



UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

SKRIPT ZUR VORLESUNG
HÖHERE MATHEMATIK II
FÜR DAS
STUDIUM DER PHYSIK

gehalten an der
Universität Heidelberg
im
Sommersemester 2016
von
Johannes Walcher

INHALTSVERZEICHNIS

1	EINFÜHRUNG	3
§ 1	Die reellen Zahlen	3
§ 2	Die komplexen Zahlen	11
2	KONVERGENZ	15
§ 3	Folgen reeller Zahlen	15
§ 4	Metrische Räume	23
§ 5	Reihen	33
3	STETIGKEIT	43
§ 6	Stetige Funktionen	43
§ 7	Wertsätze	47
§ 8	Gleichmässigkeit	51
4	DIFFERENTIATION	55
§ 9	Eine reelle Veränderliche	55
§ 10	Mehrere Veränderliche	60
§ 11	Komplex differenzierbare Funktionen	68
5	UMKEHRUNG	73
§ 12	Gleichungen im $\mathbb{R}^n$	73
§ 13	Stammfunktionen	80
§ 14	Gewöhnliche Differentialgleichungen	88
6	FUNKTIONENTHEORIE	93
§ 15	Integration in $\mathbb{C}$	93
§ 16	Cauchy-Formeln	98
§ 17	Residuensatz	104

KAPITEL 1

EINFÜHRUNG

Diese Vorlesung richtet sich an Studierende der Physik im zweiten Semester. Das Hauptziel ist die Bereitstellung der für das weitere Studium der Physik wichtigsten Konzepte und Rechenmethoden der *Mathematischen Analysis*. Dabei baut die Vorlesung einerseits auf den im ersten Semester vermittelten Grundlagen der linearen Algebra auf (insbesondere werden die üblichen Grundgedanken zur mathematischen Beweisführung und Notation und zur elementaren Mengenlehre *nicht* im Detail besprochen), und stellt andererseits einen Bezug zu den theoretischen Physik-Vorlesungen her, indem einige der dort bereits verwendeten Begriffe und Ergebnisse hier zunächst noch einmal neu begründet, und anschliessend weiterentwickelt werden. (Dabei soll ganz und gar nicht die gewohnte Anschauung im Keime erstickt werden, sondern vielmehr die Wertschätzung und das Vertrauen in die Methoden der “anderen” Disziplin gestärkt werden.)

Für Kommentare, auch Fehlermeldungen, zum Skript, per Email an walcher@uni-heidelberg.de, bin ich sehr dankbar. Man beachte aber, dass dieses Skript keinen Anspruch auf Abgeschlossenheit erhebt: Viele Beweise werden bestenfalls angedeutet, mancher Begriff nur indirekt definiert.

· Zur Homepage der Vorlesung: www.mathi.uni-heidelberg.de/~walcher/teaching/sose16/hoemaphys/

§ 1 Die reellen Zahlen

Die reellen Zahlen bilden die Grundlage der Analysis, und damit der gesamten theoretischen Physik. Sie werden durch drei Gruppen von Eigenschaften charakterisiert.

Arithmetik

Definition 1.1. Ein *Körper* ist eine Menge $\mathbb{K}$ mit zwei ausgezeichneten und verschiedenen Elementen $0 \in \mathbb{K}$ und $1 \in \mathbb{K}$, $0 \neq 1$, zusammen mit zwei binären Verknüpfungen, welche wir auch “Rechenoperationen” nennen und wie gewöhnlich notieren:

$$\begin{aligned} \text{Addition: } \mathbb{K} \times \mathbb{K} \ni (a, b) &\mapsto a + b \in \mathbb{K} \\ \text{Multiplikation: } \mathbb{K} \times \mathbb{K} \ni (a, b) &\mapsto a \cdot b := ab \in \mathbb{K} \end{aligned} \tag{1.1}$$

und welche die folgenden Rechenregeln erfüllen:

$$\text{Kommutativität: } \forall a, b \in \mathbb{K} : a + b = b + a, \quad a \cdot b = b \cdot a$$

$$\text{Assoziativität: } \forall a, b, c \in \mathbb{K} : (a + b) + c = a + (b + c), \quad (a \cdot b) \cdot c = a \cdot (b \cdot c)$$

$$\text{Neutrale Elemente: } \forall a \in \mathbb{K} : a + 0 = a, \quad a \cdot 1 = a$$

$$\text{Distributivgesetz: } \forall a, b, c \in \mathbb{K} : a \cdot (b + c) = a \cdot b + a \cdot c \quad (1.2)$$

derart, dass die Gleichungen

$$a + x = 0, \quad b \cdot y = 1 \quad (1.3)$$

für alle $a \in \mathbb{K}$ und $b \in \mathbb{K} \setminus \{0\}$ Lösungen $x, y \in \mathbb{K}$ besitzen.

Lemma 1.2. (i) Die Lösungen der Gleichungen (1.3) sind eindeutig. Wir bezeichnen sie mit $-a$ bzw. $b^{-1} =: \frac{1}{b}$, und schreiben natürlich auch $b - a := b + (-a)$ etc..

(ii) Insbesondere sind 0 und 1 eindeutig.

(iii) Ausserdem impliziert $xy = 0$, dass mindestens $x = 0$ oder $y = 0$.

Beweis. (i) Aus $a + x = 0$ und $a + x' = 0$ folgt unter Anwendung nur der aufgeführten Regeln (1.2):

$$x = x + 0 = x + (a + x') = (x + a) + x' = (a + x) + x' = 0 + x' = x' \quad (1.4)$$

· Ein ähnliches Argument zeigt die Eindeutigkeit der Lösung von $by = 1$. (Die Existenz einer Lösung für $b = 0$ stünde im Widerspruch zu $0 \cdot y = 0 \neq 1 \quad \forall y \in \mathbb{K}$.)

(ii) Aus $a + 0' = a \quad \forall a$ folgt $0' = 0 + 0' = 0$ und ähnlich für die 1.

(iii) Ist $xy = 0$ und $x \neq 0$ so folgt $0 = x^{-1}(xy) = (x^{-1}x)y = 1 \cdot y = y$. □

Bemerkungen. · Die Motivation zur Forderung dieser Gesetze der Arithmetik sind zunächst einmal die Eigenschaften der natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \dots\}$, in denen man Assoziativität, Kommutativität, die Eigenschaft der 1, und Distributivität axiomatisch begründen kann, nach Geschmack auch mit der 0.

· Die Erweiterung von den natürlichen Zahlen $\mathbb{N}$ zu den ganzen Zahlen $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$, macht die erste der Gleichungen (1.3) lösbar, der Übergang zu den rationalen Zahlen $\mathbb{Q} = \{\frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}\}$ dann die zweite. Es ist bemerkenswert, dass dieser Prozess an dieser Stelle abgeschlossen werden kann: $\mathbb{Q}$ ist ein Körper und muss dazu nicht etwa noch weitererweitert werden.

· Es wäre dies jedoch nicht die einzige Möglichkeit. Anstatt zu erweitern kann man auch zu Restklassen übergehen: Für jede Primzahl p ist $\mathbb{F}_p := \mathbb{Z}/(p\mathbb{Z})$ ein endlicher Körper.

· Die Bezeichnung “Körper” geht auf R. Dedekind (1831-1916) zurück, der englische Begriff “field” auf E. H. Moore (1862-1932).

· In den Naturwissenschaften dienen Körper zur abstrakten Buchhaltung konkret ausgeführter endlicher Messoperationen. Genauer gesagt enthalten die Bildmengen physikalischer Theorien in der Regel¹ Vektorräume, in denen physikalische Grössen

¹und in einer gewissen Näherung. Der Leitgedanke dieser Vorlesung ist es, den Begriff der Näherung mathematisch genauer zu fassen.

§ 1. DIE REELLEN ZAHLEN

mit Zahlen multipliziert und untereinander addiert werden können, während die Multiplikation zweier physikalischer Größen miteinander von anderer Natur ist. (Das Produkt zweier Längen ist z.B. gar keine Länge!) Man vergleiche dies mit dem operationellen Erlernen der Grundrechenarten.

• Eine wichtige Eigenschaft von Messprozessen ist die quantitative Vergleichbarkeit physikalischer Größen. Im mathematischen Bild führt man diese Vergleiche (eventuell mit weiteren Zwischenschritten, siehe unten!) auf den folgenden Begriff zurück.

Anordnung

Definition 1.3. Ein *angeordneter Körper* ist ein Körper $\mathbb{K}$ zusammen mit einer Teilmenge P (den “positiven” Elementen) derart, dass

$$\begin{aligned}\mathbb{K} &= P \cup \{0\} \cup (-P) \\ P + P &\subset P \\ P \cdot P &\subset P\end{aligned}\tag{1.5}$$

Wir schreiben auch $a > 0$ oder $0 < a$ für $a \in P$ und $a < 0$ oder $0 > a$ für $a \in -P$ ($\Leftrightarrow -a \in P$).

Lemma 1.4. (i) Durch die Vorschrift $a > b : \Leftrightarrow a - b \in P$ wird auf einem angeordneten Körper eine strenge Totalordnung definiert (transitiv und trichotomisch), bzw. durch $a \geq b : \Leftrightarrow (a > b \text{ oder } a = b)$ eine Totalordnung (transitiv, antisymmetrisch und total) und es gelten die üblichen (zum Teil empfindlichen) Rechenregeln für Ungleichungen, wie z.B.

$$\begin{aligned}\text{(ii) Aus } a > b \text{ folgen } &\begin{cases} \frac{1}{a} < \frac{1}{b} & \text{falls } b > 0 \\ a + c > b + c & \forall c \in \mathbb{K} \\ ac \geq bc & \text{je nach dem } c \geq 0 \end{cases} \\ \text{(iii) Aus } a > b \text{ und } c > d \text{ folgen } &\begin{cases} a + c > b + d & \text{in jedem Fall} \\ ac > bd & \text{falls } b > 0 \text{ und } d > 0 \end{cases} \\ \text{(iv) Für } a \neq 0 \text{ gilt } a^2 > 0. \text{ Insbesondere ist} &\end{aligned}$$

$$1 > 0\tag{1.6}$$

Beweis. (i) Die Trichotomie: $\forall a, b \in \mathbb{K}$ gilt entweder $a > b$ oder $a = b$ oder $b > a$ folgt aus der dreiteiligen Zerlegung in (1.5). Die Transitivität sehen wir so:

$$\begin{aligned}a > b \text{ und } b > c &\Leftrightarrow a - b \in P \text{ und } b - c \in P \Rightarrow \\ &(\text{wegen } P + P \subset P) \Rightarrow (a - b) + (b - c) = a - c \in P \Leftrightarrow a > c\end{aligned}\tag{1.7}$$

(ii), (iii) Übungsaufgaben

(iv) Falls $a > 0$ so folgt $a^2 > 0$ direkt aus $P \cdot P \subset P$. Falls $a < 0$ so ist per Definition $-a > 0$. Wegen $(-1) \cdot a + a = (-1 + 1) \cdot a = 0$ gilt $-a = (-1) \cdot a$ und wegen $(-a)^2 - a^2 = (-a)^2 + (-a) \cdot a = (-a) \cdot (-a + a) = 0$ gilt $a^2 = (-a)^2 > 0$, wieder wegen $P \cdot P \subset P$. $\square$

Bemerkungen. · Die Forderung nach Anordnung schliesst endliche Körper wie $\mathbb{F}_p$ aus: In der Tat enthält jeder angeordnete Körper die natürlichen Zahlen, d.h. die eindeutige strukturerhaltende Abbildung $\phi : \mathbb{N} \rightarrow \mathbb{K}$ (mit $\mathbb{N} \ni 1 \mapsto \phi(1) := 1 \in \mathbb{K}$, $\mathbb{N} \ni 2 \mapsto \phi(2) := 1 + 1 \in \mathbb{K}$ etc.) ist injektiv: Wegen (1.6) (in $\mathbb{K}$) und den Anordnungsaxiomen (1.5) gilt auch $\phi(n) > 0$, also $\phi(n) \neq 0 \forall n \in \mathbb{N}$ (vollständige Induktion). Wir schreiben natürlich $\phi(n) =: n \in \mathbb{K}$. Ebenso enthält jeder angeordnete Körper die ganzen Zahlen und die rationalen Zahlen als angeordneten Unterkörper. · Für Mathematiker wäre es aber, wie schon bei den Rechenregeln zuvor, wichtig zu betonen, dass nicht alle möglichen angeordneten Körper alle von den rationalen Zahlen her gewohnten Eigenschaften besitzen.

Ein einfaches Hilfsmittel:

Lemma 1.5 (Bernoullische Ungleichung). *Für jedes $x \geq -1$ in einem angeordneten Körper $\mathbb{K}$ gilt für alle $n \in \mathbb{N}$:*

$$(1+x)^n \geq 1+nx \quad (1.8)$$

Beweis. Auf der rechten Seite ist natürlich $(1+x)^n$ durch Rekursion definiert, und daher beweisen wir die Aussage auch durch vollständige Induktion. Für $n = 1$ ist die Aussage klar, der Schluss von n auf $n+1$ ergibt sich wegen $1+x \geq 0$ so:

$$(1+x)^{n+1} \geq (1+nx)(1+x) = 1+(n+1)x+nx^2 \geq 1+(n+1)x \quad (1.9)$$

□

Eine Sprachregelung:

Definition 1.6. Sei $\mathbb{K}$ ein angeordneter Körper. Eine Teilmenge $T \subset \mathbb{K}$ heisst *nach oben (unten) beschränkt*, falls eine *obere (untere) Schranke* von T existiert, d.h. ein $B \in \mathbb{K}$ s.d.

$$\forall x \in T : B \geq x \quad (B \leq x) \quad (1.10)$$

Wir sagen “beschränkt” für eine Teilmenge, die sowohl nach unten als auch nach oben beschränkt ist.

Zur Betonung bemerken wir, dass falls T (durch B) nach oben beschränkt ist, so ist $-T$ (durch $-B$) nach unten beschränkt. Insbesondere sind “untere Schranken” im Sinne der Anordnung nicht notwendig “klein” im Sinne des Absolutbetrags: Setzen wir für $a \in \mathbb{K}$

$$|a| := \max\{a, -a\} = \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0 \end{cases} \quad (1.11)$$

so gelten die Regeln:

Lemma 1.7.

$$\begin{aligned} \text{Multiplikativität: } & \forall a, b \in \mathbb{K} : |a \cdot b| = |a| \cdot |b|, \\ \text{Dreiecksungleichung: } & \forall a, b \in \mathbb{K} : |a+b| \leq |a| + |b|, \\ \text{“Umgekehrte” Dreieckungl.: } & \forall a, b \in \mathbb{K} : |a-b| \geq ||a| - |b||, \\ \text{ausserdem: } & |a| = 0 \Leftrightarrow a = 0 \end{aligned} \quad (1.12)$$

§ 1. DIE REELLEN ZAHLEN

Beweis. Die erste Regel verifiziert man leicht anhand einer Fallunterscheidung. Die Dreiecksungleichung folgt aus $|a + b| = \max\{a + b, -(a + b)\}$ und $a \leq |a|$, $b \leq |b|$, die Umkehrung aus $|a| = |a - b + b| \leq |a - b| + |b|$ und $|b| \leq |a - b| + |a|$. $\square$

Für die Beweise der Analysis, sowie ihre Anwendungen in der Physik sind Abschätzungen über den Absolutbetrag von zentraler Bedeutung. Den feinen Unterschied zum Größenvergleich mittels Anordnung werden wir noch zu schätzen lernen. In der Physik ist eine wichtigerweise (im Sinne der Anordnung) nach unten beschränkte Größe die Gesamtenergie eines physikalischen Systems. Andernfalls droht Instabilität! Die vielleicht bekannteste kleinste physikalische Größe ist das Wirkungsquantum.

Definition 1.8. Ein angeordneter Körper $\mathbb{K}$ heisst *Archimedisch* falls $\mathbb{N} \subset \mathbb{K}$ nicht nach oben beschränkt ist. M.a.W.

$$\forall x \in \mathbb{K} : \exists N \in \mathbb{N} \text{ s.d. } N > x \quad (1.13)$$

Bemerkungen. · Der Körper der rationalen Zahlen $\mathbb{Q}$ ist Archimedisch.

· Die Konstruktion von angeordneten nicht-Archimedischen Körpern leuchtet keinem Anfänger sofort ein, weshalb die Wichtigkeit dieses Archimedischen Axioms von allen hier gegebenen Definitionen am schwierigsten einzusehen ist.

· Physikalisch gesehen ist die Unbeschränktheit der natürlichen Zahlen aber sehr natürlich (für jeden einzelnen Physiker etwa symbolisiert durch den Herzschlag, für die Menschheit durch die Revolution der Erde um die Sonne).

· Das Archimedische Axiom dient dazu, dass wir unsere Zahlenbereiche nicht zu weit erweitern, indem es Anordnungsunendlichkeiten, welche grösser sind als die natürlichen Zahlen, ausschliesst. (Wie wir wissen und evtl. sehen werden, sind allerdings im Sinne der Mächtigkeit von Mengen die reellen Zahlen eine grössere Unendlichkeit als die natürlichen...)

Proposition 1.9. Sei $\mathbb{K}$ ein Archimedisch angeordneter Körper, und $q, Q \in \mathbb{K}$.

(i) Ist $Q > 1$ so gibt es zu jedem $B \in \mathbb{K}$ ein $n \in \mathbb{N}$ so, dass $Q^n > B$.

(ii) Ist $0 < q < 1$ so gibt es zu jedem $\epsilon > 0$ ein $n \in \mathbb{N}$ so, dass $q^n < \epsilon$.

Beweis. (i) Wir können in eindeutiger Weise schreiben $Q = 1 + x$ mit $x > 0$. Die Bernoullische Ungleichung 1.5 liefert $Q^n \geq 1 + nx$. Weil $\mathbb{K}$ Archimedisch ist, existiert ein $n \in \mathbb{N}$ mit $nx > B$. Damit gilt $Q^n > B$.

(ii) Folgt aus (i) durch $Q = q^{-1}$ und $B = \epsilon^{-1}$. $\square$

Vollständigkeit

Bevor wir nun zur entscheidenden Charakterisierung der reellen Zahlen kommen, gehen wir noch einmal zurück zur Erweiterung der natürlichen Zahlen zu den rationalen Zahlen durch "Hinzufügen" der Lösungen von (1.3). Es ist leicht einzusehen, dass die Lösbarkeit dieser Gleichungen äquivalent ist zur Aussage:

$$\forall (a', b') \in \mathbb{K} \times \mathbb{K}^* : \exists x : b' \cdot x + a' = 0 \quad (1.14)$$

In der Linearen Algebra lernt man Vektorräume über Körpern als die natürlichen Kategorien kennen, in denen sich eine sinnvolle Theorie der linearen Gleichungen

als Verallgemeinerung von (1.14) und ihrer Lösungen entwickeln lässt. (Die Vektorraumstruktur bei physikalischen Messgrößen passt damit in wunderbarer Weise zur Idee, dass Lineare Algebra als mathematische Begleitvorlesung zur Theoretischen Physik I ausreicht!)

· Andererseits haben, wie hinreichend bekannt, gewisse natürlich auftretende nicht-lineare Aufgaben wie z.B.

$$x^2 = n, \quad n \in \mathbb{N} \setminus \{m^2 \mid m \in \mathbb{N}\} \quad (1.15)$$

keine Lösung in $\mathbb{Q}$.²

· Eine Möglichkeit zur Abhilfe ist die weitere Erweiterung der rationalen Zahlen um die fehlenden Lösungen von Gleichungen wie (1.15) und deren Verwandten höheren Grades. Diese Option ist vom mathematischen Standpunkt her ökonomisch und wird von grossen Teilen jener Zunft bevorzugt.

· In dieser Vorlesung folgen wir hingegen der Mutmassung, dass es für die Physik ja ausreichen könnte, Gleichungen wie (1.15) *näherungsweise* zu lösen (auch wenn die physikalische Bedeutung solcher Gleichungen vielleicht zunächst gar nicht klar ist...)

· Die Idee, dass man solche Näherungen “*im Prinzip* beliebig verbessern können muss”, führt ausgehend von den rationalen direkt zu den reellen Zahlen $\mathbb{R}$. Sei z.B. $x_0 \in \mathbb{Q}$ eine “näherungsweise” Lösung der Gleichung $x^2 = 2$, d.h. $\delta_0 := x_0^2 - 2 \in \mathbb{Q}$ sei “klein” in einem geeigneten Sinne. Der Einfachheit halber nehmen wir an, dass $x_0 > 0$ und $x_0^2 > 2$, also auch $\delta_0 > 0$. Es ist leicht zu sehen, dass solch ein x_0 existiert.) Definieren wir dann

$$x_1 := x_0 - \frac{\delta_0}{2x_0} \quad (1.16)$$

so gilt

(i) $x_1 > 0$ ($\Leftrightarrow 0 < 2x_0^2 - \delta_0 = x_0^2 + 2 \checkmark$)

(ii)

$$\delta_1 := x_1^2 - 2 = \left(x_0 - \frac{\delta_0}{2x_0}\right)^2 - 2 = x_0^2 - 2 - \delta_0 + \left(\frac{\delta_0}{2x_0}\right)^2 = \frac{\delta_0}{x_0^2} \cdot \frac{\delta_0}{4} \quad (1.17)$$

Wegen $\delta_0 = x_0^2 - 2 < x_0^2$ ist damit $\delta_1 < \frac{\delta_0}{4}$ und daher x_1 eine (noch) “bessere” approximative Lösung als x_0 .

Wegen (i) und $\delta_1 > 0$ lässt sich diese Prozedur wiederholen. Allerdings ist $x_1 \in \mathbb{Q}$ und keine noch so lange Iteration kann eine exakte Lösung liefern, da ja kein $x \in \mathbb{Q}$ existiert mit $x^2 = 2$. Was tun?

· Die reellen Zahlen entstehen durch “Füllen aller Lücken”, die durch solche Iterationen in den rationalen Zahlen aufgetan werden. Das weitere Ziel der Analysis ist dann die Identifikation reeller Zahlen, die durch solche Prozesse auseinander hervorgehen, und die Untersuchung von deren Eigenschaften. Die ursprüngliche “Vollständigkeit der reellen Zahlen” lässt sich auf mindestens drei verschiedene äquivalente

²Zur Erinnerung: Jede rationale Zahl lässt sich in eindeutiger Weise als $x = \frac{p}{q}$ darstellen mit $(p, q) \in \mathbb{Z} \times \mathbb{N}$ teilerfremd. Dann sind auch p^2 und q^2 teilerfremd und $\frac{p^2}{q^2}$ die Darstellung von x^2 . Falls $\frac{p^2}{q^2} = n \in \mathbb{N}$ so ist $q = 1$ und $n = p^2 \in \{m^2 \mid m \in \mathbb{N}\}$.

§ 1. DIE REELLEN ZAHLEN

Weisen erfassen. Zwei davon resonieren besonders gut mit der physikalischen Idee der Approximation, und wir werden sie im nächsten Kapitel ausführlich besprechen. Am elegantesten zum Ziel führt jedoch die folgende Formulierung, welche ohne das Archimedische Axiom und ohne Folgen auskommt.

Definition 1.10. Sei $\mathbb{K}$ ein angeordneter Körper, und $T \subset \mathbb{K}$ eine Teilmenge. Eine Zahl $B \in \mathbb{K}$ heisst *kleinste obere Schranke* falls

- (i) B eine obere Schranke ist und
- (ii) Für jede andere obere Schranke B' von T gilt, dass $B' \geq B$.

Ein angeordneter Körper $\mathbb{K}$ heisst (*ordnungs-*)*vollständig*, falls jede nicht-leere nach oben beschränkte Menge eine kleinste obere Schranke besitzt.

(Die analoge Definition über grösste untere Schranken ist äquivalent.)

Theorem 1.11. (i) *Es existiert ein bis auf Isomorphismus eindeutig bestimmter vollständiger angeordneter Körper, genannt der Körper der reellen Zahlen $\mathbb{R}$.*

(ii) *$\mathbb{R}$ ist Archimedisch.*

(iii) *Jeder angeordnete Archimedische Körper ist in $\mathbb{R}$ enthalten.*

Einige Beweisideen. (i) Wir illustrieren die Konstruktion von $\mathbb{R}$ am obigen Beispiel: x_0 ist eine obere Schranke der nichtleeren Teilmenge $A_{x^2=2} := \{x \mid x < 0 \text{ oder } x^2 < 2\} \subset \mathbb{Q}$. Aber kein solches x_0 ist eine kleinste obere Schranke (denn x_1 ist ja wegen $\delta_0 > 0$ kleiner als x_0). Um die Existenz einer kleinsten oberen Schranke als Lösung der gestellten Aufgabe sicherzustellen, definiert man einen *Dedekindschen Schnitt* als eine Teilmenge $A \subset \mathbb{Q}$ mit den Eigenschaften, dass

- (α) $A \neq \emptyset$ und $A \neq \mathbb{Q}$
- (β) A ist nach unten ordnungs-abgeschlossen, d.h. $\forall x < y : y \in A \Rightarrow x \in A$ und
- (γ) A enthält kein grösstes Element: $\forall x \in A : \exists y \in A : y > x$

Zu solch einer Teilmenge gehört in ein-eindeutiger Weise eine Teilmenge $B := \mathbb{Q} \setminus A$ mit den Eigenschaften:

- (α') $B \neq \mathbb{Q}$ und $A \neq \emptyset$
- (β) B ist nach oben ordnungs-abgeschlossen.
- (γ') Falls B eine grösste untere Schranke hat, so liegt diese in B .

Beispiele für solche Schnitte sind für $y \in \mathbb{Q}$: $A_y = \{x \mid x < y\}$. Die wesentliche Beobachtung ist aber, dass es Schnitte von $\mathbb{Q}$ gibt, die nicht von der Form A_y für ein $y \in \mathbb{Q}$ sind.

Es erfordert dann etwas Arbeit, auf der Menge der Dedekindschen Schnitte Rechenoperationen zu definieren und ihre Eigenschaften zu verifizieren, aber das Endergebnis ist tatsächlich ein vollständiger angeordneter Körper. Insbesondere ist bemerkenswert, dass wir durch das Schliessen der Lücken in $\mathbb{Q}$ keine weiteren Lücken aufgetan haben. Die zum Schnitt $A_{x^2=2}$ gehörige reelle Zahl $\sqrt{2}$ erfüllt dabei $\sqrt{2} > 0$ und $(\sqrt{2})^2 = 2$.

(ii) Beachte, dass man für die Definition von “archimedisch” ($\forall x \exists N > x$) braucht, dass jeder angeordnete Körper die ganzen Zahlen enthält.

Angenommen, es gäbe ein $x \in \mathbb{K}$ so, dass $x \geq N \forall N \in \mathbb{N}$. Dann wäre x eine obere Schranke für $\mathbb{N}$, und es gäbe eine kleinste obere Schranke, x_0 . $\Rightarrow x_0 > N \forall N \in \mathbb{N}$. $\Rightarrow x_0 - 1 > N \forall N \in \mathbb{N}$. Dann aber wäre $x_0 - 1 < x_0$ ebenfalls eine obere Schranke, kleiner als x_0 .

(iii) Siehe Ethan Bloch, Real Numbers and Real Analysis für eine sehr benutzerfreundliche und ausführliche Behandlung dieses Zugangs zu den reellen Zahlen. $\square$

Für das praktische Rechnen sind die Dedekindschen Schnitte recht unhandlich. Im Folgenden setzen wir die Existenz der reellen Zahlen mit den oben genannten Eigenschaften *voraus*, und formulieren die Vollständigkeit weiter unten auf so viele verschiedenen Weisen um, wie es für unsere Zwecke nützlich ist.

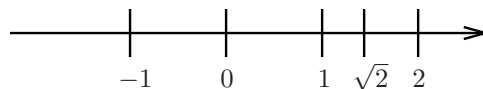
- Es ist letztlich dem Genie Isaac Newtons die Idee zu verdanken, dass die idealisierte Einführung der reellen Zahlen es nicht nur erlaubt, (gewisse, aber bei weitem (noch) nicht alle, siehe nächster Abschnitt) algebraische Gleichungen zu lösen, sondern auch, Differentialgleichungen zu formulieren, deren Lösungen (“Integrale”) die Bewegungen von Massenpunkten, die Ausbreitung von Wellen und klassischen Feldern, und dergleichen mehr, in einer geeigneten Näherung hervorragend beschreiben.
- Das Ziel dieser Vorlesung ist der theoretische Nachweis der Existenz solcher Lösungen in diesem Rahmen sowie die Entwicklung geeigneter Lösungsmethoden.

Definition 1.12. · Für eine nach oben beschränkte nichtleere Teilmenge $T \subset \mathbb{R}$ nennen wir die kleinste obere Schranke von T das Supremum von T ($\sup T$). Falls $\sup T \in T$, so ist $\sup T = \max T$, dem Maximum von T .

· Für eine nach unten beschränkte nichtleere Teilmenge $T \subset \mathbb{R}$ nennen wir die größte untere Schranke das Infimum von T ($\inf T$). Falls $\inf T \in T$, so ist $\inf T = \min T$, dem Minimum von T .

Visualisierung

Es ist zweckmässig, sich den Bereich der reellen Zahlen mit seiner Anordnung bildlich darzustellen wie aus der Grundschule bekannt. Dabei werden insbesondere Addition, Anordnung und in gewissem Sinne die Vollständigkeit erfasst, die Multiplikation nur durch Anwendung eines Rechenschiebers.



Definition 1.13. Besonders interessant und als Definitionsbereiche von Funktionen wichtige Teilmengen von $\mathbb{R}$ sind die Intervalle. Für $a < b$ unterscheiden wir:

abgeschlossen: $[a, b] := \{x | x \geq a, b \geq x\}$

“halboffen”: $\begin{cases} [a, b) := \{x | x \geq a, b > x\} \\ (a, b] := \{x | x > a, b \geq x\} \end{cases}$

offen: $(a, b) := \{x | x > a, b > x\}$

Beachte: Die “unbeschränkten Intervalle” $[a, \infty)$, $(-\infty, a]$ gelten als abgeschlossen, (a, ∞) , $(-\infty, a)$ als offen. Die “Grenze” ∞ ($\notin \mathbb{R}!$) ist also sowohl offen als auch abgeschlossen.³

³Man beachte auch, dass dem Wort “abgeschlossen” hier ein recht anderer Sinn gegeben wird als in (β).

§ 2. DIE KOMPLEXEN ZAHLEN

Proposition 1.14. · Zu jeder reellen Zahl $a > 0$ und jeder natürlichen Zahl $n \in \mathbb{N}$ gibt es genau eine reelle Zahl $x > 0$ mit $x^n = a$ (Wir schreiben auch $x = a^{1/n}$.)
· Es gilt $|x| = (x^2)^{1/2} \forall x > 0$.

§ 2 Die komplexen Zahlen

Lemma/Definition 2.1. Die Menge aller Paare $(x, y) \in \mathbb{R} \times \mathbb{R}$ reeller Zahlen, versehen mit den Operationen:

$$\text{Addition: } (x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2)$$

$$\text{Multiplication: } (x_1, y_1) \cdot (x_2, y_2) := (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1)$$

und der Auszeichnung $0_{\mathbb{C}} = (0, 0)$ und $1_{\mathbb{C}} = (1, 0)$, ist ein Körper, genannt der Körper der komplexen Zahlen $\mathbb{C}$

Bemerkungen. · Von den Körperaxiomen ist nur die Existenz des multiplikativen Inversen etwas nicht-trivial: Für $(x, y) \neq (0, 0)$ ist $x^2 + y^2 > 0$, daher

$$(x, y) \cdot \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right) = (1, 0) = 1_{\mathbb{C}} \quad (2.1)$$

Gegebenenfalls überprüfe man auch alle anderen Körperaxiome als Übung.

· Die Begründung für die Einführung der komplexen Zahlen liegt in der Möglichkeit der Lösung algebraischer Gleichungen, die in $\mathbb{Q}$ aus *Anordnungsgründen* nicht lösbar sind, und zwar nicht einmal näherungsweise (sodass das Problem auch nicht durch Übergang zu $\mathbb{R}$ gelöst wird). So gilt ja z.B. für $\mathbb{K}$ einen angeordneten Körper $x^2 \geq 0 \forall x \in \mathbb{K}$, und daraus folgt $x^2 + 1 \geq 1$, mithin hat die Gleichung $z^2 + 1 = 0$ keine Lösung in $\mathbb{K}$.

· In $\mathbb{C}$ hingegen gilt

$$(0, 1)^2 = (-1, 0) = -1_{\mathbb{C}} \quad (2.2)$$

und wie wir später sehen werden, hat sogar jede nicht-konstante monisch polynomiale Gleichung mit Koeffizienten in $\mathbb{C}$, also

$$z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0 = 0 \quad a_i \in \mathbb{C}, n \in \mathbb{N} \quad (2.3)$$

eine Lösung $z \in \mathbb{C}$ (Mit Multiplizität gezählt gibt es dann n Lösungen.) (Fundamentalsatz der Algebra.) Diese *algebraische Abgeschlossenheit von $\mathbb{C}$* ist auch für dessen zentrale Bedeutung in der Quantenmechanik verantwortlich.

· Hingegen ist es als Folgerung aus den obigen Betrachtungen unmöglich, $\mathbb{C}$ in einer Weise anzuordnen, die mit den Körperoperationen verträglich ist. Schreibt man im komplexen Kontext eine Ungleichung wie $B > 0$ so meint dies " $B \in \mathbb{R} \subset \mathbb{C}$ und positiv im Sinner der Anordnung auf $\mathbb{R}$."

· Definieren wir nun $i := (0, 1) \in \mathbb{C}$ und denken uns $\mathbb{R}$ als in $\mathbb{C}$ eingebettet via

$$\mathbb{R} \ni x \mapsto (x, 0) \in \mathbb{C} \quad (2.4)$$

(d.h., wir identifizieren $x \in \mathbb{R}$ mit $(x, 0) \in \mathbb{C}$...) dann können wir jede komplexe Zahl in eindeutiger Weise schreiben als

$$z = (x, y) = x + iy \quad x, y \in \mathbb{R} \subset \mathbb{C} \quad (2.5)$$

und “wie gewöhnlich” rechnen.

· M.a.W. ist $\mathbb{C}$ ein 2-dimensionaler Vektorraum über $\mathbb{R}$ mit Basis $1_{\mathbb{C}} = (1, 0)$ und $i = (0, 1)$. Allerdings wollen wir uns möglichst bald daran gewöhnen, komplexe Zahlen als eine Einheit aufzufassen.

· “Statt” der Anordnung haben wir die Operation der komplexen Konjugation auf $\mathbb{C}$:

$$x + iy = z \mapsto \bar{z} = x - iy \quad (2.6)$$

Sie erfüllt Rechenregeln wie $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$, $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$, $\bar{\bar{z}} = z$. Real- und Imaginärteil sind

$$\operatorname{Re}(z) := \frac{z + \bar{z}}{2} = x, \quad \operatorname{Im}(z) := \frac{z - \bar{z}}{2i} = y \quad (2.7)$$

und es gilt $\bar{z}z = x^2 + y^2 > 0$ für $z \neq 0$. Wir definieren den *Absolutbetrag*

$$|z| := \begin{cases} (\bar{z}z)^{1/2} & z \neq 0 \\ 0 & z = 0 \end{cases} \quad (2.8)$$

welcher wegen Prop. 1.14 auf $\mathbb{R}$ mit (1.11) übereinstimmt. Es gelten $|z| = 0 \Leftrightarrow z = 0$, Abschätzungen der Art

$$|\operatorname{Re}(z)| \leq |z| \quad |\operatorname{Im}(z)| \leq |z| \quad (2.9)$$

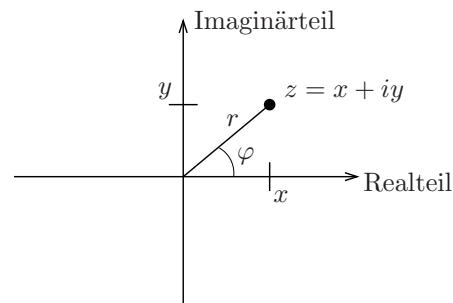
sowie insbesondere die Dreiecksungleichung

$$|z + w| \leq |z| + |w| \quad (2.10)$$

(vgl. Lemma 1.7).

Visualisierung

Die Eigenschaften dieser Definitionen erlauben es, sich $\mathbb{C}$ (als Ergänzung zu unserer algebraischen Absicht eben doch) als die Euklidische Ebene vorzustellen, in der die Addition von komplexen Zahlen genau der Addition von Vektoren entspricht.



· In diesem Bild entspricht komplexe Konjugation der Spiegelung an der reellen Achse, und der Absolutbetrag ist der Abstand vom Ursprung.

· Schreiben wir dann für $z \neq 0$

$$z = r \left(\underbrace{\frac{x}{r}}_{=:c} + i \underbrace{\frac{y}{r}}_{=:s} \right) \quad (2.11)$$

und nehme die aus der Elementargeometrie bekannte Tatsache vorweg, dass es für zwei reelle Zahlen c und s mit $c^2 + s^2 = 1$ genau einen Winkel $\varphi \in [0, 2\pi)$ gibt mit $c = \cos \varphi$ und $s = \sin \varphi$, so sehen wir, dass $z = r(\cos \varphi + i \sin \varphi)$. Mithilfe der

§ 2. DIE KOMPLEXEN ZAHLEN

Additionstheoreme für $\sin$ und $\cos$ folgt leicht, dass Multiplikation mit z geometrisch interpretiert werden kann als Drehung um den Winkel φ , gefolgt von Streckung um den Faktor r .

· Daraus folgt dann z.B. auch die Formel von deMoivre

$$z^n = r^n(\cos(n\varphi) + i \sin(n\varphi)) \quad (2.12)$$

sowie die $\mathbb{C}$ -Version von Prop. 1.14, einem Spezialfall von (2.3).

Proposition 2.2. Für jede komplexe Zahl $a \neq 0$ und jede natürliche Zahl n existieren genau n Lösungen der Gleichung

$$z^n = a$$

Beweis. Für $a = r(\cos \varphi + i \sin \varphi)$ sind dies

$$z_k = r^{1/n} \left(\cos \left(\frac{\varphi}{n} + \frac{2\pi k}{n} \right) + i \sin \left(\frac{\varphi}{n} + \frac{2\pi k}{n} \right) \right) \quad (2.13)$$

mit $k = 0, 1, \dots, n-1$. Dass dies alle sind, folgt aus der aus der Algebra bekannten Tatsache, dass ein Polynom vom Grad n höchstens n verschiedene Wurzeln besitzt. $\square$

· Wir bemerken zum weiteren Vergleich mit $\mathbb{R}$ noch, dass wir die Vollständigkeit von $\mathbb{C}$ (die wir intuitiv aus der geometrischen Darstellung als $\mathbb{R} \times \mathbb{R}$ natürlich schon jetzt erfassen) wegen der fehlenden Anordnung *nicht* über die Supremumseigenschaft ausdrücken können, sondern erst über eine der im folgenden § gegebenen Formulierungen. (Für den Beweis des Fundamentalsatzes der Algebra spielt dann die Vollständigkeit eine wesentliche Rolle.)

· Obwohl also in der komplexen Welt Aussagen wie “ $z > w$ ” im Allgemeinen keinen Sinn machen (vielmehr impliziert man für gewöhnlich mit so einer Schreibweise, dass $z, w \in \mathbb{R} \subset \mathbb{C}$), so können wir dennoch vereinbaren, dass

Definition 2.3. Eine Teilmenge $T \subset \mathbb{C}$ heisst beschränkt, falls eine $B \in \mathbb{R}$ existiert mit

$$|z| < B \quad \forall z \in T$$

· Zuletzt erwähnen wir noch als Analogon der Intervalle als “elementare” Definitionsbereiche von Funktionen im Komplexen die Kreisscheiben vom Radius $R > 0$ um den Punkt $z \in \mathbb{C}$:

$$D_R(z) := \{w \in \mathbb{C} \mid |z - w| < R\} \quad (2.14)$$

$\uparrow$
 Beachte hier $\neq$. Dies entspricht
 den offenen Intervallen.

KAPITEL 2

KONVERGENZ

In der Analysis formuliert und löst man mathematische Probleme (die wir uns als aus der Physik stammend denken) unter Ausnutzung der Vollständigkeit im Rechenbereich der reellen oder komplexen Zahlen anstatt mit expliziter Algebra oder durch Vorzeigen. Dazu ist allerdings, wie bereits angedeutet, unsere bisherige Formulierung via Supremumseigenschaft zu glatt (unter anderem macht sie ja im Komplexen keinen Sinn). In diesem Kapitel entwickeln wir zunächst aus der Supremumseigenschaft die Grundlagen der Grenzprozesse, die sich anschliessend als wesentlich allgemeiner und bequemer herausstellen. Konzeptionell handelt es sich aber zunächst nur um relativ banale Umformulierungen, wie etwa das folgende häufig eingesetzte Argument.

Lemma 3.-1. *Über eine reelle Zahl $x \in \mathbb{R}$ sei bekannt, dass für jedes $\epsilon > 0$, $x \in [-\epsilon, \epsilon]$ (m.a.W., $|x| \leq \epsilon$). Dann ist $x = 0$.*

Beweis. Angenommen $x > 0$. Dann gilt für $\epsilon_* := \frac{x}{2}$: $\epsilon_* > 0$ und $x > \epsilon_*$, d.h. $x \notin [-\epsilon_*, \epsilon_*]$, im Widerspruch zur Voraussetzung. Der Fall $x < 0$ wird analog ausgeschlossen. Es bleibt also nur $x = 0$. $\square$

§ 3 Folgen reeller Zahlen

Sei $X \neq \emptyset$ eine Menge. Unter einer Folge in X (oder auch mit Werten in X oder auch X -ige Folge) versteht man eine Abbildung

$$a : \mathbb{N} \rightarrow X \quad n \mapsto a(n) = a_n \in X \quad (3.1)$$

von der Menge der natürlichen Zahlen in X . Man schreibt auch $(a_n)_{n \in \mathbb{N}}$, $(a_n)_{n=1,2,\dots}$, $(a_n) \subset X$.

· Mit anderen Worten ist eine Folge eine Aufzählung von (wohlgemerkt nicht notwendigerweise verschiedenen) Elementen $a_1, a_2, \dots$ von X . Manche Aussagen über Folgen (wie etwa die Beschränktheit einer reellen Folge) sind als Aussagen über die Menge $\{a_n | n \in \mathbb{N}\} \subset X$ aufzufassen. Andere Aussagen betreffen die Abbildung (3.1).

Definition 3.1. (i) Eine Folge (x_n) reeller Zahlen heisst monoton wachsend, falls für alle $n, m \in \mathbb{N}$

$$n > m \Rightarrow x_n \geq x_m$$

(Offenbar ist dies gleichwertig zu $x_{n+1} \geq x_n \forall n \in \mathbb{N}$.)

(ii) (x_n) heisst streng monoton wachsend falls auf der rechten Seite $x_n \geq x_m$ (d.h.

$x_n > x_m$) steht. Sie heisst (streng) monoton fallend, falls $x_n \leq x_m$ (bzw. $x_n \not\leq x_m$).
 (iii) Unter einer Teilfolge einer Folge $(a_n)_{n \in \mathbb{N}} \subset X$ (für beliebiges X) versteht man eine Folge $(b_k)_{k \in \mathbb{N}}$ in X , welche aus (a_n) durch Angabe einer streng monoton wachsenden Folge $(n_k)_{k \in \mathbb{N}}$ natürlicher Zahlen via

$$b_k := a_{n_k} \quad (3.2)$$

gewonnen worden ist. (Das heisst, man zählt nur einen Teil der Folgenglieder von (a_n) ab, aber unter Berücksichtigung von deren ursprünglichen Reihenfolge.)

Die Idee ist nun, reelle Zahlen nicht explizit (wie etwa durch Dedekindsche Schnitte) anzugeben, sondern ihnen durch *sukzessive Approximation*, nämlich durch Folgen, allmählich immer näher zu kommen. Dies ist voll an der physikalischen Intuition ausgerichtet.

Definition 3.2. Eine Folge (x_n) reeller Zahlen heisst *konvergent* falls ein $x \in \mathbb{R}$ existiert mit der Eigenschaft, dass $\forall \epsilon > 0$ ein $N_\epsilon \in \mathbb{N}$ existiert, sodaß

$$\begin{aligned} & |x_n - x| < \epsilon \quad \forall n \geq N_\epsilon \\ \Leftrightarrow & x - x_n \in (-\epsilon, +\epsilon) \quad \forall n \geq N_\epsilon \\ \Leftrightarrow & x_n \in (x - \epsilon, x + \epsilon) \quad \forall n \geq N_\epsilon \end{aligned} \quad (3.3)$$

(Wir sagen auch: (x_n) konvergiert “gegen”, oder “mit Grenzwert” x .) Eine nicht-konvergente reelle Folge heisst “divergent”.

Bemerkungen. · Es ist oft nützlich, leichte Variationen der Definition als gleichwertig zu erkennen, wie z.B.: “falls ein $x \in \mathbb{R}$ und ein $c > 0$ existieren so, dass $\forall \epsilon > 0$ ein $N_\epsilon \in \mathbb{N}$ existiert so, dass

$$|x - x_n| \leq c \cdot \epsilon \quad \forall n > N_\epsilon” \quad (3.4)$$

· Falls (x_n) konvergiert, so ist der Grenzwert eindeutig:

Beh.: Falls (x_n) gegen x und y konvergiert, dann gilt $x = y$.

Bew.: Für $\epsilon > 0$ seien N_ϵ und M_ϵ so, dass

$$\begin{aligned} & |x_n - x| < \epsilon \quad \forall n \geq N_\epsilon \\ \text{und } & |x_n - y| < \epsilon \quad \forall n \geq M_\epsilon \end{aligned} \quad (3.5)$$

Dann gilt mit einem $n_0 \geq \max\{N_\epsilon, M_\epsilon\}$:

$$|x - y| \leq |x - x_{n_0}| + |x_{n_0} - y| < 2\epsilon \quad (3.6)$$

Wegen Lemma 3.1 folgt aus diesem Schluss $x = y$. □

· Für eine konvergente Folge macht also die Schreibweise

$$\lim_{n \rightarrow \infty} x_n = x \text{ oder auch } x_n \xrightarrow[n \rightarrow \infty]{} x \quad (3.7)$$

Sinn.

§3. FOLGEN REELLER ZAHLEN

Definition 3.3. Eine Nullfolge ist eine konvergente Folge (p.t. in $\mathbb{R}$) mit Grenzwert 0.

Lemma 3.4. Für eine reelle Folge (x_n) und $x \in \mathbb{R}$ sind die folgenden Aussagen äquivalent.

- (i) $\lim_{n \rightarrow \infty} x_n = x$
- (ii) $(x_n - x)_{n \in \mathbb{N}}$ ist eine Nullfolge.
- (iii) $(|x_n - x|)_{n \in \mathbb{N}}$ ist eine Nullfolge.

Beispiel 3.5. Wichtige Folgengrenzwerte:

- (i) Für jedes positive $s \in \mathbb{Q}$ ist $\lim_{n \rightarrow \infty} \frac{1}{n^s} = 0$
- (ii) Für $a > 0$ gilt $\lim_{n \rightarrow \infty} a^{1/n} = 1$
- (iii) $\lim_{n \rightarrow \infty} n^{1/n} = 1$
- (iv) Für $0 < q < 1$ ist $\lim_{n \rightarrow \infty} q^n = 0$
- (v) Für $k \in \mathbb{N}$ und $x > 1$ ist $\lim_{n \rightarrow \infty} \frac{n^k}{x^n} = 0$

Bew.: (i) Für $\epsilon > 0$ gilt mit $N_\epsilon > \frac{1}{\epsilon^{1/s}}$ (ein solches existiert nach dem Archimedisches Axiom) für $n \geq N_\epsilon$:

$$\frac{1}{n^s} \leq \frac{1}{N_\epsilon^s} < \left(\frac{1}{\epsilon^{1/s}} \right)^{-s} = \epsilon \quad (3.8)$$

(ii) Wir behandeln zunächst den Fall $a > 1$: Nach der Bernoullischen Ungleichung 1.5 gilt

$$\begin{aligned} a &= \left(1 + \underbrace{a^{1/n} - 1}_{>0} \right)^n \geq 1 + n(a^{1/n} - 1) \\ \Rightarrow \quad 0 &< a^{1/n} - 1 \leq \frac{a - 1}{n} \end{aligned} \quad (3.9)$$

Die Aussage folgt dann aus (i) und dem sogenannten Sandwich-Theorem:

Lemma 3.6. Für eine reelle Folge (x_n) seinen (A_n) und (B_n) konvergente Folge mit

$$\lim A_n = \lim B_n =: x \quad \text{und} \quad A_n \leq x_n \leq B_n \quad \forall n \quad (3.10)$$

Dann gilt $\lim_{n \rightarrow \infty} x_n = x$.

Beweis. Übungsaufgabe. □

Bemerkungen. · Es genügt wenn $x_n \in [A_n, B_n]$ für “fast alle n ”, d.h. $\exists N$ s.d. $x_n \in [A_n, B_n]$ für $n \geq N$.

• Aufmerksamen Leserinnen und Lesern sei an dieser Stelle noch (einmal) verraten, dass für die Abschätzungen in den meisten Beweisen die Unterscheidung zwischen $<$ und $\leq$ unerheblich ist. Auf der anderen Seite ergeht die *Warnung*, dass beim Grenzübergang selbst, Ungleichungen zu Gleichungen werden können. Beispielsweise gilt: Sind (A_n) und (B_n) konvergente Folgen mit $A_n < B_n$ für alle (oder auch “fast alle”) n , so folgt $\lim A_n \leq \lim B_n$, aber im Allgemeinen gilt nicht $<$!

Bevor mit den Beispielen fortfahren, halten wir die folgenden Rechenregeln fest:

Proposition 3.7. Für reelle Folgen (x_n) und (y_n) gelte $x_n \rightarrow x$ und $y_n \rightarrow y$. Dann gilt

$$(\alpha) \lim(x_n + y_n) = x + y$$

$$(\beta) \lim x_n \cdot y_n = x \cdot y$$

$$(\gamma) \lim |x_n| = |x|$$

$$(\delta) \text{ falls } y \neq 0, \text{ so sind fast alle } y_n \neq 0 \text{ und } \lim \frac{x_n}{y_n} = \frac{x}{y}$$

Beweis. siehe Lehrbücher. □

Bemerkungen. Natürlich gelten hier keine Umkehrungen. Insbesondere folgt aus $|x_n| \rightarrow |x|$ nicht $x_n \rightarrow x$.

Die Aussage des Beispiels (ii) im Falle $a < 1$ folgt nun aus dem bereits bewiesenen Fall $a > 1$ zusammen mit der Regel (δ) . (Für $a = 1$ ist die Aussage sowieso trivial.) Zum Nachweis von (iii) betrachten wir zunächst, wieder mit der Bernoullischen Ungleichung

$$\begin{aligned} n^{1/2} &= (1 + n^{1/2n} - 1)^n \geq 1 + n(n^{1/2n} - 1) \\ \Rightarrow 0 &\leq n^{1/2n} - 1 \leq \frac{n^{1/2} - 1}{n} = \frac{1}{n^{1/2}} - \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0 \end{aligned} \quad (3.11)$$

Also gilt $\lim_{n \rightarrow \infty} n^{1/2n} = 1$, und wegen Regel (β) folgt

$$\lim_{n \rightarrow \infty} n^{1/n} = \lim_{n \rightarrow \infty} (n^{1/2n} \cdot n^{1/2n}) = 1 \quad (3.12)$$

- Beispiel (iv) ist genau die Aussage von 1.9.
- Zum Beweis von (v) betrachten wir, schon wieder mit Bernoulli:

$$\begin{aligned} x^n &\geq 1 + n(x - 1) > n^{1/2} \cdot n^{1/2} \underbrace{(x - 1)}_{>0} \\ \Rightarrow \frac{n^{1/2}}{x^n} &< \frac{1}{n^{1/2}(x - 1)} \xrightarrow{n \rightarrow \infty} 0 \end{aligned} \quad (3.13)$$

Dann folgt die Aussage wieder mit (β) :

$$\frac{n^k}{x^n} = \left(\frac{n^{1/2}}{\underbrace{(x^{1/2k})^n}_{>1}} \right)^{2k} \xrightarrow{n \rightarrow \infty} 0 \quad (3.14)$$

· Alle diese Folgen (i)–(v), sowie die Rechenregeln, sind zwar sehr nützlich, illustrieren aber gerade *nicht* die wesentliche Idee, Folgen zum Herzeigen “neuer” reeller Zahlen (insbesondere solche in $\mathbb{R} \setminus \mathbb{Q}$) zu benutzen. Dafür müssen wir aus der Vollständigkeit Kriterien entwickeln, die die Konvergenz von gewissen Folgen garantieren, gerade auch ohne, dass wir den Grenzwert a priori explizit kennen. (Vielmehr wird dann die reelle Zahl als der Grenzwert der fraglichen Folge definiert.)

§ 3. FOLGEN REELLER ZAHLEN

· Das Paradebeispiel hierfür ist die rekursiv definierte Folge

$$x_{n+1} = x_n - \frac{x_n^2 - 2}{2x_n} = \frac{x_n}{2} + \frac{1}{x_n} \quad (3.15)$$

(vgl. (1.16)), welche für beliebigen Startwert $x_0 > \sqrt{2}$ die Eigenschaft hat, dass

$$\lim_{n \rightarrow \infty} x_n = \sqrt{2}, \quad (3.16)$$

was wir als Konsequenz aus unserem ersten grenzwertfreien Konvergenzkriterium bald beweisen.

Lemma 3.8. *Eine Folge $(x_n)_{n \in \mathbb{N}}$ konvergiert genau dann, wenn jede ihrer Teilfolgen $(x_{n_k})_{k \in \mathbb{N}}$ konvergiert. Es gilt dann*

$$\lim_{k \rightarrow \infty} x_{n_k} = \lim_{n \rightarrow \infty} x_n \quad (3.17)$$

für jede solche Teilfolge.

Beweis. trivial □

Theorem 3.9 (Prinzip der monotonen Konvergenz). *Jede monoton wachsende und nach oben beschränkte Folge reeller Zahlen ist konvergent.*

Bemerkungen. Ditto jede monoton fallende und nach unten beschränkte Folge. Wenn wir “bestimmte Divergenz” mit Grenzwert $\pm\infty$ zulassen, so kann die Beschränktheit fallengelassen werden.

Beweis. Sei (x_n) eine monoton wachsende nach oben beschränkte reelle Folge. Die Menge $T = \{x_n | n \in \mathbb{N}\} \subset \mathbb{R}$ ist dann nichtleer und nach oben beschränkt. Nach der Supremumseigenschaft existiert eine kleinste obere Schranke,

$$s := \sup\{x_n | n \in \mathbb{N}\} \quad (3.18)$$

Beh.: $\lim_{n \rightarrow \infty} x_n = s$

Bew.: Für jedes $\epsilon > 0$ ist $s - \epsilon < s$, daher keine obere Schranke für T . Es existiert also ein N_ϵ mit $x_{N_\epsilon} > s - \epsilon$. Wegen der Monotonie gilt dann für all $n \geq N_\epsilon$

$$s \geq x_n \geq x_{N_\epsilon} > s - \epsilon \quad (3.19)$$

und daraus folgt die Behauptung. □

Bew.: [von (3.16)] Aus den Betrachtungen um (1.17) folgt, dass (x_n) monoton fällt und nach unten beschränkt ist. Daher existiert jedenfalls $x = \lim x_n$, und nach den obigen Rechenregeln gilt

· $x > 0$ (denn $x_n > \sqrt{2} \forall n$)

· Die Folge (y_n) mit $y_n := x_{n+1}$ konvergiert als Teilfolge von (x_n) ebenfalls, mit dem gleichen Grenzwert x .

· Andererseits gilt wegen (α) , (δ) :

$$\begin{aligned} x &= \lim y_n = \lim \frac{x_n}{2} + \lim \frac{1}{x_n} = \frac{x}{2} + \frac{1}{x} \\ \Rightarrow x &= \sqrt{2} \end{aligned} \quad (3.20)$$

Die zweite wichtige Methode zur Vorstellung reeller Zahlen ist:

Theorem/Definition 3.10 (Intervallschachtelungsprinzip). *Eine Intervallschachtelung ist eine Folge von abgeschlossenen und beschränkten (kurz: kompakten) Intervallen $([a_n, b_n])_{n \in \mathbb{N}}$ mit*

$$\begin{aligned} & [a_{n+1}, b_{n+1}] \subset [a_n, b_n] \quad \forall n \\ \text{und} \quad & \lim_{n \rightarrow \infty} (b_n - a_n) = \lim_{n \rightarrow \infty} |b_n - a_n| = 0 \end{aligned} \quad (3.21)$$

(i) *Zu jeder Intervallschachtelung $([a_n, b_n])$ existiert genau eine reelle Zahl x , die in allen $[a_n, b_n]$ enthalten ist, d.h.*

$$\bigcap_{n=1}^{\infty} [a_n, b_n] = \{x\} \quad (3.22)$$

(ii) *Zu jeder reellen Zahl x existieren (sehr viele!) Intervallschachtelungen mit Durchschnitt x .*

Beweis. (i) Die Folge (a_n) ist monoton wachsend und nach oben beschränkt (durch irgendein b_n). Setze $x = \lim a_n$. Dann gilt $x \geq a_n \forall n$. Für $\epsilon > 0$ sei nun N_ϵ derart, dass

$$\text{und} \quad \left. \begin{aligned} & b_n - a_n < \epsilon \\ & x - a_n < \epsilon \end{aligned} \right\} \forall n \geq N_\epsilon \quad (3.23)$$

Dann gilt $\forall n \geq N_\epsilon$

$$-\epsilon < a_n - x \leq b_n - x \leq b_n - a_n < \epsilon \quad (3.24)$$

Es folgt $\lim b_n = x$, und wegen der Monotonie von (b_n) , dass $x \leq b_n$, also $x \in [a_n, b_n] \forall n$. Eindeutigkeit von x , und (ii), als Übung. $\square$

Warnung: Das Prinzip gilt *nicht* mit offenen Intervallen. Z.B. ist $\bigcap_{n=1}^{\infty} (0, \frac{1}{n}) = \emptyset$. Besonders mächtig ist

Theorem 3.11 (Satz von Bolzano-Weierstrass). *Jede (nach oben und unten) beschränkte reelle Folge (x_n) besitzt eine konvergente Teilfolge.*

Beweis. · Da (x_n) beschränkt ist, existiert ein $B > 0$ so, dass

$$\{x_n | n \in \mathbb{N}\} \subset [-B, B] \quad (3.25)$$

· Zur Identifikation eines mutmasslichen Grenzwertes definieren wir rekursiv eine Intervallschachtelung $([a_k, b_k])_{k \in \mathbb{N}}$ so, dass für jedes k gilt:

$$\text{Bed}(k) : [a_k, b_k] \text{ enthält unendlich viele } x_n \quad (3.26)$$

Dazu der 1. Schritt: Falls in $[0, B]$ unendlich viele x_n liegen, setze $[a_1, b_1] := [0, B]$. Andernfalls liegen in $[-B, 0]$ unendlich viele x_n und wir setzen $[a_1, b_1] := [-B, 0]$.

§ 3. FOLGEN REELLER ZAHLEN

k -ter Schritt: Sei $M := \frac{a_{k-1} + b_{k-1}}{2}$. Da $[a_{k-1}, b_{k-1}]$ wegen $\text{Bed}(k-1)$ unendlich viele x_n enthält, enthält mindestens eines von $[a_{k-1}, M]$ oder $[M, b_{k-1}]$ unendlich viele. Wir setzen $[a_k, b_k] = [M, b_{k-1}]$ falls dieses Intervall unendlich viele x_n enthält, sonst $[a_k, b_k] = [a_{k-1}, M]$. Damit ist $\text{Bed}(k)$ erfüllt und wir können fortfahren.

• $[a_{k+1}, b_{k+1}] \subset [a_k, b_k]$ ist klar und $b_k - a_k = \frac{B}{2^{k-1}} \xrightarrow{k \rightarrow \infty} 0$. Es liegt also eine Intervallschachtelung vor. Sei $x \in \mathbb{R}$ ihr Durchschnitt.

• Zur Konstruktion einer Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$ mit Grenzwert x sei n_1 ein Index so, dass $x_{n_1} \in [a_1, b_1]$ und dann rekursiv n_k ein Index, grösser als n_{k-1} so, dass $x_{n_k} \in [a_k, b_k]$. (Der Punkt ist, dass, falls es kein solches n_k gäbe, dann lägen höchstens endlich viele x_n in $[a_k, b_k]$, nämlich höchstens alle mit $n \leq n_{k-1}$, im Widerspruch zu $\text{Bed}(k)$.)

Beh.: $x_{n_k} \xrightarrow{k \rightarrow \infty} x$

Bew.: Für $\epsilon > 0$ sei $K \in \mathbb{N}$ so, dass $\frac{B}{2^{K-1}} < \epsilon \ \forall k \geq K$. Es gilt $x_{n_k} \in [a_k, b_k] \subset [a_K, b_K] \ \forall k \geq K$, ausserdem ist per Definition $x \in [a_K, b_K]$. Wegen $b_K - a_K < \epsilon$ folgt $|x - x_{n_k}| < \epsilon \ \forall k \geq K$. $\square$

Bemerkungen. Das im Beweis konstruierte x erfüllt

$$\begin{aligned} x &= \sup\{y \mid \text{Es existiert eine gegen } y \text{ konvergente Teilfolge von } (x_n)\} \\ &= \max\{y \mid \text{Es existiert eine gegen } y \text{ konvergente Teilfolge von } (x_n)\} \end{aligned} \quad (3.27)$$

Man nennt die Elemente der Menge auf der rechten Seite auch Häufungspunkte der Folge (x_n) und deren Supremum x den “limes superior”, geschrieben $x = \limsup(x_n)$.

⁴ Der “limes inferior” ist analog der kleinste Häufungswert.

Zuletzt erreichen wir das Ziel des am meisten eingesetzten Konvergenzkriteriums.

Definition 3.12. Eine Folge reeller Zahlen (x_n) heisst Cauchy-Folge (CF), falls für jedes $\epsilon > 0$ ein $N_\epsilon \in \mathbb{N}$ existiert derart, dass

$$|x_n - x_m| < \epsilon \quad \text{falls } n \geq N_\epsilon \text{ und } m \geq N_\epsilon \quad (3.29)$$

Theorem 3.13 (Cauchy-Kriterium). *Eine Folge reeller Zahlen konvergiert dann und nur dann wenn sie eine Cauchy-Folge ist.*

Bemerkungen. • Der Vorteil des Cauchy-Kriteriums gegenüber der ursprünglichen Definition 3.2 ist, dass es ohne explizite Erwähnung des Grenzwertes auskommt.

• Das Prinzip der monotonen Konvergenz kommt auch ohne den Grenzwert aus. Allerdings braucht dieses Prinzip die Anordnung von $\mathbb{R}$ in direkter Weise. Das Cauchy-Kriterium tut dies nur “indirekt”, nämlich über den Umweg des “Absolutbetrags”. Dies wird es uns ermöglichen, das Cauchy-Kriterium in einem sehr viel allgemeineren Kontext anzuwenden, in dem wir keine Anordnung, aber immer noch ein Analogon eines Absolutbetrags zur Verfügung haben.

⁴Äquivalente Charakterisierungen sind:

$$\begin{aligned} \limsup(x_n) &= \min\{y \mid \forall \epsilon > 0 \text{ ist } x_n > y + \epsilon \text{ für höchstens endlich viele } n.\} \\ &= \inf\{y \mid x_n > y \text{ für höchstens endlich viele } n.\} \end{aligned} \quad (3.28)$$

Im nach oben/unten unbeschränkten Fall macht $\limsup = \infty$, $\liminf = -\infty$ Sinn.

· In der Definition einer Cauchy-Folge würde die Forderung $|x_n - x_{N_\epsilon}| < \epsilon \forall n \geq N_\epsilon$ ausreichen (Übungsaufgabe). Hingegen ist etwa schon die Forderung $|x_{n+1} - x_n| < \epsilon \forall n \geq N_\epsilon$ zu schwach.

Der Beweis des Cauchy-Kriteriums ist mit unseren Vorbereitungen relativ einfach. Wir verteilen ihn auf drei Lemmas.

Lemma 3.14. *Jede konvergente Folge ist eine Cauchy-Folge.*

Beweis. Sei $(x_n)_{n \in \mathbb{N}}$ konvergent mit Grenzwert x . Für $\epsilon > 0$ sei N_ϵ s.d. $|x - x_n| < \frac{\epsilon}{2} \forall n \geq N_\epsilon$. Dann gilt für $n \geq N_\epsilon$ und $m \geq N_\epsilon$,

$$|x_n - x_m| \leq |x_n - x| + |x - x_m| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \quad (3.30)$$

□

Lemma 3.15. *Cauchy-Folgen sind beschränkt.*

Beweis. Sei $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge. Für $\epsilon = 1$ sei N_1 s.d. $|x_n - x_m| < 1 \forall n, m \geq N_1$. Insbesondere gilt mit $m = N_1$ für $n \geq N_1$: $|x_n| \leq |x_{N_1}| + 1 \forall n \geq N_1$. Mit $B := \max\{|x_1|, |x_2|, \dots, |x_{N_1-1}|, |x_{N_1}| + 1\}$ gilt dann $|x_n| \leq B \forall n$. □

Lemma 3.16. *Jede Cauchy-Folge ist konvergent.*

Beweis. Nach Lemma 3.15 ist eine Cauchy-Folge (x_n) jedenfalls beschränkt. Nach 3.11 existiert eine konvergente Teilfolge (x_{n_k}) mit Grenzwert $x := \lim_{k \rightarrow \infty} x_{n_k}$. Wir behaupten, dass $x_n \xrightarrow{n \rightarrow \infty} x$.

Bew.: Für $\epsilon > 0$ sei N_ϵ derart, dass $|x_n - x_m| < \epsilon$ für $n \geq N_\epsilon$ und $m \geq N_\epsilon$. ((x_n) ist eine Cauchy-Folge.) Sodann sei K_ϵ derart, dass $n_{K_\epsilon} \geq N_\epsilon$ und $|x_{n_{K_\epsilon}} - x| < \epsilon$. (Konvergenz der Teilfolge). Zusammen folgt für $n \geq N_\epsilon$:

$$|x_n - x| \leq |x_n - x_{n_{K_\epsilon}}| + |x_{n_{K_\epsilon}} - x| < 2\epsilon \quad (3.31)$$

□

Eine weitere Anwendung: Überabzählbarkeit von $\mathbb{R}$

Theorem 3.17. *Die Menge der reellen Zahlen ist nicht abzählbar.*

Beweis. Wir zeigen, dass für jede Folge (x_n) paarweise verschiedener reeller Zahlen mindestens eine reelle Zahl h existiert mit $h \neq x_n \forall n \in \mathbb{N}$. (Insbesondere kann also keine Abzählung *aller* reeller Zahlen existieren.)

Bew.: · Sei $a_1 = \min\{x_1, x_2\}$, $b_1 = \max\{x_1, x_2\}$ und $I_1 = (a_1, b_1)$ (das offene Intervall!) Wir setzen $k_1 = 1$ oder $k_1 = 2$ so, dass $a_1 = x_{k_1}$, und entsprechend l_1 so, dass $b_1 = x_{l_1}$. Wegen $a_1 < b_1$ ist I_1 nicht leer und enthält unendlich viele reelle Zahlen. Falls nur endlich viele Glieder von (x_n) in I_1 liegen, so existiert bereits ein $h \in I_1$, welches nicht in (x_n) vorkommt.

· Wir nehmen daher an, dass unendlich viele Glieder von (x_n) in I_1 liegen, und bezeichnen die ersten beiden mit y_1 und z_1 . Sei $a_2 = \min\{y_1, z_1\}$ und $b_2 = \max\{y_1, z_1\}$. Dann gilt $a_1 < a_2 < b_2 < b_1$. Es existiert (genau) ein k_2 mit $a_2 = x_{k_2}$ und (genau) ein l_2 mit $b_2 = x_{l_2}$.

§ 4. METRISCHE RÄUME

- Wir wiederholen das Verfahren und erhalten entweder ein h nicht in (x_n) oder aber eine streng monoton wachsende Teilfolge $(a_m) = (x_{k_m})$ und eine streng monoton fallende Teilfolge $(b_m) = (x_{l_m})$. (Die Monotonie von (k_m) und (l_m) folgt aus der Wahl von y_m, z_m als die ersten Glieder von (x_n) in I_m und der Schachtelung der Intervalle.)
- Wegen 3.9 existiert $a = \lim_{m \rightarrow \infty} a_m$ und $b = \lim b_m$, und es gilt $a \leq b$ und es existiert ein $h \in [a, b]$.
- Wegen der strengen Monotonie von (a_m) und (b_m) liegt h in keiner der beiden Folgen (Übung). Wir behaupten, h liegt auch nicht in der ursprünglichen Folge (x_n) .
- Denn angenommen $h = x_{n_*}$, dann treten nur endlich viele $a_m = x_{k_m}$ vor h in der Folge (x_n) auf, d.h. es gibt ein letztes d mit $k_d < n_*$ aber $k_m > n_* \quad \forall m > d$.
- Insbesondere gilt $k_{d+1} > n_*$, d.h. a_{d+1} tritt in der Folge (x_n) nach $x_{n_*} = h$ auf. Wegen $h \in (a_m, b_m) \quad \forall m$ steht dies aber im Widerspruch dazu, dass a_{d+1} und b_{d+1} die ersten Glieder von (x_n) im Intervall (a_d, b_d) sind. $\square$
- Es gibt einfachere Beweise mittels Entwicklung reeller Zahlen in Brüche mit einer festen Basis (z.B. Dezimal- oder Binärbrüche).

§ 4 Metrische Räume

Im letzten § haben wir das Cauchy-Kriterium für die Konvergenz reeller Folgen aus der Supremumseigenschaft von $\mathbb{R}$ hergeleitet, mit der Ankündigung, dass

- (i) für einen angeordneten Körper das Cauchy-Kriterium (zusammen mit dem Archimedischen Axiom) eine zur Supremumseigenschaft äquivalente Charakterisierung der Vollständigkeit der reellen Zahlen liefert, und
- (ii) wir das Cauchy-Kriterium sodann nutzen wollen, um den Begriff der Vollständigkeit auf Situationen auszudehnen, in denen eine Anordnung abgeht.

Der Ankündigung (ii) wollen wir nun Folge leisten. Die wesentliche Beobachtung ist, dass im Cauchy-Kriterium die Anordnung nur mittels des Absolutbetrags $|\cdot| : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ eingeht.

Definition 4.1. Ein metrischer Raum (sc. über $\mathbb{R}$) ist eine Menge X zusammen mit einer Abbildung (“Abstandsfunktion”)

$$d_X = d : X \times X \rightarrow \mathbb{R}_{\geq 0} = \{x \in \mathbb{R} | x \geq 0\} \quad (4.1)$$

mit den Eigenschaften:

$$\begin{aligned} \text{Symmetrie:} \quad & d(x, y) = d(y, x) \quad \forall x, y \in X \\ \text{Positivität:} \quad & d(x, y) = 0 \Leftrightarrow x = y \\ \text{Dreiecksungleichung:} \quad & d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in X \end{aligned} \quad (4.2)$$

Wem dieser Begriff zunächst zu abstrakt erscheint, möge sich in der Folge immer eines (oder besser jedes) der unten folgenden Beispiele vor Augen führen. Wir erklären dann

Definition 4.2. · Eine Folge $(x_k)_{k \in \mathbb{N}}$ in einem metrischen Raum (X, d) heisst konvergent gegen $x \in X$ ⁵ falls $\forall \epsilon > 0$ ein K_ϵ existiert so, dass

$$d(x, x_k) < \epsilon \quad \forall k \geq K_\epsilon \quad (4.3)$$

· Eine Folge (x_k) heisst Cauchy-Folge, falls $\forall \epsilon > 0$ ein K_ϵ existiert so, dass

$$d(x_k, x_l) < \epsilon \quad \forall k, l \geq K_\epsilon \quad (4.4)$$

· Ein metrischer Raum heisst vollständig, falls jede Cauchy-Folge in X konvergiert.

Bemerkungen. Wie bei $\mathbb{R}$ zeigt man leicht, dass Grenzpunkte eindeutig sind (dies basiert auf $d(x, y) < \epsilon \quad \forall \epsilon > 0 \Rightarrow x = y$), und dass jede konvergente Folge eine Cauchy-Folge ist (Dreiecksungleichung). Der wesentliche Punkt ist also Cauchy-Folge $\Rightarrow$ Konvergenz.

Beispiel 4.3. Die Menge der reellen Zahlen $\mathbb{R}$, ausgestattet mit der Abstandsfunktion $d_{\mathbb{R}}(x, y) = |x - y|$ ist ein vollständiger metrischer Raum im Sinne von 4.2 (siehe Theorem 3.13). Um mit der Ordnungsvollständigkeit aus 1.10 zu vergleichen, und damit der Ankündigung (i) etwas Substanz zu verleihen, bemerken wir Folgendes: Ist $\mathbb{K}$ ein angeordneter Archimedischer Körper, so können wir den Absolutbetrag $|\cdot|_{\mathbb{K}}$, definiert über (1.11), nach Thm. 1.11 auffassen als Abbildung $\mathbb{K} \rightarrow \mathbb{R}_{\geq 0}$, und man prüft sofort, dass durch $d_{\mathbb{K}}(x, y) = |x - y|_{\mathbb{K}}$ eine Abstandsfunktion im Sinne von 4.1 erklärt wird. Es gilt dann:

Beh.: Ist $\mathbb{K}$ mit dieser Abstandsfunktion ein vollständiger metrischer Raum, so ist $\mathbb{K}$ ordnungsvollständig, d.h. jede nicht-leere nach oben beschränkte Menge besitzt eine kleinste obere Schranke. M.a.W. ist $\mathbb{K} = \mathbb{R}$.

Bew.: (Skizze) Sei $T \subset \mathbb{K}$ nicht leer und $|T|_{\mathbb{K}} \leq B$ für ein $B \in \mathbb{K}$. Mit $x \in T$ setze $[a_1, b_1] = [x - 1, B]$ und definiere rekursiv für $k = 1, 2, \dots$:

$$m_k := \frac{a_k + b_k}{2} \quad (4.5)$$

$$[a_{k+1}, b_{k+1}] := \begin{cases} [a_k, m_k] & \text{falls } [m_k, b_k] \cap T = \emptyset \\ [m_k, b_k] & \text{falls } [m_k, b_k] \cap T \neq \emptyset \end{cases}$$

Die Folge $([a_k, b_k])_{k \in \mathbb{N}}$ ist dann eine Intervallschachtelung in $\mathbb{K}$ (!) (Beachte dass $|b_k - a_k|_{\mathbb{K}} = 2^{1-k} |b_1 - a_1|_{\mathbb{K}}$ wegen 1.9 bereits in $\mathbb{K}$ eine Nullfolge ist, da $\mathbb{K}$ Archimedisches ist.), und man zeigt ohne grosse Mühe, dass (a_k) , (b_k) Cauchy-Folgen sind, deren gemeinsamer Grenzwert $s = \lim a_k = \lim b_k$ eine kleinste obere Schranke von T ist. (Alle b_k sind obere Schranken von T , und für jedes $\epsilon > 0$ ist $b_k - \epsilon < a_k$ für k gross genug keine obere Schranke, da $[a_k, b_k] \cap T \neq \emptyset \quad \forall k$.)

Ganz allgemein ist:

Beispiel 4.4. Sei X irgendeine Menge. Setze

$$d(x, y) = \begin{cases} 0 & \text{falls } x = y \\ 1 & \text{sonst} \end{cases} \quad (4.6)$$

Dann ist (X, d) ein vollständiger metrischer Raum.

⁵Elemente von X werden normalerweise als Punkte bezeichnet, Limites von Folgen aber oft wieder als "Grenzwerte".

§ 4. METRISCHE RÄUME

Für uns sind besonders metrische Räume mit zusätzlicher Struktur (z.B. linearen Strukturen) wichtig und wir verlangen Kompatibilität der Abstandsfunktion.

Beispiel 4.5. Für $n \geq 1$ ist der Raum $\mathbb{R}^n \ni x = (x^1, \dots, x^n)^T$, ausgestattet mit dem Euklidischen Abstand

$$d(x, y) = \left(\sum_{i=1}^n (x^i - y^i)^2 \right)^{1/2} \quad (4.7)$$

(nicht-negative Quadratwurzel) ein vollständiger metrischer Raum.

1. Teil: $(\mathbb{R}^n, d)$ ist ein metrischer Raum: Symmetrie und Positivität von d sind klar, für die Dreiecksungleichung betrachten wir für $x, y, z \in \mathbb{R}^n$:

$$\begin{aligned} d(x, y)^2 &= \sum_{i=1}^n (x^i - z^i + z^i - y^i)^2 \\ &= \sum_{i=1}^n (x^i - z^i)^2 + \sum_{i=1}^n (z^i - y^i)^2 + 2 \sum_{i=1}^n (x^i - z^i)(z^i - y^i) \\ &= d(x, z)^2 + d(z, y)^2 + 2 \langle x - z, z - y \rangle \end{aligned} \quad (4.8)$$

wobei $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ das euklidische (oder standard) innere Produkt auf $\mathbb{R}^n$ ist. Dieses erfreut sich bekanntermassen der folgenden Eigenschaften:

$$\text{Symmetrie: } \langle \xi, \eta \rangle = \langle \eta, \xi \rangle$$

$$\text{(Bi-)Linearität: } \langle \xi + \lambda \eta, \zeta \rangle = \langle \xi, \zeta \rangle + \lambda \langle \eta, \zeta \rangle \quad (4.9)$$

$$\text{Positiv Definitheit: } \|\xi\|_2^2 := \langle \xi, \xi \rangle > 0 \quad \text{falls } \xi \neq 0$$

$$\text{Cauchy-Schwarz-Ungleichung: } |\langle \xi, \eta \rangle| \leq \|\xi\|_2 \|\eta\|_2 \quad (\text{pos. Wurzel})$$

Beweis der C-S-Ungleichung. Für $\eta = 0$ ($\Leftrightarrow \|\eta\|_2 = 0$) ist die Aussage klar. Für $\eta \neq 0$, beliebiges ξ , sei

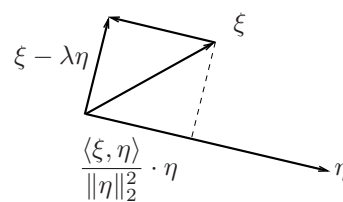
$$\lambda := \frac{\langle \xi, \eta \rangle}{\|\eta\|_2^2} \quad (4.10)$$

Dann gilt:

Deutung:

$$\begin{aligned} 0 &\leq \|\xi - \lambda \eta\|_2^2 = \|\xi\|_2^2 + \lambda^2 \|\eta\|_2^2 - 2\lambda \langle \xi, \eta \rangle \\ &= \|\xi\|_2^2 + \frac{\langle \xi, \eta \rangle^2}{\|\eta\|_2^2} - 2 \frac{\langle \xi, \eta \rangle^2}{\|\eta\|_2^2} \\ &= \|\xi\|_2^2 - \frac{\langle \xi, \eta \rangle^2}{\|\eta\|_2^2} \\ \Rightarrow \quad \langle \xi, \eta \rangle^2 &\leq \|\xi\|_2^2 \cdot \|\eta\|_2^2 \end{aligned}$$

$$(4.11)$$



Eingesetzt in (4.8) ergibt sich mit $\xi = x - z$, $\eta = z - y$:

$$d(x, y)^2 \leq d(x, z)^2 + d(z, y)^2 + 2d(x, z) \cdot d(z, y) = (d(x, y) + d(z, y))^2 \quad (4.12)$$

und daraus die Behauptung. $\square$

Intermezzo zur Vorbereitung des Beweises der Vollständigkeit: Die soeben benutzte Beziehung $d(x, y) = \|x - y\|_2$ zwischen der Abstandsfunktion (4.7) auf der dem $\mathbb{R}^n$ zugrundeliegenden Menge und dem Euklidischen inneren Produkt auf $\mathbb{R}^n$ als Vektorraum hat die folgenden Eigenschaften zur Folge ($x, y, z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$):

$$\begin{aligned} \text{Translationsinvarianz: } d(x + z, y + z) &= d(x, y) \\ \text{Homogenität: } d(\lambda x, \lambda y) &= |\lambda| d(x, y) \end{aligned} \quad (4.13)$$

Es liegt damit ein Beispiel einer etwas allgemeineren Konstruktion vor:

Definition 4.6. Sei V ein reeller Vektorraum. Eine Norm auf V ist eine Funktion

$$\|\cdot\| : V \rightarrow \mathbb{R}_{\geq 0} \quad (4.14)$$

mit den Eigenschaften

$$\begin{aligned} \|\lambda v\| &= |\lambda| \cdot \|v\| \quad \forall \lambda \in \mathbb{R}, v \in V \\ \|v\| &> 0 \quad \forall v \neq 0 \\ \|v + w\| &\leq \|v\| + \|w\| \end{aligned} \quad (4.15)$$

Ein solches Paar $(V, \|\cdot\|)$ heisst normierter Vektorraum.

Bemerkungen. Man trifft die gleiche Verabredung für einen komplexen Vektorraum, mit $|\cdot|$ = Absolutbetrag auf $\mathbb{C}$.

Lemma 4.7. Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Mit $d_{\|\cdot\|}(v, w) = \|v - w\|$ wird (die) V (zugrundeliegende Menge) zu einem metrischen Raum. $d_{\|\cdot\|}$ ist translationsinvariant und homogen (vgl. (4.13)).

· Ist umgekehrt V ein Vektorraum mit Abstandsfunktion d mit diesen Eigenschaften, so wird V durch $\|v\|_d := d(v, 0)$ zu einem normierten Vektorraum.

Beweis. Übungsaufgabe □

Beispiel 4.8. $\cdot \mathbb{R}^n$ mit $\|\cdot\| = \|\cdot\|_2$ ist ein normierter Vektorraum.

· Ebenso ist $\|\cdot\|_\infty$, definiert durch

$$\|x\|_\infty := \max\{|x^i| \mid i = 1, \dots, n\} \quad (4.16)$$

eine Norm auf $\mathbb{R}^n$. Denn es gilt ja

$$\|x + y\|_\infty = \max\{|x^i + y^i|\} \leq \max\{|x^i| + |y^i|\} \leq \max\{|x^i|\} + \max\{|y^i|\} \quad (4.17)$$

Lemma 4.9. Es existieren Konstanten $c > 0$, $C > 0$ so dass $\forall x \in \mathbb{R}^n$:

$$c \cdot \|x\|_\infty \leq \|x\|_2 \leq C \cdot \|x\|_\infty \quad (4.18)$$

Beweis. · Aus $|x^i|^2 \leq \sum_{j=1}^n |x^j|^2 \forall i$ folgt $|x^i| \leq \|x\|_2 \forall i$ und daher $\|x\|_\infty \leq \|x\|_2$ (d.h. $c = 1$)

· Andererseits gilt

$$\sum_{j=1}^n |x^j|^2 \leq n \cdot \max\{|x^i|^2\} = n \|x\|_\infty^2 \quad (4.19)$$

also $\|x\|_2 \leq \sqrt{n} \|x\|_\infty$, d.h. $C = \sqrt{n}$. □

Man sagt auch, die Normen $\|\cdot\|_\infty$ und $\|\cdot\|_2$, oder auch die zugehörigen Abstandsfunktionen, seien äquivalent. Beispielsweise ist es zur Feststellung, ob eine Folge eine Cauchy-Folge ist, egal, welche von den zwei Abstandsfunktionen man zu Grunde legt. Damit kommen wir zum

§ 4. METRISCHE RÄUME

2. Teil von (4.5): Vollständigkeit von $\mathbb{R}^n$: Sei $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^n$ eine Cauchy-Folge, sagen wir bzgl. $\|\cdot\|_\infty$. Sei für $\epsilon > 0$ $K_\epsilon \in \mathbb{N}$ s.d. $\|x_k - x_l\|_\infty < \epsilon \ \forall k, l \geq K_\epsilon$. Dann folgt für jedes $i = 1, \dots, n$:

$$|x_k^i - x_l^i| \leq \|x_k - x_l\|_\infty < \epsilon \quad \forall k, l \geq K_\epsilon \quad (4.20)$$

Also ist für jedes $i = 1, \dots, n$ die Folge $(x_k^i) \subset \mathbb{R}$ eine Cauchy-Folge, konvergiert also wegen der Vollständigkeit von $\mathbb{R}$. Wir setzen $x^i = \lim_{k \rightarrow \infty} x_k^i$ und $x = (x^1, \dots, x^n)^T$. Für $\epsilon > 0$ sei K_ϵ so, dass $|x^i - x_k^i| < \epsilon$ für $k \geq K_\epsilon$. Dann gilt mit $K_\epsilon = \max\{K_\epsilon^i\}$, $\|x - x_k\|_\infty < \epsilon$ für $k \geq K_\epsilon$. Wegen der Äquivalenz der Normen folgt

$$\lim_{k \rightarrow \infty} \|x_k - x\|_2 = \lim_{k \rightarrow \infty} \|x_k - x\|_\infty = 0 \quad (4.21)$$

□_{4.5}

· Im Spezialfall $n = 2$ sehen wir: Der Körper $\mathbb{C}$ der komplexen Zahlen, versehen mit der Abstandsfunktion

$$d(z, w) = |z - w| \in \mathbb{R}_{\geq 0} \quad (4.22)$$

ist ein vollständiger metrischer Raum (über $\mathbb{R}$!). In der Tat ist ja per Definition für $z = x + iy \in \mathbb{C}$ mit $(x, y) \in \mathbb{R}^2$,

$$|z| = (x^2 + y^2)^{1/2} = \|(x, y)\|_2 \quad (4.23)$$

· Zusätzlich zur für jeden $\mathbb{R}^n$ geltenden Aussage, dass (nach jetzigem Wissensstand zumindest) die Normen $\|\cdot\|_2$ und $\|\cdot\|_\infty$ äquivalent sind, ist bei $\mathbb{R}^2 = \mathbb{C}$ die euklidische Norm = Absolutbetrag durch die Verträglichkeit mit der Multiplikation ausgezeichnet. Für $z_1, z_2 \in \mathbb{C}$ gilt $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$. (Hingegen ist beispielsweise $\|(3, 2)(2, -1)\|_\infty = 8 \neq 6 = \|(3, 2)\|_\infty \cdot \|(2, -1)\|_\infty$.)

· Ein vergleichbares Beispiel sind die aus den Übungen bekannten Quaternionen (mit dem Unterschied, dass die quaternionische Multiplikation auf $\mathbb{R}^4$ nicht kommutativ ist.)

Beispiel 4.10. Seien $V \cong \mathbb{R}^n \ni x = (x^i)_{i=1, \dots, n}$, $W \cong \mathbb{R}^m \ni y = (y^j)_{j=1, \dots, m}$ (nicht-triviale) endlich-dimensionale Vektorräume über $\mathbb{R}$, mit Normen $\|\cdot\|_V$ und $\|\cdot\|_W$, von denen wir annehmen, dass sie unter der gegebenen Identifikation zu $\|\cdot\|_\infty$ auf $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ äquivalent sind. (Wir werden zeigen können, dass diese Annahme für endlich-dimensionale Vektorräume automatisch erfüllt ist. Die Identifikation mit $\mathbb{R}^n$ entspricht der Wahl einer Basis in V . Man kann sich bei Bedarf unter $\|\cdot\|_V$ z.B. $\|\cdot\|_2$ auf $\mathbb{R}^n$ vorstellen.) Identifizieren wir dann eine lineare Abbildung $A : V \rightarrow W$ mit der darstellenden Matrix

$$(A_i^j)_{\substack{j=1, \dots, m \\ i=1, \dots, n}} \in \text{Mat}_{m \times n}(\mathbb{R}) \quad (4.24)$$

bzgl. den gewählten Basen, so gilt für $x \in \mathbb{R}^n \cong V$:

$$\begin{aligned}
 \|Ax\|_\infty &= \max \left\{ \left| \sum_{i=1}^n A_i^j x^i \right| \mid j = 1, \dots, m \right\} \\
 &\leq \max \left\{ \sum_{i=1}^n |A_i^j x^i| \mid j = 1, \dots, m \right\} \\
 &\leq \max \left\{ \sum_{i=1}^n |A_i^j| \max\{|x^i| \mid i = 1, \dots, n\} \mid j = 1, \dots, m \right\} \\
 &\leq \max \left\{ \sum_{i=1}^n |A_i^j| \mid j = 1, \dots, m \right\} \cdot \|x\|_\infty
 \end{aligned} \tag{4.25}$$

Wegen der Äquivalenz von $\|\cdot\|_{V,W}$ mit $\|\cdot\|_\infty$ ist daher die Menge

$$\left\{ \|Ax\|_W \mid \|x\|_V = 1 \right\} = \left\{ \frac{\|Ax\|_W}{\|x\|_V} \mid x \neq 0 \right\} \tag{4.26}$$

(nicht-leer und) nach oben beschränkt und wir setzen

$$\|A\| := \sup \{ \|Ax\|_W \mid \|x\|_V = 1 \} \tag{4.27}$$

Es gilt dann: (Übungsaufgabe)

$$\begin{aligned}
 \text{(i)} \quad & \|Ax\|_W \leq \|A\| \cdot \|x\|_V \quad \forall x \in V \\
 \text{(ii)} \quad & \|A + B\| \leq \|A\| + \|B\| \text{ für } A, B \in \text{Hom}_{\mathbb{R}}(V, W) \\
 & \|\lambda A\| = |\lambda| \cdot \|A\| \text{ für } \lambda \in \mathbb{R} \\
 & \|A\| = 0 \Leftrightarrow A = 0,
 \end{aligned} \tag{4.28}$$

mit anderen Worten definiert (4.27) eine Norm auf $\text{Hom}_{\mathbb{R}}(V, W)$, die sog. Matrix- oder Operatornorm. Diese Norm hängt von der Wahl der Normen auf V und W , ist aber in jedem Fall äquivalent zu $\|\cdot\|_\infty$ auf $\mathbb{R}^{m \cdot n}$ (als Vektorraum). $(\text{Hom}_{\mathbb{R}}(V, W), d_{\|\cdot\|})$ ist ein vollständiger metrischer Raum.

• Die Operatornorm ist submultiplikativ:

$$\begin{aligned}
 \|ABx\| &\leq \|A\| \cdot \|Bx\| \leq \|A\| \cdot \|B\| \cdot \|x\| \\
 &\Rightarrow \|AB\| \leq \|A\| \cdot \|B\|
 \end{aligned} \tag{4.29}$$

(Im Allgemeinen kann hier aber nicht = gelten, denn es gibt lineare Abbildungen $A, B \neq 0$ mit $AB = 0$.)

Eine weitere Klasse von Beispielen, die uns in dieser Vorlesung noch häufiger beschäftigen wird, wird durch die folgende Verallgemeinerung von $\|\cdot\|_\infty$ auf $\mathbb{R}^n$ illustriert.

Beispiel 4.11. Sei $X \neq \emptyset$ irgendeine Menge. Die Menge

$$\mathcal{B}(X, \mathbb{R}) := \{ f : X \rightarrow \mathbb{R} \mid \exists M > 0 : |f(x)| \leq M \quad \forall x \in X \} \tag{4.30}$$

§ 4. METRISCHE RÄUME

der beschränkten Funktionen auf X wird durch punktweise Addition und Skalarmultiplikation zu einem (im Allgemeinen ∞ -dimensionalen) reellen Vektorraum. ($\mathbb{R}^n$ ist der Spezialfall $X = \{1, 2, \dots, n\}$.)

· Die Abbildung

$$\mathcal{B}(X, \mathbb{R}) \ni f \mapsto \|f\|_\infty = \sup\{|f(x)| \mid x \in X\} \in \mathbb{R}_{\geq 0} \quad (4.31)$$

ist eine Norm auf $\mathcal{B}(X, \mathbb{R})$ (die sup-Norm) und $(\mathcal{B}(X, \mathbb{R}), d_{\|\cdot\|_\infty})$ ist ein vollständiger metrischer Raum.

Bew.: Sei $(f_k) \subset \mathcal{B}(X, \mathbb{R})$ eine Cauchy-Folge. Fixiere zunächst $x \in X$. Für jedes $\epsilon > 0$ existiert ein K_ϵ so, dass $\|f_k - f_l\|_\infty < \epsilon \forall k, l \geq K_\epsilon$. Insbesondere ist dann $|f_k(x) - f_l(x)| < \epsilon \forall k, l \geq K_\epsilon$. Daraus folgt, dass für jedes x , $(f_k(x)) \subset \mathbb{R}$ eine Cauchy-Folge ist. Wir definieren $f : X \rightarrow \mathbb{R}$ durch $f(x) = \lim f_k(x)$, und zeigen

(i) $f \in \mathcal{B}(X, \mathbb{R})$ und

(ii) $\lim \|f_k - f\|_\infty = 0$

Für $\epsilon > 0$ und K_ϵ wie oben gilt für jedes $k \geq K_\epsilon$: Für jedes x ist $|f_k(x) - f_l(x)| < \epsilon \forall l \geq K_\epsilon$. Mit $l \rightarrow \infty$ folgt $|f_k(x) - f(x)| \leq \epsilon$. Daraus folgt (i) $|f(x)| \leq |f_k(x)| + \epsilon \forall x$, d.h. f ist durch $\|f_k\|_\infty + \epsilon$ beschränkt, sowie (ii) $\|f_k - f\|_\infty < 2\epsilon$. $\square$

Alle in diesem § gegebenen Beispiele (ausser 4.4) basierten auf normierten Vektorräumen (BTW, ersetzt man $\mathbb{R}^n$ durch $\mathbb{Q}^n$ so erhält man sofort nicht-vollständige metrische Räume). Die Rechenregeln (α) und (γ) aus 3.7 übertragen sich sinngemäss auf solche Situationen, (β) auf normierte Algebren, und (δ) falls Inverse existieren (wie z.B. $\mathbb{C}$).

Ein normierter Vektorraum, der bezüglich der induzierten Metrik vollständig ist, heisst *Banach-Raum*. Ist ausserdem eine Multiplikation erklärt, die unter der Norm eine Abschätzung der Form (4.29) erfüllt, so spricht man von einer *Banach-Algebra*.

Für den Konvergenzbegriff benötigt man aber nicht unbedingt die lineare Struktur, wie wir noch anhand einiger Definitionen illustrieren.

Topologische Grundbegriffe

Die grundlegende Beobachtung ist die folgende: Sei (X, d) ein metrischer Raum, und $T \subset X$ eine Teilmenge. Dann erfüllt die Einschränkung

$$d = d|_{T \times T} : T \times T \rightarrow \mathbb{R} \quad (4.32)$$

trivialerweise wieder alle Eigenschaften einer Abstandsfunktion, d.h. (T, d) ist in natürlicher Weise ein metrischer Raum. Beim Vergleich zwischen Konvergenz in T und Konvergenz in X trifft man dann auf die folgende Unterscheidung.

Definition 4.12. (i) Eine Teilmenge A eines metrischen Raumes heisst abgeschlossen, falls für jede Folge $(x_k)_{k \in \mathbb{N}}$ in A , welche als Folge in X konvergiert, der Grenzwert $x = \lim x_k$ in A liegt.

(ii) Eine Teilmenge U eines metrischen Raumes heisst offen, falls $\forall x \in U$ ein $\epsilon > 0$ existiert so, dass die “ ϵ -Kugel um x ” voll in U enthalten ist, d.h.

$$B_\epsilon(x) = \{y \in X \mid d(x, y) < \epsilon\} \subset U \quad (4.33)$$

· Insbesondere sind für $x \in X$ und $R > 0$ die offenen Kugeln

$$B_R(x) = \{y \in X \mid d(x, y) < R\} \quad (4.34)$$

offen in X : Wegen der Dreiecksungleichung ist für $y \in B_R(x)$ und $0 < \epsilon < R - d(x, y)$, $B_\epsilon(y) \subset B_R(x)$.

· Typische Beispiele abgeschlossener Mengen sind die “abgeschlossenen Kugeln”

$$\overline{B_R(x)} = \{y \in X \mid d(x, y) \leq R\} \quad (4.35)$$

· Man beachte, dass “offen” und “abgeschlossen” sich nicht gegenseitig ausschliessen. Insbesondere sind X und die leere Menge (bei geeigneter Verabredung) beides, so dass gilt:

Lemma 4.13. $U \subset X$ ist offen $\iff X \setminus U$ ist abgeschlossen.

Beweis. Für $U = \emptyset$ oder $X \setminus U = \emptyset$ ist die Aussage wahr. Andernfalls:

“ $\Rightarrow$ ”: Sei $(x_k) \subset X \setminus U$ eine Folge mit Grenzpunkt $x = \lim x_k \in X$. Wäre $x \in U$ so gäbe es ein $\epsilon_* > 0$ so, dass $B_{\epsilon_*}(x) \subset U$, ausserdem ein K_{ϵ_*} so, dass $d(x, x_k) < \epsilon$, d.h. $x_k \in B_{\epsilon_*}(x) \forall k \geq K_{\epsilon_*}$. Dies steht im Widerspruch zu $x_k \in X \setminus U$.

“ $\Leftarrow$ ”: Sei $x \in U$. Gäbe es kein $\epsilon > 0$ mit $B_\epsilon(x) \subset U$, d.h. wäre für jedes $\epsilon > 0$ $B_\epsilon(x) \cap X \setminus U \neq \emptyset$, so gäbe es für jedes $k \in \mathbb{N}$ ein $x_k \in B_{1/k}(x) \cap X \setminus U$. Die Folge $(x_k) \subset X \setminus U$ konvergiert dann gegen $x \in U$, im Widerspruch zur Abgeschlossenheit von $X \setminus U$. $\square$

Die folgenden zwei Aussagen sind unmittelbar:

Lemma 4.14. $U \subset X$ ist offen $\iff$ Für jede Folge $(x_k) \subset X$, welche gegen $x \in U$ konvergiert, liegen fast alle x_k in U .

Lemma 4.15. Eine abgeschlossene Teilmenge eines vollständigen metrischen Raumes ist vollständig (als metrischer Raum für sich).

Wir nennen einen metrischen Raum (X, d) beschränkt, falls die Teilmenge

$$\{d(x, y) \mid x, y \in X\} \quad (4.36)$$

von $\mathbb{R}$ beschränkt ist. Falls $X \neq \emptyset$, so heisst das Supremum dieser Menge der Durchmesser von X , $\text{diam } X$. Wir wenden dies auch auf (nicht-leere) Teilmengen von X an und erhalten so eine nützliche Verallgemeinerung des “Intervallschachtelungsprinzips” 3.10:

Proposition 4.16. Sei $(A_k)_{k \in \mathbb{N}}$ eine Folge von Teilmengen eines vollständigen metrischen Raumes (X, d) derart, dass gilt:

(i) $\forall k \in \mathbb{N}$ ist A_k nicht-leer, abgeschlossen und beschränkt.

(ii) $A_{k+1} \subset A_k \forall k$

(iii) $\text{diam } A_k \rightarrow 0$ für $k \rightarrow \infty$.

Dann existiert genau ein $x \in \bigcap_{l=1}^{\infty} A_l$.

§ 4. METRISCHE RÄUME

Beweis. Wähle für jedes $k \in \mathbb{N}$ ein $x_k \in A_k$. Wegen der Schachtelung ist (x_k) eine Cauchy-Folge, und für jedes $l \in \mathbb{N}$ liegen fast alle Glieder von (x_k) in A_l . Wegen der Abgeschlossenheit von A_l liegt dann der Grenzpunkt $x = \lim x_k$ in jedem A_l . Eindeutigkeit von x ist klar. $\square$

Zum Schluss kehren wir zu $\mathbb{R}^n$ zurück und beweisen die folgende Verallgemeinerung des Satzes von Bolzano-Weierstrass:

Proposition 4.17. *Jede beschränkte Folge $(x_k) \subset \mathbb{R}^n$ besitzt eine konvergente Teilfolge.*

Bemerkungen. Wir hatten im Laufe von 4.5 gesehen, dass $\mathbb{R}^n$ mit der durch die Norm $\|\cdot\|_\infty$ gegebenen Abstandsfunktion vollständig ist und zeigen die Aussage zunächst in diesem Fall. Anschliessend benutzen wir das Resultat, um zu zeigen, dass alle Normen auf $\mathbb{R}^n$ äquivalent sind, so dass 4.17 dann mit jeder Norm gilt.

Beweis. In Imitation des Beweises von 3.11 konstruieren wir zunächst eine Folge $(A_l)_{l \in \mathbb{N}}$ wie in 4.16 mit der Eigenschaft, dass jedes A_l unendlich viele x_k enthält. Da $\{x_k \mid k \in \mathbb{N}\}$ beschränkt ist, existiert ein $R > 0$ so, dass $\|x_k\|_\infty \leq R \forall k$, d.h.,

$$(x_k) \subset \overline{B_R(0)} = \{y \in \mathbb{R}^n \mid \|y\|_\infty \leq R\} = [-R, R]^n \quad (4.37)$$

Laut (4.35) sind die $\overline{B_R(0)}$ (geometrisch gesehen Quader) abgeschlossen. Im ersten Schritt teilen wir $\overline{B_R(0)}$ in 2^n Teile,

$$\overline{B_{R/2}((\pm \frac{R}{2}, \pm \frac{R}{2}, \dots, \pm \frac{R}{2})^T)} \quad (4.38)$$

vom Durchmesser R , und wählen für A_1 einen dieser Quader, welcher unendlich viele Folgenglieder enthält. Wir verfahren analog im l -ten Schritt, erhalten eine Schachtelung mit den gewünschten Eigenschaften, und setzen $\{x\} = \cap_{l=1}^\infty A_l$. Eine Wahl $x_{k_l} \in A_l$ so, dass (k_l) streng monoton wächst, liefert eine gegen x konvergente Teilfolge (x_{k_l}) . $\square$

Es ist instruktiv, sich zu überlegen, an welcher Stelle dieser Beweis im Fall eines unendlichen metrischen Raumes der Art 4.4 scheitert (und dass auch die Aussage dann falsch ist!).

Proposition 4.18. *Sei $\|\cdot\|$ eine beliebige Norm auf $\mathbb{R}^n$. Dann existieren Konstanten $c > 0$, $C > 0$ so, dass für alle $x \in \mathbb{R}^n$*

$$c \cdot \|x\|_\infty \leq \|x\| \leq C \cdot \|x\|_\infty \quad (4.39)$$

Beweis. Sind für $i = 1, \dots, n$ $e_i = (0, \dots, \overset{i\text{-te Stelle}}{1}, \dots, 0)^T$ die Elemente der Standardbasis des $\mathbb{R}^n$, so folgt für $x = \sum x^i e_i$ mit Hilfe der Dreiecksungleichung und der Homogenität von $\|\cdot\|$:

$$\|x\| \leq \sum_{i=1}^n |x^i| \cdot \|e_i\| \leq n \cdot \max\{\|e_i\|\} \cdot \|x\|_\infty \quad (4.40)$$

Damit haben wir C gefunden. Angenommen, es gäbe kein c wie verlangt. Dann gibt es insbesondere für jedes k ein $\tilde{x}_k \in \mathbb{R}^n$ mit $\|\tilde{x}_k\| < \frac{1}{k} \|\tilde{x}_k\|_\infty$. Es folgt $\tilde{x}_k \neq 0$ und mit $x_k := \tilde{x}_k / \|\tilde{x}_k\|_\infty$ erhalten wir eine Folge (x_k) mit

$$\|x_k\| < \frac{1}{k} \quad \text{und} \quad \|x_k\|_\infty = 1 \quad (4.41)$$

Die Folge (x_k) ist beschränkt (in der durch $\|\cdot\|_\infty$ gegebenen Metrik!) und besitzt nach 4.17 eine konvergente Teilfolge (x_{k_l}) . Deren Grenzpunkt x erfüllt noch $\|x\|_\infty = 1$. Andererseits folgt aus $\|\cdot\| \leq C \cdot \|\cdot\|_\infty$, (4.40), dass $\|x - x_{k_l}\| \rightarrow 0$ für $l \rightarrow \infty$, und dann aus

$$\|x\| \leq \|x - x_{k_l}\| + \|x_{k_l}\| \rightarrow 0 \quad (4.42)$$

dass $\|x\| = 0$, d.h. $x = 0$, ein Widerspruch. $\square$

Theorem/Definition 4.19 (Banachscher Fixpunktsatz). *Es sei (X, d) ein metrischer Raum. Eine Abbildung $T : X \rightarrow X$ heißt kontraktiv (oder eine Kontraktion) falls eine (konstante) reelle Zahl $0 \leq c < 1$ existiert, mit der*

$$d(T(x), T(y)) \leq c \cdot d(x, y) \quad \forall x, y \in X \quad (4.43)$$

· Eine Kontraktion $T : X \rightarrow X$ auf einem nicht-leeren und vollständigen metrischen Raum besitzt genau einen Fixpunkt, d.h. $\exists_1 x_\infty \in X$ mit $T(x_\infty) = x_\infty$.

Beweis. Sei $x_0 \in X$ beliebig. Wir definieren rekursiv für $k \geq 1$ $x_k = T(x_{k-1})$.

Beh.: (x_k) ist eine Cauchy-Folge, d.h. konvergent wegen der Vollständigkeit von X .

Bew.: · Für alle k gilt $d(x_k, x_{k-1}) \leq c^{k-1} \underbrace{d(x_1, x_0)}_{=:D}$

· Daraus folgt für $k > l$ mit Hilfe der Dreieckungleichung

$$\begin{aligned} d(x_k, x_l) &\leq d(x_k, x_{k-1}) + d(x_{k-1}, x_{k-2}) + \cdots + d(x_{l+1}, x_l) \leq \sum_{i=l}^{k-1} c^i \cdot D \\ &\leq c^l \cdot \frac{D}{1-c} \quad (\text{siehe geometrische Reihe im nächsten §}) \end{aligned} \quad (4.44)$$

Wegen $c < 1$ folgt $\lim c^l = 0$ und daraus sofort die Behauptung.

Beh.: Der Grenzpunkt $x_\infty := \lim x_k$ dieser Cauchy-Folge erfüllt $T(x_\infty) = x_\infty$.

Bew.: Für beliebiges k gilt

$$\begin{aligned} d(x_\infty, T(x_\infty)) &\leq d(x_\infty, x_k) + d(x_k, T(x_\infty)) \\ &\leq d(x_\infty, x_k) + c \cdot d(x_{k-1}, x_\infty) \rightarrow 0 \quad \text{für } k \rightarrow \infty \\ &\Rightarrow d(x_\infty, T(x_\infty)) = 0 \end{aligned} \quad (4.45)$$

· Zum Nachweis der Eindeutigkeit sei $\tilde{x}_\infty = T(\tilde{x}_\infty)$. Aus

$$d(x_\infty, \tilde{x}_\infty) = d(T(x_\infty), T(\tilde{x}_\infty)) \leq c \cdot d(x_\infty, \tilde{x}_\infty) \quad (4.46)$$

folgt wegen $c < 1$ $d(x_\infty, \tilde{x}_\infty) = 0$. $\square$

§ 5 Reihen

Begriffsklärung:⁶ Reihen sind ein wichtiger Fall der Vorstellung mathematischer Objekte durch Grenzprozesse. Formal gesehen sind Reihen nichts anderes als Folgen $(s_n)_{n \in \mathbb{N}}$ in normierten Vektorräumen, bei denen man das Augenmerk auf die Zuwächse

$$a_k = s_k - s_{k-1} \quad (5.1)$$

anstatt auf die s_n selber richtet. Dies drückt sich unter anderem darin aus, dass man die a_k als Glieder der Reihe bezeichnet (und nicht die s_n). Die Reihe allerdings heisst konvergent, wenn die Folge der s_n konvergiert. Tatsächlich folgt ja aus unseren Rechenregeln für Folgen

$$\lim a_k = \lim s_k - \lim s_{k-1} = 0 \quad (5.2)$$

falls $\lim s_k = s$ existiert. Die Folge der Glieder einer konvergenten Reihe ist also immer eine Nullfolge. (Die Umkehrung gilt nicht: Eine Reihe kann divergieren, auch wenn die a_k eine Nullfolge bilden.)

· Ein weiterer Unterschied zu Folgen liegt in der Flexibilität bei Index und Startwert. Aus (5.1) folgt für $n = 1, 2, \dots$

$$s_n = a_n + a_{n-1} + \dots + a_1 + s_0 \quad (5.3)$$

Wir werden zunächst meistens $a_0 = s_0$ und $s_n = 0$ für $n < 0$ annehmen, später einmal aber auch Reihen mit nicht-trivialen Gliedern mit negativem Index zulassen (sog. Laurent-Reihen). Änderung von endlich vielen Reihengliedern beeinflusst nicht die Konvergenz, wohl aber den Wert der Reihe.

· Man nennt die $s_n = \sum_{k=0}^n a_k$ auch die Partialsummen der Reihenglieder und schreibt $\sum_{k=0}^{\infty} a_k$ für die Folge (s_n) oder, bei Konvergenz, auch für deren Grenzwert. Der heisst dann auch Summe der Folge (a_k) .

· Unser Ziel in diesem § ist eine Übersicht über Konvergenzkriterien für Reihen sowie die Feststellung einiger Rechenregeln. Wir behandeln zunächst komplexe Reihen, mit einigen anordnungsspezifischen Bemerkungen im rein reellen Fall. Anschliessend diskutieren wir die für die Physik wichtigen Potenzreihen, im etwas allgemeineren Rahmen von Banach-Algebren.

Ein wichtiger Aspekt der Diskussion ist die Abhängigkeit der Konvergenz von Variationen der Reihenglieder.

Beispiel 5.1. Für $z \in \mathbb{C}$ ist die geometrische Reihe

$$\sum_{k=0}^{\infty} z^k = \left(s_n = \sum_{k=0}^n z^k \right)_{n=0,1,\dots} = \left(\frac{1 - z^{n+1}}{1 - z} \right)_{n=0,1,\dots} \quad (5.4)$$

· für $|z| < 1$ konvergent mit Wert $\sum_{k=0}^{\infty} z^k = \frac{1}{1 - z}$. ($z^n \rightarrow 0$ für $n \rightarrow \infty$)

· sonst divergent ((z^k) ist keine Nullfolge für $|z| \geq 1$).

⁶Dieser § ist stark an Kapitel 6 der Analysis I von Königsberger angelehnt.

Beispiel 5.2. Für $s \in \mathbb{Q}$ ist die Dirichlet-Riemann-Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k^s} = \left(\begin{array}{l} \text{Es gibt im Allgemeinen keine geschlos-} \\ \text{sene Form für die Partialsummen.} \end{array} \right) \quad (5.5)$$

- für $s > 1$ konvergent.
- für $s \leq 1$ divergent.

Bew.: Da alle Glieder positiv sind, wächst die Folge der Partialsummen streng monoton.

- Für $s > 1$, $n \in \mathbb{N}$ sei l so, dass $2^l > n$. Dann gilt

$$\begin{aligned} \sum_{k=1}^n \frac{1}{k^s} &\leq \sum_{k=1}^{2^l-1} \frac{1}{k^s} \\ &= 1 + \underbrace{\frac{1}{2^s} + \frac{1}{3^s}}_{\leq 2 \cdot \frac{1}{2^s}} + \underbrace{\frac{1}{4^s} + \dots + \frac{1}{7^s}}_{\leq 2^2 \cdot \frac{1}{2^{2s}}} + \dots + \underbrace{\frac{1}{2^{(l-1)s}} + \dots + \frac{1}{(2^l-1)^s}}_{\leq 2^{l-1} \cdot \frac{1}{2^{(l-1)s}}} \\ &\leq 1 + \frac{1}{2^{s-1}} + \left(\frac{1}{2^{s-1}}\right)^2 + \dots + \left(\frac{1}{2^{s-1}}\right)^{l-1} \\ &\leq \frac{1}{1 - \frac{1}{2^{s-1}}} \quad (\text{geometrische Reihe; } 2^{1-s} < 1 \text{ für } s > 1) \end{aligned} \quad (5.6)$$

Die Folge der Partialsummen ist also monoton wachsend und nach oben beschränkt, also konvergent nach 3.9.

- Für $s \leq 1$ gilt für alle k : $\frac{1}{k^s} \geq \frac{1}{k}$. Dann gilt für $l \in \mathbb{N}$ und $n \geq 2^l$

$$\begin{aligned} \sum_{k=1}^n \frac{1}{k^s} &\geq \sum_{k=1}^{2^l} \frac{1}{k^s} \\ &= 1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq 2 \cdot \frac{1}{4}} + \underbrace{\frac{1}{5} + \dots + \frac{1}{8}}_{\geq 4 \cdot \frac{1}{8}} + \dots + \underbrace{\frac{1}{2^{l-1}} + \dots + \frac{1}{2^l}}_{\geq 2^{l-1} \cdot \frac{1}{2^l}} \\ &= 1 + \frac{l}{2} \rightarrow \infty \quad \text{für } l \rightarrow \infty \end{aligned} \quad (5.7)$$

$\Rightarrow$ Die harmonische Reihe ist divergent.

Diese Beispiele illustrieren bereits die zwei wesentlichen Ideen zum Nachweis von Konvergenz/Divergenz von Reihen: (i) Vergleich mit bekannten Reihen, und (ii) Monotonie für Reihen mit positiven Gliedern. Die meisten Aussagen basieren auf dem

Lemma 5.3 (Dominierte Konvergenz). Ist $(a_k) \subset \mathbb{C}$ und $(c_k) \subset \mathbb{R}_{\geq 0}$ mit:

- (1) $|a_k| \leq c_k \quad \forall k$
- (2) $\sum c_k$ ist konvergent.

Dann gilt: $\sum a_k$ ist konvergent und $\left| \sum_{k=0}^{\infty} a_k \right| \leq \sum_{k=0}^{\infty} c_k$.

§ 5. REIHEN

Beweis. Man prüft das Cauchy-Kriterium für $s_n = \sum_{k=0}^n a_k$, gestützt auf dasjenige für $u_n = \sum_{k=0}^n c_k$: Für $\epsilon > 0$ sei N_ϵ so, dass $|u_n - u_m| < \epsilon$ für $n, m \geq N_\epsilon$. Dann gilt für solche n, m , oBdA mit $n \geq m$:

$$|s_n - s_m| = \left| \sum_{k=m+1}^n a_k \right| \leq \sum_{k=m+1}^n c_k = |u_n - u_m| < \epsilon \quad (5.8)$$

Daraus folgt Konvergenz der Reihe, die Abschätzung für ihren Wert folgt aus $\left| \sum_{k=0}^n a_k \right| \leq \sum_{k=0}^n c_k$ zusammen mit den Rechenregeln für Folgen. $\square$

Definition 5.4. Eine Reihe $\sum a_k$ heisst absolut konvergent, falls die Reihe der Absolutbeträge $\sum |a_k|$ konvergiert.

Nach 5.3 impliziert absolute Konvergenz gewöhnliche Konvergenz. Die Umkehrung gilt aber nicht, wie das folgende Beispiel zeigt:

Beispiel 5.5. Die alternierende harmonische Reihe,

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots \quad (5.9)$$

konvergiert, aber nicht absolut.

Hierbei folgt die Divergenz der Reihe der Absolutbeträge aus (5.7), die Konvergenz von (5.9) aus dem sog. Leibnitz-Kriterium für alternierende Reihen:

Lemma 5.6. Sei (a_k) eine monoton fallende (oder auch wachsende) Nullfolge. Dann konvergiert

$$\sum_{k=1}^{\infty} (-1)^k a_k \quad (5.10)$$

Beweis. Sei $s_n = \sum_{k=1}^n (-1)^k a_k$ die Folge der Partialsummen. Für $n = 2m$ gilt $s_{2m-1} \leq s_{2m}$ sowie

$$[s_{2m+1}, s_{2m+2}] \subset [s_{2m-1}, s_{2m}] \quad \text{und} \quad s_{2m} - s_{2m-1} = a_{2m} \rightarrow 0 \quad (5.11)$$

Die $([s_{2m-1}, s_{2m}])_{m=1,2,\dots}$ bilden also eine Intervallschachtelung.⁷ Es folgt sofort

$$\bigcap_{m=1}^{\infty} [s_{2m-1}, s_{2m}] = \left\{ \lim s_{2m} = \lim s_{2m-1} = \sum_{k=1}^{\infty} (-1)^k a_k \right\} \quad (5.12)$$

$\square$

⁷Eventuell leicht verallgemeinert im Sinne von 4.16. Ein-Punkt Teilmengen von $\mathbb{R}$ werden meistens nicht als Intervalle zugelassen.

Absolute Konvergenz

Witzigerweise klingt im Adjektiv “absolut” noch die Bedeutung an, dass man die Glieder einer absolut konvergenten Reihe beliebig umordnen oder umgruppieren kann, ohne die Konvergenz oder den Wert der Reihe zu beeinflussen. Dies folgt aus dem einfachen

Lemma 5.7. *Eine Reihe $\sum a_k$ ist genau dann absolut konvergent, wenn die Menge*

$$FS = \left\{ \sum_{k \in J} |a_k| \mid J \subset \mathbb{N} \text{ endliche Teilmenge} \right\} \quad (5.13)$$

nach oben beschränkt ist. (In diesem Fall gilt $\sum_{k=1}^{\infty} |a_k| = \sup FS$.)

Beweis. Ist die Reihe absolut konvergent, so ist der Wert von $\sum |a_k|$ klarerweise eine obere Schranke für FS : Für jedes $J \subset \mathbb{N}$ endlich gibt es ein N_ϵ mit $N_\epsilon \geq J$. Dann ist

$$\sum_{k=1}^{\infty} |a_k| \geq \sum_{k=1}^{N_\epsilon} |a_k| \geq \sum_{k \in J} |a_k| \quad (5.14)$$

(Umgekehrt ist für jedes $\epsilon > 0$ $\sum_{k=1}^{\infty} |a_k| - \epsilon$ keine obere Schranke, denn es existiert ein N_ϵ so, dass $\sum_{k \in \{1, \dots, N_\epsilon\}} |a_k| > \sum_{k=1}^{\infty} |a_k| - \epsilon$.)

• Ist umgekehrt FS nach oben beschränkt, so wächst $(\sum_{k=1}^n |a_k|)_{n=1,2,\dots}$ monoton und ist nach oben beschränkt, also konvergent. $\square$

Sei nun $\sum_{k=1}^{\infty} a_k$ eine absolut konvergente Reihe und $(J^{(i)})_{i=1,2,\dots}$ eine (endliche oder unendliche) Familie von (endlichen oder unendlichen) Teilmengen von $\mathbb{N}$ mit $J^{(i)} \cap J^{(j)} = \emptyset$ für $i \neq j$ und $\cup_i J^{(i)} = \mathbb{N}$. Wegen des Lemmas ist dann für jedes i , und jede endliche Teilmenge $J \subset J^{(i)}$ die Summe $\sum_{k \in J} |a_k|$ durch $\sum_{k=1}^{\infty} |a_k|$ beschränkt. Ebenfalls mit dem Lemma folgt daraus, dass bei beliebiger Durchnummerierung $(k_l^{(i)})_{l=1,2,\dots}$ der Elemente von $J^{(i)}$ die Summe über die zugehörigen Glieder von (a_k) (nicht notwendigerweise eine Teilfolge) absolut konvergiert. Wir schreiben $S^{(i)} = \sum_l a_{k_l^{(i)}}$ für deren Wert (Es kann sich hier um eine endliche Summe handeln. Dann ist Konvergenz trivial.), und behaupten $\sum_i S^{(i)} = \sum_{k=1}^{\infty} a_k =: S$, mit absoluter Konvergenz im Reihenfall (unendlich viele i 's).

Bew.: Für $\epsilon > 0$ sei N_ϵ so gross, dass $\sum_{k=N_\epsilon}^{\infty} |a_k| < \epsilon$. Dann gilt $|S - \sum_{k \in J} a_k| < \epsilon$ für jede endliche Teilmenge J von $\mathbb{N}$ mit $\{1, \dots, N_\epsilon\} \subset J$. Für I und L so gross, dass $\cup_{i=1}^I \cup_{l \leq L} \{k_l^{(i)}\} \supset \{1, \dots, N_\epsilon\}$ ist dann

$$\left| S - \sum_{i=1}^I \sum_{l \leq L} a_{k_l^{(i)}} \right| < \epsilon \quad (5.15)$$

Mit $L \rightarrow \infty$ folgt daraus wegen der Konvergenz der $\sum_l a_{k_l^{(i)}}$, dass $S = \sum_i S^{(i)}$. Die Konvergenz ist absolut, da $\sum_{i=1}^I \sum_{l \leq L} |a_{k_l^{(i)}}| \leq \sum_{k=1}^{\infty} |a_k|$ im Grenzübergang $L \rightarrow \infty$ $\sum_{i=1}^I |S^{(i)}| \leq \sum_{k=1}^{\infty} |a_k|$ impliziert. $\square$

§ 5. REIHEN

· In der Vorlesung haben wir den eben gegebenen Beweis des Umordnungssatzes für absolut konvergente Reihen nicht vorgeführt. Stattdessen illustrieren wir die Einschränkung der Aussage durch das folgende Beispiel: Für die umgeordnete alternierende harmonische Reihe

$$1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \cdots + \frac{1}{2k-1} - \frac{1}{2(2k-1)} - \frac{1}{4k} + \cdots \quad (5.16)$$

gilt wegen $\frac{1}{2k-1} - \frac{1}{2(2k-1)} - \frac{1}{4k} = \frac{1}{2} \left(\frac{1}{2k-1} - \frac{1}{2k} \right)$, dass die Summe der ersten $3n$ Glieder gleich

$$\frac{1}{2} \left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2n-1} - \frac{1}{2n} \right) \quad (5.17)$$

d.h. der Hälfte der Summe der ersten $2n$ Glieder der alternierenden harmonischen Reihe ist. Die dazwischenliegenden Partialsummen liegen nicht weiter als $\frac{1}{3n-1}$ bzw. $\frac{1}{3n+1}$ davon entfernt, so dass (5.16) gegen die Hälfte von (5.9) konvergiert.

Konvergenzkriterien

Proposition 5.8 (Wurzelkriterium). *Für eine Folge komplexer Zahlen $(a_k)_{k \in \mathbb{N}}$ sei*

$$L := \limsup |a_k|^{1/k} \quad (5.18)$$

Dann gilt

(i) Für $L < 1$ konvergiert die Reihe $\sum a_k$ absolut.

(ii) Für $L > 1$ divergiert die Reihe.

Zur Erinnerung: Falls $(|a_k|^{1/k})$ unbeschränkt ist, so ist formal $L = \infty$ und die Divergenz der Reihe klar. Sonst ist nach unserer Konstruktion des $\limsup$ in 3.11 für jedes $q > L$

$$|a_k|^{1/k} > q \quad (5.19)$$

für nur endlich viele k , und keine kleinere Zahl hat diese Eigenschaft (siehe (3.28)).

Beweis. · Für $L < q < 1$ sei N so, dass $|a_k|^{1/k} \leq q \ \forall n \geq N$. Dann ist

$$\sum_{k=N}^{\infty} q^k = \frac{q^N}{1-q} \quad (5.20)$$

eine konvergente Majorante für $\sum_{k=N}^{\infty} a_k$. Mit 5.3 folgt daraus (i).

· Falls $\infty > L > 1$, so gibt es $\epsilon > 0$ mit $L - \epsilon > 1$, und unendlich vielen k mit $|a_k|^{1/k} \geq L - \epsilon > 1$. Die a_k sind dann nicht einmal eine Nullfolge, also kann die Reihe $\sum a_k$ nicht konvergieren. Dies impliziert (ii). $\square$

Korollar 5.9 (Quotientenkriterium). *Es existiere $\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| =: q$. Dann gilt*

(i) Für $q < 1$ konvergiert $\sum a_k$ absolut.

(ii) Für $q > 1$ divergiert die Reihe.

Beweis. (Versteckt in der Annahme ist die Aussage $a_k \neq 0$ für fast alle k ; $q = 0$ ist hingegen erlaubt.)

Für (i) zeigen wir $\limsup |a_k|^{1/k} \leq q$. Für $q' > q$ sei K so, dass

$$\left| \frac{a_{k+1}}{a_k} \right| < q' \quad \text{für } k \geq K \quad (5.21)$$

Dann gilt für solche k

$$\begin{aligned} |a_k| &= \left| \frac{a_k}{a_{k-1}} \right| \cdots \left| \frac{a_{K+1}}{a_K} \right| \cdot |a_K| \leq (q')^{k-K} \cdot |a_K| \\ \Rightarrow |a_k|^{1/k} &\leq q' \cdot (q'^{-K} |a_K|)^{1/k} \end{aligned} \quad (5.22)$$

Wegen 3.5(ii) folgt daraus $\limsup |a_k|^{1/k} \leq q'$. Da dies für alle $q' > q$ gilt, folgt $\limsup |a_k|^{1/k} \leq q$.

· Falls $q > 1$, so wachsen die $|a_k|$ ab einem bestimmten Index streng monoton, die a_k bilden also nicht einmal eine Nullfolge. $\square$

Potenzreihen

Es handelt sich um Reihen der Form

$$P(z) = \sum_{n=0}^{\infty} a_n z^n \quad (5.23)$$

Sie dienen der Vorstellung von Funktionen einer als “klein” gedachten Veränderlichen z durch sukzessives Aufsummieren einfacher Potenzen von z mit geeigneten Koeffizienten. (Wir interessieren uns zunächst wieder für $z, a_n \in \mathbb{C}$.)

· Die Bedeutung von Potenzreihenentwicklungen in der Physik kann unmöglich (die Betonung liegt auf unmöglich) überschätzt werden. Irgendwann einmal stösst man dort auf divergente Potenzreihen, während wir hier nur den konvergenten Fall behandeln.

· Die zentrale Eigenschaft von Potenzreihen ist die Existenz des sog. Konvergenzkreises.

Lemma 5.10. *Angenommen, $P(z)$ konvergiert in einem Punkt $z_0 \in \mathbb{C}$ mit $z_0 \neq 0$. Dann konvergiert $P(z)$ in jedem Punkt $z \in \mathbb{C}$ mit $|z| < |z_0|$ absolut.*

Beweis. Es existiert ein S mit $|a_n z_0^n| \leq S$ für all n . Daraus folgt für $|z| < |z_0|$ mit $q = \left| \frac{z}{z_0} \right| < 1$:

$$|a_n z^n| = |a_n z_0^n| \cdot \left| \frac{z}{z_0} \right|^n \leq S q^n \quad (5.24)$$

$P(z)$ besitzt also die konvergente Majorante $\sum S q^n = \frac{S}{1-q}$ und konvergiert daher absolut. $\square$

Für gegebene Potenzreihe $P(z)$ setzen wir dann

$$R = \sup \{ r \in \mathbb{R} \mid P(r) \text{ konvergiert} \} \quad (5.25)$$

(Für $r = 0$ konvergiert die Reihe trivialerweise. Daher ist $R \geq 0$. Ist die Menge rechts unbeschränkt, so setzen wir formal $R = \infty$.) Wir nennen R den Konvergenzradius und $B_R(0)$ die Konvergenzscheibe, denn es gilt:

§ 5. REIHEN

Theorem 5.11. Für $|z| < R$ konvergiert $P(z)$ absolut.

· Für $|z| > R$ divergiert $P(z)$.

Beweis. Für $|z| < R$ existiert ein $r' > 0$ mit $|z| < r' < R$. Da $P(r')$ für jedes $0 < r' < R$ konvergiert, impliziert 5.10 die absolute Konvergenz von $P(z)$.

· Wäre für ein $|z| > R$ die Reihe $P(z)$ konvergent, so wäre wegen Lemma 5.10 für jedes $R' < |z|$ die Reihe $P(R')$ ebenfalls konvergent (sogar absolut). Wählt man ein solches R' mit $|z| > R' > R$, so erhält man einen Widerspruch zur Supremumsdefinition von R . $\square$

Durch Anwenden von Wurzel-/Quotientenkriterium 5.8, 5.9 erhalten wir dann die nützlichen Formeln zur Berechnung des Konvergenzradiuses einer Potenzreihe $\sum a_n z^n$

$$R = \frac{1}{\limsup |a_n|^{1/n}}$$

$$R = \frac{1}{\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|} \quad \text{falls der Grenzwert existiert} \quad (5.26)$$

In diesen Formeln gilt die Vereinbarung, dass $\frac{1}{0} = \infty$ und $\frac{1}{\infty} = 0$.

Lemma 5.12 (Restabschätzung). Für jedes $N \in \mathbb{N}$ und $r < R$ existiert eine Konstante $C \geq 0$ so, dass für alle $|z| \leq r$

$$\left| \sum_{n=0}^{\infty} a_n z^n - \sum_{n=0}^N a_n z^n \right| \leq C \cdot |z|^{N+1} \quad (5.27)$$

Beweis. Wähle ρ mit $r < \rho < R$. Dann existiert eine Konstante S so, dass $|a_n \rho^n| \leq S \forall n$. Es folgt

$$\sum_{n=N+1}^{\infty} |a_n z^n| \leq \sum_{n=N+1}^{\infty} |a_n \rho^n| \frac{|z|^n}{\rho^n} \leq S \cdot \frac{|z|^{N+1}}{\rho^{N+1}} \sum_{n=0}^{\infty} \frac{r^n}{\rho^n} = |z|^{N+1} \cdot \frac{S}{\rho^{N+1}(1 - r/\rho)} \quad (5.28)$$

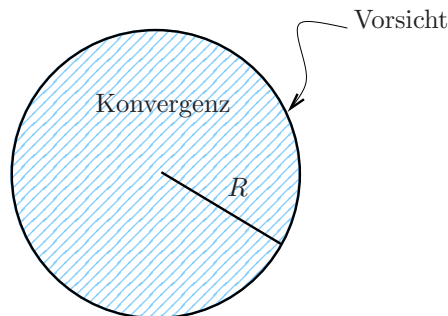
$\square$

Merkwissen

I. Unter absoluter (dominierter) Konvergenz sind alle vernünftigen Manipulationen erlaubt, im Allgemeinen aber muss man vorsichtig sein.

II. Jede Potenzreihe kommt mit einem Konvergenzradius (der auch 0 oder ∞ sein kann). Innerhalb der Konvergenzscheibe konvergiert eine Potenzreihe absolut, und der Fehler lässt sich durch den ersten weggelassenen Term abschätzen. Auf der Konvergenzkreislinie ist wieder Vorsicht geboten.

Divergenz



III. Die Definitionen und Konvergenz-Kriterien übertragen sich mit notwendigen Änderungen auf Reihen in einem vollständigen normierten Vektorraum $(V, \|\cdot\|)$ über $\mathbb{R}$ oder $\mathbb{C}$. In diesem Zusammenhang heisst etwa eine Reihe $\sum v_k$ mit $v_k \in V$ absolut konvergent (man sagt auch konvergent in der Norm), wenn die Reihe der Normen $\sum \|v_k\|$ konvergiert. Insbesondere gelten Lemma 5.3 und die Umordnungsregeln sinngemäss. Das Leibnitz-Kriterium 5.6 gilt natürlich nur in $\mathbb{R}$!

Häufig anzutreffen sind Potenzreihen, in denen die Veränderliche einer nicht notwendig kommutativen Banach-Algebra $(\mathcal{A}, \|\cdot\|)$ entstammt (z.B. $\text{Mat}_{n \times n}(\mathbb{R})$), während die Koeffizienten im Grundkörper bleiben.⁸ Für eine solche Potenzreihe ist der Konvergenzradius weiter durch (5.26) gegeben. Ein Unterschied ist, dass die Potenzreihe auch einmal ausserhalb der Konvergenzkugel $B_R(0) \subset \mathcal{A}$ konvergieren kann (Übungsaufgabe).

Beispiele

- Geometrische (oder Neumann-)Reihe

$$\sum_{n=0}^{\infty} x^n = (1 - x)^{-1} \quad \text{für } \|x\| < 1 \quad (5.29)$$

- Logarithmus-(oder Mercator-)Reihe

$$\sum_{n=1}^{\infty} (-1)^{k-1} \frac{x^k}{k} \quad R = 1 \quad (5.30)$$

- Exponentialreihe

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad R = \infty \quad (5.31)$$

- Sinus-Reihe

$$\sin(x) := \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad R = \infty \quad (5.32)$$

- Cosinus-Reihe

$$\cos(x) := \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \quad R = \infty \quad (5.33)$$

- Euler-Formel (für Grundkörper = $\mathbb{C}$)

$$\exp(ix) = \cos(x) + i \sin(x) \quad (5.34)$$

Proposition 5.13. *Für all $x, y \in \mathcal{A}$ mit $xy = yx$ gilt*

$$\exp(x) \exp(y) = \exp(x + y) \quad (5.35)$$

und daraus abgeleitet die bekannten Additionstheoreme der via (5.34) definierten trigonometrischen Funktionen.

⁸Wir nehmen die Existenz eines Einselements $1 \in \mathcal{A}$ an, sodass $x^0 = 1$ erklärt ist.

§ 5. REIHEN

Lemma 5.14. Sei $(x_n) \subset \mathcal{A}$ eine Folge mit $\lim_{n \rightarrow \infty} x_n = x$. Dann gilt

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x_n}{n}\right)^n = \exp(x) \quad (5.36)$$

Beweis von 5.13. (unter der Annahme, dass 5.14 gilt)

$$\begin{aligned} \exp(x) \exp(y) &= \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \lim_{n \rightarrow \infty} \left(1 + \frac{y}{n}\right)^n = (\text{Rechenregeln für Folgen}) \\ &= \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \left(1 + \frac{y}{n}\right)^n \quad (\text{wegen } xy = yx) \\ &= \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n} + \frac{y}{n} + \frac{xy}{n^2}\right)^n \end{aligned} \quad (5.37)$$

Wegen $\lim_{n \rightarrow \infty} \left(x + y + \frac{xy}{n}\right) = x + y$ folgt die Behauptung. $\square$

Beweis von 5.14. Für $\epsilon > 0$ sei N so gross, dass

$$\sum_{k=N+1}^{\infty} \frac{(\|x\| + 1)^k}{k!} < \epsilon \quad \text{und} \quad \|x_n\| < \|x\| + 1 \quad \forall n > N \quad (5.38)$$

Dann gilt für $n \geq N$:

$$\begin{aligned} &\left\| \left(1 + \frac{x_n}{n}\right)^n - \exp(x) \right\| \\ &\leq \underbrace{\sum_{k=0}^N \left\| \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \right\|}_{\rightarrow 0 \text{ für } n \rightarrow \infty} + \underbrace{\sum_{k=N+1}^n \left\| \binom{n}{k} \frac{x_n^k}{n^k} \right\|}_{\leq \frac{(\|x\|+1)^k}{k!} < \epsilon} + \underbrace{\sum_{k=N+1}^{\infty} \frac{\|x^k\|}{k!}}_{\leq \frac{(\|x\|+1)^k}{k!} < \epsilon} \end{aligned} \quad (5.39)$$

wobei wir benutzt haben, dass

$$\binom{n}{k} \frac{1}{n^k} = \frac{1}{k!} \prod_{i=0}^{k-1} \left(1 - \frac{i}{n}\right) \begin{cases} \leq \frac{1}{k!} & \text{für den mittleren Term} \\ \xrightarrow{n \rightarrow \infty} \frac{1}{k!} & \text{für den ersten Term} \end{cases} \quad (5.40)$$

Es gibt also ein $\tilde{N} > N$ so, dass für $n \geq \tilde{N}$ der erste Term auf der rechten Seite von (5.39) $< \epsilon$ ist, alles zusammen $< 3\epsilon$. $\square$

• Insbesondere ist $\exp(x) \exp(-x) = \exp(0) = 1$ und daher

$$\exp(x) \neq 0 \quad \forall x \quad (5.41)$$

KAPITEL 3

STETIGKEIT

Anfangen in diesem Kapitel gilt unser Hauptinteresse jetzt nicht mehr nur einzelnen reellen oder komplexen Zahlen oder Punktfolgen in metrischen Räumen, sondern vielmehr denjenigen Zusammenhängen zwischen solchen Objekten, welche durch mathematische Funktionen (a.k.a. Abbildungen) $F : (D, d_X|_D) \rightarrow (Y, d_Y)$ gegeben sind, nämlich als Zuordnung eines einzigen und eindeutigen Wertes $F(x)$ (in einem metrischen Raum Y) zu jedem Element x eines Definitionsbereichs D , üblicherweise Teilmenge eines metrischen Raumes X . Recht bald werden wir dann Funktionen als eigenständige mathematischen Objekte auffassen, und in den zugehörigen Funktionenräumen Analysis betreiben.

§ 6 Stetige Funktionen

Es ist eine Grundidee der Physik, dass “kleine Änderungen der Ursachen oder Parameter kleine Folgen” haben. Diese Vorstellung lässt sich ziemlich präzise mit der ersten Formulierung der Stetigkeit fassen.

Definition 6.1. Für metrische Räume X, Y heisst eine Funktion $F : X \rightarrow Y$ *Lipschitz-stetig*, falls eine Konstante $L \geq 0$ existiert, mit der für alle $x_1, x_2 \in X$ gilt:

$$d_Y(F(x_1), F(x_2)) \leq L \cdot d_X(x_1, x_2) \quad (6.1)$$

In der Tat garantiert ja dann ein kleiner Abstand der Urbilder x_1 und x_2 Kontrolle über den Abstand zwischen den Bildpunkten. Je nach Wert der “Lipschitz-Konstanten” L kann es aber vorkommen, dass $d_Y(F(x_1), F(x_2))$ noch nicht für alle Zwecke “klein genug” ist. Im Allgemeinen ist es daher besser, die Beweislast gleich umzukehren, und auch die Definition zunächst lokal zu treffen. Wir lassen ab jetzt häufiger mal die Indices von den Abstandsfunktionen weg.

Definition 6.2. Eine Funktion $F : X \rightarrow Y$ zwischen metrischen Räumen heisst stetig im Punkt $x_0 \in X$, falls für jedes $\epsilon > 0$ ein $\delta > 0$ existiert so, dass

$$d(F(x), F(x_0)) < \epsilon \quad \text{falls } d(x, x_0) < \delta \quad (6.2)$$

· Man nennt F stetig auf X , wenn F stetig in jedem $x_0 \in X$ ist.

Lemma 6.3. *Eine Lipschitz-stetige Funktion ist stetig.*

Beweis. Für $L = 0$ gibt es nichts zu tun. Andernfalls erfüllt, für jedes x_0 , $\delta = \frac{\epsilon}{L}$ die geforderte Bedingung. $\square$

Beispiel 6.4. · Die Stufenfunktion

$$\mathbb{R} \ni x \mapsto \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (6.3)$$

ist ausser in $x_0 = 0$ überall stetig..

· Für $X = \mathbb{R}, \mathbb{C}$, oder eine normierte Algebra sind Polynome stetig. Zum Beispiel gilt für gegebenes x_1, x_2 die Abschätzung

$$\begin{aligned} \|x_1^2 - x_2^2\| &= \|x_1(x_1 - x_2) + (x_1 - x_2)x_2\| \\ &\leq \|x_1\| \cdot \|x_1 - x_2\| + \|x_1 - x_2\| \cdot \|x_2\| \\ &\leq (\|x_1\| + \|x_2\|) \cdot \|x_1 - x_2\| \end{aligned} \quad (6.4)$$

Dies zeigt: $x \mapsto x^2$ ist Lipschitz-stetig in jeder Kugel $B_R(0)$. Allerdings wächst die Lipschitz-Konstante mit R und die Funktion ist daher nicht “global” Lipschitz-stetig. Insbesondere aber ist sie überall stetig.

· Die Verkettung $X \xrightarrow{F} Y \xrightarrow{G} Z$ zweier stetiger Funktionen ist stetig: Für $x_0 \in X$, $\zeta > 0$ sei $y_0 := F(x_0) \in Y$ und $\epsilon > 0$ so, dass $d_Z(G(y), G(y_0)) < \zeta$ für $d_Y(y, y_0) < \epsilon$. Sodann sei $\delta > 0$ so, dass $d_Y(F(x), F(x_0)) < \epsilon$ für $d_X(x, x_0) < \delta$. Dann folgt $d_Z(G(F(x)), G(F(x_0))) < \zeta$ falls $d_X(x, x_0) < \delta$.

· Lineare Abbildungen zwischen endlich-dimensionalen normierten Vektorräumen über $\mathbb{R}$ oder $\mathbb{C}$ sind Lipschitz-stetig: Die Abschätzung (i) in (4.28) besagt nichts anderes, als dass die Operatornorm eine globale Lipschitz-Konstante ist.

· Der Absolutbetrag $|\cdot| : \mathbb{C} \rightarrow \mathbb{R}_{\geq 0}$ ist die wohl wichtigste nicht-algebraische stetige Funktion: Nach der Dreiecksungleichung gilt $||z_1| - |z_2|| \leq |z_1 - z_2|$, also ist 1 eine globale Lipschitz-Konstante.

· Allgemeiner ist für jeden metrischen Raum die Abstandsfunktion $d_X : X \times X \rightarrow \mathbb{R}_{\geq 0}$ stetig als Funktion auf dem Cartesischen Produkt $X \times X$, ausgestattet mit der Produktmetrik

$$d_{X \times X}((x_1, y_1), (x_2, y_2)) := \|(d_X(x_1, x_2), d_X(y_1, y_2))\|_{\mathbb{R}^2} \quad (6.5)$$

für irgendeine Norm $\|\cdot\|_{\mathbb{R}^2}$ auf $\mathbb{R}^2$:

$$\begin{aligned} |d_X(x_1, y_1) - d_X(x_2, y_2)| &\leq |d_X(x_1, y_1) - d_X(x_2, y_1)| + |d_X(x_2, y_1) - d_X(x_2, y_2)| \\ (\text{Dreiecksungleichung in } X) \quad &\leq d_X(x_1, x_2) + d_X(y_1, y_2) \\ \left(\begin{array}{l} (\xi^1, \xi^2) \mapsto |\xi^1| + |\xi^2| \text{ ist eine Norm} \\ \text{auf } \mathbb{R}^2 \text{ und alle Normen auf } \mathbb{R}^2 \\ \text{sind äquivalent} \end{array} \right) &\leq C \cdot \|(d_X(x_1, x_2), d_X(y_1, y_2))\|_{\mathbb{R}^2} \end{aligned} \quad (6.6)$$

für eine geeignete Konstante C .

· Ein exotischeres Beispiel: Die Funktion

$$\mathbb{R} \ni x \mapsto f(x) := \begin{cases} \frac{1}{q} & \text{für } x = \frac{p}{q} \in \mathbb{Q} \setminus \{0\} \text{ mit } (p, q) = 1, q > 0 \\ 1 & \text{für } x = 0 \\ 0 & \text{sonst (i.e., } x \in \mathbb{R} \setminus \mathbb{Q}) \end{cases} \quad (6.7)$$

§6. STETIGE FUNKTIONEN

ist stetig in jedem irrationalen Punkt $x_0 \in \mathbb{R} \setminus \mathbb{Q}$ (wähle δ so klein, dass $(x_0 - \delta, x_0 + \delta)$ keine rationalen Zahlen mit Nenner $1/q \geq \epsilon$ enthält), aber unstetig in jedem $x_0 \in \mathbb{Q}$ (für jedes $\delta > 0$ ist $x_0 + \frac{\sqrt{2}}{N} \in (x_0 - \delta, x_0 + \delta) \setminus \mathbb{Q}$ für $N > 2/\delta$).

• Die Restabschätzung 5.12 zeigt, dass für eine Potenzreihe mit Konvergenzradius R die Zuordnung $B_R(0) \ni x \mapsto \sum a_n x^n$ zumindest in $x_0 = 0$ stetig ist. [Tatsächlich sind Potenzreihen, wie wir sehen werden, sogar in ganz $B_R(0)$ stetig. Wegen der für $z, w \in \mathbb{C}$ gültigen Funktionalgleichung $\exp(z + w) = \exp(z) \exp(w)$ können wir aber schon jetzt folgern, dass $\exp$ auf ganz $\mathbb{C}$ stetig ist: $|\exp(z_1) - \exp(z_2)| = |\exp(z_1)| \cdot |\exp(z_2 - z_1) - 1| \leq |\exp(z_1)| \cdot C \cdot |z_1 - z_2|$.]

Man beachte, dass die obige Definition der Stetigkeit auch ohne Vollständigkeit sinnvoll ist. Das Folgenkriterium zeigt, dass Rechenregeln für konvergente Folgen sehr eng mit Stetigkeit von Funktionen verwandt ist.

Proposition 6.5. *Eine Funktion $F : X \rightarrow Y$ ist genau dann stetig in $x_0 \in X$, wenn für jede gegen x_0 konvergente Folge $(x_n)_{n \in \mathbb{N}} \subset X$ die Bildfolge $(F(x_n))_{n \in \mathbb{N}} \subset Y$ konvergiert mit dem Grenzwert*

$$\lim F(x_n) = F(x_0) = F(\lim x_n) \quad (6.8)$$

Beweis. “ $\Rightarrow$ ”: Für $\epsilon > 0$ sei $\delta > 0$ so, dass $d(F(x_0), F(y)) < \epsilon$ für $d(x_0, y) < \delta$. Sodann sei N_ϵ so, dass $d(x_0, x_n) < \delta$ für $n \geq N_\epsilon$. Dann gilt für alle $n \geq N_\epsilon$: $d(F(x_0), F(x_n)) < \epsilon$.

“ $\Leftarrow$ ”: Angenommen, es gäbe für ein gewisses $\epsilon_* > 0$ kein $\delta > 0$ wie benötigt. Dann gäbe es insbesondere für jedes n ein x_n mit $d(x_0, x_n) < \frac{1}{n}$, aber $d(F(x_0), F(x_n)) \geq \epsilon_*$. Dann konvergiert die Folge (x_n) gegen x_0 , aber $(F(x_n))$ nicht gegen $F(x_0)$, im Widerspruch zur Voraussetzung. $\square$

Für die Definition der Differenzierbarkeit im nächsten Kapitel werden wir den folgenden Begriff der “stetigen Fortsetzung in einen Punkt” benutzen.

Definition 6.6. Seien X, Y metrische Räume, $D \subset X$ eine beliebige Teilmenge, $F : D \rightarrow Y$ eine Funktion und $x_0 \in X$. Man sagt, F hat eine stetige Fortsetzung in x_0 , falls eine Funktion $\bar{F} : D \cup \{x_0\} \rightarrow Y$ existiert, welche in x_0 stetig ist und auf $D \setminus \{x_0\}$ mit F übereinstimmt.

Bemerkungen. Typischerweise ist $x_0 \notin D$, muss aber nicht. Häufig ist F auf $D \setminus \{x_0\}$ bereits stetig.

• F besitzt eine stetige Fortsetzung in x_0 genau dann, wenn für jede Folge $(x_n) \subset D \setminus \{x_0\}$ mit $x_n \rightarrow x_0$ die Bildfolge $F(x_n)$ in Y konvergiert.

Beweis. “ $\Rightarrow$ ”: Ist $\bar{F}$ eine stetige Fortsetzung nach x_0 , so gilt für jede Folge $(x_n) \subset D \setminus \{x_0\}$: $\bar{F}(x_n) = F(x_n) \forall n$. Konvergiert dann (x_n) gegen x_0 , so folgt $F(x_n) = \bar{F}(x_n) \rightarrow \bar{F}(x_0)$ aus 6.5.

“ $\Leftarrow$ ”: Bei der Umkehrung unterscheiden wir, ob eine gegen x_0 konvergente Folge $(x_n) \subset D \setminus \{x_0\}$ existiert oder nicht. Falls nicht, so wird F mit beliebigem $\bar{F}(x_0)$ stetig in x_0 : Fast alle Glieder einer gegen x_0 konvergente Folge in D müssen mit x_0 zusammenfallen.

Falls eine gegen x_0 konvergente Folge $(x_n) \subset D \setminus \{x_0\}$ existiert, so setzen wir $\bar{F}(x_0) := \lim F(x_n)$ für eine beliebige solche Folge.

Beh.: $\bar{F}$ ist stetig in x_0 .

Bew.: (i) Sei zunächst $(\tilde{x}_n) \subset D \setminus \{x_0\}$ eine weitere gegen x_0 konvergente Folge. Dann konvergiert die “Reissverschlussfolge” (r_n) mit $r_{2n-1} = x_n$, $r_{2n} = \tilde{x}_n$ ebenfalls gegen x_0 . Nach Voraussetzung konvergiert also die Folge $(F(r_n))$, und zwar wegen 3.8 gegen $\bar{F}(x_0)$. Dann konvergiert auch $F(\tilde{x}_n)$ gegen diesen Wert.

(ii) Sei schliesslich $(\bar{x}_n)_{n \in \mathbb{N}} \subset D$ eine beliebige gegen x_0 konvergente Folge. Gilt $\bar{x}_n = x_0$ für fast alle n , so konvergiert $\bar{F}(\bar{x}_n)$ trivialerweise gegen $\bar{F}(x_0)$. Andernfalls sei $(\tilde{x}_k)$ die Teilfolge von $(\bar{x}_n)$, bei welcher die Folgenglieder weggelassen wurden, die mit x_0 zusammenfallen. Dann konvergiert wegen (i) $\bar{F}(\tilde{x}_k) = F(\tilde{x}_k)$ gegen $\bar{F}(x_0)$. Nach Wiedereinfügen der Glieder welche gleich x_0 sind, folgt $\bar{F}(\bar{x}_n) \rightarrow \bar{F}(x_0)$.

Mit 6.5 folgt, dass $\bar{F}$ in x_0 stetig ist. $\square$

· Aus dem Beweis ist klar, dass falls überhaupt eine Folge $(x_n) \subset D \setminus \{x_0\}$ mit $x_n \rightarrow x_0$ existiert (man sagt dann: x_0 ist ein Häufungspunkt von D), so ist eine stetige Fortsetzung eindeutig. Wir reden in diesem Fall vom “Grenzwert von F in x_0 ” und schreiben

$$\lim_{x \rightarrow x_0} F(x) = \bar{F}(x_0) \quad (6.9)$$

· Ein paar Beispiele: ($x \in \mathbb{R}$)

$$\lim_{x \rightarrow 1} \frac{\sqrt{x} - 1}{x - 1} = \frac{1}{2} \quad (6.10)$$

(Denn $\frac{1}{\sqrt{x+1}}$ ist eine stetige Fortsetzung in 1!)

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1 \quad (6.11)$$

(Benutze die Potenzreihe (5.32)!)

· Die Funktion f aus (6.7) hat den Grenzwert 0 in jedem rationalen Punkt $x \in \mathbb{Q}$ (und natürlich in jedem irrationalen, wo sie stetig ist). Eine Variante ist

$$g(x) := \begin{cases} x & \text{für } x \in \mathbb{R} \setminus \mathbb{Q} \\ 0 & \text{für } x \in \mathbb{Q} \end{cases} \quad (6.12)$$

Dann ist 0 der einzige Punkt, in dem g einen Grenzwert hat oder stetig ist.

· Für metrische Räume X, Y schreiben wir $\mathcal{C}(X, Y)$ für die Menge der stetigen Funktion auf X mit Werten in Y . Sind auf Y mit der metrischen Struktur verträgliche algebraische Operationen definiert (ist z.B. $Y = V$ ein normierter Vektorraum, ausgerüstet mit der induzierten Abstandsfunktion), so übertragen sich diese Operationen auf $\mathcal{C}(X, Y)$, und es gelten relativ offensichtliche Rechenregeln für Grenzwerte. Die Idee ist einfach, dass “Verträglichkeit” der algebraischen Operationen mit der Abstandsfunktion ihre Stetigkeit impliziert. Für präzisere Formulierungen und weitere Rechenregeln für stetige Funktionen siehe Lehrbücher.

· Im Falle eines angeordneten Definitionsbereichs, insbesondere für $X = \mathbb{R}$, machen die Begriffe des links- bzw. rechtsseitigen Grenzwertes Sinn: Für $D \subset \mathbb{R}$, $f : D \rightarrow Y$ und $x_0 \in \mathbb{R}$ schreiben wir (falls x_0 ein links-/rechtsseitiger Häufungspunkt von D ist, und der Grenzwert existiert) $\lim_{x \uparrow x_0} f$ für $\lim_{x \rightarrow x_0} f|_{D \cap (-\infty, x_0)}$ und $\lim_{x \downarrow x_0} f$ für $\lim_{x \rightarrow x_0} f|_{D \cap (x_0, \infty)}$

§ 7 Wertsätze

Zu den wichtigsten Anwendungen der Stetigkeit von Funktionen zählt der Nachweis, dass solche Funktionen bestimmte Werte annehmen müssen. Insbesondere kann die Stetigkeit für den Beweis der Existenz von Lösungen gewisser Klassen von Problemstellungen benutzt werden.

Unsere erste Klasse betrifft stetige Funktionen mit reellen Argumenten und reellen Werten.

Theorem 7.1 (Zwischenwertsatz). *Sei $[a, b]$ ein Intervall und $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann existiert für jedes $t \in [0, 1]$ ein $c \in [a, b]$ mit*

$$f(c) = tf(a) + (1 - t)f(b) =: f_t \quad (7.1)$$

Beweis. Für $t \in [0, 1]$ durchläuft f_t alle Werte “zwischen” $f(a)$ und $f(b)$. OBdA sei $f(a) \leq f_t \leq f(b)$. Setze $[a_0, b_0] = [a, b]$ und rekursiv für $n > 0$: $m_{n-1} := \frac{a_{n-1} + b_{n-1}}{2}$ und $[a_n, b_n] = [a_{n-1}, m_{n-1}]$ oder $[m_{n-1}, b_{n-1}]$ je nachdem, ob $f(m_{n-1}) \geq f_t$ oder $f(m_{n-1}) < f_t$. Dies ergibt eine Intervallschachtelung $([a_n, b_n])$ mit der zusätzlichen Eigenschaft, dass

$$f(a_n) \leq f_t \leq f(b_n) \quad \forall n \quad (7.2)$$

Für $c = \lim a_n = \lim b_n$ folgt aus der Stetigkeit $f(c) = \lim f(a_n) = \lim f(b_n)$ und daher wegen (7.2) $f(c) = f_t$ (Sandwich-Lemma 3.6). $\square$

Korollar 7.2. *Sei $I \subset \mathbb{R}$ ein Intervall (geschlossen, offen, beschränkt oder unbeschränkt) und $f : I \rightarrow \mathbb{R}$ streng monoton und stetig. Dann bildet f I bijektiv auf ein Intervall J ab, und die Umkehrfunktion $f^{-1} : J \rightarrow I$ ist ebenfalls streng monoton und stetig.*

Beweis. Injektivität folgt aus der streng oBdA wachsenden Monotonie: $x_1 > x_2 \Rightarrow f(x_1) > f(x_2)$. Dass $J := f(I)$ ein Intervall ist, folgt aus 7.1 (sogar unabhängig von der Monotonie). Zum Nachweis der Stetigkeit der Umkehrfunktion f^{-1} bei $y_0 = f(x_0) \in J$ zeigen wir, dass für jedes $\epsilon > 0$ ein $\delta > 0$ existiert so, dass

$$f^{-1}((y_0 - \delta, y_0 + \delta) \cap J) \subset (x_0 - \epsilon, x_0 + \epsilon) \cap I =: I_{x_0, \epsilon} \quad (7.3)$$

Ist x_0 kein Randpunkt von I , so enthält das Intervall $I_{x_0, \epsilon}$ Punkte sowohl grösser als auch kleiner als x_0 . Wegen der Monotonie enthält daher das Intervall $f(I_{x_0, \epsilon}) \subset J$ Werte sowohl kleiner als auch grösser als y_0 , und damit ein ganzes offenes Intervall $J_{y_0, \delta} := (y_0 - \delta, y_0 + \delta)$ für ein $\delta > 0$.

Ist x_0 ein unterer Randpunkt von I , enthält also $I_{x_0, \epsilon}$ zwar Punkte grösser als x_0 , aber keine kleineren, so ist y_0 ein unterer Randpunkt von J , so dass ein $\delta > 0$ existiert mit $J_{y_0, \delta} = (y_0 - \delta, y_0 + \delta) \cap J = [y_0, y_0 + \delta) \subset f(I_{x_0, \epsilon})$.

Mit den notwendigen Änderungen für einen oberen Randpunkt gilt in jedem Fall $f^{-1}(J_{y_0, \delta}) \subset I_{x_0, \epsilon}$. $\square$

· Beispielsweise sind für $n \in \mathbb{N}$ die Wurfelfunktionen $[0, \infty) \ni x \mapsto x^{1/n} \in [0, \infty)$ als Umkehrung der Potenzfunktionen bijektiv und stetig, und durch Verkettung damit für jedes $s \in \mathbb{Q}$ auch $x \mapsto x^s$.

· Das bisher Gelernte lässt sich aber auch schon für die Rekonstruktion einiger interessanter Aussagen über die wichtigsten transzendenten Funktionen anwenden. Wir erinnern dazu an die Definition der Exponentialfunktion $\exp : \mathbb{C} \rightarrow \mathbb{C}$ durch die auf ganz $\mathbb{C}$ konvergente Potenzreihe (5.31).

Proposition/Definition 7.3. (i) Für $x \in \mathbb{R}$ ist $\exp(x) > 0$

(ii) $\exp : \mathbb{R} \rightarrow \mathbb{R}_{>0}$ ist streng monoton wachsend, surjektiv und stetig.

Die gemäss 7.2 stetige Umkehrfunktion $\ln : \mathbb{R}_{>0} \rightarrow \mathbb{R}$ heisst der natürliche Logarithmus. Er ist ebenfalls streng monoton wachsend und erfüllt die Funktionalgleichung

$$\ln x_1 x_2 = \ln x_1 + \ln x_2 \quad \text{für } x_1, x_2 > 0 \quad (7.4)$$

und erlaubt die Fortsetzung der Potenzfunktionen $x \mapsto x^s = \exp(\ln x^s) = \exp(s \ln x)$ ($s \in \mathbb{Q}$) auf reelle Exponenten. Für $x > 0$, $r \in \mathbb{R}$ setzen wir

$$x^r := \exp(r \ln x) \quad (7.5)$$

Beweis. (i) Ganz allgemein nehmen Potenzreihen mit reellen Koeffizienten bei reellem Argument reelle Werte an. $\exp(x) = (\exp(x/2))^2 > 0$ folgt damit aus (1.6) zusammen mit (5.41).

(ii) Für $x > 0$ folgt direkt aus der Potenzreihe $\exp(x) > 1 + x > 1$. Mit der Funktionalgleichung $\exp(x_1) = \exp(x_1 - x_2) \exp(x_2)$ folgt daraus die Monotonie. Mit $e := \exp(1) > 1^9$ folgt mittels 1.9, dass $\{e^n | n \in \mathbb{N}\}$ nicht nach oben beschränkt ist, und weil noch $\exp(-x) = 1/\exp(x)$ folgt $\exp(\mathbb{R}) = (0, \infty)$. (Stetigkeit hatten wir schon am Ende von 6.4 festgehalten.) Die Funktionalgleichung ist die Umkehrung derer für $\exp$. $\square$

Post quantitates exponentiales considerari debent arcus circulares eorumque sinus et cosinus, quia ex ipsis exponentialibus, quando imaginariis quantitatibus involvuntur, proveniunt.¹⁰ Auch diese Funktionen hatten wir ja für ganz allgemeine Argumente via Potenzreihen (5.32), (5.33) eingeführt und bemerkt, dass aus der Funktionalgleichung für die Exponentialfunktion bereits die Additionstheoreme für sin und cos folgen. Ein Version davon lautet:

$$\begin{aligned} \cos x_1 - \cos x_2 &= \left(\cos \frac{x_1+x_2}{2} \cos \frac{x_1-x_2}{2} - \sin \frac{x_1+x_2}{2} \sin \frac{x_1-x_2}{2} \right) \\ &\quad - \left(\cos \frac{x_1+x_2}{2} \cos \frac{x_1-x_2}{2} + \sin \frac{x_1+x_2}{2} \sin \frac{x_1-x_2}{2} \right) \\ &= -2 \sin \frac{x_1+x_2}{2} \sin \frac{x_1-x_2}{2} \end{aligned} \quad (7.6)$$

· Wir wollen nun durch Einschränkung auf reelle Argument die aus der Trigonometrie bekannten Eigenschaften wiederfinden. Zunächst folgt für $x \in \mathbb{R}$ aus $|\exp(ix)|^2 = \exp(ix) \exp(-ix) = 1$ mit Hilfe der Eulerschen Formel, dass

$$\cos^2 x + \sin^2 x = 1 \quad (7.7)$$

Periodizität und Kreiszahl π sind aus der Reihendarstellung hingegen nicht offensichtlich, und wir müssen etwas ausholen.

⁹Wir schreiben ab jetzt auch e^x für $\exp(x)$ für jedes x .

¹⁰Euler, Introductio in analysin infinitorum, Cap. VIII

§ 7. WERTSÄTZE

· Für $x \in (0, 2]$ sind die Glieder der Sinus-Reihe $\frac{x^{2n+1}}{(2n+1)!}$ ab $n = 1$ betragsmässig streng monoton fallend. (Erstmalig nämlich: $\frac{x^3}{6} \leq x \cdot \frac{4}{6} < x$.) Die Leibnitz-Betrachtungen für alternierende Reihen implizieren daher für $0 < x \leq 2$:

$$x - \frac{x^3}{6} < \sin x < x \quad (7.8)$$

In ähnlicher Weise gilt für die gleichen x :

$$1 - \frac{x^2}{2} < \cos x < 1 - \frac{x^2}{2} + \frac{x^4}{24} \quad (7.9)$$

Proposition/Definition 7.4. *Auf dem Intervall $[0, 2]$ ist $\cos$ streng monoton fallend und besitzt genau eine Nullstelle, welche wir mit $\frac{\pi}{2}$ bezeichnen.*

Beweis. Aus (7.8) folgt $\sin x > \frac{1}{3}x > 0$ für $x \in (0, 2]$. Sind nun $x_1, x_2 \in [0, 2]$ mit $x_1 > x_2$ so gilt $\frac{x_1+x_2}{2} \in (0, 2]$ und $\frac{x_1-x_2}{2} \in (0, 2]$. Also folgt mit Hilfe von (7.6) $\cos x_1 - \cos x_2 < 0$. Ausserdem ist $\cos 0 = 1$ und wegen (7.8) gilt $\cos 2 < -\frac{1}{3}$. Mit dem Zwischenertsatz 7.1¹¹ folgt die Existenz genau einer Nullstelle. $\square$

· Alles Übrige ergibt sich durch geschicktes Kombinieren von Additionstheoremen und (7.7). (Übungsaufgaben?) Insbesondere verifizieren wir damit die um (2.11) herum benutzen Eigenschaften und können somit Proposition 2.2 nun voll vertrauen, was wir bald für den Beweis des Fundamentalsatzes der Algebra benutzen werden.

Minimaxsatz

Die andere Zutat zu diesem Beweis wollen wir aber in ihren natürlichen allgemeinen Kontext stellen. Wir nennen einen metrischen Raum X *folgenkompakt*, falls jede Folge $(x_n) \subset X$ eine (in X !) konvergente Teilfolge besitzt. Beispiel: Beschränkte und abgeschlossene Teilmengen des $\mathbb{R}^n$ sind nach 4.17 folgenkompakt. Nicht-leere offene Teilmengen von $\mathbb{R}^n$ sind nicht folgenkompakt.

Theorem 7.5. *Sei X ein folgenkompakter metrischer Raum, und $f : X \rightarrow \mathbb{R}$ stetig. Dann ist*

$$f(X) := \{f(x) \mid x \in X\} \quad (7.10)$$

beschränkt und (falls $X \neq \emptyset$) existiert ein $x_M \in X$ mit $f(x_M) = M := \sup f(X)$, und ein $x_m \in X$ mit $f(x_m) = m := \inf f(X)$.

Beweis. Wir zeigen, dass $f(X)$ nach oben beschränkt ist durch Widerspruch, und konstruieren gleichzeitig eine Maximumsstelle:

(i) Falls $f(X)$ nicht nach oben beschränkt ist, so existiert für jedes n ein $x_n \in X$ mit $f(x_n) > n$.

(ii) Falls $f(X)$ nach oben beschränkt ist, dann existiert für jedes n ein x_n mit $f(x_n) > M - \frac{1}{n}$.

Da X folgenkompakt ist, besitzt $(x_n) \subset X$ eine konvergente Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$ mit Grenzpunkt $x_M \in X$. Wegen der Stetigkeit von f gilt gemäss 6.5

$$\lim_{k \rightarrow \infty} f(x_{n_k}) = f(x_M) \quad (7.11)$$

¹¹Stetigkeit folgt natürlich aus dem Zusammenhang mit der Exponentialfunktion.

Klarerweise ist dies nicht mit (i) verträglich, also ist $f(X)$ nach oben beschränkt. Dann aber impliziert (ii) $f(x_M) \geq M$, d.h. wegen $f(X) \leq M$ gilt $f(x_M) = M$. $\square$

Fundamentalsatz der Algebra

Der folgende Beweis ist elementar und m.E. enorm aufschlussreich.

Theorem 7.6. *Jedes nicht-konstante Polynom mit komplexen Koeffizienten besitzt mindestens eine komplexe Nullstelle.*

Beweis. Wir betrachten das Polynom

$$P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0 \quad n \in \mathbb{N}, a_i \in \mathbb{C} \quad (7.12)$$

und behaupten zunächst, dass die reellwertige und auf ganz $\mathbb{C}$ stetige Funktion $f(z) := |P(z)|$ eine globale Minimumsstelle besitzt.

Bew.: Für $|z| > 1$ und für $i = 0, \dots, n-1$ gilt $|z^i| \leq |z|^{n-1}$. Sei nun $R > 1$ und ausserdem $R > 2 \sum_{i=0}^{n-1} |a_i|$. Dann gilt für $|z| > R$:

$$|a_{n-1}z^{n-1} + \cdots + a_1z + a_0| \leq |z|^{n-1} \cdot \sum_{i=0}^{n-1} |a_i| < |z|^{n-1} \frac{R}{2} < \frac{|z|^n}{2} \quad (7.13)$$

Daraus folgt, weiter für $|z| > R$:

$$f(z) = |P(z)| \geq |z|^n - |a_{n-1}z^{n-1} + \cdots + a_0| > |z|^n - \frac{|z|^n}{2} = \frac{|z|^n}{2} > \frac{R^n}{2} \quad (7.14)$$

Nun ist die abgeschlossene Kreisscheibe $\overline{D_R(0)}$ wie als beschränkte und abgeschlossene Teilmenge von $\mathbb{C}$ folgenkompakt. Daher nimmt gemäss Thm. 7.5 die stetige Funktion $f(z)$ auf $\overline{D_R(0)}$ ein Minimum an. Es existiert also ein $z_m \in \overline{D_R(0)}$ mit $f(z) \geq f(z_m) =: m \quad \forall z \in \overline{D_R(0)}$. Wegen $|f(0)| = |a_0| < \frac{R}{2} < \frac{R^n}{2}$ ist sicher $f(z_m) < \frac{R^n}{2}$, also wegen (7.14) kleiner als jeder Wert ausserhalb der Kreisscheibe.

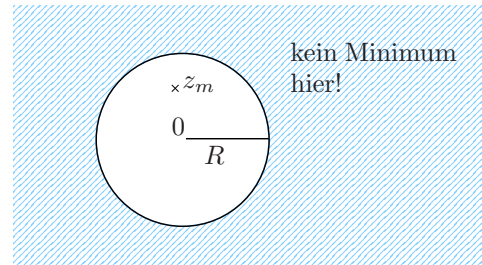
Also ist z_m eine globale Minimumsstelle auf $\mathbb{C}$:

$$f(z) \geq f(z_m) = m \quad \text{für alle } z \in \mathbb{C}. \quad (7.15)$$

· Nun wechseln wir die Perspektive und schauen uns $f(z)$ in der Umgebung von z_m an. Wir schreiben dazu zunächst P als Polynom in der Variablen $w = z - z_m$ um:

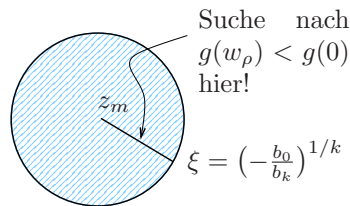
$$Q(w) := P(z_m + w) = b_0 + b_1w + \cdots + b_{n-1}w^{n-1} + b_nw^n \quad (7.16)$$

für gewisse $b_i \in \mathbb{C}$. (Tatsächlich ist $b_n = 1$, also insbesondere $\neq 0$.) Dann ist die Aussage (7.15) äquivalent dazu, dass die stetige Funktion $g(w) := |Q(w)|$ ein globales Minimum bei $w_m = 0$ besitzt. Wir behaupten, dass $m = |b_0| = 0$.



§ 8. GLEICHMÄSSIGKEIT

Bew.: Unter der Annahme, dass $b_0 \neq 0$, sei $k = \min\{i > 0 \mid b_i \neq 0\} \leq n$, m.a.W. ist $Q(w) = b_0 + b_k w^k + \text{höhere Potenzen}$. Gemäss 5.12 (angewandt auf eine endliche Reihe) existiert dann für $r = 1$ eine Konstante C so, dass $|Q(w) - b_0 - b_k w^k| \leq C \cdot |w|^{k+1} \forall |w| \leq 1$. Gemäss 2.2 existiert ein $\xi \in \mathbb{C}^\times$ mit $\xi^k = -b_0/b_k$.



Für $0 \leq \rho < \min\{1, |\xi|^{-1}\}$ sei $w_\rho = \rho\xi$. Dann gilt $|w_\rho| < 1$ und folglich

$$g(w_\rho) = |Q(w_\rho)| \leq |b_0|(1 - \rho^k) + C \cdot \rho^{k+1}|\xi|^{k+1} = m - \rho^k(m - \rho \cdot |\xi|^{k+1}C) \quad (7.17)$$

Ist dann noch $0 < \rho < m/(|\xi|^{k+1}C)$, so folgt $g(w_\rho) < m$, im Widerspruch zu (7.15). $\square$

§ 8 Gleichmässigkeit

Wir hatten uns bei unseren bisherigen Betrachtungen nur selten darum gekümmert, wie “schnell” Folgen, Reihen, etc. gegen ihre Grenzwerte konvergieren. Dies ist insbesondere dann von Nachteil, wenn verschiedene Grenzprozesse miteinander in Konkurrenz treten. Als einfachstes Beispiel diene die “Doppelfolge”

$$\mathbb{N} \times \mathbb{N} \mapsto a_{n,m} := \begin{cases} 1 & \text{für } n \geq m \\ -1 & \text{für } m > n \end{cases} \quad (8.1)$$

Klarerweise ist für jedes m $\lim_{n \rightarrow \infty} a_{n,m} = 1$, also $\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} a_{n,m} = 1$. Auf der anderen Seite ist $\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} a_{n,m} = -1$.

Ein anderes Beispiel ist die Funktionenfolge $f_n : [0, 1] \rightarrow \mathbb{R}$,

$$f_n(x) = x^n \quad (8.2)$$

Es gilt $\lim_{x \rightarrow 1} f_n(x) = 1$ für alle n . Andererseits ist

$$\lim_{x \rightarrow 1} \lim_{n \rightarrow \infty} f_n(x) = 0 \neq 1 = \lim_{n \rightarrow \infty} f_n(1) \quad (8.3)$$

Die punktweise berechnete Grenzfunktion ist also auch gar nicht stetig.

Der Begriff der Gleichmässigkeit von Grenzprozessen (aller Art) ist (nebst der dominierten Konvergenz) das wichtigste Hilfsmittel, das solche Sorgen beim Umgang mit Unendlichkeiten entschärft. Wir präsentieren zwei Beispiele. Das erste verhindert Verhalten der Art (8.3) und dient dem Nachweis der Stetigkeit von Potenzreihen. Das zweite Beispiel wird uns bei der Integration stetiger Funktionen dienlich sein.

Gleichmässige Konvergenz

Wir betrachten der Einfachheit halber komplexwertige Funktionen auf einer Menge X .

Definition 8.1. · Eine Folge $(f_n)_{n \in \mathbb{N}}$ von Funktionen $f_n : X \rightarrow \mathbb{C}$ heisst *punktweise konvergent* gegen $f : X \rightarrow \mathbb{C}$, falls für jedes $x \in X$ die Folge $(f_n(x))_{n \in \mathbb{N}} \subset \mathbb{C}$ gegen $f(x)$ konvergiert.

· Eine solche Folge heisst *gleichmässig konvergent*, falls für jedes $\epsilon > 0$ ein $N_\epsilon \in \mathbb{N}$ existiert so, dass für alle $x \in X$ $|f_n(x) - f(x)| < \epsilon$ falls $n \geq N_\epsilon$.

Bemerkungen. · Beachte den Unterschied: Bei der gleichmässigen Konvergenz müssen wir für gegebenes ϵ zuerst ein N_ϵ finden, ab welchem Index f_n an allen Stellen $x \in X$ nicht weiter als ϵ von f entfernt liegt. Bei der punktweisen Konvergenz muss ein solcher Index erst nach Bekanntgabe von x gefunden werden.

· Gleichmässige Konvergenz impliziert punktweise Konvergenz. Die Umkehrung gilt aber im Allgemeinen nicht. Z.B. gibt es bei (8.2) für jedes n Punkte $x > 0$, für die $|f_n(x) - f(x)| \geq 1/2$.

· Eine bequeme Formulierung der gleichmässigen Konvergenz benutzt die Supremumsnorm aus 4.11: (f_n) konvergiert gleichmässig gegen f genau dann, wenn für fast alle n $f_n - f \in \mathcal{B}(X, \mathbb{C})$ und $\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0$. (Man beachte aber, dass wir nicht $f_n \in \mathcal{B}(X, \mathbb{C})$ voraussetzen!) (Gegebenenfalls verzeichne man auch, dass der Begriff der punktweisen Konvergenz im Allgemeinen nicht von einer Metrik stammt.)

Theorem 8.2. Sei nun X ein metrischer Raum, und $(f_n) \subset \mathcal{C}(X, \mathbb{C})$ eine Folge von stetigen Funktionen, welche gleichmässig gegen $f : X \rightarrow \mathbb{C}$ konvergiert. Dann ist f ebenfalls stetig.

Beweis. Sei $x_0 \in X$ und $\epsilon > 0$. Sei dann N so gross, dass für $n \geq N$

$$|f_n(x) - f(x)| < \epsilon \quad \forall x \in X \quad (8.4)$$

(gleichmässige Konvergenz). Sei nun $\delta > 0$ so klein, dass $|f_N(x_0) - f_N(x)| < \epsilon$ für $d_X(x, x_0) < \delta$ (Stetigkeit von f_N). Dann folgt für solche x

$$|f(x) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| < 3\epsilon \quad (8.5)$$

da jeder der drei Terme kleiner ist als ϵ . □

Bemerkungen. · Wir hatten in 4.11 gesehen, dass der Vektorraum der beschränkten Funktionen $\mathcal{B}(X, \mathbb{R})$ auf einer beliebigen Menge mit der sup-Norm ein vollständiger metrischer Raum ist, und dies gilt auch für $\mathcal{B}(X, \mathbb{C})$. In 7.5 hatten wir gesehen, dass reellwertige stetige Funktionen auf einem folgenkompakten metrischen Raum automatisch beschränkt sind. In Verbindung mit 8.2 können wir also nun verzeichnen, dass $\mathcal{C}(X, \mathbb{C})$ eine abgeschlossene Teilmenge von $\mathcal{B}(X, \mathbb{C})$ ist.

· Eine wichtige Anwendung ist der Nachweis der Stetigkeit von durch Potenzreihen innerhalb ihrer Konvergenzkugel definierten Funktionen.

Beh.: Sei $P(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe, der Einfachheit halber mit $a_n, z \in \mathbb{C}$. Sei R der Konvergenzradius von P . Dann konvergiert für jedes $r < R$ die Folge $(p_n(z) := \sum_{k=0}^n a_k z^k)_{n \in \mathbb{N}}$ auf $\overline{B_r(0)}$ gleichmässig gegen $P(z)$.

Bew.: Sei $\epsilon > 0$. Da $P(r)$ definitionsgemäss absolut konvergiert, existiert ein N so, dass

$$\sum_{k=n+1}^{\infty} |a_k| r^k < \epsilon \quad \forall n \geq N \quad (8.6)$$

Daraus folgt aber schon für alle $z \in \overline{B_r(0)}$ und alle $n \geq N$:

$$|P(z) - p_n(z)| = \left| \sum_{k=n+1}^{\infty} a_k z^k \right| \leq \sum_{k=n+1}^{\infty} |a_k| \cdot |z|^k \leq \sum_{k=n+1}^{\infty} |a_k| r^k < \epsilon \quad (8.7)$$

§ 8. GLEICHMÄSSIGKEIT

Da jedes p_n stetig ist, ist also wegen 8.2 P eine stetige Funktion auf der Scheibe $\overline{B_r(0)}$ für jedes $r < R$. Zusammen folgt, dass $P(z)$ auf ganz $B_R(0)$ eine stetige Funktion definiert. (Im Allgemeinen konvergiert aber p_n nicht gleichmässig gegen P auf ganz $B_R(0)$!) $\square$

Gleichmässige Stetigkeit

Seien nun X, Y zwei metrische Räume.

Definition 8.3. Eine Funktion $F : X \rightarrow Y$ heisst *gleichmässig stetig*, falls für jedes $\epsilon > 0$ ein $\delta > 0$ existiert so, dass

$$d_Y(F(x_1), F(x_2)) < \epsilon \quad \text{falls } d_X(x_1, x_2) < \delta \quad (8.8)$$

Bemerkungen. Was ist der Unterschied von 6.2? Bei der gleichmässigen Stetigkeit müssen wir für gegebenes ϵ *zuerst* ein δ finden, für welches die Abschätzung für jedes nahe Paar $x_1, x_2 \in X$ gilt. Bei der gewöhnlichen Stetigkeit auf ganz X muss das δ erst nach der Bekanntgabe eines der beiden Punkte x_0 gefunden werden.

Beispiel: Die Funktion $\mathbb{C}^\times = \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C}$, $z \mapsto 1/z$ ist (als algebraische Funktion) auf dem gesamten Definitionsbereich stetig, aber nicht gleichmässig stetig: Für $\epsilon_* = 1$ gibt es für jedes $\delta > 0$ Punkte $z_1 \in \mathbb{C}^\times$ mit $|z_1| < \min\{1, \delta\}$. Mit $z_2 = z_1/2$ gilt dann $|z_1 - z_2| = |z_1|/2 < \delta$ aber

$$\left| \frac{1}{z_1} - \frac{1}{z_2} \right| = \frac{1}{|z_1|} > 1 = \epsilon_* \quad (8.9)$$

Ist hingegen zuerst $z_0 \neq 0$ bekannt und $\epsilon > 0$, so wählen wir $0 < \delta < \min\{\frac{|z_0|}{2}, \frac{|z_0|^2}{2}\epsilon\}$. Für $|z - z_0| < \delta$ ist dann $|z| > |z_0|/2$ und daher

$$\left| \frac{1}{z} - \frac{1}{z_0} \right| = \frac{|z - z_0|}{|z| \cdot |z_0|} < \frac{2}{|z_0|^2} \cdot \delta < \epsilon \quad (8.10)$$

Für folgenkompakte Räume kann so etwas aber nicht passieren.

Theorem 8.4. Sei X ein folgenkompakter metrischer Raum, und Y ein beliebiger metrischer Raum. Dann ist jede Funktion $F : X \rightarrow Y$ gleichmässig stetig.

Beweis. Andernfalls gibt es ein $\epsilon_* > 0$ für welches (8.8) für kein $\delta > 0$ erfüllt ist. Insbesondere existiert für jedes $\delta_n = \frac{1}{n}$ ein Paar $x_{1,n}, x_{2,n}$ mit

$$d_X(x_{1,n}, x_{2,n}) < \frac{1}{n} \quad \text{aber} \quad d_Y(F(x_{1,n}), F(x_{2,n})) \geq \epsilon_* \quad (8.11)$$

Da X folgenkompakt ist, besitzt $(x_{1,n})$ eine konvergente Teilfolge (x_{1,n_k}) . Sei $x \in X$ ihr Grenzwert. Da F in x stetig ist, existiert ein $\delta_* > 0$ so, dass $d_Y(F(x), F(x')) < \frac{\epsilon_*}{2}$ falls $d_X(x, x') < \delta_*$. Ist dann K so, dass $d_X(x, x_{1,n_k}) < \frac{\delta_*}{2}$ und $d_X(x_{1,n_k}, x_{2,n_k}) < \frac{\delta_*}{2}$ falls $k \geq K$, dann folgt für solche k

$$\begin{aligned} d_Y(F(x), F(x_{1,n_k})) &< \frac{\epsilon_*}{2} \\ \text{und } d_X(x, x_{2,n_k}) &< \delta_* \text{ also auch } d_Y(F(x), F(x_{2,n_k})) < \frac{\epsilon_*}{2} \end{aligned} \quad (8.12)$$

Daraus folgt aber $d_Y(F(x_{1,n_k}), F(x_{2,n_k})) < \epsilon_*$, im Widerspruch zu (8.11). $\square$

KAPITEL 4

DIFFERENTIATION

Stetige Funktionen haben die charakteristische Eigenschaft, dass unter (“genügend”) kleinen Variationen ihrer Argumente ihre Werte sich nur (“beliebig”) wenig ändern. Bei Lipschitz-stetigen Funktionen nimmt diese Kontrolle die Form einer oberen *linearen* Abschätzung an. Regelmässige Zusammenhänge zwischen physikalischen Grössen sind normalerweise nicht nur stetig, sondern lassen sich vielmehr in einer Weise *linear* approximieren, dass der “relative Fehler” im Grenzfall hinreichend kleiner Änderungen beliebig klein wird. Die Differentialrechnung verbindet auf diese Weise die bisher entwickelten Ideen zu Grenzprozessen mit der aus der linearen Algebra bekannten Theorie der Vektorräume.

Wir diskutieren zunächst den Spezialfall einer reellen Veränderlichen, dessen Rechenregeln bereits vertraut sind. Die Anordnung des Definitionsbereichs (der physikalisch als “Zeit” interpretiert werden könnte) ist später grundlegend für die Lösung (Integration) gewöhnlicher Differentialgleichungen. (Noch etwas spezieller wird es, wenn auch der Wertebereich angeordnet ist, siehe z.B. 9.5.)

Wir besprechen danach die allgemeine Differentialrechnung mehrerer reeller Veränderlicher, zu deren voller Umkehrung wir allerdings erst im WS kommen, um uns für de Rest des Sommers noch etwas mit den Besonderheiten eines (zwar nicht angeordneten, dafür aber) algebraisch abgeschlossenen Definitions- und Wertebereichs (*i.e.*, Teilmengen von $\mathbb{C}$) zu beschäftigen.

§ 9 Eine reelle Veränderliche

Definition 9.1. Sei $I \subset \mathbb{R}$ ein offenes Intervall, und $(V, \|\cdot\|)$ ein vollständiger normierter Vektorraum über $\mathbb{R}$. Eine Funktion $f : I \rightarrow V$ heisst differenzierbar im Punkt $t_0 \in I$, falls der Differenzenquotient

$$I \setminus \{t_0\} \ni t \mapsto \Delta_{f,t_0}(t) := \frac{f(t) - f(t_0)}{t - t_0} \quad (9.1)$$

stetig in t_0 fortsetzbar ist. Man nennt den Grenzwert $\dot{f}(t_0) := \lim_{t \rightarrow t_0} \Delta_{f,t_0}(t)$ die Ableitung von f in t_0 .

- f heisst differenzierbar auf I , falls f in jedem Punkt von I differenzierbar ist.
- f heisst stetig differenzierbar auf I , falls f differenzierbar auf I ist, und die Abbildung

$$\dot{f} : I \rightarrow V \quad t \mapsto \dot{f}(t) \quad (9.2)$$

stetig ist.

Bemerkungen/Beispiele 9.2. · Andere Schreibweisen für $\dot{f}(t_0)$ sind $\frac{df}{dt}(t_0)$ und $f'(t_0)$.

· Eine konstante Funktion hat die Ableitung 0, für eine affin-lineare Funktion der Form $f(t) = x_0 + vt$ ($x_0, v \in V$) gilt $\dot{f}(t) = v$.

· Mit punktweiser Addition und Skalarmultiplikation wird der Raum der differenzierbaren Funktionen $\mathcal{D}(I, V)$ (bzw. der stetig differenzierbaren Funktionen, $\mathcal{C}^1(I, V)$) zu einem reellen Vektorraum (über eine Norm unterhalten wir uns später...)

· Ist auf V eine Multiplikation erklärt, so können wir Funktionen punktweise multiplizieren und es gilt die Produktregel

$$\begin{aligned} \frac{d(f \cdot g)}{dt}(t_0) &= \lim_{t \rightarrow t_0} \frac{f(t)g(t) - f(t_0)g(t) + f(t_0)g(t) - f(t_0)g(t_0)}{t - t_0} \\ &= \frac{df}{dt}(t_0) \cdot g(t_0) + f(t_0) \cdot \frac{dg}{dt}(t_0) \end{aligned} \quad (9.3)$$

(Daraus folgen dann die bekannten Rechenregeln für Polynome in einer reellen Variablen.)

· Ist $V \cong \mathbb{R}^m$ für ein $m \in \mathbb{N}$ (m.a.W., ist V endlich-dimensional und hat eine ausgezeichnete Basis), so können wir $f = (f^1, \dots, f^m)^T$ als m -Tupel von reellwertigen Funktionen $f^j : I \rightarrow \mathbb{R}$, $j = 1, \dots, m$ schreiben. Es gilt dann: f ist genau dann differenzierbar wenn alle f^j differenzierbar sind, und $(f(t))^j = \dot{f}^j(t)$.

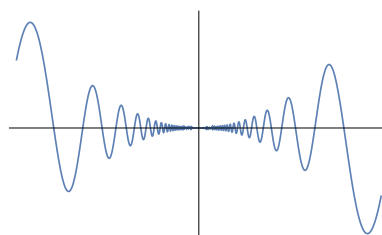
· Die Ableitung der Exponentialfunktion $t \mapsto \exp(t)$ bei $t_0 = 0$ ist

$$\frac{d}{dt} \exp(t)|_{t=0} = \lim_{t \rightarrow 0} \frac{\exp(t) - 1}{t} = 1 \quad (9.4)$$

(vgl. Übungen), woraus mit Hilfe der Funktionalgleichung folgt, dass $\frac{d}{dt} \exp(tx) = \exp(tx) \cdot x$. (Und daraus folgen dann die bekannten Ableitungen der trigonometrischen Funktionen.)

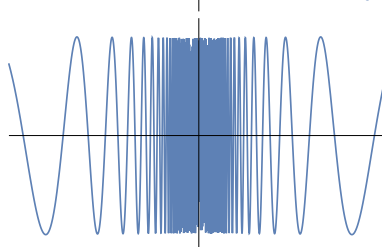
· Beispiel einer differenzierbaren, aber nicht stetig differenzierbaren Funktion:

$$f(t) = \begin{cases} t^2 \sin \frac{1}{t} & t \neq 0 \\ 0 & t = 0 \end{cases} \quad (9.5)$$



Es ist (unter Vorwegnahme der Kettenregel...)

$$\dot{f}(t) = \begin{cases} 2t \sin \frac{1}{t} - \cos \frac{1}{t} & \text{für } t \neq 0 \\ \lim_{t \rightarrow 0} t^2 \sin \frac{1}{t} = 0 & \text{für } t = 0 \end{cases} \quad (9.6)$$



die Ableitung aber nicht stetig in $t_0 = 0$.

· Je nach Sichtweise auf Grenzwerte von Funktionen gibt es verschiedene äquivalente Formulierungen der Differenzierbarkeit. Den Einstieg in die Verallgemeinerung auf mehrere Variablen ermöglicht die folgende: f ist genau dann in t_0 differenzierbar, falls ein $\dot{f}(t_0) \in V$ existiert so, dass $\forall \epsilon > 0$ ein $\delta > 0$ existiert so, dass

$$\|f(t) - f(t_0) - \dot{f}(t_0)(t - t_0)\| < \epsilon |t - t_0| \quad \forall t \in I \cap (t_0 - \delta, t_0 + \delta) \setminus \{t_0\} \quad (9.7)$$

§ 9. EINE REELLE VERÄNDERLICHE

“Die Differenz zwischen Funktion und linearer Näherung geht (in der Norm) schneller als linear in $t - t_0$ gegen Null.”

· Es ist auch sinnvoll, den Begriff auf andere Teilmengen von $\mathbb{R}$ als Definitionsbereiche zu verallgemeinern. Ist I ein abgeschlossenes oder halb-offenes Intervall, so kann man an einem Randpunkt weiter (9.7) benutzen bzw. den am Ende von § 6 eingeführten Begriff des links-/rechtsseitigen Grenzwertes. Bevor man für einen allgemeinen Definitionsbereich D auf Differenzierbarkeit in $x_0 \in D$ untersucht, verlangt man, dass für hinreichend kleines δ der Durchschnitt $D \cap (x_0 - \delta, x_0 + \delta)$ ein Intervall der Form $(x_0 - \delta, x_0 + \delta)$, $[x_0, x_0 + \delta)$ oder $(x_0 - \delta, x_0]$ ist.

· Für ein kompaktes Intervall $[a, b]$ heisst eine differenzierbare Abbildung $\gamma : [a, b] \rightarrow V$ auch ein *Weg* von $x_1 = \gamma(a)$ nach $x_2 = \gamma(b)$.

· Aus (9.7) folgt sofort, dass eine differenzierbare Funktion insbesondere stetig ist:

$$\|f(t) - f(t_0)\| \leq (\epsilon + \|\dot{f}(t_0)\|) \cdot |t - t_0| \quad (9.8)$$

für $|t - t_0| < \delta$. Allgemeiner gilt:

Proposition 9.3 (Schranksatz). *Sei $f : [t_1, t_2] \rightarrow V$ eine stetige Funktion, die auf (t_1, t_2) differenzierbar ist derart, dass $\|\dot{f}\|_{(t_1, t_2)} := \sup\{\|\dot{f}(t)\| \mid t \in (t_1, t_2)\} < \infty$. (Beispiel: f ist auf $[t_1, t_2]$ stetig differenzierbar.) Dann gilt*

$$\|f(t_2) - f(t_1)\| \leq \|\dot{f}\|_{(t_1, t_2)} \cdot |t_2 - t_1| \quad (9.9)$$

Beweis. Angenommen, die linke Seite von (9.9) wäre grösser als die rechte. Dann existierte ein $\epsilon > 0$ so, dass noch $\|f(t_2) - f(t_1)\| - (\|\dot{f}\|_{(t_1, t_2)} + \epsilon)(t_2 - t_1) > 0$. Mit solch einem ϵ betrachten wir für $t \in [t_1, t_2]$ die reellwertige Funktion

$$g_\epsilon(t) := \|f(t) - f(t_1)\| - (\|\dot{f}\|_{(t_1, t_2)} + \epsilon)(t - t_1) \quad (9.10)$$

Wegen $g_\epsilon(t_1) = 0$ ist $t_0 := \sup\{t \in [t_1, t_2] \mid g_\epsilon(t) = 0\}$ wohldefiniert. Aus der Stetigkeit von g_ϵ folgt $g_\epsilon(t_0) = 0$ und wegen $g_\epsilon(t_2) > 0$ folgt $t_0 < t_2$, und $g_\epsilon(t) > 0$ für $t \in (t_0, t_2) \neq \emptyset$. (Wäre $g_\epsilon(t') < 0$ für ein $t' \in (t_0, t_2)$ so gäbe es wegen des Zwischenwertsatzes 7.1 ein $t'' \in (t', t_2)$ mit $g_\epsilon(t'') = 0$, im Widerspruch zur Definition von t_0 .) Für alle solche t gilt:

$$\begin{aligned} 0 < g_\epsilon(t) - g_\epsilon(t_0) &= \|f(t) - f(t_1)\| - \|f(t_0) - f(t_1)\| - (\|\dot{f}\|_{(t_1, t_2)} + \epsilon)(t - t_0) \\ &\leq \|f(t) - f(t_0)\| - (\|\dot{f}\|_{(t_1, t_2)} + \epsilon)(t - t_0) \end{aligned} \quad (9.11)$$

Da aber f in t_0 differenzierbar ist, existiert ein $\delta > 0$ so, dass für solche t , die ausserdem kleiner als $t_0 + \delta$ sind, $\|f(t) - f(t_0)\| < (\|\dot{f}(t_0)\| + \epsilon)|t - t_0|$. (Siehe (9.8).) Wegen $\|\dot{f}(t_0)\| \leq \|\dot{f}\|_{(t_1, t_2)}$ steht dies im Widerspruch zu (9.11). $\square$

Korollar 9.4. *Sei $f : D \rightarrow V$ differenzierbar. Gilt $\dot{f}(t) = 0 \ \forall t \in D$, so ist f auf jedem Intervall $I \subset D$ konstant.*

Wertebereich $\mathbb{R}$

Wie bei stetigen Funktionen auch erlauben reellwertige differenzierbare Funktionen einige zusätzliche Betrachtungen. Beispielsweise lässt sich der Begriff der Differenzierbarkeit direkt auf die Anordnung und Vollständigkeit von $\mathbb{R}$ zurückführen. Übungsaufgabe:

Beispiel 9.5 (Ableiten ohne Grenzen). Für eine reellwertige Funktion einer reellen Veränderlichen soll der Begriff der Ableitung direkt auf die Anordnungs- und Vollständigkeitsstruktur der reellen Zahlen zurückgeführt werden. Sei $x_0 \in I$, einem offenen Intervall, und $f : I \rightarrow \mathbb{R}$. Eine reelle Zahl o heisst *Obersteigung* von f in x_0 , falls ein offenes Intervall $I_o \subset I$ existiert mit $x_0 \in I_o$, sodass für alle $x \in I_o$ gilt, dass

$$\begin{aligned} f(x_0) + o \cdot (x - x_0) &> f(x) \quad \text{falls } x > x_0 \\ \text{und} \quad f(x_0) + o \cdot (x - x_0) &< f(x) \quad \text{falls } x < x_0 \end{aligned} \quad (9.12)$$

Analog definiert man den Begriff einer *Untersteigung*. Sei O die Menge der Obersteigungen von f in x_0 , und U die Menge der Untersteigungen. Überlegen Sie sich zunächst (durch kurze Beweise bzw. geeignete Gegenbeispiele), dass

- (i) O nach oben ordnungsabgeschlossen ist, d.h. $o \in O \Rightarrow o' \in O \forall o' > o$, aber u.U. leer sein kann, und
- (ii) $U < O$, d.h. $u < o$ für alle $u \in U, o \in O$ (also insbesondere ist O nach unten beschränkt, falls $U \neq \emptyset$).
- (iii) Zeigen Sie dann: f ist genau dann in x_0 differenzierbar, wenn O und U nicht-leer sind, und $\inf O = \sup U =: s$. (In diesem Fall gilt $s = f'(x_0)$.)
- (iv) Geben Sie ein Beispiel einer Funktion mit Stelle, für welche U und O nicht leer sind, aber $\inf O \neq \sup U$.

Diese Formulierung ist der eigentliche Ursprung des Zusammenhangs zwischen Differentialrechnung und Monotonie-/Extremwertaufgaben. Ganz allgemein sagt man, eine Funktion $F : X \rightarrow \mathbb{R}$ auf einem metrischen Raum X habe in einem Punkt $x_0 \in X$ ein lokales Maximum (besser wäre: lokale Maximumsstelle), falls ein $\delta > 0$ existiert so, dass $F(x) \leq F(x_0)$ für alle $x \in B_\delta(x_0)$. Ein lokales Minimum ist analog definiert.

Lemma 9.6 (Notwendiges Extremstellenkriterium). *Sei I ein offenes Intervall und $f : I \rightarrow \mathbb{R}$ differenzierbar. Hat f in $x_0 \in I$ ein lokales Maximum oder Minimum, dann gilt $f'(x_0) = 0$.*

Beweis. Im Falle eines Maximums existiert ein $\delta > 0$ so, dass $f(x) \leq f(x_0)$ für $x \in I \cap (x_0 - \delta, x_0 + \delta)$. Für die $x > x_0$ ist der Differenzenquotient

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0 \quad (9.13)$$

und im Grenzfall $x \downarrow x_0$ folgt $f'(x_0) \leq 0$. Durch Betrachtung der $x_0 - \delta < x < x_0$ zeigt man $f'(x_0) \geq 0$, zusammen folgt $f'(x_0) = 0$. Der Fall eines Minimums geht analog. $\square$

§ 9. EINE REELLE VERÄNDERLICHE

Proposition 9.7 (Mittelwertsatz). Sei $[a, b] \subset \mathbb{R}$ und $f : [a, b] \rightarrow \mathbb{R}$ stetig, und auf (a, b) differenzierbar. Dann existiert ein $\xi \in (a, b)$ für das $f(b) - f(a) = (b - a)f'(\xi)$.

Beweis. Wir betrachten die Funktion $g : [a, b] \rightarrow \mathbb{R}$

$$g(t) = f(t) - f(a) - \frac{f(b) - f(a)}{b - a}(t - a) \quad (9.14)$$

g ist auf (a, b) differenzierbar und auf $[a, b]$ stetig mit $g(a) = g(b) = 0$. Nach dem Zwischenwertsatz 7.1 nimmt g ein Maximum M und ein Minimum m an. Gilt $M = m$, so verschwindet g identisch auf $[a, b]$, f ist affin-linear und die Aussage unmittelbar. Andernfalls liegt wegen $g(a) = g(b)$ entweder die Maximumstelle oder die Minimumstelle in (a, b) , und wir schliessen mit dem Kriterium 9.6. $\square$

Definition 9.8. Eine reellwertige Funktion $I \rightarrow \mathbb{R}$ auf einem Intervall heisst konvex, falls für alle $x_1, x_2 \in I$, und alle $t \in (0, 1)$ gilt

$$f((1 - t)x_1 + tx_2) \leq (1 - t)f(x_1) + tf(x_2) \quad (9.15)$$

Gilt die strenge Ungleichung, so heisst f streng konvex. Gilt die umgekehrte (strenge) Ungleichung, so heisst f (streng) konkav.

Bemerkungen. Schreibt man $x = (1 - t)x_1 + tx_2$, also $t = \frac{x - x_1}{x_2 - x_1}$ so wird aus (9.15)

$$\begin{aligned} f(x) &\leq \frac{x_2 - x}{x_2 - x_1}f(x_1) + \frac{x - x_1}{x_2 - x_1}f(x_2) \\ \Leftrightarrow f(x) &\leq f(x_1) + \frac{x - x_1}{x_2 - x_1}(f(x_2) - f(x_1)) \quad \forall x \in [x_1, x_2] \end{aligned} \quad (9.16)$$

“Zwischen je zwei Stellen x_1 und x_2 liegt die Funktion f unterhalb ihrer linearen Interpolation.”

Konvexe Funktionen spielen eine sehr wichtige Rolle bei der Untersuchung der Stabilität physikalischer Systeme. Bei differenzierbaren Funktionen gebraucht man das folgende Ergebnis:

Proposition 9.9. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) zweimal differenzierbar (d.h. f ist differenzierbar und die Ableitung f' nochmal). Dann ist f genau dann konvex, wenn $f''(x) \geq 0 \quad \forall x \in (a, b)$.

Beweisskizze. Man schreibt (9.16) um als

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x} \quad \forall x_1 < x < x_2 \quad (9.17)$$

und stellt daraus mit Hilfe des Mittelwertsatzes den Zusammenhang zum monotonen Wachstum von f' her. Dann benutzt man den Zusammenhang zwischen Monotonie einer Funktion und dem Vorzeichen ihrer Ableitung. $\square$

Wir überlassen alle weiteren Feinheiten solcher Kurvendiskussionen den Übungen.¹²

¹²Beispielsweise gilt die Verschärfung von 9.9 nur in der Richtung: $f'' > 0 \Rightarrow f$ streng konvex. f kann streng konvex sein, und f'' trotzdem Nullstellen besitzen.

§ 10 Mehrere Veränderliche

Hängt eine bestimmte (physikalische) Grösse von mehreren, voneinander unabhängigen, Variablen ab, so macht es zunächst Sinn, die Abhängigkeit von *jeder dieser Variablen getrennt* mittels der Differentialrechnung zu untersuchen. Dafür reichen im Prinzip die Methoden aus dem letzten §. Stehen aber die unabhängigen Variablen in einer engeren Beziehung zueinander (wie etwa die Koordinaten in drei räumliche Richtungen, die unter Drehungen ineinander übergehen), so ist es richtig und besser, die Abhängigkeit von *all diesen Variablen gemeinsam* zu beschreiben. Wir entwickeln in diesem § die Grundbegriffe dieser mehr-dimensionalen Differentialrechnung (Vektoranalysis).

Definition 10.1. Seien $n, m \in \mathbb{N}$ (Dimension 0 ist weniger interessant), und $U \subset \mathbb{R}^n$ offen. Eine Funktion $F : U \rightarrow \mathbb{R}^m$ heisst differenzierbar in $x_0 \in U$, falls eine lineare Abbildung $DF(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ (das “Differential von F in x_0 ”) existiert so, dass

$$\lim_{x \rightarrow x_0} \frac{\|F(x) - F(x_0) - DF(x_0)(x - x_0)\|}{\|x - x_0\|} = 0 \quad (10.1)$$

- F heisst differenzierbar auf U , falls F in jedem $x_0 \in U$ differenzierbar ist
- F heisst stetig differenzierbar auf U , falls F differenzierbar ist und die Abbildung $U \rightarrow \text{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m)$, $x \mapsto DF(x)$ stetig ist.

Bemerkungen. • In (10.1) steht im Nenner eine Norm auf $\mathbb{R}^n$, im Zähler eine auf $\mathbb{R}^m$. Welche Normen zu Grunde liegen spielt wie bereits mehrfach betont keine Rolle. (Weder für den Begriff der Differenzierbarkeit noch für den Wert des Differentials, s. unten.) Für den Begriff der stetigen Differenzierbarkeit benutzt man dann am besten die zugehörige Operatornorm auf $\text{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m) \cong \text{Mat}_{m \times n}(\mathbb{R})$.

• Die ein-dimensionale Definition 9.1 benutzt explizit die Körperstruktur von $\mathbb{R}$, um in (9.1) direkt durch den Zuwachs $t - t_0$ zu teilen. Im Unterschied dazu bewertet man in (10.1) die Güte der linearen Approximation über den Vergleich in der Norm. Dennoch reduziert sich natürlich die Definition 10.1 im Fall $n = 1$ wieder auf 9.1. Die Aussage (10.1) ist nämlich äquivalent zu: $\forall \epsilon > 0$ existiert ein $\delta > 0$ so, dass

$$\|F(x) - F(x_0) - DF(x_0)(x - x_0)\| \leq \epsilon \cdot \|x - x_0\| \quad \text{falls } \|x - x_0\| < \delta \quad (10.2)$$

Diese Formulierung reduziert sich direkt auf (9.7), bis auf die Benutzung von $\leq$ statt $<$ vor dem ϵ , welche den Ausschluss von $x = x_0$ vermeidet. In diesem Fall ist die Ableitung $\dot{f}$ (ein Vektor in $V \cong \mathbb{R}^m$) die Auswertung des Differentials (einer linearen Abbildung) auf dem Standardbasisvektor von $\mathbb{R}$.

- Dies gibt eine Möglichkeit zu sehen, dass $DF(x_0)$ durch die Gleichung (10.1) eindeutig bestimmt ist: Nach Ergebnissen der linearen Algebra genügt es ja, die Auswertung von $DF(x_0)$ auf der Standard-Basis $(e_1, \dots, e_n)$ von $\mathbb{R}^n$ zu kennen. Diese erhält man durch Einschränkung des Limes auf die Teilmenge $\{x \in U \mid x - x_0 \in \mathbb{R}e^i\}$, wo sich (10.1) auf eine ein-dimensionale Ableitung reduziert. (Mehr dazu später....)
- Ein etwas heiklerer Punkt ist die Verallgemeinerung des mehr-dimensionalen Begriffs auf nicht-offene Definitionsbereiche (was für die Integralrechnung durchaus brauchbar ist). Im ein-dimensionalen Fall kommt man mit einseitigen Grenzwerten

§ 10. MEHRERE VERÄNDERLICHE

ganz gut auf den Rand von Intervallen. In höheren Dimensionen ist es vernünftig, von den Rändern eine gewisse Glattheit zu verlangen. Für die meisten Zwecke reicht es aus, differenzierbare Funktionen von offenen Definitionsbereichen aus auf die gewünschten Mengen einzuschränken.

Proposition 10.2 (Kettenregel). *Sei $U \subset \mathbb{R}^n$ offen und $F : U \rightarrow \mathbb{R}^m$ in $x_0 \in U$ differenzierbar. Sei sodann $V \subset \mathbb{R}^m$ offen mit $F(x_0) \in V$ und $G : V \rightarrow \mathbb{R}^l$ in $F(x_0)$ differenzierbar. Dann ist $G \circ F$ in x_0 differenzierbar mit*

$$D(G \circ F)(x_0) = DG(F(x_0)) \circ DF(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^l \quad (10.3)$$

Beweis. Die Behauptung macht nur Sinn, wenn $G \circ F$ in einer offenen Umgebung von x_0 definiert ist. Dies folgt aus der Stetigkeit von F in x_0 . ($F^{-1}(V)$ ist offen.)

• Für $\zeta > 0$ gibt es wegen der Differenzierbarkeit von G in $y_0 := F(x_0)$ ein $\epsilon > 0$ so, dass $B_\epsilon(y_0) \subset V$ und

$$\|G(y) - G(y_0) - DG(y_0)(y - y_0)\|_{\mathbb{R}^l} \leq \zeta \cdot \|y - y_0\|_{\mathbb{R}^m} \quad \text{für } y \in B_\epsilon(y_0). \quad (10.4)$$

Wegen der Offenheit von U , der Stetigkeit und der Differenzierbarkeit von F in x_0 gibt es dann ein $\delta > 0$ so, dass gleichzeitig $B_\delta(x_0) \subset U$, $F(B_\delta(x_0)) \subset B_\epsilon(y_0)$ und

$$\|F(x) - F(x_0) - DF(x_0)(x - x_0)\|_{\mathbb{R}^m} \leq \zeta \cdot \|x - x_0\|_{\mathbb{R}^n} \quad \text{für } x \in B_\delta(x_0). \quad (10.5)$$

Dann ist $G \circ F$ für solche x definiert. Wir schätzen ab

$$\begin{aligned} \|G \circ F(x) - G \circ F(x_0) - DG(y_0) \circ DF(x_0)(x - x_0)\|_{\mathbb{R}^l} \\ \leq \|G(F(x)) - G(y_0) - DG(y_0)(F(x) - y_0)\|_{\mathbb{R}^l} \\ + \|DG(y_0)(F(x) - F(x_0) - DF(x_0)(x - x_0))\|_{\mathbb{R}^l} \end{aligned} \quad (10.6)$$

Der erste Term auf der rechten Seite ist wegen $F(x) \in B_\epsilon(y_0)$, (10.4) und (10.5)

$$\leq \zeta \cdot \|F(x) - F(x_0)\|_{\mathbb{R}^m} \leq \zeta \cdot (\|DF(x_0)\| + \zeta) \cdot \|x - x_0\|_{\mathbb{R}^n} \quad (10.7)$$

während der zweite Term wegen (10.5) durch

$$\|DG(y_0)\| \cdot \zeta \cdot \|x - x_0\|_{\mathbb{R}^n} \quad (10.8)$$

abgeschätzt werden kann. (Hier haben wir zweimal die Abschätzung (4.28) (i) von Vektor- durch Operatornorm benutzt.) Zusammen folgt

$$\begin{aligned} \|G \circ F(x) - G \circ F(x_0) - DG(y_0) \circ DF(x_0)(x - x_0)\|_{\mathbb{R}^l} \\ \leq \zeta \cdot (\|DF(x_0)\| + \zeta + \|DG(y_0)\|) \cdot \|x - x_0\|_{\mathbb{R}^n} \end{aligned} \quad (10.9)$$

und damit die Behauptung. $\square$

Wir stellen nun den Zusammenhang zur eingangs erwähnten Alternative her, die unabhängigen Variablen “getrennt” zu variieren. Die Idee hatten wir schon bei der Begründung der Eindeutigkeit von $DF(x_0)$ unterhalb von 10.1 verwendet. Allgemeiner definiert man

Definition 10.3. Für $U \subset \mathbb{R}^n$ und $v \in \mathbb{R}^n$ heisst eine Funktion $F : U \rightarrow \mathbb{R}^m$ “differenzierbar in Richtung v ” in $x_0 \in U$, falls $\{x_0 + tv \mid |t| < \delta\} \subset U$ für ein $\delta > 0$ und die Funktion $t \mapsto F(x_0 + tv)$ in $t_0 = 0$ differenzierbar ist. Bei Gebrauch schreibe wir $D_v F(x_0)$ für diese Richtungsableitung.

Bemerkungen/Beispiele 10.4. · Falls F in x_0 differenzierbar ist, so folgt aus der Linearität von $DF(x_0)$, dass dann F in x_0 auch in alle Richtungen v differenzierbar ist, mit Richtungsableitung $D_v F(x_0) = DF(x_0)(v)$. Alternativ folgt dies aus der Kettenregel: $t \mapsto F(x_0 + tv)$ ist die Verkettung von $t \mapsto x_0 + tv$ mit F und

$$\frac{d}{dt} F(x_0 + tv) = DF(x_0) \circ v = DF(x_0)(v) \quad (10.10)$$

· Für $j = 1, \dots, m$ sei F^j die j -te Komponente von F , das heisst für $x \in U$ ist $F(x) = (F^1(x), \dots, F^m(x))^T$. Wie im ein-dimensionalen Fall gilt dann: F ist genau dann differenzierbar, wenn all Komponenten $F^j : U \rightarrow \mathbb{R}$ differenzierbar sind. In der einen Richtung folgt dies z.B. daraus, dass man die j -te Komponente auffassen kann als Bild unter der linearen Abbildung $\lambda^j : \mathbb{R}^m \rightarrow \mathbb{R}$, $w \mapsto w^j$. Es gilt nämlich $D\lambda^j = \lambda^j$, und daher $DF^j = D(\lambda^j \circ F) = D\lambda^j \circ DF = (DF)^j$

· Die Darstellung der linearen Abbildung $DF(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ bezüglich den Standard-Basen von $\mathbb{R}^n$ und $\mathbb{R}^m$ ist dann die Matrix der Richtungsableitungen der Komponenten von F (Jacobi-Matrix oder Funktionalmatrix)

$$DF(x_0)_i^j = \lambda^j(DF(x_0))(e_i) = \frac{\partial F^j}{\partial x^i}(x_0) = \partial_i F^j(x_0) \quad (10.11)$$

· Eine Abbildung heisst partiell differenzierbar, wenn sie in alle Koordinatenrichtungen differenzierbar ist, und stetig partiell differenzierbar, wenn diese Richtungsableitungen stetige Funktionen sind.

· Im Allgemeinen ist aber die Existenz selbst aller Richtungsableitungen im Sinne von 10.3 nicht hinreichend für die (“totale”) Differenzierbarkeit im Sinne von 10.1. Ist beispielsweise $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ homogen vom Grad 1, d.h. $F(tx, ty) = tF(x, y)$, dann ist die Ableitung in Richtung (ξ, η) genau $F(\xi, \eta)$, aber nicht unbedingt linear.

Proposition 10.5. Sei $U \subset \mathbb{R}^n$ offen, und $F : U \rightarrow \mathbb{R}^m$ in einer Umgebung von $x_0 \in U$ stetig partiell differenzierbar. Dann ist F in x_0 differenzierbar.

Beweis. Wegen 10.4 genügt es, den Fall $m = 1$ zu betrachten. Es sei $\delta > 0$ so, dass $\forall x \in B_\delta(x_0)$ der abgeschlossene Quader $\{((1 - t_i)x_0^i + t_i x^i)_{i=1, \dots, n} \mid t_i \in [0, 1]\} \subset U$, und alle partiellen Ableitungen $\partial_i F$ dort existieren und stetig sind. Die Funktion $[0, 1] \ni t \mapsto F((1 - t)x_0^1 + tx^1, x_0^2, \dots, x_0^n)$ ist dann wohldefiniert und stetig differenzierbar mit Ableitung $\partial_1 F((1 - t)x_0^1 + tx^1, x_0^2, \dots, x_0^n)(x^1 - x_0^1)$. Daher existiert nach dem Mittelwertsatz 9.7 eine Stelle ξ^1 zwischen x_0^1 und x^1 so, dass

$$F(x^1, x_0^2, \dots, x_0^n) - F(x_0^1, x_0^2, \dots, x_0^n) = \partial_1 F(\xi^1, x_0^2, \dots, x_0^n)(x^1 - x_0^1) \quad (10.12)$$

Durch Wiederholung finden wir auch für $i = 2, \dots, n$ Stellen ξ^i zwischen x_0^i und x^i so, dass

$$F(x) - F(x_0) = \sum_{i=1}^n \partial_i F(x^1, \dots, x^{i-1}, \xi^i, x_0^{i+1}, \dots, x_0^n)(x^i - x_0^i) \quad (10.13)$$

§ 10. MEHRERE VERÄNDERLICHE

Mit der Abkürzung $x_i := (x^1, \dots, x^{i-1}, \xi^i, x_0^{i+1}, \dots, x_0^n)$ gilt dann

$$\begin{aligned} |F(x) - F(x_0) - \sum_{i=1}^n \partial_i F(x_0)(x^i - x_0^i)| &= \left| \sum_{i=1}^n (\partial_i F(x_i) - \partial_i F(x_0))(x^i - x_0^i) \right| \\ &\leq \sum_{i=1}^n |\partial_i F(x_i) - \partial_i F(x_0)| \cdot \|x - x_0\|_\infty \end{aligned} \quad (10.14)$$

Aus der Stetigkeit der $\partial_i F$, der Äquivalenz der Normen, und $\|x_i - x_0\|_\infty \leq \|x - x_0\|_\infty$ folgt, dass die rechte Seite für jedes $\epsilon > 0$ kleiner als $\epsilon \|x - x_0\|$ wird für $\|x - x_0\| < \delta$ klein genug, und damit die Behauptung. $\square$

Lemma 10.6 (Schranksatz). *Sei $U \subset \mathbb{R}^n$ offen und $F : U \rightarrow \mathbb{R}^m$ stetig differenzierbar. Für je zwei Punkte $x_1, x_2 \in U$, für welche die Verbindungsstrecke $\{(1-t)x_1 + tx_2 \mid t \in [0, 1]\}$ ganz in U liegt, gilt*

$$\|F(x_2) - F(x_1)\| \leq \sup\{\|DF((1-t)x_1 + tx_2)\| \mid t \in [0, 1]\} \cdot \|x_2 - x_1\| \quad (10.15)$$

Insbesondere gilt: Ist $K \subset U$ konvex und folgenkompakt, so ist

$$\|F(x_2) - F(x_1)\| \leq \|DF\|_K \cdot \|x_2 - x_1\| \quad (10.16)$$

Beweis. Die Funktion $f : [0, 1] \rightarrow \mathbb{R}^m$, $f(t) = F((1-t)x_1 + tx_2)$ ist wegen der Kettenregel 10.2 differenzierbar mit $\dot{f}(t) = DF((1-t)x_1 + tx_2)(x_2 - x_1)$ und erfüllt auch die sonstigen Voraussetzungen des Schranksatzes 9.3. Mit (4.28) (i), d.h. $\|DF(x)(x_2 - x_1)\| \leq \|DF(x)\| \cdot \|x_2 - x_1\|$ folgt die Behauptung. $\square$

Wertebereich $\mathbb{R}$

Die sich aus 10.4 und 10.5 ergebende Rückführung auf die Ableitung nach einer einzelnen Veränderlichen erlaubt eine bequeme Diskussion reellwertiger Funktionen mehrerer Veränderlicher. Beispielsweise gilt: Ist $F : U \rightarrow \mathbb{R}$ differenzierbar und x_0 ein lokales Maximum oder Minimum, dann ist $DF(x_0) = 0$.

Bew.: Es genügt zu zeigen, dass $DF(x_0)(v) = 0 \forall v \in \mathbb{R}^n$. Für $\delta > 0$ klein genug ist $\{x_0 + tv \mid t \in (-\delta, \delta)\} \subset U$ und die Einschränkung $t \mapsto F(x_0 + tv)$ hat ein lokales Maximum (bzw. Minimum) in $t_0 = 0$. Dann folgt die Aussage aus 9.6. $\square$

· Zwar macht wegen der fehlenden Anordnung im Definitionsbereich der Begriff der monotonen Funktion für $n > 1$ reelle Variablen keinen Sinn.

· Demgegenüber ist aber die Diskussion der Konvexität sinnvoll und weiterhin sehr wichtig. Wir zeigen zunächst

Lemma 10.7 (Vertauschbarkeit partieller Ableitungen). *Sei $U \subset \mathbb{R}^n$ offen und $F : U \rightarrow \mathbb{R}$. Angenommen, für ein Paar $i, j \in \{1, \dots, n\}$ existieren $\partial_i F$, $\partial_j F$ und $\partial_{ji} F := \partial_j \partial_i F$ auf U . Ausserdem sei $\partial_{ji} F$ in x_0 stetig. Dann existiert auch $\partial_{ij} F(x_0) = \partial_i \partial_j F(x_0)$ und ist gleich $\partial_{ji} F(x_0)$.*

Beweis. Zur Vereinfachung der Notation sei $n = 2$, $i = 1$, $j = 2$. Wir betrachten für (x^1, x^2) in einem Quadrat um (x_0^1, x_0^2) den “Plaquette-Quotienten”

$$\Xi(x^1, x^2) := \frac{F(x^1, x^2) - F(x_0^1, x^2) - F(x^1, x_0^2) + F(x_0^1, x_0^2)}{(x^1 - x_0^1)(x^2 - x_0^2)} \quad (10.17)$$

Wegen partieller Differenzierbarkeit in die $(1, 0)$ -Richtung gilt

$$\lim_{x^1 \rightarrow x_0^1} \Xi(x^1, x^2) = \frac{\partial_1 F(x_0^1, x^2) - \partial_1 F(x_0^1, x_0^2)}{x^2 - x_0^2} \quad (10.18)$$

und nach Voraussetzung existiert

$$\lim_{x^2 \rightarrow x_0^2} \left(\lim_{x^1 \rightarrow x_0^1} \Xi(x^1, x^2) \right) = \partial_{21} F(x_0^1, x_0^2) \quad (10.19)$$

Um zu zeigen, dass die umgekehrte Grenzwertreihenfolge das gleiche Ergebnis liefert, gehen wir zunächst mit dem Mittelwertsatz 9.7 in (10.17) zurück, um eine Stelle ξ^1 zwischen x_0^1 und x^1 zu finden, mit der

$$\Xi(x^1, x^2) = \frac{\partial_1 F(\xi^1, x^2) - \partial_1 F(\xi^1, x_0^2)}{x^2 - x_0^2} \quad (10.20)$$

Da nun $\partial_1 F$ auf U partiell in die $(0, 1)$ -Richtung differenzierbar ist, existiert, wieder nach dem Mittelwertsatz, ein ξ^2 zwischen x_0^2 und x^2 so, dass

$$\Xi(x^1, x^2) = \partial_2(\partial_1 F(\xi^1, \xi^2)) = \partial_{21} F(\xi^1, \xi^2) \quad (10.21)$$

Für gegebenes $\epsilon > 0$ existiert dann wegen der Stetigkeit von $\partial_{21} F$ ein $\delta > 0$ so, dass

$$|\Xi(x^1, x^2) - \partial_{21} F(x_0^1, x_0^2)| = |\partial_{21} F(\xi^1, \xi^2) - \partial_{21} F(x_0^1, x_0^2)| < \epsilon \quad (10.22)$$

falls $\|x - x_0\| < \delta$ (was $\|\xi - x_0\| < C \cdot \delta$ für eine Konstante C impliziert). Im Limes $x^2 \rightarrow x_0^2$ folgt wegen der partiellen Differenzierbarkeit von F in die $(0, 1)$ -Richtung

$$\left| \frac{\partial_2 F(x^1, x_0^2) - \partial_2 F(x_0^1, x_0^2)}{x^1 - x_0^1} - \partial_{21} F(x_0^1, x_0^2) \right| < \epsilon \quad (10.23)$$

falls $\|(x^1 - x_0^1, 0)\| < \delta$. Dies zeigt Existenz und Wert von $\partial_1 \partial_2 F(x_0^1, x_0^2)$ □

· Für $k \geq 0$ heisst eine Funktion $F : U \rightarrow \mathbb{R}$ k mal stetig partiell differenzierbar, wenn alle beliebig gemischten partiellen Ableitungen k -ter Ordnung existieren und stetig sind. Notation für den Vektorraum solcher Funktionen: $\mathcal{C}^k(U)$. Für höherdimensionalen Wertebereich gilt 10.7 sinngemäss weiter und man schreibt $\mathcal{C}^k(U, \mathbb{R}^m)$.

· Jedenfalls ist also für $F \in \mathcal{C}^2(U)$ die “Hesse-Matrix” $H(x) = (H_{ij}(x))_{i,j=1,\dots,n}$,

$$H_{ij}(x) = \partial_i \partial_j F(x) \quad (10.24)$$

der zweiten partiellen Ableitungen symmetrisch: $H(x)^T = H(x)$.

§ 10. MEHRERE VERÄNDERLICHE

Zur Meditation: · An jeder (unterdrückten) Stelle $x \in U$ ist $H = (H_{ij})_{i,j=1,\dots,n}$ die Matrix-Darstellung einer symmetrischen bilinearen Form

$$H : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

$$(v, w) \mapsto \sum_{i,j=1}^n H_{ij} v^i w^j = H(v, w) := D(DF(v))(w) \quad (10.25)$$

· Demgegenüber ist für $F \in \mathcal{C}^1(U, \mathbb{R}^m)$ die Jacobi-Matrix $(DF_i^j)_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$ die Matrix-Darstellung einer linearen Abbildung (dem Differential von F), siehe (10.11)

$$DF : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$v \mapsto \left(\sum_{i=1}^n DF_i^j v^i \right)_{j=1,\dots,m} = DF(v) \quad (10.26)$$

Zur Erinnerung: · Eine symmetrische Bilinearform $H : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ (oder eine sie darstellende Matrix) heisst positiv definit, falls für alle $v \in \mathbb{R}^n \setminus \{0\}$ gilt, dass $H(v, v) > 0$. (Positiv semi-definit: $\geq$). Vgl. auch (4.9).

· Eine Teilmenge $U \subset \mathbb{R}^n$ heisst konvex, falls $\forall x_1, x_2 \in U$ und alle $t \in (0, 1)$, $(1-t)x_1 + tx_2 \in U$. Dann heisst eine Funktion $F : U \rightarrow \mathbb{R}$ konvex, falls für alle solche x_1, x_2, t gilt, dass $F((1-t)x_1 + tx_2) \leq (1-t)F(x_1) + tF(x_2)$.

Ohne Beweis: In Verallgemeinerung von 9.9 gilt:

Proposition 10.8. *Sei $U \subset \mathbb{R}^n$ offen und konvex. Dann ist $F \in \mathcal{C}^2(U, \mathbb{R})$ genau dann konvex auf U falls $H(x)$ an jeder Stelle $x \in U$ positiv semi-definit ist.*

Proposition/Definition 10.9. *Ist $F : U \rightarrow \mathbb{R}$ differenzierbar, so heisst eine Stelle $x_0 \in U$ mit $DF(x_0) = 0$ ein kritischer Punkt von F . Ist $F \in \mathcal{C}^2(U)$, so gilt: Ein kritischer Punkt ist ein isoliertes lokales Maximum/Minimum, falls die Hesse-Matrix dort negativ/positiv definit ist. Ist die Hesse-Matrix indefinit,¹³ dann ist x_0 kein lokales Extremum.¹⁴*

Beweisidee. Benutze die Hesse-Matrix zur quadratischen Approximation der Funktion im Sinne einer mehr-dimensionalen Taylor-Formel. $\square$

Gradient, Divergenz, und Rotation

Wie bereits erwähnt ist, ausser der technischen Spitzfindigkeit, dass nicht jede partiell differenzierbare Funktion differenzierbar ist, eine wesentliche Motivation für die Einführung des totalen Differentials die gleichberechtigte Behandlung aller unabhängigen Variablen. In der theoretischen Physik kommt diese Sichtweise spätestens in der Hydro- und Elektrodynamik in der Form der Vektordifferentialoperatoren im $\mathbb{R}^3$ zum Tragen. Wir erklären hier kurz den Zusammenhang zwischen diesen scheinbar recht verschiedenen Ableitungsbegriffen.

¹³D.h., es existieren sowohl Richtungen v , für die $H(v, v) > 0$ als auch solche, für die $H(v, v) < 0$.

¹⁴D.h., es existieren in jeder Umgebung von x_0 sowohl Punkte x , für die $F(x) > F(x_0)$ als auch solche, für die $F(x) < F(x_0)$.

Dazu ist es unerlässlich sich klarzumachen, dass im mathematischen Bild des physikalischen Raumes (üblicherweise zunächst stillschweigend) mehr Strukturen ausgezeichnet sind, als wir bisher benutzt haben. (Tatsächlich beruht ein Grossteil des Fortschritts in die moderne Physik darauf, solche Einsichten in die Trennung von Strukturen ernstzunehmen.) Wir können hier zwar noch nicht auf alle Feinheiten eingehen, in gewisser Näherung aber verhält es sich folgendermassen.

- Als Definitionsbereiche (“Laboratorien”) dienen weiterhin offene Teilmengen $U \subset \mathbb{R}^3$. Da U im Allgemeinen nicht unter den Vektorraumoperationen abgeschlossen ist, können wir zur Diskussion der Argumente eigentlich nur die topologische Struktur eines metrischen Raumes verwenden. Denken wir bei Punkten $x \in U$ an Tripel von reellen Zahlen (x^1, x^2, x^3) , so sind dies bereits “affine Koordinaten bezüglich eines ausgezeichneten Bezugssystems”. Andere Bezugssysteme erhält man durch “passive” Translationen und Drehungen, allenfalls auch allgemeine lineare Transformationen. Zu sogenannten krummlinigen Koordinatensystemen später mehr.

- Auf dem Raum der (infinitesimalen) Variationen hingegen existiert eine Vektorraumstruktur: An jeder Stelle $x_0 \in U$ ist der Weg $(-\delta, \delta) \ni t \mapsto x_0 + tv$ in jede Richtung $v \in \mathbb{R}^3$ für ein $\delta > 0$ wohldefiniert, und wir können solche Richtungen addieren und mit Skalaren multiplizieren. (Unter Umständen aber ändert sich dabei das δ !)

- Man nennt den U umgebenden Vektorraum $\mathbb{R}^3$, angeheftet an die Stelle $x_0 \in U$, den “Tangentialraum an/zur U in x_0 ”. Wir schreiben $T_{x_0}U$ für den abstrakten Vektorraum und (e_1, e_2, e_3) für die Standardbasis in der kanonischen Identifikation $T_{x_0}U \cong \mathbb{R}^3$.

$$T_{x_0}U \ni v = \sum_{i=1}^3 v^i e_i = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \in \mathbb{R}^3 \quad (10.27)$$

- Der Dualraum $T_{x_0}^*U = \text{Hom}_{\mathbb{R}}(T_{x_0}U, \mathbb{R})$ heisst Kotangentialraum, und sollte sorgfältig von $T_{x_0}U$ unterschieden werden. Und zwar dies, obwohl er mit der Dualbasis $(\lambda^1, \lambda^2, \lambda^3)$ (mit $\lambda^j(e_i) = \delta_i^j$) ausgerüstet wieder kanonisch isomorph zu einem Vektorraum von Tripeln von reellen Zahlen ist. Wir wollen ihn zur Unterscheidung mit $\mathbb{R}^{3*}$ anschreiben.

$$T_{x_0}^*U \ni \alpha = \sum_{j=1}^3 \alpha_j \lambda^j = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^{3*} \quad (10.28)$$

- Erst die Auszeichnung eines inneren Produkts $\langle v, w \rangle := \sum_{i,j=1}^3 v^i g_{ij} w^j$ durch Angabe einer symmetrischen, positiv-definiten Matrix $(g_{ij})_{i,j=1,2,3}$ erlaubt uns eine Identifikation von $\mathbb{R}^3$ mit $\mathbb{R}^{3*}$. Wir definieren für $v \in \mathbb{R}^3$, $\lambda \in \mathbb{R}^{3*}$ musikalisch $v^\flat \in \mathbb{R}^{3*}$ und $\lambda^\sharp \in \mathbb{R}^3$ durch

$$v^\flat(w) := \langle v, w \rangle, \quad \langle \lambda^\sharp, w \rangle := \lambda(w) \quad \forall w \in \mathbb{R}^3 \quad (10.29)$$

Dass es sich hier tatsächlich um Isomorphismen handelt, wird in der linearen Algebra gezeigt. Man sieht auch leicht, dass bezüglich den oben ausgezeichneten Basen

$$(v^\flat)_j = \sum_{i=1}^3 v^i g_{ij} \quad (\lambda^\sharp)^i = \sum_{j=1}^3 g^{ij} \lambda_j \quad (10.30)$$

§ 10. MEHRERE VERÄNDERLICHE

wobei (g^{ij}) die zu (g_{ij}) inverse Matrix ist.

· Wie ebenfalls aus der linearen Algebra bekannt, könnte man durch geeignete Wahl der Koordinaten erreichen, dass die “Metrik” g durch die Einheitsmatrix dargestellt wird. Dann wären wir zurück bei (4.7) und der vertrauten 3-dimensionalen euklidischen Geometrie. In krummlinigen Koordinaten wäre g nicht konstant, was zu wichtigen Ergänzungen der folgenden Formeln führt. Wir bleiben hier bei konstantem, aber sonst beliebigem g , und bezeichnen mit $\|\cdot\|$ die zugehörige Norm auf $\mathbb{R}^3$.

· Sei nun $F \in \mathcal{C}^1(U)$. Dann ist für jedes $x \in U$ das Differential von F eine lineare Abbildung $T_x U = \mathbb{R}^3 \rightarrow \mathbb{R}$, also kanonisch ein Element von $T_x^* U = \mathbb{R}^{3*}$. Durch Nachschalten des Isomorphismuses (10.29) erhalten wir einen Tangentialvektor. Dies ist der Gradient von F .

$$\text{grad } F(x) = \nabla F(x) = (DF(x))^\sharp \quad (10.31)$$

Anschaulich gibt der Gradient die “Richtung der grössten Steigung” von F , denn (für $DF(x) \neq 0$)

$$\begin{aligned} DF(x) \left(\frac{\text{grad } F(x)}{\|\text{grad } F(x)\|} \right) &= \langle \text{grad } F(x), \text{grad } F(x) \rangle^{1/2} \\ &= \max\{DF(x)(v) \mid v \in T_x U, \langle v, v \rangle = 1\} \end{aligned} \quad (10.32)$$

· Eine Abbildung $Y : U \rightarrow \mathbb{R}^3$ (für fixiertes $U \subset \mathbb{R}^3$) heisst (Tangential-)Vektorfeld auf U . Der Gradient einer Funktion $F \in \mathcal{C}^2(U)$ ist ein Beispiel eines differenzierbaren Vektorfeldes.

· Das Differential eines Vektorfeldes $Y \in \mathcal{C}^1(U, \mathbb{R}^3)$ an der Stelle $x \in U$ ist eine lineare Abbildung $DY(x) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Die Divergenz des Vektorfeldes ist die Spur dieser linearen Abbildung. Durch partielle Ableitungen der Komponenten ausgedrückt ist dies einfach

$$\text{div } Y(x) = \text{tr } DY(x) = \sum_{i=1}^3 \partial_i Y^i(x) \quad (10.33)$$

Anschaulich gibt die Divergenz (im geeigneten Kontext) die “Quellendichte” des Vektorfeldes an, was wir aber erst durch die Integralsätze werden begründen können. Jedenfalls klingt hier ein Begriff des Volumens an.

· Zuletzt ist die Rotation eines Vektorfeldes über zwei Ecken definiert als das “Duale des antisymmetrischen Teils des Differentials bzgl. dem orientierten Volumenelement”. Dazu erinnern wir daran, dass das Kreuzprodukt von zwei linear unabhängigen Vektoren $v_1, v_2 \in \mathbb{R}^3$ definiert ist als “derjenige Vektor $v_1 \times v_2 =: v_3 \in \mathbb{R}^3$ mit $\langle v_3, v_1 \rangle = \langle v_3, v_2 \rangle = 0$, $\|v_3\| = \text{Flächeninhalt des von } v_1, v_2 \text{ aufgespannten Parallelgramms, und } (v_1, v_2, v_3) \text{ positiv orientiert}”$. Dadurch ist $v_1 \times v_2$ eindeutig bestimmt, ebenso wie $\text{rot } Y$ durch

$$\langle \text{rot } Y(x), v_1 \times v_2 \rangle = \langle DY(x)(v_1), v_2 \rangle - \langle DY(x)(v_2), v_1 \rangle \quad (10.34)$$

für alle $v_1, v_2 \in \mathbb{R}^3$. Für $g = \text{id}$ und Standard-Orientierung ergibt dies die bekannte Formel

$$\text{rot } Y(x) = \begin{pmatrix} \partial_2 Y^3 - \partial_3 Y^2 \\ \partial_3 Y^1 - \partial_1 Y^3 \\ \partial_1 Y^2 - \partial_2 Y^1 \end{pmatrix} \quad (10.35)$$

Anschaulich gibt die Rotation die “gerichtete Wirbeldichte” des Vektorfeldes an.
 · Im Übrigen gelten die in der theoretischen Physik aufgestellten Rechenregeln und Gesetze. $\operatorname{rot} \operatorname{grad} F = 0$, $\operatorname{div} \operatorname{rot} Y = 0$. Der Laplace-Operator ist

$$\Delta(F) := \operatorname{div} \operatorname{grad} F = \sum_{i,j=1}^3 g^{ij} \partial_i \partial_j F \quad (10.36)$$

§ 11 Komplex differenzierbare Funktionen

Wir diskutieren in diesem § den Begriff der Differenzierbarkeit für komplexwertige Funktionen $f : U \rightarrow \mathbb{C}$ einer komplexen Veränderlichen $z \in U \subset \mathbb{C}$. Da per Definition $\mathbb{C} \cong \mathbb{R}^2$ als zwei-dimensionaler reeller Vektorraum (sogar mit ausgezeichnete Basis, und die metrische Struktur/Norm stimmt auch), mag es scheinen, dass es sich hier allenfalls um einen Spezialfall der mehr-dimensionalen Differentiation handeln kann, wenn auch vielleicht einer mit etwas physikalischer Relevanz, da $2 < 3$ vom Ende des § 10. Es könnte kaum ein Schein mehr trügen.

Benutzt man nämlich zur Differentiation von Funktionen einer komplexen Veränderlichen nicht nur die reell-lineare Struktur des $\mathbb{R}^2$, sondern ausserdem die Körperstruktur von $\mathbb{C}$, so erhält man erstaunlicher Weise einen Ableitungsbegriff, der sich einerseits praktisch aller Eigenschaften der reellen Ableitung erfreut (natürlich nicht derjenigen, die von der Anordnung abhängen), andererseits aber die Funktionen derartig einschränkt, dass viele der technischen Komplikationen (wie Stetigkeit der Ableitung etc.) gleichsam von selbst verschwinden.

Spätestens in der Quantenmechanik lernt man, dass komplexe Funktionswerte in der Physik eine enorm wichtige Rolle spielen. Komplexe Definitionsbereiche treten hingegen nach wie vor nicht direkt in den Bildmengen physikalischer Theorien auf, sie bieten aber ein unverzichtbares praktisches Hilfsmittel. Wir werden Beispiele solcher “Lösung reeller Probleme über den Umweg ins Komplexe” im Kapitel 6 vorstellen.

Definition 11.1. Sei $U \subset \mathbb{C}$ eine offene Menge. Eine Funktion $f : U \rightarrow \mathbb{C}$ heisst komplex differenzierbar in $z_0 \in U$, falls der Differenzenquotient

$$\frac{f(z) - f(z_0)}{z - z_0} \quad (11.1)$$

sich stetig in z_0 fortsetzen lässt. In diesem Fall heisst der Wert der Fortsetzung in z_0 die Ableitung von f in z_0 . Schreibweise: $f'(z_0) \in \mathbb{C}$.

· Eine Funktion heisst holomorph in z_0 , wenn sie in einer offenen Umgebung von z_0 komplex differenzierbar ist, und holomorph auf U , wenn sie dort überall komplex differenzierbar ist.

Bemerkungen. · Wie bereits angedeutet ist diese Definition formal identisch zu 9.1, und erlaubt eine offensichtliche Verallgemeinerung zu normierten *komplexen* Vektorräumen als Wertebereich.

· Es gelten die gewohnten Produkt- und Kettenregeln. Polynome sind überall komplex differenzierbar, und rationale Funktionen ausserhalb der Nullstellen des Nenners. Exponential- und trigonometrische Funktionen behalten ebenfalls ihre gewohnten Ableitungen.

§ 11. KOMPLEX DIFFERENZIERBARE FUNKTIONEN

· Insbesondere ist eine komplex differenzierbare Funktion stetig. Es stellt sich heraus, dass die Ableitung einer holomorphen Funktion automatisch stetig (und sogar komplex differenzierbar) ist, weshalb der Begriff der “stetigen komplexen Differenzierbarkeit” nicht separat eingeführt wird.

· Als Beispiel betrachten wir die naive Fortsetzung¹⁵ der differenzierbaren, aber nicht stetig differenzierbaren Funktion aus Beispiel (9.5) auf komplexe Argumente, d.h.

$$\mathbb{C} \setminus \{0\} \ni z \mapsto f(z) = z^2 \sin \frac{1}{z} \quad (11.2)$$

Da für $z = iy$ rein imaginär $\sin \frac{1}{iy} = i \sinh \frac{1}{y} \rightarrow \pm i\infty$ für $y \rightarrow 0_{\pm}$ hat f noch nicht mal einen Grenzwert bei $z_0 = 0$, kann also dort auch nicht komplex differenzierbar sein. (Dieses Argument zeigt noch nicht die (wahre) Aussage, dass sich f überhaupt nicht als in 0 differenzierbare Funktion von der reellen Achse weg fortsetzen lässt.)

· Es gibt auch ganz vertraute komplexe Funktionen, die aus viel plumperen Gründen nirgends komplex differenzierbar sind. Wichtigstes Beispiel hierfür ist die komplexe Konjugation, $z \mapsto \bar{z}$. Mit $z - z_0 = r e^{i\varphi}$ für $\varphi \in [0, 2\pi)$ hat der Differenzenquotient

$$\frac{\bar{z} - \bar{z}_0}{z - z_0} = e^{-2i\varphi} \quad (11.3)$$

offensichtlich keinen Grenzwert für $z \rightarrow z_0$. (Im Unterschied zum vorherigen Beispiel *hat* die Einschränkung von $z \mapsto \bar{z}$ auf die reelle Achse, nämlich die Identität $x \mapsto x$ eine andere, komplex differenzierbare Fortsetzung ins Komplexe, nämlich $z \mapsto z$.)

· Durch Umschreiben von (11.1) als

$$\lim_{z \rightarrow z_0} \frac{|f(z) - f(z_0) - f'(z_0)(z - z_0)|}{|z - z_0|} = 0 \quad (11.4)$$

und Vergleich mit 10.1 sieht man aber auch, dass man die komplexe Differenzierbarkeit als Spezialfall des zwei-dimensionalen reellen Begriffs auffassen kann: Multiplikation mit $f'(z_0)$ ist eine lineare Abbildung $\mathbb{R}^2 \rightarrow \mathbb{R}^2$, und der Grenzwert $z \rightarrow z_0$ in $\mathbb{C}$ ist der gleiche wie $(x, y) \rightarrow (x_0, y_0)$ in $\mathbb{R}^2$ (wobei natürlich $z = x + iy$, $z_0 = x_0 + iy_0$).

· Zur Untersuchung des genauen Zusammenhangs zwischen den beiden Begriffen schreibt man $f(z) = u(x, y) + iv(x, y)$ mit reellwertigen Funktionen u, v von $(x, y) \in U \subset \mathbb{R}^2$, und nutzt aus, dass falls der Grenzwert von (11.1) existiert, er (per Definition) unabhängig davon ist, aus welcher (“reellen”) Richtung man sich z_0 nähert. Entlang der reellen Achse, d.h. $z = z_0 + t$ für $t \in \mathbb{R}$ findet man

$$f'(z_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t} = \frac{\partial u}{\partial x}(x_0, y_0) + i \frac{\partial v}{\partial x}(x_0, y_0) \quad (11.5)$$

Entlang der imaginären Achse ist $z = z_0 + it$ für $t \in \mathbb{R}$ und

$$f'(z_0) = \lim_{t \rightarrow 0} \frac{f(x_0, y_0 + t) - f(x_0, y_0)}{it} = -i \frac{\partial u}{\partial y}(x_0, y_0) + \frac{\partial v}{\partial y}(x_0, y_0) \quad (11.6)$$

Durch Koeffizientenvergleich ergibt sich:

¹⁵Wir hatten im Zusammenhang der Stetigkeit bereits den Begriff der Fortsetzung in einen Punkt eingeführt. Ganz allgemein ist die Fortsetzung einer Abbildung $F : X \rightarrow Y$ von einer Teilmenge $X \subset \tilde{X}$ auf $\tilde{X}$ eine Abbildung $\tilde{F} : \tilde{X} \rightarrow Y$ die auf X mit F übereinstimmt.

Lemma 11.2. Eine Funktion $f : U \rightarrow \mathbb{C}$ ist genau dann in $z_0 \in U$ komplex differenzierbar, wenn die zugehörige $\mathbb{R}^2$ -wertige Funktion in (x_0, y_0) (total) differenzierbar ist und die partiellen Ableitungen die Cauchy-Riemann-Gleichungen erfüllen:

$$\frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0) \quad \text{und} \quad \frac{\partial u}{\partial y}(x_0, y_0) = -\frac{\partial v}{\partial x}(x_0, y_0) \quad (11.7)$$

In diesem Fall ist die komplexe Ableitung gleichwertig durch (11.5) oder (11.6) gegeben.

Beweis. Es bliebe allenfalls noch nachzureichen, dass eine lineare Abbildung mit einer Matrix-Darstellung der Form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \text{Mat}_{2 \times 2}(\mathbb{R}) \quad (11.8)$$

sich als Multiplikation von $\mathbb{C}$ mit $a + ib$ identifizieren lässt. $\square$

Geometrische Interpretation

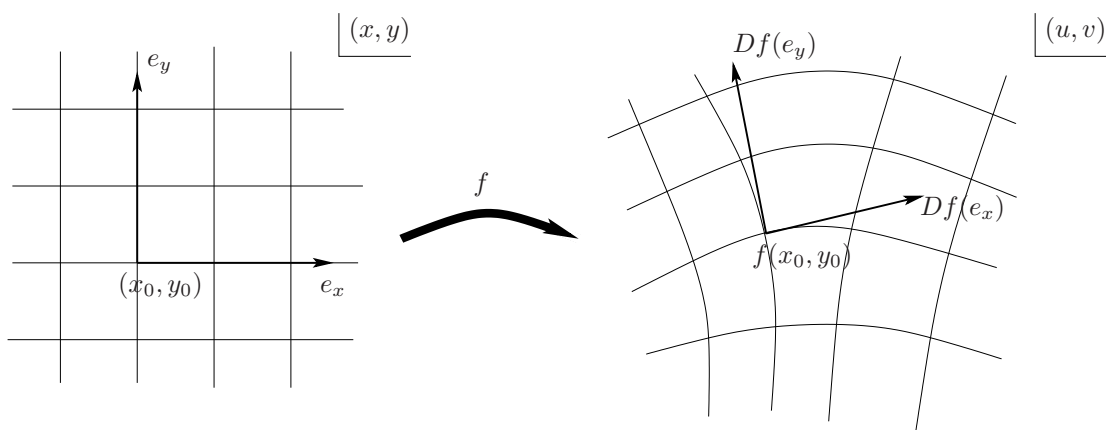
Wir nehmen (11.7) als Gelegenheit zur Schärfung unserer vektoranalytischen Anschauung. Dazu betrachten wir in der (u, v) -Ebene die Bilder der x - und y -Koordinatenrichtungen. Es ist

$$Df(e_x) = (\partial_x u, \partial_x v), \quad Df(e_y) = (\partial_y u, \partial_y v) \quad (11.9)$$

Als Konsequenz der Cauchy-Riemann-Gleichungen gilt bezüglich dem euklidischen inneren Produkt in Definitions- und Wertebereich

$$\langle Df(e_x), Df(e_y) \rangle = \langle e_x, e_y \rangle = 0, \quad \|Df(e_x)\| = \|Df(e_y)\| \quad (11.10)$$

und daraus folgt (im Fall $f'(z_0) \neq 0$), dass das Differential den Winkel zwischen je zwei beliebigen Richtungen im Tangentialraum $T_{(x_0, y_0)}U \cong \mathbb{R}^2$ erhält, Längen aber um einen gemeinsamen Faktor vergrößert oder verkleinert. Solche Abbildungen heißen “konform”. Vom komplexen Standpunkt ist dies aber nichts anderes als die in § 2 gegebene Interpretation der Multiplikation mit einer komplexen Zahl (hier: $f'(z_0)$) als “Drehstreckung”. (Die Umkehrung konform $\Rightarrow$ komplex differenzierbar gilt bis auf Orientierungswechsel/komplexe Konjugation.)



§ 11. KOMPLEX DIFFERENZIERBARE FUNKTIONEN

· Eine weitere von Ingenieuren und Physikern häufig benutzte Konsequenz aus den Cauchy-Riemann-Gleichungen ergibt sich unter der (nachträglich redundanten) Annahme, dass u und v mindestens zweimal stetig differenzierbar sind:

$$\begin{aligned}\partial_x^2 u &\stackrel{(11.7)}{=} \partial_x \partial_y v \stackrel{10.7}{=} \partial_y \partial_x v \stackrel{(11.7)}{=} -\partial_y^2 u \\ \Rightarrow \quad \partial_x^2 u + \partial_y^2 u &= 0\end{aligned}\tag{11.11}$$

Der Real- (und auch Imaginär-)Teil einer holomorphen Funktion löst also die zweidimensionale Laplace-Gleichung.

Analytische Funktionen

Die aber vielleicht wichtigste Quelle holomorpher Funktionen sind die im § 5 eingeführten Potenzreihen mit komplexen Koeffizienten. Es gilt nämlich

Proposition 11.3. *Eine Potenzreihe $P(z) = \sum_{n=0}^{\infty} a_n z^n$ ist an jeder Stelle $z_0 \in B_R(0)$ innerhalb ihrer Konvergenzscheibe komplex differenzierbar mit der (gliedweisen) Ableitung $P'(z) = \sum_{n=0}^{\infty} a_n n z^{n-1}$*

Beweis. Wir hatten bereits im § 8 gesehen, dass $P(z)$ auf $B_R(0)$ eine stetige Funktion darstellt. Der Nachweis der Differenzierbarkeit geht ganz ähnlich.

· Zunächst halten wir fest, dass die Reihe der gliedweisen Ableitungen $\sum a_n n z^{n-1}$ wegen $\lim_{n \rightarrow \infty} n^{1/n} = 1$ den gleichen Konvergenzradius hat wie P (siehe (5.26)) und daher ebenso wie P für jedes $|z| < R$ absolut konvergiert. Insbesondere ist Umordnen erlaubt.

· Es sei r so, dass $|z_0| < r < R$. Für gegebenes $\epsilon > 0$ sei dann N so gross, dass

$$\sum_{n=N+1}^{\infty} |a_n| n r^{n-1} < \frac{\epsilon}{3}\tag{11.12}$$

Dann ist für alle $|z| < r$

$$\begin{aligned}\left| \sum_{n=0}^{\infty} a_n z^n - \sum_{n=0}^{\infty} a_n z_0^n - \sum_{n=0}^{\infty} a_n n z_0^{n-1} (z - z_0) \right| \\ \leq \left| \sum_{n=0}^N (a_n z^n - a_n z_0^n - a_n n z_0^{n-1} (z - z_0)) \right| \\ + \underbrace{\sum_{n=N+1}^{\infty} |a_n| \cdot \underbrace{|z^n - z_0^n|}_{\leq n r^{n-1} |z - z_0|} + \sum_{n=N+1}^{\infty} |a_n| n r^{n-1} |z - z_0|}_{< \frac{2}{3} \epsilon |z - z_0|}\end{aligned}\tag{11.13}$$

Da Polynome komplex differenzierbar sind, existiert ein $\delta > 0$ so, dass die endliche Summe für $|z - z_0| < \delta$ kleiner als $\frac{\epsilon}{3} \cdot |z - z_0|$ ist. Zusammen ist alles kleiner als $\epsilon \cdot |z - z_0|$, falls $z \in B_\delta(z_0) \cap B_r(0)$ d.h. es folgt die Behauptung. $\square$

Definition 11.4. Eine Funktion $f : U \rightarrow \mathbb{C}$ heisst (komplex) analytisch, falls es für jeden Punkt $z_0 \in U$ eine Folge $(a_n^{(z_0)})$ komplexer Zahlen und ein $\delta > 0$ gibt so, dass $B_\delta(z_0) \subset U$ und für alle $z \in B_\delta(z_0)$

$$f(z) = \sum_{k=0}^{\infty} a_k^{(z_0)} (z - z_0)^k \quad (11.14)$$

- Insbesondere hat die Potenzreihe auf der rechten Seite positiven Konvergenzradius $R^{(z_0)}$. Es muss aber nicht $B_{R^{(z_0)}}(z_0) \subset U$ gelten.
- Durch wiederholte Anwendung von 11.3 folgt:

Korollar 11.5. Eine analytische Funktion ist in ihrem Definitionsbereich unendlich oft komplex differenzierbar, und es gilt die “Taylor-Formel”

$$f^{(k)}(z_0) = k! a_k^{(z_0)} \quad (11.15)$$

Bemerkungen/Beispiele 11.6. · Umgekehrt ist eine durch eine Potenzreihe $P(z) = \sum a_n z^n$ dargestellte Funktion analytisch: Für $z_0 \in B_R(0)$ erhält man durch “Umentwickeln”¹⁶ von $P(z_0 + (z - z_0))$ eine Potenzreihendarstellung der Form (11.14), welche mindestens für $z \in B_\delta(z_0) \subset B_R(0)$ auch wieder die gegebene Funktion darstellt.

- Beispiel: Geometrische Reihe zur Illustration der Idee der analytischen Fortsetzung.
- Beispiel: $Q(z) = z^2$ als biholomorphe Abbildung $\{z \mid \operatorname{Re}(z) > 0\} \rightarrow \mathbb{C}_- = \mathbb{C} \setminus \mathbb{R}_{\leq 0}$

Wichtige Ableitungen als Übungsaufgaben

Potenzreihen mit Argumenten aus endlich-dimensionalen Banach-Algebren sind innerhalb ihrer Konvergenzkugel ebenfalls gliedweise ableitbar. Allerdings gilt für die Ableitung im nicht-kommutativen Fall die etwas allgemeinere Formel (vgl. schon (9.3))

$$DP(x_0)(x - x_0) = \sum_{n=1}^{\infty} a_n \sum_{k=0}^{n-1} x_0^k (x - x_0) x_0^{n-k-1} \quad (11.17)$$

(Annahme: $a_n \in \mathbb{C}$) Jedenfalls ist immer

$$D \exp(0) = \operatorname{id}_{\mathcal{A}} \quad (11.18)$$

- Für $x \in \operatorname{Mat}_{n \times n}(\mathbb{R})$ mit $\det x \neq 0$ gilt

$$D \det(x)(v) = \det x \operatorname{tr}(x^{-1}v) \quad (11.19)$$

- Quaternionische Ableitung

¹⁶Dies ist nicht strenggenommen eine direkte Konsequenz aus dem Umordnungssatz von S. 36. Auf jeden Fall konvergiert aber

$$a_k^{(z_0)} = \frac{1}{k!} P^{(k)}(z_0) = \sum_{n=k}^{\infty} \binom{n}{k} a_n z_0^{n-k} \quad (11.16)$$

Die Aussage über Konvergenz der umentwickelten Reihe lässt sich am einfachsten mit den Methoden aus Kapitel 6 zeigen.

KAPITEL 5

UMKEHRUNG

Wir hatten die Einführung der reellen Zahlen und der Grenzprozesse motiviert durch die Feststellung, dass gewisse natürlich auftretende mathematische Probleme sich mit endlich vielen Rechenschritten höchstens “beliebig genau”, aber nie “exakt” lösen lassen. Mit dem Fundamentalsatz der Algebra hatten wir dann gesehen, dass es in diesem idealisierten Rahmen manchmal möglich ist, die Existenz von Lösungen nachzuweisen, auch ohne ein explizites Verfahren anzugeben, mit dem man solche Lösungen, und sei es nur approximativ, finden kann.

Wir stellen in diesem Kapitel zwei Klassen von Problemen vor, die im Rahmen unserer Differentialrechnung formuliert und lösbar sind. Die Resultate sind zwar einerseits recht allgemeine und abstrakte Existenzaussagen, wie sie auch in der weiteren theoretischen Entwicklung häufig verwendet werden. Die Beweise beruhen aber andererseits auf iterativen Näherungsverfahren, wie sie auch in konkreten Beispielen (zumindestens sinngemäss) implementiert werden können.

Die grundlegende Strategie hinter diesen Lösungsverfahren ist dabei recht simpel. Man identifiziert und löst zunächst eine linearisierte Version des Problems im Rahmen geeignet normierter Vektorräume. Dann stellt man durch geschickte Abschätzungen sicher, dass, ausgehend von einer approximativen Lösung des gegebenen allgemeinen (nicht-linearen) Problems, die Lösung des linearen Problems die approximative Lösung hinreichend verbessert. Zuletzt garantiert dann die Vollständigkeit des umgebenden Raumes, dass die Iteration des linearen Lösungsverfahrens zu einer Lösung des nicht-linearen Problems konvergiert.

§ 12 Gleichungen im $\mathbb{R}^n$

Bei der ersten Klasse von Problemen ist unsere Aufgabe die Umkehrung einer gegebenen Abbildung $F : U \rightarrow \mathbb{R}^n$ (für $U \subset \mathbb{R}^n$ offen), m.a.W. fragen wir nach der Lösbarkeit der Gleichung $F(x) = y$ für $y \in \mathbb{R}^n$ und nach den Eigenschaften einer etwaigen Umkehrabbildung in Abhängigkeit derer von F .

Das Kriterium, das wir geben wollen, ist eine mehr-dimensionale Alternative zum strengen Monotonie-Kriterium 7.2 für die Umkehrbarkeit stetiger Funktionen. Dieses Kriterium steht uns ja im Allgemeinen nicht zur Verfügung, da $\mathbb{R}^n$ nicht angeordnet ist. Stattdessen halten wir fest, dass für eine differenzierbare Funktion strenge Monotonie gesichert ist, falls die Ableitung ein konstantes Vorzeichen hat, oder als Konsequenz dessen ausgedrückt, als reelle Zahl an jeder Stelle invertierbar ist. Diese Aussage lässt sich nämlich verallgemeinern zur Invertierbarkeit des Differentials von F als lineare Abbildung $DF(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Ein zweites Zugeständnis müssen wir der Beobachtung aus den Übungen machen, dass für einen allgemeinen Definitionsbereich selbst in dem Fall, in dem eine stetige Funktion global invertierbar ist, die Umkehrfunktion nicht notwendig überall stetig ist. Wir erwarten daher auch im Allgemeinen eine Übertragung der Eigenschaften von F (hier: stetige Differenzierbarkeit) nur nach Einschränkung auf geeignet kleine Umgebungen.

Zur Einstimmung auf die Methode erinnern wir an das Rekursionsverfahren zur Bestimmung von Quadratwurzeln positiver reeller Zahlen (vgl. 3.15). Für $a > 0$ definieren wir eine Abbildung $S^{(a)} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ durch

$$S^{(a)}(x) := x - \frac{1}{2x}(x^2 - a) = \frac{1}{2}\left(x + \frac{a}{x}\right) \quad (12.1)$$

Mit anderen Worten ist $S^{(a)}$ die Lösung ξ der linearen Gleichung

$$x^2 - a + 2x(\xi - x) = 0 \quad (12.2)$$

oder anschaulich gesagt, die Koordinate des Schnittpunkts der Tangenten an den Graphen von $f^{(a)}(x) := x^2 - a$ im Punkte x mit der x -Achse. Dann gilt $S^{(a)}(x) > \sqrt{a} \forall x > 0$, und $S^{(a)}(x) < x$ falls $x > \sqrt{a}$. Daraus leitet man ab, dass für jeden Startpunkt x_0 die Folge $(x_n := (S^{(a)})^n(x_0))_{n=1,2,\dots}$ monoton fällt und für $n \rightarrow \infty$ gegen $\sqrt{a}$ konvergiert.

Die Konvergenz dieser "Newton-Methode" ist zwar nicht direkt ein Spezialfall des folgenden Beweises, die Stellung der Ableitung $2x = f^{(a)'}(x)$ in (12.1) ist aber analog zu der des Differentials in (12.5) unten. Unser allgemeines Arbeitstier zur Sicherung der Konvergenz ist der Banachsche Fixpunktsatz 4.19, den wir an dieser Stelle kurz wiederholen. Dann aber gehen wir zum

Theorem 12.1 (Satz von der Umkehrabbildung). *Sei $U \subset \mathbb{R}^n$ offen und $F : U \rightarrow \mathbb{R}^n$ stetig differenzierbar. Für ein $x_0 \in U$ sei $DF(x_0)$ nicht entartet (d.h., eine invertierbare lineare Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^n$). Dann existiert eine offene Menge $V_0 \subset \mathbb{R}^n$ mit $F(x_0) \in V_0$, eine offene Menge $U_0 \subset U$ mit $x_0 \in U_0$ und eine stetig differenzierbare Abbildung $G : V_0 \rightarrow U_0$ so, dass $F \circ G = \text{id}_{V_0}$, $G \circ F|_{U_0} = \text{id}_{U_0}$. In Worten: Die Einschränkung von F auf eine hinreichend kleine Umgebung von x_0 ist ein Diffeomorphismus auf das Bild dieser Einschränkung.*

Beweis. Wir konstruieren zunächst eine Umkehrabbildung zu F in einer geeigneten Umgebung von $y_0 := F(x_0)$ mit Hilfe des Banachschen Fixpunktsatzes 4.19 und zeigen anschliessend, dass diese Abbildung differenzierbar ist.

Zur Abkürzung schreiben wir $D_0 := DF(x_0) \in \text{End}(\mathbb{R}^n)$ für das Differential von F in x_0 . Gemäss Annahme existiert $D_0^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Wir schreiben die Operatornorm wie auch die (beliebig gewählte) Norm auf $\mathbb{R}^n$, auf der sie beruht, als $\|\cdot\|$.

Auswahl der Umgebungen: Da F in x_0 differenzierbar ist, existiert (zu $\epsilon_* = \frac{1}{2\|D_0^{-1}\|}$) ein $r > 0$ so, dass

$$\|F(x) - F(x_0) - D_0(x - x_0)\| \leq \frac{1}{2\|D_0^{-1}\|} \cdot \|x - x_0\| \quad (12.3)$$

§ 12. GLEICHUNGEN IM $\mathbb{R}^n$

$\forall x \in \overline{B_r(x_0)} \subset U$.¹⁷ Da ferner $DF : U \rightarrow \text{End}(\mathbb{R}^n)$ stetig ist, kann r so gewählt werden, dass ausserdem

$$\|DF(x) - D_0\| < \frac{1}{2\|D_0^{-1}\|} \quad (12.4)$$

$\forall x \in \overline{B_r(x_0)}$. Insbesondere ist $DF(x)$ invertierbar $\forall x \in B_r(x_0)$ (geometrische Reihe).

Wir setzen nun $s = \frac{r}{2\|D_0^{-1}\|}$ und $V_0 = B_s(y_0)$.

Konstruktion von G : Für $y \in V_0$ betrachten wir die Abbildung

$$\begin{aligned} T^{(y)} : \overline{B_r(x_0)} &\rightarrow \mathbb{R}^n \\ T^{(y)}(x) &:= x - D_0^{-1}(F(x) - y) \end{aligned} \quad (12.5)$$

(Bemerke, dass Lösungen von $F(x) = y$ genau die Fixpunkte von $T^{(y)}$ sind!) Es gilt

$$\begin{aligned} \|T^{(y)}(x) - x_0\| &= \|x - x_0 - D_0^{-1}(F(x) - F(x_0)) + D_0^{-1}(y_0 - y)\| \\ &\leq \underbrace{\|D_0^{-1}(F(x) - F(x_0) - D_0(x - x_0))\|}_{\leq \|D_0^{-1}\| \cdot \frac{1}{2\|D_0^{-1}\|} \cdot \|x - x_0\|} + \underbrace{\|D_0^{-1}\| \cdot \|y_0 - y\|}_{< \|D_0^{-1}\| \cdot s} \\ &< \frac{r}{2} + \frac{r}{2} = r \end{aligned} \quad (12.6)$$

$T^{(y)}$ bildet also $\overline{B_r(x_0)}$ auf $B_r(x_0)$ ab. $T^{(y)}$ ist allerdings auf ganz U differenzierbar mit der Ableitung $\text{id}_{\mathbb{R}^n} - D_0^{-1} \circ DF(x)$, und für $x_1, x_2 \in \overline{B_r(x_0)}$ (einer konvexen Menge) gilt daher wegen des Schrankensatzes 10.6 mit (12.4)

$$\begin{aligned} \|T^{(y)}(x_2) - T^{(y)}(x_1)\| &\leq \sup\{\|\text{id}_{\mathbb{R}^n} - D_0^{-1}DF(x)\| \mid x \in \overline{B_r(x_0)}\} \cdot \|x_2 - x_1\| \\ &\leq \frac{1}{2}\|x_2 - x_1\| \end{aligned} \quad (12.7)$$

$T^{(y)}$ ist also eine Kontraktion auf dem vollständigen metrischen Raum $\overline{B_r(x_0)}$ und hat gemäss 4.19 einen eindeutigen Fixpunkt $G(y)$, der wegen obiger Bemerkung die eindeutige Lösung von $F(x) = y$ in $\overline{B_r(x_0)}$ ist und wegen (12.6) in $B_r(x_0)$ liegt. Dies definiert die Umkehrabbildung $G : V_0 \rightarrow U_0 := F^{-1}(V_0) \cap B_r(x_0)$. ($F^{-1}(V_0)$ ist offen wegen der Stetigkeit von F , und U_0 als Durchschnitt zweier offener Mengen.)

Differenzierbarkeit von G : Es genügt, die Differenzierbarkeit von G in $y_0 = F(x_0)$ nachzuweisen. (Jeder andere Punkt $\tilde{y} \in V_0$ ist ja von der Form $\tilde{y} = F(\tilde{x})$ für ein $\tilde{x} \in B_r(x_0)$ und wegen der Invertierbarkeit von $DF(\tilde{x})$ liefert die obige Konstruktion eine Umkehrung $\tilde{G}$ von F in einer Kugel um $\tilde{y}$. Wegen der Eindeutigkeit der Fixpunkte muss $\tilde{G}$ mit G im Schnitt der beiden Kugeln übereinstimmen. Differenzierbarkeit von $\tilde{G}$ in $\tilde{y}$ impliziert dann die von G .) Wegen der Kettenregel (die wir selbst erst benutzen dürfen, wenn G differenzierbar ist) erwarten wir, dass die Ableitung von G in y_0 durch D_0^{-1} gegeben ist.

Auswerten von (12.7) für $x_1 = x_0$ und $x_2 = x = G(y)$ ergibt

$$\begin{aligned} \|x - x_0\| &= \|T^{(y)}(x) - T^{(y)}(x_0) - D_0^{-1}(y_0 - y)\| \\ &\leq \frac{1}{2}\|x - x_0\| + \|D_0^{-1}\| \cdot \|y - y_0\| \\ \Rightarrow \|x - x_0\| &\leq 2\|D_0^{-1}\| \cdot \|y - y_0\| \end{aligned} \quad (12.8)$$

¹⁷Die traditionsgemäss in der offenen Kugel geschriebene Ungleichung (10.2) setzt sich aus Stetigkeitsgründen auf den Abschluss fort, falls dieser in U enthalten ist.

d.h. $\|G(y) - G(y_0)\| \leq 2\|D_0^{-1}\| \cdot \|y - y_0\| \quad \forall y \in V_0$. Insbesondere ist G stetig in y_0 .

Ist nun $\epsilon > 0$, so existiert wegen der Differenzierbarkeit von F in x_0 ein $\zeta > 0$ so, dass

$$\|F(x) - F(x_0) - D_0(x - x_0)\| \leq \epsilon \cdot \|x - x_0\| \quad \forall x \in B_\zeta(x_0) \subset U_0 \quad (12.9)$$

Diese Aussage ist (trivialerweise) gleichbedeutend mit

$$\|D_0(G(y) - G(y_0)) - (y - y_0)\| \leq \epsilon \cdot \|G(y) - G(y_0)\| \quad \forall y \in G^{-1}(B_\zeta(x_0)) \quad (12.10)$$

Für $\delta \leq \frac{\zeta}{2\|D_0^{-1}\|}$ ist als Konsequenz von (12.8) $G(B_\delta(y_0)) \subset B_\zeta(x_0)$, und es gilt dann $\forall y \in B_\delta(y_0)$:

$$\begin{aligned} \|G(y) - G(y_0) - D_0^{-1}(y - y_0)\| &\leq \|D_0^{-1}\| \cdot \epsilon \cdot \|G(y) - G(y_0)\| \\ &\leq 2\|D_0^{-1}\|^2 \cdot \epsilon \cdot \|y - y_0\| \end{aligned} \quad (12.11)$$

Dies zeigt die Differenzierbarkeit von G in y_0 . Die Stetigkeit von $y \mapsto DG(y) = DF(G(y))^{-1}$ folgt aus der Tatsache, dass dies eine Verkettung der stetigen Funktionen $y \mapsto G(y) = x$, $x \mapsto DF(x)$ und $DF(x) \mapsto (DF(x))^{-1}$ ist. $\square$

Beispiel 12.2. Sei $\Sigma = (\sigma^1, \sigma^2, \sigma^3) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definiert durch die *elementarsymmetrischen* Funktionen

$$\begin{aligned} \sigma^1(x^1, x^2, x^3) &= x^1 + x^2 + x^3 \\ \sigma^2(x^1, x^2, x^3) &= x^1x^2 + x^1x^3 + x^2x^3 \\ \sigma^3(x^1, x^2, x^3) &= x^1x^2x^3 \end{aligned} \quad (12.12)$$

(Vgl. Aufgabe 4 auf Blatt 7.). Dann hat das Differential von Σ bezüglich den Standard-Basen die Matrix-Darstellung

$$D\Sigma(x) = \begin{pmatrix} 1 & 1 & 1 \\ x^2 + x^3 & x^1 + x^3 & x^1 + x^2 \\ x^2x^3 & x^1x^3 & x^1x^2 \end{pmatrix} \quad (12.13)$$

mit Determinante

$$\begin{aligned} \Delta = \det D\Sigma(x) &= (x^1)^2(x^2 - x^3) - (x^2)^2(x^1 - x^3) + (x^3)^2(x^1 - x^2) \\ &= (x^1 - x^2)((x^3)^2 + x^1x^2) + (x^2)^2x^3 - (x^1)^2x^3 \\ &= (x^1 - x^2)((x^3)^2 + x^1x^2 - (x^1 + x^2)(x^3)) \\ &= (x^1 - x^2)(x^1 - x^3)(x^2 - x^3). \end{aligned} \quad (12.14)$$

$D\Sigma(x)$ ist also an allen Stellen $x_0 \in \mathbb{R}^3$ invertierbar, an denen nicht zwei (oder mehr) Komponenten gleich sind. In einer (geeignet kleinen) Umgebung eines jeden solchen Punktes lässt sich Σ also lokal eindeutig und differenzierbar invertieren. Mit anderen Worten: Für gegebene $(\sigma_0^1, \sigma_0^2, \sigma_0^3) \in \Sigma(\mathbb{R}^3)$ ¹⁸ gibt es im Allgemeinen zwar mehrere (nämlich 6) mögliche (x^1, x^2, x^3) mit diesen Werten der elementarsymmetrischen Funktionen. Ist aber eine solche Wahl $x_0 = (x_0^1, x_0^2, x_0^3)$ getroffen, so legt diese

¹⁸ Σ ist nicht surjektiv, siehe Übungen.

§ 12. GLEICHUNGEN IM $\mathbb{R}^n$

weitere Wahlen in einer kleinen Umgebung eindeutig fest. Die “schlechten” Werte von $(\sigma^1, \sigma^2, \sigma^3)$ mit weniger als 6 verschiedenen Urbildpunkten sind die Nullstellen von

$$\Delta^2 = (\sigma^1 \sigma^2)^2 - 4(\sigma^2)^3 - 4(\sigma^1)^3 \sigma^3 + 18\sigma^1 \sigma^2 \sigma^3 - 27(\sigma^3)^2 \quad (12.15)$$

· Für eine endlich-dimensionale vollständige normierte Algebra $\mathcal{A}$ (z.B. $\text{Mat}_{n \times n}(\mathbb{R}) \cong \mathbb{R}^{n^2}$ als Vektorraum, versehen mit Matrixmultiplikation und einer submultiplikativen Norm) hatten wir die Exponentialabbildung $\exp : \mathcal{A} \rightarrow \mathcal{A}$ durch die Reihe (5.31) definiert. Wegen $\exp(x) \exp(-x) = 1$ liegt das Bild von $\mathcal{A}$ unter $\exp$ in der Gruppe der invertierbaren Elemente von $\mathcal{A}$. Gemäss (11.18) ist das Differential

$$D \exp(0) = \text{id}_{\mathcal{A}} \quad (12.16)$$

im Ursprung $0 \in \mathcal{A}$ invertierbar. Der Umkehrsatz besagt dann, dass $\exp$ eine geeignete Umgebung von $0 \in \mathcal{A}$ diffeomorph auf eine geeignete Umgebung von $1 \in \mathcal{A}$ abbildet.

· Der Satz über die Umkehrfunktion ist eine nicht-lineare Version der Aussage, dass “ n unabhängige Variable durch n unabhängige lineare Gleichungen eindeutig bestimmt sind”, oder in der Sprache der linearen Algebra ausgedrückt, dass für eine nicht-entartete lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ zu jedem $y \in \mathbb{R}^n$ ein eindeutiges x gehört mit $Ax = y$ (nämlich, $x = A^{-1}y$).

· Ebenfalls aus der linearen Algebra bekannt ist die Aussage, dass im Falle von $n - k < n$ unabhängigen Gleichungen, d.h. einer linearen Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^{n-k}$ von maximalem Rang $n - k$, sich die Lösungsmenge der Gleichung $Ax = y$ (für $y \in \mathbb{R}^{n-k}$) parametrisieren lässt durch einen Unterraum von $\mathbb{R}^n$ der Dimension k , dem “Kern” von A . Wir benutzen im Folgenden die Notation

$$\mathbb{R}_{\parallel}^k := \text{Ker } A = \{x_{\parallel} \in \mathbb{R}^n \mid Ax_{\parallel} = 0\} \quad (12.17)$$

und eine Identifikation $\mathbb{R}^n \cong \mathbb{R}_{\parallel}^k \times \mathbb{R}_{\perp}^{n-k}$, welche beispielsweise durch Vervollständigen einer Basis von $\mathbb{R}_{\parallel}^k$ zu einer Basis von $\mathbb{R}^n$ definiert werden kann. Bezüglich dieser Basis hat A die Block-Gestalt

$$A = (0_{(n-k) \times k}, A_{\perp}) \quad (12.18)$$

mit einer invertierbaren $(n - k) \times (n - k)$ Matrix $A_{\perp} \cong A|_{\mathbb{R}_{\perp}^{n-k}}$. Es gilt dann

$$\{x \in \mathbb{R}^n \mid Ax = y\} = \{(x_{\parallel}, (A_{\perp})^{-1}y) \mid x_{\parallel} \in \mathbb{R}_{\parallel}^k\} \quad (12.19)$$

Eine nicht-lineare Version dieser Aussage (mit Unterdrückung von $y = 0$) ist der folgende Satz.

Theorem 12.3 (Implizite Funktionen). *Seien $n > k$ natürliche Zahlen. Sei $U \subset \mathbb{R}^n$ offen und $F : U \rightarrow \mathbb{R}^{n-k}$ stetig differenzierbar. Sei $x_0 \in U$ eine Lösung der Gleichung $F(x_0) = 0$, in der das Differential $DF(x_0)$ maximalen Rang hat. Wir nehmen oBdA an, dass bezüglich einer Identifikation $\mathbb{R}^n = \mathbb{R}_{\parallel}^k \times \mathbb{R}_{\perp}^{n-k}$, $x = (x_{\parallel}, x_{\perp})$ die Einschränkung $D_{\perp}F(x_0)$ von $DF(x_0)$ auf $\mathbb{R}_{\perp}^{n-k}$ invertierbar ist. Dann existieren*

offene Umgebungen $V_0 \subset \mathbb{R}_{\parallel}^k$ von $x_{0,\parallel}$ und $W_0 \subset \mathbb{R}_{\perp}^{n-k}$ von $x_{0,\perp}$ mit $V_0 \times W_0 \subset U$, und eine stetig differenzierbare Abbildung $f : V_0 \rightarrow W_0$ so, dass für alle $x = (x_{\parallel}, x_{\perp}) \in V_0 \times W_0$ gilt:

$$F(x_{\parallel}, x_{\perp}) = 0 \iff x_{\perp} = f(x_{\parallel}) \quad (12.20)$$

M.a.W.,

$$F^{-1}(0) \cap V_0 \times W_0 = \{(x_{\parallel}, f(x_{\parallel})), x \in V_0\} \quad (12.21)$$

In Worten: Die Lösungsmenge von $n - k$ "unabhängigen" differenzierbaren Gleichungen lässt sich lokal als Graph einer Funktion durch k unabhängige Variable parametrisieren.

Beweis. · Die oBdA gemachte Voraussetzung über $D_{\perp}F(x_0)$ ist die Aussage, dass mit der Indizierung $x = (x_{\parallel}, x_{\perp}) = (x_{\parallel}^1, \dots, x_{\parallel}^k, x_{\perp}^{k+1}, \dots, x_{\perp}^n)$ die Untermatrix

$$\left(\frac{\partial F^j}{\partial x_{\perp}^i}(x_0) \right)_{\substack{i=k+1, \dots, n \\ j=1, \dots, n-k}} \quad (12.22)$$

der Jacobi-Matrix von F in x_0 invertierbar ist. Im Allgemeinen kann man dies bereits durch einfaches Umm Nummerieren der Koordinaten von $\mathbb{R}^n$ erreichen (s. Beispiel unten.)

· Wir erklären nun eine auxiliäre Abbildung $\Phi : U \rightarrow \mathbb{R}_{\parallel}^k \times \mathbb{R}^{n-k}$ durch

$$\Phi(x_{\parallel}, x_{\perp}) := (x_{\parallel}, F(x_{\parallel}, x_{\perp}))^T \quad (12.23)$$

Dann ist Φ auf U stetig differenzierbar, das Differential hat die Block-Gestalt

$$D\Phi(x) = \begin{pmatrix} \text{id}_{k \times k} & 0 \\ D_{\parallel}F(x) & D_{\perp}F(x) \end{pmatrix} \quad (12.24)$$

und ist wegen (12.22) in x_0 invertierbar. Nach dem Umkehrsatz 12.1 existieren daher offene Umgebungen $\tilde{V}_0$ von $\Phi(x_0) = (x_{0,\parallel}, 0)$ und U_0 von x_0 und eine stetig differenzierbare Abbildung $\Psi : \tilde{V}_0 \rightarrow U_0$, welche invers zu $\Phi|_{U_0}$ ist. Aus allgemeinen topologischen Betrachtungen folgt, dass wir oBdA annehmen können, dass $U_0 = U_{0,\parallel} \times U_{0,\perp}$ für offene Umgebungen $U_{0,\parallel}$ von $x_{0,\parallel}$ und $U_{0,\perp}$ von $x_{0,\perp}$. Wir schreiben bezüglich dieser Identifikation $\Psi = (\Psi_{\parallel}, \Psi_{\perp})$.

· Wir setzen $V_0 := \{x_{\parallel} \in U_{0,\parallel} \mid (x_{\parallel}, 0) \in \tilde{V}_0\}$ und $W_0 := U_{0,\perp}$. Dann ist V_0 eine offene Umgebung von $x_{0,\parallel}$, und W_0 eine offene Umgebung von $x_{0,\perp}$ und die Definition $f : V_0 \rightarrow W_0$ durch $f(x_{\parallel}) := \Psi_{\perp}(x_{\parallel}, 0)$ macht Sinn.

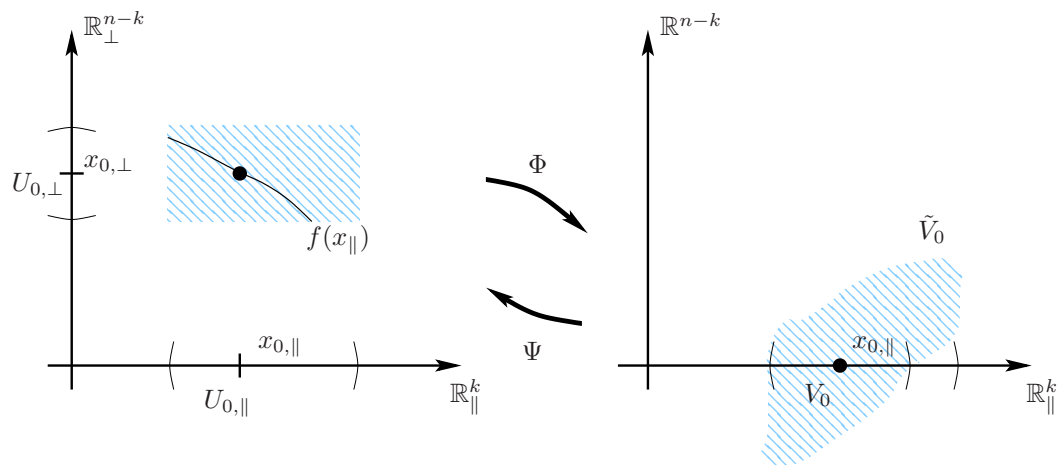
· f ist stetig partiell differenzierbar da Ψ dies ist, also nach 10.5 stetig differenzierbar.

· Ist für $(x_{\parallel}, x_{\perp}) \in V_0 \times W_0$, $F(x_{\parallel}, x_{\perp}) = 0$, so folgt $(x_{\parallel}, 0) = \Phi(x_{\parallel}, x_{\perp}) \in \tilde{V}_0$, d.h. $(x_{\parallel}, x_{\perp}) = \Psi(x_{\parallel}, 0)$, und daher $x_{\perp} = f(x_{\parallel})$.

· Gilt umgekehrt $x_{\perp} = f(x_{\parallel})$, so ist $(x_{\parallel}, x_{\perp}) = \Psi(x_{\parallel}, 0)$, d.h. $(x_{\parallel}, 0) = \Phi(x_{\parallel}, x_{\perp})$, und daher $F(x_{\parallel}, x_{\perp}) = 0$.

Zusammen folgt (12.20) □

§ 12. GLEICHUNGEN IM $\mathbb{R}^n$



Bemerkungen/Beispiele 12.4. · Aus dem Beweis folgt durch Anwendung der Kettenregel auf die Identität $F(x_{\parallel}, f(x_{\parallel})) = 0$ eine Formel für das Differential der implizit definierten Funktion f :

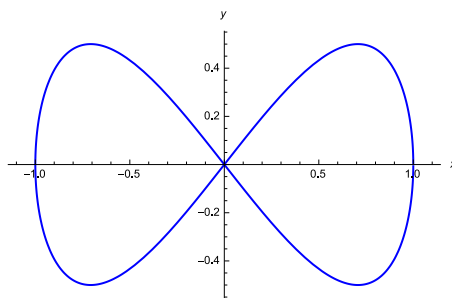
$$\begin{aligned} D_{\parallel}F(x_{\parallel}, f(x_{\parallel})) + D_{\perp}F(x_{\parallel}, f(x_{\parallel})) \circ Df(x_{\parallel}) &= 0 \\ \Rightarrow Df(x_{\parallel}) &= -D_{\perp}F(x_{\parallel}, f(x_{\parallel}))^{-1} \circ D_{\parallel}F(x_{\parallel}, f(x_{\parallel})) \end{aligned} \quad (12.25)$$

da $D_{\perp}F$ wegen der Invertierbarkeit von $D\Phi$ und (12.24) an den angegebenen Stellen invertierbar ist.

· Als Beispiel betrachten wir von der Funktion $F : \mathbb{R}^2 \rightarrow \mathbb{R}$, definiert durch $F(x, y) := x^2(1 - x^2) - y^2$ die Nullstellenmenge $N = \{(x, y) \in \mathbb{R}^2 \mid F(x, y) = 0\}$. Das Differential von F ist dargestellt durch die Jacobi-Matrix

$$DF(x, y) = (2x - 4x^3, -2y) \quad (12.26)$$

Man sieht leicht, dass ausser in $(x, y) = 0 \in N$ DF überall auf N maximalen Rang hat. Allerdings kann man nicht überall mit der gleichen Zerlegung von $\mathbb{R}^2$ in “parallele” und “orthogonale” Richtungen arbeiten. In der Umgebung von $y = 0$ muss man zur Parametrisierung $x_{\parallel}^1 = y$ benutzen, in der Umgebung von $x = \pm\sqrt{\frac{1}{2}}$ hingegen x .



An allen anderen Punkten kann man N lokal entweder mit der x -Achse oder mit der y -Achse identifizieren. In $(x, y) = 0$ ist überhaupt keine stetige oder differenzierbare ein-dimensionale Parametrisierung von N möglich.

· Der grosse Nutzen des Satzes liegt genau in der soeben illustrierten Möglichkeit, die Parametrisierung der Lösungsmenge zu variieren. Man nennt eine Teilmenge M von $\mathbb{R}^n$, die sich lokal als Nullstellenmenge einer Funktion $F : U \rightarrow \mathbb{R}^{n-k}$ mit DF von maximalem Rang darstellen lässt, eine k -dimensionale “differenzierbare Untermannigfaltigkeit” von $\mathbb{R}^n$. Über eine von 12.3 gelieferte lokale Parametrisierung von M , d.h. $V_0 \xrightarrow{\cong} M \cap V_0 \times W_0$, (die $x_{\parallel} \mapsto (x_{\parallel}, f(x_{\parallel}))$ heissen auch lokale Koordinaten) lässt

sich die Differentialrechnung von $V_0 \subset \mathbb{R}_{||}^k$ auf M hochheben. Dabei wird $\mathbb{R}_{||}^k$, aufgefasst als Tangentialraum zu V_0 in $x_{||}$, durch $(\text{id}_{k \times k}, Df(x_{||}))$ mit einem Unterraum des Tangentialraums zu $\mathbb{R}^n$ in $(x_{||}, f(x_{||}))$ identifiziert:

$$T_{x_{||}}V_0 \cong \mathbb{R}_{||}^k \ni v \mapsto (v, Df(x_{||})v) \in \mathbb{R}^n \cong T_{(x_{||}, f(x_{||}))}\mathbb{R}^n \quad (12.27)$$

Im obigen Beispiel ist N (wegen der ‘‘Singularität’’ im Ursprung) keine Untermannigfaltigkeit, $N \setminus \{(0, 0)\}$ hingegen ist.

§ 13 Stammfunktionen

Die zweite Klasse von Problemen, über die wir allgemeine Existenzaussagen machen wollen, sind Anfangswertprobleme für gewöhnliche Differentialgleichungen. Sie haben zahlreiche Beispiele schon in der klassischen Mechanik kennengelernt, wir wiederholen noch einmal die grobe Klassifikation am Anfang des nächsten §.

Zur Vorbereitung führen wir in diesem § das gewöhnliche ein-dimensionale Integral ein zur Lösung der einfachsten Sorte *linearer* Differentialgleichungen. Der entscheidende Schritt zur Lösung gewöhnlicher Differentialgleichungen entlang der allgemeinen Strategie von Seite 73 ist nämlich die Umkehrung der linearen Abbildung, welche einer Funktion $f \in \mathcal{C}^1(I, V)$ ihre Ableitung $\dot{f} \in \mathcal{C}^0(I, V)$ aus § 9 zuordnet: Wir hatten schon festgehalten, dass $\mathcal{C}^1(I, V)$ und $\mathcal{C}^0(I, V)$ reelle Vektorräume sind, und die Linearität der Zuordnung $f \mapsto \dot{f}$ ist einfach die Aussage, dass $(f_1 + \lambda f_2) = \dot{f}_1 + \lambda \dot{f}_2 \forall f_1, f_2 \in \mathcal{C}^1(I, V)$ und $\lambda \in \mathbb{R}$.

Für die Abschätzung der ‘‘Verbesserung durch Iteration’’ benötigen wir dann auch geeignete Normen auf $\mathcal{C}^0(I, V)$ und $\mathcal{C}^1(I, V)$. Da diese Vektorräume unendlich-dimensional sind, ist ihre Normierung etwas subtiler als die von $\mathbb{R}^n$, welche wir beim Satz von der Umkehrabbildung benutzt haben. Eigentlich haben wir aber auch schon alle Hilfsmittel im Kapitel 3 vorbereitet. Zunächst einmal benötigen wir, dass für eine stetige Funktion $f : I \rightarrow V$ und jedes kompakte Intervall $[a, b] \subset I$ die Menge $\{\|f(t)\|_V \mid t \in [a, b]\}$ (mit $\|\cdot\|_V$ die Norm auf V) wegen 7.5 beschränkt ist. ($[a, b]$ ist nach dem Satz von Bolzano-Weierstrass 3.11 folgenkompakt und $t \mapsto \|f(t)\|_V \in \mathbb{R}$ ist als Verkettung stetiger Funktionen stetig.) Wir bezeichnen ihr Supremum (in diesem Fall, gleich dem Maximum) mit

$$\|f\|_{[a,b]} := \sup\{\|f(t)\|_V \mid t \in [a, b]\} \quad (13.1)$$

Dann ist die Abbildung $f \mapsto \|f\|_{[a,b]}$ eine Norm auf $\mathcal{C}^0([a, b], V)$.

Definition 13.1. Sei I ein offenes Intervall und $(V, \|\cdot\|)$ ein vollständiger normierter Vektorraum. Eine Funktion $F : I \rightarrow V$ heisst Stammfunktion von $f : I \rightarrow V$, falls F auf I differenzierbar ist und $\dot{F}(t) = f(t) \forall t \in I$.

Bemerkungen. · Eine Stammfunktion ist auf jeden Fall stetig, die gegebene Funktion muss es im Allgemeinen aber nicht sein (vgl. Beispiel (9.5)). Ist jedoch $f \in \mathcal{C}^0(I, V)$, so ist eine Stammfunktion nach dem Schrankensatz 9.3 auf jedem kompakten Teilintervall $K \subset I$ Lipschitz-stetig mit Lipschitz-Konstante $L = \|f\|_K$.

· Sind $F_{1,2}$ Stammfunktionen zu $f_{1,2}$, so ist für $\lambda \in \mathbb{R}$ $F_1 + \lambda F_2$ eine Stammfunktion von $f_1 + \lambda f_2$.

§ 13. STAMMFUNKTIONEN

· Gemäss 9.4 ist jede Stammfunktion der Nullfunktion auf einem Intervall konstant. Sind daher F und $\tilde{F}$ zwei Stammfunktionen von f , so gilt $\tilde{F}(t) = F(t) + v_0 \forall t \in I$ für ein $v_0 \in V$.

· Jede jemals errechnete Ableitung gibt auch ein Beispiel einer Stammfunktion. Jedenfalls können Sie schon Polynome, die Exponential- und trigonometrischen Funktion, sowie gewisse Kombinationen dieser Klassen integrieren. Potenzreihen kann man innerhalb ihres Konvergenzintervalls (Schnitt der Konvergenzscheibe mit der reellen Achse) gliedweise integrieren. Im Allgemeinen gilt:

Theorem 13.2. *Sei $I \subset \mathbb{R}$ ein offenes Intervall und V ein vollständiger normierter Vektorraum. Dann existiert von jeder stetigen Funktion $f : I \rightarrow V$ eine bis auf die Addition einer Konstanten $v_0 \in V$ eindeutige Stammfunktion.*

Die Eindeutigkeit bis auf eine Konstante hatten wir gerade schon festgehalten. Zur Existenz entwickeln wir den Begriff des “bestimmten” Integrals einer stetigen Funktion. Die Intuition dabei ist (*nicht mehr* irgendein “Inhalt der Fläche unter der Kurve”, sondern), dass uns die Ableitung die infinitesimale Veränderung der gesuchten Funktion angibt, wir also nur diese Änderungen über genügend kleine Schritte aufsummieren brauchen.

Proposition/Definition 13.3. *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \rightarrow V$ eine stetige Funktion. Für $a, b \in I$ konvergiert die durch*

$$S_n(f; a, b) := \sum_{k=0}^{n-1} \left(\frac{b-a}{n} \right) \cdot f\left(a + \frac{k}{n}(b-a)\right) \quad (13.2)$$

definierte Folge $(S_n(f; a, b)) \subset V$. Ihr Grenzwert heisst das Integral von f von a nach b . Schreibweise:

$$\int_a^b f(s) ds := \lim_{n \rightarrow \infty} S_n(f; a, b) \quad (13.3)$$

Für $a = b$ ist offensichtlich $\int_a^a f(s) ds = 0$. Sonst aber erlauben wir in dieser Definition sowohl $a < b$ als auch $b < a$. Aus der Tatsache, dass für alle n

$$\begin{aligned} S_n(f; a, b) &= \sum_{k=0}^{n-1} \left(\frac{b-a}{n} \right) \cdot f\left(\frac{n-k}{n}a + \frac{k}{n}b\right) \\ &= \sum_{k=1}^n \left(-\frac{a-b}{n} \right) \cdot f\left(\frac{k}{n}a + \frac{n-k}{n}b\right) \\ &= -S_n(f; b, a) + \frac{b-a}{n} (f(a) - f(b)) \end{aligned} \quad (13.4)$$

folgt dann sofort eine der wichtigen Rechenregeln für das Integral. Zum Nachweis der Konvergenz von $(S_n(f; a, b))$ holen wir für $a < b$ noch ein wenig weiter aus:

· Eine *Zerlegung* $\mathcal{Z}$ eines Intervalls $[a, b] \subset \mathbb{R}$ ist eine geordnet indizierte endliche Teilmenge von $[a, b]$ mit $a, b \in \mathcal{Z}$, d.h. $\mathcal{Z} = \{a = s_0 < s_1 < \dots < s_r = b\}$ für ein $r \in \mathbb{N}$.

· Die *Feinheit* einer Zerlegung $\mathcal{Z}$ ist die positive Zahl

$$\mu(\mathcal{Z}) := \max\{s_{i+1} - s_i \mid i = 0, \dots, r-1\} \quad (13.5)$$

· Die *Riemannsche Summe*¹⁹ von f zur Zerlegung $\mathcal{Z}$ ist

$$S(f; \mathcal{Z}) := \sum_{i=0}^{r-1} (s_{i+1} - s_i) f(s_i) \in V \quad (13.6)$$

Offensichtlich ist (für $a < b$) $S_n(f; a, b)$ aus (13.2) die Riemannsche Summe zur äquidistanten Zerlegung $\{s_k = \frac{n-k}{n}a + \frac{k}{n}b \mid k = 0, \dots, n\}$ der Feinheit $\frac{b-a}{n}$.

· Man sagt, eine Zerlegung $\tilde{\mathcal{Z}}$ ist eine *Verfeinerung* der Zerlegung $\mathcal{Z}$ (Schreibweise: $\tilde{\mathcal{Z}} \triangleleft \mathcal{Z}$), falls $\mathcal{Z} \subset \tilde{\mathcal{Z}}$. Klarerweise gilt in so einem Fall $\mu(\tilde{\mathcal{Z}}) \leq \mu(\mathcal{Z})$.

· Für zwei Zerlegungen $\mathcal{Z}_1$ und $\mathcal{Z}_2$ ist die *gemeinsame Verfeinerung* $\mathcal{Z}_1 \vee \mathcal{Z}_2$ die Vereinigung $\mathcal{Z}_1 \cup \mathcal{Z}_2$ nach einer geeigneten Umindizierung. Es gilt $\mathcal{Z}_1 \vee \mathcal{Z}_2 \triangleleft \mathcal{Z}_1$ und $\mathcal{Z}_1 \vee \mathcal{Z}_2 \triangleleft \mathcal{Z}_2$.

Lemma 13.4. Sei nun $f : [a, b] \rightarrow V$ stetig. Dann existiert für jedes $\epsilon > 0$ ein $\delta > 0$ so, dass für zwei Zerlegungen $\tilde{\mathcal{Z}} \triangleleft \mathcal{Z}$ mit $\mu(\mathcal{Z}) < \delta$

$$\|S(f; \tilde{\mathcal{Z}}) - S(f; \mathcal{Z})\| < \epsilon \quad (13.7)$$

Beweis. Wir nutzen aus, dass gemäss 8.4 f als stetige Funktion auf dem folgenkompakten Intervall $[a, b]$ gleichmässig stetig ist. Es existiert also zu gegebenem $\epsilon > 0$ ein $\delta > 0$ so, dass

$$\|f(t_1) - f(t_2)\| < \frac{\epsilon}{b-a} \quad \forall t_1, t_2 \in [a, b] \text{ mit } |t_1 - t_2| < \delta \quad (13.8)$$

Sei nun $\mathcal{Z} = \{a = s_0 < s_1 < \dots < s_r = b\}$ eine Zerlegung von $[a, b]$ mit $\mu(\mathcal{Z}) < \delta$ und $\tilde{\mathcal{Z}} = \{a = \tilde{s}_0 < \tilde{s}_1 < \dots < \tilde{s}_{\tilde{r}} = b\} \triangleleft \mathcal{Z}$ eine Verfeinerung von $\mathcal{Z}$. Dann gehört zu jedem $\tilde{s}_j \in \tilde{\mathcal{Z}}$ ein eindeutiges $s_{i(j)} \in \mathcal{Z}$ mit $s_{i(j)} \leq \tilde{s}_j < s_{i(j)+1}$, und man überlegt sich (am besten mit einer Skizze), dass

$$(i) \quad s_{i+1} - s_i = \sum_{j \mid i(j)=i} (\tilde{s}_{j+1} - \tilde{s}_j) \quad \forall i = 0, \dots, r-1 \quad (13.9)$$

$$(ii) \quad |\tilde{s}_j - s_{i(j)}| < \delta \quad \forall j = 0, 1, \dots, \tilde{r} \quad (13.10)$$

Daraus folgt

$$\begin{aligned} \|S(f; \tilde{\mathcal{Z}}) - S(f; \mathcal{Z})\| &= \left\| \sum_{j=0}^{\tilde{r}-1} (\tilde{s}_{j+1} - \tilde{s}_j) f(\tilde{s}_j) - \sum_{i=0}^{r-1} (s_{i+1} - s_i) f(s_i) \right\| \\ &\quad (\text{mit (i)}) = \left\| \sum_{j=0}^{\tilde{r}-1} (\tilde{s}_{j+1} - \tilde{s}_j) (f(\tilde{s}_j) - f(s_{i(j)})) \right\| \\ &\quad (\text{wegen (ii) und (13.8)}) < \sum_{j=0}^{\tilde{r}-1} (\tilde{s}_{j+1} - \tilde{s}_j) \frac{\epsilon}{b-a} = \epsilon \end{aligned} \quad (13.11)$$

□

¹⁹Dies ist ein Spezialfall der Riemannschen Summen. Im Allgemeinen wertet man f an beliebigen “Stützstellen” $\tau_i \in [s_i, s_{i+1}]$ aus.

§ 13. STAMMFUNKTIONEN

Korollar 13.5. Für jedes $\epsilon > 0$ existiert ein $\delta > 0$ so, dass für je zwei Zerlegungen $\mathcal{Z}_1$ und $\mathcal{Z}_2$ mit $\mu(\mathcal{Z}_1) < \delta$ und $\mu(\mathcal{Z}_2) < \delta$

$$\|S(f; \mathcal{Z}_1) - S(f; \mathcal{Z}_2)\| < \epsilon \quad (13.12)$$

Beweis. Sei $\delta > 0$ nach 13.4 so, dass $\|S(f; \tilde{\mathcal{Z}}) - S(f; \mathcal{Z})\| < \frac{\epsilon}{2}$ für $\tilde{\mathcal{Z}} \triangleleft \mathcal{Z}$ und $\mu(\mathcal{Z}) < \delta$. Dann gilt für $\mu(\mathcal{Z}_{1,2}) < \delta$:

$$\begin{aligned} \|S(f; \mathcal{Z}_1) - S(f; \mathcal{Z}_2)\| \\ \leq \|S(f; \mathcal{Z}_1) - S(f; \mathcal{Z}_1 \vee \mathcal{Z}_2)\| + \|S(f; \mathcal{Z}_1 \vee \mathcal{Z}_2) - S(f; \mathcal{Z}_2)\| < \epsilon \end{aligned} \quad (13.13)$$

□

Beweis von 13.3. Für $a < b$ folgt aus den obigen Betrachtungen, dass $S_n(f; a, b)$ eine Cauchy-Folge ist (s. Def. 4.2): Wähle für $\epsilon > 0$ mit $\delta > 0$ gemäss 13.5 N_ϵ so, dass $\frac{b-a}{N_\epsilon} < \delta$. Dann folgt $\|S_n(f; a, b) - S_m(f; a, b)\| < \epsilon$ für alle $n, m \geq N_\epsilon$. Wegen der Vollständigkeit von V existiert also der Grenzwert $\lim_{n \rightarrow \infty} S_n(f; a, b)$.

• Für $b < a$ folgt die Behauptung dann aus (13.4). □

Lemma 13.6. (Eigenschaften des Integrals).

(i) Orientierungswechsel: $\int_a^b f(s)ds = - \int_b^a f(s)ds$

(ii) Linearität und Integration von Konstanten: Für $f_1, f_2 \in \mathcal{C}^0(I, V)$ und $\lambda \in \mathbb{R}$, $v \in V$ gilt

$$\int_a^b (f_1 + \lambda f_2 + v)(s)ds = \int_a^b f_1(s)ds + \lambda \int_a^b f_2(s)ds + (b-a)v$$

(iii) Standardabschätzungen ($a < b$): $\left\| \int_a^b f(s)ds \right\| \leq \int_a^b \|f(s)\|ds \leq (b-a) \cdot \|f\|_{[a,b]}$

(iv) Intervalladditivität: Für $a, b, c \in I$ gilt

$$\int_a^b f(s)ds + \int_b^c f(s)ds = \int_a^c f(s)ds$$

Beweis. (i) folgt aus (13.4), (ii) und (iii) aus den entsprechenden Eigenschaften der endlichen Summen (13.2).

(iv) Durch Ausnutzen von (i) können wir die Aussage auf den Fall $a < b < c$ zurückführen. Für $\epsilon > 0$ sei $\delta > 0$ gemäss 13.5 so, dass $\|S(f; \mathcal{Z}_1) - S(f; \mathcal{Z}_2)\| < \epsilon$ für je zwei Zerlegungen von $[a, c]$ mit $\mu(\mathcal{Z}_{1,2}) < \delta$. Sei dann N so, dass $S_n(f; a, b)$, $S_n(f; b, c)$ und $S_n(f; a, c)$ für $n \geq N$ in der Norm weniger als ϵ von den entsprechenden Integralen abweichen, sowie $\frac{c-a}{N} < \delta$. Dann sind $S_n(f; a, b) + S_n(f; b, c)$ und $S_n(f; a, c)$ Riemannsche Summen zu Zerlegungen von $[a, c]$ der Feinheit kleiner als δ . Es folgt für $n \geq N$:

$$\begin{aligned} \left\| \int_a^b f(s)ds + \int_b^c f(s)ds - \int_a^c f(s)ds \right\| &\leq \left\| \int_a^b f(s)ds - S_n(f; a, b) \right\| \\ &+ \left\| \int_b^c f(s)ds - S_n(f; b, c) \right\| + \left\| \int_a^c f(s)ds - S_n(f; a, c) \right\| \\ &+ \|S_n(f; a, b) + S_n(f; b, c) - S_n(f; a, c)\| \\ &< 4\epsilon \end{aligned} \quad (13.14)$$

□

Beweis von 13.2 (Hauptsatz der Differential- und Integralrechnung). Wähle $t_* \in I$ beliebig und setze für $t \in I$

$$F(t) := \int_{t_*}^t f(s) ds \quad (13.15)$$

Beh.: für alle $t_0 \in I$ ist $\dot{F}(t_0) = f(t_0)$.

Bew.: Für $\epsilon > 0$ sei $\delta > 0$ so, dass $\|f(s) - f(t_0)\| < \epsilon$ falls $|s - t_0| < \delta$ (Stetigkeit von f). Dann folgt unter Ausnutzung von 13.6, (ii), (i), (iv), und (iii) für alle $|t - t_0| < \delta$

$$\begin{aligned} \|F(t) - F(t_0) - (t - t_0)f(t_0)\| &= \left\| \int_{t_*}^t f(s) ds - \int_{t_*}^{t_0} f(s) ds - \int_{t_0}^t f(t_0) ds \right\| \\ &= \left\| \int_{t_0}^t (f(s) - f(t_0)) ds \right\| \\ &\leq |t - t_0| \cdot \epsilon \end{aligned} \quad (13.16)$$

Die Behauptung folgt in der Formulierung (9.7). □

· Die äquidistanten Riemannschen Summen (13.2) können zur approximativen Berechnung von Integralen benutzt werden (besser ist die Anpassung der Feinheit wie in (13.6)), die Auswertung des Limes (13.3) gelingt aber nur selten. Üblicherweise “rät” man eine Stammfunktion oder schaut sie in einer vorsorglich angelegten Tabelle nach, oder wendet einen der bekannten Rechentricks an. Die folgenden zwei folgen als Umkehrung von Produkt- bzw. Kettenregel:

Partielle Integration: Ist auf V eine Multiplikation definiert, $F \in \mathcal{C}^1(I, V)$ eine Stammfunktion von $f \in \mathcal{C}^0(I, V)$, und $G \in \mathcal{C}^1(I, V)$, so ist

$$t \mapsto F(t) \cdot G(t) - \int_{t_*}^t F(s) \cdot \dot{G}(s) ds \quad (13.17)$$

eine Stammfunktion von $f \cdot G$.

Substitutionsregel: Ist $F \in \mathcal{C}^1(I, V)$ eine Stammfunktion von $f \in \mathcal{C}^0(I, V)$ und $\tau : J \rightarrow I$ stetig differenzierbar, dann ist $F \circ \tau$ eine Stammfunktion von $(f \circ \tau)\tau'$. Der Umgang dieser Regel mit den Integrationsgrenzen erfordert etwas Konzentration: Für $a, b \in I$ und $\alpha, \beta \in J$ mit $a = \tau(\alpha)$, $b = \tau(\beta)$ gilt:

$$\int_a^b f(s) ds = \int_\alpha^\beta f(\tau(\sigma))\tau'(\sigma) d\sigma \quad (13.18)$$

“Man ersetzt $s = \tau(\sigma)$ und $ds = \frac{ds}{d\sigma} d\sigma = \tau'(\sigma) d\sigma$ und findet Stellen $\alpha, \beta \in J$ so, dass s von a nach b läuft, während σ von α nach β läuft.” Interessant ist, dass in (13.18) τ nicht injektiv sein muss, und noch nicht einmal $\tau([\alpha, \beta]) \subset [a, b]$ gelten muss. Die “überschüssigen” Teile heben sich wegen (13.6) gegenseitig auf. Beispiel: $s = \sigma^2$ in

$$\int_1^2 s ds = \left. \frac{1}{2} t^2 \right|_1^2 = \frac{3}{2} = \int_{-1}^{\sqrt{2}} \sigma^2 \cdot 2\sigma \cdot d\sigma \quad (13.19)$$

§ 13. STAMMFUNKTIONEN

Allerdings ist es nicht möglich, die Anordnung der Zerlegungsstellen in den Riemannschen Summen (13.6) aufzugeben: Betrachte etwa $f(t) = t$ auf dem Intervall $[0, 1]$ und für $n \in \mathbb{N}$ die “Zitterzerlegung” (*kein* sanktionierter Begriff)

$$(s_i) = \left(0, \underbrace{\frac{1}{n}, \frac{2}{n}, \dots, 0, \frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \frac{4}{n}, \dots, 1}_{n^2 \text{ Mal } 3} \right) \quad (13.20)$$

(mit $r = 3n^2 + n - 2$ Elementen). Es gilt $\lim_{n \rightarrow \infty} \max\{|s_{i+1} - s_i| \mid i = 0, \dots, r-1\} = 0$ aber

$$\begin{aligned} \sum_i (s_{i+1} - s_i) f(s_i) &= (n^2 - 1) \cdot \left(\frac{1}{n} \cdot \frac{1}{n} - \frac{2}{n} \cdot \frac{2}{n} \right) + \sum_{k=0}^{n-1} \frac{1}{n} \cdot \frac{k}{n} \\ &= \frac{-3(n^2 - 1)}{n^2} + \frac{1}{n^2} \cdot \frac{n(n-1)}{2} \xrightarrow{n \rightarrow \infty} -3 + \frac{1}{2} \neq \int_0^1 s ds \end{aligned} \quad (13.21)$$

Das Problem mit den Zerlegungen (13.20) ist, dass ihre absolute Länge, nämlich $\sum |s_{i+1} - s_i| = (n^2 - 1) \frac{4}{n} + 1 \rightarrow \infty$ nicht beschränkt bleibt, sie in gewissem Sinne also gar nicht “von a nach b kommen”.

Parameterabhängige Integrale

Ein weiterer häufig benutzter (und auch oft übersehener) Trick bei der Auswertung bestimmter Integrale ohne Auffinden einer Stammfunktion ist die Untersuchung der Abhängigkeit von zusätzlichen Parametern im Integranden. Grundlage dafür ist das folgende Resultat.

Proposition 13.7. *Sei $[a, b] \subset \mathbb{R}$, $U \subset \mathbb{R}^n$ offen und $F : [a, b] \times U \rightarrow V$ stetig. Dann ist die Funktion*

$$G : U \rightarrow V, \quad G(x) := \int_a^b F(s, x) ds \quad (13.22)$$

stetig. Ist V endlich-dimensional, für jedes $t \in [a, b]$ die Abbildung $x \mapsto F(t, x)$ auf U differenzierbar in Richtung $v \in \mathbb{R}^n$, und die Abbildung $[a, b] \times U \ni (t, x) \mapsto D_v F(t, x) \in V$ stetig, dann ist G auf U ebenfalls differenzierbar in Richtung v mit Richtungsableitung

$$D_v G(x) = \int_a^b D_v F(s, x) ds \quad (13.23)$$

Ausserdem übertragen sich diese Aussagen sinngemäss auf (totale) Differenzierbarkeit sowie auf Holomorphie. Die Aussage zur Stetigkeit bleibt gültig, wenn man U durch einen beliebigen metrischen Raum ersetzt.

Beweis. · Die erste Voraussetzung besagt, dass F als Funktion der beiden Variablen gemeinsam stetig ist. Insbesondere ist aber für jedes $x \in U$ die Abbildung $t \mapsto F(t, x)$ stetig, und daher G wohldefiniert, ebenso wie die rechte Seite in (13.23).

· Für $x_0 \in U$ folgt als Korollar von 8.4 aus der Folgenkompaktheit von $[a, b]$, dass für jedes $\epsilon > 0$ ein $\delta > 0$ existiert so, dass $\|F(t, x) - F(t, x_0)\| < \frac{\epsilon}{b-a}$ für alle $t \in [a, b]$

falls nur $\|x - x_0\| < \delta$.²⁰ Daraus folgt mit 13.6, dass für $\|x - x_0\| < \delta$

$$\begin{aligned} \|G(x) - G(x_0)\| &= \left\| \int_a^b (F(s, x) - F(s, x_0)) ds \right\| \\ &\leq (b - a) \|F(\cdot, x) - F(\cdot, x_0)\|_{[a, b]} < \epsilon \end{aligned} \quad (13.24)$$

Dies bedeutet, dass G in x_0 stetig ist.

· Mit Hilfe des Mittelwertsatzes 9.7 (getrennt angewendet auf jede der endlich vielen Komponenten von F) und der Stetigkeit von $D_v F$ folgert man recht analog, dass ein $\delta > 0$ existiert so, dass

$$\|F(t, x_0 + \tau v) - F(t, x_0) - D_v F(t, x_0) \tau\| < \frac{\epsilon}{b - a} \cdot |\tau| \quad (13.25)$$

für alle $t \in [a, b]$ falls nur $|\tau| < \delta$. Daraus folgt durch eine Abschätzung wie in (13.24) die Behauptung zur Differenzierbarkeit von G . $\square$

· Als Beispiel behandeln wir das Integral

$$G = \int_0^1 \frac{s^2 - 1}{\ln s} ds \quad (13.26)$$

(Der Integrand ist an den Grenzen zunächst nicht definiert, lässt sich stetig fortsetzen. Siehe auch uneigentliche Integrale weiter unten.) Wir identifizieren $G = G(2)$, wobei

$$G(x) = \int_0^1 \frac{s^x - 1}{\ln s} ds \quad (13.27)$$

und der Integrand für $x \geq 0$ die Voraussetzungen von 13.7 erfüllt. (Hier könnte man $G(x)$ als uneigentliches Integral auf $x > -1$ fortsetzen.) Es gilt dann

$$\begin{aligned} G'(x) &= \int_0^1 \frac{s^x \ln s}{\ln s} ds \\ &= \int_0^1 s^x ds = \frac{s^{x+1}}{x+1} \Big|_0^1 = \frac{1}{x+1} \end{aligned} \quad (13.28)$$

Wegen $G(0) = 0$ folgt daraus durch Integration

$$G(x) = \ln(x + 1) \quad (13.29)$$

und daher $G = G(2) = \ln 3$.

²⁰In unserer Herangehensweise reicht es für diese Aussage nicht vorauszusetzen, dass $F(t, \cdot)$ für alle t in x_0 stetig ist, sondern es muss tatsächlich F als Funktion von (t, x) in einer Umgebung von $[a, b] \times \{x_0\}$ stetig sein. Ebenso reicht es für (13.25) nicht aus vorauszusetzen, dass $F(t, \cdot)$ für alle t in x_0 differenzierbar ist.

§ 13. STAMMFUNKTIONEN

Uneigentliche Integrale

Ist $I = (\alpha, \beta)$ ein offenes, evtl. auch unbeschränktes Intervall (d.h. $\alpha, \beta \in \mathbb{R} \cup \{\pm\infty\}$), so nennt man eine stetige Funktion $f : (\alpha, \beta) \rightarrow \mathbb{R}$ ("uneigentlich") integrierbar über (α, β) , falls für ein $t_0 \in I$ die rechts-/linksseitigen Limites

$$\lim_{a \downarrow \alpha} \int_a^{t_0} f(s) ds \quad \text{und} \quad \lim_{b \uparrow \beta} \int_{t_0}^b f(s) ds \quad (13.30)$$

beide und getrennt existieren. Man schreibt in diesem Fall weiter

$$\int_{\alpha}^{\beta} f(s) ds = \lim_{a \downarrow \alpha} \int_a^{t_0} f(s) ds + \lim_{b \uparrow \beta} \int_{t_0}^b f(s) ds \quad (13.31)$$

denn das Resultat ist unabhängig von t_0 und stimmt mit der alten Definition überein, falls f in α und/oder β stetig fortsetzbar ist.

- Es existieren Kriterien, die die Existenz solcher uneigentlicher Integrale sicherstellen und gegebenenfalls die Differentiation unter dem Integralzeichen rechtfertigen. (Insbesondere: Majorantenkriterium) Beispiele:
- Die Γ -Funktion ist für $x > 0$ definiert durch

$$\Gamma(x) := \int_0^{\infty} s^{x-1} e^{-s} ds \quad (13.32)$$

und stellt dort eine differenzierbare Funktion dar. Definiert man für $z \in \mathbb{C}$, $s > 0$ $s^z = \exp(z \ln s)$, so gibt für $\operatorname{Re}(z) > 0$

$$\Gamma(z) := \int_0^{\infty} s^{z-1} e^{-s} ds \quad (13.33)$$

eine holomorphe Fortsetzung der Γ -Funktion auf die rechte Halbebene.

- Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion, von der bekannt ist: (i) f hat endlich viele Nullstellen $x_1, \dots, x_r$ mit $f'(x_i) \neq 0$ an jeder dieser Stellen, und (ii) $f(x) \geq O(x)$ für $x \rightarrow \infty$ (d.h. für x gross genug ist $|f(x)| \geq C \cdot |x|$ für eine geeignete Konstante $C > 0$).

Beh.: Dann existiert

$$G = \int_{-\infty}^{\infty} f'(x) e^{-f(x)^2} dx \quad (13.34)$$

und es gilt:

$$G = \sum_i \operatorname{sgn} f'(x_i) \cdot \int_{-\infty}^{\infty} e^{-x^2} dx = \sum_i \operatorname{sgn} f'(x_i) \cdot \sqrt{\pi} \quad (13.35)$$

Bew.: (Skizze) Für $a > 0$ erfüllt der Integrand von

$$G(a) = \int_{-\infty}^{\infty} a f'(x) e^{-a^2 f(x)^2} dx \quad (13.36)$$

die Voraussetzungen zur Differentiation unter dem Integral und es gilt

$$\begin{aligned} G'(a) &= \frac{dG}{da}(a) = \int_{-\infty}^{\infty} (f'(x) - 2a^2 f'(x) f(x)^2) e^{-a^2 f(x)^2} dx \\ &= \int_{-\infty}^{\infty} \frac{d}{dx} (f(x) e^{-a^2 f(x)^2}) dx \\ &= 0 \end{aligned} \quad (13.37)$$

$G(a)$ ist also konstant gleich $G(1) = G$. Andererseits zeigt man unter den gegebenen Voraussetzungen, dass für jedes $\epsilon > 0$

$$\left| G(a) - \sum_{i=1}^r \int_{-\infty}^{\infty} a f'(x_i) e^{-a^2 f'(x_i)^2 (x-x_i)^2} dx \right| < \epsilon \quad (13.38)$$

für a gross genug. (Der Punkt ist, dass der Integrand von $G(a)$ ausserhalb von immer kleiner werdenden Umgebungen der x_i gegen 0 strebt.) Mit der Substitution $x \rightarrow x_i + \frac{x}{a|f'(x_i)|}$ folgt

$$\int_{-\infty}^{\infty} a f'(x_i) e^{-a^2 f'(x_i)^2 (x-x_i)^2} dx = \frac{f'(x_i)}{|f'(x_i)|} \cdot \int_{-\infty}^{\infty} e^{-x^2} dx \quad (13.39)$$

und daraus die Behauptung. Der Beweis von (13.38) wird nachgeliefert. Für die Berechnung des Gaussischen Integrals, siehe § 17.

§ 14 Gewöhnliche Differentialgleichungen

Problemstellung: · Gegeben sind:

- $\mathbb{R}^N$ ein endlich-dimensionaler Vektorraum mit Norm $\|\cdot\|$.
- Ein (Zeit-)Intervall $I \subset \mathbb{R}$
- Eine stetige Funktion (Vektorfeld) $Y : I \times U \rightarrow \mathbb{R}^N$ auf dem Produkt von I mit einer offenen Menge $U \subset \mathbb{R}^N$.
- Eine Startzeit $t_0 \in I$ und ein Startwert $x_0 \in U$
- Gesucht sind: ein Intervall $[t_0, t_1] \subset I$ und eine stetig differenzierbare Funktion $x : [t_0, t_1] \rightarrow U$, für die gilt:

$$\dot{x}(t) = Y(t, x(t)) \quad \forall t \in [t_0, t_1] \quad \text{und} \quad x(t_0) = x_0 \quad (14.1)$$

Man nennt (14.1) ein explizites²¹ Anfangswertproblem (AWP) erster Ordnung, die Funktion x eine Lösung des AWP.

Fragen: · Ist das AWP (14.1) überhaupt lösbar?

- Ist die Lösung eindeutig, zumindest in dem Sinne, dass falls $\tilde{x} : [t_0, \tilde{t}_1] \rightarrow U$ eine zweite Lösung ist, $x(t) = \tilde{x}(t) \quad \forall t \in [t_0, \min(t_1, \tilde{t}_1)]$?
- Wie weit lässt sich die Lösung maximal fortsetzen? Läuft sie etwa vor dem Ende der Zeit aus U heraus?

²¹Dies bedeutet, dass die Gleichung nach $\dot{x}(t)$ aufgelöst gegeben ist. Bei impliziten Differentialgleichungen der Form $F(\dot{x}(t), x(t)) = 0$ kann man mal zunächst den Satz über implizite Funktionen 12.3 anrufen.

§ 14. GEWÖHNLICHE DIFFERENTIALGLEICHUNGEN

· Hängt die Lösung in stetiger und/oder differenzierbarer Weise von den Anfangswerten, sowie gegebenenfalls noch von weiteren “externen” Parametern ab? Mit anderen Worten, existiert mit einer offenen Umgebung V von x_0 eine stetige Funktion $\Phi : [t_0, t_1] \times V \rightarrow U$ mit

$$\frac{\partial \Phi}{\partial t}(t, x) = Y(t, \Phi(t, x)), \quad \Phi(t_0, x) = x \quad (14.2)$$

Ist Φ differenzierbar, falls Y dies ist?

Beispiel 14.1. · Siehe theoretische Physik-Vorlesungen!

· Tatsächlich ist Stetigkeit des Vektorfeldes hinreichend für die lokale Existenz von Lösungen. Wir werden diese Aussage nicht beweisen, da Stetigkeit für die Eindeutigkeit im Allgemeinen nicht ausreicht. Als Gegenbeispiel dient hier üblicherweise das Anfangswertproblem mit $I = \mathbb{R}$, $N = 1$, $Y(x) = \sqrt{|x|}$, $t_0 = 0$, $x_0 = 0$, oder kürzer

$$\dot{x} = \sqrt{|x|}, \quad x(0) = 0 \quad (14.3)$$

Offensichtlich ist $x(t) = 0 \forall t$ eine Lösung, aber auch für jedes $t_* \geq 0$:

$$x(t) = \begin{cases} 0 & t \leq t_* \\ \frac{(t - t_*)^2}{4} & \text{für } t > t_* \end{cases} \quad (14.4)$$

Die fehlende Eindeutigkeit hat damit zu tun, dass Y zwar stetig, aber nicht Lipschitzstetig ist, siehe Voraussetzung in 14.2. unten).

· Ein Standardbeispiel für endliche Existenz ist das AWP

$$\dot{x} = x^2, \quad x(0) = x_0 > 0 \quad (14.5)$$

mit Lösung

$$x(t) = \frac{1}{x_0^{-1} - t} \quad t \in [0, x_0^{-1}) \quad (14.6)$$

· Das AWP (14.5) ist vom Typ der “getrennten Veränderlichen”. Allgemein hat (hoffentlich bekanntlich) für zwei offene Intervalle $I, U \subset \mathbb{R}$, eine stetige Funktion $f : I \rightarrow \mathbb{R}$ und eine stetige Funktion $g : U \rightarrow \mathbb{R}$ mit $g(x) \neq 0 \forall x \in U$ das AWP

$$\dot{x}(t) = f(t)g(x(t)), \quad x(t_0) = x_0 \quad (14.7)$$

($t_0 \in I$, $x_0 \in U$) die Lösung

$$x(t) = H^{-1}(F(t)) \quad (14.8)$$

wobei $F = \int_{t_0}^t f(s)ds$ die Stammfunktion von f mit $F(t_0) = 0$ ist und H^{-1} die Umkehrfunktion der Stammfunktion

$$H(x) = \int_{x_0}^x \frac{1}{g(y)} dy \quad (14.9)$$

von $1/g$ mit $H(x_0) = 0$: H ist wegen $\frac{1}{g} \neq 0$ streng monoton und daher nach 7.2 global invertierbar. Die Lösung (14.8) ist definiert auf dem Teilintervall J von

$$\{t \in I \mid F(t) \in H(U)\} = F^{-1}(H(U)) \quad (14.10)$$

welches t_0 enthält. ($F^{-1}(H(U))$ ist offen, da $H(U)$ als monotones Bild eines offenen Intervalls offen ist, und F stetig ist.)

· Ebenfalls bekannt sein sollte die Rückführung von Differentialgleichungen höherer Ordnung auf die Form (14.1). So ist etwa für $F : \mathbb{R}^k \rightarrow \mathbb{R}$ die gewöhnliche skalare Differentialgleichung

$$y^{(k)}(t) = F(y(t), \dot{y}(t), \dots, y^{(k-1)}(t)) \quad (14.11)$$

äquivalent zu einer vektoriellen Gleichung für $x = (y, \dot{y}, \dots, y^{(k-1)})^T \in \mathbb{R}^k$:

$$\frac{d}{dt} \begin{pmatrix} y(t) \\ \dot{y}(t) \\ \vdots \\ y^{(k-1)}(t) \end{pmatrix} = \begin{pmatrix} \dot{y}(t) \\ \ddot{y}(t) \\ \vdots \\ F(y(t), \dots, y^{(k-1)}(t)) \end{pmatrix} \quad (14.12)$$

nämlich, $\dot{x} = Y(x(t))$, wobei $Y(x^1, \dots, x^k) = (x^2, \dots, F(x^1, \dots, x^k))^T$. (Anfangswerte und explizite Zeitabhängigkeit wurden unterdrückt.)

· Eine andere “Reduktion der Ordnung”, welche in der Physik eine grosse Rolle spielt, ergibt sich aus Erhaltungssätzen. Beispielsweise ist auf Lösungen von

$$\ddot{x}(t) = -\text{grad } V(x(t)) \quad (14.13)$$

die Energie $E = \frac{1}{2} \langle \dot{x}, \dot{x} \rangle + V(x)$ konstant. (Wobei grad gemäss (10.31) bezüglich eines euklidischen inneren Produkts $\langle \cdot, \cdot \rangle$ auf $\mathbb{R}^N$ definiert ist, und $V \in C^1(\mathbb{R}^N, \mathbb{R})$, hier also kein Vektorraum.) Im Falle $N = 1$ erhält man dann durch Auflösen nach $\dot{x}$ eine Gleichung mit getrennten Veränderlichen und Lösungen der Form

$$t - t_0 = \int_{x_0}^x \frac{1}{\sqrt{2(E - V(x))}} dx \quad (14.14)$$

(Konvergenz solcher Integrale als Übungsaufgabe.)

· Für eine (fixierte) lineare Abbildung $A : \mathbb{R}^N \rightarrow \mathbb{R}^N$ hat für jedes $x_0 \in \mathbb{R}^N$ das AWP

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0 \quad (14.15)$$

die Lösung

$$x(t) = \exp(tA)x_0 \quad (14.16)$$

Beispiel: Schwingungsgleichung. Zurück zur Theorie.

Theorem 14.2 (Lokale Existenz und Eindeutigkeit). *Seit $I \subset \mathbb{R}$ ein offenes Intervall, $U \subset \mathbb{R}^N$ offen und $Y : I \times U \rightarrow \mathbb{R}^N$ stetig. Angenommen, es existiert eine Lipschitz-Konstante $L > 0$ so, dass für alle $t \in I$ und alle $x_1, x_2 \in U$*

$$\|Y(t, x_1) - Y(t, x_2)\| \leq L \cdot \|x_1 - x_2\| \quad (14.17)$$

Dann existiert für alle $t_0 \in I$, $x_0 \in U$ ein $\delta > 0$ mit $t_0 + \delta \in I$ und so, dass $\forall t_1 \in (t_0, t_0 + \delta]$ eine eindeutige Lösung $x : [t_0, t_1] \rightarrow U$ des AWP

$$\dot{x}(t) = Y(t, x(t)), \quad x(t_0) = x_0 \quad (14.18)$$

existiert.

§ 14. GEWÖHNLICHE DIFFERENTIALGLEICHUNGEN

Beweis. Plan: Wir emulieren den Beweis des Satzes von der Umkehrabbildung 12.1 unter Zuhilfenahme des $\int_a^b$ aus 13.3 als Lösung eines linearen Problems sowie des Banachschen Fixpunktsatzes 4.19 auf einem geeigneten metrischen Raum (von Funktionen $[t_0, t_1] \rightarrow U$).

1. Schritt: (Umschreiben als Fixpunktgleichung) Die Suche nach einer differenzierbaren Funktion $x(\cdot)$ mit

$$\dot{x}(t) = Y(t, x(t)) \quad \text{und} \quad x(t_0) = x_0 \quad (14.19)$$

ist äquivalent zur Suche nach einer stetigen Funktion $x(\cdot)$ mit

$$x(t) = x_0 + \int_{t_0}^t Y(s, x(s)) ds \quad (14.20)$$

(Wegen unseres Hauptsatzes 13.2 ist nämlich $x(\cdot)$ dann schon differenzierbar, wenn sie stetig ist und die Gleichung (14.20) erfüllt.) M.a.W. suchen wir Fixpunkte der Abbildung

$$\begin{aligned} T : \mathcal{C}^0(I, U) &\rightarrow \mathcal{C}^0(I, \mathbb{R}^N) \\ T(x(\cdot))(t) &:= x_0 + \int_{t_0}^t Y(s, x(s)) ds \end{aligned} \quad (14.21)$$

evtl. nach Einschränkung auf geeignete Teilintervalle von I .

2. Schritt: (Vollständige metrische Räume) Für $R > 0$ so, dass $\overline{B_R(x_0)} \subset U$ und $t_1 \in I$, $t_1 > t_0$ betrachten wir

$$X = \{x : [t_0, t_1] \rightarrow \overline{B_R(x_0)} \text{ stetig, } x(t_0) = x_0\} \quad (14.22)$$

mit der von der Supremumsnorm auf $\mathcal{C}^0([t_0, t_1], \mathbb{R}^N)$ induzierten Abstandsfunktion

$$\|x_1(\cdot) - x_2(\cdot)\|_{[t_0, t_1]} := \max\{\|x_1(t) - x_2(t)\|_{\mathbb{R}^N} \mid t \in [t_0, t_1]\} \quad (14.23)$$

(X selbst ist kein Vektorraum!) Gemäss (einer leichten Verallgemeinerung von) 8.2 ist X ein vollständiger metrischer Raum. ($\overline{B_R(x_0)}$ ist als abgeschlossene Teilmenge eines vollständigen normierten Raumes vollständig.) Wir müssen jetzt nur noch unsere beiden Schranken R und δ so wählen, dass falls $t_1 \leq t_0 + \delta$:

(i) $T(X) \subset X$, d.h. T ist eine Selbstabbildung von X .

(ii) T ist kontraktiv, d.h.

$$\|T(x_1(\cdot)) - T(x_2(\cdot))\|_{[t_0, t_1]} \leq c \cdot \|x_1(\cdot) - x_2(\cdot)\|_{[t_0, t_1]} \quad (14.24)$$

für ein $c < 1$. Es stellt sich heraus, dass es genügt, δ anzupassen.

3. Schritt: (Selbstabbildung) Angenommen, $x(\cdot) \in X$, insbesondere $x(t) \in \overline{B_R(x_0)}$ $\forall t \in [t_0, t_1]$. Dann ist

$$\begin{aligned} \|T(x(\cdot))(t) - x_0\| &= \left\| \int_{t_0}^t Y(s, x(s)) ds \right\| \\ &\leq |t - t_0| \cdot \max\{\|Y(s, x(s))\| \mid s \in [t_0, t]\} \end{aligned} \quad (14.25)$$

nach unseren Integralabschätzungen 13.6. Wir fixieren nun $R > 0$ mit $\overline{B_R(x_0)} \subset U$ und $t_* > t_0$ mit $t_* \in I$ und setzen

$$M := \max\{\|Y(t, x)\| \mid t \in [t_0, t_*], x \in \overline{B_R(x_0)}\} < \infty \quad (14.26)$$

denn $[t_0, t_*] \times \overline{B_R(x_0)}$ ist folgenkompakt und stetige Funktionen auf folgenkompakten Mengen sind beschränkt, s. 7.5. Dann ist für $t_1 \leq \max\{t_*, t_0 + \frac{R}{M}\}$ und alle $t \in [t_0, t_1]$ als Folge von (14.25)

$$\|T(x(\cdot))(t) - x_0\| \leq \frac{R}{M} \cdot M = R \quad (14.27)$$

also $T(x) \in X$. (Stetigkeit von $T(x)$ war ja schon klar, ebenso $(t_0) = x_0$.)

4. Schritt (Kontraktivität) Für $x_1(\cdot), x_2(\cdot) \in X$ ist $\forall t \in [t_0, t_1]$

$$\begin{aligned} \|T(x_1(\cdot))(t) - T(x_2(\cdot))(t)\| &= \left\| \int_{t_0}^t (Y(s, x_1(s)) - Y(s, x_2(s))) ds \right\| \\ &\leq |t - t_0| \cdot \max\{\|Y(s, x_1(s)) - Y(s, x_2(s))\| \mid s \in [t_0, t]\} \\ \text{Lipschitz-Bed. (14.17)} &\leq |t - t_0| \cdot L \cdot \max\{\|x_1(s) - x_2(s)\| \mid s \in [t_0, t]\} \\ &\leq \underbrace{|t_1 - t_0| \cdot L}_{=: c} \cdot \|x_1(\cdot) - x_2(\cdot)\|_{[t_0, t_1]} \end{aligned} \quad (14.28)$$

Ist dann t_1 ausserdem noch $< t_0 + \frac{1}{L}$, so folgt $c < 1$, und dann ist T eine Kontraktion.

5. Schritt: (Synthese) Wir sehen: Mit

$$0 < \delta < \max\left\{t_* - t_0, \frac{R}{M}, \frac{1}{L}\right\} \quad (14.29)$$

ist für alle $t_1 \in (t_0, t_0 + \delta]$ T eine Kontraktion. Da die konstante Abbildung $x(t) = x_0$ $\forall t \in [t_0, t_1]$ auf jeden Fall in X liegt, ist X nicht-leer, und wir können mit 4.19 schliessen. $\square$

Beispiel 14.3. Im linearen AWP (14.15) ist $Y(t, x) = Ax$. Ausgehend von der Konstanten $x_0(t) = x_0 \forall t$ erhalten wir durch Iteration von T :

$$\begin{aligned} x_1(t) &= T(x_0(\cdot))(t) = x_0 + \int_0^t Ax_0 ds = x_0 + tAx_0 \\ x_2(t) &= T(x_1(\cdot))(t) = x_0 + \int_0^t A(x_0 + sAx_0) ds = x_0 + tAx_0 + \frac{1}{2}t^2A^2x_0 \\ &\vdots \quad (\text{vollständige Induktion}) \quad \vdots \\ x_n(t) &= \sum_{k=0}^n \frac{t^k}{k!} A^k \end{aligned} \quad (14.30)$$

was für $n \rightarrow \infty$ gegen die Lösung $\exp(tA)x_0$ konvergiert.

KAPITEL 6

FUNKTIONENTHEORIE

Teil unserer Aufgaben im Wintersemester werden die Umkehrung der mehrdimensionalen Differentialrechnung aus § 10 und die Lösung gewisser partieller Differentialgleichungen sein. Dies wird einige Vorbereitungen aus der linearen Algebra erfordern, sowie eine erweiterte Uminterpretation des Integralbegriffs. In diesem Kapitel stellen wir uns zunächst die durchaus einfacher scheinende Frage nach der Umkehrung der komplexen Ableitung aus § 11. Die Antwort ist höchst überraschend. Es stellt sich nämlich heraus, dass für $U \subset \mathbb{C}$ offen eine Funktion $f : U \rightarrow \mathbb{C}$ lokal genau dann eine Stammfunktion besitzt (d.h., komplex integrierbar ist), wenn sie holomorph (nämlich, komplex differenzierbar) ist. Stetigkeit reicht nicht mehr aus! Eine globale Konsequenz aus dieser “Vereinheitlichung” von Differentiation und Integration ist, dass sich bestimmte komplexe Integrale in viel systematischerer Weise als etwa die Beispiele (13.26), (13.35) durch Studium der Ableitungen des Integranden auswerten lassen.

In diesem Kapitel steht damit unser bisheriges Leitziel, dass wir nach Lösungen von physikalischen Problemen durch Approximation suchen, etwas zurück. Vielmehr bietet die Funktionentheorie (in eigentlich unerwarteter Weise) ein Hilfsmittel für *exakte* Aussagen über Strukturen in physikalischen Theorien. Ich hoffe, Sie werden davon in absehbarer Zeit profitieren!

§ 15 Integration in $\mathbb{C}$

Ein weiterer Unterschied: Anstatt wie bisher von Anfang an alles richtig hinzuschreiben, gehen wir in diesem § entdeckend vor und tasten uns schrittweise an die endgültigen Definitionen heran.²²

Zu Beginn müssen wir uns auf eine Klasse von Teilmengen von $\mathbb{C}$ als Definitionsbereiche einigen, auf denen wir nach Stammfunktionen suchen sollen. Ein Blick zurück auf 13.3 zeigt, dass die für die Definition des bestimmten Integrals zentrale Eigenschaft von Intervallen ihre Konvexität ist: Für je zwei Punkte $a, b \in I$ liegen auch alle Zwischenstellen $(1 - t)a + tb$, $t \in [0, 1]$ in I . Eine sinnvolle komplexe Verallgemeinerung sind also offene und konvexe²³ Teilmengen $U \subset \mathbb{C}$, und mit dieser Erkenntnis starten wir sofort unseren ersten Versuch, ein “Integral von a nach b in U ” zu definieren. Der Einfachheit halber beschränken wir uns auf einen (komplex) ein-dimensionalen Wertebereich. Wir nehmen also an, $f : U \rightarrow \mathbb{C}$ sei stetig und

²²Wir sagen natürlich auch nicht absichtlich etwas Falsches!

²³Zur Definition der Konvexität wird $\mathbb{C}$ mit $\mathbb{R}^2$ identifiziert: $U \subset \mathbb{C}$ heisst konvex, falls $\forall a, b \in U$, $\{a + t(b - a) \mid t \in [0, 1]\} \subset U$, s. S. 65.

setzen für $n \in \mathbb{N}$:

$$S_n(f; a, b) := \sum_{k=0}^{n-1} \left(\frac{b-a}{n} \right) \cdot f\left(a + \frac{k}{n}(b-a)\right) \quad (15.1)$$

So wie sich 11.1 von 9.1 nur im Namen der Variablen unterscheidet, ist diese Definition formal identisch zu (13.2). What could possibly go wrong?

Beh.: Es existiert $\lim_{n \rightarrow \infty} S_n(f; a, b)$

Bew.: Wir betrachten auf dem Intervall $[0, 1] \subset \mathbb{R}$ die stetige, komplexwertige Funktion

$$[0, 1] \ni t \mapsto \tilde{f}(t) := f(a + t(b-a)) = f((1-t)a + tb) \in \mathbb{C} \quad (15.2)$$

Dann ist

$$S_n(f; a, b) = (b-a) \cdot S_n(\tilde{f}; 0, 1) \quad (15.3)$$

wo die rechte Seite nun echt mit (13.2) zusammentrifft, da $0, 1 \in \mathbb{R}$. Die Konvergenz folgt also aus der Stetigkeit von $\tilde{f}$ und es gilt

$$\int_a^b f(w)dw := \lim_{n \rightarrow \infty} S_n(f; a, b) = (b-a) \int_0^1 \tilde{f}(t)dt \quad (15.4)$$

□

Beachte: Die Notation auf der linken Seite ist im Komplexen obsolet. In Anbetracht der endgültigen Definition 15.2 sollten wir schreiben

$$\int_{l(a,b)} f(w)dw \quad (15.5)$$

wobei $l(a, b) = (t \mapsto a + t(b-a))$ für die “gerichtete Strecke von a nach b ” steht.

Beispiel: Angenommen, $F : U \rightarrow \mathbb{C}$ ist eine holomorphe Funktion mit $F'(z) = f(z)$ $\forall z \in U$. Dann folgt aus der Kettenregel, dass

$$[0, 1] \ni t \mapsto \tilde{F}(t) := F(a + t(b-a)) \in \mathbb{C} \quad (15.6)$$

(reell) differenzierbar ist mit der Ableitung

$$\dot{\tilde{F}}(t) = \frac{d}{dt} F(a + t(b-a)) = (b-a) \cdot F'(a + t(b-a)) = (b-a) \cdot \tilde{f}(t) \in \mathbb{C} \quad (15.7)$$

$\tilde{F}$ ist also auf $[0, 1]$ eine Stammfunktion von $(b-a)\tilde{f}$ (oder, wenn gewollt, fortgesetzt auf ein offenes Intervall $(-\epsilon, 1+\epsilon)$ für $\epsilon > 0$). Nach dem Hauptsatz der (reellen) Differential- und Integralrechnung folgt daher

$$\int_a^b f(w)dw = \tilde{F}(1) - \tilde{F}(0) = F(b) - F(a) \quad (15.8)$$

Explizitere Beispiele sind Konstanten, Polynome, Potenzreihen, etc.

Können wir nun umgekehrt (15.4) benutzen, um für jede solche stetige Funktion $f : U \rightarrow \mathbb{C}$ eine komplexe Stammfunktion zu konstruieren? NEIN! — In unserem

§ 15. INTEGRATION IN $\mathbb{C}$

Beweis des Hauptsatzes (13.16) haben wir nämlich *alle* Eigenschaften in 13.6 benutzt und man sieht:

(i) (Orientierungswechsel) gilt auch für (15.4) ✓

(ii) (Linearität und Integration von Konstanten) ✓

(iii) (Standardabschätzung) (mit Norm = Absolutbetrag) ✓

ABER (iv) (Additivität im Integrationsbereich) gilt im Allgemeinen nicht! Der Vorfaktor $(b - a)$ in (15.4) verhindert, dass der Beweis von Seite 83 ohne weiteres durchgeht, ein einfaches Gegenbeispiel ist: $f(z) = \bar{z}$, $a = 0$, $b = 1$, $c = 1 + i$.

$$\begin{aligned}\int_0^1 \bar{w} dw &= \int_0^1 t dt = \frac{1}{2} \\ \int_1^{1+i} \bar{w} dw &= i \int_0^1 \overline{(1+it)} dt = i \int_0^1 (1-it) dt = i + \frac{1}{2} \\ \int_0^{1+i} \bar{w} dw &= (1+i) \int_0^1 (1-i)t dt = 2 \cdot \int_0^1 t dt = 1 \\ &\Rightarrow \int_0^1 + \int_1^{1+i} \neq \int_0^{1+i} \text{!!!!}\end{aligned}\tag{15.9}$$

Da aber sonst alle Eigenschaften gelten, und da aus (15.8) umgekehrt folgt, dass die Additivität notwendig ist für die Existenz einer Stammfunktion, können wir folgern:

Fazit: Eine stetige Funktion $f : U \rightarrow \mathbb{C}$ auf einer offenen und konvexen Menge $U \subset \mathbb{C}$ besitzt dann und nur dann eine komplexe Stammfunktion, d.h. eine holomorphe Funktion $F : U \rightarrow \mathbb{C}$ mit $F'(z) = f(z) \forall z \in U$ falls für alle $a, b, c \in U$ gilt:

$$\int_a^b f(w) dw + \int_b^c f(w) dw + \int_c^a f(w) dw = 0 \tag{15.10}$$

Wie in $\mathbb{R}$ ist eine solche Stammfunktion durch das unbestimmte Integral gegeben und bis auf eine additive Konstante eindeutig.

Lemma 15.1 (Lemma von Cauchy-Goursat-Pringsheim). *Sei U wie eben offen und konvex. Dann gilt die Aussage (15.10), falls f auf U holomorph ist!*

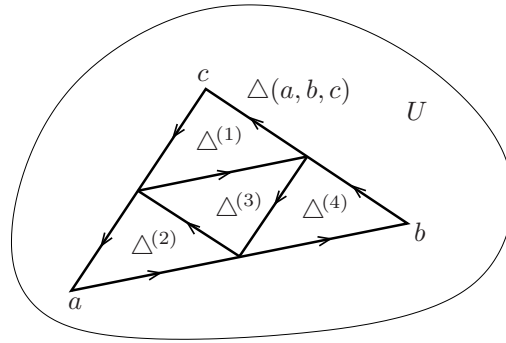
Beweis. Auf der Suche nach einem Ersatz für die gemeinsame Verfeinerung der Intervall-Zerlegungen aus 13.5 benutzen wir die Visualisierung von $\mathbb{C}$ als zweidimensionale euklidische Ebene und daraus abgeleitete geometrische Sprechweisen. Für drei Punkte $o, p, q \in \mathbb{C}$ bezeichne $\blacktriangle(o, p, q)$ das (allenfalls entartete²⁴) Dreieck mit diesen Punkten als Ecken und $\triangle(o, p, q)$ seinen Rand. Beides sind abgeschlossene Teilmengen von $\mathbb{C}$ im Sinne von 4.12. Der Durchmesser ist $\text{diam } \blacktriangle(o, p, q) = \max\{|p - o|, |q - p|, |o - q|\}$ und der Umfang $|\triangle(o, p, q)| = |p - o| + |q - p| + |o - q|$. Für $o, p, q \in U$ ist wegen der Konvexität von U $\blacktriangle(o, p, q) \subset U$. Wir schreiben

$$\int_{\triangle(o, p, q)} f(w) dw = \int_o^p f(w) dw + \int_p^q f(w) dw + \int_q^o f(w) dw \tag{15.11}$$

²⁴Sind a, b, c kollinear oder fallen zusammen, so gilt (15.10) auf jeden Fall.

für das “Integral über der Rand” in der angegebenen Orientierung. (Es gilt $\int_{\Delta(o,p,q)} f(w)dw = \int_{\Delta(p,q,o)} f(w)dw = -\int_{\Delta(o,q,p)} f(w)dw$ gemäss Eigenschaft (i).)

Die Behauptung ist also, dass für je drei Punkte $a, b, c \in U$, $\int_{\Delta(a,b,c)} f(w)dw = 0$, falls f holomorph ist. Wir zeigen dafür, dass für jedes $\epsilon > 0$ $|\int_{\Delta(a,b,c)} f(w)dw| < \epsilon$. Dazu zerteilen wir $\blacktriangle(a, b, c)$ durch Halbieren der Seiten in vier kongruente Dreiecke, deren Ränder $\Delta^{(1)}$, $\Delta^{(2)}$, $\Delta^{(3)}$ und $\Delta^{(4)}$, wir in einer Weise orientieren, dass



$$\int_{\Delta(a,b,c)} f(w)dw = \sum_{i=1}^4 \int_{\Delta^{(i)}} f(w)dw \quad (15.12)$$

Gemäss Behauptung ist jeder der vier Terme auf der rechten Seite ebenfalls null. Ohne dies zu wissen, suchen wir ein jenes Dreieck aus, für das der Wert des Integrals (im Absolutbetrag) das Maximum von allen vieren ist, und bezeichnen es mit Δ_1 . Mit $\Delta_0 := \Delta(a, b, c)$ gilt dann

$$\left| \int_{\Delta(a,b,c)} f(w)dw \right| = \left| \int_{\Delta_0} f(w)dw \right| \leq 4 \cdot \left| \int_{\Delta_1} f(w)dw \right| \quad (15.13)$$

Nun zerteilen wir analog $\blacktriangle_1$, suchen das Maximum der vier Integrale und bezeichnen das zugehörige Dreieck mit $\blacktriangle_2$. Durch Wiederholen des Verfahrens erhalten wir eine Folge $(\blacktriangle_n)$ von Dreiecken mit Rändern (Δ_n) über die

$$\left| \int_{\Delta(a,b,c)} f(w)dw \right| \leq 4^n \cdot \left| \int_{\Delta_n} f(w)dw \right| \quad (15.14)$$

Wegen der Kongruenz gilt $\text{diam } \blacktriangle_n = 2^{-n} \text{diam } \blacktriangle_0$ und $|\Delta_n| = 2^{-n} |\Delta_0|$, bilden also Nullfolgen. Ausserdem ist $\blacktriangle_{n+1} \subset \blacktriangle_n$ und nach 4.16 existiert daher genau ein Punkt $z_0 \in \bigcap_{n=0}^{\infty} \blacktriangle_n$.

· Da f nach Annahme in z_0 komplex differenzierbar ist, existiert ein $\delta > 0$ so, dass

$$|f(z) - f(z_0) - f'(z_0)(z - z_0)| \leq \epsilon \cdot |z - z_0| \quad (15.15)$$

falls $|z - z_0| < \delta$. Sei nun n so gross, dass $\text{diam } \blacktriangle_n < \delta$. Dann ist für $w \in \Delta_n$ $|w - z_0| \leq \text{diam } \blacktriangle_n < \delta$. Wegen der Eigenschaft (ii) des Integrals (15.4) gilt

$$\int_{\Delta_n} f(w)dw = \int_{\Delta_n} (f(w) - f(z_0) - f'(z_0)(w - z_0))dw \quad (15.16)$$

und mit Hilfe von Eigenschaft (iii) und (15.15) folgt

$$\begin{aligned} \left| \int_{\Delta_n} f(w)dw \right| &\leq \sup\{|f(w) - f(z_0) - f'(z_0)(w - z_0)| \mid w \in \Delta_n\} \cdot |\Delta_n| \\ &\leq \epsilon \cdot \text{diam } \blacktriangle_n \cdot |\Delta_n| \\ &\leq \epsilon \cdot \frac{\text{diam } \blacktriangle_0}{2^n} \cdot \frac{|\Delta_0|}{2^n} \end{aligned} \quad (15.17)$$

§ 15. INTEGRATION IN $\mathbb{C}$

Mit (15.14) folgt

$$\left| \int_{\Delta(a,b,c)} f(w)dw \right| \leq \epsilon \cdot \text{diam } \blacktriangle(a,b,c) \cdot |\Delta(a,b,c)| \quad (15.18)$$

und daraus die Behauptung. $\square$

Wir zeigen im nächsten §, dass falls f umgekehrt eine Stammfunktion besitzt, d.h. (15.10) gilt, dann f bereits holomorph ist. Zunächst aber flexibilisieren wir etwas Integral und Definitionsbereiche.

Die Motivation für die folgende Definition ist die bereits gemachte Beobachtung, dass *falls eine Stammfunktion existiert*, es zu ihrer Rekonstruktion nicht darauf ankommt, “wie genau wir von a nach b kommen”, Hauptsache, wir kommen an (vgl. dazu die Warnungen auf Seite 85).

Definition 15.2. Sei $U \subset \mathbb{C}$ offen. Ein stetig differenzierbarer Weg in U von a nach b , beides Punkte in U , ist eine stetig differenzierbare Abbildung $\gamma : [\alpha, \beta] \rightarrow U$ mit $\gamma(\alpha) = a$, $\gamma(\beta) = b$. Für eine stetige Funktion $f : U \rightarrow \mathbb{C}$ definieren wir dann das Integral von f über γ mit 13.3 als

$$\int_{\gamma} f(w)dw := \int_{\alpha}^{\beta} f(\gamma(t)) \cdot \dot{\gamma}(t)dt \quad (15.19)$$

$\uparrow$
 hier ist die Multiplikation
 in $\mathbb{C}$ versteckt!

• Die Länge eines Weges ist das reellwertige Integral

$$l(\gamma) := \int_{\alpha}^{\beta} |\dot{\gamma}(t)|dt \quad (15.20)$$

• Ist $\gamma_1 : [\alpha, \beta] \rightarrow U$ ein Weg von a nach b und $\gamma_2 : [\beta, \gamma] \rightarrow U$ ein Weg von b nach c so definieren wir die Verkettung (oder Aneinanderreihung) von γ_1 und γ_2 durch

$$[\alpha, \gamma] \ni t \mapsto \begin{cases} \gamma_1(t) & t \in [\alpha, \beta] \\ \gamma_2(t) & t \in (\beta, \gamma] \end{cases} \quad (15.21)$$

Dies ist wegen $\gamma_1(\beta) = b = \gamma_2(\beta)$ einer stetiger Weg von a nach c , aber in β nicht unbedingt differenzierbar, was aber zur Not auch repariert werden kann. Eine Verkettung von endlich vielen stetig differenzierbaren Wegen nennt man auch stückweise stetig differenzierbar. Integral und Länge sind als Summen über die Stücke definiert.

• Der *Orientierungswechsel* eines Wegs γ ist $[\alpha, \beta] \ni t \mapsto \gamma(\beta + \alpha - t)$. Unter Orientierungswechsel vertauschen Anfangs- und Endpunkt ihre Rolle und das Integral schnappt sich ein Vorzeichen. Die Länge bleibt sich gleich.

• Eine *Reparametrisierung* eines Wegs $\gamma : [\alpha, \beta] \rightarrow U$ ist der Weg $\tilde{\gamma} = \gamma \circ \tau$ für eine stetig differenzierbare Abbildung $\tau : [\tilde{\alpha}, \tilde{\beta}] \rightarrow [\alpha, \beta]$ mit $\tau(\tilde{\alpha}) = \alpha$, $\tau(\tilde{\beta}) = \beta$. Integral und Länge ändern sich dabei nicht.

²⁵Das Alphabet hat hier kapitulieren müssen, die Bedeutung von γ sollte aber jeweils klar sein. Noch etwas allgemeiner muss der Definitionsbereich von γ_2 nicht direkt an den von γ_1 anschliessen, formal schreibe man etwa $\gamma_2 : [\beta + \delta, \gamma + \delta] \rightarrow U$, und ersetze $\gamma_2(t)$ durch $\gamma_2(t + \delta)$ in (15.21).

· Statt Wege sagt man auch Kurven oder Konturen, insbesondere dann, wenn man betonen will, dass Integrale nur von dem Bild $\gamma([\alpha, \beta])$ (welches auch “Spur des Wegs” genannt wird) und der Orientierung, und nicht einer expliziten Parametrisierung abhängen. Ein Weg heisst geschlossen, wenn $a = b$.

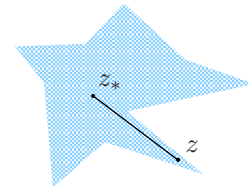
Bemerkungen. · Man prüfe nach, dass das Integral (15.19) die Regeln (i) und (ii) erfüllt und (iii) in der Form

$$\left| \int_{\gamma} f(w) dw \right| \leq \sup\{|f(w)| \mid w \in \gamma([\alpha, \beta])\} \cdot l(\gamma) \quad (15.22)$$

(Hier geht die Verträglichkeit der Multiplikation mit der euklidischen Norm = Absolutbetrag ein.)

· Bezüglich (iv) folgt wie in (15.8): Besitzt f auf U eine komplexe Stammfunktion F , so ist $\int_{\gamma} f(w) dw = F(b) - F(a)$. Dies ist gemäss 15.1 insbesondere dann der Fall, wenn f holomorph ist und U konvex. Im Allgemeinen aber *hängt das Integral tatsächlich vom Weg ab*.

· Etwas allgemeiner kann man aus 15.1 folgern: Ist U ein Sterngebiet, d.h. existiert ein “Sternmittelpunkt” $z_* \in U$ so, dass $\forall z \in U$ die Strecke $l(z_*, z) \subset U$, dann besitzt jede holomorphe Funktion auf U eine komplexe Stammfunktion.



· Man überlege sich so ausführlich wie möglich, dass die Definition 15.2 eine zufriedenstellende Verallgemeinerung des reellen Integrals 13.3 darstellt. Insbesondere treffen beide Definitionen zusammen, wenn $a < b \in \mathbb{R}$ und f eine holomorphe Fortsetzung von einem offenen Intervall $I \supset [a, b]$ auf eine offene Umgebung $U \supset I$ hinaus in die komplexe Ebene besitzt. Man beachte in diesem Zusammenhang auch die formale Ähnlichkeit zwischen (15.19) und der “Substitutionsregel” (13.18).

§ 16 Cauchy-Formeln

Beispiel 16.1. In der Praxis werden nebst den geraden Strecken $t \mapsto a + t(b - a)$ vor allem Kreise und Kreisbögen als Wege benötigt. Bereits in (2.14) hatten wir die offene Scheibe vom Radius R um $z_0 \in \mathbb{C}$ definiert als $B_R(z_0) = \{z \in \mathbb{C} \mid |z - z_0| < R\}$, die abgeschlossene Scheibe durch $\overline{B_R(z_0)} = \{z \in \mathbb{C} \mid |z - z_0| \leq R\}$. Der Kreis

$$C_R(z_0) := \{z \in \mathbb{C} \mid |z - z_0| = R\} = \overline{B_R(z_0)} \setminus B_R(z_0) \quad (16.1)$$

ist der Rand dieser Scheiben. Wir schreiben jetzt für den geschlossenen Weg mit Spur $C_R(z_0)$ in der positiven Orientierung (d.h. gegen den Uhrzeigersinn) $C_R^+(z_0)$ und in der negativen Orientierung $C_R^-(z_0)$. Gebräuchliche Parametrisierungen sind

$$[0, 2\pi] \ni t \mapsto z_0 + Re^{\pm it} \quad (16.2)$$

und

$$[0, 1] \ni t \mapsto z_0 + Re^{\pm 2\pi it} \quad (16.3)$$

§ 16. CAUCHY-FORMELN

Man greift auch häufig auf Bilder zurück, ohne Parametrisierungen explizit anzugeben. Die an die Definition 15.2 angepasste Verallgemeinerung des Hauptsatzkriteriums (15.10) lautet:

Proposition 16.2. *Sei $U \subset \mathbb{C}$ offen mit der Eigenschaft, dass für je zwei Punkte $a, b \in U$ ein Weg von a nach b in U existiert. Dann besitzt eine stetige Funktion $f : U \rightarrow \mathbb{C}$ genau dann eine komplexe Stammfunktion, wenn für jeden geschlossenen stückweise stetig differenzierbaren Weg γ*

$$\int_{\gamma} f(w)dw = 0 \quad (16.4)$$

In diesem Fall kann für jedes $z_ \in U$ eine Stammfunktion von f rekonstruiert werden durch $F(z) = \int_{\gamma(z_*, z)} f(w)dw$ für einen beliebigen Weg $\gamma(z_*, z)$ von z_* nach z in U und es gilt für jeden (stückweise stetig differenzierbaren) Weg γ von a nach b in U :*

$$\int_{\gamma} f(w)dw = F(b) - F(a) \quad (16.5)$$

Beweis. Siehe Lehrbücher. □

Im Unterschied zu 15.1 ist es aber *nicht richtig*, dass für beliebige nicht konvexe oder nicht-sternförmige offene Mengen (16.4) für jede stetige Funktion gilt:

Beispiel 16.3. Das Integral der Funktion $\mathbb{C} \setminus \{0\} \ni z \mapsto \frac{1}{z}$ über den Kreis vom Radius $R > 0$ um den Nullpunkt ist

$$\begin{aligned} \int_{C_R^+(0)} \frac{1}{w} dw &= \int_0^{2\pi} \frac{1}{Re^{it}} \frac{d}{dt}(Re^{it}) dt = \int_0^{2\pi} \frac{1}{Re^{it}} Re^{it} i dt = \int_0^{2\pi} i dt \\ &= 2\pi i \end{aligned} \quad (16.6)$$

und insbesondere nicht Null. Man bemerke, dass (16.6) von R unabhängig ist, und zur Sicherheit rechne man noch einmal für beliebiges $z_0 \in \mathbb{C}$:

$$\int_{C_R^+(z_0)} \frac{1}{w - z_0} dw = 2\pi i \quad \text{und auch noch} \quad \int_{C_R^-(z_0)} \frac{1}{w - z_0} dw = -2\pi i \quad (16.7)$$

Fazit: $z \mapsto \frac{1}{z}$ hat auf $\mathbb{C} \setminus \{0\}$ keine Stammfunktion. Hingegen hatten Sie in Aufgabe 10.3 gesehen, dass auf der geschlitzten Ebene $\mathbb{C} \setminus (-\infty, 0]$ der Hauptzweig des Logarithmus, definiert durch

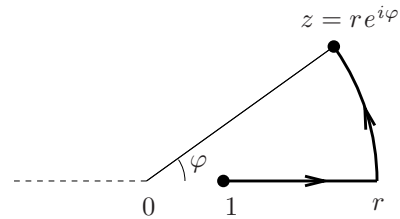
$$\text{Log}(z) = \ln r + i\varphi \quad (16.8)$$

für $z = re^{i\varphi}$ mit $r > 0$, $\varphi \in (-\pi, \pi)$, holomorph ist mit Ableitung $\text{Log}'(z) = \frac{1}{z}$. Auf dieser kleineren Menge also *existiert* eine Stammfunktion. In der Tat ist ja die geschlitzte Ebene sternförmig mit Sternmittelpunkt $1 \in z$ und es gilt

$$\text{Log}(z) = \int_{l(1, z)} \frac{1}{w} dw \quad (16.9)$$

Statt auf der direkten Strecke $l(1, z)$ kann auch entlang des rechts skizzierten Weges integriert werden. Das gerade Stück ist $[1, r] \ni t \mapsto t$, das gebogene $[0, \varphi] \ni t \mapsto re^{it}$, und

$$\int_1^r \frac{1}{s} ds + \int_0^\varphi \frac{1}{re^{it}} ire^{it} dt = \ln r + i\varphi \quad (16.10)$$



Es gehören zur Vollständigkeit der Diskussion jetzt noch Antworten auf die folgenden zwei Fragen:

- (1) Ist es *notwendig*, dass f holomorph ist, damit sie auf konvexen Mengen eine Stammfunktion besitzt?
- (2) Gibt es notwendige und hinreichende Charakterisierungen von Mengen U dafür, dass jede holomorphe Funktion auf U eine Stammfunktion besitzt?

Die Intention der Frage (1) ist *lokal*, d.h. man denkt an “kleine” Umgebungen eines gegebenen Punktes $z_0 \in U$, insbesondere Kreisscheiben $B_R(z_0)$. Die Antwort ist Ja, siehe 16.5.

Die Antwort auf die globale Frage (2) ist, dass genau dann jede auf einer offenen zusammenhängenden Menge U holomorphe Funktion eine Stammfunktion besitzt, wenn U “einfach zusammenhängend” ist. Anschaulich gesprochen bedeutet dies, dass U keine “Löcher im Inneren” hat oder, etwas präziser gesagt, dass sich jeder geschlossene Weg in U stetig auf einen Punkt zusammenziehen lässt. Eine vollständige Untersuchung würde uns hier zu weit führen. Ein Teil der Probleme rührt auch von der Mathematiker Wunsch her, möglichst allgemeine (auch nicht differenzierbare...) Wege zur Integration zuzulassen, und für solche ist es nicht so einfach zu zeigen, dass Inneres und Äusseres wohldefiniert sind.

In der Praxis werden stückweise stetig differenzierbare Wege genügen, und wir beschränken uns ausserdem auf solche, die nie stoppen und sich nicht selbst schneiden (sog. “einfache Kurven”). Dann lässt sich intuitiv bequem fassen die

Aussage: Sei $\gamma : [\alpha, \beta] \rightarrow \mathbb{C}$ einfach geschlossen (d.h. $\gamma(\alpha) = \gamma(\beta)$, aber $\gamma|_{[\alpha, \beta]}$ ist injektiv) und stückweise stetig differenzierbar mit $\dot{\gamma} \neq 0$ auf allen Stücken. Sei $U \subset \mathbb{C}$ eine offene Menge, die $\gamma([\alpha, \beta])$ und ihr gesamtes Innere enthält. Dann gilt für jede holomorphe Funktion $f : U \rightarrow \mathbb{C}$

$$\int_\gamma f(w) dw = 0 \quad (16.11)$$

Der nächste entscheidende Schritt ergibt sich aus der Verbindung von (16.11) mit dem Integral (16.6):

Theorem 16.4. Sei $U \subset \mathbb{C}$ offen und $f : U \rightarrow \mathbb{C}$ holomorph. Angenommen, für $z_0 \in U$ und $R > 0$ liegt $\overline{B_R(z_0)} \subset U$. Dann gilt für alle $z \in B_R(z_0)$:

$$f(z) = \frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{w - z} dw \quad (16.12)$$

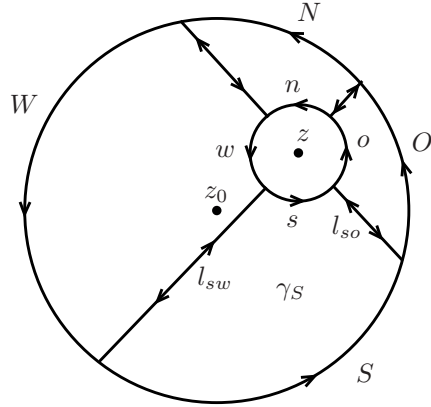
§ 16. CAUCHY-FORMELN

Beweis: Beh.: Für $r > 0$ mit $B_r(z) \subset B_R(z_0)$ gilt

$$\int_{C_R^+(z_0)} \frac{f(w)}{w-z} dw = \int_{C_r^+(z)} \frac{f(w)}{w-z} dw \quad (16.13)$$

Bew.: Wir zerlegen wie skizziert

$$\begin{aligned} \int_{C_R^+(z_0)} &= \int_O + \int_N + \int_W + \int_S \\ \int_{C_r^+(z)} &= \int_o + \int_n + \int_w + \int_s \end{aligned} \quad (16.14)$$



und addieren und subtrahieren die geraden Stücke in die Nebenhimmelsrichtungen, sodass

$$\int_{C_R^+(z_0)} - \int_{C_r^+(z)} = \int_{\gamma_O} + \int_{\gamma_N} + \int_{\gamma_W} + \int_{\gamma_S} \quad (16.15)$$

wo z.B.

$$\int_{\gamma_S} = \int_S - \int_{l_{so}} - \int_s + \int_{l_{sw}} \quad (16.16)$$

Nun liegt jeder der vier Wege mit seinem Inneren in einer (anderen) offenen konvexen Menge $\subset U$, auf der der Integrand

$$w \mapsto \frac{f(w)}{w-z} \quad (16.17)$$

jeweils holomorph ist. Damit folgt aus (16.11), dass $\int_{\gamma_S} = 0$ etc., und daraus dann (16.13).

2. Schritt: Aus (16.7) folgt, dass für jedes $r > 0$ wie oben

$$f(z) = f(z) \cdot \underbrace{\frac{1}{2\pi i} \int_{C_r^+(z)} \frac{1}{w-z} dw}_{=1} = \frac{1}{2\pi i} \int_{C_r^+(z)} \frac{f(z)}{w-z} dw \quad (16.18)$$

sodass zusammen mit (16.13) für jedes solches r

$$\frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{w-z} dw - f(z) = \frac{1}{2\pi i} \int_{C_r^+(z)} \frac{f(w) - f(z)}{w-z} dw \quad (16.19)$$

Zum Beweis der Behauptung reicht es daher aus, wenn wir für jedes $\epsilon > 0$ ein $r > 0$ angeben, für das wir zeigen können, dass der Betrag der rechten Seite kleiner als ϵ ist.

Bew.: Da f in z holomorph ist, ist der Differenzenquotient $w \mapsto \frac{f(w)-f(z)}{w-z}$ per Definition 11.1 stetig in z fortsetzbar, und daher nach 7.5 auf der folgenkompakten Menge $\overline{B_R(z_0)}$ beschränkt, sagen wir durch $M > 0$. Sei nun für $\epsilon > 0$ $r > 0$ so klein, dass $Mr < \epsilon$. Dann folgt aus der Standardabschätzung (15.22) des Integrals

$$\left| \frac{1}{2\pi i} \int_{C_r^+(z)} \frac{f(w) - f(z)}{w-z} dw \right| \leq \frac{1}{2\pi} M \cdot \underbrace{l(C_r^+(z))}_{=2\pi r} < \epsilon \quad (16.20)$$

□

Folgerungen

Man prüft nach, dass der Integrand von (16.12), z.B. in der Standardparametrisierung $w(t) = z_0 + Re^{it}$, $\dot{w}(t) = Rie^{it}$,

$$f(z) = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + Re^{it})}{z_0 + Re^{it} - z} \cdot Rie^{it} dt \quad (16.21)$$

die Voraussetzungen der Proposition 13.7 über parameterabhängige Integrale erfüllt. Insbesondere ist für festes t der Integrand in jedem $z \in B_R(z_0)$ komplex differenzierbar und die Abbildung auf die Ableitung nach z

$$(t, z) \mapsto \frac{d}{dz} \left(\frac{f(z_0 + Re^{it})}{z_0 + Re^{it} - z} \cdot Rie^{it} \right) = \frac{f(z_0 + Re^{it})}{(z_0 + Re^{it} - z)^2} \cdot Rie^{it} \quad (16.22)$$

ist stetig. Es folgt zunächst

$$f'(z) = \frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{(w - z)^2} dw, \quad (16.23)$$

und daraus durch vollständige Induktion, dass f unendlich oft komplex differenzierbar ist mit

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{(w - z)^{n+1}} dw, \quad (16.24)$$

Insbesondere folgt daraus wie angekündigt: Besitzt eine stetige Funktion $f : U \rightarrow \mathbb{C}$ eine komplexe Stammfunktion, so ist f bereits holomorph. (Denn: Eine Stammfunktion F von f ist holomorph und daher wie eben gezeigt unendlich oft komplex differenzierbar. Insbesondere ist $f = F'$ holomorph.)

Holomorphie ist also notwendig und hinreichend für die lokale Existenz einer Stammfunktion, und damit ist die Frage (1) auf Seite 100 beantwortet. Aus der folgenden Verbindung zu 11.4 wird klar, warum Kreise als Integrationswege eine besondere Rolle spielen.

Proposition 16.5 (Holomorphe Funktionen sind analytisch). *Sei $U \subset \mathbb{C}$ offen und $f : U \rightarrow \mathbb{C}$ holomorph. Sei $z_0 \in U$ und $R > 0$ mit $\overline{B_R(z_0)} \subset U$. Dann besitzt f eine für alle $z \in B_R(z_0)$ konvergente Potenzreihenentwicklung*

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k \quad (16.25)$$

mit den Koeffizienten

$$a_k = a_k^{(z_0)} = \frac{f^{(k)}(z_0)}{k!} = \frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{(w - z_0)^{k+1}} dw \quad (16.26)$$

Bemerkungen. · Hier ist R stets kleiner (und nie gleich) als der Konvergenzradius der Potenzreihe (16.25): Da U offen ist, existiert stets noch eine etwas grössere Kreisscheibe $\overline{B_{R'}(z_0)} \subset U$ mit $R' > R$, in deren Inneren die Potenzreihe (16.25) immer noch konvergiert.

§ 16. CAUCHY-FORMELN

Beweis. Wie oben nutzen wir die Trennung zwischen f und z in der Integraldarstellung (16.12) aus. Für festes $z \in B_R(z_0)$ ist

$$\left| \frac{z - z_0}{w - z_0} \right| = \frac{|z - z_0|}{R} =: \rho < 1 \quad (16.27)$$

konstant für alle $w \in C_R(z_0)$. Dort gilt daher (geometrische Reihe)

$$\left| \frac{1}{w - z} - \sum_{k=0}^n \frac{(z - z_0)^k}{(w - z_0)^{k+1}} \right| = \left| \frac{1}{w - z_0} \sum_{k=n+1}^{\infty} \frac{(z - z_0)^k}{(w - z_0)^k} \right| \leq \frac{1}{R} \cdot \frac{\rho^{n+1}}{1 - \rho} \quad (16.28)$$

Mit (16.12) und der Standardabschätzung des Integrals folgt

$$\begin{aligned} & \left| f(z) - \sum_{k=0}^n a_k (z - z_0)^k \right| \\ &= \left| \frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{w - z} dw - \sum_{k=0}^n \frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{(w - z_0)^{k+1}} dw \cdot (z - z_0)^k \right| \\ &= \left| \frac{1}{2\pi i} \int_{C_R^+(z_0)} f(w) \left(\frac{1}{w - z} - \sum_{k=0}^n \frac{(z - z_0)^k}{(w - z_0)^{k+1}} \right) dw \right| \\ &\leq \frac{1}{2\pi} \cdot \max\{|f(w)| \mid w \in C_R(z_0)\} \cdot \frac{1}{R} \cdot \frac{\rho^{n+1}}{1 - \rho} \cdot (2\pi R) \\ &\leq \max\{|f(w)| \mid w \in C_R(z_0)\} \cdot \frac{\rho^{n+1}}{1 - \rho} \end{aligned} \quad (16.29)$$

Für $n \rightarrow \infty$ konvergiert also die Reihe wie behauptet gegen $f(z)$. $\square$

Bemerkungen. Mittels (16.11) und (16.7) lässt sich als Verallgemeinerung von 16.4 zeigen, dass für einen einfach geschlossenen Weg γ , der mitsamt seinem Inneren ganz in U liegt, und einen Punkt $z \in U$ im Inneren von γ

$$\int_{\gamma} \frac{f(w)}{w - z} dw = \pm 2\pi i \cdot f(z) \quad (16.30)$$

je nach Orientierung von γ . Für z im Äusseren von γ ist das Integral 0. Und für einen ganz neugierigen geschlossenen Weg, der sich in U zu einem Punkt zusammenziehen lässt und z nicht trifft, gilt

$$\int_{\gamma} \frac{f(w)}{w - z} dw = 2\pi i \cdot f(z) \cdot W(\gamma, z) \quad (16.31)$$

wobei für jedes $z \in \mathbb{C} \setminus \gamma([\alpha, \beta])$

$$W(\gamma, z) := \frac{1}{2\pi i} \int_{\gamma} \frac{1}{w - z} dw \in \mathbb{Z} \quad (16.32)$$

die “Windungszahl” von γ um z berechnet.

§ 17 Residuensatz

Im vorigen § hatten wir aus der Cauchy-Formel 16.4 gefolgert, dass holomorphe Funktionen unendlich oft differenzierbar sind, und sich die Ableitungen nach den Formeln (16.24) durch Mittelung über geeignete Kreise berechnen lassen. Nun wollen wir die Logik umkehren, um Integrale durch Ableiten zu berechnen. Dabei beruht nach wie vor alles auf

- (i) der Invarianz des Integrals einer holomorphen Funktion unter stetigen Deformationen des Integrationsweges, solange entweder die Endpunkte festgehalten werden oder der Weg geschlossen ist und ausserdem die Deformation innerhalb der Menge stattfindet, in der die Funktion holomorph ist, und
- (ii) der nun elementaren Feststellung, dass für $n \in \mathbb{Z}$ und $R > 0$

$$\int_{C_R^+(0)} w^n dw = \begin{cases} 0 & n \neq -1 \\ 2\pi i & n = -1 \end{cases} \quad (17.1)$$

Denn für $n \neq -1$ hat die Funktion $z \mapsto z^n$ auf $\mathbb{C} \setminus \{0\}$ die Stammfunktion $\frac{1}{n+1} z^{n+1}$, und für $n = -1$ gilt (16.6).

· Ist beispielsweise eine Funktion f durch eine auf $\overline{B_R(0)}$ konvergente Potenzreihe $\sum a_n z^n$ gegeben, so lässt sich die Formel (16.26) als direkte Konsequenz von (17.1) auffassen:

$$\begin{aligned} \int_{C_R^+(0)} \frac{f(w)}{w^{k+1}} dw &= \int_{C_R^+(0)} \sum_{n=0}^{\infty} a_n w^{n-k-1} dw = \int_{C_R^+(0)} \sum_{n=-k-1}^{\infty} a_{n+k+1} w^n dw \\ &= \underbrace{\sum_{n=-k-1}^{-2} a_{n+k+1} \int_{C_R^+(0)} w^n dw}_{=0} + \int_{C_R^+(0)} \frac{a_k}{w} dw + \underbrace{\int_{C_R^+(0)} \sum_{n=0}^{\infty} a_{n+k+1} w^n dw}_{=0 \text{ s. (16.11)}} \\ &= 2\pi i \cdot a_k \end{aligned} \quad (17.2)$$

da der Rest der Potenzreihe eine holomorphe Funktion darstellt.

· Und noch ein Physiker-Argument für (17.1): Das Integral ist wegen der im Beweis von 16.4 benutzten Version von (i) unabhängig von R , andererseits transformiert sich der Integrand unter der Substitution $w \rightarrow Rw$ mit einem Faktor R^{n+1} . Das Integral kann also nur für $n = -1$ von null verschieden sein.

Definition 17.1. Für $z_0 \in \mathbb{C}$ und zwei komplexe Folgen $(a_n)_{n=0,1,\dots}$, $(a_{-n})_{n=1,2,\dots}$ heisst das Paar von Reihen

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n, \quad \sum_{n=1}^{\infty} a_{-n} (z - z_0)^{-n} \quad (17.3)$$

Laurent-Reihe mit Entwicklungspunkt z_0 . Sie heisst konvergent, wenn beide Reihen in (17.3) getrennt konvergieren. In diesem Fall ist der Wert der Laurent-Reihe die Summe der Grenzwerte. Man schreibt für die Laurent-Reihe bzw. ihre Werte auch $\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$.

§ 17. RESIDUENSATZ

· Gemäss 5.10, (5.26) konvergiert die Reihe mit positiven Potenzen von $z - z_0$ für

$$|z - z_0| < R := (\limsup |a_n|^{1/n})^{-1} \quad (17.4)$$

und analog die mit negativen Potenzen für

$$|z - z_0|^{-1} < S^{-1} := (\limsup |a_{-n}|^{1/n})^{-1} \quad (17.5)$$

und nach Anwendung von Proposition 11.3 folgt daraus, dass die Laurent-Reihe $L(z) := \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$ auf dem Kreisring

$$A_{R,S}(z_0) := \{z \in \mathbb{C} \mid S < |z - z_0| < R\} \quad (17.6)$$

konvergiert, dort eine holomorphe Funktion darstellt, und gliedweise abgeleitet werden darf. (Interessant ist dies natürlich eigentlich nur dann, wenn $S < R$.) Ein wesentlicher Unterschied zu Potenzreihen ist, dass Laurent-Reihen mit $a_{-1} \neq 0$ auf ihrem Konvergenz-Kreisring keine Stammfunktion besitzen.

Beispiel: Für $|z| > 1$: $\sum_{n=0}^{\infty} z^{-n} = \frac{1}{1 - \frac{1}{z}} = \frac{z}{1 - z}$

· Für $|z| > 0$: $\sum_{n=0}^{\infty} \frac{z^{-n}}{n!} = \exp\left(\frac{1}{z}\right)$

Umgekehrt gilt aber auch:

Proposition 17.2. Sei $U \subset \mathbb{C}$ offen und $f : U \rightarrow \mathbb{C}$ holomorph. Sei $z_0 \in \mathbb{C}$ und $\infty \geq R > S \geq 0$ mit $A_{R,S}(z_0) \subset U$. (Im Allgemeinen muss z_0 nicht in U sein!) Dann besitzt f eine für alle $z \in A_{R,S}(z_0)$ konvergente Entwicklung in eine Laurentreihe:

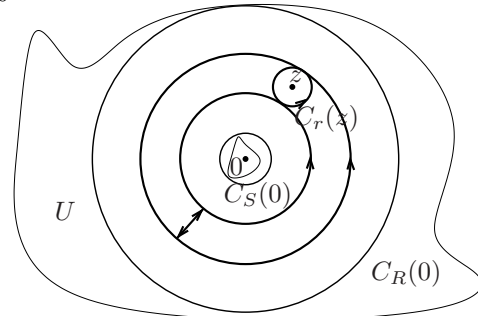
$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n \quad \forall z \in A_{R,S}(z_0) \quad (17.7)$$

Beweis. Wir nehmen der Einfachheit halber $z_0 = 0$.

Sei $r > 0$ so, dass $\overline{B_r(z)} \subset A_{R,S}(0)$. Dann gilt nach 16.4

$$f(z) = \frac{1}{2\pi i} \int_{C_r^+(z)} \frac{f(w)}{w - z} dw \quad (17.8)$$

Durch Nachahmen von (16.13) anhand der Skizze wird dies gleich



$$f(z) = \frac{1}{2\pi i} \int_{C_{|z|+r}^+(0)} \frac{f(w)}{w - z} dw + \frac{1}{2\pi i} \int_{C_{|z|-r}^+(0)} \frac{f(w)}{w - z} dw \quad (17.9)$$

Im ersten Integranden gilt $|z| < |w|$, d.h. $\frac{1}{w-z} = \sum_{n=0}^{\infty} \frac{z^n}{w^{n+1}}$ im zweiten $|z| > |w|$ und $\frac{1}{w-z} = -\sum_{n=1}^{\infty} \frac{w^{n-1}}{z^n} = \sum_{n=-\infty}^{-1} \frac{z^n}{w^{n+1}}$. Mit dem gleichen Konvergenzargument wie im Beweis von 16.5 folgt die Behauptung, wobei

$$a_n = \frac{1}{2\pi i} \cdot \left\{ \begin{array}{ll} \int_{C_{|z|+r}^+(0)} \frac{f(w)}{w^{n+1}} dw & n \geq 0 \\ \int_{C_{|z|-r}^+(0)} \frac{f(w)}{w^{n+1}} dw & n < 0 \end{array} \right\} = \frac{1}{2\pi i} \int_{C_\rho^+(0)} \frac{f(w)}{w^{n+1}} dw \quad \forall n \in \mathbb{Z} \quad (17.10)$$

für jedes $S < \rho < R$, da die Integrale von solchen ρ unabhängig sind. $\square$

Mit Zuversicht erweitern wir nun erstmals unseren Begriff von Funktionen auf solche, die streng genommen gar nicht überall definiert sind.

Definition 17.3. Sei $U \subset \mathbb{C}$ offen. Eine *holomorphe Funktion mit isolierten Singularitäten* auf U ist eine endliche oder abzählbare Teilmenge $\{z_1, z_2, \dots\}$ von (verschiedenen) Punkten $z_i \in U$, welche keinen Häufungspunkt in U hat, zusammen mit einer holomorphen Funktion $f : U \setminus \{z_1, z_2, \dots\} \rightarrow \mathbb{C}$.

· Aus der Tatsache, dass die Ausnahmemenge $\{z_1, z_2, \dots\}$ keinen Häufungspunkt in U hat, folgt, dass für jedes z_i ein $R_i > 0$ existiert so, dass $B_{R_i}(z_i) \subset U$ und f auf der “punktierten Scheibe” $B_{R_i}(z_i) \setminus \{z_i\}$ definiert und holomorph ist. Aus 17.2 folgt dann, dass f auf $B_{R_i}(z_i) \setminus \{z_i\} = A_{R_i,0}(z_i)$ durch eine Laurentreihe dargestellt wird, d.h.

$$f(z) = \sum_{n=-\infty}^{\infty} a_n^{(i)}(z - z_i)^n \quad \forall z \in B_{R_i}(z_i) \setminus \{z_i\} \quad (17.11)$$

Man nennt dabei

$$H_i(z) := \sum_{n=-\infty}^{-1} a_n^{(i)}(z - z_i)^n \quad (17.12)$$

den *Hauptteil* der Laurentreihe ($\sum_{n=0}^{\infty}$ den Nebenteil), und

$$a_{-1}^{(i)} =: \text{Res}_{z_i}(f) \quad (17.13)$$

das *Residuum* von f in z_i . Gemäss (17.10) gilt nämlich

$$\text{Res}_{z_i}(f) = \frac{1}{2\pi i} \int_{C_\rho^+(z_i)} f(w) dw \quad (17.14)$$

für alle $0 < \rho < R_i$ (was man aber für die Berechnung selten benutzt).

· Verschwindet der Hauptteil, d.h. gilt $a_n^{(i)} = 0 \quad \forall n < 0$, so nennt man die zugehörige Singularität *hebbar*. Das ist genau dann der Fall, wenn sich f in z_i holomorph fortsetzen lässt. (Es liegt also gar keine Singularität vor.)

· Ist der Hauptteil endlich, d.h. gilt für ein $k_i > 0$ $a_n^{(i)} = 0$ für $n < -k_i$ und $a_{-k_i}^{(i)} \neq 0$, so nennt man die zugehörige Singularität eine *Pol der Ordnung k_i* .

· Ist der Hauptteil unendlich, so heisst die Singularität *wesentlich* (oder auch “von unendlicher Ordnung”).

· Im Folgenden interessieren uns nur Pole endlicher Ordnung. Eine holomorphe Funktion mit isolierten Singularitäten auf U , welche allesamt Pole (sc. endlicher Ordnung) sind, heisst *meromorphe Funktion* auf U .

· Bei meromorphen Funktionen macht es Sinn, den Wertebereich von $\mathbb{C}$ auf $\mathbb{C} \cup \{\infty\}$ zu erweitern und den Polen formal den Wert ∞ zuordnend $f : U \rightarrow \mathbb{C} \cup \{\infty\}$ zu schreiben.

Diese Vorbereitungen versetzen uns nun in die Lage, eine für alle praktischen Zwecke hinreichende Version des Residuensatzes zu formulieren. Zuvor noch eine Formel zur Berechnung von Residuen meromorpher Funktionen, welche direkt aus der Laurententwicklung und der Taylorformel (16.26) herzuleiten ist.

§ 17. RESIDUENSATZ

Lemma 17.4 (Berechnung von Residuen). *Ist $z_0 \in U$ ein Pol der Ordnung k einer meromorphen Funktion f auf U , so gilt*

$$\operatorname{Res}_{z_0}(f) = \lim_{z \rightarrow z_0} \frac{1}{(k-1)!} \left(\frac{d}{dz} \right)^{k-1} ((z - z_0)^k f(z)) \quad (17.15)$$

Persönlich finde ich das direkte Einsetzen von Potenzreihen ineinander ökonomischer, da es die für die Benutzung der Formel (17.15) sowieso benötigte Ordnung der Polstelle gleich mitliefert.

Beispiel: Für kleines z ist

$$\begin{aligned} \frac{1}{z(\exp(z) - 1)} &= \frac{1}{z(z + \frac{z^2}{2} + \dots)} = \frac{1}{z^2 + \frac{z^3}{2} + \dots} = \frac{1}{z^2} \cdot \frac{1}{1 + \frac{z}{2} + \dots} \\ &= \frac{1}{z^2} (1 - \frac{z}{2} + \mathcal{O}(z^2)) = \frac{1}{z^2} - \frac{1}{2z} + \mathcal{O}(1) \\ \Rightarrow \operatorname{Res}_0 \frac{1}{z(\exp(z) - 1)} &= -\frac{1}{2} \end{aligned} \quad (17.16)$$

Theorem 17.5. *Sei f eine meromorphe Funktion auf einer offenen Menge $U \subset \mathbb{C}$. Dann gilt für jeden positiv orientierten einfach geschlossenen stückweise stetig differenzierbaren Weg γ , der mitsamt seinem Inneren ganz in U liegt und nicht die Pole von f trifft,*

$$\int_{\gamma} f(w) dw = 2\pi i \cdot \sum_{\substack{\text{Pole } z_i \text{ im} \\ \text{Inneren von } \gamma}} \operatorname{Res}_{z_i}(f) \quad (17.17)$$

Bemerkungen. Der in der Aussage vorkommende Begriff des “Inneren” eines einfach geschlossenen Wegs kann mit Hilfe der Windungszahl (16.32) sauber definiert werden. In den Anwendungen ist es aber wichtiger, sich über die Orientierung Rechenschaft abzulegen. Es sei auch darauf hingewiesen, dass wir im Beweis die Aussagen (16.11), (16.30) benutzen, die wir nicht für alle fraglichen Wege streng bewiesen haben.

Beweis. · Zunächst wollen wir begründen, dass die Summe auf der rechten Seite von (17.17) endlich ist: Da γ zusammen mit seinem Inneren beschränkt und abgeschlossen sind (das folgt aus der genauen Definition des “Inneren”, siehe etwa (16.30)), so müssten, falls unendlich viele Polstellen von f im Inneren lägen, diese einen Häufungspunkt in U haben (s. 4.17), im Widerspruch zur Definition einer meromorphen Funktion.

· Es seien für $r \in \mathbb{N}$ (oder allenfalls 0) $z_1, \dots, z_r \in U$ die Polstellen von f im Inneren von γ , und für $i = 1, \dots, r$

$$H_i(z) = \sum_{n=-k_i}^{-1} a_n^{(i)} (z - z_i)^n \quad (17.18)$$

$a_n^{(i)} \in \mathbb{C}$, der Hauptteil von f am i -ten Pol. $H_i(z)$ ist eine meromorphe Funktion auf $\mathbb{C}$ mit einem Pol bei z_i , insbesondere holomorph auf $\mathbb{C} \setminus \{z_i\}$.

· In einer Umgebung eines jeden z_i verschwindet dann der Hauptteil von

$$\tilde{f}(z) := f(z) - \sum_{i=1}^r H_i(z) \quad (17.19)$$

so dass sich $\tilde{f}$ zu einer holomorphen Funktion auf einer Umgebung von γ mitsamt seinem Inneren fortsetzen lässt. Aus (16.11) folgt $\int_{\gamma} \tilde{f}(w)dw = 0$ und daher

$$\int_{\gamma} f(w)dw = \sum_{i=1}^r \int_{\gamma} H_i(w)dw \quad (17.20)$$

· Nun ist für jedes i

$$\int_{\gamma} \sum_{n=-k_i}^{-2} a_n^{(i)} (z - z_i)^n = 0 \quad (17.21)$$

denn der Integrand besitzt auf ganz $\mathbb{C}$ die Stammfunktion

$$\sum_{n=-k_i}^{-2} \frac{a_n^{(i)}}{n+1} (z - z_i)^{n+1} \quad (17.22)$$

Es folgt

$$\int_{\gamma} f(w)dw = \sum_{i=1}^r \int_{\gamma} \frac{a_{-1}^{(i)}}{w - z_i} dw \stackrel{(16.30)}{=} 2\pi i \cdot \sum_{i=1}^r \text{Res}_{z_i}(f) \quad (17.23)$$

□

Bemerkungen. Die Verallgemeinerung von 17.5 für einen nicht notwendig einfach geschlossenen Weg γ , der sich in U zu einem Punkt zusammenziehen lässt und die isolierten Singularitäten der (sonst) holomorphen Funktion f auf U vermeidet, lautet

$$\int_{\gamma} f(w)dw = 2\pi i \cdot \sum_i \text{Res}_{z_i}(f) \cdot W(\gamma, z_i) \quad (17.24)$$

wobei in der Summe die Windungszahl (16.32) weiter nur für endlich viele Singularitäten z_i von null verschieden ist.

Auswertung reeller Integrale

Die folgenden Beispiele enthalten als gemeinsame Zutaten:

- (1) die Identifikation eines gegebenen reellen Integrals als Kurvenintegral einer meromorphen Funktion. Dabei muss der reelle Integrand nicht unbedingt “erhalten” bleiben, sondern kann als Real-/Imaginärteil, oder auch eine sonstige Linearkombination von komplexen Integranden auftauchen.
 - (2) das Schliessen der Wege in der komplexen Ebene unter Berücksichtigung von (3). Bei uneigentlichen Integralen werden häufig Grenzwerte über Familien von Wegen gebildet.
 - (3) das Abschätzen von bei (2) eingeführten zusätzlichen Beiträgen, gegebenenfalls mit einer Anpassung der gewählten Schliessung.
 - (4) die Auswertung des geschlossenen Kurvenintegrals mit dem Residuensatz.
- Die genauen Rezepte der gängigen Klassen unterscheiden sich allerdings in den Details, was man am besten durch eigenes Ausprobieren schätzen lernt.

§17. RESIDUENSATZ

Beispiel 17.6. Seien $a, b \in \mathbb{R}$ mit $a > |b|$. Dann erkennen wir in

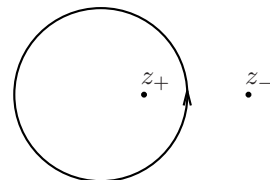
$$I(a, b) = \int_0^{2\pi} \frac{1}{a + b \cos \varphi} d\varphi \quad (17.25)$$

die Parametrisierung eines Integrals über den positiv orientierten Einheitskreis

$$\int_0^{2\pi} \frac{1}{a + \frac{b}{2}(e^{i\varphi} + e^{-i\varphi})} \frac{1}{i} e^{-i\varphi} \frac{d}{d\varphi} e^{i\varphi} = \frac{1}{i} \int_{C_1^+(0)} \frac{1}{a + \frac{b}{2}(w + w^{-1})} \cdot \frac{1}{w} dw \quad (17.26)$$

eines auf $\mathbb{C}$ meromorphen Integranden mit Polen bei

$$\frac{b}{2}z^2 + az + \frac{b}{2} = 0 \quad (17.27)$$



Für $b = 0$ ist dies genau der Ursprung mit Residuum 1, d.h.

$$I(a, 0) = \frac{2\pi}{a} \quad (17.28)$$

Für $b > 0$ lösen wir $\frac{b}{2}z^2 + az + \frac{b}{2} = \frac{b}{2}(z - z_+)(z - z_-)$ mit $z_{\pm} = \frac{-a \pm \sqrt{a^2 - b^2}}{b}$. Wegen $z_+ - z_- = \frac{2\sqrt{a^2 - b^2}}{b} \neq 0$ und $z_+ z_- = 1$ liegt genau ein Pol (nämlich z_+) in $B_1(0)$, und das Residuum dort ist

$$\lim_{z \rightarrow z_+} \frac{z - z_+}{\frac{b}{2}(z - z_+)(z - z_-)} = \frac{1}{\frac{b}{2}(z_+ - z_-)} = \frac{1}{\sqrt{a^2 - b^2}} \quad (17.29)$$

Es folgt

$$I(a, b) = \frac{2\pi}{\sqrt{a^2 - b^2}} \quad (17.30)$$

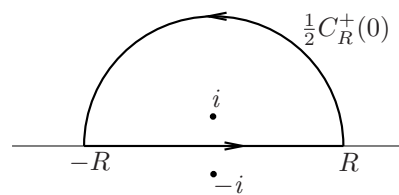
(was bei $b \rightarrow 0$ noch in (17.28) übergeht).

Beispiel 17.7. Das uneigentliche Integral (vgl. (13.31))

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \lim_{R \rightarrow \infty} \int_{-R}^R \frac{1}{1+x^2} dx \quad (17.31)$$

dessen Konvergenz man durch Vergleich mit $\int_1^{\infty} 1/x^2$ leicht verifiziert, und das sich auch elementar berechnen lässt, betrachten wir als den Grenzwert eines komplexen Kurvenintegrals von $-R$ nach R der auf $\mathbb{C}$ meromorphen Funktion $z \mapsto \frac{1}{1+z^2}$ mit Polen bei $\pm i$.

Alsdann schliessen wir den Weg durch einen Halbkreis in der oberen Halbebene, auf dem (für $R > 1$)



$$\left| \frac{1}{1+w^2} \right| \leq \frac{1}{R^2 - 1} \quad (17.32)$$

so dass durch Standardabschätzung

$$\left| \int_{\frac{1}{2}C_R^+(0)} \frac{1}{1+w^2} dw \right| \leq \frac{1}{R^2 - 1} \cdot \pi R \rightarrow 0 \quad \text{für } R \rightarrow \infty \quad (17.33)$$

Es folgt

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = 2\pi i \cdot \text{Res}_i \frac{1}{1+z^2} = 2\pi i \cdot \frac{1}{2i} = \pi \quad (17.34)$$

Beispiel 17.8. Das Integral

$$S := \int_{-1}^1 \sqrt{1-x^2} dx \quad (17.35)$$

ist Prototyp einer Klasse, die in der kanonischen Mechanik als Wirkungs- oder Periodenintegrale eine grosse Rolle spielen. Zur seiner Berechnung schieben wir (ganz unintuitiv) als erstes das Integral hinaus in die komplexe Ebene, d.h. wir schreiben unter Berufung auf 13.7

$$S = \lim_{\epsilon \rightarrow 0} \int_{-1}^1 \sqrt{1-(x-i\epsilon)^2} dx \quad (17.36)$$

wobei $\sqrt{\cdot}$ die auf S. 72 diskutierte Fortsetzung der Wurzelfunktion in die komplexe Ebene ist: $\sqrt{1-z^2} > 0$ für $1-z^2 > 0$, d.h. $z \in (-1, 1)$, und nicht definiert für $z \in [1, \infty) \cup (-\infty, -1] \subset \mathbb{C}$. Nach dem Hinausschieben ändert sich dann nichts, wenn wir den Integranden durch $i\sqrt{z^2-1}$ ersetzen, wo jetzt $\sqrt{z^2-1} > 0$ für $z > 1$, $\sqrt{z^2-1} < 0$ für $z < -1$, und nicht definiert für $z \in [-1, 1]$. Nach dem Rotieren des Schnitts ist also

$$S = \lim_{\epsilon \rightarrow 0} \int_{-1-i\epsilon}^{1-i\epsilon} i\sqrt{w^2-1} dw \quad (17.37)$$

oder bildlich

$$\int_{-1}^1 \sqrt{1-x^2} dx = \int_{-1-i\epsilon}^{1-i\epsilon} i\sqrt{w^2-1} dw$$

Nun unterscheidet sich der Integrand auf der gegenüberliegenden Seite des Schnittes im Limes $\epsilon \rightarrow 0$ gerade durch ein Vorzeichen, sodass

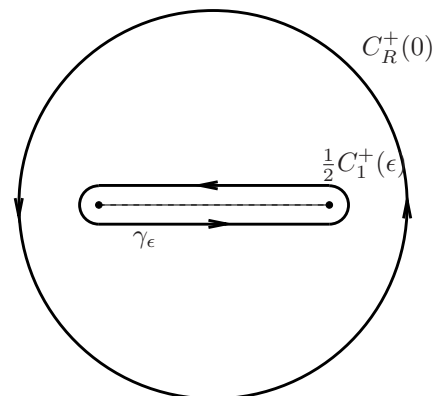
$$S = \frac{1}{2} \lim_{\epsilon \rightarrow 0} \left(\int_{-1-i\epsilon}^{1-i\epsilon} i\sqrt{w^2-1} dw - \int_{-1+i\epsilon}^{1+i\epsilon} i\sqrt{w^2-1} dw \right) \quad (17.38)$$

Auf den kleinen Halbkreisen $\frac{1}{2}C_{\pm 1}^+(\epsilon)$ bleibt im Limes $\epsilon \rightarrow 0$ der Integrand beschränkt, sodass

$$\lim_{\epsilon \rightarrow 0} \int_{\frac{1}{2}C_{\pm 1}^+(\epsilon)} i\sqrt{w^2-1} dw = 0 \quad (17.39)$$

Es folgt durch Schliessen des Weges

$$S = \frac{i}{2} \lim_{\epsilon \rightarrow 0} \int_{\gamma_\epsilon} \sqrt{w^2-1} dw \quad (17.40)$$



§17. RESIDUENSATZ

Nun liegt aber ein Integral einer holomorphen Funktion über einen geschlossenen Weg in $\mathbb{C} \setminus [-1, 1]$ vor, sodass dieses unabhängig von ϵ ist, und auch gleich dem Integral über den Kreis $C_R^+(0)$ für $R > 1$. Es konvergiert dort die Entwicklung

$$\sqrt{w^2 - 1} = w - \frac{1}{2w} - \frac{1}{8w^3} + \dots \quad (17.41)$$

die wir gliedweise integrieren können, um als Endergebnis

$$S = \frac{i}{2} \int_{C_R^+(0)} \left(w - \frac{1}{2w} - \frac{1}{8w^3} + \dots \right) dw = \frac{i}{2} \int_{C_R^+(0)} \frac{-1}{2w} dw = -\frac{i}{2} \cdot \frac{2\pi i}{2} = \frac{\pi}{2} \quad (17.42)$$

zu erhalten. (Auch dieses Integral geht natürlich mit elementaren Mitteln...)

Beispiel 17.9. Das Gausssche Integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad (17.43)$$

spielt eine grundlegende Rolle unter anderem in der Fouriertransformation, der Wellenmechanik, der Wahrscheinlichkeitstheorie und dem Pfadintegralzugang zur Quantenfeldtheorie. Es gibt eine Vielzahl von Tricks zur Berechnung von (17.43), von denen der Residuensatz einer der weniger bekannten (und auch wenigst offensichtlichen) ist. Andererseits ist die Methode ziemlich rein ein-dimensional²⁶, sodass wir sie hier vorführen können.

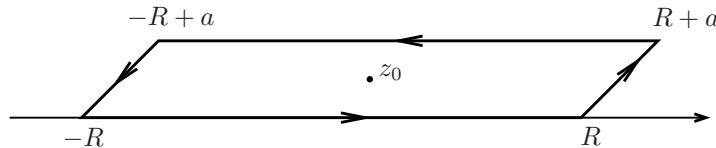
Wir setzen unser Vorwissen über das Ergebnis ein in die Zahl

$$a := \sqrt{\pi} e^{\pi i/4} = \sqrt{\frac{\pi}{2}}(1 + i) \quad \text{mit der Eigenschaft} \quad a^2 = \pi i \quad (17.45)$$

Dann erfüllt die auf $\mathbb{C}$ meromorphe Funktion $f(z) = \frac{e^{-z^2}}{1 + e^{-2az}}$ die Gleichung

$$\begin{aligned} f(z) - f(z + a) &= \frac{e^{-z^2}}{1 + e^{-2az}} - \frac{e^{-z^2-2az-a^2}}{1 + e^{-2az-2a^2}} = \frac{e^{-z^2}}{1 + e^{-2az}} - \frac{e^{-z^2-2az-\pi i}}{1 + e^{-2az-2\pi i}} \\ &= \frac{e^{-z^2}(1 + e^{-2az})}{1 + e^{-2az}} = e^{-z^2} \end{aligned} \quad (17.46)$$

Wir betrachten nun für $R > 0$ das Integral von f über den Rand $\square(R)$ des Parallelogramms mit den Ecken $-R, R, R + a, -R + a$.



²⁶Die schnellste Methode wertet das Quadrat des Integrals durch zwei-dimensionale Integration in Polarkoordinaten aus:

$$\left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 = \int_{\mathbb{R}^2} e^{-x^2-y^2} dx dy = \int_0^{\infty} \int_0^{2\pi} e^{-r^2} r d\varphi dr = 2\pi \int_0^{\infty} e^{-r^2} r dr = \pi \quad (17.44)$$

Einerseits enthalten dann wegen (17.46) die langen Seiten das gewünschte Integral:

$$(17.47) \quad \int_{-R}^R f(w)dw + \int_{R+a}^{-R+a} f(w)dw = \int_{-R}^R f(x)dx - \int_{-R}^R f(x+a)dx = \int_{-R}^R e^{-x^2}dx$$

Auf der rechten kurzen Seite ist mit $w = R + ta$, $t \in [0, 1]$

$$(17.48) \quad \begin{aligned} |e^{-2aw}| &= e^{\operatorname{Re}(-2a(R+ta))} = e^{-R\sqrt{2\pi}} \\ \text{und } |e^{-w^2}| &= e^{-\operatorname{Re}(R+ta)^2} = e^{-R^2 - tR\sqrt{2\pi}} \leq e^{-R^2} \end{aligned}$$

und daher gilt für R gross genug

$$(17.49) \quad \left| \frac{e^{-w^2}}{1 + e^{aw}} \right| \leq \frac{e^{-R^2}}{1 - e^{-R\sqrt{2\pi}}} \longrightarrow 0 \text{ für } R \rightarrow \infty$$

Nach einer analogen (bitte nachzuprüfenden) Abschätzung auf der linken kurzen Seite von $\square(R)$ folgt mittels Standardabschätzung des Integrals

$$(17.50) \quad \lim_{R \rightarrow \infty} \left(\int_R^{R+a} f(w)dw + \int_{-R+a}^{-R} f(w)dw \right) = 0$$

Andererseits liegen die Pole von $f(z)$ bei $-2az_n = (2n+1)\pi i = (2n+1)a^2$ für $n \in \mathbb{Z}$, d.h. $z_n = \frac{2n+1}{2}a$. Im Inneren des Parallelogramms befindet sich genau $z_0 = \frac{a}{2}$, und das Residuum ist

$$(17.51) \quad \operatorname{Res}_{z_0} f = \frac{e^{-z_0^2}}{\frac{d}{dz}\big|_{z=z_0}(1 + e^{-2az})} = \frac{e^{-\pi i/4}}{-2ae^{-a^2}} = -\frac{i}{2\sqrt{\pi}}$$

Es folgt aus obigen Beziehungen und dem Residuensatz

$$(17.52) \quad \begin{aligned} \int_{-\infty}^{\infty} e^{-x^2}dx &= \lim_{R \rightarrow \infty} \int_{\square(R)} f(w)dw = 2\pi i \cdot \operatorname{Res}_{z_0} f = 2\pi i \cdot \frac{-i}{2\sqrt{\pi}} \\ &= \sqrt{\pi} \end{aligned}$$

wie behauptet. □