Scriptum zur Vorlesung Analysis I und II

Prof. W. Hoffmann
Universität Bielefeld
SS 2022 und WS 2022/23

Stichwortverzeichnis

1 Grundlagen

Die moderne Mathematik gründet sich auf der Mengenlehre und diese wiederum auf der Logik. Darum müssen wir uns zunächst mit diesen Grundlagen bekannt machen.

1.1 Aussagen

Vorlesung 1b
8.4.2022

Es werden nur Aussagen betrachtet, die einen der beiden Wahrheitswerte „wahr“ oder „falsch“, abgekürzt w oder f, haben. Aus gegebenen Aussagen kann man durch Verknüpfung neue Aussagen bilden. So ist die *Negation* (auch Verneinung genannt) einer Aussage A, abgekürzt $\neg A$, gegeben durch die Wahrheitstafel

A	$\neg A$
w	f
f	w

Die Verneinung der Aussage

„Du hast immer Zeit für mich“

ist die Aussage

„Du hast nicht immer Zeit für mich“.

Verknüpfungen von zwei Aussagen A und B sind z. B. die *Konjunktion* $A \wedge B$ (auch Und-Verknüpfung genannt) sowie die *Disjunktion* $A \vee B$ (auch Oder-Verknüpfung genannt), gegeben durch die Wahrheitstafeln

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

A	B	$A \vee B$
w	w	w
w	f	w
f	w	w
f	f	f

Beispiele sind die Aussagen

„Der Beschuldigte hatte ein Motiv und die Gelegenheit für die Tat.“
„Dieser Neureiche hat eine Erbschaft gemacht oder im Lotto gewonnen.“

Man beachte, dass in der Logik das Wort „oder“ im einschließenden Sinne gebraucht wird.

Die *Implikation* (auch Subjunktion genannt) ist die Aussage $A \Rightarrow B$ (gelesen „wenn A, dann B“), die gegeben ist durch

A	B	$A \Rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

Sie drückt keine Ursache-Wirkung-Beziehung aus, z. B.:

> „Wenn du diese Aufgabe lösen kannst, dann bist du ein Genie.“

Manchmal wird der Bedingungssatz auch durch das Wort „falls“ oder gar nicht eingeleitet und der Hauptsatz durch das Wort „so“:

> „Kräht der Hahn auf dem Mist, so ändert sich das Wetter, oder es bleibt wie es ist.“

Schließlich haben wir noch die *Äquivalenz* $A \Leftrightarrow B$ (gelesen „genau dann A, wenn B“) und die *Antivalenz* $A \rightarrowtail\!\!\!\!\!\leftarrowtail B$ (gelesen „entweder A oder B“) mit den Wahrheitstafeln

A	B	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

A	B	$A \rightarrowtail\!\!\!\!\!\leftarrowtail B$
w	w	f
w	f	w
f	w	w
f	f	f

Im täglichen Leben werden diese oft durch die Worte „wenn“ bzw. „oder“ ausgedrückt, z. B.

> „Wenn du mir dein Fahrrad leihst, kannst du mit meinem Ball spielen.“
> „Du gibst mir jetzt den Ball zurück oder ich sag’ es Mutti.“

Ein logisches Gesetz ist eine Verknüpfung von Variablen, die bei jeder Belegung der Variablen mit Aussagen zu einer wahren Aussage wird, z. B.

$$A \Rightarrow A \vee B, \qquad A \wedge B \Rightarrow B.$$

Viele logische Gesetze haben die Struktur einer Äquivalenz, z. B.

$$\neg\neg A \Leftrightarrow A, \qquad (A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A),$$
$$(A \Leftrightarrow B) \Leftrightarrow ((A \Rightarrow B) \wedge (B \Rightarrow A)).$$

(Die Aussage $\neg B \Rightarrow \neg A$ nennt man übrigens die *Kontraposition* der Implikation $A \Rightarrow B$). Weitere Beispiele sind die *Kommutativgesetze*

$$A \wedge B \Leftrightarrow B \wedge A, \qquad A \vee B \Leftrightarrow B \vee A,$$

die *Assoziativgesetze*

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

die *Distributivgesetze*

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

sowie die *de Morganschen Gesetze*

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

Um ein logisches Gesetz zu beweisen, müssen wir nachprüfen, dass es für alle möglichen Wahrheitswerte der vorkommenden Variablen wahr ist. Dabei sind die Wahrheitswerte von verschachtelten Verknüpfungen schrittweise zu bestimmen, z. B.

A	B	$A \Rightarrow B$	$B \Rightarrow A$	$(A \Rightarrow B) \wedge (B \Rightarrow A)$
w	w	w	w	w
w	f	f	w	f
f	w	w	f	f
f	f	w	w	w

Durch Vergleich mit der Wahrheitstafel für die Äquivalenz $A \Leftrightarrow B$ finden wir, dass diese zu der Aussage

$$(A \Rightarrow B) \wedge (B \Rightarrow A)$$

äquivalent ist.

1.2 Mengen und Prädikate

Vorlesung 2
13.4.2022

Eine *Menge* ist, naiv gesprochen, eine Zusammenfassungen von Objekten. Manchmal kann man ihre Elemente aufzählen, wobei man sie in geschweifte Klammern einschließt und die Reihenfolge gleichgültig ist, also

$$\{\text{Sonne, Erde, Mond}\} = \{\text{Erde, Mond, Sonne}\}.$$

Wenn ein Objekt zu einer Menge gehört, nennt man es ein *Element* dieser Menge. Statt dessen sagt man manchmal auch, dass die Menge das Objekt enthält. Auch eine Menge ist ein Objekt, und man kann Mengen von Mengen bilden, wie zum Beispiel die Menge aller Ehepaare und eingetragenen Partnerschaften. In der Mathematik werden Objekte häufig durch Buchstaben abgekürzt, beispielsweise

- das Verhältnis des Umfangs eines Kreises zum Durchmesser durch den griechischen Buchstaben π,
- die Menge der natürlichen Zahlen und die der reellen Zahlen mit den fettgedruckten Buchstaben **N** bzw. **R**,
- die leere Menge, also die Menge, die kein Element hat, mit $\varnothing$.

Aus der handschriftlichen Imitation $\mathbb{N}$, $\mathbb{R}$ solcher Buchstaben ist dann ein neuer Schrifttyp hervorgegangen, der sich für Zahlbereiche durchgesetzt hat.

In mathematischen Überlegungen kommen oft unbekannte Objekte (kurz „Unbekannte“) vor, die man mit kursiven Buchstaben abkürzt, am liebsten mit x, aber auch mit den Anfangsbuchstaben ihrer Bezeichnungen. Bezeichnet man beispielsweise mit r das Verhältnis (lateinisch *ratio*) der längeren Seite eines Papierbogens zu seiner kürzeren Seite, so führt die Anforderung an DIN-Formate der Reihe A, dass ein halbiertes Blatt wieder das selbe Seitenverhältnis haben soll, auf die Gleichung

$$r \cdot r = 2.$$

Bezeichnet man mit x und y die Kantenlängen des Formats A0 in Metern, so gibt es außerdem die Anforderung

$$x \cdot y = 1.$$

Traditionell wird das Multiplikationszeichen (außer zwischen zwei Zahlen) meist weggelassen, man schreibt beispielsweise $xy = 1$, weshalb mehrbuchstabige Abkürzungen in der Mathematik vermieden oder mit anderen Schrifttypen geschrieben werden.

Die Gleichung $17 \cdot 91 = 13 \cdot 119$ ist eine wahre Aussage, die Gleichung $0^0 = 0$ ist eine falsche Aussage, aber die beiden obigen Gleichungen sind gar keine Aussagen, weil sich erst beim Einsetzen von Zahlen an Stelle der Unbekannten entscheidet, ob sie wahr oder falsch sind. Statt dessen spricht man von *Aussageformen* oder *Prädikaten*. Aussagen und Prädikate müssen nicht die Form von Gleichungen haben, sondern können in Sätzen formuliert werden, etwa „das Verhältnis von Kreisumfang zu Durchmesser ist ein Element der Menge der reellen Zahlen“. Letzteres lässt sich auch in Formeln ausdrücken, nämlich

$$\pi \in \mathbb{R}.$$

Prädikate mit genau einer Unbekannten x werden oft *Eigenschaften* von x genannt. So sollten die Seitenlängen eines Papierbogens positiv sein, also die Eigenschaften

$$x > 0, \qquad y > 0$$

haben, und dann folgt auch

$$r > 0.$$

Prädikate mit zwei Unbekannten werden oft *Relationen* genannt. So können wir beispielsweise vereinbaren, dass x die größere der beiden Seitenlängen bezeichnen soll:

$$x \geq y.$$

Relationen werden in mathematischen Formeln gern durch Sonderzeichen wie $\leq$ oder $\in$ zwischen den beiden Variablen ausgedrückt. Die Negation der Aussage wird häufig durch senkrechtes[1] Durchstreichen ausgedrückt, also

$$\pi \notin \mathbb{N}, \qquad \pi \neq 3{,}14159$$

Aussagen wie „die Elementarladung ist winzig“ oder

$$\pi \approx 3{,}14159$$

hängen vom Ermessen ab und werden in der Mathematik nicht betrachtet.

Zu jeder Eigenschaft gehört eine Menge, nämlich die Menge derjenigen Objekte, auf die diese Eigenschaft zutrifft. Ist die Eigenschaft durch eine Gleichung oder Ungleichung gegeben, so bezeichnet man die zugehörige Menge als *Lösungsmenge* dieser Gleichung. Oft lassen sich die Objekte mit einer vorgegebenen Eigenschaft nicht mehr aufzählen, und man schreibt in die Mengenklammer nur eine Unbekannte, gefolgt von einem senkrechten Strich und dem Prädikat, z. B.[2]

$$\{n \mid n \text{ ist eine Primzahl}\}$$

Umgekehrt gehört zu jeder Menge eine Eigenschaft, nämlich die, ein Element dieser Menge zu sein. Die Prädikate

„n ist eine natürliche Zahl“

und

$$n \in \mathbb{N}$$

sind also gleichbedeutend.

Will man aussagen, dass ein Prädikat zu einer wahren Aussage wird, egal welches Objekt man für die Unbekannte einsetzt, dann stellt man die Worte[3] „für alle x gilt“ voran, abgekürzt

$$\forall\, x$$

Will man hingegen aussagen, dass ein Prädikat für wenigstens ein Objekt zu einer wahren Aussage wird, dann stellt man die Worte [4] „es existiert ein x,

[1]in englischsprachigen Ländern schräges

[2]gelesen: Menge aller n mit der Eigenschaft „n ist eine Primzahl“

[3]Man sagt auch „für beliebige x gilt“ oder „für jedes x gilt“.

[4]Man sagt auch „es gibt ein x, so dass“.

so dass“ voran, abgekürzt

$$\exists\, x$$

Das darauf folgende Prädikat wird meist in Klammern gesetzt. Die Zeichen $\forall$ und $\exists$ nennt man *Quantoren*. Banale Beispiele sind

$$\forall\, x\ (x = x), \qquad \exists\, M\ \forall\, x\ (x \notin M).$$

Eine *Definition* Definition legt einen neuen Begriff (eine Menge oder ein Prädikat) fest. In der Mathematik muss dies allein durch bereits bekannte Begriffe erfolgen.

Definition 1.1. *Es seien M und N Mengen.*[5]

(i) *Wir sagen, dass M eine* Teilmenge *von N ist, abgekürzt*[6] $M \subseteq N$, *wenn*[7] *für alle Objekte x gilt*

$$x \in M \Rightarrow x \in N.$$

(ii) *Wir sagen, dass M eine* echte Teilmenge *von N ist, abgekürzt*[8] $M \subset N$, *wenn*[7] $M \subseteq N$ *und* $M \neq N$ *ist.*

(iii) *Die* Schnittmenge[10] *von M und N ist*[11]

$$M \cap N = \{x \mid x \in M \wedge x \in N\}.$$

(iv) *Die* Vereinigungsmenge[12] *von M und N ist*[13]

$$M \cup N = \{x \mid x \in M \vee x \in N\}.$$

(v) *Die* Differenzmenge *von M und N ist*[14]

$$M \setminus N = \{x \mid x \in M \wedge x \notin N\}.$$

(vi) *Wir sagen, dass die Mengen M und N* disjunkt *sind, wenn* $M \cap N = \varnothing$ *ist.*

Zur Veranschaulichung von Mengenrelationen wie (i), (ii), (vi) und von Mengenoperationen wie (iii)–(iv) dienen *Venn-Diagramme*.

[5] Damit erspart man sich, jedem der folgenden Prädikate die Worte „Für alle Mengen M und N gilt…“ voranzustellen.

[6] gelesen „M (ist) enthalten in N“

[7] gemeint ist eigentlich „genau dann, wenn“.

[8] gelesen „M (ist) echt enthalten in N“

[10] auch Durchschnitt genannt

[11] gelesen „M geschnitten (mit) in N“

[12] auch Vereinigung genannt

[13] gelesen „M vereinigt (mit) in N“

[14] oft gelesen „M ohne in N“

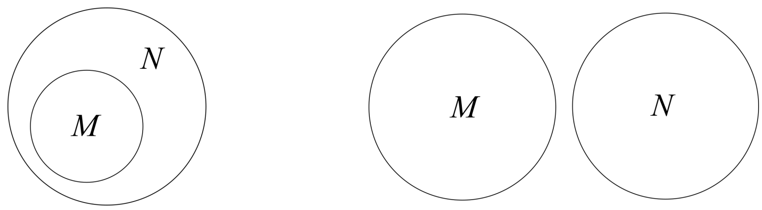


Abbildung 1: Teilmenge, disjunkte Mengen

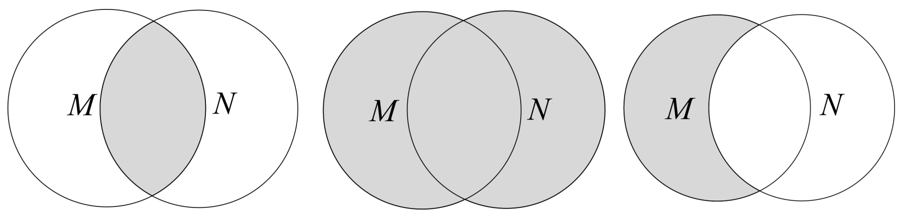


Abbildung 2: Schnittmenge, Vereinigungsmenge, Differenzmenge

Wichtige Aussagen werden in der Mathematik zur späteren Verwendung als Lehrsätze herausgestellt, die man heutzutage kurz Sätze nennt.

Satz 1.1. *Für alle Mengen K, L, M und N gelten die Kommutativgesetze*

$$M \cap N = N \cap M, \qquad M \cup N = N \cup M,$$

die Assoziativgesetze

$$(L \cap M) \cap N = L \cap (M \cap N), \qquad (L \cup M) \cup N = L \cup (M \cup N),$$

die Distributivgesetze

$$L \cup (M \cap N) = (L \cup M) \cap (L \cup N), \quad L \cap (M \cup N) = (L \cap M) \cup (L \cap N)$$

und die de Morganschen Gesetze

$$L \setminus (M \cap N) = (L \setminus M) \cup (L \setminus N), \qquad L \setminus (M \cup N) = (L \setminus M) \cap (L \setminus N).$$

Es ist lehrreich, sich diese Aussagen an Hand von Venn-Diagrammen klar zu machen. Das ist allerdings kein Beweis, sondern dieser beruht darauf, dass Mengen M und N genau dann gleich sind, wenn für alle Objekte x gilt $x \in M \Leftrightarrow x \in N$. Darum folgen die im Satz aufgeführten Gesetze der Mengenlehre aus den entsprechenden logischen Gesetzen.

Viele Prädikate sind nur für die Elemente einer gewissen Menge sinnvoll, wie z. B. die Eigenschaft „x ist ehrlich“ nur für vernunftbegabte Wesen. Man sagt, dass ein Prädikat für die Elemente einer Menge definiert ist und nennt diese Menge den *Definitionsbereich*. So ist etwa die Gleichung $r^2 = 2$ nur für reelle Zahlen r definiert, und man schreibt darum ihre Lösungsmenge als

$$\{r \in \mathbb{R} \mid r^2 = 2\}.$$

Im vorliegenden Fall ist die Lösungsmenge gleich

$$\{\sqrt{2}, -\sqrt{2}\}.$$

Allgemeiner gilt für jede nichtnegative Zahl a, dass die Lösungsmenge der Gleichung

$$r^2 = a$$

die Menge

$$\{\sqrt{a}, -\sqrt{a}\}$$

ist. In der Mengenklammer darf ein Element mehrmals vorkommen, darum erhalten wir auch, wenn wir für a die Zahl 0 einsetzen, eine wahre Aussage.

Wir wollen den Begriff der Lösungsmenge auf Gleichungen und Ungleichungen mit mehreren Unbekannten verallgemeinern. Betrachten wir beispielsweise das Prädikat

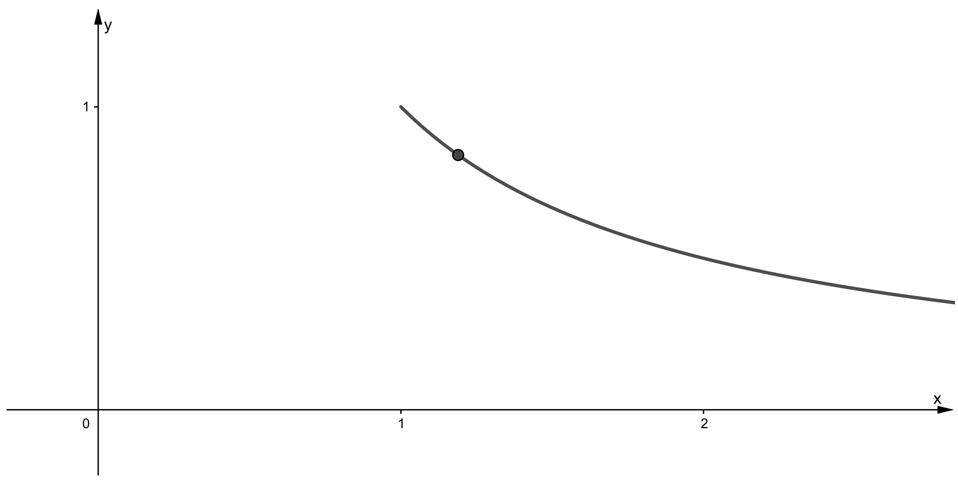
$$xy = 1 \quad \wedge \quad x \geq y.$$

Sind mehrere Gleichungen und Ungleichungen durch „und" verknüpft, dann schreibt man sie traditionell als (Un)Gleichungssystem, hier also

$$\begin{cases} xy = 1 \\ \;\; x \geq y \end{cases}$$

Eine Lösung besteht aus zwei Zahlen, wobei es darauf ankommt, welche für x und welche für y eingesetzt wird. Wegen $\{x, y\} = \{y, x\}$ kommt der Begriff der Menge hier nicht in Frage. Statt dessen definiert man Paare[15] (x, y) auf eine Weise, dass genau dann $(v, w) = (x, y)$ gilt, wenn $v = x$ und $w = y$ gilt. So ist beispielsweise das Paar $(2, \frac{1}{2})$ eine Lösung unseres Systems, das Paar $(\frac{1}{2}, 2)$ hingegen nicht, denn wenn $x = \frac{1}{2}$ und $y = 2$ ist, dann ist zwar die Gleichung erfüllt, die Ungleichung aber nicht. Interpretiert man x und y als Koordinaten in einem Koordinatensystem, so kann man die Lösungsmenge als Teilmenge der Ebene betrachten.

[15] zur Betonung der Wichtigkeit der Reihenfolge nennt man sie auch „geordnete Paare"



Diese Überlegungen führen auf den folgenden Begriff.

Definition 1.2. *Die* Produktmenge *zweier Mengen M und N (auch* cartesisches Produkt *genannt) ist die Menge*

$$M \times N = \{(x, y) \mid x \in M \wedge y \in N\}.$$

Ein anderes Beispiel betrifft eine Heiratsvermittlung (oder heutzutage eine heterosexuelle Dating-App), die eine Menge M von Kundinnen und eine Menge N von Kunden hat. Wir betrachten die Relation

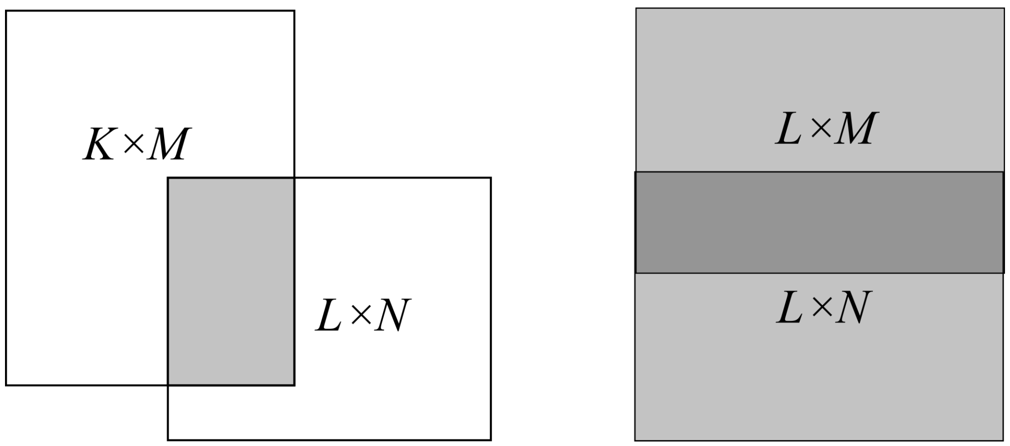
„x ist an y interessiert und y ist an x interessiert“

mit dem Definitionsbereich $M \times N$. Man kann sie darstellen, indem man die Werte „wahr“ bzw. „falsch“ in die Felder einer Tabelle einträgt, deren Spalten den Damen und deren Zeilen den Herren entsprechen.

Satz 1.2. *Es gelten die Distributivgesetze*

$$(K \cap L) \times (M \cap N) = (K \times M) \cap (L \times N), \quad L \times (M \cup N) = L \times M \cup L \times N$$

Das wird plausibel, wenn man den Spezialfall betrachtet, dass M und N Teilmengen von $\mathbb{R}$ sind, so dass man $M \times N$ als Teilmenge der Koordinatenebene darstellen kann.



Dies ist wieder kein Beweis, sondern dieser ergibt sich aus den logischen Gesetzen aufgrund der Definition der Gleichheit von geordneten Paaren.

Gleichungssysteme mit drei, vier, fünf, ... Unbekannten haben als Lösungen Tripel (x, y, z) bzw. Quadrupel (w, x, y, z) bzw. Quintupel (v, w, x, y, z) bzw. ... Im Fall von n Unbekannten spricht man von n-Tupeln.

Um noch einmal auf die DIN-Formate zurückzukommen: Wir hatten r definiert durch die Bedingung

$$rx = y.$$

Multiplizieren wir beide Seiten mit y, so erhalten wir unter Benutzung von $xy = 1$

$$r = y^2.$$

Aus dem oben ermittelten Wert von r ergibt sich wegen $y > 0$

$$y = \sqrt[4]{2}, \qquad x = \frac{1}{\sqrt[4]{2}}.$$

Im obigen Bild ist der Punkt markiert, der dieser Lösung entspricht.

Man kann einen Quantor auch einem Prädikat mit mehreren Unbekannten voranstellen. Dabei entsteht keine Aussage, sondern ein Prädikat, das immer noch von den übrigen Unbekannten abhängt. So ist z. B.

$$\exists\, m \in \mathbb{N}\, (m^2 = n)$$

die Eigenschaft einer natürlichen Zahl n, Quadratzahl zu sein. Hier können wir die gebundene Unbekannte m auch durch eine andere Unbekannte ersetzen, die noch nicht im Prädikat vorkommt. Setzt man weitere Quantoren vor, so kann man sich manche Klammern sparen, z. B. in der Binomischen Formel

$$\forall\, x \in \mathbb{R}\, \forall\, y \in \mathbb{R}\, \big((x + y)^2 = x^2 + 2xy + y^2\big).$$

Wenn wie hier alle Unbekannten durch Quantoren gebunden sind, entsteht eine Aussage. Bei der Vertauschung benachbarter Quantoren der gleichen Art ändert sich die Aussage nicht, bei der Vertauschung von solchen verschiedener Art aber schon: Die Aussage „Für jede Krankheit gibt es eine Medizin, die sie heilt“ ist nicht gleichbedeutend mit der Aussage „Es gibt eine Medizin, die jede Krankheit heilt“.

Unser naiver Ansatz der Mengenlehre führt leider zu Widersprüchen. Am bekanntesten ist die *Russellsche Antinomie*: Wenn die Menge

$$\{x \mid x \notin x\}$$

Element ihrer selbst ist, dann ist sie es nach ihrer Definition nicht, und umgekehrt. Der Ausweg besteht in der axiomatischen Mengenlehre, die aber unseren Rahmen sprengen würde.

1.3 Abbildungen

Vorlesung 3a
20.4.2022

Der Begriff der Abbildung hat zwei Quellen. Zum einen wurde die Entstehung von Bildern in Lochkameras oder auf Leinwänden durch den geometrischen Begriff der Abbildung abstrahiert, bei dem jedem Originalpunkt P ein Bildpunkt P' zugeordnet wird. Die selbe Idee wurde auf Bewegungen von Gegenständen angewendet, wo jedem Punkt P in der Ausgangsposition ebenfalls ein Punkt P' in der Endposition zugeordnet ist.

Ein scheinbar völlig anderes Phänomen ist die Veränderung von physikalischen Größen wie etwa dem Luftdruck p oder dem zurückgelegten Weg s im Laufe der Zeit. Zur Unterscheidung fügte man den Zeitpunkt in Klammern an, und die Schreibweise $s(t)$ war entstanden. Man bezeichnete Größen wie s und t als Veränderliche oder Variable.

In der Mathematik empfand man es als zu unexakt, sämtliche mögliche Abhängigkeiten mit dem selben Symbol der abhängigen Größe zu bezeichnen, und schrieb $s = f(t)$, wobei f die konkrete Rechenvorschift symbolisiert, etwa $f(t) = g \cdot t^2$, wenn es sich um den Fallweg handelt. Leibnitz führte für eine solche Vorschrift die Bezeichnung „Funktion“ ein. Im Prinzip kann man sogar Prädikate als Funktionen mit den Werten „wahr“ oder „falsch“ auffassen, weshalb der Begriff „Variable“ letztlich den Begriff „Unbekannte“ verdrängte. Schließlich wurde klar, dass im Wesentlichen das Selbe geschieht, wenn man einem Punkt P den Bildpunkt P' oder einer Größe t die abhängige Größe $f(t)$ zuordnet, so dass beides mit dem folgenden mathematischen Begriff beschrieben werden kann:

Eine *Abbildung* f von einer Menge M in eine Menge N ordnet jedem Element x von M genau[16] ein Element von N zu, das man mit $f(x)$ bezeichnet. Man nennt M den *Definitionsbereich* und N den *Zielbereich* von f. Die Menge

$$\{(x, y) \in M \times N \mid y = f(x)\}$$

wird der *Graph* von f genannt. Die Worte „Abbildung f von M in N" kürzt man durch $f : M \to N$ oder $M \xrightarrow{f} N$ ab.

Dies ist eigentlich keine Definition, da der neue Begriff „Abbildung" auf den ebenfalls noch nicht definierten Begriff „Zuordnung" zurückgeführt wird. Ein Ausweg besteht darin, die Aussage $f(x) = y$ als Relation zwischen x und y aufzufassen.

Wir schreiben $f = g$, wenn die Abbildungen f und g gleiche Definitionsbereiche und gleiche Zielbereiche haben und für alle Elemente x des Wertebereichs gilt $f(x) = g(x)$. Für jede Menge M ist die *identische Abbildung* $\mathrm{id}_M : M \to M$ durch $\mathrm{id}_M(x) = x$ gegeben.

Sind M und N Teilmengen von $\mathbb{R}$, so nennt man f eine Funktion und kann den Graphen als Teilmenge der Koordinatenebene darstellen. Den Graphen einer Abbildung zwischen endlichen Mengen kann man in Form einer Tabelle darstellen, wie z. B. bei einem Fragebogen.

Definition 1.3. *Ist $f : M \to N$ und $g : L \to M$, so ist die* Verkettung *von f mit g die Abbildung $f \circ g : L \to N$, die gegeben ist durch*

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

Dieser Begriff wird durch die Nacheinanderausführung von Bewegungen motiviert, wo einem Originalpunkt bei der ersten Bewegung der Bildpunkt P' und diesem bei der zweiten Bewegung der Bildpunkt P'' zugewiesen wird. Sind f und g Funktionen und ist $f(x)$ durch einen algebraischen Term gegeben, in dem die Variable x vorkommt, dann erhält man $f(g(u))$, indem man in diesem Term überall x durch $g(u)$ ersetzt.

Man prüft unmittelbar nach, dass für $f : M \to N$ gilt

$$f \circ \mathrm{id}_M = \mathrm{id}_N \circ f = f$$

und dass für $K \xrightarrow{h} L \xrightarrow{g} M \xrightarrow{f} N$ gilt

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

[16]Wenn aus dem Zusammenhang nicht klar ist, ob „ein" als Artikel oder Zahlwort gemeint ist, sagt man im ersten Fall „mindestens ein" und im zweiten Fall „genau ein".

Definition 1.4. (i) *Die Abbildung $f : M \to N$ heißt* injektiv[17], *wenn für alle Elemente u und v von M gilt*

$$u \neq v \quad \Longrightarrow \quad f(u) \neq f(v).$$

(ii) *Die Abbildung $f : M \to N$ heißt* surjektiv[18], *wenn es für jedes Element y von N ein Element x von M gibt, so dass $f(x) = y$.*

(iii) *Die Abbildung $f : M \to N$ heißt* bijektiv[19], *wenn sie sowohl injektiv als auch surjektiv ist.*

(iv) *Eine Abbildung $h : N \to M$ heißt* Umkehrabbildung *von $f : M \to N$, wenn für alle $x \in M$ und $y \in N$ gilt*

$$f(x) = y \quad \Longleftrightarrow \quad h(y) = x.$$

Die Bedingung für die Injektivität ist äquivalent zu ihrer Kontraposition

$$f(u) = f(v) \quad \Longrightarrow \quad u = v.$$

Definiert man den *Wertebereich* von f als

$$\{y \in N \mid \exists x \in M \; f(x) = y\},$$

so bedeutet die Surjektivität, dass der Wertebereich gleich dem Zielbereich ist. Es ist klar, dass eine Abbildung höchstens eine Umkehrabbildung haben kann.

Satz 1.3. (i) *Die Verkettung von injektiven Abbildungen ist injektiv.*

(ii) *Die Verkettung von surjektiven Abbildungen ist surjektiv.*

(iii) *Eine Abbildung $f : M \to N$ ist genau dann bijektiv, wenn sie eine Umkehrabbildung besitzt.*

Beweis. **(i)** In den Bezeichnungen von Definition 1.3 seien u und v beliebige Elemente von L. Ist $u \neq v$, so folgt wegen der Injektivität von g, dass $g(u) \neq g(v)$, und wegen der Injektivität von f, dass $f(g(u)) \neq f(g(v))$, was nach Definition 1.3 bedeutet, dass $f \circ g(u) \neq f \circ g(v)$.

(ii) Ist z ein beliebiges Element von N, so gibt es wegen der Surjektivität von f ein Element y von M, so dass $f(y) = z$, und wegen der Surjektivität

[17] oder eineindeutig

[18] oder Abbildung von M **auf** N

[19] oder umkehrbar eindeutig

von g gibt es dann ein Element x von L, so dass $g(x) = y$. Nach Definition 1.3 gilt dann $f \circ g(x) = z$.

(iii) Angenommen, h ist die Umkehrabbildung von f. Sind u und v Elemente von M und bezeichnen wir $f(u) = r$ und $f(v) = s$, so gilt $h(r) = u$ und $h(s) = v$. Aus $r = s$ folgt wegen der Eindeutigkeit von h, dass $u = v$, also ist f injektiv. Ist y ein beliebiges Element von N, so hat das Element $x = h(y)$ die Eigenschaft $f(x) = y$, also ist f surjektiv. Zusammenfassend erhalten wir, dass f bijektiv ist.

Zum Beweis der Umkehrung nehmen wir an, dass f bijektiv ist. Ist y ein beliebiges Element von N, so gibt es wegen der Surjektivität ein Element x von M mit (der Eigenschaft) $f(x) = y$, und wegen der Injektivität von f ist x eindeutig bestimmt. Wir können also festlegen, dass $h(y) = x$ ist. Offensichtlich ist die so definierte Abbildung h die Umkehrabbildung von f. □

Ist $f : M \to N$ und $K \subseteq M$ sowie $L \subseteq N$, wobei L den Wertebereich von f enthält, so definieren wir die *Einschränkung* $f|_K : K \to N$ und die *Beschränkung* ${}_L|f : M \to L$ als die Abbildungen mit der selben Abbildungsvorschrift wie f. Ihr Graph ist die Schnittmenge des Graphen von f mit der Menge $K \times N$ bzw. $M \times L$. Bei geeigneter Wahl von K und L erhält man eine bijektive Abbildung ${}_L|f|_K$. Wir werden beispielsweise später zeigen, dass aus der durch $f(x) = x^2$ gegebenen Funktion eine bijektive Funktion entsteht, wenn wir sie auf die Menge der nichtnegativen Zahlen einschränken und auf die selbe Menge beschränken.

1.4 Äquvalenz- und Ordnungsrelationen

Vorlesung 3b
22.4.2022

Eine *Partition* einer Menge M ist eine Menge $\mathcal{P}$ von nichtleeren Teilmengen von M, so dass es für jedes Element x von M genau ein Element A von $\mathcal{P}$ mit der Eigenschaft $x \in A$ gibt. So wurden beispielsweise die Teilnehmer dieser Veranstaltung in Tutorien eingeteilt. Ein wichtigeres Beispiel ist die Klassifikation von Lebewesen, Objekten, Lauten oder Begriffen als eine Methode der Wissenschaften. Bei einer Klassifikation, beispielsweise von Lebewesen, gibt es keine vorgefasste Liste von Klassen oder Arten. Vielmehr gibt es Kriterien dafür, wann zwei Lebewesen zur selben Art gehören. So ein Kriterium ist logisch betrachtet eine Relation.

Definition 1.5. *Eine Relation*[20] $\sim$ *auf einer Menge* M *wird* Äquivalenzrelation *genannt, wenn für beliebige Elemente* x, y *und* z *von* M *gilt:*

[20]Das Zeichen $\sim$ wird „Tilde“ oder „Schlange“ gelesen, und wenn es eine Äquivalenzrelation ohne eigenen Namen bezeichnet, auch als „äquivalent“.

- $x \sim x$
- $x \sim y \ \Rightarrow \ y \sim x$
- $x \sim y \ \wedge \ y \sim z \ \Rightarrow \ x \sim z$

Diese drei Axiome der Äquivalenzrelation nennt man Reflexivität, Symmetrie und Transitivität.

Satz 1.4. (i) *Ist $\mathcal{P}$ eine Partition von M und definieren wir für $x, y \in M$*

$$x \sim y \quad \Leftrightarrow \quad \exists\, A \in \mathcal{P}\,(x \in A \ \wedge \ y \in A),$$

so ist $\sim$ eine Äquivalenzrelation auf M.

(ii) *Ist $\sim$ eine Äquivalenzrelation auf einer Menge M und setzen wir für jedes Element x von M*

$$[x] = \{y \in M \mid y \sim x\},$$

genannt Äquivalenzklasse von x, so ist[21]

$$\mathcal{P} = \{[x] \mid x \in M\}$$

eine Partition von M.

Die Gleichheitsrelation ist übrigens ein Beispiel einer Äquivalenzrelation. In diesem Fall besteht die zugehörige Partition aus den Einermengen in M.

Beweis. Ist eine Partition $\mathcal{P}$ gegeben und definieren wir eine Relation $\sim$ wie in (i), so ist sie offensichtlich reflexiv und symmetrisch. Gilt $x \sim y$ und $y \sim z$, so gibt es $A, B \in \mathcal{P}$, so dass $x, y \in A$ und $y, z \in B$ ist. Da aber y in nur einem Teil der Partition enthalten sein kann, gilt $A = B$, und die Transitivität folgt.

Nun betrachten wir eine Äquivalenzrelation $\sim$ und definieren $\mathcal{P}$ wie in Behauptung (ii). Wegen der Reflexivität gehört jedes Element x von M zu einem Element von $\mathcal{P}$, nämlich $[x]$.

Angenommen, für Elemente x, y von M haben $[x]$ und $[y]$ ein gemeinsames Element z. Dann gilt $z \sim x$ und $z \sim y$. Mit der Symmetrie folgt $x \sim z$, und mit der Transitivität folgt $x \sim y$. Für jedes Element u von $[x]$ gilt $u \sim x$, und mit der Transitivität folgt $u \sim y$, so dass $[x] \subseteq [y]$. Genauso zeigt man $[y] \subseteq [x]$, und somit folgt $[x] = [y]$. Verschiedene Äquivalenzklassen sind also disjunkt, und darum kann ein Element von M nur zu einer gehören. $\square$

[21] Exakter wäre $\mathcal{P} = \{A \mid \exists\, x \in M\,(A = [x])\}$.

Eine andere wissenschaftliche Methode besteht darin, Objekte (etwa chemische Elemente, archäologische Funde) nach gewissen Kriterien in eine Reihenfolge zu bringen. Auch diese Kriterien sind Relationen.

Definition 1.6. *Eine Relation*[22] $\preceq$ *auf einer Menge* M *wird* Ordnung *genannt, wenn für beliebige Elemente* x, y, z *von* M *gilt*

- $x \preceq x$
- $x \preceq y \ \wedge \ y \preceq x \ \Rightarrow \ x = y$
- $x \preceq y \ \wedge \ y \preceq z \ \Rightarrow \ x \preceq z$
- $x \preceq y \ \vee \ y \preceq x$

Die zweite Eigenschaft nennt man Antisymmetrie, die Letzte nennt man Totalität, die anderen sind uns bereits bekannt. Aus jeder Ordnung erhält man eine strikte Ordnung

$$x \prec y \quad \Leftrightarrow \quad x \preceq y \ \wedge \ x \neq y,$$

die man ebenfalls durch vier Axiome charakterisieren kann, und aus jeder strikten Ordnung erhält man eine nichtstrikte:

$$x \preceq y \quad \Leftrightarrow \quad x \prec y \ \vee \ x = y.$$

Durch Vertauschung der Argumente entsteht eine Ordnung, die man gern durch das gespiegelte Symbol abkürzt, also

$$x \succ y \quad \Leftrightarrow \quad y \prec x.$$

Die Verknüpfung

$$x \prec y \ \wedge \ y \prec z$$

schreibt man oft in der platzsparenden Form

$$x \prec y \prec z.$$

Auch die Menge der reellen Zahlen und Alphabete werden mit Ordnungen versehen. Eine Menge zusammen mit einer Ordnung nennt man geordnete Menge.

[22] oft gelesen „vor oder gleich“

Definition 1.7. *Eine Teilmenge I einer geordneten Menge M heißt* Intervall *in M, wenn für beliebige Elemente x und z von I und jedes Element y von M mit der Eigenschaft $x \prec y \prec z$ gilt $y \in I$.*

Eine Teilmenge I einer geordneten Menge M heißt Endintervall *von M, wenn für jedes Element x von I und jedes Element y von M mit der Eigenschaft $x \prec y$ gilt $y \in I$.*

Analog definiert man Anfangsintervalle. Man sollte sich klar machen, dass jedes Endintervall wie auch jedes Anfangsintervall ein Intervall ist.

1.5 Kardinalzahlen

Wenn man sich nach Abschluss der Schule fragt, was eine Zahl ist, so wird man nicht in der Lage sein, eine Definition anzugeben. Nach den Ansprüchen der Mathematik müssen die Zahlbereiche von Neuem auf Grundlage der Mengenlehre aufgebaut werden.

Ohne Zahlen kann man nicht einmal sagen, wie viele Elemente eine Menge hat, aber es ist leicht zu sagen, was es bedeutet, dass zwei Mengen gleich viele Elemente haben.

Definition 1.8. *Wir sagen, dass eine Menge M* gleichmächtig *zu einer Menge N ist (abgekürzt $M \sim N$), wenn es eine bijektive Abbildung von M auf N gibt.*

Wegen der bereits erwähnten Russelschen Antinomie führt es zu Widersprüchen, wenn man die Menge aller Mengen betrachtet. In der Mengenlehre nach Zermelo und Fraenkel werden die Axiome so gewählt, dass diese Bildung ausgeschlossen ist. In der Mengenlehre nach von Neumann, Bernais und Gödel wird der allgemeinere Begriff der Klasse betrachtet, wobei diejenigen Klassen als Mengen bezeichnet werden, die Element einer Klasse sind. Dann darf man von der Klasse aller Mengen und auch von Relationen auf Klassen sprechen. Wir nehmen das einmal zur Kenntnis, ohne näher auf die Einzelheiten einzugehen.

Satz 1.5. *Die Relation der Gleichmächtigkeit ist eine Äquivalenzrelation auf der Klasse der Mengen.*

Beweis. Die Reflexivität folgt aus der Bijektivität der identischen Abbildung, die Symmetrie aus Satz 1.3(iii) und die Transitivität aus Satz 1.3(i),(ii). □

Die Mächtigkeitsklasse einer Menge M umfasst all diejenigen Mengen, die gleichmächtig zu M sind. Aus Satz 1.4 (verallgemeinert auf Klassen) folgt, dass jede Menge M zu genau einer Mächtigkeitsklasse gehört. Wir sparen uns die Arbeit, Zahlen als zusätzliche Objekte einzuführen.

Definition 1.9. *Die* Mächtigkeit *einer Menge M, abgekürzt $|M|$, ist die Klasse aller Mengen, die gleichmächtig zu M sind. Jede solche Klasse nennen wir eine* Kardinalzahl. *Wir definieren die Kardinalzahlen*

$$0 = |\varnothing|, \qquad 1 = |\{\varnothing\}|.$$

Die Menge aller Abbildungen von einer Menge M in eine Menge N bezeichnen wir mit N^M.

Satz 1.6. *Es seien K, L, M und N Mengen, so dass*

$$K \sim M, \qquad L \sim N.$$

Dann gilt

$$K \times L \sim M \times N, \qquad L^K \sim N^M.$$

Ist außerdem K disjunkt zu L und M disjunkt zu N, so ist

$$K \cup L \sim M \cup N.$$

Beweis. Sind bijektive Abbildungen $p : K \to M$ und $q : L \to N$ gegeben, so können wir Abbildungen

$$f : K \times L \to M \times N, \qquad g : L^K \to N^M$$

durch die Vorschriften

$$f(x, y) = (p(x), q(y)), \qquad g(r) = q \circ r \circ s$$

definieren, wobei s die Umkehrabbildung von p bezeichnet. Ist außerdem K disjunkt zu L und M disjunkt zu N, so können wir eine Abbildung $h : K \cup L \to M \cup N$ durch

$$h(x) = \begin{cases} p(x) & \text{wenn } x \in K, \\ q(x) & \text{wenn } x \in L \end{cases}$$

definieren. Die Nachprüfung, dass f, g und h bijektiv sind, überlassen wir den Teilnehmern. $\square$

Der Satz rechtfertigt die folgende Definition.

Definition 1.10. *Sind m und n Kardinalzahlen, also $m = |M|$ und $n = |N|$ für gewisse Mengen M und N, so definieren wir ihr* Produkt *bzw. die m-te* Potenz *von n*

$$m \cdot n = |M \times N|, \qquad n^m = |N^M|.$$

Sind außerdem M und N disjunkt[23]*, so definieren wir die* Summe

$$m + n = |M \cup N|.$$

[23] Man kann gegebene Mengen M und N durch disjunkte gleichmächtige Mengen ersetzen.

Jetzt erkennt man den Grund für die Einführung der Schreibweise N^M.

Satz 1.7. *Für beliebige Kardinalzahlen l, m und n gelten die Kommutativgesetze*

$$m + n = n + m, \qquad m \cdot n = n \cdot m,$$

die Assoziativgesetze

$$(l + m) + n = l + (m + n), \qquad (l \cdot m) \cdot n = l \cdot (m \cdot n),$$

das Distributivgesetz

$$l \cdot (m + n) = l \cdot m + l \cdot n$$

und die Potenzgesetze

$$(m \cdot n)^l = m^l \cdot n^l, \qquad n^{l+m} = n^l \cdot n^m, \qquad n^{l \cdot m} = (n^l)^m.$$

Beweis. Wir wählen Mengen L, M und N, so dass $|L| = l$, $|M| = m$ und $|N| = n$. Das Kommutativ- und Assoziativgesetz der Addition sowie das Distributivgesetz folgen unmittelbar aus den entsprechenden Gesetzen in Satz 1.1. Dabei muss man im letzten Fall voraussetzen, dass M und N disjunkt sind, und im zweiten Fall, dass L, M und N paarweise disjunkt sind.

Zum Beweis der Gesetze für die Multiplikation genügt es zu zeigen, dass

$$M \times N \sim N \times M, \qquad (L \times M) \times N \sim L \times (M \times N).$$

Die dazu benötigten bijektiven Abbildungen sind durch

$$f(x, y) = (y, x), \qquad g((x, y), z) = (x, (y, z))$$

gegeben. Der Beweis der Potenzgesetze ist eine Übungsaufgabe. □

Die Relation $\subseteq$ ist eine Halbordnung auf der Klasse der Mengen, d. h. sie hat alle Eigenschaften einer Ordnung mit Ausnahme der Totalität. Wir führen nun eine Ordnung unter den Kardinalzahlen ein.

Definition 1.11. *Sind m und n Kardinalzahlen, also $m = |M|$ und $n = |N|$ für gewisse Mengen M und N, so sagen wir, dass m kleiner oder gleich n ist, abgekürzt $m \leq n$, wenn es eine injektive Abbildung $f : M \to N$ gibt.*

Aus Satz 1.3(i) folgt, dass dies nicht von der Wahl der Vertreter M und N abhängt.

Satz 1.8 (Schröder-Bernstein)**.** *Gibt es eine injektive Abbildung $f : M \to N$ und eine injektive Abbildung $g : N \to M$, so gibt es eine bijektive Abbildung $h : M \to N$.*

Beweisidee. Wir können annehmen dass M und N disjunkt sind und betrachten die maximalen Folgen

$$\ldots, x_i, y_i, x_{i+1}, y_{i+1}, \ldots$$

bei denen für alle i gilt $x_i \in M$, $y_i \in N$,

$$f(x_i) = y_i, \qquad g(y_i) = x_{i+1}.$$

Aus der Injektivität von f und g folgt, dass diese Folgen eine Partition von $M \cup N$ bilden. In den Folgen, die mit einem Element von M beginnen (das nicht im Wertebereich von g liegt), setzen wir $h(x_i) = y_i$, und in den Folgen, die mit einem Element von N beginnen (das nicht im Wertebereich von f liegt), setzen wir $h(x_i) = y_{i-1}$. In den beiderseits unbeschränkten Folgen (die sich auch zum Zykel schließen können) wählen wir eine der beiden Möglichkeiten. □

Satz 1.9. (i) *Die Kleiner-Gleich-Relation zwischen Kardinalzahlen ist eine Ordnung.*

(ii) *Für Kardinalzahlen l, m und n gelten die* Monotoniegesetze

$$l \leq n \quad \Rightarrow \quad l + m \leq n + m \ \wedge \ l \cdot m \leq n \cdot m \ \wedge \ l^m \leq n^m.$$

Beweis. Die Reflexivität folgt aus der Injektivität der identischen Abbildung, die Transitivität aus Satz 1.3(i). Die Antisymmetrie folgt aus Satz 1.8, aber für die Totalität benötigt man den Begriff der Ordinalzahl, deren Einführung unseren Rahmen sprengen würde.

Zum Beweis von (ii) wählen wir eine injektive Abbildung $q : L \to N$ und argumentieren wie im Beweis von Satz 1.6, wobei diesmal $p = \mathrm{id}_M$ ist. Sind beispielsweise r_1, $r_2 : M \to L$ und ist $q \circ r_1 = q \circ r_2$, also $q(r_1(x)) = q(r_2(x))$ für jedes $x \in M$, so folgt aus der Injektivität von q, dass $r_1(x) = r_2(x)$ ist, und die Injektivität der Abbildung $g : L^M \to N^M$ folgt. □

Folgerung 1.1. *Ist $f : M \to N$ eine Abbildung und $|M| > |N|$, so ist f nicht injektiv.*

Dies ist einfach die Kontraposition der Definition der Kleiner-Gleich-Relation. Man nennt dieses Argument den *Schubfachschluss:* Verteilt man m Gegenstände auf n Fächer, wobei $m > n$ ist, so gibt es ein Fach, in dem mehr als ein Gegenstand landet.

Satz 1.10 (Cantor)**.** *Ist $\mathcal{P}$ die Menge aller Teilmengen der Menge M, dann gibt es keine surjektive Abbildung $f : M \to \mathcal{P}$.*

Man nennt $\mathcal{P}$ übrigens die *Potenzmenge* von M.

Beweis. Es sei F eine beliebige Abbildung von M in $\mathcal{P}$. Wir betrachten die Teilmenge

$$L = \{x \in M \mid x \notin F(x)\}$$

von M. Nun sei u irgend ein Element von M. Wenn $u \in F(u)$ ist, so ist laut Definition $u \notin L$, ist hingegen $u \notin F(u)$, so gilt $u \in L$. Auf jeden Fall können die Mengen $F(u)$ und L nicht übereinstimmen, das heißt, es gibt kein $u \in M$ mit der Eigenschaft $F(u) = L$, und somit ist F nicht surjektiv. □

Es gibt eine injektive Abbildung $G : M \to \mathcal{P}$, zum Beispiel $G(x) = \{x\}$, und somit ist $|\mathcal{P}| \geq \mathcal{M}$. Aus dem Satz ergibt sich

Folgerung 1.2. *Ist $\mathcal{P}$ die Potenzmenge von M, so gilt*

$$|\mathcal{P}| > |M|.$$

2 Zahlbereiche

2.1 Natürliche Zahlen

Vorlesung 4a
27.4.2022

Naiv gesagt sind natürliche Zahlen endliche Kardinalzahlen. Dazu muss man erst einmal definieren, was endlich heißt. Überraschender Weise ist es einfacher, das Gegenteil zu definieren: Eine Menge heißt *unendliche Menge*, wenn sie gleichmächtig zu einer ihrer echten Teilmengen ist. Die Existenz wenigstens einer solchen Menge ist übrigens ein Axiom der Mengenlehre.

Ist U eine unendliche Menge, so gibt es also eine injektive Abbildung $i : U \to U$ und ein Element a von U, das nicht im Wertebereich von i liegt. (Wenn U Teilmenge einer anderen Menge W ist, so ist auch diese unendlich, denn man kann i durch die identische Abbildung von $W \setminus U$ fortsetzen.) Die Menge

$$N = \{a, i(a), i(i(a)), i(i(i(a))), \dots\}$$

ist ein Prototyp der Menge der natülichen Zahlen. Sie hat ein Anfangselement a, und die Einschränkung (und Beschränkung)

$$s : N \to N$$

der Abbildung i auf die Menge N ist ebenfalls injektiv, aber nicht surjektiv. (Hilbert hat dies in populärwissenschaftlichen Vorträgen als unendliche Menge von Zimmern eines Hotel veranschaulicht. Wenn alle belegt sind und ein weiterer Gast untergebracht werden soll, verlegt man jeweils den Gast aus Zimmer n in Zimmer $s(n)$, wodurch Zimmer a frei wird.)

Genau genommen ist eine Aufzählung mit Fortsetzungspunkten mathematisch nicht exakt, sondern man muss N als kleinste Teilmenge von M (bezüglich der Halbordnung $\subseteq$) mit folgender Eigenschaft definieren:

($*$) Sie enthält a, und wenn sie ein Element x enthält,
so enthält sie auch $i(x)$.

Ihre Existenz ist dadurch gesichert, dass sie die Schnittmenge aller Teilmengen der Eigenschaft ($*$) ist, denn diese hat ebenfalls die Eigenschaft ($*$). Dazu müssen vorher die Mengenoperationen auf beliebig viele Argumente verallgemeinert werden: Ist X eine Menge von Mengen, so definiert man die *Vereinigungsmenge* der Elemente von X als

$$\{x \mid \exists\, M \in X\,(x \in M)\}$$

und, falls X nicht leer ist, die *Schnittmenge* der Elemente von X als

$$\{x \mid \forall\, M \in X\,(x \in M)\}.$$

Die natürlichen Zahlen wurden zuerst von Giuseppe Peano axiomatisch definiert. In mengentheoretischer Sprache läuft das auf Folgendes hinaus.

Definition 2.1. *Unter einer* Peanomenge *verstehen wir eine Menge N mit einer injektiven, aber nicht surjektiven Abbildung $s : N \to N$, genannt* Nachfolgerabbildung, *und einem Element a, genannt* Anfangselement, *mit folgender Eigenschaft:*

Ist K eine Teilmenge von N, so dass

$$a \in K, \qquad \forall\, x \in K\,(s(x) \in K),$$

so ist $K = N$.

Solche Mengen existieren tatsächlich und haben eine weitere Eigenschaft:

Satz 2.1. (i) *Entstehen N und s wie oben aus einer injektiven, aber nicht surjektiven Abbildung $i : U \to U$ und dem Anfangselement a, so ist N eine Peanomenge mit der Nachfolgerabbildung s.*

(ii) *Ist N eine Peanomenge mit der Nachfolgerabbildung s, so gibt es für jede Menge M, jede Abbildung $r : M \to M$ und jedes Element $b \in M$ genau eine Abbildung $f : N \to M$, so dass $f(a) = b$ ist und für alle $x \in N$ gilt*

$$f(s(x)) = r(f(x)).$$

Beweis. **(i)** Als Einschränkung von i ist s injektiv, und a gehört nicht zum Wertebereich. Ist K wie in der Definition, so ist $K \subseteq N$ und $N \subseteq U$, also $K \subseteq U$, und K hat die Eigenschaft $(*)$. Da N die kleinste Teilmenge von U mit dieser Eigenschaft ist, gilt $N \subseteq K$, und die Behauptung folgt.
(ii) Wir definieren eine Abbildung h der Menge $N \times M$ in sich selbst durch

$$h(x, y) = (s(x), r(y)).$$

Wenn es die behauptete Abbildung f gibt, so ist ihr Graph G eine Teilmenge von $N \times M$ mit der Eigenschaft, dass sie das Element (a, b) und mit jedem ihrer Elemente auch dessen Bild unter h enthält. Um die Existenz von f zu zeigen, betrachten wir die kleinste Teilmenge G mit dieser Eigenschaft.

Die Menge

$$K = \{x \in N \bigm| |\{y \in M \bigl((x, y) \in G\bigr)\}| = 1\}$$

enthält a, denn wenn es außer b ein weiteres Element $c \in M$ mit $(a, c) \in G$ gäbe, hätte auch $G \setminus \{(a, c)\}$ die Eigenschaft $(*)$ in Bezug auf die Abbildung h, aber dann wäre G nicht die kleinste Teilmenge mit dieser Eigenschaft.

Ist $x \in K$, gibt es also ein $z \in M$ mit der Eigenschaft

$$\{y \in M \bigl((x, y) \in G\bigr)\} = \{(x, z)\},$$

so ist $(s(x), r(z)) = h(x, z) \in G$, und wenn es außer $r(z)$ ein weiteres Element $y \in M$ mit $(s(x), y) \in G$ gäbe, hätte $G \setminus \{(s(x), y)\}$ ebenfalls die Eigenschaft $(*)$. Dann wäre G nicht die kleinste Teilmenge mit diesere Eigenschaft, also ist $s(x) \in K$.

Nach (i) ist also $K = N$, und somit ist G der Graph einer Abbildung. Eine größere Menge ist kein Graph, also ist f eindeutig bestimmt. □

Es ist verlockend, einfach eine Peanomenge als Menge der natürlichen Zahlen zu bezeichnen, aber leider gibt es viele derartige Mengen. Zumindest gilt:

Folgerung 2.1. *Sind N_1 und N_2 Peanomengen mit Nachfolgerabbildungen $s_1 : N_1 \to N_1$ und $s_2 : N_2 \to N_2$, so gibt es eine eindeutig bestimmte bijektive Abbildung $f_1 : N_1 \to N_2$, so dass für alle $x \in N_1$ gilt*

$$f_1(s_1(x)) = s_2(f_1(x)).$$

Eine Folgerung ist eigentlich auch ein Satz, dessen Beweis aber offensichtlich ist. Für Anfänger sollte man vielleicht einige Details angeben: Natürlich ist f_1 die Abbildung aus Satz 2.1 angewendet auf $N = N_1$ und $M = N_2$. Anwendung auf $N = N_2$ und $M = N_1$ liefert eine Abbildung $f_2 : N_2 \to N_1$. Die Abbildung $f = f_2 \circ f_1$ erfüllt genau wie id_{N_1} die Gleichung in Satz 2.1(i), und aus der Eindeutigkeit folgt $f_2 \circ f_1 = \mathrm{id}_{N_1}$. Analog zeigt man $f_1 \circ f_2 = \mathrm{id}_{N_2}$.

Laut Folgerung sind die Nachfolgerabbildungen also miteinander verträglich, und es spielt keine Rolle, welche dieser Mengen man als die Menge der natürlichen Zahlen betrachtet. Die meisten Mathematiker ziehen es vor, eindeutig bestimmte natürliche Zahlen zu haben. Dazu nimmt man für U die

Klasse aller Mengen, für a ihr Element $\varnothing$, und benutzt die Abbildung $i(x) = x \cup \{x\}$. (Es ist klar, dass $\varnothing$ nicht im Wertebereich liegt, aber warum diese Abbildung injektiv ist, wollen wir hier nicht diskutieren.) Die resultierende Menge der natürlichen Zahlen bezeichnet man mit $\mathbb{N}$. Solange nicht geklärt ist, welche Zahl die Rolle des Anfangselements a spielt, bleiben wir vorerst bei der alten Bezeichnung.

Folgerung 2.2. *Es sei $P(n)$ ein Prädikat, das für $n \in N$ definiert ist. Angenommen, es gilt*

$$P(a) \qquad \text{und} \qquad \forall\, n \in N\,\big(P(n) \Rightarrow P(s(n))\big).$$

Dann gilt

$$\forall\, n \in N\,(P(n)).$$

Das folgt aus Satz 2.1(i), wenn man die Teilmenge

$$K = \{n \in N \mid P(n)\}$$

betrachtet. Daraus ergibt sich die Methode der *vollständigen Induktion*, mit der man die Gültigkeit einer Aussage über alle natürlichen Zahlen beweisen kann.

> Es genügt, ihre Gültigkeit für das Anfangselement zu zeigen
> (Induktionsanfang)
> und für eine beliebige natürliche Zahl n zu beweisen:
> Gilt sie für die Zahl n, so gilt sie auch für ihren Nachfolger
> (Induktionsschritt).

Bei Satz 2.1(ii) haben wir übrigens, genau besehen, die Aussage

$$|\{y \in M\,\big((x, y) \in G\big)\}| = 1$$

für alle $x \in N$ durch vollständige Induktion bewiesen.

An Stelle von U, i und a könnte man auch N, s und ein beliebiges Element n von N betrachten. Die kleinste Teilmenge von N, die n und mit jedem Element x auch $s(x)$ enthält, bezeichnen wir mit N_n. Dann zeigt man durch vollständige Induktion, dass

$$\{y \in N_n \mid \exists\, x \in N_n\,(s(x) = y)\} = N_n \setminus \{n\} = N_{s(n)}.$$

Wir setzen

$$I_n = N \setminus N_n,$$

dann ist beispielsweise

$$I_a = \varnothing, \qquad I_{s(a)} = \{a\}, \qquad I_{s(s(a))} = \{a, s(a)\}, \ldots$$

Definition 2.2. *Wir definieren eine Relation auf N durch*

$$x \preceq y \quad \Leftrightarrow \quad y \in N_x.$$

Satz 2.2. (i) *Diese Relation ist eine Wohlordnung, d. h. eine Ordnung, bezüglich derer jede nicht leere Teilmenge ein kleinstes Element besitzt.*

(ii) *Die echten Anfangsintervalle von N sind die Mengen*

$$I_n = \{x \in N \mid x \prec n\}.$$

Sie sind endlich (d. h. nicht unendlich im Sinne unserer Definition).

(iii) *Für jede Teilmenge L von N gibt es ein Anfangsintervall I von N und eine monoton wachsende bijektive Abbildung $f : I \to L$.*

Beweis. **(i)** Die Reflexivität ist klar, die Transitivität und Antisymmetrie gelten wegen $y \in N_x \Leftrightarrow N_y \subseteq N_x$.

Angenommen, die Teilmenge L besitzt kein kleinstes Element. Wir beweisen durch vollständige Induktion, dass dann für alle n gilt

$$L \subseteq N_n.$$

Wegen $N_a = N$ gilt die Aussage für $n = a$ (Induktionsanfang). Gilt die Aussage für ein Element n, so ist n kein Element von L, weil es sonst das kleinste wäre. Somit folgt $L \subseteq N_{s(n)}$ (Induktionsschritt). Hat L ein Element c, so ist steht das im Widerspruch zur bewiesenen Aussage im Fall $n = s(c)$, denn $c \notin N_{s(c)}$.

Durch Anwendung auf $L = \{x, y\}$ folgt die Totalität.

(ii) Ist I eine echte Teilmenge, so ist $N \setminus I$ nicht leer, hat also ein kleinstes Element n. Ist I außerdem Anfangsintervall, so kann kein Element, das größer als n ist, zu I gehören, also ist $N \setminus I = N_n$.

Wir beweisen die Endlichkeit des Anfangsintervalls I_n durch vollständige Induktion. Dies ist klar für $I_a = \varnothing$. Angenommen, die Aussage gilt für das Element n, aber nicht für $s(n)$, d. h. es gibt eine injektive Abbildung f von $I_{s(n)}$ in sich, die nicht surjektiv ist. Wenn es ein $b \in I_{s(n)}$ mit der Eigenschaft $f(b) = n$ gibt, dann muss es ein anderes Element $c \in I_{s(n)}$ geben, das nicht im Wertebereich liegt, und dann ändern wir f ab, indem wir $f(b) = c$ setzen. Der Wertebereich ist nun in I_n enthalten, und die Einschränkung und Beschränkung von f auf I_n ist eine injektive Abbildung von I_n in sich. Da aber $f(n)$ nun im Wertebereich fehlt, ist sie nicht surjektiv (Widerspruch).

(iii) Für $x \in L$ bezeichnen wir mit $r(x)$ das kleinste Element von $L \cap N_{s(x)}$, falls diese Menge nicht leer ist. Andernfalls setzen wir $r(x)$ gleich einem

Symbol z, das von allen Elementen von N verschieden ist. Setzen wir noch $r(z) = z$, so erhalten wir eine Abbildung r von $L \cup \{z\}$ in sich. Laut Satz 2.1 existiert eine Abbildung $f : N \to L \cup \{z\}$, so dass $f(a)$ das kleinste Element von L ist (oder z, falls $L = \varnothing$) und

$$f(s(x)) = r(f(x)).$$

Wir setzen $I = \{x \in N \mid f(x) \neq z\}$. Dann ist die Behauptung erfüllt (Übungsaufgabe). □

Satz 2.3. (i) *Jede endliche Kardinalzahl ist von der Form*

$$|I_n|$$

für ein eindeutig bestimmtes Element $n \in N$. Insbesondere ist $|I_a| = 0$.

(ii) *Für alle $n \in N$ gilt*

$$|I_{s(n)}| = |I_n| + 1.$$

(iii) *Für alle k, $n \in N$ gilt*

$$|I_k| \leq |I_n| \quad \Leftrightarrow \quad k \preceq n,$$

Beweis. **(i)** Es sei M eine Menge und $m = |M|$. Ist $m \geq |N|$, so ist M unendlich. Andernfalls ist nach Satz 1.9 $m \leq |N|$, also gibt es eine injektive Abbildung $M \to N$. Ist L ihr Wertebereich, so erhalten wir durch Verkettung mit der Abbildung aus Satz 2.1(iii) eine bijektive Abbildung von M auf ein Anfangsintervall I von N. Da M endlich und N unendlich ist, muss I ein echtes Anfangsintervall sein. Wäre $|I_k| = |I_n|$ für $k \prec n$, so wäre I_n gleichmächtig zu der echten Teilmenge I_k im Widerspruch zu Satz 2.1(ii).
(ii) folgt daraus, dass $I_{s(n)}$ die Vereinigung der disjunkten Mengen I_n und $\{n\}$ ist.
(iii) folgt aus

$$I_k \subseteq I_n \quad \Leftrightarrow \quad N_k \supseteq N_n \quad \Leftrightarrow \quad k \preceq n.$$ □

Da es nach dem Satz eine bijektive Beziehung zwischen endlichen Kardinalzahlen und den Elementen der Menge N gibt, wird im Tagesgeschäft der Mathematiker nicht streng zwischen beiden unterschieden. Man benutzt also die selben Bezeichnungen $0, 1, \ldots$ für beide Arten von Zahlen. Insbesondere bezeichnet man das Anfangselement a mit 0.

Satz 2.4. *Sind m und n natürliche Zahlen, so sind auch*

$$m + n, \qquad m \cdot n, \qquad n^m$$

natürliche Zahlen.

Die erste Behauptung gilt für $n = 0$. Angenommen, sie gilt für eine natürliche Zahl n, das heißt, $m + n$ ist eine natürliche Zahl. Dann gilt das Gleiche für $s(m + n)$, und nach Satz 2.3(ii) ist

$$s(m + n) = (m + n) + 1.$$

Nach dem selben Satz gilt

$$m + s(n) = m + (n + 1),$$

und Satz 1.7 sind die rechten Seiten gleich. Somit folgt die erste Behauptung durch vollständige Induktion. Der Beweis der anderen ist eine Übungsaufgabe.

Nun wollen wir die Subtraktion einführen.

Satz 2.5. *Sind l und m Kardinalzahlen und ist*

$$l \leq m,$$

so gibt es eine Kardinalzahl k mit der Eigenschaft

$$k + l = m.$$

Sind l und m natürliche Zahlen, so ist auch k eine natürliche Zahl und eindeutig bestimmt.

Beweis. Ist $l = |L|$ und $m = |M|$, so gibt es eine injektive Abbildung $L \to M$, und wir können annehmen, dass L eine Teilmenge von M ist. Nun gilt die behauptete Gleichung für $k = |M \setminus L|$. Gilt die Gleichung für irgendein $k = |K|$, wobei K und L disjunkt sind, so gibt es eine bijektive Abbildung $K \cup L \to M$, und durch Einschränkung erhalten wir eine injektive Abbildung $K \to M$, so dass $k \leq m$ ist. Ist also m eine natürliche Zahl, so auch k.

Noch zu zeigen:
Für alle natürlichen Zahlen k_1, k_2 und l gilt

$$k_1 + l = k_2 + l \quad \Rightarrow \quad k_1 = k_2.$$

Wir tun dies durch vollständige Induktion nach l. Für $l = 0$ ist nichts zu zeigen. Damit ist der Induktionsanfang erledigt. Angenommen, die Behauptung gilt für eine Zahl l. Nun sei

$$k_1 + s(l) = k_2 + s(l).$$

Nach Satz 2.3(ii) bedeutet das

$$k_1 + (l + 1) = k_2 + (l + 1)$$

und nach dem Assoziativgesetz aus Satz 1.7

$$(k_1 + l) + 1 = (k_2 + l) + 1.$$

Nach Satz 2.4 sind $k_1 + l$ und $k_2 + l$ natürliche Zahlen, also folgt

$$s(k_1 + l) = s(k_2 + l).$$

Wegen der Injektivität von s folgt

$$k_1 + l = k_2 + l$$

und mit der Induktionsvoraussetzung

$$k_1 = k_2.$$

□

Die letzte Aussage des Satzes rechtfertigt die Schreibweise $k = m - l$ für die *Differenz* natürliche Zahlen. Für unendliche Kardinalzahlen ist die Aussage falsch: Offensichtlich gilt $|N| = |I_n| + |N_n|$ für alle $n \in N$, und wegen Folgerung 2.1 ist $|N_n| = |N|$.

In Satz 1.10 hatten wir gesehen, dass gewisse Mengen sehr mächtig sein können. Nun zeigen wir, dass gewisse Mengen weniger mächtig sind, als man denken könnte.

Satz 2.6. *Die Mengen* $\mathbb{N}$ *und* $\mathbb{N} \times \mathbb{N}$ *sind gleichmächtig.*

Beweis. Wir definieren eine Abbildung r der letzteren Menge in sich durch

$$r(x, y) = \begin{cases} (x - 1, y + 1), & \text{falls } x > 0, \\ (y + 1, 0), & \text{falls } x = 0. \end{cases}$$

Nach Satz 2.1(ii) gibt es eine eindeutig bestimmte Abbildung $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ mit den Eigenschaften

$$f(0) = (0, 0), \qquad \forall n \in \mathbb{N} \big(f(n + 1) = r(f(n))\big).$$

Es ist nicht schwer zu sehen, dass f bijektiv ist. □

Man nennt diese Beweismethode das *erste Cantorsche Diagonalverfahren*.

$$\begin{array}{ccccc} \ddots & \ddots & \ddots & \ddots & \ddots \\ f(9) & f(13) & f(18) & f(24) & \ddots \\ f(5) & f(8) & f(12) & f(17) & \ddots \\ f(2) & f(4) & f(7) & f(11) & \ddots \\ f(0) & f(1) & f(3) & f(6) & \ddots \end{array}$$

Definition 2.3. *Eine* Folge *in einer Menge M ist eine Abbildung von einem Intervall I in der Menge $\mathbb{N}$ der natürlichen Zahlen in die Menge M. Man nennt die Folge endlich, wenn I endlich ist.*

Es ist üblich, die Werte einer Folge nicht mit $f(n)$, $g(n)$ o. ä., sondern mit x_n, y_n o. ä. zu bezeichnen und die Glieder der Folge zu nennen. Man nennt tiefgestellte Buchstaben oder Zahlen Indizes und I die Indexmenge. Dann kann man eine Folge mit der Indexmenge $\{1, 2, \ldots, n\}$ als n-Tupel

$$(x_1, x_2, \ldots, x_n)$$

betrachten. Für eine Folge mit Indexmenge I schreibt man $(x_n)_{n\in I}$.

Falls jemand darüber beunruhigt ist, dass der Beweis von Satz 1.8 auf den Begriff der Folge vorgreift: Er kann (und sollte) ganz ohne Verwendung von Folgen geführt werden. Sie sind lediglich geeignet, die Beweisidee besser zu verstehen.

Folgerung 2.3. *Ist $r : M \to M$ eine Abbildung und $b \in M$, so gibt es genau eine Folge $(x_n)_{n\in\mathbb{N}}$ mit den Eigenschaften*

$$x_0 = b, \qquad \forall\, n \in \mathbb{N}\, \big(x_{n+1} = r(x_n)\big).$$

Dies folgt aus Satz 2.1(ii). Man sagt dann, dass die Folge *rekursiv definiert* ist.

2.2 Angeordnete Körper

Vorlesung 4b
29.4.2022

Aus der Schule ist der Bereich $\mathbb{Q}$ der rationalen Zahlen bekannt. Dieser ist ein Körper im folgenden Sinne.

Definition 2.4. *Ein* Körper *ist eine Menge K zusammen mit zwei Abbildungen $K \times K \to K$, genannt Addition und Multiplikation und abgekürzt in der bekannten Weise, mit folgenden Eigenschaften:*

(i) *Es gelten die Kommutativ- und Assoziativgesetze für Addition und Multiplikation sowie das Distributivgesetz.*

(ii) *K hat verschiedene Elemente 0 und 1, so dass für alle Elemente x von K gilt $x + 0 = x$ und $x \cdot 1 = x$.*

(iii) *Für beliebige Elemente x und y von K gibt es ein Element d von K, so dass $x = y + d$.*

(iv) *Für beliebige Elemente x von K und y von $K \setminus \{0\}$ gibt es ein Element q von K, so dass $x = y \cdot q$.*

Es ist leicht zu sehen, dass die Elemente 0, 1 und (für gegebene x und y) d und q eindeutig bestimmt sind. Man nennt $x + y$ die *Summe*, $x \cdot y$ das *Produkt*, d die *Differenz* und q den *Quotienten* von x und y. Man schreibt $d = x - y$, $q = x : y$, und statt $0 - y$ schreibt man auch $-y$ und nennt es das *entgegengesetzte Element* von y. Schließlich nennt man $1 : y$ das *inverse Element* von y. Abbildungen von einer Menge in einen Körper nennt man *Funktionen* auf dieser Menge.

In dieser Vorlesung werden wir die Konstruktion von $\mathbb{Q}$, die algebraischer Natur ist, nicht durchführen. Die Elemente entstehen als Äquivalenzklassen $\frac{a}{b}$ von Paaren (a, b) ganzer Zahlen mit der Eigenschaft $b \neq 0$. Eine ganze Zahl a wird mit der rationalen Zahl $\frac{a}{1}$ identifiziert, so dass der Bereich $\mathbb{Z}$ der ganzen Zahlen zu einer Teilmenge von $\mathbb{Q}$ wird, und dann ist $\frac{a}{b}$ das Selbe wie $a : b$. Darum wird der Bruchstrich gern als platzsparendes Synonym für das Divisionszeichen missbraucht.

In der Schule werden auch reelle Zahlen besprochen, allerdings nicht rigoros konstruiert. Dies soll hier später ausgeführt werden. Um danach nicht alles zu wiederholen, was für rationale und reelle Zahlen gleichermaßen gilt, werden wir solche Aussagen allgemein für Körper formulieren.

Da die natürlichen Zahlen im Algemeinen nicht in einem Körper K enthalten sind, definieren wir rekursiv für $n \in \mathbb{N}$ und $x \in K$

$$\begin{aligned} 0 \cdot x &= 0, \qquad & (n+1) \cdot x &= n \cdot x + x, \\ x^0 &= 1, \qquad & x^{n+1} &= x^n \cdot x, \end{aligned}$$

wobei 0 und 1 auf der linken Seite der Gleichungen Elemente von $\mathbb{N}$ bezeichnen und auf der rechten Seite der Gleichungen Elemente von K. Das Multiplikationszeichen wird meist weggelassen, wenn es nicht zwischen zwei Zahlen steht. Damit Distributiv- und Potenzgesetze weiter gelten, definieren wir außerdem

$$(-n) \cdot x = -(n \cdot x), \qquad x^{-n} = 1 : x^n.$$

Im Unterschied zur Algebra spielt in der Analysis die Anordnung der Körperelemente eine wichtige Rolle.

Definition 2.5. *Ein* angeordneter Körper *ist ein Körper K zusammen mit einer Teilmenge K_+, die folgende Eigenschaften hat:*

(i) *Für jedes Element x von K gilt genau eine der Aussagen $x \in K_+$, $-x \in K_+$, $x = 0$.*

(ii) *Für alle Elemente x und y von K_+ ist $x + y \in K_+$ und $x \cdot y \in K_+$.*

Wir sagen, dass x kleiner als y ist, abgekürzt $x < y$, wenn $y - x \in K_+$ ist. Die Elemente mit den drei Eigenschaften in (i) nennen wir positive, negative *bzw.* verschwindende *Elemente von K.*

Aus der Schule ist bekannt, dass die Menge $\mathbb{Q}_+$, die aus allen rationalen Zahlen $\frac{m}{n}$ besteht, bei denen m und n positive ganze Zahlen sind, diese Eigenschaften hat, so dass $\mathbb{Q}$ zu einem angeordneten Körper wird.

Aus der Definition folgt sofort für alle Elemente x, y, z von K:

$$\begin{aligned} x < y \ \wedge \ y < z \quad &\Rightarrow \quad x < z, \\ x < y \quad &\Rightarrow \quad -x > -y, \\ x < y \quad &\Rightarrow \quad x + z < y + z, \\ x < y \ \wedge \ z > 0 \quad &\Rightarrow \quad x \cdot z < y \cdot z. \end{aligned}$$

Mit Hilfe der ersten Regel (Transitivität) können wir die letzten beiden (genannt *Monotoniegesetze*) verallgemeinern:

$$\begin{aligned} x < y \ \wedge \ z < w \quad &\Rightarrow \quad x + z < y + w, \\ 0 < x < y \ \wedge \ 0 < z < w \quad &\Rightarrow \quad x \cdot z < y \cdot w, \end{aligned} \tag{1}$$

und aus der zweiten und der letzten Regel folgt

$$x < y \ \wedge \ z < 0 \quad \Rightarrow \quad x \cdot z > y \cdot z.$$

Unterscheiden wir die Fälle in Definition 2.5(i), so folgt nun

$$x \neq 0 \quad \Rightarrow \quad x \cdot x > 0.$$

Schließlich gilt

$$x > 0 \quad \Rightarrow \quad \frac{1}{x} > 0,$$

weil die anderen Fälle für $\frac{1}{x}$ zum Widerspruch führen würden.

Die Relation $<$ ist eine strikte Ordnung. Eigenschaft (i), auch Trichotomie genannt, impliziert sowohl Antisymmetrie als auch Totaltiät der zugehörigen nicht-strikten Ordnung $\leq$.

Wir beweisen nun zwei grundlegende Ungleichungen, die für angerodnete Körper gelten.

Satz 2.7 (Bernoulli-Ungleichung)**.** *Für alle ganzen Zahlen n und alle Elemente x von K mit der Eigenschaft $x > -1$ gilt*

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

Beweis. Für $n = 0$ ist nichts zu beweisen. Angenommen, die Ungleichung gilt für eine natürliche Zahl n. Dann ist laut rekursiver Definition

$$(1 + x)^{n+1} = (1 + x)^n(1 + x).$$

Aus Induktionsvoraussetzung und Monotoniegesetz folgt wegen $1 + x > 0$

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

Nach den Rechengesetzen erhalten wir

$$(1 + nx)(1 + x) = (1 + nx) + (x + nx^2) = 1 + (n + 1)x + nx^2.$$

Wegen $n \geq 0$ und $x^2 \geq 0$ ist nach den Monotoniegesetzen

$$1 + (n + 1)x + nx^2 \geq 1 + (n + 1)x.$$

Nun folgt die Behauptung für $n + 1$ statt n mit der Transitivität.

Der Beweis für $n < 0$ ist eine Übungsaufgabe. □

Satz 2.8. *Für jede natürliche Zahl n, die größer oder gleich 1 ist, und alle Elemente $x_1, \ldots, x_n$ von K_+ mit der Eigenschaft*

$$x_1 \cdots x_n = 1$$

gilt

$$x_1 + \ldots + x_n \geq n \cdot 1.$$

Das Einselement von K kann auf der rechten Seite der letzten Ungleichung weggelassen werden, wenn $\mathbb{N}$ in K enthalten ist.

Beweis. Der Induktionsanfang $n = 1$ ist trivial. Angenommen, die Behauptung gilt für n Zahlen. Haben wir nun $x_1 \cdots x_n x_{n+1} = 1$, so wenden wir die Induktionsvoraussetzung auf die Zahlen $x_1, \ldots, x_{n-1}$ und $x_n x_{n+1}$ an und erhalten

$$x_1 + \cdots + x_{n-1} + x_n x_{n+1} \geq n.$$

Addiere 1 auf beiden Seiten. Es genügt zu zeigen, dass

$$x_n + x_{n+1} \geq x_n x_{n+1} + 1,$$

d. h. $x_n + x_{n+1} - x_n x_{n+1} - 1 \geq 0$, d. h.

$$(x_n - 1)(1 - x_{n+1}) \geq 0.$$

Wären alle Faktoren kleiner als 1, so auch das Produkt. Somit muss es einen Index i geben, so dass $x_i \geq 1$ ist, und analog muss es ein j geben, so dass $x_j \leq 1$ ist. Ohne Beschränkung der Allgemeinheit ist $i = n$ und $j = n + 1$. □

Mit etwas mehr Sorgfalt kann man sogar beweisen, dass in der behaupteten Ungleichung nur dann Gleichheit eintreten kann, wenn alle n Elemente gleich 1 sind.

Die Verwendung von Fortsetzungspunkten im letzten Satz und Beweis war nicht rigoros. Darum definiert man Summen- und Produktzeichen für festen Anfangsindex m rekursiv durch

$$\sum_{i=m}^{m} x_i = x_m, \qquad \prod_{i=m}^{m} x_i = x_m,$$
$$\sum_{i=m}^{n+1} x_i = \sum_{i=m}^{n} x_i + x_{n+1}, \qquad \prod_{i=m}^{n+1} x_i = \prod_{i=m}^{n} x_i \cdot x_{n+1}$$

Dabei ist der Index (hier mit i bezeichnet) eine Variable, die außerhalb des Summen- oder Produktterms keine Bedeutung hat. Die Aussage von Satz 2.8 schreibt sich nun als

$$\prod_{i=1}^{n} x_i = 1 \quad \Rightarrow \quad \sum_{i=1}^{n} x_i \geq n.$$

Als weitere Anwendung beweisen wir die *verallgemeinerte dritte binomische Formel*

$$(x-y)\sum_{i=1}^{n} x^{n-i}y^{i-1} = x^n - y^n.$$

für $x, y \in K$ und $n \in \mathbb{N}\backslash\{0\}$. Die linke Seite ist nach einem verallgemeinerten Distributivgesetz gleich

$$\sum_{i=1}^{n} x^{n+1-i}y^{i-1} - \sum_{i=1}^{n-1} x^{n-i}y^{i}.$$

In der zweiten Summe substituieren wir $i = j - 1$ und erhalten

$$\sum_{i=1}^{n} x^{n+1-i}y^{i-1} - \sum_{j=2}^{n+1} x^{n+1-j}y^{j-1}.$$

Dann substituieren wir wieder j durch i und bemerken, dass sich die Terme mit $i \in \{2, \ldots, n\}$ aufheben. Die hier verwendeten Regeln müssen streng genommen durch vollständige Induktion beweisen werden. Die bewiesene Formel zeigt, dass für $n > 0$ die durch

$$f(x) = x^n$$

definierte Funktion $f : K_+ \to K_+$ streng monoton wachsend ist.

2.3 Schranken und Grenzen

Vorlesung 5a
4.5.2022

In technischen und medizinischen Anwendungen sind oft Schranken für zulässige Werte vorgegeben.

Definition 2.6. *Es sei M eine Teilmenge eines angeordneten Körpers K. Ein Element a von K heißt* obere Schranke *von M, wenn für alle Elemente x von M gilt $x \leq a$. Besitzt M eine obere Schranke, so sagt man, die Menge M sei* von oben beschränkt.

Analog definiert man untere Schranken *und* von unten beschränkte *Mengen.*

Eine Teilmenge von K heißt beschränkt*, wenn sie sowohl von oben als auch von unten beschränkt ist.*

Beispiel. Hat die Menge M ein größtes Element (auch als *Maximum* von M bezeichnet, abgekürzt $\max M$), so ist dieses auch eine obere Schranke von M, und hat sie ein kleinstes Element (auch als *Minimum* von M bezeichnet, abgekürzt $\min M$), so ist dieses eine untere Schranke. $\triangleleft$

Beispiel. Ist n eine natürliche Zahl, so ist n nach Satz 2.8 eine untere Schranke (und sogar das kleinste Element) für die Menge

$$M = \{x_1 + \ldots + x_n \mid x_1 > 0, \ldots, x_n > 0, x_1 x_2 \cdots x_n = 1\}.$$

Sie hat für $n > 1$ übrigens keine obere Schranke. $\triangleleft$

Wir sagen, dass eine Folge in K von oben/unten beschränkt ist, wenn das für die Menge ihrer Glieder gilt, und wir sagen, dass eine Funktion mit Werten in K von oben/unten beschränkt ist, wenn das für ihren Wertebereich gilt. So ist z. B. $(q^n)_{n\in\mathbb{N}}$ für $q = 1 + x > 1$ wegen Satz 2.7 nicht beschränkt.

Satz 2.9. *Ist $q \in K$ und $k \in \mathbb{N}$, wobei $0 < q < 1$ ist, so ist die Folge*

$$(n^k q^n)_{n\in\mathbb{N}}$$

beschränkt.

Beweis. Nach Satz 2.7 gilt für alle $x > -1$ und $n \in N$

$$(1 + x)^n \geq 1 + nx > nx, \qquad (1 + x)^{-k} \geq 1 - kx.$$

Wir können x so wählen, dass $1 - kx = q$ ist. Dann ist $x > 0$, und mit den Monotonie- und Potenzgesetzen folgt

$$q^{-n} \geq ((1 + x)^{-k})^{-n} = (1 + x)^{kn} = ((1 + x)^n)^k > (nx)^k.$$

Multiplizieren wir beide Seiten mit q^n und dividieren durch x^k, so sehen wir, dass x^{-k} eine obere Schranke ist. Nach den Monotoniegesetzen ist 0 eine untere Schranke. $\square$

Das Finden von Schranken nennt man *Abschätzen*, es ist eine typische Methode der Analysis. Ist a eine obere Schranke und $b > a$, so ist auch b eine. Die Menge der oberen Schranken einer Menge M ist also ein Endintervall von K, die Menge der unteren Schranken ein Anfangsintervall. Je kleiner die obere Schranke, desto mehr Information enthält sie also. Manchmal genügen grobe Abschätzungen oder die bloße Beschränktheit, manchmal braucht man bestmögliche Ergebnisse.

Definition 2.7. *Gibt es unter allen oberen Schranken einer Menge M eine kleinste, so heißt sie* obere Grenze *oder* Supremum *von M, abgekürzt* $\sup M$*. Analog definiert man die* untere Grenze *oder das* Infimum*, abgekürzt* $\inf M$*, als größte untere Schranke.*

So ist z. B. $\inf K_+ = 0$, aber K_+ hat kein kleinstes Element, denn mit jedem Element x enthält es das kleinere Element $x : 2$. Hat eine Menge M ein größtes Element, so ist dieses auch ihr Supremum.

Beispiel. Wir halten $a \in K_+$ fest betrachten die Menge

$$M = \{x \in K_+ \mid x^2 < a\}.$$

Wegen der strengen Monotonie der Quadratfunktion auf K_+ besteht

$$A = \{y \in K_+ \mid y^2 > a\}$$

aus oberen Schranken von M, und es kann höchstens ein Element $c \in K$ mit der Eigenschaft

$$c^2 = a$$

geben. Es gilt $\min\{\frac{1}{2}, a\} \in M$ und, nach Satz 2.7, $1 + \frac{a}{n} \in A$.

Ist $x \in M$ gegeben, so ist $y = a : x$ in A, ist hingegen $y \in A$ gegeben, so ist $x = a : y$ in M, und in beiden Fällen ist $xy = a$. Das arithmetische Mittel $y' = \frac{x+y}{2}$ ist in A, und $y' \leq y$, weil

$$y'^2 - a = \left(\frac{y-x}{2}\right)^2, \qquad y - y' = \frac{y-x}{2},$$

während $x' = a : y'$ in M ist. Es folgt also

$$x < x' < y' < y,$$

und somit hat M kein größtes Element und A kein kleinstes Element. Wenn c existiert, so ist es die obere Grenze von M, aber wenn c nicht existiert, so hat M keine obere Grenze.

Für den Körper $K = \mathbb{Q}$ kann tatsächlich Letzteres eintreten, zum Beispiel für $a = 2$. Nach Definition der rationalen Zahlen müsste $c = \frac{m}{n}$ mit m, $n \in \mathbb{N} \setminus \{0\}$ sein,wobei wir annehmen können, dass m und n teilerfremd sind. Die Gleichung $c^2 = 2$ bedeutet dann

$$m^2 = 2n^2.$$

Nun müsste m gerade sein, also $m = 2k$, und wegen $4k^2 = 2n^2$ folgt dann

$$2k^2 = n^2.$$

Dann müsste auch n gerade sein, was der Teilerfremdheit widerspricht. $\triangleleft$

Es gibt Rechenregeln für Grenzen wie z. B. die folgende.

Lemma 2.1. *Hat M das Supremum c und N das Supremum d, so hat die Menge*

$$M + N = \{x + y \mid x \in M,\, y \in N\}$$

das Supremum $c + d$.

Beweis. Es folgt aus dem Monotoniegesetz der Addition, dass $c + d$ eine obere Schranke ist. Wir behaupten, dass eine beliebige Zahl e, die kleiner als $c + d$ ist, keine obere Schranke von $M + N$ sein kann. Wir können nämlich $e = c + d - 2\varepsilon$ schreiben, wobei $\varepsilon > 0$ ist. Da $c - \varepsilon$ keine obere Schranke von M ist, gibt es ein $x \in M$ mit der Eigenschaft $x > c - \varepsilon$, und analog findet man $y \in N$ mit $y > d - \varepsilon$. Mit dem Monotoniegesetz folgt $x + y > e$. $\square$

Das Arbeiten in unvollständigen Körpern wie $\mathbb{Q}$ ist anstrengend, denn man muss ständig prüfen, ob Suprema und Infima tatsächlich existieren.

Definition 2.8. *Ein angeordneter Körper K heißt* vollständig*, wenn jede von oben beschränkte Teilmenge ein Supremum hat.*

Man hätte genauso gut verlangen können, dass jede von unten beschränkte Menge ein Infimum hat, denn ist c das Infimum von M, so ist $-c$ das Supremum von $-M = \{-x \mid x \in M\}$. Es gibt eine weitere Umformulierung, die den folgenden Begriff benutzt.

Definition 2.9. *Ein* Schnitt *in einem angeordneten Körper ist eine Menge $\{A, B\}$ mit folgenden Eigenschaften:*

- *A, B sind nichtleere Teilmengen von K,*
- *A ist die Menge der oberen Schranken von B außer der kleinsten,*
- *B ist die Menge der unteren Schranken von A außer der größten.*

Wir nennen A die Obermenge *und B die* Untermenge *des Schnitts. Ein Schnitt $\{A, B\}$ heißt* trivial, *wenn es ein c gibt, so dass*

$$A = \{x \in K \mid x > c\}, \qquad B = \{x \in K \mid x < c\}.$$

Wenn es keine kleinste obere Schranke von B bzw. keine größte untere Schranke von A gibt, ist der jeweilige Zusatz überflüssig. Er dient dem Zweck, dass in jedem Schnitt einheitlich die Obermenge ein Endintervall ohne kleinstes Element und die Untermenge ein Anfangsintervall ohne größtes Element ist.

Satz 2.10. *Ein angeordneter Körper ist genau dann vollständig, wenn jeder Schnitt trivial ist.*

Beweisidee: Ist M eine von oben beschränkte nichtleere Teilmenge, A die Menge ihrer oberen Schranken außer der kleinsten und B die Menge der unteren Schranken von A außer der größten, so ist $\{A, B\}$ ein Schnitt. Ist er trivial, so ist das Element c das Supremum von M.

Ist umgekehrt $\{A, B\}$ ein Schnitt und existieren $c = \sup B$ und $d = \inf A$, so ist $c = d$, und der Schnitt ist trivial. $\square$

Die Einzelheiten sind Inhalt einer Übungsaufgabe. Wir merken an, dass die Definitionen und Sätze auf beliebige geordnete Mengen anwendbar sind.

2.4 Reelle Zahlen

Vorlesung 5b
6.5.2022

Wir wollen der Körper $\mathbb{Q}$ der rationalen Zahlen zu einem vollständigen angeordneten Körper erweitern. Eine Möglichkeit wäre, für jeden nichttrivialen Schnitt ein zusätzliches Element zwischen Untermenge und Obermenge einfügen, um den Schnitt trivial zu machen. Dabei entsteht allerdings die Frage, woher man diese Elemente nehmen sollte, und zweitens müsste man in den späteren Beweisen eine unschöne Fallunterscheidung vornehmen. Geschickter ist es, da ja die Menge der hinzukommenden Elemente in Bijektion zur Menge der nichttrivialen Schnitte stünde, einfach die letztere Menge dafür zu nutzen, und der Einheitlichkeit halber auch die Elemente von $\mathbb{Q}$ durch die zugehörigen Schnitte zu ersetzen.

Definition 2.10. *Die Menge $\mathbb{R}$ der reellen Zahlen ist die Menge der Schnitte in der geordneten Menge $\mathbb{Q}$. Sind $\{A, B\}$ und $\{C, D\}$ Schnitte, wobei B und D die Untermengen sind, so setzen wir*

$$\{A, B\} \leq \{C, D\} \quad \Leftrightarrow \quad B \subseteq D.$$

Statt $B \subseteq D$ hätte man auch die äquivalente Bedingung $A \supseteq C$ benutzen können. Jetzt ist $\mathbb{Q}$ zwar keine Teilmenge von $\mathbb{R}$, aber die einfache Struktur von $\mathbb{R}$ ist uns wichtiger. Dies ist eine gewohnte Situation, denn auch die natürlichen Zahlen sind zunächst keine Kardinalzahlen und die ganzen Zahlen keine rationalen Zahlen.

Satz 2.11. *Die Menge $\mathbb{R}$ mit der angegebenen Relation ist eine vollständige geordnete Menge, und die Abbildung $\mathbb{Q} \to \mathbb{R}$, die jeder rationalen Zahl den zugehörigen trivialen Schnitt zuordnet, ist streng monoton wachsend, also insbesondere injektiv.*

Die letzte Aussage rechtfertigt es, dass man jede rationale Zahl mit dem zugehörigen Schnitt identifiziert.

Beweis. Die Antisymmetrie und die Transitivität der genannten Relation folgen aus den analogen Eigenschaften der Teilmengenrelation. Zum Beweis der Totalität nehmen wir an, in den Bezeichnungen der Definition sei die Bedingung $B \subseteq D$ nicht erfüllt. Dann gibt es also ein Element b von $B \setminus D$. Da D ein Anfangsintervall ist, kann es kein Element d von D mit der Eigenschaft $d > b$ geben, mit anderen Worten, b ist eine obere Schranke von D. Da B ein Anfangsintervall ist, folgt daraus $D \subseteq B$. Aus der Negation von $\{A, B\} \leq \{C, D\}$ folgt also $\{A, B\} \geq \{C, D\}$.

Die Monotonie der Abbildung $\mathbb{Q} \to \mathbb{R}$ folgt direkt aus der Definition von trivialen Schnitten und der Relation $\leq$ auf $\mathbb{R}$. Damit ist die letzte Aussage des Satzes bewiesen.

Zum Beweis der Vollständigkeit betrachten wir einen Schnitt $\{\mathcal{E}, \mathcal{F}\}$ in $\mathbb{R}$, wobei $\mathcal{E}$ die Obermenge ist. Wir setzen

E gleich der Vereinigung aller Obermengen A von Schnitten $\{A, B\} \in \mathcal{E}$,
F gleich der Vereinigung aller Untermengen D von Schnitten $\{C, D\} \in \mathcal{F}$

und behaupten, dass $\{E, F\}$ ein Schnitt mit der Obermenge E und der Untermenge F ist. Ebenso wie die Mengen A bzw. D hat E kein kleinstes und F kein größtes Element, und sie sind nicht leer.

Ist $e \in E$ und $f \in F$, so gibt es $\{A, B\} \in \mathcal{E}$ und $\{C, D\} \in \mathcal{F}$, so dass $e \in A$ und $f \in D$. Da $\{\mathcal{E}, \mathcal{F}\}$ ein Schnitt ist, gilt $\{A, B\} \geq \{C, D\}$, d. h. $B \supseteq D$ und somit $f \in B$, und da $\{A, B\}$ ein Schnitt ist, gilt $e > f$. Somit sind E und F disjunkt, E besteht aus oberen Schranken von F und F aus unteren Schranken von E.

Angenommen, es gibt rationale Zahlen $y < x$, die weder zu E noch zu F gehören, und es seien $\{C, D\} < \{A, B\}$ die zugehörigen trivialen Schnitte. Jedes Element von $\mathcal{E}$ ist größer oder gleich $\{A, B\}$, denn x ist eine untere Schranke seiner Obermenge, und analog ist jedes Element von $\mathcal{F}$ kleiner als $\{C, D\}$, denn y ist eine obere Schranke seiner Untermenge. Dies ist ein Widerspruch, denn zwischen Unter- und Obermenge von $\{\mathcal{E}, \mathcal{F}\}$ kann nicht mehr als ein Element liegen. Somit ist auch $\{E, F\}$ ein Schnitt. $\square$

Wir wollen nun die Menge $\mathbb{R}$ mit Rechenoperationen versehen.

Definition 2.11. *Sind $\{A, B\}$ und $\{C, D\}$ Schnitte mit den Obermengen A und C, so setzen wir*

$$\{A, B\} + \{C, D\} = \{A + C, B + D\}.$$

Satz 2.12. (i) *Definition 2.11 ist korrekt.*

(ii) *Die so definierte Addition auf $\mathbb{R}$ ist kommutativ und assoziativ.*

(iii) *Der Nullschnitt hat die Eigenschaft aus Definition 2.4(ii), und Eigenschaft (iii) aus Definition 2.4 ist erfüllt.*

(iv) *Es gilt das Monotoniegesetz der Addition.*

(v) *Die Abbildung $\mathbb{Q} \to \mathbb{R}$ ist verträglich mit der Addition.*

Zum Beweis benötigen wir folgende Eigenschaft von Schnitten:

Lemma 2.2. *Ist A die Obermenge und B die Untermenge eines Schnitts, so gibt es für jede positive Zahl $\varepsilon \in \mathbb{Q}$ Elemente $a \in A$ und $b \in B$, so dass $a - b < \varepsilon$.*

Beweis. Wir wählen $x \in A$ und $d \in \,]0, \varepsilon/2[$ und setzen $x_n = x + nd$. Dann ist $J = \{i \in \mathbb{N} \mid x_i \in B\}$ ein nichtleeres Anfangsinterall und $I = \{i \in \mathbb{N} \mid x_i \in A\}$ ein nichtleeres Endintervall von $\mathbb{N}$, und es gibt höchstens einen Index i, so dass $x_i \notin A \cup B$. Setzen wir $k = \max J$, $b = x_k$, $l = \min I$ und $a = x_l$, so ist $a - b \leq 2d < \varepsilon$. □

Beweis des Satzes. **(i)** Ist $a \in A$, $c \in C$ und $x > a + c$, so ist $x - c > a$. Da A ein Endintervall ist, folgt $x - c \in A$, also $x = (x - c) + c \in A + C$. Somit ist auch $A + C$ ein Endintervall. Analog sieht man, dass $B + D$ ein Anfangsintervall ist.

Ist $a \in A$, $c \in C$, $b \in B$ und $d \in D$, so ist $a > b$ und $c > d$, also $a + c > b + d$. Angenommen, es gäbe Elemente $x > y$ zwischen $A + C$ und $B + D$. Nach dem Lemma gäbe es dann Elemente $a \in A$, $b \in B$, $c \in C$ und $d \in D$, so dass $a - b < (x - y) : 2$ und $c - d < (x - y) : 2$. Dann wäre $(a + c) - (b + d) < x - y$ im Widerspruch zu $a + c > x$ und $b + d < y$.

(ii) Die Kommutativität der Addition von Schnitten ist offensichtlich. Ihre Assoziativität folgt aus der Assoziativität der Addition von Teilmengen, die wiederum aus der Assoziativität der Addition in $\mathbb{Q}$ folgt.

(iii) Der zur Zahl 0 ehörige triviale Schnitt ist offensichtlich der Nullschnitt. Für jeden Schnitt $\{A, B\}$ ist $\{A, B\} + \{-B, -A\}$ gleich dem Nullschnitt, und das genügt für die Existenz der Subtraktion.

(iv) ergibt sich daraus, dass für beliebige Teilmengen B, D und F aus $B \subseteq D$ folgt $B + F \subseteq D + F$.

(v) folgt aus dem Monotoniegesetz der Addition. □

Bei der Definition der Multiplikation ergibt sich eine Schwierigkeit. Um sie zu beschreiben, nennen wir eine Schnittkomponente gemischt, wenn sie sowohl positive als auch negative Elemente enthält. Andernfalls nennen wir sie rein. Beim Nullschnitt sind beide Komponenten rein, alle anderen Schnitte haben je eine reine und eine gemischte Komponente. Ist B gemischt, so ist

$$B \cdot D = \{b \cdot d \mid b \in B,\, d \in D\}$$

gleich der gesamten Menge $\mathbb{Q}$ unabhängig von D. Selbst wenn man nur Komponenten gleicher Sorte miteinander multipliziert, bleibt das Problem, wenn beide gemischt sind. Die Ursache ist, dass eine gemischte Komponente auch Elemente mit dem falschen Vorzeichen enthält. Die mit dem richtigen Vorzeichen bilden eine Teilmenge

$$\tilde{B} = \{b \in B \mid \forall\, x \in \mathbb{Q} \setminus B \; (b \cdot x > 0)\}.$$

Es gilt

$$B \cdot \tilde{D} = \tilde{B} \cdot D,$$

und wir bezeichnen diese Menge als modifiziertes Produkt $B * D$ der gemischten Komponenten B und D.

Definition 2.12. *Sind $\{A, B\}$ und $\{C, D\}$ Schnitte mit den reinen Komponenten A und C sowie den gemischten Komponenten B und D, so setzen wir*

$$\{A, B\} \cdot \{C, D\} = \{A \cdot C, B * D\}.$$

Ist außerdem $\{E, F\}$ der Nullschnitt, so setzen wir

$$\{A, B\} \cdot \{E, F\} = \{E, F\}.$$

Wir haben Schnitte nicht als geordnete Paare (A, B) definiert, weil in der Schreibweise $\{A, B\}$ nicht unbedingt die Obermenge an der ersten und die Untermenge an der zweiten Stelle stehen muss. Damit vermeiden wir eine weitere Fallunterscheidung in der Definition.

Satz 2.13. (i) *Definition 2.12 ist korrekt.*

(ii) *Die so definierte Multiplikation auf $\mathbb{R}$ ist kommutativ und assoziativ, und es gilt das Distributivgesetz.*

(iii) *Der dem Einselement entsprechende triviale Schnitt hat die Eigenschaft aus Definition 2.4(ii), und Eigenschaft (iv) aus Definition 2.4 ist erfüllt.*

(iv) *Es gilt das Monotoniegesetz der Multiplikation.*

(v) *Die Abbildung $\mathbb{Q} \to \mathbb{R}$ ist verträglich mit der Multiplikation.*

[Beweis wird noch eingefügt.]

Fassen wir die Sätze 2.11, 2.12 und 2.13 zusammen, so ergibt sich

Folgerung 2.4. *Die Menge $\mathbb{R}$ mit den oben beschriebenen Rechenoperationen und der Ordnung ist ein vollständiger angeordneter Körper.*

Man kann sogar zeigen, dass $\mathbb{R}$ bis auf Isomorphie der einzige vollständige angeordnete Körper ist, der dem *Archimedischen Axiom* genügt, welches besagt, dass es für jedes Element x ein Vielfaches des Einselements gibt, das größer als x ist. Die konkrete Konstruktion mit Hilfe von Schnitten ist dabei unwichtig. In der Tat gibt es andere Konstruktionen. Man spricht also einfach vom Körper der reellen Zahlen, und der Körper $\mathbb{Q}$ ist der kleinste Teilkörper. Die Elemente von $\mathbb{R}$, die nicht in $\mathbb{Q}$ liegen, nennt man *irrationale Zahlen*.

Wenn man in Definition 2.9 die Bedingung (i), dass Ober- und Untermenge nicht leer sein sollen, fallen lässt, dann sind zwei uneigentliche Schnitte zugelassen, nämlich ∞ mit der Obermenge $\varnothing$ und der Untermenge $\mathbb{Q}$ sowie $-\infty$ mit der Obermenge $\mathbb{Q}$ und der Untermenge $\varnothing$. (Streng genommen können wir sie mit unserer Definition von Schnitten als Zweiermengen nicht unterscheiden.) Nun ist $\mathbb{R} \cup \{\infty, -\infty\}$ eine geordnete Menge mit dem größten Element ∞ und dem kleinsten Element $-\infty$.

Folgerung 2.5. *Jedes Intervall in* $\mathbb{R}$ *hat eine der Formen*

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

Ist nämlich ein nicht leeres Intervall von unten beschränkt, so hat es nach Satz 2.11 eine untere Grenze a, und ist es von oben beschränkt, so hat es eine obere Grenze b. Dann muss man noch unterscheiden, ob es ein kleinstes und ob es ein größtes Element besitzt. Für den Körper $\mathbb{Q}$ wäre die entsprechende Aussage falsch, wie Beispiel 2.3 zeigt.

Satz 2.14. *Für jede natürliche Zahl* $n \ge 1$ *und jede reelle Zahl* $a \ge 0$ *existiert genau eine reelle Zahl* $c \ge 0$*, so dass* $c^n = a$.

Wir nennen c die n-te *Wurzel* aus a, abgekürzt $\sqrt[n]{a}$.

Beweis. Wir können annehmen, dass $x > 0$ ist. Wie in Beispiel 2.3 betrachten wir die Menge

$$M = \{x \in \mathbb{R}_+ \mid x^n < a\}.$$

Wegen der verallgemeinerten dritten binomischen Formel ist die Funktion $f(x) = x^n$ auf $\mathbb{R}_+$ streng monoton wachsend, also kann es höchstens ein Element c mit der verlangten Eigenschaft geben, und die Menge

$$A = \{y \in \mathbb{R}_+ \mid y^n > a\}$$

besteht aus oberen Schranken von M. Es gilt $\min\{\frac{1}{2}, a\} \in M$, und nach der Bernoulli-Ungleichung ist $1 + \frac{a}{n} \in A$, also ist M beschränkt. Ist $y \in A$, so gilt für $0 < y' < y$ nach der verallgemeinerten dritten binomischen Formel und den Monotoniegesetzen

$$y^n - y'^n = (y - y')(y^{n-1} + y^{n-2}y' + \ldots + y'^{n-1}) < ny^{n-1}(y - y').$$

Können wir y' so wählen, dass die rechte Seite zu $y^n - a$ wird, so folgt $y'^n > a$, also $y' \in A$. Dazu müssen wir

$$y' = y - \frac{y^n - a}{ny^{n-1}}$$

setzen, was in der Tat im Intervall $]0, y[$ ist. Somit hat A kein kleinstes Element. Analog zeigt man für $x \in M$ und $x < x' < 2x$

$$x'^n - x^n < (x' - x)x^{n-1}(2^{n-1} + \ldots + 2 + 1) = (2^n - 1)x^{n-1}(x' - x).$$

Wir können x' so wählen, dass die rechte Seite kleiner als $a - x^n$ und folglich $x'^n < a$ wird. Somit hat M kein größtes Element.

Nach Satz 2.11 existiert

$$c = \sup M.$$

Wäre $c \in M$, so wäre $c = \max M$, und wäre $c \in A$, könnnte es nicht die Kleinste unter den oberen Schranken sein. Da beides nicht möglich ist, folgt $c^n = a$. □

Definition 2.13. *Für alle $x \in \mathbb{R}_+$, $m \in \mathbb{Z}$ und $n \in \mathbb{N} \setminus \{0\}$ setzen wir*

$$x^{\frac{m}{n}} = \sqrt[n]{x^m}.$$

Die Korrektheit dieser Definition und die Potenzgesetze für rationale Exponenten wurden schon in der Vorlesung 0 diskutiert.

Folgerung 2.6. *Für beliebige positive reelle Zahlen $a_1, \ldots, a_n$ gilt*

$$\frac{n}{\frac{1}{a_1} + \ldots + \frac{1}{a_n}} \leq \sqrt[n]{a_1 \ldots a_n} \leq \frac{a_1 + \ldots + a_n}{n},$$

d. h. das harmonische Mittel *ist kleiner oder gleich dem* geometrischen Mittel *und dieses wiederum kleiner oder gleich dem* arithmetischen Mittel.

Bezeichnen wir nämlich das geometrische Mittel mit g, so gilt

$$\frac{a_1}{g} \cdots \frac{a_n}{g} = 1,$$

und aus Satz 2.8 folgt

$$\frac{a_1}{g} + \ldots + \frac{a_n}{g} \geq n.$$

Daraus folgt die rechte Ungleichung. Ersetzen wir jede Zahl a_i durch ihren Kehrwert, so folgt die linke Ungleichung.

Zum Abschluss dieses Abschnitts erinnern wir daran, dass durch

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

der *Absolutbetrag* einer reellen Zahl definiert ist. Die selbe Definition ist in einem beliebigen angeordneten Körper anwendbar. Es gilt

$$|x| < y \;\Leftrightarrow\; -y < x < y.$$

Eine Menge in einem angeordneten Körper ist darum genau dann beschränkt, wenn die Menge der Absolutbeträge ihrer Elemente von oben beschränkt ist.

2.5 Komplexe Zahlen

Vorlesung 6a
11.5.2022

In einem angeordneten Körper haben manche quadratische Gleichungen keine Lösung, z. B.

$$z^2 + 1 = 0. \tag{2}$$

Wir nehmen einmal an, dass es einen Körper gibt, der $\mathbb{R}$ als Teilkörper enthält und in dem die Gleichung (2) eine Lösung i besitzt. Dann gilt für alle reellen Zahlen x, y, u und v

$$(u + iv) + (x + iy) = (u + x) + i(v + y),$$
$$(u + iv) \cdot (x + iy) = (ux - vy) + i(uy + vx).$$

Die Teilmenge aller Elemente der Form $x + iy$ ist also abgeschlossen unter Addition und Multiplikation. Könnte man ein Element auf zwei Arten darstellen, z. B.

$$u + iv = x + iy,$$

so wäre $u - x = i(y - v)$, also

$$(u - x)^2 = -(v - y)^2$$

und somit $u = x$ und $v = y$. In dem Erweiterungskörper hat für gegebene u und v die Gleichung

$$(u + iv) + (x + iy) = 0$$

eine Lösung, nämlich $x = -u$, $y = -v$, und für $u + iv \neq 0$ hat auch die Gleichung

$$(u + iv) \cdot (x + iy) = 1$$

eine Lösung. Sie ist ja äquivalent zu dem Gleichungssystem

$$\begin{aligned} ux - vy &= 1, \\ uy + vx &= 0, \end{aligned}$$

und man findet leicht

$$x = \frac{u}{u^2 + v^2}, \qquad y = \frac{-v}{u^2 + v^2},$$

wobei $u^2 + v^2 > 0$, weil nach Voraussetzung u und v nicht beide gleich Null sind. Es folgt, dass die Teilmenge der Elemente der Form $x + iy$ bereits einen Körper bildet, in dem die obige quadratische Gleichung lösbar ist, und dass die Abbildung $(x, y) \mapsto x + iy$ von $\mathbb{R} \times \mathbb{R}$ auf diesen Teilkörper bijektiv ist.

All dies beruhte auf einer unbewiesenen Annahme, aber es zeigt, wie wir den fraglichen Körper konstruieren können.

Definition 2.14. *Es sei $\mathbb{C}$ die Menge aller geordneten Paare reeller Zahlen mit den Operationen*

$$\begin{aligned} (u, v) + (x, y) &= (u + x, v + y), \\ (u, v) \cdot (x, y) &= (ux - vy, uy + vx). \end{aligned}$$

Satz 2.15. *Die Menge $\mathbb{C}$ mit diesen Operationen ist ein Körper mit dem Nullelement $(0, 0)$ und dem Einselement $(1, 0)$.*

Beweis. Wir führen die Nachprüfung der Körperaxiome nur am Beispiel des Assoziativgesetzes der Multiplikation vor:

$$\begin{aligned} (s, t) \cdot \big((u, v) \cdot (x, y)\big) &= (s, t) \cdot (ux - vy, uy + vx) \\ &= \big(s(ux - vy) - t(uy + vx), s(uy + vx) + t(ux - vy)\big), \\ \big((s, t) \cdot (u, v)\big) \cdot (x, y) &= (su - tv, sv + tu) \cdot (x, y) \\ &= \big((su - tv)x - (sv + tu)y, (su - tv)y + (sv + tu)x\big). \end{aligned}$$

Die geordneten Paare auf der rechten Seite stimmen überein.

Die motivierenden Betrachtungen vor der Definition lassen sich in die Sprache geordneter Paare umformulieren und zeigen die Existenz der entgegengesetzten Zahl und des Kehrwertes, woraus die Existenz der Differenz und des Quotienten folgt. $\square$

Wir nennen die Elemente des Körpers $\mathbb{C}$ *komplexe Zahlen*. Die durch $x \mapsto (x, 0)$ gegebene Abbildung $\mathbb{R} \to \mathbb{C}$ ist ein injektiver Homomorphismus. Wir identifizieren darum $\mathbb{R}$ mit dem Wertebereich dieser Abbildung. Das Element $(0, 1)$ ist eine Lösung der Gleichung 2. Wir bezeichnen es mit i und nennen es *imaginäre Einheit*. Die eingangs angenommene Situation ist jetzt vollständig realisiert. Man veranschaulicht komplexe Zahlen z als Punkte mit Koordinaten $\operatorname{Re} z$ und $\operatorname{Im} z$ in einer Ebene (benannt nach Gauß oder Argand).

Jede komplexe Zahl lässt sich eindeutig in der Form $z = x + iy$ schreiben, und wir werden die Schreibweise als Paar nicht mehr benutzen. Man bezeichnet x als den *Realteil* und y als den *Imaginärteil* von z, abgekürzt $x = \operatorname{Re} z$ und $y = \operatorname{Im} z$, und mann nennt $\bar{z} = x - iy$ die zu z *konjugierte Zahl*.

Nach diesem Ausflug in die Algebra kehren wir wieder zur Analysis zurück. Der Begriff des Absolutbetrags aus dem vorigen Abschnitt ist nur für angeordnete Körper sinnvoll und muss verallgemeinert werden.

Definition 2.15. *Ein* Absolutbetrag *auf einem Körper K ist eine Abbildung von K in einen angeordneten Körper L, abgekürzt*[24] $x \mapsto |x|$, *so dass für alle Elemente x und y von K gilt*

$$\begin{aligned} |x| &\geq 0, & |x| = 0 &\Leftrightarrow x = 0, \\ |x \cdot y| &= |x| \cdot |y|, & |x + y| &\leq |x| + |y|. \end{aligned}$$

Der Absolutbetrag heißt archimedisch, *wenn es für jedes Element y von L_+ eine natürliche Zahl n gibt, so dass $|n \cdot 1_K| > y$.*

Manche Körper, beispielsweise $\mathbb{Q}$, besitzen auch nichtarchimedische Absolutbeträge.

Satz 2.16. *Die einzigen* Automorphismen *des Körpers $\mathbb{C}$ (d. h. Isomorphismen $\mathbb{C} \to \mathbb{C}$), die alle Elemente von $\mathbb{R}$ fest lassen, sind die identische Abbildung und die Konjugation. Durch*

$$|z| = \sqrt{z\bar{z}}$$

wird ein archimedischer Absolutbetrag komplexer Zahlen definiert.

Beweis. Gleichung (2) nimmt jetzt die Form

$$(z + i)(z - i) = 0$$

an, sie hat folglich nur die Lösungen i und $-i$. Ein Automorphismus f muss eine Lösung auf eine Lösung abbilden, also ist $f(i) = i$ oder $f(i) = -i$. Es

[24] Nicht zu verwechseln mit der Mächtigkeit!

gibt also nur die genannten Möglichkeiten, und man rechnet leicht nach, dass diese tatsächlich Automorphismen sind.

Wir prüfen die Eigenschaften eines Absolutbetrags nach. Für $z = x + iy$ ist $z\bar{z} = x^2 + y^2 \geq 0$, wobei Gleichheit nur für $x = y = 0$ eintritt. Also ist $|z| = \sqrt{x^2 + y^2}$ korrekt definiert und hat Eigenschaft (i) aus Definition 2.15. Für $y = 0$ erhalten wir den reellen Absolutbetrag $\sqrt{x^2} = |x|$. Eigenschaft (ii), nämlich

$$|zw| = |z||w|$$

folgt für alle komplexen Zahlen z und w aus den Potenzgesetzen und der Tatsache, dass die Konjugation ein Automorphismus ist.

Ist z rein imaginär, d. h. $\operatorname{Re} z = 0$, so ist z^2 reell und nicht positiv. Wegen $\operatorname{Re}(z\bar{w} - w\bar{z}) = \operatorname{Re} z\bar{w} - \operatorname{Re} \overline{z\bar{w}} = 0$ folgt also

$$(z\bar{w} - w\bar{z})^2 \leq 0.$$

Addieren wir $4z\bar{z}w\bar{w}$, so erhalten wir

$$(z\bar{w} + w\bar{z})^2 \leq 4|z|^2|w|^2.$$

Wegen $\operatorname{Im}(z\bar{w} + w\bar{z}) = \operatorname{Im} z\bar{w} + \operatorname{Im} \overline{z\bar{w}} = 0$ steht links das Quadrat einer reellen Zahl, und mit der Monotonie der Wurzelfunktion ist

$$|z\bar{w} + w\bar{z}| \leq 2|z||w|.$$

(Dies ist ein Spezialfall der Cauchy-Schwarz-Ungleichung.) Es folgt

$$(z + w)\overline{(z + w)} = z\bar{z} + z\bar{w} + w\bar{z} + w\bar{w} \leq |z|^2 + 2|z||w| + |w|^2.$$

Wegen der Monotonie der Wurzelfunktion folgt schließlich die Dreiecksungleichung

$$|z + w| \leq |z| + |w|.$$

Die archimedische Eigenschaft der Absolutbetrages überträgt sich von $\mathbb{R}$ auf $\mathbb{C}$. □

Wir beenden den Abschnitt mit zwei Bemerkungen. Da die Summe zweier Zahlen größer oder gleich jeder von beiden und andererseits kleiner oder gleich dem Doppelten der größeren von beiden ist, gilt

$$\max\{x^2, y^2\} \leq x^2 + y^2 \leq 2\max\{x^2, y^2\}.$$

Wenden wir das auf eine komplexe Zahl $z = x + iy$ an, so folgt mit der Monotonie der Wurzelfunktion

$$\max\{|\operatorname{Re} z|, |\operatorname{Im} z|\} \leq |z| \leq \sqrt{2}\max\{|\operatorname{Re} z|, |\operatorname{Im} z|\}. \tag{3}$$

Wenn bei einer Menge komplexer Zahlen sowohl die Menge ihrer Realteile als auch die Menge ihrer Imaginärteile beschränkt ist, dann folgt aus der rechten Ungleichung, dass auch die Menge ihrer Absolutbeträge beschränkt ist. Aus der linken Ungleichung folgt die Umkehrung.

Die zweite Bemerkung bezieht sich auf eine Methode, mit der man den Quotienten zweier komplexer Zahlen als Quotienten einer komplexen Zahl und einer reellen Zahl schreiben kann. Da der Bruchstrich meist als Synonym für das Divisionszeichen verwendet wird, spricht man von *Reellmachen des Nenners*. Motiviert durch die Formel für das Inverse einer komplexen Zahl erweitert man mit der konjugierten Zahl des Nenners:

$$\frac{z}{u+iv} = \frac{z(u-iv)}{u^2+v^2}.$$

Dabei sind natürlich u und v reelle Zahlen. Eine ähnliche Methode erlaubt das *Rationalmachen des Nenners*, wenn dort als Irrationalzahlen nur zwei Wurzeln aus rationalen Zahlen a und b vorkommen: Für rationale Zahlen u und v und eine beliebige reelle Zahl z gilt

$$\frac{z}{u\sqrt{a}+v\sqrt{b}} = \frac{z(u\sqrt{a}-v\sqrt{b})}{au^2-bv^2}.$$

3 Grenzwerte

3.1 Motivation

Vorlesung 6b 13.5.2022

Heron von Alexandria hat ein Verahren zur näherungsweisen Berechnung von Quadratwurzel aus einer Zahl A beschrieben, das bereits in Mesopotamien bekannt war. Ausgehend von einem Rechteck mit dem Flächeninhalt A (beispielsweise mit den Seitenlängen A und 1) sucht man ein flächengleiches Quadrat. Man kommt dem Quadrat näher, indem man zu einem flächengleichen Rechteck übergeht, dessen eine Seite gleich dem arithmetischen Mittel der Seiten des Ausgangsrechtecks ist. Dieses Verfahren wendet man immer wieder auf das entstehende Rechteck an.

Die Seiten des Ausgangsrechtecks seien x_0 und y_0. Wir können annehmen, dass $x_0 < y_0$ ist. Nach Voraussetzung ist $x_0 y_0 = A$. Die Folge der Seitenlängen der entstehenden Rechtecke ist dann rekursiv definiert durch

$$y_{n+1} = \frac{1}{2}(x_n + y_n), \qquad x_{n+1} = \frac{A}{y_{n+1}},$$

so dass für alle n gilt $x_n y_n = A$. Bereits in Beispiel 2.3 hatten wir gesehen, dass

$$x_n < x_{n+1} < \sqrt{A} < y_{n+1} < y_n.$$

Es folgt

$$y_{n+1} - x_{n+1} = \frac{1}{2}(y_n - x_n) + x_n - x_{n+1} < \frac{1}{2}(y_n - x_n),$$

und mit Hilfe der vollständigen Induktion zeigt man leicht

$$y_n - x_n \leq \frac{1}{2^n}(y_0 - x_0).$$

Um bei der Bestimmung von Abständen nicht ständig darauf achten zu müssen, die kleinere Zahl von der Größeren abzuziehen, verwendet man gern den Absolutbetrag. Es gilt nun

$$\left|x_n - \sqrt{A}\,\right| < \frac{1}{2^n}(y_0 - x_0), \qquad \left|y_n - \sqrt{A}\,\right| < \frac{1}{2^n}(y_0 - x_0).$$

Wenn man also einen Näherungswert für $\sqrt{A}$ mit einem Fehler kleiner als eine vorgegebene Zahl ε benötigt, kann man x_n oder y_n nehmen, vorausgesetzt,

$$\frac{1}{2^n}(y_0 - x_0) \leq \varepsilon.$$

Nach Satz 2.7 genügt dafür

$$n \geq \frac{y_0 - x_0}{\varepsilon} - 1.$$

Da wir von nun an viel mit Folgen zu tun haben werden, sollten wir ein paar traditionelle Bezeichnungen in Erinnerung bringen. Eine Folge heißt *arithmetische* bzw. *geometrische* bzw. *harmonische Folge*, wenn jedes Glied (mit Aunahme des Anfangsgliedes und, falls vorhanden, des Endgliedes) das arithmetische bzw. geometrische bzw. harmonische Mittel der beiden benachbarten Glieder ist.

3.2 Definition von Grenzwerten

Im Folgenden sei K ein Körper mit Absolutbetrag, der Werte im Körper L annimmt, also z. B. $K = L = \mathbb{Q}$ oder $K = L = \mathbb{R}$ oder $K = \mathbb{C}$, $L = \mathbb{R}$.

Definition 3.1. *Eine unendliche Folge von Elementen x_n in K* konvergiert *gegen das Element a, abgekürzt $x_n \to a$ $(n \to \infty)$, wenn es für jedes positive Element ε von L eine natürliche Zahl n_0 gibt, so dass für alle natürlichen Zahlen n mit der Eigenschaft $n \geq n_0$ gilt*

$$|x_n - a| < \varepsilon.$$

In Formeln ausgedrückt bedeutet $x_n \to a$ $(n \to \infty)$ also

$$\forall \varepsilon \in L_+ \, \exists n_0 \in \mathbb{N} \, \forall n \in \mathbb{N} \, (n \geq n_0 \Rightarrow |x_n - a| < \varepsilon).$$

An Stelle von n und ε kann jede andere Variable stehen, da es sich um gebundene Variablen handelt.

Beispiel. Die *harmonische Folge* konvergiert gegen Null, d. h. für alle $c \in K$ gilt

$$\frac{c}{n} \to 0 \;\; (n \to \infty).$$

Ist nämlich $\varepsilon > 0$, so gibt es eine natürliche Zahl n_0, so dass $n_0 > \frac{|c|}{\varepsilon}$, und dann gilt für alle n mit $n \geq n_0$

$$\left|\frac{c}{n}\right| = \frac{|c|}{n} \leq \frac{|c|}{n_0} < \varepsilon. \qquad \triangleleft$$

Satz 3.1. *Konvergiert eine Folge x_n sowohl gegen a als auch gegen b, so ist $a = b$.*

Beweis. Angenommen, $a \neq b$. Dann ist $\varepsilon = \frac{|a-b|}{2}$ positiv, also gibt es n_0 und n_1, so dass für $n \geq n_0$ gilt $|x_n - a| < \varepsilon$ und für $n \geq n_1$ gilt $|x_n - b| < \varepsilon$. Für $n \geq \max\{n_0, n_1\}$ gelten beide Ungleichungen, und mit der Dreiecksungleichung folgt

$$|a - b| = |(a - x_n) + (x_n - b)| \leq |a - x_n| + |x_n - b| < 2\varepsilon$$

(Widerspruch). Also war unsere Annahme falsch. □

Damit ist die folgende Definition gerechtfertigt.

Definition 3.2. *Konvergiert eine Folge x_n gegen a, so nennt man a den* Grenzwert *der Folge und bezeichnet ihn mit $\lim\limits_{n\to\infty} x_n$. Besitzt eine Folge einen Grenzwert in K, so heißt sie* konvergent *in K, andernfalls* divergent *in K.*

Die Folge in Abschnitt 3.1 ist also konvergent in $\mathbb{R}$, aber für $a = 2$ nicht in $\mathbb{Q}$.

Satz 3.2. *Jede konvergente Folge x_n ist beschränkt, d. h. es existiert ein Element c von L, so dass für alle natürlichen Zahlen n gilt $|x_n| \leq c$.*

Beweis. Ist $\lim\limits_{n\to\infty} x_n = a$, so gibt es ein n_0, so dass für $n \geq n_0$ gilt $|x_n - a| < 1$, also nach der Dreiecksungleichung

$$|x_n| = |(x_n - a) + a| \leq |x_n - a| + |a| < 1 + |a|,$$

und die Behauptung folgt mit

$$c = \max\{|x_0|, \ldots, |x_{n-1}|, 1 + |a|\},$$

falls die Folge mit dem nullten Glied beginnt, wobei das Maximum existiert (vgl. Aufgabe 16). $\square$

Beispiel. Wir betrachten die *arithmetische Folge*

$$x_n = x_0 + nd,$$

wobei $d \in K$.

Ist $d = 0$, so ist $x_n = x_0$ für alle n, also $x_n \to x_0 \ (n \to \infty)$.

Nach der Dreiecksungleichung gilt

$$n|d| = |x_n - x_0| \leq |x_n| + |x_0|.$$

Ist $d \neq 0$, dann kann es kein $c \in L$ geben, so dass $|x_n| \leq c$ für alle n. Also ist die Folge dann unbeschränkt und nach Satz 3.2 divergent. $\triangleleft$

Beispiel. Wir betrachten die *geometrische Folge*

$$x_n = x_0 q^n,$$

wobei x_0, $q \in K \setminus \{0\}$.

Ist $q = 1$, so ist $x_n = x_0$ für alle n, also $x_n \to x_0 \ (n \to \infty)$.

Ist $|q| > 1$, also $|q| = 1 + h$ mit $h > 0$, so ist $|x_n| \geq |x_0|(1 + nh)$ nach Bernoulli-Ungleichung, also ist x_n unbeschränkt und somit divergent. $\triangleleft$

Satz 3.3. *Ist $|q| < 1$ und k eine natürliche Zahl, so gilt*

$$n^k q^n \to 0 \ (n \to \infty).$$

Beweis. Es sei $l = k + 1$. Dann ist die Folge der Zahlen $n^l |q|^n$ nach Satz 2.9 beschränkt, sagen wir durch die Zahl c. Nun ist

$$|n^k q^n| = \frac{1}{n} |n^l q^n| \leq \frac{c}{n},$$

und die Behauptung folgt wie bei der harmonischen Folge. $\square$

Folgerung 3.1. *Für jede natürliche Zahl k gilt*

$$\sqrt[n]{n^k} \to 1 \ (n \to \infty).$$

Für jedes $\varepsilon > 0$ gilt ja $n^k(1+\varepsilon)^{-n} \to 0$ $(n \to \infty)$ nach Satz 3.3, also existiert ein n_0, so dass für $n \geq n_0$ gilt $n^k(1+\varepsilon)^{-n} < 1$. Letzteres ist wegen der Monotonie der Wurzelfunktion äquivalent zu $\sqrt[n]{n^k} < 1+\varepsilon$, und aus dem selben Grund gilt $\sqrt[n]{n^k} \geq 1$, so dass $|\sqrt[n]{n^k} - 1| < \varepsilon$.

Beispiel. Zur Untersuchung der Folge $x_n = \frac{b^n}{n!}$ für beliebiges $b \in K$ wählen wir eine natürliche Zahl m, so dass $m \geq 2|b|$. Dann gilt für $n \geq m$

$$|x_n| = \prod_{i=1}^{n} \frac{|b|}{i} = \prod_{i=1}^{m} \frac{|b|}{i} \prod_{i=m+1}^{n} \frac{|b|}{i} \leq \frac{|b|^m}{m!} \left(\frac{1}{2}\right)^{n-m} = \frac{(2|b|)^m}{m!} \left(\frac{1}{2}\right)^{n},$$

und es folgt leicht

$$\frac{b^n}{n!} \to 0 \ \ (n \to \infty)$$

mit Hilfe von Satz 3.3 für $q = \frac{1}{2}$ und $k = 0$. ◁

Bemerkung. Für eine Folge komplexer Zahlen z_n und eine komplexe Zahl c gilt genau dann Vorlesung 7a 18.05.2022

$$z_n \to c \ \ (n \to \infty),$$

wenn

$$\operatorname{Re} z_n \to \operatorname{Re} c \quad \wedge \quad \operatorname{Im} z_n \to \operatorname{Im} c \ \ (n \to \infty).$$

Der Kürze halbe schreiben wir $\operatorname{Re} z_n = x_n$, $\operatorname{Im} z_n = y_n$, $\operatorname{Re} c = a$, $\operatorname{Im} c = b$. In der Tat, gilt $x_n \to a \ \ (n \to \infty)$, so gibt es für jedes positive ε_0 eine natürliche Zahl n_0, so dass für $n \geq n_0$ gilt $|x_n - a| < \varepsilon_0$, und gilt $y_n \to b \ \ (n \to \infty)$, so gibt es für jedes positive ε_1 ein n_1, so dass für $n \geq n_1$ gilt $|y_n - b| < \varepsilon_1$. Für $n \geq \max\{n_0, n_1\}$ gilt dann nach der rechten Ungleichung vom Ende des vorigen Abschnitts

$$|z_n - c| = |(x_n - a) + i(y_n - b)| \leq \sqrt{2} \max\{|x_n - a|, |y_n - b|\} < \sqrt{2} \max\{\varepsilon_0, \varepsilon_1\}.$$

Ist nun eine positive Zahl ε gegeben, so wenden wir dies mit $\varepsilon_1 = \varepsilon_2 = \frac{\varepsilon}{\sqrt{2}}$ an, wonach die Konvergenz der Folge z_n klar ist. Die Umkehrung zeigt man ähnlich.

3.3 Rechenregeln für Grenzwerte

Aus konvergenten Folgen in einem Körper K mit Absolutbetrag kann man mit Hilfe der Rechenoperationen weitere konvergente Folgen bilden.

Satz 3.4. *Wenn $x_n \to a$ und $y_n \to b$ für $n \to \infty$, dann $x_n + y_n \to a + b$ und $x_n \cdot y_n \to a \cdot b$ für $n \to \infty$.*

Beweis. Laut Dreiecksungleichung gilt für alle n

$$|(x_n + y_n) - (a + b)| \leq |x_n - a| + |y_n - b|.$$

Wegen der Konvergenz von x_n und y_n gibt es für jedes $\varepsilon > 0$ natürliche Zahlen n_0 und n_1, so dass für $n \geq n_0$ gilt $|x_n - a| < \frac{\varepsilon}{2}$, während für $n \geq n_1$ gilt $|y_n - b| < \frac{\varepsilon}{2}$. Für $n \geq \max\{n_0, n_1\}$ folgt

$$|(x_n + y_n) - (a + b)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

und die Behauptung über die Summe ist bewiesen.

Für das Produkt bemerken wir, dass nach den Eigenschaften des Absolutbetrags gilt

$$\begin{aligned}|x_n y_n - ab| &= |x_n(y_n - b) + (x_n - a)b| \\ &\leq |x_n(y_n - b)| + |(x_n - a)b| = |x_n||y_n - b| + |x_n - a||b|.\end{aligned}$$

Nach Satz 3.2 gibt es ein $c \in L_+$, so dass

$$|x_n| \leq c \quad \text{für alle } n, \qquad |b| \leq c,$$

also

$$|x_n y_n - ab| \leq c\big(|y_n - b| + |x_n - a|\big).$$

Für jedes $\varepsilon > 0$ gibt es n_0' und n_1', so dass für $n \geq n_0'$ gilt $|x_n - a| < \frac{\varepsilon}{2c}$, während für $n \geq n_1'$ gilt $y_n - b| < \frac{\varepsilon}{2c}$. Also folgt für $n \geq \max\{n_0', n_1'\}$

$$|x_n y_n - ab| < \varepsilon.$$

□

Folgerung 3.2. *Wenn $x_n \to a$ und $y_n \to b$ für $n \to \infty$, dann $x_n - y_n \to a - b$ für $n \to \infty$.*

Dies ergibt sich, indem wir $x_n - y_n = x_n + (-1) \cdot y_n$ schreiben.

Satz 3.5. *Wenn $x_n \to a$ für $n \to \infty$ und $a \neq 0$, dann gibt es ein n_1, so dass $x_n \neq 0$ für $n \geq n_1$, und es gilt $\frac{1}{x_n} \to \frac{1}{a}$ für $n \to \infty$.*

Beweis. Wegen der Konvergenz von x_n gibt es ein n_1, so dass für $n \geq n_1$ gilt $|x_n - a| < \frac{|a|}{2}$. Mit der Dreiecksungleichung folgt für diese n

$$|a| = |(a - x_n) + x_n| \leq |a - x_n| + |x_n| < \frac{|a|}{2} + |x_n|,$$

also $|x_n| > \frac{|a|}{2}$, insbesondere $x_n \neq 0$, und

$$\left|\frac{1}{x_n} - \frac{1}{a}\right| = \left|\frac{a - x_n}{x_n a}\right| = \frac{|a - x_n|}{|x_n||a|} < \frac{2}{|a|^2}|x_n - a|.$$

Für jedes $\varepsilon > 0$ gibt es ein n_0, so dass für $n \geq n_0$ gilt $|x_n - a| < \frac{|a|^2}{2}\varepsilon$, und für $n \geq \max\{n_0, n_1\}$ folgt

$$\left|\frac{1}{x_n} - \frac{1}{a}\right| < \varepsilon. \qquad \square$$

Folgerung 3.3. *Wenn $x_n \to a$ und $y_n \to b$ für $n \to \infty$, wobei $b \neq 0$, dann gibt es ein n_1, so dass $y_n \neq 0$ für $n \geq n_1$, und es gilt $\frac{x_n}{y_n} \to \frac{a}{b}$ für $n \to \infty$.*

Dies ergibt sich, indem wir $\frac{x_n}{y_n} = x_n \cdot \frac{1}{y_n}$ schreiben.
Die eben gezeigten Rechenregeln kann man auch so formulieren: Es gilt

$$\lim_{n\to\infty}(x_n \pm y_n) = \lim_{n\to\infty} x_n \pm \lim_{n\to\infty} y_n,$$
$$\lim_{n\to\infty}(x_n \cdot y_n) = \lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} y_n,$$
$$\lim_{n\to\infty}\frac{x_n}{y_n} = \frac{\lim_{n\to\infty} x_n}{\lim_{n\to\infty} y_n},$$

vorausgesetzt, die Ausdrücke auf der rechten Seite sind definiert.
Warnung. Aus $\lim_{n\to\infty} x_n = a$ folgt nicht $\lim_{n\to\infty}(x_n y_n) = \lim_{n\to\infty}(a y_n)$, wie das Beispiel $x_n = \frac{1}{n}$, $y_n = n$ zeigt.
Beispiel. Es sei

$$x_n = \frac{(3n+5)^2}{2n^2+1}.$$

Die Sätze 3.4 und 3.5 sind nicht unmittelbar anwendbar, da Zähler und Nenner jeweils divergent sind. Man kürze durch geeignete Potenz von n, so dass der Nenner einen von Null verschiedenen Grenzwert hat:

$$x_n = \frac{\left(\frac{3n+5}{n}\right)^2}{\frac{2n^2+1}{n^2}} = \frac{\left(3+\frac{5}{n}\right)^2}{2+\frac{1}{n^2}} \to \frac{(3+0)^2}{2+0^2} = \frac{9}{2} \qquad (n \to \infty). \qquad \triangleleft$$

3.4 Der Satz von Bolzano-Weierstraß

Üblicherweise sagt man, dass man aus einer Folge von Gliedern x_n eine *Teilfolge* erhält, indem man beliebig viele Folgeglieder streicht, so dass unendlich viele verbleiben. Richtiger wäre es zu sagen, dass man Elemente aus der Indexmenge streicht. Satz 2.2(iii) zeigt, dass wir die verbleibenden Indizes in aufsteigender Reihenfolge nummerieren können. Es gibt also eine streng monoton wachsende Folge n_k von natürlichen Zahlen, und die Teilfolge besteht aus den Gliedern x_{n_k}, die durch k nummeriert werden. Genaugenommen ist dies die Verkettung der Abbildungen[25] $n \mapsto x_n$ und $k \mapsto n_k$. Dieses Umnummerieren wäre unnötig, wenn man als Indexmenge eine beliebige Teilmenge von $\mathbb{N}$ zulassen würde, aber das ist in der elementaren Analysis unüblich. Offensichlich ist eine Teilfolge einer Teilfolge auch eine Teilfolge der Ausgangsfolge.

Es ist auch klar, dass aus $x_n \to a$ für $n \to \infty$ folgt, dass $x_{n_k} \to a$ für $k \to \infty$.

Beispiel. Die durch $q = \frac{1}{2}\big(1 + i\sqrt{3}\big)$ und $z_n = q^n$ definierte Folge ist divergent (siehe Aufgabe 29), aber sie hat die konvergenten Teilfolgen $(z_{6k})_{k\in\mathbb{N}}$, $(z_{6k+1})_{k\in\mathbb{N}}$, ..., $(z_{6k+5})_{k\in\mathbb{N}}$, denn $q^6 = 1$. $\triangleleft$

Satz 3.6. *Jede beschränkte Folge in $\mathbb{R}$ oder $\mathbb{C}$ hat eine konvergente Teilfolge.*

Beweis. Wir betrachten zunächst eine Folge von reellen Zahlen x_n mit einer unteren Schranke a_0 und einer oberen Schranke b_0. Dann ist x_n eine Folge im Intervall $[a_0, b_0]$. Wir definieren nun rekursiv eine Folge von Intervallen $[a_m, b_m]$, von denen jedes unendlich viele Folgeglieder x_n enthält, also genauer gesagt, so dass es für jedes m unendlich viele n mit der Eigenschaft $x_n \in [a_m, b_m]$ gibt.

Angenommen, das m-te Intervall ist bereits definiert und enthält unendlich viele x_n. Es sei c_m das arithmetische Mittel von a_m und b_m. Gäbe es nur endlich viele Indizes n mit $x_n \in [a_m, c_m]$ und auch nur endlich viele Indizes n mit $x_n \in [c_m, b_m]$, so gäbe es auch nur enlich viele Indizes n mit $x_n \in [a_m, b_m]$. Da dies nicht der Fall ist, muss eines der Intervalle $[a_m, c_m]$ oder $[c_m, b_m]$ unendlich viele Folgeglieder entalten. Wir wählen ein solches Intervall aus und bezeichnen es mit $[a_{m+1}, b_{m+1}]$.

Man beachte, dass seine Länge $b_{m+1} - a_{m+1}$ genau halb so groß ist wie die Länge des vorigen Intervalls $[a_m, b_m]$. Hieraus folgt durch vollständige Induktion

$$b_m - a_m = \frac{1}{2^m}(b_0 - a_0).$$

Wir setzen nun $n_0 = 1$ und definieren rekursiv eine streng monoton wachsende Folge natürlicher Zahlen n_m, so dass $x_{n_m} \in [a_m, b_m]$ ist. Es gibt ja nach Konstruktion unendlich viele Indizes n mit der Eigenschaft $x_n \in [a_{m+1}, b_{m+1}]$, und ist n_m bereits definiert, so setzen wir n_{m+1} gleich dem Kleinsten unter ihnen, der größer als n_m ist. Wegen $[a_{m+1}, b_{m+1}] \subseteq [a_m, b_m]$ gilt sogar

$$x_{n_k} \in [a_m, b_m] \qquad \text{für alle } k \geq m.$$

Da jedes b_m eine obere Schranke der Menge $\{a_m \mid m \in \mathbb{N}\}$ ist, hat diese nach Satz 2.11 eine kleinste obere Schranke c, und für alle m gilt

$$a_m \leq c \leq b_m.$$

[25] Das erste n bezeichnet eine Variable, das zweite n eine Folge.

Ist nun eine positive Zahl ε gegeben, so gibt es wegen obiger Formel für die Intervalllänge und Satz 2.7 eine natürliche Zahl m, so dass $b_m - a_m < \varepsilon$. Für $k \geq m$ gehören dann sowohl x_{n_k} als auch c zum Intervall $[a_m, b_m]$, und es folgt

$$|x_{n_k} - c| < \varepsilon.$$

Wir haben bewiesen, dass $x_{n_k} \to c \ (k \to \infty)$.

Ist eine beschränkte Folge komplexer Zahlen z_n gegeben, so sind die Folgen der reellen Zahlen $\operatorname{Re} z_n$ und $\operatorname{Im} z_n$ ebenfalls beschränkt. Es gibt nach dem Bewiesenen eine Teilfolge, für die die Folge der Realteile konvergiert, und davon wiederum eine Teilfolge, bei der auch die Folge der Imaginärteile konvergiert. Nun konvergiert auch die zugehörige Teilfolge komplexer Zahlen. □

Eine Folge von Intervallen, bei der jedes folgende Intervall in dem vorangehenden enthalten ist, und wo die Länge der Intervalle jede vorgegebene positive Zahl unterschreitet, nennt man übrigens eine *Intervallschachtelung*. Auch das Heronverfahren erzeugt eine Intervallschachtelung $[x_n, y_n]$. Man hätte die Beweismethode des reellen Falls auch direkt auf den komplexen Fall übertragen können, indem man die Folge in ein genügend großes Quadrat einschließt und dieses in jedem Schritt in vier Teilquadrate unterteilt.

3.5 Grenzwerte und Ungleichungen

Vorlesung 7b
20.05.2022

In einem angeordneten Körper gelten für Grenzwerte neben den Regeln aus dem Abschnitt 3.3 weitere Gesetzmäßigkeiten.

Satz 3.7. *Ist $x_n \to a$ und $y_n \to b$ für $n \to \infty$ und gilt $x_n \leq y_n$ für alle (genügend großen) n, so gilt $a \leq b$.*

Warnung: $\leq$ ist nicht durch $<$ ersetzbar!

Beweis. Für jedes $\varepsilon > 0$ gibt es ein n_0, so dass für $n \geq n_0$ gilt $|x_n - a| < \varepsilon$ und $|y_n - b| < \varepsilon$, also

$$a - \varepsilon < x_n \leq y_n < b + \varepsilon, \qquad a - b < 2\varepsilon.$$

Wäre $a > b$, so ergäbe sich mit $\varepsilon = \frac{a-b}{2}$ ein Widerspruch. □

Satz 3.8 (Einschließungskriterium). *Ist $x_n \to a$ und $z_n \to a$ für $n \to \infty$ und gilt $x_n \leq y_n \leq z_n$ für alle (genügend großen) n, so ist $y_n \to a$ für $n \to \infty$.*

Beweis. Für jedes $\varepsilon > 0$ gibt es ein n_0, so dass für $n \geq n_0$ gilt $|x_n - a| < \varepsilon$ und $|z_n - a| < \varepsilon$, also

$$a - \varepsilon < x_n \leq y_n \leq z_n < a + \varepsilon, \qquad |y_n - a| < \varepsilon.$$ □

Beispiel. Ist $c \geq 1$, so gilt $1 \leq \sqrt[n]{c} \leq \sqrt[n]{n}$ für genügend große n, und mit Folgerung 3.1 ergibt sich

$$\sqrt[n]{c} \to 1 \ \ (n \to \infty).$$

Mit Satz 3.5 und einem Potenzgesetz folgt das auch für $0 < c \leq 1$. $\triangleleft$

Beispiel. Für $p \geq q \geq 0$ gilt

$$\sqrt[n]{p^n} \leq \sqrt[n]{p^n + q^n} \leq \sqrt[n]{p^n + p^n},$$

und es folgt

$$\sqrt[n]{p^n + q^n} \to p \ \ (n \to \infty)$$

mit Hilfe des vorigen Beispiels für $c = 2$. $\triangleleft$

Beispiel. Es gilt

$$0 < \sqrt{1 + \frac{1}{n}} - 1 = \frac{(\sqrt{1 + \frac{1}{n}} - 1)(\sqrt{1 + \frac{1}{n}} + 1)}{\sqrt{1 + \frac{1}{n}} + 1} = \frac{\frac{1}{n}}{\sqrt{1 + \frac{1}{n}} + 1} < \frac{1}{2n},$$

also

$$\sqrt{1 + \frac{1}{n}} - 1 \to 0, \qquad \sqrt{1 + \frac{1}{n}} \to 1 \ \ (n \to \infty)$$

nach Satz 3.8. $\triangleleft$

Satz 3.9 (Monotoniekriterium). *Jede beschränkte monotone Folge reeller Zahlen ist konvergent in* $\mathbb{R}$.

Dies wäre im Körper $\mathbb{Q}$ falsch, wie man am Heronverfahren für $a = 2$ sieht.

Beweis. Wir betrachten o. B. d. A. den Fall monotonen Wachstums. Wegen der Beschränktheit existiert $a = \sup\{x_n \mid n \in \mathbb{N}\}$ nach Satz 2.11. Ist $\varepsilon > 0$, so kann $a - \varepsilon$ keine obere Schranke sein, also gibt es ein n_0, so dass $x_{n_0} > a - \varepsilon$. Aus der Monotonie folgt $x_n > a - \varepsilon$ für $n \geq n_0$. Da a obere Schranke ist, gilt $x_n \leq a$, also $|x_n - a| < \varepsilon$. $\square$

Als Anwendung untersuchen wir das Wachstum eines Anfangskapitals a bei einem nominalen Zinssatz x. Nach einem Jahr beträgt das Guthaben (unter Vernachlässigung von Bearbeitungsgebühren)

$(1 + x)\,a$	bei jährlicher Verzinsung
$\left(1 + \frac{x}{12}\right)^{12} a$	bei (naiver) monatlicher Verzinsung
$\left(1 + \frac{x}{365}\right)^{365} a$	bei (naiver) täglicher Verzinsung
$\ldots$	

Steigt die Auszahlung bei häufigerer Verzinsung? Kann sie ins Unermessliche steigen?

Satz 3.10. *Für alle reellen Zahlen x existiert der Grenzwert*

$$\exp x = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n .$$

Für alle x, $y \in \mathbb{R}$ gilt

(i) $\exp x > 0$, $\exp x \geq 1 + x$,

(ii) *wenn* $x \leq y$, *dann* $\exp x \leq \exp y$,

(iii) $\exp(x + y) = \exp x \cdot \exp y$.

Beweis. Zunächst halten wir x fest und setzen für $n > 0$

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

Behauptung 1: Für $n > -x$ ist die Folge a_n monoton wachsend.
Dazu genügt es wegen $a_n > 0$ zu zeigen, dass $a_{n+1}/a_n \geq 1$. Es gilt

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

Für $n > -x$ ist $\frac{x}{x+n} < 1$, also nach Bernoulli

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

Behauptung 2: Ist $x \in \mathbb{Z}$, so ist für $n > -x$

$$a_n \leq b_n = \left(1 + \frac{x}{n}\right)^{n+x} ,$$

und die Folge b_n ist monoton fallend.

Es folgt nämlich aus Monotonie der Potenzfunktion (Fallunterscheidung $x > 0$, $x = 0$, $x < 0$)

$$\frac{b_n}{a_n} = \left(1 + \frac{x}{n}\right)^x \geq 1^x = 1.$$

Zum Beweis von $b_n/b_{n+1} \geq 1$ benutzen wir die obige Rechnung und die Bernoulli-Ungleichung:

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

Behauptung 3: a_n ist beschränkt, also nach Satz 3.9 konvergent. (Wegen $a_1 = 1 + x$ folgt dann (i) aus Behauptung 1 und Satz 3.7.)

Für $x \in \mathbb{Z}$ ist jedes b_n mit $n > -x$ obere Schranke. Für beliebige $x, y \in \mathbb{R}$ mit $x \leq y$ ist

$$\left(1 + \frac{x}{n}\right)^n \leq \left(1 + \frac{y}{n}\right)^n,$$

sobald $n > \max\{-x, -y\}$. Wählen wir $y \in \mathbb{Z}$, so folgt Beschränktheit für x. Mit Satz 3.7 folgt (ii).

Behauptung 4: Es gilt

$$\lim_{n\to\infty} \left(1 + \frac{x}{n^2}\right)^n = 1.$$

Ist nämlich $0 < c < \exp x < C$, so gilt für große m

$$c \leq \left(1 + \frac{x}{m}\right)^m \leq C.$$

Ersetzen wir m durch n^2 und wenden die (monotone) Wurzelfunktion an, so ergibt sich

$$\sqrt[n]{c} \leq \left(1 + \frac{x}{n^2}\right)^n \leq \sqrt[n]{C},$$

und die Behauptung folgt aus dem Einschließungskriterium.

Zum Beweis von (iii) beachte man, dass

$$\frac{(1 + \frac{x}{n})(1 + \frac{y}{n})}{1 + \frac{x+y}{n}} = 1 + \frac{1}{1 + \frac{x+y}{n}} \cdot \frac{xy}{n^2}.$$

Für alle $n \geq 2|x + y|$ gilt $1 + \frac{x+y}{n} \geq \frac{1}{2}$, also

$$1 - \frac{2|xy|}{n^2} \leq \frac{(1 + \frac{x}{n})(1 + \frac{y}{n})}{1 + \frac{x+y}{n}} \leq 1 + \frac{2|xy|}{n^2}.$$

Für große n sind alle Terme positiv, und die Ungleichungen gelten auch für die $n$ten Potenzen. Mit dem Einschließungskriterium folgt (iii) aus Behauptung 4. □

Folgerung 3.4. *Für $x < 1$ gilt*

$$\exp x \leq \frac{1}{1 - x}.$$

Dies folgt aus (iii) für $y = -x$ und aus (i).

Folgerung 3.5. *Für alle reellen Zahlen x und alle rationalen Zahlen r gilt*

$$\exp rx = (\exp x)^r.$$

Dies folgt aus (iii) für $r = -1$ und (durch vollständige Induktion) für alle $r \in \mathbb{N}$, also gilt es für $r \in \mathbb{Z}$. Für beliebige $r = \frac{m}{n}$ folgt es aus

$$\left(\exp \frac{m}{n} x\right)^n = \exp mx = (\exp x)^m.$$

Definition 3.3. *Die Funktion* exp *wird* Exponentialfunktion *genannt, die Zahl $e = \exp 1$ heißt* Eulersche Zahl.

Für $r \in \mathbb{Q}$ ist also $\exp r = e^r$. Man schreibt $\exp x$ auch im Fall von irrationalen Zahlen oft als e^x.

Folgerung 3.6. *Die Funktion* exp *ist streng monoton wachsend.*

Sie ist monoton wachsend nach (ii), und gäbe es reelle Zahlen $x < y$, so dass $\exp x = \exp y$, so wäre nach (iii) dann $\exp(y - x) = 1$, aber dies widerspräche Eigenschaft (i).

3.6 Uneigentliche Grenzwerte

Vorlesung 8a
25.05.2022

Bereits in Abschnitt 2.4 hatten wir die uneigentlichen Zahlen ∞ und $-\infty$ eingeführt. Das hatte den Vorteil, dass in der erweiterten geordneten Menge $\mathbb{R} \cup \{\infty, -\infty\}$ jede Teilmenge eine obere und eine untere Grenze hat. Auch im Zusammenhang mit Grenzwerten sind diese Elemente von Nutzen.

Definition 3.4. *Wir sagen, dass eine Folge x_n in $\mathbb{R} \cup \{\infty, -\infty\}$* gegen ∞ divergiert, *abgekürzt $x_n \to \infty \ (n \to \infty)$, wenn es für jede reelle Zahl c eine natürliche Zahl n_0 gibt, so dass für alle Indizes n mit der Eigenschaft $n \geq n_0$ gilt $x_n > c$.*

Analog definiert man, wann $x_n \to -\infty \ (n \to \infty)$.

In diesen Fällen nennt man ∞ bzw. $-\infty$ den uneigentlichen Grenzwert *der Folge und bezeichnet ihn ebenfalls mit $\lim\limits_{n\to\infty} x_n$.*

Satz 3.11. *Gegeben sei eine Folge x_n in $\mathbb{R}$.*

(i) *Wenn $x_n \to \infty \ (n \to \infty)$ und ein $a \in \mathbb{R}$ existiert, so dass $y_n \geq a$ für alle n, dann ist $x_n + y_n \to \infty \ (n \to \infty)$.*

(ii) *Wenn $x_n \to \infty \ (n \to \infty)$ und ein $a \in \mathbb{R}_+$ existiert, so dass $y_n \geq a$ für alle n, dann ist $x_n \cdot y_n \to \infty \ (n \to \infty)$.*

(iii) *Es gilt genau dann $x_n \to \infty \ (n \to \infty)$, wenn $-x_n \to -\infty \ (n \to \infty)$.*

(iv) *Es gilt genau dann $x_n \to \infty \ (n \to \infty)$, wenn $x_n > 0$ für genügend große n und $\frac{1}{x_n} \to 0 \ (n \to \infty)$.*

Beweis. Wir führen ihn nur für (iv) vor. Angenommen, $x_n \to \infty \;\; (n \to \infty)$. Dann gibt es laut Definition ein n_0, so dass für $n \geq n_0$ gilt $x_n > 0$, und für jedes $\varepsilon > 0$ gibt es ein n_1, so dass für $n \geq n_1$ gilt $x_n > \frac{1}{\varepsilon}$. Für $n \geq \max\{n_0, n_1\}$ gilt dann $0 < \frac{1}{x_n} < \varepsilon$.

Ist umgekehrt $x_n > 0$ für $n \geq n_0$ und $\frac{1}{x_n} \to 0 \;\; (n \to \infty)$, so gibt es für jedes $c > 0$ ein n_1, so dass für $n \geq n_1$ gilt $\left|\frac{1}{x_n}\right| < \frac{1}{c}$. Für $n \geq \max\{n_0, n_1\}$ gilt dann $x_n > c$. □

Beispiel. Für $q > 1$ und $k \in \mathbb{N}$ gilt

$$\frac{q^n}{n^k} \to \infty \;\; (n \to \infty).$$

Die Folge der Kehrwerte $n^k \left(\frac{1}{q}\right)^n$ konvergiert nämlich nach Satz 3.3 gegen Null. ◁

Beispiel. Wenn $x_n \to \infty$ und $y_n \to -\infty \;\; (n \to \infty)$, so folgt über die Konvergenz von $x_n + y_n$ gar nichts, wie die Beispiele $x_n = n$ und $y_n = -n$ oder $y_n = -n^2$ oder $y_n = -n + (-1)^n$ zeigen. Wenn $x_n \to \infty$ und $y_n \to 0 \;\; (n \to \infty)$, so folgt über die Konvergenz von $x_n \cdot y_n$ gar nichts, wie die Beispiele $x_n = n$ und $y_n = \frac{1}{n}$ oder $y_n = \frac{1}{\sqrt{n}}$ oder $y_n = \frac{(-1)^n}{n}$ zeigen. ◁

Folgerung 3.7. *Die Aussagen von Satz 3.4 gelten sinngemäß für uneigentliche Grenzwerte außer für die Addition zweier Folgen mit entgegengesetzten uneigentlichen Grenzwerten und das Produkt zweier Folgen, von denen eine gegen Null konvergiert und die andere einen uneigentlichen Grenzwert hat, wenn wir definieren*

$$\begin{aligned} \infty + \infty = \infty, &\qquad (-\infty) + (-\infty) = -\infty, \\ \infty \cdot \infty = (-\infty) \cdot (-\infty) = \infty, &\qquad \infty \cdot (-\infty) = (-\infty) \cdot \infty = -\infty \end{aligned}$$

sowie für alle a, $b \in \mathbb{R}$

$$a + \infty = \infty + b = \infty, \qquad a + (-\infty) = (-\infty) + b = -\infty,$$

für alle $a > 0$, $b > 0$

$$a \cdot \infty = \infty \cdot b = \infty, \qquad a \cdot (-\infty) = (-\infty) \cdot b = -\infty,$$

und schließlich für alle $a < 0$, $b < 0$

$$a \cdot \infty = \infty \cdot b = -\infty, \qquad a \cdot (-\infty) = (-\infty) \cdot b = \infty,$$

Man beachte, dass die Aussagen von Satz 3.11 universeller anwendbar sind als die entsprechenden Aussagen der Folgerung.

3.7 Limes superior und limes inferior

Aus Satz 3.9 ergibt sich, dass jede monotone Folge in $\mathbb{R} \cup \{\infty, -\infty\}$ einen eigentlichen oder uneigentlichen Grenzwert hat. Für eine beliebige Folge x_n betrachten wir die monoton fallende Folge

$$x_n^+ = \sup\{x_n, x_{n+1}, \dots\} = \sup\{x_k \mid k \geq n\}$$

und die monoton wachsende Folge

$$x_n^- = \inf\{x_n, x_{n+1}, \dots\} = \inf\{x_k \mid k \geq n\}.$$

Definition 3.5. *Wir definieren*[26]

$$\varlimsup_{n\to\infty} x_n = \lim_{n\to\infty} x_n^+, \qquad \varliminf_{n\to\infty} x_n = \lim_{n\to\infty} x_n^-,$$

und nennen dies den limes superior *bzw. den* limes inferior *der Folge x_n.*

Beispiel. Für die Folge $x_n = (-1)^n + \frac{1}{n}$ gilt $x_{2k}^+ = x_{2k-1}^+ = 1 + \frac{1}{2k}$ und $x_n^- = -1$, also $\varlimsup_{n\to\infty} x_n = 1$ und $\varliminf_{n\to\infty} x_n = -1$. ◁

Satz 3.12. *Eine Folge x_n ist genau dann konvergent, wenn sie beschränkt ist und*

$$\varlimsup_{n\to\infty} x_n = \varliminf_{n\to\infty} x_n.$$

In diesem Fall gilt $\lim_{n\to\infty} x_n = \varlimsup_{n\to\infty} x_n = \varliminf_{n\to\infty} x_n$.

Beweis. Angenommen, $\varlimsup_{n\to\infty} x_n = \varliminf_{n\to\infty} x_n = a \in \mathbb{R}$. Zu gegebenem $\varepsilon > 0$ gibt es dann ein n_0, so dass für $n \geq n_0$ gilt

$$a - \varepsilon < x_n^-, \qquad x_n^+ < a + \varepsilon.$$

Nach Definition von $x_n^\pm$ ist $x_n^- \leq x_n \leq x_n^+$, und die Konvergenz folgt.

Umgekehrt sei $\lim_{n\to\infty} x_n = a$. Zu gegebenem $\varepsilon > 0$ gibt es dann ein n_0, so dass für $n \geq n_0$ gilt

$$a - \varepsilon < x_n < a + \varepsilon,$$

also nach Definition der Grenzen

$$a - \varepsilon \leq x_n^-, \qquad x_n^+ \leq a + \varepsilon.$$

[26] Statt $\varlimsup$ und $\varliminf$ sind auch die Abkürzungen $\limsup$ und $\liminf$ üblich.

Mit Satz 3.7 folgt

$$a - \varepsilon \leq \varliminf_{n\to\infty} x_n, \qquad \varlimsup_{n\to\infty} x_n \leq a + \varepsilon$$

und, da $\varepsilon > 0$ beliebig war,

$$\varlimsup_{n\to\infty} x_n \leq a \leq \varliminf_{n\to\infty} x_n.$$

Andererseits folgt aus $x_n^- \leq x_n^+$, dass

$$\varliminf_{n\to\infty} x_n \leq \varlimsup_{n\to\infty} x_n$$

und somit Gleichheit gilt. □

3.8 Häufungspunkte

Vorlesung 8b 27.05.2022

Auf einem Körper K sei ein Absolutbetrag mit Werten in einem angeordneten Körper L gegeben.

Definition 3.6. *Ein Element a von K heißt* Häufungspunkt *einer Folge von Elementen x_n von K, wenn es für jedes $\varepsilon \in L_+$ unendlich viele n gibt, so dass $|x_n - a| < \varepsilon$.*

Der Grenzwert einer konvergenten Folge ist offenbar ein Häufungspunkt.

Beispiel. Da die Mengen $\mathbb{N}$ und $\mathbb{Q}$ gleichmächtig sind, gibt es eine Folge in $\mathbb{R}$, in der jede rationale Zahl genau einmal vorkommt. Offenbar ist jede reelle Zahl Häufungspunkt dieser Folge. ◁

Satz 3.13. *Ein Element a ist genau dann Häufungspunkt einer Folge, wenn es Grenzwert einer ihrer Teilfolgen ist.*

Beweis. Es sei a ein Häufungspunkt. Wähle n_1, so dass $|x_{n_1} - a| < 1$, dann wähle $n_2 > n_1$, so dass $|x_{n_2} - a| < \frac{1}{2}$, dann $n_3 > n_2$, so dass $|x_{n_3} - a| < \frac{1}{3}$ usw. Dies ist möglich, da es für jedes k unendlich viele n gibt, so dass $|x_n - a| < \frac{1}{k}$. Offensichtlich gilt $x_{n_k} \to a \ (k \to \infty)$.

Umgekehrt sei $x_{n_k} \to a \ (k \to \infty)$. Ist $\varepsilon > 0$, so gibt es ein k_0, so dass für $k \geq k_0$ gilt $|x_{n_k} - a| < \varepsilon$, also $|x_n - a| < \varepsilon$ für unendlich viele n. □

Für einen angeordneten Körper besteht ein Zusammenhang mit den Begriffen aus dem vorigen Abschnitt:

Satz 3.14. *Ist der limes superior endlich, so ist er der größte Häufungspunkt. Ist der limes inferior endlich, so ist er der kleinste Häufungspunkt.*

Beweis. Es sei $\overline{\lim_{n\to\infty}} x_n = a$. Zu gegebenem $\varepsilon > 0$ und n_0 finden wir $n_1 > n_0$ mit der Eigenschaft $|x_{n_1} - a| < \varepsilon$ wie folgt.

Wegen $x_n^+ \to a$ $(n \to \infty)$ gibt es $n > n_0$, so dass $a - \varepsilon < x_n^+ < a + \varepsilon$. Insbesondere ist $x_k < a+\varepsilon$ für alle $k \geq n$, und $a-\varepsilon$ ist keine obere Schranke von $\{x_n, x_{n+1}, \dots\}$. Somit gibt es ein $n_1 \geq n$, so dass $a - \varepsilon < x_{n_1}$.

Wenden wir dasselbe auf n_1 anstelle n_0 an, so finden wir $n_2 > n_1$ mit der selben Eigenschaft usw., also ist a ein Häufungspunkt.

Nun sei $b > a$. Dann gibt es zu $\varepsilon = \frac{b-a}{2}$ ein n_0, so dass für $n \geq n_0$ gilt $|x_n^+ - a| < \varepsilon$, also $x_n \leq x_n^+ < b - \varepsilon$. Für alle n mit der Eigenschaft $|x_n - b| < \varepsilon$ gilt also $n < n_0$, und somit ist b kein Häufungspunkt. □

Definition 3.7. *Eine Teilmenge A von K heißt* abgeschlossen *in K, wenn jeder Häufungspunkt einer jeden Folge von Elementen von A ebenfalls in A liegt.*

Eine Teilmenge U von K heißt offen *in K, wenn es für jedes Element x von U ein $\varepsilon \in L_+$ gibt, so dass alle Elemente y von K mit der Eigenschaft $|x - y| < \varepsilon$ in U liegen.*

Beispiel. Nach Satz 3.7 und 3.13 ist für jede reelle Zahl a die Menge $[a, \infty[$ abgeschlossen in $\mathbb{R}$. ◁

Beispiel. Die Menge $[0, 1[$ ist weder offen noch abgeschlossen in $\mathbb{R}$. ◁

Beispiel. Die Mengen K und $\varnothing$ sind sowohl offen als auch abgeschlossen in K. ◁

Beispiel. Die Menge $]0, 1[$ offen in $\mathbb{R}$, aber nicht in $\mathbb{C}$. ◁

Satz 3.15. *Eine Menge A ist genau dann abgeschlossen in K, wenn ihr Komplement $U = K \setminus A$ offen in K ist.*

Beweis. Angenommen, U ist nicht offen. Dann hat U ein Element x, so dass es für jedes $\varepsilon > 0$ ein $y \in A$ gibt, so dass $|x - y| < \varepsilon$. Insbesondere gibt es für jedes n ein y_n, so dass $|x - y_n| < \frac{1}{n}$. Dann ist $y_n \to x$ $(n \to \infty)$, also A nicht abgeschlossen.

Angenommen, U ist offen. Ist nun x_n eine Folge in $A = K \setminus U$, so kann ein Element x von U nicht ihr Häufungspunkt sein. Es gibt nämlich ein $\varepsilon > 0$, so dass für alle $y \in U$ mit $|y - x| < \varepsilon$ gilt $y \in U$. Für alle n gilt somit $|x_n - x| \geq \varepsilon$. Somit muss jeder Häufungspunkt von x_n in A liegen, so dass A abgeschlossen ist. □

3.9 Cauchy-Folgen

Für angeordnete Körper kennen wir ein Konvergenzkriterium, für dessen Anwendung noch kein Grenzwert bekannt sein muss, nämlich das Monotoniekriterium. Auch für Körper mit Absolutbetrag gibt es ein Konvergenzkriterium mit dieser Eigenschaft.

Definition 3.8. *Es sei K ein Körper mit Absolutbetrag mit Werten in einem angeordneten Körper L. Eine Folge x_n in K heißt* Cauchy-Folge[27], *wenn es für alle $\varepsilon \in L_+$ eine natürliche Zahl n_0 gibt, so dass für alle natürlichen Zahlen m und n mit den Eigenschaften $m \geq n_0$ und $n \geq n_0$ gilt $|x_m - x_n| < \varepsilon$.*

Satz 3.16. *Es sei $K = \mathbb{R}$ oder $K = \mathbb{C}$. Eine Folge in K ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.*

Beweis. Ist $x_n \to a \ (n \to \infty)$, so folgt aus

$$|x_m - x_n| \leq |x_m - a| + |a - x_n|,$$

dass x_n eine Cauchy-Folge ist.

Nun sei umgekehrt x_n eine Cauchy-Folge. Wir beweisen zunächst, dass sie beschränkt ist. Dazu wählen wir $\varepsilon = 1$ in der Definition der Cauchy-Folge und setzen

$$c = \max\{|x_0|, \ldots, |x_{n_0-1}|, |x_{n_0}| + 1\}.$$

Nun folgt die Behauptung wie im Beweis von Satz 3.2.

Wir können also Satz 3.6 anwenden und sehen, dass es eine Teilfolge x_{n_k} gibt, die einen Grenzwert a besitzt. Wir behaupten, dass a auch Grenzwert der Gesamtfolge x_n ist. Ist $\varepsilon > 0$, so gibt es wegen $x_{n_k} \to a \ (n \to \infty)$ ein k_0, so dass für $k \geq k_0$ gilt

$$|x_{n_k} - a| < \frac{\varepsilon}{2}.$$

Da x_n eine Cauchy-Folge ist, gibt es außerdem ein n_0, so dass für $m \geq n_0$ und $n \geq n_0$ gilt

$$|x_m - x_n| < \frac{\varepsilon}{2}.$$

Da die Folge n_k streng monoton wachsend ist, gibt es ein $k_1 \geq k_0$, so dass $n_{k_1} \geq n_0$ ist. Setzen wir in der vorigen Ungleichung $m_1 = n_{k_1}$, so folgt für $n \geq n_0$ mit der Dreiecksungleichung

$$|x_n - a| \leq |x_n - x_{m_1}| + |x_{m_1} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Da ε eine beliebige positive Zahl war, folgt $x_n \to a \ (n \to \infty)$. □

[27] oder *Fundamentalfolge*

Beispiel. Wir müssen bei der geometrischen Folge $x_n = x_0 q^n$ noch den Fall betrachten, dass $|q| = 1$, aber $q \neq 1$. Dann ist

$$|x_{n+1} - x_n| = |x_0||q - 1|$$

konstant, also ist die Folge keine Cauchy-Folge und somit divergent. $\triangleleft$

Definition 3.9. *Ein Körper K mit Absolutbetrag heißt* vollständig*, wenn jede Cauchy-Folge konvergent ist.*

Wir sehen also, dass $\mathbb{R}$ und $\mathbb{C}$ vollständige Körper mit Absolutbetrag sind. Der Körper $\mathbb{Q}$ ist nicht vollständig, denn das Heronverfahren erzeugt z. B. für $A = 2$ eine divergente Cauchy-Folge.

Man kann eine Vervollständigung von $\mathbb{Q}$ auch auf andere Weise konstruieren, nämlich als Menge der Äquivalenzklassen von Cauchy-Folgen, wobei zwei Cauchy-Folgen x_n und y_n äquivalent genannt werden, wenn es für jede positive rationale Zahl ε ein n_0 gibt, so dass für $n \geq n_0$ gilt $|x_n - y_n| < \varepsilon$. Auch auf dieser Menge kann man wieder die Strurktur eines Körpers mit Absolutbetrag einführen. Der so entstehende Körper ist zu isomorph $\mathbb{R}$.

Die Vervollständigung von $\mathbb{Q}$ bezüglich der Anordnung und bezüglich des Absolutbetrags liefern also das selbe Ergebnis. Es gibt allerdings angeordnete Körper, z. B. den Levi-Civita-Körper, die bezüglich des Absolutbetrags vollständig sind, bezüglich der Anordnung jedoch nicht. In diesem Fall lässt sich die Körperstruktur nicht fortsetzen auf die Vervollständigung durch Schnitte.

4 Reihen

Vorlesung 9a
01.06.2022

Der antike Philosoph Zenon kam auf das Paradoxon, dass der Held Achilles einen Wettlauf mit einer Schildkröte nicht gewinnen könnte, falls ihr beim Start ein Vorsprung gewährt würde. Wenn Achilles beim Startpunkt der Schildkröte ankommt, ist sie bereits an einem anderen Ort, wenn er dort ankommt, wieder an einem anderen usw.

Nach der mengentheoretischen Begründung der Mathematik ist es aber gar nicht paradox, dass ein beschränktes Zeitintervall unendlich viele Zeitpunkte enthält. Hat die Schildkröte den Vorsprung a und die q-fache Geschwindigkeit von Achilles, wobei $0 < q < 1$, und bezeichnen wir den Weg von Achilles bis zum Ort des Überholens mit b, so gilt offenbar $a + qb = b$, also $b = \frac{a}{1-q}$. Die Länge der n-ten von Achilles durchlaufenen Teilstrecke ist aq^n. Wenn man all diese unendlich vielen Teilstrecken in irgend einem Sinne summieren könnte, sollte sich b ergeben.

Ein anderes Beispiel ergibt sich bei der Betrachtung von Stellenwertsystemen. Will man einem unendlichen Dezimalbruch wie

$$3{,}14159265\ldots$$

einen Sinn geben, so muss man ebenfalls unendlich viele Terme summieren, nämlich

$$3+1\cdot10^{-1}+4\cdot10^{-2}+1\cdot10^{-3}+5\cdot10^{-4}+9\cdot10^{-5}+2\cdot10^{-6}+6\cdot10^{-7}+5\cdot10^{-8}+\ldots$$

4.1 Konvergenz von Reihen

Definition 4.1. *Für eine Folge reeller oder komplexer Zahlen x_n sei*

$$x_{n_1} + x_{n_1+1} + \ldots = \lim_{m\to\infty} (x_{n_1} + x_{n_1+1} + \ldots + x_m),$$

oder in anderer Schreibweise

$$\sum_{n=n_1}^{\infty} x_n = \lim_{m\to\infty} \sum_{n=n_1}^{m} x_n.$$

Der neu definierte Term auf der linken Seite heißt Reihe*, sein Wert* Summe der Reihe*. Die Summe auf der rechten Seite heißt m*-te Partialsumme*. Die Reihe heißt konvergent, wenn die Folge ihrer Partialsummen konvergent ist, sonst heißt sie divergent. (Im letzteren Fall haben beide Seiten eigentlich keinen Sinn.)*

Beispiel. Aus der dritten binomischen Formel folgt die Summenformel der geometrischen Folge für $q \neq 1$:

$$1 + q + q^2 + \ldots + q^m = \frac{1-q^{m+1}}{1-q}.$$

Für $|q| < 1$ ist die *geometrische Reihe* konvergent:

$$\sum_{n=0}^{\infty} q^n = \lim_{m\to\infty} \frac{1-q^{m+1}}{1-q} = \frac{1}{1-q}.$$

(Dies löst das Zenonsche Paradoxon auf.) Andernfalls ist sie divergent (auch im Fall $q = 1$).

Mit $q = \frac{1}{10^k}$ können wir periodische Dezimalbrüche in gemeine Brüche umwandeln:

$$\begin{aligned}
0{,}\overline{428571} &= \frac{428571}{1000000} + \frac{428571}{1000000^2} + \dots \\
&= \frac{428571}{1000000}\left(1 + \frac{1}{1000000} + \frac{1}{1000000^2} + \dots\right) \\
&= \frac{428571}{1000000} \cdot \frac{1}{1 - \frac{1}{1000000}} = \frac{428571}{999999} = \frac{3}{7}.
\end{aligned}$$

Alle periodischen Dezimalbrüche stellen also rationale Zahlen dar. Das Selbe gilt in anderen Stellenwertsystemen. Übrigens ist

$$0{,}\overline{9} = \frac{9}{10}(1 + 10^{-1} + 10^{-2} + \dots) = \frac{9}{10} \cdot \frac{1}{1 - \frac{1}{10}} = 1.$$

Dezimalbrüche mit Neunerperiode (oder allgemein Systembrüche mit Periode $g-1$ im Fall der Grundzahl g) sind also überflüssig. $\triangleleft$

Beispiel. Nach den Rechenregeln für das Summenzeichen gilt

$$\begin{aligned}
\sum_{n=1}^{m} \frac{1}{n(n+1)} = \sum_{n=1}^{m}\left(\frac{1}{n} - \frac{1}{n+1}\right) &= \sum_{n=1}^{m} \frac{1}{n} - \sum_{n=1}^{m} \frac{1}{n+1} \\
&= \sum_{n=1}^{m} \frac{1}{n} - \sum_{n=2}^{m+1} \frac{1}{n} = 1 - \frac{1}{m+1}.
\end{aligned}$$

Durch Grenzübergang folgt

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1.$$

Summen der Form $\sum\limits_{n=n_1}^{n_2} (x_n - x_{n+1})$ nennt man *Teleskopsummen*. Eine solche trat auch im Beweis der dritten binomischen Formel auf. In der Lösung zu Aufgabe 33(c) benutzt man ein Teleskopprodukt. $\triangleleft$

Beispiel. Die *harmonische Reihe* ist

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots$$

Wir zerlegen die Partialsummen für den Fall, dass m eine Zweierpotenz ist, in Teilstücke:

$$1 + \left(\frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^k+1} + \dots + \frac{1}{2^{k+1}}\right)$$

Schätzen wir die Glieder in jedem Teilstück durch das Letzte ab, so folgt

$$1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{2^{k+1}} \geq 1+\frac{1}{2}+\frac{2}{4}+\frac{4}{8}+\ldots+\frac{2^k}{2^{k+1}}=1+\frac{k+1}{2}.$$

Die harmonische Reihe ist also divergent. $\triangleleft$

Aus Satz 3.4 und 3.7 ergibt sich:

Folgerung 4.1. *Es sei*

$$\sum_{n=n_1}^{\infty} x_n = a, \qquad \sum_{n=n_1}^{\infty} y_n = b, \qquad c \in \mathbb{C}.$$

Dann gilt:

$$\sum_{n=n_1}^{\infty} (x_n + y_n) = a+b, \qquad \sum_{n=n_1}^{\infty} c x_n = ca.$$

Ist $x_n \leq y_n$ für alle n, so gilt $a \leq b$.

Warnung: Es gibt keine Formel für $\sum\limits_{n=n_1}^{\infty} x_n y_n$.

Satz 4.1 (Cauchy-Kriterium). *Die Reihe $\sum\limits_{n=n_1}^{\infty} x_n$ konvergiert genau dann, wenn es für jedes $\varepsilon > 0$ ein m_0 gibt, so dass für $m \geq m_0$ und $p \geq 0$ gilt*

$$\left| \sum_{n=m+1}^{m+p} x_n \right| < \varepsilon.$$

Für die Konvergenz der Reihe ist also notwendig (aber nicht hinreichend, siehe harmonische Folge), dass die Folge der Glieder x_n eine Nullfolge ist.

Beweis. Mit der Bezeichnung $s_m = \sum\limits_{n=n_1}^{m} x_n$ gilt

$$\sum_{n=m+1}^{m+p} x_n = s_{m+p} - s_m.$$

Die im Satz genannte Bedingung bedeutet also, dass die Folge der Partialsummen s_m eine Cauchy-Folge ist. Damit folgt die Behauptung aus Satz 3.16. $\square$

Nun betrachten wir Reihen mit reellen Gliedern.

Folgerung 4.2. *Eine Reihe mit nichtnegativen Gliedern ist genau dann konvergent, wenn die Folge ihrer Partialsummen von oben beschränkt ist.*

Dies folgt aus Satz 3.2 und dem Monotoniekriterium (Satz 3.9), denn die Folge s_m ist genau dann monoton wachsend, wenn $x_n \geq 0$ für alle n gilt.

Beispiel. Für alle $m \geq 1$ erhalten wir mit einem früheren Beispiel

$$\sum_{n=1}^{m} \frac{1}{n^2} \leq 1 + \sum_{n=2}^{m} \frac{1}{n(n-1)} = 1 + \sum_{n=1}^{m-1} \frac{1}{(n+1)n} \leq 2,$$

also ist die Reihe $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ konvergent. $\triangleleft$

Eine Reihe heißt *alternierend*, wenn ihre Glieder abwechselnd nichtnegativ und nichtpositiv sind, z. B.

$$x_{2k} \geq 0, \qquad x_{2k+1} \leq 0.$$

Satz 4.2 (Leibniz-Kriterium). *Ist die Reihe* $\sum\limits_{n=0}^{\infty} x_n$ *alternierend und ist* $|x_n|$ *eine monoton fallende Folge mit dem Grenzwert* 0*, so ist die Reihe konvergent.*

Beweis. Wir betrachten die Teilfolgen s_{2k} und s_{2k+1} der Partialsummenfolge s_m. Sind die Vorzeichen wie oben, so gilt

$$\begin{aligned} s_{2k+1} - s_{2k} &= x_{2k+1} \leq 0, \\ s_{2(k+1)} - s_{2k} &= x_{2k+1} + x_{2k+2} = -|x_{2k+1}| + |x_{2k+2}| \leq 0, \\ s_{2k+1} - s_{2k-1} &= x_{2k} + x_{2k+1} = |x_{2k}| - |x_{2k+1}| \geq 0. \end{aligned}$$

Also ist die Folge s_{2k} monoton fallend mit unterer Schranke s_1 und die Folge s_{2k+1} monoton wachsend mit oberer Schranke s_0. Nach dem Monotoniekriterium sind beide konvergent, und wegen $s_{2k} - s_{2k+1} = |x_{2k+1}| \to 0 \ (k \to \infty)$ haben beide den selben Grenzwert a. Mit der Präsenzaufgabe 17 folgt nun $s_m \to a \ (m \to \infty)$. $\square$

Wir erhalten sogar eine Fehlerabschätzung: Für jedes m liegt a zwischen s_m und s_{m+1}, also

$$|s_m - a| \leq |x_{m+1}|.$$

Beispiel. Die Reihe

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + - \dots$$

ist konvergent, die Reihensumme a liegt zwischen $\frac{1}{2}$ und 1.

Nun ordnen wir die Reihe um, indem wir auf ein positives Glied immer zwei negative folgen lassen, und betrachten nur eine Teilfolge der Partialsummenfolge:

$$\begin{aligned}\left(1-\frac{1}{2}\right)-\frac{1}{4}+\left(\frac{1}{3}-\frac{1}{6}\right)-\frac{1}{8}+\left(\frac{1}{5}-\frac{1}{10}\right)-\frac{1}{12}+\left(\frac{1}{7}-\frac{1}{14}\right)-\frac{1}{16}+\dots\\ =\frac{1}{2}-\frac{1}{4}+\frac{1}{6}-\frac{1}{8}+\frac{1}{10}-\frac{1}{12}+\frac{1}{14}-\frac{1}{16}+-\dots=\frac{a}{2},\end{aligned}$$

wobei wir Folgerung 4.1 benutzt haben. Wegen $a \neq 0$ ist $\frac{a}{2} \neq a$, die Reihensumme hat sich also durch das Umordnen geändert. $\triangleleft$

4.2 Umordnungen und absolute Konvergenz

Vorlesung 9b
03.06.2022

Wir betrachten wieder Reihen mit reellen oder komplexen Gliedern.

Definition 4.2. *Die Reihe* $\sum\limits_{n=0}^{\infty} y_n$ *heißt* Umordnung *der Reihe* $\sum\limits_{n=0}^{\infty} x_n$*, wenn es eine bijektive Abbildung* $\sigma : \mathbb{N} \to \mathbb{N}$ *gibt, so dass* $y_n = x_{\sigma(n)}$ *für alle* n *gilt.*

Dies definiert offensichtlich eine Äquivalenzrelation zwischen Reihen.

Definition 4.3. *Die Reihe* $\sum\limits_{n=0}^{\infty} x_n$ *heißt* absolut konvergent*, wenn die Reihe* $\sum\limits_{n=0}^{\infty} |x_n|$ *konvergent ist. Eine Reihe, die konvergent, aber nicht absolut konvergent ist, nennt man* bedingt konvergent*.*

Der Begriff „absolute Konvergenz“ ist gerechtfertigt, denn aus Satz 4.1 erhalten wir:

Folgerung 4.3. *Jede absolut konvergente Reihe ist konvergent.*

Nach der Dreiecksungleichung gilt nämlich für alle natürlichen Zahlen m und p

$$\left|\sum_{n=m+1}^{m+p} x_n\right| \leq \sum_{n=m+1}^{m+p} |x_n|.$$

Erfüllt $\sum\limits_{n=0}^{\infty} |x_n|$ das Cauchy-Kriterium, so auch $\sum\limits_{n=0}^{\infty} x_n$.

Für Folgen gibt es keinen Begriff der absoluten Konvergenz.

Satz 4.3. *Ist eine Reihe absolut konvergent, so ist auch jede Umordnung absolut konvergent und hat die selbe Summe.*

Beweis. Es sei $\sigma : \mathbb{N} \to \mathbb{N}$ bijektiv und τ die Umkehrabbildung. Ist nun $\varepsilon > 0$, dann gibt es wegen der absoluten Konvergenz ein m_0, so dass

$$\sum_{n=m_0+1}^{\infty} |x_n| = \sum_{n=0}^{\infty} |x_n| - \sum_{n=0}^{m_0} |x_n| < \frac{\varepsilon}{2}.$$

Setzen wir $l_0 = \max\{\tau(0), \ldots, \tau(m_0)\}$, dann gilt

$$\{0, 1, 2, \ldots, m_0\} \subseteq \{\sigma(0), \sigma(1), \ldots, \sigma(l_0)\}.$$

Für $l \geq l_0$ und $m \geq m_0$ folgt mit der Dreiecksungleichung

$$\left|\sum_{k=0}^{l} x_{\sigma(k)} - \sum_{n=0}^{m} x_n\right| = \left|\sum_{\substack{n=0\\ \tau(n)\leq l}}^{m'} x_n - \sum_{n=0}^{m} x_n\right| \leq \sum_{n=m_0+1}^{\max\{m,m'\}} |x_n| \leq \frac{\varepsilon}{2},$$

wobei $m' = \max\{\sigma(0), \ldots, \sigma(l)\}$, denn unter den Gliedern, die sich wegkürzen, sind $x_0, \ldots, x_{m_0}$.

Wegen der Konvergenz der ursprünglichen Reihe können wir m so wählen, dass

$$\left|\sum_{n=0}^{m} x_n - a\right| < \frac{\varepsilon}{2},$$

wobei a die Reihensumme bezeichnet. Mit der Dreiecksungleichung folgt

$$\left|\sum_{k=0}^{l} x_{\sigma(k)} - a\right| \leq \left|\sum_{k=0}^{l} x_{\sigma(k)} - \sum_{n=0}^{m} x_n\right| + \left|\sum_{n=0}^{m} x_n - a\right| < \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, haben wir bewiesen, dass

$$\sum_{k=0}^{\infty} x_{\sigma(k)} = a.$$

Die selbe Überlegung mit $|x_n|$ an Stelle von x_n beweist die absolute Konvergenz der umgeordneten Reihe. □

Satz 4.4. *Ist eine Reihe mit reellen Gliedern bedingt konvergent, so gibt es für jedes abgeschlossene Intervall I eine Umordnung, so dass die Menge der Häufungspunkte der Partialsummenfolge gleich I ist, und dass für beschränktes I auch diese Folge beschränkt ist.*

Da eine Folge genau dann konvergent ist, wenn sie genau einen Häufungspunkt besitzt, sehen wir insbesondere, dass man durch Umordnen einer bedingt konvergenten Reihe jede beliebige Reihensumme erzeugen kann.

Beweisskizze. Wir setzen

$$x_n^+ = \max\{x_n, 0\}, \qquad x_n^- = \min\{x_n, 0\},$$

so dass $x_n = x_n^+ + x_n^-$. Wir behaupten, dass die Reihen

$$\sum_{n=1}^{\infty} x_n^+, \qquad \sum_{n=1}^{\infty} x_n^-$$

divergent sind. Angenommen, die erste der beiden wäre konvergent. Wegen $x_n^- = x_n^+ - x_n$ wäre nach Folgerung 4.1 auch die zweite konvergent, und wegen $|x_n| = x_n^+ - x_n^-$ wäre die gegebene Reihe absolut konvergent (Widerspruch).

Wir können annehmen, dass $x_n \neq 0$ für alle n. Es sei u_n die Teilfolge der positiven und v_n die der negativen Glieder in der Folge x_n. Man erhält diese aus den Folgen x_n^+ bzw. x_n^-, indem man alle Nullen entfernt. Es sei $c = \inf I$ und $d = \sup I$. Sind c und d endlich, dann bilden wir eine neue Folge x'_n mit Partialsummen s'_m wie folgt:

$u_1, u_2, \ldots, u_{n_1}$ solange, bis $s'_{n_1} \geq d$, dann

$v_1, v_2, \ldots, v_{n_2}$ solange, bis $s'_{n_1+n_2} \leq c$, dann

$u_{n_1+1}, u_{n_1+2}, \ldots u_{n_1+n_3}$ solange, bis $s'_{n_1+n_2+n_3} \geq d$

usw.

Ist hingegen $d = \infty$, dann wählt man n_{2k+1} jeweils so, dass

$$s'_{n_1+\ldots+n_{2k+1}} \geq k.$$

Ist $c = \infty$, also $I = \varnothing$, dann wählt man n_{2k+1} ebenso und setzt $n_{2k} = 1$ für alle k. Ist schließlich $c = -\infty$, dann wählt man n_{2k} so, dass

$$s'_{n_1+\ldots+n_{2k}} \leq -k.$$

All dies ist nach Folgerung 4.2 möglich wegen der Divergenz der Reihen

$$\sum_{n=1}^{\infty} u_n, \qquad \sum_{n=1}^{\infty} v_n.$$

Für jedes $a \in I$ gibt es beliebig große n, so dass $s'_{n-1} < a \leq s'_n$, also $|s'_n - a| < x'_n$. Da x_n eine Nullfolge ist, sind u_n und v_n Nullfolgen, also ist auch x'_n eine, und somit ist a ein Häufungspunkt.

Ist d endlich, so gilt für $n_{2k-1} \leq n < n_{2k+1}$, dass $s'_n \leq d + u_{n_1+\ldots+n_{2k-1}}$. Da u_n eine Nullfolge ist, ist kein Häufungspunkt größer als d. Analog zeigt man, dass für endliches c auch kein Häufungspunkt kleiner als c sein kann, und nebenbei ergibt sich, dass für beschränktes I die Folge s'_n beschränkt ist. □

Folgerung 4.4. *Eine Reihe mit reellen oder komplexen Gliedern ist genau dann absolut konvergent, wenn jede ihrer Umordnungen konvergent ist.*

Für reelle Glieder folgt eine Richtung aus Satz 4.3 und die andere aus Satz 4.4. Sind die Glieder $z_n = x_n + iy_n$ komplex und ist jede Umordnung von $\sum_{n=1}^{\infty} z_n$ konvergent, so ist auch jede Umordnung von $\sum_{n=1}^{\infty} x_n$ und jede von $\sum_{n=1}^{\infty} y_n$ konvergent. Nach Satz 4.4 sind beide absolut konvergent, und wie in der Bemerkung am Ende von Abschnitt 3.2 folgt die absolute Konvergenz von $\sum_{n=1}^{\infty} z_n$.

4.3 Kriterien der absoluten Konvergenz

Vorlesung 10a 08.06.2022

Folgerung 4.4 zeigt die Bedeutung der absoluten Konvergenz. Man möchte feststellen, ob sie bei einer gegebenen Reihe vorliegt. Die absolute Konvergenz der Reihe $\sum_{n=0}^{\infty} x_n$ ist nach Folgerung 4.2 gleichbedeutend mit der Beschränktheit der Partialsummenfolge der Reihe $\sum_{n=0}^{\infty} |x_n|$.

Folgerung 4.5 (Vergleichskriterium)**.** *Für alle n sei $0 \leq x_n \leq y_n$. Ist die Reihe $\sum_{n=0}^{\infty} y_n$ konvergent, so auch die Reihe $\sum_{n=0}^{\infty} x_n$.*

In der Tat ist in dieser Situation für alle m

$$\sum_{n=0}^{m} x_n \leq \sum_{n=0}^{m} y_n.$$

Wir haben dieses Kriterium praktisch bereits im Fall $x_n = \frac{1}{n^2}$, $y_n = \frac{1}{n(n+1)}$ angewendet.

Beispiel. Durch Vergleich mit der geometrischen Reihe mit dem Quotienten g sind Systembrüche im Stellenwertsystem zur Grundzahl g (also im Fall $g = 10$ Dezimalbrüche) konvergent. ◁

Beispiel. Für $y_n = \frac{(n+1)^n}{n^{n+1}}$ gilt $y_n = \left(1 + \frac{1}{n}\right)^n \frac{1}{n} \geq \frac{2}{n}$ nach der Bernoulli-Ungleichung, also ist die Reihe $\sum_{n=1}^{\infty} y_n$ divergent. ◁

Satz 4.5 (Wurzelkriterium)**.** *Es sei*

$$a = \overline{\lim_{n\to\infty}} \sqrt[n]{|x_n|}.$$

Ist $a < 1$, so ist die Reihe $\sum_{n=0}^{\infty} x_n$ absolut konvergent, ist hingegen $a > 1$, so ist die Reihe divergent. (Letzteres auch, wenn $a = \infty$.)

Beweis. Ist $a < 1$, so gibt es ein $q \in \mathbb{R}$, so dass $a < q < 1$, und ein n_0, so dass für $n \geq n_0$ gilt $\sqrt[n]{|x_n|} \leq q$, also wegen der Monotonie der Wurzelfunktion $|x_n| \leq q^n$, und die Behauptung folgt durch Vergleich mit der geometrischen Reihe.

Ist $a > 1$, so gibt es unendlich viele n mit der Eigenschaft $\sqrt[n]{|x_n|} \geq 1$, d. h. $|x_n| \geq 1$. Also ist x_n keine Nullfolge und die Reihe nach Satz 4.1 divergent. □

In dem obigen Beispiel $x_n = \frac{(n+1)^n}{n^{n+1}}$ versagt das Wurzelkriterium:

$$\sqrt[n]{x_n} = \frac{n+1}{n\sqrt[n]{n}} = \frac{1+\frac{1}{n}}{\sqrt[n]{n}} \to 1 \ \ (n \to \infty).$$

Das Wurzelkriterium ist oft schwer nachzuprüfen. Manchmal hilft

Satz 4.6 (Quotientenkriterium)**.** *Es sei $x_n \neq 0$ für genügend große n.*

(i) *Ist $\overline{\lim\limits_{n\to\infty}} \left|\frac{x_{n+1}}{x_n}\right| < 1$, so ist die Reihe $\sum_{n=0}^{\infty} x_n$ absolut konvergent.*

(ii) *Ist $\left|\frac{x_{n+1}}{x_n}\right| \geq 1$ für genügend große n, so ist die Reihe divergent.*

Beweis. **(i)** Ist $\overline{\lim\limits_{n\to\infty}} \left|\frac{x_{n+1}}{x_n}\right| < q < 1$, dann gibt es ein n_0, so dass für $n \geq n_0$ gilt $\left|\frac{x_{n+1}}{x_n}\right| \leq q$, also $|x_{n+1}| \leq q|x_n|$. Durch vollständige Induktion zeigt man $|x_n| \leq q^{n-n_0}|x_{n_0}|$, und die Behauptung folgt durch Vergleich mit der geometrischen Reihe.
(ii) Für alle n ab einem gewissen n_0 gilt $|x_{n+1}| \geq |x_n|$, also ist x_n keine Nullfolge. □

Bemerkung. Man sieht am Beweis, dass

$$\overline{\lim_{n\to\infty}} \sqrt[n]{|x_n|} \leq \overline{\lim_{n\to\infty}} \left|\frac{x_{n+1}}{x_n}\right|,$$

also ist das Wurzelkriterium mindestens so stark wie das Quotientenkriterium. Stärker ist es z. B. für

$$1 + 1 + q + q + q^2 + q^2 + \dots \qquad (|q| < 1).$$

Bemerkung. Die Bedingung in (ii) ist z. B. für $\lim\limits_{n\to\infty} \left|\frac{x_{n+1}}{x_n}\right| > 1$ erfüllt. Man kann sie nicht durch den limes superior ausdrücken, wie folgendes Beispiel zeigt:

$$1 + 2 + q + 2q + q^2 + 2q^2 + \ldots \qquad (|q| < 1).$$

Beispiel. Für $x_n = \frac{n!^2}{(2n)!}$ ist

$$\frac{x_{n+1}}{x_n} = \frac{(n+1)^2}{(2n+1)(2n+2)} \to \frac{1}{4} \ \ (n \to \infty),$$

also ist $\sum\limits_{n=0}^{\infty} x_n$ (absolut) konvergent. ◁

Das folgende Ergebnis kennen wir bereits in den Spezialfällen $s = 1$ und $s = 2$.

Satz 4.7. *Es sei $s \in \mathbb{Q}$. Die Reihe*

$$\sum_{n=1}^{\infty} \frac{1}{n^s}$$

ist für $s > 1$ konvergent und für $s \leq 1$ divergent.

Beweis. Für $s \leq 0$ bilden die Glieder keine Nullfolge, und die Behauptung ist klar. Nun sei $s > 0$. Wegen der Positivität der Glieder genügt es, eine Teilfolge der Partialsummenfolge zu betrachten. Dazu schätzen wir Teilstücke der Reihe nach oben und unten ab. Für $k \in \mathbb{N}$ ist wegen der Monotonie der Potenzfunktion

$$\frac{2^k}{(2^{k+1})^s} \leq \frac{1}{(2^k+1)^s} + \frac{1}{(2^k+2)^s} + \ldots + \frac{1}{(2^{k+1})^s} \leq \frac{2^k}{(2^k)^s}.$$

Summieren wir über $k \in \{0, 1, \ldots, l-1\}$, so erhalten wir

$$2^{-s} \sum_{k=0}^{l-1} \left(2^{1-s}\right)^k \leq \sum_{n=2}^{2^l} \frac{1}{n^s} \leq \sum_{k=0}^{l-1} \left(2^{1-s}\right)^k.$$

Wegen der Monotonie der Potenzfunktion ist $2^{1-s} < 1$ für $s > 1$ und $2^{1-s} \geq 1$ für $s \leq 1$. Nun folgt die Behauptung durch Vergleich mit der geometrischen Reihe. □

Beispiel. Wegen

$$\frac{\sqrt{n+1} - \sqrt{n}}{n} = \frac{(n+1) - n}{n(\sqrt{n+1} + \sqrt{n})} \leq \frac{1}{n(\sqrt{n} + \sqrt{n})}$$

ist die Reihe $\sum\limits_{n=0}^{\infty} \frac{\sqrt{n+1}-\sqrt{n}}{n}$ konvergent. ◁

4.4 Potenzreihen

Es sei $K = \mathbb{R}$ oder $K = \mathbb{C}$. Jedes Polynom mit Koeffizienten a_n in K definiert eine Funktion f auf K, genannt ganzrationale Funktion:

$$f(z) = \sum_{n=0}^{m} a_n z^n.$$

Als Verallgemeinerung betrachten wir *Potenzreihen*

$$\sum_{n=0}^{\infty} a_n z^n. \tag{4}$$

Satz 4.8. (i) *Es sei $\overline{\lim\limits_{n\to\infty}} \sqrt[n]{|a_n|} = a \in \mathbb{R}$. Dann ist die Potenzreihe (4) für $a|z| < 1$ absolut konvergent und für $a|z| > 1$ divergent.*

(ii) *Ist die Folge $\sqrt[n]{|a_n|}$ unbeschränkt, so ist die Potenzreihe (4) für $z \neq 0$ divergent.*

Beweis. Aussage (i) folgt wegen $\sqrt[n]{|a_n z^n|} = \sqrt[n]{|a_n|} \cdot |z|$ aus dem Wurzelkriterium.

Im Fall (ii) ist die Folge $\sqrt[n]{|a_n|} \cdot |z|$ unbeschränkt, also auch die Folge der n-ten Potenzen $|a_n z^n|$. □

Definition 4.4. *Der* Konvergenzradius *der Potenzreihe ist im Fall (i) die Zahl $r = \frac{1}{a}$ (wobei wir $\frac{1}{0} = \infty$ setzen) und im Fall (ii) die Zahl $r = 0$. Die Menge $\{z \in K \,\big|\, |z| < r\}$ nennt man im Fall $K = \mathbb{R}$ das* Konvergenzintervall *und im Fall $K = \mathbb{C}$ den* Konvergenzkreis *der Potenzreihe.*

Folgerung 4.6. *Die Potenzreihe (4) ist für $|z| < r$ absolut konvergent und für $|z| > r$ divergent.*

Aus dem Quotientenkriterium erhalten wir:

Folgerung 4.7. *Ist die Folge $\left|\frac{a_{n+1}}{a_n}\right|$ konvergent, so gilt*

$$a = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|.$$

Gilt $\left|\frac{a_{n+1}}{a_n}\right| \to \infty$ $(n \to \infty)$, so ist $r = 0$.

Aus Folgerung 4.1 erhalten wir

Folgerung 4.8. *Ist eine weitere Potenzreihe* $\sum\limits_{n=0}^{\infty} b_n z^n$ *mit dem Konvergenzradius* s *gegeben, so gilt für* $|z| < \min\{r, s\}$

$$\sum_{n=0}^{\infty}(a_n + b_n)z^n = \sum_{n=0}^{\infty} a_n z^n + \sum_{n=0}^{\infty} b_n z^n.$$

Beispiel. Die geometrische Reihe ist eine Potenzreihe. Für $|z| < 1$ ist

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z},$$

für $|z| \geq 1$ ist die Reihe divergent. $\triangleleft$

Für das nächste Beispiel benötigen wir eine Verallgemeinerung der ersten binomischen Formel.

Satz 4.9 (Binomialtheorem). *Für beliebige Elemente* a, b *eines Körpers und jede natürliche Zahl* m *gilt*

$$(a+b)^m = \sum_{n=0}^{m} \binom{m}{n} a^{m-n} b^n,$$

wobei

$$\binom{m}{n} = \frac{m(m-1)\ldots(m-n+1)}{n!}$$

Dies kann man durch vollständige Induktion beweisen.

Beispiel. Wir haben die Potenzreihe

$$\sum_{n=0}^{\infty} \binom{m}{n} z^n = (1+z)^m,$$

weil die Binomialkoeffizienten $\binom{m}{n}$ für $n > m$ laut Definition verschwinden. $\triangleleft$

Beispiel. Die Reihe

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} \tag{5}$$

ist nach Folgerung 4.7 für alle $z \in \mathbb{C}$ absolut konvergent $\triangleleft$

Satz 4.10. *Für alle komplexen Zahlen* z *gilt*

$$\left(1 + \frac{z}{m}\right)^m \to \sum_{n=0}^{\infty} \frac{z^n}{n!} \qquad (m \to \infty).$$

Beweis. Nach dem Binomialtheorem gilt

$$\left(1+\frac{z}{m}\right)^m = \sum_{n=0}^{\infty} a_{m,n}\frac{z^n}{n!},$$

wobei

$$a_{m,n} = \frac{m(m-1)\dots(m-n+1)}{m^n} = \left(1-\frac{1}{m}\right)\left(1-\frac{2}{m}\right)\dots\left(1-\frac{n-1}{m}\right).$$

Wir müssen zeigen, dass die Folge der Zahlen

$$\sum_{n=0}^{\infty}\frac{z^n}{n!} - \left(1+\frac{z}{m}\right)^m = \sum_{n=0}^{\infty}(1-a_{m,n})\frac{z^n}{n!}$$

für $m \to \infty$ gegen Null konvergiert. Zunächst merken wir an, dass

$$1 = a_{m,0} = a_{m,1} > a_{m,2} > \dots > a_{m,m} > a_{m,m+1} = \dots = 0.$$

Mit der Dreiecksungleichung folgt für beliebiges n_0

$$\left|\sum_{n=0}^{\infty}\frac{z^n}{n!} - \left(1+\frac{z}{m}\right)^m\right| \le \left|\sum_{n=0}^{n_0}(1-a_{m,n})\frac{z^n}{n!}\right| + \sum_{n=n_0+1}^{\infty}\frac{|z|^n}{n!}.$$

Wegen der vor dem Satz gezeigten absoluten Konvergenz gibt es zu jedem $\varepsilon > 0$ ein n_0, so dass der zweite Term kleiner als $\frac{\varepsilon}{2}$ ist. Angesichts von $a_{m,n} \to 1$ $(m \to \infty)$ und den Rechenregeln für Grenzwerte gibt es ein m_0, so dass der erste Term für $m \ge m_0$ kleiner als $\frac{\varepsilon}{2}$ ist. $\square$

Definition 4.5. *Wir bezeichnen die Potenzreihe* (5) *als* Exponentialreihe *und kürzen ihre Summe mit* $\exp z$ *ab. Die dadurch dargestellte Funktion nennen wir* Exponentialfunktion.

Dies ist natürlich verträglich mit Definition 3.3, und es gilt $\overline{\exp z} = \exp \bar{z}$. Die Exponentialfunktion ist nicht die einzige Funktion, die man mit Hilfe von Potenzreihen definieren kann. Wir beweisen einige Eigenschaften solcher Funktionen.

Lemma 4.1. *Definieren wir auf der Menge* $\{z \in K \mid |z| \le c\}$ *eine Funktion* f *durch*

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \tag{6}$$

wobei c *kleiner als der Konvergenzradius* r *ist, so ist die Funktion beschränkt.*

Beweis. Für z im Definitionsbereich von f gilt nach der Dreiecksungleichung

$$\left|\sum_{n=0}^{m} a_n z^n\right| \leq \sum_{n=0}^{m} |a_n||z|^n \leq \sum_{n=0}^{m} |a_n|c^n.$$

Nach Folgerung 4.6 konvergiert die rechte Seite für $m \to \infty$ gegen eine Zahl d. Die Menge $\{w \in K \mid |w| \leq d\}$, die alle Partialsummen enthält, ist abgeschlossen (Übungsaufgabe). □

Definition 4.6. *Eine Zahl $a \in K$ heißt* Häufungspunkt der Teilmenge M *von K, wenn es für jedes $\varepsilon > 0$ ein Element x von M gibt, so dass $x \neq a$ und $|x - a| < \varepsilon$.*

(Es gibt dann unendlich viele solche x.)

Satz 4.11. *Hat die Potenzreihe (6) einen positiven Konvergenzradius und ist 0 ein Häufungspunkt der Nullstellenmenge von f, so ist $a_n = 0$ für alle n.*

Beweis. Angenommen, die Behauptung ist falsch. Dann gibt es eine kleinste natürliche Zahl m, so dass $a_m \neq 0$. Nun ist

$$f(z) = \sum_{n=m}^{\infty} a_n z^n = z^m \left(a_m + z \sum_{k=0}^{\infty} a_{m+1+k} z^k\right).$$

Für jede Nullstelle $z \neq 0$ mit $|z| < c$ folgt

$$0 = a_m + z \sum_{k=0}^{\infty} a_{m+1+k} z^k$$

Wir wählen c, so dass $0 < c < r$. Nach dem Lemma gibt es ein $d > 0$, so dass für $|z| \leq c$ gilt

$$\left|\sum_{k=0}^{\infty} a_{m+1+k} z^k\right| \leq d,$$

und wenn wir die obige Gleichung nach a_m auflösen, folgt

$$|a_m| \leq d|z|.$$

Dann kann 0 aber kein Häufungspunkt der Nullstellenmenge sein (Widerspruch). □

4.5 Doppelreihen

Vorlesung 10b
10.06.2022

Wollen wir alle Einträge in einer Tabelle von Zahlen addieren, so können wir erst die Spaltensummen bilden und diese addieren:

$$\begin{array}{cccc} z_{00} & z_{01} & z_{02} & \cdots \\ z_{10} & z_{11} & z_{12} & \cdots \\ \vdots & \vdots & \vdots & \\ \hline s_0 & s_1 & s_2 & \cdots \end{array}$$

Wir können auch erst die Zeilensummen bilden und diese addieren. Schließlich können wir auch die Tabelleneinträge in irgend einer Reihenfolge addieren. Dies läuft auf eine Umordnung hinaus und kann bei „Tabellen“ mit unendlich vielen Zeilen und Spalten verschiedene Ergebnisse liefern.

Satz 4.12. *Für alle natürlichen Zahlen m und n seien komplexe Zahlen $z_{m,n}$ gegeben. Es seien m_k und n_k Folgen natürlicher Zahlen, so dass die durch $k \mapsto (m_k, n_k)$ gegebene Abbildung $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$ bijektiv ist. Dann sind folgende Aussagen äquivalent:*

(i) *Die Reihe*
$$\sum_{k=0}^{\infty} z_{m_k,n_k}$$
ist absolut konvergent.

(ii) *Es gibt eine reelle Zahl d, so dass für alle natürlichen Zahlen M und N gilt*
$$\sum_{m=0}^{M} \sum_{n=0}^{N} |z_{m,n}| \leq d.$$

Sind diese Bedingungen erfüllt, so gilt

$$\sum_{k=0}^{\infty} z_{m_k,n_k} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} z_{m,n}.$$

Beweis. Wir betrachten die Partialsummen

$$\sum_{k=0}^{K} |z_{m_k,n_k}|, \qquad \sum_{m=0}^{M} \sum_{n=0}^{N} |z_{m,n}|.$$

Ist K gegeben, so ist die linke Summe eine Teilsumme der rechten für $M = \max\{m_k \mid k \leq K\}$ und $N = \max\{n_k \mid k \leq K\}$. Sind hingegen M und

N gegeben, so ist die rechte Summe eine Teilsumme der linken für $K = \max\{k \mid m_k \leq M, n_k \leq N\}$. Ist also die eine Menge von Partialsummen beschränkt, so auch die andere, und die Äquivalenz von (i) und (ii) ergibt sich mit Folgerung 4.2.

Nun seien (i) und (ii) erfüllt und $\varepsilon > 0$ gegeben. Nach dem Cauchykriterium gibt es ein k_0, so dass für $K \geq k_0$ gilt

$$\sum_{k=k_0+1}^{K} |z_{m_k,n_k}| < \varepsilon.$$

Für

$$M \geq \max\{m_k \mid k \leq k_0\}, \qquad N \geq \max\{n_k \mid k \leq k_0\},$$
$$K \geq \max\{k \mid m_k \leq M, n_k \leq N\}$$

folgt ähnlich wie im Beweis von Satz 4.3

$$\left| \sum_{n=0}^{N} \sum_{m=0}^{M} z_{m,n} - \sum_{k=0}^{K} z_{m_k,n_k} \right| \leq \sum_{k=k_0+1}^{K} |z_{m_k,n_k}| < \varepsilon.$$

Für festes M und N können wir zum Grenzwert $K \to \infty$ übergehen, und wegen der Abgeschlossenheit von $\{z \in K \,\big|\, |z| \leq \varepsilon\}$ erhalten wir

$$\left| \sum_{n=0}^{N} \sum_{m=0}^{M} z_{m,n} - \sum_{k=0}^{\infty} z_{m_k,n_k} \right| \leq \varepsilon.$$

Da die innere Reihe nach (ii) absolut konvergent ist, können wir auch zum Grenzwert $M \to \infty$ übergehen und erhalten

$$\left| \sum_{n=0}^{N} \sum_{m=0}^{\infty} z_{m,n} - \sum_{k=0}^{\infty} z_{m_k,n_k} \right| \leq \varepsilon.$$

Nun folgt die Behauptung mit der Grenzwertdefinition. □

Folgerung 4.9 (Doppelreihensatz). *Unter der Bedingung (i) oder (ii) gilt*

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} z_{m,n} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} z_{m,n}.$$

Wenden wir nämlich den Satz zusätzlich auf die Doppelreihe mit den Gliedern $w_{mn} = z_{nm}$ an, so sehen wir, dass die beiden Doppelreihen gleich zwei einfachen Reihen sind, die bei geeigneter Wahl der Bijektionen übereinstimmen.

Folgerung 4.10 (Reihenmultiplikation)**.** *Es gilt*

$$\sum_{m=0}^{\infty} x_m \sum_{n=0}^{\infty} y_n = \sum_{k=0}^{\infty} x_{m_k} y_{n_k},$$

falls die Reihen auf der linken Seite absolut konvergent sind.

Wegen

$$\sum_{m=0}^{M} \sum_{n=0}^{N} x_m y_n = \sum_{m=0}^{M} x_m \sum_{n=0}^{N} y_n$$

ist nämlich die Bedingung (ii) für $z_{m,n} = x_m y_n$ erfüllt.

Nun kommen wir zu den Anwendungen.

Satz 4.13. *Haben die Potenzreihen*

$$\sum_{m=0}^{\infty} a_m z^m, \qquad \sum_{n=0}^{\infty} b_n z^n$$

die Konvergenzradien r bzw. s, so gilt für $|z| < \min\{r, s\}$

$$\sum_{m=0}^{\infty} a_m z^m \sum_{n=0}^{\infty} b_n z^n = \sum_{l=0}^{\infty} c_l z^l,$$

wobei

$$c_l = \sum_{m=0}^{l} a_m b_{l-m} = \sum_{n=0}^{l} a_{l-n} b_n.$$

Beweis. Nach Folgerung 4.6 sind beide Reihen für die angegebenen Argumente z absolut konvergent, und Folgerung 4.10 ist anwendbar. Es sei $L_l = \{(m, n) \in \mathbb{N} \times \mathbb{N} \mid m + n = l\}$. Die Summe der Glieder der Doppelreihe über die Elemente von L_l ist $c_l z^l$. Wir lassen die Folge (m_k, n_k) der Reihe nach die Mengen L_0, L_1, L_2, … durchlaufen. Betrachten wir nur die Partialsummen nach Erledigung einer dieser Mengen, so erhalten wir eine Teilfolge, die gerade die Partialsummenfolge der rechten Seite in Satz 4.13 ist. □

Satz 4.14 (Funktionalgleichung)**.** *Für alle komplexen Zahlen z und w gilt*

$$\exp(z + w) = \exp z \cdot \exp w.$$

Beweis. Wir wenden Folgerung 4.10 an und berechnen die Teilsumme über L_l mit Hilfe des Binomialtheorems:

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

Satz 4.15. *Die Potenzreihe*

$$f(z) = \sum_{m=0}^{\infty} a_m z^m$$

habe den Konvergenzradius r, und es sei $c \in \mathbb{C}$, so dass $|c| < r$. Dann gilt für $|z - c| < r - |c|$

$$f(z) = \sum_{n=0}^{\infty} b_n (z - c)^n,$$

wobei

$$b_n = \sum_{m=n}^{\infty} \binom{m}{n} a_m c^{m-n}.$$

Beweis. Nach dem Binomialtheorem ist

$$f(z) = \sum_{m=0}^{\infty} a_m \big(c + (z - c)\big)^m = \sum_{m=0}^{\infty} \sum_{n=0}^{m} a_m \binom{m}{n} c^{m-n} (z - c)^n.$$

(Hier wie auch im Satz können wir alle Summen von 0 bis ∞ erstrecken.) Nach Assoziativgesetz und Binomialtheorem gilt

$$\begin{aligned} \sum_{m=0}^{N} \sum_{n=0}^{N} \left| a_m \binom{m}{n} c^{m-n} (z - c)^n \right| &\leq \sum_{m=0}^{N} |a_m| \sum_{n=0}^{m} \binom{m}{n} |c|^{m-n} |z - c|^n \\ &\leq \sum_{m=0}^{\infty} |a_m| \big(|c| + |z - c|\big)^m. \end{aligned}$$

Wegen $|c| + |z - c| < r$ ist die rechts stehende Reihe konvergent, also ist Bedingung (ii) von Satz 4.12 erfüllt und Folgerung 4.9 anwendbar. □

Definition 4.7. *Es sei $K = \mathbb{R}$ oder $K = \mathbb{C}$ und D eine offene Teilmenge von K. Eine Funktion $f : D \to K$ heißt* K-analytisch, *wenn es für jedes Element c von D ein $r > 0$ und eine Folge von Zahlen $a_n \in K$ gibt, so dass für $z \in D$ mit der Eigenschaft $|z - c| < r$ gilt*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - c)^n.$$

So ist z. B. die Funktion $f(z) = \frac{1}{z}$ auf $\mathbb{C} \setminus \{0\}$ analytisch, denn für $|z - c| < |c|$ gilt

$$\frac{1}{z} = \frac{1}{c} \cdot \frac{1}{1 - \frac{c-z}{c}} = \frac{1}{c} \sum_{n=0}^{\infty} \left(\frac{c - z}{c} \right)^n = \sum_{n=0}^{\infty} \frac{(-1)^n}{c^{n+1}} (z - c)^n.$$

Nach Satz 4.15 definiert jede Potenzreihe mit positivem Konvergenzradius eine analytische Funktion auf ihrem Konvergenzkreis. Nach Folgerung 4.8, Satz 4.13 und Aufgabe 50* sind die Summe, das Produkt und die Verkettung zweier analytischer Funktionen ebenfalls analytische Funktionen, und aus Satz 4.11 folgt: Ist $a \in D$ ein Häufungspunkt der Nullstellenmenge einer analytischen Funktion f auf D, dann gibt es ein $r > 0$, so dass $f(z) = 0$ für alle z mit $|z - a| < r$.

5 Stetige Funktionen

Vorlesung 11a
15.06.2022

Intuitiv unterscheidet man bei Funktionen zwischen sprunghafter und stetiger Änderung. Wir wollen diese Eigenschaft exakt beschreiben. Wir betrachten Funktionen $f : D \to L$ mit einem Definitionsbereich $D \subset K$, wobei K und L vollständige Körper mit Absolutbetrag sind. Für uns kommen dafür nur $\mathbb{R}$ und $\mathbb{C}$ in Frage.

5.1 Lipschitz-Stetigkeit

Wir beginnen mit einem einfacheren Begriff.

Definition 5.1. *Eine Funktion $f : D \to L$ heißt* Lipschitz-stetig, *wenn es eine Zahl $d > 0$ gibt, so dass für alle Elemente x und y von D gilt*

$$|f(x) - f(y)| \leq d|x - y|.$$

Satz 5.1. *Eine Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ mit dem Konvergenzradius r stellt auf der Menge $D = \{x \in K \,\big|\, |x| \leq c\}$ eine Lipschitz-stetige Funktion f dar, falls $c < r$.*

Beweis. Nach der verallgemeinerten dritten binomischen Formel und der Dreiecksungleichung gilt für x und y in D

$$|x^n - y^n| \leq |x - y| \sum_{k=1}^{n} |x|^{n-k}|y|^{k-1} \leq nc^{n-1}|x - y|,$$

und durch Summation erhalten wir

$$\left|\sum_{n=0}^{m} a_n(x^n - y^n)\right| \leq |x - y| \sum_{n=1}^{m} n|a_n|c^{n-1}.$$

Rechts stehen die Partialsummen einer Potenzreihe, die nach Satz 4.8 ebenfalls den Konvergenzradius r hat. Wir können diese durch ihre Reihensumme d ersetzen. Wegen der Abgeschlossenheit von $\{z \in K \,\big|\, |z| \leq d|x-y|\}$ erhalten wir

$$|f(x) - f(y)| \leq d|x - y|. \qquad \square$$

Es folgt, dass die Einschränkungen der ganzrationalen Funktionen und der Exponentialfunktion auf jede beschränkte Menge Lipschitz-stetig sind. Der Absolutbetrag ist wegen der Dreiecksungleichung Lipschitz-stetig. Die Signum-Funktion auf $\mathbb{R}$

$$\operatorname{sgn} x = \begin{cases} 1, & \text{falls } x > 0, \\ 0, & \text{falls } x = 0, \\ -1, & \text{falls } x < 0 \end{cases}$$

ist nicht Lipschitz-stetig, auch nicht die Quadratwurzelfunktion, denn die Ungleichung $\left|\sqrt{x} - \sqrt{0}\right| \leq d|x - 0|$ ist für $0 < x < d^{-2}$ verletzt.

5.2 Stetigkeit

Definition 5.2. *Es sei $D \subseteq K$ und $a \in D$. Eine Funktion $f : D \to L$ heißt* stetig *an der Stelle a, wenn es für jede positive reelle Zahl ε eine positive reelle Zahl δ gibt, so dass für alle $x \in D$ mit der Eigenschaft $|x - a| < \delta$ gilt $|f(x) - f(a)| < \varepsilon$.*

Eine Funktion heißt stetig, wenn sie an jeder Stelle ihres Definitionsbereiches stetig ist.

Offensichtlich ist jede Lipschitz-stetige Funktion stetig, denn man kann $\delta = \varepsilon/d$ wählen. Insbesondere sind alle ganzrationalen Funktionen und die Exponentialfunktion stetig. Die Signumfunktion ist an jeder Stelle mit Ausnahme von 0 stetig.

Beispiel. Es sei

$$f\left(\tfrac{m}{n}\right) = \tfrac{1}{n}$$

für beliebige teilerfremde Zahlen $m \in \mathbb{Z}$ und $n \in \mathbb{N} \setminus \{0\}$, und es sei $f(x) = 0$ für jede irrationale Zahl x.

Ist $\varepsilon = \frac{1}{n}$ und $\delta > 0$, so gibt es immer eine irrationale Zahl x, so dass $\left|x - \frac{m}{n}\right| < \delta$, aber $\left|f(x) - f\left(\frac{m}{n}\right)\right| = \varepsilon$. Somit ist f an allen rationalen Stellen unstetig.

Ist a irrational und $\varepsilon > 0$, so ist $N = \left\{n \in \mathbb{N} \setminus \{0\} \,\middle|\, \frac{1}{n} \geq \varepsilon\right\}$ endlich, also auch

$$M = \left\{\tfrac{m}{n} \,\middle|\, n \in N,\ \left|a - \tfrac{m}{n}\right| \leq 1\right\}.$$

Diese Menge ist nicht leer, und nach Aufgabe 16 existiert

$$\delta = \min\{|a-r| \mid r \in M\},$$

es ist $0 < \delta < 1$, und für $|x-a| < \delta$ ist $|f(x) - f(a)| < \varepsilon$. Somit ist f an allen irrationalen Stellen stetig. $\lhd$

Satz 5.2 (Folgenkriterium der Stetigkeit)**.** *Eine Funktion $f : D \to L$ ist genau dann an der Stelle a stetig, wenn für jede Folge x_n in D mit Grenzwert a die Folge $f(x_n)$ gegen $f(a)$ konvergiert.*

Beweis. Es sei f stetig in a und $x_n \to a$ $(n \to \infty)$. Ist $\varepsilon > 0$, so gibt es wegen der Stetigkeit ein $\delta > 0$, so dass für $|x-a| < \delta$ gilt $|f(x) - f(a)| < \varepsilon$, und wegen der Konvergenz der Folge gibt es ein n_0, so dass für $n \geq n_0$ gilt $|x_n - a| < \delta$. Für diese n gilt dann $|f(x_n) - f(a)| < \varepsilon$, und weil $\varepsilon > 0$ beliebig war, folgt $f(x_n) \to f(a)$ $(n \to \infty)$.

Nun sei f nicht stetig in a, d. h. es gibt ein $\varepsilon > 0$, so dass für jedes $\delta > 0$ ein $x \in D$ existiert, für das $|x-a| < \delta$, aber $|f(x) - f(a)| \geq \varepsilon$ gilt. Wir halten ε fest und wählen für jedes $\delta = \frac{1}{n}$ ein x_n wie oben. Dann gilt $|x_n - a| < \frac{1}{n}$, also $x_n \to a$ $(n \to \infty)$. Andererseits ist $|f(x_n) - f(a)| \geq \varepsilon$, also $f(x_n) \not\to f(a)$ $(n \to \infty)$. $\square$

Die Anwendung dieses Kriteriums auf den Lipschitz-stetigen Absolutbetrag kann im Beweis der Sätze 4.12 und 5.1 die Abgeschlossenheit ersetzen.

Sind Funktionen $f : D_1 \to L$ und $D_2 \to L$ gegeben, so definieren wir Funktionen $f+g$ und $f \cdot g$ mit Definitionsbereich $D_1 \cap D_2$ sowie eine Funktion $\frac{1}{f}$ mit Definitionsbereich $\{x \in D \mid f(x) \neq 0\}$ durch

$$(f+g)(x) = f(x) + g(x), \qquad (f \cdot g)(x) = f(x) \cdot g(x), \qquad \frac{1}{f}(x) = \frac{1}{f(x)}.$$

Satz 5.3. *Die Funktionen f und g seien an der Stelle a stetig. Dann sind auch die Funktionen $f+g$ und $f \cdot g$ an der Stelle a stetig. Ist außerdem $f(a) \neq 0$, so ist auch die Funktion $\frac{1}{f}$ an der Stelle a stetig.*

Beweis. Für jede Folge von Elementen x_n von D mit Grenzwert a gilt $f(x_n) \to f(a)$ und $g(x_n) \to g(a)$ nach Satz 5.2. Mit Satz 3.4 folgt

$$f(x_n) + g(x_n) \to f(a) + g(a), \qquad f(x_n) \cdot g(x_n) \to f(a) \cdot g(a) \quad (n \to \infty).$$

Ist $f(a) \neq 0$, so folgt mit Satz 3.5, dass

$$\frac{1}{f(x_n)} \to \frac{1}{f(a)} \quad (n \to \infty).$$

Nun folgt die Behauptung mit Hilfe der anderen Richtung von Satz 5.2. $\square$

Man hätte den Beweis auch direkt nach dem Muster der Beweise von Satz 3.4 und 3.5 führen können. Zur Abwechslung werden wir das beim nächsten Satz tun. Hier geht es um die Verkettung $f \circ g$ von Funktionen $f : D \to L$ und $g : E \to K$, wobei E in einem Körper M liegt. Die Verkettung existiert, wenn der Wertebereich von g in D liegt. (In der Analysis unterscheidet man nicht zwischen einer Funktion g und ihrer Beschränkung ${}_D|g$.)

Satz 5.4. *Die Funktion g sei stetig an der Stelle a, und die Funktion f sei stetig an der Stelle $g(a)$. Existiert die Verkettung $f \circ g$, so ist sie stetig an der Stelle a.*

Beweis. Es sei $\varepsilon > 0$. Wegen der Stetigkeit von f in $g(a)$ gibt es ein $\delta > 0$, so dass für $x \in D$ mit der Eigenschaft $|x - g(a)| < \delta$ gilt $|f(x) - f(g(a))| < \varepsilon$. Wegen der Stetigkeit von g in a gibt es ein $\eta > 0$, so dass für $u \in E$ mit $|u - a| < \eta$ gilt $|g(u) - g(a)| < \delta$. Für diese u folgt $|f \circ g(u) - f \circ g(a)| < \varepsilon$. $\square$

Satz 5.5. *Eine streng monotone Funktion auf einem Intervall, deren Wertebereich ebenfalls ein Intervall ist, ist stetig.*

Beweis. Es seien I und J Intervalle und $f : I \to J$ streng monoton wachsend und bijektiv. Wir beweisen die Stetigkeit von f an einer Stelle $a \in I$. Es sei $\varepsilon > 0$ gegeben.
<u>1. Fall:</u> Es gibt ein $c \in J$ mit $c < f(a)$. Setzen wir $d = \max\{c, f(a) - \varepsilon\}$, dann ist $d < f(a)$. Da J ein Intervall ist, folgt $d \in J$, und wegen der Surjektivität von f gibt es ein $b \in I$, so dass $f(b) = d$. Wegen der Monotonie folgt aus $f(b) < f(a)$, dass $b < a$, und für alle $x \in I$ gilt

$$b < x \leq a \quad \Rightarrow \quad f(a) - \varepsilon < f(x) \leq f(a).$$

<u>2. Fall:</u> $f(a) = \min J$. Dann ist wegen der strengen Monotonie $a = \min I$. Also erhalten wir das selbe Ergebnis wie im 1. Fall, wenn wir $b = a - 1$ setzen.

Analog findet man ein $b' > a$, so dass für alle $x \in I$ gilt

$$a \leq x < b' \quad \Rightarrow \quad f(a) \leq f(x) < f(a) + \varepsilon.$$

Setzen wir $\delta = \min\{a - b, b' - a\}$, so folgt für $x \in I$

$$a - \delta < x < a + \delta \quad \Rightarrow \quad f(a) - \varepsilon < f(x) < f(a) + \varepsilon. \qquad \square$$

Folgerung 5.1. *Für jede rationale Zahl r ist die Potenzfunktion $f(x) = x^r$ auf $\mathbb{R}_+$ stetig, und für $r \geq 0$ ist sie sogar auf $[0, \infty[$ stetig.*

Für $r \neq 0$ ist diese Funktion nämlich streng monoton, ihre Umkehrfunktion ist die Potenzfunktion mit dem Exponenten $\frac{1}{r}$, und beide sind nach Satz 2.14 auf dem Intervall $\mathbb{R}_+$ bzw. $[0, \infty[$ definiert.

5.3 Eigenschaften stetiger Funktionen

Vorlesung 11b
17.06.2022

Das Supremum des Wertebereichs einer Funktion f nennt man das Supremum der Funktion, abgekürzt $\sup f$, und $\sup_A f = \sup\bigl(f|_A\bigr)$. Gleiches gilt für das Infimum und, wenn sie existieren, das Maximum und das Minimum.

Satz 5.6. *Ist $D \neq \varnothing$ eine beschränkte abgeschlossene Teilmenge von $\mathbb{R}$ oder $\mathbb{C}$, so hat jede stetige Funktion $f : D \to \mathbb{R}$ ein Maximum und ein Minimum.*

Beweis. Es sei $d = \sup f$. Dann gibt es eine streng monoton wachsende Folge y_n, die den eigentlichen oder uneigentlichen Grenzwert d hat. Für jedes n gibt es wegen $y_n < d$ ein $x_n \in D$, so dass $y_n < f(x_n) \leq d$. Nach dem Satz von Bolzano-Weierstraß hat die Folge x_n einen Häufungspunkt a, der wegen der Abgeschlossenheit in D liegt. Nach Satz 3.13 konvergiert eine Teilfolge x_{n_k} gegen a, und nach Satz 5.2 gilt $f(x_{n_k}) \to f(a)$ $(k \to \infty)$. Die Teilfolge y_{n_k} hat den selben Grenzwert d wie y_n, und mit Satz 3.8 folgt $f(a) = d$.

Durch Anwendung des Bewiesenen auf die Funktion $-f$ erhalten wir die Aussage über das Minimum. □

Satz 5.7 (Zwischenwertsatz)**.** *Ist $f : I \to \mathbb{R}$ stetig, wobei I ein Intervall ist, und gilt für Elemente $a \leq b$ von I, dass $f(a) \leq d \leq f(b)$, dann gibt es ein Element $c \in [a, b]$, so dass $f(c) = d$.*

Beweis. Die Menge

$$M = \{x \in [a, b] \mid f(x) \leq d\}$$

enthält a und hat die obere Schranke b. Setzen wir $c = \sup M$, so gilt also $a \leq c \leq b$ und somit $c \in I$.

Wäre $f(c) > d$, so wäre $c \neq a$, und es gäbe ein $\delta > 0$, so dass für $x \in I$ mit $|x - c| < \delta$ gilt $f(x) > d$. All diese x wären obere Schranken von M (Widerspruch).

Wäre $f(c) < d$, so wäre $c \neq b$, und es gäbe ein $\delta > 0$, so dass für $x \in I$ mit $|x-c| < \delta$ gilt $f(x) < d$. All diese x würden zu M gehören (Widerspruch). □

Wenden wir den Satz auf die Funktion $-f$ an, so sehen wir, dass er auch unter der Vorausstzung $f(a) \geq d \geq f(b)$ gilt. Wir sehen also:

Folgerung 5.2. *Ist I ein Intervall und $f : I \to \mathbb{R}$ stetig, so ist der Wertebereich von f ein Intervall.*

Zusammen mit Satz 5.5 ergibt dies:

Folgerung 5.3. *Ist I ein Intervall und $f : I \to \mathbb{R}$ stetig und streng monoton mit dem Wertebereich J, so ist die Umkehrfunktion $g : J \to \mathbb{R}$ ebenfalls stetig und streng monoton.*

Nach Folgerung 3.6 ist die Exponentialfunktion streng monoton wachsend, aus Satz 3.10(i) und Folgerung 3.4 ergibt sich der Wertebereich $]0, \infty[$.

Folgerung 5.4. *Die Funktion* $\exp : \mathbb{R} \to \mathbb{R}$ *hat eine stetige streng monoton wachsende Umkehrfunktion*[28] $\ln :]0, \infty[\to \mathbb{R}$, *genannt* natürlicher Logarithmus. *Für alle* $s > 0$, $t > 0$ *gilt*

$$\ln(s \cdot t) = \ln s + \ln t, \qquad 1 - \frac{1}{s} \leq \ln s \leq s - 1.$$

Mit Hilfe der Substitution $x = \ln s$, $y = \ln t$ folgt diese Funktionalgleichung aus derjenigen in Satz 3.10(iii), und die Ungleichungen folgen aus Satz 3.10(i) und Folgerung 3.4 (Übungsaufgabe).

Definition 5.3. *Für* $x > 0$ *und* $s \in \mathbb{C}$ *setzen wir*

$$x^s = \exp(s \ln x).$$

Man überzeugt sich leicht, dass dies mit dem schon betrachteten Fall $s \in \mathbb{Q}$ verträglich ist und die Potenzgesetze weiterhin gültig sind. Insbesondere gilt $|x^s| = x^{\operatorname{Re} s}$. Die Reihe in Satz 4.7 ist also für $\operatorname{Re} s > 1$ konvergent. Ihre Reihensumme $\zeta(s)$ nennt man Riemannsche Zetafunktion, sie spielt eine wichtige Rolle im Beweis des Primzahlsatzes.

5.4 Winkelfunktionen

Wir betrachten die Exponentialfunktion mit komplexen Argumenten.

Definition 5.4. *Der* Sinus *und der* Kosinus *einer reellen Zahl* t *sind*

$$\sin t = \operatorname{Im} \exp it, \qquad \cos t = \operatorname{Re} \exp it.$$

Hieraus ergibt sich die Eulersche Formel

$$\exp it = \cos t + i \sin t.$$

Satz 5.8. (i) *Die Sinusfunktion ist ungerade, die Kosinusfunktion ist gerade, beide sind stetig, und es gelten die* Additionstheoreme

$$\begin{aligned} \sin(s+t) &= \sin s \, \cos t + \cos s \, \sin t, \\ \cos(s+t) &= \cos s \, \cos t - \sin s \, \sin t \end{aligned}$$

für alle reellen Zahlen s *und* t.

[28] ln ist die Abkürzung von *logarithmus naturalis.*

(ii) *Es gibt eine kleinste positive Zahl* π *mit der Eigenschaft* $\exp \pi i = -1$, *und es gilt*

$$\sin\left(t + \frac{\pi}{2}\right) = \cos t, \qquad \cos\left(t + \frac{\pi}{2}\right) = -\sin t.$$

(iii) *Jede komplexe Zahl* $z \neq 0$ *lässt sich eindeutig in der Form*

$$z = r \exp it$$

mit $r > 0$ *und* $t \in [0, 2\pi[$ *schreiben. (Man nennt* t *das* Argument *von* z, *abgekürzt* $\arg z$.*)*

(iv) *Die Sinusfunktion ist auf* $[-\frac{\pi}{2}, \frac{\pi}{2}]$ *streng monoton wachsend und die Kosinusfunktion auf* $[0, \pi]$ *streng monoton fallend.*

Beweis. Die Stetigkeit folgt aus den Sätzen 5.1 und 5.4. Aus Satz 4.14 folgt

$$\exp(-it) = \overline{\exp it}, \qquad \exp(is + it) = \exp is \cdot \exp it,$$

also

$$\begin{aligned} \cos(-t) + i\sin(-t) &= \cos t - i \sin t, \\ \cos(s+t) + i\sin(s+t) &= (\cos s + i \sin s)(\cos t + i \sin t), \end{aligned}$$

und Aussage (i) ergibt sich durch Vergleich von Real- und Imaginärteil. Nebenbei ergibt sich

$$|\exp it| = 1, \qquad \sin^2 t + \cos^2 t = 1. \tag{7}$$

Wegen

$$\operatorname{Re} z = \frac{z + \bar{z}}{2}, \qquad \operatorname{Im} z = \frac{z - \bar{z}}{2i}$$

folgt

$$\cos t = \frac{\exp it + \exp(-it)}{2}, \qquad \sin t = \frac{\exp it - \exp(-it)}{2i},$$

und wir erhalten aus Satz 4.10 die absolut konvergenten alternierenden Reihenentwicklungen

$$\cos t = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!}, \qquad \sin t = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!}.$$

Die Absolutbeträge der Glieder sind streng monoton fallend, falls die Quotienten

$$\left|\frac{t^{2k}}{(2k)!}\right| : \left|\frac{t^{2k-2}}{(2k-2)!}\right| = \frac{t^2}{2k(2k-1)}, \qquad \left|\frac{t^{2k+1}}{(2k+1)!}\right| : \left|\frac{t^{2k-1}}{(2k-1)!}\right| = \frac{t^2}{2k(2k+1)}$$

kleiner als 1 sind. Dies gilt ab $k = 1$ für $|t| < \sqrt{1 \cdot 2}$ bzw. $|t| < \sqrt{2 \cdot 3}$ und ab $k = 2$ für $|t| < \sqrt{3 \cdot 4}$ bzw. $|t| < \sqrt{4 \cdot 5}$. Es folgt

$$0 < \cos t < 1 \quad \text{für } |t| < \sqrt{2},$$
$$0 < \sin t < t \quad \text{für } 0 < t < \sqrt{6},$$
$$\cos 2 < 1 - \frac{2^2}{2!} + \frac{2^4}{4!} = -\frac{1}{3}.$$

Nach dem Zwischenwertsatz hat cos im Intervall $\left]\sqrt{2}, 2\right[$ eine Nullstelle. Wir bezeichnen das Doppelte der kleinsten positiven Nullstelle von cos mit π. Wegen $\sin \frac{\pi}{2} > 0$ folgt

$$\exp \frac{\pi i}{2} = i, \qquad \exp\left(it + \frac{\pi i}{2}\right) = i \exp it, \tag{8}$$

und Aussage (ii) ist bewiesen.

Aus den Additionstheoremen folgt

$$\cos(s + t) - \cos(s - t) = -2 \sin s \, \sin t$$

und, mit der Substitution $u = s + t$, $v = s - t$,

$$\cos u - \cos v = -2 \sin \frac{u + v}{2} \sin \frac{u - v}{2}.$$

Für $0 \leq v < u \leq 2$ ist dies negativ, also ist die Kosinusfunktion auf dem Intervall $[0, 2]$ streng monoton fallend und definiert eine bijektive Abbildung

$$\cos : \left[0, \frac{\pi}{2}\right[\to \left]0, 1\right].$$

Auf dem selben Intervall ist die Sinusfunktion wegen (7) streng monoton wachsend, und zusammen mit Teil (ii) folgt Aussage (iv).

Die Werte der Funktion $f(t) = \exp it$ eingeschränkt auf $\left[0, \frac{\pi}{2}\right[$ liegen auf dem Viertelkreisbogen

$$V = \{x + iy \in \mathbb{C} \mid x^2 + y^2 = 1,\ x > 0,\ y \geq 0\},$$

und die Abbildung $\mathrm{Re} : V \to \left]0, 1\right]$ ist umkehrbar. Wegen $\mathrm{Re} \circ f = \cos$ definiert f eine bijektive Abbildung

$$\left[0, \frac{\pi}{2}\right[\to V,$$

und mit Hilfe von (8) sehen wir, dass f auch eine bijektive Abbildung

$$[0, 2\pi[\to \{z \in \mathbb{C} \,\big|\, |z| = 1\}$$

definiert. Ist $z \neq 0$ eine komplexe Zahl und setzen wir $r = |z|$, so liegt $\frac{z}{r}$ im Wertebereich dieser Abbildung, und Aussage (iii) folgt. $\square$

Aus dem Beweis folgt auch, dass für alle t gilt

$$|\sin t| \leq |t|.$$

Mit der Identität $\exp nz = (\exp z)^n$, die aus Satz 4.14 durch vollständige Induktion folgt, ergibt sich[29] aus der Eulerschen Formel und dem Binomialtheorem durch Vergleich der Real- und Imaginärteile

$$\sin nt = \sum_{\substack{k=1 \\ k \text{ ungerade}}}^{n} \binom{n}{k} i^{k-1} \cos^{n-k} t \, \sin^k t,$$

$$\cos nt = \sum_{\substack{k=0 \\ k \text{ gerade}}}^{n} \binom{n}{k} i^{k} \cos^{n-k} t \, \sin^k t.$$

Definition 5.5. *Der* Tangens *und der* Kotangens *einer reellen Zahl t sind definiert als*

$$\tan t = \frac{\sin t}{\cos t}, \qquad \cot t = \frac{\cos t}{\sin t},$$

falls der Nenner jeweils von Null verschieden ist.

Folgerung 5.5. *Es gilt*

$$\tan\left(t + \frac{\pi}{2}\right) = -\cot t, \qquad \cot\left(t + \frac{\pi}{2}\right) = -\tan t.$$

Die Tangensfunktion ist auf dem Intervall $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ streng monoton wachsend, die Kotangensfunktion auf $]0, \pi[$ streng monoton fallend. Beide sind stetig auf ihrem Definitionsbereich, ungerade und haben den Wertebereich $\mathbb{R}$.

Auf $\left[0, \frac{\pi}{2}\right[$ ist die Sinusfunktion nämlich streng monoton wachsend und die Kosinusfunktion streng monoton fallend, also nach den Monotoniegesetzen die Tangensfunktion streng monoton wachsend und ihr Kehrwert streng monoton fallend. Nach dem Zwischenwertsatz gibt es für jedes $d > 0$ ein $x \in \left]1, \frac{\pi}{2}\right[$, so dass $\cos x < \frac{1}{d}$, also $\tan x > d \sin 1$, und somit ist die Tangensfunktion nicht von oben beschränkt. Der Rest ist offensichtlich.

Definition 5.6. *Die Umkehrfunktionen der* Winkelfunktionen

$$\sin\Big|_{\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]}, \qquad \cos\Big|_{[0,\pi]},$$

$$\tan\Big|_{\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[}, \qquad \cot\Big|_{]0,\pi[}$$

[29] Traditionell steht $\sin^k t$ für $(\sin t)^k$ und nicht für die k-fache Verkettung der Sinusfunktion angewendet auf t.

bezeichnet man als zyklometrische Funktionen

$$\arcsin, \quad \arccos : [-1,1] \to \mathbb{R}, \qquad \arctan, \quad \operatorname{arccot} : \mathbb{R} \to \mathbb{R},$$

genannt Arkussinus, Arkuskosinus, Arkustangens *und* Arkuskotangens[30].

5.5 Grenzwerte von Funktionen

Vorlesung 12a
22.06.2022

Die Funktion

$$f(x) = \frac{x^2 - 1}{x - 1}$$

ist an der Stelle 1 nicht definiert. Allerdings stimmt sie an allen anderen Stellen mit der stetigen Funktion

$$g(x) = x + 1$$

überein. Dieses Phänomen erfassen wir mit dem folgenden Begriff. Dabei seien wieder K und L vollständige Körper mit Absolutbetrag.

Definition 5.7. *Es sei $a \in K$ ein Häufungspunkt der Teilmenge $D \subseteq K$ und $f : D \to L$. Wir sagen, die Funktion $f(x)$* konvergiere *gegen die Zahl c für x gegen a, abgekürzt*[31]

$$f(x) \to c \ \ (x \to a),$$

wenn die Funktion

$$g(x) = \begin{cases} f(x), & \text{wenn } x \in D \setminus \{a\}, \\ c, & \text{wenn } x = a \end{cases}$$

an der Stelle a stetig ist.

Beispiel. Es sei $f : \mathbb{C} \setminus \{0\} \to \mathbb{C}$ definiert durch

$$f(z) = \frac{\exp z - 1}{z}.$$

Aus Folgerung 4.1 ergibt sich

$$f(z) = \sum_{n=1}^{\infty} \frac{z^{n-1}}{n!} = \sum_{k=0}^{\infty} \frac{z^k}{(k+1)!},$$

[30] ursprünglich *arcus sinus, arcus cosinus, arcus tangentis, arcus cotangentis.*

[31] An Stelle von x kann auch eine andere Variable stehen.

wobei die entstehende Potenzreihe z. B. nach dem Vergleichskriterium für alle $z \in \mathbb{C}$ absolut konvergiert, also nach Satz 5.1 auf ganz $\mathbb{C}$ eine stetige Funktion g darstellt. Es folgt

$$\frac{\exp z - 1}{z} \to 1 \quad (z \to 0).$$

Dies gilt natürlich auch für die Einschränkung auf die Menge $\mathbb{R} \setminus \{0\}$. $\triangleleft$

Beispiel. Es sei $K = \mathbb{R}$ oder $K = \mathbb{C}$, $q \in K$ und $n \in \mathbb{N}$. Dann ist die Funktion

$$f(q) = \frac{1 - q^n}{1 - q}$$

für $q \neq 1$ definiert. Aus der Summenformel der geometrischen Folge erhalten wir die alternative Darstellung

$$f(q) = 1 + q + \ldots + q^{n-1},$$

die für alle q definiert und nach Satz 5.3 an der Stelle 0 stetig ist. Somit gilt $f(q) \to n \ \ (q \to 1)$. $\triangleleft$

Satz 5.9. *Folgende Aussagen über eine Funktion $f : D \to L$ sind äquivalent.*

(i) $f(x) \to c \ \ (x \to a)$,

(ii) *Für jede Folge $(x_n)_{n \in \mathbb{N}}$ in $D \backslash \{a\}$ mit der Eigenschaft $x_n \to a \ \ (n \to \infty)$ gilt $f(x_n) \to c \ \ (n \to \infty)$,*

(iii) *Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in D \setminus \{a\}$ gilt*

$$|x - a| < \delta \quad \Rightarrow \quad |f(x) - c| < \varepsilon.$$

Gilt $f(x) \to b$ und $f(x) \to c \ \ (x \to a)$, so ist $b = c$.

Beweis. In den Aussagen (ii) und (iii) können wir f durch die Funktion g aus der Definition sowie die Zahl c durch $g(a)$ ersetzen. Nun drückt Aussage (iii), in der wir $x = a$ nicht mehr auszuschließen brauchen, genau die Stetigkeit von g an der Stelle a aus. Mit Satz 5.2 folgt also, dass (i) $\Leftrightarrow$ (iii) $\Rightarrow$ (ii).

Angenommen, es gilt (ii), und in D ist eine Folge x_n gegeben, die gegen a konvergiert. Wenn unendlich viele Glieder von a verschieden sind, so bilden sie eine Teilfolge x_{n_k} in $D \setminus \{a\}$, für die nach (ii) gilt $g(x_{n_k}) \to c \ \ (k \to \infty)$. Natürlich folgt dann $g(x_n) \to c \ \ (n \to \infty)$, und letzteres gilt auch, wenn nur endlich viele Glieder von a verschieden sind. Nach Satz 5.2 ist dann g stetig an der Stelle a.

Angenommen, es gilt $f(x) \to b$ und $f(x) \to c$ $(x \to a)$, wobei $b \neq c$. Wir setzen $\varepsilon = \frac{|b-c|}{2}$. Dann gibt es ein $\delta > 0$, so dass für $x \in D \setminus \{a\}$ mit $|x-a| < \delta$ gilt

$$|b - f(x)| < \varepsilon, \qquad |f(x) - c| < \varepsilon.$$

Da a ein Häufungspunkt von D ist, gibt es solche x tatsächlich, und mit der Dreiecksungleichung folgt $|b-c| < 2\varepsilon$ (Widerspruch). □

Häufig wird Eigenschaft (iii) als Definition angegeben. Wegen der Bemerkungen am Ende von Abschnitt 2.5 erhalten wir:

Folgerung 5.6. *Für eine Funktion f mit dem Zielbereich $\mathbb{C}$ gilt genau dann $f(x) \to c$ $(x \to a)$, wenn $\operatorname{Re} f(x) \to \operatorname{Re} c$ und $\operatorname{Im} f(x) \to \operatorname{Im} c$ $(x \to a)$.*

Beispiel. Aus

$$\frac{\exp it - 1}{it} \to 1 \quad (t \to 0)$$

erhalten wir

$$\frac{\sin t}{t} \to 1, \quad \frac{\cos t - 1}{t} \to 0 \quad (t \to 0),$$

durch Einsetzen der Eulerschen Formel. ◁

Wegen der letzten Aussage von Satz 5.9 ist folgender Begriff sinnvoll.

Definition 5.8. *Gilt $f(x) \to c$ $(x \to a)$, so nennen wir c den* Grenzwert der Funktion *f an der Stelle a, abgekürzt*

$$\lim_{x \to a} f(x).$$

Viele Aussagen über Grenzwerte von Folgen haben Analoga für Grenzwerte von Funktionen.

Satz 5.10 (Einschließungskriterium). *Es seien f, g und h Funktionen auf D mit Werten in $\mathbb{R}$, so dass für alle $x \in D \setminus \{a\}$ gilt*

$$f(x) \leq g(x) \leq h(x).$$

Wenn $f(x) \to c$ und $h(x) \to c$ $(x \to a)$, so auch $g(x) \to c$ $(x \to a)$.

Beweis. Dies folgt vermittels Satz 5.9 aus Satz 3.8 oder kann direkt wie jener bewiesen werden. □

Beispiel. Aus Folgerung 5.4 erhalten wir, falls $s > 0$ und $s \neq 1$,

$$\min\{1, \tfrac{1}{s}\} \leq \frac{\ln s}{s-1} \leq \max\{1, \tfrac{1}{s}\},$$

und es folgt

$$\frac{\ln s}{s-1} \to 1 \quad (s \to 1)$$

mit Hilfe von Aufgabe 52. $\triangleleft$

Mit Hilfe von Satz 5.9 erhalten wir aus den Sätzen 3.4, 3.5 und 5.4:

Folgerung 5.7. *Es seien $f,\, g : D \to L$ und a ein Häufungspunkt von D. Wenn*

$$f(x) \to b, \quad g(x) \to c \quad (x \to a),$$

dann gilt

$$f(x) + g(x) \to b + c, \quad f(x) \cdot g(x) \to b \cdot c \quad (x \to a).$$

Ist außerdem $b \neq 0$, so folgt auch

$$\frac{1}{f(x)} \to \frac{1}{b} \quad (x \to a).$$

Ist außerdem $E \subseteq L$ und $h : E \to M$ stetig an der Stelle b, dann folgt

$$h(f(x)) \to h(b) \quad (x \to a),$$

falls die Verkettung $h \circ f$ existiert.

Beispiel. Nach den Potenzgesetzen gilt

$$(1+x)^{\frac{1}{x}} = \big(\exp \ln(1+x)\big)^{\frac{1}{x}} = \exp \frac{\ln(1+x)}{x}.$$

Nun folgt

$$(1+x)^{\frac{1}{x}} \to \exp 1 = e \quad (x \to 0)$$

wegen der Stetigkeit der Exponentialfunktion. $\triangleleft$

Eine weitere Spielart von Grenzwerten ist die Folgende.

Definition 5.9. *Es sei $f : D \to L$, wobei $D \subseteq \mathbb{R}$ von oben unbeschränkt ist. Wir schreiben*

$$f(x) \to c \quad (x \to \infty),$$

falls es für jedes $\varepsilon > 0$ ein $x_0 \in \mathbb{R}$ gibt, so dass für alle $x \in D$ mit der Eigenschaft $x \geq x_0$ gilt

$$|f(x) - c| < \varepsilon.$$

Analog definiert man die Aussage $f(x) \to c$ $(x \to -\infty)$.

Wie oben sieht man, dass es nur eine solche Zahl c geben kann, und man bezeichnet sie mit

$$\lim_{x\to\infty} f(x).$$

Wenn $f(x) \to c$ $(x \to \infty)$, so hat offensichtlich auch die Folge $y_n = f(n)$ den Grenzwert c, aber die Umkehrung braucht nicht zu gelten, wie das Beispiel $f(x) = \sin(\pi x)$ zeigt. Man müsste schon wie in Satz 5.9 die Konvergenz von $f(x_n)$ für alle Folgen x_n mit dem uneigentlichen Grenzwert ∞ verlangen. Allerdings gilt:

Folgerung 5.8. *Für q, s, $t \in \mathbb{R}$ mit $0 < q < 1$ und $t > 0$ gilt*

$$x^s q^x \to 0 \quad (x \to \infty), \qquad \frac{(\ln y)^s}{y^t} \to 0 \quad (y \to \infty), \qquad z^t\,|\ln z|^s \to 0 \quad (z \to 0).$$

Da nämlich x^s für $x \geq 1$ monoton in s wächst, genügt es wegen des Einschließungskriteriums, dies für $s \in \mathbb{N}$ nachzuprüfen. Ist nun n die kleinste natürliche Zahl mit der Eigenschaft $n \geq x$, so gilt

$$0 < x^s q^x \leq n^s q^{n-1},$$

und nach Satz 3.3 gibt es für jedes $\varepsilon > 0$ ein n_0, so dass die rechte Seite für $n \geq n_0$ kleiner als ε ist. Die erste Behauptung folgt, wenn wir im Sinne der Definition $x_0 = n_0$ setzen. Schreiben wir $y = \exp x$, so ist $y^{-t} = q^x$ mit $q = \exp(-t) < 1$, und die zweite Behauptung folgt. Die Substitution $z = \frac{1}{y}$ zeigt schließlich die dritte Behauptung.

Man kann auch uneigentliche Grenzwerte von Funktionen einführen:

Definition 5.10. *Es sei $f : D \to \mathbb{R}$ und a ein Häufungspunkt von D. Wir sagen, die Funktion f divergiere gegen ∞ für x gegen a, abgekürzt*

$$f(x) \to \infty \ \ (x \to a),$$

wenn es für jedes $c \in \mathbb{R}$ ein $\delta > 0$ gibt, so dass für alle $x \in D$ mit $|x - a| < \delta$ gilt $f(x) > c$. Analog definiert man die Aussagen $f(x) \to -\infty$ für $x \to a$ bzw. für $x \to \infty$.

Es gelten ähnliche Rechenregeln wie in Satz 3.11. Daraus erhalten wir

Folgerung 5.9. *Für q, s, $t \in \mathbb{R}$ mit $q > 1$ und $t > 0$ gilt*

$$x^s q^x \to \infty \ \ (x \to \infty), \qquad y^t (\ln y)^s \to \infty \ \ (y \to \infty), \qquad \frac{|\ln z|^s}{z^t} \to \infty \ \ (z \to 0).$$

Die Funktion $\operatorname{sgn} x$ hat an der Stelle 0 keinen Grenzwert, wohl aber einseitige Grenzwerte in dem folgenden Sinne.

Definition 5.11. *Es sei $f : D \to \mathbb{R}$, wobei $D \subseteq \mathbb{R}$, und $a \in \mathbb{R}$. Wir setzen*

$$\begin{aligned} D_{<a} &= D \cap \,]-\infty, a[\,, \qquad & D_{>a} &= D \cap \,]a, \infty[\,, \\ f_{<a} &= f|_{D_{<a}}, & f_{>a} &= f|_{D_{>a}}, \end{aligned}$$

und definieren den linksseitigen bzw. rechtsseitigen Grenzwert als

$$\begin{aligned} f(x) \to c \ \ (x \to a - 0), &\quad \textit{wenn} \quad f_{<a}(x) \to c \ \ (x \to a), \\ f(x) \to c \ \ (x \to a + 0), &\quad \textit{wenn} \quad f_{>a}(x) \to c \ \ (x \to a) \end{aligned}$$

(wozu a ein Häufungspunkt der jeweiligen Menge sein muss).

Wenn a Häufungspunkt von $D_{<a}$ und $D_{>a}$ ist, so existiert $\lim\limits_{x\to a} f(x)$ offensichtlich genau dann, wenn die beiden einseitigen Grenzwerte existieren und gleich sind. Außerdem ist

$$\lim_{x\to\pm 0} f(x) = \lim_{x\to\pm\infty} f\left(\frac{1}{x}\right).$$

So gilt z. B.

$$\lim_{x\to\pm 0} \arctan\frac{1}{x} = \pm\frac{\pi}{2}.$$

Die obigen Existenzkriterien setzen voraus, dass man den potentiellen Grenzwert c bereits kennt. Nicht so das folgende: Vorlesung 12b 24.06.2022

Satz 5.11 (Monotoniekriterium)**.** *Es sei $f : D \to \mathbb{R}$ monoton und beschränkt, und a sei Häufungspunkt von $D_{<a}$ bzw. $D_{>a}$. Dann existiert $\lim\limits_{x\to a-0} f(x)$ bzw. $\lim\limits_{x\to a+0} f(x)$.*

Das Beispiel der Signum-Funktion zeigt, dass es keine analoge Aussage über zweiseitige Grenzwerte gibt.

Beweis. Es sei z. B. f monoton wachsend und $c = \sup f_{<a}$, was wegen der Vollständigkeit von L existiert. Wegen der Beschränktheit ist $c < \infty$, wegen $D_{<a} \neq \varnothing$ ist $c > -\infty$. Ist nun $\varepsilon > 0$, so gibt es ein $b \in D_{<a}$, so dass $f(b) > c - \varepsilon$, und wegen der Monotonie folgt $c - \varepsilon < f(x) \leq c < c + \varepsilon$ für $b < x < a$. □

Für das folgende Kriterium braucht f nicht monoton und K kein angeordneter Körper zu sein.

Satz 5.12 (Cauchy-Kriterium)**.** *Es sei a ein Häufungspunkt von $D \subseteq K$. Eine Funktion $f : D \to L$ hat genau dann eine Grenzwert an der Stelle a, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle x und y in $D \setminus \{a\}$ gilt:*

$$|x - a| < \delta \quad \wedge \quad |y - a| < \delta \quad \Longrightarrow \quad |f(x) - f(y)| < \varepsilon.$$

Beweis. Gilt $f(x) \to c \ (x \to a)$, so folgt die Cauchy-Bedingung aus

$$|f(x) - f(y)| \leq |f(x) - c| + |c - f(y)|.$$

Umgekehrt sei die Cauchy-Bedingung erfüllt. Es seien x_n und y_n Folgen in $D \setminus \{a\}$, so dass $x_n \to a$ und $y_n \to a \ (n \to \infty)$, und es sei $\varepsilon > 0$ gegeben. Dann liefert die besagte Bedingung ein δ, und zu diesem gibt es ein n_0, so dass für $m \geq n_0$ und $n \geq n_0$ gilt $|x_m - a| < \delta$ und $|y_n - a| < \delta$. Für diese m und n folgt also $|f(x_m) - f(y_n)| < \varepsilon$. Wir können das im Fall $y_n = x_n$ anwenden und sehen, dass $f(x_n)$ eine Cauchy-Folge ist, die wegen der Vollständigkeit von L einen Grenzwert c hat. Andererseits sehen wir, dass für eine zweite Folge y_n wie oben die Folge $f(x_n) - f(y_n)$ eine Nullfolge ist, so dass auch $f(y_n) \to c \ (n \to \infty)$. Nach Satz 5.9 folgt dann, dass $f(x) \to c \ (x \to a)$. □

5.6 Gleichmäßige Stetigkeit

Das Cauchy-Kriterium gibt an, wann sich eine Funktion $f : D \to L$ in einem Häufungspunkt von D definieren lässt, so dass die entstehende Funktion dort stetig ist. Wir bezeichnen die Menge der Häufungspunkte mit D' und nennen $\bar{D} = D \cup D'$ den *Abschluss* von D. Ist es möglich, f zu einer Funktion g auf $\bar{D}$ fortzusetzen, die in jedem Punkt stetig ist? Dazu ist offensichtlich notwendig, dass f selbst stetig ist, aber das genügt im Allgemeinen nicht.

Definition 5.12. *Eine Funktion $f : D \to L$ heißt* gleichmäßig stetig*, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle x, y in D gilt*

$$|x - y| < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon.$$

Offenbar ist jede Lipschitz-stetige Funktion gleichmäßig stetig und jede gleichmäßig stetige Funktion stetig.

Beispiel. Die Funktion $f(x) = \frac{1}{x}$ ist auf $D = \,]0, \infty]$ stetig, aber nicht gleichmäßig stetig. Ist $x = 2y > 0$, so ist $|x - y| = \frac{x}{2}$ und $|f(x) - f(y)| = \frac{1}{x}$. Geben wir $\varepsilon = 1$ vor, so gibt es kein geeignetes δ, denn wir finden immer $x \in \,]0, 2\delta[$, so dass $\frac{1}{x} \geq 1$. ◁

Beispiel. Die Funktion $f(x) = \sqrt{x}$ ist auf $[0, \infty[$ gleichmäßig stetig (aber bekanntlich nicht Lipschitz-stetig): Ist $\varepsilon > 0$, so gilt

$$|x - y| < \varepsilon^2 \quad \Rightarrow \quad |\sqrt{x} - \sqrt{y}| < \varepsilon,$$

denn

$$|\sqrt{x} - \sqrt{y}|^2 \leq |\sqrt{x} + \sqrt{y}||\sqrt{x} - \sqrt{y}| = |x - y|.$$

Satz 5.13 (Heine). *Ist die Funktion $f : D \to L$ stetig und der Definitionsbereich D abgeschlossen und beschränkt, so ist f gleichmäßig stetig.*

Beweis. Angenommen, f ist nicht gleichmäßig stetig. Dann gibt es ein $\varepsilon > 0$, so dass für jedes $\delta > 0$ Elemente x, $y \in D$ existieren, so dass $|x - y| < \delta$ und $|f(x) - f(y)| \geq \varepsilon$. Für $\delta = \frac{1}{n}$ wählen wir solche Elemente x_n und y_n. Auf Grund der Beschränktheit von D hat die Folge x_n nach Satz 3.6 einen Häufungspunkt a, der wegen der Abgeschlossenheit in D liegt. Nach Satz 3.13 gibt es also eine Teilfolge x_{n_k}, die gegen a konvergiert. Wegen $|x_n - y_n| < \frac{1}{n}$ ist $x_n - y_n$ eine Nullfolge, also auch $y_{n_k} \to a \ \ (k \to \infty)$. Da f stetig ist, folgt

$$f(x_{n_k}) \to f(a), \qquad f(y_{n_k}) \to f(a) \qquad (k \to \infty),$$

so dass

$$f(x_{n_k}) - f(y_{n_k}) \to 0 \qquad (k \to \infty)$$

(Widerspruch). □

Satz 5.14. *Ist $f : D \to L$ gleichmäßig stetig, so gibt es genau eine stetige Fortsetzung $g : \bar{D} \to L$, und diese ist ebenfalls gleichmäßig stetig.*

Beweis. Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle Elemente x, y von D gilt

$$|x - y| < 3\delta \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon.$$

Ist nun $a \in \bar{D}$, so folgt für alle x und y mit $|x - a| < \delta$ und $|y - a| < \delta$, dass $|x - y| < 2\delta$ und somit $f(x) - f(y)| < \varepsilon$. Ist also a ein Häufungspunkt, so existiert $\lim\limits_{x \to a} f(x)$ nach dem Cauchykriterium, was gleich $g(a)$ sein muss, andernfalls ist $g(a) = f(a)$.

Es bleibt die gleichmäßige Stetigkeit von g zu zeigen. Zu gegebenem $\varepsilon > 0$ sei $\delta > 0$ wie oben bestimmt. Sind a, b Elemente von $\bar{D}$, so gibt es nach Definition von g ein $\eta > 0$, so dass für alle x und y in D gilt

$$\begin{aligned} |x - a| < \eta \quad &\Rightarrow \quad |f(x) - g(a)| < \varepsilon, \\ |y - b| < \eta \quad &\Rightarrow \quad |f(y) - g(b)| < \varepsilon, \end{aligned}$$

wobei wir annehmen können, dass $\eta \leq \delta$. Wegen a, $b \in \bar{D}$ gibt es solche x und y tatsächlich, und für $|a - b| < \delta$ gilt

$$|x - y| \leq |x - a| + |a - b| + |b - y| < \eta + \delta + \eta \leq 3\delta,$$

also

$$|g(a) - g(b)| \leq |g(a) - f(x)| + |f(x) - f(y)| + |f(y) - g(b)| < 3\varepsilon. \qquad \square$$

Für beschränktes D hat eine Funktion $f : D \to L$ also genau dann eine stetige Fortsetzung auf $\bar{D}$, wenn sie gleichmäßig stetig ist. Für unbeschränktes D gilt das nicht, wie das Beispiel $f(x) = x^2$ auf $D = \,]0, \infty[$ zeigt.

6 Differentialrechnung

6.1 Definition und Berechnung der Ableitung

Wir betrachten einen vollständigen Körper K mit Absolutbetrag. Für uns kommen wieder nur $\mathbb{R}$ und $\mathbb{C}$ in Frage.

Definition 6.1. *Es sei $D \subseteq K$ und $a \in D$ ein Häufungspunkt von D. Eine Funktion $f : D \to K$ heißt* differenzierbar *an der Stelle a, wenn der so genannte* Differenzenquotient

$$\frac{f(x) - f(a)}{x - a}$$

einen Grenzwert für $x \to a$ besitzt. Dieser heißt dann Ableitung *der Funktion f an der Stelle a, abgekürzt*[32] $f'(a)$.

Dies bedeutet laut Definition 5.7, dass sich der auf $D \setminus \{a\}$ definierte Differenzenquotient zu einer Funktion $\tilde{f}$ auf D fortsetzen lässt, die an der Stelle a stetig ist. Wir erhalten somit folgende Umformulierung:

Eine Funktion $f : D \to K$ ist genau dann an der Stelle a differenzierbar, wenn es eine an der Stelle a stetige Funktion $\tilde{f}$ auf D gibt, so dass für $x \in D$ gilt

$$f(x) = f(a) + \tilde{f}(a)(x - a),$$

und dann ist $f'(a) = \tilde{f}(a)$.

Aus Satz 5.3 ergibt sich:

Folgerung 6.1. *Jede an einer Stelle differenzierbare Funktion ist dort stetig.*

Beispiel. Für $n \in \mathbb{N}$ gilt

$$\frac{x^n - a^n}{x - a} = \sum_{k=1}^{n} x^{n-k} a^{k-1} \to n a^{n-1} \quad (x \to a),$$

also hat die Funktion $f(x) = x^n$ an der Stelle a die Ableitung $f'(a) = na^{n-1}$. Des Weiteren ist für $a \neq 0$

$$\frac{\frac{1}{x^n} - \frac{1}{a^n}}{x - a} = -\frac{1}{a^n x^n} \cdot \frac{x^n - a^n}{x - a} \to -na^{-n-1} \quad (x \to a),$$

also gilt die obige Aussage sogar für $n \in \mathbb{Z}$. $\triangleleft$

[32] und gelesen „f Strich von a"

Beispiel. Nun sei $K = \mathbb{R}$, $a \in D = [0, \infty[$. Für $a > 0$ gilt

$$\frac{\sqrt{x} - \sqrt{a}}{x - a} = \frac{1}{\sqrt{x} + \sqrt{a}} \to \frac{1}{2\sqrt{a}} \quad (x \to a),$$

also gilt obige Formel sogar für $n = \frac{1}{2}$. An der Stelle $a = 0$ existiert die Ableitung nicht. $\triangleleft$

Beispiel. Für $a > 0$ gilt nach Folgerung 5.4 und dem Beispiel nach Satz 5.10

$$\frac{\ln x - \ln a}{x - a} = \frac{1}{a} \cdot \frac{\ln \frac{x}{a}}{\frac{x}{a} - 1} \to \frac{1}{a} \quad (x \to a)$$

also ist $\ln' a = \frac{1}{a}$. $\triangleleft$

Beispiel. Aus Satz 4.14 und dem Bespiel nach Definition 5.7 folgt, wenn wir $x = a + h$ schreiben,

$$\frac{\exp(a + h) - \exp a}{h} = \exp a \cdot \frac{\exp h - 1}{h} \to \exp a \quad (h \to 0),$$

also ist $\exp' = \exp$. Dies gilt für $K = \mathbb{R}$ und auch $K = \mathbb{C}$. $\triangleleft$

Beispiel. Aus Satz 5.8(i) und dem Beispiel nach Folgerung 5.6 ergibt sich

$$\begin{aligned}
\frac{\sin(a + h) - \sin a}{h} &= \frac{\sin a\,(\cos h - 1) + \cos a\,\sin h}{h} \to \cos a \quad (h \to 0),\\
\frac{\cos(a + h) - \cos a}{h} &= \frac{\cos a\,(\cos h - 1) - \sin a\,\sin h}{h} \to -\sin a \quad (h \to 0),
\end{aligned}$$

also ist $\sin' = \cos$ und $\cos' = -\sin$. $\triangleleft$

Satz 6.1. *Sind die Funktionen f und g an der Stelle a differenzierbar, so auch $f + g$, $f \cdot g$ und, falls $f(a) \neq 0$, die Funktion $\frac{1}{f}$. Es gilt*

$$\begin{aligned}
(f + g)'(a) &= f'(a) + g'(a) && \textit{(Summenregel)},\\
(f \cdot g)'(a) &= f'(a)g(a) + f(a)g'(a) && \textit{(Produktregel}^{33}\textit{)},\\
\left(\frac{1}{f}\right)'(a) &= -\frac{f'(a)}{f(a)^2}.
\end{aligned}$$

Beweis. Schreiben wir

$$f(x) = f(a) + \tilde{f}(x)(x - a), \qquad g(x) = g(a) + \tilde{g}(x)(x - a)$$

[33] auch Leibniz-Regel genannt

wie in der alternativen Definition, so folgt

$$\begin{aligned} f(x) + g(x) &= f(a) + g(a) + \big(\tilde{f}(x) + \tilde{g}(x)\big)(x - a), \\ f(x) \cdot g(x) &= f(a) \cdot g(a) \\ &\qquad + \big(\tilde{f}(x)g(a) + f(a)\tilde{g}(x) + \tilde{f}(x)\tilde{g}(x)(x - a)\big)(x - a). \end{aligned}$$

Im Fall $f(a) \neq 0$ ist wegen der Stetigkeit im Punkt a auch $f(x) \neq 0$ für x in einer Umgebung von a, und dort gilt

$$\frac{1}{f(x)} - \frac{1}{f(a)} = -\frac{f(x) - f(a)}{f(x)f(a)} = -\frac{\tilde{f}(x)}{f(x)f(a)} \cdot (x - a).$$

Nun folgen die Behauptungen mit der alternativen Definition. $\square$

Folgerung 6.2. *In der Situation des Satzes gilt, falls $g(a) \neq 0$,*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2} \qquad \textit{(Quotientenregel).}$$

Satz 6.2 (Kettenregel)**.** *Gegeben seien Funktionen $f : D \to K$ und $g : E \to K$, deren Verkettung existiert. Ist f an der Stelle a und g an der Stelle $f(a)$ differenzierbar, so ist $g \circ f$ an der Stelle a differenzierbar, und*

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

Beweis. Es sei $b = f(a)$. Schreiben wir wie in der Definition

$$f(x) = f(a) + \tilde{f}(x)(x - a), \qquad g(y) = g(b) + \tilde{g}(y)(y - b)$$

für $x \in D$ und $y \in E$, so erhalten wir wegen $f(x) \in E$

$$g(f(x)) = g(b) + \tilde{g}(f(x))\big(f(x) - f(a)\big) = g(b) + \tilde{g}(f(x))\tilde{f}(x)(x - a).$$

Die Funktion $\tilde{g}(f(x))\tilde{f}(x)$ ist nach den Sätzen 5.3 und 5.4 und Folgerung 6.1 an der Stelle a stetig und hat dort den gewünschten Wert. $\square$

Beispiel. Für $s \in \mathbb{R}$ ist die Verkettung von $f(x) = s \ln x$ und $g(y) = \exp y$ gleich

$$g \circ f(x) = x^s.$$

Wegen $f'(a) = \frac{s}{a}$ und $g'(b) = \exp b$ erhalten wir für $a > 0$

$$(g \circ f)'(a) = \exp(s \ln a) \cdot \frac{s}{a} = sa^{s-1},$$

was unser erstes Beispiel verallgemeinert. $\triangleleft$

Satz 6.3. *Es sei $f : D \to K$ injektiv und an der Stelle a differenzierbar, wobei $f'(a) \neq 0$. Ist die Umkehrfunktion $g : E \to K$ an der Stelle $b = f(a)$ stetig, so ist sie an dieser Stelle differenzierbar, und*

$$g'(b) = \frac{1}{f'(a)}.$$

Beweis. Wie in der Definition schreiben wir für $x \in D$

$$f(x) - f(a) = \tilde{f}(x)(x-a).$$

Da f injektiv ist, gilt für $y \in E$

$$y - b = \tilde{f}(g(y))\big(g(y) - g(b)\big).$$

Da g in b stetig ist, ist es nach Satz 5.4 auch $\tilde{f} \circ g$. Wegen

$$\tilde{f}(g(b)) = \tilde{f}(a) = f'(a) \neq 0$$

ist nach Satz 5.3 auch $\frac{1}{\tilde{f} \circ g}$ in b stetig, und

$$g(y) - g(b) = \frac{1}{\tilde{f}(g(y))}\,(y-b).$$

Es folgt die Differenzierbarkeit und die Formel für $g'(b)$. $\square$

Beispiel. Die Umkehrfunktion von $f(x) = \sin x$ ist $g(y) = \arcsin y$. Wenn $a \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ und $b = \sin a$ ist, so ist $\cos a > 0$, also

$$g'(b) = \frac{1}{\cos a} = \frac{1}{\sqrt{1-\sin^2 a}} = \frac{1}{\sqrt{1-b^2}}. \qquad \triangleleft$$

Bemerkung. Aus $f'(a) \neq 0$ folgt nicht die Injektivität von f in einer Umgebung von a, wie das Beispiel der überall differenzierbaren Funktion

$$f(x) = \begin{cases} x + x^2 \cos \frac{\pi}{x}, & \text{wenn } x \neq 0, \\ 0, & \text{wenn } x = 0 \end{cases}$$

zeigt. Es ist nämlich

$$\begin{aligned} f\left(\frac{1}{n}\right) - f\left(\frac{1}{n+1}\right) &= \frac{1}{n} - \frac{1}{n+1} + \frac{\cos \pi n}{n^2} - \frac{\cos \pi(n+1)}{(n+1)^2} \\ &= \frac{1}{n(n+1)} + (-1)^n \left(\frac{1}{n^2} + \frac{1}{(n+1)^2}\right), \end{aligned}$$

und somit ist f in keiner Umgebung von 0 monoton.

Bemerkung. Die Stetigkeit von g an der Stelle b folgt nicht aus den anderen Bedingungen (siehe Aufgabe 60). Man kann sie durch Einschränkung von f erzwingen. Dazu wähle man $0 < \varepsilon < |f'(a)|$. Dann gibt es ein $\delta > 0$, so dass für $|x-a| < \delta$ gilt $|\tilde{f}(x) - \tilde{f}(a)| < \varepsilon$. Wir schränken f auf die Menge $D_0 = \{x \in D \,\big|\, |x-a| < \delta\}$ ein. Wegen

$$f'(a)(x-a) = \big(f(x) - f(a)\big) - \big(\tilde{f}(x) - \tilde{f}(a)\big)(x-a),$$

ist für $x \in D_0$

$$|f'(a)||x-a| \leq |f(x) - f(a)| + \varepsilon |x-a|,$$

das heißt

$$|x-a| \leq \frac{|f(x) - f(a)|}{|f'(a)| - \varepsilon},$$

also für y im Wertebereich von $f|_{D_0}$

$$|g(y) - g(b)| \leq \frac{|y-b|}{|f'(a)| - \varepsilon}.$$

Definition 6.2. *Eine Funktion f heißt* differenzierbar*, wenn sie an jeder Stelle ihres Definitionsbereiches differenzierbar ist. In diesem Fall nennt man die Funktion f' ihre Ableitung.*

Jeder Punkt von D muss dann ein Häufungspunkt von D sein. Traditionell wird der Strich, der die Ableitung kennzeichnet, nicht direkt an Wortfunktionen wie exp, ln, sin oder cos angefügt. Ist $f(x)$ durch einen Term gegeben, so bezeichnet dieser Term, eingeklammert und versehen mit einem Strich, einen Term, der $f'(x)$ angibt. In diesem Sinne ergibt sich aus den obigen Beispielen:

$$\begin{aligned} (x^s)' &= s\,x^{s-1}, & (\sin x)' &= \cos x, \\ (\exp x)' &= \exp x, & (\cos x)' &= -\sin x, \\ (\ln x)' &= \frac{1}{x}, & (\arcsin x)' &= \frac{1}{\sqrt{1-x^2}}. \end{aligned}$$

6.2 Eigenschaften differenzierbarer Funktionen

Vorlesung 13a
29.06.22

Definition 6.3. *Es sei $D \subseteq \mathbb{R}$. Eine Funktion $f : D \to \mathbb{R}$ hat an der Stelle a ein* lokales Maximum*, wenn $a \in D$ ist und es ein $\delta > 0$ gibt, so dass für alle $x \in D$ gilt*

$$|x-a| < \delta \quad \Rightarrow \quad f(x) \leq f(a).$$

Analog definiert man ein lokales Minimum. *Eine Funktion hat an der Stelle a ein* lokales Extremum, *wenn sie dort ein lokales Minimum oder lokales Maximum hat.*

Der folgende Satz ist nützlich bei der Bestimmung lokaler Extrema.

Satz 6.4. *Es sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ an der Stelle a differenzierbar. Hat f ein lokales Extremum an der Stelle a und ist a ein Häufungspunkt von $D_{<a}$ und von $D_{>a}$, so gilt $f'(a) = 0$.*

Beweis. Hat f an der Stelle a beispielsweise ein lokales Maximum, dann gibt es $\delta > 0$, so dass der Differenzenquotient $\frac{f(x)-f(a)}{x-a}$ für $x \in D \cap]a, a+\delta[$ nicht positiv und für $x \in D \cap]a-\delta, a[$ nicht negativ ist, also nach Satz 3.7 und 5.9

$$\lim_{x\to a+0} \frac{f(x)-f(a)}{x-a} \leq 0, \qquad \lim_{x\to a-0} \frac{f(x)-f(a)}{x-a} \geq 0.$$

(Man bezeichnet dies auch als die einseitigen Ableitungen.) Da f an der Stelle a differenzierbar ist, existieren diese Grenzwerte und stimmen mit $f'(a)$ überein, und die Behauptung folgt. □

Beispiel. In Satz 2.9 haben wir gesehen, dass für $k \in \mathbb{N}$ und $0 < q < 1$ die Folge der Zahlen $n^k q^n$ die obere Schranke $\left(\frac{k}{1-q}\right)^k$ hat. Nun können wir das Supremum m der Funktion

$$f(x) = x^k q^x$$

auf $[0, \infty[$ bestimmen. Da m positiv ist, gibt es wegen Folgerung 5.8 ein $c > 0$, so dass für $x > c$ gilt $f(x) < m$. Nach Satz 5.6 hat f auf $[0, c]$ ein Maximum, und dieses ist auch das Maximum von f auf ganz $[0, \infty[$. Für $k > 0$ wird es wegen $f(0) = 0$ nach Satz 6.4 in einer Nullstelle von

$$f'(x) = kx^{k-1}q^x + x^k q^x \ln q$$

angenommen, also einer Lösung der Gleichung

$$k + x \ln q = 0.$$

Diese hat nur die Lösung $x = -k/\ln q$, und wir erhalten

$$\max\{x^k q^x \mid x \geq 0\} = \left(-\frac{k}{e \ln q}\right)^k,$$

was auch für $k = 0$ richtig ist. Die grobe Schranke aus Satz 2.9 gilt wegen $1 - q \geq \ln q \geq e \ln q$ auch für die Funktion f. ◁

Satz 6.5 (Satz von Rolle)**.** *Es sei* $f : [a, b] \to \mathbb{R}$ *stetig und* $f|_{]a,b[}$ *differenzierbar, wobei* $a < b$*. Wenn* $f(a) = f(b)$ *ist, dann gibt es ein* $c \in\,]a, b[$*, so dass* $f'(c) = 0$ *ist.*

Beweis. Nach Satz 5.6 hat f ein Maximum und ein Minimum. Ist wenigstens eines von beiden verschieden von $f(a)$, so wird es in einem inneren Punkt c angenommen, und mit Satz 6.4 folgt $f'(c) = 0$. Andernfalls ist $\max f = \min f = f(a) = f(b)$, und dann ist f konstant. □

Die Definition der Ableitung ist auch auf Funktionen $[a, b] \to \mathbb{C}$ anwendbar, aber für diese gilt der Satz nicht, wie das Beispiel $f(x) = \exp ix$ auf $D = [0, 2\pi]$ zeigt.

Satz 6.6 (verallgemeinerter Mittelwertsatz der Differentialrechnung)**.** *Die Funktionen* f*,* $g : [a, b] \to \mathbb{R}$ *seien stetig und auf* $]a, b[$ *differenzierbar, wobei* $a < b$*. Dann gibt es ein* $c \in\,]a, b[$*, so dass*

$$\big(f(b) - f(a)\big)g'(c) = \big(g(b) - g(a)\big)f'(c).$$

Natürlich ist dann

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)},$$

falls die Nenner nicht verschwinden.

Beweis. Wir wenden den Satz von Rolle auf die Funktion

$$h(x) = \big(f(b) - f(a)\big)\big(g(x) - g(a)\big) - \big(g(b) - g(a)\big)\big(f(x) - f(a)\big)$$

an und beachten, dass $h(a) = h(b) = 0$ und

$$h'(x) = \big(f(b) - f(a)\big)g'(x) - \big(g(b) - g(a)\big)f'(x).$$ □

Im Spezialfall $g(x) = x$ erhalten wir:

Folgerung 6.3 (Mittelwertsatz der Differentialrechnung)**.** *Unter den Bedingungen des Satzes gibt es ein* $c \in\,]a, b[$*, so dass*

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

Dies gilt sinngemäß auch für $b < a$. Man kann beide Fälle mit einer alternativen Formulierung erfassen. Dazu brauchen wir folgenden Begriff.

Definition 6.4. *Wir nennen* a *einen* inneren Punkt *der Teilmenge* M *eines Körpers* K *mit Absolutbetrag, wenn es ein* $\delta > 0$ *gibt, so dass alle Punkte* x *von* K *mit der Eigenschaft* $|x - a| < \delta$ *in* M *liegen. Die Teilmenge der inneren Punkte bezeichnen wir mit* $\mathring{M}$*.*

Offensichtlich ist eine Teilmenge von K genau dann offen, wenn sie nur innere Punkte hat.

Folgerung 6.4. *Es sei I ein Intervall, $f : I \to \mathbb{R}$ stetig und in jedem inneren Punkt differenzierbar. Dann gibt es für $a \in I$ und $h \in \mathbb{R}$ mit der Eigenschaft $a + h \in I$ ein $\vartheta \in \,]0, 1[$, so dass*

$$f(a + h) - f(a) = f'(a + \vartheta h)h.$$

Man zeigt dies durch Anwendung des Mittelwertsatzes auf die Funktion $g(t) = f(a + th)$ auf $[0, 1]$.

Folgerung 6.5. *Unter den Bedingungen von Folgerung 6.4 ist f genau dann Lipschitz-stetig, wenn f' beschränkt ist.*

Ist nämlich d eine obere Schranke von $|f'|$, so gibt es nach dem Mittelwertsatz zwischen beliebigen Punkten a und b in I ein c, so dass

$$|f(a) - f(b)| = |f'(c)||b - a| \leq d|b - a|.$$

Ist umgekehrt d eine Lipschitz-Konstante, so gilt für innere Punkte a und beliebige Punkte $x \neq a$ in I

$$\left| \frac{f(x) - f(a)}{x - a} \right| \leq d,$$

und durch Grenzübergang $x \to a$ ergibt sich $|f'(a)| \leq d$.

Beispiel. Man kann zeigen, dass die Funktion $f(x) = x^s \cos \frac{\pi}{x}$ für eine reelle Zahl $s < 2$ nicht Lipschitz-stetig ist, indem man ihre Werte an den Stellen $\frac{1}{n}$ für aufeinanderfolgende $n \in \mathbb{N} \setminus \{0\}$ betrachtet. Wegen

$$f'(x) = x^{s-2} \left(\pi \sin \frac{\pi}{x} + sx \cos \frac{\pi}{x} \right)$$

sieht man, dass f genau für $s \geq 2$ Lipschitz-stetig ist. Sie setzt sich genau dann zu einer stetigen Funktion auf ganz $\mathbb{R}$ fort, wenn $s > 0$ ist, und diese ist genau dann auch an der Stelle 0 differenzierbar, wenn $s > 1$ ist. $\triangleleft$

Folgerung 6.6. *Unter den Bedingungen von Folgerung 6.4 ist f genau dann konstant, wenn für alle inneren Punkte x von I gilt $f'(x) = 0$.*

Bei der Differentiation geht man von einer Funktion zu einer (von ihr) abgeleiteten Funktion über. Manchmal will man diese Operation umkehren.

Definition 6.5. *Eine differenzierbare Funktion F auf einem Intervall I heißt* Stammfunktion *der Funktion f, wenn f die Ableitung von F ist.*

Wir werden im Rahmen der Integralrechnung näher darauf eingehen, wie man F findet. Ist F eine Stammfunktion von f und C eine Zahl (betrachtet als konstante Funktion), so ist $F + C$ ebenfalls eine Stammfunktion von f. Umgekehrt erhalten wir aus der vorigen Folgerung:

Folgerung 6.7. *Sind F und G Stammfunktionen der Funktion f, so gibt es eine Konstante C derart, dass $G = F + C$.*

Das Monotonieverhalten differenzierbarer Funktionen lässt sich leicht beschreiben.

Folgerung 6.8. *Unter den Bedingungen von Folgerung 6.4 ist f genau dann monoton wachsend, wenn für alle inneren Punkte x gilt $f'(x) \geq 0$, und f ist streng monoton wachsend, wenn darüber hinaus f' auf keinem offenen Teilintervall verschwindet.*

Ist nämlich f monoton wachsend, so sind die Differenzenquotienten nichtnegativ, und dies gilt nach Satz 3.7 und 5.9 auch für die Ableitung. Die Umkehrung folgt aus dem Mittelwertsatz. Die Funktion ist nach Folgerung 6.6 genau dann konstant auf einem Teilintervall, wenn die Ableitung auf seinem Inneren verschwindet.

Aus der letzten Folgerung ergibt sich sofort:

Folgerung 6.9. *Unter den Bedingungen von Folgerung 6.4 sei $a \in I$. Gibt es ein $\delta > 0$, so dass $f'(x) \geq 0$ für $x \in I \cap \,]a - \delta, a[$ und $f'(x) \leq 0$ für $x \in I \cap \,]a, a + \delta[$, so hat f an der Stelle a ein lokales Maximum.*

Die Monotonie der Ableitung hängt eng mit folgendem Begriff zusammen.

Definition 6.6. *Es sei I ein Intervall. Eine Funktion $f : I \to \mathbb{R}$ heißt* konvex, *wenn für alle x, y und z in I gilt*

$$x \leq y \leq z \quad \Rightarrow \quad (z - y)f(x) + (x - z)f(y) + (y - x)f(z) \geq 0.$$

Sie heißt konkav, *wenn rechts die umgekehrte Ungleichung gilt.*

Die obige Ungleichung ist trivialerweise erfüllt, wenn $x = y$ oder $y = z$ ist. Sie ist im Fall $x < y < z$ äquivalent zu jeder der beiden Ungleichungen

$$f(z) \geq \frac{(y - z)f(x) + (z - x)f(y)}{y - x}, \qquad f(x) \geq \frac{(z - x)f(y) + (x - y)f(z)}{z - y} \tag{9}$$

und auch zu der Ungleichung

$$f(y) \leq \frac{(z - y)f(x) + (y - x)f(z)}{z - x}, \tag{10}$$

Die Ungleichungen (9) und (10) besagen, dass der Graph von f unterhalb jeder Sehne und oberhalb ihrer Verlängerung (Sekante) verläuft. Durch Vorzeichenwechsel von f erhält man die analogen Aussagen für konkave Funktionen. Die Bedingung der Konvexität ist auch äquivalent dazu, dass für $x < y < z$ eine der Ungleichungen

$$\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x}, \qquad \frac{f(z)-f(x)}{z-x} \leq \frac{f(z)-f(y)}{z-y} \tag{11}$$

oder die Ungleichung

$$\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(y)}{z-y} \tag{12}$$

gilt.

Satz. *Unter den Bedingungen von Folgerung 6.4 sind folgende Aussagen äquivalent:*

(i) *f ist konvex.*

(ii) *f' ist monoton wachsend.*

(iii) *Für alle $a \in \mathring{I}$ und alle $x \in I$ gilt*

$$f(x) \geq f(a) + f'(a)(x-a).$$

Aussage (iii) besagt, dass der Graph oberhalb jeder Tangente verläuft.

Beweis. Angenommen, es gilt Aussage (i). Wenden wir Ungleichung (12) auf $a < x < b$ und $x < b < y$ an, so erhalten wir wegen der Transitivität

$$\frac{f(x)-f(a)}{x-a} \leq \frac{f(y)-f(b)}{y-b},$$

und für $a, b \in \mathring{I}$ folgt wegen der Differenzierbarkeit $f'(a) \leq f'(b)$, also Aussage (ii). Wenden wir hingegen die linke Ungleichung in (11) auf $a < y < z$ und die rechte auf $x < y < a$ an und bilden die Grenzwerte für $y \to a+0$ bzw. $y \to a-0$, so folgt

$$f'(a) \leq \frac{f(z)-f(a)}{z-a}, \qquad \frac{f(a)-f(x)}{a-x} \leq f'(a),$$

also Aussage (iii). Umgekehrt folgt aus den beiden letzten Ungleichungen wegen der Transitivität die Ungleichung (12) für $x < a < z$, also Aussage (i).

Für $x < y < z$ gibt es nach dem Mittelwertsatz $u \in \,]x,y[$ und $v \in \,]y,z[$, so dass

$$\frac{f(y)-f(x)}{y-x} = f'(u), \qquad \frac{f(z)-f(y)}{z-y} = f'(v).$$

Ist also (ii) erfüllt, so folgt (i). □

Wir untersuchen nun Grenzwerte von Quotienten, bei denen die Rechenregeln versagen.

Satz 6.7 (Regel von de l'Hospital im Fall $\frac{0}{0}$)**.** *Die Funktionen f und g seien differenzierbar auf $]a,b[$, es sei $f(x) \to 0$ und $g(x) \to 0$ $(x \to a)$, und für $x \in \,]a,b[$ sei $g(x) \neq 0$ und $g'(x) \neq 0$. Wenn*

$$\frac{f'(x)}{g'(x)} \to d \quad (x \to a),$$

so auch

$$\frac{f(x)}{g(x)} \to d \quad (x \to a).$$

Beweis. Wir setzen f und g durch $f(a) = g(a) = 0$ stetig fort. Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $z \in \,]a, a+\delta[$ gilt

$$\left| \frac{f'(z)}{g'(z)} - d \right| < \varepsilon.$$

Ist nun $x \in \,]a, a+\delta[$, so gibt es nach Satz 6.6 ein $z \in \,]a, x[$, so dass

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(z)}{g'(z)}. \qquad \square$$

Satz 6.8 (Regel von de l'Hospital im Fall $\frac{\infty}{\infty}$)**.** *Die Funktionen f und g seien differenzierbar auf $]a, b[$, es sei $\frac{1}{f(x)} \to 0$ und $\frac{1}{g(x)} \to 0$ $(x \to a)$, und für $x \in \,]a, b[$ sei $g'(x) \neq 0$. Wenn*

$$\frac{f'(x)}{g'(x)} \to d \quad (x \to a),$$

so auch

$$\frac{f(x)}{g(x)} \to d \quad (x \to a).$$

Beweis. Es seien ε und δ wie im vorigen Beweis. Da keine Werte an der Stelle a existieren, betrachten wir statt dessen einen Punkt $y \in \,]a, a+\delta[$. Dann gilt für x nah bei a

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(y)}{g(x) - g(y)} h(x)$$

mit einem Korrekturfaktor

$$h(x) = \frac{1 - \frac{g(y)}{g(x)}}{1 - \frac{f(y)}{f(x)}}.$$

Aus den Annahmen über f und g folgt, dass $h(x) \to 1$ $(x \to a)$, also gibt es ein $\eta > 0$, so dass $|h(x) - 1| < \varepsilon$ für $x \in \,]a, a+\eta[$. Nach Satz 6.6 gibt es ein $z \in \,]x, y[$, so dass

$$\frac{f(x)}{g(x)} = \frac{f'(z)}{g'(z)} h(x)$$

also gilt für $x \in \,]a, a+\eta[$ nach der Dreiecksungleichung

$$\left| \frac{f(x)}{g(x)} - d \right| \leq \left| \frac{f'(z)}{g'(z)} - d \right| |h(x)| + |d||h(x) - 1| < \varepsilon(1+\varepsilon) + |d|\varepsilon. \qquad \square$$

Folgerung 6.10. *Die Regeln von de l'Hospital gelten auch für $x \to \infty$.*

Mit den Bezeichnungen $f_1(x) = f(\frac{1}{x})$ und $g_1(x) = g(\frac{1}{x})$ ist nämlich $f_1'(x) = -\frac{1}{x^2} f'(\frac{1}{x})$ und somit

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{y\to 0} \frac{f_1(y)}{g_1(y)} = \lim_{y\to 0} \frac{f_1'(y)}{g_1'(y)} = \lim_{y\to 0} \frac{f'(\frac{1}{y})}{g'(\frac{1}{y})} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)}.$$

Bemerkung 1. Die Regeln von de l'Hospital gelten natürlich sinngemäß für linksseitige Grenzwerte und, indem man beides zusammenfügt, auch für beidseitige Grenzwerte. Letzteres kann man auch direkt mit Hilfe von Folgerung 6.4 zeigen.
Bemerkung 2. In Satz 6.8 ist wegen $\frac{1}{g(x)} \to 0$ automatisch $g(x) \neq g(y)$ für festes y und alle x nahe bei a. Daraus folgt aber nicht $g'(z) \neq 0$, denn f' und g' könnten gleichzeitig verschwinden.

6.3 Höhere Ableitungen

Vorlesung 13b
01.07.2022

Wir betrachten wieder Funktionen f auf einer Teilmenge D eines Körpers K, der $\mathbb{R}$ oder $\mathbb{C}$ sein kann, mit Werten im selben Körper. Dabei soll die Bezeichnung f nicht nur die Zuordnung, sondern auch den Definitionsbereich festlegen. Die Ableitung einer beliebigen Funktion $f : D \to K$ ist eine Funktion $f' : D_1 \to K$, wobei D_1 die Menge der Stellen in D ist, an denen f differenzierbar ist, was auch die leere Menge sein kann. Insbesondere ist $D_1 \subseteq D$ und besteht aus Häufungspunkten von D.

Bezeichnet $\mathcal{F}$ die Menge der Funktionen auf Teilmengen von K mit Werten in K, so ist die Differentiation eine Operation, also schlicht gesagt eine Abbildung $\mathcal{F} \to \mathcal{F}$, die wir mit $f \mapsto f'$ bezeichnen. Diese können wir auch mehrmals anwenden, und die zweite Ableitung $(f')'$ schreiben wir einfach f'' (gelesen f zwei Strich), die dritte Ableitung $((f')')'$ schreiben wir f''' (gelesen f drei Strich) usw.

Definition 6.7. *Wir definieren die n-te* Ableitung[34] $f^{(n)}$ *von f rekursiv durch*

$$f^{(0)} = f, \qquad f^{(n+1)} = \left(f^{(n)}\right)'.$$

Eine Funktion f heißt n-mal differenzierbar *an der Stelle a, wenn a zum Definitionsbereich D_n von $f^{(n)}$ gehört.*

Eine Funktion heißt unendlich oft[35] differenzierbar *an der Stelle a, wenn sie für jede natürliche Zahl n an dieser Stelle n-mal differenzierbar ist.*

[34] oder Ableitung der Ordnung n
[35] richtiger wäre „beliebig oft“

Eine Funktion heißt n-mal differenzierbar*, wenn dies an jeder Stelle ihres Definitionsbereichs der Fall ist. Sie heißt n-mal* stetig differenzierbar*, wenn außerdem ihre n-te Ableitung stetig ist.*

Man beachte, dass im letzteren Fall ihre Ableitungen kleinerer Ordnung als n differenzierbar, also ebenfalls stetig sind. Im Allgemeinen haben die Ableitungen verschiedener Ordnung Definitionsbereiche

$$D = D_0 \supseteq D_1 \supseteq D_2 \dots$$

Wegen der Assoziativität der Verkettung gilt für beliebige natürliche Zahlen m und n

$$f^{(m+n)} = (f^{(m)})^{(n)},$$

was man leicht durch vollständige Induktion nach n beweisen kann. Eine Funktion ist also an einer Stelle genau dann $m+n$-mal differenzierbar, wenn ihre m-te Ableitung an dieser Stelle n-mal differenzierbar ist.

Beispiel. Durch vollständige Induktion beweist man, dass für $s \in \mathbb{R}$ und $a > 0$ die Funktionen

$$f(x) = x^s, \qquad g(x) = a^x,$$

wobei f nur für $x > 0$ definiert ist, die höheren Ableitungen

$$f^{(n)}(x) = s(s-1)\cdots(s-n+1)x^{s-n}, \qquad g^{(n)}(x) = (\ln a)^n a^x$$

haben und unendlich oft differenzierbar sind. Für $s \in \mathbb{N}$ ist die Potenzfunktion f sogar auf ganz $\mathbb{C}$ definiert, und für $n > s$ ist $f^{(n)}(x) = 0$, während die obige Formel bei $x = 0$ nicht mehr definiert ist. Wegen $\ln' x = x^{-1}$ ist auch die Logarithmusfunktion unendlich oft differenzierbar. $\triangleleft$

Satz 6.9. *Sind Funktionen f und g mit dem selben Definitionsbereich n-mal (stetig) differenzierbar, so auch die Funktionen $f+g$, $f \cdot g$ und, falls g auf D keine Nullstelle hat, auch $\frac{f}{g}$. Außerdem ist*

$$(f+g)^{(n)} = f^{(n)} + g^{(n)}, \qquad (fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)} g^{(k)}.$$

Beweis. Im Induktionsanfang, also für $n = 0$, ist nichts zu beweisen bzw. er folgt aus Satz 5.3. Angenommen, die Behauptungen gelten für eine Zahl n, und wir betrachten nun $n+1$-mal (stetig) differenzierbare Funktionen f und g, die dann natürlich auch n-mal (stetig) differenzierbar sind. Nach Satz 6.1 und Folgerung 6.2 gilt

$$(f+g)' = f' + g', \qquad (fg)' = f'g + fg', \qquad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2},$$

und die rechten Seiten sind nach Induktionsvoraussetzung n-mal (stetig) differenzierbar. Damit folgt die Induktionsbehauptung entsprechend der Bemerkung nach der Definition. Die Formeln beweist man ebenfalls durch vollständige Induktion, wobei man bei der Produktregel wie im Beweis der binomischen Formel vorgeht. □

Satz 6.10. (i) *Sind die Funktionen f und g beide n-mal (stetig) differenzierbar und existiert ihre Verkettung, so ist sie ebenfalls n-mal (stetig) differenzierbar.*

(ii) *Ist f injektiv und n-mal (stetig) differenzierbar, wobei $n \geq 1$, hat f' keine Nullstellen und ist die Umkehrfunktion von f stetig, so ist sie ebenfalls n-mal (stetig) differenzierbar.*

Beweis. **(i)** Im Induktionsanfang ist nichts zu zeigen, oder er folgt aus Satz 5.4. Angenommen, die Behauptung gilt für eine Zahl n, und f, g sind $n+1$-mal (stetig) differenzierbar. Nach Satz 6.2 ist

$$(g \circ f)' = (g' \circ f) \cdot f',$$

und nach Satz 6.9 und der Induktionsvoraussetzung ist die rechte Seite n-mal (stetig) differenzierbar.
(ii) Der Induktionsanfang folgt aus Satz 6.3. Angenommen, die Behauptung gilt für eine Zahl n, und f ist $n+1$-mal (stetig) differenzierbar. Nach Satz 6.3 gilt für die Umkehrfunktion h

$$h' = \frac{1}{f' \circ h},$$

und nach Satz 6.10(i), Satz 6.9 und der Induktionsvoraussetzung ist die rechte Seite n-mal (stetig) differenzierbar. □

Im Fall $K = \mathbb{R}$ kann man höhere Ableitungen benutzen, um festzustellen, ob an einem stationären Punkt ein lokales Extremum vorliegt.

Satz 6.11. *Für eine natürliche Zahl $n \geq 1$ sei die Funktion f auf einem Intervall n mal stetig differenzierbar, und es gelte für einen inneren Punkt a*

$$f'(a) = f''(a) = \ldots = f^{(n-1)}(a) = 0, \qquad f^{(n)}(a) \neq 0.$$

Dann hat f an der Stelle a

- *kein lokales Extremum, falls n ungerade,*
- *ein lokales Minimum, falls n gerade und $f^{(n)}(a) > 0$,*

- *ein lokales Maximum, falls n gerade und $f^{(n)}(a) < 0$.*

Beweis. Wir sagen, dass eine Eigenschaft von $f : I \to \mathbb{R}$ lokal rechts von a vorliegt, wenn es ein $\delta > 0$ gibt, so dass sie auf $I \cap [a, a + \delta[$ vorliegt. Analog definieren wir, wann sie lokal links von a vorliegt bzw. wann sie lokal bei a vorliegt. Wir machen die folgenden stärkeren Behauptungen: In der Situation des Satzes ist f

(i) lokal bei a streng monoton wachsend, falls n ungerade und $f^{(n)}(a) > 0$,

(ii) lokal bei a streng monoton fallend, falls n ungerade und $f^{(n)}(a) < 0$,

(iii) lokal links von a streng monoton fallend und rechts von a streng monoton wachsend, falls n gerade und $f^{(n)}(a) > 0$,

(iv) lokal links von a streng monoton wachsend und rechts von a streng monoton fallend, falls n gerade und $f^{(n)}(a) < 0$.

Dies beweisen wir durch vollständige Induktion. Im Fall $n = 1$ folgt aus der Stetigkeit von f', dass es ein $\delta > 0$ gibt, so dass f' auf $]a - \delta, a + \delta[$ nicht das Vorzeichen wechselt, und die Aussage (i) bzw. (ii) folgt.

Angenommen, die Aussage gilt für eine Zahl $n \geq 1$, und f ist $n + 1$-mal stetig differenzierbar, wobei $f^{(k)}(a) = 0$ für $k \leq n$. Dann ist die Induktionsvoraussetzung auf f' anwendbar. Ist die Ableitung f' lokal rechts von a monoton wachsend, so ist sie wegen $f'(a) = 0$ lokal streng rechts von a positiv, also f lokal rechts von a monoton wachsend. Analog behandelt man die anderen Fälle, und die Aussage über f folgt. □

Nun sei K wieder einer der Körper $\mathbb{R}$ oder $\mathbb{C}$. Aus den uns bekannten unendlich oft differenzierbaren Funktionen können wir mit Hilfe der Sätze 6.9 und 6.10 weitere solche Funktionen erzeugen. Es gibt noch eine andere Quelle von unendlich oft differenzierbaren Funktionen. Dazu zeigen wir zunächst:

Satz 6.12. *Die Ableitung einer Potenzreihe ist eine Potenzreihe mit dem selben Konvergenzradius, die durch gliedweise Differentiation entsteht.*

Beweis. Ist

$$f(z) = \sum_{m=0}^{\infty} a_m z^m$$

mit dem Konvergenzradius r und ist $|c| < r$, so gilt nach Satz 4.15 für $|z - c| + |c| < r$

$$f(z) = \sum_{n=0}^{\infty} b_n (z - c)^n,$$

wobei

$$b_n = \sum_{m=n}^{\infty} \binom{m}{n} a_m c^{m-n}.$$

Wegen $f(c) = b_0$ folgt

$$\frac{f(z) - f(c)}{z - c} = \sum_{n=1}^{\infty} b_n (z - c)^{n-1},$$

so dass

$$f'(c) = b_1 = \sum_{m=1}^{\infty} m a_m c^{m-1}.$$

Nach Satz 4.8 hat auch diese Potenzreihe den Konvergenzradius r. □

Durch vollständige Induktion erhalten wir:

Folgerung 6.11. *Jede analytische Funktion ist unendlich oft differenzierbar.*

Die Umkehrung gilt allerdings nicht. Es gibt nämlich unendlich oft reell differenzierbare Funktionen, die nicht reell analytisch sind (siehe Aufgabe 64). Hingegen kann man zeigen, dass eine beliebige auf einer offenen Teilmenge von $\mathbb{C}$ definierte komplex differenzierbare Funktion automatisch komplex analytisch ist. Solchen Funktionen ist die Veranstaltung „Funktionentheorie“ gewidmet.

6.4 Die Taylorsche Formel

Vorlesung 14a 06.07.2022

Die Bedingung der Differenzierbarkeit einer Funktion $f : D \to K$ an der Stelle $a \in D$, nämlich

$$f(x) = f(a) + \tilde{f}(x)(x - a), \qquad \tilde{f}(x) \to f'(a) \quad (x \to a),$$

können wir so umdeuten, dass f durch die ge Funktion

$$p(x) = f(a) + f'(a)(x - a)$$

angenähert wird, d. h.

$$f(x) = p(x) + r(x),$$

wobei das Restglied die Eigenschaft

$$\frac{r(x)}{x - a} = \tilde{f}(x) - \tilde{f}(a) \to 0 \quad (x \to a)$$

hat. Die lineare Funktion p ist durch die Eigenschaften

$$p(a) = f(a), \qquad p'(a) = f'(a)$$

eindeutig bestimmt.

Für eine n-mal an der Stelle a differenzierbare Funktion f können wir eine ganzrationale Funktion

$$p_n(x) = c_0 + c_1(x-a) + \ldots + c_n(x-a)^n = \sum_{k=0}^{n} c_k(x-a)^k$$

vom Grad höchstens n finden, die an der Stelle a die selben Ableitungen wie f bis zur Ordnung n hat. Wegen

$$p_n^{(m)}(x) = \sum_{k=m}^{n} c_k k(k-1)\cdots(k-m+1)(x-a)^{k-m}, \qquad p_n^{(m)}(a) = m!c_m$$

müssen wir nämlich

$$c_k = \frac{f^{(k)}(a)}{k!}$$

setzen, damit $p_n^{(k)}(a) = f^{(k)}(a)$ für k von 0 bis n gilt. Man nennt

$$p_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

das n-te *Taylor-Polynom* von f. Wenn wir nun

$$f(x) = p_n(x) + r_n(x)$$

schreiben, wie groß ist dann das Restglied r_n? Wir geben die Antwort nur für $K = \mathbb{R}$.

Satz 6.13. *Eine Funktion f auf einem Intervall I sei n-mal stetig differenzierbar und auf der Menge $\mathring{I}$ der inneren Punkte sogar $n+1$-mal differenzierbar. Dann gibt es für jedes $x \in I$ ein z zwischen a und x, so dass*

$$r_n(x) = \frac{f^{(n+1)}(z)}{(n+1)!}(x-a)^{n+1}$$

(Restglied in der Form von Lagrange).

Beweis. Wir halten $x > a$ fest und betrachten die Hilfsfunktion

$$g(y) = \sum_{k=0}^{n} \frac{f^{(k)}(y)}{k!}(x-y)^k.$$

Dann gilt $g(x) = f(x)$, $g(a) = p_n(x)$ und

$$g'(y) = \sum_{k=0}^{n} \frac{f^{(k+1)}(y)}{k!}(x-y)^k - \sum_{k=1}^{n} \frac{f^{(k)}(y)}{(k-1)!}(x-y)^{k-1} = \frac{f^{(n+1)}(y)}{n!}(x-y)^n.$$

Ist h auf $[a,x]$ stetig und auf $]a,x[$ differenzierbar, so dass h' keine Nullstellen hat, dann gibt es nach dem verallgemeinerten Mittelwertsatz ein $z \in \,]a,x[$, so dass

$$\frac{g(x)-g(a)}{h(x)-h(a)} = \frac{g'(z)}{h'(z)},$$

und es folgt

$$r_n(x) = g(x) - g(a) = \frac{h(x)-h(a)}{h'(z)} \cdot \frac{f^{(n+1)}(z)}{n!}(x-z)^n.$$

Setzen wir $h(y) = (x-y)^{n+1}$, so gilt $h'(z) = -(n+1)(x-z)^n$, und die Behauptung folgt. Im Fall $x < a$ betrachtet man das Selbe auf $[x,a]$. $\square$

Das Taylorpolynom p_1 vom Grad höchstens 1 ist die eingangs erwähnte lineare Funktion p. Damals genügte die Differenzierbarkeit von f an der Stelle a, um eine Aussage über das Restglied $r = r_1$ zu treffen. Satz 6.13 hingegen verlangt im Fall $n = 1$ die zweimalige Differenzierbarkeit in einer Umgebung von a und liefert dafür genauere Informationen über das Restglied. Es gibt aber eine Version des Satzes, die das frühere Ergebnis getreu verallgemeinert:

Satz 6.14. *Ist f auf einem Intervall $n-1$-mal differenzierbar und an der Stelle a sogar n-mal differenzierbar, so gilt*

$$\frac{r_n(x)}{(x-a)^n} \to 0 \quad (x \to a).$$

Beweis. Wir können $n \geq 2$ annehmen. Dann erfüllt die Funktion r_n wegen Folgerung 6.1 die Anforderungen von Satz 6.13 für $n-2$ statt n, wobei $r_n^{(k)}(a) = 0$ für $k \leq n$, und wir erhalten ein $z \in \mathring{I}$ mit der Eigenschaft $|z-a| < |x-a|$, so dass

$$r_n(x) = \frac{r_n^{(n-1)}(z)}{(n-1)!}(x-a)^{n-1}.$$

Es folgt

$$\left|\frac{r_n(x)}{(x-a)^n}\right| = \frac{1}{(n-1)!}\left|\frac{r_n^{(n-1)}(z)}{x-a}\right| < \frac{1}{(n-1)!}\left|\frac{r_n^{(n-1)}(z)}{z-a}\right|.$$

Wegen der n-maligen Differenzierbarkeit an der Stelle a ist

$$\frac{r_n^{(n-1)}(z)}{z-a} = \frac{r_n^{(n-1)}(z) - r_n^{(n-1)}(a)}{z-a} \to r_n^{(n)}(a) = 0 \quad (z \to a),$$

und die Behauptung folgt. □

Man kann die Taylorsche Formel zum Einen zur Berechnung von Grenzwerten benutzen, zum Anderen um festzustellen, ob an einem stationären Punkt ein lokales Extremum vorliegt. Das folgende Kriterium kommt mit schwächeren Voraussetzungen als Satz 6.11 aus.

Satz 6.15. *Für eine natürliche Zahl $n \geq 1$ sei die Funktion f auf einem Intervall $n-1$-mal differenzierbar und an der Stelle a sogar n-mal differenzierbar, und es gelte*

$$f'(a) = f''(a) = \ldots = f^{(n-1)}(a) = 0, \qquad f^{(n)}(a) \neq 0.$$

Dann hat f an der Stelle a

(i) *kein lokales Extremum, falls n ungerade,*

(ii) *ein lokales Maximum, falls n gerade und $f^{(n)}(a) < 0$,*

(iii) *ein lokales Minimum, falls n gerade und $f^{(n)}(a) > 0$.*

Beweis. Die Taylorsche Formel vereinfacht sich zu

$$f(x) = f(a) + \frac{f^{(n)}(a)}{n!}(x-a)^n + r_n(x),$$

und es folgt

$$f(x) - f(a) = (x-a)^n \left(\frac{f^{(n)}(a)}{n!} + \frac{r_n(x)}{(x-a)^n}\right).$$

Zu $\varepsilon = \frac{|f^{(n)}(a)|}{n!}$ gibt es nach Satz 6.14 ein $\delta > 0$, so dass die Klammer auf der rechten Seite für $|x-a| < \delta$ konstantes Vorzeichen hat. □

Beispiel. Ist f eine analytische Funktion, so ist p_n die n-te Partialsumme der Potenzreihe mit dem Zentrum a:

$$f(x) = \sum_{k=0}^{\infty} a_k (x-a)^k, \qquad p_n(x) = \sum_{k=0}^{n} a_k (x-a)^k.$$

Die ist z. B. auf die Funktionen exp, sin und cos anwendbar. $\triangleleft$

Die im vorigen Abschnitt berechneten Beispiele von höheren Ableitungen versetzen uns in die Lage, weitere Taylor-Polnome und Restglieder zu bestimmen:

Beispiel. Ist $f(x) = \ln(1+x)$, so gilt

$$f^{(k)}(x) = (-1)^{k-1}(k-1)!(1+x)^{-k}$$

für $k \geq 1$ und somit

$$p_n(x) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} x^k, \qquad r_n(x) = \frac{(-1)^n}{n+1} \left(\frac{x}{1+z} \right)^{n+1}$$

für ein z zwischen 0 und x. $\triangleleft$

Beispiel. Ist $s \in \mathbb{R}$ und $f(x) = (1+x)^s$, so gilt

$$f^{(k)}(x) = s(s-1)\cdots(s-k+1)(1+x)^{s-k}$$

und somit

$$p_n(x) = \sum_{k=0}^{n} \binom{s}{k} x^k, \qquad r_n(x) = \binom{s}{n+1} (1+z)^{s-n-1} x^{n+1}$$

für ein z zwischen 0 und x. $\triangleleft$

Definition 6.8. *Die* Taylor-Reihe *einer auf einem Intervall I unendlich oft differenzierbaren Funktion f an der Stelle $a \in I$ ist die Potenzreihe*

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k.$$

Was ist der Konvergenzradius dieser Reihe, welche Funktion stellt sie dar? Im ersten Beispiel erhalten wir die schon bekannten Reihen aus Satz 4.10 und dem Beweis von Satz 5.8 mit dem Konvergenzradius ∞. In den letzten beiden Beispielen ist der Konvergenzradius nach Satz 4.7 und Aufgabe 47 gleich 1. Es kommt auch vor, dass die Taylor-Reihe den Konvergenzradius 0 hat.

Die Taylor-Reihe von f konvergiert natürlich genau dann gegen f, wenn das Restglied gegen 0 konvergiert. Im Beispiel $f(x) = \ln(1+x)$ haben wir

$$|r_n(x)| < \begin{cases} \frac{x^{n+1}}{n+1}, & \text{wenn } x > 0, \\ \frac{1}{n+1}\left(\frac{|x|}{1-|x|}\right)^{n+1}, & \text{wenn } x < 0. \end{cases}$$

Für $-\frac{1}{2} \leq x \leq 1$ folgt also

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$$

(*Logarithmusreihe*), und für $x = 1$ erhalten wir nebenbei den Wert der alternierenden harmonischen Reihe:

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \ln 2.$$

Auch im Beispiel $f(x) = (1+x)^s$ folgt $r_n(x) \to 0 \ (n \to \infty)$ und somit

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

(*Binomialreihe*) nur für $-\frac{1}{2} \leq x \leq 1$. Was passiert im Rest des Konvergenzintervalls?

Satz 6.16. *In der Situation von Satz 6.13 gibt es ein w zwischen a und x, so dass*

$$r_n(x) = \frac{f^{(n+1)}(w)}{n!}(x-a)(x-w)^n$$

(Restglied in der Form von Cauchy).

Der Beweis ist identisch mit dem von Satz 6.13 außer dass man diesmal die Funktion $h(y) = x-y$ wählt, also den gewöhnlichen Mittelwertsatz anwendet.

Beispiel. Schreiben wir das Restglied für $f(x) = \ln(1+x)$ in der Form von Cauchy, so erhalten wir

$$r_n(x) = \frac{(-1)^n}{(1+w)^{n+1}} x(x-w)^n.$$

Für $-1 < x < 0$ ist $x < w < 0$, also $w(x+1) < 0$. Durch Addition von $x - w$ auf beiden Seiten erhalten wir $x(1+w) < x - w$, also

$$\frac{x-w}{1+w} > x, \qquad \left|\frac{x-w}{1+w}\right| < |x|$$

und schließlich

$$r_n(x)| < \frac{|x|^{n+1}}{1+z} < \frac{|x|^{n+1}}{1+x},$$

so dass auch hier gilt $r_n(x) \to 0 \ (n \to \infty)$. Die Logarithmusreihe konvergiert also für $-1 < x \leq 1$ gegen $\ln(1+x)$. Ähnliches kann man auch für die Binomialreihe zeigen. $\lhd$

Beispiel. Es gibt keine allgemeine Formel für die höheren Ableitungen der Funktion

$$f(x) = \arctan x$$

auf ganz $\mathbb{R}$. Wir können aber die in der Präsenzaufgabe 37 bestimmte erste Ableitung für $|x| < 1$ als geometrische Reihe schreiben:

$$f'(x) = \frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-x^2)^n.$$

Die für $|x| \leq 1$ konvergente Potenzreihe

$$g(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$$

ist nach Satz 6.12 auf $]-1, 1[$ eine Stammfunktion von f', das heißt, es gilt $f' = g'$. Wegen $f(0) = 0 = g(0)$ zeigt Folgerung 6.5, dass $f = g$ ist, und dies gilt auch an den Intervallgrenzen wegen der Einschließung aus dem Beweis des Leibniz-Kriteriums und der Stetigkeit der beteiligten Funktionen. Wir erhalten insbesondere

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \arctan 1 = \frac{\pi}{4}$$

durch Einsetzen von $x = 1$. $\lhd$

Beispiel. Dieser Trick hilft auch bei der Funktion

$$f(x) = \arcsin x$$

weiter. Für $|x| < 1$ gilt nämlich

$$f'(x) = \frac{1}{\sqrt{1-x^2}} = \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} (-x^2)^n = \sum_{n=0}^{\infty} \frac{(2n-1)!!}{(2n)!!} x^{2n},$$

wobei rekursiv

$$0!! = (-1)!! = 1, \qquad (n+2)!! = n!! \cdot (n+2)$$

definiert wird, und man findet

$$\arcsin x = \sum_{n=0}^{\infty} \frac{(2n-1)!!}{(2n)!!} \cdot \frac{x^{2n+1}}{2n+1}$$

ebenfalls für $|x| < 1$. $\triangleleft$

6.5 Das Landausymbol

Vorlesung 15
13.07.2022

In der alternativen Definition der Differenzierbarkeit und in den Taylorschen Formeln kommen Restglieder vor. Um die lästige Einführung immer neuer Bezeichnungen zu umgehen, wurde folgende Schreibweise eingeführt.

Definition 6.9. *Kommt in einer Gleichung das Symbol*[36]

$$O(g(x))$$

vor, wobei nach der Gleichung die Angabe $x \to a$ folgt, so ist gemeint, dass die Formel zu einer wahren Aussage wird, wenn man besagtes Symbol durch einen geeigneten Term $f(x)$ ersetzt, für den gilt

$$\overline{\lim_{x \to a}} \left| \frac{f(x)}{g(x)} \right| < \infty.$$

Kommt in einer Gleichung das Symbol[37]

$$o(g(x))$$

vor, wobei nach der Gleichung die Angabe $x \to a$ folgt, so ist gemeint, dass die Gleichung zu einer wahren Aussage wird, wenn man besagtes Symbol durch einen geeigneten Term $f(x)$ ersetzt, für den gilt

$$\lim_{x \to a} \frac{f(x)}{g(x)} = 0.$$

Das Symbol O bedeutet grob gesprochen „von der Ordnung …“ und wurde von Paul Bachmann eingeführt. Das Symbol o bedeutet grob gesprochen „von kleinerer Ordnung als …“ und wurde kurz darauf von Edmund Landau eingeführt. Dabei kann $a \in \mathbb{R} \cup \{\infty, -\infty\}$ sein. Ein Landausymbol kann in einer Gleichung mehrmals vorkommen und für verschiedene Terme

[36] gelesen „(groß) O von $g(x)$“
[37] gelesen „klein o von $g(x)$“

stehen. Die Eigenschaft des Terms $f(x)$, der für $O(g(x))$ stehen kann, lässt sich einfacher ohne limes superior ausdrücken, nämlich im Fall $a \in \mathbb{R}$

$$\exists C > 0 \, \exists \delta > 0 \, \forall x \, \big(0 < |x - a| < \delta \;\Rightarrow\; |f(x)| \leq C|g(x)| \big)$$

und im Fall $a = \infty$

$$\exists C > 0 \, \exists c \in \mathbb{R} \, \forall x \, \big(x > c \;\Rightarrow\; |f(x)| \leq Cg(x) \big).$$

Hier durchläuft x den Definitionsbereich der vorkommenden Terme.

Beispiel. Die Taylorsche Formel mit Restglied in qualitativer Form lautet

$$\begin{aligned} f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots \\ \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + o\big((x-a)^n\big) \quad (x \to a) \end{aligned}$$

(siehe Satz 6.14). Im Spezialfall $n = 1$ ist dies die alternative Definition der Ableitung. Ist die Funktion für x nahe bei a sogar $n+1$-mal stetig differenzierbar, dann gilt die stärkere Aussage

$$\begin{aligned} f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots \\ \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + O\big((x-a)^{n+1}\big) \quad (x \to a), \end{aligned}$$

die aus Satz 6.13 folgt. $\triangleleft$

Beispiel. Es gilt

$$\begin{aligned} \sin x &= x - \frac{x^3}{6} + O(x^5) \quad (x \to 0), \\ \ln(1-y) &= -y - \frac{y^2}{2} - \frac{y^3}{3} - \frac{y^4}{4} + O(y^5) \quad (y \to 0). \end{aligned}$$

Wir wollen $\sin x$ für y substituieren. Es gilt

$$\sin^2 x = x^2 - \frac{x^4}{3} + O(x^5) \quad (x \to 0),$$

denn die Terme $x^3 \cdot x^3$, $x \cdot O(x^5)$ und $O(x^5) \cdot O(x^5)$ werden von $O(x^5)$ absorbiert. Analog finden wir

$$\sin^3 x = x^3 + O(x^5), \quad \sin^4 x = x^4 + O(x^5) \quad (x \to 0).$$

Schließlich ist

$$|\sin^5 x| \le |x^5|.$$

Somit ergibt sich

$$\begin{aligned}\ln(1-\sin x) &= -\left(x - \frac{x^3}{6}\right) - \frac{1}{2}\left(x^2 - \frac{x^4}{3}\right) - \frac{x^3}{3} - \frac{x^4}{4} + O(x^5) \\ &= -x - \frac{x^2}{2} - \frac{x^3}{6} + \frac{x^4}{12} + O(x^5) \quad (x \to 0).\end{aligned}$$

◁

7 Integralrechnung

7.1 Das Riemannsche Integral

Vorlesung 1a
12.10.2022

Verschiedenste geometrische und physikalische Probleme, etwa die Bestimmung der von einer Kurve umschlossenen Fläche oder der Masse bei gegebener Dichteverteilung, lassen sich durch ein und den selben mathematischen Begriff des Integrals einer Funktion modellieren. Grob gesprochen geht es darum, den Inhalt der Fläche zwischen dem Graphen und der Abszissenachse zu bestimmen.

Definition 7.1. *Unter einer* Teilung *eines beschränkten abgeschlossenen Intervalls* $[a, b]$ *positiver Länge verstehen wir eine Teilmenge* $T = \{x_0, \dots, x_n\}$, *so dass*

$$a = x_0 < x_1 < \dots < x_n = b.$$

Das Maximum der Längen $x_k - x_{k-1}$ *der Teilungsintervalle bezeichnen wir als* Feinheit *der Teilung.*

Für eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ *nennen wir*

$$\overline{S}(f, T) = \sum_{k=1}^{n} \sup_{[x_{k-1}, x_k]} f \cdot (x_k - x_{k-1}),$$
$$\underline{S}(f, T) = \sum_{k=1}^{n} \inf_{[x_{k-1}, x_k]} f \cdot (x_k - x_{k-1})$$

die Riemannsche Ober- *bzw.* Untersumme *von* f *bezüglich* T.

Die hier vorkommenden Grenzen existieren nach Satz 2.11.

Lemma 7.1. (i) *Für beliebige Teilungen* T_1 *und* T_2 *gilt*

$$\underline{S}(f, T_1) \leq \overline{S}(f, T_2).$$

(ii) *Für Teilungen* $T \subseteq T'$ *gilt*

$$\underline{S}(f, T) \leq \underline{S}(f, T'), \qquad \overline{S}(f, T) \geq \overline{S}(f, T').$$

Beweis. Ungleichung (i) im Fall $T_1 = T_2 = T$ folgt aus der Tatsache

$$\inf_{[x_{k-1}, x_k]} f \leq \sup_{[x_{k-1}, x_k]} f.$$

(ii) folgt daraus, dass für $x_{k-1} < z < x_k$ gilt

$$\begin{aligned}\inf_{[x_{k-1},x_k]} f \cdot (x_k - x_{k-1}) &= \inf_{[x_{k-1},x_k]} f \cdot \big((x_k - z) + (z - x_{k-1})\big)\\ &\leq \inf_{[x_{k-1},z]} f \cdot (z - x_{k-1}) + \inf_{[z,x_k]} f \cdot (x_k - z),\end{aligned}$$

und der analogen Ungleichung für obere Grenzen. Nun folgt der allgemeine Fall von (i) mittels Transitivität aus dem bewiesenen Spezialfall für $T = T_1 \cup T_2$ und aus (ii). $\square$

Definition 7.2. *Für eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ heißen die Zahlen*

$$\overline{\int_a^b} f(x)\,dx = \inf\{\overline{S}(f,T) \mid T \text{ ist Teilung von } [a,b]\},$$
$$\underline{\int_a^b} f(x)\,dx = \sup\{\underline{S}(f,T) \mid T \text{ ist Teilung von } [a,b]\}$$

das Riemannsche Ober- *bzw.* Unterintegral *von f über das Intervall $[a,b]$. Die Funktion f heißt* integrierbar[38]*, falls beide übereinstimmen, und den gemeinsame Wert nennt man dann das* Integral *von f über $[a,b]$, abgekürzt*[39]

$$\int_a^b f(x)\,dx.$$

Die Funktion f nennt man in diesem Zusammenhang den Integranden.

Bemerkung. Nach Lemma 7.1 und Satz 2.11 existieren Ober- und Unterintegral, und es gilt

$$\underline{\int_a^b} f(x)\,dx \leq \overline{\int_a^b} f(x)\,dx.$$

Bemerkung. Die Funktion f ist genau dann integrierbar, wenn es für jedes $\varepsilon > 0$ eine Teilung T gibt, so dass $\overline{S}(f,T) - \underline{S}(f,T) < \varepsilon$.

[38] im Riemannschen Sinne

[39] Das von Leibniz eingeführte Integralzeichen ist ein kursives langes *s*, wie es damals im Wortinneren üblich war.

Bemerkung. Würden wir in der Definition $[x_{k-1}, x_k]$ durch $]x_{k-1}, x_k[$ ersetzen, so würden sich die Werte von Ober- und Unterintegral nicht ändern.

Beispiel. Die Riemannschen Summen der Funktion $f(x) = x^3$ auf dem Intervall $[0, 1]$ bezüglich der äquidistanten Teilung

$$T_n = \left\{0, \tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{n-1}{n}, 1\right\},$$

also mit $x_k = \frac{k}{n}$, sind

$$\overline{S}(f, T_n) = \sum_{k=1}^{n} x_k^3 \cdot \frac{1}{n} = \frac{1}{n^4} \sum_{k=1}^{n} k^3, \qquad \underline{S}(f, T_n) = \sum_{k=1}^{n} x_{k-1}^3 \cdot \frac{1}{n} = \frac{1}{n^4} \sum_{k=1}^{n} (k-1)^3.$$

Man zeigt durch vollständige Induktion, dass

$$\sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2,$$

und es folgt

$$\lim_{n\to\infty} \overline{S}(f, T_n) = \lim_{n\to\infty} \underline{S}(f, T_n) = \frac{1}{4}.$$

Somit ist f integrierbar, und

$$\int_0^1 x^3 \, dx = \frac{1}{4}. \qquad \triangleleft$$

Beispiel. Es sei

$$f(x) = \begin{cases} 1, & \text{wenn } x \in \mathbb{Q}, \\ 0, & \text{wenn } x \notin \mathbb{Q}. \end{cases}$$

Da jedes Intervall positiver Länge sowohl rationale als auch irrationale Zahlen enthält, ist

$$\sup_{[x_{k-1}, x_k]} f = 1, \qquad \inf_{[x_{k-1}, x_k]} f = 0,$$

so dass

$$\overline{\int_a^b} f(x) \, dx = b - a, \qquad \underline{\int_a^b} f(x) \, dx = 0.$$

Somit ist f nicht integrierbar. $\triangleleft$

Wir geben nun zwei allgemeine Kriterien für die Integrierbarkeit an.

Satz 7.1. *Jede stetige Funktion auf einem beschränkten abgeschlossenen Intervall ist integrierbar.*

Beweis. Es sei $\varepsilon > 0$ gegeben. Nach Satz 5.13 ist f gleichmäßig stetig, es gibt also ein $\delta > 0$, so dass für x und y aus $[a, b]$ mit der Eigenschaft $|x - y| < \delta$ gilt $|f(x) - f(y)| < \varepsilon$. Ist nun T eine Teilung mit einer Feinheit kleiner als δ, so gilt das Gesagte insbesondere für x und y aus dem selben Teilintervall $[x_{k-1}, x_k]$, so dass

$$\sup_{[x_{k-1},x_k]} f - \inf_{[x_{k-1},x_k]} f \leq \varepsilon.$$

Multiplizieren wir dies mit $x_k - x_{k-1}$ und summieren über k, so erhalten wir

$$\overline{S}(f,T) - \underline{S}(f,T) \leq \sum_{k=1}^{n} \varepsilon(x_k - x_{k-1}) = \varepsilon(b - a).$$

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

Satz 7.2. *Jede beschränkte monotone Funktion auf einem beschränkten abgeschlossenen Intervall ist integrierbar.*

Beweis. Ist f beispielsweise monoton wachsend, so gilt

$$\inf_{[x_{k-1},x_k]} f = f(x_{k-1}), \qquad \sup_{[x_{k-1},x_k]} f = f(x_k).$$

Multiplizieren wir dies mit $x_k - x_{k-1}$ und summieren über k, so erhalten wir

$$\overline{S}(f,T) - \underline{S}(f,T) = \sum_{k=1}^{n} \big(f(x_k) - f(x_{k-1})\big)(x_k - x_{k-1}),$$

wobei alle Terme nichtnegativ sind. Hat T eine Feinheit kleiner als δ, so ist die rechte Seite beschränkt durch

$$\sum_{k=1}^{n} (f(x_k) - f(x_{k-1}))\,\delta = \big(f(x_n) - f(x_0)\big)\delta = \big(f(b) - f(a)\big)\delta.$$

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

7.2 Die Additivität bezüglich des Integrationsintervalls

Vorlesung 1b
14.10.2022

Abweichend von unseren bisherigen Formulierungen lässt man zu, dass das Integrationsintervall $[a, b]$ nicht gleich dem Definitionsbereich D des Integranden f ist, solange es in D enthalten ist. Anstelle von „$[a, b] \subseteq D$ und $f|_{[a,b]}$ ist integrierbar“ sagt man traditionell „f ist über das Intervall $[a, b]$ integrierbar“. Die entsprechende Bezeichnung „Integral von f über das Intervall $[a, b]$“ haben wir bereits in Definition 7.1 benutzt.

Satz 7.3. *Es sei $a < b < c$. Eine beschränkte Funktion $f : [a, c] \to \mathbb{R}$ ist genau dann integrierbar, wenn die Funktionen $f|_{[a,b]}$ und $f|_{[b,c]}$ integrierbar sind, und in diesem Fall gilt*

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx.$$

Beweis. Ist T eine Teilung von $[a, b]$ und $T' = T \cup \{b\}$, so erhalten wir Teilungen $T_1 = T' \cap [a, b]$ und $T_2 = T' \cap [b, c]$ von $[a, b]$ bzw. $[b, c]$. Geht man umgekehrt von Teilungen T_1 von $[a, b]$ und T_2 von $[b, c]$ aus, so ist $T' = T_1 \cup T_2$ eine Teilung von $[a, c]$. In jedem Fall gilt

$$\overline{S}(f, T_1) + \overline{S}(f, T_2) = \overline{S}(f, T'), \qquad \underline{S}(f, T_1) + \underline{S}(f, T_2) = \underline{S}(f, T').$$

Bilden wir die unteren bzw. oberen Grenzen über alle Teilungen T_1, T_2 oder auch über alle Teilungen T, so folgt mit Lemma 7.1

$$\overline{\int_a^c} f(x)\,dx = \overline{\int_a^b} f(x)\,dx + \overline{\int_b^c} f(x)\,dx, \quad \underline{\int_a^c} f(x)\,dx = \underline{\int_a^b} f(x)\,dx + \underline{\int_b^c} f(x)\,dx.$$

Ist f über $[a, b]$ und über $[b, c]$ integrierbar, so stimmen die rechten Seiten überein, also auch die linken Seiten, das heißt, f ist integrierbar. Die Umkehrung wird klar, wenn man die beiden Gleichungen voneinander abzuziehen und bemerkt, dass die Differenz aus Oberintegral und Unterintegral nicht negativ ist. □

Unter Berücksichtigung der Sätze 7.1 und 7.2 ergibt sich:

Folgerung 7.1. *Stückweise stetige Funktionen (d. h. Funktionen mit endlich vielen Unstetigkeitsstellen, in denen aber einseitige Grenzwerte existieren) und stückweise monotone beschränkte Funktionen auf beschränkten abgeschlossenen Intervallen sind integrierbar.*

Bisher musste die obere Integrationsgrenze größer als die untere sein.

Definition 7.3. *Für eine integrierbare Funktion f auf $[a, b]$ setzen wir*

$$\int_a^a f(x)\,dx = 0, \qquad \int_b^a f(x)\,dx = -\int_a^b f(x)\,dx.$$

Dies ist die einzige Möglichkeit, damit gilt (Übungsaufgabe):

Folgerung 7.2. *Satz 7.3 gilt ohne die Voraussetzung $a < b < c$, wenn f auf dem größten vorkommenden Intervall integrierbar ist.*

7.3 Das Riemannsche Integral für komplexwertige Funktionen

Unsere Definition des Integrals ist nicht auf komplexwertige Funktionen anwendbar, da der Körper der komplexen Zahlen keine Anordnung besitzt. Man könnte sich aus der Affäre ziehen, indem man Real- und Imaginärteil getrennt betrachtet. Wir wählen aber einen anderen Zugang, mit dem sich im folgenden Abschnitt die Eigenschaften des Integrals besser beweisen lassen, und der sich auf vektorwertige Funktionen verallgemeinert.

Definition 7.4. *Eine Menge* $Z = \{z_1, z_2, \ldots, z_n\}$ *heißt Menge von* Stützstellen *für die Teilung* $T = \{x_0, x_1, \ldots, x_n\}$ *des Intervalls* $[a, b]$*, wenn*

$$x_0 \leq z_1 \leq x_1 \leq \ldots \leq z_n \leq x_n.$$

Ist f *eine Funktion auf dem Intervall* $[a, b]$ *mit Werten in einem Vektorraum über* $\mathbb{R}$*, so nennen wir*

$$S(f, T, Z) = \sum_{k=1}^{n} f(z_k)(x_k - x_{k-1})$$

die Riemannsche Zwischensumme *(oder einfach Riemannsche Summe) von* f *bezüglich* T *und* Z*.*

Man kann dies natürlich auch in dem Fall anwenden, dass die Werte von f reelle Zahlen sind. Dann liegt die Zwischensumme zwischen der Ober- und der Untersumme, was ihren Namen erklärt.

Satz 7.4. *Für eine Funktion* $f : [a, b] \to \mathbb{R}$ *und eine Zahl* $I \in \mathbb{R}$ *gilt genau dann*

$$\int_a^b f(x)\, dx = I,$$

wenn es für jedes $\varepsilon > 0$ *ein* $\delta > 0$ *gibt, so dass für jede Teilung* T *von* $[a, b]$ *mit einer Feinheit kleiner als* δ *und für jede Menge* Z *von Stützstellen für* T *gilt*

$$|S(f, T, Z) - I| < \varepsilon.$$

Beweis. Angenommen, für ein gegebenes $\varepsilon > 0$ gibt es ein $\delta > 0$ mit der besagten Eigenschaft, und es sei T eine Teilung, deren Feinheit kleiner als δ ist. Für jedes k gibt es ein $z_k \in [x_{k-1}, x_k]$, so dass

$$f(z_k) > \sup_{[x_{k-1}, x_k]} f - \frac{\varepsilon}{b - a}.$$

Dann ist $S(f,T,Z) > \overline{S}(f,T) - \varepsilon$ und somit $\overline{S}(f,T) < I + 2\varepsilon$. Ähnlich zeigt man, das $\underline{S}(f,T) > I - 2\varepsilon$. Da wir $\varepsilon > 0$ beliebig wählen können, folgt $\int_a^b f(x)\,dx = I$.

Umgekehrt sei $\int_a^b f(x)\,dx = I$, und es sei ein $\varepsilon > 0$ gegeben. Dann gibt es eine Teilung T', so dass

$$\overline{S}(f,T') < I + \varepsilon, \qquad \underline{S}(f,T') > I - \varepsilon. \tag{13}$$

Wir schreiben $T' = \{x'_0, \ldots, x'_{n'}\}$ mit $a = x'_0 < \ldots < x'_{n'} = b$. Nun sei T eine beliebige Teilung und Z eine Menge von Stützstellen dafür. Liegt ein Teilintervall $[x_{k-1}, x_k]$ von T innerhalb eines Teilintervalls $[x'_{l-1}, x'_l]$ von T', so gilt

$$\inf_{[x'_{l-1},x'_l]} f \le f(z_k) \le \sup_{[x'_{l-1},x'_l]} f.$$

Dies ist für höchstens $n' - 1$ Teilintervalle von T nicht erfüllt, nämlich diejenigen, die ein Element von T' im Inneren enthalten. Für beliebige k und l gilt zumindest

$$\inf_{[x'_{l-1},x'_l]} f - c \le f(z_k) \le \sup_{[x'_{l-1},x'_l]} f + c,$$

wobei $c = \sup f - \inf f$. Multiplizieren wir dies mit den Längen L_{kl} der jeweiligen Intervalle $[x_{k-1}, x_k] \cap [x'_{l-1}, x'_l]$, so erhalten wir nach Summation über k und l wegen

$$\sum_{k=1}^{n} L_{kl} = x_k - x_{k-1}, \qquad \sum_{l=1}^{n'} L_{kl} = x'_l - x'_{l-1},$$

dass

$$\underline{S}(f,T') - n'c\delta \le S(f,T,Z) \le \overline{S}(f,T') + n'c\delta.$$

Ist nun $\delta < \frac{\varepsilon}{n'c}$, so folgt unter Berücksichtigung der Ungleichungen (13), dass $|S(f,T,Z) - I| < 2\varepsilon$. □

Nun kommen wir zu komplexwertigen Funktionen.

Definition 7.5. *Eine Funktion $f : [a,b] \to \mathbb{C}$ heißt integrierbar, wenn es eine Zahl $I \in \mathbb{C}$ gibt, die der Bedingung aus Satz 7.4 genügt, und in diesem Fall nennt man I das Integral von f über das Intervall $[a,b]$, abgekürzt $I = \int_a^b f(x)\,dx$.*

Man beachte, dass sich dieses Integral nicht mehr als (vorzeichenbehafteter) Flächeninhalt deuten lässt. Satz 7.4 zeigt, dass beide Definitionen kompatibel sind. Außerdem erhalten wir mit Hilfe der Ungleichungen (3) auf S. 46 wegen $\operatorname{Re} S(f,T,Z) = S(\operatorname{Re} f, T, Z)$ und $\operatorname{Im} S(f,T,Z) = S(\operatorname{Im} f, T, Z)$:

Folgerung 7.3. *Eine Funktion $f : [a,b] \to \mathbb{C}$ ist genau dann integrierbar, wenn die Funktionen $\operatorname{Re} f$ und $\operatorname{Im} f$ integrierbar sind, und dann gilt*

$$\operatorname{Re} \int_a^b f(x)\,dx = \int_a^b \operatorname{Re} f(x)\,dx, \qquad \operatorname{Im} \int_a^b f(x)\,dx = \int_a^b \operatorname{Im} f(x)\,dx.$$

Damit übertragen sich die Sätze 7.1 und 7.3 auf komplexwertige Funktionen. Man könnte sie aber auch direkt beweisen.

7.4 Eigenschaften des Integrals

Vorlesung 2a
19.10.22

Für das Integral gelten naheliegende Rechenregeln.

Satz 7.5. *Sind die Funktionen f, $g : [a,b] \to \mathbb{C}$ integrierbar und ist $c \in \mathbb{C}$, so sind auch $f+g$, $c \cdot f$ und $|f|$ integrierbar, und es gilt*

(a) $$\int_a^b \big(f(x)+g(x)\big)dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx,$$

(b) $$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx,$$

(c) $$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx.$$

Beweis. Offensichtlich gilt für jede Teilung T von $[a,b]$ und jede Menge Z von Stützstellen für T

$$S(f+g,T,Z) = S(f,T,Z) + S(g,T,Z)$$

(was für Ober- und Untersummen falsch gewesen wäre). Mit der Dreiecksungleichung folgt

$$
\begin{aligned}
&\left| S(f+g,T,Z) - \int_a^b f(x)\,dx - \int_a^b g(x)\,dx \right| \\
&\qquad \le \left| S(f,T,Z) - \int_a^b f(x)\,dx \right| + \left| S(g,T,Z) - \int_a^b g(x)\,dx \right|.
\end{aligned}
$$

Wegen der Integrierbarkeit von f und g finden wir für jedes $\varepsilon > 0$ ein $\delta > 0$, so dass die rechte Seite kleiner als ε ist, wenn die Feinheit von T kleiner als δ ist. Somit folgt die Integrierbarkeit von $f+g$ und die Formel (a). Analog behandelt man cf.

Aus der Dreiecksungleichung folgt

$$|f(x)| - |f(y)| \le |f(x) - f(y)| \le |\operatorname{Re} f(x) - \operatorname{Re} f(y)| + |\operatorname{Im} f(x) - \operatorname{Im} f(y)|.$$

Bilden wir das Supremum über alle x, $y \in [a,b]$, so folgt wegen $|r-s| = \max\{r-s, s-r\}$ für reelle Zahlen r und s

$$\sup|f| - \inf|f| \le (\sup \operatorname{Re} f - \inf \operatorname{Re} f) + (\sup \operatorname{Im} f - \inf \operatorname{Im} f).$$

Das Selbe gilt für die Einschränkung von f auf ein Teilintervall von T, so dass

$$\overline{S}(|f|,T) - \underline{S}(|f|,T) \le \big(\overline{S}(\operatorname{Re} f,T) - \underline{S}(\operatorname{Re} f,T)\big) + \big(\overline{S}(\operatorname{Im} f,T) - \underline{S}(\operatorname{Im} f,T)\big).$$

Ist also f integrierbar, so auch $|f|$. Für jedes $\varepsilon > 0$ finden wir eine Teilung T, so dass

$$\left| S(f,T,Z) - \int_a^b f(x)\,dx \right| < \varepsilon, \qquad \left| S(|f|,T,Z) - \int_a^b |f(x)|\,dx \right| < \varepsilon,$$

und folglich

$$\left| \int_a^b f(x)\,dx \right| - \varepsilon < |S(f,T,Z)|, \qquad S(|f|,T,Z) < \int_a^b |f(x)|\,dx + \varepsilon.$$

Die Anwendung der Dreiecksungleichung auf die Riemannschen Summen ergibt

$$|S(f,T,Z)| \le S(|f|,T,Z).$$

Da $\varepsilon > 0$ beliebig war, folgt die Ungleichung (c). □

Für reellwertige Funktionen hat das Integral Monotonieeigenschaften.

Satz 7.6. *Ist $f : [a, b] \to \mathbb{R}$ integrierbar und $f(x) \geq 0$ für alle $x \in [a, b]$, so gilt*

$$\int_a^b f(x)\,dx \geq 0.$$

Ist f außerdem stetig und gibt es ein $c \in [a, b]$, so dass $f(c) > 0$, so gilt hier die strenge Ungleichung.

Beweis. Auf jedem Teilintervall einer Teilung T gilt

$$\sup_{[x_{k-1}, x_k]} f \geq 0,$$

so das $\overline{S}(f, T) \geq 0$, und durch Bildung des Supremums folgt die erste Behauptung. Ist außerdem $f(c) = \varepsilon > 0$, so gibt es im Fall der Stetigkeit ein $\delta > 0$, so dass für x mit der Eigenschaft $|x - c| < \delta$ gilt $f(x) > \frac{\varepsilon}{2}$. Definieren wir $g : [a, b] \to \mathbb{R}$ durch

$$g(x) = \begin{cases} \frac{\varepsilon}{2} & \text{für } |x - c| < \delta, \\ 0 & \text{andernfalls,} \end{cases}$$

so ist $f - g \geq 0$, also nach dem Bewiesenen $\int_a^b f(x)\,dx \geq \int_a^b g(x)\,dx = \varepsilon\delta$. $\square$

Der folgende Begriff kommt häufig in Anwendungen vor.

Definition 7.6. *Der* Mittelwert *einer integrierbaren Funktion f auf einem beschränkten abgeschlossenen Intervall $[a, b]$ positiver Länge ist*

$$\frac{1}{b - a} \int_a^b f(x)\,dx.$$

Satz 7.7 (verallgemeinerter Mittelwertsatz der Integralrechnung)**.** *Sind f, $g : [a, b] \to \mathbb{R}$ stetig und gilt $g(x) > 0$ für alle $x \in [a, b]$, so gibt es ein $c \in \,]a, b[$, so dass*

$$\int_a^b f(x)g(x)\,dx = f(c) \int_a^b g(x)\,dx.$$

Für die Funktion $g(x) = 1$ erhält man den gewöhnlichen Mittelwertsatz, der besagt, dass eine stetige Funktion ihren Mittelwert annimmt.

Beweis. Nach Satz 5.6 existieren $m = \min f$ und $M = \max f$, und aus Satz 7.6 folgt, dass

$$m \int_a^b g(x)\, dx \leq \int_a^b f(x)g(x)\, dx \leq M \int_a^b g(x)\, dx.$$

Nach Satz 7.6 ist $\int_a^b g(x)\, dx > 0$, also existiert die Zahl

$$d = \int_a^b f(x)g(x)\, dx \Bigg/ \int_a^b g(x)\, dx.$$

Nach dem Bewiesenen liegt sie im Intervall $[m, M]$, und nach Satz 5.7 wird sie von f angenommen.

Es bleibt zu zeigen, dass man c verschieden von a und b wählen kann. Ist $d = m$ oder $d = M$, so zeigt die Anwendung von Satz 7.6 auf die Funktion $(f - m)g$ bzw. $(M - f)g$, dass diese verschwinden und somit f konstant sein muss, und dann ist c in $[a, b]$ beliebig. Andernfalls ist $m < d < M$. Nach der Definition von Minimum und Maximum existieren x, $y \in [a, b]$, so dass $f(x) = m$ und $f(y) = M$. Nach Satz 5.7 gibt es c im Intervall mit den Grenzen x und y, so dass $f(c) = d$ ist. Nun ist c verschieden von x und y, also auch von a und b. $\square$

7.5 Unbestimmte Integrale

Vorlesung 2b
21.10.22

In gewissem Sinne ist die Integration die Umkehroperation der Differentiation.

Satz 7.8. *Es sei I ein Intervall positiver Länge, $f : I \to \mathbb{R}$ eine stetige Funktion und $a \in I$. Dann ist die durch*

$$F(x) = \int_a^x f(t)\, dt$$

definierte Funktion $F : I \to \mathbb{R}$ differenzierbar, und $F' = f$.

Beweis. Wir halten $x \in I$ fest. Ist $h \in \mathbb{R}$ derart, dass $x + h \in I$, so gilt nach Satz 7.3

$$F(x + h) - F(x) = \int_x^{x+h} f(t)\, dt,$$

also

$$F(x+h) - F(x) - hf(x) = \int\limits_x^{x+h} \big(f(t) - f(x)\big)dt.$$

Ist $\varepsilon > 0$ gegeben, so gibt es wegen der Stetigkeit von f an der Stelle x ein $\delta > 0$, so dass für $|t - x| < \delta$ gilt $|f(t) - f(x)| < \varepsilon$. Mit Satz 7.5, verallgemeinert auf den Fall $b < a$, und Satz 7.6 erhalten wir für $|h| < \delta$

$$|F(x+h) - F(x) - hf(x)| \leq \left| \int\limits_x^{x+h} |f(t) - f(x)|\, dt \right| \leq |h|\varepsilon,$$

also für $0 < |h| < \delta$

$$\left| \frac{F(x+h) - F(x)}{h} - f(x) \right| \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt

$$\frac{F(x+h) - F(x)}{h} \to f(x) \qquad (h \to 0). \qquad \square$$

Nun wird die Definition 6.5 mit Leben erfüllt.

Folgerung 7.4. *Jede stetige Funktion besitzt eine Stammfunktion.*

Satz 7.9 (Hauptsatz der Differential- und Integralrechnung[40])**.** *Ist $f : I \to \mathbb{R}$ eine stetige Funktion und F eine Stammfunktion von f, so gilt für Elemente a und b des Intervalls I*

$$\int\limits_a^b f(x)\, dx = F(b) - F(a).$$

Beweis. Sind F und G zwei Stammfunktionen, so ist wegen Folgerung 6.7 $F(b) - F(a) = G(b) - G(a)$. Es genügt also, den Satz für eine einzige Stammfunktion zu beweisen. Für die Stammfunktion aus Satz 7.8 ist die Behauptung aber offensichtlich. $\square$

Beispiel. Wir betrachten die Funktion $f(x) = x^s$, die für $s \in \mathbb{N}$ auf der Menge $\mathbb{R}$, für $s \in \mathbb{Z} \setminus \mathbb{N}$ auf $\mathbb{R} \setminus \{0\}$ und andernfalls auf $]0, \infty[$ definiert ist. Für $s \in \mathbb{R}_+ \setminus \mathbb{N}$ lässt sie sich stetig auf $[0, \infty[$ fortsetzen. Man prüft leicht

[40] Statt „Differential- und Integralrechnung“ sagt man traditionell auch „Infinitesimalrechnung“.

nach, dass für $s \neq -1$ die Funktion $F(x) = \frac{x^{s+1}}{s+1}$ eine Stammfunktion von f ist, also gilt für $[a, b]$ im Definitionsbereich

$$\int_a^b x^s \, dx = \frac{b^{s+1} - a^{s+1}}{s+1}.$$

Im Fall $s = -1$ ist $\ln x$ eine Stammfunktion auf $]0, \infty[$ und $\ln(-x)$ eine Stammfunktion auf $]-\infty, 0[$. $\triangleleft$

Bemerkung. Man bezeichnet eine beliebige Stammfunktion der Funktion f mit dem Integralzeichen ohne Integrationsgrenzen und nennt dies das *unbestimmte Integral* von f, also

$$\int f(x) \, dx = F(x) + C$$

mit einer Konstanten C. Ist $F(x)$ durch einen umfangreichen Term gegeben, so vermeidet man das zweimalige Hinschreiben, indem man die Abkürzungen

$$F(b) - F(a) = F\Big|_a^b = \Big[F\Big]_a^b$$

vereinbart. Steht hier statt F ein Term $F(x)$, dann müsste man eigentlich $F(x)\big|_{x=a}^{x=b}$ schreiben (besonders wenn weitere Variablen vorkommen).

Beispiel. Man findet oft die Angabe

$$\int \frac{dx}{x} = \ln|x| + C \qquad \text{für } x \neq 0,$$

aber eigentlich kann man auf den beiden Komponenten des Definitionsbereiches verschiedene Konstanten wählen. Haben a und b das selbe Vorzeichen, so gilt

$$\int_a^b \frac{dx}{x} = \Big[\ln|x|\Big]_a^b.$$

Indem man die Liste der bekannten Ableitungen umkehrt, erhält man außerdem die weiteren so genannten *elementaren Integrale*

$$
\begin{aligned}
\int e^x dx &= e^x + C,\\
\int \sin x \, dx &= -\cos x + C,\\
\int \cos x \, dx &= \sin x + C,\\
\int \frac{dx}{\sqrt{1-x^2}} &= \arcsin x + C \qquad \text{für } |x| < 1,\\
\int \frac{dx}{\sqrt{x^2-1}} &= \operatorname{arcosh} x + C \qquad \text{für } x > 1,\\
\int \frac{dx}{\sqrt{1+x^2}} &= \operatorname{arsinh} x + C,\\
\int \frac{dx}{1+x^2} &= \arctan x + C,\\
\int \frac{dx}{1-x^2} &= \operatorname{artanh} x + C \qquad \text{für } |x| < 1,\\
\int \frac{dx}{1-x^2} &= \operatorname{arcoth} x + C \qquad \text{für } |x| > 1,\\
\int \frac{dx}{\cos^2 x} &= \tan x + C \qquad \text{für } x \notin \left\{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\right\},\\
\int \frac{dx}{\sin^2 x} &= -\cot x + C \qquad \text{für } x \notin \{k\pi \mid k \in \mathbb{Z}\}.
\end{aligned}
$$

Genaugenommen gelten die Formeln wieder nur auf jedem Teilintervall der Definitionsbereiche. Die unbekannten Funktionsnamen werden am Ende des folgenden Abschnitts erklärt. $\triangleleft$

7.6 Integrationstechniken

Vorlesung 3a
26.10.22

Aus den in Satz 6.1 bewiesenen Regeln der Differentiation ergibt sich mit Hilfe der Definition 6.5

Folgerung 7.5. *Sind f und g stetige Funktionen auf einem Intervall I und ist c eine Zahl, so gilt*

$$\int (f(x) + g(x)) \, dx = \int f(x) \, dx + \int g(x) \, dx, \qquad \int c\, f(x) \, dx = c \int f(x) \, dx$$

bei geeigneter Wahl der jeweiligen Stammfunktionen.

Natürlich spielt auch Folgerung 7.4 aus Satz 7.8 eine Rolle, die die Schreibweise des unbestimmten Integrals rechtfertigt. Wenden wir Satz 7.9 auf die obigen Gleichungen an, so ergibt sich wiederum Satz 7.5(a,b) im Spezialfall stetiger Funktionen.

Man beachte, dass wir die Produktregel nur in dem Spezialfall benutzt haben, in dem einer der beiden Faktoren konstant ist. Mit Hilfe des Hauptsatzes lassen sich auch die Kettenregel und die allgemeine Produktregel der Differentiation in Aussagen über Stammfunktionen ummünzen.

Satz 7.10 (Substitutionsregel)**.** *Es sei f auf einem Intervall I stetig und $g : [a, b] \to I$ stetig differenzierbar. Dann gilt*

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(y)\,dy.$$

Beweis. Es sei F eine Stammfunktion von f. Nach der Kettenregel ist

$$(F \circ g)'(x) = F'(g(x))g'(x) = f(g(x))g'(x),$$

also nach dem Hauptsatz

$$\int_a^b f(g(x))g'(x)\,dx = F \circ g\Big|_a^b = F\Big|_{g(a)}^{g(b)} = \int_{g(a)}^{g(b)} f(y)\,dy. \qquad \square$$

Ist $g(x)$ durch einen Term gegeben, so wird traditionell die Abbildung g nicht erwähnt, sondern man sagt, dass y in $f(y)$ durch diesen Term ersetzt (substituiert) wird. In der Leibnizschen Schreibweise ist dann $g'(x) = \frac{dy}{dx}$, und man schreibt formal $dy = g'(x)dx$, was man für dy substituieren kann. Auf diese Weise ergibt sich die Substitutionsregel von selbst, wenngleich das kein Beweis ist.

Beispiel. Wir wollen $\int_a^b c^x dx$ berechnen, wobei $c > 0$. Substituieren wir

$$y = x \ln c, \qquad \frac{dy}{dx} = \ln c,$$

so ist der Integrand $c^x = e^y$. Die (konstante) Ableitung kommt zwar in unserem Integral nicht vor, aber für $c \neq 1$ ist $dx = \frac{1}{\ln c} dy$, so dass

$$\int_a^b c^x\,dx = \frac{1}{\ln c}\int_{a \ln c}^{b \ln c} e^y\,dy = \frac{e^y}{\ln c}\Big|_{a \ln c}^{b \ln c} = \frac{c^b - c^a}{\ln c}.$$

Natürlich hätte man auch einfach bemerken können, dass $\frac{c^x}{\ln c}$ eine Stammfunktion von c^x ist. $\triangleleft$

Beispiel. Mit der Methode der quadratischen Ergänzung findet man

$$x^2 - 4x + 7 = (x-2)^2 + 3 = 3\left(1 + \left(\frac{x-2}{\sqrt{3}}\right)^2\right).$$

Mit der Substitution

$$y = \frac{x-2}{\sqrt{3}}, \qquad \frac{dy}{dx} = \frac{1}{\sqrt{3}}$$

erhalten wir

$$\int_a^b \frac{dx}{x^2 - 4x + 7} = \int_{\frac{a-2}{\sqrt{3}}}^{\frac{b-2}{\sqrt{3}}} \frac{\sqrt{3}\,dy}{3(1+y^2)} = \frac{1}{\sqrt{3}} \arctan y \bigg|_{\frac{a-2}{\sqrt{3}}}^{\frac{b-2}{\sqrt{3}}}. \qquad \triangleleft$$

Beispiel. Ist $f(y) = \frac{1}{y}$, so wird die Substitutionsregel zu

$$\int_a^b \frac{g'(x)}{g(x)}\,dx = \int_{g(a)}^{g(b)} \frac{dy}{y} = \big[\ln|y|\big]_{g(a)}^{g(b)},$$

vorausgesetzt, $g(x) \neq 0$ für $x \in [a, b]$. So ergibt sich beispielsweise

$$\int_a^b \tan x\,dx = -\int_a^b \frac{\cos' x\,dx}{\cos x} = -\int_{\cos a}^{\cos b} \frac{dy}{y} = \ln|\cos a| - \ln|\cos b|,$$

vorausgesetzt, die Tangensfunktion ist auf $[a, b]$ definiert. $\triangleleft$

Im Beweis von Satz 7.10 haben wir die Stammfunktion $f \circ g$ gefunden. Zur Bestimmung einer Formel müssen wir in dem Term $f(y)$ die Variable y durch den Term $g(x)$ substituieren. Im letzten Beispiel erhalten wir so

$$\int \tan x\,dx = \ln|\cos x| + C.$$

Satz 7.11 (partielle Integration)**.** *Sind f und g auf einem Intervall I stetig differenzierbar, so gilt für a, $b \in I$*

$$\int_a^b f(x)g'(x)\,dx = f(x)g(x)\Big|_a^b - \int_a^b f'(x)g(x)\,dx.$$

Beweis. Nach dem Hauptsatz gilt

$$\int_a^b (fg)'(x)\,dx = f(x)g(x)\Big|_a^b,$$

und nach der Produktregel ist der Integrand gleich $f'(x)g(x)+f(x)g'(x)$. $\square$

Beispiel. Es sei $s \neq -1$ und a, $b > 0$. Dann ist

$$\int_a^b x^s \ln x\,dx = \frac{x^{s+1}}{s+1}\ln x\Big|_a^b - \frac{1}{s+1}\int_a^b x^s\,dx = \frac{x^{s+1}}{(s+1)^2}\big((s+1)\ln x - 1\big)\Big|_a^b.$$

Ist $s > 0$, so gilt das durch Grenzübergang auch für $a = 0$. Im verbleibenden Fall $s = -1$ liefert die partielle Integration

$$\int_a^b \frac{\ln x}{x}\,dx = (\ln x)^2\Big|_a^b - \int_a^b \frac{\ln x}{x}\,dx.$$

Wir scheinen uns im Kreis gedreht zu haben, weil rechts wieder das gesuchte unbestimmte Integral auftaucht. Man kann es aber auf die linke Seite bringen und dann beide Seiten durch 2 dividieren. Das selbe Ergebnis erhält man auch mit der Substitutionsregel:

$$\int_a^b \frac{\ln x}{x}\,dx = \int_{\ln a}^{\ln b} y\,dy = \frac{y^2}{2}\Big|_{\ln a}^{\ln b}. \qquad \triangleleft$$

Man kann die partielle Integration natürlich auch zur Bestimmung von Stammfunktionen benutzen, denn der Beweis von Satz 7.11 ergibt

$$\int f(x)g'(x)\,dx = f(x)g(x) - \int f'(x)g(x)\,dx.$$

Beispiel. Es gilt

$$\int xe^x dx = xe^x - \int e^x dx = (x-1)e^x,$$
$$\int x^2e^x dx = x^2e^x - \int 2xe^x dx = \big(x^2 - 2(x-1)\big)e^x$$

usw. Mit der selben Methode, den Exponenten durch partielle Integration zu verringern, berechnet man auch $\int x^n \sin x\,dx$ und $\int x^n \cos x\,dx$. $\triangleleft$

Beispiel. Manchmal hilft es weiter, den Faktor 1 als Ableitung aufzufassen:

$$\begin{aligned}\int \sqrt{1-x^2}\,dx &= x\sqrt{1-x^2} - \int \frac{x(-2x)}{2\sqrt{1-x^2}}\,dx \\ &= x\sqrt{1-x^2} - \int \frac{1-x^2}{\sqrt{1-x^2}}\,dx + \int \frac{dx}{\sqrt{1-x^2}} \\ &= x\sqrt{1-x^2} - \int \sqrt{1-x^2}\,dx + \arcsin x + C.\end{aligned}$$

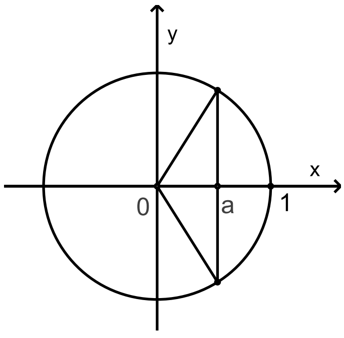
Bringen wir das letzte Integral auf die linke Seite und teilen durch 2, so erhalten wir die Antwort. Für $a,\, b \in\,]-1,1[$ gilt also

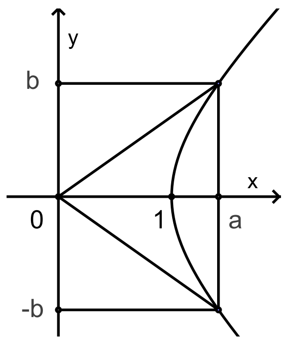
$$\int_a^b \sqrt{1-x^2}\,dx = \frac{1}{2}\big[x\sqrt{1-x^2} + \arcsin x\big]_a^b,$$

und durch Grenzübergang gilt dies sogar für $a,\, b \in [-1,1]$. Eine andere Methode besteht in der Substitution $x = \sin\varphi$. Setzen wir $b = 1$, so erhalten wir

$$a\sqrt{1-a^2} + 2\int_a^1 \sqrt{1-x^2}\,dx = \arcsin 1 - \arcsin a = \arccos a.$$

Man kann $\arccos a$ also als Fläche des Kreissektors interpretieren, dessen Ecken die Punkte des Einheitskreises mit Abszisse a sind. $\triangleleft$





In Analogie zu den Winkelfunktionen bildet man die *Hyperbelfunktionen*

$$\sinh x = \frac{e^x - e^{-x}}{2}, \qquad \cosh x = \frac{e^x + e^{-x}}{2},$$
$$\tanh x = \frac{\sinh x}{\cosh x}, \qquad \coth x = \frac{\cosh x}{\sinh x},$$

genannt[41] Hyperbelsinus, Hyperbelkosinus, Hyperbeltangens und Hyperbelkotangens, weil der Punkt mit den Koordinaten $\cosh x$ und $\sinh x$ wegen $\cosh^2 x - \sinh^2 x = 1$ auf einer Hyperbel liegt. Ihre Ableitungen sind

$$\begin{aligned} \sinh' x &= \cosh x = \sqrt{1+\sinh^2 x}, & \cosh' x &= \sinh x = \operatorname{sgn}(x)\sqrt{\cosh^2 x - 1}, \\ \tanh' x &= \frac{1}{\cosh^2 x} = 1 - \tanh^2 x, & \coth' x &= -\frac{1}{\sinh^2 x} = 1 - \coth^2 x. \end{aligned}$$

Ihre Umkehrfunktionen arsinh, arcosh (von $\cosh|_{[0,\infty[}$), artanh (definiert auf $]-1,1[$) und arcoth (definiert auf $]-\infty,-1[\,\cup\,]1,\infty[$) nennt man *Areafunktionen*[42], sie haben nach Satz 6.3 die Ableitungen

$$\begin{aligned} \operatorname{arsinh}' x &= \frac{1}{\sqrt{1+x^2}}, & \operatorname{arcosh}' x &= \frac{1}{\sqrt{x^2-1}}, \\ \operatorname{artanh}' x &= \frac{1}{1-x^2}, & \operatorname{arcoth}' x &= \frac{1}{1-x^2}, \end{aligned}$$

woraus sich die im vorigen Absatz erwähnten elementaren Integrale ergeben.

Beispiel. Ähnlich wie oben findet man

$$\int_a^b \sqrt{x^2+1}\,dx = \frac{1}{2}\big[x\sqrt{x^2+1} + \operatorname{arsinh} x\big]_a^b$$

sowie für a, $b \in [1,\infty[$

$$\int_a^b \sqrt{x^2-1}\,dx = \frac{1}{2}\big[x\sqrt{x^2-1} - \operatorname{arcosh} x\big]_a^b.$$

Die Spezialfälle

$$2\int_0^b \sqrt{y^2+1}\,dy - b\sqrt{b^2+1} = \operatorname{arsinh} b,$$

$$a\sqrt{a^2-1} - 2\int_1^a \sqrt{x^2-1}\,dx = \operatorname{arcosh} a$$

[41] lateinisch *sinus hyperbolicus, cosinus hyperbolicus, tangens hyperbolicus* und *cotangens hyperbolicus*

[42] lateinisch ursprünglich *area sinus hyperbolici, area cosinus hyperbolici, area tangentis hyperbolici* und *area cotangentis hyperbolici*

zeigen, dass man $\operatorname{arsinh} b$ und $\operatorname{arcosh} a$ als Fläche des Hyperbelsektors interpretieren kann, der von dem Bogen der Hyperbel $x^2 - y^2 = 1$ begrenzt wird, dessen Endpunkte die Abszisse a und die Ordinate $\pm b$ haben. Daher rührt die Bezeichnung Areafunktionen für die Umkehrungen der Hyperbelfunktionen. $\triangleleft$

7.6.1 Integration rationaler Funktionen

Vorlesung 3b
28.10.22

Eine Funktion heißt *rational*, wenn sie durch einen Term gegeben ist, in dem nur die vier Grundrechenoperationen vorkommen. Kommt dabei keine Division vor, so heißt sie *ganzrationale Funktion*. Jede solche Funktion kann als Quotient zweier ganzrationaler Funktionen geschrieben werden. Um zu zeigen, dass diese Funktionen elementare Stammfunktionen besitzen, benötigt man den

Satz 7.12 (Hauptsatz der Algebra)**.** *Jede ganzrationale Funktion einer Variablen mit komplexen Koeffizienten hat eine komplexe Nullstelle.*

Obwohl die Formulierung algebraisch erscheint, geht doch die analytische Eigenschaft der Vollständigkeit des Körpers der komplexen Zahlen ein. Der Beweis ist darum Gegenstand der Analysis. Wir werden ihn hier nicht führen, da im Rahmen der (komplexen) Funktionentheorie ein kurzer und eleganter Beweis gegeben werden kann.

Lemma 7.2. *Eine Zahl c ist genau dann Nullstelle einer ganzrationalen Funktion f, wenn es eine ganzrationale Funktion g gibt, so dass*

$$f(x) = (x - c)g(x).$$

Beweis. Definieren wir eine neue rationale Funktion durch $f_1(y) = f(c + y)$, dann können wir die ursprüngliche als $f(x) = f_1(x - c)$ zurückgewinnen. Ist $f_1(0) = 0$, dann gibt es eine ganzrationale Funktion g_1, so dass $f_1(y) = yg_1(y)$, und umgekehrt. $\square$

Man kann g auch durch Polynomdivision bestimmen. Nun kommen wir zu dem für die Integration wichtigen Ergebnis.

Satz 7.13. *Jede rationale Funktion R von einer Variablen z mit komplexen Koeffizienten kann auf eindeutige Weise als Summe einer ganzrationalen Funktion und von Funktionen der Form $\frac{C}{(z-c)^k}$ (genannt die* Partialbrüche *von R) mit $c, C \in \mathbb{C}$ und $k \in \mathbb{N} \setminus \{0\}$ geschrieben werden.*

Beweis. Es sei $R = \frac{f}{g}$ mit ganzrationalen Funktionen f und g. Nach dem Hauptsatz der Algebra hat g eine Nullstelle c. Nach dem Lemma gibt es eine natürliche Zahl $k \geq 1$ und eine ganzrationale Funktion h, so dass $g(z) = (z-c)^k h(z)$ und $h(c) \neq 0$. Für jedes $C \in \mathbb{C}$ gilt

$$\frac{f(z)}{g(z)} = \frac{C}{(z-c)^k} + \frac{f(z) - Ch(z)}{(z-c)^k h(z)}.$$

Wegen $h(c) \neq 0$ können wir C so wählen, dass $f(c) - Ch(c) = 0$, und dann ist nach dem Lemma

$$f(z) - Ch(z) = (z-c) f_1(z)$$

mit einer ganzrationalen Funktion f_1. Der Grad von $g_1(z) = (z-c)^{k-1} h(z)$ ist kleiner als der von g, und

$$\frac{f(z)}{g(z)} = \frac{C}{(z-c)^k} + \frac{f_1(z)}{g_1(z)}.$$

Nun folgt die Existenz der behaupteten Darstellung durch Induktion nach dem Grad von g.

Die Eindeutigkeit beweisen wir ebenfalls durch vollständige Induktion. Der Koeffizient C beim höchsten vorkommenden Exponenten k ist eindeutig bestimmt als

$$C = \lim_{z \to c} (z-c)^k \frac{f(z)}{g(z)}$$

und die Zahl k als die kleinste natürliche Zahl, für die dieser Grenzwert existiert. Laut Induktionsvoraussetzung ist die Partialbruchzerlegung von $\frac{f_1(z)}{g_1(z)}$ eindeutig bestimmt. $\square$

Beispiel. Um die Partialbruchzerlegung der rationalen Funktion

$$\frac{x^4 + 3x^3 - 6x + 5}{x^3 - 3x + 2}$$

zu finden, bestimmen wir zunächst die Nullstellen des Nenners:

$$x^3 - 3x + 2 = (x-1)^2 (x+2)$$

Nun wenden die Methode aus dem Beweis auf $c = 1$ und $k = 2$ an:

$$\frac{x^4 + 3x^3 - 6x + 5}{(x-1)^2 (x+2)} = \frac{1}{(x-1)^2} + \frac{x^3 + 4x^2 + 4x - 3}{(x-1)(x+2)}.$$

Auf den Rest können wir sie noch einmal anwenden:

$$\frac{x^3+4x^2+4x-3}{(x-1)(x+2)} = \frac{2}{x-1} + \frac{x^2+5x+7}{x+2}.$$

Eine nochmalige Anwendung (oder einfach Polynomdivision) ergibt

$$\frac{x^2+5x+7}{x+2} = \frac{1}{x+2} + x + 3.$$

Nun braucht man nur noch einzusetzen.

Anstatt die Methode aus dem Beweis zu benutzen, setzt man oft eine Partialbruchzerlegung mit unbestimmten Koeffizienten an und bestimmt letztere durch Beseitigung der Nenner und Koeffizientenvergleich. $\triangleleft$

Hat man die Partialbruchzerlegung bestimmt, so kann man jeden Partialbruch als möglicherweise komplexwertige Funktion einzeln integrieren. Da wir aber noch keine komplexe Logarithmusfunktion kennen, benutzen wir folgende Version der Partialbruchzerlegung.

Satz 7.14. *Jede rationale Funktion R von einer Variablen x mit reellen Koeffizienten kann auf eindeutige Weise als Summe einer ganzrationalen Funktion und von Funktionen der Formen*

$$\frac{C}{(x-c)^k}, \qquad \frac{Ax+B}{\big((x-a)^2+b^2\big)^k}$$

mit a, b, c, A, B, $C \in \mathbb{R}$ und $k \in \mathbb{N} \setminus \{0\}$ geschrieben werden, wobei $b > 0$.

Beweis. Aus $\overline{R(z)} = R(\bar{z})$ folgt dann wegen der Eindeutigkeit, dass mit jedem Partialbruch $\frac{C}{(z-c)^k}$ auch $\frac{\bar{C}}{(z-\bar{c})^k}$ auftritt. Für $c = a + ib$ mit $b > 0$ erhalten wir

$$\frac{C}{(x-c)^k} + \frac{\bar{C}}{(x-\bar{c})^k} = \frac{p(x)}{q(x)^k},$$

wobei p ein Polynom mit reellen Koeffizienten ist und

$$q(x) = (z-c)(z-\bar{c}) = (x-a)^2 + b^2.$$

Durch Polynomdivision findet man ein Polynom p_1 kleineren Grades und reelle Zahlen A und B, so dass

$$p(x) = p_1(x)q(x) + Ax + B.$$

Nun folgt die Behauptung durch vollständige Induktion nach k. $\square$

Partialbrüche mit Potenzen von q im Nenner integriert man, indem man den Zähler in ein Vielfaches von $x-a$ und eine Konstante aufteilt. Für $k=1$ gilt

$$\int \frac{x-a}{q(x)}\,dx = \frac{1}{2}\ln q(x) + C, \qquad \int \frac{dx}{q(x)} = \frac{1}{b}\arctan\frac{x-a}{b} + C,$$

wobei C diesmal eine Integrationskonstante bezeichnet. Die im linken Integral vorgenommene Substitution funktioniert auch für höheres k. Beim rechten verringert man den Exponenten mit Hilfe der Formel

$$2(k-1)b^2 \int \frac{dx}{q(x)^k} = \frac{x-a}{q(x)^{k-1}} + (2k-3)\int \frac{dx}{q(x)^{k-1}}.$$

7.6.2 Zurückführung auf Integrale rationaler Funktionen

Vorlesung 4a 2.11.22

Integrale von Verkettungen rationaler Funktionen R mit einigen weiteren Funktionen lassen sich durch geschickte Substitutionen auf Integrale rationaler Funktionen zurückführen. Dabei kann R eine Funktion von mehrerer Variablen sein, aber die Verkettung darf nur von einer Variablen abhängen. Wir schreiben nur die Version für unbestimmte Integrale auf. Manchmal liegt es näher, zunächst die Substitutionsvariable durch die vorhandene Variable auszudrücken, also in den Bezeichnungen von Satz 7.10 die Umkehrfunktion von g anzugeben.

(i) Bei Integralen der Form

$$\int R\big(x, \sqrt[n]{ax+b}\big)\,dx$$

mit $a \neq 0$ führt die Substitution

$$t = \sqrt[n]{ax+b}, \qquad x = \frac{t^n - b}{a}, \qquad dx = \frac{n}{a}\,t^{n-1}dt$$

auf das Integral einer rationalen Funktion

$$\frac{n}{a}\int R\left(\frac{t^n-b}{a}, t\right) t^{n-1}dt.$$

(ii) Bei Integralen der Form

$$\int R(e^{ax})\,dx$$

mit $a \neq 0$ führt die Substitution

$$t = e^{ax}, \qquad x = \frac{1}{a} \ln t, \qquad dx = \frac{dt}{at}$$

auf das Integral einer rationalen Funktion

$$\frac{1}{a} \int \frac{R(t)}{t}\, dt.$$

(iii) Bei Integralen der Form

$$\int R(\cos x, \sin x)\, dx$$

gibt es zwei einfach zu behandelnde Spezialfälle. Gilt $R(-u, v) = -R(u, v)$, dann gibt es eine rationale Funktion R_1, so dass

$$R(u, v) = R_1(u^2, v)u,$$

Gilt hingegen $R(u, -v) = -R(u, v)$, dann gibt es eine rationale Funktion R_2, so dass

$$R(u, v) = R_2(u, v^2)v.$$

Man substituiert

$$t = \sin x, \quad dt = \cos x\, dx \qquad \text{bzw} \qquad t = \cos x, \quad dt = \sin x\, dx$$

und erhält

$$\int R_1(1 - t^2, t)\, dt \qquad \text{bzw.} \qquad \int R_2(t, 1 - t^2)\, dt.$$

Im allgemeinen Fall hilft die Substitution

$$t = \tan \frac{x}{2}, \qquad dt = \frac{dx}{2 \cos^2 \frac{x}{2}}.$$

Wegen

$$\begin{aligned} 1 + t^2 &= \frac{\cos^2 \frac{x}{2} + \sin^2 \frac{x}{2}}{\cos^2 \frac{x}{2}} = \frac{1}{\cos^2 \frac{x}{2}}, \\ 1 - t^2 &= \frac{\cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}}{\cos^2 \frac{x}{2}} = \frac{\cos x}{\cos^2 \frac{x}{2}}, \\ 2t &= \frac{2 \sin \frac{x}{2} \cos \frac{x}{2}}{\cos^2 \frac{x}{2}} = \frac{\sin x}{\cos^2 \frac{x}{2}} \end{aligned}$$

ist nämlich

$$\sin x = \frac{2t}{1+t^2}, \qquad \cos x = \frac{1-t^2}{1+t^2}, \qquad dx = \frac{2dt}{1+t^2},$$

und wir erhalten das Integral einer rationalen Funktion

$$\int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right) \frac{2dt}{1+t^2}.$$

Natürlich muss die Substitutionsfunktion auf ein Intervall eingeschränkt werden, auf dem sie monoton und somit umkehrbar ist.

(iv) Bei Integralen der Form

$$\int R\big(x, \sqrt{ax^2+2bx+c}\,\big)\, dx$$

mit $a \neq 0$ bildet man im Radikanden zunächst die quadratische Ergänzung:

$$ax^2+2bx+c = \frac{(ax+b)^2+ac-b^2}{a} = A(\pm t^2 \pm 1),$$

wobei

$$A = \left|\frac{ac-b^2}{a}\right|, \qquad t = \frac{ax+b}{\sqrt{|ac-b^2|}}.$$

und durch die offensichtliche Substitution erhalten wir ein Integral in einer der Formen

$$\int R_3\big(t, \sqrt{t^2+1}\,\big)\, dt, \quad \int R_3\big(t, \sqrt{t^2-1}\,\big)\, dt, \quad \int R_3\big(t, \sqrt{1-t^2}\,\big)\, dt.$$

Die weitere Substitution

$$t = \sinh u, \qquad t = \pm\cosh u \qquad \text{bzw.} \qquad t = \cos u$$

führt dann auf ein schon unter (ii) bzw. (iii) behandeltes Integral.

7.7 Uneigentliche Integrale

Vorlesung 4b
4.11.22

Das Riemannsche Integral ist nur für beschränkte Funktionen auf beschränkten Intervallen definiert, da sonst die Ober- oder Untersummen nicht existieren.

Definition 7.7. *Es sei $-\infty \le a < b \le \infty$ und $f :]a,b[\to \mathbb{C}$, wobei f über jedes beschränkte abgeschlossene Teilintervall $[\alpha,\beta]$ integrierbar ist. Man nennt*

$$\lim_{\alpha\to a}\lim_{\beta\to b}\int\limits_{\alpha}^{\beta} f(x)\,dx$$

das uneigentliche Integral *von f über $]a,b[$, falls beide Grenzwerte existieren, und bezeichnet es mit $\int\limits_a^b f(x)\,dx$. An Stelle von „das uneigentliche Integral existiert" sagt man traditionell „dass das uneigentliche Integral ist konvergent".*

Ist f zu einer integrierbaren Funktion auf $[a,b]$ fortsetzbar (wofür $[a,b]$ und f beschränkt sein müssen), so ist das (eigentliche) Integral aufgrund von Satz 7.5(c) gleich dem uneigentlichen Integral, weshalb man für beide die selbe Schreibweise verwenden darf. Wegen Folgerung 5.7 und Satz 7.3 ist für jedes $c \in]a,b[$

$$\int\limits_a^b f(x)\,dx = \lim_{\alpha\to a}\int\limits_{\alpha}^{c} f(x)\,dx + \lim_{\beta\to b}\int\limits_{c}^{\beta} f(x)\,dx.$$

Das uneigentliche Integral über $]a,b[$ ist also genau dann konvergent, wenn die uneigentlichen Integrale über $]a,c[$ und über $]c,b[$ konvergent sind. Außerdem ist die Reihenfolge der Grenzwerte in der Definition gleichgültig.

Man nennt die Integrationsgrenze a *kritisch*, wenn sich f nicht zu einer integrierbaren Funktion auf $[a,c]$ fortsetzt, und analog für b. Die Grenze $\pm\infty$ ist offensichtlich immer kritisch.

Beispiel. Für $s \in \mathbb{C}$ ist

$$\int\limits_1^\infty \frac{dx}{x^s} = \lim_{\beta\to\infty}\int\limits_1^\beta \frac{dx}{x^s} = \lim_{\beta\to\infty}\begin{cases}\frac{1-\beta^{1-s}}{s-1} & \text{für } s \neq 1,\\ \ln\beta & \text{für } s = 1.\end{cases}$$

Das uneigentliche Integral konvergiert genau dann, wenn $\operatorname{Re} s > 1$, und zwar gegen $\frac{1}{s-1}$. Weiter gilt

$$\int\limits_0^1 \frac{dx}{x^s} = \lim_{\alpha\to 0}\int\limits_\alpha^1 \frac{dx}{x^s} = \lim_{\alpha\to 0}\begin{cases}\frac{1-\alpha^{1-s}}{1-s} & \text{für } s \neq 1,\\ -\ln\alpha & \text{für } s = 1.\end{cases}$$

Dieses uneigentliche Integral konvergiert genau dann, wenn $\operatorname{Re} s < 1$, und zwar gegen $\frac{1}{1-s}$. $\triangleleft$

Definition 7.8. *Das uneigentliche Integral* $\int\limits_a^b f(x)\,dx$ *heißt* absolut konvergent, *wenn das Integral* $\int\limits_a^b |f(x)|\,dx$ *konvergent ist.*

Satz 7.15. (i) *Ein absolut konvergentes Integral ist konvergent.*

(ii) *Sind* $f,\, g :]a,b[\to \mathbb{R}$ *und gilt* $0 \le f(x) \le g(x)$ *für alle* $x \in]a,b[$*, so folgt aus der Konvergenz von* $\int\limits_a^b g(x)\,dx$ *die von* $\int\limits_a^b f(x)\,dx$*.*

Beweis. **(i)** Es sei z. B. nur b kritisch. Wir betrachten

$$F(\beta) = \int\limits_a^\beta f(x)\,dx, \qquad H(\beta) = \int\limits_a^\beta |f(x)|\,dx.$$

Wenn $\lim\limits_{\beta\to b} H(\beta)$ existiert, so gibt es nach dem Cauchykriterium (Satz 5.12) für jedes $\varepsilon > 0$ ein $x_0 < b$, so dass für alle $\beta, \gamma \in]x_0, b[$ gilt $|H(\gamma) - H(\beta)| < \varepsilon$. Nach den Sätzen 7.3 und 7.5(c) ist

$$|F(\gamma) - F(\beta)| \le |H(\gamma) - H(\beta)|,$$

also existiert dann auch $\lim\limits_{\beta\to b} F(\beta)$.

(ii) Es sei z. B. nur b kritisch für f oder g. Mit der Bezeichnung $G(\beta) = \int\limits_a^\beta g(x)\,dx$ gilt nach Satz 7.6

$$0 \le F(\beta) \le G(\beta),$$

und aus der Beschränktheit von G folgt die von F. Ebenfalls nach Satz 7.6 ist F monoton wachsend, und die Konvergenz folgt mit dem Monotoniekriterium (Satz 5.11). □

Beispiel. Durch partielle Integration erhalten wir

$$\int\limits_1^\beta \frac{\sin x}{x}\,dx = -\left.\frac{\cos x}{x}\right|_1^\beta - \int\limits_1^\beta \frac{\cos x}{x^2}\,dx.$$

Das für $\beta = \infty$ uneigentliche Integral auf der rechten Seite ist nach Satz 7.15 absolut konvergent durch Vergleich mit $\int\limits_1^\infty \frac{dx}{x^2}$. Also ist das uneigentliche Integral

$$\int\limits_0^\infty \frac{\sin x}{x}\,dx = \int\limits_0^1 \frac{\sin x}{x}\,dx + \int\limits_1^\infty \frac{\sin x}{x}\,dx$$

(bei dem nur die Grenze ∞ kritisch ist) konvergent. Für jedes $k \in \mathbb{N}$ gilt aber

$$\int_{(k-1)\pi}^{k\pi} \left|\frac{\sin x}{x}\right| dx \geq \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\sin x|\, dx = \frac{2}{k\pi},$$

also

$$\int_0^{n\pi} \left|\frac{\sin x}{x}\right| dx \geq \frac{2}{\pi} \sum_{k=1}^{n} \frac{1}{k},$$

und wegen der Divergenz der harmonischen Reihe ist das uneigentliche Integral nicht absolut konvergent. $\triangleleft$

Satz 7.16 (Integralkriterium für Reihen)**.** *Es sei $n_0 \in \mathbb{N}$ und $f : [1, \infty[\to \mathbb{R}$ monoton fallend und nichtnegativ. Die Reihe $\sum\limits_{n=n_0}^{\infty} f(n)$ ist genau dann konvergent, wenn das Integral $\int\limits_{n_0}^{\infty} f(x)\, dx$ konvergent ist.*

Beweis. Nach Satz 7.2 ist f auf jedem beschränkten Teilintervall integrierbar. Für $x \in [k, k+1]$ gilt $f(k) \geq f(x) \geq f(k+1)$, also nach Satz 7.6

$$f(k) \geq \int_k^{k+1} f(x)\, dx \geq f(k+1),$$

$$\sum_{k=n_0}^{n-1} f(k) \geq \int_{n_0}^{n} f(x)\, dx \geq \sum_{k=n_0+1}^{n} f(k),$$

und die Behauptung folgt aus den Monotoniekriterien (Sätze 3.9 und 5.11). $\square$

Beispiel. Ist $s \in \mathbb{R}$, so erhalten wir mit der Substitution $y = \ln x$, $dy = \frac{dx}{x}$

$$\int_2^{\infty} \frac{dx}{x(\ln x)^s} = \int_{\ln 2}^{\infty} \frac{dy}{y^s},$$

was nach dem obigen Beispiel genau für $s > 1$ konvergent ist. Wir können den Satz mit $f(x) = \frac{1}{x(\ln x)^s}$ anwenden, denn

$$f'(x) = -\frac{s + \ln x}{x^2(\ln x)^{s+1}}$$

ist negativ für genügend große x. Also ist auch

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^s}$$

genau dann konvergent, wenn $s > 1$. $\triangleleft$

8 Metrische Räume

Vorlesung 5a
9.11.22

Nachdem wir in der Vorlesung Analysis I Funktionen auf Teilmengen eines Körpers K mit Absolutbetrag untersucht haben, wollen wir uns in der Vorlesung Analysis II mit Funktionen auf Teilmengen des Raumes

$$K^n = \underbrace{K \times K \times \ldots \times K}_{n}$$

der n-Tupel

$$x = (x_1, x_2, \ldots, x_n)$$

von Elementen x_i aus K befassen. Diese lassen sich addieren und mit Elementen aus K multiplizieren:

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n), \qquad a \cdot x = (a \cdot x_1, a \cdot x_2, \ldots, a \cdot x_n).$$

Die Menge K^n mit diesen Verknüpfungen ist ein Beispiel für den folgenden Begriff.

Definition 8.1. *Ein* Vektorraum *über einem Körper K ist eine Menge V versehen mit zwei Verknüpfungen $V \times V \to V$ und $K \times V \to V$, genannt Addition und Skalarmultiplikation und abgekürzt durch die Zeichen „+“ und „·“, die folgende Eigenschaften haben:*

(i) *Für alle x, y und z in V gilt*

$$x + y = y + x, \qquad (x + y) + z = x + (y + z).$$

(ii) *Es gibt ein Element 0_V in V, so dass für alle x in V gilt*

$$x + 0_V = x.$$

(iii) *Für beliebige Elemente x, y von V gibt es ein Element z in V, so dass*

$$x = y + z.$$

(iv) *Für alle a, b in K und alle x, y in V gilt*

$$(a + b) \cdot x = a \cdot x + b \cdot x, \qquad a \cdot (x + y) = a \cdot x + a \cdot y,$$
$$(a \cdot b) \cdot x = a \cdot (b \cdot x), \qquad 1_K \cdot x = x.$$

Das Element 0_V nennt man den Nullvektor in V, und das Element z in (iii) bezeichnet man mit $x - y$. Man kann sich überzeugen, dass sie jeweils eindeutig bestimmt sind. Statt $0_V - x$ schreibt man meist $-x$. Der Kürze halber bezeichnet man oft das Einselement 1_K mit 1 sowie die Nullelemente 0_K und 0_V mit 0, und man lässt das Multiplikationszeichen (außer zwischen Zahlen) weg.

Einen Vektorraum über einem Körper K nennt man auch einen K-Vektorraum. Vektorräume werden in der Vorlesung zur Linearen Algebra ausführlich behandelt, und dort wird auch der Begriff der Dimension eingeführt. Teilnehmer/innen, die diese Veranstaltung erst später belegen, mögen sich unter einem n-dimensionalen Vektorraum immer das obige Beispiel K^n vorstellen.

Es gibt aber auch unendlichdimensionale Vektorräume. So folgt aus den Sätzen 5.3, 6.1, 6.9 und 7.5, dass nicht nur die Menge aller Funktionen, sondern auch die Mengen der stetigen Funktionen, der differenzierbaren Funktionen, der k Mal differenzierbaren Funktionen sowie der integrierbaren Funktionen auf einem gegebenen Definitionsbereich jeweils einen Vektorraum bilden.

8.1 Normen

Es sei K der Körper $\mathbb{R}$ der reellen Zahlen oder der Körper $\mathbb{C}$ der komplexen Zahlen. Anstelle des Absolutbetrages benötigen wir im Fall einer beliebigen Zahl von Variablen den folgenden Begriff.

Definition 8.2. *Eine* Norm *auf einem K-Vektorraum V ist eine Abbildung $V \to \mathbb{R}$ (deren Wert an der Stelle x wir mit $\|x\|$ bezeichnen, gelesen „Norm von x“), die folgenden Eigenschaften für alle x, $y \in V$ und $a \in K$ hat:*

(i) $\|a \cdot x\| = |a| \cdot \|x\|$,

(ii) $\|x + y\| \leq \|x\| + \|y\|$ *(Dreiecksungleichung),*

(iii) wenn $\|x\| = 0$, *dann* $x = 0_V$.

Aus (i) folgt die Umkehrung von (iii), nämlich

$$\|0_V\| = \|0 \cdot 0_V\| = |0| \cdot \|0_V\| = 0,$$

und daraus sowie aus (ii) und (i) folgt

$$0 = \|x - x\| \leq \|x\| + \|-x\| = 2\|x\|,$$

so dass eine Norm nur nichtnegative Werte annehmen kann.

Der folgende Satz liefert Beispiele von Normen auf K^n.

Satz 8.1. *Für eine reelle Zahl* $p > 0$ *und für* $x \in K^n$ *sei*

$$\|x\|_p = \left(|x_1|^p + \ldots + |x_n|^p\right)^{1/p}, \qquad \|x\|_\infty = \lim_{p\to\infty} \|x\|_p.$$

Dies ist für alle $p \in [1, \infty]$ *eine Norm*[43]*. Für alle* p *und* q *mit der Eigenschaft* $\frac{1}{p} + \frac{1}{q} = 1$ *gilt die* Hölder-Ungleichung

$$\sum_{i=1}^{n} |x_i y_i| \leq \|x\|_p \|y\|_q.$$

Bemerkungen. Mit dem Einschließungskriterium erhält man (siehe S. 56)

$$\|x\|_\infty = \max\big\{|x_i| : i \in \{1, 2, \ldots, n\}\big\}.$$

Die Bedingungen $q > 0$ und $\frac{1}{p} + \frac{1}{q} = 1$ erzwingen $p > 1$ für die Hölder-Ungleichung, aber durch Grenzübergang erhält man auch

$$\sum_{i=1}^{n} |x_i y_i| \leq \|x\|_1 \|y\|_\infty,$$

was sich übrigens leicht direkt nachprüfen lässt. Im Fall $p < 1$ ist die Eigenschaft (ii) einer Norm nicht erfüllt, wie man am Beispiel $x = (1, 0)$, $y = (0, 1)$ sieht.

Für das *Skalarprodukt*

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y}_i$$

zweier Elemente x und y von K^n (wobei die komplexe Konjugation im Fall $K = \mathbb{R}$ unnötig ist) gilt offensichtlich

$$\langle x, x\rangle = \|x\|_2^2,$$

und aus der Dreiecksungleichung in K sowie der Hölderungleichung im Fall $p = q = 2$ folgt die *Cauchy-Schwarz-Ungleichung*

$$|\langle x, y\rangle| \leq \|x\|_2 \|y\|_2.$$

Zum Beweis von Satz 8.1 benötigen wir weitere Ungleichungen. In Verallgemeinerung eines bekannten Begriffs nennen wir

$$\frac{m_1 x_1 + \ldots + m_k x_k}{m}$$

[43] $\|x\|_p$ wird „p-Norm von x“ gelesen. Die Dreiecksungleichung für diese Norm heißt auch *Minkowski-Ungleichung*.

das *gewichtete arithmetische Mittel* der Zahlen $x_1, \ldots, x_k$ mit den positiven Gewichten $m_1, \ldots, m_k$, wobei $m = m_1 + \ldots + m_k$, und für positive x_i nennen wir

$$\left(x_1^{m_1} \cdots x_k^{m_k}\right)^{1/m}$$

das *gewichtete geometrische Mittel* dieser Zahlen.

Lemma 8.1. *Das gewichtete geometrische Mittel ist nicht größer als das gewichtete arithmetische Mittel mit den selben Gewichten, wobei Gleichheit genau dann eintritt, wenn die Zahlen x_i zueinander gleich sind.*

Beweis. Sind die Gewichte m_i ganze Zahlen, so folgt die Behauptung aus Folgerung 2.6, also der Ungleichung zwischen dem gewöhnlichen geometrischen und arithmetischen Mittel der Zahlen

$$\underbrace{x_1, \ldots, x_1}_{m_1}, \underbrace{x_2, \ldots, x_2}_{m_2}, \ldots \underbrace{x_k, \ldots, x_k}_{m_k}.$$

Da sich das gewichtete arithmetische bzw. geometrische Mittel nicht ändert, wenn man alle m_i durch die selbe positive Zahl dividiert, folgt die Behauptung auch für rationale Gewichte. Sind schließlich die Gewichte m_i beliebige positive reelle Zahlen, so wählen wir für jedes i eine Folge m_{ij} von positiven rationalen Zahlen, so dass $m_{ij} \to m_i$ für $j \to \infty$. Die Ungleichung gilt bereits für jedes j, und nach Satz 3.7 bleibt sie gültig, wenn wir auf beiden Seiten zum Grenzwert $j \to \infty$ übergehen. Nach den Rechenregeln existieren diese Grenzwerte und sind gleich dem geometrischen bzw. arithmetischen Mittel mit den Gewichten m_i. □

Folgerung 8.1. *Für beliebige positive relle Zahlen $a_1, \ldots, a_k$ und $p_1, \ldots, p_k$ mit der Eigenschaft $\frac{1}{p_1} + \ldots + \frac{1}{p_k} = 1$ gilt die* Young-Ungleichung

$$a_1 \cdots a_k \le \frac{a_1^{p_1}}{p_1} + \ldots + \frac{a_k^{p_k}}{p_k}.$$

Dazu setzen wir einfach $x_i = a_i^{p_i}$ und $m_i = \frac{1}{p_i}$.

Beweis von Satz 8.1. Behauptung (i) folgt aus der analogen Eigenschaft des Absolutbetrages und dem Distributivgesetz. Zum Beweis von (iii) betrachten wir ein Element x mit der Eigenschaft $\|x\|_p = 0$. Im Fall $p = \infty$ ist die Behauptung offensichtlich. Andernfalls bilden wir die $p$te Potenz und erhalten $|x_1|^p + \ldots + |x_n|^p = 0$. Da alle Summanden nicht negativ sind, folgt $|x_i|^p = 0$, also $x_i = 0$, für alle i.

Als Nächstes beweisen wir die Hölder-Ungleichung, wobei wir zunächst annehmen, dass $\|x\|_p = \|y\|_q = 1$. Nach dem Spezialfall $k = 2$ der Young-Ungleichung gilt für jedes i

$$|x_i y_i| \leq \frac{|x_i|^p}{p} + \frac{|y_i|^q}{q},$$

und durch Summation über i erhalten wir wegen $\frac{1}{p} + \frac{1}{q}$ die Behauptung. Sind x und y beliebig, aber von 0_V verschieden, so können wir das Bewiesene auf die Vektoren

$$\frac{x}{\|x\|_p}, \qquad \frac{y}{\|y\|_q}$$

anwenden und erhalten mit dem Distributivgesetz

$$\frac{1}{\|x\|_p \|y\|_q} \sum_{i=1}^{n} |x_i y_i| \leq 1.$$

Ist schließlich $x = 0$ oder $y = 0$, so ist die Behauptung offensichtlich.

Nun beweisen wir die Eigenschaft (ii) für $1 < p < \infty$. Zu diesem gibt es ein eindeutig bestimmtes q, so dass $\frac{1}{p} + \frac{1}{q} = 1$, und mit der Hölder-Ungleichung folgt

$$\sum_{i=1}^{n} |x_i||x_i + y_i|^{p-1} \leq \|x\|_p \left(\sum_{i=1}^{n} |x_i + y_i|^{(p-1)q} \right)^{\frac{1}{q}}.$$

Wegen $\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p}$ bedeutet dies

$$\sum_{i=1}^{n} |x_i||x_i + y_i|^{p-1} \leq \|x\|_p \|x + y\|_p^{p-1}$$

und analog

$$\sum_{i=1}^{n} |y_i||x_i + y_i|^{p-1} \leq \|y\|_p \|x + y\|_p^{p-1}.$$

Addieren wir beide Ungleichungen und benutzen die Dreiecksungleichung in K, so folgt

$$\|x + y\|_p^p \leq \sum_{i=1}^{n} (|x_i| + |y_i|)|x_i + y_i|^{p-1} \leq (\|x\|_p + \|y\|_p)\|x + y\|_p^{p-1}.$$

Für $\|x + y\|_p = 0$ ist nichts zu beweisen, und andernfalls können wir durch $\|x + y\|_p^{p-1}$ dividieren. Damit ist die Minkowski-Ungleichung (ii) bewiesen.

Der Beweis in den Fällen $p = 1$ und $p = \infty$ ist einfacher und wird den Teilnehmern als Übungsaufgabe überlassen. Die Minkowski-Ungleichung folgt natürlich auch durch Grenzübergang. $\square$

Definition 8.3. *Zwei Normen $\| \, . \, \|$ und $\| \, . \, \|'$ auf einem Vektorraum V heißen* äquivalent, *wenn es Konstanten c und c' gibt, so dass für alle $x \in V$ gilt*

$$\|x\|' \leq c\|x\|, \qquad \|x\| \leq c'\|x\|'.$$

Lemma 8.2. *Für alle $p \in \,]1, \infty[$ und $x \in K^n$ gilt*

$$\|x\|_\infty \leq \|x\|_p \leq \|x\|_1 \leq n\|x\|_\infty,$$

die Normen $\| \, . \, \|_p$ für $p \in [1, \infty]$ auf K^n sind also sämtlich äquivalent.

Beweis. Ersetzen wir in x alle Koordinaten außer der betragsmäßig größten durch Null, so verkleinert sich $\|x\|_p$, und es folgt $\|x\|_\infty \leq \|x\|_p$. Schreiben wir hingegen x als Summe der Vektoren $(0, \ldots, 0, x_i, 0, \ldots, 0)$, so ergibt sich aus der Minkowskiungleichung $\|x\|_p \leq \|x\|_1$. Ersetzen wir schließlich in x alle Koordinaten durch die betragsmäßig größte, so ergibt sich $\|x\|_1 \leq n\|x\|_\infty$. □

Mit Hilfe der so genannten Jensen-Ungleichung kann man sogar zeigen, dass für beliebige $p \geq q$ gilt $\|x\|_p \leq \|x\|_q$.

8.2 Metriken

Vorlesung 5b
11.11.22

Viele Objekte der realen Welt lassen sich nicht durch Vektorräume beschreiben. Trotzdem kann man Abstände zwischen Punkten betrachten.

Definition 8.4. *Es sei X eine Menge. Eine Funktion $d : X \times X \to \mathbb{R}$ heißt* Metrik, *wenn für alle x, y und $z \in X$ gilt:*

(i) *$d(x, y) = d(y, x)$,*

(ii) *$d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung),*

(iii) *genau dann $d(x, y) = 0$, wenn $x = y$.*

Eine Menge, die mit einer Metrik versehen ist, nennt man metrischen Raum[44], *ihre Elemente nennt man* Punkte.

Durch Anwendung der Eigenschaften (iii), (ii) und (i) sehen wir, dass

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y),$$

d. h. dass eine Metrik nur nichtnegative Werte annimmt. Manchmal ist es sinnvoll, Metriken mit Werten in $[0, \infty]$ zu betrachten, wobei die Addition durch die Festlegung $a + \infty = \infty + a = \infty$ ausgedehnt wird.

[44]Strenggenommen ist ein metrischer Raum ein geordnetes Paar (X, d).

Beispiel. Ist V ein Vektorraum mit einer Norm, so ist

$$d(x, y) = \|x - y\|$$

eine Metrik auf V. Die Norm $\| \,.\, \|_2$ auf $\mathbb{R}^n$ liefert die euklidische Metrik. $\triangleleft$

Das richtige mathematische Modell für den Weltraum der Newtonschen Mechanik ist nicht ein Vektorraum, sondern ein affiner Raum.

Definition 8.5. *Ein* affiner Raum *ist eine Menge X zusammen mit einem Vektorraum V und einer Abbildung $X \times V \to X$, geschrieben $(x, v) \mapsto x + v$, mit folgenden Eigenschaften:*

(a) Für alle $x \in X$ und u, $v \in V$ gilt $x + (u + v) = (x + u) + v$.

(b) Für beliebige x, $y \in X$ gibt es genau ein $v \in V$, so dass $x + v = y$. (Man bezeichnet v mit $\overrightarrow{xy}$.) Die durch ein Element $v \in V$ gegebene Abbildung $X \to X$, $x \mapsto x + v$, nennt man Translation.

Man bezeichnet die Elemente von X als Punkte und die von V als Vektoren. Wählt man einen Punkt o von X als so genannten Ursprung, dann ist die durch $v \mapsto o + v$ gegebene Abbildung $V \mapsto X$ bijektiv. Die Umkehrabbildung ordnet jedem Punkt x seinen Ortsvektor $\overrightarrow{ox}$ zu. Viele Autoren finden affine Räume darum entbehrlich.

Definition 8.6. *Eine Metrik d auf einem affinen Raum X heißt* translationsinvariant, *wenn für alle x, $y \in X$ und v im zugehörigen Vektorraum V gilt*

$$d(x + v, y + v) = d(x, y).$$

Es ist klar, dass für jede Norm $\| \,.\, \|$ auf V durch

$$d(x, y) = \|\overrightarrow{xy}\|$$

eine translationsinvariante Metrik d auf X definiert wird.

Beispiel. Ist d eine Metrik auf der Menge X und Y eine Teilmenge von X, so ist die Einschränkung $d|_{Y \times Y}$ eine Metrik auf Y, genannt die von d *induzierte Metrik*. Die Menge Y, versehen mit dieser Metrik, nennt man einen metrischen Teilraum von X. Betrachtet man z. B. die Erdoberfläche als Teilmenge des euklidischen Raumes, so beschreibt $d(x, y)$ die Länge der Sehne durch das Erdinnere, die x und y verbindet. $\triangleleft$

Beispiel. Es sei X die Menge aller Stellungen, die ein Körper im Raum einnehmen kann, ohne seinen Schwerpunkt zu verlagern. Es ist bekannt, dass man von einer Stellung x zu einer anderen Stellung y immer durch eine Drehung um eine geeignete Achse gelangen kann. Bezeichnet $d(x, y)$ den minimalen erforderlichen Drehwinkel, so ist d eine Metrik. $\triangleleft$

Beispiel. Es sei A eine Menge und $d : A^n \times A^n \to \mathbb{R}$ gegeben durch

$$d(x, y) = \left| \left\{ i \in \{1, 2, \ldots, n\} : x_i \neq y_i \right\} \right|.$$

Dann ist d eine Metrik auf A^n. Man kann sich A als ein Alphabet und A^n als Menge der Zeichenketten der Länge n vorstellen. Dann gibt $d(x, y)$ an, an wie vielen Stellen sich die Zeichenketten x und y unterscheiden. In A^n kann man die Teilmenge Y der sinnvollen Wörter betrachten. $\triangleleft$

Metrische Räume bilden einen passenden Rahmen für die Analysis.

Definition 8.7. *Es sei X ein metrischer Raum.*

(i) Eine Folge von Punkten x_k von X konvergiert *gegen einen Punkt a von X, wenn es für jede positive reelle Zahl ε eine natürliche Zahl k_0 gibt, so dass für alle natürlichen Zahlen k mit der Eigenschaft $k \geq k_0$ gilt*

$$d(x_k, a) < \varepsilon.$$

(ii) Ein Punkt a von X heißt Häufungspunkt *der Folge von Punkten x_k von X, wenn es für jede positive reelle Zahl ε und jede natürliche Zahl k_0 eine natürliche Zahl k gibt, so dass $k \geq k_0$ und*

$$d(x_k, a) < \varepsilon.$$

(iii) Eine Teilmenge A von X heißt abgeschlossen *in X, wenn jeder Häufungspunkt einer Folge von Elementen von A ebenfalls in A liegt.*

(iv) Eine Teilmenge U von X heißt offen *in X, wenn es für jedes Element a von U eine positive reelle Zahl ε gibt, so dass alle Elemente x von X mit der Eigenschaft $d(x, a) < \varepsilon$ in U liegen.*

In dem Spezialfall, dass X ein Körper mit Absolutbetrag ist, stimmen diese Begriffe mit den früher betrachteten überein. Die Sätze 3.1, 3.2, 3.13 und 3.15 übertragen sich einschließlich ihrer Beweise. Wir fassen die Versionen für metrische Räume hier zusammen:

Satz 8.2. *(i) Konvergiert eine Folge sowohl gegen a als auch gegen b, so ist $a = b$.*

(ii) Jede konvergente Folge ist beschränkt, d. h. es gibt eine Zahl $c > 0$, so dass für alle k und l gilt $d(x_k, x_l) \leq c$.

(iii) Ein Punkt a ist genau dann Häufungspunkt einer Folge, wenn er Grenzwert einer ihrer Teilfolgen ist.

(iv) Eine Menge A ist genau dann abgeschlossen in X, wenn ihr Komplement $X \setminus A$ offen in X ist.

Wegen Aussage (i) können wir sagen, dass a der *Grenzwert* der Folge x_k ist, wenn diese gegen a konvergiert.

Mit dem folgenden Begriff kann man die Definitionen von Grenzwert und Häufungspunkt anschaulicher formulieren.

Definition 8.8. *Eine offene Menge in X, die einen Punkt a enthält, nennt man eine* Umgebung *von a in X.*

Satz 8.3. *(i) Für jedes $\varepsilon > 0$ ist die Menge*

$$U_\varepsilon(a) = \{x \in X \mid d(x, a) < \varepsilon\}$$

eine Umgebung von a, genannt ε-Umgebung von a.

(ii) Eine Folge von Punkten x_k konvergiert genau dann gegen einen Punkt a, wenn es für jede Umgebung U von a ein k_0 gibt, so dass für alle k mit der Eigenschaft $k \geq k_0$ gilt $x_k \in U$.

(iii) Ein Punkt a ist genau dann Häufungspunkt einer Folge von Punkten x_k, wenn es für jede Umgebung U von a und jedes k_0 ein k gibt, so dass $k \geq k_0$ und $x_k \in U$.

Beweis. **(i)** Es sei $b \in U_\varepsilon(a)$, also $d(b, a) < \varepsilon$. Setzen wir $\delta = \varepsilon - d(b, a)$, so ist $\delta > 0$, und es gilt für jedes $x \in X$ mit $d(x, b) < \delta$, dass $d(x, a) \leq d(x, b) + d(b, a) < \delta + d(b, a) = \varepsilon$, also $x \in U_\varepsilon(a)$.
(ii) Angenommen, die Folge x_k konvergiert gegen a. Ist U eine Umgebung von a, so gibt es nach Definition der Offenheit ein $\varepsilon > 0$, so dass alle $x \in X$ mit $d(x, a) < \varepsilon$ in U liegen. Wegen der Konvergenz gibt es dann ein k_0, so dass für alle $k \geq k_0$ gilt $d(x_k, a) < \varepsilon$. Für diese k ist somit $x_k \in U$. Die Umkehrung ist offensichtlich, denn wenn eine Eigenschaft für alle Umgebungen gilt, so gilt sie insbesondere für alle ε-Umgebungen.

Der Beweis von (iii) ist ähnlich. □

Bemerkungen.

(i) Metriken auf einem Vektorraum, die von äquivalenten Normen induziert werden, definieren den selben Begriff der Konvergenz sowie der Offenheit und Abgeschlossenheit.

(ii) In Analogie zur Bemerkung am Ende von Abschnitt 3.2 gilt: Eine Folge von Elementen

$$x_k = (x_{1k}, \ldots, x_{nk})$$

von K^n konvergiert (bezüglich einer der in Satz 8.1 betrachteten Normen) genau dann gegen einen Punkt $a = (a_1, \ldots, a_n)$, wenn jede der

Koordinatenfolgen x_{ik} gegen die jeweilige Koordinate a_i von a konvergiert. Ist nämlich $\|x_k - a\|_\infty$ eine Nullfolge, so sind offensichtlich die $|x_{ik} - a_i|$ Nullfolgen, und sind umgekehrt die $|x_{ik} - a_i|$ Nullfolgen, so ist offensichtlich $\|x_k - a\|_1$ eine Nullfolge.

Satz 8.4. *Es sei X ein metrischer Raum.*

(i) Die Mengen X und $\varnothing$ sind offen in X.

(ii) Sind U und V offen in X, so ist auch der Durchschnitt $U \cap V$ offen in X.

(iii) Ist für jeden Index i aus einer Menge I eine offene Teilmenge U_i von X gegeben, so ist die Vereinigung

$$U = \bigcup_{i \in I} U_i$$

offen in X.

Beweis. **(i)** Der gesamte Raum X ist offen, da für jeden Punkt x gilt $U_1(x) \subseteq X$, und $\varnothing$ ist offen, da es keinen Punkt gibt, für den eine Bedingung zu erfüllen wäre.
(ii) Ist $x \in U \cap V$, so gibt es wegen der Offenheit von U und V Zahlen $\varepsilon > 0$ und $\delta > 0$, so dass $U_\varepsilon(x) \subseteq U$ und $U_\delta(x) \subseteq V$. Setzen wir $\eta = \min\{\varepsilon, \delta\}$, so gilt $U_\eta(x) \subseteq U \cap V$.
(iii) Ist x Element der Vereinigung, so gibt es einen Index $i \in I$, so dass $x \in U_i$. Da U_i offen ist, gibt es ein $\varepsilon > 0$, so dass $U_\varepsilon(x) \subseteq U_i$, also $U_\varepsilon(x) \subseteq U$. □

Bemerkung. Für das Komplement[45] $B^{\mathrm{c}} = X \setminus B$ gelten die de Morganschen Regeln

$$\left(\bigcup_{i \in I} B_i\right)^{\mathrm{c}} = \bigcap_{i \in I} B_i^{\mathrm{c}}, \qquad \left(\bigcap_{i \in I} B_i\right)^{\mathrm{c}} = \bigcup_{i \in I} B_i^{\mathrm{c}}.$$

Mit Hilfe von Satz 8.2(iv) erhält man also eine Version von Satz 8.4 für abgeschlossene Mengen.

Ist Y ein metrischer Teilraum von X, so braucht eine offene Teilmenge U von Y nicht offen in X und eine abgeschlossene Teilmenge A von Y nicht abgeschlossen in X zu sein. Man betrachte z. B. $X = \mathbb{R}$, $Y = U = A = [0, 1[$.

Satz 8.5. *Es sei Y ein metrischer Teilraum von X.*

[45] Bei dieser Schreibweise muss X aus dem Zusammenhang erkennbar sein.

(i) Die in Y offenen Mengen sind genau die Mengen der Form $U \cap Y$, wobei U eine offene Menge in X ist.

(ii) Die in Y abgeschlossenen Mengen sind genau die Mengen der Form $A \cap Y$, wobei A eine abgeschlossene Menge in X ist.

Beweis. Wegen $Y \setminus (A \cap Y) = Y \cap (X \setminus A)$ genügt es angesichts von Satz 8.2(iv), die Aussage (i) zu beweisen. Zuvor bemerken wir: Sind $U_\varepsilon^X(b)$ und $U_\varepsilon^Y(b)$ die ε-Umgebung eines Punktes $b \in Y$ in den Räumen X bzw. Y, so ist $U_\varepsilon^X(b) \cap Y = U_\varepsilon^Y(b)$.

Ist U offen in X und $b \in U \cap Y$, so gibt es ein $\varepsilon > 0$, so dass $U_\varepsilon^X(b) \in U$, und dann ist $U_\varepsilon^Y(b) \in U \cap Y$. Also ist $U \cap Y$ offen in Y.

Nun sei V eine offene Teilmenge von Y. Für jedes $b \in V$ gibt es ein $\varepsilon_b > 0$, so dass $U_{\varepsilon_b}^Y(b) \subseteq V$. Nach Satz 8.3(i) und 8.4(iii) ist die Menge

$$U = \bigcup_{b \in V} U_{\varepsilon_b}^X(b)$$

offen in X. Es gilt $U \cap Y \subseteq V$, und für jedes $b \in V$ ist $b \in U_{\varepsilon_b}(b) \cap Y \subseteq U \cap Y$, also $V \subseteq U \cap Y$. □

Um für jeden Punkt $b \in Y$ ein ε_b zu wählen, scheint man das so genannte Auswahlaxiom zu benötigen. Im vorliegenden Fall ist das aber nicht nötig, denn

$$\varepsilon_b = \min\{1, \sup\{\varepsilon > 0 \mid U_\varepsilon^Y(b) \subseteq V\}\} > 0.$$

Eine Familie von Teilmengen einer beliebigen Menge X, die die Eigenschaften der Familie der offenen Mengen aus Satz 8.4 hat, nennt man übrigens eine *Topologie* auf X. Eine Menge X, die mit einer Topologie versehen ist, nennt man *topologischen Raum*. Für solche Räume betrachtet man die Aussagen (ii) und (iii) von Satz 8.3 und den Satz 8.5 als Definitionen.

8.3 Vollständigkeit

Vorlesung 6a
16.11.22

Wie schon im Fall von Körpern spielt der Begriff der Vollständigkeit auch für metrische Räume eine wichtige Rolle.

Definition 8.9. *Eine Folge von Elementen x_k eines metrischen Raumes X heißt* Cauchy-Folge, *wenn es für jede positive Zahl ε eine natürliche Zahl k_0 gibt, so dass für alle natürlichen Zahlen k und l mit der Eigenschaft $k \geq k_0$ und $l \geq k_0$ gilt*

$$d(x_k, x_l) < \varepsilon.$$

Ein metrischer Raum heißt vollständig, *wenn jede Cauchy-Folge in diesem Raum konvergent ist.*

Wegen $d(x_k, x_l) \leq d(x_k, a) + d(a, x_l)$ ist jede konvergente Folge eine Cauchy-Folge.

Ähnlich wie beim Grenzwertbegriff sieht man: Eine Folge von Elementen

$$x_k = (x_{1k}, \ldots, x_{nk})$$

von K^n ist genau dann eine Cauchy-Folge (bezüglich einer der in Satz 8.1 betrachteten Normen), wenn jede der Koordinatenfolgen x_{ik} eine Cauchy-Folge ist. Da die Körper $\mathbb{R}$ und $\mathbb{C}$ nach Satz 3.16 vollständig sind, erhalten wir

Folgerung 8.2. *Die Vektorräume $\mathbb{R}^n$ und $\mathbb{C}^n$ sind bezüglich jeder der Normen $\| \,.\, \|_p$ vollständig.*

Lemma 8.3. *Ein abgeschlossener Teilraum Y eines vollständigen metrischen Raumes X ist vollständig. Ein vollständiger Teilraum eines metrischen Raumes X ist abgeschlossen in X.*

Beweis. Angenommen, Y ist abgeschlossen in X. Ist y_k eine Cauchy-Folge in Y, so ist sie auch eine Cauchy-Folge in X, hat dort also einen Grenzwert a. Da Y abgeschlossen ist, gilt $a \in Y$, also ist die Folge in Y konvergent.

Angenommen, Y ist vollständig. Ist y_k eine Folge in Y, die gegen einen Punkt a von X konvergiert, so ist sie eine Cauchy-Folge in X, also auch eine Cauchy-Folge in Y. Wegen der Vollständigkeit von Y hat sie einen Grenzwert b in Y, und nach Satz 8.2(i) ist $a = b$. □

Beispiel. Wir definieren für jede integrierbare Funktion f auf dem beschränkten Intervall $[a, b]$

$$\|f\|_1 = \int_a^b |f(x)| \, dx.$$

Dies ist keine Norm auf dem Vektorraum solcher Funktionen, weil z. B. für die Funktion

$$f(x) = \begin{cases} 1, & \text{wenn } x = a, \\ 0 & \text{sonst} \end{cases}$$

gilt $\|f\|_1 = 0$. (Dem kann man abhelfen, in dem man den Raum der Äquivalenzklassen von Funktionen betrachtet, wobei $f \sim g$, wenn $\|f - g\|_1 = 0$.) Mit Hilfe der Sätze 7.5 und 7.6 kann man leicht zeigen, dass die Einschränkung von $\| \,.\, \|_1$ auf den Unterraum der stetigen Funktionen eine Norm ist.

Betrachten wir nun auf $[-1, 1]$ die Funktionen

$$f_k(x) = \frac{kx}{|kx| + 1}.$$

Weil die Funktionen ungerade sind, ist

$$\|f_k - \operatorname{sgn}\|_1 = 2\int_0^1 (\operatorname{sgn} x - f_k(x))\,dx,$$

und durch die Substitution $u = kx + 1$ erhalten wir

$$\|f_k - \operatorname{sgn}\|_1 = \frac{2}{k}\int_1^{k+1}\left(1 - \frac{u-1}{u}\right)du = 2\frac{\ln(k+1)}{k}.$$

Die Folge konvergiert also im Raum der Äquivalenzklassen gegen die Signumfunktion. Der Unterraum der stetigen Funktionen ist somit nicht abgeschlossen und folglich auch nicht vollständig bezüglich der Norm $\|\,.\,\|_1$. Der Raum der Äquivalenzklassen aller integrierbaren Funktionen im Riemannschen Sinne ist übrigens auch nicht vollständig, aber den Beweis lernen Sie erst in der Veranstaltung zur Maßtheorie kennen. $\triangleleft$

Definition 8.10. *Es sei X eine Menge. Für jede beschränkte reell- oder komplexwertige Funktion $f : X \to K$ nennen wir*

$$\|f\| = \sup|f|$$

die Supremumsnorm *von f. Eine Folge, die bezüglich der Supremumsnorm konvergiert, nennen wir* gleichmäßig konvergent*.*

Eine Folge von Funktionen f_k konvergiert also genau dann gleichmäßig gegen eine Funktion f, wenn es für jede positive Zahl ε eine natürliche Zahl k_0 gibt, so dass für alle natürlichen Zahlen k mit der Eigenschaft $k \geq k_0$ und alle Elemente x von X gilt

$$|f_k(x) - f(x)| < \varepsilon.$$

Diese Umformulierung ist sogar auf unbeschränkte Funktionen anwendbar, wobei die Differenzen $f_k - f$ natürlich beschränkt sein müssen.

Lemma 8.4. *Der Raum der beschränkten Funktionen auf einer beliebigen Menge ist bezüglich der Supremumsnorm vollständig.*

Beweis. Es sei f_n eine Cauchy-Folge. Für jedes Element x von X gilt

$$|f_k(x) - f_l(x)| \leq \|f_k - f_l\|,$$

also ist $f_k(x)$ eine Cauchy-Folge. Wegen der Vollständigkeit von K hat diese einen Grenzwert, den wir mit $f(x)$ bezeichnen. So erhalten wir eine Funktion $f : X \to K$.

Nun sei $\varepsilon > 0$. Dann gibt es ein k_0, so dass für $k \geq k_0$ und $l \geq k_0$ gilt $\|f_k - f_l\| < \varepsilon$, also für jedes $x \in X$

$$|f_k(x) - f_l(x)| < \varepsilon.$$

Durch Grenzübergang $l \to \infty$ folgt

$$|f_k(x) - f(x)| \leq \varepsilon,$$

und mit der Dreiecksungleichung

$$|f(x)| \leq |f(x) - f_k(x)| + |f_k(x)| \leq \varepsilon + \|f_k\|.$$

Da x beliebig war, ist f beschränkt und

$$\|f_k(x) - f(x)\| \leq \varepsilon.$$

Da ε beliebig war, konvergiert die Folge f_k gleichmäßig gegen f. $\square$

Man kann die Supremumsnorm auch für Funktionen mit Werten in einem Vektorraum V definieren, wenn auf diesem eine Norm definiert ist. Ist V vollständig, so gilt das Lemma auch hier.

Noch allgemeiner kann man Abbildungen mit Werten in einem metrischen Raum Y mit einer Metrik e betrachten. Eine solche Abbildung heißt beschränkt, wenn ihr Bild (Wertevorrat) beschränkt ist, d. h. wenn es eine Zahl c gibt, so dass für alle u, $v \in X$ gilt $e(f(u), f(v)) \leq c$. Auf dem Raum aller beschränkten Abbildungen $X \to Y$ definiert man die Supremumsmetrik

$$d(f, g) = \sup\{e(f(x), g(x)) \mid x \in X\}.$$

Die allgemeine Version des Lemmas besagt, dass d vollständig ist, wenn e vollständig ist. Lässt man für Metriken den Wert ∞ zu, dann kann man sogar die Bedingung der Beschränktheit fallen lassen.

Satz 8.6. *Ist X ein metrischer Raum, so gibt es einen vollständigen metrischen Raum Y und eine injektive Abbildung $j : X \to Y$ mit folgender Eigenschaft. Ist d die Metrik auf X und e die auf Y, so ist $e(j(x), j(y)) = d(x, y)$, und jeder Punkt von Y ist ein Häufungspunkt des Wertebereichs von j.*

Meist identifiziert man die Elemente von X mit ihren Bildern unter j und nennt Y die *Vervollständigung* von X.

In der Vorlesung Analysis I haben wir den Körper $\mathbb{R}$ der reellen Zahlen als Vervollständigung der geordneten Menge $\mathbb{Q}$ definiert. Man kann ihn genausogut als Vervollständigung des metrischen Raums $\mathbb{Q}$ definieren. In beiden Fällen muss man zusätzlich die Rechenoperationen auf $\mathbb{R}$ definieren und die Körperaxiome nachprüfen.

Beweis. Wir nennen Cauchy-Folgen $x' = (x'_k)_{k\in\mathbb{N}}$ und $x'' = (x''_k)_{k\in\mathbb{N}}$ in X äquivalent, wenn es für jedes $\varepsilon > 0$ ein k_0 gibt, so dass für $k,\, l \geq k_0$ gilt

$$d(x'_k, x''_k) < \varepsilon.$$

Dies ist offensichtlich eine Äquivalenzrelation, und die Menge der Äquivalenzklassen bezeichnen wir mit Y. Die Abbildung j, die jedem Element x die konstante Folge $(x)_{k\in\mathbb{N}}$ zuordnet, ist offensichtlich injektiv.

Für beliebige Folgen x' und x'' und beliebige Indizes k und l gilt nach der Dreiecksungleichung

$$d(x'_k, x''_n) \leq d(x'_k, x'_l) + d(x'_l, x''_m) + d(x''_m, x''_n)$$

und das Selbe mit vertauschten Rollen von k und l, also

$$|d(x'_k, x''_n) - d(x'_l, x''_m)| \leq d(x'_k, x'_l) + d(x''_m, x''_n).$$

Sind x' und x'' Cauchy-Folgen in X, so gibt es für jedes $\varepsilon > 0$ natürliche Zahlen k_0 und n_0, so dass die rechte Seite für $k,\, l \geq k_0$ und $m,\, n \geq n_0$ kleiner als ε ist. In diesem Sinne ist $d(x'_k, x''_n)$ eine Cauchysche Doppelfolge in $\mathbb{R}$, und somit existiert

$$\tilde{e}(x', x'') = \lim_{k\to\infty} \lim_{n\to\infty} d(x'_k, x''_n).$$

Die Eigenschaften (i) und (ii) einer Metrik folgen für $\tilde{e}$ durch Grenzübergang aus denen für d, und aus $\tilde{e}(x', x'') = 0$ folgt die Äquivalenz von x' und x''. Außerdem hängt $\tilde{e}(x', x'')$ nur von den Äquivalenzklassen der Folgen x' und x'' ab und definiert somit eine Metrik e auf Y. Ordnen wir jedem Glied x_k einer Folge x in X die konstante Folge $(x_k)_{l\in\mathbb{N}}$ zu, so erhalten wir eine Folge in Y, die gegen die Äquivalenzklasse von x konvergiert.

Zum Beweis der Vollständigkeit müssen wir eine Cauchy-Folge $x = (x_k)_{k\in\mathbb{N}}$ in Y betrachten. Ihre Glieder $x_k = (x_{k,n})_{n\in\mathbb{N}}$ sind ihrerseits (Äquivalenzklassen von) Cauchy-Folgen, und darum gibt es für jedes k ein n_k, so dass für $m \geq n_k$ gilt

$$d(x_{k,n_k}, x_{k,m}) < \frac{1}{k}.$$

Wir behaupten, dass auch $(x_{k,n_k})_{k\in\mathbb{N}}$ eine Cauchy-Folge in X ist und somit ein Element y von Y definiert. Ist $\varepsilon > 0$ gegeben, so gibt es ein k_0, so dass für $k,\, l \geq k_0$ gilt

$$\tilde{e}(x_k, x_l) < \varepsilon.$$

Für gegebene k und l gibt es also ein $m_{k,l}$, so dass für $m,\, n \geq m_{k,l}$ gilt

$$d(x_{k,m}, x_{l,n}) < 2\varepsilon.$$

Wir können n und l so wählen, dass außerdem $m \geq n_k$ und $n \geq n_l$ gilt, und dann folgt

$$d(x_{k,n_k}, x_{l,n_l}) \leq d(x_{k,n_k}, x_{k,m}) + d(x_{k,m}, x_{l,n}) + d(x_{l,n}, x_{l,n_l}) \leq \frac{1}{k} + 2\varepsilon + \frac{1}{l}.$$

Ist außerdem $k,\, l \geq \frac{1}{\varepsilon}$, so ist die rechte Seite höchstens 4ε, und die Behauptung folgt. Die Nachprüfung, dass die Folge x gegen y konvergiert, bleibt den Teilnehmern als Übungsaufgabe überlassen. □

8.4 Kontraktionen

Der uns bekannte Begriff der Lipschitz-Stetigkeit lässt sich verallgemeinern.

Definition 8.11. *Es seien X und Y metrische Räume mit Metriken d bzw. e, und es sei f eine Abbildung von X nach Y.*

(i) *Die Abbildung f heißt* Lipschitz-stetig, *wenn es eine Zahl $c \geq 0$ gibt, so dass für alle Punkte u und v von X gilt*

$$e(f(u), f(v)) \leq cd(u, v).$$

Die Zahl c nennt man eine Lipschitzkonstnte für f.

(ii) *Die Abbildung f heißt* Kontraktion, *wenn sie Lipschitz-stetig mit einer Lipschitzkonstanten $c < 1$ ist.*

Für jeden Punkt $a \in X$ ist die Funktion

$$f(x) = d(x, a)$$

Lipschitz-stetig mit der Lipschitzkonstanten 1, denn aus der Dreiecksungleichung folgt

$$|d(u, a) - d(v, a)| \leq d(u, v).$$

Der folgende Satz zeigt die Nützlichkeit des Begriffs der Vollständigkeit.

Satz 8.7 (Banachscher Fixpunktsatz)**.** *Ist $f : X \to X$ eine Kontraktion von einem nichtleeren vollständigen metrischen Raum X in sich selbst, so hat f einen Fixpunkt, d. h. einen Punkt $a \in X$, so dass $f(a) = a$.*

Ist b ebenfalls ein Fixpunkt, so gilt $d(a, b) = d(f(a), f(b)) \leq c\, d(a, b)$, also $d(a, b) = 0$ und somit $a = b$. Es kann also nur einen Fixpunkt geben.

Beweis. Wegen $X \neq \varnothing$ gibt es einen Punkt $x_0 \in X$. Wir definieren eine Folge von Punkten x_k rekursiv durch

$$x_{k+1} = f(x_k).$$

Wir beweisen durch vollständige Induktion nach k, dass

$$d(x_{k+1}, x_k) \leq c^k d(x_1, x_0).$$

Dies gilt offenbar für $k = 0$, und gilt es für eine Zahl k, so folgt

$$d(x_{k+2}, x_{k+1}) = d(f(x_{k+1}), f(x_k)) \leq c\, d(x_{k+1}, x_k) \leq c^{k+1} d(x_1, x_0).$$

Für $l \geq k$ folgt mit der Dreiecksungleichung

$$d(x_l, x_k) \leq \sum_{i=k}^{l-1} d(x_{i+1}, x_i) \leq \sum_{i=k}^{l-1} c^i d(x_1, x_0) \leq \frac{c^k}{1-c} d(x_1, x_0).$$

Wegen $|c| < 1$ konvergiert die rechte Seite für $k \to \infty$ gegen Null, und somit ist x_k eine Cauchy-Folge. Da X vollständig ist, konvergiert sie gegen einen Punkt a. Wegen der Lipschitz-Stetigkeit von f folgt aus der Rekursionsformel durch Grenzübergang, dass $a = f(a)$. □

8.5 Stetigkeit

Vorlesung 6b
18.11.22

Der Begriff der Stetigkeit von Abbildungen ist eine offensichtliche Verallgemeinerung des Begriffs der Stetigkeit von Funktionen.

Definition 8.12. *Es seien X und Y Mengen mit Metriken d bzw. e, es sei $a \in X$ und f eine Abbildung von X in Y.*

(i) *Die Abbildung f heißt* stetig *an der Stelle a, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in X$ mit der Eigenschaft $d(x, a) < \delta$ gilt $e(f(x), f(a)) < \varepsilon$.*

(ii) *Die Abbildung f heißt stetig, wenn sie an jeder Stelle stetig ist.*

Bemerkungen.

(i) Wird die Metrik d oder e auf einem Vektorraum über K von einer Norm induziert, so ändert sich der Begriff der Stetigkeit nicht, wenn man zu einer äquivalenten Norm übergeht. Dabei ist wie immer $K \in \{\mathbb{R}, \mathbb{C}\}$.

(ii) Jede Lipschitz-stetige Funktion ist stetig. Für jeden Punkt a ist die Funktion
$$f(x) = d(x, a)$$
Lipschitz-stetig, denn aus der Dreiecksungleichung folgt
$$|d(u, a) - d(v, a)| \leq d(u, v).$$

(iii) Ist f eine Abbildung von einer Menge X in den Raum K^n, so ist für jedes $x \in X$ das Element $f(x)$ ein n-Tupel, das wir mit
$$(f_1(x), \ldots, f_n(x))$$
bezeichnen. Auf diese Weise erhalten wir aus einer Abbildung $f : X \to K^n$ Koordinatenfunktionen $f_i : X \to K$ und umgekehrt.

(iv) Eine Abbildung f von einem metrischen Raum X in den Raum K^n ist genau dann stetig, wenn ihre Koordinatenfunktionen stetig sind, denn es gilt für jedes $i \in \{1, \ldots, n\}$

$$|f_i(x) - f_i(a)| \leq \|f(x) - f(a)\|_\infty \leq \|f(x) - f(a)\|_1 \leq \sum_{j=1}^{n} |f_j(x) - f_j(a)|.$$

Der Begriff der Stetigkeit hängt eng mit dem des Grenzwertes zusammen.

Definition 8.13. *Es seien X und Y Mengen mit Metriken d bzw. e.*

(i) *Ein Punkt a heißt* Häufungspunkt *einer Teilmenge D von X, wenn in jeder Umgebung von a ein Punkt von $D \setminus \{a\}$ liegt.*

(ii) *Es sei a ein Häufungspunkt der Menge D und $f : D \to Y$. Wir sagen, dass für x gegen a die Funktion f gegen ein Element b von Y konvergiert, abgekürzt $f(x) \to b$ $(x \to a)$, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in D \setminus \{a\}$ mit der Eigenschaft $d(x, a) < \delta$ gilt $e(f(x), b) < \varepsilon$.*

Bemerkungen.

(i) Weil a Häufungspunkt von D ist, kann in Analogie zu Satz 8.2(i) eine Abbildung f für $x \to a$ gegen höchstens ein Element b konvergieren. Wir bezeichnen b mit $\lim\limits_{x \to a} f(x)$ und nennen es den *Grenzwert* von f für x gegen a.

(ii) Ein Punkt von D, der kein Häufungspunkt von D ist, hießt isolierter Punkt von D. Er ist trotzdem Häufungspunkt einer Folge in D, z. B. $a, a, a, \ldots$

(iii) Die Begriffe Grenzwert und Stetigkeit hängen wie folgt zusammen. Ist a ein Häufungspunkt von D, so für gilt eine Abbildung $f : D \to Y$ genau dann

$$f(x) \to b \ (x \to a),$$

wenn die durch

$$g(x) = \begin{cases} f(x), & \text{wenn } x \in D \setminus \{a\}, \\ b, & \text{wenn } x = a \end{cases}$$

definiere Abbildung $g : D \cup \{a\} \to Y$ an der Stelle a stetig ist. Dabei wird $D \cup \{a\}$ als metrischer Teilraum von X betrachtet.

Beispiel. Es sei $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ definiert durch

$$f(x_1, x_2) = \frac{x_1 x_2}{x_1^2 + x_2^2}.$$

Für festes x_2 gilt

$$\lim_{x_1 \to 0} f(x_1, x_2) = 0,$$

und für festes x_1 gilt

$$\lim_{x_2 \to 0} f(x_1, x_2) = 0.$$

Würde der Grenzwert

$$\lim_{x \to (0,0)} f(x) = b$$

existieren, dann müsste es für jedes $\varepsilon > 0$ ein $\delta > 0$ geben, so dass für $\|(t,0)\| < \delta$ und $\|(t,t)\| < \delta$ gilt $|f(t,0) - b| < \varepsilon$ und $|f(t,t) - b| < \varepsilon$. Wir haben aber

$$f(t,0) = 0, \qquad f(t,t) = \frac{1}{2},$$

und für $\varepsilon = \frac{1}{4}$ ergibt sich ein Widerspruch. Der besagte Grenzwert existiert also nicht. $\lhd$

Sind V und W Vektorräume über K mit Normen, die wir jetzt beide mit $\|\,.\,\|$ bezeichnen, so können wir auf dem Produktraum $V \times W$ für $p \in [1, \infty]$ die Normen

$$\|(x_1, x_2)\|_p = \sqrt[p]{\|x_1\|^p + \|x_2\|^p} \qquad (x_1 \in V,\ x_2 \in W)$$

betrachten, die alle äquivalent sind.

Satz 8.8. *Für einen K-Vektorraum V mit Norm sind die Addition und die Skalarmultiplikation stetige Abbildungen*

$$V \times V \to V, \qquad K \times V \to V.$$

Beweis. Die Addition ist sogar Lipschitz-stetig, denn nach der Dreiecksungleichung gilt für (u_1, u_2), $(v_1, v_2) \in V \times V$

$$\begin{aligned}\|(u_1 + u_2) - (v_1 + v_2)\| &= \|(u_1 - v_1) + (u_2 - v_2)\| \\ &\leq \|u_1 - v_1\| + \|u_2 - v_2\| = \|(u_1, u_2) - (v_1, v_2)\|_1.\end{aligned}$$

Bei der Skalarmultiplikation gilt für (a, x), $(a_0, x_0) \in K \times V$

$$\|a \cdot x - a_0 \cdot x_0\| = \|(a - a_0)x + a_0(x - x_0)\| \leq |a - a_0|\|x\| + |a_0|\|x - x_0\|.$$

Zu gegebenem $(a_0, x_0) \in K \times V$ und $\varepsilon > 0$ wählen wir $\delta = \min\left\{1, \frac{\varepsilon}{1+\|x_0\|+|a_0|}\right\}$. Dann gilt für $\|(a,x) - (a_0, x_0)\|_\infty < \delta$

$$\|x\| \leq \|x - x_0\| + \|x_0\| \leq 1 + \|x_0\|$$

und somit

$$\|a \cdot x - a_0 \cdot x_0\| \leq (1 + \|x_0\| + |a_0|)\delta \leq \varepsilon.$$

Also ist die Skalarmultiplikation an der Stelle (a_0, x_0) stetig. □

In Analogie zu Satz 5.2 gilt das Folgenkriterium:

Satz 8.9. *Eine Abbildung $f : X \to Y$ ist genau dann stetig an der Stelle a, wenn für jede Folge x_k, die in X gegen a konvergiert, die Folge $f(x_k)$ in Y gegen $f(a)$ konvergiert.*

Beweis. Angenommen, f ist nicht stetig an der Stelle a. Dann gibt es ein $\varepsilon > 0$, so dass für jedes $\delta > 0$ ein $x \in X$ existiert, so dass $d(x,a) < \delta$, aber $e(f(x), f(a)) \geq \varepsilon$.

Insbesondere gibt es also für jedes $k \in \mathbb{N}$ ein x_k, so dass $d(x_k, a) < \frac{1}{k}$, aber $e(f(x_k), f(a)) \geq \varepsilon$. Dies bedeutet, dass x_k gegen a, aber $f(x_k)$ nicht gegen $f(a)$ konvergiert. Das Folgenkriterium ist also nicht erfüllt.

Die Umkehrung ist einfach. □

Man kann die Stetigkeit einer Abbildung $f : X \to Y$ zwischen metrischen Räumen auch mit Hilfe von Umgebungen in Analogie zu Satz 8.3 ausdrücken. Dazu definieren wir das Bild einer Teilmenge A von X und das Urbild einer Teilmenge B von Y durch

$$f(A) = \{f(x) \mid x \in a\}, \qquad f^{-1}(B) = \{x \in X \mid f(x) \in B\}.$$

Satz 8.10. *Es sei $f : X \to Y$ eine Abbildung zwischen metrischen Räumen.*

(i) *f ist genau dann stetig an der Stelle a, wenn es für jede Umgebung V von $f(a)$ eine Umgebung U von a gibt, so dass $f(U) \subseteq V$.*

(ii) *f ist genau dann stetig, wenn für jede offene Menge V in Y das Urbild $f^{-1}(V)$ offen in X ist.*

Beweis. **(i)** Es sei V eine Umgebung von $f(a)$. Dann gibt es ein $\varepsilon > 0$, so dass $U_\varepsilon^Y(f(a)) \subseteq V$. Ist f stetig an der Stelle a, so gibt es ein $\delta > 0$, so dass $f(U_\delta^X(a)) \subseteq U_\varepsilon^Y(f(a))$. Also ist das Kriterium mit $U = U_\delta^X(a)$ erfüllt.

Ist umgekehrt das Kriterium erfüllt und $\varepsilon > 0$ gegeben, so wenden wir es auf $V = U_\varepsilon^Y(f(a))$ an und erhalten eine Umgebung U von a, so dass

$f(U) \subseteq U_\varepsilon^Y(f(a))$. Wegen der Offenheit von U gibt es ein $\delta > 0$, so dass $U_\delta^X(a) \subseteq U$. Somit gilt $f(U_\delta^X(a)) \subseteq U_\varepsilon^Y(f(a))$, d. h. f ist stetig an der Stelle a.
(ii) Angenommen, f ist stetig. Für eine offene Menge V von Y wollen wir zeigen, dass $f^{-1}(V)$ offen ist. Ist a in diesem Urbild, so ist $f(a) \in V$, und nach Teil (i) gibt es eine Umgebung U von a, so dass $f(U) \subseteq V$. Das bedeutet $U \subset f^{-1}(V)$, und da a in der Menge $f^{-1}(V)$ beliebig gewählt war, ist sie offen.

Umgekehrt sei das Kriterium erfüllt. Ist nun $a \in X$, so ist für jede Umgebung V von $f(a)$ die Menge $U = f^{-1}(V)$ eine Umgebung von a mit der Eigenschaft $f(U) \subseteq V$. Somit ist f nach Teil (i) an der Stelle a stetig. $\square$

Bemerkung. Wegen $f^{-1}(Y \setminus V) = X \setminus f^{-1}(V)$ sehen wir unter Verwendung von Satz 8.2(iv), dass eine Abbildung genau dann stetig ist, wenn die Urbilder aller abgeschlossenen Mengen abgeschlossen sind.

Beispiel. Ist $f : X \to \mathbb{R}$ eine stetige Funktion, so ist für jedes $b \in \mathbb{R}$ die Menge

$$\{x \in X \mid f(x) < b\} = f^{-1}\big(]-\infty, b[\big)$$

offen und die sogenannte *Niveaufläche*

$$\{x \in X \mid f(x) = b\} = f^{-1}(\{b\})$$

abgeschlossen. Wenden wir dies auf eine Norm auf einem Vektorraum an, so erhalten wir eine offene Einheitskugel $\{x \in V \mid \|x\| < 1\}$ und eine abgeschlossene Einheitssphäre $\{x \in V \mid \|x\| = 1\}$. $\triangleleft$

Für die Verkettung gilt in Analogie zu Satz 5.4:

Satz 8.11. *Ist die Abbildung $f : X \to Y$ stetig an der Stelle a und die Abbildung $g : Y \to Z$ stetig an der Stelle $b = f(a)$, so ist die Abbildung $g \circ f : X \to Z$ stetig an der Stelle a.*

Beweis. Ist W eine Umgebung von $g(f(a)) = g(b)$ in Z, so gibt es wegen der Stetigkeit von g an der Stelle b eine Umgebung V von b in Y, so dass $g(V) \subseteq W$, und wegen der Stetigkeit von f an der Stelle a gibt es eine Umgebung U von a in X, so dass $f(U) \subseteq V$. Es folgt $g \circ f(U) \subseteq W$, und weil W beliebig war, ist $g \circ f$ an der Stelle a stetig. $\square$

Alternativ hätte man den Beweis von Satz 5.4 kopieren oder Satz 8.9 benutzen können. Aus den Sätzen 5.3, 8.8 und 8.11 ergibt sich:

Folgerung 8.3. *Jede Abbildung $D \to K^n$, deren Koordinatenfunktionen auf D durch Terme gegeben sind, in denen nur arithmetische Operationen und stetige Funktionen vorkommen, stetig.*

Nun kommen wir auf die Frage nach der Vollständigkeit von Funktionenräumen zurück. Es sei weiterhin $K \in \{\mathbb{R}, \mathbb{C}\}$.

Vorlesung 7a
23.11.22

Satz 8.12. *Es sei X ein metrischer Raum. Wenn eine Folge stetiger Abbildungen $f_k : X \to K^n$ gleichmäßig gegen eine Abbildung $f : X \to K^n$ konvergiert, so ist f stetig.*

Beweis. Zur Vereinfachung der Bezeichnungen sei $n = 1$. Ist $\varepsilon > 0$, so gibt es wegen der gleichmäßigen Konvergenz eine natürliche Zahl k, so dass

$$\sup_{x \in X} |f_k(x) - f(x)| < \frac{\varepsilon}{3}.$$

Nun sei $a \in X$. Wegen der Stetigkeit von f_k gibt es eine Umgebung U von a, so dass für $x \in U$ gilt

$$|f_k(x) - f_k(a)| < \frac{\varepsilon}{3}.$$

Für $x \in U$ folgt nun mit der Dreiecksungleichung

$$|f(x) - f(a)| \leq |f(x) - f_k(x)| + |f_k(x) - f_k(a)| + |f_k(a) - f(a)| \leq \varepsilon.$$

Da ε beliebig war, folgt die Stetigkeit von f an der Stelle a. $\square$

Für allgemeines n muss man den Absolutbetrag durch eine Norm auf K^n ersetzen, die aber bezeichnungsmäßig von der Supremumsnorm auf dem Funktionenraum zu unterscheiden ist.

Der selbe Beweis zeigt, dass für eine Folge stetiger Abbildungen $f_k : X \to Y$ zwischen metrischen Räumen, die bezüglich der Supremumsmetrik gleichmäßig gegen eine Abbildung $f : X \to Y$ konvergiert, die Abbildung f stetig ist.

Nach dem Satz ist der Raum der stetigen beschränkten Funktionen $X \to K^n$ abgeschlossen im Raum aller beschränkten Funktionen $X \to K^n$. In Verbindung mit Lemma 8.3 und Lemma 8.4 erhalten wir:

Folgerung 8.4. *Der Raum der beschränkten stetigen Funktionen auf einem metrischen Raum mit Werten in $\mathbb{R}^n$ oder $\mathbb{C}^n$ ist vollständig.*

Die Umkehrabbildung einer bijektiven stetigen Abbildung braucht natürlich nicht stetig zu sein.

Definition 8.14. *Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen heißt* Homöomorphismus, *wenn sie eine stetige Umkehrabbildung besitzt. Zwei metrische Räume heißen* homöomorph, *wenn es zwischen ihnen einen Homöomorphismus gibt.*

Beispiel. Es sei $\| \,.\, \|$ eine Norm auf einem Vektorraum V und

$$f(x) = \frac{x}{1 + \|x\|}.$$

für $x \in V$. Dann ist

$$\|f(x)\| = \frac{\|x\|}{1 + \|x\|} < 1,$$

wir erhalten also eine Abbildung $f : V \to B = \{x \in V \mid \|x\| < 1\}$. Außerdem folgt

$$1 - \|f(x)\| = \frac{1}{1 + \|x\|}.$$

Wir können nun die Definition von f nach x auflösen. Setzen wir für $y \in B$

$$g(y) = \frac{y}{1 - \|y\|},$$

so ist g die Umkehrabbildung von f. Da die Norm Lipschitz-stetig ist, sind f und g stetig, und somit sind V und B homöomorph. ◁

8.6 Zusammenhang

Der Zwischenwertsatz ist äquivalent zu der Aussage, dass der Wertebereich einer stetigen Funktion auf einem Intervall ebenfalls ein Intervall ist (vgl. Folgerung 5.2). Dies wollen wir nun verallgemeinern.

Definition 8.15. *Es sei X ein metrischer Raum. Für Punkte a und b von X verstehen wir unter einem* Weg *von a nach b in X eine stetige Abbildung $g : [0, 1] \to X$, so dass $g(0) = a$ und $g(1) = b$. Der Raum X heißt* wegzusammenhängend[46]*, wenn es für beliebige Punkte a und b von X einen Weg von a nach b in X gibt.*

Wenn wir eine Teilmenge eines metrischen Raumes wegzusammenhängend nennen, so betrachten wir sie natürlich als Teilraum. Nach dem Zwischenwertsatz muss eine wegzusammenhängende Teilmenge von $\mathbb{R}$ ein Intervall sein, und die Umkehrung ist offensichtlich.

Man kann auf einem metrischen Raum X eine Relation $\sim$ definieren, indem man festlegt, dass $a \sim b$ ist, wenn es einen Weg von a nach b gibt. Dies ist offensichtlich eine Äquivalenzrelation, und ihre Äquivalenzklassen nennt man *Wegzusammenhangskomponenten*.

[46] auch *bogenzusammenhängend* oder *linear zusammenhängend* genannt

Beispiel. Die Teilräume

$$Y = \left\{(x, y) \in \mathbb{R}^2 \;\middle|\; x > 0,\ y = \cos\frac{1}{x}\right\}, \qquad Z = \{(0, 0)\}$$

von $\mathbb{R}^2$ sind offensichtlich wegzusammenhängend. Wir behaupten, dass der Teilraum $X = Y \cup Z$ nicht wegzusammenhängend ist, also die Wegzusammenhangskomponenten Y und Z hat. Angenommen, es gibt einen Weg g von $(0, 0)$ nach $(1, \cos 1)$. Dann ist $g^{-1}(0, 0)$ abgeschlossen, hat also ein Maximum c. Nach Stetigkeit gibt es ein $\delta > 0$, so dass für $c \leq t < c + \delta$ gilt $|g_2(t)| < 1$. Dann kann g_1 auf $[c, c + \delta[$ nicht die Werte $1/k\pi$ mit $k \in \mathbb{N} \setminus \{0\}$ annehmen, also wegen dem Zwischenwertsatz keine positiven Werte haben. Das widerspricht der Definition von c. $\triangleleft$

Der Punkt $(0, 0)$ der einen Komponente Z ist hier ein Häufungspunkt der anderen Komponente Y, was der Intuition widerspricht. Darum betrachtet man noch einen anderen Zusammenhangsbegriff.

Definition 8.16. *Ein metrischer Raum X heißt unzusammenhängend, wenn er zwei nichtleere abgeschlossene Teilmengen A und B besitzt, so dass*

$$A \cup B = X, \qquad A \cap B = \varnothing.$$

Andernfalls heißt er zusammenhängend[47].

Wegen Satz 8.2(iv) hätte man genauso gut verlangen können, dass A und B offen sind oder dass A offen und abgeschlossen ist.

Lemma 8.5. *Ein metrischer Teilraum von $\mathbb{R}$ ist genau dann zusammenhängend, wenn er ein Intervall ist.*

Beweis. Ist $X \subseteq \mathbb{R}$ kein Intervall, so gibt es a, $b \in X$ und $c \in \mathbb{R} \setminus X$, so dass $a < c < b$. Die Mengen

$$A =]-\infty, c[\, \cap X, \qquad B =]c, \infty[\, \cap X$$

sind nach Satz 8.5 offen in X, es gilt $a \in A$, $b \in B$, und man sieht, dass X unzusammenhängend ist.

Nun sei I ein Intervall. Angenommen, $I = A \cup B$, wobei A und B disjunkt, nicht leer und abgeschlossen im Teilraum I sind. Es sei b ein innerer Punkt von I. Nach eventueller Umbenennung können wir annehmen, dass $b \in B$ ist, und nach eventueller Multiplikation mit -1 können wir annehmen, dass $A \cap\,]-\infty, b]$ nicht leer ist und folglich ein Supremum a besitzt. Nach dessen

[47] auch *Hausdorff-zusammenhängend* genannt

Definition gibt es für jedes $\varepsilon > 0$ ein Element $x \in A$ mit $a-\varepsilon < x \leq a$, und da A abgeschlossen ist, folgt $a \in A$. Wegen der Disjunktheit von A und B folgt $a < b$, so dass die Teilmenge $]a, b[$ von B nicht leer ist. Da B abgeschlossen ist, folgt $a \in B$, und wir haben einen Widerspruch zur Disjunktheit. $\square$

Im Allgemeinen ist es schwer unmittelbar nachzuweisen, dass ein metrischer Raum nicht wegzusammenhängend ist oder dass er zusammenhängend ist. Der folgende Satz erleichtert das.

Satz 8.13. *Jeder wegzusammenhängende metrische Raum ist zusammenhängend.*

Beweis. Angenommen, wir haben eine Zerlegung $X = A \cup B$ in disjunkte nichtleere abgeschlossene Teilmengen. Dann wählen wir $a \in A$ und $b \in B$. Ist X wegzusammenhängend, so gibt es einen Weg g von a nach b, und die Mengen $g^{-1}(A)$ und $g^{-1}(B)$ sind offensichtlich disjunkt und nach Satz 8.10 offen. Wegen $0 \in g^{-1}(A)$ und $1 \in g^{-1}(B)$ sind sie nicht leer, was Lemma 8.5 widerspricht. $\square$

Es gibt zusammenhängende metrische Räume, die nicht wegzusammenhängend sind.

Beispiel. Wir behaupten, dass der Teilraum $X = Y \cup Z$ aus dem vorigen Beispiel zusammenhängend ist. Hätten wir nämlich eine Zerlegung $X = A \cup B$ in nichtleere disjunkte offene Teilmengen, so wäre $Y = (Y \cap A) \cup (Y \cap B)$ eine Zerlegung in disjunkte offene Teilmengen. Da Y zusammenhängend ist, muss dann einer der beiden Teile leer sein, also $Y \subseteq A$ oder $Y \subseteq B$. Das Gleiche gilt für Z, so dass nur $A = Y$, $B = Z$ oder umgekehrt in Frage kommt. Die Menge Y ist zwar offen in X, die Menge Z jedoch nicht, wie schon bemerkt. $\triangleleft$

Der Zwischenwertsatz verallgemeinert sich wie folgt.

Satz 8.14. *Es sei $f : X \to Y$ eine surjektive stetige Abbildung.*

(i) *Ist X wegzusammenhängend, so auch Y.*

(ii) *Ist X zusammenhängend, so auch Y.*

Bei nicht surjektiven Abbildungen sagt der Satz etwas über den Wertebereich $f(X)$ aus.

Beweis. **(i)** Für $a, b \in Y$ gibt es $u, v \in X$, so dass $f(u) = a$ und $f(v) = b$. Ist X wegzusammenhängend, so gibt es einen Weg g von u nach v, und dann ist $f \circ g$ nach Satz 8.11 ein Weg von a nach b.

(ii) Ist Y nicht zusammenhängend, so gibt es eine Zerlegung $Y = A \cup B$ wie in der Definition. Nun sind die Mengen $f^{-1}(A) = f^{-1}(U)$ und $f^{-1}(B) = f^{-1}(V)$ nach Satz 8.10 abgeschlossen und bilden eine Zerlegung von X, die zeigt, dass auch X nicht zusammenhängend ist. □

Als Anwendung beweisen wir die Starrheit von K-analytischen Funktionen, wobei $K = \mathbb{R}$ oder $K = \mathbb{C}$.

Satz 8.15. *Es sei f eine K-analytische Funktion auf einer zusammenhängenden offenen Teilmenge U von K. Hat die Nullstellenmenge von f einen Häufungspunkt in U, so ist $f(x) = 0$ für alle $x \in U$.*

Beweis. Es sei N die Nullstellenmenge von f und A die Menge der Häufungspunkte von N, die nach Voraussetzung nicht leer ist. Wie in der Lösung von Aufgabe 24(b) sieht man, dass die Menge der Häufungspunkte einer beliebigen Menge abgeschlossen ist. Laut Satz 4.11 ist jeder Punkt von A ein innerer Punkt von N, und jeder innere Punkt von N liegt offensichtlich in A. Also ist A auch offen, und da U zusammenhängend ist, muss $A = U$ sein. □

Nach dem Satz ist eine K-analytische Funktion auf einer zusammenhängenden offenen Menge durch die Vorgabe ihrer Werte auf einer beliebig kleinen Teilmenge bestimmt, vorausgesetzt, diese hat wenigstens einen Häufungspunkt.

8.7 Kompaktheit

Vorlesung 7b
25.11.22

Ein wichtiges Ergebnis der Analysis ist der Satz von Bolzano-Weierstraß, der in unserer Vorlesung als Satz 3.6 vorkommt. Metrische Räume, in denen eine entsprechende Aussage gilt, erhalten einen besonderen Namen.

Definition 8.17. *Ein metrischer Raum X heißt* kompakt*, wenn jede Folge in X einen Häufungspunkt besitzt.*

Wenn eine Teilmenge eines metrischen Raumes kompakt genannt wird, so ist sie als metrischer Teilraum gemeint. Das bedeutet also, dass jede Folge in ihr einen Häufungspunkt hat, der ebenfalls zu dieser Teilmenge gehört. Aus dem Satz von Bolzano-Weierstraß ergibt sich:

Folgerung 8.5. *Eine Teilmenge von $\mathbb{R}$ ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.*

Jede Folge in einer solchen Menge ist nämlich eine beschränkte Folge, hat also nach dem Satz einen Häufungspunkt, und dieser gehört nach Definition der Abgeschlossenheit ebenfalls zur gegebenen Menge. Die Umkehrung folgt aus der Kontraposition von Satz 3.2 und der Tatsache, dass der Grenzwert der einzige Häufungspunkt einer konvergenten Folge ist.

Folgender Satz beschert uns weitere Beispiele kompakter metrischer Räume.

Satz 8.16. *Sind X und Y kompakte metrische Räume mit Metriken d bzw. e, so ist auch $X \times Y$ kompakt bezüglich der Metrik*

$$g\big((x',y'),(x'',y'')\big) = \max\{d(x',x''), e(y',y'')\}.$$

Beweis. Es sei (x_k, y_k) eine Folge in $X \times Y$. Da X kompakt ist, hat die Folge x_k einen Häufungspunkt a, und nach Satz 8.2(iii) hat sie eine Teilfolge x_{k_l}, die gegen a konvergiert. Da Y kompakt ist, hat die Folge y_{k_l} einen Häufungspunkt b. Nun hat die Folge (x_{k_l}, y_{k_l}) bezüglich der Metrik g den Häufungspunkt (a, b). □

Hier sind einige Eigenschaften kompakter Mengen:

Satz 8.17. (i) *Jede abgeschlossene Teilmenge eines kompakten metrischen Raumes ist kompakt.*

(ii) *Jeder kompakte metrische Raum ist vollständig.*

(iii) *Jede kompakte Teilmenge eines metrischen Raumes ist abgeschlossen in diesem Raum.*

(iv) *Ist $f : X \to Y$ eine surjektive stetige Abbildung zwischen metrischen Räumen und ist X kompakt, so ist auch Y kompakt.*

Beweis. **(i)** Ist X kompakt, so ist jede Folge in einer Teilmenge A auch eine in X und hat somit einen Häufungspunkt, der zu A gehört, wenn A in X abgeschlossen ist.
(ii) Es sei x_k eine Cauchyfolge in X. Ist $\varepsilon > 0$ gegeben, so gibt es ein k_0, so dass für k, $l \geq k_0$ gilt $d(x_k, x_l) < \frac{\varepsilon}{2}$. Ist X kompakt, so hat die Folge einen Häufungspunkt a. Also kann man l so wählen, dass $d(x_l, a) < \frac{\varepsilon}{2}$. Mit der Dreiecksungleichung folgt $d(x_k, a) < \varepsilon$. Da ε beliebig war, konvergiert die Folge x_k gegen a.
(iii) Ist A eine kompakte Teilmenge von X, so ist A nach (ii) vollständig und nach Lemma 8.3 abgeschlossen in X.
(iv) Es sei y_k eine Folge in Y. Wegen der Surjektivität existiert für jedes k ein Element x_k von X, so dass $f(x_k) = y_k$. Wegen der Kompaktheit von X hat die Folge x_k einen Häufungspunkt a, und nach Satz 8.2(iii) gibt es eine Teilfolge x_{k_l}, die gegen a konvergiert. Nach Satz 8.9 konvergiert dann die Teilfolge y_{k_l} gegen $b = f(a)$, und nach Satz 8.2(iii) ist b ein Häufungspunkt der Folge y_k. □

Definition 8.18. *Ein metrischer Raum X heißt* beschränkt*, wenn sein* Durchmesser

$$\operatorname{diam} X = \sup\{d(x,y) \mid x,\, y \in X\}$$

endlich ist. Ein metrischer Raum X heißt absolut beschränkt*, wenn für jedes $\varepsilon > 0$ eine endliche Teilmenge A existiert, so dass es für jedes $x \in X$ ein $a \in A$ mit $d(x,a) < \varepsilon$ gibt.*

Auch diese Eigenschaften kann man auf Teilmengen metrischer Räume anwenden, indem man sie als Teilräume betrachtet. Mit Hilfe der Dreiecksungleichung sieht man leicht, dass ein metrischer Raum genau dann beschränkt ist, wenn es einen Punkt a gibt, so dass $d(x,a)$ als Funktion von x beschränkt ist, und dass dies dann für jeden Punkt a gilt. Eine Teilmenge A wie in der Definition nennt man ein ε-Netz, für sie gilt $\operatorname{diam} X \leq \operatorname{diam} A + 2\varepsilon$.

Aus Satz 8.17(iv) und Folgerung 8.5 ergibt sich

Folgerung 8.6. *Jeder kompakte metrische Raum ist beschränkt.*

Stärker ist folgendes Kompaktheitskriterium.

Satz 8.18. *Ein metrischer Raum X ist genau dann kompakt, wenn er vollständig und absolut beschränkt ist.*

Beweis. Ist der metrische Raum X kompakt, so ist er nach Satz 8.17(ii) vollständig. Angenommen, er ist nicht absolut beschränkt. Dann gibt es ein $\varepsilon > 0$, für das kein ε-Netz existiert. Wir konstruieren nun rekursiv eine Folge $x_1,\, x_2,\, \ldots$ mit $d(x_k, x_l) \geq \varepsilon$ für alle $k \neq l$. Da X nicht leer ist, existiert ein $x_1 \in X$, und sind $x_1,\, \ldots,\, x_n$ bereits konstruiert, so bilden sie kein ε-Netz, also gibt es einen Punkt $x_{n+1} \in X$ mit $d(x_{n+1}, x_k) \geq \varepsilon$ für alle $k \in \{1, \ldots, n\}$. Die Folge der x_k besitzt dann keinen Häufungspunkt im Widerspruch zur Kompaktheit von X.

Umgekehrt sei X vollständig und absolut beschränkt. Wir müssen dann für eine beliebige Folge von Punkten x_k einen Häufungspunkt finden. Dazu betrachten wir für jedes $n \in \mathbb{N} \setminus \{0\}$ ein $\frac{1}{n}$-Netz A_n. Wir finden dann rekursiv Punkte $a_n \in A_n$, so dass es für jedes n unendlich viele $k \in \mathbb{N}$ gibt, für die gilt

$$d(x_k, a_1) < 1, \quad d(x_k, a_2) < \frac{1}{2}, \quad \ldots \quad d(x_k, a_n) < \frac{1}{n}.$$

Für beliebige m, n folgt mit der Dreiecksungleichung

$$d(a_m, a_n) < \frac{1}{m} + \frac{1}{n},$$

also ist die Folge der a_n eine Cauchyfolge und hat somit einen Grenzwert a, und wegen der Stetigkeit von d folgt

$$d(a, a_n) \leq \frac{1}{n}.$$

Mit den obigen Ungleichungen folgt, dass es für jedes n ein k gibt, so dass

$$d(x_k, a) < \frac{2}{n},$$

also ist a auch ein Häufungspunkt der Folge der x_k. □

Wir erhalten jetzt eine Verallgemeinerung von Folgerung 8.5.

Folgerung 8.7. *Eine Teilmenge von $\mathbb{R}^n$ oder $\mathbb{C}^n$ ist genau dann kompakt, wenn sie bezüglich $\| \,.\, \|_\infty$ beschränkt und abgeschlossen ist.*

Eine Teilmenge in $\mathbb{R}^n$ ist nämlich genau dann beschränkt, wenn sie in einem kartesischen Produkt beschränkter Intervalle enthalten ist.

Beispiel. Es sei V der Vektorraum der beschränkten Folgen in $\mathbb{R}$ mit der Norm $\|x\|_\infty = \sup\big\{|x_k| \,\big|\, k \in \mathbb{N}\big\}$. Nach Lemma 8.4 ist V vollständig. Die Teilmenge

$$S = \{x \in V \mid \|x\|_\infty = 1\}$$

ist beschränkt und abgeschlossen, also nach Lemma 8.3 ebenfalls vollständig. Sie ist aber nicht kompakt, denn für die Folge der Vektoren e_k, bei denen das k-te Glied gleich 1 und alle anderen gleich 0 sind, gilt $\|e_k - e_l\|_\infty = 1$ für $k \neq l$. $\triangleleft$

Nun kommen wir zu den Anwendungen der Kompaktheit. Als Erstes formulieren wir das Analogon von Satz 5.6.

Folgerung 8.8. *Ist X ein kompakter metrischer Raum und $f : X \to \mathbb{R}$ eine stetige Funktion, so besitzt f ein Maximum und ein Minimum.*

Beweis. Der Wertebereich ist nämlich nach Satz 8.17(iv) kompakt, also nach Folgerung 8.5 beschränkt und abgeschlossen, und seine obere und untere Grenze sind Häufungspunkte. $\square$

Folgerung 8.9. *Ist $f : X \to Y$ eine bijektive stetige Abbildung zwischen metrischen Räumen und ist X kompakt, so ist f ein Homöomorphismus.*

Beweis. Ist A eine abgeschlossene Teilmenge von X, so ist sie nach Satz 8.17(i) kompakt. Ihr Bild $f(A)$ ist nach Satz 8.17(iv) kompakt und nach Satz 8.17(iii) abgeschlossen. Das Bild unter der Abbildung f ist gleich dem Urbild unter der Umkehrabbildung f^{-1}. Entsprechend der Bemerkung nach Satz 8.10 ist f^{-1} stetig. $\square$

Definition 8.19. *Es sei X ein metrischer Raum. Für Punkte x und Teilmengen A und B von X definieren wir*

$$\begin{aligned} d(x,B) = d(B,x) &= \inf\{d(x,y) \mid y \in B\}, \\ d(A,B) &= \inf\{d(x,y) \mid x \in A,\ y \in B\}. \end{aligned}$$

Beispiel. Für die Teilmengen $A = \mathbb{N} \setminus \{0\}$ und $B = \{k - \frac{1}{k} \mid k \in A\}$ von $\mathbb{R}$ gilt $d(A,B) = 0$, obwohl sie abgeschlossen und disjunkt sind. $\triangleleft$

Satz 8.19. *Es sei X ein metrischer Raum.*

(i) *Für jede Teilmenge B ist die durch $f(x) = d(x, B)$ definierte Funktion $f : X \to \mathbb{R}$ Lipschitz-stetig.*

(ii) *Ist A abgeschlossen, B kompakt und $A \cap B = \varnothing$, so ist $d(A, B) > 0$.*

Beweis. **(i)** Für alle x, y, $z \in X$ gilt nach der Dreiecksungleichung

$$d(x, z) \leq d(x, y) + d(y, z).$$

Bilden wir das Infimum über alle $z \in B$, so folgt

$$d(x, B) \leq d(x, y) + d(y, B).$$

Durch Vertauschung von x und y ergibt sich eine weitere Ungleichung. Fassen wir beide zusammen, so erhalten wir

$$|f(x) - f(y)| \leq d(x, y).$$

(ii) Wegen

$$\{d(x, y) \mid x \in A,\ y \in B\} = \bigcup_{y \in B} \{d(x, y) \mid x \in A\}$$

gilt $d(A, B) = \inf\{d(A, y) \mid y \in B\}$. Ist B kompakt, so gibt es nach Folgerung 8.8 ein $b \in B$, so dass $d(A, B) = d(A, b)$. Ist A abgeschlossen, so gibt es wegen $b \notin A$ ein $\varepsilon > 0$, so dass $U_\varepsilon(b) \cap A = \varnothing$. Also gilt $d(A, b) \geq \varepsilon$ und somit $d(A, B) \geq \varepsilon$. □

Satz 8.20. *Es sei $K = \mathbb{R}$ oder $K = \mathbb{C}$.* Vorlesung 8a 30.11.22

(i) *Alle Normen auf einem endlichdimensionalen K-Vektorraum sind äquivalent.*

(ii) *Jede K-lineare Abbildung zwischen endlichdimensionalen K-Vektorräumen ist Lipschitz-stetig.*

Beweis. **(i)** Die Menge $S = \{x \in \mathbb{R}^n \mid \|x\|_\infty = 1\}$ ist als Urbild einer abgeschlossenen Menge bezüglich $\|\,.\,\|_\infty$ abgeschlossen und offensichtlich beschränkt, also nach Folgerung 8.7 kompakt. Ist $\|\,.\,\|'$ eine beliebige Norm auf $\mathbb{R}^n$, so gilt nach der Dreiecksungleichung

$$\|x\|' \leq \sum_{i=1}^{n} |x_i|\, \|e_i\|'$$

mit den Basisvektoren e_i wie oben, also finden wir wie in Lemma 8.2 ein c, so dass für alle $x \in V$ gilt

$$\|x\|' \leq c\|x\|_\infty.$$

Die Norm $\|\,.\,\|'$ ist also bezüglich $\|\,.\,\|_\infty$ Lipschitz-stetig, und ihre Einschränkung auf S hat nach Satz 8.8 ein Minimum ε, welches nach Eigenschaft (iii) der Norm positiv ist. Ist $x \neq 0$ und $t = \|x\|_\infty$, so ist $t^{-1}x \in S$, also

$$\|x\|' = t\|t^{-1}x\|' \geq t\varepsilon,$$

und mit $c' = \frac{1}{\varepsilon}$ folgt

$$\|x\|_\infty \leq c'\|x\|'.$$

Dies gilt offensichtlich auch für $x = 0$, also ist jede Norm auf $\mathbb{R}^n$ äquivalent zu $\|\,.\,\|_\infty$. Da jeder endlichdimensionale $\mathbb{R}$-Vektorraum V isomorph zu $\mathbb{R}^n$ ist, sind somit sind alle Normen auf V äquivalent. Das Selbe gilt auch für einen beliebigen $\mathbb{C}$-Vektorraum W, denn jede Norm auf W ist auch eine Norm bezüglich der unterliegenden Struktur[48] eines $\mathbb{R}$-Vektorraums.
(ii) Es sei $f : K^n \to K^m$ K-linear. Dann gibt es Elemente $a_{ij} \in K$, so dass die Koordinatenfunktionen von f gegeben sind durch

$$f_i(x) = \sum_{j=1}^{n} a_{ij}x_j.$$

Nun gilt

$$\|f(x)\|_1 \leq \|x\|_\infty \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|,$$

also ist f Lipschitz-stetig bezüglich gewisser Normen. Mit Teil (i) folgt die Behauptung für beliebige Normen auf beliebigen endlichdimensionalen K-Vektorräumen. $\square$

Definition 8.20. *Man nennt die Lipschitz-Konstante*

$$\|f\| = \sup_{x \in V \setminus \{0\}} \frac{\|f(x)\|'}{\|x\|}$$

die Norm der linearen Abbildung $f : V \to V'$ *bezüglich der gegebenen Normen auf V und V'.*

Auch der Begriff der gleichmäßigen Stetigkeit verallgemeinert sich auf metrische Räume.

[48] Diese Struktur besteht aus der selben Addition und aus der Skalarmultiplikation lediglich mit Elementen des Teilkörpers $\mathbb{R}$.

Definition 8.21. *Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen heißt* gleichmäßig stetig*, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle u, $v \in X$ gilt*

$$d(u,v) < \delta \quad \Longrightarrow \quad e(f(u), f(v)) < \varepsilon.$$

Zwei Metriken d und d^* auf einer Menge heißen äquivalent*, wenn die identischen Abbildungen $(X,d) \to (X,d^*)$ und $(X,d^*) \to (X,d)$ gleichmäßig stetig sind.*

Offensichtlich bildet eine gleichmäßig stetige Abbildung konvergente Folgen auf konvergente Folgen und Cauchy-Folgen auf Cauchy-Folgen ab. Diese Begriffe ändern sich also nicht, wenn man zu einer äquivalenten Metrik übergeht.

Auch Satz 5.13 verallgemeinert sich.

Satz 8.21. *Ist X kompakt und $f : X \to Y$ stetig, so ist f gleichmäßig stetig.*

Der damalige Beweis überträgt sich, wobei der Satz von Bolzano-Weierstraß durch die Kompaktheit von X zu ersetzen ist.

Zum Abschluss noch eine Warnung: Der hier eingeführte Begriff der Kompaktheit wird oft Folgenkompaktheit genannt, denn für topologische Räume gibt es einen besser geeigneten Begriff, der zunächst Bikompaktheit genannt wurde und der die Folgenkompaktheit als Standardbedeutung der Kompaktheit bald verdrängt hat. Im Spezialfall metrischer Räume sind beide Begriffe aber äquivalent.

9 Differentiation und Integration vektorwertiger Funktionen

9.1 Definition und Eigenschaften

Der Begriff der Ableitung verallgemeinert sich auf Funktionen mit Werten in einem K-Vektorraum V, wobei $K = \mathbb{R}$ oder $K = \mathbb{C}$.

Definition 9.1. *Es sei $D \subset K$ und $a \in D$ ein Häufungspunkt von D. Eine Funktion $f : D \to V$ heißt* differenzierbar *an der Stelle a, wenn der* Differenzenquotient

$$\frac{1}{x-a}(f(x) - f(a))$$

einen Grenzwert für $x \to a$ besitzt. Dieser heißt dann Ableitung *der Funktion f an der Stelle a, abgekürzt $f'(a)$.*

Man kann auch die Ableitung einer Funktion f mit Werten in einem affinen Raum X über K als

$$f'(a) = \lim_{x \to a} \frac{1}{x-a} \overrightarrow{f(a)f(x)}$$

definieren, falls der Grenzwert in dem zu X gehörigen Vektorraum existiert.

Bisher kennen wir nur das Integral von Funktionen $f : [a, b] \to K$. Dabei haben wir Teilungen $T = \{x_0, \dots, x_m\}$ von $[a, b]$ und zugehörige Mengen $Z = \{z_1, \dots, z_m\}$ von Stützstellen für T betrachtet, d. h.

$$a = x_0 \le z_1 \le x_1 \le z_2 \le x_2 \le \dots \le z_m \le x_m = b.$$

Die Feinheit von T war als $\max\bigl\{x_k - x_{k-1} \mid k \in \{1, \dots, m\}\bigr\}$ definiert. Die *Riemannsche Summe*

$$S(f, T, Z) = \sum_{k=1}^{m} (x_k - x_{k-1}) f(z_k)$$

ergibt auch für vektorwertige Funktionen einen Sinn, und in Anlehnung an Definition 7.5 setzen wir fest:

Definition 9.2. *Es sei V ein endlichdimensionaler Vektorraum über K und $f : [a, b] \to V$. Ein Element $I \in V$ heißt* Integral *der Funktion f über das Intervall $[a, b]$, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für jede Teilung T von $[a, b]$ mit einer Feinheit kleiner als δ und für jede Menge von Stützstellen Z für T gilt*

$$\|S(f, T, Z) - I\| < \varepsilon.$$

Die Funktion f heißt integrierbar*, wenn ein Integral existiert.*

Wenn f integrierbar ist, so gibt es offenbar nur ein Integral I, und wir bezeichnen es mit

$$I = \int_a^b f(x)\, dx.$$

Wegen Satz 8.20(i) hängen Ableitung und Integral einer vektorwertigen Funktion nicht von der Wahl der Norm auf V ab. Am einfachsten, wenn auch nicht besonders elegant, ist es, eine Basis zu wählen und alles auf den skalaren Fall zurückzuführen.

Lemma 9.1. (i) *Eine Funktion $f : D \to K^n$ ist genau dann differenzierbar an der Stelle a, wenn ihre Koordinatenfunktionen $f_1, \dots, f_n : D \to K$ an der Stelle a differenzierbar sind, und dann ist*

$$f'(a) = \bigl(f_1'(a), \dots, f_n'(a)\bigr).$$

(ii) *Eine Funktion $f : [a,b] \to K^n$ ist genau dann integrierbar, wenn ihre Koordinatenfunktionen integrierbar sind, und dann ist*

$$\int_a^b f(x)\,dx = \left(\int_a^b f_1(x)\,dx, \ldots, \int_a^b f_1(x)\,dx\right).$$

Beweis. Nach den Definitionen gilt

$$\frac{1}{x-a}(f(x)-f(a)) = \left(\frac{f_1(x)-f_1(a)}{x-a}, \ldots, \frac{f_n(x)-f_n(a)}{x-a}\right),$$
$$S(f,T,Z) = \big(S(f_1,T,Z), \ldots, S(f_n,T,Z)\big),$$

und die Behauptungen folgen aus den früheren Bemerkungen über Grenzwerte vektorwertiger Funktionen. □

Die Ableitungsregeln aus Satz 6.1 verallgemeinern sich auf den vektorwertigen Fall, wobei wir allerdings für die Produktregel einen Begriff aus der linearen Algebra benötigen.

Definition 9.3. *Sind U, V und W Vektorräume über K, so sagen wir, eine Abbildung $b : U \times V \to W$ sei* bilinear, *wenn für jedes $u_0 \in U$ und jedes $v_0 \in V$ die Abbildungen*

$$v \mapsto b(u_0, v), \qquad u \mapsto b(u, v_0)$$

linear sind.

So ist z. B. ein Skalarprodukt im Fall $K = \mathbb{R}$ eine bilineare Abbildung $V \times V \to \mathbb{R}$. Auch die Aussagen der Sätze 7.1, 7.3 und 7.5 über Integrale verallgemeinern sich ohne Schwierigkeiten.

Satz 9.1. *Es seien U, V und W endlichdimensionale Vektorräume über K.*

(i) *Sind f, $g : D \to V$ an der Stelle a differenzierbar, so ist $f+g$ an der Stelle a differenzierbar, und es gilt die Summenregel*

$$(f+g)'(a) = f'(a) + g'(a).$$

(ii) *Sind $f : D \to U$ und $g : D \to V$ an der Stelle a differenzierbar und ist $b : U \times V \to W$ bilinear, so ist $b \circ (f,g)$ an der Stelle a differenzierbar, und es gilt die Produktregel*

$$\big(b \circ (f,g)\big)'(a) = b \circ (f',g)(a) + b \circ (f,g')(a).$$

(iii) *Ist $f'(x) = 0$ für alle $x \in [a, b]$, so ist f konstant auf $[a, b]$.*

(iv) *Jede stetige Funktion auf einem kompakten Intervall ist integrierbar.*

(v) *Ist $a \le b \le c$, so ist eine Funktion $f : [a, c] \to V$ genau dann integrierbar, wenn f über $[a, b]$ und über $[b, c]$ integrierbar ist, und dann gilt*

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx.$$

(vi) *Sind f, $g : [a, b] \to V$ integrierbar und $l : V \to W$ eine lineare Abbildung, so sind auch $f + g$, $l \circ f$ und $\|f\|$ integrierbar, und es gilt*

$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx,$$

$$\int_a^b l(f(x))\,dx = l\left(\int_a^b f(x)\,dx\right),$$

$$\left\|\int_a^b f(x)\,dx\right\| \le \int_a^b \|f(x)\|\,dx.$$

Beweis. Beim Beweis von Aussage (ii) benutzt man die Identität

$$b\big(f(x), g(x)\big) - b\big(f(a), g(a)\big) = b\big(f(x) - f(a), g(x)\big) + b\big(f(a), g(x) - g(a)\big).$$

Man kann b als Element von $\mathrm{Hom}(U, \mathrm{Hom}(V, W))$ auffassen, und durch zweimalige Anwendung von Satz 8.20(ii) finden wir ein $c > 0$, so dass für alle $u \in U$ und $v \in V$ gilt

$$\|b(u, v)\| \le c\|u\|\|v\|.$$

Aussage (iii) folgt mit Lemma 9.1(i) aus Folgerung 6.6, obwohl der Mittelwertsatz für $n > 1$ nicht gilt.

Aussage (iv) folgt mit Lemma 9.1(ii) aus Satz 7.1 und der Tatsache, dass die Koordinatenfunktionen einer stetigen Funktion stetig sind.

Die ersten beiden Identitäten in Teil (vi) ergeben sich aus den Eigenschaften

$$S(f + g, T, Z) = S(f, T, Z) + S(g, T, Z), \qquad S(l \circ f, T, Z) = l(S(f, T, Z)).$$

Beim Beweis der Integrierbarkeit auf der rechten Seite der Ungleichung in Teil (vi) benutzt man, dass im Fall $V = K^n$ und der Norm $\| \,.\, \|_1$ gilt

$$\sup_{x \in I_k} \|f(x)\|_1 - \inf_{x \in I_k} \|f(x)\|_1 \leq \sum_{j=1}^{n} (\sup_{x \in I_k} f_j(x) - \inf_{x \in I_k} f_j(x)),$$

wobei I_k das k-te Teilintervall der Teilung T bezeichnet, so dass

$$\overline{S}(\|f\|_1, T) - \underline{S}(\|f\|_1, T) \leq \sum_{j=1}^{n} \big(\overline{S}(f_j, T) - \underline{S}(f_j, Z)\big).$$

Ansonsten sind die Beweise identisch mit den früheren. □

Auf Grund von Satz 9.1(v) können wir $\int\limits_a^b f(x)\,dx$ auch wieder ohne die Voraussetzung $a \leq b$ definieren. Der Hauptsatz der Infinitesimalrechnung hat ebenfalls eine vektorwertige Version:

Satz 9.2. *Ist $f : [a,b] \to V$ stetig differenzierbar, so gilt*

$$\int_a^b f'(x)\,dx = f(b) - f(a).$$

Die Beweise der Sätze 7.8 und 7.9 übertragen sich nämlich wortwörtlich, wobei der Absolutbetrag durch eine Norm zu ersetzen ist.

Hat f Werte in einem affinen Raum, so ist die rechte Seite durch $\overrightarrow{f(a)f(b)}$ zu ersetzen.

9.2 Stetigkeit parameterabhängiger Integrale

Wenn eine Funktion f von vielen Variablen abhängt, so hält man oft einige Variablen fest und betrachtet die entstehende *partielle Funktion* der übrigen Variablen. Die festgehaltenen Variablen, hier zu einer Variablen t zusammengefasst, nennt man *Parameter*.

Im Folgenden sei T ein metrischer Raum, $[a,b]$ ein kompaktes Intervall und V ein endlichdimensionaler Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$. Wir wollen wissen, welche Bedingungen wir an eine Funktion $f : [a,b] \times T \to V$ stellen müssen, damit durch

$$I(t) = \int\limits_a^b f(x,t)\,dx$$

eine stetige Funktion I auf T definiert wird. Wir setzen

$$F(t)(x) = f(x,t),$$

d. h. der Wert $F(t)$ der Funktion F an der Stelle t ist seinerseits eine Funktion[49] auf $[a, b]$ mit Werten in V. Natürlich muss letztere Funktion auf $[a, b]$ integrierbar sein.

Satz 9.3. *Es sei F eine Abbildung von T in den Raum der integrierbaren Funktionen $[a, b] \to V$. Ist F an einer Stelle $u \in T$ stetig bezüglich der Supremumsnorm, so ist die oben definierte Funktion I stetig an der Stelle u.*

Beweis. Aus der Stetigkeit von F an der Stelle u folgt, dass es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $t \in T$ mit der Eigenschaft $d(t, u) < \delta$ gilt $\|F(t) - F(u)\| < \frac{\varepsilon}{b-a}$, d. h.

$$|f(x, t) - f(x, u)| < \frac{\varepsilon}{b - a}$$

für alle $x \in [a, b]$. Nach Satz 9.1(iv) ist

$$I(t) - I(u) = \int_a^b (f(x, t) - f(x, u))\, dx,$$

und für $d(t, u) < \delta$ folgt nach Satz 9.1(vi) und Satz 7.6

$$\|I(t) - I(u)\| \leq \int_a^b \|f(x, t) - f(x, u)\|\, dx < \varepsilon. \qquad \square$$

Vorlesung 8b
2.12.22

Man kann die Stetigkeit von I an der Stelle u auch durch

$$I(t) \to I(u) \ \ (t \to u)$$

ausdrücken. Wir fragen nun nach der Existenz dieses Grenzwertes, wenn $I(u)$ noch gar nicht definiert ist. Statt des Parameters t betrachten wir der Einfachheit halber einen Parameter $k \in \mathbb{N}$.

Satz 9.4. *Konvergiert eine Folge integrierbarer Funktionen $f_k : [a, b] \to V$ gleichmäßig gegen eine Funktion $f : [a, b] \to V$, so ist auch f integrierbar, und*

$$\int_a^b f_k(x)\, dx \to \int_a^b f(x)\, dx \ \ (k \to \infty).$$

[49] Dies hängt mit der Gleichmächtigkeit $V^{X \times T} \sim (V^X)^T$ aus Aufgabe I.15* zusammen.

Beweis. Wir bezeichnen das Integral auf der linken Seite mit I_k. Da die Folge f_k eine Cauchy-Folge bezüglich der Supremumsnorm ist, gibt es für jedes $\varepsilon > 0$ ein k_0, so dass für $k \geq k_0$ und $l \geq k_0$ und alle $x \in [a,b]$ gilt

$$\|f_k(x) - f_l(x)\| < \frac{\varepsilon}{b-a},$$

und wie im vorigen Beweis folgt $\|I_k - I_l\| < \varepsilon$. Die Folge I_k ist also auch eine Cauchy-Folge und hat wegen der Vollständigkeit von V einen Grenzwert I. Außerdem konvergiert nach Voraussetzung f_k gleichmäßig gegen f. Es gibt also für jedes $\varepsilon > 0$ ein k, so dass

$$\|I_k - I\| < \frac{\varepsilon}{3}, \qquad \|f_k(x) - f(x)\| < \frac{\varepsilon}{3(b-a)}$$

für alle $x \in [a,b]$. Daraus folgt, dass für alle Teilungen T von $[a,b]$ und alle zugehörigen Mengen Z von Stützstellen gilt

$$\|S(f_k,T,Z) - S(f,T,Z)\| < \frac{\varepsilon}{3}.$$

Laut Definition von I_k gibt es schließlich ein $\delta > 0$, so dass

$$\|S(f_k,T,Z) - I_k\| < \frac{\varepsilon}{3},$$

falls T eine Feinheit kleiner als δ hat. Mit der Dreiecksungleichung erhalten wir für solche Teilungen T

$$\|S(f,T,Z) - I\| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Da ε beliebig war, folgt die Integrierbarkeit von f und

$$\int_a^b f(x)\,dx = I. \qquad \square$$

Beispiel. Die durch

$$g(x) = \begin{cases} \frac{1}{q}, & \text{falls } x = \frac{p}{q} \text{ mit teilerfremden } p,\, q \in \mathbb{Z},\, q > 0, \\ 0 & \text{falls } x \notin \mathbb{Q} \end{cases}$$

definierte Funktion $g : \mathbb{R} \to \mathbb{R}$ ist auf jedem kompakten Intervall $[a,b]$ integrierbar, weil es für jedes $\varepsilon > 0$ nur endlich viele $x \in [a,b]$ gibt, so dass $g(x) > \varepsilon$. Das Gleiche gilt für die Funktionen $f_k = \sqrt[k]{g}$ mit $k \in \mathbb{N} \setminus \{0\}$. Die Folge konvergiert für $k \to \infty$ punktweise gegen die Funktion

$$f(x) = \begin{cases} 1, & \text{falls } x \in \mathbb{Q}, \\ 0 & \text{andernfalls,} \end{cases}$$

die nicht integrierbar ist. $\triangleleft$

9.3 Variation

Die Bewegung eines Punktes in einem Raum X während eines Zeitintervalls I kann man durch eine Abbildung $f : I \to X$ beschreiben. Wir interessieren uns für die Länge des zurückgelegten Weges.

Definition 9.4. *Ist I ein Intervall, X ein metrischer Raum und $f : I \to X$, so setzen wir für jede Teilung $T = \{t_0, \ldots, t_m\}$ eines Teilintervalls $[a, b] \subseteq I$*

$$V(f, T) = \sum_{k=1}^{m} d(f(t_{k-1}), f(t_k)).$$

Wir nennen

$$\operatorname{Var}_a^b(f) = \sup\{V(f, T) \mid T \textit{ ist Teilung von } [a, b].\}$$

die Variation *von f über $[a, b]$. Wir sagen, dass f von beschränkter Variation ist, wenn $\operatorname{Var}_a^b(f) < \infty$ für alle $a \leq b$ in I.*

Wir bemerken, dass

$$\operatorname{Var}_a^b(f) \geq V(f, \{a, b\}) = d(f(a), f(b)).$$

Beispiel. Die Abbildung $f : \mathbb{R} \to \mathbb{C}$ sei durch

$$f(t) = \begin{cases} t \exp \dfrac{i}{|t|}, & \text{wenn } t \neq 0, \\ 0 & \text{wenn } t = 0 \end{cases}$$

gegeben. Nach dem Einschließungskriterium ist $\lim\limits_{t \to 0} f(t) = 0$, also ist f stetig. Es gilt

$$\left| f\left(\frac{1}{k\pi}\right) - f\left(\frac{1}{(k-1)\pi}\right) \right| = \left| \frac{(-1)^k}{k\pi} - \frac{(-1)^{k-1}}{(k-1)\pi} \right| = \frac{1}{k\pi} + \frac{1}{(k-1)\pi} > \frac{2}{k\pi},$$

also

$$V\left(f, \left\{0, \frac{1}{m\pi}, \frac{1}{(m-1)\pi}, \ldots, \frac{1}{\pi}, 1\right\}\right) > \frac{2}{\pi} \sum_{k=2}^{m} \frac{1}{k},$$

und wegen der Divergenz der harmonischen Reihe ist $\operatorname{Var}_0^1(f) = \infty$. $\triangleleft$

Satz 9.5. *Es sei X ein metrischer Raum, I ein Intervall und $f : I \to X$.*

(i) *Sind I und J kompakte Intervalle und $h : J \to I$ monoton und bijektiv, so ist für $[a, b] \subseteq J$*

$$\operatorname{Var}_a^b(f \circ h) = \operatorname{Var}_{h(a)}^{g(b)} f.$$

(ii) *Für alle a, b, $c \in I$, wobei $a \le b \le c$ ist, gilt*

$$\operatorname{Var}_a^b f + \operatorname{Var}_b^c f = \operatorname{Var}_a^c f.$$

(iii) *Ist $X = V$ ein K-Vektorraum, so gilt für f, $g : I \to V$ und $c \in K$*

$$\operatorname{Var}_a^b(f+g) \le \operatorname{Var}_a^b(f) + \operatorname{Var}_a^b(g), \qquad \operatorname{Var}_a^b(c \cdot f) = |c| \operatorname{Var}_a^b(f).$$

(iv) *Ist $X = K^n$, so ist f genau dann von beschränkter Variation, wenn alle Koordinatenfunktionen f_j von beschränkter Variation sind.*

Beweis. **(i)** Ist T eine Teilung von $[h(a), h(b)]$, so ist $T' = h^{-1}(T)$ eine Teilung von $[a, b]$, und ist T' eine Teilung von $[a, b]$, so ist $T = h(T')$ eine von $[h(a), h(b)]$. In beiden Fällen gilt

$$V(f \circ h, T') = V(f, T).$$

Wir haben also das Supremum der selben Menge zu bilden.
(ii) Sind T_1 und T_2 Teilungen von $[a, b]$ bzw. $[b, c]$, so ist $T = T_1 \cup T_2$ eine Teilung von $[a, c]$, und

$$V(f, T_1) + V(f, T_2) = V(f, T).$$

Nun folgt die Behauptung ähnlich wie bei Satz 7.3.
(iii) Laut Dreiecksungleichung und Eigenschaft (i) der Norm gilt

$$V(f+g, T) \le V(f, T) + V(g, T), \qquad V(cf, T) = |c| V(f, T).$$

(iv) Ist f bezüglich einer Norm auf K^n von beschränkter Variation, so auch bezüglich jeder äquivalenten Norm. Im Falle $\|\,.\,\|_1$ ist

$$V(f, T) = V(f_1, T) + \ldots + V(f_n, T),$$

und die linke Seite ist genau dann unabhängig von T beschränkt, wenn es jeder Summand auf der rechten Seite ist. □

Wie schon bei Satz 7.3 bleibt die Behauptung von Satz 9.5(ii) ohne die Voraussetzung $a \le b \le c$ gültig, wenn wir für $a \ge b$ definieren $\operatorname{Var}_a^b(f) = -\operatorname{Var}_b^a(f)$.

Nun wollen wir den Begriff der Kurve[50] definieren. Dabei sollen der Einheitskreis S und die Hyperbel H, die als Lösungsmengen der Gleichungen

$$x^2 + y^2 = 1 \qquad \text{bzw.} \qquad x^2 - y^2 = 1$$

gegeben sind, Beispiele von Kurven in $\mathbb{R}^2$ sein.

[50]Die lateinischen Worte *linea curva* und *linea recta* bedeuten ursprünglich „gekrümmter Faden“ bzw. „gerader Faden“.

Definition 9.5. *Eine* einfache Kurve *in einem metrischen Raum X ist eine Teilmenge, die homöomorph zu einer Vereinigung von endlich vielen disjunkten Intervallen und Kreisen ist.*

Es genügt natürlich, jede Zusammenhangskomponente einer Kurve einzeln zu studieren. Darum befassen wir uns im Folgenden mit zusammenhängenden Kurven. Einen Homöomorphismus f von einem Intervall I oder der Kreislinie S auf eine zusammenhängende Kurve nennt man *Parametrisierung* der Kurve. Man kann sie als stetige injektive Abbildung in den Raum X auffassen. Im ersten Fall sprechen wir von einer offenen, im zweiten Fall von einer geschlossenen Kurve. Aus Folgerung 8.9 und Satz 8.14 ergibt sich:

Folgerung 9.1. *Jede stetige injektive Abbildung von S oder einem kompakten Intervall I in einen metrischen Raum ist eine Parametrisierung einer zusammenhängenden einfachen Kurve.*

Die Abbildung in dem Beispiel am Anfang des Abschnitts parametrisiert eine einfache offene Kurve.

Aus der Transitivität der Homöomorphie und den Sätzen 5.5 und 5.7 erhalten wir:

Folgerung 9.2. (i) *Sind $f : I \to X$ und $g : J \to X$ Parametrisierungen der selben offenen Kurve, so gibt es eine monotone bijektive Abbildung $h : I \to J$, so dass $f = g \circ h$.*

(ii) *Ist $g : J \to X$ eine Parametrisierung einer offenen Kurve, I ein Intervall und $h : I \to J$ eine monotone bijektive Abbildung, so ist auch $g \circ h$ eine Parametrisierung dieser Kurve.*

Man nennt h eine *Umparametrisierung*.

Die Bahn der Bewegung eines Massenpunktes, dessen Position zum Zeitpunkt t durch $f(t)$ gegeben ist, braucht keine einfache Kurve zu sein, weil f nicht injektiv sein muss. Darum definiert man noch einen weiteren Kurvenbegriff, wobei wir uns der Einfachheit halber auf offene Kurven beschränken.

Definition 9.6. *Es sei X ein metrischer Raum.*

(i) *Zwei auf Intervallen definierte stetige Abbildungen $f : I \to X$ und $g : J \to X$ heißen äquivalent, wenn es eine monotone bijektive Abbildung $h : I \to J$ gibt, so dass $f = g \circ h$.*

(ii) *Unter einer offenen* Kurve *in X verstehen wir eine Äquivalenzklasse von solchen Abbildungen, und jeden Repräsentanten nennen wir eine* Parametrisierung *der Kurve.*

(iii) *Wird eine Kurve C durch eine Abbildung $f : [a, b] \to X$ parametrisiert, so nennen wir* $\operatorname{Var}_a^b(f)$ *die* Länge *von C. Eine Kurve, bei der jedes kompakte Teilstück eine endliche Länge hat nennt man* rektifizierbar[51].

Die beschriebene Relation ist tatsächlich eine Äquivalenzrelation. Wegen Satz 9.5(i) hängt die Länge einer Kurve nicht von der Parametrisierung ab.

Nun können wir zu jeder Metrik d die zugehörige innere Metrik d^* auf dem selben metrischen Raum X definieren: Für zwei Punkte x und y ist $d^*(x, y)$ das Infimum der Längen aller Kurven, die durch Wege von x nach y in X parametrisiert werden. Eigenschaft (i) ist offensichtlich, Eigenschaft (ii), also die Dreiecksungleichung, folgt aus Satz 9.5(ii), und Eigenschaft (iii) aus der Eigenschaft $d^*(x, y) \geq d(x, y)$. Liegen x und y in verschiedenen Wegzusammenhangskomponenten, so ist $d^*(x, y) = \infty$. Im Spezialfall der Erdoberfläche erhalten wir den Abstand in Luftlinie.

Satz 9.6. *Ist V ein endlichdimensionaler Vektorraum (oder affiner Raum) und $f : [a, b] \to V$ stetig sowie auf $]a, b[$ stetig differenzierbar, so gilt*

$$\operatorname{Var}_a^b(f) = \int_a^b \|f'(t)\|\, dt,$$

wobei auf beiden Seiten die selbe Norm zu benutzen ist.

Dabei ist auf der rechten Seite das uneigentliche Integral gemeint, das notfalls als uneigentlicher Grenzwert definiert ist, also auch den Wert ∞ haben kann. Dies geschieht bei der Kurve im Beispiel am Anfang dieses Abschnitts für $0 = a < b$. Ein kompaktes Teilstück dieser Kurve ist also genau dann rektifizierbar, wenn es nicht den Ursprung enthält.

Beweis. Zunächst sei f auf ganz $[a, b]$ stetig differenzierbar. Es sei $\varepsilon > 0$. Da f' nach Satz 8.21 gleichmäßig stetig ist, gibt es ein $\delta_1 > 0$, so dass für $u, v \in [a, b]$ mit $|u - v| < \delta_1$ gilt $\|f'(u) - f'(v)\| < \frac{\varepsilon}{2(b-a)}$. Nun sei $T = \{t_0, \ldots, t_m\}$ eine Teilung von $[a, b]$ mit einer Feinheit kleiner als δ_1 und dazu $Z = \{z_1, \ldots, z_m\}$ eine Menge von Stützstellen. Nach Satz 9.2 gilt

$$f(t_k) - f(t_{k-1}) - (t_k - t_{k-1}) f'(z_k) = \int_{t_{k-1}}^{t_k} (f'(t) - f'(z_k))\, dt.$$

Mit der Dreiecksungleichung und Satz 9.1(iv) folgt

$$\begin{aligned}\big|\|f(t_k) - f(t_{k-1})\| - (t_k - t_{k-1})\|f'(z_k)\|\big| \\ \leq \|f(t_k) - f(t_{k-1}) - (t_k - t_{k-1}) f'(z_k)\| \\ \leq \int_{t_{k-1}}^{t_k} \|f'(t) - f'(z_k)\|\, dt \leq \frac{\varepsilon}{2(b-a)}(t_k - t_{k-1}).\end{aligned}$$

[51] Das lateinische Wort *rektificare* bedeutet „gerade machen“ oder „strecken“.

Durch Summation ergibt sich

$$\left|V(f,T) - S(\|f'\|,T,Z)\right| \leq \frac{\varepsilon}{2}.$$

Nach Definition des Integrals gibt es ein $\delta_2 > 0$, so dass

$$\left|S(\|f'\|,T,Z) - \int_a^b \|f'(x)\|\,dx\right| < \frac{\varepsilon}{2},$$

wenn die Feinheit von T kleiner als δ_2 ist. Ist sie kleiner als $\delta = \min\{\delta_1, \delta_2\}$, so folgt mit der Dreiecksungleichung

$$\left|V(f,T) - \int_a^b \|f'(x)\|\,dx\right| < \varepsilon.$$

Ist T' eine beliebige Teilung, so gibt es eine Verfeinerung $T \supseteq T'$ mit einer Feinheit kleiner als δ, und $V(f,T) \geq V(f,T')$. Also ist $\mathrm{Var}_a^b(f)$ gleich dem Supremum über alle Teilungen mit einer Feinheit kleiner als δ, und es folgt

$$\left|\mathrm{Var}_a^b(f) - \int_a^b \|f'(x)\|\,dx\right| \leq \varepsilon.$$

Da ε beliebig war, folgt die Behauptung.

Nun kommen wir zum allgemeinen Fall. Für $a < \alpha < \beta < b$ gilt nach Satz 9.5(ii)

$$\mathrm{Var}_\alpha^\beta f \leq \mathrm{Var}_a^b f,$$

und die linke Seite ist monoton wachsend in β und monoton fallend in α. Daraus folgt

$$\lim_{\alpha \to a+0} \lim_{\beta \to b-0} \mathrm{Var}_\alpha^\beta f \leq \mathrm{Var}_a^b f.$$

Angenommen, die linke Seite, die wir mit c bezeichnen, sei kleiner als die rechte Seite, also insbesondere verschieden von ∞. Dann gibt es eine Teilung T von $[a,b]$, so dass

$$V(f,T) > c.$$

Wir können annehmen, dass $T' = T \setminus \{a,b\}$ nicht leer ist, und setzen $\alpha = \min T'$ sowie $\beta = \max T'$. Dann gilt

$$V(f,T) = V(f,T') + \|f(\alpha) - f(a)\| + \|f(b) - f(\beta)\|.$$

Wegen der Stetigkeit von f gibt es ein $\delta > 0$, so dass die Summe der letzten beiden Terme kleiner als $V(f,T) - c$ ist, falls $\alpha - a < \delta$ und $b - \beta < \delta$ ist. Wir können annehmen, dass diese Voraussetzung erfüllt ist, weil das Bewiesene richtig bleibt, wenn wir zu T Punkte hinzufügen. Dann folgt

$$V(f,T') > c.$$

Andererseits ist wegen der Monotonie

$$\operatorname{Var}_{\alpha}^{\beta}(f) \leq c,$$

und beides zusammen widerspricht der Definition der Variation als (kleinster) oberer Schranke. Somit war unsere Annahme falsch, und aus der resultierenden Gleichheit und der Definition uneigentlicher Integrale ergibt sich die Formel im Satz durch Grenzübergang. □

Beispiel. Für die durch

$$f(t) = \exp it$$

definierte Funktion $f : \mathbb{R} \to \mathbb{C}$ erhalten wir

$$\operatorname{Var}_a^b f = b - a.$$

Somit hat der Einheitskreis den Umfang 2π, und das Bogenmaß von Winkeln lässt sich nun mathematisch korrekt behandeln, was wir bei der Einführung der Winkelfunktionen in Abschnitt 5.3 noch aufgeschoben hatten. ◁

Man beachte, dass eine Funktion von beschränkter Variation nicht stetig zu sein braucht.

Satz 9.7. *Es sei I ein Intervall. Eine Funktion $f : I \to \mathbb{R}$ ist genau dann von beschränkter Variation, wenn es eine monoton wachsende Funktion g und eine monoton fallende Funktion h auf I gibt, so dass $f = g + h$.*

Beweis. Ist g monoton wachsend und h monoton fallend, so gilt offenbar

$$\operatorname{Var}_a^b(g) = g(b) - g(a), \qquad \operatorname{Var}_a^b(h) = h(a) - h(b),$$

und nach Satz 9.5(iii) ist $g + h$ von beschränkter Variation.

Ist umgekehrt f von beschränkter Variation und $a \in I$, so ist

$$g(t) = \operatorname{Var}_a^t(f)$$

monoton wachsend, denn für $s \leq t$ in I gilt nach Satz 9.5(ii)

$$g(t) - g(s) = \operatorname{Var}_s^t(f) \geq 0.$$

Genauer gilt

$$\operatorname{Var}_s^t(f) \geq |f(t) - f(s)| \geq f(t) - f(s).$$

Setzen wir also $h = f - g$, so folgt $h(s) \geq h(t)$. □

Diese Idee ist auch für rektifizierbare Kurven C in beliebigen metrischen Räumen X von Nutzen. Ist $f : I \to X$ eine Parametrisierung und $a \in I$, so ist $g(t) = \operatorname{Var}_a^t(f)$ monoton wachsend. Gilt für zwei Zahlen u, $v \in I$ die Gleichheit $g(u) = g(v)$, so gilt auch $f(u) = f(v)$. Somit gibt es eine Abbildung $h : J \to X$, so dass $h \circ g = f$, wobei $J = g(I)$. Man nennt h eine natürliche Parametrisierung von C, weil für beliebige $s, t \in J$ gilt $\operatorname{Var}_s^t(h) = t - s$. Ist $\tilde{h} : \tilde{J}$ eine weitere natürliche Parametrisierung, so gibt es Zahlen $c \in \{1, -1\}$ und $d \in \mathbb{R}$, so dass $\tilde{h}(s) = h(cs + d)$ für alle $s \in \tilde{J}$.

10 Differentiation von Funktionen mehrerer Variablen

Vorlesung 9a
7.12.22

10.1 Definition der Ableitung

Häufig hängt eine Größe von mehreren anderen ab. Deshalb benötigt man den Begriff der Ableitung von Funktionen mehrerer Variablen. Tatsächlich gibt es verschiedene Versionen dieses Begriffs.

Es sei wieder $K = \mathbb{R}$ oder $K = \mathbb{C}$. Wir betrachten Funktionen, deren Definitionsbereich D eine Teilmenge von K^n ist. Durch Einschränkung erhält man Funktionen von einer Variablen und kann die bekannten Begriffe übertragen.

Definition 10.1. *Es sei D eine Teilmenge von K^n und $a \in D$. Die* partielle Ableitung *einer Funktion $f : D \to K^m$ nach dem j-ten Argument an der Stelle a ist die Ableitung der* partiellen Funktion *einer Variablen*

$$x_j \mapsto f(a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_n)$$

an der Stelle a_j. Wir bezeichnen sie mit $\delta_j f(a)$. Wenn sie existiert, heißt f an der Stelle a partiell differenzierbar *nach dem j-ten Argument.*

Schreiben wir $x_j = a_j + t$, so erhalten wir

$$\delta_j f(a) = \lim_{t \to 0} \frac{f(a + te_j) - f(a)}{t},$$

wobei $e_j \in K^n$ der j-te Vektor der Standardbasis ist. Damit $\delta_j f(a)$ definiert ist, muss insbesondere a ein Häufungspunkt von $\{a + te_j \in D \mid t \in K\}$ sein.

Man kann die j-te partielle Ableitung als Funktion auf der Menge aller Stellen betrachten, an denen sie existiert, und erhält eine Funktion $\delta_j f$, die traditionell als

$$\frac{\partial f}{\partial x_j}$$

bezeichnet wird, wobei das j-te Argument immer mit der selben Variablen bezeichnet werden muss.[52]

Beispiel. Es sei $K = \mathbb{R}$ und $f(x) = \sqrt{x_1^2 + \ldots + x_n^2}$. Dann ist für $x \in \mathbb{R}^n \setminus \{0\}$

$$\delta_j f(x) = \frac{1}{2}(x_1^2 + \ldots + x_n^2)^{-\frac{1}{2}} \cdot 2x_j = \frac{x_j}{f(x)}. \qquad \triangleleft$$

[52] Der Buchstabe ∂ ist übrigens ein kursives kyrillisches d.

Häufig hängt eine Größe nicht von Skalaren ab, sondern von einem Vektor oder einem Punkt. Man kann zwar in einem Vektorraum Koordinaten einführen, aber dies ist willkürlich und der Begriff der partiellen Ableitung dann unnatürlich.

Definition 10.2. *Es sei D eine Teilmenge eines K-Vektorraumes V und W ein weiterer K-Vektorraum. Die* Richtungsableitung[53] *einer Funktion $f : D \to W$ bezüglich eines Vektors $v \in V$ an der Stelle $a \in D$ ist*

$$\delta_v f(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t}.$$

Offensichtlich gilt $\delta_{tv} f(a) = t\delta_v f(a)$ für $t \in K$. Damit $\delta_v f(a)$ definiert ist, muss insbesondere a ein Häufungspunkt der Menge $\{a + tv \in D \mid t \in K\}$ sein. Im Fall $V = K^n$ sind die partiellen Ableitungen nichts Anderes als die Richtungsableitungen bezüglich der Vektoren der Standardbasis. Im Spezialfall $V = K$ ist $\delta_1 f(a) = f'(a)$, wobei man den Index 1 als Vektor in $\mathbb{R}$ interpretieren kann.

Beispiel. Es sei f wie im vorigen Beispiel und $v = (v_1, \dots, v_n)$. Dann ist $\partial_v f(x)$ die Ableitung von

$$\sqrt{(x_1 + tv_1)^2 + \ldots + (x_n + tv_n)^2}$$

als Funktion von t an der Stelle 0, also gilt für $x \neq 0$

$$\delta_v f(x) = \frac{1}{2}(x_1^2 + \ldots + x_n^2)^{-\frac{1}{2}}(2x_1v_1 + \ldots + 2x_nv_n) = \frac{\langle x, v\rangle}{f(x)}$$

mit dem Skalarprodukt $\langle\, .\, ,\, .\,\rangle$ aus Abschnitt 8.1. ◁

Wir erinnern uns, dass die Ableitung einer Funktion f einer Variablen den Anstieg der Tangente an den Graphen angibt. Die Tangente mit der Gleichung $y = f(a) + f'(a)(x - a)$ ist der Graph einer linearen Funktion. Bezeichnet man die Abweichung der Funktion f von dieser Linearisierung mit $r(x - a)$, ist also

$$f(x) = f(a) + f'(a)(x - a) + r(x - a),$$

so gilt

$$\frac{r(x-a)}{x-a} = \frac{f(x) - f(a)}{x - a} - f'(a) \to 0 \;\; (x \to a).$$

[53] für unendlichdimensionales V auch *Variation* genannt. Unsere Schreibweise $\delta_v f$ entstammt der Variationsrechnung.

Mit der Substitution $x = a + v$ werden diese Gleichungen zu

$$f(a+v) = f(a) + f'(a)v + r(v), \qquad \frac{r(v)}{v} \to 0 \ \ (v \to 0).$$

In der Leibnizschen Schreibweise

$$\frac{df}{dx} = f'(a)$$

wurden die Differentiale als unendlich kleine Größen betrachtet, und man schrieb auch

$$df = f'(a)\, dx.$$

Man kann dieser Gleichung einen korrekten Sinn geben, indem man dx als unabhängige Größe betrachtet (die man dann besser mit einer einzelnen Variablen wie v bezeichnen sollte), die reelle Werte annehmen kann und von der df proportional abhängt.

Dies kann man auf Abbildungen zwischen Vektorräumen und sogar zwischen affinen Räumen verallgemeinern. Der Begriff der Proportionalität ist dabei durch den der Linearität zu ersetzen. Im affinen Fall müsste man $x - a$ als $\overrightarrow{ax}$ schreiben. Bequemer ist die letzte Schreibweise mit einen Vektor v.

Definition 10.3. *Es seien V und W endlichdimensionale K-Vektorräume, D eine Teilmenge von V und a ein innerer Punkt von D. Eine K-lineare Abbildung $l : V \to W$ heißt* Differential *einer Funktion $f : D \to W$ an der Stelle a, wenn mit der Schreibweise*

$$f(a+v) = f(a) + l(v) + r(v)$$

gilt

$$\frac{r(v)}{\|v\|} \to 0_W \ \ (v \to 0_V).$$

*Die Funktion f heißt K-*differenzierbar *an der Stelle a, wenn ein K-lineares Differential von f an dieser Stelle existiert.*

Bemerkung. Da nach Satz 8.20(i) alle Normen auf V wie auch auf W äquivalent sind, hängt die Differenzierbarkeit nicht von der Wahl der Normen ab. Das so genannte Restglied r ist eindeutig bestimmt, denn man kann die obige Gleichung nach ihm auflösen. Es misst den Fehler bei der Näherung der Funktion $f(a+v)$ durch die Linearisierung $f(a)+l(v)$. In der linearen Algebra wird gezeigt, dass der Graph der letzteren Abbildung ein affiner Unterraum von $V \times W$ ist.

Wir konnten natürlich nicht mehr durch v dividieren, aber die neue Bedingung wird im Spezialfall $V = K$ zu $r(v)/|v| \to 0 \ (v \to 0)$, was äquivalent zur alten Bedingung $r(v)/|v| \to 0 \ (v \to 0)$ ist.

Bemerkung. Jede $\mathbb{C}$-lineare Abbildung zwischen $\mathbb{C}$-Vektorräumen ist $\mathbb{R}$-linear als Abbildung zwischen den unterliegenden $\mathbb{R}$-Vektorräumen. Somit ist jede $\mathbb{C}$-differenzierbare Abbildung auch $\mathbb{R}$-differenzierbar mit dem selben Differential.

Bemerkung. Ist f an der Stelle a differenzierbar, so ist f an dieser Stelle wegen $r(v) \to 0_W \ (v \to 0_V)$ auch stetig. Außerdem existieren dort die Richtungsableitungen bezüglich aller Vektoren $v \in V$, denn weil a ein innerer Punkt von D ist, ist $f(a+tv)$ ist für t in einer Umgebung der Null definiert, und

$$\delta_v f(a) = \lim_{t\to 0} \frac{l(tv) + r(tv)}{t} = l(v) + \lim_{t \to 0} \frac{r(tv)}{t} = l(v),$$

was für $v = 0_V$ offensichtlich ist, während für $v \neq 0_V$ nach Definition gilt

$$\left\| \frac{r(tv)}{t} \right\| = \|v\| \frac{\|r(tv)\|}{\|tv\|} \to 0 \ (t \to 0).$$

Somit ist das Differential von f an einer Stelle a eindeutig bestimmt, und wir bezeichnen es mit

$$l = df(a).$$

Aus ihm ergeben sich die Richtungsableitungen als

$$\delta_v f(a) = df(a)(v).$$

Umgekehrt folgt aus der Existenz der Richtungsableitungen und einer linearen Abbildung l mit der Eigenschaft $\delta_v f(a) = l(v)$ für alle $v \in V$ aber nicht die Differenzierbarkeit an der Stelle a!

Bemerkung. Eine Abbildung $f : D \to K^m$ ist genau dann differenzierbar an der Stelle a, wenn alle ihre Koordinatenfunktionen an dieser Stelle differenzierbar sind.

Sind nämlich die Koordinatenfunktionen f_j differenzierbar, so gibt es lineare Abbildungen $l_j : D \to K$, so dass

$$f_j(a+v) = f_j(a) + l_j(v) + r_j(v) \quad \Rightarrow \quad \frac{r_j(v)}{\|v\|} \to 0 \ (v \to 0_V).$$

Setzen wir $l(v) = (l_1(v), \ldots, l_m(v))$, so ist $r(v) = (r_1(v), \ldots, r_m(v))$, und die Existenz des Grenzwertes in der Definition folgt wie in Lemma 9.1. Die Umkehrung zeigt man analog.

Allgemeiner kann man Abbildungen in Produkträume betrachten, deren Faktoren an Stelle von K selbst Vektorräume sind. Der Einfachheit betrachten wir nur den Fall $m = 2$, also $W = W_{\mathrm{I}} \times W_{\mathrm{II}}$. Zur Nummerierung der Faktoren benutzen wir hier römische Zahlen, denn im Fall $W = K^m$ könnten sie durch Zusammenfassung mehrerer Faktoren K entstehen, so dass man beide Nummerierungen gleichzeitig betrachten muss. Mit der Bezeichnung

$$f(x) = (f_{\mathrm{I}}(x), f_{\mathrm{II}}(x)) \qquad \text{gilt dann} \qquad df(a) = (df_{\mathrm{I}}(a), df_{\mathrm{II}}(a)).$$

Bemerkung. Im Falle $V = K^n$ ist

$$l(v) = l(v_1 e_1 + \ldots + v_n e_n) = l(e_1)v_1 + \ldots + l(e_n)v_n,$$

also mit den obigen Bezeichnungen

$$df(a)(v) = \delta_1 f(a) v_1 + \ldots + \delta_n f(a) v_n.$$

Traditionell schreibt man

$$df = \frac{\partial f}{\partial x_1} dx_1 + \ldots + \frac{\partial f}{\partial x_n} dx_n,$$

wobei meist vorausgesetzt wird, dass f an jeder Stelle a von D differenzierbar ist, und die Abhängigkeit von a weggelassen wird.

Allgemeiner kann man als Definitionsbereich einen Produktraum betrachten, dessen Faktoren an Stelle von K selbst Vektorräume sind. Der Einfachheit halber betrachten wir nur den Fall $n = 2$, also $V = V_{\mathrm{I}} \times V_{\mathrm{II}}$ und darin eine Stelle $a = (a_{\mathrm{I}}, a_{\mathrm{II}})$. Dann haben wir partielle Funktionen $x_{\mathrm{I}} \mapsto f(x_{\mathrm{I}}, a_{\mathrm{II}})$ und $x_{\mathrm{II}} \mapsto f(a_{\mathrm{I}}, x_{\mathrm{II}})$. Ihre Differentiale nennt man *partielle Differentiale* von f; wir bezeichnen sie mit

$$\partial_{\mathrm{I}} f(a) : V_{\mathrm{I}} \to W, \qquad \partial_{\mathrm{II}} f(a) : V_{\mathrm{II}} \to W.$$

Für $v = (v_{\mathrm{I}}, v_{\mathrm{II}}) \in V$ gilt dann

$$df(a)(v) = \partial_{\mathrm{I}} f(a)(v_{\mathrm{I}}) + \partial_{\mathrm{II}} f(a)(v_{\mathrm{II}}),$$

und zur Unterscheidung nennt man df traditionell das *totale Differential* von f.

Bemerkung. Ist sowohl der Definitionsbereich $V = K^n$ als auch der Zielbereich $W = K^m$, so können wir die letzten Bemerkungen kombinieren und erhalten

$$\begin{aligned}
df_1(a)(v) &= \delta_1 f_1(a) v_1 + \delta_2 f_1(a) v_2 + \ldots + \delta_n f_1(a) v_n, \\
df_2(a)(v) &= \delta_1 f_2(a) v_1 + \delta_2 f_2(a) v_2 + \ldots + \delta_n f_2(a) v_n, \\
\vdots \qquad & \qquad\qquad \vdots \\
df_m(a)(v) &= \delta_1 f_m(a) v_1 + \delta_2 f_m(a) v_2 + \ldots + \delta_n f_m(a) v_n.
\end{aligned}$$

Schreiben wir die Elemente von V und W als Spaltenvektoren, so erhalten wir die übersichtliche Schreibweise

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}, \qquad df(a)v = \begin{pmatrix} \delta_1 f_1(a) & \dots & \delta_n f_1(a) \\ \vdots & & \vdots \\ \delta_1 f_m(a) & \dots & \delta_n f_m(a) \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Die $m \times n$-Matrix auf der rechten Seite nennt man die *Funktionalmatrix* oder *Jacobi-Matrix* von f an der Stelle a. Wenn man sie mit $f'(a)$ bezeichnet, so gilt die selbe Formel wie im Fall $n = 1$, nur dass die Multiplikation jetzt die von Matrizen ist. Abstrakte Vektorräume V und W muss man erst durch Wahl von Basen mit K^n bzw. K^m identifizieren, um eine Jacobi-Matrix zu erhalten, die dann natürlich von der Wahl der Basen abhängt.

Satz 10.1. *Ist $D \subseteq K^n$, W ein endlichdimensionaler K-Vektorraum und hat $f : D \to W$ in einer Umgebung eines inneren Punktes a von D bezüglich aller Argumente partielle Ableitungen, die an der Stelle a stetig sind, so ist f an dieser Stelle differenzierbar.*

Beweis. Wir müssen zeigen, dass die K-lineare Abbildung

$$l(v) = \sum_{j=1}^{n} \delta_j f(a) v_j$$

das K-lineare Differential von f an der Stelle a ist.

Zunächst sei $K = \mathbb{R}$. Durch Wahl einer Basis können wir annehmen, dass $W = \mathbb{R}^m$ ist, und nach einer obigen Bemerkung genügt es, den Fall $W = \mathbb{R}$ zu betrachten. Laut Definition einer Umgebung gibt es ein $\eta_0 > 0$, so dass die partiellen Ableitungen von f auf der η_0-Umgebung von a definiert sind. Für $v \in \mathbb{R}^n$ mit $\|v\| < \eta_0$ liegen die Punkte

$$a_0 = a, \quad a_1 = a_0 + v_1 e_1, \quad a_2 = a_1 + v_2 e_2, \quad \dots, \quad a_n = a_{n-1} + v_n e_n$$

in $U_{\eta_0}(a)$, ebenso die Verbindungsstrecken $[a_{j-1}, a_j]$. Wenden wir den Mittelwertsatz der Differentialrechnung auf $f(a_{j-1}+te_j)$ als Funktion von $t \in [0, v_j]$ an, so erhalten wir ein $b_j \in [a_{j-1}, a_j]$, so dass

$$f(a_j) - f(a_{j-1}) = \delta_j f(b_j) v_j.$$

Durch Summation ergibt sich

$$f(a + v) - f(a) = \sum_{j=1}^{n} \delta_j f(b_j) v_j, \tag{14}$$

wobei die Punkte b_j von v abhängen, aber immer gilt $\|b_j - a\| < \|v\|$. Wegen der Stetigkeit der partiellen Ableitungen gibt es für jedes $\varepsilon > 0$ ein $\eta \in \,]0, \eta_0[$, so dass für $\|x - a\| < \eta$ gilt

$$|\delta_j f(x) - \delta_j f(a)| < \varepsilon.$$

Ist $\|v\| < \eta$, so gilt dies insbesondere für $x = b_j$, also ist nach der Dreiecksungleichung

$$|f(a+v) - f(a) - l(v)| \leq \sum_{j=1}^{n} \bigl|\bigl(\delta_j f(b_j) - \delta_j f(a)\bigr)v_j\bigr| < \varepsilon\|v\|_1.$$

Wir können annehmen, dass $\|\,.\,\| = \|\,.\,\|_1$. Da ε beliebig war, folgt

$$\frac{|f(a+v) - f(a) - l(v)|}{\|v\|} \to 0 \;\; (v \to 0_V).$$

Im Fall $K = \mathbb{C}$ ist jede reelle Richtungsableitung gleich der komplexen Richtungsableitung bezüglich des selben Vektors, und es gilt $\delta_{iv} f = i\delta_v f$. Darum ergibt sich die selbe Abbildung l, wenn wir statt der Standardbasis von $\mathbb{C}^n$ die reelle Basis $(e_1, ie_1, \ldots, e_n, ie_n)$ des unterliegenden reellen Vektorraums benutzen. Sind die komplexen partiellen Ableitungen an der Stelle a stetig, so gilt das auch für die reellen. Somit folgt die Behauptung aus dem Bewiesenen. □

Definition 10.4. *Es seien V und W Vektorräme über K und U eine offene Teilmenge von V. Eine Abbildung $f : U \to W$ heißt K-differenzierbar, wenn sie es an jeder Stelle von U ist, und sie heißt stetig K-differenzierbar, wenn sie differenzierbar ist und die Funktion $df : U \to \mathrm{Hom}(V, W)$ stetig ist.*

Aus Satz 10.1 und dem vorher erwähnten Zusammenhang zwischen totalem Differential und partiellen Ableitungen erhalten wir:

Folgerung 10.1. *Sind alle partiellen Ableitungen von f auf der offenen Teilmenge $U \subseteq \mathbb{R}^n$ stetig, so ist f auf U stetig differenzierbar.*

Beispiel. Es seien U, V und W endlichdimensionale Vektorräume sowie

$$f : U \times V \to W$$

eine bilineare Abbildung. Wir behaupten, dass diese Abbildung stetig differenzierbar ist und ihr Differential an einer Stelle $(a, b) \in U \times V$ gegeben ist durch

$$df(a,b)(u,v) = f(u,b) + f(a,v).$$

Im Beweis von Satz 9.1(ii) hatten wir gesehen, dass ein $c > 0$ existiert, so dass $\|f(u,v)\| \leq c\|u\|\|v\|$. Verwenden wir auf $U \times V$ die Norm $\|(u,v)\| = \sqrt{\|u\|^2 + \|v\|^2}$, so gilt nach der Ungleichung zwischen dem arithmetischen und dem geometrischen Mittel der Zahlen $\|u\|^2$ und $\|v\|^2$, dass

$$\|u\|\|v\| \leq \frac{1}{2}\|(u,v)\|^2,$$

und die Bedingung

$$\frac{r(u,v)}{\|(u,v)\|} \to 0_W \quad \big((u,v) \to (0_U, 0_V)\big)$$

folgt wegen $r(u,v) = f(u,v)$. $\triangleleft$

Unter einem Operator versteht man eine Abbildung, deren Definitions- und Zielbereich Mengen von Funktionen sind. In der Physik sind folgende Differentialoperatoren von Bedeutung.

Vorlesung 9b
9.12.22

Beispiel. Ist auf einem reellen Vektorraum V ein Skalarprodukt gegeben und die skalarwertige Funktion f auf $D \subseteq V$ an der Stelle $a \in \mathring{D}$ differenzierbar, so ist der *Gradient* $\operatorname{grad} f(a) \in V$ charakterisiert durch

$$\langle \operatorname{grad} f(a), v \rangle = df(a)(v)$$

für alle $v \in V$. Im Fall $V = \mathbb{R}^n$ mit dem Standardskalarprodukt erhalten wir

$$\operatorname{grad} f(a) = \begin{pmatrix} \delta_1 f(a) \\ \vdots \\ \delta_n f(a) \end{pmatrix}.$$

Mit Hilfe des *Nabla-Operators*

$$\nabla = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_n \end{pmatrix}$$

schreibt man dies symbolisch in der Form $\operatorname{grad} f = \nabla f$. $\triangleleft$

Beispiel. Ist X ein Vektorfeld auf $D \subseteq \mathbb{R}^n$, also $X : D \to \mathbb{R}^n$, das an der Stelle $a \in \mathring{D}$ differenzierbar ist, so definiert man die *Divergenz* $\operatorname{div} X(a) \in \mathbb{R}$ durch

$$\operatorname{div} X(a) = \delta_1 X_1(a) + \ldots + \delta_n X_n(a).$$

Dies drückt man auch symbolisch durch $\operatorname{div} X = \langle \nabla, X \rangle$ aus. $\triangleleft$

Beispiel. Ist ein Vektorfeld X auf $D \subseteq \mathbb{R}^3$ an der Stelle $a \in \mathring{D}$ differenzierbar, so definiert man die *Rotation* $\operatorname{rot} X(a) \in \mathbb{R}^3$ durch

$$\operatorname{rot} X(a) = \begin{pmatrix} \delta_2 X_3(a) - \delta_3 X_2(a) \\ \delta_3 X_1(a) - \delta_1 X_3(a) \\ \delta_1 X_2(a) - \delta_2 X_1(a) \end{pmatrix}.$$

Dies drückt man symbolisch durch $\operatorname{rot} X = \nabla \times X$ aus, wobei $\times$ das *Vektorprodukt* bezeichnet. $\triangleleft$

Für den Beweis der Ableitungsregeln ist eine Vorbetrachtung nötig. Bezeichnen wir die rechte Seite der Gleichung (14) im Beweis von Satz 10.1 mit $\tilde{f}(a+v)(v)$, so gilt

$$f(x) - f(a) = \tilde{f}(x)(x-a),$$

und im Fall $n = 1$ ist $\tilde{f}(x)$ nichts anderes als der Differenzenquotient. Ein solcher verallgemeinerter Differenzenquotient existiert immer, ist aber im Allgemeinen nicht eindeutig bestimmt.

Lemma 10.1 (Carathéodory)**.** *Es seien V und W endlichdimensionale Vektorräume und a ein innerer Punkt einer Teilmenge $D \subseteq V$. Eine Abbildung $f : D \to W$ ist genau dann an der Stelle a differenzierbar, wenn eine Abbildung*

$$\tilde{f} : D \to \operatorname{Hom}(V, W)$$

existiert, die an der Stelle a stetig ist, so dass

$$f(x) = f(a) + \tilde{f}(x)(x-a).$$

Beweis. Angenommen, $\tilde{f}$ existiert. Setzen wir $l = \tilde{f}(a)$, so gilt

$$r(v) = \big(\tilde{f}(a+v) - \tilde{f}(a)\big)v, \qquad \|r(v)\| \le \big\|\tilde{f}(a+v) - \tilde{f}(a)\big\|\|v\|$$

im Sinne von Definition 8.20, und aus der Stetigkeit von $\tilde{f}$ an der Stelle a folgt

$$\frac{\|r(v)\|}{\|v\|} \to 0 \qquad (v \to 0_V),$$

also ist f an der Stelle a differenzierbar.

Nun sei umgekehrt f an der Stelle a differenzierbar mit Ableitung l. Wir können annehmen, dass $\|v\|^2 = \langle v, v\rangle$ für ein Skalarprodukt auf V. Im Fall $K = \mathbb{C}$ ist dieses übrigens nur bezüglich des ersten Arguments linear, aber bezüglich des zweiten Arguments semilinear, d. h. $\langle u, av\rangle = \bar{a}\langle u, v\rangle$, und es gilt $\langle v, u\rangle = \overline{\langle u, v\rangle}$. Für u, $v \in V$ mit der Eigenschaft $a + v \in D$ setzen wir

$$\tilde{f}(a+v)(u) = \begin{cases} l(u) + \dfrac{\langle u, v\rangle}{\|v\|^2} r(v), & \text{wenn } v \neq 0_V, \\ l(u), & \text{wenn } v = 0_V. \end{cases}$$

Dann ist $\tilde{f}(a+v) \in \operatorname{Hom}(V, W)$ und

$$\tilde{f}(a+v)(v) = l(v) + r(v) = f(a+v) - f(a).$$

Außerdem gilt für $v \neq 0_V$

$$(\tilde{f}(a+v) - \tilde{f}(a))(u) = \frac{\langle u, v\rangle}{\|v\|^2} r(v),$$

also nach der Cauchy-Schwarz-Ungleichung

$$\|(\tilde{f}(a+v) - \tilde{f}(a))(u)\| \leq \frac{\|r(v)\|}{\|v\|}\|u\|.$$

Laut Definition der Norm einer linearen Abbildung folgt

$$\|\tilde{f}(a+v) - \tilde{f}(a)\| \leq \frac{\|r(v)\|}{\|v\|},$$

also ist $\tilde{f}$ an der Stelle a stetig. □

Nun verallgemeinern wir die Sätze 6.1 und 6.2 samt ihren Beweisen.

Satz 10.2. *Es seien U, V und W Vektorräume über K.*

(i) Ist $D \subseteq V$ und sind f, $g : D \to W$ an der Stelle $a \in \mathring{D}$ differenzierbar, so ist auch $f + g$ an der Stelle a differenzierbar, und es gilt die Summenregel

$$d(f+g)(a) = df(a) + dg(a).$$

(ii) Es sei $D \subseteq U$, $E \subseteq V$. Ist $f : D \to E$ an der Stelle $a \in \mathring{D}$ und $g : E \to W$ an der Stelle $b = f(a) \in \mathring{E}$ differenzierbar, so ist $g \circ f$ an der Stelle a differenzierbar, und es gilt die Kettenregel

$$d(g \circ f)(a) = dg(b) \circ df(a).$$

Beweis. **(i)** Nach Lemma 10.1 existieren Abbildungen $\tilde{f}, \tilde{g} : D \to \operatorname{Hom}(V, W)$, die an der Stelle a stetig sind, so dass für $v \in V$ mit der Eigenschaft $a+v \in D$ gilt

$$f(a+v) = f(a) + \tilde{f}(a+v)(v), \qquad g(a+v) = g(a) + \tilde{g}(a+v)(v),$$

also

$$f(a+v) + g(a+v) = f(a) + g(a) + \big(\tilde{f}(a+v) + \tilde{g}(a+v)\big)(v).$$

Eine Abbildung $D \to \mathrm{Hom}(V,W) \times \mathrm{Hom}(V,W)$ ist genau dann stetig, wenn ihre Komponenten $D \to \mathrm{Hom}(V,W)$ stetig sind, und nach Satz 8.8 ist die Addition $\mathrm{Hom}(V,W) \times \mathrm{Hom}(V,W) \to \mathrm{Hom}(V,W)$ Lipschitz-stetig. Die Abbildung $\widetilde{f+g} = \tilde{f} + \tilde{g}$ ist also an der Stelle a stetig, und sie hat dort den Wert $df(a) + dg(a)$. Nun folgt die Behauptung nach Lemma 10.1.

(ii) Diesmal ist $\tilde{f} : D \to \mathrm{Hom}(U,V)$, $\tilde{g} : E \to \mathrm{Hom}(V,W)$, und

$$g(f(a+u)) = g\big(f(a) + \tilde{f}(a+u)(u)\big) = g(f(a)) + \widetilde{g \circ f}(a+u)(u),$$

wobei

$$\widetilde{g \circ f}(a+u) = \tilde{g}\big(f(a) + \tilde{f}(a+u)(u)\big) \circ \tilde{f}(a+u).$$

Nach Satz 8.8 ist die Verkettung $\mathrm{Hom}(V,W) \times \mathrm{Hom}(U,V) \to \mathrm{Hom}(U,W)$ eine stetige Abbildung. Außerdem ist $v \mapsto \tilde{g}(f(a) + \tilde{f}(a+u)(u))$ nach Satz 8.11 an der Stelle 0_V stetig. Es folgt, dass $\widetilde{g \circ f}$ an der Stelle a stetig ist und den Wert $\tilde{g}(f(a)) \circ \tilde{f}(a)$ hat. Die Behauptung folgt wieder mit Lemma 10.1. $\square$

Dass im Satz keine Produktregel auftaucht, ist kein Versehen, denn sie folgt aus der Kettenregel und der Formel für die Ableitung einer bilinearen Abbildung $g : V_{\mathrm{I}} \times V_{\mathrm{II}} \to W$. Sind $f_{\mathrm{I}} : U \to V_{\mathrm{I}}$ und $f_{\mathrm{II}} : U \to V_{\mathrm{II}}$ an einer Stelle a differenzierbar, so gilt dies auch für die durch $f(x) = (f_{\mathrm{I}}(x), f_{\mathrm{II}}(x))$ definierte Abbildung $f : U \to V_{\mathrm{I}} \times V_{\mathrm{II}}$. Nun folgt

$$d(g \circ f)(a)(u) = g(df_{\mathrm{I}}(a)(u), f_{\mathrm{II}}(a)) + g(f_{\mathrm{I}}(a), df_{\mathrm{II}}(a)(u)).$$

Wir benötigen auch keine neue Quotientenregel, weil $1/f$ nur für skalarwertige Funktionen f definiert ist und als Verkettung von f mit der Kehrwertfunktion angesehen werden kann.

Folgerung 10.2. *Für eine offene Menge D von K^n ist jede Abbildung $D \to K^m$ differenzierbar, deren Koordinatenfunktionen auf D durch Terme gegeben sind, in denen nur arithmetische Operationen und differenzierbare Funktionen vorkommen.*

Die Differentialrechnung entstand aus den Bedürfnissen der Newtonschen Mechanik. Diese findet allerdings nicht in einem Vektorraum, sondern einem affinen Raum Statt.

Definition 10.5. *Sind X und Y affine Räume mit den zugehörigen Vektorräumen von Translationen V bzw. W, so heißt $g : X \to Y$* affine Abbildung, *wenn es eine lineare Abbildung $l : V \to W$ gibt, so dass für alle $a \in X$ und $v \in V$ gilt*

$$g(a+v) = g(a) + l(v).$$

In dieser Situation werden durch Normen auf V und W Metriken d bzw. e auf X und Y definiert. Ist $D \subseteq X$, so kann man die Richtungsableitung einer Funktion $f : D \to Y$ bezüglich $v \in V$ an einer Stelle $a \in \mathring{D}$ als Element von W definieren:

$$\delta_v f(a) = \lim_{t \to 0} \frac{1}{t} \overrightarrow{f(a)f(a+tv)}.$$

Das Differential $df(a) \in \operatorname{Hom}(V, W)$ ist charakterisiert durch

$$\overrightarrow{f(a)f(a+v)} = df(a)(v) + r(v), \quad \text{wobei} \quad \frac{r(v)}{\|v\|} \to 0_W \ \ (v \to 0_V).$$

Betrachtet man die durch

$$g(a+v) = f(a) + df(a)(v)$$

gegebene affine Abbildung $g : D \to W$, so kann man dies auch durch

$$\frac{e(f(x), g(x))}{d(x,a)} \to 0 \qquad (x \to a)$$

ausdrücken, d. h. f wird in der Umgebung von a durch die affine Abbildung g angenähert. In den obigen Bezeichnungen gilt übrigens $dg(a) = l$ für jedes $a \in X$.

10.2 Höhere Differentiale

Vorlesung 10a
14.12.22

Es seien V und W endlichdimensionale Vektorräume über $K \in \{\mathbb{R}, \mathbb{C}\}$. Wir betrachten wieder Abbildungen f von Teilmengen D von V in den Vektorraum W. Dabei soll die Bezeichnung f nicht nur die Zuordnung, sondern auch den Definitionsbereich festlegen. Das Differential einer beliebigen Funktion $f : D \to W$ ist eine Funktion $df : D_1 \to \operatorname{Hom}(V, W)$, wobei D_1 die Menge der inneren Punkte von D ist, in denen f differenzierbar ist, was auch die leere Menge sein kann. Insbesondere ist $D_1 \subseteq D$. Bezeichnet $\mathcal{F}(V, W)$ die Menge der Funktionen auf Teilmengen von V mit Werten in W, so definiert das Differential eine Abbildung

$$d : \mathcal{F}(V, W) \to \mathcal{F}(V, \operatorname{Hom}(V, W)).$$

Diese können wir auch mehrmals anwenden. So erhalten wir z. B. das zweite Differential

$$d^2 f : D_2 \to \operatorname{Hom}(V, \operatorname{Hom}(V, W))$$

In Analogie zu Definition 6.7 legen wir fest:

Definition 10.6. *Wir definieren das k-te* Differential *der Abbildung f rekursiv durch*

$$d^0 f = f, \qquad d^{k+1} f = d(d^k f).$$

Eine Abbildung f heißt k-mal differenzierbar *an der Stelle a, wenn a zum Definitionsbereich D_k von $d^k f$ gehört.*

Eine Abbildung heißt unendlich oft[54] differenzierbar *an der Stelle a, wenn sie für jede natürliche Zahl k an dieser Stelle k-mal differenzierbar ist.*

Eine Abbildung heißt k-mal differenzierbar*, wenn dies an jeder Stelle ihres Definitionsbereichs der Fall ist. Sie heißt* k-mal stetig differenzierbar*, wenn außerdem ihre k-te Ableitung stetig ist.*

Man beachte, dass im letzteren Fall ihre Ableitungen kleinerer Ordnung als k differenzierbar, also ebenfalls stetig sind. Im Allgemeinen haben die Ableitungen verschiedener Ordnung Definitionsbereiche

$$D = D_0 \supseteq D_1 \supseteq D_2 \supseteq \ldots$$

Wegen der Assoziativität der Verkettung gilt für beliebige natürliche Zahlen k und l

$$d^{k+l}f = d^l(d^k f),$$

und eine Abbildung ist an einer Stelle genau dann $k+l$-mal differenzierbar, wenn sie an dieser Stelle k-mal differenzierbar ist und ihr k-tes Differential an dieser Stelle l-mal differenzierbar ist.

Statt $d^2f(a)(u)(v)$ schreiben wir $d^2f(a)(u,v)$, was bilinear von u, $v \in V$ abhängt. Anstelle von $df(x)$ kann man auch die Abbildung $x \mapsto df(x)(v) = \delta_v f(x)$ für festes v in Richtung eines Vektors u ableiten und erhält nach der Produktregel $d(df)(a)(u)(v)$. Somit ergibt sich

$$d^2f(a)(u,v) = \delta_u \delta_v f(a).$$

Analog ist das dritte Differential $d^3f(a)$ eine *trilineare Abbildung* $V^3 \to W$ usw. Allgemein nennen wir eine Abbildung $V^k \to W$ *multilinear*, wenn die partiellen Funktionen, die beim Festhalten aller Argumente bis auf eines entstehen, linear sind. Die multilinearen Abbildungen $V^k \to W$ bilden einen endlichdimensionalen Vektorraum, den wir einmal mit $\mathrm{Mult}^k(V,W)$ bezeichnen wollen, auch wenn das nicht allgemein üblich ist. Außerdem setzten wir $\mathrm{Mult}^0(V,W) = W$.

Wir betrachten $d^kf(a)$ als Element von $\mathrm{Mult}^k(V,W)$. Nach der rekursiven Definition ist dann $d^{k+1}f(a) \in \mathrm{Mult}^k(V, \mathrm{Hom}(V,W))$, was wir aber mit $\mathrm{Mult}^{k+1}(V,W)$ identifizieren können. Nun ergibt sich durch vollständige Induktion

$$d^kf(a)(v_1, \ldots, v_k) = \delta_{v_1} \cdots \delta_{v_k} f(a).$$

Ist $V = K^n$, so gilt für $u = (u_1, \ldots, u_n)$ und $v = (v_1, \ldots, v_n)$

$$d^2f(a)(u,v) = \sum_{i,j=1}^{n} \delta_i \delta_j f(a) u_i v_j.$$

[54]richtiger wäre „beliebig oft“

Allgemeiner erhalten wir für Vektoren $v_j = (v_{j,1}, \ldots, v_{j,n}) \in K^n$

$$d^k f(a)(v_1, \ldots, v_k) = \sum_{i_1, \ldots, i_k = 1}^{n} \delta_{i_1} \ldots \delta_{i_k} f(a) v_{1,i_1} \cdots v_{k,i_k}. \tag{15}$$

In traditioneller Schreibweise wäre das für $n = k = 2$ und $v_1 = v_2$

$$d^2 f = \frac{\partial^2 f}{\partial x_1^2}\, dx_1^2 + \frac{\partial^2 f}{\partial x_1 \partial x_1}\, dx_1\, dx_2 + \frac{\partial^2 f}{\partial x_2 \partial x_1}\, dx_2\, dx_1 + \frac{\partial^2 f}{\partial x_2^2}\, dx_2^2.$$

Ist $W = K^m$, so ist eine Abbildung $f : D \to W$ genau dann k-mal an der Stelle a differenzierbar, wenn dies für ihre Koordinatenfunktionen f_j gilt, und dann ist

$$d^k f(a) = \big(d^k f_1(a), \ldots, d^k f_m(a)\big) : V^k \to K^m.$$

Eine analoge Aussage gilt für Zerlegungen $W = W_{\mathrm{I}} \times W_{\mathrm{II}}$.

Beispiel. Jede bilineare Abbildung $b : V_{\mathrm{I}} \times V_{\mathrm{II}} \to W$ zwischen endlichdimensionalen Vektorräumen ist unendlich oft differenzierbar, und es gilt für a, u, $v \in V = V_{\mathrm{I}} \times V_{\mathrm{II}}$

$$\begin{aligned} db(a)(u) &= b(u_{\mathrm{I}}, a_{\mathrm{II}}) + b(a_{\mathrm{I}}, u_{\mathrm{II}}), \\ d^2 b(a) &= b(u_{\mathrm{I}}, v_{\mathrm{II}}) + b(v_{\mathrm{I}}, u_{\mathrm{II}}). \end{aligned}$$

Da $d^2 b(a)$ nicht von a abhängt, verschwindet $d^k b$ für $k > 2$. $\triangleleft$

Satz 6.9 und Satz 6.10(i) übertragen sich problemlos.

Satz 10.3. *Es seien U, V und W endlichdimensionale Vektorräume über K.*

(i) *Ist $D \subseteq V$ und sind f, $g : D \to W$ an der Stelle $a \in \mathring{D}$ beide k-mal differenzierbar, so auch $f + g$, und es gilt die* Summenregel

$$d^k(f + g)(a) = d^k f(a) + d^k g(a).$$

(ii) *Sind f und g beide k-mal stetig differenzierbar, so auch $f + g$.*

(iii) *Es sei $D \subseteq U$ und $E \subseteq V$. Ist $f : D \to E$ an der Stelle $a \in \mathring{D}$ und $g : E \to W$ an der Stelle $f(a) \in \mathring{E}$ jeweils k-mal differenzierbar, so ist $g \circ f$ an der Stelle a ebenfalls k-mal differenzierbar.*

(iv) *Sind f und g beide k-mal stetig differenzierbar, so auch $f \circ g$.*

Die Aussagen (iii) und (iv) kann man als qualitative *Kettenregel* auffassen. Aus ihr folgt wieder die Produktregel, für die man sogar eine Formel finden kann.

Beweis. Wir übergehen den einfachen Beweis der Aussagen (i) und (ii) und kommen gleich zu (iii). Im Fall $k = 0$ ist nichts zu zeigen. Angenommen, die Behauptung gilt bereits für eine Zahl k. Die Funktionen f und g seien nun $k+1$-mal an den Stellen a bzw. $f(a)$ differenzierbar. Dann sind sie in Umgebungen dieser Stellen differenzierbar, und laut Satz 10.2 gilt in einer Umgebung von a

$$d(g \circ f) = b \circ (dg \circ f, df),$$

wobei $b : \mathrm{Hom}(V, W) \times \mathrm{Hom}(U, V) \to \mathrm{Hom}(U, W)$ die Verkettung ist. Außerdem sind df und dg bereits k-mal an den Stellen a bzw. $f(a)$ differenzierbar, und f ist nach einer obigen Bemerkung k-mal an der Stelle a differenzierbar. Nach Induktionsvoraussetzung gilt dies auch für $dg \circ f$ und somit für $(dg \circ f, df)$. Da die bilineare Abbildung b nach obigem Beispiel k-mal differenzierbar ist, gilt das nach Induktionsvoraussetzung auch für die linke Seite an der Stelle a.

Für $k = 0$ folgt Aussage (iv) aus Satz 8.11. Angenommen, sie gilt für eine Zahl k, und f und g seien nun $k+1$-mal stetig differenzierbar. Dann sind df, dg, f und b alle k-mal stetig differenzierbar, und das Gleiche gilt nach obiger Formel und der Induktionsvoraussetzung für $d(g \circ f)$. □

Satz 10.4 (Schwarz). *Es seien V und W endlichdimensionale K-Vektorräume, $D \subseteq V$ und $f : D \to W$. Angenommen, für zwei Vektoren u, $v \in V$ existieren die zweifachen Richtungsableitungen*

$$\delta_u \delta_v f, \quad \delta_v \delta_u f$$

in einer Umgebung von a und sind an der Stelle a stetig. Dann gilt

$$\delta_u \delta_v f(a) = \delta_v \delta_u f(a).$$

Beweis. Da reelle und komplexe Richtungsableitung übereinstimmen, können wir annehmen, dass $K = \mathbb{R}$ ist. Indem wir eine Basis von W wählen und die Komponenten von f einzeln betrachten, können wir annehmen, dass $W = \mathbb{R}$ ist. Setzen wir

$$g(x_1, x_2) = f(a + x_1 u + x_2 v),$$

dann gibt ein $\eta > 0$, so dass

$$\delta_1 \delta_2 g(x_1, x_2) = \delta_u \delta_v f(a + x_1 u + x_2 v), \quad \delta_2 \delta_1 g(x_1, x_2) = \delta_v \delta_u f(a + x_1 u + x_2 v)$$

in der η-Umgebung des Nullpunktes bezüglich $\| \, . \, \|_\infty$ existieren und im Nullpunkt stetig sind. Es seien x_1, $x_2 \in \,]0, \eta[$. Dann ist die Funktion

$$h(t) = g(t, x_2) - g(t, 0)$$

für $t \in [0, x_1]$ definiert. Wenden wir auf sie den Mittelwertsatz der Differentialrechnung an, so erhalten wir ein $b_1 \in \,]0, x_1[$, so dass $h(x_1) - h(0) = h'(b_1)x_1$, d. h.

$$g(x_1, x_2) - g(x_1, 0) - g(0, x_2) + g(0, 0) = \big(\delta_1 g(b_1, x_2) - \delta_1 g(b_1, 0)\big)x_1.$$

Dabei hängt b_1 von $x = (x_1, x_2)$ ab. Wenden wir den Mittelwertsatz auf die Funktion $\delta_1 g(b_1, t)$ an, so erhalten wir ein $b_2 \in \,]0, x_2[$, so dass

$$\delta_1 g(b_1, x_2) - \delta_1 g(b_1, 0) = \delta_2 \delta_1 g(b_1, b_2)x_2,$$

wobei b_2 von b_1 und x_2, also letztlich von x abhängt. Wir erhalten also für jedes x wie oben ein $b \in V$, so dass $\|b\| < \|x\|$ und

$$g(x_1, x_2) - g(x_1, 0) - g(0, x_2) + g(0, 0) = \delta_2 \delta_1 g(b_1, b_2)x_1 x_2.$$

Vertauschen wir die Rollen von x_1 und x_2, so erhalten wir für jedes x ein $c \in V$, so dass $\|c\| < \|x\|$ und

$$g(x_1, x_2) - g(0, x_2) - g(x_1, 0) + g(0, 0) = \delta_1 \delta_2 g(c_1, c_2)x_1 x_2.$$

Wegen $x_1 \neq 0$ und $x_2 \neq 0$ folgt

$$\delta_2 \delta_1 g(b) = \delta_1 \delta_2 g(c),$$

wobei b und c von x abhängen und nach dem Einschließungskriterium gilt

$$b \to 0, \quad c \to 0 \quad (x \to 0).$$

Wegen der Stetigkeit der partiellen Ableitungen folgt

$$\delta_2 \delta_1 g(0) = \delta_1 \delta_2 g(0). \qquad \square$$

Folgerung 10.3. *Ist $f : D \to W$ eine k-mal stetig differenzierbare Abbildung, so gilt für jede Permutation σ von $\{1, \ldots, k\}$ und alle Vektoren $v_1, \ldots, v_k \in V$*

$$\delta_{v_1} \cdots \delta_{v_k} f = \delta_{v_{\sigma(1)}} \cdots \delta_{v_{\sigma(k)}} f,$$

d. h. $d^k f(a)$ ist eine symmetrische Multilinearform $V^k \to W$.

Für $1 \leq l < k$ ist nämlich $\delta_{v_{l+2}} \cdots \delta_{v_k} f$ noch $l + 1 \geq 2$-mal stetig differenzierbar, und wir können δ_{v_l} und $\delta_{v_{l+1}}$ vertauschen. Jede Permutation σ lässt sich aus Transpositionen benachbarter Elemente zusammensetzen.

Beispiel. Ist f zweimal stetig differenzierbar auf einer offenen Teilmenge D von $\mathbb{R}^3$, so gilt

$$\operatorname{rot}\operatorname{grad} f = \begin{pmatrix} \delta_2\delta_3 f - \delta_3\delta_2 f \\ \delta_3\delta_1 f - \delta_1\delta_3 f \\ \delta_1\delta_2 f - \delta_2\delta_1 f \end{pmatrix} = 0.$$

Ist $X : D \to \mathbb{R}^3$ zweimal stetig differenzierbar, so gilt

$$\operatorname{div}\operatorname{rot} X = \delta_1(\delta_2 X_3 - \delta_2 X_2) + \delta_2(\delta_3 X_1 - \delta_1 X_3) + \delta_3(\delta_1 X_2 - \delta_2 X_1) = 0.$$

Ein Gradientenfeld ist also rotationsfrei und die Rotation eines Vektorfeldes ist divergenzfrei. $\triangleleft$

Man definiert den *Laplace-Operator* auf einer offenen Teilmenge D von $\mathbb{R}^n$ für zweimal stetig differenzierbares $f : D \to \mathbb{R}$ durch

$$\Delta f = \operatorname{div}\operatorname{grad} f,$$

symbolisch $\Delta = \langle \nabla, \nabla \rangle$. Dann gilt

$$\Delta f = \delta_1^2 f + \ldots + \delta_n^2 f.$$

Aus Satz 10.1 ergibt sich folgendes Kriterium.

Folgerung 10.4. *Existieren alle höheren partiellen Ableitungen einer Funktion $f : D \to W$ bis zur Ordnung k in einer Umgebung eines inneren Punktes a der Teilmenge D von K^n und sind sie an der Stelle a stetig, so ist f an dieser Stelle k-mal stetig differenzierbar.*

Zum Abschluss führen wir noch ein Analogon des Begriffs des Homöomorphismus ein.

Definition 10.7. *Eine bijektive Abbildung f zwischen offenen Teilmengen von endlichdimensionalen Vektorräumen heißt* Diffeomorphismus *der Klasse C^k, wenn die Abbildung f und ihre Umkehrabbildung k-mal stetig differenzierbar sind.*

Aufgabe 47 liefert ein Beispiel für einen Diffeomorphismus der Klasse C^∞.

10.3 Die Taylorsche Formel

Vorlesung 10b
16.12.22

Wir wollen Satz 6.13 zunächst auf vektorwertige Funktionen von einer Variablen verallgemeinern. Da der Mittelwert der Differentialrechnung nur für $\mathbb{R}$-wertige Funktionen gilt, benutzen wir statt dessen den Hauptsatz der Infinitesimalrechnung, der uns ein Restglied in Integralform liefert. Dabei muss die letzte vorkommende Ableitung allerdings stetig sein. Dies dient der Vorbereitung auf den Fall von Funktionen f mehrerer Variablen.

Wir betrachten also eine Funktion g auf einem Intervall I mit Werten in einem endlichdimensionalen reellen Vektorraum W. Ist g stetig differenzierbar, so gilt nach Satz 9.2

$$g(x) = g(a) + \int_a^x g'(t)\, dt.$$

Ist g zweimal stetig differenzierbar, so können wir partiell integrieren, wenn wir einen Faktor 1 einfügen. Seine Stammfunktion hat die Form $t + C$, und wir erhalten

$$\int_a^x g'(t)\, dt = (t + C)g'(t)\Big|_a^x - \int_a^x (t + C)g''(t)\, dt.$$

Die geschickteste Wahl ist $C = -x$, denn dann erhalten wir durch Einsetzen

$$g(x) = g(a) + g'(a)(x - a) + \int_a^x (x - t)g''(t)\, dt,$$

also die Formel aus der Definition 10.3 mit dem Restglied $r(x - a)$ in Integralform. Ist g dreimal stetig integrierbar, so können wir ein weiteres Mal partiell integrieren und erhalten

$$g(x) = g(a) + g'(a)(x - a) + \frac{(x - a)^2}{2} g'(a) + \int_a^x \frac{(x - t)^2}{2} g'''(t)\, dt.$$

Nun ist die folgende Aussage keine Überraschung.

Lemma 10.2. *Ist g eine $k + 1$-mal stetig differenzierbare Funktion auf einem Intervall I mit Werten in einem endlichdimensionalen reellen Vektorraum W, dann gilt für a, $x \in I$*

$$g(x) = \sum_{j=0}^{k} \frac{(x - a)^j}{j!} g^{(j)}(a) + \int_a^x \frac{(x - t)^k}{k!} g^{(k+1)}(t)\, dt.$$

Beweis durch vollständige Induktion. Die Fälle $k \in \{0, 1, 2\}$ haben wir bereits erledigt. Angenommen, die Behauptung gilt für eine Zahl $k - 1$, und g

sei $k+1$-mal stetig differenzierbar. Dann ist g auch k-mal stetig differenzierbar, und

$$g(x) = \sum_{j=0}^{k-1} \frac{(x-a)^j}{j!} g^{(j)}(a) + \int\limits_a^x \frac{(x-t)^{k-1}}{(k-1)!} g^{(k)}(t)\,dt.$$

Da $g^{(k)}$ stetig differenzierbar ist, finden wir durch partielle Integration

$$\int\limits_a^x \frac{(x-t)^{k-1}}{(k-1)!} g^{(k)}(t)\,dt = -\frac{(x-t)^k}{k!} g^{(k)}(t)\Big|_a^x + \int\limits_a^x \frac{(x-t)^k}{k!} g^{(k+1)}(t)\,dt.$$

Der Term außerhalb des Integrals verschwindet für $t = x$ und liefert für $t = a$ den k-ten Summanden der behaupteten Formel. □

Indem man Satz 7.7 auf das Restglied in Integralform anwendet, erhält man die verschiedenen Formen des Restgliedes in den Sätzen 6.13 und 6.16.

Nun betrachten wir vektorwertige Funktionen mehrerer Variablen, oder besser gesagt, Abbildungen zwischen Vektorräumen.

Satz 10.5. *Es seien V und W endlichdimensionale reelle Vektorräume sowie D eine offene Teilmenge von V und $a \in D$. Weiter sei $f : D \to W$ eine $k+1$-mal stetig differenzierbare Abbildung und $v \in V$, so dass die Strecke $[a, a+v]$ in D enthalten ist. Dann gilt*

$$f(a+v) = \sum_{j=0}^{k} \frac{1}{j!} \delta_v^j f(a) + \int\limits_0^1 \frac{(1-t)^k}{k!} \delta_v^{k+1} f(a+tv)\,dt.$$

Dies ist eine Version der Taylorschen Formel. Für festes a nennt man die Summe auf der rechten Seite das *Taylor-Polynom* der Ordnung k von f an der Stelle a. Bezeichnen wir es mit $p_k(v)$, so hat die Formel die Gestalt

$$f(a+v) = p_k(v) + r_k(v),$$

wobei man $r_k(v)$ das *Restglied* nennt.

Beweis. Wir halten a, k und v fest. Wegen $[a, a+v] \subset D$ wird durch

$$g(t) = f(a+tv)$$

eine Hilfsfunktion $g : [0,1] \to W$ definiert. Nach Satz 10.3(iv) ist sie $k+1$-mal stetig differenzierbar, und nach der Definition der Richtungsableitung gilt

$$g'(t) = \delta_v f(a+tv).$$

Durch mehrmalige Anwendung dieser Formel folgt für $j \leq k+1$

$$g^{(j)}(t) = \delta_v^j f(a+tv).$$

Nun ergibt sich der Satz aus Lemma 10.2 mit $a = 0$ und $x = 1$ durch Einsetzen. $\square$

Wie wir wissen, lässt sich die im Satz auftretende k-fache Richtungsableitung durch das k-te Differential ausdrücken, nämlich

$$\delta_v^k f(a) = d^k f(a)(\underbrace{v, \ldots, v}_{k}).$$

Das Restglied r_1 stimmt mit dem Restglied in der Definition 10.3 überein. Ist also f an der Stelle a differenzierbar, so gilt $\frac{r_1(v)}{\|v\|} \to 0 \;\; (v \to 0)$. Wir wollen eine analoge Aussage für beliebige k beweisen, die man qualitative Taylorsche Formel nennen könnte.

Satz 10.6. *Es seien V, W endlichdimensionale reelle Vektorräume, $D \subseteq V$, und $f : D \to W$ sei k-mal differenzierbar in dem inneren Punkt a von D, wobei $k \geq 1$. Dann gilt für $v \in V$ mit $a+v \in D$*

$$f(a+v) = p_k(v) + r_k(v), \qquad \frac{r_k(v)}{\|v\|^k} \to 0_W \qquad (v \to 0_V).$$

Die erste Gleichung ist eigentlich nur eine Definition von r_k. Für den Beweis benötigen wir folgende Aussage: Für jede multilineare Abbildung $m \in \mathrm{Mult}^k(V,W) = \mathrm{Hom}(V, \mathrm{Mult}^{k-1}(V,W))$ existiert für vorgegebene Normen auf V und W eine reelle Zahl c, so dass für alle $v_1, \ldots, v_k \in V$ gilt

$$\|m(v_1, \ldots, v_k)\| \leq c\|v_1\| \cdots \|v_k\|.$$

Dies zeigt man durch vollständige Induktion mit Hilfe von Satz 8.20(ii). Man kann die kleinstmögliche solche Konstante mit $\|m\|$ bezeichnen und erhält so eine Norm auf $\mathrm{Mult}^k(V,W)$.

Beweis. Da die Behauptung für $k = 1$ nach Definition bereits gilt, sei nun $k > 1$. Laut Definition muss $f^{(k-1)}$ in einer Umgebung U von a existieren und an der Stelle a differenzierbar sein. Es gibt ein $\eta > 0$, so dass für $\|v\| < \eta$ gilt $a+v \in U$, also auch $[a, a+v] \subseteq U$. Da $f^{(k-2)}$ auf U stetig ist, können wir Satz 10.5 mit $k-2$ an Stelle von k anwenden und erhalten

$$f(a+v) = p_{k-2}(v) + \int_0^1 \frac{(1-t)^{k-2}}{(k-2)!} \delta_v^{k-1} f(a+tv)\, dt.$$

Wegen der Differenzierbarkeit von $f^{(k-1)} : U \to \mathrm{Mult}^{k-1}(V,W)$ an der Stelle a gilt für $\|u\| < \eta$

$$d^{k-1}f(a+u) = d^{k-1}f(a) + d^k f(a)u + r(u), \qquad \frac{r(u)}{\|u\|} \to 0 \qquad (u \to 0),$$

wobei $d^k f(a) \in \mathrm{Hom}(V, \mathrm{Mult}^{k-1}(V,W)) = \mathrm{Mult}^k(V,W)$. Setzen wir $u = tv$, so folgt

$$\delta_v^{k-1} f(a+tv) = \delta_v^{k-1} f(a) + t\delta_v^k f(a) + r(tv)(\underbrace{v,\dots,v}_{k-1}).$$

Beim Einsetzen ergeben angesichts von Satz 9.1(iv) und

$$\int_0^1 \frac{(1-t)^{k-2}}{(k-2)!}\, dt = \frac{1}{(k-1)!}, \qquad \int_0^1 \frac{t(1-t)^{k-2}}{(k-2)!}\, dt = \frac{1}{k!}$$

die ersten beiden Terme gerade die beiden restlichen Glieder von $p_k(v)$, und es folgt

$$r_k(v) = \int_0^1 \frac{(1-t)^{k-2}}{(k-2)!}\, r(tv)(v,\dots,v)\, dt.$$

Nach Satz 9.1 ist

$$\|r_k(v)\| \le \int_0^1 \frac{(1-t)^{k-2}}{(k-2)!}\, \|r(tv)(v,\dots,v)\|\, dt.$$

Mit den Bemerkungen vor dem Beweis folgt

$$\|r(u)(v_1,\dots,v_{k-1})\| \le \|r(u)\| \|v_1\| \cdots \|v_{k-1}\|.$$

Somit ist

$$\frac{\|r_k(v)\|}{\|v\|^k} \le \int_0^1 \frac{(1-t)^{k-2}}{(k-2)!} \cdot \frac{\|r(tv)\|}{\|v\|} dt.$$

Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für $\|u\| < \delta$ gilt

$$\frac{\|r(u)\|}{\|u\|} < k!\varepsilon.$$

Setzen wir $u = tv$ mit $\|v\| < \delta$ und $t \in [0,1]$, so ist dies erfüllt, und es folgt

$$\frac{\|r_k(v)\|}{\|v\|^k} < k!\varepsilon \int_0^1 \frac{t(1-t)^{k-2}}{(k-2)!}\, dt = \varepsilon. \qquad \square$$

Vorlesung 11a
21.12.22

Ist $V = \mathbb{R}^n$, so können wir die in der Taylorschen Formel vorkommenden mehrfachen Richtungsableitungen bezüglich eines Vektors $v = (v_1, \dots, v_n)$ durch partielle Ableitungen ausdrücken. Gleichung (15) spezialisiert sich nämlich zu

$$\delta_v^k f(a) = \sum_{i_1, \dots, i_k = 1}^{n} \delta_{i_k} \cdots \delta_{i_1} f(a) v_{i_1} \cdots v_{i_k}.$$

Ist f eine k-mal stetig differenzierbare Funktion in einer Umgebung von a, so können wir die partiellen Ableitungen nach Satz 10.4 umordnen und gleiche Ableitungen zusammenfassen. Ist α_i die Anzahl, wie oft die Zahl i unter den Zahlen $i_1, \dots, i_k$ vorkommt, so ist jedem k-Tupel $(i_1, \dots, i_k)$ ein n-Tupel $(\alpha_1, \dots, \alpha_n)$ mit der Eigenschaft

$$\alpha_1 + \ldots + \alpha_n = k$$

zugeordnet. Ein n-Tupel $(\alpha_1, \dots, \alpha_n)$ entsteht aus

$$\frac{k!}{\alpha_1! \cdots \alpha_n!}$$

verschiedenen k-Tupeln $(i_1, \dots, i_k)$, und es folgt

$$\frac{1}{k!} \delta_v^k f(a) = \sum_{\alpha_1 + \ldots + \alpha_n = k} \frac{\delta_1^{\alpha_1} \cdots \delta_n^{\alpha_n} f(a)}{\alpha_1! \cdots \alpha_n!} v_1^{\alpha_1} \cdots v_n^{\alpha_n}.$$

10.4 Lokale Extrema

Wir wollen die Kriterien für lokale Extrema auf den Fall von Funktionen von mehreren Variablen verallgemeinern.

Definition 10.8. *Es sei f eine reellwertige Funktion auf einem metrischen Raum X und a ein Punkt von X.*

(i) Die Funktion f hat an der Stelle a ein lokales Minimum*, wenn es eine Umgebung U von a in X gibt, so dass für alle $x \in U$ gilt $f(x) \geq f(a)$.*

(ii) Die Funktion f hat an der Stelle a ein striktes lokales Minimum, wenn es eine Umgebung U von a gibt, so dass für $x \in U \setminus \{a\}$ gilt $f(x) > f(a)$.

(iii) Analog definiert man ein (striktes) lokales Maximum.

(iv) Wir sagen, dass f an der Stelle a ein lokales Extremum *hat, wenn f dort ein lokales Minimum oder ein lokales Maximum hat.*

Im Folgenden sei D eine Teilmenge eines endlichdimensionalen reellen Vektorraums V. Wir formulieren nun ein notwendiges Kriterium, das Satz 6.4 verallgemeinert.

Satz 10.7. *Die Funktion $f : D \to \mathbb{R}$ habe an der Stelle $a \in \mathring{D}$ ein lokales Extremum. Existiert die Richtungsableitung bezüglich eines Vektors v, so ist*

$$\delta_v f(a) = 0.$$

Ist f an der Stelle a differenzierbar, so gilt

$$df(a) = 0.$$

Beweis. Es sei $g(h) = f(a + tv)$. Da die Abbildung $t \mapsto a + tv$ stetig ist, ist g in einer Umgebung der Stelle 0 definiert und hat dort ein lokales Extremum. Exixtiert $\delta_v f(a)$, so ist g an der Stelle 0 differenzierbar. Nach Satz 6.4 gilt $g'(0) = 0$, und die erste Behauptung folgt. Wegen $df(a)(v) = \delta_v f(a)$ folgt die zweite. □

Einen Punkt a, in dem die Richtungsableitungen von f bezüglich aller Vektoren verschwinden, nennt man *stationären Punkt* von f. Im Fall $V = \mathbb{R}^n$ müssen dazu notwendigerweise alle partiellen Ableitungen verschwinden. Ist f an der Stelle a differenzierbar, so ist das auch hinreichend für das Vorliegen eines stationären Punktes.

Beispiel. Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ sei gegeben durch

$$f(x, y) = xye^{-(x^2+y^2)/2}.$$

Die ersten partiellen Ableitungen sind

$$\begin{aligned}\delta_x f(x, y) &= \frac{\partial f}{\partial x} = y(1 - x^2)e^{-(x^2+y^2)/2},\\ \delta_y f(x, y) &= \frac{\partial f}{\partial y} = x(1 - y^2)e^{-(x^2+y^2)/2}.\end{aligned}$$

Ist (x, y) ein stationärer Punkt von f, so gilt also

$$\begin{aligned}y(1 - x^2) &= 0,\\ x(1 - y^2) &= 0.\end{aligned}$$

Dieses nichtlineare Gleichungssystem lässt sich ausnahmsweise leicht lösen. Ist $y \neq 0$, so muss $x = \pm 1$ und $y = \pm 1$ sein, ist aber $y = 0$, so muss auch $x = 0$ sein. Eine Probe zeigt, dass die stationären Punkte von f genau die Punkte

$$(1, 1), \quad (1, -1), \quad (-1, 1), \quad (-1, -1), \quad (0, 0)$$

sind. $\triangleleft$

Wir wollen nun ein hinreichendes Kriterium für lokale Extrema formulieren.

Definition 10.9. *Eine Multilinearform* $m \in \operatorname{Mult}^k(V, \mathbb{R})$ *heißt* positiv definit[55], *wenn für alle* $v \in V \setminus \{0\}$ *gilt* $m(v, \dots, v) > 0$.

Analog definiert man, wann eine Multilinearform negativ definit *ist.*

Eine Multilinearform m *heißt* indefinit, *wenn es sowohl Vektoren* $v \in V$ *mit der Eigenschaft* $m(v, \dots, v) > 0$ *als auch solche mit der Eigenschaft* $m(v, \dots, v) < 0$ *gibt.*

Wegen

$$m(tv, \dots, tv) = t^k m(v, \dots, v)$$

für $t \in \mathbb{R}$ kann es definite Multilinearformen vom Grad k nur geben, wenn k gerade ist. Es gibt Multilinearformen, die weder positiv definit noch negativ definit noch indefinit sind. Nun beweisen wir eine Verallgemeinerung von Satz 6.15.

Satz 10.8. *Die Funktion* $f : D \to \mathbb{R}$ *sei* k*-mal an der Stelle* $a \in \mathring{D}$ *differenzierbar, wobei* $k \geq 2$, *und es sei*

$$df(a) = 0, \quad \dots, \quad d^{k-1}f(a) = 0.$$

(i) Ist $d^k f(a)$ *positiv (bzw. negativ) definit, so hat* f *an der Stelle* a *ein lokales Minimum (bzw. Maximum).*

(ii) Ist $d^k f(a)$ *indefinit, so hat* f *an der Stelle* a *kein lokales Extremum.*

Beweis. Nach Satz 10.6 gilt für $a + v \in D$

$$f(a+v) = f(a) + m(v, \dots, v) + r_k(v), \qquad \frac{r_k(v)}{\|v\|^k} \to 0 \ \ (v \to 0_V),$$

wobei $m = \frac{1}{k!} f^{(k)}(a)$ ist und die Terme der Ordnung 1 bis $k-1$ nach Voraussetzung verschwinden.

Nun sei z. B. m positiv definit. Bereits im Beweis von Satz 8.20(i) hattten wir gesehen, dass die Menge $S = \{v \in V \mid \|v\| = 1\}$ kompakt ist. Durch mehrmalige Anwendung von Satz 8.8(ii) sieht man, dass m stetig ist, also nach Satz 8.11 auch die Abbildung $v \mapsto m(v, \dots, v)$. Ihre Einschränkung auf S hat nach Folgerung 8.8 ein Minimum c. Da m positiv definit ist, gilt

[55] Dies ist das lateinische Partizip von *definire* (bestimmen).

$c > 0$, und wir können $\varepsilon \in\,]0, c[$ wählen. Nach obiger Konvergenzaussage gibt es ein $\delta > 0$, so dass für $a + v \in D$ und $\|v\| < \delta$ gilt

$$|r_k(v)| < \varepsilon \|v\|^k.$$

Für v mit diesen Eigenschaften folgt

$$f(a+v) - f(a) = m\left(\frac{v}{\|v\|}, \dots, \frac{v}{\|v\|}\right)\|v\|^k + r_k(v) \geq c\|v\|^k - |r_k(v)|,$$

also

$$f(a+v) \geq f(a) + (c - \varepsilon)\|v\|^k,$$

und für $v \neq 0$ ist die rechte Seite größer als $f(a)$.

Nun sei m indefinit. Dann gibt es Vektoren u und v mit $m(u, \dots, u) < 0$ und $m(v, \dots, v) > 0$. Definieren wir $g(t) = f(a + tu)$ und $h(t) = f(a + tv)$, so gilt

$$g'(0) = h'(0) = \cdots = g^{(k-1)}(0) = h^{(k-1)}(0) = 0, \quad g^{(k)}(0) < 0, \quad h^{(k)}(0) > 0.$$

Wie eben zeigt man, dass in einer rechtsseitigen Umgebung der Null g kleiner und h größer als $f(a)$ ist. Folglich hat f an der Stelle a kein lokales Extremum.

□

Beispiel. Es sei f wie oben. Die zweiten partiellen Ableitungen sind

$$\begin{aligned}
\delta_x^2 f(x,y) &= \frac{\partial^2 f}{\partial x^2} = xy(x^2-3)e^{-(x^2+y^2)/2}, \\
\delta_x\delta_y f(x,y) &= \frac{\partial^2 f}{\partial x \partial y} = (1-x^2)(1-y^2)e^{-(x^2+y^2)/2}, \\
\delta_y^2 f(x,y) &= \frac{\partial^2 f}{\partial y^2} = xy(y^2-3)e^{-(x^2+y^2)/2}.
\end{aligned}$$

Das zweite Differential, angewendet auf zwei gleiche Vektoren, ist

$$d^2 f(a,b)\left(\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right), \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)\right) = \delta_x^2 f(a,b)u^2 + \delta_x\delta_y f(a,b)uv + \delta_y\delta_x f(a,b)vu + \delta_y^2 f(a,b)v^2,$$

wobei $\delta_x\delta_y f = \delta_y\delta_x f$ nach Satz 10.4. In den stationären Punkten erhalten wir

$$\begin{aligned}
d^2 f(\pm1, \pm1)\left(\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right), \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)\right) &= -\tfrac{2}{e}(u^2+v^2), \\
d^2 f(\pm1, \mp1)\left(\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right), \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)\right) &= \tfrac{2}{e}(u^2+v^2), \\
d^2 f(0,0)\left(\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right), \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)\right) &= \tfrac{2}{e}uv.
\end{aligned}$$

Somit hat f an den Stellen $(1,1)$ und $(-1,-1)$ lokale Maxima, an den Stellen $(1,-1)$ und $(-1,1)$ lokale Minima, und an der Stelle $(0,0)$ liegt kein lokales Extremum vor. $\triangleleft$

In den Anwendungen ist f meist in einer Umgebung des stationären Punktes k-mal stetig differenzierbar, so dass $d^k f$ nach Folgerung 10.3 eine symmetrische Multilinearform ist. Die Bilinearform $d^2 f(a)$ nennt man übrigens die *Hessesche Form* von f an der Stelle a.

Bei symmetrischen Multilinearformen m von beliebigem Grad k ist die Definitheit schwer zu entscheiden, aber im Fall $k=2$ (wenn also $m=b$ eine symmetrische Bilinearform ist) gibt es einen einfachen Algorithmus. Es genügt, die zugehörige quadratische Form $q(v)=b(v,v)$ zu betrachten, aus der sich die Bilinearform durch sogenannte Polarisierung

$$2b(u,v)=q(u+v)-q(u)-q(v)$$

zurückgewinnen lässt. Die meist nach Gram und Schmidt benannte Methode (die aber schon vor ihnen bekannt war) liefert eine Basis $e_1, \ldots, e_n$ von V, so dass

$$b(e_i,e_j)=0 \quad \text{für } i\neq j, \qquad b(e_i,e_i)\in\{1,-1,0\}.$$

Daran lässt sich die Definitheit leicht ablesen.

An Stelle dieser Methode beschreiben wir die analoge Methode der quadratischen Ergänzung. Ist V durch Wahl einer Basis bereits mit $\mathbb{R}^n$ identifiziert, so ist b in Matrizenschreibweise durch

$$b(u,v)=\begin{pmatrix}u_1 & \ldots & u_n\end{pmatrix}\begin{pmatrix}a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \ldots & a_{nn}\end{pmatrix}\begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}$$

und q in der Form

$$\begin{aligned}q(v)= \ & a_{11}v_1^2+a_{12}v_1v_2+\cdots+a_{1n}v_1v_n\\ &+a_{21}v_2v_1+a_{22}v_2^2+\cdots+a_{2n}v_2v_n\\ &\qquad\ldots\\ &+a_{n1}v_nv_1+a_{n2}v_nv_2+\cdots+a_{2n}v_n^2\end{aligned}$$

gegeben, wobei $a_{ij}=a_{ji}$. Man kann $a_{ij}v_iv_j$ und $a_{ji}v_jv_i$ für $i\neq j$ zu $2a_{ij}v_iv_j$ zusammenfassen.

Ist $a_{11}\neq 0$, so gehen wir zu den Koordinaten $u_1, v_2, \ldots, v_n$ über, wobei

$$u_1=v_1+\frac{a_{12}}{a_{11}}v_2+\ldots+\frac{a_{1n}}{a_{11}}v_n.$$

Ist hingegen $a_{11} = 0$, so vertauschen wir die Nummerierung der Variablen, um in die obige Situation zu kommen. Verschwinden alle a_{ii} (wie in unserem Beispiel), so hilft das nichts. In diesem Fall wählen wir Indizes $i \neq j$, so dass $a_{ij} \neq 0$, und ersetzen die Koordinaten v_i und v_j durch

$$v_i' = v_i + v_j, \qquad v_j' = v_i - v_j.$$

Dann ist

$$2v_i = v_i' + v_j', \qquad 2v_j = v_i' - v_j',$$

also

$$2a_{ij}v_iv_j = \frac{a_{ij}}{2}(v_i'^2 - v_j'^2).$$

Nun können wir die obige Substitution vornehmen.

Im Ergebnis kommt u_1 nur in dem einen Term $a_{11}u_1^2$ vor, und die übrigen Terme bilden eine quadratische Form in den restlichen Variablen, auf die man die selbe Methode rekursiv anwenden kann. Schließlich erhält man eine quadratische Form

$$b_1u_1^2 + \ldots + b_nu_n^2.$$

Diese ist offensichtlich genau dann definit, wenn alle Koeffizienten b_i von Null verschieden sind und das selbe Vorzeichen haben, und sie ist genau dann indefinit, wenn sowohl positive als auch negative Koeffizienten vorkommen.

10.5 Differentiation parameterabhängiger Integrale

Vorlesung 11b
23.12.22

Wir untersuchen jetzt, wann die durch ein parameterabhängiges Integral definierte Funktion

$$I(t) = \int_a^b f(x,t)\,dx$$

für feste reelle Zahlen $a \leq b$ differenzierbar vom Parameter t abhängt.

Satz 10.9. *Es sei $T \subseteq \mathbb{R}$ offen, die Funktion $f : [a,b] \times T \to \mathbb{R}$ sei nach dem ersten Argument integrierbar und nach dem zweiten Argument partiell differenzierbar, und die partielle Ableitung*

$$\delta_2 f = \frac{\partial f}{\partial t} : [a,b] \times T \to \mathbb{R}$$

sei für ein festes $u \in T$ an allen Stellen der Form (x,u) stetig. Dann ist I an der Stelle u differenzierbar, und

$$I'(u) = \int_a^b \delta_2 f(x,u)\,dx.$$

Beweis. Wir definieren $g : [a, b] \times T \to \mathbb{R}$ durch

$$g(x,t) = \begin{cases} \dfrac{f(x,t) - f(x,u)}{t-u}, & \text{falls } t \neq u, \\ \delta_2 f(x,u), & \text{falls } t = u. \end{cases}$$

Wir wollen die Stetigkeit von g an einer Stelle (c, u) mit $c \in [a, b]$ zeigen. Nach dem Mittelwertsatz der Differentialrechnung gibt es für jedes $(x, t) \in [a, b] \times T$ ein $\vartheta \in]0, 1[$, so dass

$$f(x,t) - f(x,u) = \delta_2 f(x, u + \vartheta(t-u))(t-u)$$

und somit

$$g(x,t) - g(c,u) = \delta_2 f(x, u + \vartheta(t-u)) - \delta_2 f(c,u).$$

Ist $\varepsilon > 0$, so gibt es wegen der Stetigkeit von $\delta_2 f$ an der Stelle (c, u) ein $\eta > 0$, so dass für $(x, t) \in [a, b] \times T$ mit den Eigenschaften $|x - c| < \eta$ und $|t - u| < \eta$ gilt

$$|\delta_2 f(x,t) - \delta_2 f(c,u)| < \varepsilon.$$

Letzteres gilt dann auch für $u + \vartheta(t - u)$ an Stelle von t, und es folgt

$$|g(x,t) - g(c,u)| < \varepsilon.$$

Da ε beliebig war, ist g an der Stelle (c, u) stetig.

Nach Aufgabe 36 erfüllt g die Voraussetzungen von Satz 9.3, so dass

$$\lim_{t \to u} \int_a^b g(x,t)\,dx = \int_a^b g(x,u)\,dx.$$

Setzen wir die Definition von g ein, so ergibt sich

$$\lim_{t \to u} \frac{I(t) - I(u)}{t-u} = \int_a^b \delta_2 f(x,u)\,dx,$$

also ist I an der Stelle u differenzierbar, und die behauptete Formel gilt. □

Ist $\delta_2 f$ auf ganz $[a, b] \times T$ stetig, so ist I nach Satz 9.3 auf T stetig differenzierbar.

Beispiel. Die Funktion

$$f(x,t) = \frac{\ln(x+t)}{x}$$

ist stetig auf $\{(x,t) \in \mathbb{R}^2 \mid x \neq 0,\ x+t > 0\}$. Für feste $a \leq b$ ist die Funktion

$$I(t) = \int\limits_a^b \frac{\ln(x+t)}{x}\,dx$$

für $t > -a$ definiert. Man kann beweisen, dass I keine elementare Funktion ist, also helfen die Ableitungsregeln nicht weiter. Mit Satz 10.9 folgt

$$I'(t) = \int\limits_1^a \frac{dx}{x(x+t)}.$$

Für $t \neq 0$ ist

$$I'(t) = \frac{1}{t}\int\limits_1^a \left(\frac{1}{x} - \frac{1}{x+t}\right) dx = \frac{\ln x - \ln(x+t)}{t}\bigg|_{x=a}^{x=b} = \frac{1}{t}\ln\frac{a(t+b)}{b(t+a)},$$

während

$$I'(0) = \int\limits_a^b \frac{dx}{x^2} = -\frac{1}{x}\bigg|_{x=a}^{x=b} = \frac{1}{a} - \frac{1}{b}.$$

Nach dem Satz ist I' stetig, was sich natürlich auch aus den Rechenregeln für Grenzwerte ergibt. $\triangleleft$

Ein Integral kann auch von mehreren Parametern abhängen und Werte in einem Vektorraum über einem Körper K annehmen. Dabei sei $K = \mathbb{R}$ oder $K = \mathbb{C}$, und Differenzierbarkeit bedeute K-Differenzierbarkeit.

Satz 10.10. *Es seien V und W endlichdimensionale K-Vektorräume, T eine offene Teilmenge von V und $f : [a,b] \times T \to W$ nach dem ersten Argument integrierbar. Das k-te partielle Differential*

$$\partial_{\mathrm{II}}^k f : [a,b] \times T \to \mathrm{Mult}^k(V,W)$$

nach dem zweiten Argument existiere und sei stetig. Die oben definierte Abbildung $I : T \to W$ ist dann k-mal stetig differenzierbar, und es gilt

$$d^k I(t) = \int\limits_a^b \partial_{\mathrm{II}}^k f(x,t)\,dx.$$

Beweis. Für $k = 0$ folgt dies aus Satz 9.3. Nun sei $k = 1$. Indem wir eine reelle Basis von W wählen, können wir Satz 10.9 auf W-wertige Funktionen verallgemeinern und für festes $u \in T$ und $v \in V$ auf $f(x, u+tv)$ als Funktion von t anwenden. Wir erhalten

$$\delta_v I(u) = \int_a^b \delta_{(0,v)} f(x,u)\, dx.$$

Dies gilt für alle $u \in T$, und da $\delta_{(0,v)} f$ stetig ist, folgt mit Satz 9.3 die Stetigkeit von $\delta_v I$. Dies gilt für alle v, und mit Satz 10.1 folgt die stetige reelle Differenzierbarkeit von I sowie die behauptete Formel für $k = 1$. Mit Satz 9.1(vi) folgt die K-Linearität von $dI(t)$.

Den Beweis für beliebige k führen wir durch vollständige Induktion. Angenommen, die Behauptung gilt für eine Zahl k. Wenn f nun $k+1$-mal nach dem zweiten Argument differenzierbar und $\partial_{\mathrm{II}}^{k+1} f$ stetig ist, so ist nach dem Bewiesenen zunächst

$$dI(t) = \int_a^b \partial_{\mathrm{II}} f(x,t)\, dx.$$

Da die Funktion $\partial_{\mathrm{II}} f$ nun k-mal nach dem zweiten Argument differenzierbar und $\partial_{\mathrm{II}}^k(\partial_{\mathrm{II}} f)$ stetig ist, können wir die Induktionsvoraussetzung darauf anwenden. Die Funktion dI ist also k-mal stetig differenzierbar, und

$$d^k(dI)(t) = \int_a^b \partial_{\mathrm{II}}^k(\partial_{\mathrm{II}} f)(x,t)\, dx.$$

Damit gilt die Behauptung auch für $k+1$. $\square$

Beispiel. Durch das *Eulersche Integral* erster Art

$$\mathrm{B}(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1} dx$$

wird die *Betafunktion*[56] definiert. Sie ist auf der offenen Teilmenge

$$\{(p,q) \in \mathbb{C}^2 \mid \operatorname{Re} p > 1,\ \operatorname{Re} q > 1\}$$

von $\mathbb{C}^2$ nach Folgerung 10.10 unendlich oft $\mathbb{C}$-differenzierbar. Man kann sie sogar für $\operatorname{Re} p > 0$ und $\operatorname{Re} q > 0$ als absolut konvergentes uneigentliches Integral definieren. Durch partielle Integration kann man zeigen, dass

$$p\mathrm{B}(p,q+1) = q\mathrm{B}(p+1,q),$$

[56]Der Buchstabe B ist ein großes Beta.

und durch Substitution sieht man, dass $\mathrm{B}(p,q) = \mathrm{B}(q,p)$. $\triangleleft$

Es gibt auch ein Eulersches Integral zweiter Art, nämlich

$$\Gamma(s) = \int_0^\infty e^{-x} x^{s-1}\, dx,$$

das für $\operatorname{Re} s > 0$ absolut konvergent ist und die *Gammafunktion* definiert. Es ist besser, ihre Differenzierbarkeit erst später mit Hilfe des Lebesgueschen Integrals zu behandeln, weil die Verallgemeinerung von Satz 10.9 auf uneigentliche Riemannsche Integrale unhandlich ist.

11 Nichtlineare Gleichungen

11.1 Das Newtonverfahren

Vorlesung 12a
11.01.23

Eine nichtlineare Gleichung mit n Unbekannten kann man in der Form

$$f(x_1, \ldots, x_n) = 0$$

schreiben, wobei f eine Funktion von n Variablen ist. Ihre Lösungen zu finden bedeutet, das Urbild der Null zu bestimmen. Wir fassen die Variablen zu einem Punkt $x \in K^n$ zusammen, wobei $K \in \{\mathbb{R}, \mathbb{C}\}$ ist. Fragt man nach den Lösungen von

$$f(x) = y$$

für gegebenes y, so ist dies zwar nicht allgemeiner, aber man kann die Abhängigkeit von y untersuchen. Auch ein System von nichtlinearen Gleichungen kann man in der selben Weise verstehen, wenn man vektorwertige Funktionen f zulässt. Wir wollen zunächst Bedingungen finden, unter denen die Lösung in einer Teilmenge des Definitionsbereichs von f eindeutig ist.

Ein Spezialfall sind Systeme von linearen Gleichungen, die man in der Form

$$l(x) = y$$

schreiben kann, wobei $l : V \to W$ eine lineare Abbildung ist. In diesem Fall ist aus der linearen Algebra bekannt, dass die Lösung nur eindeutig sein kann, wenn l invertierbar ist, und dazu müssen V und W die selbe Dimension haben.

Wir wollen den Fall differenzierbarer Abbildungen $f : D \to W$ betrachten, wobei D eine offene Teilmenge von V ist. Für einen Punkt $a \in D$, in dem $df(a) \in \operatorname{Hom}(V, W)$ invertierbar ist, ersetzen wir f durch die affine Abbildung

$$f(a) + df(a)(x - a).$$

Die Lösung der linearisierten Gleichung

$$f(a) + df(a)(x - a) = y$$

ist dann

$$x_1 = a + df(a)^{-1}(y - f(a)).$$

Dies ist zwar im allgemeinen keine Lösung der Ausgangsgleichung, aber wir hoffen, dass x_1 näher an der Lösung liegt als a. Durch Iteration erhält man das *Newtonverfahren*: Man definiert rekursiv eine Folge x_k durch

$$x_0 = a, \qquad x_{k+1} = x_k + df(x_k)^{-1}(y - f(x_k))$$

(solange $x_k \in D$ ist).

Beispiel. Wir suchen eine Lösung der Gleichung

$$x^2 = y.$$

Hier ist $df(a)$ durch Multiplikation mit $f'(a) = 2a$ gegeben, also

$$x_1 = a + \frac{y - a^2}{2a} = \frac{1}{2}\left(a + \frac{y}{a}\right).$$

Das Newtonverfahren ist also in diesem Fall nichts anderes als das Heronverfahren. $\triangleleft$

Das numerische Invertieren einer linearen Abbildung ist sehr aufwendig. Darum benutzt man manchmal das modifizierte Newtonverfahren

$$x_{k+1} = x_k + df(a)^{-1}(y - f(x_k)).$$

Lemma 11.1. *Die Abbildung $f : D \to W$ sei auf der offenen Teilmenge D von V stetig differenzierbar, und $df(a)$ sei invertierbar. Dann gibt es eine Umgebung U von $b = f(a)$ und eine Teilmenge X von D, so dass für jedes $y \in U$ genau eine Lösung der Gleichung*

$$f(x) = y$$

in $\mathring{X}$ existiert und das modifizierte Newtonverfahren mit beliebigem Anfangswert $x_0 \in X$ gegen diese Lösung konvergiert.

Beweis. Zur Abkürzung sei $df(a)^{-1} = l \in \mathrm{Hom}(W, V)$. Wir schreiben die Rekursionsformel im modifizierten Newtonverfahren in der Form

$$x_{k+1} = h_y(x_k),$$

wobei die Abbildung $h_y : D \to V$ für jedes $y \in W$ durch

$$h_y(x) = x + l(y - f(x))$$

gegeben ist. Ein Punkt $x \in D$ ist genau dann Fixpunkt von h_y, wenn $f(x) = y$ ist.

Nach Satz 10.2 ist h_y stetig differenzierbar, und

$$dh_y(x) = \mathrm{id}_V - df(a)^{-1} \circ df(x).$$

Offensichtlich ist $dh_y(a) = 0$. Wegen der Offenheit von D und der Stetigkeit von dh_y existiert ein $\delta > 0$, so dass für $x \in V$ mit der Eigenschaft $\|x - a\| \leq \delta$ gilt

$$x \in D, \qquad \|dh_y(x)\| \leq \frac{1}{2}.$$

Es sei $X = \{x \in V \mid \|x - a\| \leq \delta\}$. Für $x_1, x_2 \in X$ können wir Satz 9.2 und Satz 9.1(iv) auf $h_y(x_1 + t(x_2 - x_1))$ anwenden und erhalten

$$\|h_y(x_1) - h_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\|.$$

Um zu sehen, ob h_y die Menge X in sich selbst abbildet, schreiben wir

$$h_y(x) - a = (h_y(x) - h_y(a)) + l(y - b).$$

Setzen wir $U = \{y \in W \mid \|y - b\| < \frac{\delta}{2\|l\|}\}$, so folgt für $x \in X$ und $y \in U$

$$\|h_y(x) - a\| \leq \frac{1}{2}\|x - a\| + \|l\|\|y - b\| < \delta,$$

also $h_y(x) \in \mathring{X}$. Somit ist h_y für $y \in U$ eine Kontraktion von X.

Da X nach Folgerung 8.7 kompakt und nach Satz 8.17(ii) vollständig ist, hat h_y nach Satz 8.7 genau einen Fixpunkt in X, der offenbar ein innerer Punkt ist. Nach dem Beweis von Satz 8.7 konvergiert das modifizierte Newtonverfahren gegen diesen Fixpunkt. $\square$

Wir können nun endlich Satz 6.3 auf Funktionen von mehreren Variablen verallgemeinern.

Satz 11.1. *Es seien V und W endlichdimensionale K-Vektorräume.*

(i) *Ist D eine offene Teilmenge von V, $a \in D$ und $f : D \to W$ eine stetig differenzierbare Abbildung, wobei $df(a)$ invertierbar ist, dann gibt es Umgebungen U_1 von a in V und U von $b = f(a)$ in W, so dass U_1 von f bijektiv auf U abgebildet wird.*

(ii) *Bildet die k-mal stetig differenzierbare Abbildung f die offene Teilmenge U_1 von V bijektiv auf die offene Teilmenge U von W ab, wobei $k \geq 1$ und $df(x)$ für alle $x \in U_1$ invertierbar ist, so ist die Umkehrabbildung $g : U \to U_1$ ebenfalls k-mal stetig differenzierbar, und für alle $y \in U$ gilt*

$$dg(y) = df(g(y))^{-1}.$$

In der Situation von Aussage (ii) ist f also ein Diffeomorphismus der Klasse C^k von U_1 auf U.

Beweis. **(i)** Es seien U und X wie in Lemma 11.1. Da es für jedes $y \in U$ genau eine Lösung von $f(x) = y$ in $\mathring{X}$ gibt, wird $U_1 = f^{-1}(U) \cap \mathring{X}$ bijektiv auf U abgebildet.

Zur Vorbereitung auf (ii) beweisen wir noch die Lipschitz-Stetigkeit der Umkehrabbildung $g : U \to U_1$ mit X und U wie im Beweis von Lemma 11.1. Für x_1, $x_2 \in X$ gilt laut Definition von h_y

$$x_1 - x_2 = h_y(x_1) - h_y(x_2) + l(f(x_1) - f(x_2)),$$

also

$$\|x_1 - x_2\| \leq \frac{1}{2}\|x_1 - x_2\| + \|l\|\|f(x_1) - f(x_2)\|$$

und schließlich

$$\|x_1 - x_2\| \leq 2\|l\|\|f(x_1) - f(x_2)\|.$$

Sind y_1, $y_2 \in U$ und setzen wir $x_i = g(y_i)$, so folgt

$$\|g(y_1) - g(y_2)\| \leq 2\|l\|\|y_1 - y_2\|.$$

(ii) Auf der Menge U in der neuen Bedeutung ist g nach dem Bewiesenen lokal Lipschitz-stetig, also stetig. Wegen der Differenzierbarkeit von f an einer beliebigen Stelle $a \in U_1$ gibt es nach Lemma 10.1 eine Funktion $\tilde{f} : D \to \mathrm{Hom}(V, W)$, die an der Stelle a stetig ist, so dass

$$f(x) - f(a) = \tilde{f}(x)(x - a).$$

Sind nun b, $y \in U$, so folgt durch Anwendung auf $a = g(b)$ und $x = g(y)$

$$y - b = \tilde{f}(g(y))(g(y) - g(b)).$$

Die Teilmenge der invertierbaren Elemente von $\mathrm{Hom}(V, W)$ ist nach Aufgabe 55 offen, also eine Umgebung von $\tilde{f}(a) = df(a)$. Nach Satz 8.10 gibt es eine Umgebung U' von a in U_1, so dass $\tilde{f}(x)$ für $x \in U'$ invertierbar ist. Für y in $g^{-1}(U')$, was hach Satz 8.10 eine Umgebung von b ist, folgt

$$g(y) - g(b) = \tilde{f}(g(y))^{-1}(y - b),$$

und $\tilde{f} \circ g$ ist nach Satz 8.11 an der Stelle b stetig. Nach Lemma 10.1 ist g also an der Stelle b differenzierbar, und wegen $\tilde{f}(a) = df(a)$ folgt die behauptete Formel für dg an der Stelle b.

Wir beweisen nun die Behauptung für allgemeines $k \geq 1$ durch vollständige Induktion. Da die Abbildung, die jeder invertierbaren linearen Abbildung $V \to W$ ihr Inverses zuordnet, nach Aufgabe 55 stetig ist, folgt mit Satz 10.2(iii), dass g stetig differenzierbar ist. Damit ist der Induktionsanfang erledigt. Angenommen, die Aussage gilt für eine Zahl k. Ist f nun $k+1$-mal stetig differenzierbar, so ist $df : D \to \mathrm{Hom}(V, W)$ nach Definition k-mal stetig differenzierbar und g nach Induktionsvoraussetzung k-mal stetig differenzierbar. Da die Abbildung, die jedem invertierbaren Element von $\mathrm{Hom}(V, W)$ sein Inverses zuordnet, nach Aufgabe 55 unendlich oft differenzierbar ist, folgt mit Satz 10.3, dass dg nach der obigen Formel k-mal stetig differenzierbar, also g nach Definition $k+1$-mal stetig differenzierbar ist. $\square$

Beispiel. Die durch $f(x) = x^2$ gegebene Funktion $f : K \to K$ hat die Ableitung $f'(x) = 2x$, die für $x \neq 0$ invertierbar ist. Also hat jeder Punkt $a \in K \setminus \{0\}$ eine Umgebung U_1, so dass die Einschränkung von f auf U_1 invertierbar ist. Die Lösung der Gleichung $x^2 = y$ ist bis auf das Vorzeichen bestimmt. Während man im Fall $K = \mathbb{R}$ üblicherweise $U_1 =]0, \infty[$ wählt, so dass nur ein Zweig der Parabel den Graphen der Wurzelfunktion darstellt, gibt es im Fall $K = \mathbb{C}$ mehrere naheliegende Möglichkeiten. Wählt man die offene rechte Halbebene als U_1, so erhält man den sogenannten *Hauptzweig* der Wurzel. $\triangleleft$

Beispiel. Die Funktion $\exp : K \to K$ hat die Ableitung exp, die in jedem Punkt invertierbar ist. Während die Exponentialfunktion im Fall $K = \mathbb{R}$ injektiv ist, ist die Lösung der Gleichung $\exp z = w$ für gegebenes $w \in \mathbb{C} \setminus \{0\}$ nach den Sätzen 4.14 und 5.8(iii) nur bis auf Addition von Vielfachen von $2\pi i$ bestimmt. Ist die Einschränkung von exp auf eine offene Teilmenge $U_1 \subseteq \mathbb{C}$ injektiv und stetig, so nennt man ihre Umkehrfunktion einen *Zweig* des Logarithmus. Den Hauptzweig erhält man bei der Wahl der Menge $U_1 = \{z \in \mathbb{C} \bigm| |\operatorname{Im} z| < \pi\}$. $\triangleleft$

Beispiel. Jede $\mathbb{C}$-differenzierbare Abbildung ist ja auch $\mathbb{R}$-differenzierbar. Schreiben wir im vorigen Beispiel $w = x + iy$ und $z = s + it$, so erhalten wir

$$x = e^s \cos t, \qquad y = e^s \sin t.$$

Bezeichnet man noch $e^s = r$, so erhält man eine unendlich oft differenzierbare Abbildung $]0, \infty[\times \mathbb{R} \to \mathbb{R}^2 \setminus \{(0,0)\}$. Durch Einschränkung auf eine Teilmenge U_1 erhält man einen Diffeomorphismus auf eine Teilmenge U. Man nennt (r, t) dann die *Polarkoordinaten* des Punktes $(x, y) \in U$. $\triangleleft$

11.2 Implizite Funktionen

Vorlesung 12b 13.01.23

Wir wollen nun Gleichungen betrachten, die sich nicht einmal lokal eindeutig lösen lassen. Dieses Phänomen tritt schon bei linearen Gleichungssystemen auf. Dort stellt sich heraus, dass man einen Teil der Variablen frei wählen kann und dann die restlichen Variablen eindeutig bestimmt sind. Wir wollen dieses Ergebnis auf nichtlineare Gleichungen verallgemeinern.

Beispiel. Betrachten wir die Gleichung

$$x^2 + y^2 = 1,$$

so gibt es für gegebenes $x \in \left]-1, 1\right[$ genau zwei Lösungen $y = \pm\sqrt{1-x^2}$. Wählt man ein Vorzeichen aus, so erhält man y als stetige Funktion von x. Für diese Funktion haben wir eine *explizite*[57] (das heißt, nach der abhängigen Variablen aufgelöste) Formel. $\triangleleft$

Beispiel. Bei der Gleichung

$$e^{xy} = x^2 + y$$

können wir die Lösungen $(0, 1)$ und $(\pm 1, 0)$ erraten, aber es ist unmöglich, die Gleichung nach y aufzulösen. $\triangleleft$

Allgemein betrachten wir Gleichungen der Form $G(x, y) = 0$. Für festes x haben wir eine Gleichung mit der Unbekannten y, die man vielleicht mit Hilfe des modifizierten Newtonverfahrens lösen kann. Dazu muss das Differential der zugehörigen partiellen Funktion $y \mapsto G(x, y)$, also das partielle Differential $\partial_{\mathrm{II}} G$ bezüglich des zweiten Arguments, an einer Stelle invertierbar sein, die einer Lösung genügend nahe kommt. Gibt es (eventuell nach Einschränkung von G auf einen kleineren Definitionsbereich) für jedes x genau eine Lösung y, dann hängt diese von x ab. Man sagt, dass die so entstehende Funktion *implizit*[58] durch die Gleichung $G(x, y) = 0$ gegeben ist. Wir fragen uns, unter welchen Voraussetzungen diese Funktion existiert sowie stetig oder sogar differenzierbar ist.

Satz 11.2. *Es seien V_{I}, V_{II} und W endlichdimensionale K-Vektorräume, D eine offene Teilmenge von $V_{\mathrm{I}} \times V_{\mathrm{II}}$ und $k \geq 1$. Weiter seien eine k-mal stetig differenzierbare Abbildung $G : D \to W$ und ein Punkt $(a, b) \in D$ gegeben, so dass $G(a, b) = 0$ und $\partial_{\mathrm{II}} G(a, b) \in \mathrm{Hom}(V_{\mathrm{II}}, W)$ invertierbar ist.*

Dann gibt es eine Umgebung U von (a, b) in D und eine k-mal stetig differenzierbare Abbildung $f : D_{\mathrm{I}} \to V_{\mathrm{II}}$, wobei D_{I} eine offene Teilmenge von V_{I} ist, so dass

$$\{(x, y) \in U \mid G(x, y) = 0\} = \{(x, f(x)) \mid x \in D_{\mathrm{I}}\}.$$

[57] lat. „entfaltete", von griechisch $\pi\lambda\acute{\varepsilon}\kappa\omega$ (falten)

[58] lat. „verwickelt"

Die Menge auf der rechten Seite ist natürlich der Graph von f, und da (a,b) in der linken Menge enthalten ist, gilt $f(a) = b$.

Ist eine Funktion $g : V_{\mathrm{II}} \to V_{\mathrm{I}}$ gegeben und wenden wir den Satz auf die Funktion

$$G(x, y) = g(y) - x$$

an, so erhalten wir Satz 11.1 (mit vertauschten Bezeichnungen $f(x)$ und $g(y)$) als Spezialfall. Man könnte den damaligen Beweis auf die jetzige Situation verallgemeinern, aber durch einen Kunstgriff können wir uns diese Arbeit ersparen.

Beweis. Wir definieren eine Abbildung $F : D \to V_{\mathrm{I}} \times W$ durch

$$F(x, y) = (x, G(x, y)).$$

Diese ist (komponentenweise) k-mal stetig differenzierbar, und nach Satz 10.2(ii) gilt

$$dF(a, b)(u, v) = (u, \partial_{\mathrm{I}} G(a, b)(u) + \partial_{\mathrm{II}} G(a, b)(v)).$$

Ist dieser Wert vorgegeben, so bestimmt man aus der ersten Komponente zunächst u, und wegen der Invertierbarkeit von $\partial_y G(a, b)$ ist auch v bestimmt. Die lineare Abbildung $dF(a, b)$ ist also invertierbar.

Nach Satz 11.1 ist die Einschränkung von F auf eine geeignete Umgebung von (a, b) in D ein Diffeomorphismus der Klasse C^k. Wählen wir die Norm $\|(x, y)\| = \max\{\|x\|, \|y\|\}$ auf $V_{\mathrm{I}} \times V_{\mathrm{II}}$, so sehen wir, dass diese Umgebung von (a, b) eine Menge der Form $U_{\mathrm{I}} \times U_{\mathrm{II}}$ enthält, wobei U_{I} eine Umgebung von a und U_{II} eine Umgebung von b ist. Die Umkehrabbildung von $F|_{U_{\mathrm{I}} \times U_{\mathrm{II}}}$ bezeichnen wir mit H. Diese ist also k-mal stetig differenzierbar.

Es sei $D_{\mathrm{I}} = \{x \in U_{\mathrm{I}} \mid (x, 0) \in F(U_{\mathrm{I}} \times U_{\mathrm{II}})\}$. Dies ist eine Umgebung von a, denn $(a, 0) = F(a, b)$. Somit ist $U = D_{\mathrm{I}} \times U_{\mathrm{II}}$ eine Umgebung von (a, b) in D. Für $x \in D_{\mathrm{I}}$ bezeichnen wir die zweite Komponente von $H(x, 0) \in D_{\mathrm{I}} \times U_{\mathrm{II}}$ mit $f(x)$. Dann ist $f : D_{\mathrm{I}} \to U_{\mathrm{II}}$ eine k-mal stetig differenzierbare Abbildung.

Ist $(x, y) \in U$ derart, dass $G(x, y) = 0$ ist, so ist $F(x, y) = (x, 0)$, also $(x, y) = H(x, 0)$ und $f(x) = y$. Ist umgekehrt $x \in D_{\mathrm{I}}$ und $y = f(x)$, so ist $H(x, 0) = (x, y)$, also $F(x, y) = (x, 0)$ und $G(x, y) = 0$. □

Bemerkung. Bilden wir in den Bezeichnungen des Satzes das Differential von $G(x, f(x)) = 0$ als Funktion von $x \in D_{\mathrm{I}}$, so ergibt sich nach Satz 10.2

$$\partial_{\mathrm{I}} G(a, b) + \partial_{\mathrm{II}} G(a, b) \circ df(a) = 0.$$

Man kann also

$$df(a) = -\partial_{\mathrm{II}} G(a, b)^{-1} \circ \partial_{\mathrm{I}} G(a, b)$$

bestimmen, auch wenn man keine explizite Formel für f hat. Dieses Verfahren, das sich auf höhere Ableitungen verallgemeinern lässt, nennt man *implizite Differentiation.*

Beispiel. Die durch
$$G(x, y) = e^{xy} - x^2 - y$$
gegebene Funktion $G : \mathbb{R}^2 \to \mathbb{R}$ hat die partiellen Ableitungen

$$\delta_1 G(x, y) = y e^{xy} - 2x, \qquad \delta_2 G(x, y) = x e^{xy} - 1.$$

Insbesondere ist $\delta_2 G(0, 1) = -1$ invertierbar, also gibt es Zahlen $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ und eine Funktion $f :]-\varepsilon_1, \varepsilon_1[\to \mathbb{R}$, so dass für $|x| < \varepsilon_1$ und $y \in \mathbb{R}$ genau dann $f(x) = y$ gilt, wenn

$$e^{xy} = x^2 + y, \qquad |y - 1| < \varepsilon_2.$$

Außerdem ist
$$f'(0) = -\delta_2 G(0, 1)^{-1} \delta_1 G(0, 1) = 1. \qquad \triangleleft$$

11.3 Stationäre Punkte unter Nebenbedingungen

Vorlesung 13a
18.01.23

Wir suchen nach lokalen Extrema der Einschränkung einer Funktion f auf die Lösungsmenge eines nichtlinearen Gleichungssystems, das wir in der Form $g(x) = 0$ schreiben. Die Gleichungen dieses Systems nennt man auch *Nebenbedingungen*. Hierfür ist das Kriterium aus Satz 10.7 nicht anwendbar.

Das folgende Ergebnis lässt sich am besten formulieren, wenn man einen Begriff aus der linearen Algebra benutzt. Dort nennt man das Urbild des Nullvektors unter einer linearen Abbildung l den *Kern* von l, abgekürzt $\operatorname{Ker} l$. Dies ist ein linearer Unterraum. Es gilt genau dann $\operatorname{Ker} l = \{0\}$, wenn l injektiv ist.

Satz 11.3. *Es seien V und W endlichdimensionale reelle Vektorräume und D eine offene Teilmenge von V. Weiter seien Abbildungen*

$$f : D \to \mathbb{R}, \qquad g : D \to W$$

gegeben, wobei g stetig differenzierbar ist. Hat die Einschränkung von f auf die Menge
$$M = \{x \in D \mid g(x) = 0\}$$
an der Stelle c ein lokales Extremum und ist f an dieser Stelle differenzierbar sowie $dg(c) : V \to W$ surjektiv, so gilt

$$df(c)|_{\operatorname{Ker} dg(c)} = 0.$$

Die letzte Zeile kann man auch durch

$$\operatorname{Ker} dg(c) \subseteq \operatorname{Ker} df(c)$$

ausdrücken. Der Kürze halber werden wir eine Nullstelle c von g, für welche diese Bedingung erfüllt ist, einen stationären Punkt von f unter der Nebenbedingung $g(x) = 0$ nennen. Im Spezialfall $W = \{0\}$ handelt es sich um einen stationären Punkt im üblichen Sinne.

Beweis. Es sei c wie im Satz und $V_{\mathrm{I}} = \operatorname{Ker} dg(c)$. Dann gibt es einen Unterraum V_{II} von V, so dass die durch $(u, v) \mapsto u + v$ gegebene lineare Abbildung $V_{\mathrm{I}} \times V_{\mathrm{II}} \to V$ ein Isomorphismus ist. Indem wir f und g mit diesem Isomorphismus verketten, erhalten wir Abbildungen mit den selben Differenzierbarkeitseigenschaften. Wir können also annehmen, dass V selbst die Form $V_{\mathrm{I}} \times V_{\mathrm{II}}$ hat, also für $u \in V_{\mathrm{I}}$ und $v \in V_{\mathrm{II}}$ gilt

$$df(c)(u, v) = \partial_{\mathrm{I}} f(c)(u) + \partial_{\mathrm{II}} f(c)(v), \quad dg(c)(u, v) = \partial_{\mathrm{I}} g(c)(u) + \partial_{\mathrm{II}} g(c)(v).$$

Nach Definition von V_{I} ist $\partial_{\mathrm{I}} g(c)(u) = 0$ für alle u, während das partielle Differential $\partial_{\mathrm{II}} g(c) : V_{\mathrm{II}} \to W$ den Kern $\{0\}$ hat und somit injektiv ist. Da es ebenso wie $dg(c)$ nach Voraussetzung surjektiv ist, ist es umkehrbar.

Schreiben wir $c = (a, b)$, so existiert nach Satz 11.2 eine Umgebung D_{I} von a sowie eine stetig differenzierbare Abbildung $h : D_{\mathrm{I}} \to V_{\mathrm{II}}$, so dass $h(a) = b$ und

$$\{(x, h(x)) \mid x \in D_{\mathrm{I}}\} \subseteq M.$$

Durch implizite Differentiation ergibt sich

$$dh(a) = -\partial_{\mathrm{II}} g(c)^{-1} \partial_{\mathrm{I}} g(c) = 0.$$

Nun sei $(u, v) \in \operatorname{Ker} dg(c)$. Dann ist $v = 0$, und die Menge

$$D_u = \{t \in \mathbb{R} \mid a + tu \in D_{\mathrm{I}}\}$$

ist nach Satz 8.10 eine Umgebung der Null. Setzen wir

$$f_u(t) = f(a + tu, h(a + tu)),$$

so erhalten wir eine Abbildung $f_u : D_u \to \mathbb{R}$, wobei das Argument auf der rechten Seite in M liegt und für $t = 0$ gleich c ist. Nach Satz 10.2 ist f_u an der Stelle $t = 0$ differenzierbar, und

$$f_u'(0) = \partial_{\mathrm{I}} f(c)(u) + \partial_{\mathrm{II}} f(c)\big(dh(a)(u)\big) = \partial_{\mathrm{I}} f(c)(u).$$

Außerdem hat f_u an der Stelle $t = 0$ ein lokales Extremum, und mit Satz 6.4 folgt $f_u'(0) = 0$. Es ergibt sich

$$df(c)(u, 0) = \partial_{\mathrm{I}} f(c)(u) = 0.$$

Da $(u, 0) \in \operatorname{Ker} dg(c)$ beliebig war, haben wir bewiesen, dass $df(c)$ auf $\operatorname{Ker} dg(c)$ verschwindet. □

Um stationäre Punkte unter Nebenbedingungen zu finden, benutzt man folgende Aussage aus der linearen Algebra.

Satz 11.4. *Es seien V und W Vektorräume über K sowie $l : V \to K$ und $m : V \to W$ lineare Abbildungen, wobei m surjektiv ist. Es ist genau dann $\operatorname{Ker} m \subseteq \operatorname{Ker} l$, wenn es eine lineare Abbildung $\lambda : W \to K$ mit der Eigenschaft $l = \lambda \circ m$ gibt.*

Beweis. Es sei $\operatorname{Ker} m \subseteq \operatorname{Ker} l$. Für jedes $w \in W$ gibt es wegen der Surjektivität von m ein $v \in V$, so dass $m(v) = w$. Wir behaupten, dass $l(v)$ nur von w abhängt. Gilt nämlich für $v' \in V$ ebenfalls $m(v') = w$, so ist $m(v - v') = 0$, also $l(v - v') = 0$ und schließlich $l(v) = l(v')$. Wir setzen $\lambda(w) = l(v)$. Verfahren wir so für alle w, erhalten wir eine Abbildung λ mit der Eigenschaft $l = \lambda \circ m$.

Sind w_1, $w_2 \in W$ gegeben, so können wir v_1, $v_2 \in V$ mit der Eigenschaft $m(v_1) = w_1$, $m(v_2) = w_2$ wählen, und dann gilt $m(v_1 + v_2) = w_1 + w_2$, also

$$\lambda(w_1 + w_2) = l(v_1 + v_2) = l(v_1) + l(v_2) = \lambda(w_1) + \lambda(w_2).$$

Analog beweist man, dass für $w \in W$ und $t \in K$ gilt $\lambda(tw) = t\lambda(w)$, und somit ist λ linear.

Die Umkehrung ist offensichtlich. □

Ist $f : D \to \mathbb{R}$ differenzierbar, $g : D \to W$ stetig differenzierbar und $dg(x)$ für alle $x \in D$ surjektiv, so findet man die stationären Punkte von f unter der Nebenbedingung $g(x) = 0$, indem man das Gleichungssystem

$$df(a) = \lambda \circ dg(x), \qquad g(x) = 0$$

mit den Unbekannten $x \in D$ und $\lambda \in \operatorname{Hom}(W, \mathbb{R})$ löst. Wegen $d\lambda(y) = \lambda$ für alle $y \in W$ hat die erste Gleichung die äquivalente Form

$$d(f - \lambda \circ g)(x) = 0.$$

Im Spezialfall $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$ gibt es für jedes $\lambda \in \operatorname{Hom}(\mathbb{R}^m, \mathbb{R})$ Vorlesung 13b 20.01.23

Zahlen $\lambda_1, \ldots, \lambda_m$, so dass für jeden Vektor $w = (w_1, \ldots, w_m) \in \mathbb{R}^m$ gilt

$$\lambda(w) = \lambda_1 w_1 + \ldots + \lambda_m w_m.$$

Wenn wir die Koordinaten von x mit x_i und die Koordinatenfunktionen von g mit g_j bezeichnen, so erscheint unser Gleichungssystem in der Form

$$\begin{aligned}
\delta_1 f(x_1, \ldots, x_n) &= \lambda_1 \delta_1 g_1(x_1, \ldots, x_n) + \ldots + \lambda_m \delta_1 g_m(x_1, \ldots, x_n)x, \\
&\vdots \\
\delta_n f(x_1, \ldots, x_n) &= \lambda_1 \delta_n g_1(x_1, \ldots, x_n) + \ldots + \lambda_m \delta_n g_m(x_1, \ldots, x_n), \\
g_1(x_1, \ldots, x_n) &= 0, \\
&\vdots \\
g_m(x_1, \ldots, x_n) &= 0
\end{aligned}$$

mit $m + n$ Unbekannten und ebenso vielen Gleichungen. Für jede Lösung $(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m)$ ist $(x_1, \ldots, x_n)$ ein stationärer Punkt unter den Nebenbedingungen $g_1(x_1, \ldots, x_n) = 0, \ldots, g_n(x_1, \ldots, x_n) = 0$. Die Hilfsgrößen $\lambda_1, \ldots, \lambda_m$ nennt man *Lagrange-Multiplikatoren*.

Beispiel. Wir suchen die Punkte in

$$M = \{(x, y) \in \mathbb{R}^2 \mid x^4 + y^4 = 4xy + 8\},$$

die den größten bzw. kleinsten Abstand vom Koordinatenursprung haben (vgl. T. Bröcker, Analysis II, Aufgabe II.15). Dazu setzen wir

$$f(x, y) = x^2 + y^2, \qquad g(x, y) = x^4 + y^4 - 4xy - 8.$$

Es gilt

$$\begin{aligned}
\frac{\partial f}{\partial x} &= 2x, & \frac{\partial g}{\partial x} &= 4x^3 - 4y, \\
\frac{\partial f}{\partial y} &= 2y, & \frac{\partial g}{\partial y} &= 4y^3 - 4x,
\end{aligned}$$

und $dg(x, y)$ ist nur dann nicht surjektiv, wenn

$$x^3 = y, \qquad y^3 = x.$$

Dies impliziert $x^9 = x$, also $(x, y) = (0, 0)$ oder $(x, y) = \pm(1, 1)$. Diese Punkte liegen aber nicht in M. Wir erhalten das Gleichungssystem

$$\begin{aligned}
\lambda(4x^3 - 4y) &= 2x, \\
\lambda(4y^3 - 4x) &= 2y, \\
x^4 + y^4 &= 4xy + 8.
\end{aligned}$$

Ist $\lambda = 0$, so ist $(x, y) = (0, 0)$, und die dritte Gleichung ist nicht erfüllt. Multiplizieren wir die erste Gleichung mit y und die zweite mit x, so erhalten wir nach Subtraktion und Kürzen von 4λ

$$(x^3 - y)y - (y^3 - x)x = 0,$$

also

$$(x + y)(x - y)(xy + 1) = 0.$$

Ist $x + y = 0$, so erhalten wir aus der dritten Gleichung

$$2(x^2)^2 + 4x^2 - 8 = 0,$$

also

$$x^2 = \sqrt{5} - 1.$$

Ist $x - y = 0$, so folgt hingegen

$$2(x^2)^2 - 4x^2 - 8 = 0,$$

also

$$x^2 = \sqrt{5} + 1.$$

Ist schließlich $xy + 1 = 0$, so folgt

$$x^4 + x^{-4} - 4 = 0,$$

also

$$(x^4)^2 - 4x^4 + 1 = 0$$

und somit

$$x^4 = 2 \pm \sqrt{3}.$$

Die stationären Punkte unter der Nebenbedingung $g(x, y) = 0$ sind also

$$\begin{array}{ll} \pm\left(\sqrt{\sqrt{5}+1}, \sqrt{\sqrt{5}+1}\right), & \pm\left(\sqrt{\sqrt{5}-1}, \sqrt{\sqrt{5}-1}\right), \\ \left(\pm\sqrt[4]{2+\sqrt{3}}, \mp\sqrt[4]{2-\sqrt{3}}\right), & \left(\pm\sqrt[4]{2-\sqrt{3}}, \mp\sqrt[4]{2+\sqrt{3}}\right). \end{array}$$

Wir behaupten, dass M beschränkt ist. Für $(x, y) \in M$ gilt nämlich

$$8 = x^4 + y^4 - 4xy \geq x^4 + y^4 - 2x^2 - 2y^2,$$

und wegen $y^4 - 2y^2 = (y^2 - 1)^2 - 1 \geq -1$ folgt $x^4 - 2x^2 \leq 9$. Für $x \geq 2$ ist also

$$9 \geq x^4 \left(1 - \frac{2}{x^2}\right) \geq \frac{x^4}{2},$$

d. h. $|x| \leq \sqrt{3}\sqrt[4]{2}$ und analog $|y| \leq \sqrt{3}\sqrt[4]{2}$. Außerdem ist M abgeschlossen, also kompakt, und somit besitzt die Einschränkung von f auf M ein Maximum und ein Minimum. Die Werte von f in den stationären Punkten sind

$$2(\sqrt{5}+1), \qquad 2(\sqrt{5}-1), \qquad \sqrt{2+\sqrt{3}}+\sqrt{2-\sqrt{3}},$$

und ihre Quadrate sind

$$4(6+\sqrt{5}), \qquad 4(6-\sqrt{5}), \qquad 6.$$

Wegen $\sqrt{20} < \sqrt{81}$, d. h. $2\sqrt{5} < 9$, ist $6 < 4(6-\sqrt{5})$. Der größte Abstand ist also $\sqrt{2+2\sqrt{5}}$, der kleinste ist $\sqrt[4]{6}$. $\triangleleft$

11.4 Untermannigfaltigkeiten affiner Räume

Wir wollen nun die Struktur der Lösungen nichtlinearer Gleichungen näher untersuchen. Den natürlichen Rahmen für unsere Betrachtungen bieten die in Definition 8.5 eingeführten affine Räume. Wir erinnern daran, dass zu jedem affinen Raum X ein Vektorraum V von Translationen gehört. Unter der Dimension von X verstehen wir die Dimension von V. Man kann Vektorräume und affine Räume über einem beliebigen Körper K betrachten, aber für die Zwecke der Differentialrechnung setzen wir $K = \mathbb{R}$ oder $K = \mathbb{C}$. Wir können dann V mit einer Norm und X mit der zugehörigen Metrik versehen.

Indem man Translationen auf einen beliebig gewählten festen Punkt o anwendet (den man dann den Ursprung nennt), erhält man eine bijektive Abbildung $V \to X$, wobei die Umkehrabbildung einem Punkt x seinen Ortsvektor $\overrightarrow{ox}$ zuordnet. Wer will, kann daher den Begriff des affinen Raumes ignorieren und einfach $X = V$ setzen.

Was man aber nicht ignorieren kann, ist der Begriff einer affinen Abbildung. Darunter versteht man eine Abbildung f von einem affinen Raum X in einen affinen Raum Y, zu der es eine lineare Abbildung $g : V \to W$ zwischen den zugehörigen Vektorräumen gibt, so dass für alle $x \in X$ und $v \in V$ gilt

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

(Affine Abbildungen $\mathbb{R} \to \mathbb{R}$ werden in der Schule lineare Abbildungen genannt, wenngleich sie keine linearen Abbildungen im Sinne der linearen Algebra sind.)

Eine Teilmenge Z von X heißt affiner Unterraum, wenn es einen linearen Unterraum U von V gibt, so dass für jeden Punkt $z \in Z$ und jeden Vektor $u \in U$ auch $z + u$ in Z ist und auf diese Weise Z zu einem affinen Raum mit dem zugehörigen Vektorraum U wird. Das Urbild eines Punktes unter einer affinen Abbildung ist ein affiner Unterraum des Definitionsbereichs, und der Wertebereich einer affinen Abbildung ist ein affiner Unterraum des Zielbereichs. Im Spezialfall $V = K^n$ und $W = K^m$ ergibt sich aus der ersteren Aussage, dass die Lösungsmenge eines linearen Gleichungssystems ein affiner Unterraum ist. In diesem Sinne ist die lineare Algebra also recht eintönig.

Die Lösungsmengen nichtlinearer Gleichungssysteme hingegen haben mannigfaltige Formen. Um eine analoge Aussage zu machen, führt man folgenden Begriff ein.

Definition 11.1. *Es sei $K \in \{\mathbb{R}, \mathbb{C}\}$ und $k \in \mathbb{N}$. Eine Teilmenge M eines affinen Raumes X heißt* Untermannigfaltigkeit *von X der Klasse C^k, wenn es für jeden Punkt x von M einen Diffeomorphismus F der Klasse C^k von einer Umgebung U von x in X auf eine offene Teilmenge U' eines affinen Raumes Y sowie einen affinen Unterraum Z von Y gibt, so dass*

$$F(M \cap U) = Z \cap U'.$$

Beispiel. Es sei E ein euklidischer Raum, also ein affiner Raum mit euklidischer Metrik. Wir betrachten den Konfigurationsraum der Dreiecke in E mit den Kantenlängen a, b und c, also

$$M_3 = \{(x, y, z) \in E^3 \mid d(y, z) = a,\ d(z, x) = b,\ d(x, y) = c\}.$$

Natürlich ist M_3 nur dann nichtleer, wenn a, b und c nicht negativ sind und die Dreiecksungleichungen

$$a + b \geq c, \qquad b + c \geq a, \qquad c + a \geq b$$

gelten. Wir werden später sehen, dass M_3 eine Untermannigfaltigkeit von $X_3 = E^3$ der Klasse C^∞ ist. Es ist jedenfalls klar, dass M_3 das Urbild des Punktes (a^2, b^2, c^2) unter der durch

$$g_3(x, y, z) = \big(d(y, z)^2, d(z, x)^2, d(x, y)^2\big)$$

definierten Abbbildung $g_3 : X_3 \to \mathbb{R}^3$ ist. Wir haben die Abstände quadriert, damit g unendlich oft differenzierbar ist. Führt man Koordinaten ein, so wird M_3 zur Lösungsmenge eines Gleichungssystems. $\triangleleft$

Die Lösungsmenge eines nichtlinearen Gleichungssystems ist nicht immer eine Untermannigfaltigkeit. Um Bedingungen prägnant formulieren zu können, führen wir geeignete Begriffe ein.

Definition 11.2. *Es seien X und Y affine Räume und D eine offene Teilmenge von X.*

(i) *Eine differenzierbare Abbildung $f : D \to Y$ heißt* Immersion, *wenn an jeder Stelle $x \in D$ das Differential $df(x)$ injektiv ist.*

(ii) *Eine differenzierbare Abbildung $g : D \to Y$ heißt* Submersion, *wenn an jeder Stelle $x \in D$ das Differential $dg(x)$ surjektiv ist.*

(iii) *Eine Abbildung f von einem metrischen Raum D in einen metrischen Raum Y heißt* Einbettung, *wenn die Beschränkung von f zu einer Abbildung $D \to f(D)$ ein Homöomorphismus ist.*

Man beachte, dass eine stetige injektive Abbildung keine Einbettung zu sein braucht, wie man am Beispiel der durch

Vorlesung 14a
25.01.23

$$f(t) = (\sin t, \sin t \cos t)$$

definierten Abbildung $f :]-\pi, \pi[\to \mathbb{R}^2$ sieht. Ihr Wertebereich ist gleich der Lösungsmenge der Gleichung

$$y^2 = x^2(1 - x^2),$$

aber die Umkehrabbildung ist an der Stelle $(0, 0)$ unstetig.

Satz 11.5. *Es seien X und Y affine Räume und D eine offene Teilmenge von X sowie $k \geq 1$ eine natürliche Zahl.*

(i) *Ist $g : D \to Y$ eine k-mal stetig differenzierbare Submersion, so ist für jedes Element b von Y das Urbild $g^{-1}(b)$ eine Untermannigfaltigkeit von X der Klasse C^k.*

(ii) *Ist $f : D \to Y$ eine k-mal stetig differenzierbare Immersion, die gleichzeitig eine Einbettung ist, so ist $f(D)$ eine Untermannigfaltigkeit von Y der Klasse C^k.*

Beweis. O. B. d. A. sind $X = V$ und $Y = W$ Vektorräume und ist $b = 0$.
(i) Es sei $g : D \to W$ eine k-mal stetig differenzierbare Submersion, $M = g^{-1}(0)$ und c ein Punkt von M. Im Beweis von Satz 11.3 haben wir gesehen, dass man dann V als Produkt zweier Vektorräume V_{I} und V_{II} ansehen kann, wobei $V_{\mathrm{I}} = \operatorname{Ker} dg(c)$ ist, und dass wir uns in der Situation von Satz 11.2 befinden, wo g mit G bezeichnet wurde. Im Beweis wurde gezeigt, dass durch

$$F(x, y) = (x, g(x, y))$$

ein Diffeomorphismus F der Klasse C^k von einer Umgebung U von c auf eine offene Teilmenge des Raumes $V_{\mathrm{I}} \times W$ gegeben ist, der die Rolle von Y aus der

Definition einer Untermannigfaltigkeit spielt. Es ist offensichtlich, dass genau dann $F(x,y) \in V_{\mathrm{I}} \times \{0\}$ ist, wenn $g(x,y) = 0$ gilt. Wir brauchen also nur $Z = V_{\mathrm{I}} \times \{0\}$ zu setzen. Da c beliebig war, ist M eine Untermannigfaltigkeit der Klasse C^k.

(ii) Nun sei $f : D \to W$ eine k-mal stetig differenzierbare Immersion, $N = f(D)$ und c ein Punkt von N. Wir wählen $a \in D$ mit $f(a) = c$. Diesmal können wir W als Produkt zweier Unterräume W_{I} und W_{II} ansehen, wobei W_{I} der Wertebereich von $df(a)$ ist. Wir definieren eine k-mal stetig differenzierbare Abbildung $H : D \times W_{\mathrm{II}} \to W$ durch

$$H(x,z) = f(x) + (0,z).$$

Dann ist

$$dH(x,z)(v,u) = df(x)(v) + (0,u)$$

und insbesondere

$$dH(a,0) = (df_{\mathrm{I}}(a), \mathrm{id}).$$

Diese lineare Abbildung ist invertierbar, und nach Satz 11.1 schränkt sich H zu einem Diffeomorphismus der Klasse C^k von einer Umgebung U_1 von $(a,0)$ auf eine Umgebung U_2 von c ein. Wegen $H(x,0) = f(x)$ gilt

$$N \cap U_2 \supseteq H((V \times \{0\}) \cap U_1),$$

und bezeichnen wir die Umkehrabbildung von H mit F, so folgt

$$F(N \cap U_2) \supseteq (V \times \{0\}) \cap U_1,$$

aber i. Allg. nicht die gewünschte Gleichheit. Die Menge $H((V \times \{0\}) \cap U_1)$ ist das Bild einer offenen Menge unter f, also eine offene Teilmenge des metrischen Raumes N, wenn f eine Einbettung ist. Nach Satz 8.5(i) ist sie von der Form $N \cap U$ für eine offene Menge U in W, wobei o. B. d. A. $U \subseteq U_2$ ist.

Die Einschränkung von F auf die Umgebung U von c ist ein Diffeomorphismus auf $U' = F(U)$ mit der Eigenschaft

$$F(N \cap U) = (V \times \{0\}) \cap U_1 = (V \times \{0\}) \cap U'.$$

Da c beliebig war, ist N eine Untermannigfaltigkeit der Klasse C^k. □

Beispiel. Wir betrachten wieder den Konfigurationsaum M_3 der Dreiecke mit vorgegebenen Seitenlängen im euklidischen Raum E. Er ist das Urbild eines Punktes unter der Abbildung $g_3 : E^3 \to \mathbb{R}^3$, die auch in der Form

$$g_3(x,y,z) = \left(\|\overrightarrow{yz}\|^2, \|\overrightarrow{zx}\|^2, \|\overrightarrow{xy}\|^2\right)$$

geschrieben werden kann. Die Abbildung

$$h(x,y) = \|\overrightarrow{xy}\|^2 = \langle \overrightarrow{xy}, \overrightarrow{xy} \rangle$$

ist eine Verkettung einer bilinearen Abbildung mit affinen Abbildungen vom affinen Raum E in den zugehörigen Vektorraum, für die gilt

$$\overrightarrow{x(y+v)} = \overrightarrow{xy} + v,$$

so dass nach den Ableitungsregeln folgt

$$dh(x,y)(u,v) = \langle v, \overrightarrow{xy} \rangle + \langle \overrightarrow{xy}, v \rangle - \langle u, \overrightarrow{xy} \rangle - \langle \overrightarrow{xy}, u \rangle = 2\langle \overrightarrow{xy}, v-u \rangle.$$

Es ergibt sich

$$dg_3(x,y,z)(u,v,w) = 2\big(\langle \overrightarrow{yz}, w-v \rangle, \langle \overrightarrow{zx}, u-w \rangle, \langle \overrightarrow{xy}, v-u \rangle\big).$$

Diese lineare Abbildung kann nicht surjektiv sein, wenn $x = y$ oder $y = z$ oder $z = x$ ist. Nehmen wir also an, die Punkte seien paarweise verschieden. Mit den Bezeichnungen $s = w - v$ und $t = u - w$ wird die rechte Seite zu

$$2\big(\langle \overrightarrow{yz}, s \rangle, \langle \overrightarrow{zx}, t \rangle, \langle \overrightarrow{yz} + \overrightarrow{zx}, s+t \rangle\big).$$

Dies ist gleich einem vorgegebenen Wert $(p,q,r) \in \mathbb{R}^3$, wenn

$$2\langle \overrightarrow{yz}, s \rangle = p, \quad 2\langle \overrightarrow{zx}, t \rangle = q, \quad 2\langle \overrightarrow{yz}, t \rangle + 2\langle \overrightarrow{zx}, s \rangle = r.$$

Durch die ersten beiden Gleichungen ist s bis auf einen zu $\overrightarrow{yz}$ orthogonalen Vektor festgelegt und t bis auf einen zu $\overrightarrow{zx}$ orthogonalen Vektor. Sind $\overrightarrow{yz}$ und $\overrightarrow{zx}$ nicht zueinander proportional, so kann man s so anpassen, dass $\langle \overrightarrow{zx}, s \rangle$ einen beliebig vorgegebenen Wert hat. Andernfalls ist die linke Seite der dritten Gleichung festgelegt.

Wir sehen also, dass dg_3 in einem Punkt (x,y,z) genau dann surjektiv ist, wenn die Punkte x, y, z nicht kollinear sind, d. h. nicht auf einer Geraden liegen. Schränken wir also g_3 auf die Menge $D_3 \subset E^3$ solcher Tripel ein, so ist Satz 11.5 anwendbar und zeigt, dass M_3 im Fall

$$a + b > c, \qquad b + c > a, \qquad c + a > b$$

eine Untermannigfaltigkeit der Klasse C^∞ ist. Sind die Punkte kollinear, aber nicht alle gleich, so hängt für vorgegebene a, b und c der dritte Punkt auf unendlich oft differenzierbare Weise von den zwei anderen ab. Tritt also bei einer der Dreiecksungleichungen Gleichheit ein, so ist M_3 isomorph zur analogen Untermannigfaltigkeit M_2 von $X_2 = E^2$, also auch eine Untermannigfaltigkeit. Es bleibt der Fall $a = b = c = 0$, in dem $M_3 = \{(x,x,x) \mid x \in E\}$ ebenfalls eine Untermannigfaltigkeit ist, die dann isomorph zu $M_1 = X_1 = E$ ist. All dies lässt sich auf eine beliebige Anzahl von Punkten verallgemeinern. $\triangleleft$

11.5 Tangentialvektoren

Vorlesung 14b
27.01.23

Die erste Einführung in die Differentialrechnung wird meist durch die Anschauung unterstützt. Dabei wird der Differenzenquotient als Anstieg einer Sekante des Graphen der betrachteten Funktion interpretiert und der Differentialquotient, also die Ableitung, als Anstieg der Tangente. Bis jetzt haben wir die letztere aber immer noch nicht definiert. Das soll nun in einem allgemeineren Rahmen erfolgen. Dabei mögen alle Mannigfaltigkeiten mindestens von der Klasse C^1 sein, der Grundkörper gleich $\mathbb{R}$ oder $\mathbb{C}$.

Definition 11.3. *Es sei X ein affiner Raum mit zugehörigem Vektorraum V und M eine Untermannigfaltigkeit von X. Einen Vektor $v \in V$ nennen wir einen* Tangentialvektor *von M an der Stelle $a \in M$, wenn es eine differenzierbare Abbildung h einer offenen Menge I von $\mathbb{R}$ in die Menge M und eine Stelle $c \in I$ gibt, so dass*

$$h(c) = a, \qquad h'(c) = v.$$

Die Menge aller Tangentialvektoren von M an der Stelle a nennen wir den Tangentialraum *von M an dieser Stelle, abgekürzt $T_a(M)$.*

Genauer gesagt handelt es sich hier um den *linearen* Tangentialraum. Manchmal betrachtet man auch den *affinen* Tangentialraum

$$a + T_a(M) = \{a + v \mid v \in T_a(M)\} \subseteq X.$$

Ist insbesondere M eine Kurve, so nennt man den affinen Tangentialraum einfach *Tangente*[59]. Man kann o. B. d. A annehmen, dass I ein Intervall ist, und h als differenzierbaren Weg betrachten.

Neben dem Kern einer linearen Abbildung $l : V \to W$ zwischen Vektorräumen betrachtet man auch ihr Bild, abgekürzt[60] $\operatorname{Im} l = l(V)$.

Satz 11.6. *Es sei M eine Untermannigfaltigkeit des affinen Raums X mit zugehörigem Vektorraum V sowie $a \in M$.*

(i) *Der lineare Tangentialraum von M an der Stelle a ist ein Untervektorraum von V, der affine Tangentialraum ist ein affiner Unterraum von X.*

(ii) *Ist g eine Submersion einer offenen Teilmenge U von X in einen affinen Raum Y und $M = g^{-1}(b)$ für ein $b \in Y$, so ist*

$$T_a(M) = \operatorname{Ker} dg(a).$$

[59] lat. „Berührende“
[60] engl. *image*

(iii) *Ist f eine Immersion und Einbettung einer offenen Teilmenge D eines affinen Raums X in den Raum Y und $N = f(D)$, wobei $b = f(a)$, so ist*
$$T_b(N) = \operatorname{Im} df(a).$$

Beweis. **(i)** Entsprechend Definition 11.1 gibt es einen Diffeomorphismus F einer Umgebung D von a auf eine offene Teilmenge D' eines affinen Raums Y sowie einen affinen Unterraum Z von Y, so dass

$$F(M \cap D) = Z \cap D'.$$

Dann ist D' eine Umgebung von $b = F(a)$. Es seien V, W und U die zu X, Y und Z gehörigen Vektorräume, und es sei H die Umkehrabbildung von F. Wir behaupten, dass
$$T_a(M) = dH(b)(U),$$
was natürlich ein Unterraum von V ist. Für jeden Vektor $u \in U$ ist nämlich $h(t) = H(b + tu)$ für t in einer offenen Teilmenge I von $\mathbb{R}$ definiert, stetig differenzierbar, und

$$h(0) = a, \qquad h'(0) = dH(b)(u),$$

also $dH(b)(u) \in T_a(M)$. Ist umgekehrt $v \in T_a(M)$, so gibt es eine Abbildung $h : I \to M$ mit $h(c) = a$ und $h'(c) = v$. Dann ist die Abbildung

$$k = F \circ h : h^{-1}(D) \to Z$$

definiert, und für $u = k'(c) \in U$ gilt nach Satz 10.2(ii)

$$u = dF(a)(h'(c)) = dF(a)(v), \quad \text{also} \quad dH(b)(u) = v.$$

(ii) Da jeder Weg in M von g auf einen konstanten Weg abgebildet wird, ist zumindest $T_a(M) \subseteq \operatorname{Ker} dg(a)$. Nun sei o. B. d. A. $X = V$ und $Y = W$. Wir können g wie im Beweis von Satz 11.2 lokal zu einem Diffeomorphismus $F : D \to V_{\mathrm{I}} \times W$ ergänzen, wobei $V_{\mathrm{I}} = \operatorname{Ker} dg(a)$, und die Umkehrabbildung H von F erfüllt $dH(b)(v, 0) = v$ für $v \in V_{\mathrm{I}}$. Nun folgt die Behauptung aus dem Beweis von Teil (i).
(iii) Da jeder Weg in D von f auf einen Weg in N abgebildet sird, ist zumindest $T_b(N) \supseteq \operatorname{Im} df(a)$. Nun sei wieder $X = V$, $Y = W$. Wir können f wie im Beweis von Satz 11.5(ii) lokal zu einem Diffeomorphismus $H : D \times W_{\mathrm{II}} \to W$ ergänzen, wobei $H(x, 0) = f(x)$ ist. Für $v \in V$ ist somit $dH(a, 0)(v, 0) = df(a)(v)$, und die Behauptung folgt wieder aus dem Beweis von Teil (i). □

Mit dem Begriff des Tangentialraums werden ein paar frühere Definitionen geometrisch verständlicher:

- In der Definition 10.3 der Differenzierbarkeit einer Funktion $f : D \to W$ an einer Stelle a haben wir sie durch eine affine Funktion

$$g(a+v) = f(a) + df(a)(v)$$

approximiert. Ist f stetig differenzierbar, so ist der Graph von g gleich dem affinen Tangentialraum des Graphen

$$M = \{(x, f(x)) \mid x \in D\}$$

von f an der Stelle $(a, f(a))$ (Übungsaufgabe).

- Ein stationärer Punkt a einer Funktion $f : D \to \mathbb{R}$ unter der Nebenbedingung $g(x) = 0$ (siehe Satz 11.3) ist in dem Fall, wenn g eine Submersion und somit $M = g^{-1}(0)$ eine Untermannigfaltigkeit ist, charakterisiert durch die Bedingung

$$df(a)|_{T_a(M)} = 0.$$

Manchmal ist der umgebende Raum X nur ein Hilfsobjekt ohne Bedeutung für die Anwendungen. Wir gehen jetzt den ersten Schritt eines Weges, auf dem man sich von ihm befreien kann.

Definition 11.4. *Es sei M eine Untermannigfaltigkeit der Klasse C^k des affinen Raums X mit zugehörigem Vektorraum V. Eine Abbildung f von M in einen affinen Raum Y heißt k-mal stetig differenzierbar, wenn es für jeden Punkt a von M eine Umgebung U von a in X und eine k-mal stetig differenzierbare Abbildung $g : U \to Y$ gibt, so dass*

$$f|_{M\cap U} = g|_{M\cap U}.$$

Lemma 11.2. *Sind $f : M \to Y$ und $g : U \to Y$ wie in der Definition und ist W der zu Y gehörige Vektorraum, so wird durch*

$$df(a)(v) = dg(a)(v)$$

eine lineare Abbildung

$$df(a) : T_a(M) \to W$$

definiert, die nicht von der Wahl von U und g abhängt.

Beweis. Als Einschränkung der linearen Abbildung $dg(a) : V \to W$ ist auch $df(a)$ linear. Für einen beliebigen Tangentialvektor $v \in T_a(M)$ gibt es laut Definition eine offene Teilmenge I von $\mathbb{R}$, einen Punkt $c \in I$ und eine Abbildung $h : I \to M$, so dass

$$h(c) = a, \qquad h'(c) = v.$$

Nach Satz 10.2(ii) gilt

$$dg(a)(h'(c)) = (g \circ h)'(c),$$

also

$$df(a)(v) = (f \circ h)'(c),$$

was offensichtlich nicht von U und g abhängt. □

Man nennt $df(a)$ das Differential von f an der Stelle a. Hier ist eine erste Anwendung.

Satz 11.7. *Ist M eine Untermannigfaltigkeit der Klasse C^1 eines affinen Raums X und $f : M \to \mathbb{R}$ eine stetig differenzierbare Abbildung, die an einer Stlle a ein lokales Extremum hat, so gilt*

$$df(a) = 0.$$

Beweis. Ist h wie im Beweis des Lemmas, so ist c ein lokales Extremum von $f \circ h$, und die Behauptung folgt aus Satz 6.4. □

Man kann differenzierbare Mannigfaltigkeiten definieren, ohne einen umgebenden Raum zu verwenden. Das ist eines der Themen der Varanstaltung „Geometrie/Topologie“.