := \frac{d(x,y)}{3}$, dann gilt für $U(x) := U_r(x)$ und $U(y) := U_r(y)$ offensichtlich

$$U_r(x) \cap U_r(y) = \emptyset$$



„Ungläubige“ mögen sich davon überzeugen, dass auch schon $U_\rho(x) \cap U_\rho(y) = \emptyset$ gilt mit

$$\rho := \frac{d(x, y)}{2}$$

22.5 Theorem (Grundeigenschaften offener Mengen)

Ist (X, d) ein metrischer Raum und

$$\mathcal{F}_d := \{O;\ O \subset X,\ O \text{ offen } \}$$

das System der offenen Mengen in X, dann gilt:

(0) $\emptyset,\ x \in \mathcal{F}_d$, d.h. die leere Menge und der ganze Raum sind offen.

(1) Ist Λ irgendeine Indexmenge und $(O_\lambda)_{\lambda\in\Lambda}$ ein System von Teilmengen mit $O_\lambda \in \mathcal{F}_d$, dann gilt auch

$$\bigcup_{\lambda\in\Lambda} O_\lambda \in \mathcal{F}_d,$$

d.h. eine beliebige Vereinigung offener Mengen ist wieder offen.

(2) Sind $O_1, O_2, \ldots, O_n \in \mathcal{F}_d$ $(n \in \mathbb{N})$, so ist auch der Durchschnitt

$$O_1 \cap O_2 \cap O_3 \cap \ldots \cap O_n \in \mathcal{F}_d,$$

d.h. *endliche* Durchschnitte offener Mengen sind offen.

Beweis (0): Da die leere Menge überhaupt kein Element enthält, ist die Aussage:

Für jedes $x \in \emptyset$ und jedes $r > 0$ ist $U_r(x) \in \emptyset$

richtig.

Für jedes $x \in X$ und jedes $r > 0$ ist natürlich $U_r(x) \subset X$, d.h. X ist offen.

□

Beweis (1): Ist $x \in \bigcup\limits_{\lambda\in\Lambda} O_\lambda$, dann gibt es ein $\lambda_0 \in \Lambda$ mit $x \in O_{\lambda_0}$. Da O_{λ_0} offen ist, gibt es ein $r > 0$ mit $U_r(x) \subset O_{\lambda_0}$ und damit $U_r(x) \subset O_{\lambda_0} \subset \bigcup\limits_{\lambda\in\Lambda} O_\lambda$, d.h. $\bigcup\limits_{\lambda\in\Lambda} O_\lambda$ ist offen.

□

Beweis (2): Sei $O := O_1 \cap O_2 \cap \ldots \cap O_n$ und $x \in O$.

Für jedes $j \in \{1, 2, \ldots, n\}$ gibt es dann ein $r_j > 0$ mit $U_{r_j}(x) \subset O_j$.

Setzt man $r = \min\{r_1, \ldots, r_n\}$, dann ist $U_r(x) \subset O := O_1 \cap O_2 \cap \ldots \cap O_n$, also ist O offen.

□

Unter Verwendung der sog. de Morganschen Regeln ergeben sich aus §21.5 folgende Grundeigenschaften abgeschlossener Mengen in einem metrischen Raum.

22.5.1 Theorem (Grundeigenschaften abgeschlossener Mengen)

Ist (X, d) ein metrischer Raum, dann gilt

(0') $\emptyset$ und X sind abgeschlossen.

(1') Ist Λ eine beliebige Indexmenge und $A_\lambda \subset X$ abgeschlossen für alle $\lambda \in \Lambda$, dann ist auch $\bigcap\limits_{\lambda\in\Lambda} A_\lambda$ abgeschlossen, d.h. beliebige Durchschnitte abgeschlossener Mengen sind abgeschlossen.

(2') Sind $A_1, \ldots, A_n \subset X$ abgeschlossen ($n \in \mathbb{N}$), dann ist auch $A_1 \cup A_2 \cup \ldots \cup A_n$ abgeschlossen, d.h. *endliche* Vereinigungen abgeschlossener Mengen sind wieder abgeschlossen.

Zu den de Morganschen Regeln:
Ist X eine nicht leere Menge und $(X_\lambda)_{\lambda\in\Lambda}$ eine nicht leeres System von Teilmengen $X_\lambda \subset X$, dann gelten die

22.6 De Morganschen Regeln

$$X - \left(\bigcup_{\lambda\in\Lambda} X_\lambda\right) = \bigcap_{\lambda\in\Lambda} (X - X_\lambda) \quad \text{und} \quad X - \left(\bigcap_{\lambda\in\Lambda} X_\lambda\right) = \bigcup_{\lambda\in\Lambda} (X - X_\lambda)$$

Der **Beweis** ist elementar: Es ist

$$\begin{aligned} x \in X - \left(\bigcup_{\lambda\in\Lambda} X_\lambda\right) &\iff \text{Es gilt nicht (für alle } \lambda \in \Lambda \text{ ist } x \in X_\lambda) \\ &\iff \text{Es gibt ein } \lambda \in \Lambda \text{ mit } x \notin X_\lambda \\ &\iff \text{Es gibt ein } \lambda \in \Lambda \text{ mit } x \in X - X_\lambda \\ &\iff x \in \bigcup_{\lambda\in\Lambda} (X - X_\lambda)\,. \end{aligned}$$

Die zweite de Morgansche Regel beweist man analog, man kann sie auf die erste zurückführen, wenn man die Relation
$$X - (X - A) = A$$
für jede Teilmenge A von X beachtet.
Da $A \subset X$ genau dann abgeschlossen ist, wenn $X - A$ offen ist, ergeben sich die Grundeigenschaften abgeschlossener Mengen ist denen der offenen Mengen unter Verwendung der de Morganschen Regeln.

Man beachte jedoch:

22.7 Feststellungen

(a) Beliebige Durchschnitte offener Mengen brauchen nicht offen zu sein: Nimmt man für $k \in \mathbb{N}$ und $a \in X$
$$U_k := U_{\frac{1}{k}}(a)$$
die offenen Kugeln vom Radius $\frac{1}{k}$, dann ist $\bigcap\limits_{k=1}^{\infty} U_k = \{a\}$. Eine einpunktige Menge in einem metrischen Raum ist immer abgeschlossen (kann sie manchmal auch offen sein?).

(b) Beliebige Vereinigungen abgeschlossener Teilmengen brauchen nicht abgeschlossen zu sein. nimmt man z.B. $X = \mathbb{R}$ mit der natürlichen Metrik und betrachtet die abgeschlossenen Intervalle
$$A_k := \left[-1 + \frac{1}{k},\, 1 - \frac{1}{k}\right], \quad k \in \mathbb{N},$$

dann ist $\bigcup\limits_{k=1}^{\infty} A_k =]-1,1[$.

Die Grundeigenschaften offener bzw. abgeschlossener Teilmengen eines metrischen Raumes (X,d) haben sich also so fundamental erwiesen, dass man sie in der mathematischen Disziplin *Topologie* als Axiome verwendet. Man gelangt so zu einem abstrakten Typ von Räumen, den sog. *topologischen Räumen*. Obwohl wir uns nicht tiefer mit diesem Begriff beschäftigen werden, geben wir die Definition an, da zumindest der Begriff des topologischen Raumes zur mathematischen Allgemeinbildung zählt.

22.8 Definition (Topologie, topologischer Raum)

(a) Sei X eine nicht leere Menge. Ein System $\mathcal{O}$ von Teilmengen von X heißt *Topologie auf X*, wenn folgendes gilt:

(O_0) $\emptyset \in \mathcal{O},\ X \in \mathcal{O}$.

(O_1) Ist Λ eine beliebige Indexmenge und gilt $O_\lambda \in \mathcal{O}$ für alle $\lambda \in \Lambda$, dann ist auch $\bigcup\limits_{\lambda\in\Lambda} O_\lambda \in \mathcal{O}$.

(O_2) Sind $O_1, \ldots, O_n \in \mathcal{O}$ $(n \in \mathbb{N})$, dann gilt auch $O_1 \cap O_2 \cap \ldots \cap O_n \in \mathcal{O}$.

(b) Ein *topologischer* Raum ist ein Paar $(X,\mathcal{O})$, wobei X eine (nichtleere) Menge und $\mathcal{O}$ eine Topologie auf X ist. Die Teilmengen von X, die zu $\mathcal{O}$ gehören, nennt man *offene Mengen* von $(X,\mathcal{O})$

(c) Die Komplemente (in X) von offenen Mengen heißen *abgeschlossene Mengen* (in X).

22.9 Beispiele und Bemerkungen

(a) Die Bedingung (O_0) ist eine formale Folgerung aus (O_1) und (O_2), wenn man die Konventionen

$$\bigcup_{\lambda\in\emptyset} O_\lambda = \emptyset \quad \text{und} \quad \bigcap_{\lambda\in\emptyset} O_\lambda = X$$

verwendet.

(b) Jeder metrische Raum (X,d) wird zu einem topologischen Raum, wenn man für $\mathcal{O}$ das System $\mathcal{F}_d$ der offenen Mengen in einem metrischen Raum wählt. Man nennt $\mathcal{F}_d$ die von der Metrik d *induzierte Topologie*.

Ist $X_0 \subset X$ eine nicht leere Teilmenge von X und betrachtet man auf X_0 die durch $d|X_0 \times X_0$ definierte Metrik d_0, so ist (X_0, d_0) ein eigenständiger metrische Raum und damit auch topologischer Raum mit dem System $\mathcal{F}_{d_0}$ der offenen Mengen (in X_0).

Eine Teilmenge $U_0 \subset X_0$, die offen bezüglich $\mathcal{F}_{d_0}$ ist, nennen wir auch X_0-offen (offen oder relativ offen in X_0).

Für die bezüglich d_0 gebildete Kugel

$$U_r^{d_0}(a) := \{x \in X_0;\ d_0(x,a) = d(x,a) < r\}$$

gilt offensichtlich

$$U_r^{d_0}(a) = U_r(a) \cap X_0.$$

Aus diesem Grund ist eine Teilmenge $U_0 \subset X_0$ genau dann X_0-offen, wenn es eine in X offene Menge U gibt mit

$$U_0 = U \cap X.$$

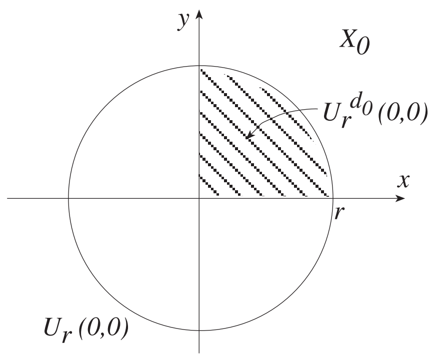
Analog definiert man der Begriff X_0-abgeschlossen. Eine Teilmenge $A_0 \subset X_0$ ist genau dann X_0-abgeschlossen, wenn es eine in X abgeschlossene Teilmenge A gibt mit $A_0 = A \cap X_0$.

Zur Illustration:
Ist etwa $X = \mathbb{R}^2$ (mit der euklidischen Metrik) und

$$X_0 = \{(x, y) \in \mathbb{R}^2;\ x \geq 0 \text{ und } y \geq 0\}$$

der „erste Quadrant“, dann ist (für $r > 0$)

$$\begin{aligned} U_r^{d_0}(0,0) &= \{(x, y) \in \mathbb{R}^2;\ x \geq 0,\ y \geq 0;\ d_0(x, y) < r\} \\ &= U_r(0,0) \cap X_0. \end{aligned}$$



Man beachte, dass $U_r^{d_0}(0,0)$ offen in X_0 ist, dass aber $U_r^{d_0}(0,0)$ nicht offen in $X = \mathbb{R}^2$ ist: Der Punkt $\left(\frac{r}{2}, 0\right) \in X_0$ ist innerer Punkt bezüglich X_0, aber nicht bezüglich X.

Dass eine Teilmenge $U_0 \subset X_0$ X_0-offen ist, bedeutet also nicht, dass U_0 offen in X ist und $U_0 \subset X_0$ gilt.

Ist jedoch die Teilmenge X_0 selber offen in X, dann ist eine Teilmenge $U_0 \subset X_0$ genau dann X_0-offen, wenn sie offen in X ist.

Weitere Beispiele:
Betrachtet man $\mathbb{R}$ mit der natürlichen Metrik, dann ist das Intervall $X_0 := [0, 1[\subset \mathbb{R}$ weder offen noch abgeschlossen in $\mathbb{R}$, aber X_0 ist offen und abgeschlossen in X_0 (der ganze Raum ist immer offen und abgeschlossen).

Ebenso ist das Intervall $[0, \frac{1}{2}[$ offen in X und das Intervall $[\frac{1}{2}, 1[$ abgeschlossen in X.

Übungsaufgabe:
Geben Sie eine in $\mathbb{R}$ offene Menge U bzw. eine in $\mathbb{R}$ abgeschlossene Menge A an mit

$$\left[0,\ \frac{1}{2}\right[= U \cap X_0 \text{ bzw. } \left[\frac{1}{2},\ 1\right[= A \cap X_0.$$

(c) Auf jeder nicht leeren Menge X ist $\mathcal{O}_{ind} := \{X, \emptyset\}$ eine Topologie auf X, man spricht von der *indiskreten (chaotischen) Topologie* auf X: Es gibt nur zwei offene Mengen X und $\emptyset$.

(d) Ein anderes Extrem: Man nimmt für $\mathcal{O}$ *alle* Teilmengen von X, d.h.

$$\mathcal{O} := \mathcal{O}_{diskret} := \mathcal{P}(X) = \text{Potenzmenge von } X.$$

$(X, \mathcal{O}_{diskret})$ heißt *diskreter* topologische Raum.

Die diskrete Metrik auf einer Menge X induziert gerade die diskrete Topologie. Bei der diskreten Topologie sind z.B. auch einelementigen Mengen $\{x\},\ (x \in X)$ offen.

(e) *Die natürliche Topologie* O_n auf $\mathbb{R}$ erhält man, wenn man als offene Mengen die Vereinigungen offener Intervalle $]a, b[$ nimmt.

(f) Für zwei Metriken d und d' auf einer Menge X können die Topologien $\mathcal{F}_d$ und $\mathcal{F}_{d'}$ übereinstimmen. Man nennt die Metriken dann *topologisch äquivalent*, z.B. sind die Metriken

$$d : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \ \text{ und } \ d' : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

mit

$$d(x, y) = |x - y| \ \text{ bzw. } \ d'(x, y) = \left| \frac{x}{1 + |x|} - \frac{y}{1 + |y|} \right|$$

topologisch äquivalent.

(g) Ein *hinreichendes* Kriterium für die topologische Äquivalenz zweier Metriken:
Sind d und d' Metriken auf X und gibt es positive reelle Konstanten c und C, so dass für alle $x, y \in X$ gilt

$$\boxed{cd(x, y) \leq d'(x, y) \leq Cd(x, y)}$$

dann sind d und d' topologisch äquivalent. Solche Metriken nennt man auch *streng äquivalent*.

Zum Beweis muss man zeigen, dass jede offene Kugel bezüglich der einen Metrik eine offene Kugel bezüglich der anderen Metrik enthält und umgekehrt. Das ist aber evident.

22.9.1 Anwendungsbeispiel

Im $\mathbb{K}^n$ sind die gängigen Metriken $d_1,\ d_2$ und d_∞ äquivalent. Dies folgt aus den bekannten Abschätzungen (vgl. ???)

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\, \|x\|_\infty \qquad \text{bzw.} \qquad \tfrac{1}{\sqrt{n}} \|x\|_1 \leq \|x\|_2 \leq \|x\|_1$$

Abbildung 25: Veranschaulichung im Fall $n = 2$.

Die aus diesen Normen abgeleiteten Metriken d_1, d_2, d_∞ ergeben also alle das gleiche System von offenen Mengen auf der $\mathbb{K}^n$ oder vornehmer: Sie erzeugen dieselbe Topologie.
Wie wir im nächsten Paragraphen sehen werden, ergeben sie dann auch den gleichen *Konvergenzbegriff* (für Folgen) auf dem $\mathbb{K}^n$.

Wir schließen diesen Paragraphen mit der Definition *besonderer Punkte*.

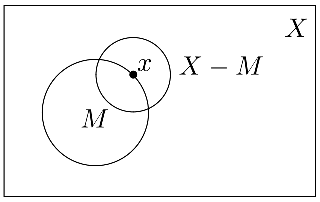
Dazu sei (X, d) ein metrischer Raum und $M \subset X$ eine nicht leere Teilmenge. Bezüglich M lassen sich die Punkte von X in Klassen einteilen:

a. Die Punkte, für welche M Umgebung ist. Das sind gerade die *inneren Punkte*, sie bilden das sog. *Innere* von M. Bezeichnung: M^0. Offensichtlich ist $M^0 \subset M$.

b. Die Punkte, die von $X - M$ umgeben werden, bilden die äußeren Punkte von M oder bezüglich M, ihre Gesamtheit heißt das *Äußere* von (bezüglich) M.

c. Alle übrigen Punkte von X heißen *Randpunkte* von M (besser bezüglich M) und ihre Gesamtheit heißt der *Rand von* M. Bezeichnung: ∂M.

   Ein Randpunkt x von M ist dadurch gekennzeichnet, dass es in jeder (noch so kleinen) Umgebung von x sowohl Punkte von M als auch Punkte vom Komplement $X - M$ gibt:



   ∂M ist anschaulich gesprochen der Ort, wo sich M und $X - M$ nicht treffen.

$\overline{M} := M \cup \partial M$ ist abgeschlossen. $\overline{M}$ heißt die *abgeschlossene Hülle* von M.

Für eine Teilmenge $A \subset X$ gilt

$$A \text{ abgeschlossen} \iff A = \overline{A}.$$

Ein Punkt $x \in \overline{M}$ lässt sich offensichtlich auch dadurch charakterisieren, dass für jede Umgebung U von x gilt

$$M \cap U \neq \emptyset.$$

Jede Umgebung U von x trifft also M, man nennt die Punkte aus $\overline{M}$ daher auch *Berührpunkte* von M.
Man beachte, dass ein Berührpunkt von M nicht zu M zu gehören braucht, dass aber umgekehrt die Punkte von M Berührpunkte von M sind: $M \subset \overline{M}$.

Verwandt mit dem Begriff der Berührpunkte ist der der *Häufungspunkte* einer Teilmenge $M \subset X$.
$x \in X$ heißt *Häufungspunkt* von M, wenn in jede Umgebung U von x ein von x verschiedener Punkt von M liegt, m.a.W., wenn x Berührpunkt von $M - \{x\}$ ist.
Es kann auch vorkommen, dass es zu einem $x \in M$ eine Umgebung U von x gibt mit

$$U \cap M = \{x\},$$

dann heißt x *isolierter Punkt* von M.

Wir fassen unsere Überlegungen zusammen in

22.10 Definition (Besondere Punkte in einem metrischen Raum)

Sei (X, d) ein metrischer Raum, $M \subset X,\ x \in X$ heißt

(a)	*Berührpunkt* von M	$\iff$	Für jede Umgebung U von x gilt $U \cap M \neq \emptyset$
(b)	*Häufungspunkt* von M	$\iff$	Für jede Umgebung U von x gilt $U \cap (M - \{x\}) \neq \emptyset$
(c)	*innerer Punkt* von M	$\iff$	M ist Umgebung von x
(d)	*Randpunkt* von M	$\iff$	x ist Berührpunkt von M und $X - M$
(e)	*äußerer Punkt* von (bzgl.) M	$\iff$	x ist innerer Punkt von $X - M$ $\iff X - M$ ist Umgebung von x
(f)	*isolierter Punkt* von M	$\iff$	Wenn $x \in M$ und wenn es eine Umgebung U von x gibt mit $U \cap M = \{x\}$.

Die Menge $\overline{M}$ aller Berührpunkte heißt der *Abschluss* von M (die abgeschlossene Hülle von M).

M heißt *dicht* in X, wenn $\overline{M} = X$.

Die Menge $\overset{\circ}{M}$ aller inneren Punkte von M heißt das Innere von M (der offene Kern von M).

Beachte: Es gilt stets $\overset{\circ}{M} \subset M$.

M heißt nirgends dicht in X, wenn $\overset{\circ}{\overline{M}} = \emptyset$.

Die Menge aller Randpunkte von M heißt der *Rand* von M; Bezeichnung: ∂M.

Es ist also $\partial M = \overline{M} \cap \overline{X - M}$ (Der Rand einer Menge ist als Durchschnitt zweier abgeschlossenen Mengen also stets abgeschlossen.

Mit M' oder $H(M)$ wird häufig die Menge aller Häufungspunkte von M bezeichnet.
Beachte: $M' = H(M) \subset \overline{M}$.

Der Beweis des folgenden Satzes sei als Übungsaufgabe im Umgang mit den eingeführten Begriffen gestellt:

22.11 Satz

Sei (X, d) ein metrischer Raum, $M \subset X$. Dann gilt:

(a) $\overset{\circ}{M}$ ist die größte in M enthaltene offene Menge von X.

$$M \text{ offen } \iff M = \overset{\circ}{M}.$$

(b) $\overline{M}$ ist abgeschlossen und $\overline{M}$ ist die kleinste abgeschlossene Teilmenge von X, die M umfasst, d.h. ist F irgendeine abgeschlossene Teilmenge von X mit $M \subset F$, dann gilt $\overline{M} \subset F$.

$$M \text{ abgeschlossen } \iff M = \overline{M}$$

$$\overline{M} = \overset{\circ}{M} \cup \partial M = M \cup \partial M = M \cup H(M)$$

$\uparrow$
(disjunkte Vereinigung)

$$M \text{ abgeschlossen } \iff \partial M \subset M \iff H(M) \subset M.$$

Wir erläutern die Begriffe an einigen Beispielen:

22.11.1 Beispiele

(a) Ist $X := (\mathbb{R}^n, \| \|_2),\ r > 0$ und $Y := \overline{U_r}(a) = \{x \in \mathbb{R}^n;\ \|x\|_2 \leq 1\},\ S^{n-1} = \{x \in \mathbb{R}^n;\ \|x\|_2 = 1\}$, dann ist $\partial Y = S^{n-1}$.

In den folgenden Beispielen betrachten wir $X = \mathbb{R}$ und Teilmengen von $\mathbb{R}$.

(b) Sei $M = [0,1] \cup \{3\} \subset M$. Dann ist $\overset{\circ}{M} =]0,1[,\ \overline{M} = [0,1] \cup \{3\} = M$. $\partial M = \{0,1,3\}$, 3 ist isolierter Punkt, denn $U_{\frac{1}{2}}(3) \cap M = \{3\}$

(c) $M := \{\frac{1}{n};\ n \in \mathbb{N}\} \subset \mathbb{R},\quad \overset{\circ}{M} = \emptyset,\ \overline{M} = \{0\} \cup M = \partial M.$
Besitzt M isolierte Punkte?

(d) $M := \mathbb{Q} \subset \mathbb{R}$. Dann ist $\overset{\circ}{\mathbb{Q}} = \emptyset,\ \overline{\mathbb{Q}} = \mathbb{R},\ \mathbb{Q}$ ist also dicht in $\mathbb{R}$, $\partial\mathbb{Q} = \mathbb{R}$. Jede reelle Zahl ist also Randpunkt von $\mathbb{Q}$.

Eine Menge (hier $\mathbb{Q}$) kann also in ihren Rand (hier $\mathbb{R}$) enthalten sein.

Es gilt auch $\overline{\mathbb{R} - \mathbb{Q}} = \mathbb{R}$.

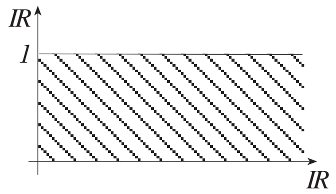
Beides gilt, weil in jeder Umgebung eine reelle Zahl sowohl rationale als auch irrationale Zahlen liegen.

(e) $M = [0,1] \subset \mathbb{R},\ N : [1,2] \subset \mathbb{R}$, dann ist $M \cup N = [0,2],\ (M \cup N)^\circ =]0,2[\neq]0,1[\cup]1,2[= \overset{\circ}{M} \cup \overset{\circ}{N}$.

(f) Ist $M = \mathbb{Z} \subset \mathbb{R}$, dann ist jeder Punkt $k \in \mathbb{Z}$ isolierter Punkt, denn es gilt etwa $]k - \frac{1}{3}, k + \frac{1}{3}[= \{k\}$ für jedes $k \in \mathbb{Z}$.

(g) Sei $M := \{(x,y) \in \mathbb{R}^2;\ 0 \leq x \text{ und } 0 < y \leq 1\} = \mathbb{R}_+ \times]0,1]$, dann ist

$$\begin{aligned} \overset{\circ}{M} &= \{(x,y) \in \mathbb{R}^2;\ x > 0,\ 0 < y < 1\} \\ \bar{M} &= \{(x,y) \in \mathbb{R}^2;\ x \geq 0,\ 0 \leq y \leq 1\} \\ &\quad \text{und} \\ \partial M &= \mathbb{R}_+ \cup \{(0,y);\ 0 \leq y \leq 1\} \cup \{(x,y);\ x \geq 0,\ y = 1\} \end{aligned}$$



(h) Ist jetzt wieder (X,d) ein metrischer Raum und $M = U_r(a)$ die offene Kugel mit dem Mittelpunkt a und Radius $r > 0$, dann gilt $\overline{M} = \overline{U_r(a)} \subset \overline{U_r}(a)$ (abgeschlossene Kugel).

Das ist offensichtlich, da $\overline{U_r}(a)$ abgeschlossen ist (und $U_r(a)$ enthält), während $\overline{M} = \overline{U_r(a)}$ die kleinste abgeschlossene Menge ist, welche $U_r(a)$ enthält: Also ist

$$\overline{U_r(a)} \subset \overline{U_r}(a).$$

Eigentlich würde man das Gleichheitszeichen erwarten, das dies i.A. nicht gilt zeigt das bizarre Beispiel der diskreten Metrik auf einer zumindest zwei elementigen Menge:

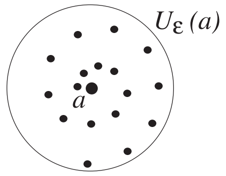
Hier gilt (vgl. $\S$21.2.1(e))
$U_1(x) = \{x\}$, also $\overline{U_r(x)} = \overline{\{x\}} = \{x\}$, aber $\overline{U_1}(x) = X$, also ist $\overline{U_1(x)}$ echt in $\overline{U_1}(x)$ enthalten.

23 Konvergenz und Stetigkeit in metrischen Räumen

In diesem Abschnitt behandeln wir die Konvergenz von Folgen in metrischen Räumen, ferner wird der Begriff der stetigen Abbildung zwischen metrischen Räumen diskutiert. Dies verallgemeinert die Begriffsbildungen über Konvergenz und Stetigkeit aus der Analysis 1. Die im ersten Abschnitt eingeführten Begriffe (wie etwa r-Umgebung, offen, abgeschlossen etc) werden uns dabei nützlich sein.

23.1 Definition (Folgen-Konvergenz in einem metrischen Raum)

Sei (X, d) ein metrischer Raum. Eine Folge $(x_k)_{k\in\mathbb{N}}$ von Elementen $x_k \in X$ heißt konvergent, wenn es ein $a \in X$ mit folgenden Eigenschaften gibt:
Zu jeder ε-Umgebung $U_\varepsilon(a)$ von a gibt es ein $N \in \mathbb{N}$, so dass für alle $k \in \mathbb{N}$ mit $k \geq N$ gilt



$$x_k \in U_\varepsilon(a) \ .$$

Wegen $x_k \in U_\varepsilon(a) \iff d(x_k, a) < \varepsilon$ ist die Konvergenzbedingung gleich bedeutend mit:
Die Folge (nicht negativer) reeller Zahlen (der Abstände) $(d(x_k, a))$ ist eine Nullfolge.

Man kann die Konvergenzbedingung auch so ausdrücken: In einer beliebig vorgegebenen $\varepsilon-$Umgebung $U_\varepsilon(a)$ liegen *fast alle* (oder schließlich alle) Folgenglieder (das heißt wie früher: Alle, bis auf endlich viele Ausnahmen oder von einer bestimmten Stelle an alle nachfolgenden Folgenglieder).

Wegen der Hausdorff-Eigenschaft eines metrischen Raumes ist a im Fall der Existenz eindeutig bestimmt und man nennt a den *Grenzwert der Folge* (x_k).

Man kann deshalb wieder kurz

$$a = \lim_{k\to\infty} x_k$$

schreiben, wenn (x_k) gegen a konvergiert. Weitere Schreibweise: $x_k \to a \ (k \to \infty)$.

23.2 Beispiele und Bemerkungen

(a) Man kann in der Definition 22.1 „jede r-Umgebung $U_r(a)$" ersetzen durch: „jede Umgebung U von a". Denn wenn die Bedingung für *jede* Umgebung von a gilt, gilt sie insbesondere für jede r-Umgebung $U_r(a)$.

Weil eine beliebige Umgebung U von a stets eine r-Umgebung $U_r(a)$ enthält, gilt auch die Umkehrung.

(b) Sind d und d' äquivalente Metriken auf X, dann gilt für einen Punkt $a \in X$ und eine Folge $(x_k),\ x_k \in X$:

$$\lim_{k\to\infty} d(x_k, a) = 0 \iff \lim_{k\to\infty} d'(x_k, a) = 0 \ .$$

(c) *Klar ist*: Gilt $\lim\limits_{k\to\infty} x_k = a$, dann gilt auch $\lim\limits_{j\to\infty} x_{k_j} = a$ für jede Teilfolge (x_{k_j}) von (x_k). Dabei ist $1 \le k_1 < k_2 < \ldots < k_j < \ldots$ eine streng monoton wachsende Folge natürlicher Zahlen.

(d) Ist $X = \mathbb{K}^n$ und $d = d_\infty$, dann gilt:

Eine Folge (x_k) mit $x_k = (x_{k1}, \ldots, x_{kn}) \in \mathbb{R}^n$ konvergiert genau dann gegen $a = (a_1, \ldots, a_n) \in \mathbb{K}^n$, wenn für jedes $j \in \{1, \ldots, n\}$ die „Koordinatenfolge“ (x_{k_j}) gegen a_j konvergiert:

$$\lim_{k\to\infty} x_{kj} = a_j \ \text{ für } \ j \in \{1, \ldots, n\} \ .$$

Beweis : Sei (x_k) konvergent gegen a (in der d_∞-Metrik). Dann gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $k \in \mathbb{N}$ mit $k \ge N$ gilt

$$\|x_k - a\|_\infty < \varepsilon \ .$$

Wegen

$$|x_{kj} - a_j| \le \max_{1\le j\le n} \{|x_{kj} - a_k\} = d_\infty(x_k, a) = \|x_k - a\|_\infty$$

gilt dann auch $\lim\limits_{k\to\infty} x_{kj} = a_j \ (1 \le j \le n)$.

Ist umgekehrt für jedes $j \in \{1, \ldots, n\}$ diese Bedingung erfüllt, dann gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $k \ge N$ und alle $j \in \{1, \ldots, n\}$ gilt

$$|x_{kj} - a_j| < \varepsilon \ .$$

Ist für $j_0 \in \{1, \ldots, n\}$

$$|x_{kj_0} - a_{j_0}| = \max_{1\le j\le n} \{|x_{kj} - a_j|\} = \|x_k - a\|_\infty \ ,$$

dann gilt auch

$$\|x_k - a\|_\infty < \varepsilon \ ,$$

für $k \ge N$, d.h.

$$\lim_{k\to\infty} x_k = a \ .$$

□

Im $\mathbb{R}^3$ ist z.B. die Folge (x_k) mit

$$x_k = \left(\sqrt[k]{k},\ \frac{1}{k!},\ \left(1 + \frac{3}{k}\right)^k \right)$$

konvergent mit dem Grenzwert $(1, 0, e^3)$.

Da die euklidische Metrik d_2 und die Maximumsmetrik auf $\mathbb{K}^n$ äquivalent sind, gilt die entsprechende Aussage auch für die euklidische Metrik:
Auch Konvergenz einer Folge (x_k) bezüglich der euklidischen Metrik ist äquivalent mit der koordinatenweisen (komponentenweisen) Konvergenz.

Im Spezialfall $\mathbb{K} = \mathbb{C}$ haben wir das früher direkt gezeigt.

Im Fall $\mathbb{K} = \mathbb{R}$ liefert die obige Feststellung die Möglichkeit die klassischen Konvergenzkriterien (welche sind das?) anzuwenden (man wendet sie auf die Koordinatenfolgen an, die reelle Zahlfolgen sind).

Im Paragraphen über Kompaktheit werden wir zeigen, dass im Standardvektorraum $\mathbb{K}^n,\ k \in \mathbb{N}$, sogar *alle* Normen äquivalent sind, d.h. den gleichen Konvergenzbegriff liefern.

(e) *Rechenregeln* für konvergente Folgen erhält man für einen *normierten Vektorraum*:
Ist $(V, \| \ \|)$ ein normierter $\mathbb{K}$-Vektorraum und $d(x,y) = \|x-y\|$ die aus der Norm abgeleitete Metrik.

Dann gilt: Sind (x_k) und (y_k) konvergente Folgen in V und ist (t_k) eine konvergente Folge reeller oder komplexer Zahlen, dann sind auch die Folgen

$$(x_k + y_k) \quad \text{und} \quad (t_k x_k)$$

konvergent, und es gilt

$$\lim_{k\to\infty}(x_k+y_k) = \lim_{k\to\infty} x_k + \lim_{k\to\infty} y_k$$

und

$$\lim_{k\to\infty}(t_k x_k) = \lim_{k\to\infty} t_k \cdot \lim_{k\to\infty} x_k.$$

Ferner gilt: Auch die Folge $(\|x_k\|)$ der Normen ist konvergent, und so ist

$$\lim_{k\to\infty} \|x_k\| = \| \lim_{k\to\infty} x_k\| \ .$$

Das folgt aus der Ungleichung $|\ \|x\| - \|y\|\ | \leq \|x-y\|$.

Wird die Norm $\| \ \|$ von einem Skalarprodukt $\langle \ , \ \rangle$ auf V induziert, gilt außerdem

$$\lim_{k\to\infty} \langle x_k, y_k\rangle = \langle \lim_{k\to\infty} x_k, \lim_{k\to\infty} y_k\rangle \ .$$

Die Beweise ergeben sich wie im Fall $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, man braucht der Betrag $| \ |$ nur durch die Norm $\| \ \|$ zu ersetzen.

Die letzte Behauptung über das Skalarprodukt folgt aus Dreiecksungleichung und der Cauchy-Schwarzschen Ungleichung:

Ist $x := \lim\limits_{k\to\infty} x_k$ und $y := \lim\limits_{k\to\infty} y_k$, dann ist

$$\begin{aligned} |\langle x_k, y_k\rangle - \langle x, y\rangle| &= |\langle x_k - x, y_k\rangle - \langle x, y - y_k\rangle| \\ &\leq |\langle x_k - x, y_k\rangle| + |\langle x, y-y_k\rangle| \\ &\leq \|x_k - x\| \, \|y_k\| + \|x\| \, \|y - y_k\|, \end{aligned}$$

wenn man noch beachtet, dass eine konvergente Folge stets beschränkt ist, d.h. alle Folgenglieder in einer abgeschlossenen Kugel liegen.

Abgeschlossene Teilmengen eines metrischen Raumes (X,d) können eine sehr komplexe Struktur haben. Hierzu vergleiche man etwa das Beispiel zur fraktalen Geometrie in Königsberger, Analysis 2, Seite 4 (4.Auflage), das ein zweidimensionales Analogon zum Cantorschen Discontinuum (Cantorsche „Wischmenge“) darstellt. Mit Hilfe des Begriffs der konvergenten Folge lassen sich jedoch abgeschlossene Teilmengen eines metrischen Raums einfach charakterisieren:
„Abgeschlossenheit“ bedeutet das Gleiche wie „Folgenabgeschlossenheit“, d.h. es gilt

23.2.1 Satz

Ist (X,d) ein metrischer Raum und $A \subset X$ eine Teilmenge, dann gilt: A ist genau dann abgeschlossen (in X), wenn für jede konvergente Folge (a_k) mit $a_k \in A$ und $\lim\limits_{k\to\infty} a_k := x \in X$ bereits $x \in A$ gilt.

Beweis : Sei $A \subset X$ abgeschlossen und (a_k) eine Folge mit $a_k \in A$ und $x := \lim\limits_{k\to\infty} a_k \in X$. Wir nehmen an, dass x *nicht* in A liegt. Da das Komplement $X - A$ offen ist, ist dann $X - A$ eine Umgebung von x, in welcher fast alle Folgenglieder liegen müssen: Es gibt also ein $N \in \mathbb{N}$, so dass für alle $k \geq N$ gilt $a_k \in X - A$. Das steht aber im Widerspruch zur Voraussetzung, dass $a_k \in A$ für alle $k \in \mathbb{N}$ gilt.

Umgekehrt sei die „Folgenbedingung" erfüllt. Wir müssen zeigen, dass A dann abgeschlossen ist. Wir nehmen an, dass A *nicht* abgeschlossen ist. Dann ist $U := X - A$ nicht offen. Es gibt daher einen Punkt $x \in U$, so dass keine offene Kugel mit Mittelpunkt x in U liegt. Insbesondere enthält dann jede Kugel $U_{\frac{1}{k}}(x)$ einen Punkt a_k mit $a_k \notin U$. Für die Folge (a_k) gilt also $a_k \in A$ und wegen $d(a_k, x) < \frac{1}{k}$ für alle k konvergiert sie gegen x. Wegen der vorausgesetzten Folgenbedingung muss also $x \in A$ gelten, das steht aber im Widerspruch zu $x \in X - A$. Also war unsere Annahme, dass A nicht abgeschlossen ist, falsch.

□

Bemerkung: Im Fall $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ haben wir mit der „Folgenabgeschlossenheit" gearbeitet.

Im Fall der Standardräume $\mathbb{K}^n,\ n \in \mathbb{N}$, konnten wir die Konvergenz einer Folge einfach auf die Konvergenz in $\mathbb{R}$ reduzieren. Es stellt sich nun die Frage nach *Konvergenzkriterien* in beliebigen metrischen Räumen, d.h. nach Kriterien, die Aussagen über die Konvergenz einer Folge gestatten, ohne dass man den Grenzwert kennt.

Wie im Fall $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ lässt sich der Begriff der *Cauchy-Folge* in einem beliebigen metrischen Raum definieren.

23.3 Definition (Cauchy-Folge)

Ist (X, d) ein metrischer Raum. Eine Folge (x_k) von Elementen $x_k \in X$ heißt *Cauchy-Folge* (in X), wenn folgendes gilt:
Zu jedem $\varepsilon > 0$ gibt es einen Index $N \in \mathbb{N}$, so dass für alle $k, l \in \mathbb{N}$ mit $k, l \geq N$ gilt

$$\boxed{d(x_k, x_l) < \varepsilon}$$

Diese Definition ist nicht überraschend, ebenso wenig die folgende Feststellung:

23.4 Satz (konvergente Folgen sind Cauchy-Folgen)

In einem metrischen Raum (X, d) ist jede konvergente Folge eine Cauchy-Folge.

Beweis : Sei (x_k) eine konvergente Folge in X und $a := \lim\limits_{k\to\infty} x_k$ ihr Grenzwert in X.

Dann gibt es nach Definition zu jedem $\varepsilon > 0$ einen Index $N \in \mathbb{N}$, so dass für alle $k \in \mathbb{N}$ mit $k \geq N$ gilt:

$$d(x_k, a) < \frac{\varepsilon}{2}\ .$$

Ist nun $l \in \mathbb{N}$ und gilt auch $l \geq N$, dann ist auch $d(x_k, a) < \frac{\varepsilon}{2}$.
Also nach der Dreiecksungleichung

$$\begin{aligned} d(x_k, x_l) &= d(x_k, a) + d(a, x_l) \\ &= d(x_k, a) + d(x_l, a) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

für alle $k, l \geq N$, d.h. (x_k) ist eine Cauchy-Folge in X.

□

Eine Cauchy-Folge in X braucht nicht konvergent (in X) zu sein. Räume X, für welche die Umkehrung von Satz 22.4 gilt, in denen also jede Cauchy-Folge einen Grenzwert in X besitzt, erhalten einige Namen.

23.5 Definition (vollständiger metrische Raum, Banach-Raum, Hilbert-Raum)

Ein metrischer Raum (X, d) heißt *vollständig* (oder komplett), wenn jede Cauchy-Folge in X einen Grenzwert in X hat.
Ein normierter Raum $(V, \| \,\|)$, der bezüglich der aus der Norm abgeleiteten Metrik $d(u, v) = \|u-v\|$ vollständig ist, heißt ein *Banach-Raum* (S.Banach, 1892-1945).
Stammt die Norm sogar von einem Skalarprodukt $\langle\ ,\ \rangle$ (ist $\|v\| = \sqrt{\langle v, v\rangle}$ für $v \in V$), dann heißt der vollständige normierte Raum (Banach-Raum) auch *Hilbert-Raum* (D.Hilbert, 1862-1943).

23.6 Bemerkungen und Beispiele

(a) Hilbert-Räume sind also spezielle Banach-Räume.

(b) Der $\mathbb{K}^n,\ n \in \mathbb{N}$ ist bezüglich der euklidischen Norm $\| \,\|_2$ ein Banach-Raum und da $\|x\|_2 = \sqrt{\langle x, x\rangle}$ ($\langle\ ,\ \rangle$ das Standardskalarprodukt auf $\mathbb{K}^n$) gilt, sogar ein *Hilbert-Raum*.

Der $\mathbb{K}^n$ ist auch bezüglich der Normen $\| \,\|_1$ und $\| \,\|_\infty$ ein Banach-Raum, jedoch stammt nur die euklidische Norm $\| \,\|_2$ von einem Skalarprodukt (Begründen Sie, warum die Maximums-Norm $\| \,\|_\infty$ nicht von einem Skalarprodukt induziert wird).

Der Beweis ist klar, denn bezüglich jeder der aus diesen Normen abgeleiteten Metriken gilt, dass eine Folge $(x_k) = (x_{k1}, \ldots, x_{kn} \in \mathbb{K}^n$ genau dann gegen $a = (a_1, \ldots, a_n) \in \mathbb{K}^n$ konvergiert, wenn für jedes $j \in \{1, \ldots, n\}$ die Koordinatenfolge (x_{kj}) gegen a_j konvergiert ($k \in \infty$).
(x_k) ist also genau dann eine Cauchy-Folge in $\mathbb{K}^n$, wenn für jedes $j \in \{1, \ldots, n\}$ die Zahlenfolgen (x_{kj}) Cauchy-Folgen in $\mathbb{K}$ sind.
In $\mathbb{K}$ ist aber jede Cauchy-Folge konvergent. Man beachte dabei, dass im Fall $\mathbb{K} = \mathbb{C} = \mathbb{R}^2$ dies auf der Tatsache beruht, dass in $\mathbb{R}$ jede Cauchy-Folge konvergiert.

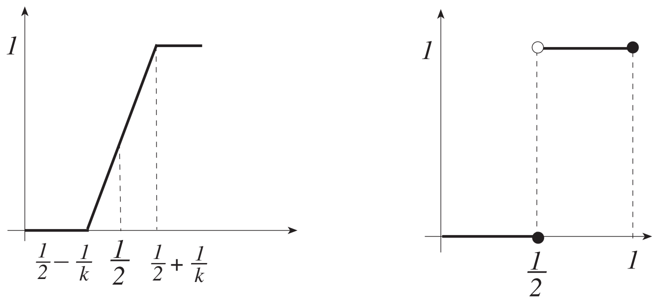
(c) Beispiele für *nicht* vollständige Räume lassen sich einfach angeben:

Man braucht nur aus einem vollständigen Raum einen Grenzwert einer Folge herauszunehmen:
Ist $a \in \mathbb{R}$, dann ist $X := \mathbb{R} - \{a\}$ (mit der von $\mathbb{R}$ induzierten Metrik) für jedes a nicht vollständig, denn etwa die Folge $(x_k) = (a + \frac{1}{k})$ ist eine Cauchy-Folge in X, ihr Grenzwert a liegt aber nicht in X.

Der Raum $\mathcal{C}([0, 1])$ der stetigen $\mathbb{R}$-wertigen Funktionen auf $[0, 1]$ mit der L^2-Norm

$$\|f\|_2 := \left(\int_0^1 f(t)^2 dt \right)^{\frac{1}{2}}$$

ist kein Banach-Raum (Hilbert-Raum), wie etwa das in der Abbildung skizzierte Beispiel zeigt.



(f_k) ist zwar eine Cauchy-Folge in $V := \mathcal{C}([0,1])$ bezüglich der L^2-Norm, die Grenzfunktion f (bezüglich der L^2-Norm) ist aber in $\frac{1}{2}$ nicht stetig.

$\mathcal{C}([0,1])$ ist auch bezüglich der L^1-Norm

$$\|f\|_1 := \int_0^1 |f(t)|dt$$

kein Banach-Raum.

Die Menge $\mathbb{Q}$ der rationalen Zahlen ist mit der von $\mathbb{R}$ induzierten Metrik nicht vollständig. Geben Sie eine Cauchy-Folge von rationalen Zahlen an, die nicht gegen eine rationale Zahl konvergiert.

Häufig kann man die Vollständigkeit eines Raumes A daraus schließen, dass man ihn als *abgeschlossenen* Teilraum eines vollständigen metrischen Raumes erkennt.

(d) Seien (X, d) ein metrischer Raum und $A \subset X$ ein Teilraum (mit der induzierten Metrik)

a) Ist A vollständig, dann ist A abgeschlossen in X.

b) Ist X vollständig und A abgeschlossen in X, dann ist auch A vollständig.

Beweis a): Sei (a_k) ein Folge in A, die in X konvergiert: $a = \lim\limits_{k\to\infty} a_k \in X$. Wir haben zu zeigen, dass $a \in A$ gilt. Da (a_k) eine Cauchy-Folge in A ist, ist (a_k) konvergent in A. Wegen der Eindeutigkeit des Grenzwertes folgt

$$a = \lim_{k\to\infty} a_k \in A\ .$$

□

Beweis b): Sei (a_k) eine Cauchy-Folge in A. Dann ist (a_k) auch eine Cauchy-Folge in X $(a_k \in A \subset X)$ und besitzt daher einen Grenzwert $a \in X$. Da A aber abgeschlossen ist, gilt

$$a = \lim_{k\to\infty} a_k \in A\ ,$$

A ist also vollständig.

□

Weitere Beispiele von vollständigen metrischen Räumen, Banach- und Hilbert-Räumen:

(e) Der Raum $B(M,\mathbb{K})$ der beschränkten $\mathbb{K}$-wertigen Funktionen ($\mathbb{K}=\mathbb{R}$ oder $\mathbb{C}$) auf einem echten Intervall $M\subset\mathbb{R}$ (allgemeiner eine nicht leere Teilmenge $M\subset\mathbb{R}$) ist ein Banach-Raum bezüglich der Supremums-Norm $\|f\|_\infty=\sup\{|f(x)|;\ x\in M\}$.

Zur Erinnerung: Konvergenz einer Folge (f_k) in der Supremums-Norm bedeutet gleichmäßige Konvergenz.

Ist (f_k) eine Cauchy-Folge bezüglich $\|\ \|_\infty$, dann *konvergiert* f_k auf M gleichmäßig gegen eine Grenzfunktion $f:M\to\mathbb{K}$. Zu jedem $\varepsilon>0$ gibt es daher ein $K\in\mathbb{N}$, so dass für alle $x\in M$ und alle $k\geq K$ gilt

$$|f(x)-f_k(x)|<\varepsilon\ .$$

Setzt man $\varepsilon=1$, so ist also für alle $x\in M$ und alle $k\geq K$

$$|f(x)|-|f_k(x)|\leq|f(x)-f_k(x)|<1\ ,$$

daher (wähle $k=K$)

$$|f(x)|<1+|f_K(x)|\leq 1+\|f_K\|_\infty$$

für alle $x\in M$, d.h. f ist ebenfalls beschränkt:

$$f=\lim_{k\to\infty}f_k\in B(M,\mathbb{K})\ .$$

Nimmt man $V:=\mathcal{C}([a,b],\mathbb{K})=\{f:[a,b]\to\mathbb{K};\ f\ \text{ stetig }\}$, mit der Supremums-Norm (= Maximums-Norm), dann ist V ein abgeschlossener Unterraum von $B([a,b],\mathbb{K})$, daher nach dem Kriterium ebenfalls vollständig.

Ist $V=R([a,b],\mathbb{K})$ der $\mathbb{K}$-Vektorraum der $\mathbb{K}$-wertigen Regelfunktionen auf dem kompakten Intervall $[a,b]$ und ist (f_k) eine Folge von Regelfunktionen $f_k:[a,b]\to\mathbb{K}$, die gleichmäßig gegen $f:[a,b]\to\mathbb{K}$ konvergiert, dann ist nach dem Stabilitätssatz (vgl. §17.2.9) f ebenfalls eine Regelfunktion, daher $R([a,b],\mathbb{K})$ als abgeschlossener Teilraum des Banach-Raumes $B([a,b],\mathbb{K})$ ebenfalls ein Banach-Raum.

Der Raum $T([a,b])$ der Treppenfunktionen ist bezüglich der Supremums-Norm kein Banach-Raum.

Die Unterräume $\mathbb{Q}$ und $]0,1[$ von $\mathbb{R}$ können nicht vollständig sein, da sie nicht abgeschlossen sind.

Der Vektorraum $\mathcal{C}^\infty(\mathbb{K})$ aller *beschränkten* reellen oder komplexen Zahlenfolgen ist ein Banach-Raum bezüglich der Norm

$$\|x\|_\infty=\max\{|x_1|,|x_2|,|x_3|,\ldots\}\ ,$$

dabei sei $x=(x_1,x_2,x_3,\ldots,x_j,\ldots),\ x_j\in\mathbb{K}$.

Der Hilbertsche-Folgenraum $l^2(\mathbb{K})$ ist ein Hilbert-Raum (vgl. Beispiel 21.1.1(g)). $l^2(\mathbb{K})$ ist in natürlicherweise ein $\mathbb{K}$-Vektorraum:
Ist $x=(x_1,x_2,\ldots)\in l^2(\mathbb{K})$ und $y=(y_1,y_2,\ldots)\in l^2(\mathbb{K})$, dann sind auch

$$\lambda x+\mu y\in l^2(\mathbb{K})$$

für beliebige $\lambda,\mu\in\mathbb{K}$.

Wir zeigen dies für $\lambda=\mu=1$. Sind nämlich $x=(x_j)$ und $y=(y_j)\in l^2(\mathbb{K})$, dann ergibt sich aus der Cauchy-Schwarzschen Ungleichung

$$\sum_{j=1}^{n}|\bar{x}_jy_j|\leq\sqrt{\sum_{j=1}^{n}|x_j|^2\cdot\sum_{j=1}^{n}|y_j|^2}\leq\sqrt{\sum_{j=1}^{\infty}|x_j|^2\cdot\sum_{j=1}^{\infty}|y_j|^2}$$

die absolute Konvergenz der Reihe $\sum\limits_{j=0}^{\infty}\bar{x}_jy_j$ (die Teilsummen der Absolutreihe sind monoton und beschränkt!).

Damit ist die Definition (Skalarprodukt)

$$\langle x, y\rangle := \sum_{j=0}^{\infty} \bar{x}_j y_j$$

für $x = (x_j),\ y = (y_j) \in l^2(\mathbb{K})$ sinnvoll.
Damit ist dann $\sqrt{\sum\limits_{j=0}^{\infty} |x_j|^2} = \sqrt{\langle x, x\rangle} =: \|x\|_2$.

Ferner ist dann

$$\sum_{j=1}^{\infty} |x_j + y_j|^2 \le \|x\|_2^2 + 2\sum_{j=0}^{\infty} |\bar{x}_j y_j| + \|y\|_2^2 < \infty$$

und damit

$$(x_j) + (y_j) = (x_j + y_j) \in l^2(\mathbb{K}) \ .$$

Die hätte man auch einfacher aus der Ungleichung $|x_j y_j|^2 \le 2(|x_j|^2 + |y_j|^2)$ schließen können.

Wir zeigen, dass $l^2(\mathbb{K})$ vollständig, d.h. hier ein Hilbert-Raum ist.

Die Folge (x_k) mit $x_k = (x_{k_1}, x_{k_2}, \ldots) \in l^2(\mathbb{K})$ sei eine Cauchy-Folge in $l^2(\mathbb{K})$.
Wir werden zeigen, dass für jedes j die Komponentenfolge $(x_{kj})_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{K}$ sein muss.

Beachte: Wir müssen „Folgen von Folgen" betrachten.

Sei $\varepsilon > 0$ beliebig vorgegeben. Dann gibt es ein $N \in \mathbb{N}$, so dass für alle $k, l \in \mathbb{N}$ mit $k, l \ge N$ gilt

$$(*) \qquad \|x_k - x_l\|_2^2 = \sum_{j=1}^{\infty} |x_{kj} - x_{lj}|^2 < \varepsilon^2 \ .$$

Dann gilt erst recht für alle $k, l \ge N$ und alle $j \in \mathbb{N}_0$:

$$|x_{kj} - x_{lj}| < \varepsilon \ .$$

Für jedes j ist die Kompenentenfolge $(x_{kj})_{k\in\mathbb{N}}$ also eine Cauchy-Folge in $\mathbb{K}$, besitzt also einen Grenzwert $x_j \in \mathbb{K}$.

Aus $(*)$ ergibt sich ferner für jedes $n \in \mathbb{N}_0$ und alle $k, l \ge N$ die Ungleichungen

$$\sum_{k=1}^{n} |x_{kj} - x_{lj}|^2 < \varepsilon^2$$

und hieraus für $l \to \infty$

$$\sum_{k=0}^{n} |x_{kj} - x_j|^2 \le \varepsilon^2$$

für $k \ge N$ und alle n und damit auch wieder für $n \to \infty$

$$(**) \qquad \sum_{j=0}^{\infty} |x_{kj} - x_j|^2 \le \varepsilon^2$$

für alle $k \ge N$.

Die Folge $x := (x_1, x_2, x_3, \ldots)$ hat die Eigenschaft, dass $x - x_N$ in $l^2(\mathbb{K})$ liegt ($x_N := (x_{N1}, x_{N2}, \ldots)$).

Wegen $x = x - x_N + x_N$ liegt x selber in $l^2(\mathbb{K})$.

Die Ungleichung $(**)$ kann man jetzt in der Form $\|x_k - x\|_2^2 \le \varepsilon$ für $k \ge N$ schreiben, in der sie besagt, dass $\lim\limits_{k\to\infty} x_k = x$ gilt.

Der Raum $l^2(\mathbb{K})$ ist ein fundamentales Beispiel für einen nicht endlich dimensionalen Hilbert-Raum.

Als Übungsaufgabe zeige man:

Übungsaufgabe: Ist (X, d) ein diskreter metrische Raum, dann ist X vollständig.

Tipp: Was sind die konvergenten Folgen in (X, d)?

Die Vollständigkeit hängt also von der *Metrik* ab. Obwohl äquivalente Metriken den gleichen Konvergenzbegriff ergeben können, hängt der Begriff der Vollständigkeit nicht nur von der Äquivalenzklasse der Metrik ab:
Betrachtet man als Beispiel $X = \mathbb{R}$ mit der natürlichen Metrik $d(x, y) := |x - y|$ und mit der zweiten Metrik

$$d'(x, y) = |\exp(x) - \exp(y)| \ ,$$

dann sind d und d' äquivalent (man weise dies nach).
Betrachtet man jedoch die Folge

$$(y_k) = \left(\log\frac{1}{k}\right), \ k \in \mathbb{N} \ ,$$

dann ist (y_k) eine Cauchy-Folge bezüglich d', nicht aber bezüglich d!
Gilt jedoch für zwei Metriken d und d' auf X Abschätzungen vom Typ (vgl. §21.9(g))

$$\boxed{cd(x, y) \leq d'(x, y) \leq Cd(x, y)}$$

für alle $x, y \in X$ mit positiven Konstanten C und c, dann ist (X, d) genau dann vollständig, wenn (X, d') vollständig ist.

Die Standardmetriken in $\mathbb{K}^n$ erfüllen solche Abschätzungen, daher ist $\mathbb{K}^n$ bezüglich *jeder* der Metriken $d_1, \ d_2, \ d_\infty$ vollständig (bzw. ein Banach-Raum).

Übungsaufgabe: Man zeige: Für $p, q \in \mathbb{N} \cup \{\infty\}$ mit $1 \leq p < q \leq \infty$ gilt

$$\|x\|_q \leq \|x\|_p \ \text{ und } \ l^p(\mathbb{K}) \subset l^q(\mathbb{K}) \ ,$$

daher sei für $x = (x_j) \in l^p(\mathbb{K})$ und $p \in \mathbb{N}$

$$\|x\|_p := \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}} \qquad \text{und} \qquad \|x\|_\infty := \sup\{|x_1|, |x_2|, \ldots\} \ .$$

Ferner zeige man, dass alle Räume $l^p(\mathbb{K})$ Banach-Räume sind.

Tipp: Beim letzten Teil orientiere man sich am Beispiel $p = 2$, wo die Behauptung gezeigt wurde.

Wir schließen unsere Betrachtungen über Konvergenz in metrischen Räumen zunächst mit einem *Fixpunktsatz*, der den Banachschen Fixpunktsatz aus Analysis 1 auf vollständige metrische Räume verallgemeinert.

23.6.1 Theorem (allg. Banachscher Fixpunktsatz)

Sei X ein vollständiger metrische Raum und $f : X \to X$ eine kontrahierernde Selbstabbildung, d.h. es gebe ein $q \in [0, 1[$, so dass für alle $x, y \in X$ gilt

$$(*) \qquad \boxed{d(f(x), f(y)) \le qd(x, y)}$$

dann besitzt f genau einen Fixpunkt, d.h. es gibt genau ein $x_* \in X$ mit $f(x_*) = x_*$.

Zusätze:

(a) Definiert man mit einem beliebigen Startpunkt $x_0 \in X$ rekursiv $x_k = f(x_{k-1}),\ k \in \mathbb{N}$, also

$$x_1 = f(x_0), \quad x_3 = f(x_1), \quad x_3 = f(x_2), \quad \text{etc.}$$

dann ist die Folge $(x_k)_{k \in \mathbb{N}}$ konvergent und es gilt

$$\lim_{k \to \infty} x_k = x_* \ .$$

(b) Für alle $k \in \mathbb{N}$ gelten die Fehlerabschätzungen

$$\boxed{d(x_k, x_*) = \tfrac{d(x_{k+1}, x_k)}{1-q} \le \tfrac{q^k}{1-q} d(x_1, x_0)}$$

und für beliebiges $x \in X$ gilt

$$\boxed{d(f(x), x_*) \le \tfrac{q}{1-q} d(f(x), x)}$$

und für alle $k \in \mathbb{N}_0$

$$\boxed{d(x_{k+1}, x_*) \le \tfrac{q}{1-q} d(x_{k+1}, x_k)}$$

(c) Die Folge (x_k) konvergiert *linear* gegen x_*, d.h. hier: Es gilt für alle $k \in \mathbb{N}_0$

$$d(x_{k+1}, x_*) \le qd(x_k, x_*) \ .$$

Der **Beweis** dieses weitreichenden Satzes ist relativ einfach. Wir werden (mit Hilfe des Cauchy-Kriteriums) zeigen, dass die Folge (x_k) konvergiert (die Bedingung, dass f eine Selbstabbildung von X ist, bewirkt, dass alle x_k wieder in X liegen), ihr Grenzwert x_* ist dann notwendig ein Fixpunkt von f (und zwar der einzige), denn es gilt für beliebige $k \in \mathbb{N}$

$$\begin{aligned} d(x_*, f(x_*)) &\le d(x_*, x_k) + d(x_k, f(x_*)) \\ &= d(x_*, x_k) + d(f(x_{k-1}), f(x_*)) \\ &\le d(x_*, x_k) + qd(x_{k-1}, x_*) \end{aligned}$$

und wegen $\lim\limits_{k \to \infty} d(x_k, x_*) = 0$ folgt

$$d(x_*, f(x_*)) \le 0 \ .$$

Andererseits ist stets $d(x_*, f(x_*)) \ge 0$, also ist $d(x_*, f(x_*)) = 0$ und dies ist äquivalent mit

$$f(x_*) = x_* \ .$$

Zur Eindeutigkeit des Fixpunktes:
Wäre $y_* \in X$ ein weiterer von x_* verschiedener Fixpunkt von f, dann wäre

$$d(x_*, y_*) = d(f(x_*), f(y_*)) \leq q d(x_*, y_*)$$

und damit $q \geq 1$, im Widerspruch zu $0 \leq q < 1$.

Zur Existenz und Konvergenz:
Wir benutzen zum Beweis die sog. „*Defektungleichung*“:

$$(D): \qquad \boxed{d(x,y) \leq \tfrac{1}{1-q}\left(d(f(x),x) + d(f(y),y)\right)}$$

die sich aus der Ungleichung

$$d(x,y) \leq d(x, f(x)) + d(f(x), f(y)) + d(f(y), y)$$

(man muss die Dreiecksungleichung zweimal anwenden) ergibt, wenn man im mittleren Term

$$d(f(x), f(y)) \leq q d(x,y)$$

beachtet.
Aus der Dreiecksungleichung folge übrigens sofort auch die Eindeutigkeit des Fixpunktes.
Mit Induktion folgt für alle $k \in \mathbb{N}_0$

$$(**) \qquad \boxed{d(x_{k+1}, x_k) \leq q^k d(x_1, x_0)}$$

Wir setzen in der Defektungleichung(D) $x := x_{k+p}$ und $y := x_k,\ k, p \in \mathbb{N}_0$, und erhalten mit der Dreiecksungleichung

$$\begin{aligned}(***) \qquad d(x_{k+p}, x_k) &\leq \frac{1}{1-q}\left(d(x_{k+p+1}, x_{k+p}) + d(x_{k+1}, x_k)\right)\\ &\overset{(**)}{\leq} \frac{1}{1-q}\left(q^{k+p} + q^k\right) d(x_1, x_0)\\ &\leq \frac{2q^k}{1-q} d(x_1, x_0) \qquad \text{für beliebiges } p > 0\end{aligned}$$

Da die Folge (q^k) ein reelle Nullfolge ist, folgt daher, dass die Folge (x_k) eine Cauchy-Folge in X ist und wegen der vorausgesetzten Vollständigkeit von X also einen Grenzwert x_* in X hat:

$$x_* = \lim_{k \to \infty} x_k \; .$$

Dass dieses x_* dann ein Fixpunkt von f ist (und zwar der einzige), haben wir bereits eingangs gezeigt.

23.6.2 Übungsaufgabe

Man beweise die Ungleichung $(***)$ direkt (d.h. ohne Benutzung der Defektungleichung (D)) durch fortgesetzte Anwendung der Dreiecksungleichung.

Die Fehlerabschätzungen aus (b) erhält man, wenn man in der Defektungleichung (D) $x := x_*$ und $y := y_*$ setzt.

$$\begin{aligned} d(x_*, x_k) &\leq \frac{1}{1-q}(\underbrace{d(f(x_*), x_*)}_{=0} + d(f(x_k), x_k)) \\ &= \frac{1}{1-q} d(f(x_k), x_k) \\ &= \frac{1}{1-q} d(x_{k+1}, x_k) \\ &\leq \frac{q^k}{1-q} d(x_1, x_0). \end{aligned}$$

Setzt man hier $k := 1,\ x := x_0,\ f(x) = f(x_0) = x_1$, so folgt

$$d(x_*, f(x)) \leq \frac{q}{1-q} d(f(x), x) \ .$$

Setzt man in der letzten Ungleichung $x = x_k$ und $f(x) = x_{k+1}$, so folgt

$$d(x_*, x_{k+1} \leq \frac{q}{1-q} d(x_{k+1}, x_k) \ .$$

Weil man die letzten beiden Ungleichungen im „Nachhinein" gewonnen hat, nennt man sie auch *a-posteriori-Abschätzungen*, während man die Ungleichung

$$d(x_*, x_k) \leq \frac{q^k}{1-q} d(x_1, x_0)$$

eine *a-priori-Abschätzung* nennt.

Die lineare Konvergenz ist offensichtlich:

$$d(x_*, x_{k+1}) = d(f(x_*), f(x_k)) \leq q d(x_*, x_k) \ .$$

23.6.3 Bemerkungen und Beispiele

(a) Weil unter den Voraussetzungen des Banachschen Fixpunktsatzes die rekursiv definierte Folge $x_k := f(x_{k-1}),\ k \geq 1$, ($x_0 \in X$ beliebig) gegen den Fixpunkt konvergiert, sagt man auch: x_* wird mit der *Methode der sukzessiven Approximation* näherungsweise bestimmt.

(b) In der Theorie einer Veränderlichen hat man mit dem Mittelwertsatz der Differenzialrechnung und Monotonieüberlegungen schlagkräftige Werkzeuge, um festzustellen, ob eine Selbstabbildung vorliegt und ab sie eine Kontraktion ist.

Ist $M \subset \mathbb{R}$ z.B. ein kompaktes Intervall und $f : M \to \mathbb{R}$ eine differenzierbare Funktion mit $q := \sup\{|f'(x)|;\ x \in M\} < 1$, dann gilt nach dem MWS der Differenzialrechnung

$$f(x) - f(y) = f'(\xi)(x - y) \ ,$$

also

$$|f(x) - f(y)| \leq q|x - y| \ ,$$

also hat die Gleichung $x = f(x)$ eine eindeutig bestimmte Lösung $x_* \in M$, die man „sukzessive Approximation" erhält, wenn man mit einem beliebigen $x_0 \in M$ startet und

$$x_1 = f(x_0), \quad x_2 = f(x_1), \quad x_3 = f(x_2) \quad \text{etc.}$$

bildet.

(c) Wie schon festgestellt, ist es im allgemeinen Fall manchmal schwierig festzustellen, ob die gegebene Abbildung $f : X \to X$ eine Selbstabbildung ist und ob sie kontrahierend ist.

In vielen Fällen kann die Voraussetzung $f : X \to X$, d.h. f ist eine Selbstabbildung, abgeschwächt werden. Gibt es nämlich einen Startwert $x_0 \in X$, so dass die Iteration $x_{k+1} = f(x_k)$ unbeschränkt durchführbar ist, so gelten die Aussagen des Fixpunktsatzes für dieses x_0.

In der Anwendung liegt häufig folgende Situation vor:
Sei $(V, \| \ \|)$ ein Banach-Raum (über $\mathbb{R}$ oder $\mathbb{C}$),

$$X := \overline{U_r}(x_0) = \{x \in X;\ d(x, y_0) \leq r\}$$

die abgeschlossene Kugel mit Mittelpunkt $x_0 \in V$ und Radius $r > 0$ und $f : X \to V$ eine Abbildung mit

$$\|f(x) - f(y)\| \leq q\|x - y\|$$

für alle $x, y \in X$ und $q \in [0, 1[$.

Ferner sei

$$\|f(x_0) - x_0\| \leq (1 - q)r \ .$$

Dann hat f genau einen Fixpunkt in X und mit dem Startwert x_0 konvergiert die durch $x_k = f(x_{k-1}),\ k \in \mathbb{N}$, definierte rekursive Folge gegen den Fixpunkt.

Zum **Beweis** muss man nur nachweisen, dass alle x_k in X liegen. Denn X ist ein abgeschlossener Teil eines Banach-Raumes und damit selbst vollständig. Wenn die Folge (x_k) sinnvoll definiert ist, ist sie eine Cauchy-Folge und ihr Grenzwert x_* ist der Fixpunkt.

Nach Definition ist $x_1 = f(x_0)$ und damit ist

$$\|f(x_0) - x_0\| = \|x_1 - x_0\| \leq (1 - q)r \leq r \ ,$$

also auch $x_1 = f(x_0) \in X$.

Nehmen wir an, es seien $x_1, \ldots, x_k \in X$, dann gilt in Analogie zu $(**)$

$$\|x_{k+1} - x_0\| \leq \frac{1 - q^{k+1}}{1 - q}\|x_1 - x_0\| \leq (1 - q^{k+1})r \leq r \ ,$$

also leigt auch $x_{k+1} \in X$ und die Iteration ist unbeschränkt durchführbar.

(d) Auf dem Banachschen Fixpunktsatz beruhen zahlreiche Existenz- und Eindeutigkeitssätze der höheren Analysis. Wir werden ihn anwenden z.B. um

- den Umkehrsatz für differenzierbare Abbildungen zu beweisen;
- Existenz- und Eindeutigkeitssätze für Differentialgleichungen (Picard-Lindelöf) zu beweisen;
- Integralgleichungen zu lösen;
- ein Newton-Verfahren in $\mathbb{R}^n$ zu entwickeln;

  etc.

Hier nur eine Anwendung in der Linearen Algebra, wo man den Banachschen Fixpunktsatz auch zum Lösen von linearen Gleichungssystemen verwenden kann.

Seien dazu $A \in \mathbb{R}^{n\times n}$ also eine $n \times n-$Matrix und $b \in \mathbb{R}^n$ vorgegeben.
Wir wollen das Gleichungssystem

$$\boxed{x = Ax + b}$$

mit dem Banachschen Fixpunktsatz lösen und betrachten dazu die (affine) Abbildung

$$f : \mathbb{R}^n \to \mathbb{R}^n \ \text{ mit } \ f(x) = Ax + b \ .$$

Für $x, y \in \mathbb{R}^n$ ist z.B. mit der Maximums-Norm $\| \ \|_\infty$ auf $\mathbb{R}^n$

$$\begin{aligned} \|f(x) - f(y)\|_\infty &= \|Ax + b - (Ay + b)\|_\infty \\ &= \|A(x - y)\|_\infty \\ &\leq L\|x - y\|_\infty \end{aligned}$$

mit $L := \max\limits_{1 \leq j \leq k} \left\{ \sum\limits_{k=1}^{n} |a_{jk}| \right\}$.

L ist die sogenannte *Zeilensummennorm* der Matrix $A = (a_{jk})$.

Ist nun $L < 1$ und $x_0 \in \mathbb{R}^n$ ein beliebiger Startvektor, so konvergiert die durch $x_{k+1} = Ax_k + b$ für $k \in \mathbb{N}_0$ definierte Folge (x_k) gegen eine Lösung a der Gleichung $a = Aa + b$.

Dieses iterative Lösungsverfahren von linearen Gleichungssystemen nennt man in der Numerik auch *Gesamtschrittverfahren*.

23.7 Stetigkeit

Wie früher festgestellt, besagt unsere alltägliche Stetigkeitsvorstellung in etwa, dass bei „stetigen" Veränderungen oder Abläufen keine abrupte, zähe Schwankungen vorkommen. Durch die Stetigkeit einer Funktion (Abbildung) f an einer Stelle a ihres Definitionsbereichs sind die Werte $f(x)$ für nahe bei a gelegene Argumente x (aus dem Definitionsbereich) an den Wert $f(a)$ in gewisser Weise gebunden: Sie weichen wenig von $f(a)$ ab, wenn x um wenig von a abweicht:

Präziser: $f(x)$ weicht *beliebig wenig* von $f(a)$ ab, wenn nur x *hinreichend wenig* von a abweicht.

Das Beispiel der Riemann-Funktion (vgl. §10.2.10(b)) hat uns jedoch gezeigt, dass uns die Anschauung bei Stetigkeitsfragen leicht trügen kann. Die obige qualitative Beschreibung lässt sich auch quantifizieren: Da mann in metrischen Räumen Abstände zwischen den Elementen messen kann (durch die Metrik), lässt sich die obige qualitative Beschreibung in eine quantitative übersetzen, nämlich in die ε-δ-Sprache.

Die Stetigkeit in metrischen Räumen lässt sich auch als „Folgenstetigkeit" charakterisieren (wie in $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$).
Wir betrachten im Folgenden Abbildungen $f : X \to Y$, wobei X und Y beliebige metrische Räume sind. Da die Metriken in X und Y eine Rolle spielen, wollen wir eine Metrik in X mit d_X und eine Metrik in Y mit d_Y bezeichnen.

23.7.1 Satz (Äquivalenzsatz für Stetigkeit)

Seien $X = (X, d_X)$ und $Y = (Y, d_Y)$ metrische Räume und $f : X \to Y$ eine Abbildung und $a \in X$ und $b := f(a)$.
Dann sind folgende Aussagen äquivalent:

(a) (ε-δ-Stetigkeit)
Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in X$ mit $x \in U_\delta(a)$ (d.h. $d_X(x, a) < \delta$) gilt

$$f(x) \in U_\varepsilon(f(a)) \quad \text{(d.h. } d_Y(f(y), f(a)) < \varepsilon) \ ;$$

(b) Für jede ε-Umgebung $U_\varepsilon(f(a))$ von $f(a)$ (in Y) gibt es eine δ-Umgebung $U_\delta(a)$ von a (in X) mit

$$f(U_\delta(a)) \subset U_\varepsilon(f(a)) \ ;$$

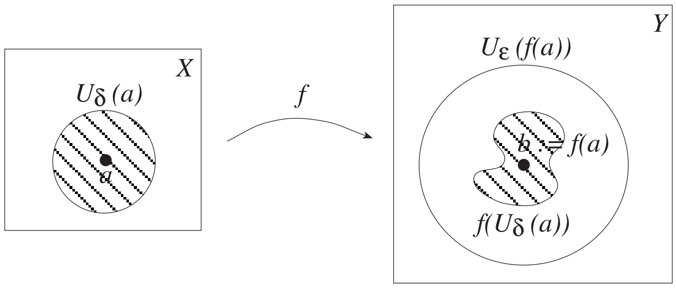
(c) (Umgebungsstetigkeit)
Für jede Umgebung V von $f(a)$ (in Y) gibt es eine Umgebung U von a (in X) mit

$$f(U) \subset V \;;$$

(d) (Folgenstetigkeit)
Für *jede* Folge (x_k) mit $x_k \in X$ und $\lim\limits_{k\to\infty} x_k = a$ konvergiert die Bildfolge $(f(x_k))$ (in Y) gegen $b := f(a)$.

f heißt *stetig in* a (an der Stelle a oder im Punkt a), wenn ein (und damit jede) der Bedingungen (a), (b), (c), (d) erfüllt ist.
f heißt *stetig* (auf X), wenn f in jedem Punkt $a \in X$ stetig ist.

Geometrische Interpretation:



Mit der Vorgabe eines beliebigen $\varepsilon > 0$ wird eine „Toleranzumgebung" $U_\varepsilon(f(a))$ festgelegt, zu der es dann eine δ-Umgebung von a (in X) geben muss, deren Bild in der beliebig vorgegebener ε-Umgebung des Bildpunktes $f(a)$ liegt.

Beweis : Dass (a) und (b) äquivalent sind, ist offensichtlich auf Grund der Definition von ε- bzw. δ-Umgebung.

Wir zeigen: Aus (c) $\Longrightarrow$ (b).

Sei also (c) erfüllt und $\varepsilon > 0$ beliebig vorgegeben. Dann ist $V := U_\varepsilon(f(a))$ eine Umgebung von $f(a)$. Nach Voraussetzung gibt es dann eine Umgebung U von a (in X) mit

$$f(U) \subset V = U_\varepsilon(f(a)) \;.$$

Wieder nach Definition enthält U eine δ-Umgebung von a: $U_\delta(a) \subset U$.
Wegen

$$f(U_\delta(a)) \subset f(U) \subset V = U_\varepsilon(f(a))$$

gilt also

$$f(U_\delta(a)) \subset U_\varepsilon(f(a)) \;.$$

Wir zeigen: Aus (b) $\Longrightarrow$ (c).

Sei umgekehrt die ε-δ-Bedingung erfüllt:

Sei V eine beliebige Umgebung von $b = f(a)$, dann enthält V eine Kugelumgebung $U_\varepsilon(b) \subset V$ von b. Zu dieser ε-Umgebung gibt es nach Voraussetzung eine δ-Umgebung $U_\delta(a)$ von a mit

$$f(U_\delta(a)) \subset U_\varepsilon(b) \subset V\ .$$

Jetzt wähle man nur noch $U := U_\delta(a)$, so sind U bzw. V Umgebungen von a bzw. b mit $f(U) \subset V$.

Wir zeigen: Aus (b) $\Longrightarrow$ (d).

Dazu sei (x_k) eine Folge mit $x_k \in X$ und $\lim\limits_{k\to\infty} x_k = a$. Zu vorgegebenem $\varepsilon > 0$ können wir nach Voraussetzung ein $\delta = \delta(\varepsilon) > 0$, so bestimmen, dass $f(U_\delta(a)) \subset U_\varepsilon(f(a))$ gilt.

Weil aber die Folge (x_k) gegen a konvergiert, liegen fast alle x_k in $U_\delta(a)$:
Es gibt also ein $N = N(\delta) \in \mathbb{N}$, so dass für alle $k \in \mathbb{N}$ mit $k \geq N$ gilt

$$x_k \in U_\delta(a)\ .$$

Dann ist aber wegen $f(U_\delta(a)) \subset U_\varepsilon(f(a))$ für alle $k \geq N$ auch

$$f(x_k) \in U_\varepsilon(f(a))\ ,$$

d.h. die Bildfolge $(f(x_k))$ konvergiert gegen $f(a)$.

Es bleibt noch etwa (d) $\Longrightarrow$ (c) zu zeigen.
Wir zeigen die äquivalente Aussage: Wenn (c) nicht gilt, kann auch (d) nicht gelten.

Wenn (c) nicht gilt, existiert eine „Ausnahmeumgebung“ V_0 von $f(a)$, so dass für *jede* Umgebung U von a gilt

$$f(U) \not\subset V_0\ .$$

Für jedes $k \in \mathbb{N}$ sei $U_k := U_{\frac{1}{k}}(a)$.
Für jedes $k \in \mathbb{N}$ gibt es dann ein $x_k \in U_k \subset X$ mit $f(x_k) \notin V_0$. Die Folge (x_k) konvergiert gegen a, aber kein Bildelement $f(x_k)$ gehört zu V_0. Damit haben wir eine Folge konstruiert, für welche Bedingung (d) nicht erfüllt ist.

Wegen $A \Longrightarrow B \iff \neg B \Longrightarrow \neg A$, ist also (d) $\Longrightarrow$ (c) gezeigt.

□

Der Beweis ist lediglich eine Variante des für $\mathbb{K} = \mathbb{R}$ geführten Beweises (vgl. ???).

23.7.2 Beispiele und Bemerkungen

(a) Das folgende *Unstetigkeitskriterium* ist eine direkte Folge des Äquivalenzsatzes:
$f : X \to Y$ ist nicht stetig in $a \in X$, wenn es eine Folge (x_k) gibt mit $x_k \in X$ und $\lim\limits_{k\to\infty} x_k = a$, für welche die Bildfolge $(f(x_k))$ nicht konvergiert oder für welche die Bildfolge $(f(x_k))$ gegen einen Wert $\neq f(a)$ konvergiert.

(b) Versieht man den Ausgangsraum X mit der diskreten Metrik $d = \delta$ mit

$$\delta(x, y) = \begin{cases} 1, & \text{für } x \neq y, \\ 0, & \text{für } x = y, \end{cases}$$

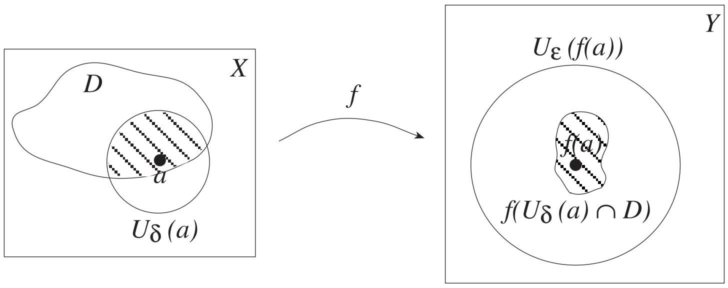
dann ist *jede* Abbildung $f : X \to Y$ (Y beliebiger metrischer Raum) stetig.
Dann ist $a \in X$ und (x_k) eine Folge aus X, die gegen a konvergiert, dann gilt wegen $U_{\frac{1}{2}}(a) = \{a\}$ für fast alle x_k auch $x_k = a$.
Trivialerweise konvergiert dann die Bildfolge $(f(x_k))$ gegen $f(a)$.

(c) Häufig ist eine Abbildung f nicht auf dem ganzen metrischen Raum X definiert, sondern nur auf einer nicht leeren Teilmenge $D \subset X$.

Nun ist aber D mit der von x induzierten Metrik d_0 ($d_0(x,y) = d(x,y)$) für $(x,y) \in D \times D$ ein selbstständiger metrische Raum, so dass man also die äquivalenten Eigenschaften aus Satz 22.7.1 anwenden kann.

Wenn man beachtet, dass die Umgebungen eines Punktes $a \in D$ bezüglich d_0 genau diejenigen Teilmengen U_0 von D sind, die sich in der Gestalt $U_0 = U \cap D$ schreiben lassen, wobei U eine Umgebung von a in X ist, so ist offensichtlich, dass die Stetigkeitsdefinition (entsprechend der Bedingung (b) von Satz 22.8.1) so lautet:

> $f : D \to Y$ ist genau dann stetig in $a \in D$, wenn es zu jeder ε-Umgebung $U_\varepsilon(f(a))$ (in Y) eine δ-Umgebung $U_\delta(a)$ von a (in X) gibt mit
>
> $$f(U_\delta(a) \cap D) \subset U_\varepsilon(f(a)).$$



Mit dieser relativen Situation haben wir schon bei Funktionen einer Variablen gearbeitet.

(d) Die *globale Stetigkeit* einer Abbildung $f : X \to Y$, d.h. f ist stetig auf ganz X, lässt sich elegant charakterisieren mit Hilfe des Begriffes des Urbildes und mit offenen Mengen.

Zur Erinnerung:
Ist $f : X \to Y$ eine Abbildung ($X,\ Y$ beliebige nicht leere Mengen) und $T \subset Y$ eine Teilmenge, dann heißt bekanntlich

$$f^{-1}(T) := \{x \in X;\ f(x) \in T\}$$

das *Urbild* von T unter f.

Offensichtlich gilt für $T_1 \subset T_2 \subset Y$

$$f^{-1}(T_1) \subset f^{-1}(T_2)$$

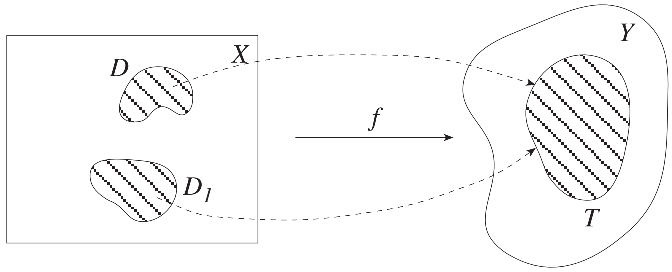
und für $D \subset X$ und $T \subset Y$ gilt:

(i) $D = f^{-1}(T) \Longrightarrow f(D) \subset T$ und
(ii) $T = f(D) \Longrightarrow D \subset f^{-1}(T)$.

Jedoch gilt i.A. nicht

Aus $T = f(D)$ folgt $D = f^{-1}(T)$,

wie das folgende Beispiel zeigt:



Hier ist $f^{-1}(T) = D \cup D_1$.

(c) und (b) lassen such auch so ausdrücken:

$$f(f^{-1}(T)) \subset T \text{ und } D \subset f^{-1}(f(D))$$

Die folgende *Charakterisierung der globalen Stetigkeit* hat den Vorteil, dass sie auch für *beliebige topologische* Räume gilt:

23.7.3 Theorem (Charakterisierung der globalen Stetigkeit)

Sei (X, d_X) und (Y, d_Y) metrische Räume.

(a) Eine Abbildung $f : X \to Y$ ist genau dann stetig, wenn für jede offene Menge V in Y, das Urbild $f^{-1}(V)$ offen in X ist.

(b) Ist $D \subset X$ eine nicht leere Teilmenge, dann ist eine Abbildung $f : D \to Y$ genau dann stetig, wenn für jede offene Teilmenge $V \subset Y$ das Urbild $f^{-1}(V)$ D-offen ist, d.h. es gibt eine offene Menge $U \subset X$ mit
$$f^{-1}(V) = U \cap D.$$

(c) Eine Abbildung $f : X \to Y$ ist genau dann stetig, wenn für jede abgeschlossene Teilmenge $A \subset Y$ das Urbild $f^{-1}(A)$ abgeschlossen in X ist.

(d) Ist $D \subset X$ und $f : D \to Y$ eine Abbildung, dann ist f genau dann stetig, wenn für jede abgeschlossene Teilmenge $A \subset Y$ das Urbild D-abgeschlossen ist, d.h. es gibt eine abgeschlossenen Teilmenge $B \subset X$ mit $f^{-1}(A) = B \cap D$.

Beweis von (a): Sei $f : X \to Y$ stetig und $V \subset Y$ offen und $a \in f^{-1}(V)$. (Wir können $f^{-1}(V) \neq \emptyset$ voraussetzen, denn die leere Menge $\emptyset$ ist offen.) Wir zeigen, dass a ein innerer Punkt von $f^{-1}(V)$ ist.

Sei $b := f(a) \in Y$. Da V offen ist, gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(b) \subset V$. Da f stetig in a ist, gibt es ein $\delta > 0$ mit $f(U_\delta(a)) \subset U_\varepsilon(f(a))$.

Daher ist
$$U_\delta(a) \subset f^{-1}(f(U_\delta(a))) \subset f^{-1}(U_\varepsilon(f(a))) \subset f^{-1}(V),$$

also ist a innerer Punkt von $f^{-1}(V)$.

Da dies für jedes $a \in f^{-1}(V)$ gilt, ist also $f^{-1}(V)$ offen (in X).

Sei umgekehrt $f^{-1}(V)$ offen für jede offene Menge $V \subset Y$. Wir zeigen, dass dann f stetig ist. Dazu sei $a \in X$ und $b = f(a)$.

Für jedes $\varepsilon > 0$ ist die Kugel $V := U_\varepsilon(b)$ offen in Y. Nach Voraussetzung ist dann $f^{-1}(V)$ offen in X. Nun ist aber $a \in f^{-1}(V) = f^{-1}(U_\varepsilon(f(a)))$, daher gibt es ein $\delta > 0$ mit

$$U_\delta(a) \subset f^{-1}(U_\varepsilon(f(a))) \ ,$$

das heißt aber

$$f(U_\delta(a)) \subset f(f^{-1}(U_\varepsilon(f(a)))) \subset U_\varepsilon(f(a)) \ ,$$

das bedeutet f ist stetig in a.

□

Beweis von (b): Der Beweis in der relativen Situation $D \subset X$ ist völlig analog (man muss nur mit der induzierten Metrik d_0 arbeiten).

Beweis von (c): Da für beliebige Teilmengen A von Y gilt

$$f^{-1}(Y - A) = X - f^{-1}(A)$$

und da nach Definition A in Y genau dann abgeschlossen ist, wenn $Y - A$ offen ist, folgt (c) aus (a).

□

Analog beweist man (d).

Anwendung:
Ist (X, d) ein metrischer Raum und $f : X \to \mathbb{R}$ (mit der Standardmetrik) stetig, so gilt für jedes $c \in \mathbb{R}$: die Mengen

(1) $U_1 := \{x \in X;\ f(x) < c\}$ bzw.
$U_2 := \{x \in X;\ f(x) > c\}$ sind offen in X;

(2) $A_1 := \{x \in X;\ f(x) \leq c\}$ bzw.
$A_2 := \{x \in X;\ f(x) \geq c\}$ sind abgeschlossen (in X).

Insbesondere sind Nullstellenmengen $N := \{x \in X;\ f(x) = 0\}$ stetiger Funktionen wegen $N = f^{-1}(\{0\})$ stets abgeschlossen.

Als Übung zeige man: Die orthogonale Gruppe $O(n; \mathbb{R}) = \{A \in \mathbb{R}^{n \times n};\ AA^t = E_n\}$ ist eine *abgeschlossene* Teilmenge des $\mathbb{R}^{n \times n}$ (den man z.B. mit der euklidischen Norm und der daraus abgeleiteten Metrik versieht).
$O(n; \mathbb{R})$ ist auch eine beschränkte Teilmenge von $\mathbb{R}^{n \times n}$ (denn alle Elemente von $A = (a_{jk})$ haben einen Betrag $|a_{jk}| \leq 1$) und damit nach dem Satz von *Heine-Borel* (vgl. §23.13) auch *kompakt*.
$O(n; \mathbb{R})$ ist ein wichtiges Beispiel einer kompakten Gruppe. Sie trägt weitere Strukturen, so ist sie z.B. auch eine *Lie-Gruppe*.

Eine wichtige Klasse stetiger Abbildungen sind sog. *Lipschitz-stetige Abbildungen*.
Dabei heißt eine Abbildung $f : X \to Y$ zwischen metrischen Räumen Lipschitz-stetig (auch beschränkt), wenn es eine Konstante $L \geq 0$ gibt, so dass für alle $x, x' \in X$ gilt

$$\boxed{d_Y(f(x), f(x')) \leq L\ d_X(x, x')}$$

Hier kann man zu vorgegebenem $\varepsilon > 0$ immer ein Universal-δ wählen, nämlich z.B. $\delta := \frac{\varepsilon}{L+1}$, mit welchem die ε-δ-Bedingung erfüllt ist.

23.7.4 Beispiele zur Lipschitz-Stetigkeit

Jede lineare Abbildung $f : \mathbb{K}^n \to \mathbb{K}^m$ $(n, m \in \mathbb{N})$ ist stetig, dabei seien $\mathbb{K}^n$ bzw. $\mathbb{K}^m$ jeweils mit der Maximumsmetrik versehen.

Beweis : Sei $(e_1, \ldots, e_n)$ die Standardbasis von $\mathbb{K}^n$, dann gilt für $x \in \mathbb{K}^n$ bzw, $a \in \mathbb{K}^n$

$$x = \sum_{j=1}^{n} x_j e_j \text{ und } a = \sum_{j=1}^{n} a_j e_j$$

mit (eindeutig bestimmten) Koeffizienten $x_j, a_j \in \mathbb{K},\ 1 \leq j \leq n.$

Aus $x - a = \sum\limits_{j=1}^{n} (x_j - a_j) e_j$ und der Linearität bzw. Eigenschaften von $\| \ \|_\infty$ folgt dann

$$\begin{aligned}
\|f(x) - f(a)\|_\infty &= \|f(x-a)\|_\infty \\
&= \left\| f\left(\sum_{j=1}^{n} (x_j - a_j) \right) e_j \right\|_\infty \\
&= \left\| \sum_{j=1}^{n} (x_j - a_j) f(e_j) \right\|_\infty \\
&\leq \sum_{j=1}^{n} |x_j - a_j| \, \|f(e_j)\|_\infty \\
&\leq \|x - a\|_\infty \cdot \underbrace{\sum_{j=1}^{n} \|f(e_j)\|_\infty}_{=L} \quad \text{(wegen } |x_j - a_j| \leq \|x - a\|_\infty) \\
&= L\|x - a\|_\infty .
\end{aligned}$$

□

Wie wir bald sehen werden, brauchen lineare Abbildungen $f : V \to W$, wobei V und W *beliebige* normierte $\mathbb{K}$-Vektorräume sind, nicht stetig zu sein.

Ist $(V, \| \ \|)$ ein beliebiger normierte $\mathbb{K}$-Vektorraum, dann ist die Norm $\| \ \| : V \to \mathbb{R}$ Lipschitz-stetig mit der Lipschitz-Konstante 1, denn es gilt für beliebige $x, a \in V$

$$|\|x\| - \|a\|| \leq \|x - a\| = 1 \cdot \|x - a\| \ .$$

In Analogie zu Folgen in $\mathbb{R}$ und $\mathbb{C}$ kann man *Rechenregeln für stetige Funktionen* beweisen, die Beweise mit Hilfe des Folgenkriteriums lassen sich im wesentlichen wörtlich übertragen.

Wir fassen uns deshalb kurz:

Seien im Folgenden $X, Y, Y_1, Y_2, \ldots, Y_m$ beliebige metrische Räume und W ein normierter $\mathbb{K}$-Vektorraum (wichtiger Spezialfall ist $W = \mathbb{K}$).

Regel I: Sind $f_1 : X \to W$ und $f_2 : X \to W$ stetig, dann ist auch $f_1 + f_2 : X \to W$ stetig.
Ist ferner $g : X \to \mathbb{K}$ stetig in $a \in X$, dann ist auch $gf : X \to W$ stetig in a und falls $g(a) \neq 0$ ist, ist auch $\frac{f}{g} : X - \{a\} \to W$ stetig in a. Ist insbesondere $g(x) \neq 0$ für alle $x \in X$, dann ist auch $\frac{f}{g} : X \to W$ stetig.

Folgerung:
Alle Polynomfunktionen $P : \mathbb{K}^n \to \mathbb{K}$ und alle rationalen Funktionen sind auf ihrem jeweiligen

Definitionsbereich stetig. Dabei ist eine Polynomfunktion auf $\mathbb{K}^n$ eine Abbildung $P : \mathbb{K}^n \to \mathbb{K}$ der Gestalt

$$x = (x_1, \ldots, x_n) \mapsto \sum_{\substack{k_1,\ldots,k_n \in \mathbb{N}_0 \\ k_1+\ldots+k_n \leq r}} c_{k_1} \ldots c_{k_n} x_1^{k_1} \ldots x_n^{k_n}$$

Eine Abbildung der Gestalt

$$\begin{array}{ccc} \mathbb{K}^n & \to & \mathbb{K} \\ (x_1, \ldots, x_n) & \mapsto & x_1^{k_1} \ldots x_n^{k_n} \end{array}$$

mit natürlichen Zahlen $k_1, \ldots, k_n$ und $k_1 + \ldots + k_n = r$ heißt ein Monom vom Grad r.
Ein Polynom ist also eine Linearkombination von Monomen vom Grad $\leq r$.
Eine rationale Funktion auf eine offene Teilmenge $D \subset \mathbb{K}^n$ ist eine Funktion, die als Quotient von Polynomfunktionen auf D darstellbar ist (dabei muss man natürlich die Nullstellen der Nennerpolynoms berücksichtigen). Da eine Polynomfunktion durch endlich viele Additionen und Multiplikationen aus der Koordinatenfunktion $x_1, \ldots, x_n$ und den Konstanten entsteht, folgt die Behauptung durch mehrfache Anwendung von Regel 1. Ebenso folgt die Behauptung über rationale Funktionen.

Beispiele:

(a) Die Abbildung $p : \mathbb{R}^{n+1} - \{0\} \to \mathbb{R}^{n+1}$ mit

$$p(x) = \frac{x}{\|x\|_2} := \frac{1}{\|x\|_2} \cdot x$$

ist stetig. Ihr Bild ist die euklidische n-Sphäre $S^n = \{x \in \mathbb{R}^{n+1};\ \|x\|_2 = 1\}$.

(b) Die Polynomfunktion

$$\begin{array}{rcl} p:\ \mathbb{R}^2 & \to & \mathbb{R}\,, \\ (x, y) & \mapsto & ax^2 + by^2 + cxy + dx + ey + f \end{array}$$

$(a, b, \ldots, f \in \mathbb{R})$ ist stetig.

(c) Die rationale Funktion

$$\begin{array}{rcl} R:\ \mathbb{R}^2 & \to & \mathbb{R}\,, \\ (x, y) & \mapsto & \dfrac{x^2 + y^2 - 1}{x^2 + y^2 + 1} \end{array}$$

ist stetig.

Nicht verwunderlich ist die folgende

Regel II: Ist $f : X \to Y$ stetig in a und $g : Y \to Z$ stetig in $b \in Y$ und ist $f(X) \subset Y$ und $b = f(a)$, dann ist $g \circ f : X \to Z$ stetig in a.

Kurz: Die Zusammensetzung stetiger Funktionen ist stetig.

Wir wiederholen den Beweis (mit dem Folgenkriterium):
Sei (x_k) eine Folge mit $x_k \in X$ uns $\lim\limits_{k\to\infty} x_k = a$. Dann gilt wegen der Stetigkeit von f in a:

$$\lim_{k\to\infty} f(x_k) = f(a) = b\,.$$

Wegen der Stetigkeit von g in b folgt

$$\lim_{k\to\infty} g(f(x_k)) = g(b) = g(f(a))\,, \text{ also}$$

$$\lim_{k\to\infty}(g\circ f)(x_k) = (g\circ f)(a)\ .$$

Übungsaufgabe: Geben Sie einen weiteren Beweis mit Hilfe der Umgebungscharakterisierung der Stetigkeit.

Beispiel:
Sei $X = (\mathbb{K}^n, \|\ \|_2)$. Dann ist die Abbildung

$$\begin{aligned} h:\ X &\ \rightarrow\ \mathbb{R}\ , \\ x &\ \mapsto\ \exp(\|x\|_2^2) = \exp(x_1^2 + \ldots + x_n^2) \end{aligned}$$

als Zusammensetzung zweier stetiger Abbildungen:

$$\begin{aligned} f:\ X &\ \rightarrow\ \mathbb{R}\ , \\ x &\ \mapsto\ \|x\|_2^2 \end{aligned}$$

und

$$\begin{aligned} g:\ \mathbb{R} &\ \rightarrow\ \mathbb{R}\ , \\ y &\ \mapsto\ \exp y \end{aligned}$$

stetig.

Vor der nächsten Regel eine Vorbemerkung:
Sind (X,d) und $(Y_1,d_1),\ldots,(Y_m,d_m)$ metrische Räume und ist

$$f : X \rightarrow Y := Y_1\times\ldots\times Y_m$$

eine gegebene Abbildung, so kann man diese in „Komponenten“ $f_j :\rightarrow X \rightarrow Y_j$ zerlegen, für die dann gilt $f = (f_1,\ldots,f_m)$.
Jede Komponente kann man auffassen als Zusammensetzung $pr_j\circ f$ von f mit der Projektion

$$\begin{aligned} pr_j:\ Y := Y_1\times\ldots\times Y_m &\ \rightarrow\ Y_j\ , \\ (y_1,\ldots,y_m) &\ \mapsto\ y_j\ . \end{aligned}$$

Man erhält ein kommutatives Diagramm (von Abbildungen)

$$\begin{tikzcd} f : X \arrow[r] \arrow[dr, "f_j = pr_j \circ f"'] & Y := Y_1\times\ldots\times Y_j\times\ldots\times Y_m \arrow[d, "pr_j"] \\ & Y_j \end{tikzcd}$$

Da eine Folge (y_k) mit $y_k\in Y = Y_1\times\ldots\times Y_m$ genau dann konvergiert, wenn jede Komponentenfolge in Y_j konvergiert, sind die Projektionen $pr_j : Y_1\times\ldots\times Y_m \rightarrow Y_j$ alle stetig ($1\le j\le m$).

Sind umgekehrt für $j\in\{1,\ldots,n\}$ Abbildungen $f_j : X\rightarrow Y_j$ gegeben, so kann man diese durch

$$f := (f_1,\ldots,f_m) : X\rightarrow Y := Y_1\times\ldots\times Y_m$$

zu einer Abbildung von X nach Y zusammenfassen.

Regel III:

(a) Sind X und $Y_1,\ldots,Y_m$ metrische Räume und ist $Y := Y_1\times\ldots\times Y_m$ ihr Produkt (versehen mit der Produkt-Metrik), dann ist eine Abbildung

$$f = (f_1,\ldots,f_m) : X\rightarrow Y := Y_1\times\ldots\times Y_m$$

genau dann stetig in $a\in X$, wenn alle Komponenten $f_j : X\rightarrow Y$ stetig in a sind.

(b) Ist insbesondere jedes $Y_j = \mathbb{K}$ $(1 \leq j \leq m)$, dann ist eine Abbildung

$$f = f(f_1, \ldots, f_m) : X \to \mathbb{K}^m$$

genau dann stetig in a, wenn jede ´Komponentenfunktion $f_j : X \to \mathbb{K}$ $(1 \leq j \leq m)$ stetig in a ist.

Der Beweis ist nach der Vorbemerkung klar, denn die Konvergenz von $f(x_k) = (f_1(x_k), \ldots, f_m(x_k))$ gegen $(f_1(a), \ldots, f_m(a))$ ist äquivalent mit der komponentenweisen Konvergenz.

Die Teilaussage (b) kann man als *Reduktionslemma* betrachten: Bei der Untersuchung stetiger Abbildungen $f = (f_1, \ldots, f_m) : X \to \mathbb{K}^m$ (mit einer der Standardmetriken), kann man sich immer auf den Fall $m = 1$ zurückziehen.
Ist $f : X \to Y$ stetig und ist $D \subset X$, dann ist natürlich $f|D : D \to Y$ auch stetig.
Man könnte auf die Idee kommen, dass z.B. eine Abbildung (Funktion) $f : \mathbb{R}^2 \to \mathbb{R}$ bereits schon stetig ist in einem Punkt $a \in \mathbb{R}^2$, wenn ihre Einschränkungen auf die achsenparallelen Geraden durch den Punkt stetig sind. Dass es nicht so ist, zeigt das folgende

Gegenbeispiel: $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x,y) = \begin{cases} \frac{2xy}{x^2+y^2}, & \text{für } (x,y) \neq (0,0)\ ; \\ 0, & \text{für } (x,y) = (0,0)\ . \end{cases}$$

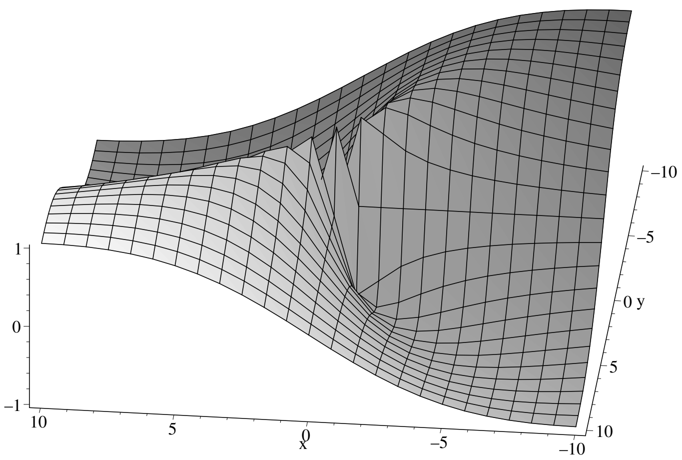
Für $(x,y) \neq (0,0)$ ist f als rationale Funktion stetig. Auch sind die Einschränkungen von f auf $\mathbb{R} \times \{0\}$ und $\{0\} \times \mathbb{R}$ stetig:

$$\begin{array}{rcl} x & \mapsto & f(x,0) = 0 \ \text{ und} \\ y & \mapsto & f(0,y) = 0 \end{array}$$

f ist aber im Nullpunkt nicht stetig, denn es ist $f(t,t) - f(0,0) = 1$ für alle $t \neq 0$. Etwas anders ausgedrückt:
Die Folge $(x_k, y_k) := (\frac{1}{k}, \frac{1}{k})$ konvergiert zwar gegen Null, die Bildfolge $f(x_k, y_k) = \frac{2\frac{1}{k}\frac{1}{k}}{\frac{1}{k^2}+\frac{1}{k^2}} = 1$ hat den konstanten Wert 1, also gilt

$$\lim_{k\to\infty} f(x_k, y_k) = 1 \neq 0 = f(0,0)\ .$$

Die Bildfolge konvergiert zwar, aber nicht gegen den Funktionswert.
Plottet man mit MAPLE den Graphen von f in einer Umgebung der Nullpunktes, so kann man die Unstetigkeit von f im $(0,0)$ erkennen.
Das Maximum von f hat den Wert 1, das Minimum den Wert -1, diese Werte werden für alle $(x,y) \in \mathbb{R}^2$ mit $x = y$ bzw. $x = -y$ angenommen.



Man könnte auf die Idee kommen, dass auch die Bilder offener bzw. abgeschlossener Mengen unter einer stetigen Abbildung wieder offen bzw. abgeschlossen sind.
Aber schon aus der Analysis 1 sind *Gegenbeispiele* bekannt:

(a)

$$\begin{aligned} f :]0, 2\pi[&\to \mathbb{R}\,, \\ x &\mapsto \sin x \end{aligned}$$

bildet das offene Intervall $]0, 2\pi[$ auf das abgeschlossene Intervall $[-1, 1]$ ab.

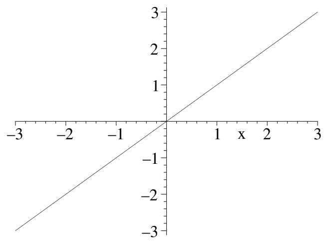
(b)

$$\begin{aligned} f : \ [1, \infty] &\to \mathbb{R}\,, \\ x &\mapsto \frac{1}{x} \end{aligned}$$

bildet die im $X_0 := [1, \infty[$ abgeschlossene Mengen $[1, \infty[$ auf die in $\mathbb{R}$ nicht abgeschlossene Menge $]0, 1]$ ab.

(c) Weiteres Gegenbeispiel: Wir betrachten

$$\begin{aligned} f : \ \mathbb{R} &\to \mathbb{R}^2\,, \\ t &\mapsto (t, t) \end{aligned}$$



bildet $\mathbb{R}$ auf die Gerade G mit der Gleichung $y = x$ ab.
Daher ist für keine offene Menge $U \subset \mathbb{R}$, $U \neq \emptyset$, das Bild $f(U)$ offen in $\mathbb{R}^2$.

(d) Wir nehmen an, dass (X, d_X) und $(Y.d_Y)$ metrische Räume sind und dass die Abbildung $f : X \to Y$ bijektiv ist. Dann existiert die Umkehrabbildung $g : Y \to X$, und es gilt

$$\begin{aligned} g(f(x)) = x \qquad & \text{für alle } x \in X \text{ und} \\ f(g(y)) = y \qquad & \text{für alle } y \in Y. \end{aligned}$$

Ist nun $U \subset X$ eine Teilmenge, so stimmt das Bild $f(U)$ mit dem Urbild $g^{-1}(U)$ überein:

$$\begin{aligned} g^{-1}(U) &= \{y \in Y;\ g(y) \in U\} \\ &= \{y \in Y;\ x := g(y) \in U\} \\ &= \{f(x);\ x \in U\} \\ &= f(U)\,. \end{aligned}$$

Daher ist die Umkehrabbildung g von f genau dann stetig auf Y, wenn das Bild jeder offenen Teilmenge $U \subset X$ unter f wieder offen (in Y) ist.

23.8 Definition

Sind (X, d_X) und (Y, d_Y) metrische Räume.
Eine stetige, bijektive Abbildung $f : X \to Y$, für welche die Umkehrabbildung $g : Y \to X$ ebenfalls stetig ist, heißt ein *Homöomorphismus*.

Eine bijektive Abbildung $f : X \to Y$ ist also genau dann ein Homöomorphismus, wenn für eine beliebige Teilmenge $U \subset X$ gilt:

$$U \text{ offen (in } X) \iff f(U) \text{ offen (in } Y).$$

23.8.1 Beispiele und Bemerkungen

(a) Mann nennt zwei metrische Räume X und Y *homöomorph*, wenn es einen Homöomorphismus $f : X \to Y$ gibt.

Die Begriffsbildung kann man auch sofort auf *topologische Räume* übertragen.

Es lassen sich dann topologische Aussagen über Teilmengen von X mittels eines Homöomorphismus übersetzen in entsprechende Aussagen über Teilmengen von Y. Man sagt auch, dass X und Y dieselbe *topologische Struktur* haben.

(b) Betrachtet man $\mathbb{R}$ mit der natürlichen Metrik. Dann sind die Abbildungen

$$\tan : \left]-\frac{\pi}{2},\ \frac{\pi}{2}\right[\to \mathbb{R} \qquad \text{und} \qquad \arctan : \mathbb{R} \to \left]-\frac{\pi}{2},\ \frac{\pi}{2}\right[$$

jeweils stetig und Umkehrabbildungen voneinander. Daher ist $\mathbb{R}$ homöomorph zum Intervall $]-\frac{\pi}{2},\ \frac{\pi}{2}[$.

Frage: Können Sie einen Homöomorphismus $f : \mathbb{R} \to]-1, 1[$ angeben?

(c) Wir betrachten im $\mathbb{R}^n$ die euklidische Metrik $\|\ \|_2$ und die Abbildung

$$\begin{array}{rcl} f :\ \mathbb{R}^n & \to & U_1(0) := \{x \in \mathbb{R}^n;\ \|x\|_2 < 1\}\ , \\ x & \mapsto & \dfrac{x}{1 + \|x\|_2} \end{array}$$

bzw.

$$\begin{array}{rcl} g :\ U_1(0) & \to & \mathbb{R}^n\ , \\ y & \mapsto & \dfrac{y}{1 - \|y\|_2} \end{array}$$

Beide Abbildungen sind stetig und Umkehrungen voneinander. Daher sind der ganze $\mathbb{R}^n$ und die (offene) euklidische Einheitskugel $U_1(0)$ homöomorph.

Testfragen:

(α) Gilt dies auch, wenn $\|\ \|$ eine beliebige Norm auf $\mathbb{R}^n$ ist?

(β) Sind die (offenen) Einheitskugeln bezüglich euklidischer Norm, der Maximumsnorm bzw. der Eins-Norm in $\mathbb{R}^n$ homöomorph?

(d) Wir identifizieren den $\mathbb{R}^3$ mit $\mathbb{C} \times \mathbb{R}$ und identifizieren $\mathbb{C}$ mit $\mathbb{C} \times \{0\} \subset \mathbb{R}^3$ und schreiben die für (euklidische) Einheitssphäre in $\mathbb{R}^3$ als

$$S^2 := \{(w, t) \in \mathbb{C} \times \mathbb{R};\ |w|^2 + t^2 = 1\}\ .$$

Ferner sei $N := (0,1) \in \mathbb{R}^2$ der „Nordpol“ der Sphäre S^2.

Dann werden durch

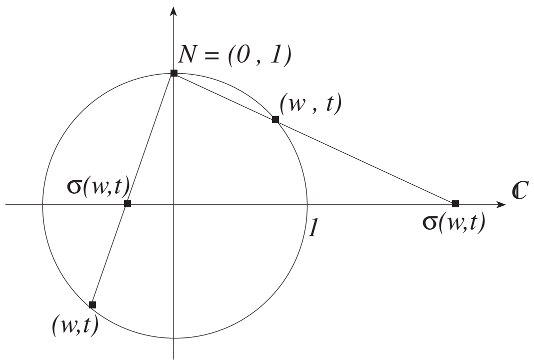
$$\begin{array}{rcl} \sigma: \ S^2 - \{N\} & \to & \mathbb{C}\ , \\ (w,t) & \mapsto & \dfrac{w}{1-t} \end{array}$$

und

$$\begin{array}{rcl} \tau: \ \mathbb{C} & \to & S^2 - \{N\}\ , \\ z & \mapsto & \left(\dfrac{2z}{|z|^2+1}, \dfrac{|z|^2-1}{|z|^2+1}\right) \end{array}$$

stetige Abbildungen definiert, die Umkehrungen voneinander sind. Man rechne dies nach.

Daher sind die im Nordpol punktierte Sphäre $S^2 - \{N\}$ und die komplexe Zahlenebene $\mathbb{C}$ homöomorph.



Geometrische Veranschaulichung

σ nennt man auch *stereographische Projektion*.

Frage: Wenn der Punkt $(w,t) \in S^2$ in den Nordpol „wandert“, wohin wandert dann der Bildpunkt $\sigma(w,t) \in \mathbb{C}$?

(e) Eine stetige, bijektive Abbildung $f: X \to Y$ braucht keine stetige Umkehrabbildung zu haben:

Dazu betrachten wir $X := \{(t,0) \in \mathbb{R}^2;\ 0 \le t \le 1\}$ und die Einheitskreislinie $S^1 := \{(x,y) \in \mathbb{R}^2;\ x^2+y^2 = 1\}$ (jeweils mit der von $\mathbb{R}^2$ induzierten Metrik) und die Abbildung

$$\begin{array}{rcl} f: \ X & \to & Y\ , \\ (t,0) & \mapsto & (\cos 2\pi t,\ \sin 2\pi t). \end{array}$$

Dann ist f stetig und bijektiv, also existiert die Umkehrabbildung $g: S^2 \to X$. Diese ist jedoch unstetig im Punkt $(1,0)$. Denn betrachtet man eine Folge $z_k \in S^1;\ Im\ z_k > 0$ und $\lim\limits_{k\to\infty} z_k = 1 + 0i = (1,0)$ und die Folge $w_k \in S^1$ mit $Im\ w_k > 0$ und $\lim\limits_{k\to\infty} w_k = 1$, dann konvergieren die Bildfolgen $(g(z_k))$ und $(g(w_k))$ gegen 0 bzw. 1.

Die Tatsache, dass f zwar bijektiv und stetig aber kein Homöomorphismus ist, schließt nicht aus, dass es vielleicht doch einen Homöomorphismus $h: [0,1[\to S^1$ geben könnte.

Mit Hilfe des *Kompaktheitsbegriffes* können wir das aber ausschließen (vgl. ???). Dort werden wir auch ein sehr einfaches und brauchbares *Homöomorphiekriterium* kennen lernen.

Wir wissen, dass eine Folge (f_k) von *stetigen* Funktionen f_k mit einem Intervall $D \subset \mathbb{R}$ als Definitionsbereich, die gleichmäßig auf D konvergiert, eine stetige Grenzfunktion f auf D besitzt.

Wir wollen diesen Satz verallgemeinern und benötigen dazu den Begriff der *gleichmäßigen Konvergenz in metrischen Räumen.*

Seien $X = (X, d_X)$ und $Y = (Y, d_Y)$ metrische Räume, $f_k : X \to Y,\ k \in \mathbb{N}$, und $f : X \to Y$ Abbildungen.
Man sagt: Die Folge (f_k) konvergiert gleichmäßig gegen f, falls zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass für alle $x \in X$ und alle $k \geq N$ gilt

$$d_Y(f_n(x), f(x)) < \varepsilon\ .$$

23.9 Satz

Sind (X, d_X) und (Y, d_Y) metrische Räume und ist (f_k) eine Folge stetiger Funktionen, $f_k : X \to Y$, die gleichmäßig gegen die Funktion $f : X \to Y$ konvergiert, dann ist auch f stetig.

Beweis : (Analog zu Analysis 1, ???)

Wir zeigen die Stetigkeit von f in einem beliebigen Punkt $a \in X$. Die Schlüsselungleichung ist die folgende:

$$d_Y(f(x), f(a)) \leq d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(a)) + d_Y(f_N(a), f(a))\ .$$

Die beiden äußeren Terme werden klein etwa $< \frac{\varepsilon}{2}$, wegen der gleichmäßigen Konvergenz (N hinreichend groß), der mittlere Term (hier ist N fest) wir ebenfalls $< \frac{\varepsilon}{2}$ für alle $x \in X$ mit $x \in U_\delta(a),\ \delta > 0$ geeignet, (wegen der Stetigkeit von f_N in a).

□

Bemerkung:
Ist $Y = \mathbb{K}$ (dabei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) und

$$B(X, \mathbb{K}) = \{f : X \to \mathbb{K};\ \text{Es gibt } c \geq 0 \text{ mit } |f(x)| \leq c \text{ für alle } x \in X\}$$

und $\|f\|_\infty = \sup\{|f(x)|;\ x \in X\}$, dann konvergiert eine Folge $(f_k),\ f_k : X \to \mathbb{K}$ genau dann gleichmäßig gegen $f : X \to \mathbb{K}$, wenn $\lim\limits_{k\to\infty} \|f_k - f\|_\infty = 0$ gilt.

23.10 Stetigkeit linearer Abbildungen, die Operatornorm

Im Beispiel 22.7.4(a) haben wir gesehen, dass jede *lineare* Abbildung $f : \mathbb{K}^n \to \mathbb{K}^m$ $(n, m \in \mathbb{N})$ stetig, sogar Lipschitz-stetig ist, wenn man $\mathbb{K}^n$ bzw. $\mathbb{K}^m$ mit irgendeiner der Standardnormen versieht. Eine lineare Abbildung zwischen normierten Räumen braucht aber nicht stetig zu sein. Zum Beispiel ist die Abbildung

$$\begin{aligned} D:\ \mathcal{C}^1[0,1] &\to \mathbb{R}\,, \\ f &\mapsto f'(0) \end{aligned}$$

linear, aber nicht stetig, wenn man $\mathcal{C}^1([0,1])$ mit der Supremums-Norm und $\mathbb{R}$ mit der natürlichen Norm (dem Betrag) versieht, denn die Supremumsnormen von $f_k(x) = \frac{\sin(k^2 x)}{k}$ bilden eine Nullfolge, während die Folge $(f_k'(0))$ der Ableitungen divergiert.

Wir beweisen zunächst für *lineare* Abbildungen zwischen beliebigen normierten $\mathbb{K}$-Vektorräumen $(V, \| \ \|_V)$ und $(W, \| \ \|_W)$ den folgenden

23.10.1 Äquivalenzsatz für Stetigkeit

Seinen $(V, \| \ \|_V)$ und $(W, \| \ \|_W)$ normierte $\mathbb{K}$-Vektorräume und $u : V \to W$ $\mathbb{K}$-linear.
Dann sind äquivalent

(1) u ist (überall) stetig;

(2) u ist stetig in $0 \in V$;

(3) Es gibt eine Konstante $C \geq 0$ mit $\|u(x)\|_W \leq C\|x\|_V$ für alle $x \in V$.

Hieraus folgt, dass eine stetige lineare Abbildung $u : V \to W$ sogar Lipschitz-stetig ist.

Beweis : Wir zeigen $(2) \Longrightarrow (3) \Longrightarrow (1)$.

Da $(1) \Longrightarrow (2)$ evident ist, sind damit alle drei Aussagen äquivalent.

(a) $(2) \Longrightarrow (3)$
u sei stetig in $0 \in V$. Dann existiert insbesondere zu $\varepsilon := 1$ ein $\delta > 0$ mit

$$\|u(x) - u(0)\|_W = \|u(x)\|_W < 1$$

für alle $x \in V$ mit $\|x - 0\|_V = \|x\|_V < \delta$.
Für beliebige $y \in V,\ y \neq 0$, setzen wir $x := \frac{\delta}{2} \cdot \frac{y}{\|y\|_V}$, dann gilt $\|x\|_V < \delta$, und es folgt

$$1 > \|u(x)\|_W = \frac{\delta}{2} \cdot \frac{1}{\|y\|_V} \|u(y)\|_W,$$

oder

$$\|u(y)\|_W < \frac{2}{\delta} \|y\|_V.$$

Lassen wir jetzt auch $y = 0$ zur Konkurenz zu, so sehen wir wegen $u(0) = 0$, dass wir $C := \frac{2}{\delta}$ wählen können.

(b) $(3) \Longrightarrow (1)$.
Aus (3) folgt für $x := y - z$ wegen der Linearität von u

$$\|u(y) - u(z)\|_W \leq C\|y - z\|_V.$$

Wenn daher $\varepsilon > 0$ beliebig vorgegeben ist, so hat man für $\delta := \frac{\varepsilon}{C+1}$

$$\|y - z\|_V < \delta = \frac{\varepsilon}{C + 1} \quad \Longrightarrow \quad \|u(y) - u(z)\|_W < \varepsilon.$$

Damit haben wir auch die Lipschitz-Stetigkeit (und damit die gleichmäßige Stetigkeit(!)) von u gezeigt.

□

23.10.2 Beispiele und Bemerkungen

(a) Wegen der Eigenschaft (3) ist die Stetigkeit einer linearen Abbildung $u : V \to W$ gleichbedeutend mit der Existenz einer Konstanten für welche $\|u(x)\|_W \leq C\|x\|_V$ für alle $x \in V$ gilt.

Lineare Abbildungen mit der Eigenschaft (3) nennt man manchmal auch *beschränkt*. Man beachte jedoch, dass diese Terminologie nicht sonderlich glücklich ist, da nach bisheriger Sprachregelung eine Abbildung $u : V \to W$ beschränkt heißt, wenn ihr Bild $u(V)$ in W beschränkt ist. Ist $W \neq \{0\}$, so ist für eine beschränkte lineare Abbildung $u : V \to W,\ u \neq 0$, das Bild i.A. nicht beschränkt. Beispiel?

(b) Für die Standardvektorräume $\mathbb{K}^n$ und $\mathbb{K}^m$ mit einer der Standardnormen erhält man nochmal etwas anders die Stetigkeit einer linearen Abbildung $u : \mathbb{K}^n \to \mathbb{K}^m$, denn u ist ein m-Tupel von linearen Abbildungen (Linearformen)

$$\begin{array}{rcll} u_j : \qquad \mathbb{K}^n & \to & \mathbb{K} & \qquad (1 \leq j \leq n)\ , \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \mapsto & a_{j1} + a_{j2} + \ldots + a_{jn} \in \mathbb{K}\ . & \end{array}$$

(c) Ist $M = [a,b] \subset \mathbb{R}$ ein kompaktes Intervall und $V = \mathcal{C}([a,b])$ der $\mathbb{R}$-Vektorraum der stetigen Funktionen $f : [a,b] \to \mathbb{R}$ mit der Supremums-Norm, so zeigt die Standardabschätzung für Integrale

$$|I(f)| = \left| \int\limits_a^b f(x)dx \right| \leq (b-a)\|f\|_\infty\ ,$$

dass das Integral

$$\begin{array}{rcl} I :\ V & \to & \mathbb{R}\ , \\ f & \mapsto & I(f) := \int\limits_a^b f(x)dx \end{array}$$

eine *stetige* Linearform ist.

(d) Für normierte $\mathbb{K}$-Vektorräume $(V, \|\ \|_V)$ und $(W, \|\ \|_W)$ sei

$$L(V,W) := \{u : V \to W;\ u\ \mathbb{K}\text{-linear und stetig}\ \}\ .$$

Man verifiziert ohne Mühe, dass $L(V,W)$ bezüglich der punktweisen Verknüpfung ein $\mathbb{K}$-Vektorraum ist. Es ist unser Ziel, in diesem $\mathbb{K}$-Vektorraum eine (natürliche) Norm einzuführen. Wir beachten zunächst, dass im Spezialfall $V = \mathbb{K}^n$ und $W = \mathbb{K}^m$ und der Standardnormen gilt

$$L(\mathbb{K}^n,\ \mathbb{K}^m) = Hom(\mathbb{K}^n, \mathbb{K}^m) \cong \mathbb{K}^{m \times n} = \mathbb{K}^{mn},$$

eine lineare Abbildung kann man also in diesem Fall mit einer $m \times n$-Matrix identifizieren und damit mit einem Element aus $\mathbb{K}^{mn}$.

Wir benutzen den Äquivalenzsatz um auf $L(V,W)$ eine Norm einzuführen.

23.10.3 Definition (Abbildungsnorm oder Operatornorm)

Seien $(V, \|\ \|_V)$ und $(W, \|\ \|_W)$ normierte $\mathbb{K}$-Vektorräume und $u \in L(V,W)$. Unter der *Operatornorm* von u versteht man die Zahl

$$\|u\| := \inf\{C \geq 0;\ \|u(x)\|_W \leq C\|x\|_V \text{ für alle } x \in V\}\ .$$

23.10.4 Beispiele und Bemerkungen

(a) Auf Grund des Äquivalenzsatzes ist die Menge, deren Infimum wir betrachten, nicht leer und durch Null nach unten beschränkt, die Definition also sinnvoll ist und es gilt stets $\|u\| \geq 0$.

Wir werden bald sehen, dass durch

$$\begin{array}{rcl} L(V,W) & \to & \mathbb{R} \\ u & \mapsto & \|u\| \end{array}$$

tatsächlich eine Norm auf $L(V,W)$ erklärt ist. Mit der Operatornorm wird also $L(V,W)$ zu einem normierten und damit auch metrischen Raum.

(b) Man beachte, dass $\|u\|$ sowohl von der Norm auf V als auch von der Norm auf W abhängt, was wir in der Bezeichnung nicht ausdrücklich hervorgehoben haben.

Für den Beispiel (vgl. 22.10.2(b))

$$\begin{array}{rcl} I:\ \mathcal{C}([a,b]) & \to & \mathbb{R}\,, \\ f & \mapsto & I(f) = \int\limits_a^b f(x)dx \end{array}$$

ist die Operatornorm offensichtlich gerade $b-a$.

(c) Nach Definition des Infimums gibt es zu jedem $\varepsilon > 0$ eine Konstante $C \geq 0$ mit $\|u(x)\|_W \leq C\|x\|_V$ für alle $x \in V$, so dass $C < \|u\| + \varepsilon$ gilt. Hieraus folgt für alle $x \in V$

$$\|u(x)\|_W \leq \|u\|\ \|x\|_V.$$

Die Operatornorm ist also charakterisiert durch:

Für alle $x \in V$ *gilt* $\|u(x)\|_W \leq \|u\|\ \|x\|_V$ *und ist*
$\|u(x)\|_W \leq C\|x\|_V$ *für alle* $x \in V$, *so ist* $\|u\| \leq C$.

u ist also ein Maß für die maximale Längenverzerrung, die ein Element $x \in V$ unter der Wirkung von u erleiden kann.

Um zu sehen, dass durch $u \to \|u\|$ wirklich eine Norm auf $L(V,W)$ definiert ist, benutzen wir den folgenden

(d) **Hilfssatz:**
Für $u \in L(V,W)$ gilt

$$\begin{array}{rcl} \|u\| & = & \sup\{\|u(x)\|_W;\ \|x\|_V \leq 1\} \\ & = & \sup\{\|u(x)\|_W;\ \|x\|_V = 1\}\ . \end{array}$$

Beweis : Nach Definition von $\|u\|$ gilt für $\|x\|_V \le 1$

$$\|u(x)\|_W \le \|u\| \; \|x\|_V \le \|u\| \; .$$

Daher ist $s := \sup\{\|u(x)\|_W ; \; \|x\| \le 1\}$ endlich und es ist $s \le \|u\|$.
Nun ist $\{x \in V ; \; \|x\|_V = 1\} \subset \{x \in V ; \; \|x\|_V \le 1\}$, also ist auch

$$t := \sup\{\|u(x)\|_W ; \; \|x\|_V = 1\}$$

endlich und es gilt $t \le s$.

Es bleibt $\|u\| \le t$ zu zeigen. Nach Definition von t gilt

$$\|x\|_V = 1 \quad \Longrightarrow \quad \|u(x)\|_W \le t \; .$$

Durch Übergang von $y \neq 0$ zu $\frac{1}{\|y\|_V}\|y\|$ folgt dann $\|u(y)\|_W \le t\|y\|_V$ für alle $y \in V$, d.h. $\|u\| \le t$ nach Definition von $\|u\|$.

□

(e) Wir zeigen jetzt, dass

$$\begin{aligned} \| \; \| : \; L(V,W) \quad &\to \quad \mathbb{R} \; , \\ u \quad &\mapsto \quad \|u\| \end{aligned}$$

tatsächlich eine Norm ist.

(N_1): $\|u\| \ge 0$ gilt nach Definition.
Ist $\|u\| = 0$, dann ist $u(x) = 0$ für alle $x \in V$ mit $\|x\|_V \le 1$ nach dem Hilfssatz, also auch $u(y) = 0$ für alle $y \in V$, d.h. u ist die Nullabbildung.

(N_2): $\|\lambda u\| = |\lambda| \; \|u\|$ für $\lambda \in \mathbb{K}$ ist offensichtlich.

(N_3): Ist auch $v \in L(V,W)$, so gilt für $\|x\|_V \le 1$

$$\begin{aligned} \|(u+v)(x)\|_W \quad &\le \quad \|u(x)\|_W + \|v(x)\|_W \\ &\le \quad \|u\| \; \|x\|_V + \|v\| \; \|x\|_V \le \|u\| + \|v\| \; , \end{aligned}$$

also

$$\|u+v\| \le \|u\| + \|v\| \; .$$

(f) Sind $(V_1, \| \; \|_{V_1})$, $(V_2, \| \; \|_{V_2})$ und $(V_3, \| \; \|_{V_3})$ normierte $\mathbb{K}$-Vektorräume und $u \in L(V_1, V_2)$ und $v \in L(V_2, V_3)$, so ist $v \circ u \in L(V_1, V_3)$ und es gilt

$$\|v \circ u\| \le \|v\| \; \|u\| \; .$$

Denn es gelten für $x \in V_1$ die folgenden Abschätzungen

$$\|v(u(x))\|_{V_3} \le \|v\| \; \|u(x)\|_{V_2} \le \|v\| \; \|u\| \; \|x\|_{V_1} \; .$$

(g) Für $V = W$ sei $L(V) := L(V, V)$.
Aus unseren Überlegungen folgt, dass $L(V)$ eine normierte Algebra ist mit $\|id_V\| = 1$.

(h) In $L(V, W)$ hat man zwei natürliche Konvegenzbegriffe:

(α) Die punktweise Konvergenz:
Eine Folge $(u_k),\ u_k \in L(V, W)$ konvergiert punktweise gegen $u \in L(V, W)$, wenn für alle $x \in V$ gilt $\lim\limits_{k\to\infty} u_k(x) = u(x)$ (Konvergenz in W).

(β) Konvergenz in der Norm (gleichmäßige Konvergenz):
(u_k) konvergiert in der Norm gegen $u \iff \lim\limits_{k\to\infty} \|u_k - u\| = 0$.

Wegen $\|u_k(x) - u(x)\|_W = \|(u_k - u)(x)\|_W \leq \|u_k - u\|\|x\|_V$ folgt aus der gleichmäßigen Konvergenz die punktweise Konvergenz.

(i) Wir wissen, dass $L(V, W)$ ein normierter Raum ist (bezüglich der Operatornorm).
Man kann sich fragen: Wann ist $L(V, W)$ ein Banach-Raum?

Übungsaufgabe:
Ist W ein Banach-Raum, ist auch $L(V, W)$ ein Banach-Raum.
Insbesondere ist $L(V, \mathbb{K})$ ein Banach-Raum und ist V ein Banach-Raum, so ist $L(V)$ eine Banach-Algebra.

Tipp:
Man zeige der Reihe nach

(α) Ist (u_k) eine Cauchy-Folge im $L(V, W)$, dann ist für jedes $x \in V$ $(u_k(x))$ eine Cauchy-Folge in W, also konvergent. Es gibt also ein eindeutig bestimmtes $y \in W$ mit $\lim\limits_{k\to\infty} u_k(x) = y$.
Durch

$$\begin{aligned} u: V &\to W\,, \\ x &\mapsto \lim_{k\to\infty} u_k(x) \end{aligned}$$

wird eine lineare Abbildung definiert.

(β) Als Cauchy-Folge ist (u_k) beschränkt, d.h. es gibt ein $C \geq 0$, so dass für alle $x \in V$ gilt $\|u_k(x)\|_W \leq C\|x\|_V$, hieraus folgt $(k \to \infty)$ $\|u(x)\|_W \leq C\|x\|_V$, d.h. $u \in L(V, W)$.

(γ) (u_k) konvergiert in $L(V, W)$ gegen u.

(j) Die Operatornorm ist i.A. schwierig zu berechnen. Nimmt man auf $\mathbb{K}^n$ und $\mathbb{K}^m$ jeweils die Maximumsnorm und ist

$$u := A \in L(\mathbb{K}^n, \mathbb{K}^m) = Hom(\mathbb{K}^n, \mathbb{K}^m) = \mathbb{K}^{m\times n},$$

also eine $m \times n$-Matrix, so ist $\|A\| = \max\limits_{1\leq j\leq n} \left\{ \sum\limits_{k=1}^{n} |a_{jk}| \right\}$ die sog. *Zeilensummennorm von A*.

Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$ mit $|x_k| \leq 1$ gilt

$$\|Ax\|_\infty = \max_{1\leq j\leq n} \left|\sum_{k=1}^n a_{jk}x_k\right| \leq \max_{1\leq j\leq n} \left(\sum_{k=1}^n |a_{jk}|\right) =: M \ .$$

Der Wert M wird angenommen:
Wir wählen j_0 so, dass $M = \sum\limits_{k=1}^n |a_{j_o k}|$ gilt und setzen $\xi_k := \frac{|a_{j_0 k}|}{a_{j_0 k}}$, falls $a_{j_0 k} \neq 0$ ist und $\xi_k = 1$ sonst.

Dann gilt für $\xi := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$, $\|\xi\|_\infty = 1$ und es ist

$$\|A\xi\|_\infty = M \ .$$

Eine weitere wichtige Norm auf $\mathbb{K}^{m\times n}$ ist die Hilbert-Schmidt-Norm (auch Frobenius-Norm genannt):

$$\|A\|_{H.S.} = \sqrt{\sum_{j=1}^m \sum_{k=1}^n |a_{jk}|^2}.$$

Das ist die euklidische Norm des Elementes

$$(a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{m1}, \ldots, a_{mn}) \in \mathbb{K}^{mn} \ .$$

Da –wie wir bald zeigen werden– auf einem endlich-dimensionalen normierten $\mathbb{K}$-Vektorraum alle Normen äquivalent sind, ist es gleichgültig mit welcher Norm man im Raum

$$L(\mathbb{K}^n, \mathbb{K}^m) = Hom(\mathbb{K}^n, \mathbb{K}^m) \cong \mathbb{K}^{m\times n} \cong \mathbb{K}^{mn}$$

arbeitet.

Da in $\mathbb{K}^{m\times n}$ alle Normen äquivalent sind, muss es –falls $\|A\|$ eine beliebige Operatornorm von $A \in \mathbb{K}^{m\times n}$ bezeichnet- Konstanten $c_1,\ c_2$ bzw. C_1 und C_2 geben, so dass gilt

$$\begin{array}{rcl} c_1\|A\|_{H.S.} \leq & \|A\| & \leq c_2\|A\|_{H.S.} \quad \text{bzw.} \\ C_1 \max\limits_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}} \{|a_{jk}|\} \leq & \|A\| & \leq C_2 \max\limits_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}} \{|a_{jk}|\} \end{array}$$

Konvergenz in $\mathbb{K}^{m\times n} \cong \mathbb{K}^{mn}$ hätte man auch ohne den Begriff des Abbildungsraums definieren können, denn Konvergenz in $\mathbb{K}^{m\times n} \cong \mathbb{K}^{mn}$ ist äquivalent mit der komponentenweisen Konvergenz.

Halten wir fest:
Die Konvergenz einer Folge $(A^{(k)}) = (a_{jl}^{(k)})$ von Matrizen $A^{(k)} \in \mathbb{K}^{m\times n}$ gegen eine Matrix $A = (a_{jl}) \in \mathbb{K}^{m\times n}$ ist äquivalent damit, dass für jedes $(j,l) \in (\{1,\ldots,m\} \times \{1,\ldots,m\})$ gilt

$$\boxed{\lim_{k\to\infty} a_{jl}^{(k)} = a_{jl}}$$

(das ist Konvergenz in $\mathbb{K}$).

Die Algebra $\mathbb{K}^{n\times n}$ der $n \times n$-Matrizen mit Einträgen aus $\mathbb{K}$ ist eine Banach-Algebra (sogar bezüglich irgendeiner Norm auf $\mathbb{K}^{n\times n}$).

Ist $z \in \mathbb{C}$ und $|z| < 1$, dann ist bekanntlich $1 - z$ invertierbar in $\mathbb{C}$ und

$$(1 - z)^{-1} = \sum_{k=0}^{\infty} z^k \quad \text{(geometrische Reihe).}$$

Eine Verallgemeinerung ist die folgende

23.10.5 Bemerkung

Ist $(\mathcal{A}, \| \ \|)$ eine Banach-Algebra mit Einselement $1_{\mathcal{A}}$ (z.B. $\mathcal{A} = \mathbb{K}^{n \times n}$) und $A \in \mathcal{A}$ mit $\|A\| < 1$, dann ist $1_{\mathcal{A}} - A$ invertierbar (in $\mathcal{A}$), uns es gilt mit $G_{\mathcal{A}}(A) := \sum\limits_{k=0}^{\infty} A^k$ (geometrische Reihe)

$$(1_{\mathcal{A}} - A)G_{\mathcal{A}}(A) = G_{\mathcal{A}}(A)(1_{\mathcal{A}} - A) = 1 \qquad \text{(sog. Neumansche Identität).}$$

Der Beweis ist offensichtlich wegen z.B.

$$(1_{\mathcal{A}} - A)G_{\mathcal{A}}(A) = \sum_{k=0}^{\infty} A^k - \sum_{k=0}^{\infty} A^{k+1} = A^0 = 1_{\mathcal{A}} \ .$$

23.11 Exkurs: Die Matrix-Exponential-Abbildung

Bei der Lösung von Systemen linearer Differentialgleichungen (aber nicht nur dort) ist es nützlich, die Matrix-Exponential-Funktion $A \mapsto e^A := \exp(A)$ für $A \in \mathbb{K}^{n \times n}$ zur Verfügung zu haben.

Wir skizzieren einen Zugang.
Zunächst kann man in jedem (normierten) $\mathbb{K}$-Vektorraum $(V, \| \ \|)$ den Begriff der *Reihe* (von Elementen = Vektoren aus V) einführen.
Dazu sei $(a_k)_{k \in \mathbb{N}_0}$ eine Folge von Elementen $a_k \in V$. Dieser Folge kann man eine neue Folge $(s_m)_{m \in \mathbb{N}_0}$ mit $s_m := \sum\limits_{k=0}^{m} a_k$ zuordnen.
(s_m) heißt die *Folge der Partialsummen* der Folge (a_k) oder die der Folge (a_k) *zugeordnete Reihe*. Wie im Fall $V = \mathbb{R}$ oder $V = \mathbb{K}$ zeigt man:
Die Abbildung

$$\begin{aligned} \sum : \ \mathit{Abb}(\mathbb{N}_0, V) &\rightarrow \mathit{Abb}(\mathbb{N}_0, V) \ , \\ (a_k) &\mapsto (s_m) = \left(\sum_{k=0}^{m} a_k \right) \end{aligned}$$

ist bijektiv.

Jede Reihe ist eine spezielle Folge und jede Folge lässt sich auch als Reihe auffassen. Die Surjektivität der Abbildung $\sum$ benutzt der sog. „Teleskop-Trick“.
Ist nun die Folge (s_m) der Partialsummen von (a_k) konvergent, dann nennt man ihren Grenzwert (das ist ein Element aus V)

$$s := \lim_{m \to \infty} s_m = \lim_{m \to \infty} \sum_{k=0}^{m} a_k$$

auch den *Wert* (die Summe) der Reihe (s_m) und schreibt dafür auch

$$s := \sum_{k=0}^{\infty} a_k \ .$$

Leider wird auch hier das Symbol $\sum a_k$ in zweierlei Bedeutung verwendet:

1. Einmal ist es ein Synonym für die Folge $(s_m) = \left(\sum\limits_{k=0}^{m} a_k\right)$, die Folge der Partialsummen (insbesondere, wenn die Konvergenzfrage noch offen ist),
2. zweitens bedeutet es (im Fall der Konvergenz von (s_m)) auch den Grenzwert von (s_m), also ein Element aus V. Welche der beiden Bedeutungen gemeint ist, ist meist aus dem Zusammenhang klar.

Ein einfaches *hinreichendes* Kriterium für die Konvergenz von (s_m) ist die Konvergenz der (Absolut–) Reihe (σ_m) mit $\sigma_m = \sum\limits_{k=0}^{m} \|a_k\|$, vorausgesetzt, der normierte Raum $(V, \| \ \|)$ ist ein Banach-Raum.
(σ_m) ist natürlich eine Folge(Reihe) in $\mathbb{R}$ und damit genau dann konvergent, wenn sie eine Cauchy-Folge (in $\mathbb{R}$) ist:
Zu jedem $\varepsilon > 0$ gibt es also einen Index $N \in \mathbb{N}$, so dass für alle $m, l \in \mathbb{N}$ mit $m > l \geq N$ gilt

$$\begin{aligned} |\sigma_m - \sigma_n| &= \left| \sum_{k=0}^{m} \|a_k\| - \sum_{k=0}^{l} \|a_k\| \right| \\ &= \left| \sum_{k=l+1}^{m} \|a_k\| \right| = \sum_{k=l+1}^{m} \|a_k\| < \varepsilon \ . \end{aligned}$$

Nach der Dreiecksungleichung ist dann aber auch für alle $m > l \geq N$

$$\|s_m - s_l\| = \left\| \sum_{k=l+1}^{m} a_k \right\| \leq \sum_{k=l+1}^{m} \|a_k\| < \varepsilon \ ,$$

d.h. die Folge (s_m) ist eine Cauchy-Folge in $(V, \| \ \|)$ und daher konvergent, da V ein Banach-Raum ist.
Man vergleiche hierzu den Beweis aus Analysis 1, dass aus der *absoluten Konvergenz* einer Reihe in $\mathbb{R}$ oder $\mathbb{C}$ die Konvergenz der Reihe selber folgt.

Wir wollen unsere Überlegungen spezifizieren auf die Banach-Algebra $\mathbb{K}^{n\times n}$ der $n \times n$-Matrizen versehen mit *irgendeiner Operator-Norm*.

Da alle Normen auf dem Vektorraum $\mathbb{K}^{n\times n}$ äquivalent sind, gibt es insbesondere eine positive reelle Zahl c_1, so dass für $A = (a_{jk}) \in \mathbb{K}^{n\times n}$ und irgendeine Operator-Norm $\| \ \|$ auf $\mathbb{K}^{n\times n}$ gilt

$$\|A\| \leq c_1 \max_{\substack{1 \leq j \leq n \\ 1 \leq k \leq n}} \{|a_{jk}|\} =: c$$

23.11.1 Theorem (Definition zu $\exp(A)$ für $A \in \mathbb{K}^{n\times n}$)

Für $A \in \mathbb{K}^{n\times n}$ sei $\|A\| \leq c$ (mit einer nicht-negativen Konstante c, $\| \ \|$ sei irgendeine Operator-Norm auf $\mathbb{K}^{n\times n}$). Dann ist die Reihe $\sum\limits_{k=0}^{\infty} \frac{1}{k!} A^k$ konvergent (in $(\mathbb{K}^{n\times n}, \| \ \|)$) und man definiert

$$e^A := \exp(A) := \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Die Abbildung (Matrix-Exponential-Abbildung)

$$\begin{aligned} \mathbb{K}^{n\times n} &\to \mathbb{K}^{n\times n} \\ A &\mapsto \exp(A) \end{aligned}$$

hat **folgende Eigenschaften**:

(a) Für $A, B \in \mathbb{K}^{n\times n}$ mit $AB = BA$ gilt

$$\boxed{\exp(A+B) = \exp(A)\exp(B)}$$

(b) Für jede $A \in \mathbb{K}^{n\times n}$ ist $\exp(A)$ invertierbar und

$$(\exp(A))^{-1} = \exp(-A)$$

(c) Ist $\Lambda = Diag(\lambda_1, \ldots, \lambda_n)$ eine Diagonalmatrix mit Diagonalelementen $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, dann ist für $k \in \mathbb{N}_0$

$$\Lambda^k = Diag(\lambda_1^k, \ldots, \lambda_n^k) \text{ und}$$

$$\exp(\Lambda) = \sum_{k=0}^{\infty} \frac{1}{k!}\Lambda^k = Diag(\exp(\lambda_1), \ldots, \exp(\lambda_n))$$

(d) Für $I := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ und $E := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und $t \in \mathbb{K}$ ist

$$\begin{aligned}
\exp(tI) &= \left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \frac{t^6}{6!} \pm\right) E + \left(t - \frac{t^3}{3!} + \frac{t^5}{5!} \mp\right) I \\
&= \cos(t)E + \sin(t)I \\
&= \begin{pmatrix} \cos(t) & 0 \\ 0 & \cos(t) \end{pmatrix} + \begin{pmatrix} 0 & -\sin(t) \\ \sin(t) & 0 \end{pmatrix} \\
&= \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.
\end{aligned}$$

Beweisskizze:
Die Konvergenz von $\sum_{k=0}^{\infty} \frac{1}{k!}A^k$ ist klar wegen $\|A^k\| \leq \|A\|^k = c^k$ und der Konvergenz von $\sum_{k=0}^{\infty} \frac{1}{k!}c^k$ $(= \exp(c))$.
Die Funktionalgleichung beweist man im Fall der Vertauschbarkeit von A und B wie die Funktionalgleichung der *gewöhnlichen* $\exp$-Funktion.
Aus der Funktionalgleichung folgt die Invertierbarkeit von $\exp(a)$ wegen

$$\exp(A)\exp(-A) = \exp(-A)\exp(A) = \exp(0) = E\ .$$

Wegen $I^{2k} = (-1)^k E$ ist die Formel für $\exp(tI)$ klar.
Da $\exp(tI)$ die „Drehmatrix“ $\begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}$ ist, nennt man die Matrix I auch eine „infinitesimale Erzeugende“ der Drehgruppe.

23.11.2 Bemerkungen und Aufgaben

(a) Ist $AB \neq BA$, so ist i.A. $\exp(A+B) \neq \exp(A)\exp(B)$, wie das Beispiel $A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ und $B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ zeigt:

Hier ist

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{und} \\ BA &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{und} \\ \exp(A)\exp(B) &= \begin{pmatrix} e & e \\ 0 & 1 \end{pmatrix} \quad \text{und} \\ \exp(A+B) &= \begin{pmatrix} e & e-1 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

(b) Für eine beliebige Matrix $A \in \mathbb{C}^{n\times n}$ und $S \in Gl(n,\mathbb{C})$ zeige man

(α) $S^{-1} \cdot \exp(A) \cdot S = \exp(S^{-1}AS)$

(β) $det(\exp(A)) = \exp(Spur(A))$

(c) Durch

$$\begin{aligned} \gamma: \mathbb{K} &\to Gl(n,\mathbb{K}) \quad (A \in \mathbb{K}^{n\times n}), \\ t &\mapsto \exp(tA) \end{aligned}$$

wird ein Homomorphismus definiert.

Wir werden bald sehen, dass diese Abbildung sogar differenzierbar ist und dass

$$\dot\gamma(t) = A\exp(tA)$$

gilt.

24 Kompaktheit

Zur Erinnerung: Wir hatten eine Teilmenge $K \subset \mathbb{R}$ oder $K \subset \mathbb{C}$ *kompakt* genannt, wenn sie abgeschlossen und beschränkt ist. Dabei bedeutet „abgeschlossen“:
Der Grenzwert jeder konvergenten Folge $(a_k),\ a_k \in K$, liegt wieder in K (diese Eigenschaft ist die sog. „Folgenabgeschlossenheit“). In der Zwischenzeit wissen wir, dass dieser Begriff der „Folgenabgeschlossenheit“ mit dem Abgeschlossenheitsbegriff in metrischen Räumen X äquivalent ist (d.h. $X - K$ ist offen).

Der Begriff der Beschränktheit ist klar:
Es gibt ein $C \geq 0$, so dass für alle $z \in K$ gilt $|z| \leq C$.

Ferner haben wir die *Bolzano-Weierstraß-Charakterisierung* des Begriffs „kompakt“ kennengelernt:
Eine Teilmenge $K \subset \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) ist genau dann kompakt, wenn jede Folge von Elementen aus K eine Teilfolge besitzt, die gegen einen Punkt aus K konvergiert.

Übungsaufgabe: Wiederholen Sie den Beweis.

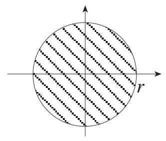
Wir werden sehen, dass für metrische Räume die Begriffe „überdeckungskompakt“ und „folgenkompakt“ äquivalent sind.

Typische Beispiele kompakter Teilmengen sind

(a) In Fall $K = \mathbb{R}$ „kompakte Intervalle“, also Intervalle vom Typ $[a, b],\ a, b \in \mathbb{R},\ a \leq b$.

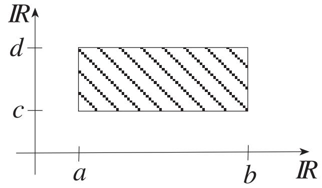
(b) Im Fall $K = \mathbb{C}$ $(= \mathbb{R}^2)$ kompakte Kreisscheiben

$$K := \{z \in \mathbb{C};\ |z| \leq r\},\ (r \geq 0)$$



oder allgemeine Kreisscheiben $\overline{U_r}(a) = \{z \in \mathbb{C};\ |z - a| \leq r\};\ a \in \mathbb{C},\ r \geq 0$, oder auch Produkte von kompakten Intervallen

$$[a, b] \times [c, d];\quad a, b, c, d \in \mathbb{R};\ a \leq b,\ c \leq d$$



Ferner wissen wir:
Eine stetige Funktion $f : [a, b] \to \mathbb{R}$ hat auf $[a, b]$ ein (globales) Maximum und ein (globales) Minimum, und ist insbesondere beschränkt auf dem kompakten Intervall $[a, b]$.

Wir wollen in diesem Abschnitt den Begriff des kompakten Intervalls so verallgemeinern, dass die zu definierenden kompakten metrischen Räume X wieder die Eigenschaft haben, dass z.B. jede stetige Funktion $f : X \to \mathbb{R}$ ein (globales) Maximum bzw. (globales) Minimum auf X besitzt.
Viele Verfahren zur Berechnung von Extremstellen sind erst dann legitim anwendbar, wenn die Existenz solcher Extremstellen feststeht. Eine Analyse der Beweise, in denen die Existenz von Maxima und Minima eingeht, führt auf den Begriff der (Überdeckungs-)Kompaktheit. Es ist zentral für die gesamte Mathematik.

Für die Standardmetriken werden wir jedoch zeigen, dass eine Teilmenge $K \subset \mathbb{K}^n$ genau dann (überdeckungs -) kompakt ist, wenn sie beschränkt und abgeschlossen ist (das ist das Theorem von Heine-Borel).

Im $\mathbb{K}^n$ stimmen also der bisher verwendete Kompaktheitsbegriff (kompakt = abgeschlossen und beschränkt) mit dem nun zu definierenden Kompaktheitsbegriff (= Überdeckungskompaktheit) überein.

Wir beschäftigen uns zunächst mit dem Begriff der „Überdeckung“.

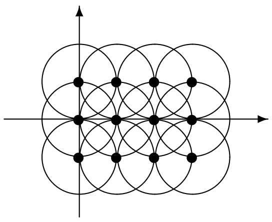
Dazu zunächst einige Beispiele:

(a) Betrachtet man in $\mathbb{R}^2$ (mit der euklidischen Metrik) alle offenen Kreisscheiben $U_1(p)$ vom Radius 1 und dem Mittelpunkt $p := (m,n) \in \mathbb{Z} \times \mathbb{Z}$, dann liegt jeder Punkt $z = (x,y) \in \mathbb{R} \times \mathbb{R}$ in mindestens einer dieser Kreisscheiben, oder äquivalent:

Die Vereinigung aller dieser Kreisscheiben überdeckt den ganzen $\mathbb{R}^2$:

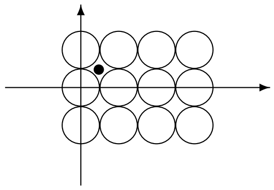
$$\mathbb{R}^2 \subset \bigcup_{p \in \mathbb{Z}\times\mathbb{Z}} U_1(p)$$

(es gilt sogar $\mathbb{R}^2 = \bigcup\limits_{p \in \mathbb{Z}\times\mathbb{Z}} U_1(p)$)



(b) Nimmt man jedoch alle Kreisscheiben mit ganzzahligen Mittelpunkten und jeweils Radius $\frac{1}{2}$, dann gibt es Punkt $z = (x,y) \in \mathbb{R}^2$, die in keiner dieser Kreisscheiben liegen, z.B. liegt der Punkt $\left(\frac{1}{2}, \frac{1}{2}\right) \in \mathbb{R}^2$ in keiner dieser (offenen) Kreisscheiben.

Die Kreisscheiben $U_{\frac{1}{2}}(p),\ p \in \mathbb{Z} \times \mathbb{Z}$, bilden keine Überdeckung von $\mathbb{R}^2$.



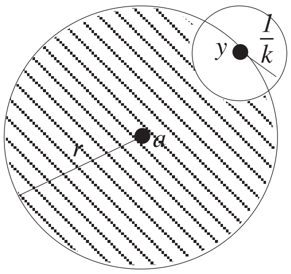
(c) Wir betrachten etwa allgemeiner eine offene Kugel $U_r(a)$ in $\mathbb{R}^n$ (mit der euklidischen Metrik) und einen festen Punkt $y \in \mathbb{R}^n$ mit $d(y,a) = r$ (also einen Randpunkt der Kugel).

Für $k \in \mathbb{N}$ setzen wir

$$U_k = \left\{ z \in U_r(a);\ d(z,y) > \frac{1}{k} \right\}.$$

Wir erhalten so ein System offener Mengen $U_k \subset U_r(a)$. Da für $z \in U_r(a)$ stets $d(z,y) > 0$ gilt, bilden die U_k eine offene Überdeckung von $U_r(a)$.

Es ist auch evident, das nicht schon endlich viele U_k ausreichen, um $U_r(a)$ zu überdecken, denn andernfalls hätte man $U_r(a) = U_k$ für ein geeignetes $k \in \mathbb{N}$, da die U_k eine aufsteigende Folge von Teilmengen von $U_r(a)$ bilden.

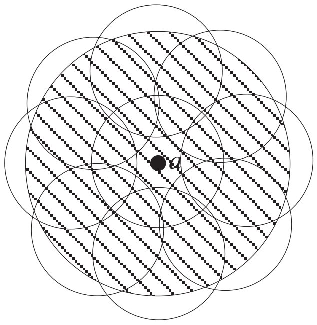


(d) Man erhält eine weitere Überdeckung von $U_r(a)$, indem man zu jedem $z \in U_r(a)$ die offene Kugel $U_z := U_{\frac{r}{2}}(z)$ wählt.

Offensichtlich gilt $U_r(a) \cup \bigcup\limits_{z \in U_r(a)} U_z$.

Hier braucht man nicht alle U_z um $U_r(a)$ zu überdecken: Es reichen sicher endlich viele geeignete U_z aus.

Diese Vorüberlegung sollen die folgenden Begriffsbildungen motivieren:



24.1 Definition (Überdeckung)

Sei (X, d) ein metrischer Raum und $D \subset X$. Ein System von Teilmengen $U_j,\ j \in J$ (J beliebige Indexmenge $\neq 0$) heißt *Überdeckung* von D, wenn

$$D \subset \bigcup_{j \in J} U_j$$

gilt, wenn also jeder Punkt $x \in D$ in mindestens einer Teilmenge U_j liegt.

Die Überdeckung heißt *offen*, wenn jede Menge U_j eine offene Teilmenge von X ist.
Ist $(U_j)_{j \in J}$ eine Überdeckung von D und $J_0 \subset J$ und gilt schon

$$D \subset \bigcup_{j \in J_0} U_j\ ,$$

dann heißt das System $(U_j)_{j \in J_0}$ eine *Teilüberdeckung* von $(U_j)_{j \in J}$.
Ist dabei J_0 eine *endliche* Menge, so spricht man von einer *endlichen* Teilüberdeckung. Unsere Beispiele (d) und (c) zeigen, dass es bei einer vorgegebenen Menge D (hier $D = U_r(a)$) sowohl offene Überdeckungen $(U_j)_{j \in J}$ gibt, bei denen endlich viele U_j ausreichen (d), um D zu überdecken, als auch offene Überdeckungen, bei denen das nicht der Fall ist (c).

Allgemein wollen wir eine Teilmenge $K \subset X$ *kompakt* nennen, wenn *jede* (wohlgemerkt jede!) offene Überdeckung von K eine endliche Teilüberdeckung enthält, d.h. präziser

24.2 Definition ((Überdeckungs-) Kompakt)

Sei (X, d) ein metrischer Raum. Eine Teilmenge $K \subset X$ heißt *kompakt*, wenn es zu *jeder* Überdeckung $(U_j)_{j \in J}$ von K durch (in X) offene Mengen U_j eine endliche Teilüberdeckung gibt, d.h. es gibt eine endliche Teilmenge $J_0 := \{j_1, \ldots, j_r\} \subset J$ mit

$$K \subset \bigcup_{k=1}^{r} U_{j_k} \qquad \left(\text{oder } K \subset \bigcup_{j \in J_0} U_j\right)$$

24.3 Beispiele und Bemerkungen

(a) Die Definition besagt *nicht*, dass K kompakt ist, wenn K eine endliche Überdeckung durch offene Mengen besitzt (eine solche gibt es immer! Wieso?), sondern, dass bei einer *beliebigen* Überdeckung durch offene Mengen bereits immer endlich viele ausreichen, um K zu überdecken!

(b) Eine offene Kreisscheibe $U_r(a)$ in $\mathbb{R}^n$ (mit der euklidischen Metrik) ist nicht kompakt (siehe Beispiel (c)).

(c) Sind $x_1, \ldots, x_r$ endlich viele (verschiedene) Punkte eines metrischen Raums X , so ist $K := \{x_1, \ldots, x_r\}$ kompakt. Beweis ist offensichtlich.

(d) Ist (X, d) ein metrischer Raum und ist (x_k) eine konvergente Folge in X und

$$a = \lim_{k \to \infty} x_k \ ,$$

dann ist die Menge

$$K := \{x_k;\ k \in \mathbb{N}\} \cup \{a\} \subset X$$

kompakt.

Beweis : Sei $(U_j)_{j \in J}$ eine beliebige offene Überdeckung von K. Da $a \in K$ gilt, gibt es ein $j^* \in J$ mit $a \in U_{j^*}$.

Da U_{j^*} offen ist, ist U_{j^*} eine Umgebung von a, in der fast alle Glieder der Folge (a_k) liegen: Es gibt also ein $N \in \mathbb{N}$, so dass für alle $k \geq N$ gilt $a_k \in U_{j^*}$.

Andererseits liegt jedes Folgenglied x_k in einem geeigneten U_{j_k} $(j_k \in J)$. Offensichtlich gilt dann bereits

$$K \subset U_{j_1} \cup U_{j_2} \cup \ldots \cup U_{j_{N-1}} \cup U_{j^*} \ ,$$

d.h. wir haben eine endliche Teilüberdeckung konstruiert.

□

(e) Lässt man in (d) den Grenzwert in der Definition von K weg, so wird die Behauptung falsch.

Zeigen Sie dies am Beispiel $X = \mathbb{R}$ und der Folge $(a_k) = \left(\frac{1}{k}\right)$ mit dem Grenzwert 0. Zeigen Sie also

$$M := \left\{\frac{1}{k}; k \in \mathbb{N}\right\}$$

ist keine kompakte Teilmenge von $\mathbb{R}$.

Tipp: Betrachten Sie dazu die offene Überdeckung von M mit

$$U_1 := \left]\frac{1}{2},\ 2\right[; \quad U_k := \left]\frac{1}{k+1},\ \frac{1}{k-1}\right[; \quad k \geq 2.$$

Das Beispiel zeigt auch:

Um zu zeigen, dass eine Teilmenge $M \subset X$ **nicht** kompakt ist, genügt es *eine* offene Überdeckung von M anzugeben, aus der man keine endliche Teilüberdeckung auswählen kann.

Weitere wichtige Klassen kompakter Mengen werden wir in den nachfolgenden Sätzen kennenlernen.

(f) Ist ein metrischer Raum X kompakt, so ist häufig folgende „Kompaktheitsschluss“ genannte Schlussweise möglich:

E bezeichne eine Eigenschaft oder einen Sachverhalt, der für jede offenen Teilmenge U von X entweder gilt oder nicht gilt. Überträgt sich E von U und V (V ebenfalls offen) auf $U \cup V$ und gibt es zu jeden Punkt $x \in X$ eine (wenn auch noch so kleine) offene Umgebung U_x, welche die Eigenschaft E ebenfalls hat, dann hat auch ganz X die Eigenschaft E.

Der Beweis ist einfach: Mittels Induktion zeigt man, dass E für die Vereinigung von endlich vielen offenen Mengen U aus X gilt und damit gilt E wegen

$$X = U_{x_1} \cup \ldots \cup U_{x_r}$$

mit geeigneten $x_j \in X$ auch für X.

Dieser „Kompaktheitsschluss" erlaubt es häufig, von lokalen Sachverhalten auf globale Sachverhalte zu schließen.

(α) Eine Funktion $f : X \to \mathbb{R}$ (X metrischer Raum) heißt *lokal beschränkt*, wenn es zu jedem $x \in X$ eine oBdA offene Umgebung U_x gibt und eine Zahl $c_x \in \mathbb{R}$ mit

$$|f(x)| \leq c_x \qquad \text{für alle } x \in U_x \ .$$

Eine lokale beschränkte Funktion braucht nicht beschränkt zu sein (so ist etwa die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ lokal beschränkt, aber nicht beschränkt).

Es gilt jedoch: Ist X kompakt und $f : X \to \mathbb{K}$ lokal beschränkt, so ist f beschränkt, denn dann ist für geeignete $x_1, \ldots, x_r \in X$

$$X = U_{x_1} \cup \ldots \cup U_{x_r}$$

und daher gilt für alle $x \in X$

$$|f(x)| \leq c := \max\{|c_{x_1}|, \ldots, |c_{x_r}|\} \ .$$

(β) Ist (f_k) eine Folge von Funktionen $f_k : X \to \mathbb{R}$ und konvergiert (f_k) lokal gleichmäßig auf X, d.h. gibt es zu jedem $x \in X$ eine oBdA offene Umgebung U_x, so dass $(f_k|U_x)$ gleichmäßig konvergiert und ist X kompakt, so konvergiert (f_k) auf ganz X (also global) gleichmäßig.

(g) Kompaktheit einer Teilmenge $K \subset X$ ist eine sogenannte „innere Eigenschaft" der Teilmenge K, d.h. man kann sie prüfen, ohne K zu verlassen, d.h. mit den K-offenen Mengen:

Eine Teilmenge $K \subset X$ ist genau dann kompakt, wenn K bezüglich der induzierten Metrik kompakt ist, d.h. wenn jede Überdeckung von K durch K-offene Teilmengen eine endliche Teilüberdeckung enthält.
Die Kompaktheit einer Teilmenge $K \subset X$ lässt sich also allein durch die induzierte Metrik entscheiden (die Lage von K in X spielt dabei keine Rolle – im Gegensatz zum Begriff der Abgeschlossenheit).

Der Beweis ist klar, da die K-offenen Mengen die Gestalt $U \cap K$ mit U offen in X haben, bedeutet

$$\begin{aligned} K &= \bigcup_{j \in J} (U_j \cap K) \quad \text{nichts anderes als} \\ K &\subset \bigcup_{j \in J} U_j \ . \end{aligned}$$

Wir bevorzugen jedoch die bequemere ursprüngliche Definition.

(h) Die Definition der Kompaktheit in *topologischen Räumen* ist identisch mit der Definition 23.2.

Die Überdeckungskompaktheit wird allerdings häufig auch als „Quasi-Kompaktheit" bezeichnet und kompakt heißt dann quasi-kompakt-hausdorffsch. Die Terminologie ist leider nicht einheitlich. Da sich unsere Betrachtungen auf metrische Räume beziehen, ist die Hausdorff-Eigenschaft immer automatisch erfüllt.

Die folgenden Sätze dienen zur Konstruktion und Kennzeichnung kompakter Mengen in metrischen Räumen. Sie erlauben häufig durch einen „einfachen Blick“ zu entscheiden, ob eine vorgegebene Menge K kompakt ist oder nicht. Da man in der Definition der Kompaktheit *jede* offene Überdeckung zu testen hat, ob sie eine endliche Teilüberdeckung enthält, sind solche Kompaktheitskriterien ausgesprochen nützlich.

24.4 Satz

Ist (X, d) ein metrischer Raum und $K \subset X$ kompakt, dann ist K abgeschlossen (in X) und beschränkt.
Beschränkt bedeutet dabei: K ist in einer geeigneten Kugel $U_r(a)$ enthalten ($a \in X,\ r > 0$).
Beweis: Abgeschlossenheit.
Da X stets abgeschlossen ist, können wir $K \neq X$ voraussetzen.
Um nachzuweisen, dass K abgeschlossen ist, müssen wir zeigen, dass das Komplement $X - K$ offen ist. Sei also $p \in X - K$ ein beliebiger Punkt. Wir konstruieren eine ε-Umgebung $U_\varepsilon(p)$ mit $U_\varepsilon(p) \in X - K$. Dazu setzen wir für $k \in \mathbb{N}$

$$U_k := \left\{\varepsilon \in X;\ d(y,p) > \frac{1}{k}\right\} .$$

Dann ist U_k offen und es gilt

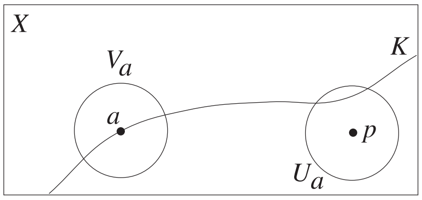
$$\bigcup_{k=1}^{\infty} U_k = X - \{p\} \supset K .$$

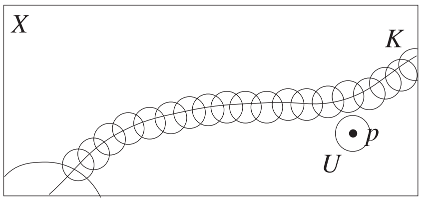
Es gibt daher wegen der Kompaktheit von K endlich viele $k_1, \ldots, k_r$ mit

$$K \subset \bigcup_{j=1}^{r} U_{k_j} .$$

Setzt man $N := \max\{k_1, \ldots, k_r\}$ und $\varepsilon := \frac{1}{N}$, so ist $U_\varepsilon(p) \subset X - K$, d.h. p ist innerer Punkt von $X - K$. K ist also abgeschlossen.

Wir geben einen Alternativbeweis für die Abgeschlossenheit, der die Hausdorff-Eigenschaft eines metrischen Raumes benutzt und auch für kompakte Teilmengen von Hausdorff-Räumen funktioniert.
Sei also wieder $p \in X - K$. Wegen der Hausdorff-Eigenschaft können wir in X zu jedem $a \in K$ disjunkte Umgebungen U_a von p und V_a von a finden.





$$\text{Aus } U_a \cap V_a = \emptyset \quad \Longrightarrow \quad \left(\bigcap_{j=1}^{r} U_{a_j}\right) \cap \left(\bigcup_{j=1}^{r} V_{a_j}\right) = \emptyset$$

Da K kompakt ist, gilt $K \subset V_{a_1} \cup \ldots \cup V_{a_r}$ für geeignete $a_1, \ldots, a_r \in K$.
Für $U := U_{a_1} \cap \ldots \cap U_{a_r}$ gilt dann: U ist Umgebung von p und $U \subset X - K$.

Die Beschränktheit von K ist klar:
Denn ist $a \in X$ beliebig, dann ist $X = \bigcup\limits_{k=1}^{\infty} U_k(a)$ und wegen $K \subset X$ gibt es endlich viele $k_1, \ldots, k_r \in \mathbb{N}$ mit

$$K \subset \bigcup_{j=1}^{r} U_{k_j}(a) .$$

Nimmt man $N := \max\{k_1, \ldots, k_r\}$, so ist $K \subset U_N(a)$, K also beschränkt.

24.5 Folgerung

Jede konvergente Folge in einem metrischen Raum ist beschränkt.

Das folgt aus 23.3(d) und dem Satz, lässt sich aber auch direkt zeigen.

Eine naheliegende Frage ist, inwiefern die Umkehrung von Satz 23.4 gilt.

24.6 Hilfssatz

Ist X ein kompakter metrische Raum und $A \subset X$ eine abgeschlossene Teilmenge, dann ist A ebenfalls kompakt.

Beweis : Sei dazu nämlich $(U_j)_{j \in J}$ irgendeine offene Überdeckung von A:

$$A \subset \bigcup_{j \in J} U_j .$$

Es ist $U := X - A$ offen und offensichtlich

$$X = U \cup \left(\bigcup_{j \in J} U_j \right) \supset A .$$

Da X kompakt ist, gibt es endlich viele $j_1, \ldots, j_r \in J$ mit

$$X = U \cup U_{j_1} \cup \ldots \cup U_{j_r} \supset A .$$

Wegen $U \cap A = \emptyset$ folgt hieraus bereits

$$A \subset U_{j_1} \cup \ldots \cup U_{j_r} .$$

□

Der klassische Satz von Bolzano-Weierstraß besagt, dass jede beschränkte Folge in $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) eine konvergente Teilfolge (= Häufungswert) besitzt. Es gilt folgende Verallgemeinerung.

24.7 Satz (von Bolzano-Weierstraß)

Ist K eine kompakte Teilmenge eines metrischen Raumes X, und (x_k) eine Folge von Punkte $x_k \in \mathbb{K}$, dann gibt es eine Teilfolge (x_{k_j}) von (x_k), die gegen einen Punkt $a \in K$ konvergiert.

Zum Beweis nehmen wir an, dass keine Teilfolge von (a_k) gegen einen Punkt von K konvergiert. Dann besitzt jeder Punkt $x \in K$ eine offene Umgebung U_x, in der höchstens endlich viele Glieder der Folge (x_k) liegen. Wenn nämlich in jeder Umgebung von x unendlich viele Folgenglieder liegen, dann könnte man eine Teilfolge von (x_k) konstruieren, die gegen x konvergiert.
Es gilt

$$K \subset \bigcup_{x \in K} U_x \ .$$

Wegen der Kompaktheit von K gibt es endlich viele Punkte $x_1, \dots, x_r \in K$ mit

$$K \subset U_{x_1} \cup \dots \cup U_{x_r} \ .$$

Damit liegen in ganz K nur endlich viele Folgenglieder.

24.8 Bemerkung

Nennt man eine Teilmenge K eines metrischen Raumes *folgenkompakt*, wenn jede Folge von Elementen aus K eine *in* K konvergente Teilfolge besitzt, dann kann man den Satz auch so aussprechen:

Eine *kompakte* Teilmenge K eines metrischen Raumes X ist *folgenkompakt*.

Hiervon gilt auch die Umkehrung, d.h. es gilt

Eine Teilmenge K eines metrischen Raumes ist genau dann kompakt, wenn sie folgenkompakt ist.

Der Beweis der Umkehrung sei als (etwa schwierigere) Übungsaufgabe gestellt. Man beweise hierzu zunächst:
Sei K eine folgenkompakte Teilmenge eines metrischen Raums X, dann gilt

(a) das *Lebesgue'sche Lemma*:
Zu jeder offenen Überdeckung $(U_j)_{j \in J}$ von K gibt es ein $\lambda > 0$, so dass für jedes $x \in K$ die λ-Umgebung $U_\lambda(x)$ in einer geeigneten der Mengen U_j enthalten ist.

Ein solches λ nennt man auch eine *Lebesgue'sche Zahl* der Überdeckung.

(b) Zu jedem $\varepsilon > 0$ gibt es endlich viele Punkte $x_1, \dots, x_r \in K$, so dass die ε-Umgebungen $U_\varepsilon(x_j)$ der x_j eine (offene) Überdeckung von K bilden.

Hierfür sagt man auch: Eine folgenkompakte Teilmenge ist totalbeschränkt.

Bemerkenswert ist der folgende

24.9 Satz

Jeder kompakte metrische Raum X ist vollständig.

Beweis : Wir müssen zeigen, dass jede Cauchy-Folge aus X einen Grenzwert in X hat.

Sei dazu (x_k) eine Cauchy-Folge in X und $\varepsilon > 0$ beliebig vorgegeben. Nach Definition gibt es dann einen Index $N_1 \in \mathbb{N}$, so dass für alle $k > m \geq N_1$ gilt

$$d(x_k, x_m) < \frac{\varepsilon}{2} .$$

□

Nach dem Satz von Bolzano-Weierstraß (§23.7) besitzt (x_k) eine konvergente Teilfolge (x_{k_j}) mit einem Grenzwert $a \in X$. a ist auch der Grenzwert der ursprünglichen Folge (x_k), denn wegen $\lim\limits_{j\to\infty} x_{k_j} = a$ gibt es ein $N_2 \in \mathbb{N}$, so dass für alle $j \geq N_2$ gilt

$$d(x_{k_j}, a) < \frac{\varepsilon}{2} .$$

Ist $N := \max\{N_1, N_2\}$, dann gilt für $k \geq N$

$$d(x_k, a) \leq d(x_k, x_{k_N}) + d(x_{k_N}, a) = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon ,$$

$$\text{d.h. } \lim_{k\to\infty} x_k = a .$$

Auch das Intervallschachtelungsprinzip aus der Analysis 1 lässt sich sinnvoll übertragen.

24.10 Theorem (allgemeines Intervallschachtelungsprinzip)

Sei X ein metrischer Raum und $(K_j)_{j\in\mathbb{N}_0}$ eine Folge nicht-leerer kompakten Teilmengen mit

$$K_0 \supset K_1 \supset K_2 \supset \ldots \supset K_j \supset K_{j+1} \supset \ldots ,$$

dann ist der Durchschnitt dieser Kompakte ebenfalls nicht leer:

$$\bigcap_{j=0}^{\infty} K_j := \{x \in X;\ x \in K_j \text{ für alle } j \in \mathbb{N}_0\} \neq \emptyset.$$

Beweis durch Widerspruch:
Wir nehmen an, es sei $\bigcap\limits_{j=0}^{\infty} K_j = \emptyset$. Dann ist (de Morgan lässt grüßen!)

$$\bigcup_{j=0}^{\infty}(X - K_j) = X .$$

Insbesondere ist dann K_0 in der Vereinigung der offenen Mengen $X - K_j$ enthalten. Da K_0 kompakt ist, gibt es eine natürliche Zahl r mit

$$K_0 \subset (X - K_{j_1}) \cup (X - K_{j_2}) \cup \ldots \cup (X - K_{j_r})$$

Ist $N := \max\{j_1, \ldots, j_r\}$, so gilt sogar $K_0 \subset X - K_N$ und daraus folgt $K_0 \cap K_N = \emptyset$.
Das ist aber ein Widerspruch zu

$$K_0 \cap K_N = K_N \neq \emptyset .$$

Zusatz: Bilden die Durchmesser der K_j eine Nullfolge, so gibt es genau $p \in \bigcap\limits_{j=0}^{\infty} K_j$, oder

$$\boxed{\{p\} = \bigcap_{j=0}^{\infty} K_j}$$

Unter dem Durchmesser einer Teilmenge K eines metrischen Raumes X versteht man dabei die Zahl

$$\delta(K) := \sup\{d(x,y);\ x,y \in K\}\ .$$

Eine weitere Möglichkeit, aus bekannten kompakten Räumen neu kompakte Räume zu konstruieren, ist die Produktbildung. Wir beschränken uns auf zwei Faktoren, der Satz gilt aber viel allgemeiner (siehe die spätere Bemerkung)

24.11 Theorem

Sind X und Y kompakte metrische Räume, dann ist auch das cartesische Produkt $X \times Y$ (versehen mit der Produktmetrik) kompakt.

Beweis : Sei $(U_j)_{j\in J}$ irgendeine offene Überdeckung von $X \times Y$, also

$$X \times Y = \bigcup_{j\in J} U_j,\ U_j \text{ offen in } X \times Y,$$

dann müssen wir eine endliche Teilüberdeckung konstruieren.

1. Schritt: Es sei $b \in Y$ ein fester Punkt, dann ist $X \times \{b\}$ (versehen mit der von $X \times Y$ induzierten Metrik) wie $X \times Y$ selbst auch ein kompakter metrische Raum, daher gibt es endlich viele Indizes $j_1, j_2, \ldots, j_r$ mit

$$X \times \{b\} \subset U_{j_1} \cup U_{j_2} \cup \ldots \cup U_{j_r}.$$

2. Schritt: Sei $U := U_{j_1} \cup \ldots \cup U_{j_r}$.
Die Menge U hängt von b ab, wir schreiben daher $U = U(b)$. Diese Menge ist jedenfalls als Vereinigung offener Mengen offen in $X \times Y$. Hieraus wollen wir schließen, dass es ein $r = r(b) > 0$ gibt mit $X \times U_r(b) \subset U$. Dies kann man so einsehen: Da U offen ist, existiert zu jedem Punkt $x \in X$ eine reelle Zahl $r(x) > 0$, so dass

$$U_{r(x)} \times U_{r(x)}(b) \subset U$$

gilt (beachte dabei, dass die Kugelumgebungen in $X \times Y$ Produkte von Kugelumgebungen in X bzw. Y sind). Es gilt also

$$X \times \{b\} \subset \bigcup_{x\in X} \left(U_{r(x)} \times U_{r(x)}(b)\right) \subset U.$$

Da aber $X \times \{b\}$ kompakt ist, gibt es $x_1, \ldots, x_n$ und Radien $r_1 := r(x_1), \ldots, r_n := r(x_n) > 0$, so dass bereits

$$X \times \{b\} \subset (U_{r_1}(x_1) \times U_{r_1}(b)) \times \ldots \times (U_{r_n}(x_n) \times U_{r_n}(b)) \subset U$$

gilt.
Setzt man jetzt

$$r := \min\{r_1, \ldots, r_n\}\ ,$$

so gilt offensichtlich

$$X \times \{b\} \subset X \times U_r(b) \subset U.$$

3. Schritt: Unsere bisherigen Überlegungen zeigen:
Zu jedem $b \in Y$ gibt es eine reelle Zahl $r > 0$, so dass $X \times U_r(b),\ r = r(b)$, von endlich vielen der Mengen U_j überdeckt sind.
Nun gilt aber

$$Y = \bigcup_{b \in Y} U_r(b).$$

Wegen der Kompaktheit gibt es dann endlich viele Punkte $b_1, \ldots, b_k \in Y$ mit

$$Y = U_{r_1}(b_1) \cup \ldots \cup U_{r_k}(b_k).$$

Daher ist

$$X \times Y = \bigcup_{\nu=1}^{k} \left(X \times U_{r_\nu}(b_\nu) \right).$$

Jede der Mengen $X \times U_{j\nu}(b_j)$ $(1 \leq \nu \leq k)$ wird aber durch endlich viele U_j überdeckt, daher genügen auch schon endlich viele U_j, um $X \times Y$ zu überdecken, das bedeutet also dass auch $X \times Y$ kompakt ist.

$\square$

Bemerkung:
Der Beweis lässt sich sofort auf endlich viele Faktoren übertragen:
Sind $X_1, \times, X_n$ kompakte metrische Räume und $X := X_1 \times \ldots \times X_n$ ihr Produkt, dann ist auch X kompakt (bezüglich der Produktmetrik).
Es gilt sogar (auch für beliebige topologische Räume):

Ein beliebiges Produkt kompakter Räume ist kompakt (Satz von Tychonoff).

Einen beweis findet man etwa in *B.Querenburg: Mengentheoretische Topologie (Springer-Verlag, 3.Aufl. 2003)*, dort im Abschnitt 8, wo auch andere Kompaktheitsbegriffe behandelt werden. Der Beweis wird meist über Ultrafilter geführt. Auch in Fall von metrischen Räumen, muss man bei unendlich vielen Faktoren erstmal auf dem Produkt eine Metrik definieren.

Als Vorstufe für den Satz von Heine-Borel in $\mathbb{K}^n$ benötigen wir

24.12 Satz (Spezialfall des Satzes von Heine-Borel)

Für beliebige $a, b \in \mathbb{R}$ mit $a \leq b$ ist das Intervall

$$D := [a, b] = \{x \in \mathbb{R};\ a \leq x \leq b\}$$

kompakt.
Das Intervall vom Typ $[a, b]$, das wir bisher auch „kompaktes“ Intervall genannt haben, ist also tatsächlich im Sinne unserer neuen allgemeineren Definition kompakt.

Beweis : Sei $(U_j)_{j \in J}$ irgendeine Überdeckung von D durch offene Mengen $U_j \subset \mathbb{R}$, also $D \subset \bigcup\limits_{j \in J} U_j$ und

$$M := \{x \in [a, b];\ [a, x] \text{ besitzt eine Überdeckung durch endlich viele } U_j\}$$

Dann ist $M \neq \emptyset$ (da $a \in M$) und nach oben beschränkt, also existiert ein $\xi := \sup M$.
Offenbar gilt: Ist $x \in M$ und ist $a \leq t \leq x \implies t \in M$, m.a.W.: M ist eine Intervall:

$$M = [a, \xi[\text{ und } M = [a, \xi]\ .$$

Wir wählen einen Index $j_0 \in J$ mit $\xi \in U_{j_0}$. Da U_{j_0} offen ist, existiert ein $\varepsilon > 0$ mit

$$U_\varepsilon(\xi) =]\xi - \varepsilon, \xi + \varepsilon[\subset U_{j_0} \ .$$

Man wähle nun irgendein x mit

$$\xi - \varepsilon < x < \xi \ \text{ und } \ a \leq x \ .$$

Dann gilt $x \in M$, also

$$[a, x] \subset U_{j_1} \cup U_{j_2} \cup \ldots \cup U_{j_r} \quad (r \in \mathbb{N} \ \text{ geeignet})$$

Daher ist auch

$$[a, \xi] \subset U_{j_0} \cup U_{j_1} \cup \ldots \cup U_{j_r},$$

also gilt auch $\xi \in M$ und damit ist $M = [a, \xi]$.

Wäre $\xi < b$, so könnte man $\varepsilon > 0$ so klein wählen, dass noch $\xi + \varepsilon \leq b$ gilt und man hätte dann

$$[a, \xi + \varepsilon] \subset U_{j_0} \cup U_{j_1} \cup \ldots \cup U_{j_r}$$

im Widerspruch zur Definition ξ als Supremum.

□

24.13 Theorem (Heine-Borel für $\mathbb{K}^n$)

Eine Teilmenge $K \subset \mathbb{K}^n,\ n \in \mathbb{N}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.

Wegen $\mathbb{C}^n = \mathbb{R}^{2n}$ genügt es den Fall $K = \mathbb{R}$ zu betrachten.

Beweis : „$\Longleftarrow$" Sei zunächst K beschränkt und abgeschlossen in $\mathbb{K}^n$. Wegen der Beschränktheit ist K in einem (möglicherweise großen) Würfel enthalten: Es gibt also ein $\lambda > 0$ mit

$$K \subset W := [-\lambda, \lambda]^n \subset \mathbb{R}^n.$$

Auch in W ist K noch abgeschlossen, denn das Komplement ist ja

$$W - K = W \cap (\mathbb{R}^n - K)$$

und damit offen in W.
Der Würfel ist als Produkt kompakter Intervalle kompakt. Als abgeschlossene Teilmenge von W ist daher auch K kompakt (vgl. ???).

„$\Longrightarrow$" Wir wissen (vgl. ???), dass eine kompakte Teilmenge eines metrischen Raumes beschränkt und abgeschlossen ist.

□

Beachte: Kompakte Teilmengen eines metrischen Raumes sind immer beschränkt und abgeschlossen. Die Umkehrung gilt für die Standardvektorräume $\mathbb{K}^n$ mit den Standardmetriken, ist i.A. aber falsch. Ein simples Beispiel ist $X =]0, 1[$ (als metrischer Raum mit der von $\mathbb{R}$ induzierten Metrik) und nimmt man $K := X$, dann ist K beschränkt und da der ganze Raum immer abgeschlossen ist (bezüglich der induzierten Metrik) ist K auch abgeschlossen, aber K ist nicht kompakt.
Man überlege sich, dass die abgeschlossene Einheitskugel in einem unendlich dimensionalen Hilbert-Raum nicht kompakt (sie ist es genau dann, wenn der Raum endlich-dimensional ist).

Viele wichtige Existenzaussagen der Analysis beruhen auf Eigenschaften stetiger Abbildungen kompakter Räume, insbesondere auf dem Satz von der Annahme eines Maximums und Minimums für stetige reellwertige Funktionen (auf kompakten Mengen) und dem Satz von der gleichmäßigen Stetigkeit (für stetige Funktionen auf kompakten Mengen).
Beide Sätze haben wir im Spezialfall $K \subset \mathbb{K}$ kompakt früher bewiesen. Wir erhalten sie als Spezialfall nochmals aus den folgenden allgemeinen Sätzen:

24.14 Satz

Sind X und Y metrische Räume, X kompakt und ist $f : X \to Y$ stetig, dann ist das Bild $f(X)$ eine kompakte Teilmenge von Y.

$\boxed{\text{Kurz: Stetige Bilder kompakter Räume sind kompakt.}}$

Der **Beweis** ist elementar:

Sei nämlich $(V_j)_{j\in J}$ irgendeine offene Überdeckung von $f(X)$, also

$$f(X) \subset \bigcup_{j\in J} V_j.$$

Wir müssen eine endliche Teilüberdeckung konstruieren.
Wegen der Stetigkeit von f sind die Urbilder $U_j = f^{-1}(V_j)$ offen in X und bilden daher eine offene Überdeckung von X

$$X = \bigcup_{j\in J} U_j.$$

Da X kompakt ist, gibt es eine endliche Teilmenge $J_0 \subset J$ mit $X = \bigcup\limits_{j\in J_0} U_j$. Dann gilt aber

$$f(X) \subset \bigcup_{j\in J_0} V_j.$$

□

Als Folgerung ergibt sich ein berühmtes Theorem:

24.15 Theorem (K.Weierstraß, 1861)

Ist X $(\neq \emptyset)$ ein kompakter metrische Raum und $f : X \to \mathbb{R}$ eine stetige Funktion, dann besitzt f ein (absolutes) Maximum und (absolutes) Minimum, d.h. es gibt Punkte $x_{max} \in X$ und $x_{min} \in X$ mit

$$f(x_{max}) \geq f(x) \quad \text{bzw.} \quad f(x_{min}) \leq f(x)$$

für alle $x \in X$.

Beweis : Nach dem vorangehenden Satz ist das Bild $A := f(X)$ eine kompakte Teilmenge von $\mathbb{R}$. A ist also insbesondere (nach oben und unten) beschränkt und nicht leer. Also existieren $M = \sup A$ und $m := \inf A$.
Wegen der Abgeschlossenheit von A gehören aber die Häufungspunkte M und m von A selber zu A.

□

24.16 Folgerung

Ist $K \subset \mathbb{R}^n$ eine (nicht-leere) beschränkte und abgeschlossene (also kompakte) Teilmenge und ist $f : K \to \mathbb{R}$ eine stetige Funktion, dann hat f auf K ein (absolutes) Maximum und (absolutes) Minimum.

24.17 Beispiel

Definiert man als Abstand zweier nicht-leerer Teilmengen K und A eines metrischen Raums (X, d) die (nicht negative) reelle Zahl

$$d(K, A) := \inf\{d(x, y);\ x \in K, y \in A\},$$

dann gilt: Ist K kompakt, A abgeschlossen und $K \cap A = \emptyset$, dann gibt es einen Punkt $p \in K$ mit

$$d(p, A) := d(\{p\}, A) = d(K, A),$$

insbesondere ist stets $d(K, A) > 0$.

Beweis : Die Funktion

$$\begin{array}{ccc} K & \to & \mathbb{R} \\ x & \mapsto & d(x, A) \end{array}$$

ist stetig und nimmt auf K ein Minimum an. Es gibt daher einen Punkt $p \in K$ mit $d(p, A) = d(K, A)$. Da (wegen $K \cap A = \emptyset$) p nicht in A liegt und A abgeschlossen ist, gibt es eine Kugel $U_r(p) \subset K$ $(r > o)$, mit $U_r(p) \cap A = \emptyset$. Folglich ist

$$d(p, a) \geq r > 0$$

für alle $a \in A$, also $d(p, A) \geq r$.

□

Zu den grundlegenden Gestaltungsprinzipien der heutigen Mathematik gehört es, dass man nicht nur Objekte betrachtet, wie z.B. Gruppen, Vektorräume, Modulen o.ä. Objekte, sondern auch Abbildungen, die mit den jeweiligen Strukturen verträglich sind, z.B. das Gruppenhomomorphismus, Vektorraumhomomorphismus (= lineare Abbildungen) etc.

Ist z.B. $f : V \to W$ eine *bijektive* lineare Abbildung zwischen endlich-dimensionalen K-Vektorräumen, dann ist die existierende Umkehrung $g : W \to V$ stets automatisch auch linear. (Können Sie diese Tatsache aus dem Stegreif beweisen?)

Ähnliches gilt für bijektive Abbildungen zwischen Gruppen.
Mann könnte daher vermuten, dass eine *bijektive*, stetige Abbildung $f : X \to Y$ zwischen metrischen Räumen X und Y stets auch eine stetige Umkehrabbildung besitzt. Das ist i.A. jedoch nicht der Fall, wie das folgende Beispiel zeigt:
Sei $M = [0, 2\pi[$ und

$$\begin{array}{rcl} f : M \to S^1 & := & \{z \in \mathbb{C};\ |z| = 1\} \\ & = & \{(x, y) \in \mathbb{R}^2;\ x^2 + y^2 = 1\} \end{array}$$

die stetige und bijektive Abbildung

$$t \mapsto \exp(it) := \cos t + i \sin t = (\cos t, \sin t)$$

(M sei mit der von $\mathbb{R}$ induzierten Metrik versehen).

Für $z \in S^1$ ist die durch

$$g(z) = t \iff t \in [0, 2\pi[\text{ und } \exp(t) = z$$

erklärte Umkehrabbildung aber im Punkt $1 = (1, 0)$ nicht stetig, denn etwa die Folgen (z_k) und (w_k) mit

$$z_k = \exp\left(i\frac{1}{k}\right) \text{ und } w_k = \exp\left(i\left(2\pi - \frac{1}{k}\right)\right)$$

konvergieren beide gegen $1 = (1, 0)$, die Bildfolgen $(g(z_k))$ bzw. $(g(w_k))$ konvergieren aber gegen 0 bzw. 2π.

Dass g nicht stetig sein kann ergibt sich auch so:
Wäre g stetig, so müsste nach ??? das Bild $g(S^1) = [0, 2\pi[$ kompakt sein, da S^1 eine kompakte Teilmenge von $\mathbb{C} = \mathbb{R}^2$ ist. Das Intervall $[0, 2\pi[$ ist aber nicht kompakt!

Bemerkung: Wir hätten die Abbildung f auf ganz $\mathbb{R}$ betrachten können, aber da die Winkelfunktionen die Periode 2π und damit auch alle Zahlen $2\pi k,\ k \in \mathbb{Z}$, als Periode haben, nimmt f bereits alle Werte in einem Intervall vom Typ $[(2k-1)\pi,\ 2k\pi[,\ k \in \mathbb{Z}$, an, speziell kann man $k = 1$ wählen.
Man sagt auch: Durch die Abbildung

$$\begin{array}{rcl} \mathbb{R} & \to & S^1 \\ t & \mapsto & \exp(it) = \cos t + i \sin t \end{array}$$

wird die reelle Gerade auf die Einheitskreislinie „aufgewickelt".

Es gilt jedoch

24.18 Satz

Ist X ein kompakter metrische Raum und $f : X \to Y$ eine bijektive stetige Abbildung auf einen metrischen Raum Y, dann ist die Umkehrabbildung $g : Y \to X$ stetig.

Der **Beweis** ist mit den uns jetzt zur Verfügung stehenden Mitteln ganz einfach.
Wir wissen (vgl. ???), dass eine Abbildung zwischen metrischen Räumen genau dann stetig ist, wenn die Urbilder offener (bzw. abgeschlossener) Mengen wieder offen (bzw. Abgeschlossen) sind.
Wir haben die Stetigkeit von $g : Y \to X$ zu zeigen.
Sei also $A \subset X$ abgeschlossen. Das Urbild von A ist dann aber gerade

$$g^{-1}(A) = f(A) \subset Y.$$

Nun ist aber A kompakt als abgeschlossener Teil von X (vgl. ???), $f(A)$ als stetiges Bild eines Kompaktum kompakt (vgl. ???) und deshalb abgeschlossen als kompakte Teilmenge des metrischen Raumes Y (vgl. ???).

$\square$

Eine weitere schöne Anwendung des Weierstraß'schen Satzes ist die Tatsache, dass in den Standardvektorräumen $\mathbb{K}^n$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) *alle Normen äquivalent sind* und damit den gleichen Konvergenzbegriff liefern:

24.19 Satz (Äquivalenzsatz für Normen)

In $\mathbb{K}^n$ sind je zwei Normen äquivalent, d.h. sind $N_1, N_2 : \mathbb{K}^n \to \mathbb{R}$ zwei beliebige Normen, dann gibt es positive reelle Zahlen a, b mit

$$\boxed{aN_1(x) \leq N_2(x) \leq bN_1(x)}$$

für alle $x \in \mathbb{K}^n$.

24.19.1 Bemerkungen

(a) Die aus den Normen abgeleiteten Metriken sind dann sogar streng äquivalent.

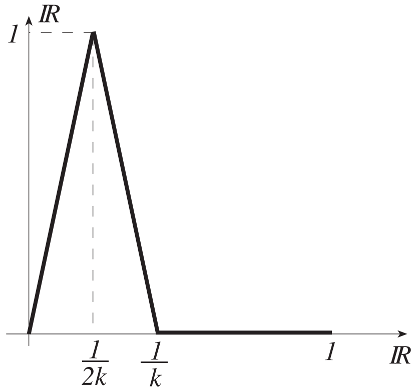
(b) Auf einem beliebigen $\mathbb{R}$- oder $\mathbb{C}$-Vektorraum V ist von vornherein nicht zu erwarten, dass je zwei Normen äquivalent sind, so sind z.B. im Vektorraum $V = \{f : [0,1] \to \mathbb{R};\ f \text{ stetig}\}$ die beiden Normen

$$N_\infty(f) := \|f\|_\infty = \max\{|f(t)|;\ t \in [0,1]\} \qquad \text{und} \qquad N_1(f) := \int_0^1 |f(t)|dt$$

nicht äquivalent.

Es gilt zwar $N_1(f) = \int_0^1 |f(t)|dt \leq N_\infty(f)$, d.h. aus der N_∞-Konvergenz (= gleichmäßige Konvergenz) einer Folge folgt stets die Konvergenz bezüglich N_1 (= Konvergenz im absoluten Mittel).

Die Umkehrung gilt jedoch i.A. nicht: Es gibt z.B. N_1-Nullfolgen, die keine N_∞-Nullfolgen sind.
Das zeigen die Funktionen $f_k : [0,1] \to \mathbb{R}$ mit denen in der Abbildung skizzierten Graphen:



(c) Für die Standard-Normen in $\mathbb{K}^n$ haben wir die Äquivalenz direkt gezeigt (vgl. ???).

Beweis des Satzes: Es genügt zu zeigen, dass eine beliebige Norm N auf $\mathbb{K}^n$ zur Würfel-Norm $\| \ \|_\infty$ äquivalent ist, hier sei also

$$\|x\|_\infty := \max\{|x_1|, |x_2|, \ldots, |x_n|\}$$

für $x = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$.
Ist $N : \mathbb{K}^n \to \mathbb{R}$ eine beliebige Norm, so gilt mit der Standard-Basis $(e_1, \ldots, e_n)$ von $\mathbb{K}^n$ zunächst

$$x = \sum_{j=1}^n x_j e_j, \quad x_j \in \mathbb{K}, \quad 1 \leq j \leq n,$$

und

$$N(x) = N\left(\sum_{j=1}^n x_j e_j\right) \leq \sum_{j=1}^n |x_j| N(e_j) \leq b\|x\|_\infty$$

mit $b := \sum_{j=1}^{n} N(e_j) > 0$.
Hieraus schließen wir, dass die Abbildung

$$\begin{array}{rcl} N: \mathbb{K}^n & \to & \mathbb{R}\,, \\ x & \mapsto & N(x) \end{array}$$

stetige ist, wenn man $\mathbb{K}^n$ mit der Würfel-Norm versieht. Es gilt ja für $x, y \in \mathbb{K}^n$

$$|N(x) - N(y)| \leq N(x - y)$$

auf Grund der Dreiecksungleichung für N, und damit

$$|N(x) - N(y)| \leq N(x - y) \leq b\|x - y\|_\infty$$

Wir wissen schon, dass $N(x) \leq b\|x\|_\infty$ gilt. Um eine Abschätzung nach unten zu erhalten, betrachten wir, dass die Würfeloberfläche

$$W := \{x \in \mathbb{K}^n;\ \|x\|_\infty = 1\}$$

kompakt ist.
Da W nicht den Nullpunkt enthält, sind die Werte der stetigen Funktion N auf W stets echt positiv, auf Grund des Weierstraß'schen Satzes gibt es daher zu $a > 0$ mit der Eigenschaft:

$$\|y\|_\infty = 1 \quad \Longrightarrow \quad N(y) \geq a$$

Hieraus folgt durch Übergang von $x \neq 0$ zu $y := \frac{x}{\|x\|_\infty}$ in der Tat

$$N(x) \geq a\|x\|_\infty$$

für alle $x \in \mathbb{K}^n$ (zunächst für $x \neq 0$, für $x = 0$ gilt das Gleichheitszeichen).

□

Übungsaufgabe: Zeigen Sie: Je zwei Normen auf einem endlich dimensionalen $\mathbb{K}$-Vektorraum sind äquivalent.

24.19.2 Bemerkung

Weitere wichtige theoretischen Anwendungen der Weierstraß'schen Satzes findet man u.a.

(a) bei manchen Beweisen des Fundamentalsatzes der Algebra, wenn an entscheidender Stelle benutzt wird, dass der Betrag einer komplexen Polynomfunktion ein (absolutes) Minimum besitzt.

(b) bei Beweisen des Spektralsatzes für symmetrische Operatoren mit Hilfe des sog. *Rayleigh-Quotienten*, wobei benutzt wird, dass die stetige Funktion

$$\begin{array}{rcl} g_S: \mathbb{R}^n & \to & \mathbb{R}\,, \\ x & \mapsto & \dfrac{< x, Sx >}{\|x\|_2^2} \end{array}$$

auf der Standard-Sphäre $S^{n-1} := \{x \in \mathbb{R}^n;\ \|x\|_2 = 1\}$ ein absolutes Minimum hat und der Funktionswert dann der kleinste Eigenwert der symmetrischen Matrix $S \in \mathbb{R}^{n \times n}$ ist.

Wir könnten den Beweis hier führen, erhalten die Aussage aber auch später ganz einfach mit Hilfe der *Lagrangeschen Multiplikatoren*.

Zum Abschluss unserer Betrachtungen über kompakte Räume zeigen wir, dass eine stetige Funktion auf einem kompakten Raum *gleichmäßig stetig* ist, eine Eigenschaft, die wir von stetigen Funktionen auf kompakten Intervallen kennen und z.B. beim Beweis, dass eine stetige Funktion $f : [a, b] \to \mathbb{R}$ eine Regelfunktion ist (also integrierbar ist), von zentraler Bedeutung war.

24.20 Definition (gleichmäßige Stetigkeit)

Eine Abbildung $f : (X, d) \to (Y, \tilde{d})$ zwischen metrischen Räumen heißt *gleichmäßig stetig*, wenn es zu jedem $\varepsilon > 0$ ein (Universal-)$\delta > 0$ gibt, so dass für alle $x, x' \in X$ mit $d(x, x') < \delta$ gilt

$$\tilde{d}(f(x), f(x')) < \varepsilon$$

Bemerkung: Eine gleichmäßig stetige Abbildung ist natürlich stetig.

24.21 Satz

Sind (X, d) und $(Y, \tilde{d})$ metrische Räume und ist X kompakt. Dann ist jede stetige Abbildung $f : X \to Y$ gleichmäßig stetig.

Beweis : Sei $\varepsilon > 0$ beliebig vorgegeben. Dann gibt es zu jedem $p \in X$ ein $\delta(p) > 0$, so dass für alle $x \in U_{\delta(p)}(p)$ gilt

$$\tilde{d}(f(x), f(p)) < \frac{\varepsilon}{2}.$$

Da $\bigcup\limits_{p \in X} U_{\frac{1}{2}\delta(p)}(p) = X$ gilt und X kompakt ist, gibt es endlich viele Punkte $p_1, \ldots, p_k \in X$ mit

$$\bigcup_{j=1}^{k} U_{\frac{1}{2}\delta(p_j)}(p_j) = X.$$

Sei $\delta := \frac{1}{2}\min\{\delta(p_1), \ldots, \delta(p_k)\}$. Sind dann $x, x' \in X$ beliebig mit $d(x, x') < \delta$, so gibt es ein $j \in \{1, \ldots, k\}$ mit $x \in U_{\frac{1}{2}\delta(p_j)}(p_j)$ und es gilt dann $x' \in U_{\delta(p_j)}(p_j)$. Es folgt

$$\tilde{d}(f(x), f(p_j)) < \frac{\varepsilon}{2} \quad \text{und} \quad \tilde{d}(f(x'), f(p_j)) < \frac{\varepsilon}{2},$$

daher

$$\tilde{d}(f(x), f(x')) \le \tilde{d}(f(x), f(p_j)) + \tilde{d}(f(p_j), f(x')) < \varepsilon.$$

□

Wir behandeln zum Abschluss des Paragraphen eine typische Anwendung des Satzes von der gleichmäßigen Stetigkeit auf *Doppelintegrale*.

24.21.1 Satz

Sei $a, b, c, d \in \mathbb{R},\ a < b,\ c < d$ und

$$\begin{array}{rcl} f:\ [a,b] \times [c,d] & \to & \mathbb{R}\,, \\ (x,y) & \mapsto & f(x,y) \end{array}$$

eine stetige Funktion.
Bei festem y ist die Funktion

$$\begin{array}{rcl} f^y:\ [a,b] & \to & \mathbb{R}\,, \\ x & \mapsto & f(x,y) \end{array}$$

dann auch stetig, kann also über $[a,b]$ integriert werden:

$$F(y) := \int_a^b f^y(x)dx = \int_a^b f(x,y)dx.$$

Die Funktion

$$\begin{array}{rcl} F:\ [c,d] & \to & \mathbb{R}\,, \\ y & \mapsto & F(y) \end{array}$$

ist stetig auf $[c,d]$, kann also über $[c,d]$ integriert werden.

Man nennt

$$\int_c^d F(y)dy = \int_c^d \left(\int_a^b f(x,y)dx \right) dy$$

ein *Doppelintegral*.

Beweis : Es ist lediglich die Stetigkeit von

$$\begin{array}{rcl} F:\ [c,d] & \to & \mathbb{R}\,, \\ y & \mapsto & F(y) = \int_a^b f(x,y)dx \end{array}$$

zu zeigen.

Dies folgt aber einfach aus der gleichmäßigen Stetigkeit von $f : [a,b] \times [a,d] \to \mathbb{R}$ (beachte: das Rechteck $[a,b] \times [c,d]$ ist kompakt).
Zu vorgegebenem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x, x' \in [a,b]$ bzw. $y, y' \in [c,d]$ mit $|x - x'| < \delta,\ |y - y'| < \delta$ gilt

$$|f(x',y') - f(x,y)| < \frac{\varepsilon}{b-a}$$

Damit ergibt sich für $|y' - y| < \delta$

$$\begin{array}{rcl} |F(y') - F(y)| & = & \left| \int_a^b (f(x,y') - f(x,y))dx \right| \\ & \leq & \int_a^b |f(x,y') - f(x,y)|dx \\ & \leq & (b-a)\dfrac{\varepsilon}{b-a} = \varepsilon. \end{array}$$

F ist also sogar Lipschitz-stetig.

□

Zusatz: Man kann den Satz benutzen um damit ein Integral von f über das Rechteck

$$R := [a,b] \times [c,d]$$

zu definieren:

$$\int_R f(x,y)d(x,y) := \int_c^d \left(\int_a^b f(x,y)dx \right) dy.$$

Mit Hilfe des Satzes über parameterabhängige Integrale (vgl. ???), kann man leicht zeigen, dass

$$\int_c^d \left(\int_a^b f(x,y)dx \right) dy = \int_a^b \left(\int_c^d f(x,y)dy \right) dx$$

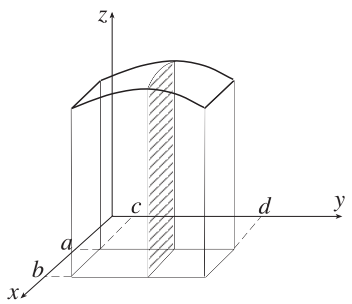
gilt, so dass also auch

$$\int_R f(x,y)d(x,y) = \int_a^b \left(\int_c^d f(x,y)dy \right) dx$$

gilt.
Hier handelt es sich um einen extremen Spezialfall des *Satzes von Fubini* (vgl. ???).

Ist $f : R \to \mathbb{R}$ eine nicht negative Funktion, d.h. $f(x) \geq 0$ für alle $x \in \mathbb{R}$, dann kann man das Integral interpretieren als das Volumen des Bereiches (Ordinatenmenge von f).

$$O(f) := \{(x,y,z) \in \mathbb{R}^3;\ a \leq x \leq b,\ c \leq y \leq d;\ 0 \leq z \leq f(x,y)\}$$



Für jedes $y \in [c,d]$ ist $F(y)$ der Flächeninhalt des Schnitts von $O(f)$ mit der zu x-z-Ebene parallelen Ebene durch y (im Bild schattiert). Alle diese Scheiben aufintegriert über $[c,d]$ ergeben das Volumen von $O(f)$.

Aufgabe: Man zeige

$$Irtf(x,y) = \begin{cases} \sqrt{1-x^2-y^2}, & \text{falls } x^2+y^2 \leq 1 \\ 0, & \text{falls } x^2+y^2 > 1 \end{cases}$$

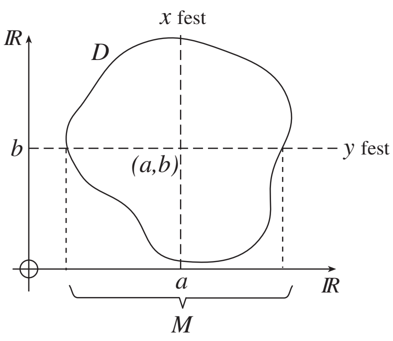
so ist

$$\int_{\substack{-1 \leq x \leq 1 \\ -1 \leq y \leq 1}} f(x,y)d(x,y) = \frac{2}{3}\pi$$

VII. Differentialrechnung in mehreren Variablen

In diesem zentralen Kapitel entwickeln wir die fundamentalen Tatsachen der Differentialrechnung für Abbildungen
$$f : D \to \mathbb{R}^m,$$
deren Definitionsbereich eine offene (nichtleere) Menge $D \subset \mathbb{R}^n$ ist. Dabei geht es zunächst darum, den „richtigen“ Begriff für die Differenzierbarkeit einer Funktion von mehreren Variablen zu finden. Die naheliegende Idee, bei einer Funktion von n Variablen $n-1$ Variablen festzuhalten und die dadurch entstehende Funktion einer Veränderlichen mit den geläufigen Methoden der Differentialrechnung einer Variablen zu behandeln, führt zum Begriff der *partiellen Ableitung* und zur *partiellen Differenzierbarkeit.*

Sei z.B. $D \subset \mathbb{R}^2$ offen und $(a,b) \in D$ ein fester Punkt und $f : D \to \mathbb{R}$ eine (reellwertige) Funktion.
Bei festem $y = b$ erhält man durch
$$\begin{array}{rcl} f^b : M & \to & \mathbb{R}\,, \\ x & \mapsto & f(x,b) \end{array}$$
eine Funktion einer Veränderlichen, deren Definitionsbereich die Menge $M := \{x \in \mathbb{R};\ (x,b) \in D\}$ ist. (siehe Abb.)



Diese Menge ist eine offene Umgebung von a (in $\mathbb{R}$). Hat nun diese Funktion in $x = a$ eine Ableitung, existiert also
$$\lim_{\substack{x \to a \\ x \neq a}} \frac{f^b(x) - f^b(a)}{x-a} = \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x,b) - f(a,b)}{x-a}\,,$$
so heißt dieser Grenzwert die *partielle Ableitung* von f nach der ersten Variablen in (a,b).

Notation: $\partial_1 f(a,b) = \partial_x f(a,b) = \dfrac{\partial f}{\partial x}(a,b)$ o.ä.

Völlig analog definiert man
$$\partial_2 f(a,b) = \partial_y f(a,b) = \frac{\partial f}{\partial y}(a,b) \quad \text{o.ä.}$$

Beispiele:

(a) Sei $f(x,y) := 3x - 6y$ und $D = \mathbb{R}^2$. Die partiellen Ableitungen existieren in allen Punkten $(x,y) \in D$ und sind konstant:
$$\partial_1 f(x,y) = 3 \qquad \text{und} \qquad \partial_2 f(x,y) = -6\,.$$

(b) Für
$$\begin{array}{rcl} f : \mathbb{R}^2 & \to & \mathbb{R}\,, \\ (x,y) & \mapsto & e^x \cos y \end{array}$$
existieren für alle $(x,y) \in \mathbb{R}^2$ die partiellen Ableitungen, und es gilt
$$\partial_1 f(x,y) = e^x \cos y \qquad \text{und} \qquad \partial_2 f(x,y) = -e^x \sin y\,.$$

(c) Für $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x,y) = \begin{cases} 0, & \text{falls } x = y = 0; \\ \frac{2xy}{x^2+y^2}, & \text{falls } (x,y) \neq (0,0); \end{cases}$$

existieren die partiellen Ableitungen im Punkt $(a,b) := (0,0)$ und es gilt

$$\partial_1 f(0,0) = 0 = \partial_2 f(0,0) \ ,$$

was sich sofort aus der Definition ergibt.

Die Funktion f ist aber an der Stelle $(0,0)$ nicht stetig, wie man etwa sieht, wenn man die Folge $(x_k, y_k) = \left(\frac{1}{k}, \frac{1}{k}\right)$ betrachtet.

Es ist $f(x_k, y_k) = \frac{2\frac{1}{k}\cdot\frac{1}{k}}{\frac{1}{k^2}+\frac{1}{k^2}} = 1$ für alle $k \in \mathbb{N}$ und $\lim\limits_{k\to\infty}(x_k, y_k) = (0,0)$, aber

$$0 = f(0,0) \neq \lim_{k\to\infty} f(x_k, y_k) = 1 \ .$$

(d) Noch einfacher ist $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = 0$, falls $xy = 0$ und $f(x,y) = 1$ sonst. Es ist $\partial_1 f(0,0) = \partial_2 f(0,0) = 0$, f ist aber unstetig in $(0,0)$.

Aus der partiellen Differenzierbarkeit folgt also i.A. nicht die Stetigkeit der Funktion an der betreffenden Stelle. Wir suchen daher noch einen stärkeren Differenzierbarkeitsbegriff, für welchen aus der Differenzierbarkeit an der betrachteten Stelle die Stetigkeit folgt.

Das ist der Begriff der *totalen Differenzierbarkeit*, wobei die Grundidee -wie in einer Veränderlichen auch- darin besteht, die Änderungsrate der Funktion durch eine lineare Abbildung geeignet zu approximieren. Die Ableitung wird keine Zahl, sondern eine *lineare Abbildung* sein. Das erfordert einen gewissen Gewöhnungsbedarf und die Verwendung der Techniken der lineare Algebra.

Die partiellen Ableitungen spielen bei der Berechnung der die lineare Abbildung repräsentierenden Matrix (bzw. der Standardbasen) -man nennt diese Matrix die *Jacobi-Matrix*- eine wichtige Rolle und sie liefern unter der zusätzlichen Voraussetzung ihrer Stetigkeit das fundamentale *hinreichende Kriterium* für die totale Differenzierbarkeit.

Partielle Ableitungen sind jedoch nur *spezielle Richtungsableitungen* (Ableitungen in Richtung der Einheitsvektoren). Wir beschäftigen uns deshalb zunächst mit

- verschiedenen Differenzierbarkeitsbegriffen,
- Rechenregeln für differenzierbare Abbildungen, von besonderer Wichtigkeit ist dabei
- die Kettenregel,
- Mittelwertsatz und Schrankensatz,
- höheren Ableitungen und dem Satz von Schwarz,
- Taylor-Formel und lokalen Extremwerten,

In einem kurzen Exkurs beschäftigen wir uns auch mit der Frage

- Wann besitzt ein Vektorfeld ein Potential?
  Dabei wird offenbar, dass sich Vektorfelder und Differetialformen ersten Grades (sog. Pfaffsche Formen) umkehrbar eindeutig entsprechen. Ob man also Kurvenintegrale längs Vektorfeldern oder Kurvenintegrale von 1-Formen betrachtet, ist nur eine Frage der Sprechweise.

Den Abschluss des Kapitels bilden

- der lokale *Umkehrsatz für differenzierbare Abbildungen* und

- der *Satz über implizite Funktionen*, so wie dessen Anwendung auf
- Extremwerte unter Nebenbedingungen (Lagrange'sche Multiplikatoren).

Zur Notation:
Im Folgenden ist D meist eine offene (nichtleere) Teilmenge des $\mathbb{R}^n$. Wir betrachten Abbildungen

$$\begin{aligned} f:\ D &\to \mathbb{R}^m\ , \\ x &\mapsto f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} = f_1(x)e_1 + \ldots + f_m(x)e_m\ , \end{aligned}$$

$(e_1, \ldots, e_m)$ die Standardbasis des $\mathbb{R}^m$.

Die Funktionen $f_j : D \to \mathbb{R}$ $(1 \leq j \leq m)$ heißen die *Komponentenfunktionen* von f. Ist $m \geq 2$, dann heißt f auch vektorwertige Funktion, für $m = 1$ Skalarfunktion oder einfach nur Funktion.

Aus Platzspargründen verwenden wir für $x \in \mathbb{R}^n$ meist die Zeilenschreibweise $x = (x_1, \ldots, x_n)$, insbesondere bei Argumenten von Funktionen.

Wenden wir jedoch Matrizen auf Elemente $x \in \mathbb{R}^n$ an, so schreiben wir diese als Spaltenvektoren, z.B.

$$Ax = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad A \in \mathbb{R}^{m \times n}\ ,$$

Ax wird hier als Matrizenprodukt aufgefasst, das Ergebnis ist also eine Spalte im $\mathbb{R}^m$. Durch Transponieren erhält man aus einem Zeilenvektor immer einen Spaltenvektor (und umgekehrt). Zur Not können wir uns auch immer auf die (Vektorraum-) Isomorphismen $\mathbb{R}^n \cong \mathbb{R}^{n\times 1} \cong \mathbb{R}^{1\times n}$ berufen.

25 Differenzierbarkeitsbegriffe

Wir erinnern an den Äquivalenzsatz für Differenzierbarkeit (§18.1.7) für Funktionen auf einem echten Intervall $D \subset \mathbb{R}$. Er besagt insbesondere, dass für Differenzierbarkeit einer Funktion $f : D \to \mathbb{R}$ in $a \in D$ folgende Aussagen äquivalent sind:

(3) Es gibt eine Zahl $l \in \mathbb{R}$ und eine in a stetige Funktion $g : D \to \mathbb{R}$ mit

$$g(a) = 0 \quad \text{und} \quad f(x) = f(a) + l(x-a) + (x-a)g(x) \qquad (\text{für } x \in D)$$

(4) Es gibt eine Zahl $l \in \mathbb{R}$, so dass für den durch die Gleichung

$$f(x) = f(a) + l(x-a) + r(x)$$

definierten Rest $r : D \to \mathbb{R}$ gilt

$$\lim_{x\to a} \frac{r(x)}{x-a} = 0$$

(5) Es gibt eine lineare Abbildung $L : \mathbb{R} \to \mathbb{R}$ mit

$$\lim_{x\to a} \frac{f(x) - f(a) - L(x-a)}{x-a} = 0\ .$$

Hier steht der Gedanke der *linearen Approximation* im Mittelpunkt.

Die „Änderungsrate“ $f(x) - f(a)$ der Funktion wird durch eine lineare Abbildung so gut approximiert, dass für den Fehler sogar $\lim\limits_{x\to a} \frac{r(x)}{x-a} = 0$ gilt.
Hier ist $l = f'(a) = L(1)$ der Wert der Ableitung an der Stelle a.

Im Mehrdimensionalen müssen wir uns an eine neue Definition der Ableitung gewöhnen.
Im Eindimensionalen ist eine lineare Abbildung $L : \mathbb{R} \to \mathbb{R}$ eindeutig bestimmt durch ihren Wert $l := L(1)$, denn es ist

$$L(l) = L(h \cdot 1) = hL(1) = hl = lh$$

und umgekehrt bewirkt jedes $l \in \mathbb{R}$ eine lineare Abbildung $L : \mathbb{R} \to \mathbb{R}$, nämlich $h \mapsto lh$.
Hier unterscheidet man deshalb meist nicht zwischen der linearen Abbildung L und ihrem Wert $l = L(1)$.

25.1 Definition ((totale) Differenzierbarkeit)

Seien $n, m \in \mathbb{N}$ und $D \subset \mathbb{R}^n$ eine offene (nicht-leere) Teilmenge. Eine Abbildung $f : D \to \mathbb{R}^m$ heißt in $a \in D$ (total) differenzierbar, wenn es eine (i.A. a abhängige) lineare Abbildung

$$L = L_a : \mathbb{R}^n \to \mathbb{R}^m$$

gibt, so dass der durch die Gleichung

$$f(x) = f(a) + L(x-a) + r(x)$$

definierte Rest $r : D \to \mathbb{R}$ die Bedingung

$$\lim_{x \to a} \frac{r(x)}{\|x-a\|} = 0$$

erfüllt. Dabei sei $\| \ \|$ irgendeine der (äquivalenten) Normen im $\mathbb{R}^n$.

Beachte: Es gilt also nicht nur $\lim\limits_{x\to a} r(x) = 0$, sondern sogar $\lim\limits_{x\to a} \frac{r(x)}{\|x-a\|} = 0$.

Man sagt: r geht schneller gegen Null als $x \mapsto \|x-a\|$ (beim Grenzübergang $x \to a$).

25.1.1 Bemerkungen

(a) Wir werden gleich sehen, dass L eindeutig bestimmt ist. Man nennt L im Fall der Existenz das *Differential* oder die *Linearisierung* von f in a.

Bezeichnung: $\quad L = df(a) = Df(a) = df|_a \qquad$ o.ä.

Manche Autoren nennen L auch die *Ableitung von f in a* und schreiben $L = f'(a)$. Autoren wie z.B. K.Königsberger verwenden den Begriff Ableitung (Bezeichnung $f'(a)$) Synonym für den Begriff der sog. Funktionalmatrix oder Jacobi-Matrix, die man erhält, wenn man die Matrix-Darstellung der linearen Abbildung L bezüglich der Standardbasen im $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ verwendet.

(b) Setzt man $h := x - a$, so sagt man auch, dass $h \mapsto f(a+h) - f(a)$ durch $h \mapsto Lh$ bei $h = 0$ in *erster Ordnung approximiert* wird.

(c) Die Eindeutigkeit der approximierenden linearen Abbildung ergibt sich einfach so:
Gilt auch $f(x) = f(a) + \tilde{L}(x-a) + \tilde{r}(x)$ mit einer linearen Abbildung $\tilde{L} : \mathbb{R}^n \to \mathbb{R}^m$ und mit $\lim\limits_{x\to a} \frac{\tilde{r}(x)}{\|x-a\|} = 0$, so folgt

$$(L - \tilde{L})(x-a) = \tilde{r}(x) - r(x)$$

und damit auch

$$\lim_{x\to a} \frac{L(x-a) - \tilde{L}(x-a)}{\|x-a\|} = 0 \, .$$

Wählt man nun einen beliebigen Einheitsvektor $v \in \mathbb{R}^n$ und betrachtet diejenigen $x \in D$, für die $x - a = tv$ mit genügend kleinem positiven t gilt, so ergibt sich

$$\lim_{t\to 0} \frac{L(tv) - \tilde{L}(tv)}{t} = 0 \, ,$$

was schon wegen der Homogenität von L natürlich $L(v) = \tilde{L}(v)$ bedeutet.

Da man für v jeden Einheitsvektor von $\mathbb{R}^n$ wählen kann, bedeutet dies $L = \tilde{L}$.

Die Eindeutigkeit der approximierenden linearen Abbildung ergibt sich auch nochmals im Zusammenhang mit den *partiellen* Ableitungen, mit deren Hilfe man die Matrixdarstellung von L (=Jacobi-Matrix von f in a) explizit berechnen kann.

(d) Eine Abbildung $f : D \to \mathbb{R}^m$ heißt *differenzierbar*, wenn f an jeder Stelle $a \in D$ differenzierbar ist.

Im Unterschied zur sog. partiellen Differenzierbarkeit, auf die wir bald zurück kommen, spricht man auch von total differenzierbar und der *totalen* Differenzierbarkeit. Wir lassen das Adjektiv „total“ jedoch meist weg.

Sowohl die lineare Abbildung L also auch der Rest r hängen i.A. von der betrachteten Stelle ab, was man in der Notation meist nicht extra zum Ausdruck bringt.

(e) Statt $f'(a)h$ verwenden manche Autoren die Schreibweise $f'(a,h)$. f' ist dann eine Abbildung von $D \times \mathbb{R}^n \to \mathbb{R}^m$, die bezüglich des zweiten Argumenten linear ist.

25.1.2 Beispiele und Bemerkungen

(a) Ist $f : \mathbb{R}^n \to \mathbb{R}^m$ eine konstante Funktion und gilt etwa $f(x) = c$ für alle $x \in \mathbb{R}^n$, so ist f für alle $x \in \mathbb{R}^n$ differenzierbar und es ist $L = df(x) = 0$ für alle $x \in \mathbb{R}$.

Das Differential ist also die Nullabbildung für alle $x \in \mathbb{R}^n$.

Die Definitionsgleichung
$$f(x) = f(a) + L(x-a) + r(x)$$
ist mit $L = 0$ und $r(x) = 0$ erfüllt.

(b) Ist
$$f : \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^m , \\ x & \mapsto & Ax + b \end{array}$$
mit $A \in \mathbb{R}^{m\times n}$ und $b \in \mathbb{R}^m$, f also eine affine Abbildung, dann gilt: f ist für alle $x \in \mathbb{R}^n$ differenzierbar und es gilt für alle $a \in \mathbb{R}^n$
$$df(a) = A \ .$$
Insbesondere ist das Differential hier unabhängig von a.

Der Beweis ist elementar. Aus $f(x) = Ax + b$ und $f(a) = Aa + b$ folgt
$$f(x) - f(a) = Ax - Aa = A(x-a) \ .$$
Also kann man $L = df(a) := A$ und $r(x) = 0$ für alle $x \in \mathbb{R}^n$ setzen.

Legt man im Fall $n = m = 1$ unsere neue Auffassung der Ableitung zu Grunde, so sind die linearen Funktion
$$\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} , \\ x & \mapsto & cx \end{array}$$
die einzigen, die mit ihrer Ableitungsfunktion übereinstimmen.

(c) Ist $S \in \mathbb{R}^{n\times n}$ eine symmetrische reelle Matrix, also $S = S^t$, so betrachten wir die quadratische Form
$$q_S : \begin{array}{ccl} \mathbb{R}^n & \to & \mathbb{R} , \\ x & \mapsto & q_S(x) = x^t S x \quad \text{(hier ist also } m = 1\text{).} \end{array}$$

Wir schreiben $x = a + h$ und erhalten

$$q_S(a+h) - q_S(a) = 2a^t Sh + h^t Sh \ .$$

Durch $Lh := 2a^t Sh$ wird eine lineare Abbildung(Linearform) $L : \mathbb{R}^n \to \mathbb{R}$ definiert und $R(h) := h^t Sh$ erfüllt die Bedingung $\lim\limits_{h\to 0} \frac{R(h)}{\|h\|} = 0$, denn es ist mit $s := \sum\limits_{j,k=1}^{n} |s_{jk}|$ nämlich

$$|R(h)| \le s\|h\|_\infty^2 \ .$$

q_S ist also in jedem Punkt $a \in \mathbb{R}^n$ differenzierbar und für das Differential gilt

$$df(a)(h) = 2a^t Sh$$

und für die Ableitung (= Matrixdarstellung von $df(a)$ bezüglich der kanonischen Basen in $\mathbb{R}^n$ und $\mathbb{R}$)

$$f'(a) = 2a^t S \ .$$

Nimmt man speziell für S die Einheitsmatrix $E = E^n$, so ergibt sich $f'(a) = 2a^t$.

(d) Ist $f : D \to \mathbb{R}$ eine in a differenzierbare Funktion, so nennt man (in Verallgemeinerung des Begriffes der Tangente im Fall $n = m - 1$) die Menge

$$\{(x, x_{n+1}) \in \mathbb{R}^{n+1};\ x_{n+1} = f(x) + df(a)(x-a)\}$$

die *Tangentialhyperebene* an den Graphen von f im Punkt $(a, f(a))$.

Wie man leicht bestätigt, ist z.B. die Ebene mit der Gleichung $x_3 = 0$ (sie ist isomorph zu $\mathbb{R}^2$) die Tangentialhyperebene an den Graphen der Funktion

$$\begin{array}{rcl} \mathbb{R}^2 & \to & \mathbb{R}\ , \\ (x_1, x_2) & \mapsto & x_1^2 + x_2^2 \end{array}$$

im Punkt $(0, 0)$.

Im $\mathbb{R}^3$ könnte man auch einfach Tangentialebene sagen, denn *jede* Ebene in $\mathbb{R}^3$ ist eine Hyperebene.

25.2 Differenzierbarkeit und Stetigkeit

Ist $f : D \to \mathbb{R}^m$ in $a \in D$ differenzierbar, so ist f stetig in a.

Ist nämlich f differenzierbar in a, so ist per Definition die Abbildung $\varrho : D \to \mathbb{R}^m$ mit

$$\varrho(x) = \begin{cases} \frac{f(x)-f(a)-L(x-a)}{\|x-a\|}, & \text{falls } x \ne a\ , \\ 0, & \text{falls } x = a\ , \end{cases}$$

stetig in a, und es gilt für alle $x \in D$

$$f(x) = f(a) + L(x-a) + \|x-a\|\varrho(x)\ .$$

Weil jede lineare Abbildung stetig ist (es gilt $\|L(x)\| \le C\|x\|$ mit einer geeigneten Konstanten C), setzt sich f aus in a stetigen Abbildungen in einfacher Weise zusammen, f ist also selber stetig in

a.

Aus der (totalen) Differenzierbarkeit folgt also die Stetigkeit. Wie unser früheres Beispiel gezeigt hat, folgt allein aus der Existenz der partiellen Ableitungen an einer Stelle a i.A. nicht die Stetigkeit an der betreffenden Stelle. Wir werden aber bald sehen, dass unter der Voraussetzung, dass die partiellen Ableitungen auch stetig sind, dann auch die (totale) Differenzierbarkeit und damit die Stetigkeit an der betreffenden Stelle folgt.

25.3 Partielle Ableitungen und Jacobi-Matrix

Schon eingangs haben wir den Begriff der partiellen Ableitung für Funktionen von zwei Veränderlichen erläutert. Wir kommen nun systematisch auf diesen Begriff zurück.

Offensichtlich ist eine Abbildung $f : D \to \mathbb{R}^m$ ($D \subset \mathbb{R}^n$ offen) genau dann in $a \in D$ (total) differenzierbar, wenn die Komponentenfunktionen $f_j : D \to \mathbb{R}$ von f $\ (f = (f_1, \dots, f_m))$ für $1 \le j \le m$ in a (total) differenzierbar sind.

Wir können uns also (wenn wir wollen) immer auf reellwertige Funktionen $(m = 1)$ mit Definitionsbereich $D \subset \mathbb{R}^n$ beschränken. Schwieriger ist es jedoch, die Dimension n des Definitionsbereichs zu erniedrigen. Man kann jedoch viele Fragen für Funktionen mehrerer Variablen auf Funktionen einer Veränderlichen zurückführen.

Wir beschreiben einen solchen Spezialisierungsprozess im Folgenden (es ist eine Verallgemeinerung unserer Eingangsüberlegungen).

Sei also $D \subset \mathbb{R}^n$ offen und $a = (a_1, \dots, a_n) \in D$ ein fester Punkt und $f : D \to \mathbb{R}$ eine Funktion. Für jedes j mit $1 \le j \le n$ kann man dann die Menge

$$D_j := \{x_j \in \mathbb{R}; (a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_n) \in D\}$$

betrachten.

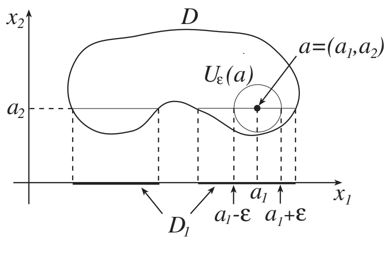
25.3.1 Bemerkung

Ist $D \subset \mathbb{R}^n$ offen, $a = (a_1, \dots, a_n) \in D$, dann ist für jedes j mit $1 \le j \le n$ die Menge

$$D_j := \{x_j \in \mathbb{R}; (a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_n) \in D\}$$

eine nicht leere Teilmenge von $\mathbb{R}$ $(a_j \in D_j)$.

Da D_j offen (und nicht leer) ist, existiert ein $\varepsilon > 0$ mit $U_\varepsilon(a_j) :=]a_j - \varepsilon, a_j + \varepsilon[$.
(siehe Abbildung: Beispiel für n=2)



Ist nun $f : D \to \mathbb{R}$ eine Funktion, so erhält man durch den Spezialisierungsprozess

$$\begin{array}{rccl} f^{[j]} : & D_j & \to & \mathbb{R}\,, \\ & x_j & \mapsto & f(a_1, a_2, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_n) \end{array}$$

Funktionen einer Veränderlichen x_j. D_j enthält insbesondere das offene Intervall $]a_j - \varepsilon, a_j + \varepsilon[$, also eine ε-Umgebung von a_j.

25.3.2 Definition (partielle Differenzierbarkeit)

Die Funktion $f : D \to \mathbb{R}$ heißt im Punkt $a = (a_1, \ldots, a_n) \in D$ nach der j-ten Variable $(1 \leq j \leq n)$ partiell differenzierbar, wenn die Funktion (einer Variablen)

$$\begin{array}{rcl} f^{[j]} : D_j & \to & \mathbb{R}\,, \\ x_j & \mapsto & f(a_1, a_2, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_n) \end{array}$$

im Punkt a_j (im Sinne der Analysis 1) differenzierbar ist.

Schreibweise: $f^{[j]'}(a_j) =: \partial_j f(a_1, \ldots, a_n) = \dfrac{\partial}{\partial_k} f(a_1, \ldots, a_n)$ o.ä.

f heißt in D *partiell differenzierbar*, wenn f in *jedem* Punkt nach *jeder* Variablen partiell differenzierbar ist.

25.3.3 Bemerkungen und Beispiele

(a) Eine direkte Definition der partiellen Ableitung von f nach der j-ten Variablen im Punkt a ist also

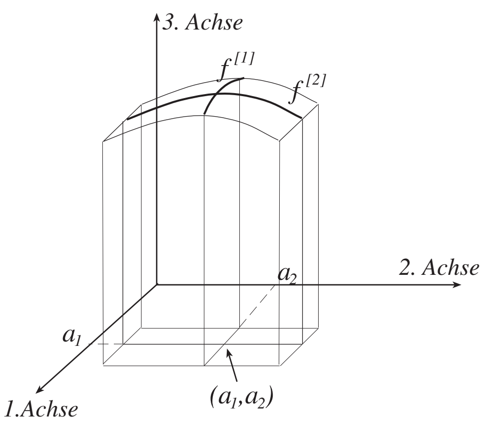
$$\partial_j f(a) := \lim_{t\to\infty} \frac{f(a+te_j) - f(a)}{t}\,,$$

wobei e_j der j-te Einheitsvektor in $\mathbb{R}^n$ ist.

Durch Addition von te_j $(t \in \mathbb{R})$ zu a wird a nur in der j-ten Komponente geändert.

Die partielle Ableitung nach der j-ten Variablen ist also die gewöhnliche Ableitung einer Funktion einer Veränderlichen (bei Festhalten der restlichen). Daher gelten dann auch analoge Rechenregeln.

Die Graphen der $f^{[j]}$ erhält man als Schnitte durch den Graphen von f. (siehe Abbildung)



Geometrisch ist $\partial_j f(a)$ der Anstieg des Graphen von $f^{[j]}$ in a.

(b) Damit man die j-te partielle Ableitung von f im Punkt $a \in D$ definieren kann, ist nicht unbedingt erforderlich, dass D offen ist. Es genügt, dass es wenigstens eine Folge $(t_\nu),\ t_\nu \in \mathbb{R},\ t_\nu \neq 0$, mit $\lim\limits_{t\to\infty} t_\nu = 0$ gibt, für welche $a + t_\nu e_j \in D$ gilt.

Die trifft z.B. zu für einen kompakten Quader

$$[a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n]$$

(wobei die $[a_j, b_j]$ echte (kompakte) Intervalle in $\mathbb{R}$ sind).

Da wir uns jedoch die Freiheit nehmen wollen, das in einer vollen zum Definitionsbereich gehörenden Umgebung des Punktes zu beweisen, setzen wir in der Regel voraus, dass a ein innerer Punkt des Definitionsbereiches ist, was bei offenem D per Definition immer der Fall ist.

(c) Ein einfaches Beispiel:

$$\begin{array}{rcl} f: \mathbb{R}^3 & \to & \mathbb{R}\,, \\ (x,y,z) & \mapsto & x^2\exp(3y)+\cos z\,. \end{array}$$

Hier ist (nach Rechenregeln einer Variablen)

$$\begin{array}{rcl} \partial_1 f(x,y,z) & = & 2x\exp(3y) \\ \partial_2 f(x,y,z) & = & 3x^2\exp(3y) \text{ und} \\ \partial_3 f(x,y,z) & = & -\sin z\,. \end{array}$$

Man verwendet hier auch die Notation

$$\frac{\partial}{\partial x}f(x,y,z),\ \frac{\partial}{\partial y}f(x,y,z) \text{ und } \frac{\partial}{\partial z}f(x,y,z)\,.$$

Aber Vorsicht! Was wäre z.B. $\frac{\partial}{\partial x}f(y,x,x)$?

(d) Existieren in $a=(a_1,\ldots,a_n)$ die partiellen Ableitungen von f nach allen Variablen, dann heißt der Zeilenvektor $(\partial_1 f(a),\ldots,\partial_n f(a))$ der *Gradient von f in a*:

$$grad\ f(a) := (\partial_1 f(a),\ldots,\partial_n f(a))\,.$$

Statt $grad\ f(a)$ findet man in der physikalischen Literatur häufig $\nabla f(a)$ (gesprochen: Nabla f in a).

(e) Existiert in allen Punkten $a\in D$ die partielle Ableitung $\partial_j f(a)$, so kann man $\partial_j f$ wieder als Funktion $\partial_j f: D\to\mathbb{R}$ auffassen ($1\le j\le n$).

f heißt *stetig partiell differenzierbar nach der j-ten Variablen*, wenn $\partial_j f: D\to\mathbb{R}$ stetig ist. Sind alle $\partial_j f$ stetig ($1\le j\le n$), dann heißt f stetig partiell differenzierbar (in D).

(f) In einer Variablen folgt aus der Differenzierbarkeit die Stetigkeit der Funktion an der betreffenden Stelle. Wie schon mehrfach angekündigt, ist dies bei partiell differenzierbaren Funktionen i.A. nicht der Fall, wie das einfache Beispiel $f:\mathbb{R}^2\to\mathbb{R}$ mit

$$f(x,y) = \begin{cases} \frac{2xy}{x^2+y^2}\,, & \text{für } (x,y)\neq(0,0)\,, \\ 0, & \text{für } (x,a)=(0,0)\,, \end{cases}$$

zeigt.

f ist in $(0,0)$ nicht stetig, also erst recht nicht (total) differenzierbar. Die partiellen Ableitungen $\partial_1 f(0,0)$ und $\partial_2 f(0,0)$ existieren jedoch und sind bei Null. Im Punkt $(0,y),\ y\neq 0$, ist $\partial_1 f(0,y)=\frac{1}{y}$. $\partial_1 f$ ist also im Nullpunkt nicht stetig. Ebenso ist $\partial_2 f$ im Nullpunkt unstetig.

Wir wollen uns nun mit dem Zusammenhang zwischen (totaler) Differenzierbarkeit und partieller Differenzierbarkeit intensiver beschäftigen. Wir zeigen zunächst, dass aus der (totalen) Differenzierbarkeit von f in a die partielle Differenzierbarkeit der Komponentenfunktionen folgt und dass das (totale) Differential $df(a)$ bezüglich der kanonischen Basen in $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ die sog. *Jacobi-Matrix* als Darstellungsmatrix besitzt. Insbesondere ergibt sich damit nochmal die Eindeutigkeit der approximierenden linearen Abbildung.

25.3.4 Satz (Zusammenhang zwischen totaler und partieller Differenzierbarkeit)

Sei $D \subset \mathbb{R}^n$ offen, $a \in D$ und $f = (f_1, \ldots, f_n)^t : D \to \mathbb{R}^m$ in $a \in D$ (total) differenzierbar. Dann sind die Komponentenfunktionen $f_k,\ 1 \leq k \leq m$ in a partiell differenzierbar und das Differential $L := df(a)$ von f in a hat bezüglich der Standardbasen in $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ die *Jacobi-Matrix* von f in a (Funktionalmatrix)

$$\mathcal{J}(f;a) := \begin{pmatrix} \partial_1 f_1(a) & \cdots & \partial_k f_1(a) & \cdots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \cdots & \partial_k f_2(a) & \cdots & \partial_n f_2(a) \\ \vdots & & \vdots & & \vdots \\ \partial_1 f_m(a) & \cdots & \partial_k f_m(a) & \cdots & \partial_n f_m(a) \end{pmatrix} = \begin{pmatrix} grad\ f_1(a) \\ grad\ f_2(a) \\ \vdots \\ grad\ f_m(a) \end{pmatrix}$$

als Darstellungsmatrix.

25.3.5 Bemerkung

Da eine lineare Abbildung und eine ihrer Darstellungmatrizen (bzgl. fest gewählter Basen) die gleiche Wirkung auf einer Vektor der Ausgangsraums haben, hätten wir in der Definition 24.1 der totalen Differenzierbarkeit statt mit einer linearen Abbildung auch gleich mit einer Matrix arbeiten können. Wir haben jedoch die koordinatenfreie Definition bevorzugt, die sich auf allgemeinere Situationen (normierte, nicht endlich-dimensionale Vektorräume) verallgemeinern lässt.

Beweis von Satz 24.3.4: Wir bezeichnen mit $S := (e_1, \ldots, e_n)$ die Standardbasis von $\mathbb{R}^n$ und mit $S' := (e'_1, \ldots, e'_m)$ die Standardbasis von $\mathbb{R}^m$.

Nach definition der Matrix einer linearen Abbildung (bzgl. der Basen S un S') ist dann

$$df(a)e_j = \sum_{k=1}^{m} \partial_j f_k(a) e'_k \quad \text{für } 1 \leq j \leq n\ .$$

Das ergibt sich wie folgt: Für $t \in \mathbb{R},\ t \neq 0$ ist

$$\begin{aligned} \sum_{k=1}^{m} \frac{f_k(a+te_j) - f_k(a)}{t} e'_k &= \frac{f(a+te_j) - f(a)}{t} \\ &= \frac{df(a)(te_j) + R(te_j)}{t} \\ &= df(a)e_j \pm \frac{R(te_j)}{\|te_j\|}\ . \end{aligned}$$

Durch Grenzübergang $t \to 0$ $(t \neq 0)$ erhält man auf der rechten Seite $df(a)e_j$ $(1 \leq j \leq n)$.

Also existiert auch der Limes auf der linken Seite (und zwar komponentenweise). Die j-te Spalte der Darstellungsmatrix ist also

$$\begin{pmatrix} \partial_j f_1(a) \\ \vdots \\ \partial_j f_k(a) \\ \vdots \\ \partial_j f_m(a) \end{pmatrix}$$

$\square$

(Man erinnere sich: In der Darstellungsmatrix stehen in der j-ten Spalte die Koordinaten des Bildes des Basisvektors e_j bzgl. der Basis S'.)

25.3.6 Bemerkungen und Beispiele

(a) Die Jacobi-Matrix für eine differenzierbare Abbildung $f : D \to \mathbb{R}^m$ ($D \subset \mathbb{R}^n$ offen) ist also vom Typ $m \times n$(m Zeilen, n Spalten).

Im Spezialfall $m = 1$ (also eine Funktion $f : D \to \mathbb{R}$) ist sie vom Typ $1 \times n$, also eine Zeile mit n Einträgen. Hier ist also

$$\mathcal{J}(f;a) = (\partial_1 f(a), \ldots, \partial_n f(a)) = grad\ f(a)\ .$$

und ihre Wirkung auf einen Spaltenvektor

$$h = (h_1, \ldots, h_n)^t \in \mathbb{R}^n \cong \mathbb{R}^{n\times 1}$$

ist also

$$\begin{aligned} df(a)h := \mathcal{J}(f;a) &= (\partial_1 f(a), \ldots, \partial_n f(a)) \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \\ &= \partial_1 f(a)h_1 + \ldots + \partial_n f(a)h_n \\ &= \langle grad\ f(a)^t\ ,\ h\rangle \end{aligned}$$

(mit dem Standardskalarprodukt $\langle\ ,\ \rangle$ in $\mathbb{R}^n$)

In der Physikliteratur findet man für die Gleichung

$$df(a)h = \partial_1 f(a)h_1 + \ldots + \partial_n f(a)h_n$$

unter Weglassen aller Argumente häufig die Kurzschreibweise

$$df = \partial_1 f dx_1 + \ldots + \partial_n f dx_n,$$

wobei die „Differentiale“ und $dx_1, \ldots, dx_n$ häufig als „unendlich kleine Größen“ interpretiert werden. Eine konkrete Interpretation ist die Interpretation als Linearformen (= Differetialformen 1.Grades oder Pfaffsche Formen) von $\mathbb{R}^n \to \mathbb{R}$.
$dx_j : \mathbb{R}^n \to \mathbb{R}$ ist dabei die Linearform

$$dx_j(h) = h_j,\ 1 \le j \le n,\ h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n.$$

Die dx_j haben die gleiche Wirkung wie die Projektionen

$$\begin{aligned} p_j :\ \mathbb{R}^n &\to \mathbb{R}\ , \\ h &\mapsto h_j\ . \end{aligned}$$

Bemerkung: Mit etwas linearer Algebra kann man $(dx_1, \ldots, dx_n)$ als die zur Standardbasis $S = (e_1, \ldots, e_n) \in \mathbb{R}^n$ duale Basis(Basis des Dualraumes $(\mathbb{R}^n)^* \cong \mathbb{R}^{1\times n}$) interpretieren.

Wir werden uns in diesem Kapitel noch etwas ausführlicher mit Differetialformen 1.Grades (Pfaffsche Formen) beschäftigen (und zwar im Zusammenhang mit Kurvenintegralen).

(b) Um eine konkrete Abbildung $f : D \to \mathbb{R}^m$ ($D \subset \mathbb{R}^n$ offen) in einem Punkt $a \in D$ auf totale Differenzierbarkeit zu untersuchen, bietet sich folgende Strategie an:

Man untersucht die Komponentenfunktionen f_k $(1 \leq k \leq m)$ von f auf partielle Differenzierbarkeit, bildet gegebenenfalls die Jacobi-Matrix $\mathcal{J}(f;a)$ und prüft nach, ob die Gleichung

$$(*) \qquad \lim_{x \to a} \frac{f(x) - f(a) - \mathcal{J}(f;a)(x-a)}{\|x-a\|} = 0$$

gilt. Ist dies der Fall, dann ist f in a (total) differenzierbar.

In vielen Beispielen ist dies der Fall, etwa bei $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x,y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix} = \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix}.$$

Für $a = (a_1, a_2) \in \mathbb{R}^2$ ist

$$\mathcal{J}(f;a) = \begin{pmatrix} grad\ f_1(a) \\ grad\ f_2(a) \end{pmatrix} = \begin{pmatrix} 2a_1 & -2a_2 \\ 2a_2 & 2a_1 \end{pmatrix}$$

Es ergibt sich hier

$$\begin{aligned} f_1(a_1+h_1, a_2+h_2) - f_1(a_1,a_2) - grad\ f_1(a) \cdot \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} &= h_1^2 - h_2^2 \\ f_2(a_1+h_1, a_2+h_2) - f_2(a_1,a_2) - grad\ f_2(a) \cdot \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} &= 2h_1h_2 \end{aligned}$$

Verwendet man die Maximumsnorm in $\mathbb{R}^2$, so sieht man leicht, dass

$$\lim_{h \to 0} \frac{r(h)}{\|h\|} = \lim_{\|h\| \to 0} \begin{pmatrix} \frac{h_1^2 - h_2^2}{\|h\|} \\ \frac{2h_1h_2}{\|h\|} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

gilt.

f ist also für alle $a \in \mathbb{R}^2$ total differenzierbar und

$$f'(a) = \mathcal{J}(f;a) = \begin{pmatrix} 2a_1 & -2a_2 \\ 2a_2 & 2a_1 \end{pmatrix}.$$

Man beachte hier die spezielle Gestalt der Jacobi-Matrix. Sie ist vom Typ

$$\begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) \end{pmatrix} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, \qquad \alpha, \beta \in \mathbb{R},$$

d.h. es ist

$$\boxed{\partial_1 f_1(a) = \partial_2 f_2(a) \qquad \text{und} \qquad \partial_2 f_1(a) = -\partial_1 f_2(a)}$$

Wir kommen auf dass Phänomen (*Cauchy-Riemannsche Differentialgleichungen*) zurück. Es handelt sich in dem Beispiel um der Real- und Imaginärteil der Abbildung

$$\begin{aligned} \mathbb{C} &\to \mathbb{C}, \\ z &\mapsto z^2. \end{aligned}$$

(Schreibe $z = x + yi,\ x, y \in \mathbb{R}$, und quadriere.)

Im Beispiel 24.3.3(e) ist die Jacobi-Matrix(= der Gradient) die Matrix $(0,0)$.

Wenn f in $a = (0,0)$ total differenzierbar wäre, müsste f in $a = (0,0)$ auch stetig sein. f ist aber unstetig in $(0,0)$. Das ist das lang angekündigte Beispiel, dass aus der bloßen Existenz der partiellen Ableitung bzw. der Jacobi-Matrix noch nicht die (totale) Differenzierbarkeit an der betreffenden Stelle folgt. Wie wir gesehen haben, sind die partiellen Ableitungen von f in diesem Beispiel unstetig in $a = (0,0)$.

Im Folgenden beschäftigen wir uns mit dem wichtigen *hinreichenden Kriterium für totale Differenzierbarkeit*, mit dem man das obige Beispiel viel einfacher behandeln kann.

Wir werden sehen:
Existieren für eine Abbildung $f := (f_1, \ldots, f_m)^t : D \to \mathbb{R}^m$ die partiellen Ableitungen der Komponentenfunktionen f_k $(1 \le k \le m)$ in $a \in D$ und sind sie stetig in a, dann ist f total differenzierbar in a.

25.4 Das Hauptkriterium für Differenzierbarkeit

Wie schon angekündigt werden wir jetzt zeigen, dass unter der zusätzlichen Voraussetzung, dass die partiellen Ableitungen der Komponentenfunktionen f_k von f an der Stelle $a \in D$ stetig sind, folgt, dass f in a total differenzierbar ist.

25.4.1 Theorem (Hauptkriterium für Differenzierbarkeit)

Ist $D \subset \mathbb{R}^n$ offen, $a \in D$ und $f = (f_1, \ldots, f_m)^t : D \to \mathbb{R}^m$ eine Abbildung. Existieren die partiellen Ableitungen $\partial_j f_k$ in einer Umgebung von a $(1 \le j \le n,\ 1 \le k \le m)$ und sind diese Funktionen stetig in a, dann ist f total differenzierbar in a.

Beweis : Wir können uns auf den Fall $m = 1$ beschränken, da f genau dann differenzierbar ist, wenn alle Komponentenfunktionen differenzierbar sind.

Die Jacobi-Matrix ist in diesem Fall der Gradient:

$$\mathcal{J}(f;a) = grad\ f(a) = (\partial_1 f(a), \ldots, \partial_n f(a))\ .$$

Die zugehörige Linearform $L : \mathbb{R}^n \to \mathbb{R}$ ist in diesem Fall gegeben durch

$$Lh = grad\ f(a) \cdot h = \sum_{j=1}^{n} \partial_j f(a) h_j \quad (h = (h_1, \ldots, h_n)^t)\ .$$

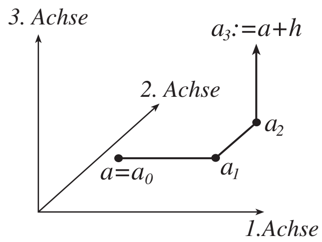
Wir müssen also zeigen:

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - L(h)}{\|h\|} = 0\ .$$

Da D offen ist, gibt es eine offene Würfelumgebung $W := W_\varepsilon(a) \subset D$. Die Idee ist nun, einen beliebigen Punkt $a+h \in W$ mit a durch einen Streckenzug zu verbinden, dessen Teilstrecken parallel zu den Koordinatenachsen verlaufen:

Dazu definieren wir rekursiv die Hilfspunkte $a_0 := a$ und $a_j := a_{j-1} + h_j e_j,\ j = 1, 2, \ldots, n$ (siehe Abbildung). Insbesondere ist dann

$$a_n = a + h_1 e_1 + h_2 e_2 + \ldots + h_n e_n = a + h\ .$$



$\sum\limits_{j=1}^{n} f(a_j) - f(a_{j-1})$ ist eine teleskopische Summen mit dem Wert $f(a+h) - f(a)$.

Also

(1) $f(a+h) - f(a) = \sum\limits_{j=1}^{n} f(a_j) - f(a_{j-1})$.

Die Punkte a_{j-1} und a_j unterscheiden sich nur in der j-ten Koordinate. Die Differenzen in dieser Summe kann man mit Hilfe des Mittelwertsatzes der Differenzialrechnung (jeweils in einer Variablen) wie folgt umformen:

Dazu betrachten wir die Funktionen ($1 \leq j \leq n$)

$$\varphi_j : [0, h_j] \to \mathbb{R} \qquad \text{mit} \qquad \varphi_j(t) = f(a_{j-1} + te_j)\ .$$

Mit diesen Funktionen gilt

(2) $f(a_j) - f(a_{j-1}) = \varphi_j(h_j) - \varphi_j(0)$.

Die Funktionen φ_j sind wegen der partiellen Differenzierbarkeit von f differenzierbar, wobei

(3) $\varphi_j'(t) = \partial_j f(a_{j-1} + te_j)$ gilt.

Auf die Differenz $\varphi_j(h_j) - \varphi_j(0)$ kann man den Mittelwertsatz der Differentialrechnung (im Intervall $[0, h_j]$) anwenden, Es gibt also Zwischenpunkte $t_j \in [0, h_j]$ mit

$$\varphi_j(h_j) - \varphi_j(0) = \varphi_j'(t_j)h_j\ .$$

Setzt man $z_j := a_{j-1} + t_j e_j$, so folgt

(4) $f(a_j) - f(a_{j-1}) = \partial_j f(z_j)h_j$.

Setzt man nun (4) in (1) ein, so folgt wegen $Lh = \sum\limits_{j=1}^{n} \partial_j f(a)h_j$

(5) $f(a+h) - f(a) - Lh = \sum\limits_{j=1}^{n} (\partial_j f(z_j) - \partial_j f(a))h_j$ und damit

(6) $|f(a+h) - f(a) - Lh| \leq \|h\|_\infty \sum\limits_{j=1}^{n} |\partial_j f(z_j) - \partial_j f(a)|$.

Für $h \to 0$ gilt $z_j \to a$ für $1 \leq j \leq n$, was unmittelbar aus der Definition von $z_j = a_{j-1} + t_j e_j$ folgt.

Wegen der Stetigkeit der partiellen Ableitungen $\partial_j f$ an der Stelle a folgt nun (genau an dieser Stelle geht die Stetigkeit ein!)

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - Lh}{\|h\|_\infty} = 0\ .$$

□

25.4.2 Bemerkungen und Beispiele

(a) Da man partielle Ableitungen relativ leicht bilden kann und ihnen häufig die Stetigkeit (ohne Rechnung) ansieht, ist das hinreichende Kriterium aus dem Theorem das am meisten benutzte um eine Abbildung auf (totale) Differenzierbarkeit zu testen.

(b) Es ist jedoch nicht notwendig für die totale Differenzierbarkeit, das heißt eine Abbildung kann auch dann (total) differenzierbar sein, wenn die partiellen Ableitungen nicht stetig sind. Dazu betrachte man z.B.

$$f : \mathbb{R}^2 \to \mathbb{R} \qquad \text{mit} \qquad f(x,y) = \begin{cases} \frac{x^2+y^2}{\sqrt{\sin(x^2+y^2)}}, & \text{falls } (x,y) \neq (0,0)\ , \\ 0, & \text{falls } (x,y) = (0,0). \end{cases}$$

Aufgabe: Zeige, dass f in $a = (0,0)$ (total) differenzierbar ist, die partiellen Ableitungen $\partial_1 f$ und $\partial_2 f$ dort aber unstetig sind.

(c) Aus §24.4.1 folgt, dass alle Polynomfunktion $p : \mathbb{R}^n \to \mathbb{R}$, also Funktionen der Gestalt

$$p(x_1, \ldots, x_n) = \sum a_{j_1,j_2,\ldots,j_n} x_1^{j_1} \cdot \ldots \cdot x_n^{j_n}$$

mit Koeffizienten $a_{j_1}, \ldots, a_{j_n} \in \mathbb{R}$ überall differenzierbar sind, da die partiellen Ableitungen wieder Polynomfunktionen und damit auch stetig sind.

Dieses Beispiel führt zu folgender

25.5 Definition

Ist $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ eine Funktion. Dann heißt f stetig partiell differenzierbar (auf D), wenn die partiellen Ableitungen $\partial_j f$ existieren und die dann erklärten Funktionen $\partial_j f : D \to \mathbb{R}$ stetige Funktionen sind ($1 \leq j \leq n$).

Eine Abbildung

$$f = (f_1, \ldots, f_n)^t : \mathrm{D} \to \mathbb{R}^m$$

heißt stetig partiell differenzierbar, wenn alle Komponentenfunktionen f_k ($1 \leq k \leq m$) stetig partiell differenzierbar sind.

Wir bezeichnen mit $\mathcal{C}^1(D, \mathbb{R}^m)$ die Gesamtheit der stetig partiell differenzierbaren Abbildungen und mit $\mathcal{C}^1(D) := \mathcal{C}^1(D, \mathbb{R})$ die Gesamtheit der stetig differenzierbaren Funktionen auf D. Nach dem Hauptkriterium ist eine $\mathcal{C}^1$-Abbildung f auf ganz D differenzierbar. Sie ist sogar stetig differenzierbar im Sinne der folgende Definition. Es ist offensichtlich, dass die betrachteten Menge $\mathbb{R}$-Vektorräume sind.

25.6 Definition (Stetige Differenzierbarkeit)

Sei $D \subset \mathbb{R}^n$ offen ($\neq \emptyset$) und $f : D \to \mathbb{R}^m$ eine Abbildung. f heißt auf D *stetig differenzierbar*, wenn f für alle $x \in D$ differenzierbar ist und die dann erklärte Abbildung

$$\begin{aligned} f' : \ D \ &\to \ Hom(\mathbb{R}^n, \mathbb{R}^m) \cong \mathbb{R}^{m \times n}\ , \\ x \ &\mapsto \ f'(x) := df(x) \cong \mathcal{J}(f; x) \end{aligned}$$

stetig auf D ist.

25.6.1 Bemerkung

Es sei daran erinnert, dass man auf dem Raum der $m \times n$-Matrizen viele (alle untereinander äquivalente) Normen erklären kann, z.B. die Operator-Norm

$$\begin{aligned} \|A\| &:= \sup\{\|Ax\|;\ \|x\| \le 1\} \quad \text{oder} \\ \|A\|_2 &:= \left(\sum_{k=1}^{m} \sum_{j=1}^{n} |a_{kj}|^2 \right)^{\frac{1}{2}} . \end{aligned}$$

Wir wissen, dass eine Folge von Matrizen genau dann konvergiert, wenn sie elementweise konvergiert. Existiert also f' $(= \mathcal{J}(f,a))$ auf D, so ist die Aussage

„aus $x_l \to a$ folgt $f'(x_l) \Longrightarrow f'(a)\ \ (l \to \infty)$“

nichts anderes als die Aussage

„aus $x_l \to a$ folgt $\partial_j f_k(x_l) \to \partial_j f_k(a)$ für $j = 1, \ldots, n$ und $k = 1, \ldots, m$“ .

f' ist also genau dann in a stetig, wenn die partiellen Ableitungen aller Komponentenfunktionen in a stetig sind.

Damit haben wir

25.6.2 Satz

Ist $D \subset \mathbb{R}^n$ offen, dann ist eine Abbildung $f : D \to \mathbb{R}^m$ genau dann stetig differenzierbar, wenn $f \in \mathcal{C}^1(D, \mathbb{R}^m)$ gilt, d.h. wenn f stetig partiell differenzierbar ist.

Für die eingeführten Differenzierbarkeitsbegriffe gilt also die folgende Äquivalenz und die folgenden Implikationen

stetig differenzierbar
$\Updownarrow$
stetig partiell differenzierbar
$\Downarrow$
(total) differenzierbar
$\Downarrow$
partiell differenzierbar

Die beiden unteren Implikationspfeile lassen sich, wie unsere Beispiele gezeigt haben, i.A. nicht umkehren.

Partiellen Ableitungen sind spezielle *Richtungsableitungen* (Ableitungen in Richtung des Koordinatenachsen). Wir werden später sehen, dass man auch Richtungsableitungen in Richtung eines beliebigen Einheitsvektors v, $\|v\|_2 = 1$, definieren kann. Dabei werden wir sehen, dass auch aus der Existenz *aller* Richtungsableitungen in einem Punkt i.A. nicht die Differenzierbarkeit in diesem Punkt folgt. Umgekehrt folgt jedoch aus der totalen Differenzierbarkeit die Existenz aller Richtungsableitungen (wie wir im nächsten Abschnitt sehen werden).

Man könnte auf die Idee verfallen, dass aus der Stetigkeit von f in a und der Existenz der partiellen Ableitungen in a die totale Differenzierbarkeit folgt. Dies ist aber i.A. auch nicht der Fall, wie das folgende Beispiel zeigt: $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x,y) = \begin{cases} \frac{xy}{\sqrt{x^2+y^2}}, & \text{für } (x,y) \neq (0,0) , \\ 0, & \text{für } x = y = 0 , \end{cases}$$

ist stetig in $a = (0,0)$, die partiellen Ableitungen $\partial_1 f(0,0)$ und $\partial_2 f(0,0)$ existieren und sind Null, also ist $grad\ f(0,0) = (0,0)$, also ist auch $grad\ f(0,0) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, aber der Rest $r(x,y)$ hat nicht die geforderte Eigenschaft. Man weise dies nach.

25.7 Exkurs: Zusammenhang zwischen (total) reeller Differenzierbarkeit in $\mathbb{R}^2$ und komplexer Differenzierbarkeit

In der Analysis 1 haben wir schon den Begriff der Ableitung für Funktionen $f : D \to \mathbb{C}$ definiert, dabei war D etwa ein echtes Intervall in $\mathbb{R}$. Da die komplexen Zahlen einen Körper bilden, kann man den Begriff der *komplexen Differenzierbarkeit* aber auch für Funktionen $f : D \to \mathbb{C}$ ($D \subset \mathbb{C}$ offen) in völliger Analogie zum reellen Fall definieren.

25.7.1 Definition

Sei $D \subset \mathbb{C}$ offen ($\neq \emptyset$). Eine Funktion $f : D \to \mathbb{C}$ heißt in $a \in D$ komplex differenzierbar, falls der Grenzwert

(1) $\lim\limits_{z \to a} \dfrac{f(z) - f(a)}{z - a}$ existiert.

Im Fall der Existenz wird dieser Grenzwert wie üblich mit $f'(a)$ bezeichnet und heißt dann (der Wert der) Ableitung von f an der Stelle a. Ist f in allen Punkten $z \in D$ komplex differenzierbar, dann heißt die Funktion

$$\begin{aligned} f' : D &\to \mathbb{C}\,, \\ z &\mapsto f'(z) \end{aligned}$$

die Ableitung(sfunktion) von f. Man nennt dann f auch *holomorph* oder *analytisch* in D.

Äquivalent mit der obigen Definition ist die Existenz einer komplexen Zahl $l\ (= f'(a))$, für welche dann der durch die Gleichung

(2) $f(z) = f(a) + l(z - a) + r(z)$

definierte Rest $r : D \to \mathbb{C}$ die Eigenschaft

$$\lim_{z \to a} \frac{r(z)}{z - a} = \lim_{z \to a} \frac{r(z)}{|z - a|} = 0$$

hat.

Ist f in a komplex differenzierbar und ist $l := f'(a)$, dann wird durch

$$\begin{aligned} L : \mathbb{C} &\to \mathbb{C}\,, \\ h &\mapsto lh \end{aligned}$$

eine $\mathbb{C}$-lineare Abbildung definiert (dabei ist $l = L(1)$), welche f in a komplex-linear approximiert (bis auf den Fehler r).

Komplexe Differenzierbarkeit von f in a ist also auch äquivalent mit der Existenz einer (von a abhängigen) $\mathbb{C}$-linearen Abbildung $L : \mathbb{C} \to \mathbb{C}$ mit

(3) $\lim\limits_{z \to a} \dfrac{f(z) - f(a) - L(z - a)}{|z - a|} = 0.$

Nun ist aber jede $\mathbb{C}$-lineare Abbildung $L : \mathbb{C} \to \mathbb{C}$ erst recht $\mathbb{R}$-linear. Identifiziert man wie üblich die komplexen Zahlen $\mathbb{C}$ mit $\mathbb{R}^2$ wie

$$x + iy \longleftrightarrow (x, y) \quad \text{oder} \quad x + iy \longleftrightarrow \begin{pmatrix} x \\ y \end{pmatrix},$$

so sieht man sofort, dass eine in $a \in \mathbb{C} = \mathbb{R}^2$ komplex differenzierbare Funktion in a auch (total) reell differenzierbar ist. Uns interessiert die umgekehrte Frage:

Wann folgt aus der (totalen) reellen Differenzierbarkeit von f in a die komplexe Differenzierbarkeit? Oder wann ist eine $\mathbb{R}$-lineare Abbildung $L : \mathbb{C} = \mathbb{R}^2 \to \mathbb{C} = \mathbb{R}^2$ auch $\mathbb{C}$-linear?

M.a.W.: Wann hat L die Wirkung $L(h) = lh$ für alle $h \in \mathbb{C} = \mathbb{R}^2$ mit geeignetem $l \in \mathbb{C} = \mathbb{R}^2$?

Schreibt man $l = \alpha + i\beta,\ \alpha, \beta \in \mathbb{R}$ und $h = x + iy,\ x, y \in \mathbb{R}$, so entspricht

(4) $h \mapsto lh = (\alpha x - \beta y) + (\alpha y + \beta x)i$ der Abbildung $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$

Zerlegt man f in Real- und Imaginärteil $f = u + iv \iff \begin{pmatrix} u \\ v \end{pmatrix}$, so sieht man, dass die Jacobi-Matrix

$$\begin{pmatrix} \partial_1 u(a) & \partial_2 u(a) \\ \partial_1 v(a) & \partial_2 v(a) \end{pmatrix}$$

von f in a eine spezielle Gestalt hat, nämlich vom Typ $\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$ sein muss, falls sie auch $\mathbb{C}$-lineare Abbildung bewirken soll.

Das bedeutet einmal, dass im Fall komplexer Differenzierbarkeit von f in a f (total) reell differenzierbar ist und der Realteil und der Imaginärteil die sog.

Cauchy-Riemannsche DGLn
$\partial_1 u(a) = \partial_2 v(a) \qquad \text{und} \qquad \partial_2 u(a) = -\partial_1 v(a)$

erfüllen müssen und wenn umgekehrt f total reell differenzierbar in a ist ($\mathbb{C} = \mathbb{R}^2$) und wenn u und v in a die Cauchy-Riemannschen DGLn erfüllen, dann ist f in a komplex differenzierbar.

Fassen wir unsere Überlegungen in einem Satz zusammen:

25.7.2 Satz (Zusammenhang zwischen reeller und komplexer Differenzierbarkeit)

Sei $D \subset \mathbb{C}$ offen ($\neq \emptyset$). Eine Funktion $f : D \to \mathbb{C}$ ist in $a \in D$ genau dann komplex differenzierbar, wenn f in a (total) reell differenzierbar ist ($\mathbb{C} = \mathbb{R}^2$) und $u = Re\, f$ und $v = Im\, f$ in a die Cauchy-Riemannschen DGLn erfüllen:

$$\boxed{\partial_1 u(a) = \partial_2 v(a) \qquad \text{und} \qquad \partial_2 u(a) = -\partial_1 v(a)}$$

In diesem Fall gilt

$$f'(a) = \partial u(a) + i\partial_1 v(a) =: \alpha + i\beta$$

sowie

$$det\, \mathcal{J}(f; a) = |f'(a)|^2 = \alpha^2 + \beta^2\ .$$

25.7.3 Bemerkungen

(a) Das systematische Studium analytischer (holomorpher) Funktionen ist Gegenstand der sog. *Funktionentheorie*.

(b) Obwohl die Definition der komplexen Differenzierbarkeit formal wie im Reellen aussieht, ergeben sich gewaltige Unterschiede (z.B. ist eine auf einer offenen Menge $D \subset \mathbb{C}$ $(D \neq \emptyset)$ einmal komplex differenzierbare Funktion beliebig oft komplex differenzierbar).

(c) Definiert man für $f = u + iv$ und $a \in D$

$$\begin{aligned} \partial_1 f(a) &:= \partial_1 u(a) + i\partial_1 v(a) \qquad \text{bzw.} \\ \partial_2 f(a) &:= \partial_2 u(a) + i\partial_2 v(a), \end{aligned}$$

so ist das System der Cauchy-Riemannschen Differentialgleichungen

$$\begin{aligned} \partial_1 u(a) &= \partial_2 v(a) \\ \partial_2 u(a) &= -\partial_1 v(a) \end{aligned}$$

äquivalent mit der Gleichung

$$\boxed{\begin{array}{c} \partial_1 f(a) + i\partial_2 f(a) = 0 \\ \text{oder} \\ \partial_1 f(a) = -i\partial_2 f(a) \end{array}}$$

(d) Häufig definiert man noch (die sog. Wirtinger-Operatoren)

$$\begin{aligned} \partial f(a) &:= \frac{1}{2}(\partial_1 f(a) - i\partial_2 f(a)) \qquad \text{und} \\ \bar{\partial} f(a) &:= \frac{1}{2}(\partial_1 f(a) + i\partial_2 f(a)) \ , \end{aligned}$$

dann gilt:
$f : D \to \mathbb{C}$ ist genau dann analytisch (holomorph), wenn f (total) reell differenzierbar ist in D und für alle $a \in D$ gilt $\bar{\partial} f(a) = 0$.

25.7.4 Einfache Beispiele

(a)

$$\begin{aligned} f : \mathbb{C} &\to \mathbb{C}\,, \\ z &\mapsto z^2 \end{aligned}$$

ist analytisch, denn hier ist $(z = x + iy = (x, y))$

$$\begin{aligned} Re\, f(z) &= u(x, y) = x^2 - y^2 \qquad \text{und} \\ Im\, f(z) &= v(x, y) = 2xy \qquad \text{und es gilt} \end{aligned}$$

$$\begin{array}{ll} \partial_1 u(x, y) = 2x & \partial_2 u(x, y) = 2y \\ \partial_1 v(x, y) = 2y & \partial_2 u(x, y) = 2x \end{array}$$

die Cauchy-Riemannschen Differentialgleichungen sind also überall erfüllt und da u und v stetig partiell differenzierbar sind, ist f (total) reell differenzierbar und für die Ableitung ergibt sich

$$f'(x) = 2x + 2yi = 2(x + yi) = 2z \ ,$$

was man natürlich auch direkt betätigen kann.

(b)

$$\begin{array}{rccl} f: & \mathbb{C} & \to & \mathbb{C}\,, \\ & z & \mapsto & \bar{z} \end{array}$$

ist nirgends komplex differenzierbar, denn die erste Cauchy-Riemannsche DGL ist nirgends erfüllt (es ist $\partial_1 u(x,y) = 1 \neq -1 = \partial_2 v(x,y)$).

(c) Die komplexe Exponentialfunktion

$$\begin{array}{rccl} \exp: & \mathbb{C} & \to & \mathbb{C}\,, \\ & z & \mapsto & \exp(z) = \exp(x)\exp(iy) = \exp(x)(\cos y + i \sin y) \end{array}$$

ist analytisch, und es gilt $\exp'(z) = \exp(z)$ für alle $z \in \mathbb{C}$.

Hier ist $u(x,y) = \exp(x)\cos y$ und $v(x,y) = \exp(x)\sin y$. Die Cauchy-Riemannschen DGLn sind überall erfüllt, u und v sind stetig partiell differenzierbar, damit ist f komplex differenzierbar (in ganz $\mathbb{C}$) und es gilt

$$\begin{aligned} \exp'(z) &= \partial_1 u(x,y) + i\partial_1 v(x,y) \\ &= \exp(x)\cos y + i\exp(x)\sin y \\ &= \exp(x)(\cos y + i \sin y) \\ &= \exp(x)\exp(iy) = \exp(x + iy) = \exp(z) \ . \end{aligned}$$

26 Differentiationsregeln

In diesem Abschnitt werden wir sehen, dass sich die uns aus der Theorie einer Variablen vertrauten Differentiationsregeln (wie z.B. die algebraischen Differentiationsregeln über Summen oder Produkte, sowie die Kettenregel) auf Funktionen von mehreren Variablen übertragen lassen. Die Beweise sind denen in einer Variablen analog, manchmal etwas umständlicher. Die wichtigste Regel ist sicherlich die *Kettenregel*, die ganz grob besagt, dass die Zusammensetzung $g \circ f$ zweier differenzierbaren Abbildungen f und g wieder differenzierbar ist und dass für die Jacobi-Matrizen

$$\mathcal{J}(g \circ f; a) = \mathcal{J}(g; f(a)) \cdot \mathcal{J}(f; a)$$

gilt.

Wir beginnen mit

26.1 Satz (algebraische Regeln)

Sei $D \subset \mathbb{R}^n$ offen, $a \in D$ und sind $f, g : D \to \mathbb{R}^m$ in a differenzierbare Abbildungen, dann sind auch

$$\begin{aligned} f + g : D &\to \mathbb{R}^m \quad \text{und} \\ \alpha f : D &\to \mathbb{R}^m, \ \ \alpha \in \mathbb{R}, \end{aligned}$$

in a differenzierbar und für die Differentiale bzw. die Jacobi-Matrizen gilt

$$\begin{aligned} d(f+g)(a) = df(a) + dg(a) \quad &\text{bzw.} \quad \mathcal{J}(f+g; a) = \mathcal{J}(f; a) + \mathcal{J}(g; a) \\ d(\alpha f)(a) = \alpha df(a) \quad &\text{bzw.} \quad \mathcal{J}((\alpha f); a) = \alpha \mathcal{J}(f; a) \end{aligned}$$

(Linearität der Ableitung)

Produkt- und Quotientenregel kann man nur für Funktionen $f : D \to \mathbb{R}$ formulieren, also im Fall $m = 1$. Dann gilt:

$f \cdot g$ ist in a differenzierbar, und es ist

$$d(f \cdot g)(a) = g(a)df(a) + f(a)dg(a) \qquad \text{(Produktregel)}$$

und falls $g(a) \neq 0$, ist

$$d\left(\frac{f}{g}\right)(a) = \frac{g(a)df(a) - f(a)dg(a)}{(g(a))^2} \qquad \text{(Quotientenregel)}$$

Für die Jacobi-Matrizen gilt also hier

$$\begin{aligned} \mathcal{J}(f \cdot g; a) &= g(a)\mathcal{J}(f; a) + f(a)\mathcal{J}(g; a) \qquad \text{bzw.} \\ \mathcal{J}\left(\frac{f}{g}; a\right) &= \frac{1}{(g(a))^2}\left(g(a)\mathcal{J}(f; a) - f(a)\mathcal{J}(g; a)\right) \end{aligned}$$

Zusatz: Sind f und g in D stetig differenzierbar, dann sind es auch $f + g$ und im Fall $m = 1$ auch $f \cdot g$, ferner $\frac{f}{g}$ in $\{x \in D;\ g(x) \neq 0\}$.

Wenn man beachtet, dass im Fall $m = 1$ $\mathcal{J}(f;a) = grad\, f(a)$ gilt, so haben wir als die *Rechenregeln für den Gradienten*

$$\begin{aligned} grad\,(f \cdot g)(a) &= g(a)\, grad\, f(a) + f(a)\, grad\, g(a) \qquad \text{bzw.} \\ grad\,\left(\frac{f}{g}\right)(a) &= \frac{1}{(g(a))^2}\,(g(a)\, grad\, f(a) - f(a)\, grad\, g(a)) \end{aligned}$$

bewiesen. Diese Rechenregeln gelten übrigens auch, wenn man nur die Existenz des Gradienten $grad\, f(a)$ bzw. $grad\, g(a)$ voraussetzt.

Beweis des Satzes: Wir wissen, dass die Differenzierbarkeit von f in a gleichbedeutend ist mit der Existenz einer linearen Abbildung $L = df(a)$ und einer i.A. stetigen Abbildung $\varrho : D \to \mathbb{R}^m$ mit

$$f(x) = f(a) + df(a)(x-a) + \|x-a\|\varrho(x) \quad \text{und} \quad \varrho(a) = 0.$$

Schreibt man entsprechend

$$g(x) = g(a) + dg(a)(x-a) + \|x-a\|\tilde{\varrho}(x)$$

mit $\tilde{\varrho} : D \to \mathbb{R}^m$, wobei $\tilde{\varrho}$ stetig in a und $\tilde{\varrho}(a) = 0$ gilt, so kann man wie in einer Variablen schließen. Die Linearitätsregeln liest man sofort ab, die Produkt- und Quotientenregel sind etwas aufwendiger zu beweisen.

Dem geneigten Leser / der geneigten Leserin sei empfohlen, etwa die Produktregel zu beweisen und die Quotientenregel im Spezialfall $f(x) = 1$ für alle $x \in D$.

Wir kommen nun zur durchaus wichtigsten Differetiationsregel, der Kettenregel, dass nämlich die Ableitung eines Kompositums das Kompositum der Ableitungen ist.

Genauer gilt

26.2 Theorem (Kettenregel)

$D \subset \mathbb{R}^n$ sei offen und $f : D \to \mathbb{R}^m$ differenzierbar in $a \in D$. $D' \subset \mathbb{R}^m$ sei offen und $g : D' \to \mathbb{R}^l$ sei differenzierbar in $b \in D'$.

Ferner sei $f(D) \subset D'$. Gilt dann $f(a) = b$, dann ist $g \circ f$ in a differenzierbar, und es gilt

$$d(g \circ f)(a) = dg(f(a)) \circ df(a)$$

bzw. für die Jacobi-Matrizen

$$\boxed{\mathcal{J}(g \circ f;a) = \mathcal{J}(g;f(a)) \cdot \mathcal{J}(f;a)} \; ,$$

wobei der Punkt rechts das Matrizenprodukt bedeutet.

Zusatz: Sind g und f stetig differenzierbar, dann ist auch $g \circ f$ stetig differenzierbar.

Diagramm:

$$\begin{tikzcd} D \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & f(D) \subset D' \arrow[r, "g"] & \mathbb{R}^l \end{tikzcd}$$

Beweis : Nach Voraussetzung gilt

$$f(a+h) = f(a) + df(a) + \|h\|\varrho_1(h)$$

mit ϱ_1 stetig in 0 und $\varrho_1(0) = 0$ bzw.

$$g(b+k) = g(b) + dg(b)k\|k\|\varrho_2(k)$$

mit ϱ_2 stetig in 0 und $\varrho_2(0) = 0$.

Setzt man speziell $k := df(a)h + \|h\|\varrho_1(h)$, so folgt

$$(g \circ f)(a+h) = (g \circ f)(a) + (dg(b) \circ df(a))h + R(h)$$

mit $R(h) = \|h\|dg(b)\varrho_1(h) + \|k\|\varrho_2(k)$.

Die Kettenregel ist bewiesen, wenn wir $\lim\limits_{h\to 0} \frac{R(h)}{\|h\|} = 0$ zeigen können.

Da die lineare Abbildung $df(a) : \mathbb{R}^n \to \mathbb{R}^m$ Lipschitz-stetig ist, gilt für k mit einer passenden Konstante C

$$\|k\| \leq \|h\|(C + \|\varrho_1(h)\|) \; .$$

Damit folgt dann $\lim\limits_{h\to 0} \frac{R(h)}{\|h\|} = 0$.

Das beweist die Differenzierbarkeit von $g \circ f$ in a, die Aussage der Kettenregel.

□

Bei Anwendungen der Kettenregel treten besonders häufig die Spezialfälle $n = 1$, m beliebig, $l = 1$ bzw. $m,\ n$ beliebig und $l = 1$.

1. Wir nehmen an, dass $D \subset \mathbb{R}$ ein offenes (echtes) Intervall ist und

$$\begin{array}{rcl} \alpha :\ D & \to & \mathbb{R}^m \; , \\ t & \mapsto & \alpha(t) = (\alpha_1(t), \ldots, \alpha_m(t))^t \end{array}$$

sei eine differenzierbare Abbildung (α entspricht dann f aus der Kettenregel). Ferner sei $g : D' \to \mathbb{R}$ eine Funktion mit

$$D' \supset \; Bild\, \alpha \; .$$

Dann wird durch

$$\begin{array}{rcl} \varphi :\ D & \to & \mathbb{R} \; , \\ t & \mapsto & g(\alpha(t)) = g(\alpha_1(t), \ldots, \alpha_m(t)) \end{array}$$

eine reellwertige Funktion definiert.

Ist nun auch g auf D' differenzierbar, dann ist auch φ auf D differenzierbar, und es gilt für $t \in D$

$$\varphi'(t) = \sum_{j=1}^{m} \partial_j g(\alpha_1(t), \ldots, \alpha_m(t))\, \alpha_j'(t) = grad\, g(\alpha(t)) \cdot \dot{\alpha}(t) \; ,$$

wobei der Punkt das Skalarprodukt in $\mathbb{R}^m$ beschreibt.

Beweis : Für $t \in D$ ist

$$\alpha'(t) = \begin{pmatrix} \alpha_1'(t) \\ \vdots \\ \alpha_m'(t) \end{pmatrix}$$

und auf D'

$$grad\, g(x) = (\partial_1 g(x), \ldots, \partial_m g(x)) \ .$$

Mit der Kettenregel finden wir also

$$\begin{aligned} \varphi'(t) = dg(\alpha(t))\,\alpha'(t) &= (\partial_1 g(\alpha(t)), \ldots, \partial_m g(\alpha(t))) \cdot \begin{pmatrix} \alpha_1'(t) \\ \vdots \\ \alpha_m'(t) \end{pmatrix} \\ &= \sum_{j=1}^{m} \partial_j g(\alpha(t)) \cdot \alpha_j'(t) \ . \end{aligned}$$

□

Aufgabe: Man überlege sich, dass diese Formel auch gilt, wenn D ein nicht offenes (echtes) Intervall in $\mathbb{R}$ ist.

2. Die Kettenregel beinhaltet insbesondere auch Formeln für die partiellen Ableitungen von $g \circ f$. Hat nämlich $g : D' \to \mathbb{R}^l$ die Komponentenfunktionen $g_1, \ldots, g_l$, so sind die Komponentenfunktionen von $g \circ f$ gegeben durch

$$g_1 \circ f,\ g_2 \circ f,\ \ldots\ ,\ g_l \circ f \ .$$

Das (i, j)-te Glied der Jacobi-Matrix $\mathcal{J}(g \circ f; a)$ berechnet sich dann zu

$$\partial_j (g_i \circ f)(a) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}(f(a)) \cdot \frac{\partial f_k}{\partial x_j}(a) \ ,$$

dabei haben wir die m unabhängigen Veränderlichen von g bzw. der g_i mit $y_1, \ldots, y_m$ bezeichnet und haben der Deutlichkeit halber $\frac{\partial f_k}{\partial x_j}(a)$ statt wie meist üblich $\partial_j f_k(a)$ geschrieben.

Insbesondere in älteren Lehrbüchern findet man die folgende *suggestive Schreibweise*: Man setzt

$$\begin{aligned} y_k &:= f_k(x_1, \ldots, x_n) \\ z_i &:= g_i(y_1, \ldots, y_m) \end{aligned}$$

und erhält dann

$$\frac{\partial z_i}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k} \frac{\partial y_k}{\partial x_j} \ .$$

Im Spezialfall $n = 1$ hat man Funktionen $y_k = f_k(t)$ einer Variablen t und man erhält

$$\frac{\partial z_i}{dt} = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k} \frac{\partial y_k}{\partial t} \ .$$

Bei der Benutzung dieser Formeln muss man allerdings die richtigen Argumente einsetzen!

Völlig analog zeigt man: Seien $D \subset \mathbb{R}^n$ und $D' \subset \mathbb{R}^m$ offen,

$$\begin{array}{rcl} g: D' & \to & \mathbb{R}\,, \\ y & \mapsto & g(y)\,, \end{array}$$

sowie

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : D \to \mathbb{R}^m$$

differenzierbare Abbildungen mit $f(D) \subset D'$. Dann ist $h : D \to \mathbb{R}$ mit $h = g \circ f$ partiell differenzierbar, und es gilt für $1 \leq j \leq n$

$$\partial_j h(x_1, \ldots, x_n) = \sum_{k=1}^{m} \frac{\partial g}{\partial y_k}(f_1(x), \ldots, f_m(x)) \cdot \frac{\partial f_k}{\partial x_j}(x_1, \ldots, x_n)\,.$$

Man betrachtet die Jacobi-Matrizen von $h,\ g$ und f. Sie sind

$$\begin{aligned} \mathcal{J}(h;x) &= grad\, h(x) = (\partial_1 h(x), \ldots, \partial_n h(x)) \\ \mathcal{J}(g;f(x)) &= \left(\frac{\partial g}{\partial y_1}(f(x)), \ldots, \frac{\partial g}{\partial y_m}(f(x))\right) \\ \mathcal{J}(f;x) &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix} \end{aligned}$$

Nach der Kettenregel gilt aber

$$\mathcal{J}(h;x) = \mathcal{J}(g \circ f;x) = \mathcal{J}(g;f(x)) \cdot \mathcal{J}(f;x)\,.$$

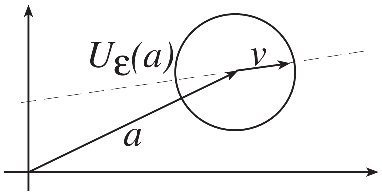
Durch obige Formel werden aber genau die Elemente dieses Matrizenprodukts geliefert.

Das Differential einer differenzierbaren Funktion $f : D \to \mathbb{R}$ ($D \subset \mathbb{R}^n$ offen) liefert eine Approximation für die Änderungsrate der Funktion beim Übergang von $a \in D$ zu $a + h \in D$:

$$f(a+h) - f(a) \approx df(a)h\,.$$

Um das Änderungsverhalten einer Funktion in einer Umgebung von $a \in D$ zu untersuchen, kann man auch eine Gerade durch a betrachten, etwa die durch

$$\begin{array}{rcl} \varphi: \mathbb{R} & \to & \mathbb{R}^n\,, \\ t & \mapsto & a + tv \end{array}$$



definierte. Dabei sei $v \in \mathbb{R}^n$ ein *Richtungsvektor* der Geraden, d.h. ein Vektor $\neq 0$, den wir auf die euklidische Länge 1 normieren: $\|v\|_2 = 1$.

Man kann nun alle möglichen Geraden durch a betrachten und das Änderungsverhalten von f auf solchen Geraden in einer Umgebung von a betrachten. Das führt zu folgender Definition:

26.3 Definition (Richtungsableitung)

Sei $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}$ eine Funktion, $a \in D$ und $v \in \mathbb{R}^n$ mit $\|v\|_2 = 1$ (Richtungsvektor). Existiert der Grenzwert

$$\lim_{t \to 0} \frac{f(a+tv) - f(a)}{t}\,,$$

dann heißt er die Richtungsableitung von f im Punkt a in Richtung v.

Bezeichnung: $\partial_v f(a) := \lim\limits_{t\to 0} \dfrac{f(a+tv)-f(a)}{t}$.

Ist $v = e_j$ der j-te Standardbasisvektor des $\mathbb{R}^n$, dann ist

$$\partial_{e_j} f(a) = \lim_{t\to 0} \frac{f(a+te_j)-f(a)}{t} = \partial_j f(a)$$

gerade die j-te partielle Ableitung von f in a. Partielle Ableitungen sind also spezielle Richtungsableitungen.

Ist $f : D \to \mathbb{R}$ nun eine differenzierbare Funktion, so existieren in allen Punkten $x \in D$ die Richtungsableitungen $\partial_v f(x)$ für jede Richtung v.

Das ist der Inhalt des folgenden Satzes:

26.3.1 Satz

Sei $D \subset \mathbb{R}^n$ offen, $a \in D$ und $f : D \to \mathbb{R}$ in a differenzierbar. Dann existiert die Richtungsableitung $\partial_v f(a)$ für jeden Richtungsvektor v, und es gilt

$$\partial_v f(a) = \partial_1 f(a)v_1 + \partial_2 f(a)v_2 + \ldots + \partial_n f(a)v_m$$

für $v = (v_1, \ldots, v_n)^t$.

Dies kann man mit Hilfe des Skalarproduktes auch so schreiben:

$$\partial_v f(a) = grad\, f(a) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \left\langle \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix}, \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\rangle$$

mit dem Standardskalarprodukt $\langle\ ,\ \rangle$ in $\mathbb{R}^n$.

Wir geben hierfür zwei Beweise:

Der erste beruht auf der Definition der Differenzierbarkeit:
Für alle $t \in \mathbb{R},\ t \neq 0$, deren Betrag hinreichend klein ist, gilt

$$\frac{f(a+tv)-f(a)}{t} = \frac{df(a)(tv) + r(tv)}{t} = df(a)v + \frac{r(tv)}{t} .$$

Wegen $0 = \lim\limits_{t\to 0} \frac{r(tv)}{\|tv\|} = \lim\limits_{t\to 0} \frac{r(tv)}{|t|}$ gilt auch

$$\lim_{t\to 0} \frac{r(tv)}{t} = 0 .$$

Beachtet man $df(a) = grad\, f(a)$, so folgt die Behauptung.

Der zweite Beweis verwendet die Kettenregel:
Ist $\varphi : \mathbb{R} \to \mathbb{R}^n$ definiert durch

$$\varphi(t) = a + tv = (a_1 + t_1, \ldots, a_n + tv_n)$$

(Gerade durch a mit dem Richtungsvektor v, $\|v\|_2 = 1$), dann ist für hinreichend kleine $\varepsilon > 0$

$$\varphi : (] - \varepsilon, \varepsilon[) \subset D,$$

also ist $h := f \circ \varphi :]-\varepsilon, \varepsilon[\to \mathbb{R}$ definiert. Nach der Definition der Richtungsableitung ist

$$\partial_v f(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t} = \lim_{t \to 0} \frac{f(\varphi(t)) - f(\varphi(0))}{t} = h'(0) \ .$$

Nach der Kettenregel ist aber (ersetze $f \to \varphi$ und $g \to f$)

$$h'(t) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\varphi(t)) \cdot \frac{\partial \varphi_j(t)}{dt} \ .$$

Nun ist aber $\varphi_j(t) = a_j + tv_j$, also

$$\varphi_j'(t) = \frac{\partial \varphi_j}{\partial t}(t) = \frac{d}{dt}(a_j + tv_j) = v_j \ ,$$

daher ist

$$h'(0) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a) v_j = grad\, f(a) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \ .$$

□

Auf Grund der Cauchy-Schwarzschen Ungleichung gibt es einen Winkel $\phi \in [0, \pi]$ zwischen den Vektoren $grad\, f(a)$ und v derart, dass

$$\boxed{\partial_v f(a) = \|grad\, f(a)\|_2 \|v\|_2 \cos\phi = \|grad\, f(a)\|_2 \cos\phi}$$

gilt.

Hieraus kann man eine Maximalitätseigenschaft des Gradienten ablesen:

(1) Seine Länge (Norm) $\|grad\, f(a)\|_2$ ist das Maximum aller Richtungsableitungen $\partial_v f(a)$, d.h.

$$\|grad\, f(a)\|_2 = \max\{\partial_v f(a),\ \|v\|_2 = 1\} := M \ ;$$

(2) Ist $M \neq 0$, dann gibt es genau einen Richtungsvektor v_0 mit $\partial_{v_0} f(a) = M$ und mit diesem ist $grad\, f(a) = M v_0$.

Man sagt deshalb: Der Gradient von f im Punkt a zeigt in die Richtung des stärksten Aufstiegs der Funktion im Punkt a.

Ist nämlich $M \neq 0$, so kann man $v_0 := \frac{grad\, f(a)}{\|grad\, f(a)\|_2}$ wählen.

Ist v ein Richtungsvektor, dann ist $-v$ ebenfalls ein solcher und es gilt offensichtlich

$$\partial_{-v} f(a) = -\partial_v f(a) \ .$$

Die Gegenrichtung des Gradienten von f in a ist die Richtung des stärksten Abstiegs und diese ist gegeben durch $-\|grad\, f(a)\|_2$.

Wir beschließen diesen Abschnitt mit zwei weiteren Beispielen:

(a) Wir betrachten $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$f(x, y) = \begin{pmatrix} \exp(x)\cos y \\ \exp(x)\sin y \end{pmatrix} \ .$$

Die Komponentenfunktionen sind also

$$f_1(x,y) = \exp(x)\cos y \qquad \text{und} \qquad f_2(x,y) = \exp(x)\sin y$$

und es gilt

$$\begin{aligned} \partial_1 f_1(x,y) &= \exp(x)\cos y, \quad \partial_2 f_1(x,y) = -\exp(x)\sin y, \\ \partial_1 f_2(x,y) &= \exp(x)\sin y, \quad \partial_2 f_2(x,y) = \exp(x)\cos y. \end{aligned}$$

Damit gilt für die Jacobi-Matrix

$$\mathcal{J}(f;(x,y)) = \begin{pmatrix} \exp(x)\cos y & -\exp(x)\sin y \\ \exp(x)\sin y & \exp(x)\cos y \end{pmatrix} .$$

Da die partiellen Ableitungen sogar $\mathcal{C}^\infty$-Funktionen sind, insbesondere also von der Klasse $\mathcal{C}^1$ sind, ist f in jedem Punkt $(x,y) \in \mathbb{R}^2$ (total) differenzierbar. Man beachte wieder die spezielle Struktur der Jacobi-Matrix, die wieder vom Typ

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$

ist.

Identifiziert man $\mathbb{R}^2$ mit $\mathbb{C}$, so handelt es sich bei der Abbildung f um die komplexe Exponentialfunktion

$$\begin{aligned} \exp : \mathbb{C} &\to \mathbb{C}, \\ z &\mapsto \exp(z) = \exp(x)(\cos y + i\sin y) \end{aligned}$$

($z = x + iy,\ x, y \in \mathbb{R}$).

(b) Ist $M \subset [0,\infty[$ ein echtes Intervall, $F : M \to \mathbb{R}$ eine Funktion und

$$K(M) = \{x \in \mathbb{R}^n;\ \|x\|_2 = \sqrt{x_1^2 + \ldots + x_n^2} \in M\}$$

(eine Kugelschale), so erhält man durch

$$\begin{aligned} f : K(M) &\to \mathbb{R}, \\ x &\mapsto F(\|x\|_2) \end{aligned}$$

eine Funktion, für welche für alle orthogonalen Matrizen A, d.h. für alle $A \in \mathcal{O}(n,\mathbb{R})$, gilt

$$f(x) = f(Ax) .$$

Eine solche Funktion nennt man auch *rotationssymmetrisch*.

Ist nun M ein *offenes* Intervall und F stetig differenzierbar, dann ist auch die Kugelschale $K(M)$ offen (in $\mathbb{R}^n$), dann ist f in jedem Punkt $x \in K(M),\ x \neq 0$, stetig partiell differenzierbar mit den partiellen Ableitungen

$$\partial_j f(x) = F'(\|x\|_2)\frac{x_j}{\|x\|_2}, \quad 1 \le j \le n .$$

Dann ist f in jedem Punkt $x \in K(M),\ x \neq 0$, (stetig) differenzierbar, und es gilt

$$f'(x) = \mathcal{J}(f;x) = \frac{F'(\|x\|_2)}{\|x\|_2} x^t .$$

Wählt man speziell $F = id_M$, dann ist $f(x) = \|x\|_2 =: r(x)$ und unsere Betrachtung zeigt, dass $r : \mathbb{R}^n - \{0\} \to \mathbb{R}$ (stetig) differenzierbar ist und dass

$$\mathcal{J}(r;x) = grad\, r(x) = \frac{1}{\|x\|_2} x^t = \frac{1}{\|x\|_2}(x_1, \ldots, x_n)$$

gilt ($x = (x_1, \ldots, x_n)^t$).

(c) Eng verwandt mit der Abbildung aus Beispiel (a) ist die Abbildung (Polarkoordinaten-Abbildung) $P_2 : \mathbb{R}^2 \to \mathbb{R}$ mit

$$P_2(r,\varphi) = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix} = \begin{pmatrix} f_1(r,\varphi) \\ f_2(r,\varphi) \end{pmatrix} .$$

Hier ist

$$\mathcal{J}(P_2;(r,\varphi)) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

und damit

$$det\mathcal{J}(P_2;(r,\varphi)) = r(\cos^2\varphi + \sin^2\varphi) = r .$$

$\mathcal{J}(P_2;(r,\varphi))$ ist also für $(r,\varphi) \in \mathbb{R}^2$ mit $r \neq 0$ invertierbar.

Berechnen Sie die inverse Matrix!

(d) Eine Abbildung $\alpha = (\alpha_1, \ldots, \alpha_n)^t : M \to \mathbb{R}^m$ eines offenen (echten) Intervalls $M \subset \mathbb{R}$ ist genau dann differenzierbar in $t \in M$, wenn dort jede Komponente α_k differenzierbar ist und dann gilt

$$\alpha'(t) := \dot\alpha(t) = (\dot\alpha_1(t), \ldots, \dot\alpha_n(t))^t ,$$

d.h. die früher gegebene ad-hoc-Definition der Ableitungen einer Kurve stimmt mit der neuen Definition überein.

26.3.2 Anwendung: Orthogonalität von Gradient und Niveaumenge

Im $\mathbb{R}^n$ betrachten wir das Standardskalarprodukt $\langle\ ,\ \rangle$. Ist $f : D \to \mathbb{R}$ eine differenzierbare Funktion auf einer offenen (nicht leeren) Menge $D \subset \mathbb{R}^n$ und $\alpha : M \to D$ eine differenzierbare Kurve ($M \subset \mathbb{R}$ ein offenes, echtes Intervall) und verläuft α ganz in einer Niveaumenge (Fasa) von f, d.h. gibt es ein $c \in \mathbb{R}$ mit $f(\alpha(t)) = c$ für alle $t \in M$. Dann gilt

$$grad\, f(\alpha(t)) \cdot \dot\alpha(t) = \left\langle \begin{pmatrix} \partial_1 f(\alpha(t)) \\ \vdots \\ \partial_n f(\alpha(t)) \end{pmatrix} \cdot \begin{pmatrix} \dot\alpha_1(t) \\ \vdots \\ \dot\alpha_n(t) \end{pmatrix} \right\rangle ,$$

d.h. der Gradient von f im Punkt $\alpha(t)$ und der Tangentialvektor $\dot\alpha(t)$ stehen aufeinander senkrecht. Kurz:

$$\boxed{grad\, f(\alpha(t)) \ \perp\ \dot\alpha(t) \text{ für alle } t \in M}$$

Der Beweis ergibt sich wegen der Konstanz von $h := f \circ \alpha$ sofort aus der Kettenregel:

$$0 = \dot h(t) = grad\, f(\alpha(t)) \cdot \dot\alpha(t) .$$

Beispiele:

(a) Sind $a_1, \ldots, a_n \in \mathbb{R}^n$ und nicht alle a_j gleich Null und ist $c = 0$, und ist $f : \mathbb{R}^n \to \mathbb{R}$ definiert durch

$$f(x_1, \ldots, x_n) = a_1x_1 + \ldots + a_nx_n ,$$

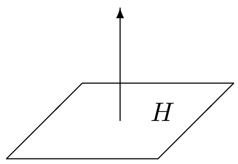
dann ist die Niveaumenge

$$N := \{x \in \mathbb{R}^2;\ a_1x_1 + \ldots + a_nx_n = 0\}$$

eine Hyperebene. Man hat

$$grad\, f(x) = (a_1, \ldots, a_n)$$

für alle $x \in \mathbb{R}^n$. $(a_1, \ldots, a_n)$ ist daher ein *Normalenvektor* für H.



(b) Ist

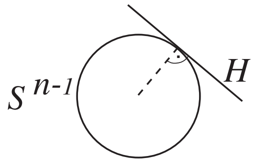
$$\begin{aligned} f:\ \mathbb{R}^n &\to \mathbb{R}\,,\\ x &\mapsto \|x\|_2 = \sqrt{x_1^2+\ldots+x_n^2}\,, \end{aligned}$$

dann ist die Niveaumenge (Ferner zu 1) $\{x \in \mathbb{R}^n;\ f(x) = 1\}$ die Sphäre S^{n-1}.

Sie ist $grad\, f(x) = \frac{1}{\|x\|_2} x^t$ für $x \neq 0$. Insbesondere ist für $x_0 \in S^{n-1}$ $grad\, f(x_0) = x_0^t$, die affine Hyperebene

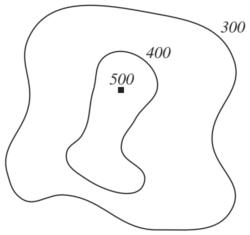
$$H := \{x_0 + v;\ v \in \mathbb{R}^n;\ \langle x_0, v\rangle = 0\}$$

gerade der sog. Tangentialraum an S^{n-1} im Punkt x_0.



Im Spezialfall $n = 2$, kann man den Graphen von $f : D \to \mathbb{R}$ als „Landschaft“ über D mit $f(x)$ als „Höhe“ über dem Punkt x interpretieren. Die Niveaulinien von f sind dann die Höhenlinien der Landschaft (man denke an eine geographische Karte).

Der Gradient von f in a steht nach unserer Feststellung senkrecht auf der Höhenlinie durch x und –wie ebenfalls oben festgestellt– zeigt $f(x)$ in die Richtung des stärksten Anstiegs von f und $-grad\, f(x)$ in die Richtung des stärksten Abstiegs. Ferner ist $\|grad\, f(x)\|_2$ ein Maß für die Steilheit am Ort x.



27 Mittelwertsätze, Schrankensätze

Ein beherrschender Satz der Differentialrechnung einer Veränderlichen ist der Mittelwertsatz (MWSD). Wichtiger noch sind seine Folgerungen (Schrankensatz, Monotonie-Kriterium, Konvexität etc.).

Man kann ihn ohne große Änderungen auf reellwertige Funktionen einer Vektorvariablen übertragen. Er liefert Aussagen über das Änderungsverhalten einer Funktion mit Hilfe von im Definitionsbereich verlaufender Kurven.

27.1 Satz (MWS für reellwertige Funktionen)

Sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ eine differenzierbare Funktion. Ferner seien $a, b \in D$ Punkte, so dass auch ihre Verbindungsstrecke

$$S := S_{ab} := \{a + t(b-a);\ 0 \leq t \leq 1\}$$

in D liegt. Dann gibt es einen Punkt $\xi \in S$ mit

$$\boxed{f(b) - f(a) = grad\, f(\xi) \cdot (b-a)}$$

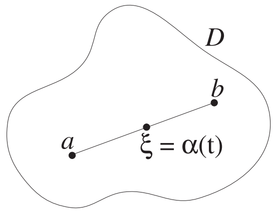
Beweis : Der Beweis ergibt sich unmittelbar aus der „eindimensionalen" Version des Mittelwertsatzes und der Kettenregel.

OBdA sei $a \neq b$. Wir betrachten die „Verbindungskurve"

$$\alpha : [0,1] \to D \quad \text{mit} \quad \alpha(t) = a + t(b-a)\ .$$

Ihre Spur ist gerade die Strecke S_{ab} und es gilt $\dot\alpha(t) = b - a$.

Wir setzen für $t \in [0,1]$ $\varphi(t) := f(\alpha(t))$.



Dann ist φ auf $[0,1]$ differenzierbar und dort gilt nach der Kettenregel

$$\varphi'(t) = (f \circ \varphi)'(t) = grad\, f(\alpha(t)) \cdot \dot\alpha(t) = grad\, f(\alpha(t)) \cdot (b-a)\ .$$

Auf $\varphi : [0,1] \to \mathbb{R}$ kann man den MWSD anwenden:
Es gibt daher ein τ mit $0 \leq \tau \leq 1$ und $\varphi(1) - \varphi(0) = \varphi'(\tau)$, also ist mit $\xi := \alpha(\tau) \in S_{ab}$

$$f(b) - f(a) = grad\, f(\alpha(\tau)) \cdot (b-a) = grad\, f(\xi) \cdot (b-a)\ .$$

□

27.2 Folgerung (Charakterisierung konstanter Funktionen)

Ist $D \subset \mathbb{R}^n$ offen und bogenweise zusammenhängend (das ist äquivalent mit polygon zusammenhängend), dann ist eine differenzierbare Funktion $f : D \to \mathbb{R}$ genau dann konstant, wenn für alle $x \in D$

$$grad\, f(x) = 0$$

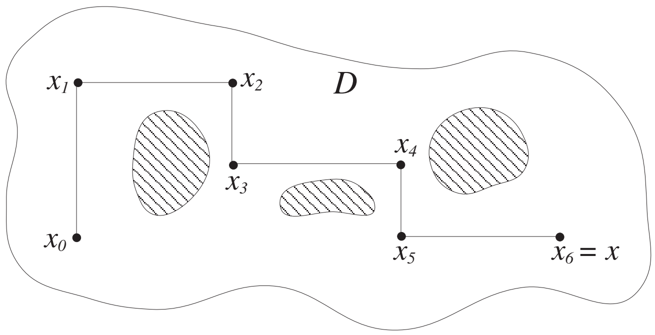
gilt.

Beweis : Wir zeigen nur die nicht triviale Richtung. f ist stetig differenzierbar, also können wir den MWSD anwenden:

Dazu sei $x_0 \in D$ ein fester Punkt. Da D polygon-zusammenhängend ist, gibt es weitere Punkte $x_1, \ldots, x_k := x$ in D, so dass die Verbindungsstrecken $S_{x_{\nu-1}x_\nu}$ für $\nu = 1, \ldots, k$ in D liegen. Nach dem Mittelwertsatz folgt

$$f(x_0) = f(x_1),\ f(x_1) = f(x_2),\ \ldots\ ,\ f(x_{k-1}) = f(x_k) = f(x)$$

und damit $f(x) = f(x_0)$ für beliebige $x \in D$.



□

27.2.1 Eine kleine Anwendung

Für den natürlichen Logarithmus $\log : \mathbb{R}_+^* \to \mathbb{R},\ x \mapsto \log x$ gilt bekanntlich die Funktionsgleichung

$$\log(xy) = \log x + \log y \quad (x, y \in \mathbb{R}_+^*)\ .$$

Wir wollen diese Formel noch etwas anders beweisen. Dazu betrachten wir die Funktion

$$\begin{array}{rcl} f :\ \mathbb{R}_+^* \times \mathbb{R}_+^* & \to & \mathbb{R}\ , \\ (x, y) & \mapsto & \log(xy) \end{array}$$

und

$$\begin{array}{rcl} g :\ \mathbb{R}_+^* \times \mathbb{R}_+^* & \to & \mathbb{R}\ , \\ (x, y) & \mapsto & \log x + \log y\ . \end{array}$$

Man beachte, dass $D := \mathbb{R}_+^* \times \mathbb{R}_+^*$ bogenweise zusammenhängend ist. Für ihre Gradienten gilt

$$\begin{array}{rcl} grad\, f(x, y) & = & \left(\dfrac{1}{x}, \dfrac{1}{y}\right) \quad \text{bzw.} \\ grad\, g(x, y) & = & \left(\dfrac{1}{x}, \dfrac{1}{y}\right)\ . \end{array}$$

Für $h := f - g$ gilt also $grad\, h(x, y) = 0$ für alle $(x, y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$, also ist für alle $(x, y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$

$$h(x, y) = const.$$

mit einer geeigneten Konstanten, d.h. $f(x, y) = g(x, y) + const.$

Zur Ermittlung der Konstanten setzen wir $(x, y) = (1, 1)$ und erhalten

$$0 = \log\, 1 = \log(1 \cdot 1) = \log\, 1 + \log\, 1 + const = 0 + const = const\ ,$$

also haben wir die Funktionalgleichung des Logarithmus mit den Rechenregeln der Differentialrechnung mehrerer Variablen bewiesen.

Der folgende Satz liefert eine Integraldarstellung der Funktionsänderung $f(b) - f(a)$, falls sich a und b durch eine in D verlaufende Kurve verbinden lassen. Der Satz stellt eine Art *Hauptsatz der Differential- und Integralrechnung* in $\mathbb{R}^n$ dar.

27.3 Satz (über die Integraldarstellung des Funktionszuwachses)

Sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ eine $\mathcal{C}^1$-Funktion und $\alpha : [0, 1] \to D$ eine $\mathcal{C}^1$-Kurve mit $\alpha(0) = a$ und $\alpha(1) = b$.

Dann gilt

$$\boxed{f(b) - f(a) = \int_0^1 grad\, f(\alpha(t)) \cdot \dot{\alpha}(t)\ dt}$$

Bemerkung: Das Integral ist das *Kurvenintegral des Vektorfeldes* $grad\, f$ *längs* α.

Beweis : Der Beweis folgt aus dem Hauptsatz der Differential- und Integralrechnung und der Kettenregel:

Es ist nämlich mit $\varphi := f \circ \alpha$

$$f(b) - f(a) = f(\alpha(1)) - f(\alpha(0)) = \int_0^1 \varphi'(t) dt = \int_0^1 grad\, f(\alpha(t)) \cdot \dot{\alpha}(t)\ dt$$

□

27.4 Folgerung (Schrankensatz)

Ist $D \subset \mathbb{R}^n$ offen und $f \in \mathcal{C}^1(D)$ und liegt für $a, b \in D$ auch die Verbindungsstrecke $S := S_{ab}$ in D, dann gilt

$$|f(b) - f(a)| \leq M\|b - a\|_2$$

mit $M := \max\{\|grad\, f(x)\|_2;\ x \in S\}$.

Beweis : Mit $\alpha(t) = a + (b - a),\ 0 \leq t \leq 1$ ergibt sich aus Satz 26.3 mit Hilfe der CSU

$$\begin{aligned} |f(b) - f(a)| = \left| \int_0^1 grad\, f(\alpha(t)) \cdot (b - a)\ dt \right| \quad &\leq \quad \int_0^1 \|grad\, f(\alpha(t))\|_2 \cdot \|b - a\|_2\ dt \\ &\leq \quad M\|b - a\|_2 \end{aligned}$$

□

Bei unseren Sätzen haben wir vorausgesetzt, dass die Dimension des Zielraums eins ist. Betrachtet man Abbildungen

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : D \to \mathbb{R}^m \,,$$

so kann man den MWS auf jede Komponenten f_j anwenden, muss aber in der Regel für jede Komponente einen anderen Zwischenwert ξ nehmen, i.A. gibt es keine Zwischenstellen $\xi \in S = S_{ab}$ mit

$$f(b) - f(a) = \mathcal{J}(f;\xi)(b-a) \,.$$

Ein einfaches Beispiel hierfür ist ($n = 1,\ m = 2$)

$$f : \mathbb{R} \to \mathbb{R}^2 \quad \text{mit} \quad f(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

und $a = 0,\ b = 2\pi$. Dann ist $f(b) - f(a) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, aber für alle $t \in [0, 2\pi]$ ist

$$\mathcal{J}(f;t)(2\pi - 0) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} 2\pi \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \,.$$

Wir formulieren für vektorwertige Funktion $f : D \to \mathbb{R}^m$ einen Mittelwertsatz und einen Schrankensatz:

27.5 Satz (MWS für vektorwertige Funktionen)

Sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}^m$ stetig differenzierbar. Mit $a, b \in D$ sei auch die Strecke

$$S = S_{ab} = \{a + t(b-a);\ 0 \le t \le 1\}$$

in D enthalten. Dann gilt

$$f(b) - f(a) = \int_0^1 \mathcal{J}(f;\, a + t(b-a))dt \cdot (b-a) \,.$$

Dabei wird das Integral über die Jacobi-Matrix komponentenweise gebildet. Das Ergebnis ist wieder eine $m \times n$-Matrix, die im Sinne der Matrizenmultiplikation mit dem Spaltenvektor $b - a$ multipliziert wird.

Beweis : Ist $\varphi_k(t) := f_k(a + t(b-a))$ für $0 \le t \le 1$ und $1 \le k \le m$. Dann ist für $1 \le k \le m$

$$\begin{aligned} f_k(b) - f_k(a) = \varphi_k(1) - \varphi_k(0) &= \int_0^1 \varphi_k'(t)dt \\ &= \int_0^1 \left(\sum_{j=1}^n \partial_j \varphi_k(a + t(b-a))(b_j - a_j) \right) dt \\ &= \sum_{j=1}^n \left(\int_0^1 \partial_j \varphi_k(a + t(b-a))dt \right) (b_j - a_j) \end{aligned}$$

Da die Jacobi-Matrix $\mathcal{J}(f;\, a+t(b-a))$ die Komponenten $\partial_j\varphi_k(a+t(b-a))$ hat, ergibt sich die Behauptung.

$\square$

27.6 Satz (Schrankensatz für Vektorwertige Funktionen)

Unter den Voraussetzungen von Satz 26.5 gilt mit

$$L := \max\{\|\mathcal{J}(f;a+t(b-a))\|_2;\ 0 \le t \le 1\}$$

die Abschätzung

$$\|f(b)-f(a)\|_2 \le L\|b-a\|_2\ .$$

Beweis : Ist $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$, so folgt nach Satz 26.5

$$|f_k(b)-f_k(a)|^2 \le \|b-a\|_2^2 \left(\int_0^1 \|grad\, f_k(\alpha(t))\|dt\right)^2$$

und mit der CSU in der Form

$$\left|\int_0^1 u\cdot 1\right|^2 \le \int_0^1 |u|^2 \cdot \int_0^1 1^2$$

$$\|f(b)-f(a)\|_2^2 \le \|b-a\|_2^2 \int_0^1 \sum_{k=1}^m \|grad\, f_k(\alpha(t))\|_2^2 dt$$

und durch Wurzelziehen ergibt sich

$$\|f(b)-f(a)\|_2 \le L\|b-a\|_2\ .$$

$\square$

28 Höhere partielle Ableitungen, der Vertauschungssatz von H.A.Schwarz

Ist $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ eine partiell differenzierbare Funktion, dann kann man sich fragen, ob die partielle Ableitungen

$$\begin{aligned} \partial_j f : \ D &\to \mathbb{R}\,, \\ x &\mapsto \partial_j f(x) \end{aligned}$$

selber wieder partiell differenzierbar sind.

Man wird f zweimal partiell differenzierbar nennen, falls die n Funktionen

$$\partial_j f : D \to \mathbb{R} \qquad (1 \leq j \leq n)$$

wieder partiell differenzierbar sind, d.h. wenn

$$\begin{array}{ccc} \partial_1\partial_1 f, & \ldots, & \partial_n\partial_1 f \\ \partial_1\partial_2 f, & \ldots, & \partial_n\partial_2 f \\ \vdots & & \vdots \\ \partial_1\partial_n f, & \ldots, & \partial_n\partial_n f \end{array}$$

existieren.

Dieses Spiel kann man weitertreiben.
Man definiert allgemein

28.1 Definition (r-mal partiell differenzierbar)

Seien $n \in \mathbb{N},\ D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ eine Funktion (von n Variablen).

Ist $r \in \mathbb{N}$, so heißt f r-mal partiell differenzierbar genau dann, wenn alle partiellen Ableitungen der Form

$$\partial_{j_r}\partial_{j_{r-1}}\ldots\partial_{j_1} f \quad (j_r, \ldots, j_1 \in \{1, \ldots, n\})$$

existieren in D.

Das bedeutet: Für $j_r, j_{r-1}, \ldots, j_1 \in \{1, 2, \ldots, n\}$ ist f nach der j_1-ten Variablen differenzierbar, $\partial_{j_1} f$ nach der j_2-ten Variablen, $\partial_{j_{r-1}} \ldots \partial_{j_2}\partial_{j_1} f$ nach der j_r-ten Variablen in D partiell differenzierbar.

Es stellt sich dabei sofort die Frage, ob das Ergebnis von der Reihenfolge der Differentiation abhängt, ob also z.B. $\partial_1\partial_2 f$ das gleiche ist wie $\partial_2\partial_1 f$, so dass dann in Wirklichkeit nicht n^2 partielle Ableitungen der Ordnung 2 existieren, sondern lediglich $\frac{1}{2}n(n+1)$, m.a.W. gilt für jeden Punkt $a \in D$, dass die Matrix (sog. Hesse-Matrix)

$$\begin{pmatrix} \partial_1\partial_1 f(a) & \cdots & \partial_n\partial_1 f(a) \\ \vdots & \ddots & \vdots \\ \partial_1\partial_n f(a) & \cdots & \partial_n\partial_n f(a) \end{pmatrix}$$

symmetrisch ist.

Einfache Beispiel zeigen, dass dies manchmal so ist:

(a) Sei z.B. die Funktion

$$\begin{aligned} f : \ \mathbb{R}^2 &\to \mathbb{R}\,, \\ (x, y) &\mapsto x^2 + xy^2 \end{aligned}$$

gegeben.
Hier gilt

$$\begin{array}{rlrl} \partial_1 f(x,y) = & 2x+y^2; & \partial_2 f(x,y) = & 2xy \\ \partial_2\partial_1 f(x,y) = & 2y; & \partial_1\partial_2 f(x,y) = & 2y \end{array}$$

also $\partial_2\partial_1 f(x,y) = \partial_2\partial_1 f(x,y)$.

(b) Das folgende Beispiel zeigt jedoch, dass die „gemischte Ableitungen" i.A. nicht vertauschbar sind.
Sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2}, & \text{falls } (x,y) \neq 0; \\ 0, & \text{falls } (x,y) = 0; \end{cases}$$

Zeigen Sie:

$$\begin{array}{rl} \partial_1 f(0,y) = -y \text{ für alle } y & \text{und} \quad \partial_2\partial_1 f(0,0) = -1, \\ \partial_2 f(x,0) = x \text{ für alle } x & \text{und} \quad \partial_1\partial_2 f(0,0) = 1. \end{array}$$

Die partiellen Ableitungen $\partial_2\partial_1 f(0,0)$ und $\partial_1\partial_2 f(0,0)$ existieren also, sind aber verschieden.

Aus unseren früheren Überlegungen wissen wir, dass allein aus der Existenz der partiellen Ableitungen an einer Stelle noch nicht die (totale) Differenzierbarkeit der betreffenden Funktion in a folgt. Sind die partiellen Ableitungen jedoch stetig in a, so konnten wir auf die (totale) Differenzierbarkeit von f in a schließen. (vgl. §24.7)

Im obigen Beispiel sind die partiellen Ableitungen $\partial_1\partial_2 f$ und $\partial_2\partial_1 f$ in Null nicht stetig (weisen Sie das nach!). Wenn man die Stetigkeit voraussetzt, kann man die Vertauschbarkeit beweisen. Dabei genügt es sogar, um die Existenz und Stetigkeit einer der gemischten Ableitungen vorauszusetzen, dann existiert auch die andere und sie sind gleich. Das ist die Aussage des *Vertauschungssatzes von H.A.Schwarz*:

28.2 Satz (Vertauschungssatz von H.A.Schwarz)

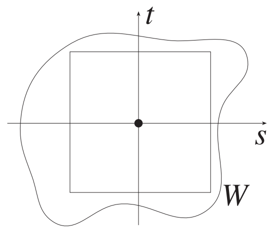
Sei $a \in \mathbb{R}^n$ und die Funktion f besitze in einer offenen Umgebung $U(a) \subset \mathbb{R}^n$ die partiellen Ableitungen $\partial_k f,\ \partial_j f$ und $\partial_k\partial_j f$.Ferner sei $\partial_k\partial_j f$ stetig in a.
Dann existiert auch $\partial_j\partial_k f(a)$ und es gilt.

$$\partial_k\partial_j f(a) = \partial_j\partial_k f(a).$$

Beweis : Um die Bezeichnungen zu vereinfachen setzen wir $\varphi(s,t) := f(a+se_k+te_t)$ $(s,t\in\mathbb{R})$.
Dann besagt die Voraussetzung, dass die partielle Ableitung $\partial_1\partial_2\varphi(0,0)$ existiert. und in $(0,0)$ stetig ist.
Zu zeigen ist die Existenz von $\partial_2\partial_1\varphi(0,0)$ und die Gleichheit $\partial_2\partial_1\varphi(0,0) = \partial_1\partial_2\varphi(0,0)$.
Die Beweisidee beruht auf mehrfacher Anwendung des MWSD auf geeignete Funktionen einer Veränderlichen.
φ ist in einer geeigneten offenen Umgebung W von $(0,0) \in \mathbb{R}^2$ definiert, welche das Quadrat

$$Q = \{(s,t);\ |s| \le r,\ |t| \le r\} \quad (r > 0 \text{ geeignet})$$

enthält.



Nach Definition ist nun

$$\partial_2\partial_1\varphi(0,0) = \lim_{t\to 0}\lim_{s\to 0}\frac{1}{s}\frac{(\varphi(s,t)-\varphi(0,t))-(\varphi(s,0)-\varphi(0,0))}{t}$$

Der Bruch rechts ist ein Differenzenquotient bezüglich der zweiten Variablen ($= t$), auf den man den MWSD anwenden kann. Man erhält

$$\frac{1}{s}\frac{(\varphi(s,t)-\varphi(0,t))-(\varphi(s,0)-\varphi(0,0))}{t} = \frac{1}{s}(\partial_2\varphi(s,\vartheta_2 t)-\partial_2\varphi(0,\vartheta_2 t)),$$

dabei ist $0 < \vartheta_2 < 1$.

Der letzte Ausdruck ist wieder ein Differenzenquotient, jetzt bezüglich der ersten Variablen, auf den man wegen der Existenz von $\partial_1\partial_2$ wieder den MWSD anwenden kann. Man erhält

$$\partial_1\partial_2\varphi(\vartheta_1 s,\vartheta_2 t),\quad 0 < \vartheta_1,\vartheta_2 < 1.$$

wegen der Stetigkeit von $\partial_1\partial_2\varphi$ an der Stelle $(0,0)$ folgt

$$\lim_{t\to 0}\lim_{s\to 0}\partial_1\partial_2\varphi(\vartheta_1 s,\vartheta_2 t) = \partial_1\partial_2\varphi(0,0),$$

also insgesamt $\partial_2\partial_1\varphi(0,0) = \partial_1\partial_2\varphi(0,0)$.

□

Eine Funktion $f : D \to \mathbb{R}$ ($D \subset \mathbb{R}^n$ offen) wird man *r-mal stetig partiell differenzierbar* nennen, wenn sie r-mal partiell differenzierbar ist und alle partiellen Ableitungen der Ordnung $\leq r$ stetig sind.

$\mathcal{C}^r(D)$ sei der Vektorraum der r-mal stetig partiell differenzierbaren Funktionen auf D und wir setzen noch

$$\mathcal{C}^\infty(D) = \bigcap_{r=0}^{\infty}\mathcal{C}^{(r)}(D) \quad \text{(dabei ist } \mathcal{C}^{(0)}(D) = \mathcal{C}(D)\text{)}.$$

Als Folgerung aus dem Satz von Schwarz erhalten wir

28.3 Korollar

Sei $D \subset \mathbb{R}^n$ offen und $f \in \mathcal{C}^r(D)$. Dann gilt

$$\partial_{j_r}\ldots\partial_{j_2}\partial_{j_1}f = \partial_{j_{\pi(r)}}\ldots\partial_{j_{\pi(2)}}\partial_{j_{\pi(1)}}f$$

für jede Permutation $\pi : \{1,\ldots,r\} \to \{1,\ldots,r\}$.

Der Beweis ergibt sich mit vollständiger Induktion nach r und der Erzeugung der symmetrischen Gruppe durch Transpositionen benachbarter Glieder (Nachbarvertauschungen).

Bemerkung: Folgende Schreibweise für die höheren partiellen Ableitungen sind ebenfalls in der Literatur üblich:

$$\partial_k\partial_j f = \frac{\partial^2 f}{\partial x_k\partial x_j},\qquad \partial_j\partial_j f = \partial_j^2 = \frac{\partial^2 f}{\partial x_j^2},\qquad \partial_{j_r}\ldots\partial_{j_1}f = \frac{\partial^r f}{\partial x_{j_r}\ldots\partial x_{j_1}}.$$

Als Anwendung des Satzes von Schwarz wollen wir uns mit der Frage befassen:
„Wann ist ein Vektorfeld ein Gradientenfeld?“

Für viele physikalische Anwendungen ist dies eine wichtige Fragestellung.

28.4 Satz

Sei $D \subset \mathbb{R}^n$ offen und $f = (f_1, \ldots, f_n)^t : D \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld. Höchstens dann gibt es eine stetig differenzierbare Funktion $\varphi : D \to \mathbb{R}$ mit $grad\,\varphi = f$, falls $f = (f_1, \ldots, f_n)$ die *Integralilitätsbedingungen*

$$\partial_k \partial_j f(a) = \partial_j \partial_k f(a), \quad 1 \le j, k \le n, \;\; a \in D,$$

erfüllt, d.h. wenn die Jacobi-Matrix $\mathcal{J}(f; a)$ *von* f *in* a *symmetrisch ist* (für alle $a \in D$).

Beweis : Gilt nämlich

$$grad\,\varphi = (\partial_1 \varphi_1, \ldots, \partial_n \varphi_n) = f = (f_1, \ldots, f_n)$$

mit einem geeigneten φ, dann ist nach dem Satz von Schwarz

$$\partial_k f_j = \partial_k \partial_j \varphi = \partial_j \partial_k \varphi = \partial_j f_k.$$

Halten wir speziell die Fälle $n = 2$ und $n = 3$ fest:

Im Fall $n = 2$, also $f = (f_1, f_2)$ lautet die Integralilitätsbedingungen einfach

$$\partial_2 f_1 = \partial_1 f_2$$

und im Fall $n = 3$

$$\partial_2 f_1 = \partial_1 f_2, \qquad \partial_3 f_1 = \partial_1 f_3, \qquad \partial_3 f_2 = \partial_2 f_3,$$

d.h. für das dem Vektorfeld f zugeordnete Vektorfeld

$$rot\, f := (\partial_2 f_3 - \partial_3 f_2, \partial_3 f_1 - \partial_1 f_3, \partial_1 f_2 - \partial_2 f_1)$$

gilt $rot\, f = 0$.

Also: Damit sich ein stetig partiell differenzierbares Vektorfeld f als Gradient einer stetig differenzierbaren Funktion $\varphi : D \to \mathbb{R}$ darstellen lässt, muss notwendigerweise $rot\, f = 0$ gelten.

□

So ist offensichtlich, dass das Vektorfeld (Wirbelfeld)

$$\begin{array}{rcl} v:\ \mathbb{R}^2 & \to & \mathbb{R}^2\,, \\ (x,y) & \mapsto & \frac{1}{2}(-y, x) = (v_1, v_2) \end{array}$$

kein Potential besitzen kann, denn es gilt

$$\partial_2 v_1 = -\frac{1}{2} \neq \frac{1}{2} = \partial_1 v_2.$$

Das sog. „Windungsfeld“

$$\begin{array}{rcl} w:\ \mathbb{R}^2 - \{(0,0)\} & \to & \mathbb{R}^2\,, \\ (x,y) & \mapsto & \left(-\frac{y}{r^2}, \frac{x}{r^2}\right), \quad r^2 = x^2 + y^2 \end{array}$$

erfüllt zwar die Integralilitätsbedingung, besitzt aber dennoch kein Potential. (Warum nicht?)

Ist jedoch $D \subset \mathbb{R}^n$ ein *Sterngebiet* (bezüglich $O \subset \mathbb{R}^n$, allgemeiner ein einfach zusammenhängendes Gebiet) und ist $v = (v_1, \ldots, v_n) \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld, das die Integralilitätsbedingungen $\partial_j v_k = \partial_k v_j$ $(1 \le j, k \le n)$ erfüllt, dann besitzt v ein Potential, nämlich: $\varphi : D \to \mathbb{R}$ mit

$$\varphi(x) := \sum_{j=1}^{n} \left(\int_0^1 v_j(tx)dt \right) x_j$$

(φ ist das Kurvenintegral von v längst der Verbindungsstrecke vom Nullpunkt zum Punkt $x \in D$ parametrisiert etwa durch $t \mapsto tx,\ 0 \leq t \leq 1$.)

Ist allgemein $D \subset \mathbb{R}^n$ offen ($\neq \emptyset$) und $F = (F_1, \ldots, F_n)^t : D \to \mathbb{R}^n$ ein stetiger Vektorfeld und $\alpha : [a, b] \to D$ eine stetig differenzierbare Kurve, dann heißt die reelle Zahl

$$\int\limits_{\alpha} F \cdot d\alpha = \int\limits_{a}^{b} \langle F(\alpha(t), \dot{\alpha}(t) \rangle dt$$

das *Kurvenintegral* von F längst α.

$\langle\ ,\ \rangle$ ist das Standardskalarprodukt in $\mathbb{R}^n$.

Ist $\alpha = (\alpha_1, \ldots, \alpha_n)^t : [a, b] \to D$ stetig differenzierbar, so ist explizit

$$\int\limits_{\alpha} F \cdot d\alpha = \sum_{j=1}^{n} \int\limits_{a}^{b} f_j(\alpha(t))\dot{\alpha}_j(t)\, dt\ .$$

Ist $\varphi : D \to \mathbb{R}$ stetig differenzierbar, dann gilt für das Vektorfeld $grad\, \varphi : D \to \mathbb{R}^n$

$$\int\limits_{\alpha} grad\, \varphi \cdot d\alpha = \varphi(\alpha(b)) - \varphi(\alpha(a))\ .$$

Wenn ein Vektorfeld also ein Gradientenfeld ist, hängt das Kurvenintegral nur von Anfangs- und Endpunkt der Kurve ab und nicht vom Verlauf dazwischen, insbesondere ist in diesem Fall für eine geschlossene Kurve α (d.h. $\alpha(a) = \alpha(b)$) stets

$$\int\limits_{\alpha} \varphi \cdot d\alpha = \varphi(\alpha(a)) - \varphi(\alpha(a)) = 0\ .$$

29 Taylor-Formel, lokale Extrema

Wir erinnern an den Mittelwertsatz für differenzierbare Funktionen $f : D \to \mathbb{R}$ ($D \subset \mathbb{R}^n$ offen) (vgl. §26.1):
Sind $a, b \in D$ Punkte, für welche auch die Verbindungsstrecke

$$S = S_{a,b} = \{a + t(b-a);\ 0 \le t \le 1\}$$

in D liegt, dann gibt es ein $\vartheta \in]0,1[$ mit

$$\boxed{f(b) - f(a) = grad\, f(a + \vartheta(b-a)) \cdot (b-a)}$$

Der Beweis ergab sich unmittelbar aus der „eindimensionalen“ Version des Mittelwertsatzes und der Kettenregel. Wir wollen nun den Mittelwertsatz zum Satz von Taylor(Taylorsche Formel) verallgemeinern und bedienen uns der gleichen Methode (Zurückführung auf die bekannte Taylorsche Formel einer Veränderlichen).

29.1 Satz (Taylor-Entwicklung zweiter Ordnung)

Sei $D \subset \mathbb{R}^n$ offen, $a \in D$ und $f \in \mathcal{C}^2(D)$. Dann gibt es zu jedem Vektor $h = (h_1, \dots, h_n)^t \in \mathbb{R}^n$, für den auch die Verbindungsstrecke $S_{a,a+h}$ in D liegt, ein $\vartheta \in]0,1[$ mit

$$f(a+h) = f(a) + \sum_{k=1}^{n} \partial_k f(a) h_k + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \partial_j \partial_k f(a + \vartheta h) h_j h_k \ .$$

Schreibt man zur Abkürzung für den Gradienten $grad\, f(x)$ einfach $f'(x)$ und für die Matrix der zweiten partiellen Ableitungen (Hesse-Matrix)

$$H_f(x) = f''(x) = (\partial_j \partial_k f(x)) \ ,$$

dann ist $H_f(x)$ symmetrisch nach dem Satz von Schwarz und die Taylor-Entwicklung erhält die besonders einprägsame Form

$$(*) \qquad \begin{aligned} f(a+h) &= f(a) + grad\, f(a) \cdot h + \frac{1}{2} h^t H_f(a + \vartheta h) h \qquad \text{oder} \\ f(a+h) &= f(a) + f'(a)h + \frac{1}{2} h^t f''(a + \vartheta h) h \ . \end{aligned}$$

Beweis : Wir betrachten die Funktion $\varphi : [0,1] \to \mathbb{R}$ mit $\varphi(t) = f(a + th)$.

Nach der Kettenregel ist zunächst

$$\varphi'(t) = \sum_{k=1}^{n} \partial_k f(a + th) \cdot h_k$$

und damit auch (nochmalige Anwendung der Kettenregel)

$$\varphi''(t) = \sum_{j=1}^{n} \partial_j \left(\sum_{k=1}^{n} \partial_k f(a + th) h_k \right) h_j = \sum_{j=1}^{n} \sum_{k=1}^{n} \partial_j \partial_k f(a + th) h_j h_k \ .$$

Nach der Taylorschen Formel in einer Veränderlichen gilt aber mit $0 < \vartheta < 1$

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2} \varphi''(\vartheta) \ .$$

Beachtet man die Definition $\varphi(t) = f(a + th)$, so folgt die Behauptung.

□

29.2 Bemerkung

Setzt man $f \in \mathcal{C}^{r+1}(D),\ r \geq 1$ voraus, so erhält man aus der Taylorschen Formel

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2}\varphi''(0) + \ldots + \frac{1}{r!}\varphi^{(r)}(0) + \frac{1}{(r+1)!}\varphi^{(r+1)}(\vartheta)$$

Taylorentwicklungen höherer Ordnung, die man mit der *Multiindexschreibweise* (vgl. etwa O.Forster: Analysis 2, §7) einigermaßen übersichtlich schreiben kann.

Für unsere Zwecke –Anwendung auf lokale Extremwerte– reicht jedoch die einfache Form der Taylorschen Formel, der wir für den anschließenden Gebrauch eine leicht veränderte Form geben wollen.

29.3 Satz

Sei $D \subset \mathbb{R}^n$ offen, $f \in \mathcal{C}^2(D)$ und U eine ganz in D liegende Umgebung von $a \in D$. Dann gilt für alle $h = (h_1, \ldots, h_n)^t \in \mathbb{R}^n$ mit $a + h \in U$ die Gleichung

$$\boxed{f(a+h) = f(a) + grad\, f(a) \cdot h + \frac{1}{2}h^t H_f(a)h + \|h\|^2 \varrho(h) \quad \text{mit } \lim_{h\to 0} \varrho(h) = 0.}$$

Bemerkung:
Diese Gleichung verallgemeinert die Definition der (totalen) Differenzierbarkeit von f in a. Ein Vorteil liegt darin, dass die Hesse-Matrix $H_f(a)$ an der Stelle a betrachtet wird und nicht an der Stelle $a + \vartheta h$. Allerdings muss man noch das Restglied $\|h\|^2\varrho(h)$ berücksichtigen.

Wegen $(*)$ gilt

$$f(a+h) = f(a) + grad\, f(a) \cdot h + \frac{1}{2}h^t H_f(a)h + r(h)$$

mit $r(h) := \frac{1}{2}h^t(H_f(a + \vartheta h) - H_f(a))h$.

Da für $0 < x \leq y$ stets $(x, y \in \mathbb{R})$

$$\frac{xy}{x^2 + y^2} \leq \frac{y^2}{x^2 + y^2} \leq 1$$

gilt, folgt für $h \neq 0$ wegen der Stetigkeit der zweiten partiellen Ableitungen, wenn man

$$r(h) = \frac{1}{2} \sum_{j,k=1}^{n} (\partial_j \partial_k f(a + \vartheta h) - \partial_j \partial_k f(a)) \cdot h_j h_k$$

beachtet,

$$\frac{\|r(h)\|_2}{\|h\|_2^2} = \frac{\|r(h)\|_2}{h_1^2 + \ldots + h_n^2} \leq \frac{1}{2} \sum_{j,k=1}^{n} |\partial_j \partial_k f(a + \vartheta h) - \partial_j \partial_k f(a)|$$

und damit

$$\lim_{h\to 0} \frac{\|r(h)\|_2}{\|h\|_2} = 0\ .$$

Wegen $\varrho(h) := \frac{r(h)}{\|h\|^2}$ und der Äquivalenz aller Normen in $\mathbb{R}^n$ folgt dann auch $\lim\limits_{h \to 0} \varrho(h) = 0$.

Lokale Extrema von Funktionen mehrerer Variablen definiert man wie im Fall einer Variablen.

Sei $D \subset \mathbb{R}^n$ beliebig ($\neq \emptyset$). Man sagt $f : D \to \mathbb{R}$ besitzt an der Stelle $a \in D$ *ein lokales Maximum* bzw. Minimum, wenn es eine r-Umgebung $U_r(a)$ von a gibt, so dass für alle $x \in U_r(a) \cap D$ stets

$$f(x) \leq f(a) \qquad \text{bzw.} \qquad f(x) \geq f(a)$$

gilt.

Lokale Maxima bzw. lokale Minima heißen auch *lokale* Extrema und die Stellen $a \in D$, an denen f ein Extremum besitzt, heißen auch *Extrem(al)stellen* von f oder genauer Maximal- oder Minimalstellen von f.

Liegt bei a ein lokales Maximum (bzw. Minimum) vor, so sagt man auch: a sei ein lokaler *Maximierer* oder *Minimierer* von f.

Gilt sogar $f(x) < f(a)$ bzw. $f(x) > f(a)$ für alle $x \in U_r(a) - \{a\}$, so spricht man von einem *strikten* lokalen Maximum bzw. Minimum bei a.

Um den Unterschied zwischen den lokalen (man sagt auch relativen) Extrema einerseits und den *Extrema* andererseits sprachlich besser hervorzuheben, nennt man die letzteren auch häufig *globale* (oder *absolute*) Extrema der betrachteten Funktion. Man mache sich klar, dass eine Funktion weder lokale noch globale Maxima zu besitzen braucht!

Für differenzierbare Funktionen in (echten) Intervallen wissen wir, dass das Verschwinden der Ableitung in einem inneren Punkt eine *notwendige* Bedingung für das Vorliegen eines Extremwertes in diesem Punkt ist (Lemma von Fermat) und dass man etwa mit Hilfe der zweiten Ableitung (falls diese existiert) auch hinreichendes Kriterium erhält.

In mehreren Veränderlichen bestehen gewisse Analogien.
Wir nehmen zunächst einmal an, dass eine Funktion $f : D \to \mathbb{R}$ ($D \subset \mathbb{R}^n$ offen, $\neq 0$) in $a \in D$ (total) differenzierbar ist und in a ein lokales Extremum hat.

Behauptung:
Alle Richtungsableitungen, insbesondere alle partiellen Ableitungen, müssen in a verschwinden.

Da f in a total differenzierbar ist, existieren zunächst einmal alle Richtungsableitungen, d.h. für alle $v \in \mathbb{R}^n$ mit $\|v\|_2 = 1$ existiert

$$\partial_v f(a) := \lim_{t \to 0} \frac{f(a + th) - f(a)}{t} \; .$$

Wir nehmen oBdA an, dass in a ein lokales Maximum vorliegt (sonst Übergang von f zu $-f$). Für alle hinreichend kleine t liegt dann $a + th$ in einer offenen Kugel um a und es gilt $f(a + th) \leq f(a)$ für diese t.

Die in einer Umgebung von 0 definierte Funktion ϱ mit $\varrho(t) = f(a + th)$ hat daher ein lokales Maximum bei $t = 0$. Daher ist $\varrho'(0) = 0$.

Wegen $\varrho'(0) = \partial_v f(a)$ ist also $\partial_v f(a) = 0$.

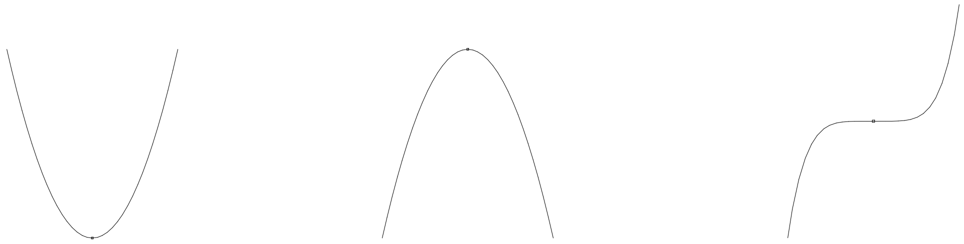
Insbesondere müssen alle partiellen Ableitungen $\partial_1 f(a), \ldots, \partial_n f(a)$ in a verschwinden, d.h. es gilt

$$\boxed{grad\, f(a) = 0}$$

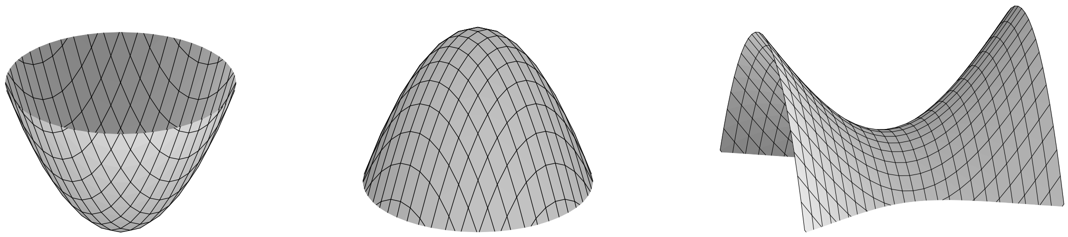
Wenn wir einen Punkt $a \in D$, in welchem die partiellen Ableitungen existieren und in welchem $grad\, f(a) = 0$ gilt, einen *kritischen Punkt* von f nennen, so haben wir also gezeigt:

Ist f in einem inneren Punkt $a \in D$ differenzierbar und besitzt a in D ein lokales Extremum, so ist a ein kritischer Punkt.

Wie in einer Variablen, kann nun in einem kritischen Punkt a ein Maximum, ein Minimum oder keines von beiden vorliegen (Sattelpunkt).



In 2 Variablen könnten die entsprechenden Graphen so aussehen:



Wir erinnern an folgende Kriterien für *positive Definitheit* einer symmetrischen Matrix $H \in \mathbb{R}^{n\times n}$:

29.3.1 Bemerkung

Eine symmetrische Matrix $H \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn ihre Eigenwerte alle positiv sind. Es genügt natürlich sogar diese Bedingung für den kleinsten Eigenwert von H zu fordern.

Wir erinnern kurz an den Beweis:
Zunächst sind Eigenwerte von symmetrischen Matrizen (selbstadjungierte Endomorphismen) stets reell.

Ist nun $\lambda \in \mathbb{R}$ ein Eigenwert von H und x ein Eigenvektor zu λ, d.h. $Hx = \lambda x,\ x \neq 0$, dann ist

$$x^t Hx = x^t \lambda x = \lambda x^t x = \lambda \|x\|_2^2$$

und da $x^t Hx > 0$ für alle $x \neq 0$ nach Voraussetzung gilt, ist auch $\lambda > 0$.

Für die umgekehrte Richtung benutzt man den Satz über die *Hauptachsentransformation*:
Es gibt eine Basis $v_1, \dots, v_n$ aus Eigenvektoren von H,

$$Hv_1 = \lambda_1 v_1, \dots, Hv_n = \lambda_n v_n \ ,$$

und schreibt man einen beliebigen Vektor $u \in \mathbb{R}^n$ in der Form $u = \sum\limits_{j=1}^{n} \gamma_j v_j,\ \gamma_j \in \mathbb{R}$, so ist

$$u^t Hu = \sum_{j=1}^{n} \lambda_j |\gamma_j|^2 \ ,$$

insbesondere ist die rechte Seite > 0 für $u \neq 0$, falls $\lambda_j > 0$ gilt für $1 \leq j \leq n$.
Ist λ_1 der kleinste Eigenwert von H, so gilt (Rayleigh-Prinzip)

$$\lambda_1 = \min\{u^t H u, \|u\|_2 = 1\} \ .$$

λ_1 ist also das Minimum der quadratischen Form $u^t H u$ auf der Einheitssphäre $u^t u = \|u\|_2^2 = 1$.

Hilfssatz:
H ist also genau dann positiv definit, wenn für *alle* $x \in \mathbb{R}^n$ gilt

$$x^t H x \geq \lambda_1 \|x\|_2^2 \ .$$

29.4 Satz

Sei $D \subset \mathbb{R}^n$ offen ($\neq \emptyset$) und $f \in \mathcal{C}^2(D)$.

(a) Hat f in $a \in D$ ein lokales Maximum(Minimum), so ist notwendig $grad\, f(a) = 0$ und die Hesse-Matrix $H = H_f(a)$ ist negativ semidefinit(positiv semidefinit).

(b) Ist $grad\, f(a) = 0$ und ist die Hesse-Matrix $H = H_f(a)$ negativ definit(positiv definit), so besitzt f in a ein lokales Maximum(lokales Minimum).

(c) Nimmt die quadratische Form $x \mapsto x^t H x$ $(H = H_f(a))$ sowohl positive als auch negative Werte an, so hat f in a keine Extremalstelle.

Beweis zu (a): Dass $grad\, f(a) = 0$ sein muss, wissen wir schon. Wegen (vgl. §28.1)

$$g''(0) = \sum_{j=1}^{n} \sum_{k=1}^{n} \partial_j \partial_k f(a) h_j h_k = h^t H_f(a) h$$

gilt im Fall eines Minimums $g''(0) \geq 0$ und daher $h^t H_f(a) h \geq 0$ für alle $h \in \mathbb{R}^n$.

Die Hesse-Matrix muss also positiv semidefinit sein im Fall eines Minimums und negativ semidefinit im Fall eines Maximums.

□

Beweis zu (b): Wegen $grad\, f(a) = 0$ gilt nach der Taylorschen Formel (vgl. §28.3)

$$f(a+h) = f(a) + \frac{1}{2} h^t H_f(a) h + \|h\|_2^2 \varrho(h) \qquad \text{mit} \qquad \lim_{h \to 0} \varrho(h) = 0 \ .$$

Sei zur Abkürzung $Q(h) := h^t H_f(a) h$ und Q positiv definit. Dann gilt für alle $h \neq 0$ mit hinreichend kleiner Norm

$$\frac{f(a+h) - f(a)}{\|h\|_2^2} = \frac{1}{2} Q\left(\frac{h}{\|h\|}\right) + \varrho(h) \qquad \text{mit} \qquad \lim_{h \to 0} \varrho(h) = 0 \ .$$

Da $H_f(a)$ positiv definit ist, gibt es nach dem Hilfssatz ein $\alpha (= \lambda_1) > 0$ mit $Q\left(\frac{h}{\|h\|}\right) \geq \alpha$ für alle $h \neq 0$. Zu diesem α existiert dann wegen $\lim\limits_{h \to 0} \varrho(h) = 0$ ein $\delta > 0$, so dass für $\|h\| < \delta$ stets

$|\varrho(h)| < \frac{\alpha}{4}$ gilt.

Das δ denken wir uns gleich so klein gewählt, dass $U := U_\delta(a) \subset D$ gilt. Da wir jedes $x \in U$ in der Form $x = a + h$ mit $\|h\| < \delta$ schreiben können, ergibt sich für alle $x = a + h \in U_\delta(a)$

$$\frac{f(x) - f(a)}{\|h\|^2} \geq \frac{1}{2}Q(h) - \varrho(h) \geq \alpha - \frac{\alpha}{4} = \frac{3}{4}\alpha > 0 \ .$$

Das Vorzeichen von $f(x) - f(a)$ ist also durch das Vorzeichen der quadratischen Form Q, also das Definitheitsverhalten der Hesse-Matrix bestimmt. f hat in diesem Fall sogar ein striktes lokales Minimum. Ist Q negativ definit, so schließt man völlig analog und erhält, dass f in diesem Fall in a ein striktes lokales Maximum hat.

□

29.5 Beispiele und Bemerkungen

(a) Um eine symmetrische Matrix $H \in \mathbb{R}^{n\times n}$ auf Definitheit zu testen, gibt es neben den schon genannten Kriterien viele weitere, aber meist in der Praxis schwierig anzuwendende, da ein hoher Rechenaufwand damit verbunden ist, z.B. das sog. *Hauptminoren-Kriterium*, das schon von Jacobi aufgestellt wurde:

Eine symmetrische Matrix $H = (h_{\mu\nu})_{\substack{1\leq\mu\leq n\\1\leq\nu\leq n}}$ ist genau dann positiv definit, wenn $det\, H_k > 0$ für $k = 1, 2, \ldots, n$ gilt. Dabei ist $H_k \in \mathbb{R}^{k\times k}$ die linke obere k-reihige Teilmatrix von H $(1 \leq k \leq n)$.

Einen Beweis findet man etwa in *G.Fischer: Lineare Algebra, 13.Aufl. S.327*.

Im Fall $n = 2$ ist dieses Kriterium aber nützlich und elementar zu beweisen:
Ist

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & \delta \end{pmatrix} \in \mathbb{R}^{2\times 2} \ ,$$

dann ist A genau dann positiv definit, wenn $\alpha > 0$ und $det\, A = \alpha\delta - \beta^2 > 0$ gilt. Ist $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, dann ist (quadratische Ergänzung)

$$\alpha \cdot x^t Ax = \alpha^2 x_1^2 + 2\alpha\beta x_1 x_2 + \alpha\delta x_2^2 = (\alpha x_1 + \beta x_2)^2 + (\alpha\delta - \beta^2)x_2^2 \ .$$

Aus dieser Identität kann man das Kriterium unmittelbar ablesen.

(b) Statt alle Eigenwerte (Nullstellen des charakteristischen Polynoms) auf Positivität zu testen, genügt es die Koeffizienten des charakteristischen Polynoms zu ermitteln und die *Vorzeichenregel von Descartes* anzuwenden.
Ist $H \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix und

$$P_H(t) = (-1)^n t^n + \alpha_{n-1}t^{n-1} + \ldots + \alpha_1 t + \alpha_0$$

ihr charakteristisches Polynom. Dann gilt

$$H \text{ ist positiv definit} \qquad \Longleftrightarrow \qquad (-1)^j\alpha_j > 0 \text{ für } j = 0, \ldots, n-1 \ .$$

Einen Beweis findet man ebenfalls in dem zitierten Buch von G.Fischer.

(c) Da man eine positiv definite Matrix durch simultane Zeilen- und Spaltentransformationen in die Einheitsmatrix überführen kann, gilt:

$H \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn es eine *invertierbare* untere Dreiecksmatrix L gibt, so dass $H = LL^t$ gilt.

Eine Variante davon ist (Cholesky-Zerlegung):

H ist genau dann positiv definit, wenn es eine untere Dreiecksmatrix W mit positiven Diagonalelementen gibt, so dass $H = WW^t$ gilt, z.B. gilt etwa

$$\begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & 180 \\ 30 & -180 & 180 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ -12 & 4\sqrt{3} & 0 \\ 10 & -5\sqrt{3} & \sqrt{5} \end{pmatrix} \begin{pmatrix} 3 & -12 & 10 \\ 0 & 4\sqrt{3} & -5\sqrt{3} \\ 0 & 0 & \sqrt{5} \end{pmatrix} ,$$

also ist die Matrix links positiv definit.

Beweise hierzu findet man in Büchern über Numerik, z.B. in G.Opfer: Numerische Mathematik für Anfänger, Vieweg-Studium, 3.Aufl. 2001.

Aus dem Satz über die Hauptachsentransformation (oder auch nur mit vollständiger Induktion über die Dimension) ergibt sich der

29.6 Sylvesterscher Trägheitssatz

Ist q eine quadratische Form auf einem n-dimensionalen $\mathbb{R}$-Vektorraum V, dann gibt es immer eine Basis $F = (v_1, \ldots, v_n)$ von V, für welche die Matrix A von q die Gestalt

$$A = \begin{pmatrix} +1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & +1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ 0 & & & & & & & & 0 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} p \\ \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} q \\ \vphantom{\begin{matrix}1\\1\\1\end{matrix}} \end{array}$$

hat ($0 \leq p, q,\ p + q \leq n$).

Eine solche Basis nennt man auch Sylvesterbasis. Ferner ist die Anzahl p der Diagonalglieder $+1$ und die Anzahl q der Diagonalglieder -1 unabhängig von der Wahl der Sylvesterbasis. Insbesondere ist p auch die Anzahl der positiven Eigenwerte, q ist die Anzahl der negativen Eigenwerte, $n - (p + q)$ ist die Anzahl der Eigenwerte, die gleich Null sind.

In einer solchen Basis hat die quadratische Form die durchsichtige Gestalt:

$$q(x_1v_1 + x_2v_2 + \ldots + x_nv_n) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2 \,.$$

q ist also genau dann positiv definit, wenn $p = n$ (also $q = 0$) ist, die Sylvestersche Normalform ist dann also die Einheitsmatrix E_n.

Es gibt einen schnellen Algorithmus zur Bestimmung des Paares (p, q). Dieses Paar nennt man häufig auch die *Signatur* der quadratischen Form, die Bezeichnungen in der Literatur sind allerdings nicht einheitlich und manchmal verwirrend.

Dazu geht man aus von der Matrix A von q bezüglich irgendeiner Basis von V. Dann kann man

A in Diagonalform transformieren (i.A. allerdings nicht orthogonal), indem man elementare Spaltenumformungen (Linksmultiplikation mit Elementarmatrizen) und die entsprechende Zeilenumformungen verwendet. Das Ergebnis ist dann wieder eine Darstellungsmatrix der quadratischen Form bezüglich einer Basis von V, auf die es aber bei der Signaturbestimmung nicht ankommt.

Beispiel:
Wir betrachten die quadratische Form $q : \mathbb{R}^2 \to \mathbb{R}$, deren Matrix bezüglich der Standardbasis (e_1, e_2) von $\mathbb{R}^2$ durch

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

gegeben ist.

Es ist also $q(x_1e_1 + x_2e_1) = 2x_1x_2$.
Welche Signatur hat A?
Zunächst addieren wir die zweite Zeile zur ersten und dann die zweite Spalte zur ersten, das ergibt

$$A' = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} .$$

Jetzt subtrahieren wir die Hälfte der ersten Zeile von der zweiten und dann die Hälfte der ersten Spalte von der zweiten, das ergibt

$$A'' = \begin{pmatrix} 2 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} .$$

Ein letzter Umformungsschritt ergibt

$$A''' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

Also hat q die Signatur $(1, -1)$; q ist daher indefinit.

29.7 Beispiele

(a) Wir wollen das *globale* Maximum und Minimum der Funktion

$$\begin{aligned} f:\ [0,1]\times[0,1] &\to \mathbb{R}\ , \\ (x,y) &\mapsto 1+4x+y-2(1+y^2)x^2 \end{aligned}$$

bestimmen.

Sinnvollerweise betrachtet man zunächst die Einschränkung von f auf die 4 Randstrecken des Quadrats $Q := [0,1] \times [0,1]$, also die vier Funktionen

$$\begin{aligned} f_1(y) &:= f(0,y) = 1+y\ , \\ f_2(y) &:= f(1,y) = 3+y-2y^2 \qquad \text{für } 0 \le y \le 1\ , \end{aligned}$$

und

$$\begin{aligned} f_3(x) &:= f(x,0) = 1+4x-2x^2\ , \\ f_4(x) &:= f(x,1) = 2+4x-4x^2 \qquad \text{für } 0 \le x \le 1\ . \end{aligned}$$

Maximum und Minimum erhält man mit den Methoden einer Variablen, wir geben das Ergebnis in Form einer Tabelle an.

	Maximum		Minimum	
	Stelle	Wert	Stelle	Wert
$f(0,y)$	$(0,1)$	2	$(0,0)$	1
$f(1,y)$	$\left(1,\frac{1}{4}\right)$	$3,125$	$(1,1)$	2
$f(x,0)$	$(1,0)$	3	$(0,0)$	1
$f(x,1)$	$\left(\frac{1}{2},1\right)$	3	$(0,1),\ (1,1)$	2

Die weiteren Kandidaten für Extremstellen sind die kritischen Punkte von f in Inneren von Q.

Es gilt

$$grad\, f(x,y) = (4 - 4(1+y^2),\ 1 - 4x^2y)\ ,$$

also $grad\, f(x,y) = 0$ genau dann, wenn

$$(1+y^2) = 1 \qquad \text{und} \qquad 4x^2y = 1\ .$$

Durch Elimination von $y = \frac{1}{4x^2}$ aus diesem nicht linearen Gleichungssystem erhält man für x eine quadratische Gleichung 4.Grades

$$x^4 - x^3 + \frac{1}{16} = 0\ .$$

$x = \frac{1}{2}$ ist offensichtlich eine Nullstelle und führt auf $y = 1$, $(x,y) = \left(\frac{1}{2},1\right)$ ist ein Randpunkt von Q.

Polynomdivision ergibt

$$x^4 - x^3 + \frac{1}{16} = \left(x - \frac{1}{2}\right)\left(x^3 - \frac{1}{2}x^2 - \frac{1}{4}x - \frac{1}{8}\right)\ .$$

Eine Visualisierung mit MAPLE oder eine Kurvendiskussion ergibt, dass das Polynom

$$x \mapsto x^3 - \frac{1}{2}x^2 - \frac{1}{4}x - \frac{1}{8}$$

nur eine reelle Nullstelle x_0 besitzt und dass $x_0 = \frac{9}{10}$ ein Näherungswert ist, den man auch mit der „Cardanischen Formel“

$$x_0 = \frac{1}{6}\left(1 + \sqrt[3]{19+\sqrt{297}} + \sqrt[3]{19-\sqrt{297}}\right)$$

erhalten kann.

$x_0 = \frac{9}{10}$ ergibt $y_0 = \frac{3}{10}$ und $f(x_0,y_0) > 3,13$. Vergleicht man dies mit der obigen Tabelle, so ergibt sich:
f hat ein absolutes Minimum im Punkt $(0,0)$ mit dem Wert $f(0,0) = 1$ und ein absolutes Maximum im Punkt (x_0,y_0) mit dem Wert $f(x_0,y_0) > 3,13$.

Insbesondere ergibt sich auch, dass in (x_0,y_0) ein lokales Maximum vorliegt, was man auch mit Hilfe der Hesse-Matrix leicht nachprüfen kann, denn diese ist im Punkt (x_0,y_0) negativ definit:

$$H_f(x_0,y_0) = \begin{pmatrix} -\frac{4}{x_0} & -\frac{2}{x_0} \\ -\frac{2}{x_0} & -\frac{1}{y_0} \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}\ .$$

Hier ist nämlich $a = -\frac{4}{x_0} < 0$ und

$$ac - b^2 = det\, H_f(x_0,y_0) = \frac{4}{x_0y_0} - \frac{4}{x_0} > 0\ .$$

Die Signatur (p, q) einer einer quadratischen Matrix A zugeordneten quadratischen Form $Q_A(x) = x^t Ax$ lässt sich mit einem schnellen Algorithmus bestimmen.

29.8 Bemerkung

Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ ist offensichtlich genau dann positiv definit, wenn für irgendeine invertierbare Matrix $P \in \mathbb{R}^{n\times n}$ gilt $P^t AP$ ist positiv definit.

Ist P eine *Elementarmatrix*, so entsteht AP aus A durch die entsprechende elementare Spaltenumformung.

$P^t AP$ entsteht aus A durch elementare Spaltenumformung und die entsprechende elementare Zeilenumformung desselben Typs. Einen derartigen doppelten Umformungsschritt wollen wir eine elementare Spalten/Zeilenumformung nennen (kurz S/Z-Umformung).

Unter einer Spalten/Zeilenumformung schlecht hin, wollen wir eine endliche Abfolge von elementaren Spalten/Zeilenumformungen verstehen.

29.9 Satz

Jede symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ lässt sich durch Spalten/Zeilenumformungen in eine Diagonalmatrix überführen, d.h. es gibt ein $P \in Gl(n, \mathbb{R})$ mit

$$P^t AP = Diag(d_1, \ldots, d_n) .$$

Der Beweis ist einfach mit Induktion nach n zu führen.

Will man entscheiden, ob A positiv definit ist, so kann man das an den Diagonalelementen $d_1, \ldots, d_n$ ablesen. A ist genau dann positiv definit, wenn alle $d_j,\ 1 \le j \le n$, positiv sind.

Ist p die Anzahl der positiven Elemente in $Diag(d_1, \ldots, d_n)$, q die Anzahl der negativen Elemente in $Diag(d_1, \ldots, d_n)$, dann ist (p, q) die *Signatur* der quadratischen Form $Q_A(x) = x^t Ax$.

Man kann den symmetrischen Diagonalisierungsalgorithmus also auch benutzen, um den Typ der durch

$$Q_A(x) = x^t Ax + \alpha^t x + \beta = 0 ,$$

$A = A^t \in \mathbb{R}^{n\times n},\ \alpha \in \mathbb{R}^n,\ \beta \in \mathbb{R}$ definierten *Quadrik* zu erkennen.

Wendet man die bei der S/Z-Diagonalisierung auftretenden Spaltenoperationen auf die Einheitsmatrix E_n an, erhält man eine Matrix $P \in Gl(n, \mathbb{R})$ mit

$$P^t AP = Diag(d_1, \ldots, d_n) .$$

Wir betrachten nochmals ein Beispiel:

$$A = \begin{pmatrix} 0 & 1 & 3 \\ 1 & 0 & 2 \\ 3 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3} .$$

Die Matrix A ist sicher nicht positiv definit, weil bei einer positiv definiten Matrix alle Diagonalelemente positiv sein müssen.

0	1	3	1	0	0	
1	0	2	0	1	0	
3	2	0	0	0	1	
1	1	3	1	0	0	
1	0	2	1	1	0	Addition der zweiten Spalte zur ersten;
5	2	0	0	0	1	
2	1	5				
1	0	2				Addition der zweiten Zeile zur ersten;
5	2	0				
2	0	5	1	$-\frac{1}{2}$	0	
1	$-\frac{1}{2}$	2	1	$\frac{1}{2}$	0	Subtraktion des $\frac{1}{2}$-fachen der ersten Spalte von der zweiten;
5	$-\frac{1}{2}$	0	0	0	1	
2	0	5				
0	$-\frac{1}{2}$	$-\frac{1}{2}$				Subtraktion des $\frac{1}{2}$-fachen der ersten Zeile von der zweiten;
5	$-\frac{1}{2}$	0				
2	0	0	1	$-\frac{1}{2}$	$-\frac{5}{2}$	
0	$-\frac{1}{2}$	$-\frac{1}{2}$	1	$\frac{1}{2}$	$-\frac{5}{2}$	Subtraktion des $\frac{5}{2}$-fachen der ersten Spalte von der dritten;
5	$-\frac{1}{2}$	$-\frac{25}{2}$	0	0	1	
2	0	0				
0	$-\frac{1}{2}$	$-\frac{1}{2}$				Subtraktion der $\frac{5}{2}$-fachen der ersten Zeile von der dritten;
0	$-\frac{1}{2}$	$-\frac{25}{2}$				
2	0	0	1	$-\frac{1}{2}$	-2	
0	$-\frac{1}{2}$	0	1	$\frac{1}{2}$	-3	Subtaktion der zweiten Spalte von der dritten;
0	$-\frac{1}{2}$	-12	0	0	1	
2	0	0				
0	$-\frac{1}{2}$	0				Subtraktion der zweiten Zeile von der dritten;
0	0	-12				

Die rechts entstehende Matrix

$$P := \begin{pmatrix} 1 & -\frac{1}{2} & -2 \\ 1 & \frac{1}{2} & -3 \\ 0 & 0 & 1 \end{pmatrix}$$

hat in der Tat die Eigenschaft

$$P^t \cdot \begin{pmatrix} 0 & 1 & 3 \\ 1 & 0 & 2 \\ 3 & 2 & 0 \end{pmatrix} \cdot P = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -12 \end{pmatrix} .$$

Die Signatur der quadratischen Form Q_A ist also $(1,2)$, Q_A ist indefinit, d.h. nach der Definition: A ist indefinit.

Ein letztes sehr beachtliches Kriterium für die positive Definitheit einer symmetrischen Matrix A ist:
Löst man das lineare Gleichungssystem $Ax = 0$ mit Hilfe des Gauß-Algorithmus, so ist A genau dann positiv definit, wenn die bei der Lösung (ohne Spaltenvertauschung) auftretende Pivot-Elemente alle positiv sind.

Zur Erinnerung: Die Pivot-Elemente sind Elemente $\neq 0$ in der Matrix, die man zum „Ausräumen“ der Spalten benutzt.

Viele der genannten Kriterien hängen eng miteinander zusammen, so liefert etwa der Gauß-Algorithmus eine sogenannte L-R-Zerlegung der Matrix A, d.h. L ist eine untere Dreiecksmatrix (mit lauter Einsen auf der Diagonalen) und R ist eine obere Dreiecksmatrix (mit Diagonalelementen $\neq 0$). A wird dabei als invertierbar vorausgesetzt.

Beispiel:

$$A := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & \frac{2}{3} & 1 \end{pmatrix}}_{=L} \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 3 \end{pmatrix}}_{=R} = LR\,.$$

Arbeitsmaterialien zur Vorlesung Analysis 3
SS 2003

Parameterabhängige Integrale

Häufig sind Funktionen definiert als Integrale der Gestalt

$$g(x) = \int f(x,t)dt,$$

ein typisches Beispiel ist etwa die Gamma-Funktion:

Für $x > 0$

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$

Wir untersuchen, unter welchen Voraussetzungen an die Funktion f die entsprechende Funktion g stetig bzw. (stetig) differenzierbar von x abhängt. Unter Benutzung des Lebesgueschen Grenzwertsatzes und unter Verwendung des Begriffs der Nullmenge ergeben sich starke Sätze, die unsere früheren Überlegungen über parameterabhängige Integrale wesentlich verallgemeinern. In de folgenden zwei Sätzen ist

$$f : X \times T \to \mathbb{R}$$

eine Funktion auf dem Produkt eines metrischen Raumes X und einer Menge $T \subset \mathbb{R}^m$. Wir machen die folgende „Generalvoraussetzung“:

Für jedes fixierte $x \in X$ sei die Funktion $t \mapsto f(x,t)$ über T integrierbar. Durch Integration über T entsteht dann eine Funktion (parameterabhängiges Integral)

$$g : \quad X \to \mathbb{R} \quad \text{mit}$$
$$g(x) = \int_T f(x,t)dt$$

I: **Stetigkeitssatz**: f habe folgende zusätzliche Eigenschaften

(1) Es gibt eine Nullmenge $N \subset T$, so dass für jedes fixierte $t \subset T - N$ die Funktion $x \mapsto f(x,t)$ stetig ist.

(2) Es gibt eine Funktion $F : T \to \mathbb{R}_+$ mit

$$\begin{array}{rcll} I^*(F) & < & \infty & \text{und} \\ |f(x,t)| & \leq & F(t) & \text{für fast alle } t \in T \text{ und } \textbf{alle } x \in X \end{array}$$

Dann ist die durch

$$g(x) := \int_T f(x,t)dt$$

für $x \in X$ definierte Funktion

$$g : X \to \mathbb{R}$$

stetig.

Kurz: Ist $f(x,t)$ Lebesgue-integrierbar bezüglich t und stetig in x, so ist $\int_T f(x,t)dt$ stetig in

x, falls f gleichmäßig in x dominiert ist.

Beweis: Zum Nachweis der Stetigkeit benutzen wir das Folgenkriterium:

Zu zeigen ist also:

Für jede Folge $(x_k),\ x_k \in X$ mit $\lim\limits_{k\to\infty} x_k = x \in X$ gilt $\lim\limits_{k\to\infty} g(x_k) = g(x)$.

Dazu betrachten wir die Funktionenfolge

$$(f_k) \text{ mit } \begin{aligned} f_k : T &\to \mathbb{R} \\ t &\mapsto f_k(t) := f(x_k, t). \end{aligned}$$

Wegen der allgemeinen Voraussetzung sind alle f_k über T integrierbar und wegen (1) konvergiert (f_k) f.ü. gegen $f(x,t)$. Wegen (2) können wir den Lebesgueschen Grenzwertsatz anwenden und erhalten

$$\begin{aligned} \lim_{k\to\infty} g(x_k) = \lim_{k\to\infty} \int\limits_T f_k(t)dt &= \lim_{k\to\infty} \int\limits_T f(x_k,t)dt \\ &= \int\limits_T f(x,t)dt = g(x). \end{aligned}$$

Bemerkung: Durch Zerlegung in Real- und Imaginärteil, erhält man sofort auch einen analogen Satz für komplexwertige Funktionen.

Beispiel: Fourier-Transformierte:

Sei $f \in \mathcal{L}^1(\mathbb{R},\mathbb{C})$.

Dann ist für jedes $x \in \mathbb{R}$ auch die Funktion

$$t \mapsto f(t)e^{-ixt}$$

über $\mathbb{R}$ integrierbar und die durch

$$\widehat{f}(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} f(t)e^{-ixt}dt$$

definierte Funktion

$$\widehat{f} : \mathbb{R} \to \mathbb{C}$$

heißt **Fourier-Transformierte** von f.

Nach dem Stetigkeitssatz ist $\widehat{f}$ stetig, da $x \mapsto f(t)e^{-ixt}$ für jedes $t \in \mathbb{R}$ stetig ist und $|f|$ eine von x unabhängige Majorante des Integranden ist.

II: **Differentiationssatz**:

Sei X jetzt eine offene Teilmenge ($\neq 0$) der $\mathbb{R}^n$ und f habe die folgenden zusätzlichen Eigenschaften:

(a) Für fast alle $t \in T$ ist $x \mapsto f(x,t)$ stetig differenzierbar.

(b) Es gibt eine Funktion

$$\begin{aligned} F : T &\to \mathbb{R}_+ && \text{mit} \\ I^*(F) &< \infty && \text{und} \\ |\partial_j f(x,t)| &\leq F(t) && \left(\partial_j := \tfrac{\partial}{\partial x_j}\right). \end{aligned}$$

Dann ist die durch

$$g(x) := \int_T f(x,t)dt$$

definierte Funktion

$$g : X \to \mathbb{R}$$

stetig differenzierbar.

Ferner ist für jedes x die Funktion $t \mapsto \partial_j f(x,t)$ über T integrierbar, und es gilt für $1 \le j \le n$

$$\boxed{\partial_j g(x) = \int_T \partial_j f(x,t)dt}$$

Diese Formel ist auch als *Leibnizsche Regel* für die Differentiation eines parameterabhängigen Integrals bekannt.

Bemerkung: Der völlig analoge Satz gilt auch für Funktionen $f : X \times T \to \mathbb{C}$.

Beweis: Sei $x_0 \in X$, wir wählen eine reelle Zahl $r > 0$, so dass die Würfelumgebung

$$W_r(x_0) = \{x \in \mathbb{R}^n;\ \|x - x_0\|_\infty < r\}$$

in X liegt (beachte: X ist offen).

Wir arbeiten mit Differenzenquotienten.

Sei dazu (h_k) eine Nullfolge reeller Zahlen mit $|h_k| < r$ und $h_k \neq 0$ für alle $k \in \mathbb{N}$. Wir setzen $x_k := x_0 + h_k e_j$ (e_j der j-te Standard-Einheitsvektor im $\mathbb{R}^n;\ 1 \le j \le n$) und betrachten

$$\varphi_k(t) := \frac{f(x_k,t) - f(x_0,t)}{h_k}$$

Die $\varphi_k : T \to \mathbb{R}$ sind integrierbare Funktionen und für fast alle $t \in T$ gilt

$$\lim_{k\to\infty} \varphi_k(t) = \partial_j f(x_0,t).$$

Ferner folgt aus dem Schrankensatz der Differentialrechnung

$$|\varphi_k(t)| \le F(t).$$

Nach dem Lebesgueschen Grenzwertsatz ist dann auch die Grenzfunktion der (φ_k), also die Funktion

$$t \mapsto \partial_j f(x_0,t)$$

integrierbar, und es gilt

$$\lim_{k\to\infty} \int_T \varphi_k(t)dt = \int_T \partial_j f(x_0,t).$$

Hieraus folgt wegen

$$\int_T \varphi_k(t)dt = \frac{g(x_k) - g(x_0)}{h_k}$$

durch Grenzübergang $k \to \infty$ die partielle Differenzierbarkeit von g und die Formel

$$(*) \qquad \partial_j g(x_0) = \int_T \partial_j f(x_0,t)dt.$$

Durch Anwendung des Stetigkeitssatzes erhält man aus $(*)$ auch die Stetigkeit der partiellen Ableitung $\partial_j g(x_0)$ und damit die stetige Differenzierbarkeit von g.

Beispiel: Fourier-Transformierte von $t \mapsto e^{-t^2/2}$

Wir berechnen das Integral ($x \in \mathbb{R}$)

$$\widehat{f}(x) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-t^2/2} e^{-ixt} dt,$$

die sog. **Fourier-Transformierte** von

$$f : t \to e^{-t^2/2} \quad (t \in \mathbb{R}).$$

Das Integral existiert und ist nach dem Stetigkeitssatz ($X = \mathbb{R}$) eine stetige Funktion von x.

Zur Berechnung setzen wir für $x \in \mathbb{R}$

$$g(x) := \int_{\mathbb{R}} e^{-\frac{1}{2}(t+ix)^2} dt$$

Wir zeigen, dass g stetig differenzierbar ist und dass für die Ableitung

$$g'(x) = -i \int_{\mathbb{R}} e^{-\frac{1}{2}(t+ix)^2} (t+ix) dt$$

gilt.

Es genügt die Behauptung für jedes beschränkte offene Integral $]-a, a[, \ a \in \mathbb{R}, \ a > 0$, zu zeigen.

Die erste Bedingung des Differetiationssatzes ist offensichtlich erfüllt und für alle $(x,t) \in]-a,a[\times \mathbb{R}$ besteht die Abschätzung

$$\frac{\partial}{\partial x} \left| e^{-\frac{1}{2}(x+it)^2} \right| \leq e^{-t^2/2} e^{a^2/2} (|t| + a) =: F(t)$$

Da F über $\mathbb{R}$ integrierbar ist, folgt die Behauptung aus dem Differentiationssatz.

Wir wollen das Integral für g' berechnen. Das ist einfach, denn $t \mapsto -e^{-\frac{1}{2}(t+ix)^2} =: G(t)$ hat die Ableitung $G'(t) = e^{-\frac{1}{2}(t+ix)^2}(t+ix)$ daher ist $g'(x) = -i(G(\infty) - G(-\infty)) = 0$ für alle $x \in \mathbb{R}$, nach dem MWSD folgt, dass g konstant auf $\mathbb{R}$ ist. Wegen $g(x) = g(0) = \sqrt{2\pi}$ (vergleiche hierzu das Beispiel über die Berechnung des Gauß-Integrals $\int_{\mathbb{R}^n} e^{-\|x\|^2} dx$ oder $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$) ergibt sich dann wegen $\widehat{f}(x) = \frac{1}{\sqrt{2\pi}} g(x)$ auch

$$\widehat{f}(x) = e^{-x^2/2} = f(x).$$

$\boxed{f \text{ stimmt also mit seiner Fourier-Transformierten überein: } f = \widehat{f}}$

Diese Gleichung lässt sich auch in der äquivalenten Form (Gauß-Verteilung)

$$\boxed{\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-t^2/2} dt = 1}$$

schreiben.

Zusatzbemerkung: Es gilt auch ein *Holomorphiesatz*, der in etwa besagt, dass im Fall $X = D =$offene nicht leere Teilmenge von $\mathbb{C}$ und der Voraussetzung, dass für jedes fixierte $t \in T$ die Abbildung

$$\begin{array}{rcl} D & \to & \mathbb{C} \\ z & \mapsto & f(z,t) \end{array}$$

holomorph (= in jedem Punkt $z \in D$ komplex differenzierbar) ist, die durch

$$\begin{array}{rcl} D & \to & \mathbb{C} \\ z & \mapsto & g(z) := \int\limits_T f(z,t)dt \end{array}$$

ebenfalls holomorph ist, und dass

$$g'(z) = \int\limits_T \frac{\partial}{\partial z} f(z,t)dt$$

gilt.

Bemerkung: Die von (b) analoge Bedingung muss natürlich ebenfalls erfüllt sein.