Vorlesungsskript Analysis I und II

Prof. Bernd Ammann

Wintersemester 2018/19

Sommersemester 2019

Universität Regensburg

Datum der aktuellen Version: 13. Oktober 2021

Großdruck

Inhaltsverzeichnis

Vorwort vii

Logischer Aufbau des ersten Kapitels ix

Warnungen ix

Kapitel 1. Elementare Logik und Grundlagen der Mathematik 1

1. Die Struktur des mathematischen Denkens 1
2. Aussagenlogik 8
3. Mengen 18
4. Quantoren 25
5. Potenzmenge und Mengensysteme 30
6. Paare und kartesische Produkte 32
7. Relationen, funktionale Relationen, Abbildungen 36
8. Familien 49

Literatur für das bisherige Kapitel 52

Kapitel 2. Zahlen 53

1. Die natürlichen Zahlen 53
2. Etwas Kombinatorik 63
3. Die ganzen Zahlen 72
4. Die rationalen Zahlen 74

5. Geordnete Körper 77
6. Die reellen Zahlen 84
6.1. Unzulänglichkeit von $\mathbb{Q}$ 85
6.2. Die Supremumseigenschaft 86
6.3. Axiome der reellen Zahlen 90
6.4. Dedekindsche Schnitte 94
7. Die komplexen Zahlen 112

Kapitel 3. Folgen und Reihen 121
1. Folgen 121
1.1. Konvergenz von Folgen 121
1.2. Monotone Folgen 129
1.3. Teilfolgen 131
1.4. Erweiterte reelle Zahlen und uneigentliche Konvergenz 132
1.5. Limes inferior und superior 137
1.6. Cauchy-Folgen 141
2. Reihen 144
2.1. Motivation von Reihen: Dezimal-Darstellung reeller Zahlen 144
2.2. Definition und elementare Eigenschaften 148
2.3. Konvergenzkriterien 151
2.4. Absolute Konvergenz 160
2.5. Alternierende Reihen 161
2.6. Umordnung von Reihen 163
3. Einige durch Reihen definierte Funktionen 172
3.1. Exponentialfunktion 172

3.2. Sinus- und Kosinus-Funktion 176
3.3. Eulersche Zahl 180
3.4. Exponentialfunktion (Fortsetzung) 181

Kapitel 4. Stetigkeit und Grenzwerte von Funktionen 185
1. Stetigkeit 185
2. Zwischenwertsatz 190
3. Stetigkeit von exp, cos und sin und Definition von log 194
4. Die Kreizahl π und Periodizität von cos und sin. 198
5. Metrische Räume und Grundbegriffe der Topologie 200
6. Grenzwerte von Funktionen 211

Kapitel 5. Differential-Rechnung für Funktionen einer Veränderlichen 215
1. Definition und elementare Eigenschaften 215
2. Lokale Extrema 222
3. Mittelwertsätze 223
4. Höhere Ableitungen und Taylorscher Satz 225

Kapitel 6. Integral-Rechnung für Funktionen einer Veränderlichen 233
1. Partitionen und Treppenfunktionen 233
2. Das Riemann-Integral 235
3. Monotone Funktionen sind Riemann-integrierbar 242
4. Stetige Funktionen sind Riemann-integrierbar 244
5. Hauptsatz der Differential- und Integralrechnung 245
6. Uneigentliche Riemann-Integrale 253

Kapitel 7. Stetigkeit und Differenzierbarkeit von Grenzwerten und Reihen 257

1. Metrische Räume: Wiederholung und Cauchy-Folgen 257
2. Punktweise und gleichmäßige Konvergenz 260
3. Differentiation von Folgen und Reihen 267
4. Potenzreihen und analytische Funktionen 271

Kapitel 8. Topologie 275
1. Normierte Vektorräume 275
2. Kontraktionen und Banachscher Fixpunktsatz 285
3. Topologische Räume 288
3.1. Wiederholung: Eigenschaften eines metrischen Raums 288
3.2. Definition topologischer Räume 289
3.3. Konvergenz in topologischen Räumen 295
4. Zusammenhang und Wegzusammenhang 299
5. Folgenkompaktheit 302
6. Kompaktheit 306

Kapitel 9. Differential-Rechnung für Funktionen in mehreren Veränderlichen 317
1. Vorbemerkungen 317
2. Differenzierbarkeit in mehreren Variablen 320
3. Höhere Ableitungen 338
3.1. Satz von Schwarz 338
3.2. Höhere Ableitungen als multi-lineare Abbildungen 344
3.3. Satz von Taylor 347
3.4. Einschub: Quadratische Formen 350
3.5. Lokale Extrema 355
4. Lokale Umkehrung differenzierbarer Abbildungen 358

5. Der Satz über implizit definierte Funktionen 367
6. Untermannigfaltigkeiten 374
6.1. Definition und erste Beispiele 374
6.2. Der Satz vom regulären Wert 378
6.3. Der Immersionssatz 385
6.4. Einbettungen von Untermannigfaltigkeiten 388
6.5. Lokale Parametrisierungen von Untermannigfaltigkeiten 391
6.6. Der Tangentialraum 396
7. Extrema mit Nebenbedingungen 398

Kapitel 10. Gewöhnliche Differentialgleichungen 403
1. Motivation 403
2. Definition und Reduktion auf autonome Gleichungen erster Ordnung 406
3. Flusslinien und Erhaltungsgrößen 411
4. Der Satz von Picard-Lindelöf 415
5. Picard-Lindelöf für gewöhnliche Differentialgleichungen höherer Ordnung 432
6. Lineare gewöhnliche Differentialgleichungen 433
7. Lineare gewöhnliche Differentialgleichungen mit konstanten Koeffizienten 441
8. Differentialgleichungen mit getrennten Variablen 442
9. Mehr zum Fluss eines zeitabhängigen Vektorfelds 445
10. Einige Kommentare zum Langzeitverhalten von gew. Diffifferentialgleichungen 449
11. Stabilität und Verhalten in der Nähe von kritischen Punkten 452

Anhang A. Mehr Details zu den Grundlagen der Logik 457
1. Das Russellsche Paradoxon 457
2. Axiomatische Mengenlehre 459

Anhang B. Die Peano-Axiome 465
1. Die Axiome und erste Konsequenzen 465
2. Vollständige Induktion und rekursive Definition 470
3. Ordnung der natürlichen Zahlen 476

Anhang C. Konstruktion von $\mathbb{R}$ mit Hilfe von Cauchy-Folgen 481
1. Mehr zu Äquivalenzrelationen 481
2. Folgen, Konvergenz und Cauchy-Folgen 484
3. Existenz und Eindeutigkeit der reellen Zahlen 496

Anhang Z. Überblick über algebraische Strukturen 503

Anhang. Literaturverzeichnis 505

Anhang. Stichworte 507

Anhang. Symbole 517

Vorwort

Dies ist das Skript zu den Vorlesungen „Analysis I“ und „Analysis II“, die ich im Wintersemester 2018/19 und Sommersemester 2019 an der Universität Regensburg halte. Das Skript sollte ungefähr das enthalten, was an der Tafel stand, und zusätzlich einige dazu gegebene Erklärungen widerspiegeln. Anhänge ergänzen die Themen, sie wurden nicht in der Vorlesung behandelt. Einige Beispiele und Erweiterungen, die aus Zeitgründen nicht in der Vorlesung behandelt wurden, sind so gedruckt, und sind natürlich gut für die Nachbereitung beziehungsweise Vertiefung geeignet. Falls Sie einen Tippfehler finden oder sonstige Verbesserungsvorschläge haben, so senden Sie bitte eine Email an mich. Leider enthält das Skript bisher nur wenige Bilder und Zeichnungen.

Die aktuelle Version ist unter dem Link

`http://www.mathematik.ur.de/ammann/lehre/2019s_analysis2/analysisi+II.pdf`

verfügbar.

Die Homepage der Vorlesung ist

`http://www.mathematik.ur.de/ammann/lehre/2010s_analysis2`.

Regensburg, April 2019,

Bernd Ammann

Logischer Aufbau des ersten Kapitels

Die Abschnitte des ersten Kapitels sind – aus Sichtweise des logischen Aufbaus – nicht optimal angeordnet. Unter anderem werden Begriffe wie Mengen, Funktionen, kartesische Produkte und ähnliches teilweise benutzt, bevor sie eingeführt werden. Dies erscheint vertretbar, da ich davon ausgehe, dass jede(r) Hörer(in) bereits eine gewisse Vorstellung von manchen der Begriffen hat, und wenn nicht, dann wird es sich in einem zweiten Durchgang klären, was gemeint ist. Die in der Vorlesung und im Skript gewählte Reihenfolge hat den Vorteil, dass man bereits früh in den Übungen auf die wichtigen Punkte eingehen kann, wie zum Beispiel die Themen Aussagenlogik und Quantoren.

Ich empfehle, mit einem gewissen zeitlichen Abstand dieses Kapitel noch einmal durchzuschauen, dann wird wahrscheinlich manches bisher unklare sich klären.

Warnungen

Legen Sie das Skript nicht in eine Ecke mit dem ruhigen Gewissen, es ja später lesen und durcharbeiten zu können. Beginnen Sie sobald wie möglich, die Lücken zu schließen. Schwierige Beweise durchschauen Sie am besten, wenn Sie sich überlegen, was der Beweis in konkreten Beispielen macht.

Bilder, Skizzen und Abschnitte die mit USW angedeutet werden, in der Vorlesung behandelt wurden, aber aus Zeitgründen noch nicht getext wurden, sind selbstverständlich auch relevant für mündliche und schriftliche Prüfungen.

KAPITEL 1

Elementare Logik und Grundlagen der Mathematik

1. Die Struktur des mathematischen Denkens

17.10.

Die natürlichen Zahlen werden seit Jahrtausenden intuitiv benutzt und untersucht. Sie sind uns vertraut, ohne dass wir aber zunächst wissen, was sie charakterisiert. So ähnlich war es mit vielen mathematischen Konzepten, zum Beispiel dem mathematischen Konzept der unendlichen Summe. Man nutzte viele Konzepte lange in einer vagen Bedeutung, ohne sich zu überlegen, wie man sie definiert. Alle Personen stimmen wohl überein, dass der Wert der unendlichen Summe

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$$

die Zahl 2 sein sollte. Und derartige unendliche Summen sind oft wichtig, zum Beispiel in der Zinseszinsrechnung, bei physikalischen Problemen und vielem mehr. Deswegen wurden Sie bereits im Mittelalter diskutiert. Leider führte eine genauere Betrachtung zu wachsenden Problemen. Es gab unter anderem viele Diskussionen, was denn der Wert der unendlichen Summe

$$1 + (-1) + 1 + (-1) + \ldots$$

sei, manche Gelehrte vertraten die Ansicht es sei 0: mit der Begündung

$$(1 + (-1)) + (1 + (-1)) + (1 + (-1)) + \ldots = 0 + 0 + 0 + \ldots = 0.$$

Dies erscheint überzeugend. Mit derselben Logik kann man aber auch begründen, dass man den Wert 1 erhält:

$$1 + ((-1) + 1) + ((-1) + 1) + \ldots = 1 + 0 + 0 + \ldots = 1.$$

Derartige Probleme motivierten die Mathematiker, die Mathematik auf solide Grundlagen zu stellen. Diese Bewegungen, die man Axiomatik nennen kann, begann im Bereich der Geometrie bereits mit Euklid von Alexandria (ca. 300 v. Chr.) und Aristoteles (384–322 v. Chr.). Wichtige Fortschritte in der Axiomatik der Geometrie und insgesamt der Axiomatik wurden im 19. Jahrhundert vollbracht.

Das Ziel der Axiomatik ist es, die gesamte Mathematik aus wenigen Grundaussagen, sogenannten Axiomen, herzuleiten.

Nehmen wir mal an, ich überlege mir, ob ich überhaupt existiere oder nicht. Mein Leben und meine Person könnte ja auch nur das Ergebnis einer Simulation eines gigantischen Großrechners sein. Was könnte ich tun, um diese Frage zu lösen? Ich kann mir selbst wehtun, ich empfinde Schmerz, also sollte ich existieren! Falsch, das könnte Teil der gigantischen Simulation sein. Ich könnte einen Studenten in der hinteren Reihe fragen, ob ich existiere, er sagt ja. Ist dadurch meine Existenz bewiesen? Nein, denn er könnte Teil derselben Simulation sein. Sie sehen, ich kann meine eigene Existenz nicht zeigen, ohne andere Grundannahmen zu machen. Dennoch ist es sehr sinnvoll anzunehmen, dass ich existiere. Solche Probleme beschäftigen die Philosophen schon seit Jahrhunderten oder besser Jahrtausenden.

Genauso wie aber unser elementares Denken Grundannahmen braucht, benötigt auch die Mathematik bestimmte Grundannahmen, und diese Grundannahmen nennt man Axiome. Im Prinzip sollte in der heutigen Mathematik alles auf den Axiomen der Mengenlehre

aufbauen. Daraus konstruiert man sich dann alle Objekte des mathematischen Denkens: Zahlen, Vektorräume, Matrizen, und vieles mehr. Oft versieht man Teilgebiete mit eigenen Axiomen, wie zum Beispiel die Axiome der klassischen Geometrie.

Axiome sind nicht mehr weiter beweisbar, sie werden einfach als gegeben hingenommen, als Grundannahmen unseres Denkens. Die Axiome erscheinen sinnvoll, entweder weil man sie als evident, also offensichtlich ansieht, oder weil man sie als Kennzeichen der Theorie ansieht. Aus diesen Axiomen werden dann Schlussfolgerungen gezogen, die ebenfalls durch weitere Axiome geregelt sind. Durch erlaubte Kombinationen von bereits bekannten wahren Aussagen erhält man neue wahre Aussagen. Eine Sammlung von so aufeinander aufbauenden wahren Aussagen, nennt man Beweis. Falls eine so erhaltene Aussage interessant erscheint, nennt man sie Theorem, Lemma, Korollar, Proposition, Satz, Hilfssatz, Folgerung oder ähnlich. Hierbei ist im allgemeinen ein Satz oder ein Theorem eine wichtige Aussage, ein Lemma oder ein Hilfssatz eine Aussage, die nur als Zwischenschritt dient, und eine Proposition hat eine Mittelstellung. Erhält man eine Aussage nahezu unmittelbar aus einem Theorem oder Satz, so nennt man dies eine Folgerung oder ein Korollar.

Damit die Aussagen nicht immer länger und länger werden, macht man Definitionen. Hierbei gibt man mathematischen Objekten oder mathematischen Sachverhalten einen Namen.

Die Mathematiker sind im Prinzip recht frei in der Wahl ihrer Definitionen. So könnte man die folgende Definition machen: Ein Auto ist eine Menge, in der die Elemente 1, 2 und 3 enthalten sind. Ein Hund ist eine Menge, in der die Elemente 1 und 2 enthalten sind. Man schließt daraus, dass jedes Auto einen Hund enthält. Diese Definitionen sind natürlich sehr irreführend, aber prinzipiell erlaubt. Wir Mathematiker bemühen uns die

Dinge so zu benennen, dass sie möglichst etwas mit der „wirklichen Welt“ zu tun haben. Um Koordinaten auf einer Kugel anzugeben, definiert man den Begriff einer „Karte“, und ein „Atlas“ ist dann definiert als Menge von Karten, so dass alles überdeckt wird. Diese Begriffe sind dann zwar nicht genau das, was man damit umgangssprachlich meint, aber auch nicht völlig ohne Zusammenhang. Die mathematischen Begriffe „Halm“ oder „Garbe“ der Mathematik haben aber keinerlei Anwendungen in der Landwirtschaft. Die „Knoten“ der Mathematik sind aber wiederum nahe an dem, was man alltagssprachlich als Knoten bezeichnet.

Um den Unterschied zwischen Definitionen und Aussagen klar zu unterscheiden, nutzen wir die folgende Notation: $a = 1,234$ ist die Aussage „a ist gleich $1,234$“. Hingegen ist $a := 1,234$ eine Definition, a ist ab sofort eine kurze Schreibweise für $1,234$.

Viele Definitionen werden von allen Mathematikern gleich gemacht, es herrscht Konsens. Man ist sich aber nicht einig, ob die Definition der natürlichen Zahlen die Null einschließen soll oder nicht. Für unsere Vorlesung gilt: die natürlichen Zahlen sind

$$\mathbb{N} = \{0, 1, 2, \ldots\}.$$

Die Menge $\mathbb{N}_{>0} := \{1, 2, 3, \ldots\}$ bezeichnen wir als natürliche Zahlen ohne Null. Dies ist eine der üblichen Definitionen. Viele Mathematiker definieren hingegen die Menge der natürlichen Zahlen als $\{1, 2 \ldots\}$. Dass verschiedene Mathematiker die Definition verschieden wählen, ist nicht weiter schlimm. Ob Null eine natürliche Zahl ist oder nicht, ist eben Definitionssache. Man schaut sich die Definition des Autors an und weiß, was er meint.

In zentralistischen Ländern wie Frankreich ist klar geregelt: Null ist eine natürliche Zahl. In Deutschland besagt DIN 5473 ebenfalls, dass Null eine natürliche Zahl ist. Lehrer in der Schule sollten sich an diese DIN-Norm halten. Außer man wohnt in Bayern. An

bayerischen Schulen ist Null keine natürliche Zahl, und daran sollte man sich als Lehrer auch halten, um die Schüler nicht zu verwirren.

Da die meisten Hörer dieser Vorlesung ja gleichzeitig Lineare Algebra bei Denis-Charles Cisinski hören, und er Franzose ist, ist bei uns Null eine natürliche Zahl. Die zukünftigen bayerischen Lehrer mögen dies bitte verzeihen und freuen sich vielleicht über diese „Zusatzqualifikation“ für andere Bundesländer.

Anders ist es bei der Zahl π. Dass der Wert dieser Zahl zwischen 3,1415 und 3,1416 liegt ist eine Aussage und keine Definition. Deswegen ist es lächerlich, dass der US-Bundesstaat Indiana 1897 den Wert von π auf $3,2$ gesetzlich festlegen wollte, um Berechnungen zu vereinfachen und um es den Schülern einfacher zu machen.

Wenn wir nun aber hier in der Vorlesung alles rigoros mit der eigentlich notwendigen mathematischen Strenge einführen würden, so müssten wir dem Aufbau der Mathematik folgen. Die Grundlagen hierbei bilden die Axiome der Logik und der Mengenlehre. Aus ihnen werden viele Aussagen gezeigt, und man kann dann irgendwann die natürlichen Zahlen, die ganzen Zahlen, die rationalen Zahlen und die reellen Zahlen einführen[1] und deren Eigenschaften studieren. Die reellen Zahlen erfüllen auch wieder einige wichtige Aussagen, die wir vorläufig die Grundeigenschaften der reellen Zahlen nennen wollen. Aus diesen Grundeigenschaften der reellen Zahlen kann man nun alle wichtigen Aussagen der Analysis herleiten. Nahezu alles, was wir zusammen in den nächsten Semestern behandeln werden, ergibt sich aus diesen Grundeigenschaften.

[1] mathematisch genauer gesagt: definieren

Nun gibt es aber ganz verschiedene Möglichkeiten, die reellen Zahlen zu definieren: in der Schule habe ich die Möglichkeit mit Intervall-Schachtelung kennengelernt, in der Vorlesung hier wollen wir Dedekindsche Schnitte behandeln, vor 5 Jahren in meiner Analysis I habe ich sie mit Hilfe von Cauchy-Folgen definiert. Jeder Zugang hat seine Vorteile und Nachteile in Bezug auf Anschaulichkeit und Verallgemeinerbarkeit. Wichtig ist hierbei: egal, wie man es macht, es ergeben sich immer die selben Grundeigenschaften. Deswegen ist es sinnvoll, die Grundeigenschaften der reellen Zahlen als die Axiome der reellen Zahlen zu bezeichnen und dann alles aus diesen Axiomen der reellen Zahlen herzuleiten. Genauso haben auch die natürliche Zahlen ihre Axiome, die sogenannten Peano-Axiome[2]. Alle von Ihnen, die gleichzeitig die Lineare Algebra I besuchen, lernen dort die Axiome eines Vektorraums kennen und leiten daraus dann alle Eigenschaften ab, die Sie über Vektorräume wissen sollten.[3]

In der Vorlesung können wir nun aber leider nicht alles streng axiomatisch einführen, da dies viel zu lange dauern würde und die meisten von Ihnen auch nicht die nötige Geduld dafür aufbringen würden – vielleicht nicht einmal ich. Wir müssen also zu Anfang einen Kompromiss zwischen der nötigen Strenge und der angemessenen Kürze finden.

Der Plan ist deswegen, zunächst Begriffe wie Mengen, Abbildungen, logische Operationen und viele ähnliche Begriffe nur intuitiv einzuführen, also ohne die richtige mathematische Strenge.

Im Anhang A im Skript gebe ich einen kleinen Einblick, wie man die Mengenlehre mathematisch stringenter einführen kann. Die zugehörigen Abschnitte A.1 und A.2 werde ich in

[2]siehe Anhang B

[3]Das Wort „Sie“ ist bewusst groß geschrieben, da ich der Meinung bin, dass alle hier im Publikum dies beherrschen sollten.

der Vorlesung nicht behandeln, aber ich empfehle ihn allen interessierten Studierenden zur Lektüre. Wenn Sie die Mengenlehre noch besser axiomatisch verstehen wollen, so ist es am besten, wenn Sie begleitend sich etwas selbst durcharbeiten, z.B. das Buch [**17**], aber auch da werden letztendlich Fragen zurückbleiben. Wenn Sie die Mengenlehre „richtig“ verstehen wollen, sollten Sie später mit einer gewissen mathematischen Reife, also in zwei oder drei Semester, die Logik nochmals systematisch studieren, z.B. an Hand der Bücher [**13**] oder [**20**].

Nachdem das einführende Kapitel beendet ist, wenden wir uns den Zahlen zu. Wir beschreiben grundlegende Eigenschaften der natürlichen, ganzen, rationalen, reellen und komplexen Zahlen. Auch da werden wir in Teilen nicht streng mathematisch vorgehen. Ich gehe zum Beispiel davon aus, dass Sie intuitiv wissen, was die natürlichen, ganzen und rationalen Zahlen sind. Eigentlich wäre es aber auch wichtig, zum Beispiel die natürlichen Zahlen axiomatisch durch die Peano-Axiome (Anhang B) zu beschreiben. Es ist nämlich zunächst nicht so recht klar, was denn die Punkte in der Definition $\mathbb{N} = \{0, 1, 2, \ldots\}$ bedeuten soll. Dies soll in der Linearen Algebra I genauer behandelt werden.

Wir werden dann zu den reellen Zahlen kommen. Da diese für die Analysis ganz zentral sind, werden wir die Axiome diskutieren und von da an alles gründlich auf den Axiomen der reellen Zahlen aufbauen: Folgen, Reihen, Differentiation, Integration und vieles mehr.

Wichtig ist auch, dass Sie sich bewusst werden, was Sie in den nächsten Wochen alles lernen sollten. Den mathematischen Inhalt der Vorlesung sollten Sie durchdringen, damit umgehen lernen und danach nie wieder vergessen. Viel wichtiger aber noch ist, dass die meisten von Ihnen Ihre gewohnte Arbeitsweise umstellen müssen. Es geht dabei beim weitem nicht nur darum, dass Sie von nun an in vieler Hinsicht für Ihr Lernen und Ihre

Lernmethoden selbst verantwortlich sind, sondern vor allem darum, dass Sie das oben diskutierte rigorose Argumentieren lernen.

Sie müssen klar unterscheiden können zwischen Axiomen, Definitionen und Aussagen (Sätzen, Theoremen, Lemmata, Propositionen,...). Sie müssen lernen, Beweise zu verifizieren. Und Sie müssen lernen, Ideen zu bekommen, wie Sie selbst Beweise führen, und auch lernen, wie Sie dann diesen Beweis gut aufschreiben. Dies zu lernen, ist nur durch eine intensive Rückmeldung möglich, und das geht natürlich nicht in der großen Vorlesung hier. Deswegen sind die Übungen für Sie ganz wichtig.

2. Aussagenlogik

Was ist eine Aussage?

Beispiele 2.1. Die folgenden Ausdrücke sind Aussagen

A_1: $2 * 3 = 6$

A_2: $2 + 2 = 1 + 3$

A_3: Die Zahl 27 hat 4 Teiler.

A_4: Alle natürliche Zahlen haben eine Primfaktor-Zerlegung

A_5: Am 13.9.2018 hatte die Donau in Regensburg Hochwasser

A_6: Christian Lindner wurde im September 2018 zum Bundeskanzler gewählt

Wir gehen davon aus, dass wir eine Sprache haben, die aus Zeichenketten besteht. Manche Zeichenketten ergeben keinen Sinn, zum Beispiel „hejekl“ oder „Hund Maus Loch“, wir

nennen sie syntaktisch nicht sinnvoll. Wenn die Zeichenkette etwas aussagt, so nennen wir sie syntaktisch sinnvoll. [4]

DEFINITION 2.2. Eine Aussage ist eine syntaktisch sinnvolle Zeichenkette, die entweder wahr (w) oder falsch (f) ist.

Eine dritte Möglichkeit gibt es nicht (lateinisch: „tertium non datur“).

A_1 bis A_6 sind Aussagen, A_1 bis A_4 sind wahr, A_5 und A_6 sind falsch. Wenn eine Aussage wahr ist, so sagen wir auch: „Die Aussage gilt.“

BEMERKUNG 2.3. Es gibt Wissenschaftsbereiche, in den das „tertium non datur“ nicht gilt, in denen also neben „wahr“ und „falsch“ weitere Möglichkeiten zugelassen werden. Z.B. in der Quantenmechanik (Schrödingers Katze ist weder tot noch lebendig), Aussagen in der Wahrscheinlichkeitstheorie (Es regnet morgen mit Wahrscheinlichkeit 0,3), Philosophie, Teilgebiete der Logik. In unserer Vorlesung sind aber gemäß obiger Definition nur Aussagen zulässig, die entweder wahr oder falsch sind.

Man kann Aussagen durch elementare logische Operationen verknüpfen. Man kann diese Operationen zum Beispiel durch Wahrheitstafeln definieren.

DEFINITION 2.4 (Negation). Die Negation $\neg$ wird durch die folgende Wahrheitstafel definiert

A	$\neg A$
w	f
f	w

[4] Dies soll hier nicht genauer definiert und spezifiziert werden, da es für unsere Zwecke unwichtig ist.

Lemma 2.5. Für alle Aussagen A gilt $\neg(\neg A) = A$

Beweis. Wir unterscheiden zwei Fälle.

1. Fall: A ist wahr

A ist wahr $\implies$ $\neg A$ ist falsch $\implies$ $\neg(\neg A)$ ist wahr

2. Fall: $A = f$

A ist falsch $\implies$ $\neg A$ ist wahr $\implies$ $\neg(\neg A)$ ist falsch □

Definition 2.6 (Aussagenlogische Verknüpfungen). Die Und-Verknüpfung $\wedge$, die Oder-Verknüpfung $\vee$, die Entweder-Oder-Verknüpfung $\veebar$, die Implikation (Wenn-Dann-Beziehung) $\rightarrow$, die umgekehrte Implikation (Dann-Wenn-Beziehung[5]) $\leftarrow$ und die Äquivalenz (Genau-Dann-Wenn-Beziehung) $\leftrightarrow$ sind durch die folgende Wahrheitstafel definiert

A	B	$A \wedge B$	$A \vee B$	$A \veebar B$	$A \rightarrow B$	$A \leftarrow B$	$A \leftrightarrow B$
w	w	w	w	f	w	w	w
w	f	f	w	w	f	w	f
f	w	f	w	w	w	f	f
f	f	f	f	f	w	w	w

Ist die Aussage $A \leftrightarrow B$ wahr, so sagen wir auch A und B sind äquivalent.

In zusammengesetzen Ausdrücken sind Negationen zuerst auszuführen, ansonsten muss man durch Klammern die Reihenfolge klären sofern nötig. Zum Beispiel gilt

$$\neg A \wedge B = (\neg A) \wedge B$$

[5]Beispiel: (Die Straße wird nass.) $\leftarrow$ (Es regnet.). In Worten: Die Straße wird (dann) nass, wenn es regnet. Man kann logisch äquivalent auch sagen: Es regnet nur dann, wenn die Straße nass wird.

und dies ist im allgemeinen nicht dasselbe wie $\neg(A \wedge B)$. Man darf aber die Klammern weglassen, wenn das Resultat nicht von der Reihenfolge abhängt, Beispiele später. **19.10.**

Bisher haben wir nun angenommen A, B seien feste Aussagen. Oft steht der Wahrheitswert von einem Ausdruck noch gar nicht fest, z.B.

A: Am 1.2.2024 wird es ein Erbeben in Japan geben

Um auch solche Ausdrücke behandeln zu können, deren Wahrheitswert noch nicht festgelegt ist, führen wir nun Aussageformen ein.

Definition 2.7 (Ausageform). Eine Aussageform ist eine eine syntaktisch sinnvolle Zeichenkette, die

- entweder eine Aussage ist, oder
- die von einer oder mehreren Variablen abhängt und die erst nach Einsetzen von Werten in diese Variablen zu einer Aussage wird.

Hierbei muss man angeben, welche Werte für die Variablen überhaupt zugelassen sind.

Eine Aussageform $A(x)$ in einer Variablen x besteht also aus der Angabe der zulässigen Werte von x, die wir zu einer Menge M zusammenfassen, und einer Abbildung $A : M \longrightarrow \{w, f\}$, $x \mapsto A(x)$. Letztere Schreibweise bedeutet, dass jedem zulässigen Wert von x ein Wert $A(x)$ zugeordnet ist, der entweder wahr (w) oder falsch (f) ist.[6] Hängt

[6]Der Pfeil $\longrightarrow$ in der Abbildungsbeschreibung ist nicht zu verwechseln mit der Implikation $\rightarrow$, deswegen wollen wir den ersteren zunächst länger und in roter Farbe schreiben. Später wird immer klar sein, welche Bedeutung der Pfeil nun hat und wir schreiben dann immer $\rightarrow$. Wir setzen hier voraus, dass die Begriffe „Menge“, „Produkte von Mengen“ und „Abbildung“ Ihnen bereits intuitiv vertraut sind, wenn nicht dann sind die Erklärungen dieses Abschnitts erst in ungefähr zwei Wochen verständlich.

$A(x_1, x_2, \ldots)$ von mehreren Variablen hat, so sollte eine Wertemenge M_1 für x_1, eine Wertemenge M_2 für x_2 etc. vorgeben werden. Aussageformen in 2 Variablen sind also Abbildungen:

$$A : M_1 \times M_2 \longrightarrow \{w, f\}, \quad (x_1, x_2) \mapsto A(x_1, x_2).$$

BEISPIELE 2.8.

$A_1(x)$:	$x > 0$	Zulässig: $x \in \mathbb{R}$
$A_2(X)$:	Das Auto X ist rot	Zulässig: Alle Autos X der Welt
$A_3(B, C)$:	$B \vee C$	Zulässig: B und C sind Aussagen.
$A_4(B, C, D, E)$:	$(B \vee C) \to (\neg D \wedge E)$	Zulässig: B, C, D und E sind Aussagen.

Sind wahr und falsch die zulässigen Werte der Variablen x, so nennt man x eine aussagenlogische Variable. Verknüpft man ausagelogische Variablen durch die Negation $\neg$ oder die logischen Verknüpfungen $\wedge$, $\vee$, $\veebar$, $\to$, $\leftarrow$ und $\leftrightarrow$, so nennt man die somit erhaltene Aussageform eine aussagenlogische Formel . Es ist hierbei erlaubt, mehrere logische Verknüpfungen oder Negationen anzuwenden. Insbesondere sind dann alle Variablen aussagenlogische Variablen. In den obigen Beispielen sind A_3 und A_4 aussagenlogische Formeln, die anderen keine.

Für aussagenlogische Formeln kann man viele Beziehungen herleiten, zum Beispiel folgende:

PROPOSITION 2.9. Für alle Belegungen von A, B und C mit Wahrheitswerten w und f sind die folgenden Aussagen wahr:

(a) $A \vee \neg A$ (Tertium non datur)

(b) $\neg(A \wedge \neg A)$ (Widerspruchsfreiheit)

(c) $A \vee w$

(d) $\neg(A \wedge f)$

(e) $(A \wedge f) \leftrightarrow f$

(f) $(A \vee f) \leftrightarrow A$

(g) $(A \wedge w) \leftrightarrow A$

(h) $(A \vee w) \leftrightarrow w$

(i) $(A \vee B) \leftrightarrow (B \vee A)$ (Kommutativität von $\vee$)

(j) $(A \wedge B) \leftrightarrow (B \wedge A)$ (Kommutativität von $\wedge$)

(k) $(A \vee (B \vee C)) \leftrightarrow ((A \vee B) \vee C)$ (Assoziativität von $\vee$)

(l) $(A \wedge (B \wedge C)) \leftrightarrow ((A \wedge B) \wedge C)$ (Assoziativität von $\wedge$)

(m) $(A \vee B) \leftrightarrow \neg(\neg A \wedge \neg B)$ (de Morgansche Regeln)

(n) $(A \wedge B) \leftrightarrow \neg(\neg A \vee \neg B)$ (de Morgansche Regeln)

(o) $(A \wedge (B \vee C)) \leftrightarrow ((A \wedge B) \vee (A \wedge C))$ (Distributivgesetze)

(p) $(A \vee (B \wedge C)) \leftrightarrow ((A \vee B) \wedge (A \vee C))$ (Distributivgesetze)

(q) $(A \leftrightarrow w) \leftrightarrow A$

(r) $(A \leftrightarrow f) \leftrightarrow \neg A$

(s) $(A \rightarrow B) \leftrightarrow (B \leftarrow A)$

(t) $(A \rightarrow B) \leftrightarrow (\neg A \vee B)$ (Kontraposition)

(u) $(A \leftarrow B) \leftrightarrow (A \vee \neg B)$

(v) $(A \rightarrow B) \leftrightarrow (\neg B \rightarrow \neg A)$

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

(x) $(A \leftrightarrow B) \leftrightarrow \neg(A \veebar B)$

(y) $(A \wedge B) \rightarrow A$

Diese Proposition wird teilweise in den Übungsgruppen bewiesen, die verbleibenden können Sie selbst mit ähnlichen Methoden leicht beweisen. Die obigen Äquivalenzen, d.h. die obigen Aussagen in der Form $D \leftrightarrow E$, beweist man oft am besten, in dem man mit Hilfe einer Wahrheitstabelle D und E für alle Belegungen von A, B und C berechnet und dann vergleicht. Nachdem man einige Aussagen gezeigt hat, kann man durch deren Kombination weitere erhalten.

Aus (l) folgt zum Beispiel, dass $A \wedge B \wedge C$ ein syntaktisch sinnvoller Ausdruck ist, da das Ergebnis nicht von der Klammerung abhängt, und analog dazu ist natürlich auch $A \vee B \vee C$ ein sinnvoller Ausdruck.

Definition 2.10. Wir nennen zwei Aussageformen äquivalent, wenn sie von denselben Variablen abhängen, die Variablen dieselbe zulässigen Werte haben und wenn beide Aussageformen für alle Belegungen der Variablen mit Werten äquivalente Aussagen ergeben.

Somit sind also die Aussageformen D und E äquivalent genau dann, wenn $D \leftrightarrow E$ für alle Belegungen der Variablen wahr ist.

Beispiel 2.11. Proposition 2.9 (i) besagt, dass $A \vee B$ äquivalent zu $B \vee A$ ist.
Proposition 2.9 (x) besagt, dass $A \leftrightarrow B$ äquivalent zu $\neg(A \veebar B)$ ist.

Lemma 2.12. Die aussagenlogischen Formeln $E_1(A, B, C) := A \to (B \to C)$ und $E_2(A, B, C) := (A \to B) \to C$ sind nicht äquivalent. (Präziser gesagt: Es ist nicht richtig, dass für alle Wahlen von A, B und C die aus E_1 und E_2 erhaltenen Aussagen äquivalent sind).

Deshalb ist also $A \to B \to C$ kein syntaktisch sinnvoller Ausdruck, man muss hier Klammern setzen.

Beweis. Angenommen, die aussagenlogischen Formeln E_1 und E_2 wären für alle A, B und C äquivalent. Dann wären sie insbesondere äquivalent im Fall, dass A, B und C falsch sind. In diesem Fall sind $A \to B$ und $B \to C$ wahr und somit ist E_1 wahr und E_2 falsch. Somit haben wir einen Widerspruch zur obigen Annahme erhalten. Die Annahme war also falsch. Also ist das Lemma bewiesen. □

Der obige Beweis ist ein Widerspruchsbeweis. Um eine Aussage F zu zeigen, nimmt man zunächst $\neg F$ an und leitet daraus einen Widerspruch her.

Hierbei ist

F: Für alle Wahlen von A, B und C sind die Aussagen $E_1(A, B, C)$ und $E_2(A, B, C)$ äquivalent.

Ein anderer Typ von Widerspruchsbeweis ist noch häufiger. Man möchte eigentlich die Aussage $E \to F$ zeigen. Man nimmt nun $\neg F$ an und leitet daraus $\neg E$ her. Man hat somit $\neg F \to \neg E$ gezeigt. Dies ist aber mit Proposition 2.9 (v) äquivalent zu $E \to F$.

Bemerkung 2.13. Die Symbole $\Longleftrightarrow$ und $\Longrightarrow$ nutzen wir ähnlich wie die Symbole $\leftrightarrow$ und $\to$, jedoch mit kleinen Unterschieden:

- Wenn wir aussagenlogische Variablen in einer aussagenlogischen Formel zusammenfügen, so wenden wir zuerst $\leftrightarrow$, $\to$, $\leftarrow$, $\wedge$,... und dann erst $\Longleftrightarrow$ und $\Longrightarrow$ an. Beispiel: Die Aussageform

$$3 + 4 = 7 \Longleftrightarrow A \vee \neg A$$

ist als $(3 + 4 = 7) \Longleftrightarrow (A \vee \neg A)$ zu lesen.

- Wir lesen Implikationsketten wie zum Beispiel $A \Longrightarrow B \Longrightarrow C$ als „(A impliziert B) und (B impliziert C)“. Als Beispiel betrachte man den Beweis von Lemma 2.5. Der Ausdruck $A \to B \to C$ ist hingegen syntaktisch gar nicht sinnvoll, da Klammern benötigt werden. Analog sind Äquivalenzketten zugelassen und zu interpretieren, zum Beispiel:

$$\begin{aligned}
\neg(A \vee \neg B) \wedge A &\overset{\text{„(n)“}}{\Longleftrightarrow} (\neg A \wedge B) \wedge A \\
&\overset{\text{(l)}}{\Longleftrightarrow} \neg A \wedge (B \wedge A) \\
&\overset{\text{(j)}}{\Longleftrightarrow} \neg A \wedge (A \wedge B) \\
&\overset{\text{(l)}}{\Longleftrightarrow} (\neg A \wedge A) \wedge B \\
&\overset{\text{„(b)“}}{\Longleftrightarrow} \mathrm{f} \wedge B \\
&\overset{\text{„(e)“}}{\Longleftrightarrow} \mathrm{f}
\end{aligned}$$

  Bei jedem $\Longleftrightarrow$ findet eine Äquivalenzumformung statt, die direkt aus bereits bekannten Aussagen folgt. Der jeweilige Buchstabe über einem der $\Longleftrightarrow$-Pfeile gibt an, welche Eigenschaft in Proposition 2.9 verwendet wurde. Die Anführungszeichen in der ersten und den beiden letzten Zeilen deuten an, dass man leicht modifizierte Versionen der zitierten Punkte hernehmen muss. Die Verwendung des Zeichens $\Longleftrightarrow$ wäre natürlich auch ohne die Erklärungen darüber gestattet. Die obige Umformungskette ist somit ein Beweis[7] der Tatsache, dass die Aussageform $\neg(A \vee \neg B) \wedge A$ für alle Belegungen von A und B mit Wahrheitswerten falsch ist.

[7]Die Worte „beweisen“ und „zeigen“ haben in der Mathematik nahezu dieselbe Bedeutung.

Eine ähnliche Verwendung wird im folgenden Beispiel verdeutlicht:

$$\begin{aligned}\neg(A \leftrightarrow B) &\iff \neg A \leftrightarrow B \\ &\iff A \leftrightarrow \neg B \\ &\iff A \veebar B \\ &\iff \neg A \veebar \neg B\end{aligned}$$

Die obigen Zeilen sind eine recht effektive Art und Weise auszudrücken, dass alle obigen Ausdrücke paarweise äquivalent sind. Allerdings muss man hier „etwas arbeiten“ um jeweils von der linken Seite von $\iff$ zur rechten zu kommen. Dies geht in diesem Beispiel mit einer Wahrheitstafel sehr gut.

- Um Missverständnisse zu vermeiden, verbieten wir es, Ausdrücke, die $\iff$ und $\implies$ enthalten, durch weitere aussagenlogische Verknüpfungen zu verbinden. Das heißt beispielsweise, dass wir $A \iff (B \implies C)$ als syntaktisch nicht sinnvoll betrachten, also vergleichbar zu „jjkwdh“.

Beispiel: Proposition 2.9 (w) und (x) ergeben dann zusammen

$$A \leftrightarrow B \iff (\neg A \vee B) \wedge (A \vee \neg B) \iff \neg(A \veebar B).$$

Bemerkung 2.14. Man könnte zunächst auch nur $\neg$ und $\wedge$ durch eine Wahrheitstafel wie oben definieren und alle danach anderen aussagenlogischen Verknüfungen durch Formeln aus Proposition 2.9, z.B. $A \vee B := \neg(\neg A \wedge \neg B)$ durch die de Morgansche Regel.

3. Mengen

Beschreibung 3.1 (Georg Cantor, 1845–1918). Unter einer Menge verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche Elemente von M genannt werden) zu einem Ganzen.

Hier wird erklärt, was eine Menge ist. Auch eine Definition erklärt, was ein neuer Begriff bedeutet. Dennoch ist die obige Erklärung keine mathematische Definition, sondern eine intuitive Erklärung. Wenn man in der Mathematik sauber definieren will, was eine Menge ist, darf man nur Worte und Strukturen benutzen, die zuvor schon ein klare Bedeutung haben. Es ist nun aber doch etwas unklar, was die Worte „Zusammenfassung“, “bestimmten“, “wohlunterschieden“ bedeuten.

Die obige „Beschreibung“ sagt uns also lediglich, was wir uns unter einer Menge vorstellen sollen. Das reicht uns aber für das folgende, und eine genaue Definition wäre zu aufwändig.

Eine Menge enthält Elemente. Zwei Mengen sind genau dann gleich, wenn sie die gleichen Elemente enthalten.

Schreibweisen:
$M = \{x, y, z\} = \{x, x, y, z\} = \{y, x, z\}$ ist die Menge, deren Elemente x, y und z sind.

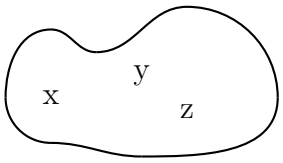
Man darf Elemente in den Klammern $\{\quad\}$ mehrfach aufzählen, sie sind dann aber nur einmal in der Menge. Ein Element kann in einer Menge sein, oder nicht darin sein, es kann aber nicht mehrfach in M enthalten sein.

$x \in M$ ist die Kurzschreibweise für „x ist ein Element von M“
$x \notin M$ ist die Kurzschreibweise für „x ist kein Element von M“

$$x \notin M \quad \Longleftrightarrow \quad \neg(x \in M)$$

Beispiele 3.2.

(1) Vorstellung von $M = \{x, y, z\}$



(2) Die Menge aller natürlichen Zahlen: $\mathbb{N} = \{0, 1, 2, 3, \ldots, \}$.
Übliche Formulierung: Die Menge $\mathbb{N}$ der natürlichen Zahlen.

(3) Die Menge $\mathbb{R}$ der reellen Zahlen [8]

(4) Die Menge $\mathbb{R}_{>0}$ der positiven reellen Zahlen. (Die Null ist weder positiv noch negativ).[9]

(5) Ist $A(x)$ eine Aussageform in einer Variablen x, so bilden die zulässigen Werte von x eine Menge, die Definitionsmenge der Aussageform. In Beispiele 2.8 ist die Definitionsmenge von A_2 die Menge aller Autos. Die Definitionsmenge von A_1 ist $\mathbb{R}$.

---

[8] Achtung: Was eine reelle Zahl ist, müssen wir eigentlich noch zuerst lernen!

[9] Sehr verbreitet ist auch die Schreibweise $\mathbb{R}_+^*$ für die positiven reellen Zahlen, und dann schreibt man $\mathbb{R}_+ := \mathbb{R}_+^* \cup \{0\}$ für alle reellen Zahlen, die nicht negativ sind. In dieser Schreibweise soll $+$ natürlich positiv bedeuten. Deswegen ist auch die Schreibweise $\mathbb{R}_+$ verwirrend, denn wir haben $0 \in \mathbb{R}_+$, obwohl 0 nicht positiv ist. Ein dritte Möglichkeit, die man gelegentlich sieht, ist $\mathbb{R}^+$ für die positiven reellen Zahlen zu benutzen, und dann definiert man $\mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}$. Um Verwirrung zu vermeiden, nutzen wir nur $\mathbb{R}_{>0}$ für die Menge der positiven reellen Zahlen und $\mathbb{R}_{\geq 0}$, wenn die Null noch dazugenommen wird.

DEFINITION 3.3. Die Menge, die keine Elemente hat, heißt die leere Menge, und wird notiert als $\emptyset$ oder $\{\}$.

!ACHTUNG!. $\emptyset \neq \{\emptyset\}$ und $1 \neq \{1\} \neq \{\{1\}\}$.

Bei Aussagen und aussagenlogischen Formeln schreiben wir ab sofort $:\Longleftrightarrow$ für „ist definiert als“.

DEFINITION 3.4 (Teilmenge).

(1) Die Menge N ist eine Teilmenge von M, genau dann wenn jedes Element von N auch ein Element von M ist. Als Formel schreiben wir $N \subset M$.
(2) Die Menge N ist eine echte Teilmenge von M, genau dann wenn N eine Teilmenge M ist und von M verschieden ist. Wir schreiben hierfür $N \subsetneq M$.

BEISPIELE 3.5.

(1) $\{1,2\} \subset \{1,2,3\}$ und $\{1,2\} \subsetneq \{1,2,3\}$.
(2) Für alle Mengen M gilt: $\emptyset \subset M$
(3) Für alle Mengen M gilt: $M \subset M$

24.10.

DEFINITION 3.6. Ist A eine auf M definierte Aussageform, so definieren wir

$$\{x \in M \mid A(x)\}$$

als die Teilmenge aller Elemente x von M, für die $A(x)$ wahr ist.

Es gilt somit für alle $y \in M$:

$$y \in \{x \in M \mid A(x)\} \quad \Longleftrightarrow \quad A(y)$$

Umgekehrt: Ist N eine Teilmenge von M, so ist

$$A(x) := \begin{cases} \text{w} & \text{falls } x \in N \\ \text{f} & \text{falls } (x \in M) \wedge (x \notin N) \end{cases}$$

eine auf M definierte Aussageform und es gilt dann

$$N = \{x \in M \mid A(x)\}.$$

Beispiele 3.7.

(1) Seien A_1 und A_2 definiert wie in Beispiele 2.8, M_2 sei die Menge aller Autos. Dann gilt

$$\{x \in \mathbb{R} \mid A_1(x)\} = \{x \in \mathbb{R} \mid x > 0\} = \mathbb{R}_{>0}$$

und

$$\{x \in M_2 \mid A_2(x)\}$$

ist die Menge aller roten Autos.

(2) Die Menge der geraden natürlichen Zahlen ist

$$2\mathbb{N} := \left\{n \in \mathbb{N} \;\middle|\; \frac{n}{2} \in \mathbb{N}\right\}.$$

Sei $k \in \mathbb{N}_{>0}$. Die Menge der durch k teilbaren ganzen Zahlen ist

$$k\mathbb{Z} := \left\{n \in \mathbb{Z} \;\middle|\; \frac{n}{k} \in \mathbb{Z}\right\}.$$

Analog $k\mathbb{N}$.

(3) Die Menge der Primzahlen ist

$$\{n \in \mathbb{N} \mid n \text{ besitzt genau zwei Teiler (in } \mathbb{N})\}.$$

Wir führen somit noch eine weitere Art und Weise ein, Mengen zu beschreiben.

DEFINITION 3.8. Sei M eine Menge und $f(x)$ ein von x abhängiger Ausdruck[10]. Dann bezeichne

$$\{f(x) \,\|\, x \in M\}$$

die Menge aller Objekte, die sich in der Form $f(x)$ für ein $x \in M$ schreiben lassen.[11] Analoges gilt auch wenn f mehrere Variablen besitzt, zum Beispiel

$$\{f(x,y) \,\|\, x \in M,\ y \in N\},$$

wenn f von den Variablen $x \in M$ und $y \in N$ abhängt.

Wir können dann $k\mathbb{Z}$ auch einführen als

$$k\mathbb{Z} := \{kz \,\|\, z \in \mathbb{Z}\}.$$

DEFINITION 3.9. Seien M und N Mengen. Wir definieren hieraus neue Mengen

(a) den Schnitt (oder die Schnittmenge) von M und N als

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

Man kann auch sagen, $M \cap N$ ist die Menge, so dass gilt:

$$x \in M \cap N \quad \Longleftrightarrow \quad (x \in M) \wedge (x \in N).$$

(b) die Vereinigung(smenge) von M und N als die Menge $M \cup N$, so dass

$$x \in M \cup N \quad :\Longleftrightarrow \quad (x \in M) \vee (x \in N).$$

[10]Die mathematisch präzisere Voraussetzung wäre: f ist eine Funktion in einer Variablen x.

[11]In einigen Wochen vereinfachen wir diese Schreibweise zu $\{f(x) \mid x \in M\}$.

(c) das Komplement von N in M als

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

oder dazu äquivalent

$$x \in M \setminus N \quad \Longleftrightarrow \quad (x \in M) \wedge (x \notin N)$$

(d) die symmetrische Differenz von N in M als

$$M \Delta N := (M \cup N) \setminus (M \cap N) = (M \setminus N) \cup (N \setminus M)$$

oder dazu äquivalent

$$x \in M \Delta N \quad \Longleftrightarrow \quad (x \in M) \veebar (x \in N)$$

BEISPIEL 3.10. Sei $A := \{1, 2, 3\}$ und $B := \{3, 4, 5\}$. Dann ist $A \cap B = \{3\}$, $A \cup B = \{1, 2, 3, 4, 5\}$, $A \setminus B = \{1, 2\}$ und , $A \Delta B = \{1, 2, 4, 5\}$

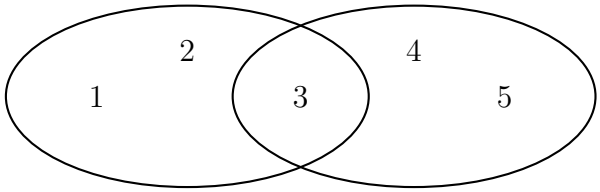


Abbildung 1: Zwei Mengen $A := \{1, 2, 3\}$ und $B := \{3, 4, 5\}$ zur Veranschaulichung von $A \cap B$, $A \cup B$ und $A \Delta B$.

DEFINITION 3.11. Zwei Mengen M und N heißen disjunkt, falls $M \cap N = \emptyset$.

Es gelten viele zu Proposition 2.9 analoge Aussagen.

PROPOSITION 3.12. Seien P, Q und R Teilmengen von M. Dann gilt

(a) $P \cup (M \setminus P) = M$
(b) $P \cap (M \setminus P) = \emptyset$
(c) $P \cup M = M$
(d) $P \cap \emptyset = \emptyset$
(e) $P \cap \emptyset = \emptyset$
(f) $P \cup \emptyset = P$
(g) $P \cap M = P$
(h) $P \cup M = M$
(i) $P \cup Q = Q \cup P$ (Kommutativität von $\cup$)
(j) $P \cap Q = Q \cap P$ (Kommutativität von $\cap$)
(k) $(P \cup (Q \cup R)) = ((P \cup Q) \cup R)$ (Assoziativität von $\cup$)

...

Jede Teilaussage dieser Proposition entspricht einer Teilaussage in Proposition 2.9 mit demselben kleinen Buchstaben. Zum Beispiel ist Proposition 3.12 (a) eine direkte Folgerung aus Proposition 2.9 (a), Proposition 3.12 (b) eine Folgerung aus Proposition 2.9 (b), etc.. Man beachte, dass die Bedeutungen von (d) und (e) in Proposition 2.9 fast diesseblen sind und deswegen dieselben „Übersetzungen“ in Proposition 3.12 haben. Analoges gilt für (c) und (h).

Übung 3.13. Übersetzen Sie die restlichen Aussagen aus Proposition 2.9, nämlich (l) bis (y) in Aussagen über Mengen. Beweisen Sie all diese Aussagen.

4. Quantoren

Aussagen[12] in der Art „Für alle Elemente $x \in M$ ist die Aussage $A(x)$ wahr.“ und „Es gibt (mindestens) ein Element $x \in M$, für das $A(x)$ wahr ist.“ sind in der Mathematik sehr häufig. Man schreibt kurz hierfür $\forall x \in M : A(x)$ und $\exists x \in M : A(x)$. Man nennt $\forall$ den Allquantor und $\exists$ den Existenzquantor. Als Beispiel betrachten wir die Definition von Stetigkeit, ein zentraler Begriff der Analysis.

Definition 4.1. Eine Abbildung $f : \mathbb{R} \longrightarrow \mathbb{R}$ ist stetig in $x_0 \in \mathbb{R}$, genau dann wenn

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists \delta \in \mathbb{R}_{>0} : \forall x \in \mathbb{R} : \underbrace{(|x - x_0| < \delta) \to (|f(x) - f(x_0)| < \epsilon)}_{A(x, x_0, f, \delta, \epsilon)}$$

$A(x, x_0, f, \delta, \epsilon)$ ist eine Aussageform, die von fünf Variablen abhängt. Wir geben die möglichen Werte durch Mengen an: $x \in \mathbb{R}$, $x_0 \in \mathbb{R}$, $f \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ (die Menge aller Abbildungen

[12] Natürlich erlauben wir hier auch aussagenlogische Formeln und Aussageformen

$\mathbb{R}\longrightarrow\mathbb{R}$), $\delta \in \mathbb{R}_{>0}$ und $\epsilon \in \mathbb{R}_{>0}$. Sobald wir am Ende des Abschnitts Abbildungen und Produkte von Mengen eingeführt haben, kann man diese Aussageform auch als Abbildung

$$A : \mathbb{R} \times \mathbb{R} \times \mathrm{Abb}(\mathbb{R}, \mathbb{R}) \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \longrightarrow \{w, f\}$$

betrachten.

Das folgende Regel ist intuitiv klar:

Regel 4.2 (Negation von Quantoren)**.** Für jede auf M definierte Aussageform gilt:

$$\neg\forall x \in M : \ A(x) \quad \Longleftrightarrow \quad \exists x \in M : \ \neg A(x) \tag{4.3}$$

Indem man $A(x)$ durch $\neg A(x)$ ersetzt und dann die gesamte Äquivalenz auf beiden Seiten negiert, erhält man

$$\neg\exists x \in M : \ A(x) \quad \Longleftrightarrow \quad \forall x \in M : \ \neg A(x) \tag{4.4}$$

Klammerregel.
Sei $A(x, y)$ eine Aussage die von den Parametern $x \in X$ und $y \in Y$, dann ist

$$\forall/\exists x \in X : \ \forall/\exists y \in Y : \ A(x, y)$$

zu lesen als

$$\forall/\exists x \in X : \ \Big(\forall/\exists y \in Y : \ A(x, y)\Big).$$

Es gilt offensichtlich

$$\forall x \in X : \forall y \in Y : \ A(x, y) \iff \forall y \in Y : \forall x \in X : \ A(x, y)$$

$$\exists x \in X : \exists y \in Y : \ A(x, y) \iff \exists y \in Y : \exists x \in X : \ A(x, y)$$

Man sagt kurz: „zwei af. (aufeinanderfolgende) Allquantoren vertauschen miteinander“, oder man sagt „zwei af. Allquantoren kommutieren“. Analog kommutieren zwei af. Existenzquantoren.

!ACHTUNG!. Allquantoren kommutieren nicht mit Existenzquantoren!

Die eine Richtung ist immer wahr:

$$\forall x \in X : \exists y \in Y : A(x,y) \iff \exists y \in Y : \forall x \in X : A(x,y)$$

Aus der rechten Seite $\exists y \in Y : \forall x \in X : A(x,y)$ folgt die linke $\forall x \in X : \exists y \in Y : A(x,y)$. Die Aussagen unterscheiden sich aber sehr. Links darf die Wahl von y von x abhängen, rechts darf sie es nicht. Man hat hier also weniger Flexibilität in der Wahl von y, da man eines finden muss, so dass $A(x,y)$ für alle $x \in X$ richtig ist.

Die Umkehrung (d.h. die Richtung von links nach rechts) ist falsch, wie das folgende Beispiel zeigt.

Beispiel 4.5. Die Aussage

$$\forall x \in \mathbb{R} : \exists y \in \mathbb{R} : x > y \tag{4.6}$$

ist wahr. Denn für ein gegebenes $x \in \mathbb{R}$ können wir die Zahl $y := x - 1$ hernehmen und diese erfüllt $y < x$. Deswegen ist für alle $x \in \mathbb{R}$ die Aussage $\exists y \in \mathbb{R}\ x > y$ wahr, also auch (4.6).

Die Aussage

$$\exists y \in \mathbb{R} : \forall x \in \mathbb{R} : x > y \tag{4.7}$$

ist falsch. Denn für ein gegebenes $y \in \mathbb{R}$ gilt nicht für alle $x \in \mathbb{R}$ die Ungleichung $x > y$. Die Ungleichung gilt zum Beispiel nicht für $x := y$. Also gibt es kein $y \in \mathbb{R}$, so dass $\forall x \in \mathbb{R} : \ x > y$ wahr ist. Also ist (4.7) falsch.

Schreibweise 4.8. Quantoren für eine Variable müssen stehen, bevor diese Variable gebraucht werden darf. (Sonst gelten sie als *nicht* syntaktisch sinnvoll.) [13]

Beispiel 4.9. Ausdrücke in der Art „Es gilt $2x = x + x$ für alle $x \in \mathbb{R}$“ erklären wir ab sofort und bis auf weiteres als syntaktisch nicht sinnvoll. Erlaubt sind nur Sätze wie „Für alle $x \in \mathbb{R}$ gilt $2x = x + x$.“

Die Regel ist deswegen wichtig, damit nicht unklar bleibt, wie die folgende Aussage verstanden werden soll:

$$\text{Für alle } x \in \mathbb{R} \text{ gilt } x > y \text{ für ein } y \in \mathbb{R}.$$

Diesen Ausdruck könnte man als die wahre Aussage (4.6) oder als die falsche Aussage (4.7) verstehen. Deswegen legen wir durch die obige Regel fest, dass Ausdrücke wie das obige A syntaktisch nicht sinnvoll sind.

Frage 4.10. Ist die folgende Aussage A wahr?

$$A :\Longleftrightarrow \quad \forall x \in \emptyset : \ x = x + 1.$$

Das didaktische Problem hier ist: obwohl uns bisher intuitiv klar war, was $\forall x \in M : \ A(x)$ bedeutet, ist $\forall x \in \emptyset : \ A(x)$ nicht intuitiv klar. In studentischen Lösungen der Aufgabenblätter steht oft in so einem Fall: „Da die Aussage $x = x + 1$ für kein einziges

[13] In späteren Semestern wird diese Regel nicht mehr immer befolgt, **falls** keine Missverständnisse zu erwarten sind.

Element in $\emptyset$ gilt, gilt sie natürlich auch nicht für alle Elemente in $\emptyset$.“ Das ist ein weit verbreiteter Fehler.

Regel 4.11. In der Mathematik ist die Aussage

$$\forall x \in \emptyset : A(x)$$

für alle Aussageformen $A(x)$ immer wahr und die Aussage

$$\exists x \in \emptyset : A(x)$$

für alle Aussageformen $A(x)$ immer falsch.

Diese Regel ist letztendlich eine Definition, man könnte auch sagen eine Konvention. Es erscheint auch intuitiv klar, dass $\exists x \in M : A(x)$ nur dann wahr sein kann, wenn in M überhaupt irgendwelche Elemente existieren. Bei der anderen Regel könnte man Zweifel haben. Aber: wenn wir hier aber eine andere Konvention wählen würden, hätten wir viele Ausnahmen zu berücksichtigen. Zum Beispiel erscheint uns die folgende Äquivalenz für alle Mengen M intuitiv klar:

$$\forall x \in M : A(x) \quad \Longleftrightarrow \quad \{x \in M \mid \neg A(x)\} = \emptyset.$$

Die rechte Seite ist für $M = \emptyset$ offensichtlich wahr.

Außerdem, ergibt sich die erste Teil in Regel 4.11 auch aus dem zweiten Grund: $\exists x \in \emptyset : \neg A(x)$ ist falsch, also auch $\neg \forall x \in \emptyset : A(x)$ falsch, also $\forall x \in \emptyset : A(x)$ wahr.

Notation 4.12. In vielen Büchern schreibt man auch oft $\exists x : A(x)$. Dies bedeutet, es existiert eine Menge, in der ein Element existiert, für das $A(x)$ wahr ist. Beispiel: die Aussage $\exists x : x \in M$ ist äquivalent zu $M \neq \emptyset$. Analog heißt $\forall x : A(x)$, dass für

alle Mengen und alle Elemente darin die Aussage $A(x)$ wahr ist.[14] Analog hierzu: ist $A(x)$ eine auf allen Elementen von allen Mengen definierte Aussageform, so bezeichnet $\{x \mid A(x)\}$ die Menge aller Elemente, für die $A(x)$ wahr ist. Diese Bezeichnungen sind vor allem in axiomatischen aufgebauten Logik-Büchern wie [**13**] üblich, sind im Zugang unserer Vorlesung eine gefährliche Fehlerquelle: zum Beispiel ist $\{x \mid A(x)\}$ oft gar keine Menge, siehe Anhang A Abschnitt 1. Wir wollen sie deswegen nicht verwenden.

5. Potenzmenge und Mengensysteme

DEFINITION 5.1. Ist X eine Menge, so ist $\mathcal{P}(X)$ definiert als die Menge aller Teilmengen von M. Man nennt $\mathcal{P}(X)$ die Potenzmenge von X.

$$x \in \mathcal{P}(X) \quad \Longleftrightarrow \quad x \subset X$$

BEISPIELE 5.2.

(1) $\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$
(2) $\mathcal{P}(\emptyset) = \{\emptyset\} \neq \emptyset$

DEFINITION 5.3. Ein Mengensystem (über einer Grundmenge X) ist eine Teilmenge von $\mathcal{P}(X)$. Dies bedeutet, dass M genau dann ein Mengensystem ist, wenn alle Elemente von M eine Teilmenge von X sind.

[14] Bei dieser „Notation“ handelt es sich eigentlich auch um eine Definition; es ist die Definition einer Schreibweise. Wir verwenden hier den Begriff Notation, um anzudeuten, dass hier kein neuer Begriff eingeführt wird, sondern eine neue Schreibweise.

Wir definieren dann

$$\bigcup M := \{x \in X \mid \exists m \in M : x \in m\}$$

und falls zusätzlich $M \neq \emptyset$ definieren wir

$$\bigcap M := \{x \in X \mid \forall m \in M : x \in m\}.$$

Offensichtlich gilt $\bigcup M \subset X$ und $\bigcap M \subset X$.

Beispiele 5.4. Beliebige Grundmenge X, $A \subset X$, $B \subset X$

(1) Für $M := \{A\}$ gilt $\bigcup M = \bigcap M = A$.
(2) Für $M := \{A, B\}$ gilt $\bigcup M = A \cup B$ und $\bigcap M = A \cap B$.

Mengensysteme über der Grundmenge $\mathbb{Z}$

(3) Für $M := \{\{1,2,3\},\{3,5,7\},\{1,3,5\}\}$ gilt $\bigcup M = \{1,2,3,5,7\}$ und $\bigcap M = \{3\}$.
(4) Sei

$$M := \{k\mathbb{Z} \,\|\, k \in \mathbb{N} \setminus \{0,1\}\} = \{2\mathbb{Z}, 3\mathbb{Z}, 4\mathbb{Z}, \ldots\}.$$

Dann gilt $\bigcup M = \mathbb{Z} \setminus \{-1,+1\}$ und $\bigcap M = \{0\}$.

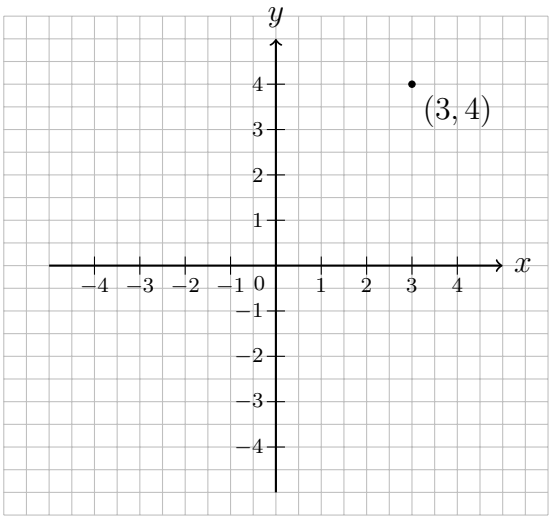
Falls $X \subset Y$, so ist jedes Mengensystem über X auch ein Mengensystem über Y und die obigen Definitionen von $\bigcup M$ und $\bigcap M$ hängen nicht von der verwendeten Grundmenge ab. Wir lassen deswegen die Angabe der Grundmenge in Zukunft weg.

$\bigcap \emptyset$ ist nicht definiert worden, denn mit obiger Definition wäre dann $\bigcap \emptyset = X$, also von der Wahl der Grundmenge abhängig. **26.10.**

6. Paare und kartesische Produkte

Es gilt $\{1, 2\} = \{2, 1\}$. Eine Menge mit genau zwei Elementen heißt ungeordnetes Paar. In der Mathematik brauchen aber auch (geordnete) Paare[15]. Wir wollen an einem Beispiel erklären, wieso sie gebraucht werden, und uns dann um eine saubere Defintion kümmern.

Réné Descartes (1596–1650):
Nach Wahl eines Ursprungs und einer Basis beschreibt man Punkte auf einer Geraden durch eine reelle Zahl, Punkte in einer Ebene durch zwei reelle Zahlen und Punkte im Raum durch drei reelle Zahlen.



$\{$ Punkte in der Ebene $\} \cong \{$ geordnete Paare (x, y) von zwei reellen Zahlen x und $y\}$
$\cong \mathbb{R} \times \mathbb{R} \cong \mathbb{R}^2$

Paar := geordnetes Paar

[15]Ein Paar ist immer ein geordnetes Paar, es sei denn wir sagen explizit ‘„ungeordnet“ dazu.

„DEFINITION“ 6.1. Sind M und N Mengen, so definieren wir

$$M \times N := \{(m, n) \,\|\, m \in M,\, n \in N\}$$

wobei (m, n) ein Paar aus m und n ist. Man nennt $M \times N$ das (kartesische) Produkt von M und N.

Für Paare gilt z.B. $(3, 4) \neq (4, 3)$.

Das Wort „kartesisch“ kommt vom Namen „Cartesius“ und dies ist die lateinische Version von Descartes.

Das problematische an obiger „Definition“ ist, dass wir den Begriff „Paar“ noch gar nicht definiert haben.

Eine Möglichkeit, wäre ein Axiom einzuführen, das uns die Existenz von Paaren liefert.

AXIOM 6.2. Zu jedem $m \in M$ und $n \in N$ gibt es ein (m, n), genannt Paar, mit der folgenden Eigenschaft: für alle $m, \tilde{m} \in M$ und alle $n, \tilde{n} \in N$ gilt

$$(m, n) = (\tilde{m}, \tilde{n}) \Longleftrightarrow m = \tilde{m} \text{ und } n = \tilde{n}.$$

Ein ähnliches Axiom wurde in der Linearen Algebra so ähnlich eingeführt (Axiom 1.1.6), es ergänzt die dort präsentierten anderen Axiome. Wenn man aber naiv argumentiert, wie bei uns der Fall, oder die zumeist verwendeten Axiomen-Systeme ZFC oder NBG (siehe Anhang A) nutzt, kann man eine Konstruktion machen, die genau solch eine Paarbildung liefert. Sie macht aus dem obigen Definitions-Versuch („Definition“ 6.1) eine richtige Definition.[16] Das heißt, man kann Axiom 6.2 weglassen und durch diese Definition ersetzen.

[16]Wir können dann also die Anführungsstriche weglassen.

DEFINITION 6.3 (Konstruktion von Paaren nach Kuratowski). Wir definieren das Paar von m und n als

$$(m, n) := \{\{m\}, \{m, n\}\}.$$

Diese Definition ist unanschaulicher als sie zunächst aussieht, wie das folgende Beispiel zeigt:

BEISPIEL 6.4. Für $M = N = \mathbb{N}$ und $m = n = 1$ gilt $(m, n) = \{\{1\}\}$.

Aus diesem Grund wählen manche Bücher andere Definitionen, die aber wieder andere Probleme haben. Wichtig ist für uns aber letztendlich gar nicht, was Produkte „wirklich sind“, sondern dass die Eigenschaft in Axiom 6.2 erfüllt ist.

LEMMA 6.5. Für das von Kuratowski definierte Paar gilt:

$$(m, n) = (\tilde{m}, \tilde{n}) \quad \Longleftrightarrow \quad m = \tilde{m} \text{ und } n = \tilde{n}.$$

Beweis. Die Richtung „$\Longleftarrow$“ ist offensichtlich.

Zu „$\Longrightarrow$“: Sei $(m, n) = (\tilde{m}, \tilde{n})$. Also gilt nach Definition

$$\{\{m\}, \{m, n\}\} = \{\{\tilde{m}\}, \{\tilde{m}, \tilde{n}\}\}. \tag{6.6}$$

Wir wissen bereits, dass $\{\{m\}, \{m, n\}\}$ ein oder zwei Elemente besitzt. Wir müssen zwei Fälle unterscheiden.

1. Fall: Die Menge $\{\{m\}, \{m, n\}\}$ besitzt ein Element.
In diesem Fall ist dann $\{m\}$ das einzige Element und somit $\{m\} = \{m, n\}$ und deswegen

$m = n$. Wegen (6.6) hat auch $\{\{\tilde{m}\}, \{\tilde{m}, \tilde{n}\}\}$ nur ein Element, also ergibt sich analog $\tilde{m} = \tilde{n}$. Es gilt nun

$$\{m\} = \{\{m\}, \{m, n\}\} = \{\{\tilde{m}\}, \{\tilde{m}, \tilde{n}\}\} = \{\tilde{m}\}.$$

Es folgt $m = n = \tilde{m} = \tilde{n}$.

2. Fall: Die Menge $\{\{m\}, \{m, n\}\}$ besitzt zwei Elemente.
Es gilt also $\{m\} \neq \{m, n\}$ und somit $m \neq n$. Analog enthält $\{\{\tilde{m}\}, \{\tilde{m}, \tilde{n}\}\}$ die verschiedenen Elemente $\{\tilde{m}\}$ und $\{\tilde{m}, \tilde{n}\}$, und daraus folgt $\tilde{m} \neq \tilde{n}$. Da $\{m\}$ und $\{\tilde{m}\}$ genau ein Element enthalten, und da $\{m, n\}$ und $\{\tilde{m}, \tilde{n}\}$ genau zwei Elemente enthalten, folgt aus (6.6), dass $\{m\} = \{\tilde{m}\}$ und somit $m = \tilde{m}$. Außerdem gilt dann $\{m, n\} = \{\tilde{m}, \tilde{n}\}$. Da diese Menge außer $m = \tilde{m}$ noch genau ein weiteres Element enthält, gilt auch $n = \tilde{n}$. □

Definition 6.7. Falls M_1, M_2, …, M_k Mengen sind, so definieren das k-fache kartesische Produkt hiervon als

$$M_1 \times M_2 \times \cdots \times M_k := M_1 \times (M_2 \times (\cdots M_{k-1} \times M_k)) \cdots).$$

Elemente hiervon nennen wir k-Tupel. Wir schreiben sie in der Form $(x_1, x_2, \ldots, x_k)$, wobei $x_i \in M_i$. Ein Paar ist also ein 2-Tupel. Ein 3-Tupel nennt man auch Tripel. 4-Tupel sind Quadrupel, 5-Tupel sind Quintupel. Wir nutzen auch die Schreibweise

$$M^k = \underbrace{M \times M \times \cdots \times M}_{k\text{-mal}}.$$

Beispiel: $\mathbb{R}^3 = \mathbb{R} \times (\mathbb{R} \times \mathbb{R})$

7. Relationen, funktionale Relationen, Abbildungen

DEFINITION 7.1. Eine Relation R auf der Grundmenge M ist eine Teilmenge von $M \times M$. Wir schreiben dann xRy für die Aussage(form) $(x, y) \in R$.

Falls $M \subset N$, so ist jede Relation auf M auch eine Relation auf N. Oft ist die Grundmenge ohne Bedeutung. Wichtig ist vor allem, dass ein solches Resservoir an möglichen Werten überhaupt existiert.

BEISPIELE 7.2.

(1) $\emptyset$ ist eine Relation auf jeder Menge.
(2) $\{\emptyset, (1, 2)\}$ ist keine Relation, da $\emptyset$ kein Paar ist.
(3) Ist M eine Menge, so ist $\delta_M := \{(x, x) \,\|\, x \in M\}$ eine Relation auf M. Man nennt sie die Diagonale von M.

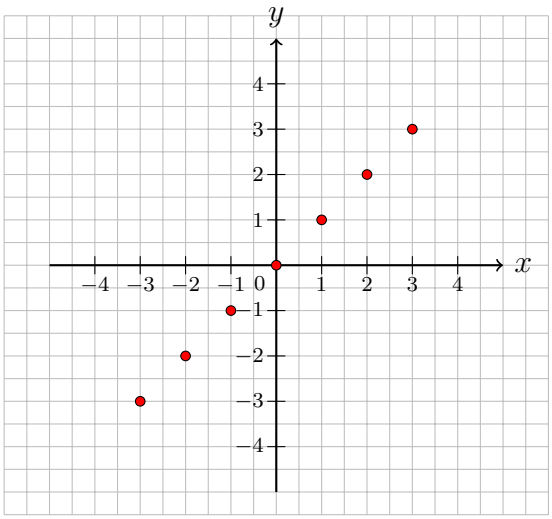
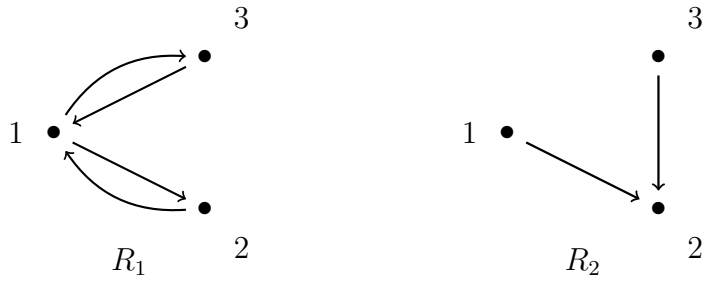


Abbildung 2: Die Diagonale der Menge $M = \{-3, -2, -1, 0, 1, 2, 3\}$.

(4) $\leq_{\mathbb{R}} := \{(x, y) \in \mathbb{R}^2 \mid x \leq y\}$ ist eine Relation auf $\mathbb{R}$.

(5) $Q := \{(x, y) \in \mathbb{R}^2 \mid y = 3x^2 + 2\}$ ist eine Relation auf $\mathbb{R}$.

(6) $M = \{1, 2, 3\}$. Wir definieren die Relationen $R_1 \subset M \times M$ und $R_2 \subset M \times M$ durch das folgende Bild oder sagen $R_1 := \{(1, 2), (1, 3), (2, 1), (3, 1)\}$ und $R_2 := \{(1, 2), (3, 2)\}$



Definition 7.3. Sei R eine Relation auf M. Die Relation heißt

(1) reflexiv auf M $:\Longleftrightarrow \forall x \in M : xRx$,

(2) symmetrisch $:\Longleftrightarrow \underbrace{\forall x \in M : \forall y \in M :}_{\forall x,y \in M:} (xRy \to yRx)$,

(3) antisymmetrisch $:\Longleftrightarrow \forall x, y \in M : \big((xRy \wedge yRx) \to x = y\big)$,

(4) transitiv $:\Longleftrightarrow \forall x, y, z \in M : \big((xRy \wedge yRz) \to xRz\big)$,

(5) Äquivalenzrelation auf M, falls R reflexiv auf M, symmetrisch und transitiv ist.

(6) Ordnung(srelation) auf M oder partielle Ordnung(srelation) auf M, falls R reflexiv, antisymmetrisch und transitiv ist.

(7) Erfüllt eine Ordnungsrelation R auf M zusätzlich die Eigenschaft

$$\forall x, y \in M : \big(xRy \vee yRx\big)$$

dann nennt man R eine totale Ordnung(srelation).[17]

BEISPIELE 7.4. Zu den Beispielen 7.2 (3), (4) und (6):

(3) δ_M ist reflexiv auf M, symmetrisch, und transitiv, also eine Äquivalenzrelation auf M.

Es gilt

$$x\delta_M y \iff x = y \wedge x \in M$$

(4) $\leq_{\mathbb{R}}$ ist reflexiv auf $\mathbb{R}$, antisymmetrisch und transitiv, also eine Ordnungsrelation auf M. Diese Ordnung ist total.

(6) R_1 ist nicht transitiv, denn $3R_1 1$ und $1R_1 2$ sind wahr, aber $3R_1 2$ ist falsch. R_1 ist auch nicht reflexiv. Also ist R_1 auch weder Ordnungsrelation noch Äquivalenzrelation. R_1 ist symmetrisch und nicht antisymmetrisch.

R_2 ist transitiv, antisymmetrisch, nicht reflexiv, nicht symmetrisch.

BEISPIELE 7.5. Ist X eine Menge, so definieren wir

$$\subset_X := \{(A, B) \in \mathcal{P}(X)^2 \mid A \subset B\}.$$

$\subset_X$ ist eine Ordnungsrelation auf $\mathcal{P}(X)$. Sie ist keine totale Ordnung, wenn X mindestens 2 Elemente besitzt. [18]

[17] Manche Quellen nennen solche eine Relation eine lineare Ordnungsrelation. Andere Quellen wiederum bezeichnen unsere „Ordnungsrelationen“ als „Halbordnungen“ und unsere „totalen Ordnungen“ als „Ordnungen“.

[18] Bei Ordnungsrelation werden immer zwei Objekte verglichen: Ist eine Zahl größer als eine andere? Ist eine Menge in der anderen enthalten? Etc. Insofern sind $\leq_{\mathbb{R}}$ und $\subset_X$ ganz wichtige Beispiele. δ_M ist auch antisymmetrisch, also auch eine Ordnungsrelation. Sobald M mindestens 2 Elemente besitzt,

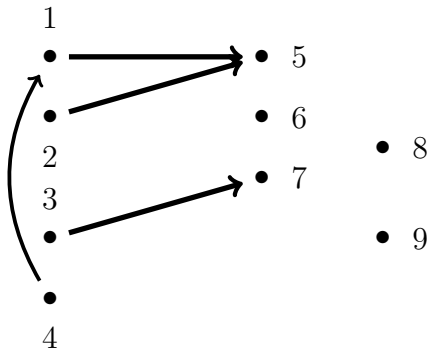
DEFINITION 7.6. Eine Relation F auf M nennen wir eine funktionale Relation, falls gilt:

$$\forall x \in M : \forall y \in M : \forall z \in M : \Big((x, y) \in F \wedge (x, z) \in F\Big) \to y = z.$$

Zu jedem $x \in M$ gibt es also höchstens ein $y \in M$ mit xFy. Wir sagen dann: x wird durch F auf y abgebildet und schreiben $x \mapsto y$ oder $y = F(x)$.

Beispiele 7.2 (1), (3) und (5) sind funktionale Relationen. Beispiel (4) ist keine funktionale Relation. In Beispiel 7.2 (6) ist R_2 funktional, aber R_1 ist nicht funktional.

BEISPIEL 7.7. Es gelte nun xRy, falls im folgenden Bild ein Pfeil von x nach y führt.



Dann ist R eine funktionale Relation auf der Grundmenge $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

---

ist es keine totale Ordnungsrelation. Dennoch ist δ_M ein atypisches Beispiel, denn δ_M ist die einzige symmetrische Ordnungsrelation auf M.

DEFINITION 7.8. Ist R eine Relation auf der Grundmenge M, so definieren wir[19]:

$$D(R) := \{x \in M \mid \exists y \in M : xRy\},$$

Definitionsbereich von R;

$$B(R) := \{y \in M \mid \exists x \in M : xRy\},$$

Bild der Relation R;

BEISPIEL 7.7. (fortgesetzt) Wir zeichnen links die Definitionsmenge $D(R)$, rechts die Grundmenge M. Die bisherigen Pfeile zeichnen wir so ein, dass sie in $D(R)$ beginnen und in M enden. Wir wählen einen „Zielbereich“ Y aus mit $B(R) \subset Y \subset M$.

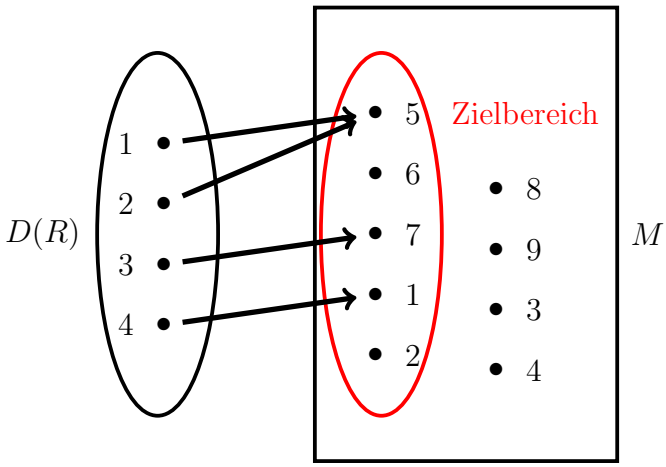


Abbildung 3: Eine funktionale Relation R mit $D(R) = \{1, 2, 3, 4\}$, $B(R) = \{1, 5, 7\}$, Zielbereich $\{1, 2, 5, 6, 7\}$ und Grundmenge $M = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

[19]Viele Bücher, z.B. [**17**] bezeichnen $B(R)$ als Wertebereich oder Wertemenge. Es gibt aber auch Bücher, die den Begriff Wertebereich im Sinne des unten definierten Zielbereichs definieren. Wir werden aus diesem Grund die Begriffe „Wertebereich“ und „Wertemenge“ vermeiden.

DEFINITION 7.9. Eine Abbildung ist ein Tripel (f, X, Y), bestehend aus Mengen X und Y und einer funktionalen Relation f auf der Grundmenge $X \cup Y$, wobei $D(f) = X$ und $B(f) \subset Y$. Zumeist schreibt man Abbildungen als $f : X \longrightarrow Y$ oder $X \xrightarrow{f} Y$. Man nennt X den Definitionsbereich der Abbildung $f : X \longrightarrow Y$, man nennt Y den Zielbereich der Abbildung und man nennt f den Graphen der Abbildung. Eine Abbildung von X nach Y ist eine Abbildung mit Definitionsbereich X und Zielbereich Y. Die Menge aller Abbildungen von X nach Y notieren wir als $\mathrm{Abb}(X, Y)$ oder als Y^X. Die Notation $x \mapsto y$ bedeutet, dass x auf y abgebildet wird, d.h. (x, y) ist ein Element der funktionalen Relation.

Der Begriff „Funktion“ ist gleichbedeutend zum Begriff „Abbildung“. Es gibt Situationen, in denen man häufiger den Begriff „Abbildung“ nutzt, in anderen ist der Begriff „Funktion“ gebräuchlicher. Zum Beispiel, wenn der Zielbereich die reellen Zahlen sind, so redet man von (reell-wertigen) Funktionen.

31.10.

BEISPIELE 7.10.

(a) $f := \{(x, x^2) \,\|\, x \in [0, 5]\} \subset [0, 5] \times \mathbb{R}$ ist eine funktionale Relation mit $D(f) = [0, 5]$ und $B(f) = [0, 25]$. Dann ist $f : [0, 5] \longrightarrow [-4, 40]$ eine Abbildung. Wir schreiben derartige Abbildungen als

$$\begin{array}{ccc} f : [0,5] & \longrightarrow & [-4,40] \\ x & \mapsto & x^2 \end{array}$$

Alternativ kann man die zweite Zeile auch durch $f(x) = x^2$ oder $f(x) := x^2$ ersetzen.

(b) Sei f wie oben. Die Abbildungen $F_1 := (f, [0, 5], [0, 25])$ und $F_2 := (f, [0, 5], [-4, 40])$ sind verschieden, obwohl dieselbe funktionale Relation f zu Grunde liegt.

(c) Die Identität id_M von M ist die Abbildung (δ_M, M, M), die wir (dem obigen Beispiel folgend) auch als $M \longrightarrow M$, $x \mapsto x$ schreiben können.

(d) Ist M eine Teilmenge von N, so nennen wir die Abbildung (δ_M, M, N) die Inklusion von M in N oder genauer Inklusion der Menge M in die Menge N. [20]

BEMERKUNG 7.11. Aus Ihrer jetzigen Sicht ist der essentielle Unterschied zwischen einer funktionalen Relation und einer Abbildung, dass im Begriff der Abbildung auch der Zielbereich festgelegt ist, während eine funktionale Relation keinen Zielbereich hat. [21] Später werden der Definitionsbereich und der Zielbereich zusätzliche Struktur tragen: z.B. eine Vektorraumstruktur, eine Gruppenstruktur oder eine Abstandsfunktion. Da wichtige Eigenschaften von Abbildungen wie z.B. Linearität und Stetigkeit von solcher Zusatzstruktur abhängt, wird es dann hilfreich sein, dass auch der Definitionsbereich im Tripel einer Abbildung enthalten ist.

Es gibt auch einige Teilbereiche der Mathematik und ihren Anwendungen, in denen man $D(f) = X$ durch $D(f) \subset X$ ersetzt. Dies ist zum Beispiel bei den „dicht definierten Operatoren„ der Funktionalanalysis und der Quantenmechanik der Fall, zum Beispiel der Orts- und der Impuls-Operator der Quantenmechanik.

---

[20] wichtig ist hier die Konjugation von „die Menge“: Genitiv (bzw. von-Konstruktion), dann Akkusativ

[21] Vergleich zu anderen Lehrbüchern: Es herrscht Uneinigkeit in der Literatur, ob eine Abbildung Information über den Zielbereich trägt oder nicht. Das Lehrbuch [**12**] unterscheidet in diesem Zusammenhang zwischen „Funktionen“, die wie unsere „funktionale Relationen“ definiert sind, und „Funktionen von X nach Y“, die im wesentlichen unseren „Abbildungen“ entsprechen. In [**17**] bedeuten die Begriffe „funktionale Relation“, „Funktion“ und „Abbildung“ dasselbe, es bleibt aber dort unklar, ob die oben definierten Abbildungen F_1 und F_2 in diesem Buch gleich oder verschieden sind. Wir werden in der Vorlesung die obigen Definitionen nutzen und zulassen, um innerhalb unserer Analysis I Klarheit zu schaffen.

BEMERKUNG 7.12. Zur Interpretation: Äquivalenz- und Ordnungs-Relationen auf der einen Seite und funktionale Relationen sind zwar ähnlich definiert (als Relationen und somit als Mengen, deren Elemente Paare sind). Dennoch interpretiert man sie anders: Äquivalenz- und Ordnungsrelationen vergleichen zwei Objekte. Beispiel: die Relation $\leq_{\mathbb{R}}$, gibt an, ob die links stehende Zahl kleiner oder gleich der rechts stehenden ist. Die Interpretation von funktionalen Relationen ist dynamischer, $x \mapsto x^2$ ordnet x den Wert x^2 zu, so wie ein(e) Platzanweiser(in) im Theater jedem Zuschauer sagt, in welchem Stockwerk er den Theatersaal betreten darf: im Parkett (0. Stockwerk) oder im k-ten Rang (k-tes Stockwerk): jedem Zuschauer wird ein Stockwerk zugeordnet. Aus diesem Grund ist die Schreibweise xRy für Ordnungs- und Äquivalenzrelationen üblich, wohingegen die Schreibweise $y = R(x)$ für funktionale Relationen üblich ist. Beides ist defniert als $(x, y) \in R$.

BEMERKUNG 7.13. Sei A eine auf M definierte Aussageform und sei $f : M \longrightarrow N$ eine Abbildung. Dann gilt

$$\{f(x) \,\|\, x \in M\} = \{y \in N \mid \exists x \in M : \, f(x) = y\} = B(f).$$

DEFINITION 7.14. Sei $f : X \longrightarrow Y$ eine Abbildung.

(1) Für $M \subset X$ definieren wir die Einschränkung oder Restriktion von $f : X \longrightarrow Y$ auf M als die Abbildung $(f|_M, M, Y)$, wobei

$$f|_M := \{(x, y) \in X \times Y \mid x \in M \wedge y = f(x)\} = \{(x, f(x)) \,\|\, x \in M\},$$

(2) Ist $g : Y \longrightarrow Z$ eine weitere Abbildung, so definieren wir die Verkettung oder die Komposition von $g : Y \longrightarrow Z$ mit $f : X \longrightarrow Y$ als die Abbildung $(g \circ f, X, Z)$, wobei

$$g \circ f(x) := g(f(x)).$$

Man liest $g \circ f$ als „g nach f“.

DEFINITION 7.15.

(1) Eine funktionale Relation f bzw. eine Abbildung $f : X \longrightarrow Y$ heißt injektiv, falls gilt

$$\forall x_1, x_2 \in X : (f(x_1) = f(x_2) \to x_1 = x_2).$$

Dies bedeutet, dass jedes Element in X unter f auf ein anderes Element in Y abgebildet wird.

(2) Eine Abbildung $f : X \longrightarrow Y$ heißt surjektiv, falls $B(f) = Y$.

(3) Eine Abbildung $f : X \longrightarrow Y$ heißt bijektiv, falls sie injektiv und surjektiv ist.

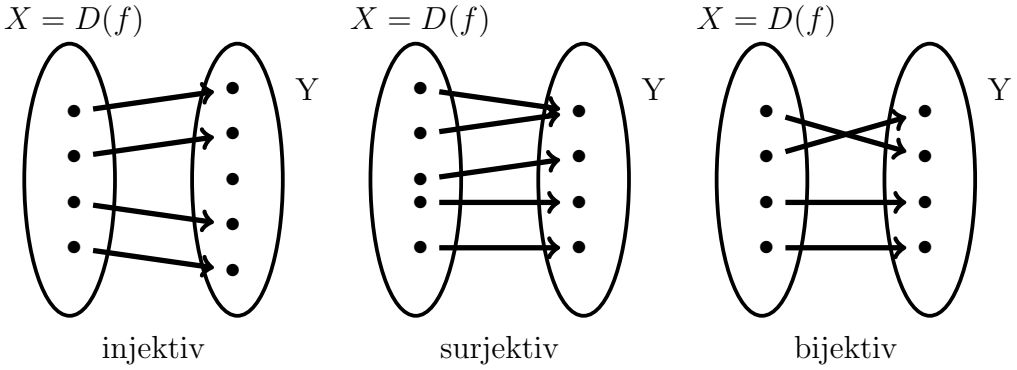


Abbildung 4: Schemtische Darstellungen von injektiv, surjektiv und bijektiv.

Wenn f eine funktionale Relation ist, so ist $f : D(f) \longrightarrow B(f)$ eine surjektive Abbildung.

BEISPIELE 7.16. (in der Vorlesung übersprungen = grau)

(a) Die Abbildung $\mathbb{R} \longrightarrow \mathbb{R}$, $x \mapsto x^2$ ist weder injektiv noch surjektiv.

(b) Die Abbildung $\mathbb{R}_{>0} \longrightarrow \mathbb{R}$, $x \mapsto x^2$ ist injektiv, aber nicht surjektiv.

(c) Die Abbildung $\mathbb{R}_{>0} \longrightarrow \mathbb{R}_{>0}$, $x \mapsto x^2$ ist bijektiv.

Um das obige Beispiel der/s Platzanweisers/in nochmals weiter zu entwickeln: Wenn der Platzanweiser nicht nur das Stockwerk zuordnet, sondern auch den konkreten Sitzplatz, so ist zu hoffen, dass diese Abbildung injektiv ist, da sonst zwei Leute auf dem selben Platz sitzen müssen. Der Kassenwart des Theaters hingegen hofft, dass die Abbildung von der Menge der Zuschauer auf die Menge der Sitzplätze auch surjektiv ist, da dann das Theater voll besetzt ist.

DEFINITION 7.17. Ist R eine Relation und $R \subset M \times M$, dann definieren wir

$$R^{-1} := \{(y, x) \in M \times M \mid xRy\},$$

Man nennt R^{-1} die Umkehrung von R.

ÜBUNG 7.18.

(a) Sei f eine funktionale Relation. Dann ist f^{-1} genau dann eine funktionale Relation, wenn f injektiv ist.
(b) Ist $f : X \longrightarrow Y$ eine Abbildung, so ist $f^{-1} : Y \longrightarrow X$ genau dann eine Abbildung, wenn $f : X \longrightarrow Y$ bijektiv ist.
(c) Ist $f : X \longrightarrow Y$ eine bijektive Abbildung, so gilt $f^{-1} \circ f = \delta_X$ und $f \circ f^{-1} = \delta_Y$.

Man nennt dann $f^{-1} : Y \longrightarrow X$ die Umkehrabbildung von $f : X \longrightarrow Y$.

ÜBUNG 7.19. Seien $f : X \longrightarrow Y$ und $g : Y \longrightarrow Z$ Abbildungen. Zeigen Sie:

(1) Ist $g \circ f : X \longrightarrow Z$ surjektiv, dann ist auch g surjektiv.
(2) Ist $g \circ f : X \longrightarrow Z$ injektiv, dann ist auch f injektiv.

(3) Geben Sie ein Beispiel an, das zeigt, dass die Umkehrung der Aussage (1) nicht gilt. Das heißt: finden Sie Abbildungen $f : X \longrightarrow Y$ und $g : Y \longrightarrow Z$, so dass $g : Y \longrightarrow Z$ surjektiv ist, aber nicht $g \circ f : X \longrightarrow Z$.

(4) Geben Sie ein Beispiel an, das zeigt, dass die Umkehrung der Aussage (2) nicht gilt.

Übung 7.20. Seien X und Y Mengen, $X \neq \emptyset$. Es gibt eine injektive Abbildung $f : X \longrightarrow Y$ genau dann, wenn es eine surjektive Abbildung $g : Y \longrightarrow X$ gibt.[22]

Siehe: Aufgabe 3 auf Übungsblatt 3

Satz 7.21 (Schröder, Bernstein)**.** Seien X und Y Mengen. Angenommen, es gibt eine injektive Abbildung $f : X \longrightarrow Y$ und eine injektive Abbildung $g : Y \longrightarrow X$, dann gibt es auch eine bijektive Abbildung $g : X \longrightarrow Y$.

Beweis: 2. Übungsblatt der Linearen Algebra I

Definition 7.22. Seien X und Y Mengen. Wir sagen:

- X und Y sind gleich mächtig, falls es eine bijektive Abbildung $f : X \longrightarrow Y$ gibt.
- Y ist mächtiger als X, falls es eine injektive Abbildung $f : X \longrightarrow Y$ und es keine bijektive Abbildung $b : X \longrightarrow Y$ gibt.
- X ist endlich, falls es eine natürliche Zahl $n \in \mathbb{N}$ gibt, so dass X und $\{m \in \mathbb{N} \mid 0 < m \leq n\}$ gleich mächtig sind. Wir schreiben dann $\#X = n$ und sagen X hat n Elemente.
- X ist unendlich, falls X nicht endlich ist. Wir schreiben dann $\#X = \infty$.
- X ist abzählbar, falls es eine injektive Abbildung $X \to \mathbb{N}$ gibt.

[22] Leider habe ich in der Vorlesung die Bedingung $X \neq \emptyset$ vergessen!

- X ist überabzählbar, wenn X nicht abzählbar ist. Dann ist X mächtiger als $\mathbb{N}$.

Übung 7.23. Sei M ein Mengensystem. Für $A, B \in M$ definieren wir

$$A \sim B \quad :\Longleftrightarrow \quad A \text{ und } B \text{ sind gleichmächtig.}$$

Zeigen Sie, dass $\sim$ eine Äquivalenzrelation auf M ist.

Beispiele 7.24.

(1) $\#X = 0 \Longleftrightarrow X = \emptyset$.
(2) Endliche Mengen sind abzählbar (Offensichtlich).
(3) X ist abzählbar unendlich $\Longleftrightarrow$ X und $\mathbb{N}$ sind gleich mächtig. (Ein Beweis benötigt vollständige Induktion, oder die Peano-Axiome oder irgend etwas, was die natürlichen Zahlen charakterisiert.)
(4) $\mathbb{Z}$ und $\mathbb{N}$ sind gleich mächtig, $\mathbb{Q}$ und $\mathbb{N}$ sind gleich mächtig (Übung: Beweisen Sie dies!). Also sind $\mathbb{Z}$ und $\mathbb{Q}$ abzählbar unendlich.
(5) $\mathbb{R}$ is mächtiger als $\mathbb{N}$ (wird später in Proposition 2.5 in Kapitel 3 bewiesen!)

!ACHTUNG!. Gilt $\#X = \infty$ und $\#Y = \infty$, so folgt daraus **nicht**, dass X und Y gleich mächtig sind.

Ist $f : X \longrightarrow Y$ eine Abbildung und $M \subset X$, so definieren wir

$$f_\#(M) := \{y \in Y \mid \exists x \in M : \, y = f(x)\} = \{f(x) \,\|\, x \in M\}.$$

Man nennt $f_\#(M)$ das Bild von M unter f. Wir erhalten eine Abbildung

$$\begin{array}{ccc} \mathcal{P}(X) & \longrightarrow & \mathcal{P}(Y) \\ M & \mapsto & f_\#(M) \end{array} \tag{7.25}$$

Analog definiert man für $N \subset Y$:

$$f^{\#}(N) := \{x \in X \mid f(x) \in N\}.$$

Man nennt $f^{\#}(N)$ das Urbild von N unter f. Wir erhalten eine Abbildung

$$\begin{array}{ccc} \mathcal{P}(Y) & \longrightarrow & \mathcal{P}(X) \\ N & \mapsto & f^{\#}(N) \end{array} \tag{7.26}$$

Die funktionalen Relationen $f_{\#}$ bzw. $f^{\#}$ werden in der Literatur (und auch in der Linearen Algebra I) oft mit f bzw. f^{-1} bezeichnet. Dies kann zu Missverständnisse führen, da f und $f_{\#}$ verschiedene funktionale Relationen sind und da f^{-1} nur dann eine funktionale Relation ist, wenn f injektiv ist. Wir nutzen deswegen in den nächsten Wochen die Bezeichnungen $f_{\#}$ und $f^{\#}$ wie oben, um Missverständnisse zu vermeiden. Später (ab Analysis II) schreiben wir auch einfach f und f^{-1}.

Übung 7.27. Sei $f : X \longrightarrow Y$ eine Abbildung, $A, B \subset X$ und $M, N \subset Y$. Zeigen Sie

(1) $f_{\#}(A \cap B) \subset f_{\#}(A) \cap f_{\#}(B)$. Geben Sie ein Beispiel an, das zeigt, dass diese Teilmenge manchmal eine echte Teilmenge ist.
(2) $f_{\#}(A \cup B) = f_{\#}(A) \cup f_{\#}(B)$.
(3) $A \subset B \to f_{\#}(A) \subset f_{\#}(B)$.
(4) $f^{\#}(M \cap N) = f^{\#}(M) \cap f^{\#}(N)$.
(5) $f^{\#}(M \cup N) = f^{\#}(M) \cup f^{\#}(N)$.
(6) $M \subset N \to f^{\#}(M) \subset f^{\#}(N)$.
(7) $A \subset f^{\#}(f_{\#}(A))$. Geben Sie ein Beispiel an, das zeigt, dass diese Teilmenge manchmal eine echte Teilmenge ist.

(8) $f_\#(f^\#(M)) \subset M$. Geben Sie ein Beispiel an, das zeigt, dass diese Teilmenge manchmal eine echte Teilmenge ist.

(9) $f^\#(M \smallsetminus N) = f^\#(M) \smallsetminus f^\#(N)$

Die Aussagen (1) und (2) werden auf dem 3. Übungsblatt dieser Vorlesung bewiesen. Die Aussagen (7), (8) und (9) wurden auf dem 2. Übungsblatt der Linearen Algebra gezeigt.

8. Familien

DEFINITION 8.1. Seien I und M Mengen. Eine (I-indizierte) Familie von Elementen von M ist dasselbe wie eine Abbildung von I nach M. Wenn man eine Abbildung $f : I \longrightarrow M$ als Familie interpretiert, so schreiben wir sie als $(f(i))_{i\in I}$ oder als $(f(i) \mid i \in M)$. Die Menge aller I-indizierten Familien von Elementen von M ist also $M^I = \mathrm{Abb}(I, M)$. Eine Familie $(f(i))_{i\in I}$ heißt nicht-leer, falls I nicht die leere Menge ist. Die Menge I heißt Indexmenge. Ist die Indexmenge $I := \mathbb{N}$ oder $I := \mathbb{N}_{>0}$, so nennt man die Elemente in $M^I = \mathrm{Abb}(I, M)$ oft auch M-wertige Folgen oder Folgen in M.

BEISPIELE 8.2. (a) Eine (reell-wertige) Folge ist definiert als $\mathbb{N}$-indizierte Familie von reellen Zahlen. Eine solche Folge ist also eine Abbildung $\mathbb{N} \longrightarrow \mathbb{R}$, die wir als $(a_i)_{i\in\mathbb{N}}$ schreiben. Alternativ darf man diese Folge auch in der Form $(a_0, a_1, a_2, a_3, \ldots)$ schreiben.

(b) Sei M eine Menge. Ein k-Tupel $(a_1, a_2, \ldots, a_k) \in M^k$ definiert dann die Familie $(a_j)_{j\in\{1,2,\ldots,k\}}$. Wir erhalten eine bijektive Abbildung $F : M^k \longrightarrow M^{\{1,\ldots,k\}}$, $(a_1, a_2, \ldots, a_k) \mapsto (a_j)_{j\in\{1,2,\ldots,k\}}$. Die Elemente von M^k wurden ganz anders „definiert“ als die Elemente von $M^{\{1,\ldots,k\}}$. Dennoch haben die Elemente von M^k und die von $M^{\{1,\ldots,k\}}$ nahezu identische Eigenschaften: beide sind durch eine geordnete Liste mit k Elementen von M

beschreibbar. Damit uns die Formalien nun nicht erdrücken, ist es nun sinnvoll so zu tun, als wäre $(a_1, a_2, \ldots, a_k)$ dasselbe wie $(a_j)_{j\in\{1,2,\ldots,k\}}$. Man sagt dazu $(a_1, a_2, \ldots, a_k)$ wird mit $(a_j)_{j\in\{1,2,\ldots,k\}}$ identifiziert. Dadurch identifizieren wir auch automatisch M^k mit $M^{\{1,\ldots,k\}}$. Um zu unterstreichen, dass dieses Identifizieren von obigem F abhängt sagen wir: wir identifizieren M^k mit $M^{\{1,\ldots,k\}}$ vermöge F. Wer das Wort „vermöge“ nicht mag, kann auch „mit Hilfe von“ sagen. [23]

(c) Die Basis eines Vektorraums ist eine Familie von Vektoren, die linear unabhängig sind und den Vektorraum erzeugen.

Bemerkung 8.3. Wenn Sie eine ähnliche Aussage oder Definition wie das obige (b) in der Schule gesehen haben, so stand dort wahrscheinlich nicht eine „Familie von Vektoren“, sondern eine „Menge von Vektoren“. Dies ist aber aus mehreren Gründen problematisch, z.B. erhalten wir mit dieser Mengendefinition, dass

$$\mathcal{B} := \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

eine Basis von $\mathbb{R}^2$ ist. Man beachte, dass $\mathcal{B}$ genau zwei Elemente enthält, da zwischen obigen Mengenklammern zweimal derselbe Vektor steht.
Die Aussage

„Seien v_1 und v_2 Vektoren in $\mathbb{R}^2$. Dann ist $\{v_1, v_2\}$ eine Basis von $\mathbb{R}^2$ genau dann wenn $\{v_1, v_2\}$ linear unabhängig sind.“

[23] Es stellt sich hier unweigerlich die Frage, wieso wir nicht gleich Elemente in $M \times M$ als Abbildungen von $\{1, 2\} \longrightarrow M$ definiert haben. Dies geht aber leider nicht, da wir zur Einführung von Abbildungen bereits wissen mussten, was ein Paar und was ein kartesisches Produkt ist.

ist falsch, denn im Fall $v_1 = v_2 \neq 0$ ist $\{v_1, v_2\} = \{v_1\}$ und somit linear unabhängig. Aber $\{v_1\}$ ist keine Basis von $\mathbb{R}^2$. Korrekt ist hingegen die Aussage:

„Seien v_1 und v_2 Vektoren in $\mathbb{R}^2$. Dann ist $(v_i)_{i\in\{1,2\}}$ eine Basis von $\mathbb{R}^2$ genau dann wenn $(v_i)_{i\in\{1,2\}}$ linear unabhängig ist."

Anschaulich gilt: Familien registrieren Mehrfachnennungen. Über $\{1, 2 \ldots, k\}$ oder über $\mathbb{N}$ indizierte Mengen registrieren auch die Ordnung der Elemente.

Definition 8.4. Ist $(M_i)_{i\in I}$ eine Familie von Mengen, (d.h. jedes M_i ist eine Menge), so definieren wir

$$\bigcup(M_i \mid i \in I) := \bigcup\{M_i \mid i \in I\} = \{x \mid \exists i \in I : x \in M_i\}$$

und falls $(M_i)_{i\in I}$ nicht-leer ist

$$\bigcap(M_i \mid i \in I) := \bigcap\{M_i \mid i \in I\} = \{x \mid \forall i \in I : x \in M_i\}.$$

Wir schreiben auch $\bigcup_{i\in I} M_i$ für $\bigcup(M_i \mid i \in I)$ und $\bigcap_{i\in I} M_i$ für $\bigcap(M_i \mid i \in I)$.

Kartesische Produkte von Mengenfamilien.
Wir haben bereits M^k mit $M^{\{1,\ldots,k\}}$ identifiziert. Die Elemente von $M^{\mathbb{N}}$ sind also eine Art „unendliches kartesisches Produkt". Solche Produkte sind seit Ende des 19. Jahrhunderts in vielen Bereichen der Mathematik ganz wichtig. Dieser Begriff soll deswegen präzisiert werden. Wir wollen auch gleich Produkte in der Form $M_1 \times M_2 \times \ldots$ zulassen, d.h. die M_i (genannt Faktoren) dürfen verschieden sein.

Definition 8.5 (Kartesische Produkte von Mengenfamilien). Gegeben sei eine Familie $(M_i)_{i\in I}$ von Mengen, d.h. gegeben sei eine Indexmenge I und für jedes $i \in I$ eine Menge

M_i. Wir definieren zunächst $M := \bigcup_{i\in I} M_i$. Das kartesische Produkt der Familie $(M_i)_{i\in I}$ ist definiert als

$$\prod_{i\in I} M_i := \left\{(m_i)_{i\in I} \in M^I \mid \forall i \in I : m_i \in M_i\right\}$$

BEISPIELE 8.6. $\prod_{i\in I} M = M^I$

Analog zu Beispiele 8.2 (b) kann man $\prod_{i\in\{1,\dots,k\}} M_i$ mit $M_1 \times \cdots \times M_k$ identifizieren.

Ein M_i in der obigen Definition nennt man den i-ten Faktor oder die i-te Komponente. Für ein fixiertes $i_0 \in I$ nennt man die Abbildung

$$\begin{array}{ccc} \prod_{i\in I} M_i & \longrightarrow & M_{i_0} \\ (m_i)_{i\in I} & \mapsto & m_{i_0} \end{array}$$

die i_0-te kanonische Projektion.

Literatur für das bisherige Kapitel

Das Buch [**12**] ist eine recht ausführliche und leicht lesbare Einführung in die mengentheoretischen Grundlagen der Mathematik. Das Buch [**17**] ist deutlich kürzer und beschränkt sich mehr auf das wesentliche, aber für manche Anfänger vielleicht etwas zu dicht geschrieben. Es fokussiert sich auch mehr auf besonders interssante Aspekte. Die wichtigsten Begriffe in Kürze finden sich auch im Buch von D. Grieser [**18**], das auf dem Skript seiner Analysis-Vorlesung [**19**] aufbaut. Weitere Literatur ist das Buch von Halmos in einer deutschen und englischen Fassung [**20**]. Einen deutlich tiefer gehenden Einstieg in die Mengenlehre findet man in [**13**].

KAPITEL 2

Zahlen

2.11.

Ab jetzt schreiben wir $\to$ oder $\longrightarrow$ für Abbildungen, statt wie bisher das rote Symbol $\longrightarrow$ zu verwenden. Wir schreiben $\{f(x) \mid x \in M\}$ statt $\{f(x) \,\|\, x \in M\}$. Die exakte Bedeutung kann man aus dem Kontext heraus ablesen.

Der einleitende Teil der Vorlesung ist nun beendet und wir – Professoren, Mitarbeiter, Tutoren **und die Hörer der Vorlesung** – sollten uns bemühen, ab jetzt deutlich rigorosere Begründungen zu geben. Wenn aus Zeitgründen nicht alles bewiesen werden kann, so müssen wir klar darstellen, wo genau wir kleinere Lücken lassen, und andeuten, wie diese geschlossen werden können (Eigenes Nachdenken, Anhang, Literatur, Zentralübung, Übungsblätter, etc.).

1. Die natürlichen Zahlen

Die Axiome für die natürlichen Zahlen bilden die Peano-Axiome, benannt nach dem italienischen Mathematiker Guiseppe Peano (1858–1932).

Axiome 1.1 („Axiome“ der natürlichen Zahlen (Peano-Axiome))**.** Die natürlichen Zahlen bestehen aus

- einer Menge $\mathbb{N}$,
- einem (ausgewählten) Element in $\mathbb{N}$, das wir 0 nennen,

- und einer Abbildung $s : \mathbb{N} \longrightarrow \mathbb{N}$, $x \mapsto s(x)$, genannt die Nachfolger-Abbildung.

Wir fordern, dass hierfür folgenden Eigenschaften, genannt die Peano-Axiome, erfüllt sind:

(P1) $0 \notin B(s)$ (das heißt: 0 ist nicht im Bild von s)

(P2) Die Abbildung $s : \mathbb{N} \longrightarrow \mathbb{N}$ ist injektiv.

(P3) (Induktionsaxiom) Erfüllt $T \subset \mathbb{N}$ die Bedingungen $0 \in T$ und $s_{\#}(T) \subset T$, dann gilt bereits $T = \mathbb{N}$.

Alles, was Sie über die natürlichen Zahlen kennen, kann man darauf aufbauen, zum Beispiel Addition, Multiplikation und Exponieren von natürlichen Zahlen, die bekannten Gesetze (zum Beispiel: Kommutativität von Addition und Multiplikation, Assoziativität von Addition und Multiplikation, Distributivgesetz, Potenzgesetz) kann man daraus herleiten. Da dies in der Linearen Algebra I detailiert besprochen werden soll, werden wir dies in der Analysis I nur erwähnen. Für Details verweisen wir auch auf Anhang B, der manche dieser Aspekte ausführt.

Auf den natürlichen Zahlen $\mathbb{N}$ definiert man eine Relation $\leq$ auf $\mathbb{N}$ wie folgt: Für alle $n, m \in \mathbb{N}$ gelte:

$$n \leq m \Longleftrightarrow \exists k \in \mathbb{N} : \; n + k = m.$$

Dies ist eine totale Ordnungsrelation auf $\mathbb{N}$, was auch ein paar hier nicht ausgeführte Beweise erfordert, wenn man alles von Grund auf beweist. Ferner definiert man

$$n \geq m :\Longleftrightarrow m \leq n$$

$$n < m :\Longleftrightarrow n \leq m \wedge \neg n = m$$

$$n > m :\Longleftrightarrow m < n$$

Aus den Peano-Axiomen leitet man zwei wichtige Sätze her: den Satz über vollständige Induktion und den Dedekindschen Rekursionssatz, beide werden in Anhang B bewiesen.

Satz 1.2 (Dedekindscher Rekursionssatz)**.** Sei M eine Menge, $a \in M$ und $g : M \times \mathbb{N} \longrightarrow M$ eine Abbildung.[1] Dann gibt es genau eine Abbildung $f : \mathbb{N} \longrightarrow M$ so dass $f(0) = a$ und

$$\forall n \in \mathbb{N} : \ f(n+1) = g(f(n), n).$$

Wenn man alles sauber aufbauen will, so sind zunächst die Summe $+$, das Produkt $\cdot$ und das Exponieren durch den Rekursionssatz zu definieren, siehe Anhang B. Wir wollen diese Operationen aber als gegeben voraussetzen.

Beispiel 1.3 (Definition der Fakultät). Wir setzen $M := \mathbb{N}$, $a = 1$ und $g(m, n) := (n+1) \cdot m$. Dann nennt man die Abbildung f, die uns Satz 1.2 liefert, die Fakultät. Wir schreiben $n!$ statt $f(n)$. Klammerkonvention: ! wird vor $+$ und $\cdot$ ausgeführt. Somit ist $! : \mathbb{N} \to \mathbb{N}$ die einzige Funktion, die die Eigenschaften

$$\begin{aligned} 0! &= 1 \\ \forall n \in \mathbb{N} \setminus \{0\} : n! &= n \cdot (n-1)! \end{aligned}$$

erfüllt. Es gilt somit $1! = 1$, $2! = 1 \cdot 2$, $3! = 1 \cdot 2 \cdot 3$, $\ldots$, $n! = 1 \cdot 2 \cdot \cdots \cdot n$.

Beispiel 1.4 (Summen- und Produktsymbole). Oft ist es wichtig, Ausdrücke der Form

$$a_1 + \ldots + a_n$$

[1] Man schreibt hier und immer in vergleichbaren Situationen dann $g(m, n)$ statt $g((m, n))$.

exakt zu beschreiben. Hierzu führen wir ein Symbol $\sum_{j=1}^{n}$ ein. Man definiert rekursiv

$$\sum_{j=1}^{0} a_j := 0$$

und für alle $n \in \mathbb{N}$

$$\sum_{j=1}^{n+1} a_j := \left(\sum_{j=1}^{n} a_j \right) + a_{n+1}.$$

Dann ist[2]

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

Analog hierzu

$$\prod_{j=1}^{0} a_j := 1$$

und für alle $n \in \mathbb{N}$

$$\prod_{j=1}^{n+1} a_j := \left(\prod_{j=1}^{n} a_j \right) \cdot a_{n+1}.$$

Insbesondere gilt also $n! = \prod_{j=1}^{n} j$.

Wir nutzen[3] auch folgende Verallgemeinerung: Es seien $k, n \in \mathbb{Z}$ und es gelte $k \leq n+1$. Dann definiert man

$$\sum_{j=k}^{n} a_j := \sum_{j=1}^{n-k+1} a_{j+k-1} = a_k + a_{k+1} + \ldots + a_n, \qquad \prod_{j=k}^{n} a_j := \prod_{j=1}^{n-k+1} a_{j+k-1} = a_k \cdot a_{k+1} \cdot \cdots \cdot a_n.$$

[2]Dies ist genau genommen keine Aussage, sondern die Definition des Ausdrucks $a_1 + \ldots + a_n$, der bisher nicht so klar definiert war, da die Bedeutung der Punkte unklar blieb.

[3]Solange wir $\mathbb{Z}$ noch nicht eingeführt haben, sollte man $k, n \in \mathbb{N}$ annehmen.

Satz 1.5 (Vollständige Induktion)**.** Sei $A(\,\cdot\,)$ eine auf $\mathbb{N}$ definierte Aussageform. Wir setzen voraus, dass Induktionsanfang und Induktionsschritt erfüllt sind:
Induktionsanfang: $A(0)$ ist wahr.
Induktionsschritt: Für alle $n \in \mathbb{N}$ gilt: $(A(n) \Longrightarrow A(n+1)))$.
Dann gilt für alle $n \in \mathbb{N}$ die Aussage $A(n)$.

Im Induktionsschritt nennt man $A(n)$ die Induktionsvoraussetzung.

Beispiel 1.6. Wir wollen zeigen, dass für alle $n \in \mathbb{N}$ die folgende Aussage wahr ist

$$A(n) : \sum_{j=1}^{n} j = \frac{1}{2} n\,(n+1).$$

Induktionsanfang $n = 0$

Für $n = 0$ steht auf der linken Seite (l.S.) nach Definition $\sum_{j=1}^{n} j = 0$.

Auf der rechten Seite (r.S.) haben wir $\frac{1}{2} 0\,(0+1) = 0$. Somit l.S. = r.S..

Deswegen ist $A(0)$ wahr.

Induktionsschritt von n auf $n+1$

Nach Induktionsvorraussetzung gilt bereits die Aussage $A(n)$, also

$$\sum_{j=1}^{n} j = \frac{1}{2} n\,(n+1).$$

Zu zeigen ist $A(n+1)$, also die Aussage

$$\sum_{j=1}^{n+1} j = \frac{1}{2}(n+1)\,\big((n+1)+1\big).$$

Wir berechnen die linke Seite:

$$\sum_{j=1}^{n+1} j = \Big(\sum_{j=1}^{n} j\Big) + (n+1) = \frac{1}{2} n\,(n+1) + (n+1) = \frac{n\,(n+1) + 2\,(n+1)}{2}) = \frac{1}{2}(n+2)\,(n+1).$$

Wir berechnen die rechte Seite:

$$\frac{1}{2}(n+1)\,\big((n+1)+1\big) = \frac{1}{2}(n+1)\,(n+2).$$

Also wieder l.S. = r.S..

Damit haben wir gezeigt, dass $A(n+1)$ aus $A(n)$ folgt.

Schlussfolgerung (wird normalerweise nicht mehr aufgeschrieben, da es bei jeder vollständigen Induktion immer dasselbe Argument ist)

Die Voraussetzungen von Satz 1.5 sind also erfüllt, deswegen gilt für alle $n \in \mathbb{N}$ die Aussage $A(n)$. □

Beispiel 1.7 (Geometrische Summe). Sei $q \in \mathbb{N} \setminus \{1\}$ (oder $q \in \mathbb{R} \setminus \{1\}$ oder $q \in \mathbb{C} \setminus \{1\}$ oder in irgendeinem unitären Ring R, so dass $(1-q)^{-1}$ existiert).

Wir wollen zeigen, dass für alle $n \in \mathbb{N}$ die folgende Aussage wahr ist

$$A(n) : \sum_{j=0}^{n} q^j = \frac{1-q^{n+1}}{1-q}.$$

Wir multiplizieren auf beiden Seiten mit $1-q$ und erhalten daraus die zu $A(n)$ äquivalente Aussageform

$$\tilde{A}(n) : \, (1-q) \sum_{j=0}^{n} q^j = 1 - q^{n+1},$$

die wir durch Induktion über n zeigen wollen.

Induktionsanfang $n = 0$

Für $n = 0$ rechnen wir

$$(1-q)\sum_{j=0}^{0} q^j = (1-q)q^0 = (1-q)\,1 = 1-q^1 = 1-q^{0+1}.$$

Deswegen ist $\tilde{A}(0)$ wahr.

Induktionsschritt von n auf $n+1$

Nach Induktionsvorraussetzung gilt bereits die Aussage $\tilde{A}(n)$, also

$$(1-q)\sum_{j=0}^{n} q^j = 1-q^{n+1}.$$

Zu zeigen ist $\tilde{A}(n+1)$, also die Aussage

$$(1-q)\sum_{j=0}^{n+1} q^j = 1-q^{(n+1)+1}.$$

Wir berechnen:

$$\begin{aligned}(1-q)\sum_{j=0}^{n+1} q^j &= (1-q)\Big(\Big(\sum_{j=0}^{n} q^j\Big)+q^{n+1}\Big) = \underbrace{(1-q)\Big(\sum_{j=0}^{n} q^j\Big)}_{=1-q^{n+1}} + (1-q)\,q^{n+1} \\ &= 1-q^{n+1}+q^{n+1}-q^{n+2} = 1-q^{n+2}.\end{aligned}$$

Also haben wir $\tilde{A}(n+1)$ gezeigt. □

Auch die folgenden elementaren Gesetze der natürlichen Zahlen müssen bei einem logisch vollständigen Aufbau durch vollständige Induktion gezeigt werden. Genau genommen, lange bevor wir obige Beispiele diskutieren können. Aus Zeitgründen werden wir dies nicht ausführen, sondern ebenfalls wieder auf Anhang B des Skripts verweisen.

Satz 1.8. $(\mathbb{N}, +, \cdot)$ erfüllt die folgenden Eigenschaften:

(Aa) **Addition ist assoziativ.**

Für alle $x, y, z \in \mathbb{N}$ gilt

$$(x + y) + z = x + (y + z).$$

(An) **Addition hat neutrales Element.**

Es gibt ein Element $0 \in \mathbb{N}$, so dass für alle $x \in \mathbb{N}$ gilt

$$x + 0 = 0 + x = x.$$

(Diese Null ist unsere bisherige!)

(Ak) **Addition ist kommutativ.**

Für alle $x, y \in \mathbb{N}$ gilt

$$x + y = y + x.$$

(Ma) **Multiplikation ist assoziativ.**

Für alle $x, y, z \in \mathbb{N}$ gilt

$$(x \cdot y) \cdot z = x \cdot (y \cdot z).$$

(Mn) **Multiplikation hat neutrales Element.**

Es gibt ein Element $1 \in \mathbb{N}$, so dass für alle $x \in \mathbb{N}$ gilt

$$x \cdot 1 = 1 \cdot x = x.$$

(Diese Eins ist unsere bisherige!)

(Mk) **Multiplikation ist kommutativ.**

Für alle $x, y \in \mathbb{N}$ gilt

$$x \cdot y = y \cdot x.$$

(AMd) **Addition und Multiplikation erfüllen das Distributivgesetz.**
Für alle $x, y, z \in \mathbb{N}$ gilt

$$x \cdot (y + z) = x \cdot y + x \cdot z$$

$$(y + z) \cdot x = y \cdot x + z \cdot x$$

Eine weitere elementare Eigenschaft von $\mathbb{N}$, die in der Linearen Algebra gezeigt wurde, ist die Wohlordnung von $\mathbb{N}$. Einen Beweis findet man in Anhang B, wo der Satz nochmals als Proposition 3.3 wiederholt ist.

Satz 1.9 (Wohlordnung von $\mathbb{N}$)**.** Jede nicht-leere Teilmenge T von $\mathbb{N}$ besitzt ein Minimum.

Hierbei nennen wir y ein Minimum von T, falls $y \in T$ und $\forall x \in T : y \leq x$. Wenn eine Menge mit Ordungsrelation ein Minimum besitzt, so ist dieses eindeutig; diese Eindeutigkeit folgt direkt aus der Antisymmetrie. Falls ein Minimum y in M existiert, so schreiben wir $y = \min M$. Wir erhalten analog die Definition eines Maximums, wenn $y \leq x$ durch $y \geq x$ ersetzen, und schreiben dann $y = \max M$.

Satz 1.10. Ist M eine abzählbar unendliche Menge, dann sind M und $\mathbb{N}$ gleich mächtig.

Beweis. Da M abzählbar ist, gibt es eine injektive Abbildung $f : M \to \mathbb{N}$. Wir definieren $Q := f_{\#}(M)$. Da f injektiv ist, ist die Abbildung

$$\begin{array}{ccc} M & \longrightarrow & Q \\ x & \mapsto & f(x) \end{array}$$

bijektiv. Also sind M und Q gleich mächtig. Zu zeigen bleibt, dass Q und $\mathbb{N}$ gleich mächtig sind.

Zeichnung mit Q in $\mathbb{N}$ und der Abbildung g, die das nächsthöhere Element sucht.

Wir definieren für $k \in \mathbb{N}$:

$$Q_{>k} := \{x \in Q \mid x > k\}.$$

Behauptung: $\forall k \in \mathbb{N} : \ Q_{>k} \neq \emptyset$.

Begründung: Angenommen $Q_{>k} = \emptyset$, dann wäre $Q \subset \{1, 2, \ldots, k\}$, also wäre Q und somit N endlich. Widerspruch.

Wir definieren $g : Q \longrightarrow Q$ als

$$g(k) := \text{das Minimum von } Q_k.$$

Die Funktion ist wohldefiniert wegen Satz 1.9. Aus der Definition von g ergibt sich insbesondere $g(k) > k$.

Definiere rekursiv mit Hilfe des Dedekindschen Rekursionssatzes die Funktion $h : \mathbb{N} \to Q$ durch

$$h(0) := \text{das Minimum von } Q,$$

$$\forall n \in \mathbb{N} : \ h(n+1) := g(h(n)).$$

Somit gilt $h(n+1) > h(n)$, und deswegen zeigt man leicht (durch eine weitere vollständige Induktion), dass $h : \mathbb{N} \to Q$ injektiv ist.[4] Da die Inklusion $Q \to \mathbb{N}$ injektiv ist, folgt aus dem Satz 7.21 von Schröder-Bernstein, dass Q und $\mathbb{N}$ gleich mächtig sind. □

[4]Die Abbildung $h : \mathbb{N} \to Q$ ist sogar bijektiv, aber wir benötigen dies im folgenden nicht.

2. Etwas Kombinatorik

Um den Umgang mit natürlichen Zahlen zu üben, zeigen wir einige Aussagen, die in der Kombinatorik wichtig sind.

Satz 2.1. Für endliche Mengen N und Q gilt $(\#N) \cdot (\#Q) = \#(N \times Q)$.

Beweis.

Sei also $s := \#N$ und $t := \#Q$. Es gibt also bijektive Abbildung $\varphi : N \to \{1, 2, \ldots, s\}$ [5] und $\psi : N \to \{1, 2, \ldots, t\}$. Dann ist die Abbildung

$$\begin{array}{rcl} \varphi \times \psi : N \times Q & \longrightarrow & \{1, 2, \ldots, s\} \times \{1, 2, \ldots, t\} \\ (n, q) & \mapsto & (\varphi(n), \psi(q)) \end{array}$$

bijektiv und man zeigt auch sofort, dass

$$\begin{array}{rcl} p : \{1, 2, \ldots, s\} \times \{1, 2, \ldots, t\} & \longrightarrow & \{1, 2, \ldots, ts\} \\ (a, b) & \mapsto & a + s(b - 1) \end{array}$$

bijektiv ist. Also

$$\#(N \times Q) = \#(\{1, 2, \ldots, s\} \times \{1, 2, \ldots, t\}) = \#(\{1, 2, \ldots, ts\}) = ts.$$

□ **7.11.**

Satz 2.2. Sei $n \in \mathbb{N} \setminus \{0\}$. Angenommen M hat k Elemente. Dann ist k^n die Anzahl der Elemente von $M^n = \underbrace{M \times \cdots \times M}_{n-\text{mal}}$.

[5] Hier gilt immer: $\{1, 2, \ldots, s\} := \{r \in \mathbb{N} \mid 0 < r \leq s\}$, insbesondere ist es die leere Menge, falls $s = 0$.

Der Beweis ist eine einfach Induktion, die nun jeder selbst durchführen können sollte.

Beweis. Wir zeigen den Satz durch vollständige Induktion über $n \in \mathbb{N} \setminus \{0\}$:

$$A(n) : \#(M^n) = (\#M)^n.$$

Induktionsanfang $n = 1$

Offensichtlich wegen $M^1 = M$ und $k^1 = k$.

Induktionsschritt von n auf $n + 1$

$$\#(M^{n+1}) = \#(M \times M^n) = (\#M) \cdot (\#(M^n)) = (\#M) \cdot (\#M)^n = (\#M)^{n+1}$$

Hierbei folgt das erste Gleichheitsszeichen direkt aus der Definition von M^{n+1}, das zweite aus dem vorangehenden Satz, das dritte nach Induktionsvoraussetzung und das letzte aus der Definition von Potenzen in $\mathbb{N}$. □

Satz 2.3. Angenommen, eine Menge M hat $n \in \mathbb{N}$ Elemente, dann gibt es 2^n Elemente in $\mathcal{P}(M)$.

Auch dies beweist man durch Induktion über n.

Beweis. **Induktionsanfang** $n = 0$

Die leere Menge $\emptyset$ ist die einzige Menge mit 0 Elementen. Sie hat einzige Teilmenge, nämlich $\emptyset$.

Induktionsschritt von n auf $n + 1$

Nun sei $\#M = n + 1$ und sei $x \in M$. Definiere $M' := M \setminus \{x\}$.

Wir definieren $F : \mathcal{P}(M') \times \{0, 1\} \longrightarrow \mathcal{P}(M)$ durch $(A, 0) \mapsto A$ und $(A, 1) \mapsto M \setminus A$.

Die Umkehr-Abbildung ist

$$\begin{aligned} G : \mathcal{P}(M) &\longrightarrow \mathcal{P}(M') \times \{0,1\} \\ B &\mapsto \begin{cases} (B, 0) & \text{falls } x \notin A \\ (M \smallsetminus B, 1) & \text{falls } x \in A \end{cases} \end{aligned}$$

Man überprüft leicht, dass G die Umkehrfunktion von F ist. Also ist F bijektiv. Wir erhalten nach Induktionsvoraussetzung

$$\#\mathcal{P}(M) = \#(M' \times \{0,1\}) \overset{Satz\ 2.1}{=} (\#M')(\#\{0,1\}) = 2^n \cdot 2 = 2^{n+1}.$$

□

Ähnlich zeigt man auch: sind M und N endliche Mengen, dann gilt

$$\#(\mathrm{Abb}(M,N)) = (\#N)^{(\#M)},$$

oder in der anderen Schreibweise

$$\#(N^M) = (\#N)^{(\#M)}.$$

Definition 2.4. Eine Permutation einer Menge M ist eine bijektive Abbildung von M nach M.

Satz 2.5. Angenommen M hat n Elemente. Dann ist $n!$ die Anzahl der Permutationen von M.

Interpretation: $n!$ ist die Anzahl der Möglichkeiten, n unterscheidbare Objekte anzuordnen.

Satz 2.5 ist ein Spezialfall von Satz 2.6, folgt also bereits aus Satz 2.6.

Satz 2.6. Seien M und N endliche Mengen mit $\#M \geq \#N$. Dann ist die Anzahl der injektiven Abbildungen von N nach M gleich

$$\frac{(\#M)!}{(\#M - \#N)!}.$$

Interpretation: $m!/(m-n)! = m \cdot (m-1) \cdot \ldots \cdot (m-n+1)$ ist die Anzahl der Möglichkeiten, aus einer Menge von m unterscheidbaren Objekten n Objekte herauszunehmen und anzuordnen.

Beweis. Wir zeigen durch Induktion über $n \in \mathbb{N}$ für festes $k \in N$.

$$A(n,k): \begin{cases} \text{Gilt } \#M = n+k \text{ und } \#N = n, \text{ dann gibt es} \\ (n+k)!/k! \text{ viele injektive Abbildungen } N \longrightarrow M. \end{cases}$$

Induktionsanfang $n = 0$.

Also $N = \emptyset$. Dann ist $\emptyset : \emptyset \to M$ die einzige Abbildung von N nach M. Sie ist injektiv. Andererseits gilt dann $(n+k)!/k! = k!/k! = 1$.

Induktionsschritt von n auf $n+1$.

Sei $\#N = n+1$ und $\#M = n+1+k$. Wir wählen ein $a \in N$. Es gibt $n+k+1$ Möglichkeiten, dem Element a ein Element von M zuzuweisen. Sobald festgelegt ist, dass a auf ein $b \in M$ abgebildet wird, kann noch eine injektive Abbildung von $N \setminus \{a\}$ auf $M \setminus \{b\}$ gewählt werden. Hierfür gibt es nach Induktionsvoraussetzung $(n+k)!/k!$ Möglichkeiten. Somit ist die Anzahl der injektiven Abbildungen von N nach M durch

$$(n+k+1) \cdot \frac{(n+k)!}{k!} = \frac{(n+k+1)\,(n+k)!}{k!} = \frac{(n+k+1)!}{k!}$$

gegeben. □

Satz 2.7. Seien $m, k \in \mathbb{N}$. Die Anzahl der k-elementigen Teilmengen von einer Menge mit m Elementen, $m \geq k$ ist:

$$\binom{m}{k} := \frac{m!}{k!\,(m-k)!}.$$

Die oben definierte Zahl $\binom{m}{k}$ heißt Binomialkoeffizient.

Beweis. Sei N eine Teilmenge mit k Elementen von M, $m = \#M$. Die Anzahl der bijektiven Abbildungen $\{1, 2, \ldots, k\} \longrightarrow N$ ist $k!$. Jede solche bijektive Abbildung $\{1, 2, \ldots, k\} \longrightarrow N$ ergibt eine injektive Abbildung $\{1, 2, \ldots, k\} \longrightarrow M$, und umgekehrt ergibt jede injektive Abbildung $f : \{1, 2, \ldots, k\} \longrightarrow M$, $x \mapsto f(x)$ eine bijektive Abbildung $f : \{1, 2, \ldots, k\} \longrightarrow B(f)$, $x \mapsto f(x)$. Ist deswegen r die Anzahl der k-elementigen Teilmengen von M, dann ist $r \cdot k!$ die Anzahl der injektive Abbildungen $f : \{1, 2, \ldots, k\} \longrightarrow M$. Nach vorigem Satz also $r \cdot k! = m!/(m-k)!$ und daraus ergibt sich

$$r = \binom{m}{k}.$$

□

Offensichtlich gilt

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

Wir setzen zudem $\binom{m}{k} := 0$ für $m, k \in \mathbb{Z}$, falls $k < 0$ oder $k > m$.

Sei $M = \{1, 2, \ldots, m\}$. Dann hat M wegen Satz 2.3 2^m Teilmengen. Da es $\binom{m}{k}$ Teilmengen mit k Elementen gibt, folgt:

FOLGERUNG 2.8. Für $m \in \mathbb{N}$ gilt

$$2^m = \sum_{k=0}^{m} \binom{m}{k}.$$

□

SATZ 2.9. Es gelte $m, k \in \mathbb{Z}$ und $m > 0$. Dann gilt:

$$\binom{m-1}{k-1} + \binom{m-1}{k} = \binom{m}{k}.$$

Beweis. In den Fällen $m = k \in \mathbb{N} \setminus \{0\}$ und den Fällen $k = 0$, $m \in \mathbb{N} \setminus \{0\}$ haben wir $\binom{m}{k} = 1$, ein Summand ist 0 und der andere 1. Die Behauptung ist trivialerweise richtig, falls $k < 0$ oder $k > m$.

In den übrigen Fällen rechnen wir

$$\begin{aligned}\binom{m-1}{k-1}+\binom{m-1}{k} &= \frac{(m-1)!}{(k-1)!\,(m-1-k+1)!}+\frac{(m-1)!}{k!\,(m-1-k)!}\\ &= \frac{k\,(m-1)!}{k\,(k-1)!\,(m-k)!}+\frac{(m-k)(m-1)!}{k!\,(m-k)\,(m-1-k)!}\\ &= \frac{(k+m-k)\,(m-1)!}{k!\,(m-k)!}\\ &= \frac{m!}{k!\,(m-k)!}\end{aligned}$$

□

Pascalsches Dreieck

Bild eines Pascalschen Dreiecks

Alternativ auch hier zu sehen.

Satz 2.10 (Binomische Formel)**.** Seien $x, y \in \mathbb{R}$ oder in einem beliebigen kommutativen Ring[6] mit 1. Sei $n \in \mathbb{N}$.

Dann gilt

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Hierbei nutzen wir die Definition: $x^0 = y^0 = (x+y)^0 = 1$.

[6]Wenn Sie diesen Begriff des nächsten Abschnitts noch nicht kennen, ignorieren Sie bitte diesen Fall und betrachten Sie zunächst den Fall $x, y \in \mathbb{N}$ oder $x, y \in \mathbb{R}$.

Beweis. **Induktionsanfang** $n = 0$

Auf beiden Seiten steht dann 1.

Induktionsschritt von $n-1$ auf n für $n \in \mathbb{N}_{>0}$

$$\begin{aligned}
(x+y)^n &= (x+y)\,(x+y)^{n-1} \\
&= (x+y)\,\left(\sum_{k=0}^{n-1}\binom{n-1}{k}x^k y^{n-k-1}\right) \\
&= \sum_{k=0}^{n-1}\binom{n-1}{k}x^{k+1}y^{n-k-1} + \sum_{k=0}^{n-1}\binom{n-1}{k}x^k y^{n-k} \\
&= \sum_{k=1}^{n}\binom{n-1}{k-1}x^k y^{n-k} + \sum_{k=0}^{n-1}\binom{n-1}{k}x^k y^{n-k} \\
&= \sum_{k=0}^{n}\binom{n-1}{k-1}x^k y^{n-k} + \sum_{k=0}^{n}\binom{n-1}{k}x^k y^{n-k} \\
&= \sum_{k=0}^{n}\left(\binom{n-1}{k-1} + \binom{n-1}{k}\right)x^k y^{n-k} \\
&= \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}.
\end{aligned}$$

Hierbei nutzen wir die Induktionsvoraussetzung in der zweiten Zeile, das Distributivgesetz in der dritten, in der vierten Zeile wird im ersten Summand k durch $k+1$ ersetzt. Danach

nutzen wir

$$\binom{n-1}{-1} = \binom{n-1}{n} = 0,$$

dann wieder das Distributivgesetz und letztendlich Satz 2.9.

□

3. Die ganzen Zahlen

Der nächste Schritt ist nun, die Menge der uns bekannten Zahlen Schritt für Schritt zu erweitern:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

Wir wollen diese Erweiterungen durch Eigenschaften charakterisieren.

Wir schreiben

$$\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\} \supset \mathbb{N}$$

für die ganzen Zahlen.[7]

Die Addition und die Multiplikation kann man zu Abbildungen

$$+ : \mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z}, \quad (a, b) \mapsto a + b$$
$$\cdot : \mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z}, \quad (a, b) \mapsto a \cdot b$$

[7]Wir wollen wiederum $\mathbb{Z}$ nur durch charakterisierende Eigenschaften beschreiben, aber kein Modell konstruieren, das diese Eigenschaften erfüllt. Ein derartige Modell kann man zum Beispiel konstruieren in dem man auf $\mathbb{N} \times \mathbb{N}$ die Äquivalenzrelation $(m, n) \sim (a, b) \iff m + b = n + a$ einführt; dann sind die Element von $\mathbb{Z}$ die Äquivalenzklassen dieser Relation. Mit einer ähnlichen Konstruktion kann man aus den ganzen Zahlen die rationalen Zahlen $\mathbb{Q}$ konstruieren. Diese beiden Konstruktion sollen in der Linearen Algebra erklärt werden.

erweitern, so dass die Eigenschaften (Aa), (Ak), (An), (Ma), (Mk), (Mn) und (AMd) weiterhin gelten, wenn man dort $\mathbb{N}$ durch $\mathbb{Z}$ ersetzt. Außerdem gilt

(Ai) **Addition hat inverse Elemente.**
Zu jedem $x \in \mathbb{Z}$ gibt es ein $y \in \mathbb{Z}$, so dass

$$x + y = y + x = 0.$$

Man nennt y das Inverse von x bezüglich der Addition und schreibt normalerweise $-x$ anstelle von y.

Schreibweise

$$x - y := x + (-y).$$

Hinzu kommt eine Eigenschaft, die anschaulich besagt, dass die ganzen Zahlen „die kleinste Erweiterung“ der natürlichen Zahlen sind, die diese Eigenschaften erfüllt. Präzise ausgedrückt:

(MEZ) **Minimale Erweiterung.**
Ist T eine Teilmenge von $\mathbb{Z}$ mit $\mathbb{N} \subset T$ und gilt

$$\forall x, y \in T : \; x + y \in T$$

und

$$\forall x \in T : \; -x \in T,$$

dann gilt bereits $T = \mathbb{Z}$.

Auch die Ordnung $\leq$ setzt sich auf $\mathbb{Z}$ fort und die Eigenschaften (1)–(5) in Lemma 3.1in Anhang B.3 gelten weiter.

Eine mit Addition und Mulitplikation versehene Menge, die die Eigenschaften (Aa), (Ak), (An), (Ai), (Ma), und (AMd) erfüllt, nennt man Ring. Gilt zusätzlich (Mn), so spricht man von einem Ring mit Eins. Ein Ring, der zusätzlich (Mk) erfüllt, nennt man einen kommutativen Ring.

Somit ist $(\mathbb{Z}, +, \cdot)$ ein kommutativer Ring mit Eins. Ein Ring $(R, \widetilde{+}, \widetilde{\cdot})$, der die natürlichen Zahlen enthält [8] und (MEZ) erfüllt, stimmt „im wesentlichen“ mit den ganzen Zahlen überein. Präziser: es gibt eine bijektive Abbildung $b : \mathbb{Z} \longrightarrow R$ mit

$$\forall x, y \in \mathbb{Z} : \; b(x + y) = b(x) \widetilde{+} b(y) \text{ und } b(x \cdot y) = b(x) \widetilde{\cdot} b(y).$$

Solche eine bijektive Abbildung, die Addition und Multiplikation erhält, nennt man einen (Ring)-Isomorphismus oder Isomorphismus von Ringen. Wir sagen: $(R, \widetilde{+}, \widetilde{\cdot})$ ist isomorph (als Ring) zu $(\mathbb{Z}, +, \cdot)$ und schreiben $(R, \widetilde{+}, \widetilde{\cdot}) \cong (\mathbb{Z}, +, \cdot)$.

Notation. Den Ring der ganzen Zahlen schreiben wir zumeist als $\mathbb{Z}$ an Stelle von $(\mathbb{Z}, +, \cdot)$, da die Addition und die Multiplikation aus dem Kontext heraus klar sind.

4. Die rationalen Zahlen

Die rationalen Zahlen sind[9]

$$\mathbb{Q} := \left\{ \frac{z}{n} \,\middle|\, z \in \mathbb{Z},\ n \in \mathbb{N}_{>0} \right\}.$$

[8] Gemeint ist hier: als Unterring enthält, das heißt es wird gefordert, dass sich die bereits auf $\mathbb{Z} \times \mathbb{Z}$ definiterten Operatoren $+$ und $\cdot$ sich zu $\widetilde{+}$ und $\widetilde{\cdot}$ fortsetzen.

[9] An dieser Stelle müsste eigentlich gesagt werden, wie $\frac{z}{n}$ definiert ist. Dies wurde in der Linearen Algebra diskutiert und wird deswegen hier nicht erklärt.

Hierbei gilt

$$\frac{z}{n} = \frac{y}{m} \Longleftrightarrow zm = yn.$$

Die Addition und die Multiplikation setzen sich zu Abbildungen

$$+ : \mathbb{Q} \times \mathbb{Q} \longrightarrow \mathbb{Q}, \quad (a, b) \mapsto a + b$$
$$\cdot : \mathbb{Q} \times \mathbb{Q} \longrightarrow \mathbb{Q}, \quad (a, b) \mapsto a \cdot b$$

fort, und die Eigenschaften (Aa), (Ak), (An), (Ai), (Ma), (Mk), (Mn) und (AMd) gelten weiterhin für $(\mathbb{Q}, +, \cdot)$. Außerdem gilt

(Mi) **Multiplikation hat inverse Elemente.**
Zu jedem $x \in \mathbb{Q} \setminus \{0\}$ gibt es ein $y \in \mathbb{Q}$, so dass

$$x \cdot y = y \cdot x = 1.$$

Man nennt y das Inverse von x bezüglich der Multiplikation und schreibt normalerweise x^{-1} anstelle von y.

9.11.

Anschaulich: $\mathbb{Q}$ ist die kleinste Erweiterung von $\mathbb{N}$, die diese Eigenschaften hat.
Präzise Bedeutung:

(MEQ) **Minimale Erweiterung.**
Ist T eine Teilmenge von $\mathbb{Q}$ mit $\mathbb{N} \subset T$ und gilt

$$\forall x, y \in T : \; x - y \in T$$
$$\forall x \in T : \forall x \in T \setminus \{0\} : \; x \cdot y^{-1} \in T,$$

dann gilt bereits $T = \mathbb{Q}$.

Auch die Ordnung setzt sich fort. Die Ordnung auf $\mathbb{Q}$ ist total.

Mit Addition und Multiplikation versehene Mengen mit mindestens 2 Elementen nennt man Körper, falls die Eigenschaften (Aa), (Ak), (An), (Ai), (Ma), (Mk), (Mn), (Mi) und (AMd) erfüllt sind.

Also ist $(\mathbb{Q}, +, \cdot)$ ein Körper. Wie bei den ganzen Zahlen schreiben wir oft $\mathbb{Q}$ für $(\mathbb{Q}, +, \cdot)$. Jeder Körper, der die natürlichen Zahlen enthält und (MEQ) erfüllt, ist isomorph zum Körper der rationalen Zahlen.

In jedem Körper K gilt $0 \neq 1$. Denn angenommen, wir hätten $0 = 1$, so folgt für alle $x \in K$: $x = 1 \cdot x = 0 \cdot x = 0$. Also $K = \{0\}$. Dies widerspricht der Forderung, dass K mindestens zwei Elemente hat.

Übung 4.1. Versehe $F_2 := \{0, 1\}$ mit der folgenden Addition und Multiplikation:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Zeigen Sie, dass $(F_2, +, \cdot)$ ein Körper ist. Man nennt ihn den Körper mit 2 Elementen. Ist $(\widetilde{K}, \widetilde{+}, \widetilde{\cdot})$ ein Körper, und hat $\widetilde{K}$ genau zwei Elemente, dann ist $(\widetilde{K}, \widetilde{+}, \widetilde{\cdot})$ isomorph zu $(F_2, +, \cdot)$.

Die Ordnung auf $\mathbb{Z}$ setzt sich zu einer Ordnung auf $\mathbb{Q}$ wie folgt fort:

Wir definieren zunächst, wann eine Zahl größer oder gleich Null ist. Für alle $m \in \mathbb{Z}$ und alle $n \in \mathbb{N}_{>0}$ definiere

$$0 \leq \frac{m}{n} \quad :\Longleftrightarrow \quad 0 \leq m \cdot n$$

Dann definieren wir die Relation $\leq$ für beliebige Zahlen $a, b \in \mathbb{Q}$:

$$a \leq b :\Longleftrightarrow 0 \leq b - a\,.$$

5. Geordnete Körper

DEFINITION 5.1. Ein geordneter Körper ist ein Quadrupel $(K, +, \cdot, \leq)$, so dass die folgenden Eigenschaften erfüllt sind:

(a) $(K, +, \cdot)$ ist ein Körper,
(b) $\leq$ ist eine totale Ordnung auf K,
(c) $\forall x, y, z \in K : (x \leq y \to x + z \leq y + z)$,
(d) $\forall x, y, z \in K : ((x \leq y \wedge 0 \leq z) \to x \cdot z \leq y \cdot z)$.

BEISPIEL 5.2. $(\mathbb{Q}, +, \cdot, \leq)$ ist ein geordneter Körper.

Nicht-Beispiel: Sei $(F_2, +, \cdot)$ der Körper mit zwei Elementen aus Übung 4.1. Dann gibt es **keine** totale Ordnung $\leq$ auf F_2, so dass $(F_2, +, \cdot, \leq)$ ein geordneter Körper ist. Vergleiche Übungsblatt 4, Aufgabe 3.

Im Rest dieses Abschnitts sei $(K, +, \cdot, \leq)$ ein geordneter Körper.

NOTATION 5.3.

$$\begin{aligned} x < y &:\iff x \leq y \wedge x \neq y \iff \neg(y \leq x) \\ x \geq y &:\iff y \leq x \\ x > y &:\iff y < x \\ x \text{ ist positiv} &:\iff x > 0 \\ x \text{ ist negativ} &:\iff x < 0 \\ K_{\leq a} &:= \{x \in K \mid x \leq a\} \end{aligned}$$

$K_{\geq a}$, $K_{<a}$, $K_{>a}$ analog.

ÜBUNG 5.4. Seien $a, x, y, z \in K$. Zeigen Sie, dass aus den Eigenschaften (a)–(d) der obigen Definition folgt:

(a) a ist genau dann positiv, wenn $-a$ negativ ist.
(b) $1 > 0$
(c) Ist $z > 0$, dann auch $z^{-1} > 0$.
(d) Ist $z < 0$, dann gilt auch $z^{-1} < 0$ und dann gilt für alle $x, y \in K$:

$$x \leq y \quad \leftrightarrow \quad x \cdot z \geq y \cdot z.$$

Beweis: Teile (a) und (b) sind Teil von Übungsblatt 4, Aufgabe 2. Teile (c) und (d) sind ähnlich zu zeigen.

Zu (c): Sei nun $z > 0$. Wir zeigen durch Widerspruch $z^{-1} > 0$. Wir nehmen hierzu an, dass $z^{-1} \leq 0$. Dann folgt mit Lemma 5.1 (d)

$$1 = z^{-1}z \leq 0,$$

was offensichtlich $1 > 0$ widerspricht.

Zu (d): $z < 0 \iff z^{-1} < 0$ ist nun offensichtlich und man zeigt (d) durch Multiplikation mit $-z$ bzw. $-z^{-1}$ unter Verwendung von Lemma 5.1 (d). □

Wir definieren rekursiv eine Abbildung $i_{\mathbb{N}} : \mathbb{N} \longrightarrow K$, $0_{\mathbb{N}} \mapsto 0_K$, $i_{\mathbb{N}}(n+1) = i_{\mathbb{N}}(n) + 1$.

Lemma 5.5.

(a) $i_{\mathbb{N}}$ erhält Addition und Multiplikation, d.h. für alle $n, m \in \mathbb{N}$ gilt:

$$i_{\mathbb{N}}(n+m) = i_{\mathbb{N}}(n) + i_{\mathbb{N}}(m) \qquad i_{\mathbb{N}}(n \cdot m) = i_{\mathbb{N}}(n) \cdot i_{\mathbb{N}}(m) \tag{5.6}$$

(b) $\forall n \in \mathbb{N}_{>0} : \; i_{\mathbb{N}}(n) > 0$,
(c) $i_{\mathbb{N}} : \mathbb{N} \longrightarrow K$ ist injektiv,

Beweis.

(a) durch Induktion über $m \in \mathbb{N}$.
(b) und (c) sind Teil von Übungsblatt 4, Aufgabe 2.

Sprechweise Wenn die Gleichungen (5.6) gelten, sagen wir: $i_{\mathbb{N}}$ erhält Addition und Multiplikation.

Lemma 5.7. Sei $(K, +, \cdot, \leq)$ ein geordneter Körper. Dann gibt es eine injektive Abbildung $i_{\mathbb{Q}} : \mathbb{Q} \longrightarrow K$, die $i_{\mathbb{N}}$ fortsetzt (das heißt: $i_{\mathbb{Q}}|_{\mathbb{N}} = i_{\mathbb{N}}$) und so dass $i_{\mathbb{Q}}$ Addition und Multiplikation erhält. Außerdem erhält $i_{\mathbb{Q}}$ die Ordnung, das heißt es gilt für $a, b \in \mathbb{Q}$

$$a \leq b \iff i_{\mathbb{Q}}(a) \leq i_{\mathbb{Q}}(b).$$

Beweisskizze: Für $m \in \mathbb{N}$, $n \in \mathbb{N}_{>0}$ definieren wir

$$i_{\mathbb{Q}}\left(\frac{m}{n}\right) := \frac{i_{\mathbb{N}}(m)}{i_{\mathbb{N}}(n)}$$

$$i_{\mathbb{Q}}\left(\frac{-m}{n}\right) := -\frac{i_{\mathbb{N}}(m)}{i_{\mathbb{N}}(n)}$$

$$i_{\mathbb{Q}}\left(\frac{0}{n}\right) := 0$$

Man muss nun zeigen, dass diese Definition Sinn ergibt. Wir dürfen einer rationalen Zahl nicht zwei verschiedene Werte zuweisen. Zu überprüfen ist hierzu

$$\frac{m}{n} = \frac{p}{q} \quad \Longrightarrow \quad \frac{i_{\mathbb{N}}(m)}{i_{\mathbb{N}}(n)} = \frac{i_{\mathbb{N}}(p)}{i_{\mathbb{N}}(q)}$$

für alle $m, n, p, q \in \mathbb{N}_{>0}$. Die entsprechenden Aussagen sind für $m \leq 0$, $p \leq 0$ zu überprüfen. Man sagt dann, die Abbildung $i_{\mathbb{Q}}$ ist wohldefiniert. Danach kann man die anderen Eigenschaften leicht beweisen.

Wir identifizieren nun $\mathbb{Q}$ mit $B(i_{\mathbb{Q}})$ vermöge $i_{\mathbb{Q}}$, das heißt wir machen keinen Unterschied mehr zwischen $r \in \mathbb{Q}$ und $i_{\mathbb{Q}}(r) \in K$.

Definition 5.8. Ein geordneter Körper $(K, +, \cdot, \leq)$ heißt archimedisch geordneter Körper, [10] falls

$$\forall a \in K : \exists n \in \mathbb{N} : a \leq n.$$

Beispiele 5.9. (a) $\mathbb{Q}$ ist archimedisch: Man zeigt leicht für $m, n \in \mathbb{N}_{>0}$: $\frac{m}{n} \leq m$ und $-\frac{m}{n} \leq m$ und $\frac{0}{n} \leq 1$.

[10] Man sagt manchmal auch: Der Körper erfüllt das archimedische Axiom. Diese Sprechweise vermeiden wir aber, da die archimedische Eigenschaft in dem von uns gewählten Zugang kein Axiom ist.

(b) In Satz 6.6 werden wir zeigen, dass jeder geordnete Körper, der die Supremumseigenschaft erfüllt, archimedisch ist. Insbesondere: Sobald wir wissen, „was $\mathbb{R}$ ist“, wissen wir auch, dass $\mathbb{R}$ archimedisch geordnet ist.

(c) Sei dann K ein Körper mit $\mathbb{Q} \subset K \subset \mathbb{R}$ mit der Addition, Multiplikation und Ordung von $\mathbb{R}$. Dann ist auch K ein archimedisch geordneter Körper. Zum Beispiel für $q \in \mathbb{Q}_{>0}$:

$$K := \mathbb{Q}[\sqrt{q}] := \{a + b\sqrt{q} \mid a, b \in \mathbb{Q}\}\,.$$

(d) Der Körper der rationalen Funktionen mit rationalen Koeffizienten [11]

$$K := \left\{ x \mapsto \frac{p(x)}{q(x)} \;\middle|\; p \text{ und } q \text{ sind Polynome mit rationalen Koeffizienten, } q \neq 0 \right\}.$$

Wir definieren Addition und Multiplikation wie bei den rationalen Zahlen:

$$\frac{p(x)}{q(x)} + \frac{\tilde{p}(x)}{\tilde{q}(x)} := \frac{p(x)\tilde{q}(x) + \tilde{p}(x)q(x)}{q(x) \cdot \tilde{q}(x)} \qquad \frac{p(x)}{q(x)} \cdot \frac{\tilde{p}(x)}{\tilde{q}(x)} := \frac{p(x) \cdot \tilde{p}(x)}{q(x) \cdot \tilde{q}(x)}.$$

Dann ist $(K, +, \cdot)$ ein Körper. Ein Polynom p wie oben schreiben wir im Fall $p \neq 0$ als $p(x) = a_{0,p} + a_{1,p}x + a_{2,p}x^2 \ldots + a_{k,p}x^k$, $k \in \mathbb{N}_0$, $a_{k,p} \neq 0$. Dann heißt $\deg(p) := k \in \mathbb{N}_0$

[11] Was wird hier eigentlich mit $x \mapsto \frac{p(x)}{q(x)}$ gemeint? Dies ist etwas subtil, es gibt verschiedene Möglichkeiten. 1.) Man kann es als Abbildung von $\mathbb{Q} \setminus q^{\#}(\{0\})$ nach $\mathbb{Q}$ auffassen, falls p und q keine gemeinsamen Nullstellen haben. Bei dieser Interpretation muss man den Definitionsbereich bei Addition und Multiplikation in naheliegnder Art und Weise evtl. durch einzelne Punkte ergänzen. 2.) Sobald wir die Körper $\mathbb{R}$ und $\mathbb{C}$ kennen, könnte man es in ähnlicher Weise auch als Abbildung $\mathbb{R} \setminus q^{\#}(\{0\}) \longrightarrow \mathbb{R}$ oder als Abbildung $\mathbb{C} \setminus q^{\#}(\{0\}) \longrightarrow \mathbb{C}$ auffassen, auch Kombinationen wie $\mathbb{Q} \setminus q^{\#}(\{0\}) \longrightarrow \mathbb{R}$ wären denkbar. 3.) Die mathematisch sinnvollste Interpretation des Ausdrucks ist aber eine andere (die formale Interpretation), die ich leider derzeit noch nicht erklären kann, und die Sie später in der Linearen Algebra kennen lernen werden. Da wir Koeffizienten in $\mathbb{Q}$ haben, ist es gar nicht so wichtig, wie wir den Ausdruck formal nun verstehen wollen. Wenn die Koeffizienten in $\mathbb{Q}$ durch Koeffizienten in einem endlichen Körper (wie zum Beispiel F_2) ersetzt werden, ist dieser Unterschied aber ganz erheblich.

der Grad des Polynoms. Wir definieren nun

$$P := \left\{ \frac{p(x)}{q(x)} \in K \;\middle|\; a_{\deg(p),p} \cdot a_{\deg(q),q} > 0 \right\}$$

Weiter definieren für $r, s \in K$:

$$r < s \quad :\Longleftrightarrow \quad s - r \in P.$$

Zu zeigen, dass $(K, +, \cdot, \leq)$ ein geordneter Körper ist, ist einige Arbeit, aber nicht sehr schwierig.

Angenommen, der Körper wäre archimedisch geordnet. Dann gibt es ein $n \in \mathbb{N}$ mit $x \leq n < n+1$. Dann wäre also $x - (n+1) < 0$. Dies ist aber ein Widerspruch zu $x - (n+1) = \frac{x-(n+1)}{1} \in P$. Also ist dieser geordnete Körper nicht archimedisch.

Sei weiterhin $(K, +, \cdot, \leq)$ ein geordneter Körper.

Definition 5.10. Die Betragsfunktion $|\cdot| : K \longrightarrow K$ ist definiert durch

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

Man nennt $|a|$ auch den Absolutbetrag von a.

Lemma 5.11.

(a) $|a| \geq 0$; $|a| = 0 \Longleftrightarrow a = 0$;
(b) $|-a| = |a|$;
(c) $|a \cdot b| = |a| \cdot |b|$; $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$ falls $b \neq 0$;
(d) $|a^2| = |a|^2 = a^2 \geq 0$

(e) $|a+b| \leq |a| + |b|$;

(f) $|a_1 + \cdots + a_k| \leq |a_1| + \cdots + |a_k|$;

(g) $|a-b| \geq \Big||a| - |b|\Big|$

Beweis. (a)-(d) folgen direkt aus den Definitionen 5.1 und 5.10.

(e): $\pm a \leq |a|$ und $\pm b \leq |b|$ ergeben $\pm(a+b) \leq |a| + |b|$.

(f): durch Induktion.

(g): Wir setzen $\tilde{a} := a - b$. Dann gilt

$$|\underbrace{\tilde{a} + b}_{a}| \leq |\tilde{a}| + |b| = |a-b| + |b|$$

Also

$$|a-b| \geq |a| - |b| \tag{5.12}$$

Durch Vertauschen von a und b erhalten wir

$$|a-b| = |b-a| \geq |b| - |a|. \tag{5.13}$$

Aus (5.12) und (5.13) erhalten wir das zu Beweisende. □

Satz 5.14. Seien $n \in \mathbb{N}_{>0}$; $a_1, \ldots, a_n \in K_{\geq 0}$ gegeben. Dann gilt

$$a_1 \cdot a_2 \cdot \ldots \cdot a_n \leq \left(\frac{a_1 + a_2 + \ldots + a_n}{n}\right)^n.$$

Bemerkung 5.15. Wenn wir in K die n-te Wurzel ziehen können (in $\mathbb{Q}$ nicht erlaubt!), bedeutet dies

$$\underbrace{\sqrt[n]{a_1 \cdot a_2 \ldots a_n}}_{\text{geometrisches Mittel}} \quad \leq \quad \underbrace{\frac{a_1 + a_2 + \ldots + a_n}{n}}_{\text{arithmetisches Mittel}}.$$

Beweis des Satzes. Wir beweisen den Satz durch Induktion nach n.

Induktionsanfang: $n = 1$ ist offensichtlich richtig.

Induktionsschritt von $n-1$ nach n für $n \geq 2$:

Seien $a_1, \ldots, a_n \in K_{\geq 0}$ gegeben.

O.B.d.A. $a_1 = \min\{a_1, \ldots, a_n\}$, $a_2 = \max\{a_1, \ldots, a_n\}$.

$M := \frac{a_1 + \ldots + a_n}{n}$. $\Longrightarrow a_1 \leq M \leq a_2$

Wir setzen $d := a_1 + a_2 - M$. Dann

$$Md - a_1 a_2 = M\left((a_1 + a_2) - M\right) - a_1 a_2 = (a_2 - M)(M - a_1) \geq 0$$

$$\Longrightarrow a_1 \cdot a_2 \leq Md \tag{5.16}$$

Das arithmetische Mittel[12] von $d, a_3, a_4, \ldots, a_n$ ist M, denn

$$d + a_3 + a_4 + \ldots + a_n = a_1 + a_2 + \ldots + a_n - M = (n-1)M.$$

Es folgt

$$a_1 a_2 \cdots a_n \overset{(5.16)}{\leq} M \underbrace{da_3 \cdots a_n}_{\substack{\leq M^{n-1} \\ \text{nach Annahme}}} = M^n$$

□

6. Die reellen Zahlen

Da die reellen Zahlen $\mathbb{R}$ für die Analysis eine zentrale Rolle spielen, wollen wir hier wieder formal vollständiger vorgehen.

[12]Im Fall $n = 2$ bedeutet ist also das arithmetische Mittel von $d, a_3, \ldots, a_n$ als das artithmetische Mittel von d zu lesen, das heißt es ist d. Es git dann auch $M = d$.

6.1. Unzulänglichkeit von $\mathbb{Q}$. Wieso reichen uns die rationalen Zahlen nicht aus?

Problem 1 (Algebraisch):

Sei p eine Primzahl. Dann ist $x^2 = p$ in $\mathbb{Q}$ nicht lösbar.

Angenommen es gäbe eine Lösung $x = m/n$, $m \in \mathbb{Z}$, $n \in \mathbb{N}_{>0}$. O.B.d.A. m, n teilerfremd.

Wir erhalten $\frac{m^2}{n^2} = x^2 = p$, also $m^2 = pn^2$.

Es folgt $p|m$ und daraus $p^2|m^2 = pn^2$, was wiederum $p|n$ impliziert. Also haben wir einen Widerspruch zur Teilerfremdheit von m und n.

Problem 2 (Geometrisch):

Bild von einem Kreis

$$\pi = \frac{\text{Umfang}}{\text{Durchmesser}} = 3,14159265... \notin \mathbb{Q}$$

Wichtige geometrische Funktionen können in $\mathbb{Q}$ nicht beschrieben werden, zum

Bild: Quadrat mit Diagonale runtergedreht

Problem 3 (Physikalisch):

Bild mit Pendel

Ein Pendel bewegt sich auf einem Kreis. Dazu muss es Punkte passieren, die keine rationalen Koordinaten haben, zum Beispiel den Punkt

$$\begin{pmatrix} -\cos(60°) \\ \sin(60°) \end{pmatrix}, \qquad \sin(60°) = \frac{\sqrt{3}}{2} \notin \mathbb{Q}\,.$$

14.11

6.2. Die Supremumseigenschaft. In diesem Abschnitt werden wir eine Eigenschaft kennenlernen, die im Körper der rationalen Zahlen $\mathbb{Q}$ nicht gilt, aber im Körper der reellen Zahlen $\mathbb{R}$. Zunächst ein paar Vorbemerkungen.

Wiederholung: Sei $\leq$ eine partielle Ordnung auf einer Menge M. Sei $x \in A \subset M$.

$$x = \min A \quad \Longleftrightarrow \quad \forall y \in A : x \leq y.$$

(Maximum analog mit $\geq$).

Gegeben sei eine totale Ordnung auf einer endlichen Menge A, dann besitzt A ein Maximum und ein Minimum.

DEFINITION 6.1. Sei $\leq$ eine (partielle) Ordnung auf einer Menge M. Eine Teilmenge $A \subset M$ heißt nach oben beschränkt in M (bzw. nach unten beschränkt in M), falls es ein $r \in M$ gibt, so dass für alle $a \in A$ die Aussage $a \leq r$ (bzw. $a \geq r$) gilt. Ein solches r heißt obere Schranke (bzw. untere Schranke) von A. Wir sagen A ist beschränkt in M, wenn A nach oben und unten beschränkt ist.

Besitzt die Menge

$$\overline{S}(A, M) := \{r \in M \mid r \text{ ist obere Schranke von } A\}$$

ein Minimum, so nennen wir $\min \overline{S}(A, M)$ das Supremum von A in M.

Besitzt die Menge

$$\underline{S}(A, M) := \{r \in M \mid r \text{ ist untere Schranke von } A\}$$

ein Maximum, so nennen wir $\max \underline{S}(A, M)$ das Infimum von A in M.

Kurzschreibweise: $\sup A$ und $\inf A$

BEISPIELE 6.2. (a) $M = \mathbb{Q}$, $A = \{2, 4, 8\}$. Dann ist r obere Schranke, gdw $r \geq 8$. Und r ist untere Schranke, gdw $r \leq 2$. Wir haben $\max A = \sup A = 8$ und $\min A = \inf A = 2$.

(b) Falls A ein Maximum besitzt, so ist dieses Maximum auch das Supremum. Denn sei $a = \max A$, dann folgt

$$\overline{S}(A, M) = \{r \in M \mid r \geq a\} =: M_{\geq a}.$$

Also $a = \min \overline{S}(A, M) = \sup A$.

(c) $M = \mathbb{Q}$, $A = \{x \in \mathbb{Q} \mid 0 < x < 1\}$ besitzt kein Minimum. Denn für alle $x \in A$ gilt: $\frac{x}{2} \in A$ und $\frac{x}{2} < x$. Ähnlich zeigt man, dass A kein Maximum besitzt.

$$\overline{S}(A, M) = \mathbb{Q}_{\geq 1} \qquad \underline{S}(A, M) = \mathbb{Q}_{\leq 0}$$

$\sup A = 1$ und $\inf A = 0$.

Bemerkungen 6.3.

(1) $A \subset \underline{S}(\overline{S}(A, M), M)$ und $A \subset \overline{S}(\underline{S}(A, M), M)$
(2) Gilt $x \in A \cap \overline{S}(A, M)$, dann gilt $\forall a \in A : a \leq x$, somit $x = \max A$ und wegen (1) auch $x = \min \overline{S}(A, M)$.
(3) Hat A ein Maximum x, dann ist $x \in \overline{S}(A, M)$.
(4) A hat ein Maximum $\iff A \cap \overline{S}(A, M) \neq \emptyset \iff \#(A \cap \overline{S}(A, M)) = 1 \iff \sup A$ existiert und $\sup A \in A$.

Definition 6.4. Sei $\leq$ eine partielle Ordnung auf M. Wir sagen $(M, \leq)$ erfüllt die Supremumseigenschaft (S) falls gilt

(S) **Supremumseigenschaft**
Sei $A \subset M$ nicht-leer und nach oben beschränkt in M. Dann besitzt A ein Supremum in M.

Beispiel 6.5. $(\mathbb{Q}, \leq)$ erfüllt die Supremumseigenschaft nicht: Die Menge

$$A := \{x \in \mathbb{Q} \,|\, x^2 \leq 2\}$$

ist nach oben beschränkt.

$$\overline{S}(A, \mathbb{Q}) := \{x \in \mathbb{Q} \,|\, x > 0 \text{ und } x^2 \geq 2\}.$$

Da es keine rationale Zahl x mit $x^2 = 2$ gibt, folgt

$$\overline{S}(A, \mathbb{Q}) := \{x \in \mathbb{Q} \,|\, x > 0 \text{ und } x^2 > 2\}.$$

Zu gegebenem $x \in \overline{S}(A, \mathbb{Q})$ definieren wir nun

$$y := x - \frac{x^2 - 2}{2x},$$

und man zeigt leicht $0 < y < x$. Weiter gilt

$$y^2 = x^2 - 2x\frac{x^2 - 2}{2x} + \left(\frac{x^2 - 2}{2x}\right)^2 > x^2 - (x^2 - 2) = 2,$$

also auch $y \in \overline{S}(A, \mathbb{Q})$. Da es zu jedem Element in $\overline{S}(A, \mathbb{Q})$ ein kleineres in $\overline{S}(A, \mathbb{Q})$ gibt, kann es in $\overline{S}(A, \mathbb{Q})$ kein Minimum geben. Also besitzt A kein Supremum in $\mathbb{Q}$. Wir werden später sehen, dass es in $\mathbb{R}$ eine positive Lösung von $x^2 = 2$ gibt, aufgrund der Supremumseigenschaft.

Satz 6.6. Ist $(K, +, \cdot, \leq)$ ein geordneter Körper, der die Supremumseigenschaft erfüllt, dann ist er archimedisch.

Die Umkehrung gilt nicht: Beispiel $K = \mathbb{Q}$.

Beweis. Angenommen $(K, +, \cdot, \leq)$ sei nicht archimedisch. Dann gibt es ein $a \in K$ mit

$$\forall n \in \mathbb{N} : a > n.$$

Das heißt, $\mathbb{N}$ ist in K nach oben beschränkt. Wenn die Supremumseigenschaft erfüllt ist, gibt es also ein Supremum s von $\mathbb{N}$ in K. Für ein beliebiges n gilt also $s \geq n + 1$ also $s - 1 \geq n$. Also ist auch $s - 1$ ein obere Schranke von $\mathbb{N}$. Dann ist aber s nicht das Minimum aller oberen Schranken, was ein Widerspruch zur Wahl von s ist. Da wir einen

Widerspruch erhalten haben, muss die obige Annahme falsch gewesen sein. Wir haben dadurch gezeigt, dass $(K, +, \cdot, \leq)$ archimedisch ist. □

6.3. Axiome der reellen Zahlen. Im nächsten Abschnitt werden wir zwei Sätze zeigen:

Satz 6.7 („Existenz der reellen Zahlen“)**.** Es gibt einen geordneten Körper, der die Supremumseigenschaft erfüllt.

Satz 6.8 („Eindeutigkeit der reellen Zahlen“)**.** Seien $(K, +, \cdot, \leq)$ und $(\widetilde{K}, \tilde{+}, \tilde{\cdot}, \widetilde{\leq})$ zwei geordnete Körper, die die Supremumseigenschaft erfüllen, dann gibt es eine eindeutige bijektive Abbildung $F : K \longrightarrow \widetilde{K}$, die Addition, Multiplikation und Ordnung erhält.
F erhält die Addition: $\forall x, y \in K : \ F(x + y) = F(x)\tilde{+}F(y)$
F erhält die Multiplikation: $\forall x, y \in K : \ F(x \cdot y) = F(x)\tilde{\cdot}F(y)$
F erhält die Ordnung: $\forall x, y \in K : \ x \leq y \Longrightarrow F(x)\widetilde{\leq}F(y)$

Man sagt dann oft: $(K, +, \cdot, \leq)$ und $(\widetilde{K}, \tilde{+}, \tilde{\cdot}, \widetilde{\leq})$ sind kanonisch isomorph und F nennt man einen Isomorphismus (von geordneten Körpern). [13].

[13]Isomorph hat hier eine andere Bedeutung als bei den natürlichen Zahlen. „Isomorph“ kommt aus dem Griechischen und heißt „von gleicher Gestalt“. Es gibt in der Mathematik viele Strukturen und zu jeder eine eigene Definition von „isomorph“. Bei der Abbildung F oben im Text handelt es sich um einen Isomorphismus von geordneten Körpern: dies ist eine bijektive Abbildung, die Addition, Multiplikation und die Ordnung erhält. Sobald es einen Isomorphismus von A nach B gibt, nennt man A und B isomorph. Das Wort kanonisch ist nicht ganz so leicht zu erklären. Es bedeutet hier, dass es einen Isomorphismus gibt, der sich aus der Struktur bereits ergibt und nicht von zusätzlichen Wahlen abhängt. In der obigen Proposition gilt $F(0) = 0$ und $F(1) = 1$ und hierdurch ist der Isomorphismus F bereits festgelegt. Er ist also bereits durch die Struktur festgelegt und deswegen kanonisch.

Die Eigenschaften „geordneter Körper“ und “Supremumseigenschaft“ sind Eigenschaften, die wir von den reellen Zahlen erwarten. Da sie nun also die reellen Zahlen bis auf Isomorphie festlegen ist es sinnvoll, diese beiden Eigenschaft als die „Axiome der reellen Zahlen “ zu betrachten. Dies verpflichtet uns dazu, alles was wir für die reellen Zahlen zeigen wollen, aus diesen Axiomen heraus zu begründen. Der große Vorteil dieses axiomatischen Zugangs ist es, dass alle danach hergeleiteten Aussagen davon unabhängig sind, wie wir ein Modell der reellen Zahlen im Beweis von Satz 6.7 konstruieren.

Axiome 6.9 (Axiome der reellen Zahlen)**.** Die reellen Zahlen bilden einen geordneten Körper $(\mathbb{R}, +, \cdot, \leq)$, der die Supremumseigenschaft erfüllt.

Um beispielhaft zu zeigen, wie wir die Supremumseigenschaft nutzen können, betrachten wir das folgende Lemma.

Lemma 6.10. Sei $a \in \mathbb{R}$, $a \geq 0$, $n \in \mathbb{N}_{>0}$, so gibt es genau ein $r \in \mathbb{R}$ mit $r^n = a$ und $r \geq 0$.

Beweis. Die Aussage ist klar für $n = 1$ und klar für $a = 0$. Sei nun $n \geq 2$ und $a > 0$.

Eindeutigkeit. Hierzu genügt es zu zeigen:

$$\forall x, y \in \mathbb{R} : 0 < x < y \Longrightarrow x^n < y^n$$

Dies folgt aus der folgenden Umformung

$$y^n - x^n = (y - x) \sum_{i=0}^{n-1} y^i x^{n-i-1} > 0,$$

da $\sum_{i=0}^{n-1} y^i x^{n-i-1}$ als Summe positiver Zahlen positiv ist.

Existenz. Wir betrachten die Menge $A := \{x \in \mathbb{R} \mid x^n \leq a\}$. Die Menge A ist nach oben beschränkt: $\max\{1, a\}$ ist eine obere Schranke von A. Man zeigt leicht, dass

$$\overline{S}(A, \mathbb{R}) = \{x \in \mathbb{R} \mid 0 < x \wedge x^n \geq a\}.$$

Da A nicht-leer ist, existiert $s := \sup A = \min \overline{S}(A, \mathbb{R})$. Offensichtlich gilt $s \geq 0$ und $s^n \geq a$.

Wir nehmen nun an $s^n > a$.

Behauptung: Es gibt ein $\delta \in \mathbb{R}$, $0 < \delta < 1$ mit

$$s^n(1 - n\delta) \geq a. \tag{6.11}$$

Mit der Behauptung argumentieren wir wie folgt:

Die Bernoulli-Ungleichung liefert $(s(1-\delta))^n \geq s^n(1-n\delta)$, siehe Übungsblatt 4, Aufgabe 1 c) mit $x := -\delta$

Es gilt also $s > s(1 - \delta) \in \overline{S}(A, \mathbb{R})$. Somit ist dann s kein Minimum von $\overline{S}(A, \mathbb{R})$. Wir haben dann einen Widerspruch zur Annahme $s^n > a$ gefunden, und haben deswegen $s^n = a$, was die Existenz liefert.

Begründung der Behauptung:

Man zeigt mit Standard-Umformungen, dass (6.11) äquivalent zu

$$\delta \leq \frac{1}{n}\left(1 - \frac{a}{s^n}\right)$$

ist. Die rechte Seite ist zwischen 0 und 1. Wenn wir als $\delta := \frac{1}{n}\left(1 - \frac{a}{s^n}\right)$ wählen, so gilt die Behauptung. □

BEMERKUNG. Für $b \in \mathbb{R}$, $b > 0$, dann können wir jetzt alle Potenzen der Form b^t, mit $t \in \mathbb{Q}$ bilden. Wir schreiben hierzu $t = p/q$ mit $p \in \mathbb{Z}$, $q \in \mathbb{N}$. Wenn wir das vorangehende Lemma für $a := b^p$ und $n := q$ anwenden, dann besagt dieses Lemma, dass es genau eine Zahl $r \in \mathbb{R}_{>0}$ gibt, so dass $r^q = b^p$. Wir würden nun gerne definieren $a^t := r$ für dieses r. Damit dies eine sinnvolle Definition ist, muss man überprüfen, dass r unahängig davon ist, wie wir t als Bruch schreiben. Dazu muss man die folgende Aussage überprüfen:[14] Es gelte $t = p/q = \tilde{p}/\tilde{q}$, $p, \tilde{p} \in \mathbb{Z}$, $q, \tilde{q} \in \mathbb{N}$. Wir wählen $r, \tilde{r} \in \mathbb{R}_{>0}$ mit $r^q = b^p$ und $\tilde{r}^{\tilde{q}} = b^{\tilde{p}}$. Dann gilt $r = \tilde{r}$.

Man beachte, dass wir den Ausdruck b^t für eine positive reelle Zahl b und eine reelle, nicht-rationale Zahl t noch nicht definieren können.

Man kann aus dem bisher Bekannten nun auch leicht zeigen, dass für $b, \tilde{b} \in \mathbb{R}_{>0}$, $t, \tilde{t} \in \mathbb{Q}$ gilt:

$$(b\tilde{b})^t = b^t\tilde{b}^t \qquad b^{t+\tilde{t}} = b^t b^{\tilde{t}}$$

Fr 16.11.

Zur Konstruktion eines Modells der reellen Zahlen, das heißt zum Beweis von Satz 6.7, gibt es verschiedene Methoden, zum Beispiel:

- Die Konstruktion als Dezimalzahlen. Hier muss man zum Beispiel darauf achten, dass $1,4\overline{9} = 1,5$. Geht genauso im Binärsystem oder bezüglich anderer Basen.
- Die Konstruktion durch Intervallschachtelungen. Recht beliebt in Schulen, da anschaulich und unabhängig davon, ob man in der Basis 10 (=Dezimaldarstellung), in der Basis 2 (=Binärdarstellung) oder noch einer anderen Basis arbeitet. Siehe zum Beispiel [**22**, Abschnitt 2.3].

[14] Diese Überprüfung sollten Sie inzwischen selbst tun können.

- Cauchy-Folgen. Dies führt unter anderem zum Begriff der Vollständigkeit, der nicht nur in diesem Kontext wichtig ist, sondern auch zum Beispiel bei Hilbert- und Banach-Räumen. Solche Räume sind z.B. wichtig, um partielle Differentialgleichungen zu lösen und um Quantenmechanik zu beschrieben.
- Dedekindsche Schnitte. Mathematisch elegant und intuitivere Definition als beim Zugang über Intervallschachtelung und Cauchy-Folgen. Wir werden diesem Zugang folgen.

6.4. Dedekindsche Schnitte. Ideen von Richard Dedekind, 1872, Link zur Orginalarbeit

DEFINITION 6.12. Sei K ein archimedisch geordneter Körper. Ein Dedekindscher Schnitt in K ist eine Teilmenge $\alpha \subset K$ mit den folgenden Eigenschaften

(1) $\alpha \neq \emptyset$, $\alpha \neq K$,
(2) für $a \in \alpha$ und $x \in K_{<a}$ gilt $x \in \alpha$,
(3) α besitzt kein Maximum.

Man nennt Dedekindsche Schnitte auch Unterklassen. Das Komplement von α schreiben wir als $\alpha' := K \setminus \alpha$ und wird die zugehörige Oberklasse genannt.

Wir definieren

$$\overline{K} := \{\alpha \in \mathcal{P}(K) \mid \alpha \text{ ist ein Dedekindscher Schnitt}\}.$$

BEISPIEL 6.13. Sei $q \in K$. Dann ist $K_{<q}$ eine Unterklasse.

(1) $q - 1 \in K_{<q}$, also $K_{<q} \neq \emptyset$. Und $q + 1 \notin K_{<q}$, also $K_{<q} \neq K$.

(2) Sei $a \in K_{<q}$ und $x \in K_{<a}$. Dann haben wir $x < a$ und $a < q$, also auch $x < q$ und somit $x \in K_{<q}$.
(3) Zu jedem $a \in K_{<q}$ ist $a' := (a+q)/2$ ein Element $a' \in K_{<q}$ mit $a' > a$.

Die Abbildung $i_K : K \longrightarrow \overline{K}$, $q \mapsto K_{<q}$ ist injektiv. Denn ist $p < q$ so ist $p \in K_{<q}$ und $p \notin K_{<p}$.

Anschauliche Motivation für diese Konstruktion. Wir nehmen mal an, dass wir bereits $\mathbb{Q} \subset \mathbb{R}$ mit den gewünschten Eigenschaften hätten, $K = \mathbb{Q}$. Definiere $\overline{\mathbb{Q}}$ wie oben. Zu $r \in \mathbb{R}$ definieren wir die Unterklasse in $\mathbb{Q}$

$$U(r) := \{q \in \mathbb{Q} \mid q < r\}.$$

Bild mit einem Zahlenstrahl mit Ober- und Unterklasse

Wir erhalten eine Abbildung $U : \mathbb{R} \longrightarrow \overline{\mathbb{Q}}$, $r \mapsto U(r)$. Man überlegt sich leicht, dass $\sup|_{\overline{\mathbb{Q}}} : \overline{\mathbb{Q}} \longrightarrow \mathbb{R}$, $\alpha \mapsto \sup \alpha$ die Umkehrabbildung von $U : \mathbb{R} \longrightarrow \overline{\mathbb{Q}}$ ist. Somit ist U bijektiv. Wir definieren nun Addition $+$, Multiplikation $\cdot$ und Ordnung $\leq$ auf $\overline{\mathbb{Q}}$ so, dass U diese Strukturen erhält; und zwar so, dass wir die Existenz von $\mathbb{R}$ gar nicht nutzen. Zu zeigen ist dann, dass $(\overline{\mathbb{Q}}, +, \cdot, \leq)$ ein geordneter Körper ist, der die Supremumseigenschaft erfüllt. Wir identifizieren wiederum $\mathbb{Q}$ mit $B(i_{\mathbb{Q}}) \subset \overline{\mathbb{Q}}$ vermöge $i_{\mathbb{Q}}$.

Es stellen sich dann unter anderem die Fragen:

- Was passiert, wenn man die Konstruktion wiederholt zu $\overline{\overline{\mathbb{Q}}}$, also $K := \overline{\mathbb{Q}}$?
- Was wissen wir über $\overline{K}$, für andere archimedisch geordnete Körper K, zum Beispiel $K := \mathbb{Q}[\sqrt{2}]$?

Es ist deswegen sinnvoll, im folgenden immer für $(K, +, \cdot, \leq)$ einen beliebigen archimedisch geordneten Körper zuzulassen, aber sich immer den Spezialfall $K = \mathbb{Q}$ vorzustellen.

Sei also von nun an $(K, +, \cdot, \leq)$ ein beliebiger archimedisch angeordneter Körper.

Lemma 6.14. Für jede Oberklasse α' gilt:

(a) $\alpha' = \{x \in K \mid \forall a \in \alpha : a < x\}$
(b) für $a' \in \alpha'$ und $x' \in K_{>a'}$ gilt $x' \in \alpha'$.

Beweis. Zu (a): Sei $x \in \alpha'$. Wenn für (mindestens) ein $a \in \alpha$ die Ungleichung $a < x$ nicht gilt, dann haben wir den Widerspruch $x \in \alpha$. Somit $\alpha' \subset \{x \in K \mid \forall a \in \alpha : a < x\}$.

Sei $x \in K$ mit $\forall a \in \alpha : a < x$. Da $x < x$ falsch ist, gilt $x \notin \alpha$. Also $\alpha' \supset \{x \in K \mid \forall a \in \alpha : a < x\}$.

Zu (b): Dies folgt nun aus der Transitivität von $<$. □

Lemma 6.15. Erfüllt $(K, +, \cdot, \leq)$ die Supremumseigenschaft, dann ist $i_K : K \longrightarrow \overline{K}$, $q \mapsto K_{<q}$ bijektiv.

Beweis. Ist α eine Unterklasse, dann ist jedes $a' \in \alpha'$ eine obere Schranke von α. Ist $b \in \overline{S}(\alpha, K)$, dann ist b nicht in α, denn sonst wäre ja b ein Maximum von α, welches nicht existiert. Somit $b \in \alpha'$. Deswegen haben wir

$$\alpha' = \overline{S}(\alpha, K).$$

Erfüllt $(K, +, \cdot, \leq)$ die Supremumseigenschaft, dann hat also jede Oberklasse ein Minimum. Wenn wir $q := \sup \alpha = \min \alpha'$ setzen, so ist dann $\alpha = K_{<q}$. □

BEISPIEL 6.16. In einem geordneten Körper K ist

$$\alpha := K_{<0} \cup \{x \in K \mid x^2 < 2\}.$$

eine Unterklasse. Wenn diese Menge ein Supremum s besitzt, so gilt $s^2 = 2$.

Im Fall $K = \mathbb{Q}$ gibt es kein solches s, das heißt α hat kein Supremum. Dann ist $i_K : K \to \overline{K}$ nicht surjektiv.

Im Fall $K = \mathbb{R}$ hat α ein Supremum in $\mathbb{R}$, das wir $\sqrt{2}$ nennen.

DEFINITION 6.17 (Addition). Zu Dedekindschen Schnitten α und β definieren wir ihre Summe als

$$\alpha + \beta := \{a + b \mid a \in \alpha, b \in \beta\}.$$

Dies ist wiederum ein Dedekindscher Schnitt, denn:

(1) Sei $a \in \alpha$ und $b \in \beta$, dann gilt $a + b \in \alpha + \beta$, also $\alpha + \beta \neq \emptyset$.
Sei $a' \in \alpha'$ und $b' \in \beta'$, dann gilt für alle $a \in \alpha$ und $b \in \beta$ zunächst $a < a'$ und $b < b'$ und deswegen auch $a + b < a' + b'$. Somit $a' + b' \notin \alpha + \beta$, und dies ergibt $\alpha + \beta \neq K$.
(2) Sei $x < a + b$, $a \in \alpha$, $b \in \beta$. Setze $a_0 := x - b < a$, somit $a_0 \in \alpha$. Deswegen gilt $x = a_0 + b \in \alpha + \beta$.
(3) Angenommen, es gäbe ein Maximum in $\alpha + \beta$, dann könnten wir es als $a + b$ mit $a \in \alpha$ und $b \in \beta$ schreiben. Da a kein Maximum von α ist, gibt es ein $a_1 \in \alpha$ mit $a_1 > a$. Wir erhalten den Widerspruch

$$a + b < a_1 + b \in \alpha + \beta.$$

Wir erhalten also eine Abbildung $+ : \overline{K} \times \overline{K} \to \overline{K}$.

PROPOSITION 6.18. $(K, +)$ ist eine kommutative Gruppe, das heißt es gelten (Aa), (Ak), (An), (Ai).

Beweis. Die Kommutativität (Ak) ist offensichtlich.

Zu (Aa):

$$\alpha + (\beta + \gamma) = \alpha + \{b + c \mid b \in \beta, c \in \gamma\} = \{a + b + c \mid a \in \alpha, b \in \beta, c \in \gamma\}$$

und man erhält ähnlich denselben Ausdruck für $(\alpha + \beta) + \gamma$.

Zu (An): Das neutrale Element ist $K_{<0}$. Zu zeigen: $\alpha = \alpha + K_{<0}$. Aus Definition 6.12 (2) folgt direkt $\alpha + K_{<0} \subset \alpha$. Um $\alpha \subset \alpha + K_{<0}$ zu zeigen, wählen wir ein $a \in \alpha$. Da a kein Maximum von α ist, gibt es ein $a_1 \in \alpha$ mit $a_1 > a$. Dann ist $b := a - a_1 < 0$ ein Element von $K_{<0}$. Somit $a = a_1 + b \in \alpha + K_{<0}$.

Zu (Ai): Sei $\alpha \in \overline{K}$. Die zugehörige Oberklasse α' kann ein Minimum besitzen oder nicht. Falls α' ein Minimum besitzt, dann definiere

$$\beta := \{-a' \mid a' \in \alpha' \setminus \{\min \alpha'\}\},$$

falls α' kein Minimum besitzt, so definiere

$$\beta := \{-a' \mid a' \in \alpha'\}.$$

Da wir in dieser Definition das Minimum von α' weggenommen haben, falls es existiert, hat β kein Maximum, also (3). Aus Lemma 6.14 ergibt sich (2). Außerdem (1) $\beta \neq \emptyset$ und $\beta \neq K$. Wir haben also $\beta \in \overline{K}$.

Nun zeigen wir $\alpha + \beta = K_{<0}$, das heißt β ist das additive Inverse zu α.

Sei $a \in \alpha$ und $b \in \beta$, dann folgt $-b \in \alpha'$ und deswegen $-b > a$. Dies ergibt $a + b < 0$, in anderen Worten $a + b \in K_{<0}$. Also $\alpha + \beta \subset K_{<0}$.

Sei nun $x \in K_{<0}$. Wir suchen $a \in \alpha$ und $b \in \beta$ mit $x = a + b$.

Ansatz: Zu einem festen $a_0 \in \alpha$ wähle

$$a := \underbrace{a_0 + (n-1)|x|}_{\in \alpha} \text{ und } b := -(\underbrace{a_0 + n|x|}_{\substack{\in \alpha' \\ \text{kein Min}}})$$

für $n \in \mathbb{N}_{>0}$. Damit dieser Ansatz erfolgreich ist, muss n das Minimum der Menge

$$M := \{m \in \mathbb{N}_{>0} \mid a_0 + m|x| \in \alpha'\}$$

sein.

Durchführung: Definiere also M wie oben. Als ersten Schritt zeigen wir $M \neq \emptyset$.

Wähle ein $a' \in \alpha'$.

$$a_0 + m|x| \geq a' \;\leftrightarrow\; m \geq \frac{a' - a_0}{|x|}$$

Da K archimedisch ist, gibt es ein $m \in \mathbb{N}_{>0}$, das die rechte Ungleichung und damit auch die linke Ungleichung erfüllt. Es folgt $m \in M$, also $M \neq \emptyset$. Auf Grund der Wohlordnung der natürlichen Zahlen (Satz 1.9) besitzt M ein Minimum. Sei $n := \min M$. Definieren wir a und b wie oben, folgt $a \in \alpha$, $-b \in \alpha'$, $a + b = x$. Wenn $-b$ nicht ein/das Minimum von α' ist, so gilt $b \in \beta$, und wir haben das zu Zeigende gezeigt.

Wenn α' ein Minimum besitzt und wenn $-b = \min \alpha'$, dann setze

$$\tilde{a} := a_0 + \left(n - \frac{1}{2}\right)|x| = -b - \frac{1}{2}|x| < -b = \min \alpha' \,,$$

und damit $\tilde{a} \in \alpha$, und

$$\tilde{b} := -\Bigg(\underbrace{a_0 + \Big(n + \frac{1}{2}\Big)|x|}_{\substack{\in \alpha' \\ \text{kein Min}}}\Bigg) = b - \frac{1}{2}|x| \in \beta.$$

Mi 21.11. Dann gilt $\tilde{a} + \tilde{b} = -|x| = x$. □

DEFINITION 6.19 (Ordnung). Wir definieren auf $\overline{K}$ eine Relation

$$\alpha \leq \beta \quad :\Longleftrightarrow \quad \alpha \subset \beta.$$

Da $\subset$ eine Ordnungsrelation auf $\mathcal{P}(M)$ ist, ist die so definierte Relation $\leq$ eine Ordnungsrelation auf $\overline{K}$.

LEMMA 6.20. Die Ordnung $\leq$ auf $\overline{K}$ ist total.

Beweis. Zu zeigen ist: für alle $\alpha, \beta \in \overline{K}$ gilt $\alpha \subset \beta$ oder $\beta \subset \alpha$.

Wenn α keine Teilmenge von β ist, dann können wir ein $a \in \alpha \setminus \beta$ wählen. Für jedes $b \in \beta$ gilt dann $b < a$, denn sonst wäre $a \in \beta$. Somit ist $b \in \alpha$. Es folgt $\beta \subset \alpha$. □

Schreibe nun $0_{\overline{K}} := 0_{\overline{K}}$. Wir schreiben $|\alpha| = \alpha$ für $\alpha \geq 0_{\overline{K}}$ und $|\alpha| = -\alpha$ für $\alpha < 0_{\overline{K}}$.

Zeichnung und Beispiel zur Motivation der Multiplikation

DEFINITION 6.21 (Multiplikation). Seien $\alpha, \beta \in \overline{K}$.

(a) Falls $\alpha \geq 0_{\overline{K}}$ und $\beta \geq 0_{\overline{K}}$ so definieren wir das Produkt von α und β als

$$\begin{aligned} \alpha \cdot \beta &:= \{c \in K \mid c < 0 \ \vee \ \exists a \in \alpha : \exists b \in \beta : (a \geq 0 \wedge b \geq 0 \wedge c = a \cdot b)\} \\ &= \{a \cdot b \mid a \in \alpha_{\geq 0} \wedge b \in \beta_{\geq 0}\} \cup K_{<0}\,. \end{aligned}$$

(b) In den anderen Fällen definieren wir

$$\alpha \cdot \beta := \pm|\alpha| \cdot |\beta|,$$

wobei das Vorzeichnen genau dann + ist, wenn $\alpha < 0_{\overline{K}}$ und $\beta < 0_{\overline{K}}$.

LEMMA 6.22. $\alpha \cdot \beta \in \overline{K}$.

Beweis. Die folgenden Argumente werden aus Zeitgründen in der Vorlesung nicht ausgeführt.

(1) $-1 \in \alpha \cdot \beta$, also $\alpha \cdot \beta \neq \emptyset$.
Wähle $a' \in \alpha'$ und $b' \in \beta'$. Für alle $a \in \alpha$ und $b \in \beta$ gilt $a < a'$ und $b < b'$, somit $ab < a'b'$. Somit ist $a'b' \notin \alpha \cdot \beta$. Also $\alpha \cdot \beta \neq K$.

(2) Sei $c = a \cdot b$ mit $a \in \alpha$ und $b \in \beta$. Wir müssen zeigen, dass jedes $d \in K_{<c}$ in $\alpha \cdot \beta$ liegt. dies ist klar, falls $d < 0$. Im Fall $0 \leq d < c$ gilt

$$d = \underbrace{a}_{\in\alpha} \cdot \underbrace{\frac{d}{a}}_{\in\beta},$$

da $\frac{d}{a} < \frac{c}{a} \in \beta$. Somit $d \in \alpha \cdot \beta$.

(3) Sei a, b, c wie oben. Da α kein Maximum besitzt, gibt es ein $a_0 \in \alpha$, $a_0 > a$. Dann ist $c_0 := a_0 \cdot b \in \alpha \cdot \beta$, $c_0 > c$. Somit hat auch $\alpha \cdot \beta$ kein Maximum.

PROPOSITION 6.23. $(\overline{K}, +, \cdot)$ ist ein Körper.

Beweisskizze. Leicht zu prüfen sind (Ma), (Mk), (AMd) und die Tatsache, dass es mindestens zwei Elemente gibt ($i_{\mathbb{Q}} : \mathbb{Q} \longrightarrow K$ ist injektiv).

Zu (Mn): Es ist einfach zu zeigen, dass $1_{\overline{K}} := K_{<1}$ ein neutrales Element für die Multiplikation ist. (Beweis ähnlich wie für das additive Inverse.)

Wir wollen zeigen, dass $1_{\overline{K}} := K_{<1}$ ein neutrales Element für die Multiplikation ist, das heißt für alle Unterklassen α gilt $\alpha \cdot K_{<1} = \alpha$.

Fall 1: $\alpha > 0$.

Es gilt dann:

$$\alpha \cdot K_{<1} \overset{\text{def}}{=} \Big\{ a \cdot b \mid a \in \alpha_{\geq 0} \wedge b \in \{x \in K \mid 0 \leq x < 1\} \Big\} \cup K_{<0} \subset \{c \in K \mid \exists a \in \alpha : \, 0 \leq c < a\} \cup K_{<0} \subset \alpha.$$

Um $\alpha \subset K_{<1} \cdot \alpha$ zu zeigen, nehmen wir ein $a \in \alpha_{\geq 0}$. Da a kein Maximum von α ist, gibt es ein $a_1 \in \alpha$ mit $a_1 > a$. Setze $b := a/a_1$, dann $0 \leq b < 1$. Somit gilt $a = a_1 \cdot b \in \alpha \cdot K_{<1}$. Hieraus folgt $\alpha \subset K_{<1} \cdot \alpha$ aus der Definition der Multiplikation.

Fall 2: $\alpha = 0_{\overline{K}}$. Für alle $\beta \in \overline{K}$

$$0_{\overline{K}} \cdot \beta = (0_{\overline{K}} - 0_{\overline{K}}) \cdot \beta = 0_{\overline{K}} \cdot \beta - 0_{\overline{K}} \cdot \beta = 0_{\overline{K}}.$$

Dies gilt insbesondere für $\beta = 1_{\overline{K}}$.

Fall 3: $\alpha < 0$.

$$\alpha \cdot K_{<1} \overset{\text{def}}{=} -|\alpha| \cdot K_{<1} \overset{\text{Fall 1}}{=} -|\alpha| = \alpha.$$

Zu (Mi): Sei $\alpha \in \overline{K}$, $\alpha \neq 0_{\overline{K}}$.
Fall 1: $\alpha > 0_{\overline{K}}$.

Wenn $\min \alpha'$ nicht existiert, so definieren wir

$$\beta := K_{\leq 0} \cup \{b \in K_{>0} \mid b^{-1} \in \alpha'\},$$

und wenn es existiert, dann

$$\beta := K_{\leq 0} \cup \{b \in K_{>0} \mid b^{-1} \in \alpha' \wedge b \neq \min \alpha'\}.$$

Es ist nicht schwer zu zeigen, dass $\beta \in \overline{K}$. Mit ähnlichen Argumenten wie im Beweis von (Ai) zeigt man $\alpha \cdot \beta = 1_{\overline{K}}$ (etwas Arbeit!). Diese Argumente werden aus Zeitgründen in der Vorlesung nicht ausgeführt.

Zu $\beta \in \overline{K}$:

(1) $0 \in \beta$, also $\beta \neq \emptyset$.
Wähle $a \in \alpha$, $a > 0$. Dann gilt $a^{-1} \notin \beta$, denn für alle $a' \in \alpha'$ gilt $(a')^{-1} < a^{-1}$.

(2) Ist $b \in \beta$ und $c \in K_{<b}$. Im Fall $c \leq 0$ gilt $c \in \beta$. Im Fall $0 < c < b$ gilt $c^{-1} > b^{-1} \in \alpha'$, somit $c \in \beta$.

(3) Angenommen b wäre ein Maximum von β. Im Fall $b = 0$ bekommt man leicht einen Widerspruch, da β positive Elemente hat. Im Fall $b > 0$ ist dann b^{-1} ein Minumum von α'. dies ist aber ausgeschlossen in der Definition von β.

Zu $\alpha \cdot \beta = 1_{\overline{K}}$:

Wir nehmen ein $c \in \alpha \cdot \beta$. O.B.d.A. $c \geq 0$. Wir können also schreiben $c = a \cdot (a')^{-1}$ mit $a \in \alpha$, $a' \in \alpha'$, $a \geq 0$. Aus $a' > a$ folgt $c < 1$. Somit $\alpha \cdot \beta \leq 1_{\overline{K}}$.

Sei nun $c \in 1_{\overline{K}}$, also $c < 1$. Wir wollen $c \in \alpha \cdot \beta$ zeigen. Dies ist trivial, falls $c \leq 0$. Deswegen sei nun $c > 0$. Wähle ein festes $a_0 \in \alpha$ mit $a_0 > 0$. Wir zeigen, die Menge

$$M := \{m \in \mathbb{N} \mid a_0 \cdot c^{-m} \in \alpha\}$$

ist endlich. Hierzu wählen wir ein $a' \in \alpha'$ und rechnen mit der Bernoulli-Ungleichung

$$a_0 \cdot c^{-m} = a_0 \cdot \left(1 + \frac{1-c}{c}\right)^m \geq a_0 \cdot \left(1 + m \cdot \frac{1-c}{c}\right) > a'$$

für alle ausreichend großen m, da K archimedisch ist. Somit gibt es ein $m_0 \in \mathbb{N}$, so dass für alle $m \geq m_0$ gilt: $a_0 \cdot c^{-m} \notin \alpha$. Also hat M höchstens m_0 Elemente und somit ein Maximum.

Sei nun $n := \max M$. Setze $a := a_0 \cdot c^{-n} \in \alpha$ und $a' := a_0 \cdot c^{-n-1} \in \alpha'$. Den Spezialfall, dass $(a')^{-1}$ das Minimum von α' ist, betrachten wir in einer Fussnote.[15] Wir setzen $b := (a')^{-1} \in \beta$, und erhalten $c = a \cdot b$.

Fall 2: $\alpha < 0_{\overline{K}}$

Bestimme ein $\beta \in \overline{K}$ mit $|\alpha| \cdot \beta = 1$. Ein multiplikatives Inverse von α ist dann $-\beta$. □

Satz 6.24. $(\overline{K}, +, \cdot, \leq)$ ist ein archimedischer geordneter Körper. Die Abbildung

$$\begin{aligned} i_K : K &\to \overline{K} \\ a &\mapsto K_{<a} \end{aligned}$$

ist eine injektive Abbildung, die Addition, Multiplikation und Ordnung erhält.

Beweis.

1.) Um zu zeigen, dass $(\overline{K}, +, \cdot, \leq)$ ein geordneter Körper ist, gehen wir (a)–(d) in der Definition 5.1 von geordneten Körpern durch.

Zu (a): $(\overline{K}, +, \cdot)$ ist ein Körper (Proposition 6.23).

[15] Man kann dann a_0 „etwas“ zu a_1 vergrößern, so dass $a := a_1 \cdot c^{-n} \in \alpha$ immer noch in α liegt, und dann ist $a' := (a_1 \cdot c^{-n-1})^{-1}$ ist dann ein nicht-minimales Element von α'.

Zu (b): $\leq$ ist totale Ordnung auf $\overline{K}$ (Lemma 6.20).

Zu (c): Gilt $\alpha \subset \beta$, dann haben wir

$$\alpha + \gamma = \{a + c \mid a \in \alpha, c \in \gamma\} \subset \{b + c \mid b \in \beta, c \in \gamma\} = \beta + \gamma.$$

Zu (d): Wir wollen zeigen

$$(\alpha \leq \beta \wedge 0_{\overline{K}} \leq \gamma) \to \alpha \cdot \gamma \leq \beta \cdot \gamma\,.$$

Fall 1: $\gamma \geq 0_{\overline{K}}$ und $0_{\overline{K}} \leq \alpha \leq \beta$

$$\begin{aligned} \alpha \cdot \gamma &= \{a \cdot c \mid a \in \alpha_{\geq 0} \wedge c \in \gamma_{\geq 0}\} \cup K_{<0} \\ &\subset \{b \cdot c \mid b \in \beta_{\geq 0} \wedge c \in \gamma_{\geq 0}\} \cup K_{<0} \\ &= \beta \cdot \gamma \end{aligned}$$

Fall 2: $\gamma \geq 0_{\overline{K}}$ und $\alpha < 0_{\overline{K}} \leq \beta$

Dann ist $\alpha \cdot \gamma \leq 0_{\overline{K}} \leq \beta \cdot \gamma$.

Fall 3: $\gamma \geq 0_{\overline{K}}$ und $\alpha \leq \beta < 0_{\overline{K}}$

Dann gilt $-\beta \leq -\alpha \Longrightarrow -\beta \cdot \gamma \leq -\alpha \cdot \gamma \Longrightarrow \alpha \cdot \gamma \leq \beta \cdot \gamma$.

Somit ist $(\overline{K}, +, \cdot)$ ein geordneter Körper.

2.) Wir wissen bereits, dass $i_K : K \longrightarrow \overline{K}$ injektiv ist, und es ist leicht zu prüfen, dass diese Abbildung Addition, Multiplikation und Ordnung erhält.

3.) Nun zeigen wir, dass $\overline{K}$ archimedisch ist, das heißt: Für jedes $\alpha \in \overline{K}$ ist ein $n \cdot 1_{\overline{K}} \in \mathbb{N}$ zu finden mit $\alpha \leq n$.

Wähle ein $a' \in \alpha'$. Da K archimedisch ist, gibt es ein $m \in \mathbb{N}$ mit $m \geq a'$. Wähle nun $\frac{1}{2} \in 1_{\overline{K}}$. Dann gilt für alle $a \in \alpha$

$$2m \cdot \frac{1}{2} \geq a' \geq a.$$

Es folgt

$$2m \cdot 1_{\overline{K}} \geq \alpha.$$

□

LEMMA 6.25. $x \in \alpha \iff i_K(x) < \alpha$

Beweis.

$$\begin{aligned} i_K(x) < \alpha &\iff i_K(x) \subsetneq \alpha \\ &\iff \text{Es existiert ein } y \text{ in } \alpha \setminus i_K(x) \\ &\iff \exists y \in \alpha : \, y \geq x \\ &\iff x \in \alpha \, . \end{aligned}$$

In der letzten Äquivalenz ist „⇒“ eine Konsequenz von Definition 6.12 (2) und „⇐“ ergibt sich mit $y := x$. □

PROPOSITION 6.26. Sei $(K_0, +, \cdot, \leq)$ ein archimedisch angeordneter Körper. Wir identifizieren $\mathbb{Q}$ mit seinem Bild in K_0 vermöge der in Lemma 5.7 definierten Abbildung $i_{\mathbb{Q}} : \mathbb{Q} \longrightarrow K_0$. Seien $a, b \in K_0$, $a < b$. Dann gibt es ein $q \in \mathbb{Q}$ mit $a < q < b$.

Interpretation für $K_0 := \mathbb{R}$: zwischen zwei verschiedenen reellen Zahlen liegt eine rationale Zahl.

Beweis. Da K_0 archimedisch ist, existieren $n, m \in \mathbb{N}$ mit $n > \frac{1}{b-a}$ und $m > -b$. Setze $\tilde{b} := b + m + 1 > 1$ und $\tilde{a} := a + m + 1$.

$$p := \max \underbrace{\{r \in \mathbb{N}_{>0} \mid r < n\tilde{b}\}}_{\text{endlich, } \neq \emptyset} .$$

Wegen $b-a>0$ erhalten wir $n(\tilde{b}-\tilde{a})=n(b-a)>1$, also $n\tilde{a}<n\tilde{b}-1$. Also $n\tilde{a}<p<n\tilde{b}$. Somit

$$a<\underbrace{\frac{p}{n}-m-1}_{\in\mathbb{Q}}<b.$$

□

Satz 6.27 (Satz von Dedekind)**.** Sei $\overline{K}=A\cup A'$, $A\neq\emptyset$, $A'\neq\emptyset$. Es gelte

$$\forall a\in A:\ \forall a'\in A':\ a<a'.$$

Dann existiert ein eindeutiges $\beta\in\overline{K}$ mit

$$\forall a\in A:\ \forall a'\in A':\ a\leq\beta\leq a'\,.$$

Man nennt β die Trennungszahl.[16]

Vorstellung: Im obigen Satz ist A so etwas ähnliches wie eine Unterklasse. Wenn man in Definition 6.12 K durch $\overline{K}$ ersetzen würde, dann wären (1) und (2) erfüllt, aber es verbleibt unbestimmt, ob A ein Maximum hat. Man stellt sich somit am besten auch dies als die Aufteilung von $\overline{K}$ in einem linken Teil A und einen rechten Teil A' vor.

Zahlenstrahl, der $\overline{K}$ symbolisiert. In der Mitte β, links davon A, rechts davon A'.

[16]Im Satz von Dedekind kann man sich nun über die gewählte Bezeichnung wundern: Wieso schreibe ich hier $a\in A$? Denn dieses a ist ja ein Element von $\overline{K}$ und damit eine Unterklasse, die bisher typischerweise mit dem Buchstaben α bezeichnet wurden. Wenn wir aber letzterer Bezeichnungslogik folgen, dann müssen wir die Elemente von A' mit α' bezeichnen. In der bisherigen Notation ist α' immer die Oberklasse zur Unterklasse α definiert. Um diese Verwechslungsgefahr zu vermeiden, schreibe ich a und a' in der Aussage des Satzes.

Beweis. **Existenz:**

Fall 1: A besitzt ein Maximum.

Dann ist $\max A$ eine Trennungszahl.

Fall 2: A besitzt kein Maximum.

Setze

$$\beta := \{b \in K \mid i_K(b) \in A\}\,.$$

β ist Dedekindscher Schnitt in K.

Denn

(1) Wähle[17] ein $\alpha \in A$ und dann ein $b \in \alpha$. Mit Lemma 6.25 folgt $i_K(b) < \alpha$. $\Longrightarrow$ $i_K(b) \in A \Longrightarrow b \in \beta \Longrightarrow \beta \neq \emptyset$.
Wähle ein $\tilde{\alpha} \in A'$ und dann ein $b' \notin \tilde{\alpha}$. Mit Lemma 6.25 folgt $i_K(b') \geq \tilde{\alpha}$. $\Longrightarrow$ $i_K(b') \in A' \Longrightarrow b' \notin \beta \Longrightarrow \beta \neq K$.

(2) Sei $b \in \beta$ und $x \in K_{<b} \Longrightarrow i_K(x) = K_{<x} < K_{<b} = i_K(b) \in A \Longrightarrow i_K(x) \in A \Longrightarrow x \in \beta$.

(3) Sei $b \in \beta$. Somit $i_K(b) \in A$. Da A kein Maximum hat, existiert ein $\alpha \in A$ mit $i_K(b) < \alpha$. Nach Proposition 6.26 mit $K_0 := \overline{K}$ gibt es ein $a \in \mathbb{Q} \subset K$ mit $i_K(b) < i_K(a) < \alpha$ und somit $b < a$. Aus $i_K(a) < \alpha \in A$ folgt $i_K(a) \in A$, also $a \in \beta$. Somit war das beliebig vorgegebene $b \in \beta$ kein Maximum.

Somit $\beta \in \overline{K}$. Zu zeigen bleibt, dass β eine Trennungszahl ist.

[17] Im Beweis benötige ich nun Elemente von Elementen von A, deswegen wechsle ich zur alten Notation $a \in \alpha \in A$. Wir schreiben $\tilde{\alpha}$ für Elemente von A', um Verwechslungen zur Oberklasse α' von α zu vermeiden.

Zu jedem $\alpha \in A$ gibt es ein $\alpha_1 \in A$ mit $\alpha < \alpha_1$. Proposition 6.26 liefert ein $b \in \mathbb{Q} \subset K$ mit $\alpha < i_K(b) < \alpha_1$. $\Longrightarrow i_K(b) \in A \Longrightarrow b \in \beta$. Mit Lemma 6.25 sehen wir $i_K(b) < \beta$ und somit $\alpha < \beta$.

Angenommen es gäbe ein $\tilde{\alpha} \in A'$ mit $\tilde{\alpha} < \beta$. Proposition 6.26 liefert ein $p \in \mathbb{Q} \subset K$ mit $\tilde{\alpha} < i_K(p) < \beta$. Wegen Lemma 6.25 haben wir $p \in \beta$, also $i_K(p) \in A$. Gleichzeitig erhalten wir aus $\tilde{\alpha} < i_K(p)$ auch $i_K(p) \in A'$, also einen Widerspruch.

Eindeutigkeit:
Angenommen es gäbe zwei Trennungszahlen β_1, β_2 mit $\beta_1 < \beta_2$. Setze $\beta_m := (\beta_1 + \beta_2)/2 \in \overline{K}$. Wir haben $\beta_1 < \beta_m$ und somit $\beta_m \notin A$. Wegen $\beta_m < \beta_2$ folgt $\beta_m \notin A'$. Dies widerspricht $A \cup A' = \overline{K}$. □ **Fr 23.11.**

SATZ 6.28. Ist K ein archimedisch geordneter Körper, dann ist $\overline{K}$ ein archimedisch geordneter Körper, der die Supremumseigenschaft erfüllt. Die Abbildung $i_{\overline{K}} : \overline{K} \longrightarrow \overline{\overline{K}}$ ist ein Isomorphismus.

Daraus folgt dann direkt Satz 6.7.

Beweis. Sei B eine nicht-leere nach oben beschränkte Teilmene von $\overline{K}$. Wenn B ein Maximum besitzt, so ist dieses auch das Supremum und wir sind fertig. Andernfalls definieren wir $A' := \overline{S}(B, \overline{K})$ und $A := \overline{K} \setminus A'$. Die Voraussetzungen in Satz 6.27 sind erfüllt. Deswegen gibt es eine Trennungszahl β zur Zerlegung $\overline{K} = A \cup A'$.

Dann haben wir $\beta \in \overline{S}(A, \overline{K})$ und damit[18] auch $\beta \in \overline{S}(B, \overline{K}) = A'$. Wegen der Folgerung in Satz 6.27 ist dann β das Minimum von $A' = \overline{S}(B, \overline{K})$ und damit das Supremum von B. Insbesondere: das Supremum existiert.

Also erfüllt $\overline{K}$ die Supremumseigenschaft. Deswegen ist $i_{\overline{K}} : \overline{K} \longrightarrow \overline{\overline{K}}$ nach Lemma 6.15 ein Isomorphismus. □

Wir wenden uns nun dem Beweis von Satz 6.8 zu.

Sei $(K, +, \cdot, \leq)$ ein geordneter Körper, der die Supremumseigenschaft erfüllt. Nach Lemma 5.7 gibt es eine injektive Abbildung $i_{\mathbb{Q}}^K : \mathbb{Q} \longrightarrow K$, die Addition, Multiplikation und Ordnung erhält. Wir identifizieren im folgenden $\mathbb{Q}$ mit dem Bild von $i_{\mathbb{Q}}^K$ vermöge $i_{\mathbb{Q}}^K$. Also $\mathbb{Q} \subset K$ und $\mathbb{Q} \subset \overline{\mathbb{Q}}$.

Sei $\alpha \in \overline{\mathbb{Q}}$, also α sei eine Unterklasse in $\mathbb{Q}$. Dann liefert jedes $a' \in \alpha'$ eine obere Schranke von α in K. Somit ist α eine nicht-leere, nach oben beschränkte Teilmenge von K. Somit existiert $\sup^K \alpha \in K$, wobei $\sup^K$ das Supremum in K bezeichnet. Wir definieren eine Abbildung

$$s : \overline{\mathbb{Q}} \longrightarrow K, \quad \alpha \mapsto \sup^K \alpha\,.$$

Lemma 6.29. Zu jedem $x \in K$ ist $U(x) := \{a \in \mathbb{Q} \mid a < x\}$ eine Unterklasse in $\mathbb{Q}$.

[18] Es gilt hier sogar $\overline{S}(A, \overline{K}) = \overline{S}(B, \overline{K})$: wir haben diese Tatsache aber weder gezeigt noch benutzt. Wir nutzen hier lediglich $C \subset D \to \overline{S}(D, \overline{K}) \subset \overline{S}(C, \overline{K})$.

Beweis. Der geordnete Körper K ist insbesondere archimedisch (Satz 6.6).

$$\begin{aligned} K \text{ archimedisch} &\overset{\text{def}}{\iff} \forall a \in K : \exists n \in \mathbb{N} : a \leq n \\ &\iff \forall a \in K : \exists n \in \mathbb{N} : a < n \\ &\iff \neg\,(\exists a \in K : \forall n \in \mathbb{N} : n \leq a) \\ &\iff \mathbb{N} \text{ unbeschränkt} \end{aligned}$$

Somit ist $\mathbb{Z}$ und damit auch $\mathbb{Q}$ nach oben und unten unbeschränkt.

Nun zeigen wir, dass $U(x)$ eine Unterklasse in $\mathbb{Q}$ ist:

(1) Nutze die Unbeschränktheit von $\mathbb{Q}$:
   $U(x) \neq \emptyset$, da $\mathbb{Q}$ nach unten unbeschränkt in K.
   $U(x) \neq \mathbb{Q}$, da $\mathbb{Q}$ nach oben unbeschränkt in K.
(2) $a \in U(x)$ und $y < a$ dann haben wir $y < a < x$.
(3) Angenommen a wäre ein Maximum von $U(x)$. Also $a < x$. Nach Proposition 6.26 mit $K_0 := K$ existiert ein $q \in \mathbb{Q}$ mit $a < q < x$. Also $q \in U(x)$. Widerspruch!

Lemma 6.30. Die Abbildung

$$U : K \longrightarrow \overline{\mathbb{Q}}, \quad x \mapsto U(x)$$

ist die Umkehrabbildung von $s : \overline{\mathbb{Q}} \longrightarrow K$.

Beweis. $\forall a \in \alpha : a < s(\alpha) \Longrightarrow \alpha \subset U(s(\alpha))$
$\forall a' \in \alpha' : a' \in \overline{S}(\alpha, \mathbb{Q}) \subset \overline{S}(\alpha, \overline{\mathbb{Q}}) \Longrightarrow a' \geq s(\alpha) \Longrightarrow a' \notin U(s(\alpha))$.
Somit $\alpha = U(s(\alpha))$.

Sei $x \in K$, $y := s(U(x)) = \sup^K\{q \in \mathbb{Q} \mid q < x\}$. Es folgt $y \leq x$. Angenommen $y < x$, dann existiert nach Proposition 6.26 mit $K_0 := K$ ein $q \in \mathbb{Q}$ mit $y < q < x$, also $q \in U(x)$. Widerspruch! □

Man kann zeigen, dass s Addition, Multiplikation und Ordnung bewahrt.[19] Nach all diesen Vorarbeiten ist der Beweis von Satz 6.8 ganz kurz.

Beweis von Satz 6.8. Sei nun $(\widetilde{K}, \tilde{+}, \tilde{\cdot}, \widetilde{\leq})$ ein weiterer geordneter Körper ist, der die Supremumseigenschaft erfüllt. Wir erhalten analog $\widetilde{s} : \overline{\mathbb{Q}} \to \widetilde{K}$. Dann erfüllt $F := \widetilde{s} \circ s^{-1} : K \longrightarrow \widetilde{K}$ die in Satz 6.8 geforderten Eigenschaften. □

Folgerung 6.31. Ist $(K, +, \cdot, \leq)$ ein archimedisch geordneter Körper, dann gibt es eine injektive Abbildung $f : K \longrightarrow \mathbb{R}$, die Addition, Multiplikation und Ordnung erhält.

Interpretation: Jeder archimedisch geordnete Körper ist im wesentlich gleich (mathematisch präziser: als geordneter Körper isomorph) zu einem Unterkörper von $\mathbb{R}$. Ein Unterkörper ist eine Teilmenge von $\mathbb{R}$, die selber ein Körper ist; auf der Teilmenge sollen Addition, Multiplikation und Ordnung wie in $\mathbb{R}$ defniert sein.

Beweis. Schränke die bijektive Abbildung $\overline{K} \longrightarrow \mathbb{R}$ auf K ein. □

7. Die komplexen Zahlen

Motivation. $x^2 = -1$ besitzt keine reelle Lösung.

[19] Der Beweis ist wiederum nicht schwer, aber besteht aus einigen Teilbehauptungen, die sich direkt aus der Definition von $+$, $\cdot$ und $\leq$ ergeben.

Idee: Erweitere die reelle Zahlengerade zu einer Ebene, der Gaußschen Zahlenebene. Die Elemente dieser Ebene sind komplexe Zahlen.

Ein Modell für die komplexen Zahlen.

DEFINITION 7.1. Auf $\mathbb{C} := \mathbb{R}^2$ definieren wir die folgenden Verknüpfungen:

$$\begin{aligned} \text{Addition: } + : \mathbb{C} \times \mathbb{C} &\longrightarrow \mathbb{C} \\ ((x_1, y_1), (x_2, y_2)) &\mapsto (x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2) \\ \text{Multiplikation: } \cdot : \mathbb{C} \times \mathbb{C} &\longrightarrow \mathbb{C} \\ ((x_1, y_1), (x_2, y_2)) &\mapsto (x_1, y_1) \cdot (x_2, y_2) := (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1) \end{aligned}$$

Gaußsche Zahlenebene und Addition

Die Abbildung $i_{\mathbb{R}} : \mathbb{R} \hookrightarrow \mathbb{C}$, $a \mapsto (a, 0)$ ist injektiv und erhält Addition und Multiplikation. Wir identifizieren a mit $(a, 0)$. Wir schreiben $\mathbf{i} := (0, 1)$. Dann gilt für $x, y \in \mathbb{R}$:

$$x + y\mathbf{i} = (x, 0) + (y, 0) \cdot (0, 1) = (x, 0) + (0, y) = (x, y).$$

Man nennt dann x den Realteil von $z = x + y\mathbf{i}$ und y den Imaginärteil, geschrieben $x = \mathbf{Re}\, z$ und $y = \mathbf{Im}\, z$.

$$\mathbb{C} = \{x + y\mathbf{i} \mid x, y \in \mathbb{R}\}.$$

Wir haben $\mathbf{i}^2 = (0 - 1) + (0 + 0)\mathbf{i} = -1$.

Komplexe Zahlen $z = x + y\mathbf{i}$, $x, y \in \mathbb{R}$ heißen

- imaginär, falls $y = \mathbf{Im}\, z \neq 0$,
- rein imaginär, falls $x = \mathbf{Re}\, z = 0$ und $y = \mathbf{Im}\, z \neq 0$.

Lemma 7.2. $(\mathbb{C}, +, \cdot)$ ist ein Körper.

Beweis. Assoziativität und Kommutativität von Addition und Multiplikation ist einfach nachzurechnen: (Aa), (Ak), (Ma), (Mk). Ebenso das Distributivgesetz (AMd). Das neutrale Element der Addition (bzw. Multiplikation) ist $0 = 0 + 0\mathbf{i}$, (bzw. $1 = 1 + 0\mathbf{i}$), (An) bzw. (Mn). Das additive Inverse von $x + y\mathbf{i}$ ist $(-x) + (-y)\mathbf{i}$, (Ai).

(Mi): Gegeben sei $z := x + y\mathbf{i} \in \mathbb{C} \setminus \{0\}$. Es folgt dann $x \neq 0 \vee y \neq 0$ und somit $x^2 + y^2 \in \mathbb{R}_{>0}$.

Wir benötigen ein multiplikatives Inverses von z, also ein $w \in \mathbb{C}$ mit $wz = 1$. Wir zeigen: Es gibt genau ein solches $w \in \mathbb{C}$.[20]

Zur Eindeutigkeit:

Angenommen $wz = 1$ mit $z = x + y\mathbf{i}$ und $w = a + b\mathbf{i}$. Multipliziere beide Seiten mit $x - \mathbf{i}y$:

$$(x - y\mathbf{i}) = (a + b\mathbf{i})(x + y\mathbf{i})(x - y\mathbf{i}) = (a + b\mathbf{i})(x^2 + y^2) = a(x^2 + y^2) + b(x^2 + y^2)\mathbf{i}\,.$$

Hieraus folgt $x = a(x^2 + y^2)$ und $-y = b(x^2 + y^2)$, also

$$a = \frac{x}{x^2 + y^2}, \tag{7.3}$$

$$b = -\frac{y}{x^2 + y^2}. \tag{7.4}$$

Zur Existenz:

[20] man muss eigentlich hier gar nicht die Eindeutigkeit zeigen, die Existenz reicht aus. Wenn man sich aber die Eindeutigkeit anschaut, bekommt man eine Idee, wie man die Existenz zeigen kann.

Wir definieren a und b wie in (7.3) und (7.4). Die Definition ergibt Sinn, da $x^2 + y^2 > 0$. Man rechnet nach

$$(x + y\mathbf{i}) \cdot (a + b\mathbf{i}) = 1\,.$$

□

Für $x, y \in \mathbb{R}$ definieren wir

$$\overline{x + y\mathbf{i}} := x - y\mathbf{i}.$$

Wir erhalten eine $\mathbb{R}$-lineare bijektive Abbildung $\mathbb{C} \longrightarrow \mathbb{C}$, $z \mapsto \overline{z}$, die komplexe Konjugation genannt wird.

Die Abbildung

$$\mathbb{C} \longrightarrow \mathbb{R}, \qquad x + y\mathbf{i} \mapsto |x + y\mathbf{i}| := \sqrt{x^2 + y^2} = \sqrt{(\overline{x + y\mathbf{i}})(x + y\mathbf{i})}$$

heißt die Betragsfunktion der komplexen Zahlen.

Eigenschaften: Seien $z, z_1, z_2 \in \mathbb{C}$.

$$\overline{z_1 \pm z_2} = \overline{z_1} \pm \overline{z_2}$$

$$\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$$

$$\overline{z_1 / z_2} = \overline{z_1} / \overline{z_2}$$

$$|z_1 + z_2| \leq |z_1| + |z_2|$$

$$|z_1 z_2| = |z_1| \cdot |z_2|$$

Es gelten auch die anderen in Lemma 5.11 aufgezählten Eigenschaften mit Ausnahme von (d), den man wie folgt abändern muss

$$|a^2| = |a|^2 = a\overline{a} \geq 0.$$

Polarkoordinaten und Veranschaulichung der Multiplikation in $\mathbb{C}$.
[21]

Bilder zur Polar-Darstellung

Schreibe z_1, z_2, z_3 in Polar-Darstellung

$$z_n = r_n(\cos\varphi_n + (\sin\varphi_n)\mathbf{i}), \qquad r_n \in \mathbb{R}_{>0}, \quad \varphi_n \in \mathbb{R}.$$

$r_n := |z_n|$; $\varphi_n = \arg(z_n)$, falls $\varphi_n \in [0; 2\pi)$.

Übung 7.5 (Übungsblatt 6 Aufgabe 4)**.** $z_3 = z_1 \cdot z_2$ ist äquivalent zu

$$r_3 = r_1 \cdot r_2 \quad \text{und} \quad \frac{\varphi_3 - (\varphi_2 + \varphi_1)}{2\pi} \in \mathbb{Z}$$

Mi 28.11.

Definition 7.6. Sei $d \in \mathbb{N}$ und seien $a_0, a_1, \ldots, a_d \in \mathbb{C}$. Eine Funktion der Form

$$\begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \mapsto & P(z) := \sum_{k=0}^{d} a_k z^k = a_0 + a_1 z^1 + a_2 z^2 + \cdots + a_d z^d \end{array}$$

[21] Ein Kommentar, um die logische Stellung dieser Veranschaulichung zu klären: um eine geometrische Vorstellung zu entwickeln, wie man komplexe Zahlen multipliziert, wurde hier die Polar-Darstellung eingeführt, obwohl wir eigentlich zum jetzigen Zeitpunkt noch gar nicht wissen, was sin und cos bedeuten. Die Übungsaufgabe (Übungsblatt 6 Aufgabe 4), die mit dieser Veranschaulichung verbunden ist, nutzt die Additionstheoreme für Sinus und Cosinus, die wir ebenfalls noch nicht bewiesen haben, die aber wahrscheinlich jeder schon in der Schule gesehen hat. Für einen exakten und streng logischen Aufbau wäre diese Veranschaulichung und die oben genannte Aufgabe besser erst dann zu behandeln, wer wir all dies eingeführt haben. Da es aber mindestens genauso wichtig ist, dass Sie möglichst bald eine gute Vorstellung von den komplexen Zahlen entwickeln, ist es dennoch richtig hier und nicht später diese Veranschaulichung zu behandeln.

nennt man eine polynomiale Funktion[22] über $\mathbb{C}$. Die Zahlen $a_0, \ldots, a_d$ nennt man die Koeffizienten. Gilt $a_d \neq 0$, so sagen wir P hat Grad d, $\deg(P) = d$. Die polynomiale Funtion mit $\mathbb{C} \longrightarrow \mathbb{C}$, $z \mapsto 0$ hat Grad $-\infty$, $\deg(0) = -\infty$. Hierbei ist $-\infty$ ein Symbol[23] mit den Eigenschaften $\forall n \in \mathbb{N} : \ -\infty < n$ und $\forall n \in \mathbb{N} : \ -\infty + n := -\infty$. Gilt $\deg(P) = d \in \mathbb{N}$, so nennt a_d den Leitkoeffizient.

Jede polynomiale Funktion P hat einen Grad $\deg(P) \in \{-\infty, 0, 1, \ldots\}$.

$$(P+Q)(z) := P(z)+Q(z), \quad (P \cdot Q)(z) := P(z) \cdot Q(z), \quad \deg(P \cdot Q) = \deg(P) + \deg(Q).$$

Theorem 7.7 (Fundamentalsatz der Algebra)**.** Sei $P : \mathbb{C} \longrightarrow \mathbb{C}$ ein Polynom von Grad $d \geq 1$. Dann gibt es ein $z \in \mathbb{C}$ mit $P(z) = 0$.

Ein solches z nennt man eine Nullstelle von P.

Mindestens einen Beweis dieses Theorems lernen wir in der Analyis III kennen.

Korollar 7.8. Zu jeder polynomialen Funktion P über $\mathbb{C}$ von Grad $d \geq 1$ gibt es komplexe Zahlen $b_1, \ldots, b_d$ mit

$$\forall z \in \mathbb{C} : \ P(z) = a_d(z - b_1)(z - b_2) \cdots (z - b_d).$$

Die Lösungsmenge der komplexen Zahlen z, die $P(z) = 0$ erfüllen, ist somit

$$\{b_j \mid j \in \{1, \ldots, d\}\}$$

[22]Sie haben im Rahmen der Schulmathematik hierfür vielleicht das Wort „Polynom“ statt „polynomialer Funtion“ benutzt, dies ist nun aber nicht mehr erlaubt, denn das Wort „Polynom“ hat in der Mathematik eine etwas andere Bedeutung, siehe Lineare Algebra.

[23]Man sollte sich nicht fragen, was $-\infty$ ist. Es ist nur ein Symbol, das $\mathbb{N}$ erweitert zu $\{-\infty\} \cup \mathbb{N}$, und die Definition der Ordnung $\leq$ und der Addition $+$ erweitern sich wie in diesen Eigenschaften beschrieben.

und hat deswegen höchstens d Elemente und mindestens 1 Element.

Beweis. Man zeigt zuerst die Zerlegung $P(z) = a_n(z - b_1)(z - b_2)\cdots(z - b_d)$ durch Induktion über d.

Induktionsanfang.
Ist P eine polynomiale Funktion von Grad 1, dann gilt $P(z) = a_1 z + a_0$ mit $a_1 \neq 0$. Daraus ergibt sich

$$P(z) = a_1 \left(z + \frac{a_0}{a_1} \right),$$

also die Aussage mit $b_1 = -a_0/a_1$.

Induktionsschritt.
Sei $d \geq 2$ gegeben, und eine polynomiale Funktion P von Grad d. Das Korollar gelte für $d - 1$. Nach dem Fundamentalsatz der Algebra gibt es eine Nullstelle b_d von P. Wir dividieren $P(z)$ durch $(z - b_d)$:

$$P(z) = Q(z)(z - b_d) + R(z),$$

wobei Q eine polynomiale Funktion von Grad $d - 1$ ist und R eine polynomiale Funktion ist, die den Divisionsrest angibt. Nun gilt[24] $\deg(R) < \deg(z - b_d) = 1$, also $\deg(R) = 0$ oder $R = 0$. Da offensichtlich $R(b_d) = 0$, folgt $R = 0$. Die polynomialen Funktionen Q und P haben denselben Leitkoeffizient, den wir a nennen. Wir wenden nun die Induktionsvoraussetzung auf Q an und schreiben:

$$Q(z) = a(z - b_1)\cdots(z - b_{n-1}),$$

[24]Beweis: Lineare Algebra oder Mathematische-Methoden-Vorlesung

woraus wir

$$P(z) = a(z - b_1) \cdots (z - b_n)$$

erhalten.

Nun rechnet man leicht $P(b_j)$ für alle $j \in \{1, \ldots, n\}$ nach. Umgekehrt, sieht man für $z \notin \{b_j \mid j \in \{1, \ldots, n\}\}$ leicht, dass $P(z)$ als Produkt von nicht-verschwindenden[25] Faktoren ebenfalls nicht-verschwindet. □

Ist w eine Nullstelle von P, dann nennt man

$$\#\{j = \Big\{1, 2, \ldots, d\} \mid w = b_j\Big\} \in \mathbb{N}$$

die Multiplizität der Nullstelle.

Bemerkung 7.9. Die Berechnung der Nullstellen ist im allgemeinen sehr schwierig. Für polynomiale Funktionen von Grad $d \leq 4$ gibt es eine Lösungsformel. In der Schule lernt man die für $d = 2$, die Formel für $d = 3$ und $d = 4$ sind zunehmend komplizierter. Für Grad $d \geq 5$ gibt es keine Lösungsformel, genauer man kann aus den Koeffizienten a_i, den Grundrechenarten $+, +, \cdot, /$ und Wurzelziehen keinen Ausdruck herleiten, der die Nullstellen angibt. Und man kann zeigen, dass es solche einen Ausdruck gar nicht gibt. Diese Aussagen werden in der Vorlesung „Algebra“ im Rahmen der Galois-Theorie behandelt. Die Galois-Theorie ist nach Évariste Galois benannt, dessen Leben und früher Tod eine sehr erstaunlich ist, siehe Wikipedia.

[25] „nicht-verschwindend“ ist eine noblere Form für „ungleich 0“

KAPITEL 3

Folgen und Reihen

1. Folgen

1.1. Konvergenz von Folgen. Im folgenden sei $K = \mathbb{Q}$, $K = \mathbb{R}$ oder $K = \mathbb{C}$. [1] Es gibt dann eine Betragsfunktion $K \to \mathbb{R}$, $x \mapsto |x|$.

$$|x + y| \leq |x| + |y|, \qquad |xy| \leq |x|\,|y|$$

DEFINITION 1.1. Eine (K-wertige) Folge ist eine Abbildung von $\mathbb{N}$ nach K.

Oft bezeichnet man auch auf $\mathbb{N}_{>k}$ definierte Abbildungen als Folgen, z.B. $\mathbb{N}_{>0} \to K$.

DEFINITION 1.2. Eine K-wertige Folge $(a_j)_{j\in\mathbb{N}}$ heißt beschränkt, wenn es ein $s \in \mathbb{R}$ gibt, so dass

$$\forall j \in \mathbb{N} : |a_j| \leq s.$$

Ist $K \subset \mathbb{R}$, dann ist $(a_j)_{j\in\mathbb{N}}$ beschränkt genau dann, wenn die Menge $\{a_j \mid j \in \mathbb{N}\}$ nach oben und unten beschränkt ist. Für $K \subset \mathbb{R}$ nennen wir $(a_j)_{j\in\mathbb{N}}$ nach oben bzw.[2] nach unten beschränkt genau dann, wenn $\{a_j \mid j \in \mathbb{N}\}$ nach oben bzw. nach unten beschränkt ist.

[1] Alles, was wir im folgenden zeigen werden, gilt auch für jeden archimedisch geordneten Körper. Da aber alle diese Körper zu einem Unterkörper von $\mathbb{R}$ isomorph sind, ist dies keine wirkliche Verallgemeinerung.

[2] Die Abkürzung „bzw.“ steht für „beziehungsweise“ und dieses Wort besagt, dass diese Definition nun in zwei Version gelesen werden kann. Zunächst mit den Worten „nach oben“ an beiden Stellen, aber

Die Begriffe obere Schranke, untere Schranke, Minimum, Maximum, Supremum, Infimum definiert man für Folgen analog, also z.B. $a \in \mathbb{R}$ ist obere Schranke von $(a_j)_{j\in\mathbb{N}}$ genau dann, wenn es obere Schranke von $\{a_j \mid j \in \mathbb{N}\}$ ist.

BEISPIELE 1.3.

(a) $(2j)_{j\in\mathbb{N}}$ ist nach unten, aber nicht nach oben beschränkt.
(b) $(3+4/j)_{j\in\mathbb{N}_{>0}}$ ist nach unten beschränkt durch 3 und nach oben beschränkt durch 7.

DEFINITION 1.4 (Konvergenz von Folgen). Eine K-wertige Folge $(a_j)_{j\in\mathbb{N}}$ konvergiert gegen $a \in K$, falls gilt

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists j_0 \in \mathbb{N} : \forall j \in \mathbb{N} : (j \geq j_0 \Longrightarrow |a_j - a| \leq \epsilon).$$

Man nennt a den Grenzwert oder Limes der Folge $(a_j)_{j\in\mathbb{N}}$ und schreibt $a = \lim_{j\to\infty} a_j$ oder $a_j \to a$ für $j \to \infty$. Wir sagen $(a_j)_{j\in\mathbb{N}}$ konvergiert (in K), falls es ein derartiges $a \in K$ gibt. Folgen, die gegen 0 konvergieren, nennt man Nullfolgen. Falls eine Folge nicht konvergiert, so sagen wir dazu sie divergiert.

!ACHTUNG!. Wenn wir $a = \lim_{j\to\infty} a_j$ schreiben, so bedeutet dies immer:

- der Grenzwert existiert, und
- der Grenzert ist a.

BEMERKUNGEN 1.5.

auch mit den Worten „nach unten“. Hierbei ist wichtig, dass im Haupt- und im Nebensatz denselben Begriff nutzt.

(a) Falls eine Folge einen Grenzwert besitzt, so ist dieser eindeutig bestimmt. Seien a und a' zwei Grenzwerte von $(a_j)_{j\in\mathbb{N}}$. Zu einem gegebenen $\epsilon \in \mathbb{R}_{>0}$ gilt also:

$$\exists j_0 \in \mathbb{N} : \forall j \in \mathbb{N} : (j \geq j_0 \Longrightarrow |a_j - a| \leq \epsilon).$$
$$\exists j_0' \in \mathbb{N} : \forall j \in \mathbb{N} : (j \geq j_0' \Longrightarrow |a_j - a'| \leq \epsilon).$$

Wähle nun so ein j_0 und j_0'. Dann gilt für alle $j \geq \max\{j_0, j_0'\}$:

$$|a - a'| \leq |a - a_j| + |a_j - a'| \leq \epsilon + \epsilon = 2\epsilon.$$

Dies gilt für alle $\epsilon \in \mathbb{R}_{>0}$. Angenommen wir haben $a \neq a'$, so gilt dies insbesondere für $\epsilon := |a - a'|/3 > 0$. Also folgt $3\epsilon \leq 2\epsilon$ und somit ergibt sich der Widerspruch $\epsilon \leq 0$. Die Annahme $a \neq a'$ war also falsch, d.h. es gilt $a = a'$.

(b) Falls $(a_j)_{j\in\mathbb{N}}$ konvergiert, so ist $(a_j)_{j\in\mathbb{N}}$ beschränkt. Begründung: sei $a := lim_{j\to\infty} a_j$. Wir wählen zu $\epsilon := 1$ ein passendes j_0. Es gilt somit für alle $j \in \mathbb{N}$ mit $j \geq j_0$:

$$|a_j| \leq |a_j - a| + |a| \leq |a| + 1.$$

Nun setzen wir

$$r := \max\{|a_0|, |a_1|, \ldots, |a_{j_0-1}|, |a| + 1\}.$$

Dann gilt für alle $j \in \mathbb{N}$: $|a_j| \leq r$. Somit ist $(a_j)_{j\in\mathbb{N}}$ beschränkt.

(c) Zu jedem $\epsilon \in \mathbb{R}_{>0}$ existiert ein $n \in \mathbb{N}_{>0}$ mit $\epsilon \geq 1/n$ (siehe Übungsblatt 5 Aufgabe 3). Deswegen gilt[3]

$$a = \lim_{j\to\infty} a_j \quad \text{gdw} \quad \forall n \in \mathbb{N}_{>0} : \exists j_0 \in \mathbb{N} : \forall j \in \mathbb{N} : (j \geq j_0 \Longrightarrow |a_j - a| \leq \frac{1}{n}).$$

Beispiele 1.6.

(a) Eine Folge $(a_j)_{j\in\mathbb{N}}$ heißt konstant, falls $a_0 = a_1 = a_2 = \ldots$. Konstante Folgen sind beschränkt und konvergieren. $a_0 = \lim_{j\to\infty} a_i$.

(b) $(\frac{1}{j})_{j\in\mathbb{N}_{>0}}$ ist eine Nullfolge. (Nutze z.B. Bem. 1.5 (c) und setze $j_0 := n$).

(c) Die Folge $\left((-1)^j\right)_{j\in\mathbb{N}}$ hat 1 als obere und -1 als untere Schranke und ist somit beschränkt. Wir werden bald sehen, dass sie nicht konvergiert.

(d) Gilt $|q| < 1$, dann ist $(q^j)_{j\in\mathbb{N}}$ eine Nullfolge. (Hierbei ist $q^0 := 1$.)
Um dies zu zeigen, bestimme mit dem folgenden Lemma zu gegebenem $\epsilon \in \mathbb{R}_{>0}$ ein $n \in \mathbb{N}$ mit $|q|^n < \epsilon$. Dann gilt für alle $j \in \mathbb{N}$, $j \geq n$:

$$|q^j| = |q|^j = \underbrace{|q|^{j-n}}_{\leq 1} |q|^j \leq \epsilon.$$

Wir haben also gezeigt:

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists n \in \mathbb{N} : \forall j \in \mathbb{N} : \left(j \geq n \Longrightarrow |q^j - 0| \leq \epsilon\right).$$

Lemma 1.7. Sei $r \in \mathbb{R}_{>0}$ mit $r < 1$ und $\epsilon \in \mathbb{R}_{>0}$. Dann gibt es ein $n \in \mathbb{N}$ mit $r^n < \epsilon$.

[3]Um die Gefahr von Verwechslungen zwischen dem Grenzwert-Pfeil (z.B. $j \to \infty$) und dem Pfeil für logische Implikation (z.B. $(a \wedge b) \to (c \vee d)$ klein z halten, benutzen wir ab nun oft $\Longrightarrow$ an Stellen, wo wir bisher $\to$ benutzt hätten, z.B. in der folgenden Formel

Beweis. Offensichtlich gilt

$$r^n < \epsilon \iff$$
$$\left(\frac{1}{r}\right)^n > \frac{1}{\epsilon} \tag{1.8}$$

Auf Übungsblatt 4 Aufgabe 1 c) haben wir für $x \in \mathbb{R}_{\geq -1}$ die Bernoulli-Ungleichung gezeigt:

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

Wende die Bernoulli für $x := (1/r) - 1 > 0$ an und erhalte

$$\left(\frac{1}{r}\right)^n \geq 1 + nx\,. \tag{1.9}$$

Aufgrund der archimedischen Eigenschaft existiert ein $n \in \mathbb{N}$ mit

$$n > \frac{1}{x}\left(\frac{1}{\epsilon} - 1\right)\,.$$

und dies ist äquivalent zu

$$1 + nx > \frac{1}{\epsilon}\,,$$

und mit (1.9) impliziert dies (1.8) und somit $r^n < \epsilon$. □ **Fr 30.11.**

PROPOSITION 1.10. Seien $(a_j)_{j\in\mathbb{N}}$ und $(b_j)_{j\in\mathbb{N}}$ konvergente Folgen. Dann gilt:

(1) $$\lim_{j\to\infty}(a_j + b_j) = \lim_{j\to\infty} a_j + \lim_{i\to\infty} b_j$$

(2) $$\lim_{j\to\infty}(a_j - b_j) = \lim_{j\to\infty} a_j - \lim_{i\to\infty} b_j$$

(3) $$\lim_{j\to\infty}(a_j \cdot b_j) = \lim_{j\to\infty} a_j \cdot \lim_{i\to\infty} b_j$$

Gilt zusätzlich: $\forall i \in \mathbb{N} : b_j \neq 0$, und ist $(b_j)_{i\in\mathbb{N}}$ keine Nullfolge, so gilt auch

$$\lim_{j\to\infty} \frac{a_j}{b_j} = \frac{\lim\limits_{j\to\infty} a_j}{\lim\limits_{i\to\infty} b_j}. \tag{4}$$

Beweis von Proposition 1.10 (1).

(1): Angenommen $a = \lim_{j\in\mathbb{N}} a_j$ und $b = \lim_{j\in\mathbb{N}} b_j$. Dies bedeutet, dass wir für jedes $\epsilon \in \mathbb{R}_{>0}$ die folgenden Aussagen haben:

$$\exists j_0 \in \mathbb{N} : \forall j \in \mathbb{N} : (j \geq j_0 \Longrightarrow |a_j - a| \leq \epsilon)$$

$$\exists k_0 \in \mathbb{N} : \forall j \in \mathbb{N} : (j \geq k_0 \Longrightarrow |b_j - b| \leq \epsilon)$$

Wir wählen nun solch ein j_0 und solch ein k_0. Dann gilt für $\ell_0 := \max\{j_0, k_0\}$.

$$\forall j \in \mathbb{N} : (j \geq \ell_0 \Longrightarrow |a_j - a| \leq \epsilon \wedge |b_j - b| \leq \epsilon).$$

Wir rechnen für $j \geq \ell_0$:

$$|(a_j + b_j) - (a + b)| = |(a_j - a) + (b_j - b)| \leq |a_j - a| + |b_j - b| \leq \epsilon + \epsilon = 2\epsilon.$$

Also ergibt sich insgesamt

$$\forall \epsilon \in \mathbb{R}_{>0} : \underbrace{\exists \ell_0 \in \mathbb{N} : \forall j \in \mathbb{N} : \Big(j \geq \ell_0 \Longrightarrow |(a_j + b_j) - (a + b)| \leq 2\epsilon.\Big)}_{A(2\epsilon):=}$$

Nach dem folgenden Lemma können wir in dieser Aussage 2ϵ durch ϵ zu ersetzen. Somit ist

$$a + b = \lim_{j\to\infty} (a_j + b_j)$$

und damit auch (1) gezeigt. □

Lemma 1.11. Sei $A(\,\cdot\,)$ eine auf $\mathbb{R}_{>0}$ definierte Aussageform, und $q \in \mathbb{R}_{>0}$. Dann gilt

$$\forall \epsilon \in \mathbb{R}_{>0} : \; A(\epsilon) \quad \Longleftrightarrow \quad \forall \epsilon \in \mathbb{R}_{>0} : \; A(q\epsilon)$$

Das Lemma gibt es vielen Variationen. Wichtiger als die Aussage des Lemmas ist es, zu verstehen, wie man das Lemma (oder eine Variation davon!) kurz beweist.

Beweis des Lemmas.
„$\Longrightarrow$“: Es gelte

$$\forall \widehat{\epsilon} \in \mathbb{R}_{>0} : \; A(\widehat{\epsilon}). \tag{1.12}$$

Für ein gegebenes $\epsilon > 0$ wollen wir nun $A(q\epsilon)$ zeigen. Wir wenden (1.12) für $\widehat{\epsilon} := q\epsilon$ an, und haben dann das gewünschte.

„$\Longleftarrow$“: Analog mit $\widehat{\epsilon} := q^{-1}\epsilon$ □

Beweis von Proposition 1.10 (Fortsetzung).
(2): Der Beweis für $(a_j - b_j)_{j\in\mathbb{N}}$ ist völlig analog zur Summe.

(3): Sei wieder $a := \lim_{j\to\infty} a_j$ und $b := \lim_{j\to\infty} b_j$. Für das Produkt $(a_j \cdot b_j)_{j\in\mathbb{N}}$ muss man etwas anders vorgehen. Zunächst nutzen wir die Tatsache, dass $(a_j)_{j\in\mathbb{N}}$ beschränkt ist. Also gibt es ein $r \in \mathbb{R}$ mit

$$\forall j \in \mathbb{N} : \; |a_j| \leq r$$

Dann argumentieren wir wie bei der Summe, rechnen aber wie folgt:

$$|a_j b_j - ab| = |a_j(b_j - b) + (a_j - a)b| \leq |a_j||b_j - b| + |a_j - a||b| \leq r\epsilon + \epsilon|b| = (r + |b|)\epsilon.$$

Nun argumentiert man wie bei der Summe, wobei man das Lemma mit $q := r + |b|$ nutzt.

(4): Sei $b := \lim_{j\to\infty} b_j \neq 0$. Wir zeigen, dass $1/b_j$ gegen $1/b$ konvergiert. Die Aussage (4) folgt dann mit (3).

Für $\epsilon_1 := |b|/2$ gibt es ein $j_0 \in \mathbb{N}$, so dass für alle $j \in \mathbb{N}_{\geq j_0}$ gilt $|b_j - b| \leq \epsilon_1 = |b|/2$. Nach Lemma 5.11 (g) haben wir dann

$$|b_j| \geq |b| - |b - b_j| \geq |b| - \frac{|b|}{2} = \frac{|b|}{2}.$$

Wir nutzen nun wieder die Definition des Grenzwerts: für jedes $\epsilon \in \mathbb{R}_{>0}$ haben wir

$$\exists k_0 \in \mathbb{N} : \forall j \in \mathbb{N} : (j \geq k_0 \Longrightarrow |b_j - b| \leq \epsilon)$$

Für $j \geq \ell_0 := \max\{j_0, k_0\}$ rechnen wir

$$\left|\frac{1}{b_j} - \frac{1}{b}\right| = \left|\frac{b - b_j}{bb_j}\right| = \frac{|b - b_j|}{|b|\,|b_j|} \leq \frac{\epsilon}{(|b|/2)\,|b|} = \frac{2}{|b|^2}\epsilon\,.$$

Wir haben also gezeigt

$$\forall \epsilon > 0 : \exists \ell_0 \in \mathbb{N} : \forall j \in \mathbb{N} : \left(j \geq \ell_0 \Longrightarrow \left|\frac{1}{b_j} - \frac{1}{b}\right| \leq \frac{2}{|b|^2}\epsilon\right).$$

Aus dem Lemma mit $q := \frac{2}{|b|^2}$ folgt dann

$$\frac{1}{b} = \lim_{j\to\infty} \frac{1}{b_j}\,.$$

1.2. Monotone Folgen.

DEFINITION 1.13. Seien $(M, \leq)$ und $(N, \leq)$ geordnete Mengen, $f : M \to N$. Wir sagen f ist

$$\begin{aligned}
\text{monoton wachsend} &:\iff \forall i, j \in M : \big(i \leq j \Longrightarrow f(i) \leq_f (j)\big)\big) \\
&\iff f \text{ erhält die Ordnung} \\
\text{monoton fallend} &:\iff \forall i, j \in M : \big(i \leq j \Longrightarrow f(i) \geq f(j)\big) \\
&\iff f \text{ ist ordnungsumkehrend} \\
\text{streng monoton wachsend} &:\iff \forall i, j \in M : \big(i < j \Longrightarrow f(i) < f(j)\big) \\
\text{streng monoton fallend} &:\iff \forall i, j \in M : \big(i < j \Longrightarrow f(i) > f(j)\big)
\end{aligned}$$

$$\begin{aligned}
\text{monoton:} &\iff f \text{ ist monoton wachsend oder monoton fallend} \\
\text{streng monoton:} &\iff f \text{ ist streng monoton wachsend oder streng monoton fallend} \\
&\iff f \text{ ist monoton und injektiv}
\end{aligned}$$

BEISPIELE 1.14.

(1) $\mathbb{R} \longrightarrow \mathbb{R}$, $x \mapsto x^3$ ist streng monoton wachsend,

(2) $(n)_{n\in\mathbb{N}}$ und $(1 - \frac{1}{n})_{n\in\mathbb{N}_{>0}}$ sind streng monoton wachsende Folgen.

(3) Für $r \in \mathbb{R}$ ist

$$\lfloor r \rfloor := \max\big(\underline{S}(\{r\}, \mathbb{R}) \cap \mathbb{Z}\big),$$

also die größte ganze Zahl $\leq r$. Analog ist $\lceil r \rceil$ als die kleinste ganze Zahl $\geq r$ definiert. Dann ist $(\lfloor \frac{n}{2} \rfloor)_{n\in\mathbb{N}} = (0, 0, 1, 1, 2, 2, 3, 3, \ldots)$ monoton wachsend, aber nicht streng monoton. Gleiches gilt für $(\lceil \frac{n}{2} \rceil)_{n\in\mathbb{N}} = (0, 1, 1, 2, 2, 3, 3, \ldots)$.

(4) Injektive monotone Funktionen sind streng monoton.

(5) Ist $(M, \leq)$ total geordnet, $(N, \leq)$ partiell geordnet und ist $f : M \longrightarrow N$ streng monoton, dann ist $f : M \longrightarrow N$ injektiv. Denn seien $x, y \in M$. Es gilt $x < y$, $x = y$ oder $x > y$. Falls f streng monoton wachsend ist und falls $x < y$, dann folgt $f(x) < f(y)$, also $f(x) \neq f(y)$. Die anderen Fälle sind analog.

Nun betrachten wir reell-wertige Folgen, also $M = \mathbb{N}$, $N = \mathbb{R}$.

Monoton wachsende (bzw. fallende) Folgen $(a_i)_{i\in\mathbb{N}}$ haben a_0 als Minimum (bzw. Maximum), sind also immer nach unten (bzw. oben) beschränkt.

Proposition 1.15. Alle beschränkten monotonen $\mathbb{R}$-wertige Folgen konvergieren (in $\mathbb{R}$).

Beweis. Sei $(a_j)_{j\in\mathbb{N}}$ eine beschränkte monotone Folge, o.B.d.A. monoton wachsend, $a_j \in \mathbb{R}$. Definieren $A := \{a_j \mid j \in \mathbb{N}\}$, $a := \sup A$. Zu jedem $\epsilon \in \mathbb{R}_{>0}$ gibt es ein $j_0 \in \mathbb{N}$ mit $a_{j_0} > a - \epsilon$, denn sonst wäre $a - \epsilon$ eine obere Schranke von A.

Also gilt für alle $j \in \{j_0, j_0 + 1, \ldots\}$: $a - \epsilon < a_{j_0} \leq a_j \leq a$. Daraus ergibt sich $|a - a_j| \leq \epsilon$. Wir erhalten $a = \lim_{j\to\infty} a_j$. □

Beispiel 1.16. Sei

$$a_j := \sum_{k=1}^{j} \frac{1}{k^2 + k^5}.$$

Die Folge $(a_j)_{j\in\mathbb{N}}$ ist streng monoton. Später werden wir sehen: $(a_j)_{j\in\mathbb{N}}$ ist beschränkt. Somit konvergiert $(a_j)_{j\in\mathbb{N}}$.

In Abschnitt 2 über Reihen im aktuellen Kapitel werden wir viele ähnliche Beispiele sehen.

Beispiel 1.17. Wallissches Produkt und ein Algorithmus zur Wurzelberechnung, siehe Zentralübung oder [**22**, Abschnitt 5.3 und 5.4].

1.3. Teilfolgen. Von nun an wieder $K = \mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$.

Definition 1.18. Ist $f : \mathbb{N} \longrightarrow \mathbb{N}$ eine streng monotone wachsende Abbildung, dann nennt man $(a_{f(k)})_{k\in\mathbb{N}}$ eine Teilfolge von $(a_j)_{i\in\mathbb{N}}$.

Beispiel 1.19. Die Folge $(j)_{j\in\mathbb{N}} = (0, 1, 2, 3, \ldots)$ hat unter anderem folgende Teilfolgen: [4]

- sich selbst, das heißt $(j)_{j\in\mathbb{N}} = (k)_{k\in\mathbb{N}}$, also $f(k) = k$
- die Folge der ungeraden Zahlen $(1, 3, 5, 7, \ldots)$, also $f(k) = 2k + 1$
- die Folge der Primzahlen $(2, 3, 5, 7, \ldots)$.

Lemma 1.20. Konvergiert $(a_j)_{j\in\mathbb{N}}$ gegen a, so konvergiert jede Teilfolge ebenfalls gegen a.

Beweis. Sei $(a_{f(k)})_{k\in\mathbb{N}}$ eine Teilfolge von $(a_j)_{j\in\mathbb{N}}$, also $f : \mathbb{N} \longrightarrow \mathbb{N}$ streng monoton.

Sei $\epsilon \in \mathbb{R}_{>0}$ gegeben. Wir haben

$$\exists j_0 \in \mathbb{N} : \ \forall j \in \mathbb{N}_{\geq j_0} : \ |a_j - a| \leq \epsilon.$$

[4]Diese Folge hat natürlich noch **viel** mehr Teilfolgen!

Wähle zu solch einem j_0 ein $k_0 \in \mathbb{N}$ mit $f(k_0) \geq j_0$, zum Beispiel $k_0 := j_0$. Dann gilt für dieses k_0

$$\forall k \in \mathbb{N}_{\geq k_0} : |a_{f(k)} - a| \leq \epsilon.$$

Insgesamt also

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists k_0 \in \mathbb{N} : \forall k \in \mathbb{N} : (k \geq k_0 \Longrightarrow |a_k - a| \leq \epsilon).$$

Ingesamt also $a = \lim_{k\to\infty} a_{f(k)}$. □

ANWENDUNG 1.21. Die Folge $\left((-1)^j\right)_{j\in\mathbb{N}}$ aus Beispiele 1.6 (c) divergiert.

Beweis. Angenommen, die Folge $\left((-1)^j\right)_{j\in\mathbb{N}}$ konvergiert gegen ein $a \in K$. Dann konvergieren auch die Teilfolgen $\left((-1)^{2k}\right)_{k\in\mathbb{N}} = (1)_{k\in\mathbb{N}}$ und $\left((-1)^{2k+1}\right)_{k\in\mathbb{N}} = (-1)_{k\in\mathbb{N}}$ gegen a.
Mi 5.12. Da diese Teilfolgen konstant sind, erhalten wir den Widerspruch $a = 1$ und $a = -1$. □

Sei $K = \mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$.

DEFINITION 1.22. Sei $(a_i)_{i\in\mathbb{N}}$ eine K-wertige Folge, $a \in K$. Wir sagen a ist ein Häufungspunkt der Folge $(a_i)_{i\in\mathbb{N}}$, falls eine Teilfolge existiert, die gegen a konvergiert.

BEISPIELE 1.23. Die Häufungspunkte der Folge $\left((-1)^n\right)_{n\in\mathbb{N}}$ sind -1 und $+1$.

$K = \mathbb{C}$. Die Häufungspunkte der Folge $\left(i^n\right)_{n\in\mathbb{N}} = (1, i, -1, -i, \ldots)$ sind -1, $+1$, $-i$ und $+i$.

1.4. Erweiterte reelle Zahlen und uneigentliche Konvergenz. In diesem Unterabschnitt ist immer $K = \mathbb{R}$.

Nach unseren bisherigen Definition besitzen zum Beispiel die folgenden Teilmenge von $\mathbb{R}$ weder Supremum noch Infimum: $\emptyset$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, ...

NOTATION 1.24. Wir definieren $\sup \emptyset := -\infty$, $\inf \emptyset := \infty$. Und $\sup M := \infty$, falls M nicht nach oben beschränkt. Und $\inf M := -\infty$, falls M nicht nach unten beschränkt.

Fortsetzung der Ordnung auf $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. $\infty = +\infty$.

$$\forall x \in \overline{\mathbb{R}} : \ -\infty \leq x \leq \infty.$$

Man nennt $\overline{\mathbb{R}}$ die erweiterten reellen Zahlen. Addition und Multiplikation sind hierauf nicht definiert.

Intervalle[5]: Seien $a, b \in \overline{\mathbb{R}}$, $a \leq b$

$(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$ offenes Intervall
$[a, b) := \{x \in \overline{\mathbb{R}} \mid a \leq x < b\}$ halboffenes Intervall
$(a, b] := \{x \in \overline{\mathbb{R}} \mid a < x \leq b\}$ halboffenes Intervall
$[a, b] := \{x \in \overline{\mathbb{R}} \mid a \leq x \leq b\}$ abgeschlossenes Intervall

DEFINITION 1.25. Sei $(a_n)_{n\in\mathbb{N}}$ eine reell-wertige Folge. Wir sagen

$$(a_n)_{n\in\mathbb{N}} \text{ konvergiert gegen unendlich}$$
$$\iff \ \forall x \in \mathbb{R} : \ \exists n_0 \in \mathbb{N} : \ \forall n \in \mathbb{N}_{\geq n_0} : \ a_n \geq x$$

Alternative Ausdrucksweisen:
$a_n \to \infty$ für $n \to \infty$
$(a_n)_{n\in\mathbb{N}}$ konvergiert gegen $\pm\infty$.
$\lim_{j\to\infty} a_j = \infty$

[5] An Stelle von $(0, 1]$ schreibt man in Schulen oft $]0, 1]$. Dies ist in der Mathematik in den Universitäten weniger üblich, da eine Klammer] am Ende eines einzuklammernden Ausdrucks stehen sollte. In den meisten Büchern schreibt man deswegen $(0, 1]$.

Wir sagen $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen $-\infty$, wenn $(-a_n)_{n\in\mathbb{N}}$ gegen $+\infty$ konvergiert. Konvergenz gegen ∞ oder $-\infty$ nennt man uneigentliche Konvergenz.[6] Wir sagen, dass ∞ (bzw. $-\infty$) ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ ist, falls eine Teilfolge gegen ∞ (bzw. $-\infty$) konvergiert.

Wir nutzen den Begriff „eigentlich konvergent“ inhaltlich äquivalent zu „konvergiert“, das Wort „eigentlich“ wird benutzt, wenn man nochmals unterstreichen will, dass es sich nicht um „uneigentliche Konvergenz“ handelt.

Wir sagen eine Folge konvergiert in $\overline{\mathbb{R}}$, falls sie eigentlich oder uneigentlich konvergiert.

Beispiel 1.26. Sei $k \in \mathbb{N}_{>0}$. Dann gilt:

$$\lim_{n\to\mathbb{N}} n^k = \infty.$$

Wenn wir $a = \lim_{n\to\infty} a_n$ schreiben, so bedeutet dies ab jetzt (für $K = \mathbb{R}$):

- der Grenzwert existiert im eigentlichen oder uneigentlichen Sinn,
- der Grenzwert ist $a \in \overline{\mathbb{R}}$.

!ACHTUNG!. Wenn wir ohne weitere Zusätze sagen: „die reell-wertige Folge konvergiert“, dann ist hier immer die Konvergenz im *eigentlichen* Sinne gemeint, d.h. gegen eine reelle Zahl. Uneigentlich konvergente Folgen sind somit immer divergent. [7]

[6] In manchen Büchern, z.B. [**14**] wird der Begriff „uneigentliche Konvergenz“ durch „bestimmte Divergenz“ ersetzt.

[7] Aber wir haben schon gesehen, dass es divergente Folgen gibt, die nicht uneigentlich konvergieren, z.B. $((-1)^n)_{n\in\mathbb{N}}$.

!ACHTUNG!. In der gesamten Analysis I und II gilt: Im Fall $K = \mathbb{C}$ gibt es keine uneigentliche Konvergenz von Folgen. Wer diesen Begriff in der Klausur oder auf dem Übungsblatt für $K = \mathbb{C}$ nutzt, macht einen Fehler.

LEMMA 1.27. Eine Folge ist genau dann nach oben (bzw. nach unten) beschränkt, wenn ∞ (bzw. $-\infty$) kein Häufungspunkt ist.

Beweis. Sei $(a_j)_{j\in\mathbb{N}}$ eine Folge.

Angenommen ∞ ist ein Häufungspunkt. Sei $f : \mathbb{N} \to \mathbb{N}$ ist streng monoton wachsend und $\lim_{n\to\infty} a_{f(n)} = \infty$. Für $x \in \mathbb{R}$ gilt dann

$$\exists n_0 \in \mathbb{N} : \forall n \in \mathbb{N}_{\geq n_0} : a_{f(n)} \geq x.$$

Also ist $x - 1$ keine obere Schranke. Also $(a_j)_{j\in\mathbb{N}}$ unbeschränkt.

Sei $(a_j)_{j\in\mathbb{N}}$ nach oben unbeschränkt.

$$\forall x \in \mathbb{R} : \exists k \in \mathbb{N} : a_k > x \tag{1.28}$$

Definiere induktiv $f : \mathbb{N} \to \mathbb{N}$. Wähle $f(0) := 0$. Seien nun $f(0), f(1), \ldots, f(n)$ für $n \in \mathbb{N}$ definiert. Wende (1.28) für $x := \max\{a_{f(0)}, a_{f(1)} \ldots, a_{f(n)}, n\}$ an und erhalte ein k. Setze $f(n+1) := k$. Dann konvergiert die Teilfolge $(a_{f(n)})_{n\in\infty}$ gegen ∞.

Also ist die Folge nach oben unbeschränkt genau dann, wenn ∞ Häufungspunkt ist.

$$
\begin{aligned}
&(a_j)_{j\in\mathbb{N}} \text{ nach unten beschränkt}\\
\Longleftrightarrow &(-a_j)_{j\in\mathbb{N}} \text{ nach oben beschränkt}\\
\Longleftrightarrow &\infty \text{ kein Häufungspunkt von } (-a_j)_{j\in\mathbb{N}}\\
\Longleftrightarrow &-\infty \text{ kein Häufungspunkt von } (a_j)_{j\in\mathbb{N}}
\end{aligned}
$$

□

PROPOSITION 1.29. Jede monotone $\mathbb{R}$-wertige Folge konvergiert eigentlich oder uneigentlich (in $\overline{\mathbb{R}}$).

Beweis. Ist die Folge beschränkt, so wissen wir bereits, dass sie eigentlich konvergiert.

Sei $(a_j)_{j\in\mathbb{N}}$ unbeschränkt und o.B.d.A. monoton wachsend. Dann ist die Folge also nach oben unbeschränkt. Dies impliziert insbesondere

$$\forall x \in \mathbb{R} : \exists j_0 \in \mathbb{N} : \ a_{j_0} > x.$$

Wegen der Monotonie gilt dann auch $a_j > x$ für alle $j \geq j_0$. Wir erhalten also $\lim_{j\to\infty} a_j = \infty$.

Wir definieren nun die Abbildung $\varphi : [-1, 1] \longrightarrow \overline{\mathbb{R}}$,

$$\varphi(x) = \begin{cases} \frac{1}{1-x} - \frac{1}{1+x} & \text{für } x \in (-1, 1) \\ \infty & \text{für } x = 1 \\ -\infty & \text{für } x = -1 \end{cases}$$

Diese Abbildung ist streng monoton und bijektiv. Sei $\psi := \varphi^{-1} : \overline{\mathbb{R}} \longrightarrow [-1, 1]$.

PROPOSITION 1.30. Sei $(a_n)_{n\in\mathbb{N}}$ eine $\mathbb{R}$-wertige Folge. Es gilt für alle $a \in \overline{\mathbb{R}}$:

$$a = \lim_{n\to\infty} a_n \iff \psi(a) = \lim_{n\to\infty} \psi(a_n)$$
$$a \text{ Häufungspunkt von } (a_n)_{n\in\mathbb{N}} \iff \psi(a) \text{ Häufungspunkt von } (\psi(a_n))_{n\in\mathbb{N}}$$

Graphen der Funktion $x \mapsto \varphi(x)$ und ihrer Umkehrfunktion

In wenigen Wochen (Bemerkung 5.14 in Kapitel 4) werden wir starke Hilfsmittel haben, um diese Proposition ganz einfach zu zeigen. Der Beweis wird deswegen hier ausgelassen.

1.5. Limes inferior und superior. In diesem Unterabschnitt ist wiederum immer $K = \mathbb{R}$.

Sei $(a_j)_{j\in\mathbb{N}}$ eine $\mathbb{R}$-wertige Folge. Definiere für $k \in \mathbb{N}$

$$b_k := \inf\{a_j \mid j \in \mathbb{N}_{\geq k}\} \in [-\infty, \infty).$$

Es gilt $b_k \leq b_{k+1}$, denn für $A \subset B \subset \mathbb{R}$ gilt $\inf B \leq \inf A$.

Somit ist $(b_k)_{k\in\mathbb{N}}$ eine monoton wachsende $\overline{\mathbb{R}}$-wertige Folge und $\lim_{k\to\infty} b_k$ existiert in $\overline{\mathbb{R}}$. Im Fall $\forall k \in \mathbb{N} : b_k = -\infty$ definiert man hierfür $\lim_{k\to\infty} b_k = -\infty$.

DEFINITION 1.31. Sei $(a_i)_{i\in\mathbb{N}}$ eine $\mathbb{R}$-wertige Folge.

(1) Wir definieren den Limes inferior als

$$\liminf_{j\to\infty} a_j := \lim_{k\to\infty} \inf\{a_j \mid j \in \mathbb{N}_{\geq k}\} \in \overline{\mathbb{R}}.$$

(2) Der Limes superior ist

$$\limsup_{j\to\infty} a_j := -\liminf_{j\to\infty}(-a_j) = \lim_{k\to\infty} \sup\{a_j \mid j \in \mathbb{N}_{\geq k}\} \in \overline{\mathbb{R}}.$$

BEISPIEL 1.32. $\liminf_{j\to\infty}(-1)^j = -1$, $\limsup_{j\to\infty}(-1)^j = 1$

$$\liminf_{j\to\infty} a_j \leq \limsup_{j\to\infty} a_j$$

SATZ 1.33. Sei $(a_i)_{i\in\mathbb{N}}$ eine $\mathbb{R}$-wertige Folge.

(1) $\liminf_{j\to\infty} a_j \in \overline{\mathbb{R}}$ ist ein Häufungspunkt von $(a_j)_{j\in\mathbb{N}}$.
(2) $\limsup_{j\to\infty} a_j \in \overline{\mathbb{R}}$ ist ein Häufungspunkt von $(a_j)_{j\in\mathbb{N}}$.
(3) Konvergiert $(a_j)_{j\in\mathbb{N}}$ in $\overline{\mathbb{R}}$, dann haben wir

$$\liminf_{j\to\infty} a_j = \lim_{j\to\infty} a_j = \limsup_{j\to\infty} a_j$$

(4) Gilt $b := \liminf_{j\to\infty} a_j = \limsup_{j\to\infty} a_j \in \overline{\mathbb{R}}$, dann konvergiert $(a_j)_{j\in\mathbb{N}}$ gegen b.
(5) Ist $(a_{f(\ell)})_{\ell\in\mathbb{N}}$ eine Teilfolge von $(a_j)_{j\in\mathbb{N}}$, dann gilt

$$\begin{aligned} \liminf_{j\to\infty} a_j &\leq \liminf_{\ell\to\infty} a_{f(\ell)} \\ \limsup_{\ell\to\infty} a_{f(\ell)} &\leq \limsup_{j\to\infty} a_j \end{aligned}$$

(6) Ist b ein Häufungspunkt der Folge $(a_j)_{i\in\mathbb{N}}$ gilt

$$\liminf_{j\to\infty} a_j \leq b \leq \limsup_{j\to\infty} a_j.$$

Beweis.

Zu (1): Wir definieren wieder $b_k := \inf\{a_j \mid j \in \mathbb{N}_{\geq k}\}$. Sei $b := \liminf_{j\to\infty} a_j = \lim_{k\to\infty} b_k$.

Wir haben dann $b_k \leq b$.

Wir betrachten zunächst den Fall $b \in \mathbb{R}$. Wir geben eine rekursive Definition von $f : \mathbb{N} \to \mathbb{N}$ an. Wir wählen $f(0) := 0$. Zu gegebenem $n \in \mathbb{N}_{>0}$ nehmen wir nun an, dass $f(n-1)$ definiert ist und wir wollen nun $f(n)$ definieren. Wegen $b = \lim_{k\to\infty} b_k$ und $b_k \leq b$ gibt es ein $k_0 \in \mathbb{N}$, so dass für alle $k \in \mathbb{N}_{\geq k_0}$ die Ungleichung $b - \frac{1}{n} \leq b_k \leq b$ gilt. Wähle $k := \max\{k_0, f(n-1)+1\}$ nun fest.

Auf Grund der Definition von b_{k_0} als Infimum gibt es ein $\ell \in \mathbb{N}_{\geq k}$ mit $b_k \leq a_\ell \leq b_k + \frac{1}{n}$. Wir setzen $f(n) := \ell$ und erhalten somit

$$b - \frac{1}{n} \leq b_k \leq a_{f(n)} \leq b_k + \frac{1}{n} \leq b + \frac{1}{n}.$$

Das heißt, das Kriterium in Bemerkung 1.5 (c) für die Konvergenz $b = \lim_{n\to\infty} a_{f(n)}$ ist erfüllt.

Die Fälle $\liminf\limits_{j\to\infty} a_j = \infty$ und/oder $\liminf\limits_{j\to\infty} a_j = -\infty$ gehen ähnlich und werden auf dem Übungsblatt 8 behandelt.

Zu (2): Wende (1) auf die Folge $(-a_j)_{j\in\mathbb{N}}$ an. Wir erhalten eine Teilfolge $(-a_{f(n)})_{n\in\mathbb{N}}$ mit

$$\lim_{n\to\infty}(-a_{f(n)}) = \liminf_{j\to\infty}(-a_j) \in \overline{\mathbb{R}}.$$

Somit

$$\lim_{n\to\infty} a_{f(n)} = -\liminf_{j\to\infty}(-a_j) = \limsup_{j\to\infty} a_j \in \overline{\mathbb{R}}.$$

Zu (3): Wähle eine Teilfolge gemäß (1) die gegen $\liminf\limits_{j\to\infty} a_j$ konvergiert. Wegen Lemma 1.20 konvergiert sie auch gegen $\lim\limits_{j\to\infty} a_j$. Also $\liminf\limits_{j\to\infty} a_j = \lim\limits_{j\to\infty} a_j$. Genauso zeigt man die die andere Gleichung mit (2). **Fr 7.12.**

Zu (4): Wir haben

$$b_j := \inf\{a_k \mid k \geq j\} \leq a_j \leq \check{b}_j := \sup\{a_k \mid k \geq j\}$$

(b_j) ist monoton wachsend, $(\check{b}_j)$ monoton fallend, also beide konvergent in $\overline{\mathbb{R}}$.

$b = \lim_{j\to\infty} b_j = \lim_{j\to\infty} \check{b}_j$, also konvergiert nach Übungsblatt 7, Aufgabe 3 b) (Sandwich-Lemma) auch (a_j) gegen diesen Grenzwert.

Zu (5): Wir haben

$$\{a_j \mid j \in \mathbb{N}_{\geq f(\ell)}\} \supset \{a_{f(m)} \mid m \in \mathbb{N}_{\geq \ell}\}$$

und somit

$$b_{f(\ell)} = \inf\{a_j \mid j \in \mathbb{N}_{\geq f(\ell)}\} \leq \inf\{a_{f(m)} \mid m \in \mathbb{N}_{\geq \ell}\} =: \tilde{b}_\ell$$

Also

$$\begin{aligned} \liminf_{j\to\infty} a_j &= \lim_{j\to\infty} b_j \overset{\text{TF}}{=} \lim_{\ell\to\infty} b_{f(\ell)} \\ &\overset{(*)}{\leq} \lim_{\ell\to\infty} \tilde{b}_\ell = \liminf_{\ell\to\infty} a_{f(\ell)} \end{aligned}$$

(b_j) und $(\tilde{b}_\ell)$ sind monoton wachsende Folgen, also konvergieren sie in $\overline{\mathbb{R}}$
Zu TF: Übergang zu Teilfolge, siehe Lemma 1.20
$(*)$: siehe Übungsblatt 7, Aufgabe 3, a)

Zu (6): Folgt nun aus (3) und (5)

Aus Satz 1.33 folgt:

KOROLLAR 1.34. Jede reell-wertige Folge besitzt eine in $\overline{\mathbb{R}}$ konvergente Teilfolge.

In anderen Worten: Jede reell-wertige Folge besitzt einen Häufungspunkt in $\overline{\mathbb{R}}$.

KOROLLAR 1.35 (Satz von Bolzano-Weierstraß)**.** Jede beschränkte reell-wertige Folge besitzt eine (in $\mathbb{R}$) konvergente Teilfolge.

In anderen Worten: Jede beschränkte reell-wertige Folge besitzt einen Häufungspunkt in $\mathbb{R}$.

Beweis. Auf Grund des Korollars 1.34 besitzt jede reell-werte Folge mindestens einen Häufungspunkt in $\overline{\mathbb{R}}$. Wegen Lemma 1.27 sind $-\infty$ und ∞ keine Häufungspunkte, wenn die Folge beschränkt ist. Wir haben also einen Häufungspunkt in $\mathbb{R}$. □

Satz 1.36 (Satz von Bolzano-Weierstraß in $\mathbb{C}$)**.** Jede beschränkte $\mathbb{C}$-wertige Folge besitzt eine (in $\mathbb{C}$) konvergente Teilfolge.

Beweis. Sei $(a_n)_{n\in\mathbb{N}}$ eine beschränkte Folge, $a_n \in \mathbb{C}$. Dann sind auch $(\mathbf{Re}\, a_n)_{n\in\mathbb{N}}$ und $(\mathbf{Im}\, a_n)_{n\in\mathbb{N}}$ beschränkt. Es gibt somit eine streng monotone Abbildung $f_1 : \mathbb{N} \longrightarrow \mathbb{N}$, so dass $(\mathbf{Re}\, a_{f_1(n)})_{n\in\mathbb{N}}$ konvergiert. Da auch $(\mathbf{Im}\, a_{f_1(n)})_{n\in\mathbb{N}}$ beschränkt ist, gibt es eine streng monotone Abbildung $f_2 : \mathbb{N} \longrightarrow \mathbb{N}$, so dass auch $(\mathbf{Im}\, a_{f_1\circ f_2(n)})_{n\in\mathbb{N}}$ konvergiert. Wir folgern, dass $(a_{f_1\circ f_2(n)})_{n\in\mathbb{N}}$ in $\mathbb{C}$ konvergiert. □

1.6. Cauchy-Folgen. In diesem Unterabschnitt sei $K = \mathbb{C}$ oder sei K ein archimedisch geordneter Körper (z.B. $K = \mathbb{R}$ oder $K = \mathbb{Q}$).

Betragsfunktion $|\cdot| : K \to \mathbb{R}$, $x \mapsto |x|$.

Definition 1.37. Eine K-wertige Folge $(a_j)_{i\in\mathbb{N}}$ heißt K-wertige Cauchy-Folge, falls

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists j_0 \in \mathbb{N} : \forall j,k \in \mathbb{N}_{\geq j_0} : |a_j - a_k| \leq \epsilon.$$

Satz 1.38. Jede konvergente Folge ist eine Cauchy-Folge.

Beweis. Es gelte $\lim_{j\to\infty} a_j = a$. Das heißt: für alle $\epsilon \in \mathbb{R}_{>0}$ gibt es ein $j_0 \in \mathbb{N}$, so dass für alle natürlichen Zahlen $j \geq j_0$ gilt: $|a_j - a| \leq \epsilon$.

Für solch ein ϵ und ein passendes j_0 nehmen wir nun natürliche Zahlen $j \geq j_0$ und $k \geq j_0$ und rechnen nach:

$$|a_j - a_k| \leq |a_j - a| + |a - a_k| \leq \epsilon + \epsilon = 2\epsilon.$$

Wir haben nun also gezeigt:

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists j_0 \in \mathbb{N} : \forall j, k \in \mathbb{N}_{\geq j_0} : |a_j - a_k| \leq 2\epsilon.$$

Wegen Lemma 1.11 ist dies äquivalent zur definierenden Eigenschaft einer Cauchy-Folge. □

Lemma 1.39. Jede Cauchy-Folge ist beschränkt.

Beweis. Sei $(a_j)_{j\in\mathbb{N}}$ eine Cauchy-Folge. Wir wählen zu $\epsilon := 1$ ein passendes j_0 wie in Definition 1.37. Es gilt somit für alle $j \in \mathbb{N}$ mit $j \geq j_0$:

$$|a_j| \leq |a_j - a_{j_0}| + |a_{j_0}| \leq |a_{j_0}| + 1.$$

Nun setzen wir

$$r := \max\{|a_0|, |a_1|, \ldots, |a_{j_0-1}|, |a_{j_0}| + 1\}.$$

Dann gilt für alle $j \in \mathbb{N}$: $|a_j| \leq r$. Somit ist $(a_j)_{j\in\mathbb{N}}$ beschränkt.

Definition 1.40. Sei $K = \mathbb{C}$ oder ein archimedisch geordneter Körper. Wir nennen K vollständig, wenn jede K-wertige Cauchy-Folge in K konvergiert.

Satz 1.41. Sei $K = \mathbb{C}$ oder sei K ein archimedisch geordneter Körper, in dem jede beschränkte Folge einen Häufungspunkt besitzt. Dann ist K vollständig. Insbesondere sind $\mathbb{R}$ und $\mathbb{C}$ vollständig.

Beweis. Sei $(a_j)_{j\in\mathbb{N}}$ eine K-wertige Cauchy-Folge. Nach Lemma 1.39 ist die Folge beschränkt und besitzt deswegen aufgrund der Annahmen im Satz eine konvergente Teilfolge $(a_{f(n)})_{n\in\mathbb{N}}$, der Limes dieser Teilfolge sei $a \in K$. Zu gegebenem $\epsilon \in \mathbb{R}_{>0}$ bestimmen wir nun ein $n_0 \in \mathbb{N}$, so dass

$$\forall n \in \mathbb{N}_{\geq n_0} : |a_{f(n)} - a| \leq \epsilon\,.$$

Wir bestimmen zudem ein $j_0 \in \mathbb{N}$, so dass

$$\forall j, k \in \mathbb{N}_{\geq j_0} : |a_j - a_k| \leq \epsilon.$$

Wir wählen nun ein $n \in \mathbb{N}_{\geq n_0}$ mit $f(n) \geq j_0$. (Eine mögliche Wahl ist $n := \max\{j_0, n_0\}$, denn dann gilt $f(n) \geq n \geq j_0$.) Für $j \geq j_0$ erhalten wir

$$|a_j - a| \leq |a_j - a_{f(n)}| + |a_{f(n)} - a| \leq 2\epsilon\,.$$

Mit Lemma 1.11 folgt $a = \lim_{j\to\infty} a_j$. □

Bemerkung 1.42. Man kann zeigen: Ist K ein vollständiger archimedisch geordneter Körper, so erfüllt er die Supremums-Eigenschaft. Somit sind für einen archimedisch geordneten Körper die folgenden Eigenschaften äquivalent:

(1) K erfüllt die Supremums-Eigenschaft
(2) Jede beschränkte Folge besitzt einen Häufungspunkt (Bolzano-Weierstraß-Eigenschaft)
(3) K ist vollständig
(4) K ist isomorph zu $\mathbb{R}$

Beispiel 1.43. Wir konstruieren rekursiv eine $\mathbb{Q}$-wertige Folge (vergleiche Zentralübung)

$$a_0 := 2, \quad a_{n+1} = \frac{1}{2}\left(a_n + \frac{2}{a_n}\right).$$

In $\mathbb{R}$ – das heißt, wenn wir sie als $\mathbb{R}$-wertige Folge interpretieren – konvergiert diese Folge gegen $\sqrt{2} \notin \mathbb{Q}$. Deswegen ist sie eine Cauchy-Folge in $\mathbb{R}$ und in $\mathbb{Q}$.

Andererseits konvergiert sie nicht in $\mathbb{Q}$, denn wenn sie in $\mathbb{Q}$ gegen $a \in \mathbb{Q}$ konvergieren würde, so hätte sie in $\mathbb{R}$ auch $a \neq \sqrt{2}$ als Grenzwert, was nicht möglich ist. Deswegen konvergiert diese Cauchy-Folge nicht in $\mathbb{Q}$.

Cauchy-Folgen sind wichtig, denn man kann hiermit die Konvergenz von Folgen zeigen, ohne den Grenzwert im voraus angeben zu müssen.

2. Reihen

2.1. Motivation von Reihen: Dezimal-Darstellung reeller Zahlen. Notationen.

Eine Dezimal-Darstellung einer positiven reellen Zahl x ist gegeben durch eine Zahl $\ell \in \mathbb{Z}$ und einer Abbildung

$$\mathbb{Z}_{\geq \ell} \to \{0, 1, 2 \ldots, 9\}, \quad n \mapsto a_n.$$

Wir erhalten hieraus die reelle Zahl

$$x = \sum_{n=\ell}^{\infty} 10^{-n} \cdot a_n.$$

Nun sind mehrere Punkte ungeklärt:

- Was ist solch eine unendliche Summe? (Siehe Abschnitt 2.2)
- Hat jede positive reelle Zahl eine Dezimal-Darstellung?
- Wenn ja, ist sie eindeutig?

Sobald dies geklärt ist, kann man „schreiben“, falls $\ell \leq 0$:

$$x = a_\ell a_{\ell+1} \dots a_0, a_1 a_2 \dots$$

bzw. falls $\ell > 0$:

$$x = 0, \underbrace{0 \dots, 0}_{\ell-1 \text{ mal}} a_\ell a_{\ell+1} \dots\dots$$

Wir sagen:

$$(a_n)_{n \in \mathbb{Z}_{\geq \ell}} \text{ ist eine Dezimal-Darstellung von } x\,.$$

Beispiele 2.1.

(1)

$$\frac{1}{7} = 0, 142857142857 \dots = 0, \overline{142857}$$

$$\frac{1}{13} = 0, \overline{076923}$$

$$\frac{1}{17} = 0, \overline{0588235294117647}$$

$$\frac{1}{28} = 0, 03\overline{571428}$$

$$3, 14159265 \dots$$

(2) Die Dezimal-Darstellung ist periodisch, genau dann wenn die dadurch beschriebene zugehörige reelle Zahl eine rationale Zahl ist.

Problem: Man kann im allgemeinen nicht alle Ziffern hinschreiben, da es ja unendlich viele sind.

Übung 2.2. Gegeben sei die Dezimal-Darstellung $(a_n)_{n\in\mathbb{Z}}$ einer positiven rationalen Zahl $\frac{p}{q}$, $p, q \in \mathbb{N}$ teilerfremd. Wir zerlegen

$$q = 2^a 5^b q_1^{c_1} \cdots q_r^{c_r}$$

wobei $2, 5q_1, \ldots, q_r$ verschiedene Primzahlen sind, $a, b, c_1, \ldots, c_r \in \mathbb{N}$. Dann teilt die Periodenlänge die Zahl $\prod_{j=1}^{r}(q_j - 1)q_j^{c_j-1}$. Die Zahl der Ziffern zwischen dem Komma und dem Beginn der Periode ist $\max\{a, b\}$.

Im Fall $r = 1$ kann man dies mit Schulwissens beweisen. Einen Beweis erhält man (im Fall $r = 1$) eigentlich schon, wenn man sich genau anschaut, wie man in der 4. Klasse das Dividieren lernt. Für einen systematischeren Beweis, der auch $r > 1$ abdeckt, sind Kongruenzen hilfreich. Ein Web-Link hierzu ist auch

http://m.schuelerlexikon.de/mobile_mathematik/Endliche_und_periodische_Dezimalbrueche.htm

Im aktuellen Abschnitt wollen wir nun solche unendlichen Summen einführen. Man benutzt aber das Wort „Reihe“ an Stelle des Wortes „Unendliche Summe“.

Bemerkung 2.3. (a) Jede Dezimal-Darstellung „beschreibt“ eine reelle Zahl.
(b) Jede positive reelle Zahl besitzt eine Dezimal-Darstellung.
(c) Die Dezimal-Darstellung ist nicht eindeutig. Sind $(a_n)_{n\in\mathbb{Z}_{\geq \ell}}$, $a_\ell \neq 0$, und $(b_n)_{n\in\mathbb{Z}_{\geq k}}$, $b_k \neq 0$, verschiedene Dezimal-Darstellungen derselben positiven Zahl, so hat mindestens eine der beiden Darstellungen eine Periode $\overline{9}$.

Zu (a): Im nächsten Abschnitt wird klar werden, was hier mit „beschreibt“ gemeint ist, nämlich der Grenzwert einer konvergenten Reihe. Die Aussage wird auch im nächsten Abschnitt gezeigt.

Zu (b) und (c): Dies sollte nach dem nächsten Abschnitt jeder als Übungsaufgabe beweisen können. Tipp zu (b): Zunächst mal überlegt man sich zu einer gegebenen reellen Zahl x, wie eine Dezimaldarstellung aussehen muss, wenn sie existiert. Danach überlegt man sich, dass diese Darstellung tatsächlich die gegebene Zahl x beschreibt. Man muss hier u.a. die Supremums-Eigenschaft von $\mathbb{R}$ nutzen.

Beispiele 2.4.

(1) $0, \overline{9}$ und $1, \overline{0}$ sind Dezimal-Darstellungen von 1
(2) Jede positive reelle Zahle hat also eine eindeutige Periode-$\overline{9}$-freie Dezimal-Darstellung.

In der Zentralübung wurde bereits das folgende gezeigt:

Proposition 2.5. Es gibt keine surjektive Abbildung von $\mathbb{N}$ nach $\mathbb{R}$.

Wir sehen somit, dass $\mathbb{R}$ überabzählbar ist.

Beweis. Angenommen es gibt eine surjektive Abbildung $f : \mathbb{N} \longrightarrow \mathbb{R}$.

Für $k \in \mathbb{N}$ schreiben $f(k)$ als Dezimalzahl

$$f(k) = \pm a^k_{\ell_k} \dots a^k_0, a^k_1 a^k_2 \dots$$

O.B.d.A. können wir annehmen, dass $j(k)$ keine Periode $\overline{9}$ hat. Wir definieren $h : \{0, 1, \ldots, 9\} \longrightarrow \{0, 1, \ldots, 9\}$ durch

$$h(m) := \begin{cases} 1 & \text{für } m = 2 \\ 2 & \text{für } m \neq 2 \end{cases}$$

Also $h(m) \notin \{m, 9, 0\}$. Für $k \in \mathbb{N}$ setze nun $b_k := h(a_k^k)$, und dann

$$x := 0, b_1 b_2 b_3 \ldots$$

Nun ist $x \neq f(k)$, den die beiden reellen Zahlen haben in ihrer Periode-$\overline{9}$-freien Darstellung eine verschiedene Ziffer an der k-ten Stelle. Also gilt $x \notin B(f) = f_{\#}(\mathbb{N})$. Somit ist f gar
Mi 12.12. nicht surjektiv, entgegen der Annahme. □

2.2. Definition und elementare Eigenschaften. Ziel: Definiere $\sum_{j=0}^{\infty} a_j$

Alle Folgen in diesem Abschnitt sind Folgen in $K = \mathbb{R}$ oder $K = \mathbb{C}$. In anderen Worten es sind $\mathbb{R}$-wertige oder $\mathbb{C}$-wertige Folgen. Der Fall $\mathbb{Q}$-wertiger Folgen ist natürlich damit auch abgedeckt, da $\mathbb{Q} \subset \mathbb{R}$, wir diskutieren aber nicht die Konvergenz in $\mathbb{Q}$, sondern nur die Konvergenz in $\mathbb{R}$.

Beachte im folgenden: eine Folge $(a_n)_{n\in\mathbb{N}}$ ist genau dann eine Nullfolge, wenn $(|a_n|)_{n\in\mathbb{N}}$ eine Nullfolge ist.

Definition 2.6. Sei $(a_j)_{j\in\mathbb{N}}$ eine Folge in K. Wir definieren die n-te Teilsumme oder Partialsumme als

$$s_n := \sum_{j=0}^{n} a_j.$$

Die Folge $(s_n)_{n\in\mathbb{N}}$ nennt man dann (unendliche) Reihe mit den Summanden a_j. Man schreibt eine Reihe in der Form

$$\sum_{j=0}^{\infty} a_j.$$

Falls die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}}$ konvergiert, so sagen wir die Reihe konvergiert und man definiert

$$\sum_{j=0}^{\infty} a_j := \lim_{n\to\infty} s_n.$$

Die Reihe divergiert, wenn die Folge der Partialsummen divergiert. Bei Folgen in $\mathbb{R}$ definiert man analog die uneigentliche Konvergenz.

Wir erlauben auch, dass Reihen nicht bei 0, sondern bei einer beliebigen anderen ganzen Zahl k_0 beginnen und schreiben dann eben $\sum_{j=k_0}^{\infty} a_j$.

!ACHTUNG!. $\sum_{j=1}^{\infty} a_j$ hat zwei Bedeutungen:
1.) die Folge der Partialsummen. Dann ist $\sum_{j=0}^{\infty} a_j \in \mathrm{Abb}(\mathbb{N}, K)$
2.) der Limes dieser Folge (falls er existiert). Dann ist $\sum_{j=0}^{\infty} a_j \in K$
Welche Bedeutung gemeint ist, wird immer aus dem Kontext heraus deutlich. Wenn man sagt „die Reihe $\sum_{j=1}^{\infty} a_j$ konvergiert oder divergiert in K“, dann ist immer die Folge der Partialsummen gemeint. Wenn man $5 = \sum_{j=1}^{\infty} a_j$ schreibt, dann ist der Limes der Folge der Partialsummen gemeint (und es ist dann implizit klar, dass wir fordern, dass dieser dann existieren muss). Um diesen Unterschied deutlicher zu machen, schreiben wir in den nächsten Zeilen aus didaktischen Gründen
1.) die Folge der Partialsummen immer in dieser Farbe
2.) den Limes dieser Folge immer in dieser Farbe

$$\sum_{j=1}^{\infty} a_j \text{ konvergent}$$

$$\iff \exists s \in K : \forall \epsilon \in \mathbb{R}_{>0} : \exists n_0 \in \mathbb{N} : \forall n \in \mathbb{N}_{\geq n_0} : \left|(\sum_{j=1}^{n} a_j) - s\right| \leq \epsilon$$

$$\iff \forall \epsilon \in \mathbb{R}_{>0} : \exists n_0 \in \mathbb{N} : \forall n \in \mathbb{N}_{\geq n_0} : \forall m \in \mathbb{N}_{\geq n} : \left|\sum_{j=n+1}^{m} a_j\right| \leq \epsilon \text{ (Cauchy-Kriterium)}$$

[8]

Jede Folge kann realisiert werden als Folge von Partialsummen. Denn sei $(s_n)_{n\in\mathbb{N}}$ gegeben. Definiere dann $a_0 := s_0$ und für $n \geq 1$: $a_n = s_n - s_{n-1}$. Dann ist s_n die n-te Partialsumme von $(a_n)_{n\in\mathbb{N}}$.

BEISPIELE 2.7.

(a) Konvergente geometrische Reihe: $\sum_{r=0}^{\infty} z^r = \frac{1}{1-z}$ für $|z| < 1$.

Denn wir rechnen:

$$(\underbrace{z^0 + z^1 + z^2 + \cdots + z^n}_{s_n=})(1 - z) = 1 - z^{n+1},$$

Nun ist $(z^{n+1})_{n\in\mathbb{N}}$ eine Nullfolge, da $(|z|^{n+1})_{n\in\mathbb{N}}$ ebenfalls eine Nullfolge ist.

(b) Divergente geometrische Reihe: $\sum_{r=0}^{\infty} z^r$ divergiert für $|z| \geq 1$: siehe unten (Prop. 2.8) und beachte, dass $(z^r)_{r\in\mathbb{N}}$ keine Nullfolge ist.

(c) Die harmonische Reihe $\sum_{r=1}^{\infty} \frac{1}{r}$ divergiert:

Dieser Teil ist noch nicht gesetzt.

(d) $\sum_{r=1}^{\infty} \frac{1}{r^2}$ konvergiert:

[8] Hier gilt die Konvention $\sum_{j=n+1}^{n} a_j = 0$

Dieser Teil ist noch nicht gesetzt.

PROPOSITION 2.8. Wenn $\sum_{j=0}^{\infty} a_j$ konvergiert, dann ist $(a_j)_{j\in\mathbb{N}}$ eine Nullfolge.

Beweis. Sei $s = \sum_{j=0}^{\infty} a_j$. Dann gilt

$$a_n = \sum_{j=0}^{n} a_j - \sum_{j=0}^{n-1} a_j \to s - s = 0 \text{ für } n \to \infty. \qquad \square$$

Die Umkehrung dieser Aussage ist falsch: Die Reihe $\sum_{r=1}^{\infty} \frac{1}{r}$ divergiert, aber $(\frac{1}{r})_{r\in\mathbb{N}\setminus\{0\}}$ ist Nullfolge.

EIGENSCHAFTEN 2.9.

(a) Wenn $\sum_{j=0}^{\infty} a_j$ und $\sum_{j=0}^{\infty} b_j$ konvergieren, dann gilt

$$\sum_{j=0}^{\infty}(a_j \pm b_j) = \sum_{j=0}^{\infty} a_j \pm \sum_{j=0}^{\infty} b_j.$$

(b) In einer konvergenten Reihe dürfen Klammern gesetzt werden (Übergang zu einer Teilfolge der Partialsummen). Man darf im allgemeinen Klammern nicht weglassen: $1-1+1-1+1\cdots$ divergiert, aber $(1-1)+(1-1)+(1-1)\cdots$ konvergiert gegen 0 und $1+(-1+1)+(-1+1)\cdots$ konvergiert gegen 1.

(c) Es gelte für alle $r \in \mathbb{N}$: $a_r \in \mathbb{R}_{\geq 0}$. Dann ist $(s_n)_{n\in\mathbb{N}}$ monoton wachsend. Also ist nach Proposition 1.15 $\sum_{r=0}^{\infty} a_r$ genau dann konvergent in K, wenn $(s_n)_{n\in\mathbb{N}}$ beschränkt ist.

2.3. Konvergenzkriterien.

Einige vereinfachende Notationen

Um in Zukunft über Folgen und Reihen effizienter sprechen zu können, führen wir noch einige Notationen ein.

DEFINITION 2.10. Sei $A(\,\cdot\,)$ eine auf $\mathbb{N}$ definierte Aussageform.

$$\begin{aligned} &\text{Für fast alle } j \in \mathbb{N} \text{ gilt } A(j) \\ :\Longleftrightarrow\ &\text{Es gibt ein } j_0 \in \mathbb{N}, \text{ so dass für alle } j \in \mathbb{N}_{\geq j_0} :\ A(j) \\ \Longleftrightarrow\ &\text{Die Menge } \{j \in \mathbb{N} \mid \neg A(j)\} \text{ ist endlich} \end{aligned}$$

Beispiel:

$$\begin{aligned} &\lim_{j\to\infty} a_j = a \\ \Longleftrightarrow\ &\text{Für alle } \epsilon \in \mathbb{R}_{>0} \text{ gilt für fast alle } j \in \mathbb{N}\text{: } |a_j - a| \leq \epsilon \\ \not\Longleftrightarrow\ &\text{Für fast alle } j \in \mathbb{N} \text{ gilt für alle } \epsilon \in \mathbb{R}_{>0}\text{: } |a_j - a| \leq \epsilon \end{aligned}$$

Die Bedingung der letzten Zeile ist genau dann erfüllt, wenn die Folge $(a_j)_{j\in\mathbb{N}}$ nach endlich vielen potentiellen Ausnahmen konstant gleich a ist. Solche Folgen konvergieren zwar gegen a, aber es gibt viel mehr Folgen, die gegen a konvergieren.

Notation:
Da wir immer wieder Ausdrücke der Art

$$\forall j \in \mathbb{N}_{\geq j_0} :\ A(j)$$

haben, schreiben wir hierfür kurz

$$\forall j \in \{j_0, j_0 + 1, \ldots\} :\ A(j).$$

PROPOSITION 2.11 (Majoranten-Kriterium)**.** Sei $(a_r)_{r\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ oder $\mathbb{C}$ und $(b_r)_{r\in\mathbb{N}}$ eine Folge in $\mathbb{R}_{\geq 0}$, so dass für fast alle $r \in \mathbb{N}$: $|a_r| \leq b_r$. Konvergiert $\sum_{r=0}^{\infty} b_r$, dann auch $\sum_{r=0}^{\infty} a_r$.

Man nennt in diesem Fall $\sum_{r=0}^{\infty} b_r$ eine konvergente Majorante.

Beweis. Für alle $\epsilon \in \mathbb{R}_{>0}$ existiert $n_0 \in \mathbb{N}$ so dass für alle $n \in \{n_0, n_0+1, \ldots\}$:

$$|a_{n+1} + \cdots + a_m| \leq |a_{n+1}| + \cdots + |a_m| \leq b_{n+1} + \cdots + b_m \leq \epsilon. \qquad \square$$

Das Majoranten-Kriterium liefert sofort: Divergiert $\sum_{r=0}^{\infty} a_r$, so divergiert auch $\sum_{r=0}^{\infty} b_r$. In diesem Fall gilt dann: $\sum_{r=0}^{\infty} b_r = \infty$ (uneigentliche Konvergenz), da die Folge der Partialsummen $(\sum_{r=0}^{n} b_r)_{n\in\mathbb{N}}$ monoton wachsend ist.

BEISPIELE 2.12. (a) $\sum_{r=1}^{\infty} \frac{1}{r^k}$ konvergiert für $k \geq 2$, denn $\sum_{r=1}^{\infty} \frac{1}{r^2}$ ist eine konvergente Majorante.

(b) $\sum_{r=1}^{\infty} \frac{1}{\sqrt{r}}$ divergiert, denn es ist eine Majorante von $\sum_{r=1}^{\infty} \frac{1}{r}$.

(c) $\sum_{r=1}^{\infty} \frac{1}{r!}$ konvergiert, weil $\sum_{r=2}^{\infty} \frac{1}{(r-1)^2} = \sum_{k=1}^{\infty} \frac{1}{k^2}$ eine konvergente Majorante von $\sum_{r=2}^{\infty} \frac{1}{r!}$ ist.

LEMMA 2.13. Sei $(b_j)_{j\in\mathbb{N}}$ eine Folge in $\mathbb{R}$.

(i) $\liminf_{j\to\infty} b_j > 1$ genau dann, wenn es ein $q \in \mathbb{R}$, $q > 1$, gibt, so dass für fast alle $j \in \mathbb{N}$: $b_j \geq q$.

(ii) $\limsup_{j\to\infty} b_j < 1$ genau dann, wenn es ein $q \in \mathbb{R}$, $q < 1$ gibt, so dass für fast alle $j \in \mathbb{N}$: $b_j \leq q$.

Beweis. Zu (i) $\Longrightarrow$:

$$c_n := \inf\{b_j \mid j \geq n\}.$$

c_n monoton wachsend, $c := \lim_{n\to\infty} c_n = \liminf_{j\to\infty} b_j > 1$.

$$q := \frac{c+1}{2}.$$

Somit gilt für fast alle $n \in \mathbb{N}$: $c_n > q$. Also für fast alle $j \in \mathbb{N}$: $b_j > q$.

Fr 14.12. Hieraus folgt die Aussage „$\Longrightarrow$“ für den Limes inferior.

Zu (i) $\Longleftarrow$: Sei wieder

$$c_n := \inf\{b_j \mid j \geq n\}.$$

Gegeben sei $q \in \mathbb{R}_{>1}$, so dass für fast alle $j \in \mathbb{N}$ gilt: $b_j \geq q$. In anderen Worten: es gibt ein $j_0 \in \mathbb{N}$, so dass für $j \in \mathbb{N}_{\geq j_0}$: $b_j \geq q$. Somit

$$q \leq c_{j_0} \leq \liminf_{j\to\infty} b_j.$$

Zu (ii): Die Aussagen für den Limes superior zeigt man ganz analog. □

Die Ausdrücke $\lim_{n\to\infty} a_n$, $\liminf_{n\to\infty} a_n$ und $\limsup_{n\to\infty} a_n$ ergeben immer noch Sinn, wenn a_n gar nicht mehr für alle $n \in \mathbb{N}$, sondern nur noch für fast alle $n \in \mathbb{N}$ definiert ist.

PROPOSITION 2.14 (Quotienten-Kriterium)**.** Sei $\sum_{j=0}^{\infty} a_j$ eine Reihe in $\mathbb{R}$ oder in $\mathbb{C}$, so dass für fast alle $j \in \mathbb{N}$: $a_j \neq 0$.

(a) Gilt $\limsup_{j\to\infty} \left|\frac{a_{j+1}}{a_j}\right| < 1$ dann konvergiert $\sum_{j=0}^{\infty} a_j$.
(b) Gilt $\liminf_{j\to\infty} \left|\frac{a_{j+1}}{a_j}\right| > 1$, dann divergiert $\sum_{j=0}^{\infty} a_j$.

Beweis. Zu (a): Lemma 2.13 liefert uns ein $j_0 \in \mathbb{N}$ und $q < 1$, so dass gilt:

$$\forall j \in \mathbb{N}_{\geq j_0} : a_j \neq 0 \text{ und } \forall j \in \mathbb{N}_{\geq j_0} : \left|\frac{a_{j+1}}{a_j}\right| \leq q.$$

Dann zeigt man durch Induktion für $n \in \mathbb{N}_{\geq j_0}$: $|a_n| \leq q^{n-j_0}|a_{j_0}|$. Somit ist $\sum_{n=0}^{\infty} q^{n-j_0}|a_{j_0}|$ eine konvergente Majorante von $\sum_{j=0}^{\infty} a_j$ und deswegen ist $\sum_{j=0}^{\infty} a_j$ konvergent.

Zu (b): Man geht ähnlich vor wie oben und erhält $|a_n| \geq q^{n-j_0}|a_{j_0}|$ für $q > 1$. Daraus folgt, dass $(|a_n|)_{n\in\mathbb{N}}$ keine Nullfolge ist, damit auch nicht $(a_n)_{n\in\mathbb{N}}$ und somit divergiert die Reihe $\sum_{n=0}^{\infty} a_n$. □

Beispiel 2.15. Sei $t = \frac{p}{q} \in \mathbb{Q}$, $p \in \mathbb{Z}$, $q \in \mathbb{N} \setminus \{0\}$. Definiere für $r \in \mathbb{R}_{\geq 0}$:

$$r^t := (\sqrt[q]{r})^p$$

Sei $x \in \mathbb{C}$. [9] Betrachte die Reihe $\sum_{n=0}^{\infty} a_n$ für $a_n = n^t x^n$. Für $n \geq 1$:

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

$$1 \leq \left(\frac{n+1}{n}\right)^t = \left(\sqrt[q]{1+\frac{1}{n}}\right)^p \overset{(*)}{\leq} \left(1+\frac{1}{n}\right)^p \to 1$$

für $n \to \infty$, wobei wir in $(*)$ genutzt haben, dass für $s \geq 1$ gilt: $1 \leq \sqrt[q]{s} \leq s$ und somit $\sqrt[q]{s^p} \leq s^p$. Wir erhalten dann

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = |x| \lim_{n\to\infty}\left(\frac{n+1}{n}\right)^t = |x|$$

Also ist $\sum_{n=0}^{\infty} n^t x^n$ konvergent für $|x| < 1$; und $\sum_{n=0}^{\infty} n^t x^n$ ist divergent für $|x| > 1$.

[9] Wir beschränken uns hier auf rationale t, da wir r^t noch nicht für beliebige reelle t definiert haben.

PROPOSITION 2.16 (Wurzel-Kriterium)**.** Sei $(a_r)_{r\in\mathbb{N}}$ eine Folge. [10]

(a) Wenn $\limsup_{r\to\infty} \sqrt[r]{|a_r|} < 1$, dann konvergiert $\sum_{r=0}^{\infty} a_r$.

(b) Gilt für unendlich viele $r \in \mathbb{N}$: $\sqrt[r]{|a_r|} \geq 1$, dann ist $\sum_{r=0}^{\infty} a_r$ divergent.

(c) Wenn $\limsup_{r\to\infty} \sqrt[r]{|a_r|} > 1$, dann ist $\sum_{r=0}^{\infty} a_r$ divergent.

Beweis. Zu (a): Unter Benutzung des Lemmas 2.13 gilt für ein geeignetes q und fast alle $r \in \mathbb{N}$

$$\sqrt[r]{|a_r|} \leq q < 1.$$

Also $|a_r| \leq q^r$. Also ist $\sum_{r=0}^{\infty} q^r$ eine konvergente Majorante von $\sum_{r=0}^{\infty} a_r$.

Zu (b): $\sqrt[r]{|a_r|} \geq 1 \iff |a_r| \geq 1$. Gilt dies für unendlich viele r, dann ist $(a_r)_{r\in\mathbb{N}}$ keine Nullfolge, also kann $\sum_{r=0}^{\infty} a_r$ nicht konvergent sein.

Zu (c): Im Fall $\limsup_{r\to\infty} \sqrt[r]{|a_r|} > 1$ gibt es eine Teilfolge $(a_{f(n)})_{n\in\mathbb{N}}$ so dass $\lim_{n\to\infty} \sqrt[n]{|a_{f(n)}|} = \limsup_{r\to\infty} \sqrt[r]{|a_r|} > 1$. Es gilt dann $\sqrt[n]{|a_{f(n)}|} > 1$ für fast alle $n \in \mathbb{N}$. Also gibt es unendlich viele $r \in \mathbb{N}$ mit $|a_r| > 1$. □

BEISPIELE 2.17. Eine Reihe der Form $\sum_{r=0}^{\infty} a_r x^r$ nennt man eine Potenzreihe *in* x.

(a) Sei $x \in \mathbb{R}_{\geq 0}$ und $a_r \geq 0$ [11]

$$\sum_{r=0}^{\infty} a_r x^r$$

$$\limsup_{r\to\infty} \sqrt[r]{a_r x^r} = x \limsup_{r\to\infty} \sqrt[r]{a_r}, \text{ falls } x \neq 0$$

[10] In der folgenden Formel ist $\sqrt[r]{|a_r|} < 1$ für $r = 0$ nicht definiert. Dies stört uns nicht, da die Folge nur für fast alle $r \in \mathbb{N}$ definiert zu sein braucht, damit wir den Limes superior definieren können.

[11] Wdh: Für $x \in \mathbb{C}$ definieren wir $x \geq 0$ als: $x \in \mathbb{R}$ und $x \geq 0$. Es ist hier im Text also implizit die Information enthalten, dass $x \in \mathbb{R}$ und $a_r \in \mathbb{R}$.

$$\limsup_{r\to\infty} \sqrt[r]{a_r x^r} = \infty \implies \forall x \in \mathbb{R}_{>0} : \sum_{r=0}^{\infty} a_r x^r \text{ divergiert}$$

$$0 < \limsup_{r\to\infty} \sqrt[r]{a_r x^r} < \infty \implies \forall x \in \left(0, \frac{1}{\limsup_{r\to\infty} \sqrt[r]{a_r}}\right) : \sum_{r=0}^{\infty} a_r x^r \text{ konvergiert}$$

$$\forall x \in \left(\frac{1}{\limsup_{r\to\infty} \sqrt[r]{a_r}}, \infty\right) : \sum_{r=0}^{\infty} a_r x^r \text{ divergiert}$$

$$\limsup_{r\to\infty} \sqrt[r]{a_r x^r} = 0 \implies \forall x \in \mathbb{R}_{\geq 0} : \sum_{r=0}^{\infty} a_r x^r \text{ konvergiert}$$

Wir definieren

$$\rho := \begin{cases} \dfrac{1}{\limsup_{r\to\infty} \sqrt[r]{a_r}} & \text{falls } \limsup_{r\to\infty} \sqrt[r]{a_r} \in \mathbb{R}_{>0} \\ 0 & \text{falls } \limsup_{r\to\infty} \sqrt[r]{a_r} = \infty \\ \infty & \text{falls } \limsup_{r\to\infty} \sqrt[r]{a_r} = 0 \end{cases}$$

$\sum_{r=0}^{\infty} a_r x^r$ konvergiert für $x < \rho$ und divergiert für $x > \rho$.

ρ heißt Konvergenzradius der Potenzreihe.

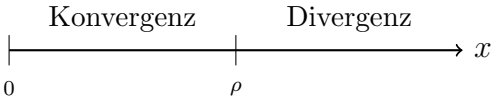


Abbildung 5:

(b) Sei nun $x \in \mathbb{C}$ und $a_i \in \mathbb{C}$. Wir definieren dann

$$\rho := \begin{cases} \dfrac{1}{\limsup\limits_{r\to\infty} \sqrt[r]{|a_r|}} & \text{falls } \limsup\limits_{r\to\infty} \sqrt[r]{|a_r|} \in \mathbb{R}_{>0} \\ 0 & \text{falls } \limsup\limits_{r\to\infty} \sqrt[r]{|a_r|} = \infty \\ \infty & \text{falls } \limsup\limits_{r\to\infty} \sqrt[r]{|a_r|} = 0 \end{cases}$$

SATZ (Satz von Cauchy und Hadamard)**.** Unter obigen Voraussetzungen konvergiert $\sum_{r=0}^{\infty} a_r x^r$ im Fall $|x| < \rho$ und divergiert im Fall $|x| > \rho$.

Die Begründung kann man völlig analog zu Beispiel (a) führen, wenn man $|x|$ an Stelle von x und $|a_r|$ an Stelle von a_r schreibt. Alternativ kann die **Konvergenz** auch aus (a) mit dem Majoranten-Kriterium gezeigt werden.

(c) Für $|x| = \rho$ kann keine Aussage getroffen werden. Die Potenzreihen $\sum_{r=0}^{\infty} r x^r$ und $\sum_{r=1}^{\infty} \frac{1}{r^2} x^r$ besitzen beide Konvergenzradius 1, denn $\lim\limits_{r\to\infty} \sqrt[r]{r} = 1$ und $\lim\limits_{r\to\infty} \sqrt[r]{r^{-2}} = (\lim\limits_{r\to\infty} \sqrt[r]{r})^{-2} = 1$. (Details hier[12]) Nun ist die Reihe $\sum_{r=0}^{\infty} r x^r$ für $x \in \mathbb{C}$, $|x| = 1$ offensichtlich divergent. Wir haben gesehen, dass für $x = 1$ die Reihe $\sum_{r=1}^{\infty} \frac{1}{r^2} x^r$ konvergiert (Beispiel 2.7 (d)), und auf Grund des Majoranten-Kriteriums konvergiert diese Reihe dann für alle $x \in \mathbb{C}$ mit $|x| = 1$.

LEMMA 2.18. Sei $a_r \in \mathbb{C} \setminus \{0\}$. Dann gilt

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} \leq \limsup_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}.$$

[12]Wir wollen zeigen, dass $\lim\limits_{r\to\infty} \sqrt[r]{r} = 1$. Hierzu ist zu zeigen, dass $x_r := \sqrt[r]{r} - 1$ eine Nullfolge ist. Man rechnet nun mit der binomischen Formel nach, dass $\underbrace{(x_r + 1)^r}_{r=} \geq \frac{r(r-1)}{2} x_r^2$. Also $0 \leq x_r^2 \leq 2/(r-1)$.

Beweis. Sei $\alpha := \limsup_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$. Fixiere ein $\beta > \alpha$. Es gilt also[13] für fast alle $n \in \mathbb{N}$:

$$\frac{|a_{n+1}|}{|a_n|} \leq \beta. \tag{2.19}$$

Bestimme also ein $n_0 \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}_{\geq n_0}$ die Ungleichung (2.19) gilt.

$$|a_n| \leq \beta^{n-n_0}|a_{n_0}|$$

$$\sqrt[n]{|a_n|} \leq \underbrace{\sqrt[n]{\beta^n}}_{\beta=} \underbrace{\sqrt[n]{\beta^{-n_0}|a_{n_0}|}}_{\to 1}$$

Wr haben genutzt, dass für jedes $s \in \mathbb{R}_{>0}$ die Aussage $\lim_{n\to\infty} \sqrt[n]{s} = 1$ gilt. [14] Daraus folgt also

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} \leq \beta.$$

Da dies für alle $\beta > \alpha$ gilt, folgt die Aussage. □

Bemerkung 2.20. Die Konvergenzaussage im Wurzel-Kriterium (Prop. 2.16 (a)) ist also stärker als die Konvergenzaussage im Quotienten-Kriterium (Prop. 2.3 (a)), denn es zeigt dieselbe Aussage mit einer schwächeren Voraussetzung.

Mi 19.12.

[13] eine Möglichkeit, dies zu sehen, ist Lemma 2.13 (ii) für $b_j = \frac{|a_{j+1}|}{\alpha|a_j|}$ anzuwenden.

[14] Beweis für $s > 1$: $s = (\sqrt[n]{s})^n = (\sqrt[n]{s} - 1 + 1)^n = \sum_{j=0}^{n} \binom{n}{j} (\sqrt[n]{s} - 1)^j 1^{n-j} \geq n(\sqrt[n]{s} - 1)$. Somit $1 \leq \sqrt[n]{s} \leq 1 + \frac{1}{n}$. Im Fall $s < 1$ wissen wir deswegen $\sqrt[n]{1/s} \to 1$ und somit konvergiert $\sqrt[n]{s} = \frac{1}{\sqrt[n]{1/s}}$ auch gegen 1.

2.4. Absolute Konvergenz.

DEFINITION 2.21 (Absolute Konvergenz). Seien $a_r \in \mathbb{C}$. Wir sagen die Reihe $\sum_{r=0}^{\infty} a_r$ konvergiert absolut, falls $\sum_{r=0}^{\infty} |a_r|$ konvergiert.

Das Majoranten-Kriterium für $b_r := |a_r|$ impliziert also: Jede absolut konvergente Reihe ist konvergent.

Majoranten-Kriterium, Quotienten-Kriterium und Wurzel-Kriterium liefern nicht nur die Konvergenz einer Reihe, sondern sogar ihre absolute Konvergenz, siehe unten.[15] Insbesondere ist eine Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ absolut konvergent, wenn $|x|$ kleiner als der Konvergenzradius ist.

PROPOSITION 2.22 (Majoranten-Kriterium mit absoluter Konvergenz)**.** Sei $(a_r)_{r\in\mathbb{N}}$ eine Folge[16] und $(b_r)_{r\in\mathbb{N}}$ eine Folge in $\mathbb{R}_{\geq 0}$, so dass für fast alle $r \in \mathbb{N}$: $|a_r| \leq b_r$. Konvergiert $\sum_{r=0}^{\infty} b_r$, dann ist $\sum_{r=0}^{\infty} a_r$ **absolut konvergent**.

PROPOSITION 2.23 (Quotienten-Kriterium mit absoluter Konvergenz)**.** Sei $\sum_{j=0}^{\infty} a_j$ eine Reihe, so dass für fast alle $j \in \mathbb{N}$: $a_j \neq 0$.

(a) Gilt $\limsup_{j\to\infty} \left|\frac{a_{j+1}}{a_j}\right| < 1$, dann konvergiert $\sum_{j=0}^{\infty} a_j$ **absolut**.
(b) wie zuvor.

PROPOSITION 2.24 (Wurzel-Kriterium mit absoluter Konvergenz)**.** Sei $(a_r)_{r\in\mathbb{N}}$ eine Folge.

---

[15] Die Beweise sind wortwörtlich gleich

[16] Wenn wir nichts anderes hinschreiben, ist unter einer Folge immer eine Folge in $\mathbb{C}$ gemeint, und eine Folge in $\mathbb{R}$ ist auch eine Folge in $\mathbb{C}$.

(a) Wenn $\limsup_{r\to\infty} \sqrt[r]{|a_r|} < 1$, dann konvergiert $\sum_{r=0}^{\infty} a_r$ **absolut**.

(b) und (c) wie zuvor.

2.5. Alternierende Reihen.

DEFINITION 2.25. Eine alternierende Reihe ist eine Reihe in der Form

$$\sum_{n=0}^{\infty}(-1)^n a_n$$

oder[17]

$$\sum_{n=0}^{\infty}(-1)^{n+1} a_n,$$

wobei $\forall n \in \mathbb{N} : \ a_n \in \mathbb{R}_{\geq 0}$.

PROPOSITION 2.26 (Regel von Leibniz)**.** Sei $(a_n)_{n\in\mathbb{N}}$ eine monoton fallende Nullfolge, $a_n \geq 0$. Dann konvergiert die Reihe

$$\sum_{n=0}^{\infty}(-1)^n a_n.$$

Beweis. Gegeben seien natürliche Zahlen m, n, n_0 mit $m \geq n \geq n_0$.

Idee (noch nicht präzise):
Um das Cauchy-Kriterium anzuwenden, wollen wir

$$(-1)^{n+1}a_{n+1} + (-1)^{n+2}a_{n+2} + \cdots + (-1)^m a_m$$

kontrollieren.

[17]Wir betrachten ab jetzt immer nur den ersten Fall, um die Notation einfach zu halten

Präzise Ausführung der Idee:

$$
\begin{aligned}
&a_{n+1} + (-1)a_{n+2} + \cdots + (-1)^{m-n+1} a_m \\
&= \begin{cases}
\underbrace{(a_{n+1} - a_{n+2})}_{\geq 0} + \cdots + \underbrace{(a_{m-1} - a_m)}_{\geq 0} \geq 0 & \text{falls } m-n \text{ gerade} \\
\underbrace{(a_{n+1} - a_{n+2})}_{\geq 0} + \cdots + \underbrace{a_m}_{\geq 0} \geq 0 & \text{falls } m-n \text{ ungerade}
\end{cases}
\end{aligned}
$$

$$
\begin{aligned}
&a_{n+1} + (-1)a_{n+2} + \cdots + (-1)^{m-n+1} a_m \\
&= \begin{cases}
a_{n+1} + \overbrace{((-1)a_{n+2} + a_{n+3})}^{\leq 0} + \cdots + \overbrace{(-1)a_m}^{\leq 0} & \\
\qquad \leq a_{n+1} & \text{falls } m-n \text{ gerade} \\
a_{n+1} + \overbrace{((-1)a_{n+2}) + a_{n+3})}^{\leq 0} + \cdots + \overbrace{((-1)a_{m-1} + a_m)}^{\leq 0} & \\
\qquad \leq a_{n+1} & \text{falls } m-n \text{ ungerade}
\end{cases}
\end{aligned}
$$

Also

$$
|\sum_{k=n+1}^{m} (-1)^k a_k| = |(-1)^{n+1} a_{n+1} + (-1)^{n+2} a_{n+2} + \cdots + (-1)^m a_m| \leq |a_{n+1}|.
$$

Da (a_n) eine Nullfolge ist, haben wir gezeigt:

$$
\forall \epsilon \in \mathbb{R}_{>0} : \exists n_0 \in \mathbb{N} : \forall n \in \mathbb{N}_{\geq n_0} : \forall m \in \mathbb{N}_{\geq n} : |\sum_{k=n+1}^{m} (-1)^k a_k| \leq \epsilon.
$$

□

Beispiele 2.27.

(a) Die harmonische alternierende Reihe

$$\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}$$

konvergiert.

(b) Die Reihe $\sum_{n=1}^{\infty}(-1)^n a_n$ mit

$$a_n := \begin{cases} \frac{1}{n} & \text{für } n \text{ gerade} \\ \frac{1}{n^2} & \text{für } n \text{ ungerade} \end{cases}$$

divergiert. Denn sei

$$\alpha := \sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}.$$

(Anschaulich ist α die unendliche Summe der a_n über alle ungeraden n.)

Dann gilt

$$\sum_{n=1}^{2r} a_n \geq (\sum_{k=1}^{r}\frac{1}{2r}) - \alpha$$

und dies konvergiert uneigentlich gegen ∞ für $r \to \infty$, da die harmonische Reihe konvergiert.

2.6. Umordnung von Reihen.

DEFINITION 2.28. Sei $\sum_{n=0}^{\infty} a_n$ eine Reihe. Ein Umordnung dieser Reihe ist eine Reihe der Form $\sum_{n=0}^{\infty} a_{f(n)}$, wobei $f : \mathbb{N} \longrightarrow \mathbb{N}$ eine Bijektion ist.

SATZ 2.29 (Umordnungssatz)**.** Ist $\sum_{n=0}^{\infty} a_n$ eine absolut konvergente Reihe in $\mathbb{R}$ oder $\mathbb{C}$, so ist jede Umordnung $\sum_{n=0}^{\infty} a_{f(n)}$ dieser Reihe ebenfalls absolut konvergent und es gilt

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} a_{f(n)}.$$

Beweis später.

Satz 2.30 (Riemannsche Umordnungssatz)**.** Ist $\sum_{n=0}^{\infty} a_n$ eine konvergente Reihe in $\mathbb{R}$, die nicht absolut konvergiert.

(1) Sei $\alpha \in \overline{\mathbb{R}}$ gegeben. Dann gibt es eine Umordnung dieser Reihe, die gegen α eigentlich oder uneigentlich konvergiert.
(2) Es gibt eine Umordnung, die beschränkt und divergent ist.

Beispiel 2.31. Die Folge $\sum_{n=1}^{\infty}(-1)^n\frac{1}{n}$ ist konvergent, aber nicht absolut konvergent. Zu jedem $\alpha \in \overline{\mathbb{R}}$ gibt es eine Umordnung von $\sum_{n=1}^{\infty}(-1)^n\frac{1}{n}$, die gegen α konvergiert.

Beweis von Satz 2.29. Sei $\sum_{n=0}^{\infty} a_n$ absolut konvergent. Dies bedeutet nach dem Cauchy-Kriterium: Zu jedem $\epsilon \in \mathbb{R}_{>0}$ gibt es ein $n_0 \in \mathbb{N}$, so dass für alle $n, \tilde{n} \in \mathbb{N}$ mit $n_0 \leq n \leq \tilde{n}$ gilt[18] : $\sum_{j=n+1}^{\tilde{n}} |a_j| \leq \epsilon$.

Wir fixieren nun ϵ und ein passendes n_0. Sei nun $\sum_{n=0}^{\infty} a_{f(n)}$ eine Umordnung dieser Reihe. Insbesondere sei $f : \mathbb{N} \to \mathbb{N}$ als bijektiv vorausgesetzt. Wir setzen nun

$$m_0 := \max f^{\#}(\{0, 1, \ldots, n_0\}) := \max\{f^{-1}(0), f^{-1}(1), \ldots, f^{-1}(n_0)\}.$$

Hier steht eine Zeichnung, die noch nicht gesetzt ist.

Dies bedeutet, für alle $\ell > m_0$ gilt $f(\ell) > n_0$. Somit gilt für $m, \tilde{m} \in \mathbb{N}$ mit $m_0 \leq m < \tilde{m}$:

$$\sum_{\ell=m+1}^{\tilde{m}} |a_{f(\ell)}| \leq \sum_{j=n_0+1}^{N} |a_j| \leq \epsilon$$

[18] Ja, die Betragsstriche sind hier um den richtigen Ausdruck, denn es handelt sich um absolute Konvergenz!

wobei $N \in \mathbb{N}$ eine "genügend große" natürliche Zahl ist, zum Beispiel

$$N := \max f_{\#}(\{m+1, m+2, \ldots, \tilde{m}\}) = \max\{f(m+1), f(m+2), \ldots, f(\tilde{m})\} > n_0 \tag{2.32}$$

oder jede Zahl die mindestens so groß wie dieses Maximum ist. Dann ist also das Cauchy-Kriterium für die Reihe $\sum_{n=0}^{\infty} |a_{f(n)}|$ erfüllt, in anderen Worten $\sum_{n=0}^{\infty} a_{f(n)}$ ist absolut konvergent.

Wir wissen nun also, dass die Reihen (absolut) konvergieren, und wir studieren nun deren Grenzwerte $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} a_{f(n)}$. Wir wählen zu $\epsilon \in \mathbb{R}_{>0}$ passende Zahlen n_0 und m_0 wie oben. Außerdem definieren wir

$$N_0 := \max\{f(0), f(1), \ldots, f(m_0)\} = \max f_{\#}(\{0, 1, \ldots, m_0\}) \geq n_0.$$

Für alle $n, \tilde{n} \in \mathbb{N}$ mit $n_0 \leq n \leq \tilde{n}$ gilt

$$|\sum_{j=0}^{\tilde{n}} a_j - \sum_{j=0}^{n} a_j| = |\sum_{j=n+1}^{\tilde{n}} a_j| \leq \sum_{j=n+1}^{\tilde{n}} |a_j| \leq \epsilon \tag{2.33}$$

und dies impliziert im Limes $\tilde{n} \to \infty$ für $n := N_0 (\geq n_0)$

$$|\sum_{j=0}^{\infty} a_j - \sum_{j=0}^{N_0} a_j| \leq \epsilon.$$

Wenn wir N wie in (2.32) definieren ergibt sich analog:

$$|\sum_{\ell=0}^{\tilde{m}} a_{f(\ell)} - \sum_{\ell=0}^{m_0} a_{f(\ell)}| = |\sum_{\ell=m_0+1}^{\tilde{m}} a_{f(\ell)}| \leq \sum_{j=n_0+1}^{N} |a_j| \leq \epsilon$$

und somit im Limes $\tilde{m} \to \infty$

$$|\sum_{\ell=0}^{\infty} a_{f(\ell)} - \sum_{\ell=0}^{m_0} a_{f(\ell)}| \leq \epsilon.$$

Außerdem

$$\sum_{j=0}^{N} a_j - \sum_{\ell=0}^{m_0} a_{f(\ell)} \;=\; \sum_{j\in I} a_j$$

wobei $I := \{0,1,\dots,N\} \smallsetminus f_\#(\{0,1,\dots,m_0\})$.[19] Nach Konstruktion von m_0 sind I und $\{0,1,\dots,n_0\}$ disjunkt[20]. Es folgt mit (2.33)

$$|\sum_{j\in I} a_j| \le \sum_{j\in I} |a_j| \le \sum_{j=n_0+1}^{N} |a_j| \le \epsilon.$$

Dies ergibt

$$\left|\sum_{j=0}^{\infty} a_j - \sum_{\ell=0}^{\infty} a_{f(\ell)}\right| \le \underbrace{\left|\sum_{j=0}^{\infty} a_j - \sum_{j=0}^{N} a_j\right|}_{\le\epsilon} + \underbrace{\left|\sum_{j=0}^{N} a_j - \sum_{\ell=0}^{m_0} a_{f(\ell)}\right|}_{\le\epsilon} + \underbrace{\left|\sum_{\ell=0}^{m_0} a_{f(\ell)} - \sum_{\ell=0}^{\infty} a_{f(\ell)}\right|}_{\le\epsilon} \le 3\epsilon$$

Fr 21.12. Da diese Aussage für alle $\epsilon \in \mathbb{R}_{>0}$ gilt, sind die Grenzwerte gleich. □

Beweis von Satz 2.30. Wir beweisen (1) im Fall $\alpha \in \mathbb{R}$. Die anderen Fälle und Ausage (2) beweist man ähnlich, wir geben hierfür aus Zeitgründen aber keine Details.

Zu (1) für $\alpha \in \mathbb{R}$: Sei $\sum_{n=0}^{\infty} a_n$ eine konvergente Reihe, die nicht absolut konvergiert. Wir definieren

$$I_+ := \{n \in \mathbb{N} \mid a_n \ge 0\} \qquad \text{und} \qquad I_- := \{n \in \mathbb{N} \mid a_n < 0\}.$$

Die Mengen I_+ oder I_- sind unendlich. (Denn: Wenn eine endlich wäre, dann wäre die Reihe absolut konvergent; Widerspruch.)

[19] An dieser Stelle gab es in der Vorlesung gegen Vorlesungsende eine Frage: es stimmt so wie es an der Tafel stand und so wie hier im Skript.

[20] Denn sei $j \in I$. Dann folgt $j \notin f_\#(\{0,1,\dots,m_0\})$ und somit $f^{-1}(j) > m_0$ und mit der Definition von m_0 erhalten wir dann, $j \notin \{0,1,\dots,n_0\}$.

Somit existieren streng monoton wachsende Abbildungen $f_\pm : \mathbb{N} \to \mathbb{N}$, mit $B(f_\pm) = I_\pm$. Wdh: $B(f_+)$ ist das Bild von f_+.

Wir setzen nun $b_n := a_{f_+(n)}$, $c_n := a_{f_-(n)}$.

$\sum_{n=0}^{\infty} b_n$ ist die Teilreihe der nicht-negativen Glieder, und $\sum_{n=0}^{\infty} c_n$ die Teilreihe der negativen Glieder. Für jede dieser Teilreihen gilt

$$\text{beschränkt} \iff \text{konvergent} \iff \text{absolut konvergent}.$$

Wenn genau eine dieser beiden Teilreihen nicht beschränkt ist, dann ist $\sum_{n=0}^{\infty} a_n$ nicht beschränkt, also nicht konvergent. Widerspruch.

Wenn beide Teilreihen beschränkt sind, dann ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent. Widerspruch.

Also sind beide Teilreihen unbeschränkt.

$$\lim_{\ell\to\infty} \sum_{n=0}^{\ell} b_n = \infty \qquad \lim_{\ell\to\infty} \sum_{n=0}^{\ell} c_n = -\infty.$$

Nach Proposition 2.8 ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge, deswegen auch $(b_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$, somit auch deren Teilfolgen $(b_{r_j})_{j\in\mathbb{N}}$ und $(c_{s_j})_{j\in\mathbb{N}}$. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $(c_n)_{n\in\mathbb{N}}$ monoton wachsend ist. (Wenn nicht, dann können wir dies durch Umordnen erreichen; jede $\mathbb{R}_{<0}$-wertige Nullfolge kann in eine monoton wachsende umgeordnet werden.) Analog können wir annehmen, dass die Teilfolge der positiven Summanden in $(b_n)_{n\in\mathbb{N}}$ monoton fallend ist.[21]

[21]Subtilität: Wir können dies nicht für die gesamte Folge erreichen: denn die Nullen machen Probleme. Da $(b_n)_{n\in\mathbb{N}}$ unbeschränkt ist, muss es unendlich viele positive b_n geben. Jede monoton fallende Nullfolge, die eine Null als Folgenglied hat, hat aber nur endlich viele positive Folgenglieder. Man lässt also die Folgenglieder, die Null sind einfach an ihrer Position stehen. Diese Folgenglieder haben sowieso

Bestimme nun zuerst das kleinste $r_1 \in \mathbb{N}$ mit

$$b_0 + b_1 + b_2 + \cdots + b_{r_1} \geq \alpha.$$

Bestimme dann das kleinste $s_1 \in \mathbb{N}$ mit

$$b_0 + b_1 + \cdots + b_{r_1} + c_0 + c_1 + c_2 \cdots + c_{s_1} < \alpha.$$

Dann gilt $b_0 + b_1 + \cdots + b_{r_1} + c_0 + c_1 + c_2 \cdots + c_{s_1-1} \geq \alpha$. [22] Hieraus folgt $|\alpha - (b_0 + b_1 + \cdots + c_{s_1})| \leq |c_{s_1}|$

Bestimme nun das kleinste $r_2 \in \mathbb{N}_{>r_1}$ mit

$$b_0 + b_1 + \cdots + b_{r_1} + c_0 + c_1 \cdots + c_{s_1} + b_{r_1+1} + b_{r_1+2} + \cdots + b_{r_2} \geq \alpha.$$

Wir erhalten $|\alpha - (b_0 + b_1 + \cdots + b_{r_2})| \leq b_{r_2}$.[23] Wir machen so weiter[24] und erhalten letztendlich eine Reihe

$$\sum_{n=0}^{\infty} d_n := b_0 + b_1 + \cdots + b_{r_1} + c_0 \cdots + c_{s_1} + b_{r_1+1} + \cdots + b_{r_2} + c_{s_1+1} + \cdots + c_{s_2} + \cdots .$$

Für $j \in \mathbb{N}$ und $\ell \geq r_j + s_j + 1$ gilt also

$$\alpha + c_{s_j} \leq \sum_{n=0}^{\ell} d_n \leq \alpha + b_{s_j}.$$

---

auf die Partialsumme keine Wirkung. Didaktische Anmerkung: die Nullen sind also zum einen ziemlich unerheblich für das Verständnis des Beweises, aber ein bisschen lästig für die Ausformulierung, da sie Spezialfälle erzeugen; deswegen habe ich die Diskussion in der Vorlesung nicht im Detail diskutiert.

[22] Im Fall $s_1 = 0$ ist dies zu lesen als $b_0 + b_1 + b_2 + \cdots + b_{r_1} \geq \alpha$.

[23] Insbesondere gilt dann $b_{r_2} > 0$.

[24] Genau genommen werden nun alle r_j und s_j rekursiv definiert und alle Aussagen dann rekursiv gezeigt

Deswegen gilt

$$\alpha = \lim_{j\to\infty}(\alpha + c_{s_j}) = \underbrace{\lim_{\ell\to\infty}\sum_{n=0}^{\ell} d_n}_{\sum_{n=0}^{\infty} d_n} = \lim_{j\to\infty}(\alpha + b_{s_j}).$$

□

Es gibt nun viele Variationen des Umordnungssatzes. Eine davon liegt der nun folgenden Multiplikation von Reihen zu Grunde. Wenn man diese (verallgemeinerten) Umordnungen systematischer verstehen will, sollte man sich mit summierbaren Familien und dem Großen Umordnungssatz (z. B. [**22**, Abschnitt 6.3]) beschäftigen. Wir überspringen aber aus Zeitgründen diesen Aspekt und konzentrieren uns auf die für uns wichtige Multiplikation.

Multiplikation von Reihen.

DEFINITION 2.34. Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ Reihen in $\mathbb{R}$ oder in $\mathbb{C}$. Wenn $\mathbb{N} \to \mathbb{N}\times\mathbb{N}$, $n \mapsto (k_n, \ell_n)$ bijektiv ist, so nennen wir die Reihe $\sum_{n=0}^{\infty} a_{k_n} b_{\ell_n}$ **eine** Produktreihe von $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$.

BEISPIELE 2.35.

(a) Eine Anordnung nach Quadraten

(b) Eine Anordnung nach Diagonalen

SATZ 2.36 (Produktreihen)**.** Sind $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ absolut konvergent, so ist jede Produktreihe von $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ absolut konvergent und hat den Grenzwert

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right).$$

Beweis.

1.) $p_n := a_{k_n} b_{\ell_n}$, $r_n := \max\{k_0, k_1, \ldots, k_n\}$, $t_n := \max\{\ell_0, \ell_1, \ldots, \ell_n\}$ Dann gilt für alle $n \in \mathbb{N}$

$$\begin{aligned} |p_0| + |p_1| + \cdots + |p_n| &\leq \left(\sum_{j=0}^{r_n} |a_j|\right)\left(\sum_{j=0}^{t_n} |b_j|\right) \\ &\leq \left(\sum_{j=0}^{\infty} |a_j|\right)\left(\sum_{j=0}^{\infty} |b_j|\right) \end{aligned}$$

Somit sind die Partialsummen der Reihe $\sum_{n=0}^{\infty} |p_n|$ beschränkt, sie bilden also eine beschränkte monoton wachsende Folge, die somit konvergiert.

Wir haben somit gesehen, dass alle Produktreihen absolut konvergieren.

2.) Wir betrachten nun zunächst den Spezialfall, dass die Produktreihe nach Quadraten geordnet ist.

Dann erhalten wir

$$p_0 + p_1 + \cdots + p_{n^2-1} = \underbrace{\left(\sum_{j=0}^{n-1} a_j\right)}_{\to \sum_{j=0}^{\infty} a_j} \underbrace{\left(\sum_{j=0}^{n-1} b_j\right)}_{\to \sum_{j=0}^{\infty} b_j}.$$

Wir haben somit eine Teilfolge der Partialsummen, die gegen

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right)$$

konvergiert. Da die Folge der Partialsummen konvergiert, konvergiert sie ebenfalls gegen diese Zahl. Die Aussage für diese spezielle Anordnung ist somit gezeigt.

3.) Durch eine Umordnung im Sinne von Definition 2.28 kann man jede Produktreihe in eine nach Quadraten angeordnete Produktreihe umformen. Deswegen ist jede Produktreihe absolut konvergent. □

Aus der Anordnung nach Diagonalen ergibt sich durch Klammersetzen[25] das folgende:

KOROLLAR 2.37 (Satz vom Cauchy-Produkt)**.** Sind $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ absolut konvergent, und setzen wir

$$c_n := \sum_{k=0}^{n} a_k b_{n-k},$$

dann ist auch $\sum_{n=0}^{\infty} c_n$ absolut konvergent und

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right).$$

Anwendung auf Potenzreihen:

Seien $\sum_{n=0}^{\infty} a_n z^n$ und $\sum_{n=0}^{\infty} b_n z^n$, Potenzreihen $a_n, b_n \in \mathbb{C}$ mit Konvergenzradien ρ_a und ρ_b. Sei $z \in \mathbb{C}$ gegeben mit $|z| < \min\{\rho_a, \rho_b\}$. Wieder $c_n := \sum_{k=0}^{n} a_k b_{n-k}$. Dann ist $|z|$ auch kleiner als der Konvergenzradius von $\sum_{n=0}^{\infty} c_n x^n$, denn wenn $\sum_{n=0}^{\infty} a_n |z|^n$ und $\sum_{n=0}^{\infty} b_n |z|^n$ konvergieren, dann konvergiert nach dem Satz vom Cauchy-Produkt (Korollar 2.37) auch $\sum_{n=0}^{\infty} c_n |z|^n$.

$$\begin{aligned} & (a_0 + a_1 z + a_2 z^2 + \cdots)(b_0 + b_1 z + b_2 z^2 + \cdots) \\ = \; & \underbrace{(a_0 b_0)}_{c_0} + \underbrace{(a_0 b_1 + a_1 b_0)}_{c_1} z + \underbrace{(a_0 b_2 + a_1 b_1 + a_2 b_0)}_{c_2} z^2 + \cdots \end{aligned}$$

[25] dies ist erlaubt, vergleiche Eigenschaften 2.9 (b)

Anschaulich: Man darf hier nach Potenzen in z ordnen. Alle Grenzwerte existieren und es gilt Gleichheit.

3. Einige durch Reihen definierte Funktionen

Sei $z \in \mathbb{C}$. Potenzreihen definieren mehrere wichtige Funktionen:

3.1. Exponentialfunktion.

$$\exp z := \sum_{n=0}^{\infty} \frac{1}{n!} z^n.$$

Diese Reihe nennt man die Exponentialreihe.

Wir bestimmen den Konvergenzradius: Wir wenden hierzu das Quotienten-Kriterium an

$$\left| \frac{a_{n+1} z^{n+1}}{a_n z^n} \right| = \frac{|z|}{n+1}$$

$$\limsup_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \frac{|z|}{n+1} = 0 < 1$$

Die Exponentialreihe konvergiert also absolut für alle $z \in \mathbb{C}$. Wir sagen sie konvergiert (absolut) auf $\mathbb{C}$. [26] Der Konvergenzradius muss dann $\rho = \infty$ sein, denn angenommen wir hätten $\rho < \infty$, dann würde die Reihe für $z := \rho + 1$ divergieren, was einen Widerspruch ergibt.

[26] Das Wort absolut wird hier normalerweise weggelassen, denn für jede Potenzreihe gilt: wenn sie in allen Punkten $z \in \mathbb{C}$ konvergiert, so ist der Konvergenzradius ∞ und deswegen konvergiert sie in allen $z \in \mathbb{C}$ sogar absolut.

Wir erhalten eine Funktion[27] $\exp : \mathbb{C} \longrightarrow \mathbb{C}$, $z \mapsto \exp(z)$, die komplexe Exponentialfunktion.

Ist $z \in \mathbb{R}$, so ist auch $\exp(z) \in \mathbb{R}$. Wir erhalten somit eine Funktion $\exp : \mathbb{R} \longrightarrow \mathbb{R}$, $x \mapsto \exp x$, die reelle Exponentialfunktion. [28]

PROPOSITION 3.1 (Funktionalgleichung der Exponentialfunktion)**.** Für alle $x, w \in \mathbb{C}$ gilt

$$\exp(z+w) = \exp z \cdot \exp w$$

Beweis. Wir rechnen:

$$\frac{1}{n!}(z+w)^n = \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k} z^k w^{n-k} = \sum_{k=0}^{n}\frac{1}{k!\,(n-k)!} z^k w^{n-k} = \sum_{k=0}^{n}\frac{1}{k!} z^k \,\frac{1}{(n-k)!} w^{n-k}$$

Wende den Satz vom Cauchy-Produkt an für

$$a_n := \frac{1}{n!}z^n \quad \text{und} \quad b_n := \frac{1}{n!}w^n \quad \text{und} \quad c_n := \frac{1}{n!}(z+w)^n.$$

□ **Mi 9.1.**

Spezialfall: $\exp(0) := \sum_{n=0}^{\infty} = 0^0 + 0^1 + 0^2 + \ldots = 1$

Man definiert immer $0^0 := 1$.

[27] Wie gesagt: wir nutzen den Begriff „Funktion“ gleichbedeutend zum Begriff „Abbildung“, allerdings ist es üblich, dass Funktionen den Zielbereich $\mathbb{R}$ oder $\mathbb{C}$ haben.

[28] Die Relation, die der reellen Exponentialfunktion unterliegt, erhält man durch Restriktion der Relation, die der komplexen Exponentialfunktion unterliegt. Es handelt sich also um zwei verschiedene Relationen, die gemäß unseren Definitionen im ersten Kapitel auch verschiedene Symbole haben sollte. Wir benutzen aber dennoch dasselbe Symbol exp, da dies so üblich ist und in der Praxis nur selten zu Missverständnissen führt.

Also $\exp 0 = 1$.

Es folgt dann auch

$$1 = \exp 0 = \exp(z - z) = (\exp z)(\exp(-z)).$$

Also

$$\exp(-z) = \frac{1}{\exp z}.$$

Es gilt (Übung) $\lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n = e^z$.

Klammer-„Regel“ für exp **und die demnächst definierten Funktionen** sin **und** cos**.**

Man lässt bei exp und anderen wichtigen Funktionen die sonst bei Funktionen üblichen Klammern weg. Man schreibt also $\exp z$ und nicht $\exp(z)$. Nun stellt sich die Frage, ob bei Ausdrücken in der Art $\exp ab$ und $\exp a + b$ zuerst exp ausgeführt wird oder zuerst verknüpft wird.

Ein Blick in die Literatur ergibt leider: das ist nicht ganz so klar geregelt, man nutzt hier eine bunte Mischung spezieller Konventionen. Im allgemeinen reiht man – wenn keine Klammern stehen – aber wie folgt, wobei die oben stehenden Operationen zuerst ausgeführt werden:

(1) Potenzen,
(2) Produkte, Vorzeichen, Quotienten[29]
(3) exp, sin, cos,

[29]Quotienten werden oft als Bruch geschrieben, dann ist die Reihenfolge zumeist klar

(4) Addition und Subtraktion

Oft wird jedoch (3) vor (2) oder hinter (4) ausgeführt. Also lieber großzügig Klammern setzen!

$\exp it = \exp(it)$

$\cos 2a = \cos(2a)$

$\cos -\alpha/2 = \cos(-\alpha/2)$

$\exp a + b = (\exp a) + b$ **meistens!** Besser also Klammern setzen.

$\sin a \sin b =$? Nach den obigen Regeln hätten wir dann $\sin(a(\sin b))$. Das interpretiert aber niemand so. Solch ein Ausdruck wird immer als $(\sin a)(\sin b)$ interpretiert.

Besser treffend wäre also

(1) Potenzen,
(2) Produkte mit kurzen Faktoren, Vorzeichen, Quotienten
(3) exp, sin, cos,
(4) Produkte mit langen Faktoren
(5) Addition

Nur: was ist lang und was ist kurz? Ab wann wird $\exp abcdefghijklm\cdots$ lang?

Die Verwirrung wird noch schlimmer: Da man ja für Funktionen $f : M \longrightarrow M$ **immer** definiert

$$f^2(x) := f(f(x))$$

würde dann $\sin^2 a$ als $\sin(\sin(a))$ zu lesen sein. Das macht aber auch niemand. Solch ein Ausdruck wird immer als $(\sin a)^2$ interpretiert.

Um das Durcheinander zu beseitigen, würde nur eine von der Kultusminister-Konferenz eingesetzte Experten-Kommission helfen, die sich zuerst mit europäischen Instanzen zusammenschließt und nach reiflicher, mehrjähriger Diskussion eine DIN-Norm erlässt.[30] Das wäre Bürokratismus pur. Und selbst, wenn die DIN erlassen wäre, ginge es dieser Reform vielleicht nicht anders als mit der letzten Rechtschreibreform: manche Bundesländer würden sie einführen, andere ablehnen. Etc. Oder es gibt sie und niemand kennt sie. [31]

Und den meisten Mathematiker sind solche Diskussionen sowieso lästig.

Also: da leben wir lieber mit weniger Ordnungen und setzen ein paar Klammern mehr.

3.2. Sinus- und Kosinus-Funktion. Sei wieder i die komplexe Zahl mit $i^2 = -1$.

Betrachte nun

$$\exp iz = \sum_{n=0}^{\infty} \frac{1}{n!}(iz)^n = \sum_{\substack{n=0 \\ n \text{ gerade}}}^{\infty} \frac{1}{n!}(iz)^n + \sum_{\substack{n=0 \\ n \text{ ungerade}}}^{\infty} \frac{1}{n!}(iz)^n$$

[32]

Wir definieren den Kosinus

$$\cos z := \sum_{\substack{n=0 \\ n \text{ gerade}}}^{\infty} \frac{1}{n!}(iz)^n = \sum_{k=0}^{\infty} \frac{1}{(2k)!}(iz)^{2k} = \sum_{k=0}^{\infty} \frac{1}{(2k)!}(-1)^k z^{2k} = 1 - \frac{1}{2}z^2 + \frac{1}{24}z^4 - \cdots$$

[30] Gibt es vielleicht sogar?

[31] Vielleicht ist es sogar so?

[32] Hier handelt es sich rechts um zwei Reihen, wo eigentlich an den ungeraden bzw. geraden Stellen n eine Null steht. Nullen in einer Reihen haben keinen Einfluss auf den Grenzwert einer Reihe, sie können beliebig hinzugefügt oder gestrichen werden. Mit dieser Interpretation ist die letzte Gleichung eine Folge von $\sum_{n=0}(a_n + b_n) = \sum_{n=0} a_n + \sum_{n=0} b_n$

Was ist der Konvergenzradius dieser Reihe?

n gerade: $|\frac{1}{n!}(iz)^n| \leq \frac{1}{n!}|z|^n$

n ungerade: $0 \leq \frac{1}{n!}|z|^n$

Nach dem Majoranten-Kriterium und der Konvergenz der Exponentialfunktion, konvergiert die Kosinusreihe auf ganz $\mathbb{C}$. Der Konvergenzradius ist also ∞.

Wir definieren den Sinus

$$\begin{aligned}\sin z &:= \frac{1}{i} \sum_{\substack{n=0 \\ n \text{ ungerade}}}^{\infty} \frac{1}{n!}(iz)^n = \frac{1}{i}\sum_{k=0}^{\infty} \frac{1}{(2k+1)!}(iz)^{2k+1} \\ &= \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}(-1)^k z^{2k+1} = z - \frac{1}{6}z^3 + \frac{1}{120}z^5 - \cdots\end{aligned}$$

Der Konvergenzradius ist ebenfalls ∞.

$$\exp iz = \cos z + i \sin z$$

$$\cos(-z) = \cos z$$

$$\sin(-z) = -\sin z$$

$$\frac{\exp iz + \exp(-iz)}{2} = \frac{\cos z + i \sin z + \cos z - i \sin z}{2} = \cos z$$

$$\frac{\exp iz - \exp(-iz)}{2i} = \sin z$$

Ist x reell, dann sind auch $\cos x$ und $\sin x$ reell. Dann gilt also

$$\cos x = \mathbf{Re}\, \exp ix \qquad \sin x = \mathbf{Im}\, \exp ix.$$

Man überlegt sich auch leicht, dass für jede Reihe in $\mathbb{C}$ gilt

$$\overline{\sum_{n=0}^{\infty} a_n} = \sum_{n=0}^{\infty} \overline{a_n}, \quad \mathbf{Re} \sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} \mathbf{Re}\, a_n, \quad \mathbf{Im} \sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} \mathbf{Im}\, a_n\,.$$

Somit gilt für $x \in \mathbb{R}$:

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

Also $|\exp(ix)|^2 = \cos^2 x + \sin^2 x = 1$.

Hier nutzen wir die Schreibwweisen $\sin^2 x := (\sin x)^2$ und $\cos^2 := (\cos x)^2$.

Proposition 3.2 (Additionstheoreme)**.** Für alle $a, b \in \mathbb{C}$ gilt

$$\cos(a+b) = \cos a \cos b - \sin a \sin b,$$

$$\sin(a+b) = \sin a \cos b + \cos a \sin b.$$

Wir geben zunächst einen Beweis für $a, b \in \mathbb{R}$ und dann einen für $a, b \in \mathbb{C}$. Der erste Beweis ist also aus Sichtweise des logischen Aufbaus der Vorlesung unnötig. Allerdings ist der erste Beweis viel besser zu merken und die Additionstheoreme sind vor allem für reelle a und b wichtig. Deswegen ist der erste für unser Verständnis der wichtigere.

Beweis für reelle a und b.

$$\begin{aligned}\cos(a+b) &= \mathbf{Re}\ \exp i(a+b) = \mathbf{Re}\ ((\exp ia)(\exp ib)) \\ &= \mathbf{Re}\ \big((\cos a + i\sin a)(\cos b + i\sin b)\big) = \cos a\cos b - \sin a\sin b\end{aligned}$$

$$\begin{aligned}\sin(a+b) &= \mathbf{Im}\ \exp i(a+b) = \mathbf{Im}\ ((\exp ia)(\exp ib)) \\ &= \mathbf{Im}\ \big((\cos a + i\sin a)(\cos b + i\sin b)\big) = \cos a\sin b + \sin a\cos b\end{aligned}$$

□

Beweis für alle a und b.

$$\cos(a+b) = \frac{\exp i(a+b) + \exp(-i(a+b))}{2} = \frac{\exp ia\exp ib + \exp(-ia)\exp(-ib)}{2}$$

$$\cos a\cos b = \frac{\exp ia\exp ib + \exp ia\exp(-ib) + \exp(-ia)\exp ib + \exp(-ia)\exp(-ib)}{4}$$

$$\sin a\sin b = \frac{\exp ia\exp ib - \exp ia\exp(-ib) - \exp(-ia)\exp ib + \exp(-ia)\exp(-ib)}{-4}$$

Dies ergibt die Formel für $\cos(a+b)$. Der Beweis der Formel für $\sin(a+b)$ geht völlig analog. □

33

[33]Wir hätten gerne natürlich noch ganz andere Eigenschaften: zum Beispiel, dass *sin* und cos die üblichen geometrischen Eigenschaften haben, und dass sie periodisch sind, genauer dass $\cos(z+2\pi) = \cos z$ und $\sin(z+2\pi) = \sin z$. Das können wir aber noch nicht sagen, denn wir müssen ja erst mal π definieren. Man definiert $\pi/2$ als die erste positive Nullstelle von sin, aber wir können leider noch gar nicht zeigen, dass die oben definierte Funktion überhaupt eine positive Nullstelle beseitzt.

3.3. Eulersche Zahl. Die Eulersche Zahl ist

$$e := \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!} > 1 + 1 + \frac{1}{2} + \frac{1}{6} = 2,6666\ldots$$

Numerisch[34]: $e = 2,718\,281\,828\,459\,045\,235\,360\,287\,471\,352\,662\,497\,757\,247\ldots$

Wdh: Für $t \in \mathbb{R}_{>0}$, $n \in \mathbb{N}$, $m \in \mathbb{N} \setminus \{0\}$ ist definiert:

t^n: rekursiv

$\sqrt[m]{t}$: durch Lemma 6.10 in Kapitel 2.

$t^{n/m} := \sqrt[m]{t^n}$.

$t^{-n/m} := \frac{1}{t^{n/m}}$:

Lemma 3.3. Für alle $q \in \mathbb{Q}$ gilt

$$e^q = \exp(q)$$

Beweis. Durch Induktion zeigt man zunächst für alle $n \in \mathbb{N}$ und alle $z \in \mathbb{C}$: $\exp(nz) = (\exp z)^n$. Angewendet auf $z := 1$ ergibt dies $\exp(n) = \exp(n \cdot 1) = e^n > 0$. Für $m \in \mathbb{N} \setminus \{0\}$ gilt dann

$$\exp(n/m) := \sum_{j=0}^{\infty} \frac{1}{j!} \left(\frac{n}{m}\right)^j > 0.$$

Nun setzen wir $z := n/m$ in $\exp(nz) = \exp(z)^n$ und erhalten

$$(\exp n/m)^m = \exp(mn/m) = e^n.$$

[34]Dieser Wert ist hier nicht vollständig angegeben, kann auch nicht vollständig als Dezimalzahl angegeben werden. Wir wollen hier nicht begründen, wieso dies die ersten Stellen der Eulerschen Zahl sind. Wir werden diesen Näherungswert im folgenden nie benötigen. Der Näherungswert dient allein der Anschauung und dem Vergleich zu anderer Literatur und anderen Konstanten!

Somit ist $x := \exp(n/m)$ eine positive reelle Zahl, die $x^m = e^n$ erfüllt. Also $x = e^{n/m}$. Wir haben somit die Aussage für alle $q \in \mathbb{Q}_{>0}$ gezeigt. Die Aussage ist trivial für $q = 0$ und folgt direkt aus der Funktionalgleichung der Exponentialfunktion (Proposition 3.1) für $q \in \mathbb{Q}_{<0}$. □

DEFINITION 3.4. Für $z \in \mathbb{C}$ definieren wir $e^z := \exp(z)$.

Funktionalgleichung: $e^{z+w} = e^z \cdot e^w$

BEMERKUNG 3.5. Die Eulersche Zahl ist irrational. Sie ist sogar transzendent, d.h. ist $p(x) = a_d x^d + \cdots + a_0 x^0$, $p \not\equiv 0$, [35] eine polynomiale Funktion mit rationalen Koeffizienten, dann gilt $p(e) \neq 0$. (Ohne Beweis)

3.4. Exponentialfunktion (Fortsetzung). In diesem Abschnitt studieren wir die reelle Exponentialfunktion $\exp : \mathbb{R} \longrightarrow \mathbb{R}$.

SATZ 3.6.

(a) $\forall r \in \mathbb{R} : \exp r > 0$, also $\exp_{\#}(\mathbb{R}) \subset \mathbb{R}_{>0}$.

(b) Die reelle Exponentialfunktion $\exp : \mathbb{R} \longrightarrow \mathbb{R}$ ist streng monoton wachsend.

(c) Sie wächst schneller als jede polynomiale Funktion im folgenden Sinn: Für jede polynomiale Funktion $p(x) = a_d x^d + \cdots + a_0 x^0$, gilt:

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists x_0 \in \mathbb{R} : \forall x \in \mathbb{R}_{\geq x_0} : \left| \frac{p(x)}{\exp(x)} \right| \leq \epsilon$$

[35] Für eine Funktion $f : M \to \mathbb{R}$ bedeutet $f \equiv 0$, dass für alle $x \in M$ gilt $f(x) = 0$. Oben gilt dann also

$$p \not\equiv 0 \quad \Longleftrightarrow \quad \neg(p \equiv 0) \quad \Longleftrightarrow \quad \neg(a_d = a_{d-1} = \cdots = a_0 = 0)$$

.

Insbesondere: ist $(a_k)_{k\in\mathbb{N}}$ eine Folge, die uneigentlich gegen ∞ konvergiert, dann ist

$$\left(\frac{p(a_k)}{\exp(a_k)}\right)_{k\in\mathbb{N}}$$

eine Nullfolge.

Beweis.

Zu (a): bereits gezeigt.

Zu (b): Es gelte $x, y \in \mathbb{R}$, $x < y$.

$$\frac{\exp y}{\exp x} = \exp(y-x) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \underbrace{(y-x)^n}_{>0} > 1.$$

Zu (c): Sei $\epsilon \in \mathbb{R}_{>0}$ gegeben. Wir setzen

$$x_0 := \frac{\max\{1!\,|a_0|, 2!\,|a_1|, \ldots, (d+1)!\,|a_d|\}}{\epsilon} > 0.$$

Somit haben wir für $\forall x \in \mathbb{R}_{\geq x_0}$

$$\left|a_\ell x^\ell\right| \leq \frac{\epsilon}{(\ell+1)!}|x_0| \cdot |x^\ell| \leq \frac{\epsilon}{(\ell+1)!} x^{\ell+1},$$

$$|p(x)| = |a_d x^d + \cdots + a_0 x^0| \leq \epsilon \sum_{n=1}^{d+1} \frac{1}{n!} x^n < \epsilon \exp(x).$$

□

KOROLLAR 3.7. Konvergiert $(a_k)_{k\in\mathbb{N}}$ uneigentlich gegen ∞, dann konvergiert $\exp(-a_k) \to 0$ und $\exp(a_k) \to \infty$.

FRAGE 3.8. Ist $\exp(\mathbb{R}) = \mathbb{R}_{>0}$?

Es erscheint uns, dass die Antwort ja sein sollte, obwohl wir es noch nicht zeigen können. Die Frage ist auch wichtig, weil wir gerne den (natürlichen) Logarithmus als Umkehrfunktion der reellen Exponentialfunktion definieren wollen, und der Logarithmus sollte auf allen positiven reellen Zahlen definiert sein.

BEISPIEL 3.9. Betrachte die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$

$$f(x) := \begin{cases} \exp x & x \leq 0 \\ 1 + \exp x & x > 0 \end{cases}$$

Die Funktion $f : \mathbb{R} \to \mathbb{R}$ aus Beispiel 3.9

Diese Funktion erfüllt alle Eigenschaften in Satz 3.6 und Korollar 3.7. Aber $f_{\#}(\mathbb{R}) = (0; 1] \cup (2; \infty) \neq \mathbb{R}_{>0}$.

Anschaulich: Da $\exp : \mathbb{R} \longrightarrow \mathbb{R}$ keine „Sprünge macht“, sollte jeder Wert zwischen 0 und ∞ von der Funktion einmal angenommen werden.

Uns fehlt also ein wichtiges Hilfsmittel. Wir wollen deswegen im folgenden Kapitel klären, was es bedeutet, dass eine Funktion „keine Sprünge macht“. Dies führt zum Begriff „Stetigkeit“.

KAPITEL 4

Stetigkeit und Grenzwerte von Funktionen

Fr 11.1.

1. Stetigkeit

Notationswiederholung $f_{\#}(D) = \{f(x) \mid x \in D\}$, siehe (7.25).

Im folgenden sei: $D \subset \mathbb{C}$, $x_0 \in D$, $f : D \longrightarrow \mathbb{C}$.

Stellen Sie sich am besten zunächst den Speziallfall $D \subset \mathbb{R}$ vor und nehmen Sie zunächst $f_{\#}(D) \subset \mathbb{R}$ an. Wenn Sie diesen Fall richtig verstanden haben, betrachten Sie den Fall $D \subset \mathbb{C}$. In späteren Semestern werden wir aber noch wesentlich allgemeinere Definitionsbereiche D benötigen, zum Beispiel 4-dimensionale gekrümmte Raumzeiten, unendlichdimensionale Vektorräume oder „topologische Räume“ und wir wollen die jetzt gemachten Definitionen möglichst so wählen, dass sich alles auf diesen allgemeineren Kontext verallgemeinert.

FRAGE 1.1. Sind Funktionen mit dem Grenzwert von Folgen verträglich?
Es gelte $a_n \to x_0$. Folgt dann $f(a_n) \to f(x_0)$?

DEFINITION 1.2. Eine Funktion $f : D \longrightarrow \mathbb{C}$ heißt folgenstetig in x_0, falls gilt: [1]

(1.3) Ist $(a_n)_{n\in\mathbb{N}}$ eine beliebige Folge in D mit $\lim\limits_{n\to\infty} a_n = x_0$,
dann folgt $\lim\limits_{n\to\infty} f(a_n) = f(x_0)$.

Eine Funktion $f : D \longrightarrow \mathbb{C}$ heißt folgenstetig, falls sie in allen $x_0 \in D$ folgenstetig ist.

BEISPIELE 1.4.

(a) Konstante Funktionen folgenstetig. Die Identitäten $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ und $\mathrm{id}_{\mathbb{C}} : \mathbb{C} \to \mathbb{C}$ sind folgenstetig.

(b) Ist $f : D \to \mathbb{C}$ folgenstetig in x_0 und $x_0 \in \widetilde{D} \subset D$, dann ist auch $f|_{\widetilde{D}} : \widetilde{D} \longrightarrow \mathbb{C}$ in x_0 folgenstetig. [2]

(c) Die Heaviside-Funktion $\Theta_1 : \mathbb{R} \to \mathbb{R}$

$$\Theta_1(x) := \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

ist nicht folgenstetig in 0. Denn die Folge $(1/n)_{n\in\mathbb{N}\setminus\{0\}}$ konvergiert gegen 0, aber $\Theta_1(1/n) = 1$ konvergiert nicht gegen $\Theta_1(0) = 0$.

Funktionsgraph der Heaviside-Funktion Θ_1

[1] Damit es keine Missverständnisse gibt, wie (1.3) zu verstehen ist, wollen wir noch eine logisch präzisere Formulierung von (1.3) geben, in der der Allquantor deutlich wird. Es ist: Für allle Folgen $(a_n)_{n\in\mathbb{N}}$ in D gilt:

$$\lim_{n\to\infty} a_n = x_0 \Longrightarrow \lim_{n\to\infty} f(a_n) = f(x_0).$$

[2] Denn: Wenn man D verkleinert, muss man weniger Folgen überprüfen.

Anderseits ist Θ_1 in allen anderen x_0 folgenstetig. Denn ist zum Beispiel $x_0 > 0$, und gilt $a_n \to x_0$, dann folgt für fast alle $n \in \mathbb{N}$: $a_n > 0$, also $\Theta_1(a_n) = 1$, und somit $\Theta_1(a_n) \to \Theta_1(x_0) = 1$. Der Fall $x_0 < 0$ ist analog.

(d) Ist $p : \mathbb{C} \longrightarrow \mathbb{C}$, $x \mapsto \sum_{k=0}^{d} b_k x^k$ eine polynomiale Funktion, so ist p in jedem $x_0 \in \mathbb{R}$ folgenstetig. Es gelte $\lim_{n\to\infty} a_n = x_0$. Dann

$$\lim_{n\to\infty} p(a_n) = \sum_{k=0}^{d} \lim_{n\to\infty} b_k (a_n)^k = \sum_{k=0}^{d} b_k (\lim_{n\to\infty} a_n)^k = p(x_0).$$

Dies folgt direkt aus Proposition 1.10 in Kapitel 3.

(e) Realteil $\mathbf{Re} : \mathbb{C} \to \mathbb{C}$, Imaginärteil $\mathbf{Im} : \mathbb{C} \to \mathbb{C}$ und Konjugation $\mathbb{C} \to \mathbb{C}$, $z \mapsto \overline{z}$ sind folgenstetig.

(f) Die Funktion f aus Beispiel 3.9 in Kapitel 3 ist in 0 nicht folgenstetig. Denn für $a_n := \frac{1}{n}$ gilt: $f(a_n) > 2$, $f(0) = f(\lim a_n) = 1$; deswegen kann $f(a_n)$ nicht gegen $f(\lim a_n)$ konvergieren. [3]

Definition 1.5.

(a) Eine Funktion $f : D \longrightarrow \mathbb{C}$ heißt stetig in x_0, falls gilt:

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists \delta \in \mathbb{R}_{>0} : \forall x \in D : (|x - x_0| < \delta \Longrightarrow |f(x) - f(x_0)| < \epsilon). \tag{1.6}$$

(b) $f : D \longrightarrow \mathbb{C}$ heißt stetig, wenn f in jedem $x_0 \in D$ stetig ist.

Bild zur Veranschaulichung der Stetigkeit

[3] Achtung: Diese Aussage ist – so wie sie steht – korrekt. Wir haben noch nicht gezeigt, dass $f(a_n)$ konvergiert, denn wir wissen noch nicht, ob die Exponentialfunktion stetig in 0 ist. Wir haben aber im obigen Satz auch nicht behauptet, dass diese Folge konvergiert!

LEMMA 1.7. $x_0 \in D \subset \mathbb{C}$, $f : D \longrightarrow \mathbb{C}$

$$f \text{ ist folgenstetig in } x_0 \quad \Longleftrightarrow \quad f \text{ ist stetig in } x_0.$$

In anderen Worten: (1.3) $\Longleftrightarrow$ (1.6).

Beweis.

Zu „$\Longrightarrow$“: Sei f nicht stetig in x_0. Dann gilt

$$\exists \epsilon \in \mathbb{R}_{>0} : \forall \delta \in \mathbb{R}_{>0} : \exists x \in D : (|x - x_0| < \delta \wedge |f(x) - f(x_0)| \geq \epsilon).$$

Wir fixieren solch ein $\epsilon \in \mathbb{R}_{>0}$. Für ein $n \in \mathbb{N} \setminus \{0\}$ setzen wir $\delta := 1/n$, und wählen ein $a_n \in D$ mit

$$|a_n - x_0| < 1/n \text{ und } |f(a_n) - f(x_0)| \geq \epsilon.$$

Dann konvergiert $(a_n)_{n \in \mathbb{N} \setminus \{0\}}$ gegen x_0, aber $(f(a_n))_{n \in \mathbb{N} \setminus \{0\}}$ konvergiert nicht gegen $f(x_0)$.

Zu „$\Longleftarrow$“: Sei f stetig in x_0, und (a_n) sei eine Folge in D, die gegen x_0 konvergiert.

Zu zeigen ist: $f(a_n) \to f(x_0)$, d.h.: [4]

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists n_0 \in \mathbb{N} : \forall n \in \mathbb{N}_{\geq n_0} : |f(a_n) - f(x_0)| < \epsilon \tag{1.8}$$

Wir fixieren also ein $\epsilon \in \mathbb{R}_{>0}$. Wir wählen ein dazu passendes $\delta \in \mathbb{R}_{>0}$ wie in (1.6), d.h. es gilt:

$$\forall x \in D : (|x - x_0| < \delta \Longrightarrow |f(x) - f(x_0)| < \epsilon). \tag{1.9}$$

[4] Bisher hatten wir in der Definition von Konvergenz von Folgen $\leq \epsilon$. Es ist aber klar, dass man hier auch $< \epsilon$ schreiben kann.

Wegen $a_n \to x_0$ gibt es ein $n_0 \in \mathbb{N}$ mit

$$\forall n \in \mathbb{N}_{\geq n_0} : |a_n - x_0| < \delta \tag{1.10}$$

Die Aussagen (1.9) und (1.10) zusammen ergeben

$$\forall n \in \mathbb{N}_{\geq n_0} : |f(a_n) - f(x_0)| < \epsilon \tag{1.11}$$

Insgesamt haben wir nun (1.8) verifiziert. □

Sprachgebrauch: Wir nennen (1.3) das Folgen-Kriterium für Stetigkeit und (1.6) das ϵ-δ-Kriterium für Stetigkeit.

Lemma 1.12. Seien $f, g : D \longrightarrow \mathbb{C}$ in x_0 stetige Funktionen, $a, b \in \mathbb{C}$. Dann sind auch die folgenden Funktionen $D \longrightarrow \mathbb{C}$ stetig in x_0:

$$(a \cdot f + b \cdot g)(x) := a \cdot f(x) + b \cdot g(x)$$

$$(f \cdot g)(x) := f(x) \cdot g(x)$$

Gilt außerdem $0 \notin g_{\#}(D)$, dann ist auch

$$\left(\frac{f}{g}\right)(x) := \frac{f(x)}{g(x)}$$

stetig in x_0.

Aus Proposition 1.10 in Kapitel 3 folgt obiges Lemma mit dem Folgenkriterium für Stetigkeit.

Lemma 1.13. Sei $f : D \longrightarrow \mathbb{C}$ stetig in $x_0 \in D$ und sei $g : \widetilde{D} \longrightarrow \mathbb{C}$ stetig in $y_0 \in \widetilde{D}$. Es gelte $f(x_0) = y_0$ und $B(f) = f_{\#}(D) \subset \widetilde{D}$. Dann ist $g \circ f : D \longrightarrow \mathbb{C}$ ebenfalls stetig in x_0.

Beweis. Wir überprüfen das Folgenkriterium. Es gelte für $n \to \infty$: $a_n \to x_0$. Aus den Voraussetzungen folgt zunächst $f(a_n) \to f(x_0) = y_0$ und daraus dann $g(f(a_n)) \to g(y_0)$, also $(g \circ f)(a_n) \to (g \circ f)(x_0)$. Somit ist $g \circ f$ folgenstetig in x_0. □

Beispiel 1.14. Sei $K = \mathbb{C}$ oder $K = \mathbb{R}$. Dann ist die Betragsfunktion

$$|\cdot| : K \to \mathbb{R}, \qquad z \mapsto |z|$$

stetig.

Begründung: Wir zeigen die Stetigkeit in $x_0 \in K$. Für $\epsilon \in \mathbb{R}_{>0}$ wähle $\delta := \epsilon$. Zu zeigen ist

$$|x - x_0| < \epsilon \Longrightarrow \big||x| - |x_0|\big| < \epsilon.$$

Dies folgt direkt aus Lemma 5.11 (g) in Kapitel 2.

Übung 1.15 (Übungsblatt 11, 2. Aufgabe)**.** Sei $D \subset \mathbb{C}$, $x_0 \in D$ und $f : D \to \mathbb{C}$ eine Funktion, und sei $r \in \mathbb{R}_{>0}$ fixiert. Wir definieren $B_r(x_0) := \{z \in \mathbb{C} \,|\, |z - x_0| < r\}$. Zeigen Sie: f ist stetig in x_0 genau dann, wenn $f|_{D \cap B_r(x_0)}$ stetig in x_0 ist.

Wir sagen hierzu: Stetigkeit ist eine lokale Eigenschaft.

2. Zwischenwertsatz

In diesem Abschnitt ist der Definitionsbereich D von f eine Teilmenge von $\mathbb{R}$, $f : D \longrightarrow \mathbb{R}$.

Satz 2.1 (Zwischenwertsatz)**.** Sei $f : [a, b] \longrightarrow \mathbb{R}$ eine stetige Funktion, und sei $f(a) \leq y \leq f(b)$ oder $f(a) \geq y \geq f(b)$. Dann gibt es ein $c \in [a, b]$ mit $f(c) = y$.

Wir haben dann also

$$[f(a), f(b)] \subset f_{\#}([a, b]) \text{ oder } [f(b), f(a)] \subset f_{\#}([a, b])$$

Bild mit der obigen Funktion f

Beweis. O.B.d.A. $f(a) \leq y \leq f(b)$.

Wir definieren rekursiv die Folgen (a_n) und (b_n).

$a_0 := a$, $b_0 := b$.

Seien a_{n-1} und b_{n-1} definiert. Setze

$$z_{n-1} := \frac{a_{n-1} + b_{n-1}}{2}$$

Wir definieren dann a_n und b_n wie folgt:

- Im Fall $f(z_{n-1}) \leq y$ setzen wir $a_n := z_{n-1}$ und $b_n := b_{n-1}$.
- Im Fall $f(z_{n-1}) > y$ setzen wir $a_n := a_{n-1}$ und $b_n := z_{n-1}$.

Die Folge (a_n) ist monoton wachsend und beschränkt. Also existiert $c := \lim_{n\to\infty} a_n$. Da $b_n - a_n = (b-a)/2^n \to 0$ haben wir auch $c = \lim_{n\to\infty} b_n$. Mit der Folgenstetigkeit von f in c sehen wir:

$$f(c) = \lim_{n\to\infty} f(a_n) \leq y \text{ und } f(c) = \lim_{n\to\infty} f(b_n) \geq y. \qquad \square$$

Mi 16.1.

Beispiel 2.2. Die Funktion $f : [-1, 2] \to \mathbb{R}$, $f(x) = x^2$ ist stetig. Es gilt

$$[f(-1), f(2)] = [1, 4] \subsetneq [0, 4] = f_{\#}([-1, 2]).$$

Beispiel 2.3. Sei $a \in \mathbb{R}_{>0}$ und $n \in \mathbb{N}_{\geq 2}$. Betrachte die stetige Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$, $x \mapsto x^n - a$. Es gilt $f(0) = -a < 0$ und

$$f(a+1) = \left(\sum_{j=0}^{n} \binom{n}{j} a^j \right) - a \geq (n-1)a + 1 > 0.$$

Also besitzt f eine Nullstelle im Intervall $[0, a+1]$. Dies ist ein alternativer Beweis der Existenz von $\sqrt[n]{a}$.

KOROLLAR 2.4. Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ stetig. Dann ist $f_{\#}(I)$ ebenfalls ein Intervall.

Beachte: Intervalle können hier offen, halboffen oder abgeschlossen sein! Sie können beschränkt oder unbeschränkt sein!

Vorüberlegung: Eine Teilmenge J von $\mathbb{R}$ ist genau dann ein Intervall, wenn gilt:

$$\forall u, v \in J : u \leq v \Longrightarrow [u, v] \subset J$$

Beweis. Seien $u, v \in f_{\#}(I)$. Wir schreiben $u = f(a)$, $v = f(b)$, $a, b \in I$. O.B.d.A. $a \leq b$. Nach dem Zwischenwertsatz gilt

$$[u, v] = [f(a), f(b)] \subset f_{\#}([a, b]) \subset f_{\#}(I).$$

Also ist $f_{\#}(I)$ ein Intervall. □

BEISPIEL 2.5. Die Funktion $f : (-1, 2] \to \mathbb{R}$, $f(x) = x^2$ ist stetig. $f_{\#}((-1, 2]) = [0, 4]$. Die Bilder halboffener Intervalle können also abgeschlossen sein.

BEISPIEL 2.6. $f : (-1, 1) \longrightarrow \mathbb{R}\ x \mapsto x^2$. Dann ist $f_{\#}((-1, 1)) = B(f) = [0, 1)$. Wir haben hier ein Intervall erhalten. Das Bild-Intervall ist nicht mehr offen, sondern nur noch halboffen. Ähnliche Beispiele zeigen: die Bilder offener Intervalle können abgeschlossen sein.

SATZ 2.7. Sei $I \subset \mathbb{R}$ ein Intervall, und $f : I \longrightarrow \mathbb{R}$ eine streng monotone, stetige Funktion, $J := f_{\#}(I)$. Dann ist $f^{-1} : J \longrightarrow I$ wohldefiniert und stetig.

Beweis. Da $f : I \longrightarrow J$ eine bijektive Funktion ist, ist die Umkehrfunktion $f^{-1} : J \longrightarrow I$ wohldefiniert.

Graph einer stetigen streng monotonen Funktion $f : I \to J$

Wir zeigen nun die Stetigkeit von f^{-1} in einem beliebigen $y_0 \in J$. O.B.d.A. sei f streng monoton wachsend.[5] Zu zeigen ist:

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists \delta \in \mathbb{R}_{>0} : \forall y \in J : (|y - y_0| < \delta \Longrightarrow |f^{-1}(y) - f^{-1}(y_0)| < \epsilon).$$

Setze $x_0 := f^{-1}(y_0)$.

1. Fall: $\inf J < y_0 < \sup J$. Dies ist äquivalent zu $\inf I < x_0 < \sup I$. In Worten anschaulich: y_0 liegt nicht am Rand des Intervalls J.

Sei $\epsilon \in \mathbb{R}_{>0}$ gegeben. Setze

$$\tilde{\epsilon} := \min\{\epsilon, (x_0 - \inf I)/2, (\sup I - x_0)/2\} > 0$$

Dann $\inf I < x_0 - \tilde{\epsilon} < x_0 < x_0 + \tilde{\epsilon} < \sup I$. Setze $y_\pm := f(x_0 \pm \tilde{\epsilon})$ und $\delta := \min\{(y_+ - y_0), (y_0 - y_-)\}$. Also

$$y_- \leq y_0 - \delta < y_0 < y_0 + \delta \leq y_+.$$

Da f^{-1} ebenfalls streng monoton wachsend ist, gilt für alle $y \in (y_0 - \delta, y_0 + \delta)$:[6]

$$x_0 - \epsilon \leq x_0 - \tilde{\epsilon} = f^{-1}(y_-) \leq f^{-1}(y_0 - \delta)$$

$$\leq f^{-1}(y) \leq f^{-1}(y_0 + \delta) \leq f^{-1}(y_+) = x_0 + \tilde{\epsilon} \leq x_0 + \epsilon\,.$$

[5] Dies bedeutet: Wir betrachten nur den Fall, dass f streng monoton wachsend ist. Der Fall, dass f streng monoton fallend ist, folgt dann durch Vorzeichenwechsel.

[6] Das heißt: für alle $y \in J$ mit $y_0 - \delta < y < y_0 + \delta$

Also

$$|f^{-1}(y) - f^{-1}(y_0)| = |f^{-1}(y) - x_0| \leq \epsilon.$$

2. Fall: $\inf J = y_0$ oder $\sup J = y_0$. Anschaulich: y_0 ist am Rand des Intervalls J.

Geht ganz ähnlich. Man muss dabei nur eine der Seiten weglassen. □

BEISPIELE 2.8. Die folgenden Funktionen sind streng monoton, stetig und surjektiv und haben deswegen eine stetige Umkehrabbildung:

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f(x) := x^3 + x$$

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f(x) = x^{2n+1},\ n \in \mathbb{N}$$

$$f : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{\geq 0}, \quad f(x) = x^n,\ n \in \mathbb{N} \setminus \{0\}$$

Also sind die Wurzelfunktionen $\sqrt[n]{\ } : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{\geq 0}$ und $\sqrt[2n+1]{\ } : \mathbb{R} \longrightarrow \mathbb{R}$ stetig.

3. Stetigkeit von exp, cos und sin und Definition von log

LEMMA 3.1. Sei $\sum_{j=1}^{\infty} a_j$ eine absolut konvergierende Reihe, $a_j \in \mathbb{C}$. Dann gilt

$$\Big|\sum_{j=1}^{\infty} a_j\Big| \leq \sum_{j=1}^{\infty} |a_j|.$$

Beweis.

$$\Big|\sum_{j=1}^{\infty} a_j\Big| \overset{\text{Bsp. 1.14}}{=} \lim_{n\to\infty} \Big|\sum_{j=1}^{n} a_j\Big|$$

$$\overset{\substack{\text{Dreiecks-}\\ \text{Ungl.}}}{\leq} \lim_{n\to\infty} \sum_{j=1}^{n} |a_j| \overset{\text{def}}{=} \sum_{j=1}^{\infty} |a_j|$$

□

SATZ 3.2. Die komplexe Exponentialfunktion $\exp : \mathbb{C} \to \mathbb{C}$ ist stetig.

Beweis. Wir zeigen zunächst die Stetigkeit in 0. Dazu rechnen wir für $x \in \mathbb{C}$ mit $|z| \leq \frac{1}{2}$:

$$\begin{aligned} |\exp z - \exp 0| &= |\exp z - 1| = \Big| \sum_{j=1}^{\infty} \frac{1}{j!} z^j \Big| \\ &\overset{\text{Lem. 3.1}}{\leq} \sum_{j=1}^{\infty} \Big| \frac{1}{j!} z^j \Big| \leq \sum_{j=1}^{\infty} \Big| z^j \Big| = \sum_{j=1}^{\infty} |z|\,|z|^{j-1} \\ &\leq |z| \sum_{k=0}^{\infty} |z|^k \leq |z| \underbrace{\sum_{k=0}^{\infty} \left| \frac{1}{2} \right|^k}_{=2} = 2\,|z|. \end{aligned}$$

Für gegebenes $\epsilon \in \mathbb{R}_{>0}$ setze nun $\delta := \min\{\frac{1}{2}, \frac{\epsilon}{2}\}$. Dann gilt

$$|z - 0| < \delta \quad \Longrightarrow \quad |\exp z - \exp 0| < \epsilon.$$

Also ist $\exp : \mathbb{C} \to \mathbb{C}$ in 0 stetig.

Ist $f : \mathbb{C} \to \mathbb{C}$ in 0 stetig, dann auch $af \colon z \mapsto a \cdot f(z)$ für $a \in \mathbb{C}$. Wir setzen $f := \exp$ und $a := \exp z_0$. Somit ist

$$g : \mathbb{C} \to \mathbb{C}, \qquad z \mapsto g(z) := (\exp z_0) \cdot (\exp z) = \exp(z + z_0)$$

stetig in 0. Da $h : \mathbb{C} \to \mathbb{C}$, $h(z) := z - z_0$ stetig ist, folgt mit Lemma 1.13 die Stetigkeit von $g \circ h$ in z_0. Offensichtlich $\exp = g \circ h$. □

Es folgt daraus, dass auch die folgenden Funktionen $\mathbb{C} \to \mathbb{C}$ stetig sind:

$$z \mapsto \exp(iz), \qquad z \mapsto \exp(-iz)$$
$$z \mapsto \cos z = \frac{\exp iz + \exp(-iz)}{2}$$
$$z \mapsto \sin z = \frac{\exp iz - \exp(-iz)}{2i}$$

Durch Einschränkung auf $\mathbb{R}$ folgt auch, dass $\exp : \mathbb{R} \to \mathbb{R}$, $\cos : \mathbb{R} \to \mathbb{R}$ und $\sin : \mathbb{R} \to \mathbb{R}$ stetig sind. Aufgrund des Zwischenwertsatzes ist $\exp(\mathbb{R})$ ein Intervall. Wegen $\lim_{n\to\infty} \exp n = \infty$ und $\lim_{n\to\infty} \exp(-n) = 0$ folgt $\exp_{\#}(\mathbb{R}) = \mathbb{R}_{>0}$.

Korollar 3.3. Die reelle Exponentialfunktion $\exp : \mathbb{R} \longrightarrow \mathbb{R}_{>0}$ ist streng monoton, stetig und bijektiv. □

Definition 3.4. Die Umkehrfunktion von $\exp : \mathbb{R} \longrightarrow \mathbb{R}_{>0}$ nennt man den (natürlichen)[7] Logarithmus $\log : \mathbb{R}_{>0} \longrightarrow \mathbb{R}$.

Wegen Satz 2.7 des aktuellen Kapitels 4 ist der Logarithmus stetig. Die Funktionalgleichung der Exponentialfunktion, siehe Satz 3.1 in Kapitel 3 übersetzt sich nun in die Funktionalgleichung der Logarithmusfunktion:

$$\forall r, s \in \mathbb{R}_{>0} : \ \log(rs) = \log r + \log s.$$

[7] In der Mathematik wird der natürliche Logarithmus fast immer mit log bezeichnet. In technischen und vielen alltäglichen Bereichen wird der natürliche Logarithhmus mit dem Symbol ln bezeichnet und das Symbol log wird dann – anders als in unserer Vorlesung – für den Zehner-Logarithmus, also den Logarithmus zur Basis 10 genutzt. Im Mathematik-Studium wird aber nur der natürliche Logarithmus eine Rolle spielen. Deswegen wird der Zehner-Logarithmus hier weder definiert, noch mit einem Symbol versehen.

$\log 1 = 0$, $\log e = 1$

LEMMA 3.5. Sei $r \in \mathbb{R}_{>0}$ und $z \in \mathbb{C}$. Es gilt

$$r^z = \exp(z \log r)$$

in den folgenden Fällen.

(1) $r = e$ und $z \in \mathbb{C}$
(2) $r \in \mathbb{R}_{>0}$ und $z \in \mathbb{Q}$

Für alle anderen Kombination von $r \in \mathbb{R}_{>0}$ und $z \in \mathbb{C}$ wurde r^z noch gar nicht definiert.[8]

Beweis.
(1) folgt direkt aus Definition 3.4 in Kapitel 3.
(2) Die Aussage ist klar für $z = 1$. Dann folgt (2) ganz analog wie Lemma 3.3 in Kapitel 3.
□

DEFINITION 3.6. Für $r \in \mathbb{R}_{>0}$ und $z \in \mathbb{C}$ definiert man nun

$$r^z := \exp(z \log r)$$

Auf Grund des obigen Lemmas verallgemeinert dies also bisherige Definitionen von r^z in den beiden obigen Fällen.

$$r^{z+w} = r^z r^w \qquad (rs)^z = r^z s^z$$

$$\log r^s = \log \exp(s \log r) = s \log r$$

[8]Wir haben auch schon $(-1)^3$ und $(1+i)^{-2}$ definiert, aber es interessieren uns derzeit nur $r \in \mathbb{R}_{>0}$.

für $r, s \in \mathbb{R}_{>0}$, $z, w \in \mathbb{C}$. Es gilt auch

$$(r^s)^t = r^{st}$$

für $r \in \mathbb{R}_{>0}$ und $s, t \in \mathbb{R}$.

4. Die Kreizahl π und Periodizität von cos und sin.

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \underbrace{\frac{1}{(2n)!} x^{2n}}_{a_n}$$

Für $x \in (0, 2]$ und $n \geq 1$ gilt $a_n > 0$ und

$$\frac{a_{n+1}}{a_n} = \frac{x^2}{(2n+1)(2n+2)} \leq 1.$$

Also

$$\cos x = a_0 - a_1 + \underbrace{(a_2 - a_3)}_{\geq 0} + \underbrace{(a_4 - a_5)}_{\geq 0} \ldots \geq a_0 - a_1 = 1 - \frac{x^2}{2}$$

$$\cos x = a_0 - a_1 + a_2 + \underbrace{(-a_3 + a_4)}_{\leq 0} + \underbrace{(-a_5 + a_6)}_{\leq 0} \ldots \leq a_0 - a_1 + a_2 = 1 - \frac{x^2}{2} + \frac{x^4}{24}$$

Also für $x \in [0, 2]$

$$\boxed{1 - \tfrac{x^2}{2} \leq \cos x \leq 1 - \tfrac{x^2}{2} + \tfrac{x^4}{24}}$$

Ganz ähnlich für $0 \leq x \leq 4$

$$\boxed{x - \tfrac{x^3}{6} \leq \sin x \leq x.}$$

$$\cos 0 = 1, \qquad \cos 2 \leq 1 - 2 + \frac{2}{3} = -\frac{1}{3} < 0.$$

Aus dem Zwischenwertsatz folgt:

$$A := \{x \in [0,2] \mid \cos x = 0\} \neq \emptyset.$$

Setze $\alpha := \inf A$. Zu jedem $n \in \mathbb{N}$ wähle ein $x_n \in A$ mit $x_n \leq \alpha + \frac{1}{n}$. Dann gilt $x_n \to \alpha$ für $n \to \infty$ und

$$\cos \alpha = \cos(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} \underbrace{\cos x_n}_{0} = 0.$$

$\pi := 2\alpha \in [0,4]$.

Definition 4.1. Die Kreiszahl π ist definiert als das Doppelte der ersten positiven Nullstelle der reellen Kosinusfunktion.

$$\pi = 3,14159265...$$

Es folgt aus den bisher bewiesenen Eigenschaften in Abschnitt 3.2 in Kapitel 3 nun durch einfache Rechnung:

- $\sin \pi/2 = 1$, da $\sin \pi/2 \geq \frac{\pi}{2}\left(1 - \left(\frac{\pi}{2}\right)^2/6\right) > 0$ und $\sin^2 \alpha + \cos^2 \alpha = 1$,
- $\sin(x + (\pi/2)) = \cos x$, $\cos(x + (\pi/2)) = -\sin x$, (Additionstheoreme)
- $\cos(2\pi + x) = -\cos(\pi + x) = \cos x$, $\sin(2\pi + x) = -\sin(\pi + x) = \sin x$,
- $e^{2\pi i} = \cos 2\pi + i \sin 2\pi = 1$

Aus

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

folgt $\cos x > \cos y$ für $0 \le x < y \le 2$.

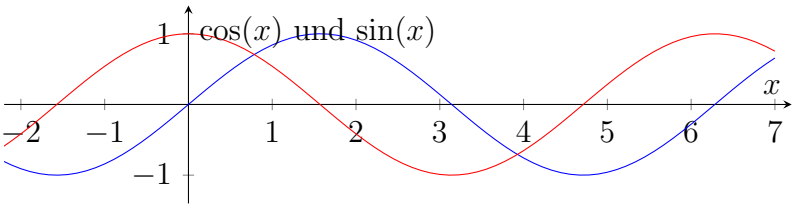


Abbildung 6: Die Graphen von cos und sin

Fr 18.1.

5. Metrische Räume und Grundbegriffe der Topologie

Der Begriff der „Stetigkeit“ gehört zum mathematischen Teilgebiet mit dem Namen „Topologie“, so wie der Begriff „lineare Abbildungen“ zum Teilgebiet „Lineare Algebra“ gehört. Im Laufe Ihres Studiums werden Sie immer wieder von stetigen Abbildungen hören, und die topologischen Begriffe werden dann immer wichtiger.

Dies betrifft zum einen die Mathematiker unter Ihnen, und zwar sowohl diejenigen die sich an Reiner Mathematik orientieren (Topologische Räume sind da fast schon omnipräsent), als auch diejenigen, die sich an Angewandter Mathematik orientieren (Beschreibung von Flächen im Raum, Mannigfaltigkeiten, partielle Differentialgleichung, Modellierung von

physikalischen Systemen). Und es trifft auch auf Physiker zu. Zum Beispiel benötigt eine präzise Formulierung selbst einfacher quantenmechanischer Systeme wie dem Wasserstoff-Atom solche topologischen Grundbegriffe (Energie-, Orts-, Impulsoperatoren sind „dicht definierte“ „unstetige“ „abgeschlossene“ Operatoren)[9] und die zugrundeliegenden Räume der klassischen Mechanik, der Allgemeinen Relativitätstheorie und der Elektrodynamik sind Mannigfaltigkeiten und damit mit einer Topologie versehen. Natürlich trifft vieles davon auch auf Computational Science und Nanoscience zu.

In diesem Abschnitt lernen wir nun einige Grundbegriffe der Topologie kennen: Stetigkeit in einem etwas allgemeineren Sinn, offene Mengen, abgeschlossene Mengen, Häufungspunkte und Umgebungen. Diese Begriffe werden (in noch allgemeinerer Form) Ihnen häufig im Studium begegnen. Wir wollen aber noch nicht das zentrale Objekt der Topologie, nämlich „topologische Räume“ einführen, sondern uns zunächst auf metrische Räume spezialisieren.

DEFINITION 5.1. Ein metrischer Raum ist ein Paar (M, d), wobei M eine Menge ist und $d : M \times M \longrightarrow \mathbb{R}_{\geq 0}$ eine Abbildung ist mit folgenden Eigenschaften, die für alle $x, y, z \in M$ gelten sollen:

(a) $d(x, y) = 0 \iff x = y$ (Definitheit)
(b) $d(x, y) = d(y, x)$ (Symmetrie)
(c) $d(x, y) \leq d(x, z) + d(z, y)$ (Dreiecks-Ungleichung)

Man nennt d eine Metrik auf M.

BEISPIELE 5.2.

[9] Es ist derzeit nicht wichtig, dass Sie diese Begriffe kennen, wichtig ist: es sind Begriffe der Topologie.

(1) Sei $M \subset \mathbb{R}$ oder $M \subset \mathbb{C}$. Für $z, w \in M$ definieren wir

$$d(z,w) := |z-w| = \sqrt{(\mathbf{Re}\,(z-w))^2 + (\mathbf{Im}\,(z-w))^2}.$$

(2) Sei $M \subset \mathbb{R}^n$. Zu $x = (x_1, x_2, \ldots, x_n) \in M \subset \mathbb{R}^n$ und $y = (y_1, y_2, \ldots, y_n) \in M \subset \mathbb{R}^n$ definieren wir den euklidischen Abstand als

$$d_{\text{eukl}}(x,y) := \|x-y\| = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_n-y_n)^2}.$$

In beiden Beispielen sind (a) und (b) klar. Für (1) wurde (c) schon diskutiert. Die Ungleichung (c) für (2) wird in der Linearen Algebra I (und dem Kurs „Mathematische Methoden“ für Physiker) gezeigt.

Beispiel 5.3. Ist (M,d) ein metrischer Raum und $N \subset M$. Dann ist $(N, d|_{N\times N})$ ebenfalls ein metrischer Raum. Wir nennen $d_{N\times N}$ die von (M,d) auf N induzierte Metrik.

Schreibweise: Sei $f : M \longrightarrow \widetilde{M}$ eine Abbildung und sind (M,d) und $(\widetilde{M}, \widetilde{d})$ metrische Räume, dann schreiben wir hierfür oft $f : (M,d) \longrightarrow (\widetilde{M}, \widetilde{d})$.

Ist eine Abbildung $f : (M,d) \longrightarrow (\widetilde{M}, \widetilde{d})$ bijektiv und gilt für alle $x, y \in M$: $d(x,y) = \widetilde{d}(f(x), f(y))$, dann nennt man die Abbildung eine Isometrie. Man sieht sofort, dass $(x,y) \mapsto x + iy$ eine Isometrie von $(\mathbb{R}^2, d_{\text{eukl}})$ nach $(\mathbb{C}, d)$ ist, wobei d die Metrik im obigen Beispiel ist. Da wir oft $\mathbb{R}^2$ mit$\mathbb{C}$ identifizieren, schreiben wir ebenfalls d_{eukl} für diese Metrik auf $\mathbb{C}$.

Definition 5.4 (ϵ-δ-Kriterium für Stetigkeit für Abbildungen zwischen metrischen Räumen). $f : (M,d) \longrightarrow (\widetilde{M}, \widetilde{d})$ ist stetig in $x_0 \in M$, falls gilt:

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists \delta \in \mathbb{R}_{>0} : \forall x \in M : (d(x,x_0) < \delta \Longrightarrow \tilde{d}(f(x), f(x_0)) < \epsilon). \tag{5.5}$$

f stetig $\iff$ f stetig in allen $x_0 \in M$.

Im Fall $M \subset \mathbb{C}$, $d(z,w) = |z-w|$ stimmt diese Definition offensichtlich mit Definition 1.5 überein.

DEFINITION 5.6. Sei (M,d) ein metrischer Raum, $x \in M$, $r \in \mathbb{R}_{>0}$. Der (offene) Ball von Radius r um x in (M,d) ist die Menge

$$B_r(x) := \{y \in M \mid d(y,x) < r\}.$$

Eine Teilmenge $U \subset M$ heißt Umgebungvon x in (M,d), falls es ein $r \in \mathbb{R}_{>0}$ gibt, so dass $B_r(x) \subset U$.

DEFINITION 5.7. Eine Teilmenge O von M heißt offen in (M,d), wenn für jeden Punkt $x \in O$ gilt, dass O eine Umgebung von x in (M,d) ist. Eine Teilmenge A von M heißt abgeschlossen in (M,d), falls $M \smallsetminus A$ offen ist.

BEISPIELE 5.8.

(a) Die leere Menge und M sind offen in (M,d). Diese beiden Mengen sind auch abgeschlossen.

(b) $O \subset \mathbb{R}^2 \cong \mathbb{C}$

Bild einer offenen Menge O in $\mathbb{R}^2$, darin ein x und ein geeigneter Ball $B_r(x)$.

(c) Alle offenen Intervalle in $\mathbb{R}$ sind offen in $(\mathbb{R}, d_{\text{eukl}})$. Alle abgeschlossenen Intervalle in $\mathbb{R}$ sind abgeschlossen in $(\mathbb{R}, d_{\text{eukl}})$.

(d) Die Menge $(0,1) \times (0,1)$ ist offen in $(\mathbb{R}^2, d_{\text{eukl}})$

(e) Für alle $r \in \mathbb{R}_{>0}$ und alle $x \in M$ ist $B_r(x)$ eine offene Teilmenge von (M, d). Begründung: Sei $y \in B_r(x)$, $\rho := d(x, y)$. Dann gilt $B_{r-\rho}(y) \subset B_r(x)$ (Dreiecks-Ungleichung).

Zeichnung von $B_{r-\rho}(y) \subset B_r(x)$ in $\mathbb{R}^2$.

(f) Ist U Umgebung von x und $U \subset V \subset M$, dann ist V ebenfalls Umgebung von x.

Lemma 5.9. Sei $x \in M$ und $U \subset M$. Dann ist U eine Umgebung von x, genau dann wenn es ein $O \subset M$ gibt, das

- in (M, d) offen ist, und
- so dass $x \in O \subset U$.

Solch ein O ist also eine offene Umgebung von x.

Beweis.
„$\Longrightarrow$“: klar, da offene Bälle offen sind.
„$\Longleftarrow$“: Wenn O offen ist und $x \in O$, dann ist O eine Umgebung von x. Aus $O \subset U$ ergibt sich, dass U Umgebung von x ist. □

Proposition 5.10.

(a) Sei $x \in M$ und seien U_1 und U_2 Umgebungen von x in (M, d). Dann ist auch $U_1 \cap U_2$ eine Umgebung von x.
(b) Sind O_1 und O_2 offene Teilmengen von (M, d), so ist $O_1 \cap O_2$ offen.
(c) Sei $(O_j)_{j \in I}$ eine Familie offener Teilmengen von (M, d), dann ist auch

$$\bigcup_{j \in I} O_j$$

eine offene Teilmenge.

Beweis.
Zu (a): Da U_1 und U_2 Umgebungen von x sind, gibt es $r_1, r_2 \in \mathbb{R}_{>0}$ mit $B_{r_1}(x) \subset U_1$ und $B_{r_2}(x) \subset U_2$. Für $r := \min\{r_1, r_2\}$ gilt $B_r(x) \subset U_1 \cap U_2$.

Zu (b): Gegeben sei ein $x \in O_1 \cap O_2$ ist. Da O_1 und O_2 offen sind, sind sie Umgebungen von x. Aus (a) folgt, dass $O_1 \cap O_2$ ebenfalls Umgebung von x ist. Da dies für jedes $x \in O_1 \cap O_2$ gilt, ist $O_1 \cap O_2$ offen.

Zu (c): Sei $x \in \bigcup_{j\in I} O_j$. Fixiere ein $k \in I$ mit $x \in O_k$. Da O_k offen ist, ist O_k eine Umgebung von x in (M, d), und damit auch $\bigcup_{j\in I} O_j \supset O_k$. Also ist $\bigcup_{j\in I} O_j$ eine Umgebung von jedem $x \in \bigcup_{j\in I} O_j$, also offen. □

!ACHTUNG!.

- Für $n \in \mathbb{N} \setminus \{0\}$ ist $(-1/n, 1/n)$ offen in $\mathbb{R}$, aber nicht offen in $\mathbb{C}$.
- Proposition 5.10 besagt: endliche Durchschnitte offener Mengen sind offen und beliebige (also auch unendliche!) Vereiningung offener Mengen sind auch wieder offen. Allerdings ist ein unendlicher Durchschnitt offener Mengen im allgemeinen nicht mehr offen: $\bigcap_{n\in\mathbb{N}}(-1/n, 1/n) = \{0\}$ ist nicht offen!
- „Abgeschlossen“ ist nicht das Gegenteil zu „offen“.

Proposition 5.11 (Umgebungskriterium für Stetigkeit)**.** Gegeben sei eine Abbildung $f : (M, d) \longrightarrow (\widetilde{M}, \widetilde{d})$ und ein $x_0 \in M$. Diese Abbildung ist stetig in x_0 genau dann, wenn gilt:

für alle Umgebungen U von $f(x_0)$ in $(\widetilde{M}, \widetilde{d})$ gilt:
$f^{\#}(U)$ ist eine Umgebung von x_0 in (M, d).

Beweis. Die umrahmte Aussage ist eine Umformulierung des ϵ-δ-Kriteriums unter Nutzung des oben definierten Begriffs „Umgebung“. □

Kurz kann man hierfür sagen: Urbilder von Umgebungen sind Umgebungen.

Proposition 5.12. Eine Funktion $f : (M, d) \longrightarrow (\widetilde{M}, \widetilde{d})$ ist genau dann stetig, wenn für alle offenen Teilmengen O von $(\widetilde{M}, \widetilde{d})$ gilt: $f^{\#}(O)$ ist offen in (M, d).

Kurz kann man hierfür sagen: Urbilder offener Mengen sind offen.

Beweis. Folgt direkt aus dem Umgebungskriterium für Stetigkeit (Proposition 5.11).

Definition 5.13. Eine Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert in (M, d) gegen $x_0 \in M$, falls für jede Umgebung U von x_0 gilt:

$$\exists n_0 \in \mathbb{N} : \forall n \in \mathbb{N}_{\geq n_0} : a_n \in U.$$

Die ist äquivalent zu

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists n_0 \in \mathbb{N} : \forall n \in \mathbb{N} : \big(n \geq n_0 \Longrightarrow d(a_n, x_0) < \epsilon\big),$$

und dies stimmt mit Definition 1.4 überein, falls $d(a_n, x_0) = |a_n - x_0|$. Die Folge konvergiert in (M, d), wenn es in M solch ein x_0 gibt.

Diese Definition verallgemeinert die Konvergenz von Folgen in $\mathbb{R}$ und $\mathbb{C}$. Man definiert $\lim_{n\to\infty}$ und die Folgenstetigkeit gleich wie zuvor. Folgenstetigkeit und Stetigkeit sind wieder äquivalente Eigenschaften. Um dies zu beweisen, ersetze man im Beweis von Lem-

Mi 23.1. ma 1.7 $|x - x_0| < \delta$ durch $d(x, x_0) < \delta$ und so weiter.

BEMERKUNG 5.14. $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$. In der Diskussion vor Proposition 1.30 in Kapitel 3 haben wir eine bijektive Abbildung $\varphi : [-1, 1] \longrightarrow \overline{\mathbb{R}}$ definiert,

$$\varphi(x) = \begin{cases} \frac{1}{1-x} - \frac{1}{1+x} & \text{für } x \in (-1, 1) \\ \infty & \text{für } x = 1 \\ -\infty & \text{für } x = -1 \end{cases}$$

Graph der Funktion φ.

Wir versehen nun $\overline{\mathbb{R}}$ mit der Metrik

$$d_{\overline{\mathbb{R}}}(x, y) := \left|\varphi^{-1}(x) - \varphi^{-1}(y)\right|.$$

Dann ist $(\overline{\mathbb{R}}, d_{\overline{\mathbb{R}}})$ ein metrischer Raum. Für eine reell-wertige Folge $(a_n)_{n\in\mathbb{N}}$ und ein $a \in \overline{\mathbb{R}}$ gilt

Für	im klassischen Sinn		im Sinn metrischer Räume
$a \in \mathbb{R}$	$\lim_{n\to\infty} a_n = a$ im eigentlichen Sinn	$\Longleftrightarrow$	$\lim_{n\to\infty} a_n = a$ im metrischen Raum $(\overline{\mathbb{R}}, d_{\overline{\mathbb{R}}})$
$a = \pm\infty$	$\lim_{n\to\infty} a_n = a$ im uneigentlichen Sinn	$\Longleftrightarrow$	$\lim_{n\to\infty} a_n = a$ im metrischen Raum $(\overline{\mathbb{R}}, d_{\overline{\mathbb{R}}})$

Proposition 1.30 aus Kapitel 3, folgt nun unmittelbar.

!ACHTUNG!. Wir wiederholen nochmals erlaubte und unerlaubte Sprechweisen klären an Hand des Beispiels $a_n := n^2$.

Erlaubt: $(a_n)_{n\in\mathbb{N}}$ konvergiert uneigentlich gegen ∞.

$(a_n)_{n\in\mathbb{N}}$ konvergiert in $(\overline{\mathbb{R}}, d_{\overline{\mathbb{R}}})$ gegen ∞.

Nicht erlaubt: $(a_n)_{n\in\mathbb{N}}$ konvergiert.

Schreibweise: Wenn aus dem Kontext heraus klar ist, welche Metrik gemeint ist, dann schreibt man oft einfach M statt (M, d). Teilmengen $X \subset M$ tragen die induzierte Metrik $d|_{X\times X}$, wenn nichts anderes angegeben.

DEFINITION 5.15. Sei X Teilmenge eines metrischen Raums (M, d). Wir nennen $x_0 \in M$ einen Häufungspunkt von X in (M, d), falls jede Umgebung von x_0 unendlich viele Punkte von X enthält.

BEISPIELE 5.16.

(a) Endliche Mengen haben keine Häufungspunkte.
(b) Die Mengen $\mathbb{N}$ und $\mathbb{Z}$ haben keine Häufungspunkte in $\mathbb{R}$ und $\mathbb{C}$.
(c) Die Menge aller Häufungspunkte von $\mathbb{Q}$ in $\mathbb{C}$ ist $\mathbb{R}$.
(d) Seien $a, b \in \mathbb{R}$, $a < b$. Die Menge aller Häufungspunkte von (a, b) in $\mathbb{R}$ (oder in $\mathbb{C}$) ist die Menge $[a, b]$.

LEMMA 5.17. Sei X Teilmenge eines metrischen Raums (M, d), und sei $y \in M$. Äquivalent sind dann die folgenden Aussagen:

(1) y ist Häufungspunkt von X
(2) Für jede Umgebung U von y gilt $U \cap X \setminus \{y\} \neq \emptyset$
(3) Es gibt eine Folge (a_n) in $X \setminus \{y\}$ mit $\lim_{n\to\infty} a_n = y$
(4) Es gibt eine Folge in $X \setminus \{y\}$, so dass y Häufungspunkt dieser Folge ist.

Beweis.

„(1)$\Longrightarrow$(2)“: klar.

„(2)$\Longrightarrow$(3)“: $B_{1/n}(y) \cap X$ besitzt ein Element $a_n \neq y$. Dann $d(a_n, y) \leq 1/n \to 0$.

„(3)$\Longrightarrow$(1)“: Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge wie in (2), und U eine Umgebung von y. Es gibt ein $n_0 \in \mathbb{N}$, so dass $\forall n \in \mathbb{N}_{\geq n_0} : a_n \in U$. Wenn $P := \{d(a_n, y) \mid n \in \mathbb{N}_{\geq n_0}\}$ nur endlich viele Elemente enthält, dann ist $\inf\{d(a_n, y) \mid n \in \mathbb{N}\} > 0$. Widerspruch. Somit hat P unendlich viele Elemente und deswegen auch $U \cap X$.

„(3)$\Longleftrightarrow$(4)“: siehe Definition 1.22 in Kapitel 3. □

!ACHTUNG!. Die Definitionen von „Häufungspunkt einer Menge“ (siehe oben) und „Häufungspunkt einer Folge“ (Definition 1.22 in Kapitel 3) sind zwar verwandt, aber doch verschieden, wie das folgende Beispiel zeigt. Die reell-wertige Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n := 1$ hat den Häufungspunkt 1, aber die Menge $\{a_n \mid n \in \mathbb{N}\}$ hat keine Häufungspunkte.

Bemerkung 5.18 (Aufgabe 4 auf Blatt 12). Eine Teilmenge eines metrischen Raumes (M, d) ist abgeschlossen, genau dann wenn sie alle ihre Häufungspunkte enthält.

Definition 5.19. Wir nennen einen metrischen Raum (M, d) folgenkompakt, wenn jede Folge in M eine in (M, d) konvergente Teilfolge besitzt. Eine Teilmenge $X \subset M$ nennt man folgenkompakt, wenn $(X, d|_{X\times X})$ folgenkompakt ist.[10].

Lemma 5.20. Folgenkompakte Teilmengen eines metrischen Raumes sind abgeschlossen.

Beweis. Sei X eine folgenkompakte Teilmenge von (M, d) und y ein Häufungspunkt von X. Dann gibt es eine Folge $(a_n)_{n\in\mathbb{N}}$ in X mit $\lim_{n\to\infty} a_n = y$. Da X folgenkompakt ist, gibt es eine Teilfolge $(a_{f(k)})_{k\in\mathbb{N}}$ in X, $f : \mathbb{N} \to \mathbb{N}$ streng monoton wachsend, mit $z := \lim_{k\to\infty} a_{f(k)} \in X$. Diese Teilfolge konvergiert auch in M, also haben wir $y = z$. Somit enthält X jeden Häufungspunkt und ist somit abgeschlossen (Bemerkung 5.18). □

[10]Achtung: Der Grenzwert muss dann in X liegen!

BEISPIEL 5.21. Sei I ein Interval mit der von d_{eukl} induzierten Metrik. Dann ist I folgenkompakt genau dann, wenn I abgeschlossen und beschränkt ist. Denn:

- Ist I folgenkompakt, dann ist I abgeschlossen in $(\mathbb{R}, d_{\text{eukl}})$ aufgrund des obigen Lemmas.
- Ist I folgenkompakt, dann ist I beschränkt. Denn wenn I nicht nach oben (bzw. unten) beschränkt ist, dann gibt es eine Folge, die uneigentlich gegen $+\infty$ (bzw. $-\infty$) konvergiert. Diese Folge hat keine konvergente Teilfolge. [11]
- Ist I beschränkt, dann besitzt jede Folge in I aufgrund des Satzes von Bolzano-Weierstraß (Satz 1.35 in Kapitel 3) einen Häufungspunkt, also eine in $\mathbb{R}$ konvergente Teilfolge. Ist I zusätzlich abgeschlossen, dann ist der Grenzwert in I. Somit ist I folgenkompakt.

SATZ 5.22. Seien (M, d) und $(\tilde{M}, \tilde{d})$ metrische Räume. Ist (M, d) folgenkompakt und $f : (M, d) \to (\tilde{M}, \tilde{d})$ stetig, dann ist $B := f_{\#}(M)$ mit der Metrik $\tilde{d}|_{B\times B}$ ebenfalls folgenkompakt.

Beweis. Sei $(y_n)_{n\in\mathbb{N}}$ eine Folge in $f_{\#}(M)$. Wähle $x_n \in M$ mit $f(x_n) = y_n$ und danach eine konvergente Teilfolge $(x_{h(k)})_{k\in\mathbb{N}}$, $h : \mathbb{N} \to \mathbb{N}$ streng monoton wachsend. Dann gilt

$$\lim_{k\to\infty} y_{h(k)} = \lim_{k\to\infty} f(x_{h(k)}) = f\Big(\underbrace{\lim_{k\to\infty} x_{h(k)}}_{\in M}\Big) \in f_{\#}(M).$$

Also ist $f_{\#}(M)$ folgenkompakt. □

[11] Alternative Beweise: a) Übungsblatt 13, Aufgabe 4. Zweiter Beweis: Ist (I, d_{eukl}) folgenkompakt,d ann ist auch $(I, d_{\overline{R}}|_{I\times I})$ folgenkompakt. Also ist I abgeschlossen in $\overline{\mathbb{R}}$. Somit sind $-\infty$ und $+\infty$ keine Häufungspunkte von I in $(\overline{\mathbb{R}}, d_{\overline{\mathbb{R}}})$, was bedeutet, dass I beschränkt ist.

FOLGERUNG 5.23. Seien $a, b \in \mathbb{R}$, $a \leq b$ und $f : [a, b] \to \mathbb{R}$ stetig. Dann ist $f_{\#}([a, b])$ ein abgeschlossenes beschränktes Intervall. Es existieren also $x_1, x_2 \in [a, b]$ mit

$$\forall x \in [a, b] : \ f(x_1) \leq f(x) \leq f(x_2).$$

Die Folgerung folgt direkt aus dem obigen Beispiel und dem obigen Satz.

6. Grenzwerte von Funktionen

LEMMA 6.1. Seien (M, d) und $(N, \widetilde{d})$ metrische Räume, $D \subset M$, und $x_0 \in M$ ein Häufungspunkt von D. Zu jeder Abbildung $f : D \to N$ gibt es höchstens eine Abbildung $F : D \cup \{x_0\} \to N$, so dass

- $F|_{D \setminus \{x_0\}} = f|_{D \setminus \{x_0\}}$
- und so dass F stetig in x_0 ist.

DEFINITION 6.2. Wenn solch ein F existiert, so sagen wir der Grenzwert oder Limes

$$\lim_{x \to x_0} f(s)$$

existiert. Wir nennen $F(x_0)$ den Grenzwert oder Limes von f in x_0 und schreiben

$$\lim_{x \to x_0} f(x) := F(x_0).$$

BEMERKUNG 6.3. Es sind die Fälle $x_0 \in D$ und $x_0 \notin D$ zugelassen. Im Fall $x_0 \in D$ muss nicht $f(x_0) = F(x_0)$ gelten.

Beweis des Lemmas. Wähle eine Folge $(a_n)_{n\in\mathbb{N}}$ in $D \setminus \{x_0\}$ mit $a_n \to x_0$. Wenn ein F wie oben existiert, dann gilt

$$F(x_0) = F(\lim_{n\to\infty} a_n) = \lim_{n\to\infty} F(a_n) = \lim_{n\to\infty} f(a_n).$$

Es gibt also höchstens einen passenden Wert für $F(x_0)$. Auf $D \setminus \{x_0\}$ ist F sowieso schon bestimmt. Also kann es höchstens eine solche Abbldung F geben. □

Definition 6.4. Sei $D \subset \mathbb{R}$, und sei $x_0 \in \mathbb{R}$ ein Häufungspunkt von $D \cap [x_0, \infty)$. Dann definiert man den rechtsseitigen Grenzwert/Limes als

$$\lim_{x\searrow x_0} f(x) := \lim_{x\to x_0} (f|_{D\cap[x_0,\infty)}).$$

Wenn wir $[x_0, \infty)$ durch $(-\infty, x_0]$ ersetzen, erhalten wir die Definition des linksseitigen Grenzwerts/Limes

$$\lim_{x\nearrow x_0} f(x).$$

Für $x_0 \in D \subset \mathbb{R}$, x_0 Häufungspunkt von $D \cap [x_0, \infty)$ und $D \cap (-\infty, x_0]$:

$$\begin{aligned} & f : D \to \mathbb{R} \text{ ist stetig in } x_0 \\ \iff & \lim_{x\to x_0} f(x) \text{ existiert und } \lim_{x\to x_0} f(x) = f(x_0) \\ \iff & \lim_{x\searrow x_0} f(x) \text{ und } \lim_{x\nearrow x_0} f(x) \text{ existieren und } \lim_{x\searrow x_0} f(x) = \lim_{x\nearrow x_0} f(x) = f(x_0) \end{aligned}$$

Beispiel 6.5. Definiere $f : \mathbb{R} \longrightarrow \mathbb{R}$.

$$f(x) := \begin{cases} 1 & \text{für } x > 0 \\ 0 & \text{für } x = 0 \\ -1 & \text{für } x < 0 \end{cases}$$

$\lim_{x\to 0} f(x)$ existiert nicht. Hingegen existieren $\lim_{x\searrow 0} f(x) = 1$ und $\lim_{x\nearrow 0} f(x) = -1$.

BEMERKUNG 6.6. Wir versehen $\mathbb{N} \cup \{\infty\}$ mit der Metrik[12]

$$d_{\text{inv}}(x, y) = \left| \frac{1}{x+1} - \frac{1}{y+1} \right|,$$

wobei wir in diesem[13] Zusammenhang $\frac{1}{\infty} := 0$ setzen. Dieses d_{inv} ist eine Metrik auf $\mathbb{N} \cup \{\infty\}$. Sei (a_n) eine Folge in (M, d), und sei $f : \mathbb{N} \longrightarrow M$ die Abbildung mit $f(n) = a_n$. Dann gilt

$$\begin{aligned} & \lim_{n\to\infty} a_n = a \text{ im Sinne von Kap. 3 Abschnitt 1.1} \\ \iff \quad & \lim_{n\to\infty} f(n) = a \text{ im Sinne des aktuellen Abschnitts} \end{aligned}$$

[12]Die folgende Metrik ist nur ein Beispiel einer Metrik, die verwendet werden kann. Die Einchränung von $d_{\overline{R}}$ auf $(\mathbb{N} \cup \{\infty\}) \times (\mathbb{N} \cup \{\infty\})$ ist zum Beispiel ebenfalls eine Metrik, für die die folgenden Aussagen richtig sind.

[13]Wenn wir es brauchen, wird $\frac{1}{\infty}$ jedesmal neu definiert, denn man kann diese Definition nicht so machen, dass alle ülichen Rechenregeln gelten. Man kommt dann immer wieder zu Fragen wie: $\frac{1}{0} = \infty$ oder $\frac{1}{0} = -\infty$. Ist $\infty - \infty = 0$ oder $\infty - \infty = 17$? „Beweis“ der letzten Aussage: $\infty + 17 = \infty$. Löse diese Gleichung nach 17 auf. Beim Rechnen mit ∞ funktioniert eben alles mögliche nicht mehr richtig.

KAPITEL 5

Differential-Rechnung für Funktionen einer Veränderlichen

1. Definition und elementare Eigenschaften

Ab sofort nutzen wir $\epsilon > 0$ oft im Sinne von $\epsilon \in \mathbb{R}_{>0}$, falls aus dem Kontext heraus klar ist, dass ϵ eine reelle Zahl ist. Wir betrachten in den folgenden Abschnitten nur Funktionen der Form $f : D \longrightarrow \mathbb{R}$, $D \subset \mathbb{R}$.

DEFINITION 1.1. Sei $D \subset \mathbb{R}$ offen, $x_0 \in D$.[1] Eine Funktion $f : D \longrightarrow \mathbb{R}$ nennt man differenzierbar in x_0, wenn der Grenzwert

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Wenn er existiert, so nennt man

$$f'(x_0) := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

die Ableitung von f an der Stelle x_0. Die Funktion f ist differenzierbar, wenn sie in jedem $x_0 \in D$ differenzierbar ist. Den Quotient

$$\frac{f(x) - f(x_0)}{x - x_0}$$

nennt man den Differenzenquotient.

[1] Damit diese Definition sinnvoll ist, reicht hier eigentlich bereits, dass D eine Umgebung von x_0, oder sogar nur, dass x_0 ein Häufungspunkt von D ist.

Graph einer Funktion f mit einem Dreieck, das den Differenzenquotient andeutet.

Interpretation: Die Zahl $f'(x_0)$ gibt an, wie stark der Graph der Funktion an der Stelle $(x_0, f(x_0))$ ansteigt.

Alternative Bezeichnungen:

$$f'(x_0) = \frac{df}{dx}(x_0) = \frac{\partial f}{\partial x}(x_0) = Df(x_0)$$

$$\begin{aligned} f \text{ differenzierbar in } x_0 &\iff \text{Die Ableitung von } f \text{ existiert in } x_0 \\ &\iff f'(x_0) \text{ existiert} \end{aligned}$$

Beispiele 1.2.

(1) Konstante Funktionen $\mathbb{R} \to \mathbb{R}$, $x \mapsto a$ sind differenzierbar und $f'(x_0) = 0$ für alle $x_0 \in \mathbb{R}$.

(2) Die Identität $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \longrightarrow \mathbb{R}$, $x \mapsto x$ ist differenzierbar und $f'(x_0) = 1$ für alle $x_0 \in \mathbb{R}$.

(3) Differenzierbarkeit ist eine lokale Eigenschaft: Ist U eine Umgebung von x_0, enthalten in den Definitionsbereichen von f und $\widetilde{f}$, und gilt $f|_U = \widetilde{f}|_U$, dann ist f differenzierbar in x_0 genau dann, wenn $\widetilde{f}$ in x_0 differenzierbar ist. Dann gilt $f'(x_0) = \widetilde{f}'(x_0)$.

(4) Die Funktion $f : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{0\}$, $x \mapsto x^{-1}$ ist differenzierbar und es gilt

$$f'(x_0) = -\frac{1}{x_0^2}.$$

Denn:

$$\frac{\frac{1}{x}-\frac{1}{x_0}}{x-x_0}=\frac{x_0-x}{xx_0(x-x_0)}=\frac{-1}{xx_0}\to-\frac{1}{x_0^2}\qquad\text{für } x\to x_0$$

Fr 25.1.

Umformungen:

$$\begin{aligned}
&f'(x_0)\text{ existiert}\\
\iff&\forall\epsilon>0:\ \exists\delta>0:\ \forall x\in D\setminus\{x_0\}:\ \left(|x-x_0|<\delta\Longrightarrow\left|\frac{f(x)-f(x_0)}{x-x_0}-f'(x_0)\right|<\epsilon\right)\\
\iff&\forall\epsilon>0:\ \exists\delta>0:\ \forall x\in D:\\
&\Big(|x-x_0|<\delta\Longrightarrow|f(x)-f(x_0)-f'(x_0)(x-x_0)|\le\epsilon\,|x-x_0|\Big)
\end{aligned}$$

Interpretation:

Differenzierbarkeit = „gute“ Approximierbarkeit durch eine lineare Funktion

Satz 1.3. Seien D und f wie oben. Ist f differenzierbar in x_0, so ist f stetig in x_0.

Die Umkehrung gilt nicht, siehe Übungen.

Beweis. Definiere

$$g(x):=\frac{f(x)-f(x_0)}{x-x_0}\quad\text{für } x\neq x_0$$

$$\lim_{x\to x_0}((x-x_0)g(x))=(\lim_{x\to x_0}(x-x_0))(\lim_{x\to x_0}g(x))=0\cdot f'(x_0)=0$$

[2]

$$f(x)=f(x_0)+(x-x_0)g(x)\quad\text{für } x\neq x_0$$

[2]Hier haben wir benutzt: sind $g:D\longrightarrow\mathbb{R}$ und $h:D\longrightarrow\mathbb{R}$ Funktionen, x_0 ein Häufungspunkt von D und existieren $\lim_{x\to x_0}g(x)$ und $\lim_{x\to x_0}h(x)$, dann existiert auch $\lim_{x\to x_0}(g(x)h(x))$ und es gilt

$$\lim_{x\to x_0}(g(x)h(x))=\left(\lim_{x\to x_0}g(x)\right)\left(\lim_{x\to x_0}h(x)\right).$$

Also

$$\lim_{x\to x_0} f(x) = f(x_0) + 0 \cdot f'(x_0) = f(x_0).$$

Somit ist f stetig in x_0. □

Regeln 1.4. Im folgenden sei D offen in $\mathbb{R}$, $f, f_j, g : D \longrightarrow \mathbb{R}$.

(1) Sind f_j differenzierbar in x_0, $a_j \in \mathbb{R}$ für $j \in \{1, 2, \ldots, r\}$, dann gilt

$$\left(\sum_{j=1}^{r} a_j f_j\right)'(x_0) = \sum_{j=1}^{r} a_j f_j'(x_0).$$

Solche Aussagen sind immer so zu verstehen, dass alle angegebenen Ableitungen in x_0 existieren und die obige Gleichung für die so erhaltenen Zahlen gilt.

(2) Produktregel.

Sind f und g differenzierbar in x_0, dann gilt

$$(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

Die Existenz der Ableitung $(f \cdot g)'(x_0)$ ist wieder eine implizite Teilaussage der obigen Formelzeile.

Begründung:

$$\frac{(f \cdot g)(x) - (f \cdot g)(x_0)}{x - x_0} = \frac{f(x)(g(x) - g(x_0))}{x - x_0} + \frac{(f(x) - f(x_0))g(x_0)}{x - x_0}$$
$$\to f(x_0)g'(x_0) + f'(x_0)g(x_0)$$

für $x \to x_0$.

Dies ist eine leichte Übungsaufgabe. Analoges gilt für Summen, Differenzen und, falls $\lim_{x\to x_0} h(x) \neq 0$, auch für die Quotientenfuntion $x \mapsto g(x)/h(x)$ nach Einschränkung des Definitionsbereichs auf $D \cap h^{\#}(\mathbb{R} \setminus \{0\}) = D \setminus h^{\#}(\{0\})$.

(3) Quotientenregel.

Sind f und g differenzierbar in x_0, $g(x_0) \neq 0$, dann gilt

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

Die Existenz der Ableitung auf der linken Seite ist wieder eine implizite Teilaussage der obigen Formelzeile.

Begründung 1: Ähnlich wie oben, siehe [**22**].

Begründung 2: Folgt unten aus Kettenregel

(4) Kettenregel.

Sei $f : D \longrightarrow \mathbb{R}$, $g : E \longrightarrow \mathbb{R}$ mit $f_\#(D) \subset E$. Sei f in $x_0 \in D$ differenzierbar und sei g in $f(x_0)$ differenzierbar. Dann gilt

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Die Existenz der Ableitung auf der linken Seite ist wieder eine implizite Teilaussage der obigen Formelzeile.

Begründung: für $y_0 := f(x_0)$ Wir rechnen

$$\frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} = \frac{g(f(x)) - g(y_0)}{f(x) - y_0} \cdot \frac{f(x) - f(x_0)}{x - x_0}$$

Aus

$$\lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = g'(y_0)$$

und Lemma 1.13 in Kapitel 4 folgt

$$\lim_{x \to x_0} \frac{g(f(x)) - g(y_0)}{f(x) - y_0} = g'(y_0)$$

Wir erhalten

$$\lim_{x \to x_0} \frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} = g'(y_0)f'(x_0)$$

Begründung der Quotientenregel: Sei $h(x) := 1/x$. Dann

$$\left(\frac{1}{g}\right)'(x_0) = (h \circ g)'(x_0) = -\frac{1}{g(x_0)^2} g'(x_0)$$

$$\begin{aligned}\left(\frac{f}{g}\right)'(x_0) &= f'(x_0)\frac{1}{g(x_0)} + f(x_0)\left(-\frac{1}{g(x_0)^2}g'(x_0)\right) \\ &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}\end{aligned}$$

BEMERKUNG 1.5. Sei $D \subset \mathbb{R}$, $f : D \longrightarrow \mathbb{R}$ injektiv und stetig.

(1) Ist D ein Intervall, dann ist f streng monoton. Denn angenommen f ist nicht streng monoton, dann gibt es $a, b, c \in D$ mit $a < b < c$ und $(f(a) < f(b) > f(c)$ oder $f(a) > f(b) < f(c))$. Dies ergibt mit dem Zwischenwertsatz einen Widerspruch zur Injektivität.

(2) Ist D offen, dann auch $f_\#(D)$. Denn sei $y \in f_\#(D)$, dann schreibe $y = f(x)$. Wähle $a, b \in D$ so dass $x \in (a, b) \subset [a, b] \subset D$. Wegen (1) ist $f|_{[a,b]}$ streng monoton. O.B.d.A. streng monoton wachsend (sonst ersetze $x \mapsto f(x)$ durch $x \mapsto f(-x)$). Somit gilt $y \in (f(a), f(b)) \subset f_\#(D)$.

(3) Ist D offen und f differenzierbar in $x_0 \in D$ und f^{-1} differenzierbar in $f(x_0)$, dann gilt

$$1 = \mathrm{id}'_D(x_0) = (f^{-1})'(f(x_0))\,(f'(x_0)).$$

Also $f'(x_0) \neq 0$.

PROPOSITION 1.6 (Ableitung der Umkehrfunktion.)**.** Sei D offen, $f : D \longrightarrow \mathbb{R}$ injektiv und stetig, und sei f differenzierbar in $x_0 \in D$, $f'(x_0) \neq 0$. Dann ist f^{-1} differenzierbar in $y_0 := f(x_0)$ und

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Beweis. Wähle $a, b \in D$, so dass $x_0 \in (a,b) \subset D$. Dann ist $f|_{(a,b)}$ streng monoton. $y = f(x)$, $y_0 = f(x_0)$.

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} \xrightarrow{y \to y_0} \frac{1}{f'(x_0)},$$

denn mit $y \to y_0$ geht auch $x \to x_0$, da f^{-1} stetig ist (Satz 2.7 und Lemma 1.13 in Kapitel 4). □

Beispiel 1.7.

(1) Wir betrachten $\exp : \mathbb{R} \to \mathbb{R}$.

$$\begin{aligned}\exp'(x) &= \lim_{h\to 0} \frac{\exp(x+h) - \exp(x)}{(x+h) - x} = \lim_{h\to 0} \left((\exp x) \frac{\exp(h) - \exp(0)}{h} \right) \\ &= (\exp x) \lim_{h\to 0} \frac{(\exp h) - 1}{h} = (\exp x) \lim_{h\to 0} \Big(\sum_{n=1}^{\infty} \frac{1}{n!} h^{n-1} \Big) \\ &\overset{(*)}{=} \exp x.\end{aligned}$$

Bei $(*)$ haben wir genutzt, dass wir wegen Lemma 3.1 in Kapitel 4 wissen, dass

$$\Big| \sum_{n=2}^{\infty} \frac{1}{n!} h^{n-1} \Big| \leq \sum_{n=2}^{\infty} \Big| \frac{1}{n!} h^{n-1} \Big| \leq \sum_{n=2}^{\infty} \big| h^{n-1} \big| = \frac{|h|}{1 - |h|},$$

falls $|h| < 1$ und dies konvergiert gegen 0 für $h \to 0$.

(2) Aus $\log = (\exp)^{-1}$ folgt dann mit Proposition 1.6 für $x \in \mathbb{R}_{>0}$

$$\log'(x) = \frac{1}{\exp(\log x)} = \frac{1}{x}.$$

Beispiel 1.8. Sei $f : \mathbb{R}_{>0} \longrightarrow \mathbb{R}$, $f(x) = x^a$, $a \in \mathbb{Q}$. Dann gilt

$$f'(x_0) = a x_0^{a-1}. \tag{1.9}$$

Dies zeigt man

- für $a \in \mathbb{N}$ mit $\mathrm{id}'_{\mathbb{R}}(x_0) = 1$ und der Produktregel.
- für $a = 1/n$, $n \in \mathbb{N} \setminus \{0\}$ mit Umkehrfunktion. Sei $f(x) = x^a = \sqrt[n]{x}$. Dann $f = g^{-1}$ mit $g(x) = x^n$.

$$f'(x_0) = \frac{1}{g'(f(x_0))} = \frac{1}{n(\sqrt[n]{x_0})^{n-1}} = a x_0^{(1-n)/n} = a x_0^{a-1}.$$

- Dann für $a \in \mathbb{Q}_{\geq 0}$ mit der Kettenregel $x^{p/q} = \sqrt[q]{x^p}$.
- Dann für $a \in \mathbb{Q}_{<0}$ mit der Quotientenregel

Ein anderer Beweis gilt sogar für alle $a \in \mathbb{R}$:[3]

$$f(x) = x^a = \exp(a \log x)$$

$$f'(x) = \exp'(a \log x) \cdot (a \log' x) = \exp(a \log x) \cdot \frac{a}{x} = x^a \cdot \frac{a}{x} = a x^{a-1},$$

und somit (1.9) für alle $a \in \mathbb{R}$.

30.1.

2. Lokale Extrema

DEFINITION 2.1. Sei (M, d) ein metrischer Raum, $f : M \longrightarrow \mathbb{R}$, $x_0 \in M$. Wir sagen f hat ein lokales Maximum (bzw. lokales Minimum) in x_0, falls es eine Umgebung U von x_0 gibt, so dass

$$\forall x \in U : f(x) \leq f(x_0) \quad (\text{bzw. } f(x) \geq f(x_0)).$$

lokales Extremum in x_0 $\quad :\Longleftrightarrow \quad$ lokales Maximum oder Minimum in x_0

SATZ 2.2. Sei $D \subset \mathbb{R}$ offen, $f : D \longrightarrow \mathbb{R}$, f differenzierbar in $x_0 \in D$. Wenn f ein lokales Extremum in x_0 besitzt, so gilt $f'(x_0) = 0$.

[3]Er gilt sogar für alle $a \in \mathbb{C}$, wenn man obige Definitionen sinngemäß auf Funktionen $f : D \to \mathbb{C}$ mit $D \subset \mathbb{R}$ verallgemeinert.

Beweis. O.B.d.A. lokales Maximum in x_0.

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0 \quad \text{für } x > x_0, \quad x \in U$$

Somit $f'(x_0) \leq 0$.

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0 \quad \text{für } x < x_0, \quad x \in U$$

Somit $f'(x_0) \geq 0$. □

3. Mittelwertsätze

Wiederholung: (Folgerung 5.23 aus Kapitel 4)

Seien $a, b \in \mathbb{R}$, $a \leq b$, $f : [a, b] \to \mathbb{R}$ stetig. Dann nimmt f ein Maximum und ein Minimum an, d.h. es gibt $x_1, x_2 \in [a, b]$, so dass für alle $x \in [a, b]$:

$$f(x_1) \leq f(x) \leq f(x_2).$$

Satz 3.1 (Satz von Rolle, benannt nach Michel Rolle, 1652–1719)**.** Sei $a < b$. Sei $f : [a, b] \longrightarrow \mathbb{R}$ differenzierbar[4] auf (a, b) und stetig auf $[a, b]$ und $f(a) = f(b)$. Dann existiert ein $c \in (a, b)$ mit $f'(c) = 0$.

Zeichnung zu Satz von Rolle

Beweis. $A := [a, b]$, wähle x_1 und x_2 wie oben.

1. Fall: $x_1 \in \{a, b\}$ und $x_2 \in \{a, b\}$. Dann gilt

$$f(a) = f(b) = \max f_{\#}([a, b]) = \min f_{\#}([a, b])$$

[4]Differenzierbar auf A (bzw. stetig auf A) bedeutet differenzierbar (bzw. stetig) in allen $x \in A$.

Also f konstant, also $f'(x_0) = 0$ für alle $x_0 \in (a, b)$.

2. Fall: $x_1 \in (a, b)$ oder $x_2 \in (a, b)$, o.B.d.A. $x_1 \in (a, b)$. Dann gilt $f'(x_1) = 0$ wegen Satz 2.2. □

Satz 3.2 (1. Mittelwertsatz)**.** Sei $a < b$. Sei $f : [a, b] \longrightarrow \mathbb{R}$ differenzierbar auf (a, b) und stetig auf $[a, b]$. Dann existiert ein $c \in (a, b)$ mit

$$f(b) - f(a) = (b - a)f'(c).$$

Zeichnung zu 1. Mittelwertsatz

Beweis.

$$F(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

$F(a) = f(a) = F(b)$. Wende den Satz von Rolle an.

$$0 = F'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}$$

□

Satz 3.3 (2. Mittelwertsatz)**.** Sei $a < b$. Seien $f, g : [a, b] \longrightarrow \mathbb{R}$ differenzierbar auf (a, b) und stetig auf $[a, b]$. Ferner gelte für alle $x \in (a, b)$: $g'(x) \neq 0$. Dann existiert ein $c \in (a, b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Beweis. Der Satz von Rolle liefert: $g(b) \neq g(a)$. Definiere

$$F(x) := f(x) - f(a) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)).$$

Also $F(a) = F(b) = 0$. Nach dem Satz von Rolle gibt es ein $c \in (a, b)$ mit

$$0 = F'(c) = f'(c) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(c).$$

□

KOROLLAR 3.4. Sei $f : (a, b) \longrightarrow \mathbb{R}$ differenzierbar, so gilt:

$$f'(x) > 0 \text{ für alle } x \in (a, b) \Longrightarrow f \text{ ist streng monoton wachsend}$$

$$f'(x) < 0 \text{ für alle } x \in (a, b) \Longrightarrow f \text{ ist streng monoton fallend}$$

$$f'(x) \geq 0 \text{ für alle } x \in (a, b) \Longleftrightarrow f \text{ ist monoton wachsend}$$

$$f'(x) \leq 0 \text{ für alle } x \in (a, b) \Longleftrightarrow f \text{ ist monoton fallend}$$

$$f'(x) = 0 \text{ für alle } x \in (a, b) \Longleftrightarrow f \text{ ist konstant}$$

Beweis. Die $\Longrightarrow$-Pfeile folgen direkt aus dem 1. Mittelwertsatz. Sei nun f monoton wachsend, dann gilt $f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$. Und analog folgt aus „monoton fallend“ dann $f'(x_0) \leq 0$. Der dritte $\Longleftrightarrow$-Pfeil folgt aus den beiden ersten. □

4. Höhere Ableitungen und Taylorscher Satz

Sei $I \subset \mathbb{R}$ ein offenes Intervall.

$$C^0(I) := \{f : I \longrightarrow \mathbb{R} \mid f \text{ stetig}\}$$

Ist f auf I differenzierbar, so erhalten wir $f' : I \longrightarrow \mathbb{R}$, $x \mapsto f'(x)$.

Falls zusätzlich f' stetig ist, so nennen wir f stetig differenzierbar.

Ist f' differenzierbar, so nennt man f zweimal differenzierbar.... Ist f' k-mal differenzierbar, so sagen wir: f ist $(k+1)$-mal differenzierbar.

$f^{(0)} := f$, $f^{(1)} := f'$, $f^{(k+1)} := (f')^{(k)}$.

f nennt man $(k+1)$-mal differenzierbar **in $\mathbf{x_0} \in I$**, wenn f k-mal differenzierbar auf einer Umgebung von x_0 ist und $f^{(k)}$ in x_0 differenzierbar ist.

Für $k \in \mathbb{N}$ nennt man

$$C^k(I) := \{f : I \longrightarrow \mathbb{R} \mid f \text{ ist } k\text{-mal differenzierbar und } f^{(k)} \in C^0(I)\}$$

die Menge der k-mal stetig differenzierbaren Funktionen.

$C^\infty(I) := \bigcap_{k\in\mathbb{N}} C^k(I)$, die Menge der unendlich oft differenzierbaren Funktionen, glatt = unendlich oft differenzierbar.

Satz 4.1. Sei $f : (a,b) \longrightarrow \mathbb{R}$ differenzierbar auf (a,b) und 2-mal differenzierbar in $x_0 \in (a,b)$.

(1) Es gelte $f'(x_0) = 0$ und $f''(x_0) < 0$. Dann besitzt f in x_0 ein lokales Maximum.
(2) Es gelte $f'(x_0) = 0$ und $f''(x_0) > 0$. Dann besitzt f in x_0 ein lokales Minimum.
(3) Besitzt f in x_0 ein lokales Maximum, so gilt $f'(x_0) = 0$ und $f''(x_0) \leq 0$.
(4) Besitzt f in x_0 ein lokales Minimum, so gilt $f'(x_0) = 0$ und $f''(x_0) \geq 0$.

Beweis.
Zu (1): Die Funktion

$$g(x) := \begin{cases} \frac{f'(x)-f'(x_0)}{x-x_0} & \text{für } x \neq x_0 \\ f''(x_0) & \text{für } x = x_0 \end{cases}$$

ist stetig in x_0. Aus $f''(x_0) < 0$ folgt dann mit dem ϵ-δ-Kriterium: Zu $\epsilon := |f''(x_0)|/2$ gibt es dann ein $\delta > 0$, so dass für alle $x \in (x_0 - \delta, x_0 + \delta)$ gilt:

$$x \in (a, b) \text{ und } \left| \frac{f'(x) - f'(x_0)}{x - x_0} - f''(x_0) \right| < \epsilon.$$

Daraus folgt

$$\frac{f'(x) - \overbrace{f'(x_0)}^{=0}}{x - x_0} < f''(x_0) + \epsilon = \frac{1}{2} f''(x_0) < 0.$$

Also gilt $f'(x) < 0$ für $x \in (x_0, x_0 + \delta)$ und $f'(x) > 0$ für $x \in (x_0 - \delta, x_0)$. Dies impliziert, zusammen mit Korollar 3.4 und der Stetigkeit von f in x_0, dass $f(x_0) = \max f_{\#}((x_0 - \delta, x_0 + \delta))$.

Zu (2): Analog.

Zu (3): Es folgt zunächst $f'(x_0) = 0$. Wenn $f''(x_0) > 0$ wäre, so hätte f in x_0 ein lokales Maximum und Minimum, wäre also konstant auf einer Umgebung von x_0. Somit $f''(x_0) = 0$. Widerspruch.

Zu (4): Analog. □

Satz 4.2 (Satz von Taylor)**.** Sei $f : (a, b) \longrightarrow \mathbb{R}$ $(n+1)$-mal differenzierbar, $x, x_0 \in (a, b)$. Definiere das Restglied $R_n^{x_0}(x)$ durch

$$\begin{aligned} f(x) &= f(x_0) + (x - x_0) f'(x_0) + \frac{(x - x_0)^2}{2!} f^{(2)}(x_0) + \cdots + \frac{(x - x_0)^n}{n!} f^{(n)}(x_0) + R_n^{x_0}(x) \\ &= \sum_{j=0}^{n} \frac{(x - x_0)^j}{j!} f^{(j)}(x_0) + R_n^{x_0}(x) \end{aligned}$$

Dann gilt:
(1) $\lim\limits_{x \to x_0} \dfrac{R_n^{x_0}(x)}{(x - x_0)^{n+1}}$ existiert und ist gleich $\dfrac{f^{(n+1)}(x_0)}{(n+1)!}$.

(2) $\dfrac{R_n^{x_0}(x)}{(x-x_0)^{n+1}} = \dfrac{f^{(n+1)}(x_0+\vartheta(x-x_0))}{(n+1)!}$ für ein $\vartheta \in (0,1)$.

Lagrangesche Restglieddarstellung

Bemerkung: Der Fall $n = 0$ ist der 1. Mittelwertsatz.

Beweis. Wir fixieren ein x_0. Setze $g(x) := R_n^{x_0}(x)$, $x \neq x_0$. Man rechnet nach $0 = g(x_0) = g'(x_0) = g^{(2)}(x_0) = \cdots = g^{(n)}(x_0)$ und $g^{(n+1)}(x_0) = f^{(n+1)}(x_0)$.

$$
\begin{aligned}
\frac{R_n^{x_0}(x)}{(x-x_0)^{n+1}} &= \frac{g(x)-g(x_0)}{(x-x_0)^{n+1}-(x_0-x_0)^{n+1}} \\
&\overset{\text{2.MWS}}{=} \frac{g'(\xi_1)}{(n+1)(\xi_1-x_0)^n} \\
&= \frac{g'(\xi_1)-g'(x_0)}{(n+1)(\xi_1-x_0)^n-(n+1)(x_0-x_0)^n} \\
&\overset{\text{2.MWS}}{=} \frac{g^{(2)}(\xi_2)}{n(n+1)\,(\xi_2-x_0)^{n-1}} \\
&\;\;\vdots \\
&= \frac{g^{(n)}(\xi_n)-g^{(n)}(x_0)}{(n+1)!\,(\xi_n-x_0)} \\
&\overset{\text{2.MWS}}{=} \frac{g^{(n+1)}(\xi_{n+1})}{(n+1)!} = \frac{f^{(n+1)}(\xi_{n+1})}{(n+1)!}
\end{aligned}
$$

für geeigenete ξ_j, mit $x_0 < \xi_{n+1} < \xi_n < \cdots < \xi_1 < x$ oder $x_0 > \xi_{n+1} > \xi_n > \cdots > \xi_1 > x$. Hieraus folgt Teil (2) des Satzes.

In der zweitletzten Zeile der Umformung steht (bis auf den Faktor $(n+1)!$) der Differenzenquotient von $g^{(n)}$, der für $\xi_n \to x_0$ gegen

$$\frac{1}{(n+1)!} g^{(n+1)}(x_0) = \frac{1}{(n+1)!} f^{(n+1)}(x_0)$$

konvergiert.

Wegen $|\xi_n - x_0| \leq |x - x_0|$ haben wir dieselbe Aussage auch im Limes $x \to x_0$. Somit

$$\lim_{x \to x_0} \frac{R_n^{x_0}(x)}{(x-x_0)^{n+1}} = \frac{1}{(n+1)!} f^{(n+1)}(x_0)$$

Somit gilt auch (1). □

DEFINITION 4.3. Sei $f : (a,b) \longrightarrow \mathbb{R}$ n-mal differenzierbar, $x_0 \in (a,b)$. Dann nennt man

$$T_n^{x_0} f\,(x) := \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j$$

das Taylor-Polynom n-ten Grades von f mit Entwicklungspunkt x_0. Ist $f \in C^\infty((a,b))$, dann nennt man

$$T^{x_0} f\,(x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j$$

die Taylorreihe von f mit Entwicklungspunkt x_0.

!ACHTUNG!. Die Konvergenz der Taylor-Reihe ist zunächst unklar. Und selbst wenn sie konvergiert, ist unklar, ob sie mit der ursprünglichen Funktion übereinstimmt. Man kann zeigen: jede Potenzreihe (also auch jede Potenzreihe mit Konvergenzradius 0) ist die Taylor-Reihe einer glatten Funktion, siehe `https://www.semanticscholar.org/paper/Peano's-Unnoticed-Proof-of-Borel's-Theorem-Besenyei/3f06a77700dfdfb470d1786fc9348b1865b` für einen Beweis.

In Analysis II und III werden wir sehen: Ist f durch eine konvergente Potenzreihe gegeben,
1.2. dann stimmt die Taylorreige $T^0 f$ mit Entwicklungspunkt 0 mit dieser Potenzreiheüberein.

Anwendung: Potenzreihendarstellung der Logarithmus-Funktion.

Entwicklungspunkt: $x_0 = 1$

$\log'(x) = \frac{1}{x}, \quad \log^{(j)}(x) = (-1)^{j-1}\frac{(j-1)!}{x^j}$

$\log(1) = 0$, $\log^{(j)}(1) = (-1)^{j-1}(j-1)!$

$T_n^1 \log(x) = \underbrace{\frac{\log(1)}{0!}(x-1)^0}_{0} + \sum_{j=1}^n \frac{(-1)^{j-1}(j-1)!}{j!}(x-1)^j = \sum_{j=1}^n \frac{(-1)^{j-1}}{j}(x-1)^j$

Fragen 4.4.

- Konvergiert die Potenzreihe $T^1\log(x)$ für manche $x \neq 1$? In anderen Worten: ist der Konvergenzradius der Reihe positiv?
- Falls ja: gilt im Konvergenzbereich $\log(x) = T^1\log(x)$?

Antwort: Ja für $x \in (0, 2)$.

Wir wollen die Aussage für $x \in [1, 2)$ zeigen. Für $x \in (0, 1)$ kann man die Ausage mit einer anderen Version des Satzes von Taylor zeigen, der Cauchyschen Restglieddarstellung, siehe [**1**, Kap. IV, Korollar 3.8 und Anwendungen].

Nach dem Satz von Taylor existiert ein $\vartheta \in (0, 1)$ mit

$$\begin{aligned}\log x - T_n^1\log(x) = R_n^1(x) &= \frac{(x-1)^{n+1}}{(n+1)!}\log^{(n+1)}(1+\vartheta(x-1)) \\ &= (-1)^n\frac{(x-1)^{n+1}}{(n+1)!}\frac{n!}{(1+\vartheta(x-1))^{n+1}}\end{aligned}$$

Also

$$|\log x - T_n^1\log(x)| = \left|\frac{(x-1)^{n+1}}{(n+1)!}\frac{n!}{(1+\vartheta(x-1))^{n+1}}\right| \leq \frac{1}{n+1}$$

Somit $T_n^1\log(x) \overset{n\to\infty}{\longrightarrow} \log(x)$ für $x \in [1,2)$.

PROPOSITION 4.5. Die Taylorreihe von $\log : \mathbb{R}_{>0} \to \mathbb{R}$ mit Entwicklungspunkt 1 ist

$$T^1\log(x) = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j}(x-1)^j.$$

Die Reihe konvergiert für $|x-1| < 1$, also für insbesondere für $x \in (0,2)$. Außerdem gilt für alle $x \in (0,2)$:

$$\log x = T^1\log(x).$$

Beweis. Dass die Reihe für $x \in [1,2)$ gegen $\log x$ konvergiert, wurde oben gezeigt. Dass die Reihe für $|x-1| < 1$ konvergiert, folgt (z.B.) aus dem Quotientenkriterium. Die Aussage $T^1\log(x) = \log(x)$ für $x \in (0,1)$ zeigen wir aus Zeitgründen nicht mehr, einen Beweis findet man wieder in [**1**, Kap. IV, Korollar 3.8 und Anwendungen]. □

Übrigens: In der Analysis III werden Sie Sätze kennenlernen, die besagen: wenn $\log x = T^1\log(x)$ für alle $x \in [1,2)$ gilt, dann auch für alle $x \in (0,1)$.

KAPITEL 6

Integral-Rechnung für Funktionen einer Veränderlichen

Ziel: Integration von Funktionen. Flächeninhalt unter dem Graphen einer Funktion.

Graph einer Funktion. Flächeninhalt darunter im Intervall $[a, b]$ markiert.

1. Partitionen und Treppenfunktionen

Sei I eine Menge. (Partielle) Ordnung auf $\mathrm{Abb}(I, \mathbb{R}) = \mathbb{R}^I \ni f, g$:

$$f \leq g \quad :\Longleftrightarrow \quad \forall x \in I : \, f(x) \leq g(x).$$

DEFINITION 1.1. Eine Zerlegung oder Partition von $I = [a, b]$, $a, b \in \mathbb{R}$ ist ein Tupel $(t_0, t_1, \ldots, t_k) \in \mathbb{R}^{k+1}$ mit $a = t_0 < t_1 < \cdots < t_k = b$. Wir schreiben $Z = \{t_0 < t_1 < \cdots < t_k\}$. Man nennt $Z_1 = \{s_0 < s_1 < \cdots < s_\ell\}$ eine Verfeinerung von $Z_2 = \{t_0 < t_1 < \cdots < t_k\}$, wenn Z_1 durch Hinzufügen von Punkten aus Z_2 entsteht, in anderen Worten falls

$$\{t_i \mid i = 0, 1, \ldots, k\} \subset \{s_i \mid i = 0, 1, \ldots, \ell\}.$$

Zu zwei Zerlegungen Z_1 und Z_2 von $[a, b]$ gibt es Zerlegung Z_3 die feiner als Z_1 und feiner als Z_2 ist: vereinige hierzu die Mengen der Zerlegungspunkte und ordne sie.

DEFINITION 1.2. Eine Funktion $f : [a, b] \to \mathbb{R}$ heißt Treppenfunktion, falls es eine Zerlegung $Z = \{t_0 < t_1 < \cdots < t_k\}$ gibt, so dass $f|_{(t_{j-1}, t_j)}$ konstant ist für alle $j \in \{1, 2, \ldots, k\}$.

Wir nennen dieses Z eine Zerlegung, die zu der Treppenfunktion passt. Jede Verfeinerung einer passenden Zerlegung passt ebenfalls. Die Menge der Treppenfunktionen $[a,b] \to \mathbb{R}$ nennen wir $\mathcal{T}[a,b]$.

Treppenfunktionen sind beschränkt, da sie nur endlich viele Werte annehmen.

LEMMA 1.3. Sind $f, f_1, f_2 \in \mathcal{T}[a,b]$, $\lambda_1, \lambda_2 \in \mathbb{R}$, dann folgt auch

(a) $\lambda_1 f_1 + \lambda_2 f_2 \in \mathcal{T}[a,b]$, d.h. $\mathcal{T}[a,b]$ ist ein Vektorraum.[1]
(b) $f_1 f_2 \in \mathcal{T}[a,b]$, d.h. $\mathcal{T}[a,b]$ ist ein Ring bzw. eine Algebra.[2]
(c) $\max\{f_1, f_2\} \in \mathcal{T}[a,b]$ und $\min\{f_1, f_2\} \in \mathcal{T}[a,b]$ [3]
(d) $|f| = \max\{f, -f\} \in \mathcal{T}[a,b]$

Beweis klar.

DEFINITION 1.4. Für eine Treppenfunktion $f \in \mathcal{T}[a,b]$ mit passender Zerlegung $Z = \{t_0 < t_1 < \cdots < t_k\}$ definieren wir das Integral

$$\int_a^b{}_{(Z)}\, f(x)\,dx := \sum_{j=1}^{k} (t_j - t_{j-1})\, f\left(\frac{t_{j-1}+t_j}{2}\right) \in \mathbb{R}.$$

Positive Treppenfunktion. Die Fläche darunter ist bunt und ergibt das Integral

Wenn Z_1 und Z_2 zu einer Treppenfunktion f passen, so gilt

$$\int_a^b{}_{(Z_1)}\, f(x)\,dx = \int_a^b{}_{(Z_2)}\, f(x)\,dx.$$

[1] Hier wird wieder wie immer definiert $(\lambda_1 f_1 + \lambda_2 f_2)(x) := \lambda_1 f_1(x) + \lambda_2 f_2(x)$.

[2] Hier wird wieder wie immer definiert $(f_1 f_2)(x) := f_1(x) f_2(x)$.

[3] Hier wird wieder wie immer definiert $\max\{f_1, f_2\}(x) := \max\{f_1(x), f_2(x)\}$ und $\min\{f_1, f_2\}(x) := \min\{f_1(x), f_2(x)\}$

Wir schreiben hierfür deswegen ab sofort einfach

$$\int_a^b f(x)\,dx.$$

PROPOSITION 1.5. Die Abbildung $\int_a^b : \mathcal{T}[a,b] \to \mathbb{R}$ ist

(1) linear, d.h.

$$\forall f_1, f_2 \in \mathcal{T}[a,b] : \forall \lambda_1, \lambda_2 \in \mathbb{R} : \int_a^b (\lambda_1 f_1(x) + \lambda_2 f_2(x))\,dx = \lambda_1 \int_a^b f_1(x)\,dx + \lambda_2 \int_a^b f_2(x)\,dx$$

(2) monoton, d.h.

$$\forall f_1, f_2 \in \mathcal{T}[a,b] : \left(f_1 \leq f_2 \quad \Longrightarrow \quad \int_a^b f_1(x)\,dx \leq \int_a^b f_2(x)\,dx\right)$$

Beweis klar aus der Konstruktion des Integrals.

2. Das Riemann-Integral

$$f \in \mathbb{R}^{[a,b]} \text{ ist nach oben beschränkt } \Longleftrightarrow \exists g \in \mathcal{T}[a,b] : g \geq f \Longleftrightarrow (\mathcal{T}[a,b])_{\geq f} \neq \emptyset$$

$$f \in \mathbb{R}^{[a,b]} \text{ ist nach unten beschränkt } \Longleftrightarrow \exists g \in \mathcal{T}[a,b] : g \leq f \Longleftrightarrow (\mathcal{T}[a,b])_{\leq f} \neq \emptyset$$

DEFINITION 2.1. Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Wir definieren das riemannsche Oberintegral als

$$\overline{\int_a^b} f(x)\,dx := \inf\left\{\int_a^b g(x)\,dx \,\middle|\, g \in (\mathcal{T}[a,b])_{\geq f}\right\}$$

und das riemannsche Unterintegral als

$$\underline{\int_a^b} f(x)\,dx := \sup\left\{\int_a^b g(x)\,dx \,\middle|\, g \in (\mathcal{T}[a,b])_{\leq f}\right\}$$

Für die beschränkte Funktion f gilt

$$-\infty \ < \int_a^b f(x)\,dx \ \leq \overline{\int_a^b} f(x)\,dx \ < \ \infty$$

Ist $f \in \mathcal{T}[a,b]$, so gilt

$$\int_a^b f(x)\,dx = \underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x).$$

DEFINITION 2.2. Man nennt $f \in \mathbb{R}^{[a,b]}$ Riemann-integrierbar, falls f beschränkt ist und falls

$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx.$$

In diesem Fall nennt man

$$\int_a^b f(x)\,dx := \underline{\int_a^b} f(x)\,dx$$

das Riemann-Integral von f.

$$\mathcal{R}[a,b] := \{f \in \mathbb{R}^{[a,b]} \mid f \text{ ist Riemann-integrierbar}\}$$

LEMMA 2.3 (Riemannsches Kriterium)**.** Sei $f \in \mathbb{R}^{[a,b]}$. Dann gilt

$$f \in \mathcal{R}[a,b] \quad \Longleftrightarrow \quad \forall \epsilon \in \mathbb{R}_{>0} : \exists g_u, g_o \in \mathcal{T}[a,b] : (\, g_u \leq f \leq g_o) \wedge \left(\int_a^b g_o(x)\,dx - \int_a^b g_u(x)\,dx \leq \epsilon \right).$$

Für g_u und g_o wie oben gilt dann wegen der Monotonie

$$\int_a^b g_u(x)\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b g_o(x)\,dx.$$

Eine stetige Funktion mit Treppenfunktionen g_u darunter und g_o darüber.

Beweis.

„$\Longrightarrow$“: Auf Grund der Definition des Oberintegrals gibt es eine Treppenfunktion $g_o \geq f$ mit $\int_a^b g_o(x)\,dx \leq \int_a^b f(x)\,dx + \epsilon/2$. Auf Grund der Definition des Unterintegrals gibt es eine Treppenfunktion $g_u \leq f$ mit $\int_a^b g_u(x)\,dx \geq \int_a^b f(x)\,dx - \epsilon/2$. Da da Oberintegral von f mit dem Unterintegral übereinstimmt, gilt

$$\int_a^b g_o(x)\,dx - \int_a^b g_u(x)\,dx \leq \epsilon$$

„$\Longleftarrow$“: Gilt die rechte Seite, so haben wir für jedes $\epsilon > 0$:

$$0 \leq \int_a^b f(x)\,dx - \int_a^b f(x)\,dx \leq \int_a^b g_o(x)\,dx - \int_a^b g_u(x)\,dx \leq \epsilon.$$

Daraus folgt die Riemann-Integrierbarkeit. □ **Mi 6.2.**

Lemma 2.4. Sei $a \leq b$. Sind $f, f_1, f_2 \in \mathcal{R}[a,b]$, $\lambda_1, \lambda_2 \in \mathbb{R}$ dann gilt:

(a) $-f \in \mathcal{R}[a,b]$ und

$$\int_a^b (-f)(x)\,dx = -\int_a^b f(x)\,dx\,.$$

(b) $\lambda_1 f_1 + \lambda_2 f_2 \in \mathcal{R}[a,b]$, d.h. $\mathcal{R}[a,b]$ ist ein Vektorraum. Die Abbildung

$$\int_a^b : \mathcal{R}[a,b] \to \mathbb{R}, \quad f \mapsto \int_a^b f(x)\,dx$$

ist linear.

(c) $f_1 f_2 \in \mathcal{R}[a,b]$, d.h. $\mathcal{R}[a,b]$ ist ein Ring bzw. eine Algebra.

(d) $\max\{f_1, f_2\} \in \mathcal{R}[a,b]$ und $\min\{f_1, f_2\} \in \mathcal{R}[a,b]$

(e) $|f| = \max\{f, -f\} \in \mathcal{R}[a,b]$

Beweis.

(a): Aus

$$g \in (\mathcal{T}[a,b])_{\leq -f} \Longleftrightarrow -g \in (\mathcal{T}[a,b])_{\geq f}$$

folgt

$$\underline{\int_a^b} (-f)(x)\,dx = -\overline{\int_a^b} f(x)\,dx$$

und genauso

$$\overline{\int_a^b} (-f)(x)\,dx = -\underline{\int_a^b} f(x)\,dx\,.$$

(b):

Fall $\lambda_1 \geq 0$ und $\lambda_2 \geq 0$: Sei $f_{1,u} \in (\mathcal{T}[a,b])_{\leq f_1}$ und $f_{2,u} \in (\mathcal{T}[a,b])_{\leq f_2}$, dann gilt wegen Lemma 1.3 (a) $\lambda_1 f_{1,u} + \lambda_2 f_{2,u} \in (\mathcal{T}[a,b])_{\leq \lambda_1 f_1 + \lambda_2 f_2}$. Es folgt

$$\begin{aligned} \underline{\int_a^b} (\lambda_1 f_1(x) + \lambda_2 f_2)(x)\,dx \qquad &\geq \qquad \int_a^b (\lambda_1 f_{1,u} + \lambda_2 f_{2,u})(x)\,dx \\ &\overset{Prop.\ 1.5(1)}{=} \lambda_1 \int_a^b f_{1,u}(x)\,dx + \lambda_2 \int_a^b f_{2,u}(x)\,dx \end{aligned}$$

und somit durch Bildung der Suprema

$$\underline{\int_a^b} (\lambda_1 f_1(x) + \lambda_2 f_2)(x)\,dx \geq \lambda_1 \int_a^b f_1(x)\,dx + \lambda_2 \int_a^b f_2(x)\,dx.$$

Analog mit oberen Schranken:

$$\overline{\int_a^b} (\lambda_1 f_1(x) + \lambda_2 f_2)(x)\,dx \leq \lambda_1 \int_a^b f_1(x)\,dx + \lambda_2 \int_a^b f_2(x)\,dx.$$

Der allgemeine Fall: folgt dann mit (a).

(d): Wir zeigen zunächst: Ist $f \in \mathcal{R}[a,b]$, dann auch $f^+ := \max\{f, 0\}$.

Sei also $f \in \mathcal{R}[a,b]$ und $\epsilon > 0$ gegeben. Wir nehmen $g, h \in \mathcal{T}[a,b]$ mit $g \le f \le h$ und $\int_a^b h(x)\,dx - \int_a^b g(x)\,dx \le \epsilon$. Wir setzen $g^+ := \max\{g, 0\} \in \mathcal{T}[a,b]$ und $h^+ := \max\{h, 0\} \in \mathcal{T}[a,b]$. Dann gilt $g^+ \le f^+ \le h^+$ und

$$h^+(x) - g^+(x) = \left.\begin{cases} h(x) - g^+(x) & \text{falls } h(x) \ge 0 \\ 0 & \text{falls } h(x) < 0 \end{cases}\right\} \le h(x) - g(x)$$

Also

$$\int_a^b h^+(x)\,dx - \int_a^b g^+(x)\,dx \le \int_a^b h(x)\,dx - \int_a^b g(x)\,dx \le \epsilon.$$

Also ist dann $f^+ \in \mathcal{R}[a,b]$.

Wegen $\max\{f_1, f_2\} = f_2 + \max\{f_1 - f_2, 0\}$ und $\min\{f_1, f_2\} = f_2 - \max\{f_2 - f_1, 0\}$ folgt dann (d).

(c):

Der Fall $f_1 \ge 0$ und $f_2 \ge 0$: Bestimme zu $j \in \{1,2\}$: $f_{j,u}, f_{j,o} \in \mathcal{T}[a,b]$ mit

$$0 \le f_{j,u} \le f_j \le f_{j,o} \le C_j := \sup\{f_j(x) \mid x \in [a,b]\}$$

und

$$\int_a^b f_{j,o}(x)\,dx - \int_a^b f_{j,u}(x)\,dx \le \epsilon.$$

Aus Lemma 1.3 (b) folgt $f_{1,u} f_{2,u} \in \mathcal{T}[a,b]$ und $f_{1,o} f_{2,o} \in \mathcal{T}[a,b]$. Ferner gilt

$$\begin{aligned} & \int_a^b f_{1,o}(x) f_{2,o}(x)\,dx - \int_a^b f_{1,u}(x) f_{2,u}(x)\,dx \\ \le\ & \int_a^b f_{1,o}(x)\big(f_{2,o}(x) - f_{2,u}(x)\big)\,dx + \int_a^b \big(f_{1,o}(x) - f_{1,u}(x)\big) f_{2,u}(x)\,dx \\ \le\ & C_2\epsilon + C_1\epsilon\,. \end{aligned}$$

Der allgemeine Fall folgt dann aus der Gleichung

$$\begin{aligned} f_1 f_2 &= \max\{f_1,0\}\max\{f_2,0\} - \max\{-f_1,0\}\max\{f_2,0\} \\ &\quad - \max\{f_1,0\}\max\{-f_2,0\} + \max\{-f_1,0\}\max\{-f_2,0\}. \end{aligned}$$

(e): folgt direkt aus (d). □

LEMMA 2.5 (Monotonie)**.** Seien, $f_1, f_2 \in \mathcal{R}[a,b]$. Ist $f_1 \le f_2$, so haben wir

$$\int_a^b f_1(x)\,dx \le \int_a^b f_2(x)\,dx.$$

Beweis. $0 \in \mathcal{T}[a,b]$, $0 \le f_2 - f_1$. Es folgt

$$0 = \int_a^b 0\,dx \overset{(*)}{\le} \int_a^b (f_2 - f_1)(x)\,dx \overset{Lemma\ 2.4(b)}{=} \int_a^b f_2(x)\,dx - \int_a^b f_1(x)\,dx\,.$$

Die Ungleichung $(*)$ gilt nach Definition des Unterintegrals: [4] es ist definiert als das Supremum der Menge

$$A := \Big\{ \int_a^b g(x)\,dx \,\Big|\, g \in (\mathcal{T}[a,b])_{\le f} \Big\} \ni \int_a^b 0\,dx = 0.$$

□

Notation: Für $f : D \to \mathbb{R}$ definiere $\|f\|_\infty := \sup\big\{|f(x)| \,\big|\, x \in D\big\} \in [0,\infty]$.

[4] An der Stelle $(*)$ kann man nicht mit der Monotonie des Integrals argumentieren. Denn die Monotonie für Treppenfunktion (Proposition 1.5 (2)) ist nicht anwendbar, da f im allgemeinen keine Treppenfunktion ist. Und die Monotonie für Riemann-integrierbare Funktionen wollen wir ja hier erst zeigen. Wenn wir sie verwenden würden, so wäre dies ein Zirkelschluss: Ein Beweis, in dem man die zu beweisende Aussage bereits verwendet und der deswegen nicht zulässig ist.

PROPOSITION 2.6. Für $f \in \mathcal{R}[a,b]$ gilt:

$$\Big|\int_a^b f(x)\,dx\Big| \le \int_a^b |f(x)|\,dx \le \|f\|_\infty (b-a)$$

Beweis. Wegen $f = \max\{f,0\} - \max\{-f,0\}$ und $|f| = \max\{f,0\} + \max\{-f,0\}$ folgt

$$\begin{aligned}\Big|\int_a^b f(x)\,dx\Big| &\le \Big|\underbrace{\int_a^b \max\{f(x),0\}\,dx}_{\ge 0} - \underbrace{\int_a^b \max\{-f(x),0\}\,dx}_{\ge 0}\Big| \\ &\overset{(*)}{\le} \int_a^b \max\{f(x),0\}\,dx + \int_a^b \max\{-f(x),0\}\,dx \\ &= \int_a^b |f(x)|\,dx,\end{aligned}$$

wobei wir an der Stelle $(*)$ die Dreiecks-Ungleichung $|a-b| \le |a| + |b|$ benutzt haben.

Wegen $|f(x)| \le \|f\|_\infty$ folgt mit der Monotonie

$$\int_a^b |f(x)|\,dx \le \int_a^b \|f\|_\infty\,dx \le \|f\|_\infty (b-a).$$

□

LEMMA 2.7. Seien $a,b,c \in \mathbb{R}$ mit $a < b < c$, und $f : [a,c] \to \mathbb{R}$. Dann gilt

$$f \in \mathcal{R}[a,c] \quad \Longleftrightarrow \quad f|_{[a,b]} \in \mathcal{R}[a,b] \text{ und } f|_{[b,c]} \in \mathcal{R}[b,c]$$

und wenn f Riemann-integrierbar ist, so gilt

$$\int_a^c f(x)\,dx = \int_a^b f|_{[a,b]}(x)\,dx + \int_b^c f|_{[b,c]}(x)\,dx.$$

Beweis. Man wählt wieder geeignete Treppenfunktionen oberhalb und unterhalb von f und setzt diese zusammen oder durchtrennt sie an der Stelle b. □

NOTATION 2.8. Die Integrationsvariable (bis jetzt immer x) kann durch jede andere noch nicht vergebene Variable ersetzt werden:

$$\int_a^b f(x)\,dx = \int_a^b f(y)\,dy = \cdots = \int_a^b f(t)\,dt\,.$$

Aber bereits vergebene Variablen dürfen nicht als Integrationsvariable benutzt werden, zum Beispiel ist

$$\int_a^b f(a)\,da$$

ist nicht erlaubt.

Gilt $[a,b] \subset [c,d]$ und $f \in \mathcal{R}[c,d]$, dann schreiben wir

$$\int_a^b f(x)\,dx := \int_a^b f|_{[a,b]}(x)\,dx.$$

Gilt $a > b$, und $f \in \mathcal{R}[b,a]$, so definieren wir

$$\int_a^b f(x)\,dx := -\int_b^a f(x)\,dx.$$

Insbesondere gilt für alle reelle Zahlen $a,b,c \in [d,e] \subset \mathbb{R}$ und für alle $f \in \mathcal{R}[d,e]$:

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx.$$

3. Monotone Funktionen sind Riemann-integrierbar

SATZ 3.1. Sei $a,b \in \mathbb{R}$, $a < b$, und $f : [a,b] \to \mathbb{R}$ eine monotone Funktion. Dann ist $f \in \mathcal{R}[a,b]$.

Beweis. O.B.d.A. sei f monoton wachsend. Sei $\epsilon \in \mathbb{R}_{>0}$ gegeben. Zu einer Zahl $k \in \mathbb{N}$, die wir noch geeignet wählen werden, definieren wir:

$$t_j := a + j\frac{b-a}{k}.$$

Insbesondere ist $\{t_0 < t_1 < \cdots < t_k\}$ eine Zerlegung von $[a, b]$. Wir definieren die Treppenfunktionen $g_o, g_u \in \mathcal{T}[a, b]$, so dass für alle $j \in \{1, 2, \ldots k\}$ und alle $x \in [t_{j-1}, t_j)$ gilt

$$g_o(x) = f(t_j), \qquad g_u(x) = f(t_{j-1});$$

und $g_o(b) = g_u(b) = f(b)$.

Bild einer monoton wachsenden Funktion f und von Treppenfunktionen g_u, g_o wie oben.

Da f monoton wachsend ist gilt $g_u \leq f \leq g_o$. Außerdem

$$\begin{aligned}
\int_a^b g_o(x)\,dx &= \sum_{j=1}^{k} (t_j - t_{j-1})\, g_o\left(\frac{t_{j-1} + t_j}{2}\right) = \sum_{j=1}^{k} \frac{b-a}{k}\, f(t_j), \\
\int_a^b g_u(x)\,dx &= \sum_{j=1}^{k} (t_j - t_{j-1})\, g_u\left(\frac{t_{j-1} + t_j}{2}\right) = \sum_{j=1}^{k} \frac{b-a}{k}\, f(t_{j-1}).
\end{aligned}$$

Somit

$$\int_a^b g_o(x)\,dx - \int_a^b g_u(x)\,dx = \frac{b-a}{k}\left(\sum_{j=1}^{k} f(t_j) - \sum_{\ell=0}^{k-1} f(t_\ell)\right) = \frac{b-a}{k}(f(b) - f(a))$$

Wenn wir also bei der obigen Konstruktion der t_j, die Zahl $k \in \mathbb{N}$ so groß wählen, dass

$$k \geq \frac{b-a}{\epsilon}(f(b) - f(a)),$$

dann gilt $\int_a^b g_o(x)\,dx - \int_a^b g_u(x)\,dx \leq \epsilon$. Mit dem riemannschen Kriterium Lemma 2.3 folgt die Riemann-Integrierbarkeit. □

4. Stetige Funktionen sind Riemann-integrierbar

LEMMA 4.1. Sei $f : [a, b] \to \mathbb{R}$ stetig. Dann ist f auch gleichmäßig stetig, d.h.

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists \delta \in \mathbb{R}_{>0} : \forall x, y \in [a, b] : \big(|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon\big). \tag{4.2}$$

Beweis. Angenommen f ist nicht gleichmäßig stetig. Es gibt dann ein $\epsilon > 0$, so dass

$$\forall \delta \in \mathbb{R}_{>0} : \exists x, y \in [a, b] : \big(|x - y| < \delta\big) \wedge \big(|f(x) - f(y)| \geq \epsilon\big). \tag{4.3}$$

Zu jedem $n \in \mathbb{N}$, und $\delta := \delta_n := 1/n$ wählen wir $x_n, y_n \in [a, b]$ wie in (4.3). Auf Grund des Satzes von Bolzano-Weierstraß (Korollar 1.35 in Kapitel 2) gibt es eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$, so dass

$$\lim_{k\to\infty} x_{n_k}$$

existiert. Sei $x := \lim_{k\to\infty} x_{n_k} \in [a, b]$. Wegen $|x_{n_k} - y_{n_k}| < 1/n_k$ gilt auch $x = \lim_{k\to\infty} y_{n_k}$. Da f stetig ist gilt dann auch

$$f(x) = \lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} f(y_{n_k}),$$

also

$$\lim_{k\to\infty} (f(x_{n_k}) - f(y_{n_k})) = 0.$$

Dies widerspricht jedoch der obigen Wahl von x_{n_k} und y_{n_k}, denn diese haben wir so gewählt, dass für alle $k \in \mathbb{N}$ die Ungleichung $|f(x_{n_k}) - f(y_{n_k})| \geq \epsilon$ gilt.[5]

Wir haben gezeigt, dass die Annahme, f sei nicht gleichmäßig stetig, einen Widerspruch impliziert. Also muss f gleichmäßig stetig sein. □

[5]Und natürlich für das feste oben gewählte $\epsilon > 0$!

Satz 4.4. Sei $f : [a, b] \to \mathbb{R}$ stetig. Dann ist f Riemann-integrierbar. In anderen Worten:

$$C^0([a, b]) \subset \mathcal{R}[a, b].$$

Beweis. Zu einem gegebenen $\epsilon > 0$ wähle $\delta > 0$ wie in (4.2). Dann wählen wir $k_\epsilon \in \mathbb{N} \setminus \{0\}$ so groß, dass $(b - a)/k_\epsilon < \delta$. Definiere ähnlich wie oben

$$t_j := a + j\frac{b - a}{k_\epsilon}.$$

Wir definieren die Treppenfunktionen $g_o, g_u \in \mathcal{T}[a, b]$, so dass für alle $j \in \{1, 2, \dots k\}$ und alle $x \in [t_{j-1}, t_j)$ gilt

$$g_o(x) = \sup\{f(x) \mid x \in [t_{j-1}, t_j)\}, \qquad g_u(x) = \inf\{f(x) \mid x \in [t_{j-1}, t_j)\}$$

und $g_o(b) = g_u(b) = f(b)$. Dann gilt $0 \le g_o(x) - g_u(x) \le \epsilon$, und deswegen

$$\int_a^b g_o(x)\,dx - \int_a^b g_u(x)\,dx \le \epsilon(b - a),$$

und offensichtlich $g_u \le f \le g_o$. Mit Lemma 2.3 folgt die Riemann-Integrierbarkeit. □

!ACHTUNG!. Wenn eine Funktion $f : (a, b) \to \mathbb{R}$ stetig ist, so kann man daraus **nicht** schließen, dass f Riemann-integrierbar ist. Beispiel: Die Funktion $f : (0, 1) \to \mathbb{R}$, $x \mapsto 1/x$, ist stetig, aber nicht beschränkt und somit auch nicht Riemann-intergierbar.

5. Hauptsatz der Differential- und Integralrechnung

Satz 5.1 (Hauptsatz der Differential- und Integralrechnung, Teil I)**.** Sei $f \in \mathcal{R}[a, b]$, $\gamma \in [a, b]$. Definiere

$$F : [a, b] \to \mathbb{R}, \qquad F(x) := \int_\gamma^x f(t)\,dt.$$

Dann gilt

(1) F ist stetig auf $[a, b]$.

(2) Ist f stetig in $x_0 \in (a, b)$, so ist F differenzierbar in x_0, und $F'(x_0) = f(x_0)$.

Fr 8.2.

Beweis.

(1) $f \in \mathcal{R}[a, b]$, also f beschränkt, somit

$$\|f\|_\infty \overset{(\text{def})}{=} \sup\left\{|f(x)| \,\middle|\, x \in [a, b]\right\} < \infty.$$

Sei $x, x_0 \in [a, b]$.

$$\left|F(x) - F(x_0)\right| = \left|\int_{x_0}^{x} f(t)\, dt\right| \leq \left|\int_{x_0}^{x} |f(t)|\, dt\right| \leq |x - x_0|\, \|f\|_\infty$$

Zu gegebenem $\epsilon > 0$ wählen wir nun $\delta := \epsilon/\|f\|_\infty > 0$ und dann ist das ϵ-δ-Kriterium für Stetigkeit in x_0 erfüllt.

(2)

$$\begin{aligned}
\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| &= \left|\frac{1}{x - x_0}\right| \left|\int_{x_0}^{x} (f(t) - f(x_0))\, dt\right| \\
&\leq \left|\frac{1}{x - x_0}\right| \left|\int_{x_0}^{x} |(f(t) - f(x_0))|\, dt\right| \\
&\leq \sup\left\{|f(t) - f(x_0)| \,\middle|\, t \in [x, x_0] \text{ bzw. } t \in [x_0, x]\right\}
\end{aligned}$$

Wir fixieren nun ein $\epsilon > 0$. Da f in x_0 stetig gibt es zu diesem $\epsilon > 0$ ein $\delta > 0$, so dass für alle $t \in [a, b]$ mit $|t - x_0| \leq \delta$ gilt:

$$|f(t) - f(x_0)| \leq \epsilon,$$

insbesondere gilt für $x \in [a, b] \cap [x_0 - \delta, x_0 + \delta]$:

$$\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| \leq \sup\left\{|f(t) - f(x_0)| \,\middle|\, t \in [x, x_0] \text{ bzw. } t \in [x_0, x]\right\} \leq \epsilon.$$

Daraus ergibt sich

$$\lim_{x\to x_0}\left(\frac{F(x)-F(x_0)}{x-x_0}-f(x_0)\right)=0$$

und somit die Behauptung im zweiten Teil.

□

Ist $I\neq\emptyset$ ein Intervall, so sei $\mathring{I}:=(\inf I,\sup I)$ das offene Intervall mit denselben Grenzen wie I.

Bsp: $I:=[a,b]$, $\mathring{I}=(a,b)$. Ist J ein offenes Intervall in $\mathbb{R}$, dann $\mathring{J}=J$.

Für offene Teilmengen U definieren wir $\mathring{U}:=U$.

Definition 5.2. Sei I ein Intervall oder eine offene Teilmenge von $\mathbb{R}$, und $f:I\to\mathbb{R}$. Wir sagen $F:I\to\mathbb{R}$ ist eine Stammfunktion von f, wenn F stetig ist, $F|_{\mathring{I}}$ differenzierbar ist und $F'(x)=f(x)$ für alle $x\in\mathring{I}$.

Beispiele 5.3.

(1) Der Hauptsatz der Diff.- und Int.-Rechnung I besagt, dass stetige Funktionen $f:[a,b]\to\mathbb{R}$ immer eine Stammfunktion haben.
(2) Auch stetige Funktionen $f:(c,d)\to\mathbb{R}$ besitzen immer eine Stammfunktion: wähle $\gamma\in(c,d)$, dann ist $F(x)=\int_\gamma^x f(t)\,dt$ eine Stammfunktion.
(3) Ist $U\subset\mathbb{R}$ offen und $f:U\to\mathbb{R}$ stetig, dann besitzt f ebenfalls eine Stammfunktion. Denn ist $U\subset\mathbb{R}$ offen, dann gibt es eine Familie offener Intervalle $(U_\alpha)_{\alpha\in A}$, so dass $U=\bigcup_{\alpha\in A}U_\alpha$, $U_\alpha\cap U_\beta=\emptyset$ für $\alpha\neq\beta$[6]. Man konstruiert nun auf jedem U_α die Stammfunktion wie in (2).

[6]Man sagt zu letzterer Eigenschaft: $(U_\alpha)_{\alpha\in A}$ ist eine Familie von paarweise disjunkten offenen Intervallen

(4) Es gibt Riemann-integrierbare Funktionen ohne Stammfunktion. Beispiel

$$f : [-1,1] \to \mathbb{R}, \quad f(x) := \begin{cases} -1 & \text{, falls } x \leq 0 \\ 1 & \text{, falls } x > 0 \end{cases}$$

$f \in \mathcal{R}[-1,1]$, da f eine Treppenfunktion ist. Angenommen f hätte eine Stammfunktion $F : [-1,1] \to \mathbb{R}$. Dann gilt $F'(x) = 1$ für $x > 0$, somit gibt es eine $c \in \mathbb{R}$, so dass für alle $x \in (0,1)$ gilt: $F(x) = c + x$. Es gilt $F'(x) = -1$ für $x \leq 0$, somit gibt es ein $d \in \mathbb{R}$ mir $F(x) = d - x$ für $d \leq 0$. Da F stetig ist, gilt $c = d$, also $F(x) = c + |x|$ für alle $x \in [-1,1]$. Diese Funktion ist nicht in 0 differenzierbar, also ist F keine Stammfunktion.

(5) Es gibt nicht-Riemann-integrierbare Funktionen mit Stammfunktion. Wir definieren:

$$F : \mathbb{R} \to \mathbb{R}, \quad F(x) := \begin{cases} x^2 \sin\left(\dfrac{1}{x^2}\right), & \text{falls } x > 0 \\ 0, & \text{falls } x \leq 0. \end{cases}$$

Dann ist F differenzierbar auf ganz $\mathbb{R}$ und somit ist $F|_{[-1,1]}$ eine Stammfunktion von $f := F'|_{[-1,1]} : [-1,1] \to \mathbb{R}$. Man rechnet nach für $j \to \infty$:

$$\begin{aligned} x_j &:= (2\pi j + \pi)^{-1/2} \to 0 \\ f(x_j) &= 2\frac{1}{\sqrt{2\pi j + \pi}} \underbrace{\sin(2\pi j + \pi)}_{=0} \\ &\quad - 2\sqrt{2\pi j + \pi} \underbrace{\cos(2\pi j + \pi)}_{=-1} \to \infty \end{aligned}$$

Deswegen ist f nicht beschränkt und deswegen auch nicht Riemann-integrierbar.

Satz 5.4 (Hauptsatz der Differential- und Integralrechnung, Teil II)**.** Sei $f : [a, b] \to \mathbb{R}$ Riemann-integierbar, und F sei eine Stammfunktion von f. Dann gilt für $x, y \in [a, b]$:

$$\int_y^x f(t)\,dt = F(x) - F(y). \tag{5.5}$$

Beweis.

1. Fall: f ist stetig.
$F_0(x) := \int_y^x f(t)\,dt$ ist eine Stammfunktion von f, und deswegen $F(x) = F_0(x) + F(y)$. Daraus folgt die Aussage.

2. Fall: allgemein
O.B.d.A. $y < x$. Sei $g_o \in \mathcal{T}[y, x]$, $g_o \geq f$, und sei $Z = \{t_0 < t_1 < \ldots < t_k\}$ eine zu g_o passende Zerlegung von $[y, x]$.

$$\begin{aligned} F(x) - F(y) &= \sum_{j=1}^{k} (F(t_{j-1}) - F(t_j)) \\ &\overset{\text{1.MWS}}{=} \sum_{j=1}^{k} (t_j - t_{j-1}) F'(\xi_j) \end{aligned}$$

für geeignete $\xi_j \in (t_{j-1}, t_j)$. Wir rechnen weiter

$$\begin{aligned} F(x) - F(y) &= \sum_{j=1}^{k} (t_j - t_{j-1}) f(\xi_j) \\ &\leq \sum_{j=1}^{k} (t_j - t_{j-1}) g_o(\xi_j) \\ &= \int_y^x g_o(t)\,dt \end{aligned}$$

Da dies für alle Treppenfunktionen oberhalb von f gilt, folgt durch Infimumsbildung:

$$F(x) - F(y) \leq \int_y^x f(t)\,dt$$

Die Gleichung

$$F(x) - F(y) \geq \int_y^x f(t)\,dt$$

zeigt man ganz analog. Da f Riemann-integrierbar ist, folgt die Aussage. □

Beispiele 5.6.

(1) Eine Stammfunktion der Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto \sum_{k=0}^{n} a_k x^k$ ist $F : \mathbb{R} \mapsto \mathbb{R}$, $x \mapsto \sum_{k=0}^{n} \frac{a_k}{k+1} x^{k+1}$.

(2) Die Stammfunktion der Funktion

$$f : \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \frac{1}{x^2+1}$$

ist $F = \arctan : \mathbb{R} \to (-\pi/2; \pi/2) \subset \mathbb{R}$, die Umkehr-Funktion des Tangens. Denn mit

$$\tan'(y) = \frac{1}{\cos^2 y} = \frac{\sin^2 y + \cos^2 y}{\cos^2 y} = \tan^2 y + 1$$

und mit Proposition 1.6 in Kapitel 5 sieht man

$$\arctan'(x) = \frac{1}{\tan^2(\arctan(x)) + 1} = \frac{1}{x^2+1}.$$

(3) Es gibt viele Tabellen mit vielen Stammfunktionen, siehe zum Beispiel:

`http://de.wikipedia.org/wiki/Tabelle_von_Ableitungs-_und_Stammfunktionen`

Notation 5.7.

$$F(x)\Big|_{x=a}^{b} := F(b) - F(a)$$

Dann schreibt sich (5.5) als

$$\int_a^b f(x)\,dx = F(x)\Big|_{x=a}^b.$$

$C^0([a,b]) := \{f : [a,b] \to \mathbb{R} \mid f \text{ stetig}\}$

$f \in C^1([a,b]) \iff f \in C^0([a,b])$, $f|_{(a,b)}$ differenzierbar und f' kann man zu einer stetigen Funktionen $f' : [a,b] \to \mathbb{R}$ fortsetzen.[7]

Satz 5.8 (Partielle Integration)**.** Seien $a, b \in \mathbb{R}$, $a < b$ und $f, g \in C^1([a,b])$. Dann

$$\int_a^b f'(x)g(x)\,dx = f(x)g(x)\Big|_{x=a}^b - \int_a^b f(x)g'(x)\,dx$$

Beweis. $x \mapsto f(x)g(x)$ ist Stammfunktion von $x \mapsto f'(x)g(x) + f(x)g'(x)$. □

Beispiel 5.9. Zu berechnen ist $A := \int_a^b \cos^3 x\,dx$. Setze $f(x) = \sin x$ und $g(x) = \cos^2 x$. Es folgt mit partieller Integration:

$$A = \int_a^b f'(x)g(x)\,dx = \sin x \cos^2 x\Big|_{x=a}^b - \int_a^b (\sin x)(-2)\sin x \cos x\,dx$$

Wenn wir nun $\sin^2 x = 1 - \cos^2 x$ nutzen, so ergibt sich

$$A = \sin x \cos^2 x\Big|_{x=a}^b + 2\int_a^b \cos x\,dx - 2A.$$

Also

$$A = \frac{1}{3}(\sin b \cos^2 b - \sin a \cos^2 a) - \frac{2}{3}(\sin b - \sin a).$$

Satz 5.10 (Integration durch Substitution)**.** Seien $a, b, c, d \in \mathbb{R}$, $a < b$ und $c < d$. Gegeben sei

[7] Wir bezeichnen diese Fortsetzung auch einfach mit f' und es gilt dann auch

$$f'(a) = \lim_{x \searrow a} \frac{f(x) - f(a)}{x - a}, \qquad f'(b) = \lim_{x \nearrow b} \frac{f(x) - f(b)}{x - b}.$$

- $f \in C^0([a,b])$,
- $\varphi \in C^1([c,d])$ mit $\varphi_\#([c,d]) \subset [a,b]$

Dann gilt

$$\int_{\varphi(c)}^{\varphi(d)} f(s)\,ds = \int_c^d f(\varphi(t))\varphi'(t)\,dt.$$

Beweis. Die Integranden, d.h. die Funktionen f und $t \mapsto f(\varphi(t))\varphi'(t) =: \tilde{f}(t)$ sind stetig, also Riemann-integrierbar und besitzen Stammfunktionen. Sei $F : [a,b] \to \mathbb{R}$ eine Stammfunktion von f, dann ist nach Kettenregel $F \circ \varphi : [c,d] \to \mathbb{R}$ eine Stammfunktion von $\tilde{f}$.

$$\int_{\varphi(c)}^{\varphi(d)} f(s)\,ds = F(\varphi(d)) - F(\varphi(c)) = \int_c^d f(\varphi(t))\varphi'(t)\,dt.$$

□

Merkregel: Die Leibnizsche Differentialschreibweise (= Verwendung von den „unendlich kleinen“ Zahlen ds, dt) liefert eine gute Merkregel. [8]

$$s = \varphi(t),$$

also auch $f(\varphi(t)) = f(s)$. Schreibe

$$\varphi'(t) = \lim_{\tilde{t}\to t} \frac{\varphi(\tilde{t}) - \varphi(t)}{\tilde{t} - t} = \lim_{\tilde{t}\to t} \frac{\tilde{s} - s}{\tilde{t} - t} = \frac{ds}{dt}$$

Man löst auf

$$ds = \varphi'(t)\,dt$$

und ersetzt die t-Integrationsgrenzen c und d durch die s-Integrationsgrenzen $\varphi(c)$ und $\varphi(d)$. Und man hat die Subsitutionsformel.

[8] Sie ist nicht mathematisch präzise, denn es bleibt ja unklar, was eine unendlich kleine Zahl denn sein soll. Sie ist aber sehr effizient als Merkregel.

Typische Anwendungen.

1.)

$$\int_a^b te^{t^2}\,dt = \frac{1}{2}\int_{a^2}^{b^2} e^s\,ds = \frac{1}{2}e^{b^2} - \frac{1}{2}e^{a^2}$$

Hier wurde $f(s) = e^s$ und $\varphi(t) = t^2$ genutzt.

2.) Sei $0 < a < b$.

$$\int_a^b \frac{1}{(e^t-1)(e^t+1)}\,dt = \int_a^b \frac{1}{e^t(e^t-1)(e^t+1)}e^t\,dt = \int_{e^a}^{e^b} \frac{1}{s(s-1)(s+1)}\,ds = \text{ siehe Zentralübung}$$

Hier wurde $f(s) = \frac{1}{s(s-1)(s+1)}$ und $\varphi(t) = e^t$ genutzt.

Wir machen einen Partialsummenansatz

$$\frac{1}{s(s-1)(s+1)} = \frac{a}{s} + \frac{b}{s-1} + \frac{c}{s+1}$$

und rechnen aus, dass dieser Ansatz für $a = 1$ und $b = c = -1/2$ funktioniert. somit hat $f : \mathbb{R}\setminus$ die Stammfunktion

$$\log(s) - \frac{1}{2}\log(s-1) - \frac{1}{2}\log(s+1)$$

und damit kann dieses Integral explizit berechnet werden.

6. Uneigentliche Riemann-Integrale

Wdh.: $F : \mathbb{R} \longrightarrow \mathbb{R}$, $a \in \mathbb{R}$

$$a = \lim_{x\to\infty} F(x) \quad \Longleftrightarrow \quad \forall \epsilon \in \mathbb{R}_{>0} : \exists x_0 \in \mathbb{R} : \forall x \in \mathbb{R}_{\geq x_0} : |F(x) - a| < \epsilon$$

Analog für $F : [a, \infty) \longrightarrow \mathbb{R}$ und andere Definitionsbereiche.

DEFINITION 6.1 (Uneigentliches Riemann-Integral, oberes Ende). Sei $a \in \mathbb{R}$, $b \in \mathbb{R} \cup \{\infty\}$, $a < b$, $f : [a, b) \longrightarrow \mathbb{R}$.
Wir sagen f ist (am oberen Ende) uneigentlich Riemann-integrierbar falls $f|_{[a,\beta]} \in \mathcal{R}[a, \beta]$ für alle $\beta \in (a, b)$ und falls

$$\int_a^b f(x)\,dx := \lim_{\beta \nearrow b} \int_a^\beta f(x)\,dx \tag{6.2}$$

existiert. Wir sagen dazu auch $\int_a^b f(x)\,dx$ konvergiert. Wir sagen: das Integral konvergiert absolut, falls $\int_a^b |f(x)|\,dx$ konvergiert.

Man beachte: die linke Seite von (6.2) ist eine Definition!

Spezialfall: Wenn $b < \infty$ und falls $f = \tilde{f}_{[a,b)}$ für ein $\tilde{f} \in \mathcal{R}[a, b]$ ist, dann gilt wegen dem Hauptsatz der Differential- und Integralrechnung[9]:

$$\int_a^b \tilde{f}(x)\,dx = \lim_{\beta \to b} \int_a^\beta \tilde{f}(x)\,dx$$

In diesem Fall gilt also

$$\int_a^b \tilde{f}(x)\,dx = \int_a^b f(x)\,dx.$$

DEFINITION 6.3 (Uneigentliches Riemann-Integral, unteres Ende und an beiden Enden). Sei $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R}$, $a < b$, $f : (a, b] \longrightarrow \mathbb{R}$.
Wir sagen f ist (am unteren Ende) uneigentlich Riemann-integrierbar falls $f|_{(\alpha,b]} \in \mathcal{R}[\alpha, b]$ für alle $\alpha \in (a, b)$ und falls

$$\int_a^b f(x)\,dx := \lim_{\alpha \searrow a} \int_\alpha^b f(x)\,dx \tag{6.4}$$

existiert.

[9]Die folgende Zeile ist keine Definition, sondern eine Aussage!

Sei $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$, $a < b$, $f : (a, b) \longrightarrow \mathbb{R}$. Wir sagen f ist (an beiden Enden) uneigentlich Riemann-integrierbar, falls für ein $\gamma \in (a, b)$ die Summe

$$\int_a^b f(x)\,dx := \int_a^\gamma f(x)\,dx + \int_\gamma^b f(x)\,dx \tag{6.5}$$

existiert.[10]

Beispiele 6.6.

(1) $\int_1^\infty x^a\,dx$ konvergiert genau dann, wenn $a < -1$. Denn

$$F(r) := \int_1^r x^a\,dx = \begin{cases} \frac{1}{a+1}(r^{a+1} - 1) & \text{für } a \neq -1 \\ \log r & \text{für } a = -1 \end{cases}$$

Im Limes $r \to \infty$ gilt

$$F(r) \to \begin{cases} \infty & \text{für } a \geq -1 \\ -\frac{1}{a+1} & \text{für } a < -1 \end{cases}$$

(2) $\int_0^1 x^a\,dx$ konvergiert genau dann, wenn $a > -1$.

ENDE DER VORLESUNG ANALYSIS I

[10]Hierzu müssen beide Integrale der rechten Seite von (6.5) konvergieren. Die Existenz und der Wert von $\int_a^b f(x)\,dx$ ist dann unabhängig von γ.

KAPITEL 7

Stetigkeit und Differenzierbarkeit von Grenzwerten und Reihen

24.4.

Vereinfachte Notationen ab jetzt:

- Quantoren hinter Aussagen sind erlaubt, falls keine Missverständnisse zu erwarten sind.
- Wir schreiben $\mathrm{im}(f)$ für $B(f)$,
- $f(A)$ für $f_{\#}(A)$,
- und $f^{-1}(A)$ für $f^{\#}(A)$.
- Wir schreiben oft $\forall \epsilon > 0$ statt $\forall \epsilon \in \mathbb{R}_{>0}$, analog $\exists \delta > 0, \ldots$

1. Metrische Räume: Wiederholung und Cauchy-Folgen

Ein metrischer Raum ist ein Paar (M, d), wobei die Metrik $d : M \times M \longrightarrow \mathbb{R}_{\geq 0}$ die folgenden Eigenschaften für alle $x, y, z \in M$ erfüllen möge

(a) $d(x, y) = 0 \iff x = y$ (Definitheit)
(b) $d(x, y) = d(y, x)$ (Symmetrie)
(c) $d(x, y) \leq d(x, z) + d(z, y)$ (Dreiecks-Ungleichung)

$r \in \mathbb{R}_{\geq 0}$, $x \in M$
offener Ball $B_r(x) := \{y \in M \mid d(x, y) < r\}$

abgeschlossener Ball $\overline{B}_r(x) := \{y \in M \mid d(x,y) \leq r\}$

Sei $(x_n)_{n\in\mathbb{N}}$ eine M-wertige Folge, $a \in M$.

Konvergenz.

$$\begin{aligned} & (x_n)_{n\in\mathbb{N}} \text{ konvergiert gegen } a \\ :\Longleftrightarrow\ & \lim_{n\to\infty} x_n = a \\ :\Longleftrightarrow\ & \lim_{n\to\infty} d(x_n, a) = 0 \\ \Longleftrightarrow\ & \forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall n \in \mathbb{N}_{\geq N} : d(x_n, a) < \epsilon \\ \Longleftrightarrow\ & \forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall n \in \mathbb{N}_{\geq N} : x_n \in B_\epsilon(a) \end{aligned}$$

Cauchy-Folge.

$$\begin{aligned} & (x_n)_{n\in\mathbb{N}} \text{ ist eine Cauchy-Folge} \\ \Longleftrightarrow\ & \forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall n, m \in \mathbb{N}_{\geq N} : d(x_n, x_m) < \epsilon \end{aligned}$$

Satz 1.1. Jede konvergente Folge ist eine Cauchy-Folge.

Den Beweis erhält man aus dem Beweis von Satz 1.38 aus Kapitel 3, indem man $|x-y|$ durch $d(x,y)$ ersetzt.

Beweis. Es gelte $\lim_{j\to\infty} a_j = a$. Das heißt: für alle $\epsilon \in \mathbb{R}_{>0}$ gibt es ein $j_0 \in \mathbb{N}$, so dass für alle natürlichen Zahlen $j \geq j_0$ gilt: $d(a_j, a) \leq \epsilon$.

Für solch ein ϵ und ein passendes j_0 nehmen wir nun natürliche Zahlen $j \geq j_0$ und $k \geq j_0$ und rechnen nach:

$$d(a_j, a_k) \leq d(a_j, a) + d(a, a_k) \leq \epsilon + \epsilon = 2\epsilon.$$

Wir haben nun also gezeigt:

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists j_0 \in \mathbb{N} : \forall j, k \in \mathbb{N}_{\geq j_0} : d(a_j, a_k) \leq 2\epsilon.$$

Wegen Lemma 1.11 aus Kapitel 3 ist dies äquivalent zur definierenden Eigenschaft einer Cauchy-Folge. □

DEFINITION 1.2. Ein metrischer Raum (M, d) heißt vollständig, wenn jede M-wertige Cauchy-Folge in (M, d) konvergiert.

Ist (M, d) ein metrischer Raum und N eine Teilmenge. Dann ist $d|_{N\times N}$ eine Metrik auf N, die von (M, d) auf N induzierte Metrik, siehe Beispiel 5.3 in Kapitel 4.

ÜBUNG 1.3. Sei (M, d) ein metrischer Raum. Dann ist $N \subset M$ abgeschlossen genau dann, wenn für alle N-wertigen Folgen, die in (M, d) konvergieren, der Grenzwert in N liegt.

ÜBUNG 1.4. Sei N eine Teilmenge eines metrischen Raumes (M, d) und sei $\tilde{d}$ die induzierte Metrik auf N.

(a) Ist $(N, \tilde{d})$ vollständig, so ist N abgeschlossen in M.
(b) Ist (M, d) vollständig, so gilt

$$(N, \tilde{d}) \text{ vollständig} \quad \Longleftrightarrow \quad N \text{ abgeschlossen in } M.$$

2. Punktweise und gleichmäßige Konvergenz

Im folgenden sei (Y, d) ein metrischer Raum

Die wichtigsten Beispiele, an die Sie im folgenden denken sollten, sind $Y = \mathbb{R}$ und $Y = \mathbb{C}$ mit $d(x, y) := |x - y|$. Später werden wir zum Beispiel für Y auch Vektorräume (mit einer Norm) oder noch andere Räume betrachten. Wir betrachten nun Funktionen $D \longrightarrow Y$.

DEFINITION 2.1. Eine Folge von Funktionen besteht aus einer Menge D, einem metrischen Raum (Y, d), und einer Folge $(f_n)_{n\in\mathbb{N}}$, wobei wir für jedes $n \in \mathbb{N}$ eine Funktion $f_n : D \longrightarrow Y$ haben.
Sei $f : D \longrightarrow Y$ eine weitere Funktion. Wir sagen dann: (f_n) konvergiert punktweise gegen $f : D \longrightarrow Y$, falls für alle $x \in D$ gilt:

$$\lim_{n\to\infty} f_n(x) = f(x).$$

Der Begriff der punktweisen Konvergenz erscheint ein naheliegender Begriff für die Konvergenz von solchen Funktionen.

BEISPIEL 2.2. $Y := \mathbb{C}$. Gegeben sei eine Potenzreihe $p(x) := \sum_{j=0}^{\infty} a_j x^j$ mit Konvergenzradius ρ.

$$f_n(x) := \sum_{j=0}^{n} a_j x^j.$$

Dann konvergiert $(f_n : B_\rho(0) \longrightarrow \mathbb{C})_{n\in\mathbb{N}}$ punktweise gegen $p : B_\rho(0) \longrightarrow \mathbb{C}$.

Da nun alle f_n stetig sind, kann man hoffen, dass wir dadurch zeigen können, dass $p : B_\rho(0) \longrightarrow \mathbb{C}$ stetig ist. Hierbei tritt aber ein Problem auf, das am folgenden Beispiel deutlich wird.

BEISPIEL 2.3. $D = [0,1]$, $Y := \mathbb{R}$, $f_n(x) := x^n$.

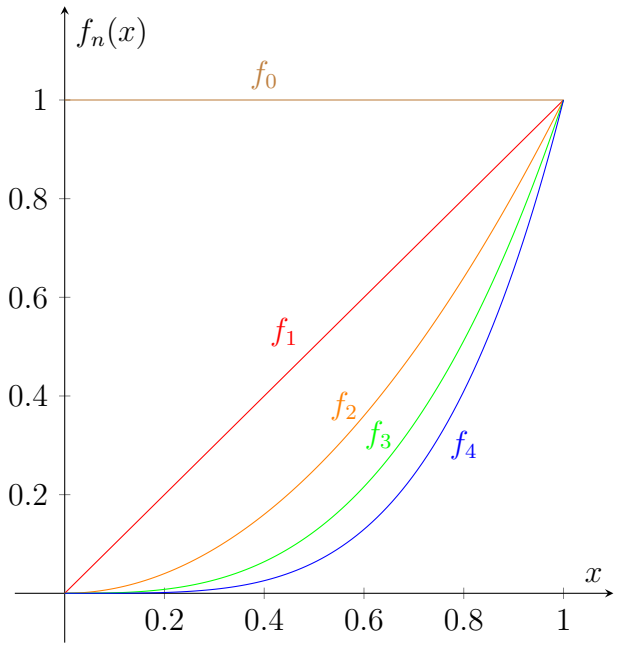


Abbildung 7: Die Graphen der Funktionen f_n

$$f_n(x) \to f(x) := \begin{cases} 0 & \text{für } x < 1 \\ 1 & \text{für } x = 1 \end{cases}$$

(f_n) konvergiert punktweise gegen $f : [0,1] \longrightarrow \mathbb{R}$. Alle Funktionen $f_n : [0,1] \longrightarrow \mathbb{R}$ sind stetig. Dennoch ist die Grenzfunktion f nicht stetig.

Lösung des Problems: Gleichmäßige Konvergenz.

Wir machen aus $\mathrm{Abb}(D,Y) = Y^D$ einen metrischen Raum. Für $f,g \in \mathrm{Abb}(D,Y)$ definieren wir die Supremums-Distanz

$$d_{\sup}(f,g) := \sup_{x\in D}\Big\{ d\big(f(x),g(x)\big)\Big\} \in [0,\infty].$$

$(\mathrm{Abb}(D,Y), d_{\sup})$ erfüllt (a) Definitheit, (b) Symmetrie, (c) Dreiecks-Ungleichung; aber es ist keine Norm, da $d_{\sup}(f,g) = \infty$ möglich ist. Wir setzen deswegen

$$d_{\mathrm{glm}}(f,g) := \min\{d_{\sup}(f,g), 289\}.$$

Dann ist $(\mathrm{Abb}(D,Y), d_{\mathrm{glm}})$ ein metrischer Raum[1].

Definition 2.4. Sei $(f_n : D \longrightarrow Y)_{n\in\mathbb{N}}$ eine Folge von Funktionen. Wir definieren

$$\begin{aligned}
&(f_n) \text{ konvergiert gleichmäßig gegen } f : D \longrightarrow Y \\
:\Longleftrightarrow\ &(f_n) \text{ konvergiert gegen } f \text{ im metrischen Raum } (\mathrm{Abb}(D,Y), d_{\mathrm{glm}}) \\
\Longleftrightarrow\ &d_{\mathrm{glm}}(f_n,f) \to 0 \text{ für } n\to\infty \\
\Longleftrightarrow\ &\sup_{x\in D}\Big\{ d\big(f(x),g(x)\big)\Big\} \to 0 \text{ für } n\to\infty \\
\Longleftrightarrow\ &\forall \epsilon\in\mathbb{R}_{>0} : \exists n_0\in\mathbb{N} : \forall n\in\mathbb{N}_{\geq n_0} : \forall x\in D : d(f_n(x),f(x)) \leq \epsilon \\
\overset{Y\subset\mathbb{C}}{\Longleftrightarrow}\ &\forall \epsilon\in\mathbb{R}_{>0} : \exists n_0\in\mathbb{N} : \forall n\in\mathbb{N}_{\geq n_0} : \forall x\in D : |f_n(x)-f(x)| \leq \epsilon
\end{aligned}$$

Logische Verknüpfung

$$\exists x\in M : \forall y\in N : A(x,y) \quad\Longrightarrow\quad \forall y\in N : \exists x\in M : A(x,y)$$

[1]Man kann in dieser Konstruktion natürlich 289 durch jede andere positive reelle Zahl ersetzen.

Somit:

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists n_0 \in \mathbb{N} : \forall n \in \mathbb{N}_{\geq n_0} : \forall x \in D : d(f_n(x), f(x)) \leq \epsilon$$
$$\implies \forall x \in D : \forall \epsilon \in \mathbb{R}_{>0} : \exists n_0 \in \mathbb{N} : \forall n \in \mathbb{N}_{\geq n_0} : d(f_n(x), f(x)) \leq \epsilon$$

Also:

$$\text{Gleichmäßige Konvergenz} \implies \text{Punktweise Konvergenz}$$

!ACHTUNG!. Die Umkehrung ist falsch. Punktweise Konvergenz impliziert nicht gleichmäßige. In Beispiel 2.3 konvergiert (f_n) punktweise gegen f, aber nicht gleichmäßig.

Von nun an sei D ein metrischer Raum mit Metrik d_D, zum Beispiel ein Intervall, eine Teilmenge von $\mathbb{R}^n$ oder eine Teilmenge von $\mathbb{C}^n$ mit der induzierten Metrik.

Satz 2.5. Sei $(f_n : D \longrightarrow Y)_{n\in\mathbb{N}}$ eine Folge von stetigen Funktionen, die gleichmäßig gegen $f : D \longrightarrow Y$ konvergiert. Dann ist auch $f : D \longrightarrow Y$ stetig.

Beispiel 2.6 (Fortsetzung Beispiel 2.2). Sei wieder ρ der Konvergenzradius von $p(x) = \sum_{j=0}^{\infty} a_j x^j$, $0 < R < \rho$. Definiere die Partialsummen

$$f_n(x) := \sum_{j=0}^{n} a_j x^j.$$

Dann gilt für $x \in \overline{B}_R(0)$:

$$|f_n(x) - p(x)| = \Big| \sum_{j=n+1}^{\infty} a_j x^j \Big| \leq \sum_{j=n+1}^{\infty} |a_j||x|^j \leq \sum_{j=n+1}^{\infty} |a_j| R^j \to 0.$$

Wir haben Konvergenz gegen 0, da $\sum_{j=0}^{\infty} a_j R^j$ absolut konvergiert. Also konvergiert $(f_n : \overline{B}_R(0) \longrightarrow \mathbb{C})_{n\in\mathbb{N}}$ gleichmäßig gegen $p : \overline{B}_R(0) \longrightarrow \mathbb{C}$. [2] Wir sehen also somit, dass

[2] Achtung: Für viele Potenzreihen konvergiert $f_n' : B_\rho(0) \to \mathbb{C}$ nicht gleichmäßig. Erst nach Einschränkung auf einen etwas kleineren Ball $\overline{B}_R(0)$ erhalten wir gleichmäßige Konvergenz.

$p|_{\overline{B}_R(0)} : \overline{B}_R(0) \longrightarrow \mathbb{C}$ stetig ist. Da dies für alle $R \in (0, \rho)$ gilt, sehen wir, dass $p|_{B_\rho(0)} : B_\rho(0) \longrightarrow \mathbb{C}$ ebenfalls stetig ist.

26.4.

Beweis des Satzes. Damit unsere Notation einfach ist, beschränken wir uns auf den Fall $D \subset \mathbb{C}$, $Y \subset \mathbb{C}$, $d_D(x, y) = |x - y|$, $d(x, y) = |x - y|$. Den allgemeinen Fall beweist man völlig analog.

Wir zeigen die Stetigkeit von f in x_0. Sei $\epsilon \in \mathbb{R}_{>0}$ gegeben. Wir wählen hierzu ein $N \in \mathbb{N}$, so dass für alle natürlichen Zahlen $n \geq N$ gilt:

$$d_{\mathrm{glm}}(f, f_n) < \tilde{\epsilon} := \frac{\epsilon}{3}$$

Da f_N stetig in x_0 stetig ist, gibt es ein $\delta \in \mathbb{R}_{>0}$, so dass

$$|x - x_0| < \delta \Longrightarrow |f_N(x) - f_N(x_0)| < \tilde{\epsilon}.$$

Wir schließen, dass für $x \in D$ mit $|x - x_0| < \delta$ gilt:

$$|f(x) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| < 3\tilde{\epsilon} = \epsilon.$$

Wir haben also die Stetigkeit von f in x_0 geprüft. Da x_0 ein beliebiges Element von D ist, ist f stetig (auf ganz D). □

Beweis des Satzes (volle Allgemeinheit). Wir zeigen die Stetigkeit von f in x_0. Sei $\epsilon \in \mathbb{R}_{>0}$ gegeben. Wir wählen hierzu ein $N \in \mathbb{N}$, so dass für alle natürlichen Zahlen $n \geq N$ gilt:

$$d_{\mathrm{glm}}(f, f_n) < \tilde{\epsilon} := \frac{\epsilon}{3}$$

Da f_N stetig in x_0 stetig ist, gibt es ein $\delta \in \mathbb{R}_{>0}$, so dass

$$d_D(x, x_0) < \delta \Longrightarrow d(f_N(x), f_N(x_0)) < \tilde{\epsilon}.$$

Wir schließen, dass für $x \in D$ mit $d(x, x_0) < \delta$ gilt:

$$d(f(x), f(x_0)) \leq d(f(x), f_N(x)) + d(f_N(x), f_N(x_0)) + d(f_N(x_0), f(x_0)) < 3\tilde{\epsilon} = \epsilon.$$

Wir haben also die Stetigkeit von f in x_0 geprüft. Da x_0 ein beliebiges Element von D ist, ist f stetig (auf ganz D). □

LEMMA 2.7. Sei D eine Menge und (Y, d) ein vollständiger metrischer Raum.[3] Sei $(f_n : D \to Y)_{n\in\mathbb{N}}$ eine Folge von Funktionen. Dann sind die folgenden Aussagen äquivalent:

(1) $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig (gegen eine Funktion $D \to \mathbb{R}$)

(2)

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists n_0 \in \mathbb{N} : \forall j, k \in \mathbb{N}_{\geq n_0} : d_{\text{glm}}(f_j, f_k) < \epsilon$$

(3)

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists n_0 \in \mathbb{N} : \forall j, k \in \mathbb{N}_{\geq n_0} : \forall x \in D : d(f_j(x), f_k(x)) < \epsilon$$

Im Fall $D = \{x_0\}$ steht hier: $(f_n(x_0))$ konvergiert genau dann, wenn $(f_n(x_0))$ eine Cauchy-Folge ist.

Offensichtlich ist:

(1) $\iff$ $(f_n)_{n\in\mathbb{N}}$ konvergiert in $(\text{Abb}(D, Y), d_{\text{glm}})$.

(2) $\iff$ $(f_n)_{n\in\mathbb{N}}$ ist Cauchy-Folge in $(\text{Abb}(D, Y), d_{\text{glm}})$.

Beweis. „(2)$\iff$(3)“ folgt direkt aus der Definition von d_{glm}.

„(1)$\Longrightarrow$(2)“: jede konvergente Folge ist eine Cauchy-Folge (Satz 1.1).

[3]Wichtig sind hier vor allem die Fälle $Y = \mathbb{R}$, $X = \mathbb{C}$ und $Y = \mathbb{R}^k$ mit der Standardmetrik.

„(3)$\Longrightarrow$(1)“: Wegen (3) ist $(f_n(x))$ eine Cauchy-Folge, sei $f(x) := \lim_{n\to\infty} f_n(x)$, $f : D \to \mathbb{R}$. Zu gegebenem $\epsilon > 0$ wähle man ein n_0 wie in (3). Für $j, k \in \mathbb{N}_{\geq n_0}$:

$$d(f(x), f_j(x)) \leq d(f(x), f_k(x)) + d(f_k(x), f_j(x)) < \underbrace{d(f(x), f_k(x))}_{\to 0} + \epsilon$$

für $k \to \infty$. Somit $d(f(x), f_j(x)) \leq \epsilon$.

Da dies für alle $x \in D$ gilt, haben wir

$$d_{\mathrm{glm}}(f, f_j) \leq \epsilon.$$

Wir haben also gezeigt:

$$\forall \epsilon > 0 : \exists n_0 \in \mathbb{N} : \forall j \in \mathbb{N}_{\geq n_0} : d_{\mathrm{glm}}(f_j, f) \leq \epsilon$$

und dies ist (1). □

KOROLLAR 2.8. Sei D eine nicht-leere Menge und (Y, d) ein metrischer Raum. Dann ist $(\mathrm{Abb}(D, Y), d_{\mathrm{glm}})$ genau dann vollständig, wenn (Y, d) vollständig ist.

Beweis. Wegen Lemma 2.7 (2)$\Longrightarrow$(1) folgt aus der Vollständigkeit von (Y, d) die Vollständigkeit von $(\mathrm{Abb}(D, Y), d_{\mathrm{glm}})$.

Umgekehrt kann jedes $y \in Y$ als konstante Funktion $\underline{y} : D \to Y$, $x \mapsto y$ betrachtet werden, die wir mit $\underline{y}$ bezeichnen. Wegen $d_{\mathrm{glm}}(\underline{x}, \underline{y}) = d(x, y)$ bildet $D \to \mathrm{Abb}(D, Y)$, $y \mapsto \underline{y}$ Cauchy-Folgen auf Cauchy-Folgen ab. Grenzwerte konstanter Funktionen in $(\mathrm{Abb}(D, Y), d_{\mathrm{glm}})$ sind wiederum konstant. Aus der Vollständigkeit von $(\mathrm{Abb}(D, Y), d_{\mathrm{glm}})$ folgt dann auch die Vollständigkeit von (Y, d). □

3. Differentiation von Folgen und Reihen

Sei $f_n : (a,b) \longrightarrow \mathbb{R}$ eine Folge differenzierbarer Funktionen, die gleichmäßig gegen $f : (a,b) \longrightarrow \mathbb{R}$ konvergiert.

FRAGE 3.1. Ist dann f ebenfalls differenzierbar?

BEISPIELE 3.2.

(1) $f_n(x) := \sqrt{x^2 + \frac{1}{n}}$ konvergiert gleichmäßig gegen $f(x) = |x|$.
Denn für reelle Zahlen $a, b \in \mathbb{R}$ gilt

$$a^2 + b^2 \leq |a|^2 + |b|^2 + 2|a|\,|b| = (|a| + |b|)^2$$

und somit

$$\sqrt{a^2 + b^2} \leq |a| + |b|.$$

Wenden wir dies auf $a := x$ und $b := 1/\sqrt{n}$ an, so erhalten wir

$$|x| \leq f_n(x) \leq |x| + \frac{1}{\sqrt{n}}.$$

Also $|f_n(x) - f(x)| \leq \frac{1}{\sqrt{n}}$, d.h. $d_{\mathrm{glm}}(f_n, f) \leq n^{-1/2}$. Nun sind alle Funktionen f_n differenzierbar (auf $\mathbb{R}$), aber f ist in 0 nicht differenzierbar. Siehe auch [**22**, Abschnitt 9.5].

(2) Wir betrachten die Reihe

$$f_n : \mathbb{R} \to \mathbb{R}, \qquad f_n(x) := \sum_{\ell=1}^{n} \frac{1}{\ell^2} \sin(\ell^2 x)$$

$(f_n : \mathbb{R} \to \mathbb{R})_{n\in\mathbb{N}}$ konvergiert gleichmäßig, denn

$$|f_j(x) - f_k(x)| = |\sum_{\ell=k+1}^{j} \frac{1}{\ell^2}\sin(\ell^2 x)| \le \sum_{\ell=k+1}^{j} \frac{1}{\ell^2} \le \sum_{\ell=k+1}^{\infty} \frac{1}{\ell^2} \to 0 \qquad \text{für } j \ge k, \quad k \to \infty$$

Gliedweises Differenzieren liefert:

$$f_n'(x) = \sum_{\ell=1}^{n} \cos(\ell^2 x)$$

Diese Reihe ist divergent, u.a. da für viele $x \in \mathbb{R}$, z.B. für $x = 1$ die Folge $(\cos(\ell^2 x))_\ell$ keine Nullfolge ist.

THEOREM 3.3. Seien $a, b, x_0 \in \mathbb{R}$ mit $a < x_0 < b$. Sei $f_n : (a, b) \longrightarrow \mathbb{R}$ eine Folge von differenzierbaren Funktionen, so dass

(1) $(f_n(x_0))_{n\in\mathbb{N}}$ konvergiert (Konvergenz in einem Punkt)
(2) $(f_n')_{n\in\mathbb{N}}$ konvergiert gleichmäßig

Dann gilt:

(i) $f_n : (a, b) \longrightarrow \mathbb{R}$ konvergiert gleichmäßig gegen eine Funktion, die wir $f : (a, b) \longrightarrow \mathbb{R}$ nennen.
(ii) f ist differenzierbar auf (a, b)
(iii) f_n' konvergiert gegen f'

Beweis. Sei $x_1 \in (a, b)$ und $j, k \in \mathbb{N}$.

$$I := \begin{cases} [x_0, x_1] & \text{falls } x_1 \ge x_0, \\ [x_1, x_0] & \text{falls } x_1 < x_0. \end{cases}$$

Wir wenden den 1. Mittelwertsatz auf die Funktion $I \to \mathbb{R}$, $x \mapsto f_j(x) - f_k(x)$ an: Es existiert ein $\xi \in I \subset (a,b)$, so dass

$$(f_j(x_0) - f_k(x_0)) - (f_j(x_1) - f_k(x_1)) = (x_0 - x_1)(f_j'(\xi) - f_k'(\xi)).$$

Dies ergibt

$$|f_j(x_1) - f_k(x_1)| \leq |x_0 - x_1|\,|f_j'(\xi) - f_k'(\xi)| + |f_j(x_0) - f_k(x_0)|$$

Wir wählen für ein gegebenes $\epsilon > 0$ die Zahl n_0 so groß, dass für alle $j,k \in \mathbb{N}_{\geq n_0}$ gilt: $|f_j(x_0) - f_k(x_0)| \leq \epsilon/2$. Dies ist möglich, da $(f_n(x_0))_{n\in\mathbb{N}}$ eine Cauchy-Folge ist, siehe (1).

Wir wählen nun eine Zahl n_1, so dass für alle $j,k \in \mathbb{N}_{\geq n_1}$ und für alle $x \in (a,b)$ gilt:

$$|f_j'(x) - f_k'(x)| < \frac{\epsilon}{2(b-a)}.$$

Dies ist möglich, da f_j' gleichmäßig konvergiert, siehe (2).

Somit gilt für $n_2 := \max\{n_0, n_1\}$ für alle $j,k \in \mathbb{N}_{\geq n_2}$ und alle $x_1 \in (a,b)$ die Ungleichung

$$|f_j(x_1) - f_k(x_1)| \leq |x_0 - x_1| \underbrace{|f_j'(\xi) - f_k'(\xi)|}_{<\frac{\epsilon}{2(b-a)}} + \underbrace{|f_j(x_0) - f_k(x_0)|}_{<\epsilon/2} < \frac{\epsilon|b-a|}{2(b-a)} + \frac{\epsilon}{2} = \epsilon$$

Also $d_{\mathrm{glm}}(f_j, f_k) \leq \epsilon$, das heißt $(f_n)_{n\in\mathbb{N}}$ ist eine Cauchy-Folge in $(\mathrm{Abb}((a,b),\mathbb{R}), d_{\mathrm{glm}})$.

Da $\mathbb{R}$ vollständig ist, ist nach Korollar 2.8 auch $(\mathrm{Abb}((a,b),\mathbb{R}), d_{\mathrm{glm}})$ vollständig; es existiert also eine Funktion $f : (a,b) \to \mathbb{R}$, so dass (f_n) gegen f gleichmäßig konvergiert. Es folgt Aussage (i). **Fr. 3.5.**

Sei nun $g(x) := \lim_{n\to\infty} f_n'(x)$. Dieser Limes ist gleichmäßig nach Voraussetzung. Wähle zu gegebenem $\epsilon > 0$ ein $n_3 \in \mathbb{N}$, so dass für $j,k \in \mathbb{N}_{\geq n_3}$ gilt: $d_{\mathrm{glm}}(f_j', f_k') < \epsilon$ und $d_{\mathrm{glm}}(f_j', g) < \epsilon$.

Für solche j und k wenden wir ähnlich wie oben wieder den 1. Mittelwertsatz auf die Funktion $x \mapsto f_j(x) - f_k(x)$ an.

$$\left| \frac{(f_j(x) - f_k(x)) - (f_j(x_2) - f_k(x_2))}{x - x_2} \right| \overset{1.\ \text{MWS}}{=} |f_j'(\xi) - f_k'(\xi)| < \epsilon$$

für ein ξ zwischen x und $_2$. Im Grenzwert $k \to \infty$ erhalten wir

$$\left| \frac{(f_j(x) - f(x)) - (f_j(x_2) - f(x_2))}{x - x_2} \right| \leq \epsilon. \tag{3.4}$$

Diese Abschätzung gilt für alle $x \in (a, b)$, $x \neq x_2$, $j \geq n_3$; und n_3 hängt nicht von x ab, sondern nur von ϵ. Manchmal schreibt man hierfür kurz $n_3 = n_3(\epsilon)$, um auszudrücken, dass n_3 nur von ϵ abhängt, auch wenn diese Formulierung leicht falsch verstanden werden kann.

Da die Funktion f_j in x_2 differenzierbar ist, gibt es [4] zu oben gegebenem $\epsilon > 0$ und $j \geq n_3(\epsilon)$ ein $\delta_j > 0$, so dass für alle $x \in (x_2 - \delta_j, x_2 + \delta_j)$ gilt:

$$\left| \frac{f_j(x) - f_j(x_2)}{x - x_2} - f_j'(x_2) \right| \leq \epsilon. \tag{3.5}$$

Aus (3.4), (3.5) und $d_{\text{glm}}(f_j', g) \leq \epsilon$ folgt für $j \geq n_3(\epsilon)$

$$\forall x \in (x_2 - \delta_j, x_2 + \delta_j) : \left| \frac{f(x) - f(x_2)}{x - x_2} - g(x_2) \right| \leq 3\epsilon. \tag{3.6}$$

Wir haben gezeigt: Für jedes $\epsilon > 0$ haben wir ein j gefunden, und dann dazu ein $\delta_j > 0$, so dass (3.6) gilt. Dies ergibt

$$\lim_{x \to x_2} \frac{f(x) - f(x_2)}{x - x_2} = g(x_2)$$

und somit die Aussagen (ii) und (iii). □

[4] Achtung: wir behaupten hier: für jedes j gibt es solch ein δ_j. Man muss hier aber davon ausgehen, dass dieses δ_j von j abhängen kann.

4. Potenzreihen und analytische Funktionen

Satz 4.1. $\sum_{j=0}^{\infty} a_j x^j$, $a_j \in \mathbb{R}$ eine Potenzreihe mit Konvergenzradius $\rho > 0$. Dann ist die Funktion

$$f : (-\rho, \rho) \to \mathbb{R}, \qquad x \mapsto \sum_{j=0}^{\infty} a_j x^j$$

differenzierbar und

$$f'(x) = \sum_{j=1}^{\infty} j a_j x^{j-1}.$$

Man kann also in dieser Situation die Ableitung in die unendliche Summe „hineinziehen“. Obige Beispiele zeigen aber, dass dies eben nicht für alle gleichmäßig konvergenten Reihen von Funktionen gilt.

Beweis. Wähle ein $R \in (0, \rho)$. Wir wollen Theorem 3.3 anwenden für: $f_n : (-R, R) \to \mathbb{R}$, $f_n(x) = \sum_{j=0}^{n} a_j x^j$, $x_0 = 0$. Die Konvergenz in $x_0 = 0$ ist klar. Hierzu überprüfen wir, dass $f_n' : (-R, R) \to \mathbb{R}$ gleichmäßig konvergiert. [5]

Wir rechnen

$$f_n'(x) = \sum_{j=0}^{n} j a_j x^{j-1}$$

Für festes $x \neq 0$ konvergiert $f_n'(x)$ für $n \to \infty$ genau dann, wenn $h_n(x) := \sum_{j=0}^{n} j a_j x^j$ konvergiert. Wir berechnen den Konvergenzradius ρ' der Potenzreihe $\sum_{j=0}^{\infty} j a_j x^j$.

$$\limsup_{j\to\infty} \sqrt[j]{j|a_j|} = \underbrace{\left(\lim_{j\to\infty} \sqrt[j]{j}\right)}_{=1} \left(\limsup_{j\to\infty} \sqrt[j]{|a_j|}\right)$$

[5] Achtung: Für viele Potenzreihen konvergiert $f_n' : (-\rho, \rho) \to \mathbb{R}$ nicht gleichmäßig. Erst nach Einschränkung auf ein etwas kleineres Intervall $(-R, R)$ erhalten wir gleichmäßige Konvergenz. Siehe die analoge Fussnote in Kapitel 4, Abschnitt 2.

Also $\rho' = \rho$, das heißt der Konvergenzradius von $\sum_{j=0}^{\infty} j a_j x^{j-1}$ ist ρ.

Wegen Beispiel 2.6 konvergiert die Potenzreihe $\sum_{j=0}^{\infty} j a_j x^{j-1}$ auf $(-R, R)$ gleichmäßig, das heißt die Folge von Funktionen

$$g_n : (-R, R) \to \mathbb{R}, \qquad x \mapsto \sum_{j=0}^{n} j a_j x^{j-1}.$$

konvergiert gleichmäßig gegen eine Funktion $g : (-R, R) \to \mathbb{R}$. □

Beispiele 4.2.

(1) $\exp : \mathbb{R} \to \mathbb{R}$, $x \mapsto \sum_{j=0}^{\infty} \frac{1}{j!} x^j$ ist differenzierbar und

$$\exp'(x) = \sum_{j=1}^{\infty} \frac{j}{j!} x^{j-1} = \sum_{j=1}^{\infty} \frac{(j-1)!^{j-1}}{x} = \exp(x).$$

(2) $\sin : \mathbb{R} \to \mathbb{R}$, $x \mapsto \sum_{j=0}^{\infty} (-1)^j \frac{1}{(2j+1)!} x^{2j+1}$ ist differenzierbar und

$$\sin'(x) = \sum_{j=0}^{\infty} \frac{1}{(2j)!} x^{2j} = \cos(x).$$

Analog

$$\cos'(x) = -\sin(x).$$

Folgerung 4.3. Potenzreihen mit Konvergenzradius $\rho > 0$ sind auf dem Intervall $(-\rho, \rho)$ glatt, d.h. unendlich oft differenzierbar. Die Taylorreihe mit Entwicklungspunkt 0 von einer Potenzreihe ist wieder genau diese Potenzreihe.

$$T^0(\sum_{j=0}^{\infty} a_j x^j)(x) = \sum_{j=0}^{\infty} a_j x^j.$$

□

FOLGERUNG 4.4. Seien $f(x) = \sum_{n=0}^{\infty} a_n x^n$ und $g(x) = \sum_{n=0}^{\infty} b_n x^n$ Potenzreihen mit Konvergenzradien $\rho_a > 0$ und $\rho_b > 0$. Gibt es ein $\epsilon \in (0, \min\{\rho_a, \rho_b\}]$ so dass

$$f(x) = g(x) \qquad \text{für alle } x \text{ mit } |x| < \epsilon,$$

dann gilt $a_n = b_n$ für alle $n \in \mathbb{N}$.

□

DEFINITION 4.5. Sei I ein offenes Intervall. Eine Funktion $f : I \to \mathbb{R}$ heißt analytisch oder reell-analytisch, falls es zu jedem $x_0 \in I$ ein $\epsilon > 0$ und eine Potenzreihe $p(x) = \sum_{j=0}^{\infty} a_j x^j$ gibt, so dass gilt

- der Konvergenzradius ρ der Potenzreihe ist in $(0, \infty]$,
- $f(x + x_0) = p(x)$ für alle $x \in \mathbb{R}$ mit $|x| < \rho$ und $x + x_0 \in I$.

$C^\omega(I) := \{f : I \longrightarrow \mathbb{R} \mid f \text{ ist analytisch.}\}$

$$C^\omega \subset C^\infty \subset \cdots \subset C^{k+1} \subset C^k \subset \cdots \subset C^1 \subset C^0.$$

LEMMA 4.6. Potenzreihen sind auf ihrem Konvergenzbereich analytisch.

Beweis. kommt noch (folgt aus dem Umordnungssatz) □

SATZ 4.7 (Eindeutige Fortsetzbarkeit analytischer Funktionen)**.** Sei I ein offenes Intervall und sind $f, g : I \longrightarrow \mathbb{R}$ analytische Funktionen. In $x_0 \in I$ gelte eine der folgenden Bedingungen:

(a) Für alle $n \in \mathbb{N}$ gilt $f^{(n)}(x_0) = g^{(n)}(x_0)$.
(b) Es gibt ein $\epsilon > 0$ mit $f|_{(x_0-\epsilon,x_0+\epsilon)} = g|_{(x_0-\epsilon,x_0+\epsilon)}$

Dann haben wir sogar $f = g$.

Die Beweise des Lemmas und des Satzes wurden in der Vorlesung ausgelassen, da wir in der Analysis III analoge Aussagen für komplex-analytische Funktionen beweisen werden.

Beweis. Mit Folgerung 4.3 sehen wir, dass aus (a) die Voraussetzung (b) folgt. Sei $I = (a, b)$, $a < b$, $a, b \in \overline{\mathbb{R}}$. Wir definieren nun

$$t_1 := \sup\{t \in [x_0, b) \mid f = g \text{ auf } [x_0, t]\}.$$

Aus dieser Definition folgt $f = g$ auf $[x_0, t_1)$. Angenommen $t_1 < b$. Da $f - g$ stetig ist, gilt auch $f(t_1) = g(t_1)$, also stimmen f und g auf $[x_0, t_1]$ überein. Hieraus ergibt sich für alle $n \in \mathbb{N}$: $f^{(n)}(t_1) = g^{(n)}(t_1)$. Da f und g analytisch sind, gibt es ein $\epsilon > 0$, so dass $f = g$ auf $[t_1, t_1 + \epsilon)$. Dies ist ein Widerspruch zur Wahl von t_1. Also $t_1 = b$. Deswegen gilt $f = g$ auf $[x_0, b)$. Auf dem Intervall $(a, x_0]$ argumentiert man analog. □

KAPITEL 8

Topologie

Als Literatur in diesem Kapitel emfehle ich das erste Kapitel von [**34**] (normierte Vektorräume und metrische Räume) und Kapitel VI von [**9**] (inklusive topologische Räume).

1. Normierte Vektorräume

Vorbemerkung zur Vektorräumen: Im Gegensatz zur Linearen Algebra ist es in der Analysis nicht üblich und zumeist auch nicht hilfreich, streng zwischen Zeilen- und Spaltenvektoren zu unterscheiden.[1]

$\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

DEFINITION 1.1. Sei V ein $\mathbb{K}$-Vektorraum. Eine Norm auf V ist eine Abbildung $\| \cdot \| : V \to \mathbb{R}$, so dass für alle $x, y \in V$, $\lambda \in \mathbb{K}$ gilt:

[1]Sobald man anfängt, Matrizen mit Vektoren zu multiplizieren, müssen wir natürlich sorgfältiger werden. Dies ist allerdings aus Sicht der Analysis nur ein kurzer Zwischenschritt auf dem Weg zu „Tensoren“. Vektoren haben einen Index, Matrizen zwei und Tensoren beliebig viele. Die in der Linearen Algebra erfolgreiche Notation mit Zeilen und Spalten versagt, sobald wir mindestens drei Indizes haben. Tensoren werden in der Analysis benötigt, um den mehrdimensionalen Satz von Taylor zu erhalten. Später sind Tensorfelder ein wichtiges Hilfsmittel in der Analysis IV. Tensoren sind omnipräsent in der Physik (Allgemeine Relativitätstheorie, Materialwissenschaft, Elektrodynamik) und vielen anderen Anwendungen in Chemie und Technik.

(a) Definitheit der Norm: $\|0\| = 0$ und für $x \neq 0$ gilt $\|x\| > 0$,

(b) Homogenität der Norm: $\|\lambda x\| = |\lambda| \, \|x\|$

(c) Dreiecksungleichung der Norm: $\|x + y\| \leq \|x\| + \|y\|$

Man nennt dann $(V, \| \cdot \|)$ einen normierten Vektorraum.

Beispiele 1.2.

(1) Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf dem $\mathbb{K}$-Vektorraum V, wie in der Linearen Algebra II behandelt. Dann ist

$$x \mapsto \|x\| := \sqrt{\langle x, x \rangle}$$

eine Norm auf V.

Ist $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt auf $\mathbb{K}^n$, so nennt man die zugehörige Norm die Standardnorm auf $\mathbb{K}^n$.

(2) Sei D eine Menge. Eine Funktion $f : D \longrightarrow \mathbb{K}$ heißt *beschränkt*, wenn

$$\|f\|_\infty := \sup_{x \in D} \{|f(x)|\}$$

endlich ist. Die Menge $B(D, \mathbb{K})$ der beschränkten Funktionen von D nach $\mathbb{K}$ bildet mit der üblichen punktweisen Addition und Multiplikation mit Skalaren einen $\mathbb{K}$-Vektorraum. Dann ist $\| \cdot \|_\infty$ eine Norm auf $B(D, \mathbb{K})$. Definitheit, Homogenität und Dreiecksungleichung sind leicht zu prüfen. Man nennt diese Norm die Supremumsnorm.Wir erhalten insbesondere eine weitere Norm auf $\mathbb{K}^n = \mathrm{Abb}(\{1, 2, \ldots, n\}, \mathbb{K}) = B(\{1, 2, \ldots, n\}, \mathbb{K})$.

(3) Sei $p \in [1, \infty)$, $n \in \mathbb{N}$. Für $z = (z_1, \ldots, z_n) \in \mathbb{K}^n$ definieren wir die p-Norm

$$\|z\|_p := \left(\sum_{j=1}^{n} |z_j|^p \right)^{1/p} \in \mathbb{R}_{\geq 0}.$$

Dann ist $\|\cdot\|_2$ die Standardnorm auf $\mathbb{K}^n$ aus Beispiel (1).

Es ist eine Norm für alle $p \in [1,\infty)$: Definitheit und Homogenität von $\|\cdot\|_p$ werden in Übungsblatt 3, Aufgabe 3 gezeigt.

Bild aller Punkte $x \in \mathbb{R}^2$ mit $\|x\|_p = 1$ für $p = 1, \frac{5}{4}, 2, 5, \infty$.

Die Dreiecksungleichung ist klar für $p = 1$.

Sei nun $p > 1$. Bestimme $q \in (1,\infty)$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Also $q = p/(p-1)$. Nach Übungsblatt 3, Aufgabe 3 gilt für alle $a = (a_1,\dots,a_n), b = (b_1,\dots,b_n) \in \mathbb{R}^n$ die Höldersche Ungleichung

$$\sum_{j=1}^{n} |a_j b_j| \le \|a\|_p \cdot \|b\|_q.$$

Wir setzen $s_j := |x_j + y_j|^{p-1}$ und $s = (s_1,\dots,s_n)$. Wir wenden die Höldersche Ungleichung zunächst auf $a = x$ und $b = s$ an und dann nochmals auf $a = y$ und $b = s$, und erhalten.

$$\|x+y\|_p^p = \sum_{j=1}^{n} |x_j + y_j|^p \le \sum_{j=1}^{n} |x_j|\,|s_j| + \sum_{j=1}^{n} |y_j|\,|s_j| \le \|x\|_p \|s\|_q + \|y\|_p \|s\|_q$$

Man rechnet nun

$$\|s\|_q = \Big(\sum_{j=1}^{n} |s_j|^q\Big)^{1/q} = \Big(\sum_{j=1}^{n} |x_j + y_j|^p\Big)^{(1/p)\cdot(p-1)} = \|x+y\|_p^{p-1}.$$

Insgesamt folgt

$$\|x+y\|_p \le \|x\|_p + \|y\|_p.$$

Diese Ungleichung nennt man die Minkowski-Ungleichung. **Mi 8.5.**

(4) Ist $(V, \|\cdot\|)$ ein normierter Vektorraum und W ein Untervektorraum von V. Dann ist $\|\cdot\|\big|_W : W \to \mathbb{R}$ ein Norm auf W, die wir die von $\|\cdot\|$ induzierte Norm auf W nennen.

(5) Ist $\|\cdot\|$ eine Norm auf einem $\mathbb{C}$-Vektorraum V. Dann ist V auch ein $\mathbb{R}$-Vektorraum, und $\|\cdot\|$ ist dann auch eine Norm auf V im Sinn von $\mathbb{R}$-Vektorräumen.

DEFINITION 1.3. Ist $(V, \|\cdot\|)$ ein normierter Raum, so definiert

$$d(x, y) := \|x - y\|$$

eine Metrik auf V. Man nennt sie die von $\|\cdot\|$ induzierte Metrik auf V. Dadurch sind die topologischen Begriffe, also Begriffe wie „offen“, „abgeschlossen“, konvergent“, „Cauchy-Folge“, „Häufungspunkte“ etc. definiert. Insbesondere: Eine V-wertige Folge $(x(k))_{k\in\mathbb{N}}$ konvergiert gegen a in $(V, \|\cdot\|)$, falls $\|x(k) - a\| \to 0$ für $k \to \infty$. Aus der Dreiecksungleichung sieht man leicht: Konvergiert $x(k)$ gegen a, so konvergiert $\|x(k)\|$ gegen $\|a\|$.[2] Die Norm-Funktion

$$\|\cdot\| : V \to \mathbb{R}, \quad x \mapsto \|x\|$$

ist also (folgen-)stetig.

BEISPIEL 1.4. Die von der Supremumsnorm $\|\cdot\|_\infty$ induzierte Metrik auf $B(D, \mathbb{K})$ ist die Supremums-Distanz $d_{\sup}$ aus Kapitel 7.

Eine Folge $(x(k))_{k\in\mathbb{N}}$ heißt beschränkt im normierten Raum $(V, \|\cdot\|)$, falls es ein $C \in \mathbb{R}_{>0}$ gibt, so dass für alle $k \in \mathbb{N}$ gilt: $\|x(k)\| \leq C$.

LEMMA 1.5. Sei $x_j(k) \in \mathbb{K}$ bzw. $a_j \in \mathbb{K}$ die j-te Komponente von $x(k) \in \mathbb{K}^n$ bzw. $a \in \mathbb{K}^n$ bezüglich der Standard-Basis. Dann gilt für jede Norm auf V:

$$\lim_{k\to\infty} x(k) = a \text{ in } (\mathbb{K}^n, \|\cdot\|_\infty) \quad \Longleftrightarrow \quad \forall j \in \{1, 2, \ldots, n\} : \lim_{k\to\infty} x_j(k) = a_j.$$

[2] Denn $|\|x(k)\| - \|a\|| \leq \|x(k) - a\|$ oder verwende Übungsblatt 3 Aufgabe 4.

Beweis.

$$\begin{aligned}
&\lim_{k\to\infty} x(k) = a \text{ in } (\mathbb{K}^n, \|\cdot\|_\infty \\
\iff\ & \forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall k \in \mathbb{N}_{\geq N} : \|x(k) - a\|_\infty < \epsilon \\
\iff\ & \forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall k \in \mathbb{N}_{\geq N} : \forall j \in \{1, \ldots, n\} : |x_j(k) - a| < \epsilon \\
\overset{(*)}{\iff}\ & \forall j \in \{1, \ldots, n\} : \forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall k \in \mathbb{N}_{\geq N} : |x_j(k) - a| < \epsilon \\
\iff\ & \forall j \in \{1, 2, \ldots, n\} : \lim_{k\to\infty} x_j(k) = a_j.
\end{aligned}$$

Für die Umformung $(*)$ haben wir die Umformung

$$\exists N \in \mathbb{N} : A(j, N) : \forall j \in \{1, \ldots, n\} : A(j, N) \quad \iff \quad \forall j \in \{1, \ldots, n\} : \exists N \in \mathbb{N} : A(j, N)$$

benutzt, und wir müssen begründen wieso wir die Quantoren vertauschen dürfen. Die Richtung „$\Longrightarrow$“ ist offensichtlich. Für „$\Longleftarrow$“ argumentieren wir wie folgt: auf der rechten Seite erhalten wir für jedes j ein potenziell j-abhängiges $N = N_j$. Da j aber nur endlich viele Werte durchläuft, ist $\widehat{N} := \max\{N_j \mid j \in \{1, 2, \ldots, n\}\}$ eine natürliche Zahl, so dass die linke Seite für $N = \widehat{N}$ gilt. □

Lemma 1.6 (Bolzano-Weierstraß auf $(\mathbb{R}^n, \|\cdot\|_\infty)$)**.** Ist $(x(k))_{k\in\mathbb{N}}$ eine beschränkte Folge in $(\mathbb{R}^n, \|\cdot\|_\infty)$, dann besitzt diese Folge eine konvergente Teilfolge.

Der folgende Beweis würde das Ergebnis auch für alle anderen p-Normen liefern. Diese Verallgemeinerung folgt aber bald ganz einfach.

Beweis. Schreibe $x(k) = (x_1(k), \ldots, x_n(k))$. Es gilt $|x_j(k)| \leq \|x(k)\|_\infty$. Somit ist $(x_j(k))_{k\in\mathbb{N}}$ eine beschränkte $\mathbb{R}$-wertige Folge. Nach dem Satz von Bolzano-Weierstraß (Korollar 1.35 in Kapitel 2) besitzt die Folge $(x_1(k))_{k\in\mathbb{N}}$ eine (in $\mathbb{R}$) konvergente Teilfolge $(x_1(f_1(k)))_{k\in\mathbb{N}}$.

Weiter besitzt $(x_2(f_1(k)))_{k\in\mathbb{N}}$ eine konvergente Teilfolge $(x_2(f_1\circ f_2(k)))_{k\in\mathbb{N}}$. Und so weiter. Nachdem wir n-mal zu einer derartigen Teilfolge übergegangen sind, konvergiert

$$(x_j(\underbrace{f_1\circ f_2\circ\cdots\circ f_n}_{f:=}(k)))_{k\in\mathbb{N}}$$

für alle j. Sei $a_j := \lim_{k\to\infty} x_j(f(k))$. Dann $\|a - x(f(k))\|_\infty \to 0$ für $k\to\infty$. □

DEFINITION 1.7. Zwei Normen $\|\cdot\|$ und $|||\cdot|||$ auf einem Vektorraum V nennt man äquivalent, falls es ein $C\in\mathbb{R}_{>0}$ gibt, so dass für alle $x\in V$ gilt:

$$\|x\|\le C|||x||| \qquad \text{und} \qquad |||x|||\le C\|x\|.$$

SATZ 1.8. Alle Normen auf $\mathbb{R}^n$ sind äquivalent.

Beweis. Es reicht, den Satz zu zeigen für $\|\cdot\| = \|\cdot\|_\infty$ und eine weitere Norm $|||\cdot|||$. Sei $(e_1,\dots,e_n)$ die Standardbasis von $\mathbb{R}$.

$$|||x||| = \Big|||\sum_{j=1}^n x_j e_j\Big||| \le \sum_{j=1}^n |x_j|\,|||e_j||| \le \|x\|_\infty \underbrace{(|||e_1||| + \ldots + |||e_n|||)}_{C_1:=}.$$

Dies liefert nun die rechte der zu zeigenden Ungleichungen.

Angenommen, die Normen sind nicht äquivalent, d.h. es gibt kein C wie oben. Dann gibt es eine Folge $(x(k))_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ mit $\|x(k)\|_\infty > k|||x(k)|||$. O.B.d.A. $\|x(k)\|_\infty = 1$ (sonst ersetze $x(k)$ durch $x(k)/\|x(k)\|_\infty$). Wir können auch annehmen, dass diese Folge konvergiert, da wir gegebenenfalls zu einer in $(\mathbb{R}^n,\|\cdot\|_\infty)$ konvergenten Teilfolge übergehen können (siehe Lemma 1.6). Sei a der Grenzwert. Wegen $|||x(k)-a|||\le C_1\|x(k)-a\|_\infty\to 0$ ist a auch der Grenzwert bezüglich der $|||\cdot|||$-Norm. Es folgt $\|a\|_\infty = \lim_{k\to\infty}\|x(k)\|_\infty = 1$ und wegen $|||x(k)||| < \frac{1}{k}$ auch $|||a||| = \lim_{k\to\infty}|||x(k)||| = 0$. Also $a=0$, was $\|a\|_\infty = 1$ widerspricht. □

Der obige Satz gilt wegen Beispiele 1.2 (5) auch, wenn wir $\mathbb{R}^n$ durch $\mathbb{C}^n$ ersetzen.

Sind zwei Normen $\|\cdot\|$ und $|||\cdot|||$ äquivalent, so gilt

$$\lim_{k\to\infty} \|x(k)-a\| = 0 \qquad \Longleftrightarrow \qquad \lim_{k\to\infty} |||x(k)-a||| = 0.$$

$$(x(k))_{k\in\mathbb{N}} \text{ beschränkt in } (V,\|\cdot\|) \qquad \Longleftrightarrow \qquad (x(k))_{k\in\mathbb{N}} \text{ beschränkt in } (V,|||\cdot|||).$$

Da jeder endlich-dimensionale $\mathbb{K}$-Vektorraum (als Vektorraum) isomorph zu $\mathbb{K}^n$ für ein geeignetes $n \in \mathbb{N}\cup\{0\}$ ist, folgt hieraus unmittelbar das folgende Korollar.

Korollar 1.9. Für einen endlich-dimensionalen $\mathbb{K}$-Vektorraum V gilt:

(a) Alle Normen auf V sind zueinander äquivalent.

(b) Bezüglich einer beliebigen Basis sei $x_j(k)$ bzw. a_j die j-te Komponente von $x(k)$ bzw. von a gilt:

$$\lim_{k\to\infty} x(k) = a \quad \Longleftrightarrow \quad \forall j \in \{1,2,\ldots,\dim V\} : \lim_{k\to\infty} x_j(k) = a_j.$$

(c) Jede beschränkte V-wertige Folge besitzt eine konvergente Teilfolge. (Satz von Bolzano-Weierstraß für endlich-dimensionale Vektorräume).

Wir betrachten nun einige unendlich-dimensionale normierte Räume.

Beispiel 1.10. Für $p \in [1,\infty)$ und $x = (x_j) \in \mathrm{Abb}(\mathbb{N},\mathbb{K})$ definieren wir

$$\|x\|_p := \Big(\sum_{j=0}^{\infty} |x_j|^p\Big)^{1/p}.$$

Hierbei weisen wir der Reihe genau dann den Wert ∞ zu, wenn sie divergiert, das heißt wenn sie unbestimmt gegen ∞ konvergiert. Wir setzen $\infty^{1/p} := \infty$. Für $p = \infty$ definieren

wir $\|x\|_\infty$ wie in Beispiele 1.2 (2). Man definiert nun

$$\ell^p(\mathbb{K}) := \{x \in \mathrm{Abb}(\mathbb{N}, \mathbb{K}) \mid \|x\|_p < \infty\}.$$

Dann ist $\| \cdot \|_p$ eine Norm auf $\ell^p(\mathbb{K})$: Die Homogenität und Definitheit ist klar. Die Dreiecks-Ungleichung ist klar für $p = \infty$. Um die Dreiecks-Ungleichung für $p < \infty$ zu zeigen, definieren wir zu $x \in \mathrm{Abb}(\mathbb{N}, \mathbb{K})$ den Vektor $x^{(n)} \in \mathrm{Abb}(\mathbb{N}, \mathbb{K})$, $n \in \mathbb{N}$ durch

$$x_j^{(n)} := \begin{cases} x_j & \text{für } j \leq n, \\ 0 & \text{für } j > n. \end{cases}$$

Für $x \in \ell^p(\mathbb{K})$ gilt

$$\|x - x^{(n)}\|_p^p = \sum_{j=n+1}^{\infty} |x_j|^p \to 0 \qquad \text{für } n \to \infty,$$

und somit gilt $x^{(n)} \to x$. Wir können $x^{(n)}$ als Vektor $(x_0, \ldots, x_n) \in \mathbb{K}^{(n+1)}$ betrachten und dann die Dreiecks-Ungleichung von Beispiele 1.2 (3) anwenden:

$$\|x^{(n)} + y^{(n)}\|_p \leq \|x^{(n)}\|_p + \|y^{(n)}\|_p.$$

Im Limes $n \to \infty$ ergibt sich dann

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Fr 10.5.

Übung 1.11. Zeigen Sie, dass $\ell^p(\mathbb{R})$ für alle $p \in [1, \infty]$ vollständig ist.

Bemerkung 1.12. Auf unendlich-dimensionalen Vektorräumen gibt es nicht-äquivalente Normen. Sei

$$\mathrm{Abb}_c(\mathbb{N}, \mathbb{K}) := \left\{ (x_j) \in \mathrm{Abb}(\mathbb{N}, \mathbb{K}) \,\middle|\, \#\{j \in \mathbb{N} \mid x_j \neq 0\} < \infty \right\}.$$

Dieser Vektorraum besitzt eine Basis $(\chi_j)_{j\in\mathbb{N}}$, wobei

$$\chi_j(k) = \begin{cases} 1 & \text{falls } j = k\,, \\ 0 & \text{falls } j \neq k\,. \end{cases}$$

Dann gilt $\mathrm{Abb}_c(\mathbb{N},\mathbb{K}) \subset \ell^p(\mathbb{K})$ für alle $p \in [1,\infty]$, und wir erhalten eine induzierte p-Norm auf $\mathrm{Abb}_c(\mathbb{N},\mathbb{K}) \subset \ell^p(\mathbb{K})$. Für $p \neq \tilde{p}$ sind die Normen $\|\cdot\|_p$ und $\|\cdot\|_{\tilde{p}}$ nicht äquivalent. Betrachten wir zum Beispiel die Folge $\left(\frac{1}{n}\sum_{j=0}^{n}\chi_j\right)_{n\in\mathbb{N}}$. Dann gilt

$$\Big\|\frac{1}{n}\sum_{j=0}^{n-1}\chi_j\Big\|_1 = 1, \qquad \Big\|\frac{1}{n}\sum_{j=0}^{n-1}\chi_j\Big\|_\infty = \frac{1}{n}.$$

Somit ist diese Folge eine Nullfolge in der Norm $\|\cdot\|_\infty$, aber keine Nullfolge in der Norm $\|\cdot\|_1$.

Beispiel 1.13. Betrachte den Vektorraum $V = C^0([a,b],\mathbb{K})$, $b > a$, $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Wir definieren: $f \in C^0([a,b],\mathbb{K})$

$$\|f\|_1 := \int_a^b |f(x)|\,dx, \qquad \|f\|_\infty := \sup\{|f(x)| \mid x \in [a,b]\}.$$

Die Normen $\|\cdot\|_1$ und $\|\cdot\|_\infty$ sind nicht äquivalent. Es gilt zwar $\|f\|_1 \leq (b-a)\|f\|_\infty$. Es gibt aber **keine** Konstante $C \in \mathbb{R}_{>0}$ mit

$$\|f\|_\infty \leq C\|f\|_1.$$

Um dies zu zeigen, betrachte man für $n \geq 1/(b-a)$ die Funktion

$$f_n : [a,b] \to \mathbb{R}, \quad f(x) := \begin{cases} 1 - n(x-a) & \text{falls } a \leq x \leq a + \frac{1}{n} \\ 0 & \text{falls } a + \frac{1}{n} < x \leq b \end{cases}$$

Es gilt $\|f_n\|_\infty = 1$ und $\|f_n\|_1 = 1/(2n) \to 0$. Die Folge $(f_n)_{n\in\mathbb{N}}$ ist also eine Nullfolge in $(C^0([a,b],\mathbb{K}), \|\cdot\|_1)$, aber nicht in $(C^0([a,b],\mathbb{K}), \|\cdot\|_\infty)$. Die Folge $(f_n)_{n\in\mathbb{N}}$ besitzt keine konvergente Teilfolge in $((C^0([a,b],\mathbb{K}), \|\cdot\|_\infty)$.

DEFINITION 1.14. Einen normierten Vektorraum $(V, \|\cdot\|)$ nennt man einen Banachraum oder einen vollständigen normierten Vektorraum, falls die induzierte Metrik auf V vollständig ist. Einen euklidischen oder unitären Vektorraum $(V, \langle\cdot,\cdot\rangle)$ nennt man einen Hilbertraum, wenn $(V, x \mapsto \|x\| := \sqrt{\langle x, x\rangle})$ ein Banachraum ist.

Aus dem Satz von Bolzano-Weierstraß erhalten wir das folgende Korollar.

KOROLLAR 1.15. Alle endlich-dimensionalen normierten Vektorräume sind Banachräume. Alle endlich-dimensionalen euklidischen oder unitären Vektorräume sind Hilberträume.

Hilberträume und Banachräume (beliebiger Dimension) sind unter anderem von zentraler Bedeutung

- in der Quantenmechanik. Hier ist insbesondere der Hilbertraum $\ell^2(\mathbb{C})$ mit dem Skalarprodukt
$$\langle (x_j), (y_j)\rangle := \sum_{j=0}^{\infty} \overline{x_j} y_j$$
ganz wichtig.[3]

[3]Beispiel: Angenommen ein quantenmechanisches Teilchen kann abzählbar viele reine Zustände annehmen, die wir mit den natürlichen Zahlen $\mathbb{N}$. Zum Beispiel die möglichen Anregungszustände eines Elektrons um einen Atomkern mit verschiedener Energie. Das Teilchen wird dann durch einen Vektor $\psi \in \ell^2(\mathbb{C})$ mit $\|\psi\|_2 = 1$ beschrieben, und $|\psi_k|^2$ ist die Wahrscheinlichkeit, dass eine Messung den k-ten reinen Zustand ergibt. Der Raum $\ell^2(\mathbb{C})$ beschreibt aber auch noch viel kompliziertere Systeme. Solange

- in der konkreten Berechung von aufwändigen mathematischen Probleme mit Hilfe des Computers (Numerik von partiellen Differentialgleichungen)

2. Kontraktionen und Banachscher Fixpunktsatz

Wiederholung:

Ein metrischer Raum ist ein Paar (X, d), wobei die Metrik $d : X \times X \longrightarrow \mathbb{R}_{\geq 0}$ die folgenden Eigenschaften für alle $x, y, z \in X$ erfüllen möge

(a) $d(x, y) = 0 \iff x = y$ (Definitheit)
(b) $d(x, y) = d(y, x)$ (Symmetrie)
(c) $d(x, y) \leq d(x, z) + d(z, y)$ (Dreiecks-Ungleichung)

$$\lim_{n\to\infty} x_n = a \qquad \Longleftrightarrow \qquad \lim_{n\to\infty} d(x_n, a) = 0$$

Beispiel 2.1. Ist $(V, \|\cdot\|)$ ein normierter Vektorraum, so definiert $d(x, y) := \|x - y\|$ eine Metrik auf V, die von $\|\cdot\|$ auf V induzierte Metrik.

Sei (X, d) ein metrischer Raum, $Y \subset X$, $f : Y \to X$.

Definition 2.2. Man nennt $x \in Y$ einen Fixpunkt von f, falls $f(x) = x$. Die Abbildung $f : Y \to X$ wird Kontraktion genannt, falls es ein $L \in [0, 1)$ gibt, so dass für alle $y, z \in Y$ gilt:

$$d(f(y), f(z)) \leq Ld(y, z).$$

man noch keine Quantenfeldtheorie betreibt sind alle Hilbert-Räume entweder isomorph zu $\mathbb{C}^k$, $k \in \mathbb{N}$ mit dem Standard-Skalarprodukt oder zu $\ell^2(\mathbb{C})$.

BEISPIEL 2.3. $a \in (-1,1)$, $b \in \mathbb{R}$, $X = Y = \mathbb{R}$. Dann ist $f(x) = ax + b$ eine Kontraktion mit $L = |a|$.

Kontraktionen sind immer stetig.

Kontraktionen besitzen höchstens einen Fixpunkt. Denn sind x_1 und x_2 Fixpunkte, so gilt $d(x_1, x_2) \leq Ld(x_1, x_2)$, also $d(x_1, x_2) = 0$.

BEISPIEL 2.4. $X = Y = \mathbb{R} \setminus \{0\}$. Dann ist $f(x) = x/2$ eine Kontraktion ohne Fixpunkte.

SATZ 2.5 (Banachscher Fixpunktsatz)**.** Sei X ein vollständiger metrischer Raum, $A \subset X$ eine abgeschlossene Teilmenge, und $Y \subset X$ eine beliebige nicht-leere Teilmenge. Sei $f : Y \to X$ eine Kontraktion mit $f(Y) \subset A \subset Y$. Dann besitzt f genau einen Fixpunkt.

Beweis. Wähle ein $x_0 \in Y$ und definiere dann rekursiv

$$x_{n+1} = f(x_n)$$

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq Ld(x_n, x_{n-1})$$

Für $k \in \mathbb{N}$:

$$\begin{aligned} d(x_{n+k}, x_n) &\leq \sum_{j=1}^{k} d(x_{n+j}, x_{n+j-1}) \leq (L + L^2 \cdots + L^k)\, d(x_n, x_{n-1}) \\ &\leq \frac{L\, d(x_n, x_{n-1})}{1 - L} \leq \frac{L^n\, d(x_1, x_0)}{1 - L}. \end{aligned}$$

Hieraus sieht man sofort, dass $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist. Da X vollständig ist, ist sie konvergent in X.

$$x := \lim_{n\to\infty} x_n = \lim_{n\to\infty} f(x_{n-1})$$

Somit ist x Häufungspunkt der Menge $f(Y)$ und deswegen (nach Aufgabe 4 Übungsblatt 12 der Analysis I) in A.

$$f(x) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x$$

Also ist x ein Fixpunkt, und nach obiger Bemerkung der einzige. □

Beispiel 2.6. Betrachte die Funktion $f : [1,2] \to \mathbb{R}$, $f(x) = x + \frac{1}{2}\cos x$. Dann gilt $f'(x) = 1 - \frac{1}{2}\sin x \geq 1/2$, also ist f streng monoton wachsend. Auf Grund der Ergebnisse in Abschnitt 4 in Kapitel 4 gilt $\cos 1 \geq 1/2$, also $f(1) \geq 5/4$, und $\cos 2 < 0$, also $f(2) < 2$. Es folgt $f([1,2]) \subset [5/4, 2]$. Wegen $\sin x \geq x - x^3/6$ gilt für alle $x \in [1,2]$ die Ungleichung $\sin x \geq 2/3$ und somit

$$f'(x) \leq 1 - \frac{1}{2}\frac{2}{3} = \frac{2}{3}.$$

Mit dem Mittelwertsatz ergibt sich daraus $|f(x) - f(y)| \leq (2/3)|x - y|$ für $x, y \in [1,2]$, d.h. f ist eine Kontraktion. Da $[1,2]$ vollständig ist, konvergiert nach dem Banachschen Fixpunktsatz (und desen Beweis) die rekursiv definierte Folge

$$x_0 := 1, \qquad x_{n+1} := f(x_n)$$

gegen den eindeutigen Fixpunkt a von f. Wegen $\pi/2 \in [1,2]$ folgt $a = \pi/2$ und wir bekommen

$$|\pi/2 - x_n| \leq \left(\frac{2}{3}\right)^n.$$

Hiermit könnte man ein Computer-Programm zur Berechnung von π schreiben. Es gibt aber deutlich effizientere Verfahren zur Berechnung von π.

Mi 15.5.

3. Topologische Räume

3.1. Wiederholung: Eigenschaften eines metrischen Raums. Wir definieren in einem metrischen Raum (M, d) für $O, U \subset M$, $x \in M$, $r > 0$:

$$B_r^{(M,d)}(x) := B_r(x) := \{y \in M \mid d(x, y) < r\}$$

$$U \text{ ist Umgebung von } x \text{ in } (M, d) \iff \exists r > 0 : B_r(x) \subset U.$$

$$U \text{ ist offen in } (M, d) \iff U \text{ ist Umgebung von allen } y \in U.$$

Wir haben gesehen (Proposition 5.10 in Kapitel 4):

(1) $\emptyset$ und M offen in (M, d).
(2) Sind U_1 und U_2 offene Teilmengen, dann ist auch $U_1 \cap U_2$ offen.
(3) Sei $(U_j)_{j \in I}$ eine Familie offener Teilmengen von (M, d), dann ist auch

$$\bigcup_{j \in I} U_j$$

eine offene Teilmenge.

In (c) ist wichtig, dass beliebige Indexmengen I erlaubt sind: auch abzählbar unendliche und überabzählbare.

Wir definieren

$$\mathcal{O}_{M,d} := \{U \subset M \mid U \text{ offen in } (M, d)\}$$

Wir nennen $A \subset M$ abgeschlossen, falls $M \setminus A$ offen ist.

Auf Grund der Dreicksungleichung ist $B_r(x)$ offen.

BEISPIEL 3.1. Die diskrete Metrik auf M:

$$d_{\text{disc}}(x,y) = \begin{cases} 1 & \text{falls } x \neq y \\ 0 & \text{falls } x = y \end{cases}.$$

Für $r = 1/2$ und alle $x \in M$ gilt $B_r(x) = \{x\}$. Daraus folgt, dass jede Teilmenge von M offen und abgeschlossen ist. Also $\mathcal{O}_{M,d_{\text{disc}}} = \mathcal{P}(M)$.

3.2. Definition topologischer Räume.

DEFINITION 3.2. Sei X eine Menge und $\mathcal{O} \subset \mathcal{P}(X)$. Man nennt $\mathcal{O}$ eine Topologie auf X, falls gilt:

(1) $\emptyset \in \mathcal{O}$, $X \in \mathcal{O}$,
(2) Aus $U_1 \in \mathcal{O}$ und $U_2 \in \mathcal{O}$ folgt $U_1 \cap U_2 \in \mathcal{O}$,
(3) Ist $(U_i)_{i\in I}$ eine Familie, $U_i \in \mathcal{O}$, dann gilt auch $\bigcup_{i\in I} U_i \in \mathcal{O}$.

Ein topologischer Raum ist ein Paar $(X, \mathcal{O})$, wobei $\mathcal{O}$ eine Topologie auf X ist. Man sagt: U ist eine offene Menge in $(X, \mathcal{O})$ genau dann, wenn $U \in \mathcal{O}$. Wenn die Topologie auf X aus dem Kontext heraus klar ist, schreibt man oft auch X für $(X, \mathcal{O})$.

BEISPIELE 3.3. (a) Ist (X, d) ein metrischer Raum, so ist $\mathcal{O}_{X,d}$ eine Topologie auf X. Man nennt sie die von d induzierte Topologie.
(b) $\mathcal{O} := \mathcal{P}(X)$ ist eine Topologie auf X, die sogenannte diskrete Topologie. Sie wird von der diskreten Metrik d_{disc} induziert.
(c) Die Klumpentopologie auf X ist definiert als $\mathcal{O}_{\text{Klump}} := \{\emptyset, X\}$.
(d) Ist $\mathcal{O}_X$ eine Topologie auf X. Sei $Y \subset X$. Dann ist

$$\mathcal{O}_Y := \{U \cap Y \mid U \in \mathcal{O}_X\}$$

eine Topologie auf Y, genannt die Spurtopologie auf Y oder einfach die auf Y induzierte Topologie.

Bild von einer kompakten Teilmenge Y mit glattem Rand in $\mathbb{R}^2$, die mit einem offenen Ball in $\mathbb{R}^2$ geschnitten wird und so eine offene Teilmenge von Y definiert.

Die Elemente von $\mathcal{O}_Y$ nennt oft auch in Y offene Mengen.

(e) Sind $\|\cdot\|$ und $|||\cdot|||$ äquivalente Normen auf dem Vektorraum V, so induzieren sie dieselbe Topologie auf V. Denn angenommen es gilt für alle $x \in V$

$$\|x\| \leq C|||x||| \qquad \text{und} \qquad |||x||| \leq C\|x\|.$$

Dann folgt für $x \in V$, $r > 0$:

$$B_r^{(V,\|\cdot\|)}(x) \subset B_{Cr}^{(V,|||\cdot|||)}(x), \qquad B_r^{(V,|||\cdot|||)}(x) \subset B_{Cr}^{(V,\|\cdot\|)}(x).$$

Also für $U \subset V$:

$$U \text{ offen in } (V, \|\cdot\|) \iff U \text{ offen in } (V, |||\cdot|||).$$

Umgekehrt gilt: Nicht-äquivalente Normen induzieren verschiedene Topologien.

(f) Die Metrik d_{glm} definiert eine Topologie $\mathcal{O}_{\text{glm}}$ auf $\text{Abb}(D, \mathbb{R})$

(g) Auf dem Raum $\text{Abb}(D, \mathbb{R})$ definieren wir die Produkttopologie wie folgt: Für $a, b \in \mathbb{R}$ und $x \in D$ definiere

$$U_{a,b}^x := \{f \in \text{Abb}(D, \mathbb{R}) \mid a < f(x) < b\} \subset \text{Abb}(D, \mathbb{R}).$$

Dann definiere

$$\begin{aligned} \mathcal{S} &:= \{U^x_{a,b} \mid a < b \text{ und } x \in D\} \subset \mathcal{P}(\mathrm{Abb}(D, \mathbb{R})). \\ \mathcal{B} &:= \{U_1 \cap \ldots \cap U_k \mid k \in \mathbb{N}_{>0} \text{ und } U_1, \ldots, U_k \in \mathcal{S}\} \cup \{\mathrm{Abb}(D, \mathbb{R})\} \\ &\quad \subset \mathcal{P}(\mathrm{Abb}(D, \mathbb{R})). \\ U \in \mathcal{O}_{\mathrm{prod}} &:\Longleftrightarrow \forall x \in U : \exists B \in \mathcal{B} : x \in B \subset U \end{aligned}$$

$\mathcal{O}_{\mathrm{prod}}$ ist eine Topologie auf $\mathrm{Abb}(D, \mathbb{R})$: Die Eigenschaften (1)–(3) in Definition 3.2 sind leicht zu prüfen. [4] Man kann mit etwas mehr Aufwand auch zeigen:

- $\mathcal{O}_{\mathrm{prod}} \subset \mathcal{O}_{\mathrm{glm}}$ und falls D unendlich: $\mathcal{O}_{\mathrm{prod}} \subsetneq \mathcal{O}_{\mathrm{glm}}$
- Ist D überabzählbar unendlich, z.B. $D = [0, 1]$, dann gibt es keine Metrik auf $\mathrm{Abb}(D, \mathbb{R})$, die die Topologie $\mathcal{O}_{\mathrm{prod}}$ induziert.
- $\mathcal{O}_{\mathrm{prod}}$ „beschreibt“ punktweise Konvergenz, wie wir bald sehen werden.

Bemerkung 3.4. Wieso sind topologische Räume wichtig?

- Es gibt wichtige topologische Räume, die nicht von irgendeiner Metrik induziert werden. Ein Beispiel dafür ist $(\mathrm{Abb}([0, 1], \mathbb{R}), \mathcal{O}_{\mathrm{prod}})$. Es gibt keine Metrik, die diese Topologie induziert, siehe hierzu auch Beispiel 3.17.
- Oft besitzt ein Raum M verschiedene Metriken, die dieselbe Topologie $\mathcal{O}$ induzieren. Zum Beispiel führen alle Normen zur selben Topologie auf $\mathbb{R}^n$. Wenn man nun Definition und Sätze erstellt, die nur die Topologie $\mathcal{O}$ nutzen, aber nicht die Metrik, dann ist klar, dass die definierten Begriffe und Sätze nicht von der Wahl der passenden Metrik abhängen.

[4] Diese Produkttopologie ist ein Spezialfall einer allgemeineren Konstruktion, siehe [**17**, Kap. 10].

- Manche topologische Räume besitzen zwar eine Metrik, diese sieht aber ziemlich unhandlich aussieht. Auf dem Vektorraum $C^\infty(\mathbb{R},\mathbb{R})$ der glatten Funktionen wird oft eine „natürliche“ Topologie definiert[5]. Die Topologie wird zwar von einer geeigneten Metrik d induziert. Diese ist aber für praktische Zwecke unhandlich. Inbesondere gibt es keine Norm auf $C^\infty(\mathbb{R},\mathbb{R})$, die diese Topologie induziert.

Es ist wichtig, dass Sie möglichst früh mit topologischen Räumen umgehen, damit Ihnen automatisch bewusst ist, welche Definitionen topologischer Natur sind, und welche Definitionen zusätzliche Struktur (Metriken, Normen, Vektorräume) benötigen.

DEFINITION 3.5. U Umgebung von $x :\Longleftrightarrow$ Es existiert eine offene Menge W mit $x \in W \subset U$.

$A \subset X$ heißt abgeschlossen in $(X,\mathcal{O}) :\Longleftrightarrow X \setminus A$ offen in $(X,\mathcal{O})$.

DEFINITION 3.6. Ist N eine Teilmenge eines topologischen Raumes $(X,\mathcal{O})$, dann definieren wir:

den Abschluss oder die abgeschlossene Hülle von N als

$$\overline{N} := \bigcap \{A \subset X \mid N \subset A \text{ und } A \text{ abgeschlossen in } X\},$$

das Innere von N oder den inneren Kern von N als

$$\mathring{N} := \bigcup \{U \subset N \mid U \text{ offen in } X\},$$

[5]Wir wollen dies hier nicht tun, sondern nur anmerken, dass sie zum Beispiel wichtig ist, wenn man die in der Physik oft benutzte „δ-Funktion“ mathematisch sauber verstehen will.

den Rand von N als

$$\partial N := \overline{N} \setminus \mathring{N}.$$

Lemma 3.7.
(a) $\overset{\circ}{N} = X \setminus \overline{(X \setminus N)}, \qquad \overline{N} = X \setminus \overbrace{(X \setminus N)}^{\circ}$

(b) $x \in \overset{\circ}{N} \iff N$ Umgebung von x

(c) $x \in \overline{N} \iff$ für alle Umgebungen U von x gilt: $U \cap N \neq \emptyset$

(d) $x \in \partial N \iff$ für alle Umgebungen U von x gilt: $U \cap N \neq \emptyset$ und $U \cap (X \setminus N) \neq \emptyset$.

Beweis. Übung. □

Beispiel 3.8. $N = [0,1] \times (0,1) \subset \mathbb{R}^2$, Standard-Norm.

$\overset{\circ}{N} = (0,1) \times (0,1)$

$\overline{N} = [0,1] \times [0,1]$

$\partial N = ([0,1] \times \{0,1\}) \cup (\{0,1\} \times [0,1])$

Zeichnung von diesen Mengen

Definition 3.9. Eine Teilmenge $N \subset X$ heißt dicht, falls $\overline{N} = X$.

Beispiel 3.10. $\mathbb{Q}$ ist dicht in $\mathbb{R}$ (versehen mit Standard-Norm und der induzierten Topologie).

Definition 3.11. Seien $(X, \mathcal{O}_X)$ und $(Y, \mathcal{O}_Y)$ topologische Räume, und $f : X \to Y$ eine Abbildung, $x_0 \in X$. Man nennt f stetig in x_0, falls gilt:

Für jede Umgebung U von $f(x_0)$ in Y ist $f^{-1}(U)$ eine Umgebung von x_0.

(Kurz: Urbilder von Umgebungen sind Umgebungen.)
Man nennt f stetig, falls gilt:

Für alle $U \in \mathcal{O}_Y$ gilt $f^{-1}(U) \in \mathcal{O}_X$.

(Kurz: Urbilder offener Mengen sind offen.)

Man sieht leicht: f ist genau dann stetig, wenn es in allen $x_0 \in X$ stetig ist.

Diese Definition verallgemeinert alle bisherigen Definitionen von Stetigkeit.

3.3. Konvergenz in topologischen Räumen.

DEFINITION 3.12. Ein topologischer Raum $(X, \mathcal{O})$ wird Hausdorffraum genannt, falls die Hausdorff-Eigenschaft gilt:

für alle $x, y \in X$ mit $x \neq y$ gibt es offene Mengen U_x und U_y, so dass $x \in U_x$, $y \in U_y$ und

$$U_x \cap U_y = \emptyset.$$

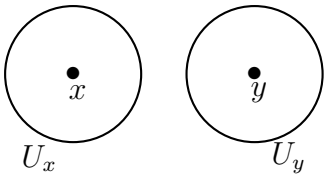


Abbildung 8: Hausdorff-Eigenschaft

[6]

LEMMA 3.13. Jeder metrische Raum ist ein Hausdorffraum.

Beweis. Seien $x, y \in X$, $x \neq y$. Zu $r := d(x, y)/3$ definiere die offenen Mengen

$$U_x := B_r(x) \qquad U_y := B_r(y).$$

[6]Der Hausdorffraum ist nach Felix Hausdorff bennannt, ein Blick in die Biographie wird empfohlen, `http://de.wikipedia.org/wiki/Felix_Hausdorff`.

Wegen der Dreiecksungleichung gilt $U_x \cap U_y = \emptyset$. □

Wiederholung (Definition 2.10 in Kapitel 3):
Sei $A(\,\cdot\,)$ eine auf $\mathbb{N}$ definierte Aussageform.

$$\begin{aligned} &\text{Für } \textit{fast alle } j \in \mathbb{N} \text{ gilt } A(j) \\ :\Longleftrightarrow\ &\text{Es gibt ein } j_0 \in \mathbb{N}, \text{ so dass für alle } j \in \mathbb{N}_{\geq j_0} : A(j) \\ \Longleftrightarrow\ &\text{Die Menge } \{j \in \mathbb{N} \mid \neg A(j)\} \text{ ist endlich} \end{aligned}$$

DEFINITION 3.14. Sei $(X, \mathcal{O})$ ein Hausdorffraum, $(x_i)_{i\in\mathbb{N}}$ eine Folge in X, $a \in X$.

$$\begin{aligned} a = \lim_{i\to\infty} x_i \ \Longleftrightarrow\ &\text{Für jede Umgebung } U \text{ von } a \text{ gilt } x_i \in U \text{ für fast alle } i \in \mathbb{N} \\ \Longleftrightarrow\ &\text{Zu jeder Umgebung } U \text{ von } a \text{ gibt es ein } i_0 \in \mathbb{N}, \\ &\text{so dass für alle } i \geq i_0\text{: } x_i \in U. \end{aligned}$$

Grenzwerte in Hausdorffräumen sind eindeutig, falls sie existieren. Denn angenommen a und b wären Grenzwert der Folge $(x_i)_{i\in\mathbb{N}}$, $a \neq b$. Da X ein Hausdorffraum ist, gibt es offene Mengen U_a und U_b mit $a \in U_a$, $b \in U_b$ und $U_a \cap U_b = \emptyset$. Da U_a eine Umgebung von a ist, gibt es ein $i_0 \in \mathbb{N}$, so dass für $i \geq i_0$ gilt: $x_i \in U_a$. Analog dazu gibt es ein $j_0 \in \mathbb{N}$, so dass für $i \geq j_0$ gilt: $x_i \in U_b$. Dann gilt für $k \geq \max\{i_0, j_0\}$ sogar $x_k \in U_a \cap U_b = \emptyset$, was offensichtlich nicht möglich ist. Also war die Annahme falsch, dass es zwei verschiedene Grenzwerte gibt. [7]

BEMERKUNG 3.15. Nicht jede Topologie wird von einer Metrik induziert.

[7] Man kann Konvergenz auch in topologischen Räumen betrachten, die nicht mehr hausdorffsch sind. Dazu benötigt man die Theorie der „Filter“, das ist aber in Ihrer Situation doch recht aufwändig im Vergleich zum Nutzen.

BEISPIEL 3.16. Sei $\mathcal{O}_{\text{Klump}} := \{\emptyset, X\}$ die Klumpentopologie auf X. Im Fall $\#X \geq 2$ ist $(X, \mathcal{O}_{\text{Klump}})$ kein Hausdorffraum, also nicht von einer Metrik induziert. [8]

BEISPIEL 3.17. Auf der Menge $\text{Abb}([0,1], \mathbb{R})$ gibt es eine Topologie (die sogenannte Produkttopologie, siehe z. B. [**17**, Kap. 10] oder siehe Beispiele 3.3 (g)). Die Produkt-Topologie erfüllt die Hausdorff-Eigenschaft. Außerdem gilt: eine Folge genau dann konvergiert, wenn sie punktweise konvergiert. Es gibt aber keine Metrik auf $\text{Abb}([0,1], \mathbb{R})$, für die eine Folge von Funktionen genau dann konvergiert, wenn sie punktweise konvergiert. Man kann sogar die stärkere Aussage[9] zeigen: Es gibt auf $C^0([0,1], \mathbb{R})$ keine Metrik d, so dass eine Folge von stetigen Funktionen genau dann punktweise konververgiert, wenn sie in $(C^0([0,1], \mathbb{R}), d)$ konvergiert: M. K. Fort, A note on pointwise convergence, Proc. AMS **2** (1951), 34–35.

BEMERKUNG 3.18. Ist X ein Hausdorffraum, $x \in X$. Dann ist $\{x\}$ abgeschlossen. (Begründung: Ist $y \in X \setminus \{x\}$, so liefert die Hausdorff-Eigenschaft eine offene Umgebung U_y von y, die in $X \setminus \{x\}$ enthalten ist. Also ist $X \setminus \{x\}$ offen.)

DEFINITION 3.19. Seien $(X, \mathcal{O}_X)$ und $(Y, \mathcal{O}_Y)$ Hausdorffräume. Eine Abbildung ist folgenstetig, wenn für alle Folgen $(x_j)_{j\in\mathbb{N}}$ in X und $a \in X$ gilt:

$$\text{Wenn } a = \lim_{j\to\infty} x_j, \text{ dann } f(a) = \lim_{j\to\infty} f(x_j)$$

[8] Man kann genau sagen, welche topologischen Räume von einer Metrik induziert werden und welche nicht. Der Metrisierungssatz von Bing, Nagata und Smirnow gibt hierfür Kriterien, die hinreichend und notwendig sind, siehe [**29**, Satz 10.14]. Die Formulierung und der Beweis dieses Satzes ist eine wichtige kulturelle Leistung der Mathematik, aber gebraucht habe ich bisher den Satz leider noch nie. Wenn man sich jedoch zu dem Satz durcharbeitet, lernt mal sehr viele nützliche Hilfsmittel kennen.

[9] Begründen Sie als Übung, wieso diese Aussage stärker ist!

LEMMA 3.20. Seien X und Y topologische Räume, $f : X \to Y$.

(1) Ist f stetig, dann ist f auch folgenstetig.
(2) Ist d_X bzw. d_Y eine Metrik auf X bzw. auf Y, die die Topologie auf X bzw. Y induziert. Dann ist f genau dann stetig, wenn f folgenstetig ist.

Beweis.
Zu (1): Es gelte $a = \lim_{j\to\infty} x_j$ in X. sei U eine Umgebung von $f(a)$ in Y. Da f stetig in a ist, ist $f^{-1}(U)$ eine Umgebung von a. Somit gilt $x_j \in f^{-1}(U)$ für fast alle $j \in \mathbb{N}$. Daraus folgt für diese j auch $f(x_j) \in U$. Es folgt $f(a) = \lim_{j\to\infty} f(x_j)$.

Zu (2): „$\Longrightarrow$“ ist nun gezeigt. Den Beweis von „$\Longleftarrow$“ erhalten wir aus dem Beweis von Lemma 1.7 „$\Longrightarrow$“ aus Kapitel 4, indem wir $|x - x_0|$ durch $d_X(x, x_0)$ und $|f(x) - f(x_0)|$ durch $d_Y(f(x), f(x_0))$ ersetzen.[10] □

!ACHTUNG!. Es gibt eine Abbildung $f : X \to Y$, X und Y topologische Räume[11], mit den folgenden Eigenschaften

- f ist folgenstetig, d.h. für $x_i, x \in X$ folgt aus $x = \lim_{i\to\infty} x_i$ auch $f(x) = \lim_{i\to\infty} f(x_i)$.
- f ist nicht stetig

[10]Solche Beweise, in denen Teile in alten Beweisen ersetzt werden, mag ich eigentlich nicht. Denn man macht da ganz schnell Fehler. Um solche Verallgemeinerungen zu vermeiden, hätten wir aber noch früher metrische Räume einführen müssen. Und dies ist auch der Grund, wieso wir jetzt topologische Räume behandeln: damit wir später die kommenden Aussagen nicht noch einmal so ähnlich zeigen müssen.

[11]Dies geht natürlich nicht mit metrischen Räumen, wie oben gezeigt!

Da solche „exotischen“ Beispiele in der Vorlesung keine Rolle spielen werden, wird diese Tatsache nur im Skript kurz erwähnt.

4. Zusammenhang und Wegzusammenhang

Fr 17.5.

Schreibweise: X für $(X, \mathcal{O}_X)$, wenn klar ist, welche Topologie auf X zu wählen ist. Beispiel: $\mathbb{R}^n$ trägt (fast immer) die von der Standard-Norm induzierte Topologie.

Schreibweise: Disjunkte Vereinigung $U \overset{\bullet}{\cup} V = X$ bedeutet $U \cup V = X$ und $U \cap V = \emptyset$.

Definition 4.1. Ein topologischer Raum X heißt nicht zusammenhängend, wenn es nicht-leere offene Mengen $U, V \subset X$ mit $U \overset{\bullet}{\cup} V = X$ gibt. Zusammenhängend := nicht nicht zusammenhängend.

Ist $A \subset X$ und ist $\mathcal{O}_X$ eine Topologie auf X, dann definiert man auch :

A ist zusammenhängend $:\Longleftrightarrow (A, \mathcal{O}_A)$ ist zusammenhängend

Hierbei ist $\mathcal{O}_A$ die auf A induzierte Topologie (Spurtopologie).

Beispiele 4.2.

(a) Die Menge $\{0,1\} \subset \mathbb{R}$ ist nicht zusammenhängend.[12] Denn $\{0\} = B_{1/2}(0) \cap \{0,1\}$ ist offen in $\{0,1\}$. Analog ist $\{1\}$ offen in $\{0,1\}$. Und

$$\{0,1\} = \{0\} \overset{\bullet}{\cup} \{1\}.$$

(b) Mehrere Bilder, u.a. eines mit zwei disjunkten Bällen in $\mathbb{R}^2$.

Nicht zusammenhängend.

[12] Wenn nichts anderes angegeben ist, hat $\mathbb{R}$ immer die Standard-Topologie und eine Teilmenge von $\mathbb{R}$ die induzierte Topologie. In diesem Spezialfall ist die diskrete Topologie auf $\{0,1\}$.

(c) Sei $\mathbb{Q}$ mit der Standardtopologie (d.h. mit der von $d(x,y) = |x-y|$ induzierte Topologie) versehen. Dann ist $\mathbb{Q}$ nicht zusammenhängend. Definiere $U := (-\infty, \sqrt{2}) \cap \mathbb{Q}$ und $V := (\sqrt{2}, \infty) \cap \mathbb{Q}$. Die Mengen U und V sind offen in $\mathbb{Q}$. [13] $U \dot{\cup} V = X$.

(d) eine Teilmenge A von $\mathbb{R}$ ist zusammenhängend, gdw A ein Intervall ist. (Übungsblatt 5 Aufgabe 1)

PROPOSITION 4.3. Seien X und Y topologische Räume, $f : X \to Y$ stetig und X zusammenhängend. Dann ist $f(X)$ (mit der Spurtopologie von Y versehen) ebenfalls zusammenhängend.

Beweis. Angenommen $f(X)$ ist nicht zusammenhängend. Dann gibt es Mengen U und V mit den folgenden Eigenschaften:

- $U \dot{\cup} V = f(X)$,
- U offen in $f(X)$, d.h. es gibt eine offene Teilmenge $\widetilde{U}$ von Y mit $U = Y \cap \widetilde{U}$ (siehe Definition der Spurtopologie),
- V offen in $f(X)$, d.h. es gibt eine offene Teilmenge $\widetilde{V}$ von Y mit $V = Y \cap \widetilde{V}$ (siehe Definition der Spurtopologie).
- $U \neq \emptyset$ und $V \neq \emptyset$

Da f stetig ist, ist auch $f^{-1}(\widetilde{U}) = f^{-1}(U)$ offen, und offensichtlich gilt $f^{-1}(U) \neq \emptyset$. Analog für $f^{-1}(V)$. Außerdem haben wir $X = f^{-1}(U) \dot{\cup} f^{-1}(V)$. □

BEISPIEL 4.4. Ist $f : X \to \mathbb{R}$ stetig und X zusammenhängend. Dann ist $f(X)$ ein Intervall.

[13] Sie sind nicht offen in $\mathbb{R}$, aber offen in $\mathbb{Q}$. Wer dies nicht versteht, muss nochmals genau die Definition der Spurtopologie lesen.

Wenn wir den Spezialfall betrachten, dass $X = I$ ein Intervall ist, und $Y = \mathbb{R}$, so erhalten wir wieder die folgende Aussage, die wir als Korollar 2.4 in Kapitel 4 bereits bewiesen haben

KOROLLAR 2.4 in Kap. 4. Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ stetig. Dann ist $f(I)$ ebenfalls ein Intervall.

Aus diesem Korollar erhält man auch den Zwischenwertsatz. Proposition 4.3 verallgemeinert also den Zwischenwertsatz.

DEFINITION 4.5. Sei X ein topologischer Raum, $x, y \in X$. Ein Weg von x nach y in X ist eine stetige Abbildung $s : [0, 1] \to X$ mit $s(0) = x$ und $s(1) = y$.

Ein topologischer Raum X ist wegzusammenhängend, wenn es zu allen Punkten $x, y \in X$ einen Weg von x nach y gibt.

PROPOSITION 4.6. Ist X wegzusammenhängend, so ist X auch zusammenhängend.

Beweis: Übungsblatt 5.

BEISPIEL 4.7. Eine Menge $A \subset \mathbb{R}^n$ heißt sternförmig bezüglich $x_0 \in A$, falls gilt: ist $x \in A$ und $t \in [0, 1]$, dann haben wir auch $(1 - t)x_0 + tx \in A$.

Bild einer sternförmigen Menge mit Zentrum x_0 und einem Punkt x, sowie der Verbindungsstrecke dieser Punkte

Sternförmige Mengen sind wegzusammenhängend, also auch zusammenhängend.

BEMERKUNG 4.8. Wir betrachten die folgende Teilmenge von $\mathbb{R}^2$

$$M := \Big(\{0\} \times [-1, 1]\Big) \cup \{(x, \sin(\pi/x)) \mid x \in \mathbb{R}_{>0}\}$$

und versehen M mit der von der Standardtopologie auf $\mathbb{R}^2$ induzierten Topologie.

Bild dieser Menge

Die Menge M ist zusammenhängend, aber nicht wegzusammenhängend.

Übung 4.9. Ist X wegzusammenhängend und $f : X \to Y$ stetig. Dann ist auch $f(X)$ wegzusammenhängend.

5. Folgenkompaktheit

Definition 5.1. Sei $(X, \mathcal{O}_X)$ ein Hausdorffraum. Wir sagen, $(X, \mathcal{O}_X)$ ist folgenkompakt, wenn jede Folge in X eine (in $(X, \mathcal{O}_X)$) konvergente Teilfolge besitzt. Eine Teilmenge $A \subset X$ nennt man folgenkompakt, wenn $(A, \mathcal{O}_A)$ folgenkompakt ist, wobei $\mathcal{O}_A$ die Spurtopologie auf A ist.

Diese Definition verallgemeinert die analoge Definition für metrische Räume, siehe Definition 5.19 in Kapitel 4.

Satz 5.2. Sei $(V, \| \cdot \|)$ ein endlich-dimensionaler reeller oder komplexer Vektorraum. Wir versehen V mit der von $\| \cdot \|$ induzierten Topologie. Eine Teilmenge A von V ist genau dann folgenkompakt, wenn sie abgeschlossen und beschränkt ist.

Wdh: A beschränkt $\iff \exists C \in \mathbb{R}_{\geq 0} : \forall x \in A : \|x\| \leq C$

Beispiele 5.3. Ein Intervall I ist genau dann folgenkompakt, wenn es abgeschlossen und beschränkt ist, d.h. wenn es $a, b \in \mathbb{R}$, $a \leq b$ gibt mit $I = [a, b]$.

Die folgenden Teilmengen von $V = \mathbb{R}^n$ sind folgenkompakt:
Abgeschlossene Bälle $\overline{B}_r(x_0) := \{x \in V \mid \|x - x_0\| \leq r\}$ bezüglich beliebiger Normen.
Quader $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$.

Ist $(V, \| \cdot \|)$ ein normierter Vektorraum. Aus der Dreiecks-Ungleichung folgt

$$\Big| \|x\| - \|y\| \Big| \leq \|x - y\|$$

und hieraus folgt die Stetigkeit der Norm $V \to \mathbb{R}$, $x \mapsto \|x\|$.

Beweis von Satz 5.2. O.B.d.A. $V = \mathbb{R}^n$.

„$\Longrightarrow$“: Sei A folgenkompakt.

Dann ist A abgeschlossen nach Lemma 5.20 aus Kapitel 4 (Es besagt: Folgen-kompakte Teilmengen A in einem metrischen Raum X sind in X abgeschlossen.)

Angenommen A ist nicht beschränkt. Dann gibt es eine Folge (a_i) mit $\|a_i\| \to \infty$. Eine Teilfolge $(a_{f(i)})$ konvergiert dann gegen $y \in A$. Es folgt

$$\|y\| = \lim_{i \to \infty} \|a_{f(i)}\| = \infty.$$

Widerspruch.

„$\Longleftarrow$“: Ist A beschränkt, so ist jede A-wertige Folge $(a_i)_{i \in \mathbb{N}}$ beschränkt. Nach dem Satz von Bolzano-Weierstraß (Lemma 1.6 bzw. Korollar 1.9) besitzt die Folge eine konvergente Teilfolge $(a_{f(i)})_{i \in \mathbb{N}}$ mit Grenzwert a. Wenn A zusätzlich abgeschlossen ist, so konvergiert die Teilfolge in A gegen a. Somit ist A folgenkompakt. □

Satz 5.4 (Verallgemeinert Satz 5.22 in Kapitel 4)**.** Seien X, Y Hausdorffräume und sei $f : X \to Y$ folgenstetig[14]. Ist X folgenkompakt, dann ist auch $B := im(f)$ (versehen mit der Spurtopologie) folgenkompakt.

Der Beweis geht völlig gleich wie der Beweis von Satz 5.22 in Kapitel 4.

Beweis. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in B. Wähle $b_n \in X$ mit $f(b_n) = a_n$. Wähle eine konvergente Teilfolge $(b_{j(n)})$, $b := \lim_{n\to\infty} b_{j(n)}$. Da f stetig ist, haben wir

$$f(b) := \lim_{n\to\infty} \underbrace{f(b_{j(n)})}_{=a_{j(n)}}.$$

Somit ist eine konvergente Teilfolge von $(a_n)_{n\in\mathbb{N}}$ gefunden. Wir haben die Folgenkompaktheit von B überprüft. □

Beispiel 5.5. Ist $f : [a, b] \to \mathbb{R}$ stetig, dann ist $f([a, b])$ ein folgenkompaktes Intervall, d.h. $f([a, b]) = [c, d]$ für geeignete $c, d \in \mathbb{R}$.

Korollar 5.6. Ist X ein folgenkompakter Hausdorffraum und $f : X \longrightarrow \mathbb{R}$ (folgen-)stetig. Dann nimmt f ein Maximum und ein Minimum an, d.h. es gibt $x_1, x_2 \in X$, so dass für alle $x \in X$:

$$f(x_1) \leq f(x) \leq f(x_2).$$

Schreibweise: $\max f := \max f(X) = \max\{f(x) \mid x \in X\}$. $\min f$ analog.

Beweis. Schreibe das folgenkompakte Intervall $\mathrm{im}(f)$ als $[a_1, a_2]$. Wähle x_i mit $f(x_i) = a_i$. □

[14]Die Proposition ist natürlich auch richtig, wenn wir „folgenstetig“ durch „stetig“ ersetzen, denn jede stetige Funktion ist folgenstetig.

BEMERKUNG 5.7. Folgenkompakte metrische Räume sind vollständig. Denn besitzt eine Cauchy-Folge eine Teilfolge, die gegen x_0 konvergiert, so konvergiert die gesamte Cauchy-Folge gegen x_0.

Den Rest des Abschnitts wollen wir zunächst in der Vorlesung aus Zeitgründen überspringen. Die Resultate werden später im Beweis von „folgenkompakt $\Longrightarrow$ kompakt“ in Theorem 6.7 nutzen.

LEMMA 5.8. Sei (X, d) ein folgenkompakter metrischer Raum, und $\epsilon > 0$. Dann gibt es $k \in \mathbb{N}$ und $\{x_1, x_2, \dots x_k\} \subset X$, so dass

$$X = B_\epsilon(x_1) \cup B_\epsilon(x_2) \cup \cdots \cup B_\epsilon(x_k).$$

Beweis. O.B.d.A. $X \neq \emptyset$. Wähle ein $x_1 \in X$. Wähle nun weitere x_i rekursiv: Falls x_1, x_2,$\dots x_j$ gewählt sind, so setze

$$A_j := X \setminus (B_\epsilon(x_1) \cup B_\epsilon(x_2) \cup \cdots \cup B_\epsilon(x_j)).$$

Im Fall $A_j = \emptyset$ ist die Behauptung gezeigt. Andernfalls wähle ein $x_{j+1} \in U_j$.

Wenn die Aussage des Lemmas nicht gilt, so erhalten wir eine Folge $(x_i)_{i\in\mathbb{N}}$ mit $d(x_i, x_j) \geq \epsilon$ für alle $i \neq j$. Solch eine Folge besitzt keine konvergente Teilfolge. Also ist X nicht folgenkompakt. □

FOLGERUNG 5.9. Ist X ein folgenkompakter metrischer Raum, so existiert eine Folge $(x_j)_{j\in\mathbb{N}}$, so dass $\{x_j \mid j \in \mathbb{N}\}$ dicht ist. In anderen Worten: X besitzt eine abzählbare dichte Teilmenge.

Beweis. Zu $n \in \mathbb{N} \setminus \{0\}$ wähle $k_n \in \mathbb{N}$ und $x_j^n \in X$, so dass

$$X = \bigcup_{j=1}^{k_n} B_{1/n}(x_j^n).$$

Wir nehmen nun die Folge

$$(x_j)_{j\in\mathbb{N}} = (x_1^1, \dots x_{k_1}^1, x_1^2, \dots, x_{k_2}^2, x_1^3, \dots)$$

Sei nun ein beliebiges $y \in X$ gegeben. Zu zeigen ist, dass y im Abschluss von $A := \{x_j \mid j \in \mathbb{N}\}$ liegt. Zu jedem $n \in \mathbb{N}$ wähle ein j_n, so dass $y \in B_{1/n}(x_{j_n}^n)$, also $d(y, x_{j_n}) < 1/n$. Also ist $(x_{j_n})_{n\in\mathbb{N}}$ eine Folge in A, die gegen y konvergiert. Somit $y \in \overline{A}$. □

6. Kompaktheit

Wir führen nun eine andere Eigenschaft ein, die sich „Kompaktheit“ nennt, und werden dann die erstaunliche und weitreichende Tatsache beweisen, dass metrische Räume genau dann kompakt sind, wenn sie folgenkompakt sind.

Definition 6.1. Sei X ein topologischer Raum.

(1) Eine offene Überdeckung von X ist eine Familie $(U_i)_{i\in I}$ von offenen Mengen[15] mit $X = \bigcup_{i\in I} U_i$.
(2) Wir sagen X erfüllt die Heine-Borel-Eigenschaft, falls gilt:
Zu **jeder** offenen Überdeckung $(U_i)_{i\in I}$ von X gibt es ein $k \in \mathbb{N}$ und $i_1, i_2, \dots, i_k \in I$ mit

$$X = U_{i_1} \cup U_{i_2} \cup \dots \cup U_{i_k}.$$

[15]Gemeint ist damit natürlich: alle U_i sind offen in X.

[16] Man nennt eine solche Überdeckung eine endliche Teilüberdeckung.[17]

(3) Man nennt X kompakt, falls X ein Hausdorffraum ist und die Heine-Borel-Eigenschaft erfüllt.[18]

Lemma 6.2. Wir versehen X mit der diskreten Topologie. Dann ist X kompakt, genau dann wenn X endlich ist.

Beweis. Offensichtlich ist jede endliche Menge kompakt.

Angenommen X ist kompakt und diskret (d.h. mit der diskreten Topologie versehen). Für $x \in X$ ist $U_x := \{x\}$ offen. Also ist $(U_x)_{x\in X}$ eine offene Überdeckung von X. Da X kompakt ist, gibt es eine endliche Menge $J \subset X$, so dass $X = \bigcup_{x\in J} U_x = J$. Also ist bereits X endlich. □

Proposition 6.3. Sei X ein kompakter topologischer Raum und sei A eine abgeschlossene Teilmenge. Dann ist A ebenfalls kompakt.

[16] Man wundert sich vielleicht, wieso wir nicht $k \in \mathbb{N} \setminus \{0\}$ genommen haben. Dies liegt am Spezialfall $X = \emptyset$, der $(U_i)_{i\in\emptyset}$ als offene Überdeckung hat. Für ihn wählt man $k = 0$. Die leere Menge erfüllt also die Heine-Borel-Eigenschaft und ist kompakt.

[17] Teilüberdeckung bedeutet hier nicht, dass der Raum nur zum Teil überdeckt wird. Er wird ganz überdeckt. Die Aussage ist vielmehr: Man kann alle bis auf endlich viele der U_i's weglassen und hat immer noch eine Überdeckung. Man überdeckt mit einem endlichen Teil der offenen Mengen.

[18] Achtung: Manche Mathematiker nennen einen topologischen Raum kompakt, sobald er die Heine-Borel-Eigenschaft erfüllt, ohne die Hausdorff-Eigenschaft einzuschließen. Wenn Sie ein Buch oder Skript zu diesem Thema in die Hand nehmen, müssen Sie immer überprüfen, welche Definition nun benutzt wird. Ähnlich wie in der Frage, ob 0 eine natürliche Zahl ist oder nicht, gibt es hier eben verschiedene übliche Konventionen.

Beweis. Sei $(U_i)_{i\in I}$ eine offene Überdeckung von A. Da A die Spurtopologie von X trägt, gibt es Mengen $\widetilde{U}_i$, offen in X mit $U_i = A \cap \widetilde{U}_i$. Sei $i_0 \notin I$, $J := I \cup \{i_0\}$, $\widetilde{U}_{i_0} := X \smallsetminus A$. Dann ist $(\widetilde{U}_i)_{i\in J}$ eine offene Überdeckung von X. Da X kompakt ist, gibt es eine endliche Menge $K \subset J$, so dass $(\widetilde{U}_i)_{i\in K}$ bereits X überdeckt. Dann überdeckt $(U_i)_{i\in K\smallsetminus\{i_0\}}$ bereits A. □

Übung 6.4. () Zeigen Sie, dass Lemma 6.2 und Proposition 6.3 auch dann gelten, wenn man „kompakt“ durch „folgenkompakt“ ersetzt.

Proposition 6.5. Sei X ein Hausdorffraum und A eine kompakte Teilmenge [19]. Dann ist A abgeschlossen.

Beweis. Wir zeigen, dass $X \smallsetminus A$ offen ist. Sei $y \in X \smallsetminus A$. Zu jedem $x \in A$ gibt es auf Grund der Hausdorff-Eigenschaft disjunkte offene Mengen U_x und U_y^x, so dass $x \in U_x$ und $y \in U_y^x$. Achtung: Wie die Notation ja bereits andeutet, hängt U_y^x von x ab! Nun ist $(A \cap U_x)_{x\in A}$ eine offene Überdeckung von A. Auf Grund der Kompaktheit von A gibt es $x_1, \ldots, x_k \in A$ mit

$$A = (A \cap U_{x_1}) \cup \ldots \cup (A \cap U_{x_k})$$

Wir setzen

$$V_y := \bigcap_{i=1}^{k} U_y^{x_i} \qquad \text{offen in } X.$$

Also ist $X \smallsetminus A = \bigcup_{y\in X\smallsetminus A} V_y$ ebenfalls offen. □

Proposition 6.6. Seien X, Y Hausdorffräume und sei $f : X \to Y$ stetig. Ist X kompakt, dann ist auch das Bild $im(f) = f(X)$ (versehen mit der Spurtopologie) kompakt.

[19] wie immer: bezüglich der Spur-Topologie

Beweis. Sei $(U_i)_{i\in I}$ eine offene Überdeckung von $f(X)$, U_i offen in $f(X)$. Dann ist $\left(f^{-1}(U_i)\right)_{i\in I}$ eine offene Überdeckung von X. Wegen der Kompaktheit von X gibt es eine endliche Menge J, so dass bereits $\left(f^{-1}(U_i)\right)_{i\in J}$ die Menge X überdeckt. Wegen $f(f^{-1}(U_i)) = U_i \cap f(X) = U_i$ folgt, dass $(U_i)_{i\in J}$ bereits $f(X)$ überdeckt. □

THEOREM 6.7. Sei X ein metrischer Raum. Dann ist X genau dann folgenkompakt, wenn X kompakt ist.

Beweis des Theorems.

„kompakt $\Longrightarrow$ folgenkompakt“: Sei X kompakt. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in X.

1. Fall: $A := \{a_n \mid n \in \mathbb{N}\}$ besitzt keinen Häufungspunkt in X
Also ist A abgeschlossen, und somit kompakt. Außerdem ist die Topologie auf A diskret. Wegen Lemma 6.2 ist dann A endlich. Nach dem Schubfach-Prinzip [20] gibt es also ein $a \in A$, so dass $I := \{n \in \mathbb{N} \mid a_n = a\}$ unendlich ist. Dann ist $(a_n)_{n\in I}$ eine Teilfolge[21] von $(a_n)_{n\in\mathbb{N}}$, die konstant gleich a ist und somit gegen a konvergiert.

2. Fall: $\{a_n \mid n \in \mathbb{N}\}$ besitzt einen Häufungspunkt x_0 in X
Es folgt mit Lemma 5.17 aus Kapitel 4 die Existenz einer konvergenten Teilfolge.

„folgenkompakt $\Longrightarrow$ kompakt“:

[20] Das sogenannte Schubfach-Prinzip besagt anschaulich: wenn ich unendlich viele Kugeln in endlich viele Schubladen einsortieren, dann habe ich einer Schublade (= Schubfach) unendlich viele Kugeln. Mathematisch präziser: Ist eine Funktion $g : \mathbb{N} \to A$, gegeben mit A endlich, so gibt es ein $a \in A$ mit unendlichem $g^{-1}(a)$. Es sollte Ihnen inzwischen klar sein, wie man des Schubfach-Prinzips beweist.

[21] Wir schreiben hier Teilfolgen zum ersten Mal auf eine andere Art und Weise. Wenn man die eindeutige streng monoton wachsende Bijektion $f : \mathbb{N} \to I$ wählt, so ist mit $(a_n)_{n\in I}$ die Teilfolge $(a_{f(n)})_{n\in\mathbb{N}}$ gemeint.

Sei X folgenkompakt. Gegeben sei eine offene Überdeckung $(U_i)_{i\in I}$ von X.

Wir zeigen in einem ersten Schritt, dass es eine abzählbare Menge $J \subset I$ gibt, so dass $(U_i)_{i\in J}$ eine (offene) Überdeckung ist. [22]. In einem zweiten Schritt zeigen wir dann, dass es eine endliche Menge $K \subset J$ gibt, so dass $(U_i)_{i\in K}$ eine (offene) Überdeckung ist.

1. Schritt:
Sei D eine abzählbare dichte Teilmenge von X, die es auf Grund von Folgerung 5.9 in X gibt. Dann ist

$$\mathcal{B} := \{B_{1/n}(p) \mid n \in \mathbb{N}, \quad p \in D\}$$

eine abzählbare Menge offener Mengen. Man sieht schnell, dass jede offene Menge in X eine Vereinigung geeigneter Elemente von $\mathcal{B}$ ist. Sei nun

$$\mathcal{B}' := \{U \in \mathcal{B} \mid \exists i \in I : U \subset U_i\}$$

und zu jedem $U \in \mathcal{B}'$ wählen wir solch ein $i = f(U)$, also $U \subset U_{i(U)}$, $f : \mathcal{B}' \to I$. Sei $J := f(\mathcal{B}')$. Da $\mathcal{B}$ abzählbar ist, ist es auch $\mathcal{B}'$ und somit J.

Wir müssen für diesen Schritt noch zeigen, dass $(U_i)_{i\in J}$ überdeckt. Sei also $x \in X$. Dann gibt es $i \in I$ mit $x \in U_i$. Wähle $n \in \mathbb{N}$, so dass $B_{2/n}(x) \subset U_i$. Wähle ein $p \in D$ mit $d(p, x) < 1/(2n)$. Dann gilt $x \in B_{1/n}(p) \subset B_{2/n}(x) \subset U_i$ und $B_{1/n}(p) \in \mathcal{B}'$. Also auch $x \in U_{f(B_{1/n}(p))}$.

2. Schritt:
Sei nun also $(U_i)_{i\in J}$ eine offene Überdeckung mit J abzählbar. O.B.d.A. $J = \mathbb{N}$. Wir nehmen an, dass $V_n := \bigcup_{i=1}^{n} U_i$ eine echte Teilmenge von X ist[23] für alle $n \in \mathbb{N}$. Wähle

[22] Der erste Schritt wird in der Vorlesung übersprungen

[23] Erinnerung: „A echte Teilmenge von X“ bedeutet: $A \subset X$ und $A \neq X$.

$x_n \in X \setminus V_n$. Wähle eine konvergente Teilfolge $(x_{f(n)})_{n\in\mathbb{N}}$ der Folge $(x_n)_{n\in\mathbb{N}}$. Es gibt nun ein $j \in \mathbb{N}$, so dass $\lim_{n\to\infty} x_{f(n)} \in U_j$. Da U_j offen ist, liegt für fast alle $n \in \mathbb{N}$ dann $x_{f(n)} \in U_j$. Unter anderem gibt es ein $n \in \mathbb{N}$, $f(n) \geq j$, mit $x_{f(n)} \in U_j$, was offensichtlich ein Widerspruch zur Konstruktion der x_j ist.

Also gibt es ein V_n mit $V_n = X$. □

Bemerkung 6.8. Es gibt Hausdorffräume, die kompakt, aber nicht folgenkompakt sind. Es gibt auch Hausdorffräume, die folgenkompakt, aber nicht kompakt sind. Siehe im Buch [**32**].

Mehr zum Thema folgenkompakte/kompakt inklusive schöner Bilder findet man hier: `http://simomaths.wordpress.com/2013/02/24/topology-sequentially-compact-spaces-and-comp`

Beispiel 6.9. Für eine Teilmenge A eines endlich-dimensionalen normierten Vektorraums sind äquivalent:

(1) A ist beschränkt und abgeschlossen
(2) A ist kompakt
(3) A ist folgenkompakt

Mi 22.5.

Definition 6.10. Seien X und Y topologische Räume. Eine Abbildung $f : X \to Y$ nennt man Homöomorphismus, falls gilt:

- f ist stetig
- f ist bijektiv
- $f^{-1} : Y \to X$ ist stetig

In anderen Worten: f ist eine Bijektion mit der Eigenschaft

$$U \text{ offen in } X \quad \Longleftrightarrow \quad f(U) \text{ offen in } Y.$$

Existiert ein Homöomorphismus von X nach Y, so nennt man X und Y homöomorph. Homöomorphie ist eine Äquivalenzrelation auf der Klasse [24] aller topologischer Räume.

Satz 6.11. Ist X ein kompakter topologischer Raum, Y ein Hausdorffraum und $f : X \to Y$ bijektiv und stetig, dann ist f ein Homöomorphismus.

Beweis. Aufgabe 1 auf Übungsblatt 6.

□

Definition 6.12. Sind (M, d^M) und (N, d^N) metrische Räume. Eine Abbildung $f : M \to N$ heißt gleichmäßig stetig, falls

$$\forall \epsilon \in \mathbb{R}_{>0} : \exists \delta \in \mathbb{R}_{>0} : \forall x, y \in M : \big(d^M(x, y) < \delta \Longrightarrow d^N(f(x), f(y)) < \epsilon\big). \tag{6.13}$$

Lemma 6.14. Sind (M, d^M) und (N, d^N) metrische Räume, und $f : M \to N$ stetig. Ist M kompakt, so ist f gleichmäßig stetig.

Insbesondere folgt hieraus das noch nicht bewiesene Lemma 4.1 in Kapitel 6.

Beweis. Angenommen f ist nicht gleichmäßig stetig. Es gibt dann ein $\epsilon > 0$, so dass

$$\forall \delta \in \mathbb{R}_{>0} : \exists x, y \in M : d^M(x, y) < \delta \wedge d^N(f(x), f(y)) \geq \epsilon. \tag{6.15}$$

[24] Die Gesamtheit aller topologischer Räume ist so groß, dass es keine Menge mehr ist, sondern eine Klasse. Siehe den Anfang der Analysis I.

Zu jedem $n \in \mathbb{N} \setminus \{0\}$ und $\delta := \delta_n := 1/n$ wählen wir $x_n, y_n \in M$ wie in (6.15). Da M kompakt und somit folgenkompakt ist, gibt es eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$, so dass

$$x := \lim_{k\to\infty} x_{n_k}$$

in M existiert. Wegen $d(x_{n_k}, y_{n_k}) < 1/n_k$ gilt auch $x = \lim_{k\to\infty} y_{n_k}$. Da f (folgen-)stetig ist gilt dann auch

$$f(x) = \lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} f(y_{n_k}),$$

also

$$\lim_{k\to\infty} d^N\big(f(x_{n_k}), f(y_{n_k})\big) = 0.$$

Dies widerspricht jedoch der obigen Wahl von x_{n_k} und y_{n_k}, denn diese haben wir so gewählt, dass für alle $k \in \mathbb{N}$ die Ungleichung $d^N\big(f(x_{n_k}), f(y_{n_k})\big) \geq \epsilon$ gilt.[25]

Wir haben gezeigt, dass die Annahme, f sei nicht gleichmäßig stetig, einen Widerspruch impliziert. Also muss f gleichmäßig stetig sein. □

Sei (X, d) ein metrischer Raum und $\emptyset \neq Y \subset X$. Definiere für $x \in X$

$$d(x, Y) := \inf\{d(x, y) \mid y \in Y\}$$

Dann gilt

$$d(x, Y) = 0 \iff x \in \overline{Y},$$
$$d(x, X \setminus U) \geq r \iff B_r(x) \subset U.$$

Übung 6.16 (Aufgabe 4 (ii) (a) auf Übungsblatt 3)**.** Die Funktion $X \to \mathbb{R}$, $x \mapsto d(x, Y)$ ist stetig.

[25]Und natürlich für das feste oben gewählte $\epsilon > 0$!

Beweis. (Lösung der Aufgabe)

$$d(x_1, y) \leq d(x_1, x_2) + d(x_2, y)$$

Nehme das Infimum über $y \in Y$:

$$d(x_1, Y) \leq d(x_1, x_2) + d(x_2, Y)$$

und analoges gilt, wenn wir x_1 und x_2 vertauschen. Es folgt

$$|d(x_1, Y) - d(x_2, Y)| \leq d(x_1, x_2).$$

Daraus folgt, dass $x \mapsto d(x, Y)$, $X \to \mathbb{R}$ stetig ist. □

PROPOSITION 6.17. Sei (X, d) ein kompakter metrischer Raum und $(U_i)_{i \in I}$ eine offene Überdeckung. Dann gibt es ein $\epsilon > 0$ mit der folgenden Eigenschaft: Zu jedem $x \in X$ gibt es ein $i \in I$ mit $B_\epsilon(x) \subset U_i$.

Fr 24.5. Solch ein ϵ nennt man eine Lebesguesche Zahl.

Eine kompakte Teilmenge in $\mathbb{R}^2$ mit Überdeckung und darin ϵ-Bälle.

Beweis. Da X kompakt ist, gibt es eine endliche Teilüberdeckung $(U_i)_{i \in J}$, also $J \subset I$ endlich. O.B.d.A. $J = \{1, 2, \ldots, k\}$. Die Funktion

$$f : X \to \mathbb{R}, \quad x \mapsto \frac{1}{k} \sum_{i=1}^{k} d(x, X \smallsetminus U_i)$$

ist stetig. Gilt $x \in U_i$, so ist $d(x, X \smallsetminus U_i) > 0$. Da jedes $x \in X$ in $U_1 \cup \ldots \cup U_k$ enthalten ist, gilt $f(x) > 0$. Da X kompakt ist, nimmt f sein Minimum in einem Punkt $x_0 \in X$ an.

Setze

$$\epsilon := \min f = f(x_0) > 0.$$

Aus $f(x) \geq \epsilon$ folgt, dass es mindestens ein $i \in \{1, 2, \ldots, k\}$ gibt mit $d(x, X \setminus U_i) \geq \epsilon$, also $B_\epsilon(x) \subset U_i$. □

KAPITEL 9

Differential-Rechnung für Funktionen in mehreren Veränderlichen

1. Vorbemerkungen

Die Worte „Veränderliche“ und „Variable“ haben im folgenden dieselbe Bedeutung.

In diesem Kapitel versehen wir $\mathbb{R}^n$ mit irgendeiner Norm $\| \cdot \|$. Die davon induzierte Topologie ist die Standard-Topologie, denn alle Normen auf $\mathbb{R}^n$ sind äquivalent.

Standard-Basis von $\mathbb{R}^n$: $(e_1, e_2, \ldots, e_n)$.

$$\mathcal{L}(\mathbb{R}^n, \mathbb{R}^k) \quad := \quad \big\{\text{Lineare Abbildungen von } \mathbb{R}^n \text{ nach } \mathbb{R}^k\big\} \quad \cong \quad \{k \times n\text{-Matrizen}\} \quad =: \quad \mathbb{R}^{k\times n}$$

Bei Prof. Cisinski wurde $M_{k\times n}(\mathbb{K})$ an Stelle von $\mathbb{K}^{k\times n}$ verwendet.

Abweichend von unserer bisherigen Schreibweise, übernehmen wir von nun an die in der Linearen Algebra übliche Konvention, dass Elemente von $\mathbb{R}^n$ als Spaltenvektoren geschrieben werden. Wir schreiben auch

$$x = (x_1, x_2, \ldots, x_n)^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

LEMMA 1.1. Ist $A : \mathbb{R}^n \to \mathbb{R}^k$ eine lineare Abbildung, dann ist A stetig.

Beweis. Schreibe $x = (x_1, \ldots, x_n)^T = \sum_{j=1}^n x_j e_j \in \mathbb{R}^n$ und

$$Ax = \sum_{i=1}^{k} \sum_{j=1}^{n} a_{ij} x_j e_i.$$

Da alle Normen auf $\mathbb{R}^n$ und $\mathbb{R}^k$ äquivalent sind, reicht es, die Behauptung für die Supremumsnorm zu zeigen.

$$\|Ax\|_\infty \leq \underbrace{\left(\max_{i \in \{1,2\ldots,k\}} \sum_{j=1}^{n} |a_{ij}| \right)}_{C:=} \|x\|_\infty \leq C\|x\|_\infty.$$

Zu gegebenem $\epsilon > 0$ setze $\delta := \epsilon / C$. Dann gilt

$$d_\infty(x, y) = \|x - y\|_\infty < \delta \quad \Longrightarrow \quad d_\infty(Ax, Ay) = \|Ax - Ay\|_\infty = \|A(x - y)\|_\infty < \epsilon.$$

□

KOROLLAR 1.2. Sind V und W endlich-dimensionale normierte Vektorräume, und $A : V \to W$ linear, dann ist A stetig.

BEMERKUNG 1.3. Sei V ein Vektorraum mit Basis $(b_n \mid n \in \mathbb{N})$, den wir mit der zugehörigen Supremumsnorm

$$\|\sum_{n=0}^{\ell} x_n b_n\|_\infty := \max_{n\in\{0,\dots,\ell\}} |x_n|$$

versehen. Dann gibt es genau eine lineare Abbildung $A : V \to \mathbb{R}$ mit $A(b_n) = n$. Diese Abbildung ist nicht stetig. Denn setzen wir $a_n := \frac{1}{\sqrt{n}} b_n \in V$, dann ist $(a_n)_{n\in\mathbb{N}\smallsetminus\{0\}}$ eine Nullfolge in $(V, \| \cdot \|_\infty)$, aber $f(a_n) = \sqrt{n} \to \infty$ für $n \to \infty$.

BEMERKUNG 1.4. Das meiste, was wir in diesem Kapitel machen, gilt ebenso (zumeist mit denselben Beweisen), wenn man

- $\mathbb{R}^n$, $\mathbb{R}^k$ und ähnliche Räume durch beliebige (auch unendlich-dimensionale) Banachräume[1] ersetzt und
- lineare Abbildungen (und endliche Matrizen) durch stetige lineare Abbildungen ersetzt.

Für Details verweisen wir auf das Buch [**2**]. Der mathematische Mehraufwand ist gering. In der Vorlesung schränken wir uns auf Räume $\mathbb{R}^n$, $\mathbb{R}^k$,... ein, damit klar ist, dass dies zunächst die wichtigen Beispiele sind, die Sie möglichst gründlich verstehen müssen. Es folgt dann auch unmittelbar[2], dass alle von uns bewiesenen Aussagen dann auch in endlich-dimensionalen normierten Vektorräumen V gelten, indem wir eine Basis von V wählen und dann mit $\mathbb{R}^n$ identifizieren.

[1] also vollständige normierte Vektorräume

[2] Das heißt hier: ohne nochmals in die Beweise schauen zu müssen.

2. Differenzierbarkeit in mehreren Variablen

DEFINITION 2.1. Sei $U \subset \mathbb{R}^n$ offen, $p \in U$. Wir sagen, $f : U \to \mathbb{R}^k$ ist total differenzierbar in p oder differenzierbar in p, falls es eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^k$ gibt mit

$$\lim_{h \to 0} \frac{f(p+h) - f(p) - Ah}{\|h\|} = 0. \tag{2.2}$$

Wir sagen, f ist differenzierbar auf $W \subset U$, wenn f in allen $p \in W$ differenzierbar ist. Und f ist differenzierbar, wenn f auf U differenzierbar ist.

[3]

PROPOSITION 2.3. Sei f, U, p wie oben. Dann gibt es höchstens ein $A \in \mathbb{R}^{k \times n}$, für das (2.2) gilt.

Falls f differenzierbar in p ist, dann nennt man das obige (eindeutig bestimmte) A die (totale) Ableitung oder das (totale) Differential von f in p und schreibt sie als $D_p f := A$ und $f'(p) := A$.

PROPOSITION 2.4. Ist f differenzierbar in p, dann ist f stetig in p.

Beweis der beiden vorangehenden Propositionen. (2.2) ist äquivalent zu der Aussage: Es gibt eine in p stetige Funktion $r : U \to \mathbb{R}^k$ mit $r(p) = 0$ und

$$f(p+h) = f(p) + Ah + r(p+h)\|h\| \qquad \text{für alle } h \in \mathbb{R}^n \text{ mit } p + h \in U.$$

[3] So wie die Definitionen bisher gewählt sind, ist das Wörtchen „total“ ein erlaubtes Vorwörtchen, hat aber keine inhaltliche Bedeutung. Später werden wir auch den Begriff „partiell differenzierbar“ kennen lernen. Man benutzt dann das Vorwörtchen “total“ um nochmals zu betonen, dass hier nicht „partiell differenzierbar“ gemeint ist. Aber selbst dann ist die Bedeutung von „differenzierbar“ und „total differenzierbar“ die gleiche.

Wir schreiben x für $p+h$. Die Funktionen $U \ni x \mapsto f(p) \in \mathbb{R}^k$, $x \mapsto A(x-p)$ und $x \mapsto r(x)\|x-p\|$ sind stetig in p und somit auch $x \mapsto f(x)$. Also ist Proposition 2.4 gezeigt.

Wir zeigen nun Proposition 2.3. Angenommen (2.2) gilt für $A := A_1 \in \mathbb{R}^{k\times n}$ und $A := A_2 \in \mathbb{R}^{k\times n}$. O.B.d.A. sei $U \supset B_\epsilon(p_0)$. Bestimme[4] nun zugehörige r_i zu A_i. Es gilt für $x = p+h \in U$:

$$A_1 h + r_1(p+h)\|h\| = A_2 h + r_2(p+h)\|h\|$$

Wenn wir nun für $t \in [0,1]$ die Variable h durch th ersetzen, ergibt sich

$$t(A_1-A_2)(h) = (A_1-A_2)(th) = (r_2(p+th)-r_1(p+th))\ \|th\| = (r_2(p+th)-r_1(p+th))\ t\|h\|.$$

Wir dividieren durch t und bekommen:

$$(A_1 - A_2)h = \underbrace{(r_2(p+th) - r_1(p+th))}_{\to 0}\ \|h\| \to 0$$

für $t \to 0$. Da dies für alle $h \in B_\epsilon(0)$ gilt, gilt es insbesondere für $h := \frac{\epsilon}{2} e_\ell$, also $(A_1 - A_2)(e_\ell) = 0$. Es folgt $A_1 = A_2$. □

Beispiel 2.5. Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, $(x_1, x_2)^T \mapsto x_1 x_2$ ist differenzierbar in jedem Punkt $p = (p_1, p_2)^T$. Denn, wenn wir $A := (p_2, p_1)$ setzen[5], so gilt für $x = (x_1, x_2)^T$ und $h = (h_1, h_2)^T$

$$f(p+h) - f(p) - Ah = (p_1+h_1)(p_2+h_2) - p_1p_2 - (p_2h_1 + p_1h_2) = h_1h_2 = r(p+h)\|h\|,$$

[4] Achtung: in der Vorlesung habe ich die Funktionen r_i etwas anders definiert. Das ändert aber wenig an der Beweisstruktur.

[5] kein Tippfehler bei A: es ist ein Zeilenvektor!

wobei wir für $h \in \mathbb{R}^2 \setminus \{0\}$ definieren

$$r(p+h) := \frac{h_1 h_2}{\|h\|}$$

und $r(p) = 0$. Es gilt im Fall $\|\cdot\| = \|\cdot\|_\infty$:

$$\left|\frac{h_1 h_2}{\|h\|_\infty}\right| \leq \left|\frac{\|h\|_\infty \|h\|_\infty}{\|h\|_\infty}\right| \leq \|h\|_\infty,$$

das heißt in diesem Fall haben wir $\lim_{h\to 0} r(p+h) = 0 = r(p)$. Wegen der Äquivalenz aller Normen auf gilt diese Aussage auch für beliebige Normen $\|\cdot\|$. Also ist f differenzierbar in p und $f'(p) = (p_2, p_1)$.

Bemerkung 2.6. Differenzierbarkeit ist eine lokale Eigenschaft. D.h. ist U offen in $\mathbb{R}^n$, $p \in U$ und stimmen die Funktionen $g, f : U \to \mathbb{R}^k$ auf einer Umgebung von p überein, dann gilt:

(a) f ist genau dann in p differenzierbar, wenn g in p differenzierbar ist.
(b) Wenn f in p differenzierbar ist, dann gilt $f'(p) = g'(p)$.

Bemerkung 2.7. Schreibe $f : U \to \mathbb{R}^k$ in Komponenten $f(x) = (f_1(x), f_2(x), \ldots, f_k(x))^T$ mit $f_i : U \to \mathbb{R}$. Sei $p \in U$. Es gilt wegen Korollar 1.9 (b) in Kapitel 8

$$f \text{ ist stetig in } p \quad \Longleftrightarrow \quad f_1,\, f_2, \ldots \text{und } f_k \text{ sind stetig in } p$$

$$\lim_{h\to 0} f(p+h) = \begin{pmatrix} \lim\limits_{h\to 0} f_1(p+h) \\ \lim\limits_{h\to 0} f_2(p+h) \\ \vdots \\ \lim\limits_{h\to 0} f_k(p+h) \end{pmatrix}$$

Diese Zeile ist so zu lesen: wenn eine der beiden Seiten existiert, dann auch die andere und dann sind die Werte gleich.

$$f \text{ ist differenzierbar in } p \iff f_1,\, f_2,\ldots \text{und } f_k \text{ sind differenzierbar in } p$$

$$f'(p) = \begin{pmatrix} f_1'(p) \\ f_2'(p) \\ \vdots \\ f_k'(p) \end{pmatrix} \in \mathbb{R}^{k\times n}$$

Also: $f_i'(p) \in \mathbb{R}^{1\times n}$ ist eine Zeile der Matrix $f'(p)$.

Bemerkung 2.8. Sind V und W Vektorräume, $\dim V = n$, $\dim W = k$, U offen in V, $f : U \to W$. Die Wahl von Basen liefert Isomorphismen $\varphi : \mathbb{R}^n \to V$ und $\psi : \mathbb{R}^k \to W$.

Wir sagen: f ist differenzierbar in $p \in U$ gdw $\tilde{f} := \psi^{-1} \circ f \circ \varphi$ ist differenzierbar in $\varphi^{-1}(p)$.

Diese Definition ist unabhängig von der Wahl der Basen. Man erhält ein

$$f'(p) \in \mathcal{L}(V, W)$$

das ebenfalls unabhängig von der Basis definiert ist, und

$$\psi^{-1} \circ f'(p) \circ \varphi = \tilde{f}'(\varphi^{-1}(p)).$$

Bemerkung 2.9. Ist $f : U \to \mathbb{R}^k$ differenzierbar, so erhalten wir eine Abbildung $f' : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k) \cong \mathbb{R}^{k\times n} \cong \mathbb{R}^{kn}$. Ist f' wieder differenzierbar (also f zweimal differenzierbar), so erhalten wir

$$f'' : U \to \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k)) \cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}^{kn}) \cong \mathbb{R}^{n^2 k}$$

Ist f r-mal differenzierbar, dann

$$f^{(r)} : U \to \mathbb{R}^{n^r k}$$

Wie berechnet man die Koeffizienten von $f'(p)$?

DEFINITION 2.10. Sei U offen in $\mathbb{R}^n$, $f : U \to \mathbb{R}^k$, $p \in U$,

(1) Zu $v \in \mathbb{R}^n \setminus \{0\}$ wähle $\epsilon > 0$ so klein, dass $f_{p,v} : (-\epsilon, \epsilon) \to \mathbb{R}^k$, $f_{p,v}(t) := f(p + tv)$ wohldefiniert ist. Wir sagen: f ist differenzierbar in p in Richtung v, falls $f_{p,v}$ in 0 differenzierbar ist. Wir nennen dann

$$\partial_{p,v} f := f'_{p,v}(0)$$

die Richtungsableitung von f an der Stelle p in Richtung v.

(2) Wir sagen f ist in p partiell differenzierbar falls f in Richtung e_i differenzierbar ist für alle $i \in \{1, 2, \ldots, n\}$. Schreibweise $p = (p_1, \ldots, p_n)$. Man nennt

$$\frac{\partial f}{\partial x_i}(p) := f'_{p,e_i}(0) = \frac{d}{dt}|_{t=0}\Big(t \mapsto f(p_1, p_2, \ldots, p_{i-1}, p_i + t, p_{i+1}, \ldots, p_n)\Big)$$

die i-te partielle Ableitung von f an der Stelle p.

Mi 29.5.

LEMMA 2.11. f, U, n, k, p wie oben. Ist f (total) differenzierbar in p, dann ist f in p in jede Richtung differenzierbar. Für alle $v \in \mathbb{R}^n$ gilt dann

$$\partial_{p,v} = f'(p) \cdot v.$$

Beweis. Gegeben sei ein $v \in \mathbb{R}^n \setminus \{0\}$. Wähle $\epsilon > 0$ so klein, dass für alle $t \in (-\epsilon, \epsilon)$ gilt: $tv \in U$. Mit $t \to 0$ gilt dann auch $tv \to 0$. Es gelte

$$\lim_{h \to 0} \frac{f(p + h) - f(p) - Ah}{\|h\|} = 0$$

für $A = f'(p) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k)$. Dann gilt auch

$$0 = \lim_{t\to 0} \frac{f(p+tv) - f(p) - A(tv)}{\|tv\|} = \frac{1}{\|v\|} \lim_{t\to 0} \frac{f(p+tv) - f(p) - A(tv)}{|t|}$$

und somit

$$0 = \left(\lim_{t\to 0} \frac{f(p+tv) - f(p)}{t} - Av \right).$$

Also existiert die Richtungsableitung und nimmt den folgenden Wert an:

$$\partial_{p,v} f = f'_{p,v}(0) = \lim_{t\to 0} \frac{f(p+tv) - f(p)}{t} = Av$$

□

KOROLLAR 2.12. Ist f in p differenzierbar, dann ist die Ableitung in p:

$$f'(p) = \left(\frac{\partial f}{\partial x_1}(p) \quad \frac{\partial f}{\partial x_2}(p) \quad \dots \quad \frac{\partial f}{\partial x_n}(p) \right) \in \mathbb{R}^{k\times n}$$

Also ist $\frac{\partial f}{\partial x_i}(p)$ die i-te Spalte der Matrix $f'(p)$.

BEMERKUNG 2.13. Es gilt also für U, p, n, k, f wie oben:

$$f \text{ diffbar in } p \quad \Longrightarrow \quad f \text{ diffbar in jede Richtung in } p \quad \Longrightarrow \quad f \text{ partiell diffbar in } p.$$

Aber man kann durch Beispiele zeigen, dass die Umkehrungen nicht gelten:

$$f \text{ diffbar in } p \quad \not\Longleftarrow \quad f \text{ diffbar in jede Richtung in } p \quad \not\Longleftarrow \quad f \text{ partiell diffbar in } p.$$

BEISPIEL 2.14 (Ausführlich diskutiert in der Zentralübung am 28.5., in der Vorlesung nur kurz).

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad f(x,y) := \begin{cases} \frac{2xy}{x^2+y^2} & \text{falls } (x,y) \neq 0 \\ 0 & \text{falls } x = y = 0 \end{cases}$$

f ist in 0 partiell differenzierbar und

$$\frac{\partial f}{\partial x}(0) = 0, \qquad \frac{\partial f}{\partial y}(0) = 0.$$

Aber für $v := \frac{1}{\sqrt{2}}(1,1)^T$ ist $t \mapsto f(tv)$ nicht einmal stetig in 0. Denn $f(tv) = 1$ für $t \neq 0$ und $f(0) = 0$. Also ist f in 0 nicht in Richtung v differenzierbar. Dadurch ist f auch nicht total differenzierbar in 0.

f ist also partiell differenzierbar in 0. Nach einem Basiswechsel ist f aber nicht mehr partiell differenzierbar.

Achsenkreuz, beschriftet x und y mit Höhenlinien der Funktion f.

Es gibt Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$, die in 0 in alle Richtungen differenzierbar sind mit $\forall v \in \mathbb{R}^2 : \partial_{p,v} f = 0$, die aber dennoch nicht in 0 stetig sind. Sie sind also auch in 0 nicht total differenzierbar.

PROPOSITION 2.15. f, U, n, k, p wie oben. Äquivalent sind:

(1) f auf U partiell differenzierbar und alle partiellen Ableitungen

$$\frac{\partial f}{\partial x_i} : U \to \mathbb{R}^k, \qquad i \in \{1, 2, \ldots, n\}$$

sind stetig.

(2) f auf U total differenzierbar und

$$f' : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k) \cong \mathbb{R}^{k \times n}$$

ist stetig

Ist (1) bzw. (2) erfüllt, so sagen wir f ist stetig differenzierbar.

Sei U offen in $\mathbb{R}^n$.

$$C^1(U,\mathbb{R}^k) := \{f : U \to \mathbb{R}^k \mid f \text{ ist stetig differenzierbar}\}.$$

$$C^{r+1}(U,\mathbb{R}^k) := \{f \in C^1(U,\mathbb{R}^k) \mid f' \in C^r(U,\mathbb{L}(\mathbb{R}^n,\mathbb{R}^k)) = C^r(U,\mathbb{R}^{k\times n})\}$$

Elemente in $C^r(U,\mathbb{R}^k)$ nennt man r-mal stetig differenzierbar.

Beweis der Proposition. „$(2) \Longrightarrow (1)$“ folgt direkt aus dem oben gesagten.

„$(1) \Longrightarrow (2)$“:

Zu zeigen ist die Differenzierbarkeit in einem beliebigen $p \in U$. Die Stetigkeit der Ableitung ergibt sich dann direkt aus der Stetigkeit der partiellen Ableitungen. Sobald wir wissen, dass f differenzierbar ist, erhalten wir aus den partiellen Ableitungen die totale Ableitung

$$A := \begin{pmatrix} \frac{\partial f}{\partial x_1}(p) & \frac{\partial f}{\partial x_2}(p) & \dots & \frac{\partial f}{\partial x_n}(p) \end{pmatrix} \in \mathbb{R}^{k\times n}. \tag{2.16}$$

Da man die Differenzierbarkeit komponentenweise überprüfen kann, können wir ohne Beschränkung der Allgemeinheit annehmen, dass $k = 1$.

Zu $\epsilon > 0$ wähle ein $\delta > 0$, so dass für alle $\eta \in \mathbb{R}^n$ mit $\|\eta\|_\infty < \delta$ gilt:

$$p + \eta \in U \text{ und für } i \in \{1,2,\dots,n\} \text{ gilt } \left|\frac{\partial f}{\partial x_i}(p+\eta) - \frac{\partial f}{\partial x_i}(p)\right| < \frac{\epsilon}{n}.$$

Zu $h = (h_1, \ldots, h_n)^T \in \mathbb{R}^n$ mit $\|h\|_\infty < \delta$ setze

$$\begin{aligned} q^{(0)} &:= p = (p_1, \ldots, p_n)^T \\ q^{(1)} &:= (p_1 + h_1, p_2, \ldots, p_n)^T, \\ &\vdots \qquad \vdots \\ q^{(\ell)} &:= (p_1 + h_1, \ldots, p_\ell + h_\ell, p_{\ell+1}, \ldots, p_n)^T \\ &\vdots \qquad \vdots \\ q^{(n)} &:= p + h = (p_1 + h_1, \ldots, p_n + h_n)^T \end{aligned}$$

Wir fixieren nun alle Variablen außer der i-ten Koordinate und wenden dann den 1. Mittelwertsatz der Analysis I auf die Funktion $t \mapsto f(q^{(i-1)} + te_i)$, definiert für $t \in [0, h_i]$ bzw. $t \in [h_i, 0]$ an. Wir erhalten

$$f(q^{(i)}) - f(q^{(i-1)}) \overset{\text{1. MWS}}{=} h_i \, \frac{\partial f}{\partial x_i} \underbrace{(p_1 + h_1, \ldots, p_{i-1} + h_{i-1}, p_i + \vartheta_i h_i, p_{i+1}, \ldots, p_n)}_{=:\xi^{(i)}}$$

für ein geeignetes $\vartheta_i \in [0, 1]$. Mit $\|\xi^{(i)} - p\|_\infty < \delta$, folgt

$$\left| f(q^{(i)}) - f(q^{(i-1)}) - h_i \frac{\partial f}{\partial x_i}(p) \right| \leq |h_i| \, \frac{\epsilon}{n}$$

Es ergibt sich für die in (2.16) definierte Matrix A

$$|f(p + h) - f(p) - Ah| \leq \sum_{i=1}^{n} \left| f(q^{(i)}) - f(q^{(i-1)}) - h_i \frac{\partial f}{\partial x_i}(p) \right| \leq \epsilon \, \|h\|_\infty.$$

Wir haben das ϵ-δ-Kriterium für den Grenzwert

$$\lim_{h \to 0} \frac{f(p + h) - f(p) - Ah}{\|h\|_\infty} = 0$$

überprüft, und dies besagt, dass f in p differenzierbar mit Ableitung A ist. [6] □

PROPOSITION 2.17 (Summenregel)**.** Seien U, n, k, p, wie oben. Seien $f, g : U \to \mathbb{R}^k$, $a, b \in \mathbb{R}$, $v \in \mathbb{R}^n \setminus \{0\}$. Sind f und g in p differenzierbar (bzw. in Richtung v differenzierbar, bzw. partiell differenzierbar), dann ist $af + bg : U \to \mathbb{R}^k$ ebenfalls in p differenzierbar (bzw. in Richtung v differenzierbar, bzw. partiell differenzierbar). Es gilt dann

$$(af + bg)'(p) = af'(p) + bg'(p)$$

bzw.

$$\partial_{p,v}(af + g) = a\partial_{p,v}f + b\partial_{p,v}g$$

bzw.

$$\frac{\partial(af + bg)}{\partial x_i}(p) = a\frac{\partial f}{\partial x_i}(p) + b\frac{\partial g}{\partial x_i}(p).$$

Beweis. Offensichtlich. □

BEMERKUNG 2.18. Sei $f = (f_1, \ldots, f_k)^T : U \to \mathbb{R}^k$ in $p \in U$ partiell differenzierbar. Wenn wir $f'(p)$ als Matrix in $\mathbb{R}^{k\times n}$ (und nicht als lineare Abbildung in $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^k)$) ansehen, dann gilt

$$f'(p) = \left(\frac{\partial f_i}{\partial x_j}(p)\right)_{\substack{i\in\{1,2,\ldots,k\}\\ j\in\{1,2,\ldots,n\}}}.$$

Man nennt diese Matrix oft auch die Jacobi-Matrix von f in p. Da wir auf $\mathbb{R}^n$ und $\mathbb{R}^k$ die Standardbasen verwenden und lineare Abbildungen mit Matrizen identifizieren $\mathbb{R}^{k\times n} =$

[6] Ist Ihnen klar, wieso es hilfreich war, uns auf den Fall $k = 1$ einzuschränken? Der Grund ist, dass wir den 1. Mittelwertsatz nutzen, den wir nur für Funktionen $I \to \mathbb{R}$, I ein Intervall, bewiesen haben; also für $k = 1$. Wir werden eine Verallgemeinerung des Mittelwertsatz auf Funktionen $f : U \to \mathbb{R}$, U offen in $\mathbb{R}^n$ zwar bald kennenlernen, eine Verallgemeinerung auf Funktionen nach $\mathbb{R}^k$, $k > 1$ ist jedoch nicht möglich, siehe Beispiel 2.24.

$\mathcal{L}(\mathbb{R}^k, \mathbb{R}^n)$, ist die Jacobi-Matrix dasselbe wie die Ableitung. Sobald man aber $\mathbb{R}^k$ und $\mathbb{R}^n$ durch beliebige Vektorräume V und W ersetzt, so ist die Ableitung $f'(p) \in \mathcal{L}(V, W)$ unabhängig von der gewählten Basis definiert. Die Jacobi-Matrix ist dann die Matrix-Darstellung von $f'(p)$ bezüglich zu wählender Basen.

Satz 2.19 (Kettenregel)**.** Sei U offen in $\mathbb{R}^n$, V offen in $\mathbb{R}^m$, $f : U \to \mathbb{R}^m$, $g : V \to \mathbb{R}^r$, $f(U) \subset V$. Ist f differenzierbar in p und g differenzierbar in $q := f(p)$, dann ist $g \circ f$ differenzierbar in p und

$$(g \circ f)'(p) = g'(q) \cdot f'(p).$$

Man kann $g'(q) \cdot f'(p)$ als Verkettung der linearen Abbildungen $f'(p) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ und $g'(q) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^r)$ ansehen, [7] oder $f'(p)$ und $g'(q)$ als Matrizen betrachten (also als Jacobi-Matrizen!) und dann gilt $f'(p) \in \mathbb{R}^{m \times n}$, $g'(q) \in \mathbb{R}^{r \times m}$, $g'(q) \cdot f'(p) \in \mathbb{R}^{r \times n}$.

!ACHTUNG!. Es gilt **nicht** $g'(q) \cdot f'(p) = f'(p) \cdot g'(q)$. Die rechte Seite ist für $r \neq n$ gar nicht definiert, und auch im Fall $n = r$ gilt hier im allgemeinen keine Gleichheit.

Beweis der Kettenregel. (Bewiesen am Fr. 31.5.)

$$f(p+h) = f(p) + f'(p) \cdot h + \|h\|\epsilon(h), \qquad \lim_{h \to 0} \epsilon(h) = 0$$

$$g(q+k) = g(q) + g'(q) \cdot k + \|k\|\delta(k), \qquad \lim_{k \to 0} \delta(k) = 0$$

[7] Bei dieser Interpretation wäre es eigentlich sauberer $g'(q) \circ f'(p)$ an Stelle von $g'(q) \cdot f'(p)$ zu schreiben. Da wir aber sowieso lineare Abbildungen mit Matrizen identifizieren, ist es auch sinnvoll die Verkettung linearer Abbildungen mit Matrizen-Multiplikation zu identifizieren.

Für $h \to 0$ geht auch $k := f(p+h) - f(p)$ gegen Null.

$$\begin{aligned} g(f(p+h)) &= g(f(p)) + g'(f(p))\big(\overbrace{f'(p)\cdot h + \|h\|\epsilon(h)}^{=k}\big) + \|k\|\delta(k) \\ &= g(q) + g'(q)\cdot f'(p)\cdot h + \|h\|\Big(g'(q)\epsilon(h) + \frac{\|k\|\delta(k)}{\|h\|}\Big) \end{aligned}$$

Die Funktion $h \mapsto \gamma(h) := g'(q)\cdot\epsilon(h)$ ist die Verkettung der in 0 stetigen Funktion ϵ mit der linearen (und somit stetigen) Funktion $k \mapsto g'(q)\cdot k$. Also ist γ stetig in 0 und $\gamma(0) = 0$.

Mit $h \to 0$ gilt auch $\delta(k) \to 0$. Wir zeigen, dass es eine Konstante $C > 0$ gibt, so dass $\|k\|/\|h\| \le C$ für alle h in einer Umgebung von 0 gilt. Sobald dies gezeigt ist, folgt

$$\lim_{h\to 0} \frac{\|k\|\delta(k)}{\|h\|} = 0,$$

wobei man k als Funktion in h betrachten sollte. Und hieraus ergibt sich dann die Aussage des Satzes (Kettenregel).

Es reicht,$\|k\|/\|h\| \le C$ für die Supremums-Norm auf $\mathbb{R}^n$, $\mathbb{R}^m$ und $\mathbb{R}^{m\times n}$ zu zeigen, da auf jedem dieser Vektorräume alle Normen äquivalent sind.

Für $A \in \mathbb{R}^{m\times n}$ und $h \in \mathbb{R}$ gilt

$$\|Ah\|_\infty \le n\|A\|_\infty\|h\|_\infty$$

Also

$$\begin{aligned} \frac{\|k\|_\infty}{\|h\|_\infty} &= \frac{\|f'(p)\cdot h + \|h\|_\infty\epsilon(h)\|_\infty}{\|h\|_\infty} \\ &\le \frac{n\|f'(p)\|_\infty\,\|h\|_\infty + \|h\|_\infty\,\|\epsilon(h)\|_\infty}{\|h\|_\infty} \\ &\le n\|f'(p)\|_\infty + \|\epsilon(h)\|_\infty \end{aligned}$$

Da $\epsilon(h) \to 0$ für $h \to 0$ gilt, ist $\|\epsilon(h)\|_\infty$ auf einer ausreichend kleinen Umgebung von 0 beschränkt. □

KOROLLAR 2.20 (Produktregel)**.** Sei U offen in $\mathbb{R}^n$, $f, g : U \to \mathbb{R}$. Sind f und g differenzierbar in $p \in U$, dann auch $f \cdot g : U \to \mathbb{R}$, $x \mapsto f(x) \cdot g(x)$, und es gilt

$$(f \cdot g)'(p) = g(p)f'(p) + f(p)g'(p)$$

Beweis. Betrachte

$$F : U \to \mathbb{R}^2, \qquad F(x) := \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}.$$

Nach Voraussetzung ist F differenzierbar in p.

$$F'(p) = \begin{pmatrix} f'(p) \\ g'(p) \end{pmatrix} \in \mathbb{R}^{2\times n}.$$

Die Abbildung $\varphi : \mathbb{R}^2 \to \mathbb{R}$, $(x_1, x_2)^T \mapsto \varphi(x) = x_1x_2$ ist (überall) differenzierbar und es gilt $\varphi'(x_1, x_2) = (x_2, x_1) \in \mathbb{R}^{1\times 2}$, siehe Beispiel 2.5 oder berechne es über die partiellen Ableitungen.

$$f \cdot g = \varphi \circ F$$

Also ist nach Kettenregel $f \cdot g$ in p differenzierbar und es gilt

$$(f \cdot g)'(p) = \varphi'(F(p)) \cdot F'(p) = (g(p), f(p)) \cdot \begin{pmatrix} f'(p) \\ g'(p) \end{pmatrix} = g(p)f'(p) + f(p)g'(p).$$

Fr 31.5. □

Übersicht: Wichtige Beispiele und Methoden, um Differenzierbarkeit zu zeigen.

- Sei $f : \mathbb{R}^n \to \mathbb{R}^k$ von der Form $f(x) = Ax + b$, $A \in \mathbb{R}^{k\times n}$, $b \in \mathbb{R}^k$. Dann ist f überall differenzierbar und es gilt $f'(x) = A$ überall, denn
$$\frac{f(x+h) - f(x) - Ah}{\|h\|} = \frac{0}{\|h\|}.$$
Also ist $f' : \mathbb{R}^n \to \mathbb{R}^{k\times n}$ konstant und deswegen $f''(x) = 0 \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k))$.
- Ist $f : U \to \mathbb{R}^k$ (k-mal) differenzierbar und $V \subset U$ offen, dann ist auch $f|_V$ (k-mal) differenzierbar. Dasselbe für k-mal stetig differenzierbar.
- Ist $f : U \to \mathbb{R}^k$ ist auf U partiell differenzierbar und die partiellen Ableitungen sind stetig, dann ist $f : U \to \mathbb{R}^k$ stetig differenzierbar[8] (Proposition 2.15).
- Summen-, Produkt-, Quotienten-, Kettenregel.
- Polynome in mehreren Variablen sind glatt (= beliebig oft differenzierbar), z.B.
$$\mathbb{R}^3 \to \mathbb{R}, (x, y, z)^T \mapsto x^3 + x^2y^4 + 3xyz.$$
- Ebenso Funktionen wie
$$f : \mathbb{R}^n \to \mathbb{R}, \quad f(x) = e^{i\langle \xi, x\rangle}, \qquad \xi \in \mathbb{R}^n;$$
$$f : \mathbb{R}^n \to \mathbb{R}, \quad f(x) = \|x\|_2^2 = \sum_{j=1}^n x_j^2;$$
$$f : \mathbb{R}^n \to \mathbb{R}, \quad f(x) = x_1 \cdot \arctan\big(1 + \|x\|_2^2 - \cos\|x\|_2\big).$$
(Es wurde insbesondere verwendet, dass $g : \mathbb{R} \to \mathbb{R}$, $g(s) := \sum_{j=0}^\infty \frac{(-1)^j}{(2j)!} s^j$ glatt ist und $g(t^2) = \cos(t)$ für $t \in \mathbb{R}$ erfüllt. Obige Aussage sind richtig für die Standard-Norm $\|\cdot\|_2$, aber nicht für alle Normen: die Funktionen $\mathbb{R}^n \ni x \mapsto \|x\|_\infty^2$ und $\mathbb{R}^n \ni x \mapsto \|x\|_1^2$ sind zum Beispiel nicht differenzierbar.)

ENDE DER ÜBERSICHT

[8] Also insbesondere differenzierbar!

KOROLLAR 2.21 (Quotientenregel)**.** Sei U offen in $\mathbb{R}^n$, $f : U \to \mathbb{R}$, $g : U \to \mathbb{R} \setminus \{0\}$. Sind f und g differenzierbar in $p \in U$, dann auch $f/g : U \to \mathbb{R}$, $x \mapsto f(x)/g(x)$, und es gilt

$$\left(\frac{f}{g}\right)'(p) = \frac{g(p)f'(p) - f(p)g'(p)}{g(p)^2}.$$

Beweis. Ähnlich wie die Produktregel, aber mit $\varphi : \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R}$, $(y, z) \mapsto y/z$.

$$\varphi'(y, z) = (1/z, -y/z^2)$$

$$\left(\frac{f}{g}\right)'(p) = \varphi'(F(p)) \cdot F'(p) = (1/g(p), -f(p)/g(p)^2) \begin{pmatrix} f'(p) \\ g'(p) \end{pmatrix} = \frac{g(p)f'(p) - f(p)g'(p)}{g(p)^2}.$$

□

BEMERKUNGEN 2.22.

(1) Produktregel, Quotientenregel, Summenregel richtig für Richtungsableitungen, insbesondere für partielle Ableitungen

(2) Die Kettenregel **gilt nicht mehr** für alle partiell differenzierbare Funktionen. Ein Beispiel wird in der Zentralübung diskutiert. Man[9] betrachte die differenzierbare Funktion

$$h : \mathbb{R}^2 \to \mathbb{R}^2, \qquad h(x, y) = \begin{pmatrix} x + y \\ x - y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}, \qquad h'(x, y) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

und die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ aus Beispiel 2.14. Dann ist f in 0 partiell differenzierbar und es gilt

$$\frac{\partial f}{\partial x}(0) = \frac{\partial f}{\partial y}(0) = 0$$

[9]In Beispielen wie diesem benutzen wir ab jetzt die Notation $h(x, y)$ als alternative, kürzere Schreibweise für $h\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = h\big((x \quad y)^T\big)$

Aber

$$(f \circ h)(x,y) = \frac{2(x+y)(x-y)}{(x+y)^2+(x-y)^2} = \frac{(x^2-y^2)}{x^2+y^2},$$

also

$$\frac{\partial(f\circ h)}{\partial x}(0) = 1, \qquad \frac{\partial(f\circ h)}{\partial x}(0) = -1.$$

$$(0,0)\cdot\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} = (0,0) \neq (1,-1)$$

Wenn die Kettenregel gültig wäre, so würde sie besagen, dass die linke und die rechte Seite gleich sind. Also ist dies ein Gegenbeispiel.

Mittelwertsatz.

Der Mittelwertsatz für Funktionen $\mathbb{R}^n \supset U \to \mathbb{R}$ ergibt sich direkt aus dem der Analysis I.

PROPOSITION 2.23 (Mittelwertsatz)**.** Sei U offen in $\mathbb{R}^n$, $p, q \in U$ mit

$$[p,q] := \{p + t(q-p) \mid t \in [0,1]\} \subset U.$$

Die Funktion $f : U \to \mathbb{R}$ sei stetig auf $[p,q]$ und differenzierbar auf

$$\{p + t(q-p) \mid t \in (0,1)\}.$$

Dann gibt es ein $\vartheta \in (0,1)$ mit

$$f(q) = f(p) + f'(p + \vartheta(q-p)) \cdot (q-p).$$

Beweis. $\varphi : [0,1] \to \mathbb{R}^n$, $t \mapsto p + t(q-p)$.

$f \circ \varphi$ stetig auf $[0,1]$ und differenzierbar auf $(0,1)$. Aus dem ersten Mittelwertsatz (Satz 3.2 aus Kapitel 5) folgt die Existenz eines $\vartheta \in (0,1)$ mit

$$f(q)-f(p) = (f\circ\varphi)(1)-(f\circ\varphi)(0) \overset{\text{1. MWS}}{=} (f\circ\varphi)'(\vartheta) = f'(\varphi(\vartheta))\cdot\varphi'(\vartheta) = f'(p+\vartheta(q-p))\cdot(q-p).$$

□

!ACHTUNG!. Man kann hier nicht ohne weiteres die Zielmenge $\mathbb{R}$ durch $\mathbb{R}^k$ ersetzen. Wenn wir $f = (f_1, \ldots, f_k)^T$ schreiben, kann man zwar den Mittelwert-Satz für jedes f_i anwenden, und wir erhalten jedesmal ein $\vartheta_i \in [0,1]$ mit

$$f_i(q) = f_i(p) + f_i'(p + \vartheta_i(q-p)) \cdot (q-p).$$

Man findet aber kein $\vartheta \in [0,1]$ mit

$$f(q) = f(p) + f'(p + \vartheta(q-p)) \cdot (q-p).$$

Beispiel 2.24. $f : \mathbb{R} \to \mathbb{R}^2$, $f(t) = (\cos t, \sin t)^T$, $p = 0$, $q = 2\pi$. Dann gilt $f(p) = f(q)$, aber $f'(t) = (-\sin t, \cos t)^T \neq 0 \in \mathbb{R}^2$. Es kann also kein solches ϑ geben.

Definition 2.25. Sei X ein topologischer Raum und Y eine Menge. Eine Funktion $f : X \to Y$ heißt lokalkonstant, wenn jedes $x \in X$ eine Umgebung U in X besitzt, so dass $f|_U$ konstant ist.

Konstante Funktionen sind lokal konstant. Ist X diskret (zum Beispiel $X = \{1,2,3\} \subset \mathbb{R}$), dann ist jede Funktion $f : X \to Y$ lokal konstant.

Ist $f : X \to Y$ lokal konstant und $A \subset X$, dann ist $f|_A$ lokal konstant.

LEMMA 2.26. Ist $f : A \to Y$ lokal konstant und ist A zusammenhängend, dann ist f konstant.

Beweis. Sei $a \in A$ gegeben. Dann folgt aus der lokalen Konstantheit, dass $U_1 := \{x \in A \mid f(x) = f(a)\}$ offen in A ist und dass $U_2 := \{x \in A \mid f(x) \neq f(a)\}$ offen ist. Aus $A = U_1 \dot{\cup} U_2$, $U_1 \neq \emptyset$ und A zusammenhängend folgt $U_2 = \emptyset$. □

Aus dem Mittelwertsatz ergibt sich das folgende Korollar:

KOROLLAR 2.27. Sei $f : U \to \mathbb{R}$ wie oben. Dann gilt

$$f'(x) = 0 \text{ für alle } x \in U \quad \Longleftrightarrow \quad f \text{ ist lokal konstant.}$$

Beweis.

„$\Longleftarrow$“ ist klar.

„$\Longrightarrow$“: Zu $p \in U$ wähle $\epsilon > 0$ mit $B_\epsilon(p) \subset U$. Für jedes $q \in B_\epsilon(p)$ gibt es ein $\vartheta \in [0, 1]$ mit

$$f(q) = f(p) + f'(p + \vartheta(q - p)) \cdot (q - p) = f(p) + 0.$$

Also ist $f|_{B_\epsilon(p)}$ konstant. □

DEFINITION 2.28. Eine Teilmenge $G \subset \mathbb{R}^n$ nennt man Gebiet, wenn G nicht-leer, offen und zusammenhängend ist.

KOROLLAR 2.29. Sei $f : U \to \mathbb{R}$ wie oben, und U ein Gebiet.

$$f'(x) = 0 \text{ für alle } x \in U \quad \Longleftrightarrow \quad f \text{ ist konstant.}$$

3. Höhere Ableitungen

3.1. Satz von Schwarz. Sei U offen in $\mathbb{R}^n$, $f : U \to \mathbb{R}^k$.

Die i-te partielle Ableitung ist

$$\frac{\partial f}{\partial x_i} : U \to \mathbb{R}^k.$$

Man nennt es auch die partielle Ableitung nach x_i.

Wenn wir diese Ableitung nun nach x_j partiell ableiten, erhalten wir

$$\frac{\partial^2 f}{\partial x_j \partial x_i} := \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) : U \to \mathbb{R}^k$$

Beispiel: $f(x_1, x_2) = x_1^2 x_2$, $\frac{\partial f}{\partial x_1}(x_1, x_2) = 2x_1 x_2$, $\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) = 2x_1$.

Zu $i_1, i_2, \ldots, i_\ell \in \{1, 2, \ldots, n\}$ definieren wir rekursiv:

$$\frac{\partial^\ell f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_\ell}} := \frac{\partial}{\partial x_{i_1}}\left(\frac{\partial^{\ell-1} f}{\partial x_{i_2} \cdots \partial x_{i_\ell}}\right) : U \to \mathbb{R}^k.$$

Die Anzahl der partiellen Ableitungen von der Ordnung ℓ ist n^ℓ. Sind alle diese Ableitungen verschieden?

Wir betrachten nun den Spezialfall $k = 1$ und $\ell = 2$.

Definition 3.1. Wenn $f : U \to \mathbb{R}$ auf U differenzierbar ist und $(f')^T : U \to \mathbb{R}^n = \mathbb{R}^{n\times 1}$ in $p \in U$ differenzierbar ist, dann nennt

$$\left(\frac{\partial^2 f}{\partial x_j \partial x_i}(p)\right)_{i,j} = \big((f')^T\big)'(p)$$

die Hesse-Matrix von f in p.

Die zugehörige bilineare Abbildung ist die Hesse-Abbildung

$$\mathrm{Hess}_p f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) \mapsto \sum_{i,j=1}^{n} x_i \frac{\partial^2 f}{\partial x_j \partial x_i}(p)\, y_j$$

Die Hesse-Matrix ist die Matrix-Darstellung der Hesse-Abbildung.

Schreibweise für Hesse-Matrix: $f''(p) := \big((f')^T\big)'(p) \in \mathbb{R}^{n\times n}$.

Wir zeigen nun folgenden Satz:

Satz 3.2 (Satz von Schwarz)**.** Sei $f : U \to \mathbb{R}^k$, U offen in $\mathbb{R}^n$, $p \in U$, $\ell \geq 2$. Ist f $(\ell-1)$-mal differenzierbar auf U und ℓ-mal differenzierbar im Punkt p, dann sind die ℓ-ten partiellen Ableitungen in p unabhängig von der Reihenfolge der Ableitungen.

10

Zum Beispiel: Ist f zweimal differenzierbar auf U, dann gilt

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x_1, \ldots, x_n) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_1, \ldots, x_n)$$

das heißt $f''(p)$ ist eine symmetrische Matrix.

Dies bedeutet, wir brauchen danach gar nicht mehr darauf zu achten, welcher Index von $\frac{\partial^2 f}{\partial x_i \partial x_j}$ die Zeile und welcher die Spalte angibt.

Im Beweis dieses Satzes von Schwarz ist es aber natürlich sehr wichtig, welche Rolle Zeilen und Spalten haben: $f'(p) \in \mathbb{R}^{1\times n}$ ist ein Zeilen-Vektor, $f'(p)^T \in \mathbb{R}^{n\times 1}$ ist ein

[10] Nach Hermann Amandus Schwarz (1843-1921) siehe `http://de.wikipedia.org/wiki/Hermann_Amandus_Schwarz`

Spalten-Vektor, und die Hesse-Matrix ist

$$\big((f')^T\big)'(p) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(p)\right)_{i,j}.$$

Mi 5.6.

Beweis. O.B.d.A. $\ell = 2$.[11] O.B.d.A. $k = 1$ (komponentenweise Argumentation). O.B.d.A. $\|\cdot\| = \|\cdot\|_2$. Sei $\epsilon > 0$ gegeben und zunächst fixiert. Wähle $\rho > 0$ so, dass

$$B_{3\rho}(p) = \{x \in \mathbb{R}^n \mid \|x - p\|_2 < 3\rho\} \subset U.$$

Wähle $u, v \in \mathbb{R}^n$ mit $0 < \|u\|_2 = \|v\|_2 \leq \rho$. Definiere

$$g : [0,1] \to \mathbb{R}, \qquad g(t) := f(p + tu + v) - f(p + tu).$$

Wir rechnen für eine reelle Zahl $\vartheta \in (0,1)$. Sie darf zunächst beliebig sein, und später mit dem 1. Mittelwertsatz so gewählt, dass $g(1) - g(0) = g'(\vartheta)$.

$$\begin{aligned} g'(\vartheta) &= \Big(f'(p + \vartheta u + v) - f'(p + \vartheta u)\Big)u \\ &= \Big(f'(p + \vartheta u + v) - f'(p) - (f'(p + \vartheta u) - f'(p))\Big)u \end{aligned}$$

Da $f''(p)$ existiert, wissen wir, dass

$$\| \underbrace{\big(f'\big)^T(p + \vartheta u + v) - \big(f'\big)^T(p) - f''(p) \cdot (\vartheta u + v)}_{a:=} \| \leq \epsilon \, \|\vartheta u + v\| \leq \epsilon \, (\|u\| + \|v\|),$$

[11] Wenn wir zwei Ableitungen vertauschen können, dann können wir alle Ableitungen miteinander sukzessive vertauschen.

gilt, falls $\rho > 0$ genügend klein gewählt wurde. [12] Und analog erhalten wir für genügend kleines $\rho > 0$:

$$\| \underbrace{\big(f'\big)^T(p+\vartheta u) - \big(f'\big)^T(p) - f''(p)\cdot(\vartheta u)}_{b:=} \| \leq \epsilon \, \|\vartheta u\| \leq \epsilon \, (\|u\| + \|v\|).$$

Dies ergibt[13]

$$\begin{aligned} g'(\vartheta) &= \big(g'(\vartheta)\big)^T \\ &= u^T \Big(\big(f'\big)^T(p+\vartheta u + v) - \big(f'\big)^T(p) - \big((f')^T(p+\vartheta u) - (f')^T(p)\big)\Big) \\ &= u^T \big(a - b + f''(p)\cdot v\big) \end{aligned}$$

Wir erhalten unter Benutzung der Cauchy-Schwarz-Ungleichung[14]

$$\forall x, y \in \mathbb{R}^n : \; x^T y = \langle x, y\rangle \leq \|x\|_2 \|y\|_2$$

$$|g'(\vartheta) - u^T f''(p) v| = |u^T(a-b)| \leq \|u\| \cdot 2\epsilon \, (\|u\| + \|v\|) \leq 2\epsilon \, (\|u\| + \|v\|)^2.$$

Mit dem 1. Mittelwertsatz können wir nun $\vartheta \in (0,1)$ so wählen, dass

$$g'(\vartheta) \overset{\text{1. MWS}}{=} g(1) - g(0) = f(p+u+v) - f(p+u) - f(p+v) + f(p).$$

[12] Genaue Logik: zu obigem $\epsilon > 0$ gibt es ein $\rho_1 \in (0, \rho]$, so dass die Aussage für $0 < \|u\|_2 = \|v\|_2 \leq \rho_1$ gilt. Im nächsten Schritt erhalten wir dann ein $\rho_2 \in (0, \rho_1]$ mit analogen Eigenschaften.

[13] Identifiziere $\mathbb{R} \cong \mathbb{R}^1 \cong \mathbb{R}^{1\times 1}$

[14] Dies ist wieder Hermann Amandus Schwarz, allerdings ist es fraglich, ob hier der Name Schwarz so angebracht ist. Schwarz wurde 1843 geboren, obwohl der Satz von Cauchy-Schwarz in $\mathbb{R}^n$ bereits 1821 von Cauchy veröffentlicht wurde. Die Integral-Version stammt von Buniakovsky (1859), siehe `http://jeff560.tripod.com/c.html`. Dennoch sollte man die Ungleichung die Cauchy-Schwarz-Ungleichung nennen, denn nur mit diesem seit ca. 1930 üblichen Namen weiß man, welche Ungleichung gemeint ist.

Somit

$$|f(p+u+v)-f(p+u)-f(p+v)+f(p)-u^T f''(p)v| = |u^T(a-b)| \leq 2\epsilon\,(\|u\|+\|v\|)^2.$$

Der Ausdruck $f(p+u+v)-f(p+u)-f(p+v)+f(p)$ ändert sich nicht, wenn wir u und v vertauschen.

Also

$$|u^T f''(p)v - v^T f''(p)u| \leq 4\epsilon(\|u\|+\|v\|)^2.$$

Wir setzen $u := \lambda e_i$, $v = \lambda e_j$, $\lambda > 0$ genügend klein.

$$\left|\lambda^2 \frac{\partial^2 f}{\partial x_j \partial x_i}(p) - \lambda^2 \frac{\partial^2 f}{\partial x_i \partial x_j}(p)\right| \leq 4\epsilon\lambda^2(1+1)^2.$$

$$\left|\frac{\partial^2 f}{\partial x_j \partial x_i}(p) - \frac{\partial^2 f}{\partial x_i \partial x_j}(p)\right| \leq 16\epsilon.$$

Da wir $\epsilon > 0$ zu Beginn des Beweises beliebig klein[15] wählen konnten[16], ergibt sich

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(p) = \frac{\partial^2 f}{\partial x_i \partial x_j}(p).$$

□[17]

[15]Hier bedeutet „beliebig klein“ „beliebig nahe an 0“

[16]und dann eben auch ρ, λ etc klein, aber positiv wählen

[17]Kommentar zu den Quellen: verschiedene Beweise in der Literatur benutzen etwas andere Voraussetzungen. Die obige Version kenne ich von einem Vorlesungsskript Prof. Helmut Klingen, Freiburg, WS 1988/89. In Königsberger [**23**] benötigt man andere Voraussetzungen an Stelle der ℓ-fachen Differenzierbarkeit in p, nämlich (im Fall $\ell = 2$), dass eine der $\frac{\partial^2 f}{\partial x_j \partial x_i}$ für ein festes i und j stetig in p sein soll. Es folgt dann ebenfalls die obige Gleichheit. Die Königsbergerschen Voraussetzungen sind weder schwächer noch stärker als unsere. Viele Bücher nehmen 2-mal stetig differenzierbar an, und erhalten dann eine Aussage, die echt schwächer als unsere obige und als die Königsbergersche ist.

BEISPIEL 3.3 (Zentralübung). Sei $f : \mathbb{R}^2 \to \mathbb{R}$,

$$(x,y) \mapsto \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & \text{falls } (x,y) \neq (0,0), \\ 0 & \text{falls } (x,y) = (0,0). \end{cases}$$

Man rechnet nach, dass f zweimal partiell differenzierbar ist. Aber es gilt

$$\frac{\partial^2 f}{\partial y\, \partial x}(0,0) = -1, \qquad \frac{\partial^2 f}{\partial x\, \partial y}(0,0) = 1$$

Schreibweise

Oft verwendet man x, y, z an Stelle von x_1, x_2, x_3. Dann

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}.$$

18

[18]Bei physikalischen Anwendungen ist oft die Reihenfolge der Variablen unklar, und bestimmte Buchstaben bedeuten immer bestimmte Größen. Eine gewisse Menge Gas soll bei einer Temperatur T und einem Volumen V den Druck p haben. Dann ist p eine Funktion in T und V, also $p : \mathbb{R}^2_{>0} \to \mathbb{R}_{>0}$, $(T,V) \mapsto p(T,V)$ oder $(V,T) \mapsto p(V,T)$. Welches von beiden, weiß man eben nicht so recht, und ist aber irrelevant. Jedenfalls ist in beiden Fällen klar, was mit $\partial p/\partial V$ gemeint ist. Hier ist die eingeführte Notation sehr effektiv und ungefährlich. Gefährlich wird es aber in der Physik, wenn wir nun den Blickwinkel wechseln und sagen, der Druck des Gases hängt von dem Volumen V und der im Gas enthaltenen Enthalpie H ab, $(V,H) \mapsto p(V,H)$. Dann erhält $\partial p/\partial V$ eine andere Bedeutung, denn nun ist damit die infinitesimale Änderung des Drucks bei Volumenänderung und konstanter Enthalpie gemeint, wohingegen zuvor damit die infinitesimale Änderung des Drucks bei Volumenänderung und konstanter Temperatur gemeint war. Dieser Sachverhalt hat mich in der Thermodynamik damals als Student etwas verwirrt.

3.2. Höhere Ableitungen als multi-lineare Abbildungen. In diesem Abschnitt sei wieder U offen in $\mathbb{R}^n$. Sei V ein reeller Vektorraum. Alles was wir bisher für Funktionen $g : U \to \mathbb{R}^k$ mit beliebigem $k \in \mathbb{N}$ definiert und gezeigt haben, können wir auch – mit leichten redaktionellen Modifikationen – für Funktionen $g : U \to V$ definieren und zeigen[19]. Ist $g : U \to V$ differenzierbar, so erhalten wir die Ableitung $g' : U \to \mathcal{L}(\mathbb{R}^n, V)$. Ein Vorteil dieser Sichtweise ist dann, dass die Differenzierbarkeit und die Ableitung von g unabhängig von der verwendeten Basis von V definiert sind.

Diese Sichtweise erlaubt uns auch, eine (mindestens) ℓ-mal differenzierbare Funktion $f : U \to \mathbb{R}$ ℓ-mal zu differenzieren. Im Fall $\ell = 2$ ist $f' : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \cong \mathbb{R}^{n\times 1}$. Wir wenden dies auf $g = f'$ und $V = \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ an und erhalten $g' = f'' : U \to \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$. Die Transposition oder äquivalent dazu die Matrix-Darstellung in der Standard-Basis liefert wie in der Linearen Algebra üblich einen Isomorphismus $\mathcal{L}(\mathbb{R}^n, \mathbb{R}) \xrightarrow{\cong} \mathbb{R}^n$, $v \mapsto v^T$. Dies ergibt

$$\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})) \cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \cong \mathbb{R}^{n\times n} \cong \mathbb{R}^{n^2}.$$

Wenn wir dies iterativ für höhere Ableitungen durchführen, dann ist es ab der dritten Ableitung nicht mehr möglich und sinnvoll, den Formalismus der Linearen Algebra, der streng zwischen Zeilen- und Spaltenvektoren unterscheidet, einzuhalten. Die ℓ-te Ableitung hat ℓ Indizes und für das übliche Matrizen-Kalkül bräuchten wir dann ein ℓ-dimensionales Blatt Papier um die Matrix zu notieren. Eine mathematisch konzeptionelle Sichtweise zur Beschreibung dieser multilinearen Abbildungen ist mit Hilfe von Tensoren (Lineare Algebra II), ein Begriff, den wir hier nur andeuten wollen. In konkreten Berechnungen arbeitet man oft am einfachsten mit iterierten partiellen Ableitungen.

[19]Eine Ausnahme ist hier natürlich der Mittelwertsatz, da das Ziel 1-dimensional sein muss.

Sei $p \in U$.

$f(p) \in \mathbb{R}$, $(0,0)$-Tensor=Skalar

$f'(p) = \left(\dfrac{\partial f}{\partial x_i}(p) \right)_{i=1,\dots,n} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \cong \mathbb{R}^{1\times n} \cong \mathbb{R}^n$, $(0,1)$-Tensor=(Zeilen-)Vektor

$f' : U \to \mathbb{R}^n$

$f''(p) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$

$\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})) \cong \{\text{bilineare Abb. } \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}\} \cong \mathbb{R}^{n\times n} \cong \mathbb{R}^{n^2}$, $(0,2)$-Tensor=Matrix

$f'' : U \to \mathbb{R}^{n\times n}$

$f''(p) = \left(\dfrac{\partial^2 f}{\partial x_i \, \partial x_j}(p) \right)_{i,j\in\{1,\dots,n\}}$

$$f'''(p) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})))$$

$$\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))) \cong \{\text{trilineare Abb. } \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}\} \cong \mathbb{R}^{n\times n\times n} \ (0,3)\text{-Tensor}$$

$$f''' : U \to \mathbb{R}^{n\times n\times n}$$

$$f'''(p) = \left(\frac{\partial^3 f}{\partial x_i \, \partial x_j \, \partial x_k}(p)\right)_{i,j,k\in\{1,\dots,n\}}$$

$$\vdots$$

$$f^{(\ell)}(p) : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{\ell\text{-mal}} \to \mathbb{R}, \ \ell\text{-linear}, (0,\ell)\text{-Tensor}$$

$$f^{(\ell)} := \left(f^{(\ell-1)}\right)' \quad \ell\text{-linear}$$

$$f^{(\ell)} = f : U \to \{\ell\text{-lineare Abb. } \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{\ell\text{-mal}} \to \mathbb{R}\}$$

$$f^{(\ell)}(p) = \left(\frac{\partial^\ell f}{\partial x_{i_1} \, \partial x_{i_2} \cdots \partial x_{i_\ell}}\right)_{i_1,\dots,i_\ell\in\{1,\dots,n\}} (p)$$

Der Satz von Schwarz besagt, dass all diese multi-linearen Abbildung in allen Einträgen symmetrisch sind.[20] Wir haben dies bereits benutzt und die Reihenfolge der Indizes getauscht.

Beispiel 3.4. Sei U offen in $\mathbb{R}^n$, seien $p = (p_1, \dots, p_n)^T \in U$, $q = (q_1, \dots, q_n)^T \in \mathbb{R}^n$, $f : U \to \mathbb{R}$ sei ℓ-mal differenzierbar. Es gilt $c(p + tq) \in U$ für $t \in (-\epsilon, \epsilon)$ für ein kleines $\epsilon > 0$. Sei $g(t) := f(c(t))$ für $t \in (-\epsilon, \epsilon)$, d.h. $g = f_{p,q}$ in der Notation von Definition 2.10.

[20] Das heißt, man darf die Einträge in der Reihenfolge tauschen, also zum Beispiel $f^{(k)}(p)(h_1, h_2, \dots, h_k) = f^{(k)}(p)f(h_2, h_1, h_3, \dots, h_k) = f^{(k)}(p)f(h_k, h_{k-1}, \dots, h_1)$

$$c(0) = p, \quad c'(t) = q, \quad c^{(\ell)}(t) = 0 \text{ für } \ell \geq 2.$$

$$\begin{aligned}
g'(t) &= f'(c(t))c'(t) = \sum_{i_1=1}^{n} \frac{\partial f}{\partial x_{i_1}}(c(t))\, q_{i_1} \\
g''(t) &= c'(t)^T f''(c(t))c'(t) + f'(c(t))\underbrace{c''(t)}_{=0} = \mathrm{Hess}_{c(t)} f\big(c'(t), c'(t)\big) = \sum_{i_1,i_2=1}^{n} \frac{\partial^2 f}{\partial x_{i_1}\partial x_{i_2}}(c(t))\, q_{i_1} q_{i_2} \\
g'''(t) &= \Big(\text{trilineare Abb. zu } f'''(c(t))\Big)\big(c'(t), c'(t), c'(t)\big) = \sum_{i_1,i_2,i_3=1}^{n} \frac{\partial^3 f}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}}(c(t))\, q_{i_1} q_{i_2} q_{i_3} \\
&\vdots \\
g^{(\ell)}(t) &= \Big(\ell\text{-lineare Abb. zu } f^{(\ell)}(c(t))\Big)\big(\underbrace{c'(t), \ldots, c'(t)}_{\ell\text{-mal}}\big) = \sum_{i_1,\ldots,i_\ell=1}^{n} \frac{\partial^\ell f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_\ell}}(c(t))\, q_{i_1}\cdots q_{i_\ell}
\end{aligned}$$

3.3. Satz von Taylor.

Satz 3.5 (Taylor)**.** Sei U offen in $\mathbb{R}^n$ und $f : U \to \mathbb{R}$. Seien $p = (p_1, \ldots, p_n)^T, x = (x_1, \ldots, x_n)^T \in U$ mit $[p, x] \subset U$. Die Funktion f sei ℓ-mal differenzierbar auf U und $(\ell + 1)$-mal differenzierbar auf

$$\{p + t(x - p) \mid t \in (0, 1)\}.$$

Wir definieren das Restglied $R_\ell^p(x)$ so, dass gilt:

$$\begin{aligned}
f(x) &= f(p) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p) \cdot (x_i - p_i) + \frac{1}{2!} \sum_{i_1,i_2=1}^{n} \frac{\partial^2 f}{\partial x_{i_1}\, \partial x_{i_2}}(p) \cdot (x_{i_1} - p_{i_1}) \cdot (x_{i_2} - p_{i_2}) \\
&\quad + \cdots + \frac{1}{\ell!} \sum_{i_1,i_2,\ldots,i_\ell=1}^{n} \frac{\partial^\ell f}{\partial x_{i_1}\, \partial x_{i_2} \cdots \partial x_{i_\ell}}(p) \cdot (x_{i_1} - p_{i_1}) \cdot (x_{i_2} - p_{i_2}) \cdots (x_{i_\ell} - p_{i_\ell}) \\
&\quad + R_\ell^p(x)
\end{aligned}$$

Dann gibt es ein $\vartheta \in (0,1)$ mit

$$R_\ell^p(x)$$

$$= \frac{1}{(\ell+1)!} \sum_{i_1,i_2,\ldots,i_{\ell+1}=1}^{n} \left(\frac{\partial^{\ell+1} f}{\partial x_{i_1}\, \partial x_{i_2} \cdots \partial x_{i_{\ell+1}}} (p + \vartheta(x-p)) \right) \cdot (x_{i_1} - p_{i_1}) \cdots (x_{i_{\ell+1}} - p_{i_{\ell+1}})$$

Die blauen Klammern geben exemplarisch an, wie man derartige Ausdrücke verstehen muss, diese Klammern werden aber normalerweise nicht gesetzt.

Beweis. Anwendung des Satzes von Taylor aus Analysis I auf die Funktion

$$g(t) := f(p + t(x-p))$$

ergibt:

$$g(t) = \sum_{j=0}^{\ell} \frac{t^j}{j!} g^{(j)}(0) + r_\ell^0(t) \tag{3.6}$$

mit

$$r_\ell^0(t) = \frac{t^{\ell+1}}{(\ell+1)!} g^{(\ell+1)}(\vartheta t)$$

für ein $\vartheta \in (0,1)$. Nun setzen wir $t = 1$ in (3.6), wenden Beispiel 3.4 für obiges p und $q = x - p$ an und erhalten mit $R_\ell^p(x) := r_\ell^0(1)$ die Behauptung. □

NOTATION 3.7 (Multi-Indizes). Sei $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$, $p \in U$, $f : U \to \mathbb{R}$. Dann definieren wir

$$\alpha! := \alpha_1!\,\alpha_2! \cdots \alpha_n!$$

$$|\alpha| := \alpha_1 + \cdots + \alpha_n$$

$$h^\alpha := h_1^{\alpha_1} \cdots h_n^{\alpha_n}$$

$$D^\alpha f|_p := \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(p) := \frac{\partial^{\alpha_1+\cdots+\alpha_n} f}{\underbrace{\partial x_1 \cdots \partial x_1}_{\alpha_1\text{-mal}} \cdots \underbrace{\partial x_n \cdots \partial x_n}_{\alpha_n\text{-mal}}}(p)$$

Die Bezeichnung $\frac{\partial^{|\alpha|} f}{\partial x^\alpha}(p)$ wird öfter verwendet, u.a. in der Physik. Die Bezeichnung $D^\alpha f|_p$ erscheint mir zweckmäßiger und effektiver.

Fr. 7.6.

Wenn wir nun den Satz von Schwarz benutzen, gleiche Terme in der obigen Taylorschen Formel zusammenfassen, und elementare Formeln der Kombinatorik nutzen[21] (siehe Zentralübung am 18.6.), dann erhalten wir den folgenden Satz.

FOLGERUNG 3.8 (Satz von Taylor in Multi-Index-Notation)**.** Sei U, p, x, f wie oben.

Wir definieren das Restglied $R_\ell^p(x)$ so, dass gilt:

$$f(x) \;=\; \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \le \ell}} \frac{1}{\alpha!} D^\alpha f|_p (x-p)^\alpha + R_\ell^p(x)\,.$$

Dann gibt es ein $\vartheta \in (0,1)$ mit

$$R_\ell^p(x) \;=\; \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = \ell+1}} \frac{1}{\alpha!} D^\alpha f|_{p+\vartheta(x-p)} (x-p)^\alpha\,.$$

[21] um abzuzählen bzw. zu berechnen, wie oft jeder Term vorkommt

3.4. Einschub: Quadratische Formen. Um den Term 2. Ordnung in der Taylor-Entwicklung genauer zu verstehen, betrachten wir quadratische Formen.

DEFINITION 3.9. Sei $\mathbb{K}$ ein Körper mit $2 \neq 0$, z.B. $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Sei V ein $\mathbb{K}$-Vektorraum. Eine quadratische Form ist eine Abbildung $Q : V \to \mathbb{K}$ in der Form

$$Q(x) = b(x, x)$$

für eine $\mathbb{K}$-bilineare Abbildung $b : V \times V \to \mathbb{K}$.

BEMERKUNGEN 3.10.

(1) Die Bilinearform b in obiger Definition kann symmetrisch gewählt werden. Denn für eine beliebige bilineare Abbildung $b : V \times V \to \mathbb{K}$ schreiben wir

$$b_{\pm}(x, y) := \frac{b(x, y) \pm b(y, x)}{2}.$$

Also $b(x, y) = b_+(x, y) + b_-(x, y)$. Nun ist b_+ ist symmetrisch: $b_+(x, y) = b_+(y, x)$. Und b_- ist antisymmetrisch $b_-(x, y) = -b_-(y, x)$. Wegen $b_-(x, x) = -b_-(x, x) = 0$ gilt $b(x, x) = b_+(x, x)$.

(2) Für $Q(x) = b(x, x)$ gilt

$$Q(x + y) - Q(x - y) = b(x + y, x + y) - b(x - y, x - y) = 2b(x, y) + 2b(y, x) = 4b_+(x, y).$$

Der symmetrische Anteil von b ist also durch Q eindeutig festgelegt.

DEFINITION 3.11. Sei $\mathbb{K} = \mathbb{R}$. Eine quadratische Form Q auf V heißt

$$\begin{aligned}
\text{positiv definit} &:\Longleftrightarrow \forall x \in V \setminus \{0\} : Q(x) > 0 \\
\text{negativ definit} &:\Longleftrightarrow \forall x \in V \setminus \{0\} : Q(x) < 0 \\
\text{positiv semi-definit} &:\Longleftrightarrow \forall x \in V : Q(x) \geq 0 \\
\text{negativ semi-definit} &:\Longleftrightarrow \forall x \in V : Q(x) \leq 0 \\
\text{indefinit} &:\Longleftrightarrow Q \text{ ist weder positiv semi-definit noch negativ semi-definit} \\
&\Longleftrightarrow \text{Es existieren } x_+, x_- \in V \text{ mit } Q(x_+) > 0 \text{ und } Q(x_-) < 0.
\end{aligned}$$

Nun[22] sei $V = \mathbb{R}^n$. Jede quadratische Matrix $A \in \mathbb{R}^{n\times n}$ definiert eine quadratische Form

$$Q_A(x) = x^T A x.$$

Umgekehrt kann man zu einer gegebenen quadratischen Abbildung $Q : \mathbb{R}^n \to \mathbb{R}$ definieren:

$$a_{ij} := \frac{Q(e_i + e_j) - Q(e_i - e_j)}{4}, \qquad A_Q := (a_{ij})_{ij} \in \mathbb{R}^{n\times n}$$

Wir erhalten eine Bijektion

$$\begin{aligned}
\{\text{Quadratische Formen auf } \mathbb{R}^n\} &\to \{A \in \mathbb{R}^{n\times n} \mid A^T = A\} \\
Q &\mapsto A_Q \\
Q_A &\mapsfrom A
\end{aligned}$$

[22]Man kann ähnlich vorgehen für beliebige endlich-dimensionale relle Vektorräume V, in dem man zuerst eine basis von V wählt, die uns dann erlaubt, V mit $\mathbb{R}^n$ zu identifizieren. Es stellt sich dann natürlich die Frage, wie sich die definierten Objekte transformieren, wenn wir die Basis wechseln. Dies sollte im Rahmen der Linearen Algebra diskutiert werden.

Wir nennen eine quadratische Matrix A positiv definit (bzw. positiv semi-definit, negativ (semi-) definit, indefinit), falls A symmetrisch ist und Q_A die gleichnamige Eigenschaft hat.

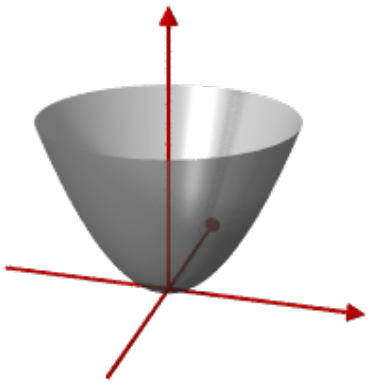
Alle symmetrische Matrizen sind diagonalisierbar (siehe Korollar 1.8.5 der Linearen Algebra II von Prof. Cisinski). Man sieht dann leicht, dass für symmetrische Matrizen gilt:

A ist positiv (bzw. negativ) definit $:\Longleftrightarrow$ alle Eigenwerte von A sind positiv (bzw. negativ)

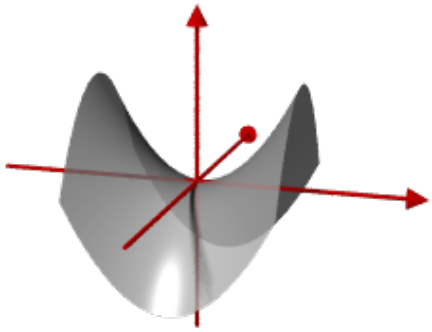
A ist positiv (bzw. negativ) semi-definit $:\Longleftrightarrow$ alle Eigenwerte sind nicht-negativ (bzw. nicht-positiv)

A ist indefinit $:\Longleftrightarrow$ A hat positive und negative Eigenwerte

Beispiele 3.12.



Elliptisches Paraboloid E



Hyperbolisches Paraboloid H

Die Bilder sind von der Webseite `http://de.wikipedia.org/wiki/Paraboloid` gezichnet von Lars Rohwedder (RokerHRO) - german wikipedia, Gemeinfrei, Paraboloid `https://commons.wikimedia.org/w/index.php?curid=213587`, Hyperbolid erzeugt mit Povray 3.6, `https://commons.wikimedia.org/w/index.php?curid=23859326`

Die Menge E der Punkte mit $z = x^2 + y^2$ ist der Graph der Funktion $f_1(x, y) = x^2 + y^2$. Die Funktion f_1 ist die quadratische Form zum euklidischen Skalarprodukt auf $\mathbb{R}^2$. Sie ist positiv definit. Die zugehörige symmetrische Matrix ist die Matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Graphen von positiv definiten quadratischen Formen nennt elliptisches Paraboloid.

Die Menge H der Punkte mit $z = x^2 - y^2$ ist der Graph der Funktion $f_2(x, y) = x^2 - y^2$. Die Funktion f_2 ist die quadratische Form zur bilinearen Abbildung $b : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $b((x, y)^T, (\tilde{x}, \tilde{y})^T) = x\tilde{x} - y\tilde{y}$. Sie ist indefinit. Die zugehörige symmetrische Matrix ist

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Graphen von indefiniten quadratischen Formen nennt hyperbolisches Paraboloid.

Wir versehen nun $\mathbb{R}^{n\times n}$ mit einer beliebigen Norm $\| \cdot \|$, und erhalten so eine Topologie auf $\mathbb{R}^{n\times n}$, die unabhängig von der Norm ist.

Lemma 3.13. Sei $A \in \mathbb{R}^{n\times n}$ eine quadratische Matrix, so dass Q_A positiv definit (bzw. negativ definit, bzw. indefinit) ist. Dann gibt es eine Umgebung U von A in $\mathbb{R}^{n\times n}$, so dass für alle $B \in U$ auch Q_B positiv definit (bzw. negativ definit, bzw. indefinit) ist.

Ersetzt man definit durch semi-definit, so ist das Lemma nicht mehr gültig!

Im indefiniten Fall zeigen wir sogar etwas mehr:

Zusatz 3.14. Gilt $Q_A(x_+) > 0$ und $Q_A(x_-) < 0$, dann gibt es eine Umgebung U von A in $\mathbb{R}^{n\times n}$, so dass für alle $B \in U$ gilt: $Q_B(x_+) > 0$ und $Q_B(x_-) < 0$.

23

Beweis. Angenommen es gebe ein $x \in \mathbb{R}^n \setminus \{0\}$ mit $x^T Ax > 0$. Indem wir x durch $x/\|x\|_2$ ersetzen, können wir annehmen, dass

$$x \in S^{n-1} := \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$$

Da die Abbildung $F : S^{n-1} \times \mathbb{R}^{n\times n} \to \mathbb{R}$, $(y, B) \mapsto y^T By$ stetig ist, gibt es zu jedem x wie oben ein $\epsilon(x) > 0$, so dass:

$$\forall y \in S^{n-1} : \forall B \in \mathbb{R}^{n\times n} : \|x - y\|_2 < \epsilon(x) \wedge \|B - A\| < \epsilon(x) \Longrightarrow F(y, B) > 0. \tag{3.15}$$

Die Zahl $\epsilon(x)$ hängt natürlich von x ab!

Eine analoge Konstruktion geht auch im Fall $Q_A(x, x) < 0$.

1. Fall: Ist nun Q_A indefinit, dann existieren $x_+, x_- \in S^{n-1}$ mit $Q_A(x_+) > 0$ und $Q_A(x_-) < 0$. Dann gilt für $\|B - A\| < \min\{\epsilon(x_+), \epsilon(x_-)\}$: $Q_B(x_+) > 0$ und $Q_B(x_-) < 0$.

2. Fall: Ist Q_A positiv definit, dann wähle zu jedem $x \in S^{n-1}$ ein $\epsilon(x) > 0$ wie in (3.15). Dann ist $(B_{\epsilon(x)}(x) \cap S^{n-1})_{x\in S^{n-1}}$ eine offene Überdeckung von S^{n-1}. Da S^{n-1} kompakt ist, kann man zu dieser Überdeckung eine Lebesguesche Zahl $\epsilon > 0$ wählen (die nun nicht mehr von x abhängt!). Dies bedeutet, dass es zu jedem $y \in S^{n-1}$ ein $x \in S^{n-1}$ gibt mit

[23] Dagegen ist die folgende Aussage **falsch**: Ist Q_A indefinit, dann gibt es eine Umgebung U von A in $\mathbb{R}^{n\times n}$, so dass gilt

$$\forall B \in U : \forall x \in \mathbb{R}^n : (Q_A(x) > 0 \Longrightarrow Q_B(x) > 0).$$

Wie so oft ist hier die Reihenfolge der Quantoren zu beachten! Wahr ist: für alle x mit $Q_A(x) > 0$, gibt es eine Umgebung U so dass .... Falsch ist hingegen: Es gibt ein U so dass für alle x mit $Q_A(x) > 0$ gilt, ... In anderen Worten: das U kann nicht unabhängig von x gewählt werden.

$B_\epsilon(y) \cap S^{n-1} \subset B_{\epsilon(x)}(x) \cap S^{n-1}$. Es gilt[24] dann

$$\forall B \in \mathbb{R}^{n\times n} : \forall y \in S^{n-1} : \|B - A\| < \epsilon \Longrightarrow F(y, B) = Q_B(y) > 0.$$

Also ist für solche B die quadratische Form Q_B positiv definit.

3. Fall: negativ analog. Folgt durch Vorzeichenwechel aus dem 2. Fall. □

3.5. Lokale Extrema.

Wiederholung.

Sei (M, d) ein metrischer Raum, $f : M \to \mathbb{R}$, $p \in M$. Wir sagen f hat ein lokales Maximum (bzw. lokales Minimum) in p, falls es eine Umgebung U von p gibt, so dass

$$\forall x \in U : f(x) \leq f(p) \quad (\text{bzw. } f(x) \geq f(p)).$$

lokales Extremum in p $\quad :\Longleftrightarrow \quad$ lokales Maximum oder Minimum in p

Wir sagen f hat ein lokales striktes Maximum (bzw. lokales striktes Minimum) in p, falls es eine Umgebung U von p gibt, so dass

$$\forall x \in U \setminus \{p\} : f(x) < f(p) \quad (\text{bzw. } f(x) > f(p)).$$

Lemma 3.16. Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$. Die Funktion f habe ein lokales Minimum in $p \in U$.

(1) Wenn die Richtungsableitung $\partial_{p,v} f$ existiert, so gilt $\partial_{p,v} f = 0$.
(2) Wenn f in p differenzierbar ist, so gilt $f'(p) = 0$.

[24]Wir schließen auch hieraus, dass alle zu y so gewählten x die Ungleichung $\epsilon \leq \epsilon(x)$ erfüllen. Achtung: Wir haben aber **nicht** behauptet $\forall x \in S^{n-1} : \epsilon \leq \epsilon(x)$. Man könnte dies durch geschicktere Konstruktion von $\epsilon(x)$ erreichen, aber wir haben keine Kontrolle, ob es mit dme oben gewählten funktioniert.

(3) Wenn f auf einer Umgebung von p differenzierbar ist und in p zweimal differenzierbar ist, so ist $f''(p) = \mathrm{Hess}_p f$ positiv semi-definit.

Bei lokalen Maxima in p gilt alles analog, wenn wir positiv semi-definit durch negativ semi-definit ersetzen.

Beweis. Wenn f ein lokales Minimum in p hat, dann hat $f_{p,v}$, $f_{p,v}(t) := f(p + tv)$ ein lokales Minimum in 0. Es folgt $\partial_{p,v} f = f'_{p,v}(0) = 0$, falls es existiert. Es folgt $f'(p) = 0$ falls es existiert. Weiter gilt (falls $f''(p)$ existiert):

$$v^T f''(p) v = \mathrm{Hess}_p f(v, v) = f''_{p,v}(0) \geq 0.$$

Lokales Maximum analog. □

Definition 3.17. Man nennt p einen kritischen Punkt oder einen stationären Punkt von f, falls $f'(p) = 0$.

Also

$$\text{lokales Extremum in } p \implies p \text{ kritischer Punkt}$$

Satz 3.18. Sei U offen in $\mathbb{R}^n$, $f \in C^2(U, \mathbb{R})$ (d.h. $f : U \to \mathbb{R}$ ist zweimal stetig differenzierbar). Sei p ein kritischer Punkt von f.

(1) Ist $f''(p)$ positiv definit, dann hat f ein lokales striktes Minimum in p.
(2) Ist $f''(p)$ negativ definit, dann hat f ein lokales striktes Maximum in p.
(3) Ist $f''(p)$ indefinit, dann hat f weder ein lokales Maximum noch ein lokales Minimum in p.

!ACHTUNG!. Keine derartige Aussage möglich, falls $f''(p)$ semi-definit ist.

Mi 12.6.

Beweis. Ist $f''(p)$ positiv definit (bzw. negativ definit), so gibt es wegen Lemma 3.13 und der Stetigkeit von $U \to \mathbb{R}^{n\times n}$, $x \mapsto f''(x)$ eine Umgebung V von p in U, so dass $f''(q)$ ebenfalls positiv definit (bzw. negativ definit) für alle $q \in V$ ist. Wir wenden die Taylorformel für $\ell = 1$ an:

$$f(p+h) = f(p) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i\, \partial x_j}(p+\vartheta h)h_i h_j = f(p) + \frac{1}{2}h^T f''(p+\vartheta h)h,$$

falls $[p, p+h] \subset V$, $h \neq 0$. Also ist $f(p+h) > f(p)$ (bzw. $f(p+h) < f(p)$) für alle $p + h \in V \smallsetminus \{p\}$.

Der indefinite Fall ist ähnlich. Wir wählen $h_+, h_- \in \mathbb{R}^n$ mit $h_+^T f''(p)h_+ > 0$ und $h_-^T f''(p)h_- < 0$. Nach Zusatz 3.14 gibt es eine Umgebung V um p in U, so dass $h_+^T f''(q)h_+ > 0$ und $h_-^T f''(q)h_- < 0$ für alle $q \in U$. Es folgt dann für genügend kleines $\lambda > 0$:

$$f(p+\lambda h_+) > f(p) \qquad f(p+\lambda h_-) < f(p).$$

□

Beispiel 3.19. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^3 - 4x - xy + \frac{1}{2}y^2$.

$f'(x,y) = (3x^2 - 4 - y, -x + y)$. Die kritische Punkte sind also $p_1 := (-1,-1)$ und $p_2 := (4/3, 4/3)$.

Wir rechnen weiter

$$f''(x,y) = \begin{pmatrix} 6x & -1 \\ -1 & 1 \end{pmatrix}.$$

Es ist also $f''(p_1)$ indefinit: dies ist ein Sattelpunkt. Und $f''(p_2)$ ist positiv definit: dies ist ein lokales striktes Minimum.

4. Lokale Umkehrung differenzierbarer Abbildungen

Frage:
Sei $\emptyset \neq U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^k$ stetig, sei $V := f(U)$ offen in $\mathbb{R}^k$. Wann besitzt f eine (stetige) Umkehrfunktion $g : V \to U$?

Teil-Antwort:
Fall: $n \neq k$. Dann hat keine Funktion f wie oben eine Umkehrfuntion! Beweis nicht ganz einfach (In der Vorlesung Topologie I)[25].

Ab jetzt $n = k$.

Wenn $U = \mathbb{R}^n$ und wenn f eine lineare Abbildung ist, sagen wir $f(x) = Ax$, $A \in \mathbb{R}^{n\times n}$, dann wurde diese Frage intensiv in der Linearen Algebra diskutiert:

$$
\begin{aligned}
& f \text{ besitzt eine Umkehrfunktion} \\
\iff\ & f \text{ besitzt eine lineare Umkehrfunktion} \\
\iff\ & f \text{ ist injektiv} \\
\iff\ & f \text{ ist surjektiv} \\
\iff\ & \det f \neq 0 \\
\iff\ & A \text{ ist eine invertierbare Matrix}
\end{aligned}
$$

Hierbei $\det f = \det A$.

[25]Und man muss hier sehr genau aufpassen, wie man argumentiert. Man könnte zum Beispiel vermuten, dass f nicht surjektiv ist, falls $n > k$. Dies ist aber nicht richtig. Es gibt zum Beispiel stetige surjektive Abbildungen $\mathbb{R} \to \mathbb{R}^n$ für alle $n \in \mathbb{N}_{>0}$. Diese Funktionen sind aber nicht injektiv.

Ähnlich gilt für $A \in \mathbb{R}^{n\times n}$, $p \in \mathbb{R}^n$ und $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(x) := Ax + p$:

$$f \text{ invertierbar} \iff A \text{ invertierbar}$$

und gegebenenfalls $f^{-1}(y) = A^{-1}(y - p)$.

Notation für die Einheitsmatrix:

$$\mathbb{1}_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}$$

$$\mathrm{GL}(n, \mathbb{R}) := \{A \in \mathbb{R}^{n\times n} \mid \det A \neq 0\}.$$

Ab jetzt sei f differenzierbar (aber wir nehmen nicht mehr an, dass f linear sein muss).

Lemma 4.1. Sei U offen in $\mathbb{R}^n$ und $f : U \to \mathbb{R}^n$ differenzierbar in $x \in U$. Sei $V := f(U)$ offen in $\mathbb{R}^n$ und $g : V \to \mathbb{R}^n$ differenzierbar in $f(x)$ mit $g \circ f = \mathrm{id}_U$. Dann gilt:

$$g'(y) \cdot f'(x) = \mathbb{1}_n, \text{ wobei } y = f(x)$$

Insbesondere ist dann $f'(x)$ eine invertierbare Matrix.

Beweis. Kettenregel. □

Aus Kapitel 5 wissen wir (Proposition 1.6 und einige verbundene Aussagen):

Proposition. Sei (a, b) ein Intervall, $f : (a, b) \to \mathbb{R}$ differenzierbar, mit $f'(x) \neq 0$ für alle $x \in (a, b)$. Dann ist f injektiv und $f((a, b))$ ist auch ein offenes Intervall. Die

Umkehrfunktion

$$g := f^{-1} : f((a,b)) \to (a,b)$$

existiert, ist differenzierbar und es gilt für alle $x \in (a,b)$

$$g'(y) = \Big(f'(x)\Big)^{-1}, \text{ wobei } y = f(x)$$

Ein ähnlicher Satz ist nicht nur für $n = 1$, sondern für alle $n \in \mathbb{N}$ richtig, der lokale Umkehrsatz.

DEFINITION 4.2. Sei U offen in $\mathbb{R}^n$, $p \in U$ und $f : U \to \mathbb{R}^n$. Wir sagen f ist bei p lokal umkehrbar, falls es eine offene Umgebung V von p in U und eine offene Umgebung W von $f(p)$ in $\mathbb{R}^n$ gibt, so dass $f|_V : V \to W$ bijektiv ist.

SATZ 4.3 (Lokaler Umkehrsatz)**.** Sei U offen in $\mathbb{R}^n$, $p \in U$. Sei $f \in C^\ell(U, \mathbb{R}^n)$, $\ell \geq 1$. Ist $\det(f'(p)) \neq 0$, dann gibt es offene Umgebung V von p in U und eine offene Umgebung W von $f(p)$ in $\mathbb{R}^n$, so dass gilt

(1) $f|_V : V \to W$ bijektiv. Somit ist f lokal umkehrbar bei p.
(2) $(f|_V)^{-1} : W \to V \subset \mathbb{R}^n$ ist ℓ-mal stetig differenzierbar. [26]
(3) $\left[(f|_V)^{-1}\right]'(y) = (f'(x))^{-1}$, für alle $x \in V$, wobei $y = f(x)$

Der Satz weicht an einigen wichtigen Stellen von der obigen Proposition ab. Zum einen muss man natürlich „$f'(x) \neq 0$“ durch „$f'(x)$ invertierbar“ ersetzen. Es gibt aber einige weitere wichtige Unterschiede

- Keine globale Aussage für Umkehrung, nur eine lokale Aussage.

[26] In anderen Worten $(f|_V)^{-1} \in C^\ell(W, \mathbb{R}^n)$.

- Wir benötigen stetige Differenzierbarkeit.[27]

BEISPIELE 4.4.

(1) Ist f affin-linear, sagen wir $f(x) = Ax+p$, mit $A \in \mathbb{R}^{n\times n}$, $p \in \mathbb{R}^n$. Dann ist $f'(p) = A$. Falls $A \in \mathrm{GL}(n, \mathbb{R})$, dann kann man $V = W = \mathbb{R}^n$ wählen, f^{-1} siehe oben.

(2) Wir betrachten die komplexe Exponentialfunktion $\exp : \mathbb{C} \to \mathbb{C}$ mit Hilfe von $\mathbb{C} \cong \mathbb{R}^2$ als Funktion $f : \mathbb{R}^2 \to \mathbb{R}^2$,

$$f(x, y) = \begin{pmatrix} \mathbf{Re}\, \exp(x + iy) \\ \mathbf{Im}\, \exp(x + iy) \end{pmatrix} = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}$$

$$f'(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} = e^x \begin{pmatrix} \cos y & -\sin y \\ \sin y & \cos y \end{pmatrix}$$

$$\det f'(x, y) = e^{2x} \cdot 1 \neq 0.$$

Der Definitionsbereich $\mathbb{R}^2$ von f ist zusammenhängend, $f'(p)$ ist invertierbar für alle $p \in D(f)$, aber dennoch ist f nicht injektiv. Deswegen: nur lokale Umkehrung möglich.

Beweis.

O.B.d.A. $p = 0$, $f(p) = 0$, $f'(p) = \mathbb{1}_n$.

Wenn f dies nicht erfüllt, dann definiere

$$g(x) = f(x + p) - f(p)$$

[27]Genauer: wir nutzen diese Voraussetzung im Beweis. Frau Astrid Bönning, Nürnberg, hat mir mitgeteilt, dass eine Version mit totaler Differenzierbarkeit in ihrer Diplomarbeit zu finden ist.

und

$$\hat{f}(x) = g'(0)^{-1} g(x).$$

Dann haben wir $\hat{f}(0) = 0$ und $\hat{f}'(0) = \mathbb{1}_n$. Und wenn wir den lokalen Umkehrsatz für $\hat{f}$ in 0 zeigen, dann folgt hieraus unmittelbar der lokale Umkehrsatz für f in p.

Zu (1); Teil (i): Bestimmung von V und W.

Im folgenden sei $\| \cdot \|$ immer die Supremumsnorm auf $\mathbb{R}^n$ und auf $\mathbb{R}^{n\times n}$, d.h.

$$\left\| \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \right\| = \max\{|a_1|, \ldots, |a_n|\}, \qquad \left\| (a_{ij})_{i,j\in\{1,\ldots,n\}} \right\| = \max\{|a_{11}|, |a_{12}|, \ldots, |a_{1n}|, |a_{21}|, \ldots, |a_{nn}|\}.$$

Also $\|Ax\| \leq n\|A\|\,\|x\|$ für $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$.

Definiere $\varphi(x) = x - f(x)$, also $\varphi'(0) = 0$.

Bestimme $r > 0$ so, dass[28]

(a) $B_{2r}(0) = \{x \in \mathbb{R}^n \mid \|x\| < 2r\} \subset U$
(b) $\|\varphi'(x)\| < \frac{1}{2n}$ für alle $x \in B_{2r}(0)$

(b) ist möglich, da φ stetig differenzierbar ist[29].

[28]Eine kleine Bemerkung zu den hier verwendeten „Bällen“ $B_r(p)$. Es sind Bälle nicht – wie zumeist üblich – bezüglich der Standardnorm, sondern bezüglich der Supremumsnorm. Deswegen ist $B_R(p)$ eigentlich ein achsenparalleler Würfel mit Seitenlänge $2R$ und mit Schwerpunkt p.

[29]und da Urbilder offener Mengen unter stetigen Abbildungen wieder offen sind

Für $x, y \in \overline{B_r(0)} = \{x \in \mathbb{R}^n \mid \|x\| \leq r\}$, wenden wir nun den Mittelwertsatz auf die j-te Komponente φ_j von φ an[30] und erhalten für ein $\vartheta \in (0,1)$:

$$|\varphi_j(y)-\varphi_j(x)| = |\varphi_j'(\underbrace{x+\vartheta(y-x)}_{z:=})\cdot(y-x)| \leq \|\varphi'(z)\cdot(y-x)\| \leq n\|\varphi'(z)\|\,\|y-x\| \leq \frac{n}{2n}\|y-x\|.$$

Daraus ergibt sich dann

$$\|\varphi(y)-\varphi(x)\| \leq \frac{1}{2}\|y-x\|. \tag{4.5}$$

Setze nun $W := B_{r/2}(0)$ und $V := B_r(0) \cap f^{-1}(W)$. Letzteres ist eine offene Umgebung von 0. Wir haben

$$f(V) \subset f(f^{-1}(W)) \subset W$$

Fr 14.6.

Zu (1); Teil (ii): $f|_V : V \to W$ ist bijektiv.

$f|_V$ injektiv: Seien $x, y \in V$ mit $f(x) = f(y)$. Dann gilt $x - \varphi(x) = y - \varphi(y)$, also

$$\|y-x\| = \|\varphi(y)-\varphi(x)\| \overset{(4.5)}{\leq} \frac{1}{2}\|y-x\|$$

Also $y = x$.

$f|_V$ surjektiv: Gegeben sei $w \in W$. Wir definieren

$$\psi_w : \overline{B_r(0)} \to \mathbb{R}^n, \qquad x \mapsto w + \varphi(x).$$

Die Abbildung ist eine Kontraktion, denn

$$\|\psi_w(y)-\psi_w(x)\| = \|\varphi(y)-\varphi(x)\| \overset{(4.5)}{\leq} \frac{1}{2}\|y-x\|.$$

[30] Also $\varphi_j : U \to \mathbb{R}$, $\varphi(x) = (\varphi_1(x), \ldots, \varphi_n(x))^T$

Es gilt $\psi_w(\overline{B_r(0)}) \subset B_r(0)$, denn aus (4.5) folgt:

$$\|\psi_w(x)\| \leq \underbrace{\|w\|}_{<r/2} + \underbrace{\|\varphi(x) - \overbrace{\varphi(0)}^{=0}\|}_{\leq r/2} < r.$$

Da $\overline{B_r(0)}$ vollständig ist, gibt es nach dem Banachschen Fixpunktsatz[31] ein $z \in \overline{B_r(0)}$ mit $\psi_w(z) = z$ und man sieht dann $z \in B_r(0)$.[32] Es gilt also

$$z = \psi_w(z) = w + \varphi(z) = w + z - f(z),$$

also $f(z) = w$. Daraus folgt dann auch $z \in f^{-1}(W)$, d.h. $z \in V$.

Zu (2); Teil (i): $(f|_V)^{-1} : W \to V \subset \mathbb{R}^n$ **ist differenzierbar auf** W**.**

Überprüfe Differenzierbarkeit in $w \in W$, $w = f(\tilde{p})$. Zu h mit $\tilde{p} + h \in V$ wähle ein k mit

$$f(\tilde{p} + h) = w + k.$$

Wir können hier aber auch h als Funktion in k betrachten, da $f|_V$ bijektiv.

$$k = f(\tilde{p} + h) - f(\tilde{p}) = f'(\tilde{p}) \cdot h + \epsilon(h)\|h\|, \qquad \lim_{h \to 0} \epsilon(h) = 0$$

Wir multiplizieren diese vektorielle Gleichung mit $(f'(\tilde{p}))^{-1}$, falls es existiert.

$$\begin{aligned} (f'(\tilde{p}))^{-1}k &= h + \|h\|(f'(\tilde{p}))^{-1}\epsilon(h) \\ &= (\tilde{p} + h) - \tilde{p} + \|h\|(f'(\tilde{p}))^{-1}\epsilon(h) \\ &= f^{-1}(w + k) - f^{-1}(w) - \|k\|\delta(k) \end{aligned}$$

[31] angewendet auf $f := \psi_w$ und $Y := X := \overline{B_r(0)}$

[32] Denn es ist ja im Bild von ψ_w.

wobei wir definieren

$$\delta(k) = -\frac{\|h\|}{\|k\|}(f'(\tilde{p}))^{-1}\epsilon(h)\,.$$

Die Behauptung folgt also, sobald wir $\lim_{k\to 0}\delta(k) = 0$ gezeigt haben.

Wir rechnen:

$$\varphi(\tilde{p}+h) - \varphi(\tilde{p}) = \tilde{p} + h - f(\tilde{p}+h) - \tilde{p} + f(\tilde{p}) = h - k$$

Also

$$\|h\| \leq \|\varphi(\tilde{p}+h) - \varphi(\tilde{p})\| + \|k\| \overset{(4.5)}{\leq} \frac{1}{2}\|h\| + \|k\|.$$

Es folgt $\|h\| \leq 2\|k\|$ und somit

$$\lim_{k\to 0}\delta(k) = 0.$$

Wir haben gezeigt: ist $f'(\tilde{p})$ invertierbar, dann ist $(f|_V)^{-1}$ differenzierbar in $w = f(\tilde{p})$ und es gilt dann $\left((f|_V)^{-1}\right)'(w) = \left(f'(\tilde{p})\right)^{-1}$.

Die Umgebung V wurde so gewählt, dass $\|\varphi'(\tilde{p})\| < \frac{1}{2n}$ für alle $\tilde{p} \in V$. Aus $\varphi(x) = x - f(x)$ folgt $\varphi'(\tilde{p}) = \mathbb{1}_n - f'(\tilde{p})$, somit

$$\|f'(\tilde{p})\cdot x\| \geq \|x\| - \|\varphi'(\tilde{p})\cdot x\| \geq \|x\| - n\|\varphi'(\tilde{p})\|\,\|x\| \geq \|x\| - \frac{n}{2n}\|x\| \geq \frac{1}{2}\|x\|.$$

Also ist $f'(\tilde{p})$ eine invertierbare Matrix. Somit ist $(f|_V)^{-1}$ differenzierbar auf W; und es gilt (3). **Zu (3).**

Diese Formel folgt nun direkt aus den Überlegungen unter (2), Teil (i).

Zu (2); Teil (ii): ℓ-fache stetige Differenzierbarkeit.

Auf Grund von (3) bildet die Komposition

$$W \xrightarrow{(f|_V)^{-1}} V \xrightarrow{f'} \mathrm{GL}(n,\mathbb{R}) \xrightarrow{\mathrm{inv}} \mathrm{GL}(n,\mathbb{R})$$

$w \in W$ auf $\left[(f|_V)^{-1}\right]'(w)$ ab. Da die Inversen-Abbildung

$$\mathrm{inv} : \mathrm{GL}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R}), \qquad A \mapsto A^{-1}$$

glatt ist,[33] folgt durch Induktion aus der ℓ-fachen stetigen Differenzierbarkeit von f auch die ℓ-fache stetige Differenzierbarkeit von $(f|_V)^{-1}$. □

Beispiel 4.6. $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) := \left(\frac{1}{2}x^2 + y,\ xy\right)^T$.

$$f'(x,y) := \begin{pmatrix} x & 1 \\ y & x \end{pmatrix}, \quad \det f'(x,y) = x^2 - y\,.$$

Also ist f bei allen Punkten (x,y) mit $x^2 \neq y$ lokal umkehrbar. In den Punkten der Form (x,x^2) macht der lokale Umkehrsatz keine Aussage, ob eine lokale Umkehrfunktion existiert. Wenn dort aber in (x,x^2) eine lokale Umkehrfunktion existiert, so ist sie nach Lemma 4.1 in $f(x,x^2)$ nicht differenzierbar.

Es gibt also sehr oft lokale Umkehrfunktionen. In den meisten Fällen kann man diese lokale Umkehrfunktion aber nicht explizit (d.h. als Formelausdruck) angeben. Man kennt dann die Existenz der Umkehrfunktion, ohne sie als Formelausdruck hinschreiben zu können.

Definition 4.7. Sind U und V offene Mengen in $\mathbb{R}^n$, $\ell \in \mathbb{N}_{\geq 1}$. Man nennt $f : U \to V$ einen C^ℓ-Diffeomorphismus (von U auf V), falls $f : U \to V$ bijektiv ist und sowohl f als

[33] Die Formel für A^{-1}, die Sie aus der Linearen Algebra inzwischen kennen sollten, hat Einträge, die rationale Funktionen in den Koeffizienten von A sind. Die Nenner dieser rationalen Funktionen sind nicht Null, da die Matrix invertierbar ist.

auch f^{-1} ℓ-mal stetig differenzierbar sind. In der Analysis II und III verwenden wir das Wort Diffeomorphismus im Sinne von C^1-Diffeomorphismus.

Der lokale Umkehrsatz liefert:

LEMMA 4.8. Sei U offen in $\mathbb{R}^n$ und $f : U \to \mathbb{R}^n$ eine Funktion, $V := f(U)$. Dann ist f genau dann ein Diffeomorphismus von U auf V, wenn

- $f \in C^1(U, \mathbb{R}^n)$,
- $\det f'(x) \neq 0$ für alle $x \in U$,
- f injektiv ist.

□

5. Der Satz über implizit definierte Funktionen

Motivation: Gegeben sei eine Funktion $F : \mathbb{R}^2 \to \mathbb{R}$. Wir wollen alle Lösungen von

$$F(x, y) = 0$$

bestimmen. Am liebsten wäre es uns, wenn wir eine Lösungsfunktion hätten, d.h. eine differenzierbare Funktion $f : (a, b) \to \mathbb{R}$, so dass $F(x, f(x)) = 0$ und so dass alle oder möglichst viele Lösungen von der Form $(x, f(x))$ sind.

BEISPIELE 5.1.

(1) $F(x, y) := y - x^3 + x^2 - 1 = 0$. Dann ist $f(x) = x^3 - x^2 + 1$ solch eine Lösungsfunktion.

Bild der zugehörigen Lösungsmenge

$$\frac{\partial F}{\partial y} = 1 \neq 0 \qquad \text{in allen Lösungen } (x, y)$$

(2) $F(x, y) := y^2 - x^2 - 1 = 0$. Dann benötigen wir zwei Funktionen, um die Lösungen zu beschreiben

$$f_+(x) = \sqrt{x^2 + 1}, \qquad f_-(x) = -\sqrt{x^2 + 1}, \qquad f_\pm : \mathbb{R} \to \mathbb{R}.$$

Bild der Graphen von f_+ und f_-.

$$\frac{\partial F}{\partial y} = 2y \neq 0 \qquad \text{in allen Lösungen } (x, y)$$

(3) $F(x, y) = x - y^2 = 0$.

Bild der zugehörigen Lösungsmenge

Eine Beschreibung der Lösungen in der Form $(x, f(x))$ ist schwierig, da wir zum einen zwei Funktionen benötigen ($x \mapsto \sqrt{x}$ und $x \mapsto -\sqrt{x}$) und da die Wurzelfunktion in 0 nicht differenzierbar ist. [34]

$$\frac{\partial F}{\partial y} = 2y \neq 0 \qquad \text{in allen Lösungen außer in } (0, 0)$$

Problemlos möglich, wenn wir x und y vertauschen. Sei $g : \mathbb{R} \to \mathbb{R}$, $g(y) = y^2$. Die Lösungsmenge von $x - y^2 = 0$ ist

$$\{(g(y), y) \mid y \in \mathbb{R}\}.$$

$$\frac{\partial F}{\partial x} = 1 \neq 0 \qquad \text{in allen Lösungen}$$

(4) $F(x, y) = x - y^3 = 0$.

[34] Da die Quadratwurzelfunktion auf $[0, \infty)$ definiert ist, haben wir streng genommen gar nicht definiert, was „differenzierbar in 0“ bedeutet, denn wir benötigen eigentlich dazu, dass die Funktion auf einer Umgebung der 0 definiert ist. Definition 1.1 in Kapitel 5 ist aber auch sinnvoll, wenn wir nur fordern dass der Punkt x_0 ein Häufungspunkt der Definitionsmenge ist. Mit dieser Definition ist die Wurzelfunktion nicht differenzierbar in 0.

Bild der zugehörigen Lösungsmenge

Eine Beschreibung der Lösungen in der Form $(x, f(x))$ ist möglich durch die Kubikwurzel $x \mapsto \sqrt[3]{x}$, aber sie ist in 0 nicht differenzierbar.

$$\frac{\partial F}{\partial y} = 3y^2 \neq 0 \qquad \text{in allen Lösungen außer in } (0,0)$$

Die Differenzierbarkeit ist aber wiederum gegeben, wenn wir zunächst x und y vertauschen. Sei $g : \mathbb{R} \to \mathbb{R}$, $g(y) = y^3$. Die Lösungsmenge von $x - y^3 = 0$ ist

$$\{(g(y), y) \mid y \in \mathbb{R}\}.$$

$$\frac{\partial F}{\partial x} = 1 \neq 0 \qquad \text{in allen Lösungen}$$

(5) $F(x, y) = x^2 - y^2 = 0$. Zwei Lösungsfunktionen $f_\pm(x) = \pm x$, die sich schneiden

Bild der zugehörigen Lösungsmenge

$$\frac{\partial F}{\partial y} = -2y \neq 0 \qquad \text{in allen Lösungen außer in } (0,0)$$

(6) $F(x, y) = x^3 - y^2 = 0$.

Bild der zugehörigen Lösungsmenge $\{(\sqrt[3]{y^2}, y)^T \mid y \in \mathbb{R}\}$

$$\frac{\partial F}{\partial y} = -2y \neq 0 \qquad \text{in allen Lösungen außer in } (0,0)$$

(7) $F(y, x) = x^2 + y^2 = 0$. Lösungsmenge $\{(0,0)\}$. Nur ein Punkt!

$$\frac{\partial F}{\partial y}(0,0) = 0$$

Vermutung: Es geht gut, wenn $\frac{\partial F}{\partial y}(x, y) \neq 0$ für alle Lösungen (x, y). Der Satz über implizite Funktionen macht genau solch eine Aussage.

Notation.

[35] Von nun an sei $U \neq \emptyset$ offen in $\mathbb{R}^{n+m}$ und $F : U \to \mathbb{R}^m$ stetig differenzierbar. Punkte in $\mathbb{R}^{n+m}$ schreiben wir zumeist als (x, y) mit $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, $y = (y_1, \ldots, y_m)^T \in \mathbb{R}^m$, $F(x, y) = (F_1(x, y), \ldots, F_m(x, y))$.

$$\frac{\partial F}{\partial x}(x, y) := \left(\frac{\partial F_i}{\partial x_j}(x, y)\right)_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} \in \mathbb{R}^{m \times n}$$
$$\frac{\partial F}{\partial y}(x, y) := \left(\frac{\partial F_i}{\partial y_j}(x, y)\right)_{\substack{i=1,\ldots,m \\ j=1,\ldots,m}} \in \mathbb{R}^{m \times m}$$

Ziel: Finde Funktionen

$$\begin{aligned} y_1 &= y_1(x_1, \ldots, x_n) \\ y_2 &= y_2(x_1, \ldots, x_n) \\ \vdots &\quad \vdots \\ y_m &= y_m(x_1, \ldots, x_n) \end{aligned}$$

so dass $(x, y(x))$ Lösungen von $F(x, y) = 0$.

Satz 5.2 (Satz über implizite Funktionen)**.** Sei $\emptyset \neq U \subset \mathbb{R}^{n+m}$ offen, und $F : U \to \mathbb{R}^m$ stetig differenzierbar. In einem Punkt $(\hat{x}, \hat{y}) \in U$ soll $F(\hat{x}, \hat{y}) = 0$ und $\det \frac{\partial F}{\partial y}(\hat{x}, \hat{y}) \neq 0$ gelten. Dann gibt es

- eine offene Umgebung V von $\hat{x}$ in $\mathbb{R}^n$,
- eine offene Umgebung W von $\hat{y}$ in $\mathbb{R}^m$,
- eine Funktion $f : V \to W$

[35] Die Indizes k, n und M wechseln etwas ihre Rollen im Vergleich zu den vorherigen Kapiteln

mit den folgenden Eigenschaften

(a) $V \times W \subset U$

(b) $\{(x, y) \in V \times W \mid F(x, y) = 0\} = \{(x, f(x)) \mid x \in V\}$

(c) f ist stetig differenzierbar und es gilt

$$f'(x) = -\left(\frac{\partial F}{\partial y}(x, f(x))\right)^{-1} \frac{\partial F}{\partial x}(x, f(x))$$

19.6.

Veranschaulichung des Satzes über implizite Funktionen

BEMERKUNG 5.3. Zu gegebenem F, U, $\hat{x}$ und $\hat{y}$ liefert der Satz: f, V, W. Diese sind nicht eindeutig, z.B. kann man V verkleinern und dann f einschränken. Erfüllen $\tilde{f}$, $\tilde{V}$, $\tilde{W}$ ebenfalls den Satz und setzen wir $Z := V \cap \tilde{V} \cap f^{-1}(W \cap \tilde{W})$, dann gilt $f|_Z = \tilde{f}|_Z$.

BEMERKUNG 5.4. Sei $U \subset \mathbb{R}^m$ offen und $f : U \to \mathbb{R}^m$ stetig differenzierbar. Wenn wir den Satz über implizite Funktionen auf $F : \mathbb{R}^m \times U \to \mathbb{R}^m$, $F(x, y) := x - f(y)$ anwenden $(n = m)$, so erhalten wir den lokalen Umkehrsatz.

BEMERKUNG 5.5. Zu gegebenem F, U, $\hat{x}$ und $\hat{y}$ liefert der Satz: f, V, W. Diese sind nicht eindeutig, z.B. kann man V verkleinern und dann f einschränken. Erfüllen $\tilde{f}$, $\tilde{V}$, $\tilde{W}$ ebenfalls den Satz und setzen wir $Z := V \cap \tilde{V} \cap f^{-1}(W \cap \tilde{W})$, dann gilt $f|_Z = \tilde{f}|_Z$.

BEMERKUNG 5.6. Sei $U \subset \mathbb{R}^m$ offen und $f : U \to \mathbb{R}^m$ stetig differenzierbar. Wenn wir den Satz über implizite Funktionen auf $F : \mathbb{R}^m \times U \to \mathbb{R}^m$, $F(x, y) := x - f(y)$ anwenden $(n = m)$, so erhalten wir den lokalen Umkehrsatz.

DEFINITION 5.7. Die Funktion f nennt man eine durch $F(x, y) = 0$ implizit definierte Funktion.

BEMERKUNG 5.8. Aussage (b) besteht eigentlich aus zwei Teilen. Zum einen besagt die Aussage, dass $(x, f(x))$ eine Lösung ist, also $F(x, f(x)) = 0$ für alle $x \in V$. Zum anderen sind lokal (d.h. heißt hier in $V \times W$) alle Lösungen von dieser Form.

BEMERKUNG 5.9. Oft kann man implizite Funktionen nicht explizit angeben. „Explizit angeben“ heißt hier durch einen Formelausdruck, der sich aus Standard-Operationen zusammensetzt.

ZUSATZ 5.10 (zum Satz5.2 über implizite Funktionen)**.** Sei $\ell \geq 1$. Ist F ℓ-mal stetig differenzierbar, so ist f ebenfalls ℓ-mal stetig differenzierbar.

BEISPIEL 5.11. Wir betrachten die Funktion

$$F : \mathbb{R}^{n+1} \to \mathbb{R}, \qquad F(z) := \langle z, z \rangle - 1.$$

$F \in C^\infty(\mathbb{R}^{n+1}, \mathbb{R})$. Dann ist die Lösungsmenge die n-dimensionale Sphäre

$$S^n := \{z \in \mathbb{R}^{n+1} \mid F(z) = 0\} = \{z \in \mathbb{R}^{n+1} \mid \|z\|_2 = 1\}.$$

Wir schreiben $z = (x, y)$ mit $x \in \mathbb{R}^n$ und $y \in \mathbb{R}$. Nun ist $(\partial F/\partial y)(x, y) = 2y$ und dies ist für $y \neq 0$ nicht Null. Zu jedem $(\hat{x}, \hat{y}) \in S^n$ mit $\hat{y} \neq 0$ kann man den Satz über implizite Funktionen anwenden.

Wir beschränken uns auf den Fall $\hat{y} > 0$. Dann sind mögliche Wahlen:

$V := B_1(0, \mathbb{R}^n) := \{x \in \mathbb{R}^n \mid \|x\| < 1\}$, $W := (0, \infty)$, $f(x) := \sqrt{1 - \|x\|_2^2}$

Eine andere Wahl wäre, falls $\|\hat{x}\| < 1 - \delta$:

$V := B_{1-\delta}(0, \mathbb{R}^n)$, $W := (0, \infty)$, $f(x) := \sqrt{1 - \|x\|_2^2}$

Eine weitere mögliche Wahl ist: $V := B_1(0, \mathbb{R}^n)$, $W := \left(0, \frac{1001}{1000}\right)$, $f(x) := \sqrt{1 - \|x\|_2^2}$

Nicht erlaubt sind:

$V := B_1(0, \mathbb{R}^n)$, $W := (0, 1)$, (W zu klein)

$V := B_1(0, \mathbb{R}^n)$, $W := \mathbb{R}$, (W zu groß)

Zeichnung des Standardkreises mit den oben erklärten Mengen.

Beweis des Satzes5.2 über implizite Funktionen und des Zusatzes. Sei F wie im Satz gegeben. Wir definieren: $G : U \to \mathbb{R}^{n+m}$, $G(x, y) := \begin{pmatrix} x \\ F(x, y) \end{pmatrix}$. Wir haben $G \in C^\ell(U, \mathbb{R}^{m+n})$ und

$$G'(x, y) = \begin{pmatrix} \mathbb{1}_n & 0 \\ * & \frac{\partial F}{\partial y}(x, y) \end{pmatrix}$$

Also

$$\det G'(\hat{x}, \hat{y}) = \underbrace{(\det \mathbb{1}_n)}_{=1} \cdot \left(\det \frac{\partial F}{\partial y}(\hat{x}, \hat{y}) \right) \neq 0 \,.$$

Wende den lokalen Umkehrsatz auf G im Punkt $p := (\hat{x}, \hat{y})$ an. Es gibt dann eine offene Umgebung $\hat{V}$ von p in U und eine offene Umgebung $\hat{W}$ von $G(p)$ in $\mathbb{R}^{n+m}$, so dass $G|_{\hat{V}} : \hat{V} \to \hat{W}$ bijektiv ist und so dass $(G|_{\hat{V}})^{-1} \in C^\ell(\hat{W}, \mathbb{R}^{n+m})$. Offensichtlich hat $(G|_{\hat{V}})^{-1}$ die Form

$$(G|_{\hat{V}})^{-1}(x, y) := \begin{pmatrix} x \\ \psi(x, y) \end{pmatrix}$$

für ein $\psi \in C^\ell(\hat{W}, \mathbb{R}^m)$. Durch Differenzieren von $\mathrm{id}_{\hat{V}} = (G|_{\hat{V}})^{-1} \circ G_{\hat{V}}$ erhalten wir

(5.12)

$$\mathbb{1}_{n+m} = \left[(G|_{\hat{V}})^{-1} \right]' (x, F(x, y)) \circ G'(x, y) = \begin{pmatrix} \mathbb{1}_n & 0 \\ * & \frac{\partial \psi}{\partial y}(x, F(x, y)) \end{pmatrix} \begin{pmatrix} \mathbb{1}_n & 0 \\ * & \frac{\partial F}{\partial y}(x, y) \end{pmatrix} .$$

Ausmultiplizieren und Vergleichen des rechten unteren Blocks auf beiden Seiten ergibt

$$\mathbb{1}_m = \left(\frac{\partial \psi}{\partial y}(x, F(x,y))\right)\left(\frac{\partial F}{\partial y}(x,y)\right).$$

O.B.d.A. können wir annehmen,[36] dass $\hat{V}$ die Gestalt $\hat{V} = V_0 \times W$ hat mit V_0 offen in $\mathbb{R}^n$ und W offen in $\mathbb{R}^m$. Die Abbildung ψ bildet $\hat{W}$ in W ab, und $\psi(\hat{x}, 0) = (\hat{x}, \hat{y})^T$. Auf Grund der Stetigkeit von ψ finden wir eine offene Umgebung V von $\hat{x}$ mit $V \subset V_0$ und $V \times \{0\} \subset \hat{W}$. Dann ist $f : V \to W$, $f(x) := \psi(x, 0)$ eine wohldefinierte C^ℓ-Abbildung und $F(x, f(x)) = 0$. Wir haben somit (a), „$\supset$“ in (b) und $f \in C^\ell$.

Ist $(x, y) \in V \times W$ mit $F(x, y) = 0$, dann ist $G(x, y) = (x, 0)^T$, also $(G|_{\hat{V}})^{-1}(x, 0) = (x, y)^T$. Dies impliziert $f(x) = y$. Wir erhalten „$\subset$“ in (b).

Wenn wir die Stern-Einträge in (5.12) berücksichtigen, sieht man

$$0 = \frac{\partial \psi}{\partial x}(x, F(x,y)) + \frac{\partial \psi}{\partial y}(x, F(x,y))\frac{\partial F}{\partial x}(x,y),$$

also

$$f'(x) = \frac{\partial \psi}{\partial x}(x, 0) = -\frac{\partial \psi}{\partial y}(x,0)\frac{\partial F}{\partial x}(x, f(x)) = -\left(\frac{\partial F}{\partial y}(x, f(x))\right)^{-1}\frac{\partial F}{\partial x}(x, f(x)).$$

□

6. Untermannigfaltigkeiten

6.1. Definition und erste Beispiele.

Definition 6.1. Sei $\ell \geq 1$, $k \geq n$. Sei M eine Teilmenge von $\mathbb{R}^k$. Wir nennen M eine n-dimensionale Untermannigfaltigkeit von $\mathbb{R}^k$ mit Regularität C^ℓ, falls es zu jedem $p \in M$

[36] Wenn dies nicht der Fall ist, ersetzen wir einfach $\hat{V}$ durch eine kleinere offene Umgebung von p, die die gewünschte Produktgestalt hat, und $\hat{W}$ ist dann entsprechend zu verkleinern.

offene Mengen $U, V \subset \mathbb{R}^k$ und einen C^ℓ-Diffeomorphismus $\varphi : U \to V$ gibt, so dass $p \in U$ und $\varphi(U \cap M) = V \cap (\mathbb{R}^n \times \{0\})$. Solch ein φ nennt man eine Untermannigfaltigkeitskarte von M.

Anschaulich: der Diffeomorphismus φ verbiegt $U \cap M$ in eine offene Teilmenge von $\mathbb{R}^n \times \{0\} \subset \mathbb{R}^k$.

Ein Bild, das eine Untermannigfaltigkeitskarte graphisch darstellt mit Untermannigfaltigkeitskarte

Untermannigfaltigkeit mit Regularität C^ℓ := C^ℓ-Untermannigfaltigkeit

In der Analysis II und III nutzen wir den Begriff „Untermannigfaltigkeit“ (ohne weitere Angabe der Regularität) immer im Sinne von „C^1-Untermannigfaltigkeit“.

Spezielle Dimensionen.
In manchen Dimensionen haben Untermannigfaltigkeiten spezielle Namen:[37]

2-dimensionale Untermannigfaltigkeiten von $\mathbb{R}^k$ =: Flächen in $\mathbb{R}^k$

$(k-1)$-dimensionale Untermannigfaltigkeiten von $\mathbb{R}^k$ =: Hyperflächen in $\mathbb{R}^k$

Beispiele 6.2.

(1) Sei $M \subset \mathbb{R}^k$ ein n-dimensionaler Untervektorraum (mit $n \leq k$). Wähle eine Basis $(v_1, \ldots, v_n)$ von M. Wie in der Linearen Algebra I gezeigt, existieren Vektoren $v_{n+1}, \ldots, v_k$ in $\mathbb{R}^k$, so dass $(v_1, \ldots, v_n, v_{n+1}, \ldots, v_k)$ eine Basis von $\mathbb{R}^k$ ist. Betrachte die lineare Abbildung $\varphi : \mathbb{R}^k \to \mathbb{R}^k$, $v_i \mapsto e_i$, wobei e_i der i-te kanonische Basisvektor

[37]Der Fall $n = 1$ is hier leider etwas kompliziert: zusammenhängende 1-dimensionale Untermannigfaltigkeit =: reguläre eingebettete Kurve

ist. Dann ist φ ein linearer Isomorphismus, insbesondere ein C^∞-Diffeomorphismus mit

$$\begin{aligned}\varphi(M) &= \varphi(\mathrm{Span}\{v_1,\dots,v_n\})\\ &= \mathrm{Span}\{e_1,\dots,e_n\}\\ &= \mathbb{R}^k \cap (\mathbb{R}^n \times \{0\}).\end{aligned}$$

Also ist M eine n-dimensionale Untermannigfaltigkeit der Regularität C^∞ von $\mathbb{R}^k$.

(2) Ist W offen in $\mathbb{R}^n$, $f \in C^\ell(W,\mathbb{R}^m)$. Dann ist der Graph von f

$$\mathrm{Graph} f := \{(x, f(x)) \mid x \in W\}$$

eine n-dimensionale Untermannigfaltigkeit von $\mathbb{R}^{n+m}$ mit Regularität C^ℓ. Denn

$$\varphi : \underbrace{W \times \mathbb{R}^m}_{U:=} \to \underbrace{W \times \mathbb{R}^m}_{V:=}, \quad (x,y) \mapsto (x, y - f(x))$$

ist eine ℓ-mal stetig differenzierbare Abbildung mit ℓ-mal stetig differenzierbarer Umkehrabbildung

$$\varphi^{-1} : W \times \mathbb{R}^m \to W \times \mathbb{R}^m, \quad (x,y) \mapsto (x, y + f(x))\,.$$

$$\begin{array}{rcl} U \cap \mathrm{Graph} f = \mathrm{Graph} f = \{(x, f(x) \mid x \in W\} & \stackrel{\varphi}{\to} & W \times \{0\} = V \cap (\mathbb{R}^n \times \{0\}) \\ (x, f(x)) & \mapsto & (x, 0) \end{array}$$

Ein Bild, das die Abbildung φ darstellt

(3) Ist W offen in $\mathbb{R}^n$, $f_-, f_+ \in C^\ell(W,\mathbb{R}^m)$ mit $f_- < f_+$. Dann ist $\mathrm{Graph} f_- \cup \mathrm{Graph} f_+$ eine n-dimensionale Untermannigfaltigkeit von $\mathbb{R}^{n+m}$. Definiere zum Beispiel

$$\varphi_- : \underbrace{\{(x,y) \mid x \in W \wedge y < f_+(x)\}}_{U:=} \to \underbrace{\{(x,y) \mid x \in W \wedge y < f_+(x) - f_-(x)\}}_{V:=}, \quad (x,y) \mapsto (x, y - f_-(x))$$

$$\varphi_+ : \underbrace{\{(x,y) \mid x \in W \wedge f_-(x) < y\}}_{U:=} \to \underbrace{\{(x,y) \mid x \in W \wedge -f_+(x) + f_-(x) < y\}}_{V:=}, \quad (x,y) \mapsto (x, y - f_+(x))$$

Ein Bild, das die Abbildungen φ_+ und φ_- darstellt

Fr. 21.6.

(4) M ist 0-dimensionale Untermannigfaltigkeit von $\mathbb{R}^k \iff M$ ist eine diskrete Teilmenge von $\mathbb{R}^k :\iff$ Die Spurtopologie auf M ist die diskrete Topologie
Zum Beispiel ist $\{1/n \mid n \in \mathbb{N}_{>0}\}$ eine 0-dimensionale Untermannigfaltigkeit von $\mathbb{R}$.

(5) M ist k-dimensionale Untermannigfaltigkeit von $\mathbb{R}^k \iff M$ ist eine offene Teilmenge von $\mathbb{R}^k$

(6) Ist $M \subset \mathbb{R}^k$ eine n-dimensionale Untermannigfaltigkeit der Regularität C^ℓ, so ist auch jede offene Teilmenge M' von M eine n-dimensionale C^ℓ-Untermannigfaltigkeit von $\mathbb{R}^k$. Denn zu gegebenem $p \in M'$ wähle eine Untermannigfaltigkeitskarte $\varphi : U \to V$ mit $p \in U$. Da M' offen in M ist, gibt es eine in $\mathbb{R}^k$ offene Teilmenge U' mit $M' = M \cap U'$. Dann ist $\varphi|_{U \cap U'} : U \cap U' \to \varphi(U \cap U')$ eine Untermannigfaltigkeitskarte von M' mit $p \in U \cap U'$.

(7) Es folgt hieraus, dass die Untermannigfaltigkeits-Eigenschaft lokal ist, das heißt: Eine Teilmenge $M \subset \mathbb{R}^k$ ist genau dann eine n-dimensionale C^ℓ-Untermannigfaltigkeit, wenn jeder Punkt $p \in M$ eine offene Umgebung U besitzt, so dass $U \cap M$ eine n-dimensionale C^ℓ-Untermannigfaltigkeit ist.

Untermannigfaltigkeiten sind insbesondere für Anwendungen in der Physik und anderen Naturwissenschaften, aber auch zum Beispiel für (nicht-lineare) Optimierungsprobleme in der Betriebswirtschaftslehre sehr wichtig.

Untervektorräume von $\mathbb{R}^k$ kann man auf verschiedene Arten beschreiben, zum Beispiel

(1) durch Gleichungen, z.B. $V := \{(x_1, \ldots, x_k) \in \mathbb{R}^k \mid x_1 + 2x_2 = 0\}$
(2) als Graph, das heißt $V := \mathrm{Graph} f$ einer linearen Funktion $f : \mathbb{R}^n \to \mathbb{R}^{k-n}$
(3) $V := \mathrm{Bild}(f)$ für eine lineare Abbildung $f : \mathbb{R}^m \to \mathbb{R}^k$
(4) durch Parametrisierung, z.B. definiere zu $v_1, v_2 \in \mathbb{R}^k$:

$$V := \{tv_1 + sv_2 \mid t, s \in \mathbb{R}\}.$$

(5) Durch eine lineare Untermannigfaltigkeitskarte wie in Beispiel 6.2 (1).

Jede dieser Beschreibungen verallgemeinert sich zu einer Beschreibung von Untermannigfaltigkeiten:

(1) Abschnitt 6.2: lokal als Urbild eines regulären Werts,
(2) Abschnitt 6.2: lokal als Graph, das heißt lokal als $\mathrm{Graph} f$ einer C^ℓ-Funktion $f : U \to \mathbb{R}^{k-n}$, U offen in $\mathbb{R}$
(3) Abschnitt 6.3: lokal als Bild einer Immersion,
(4) Abschnitt 6.5: durch eine lokale Parametrisierung
(5) durch eine Untermannigfaltigkeitskarte wie in Definition 6.1

Hierbei ist (5) für theoretische Überlegungen sehr hilfreich ist, aber oft wenig geschickt für konkrete Rechnungen. Moderne Programme, die gekrümmte Flächen modellieren müssen, z.B. Raytracer wie povray nutzen verschiedene dieser Methoden, um Flächen zu beschreiben. Sie haben zumeist Hilfsmittel, um von einer Form in die andere umrechnen zu können.

6.2. Der Satz vom regulären Wert. Ein wichtiges Kriterium, um zu zeigen, dass eine Teilmenge von $\mathbb{R}^k$ eine Untermannigfaltigkeit ist, ist der Satz vom regulären Wert.

DEFINITION 6.3. Sei U offen in $\mathbb{R}^k$, $F : U \to \mathbb{R}^m$ ℓ-mal stetig differenzierbar, $\ell \geq 1$. Wir sagen $z \in \mathbb{R}^m$ ist ein regulärer Wert von F, falls gilt

$$\forall p \in F^{-1}(\{z\}) : \ F'(p) \text{ hat Rang } m. \tag{6.4}$$

Ist z kein regulärer Wert, so nennt man z einen singulären Wert.

!ACHTUNG!. Im Fall $z \notin F(U)$ ist z ein regulärer Wert! Denn der Allquantor geht über die leere Menge, die Aussage (6.4) ist also wahr. Wenn z ein regulärer Wert von F, so impliziert dies nicht, dass z ein Wert von F ist.

BEISPIELE 6.5. (1) Ist $F : \mathbb{R}^n \to \mathbb{R}^k$ eine lineare Funktion. Dann gilt:

$$\begin{aligned} & F \text{ ist surjektiv} \\ \iff \ & 0 \text{ ist ein regulärer Wert} \\ \iff \ & \text{Alle } z \in \mathbb{R}^k \text{ sind reguläre Werte} \end{aligned}$$

(2) Wir betrachten (ähnlich wie in Beispiel 5.11) die Funktion

$$F : \mathbb{R}^{n+1} \to \mathbb{R}, \qquad F(x) := \langle x, x \rangle.$$

$$F'(x) = 2x^T.$$

Dann ist 0 kein regulärer Wert, aber jedes $z \in \mathbb{R} \setminus \{0\}$ ist ein regulärer Wert. Für $z < 0$ ist $F^{-1}(\{z\}) = \emptyset$ und dann ist z ein regulärer Wert, siehe die obige Achtung-Bemerkung. Für $z > 0$ gilt $\|F'(x)\|_2 = 2\sqrt{z}$ falls $F(x) = z$. In diesem Fall hat also $F'(x)$ Rang 1.

$$S^n := F^{-1}(\{1\}) = \{x \in \mathbb{R}^{n+1} \mid \|x\|_2 = 1\}.$$

DEFINITION 6.6. Sei

$$\mathcal{S}_k := \Big\{\underbrace{\text{Permutationen}}_{=\text{Bijektionen}}\{1,2,\ldots,k\} \to \{1,2,\ldots,k\}\Big\}$$

Zu $\sigma \in \mathcal{S}_k$ sei L_σ die lineare Abbildung $\mathbb{R}^k \to \mathbb{R}^k$ mit $e_j \mapsto e_{\sigma(j)}$.

SATZ 6.7 (Satz vom regulären Wert)**.** Sei M eine Teilmenge von $\mathbb{R}^k$, $\ell \geq 1$. Äquivalent sind:

(1) M ist n-dimensionale Untermannigfaltigkeit von $\mathbb{R}^k$ mit Regularität C^ℓ
(2) Zu jedem $p \in M$ existiert eine offene Umgebung U in $\mathbb{R}^k$ und eine Funktion $F \in C^\ell(U, \mathbb{R}^{k-n})$, so dass 0 regulärer Wert von F ist und $F^{-1}(\{0\}) = U \cap M$.
(3) Zu jedem $p \in M$ existiert eine offene Umgebung U in $\mathbb{R}^k$ und eine Funktion $F \in C^\ell(U, \mathbb{R}^{k-n})$, ein $z \in \mathbb{R}^{k-n}$, so dass z regulärer Wert von F ist und $F^{-1}(\{z\}) = U \cap M$.
(4) Lokal ist M nach Permutation der Komponenten von $\mathbb{R}^k$ der Graph einer Funktion, genauer: zu jedem $p \in M$ existiert ein $\sigma \in \mathcal{S}_k$, eine offene Teilmenge V von $\mathbb{R}^n$, eine offene Teilmenge W von $\mathbb{R}^{k-n}$ und eine C^ℓ-Funktion $f : V \to W$, so dass

$$(V \times W) \cap L_\sigma(M) = \text{Graph}(f).$$

BEISPIELE 6.8.

(1) S^n ist eine Untermannigfaltigkeit von $\mathbb{R}^{n+1}$ mit Regularität C^∞, da 1 ein regulärer Wert der Funktion F aus dem letzten Beispiel ist.
(2) Sei wieder W offen in $\mathbb{R}^n$, $f \in C^\ell(W, \mathbb{R}^m)$, Graph$f := \{(x, f(x)) \mid x \in W\}$. Dann ist

$$F : \underbrace{W \times \mathbb{R}^m}_{U:=} \to \mathbb{R}^m, \quad (x, y) \mapsto y - f(x)$$

ist eine ℓ-mal stetig differenzierbare Abbildung. Wir rechnen

$$F'(x,y) = (-f'(x) \quad \mathbb{1}_m)$$

und deswegen ist jedes $z \in \mathbb{R}^m$ ein regulärer Wert von F, insbesondere ist 0 regulärer Wert. Dies liefert einen neuen Beweis, dass Graph$f$ eine C^ℓ-Untermannigfaltigkeit von $\mathbb{R}^{n+m}$ ist.

Bemerkung 6.9. In den obigen Beispielen ist die Untermannigfaltigkeit eine Teilmenge von U. Es gibt aber auch n-dimensionale C^∞-Untermannigfaltigkeiten $M \subset \mathbb{R}^k$, für die es kein $F \in C^1(U, \mathbb{R}^{k-n})$ wie oben mit $M \subset U$ gibt. Ein Beispiel hierfür ist das Möbius-Band , das wir später näher betrachten wollen. Um solche Untermannigfaltigkeiten als Nullstellen regulärer Werte darzustellen, benötigt man mehrere offene Mengen $U \subset \mathbb{R}^k$ und Abbildungen $F : U \to \mathbb{R}^{k-n}$ wie oben.

Bemerkung 6.10. (Wiederholung aus der Linearen Algebra) Sei $A \in \mathbb{R}^{m\times k}$ und $m \leq k$, z.B. obiges $F'(x)$. Dann gilt

A hat Rang m

$\iff$ A hat maximalen Rang

$\iff$ Die Abbildung $\mathbb{R}^k \to \mathbb{R}^m, \quad x \mapsto Ax$ ist surjektiv

$\iff$ Die Abbildung $\mathbb{R}^m \to \mathbb{R}^k, \quad x \mapsto A^T x$ ist injektiv

$\iff$ Nach Umordnen der Spalten von A, bilden die letzten m Spalten eine invertierbare Matrix

Beispiel 6.11.

$$A := \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 2 \end{pmatrix}$$

Die letzten beiden Spalten sind linear abhängig. Nach Vertauschen der ersten und dritten Spalte haben wir

$$\begin{pmatrix} 4 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix}.$$

Wegen

$$\det \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} = 5 \neq 0$$

sehen wir, dass A Rang 2 besitzt.

Beweis von Satz 6.7.

„(1) $\Longrightarrow$ (2)“: Zu $p \in M$ sei eine Untermannigfaltigkeitskarte φ gegeben. Sei $\pi : \mathbb{R}^k \to \mathbb{R}^{k-n}$ die Projektion auf die letzten $k-n$ Komponenten. Dann ist $F := \pi \circ \varphi : U \to \mathbb{R}^{k-n}$ eine C^ℓ-Abbildung. Weiter ist 0 ist ein regulärer Wert, denn $F'(p) = \pi'(\varphi(p))\varphi'(p)$, der Rang von $\pi'(\varphi(p))$ ist $k-n$, und da $\varphi'(p)$ invertierbar ist, bleibt der Rang einer $(k-n)\times k$-Matrix unverändert, wenn wir sie von rechts mit $\varphi'(p)$ multiplizieren. Außerdem gilt $F^{-1}(\{0\}) = \varphi^{-1}(V \cap (\mathbb{R}^n \times \{0\})) = U \cap M$.

„(2) $\Longrightarrow$ (3)“: Offensichtlich.

„(3) $\Longrightarrow$ (4)“: Zu $p \in M$ sei eine offene Umgebung U in $\mathbb{R}^k$ gegeben, eine Abbildung $F \in C^\ell(U, \mathbb{R}^{k-n})$ und ein $z \in \mathbb{R}^{k-n}$, so dass z regulärer Wert ist und $F^{-1}(\{z\}) = U \cap M$. Insbesondere $F(p) = z$, also hat $F'(p)$ Rang $m := k-n$. Nach Umordnen der Komponenten von $\mathbb{R}^k$, können wir annehmen, dass die hintersten m Spaltenvektoren in $F'(p) \in \mathbb{R}^{m\times k}$ eine invertierbare Matrix bilden. Somit sind die Voraussetzung erfüllt, um den Satz 5.2 über implizite Funktionen im Punkt $(\hat{x}, \hat{y}) := p$ und mit $\tilde{F}(x, y) := F(x, y) - z$ statt F anzuwenden. Die damit erhaltenen V, W und f erfüllen (4).

„(4) $\Longrightarrow$ (1)“: Folgt direkt aus Beispiele 6.2 (2) „Graphen sind Untermannigfaltigkeiten“ und (7) „Die Untermannigfaltigkeits-Eigenschaft ist lokal“. □

BEISPIEL 6.12 (Torus). Für ein $R \in (0,1)$ betrachten wir

$$\mathbb{T}_R := \left\{ (x,y,z) \in \mathbb{R}^3 \mid \left(\sqrt{x^2+y^2} - 1 \right)^2 + z^2 = R^2 \right\} \subset \mathbb{R}^3.$$

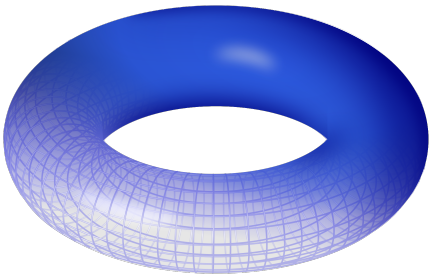


Abbildung 9: Eine Zentralprojektion von $\mathbb{T}_R$ für $R \approx 0,2$.
Quelle: Wikipedia `https://commons.wikimedia.org/wiki/File:Torus.svg`
GNU Free Documentation License, Version 1.2

ÜBUNG 6.13. Zeigen Sie, dass $\mathbb{T}_R$ eine zwei-dimensionale Untermannigfaltigkeit von $\mathbb{R}^3$ ist.

Lösung (von Nicolas Ginoux): Die Teilmenge $\mathbb{T}_R$ ist ein sogenannter *Drehtorus* im $\mathbb{R}^3$: er entsteht durch Drehung eines Kreises in der x-z-Ebene mit Mittelpunkt $(x,z) = (1,0)$ und mit Radius R um die z-Achse.
Sei

$$\begin{aligned} U &:= \left\{ (x,y,z) \in \mathbb{R}^3 \mid (x,y) \neq (0,0) \text{ und } (\|(x,y)\|_2 = 1 \implies z \neq 0) \right\} \\ &= \mathbb{R}^3 \setminus \left\{ \mathbb{R} \cdot e_3 \cup (\mathbb{S}^1 \times \{0\}) \right\} \subset \mathbb{R}^3 \end{aligned}$$

und $F\colon U \to \mathbb{R}$, $(x,y,z) \mapsto \left(\sqrt{x^2+y^2}-1\right)^2 + z^2 - R^2$. Die Teilmenge U ist offen im $\mathbb{R}^3$ mit $\mathbb{T}_R \subset U$: wegen $0 < R < 1$ hat die Gleichung $1 + z^2 = R^2$ keine (reelle) Lösung und damit gilt $(x,y) \neq (0,0)$ für alle $(x,y) \in \mathbb{T}_R$; außerdem gilt, im Fall $(x,y,z) \in \mathbb{T}_R$ mit $\|(x,y)\|_2 = 1$, die Gleichung $z^2 = R^2$, insbesondere $z \neq 0$.

Nun ist $\mathbb{T}_R = F^{-1}(\{0\})$. Um zu beweisen, dass $\mathbb{T}_R$ eine zwei-dimensionale Untermannigfaltigkeit von $\mathbb{R}^3$ ist, reicht es daher, zu beweisen, dass 0 ein regulärer Wert der Funktion F ist.

Da $(x,y) \mapsto \sqrt{x^2+y^2}$ auf $\mathbb{R}^2 \smallsetminus \{(0,0)\}$ unendlich oft differenzierbar ist, ist die Funktion $F\colon U \to \mathbb{R}$ als Verknüpfung unendlich oft differenzierbarer Abbildungen unendlich oft differenzierbar. Für alle $(x,y,z) \in U$ ist die Matrix von $F'(x,y,z)$ in den kanonischen Basen gegeben durch

$$2\begin{pmatrix} \frac{x(\sqrt{x^2+y^2}-1)}{\sqrt{x^2+y^2}} & \frac{y(\sqrt{x^2+y^2}-1)}{\sqrt{x^2+y^2}} & z \end{pmatrix}.$$

Gilt $\sqrt{x^2+y^2} = 1$, so verschwinden die beiden ersten Komponenten von $F'(x,y,z)$; ist nun $(x,y,z) \in U$, so gilt dann $z \neq 0$ und damit ist $F'(x,y,z) \neq 0$. Gilt $\sqrt{x^2+y^2} \neq 1$, so können wegen $(x,y) \neq (0,0)$ nicht beide erste und zweite Komponenten von $F'(x,y,z)$ verschwinden, d.h., es gilt auch in diesem Fall $F'(x,y,z) \neq 0$. Insgesamt haben wir bewiesen, dass für jedes $(x,y,z) \in U$ die Abbildung $F'(x,y,z)$ nicht verschwindet und ist damit – weil $F'(x,y,z)$ seine Werte im 1-dimensionalen Vektorraum $\mathbb{R}$ hat – surjektiv. Daraus folgt, dass 0 ein regulärer Wert der Funktion F ist. Dies liefert das Ergebnis.

Bemerkung 6.14. Man zeigt leicht, dass

$$\mathbb{T}_R := \{T_R(\alpha,\beta) \mid \alpha,\beta \in \mathbb{R}\},$$

wobei $T_R : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^3$,

$$T_R(\alpha, \beta) := \begin{pmatrix} (1 + R\cos\alpha)\cos\beta \\ (1 + R\cos\alpha)\sin\beta \\ R\sin\alpha \end{pmatrix}.$$

Mi 26.6.

6.3. Der Immersionssatz.

DEFINITION 6.15. Sei W eine offene Teilmenge von $\mathbb{R}^n$ und $\Psi : W \to \mathbb{R}^k$. Man nennt Ψ eine Immersion , wenn Ψ stetig differenzierbar ist und wenn für alle $p \in W$ die Matrix $f'(p)$ Rang n hat (d.h. die zu $f'(p)$ gehörende lineare Abbildung $\mathbb{R}^n \to \mathbb{R}^k$, $x \mapsto f'(p) \cdot x$ ist injektiv). Man nennt Ψ eine C^ℓ-Immersion[38], wenn zusätzlich $f \in C^\ell$, $\ell \geq 1$.

Insbesondere ist dann $k \geq n$, falls $W \neq \emptyset$. Die Eigenschaften „Ψ ist injektiv“ und „Ψ ist eine Immersion“ sind unabhängig voneinander.

BEISPIELE 6.16.

(1) Die Abbildung $\Psi : \mathbb{R} \to \mathbb{R}^2$, $\Psi(t) := (\cos t, \sin t)^T$ erfüllt $\Psi'(t) := (-\sin t, \cos t)^T \neq 0$. Also ist Ψ eine Immersion, aber nicht injektiv.
(2) $\Psi : \mathbb{R} \to \mathbb{R}$, $\Psi(t) := t^3$ ist ein Homöomorphismus, also insbesondere injektiv, aber keine Immersion, denn $\Psi'(0) = 0$.
(3) Die in Beispiel 6.14 definierte Funktion $T_R : \mathbb{R}^2 \to \mathbb{R}^3$ ist für $R \in (0, 1)$ eine C^∞-Immersion (Beweis: kleinere Rechnung!).

SATZ 6.17 (Immersionssatz)**.** Sei W eine offene Teilmenge von $\mathbb{R}^n$ und $\Psi : W \to \mathbb{R}^k$ eine C^ℓ-Immersion, $\ell \geq 1$. Dann gibt es zu jedem $p \in W$ eine offene Umgebung $U \subset W$,

[38] Also: C^1-Immersion=Immersion

so dass $\Psi(U)$ eine n-dimensionale C^ℓ-Untermannigfaltigkeit von $\mathbb{R}^k$ ist und so dass $\Psi|_U : U \to \Psi(U)$ ein Homöomorphismus ist.

Beweis des Immersionssatzes. Sei $\Psi\colon W \longrightarrow \mathbb{R}^k$ eine Immersion und $p \in W$. Da $\Psi'(x_0)$ Rang n hat, sind die Vektoren $\frac{\partial\Psi}{\partial x_1}(p), \ldots, \frac{\partial\Psi}{\partial x_n}(p)$ linear unabhängig in $\mathbb{R}^k$. Deswegen existieren $k-n$ Vektoren $v_{n+1}, \ldots, v_k$ in $\mathbb{R}^k$ so, dass $\left(\frac{\partial\Psi}{\partial x_1}(p), \ldots, \frac{\partial\Psi}{\partial x_n}(p), v_{n+1}, \ldots, v_k\right)$ eine Basis von $\mathbb{R}^k$ ist. Betrachte nun die Abbildung

$$\Phi : W \times \mathbb{R}^{k-n} \longrightarrow \mathbb{R}^k, (x, \alpha) \longmapsto \Psi(x) + \sum_{j=n+1}^{k} \alpha_j v_j,$$

wobei $x \in W$ und $\alpha = (\alpha_{n+1}, \ldots, \alpha_k) \in \mathbb{R}^{k-n}$. Offensichtlich ist $\Phi \in C^\ell$. Wir rechnen:

$$\Phi'(p,0) = \left(\frac{\partial\Psi}{\partial x_1}(p) \quad \frac{\partial\Psi}{\partial x_2}(p) \ \cdots \ \frac{\partial\Psi}{\partial x_n}(p) \quad v_{n+1} \ \cdots \ v_k\right).$$

Die Spalten bilden – wie oben konstruiert – eine Basis, d.h. $\Phi'(p,0)$ ist eine invertierbare Matrix.

Nach dem lokalen Umkehrsatz existieren eine offene Umgebung V_1 von $(p,0)$ in $W \times \mathbb{R}^{k-n}$ und eine offene Umgebung V_2 von $\Phi(p,0) = \Psi(p)$ in $\mathbb{R}^k$, so dass $\Phi|_{V_1} : V_1 \longrightarrow V_2$ ein C^ℓ-Diffeomorphismus ist.

Durch Verkleinern von V_1 und V_2 können wir erreichen, dass V_1 von der Form $V_1 = U \times U'$ ist (mit $p \in U \subset W$, $0 \in U'$).

Dann ist $\varphi := (\Phi|_{U\times U'})^{-1} : V_2 \to U \times U'$ eine Untermannigfaltigkeitskarte von $\Psi(U) = \Phi(U \times \{0\})$. Da $\Phi|_{U\times U'} : U \times U' \to V_2$ ein Homöomorphismus ist, ist auch $\Psi|_U = (\Phi|_{U\times U'})\,\big|_{U\times\{0\}} : U \times \{0\} \cong U \to \Psi(U)$ ein Homöomorphismus.

Zeichnung von $\Psi(W)$ als Kurve in $\mathbb{R}^2$ und von $\Phi(U \times U') = V_2$ mit φ Abbildungspfeil auf $U \times U'$.

□

ZUSATZ 6.18. Die Voraussetzungen seien wie im vorangehenden Satz. Außerdem nehmen wir an, $\Psi : W \to \Psi(W)$ sei ein Homöomorphismus. Sei $V \subset \mathbb{R}^r$ eine offene Teilmenge und $h : V \to \mathbb{R}^k$ eine C^ℓ-Abbildung mit $h(V) \subset \Psi(W)$. Dann ist $\Psi^{-1} \circ h : V \to W$ ebenfalls in C^ℓ.

Zeichnung zur Veranschaulichung des Zusatzes

Beweis. Wir zeigen die stetige Differenzierbarkeit auf einer Umgebung eines beliebigen Punkts $x \in V$. Wir definieren φ wieder wie im Beweis des Satzes auf einer kleinen offenen Umgebung V_2 von $h(x) = \Psi(p)$, U eine offene Umgebung von p wie oben. Wir setzen $V_0 := h^{-1}(V_2)$, dies ist eine offene Umgebung von x. Es gilt für $y \in V_0$:

$$\varphi \circ h(y) = \Psi^{-1} \circ h\,.$$

Die linke Seite ist ℓ-mal stetig differenzierbar als Funktion in y also auch die rechte, was zu zeigen war. □

BEISPIELE 6.19. Es gibt injektive Immersionen $\Psi : W \to \mathbb{R}^k$, so dass $\Psi : W \to \Psi(W)$ eine unstetige Umkehrabbildung besitzt:

(1) Die Abbildung $\Psi : (0, 2\pi) \to \mathbb{R}^2$, $\Psi(t) := (\sin t, \sin(2t))^T$ erfüllt $\Psi'(t) := (\cos t, 2\cos(2t))^T$. Aus $\cos t = 0$ folgt $t = \pi(k + \frac{1}{2})$ für $k \in \mathbb{Z}$, also $t \in \{\pi/2, 3\pi/2\}$. Dann aber gilt $\cos(2t) \neq 0$. Es folgt $\Psi'(t) \neq 0$ für alle $t \in (0, 2\pi)$, das heißt Ψ ist eine Immersion. Man sieht auch leicht, dass Ψ injektiv ist. Allerdings ist Ψ kein Homöomorphismus

von $(0, 2\pi)$ auf $\Psi((0, 2\pi))$. Denn für eine Folge von $t_i \in (0, 2\pi)$ mit $\lim_{i\to\infty} t_i = 0$ gilt $p_i := \Psi(t_i) \to 0 = (0, 0) = \Psi(\pi) \in \mathbb{R}^2$. Somit gilt $p_i \to 0 \in \mathbb{R}^2$, aber $\Psi^{-1}(p_i) = t_i \not\to \pi = \Psi^{-1}(0)$. Somit ist Ψ^{-1} nicht (folgen-)stetig.

Bild von $\Psi((0, 2\pi))$

(2) Sei wieder T_R wie in Beispiel 6.14 definiert. Zu einer Zahl $\rho \in \mathbb{R}$ definieren wir

$$\Psi_\rho : \mathbb{R} \to \mathbb{R}^3, \quad \Psi_\rho(t) := T_R(2\pi\rho t, 2\pi t).$$

Fall 1: $\rho = p/q \in \mathbb{Q}$, seien p,$q \in \mathbb{Z}$ teilerfremd. Dann ist Ψ_ρ periodisch: $\Psi_\rho(t+q) = \Psi_\rho(t)$. $\Psi(\mathbb{R})$ ist eine geschlossene Kurve[39], die sich in einem Durchlauf p-mal um die „Seele“ $S^1 \times \{0\} \subset \mathbb{R}^3$ des Torus windet und q-mal um die z-Achse. Man nennt $\Psi(\mathbb{R})$ einen (p, q)-Torus-Knoten. [40]

Fall 2: $\rho \in \mathbb{R} \setminus \mathbb{Q}$. Dann ist Ψ_ρ eine injektive Immersion. Ihr Bild $\Psi_\rho(\mathbb{R})$ liegt dicht in $\mathbb{T}_R$, aber Ψ_ρ ist kein Hoöomorphismus auf sein Bild. Denn sind 1, ρ in τ rational linear unabhängige reelle Zahlen, dann ist jedes $x \in \mathbb{T}_R$ Häufungspunkt der Folge $a_k := \psi(k\tau)$. Eine Teilfolge konvergiert dann gegen ein $x \in \Psi(\mathbb{R})$. Deswegen ist Ψ^{-1} nicht folgenstetig.

6.4. Einbettungen von Untermannigfaltigkeiten. In diesem Abschnitt sei $\pi_x : \mathbb{R}^k \to \mathbb{R}^n$ die Projektion auf die ersten n Komponenten und $\pi_y : \mathbb{R}^k \to \mathbb{R}^{k-n}$ die Projektion auf die letzten $k - n$ Komponenten.

---

[39] genauer: es ist eine zusammenhängende kompakte 1-dimensionale Untermannigfaltigkeit von $\mathbb{R}^3$

[40] Torusknoten sind derzeit ein beliebtes Kinderspielzeug und sind unter dem Namen „Flow ring“ im Handel erhältlich und auch auf Video-Plattformen findet man alles mögliche.

Lemma 6.20. Sei M eine n-dimensionale[41] Untermannigfaltigkeit von $\mathbb{R}^k$ mit einer Untermannigfaltigkeitskarte $\varphi : U \to V$, $p \in U \cap M$, und sei $\Psi : W \to \mathbb{R}^k$ eine injektive Immersion mit $\Psi(W) \subset M$ und $p \in \Psi(W)$. Sei $T_p \subset \mathbb{R}^k$ das Bild von $\Psi'(\Psi^{-1}(p))$. Dann gilt $\varphi'(p) \cdot T_p = \mathbb{R}^n \times \{0\}$.

Das Lemma gilt in leicht veränderter Form, wenn Ψ nicht mehr injektiv ist, sondern nur noch eine Immersion. Wir wollen dies aber nicht näher betrachten, da man jede Immersion durch Einschränkung auf eine genügend kleine Umgebung eines gegebenen Punktes injektiv machen kann (Immersionsatz).

Zeichnung zu obigem Satz

Beweis. Setze $W_0 := \Psi^{-1}(U) \subset W$ und $x_0 := \Psi^{-1}(p)$. Man prüft leicht, dass $\pi_y \circ \varphi \circ \Psi|_{W_0} : W_0 \to \mathbb{R}^{k-n}$ alles auf 0 abbildet. Nach Kettenregel ist dann

$$0 = \underbrace{\pi_y'(\varphi(p))}_{= (0\ \mathbb{1}_{k-n})} \cdot \varphi'(p) \cdot \Psi'(x_0)\,,$$

das heißt, das Bild der Matrix $A := \varphi'(p) \cdot \Psi'(x_0)$ liegt in $\mathbb{R}^n \times \{0\}$, und da $\Psi'(x_0)$ Rang n besitzt und $\varphi'(p)$ invertierbar ist, wissen wir, dass das Bild von A gleich $\mathbb{R}^n \times \{0\}$ ist. Da das Bild von A gleich $\varphi'(p) \cdot T_p$ ist, folgt die Behauptung. □ **Fr 28.6.**

Satz 6.21. Sei $\Psi : W \to \mathbb{R}^k$ eine injektive C^ℓ-Immersion, $\ell \geq 1$. Dann ist $\Psi(W)$ eine C^ℓ-Untermannigfaltigkeit von $\mathbb{R}^k$ genau dann, wenn $\Psi : W \to \Psi(W)$ ein Homöomorphismus ist.

[41] Das Wort „n-dimensional“ habe ich wahrscheinlch in der Vorlesung leider vergessen, ist aber sehr wichtig.

Zeichnung zu diesem Satz

Beweis. Setze $M := \Psi(W)$.

„$\Longleftarrow$“: Wir wollen zeigen, dass M eine C^ℓ-Untermannigfaltigkeit von $\mathbb{R}^k$ ist. Sei dazu $p \in M$ und $p = \Psi(x_0)$, $x_0 \in W$. Nach dem Immersionssatz 6.17 gibt es eine offene Umgebung W' von x_0 in W, so dass $M' := \Psi(W')$ eine C^ℓ-Untermannigfaltigkeit von $\mathbb{R}^k$ ist. Da $\Psi : W \to \Psi(W)$ ein Homöomorphismus ist und da W' offen in W ist, ist $M' = \Psi(W')$ eine offene Teilmenge von $M = \Psi(W)$, d.h. es existiert eine in $\mathbb{R}^n$ offene Menge U_1 mit $M' = U_1 \cap M$.

Sei $\varphi_2 : U_2 \to V_2$ eine Untermannigfaltigkeitskarte von M' mit $p \in U_2$. Setze $U := U_1 \cap U_2 \ni p$, $V := \varphi_2(U)$ und $\varphi := \varphi_2\big|_U$. Dann sind U und V offen in $\mathbb{R}^k$, und $\varphi : U \to V$ ist ein C^ℓ-Diffeomorphismus, für den gilt:

$$\varphi(U \cap M) = \varphi_2(U \cap M') = V \cap \varphi_2(M') = \underbrace{V \cap V_2}_{=V} \cap (\mathbb{R}^n \times \{0\}).$$

Also ist $\varphi : U \to V$ eine Untermannigfaltigkeitskarte von M mit $p \in U$.

„$\Longrightarrow$“: Wir zeigen, dass $\Psi^{-1} : M \to W$ folgenstetig ist. Sei hierzu eine Folge $(p_i)_{i\in\mathbb{N}}$ in M gegeben mit $\lim_{i\to\infty} p_i = p \in M$. Da nach Voraussetzung M eine C^ℓ-Untermannigfaltigkeit ist, können wir eine Untermannigfaltigkeitskarte $\varphi : U \to V$ mit $p \in U$ wählen. Es liegen fast alle p_i in $U \cap M$ und deswegen sind fast alle $q_i := \pi_x \circ \varphi(p_i) \in \pi_x(V) \subset \mathbb{R}^n$ wohldefiniert und konvergieren gegen $q := \pi_x \circ \varphi(p) \in \mathbb{R}^n$. Insbesondere $\varphi(p_i) = (q_i, 0)$.

Setze wieder $W_0 := \Psi^{-1}(U) \subset W$ und $x_0 := \Psi^{-1}(p)$. Wie im vorangehenden Lemma gezeigt, ist das Bild von $\varphi'(p)\Psi'(x_0)$ gleich $\mathbb{R}^n \times \{0\}$.

Somit ist

$$(\pi_x \circ \varphi \circ \Psi|_{W_0})'(x_0) = \underbrace{\pi_y'(\varphi(p))}_{= (\mathbb{1}_n \; 0)} \cdot \varphi'(p) \cdot \Psi'(x_0)$$

eine invertierbare Matrix in $\mathbb{R}^{n\times n}$. Also ist nach dem lokalen Umkehrsatz $\rho := \pi_x \circ \varphi \circ \Psi|_{W_0}$ lokal um x_0 umkehrbar und $x_i := \rho^{-1}(q_i) \to x_0 := \rho^{-1}(q)$. Wegen $\Psi(x_i) \in M$ haben wir $\varphi \circ \Psi(x_i) = (q_i, 0)$ und aus der Injektivität von φ folgt $\Psi(x_i) = p_i$, also $\Psi^{-1}(p_i) = x_i \to x_0 = \Psi^{-1}(p)$. Wir haben also die Folgenstetigkeit und somit die Stetigkeit von $\Psi^{-1} : M \to W$ gezeigt. □

Beispiel 6.22 (Fortsetzung von Beispiel 6.19 (1)). Sei wieder $\Psi : (0, 2\pi) \to \mathbb{R}^2$, $\Psi(t) := (\sin t, \sin(2t))^T$. Diese Abbildung ist eine injektive C^∞-Immersion, aber kein Homöomorphismus von $(0, 2\pi)$ auf $\Psi\big((0, 2\pi)\big)$: das haben wir oben bereits alles gezeigt. Das Bild $M := \Psi\big((0, 2\pi)\big)$ ist keine Untermannigfaltigkeit, da es keine Untermannigfaltigkeitskarte von M gibt, die $0 = \Psi(\pi) \in \mathbb{R}^2$ enthält. Die Nicht-Existenz einer Untermannigfaltigkeitskarte kann man auch elementar beweisen, bequemer geht es aber mit dem obigen Satz.

Nochmals dieselbe Zeichnung vom Bild von Ψ.

Definition 6.23. Eine injektive Immersion, die ein Homöomorxphismus auf ihr Bild ist, nennt man eine Einbettung einer Untermannigfaltigkeit.

6.5. Lokale Parametrisierungen von Untermannigfaltigkeiten.

Definition 6.24. Sei $M \subset \mathbb{R}^k$ eine Teilmenge. Eine n-dimensionale lokale C^ℓ-Parametrisierung von M ist eine C^ℓ-Immersion $\Psi : W \to \mathbb{R}^k$ mit den folgenden Eigenschaften:

(a) $W \subset \mathbb{R}^n$ ist offen,

(b) $\Psi(W)$ ist offen in M, das heißt $\Psi(W) = U \cap M$ für eine offene Teilmenge U von $\mathbb{R}^k$,

(c) $\Psi : W \to \Psi(W)$ ist ein Homöomorphismus.

LEMMA 6.25. Ist $\Psi : W \to \mathbb{R}^k$ eine lokale C^ℓ-Parametrisierung von $M \subset \mathbb{R}^k$ wie in obiger Definition und W_0 eine offene Teilmenge von W. Dann ist auch $\Psi|_{W_0} : W_0 \to \mathbb{R}^k$ ebenfalls eine lokale C^ℓ-Parametrisierung von M.

Beweis. Wähle zu Ψ ein U wie in (b).

Offensichtlich ist $\hat{\Psi} := \Psi|_{W_0}$ eine C^ℓ-Immersion, die (a) und (c) mit $\hat{\Psi}$ statt Ψ erfüllt.

Zu (b) für $\hat{\Psi}$: Da $\Psi : W \to \Psi(W)$ ein Homöomorphismus ist, ist $\Psi(W_0)$ offen in $\Psi(W)$.[42] Es gibt also eine offene Teilmenge $\tilde{U}$ von $\mathbb{R}^k$ mit $\Psi(W_0) = \Psi(W) \cap \tilde{U}$. Es folgt $\Psi(W_0) = \underbrace{(U \cap \tilde{U})}_{\text{offen in } \mathbb{R}^k} \cap M$. □

PROPOSITION 6.26. Sei $M \subset \mathbb{R}^k$ eine Teilmenge. Dann sind äquivalent:
(a) M ist eine n-dimensionale C^ℓ-Untermannigfaltigkeit.
(b) Für alle $p \in M$ existiert eine n-dimensionale lokale C^ℓ-Parametrisierung von M der Regularität C^ℓ, deren Bild p enthält.

[42] In der Notation für Bilder und Urbilder, die wir in der Analysis I benutzt haben, wird dies etwas klarer: es gilt $\Psi_\#(W_0) = (\Psi^{-1})^\#(W_0)$ und diese Menge ist als Urbild der offenen Menge W_0 bezüglich der stetigen Abbildung $\Psi^{-1} : \Psi(W) \to W$ wieder offen.

Beweis.
„(a) $\Longrightarrow$ (b)“ Sei $p \in M$ beliebig. Sei $\varphi : U \to V$ eine Untermannigfaltigkeitskarte von M mit $p \in U$. Definiere $W := V \cap (\mathbb{R}^n \times \{0\})$ und $\Psi := (\varphi^{-1})|_W \colon W \to \mathbb{R}^k$. Dann ist W offen in $\mathbb{R}^n$, Ψ ist C^ℓ (da $\varphi^{-1} \in C^\ell$) mit $\Psi'(x) = (\varphi^{-1})'(x)_{|\mathbb{R}^n \times \{0\}}$ injektiv $\forall x \in W$, es gilt $\Psi(W) = \varphi^{-1}(W) = U \cap M$ und $\Psi : W \to \Psi(W)$ ist die Einschränkung eines Homöomorphismus, ist also selbst ein Homöomorphismus.

„(b) $\Longrightarrow$ (a)“ Sei $p \in M$ beliebig. Sei $\Psi \colon W \to \mathbb{R}^k$ eine lokale C^ℓ-Parametrisierung von M mit $p = \Psi(x_0) \in \Psi(W)$. Nach dem Immersionssatz 6.17 gibt es eine offene Umgebung W_0 von x_0, so dass $\Psi(W_0) \ni p$ eine Untermannigfaltigkeit ist.

Da $\Psi^{-1} : \Psi(M) \to M$ stetig ist, ist $\Psi(W_0)$ offen in $\Psi(W)$ und damit wegen (b) von Definition 6.24 offen in M. Es gibt also eine in $\mathbb{R}^k$ offene Menge $U \ni p$ mit $\Psi(W_0) = U \cap M$. Wir haben also zu jedem $p \in M$ eine in $\mathbb{R}^k$ offene Umgebung U gefunden, so dass $U \cap M$ eine Untermannigfaltigkeit von $\mathbb{R}^k$ ist. Da die Untermannigfaltigkeitseigenschaft eine lokale Eigenschaft ist (Beispiele 6.2 (7)), folgt daraus, dass M eine C^ℓ-Untermannigfaltigkeit ist. □

Beispiele 6.27.

(1) Sei $U \subset \mathbb{R}^n$ offen und $f \in C^\ell(U, \mathbb{R}^m)$. Betrachte

$$\mathrm{Graph} f := M_f := \{(x, f(x)) \,|\, x \in U\} \subset \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}.$$

Dann ist die Abbildung $\Psi : U \longrightarrow \mathbb{R}^{n+m}$, $x \longmapsto (x, f(x))^T$ eine n-dimensionale globale C^ℓ-Parametrisierung von M_f: die Abbildung Ψ ist C^ℓ, es gilt $\Psi'(x) = \begin{pmatrix} \mathbb{1}_n \\ f'(x) \end{pmatrix}$, $\forall x \in U$, wobei $\mathbb{1}_n$ die $n \times n$ Einheitsmatrix bezeichnet – insbesondere gilt $\mathrm{rg}(\Psi'(x)) =$

n – und $\Psi(U) = M_f$, wobei $\Psi\colon U \longrightarrow M_f$ ein Homöomorphismus ist mit Umkehrabbildung $M_f \longrightarrow U$, $(x, f(x)) \longmapsto x$ (Projektion auf die n ersten Koordinaten).

(2) Sei $M := S^n \subset \mathbb{R}^{n+1}$. Sei $U := \{x \in \mathbb{R}^n\colon \|x\|_2 < 1\} \subset \mathbb{R}^n$. Definiere $\Psi_+\colon U \longrightarrow \mathbb{R}^{n+1}$, $x \longmapsto (x, \sqrt{1 - \|x\|_2^2})^T$ sowie $\Psi_-\colon U \longrightarrow \mathbb{R}^{n+1}$, $\Psi_-(x) := (x, -\sqrt{1 - \|x\|_2^2})^T$. Dann sind Ψ_+ und Ψ_- n-dimensionale lokale C^∞-Parametrisierungen von S^n. Es gilt

$$\Psi_+(U) = S^n \cap \left(\mathbb{R}^n \times (0, \infty)\right) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n+1} \end{pmatrix} \in S^n \;\middle|\; x_{n+1} > 0 \right\}.$$

Überdeckung von $S^1 \subset \mathbb{R}^2$ durch die Bilder von 4 Parametrisierungen

Wir vertauschen nun die Komponenten von $\mathbb{R}^{n+1}$ so, dass die ℓ-te Komponente zur letzten Komponente wird. Wir erhalten dann zu jedem $\ell = 1, \ldots, n$, zwei weitere lokale Parametrisierungen $\Psi_\pm^\ell$, setze $\Psi_\pm^{n+1} := \Psi_\pm$. Es gilt

$$\Psi_\pm^\ell(U) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n+1} \end{pmatrix} \in S^n \;\middle|\; \pm x_\ell > 0 \right\},$$

und somit

$$S^n = \Psi_+^1(U) \cup \Psi_+^2(U) \cup \cdots \cup \Psi_+^{n+1}(U) \cup \Psi_-^1(U) \cup \Psi_-^2(U) \cup \cdots \cup \Psi_-^{n+1}(U).$$

Die Sphäre wird also von den Bildern von $2(n+1)$ lokalen Parametrisierungen überdeckt.

(3) Seien $a, b \in \mathbb{R}$ und $R \in (0,1)$. Sei $\mathbb{T}_R$ der Torus und T_R definiert wie in Beispiel 6.14. Dann ist die Abbildung

$$(a, a+2\pi) \times (b, b+2\pi) \to \mathbb{R}^3, \quad (\alpha, \beta) \mapsto T_R(\alpha, \beta)$$

eine lokale C^∞-Parametrisierung von $\mathbb{T}_R$.

BEMERKUNG 6.28. Sei $i : \mathbb{R}^n \to \mathbb{R}^k$ die Abbildung $x \mapsto (x, 0)$. Ist $\varphi : U \to V$ eine C^ℓ-Untermannigfaltigkeitskarte von M, so ist $\Psi := \varphi^{-1} \circ i : W \to \mathbb{R}^k$, $W := i^{-1}(V) \cong V \cap (\mathbb{R}^n \times \{0\})$ eine lokale C^ℓ-Parametrisierung: Die C^ℓ-Eigenschaft und die Offenheit von W ist klar, ebenso $\Psi(W) = V \cap M$ und somit die Offenheit von $\Psi(W)$ in M. Da i eine Immersion ist und φ ein Diffeomorphismus ist, ist auch Ψ eine Immersion. Die Umkehrabbildung von $\Psi : W \to \Psi(W)$ ist $\pi_x \circ \varphi|_{V \cap M}$ und diese Abbildung ist offensichtlich stetig.

Umgekehrt kann man zu jeder lokalen Parametrisierung Ψ eine Untermannigfaltigkeitskarte φ finden, so dass man Ψ wie oben aus φ erhält (ohne Beweis). Die lokale Existenz von solch einem φ haben wir im Beweis des Immersionssatzes 6.17 gezeigt, dort haben wir Ψ zu einer Funktion Φ fortgesetzt, die eine lokale Umkehrfunktion φ besitzt. Man kann mit größerem Aufwand solch ein $\varphi : U \to V$ konstruieren, so dass $\Psi(M) \subset U$.

Bild zu dieser Bemerkung

BEMERKUNG 6.29. Man kann sich nun natürlich fragen: sind die Eigenschaften (b) und (c) in Definition 6.24 wirklich notwendig? Wir haben im letzten Abschnitt gesehen, dass wir (c) benötigen, damit das Bild eine Untermannigfaltigkeit ist. Wenn (c) nicht erfüllt ist, haben wir Gegenbeispiele. Bedingung (b) ist ebenfalls notwendig. Die Menge $M :=$

$[0,1]\times\mathbb{R}^n$ ist keine Untermannigfaltigkeit von $\mathbb{R}^{n+1}$.[43] Wir definieren nun zu jedem $(x,y) \in [0,1] \times \mathbb{R}^n$ ist $\Psi_x : \mathbb{R}^n \to \mathbb{R}^{n+1}$, $\Psi_x(t) = (x,t)$ eine injektive Immersion, die (a) und (c), aber nicht (b) erfüllt. Offensichtlich ist (x,y) im Bild von Ψ_x.

6.6. Der Tangentialraum.

DEFINITION 6.30. Sei $M \subset \mathbb{R}^k$ eine n-dimensionale Untermannigfaltigkeit und $p \in M$ ein Punkt. Der Tangentialraum an M im Punkt p ist

$$T_pM := \{v \in \mathbb{R}^k \,|\, \exists\ \varepsilon > 0 \quad \text{und} \quad c\colon (-\varepsilon,\varepsilon) \longrightarrow \mathbb{R}^k \text{ ist } C^1\text{-Kurve mit } c((-\varepsilon,\varepsilon)) \subset M,$$
$$c(0) = p \quad \text{und} \quad c'(0) = v\} \subset \mathbb{R}^k.$$

PROPOSITION 6.31. Sei $M \subset \mathbb{R}^k$ eine Untermannigfaltigkeit und $p \in M$ ein Punkt.

(a) Sei $\Psi\colon W \longrightarrow \mathbb{R}^k$ eine n-dimensionale lokale C^1-Parametrisierung von M um p. Sei $x_0 := \Psi^{-1}(p) \in W$. Dann gilt

$$T_pM = \mathrm{Bild}\big(\Psi'(x_0)\big).$$

(b) Sei $\varphi : U \longrightarrow V$ eine Untermannigfaltigkeitskarte von M um p. Dann gilt

$$T_pM = \varphi'(p)^{-1} \cdot (\mathbb{R}^n \times \{0\})\,.$$

(c) Sei $U \subset \mathbb{R}^k$ offene Umgebung von p und $F\colon U \longrightarrow \mathbb{R}^{k-n}$ C^ℓ mit 0 als regulärem Wert und $F^{-1}(\{0\}) = U \cap M$. Dann gilt

$$T_pM = \ker(F'(p)).$$

[43] Begründung: Es ist keine $(n+1)$-dimensionale Untermannigfaltigkeit, da es nicht offen in $\mathbb{R}^{n+1}$ ist. Es ist keine 0-dimensionale Untermannigfaltigkeit, da es keine diskrete Teilmenge ist. Somit ist die Aussage im Fall $n = 0$ gezeigt. Für $n > 0$ kann man auch alle anderen Dimensionen ausschließen (etwas Arbeit!). Für unser Beispiel reicht es aber, den Fall $n = 0$ zu betrachten.

Insbesondere ist T_pM ein n-dimensionaler Untervektorraum von $\mathbb{R}^k$.

Beweis. Zu (a): Sei $v \in T_pM$. Dann existiert $c : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R}^k$ C^1 (für ein $\varepsilon > 0$) mit $c((-\varepsilon, \varepsilon)) \subset M$, $c(0) = p$ und $c'(0) = v$. Nach möglicher Verkleinerung von ε können wir annehmen, dass $c((-\varepsilon, \varepsilon)) \subset \Psi(W)$ gilt. Dann ist $\hat{c} := \Psi^{-1} \circ c$: $(-\varepsilon, \varepsilon) \longrightarrow \mathbb{R}^n$ stetig differenzierbar wegen Zusatz 6.18.

$$v = c'(0) = (\Psi \circ \hat{c})'(0) = \Psi'(\underbrace{\hat{c}(0)}_{x_0}) \cdot \hat{c}'(0)\,,$$

also $v \in \mathrm{Bild}\big(\Psi'(x_0)\big)$ und damit $T_pM \subset \mathrm{Bild}\big(\Psi'(x_0)\big)$.

Umgekehrt sei $v \in \mathrm{Bild}(\Psi'(x_0))$, sagen wir $v = \Psi'(x_0) \cdot w$ mit $w \in \mathbb{R}^n$. Wir definieren dann $\hat{c}(t) := x_0 + tw$. Für ein kleines $\epsilon > 0$ gilt $\hat{c}\big((-\epsilon, \epsilon)\big) \subset W$. Für $|t| < \epsilon$ gilt $c(t) := \Psi\big(\hat{c}(t)\big) \in \Psi(W) \subset M$. Die Kurve c ist in C^1 und $c(0) = p$, $c'(0) = \Psi'\big(x_0\big) \cdot w = v$. Somit ist $v \in T_pM$. Insgesamt also $\mathrm{Bild}\big(\Psi'(x_0)\big) \subset T_pM$.

Zu (b): Sei wieder $i : \mathbb{R}^n \to \mathbb{R}^k$ die Inklusions-Abbildung $x \mapsto (x, 0)$. Nach Bemerkung 6.28 ist $\Psi := \varphi^{-1} \circ i$ eine lokale Parametrisierung. Sei $\varphi(p) = (x_0, 0)$, $x_0 \in \mathbb{R}^n$. Also haben wir

$$T_pM \overset{\text{(a)}}{=} \Psi'(x_0) \cdot \mathbb{R}^n = (\varphi^{-1})'(\underbrace{i(x_0)}_{=\varphi(p)}) \cdot \underbrace{i'(x_0) \cdot \mathbb{R}^n}_{=\mathbb{R}^n \times \{0\}} = \Big(\varphi'(p)\Big)^{-1} \cdot \big(\mathbb{R}^n \times \{0\}\big)\,.$$

Ein alternativer Beweis von (b) wird durch Lemma 6.20 geliefert.

Zu (c): Sei $\Psi : W \to \mathbb{R}^k$ eine lokale Parametrisierung mit $p \in \Psi(W)$, $W_0 := \Psi^{-1}(U)$, $p = \Psi(x_0)$. Also ist $F \circ \Psi|_{W_0}$ konstant 0, also $0 = F'(p) \cdot \Psi(x_0)$. Somit haben wir $T_pM = \mathrm{Bild}\big(\Psi'(x_0)\big) \subset \ker(F'(p))$. Da $F'(p)$ Rang $k - n$ hat, besagt der Rangsatz, dass $n = \dim \ker(F'(p)) = \dim \mathrm{Bild}\big(\Psi'(x_0)\big)$. Es folgt $T_pM = \ker(F'(p))$. □

BEISPIELE 6.32.

(1) Sei M eine nichtleere offene Teilmenge von $\mathbb{R}^k$. Dann ist M eine k-dimensionale C^∞-Untermannigfaltigkeit von $\mathbb{R}^k$: wähle $U = V = M$ und $\varphi := \mathrm{Id}_M$ als Untermannigfaltigkeitskarte. Aus Proposition 6.31 folgt, $\forall p \in M$:

$$T_pM = \mathbb{R}^k.$$

(2) Sei $M := M_f$ wie in den Beispielen 6.5. Dann ist $\Psi\colon U \longrightarrow \mathbb{R}^k$, $x \longmapsto (x, f(x))^T$ eine globale C^1-Parametrisierung von M_f nach den Beispielen 6.27. Nach Proposition 6.31 gilt dann, für $p = (\hat{p}, f(\hat{p}))^T \in M_f$:

$$\begin{aligned} T_pM = \Psi'(\hat{p})(\mathbb{R}^n) &= \{(v, f'(\hat{p}) \cdot v)^T \mid v \in \mathbb{R}^n\} \\ &= \mathrm{Graph}(f'(\hat{p})) \\ &(= M_{f'(\hat{p})}). \end{aligned}$$

(3) Sei $M := S^n \subset \mathbb{R}^{n+1}$. Sei $F\colon \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$, $x \longmapsto \|x\|_2^2 - 1$. Dann ist $0 \in \mathbb{R}$ regulärer Wert von F mit $F^{-1}(\{0\}) = S^n$. Aus Proposition 6.31 folgt, für jedes $p \in S^n$: $T_pM = \ker(F'(p))$, mit $F'(p) = 2p^T$, d.h.,

$$\begin{aligned} T_pM &= \{v \in \mathbb{R}^{n+1} \mid p^T v = 0\} \\ &= \{v \in \mathbb{R}^{n+1} \mid \langle p, v\rangle = 0\} \\ &= p^\perp \ (= \text{Orthogonales Komplement von } p). \end{aligned}$$

Mi 3.7.

7. Extrema mit Nebenbedingungen

DEFINITION 7.1. Sei $M \subset \mathbb{R}^k$ eine n-dimensionale Untermannigfaltigkeit von $\mathbb{R}^k$ und $U \subset \mathbb{R}^k$ eine offene Teilmenge von $\mathbb{R}^k$ mit $M \subset U$. Sei $f\colon U \longrightarrow \mathbb{R}$ eine Funktion.

(a) Ist f differenzierbar in $p \in M \subset U$, so heißt p stationärer (oder kritischer) Punkt von $f_{|M}\colon M \longrightarrow \mathbb{R}$, wenn $f'(p)_{|T_pM} = 0$ gilt.[44]

(b) Ein Punkt $p \in M$ heißt lokales Maximum (bzw. Minimum) von $f|_M\colon M \longrightarrow \mathbb{R}$, wenn eine offene Umgebung V von p in M so existiert, dass $f(p) \geq f(q)$ (bzw. $f(p) \leq f(q)$) für alle $q \in V$ gilt.

(c) Ein Punkt $p \in M$ heißt lokales Extremum von $f_{|M}\colon M \longrightarrow \mathbb{R}$, wenn p ein lokales Minimum oder ein lokales Maximum von $f|_M$ ist.

Wie findet man lokale Extrema? Wir führen zwei Möglichkeiten auf:

- Wähle eine Parametrisierung $\Psi : W \to \mathbb{R}^k$ der Untermannigfaltigkeit. Wir wissen bereits, wie man lokale Extrema von $f \circ \Psi : W \to \mathbb{R}$ findet. Und x ist lokales Extremum von $f \circ \Psi$ genau dann, wenn $\Psi(x)$ lokales Extremum von $f|_M$ ist. Insbesondere folgt für $p = \Psi(x)$ aus der Kettenregel

  $$x \text{ ist station. Punkt von } f \circ \Psi \iff f'(p) \cdot \Psi'(x) = 0 \iff f'(p)_{|T_pM} = 0.$$

  Diese Methode sollten Sie nun also durch Kombination bereits behandelter Techniken beherrschen.
- Durch Lagrange-Multiplikatoren: siehe unten.

PROPOSITION 7.2. Sei $M \subset \mathbb{R}^k$ eine n-dimensionale Untermannigfaltigkeit, $U \subset \mathbb{R}^k$ eine offene Umgebung von M und $f\colon U \longrightarrow \mathbb{R}$ differenzierbar. Ist $p \in M$ lokales Extremum von $f_{|M}$, so ist p ein stationärer Punkt von $f_{|M}$.

Beweis. Sei $v \in T_pM$. Dann existieren $\varepsilon > 0$ und eine C^1-Abbildung $c\colon (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R}^k$ mit $c((-\varepsilon, \varepsilon)) \subset M$, $c(0) = p$ und $c'(0) = v$. Die Abbildung $f \circ c\colon (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R}$ ist dann

[44] Mit $f'(p)_{|T_pM}$ ist hier die lineare Abildung $\mathbb{R}^k \subset T_pM \ni x \mapsto f'(p) \cdot x \in \mathbb{R}$ gemeint.

C^1 mit $f \circ c(0) = f(p)$. Da p ein lokales Extremum von $f_{|M}$ ist, ist $0 \in (-\varepsilon, \varepsilon)$ ebenfalls ein lokales Extremum von $f \circ c$. Daraus folgt $(f \circ c)'(0) = 0$, d.h., mit der Kettenregel, $f'(p) \cdot v = 0$. Damit bekommen wir $f'(p)_{|T_pM} = 0$, was zu beweisen war. □

Satz 7.3 (Satz über Lagrange-Multiplikatoren)**.** Sei $M \subset \mathbb{R}^k$ eine n-dimensionale Untermannigfaltigkeit, $U \subset \mathbb{R}^k$ eine offene Umgebung von M und $f\colon U \longrightarrow \mathbb{R}$ differenzierbar. Sei $p \in M$ ein stationärer Punkt (z.B. ein lokales Extremum) von $f_{|M}$ und $F\colon V \longrightarrow \mathbb{R}^{k-n}$ eine in einer offenen Umgebung V von p in $\mathbb{R}^k$ definierte C^1-Abbildung mit $0 \in \mathbb{R}^{k-n}$ als regulärem Wert und $F^{-1}(\{0\}) = V \cap M$. Schreibe $F(x) = (F_1(x), \ldots, F_{k-n}(x))\ \forall x \in V$. Dann gibt es eindeutige reelle Zahlen $\lambda_1, \ldots, \lambda_{k-n}$ mit

$$f'(p) = \sum_{j=1}^{k-n} \lambda_j F_j'(p).$$

Diese Zahlen $\lambda_1, \ldots, \lambda_{k-n}$ heißen Lagrange-Multiplikatoren von f unter den Nebenbedingungen $F_1, \ldots, F_{k-n}$ an der Stelle p.

Zusatz 7.4. Umgekehrt gilt: existieren solche Lagrange-Multiplikatoren, dann ist p ein stationärer Punkt von $f|_M$.

Die Lagrange-Multiplikatoren sind nach dem Mathematiker und Physiker Joseph-Luis Lagrange benannt, der insbesondere die Lagrangesche Formulierung der klassischen Mechanik in der Physk vorangetrieben hat. Es ist aber auch ein wichtiges Hifsmittel zum Beispiel in den Wirtschaftswissenschaften, wenn man eine Optimierung unter gegebenen nicht-linearen Zwangsbedingungen durchführen will.

Beweis des Korollars. Nach Voraussetzung gilt $f'(p)_{|T_pM} = 0$ (im Fall, wo $p \in M$ ein lokales Extremum von $f_{|M}$ ist, folgt dies aus Proposition 7.2). Nach Proposition 6.31 gilt

aber

$$\begin{aligned} T_pM &= \ker(F'(p)) \text{ mit } F'(p) = \begin{pmatrix} F_1'(p) \\ \vdots \\ F_{k-n}'(p) \end{pmatrix} \\ &= \bigcap_{j=1}^{k-n} \ker(F_j'(p)). \end{aligned}$$

Mit

$$\begin{aligned} \ker(F_j'(p)) &= \{v \in \mathbb{R}^k \mid F_j'(p) \cdot v = 0\} = \{v \in \mathbb{R}^k \mid \langle F_j'(p)^T, v\rangle = 0\} \\ &= (F_j'(p)^T)^\perp \end{aligned}$$

folgt $T_pM = \bigcap_{j=1}^{k-n} (F_j'(p)^T)^\perp = \text{Span}\{F_1'(p)^T, \ldots, F_{k-n}'(p)^T\}^\perp$ (siehe LA II). Die Bedingung $f'(p)_{|T_pM} = 0$ ist zu $\langle f'(p)^T, v\rangle = 0 \ \forall v \in T_pM$ äquivalent, d.h. zu $f'(p)^T \in (T_pM)^\perp \subset \mathbb{R}^k$. Insgesamt folgt

$$f'(p)^T \in (\text{Span}\{F_1'(p)^T, \ldots, F_{k-n}'(p)^T\}^\perp)^\perp = \text{Span}\{F_1'(p)^T, \ldots, F_{k-n}'(p)^T\}, \tag{7.5}$$

d.h., $\exists\, \lambda_1, \ldots, \lambda_{k-n} \in \mathbb{R}$ mit $f'(p)^T = \sum_{j=1}^{k-n} \lambda_j F_j'(p)^T$, d.h., $f'(p) = \sum_{j=1}^{k-n} \lambda_j F_j'(p)$. Da $F_j'(p), \ldots, F_{k-n}'(p)$ linear unabhängig sind, sind $\lambda_1, \ldots, \lambda_{k-n}$ eindeutig mit dieser Eigenschaft. □

Der Zusatz folgt, indem wir die obigen Argmente umkehren: Wenn Lagrange-Multiplikatoren existieren, dann gilt (7.5) und somit $f'(p)^T \in (T_pM)^\perp$. Dann verschwindet $v \mapsto f'(p) \cdot v$ auf T_pM.

KAPITEL 10

Gewöhnliche Differentialgleichungen

Literatur zu diesem Kapitel.

- [**15**], Kapitel 2 (gut zu lesen)
- [**7**], Kapitel 14 (Rolle des Banachschen Fixpunktsatzes klarer)
- [**11**] (umfassender, tiefer gehend)
- [**23**]Mehr zu Stabilität
- [**35**]noch mehr zu Stabilität, deutlich ausführlicher und umfassender als [**23**], sehr gut lesbar.
- [**27**]
- [**4**]

1. Motivation

Wir motivieren an Hand eines ebenen Pendels aus der Physik. Es gibt auch viele Systeme in anderen Wissenschaften (Chemie, Biologie, Medizin, Wirtschaftswissenschaften,...), die durch gewöhnliche Differentialgleichungen modelliert werden.

Bild eines Pendels mit eingezeichneten Kräften, siehe Vorlesung.

$\varphi : \mathbb{R} \to \mathbb{R}$

$$\varphi''(t) = -\frac{g}{\ell}\sin(\varphi(t)) \tag{1.1}$$

wobei $g = 9,81 N/kg$ die Gravitationskonstante, ℓ die Pendellänge und φ der Auslenkungswinkel ist.

Wichtige physikalische Fragen:

- Beschreibt dieses Modell die physikalische Situation? (Problematisch: Reibung, Drehenergie des Körpers, Masse des Stabes, Störeinflüsse, ...)
- Wie genau sind die Ausgangswerte $\varphi(0)$, $\varphi'(0)$ und die Parameter g, ℓ,...?

Dies ist nicht das Thema der Vorlesung!

Wichtige mathematische Fragen:

- Gibt es Funktionen $\varphi : \mathbb{R} \to \mathbb{R}$, die (1.1) lösen? (Wenn nicht, dann ist dies sicher kein gutes Modell!) Wir werden sehen: es gibt Lösungen!
- Wenn ja, wieviele Lösungen gibt es? Wir werden sehen: zu gegebenem $\varphi(0)$ und $\varphi'(0)$ gibt es genau eine Lösungsfunktion. Die Gleichung (1.1) hat also einen zwei-dimensionalen Lösungsraum.[1]
- Gibt es Lösungen $\varphi : (a, b) \to \mathbb{R}$, die sich nicht zu Lösungen auf ganz $\mathbb{R}$ fortsetzen lassen? Wir werden sehen: „Nein“ für diese Gleichung, aber bei vielen physikalischen Systemen „Ja“.

[1]Den Begriff „zwei-dimensional“ haben wir hier noch gar nicht sauber definiert. Der Lösungsraum ist eine zwei-dimensionale Untermannigfaltigkeit des unendlich-dimensionalen Vektorraums $C^\infty(\mathbb{R}, \mathbb{R})$. Die obige Aussage interpretieren Sie am besten so, dass der Lösungsraum durch die beiden Anfangswert $\varphi(0)$ und $\varphi'(0)$ parametrisiert wird.

- Sind alle Lösungen periodisch? Man kann zeigen: viele, aber nicht alle. [2]
- Kann man die Lösungen explizit angeben? Antwort: Es geht noch mit großem Aufwand. Man braucht „elliptische Integrale“, siehe `http://en.wikipedia.org/wiki/Pendulum_(mathematics)`. Bei komplizierteren physikalischen Systemen nahezu unmöglich.
- Wie lang ist die Periodenlänge von periodischen Lösungen? Antwort: Als Potenzreihe angebbar. Bei komplizierteren Gleichungen nicht explizit berechenbar.
- Hängt sie von der maximalen Auslenkung ab? Antwort: Die Periodenlänge wird länger, wenn die maximale Auslenkung länger wird.
- Erhaltungsgrößen? Antwort: Energieerhaltung, Drehimpulserhaltung (beim räumlichen Pendel),...

Ziel: Möglichst großes systematisches Verständnis solcher Gleichungen. Beschreibung des qualitativen Verhaltens, falls die Lösungen nicht explizit zu finden sind.

Oft hilfreich: Linearisierte Gleichung. Ersetze nicht-lineare Terme durch das Taylorpolynom ersten Grades. Im Pendel-Beispiel: ersetze

$$\sin\varphi = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!}\varphi^{2j+1} = 0 + \varphi + 0\cdot\varphi^2 - \frac{1}{6}\varphi^3 + \cdots$$

[2] Alle Lösungen, deren Gesamtenergie unterhalb von $E_0 := mg\ell$ liegt, können keine Überschläge (=loopings) machen, sie sind periodisch. Lösungen φ mit Gesamt-Energie größer als E_0 sind im Sinne von Funktionen $\varphi : \mathbb{R} \to \mathbb{R}$ nicht periodisch. Die zugehörige räumliche Bewegungskurve $\mathbb{R} \to \mathbb{R}^3$, $t \mapsto (\ell\sin\varphi(t), 0, -\ell\cos\varphi(t))^T$ ist aber periodisch. Ist die Gesamtenergie gleich E_0, dann ist weder φ noch die zugehörige räumliche Bewegungskurve periodisch. Dann gilt $\lim_{t\to\pm\infty}\varphi(t) = \pm\pi$ und die räumliche Bewegungskurve konvergiert in beide Richtungen gegen $(0,0,-\ell)^T$.

durch φ. Wir erhalten

$$\varphi''(t) = -\frac{g}{\ell}\varphi(t)$$

Viel leichter zu lösen:

$$\varphi(t) = a\cos\left(\sqrt{\frac{g}{\ell}}\,t\right) + b\sin\left(\sqrt{\frac{g}{\ell}}\,t\right)$$

Beschreiben die Lösungen der vereinfachten Gleichung auch hinreichend gut die Lösungen der ursprünglichen Gleichung?

2. Definition und Reduktion auf autonome Gleichungen erster Ordnung

DEFINITION 2.1. Seien $m, n, k \in \mathbb{N}$. Gegeben sei

$$U \subset \mathbb{R} \times (\mathbb{R}^n)^{k+1} = \mathbb{R} \times \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k+1\text{-mal}}$$

und eine Funktion $F : U \to \mathbb{R}^m$. Wir sagen $\varphi : (a, b) \to \mathbb{R}^n$, $a, b \in \mathbb{R}$, $a < b$ eine Lösung der durch F gegebenen gewöhnlichen Differentialgleichung, falls gilt:

(1) φ ist k-mal differenzierbar
(2) $(t, \varphi(t), \varphi'(t), \ldots, \varphi^{(k)}(t)) \in U$ für alle $t \in (a, b)$
(3) Für alle $t \in (a, b)$ gilt

$$F(t, \varphi(t), \varphi'(t), \ldots, \varphi^{(k)}(t)) = 0. \tag{2.2}$$

Man nennt (a, b) das Lösungsintervall der Lösung. Eine Gleichung der Form (2.2) nennt man eine gewöhnliche Differentialgleichung von Ordnung $\leq k$.[3] Wir nennen F die definierende Funktion der gewöhnlichen Differentialgleichung. Wir sagen, die gewöhnliche

[3]Man sagt, die Ordnung ist gleich k, falls F tatsächlich von $\varphi^{(k)}$ abhängt.

Differentialgleichung ist autonom, falls die definierende Gleichung nicht von t abhängt. Wir schreiben dann oft

$$F(\varphi(t), \varphi'(t), \ldots, \varphi^{(k)}(t)) = 0,$$

d.h. $F : U \to \mathbb{R}^m$, $U \subset (\mathbb{R}^n)^{k+1}$.

Beispiele 2.3.

(1) Pendel: $m = n = 1$, $k = 2$, $U = \mathbb{R}^4$,

$$F(t, \varphi_0, \varphi_1, \varphi_2) = \varphi_2 + \frac{g}{\ell} \sin \varphi_0$$

(2) Linearisiertes Pendel: wie oben aber

$$F(t, \varphi_0, \varphi_1, \varphi_2) = \varphi_2 + \frac{g}{\ell} \varphi_0$$

(3) $m = n = k = 1$, $U = \mathbb{R}^3$. Sei $a \in \mathbb{R}$

$$F(t, \varphi_0, \varphi_1) = \varphi_1 - a\varphi_0$$

Wir erhalten die Differentialgleichung

$$\varphi'(t) = a\varphi(t).$$

Dann ist $\varphi(t) = e^{at}$ eine Lösung.

(4) Sei $g : (a, b) \to \mathbb{R}$. Wir definieren

$$F(t, \varphi_0, \varphi_1) = \varphi_1 - g(t).$$

Dann ist G eine Lösung von

$$F(t, G(t), G'(t)) = 0,$$

genau dann, wenn G eine Stammfunktion zu g ist.

(5) Sei $f : \mathbb{R}^2 \to \mathbb{R}$. Für

$$F(t, \varphi_0, \varphi_1) = \varphi_1 - f(t, \varphi_0).$$

haben wir die Differentialgleichung

$$\varphi'(t) = f(t, \varphi(t)). \tag{2.4}$$

Die ersten drei Beispiele sind autonom, die letzten beiden nicht.

Bemerkungen 2.5. Wichtig sind auch Differentialgleichungen, in denen Funktionen in mehreren Veränderlichen gesucht werden, so dass die definierende Funktion von partiellen Ableitungen in mehreren Veränderlichen abhängt. Solche Differentialgleichungen nennt man partielle Differentialgleichung. Genauer gesagt: eine partielle Differentialgleichung ist eine Differentialgleichung, die keine gewöhnliche Differentialgleichung ist.

Zum Beispiel sei V offen in $\mathbb{R}^\ell$ und $\varphi : V \to \mathbb{R}^n$. Wir definieren den Laplace-Operator $\Delta : C^2(V, \mathbb{R}^n) \to C^0(V, \mathbb{R}^n)$

$$\Delta\varphi := \sum_{i=1}^{\ell} \frac{\partial^2}{\partial x_\ell^2}\varphi$$

Dann ist $\Delta\varphi = 0$ eine partielle Differentialgleichung, die in vielen Bereichen der Anwendungen (Physik, Chemie, Biologie, Finanzmathematik,...) sehr wichtig ist.

Wir behandeln in dieser Vorlesung nur gewöhnliche Differentialgleichungen. Die Lösungstheorie partieller Differentialgleichungen wird in fortgeschritteneren Vorlesungen behandelt.

Bemerkung 2.6.

(1) Man kann jede gewöhnliche Differentialgleichung von k-ter Ordnung in eine Differentialgleichung erster Ordnung überführen. Hierzu definieren wir für $\varphi : (a, b) \to$

$\mathbb{R}^n$ die Funktionen $\psi_i(t) := \varphi^{(i)}(t)$ und fassen diese Funktionen zu $\Psi : (a,b) \to (\mathbb{R}^n)^k$ als $\Psi(t) := (\psi_0(t), \psi_1(t), \ldots, \psi_{k-1}(t))$ zusammen. Wir haben genau dann eine Lösung von Gleichung (2.2), falls wir eine Lösung von

$$\begin{aligned} 0 &= F(t, \psi_0(t), \psi_1(t), \ldots, \psi_{k-1}(t), \psi'_{k-1}(t)) \\ 0 &= \psi'_0(t) - \psi_1(t) \\ \vdots &= \vdots \\ 0 &= \psi'_{k-2}(t) - \psi_{k-1}(t) \end{aligned}$$

haben. Wir können nun die gesamte rechte Seite in der Form $\widetilde{F}(t, \Psi(t), \Psi'(t))$ schreiben, wobei $\widetilde{F} : \widetilde{U} \to \mathbb{R}^m \times (\mathbb{R}^n)^{k-1}$ für eine geeignete Teilmenge $\widetilde{U}$ von $\mathbb{R} \times (\mathbb{R}^n)^k \times (\mathbb{R}^n)^k$.

$$\widetilde{F}(t, \Psi(t), \Psi'(t)) = \begin{pmatrix} F(t, \psi_0(t), \psi_1(t), \ldots, \psi_{k-1}(t), \psi'_{k-1}(t)) \\ \psi'_0(t) - \psi_1(t) \\ \vdots \\ \psi'_{k-2}(t) - \psi_{k-1}(t) \end{pmatrix} \in \mathbb{R}^{m+(k-1)n}.$$

Wir haben eine gewöhnliche Differentialgleichung von k-ter Ordnung für $\varphi : (a,b) \to \mathbb{R}^n$ durch eine gewöhnliche Differentialgleichung erster Ordnung für $\Psi : (a,b) \to \mathbb{R}^{nk}$ ausgedrückt.

Im Beispiel des Pendels führen wir also eine Funktion $\rho : (a,b) \to \mathbb{R}$ ein,[4] die φ' repräsentieren soll, und lösen das System

$$\begin{aligned} 0 &= \rho'(t) + \frac{g}{\ell} \sin \varphi(t) \\ 0 &= \varphi'(t) - \rho(t) \end{aligned}$$

[4]Also $\psi_0(t) = \varphi(t)$, $\psi_1(t) = \rho(t)$ in der Notation von oben.

Wir definieren dann

$$\widetilde{F}(t, \begin{pmatrix} \varphi \\ \rho \end{pmatrix}, \begin{pmatrix} \hat{\varphi} \\ \hat{\rho} \end{pmatrix}) = \begin{pmatrix} \hat{\rho} + \frac{g}{\ell} \sin \varphi \\ \hat{\varphi} - \rho \end{pmatrix} \in \mathbb{R}^2.$$

Die dadurch definierte gewöhnliche Differentialgleichung ist

$$0 = \widetilde{F}(t, \begin{pmatrix} \varphi(t) \\ \rho(t) \end{pmatrix}, \begin{pmatrix} \varphi(t) \\ \rho'(t) \end{pmatrix}) = \begin{pmatrix} \rho'(t) + \frac{g}{\ell} \sin \varphi(t) \\ \varphi'(t) - \rho(t) \end{pmatrix} \in \mathbb{R}^2.$$

Dies ist eine autonome gewöhnliche Differentialgleichung erster Ordnung.

(2) Man kann jede gewöhnliche Differentialgleichung in eine autonome Differentialgleichung überführen. Hierzu führen wir eine neue Funktion $\tau(t)$, die $\tau'(t) = 1$ und $\tau(t_0) = t_0$ erfüllt. Es gilt dann also $\tau(t) = t$. Wir setzen nun $\Psi(t) := (\tau(t), \varphi(t))$, $\Psi : (a, b) \to \mathbb{R}^{n+1}$ setzen und dann die definierenden Funktion $F : U \to \mathbb{R}^m$ durch eine weitere Komponente $\tau'(t) = 1$ zu einer Funktion $\widetilde{F} : U \to \mathbb{R}^{m+1}$ ergänzen.[5]

Wir betrachten dazu Beispiel 2.3 (5). Wir überführen die bisherige gewöhnliche Differentialgleichung für $\varphi(t)$ in eine für $\Psi(t) = (\tau(t), \varphi(t))$. Dazu definieren wir

$$\widetilde{F}\left(\begin{pmatrix} \tau_0 \\ u_0 \end{pmatrix}, \begin{pmatrix} \tau_1 \\ u_1 \end{pmatrix}\right) = \begin{pmatrix} \tau_1 - 1 \\ u_1 - f(\tau_0, u_0). \end{pmatrix}$$

Diese Funktion definiert die autonome gewöhnliche Differentialgleichung

$$\begin{cases} \tau'(t) = 1 \\ u'(t) = f(\tau(t), u(t)) \end{cases} \tag{2.7}$$

Ist u eine Lösung von (2.4), dann ist$\Psi(t) := (t, u(t))$ eine Lösung von (2.7). Ist umgekehrt $\Psi(t) = (\tau(t), u(t))$ eine Lösung von (2.7). Dann folgt $\tau(t) = t + t_0$ für

[5]Der Definitionsbereich von F ändert sich nicht, aber die Interpretation der t-Variable.

eine Konstante $t_0 \in \mathbb{R}$. Wir setzen dann $\tilde{u}(t) := u(t - t_0)$. Dann gilt

$$\tilde{u}'(t) = u'(t - t_0) = f(t, u(t - t_0)) = f(t, \tilde{u}(t)).$$

Also ist $\tilde{u}$ eine Lösung von (2.4).

Fr 5.7.

3. Flusslinien und Erhaltungsgrößen

Wir haben diskutiert:

- Jede gewöhnliche Differentialgleichung kann auf eine Differentialgleichung erster Ordnung reduziert werden
- Jede gewöhnliche Differentialgleichung kann auf eine autonome Differentialgleichung (der gleichen Ordnung) reduziert werden

Viele (gewöhnliche)[6] Differentialgleichungen haben keine Lösung oder können nur schwer gelöst werden. Zum Beispiel hat

$$(\varphi(t)^2 + (\varphi'(t) - 1)^2 = 0$$

keine Lösung $\varphi : (a, b) \to \mathbb{R}$.

In guten Situationen hingegen kann man Differentialgleichungen erster Ordnung nach $\varphi'(t)$ auflösen, und dann sind die Gleichungen besser zu behandeln:

$$\varphi'(t) = f(t, \varphi(t))$$

[6]Wir lassen das Wort „gewöhnlich“ nun oft weg, da alle Differentialgleichungen in der Vorlesung gewöhlich sind.

oder

$$\varphi'(t) = f(\varphi).$$

Man nennt dies eine explizite gewöhnliche Differentialgleichung.

DEFINITION 3.1. Sei U offen in $\mathbb{R}^n$. Ein (glattes) Vektorfeld ist eine (glatte) Abbildung $f : U \to \mathbb{R}^n$.

Vorstellung: Im Punkt x beginnt der Pfeil, der durch $f(x)$ beschrieben wird.

Bild eines Vektorfelds, symbolisiert duruch viele Pfeile in $\mathbb{R}^2$.

DEFINITION 3.2. Eine Flusslinie oder Integralkurve des Vektorfeldes $f : U \to \mathbb{R}^n$ ist eine Lösung von

$$\varphi'(t) = f(\varphi(t)).$$

autonome explizite gew. Diff'glng auf $U \mathrel{\widehat{=}}$ Vektorfeld auf U .

DEFINITION 3.3. Eine Erhaltungsgröße oder ein erstes Integral zum Vektorfeld f ist eine Funktion $E : U \to \mathbb{R}$, so dass $t \mapsto E(\varphi(t))$ entlang jeder Flusslinie von f konstant ist.

Schreibweise: schreibe auch $\frac{d}{dt}h$ für die Ableitung von h nach t.

BEISPIEL 3.4 (Pendel). Wir betrachten die Gleichung

$$\varphi''(t) = -\sin\varphi(t). \tag{3.5}$$

Wir formen in eine Differentialgleichung erster Ordnung um, indem wir $\rho(t) = \varphi'(t)$ setzen, $\Psi(t) = (\varphi(t), \rho(t))^T \in \mathbb{R}^2$.

$$\begin{pmatrix} \varphi(t) \\ \rho(t) \end{pmatrix}' = \begin{pmatrix} \begin{pmatrix} \rho(t) \\ -\sin\varphi(t) \end{pmatrix} \end{pmatrix}$$

Lösung sind Flusslinien von $f(\varphi, \rho) = (\rho, -\sin\varphi)^T$. Ist $\Psi(t) = (\varphi(t), \rho(t))^T$ eine Flusslinie von f, so ist $\varphi(t)$ eine Lösung von (3.5).

Multipliziere (3.5) mit $\varphi'(t)$,

$$\frac{1}{2}\frac{d}{dt}((\varphi'(t))^2) = \varphi'(t)\varphi''(t) = -\varphi'(t)\sin(\varphi(t)) = \frac{d}{dt}(\cos\varphi(t))$$

Wir definieren $E : \mathbb{R}^2 \to \mathbb{R}$, $E(\varphi, \rho) = \frac{1}{2}\rho^2 - \cos\varphi$. Dann ist

$$\frac{d}{dt}(E((\varphi(t), \rho(t))^T)) = \frac{1}{2}\left((\varphi'(t))^2 - \cos\varphi(t)\right)' = 0.$$

Also ist E eine Erhaltungsgröße.

Die Höhenlinien der Funktion $(\varphi, \rho) \mapsto E(\varphi, \rho)$. Zusammenhangskomponenten einer Höhenlinie (ohne stationäre Punkte) bilden die Orbits, das heißt die Kurven, entlang derer sich die Lösung bewegt. Siehe „Phase diagram“ auf der Webseite `http://physique.unice.fr/sem6/2011-2012/PagesWeb/PT/Pendule/En/study1_simple.html`

ÜBUNG 3.6 (Nutzt den Satz von Picard-Lindelöf, siehe unten)**.** Sei $f : U \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld. Sei $E : U \to \mathbb{R}$ differenzierbar. Dann ist E genau dann

eine Erhaltungsgröße von f, wenn [7]

$$f(x) \in \ker E'(x)$$

für alle $x \in U$.

Lösung: Zentralübung

Beweis. Es gelte $f(x) \in \ker E'(x)$ für alle $x \in U$. Sei $\varphi : (a, b) \to U$ eine Flusslinie von f, dann gilt

$$(E \circ \varphi)'(t) = E'(\varphi(t)) \cdot \varphi'(t) = E'(\varphi(t)) \cdot f(\varphi(t)) = 0.$$

Also ist E eine Erhaltungsgröße.

Umgekehrt sei E eine Erhaltungsgröße. Da f stetig differenzierbar ist, ist f Orts-Lipschitz-stetig. Zu jedem $x_0 \in U$ existiert also nach dem Satz von Picard-Lindelöf ein $\epsilon > 0$ und eine Flusslinie $\varphi : (-\epsilon, \epsilon) \to U$ von f mit $\varphi(0) = x_0$. Da $E \circ \varphi$ konstant ist, rechnen wird

$$0 = (E \circ \varphi)'(0) = E'(\varphi(0)) \cdot \varphi'(0) = E'(x_0) \cdot f(x_0).$$

Somit gilt $f(x_0) \in \ker E'(x_0)$ für alle $x_0 \in U$. □

ÜBUNG 3.7 (Nutzt den Satz von Picard-Lindelöf, siehe unten)**.** Sei M eine Untermannigfaltigkeit von $\mathbb{R}^n$, $M \subset U$, U offen in $\mathbb{R}^n$ und sei M abgeschlossen in $\mathbb{R}^n$. Sei $f : U \to \mathbb{R}^n$ ein Lipschitz-stetiges Vektorfeld. Ist f tangential an M (d.h. es gelte $f(p) \in T_pM$ für alle $p \in M$) und ist $c : (a, b) \to \mathbb{R}^n$ eine Integralkurve von f mit $a < t_0 < b$ und $c(t_0) \in M$, so ist $c(t) \in M$ für alle $t \in (a, b)$.

[7] Hier wird $E'(x)$ als lineare Abbildung $\mathbb{R}^n \to \mathbb{R}$ betrachtet. Wenn wir $E'(x)$ als Zeilenvektor betrachten, heißt die analoge Aussage $(E'(x))^T \perp f(x)$ oder äquivalent $E'(x) \cdot f(x) = 0$.

Lösung: Zentralübung

BEMERKUNG 3.8. Im allgemeinen ist es sehr schwer, eine oder mehrere Erhaltungsgrößen zu erhalten. In der Lagrange-Mechanik der klassischen Physik liefert das Noethersche Theorem zu jeder Symmetrie eines mechanischen Systems eine Erhaltungsgröße. Typische Erhaltungsgrößen in der Physik sind Energie, Komponenten des Impulsvektors, Komponenten des Drehimplusvektors, etc..

Wenn Erhaltungsgrößen $E_1, \dots, E_\ell$ existieren, dann bilden diese eine erhaltene Abbildung $E = (E_1, \dots, E_\ell) : U \to \mathbb{R}^\ell$. Jede Flussline $\varphi : I \to U$ mit $E(\varphi(t_0)) = E_0$, $t_0 \in I$ erfüllt dann also $\varphi(I) \subset E^{-1}(\{E_0\})$. Ist E_0 ein regulärer Wert von E, dann bewegt sich die Flusslinie entlang der $(n-\ell)$-dimensionalen Untermannigfaltigkeit $M := E^{-1}(\{E_0\})$. Durch Verwendung einer lokalen Parametrisierung $\Psi : W \to M$ kann man die durch f gegebene gewöhnliche Differentialgleichung in eine autonome gewöhnliche Differentialgleichung erster Ordnung auf W umwandeln.[8] Somit haben wir die Lösung der Differentialgleichung lokal auf die Lösung einer anderen Differentialgleichung reduziert, die oft einfacher zu lösen ist, da die Dimension $n - \ell$ kleiner ist.

4. Der Satz von Picard-Lindelöf

Im folgenden Abschnitt besteht die Gefahr, Intervalle $(a, b) \subset \mathbb{R}$ mit dem Paar $(a; b) \in \mathbb{R}^2$ zu verwechseln. Wir schreiben deswegen weiterhin (a, b) für das Paar und $(a; b)$ für das Intervall.

Ab jetzt: Sei U offen in $\mathbb{R} \times \mathbb{R}^n$, $f : U \to \mathbb{R}^n$.

[8]Genauer: Man definiert ein Vektorfeld $\tilde{f} : W \to \mathbb{R}^{n-\ell}$ so, dass $\Psi'(p)\tilde{f}(p) = f(\Psi(p))$ für alle $p \in W$. Dannist $\tilde{\varphi} : I \to W$ eine Flusslinie von $\tilde{f}$, wenn $\varphi := \Psi \circ \tilde{\varphi}$ eine Flusslinie von f ist.

Elemente in $\mathbb{R} \times \mathbb{R}^n \ni (t, x)$, $t \in \mathbb{R}$ Zeit, $x \in \mathbb{R}^n$ Ort.

DEFINITION 4.1. Sei f wie oben. Wir sagen f ist Orts-Lipschitz-stetig, falls es ein $L \in \mathbb{R}$ gibt, so dass für alle $(t, x), (t, y) \in U$ gilt:

$$\|f(t, x) - f(t, y)\| \leq L\|x - y\|.$$

Wir sagen auch f ist Orts-Lipschitz-stetig mit Lipschitz-Konstante L.

Wir sagen f ist lokal Orts-Lipschitz-stetig, falls jedes $(t, x) \in U$ eine offene Umgebung V in U besitzt, so dass $f|_V$ Orts-Lipschitz-stetig ist.

9

BEISPIELE 4.2.

(1) Sei $f : U \to \mathbb{R}^n$ Lipschitz-stetig, das heißt, es gebe ein $L \in \mathbb{R}$, so dass für alle $(t, x), (s, y) \in U$:

$$\|f(t, x) - f(s, x)\| \leq L\sqrt{|t - s|^2 + \|x - y\|^2}.$$

Dann ist f Orts-Lipschitz-stetig. Insbesondere ist jede lineare Funktion (Orts-)Lipschitz-stetig.

(2) $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $f(t, x) = \sqrt[3]{t}$ ist Orts-Lipschitz-stetig und stetig, aber nicht Lipschitz-stetig.

(3) Lipschitz-stetige Funktionen sind stetig, aber es gibt nicht-stetige Orts-Lipschitz-stetige Funktionen

[9]Die Eigenschaften Orts-Lipschitz-stetig und lokal Orts-Lipschitz-stetig sind unabhängig von der gewählten Norm auf $\mathbb{R}^n$. Die möglichen Konstanten L hängen aber von der Wahl der Norm ab.

(4) Schreibe $(t, x) = (t, x_1, \ldots, x_n)$. Angenommen die partiellen Ableitungen

$$\frac{\partial}{\partial x_1} f, \ldots, \frac{\partial}{\partial x_n} f : U \to \mathbb{R}^n$$

existieren und sind stetig (als Funktionen $U \to \mathbb{R}^n$). Sei $\overline{B}_\epsilon(t_0, x_0) \in U$. Wir definieren

$$C := \max_{i=1,\ldots,n} \max_{(s,y)\in\overline{B}_\epsilon(t_0,x_0)} \left\| \frac{\partial}{\partial x_i} f(s, y) \right\|_\infty < \infty.$$

Das hintere Maximum wird angenommen (und ist deswegen endlich), da $\overline{B}_\epsilon(t_0, x_0)$ kompakt ist. Schreibe $f(t, x) = (f_1(t, x), \ldots, f_n(t, x))$. Dann gilt nach dem Mittelwertsatz

$$f_j(s, y) - f_j(s, z) = \sum_{i=1}^{n} (y_i - z_i) \frac{\partial f_j}{\partial x_i}(s, w_j)$$

für ein w_j auf der Strecke von y nach z. Also

$$|f_j(s, y) - f_j(s, z)| \leq \|y - z\|_\infty nC,$$

falls $(s, y), (s, z) \in \overline{B}_\epsilon(t_0, x_0)$. Bezüglich der Norm $\| \cdot \|_\infty$ ist $f|_{B_\epsilon(t_0,x_0)}$ also Orts-Lipschitz-stetig. Es folgt, dass f lokal Orts-Lipschitz-stetig ist.

SATZ 4.3 (Satz von Picard-Lindelöf)**.** Seien $U \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ lokal [10] Orts-Lipschitz-stetig und stetig. Sei $(t_0, x_0) \in U$. Dann gibt es ein $\epsilon_0 \in \mathbb{R}_{>0}$, so dass es für jedes $\epsilon \in (0; \epsilon_0]$ genau eine stetig differenzierbare Abbildung $\varphi : (t_0 - \epsilon; t_0 + \epsilon) \to \mathbb{R}^n$ gibt mit:

- $\varphi(t_0) = x_0$ (Anfangswert)
- für alle $t \in (t_0 - \epsilon; t_0 + \epsilon)$ gelte $(t, \varphi(t)) \in U$

[10] Die Aussage ist natürlich auch noch immer richtig, wenn wir das Wort „lokal“ weglassen.

- und für alle $t \in (t_0 - \epsilon; t_0 + \epsilon)$ gelte

$$\varphi'(t) = f(t, \varphi(t)). \tag{4.4}$$

Beweis. Im folgenden sei

$$B_\delta^\infty(x_0) := \{x \in \mathbb{R}^n \mid \|x - x_0\|_\infty < \delta\}.$$

Wir wählen $\delta > 0$, so dass $V := (t_0-\delta; t_0+\delta) \times B_\delta^\infty(x_0) \subset \overline{V} := [t_0-\delta; t+\delta] \times \overline{B_\delta^\infty(x_0)} \subset U$. Wenn wir $\delta > 0$ genügend klein wählen, dann ist $f|_V$ Orts-Lipschitz-stetig, und sei $L \in \mathbb{R}$ hierzu eine Lipschitz-Konstante. Da f stetig ist, ist f beschränkt auf der kompakten Menge $\overline{V}$.

$\epsilon \in (0; \delta)$ wird später festgelegt.

Notation: $I = (t_0 - \epsilon; t_0 + \epsilon)$,

Sei $E := C_b(I, \mathbb{R}^n)$ der Vektorraum der beschränkten stetigen Funktionen von I nach $\mathbb{R}^n$. Für $\varphi \in E$ definiere[11]

$$|||\varphi|||_\infty := \sup_{t \in I} \|\varphi(t)\|_\infty$$

$$\varphi_i \to \varphi \text{ in } (E, |||\cdot|||_\infty) \quad \overset{\text{(def)}}{\iff} \quad \varphi_i \text{ konvergiert gleichmäßig gegen } \varphi$$

Lemma 4.5. $(E, |||\cdot|||_\infty)$ ist ein Banachraum, d.h. jede Cauchy-Folge in E konvergiert.

[11] Um die Norm auf $\mathbb{R}^n$ und die auf E genauer zu unterscheiden, schreiben wir für letztere im Skript $|||\cdot|||_\infty$.

Beweis des Lemmas. Wir wenden Korollar 2.8 aus Kapitel 7 an, wobei (Y, d) der metrische Raum zum normierten Raum $(\mathbb{R}^n, \|\cdot\|_\infty)$ ist und $D = I$. Wir erhalten, dass $(\mathrm{Abb}(I, \mathbb{R}^n), d_{\mathrm{glm}})$ vollständig ist. Wir haben $E = \{\varphi \in \mathrm{Abb}(I, \mathbb{R}^n) \mid \varphi \text{ ist beschränkt und stetig}\}$. Wegen Satz 2.5 aus Kapitel 7 ist ein gleichmäßiger Grenzwert stetiger Funktionen wiederum stetig. Da auch Beschränkheit unter gleichmäßigen Limites erhalten bleibt, ist E abgeschlossen in $\mathrm{Abb}(I, \mathbb{R}^n)$. Somit ist E vollständig bzgl. d_{glm}. Hieraus folgt, dass $(E, |||\cdot|||_\infty)$ vollständig ist. □

Sei φ_0 konstant x_0, d.h. $\varphi_0(t) = x_0\ \forall t \in I$. Wir definieren die folgende Teilmenge von E:

$$E_{\delta,\epsilon} := \{\varphi \in E \mid \forall t \in (t_0 - \epsilon; t_0 + \epsilon) : \|\varphi(t) - x_0\|_\infty \leq \delta\} = \{\varphi \in E \mid |||\varphi - \varphi_0|||_\infty \leq \delta\}.$$

Man sieht leicht, dass $E_{\delta,\epsilon}$ eine abgeschlossene Teilmenge von $(E, |||\cdot|||_\infty)$ ist. Insbesondere ist $E_{\delta,\epsilon}$ mit der von $|||\cdot|||_\infty$ induzierten Metrik ein vollständiger metrischer Raum.

Wir definieren[12]

$$A : E_{\delta,\epsilon} \to E, \qquad A(\varphi)(t) = x_0 + \int_{t_0}^{t} f(\tau, \varphi(\tau))\, d\tau.$$

Die Beschränktheit von $A(\varphi)$ folgt aus der Beschränktheit von $f|_{\overline{V}}$.

Lemma 4.6 (Umformulierung der Gew. Dgl. in ein Fixpunktproblem eines Integraloperators)**.** Es gilt $A(\varphi) = \varphi$ genau dann, wenn φ eine stetig differenzierbare Lösung von (4.4) ist mit $\varphi(t_0) = x_0$.

Beweis des Lemmas. Es gelte $A(\varphi) = \varphi$. Dann gilt (inklusive Existenz der Ableitung!)

$$\varphi'(t) = \frac{d}{dt} \int_{t_0}^{t} f(\tau, \varphi(\tau))\, d\tau = f(t, \varphi(t)).$$

[12]Im folgenden ist immer $A(\varphi)(t)$ als $\Big(A(\varphi)\Big)(t)$ zu lesen

Aus dieser Formel folgt auch die stetige Differenzierbarkeit von φ.

Umgekehrt: ist φ eine stetig differenzierbare Lösung von (4.4) mit $\varphi(t_0) = x_0$, dann ist

$$A(\varphi)(t) = x_0 + \int_{t_0}^{t} f(\tau, \varphi(\tau))\, d\tau = x_0 + \int_{t_0}^{t} \varphi'(\tau)\, d\tau = \varphi(t).$$

□

bis hier: Fr 5.7.

ab hier: Mi 17.7.

Im folgenden seien die Lipschitz-Konstanten immer bezüglich der Norm $\|\cdot\|_\infty$ definiert.

LEMMA 4.7. Sei $f|_{\overline{V}}$ Orts-Lipschitz-stetig mit Lipschitz-Konstante L. Dann ist die Abbildung $A : E_{\delta,\epsilon} \to E$ Lipschitz-stetig mit Konstante ϵL.

Beweis des Lemmas. Seien φ, $\psi \in E_{\delta,\epsilon}$.

$$\begin{aligned} A(\varphi)(t) - A(\psi)(t) &= \int_{t_0}^{t} \Big(f\big(\tau, \varphi(\tau)\big) - f\big(\tau, \psi(\tau)\big) \Big)\, d\tau \\ \Big\| A(\varphi)(t) - A(\psi)(t) \Big\|_\infty &\le \int_{t_0}^{t} \big\| f\big(\tau, \varphi(\tau)\big) - f\big(\tau, \psi(\tau)\big) \big\|_\infty \, d\tau \\ &\le \left| \int_{t_0}^{t} L \big\| \varphi(\tau) - \psi(\tau) \big\|_\infty \, d\tau \right| \\ &\le \epsilon L \sup_{\tau \in I} \|\varphi(\tau) - \psi(\tau)\|_\infty \end{aligned}$$

Also

$$\|\!| A(\varphi) - A(\psi) |\!\|_\infty \le \epsilon L \|\!| \varphi - \psi |\!\|_\infty.$$

□

LEMMA 4.8. Sei $C := \max\{\|f(t,x)\|_\infty \mid (t,x) \in \overline{V}\}$. Wenn wir

$$\epsilon \le \epsilon_0 := \min\Big\{ \frac{\delta}{2(C + \delta L)}, \frac{\delta}{2} \Big\} \tag{4.9}$$

wählen, dann gilt $A(E_{\delta,\epsilon}) \subset E_{\delta/2,\epsilon}$.

Beweis.

$$\begin{aligned}
\|\!|A(\varphi_0) - \varphi_0|\!\|_\infty &\leq \sup_{t\in I} \|A(\varphi_0)(t) - x_0\|_\infty \\
&= \sup_{t\in I} \left\| x_0 + \int_{t_0}^t f(\tau, x_0)\, d\tau - x_0 \right\|_\infty \\
&= \sup_{t\in I} \left| \int_{t_0}^t \|f(\tau, x_0)\|_\infty \, d\tau \right| \\
&\leq \sup_{t\in I} |t - t_0| C = \epsilon C.
\end{aligned}$$

Also gilt für beliebiges $\varphi \in E_{\delta,\epsilon}$:

$$\begin{aligned}
\|\!|A(\varphi) - \varphi_0|\!\|_\infty &\leq \|\!|A(\varphi) - A(\varphi_0)|\!\|_\infty + \|\!|A(\varphi_0) - \varphi_0|\!\|_\infty \\
&\leq \epsilon L \|\!|\varphi - \varphi_0|\!\|_\infty + \epsilon C \\
&\leq \epsilon L\delta + \epsilon C \leq \frac{\delta}{2},
\end{aligned}$$

falls $\epsilon \leq \epsilon_0 \leq \delta/2$ wie oben gewählt wird. Es folgt $A(E_{\delta,\epsilon}) \subset E_{\delta/2,\epsilon}$. □

Aus (4.9) folgt auch $\epsilon L < 1/2$. Das heißt, dann ist $A : E_{\delta,\epsilon} \to E_{\delta,\epsilon}$ eine Kontraktion. Es gilt $\varphi_0 \in E_{\delta,\epsilon} \neq \emptyset$. Nach dem Banachschen Fixpunktsatz (Satz 2.5 in Kapitel 8) besitzt also A einen eindeutigen Fixpunkt in $E_{\delta,\epsilon}$. Den Fixpunkt nennen[13] wir φ_ϵ. Insbesondere haben wir nun den Existenzteil des Theorems gezeigt.

[13]Wir schreiben φ_ϵ um klar zu machen, dass wir noch nichts darüber wissen, ob es von ϵ abhängt. Achtung: wir haben auch noch nicht geklärt, ob die Lösung von δ abhängt. Und selbst bei L können wir verschiedene Konstanten wählen, und die wählbaren Konstanten hängen auch wieder von der Wahl von δ ab. Kurzum: φ_ϵ könnte noch von vielen Wahlen abhängen, und wir werden nun sehen, dass es dies eigentlich gar nicht tut.

Ist nun $0 < \tilde{\epsilon} < \epsilon$, dann ist

$$\tilde{\varphi} := (\varphi_\epsilon)|_{(t_0-\tilde{\epsilon};t_0+\tilde{\epsilon})} \in E_{\delta,\tilde{\epsilon}}$$

eine Lösung von (4.4) mit $\tilde{\varphi}(t_0) = x_0$ und deswegen gilt $\varphi_{\tilde{\epsilon}} = \tilde{\varphi}$.

Das Theorem (=der Satz von Picard-Lindelöf) ist nun bewiesen, sobald wir das folgende Lemma haben. Dann ist nämlich auch die im Theorem behauptete Eindeutigkeit gezeigt.

LEMMA 4.10. Sei $\rho \in (0; \epsilon_0]$ und $\psi : (t_0 - \rho; t_0 + \rho) \to \mathbb{R}^n$ eine Lösung von (4.4) mit $\psi(t_0) = x_0$, $(t, \psi(t)) \in U \ \forall t$. Dann gilt $\psi \in E_{\delta,\rho}$.

Aus dem Lemma folgt nämlich, dass ψ ein und damit der einzige Fixpunkt von $A|_{E_{\delta,\rho}}$ ist.

Beweis. Wir zeigen, dass für alle $t \in (t_0 - \rho; t_0 + \rho)$ gilt:

$$\|\psi(t) - x_0\|_\infty < \delta.$$

Angenommen diese Ungleichung gilt nicht für alle solchen t. Dann ist

$$T := \left\{ s \in [0, \rho) \,\middle|\, \|\psi(t_0 + s) - x_0\|_\infty = \delta \text{ oder } \|\psi(t_0 - s) - x_0\|_\infty = \delta \right\}$$

eine nicht-leere abgeschlossene[14] Teilmenge von $[0, \rho)$.

Setze $\tilde{\rho} := \min T$, und wir haben $0 < \tilde{\rho} < \rho \leq \epsilon_0$. Wir erhalten $\|\psi(t) - x_0\|_\infty < \delta$ für $t \in (t_0 - \tilde{\rho}; t_0 + \tilde{\rho})$ und somit $\psi|_{(t_0-\tilde{\rho};t_0+\tilde{\rho})} \in E_{\delta,\tilde{\rho}}$, also insbesondere

$$\psi(t) = \varphi_{\epsilon_0}(t) \quad \forall t \in (t_0 - \tilde{\rho}; t_0 + \tilde{\rho}) \tag{4.11}$$

[14]Da $s \mapsto \|\psi(t_0 + s) - x_0\|_\infty$ stetig ist und $\{\delta\}$ abgeschlossen in $\mathbb{R}$.

Auf Grund der Stetigkeit von ψ und φ_{ϵ_0} folgt die Gültigkeit von (4.11) für $t \in [t_0-\tilde{\rho}; t_0+\tilde{\rho}]$. Es folgt der Widerspruch

$$\delta \overset{\tilde{\rho}\in T}{=} \|\psi(t_0 \pm \tilde{\rho}) - x_0\|_\infty = \|\varphi_{\epsilon_0}(t_0 \pm \tilde{\rho}) - x_0\|_\infty \leq \frac{\delta}{2}.$$

□

Aus Lemma 4.10 folgt dann auch: Lösungen von (4.4) mit dem selben Anfangswert stimmen auf dem Durchschnitt ihrer Definitionsbereiche überein.[15]

Der Beweis des Theorems liefert noch mehr

- Ein numerisches[16] Berechnungsverfahren für die Lösung. Definiere rekursiv: $\varphi_0(t) = x_0$, $\varphi_{k+1} := A(\varphi_k)$. Dann konvergiert φ_k gegen die Lösung. (Ohne Beweis!)
- Information darüber wie $\varphi(t)$ vom Anfangswert x_0 abhängt, siehe folgender Zusatz 4.13.

Beispiele 4.12.

(a) Ist $f : U \to \mathbb{R}^n$ stetig differenzierbar, dann ergibt der Satz von Picard-Lindelöf zu jedem Anfangswert $\varphi(t_0) = x_0$ eine Lösung $\varphi : I \to \mathbb{R}^n$ von $\varphi'(t) = f(\varphi(t))$. Wir werden sehen: es gibt ein maximales Intervall, auf dem die Funktion φ definiert werden kann.

[15]Die oben konstruierte Zahl ϵ_0 hängt von der Wahl von δ ab. Die Lösung φ ist aber im folgenden Sinn von δ unabhängig: Sind Lösungen zu zwei verschiedenen Wahlen von δ gegeben, so stimmen diese Lösungen auf einer Umgebung von t_0 überein.

[16]„Numerisch“ bedeutet: man bekommt keine explizite Formel für die Lösung heraus, dafür aber ein Verfahren, das in ein Computer-Programm umsetzbar ist und das dann beliebig genaue Näherungslösungen liefert.

(b) $n = 1$, $f(t,x) = \sqrt[3]{t}$ ist Orts-Lipschitz-stetig. Also erfüllt $\varphi'(t) = \sqrt[3]{t}$, $\varphi(t) \in \mathbb{R}$ die Voraussetzungen des Satzes von Picard-Lindelöf. Man kann explizit die Lösung bestimmen:

$$\varphi(t) = \frac{3}{4}t^{4/3}.$$

(c) $n = 1$, $f(t,x) = \sqrt[3]{x}$ ist nicht Orts-Lipschitz-stetig. Picard-Lindelöf ist nicht anwendbar. Dennoch gibt es Lösungen. Falls $\varphi(t) \neq 0$:

$$\begin{aligned}
& \varphi'(t) = \sqrt[3]{\varphi(t)} \\
\iff \quad & \frac{\varphi'(t)}{\sqrt[3]{\varphi(t)}} = 1 \\
\iff \quad & \frac{d}{dt}\frac{3}{2}(\varphi(t))^{2/3} = 1 \\
\iff \quad & \exists c \in \mathbb{R} : \frac{3}{2}(\varphi(t))^{2/3} = t + c \\
\iff \quad & \exists c \in \mathbb{R} : \varphi(t) = (\frac{2}{3}(t+c))^{3/2}.
\end{aligned}$$

Dies liefert Lösungen auf $(-c, \infty)$. Nun ist

$$\varphi(t) = \begin{cases} (\frac{2}{3}(t+c))^{3/2} & t > -c \\ 0 & t \leq -c \end{cases}$$

stetig differenzierbar. Zum Anfangswert $\varphi(0) = 0$ gibt es viele Lösungen (zu jedem $c \geq 0$ eine). Weitere Lösungen

$$\varphi(t) = \begin{cases} -(\frac{2}{3}(t+c))^{3/2} & t > -c \\ 0 & t \leq -c \end{cases}.$$

Existenz, aber keine Eindeutigkeit!

In der Aussage des Satzes 4.3 von Picard-Lindelöf lässt sich nicht ablesen, wie ϵ_0 von x_0 abhängt. Wenn man den Beweis des Satzes von Picard-Lindelöf genau betrachtet, so erhält man aber folgendes:

ZUSATZ 4.13 (zum Satz von Picard-Lindelöf)**.** Es sollen die Voraussetzungen von Satz 4.3 von Picard-Lindelöf gelten. Jeder Punkt (t_0, x_0) besitzt dann eine Umgebung W und ein $\epsilon_0 > 0$, so dass für jedes $(t_1, x_1) \in W$ der Satz von Picard-Lindelöf eine eindeutige Lösung $\varphi : (t_1 - \epsilon_0; t_1 + \epsilon_0) \to \mathbb{R}^n$ mit Anfangswert $\varphi(t_1) = x_1$ liefert.

Beweis des Zusatzes (skizziert). Sei wieder $f : U \to \mathbb{R}^n$, $(t_0, y_0) \in U$ wie oben. Hierzu sei δ, ϵ wie im Beweis des Satzes 4.3 von Picard-Lindelöf konstruiert gewählt. Sei $(t_1, x_1) \in U$ mit $\|(t_1, x_1) - (t_0, x_0)\|_\infty \leq \delta/4$. Wenn man den Beweis des Satzes 4.3 von Picard-Lindelöf genau anschaut, so sieht man, dass die Differentialgleichung auch eine Lösung $\varphi_{t_1,y_1} : (t_1 - \epsilon_0; t_1 + \epsilon_0) \to \mathbb{R}^n$ mit $\varphi_{t_1,x_1}(t_1) = x_1$ besitzt. Der Zusatz ist also richtig für $W := B^\infty_{\delta/4}\big((t_0, x_0)\big) \subset U$. □

Dasselbe ϵ_0 kann also zu allen Startwerten in W gewählt werden. Man beachte aber, dass wir weder hier noch im Satz von Picard-Lindelöf gefordert haben, dass ϵ_0 in irgendeinem Sinne maximal gewählt wird. Diese Abhängigkeit von den Anfangswerten wird später in Satz 4.20) wieder aufgegriffen.

Ende Mi 17.7.

Fr 19.7.

Wdh. Picard-Lindelöf.

Sei U offen in $\mathbb{R} \times \mathbb{R}^n$, $f : U \to \mathbb{R}^n$ Orts-Lipschitz-stetig und stetig. Zu jedem (t_0, x_0) gibt es

- $\epsilon > 0$ und

- Eine eindeutige C^1-Funktion $\varphi_{t_0,x_0} : (t_0 - \epsilon; t_0 + \epsilon) \to \mathbb{R}^n$, so dass für alle $t \in (t_0 - \epsilon; t_0 + \epsilon)$ gilt:

$$(t, \varphi_{t_0,x_0}(t)) \in U \text{ und } \varphi'_{t_0,x_0}(t) = f(t, \varphi_{t_0,x_0}(t))$$

Ab jetzt: Lösung = Lösung von $\varphi'(t) = f(t, \varphi(t))$.

Satz 4.14 (Satz von Picard-Lindelöf auf maximalen Intervallen)**.** Seien U, f, t_0 und x_0 wie im Satz von Picard-Lindelöf. Setze

$$t_{\max} := \sup\{t > t_0 \mid \text{Es existiert } \epsilon > 0 \text{ und eine Lösung } \varphi : (t_0 - \epsilon; t) \to \mathbb{R}^n$$
$$\text{von (4.4) mit Anfangswert } x_0\} \in (t_0; \infty]$$
$$t_{\min} := \inf\{t < t_0 \mid \text{Es existiert } \epsilon > 0 \text{ und eine Lösung } \varphi : (t; t_0 + \epsilon) \to \mathbb{R}^n$$
$$\text{von (4.4) mit Anfangswert } x_0\} \in [-\infty; t_0).$$

(1) Dann existiert auch eine Lösung $\varphi_{\max} : (t_{\min}; t_{\max}) \to \mathbb{R}^n$ mit Anfangswert $\varphi_{\max}(t_0) = x_0$. Man nennt $(t_{\min}; t_{\max})$ das maximale Lösungs-Intervall.
(2) Jede Lösung $\varphi : I \to \mathbb{R}^n$ mit demselben Anfangswert erfüllt $I \subset (t_{\min}; t_{\max})$ und $\varphi = \varphi_{\max}|_I$.

Beweis. **1. Schritt:** Sind $\varphi : I \to \mathbb{R}^n$ und $\tilde{\varphi} : \tilde{I} \to \mathbb{R}^n$ mit $\varphi(t_0) = \tilde{\varphi}(t_0) = x_0$; I, $\tilde{I}$ offene Intervalle. Wir zeigen, dass φ und $\tilde{\varphi}$ auf $I \cap \tilde{I}$ übereinstimmen.

Wir definieren

$$T := \{t \in I \cap \tilde{I} \mid \varphi(t) = \tilde{\varphi}(t)\}.$$

Da $h : t \mapsto \varphi(t) - \tilde{\varphi}(t)$ stetig ist und da $\{0\}$ abgeschlossen in $\mathbb{R}$, ist $T = h^{-1}(\{0\})$ ebenfalls abgeschlossen in $I \cap \tilde{I}$. Wegen $t_0 \in T$ wissen wir $T \neq \emptyset$. Wir zeigen nun, dass T offen in $I \cap \tilde{I}$ ist. Da $I \cap \tilde{I}$ zusammenhängend ist, folgt dann $T = I \cap \tilde{I}$ und somit der erste Schritt.

Sei also $t_1 \in T$. Wähle $\epsilon_1 > 0$ so, dass $(t_1-\epsilon_1; t_1+\epsilon_1) \subset I \cap \tilde{I}$. Nach dem Satz 4.3 von Picard-Lindelöf gibt es ein $\epsilon \in (0; \epsilon_1]$, so dass es genau eine Lösung $\hat{\varphi} : J := (t_1 - \epsilon; t_1 + \epsilon) \to \mathbb{R}^n$ mit Anfangswert $\hat{\varphi}(t_1) = \varphi(t_1)$ gibt. Also

$$\hat{\varphi} \equiv \varphi|_J \equiv \tilde{\varphi}|_J.$$

Somit $J \subset T$. Wir haben gesehen, dass T offen ist. Der erste Schritt ist somit beendet.

2. Schritt: Wir wollen nun die Lösung $\varphi_{\max} : (t_{\min}; t_{\max}) \to \mathbb{R}^n$ mit Anfangswert $\varphi_{\max}(t_0) = x_0$ definieren.

Zu $t \in (t_{\min}; t_{\max})$ wählen wir eine Lösung $\varphi : I \to \mathbb{R}^n$, I offenes Intervall, $t_0, t \in I$ mit $\varphi(t_0) = x_0$. Dann setzen wir $\varphi_{\max}(t) = \varphi(t)$. Der erste Schritt garantiert, dass die Definition von $\varphi_{\max}(t)$ nicht davon abhängt, wie wir I und φ gewählt haben. Und für derartige φ gilt dann

$$\varphi_{\max}|_I = \varphi.$$

Insbesondere ist nun $\varphi_{\max}$ auf $(t_{\min}; t_{\max})$ wohldefiniert und eine stetig differenzierbare Lösung. □

PROPOSITION 4.15. Seien U, f, (t_0, x_0) wie im Satz von Picard-Lindelöf. Sei φ eine Lösung mit $\varphi(t_0) = x_0$ mit maximalem Lösungs-Intervall $(t_{\min}; t_{\max})$.

Gilt $t_{\max} < \infty$, dann gibt es **kein** x_∞ mit $(t_{\max}, x_\infty) \in U$ und

$$x_\infty = \lim_{t \to t_{\max}} \varphi(t).$$

Analog für $t_{\min} > -\infty$.

In anderen Worten: Wenn der Grenzwert existiert, dann ist er nicht in U.[17]

Beweis. Wir betrachten nur die Aussage um $t_{\max}$.

Angenommen solch ein x_∞ existiert. Es existiert nun ein $\epsilon > 0$ und eine Umgebung V von $(t_{\max}, x_\infty)$ in $\mathbb{R} \times \mathbb{R}^n$, so dass für alle $(s, y) \in V$ gilt: es gibt eine Lösung $\varphi_{s,y} : (s - \epsilon; s + \epsilon) \to \mathbb{R}^n$ mit $\varphi_{s,y}(s) = y$.

Es existiert nun ein $\delta > 0$ so dass für alle $t \in (t_{\max} - \delta; t_{\max})$: $(t, \varphi(t)) \in V$. Wähle nun ein t_1 mit $t_{\max} - \epsilon/2 < t_1 < t_{\max}$, $t_{\max} - \delta < t_1$.

Wir erhalten nun eine Lösung

$$\hat{\varphi}(t) := \begin{cases} \varphi(t) & \text{für } t_{\min} < t \leq t_1 \\ \varphi_{t_1,\varphi(t_1)}(t) & \text{für } t \geq t_1 \end{cases}$$

Ein passendes Bild

die auf dem Intervall $[t_{\min}; t_1 + \epsilon)$ definiert ist, aber $t_1 + \epsilon > t_{\max}$. Dies ist ein Widerspruch. □

Bemerkung. Unter den Voraussetzung von Proposition 4.15 erhalten wir mit genau den gleichen Argumenten wie im obigen Beweis sogar eine etwas stärkere Aussage:

Gilt $t_{\max} < \infty$, dann gibt es **keine** Folge $t_i \nearrow t_{\max}$, so dass

$$x_\infty = \lim_{i \to \infty} \varphi(t_i)$$

[17] Es gibt ein Beispiele für alle denkbaren Fälle: $t_{\max} < \infty$ und Grenzwert existiert außerhalb U. $t_{\max} < \infty$ und Grenzwert existiert nicht. Im Falle $t_{\max} = \infty$ kann man keine Aussage machen. Ein uneigentlicher Grenzwert kann nicht existieren, außerhalb von U existieren oder in U existieren. Analog für $t_{\min}$.

existiert und $(t_{\max}, x_\infty) \in U$. Insbesondere folgt hieraus: Im Fall $t_{\max} < \infty$, gilt für $t_0 \in (t_{\min}; t_{\max})$

$$U \cap \overline{\{(t, \varphi(t)) \mid t_0 \leq t < t_{\max}\}}$$

ist nicht kompakt.

BEMERKUNG 4.16 (Unter Nutzung des nächsten Abschnitts, der in der Vorlesung zeitlich zu vor behandelt wurde.). Ist f stetig differenzierbar, und φ eine Lösung von $\varphi'(t) = f(t, \varphi(t))$. Dann erhalten wir aus der Kettenregel $\varphi \in C^2$ und

$$\varphi''(t) = \frac{\partial}{\partial t} f(t, \varphi(t)) + \frac{\partial}{\partial x} f(t, \varphi(t)) \cdot \varphi'(t)\,.$$

Dies kann man bis zur k-ten Ableitung iterieren, falls $f \in C^k$. Wenn man nun die so erhaltene gewöhnliche Differentialgleichung k-ter Ordnung studiert, erhält man wichtige Aussagen über die Ableitungen von φ. Insbesondere: ist $f \in C^k$, dann ist $\varphi \in C^{k+1}$.

FRAGE 4.17. Hängt die Lösung stetig von t_0 und x_0 ab?

LEMMA 4.18 (Gronwall)**.** Sei I ein Intervall, $C, L \in \mathbb{R}_{\geq 0}$, $g : I \to \mathbb{R}$ stetig, $a \in I$ und es gelte für alle $t \in I$

$$g(t) \leq \begin{cases} C + L \int_a^t g(\tau)\, d\tau & \text{falls } t \geq a \\ C + L \int_t^a g(\tau)\, d\tau\,, & \text{falls } t \leq a\,. \end{cases}$$

Dann gilt für alle $t \in I$:

$$g(t) \leq e^{L\,|t-a|}\, C\,.$$

Beweis. Wir betrachten zunächst $t \geq a$. Sei $\psi(t) := C + L \int_a^t g(\tau)\, d\tau$, das heißt $g(t) \leq \psi(t)$ für $t \in I$ mit $t \geq a$. Nach dem Hauptsatz der Differenial- und Integralrechnung ist ψ diffbar

mit $\psi'(t) = Lg(t) \leq L\psi(t)$. Wir rechnen

$$\frac{d}{dt}\left(\psi(t)\, e^{-L(t-a)}\right) \;=\; \underbrace{\psi'(t)}_{\leq L\psi(t)}\, e^{-L(t-a)} - L\psi(t))\, e^{-L(t-a)} \leq 0\,.$$

Durch Integration über das Intervall $[a,t]$ folgt

$$\psi(t)\, e^{-L(t-a)} \leq \psi(a)\, e^{-L(a-a)} = C\,,$$

und somit die Behauptung.

Der Fall $t \leq a$ folgt, in dem wir den bereits beweisenen Fall auf $\tilde{g}(\tau) := g(2a-\tau)$ anwenden. □

PROPOSITION 4.19 (Lösungsabschätzung)**.** Sei $U \subset \mathbb{R}\times\mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$ Orts-Lipschitz-stetig mit Lipschitz-Konstante L und stetig. Sei I ein Intervall und $a \in I$. Seien $\varphi_1, \varphi_2 : I \to \mathbb{R}^n$ Lösungen von

$$\varphi'(t) = f(t, \varphi(t))\,.$$

(Also insbesondere seien beide differenzierbar und für alle $t \in I$ gelte $(t,\varphi_1(t)) \in U$ und $(t,\varphi_2(t)) \in U$.) Dann gilt für alle $t \in I$:

$$\|\varphi_1(t) - \varphi_2(t)\|_\infty \leq e^{L\,|t-a|}\, \|\varphi_1(a) - \varphi_2(a)\|_\infty\,.$$

Beweis. Mit Lemma 4.6 erhalten wir

$$\varphi_1(t) - \varphi_2(t) = \varphi_1(a) - \varphi_2(a) + \int_a^t \underbrace{\left(f(\tau,\varphi_1(\tau)) - f(\tau,\varphi_2(\tau))\right)}_{\leq L\|\varphi_1(\tau)-\varphi_2(\tau)\|}\, d\tau\,.$$

Hieraus erhalten wir daraus

$$\|\varphi_1(t) - \varphi_2(t)\|_\infty \leq \|\varphi_1(a) - \varphi_2(a)\|_\infty + L\left|\int_a^t \|\varphi_1(\tau) - \varphi_2(\tau)\|_\infty\, d\tau\right|.$$

Wir wenden das Lemma von Gronwall für $g(t) := \|\varphi_1(t) - \varphi_2(t)\|_\infty$ und $C := \|\varphi_1(a) - \varphi_2(a)\|_\infty$ an und erhalten die Aussage. □

Satz 4.20 (Stetige Abhängigkeit von den Anfangsbedinungen)**.** Es sollen die Voraussetzungen und Definitionen aus dem Satz 4.3 von Picard-Lindelöf und dem Zusatz 4.13 gelten. Das heißt: Seien $U \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ lokal Orts-Lipschitz-stetig und stetig. Sei $(t_0, x_0) \in U$. Es sei eine offene Umgebung W von (t_0, x_0) in U und ein $\epsilon_0 > 0$ gegeben, so dass es zu jedem Anfangswert $(t_1, x_1) \in W$ eine Lösung $\varphi_{t_1,x_1} : (t_0 - \epsilon_0; t_0 + \epsilon_0) \to \mathbb{R}^n$ von

$$\varphi'(t) = f(t, \varphi(t))$$

gibt.[18] Dann hängt $\varphi_{t_1,x_1}(t)$ auch stetig von t_1 und x_1 ab. Genauer: die Funktion

$$\Phi : (t_0 - \epsilon_0; t_0 + \epsilon_0) \times W \to \mathbb{R}^n, \quad (t, t_1, x_1) \mapsto \varphi_{t_1,x_1}(t) = \varphi_{t_0,x_1}(t - t_1 + t_0)$$

ist stetig.

Beweis. Man sieht die Folgenstetigkeit der Funktion Φ ganz einfach aus der vorangehenden Proposition. □

Skizze eines alternativen Beweises.[19] (Details siehe z.B. [**11**, Chap. 1, Sec. 5])

Man nutzt Zusatz 4.13 um ein gemeinsames Definitionsintervall für alle (t_1, x_1) nahe (t_0, x_0) zu finden. Man definiert iterativ für $|t - t_1| < \epsilon/2$ (für das ϵ aus Beweis des

[18] Dieselbe Aussage gilt zunächst für $\varphi_{t_1,x_1} : (t_1 - \epsilon_0; t_1 + \epsilon_0) \to \mathbb{R}^n$ und daraus folgt obige Aussage nach Verkleinerung von ϵ_0 und W.

[19] Der alternative Beweis erfordert im Beweis des Satzes von Picard-Lindelöf etwas höheren Abstraktionsaufwand (Gleichmäßigkeit in weiteren Variablen, ist aber letztendlich effektiver geführt. Der oben aufgeführte ist anschaulicher, sobald man Picard-Lindelöf als bewiesen akzeptiert.

Satzes 4.3 von Picard-Lindelöf):

$$\Phi_0(t, t_1, x_1) := x_1, \qquad \Phi_{j+1}(t, t_1, x_1) := x_1 + \int_{t_1}^{t} f(\tau, \Phi_j(\tau, t_1, x_1))\, d\tau.$$

Man zeigt iterativ, dass Φ_j stetig ist für alle j und zeigt dann, dass Φ_j lokal gleichmäßig gegen Φ konvergiert. Die Stetigkeit ist nun für $|t - t_1| < \epsilon/2$ gezeigt. Um es für alle t, t_1 zu zeigen wählt man Zahlen $t(0), \ldots, t(L)$ mit $t(0) = t_1$, $t(L) = t$ und $|t(i) - t(i-1)| < \epsilon/2$. Die Stetigkeit von Φ folgt dann aus der Verkettung von stetigen Funktionen wie folgt

$$\Phi(t, t_1, y) = \Phi(t(L), t(L-1), \Phi(t(L-1), t(L-2), \Phi(\ldots, \Phi(t(1), t(0), x_1)\ldots).$$

□

5. Picard-Lindelöf für gewöhnliche Differentialgleichungen höherer Ordnung

Wir schreiben wieder (a, b) für das Intervall.

Aus dem Satz von Picard-Lindelöf ergibt sich unmittelbar durch Nutzen der Reduktion auf Systeme erster Ordnung (Bemerkung 2.6 (1)):

Satz 5.1. Gegeben sei eine explizite Differentialgleichung k-ter Ordnung

$$\varphi^{(k)}(t) = f(t, \varphi(t), \ldots, \varphi^{k-1}(t))$$

$f : U \to \mathbb{R}^n$ stetig und lokal Orts-Lipschitz-stetig[20], $U \subset \mathbb{R} \times (\mathbb{R}^n)^k$ offen. Sei $(t_0, \varphi_0, \varphi_1, \ldots, \varphi_{k-1}) \in U$, dann gibt es ein $\epsilon_0 \in \mathbb{R}_{>0}$, so dass es für jedes $\epsilon \in (0, \epsilon_0]$ genau eine genau eine Lösung $\varphi : (t_0 - \epsilon, t_0 + \epsilon) \to \mathbb{R}^n$ mit $\varphi(t_0) = \varphi_0$, $\varphi'(t_0) = \varphi_1$, $\ldots$ und $\varphi^{(k-1)}(t_0) = \varphi_{k-1}$.

[20]Das heißt: Lipschitz-stetig in allen Komponenten (außer der t-Komponente), und die Lipschitz-Konstante kann stetig in t gewählt werden.

Wenn wir mit maximalen Intervallen arbeiten und Satz 4.14 anwenden, dann erhalten wir diese Version:

Satz 5.2. Gegeben sei eine explizite Differentialgleichung k-ter Ordnung

$$\varphi^{(k)}(t) = f(t, \varphi(t), \dots, \varphi^{k-1}(t))$$

$f : U \to \mathbb{R}^n$ stetig und lokal Orts-Lipschitz-stetig[21], $U \subset \mathbb{R}\times(\mathbb{R}^n)^k$ offen. Sei $(t_0, \varphi_0, \varphi_1, \dots, \varphi_{k-1}) \in U$, dann gibt es genau eine auf einem maximalen Intervall definiert Lösung φ mit $\varphi(t_0) = \varphi_0$, $\varphi'(t_0) = \varphi_1$, ... und $\varphi^{(k-1)}(t_0) = \varphi_{k-1}$. Jede auf einem beliebigen Intervall definierte Lösung erhält man Restriktion aus der obigen.

6. Lineare gewöhnliche Differentialgleichungen

Definition 6.1. Sei I ein offenes Intervall, $n \in \mathbb{N}$. Gegeben seien stetige Abbildungen

$$A_0, A_1, \dots, A_{k-1} : I \to \mathbb{R}^{n\times n}$$

und $b : I \to \mathbb{R}^n$. Eine Gleichung der Form

$$\varphi^{(k)}(t) = A_{k-1}(t)\varphi^{(k-1)}(t) + \dots A_1(t)\varphi'(t) + A_0(t)\varphi(t) + b(t) \tag{6.2}$$

nennt man eine explizite lineare gewöhnliche Differentialgleichung k-ter Ordnung; und eine Lösung ist ein $\varphi \in C^k(J, \mathbb{R}^n)$, J ein offenes Intervall in I, so dass (6.2) für alle $t \in J$ gilt. Man nennt (6.2) homogen, falls $b \equiv 0$, sonst ist es inhomogen.

[21]Das heißt: Lipschitz-stetig in allen Komponenten (außer der t-Komponente), und die Lipschitz-Konstante kann stetig in t gewählt werden.

Reduktion auf Ordnung 1 (Wiederholung/Spezialfall)

Schreiben wir $\Phi(t) = \begin{pmatrix} \varphi(t) \\ \varphi'(t) \\ \vdots \\ \varphi^{(k-1)}(t)) \end{pmatrix} \in \mathbb{R}^{nk}$.

Ist φ eine Lösung von (6.2), so erfüllt das oben definierte Φ die Gleichung

$$\Phi'(t) = \begin{pmatrix} 0 & \mathbb{1}_n & 0 & \cdots & 0 \\ 0 & 0 & \mathbb{1}_n & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \cdots & \mathbb{1}_n \\ A_0(t) & A_1(t) & A_2(t) & \cdots & A_{k-1}(t) \end{pmatrix} \Phi(t) + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}. \tag{6.3}$$

Umgekehrt: ist Φ eine Lösung von (6.3), dann bilden die ersten n-Komponenten von Φ eine Lösung von (6.2). Es reicht also, sich auf Gleichungen erster Ordnung einzuschränken.

Ab jetzt $k = 1$, d.h.

$$\varphi'(t) = A(t)\varphi(t) + b(t) \qquad \text{kurz: } \varphi' = A\varphi + b \tag{6.4}$$

Das Wort „Lösung“ benutzen wir in diesem Abschnitt immer im Sinn „stetig differenzierbare Lösung“.

Zusätzlich zu den obigen Voraussetzungen seien von nun an A und b stetig.

Satz 6.5. Sei $J \neq \emptyset$ ein offenes Teil-Intervall von I. Eine Lösung $\varphi : J \to \mathbb{R}^n$ lässt sich zu einer auf I definierten Lösung fortsetzen.

Beweis später.

Eine Möglichkeit, diesen Satz zu zeigen, ist, geeignete Abschätzungen zu machen und dann Proposition 4.15 zu verwenden. Siehe zum Beispiel [**23**], Abschnitt 4.3 und die Vorarbeiten zuvor.

Wir wollen einen anderen Weg beschreiten und werden sehen, dass wir bald den Satz ohne Rechenaufwand zeigen können. Unser Weg hat auch den Vorteil, dass wir dann nicht Zusatz 4.13 nutzen, dessen Beweis nur skizziert wurde.

Da wir in den folgenden Zeilen den Satz noch nicht nutzen können, müssen wir alle nun folgenden Aussagen so formulieren, dass wir keine Probleme erhalten, falls eine Lösung nicht auf ganz I definierbar sein sollte. Sobald dann aber Satz 6.5 gezeigt ist, können wir davon ausgehen, dass dann alle Lösungen zu Lösungen $I \to \mathbb{R}^n$ erweitert werden können.

Die zu (6.4) gehörende homogene Differentialgeleichung ist:

$$\varphi'(t) = A(t)\varphi(t) \tag{6.6}$$

Lemma 6.7. Sei J ein offenes Teil-Intervall von I.

(1) Sei $\hat{\varphi}$ eine auf J definierte Lösungen der inhomogenen Gleichung (6.4). Dann gilt

$$\begin{aligned} & \tilde{\varphi} \text{ ist eine auf } J \text{ definierte Lösung der homogenen Gleichung (6.6)} \\ \iff\ & \hat{\varphi} + \tilde{\varphi} \text{ ist eine auf } J \text{ definierte Lösung der inhomogenen Gleichung (6.4)} \end{aligned}$$

(2) Die auf J definierten Lösungen von (6.6) bilden einen Untervektorraum von $C^1(J, \mathbb{R}^n)$. Wir nennen ihn $\mathcal{L}_J$.

Beweis. Nachrechnen. □

LEMMA 6.8. Zu jedem $t_0 \in I$ gibt es ein $\epsilon > 0$, so dass gilt: zu jedem $v \in \mathbb{R}^n$ existiert genau eine auf $(t_0 - \epsilon, t_0 + \epsilon)$ definierte Lösung φ von (6.4) mit $\varphi(t_0) = v$.

Beweis. Die Funktion $f(t, y) := A(t)y + b(t)$ ist Orts-Lipschitz-stetig. Nach dem Satz 4.3 von Picard-Lindelöf gibt es Zahlen $\hat{\epsilon} > 0$, $\epsilon_1 > 0$,..., $\epsilon_n > 0$ und folgende Lösungen:

- Eine C^1-Funktion $\hat{\varphi} : (t_0 - \tilde{\epsilon}, t_0 + \tilde{\epsilon}) \to \mathbb{R}^n$ mit $\hat{\varphi}' = A\hat{\varphi} + b$, $\hat{\varphi}(t_0) = 0$.
- Für $i \in \{1, \ldots, n\}$: C^1-Funktionen $\varphi_i : (t_0 - \epsilon_i, t_0 + \epsilon_i) \to \mathbb{R}^n$ mit $\varphi_i' = A\varphi_i$, $\varphi_i(t_0) = e_i$.

Setze nun $\epsilon := \min\{\tilde{\epsilon}, \epsilon_1, \ldots, \epsilon_n\} > 0$.

Zu jedem $v = (v_1, \ldots, v_n)^T \in \mathbb{R}^n$ ist

$$\varphi := \tilde{\varphi} + \sum_{i=1}^{n} v_i \varphi_i$$

eine auf $(t_0 - \epsilon, t_0 + \epsilon)$ definierte Lösung von

$$\varphi' = A\varphi + b, \quad \varphi(t_0) = v.$$

nach der Eindeutigkeitsaussage im Satz 4.3 von Picard-Lindelöf ist dies die einzige Lösung. □

Insbesondere ist die Auswertungsabbildung

$$\iota_{t_0} : \mathcal{L}_{(t_0-\epsilon, t_0+\epsilon)} \to \mathbb{R}^n, \qquad \varphi \mapsto \varphi(t_0)$$

ein Isomorphismus (von Vektorräumen) und

$$\dim \mathcal{L}_{(t_0-\epsilon, t_0+\epsilon)} = n.$$

Sind J_1, J_2 Intervalle mit $\emptyset \neq J_1 \subset J_2 \subset I$, dann ist wegen der Eindeutigkeitsaussage im Satz 4.3 von Picard-Lindelöf (bzw. der daraus gefolgerten Aussage Satz 4.14 (2)) die Restriktionsabbildung

$$\mathcal{L}_{J_2} \to \mathcal{L}_{J_1}, \qquad \varphi \mapsto \varphi|_{J_1}$$

injektiv.[22] Also

$$\dim \mathcal{L}_{J_2} \leq \dim \mathcal{L}_{J_1} \leq n. \tag{6.9}$$

Beweis von Satz 6.5. Sei $\varphi : (t_1, t_2) \to \mathbb{R}^n$ eine Lösung von (6.4), $(t_1, t_2) \subset I$.

Im Fall $t_2 \in I$, werden wir zeigen, dass sich die Lösung auf ein größeres offenes Intervall fortsetzt, das t_2 enthält. Analoges gilt für $t_1 \in I$. Wenn also (t_1, t_2) das maximale Lösungs-Intervall ist, so muss $(t_1, t_2) = I$ gelten.

Sei also $t_2 \in I$. Bestimme $\epsilon > 0$ und $\tilde{\varphi}$ wie im Beweis des letzten Lemmas für $t_0 := t_2$. Dann ist

$$\varphi|_{(t_2-\epsilon,t_2)} - \tilde{\varphi}|_{(t_2-\epsilon,t_2)} \in \mathcal{L}_{(t_2-\epsilon,t_2)}.$$

Wegen den obigen Überlegungen (6.9) gilt

$$n = \dim \mathcal{L}_{(t_2-\epsilon,t_2+\epsilon)} \leq \dim \mathcal{L}_{(t_2-\epsilon,t_2)} \leq n.$$

Also gilt überall „=“ und die Restriktion $\mathcal{L}_{(t_2-\epsilon,t_2+\epsilon)} \to \mathcal{L}_{(t_2-\epsilon,t_2)}$ ist surjektiv. Es gibt also ein $\psi \in \mathcal{L}_{(t_2-\epsilon,t_2+\epsilon)}$, mit

$$\psi|_{(t_2-\epsilon,t_2)} = \varphi|_{(t_2-\epsilon,t_2)} - \tilde{\varphi}|_{(t_2-\epsilon,t_2)}.$$

[22] Sobald der Satz gezeigt ist, steht hier überall „=“.

Somit ist

$$\hat{\varphi}(t) := \begin{cases} \varphi(t) & t \in (t_1, t_2) \\ \tilde{\varphi}(t) + \psi(t) & t \in [t_2, t_2 + \epsilon) \end{cases}$$

eine stetig differenzierbare Lösung von $\varphi' = A\varphi + b$, die auf $(t_1, t_2 + \epsilon)$ definiert ist. Nach den Ausführungen zu Beginn des Beweises folgt daraus der Satz. □

FOLGERUNG 6.10.

$$\{\text{Auf } I \text{ definierte Lösungen von } \varphi' = A\varphi + b\} \to \mathbb{R}^n, \qquad \varphi \mapsto \varphi(t_0)$$

ist bijektiv und $\dim \mathcal{L}_I = n$.

FOLGERUNG 6.11. Gegeben seien $\varphi_1, \ldots, \varphi_n \in \mathcal{L}_I$, $t_0 \in I$. Dann gilt

$$(\varphi_1, \ldots, \varphi_n) \text{ ist Basis von } \mathcal{L}_I \quad \Longleftrightarrow \quad \det(\varphi_1(t_0), \ldots, \varphi_n(t_0)) \neq 0.$$

Diese Determinante nennt man Wronski-Determinante. Sie ist hilfreich, um zu prüfen, ob man schon alle Lösungen der homogenen DGl gefunden hat.

Falls es eine Basis ist, so nennt man $\Phi = (\varphi_1, \ldots, \varphi_n) : I \to \mathbb{R}^{n \times n}$ eine Fundamentallösung von $\varphi' = A\varphi$.

~~**Lösungs-Strategie zur Lösung von** (6.4)~~

1.) Bestimme die Menge aller Lösungen der homogenen DGl.

2.) Dann finde **eine** Lösung der inhomogenen DGl.

Mit Hilfe von Lemma 6.7 erhalten wir dann alle Lösungen der inhomogenen DGl.

Wie erhält man alle Lösungen der homogenen DGl?

Fr 12.7. $n = 1$: [23] Lösungen von $\varphi' = A\varphi$ verschwinden entweder nie oder überall (Eindeutigkeit

in Picard-Lindelöf), o.B.d.A. sei $\varphi(t) \neq 0 \quad \forall t$.

$$\begin{aligned} \varphi' = A\varphi \qquad &\iff \qquad \frac{\varphi'}{\varphi} = A \\ &\iff \qquad \frac{d}{dt} \log|\varphi(t)| = A(t) \\ \varphi(t) &= C \exp\left(\int_{t_0}^{t} A(\tau)\, d\tau\right). \end{aligned}$$

Alle Lösungen sind von dieser Form mit $C \in \mathbb{R}$.

$n > 1$: im allgemeinen häufig schwer lösbar.

Spezialfall: A konstant. O.B.d.A. $I = \mathbb{R}$.

Dann ist die Menge der Lösungen $\varphi' = A\varphi$

$$\{t \mapsto \varphi(t) = \exp(At)\varphi_0 \mid \varphi_0 \in \mathbb{R}^n\}$$

Hierbei ist

$$\exp(At) := \sum_{j=0}^{\infty} \frac{1}{j!} A^j t^j \in \mathbb{R}^{n\times n}.$$

Diese Reihe konvergiert absolut: wir rechnen $\|AB\|_\infty \leq n\|A\|_\infty \|B\|_\infty$, $\|A^j\|_\infty \leq n^{j-1}\|A\|_\infty^j$

$$\begin{aligned} \left\| \frac{1}{j!} A^j t^j \right\|_\infty &\leq \frac{1}{j!} n^{j-1} \|A\|_\infty^j |t|^j := a_j \\ \frac{a_j}{a_{j-1}} &= \frac{n}{j} \|A\|_\infty |t| \to 0. \end{aligned}$$

$$\frac{d}{dt} \exp(tA) = A \exp(tA) = \exp(tA) A\,.$$

[23] Dies ist ein Spezialfall der Methode "getrennte Variable„, siehe Abschnitt 8.

$$\exp(0 \cdot A) = \exp(0_n) = \mathbb{1}_n \,.$$

$\Phi(t) := \exp(At)$ ist eine Fundamentallösung.

Wir werden dies im nächsten Abschnitt nochmals betrachten.

Wie erhält man eine Lösung der inhomogenen DGl?

PROPOSITION 6.12 (Variation der Konstanten)**.** $A \in C^0(I, \mathbb{R}^{n\times n})$, $b \in C^0(I, \mathbb{R}^n)$, $t_0 \in I$. Sei $\Phi : I \to \mathbb{R}^{n\times n}$ eine Fundamentallösung von $\varphi'(t) = A(t)\varphi(t)$. Dann ist

$$t \mapsto \Phi(t) \int_{t_0}^{t} \Phi(\tau)^{-1} b(\tau)\, d\tau$$

eine Lösung von $\varphi'(t) = A(t)\varphi(t) + b(t)$.

Beweis.

$$\begin{aligned}
& \frac{d}{dt} \left(\Phi(t) \int_{t_0}^{t} \Phi(\tau)^{-1} b(\tau)\, d\tau \right) \\
= \;& \Phi'(t) \int_{t_0}^{t} \Phi(\tau)^{-1} b(\tau)\, d\tau + \Phi(t)\Phi(t)^{-1} b(t) \\
= \;& A(t)\Phi(t) \int_{t_0}^{t} \Phi(\tau)^{-1} b(\tau)\, d\tau + b(t) \\
= \;& A(t) \left(\Phi(t) \int_{t_0}^{t} \Phi(\tau)^{-1} b(\tau)\, d\tau \right) + b(t)
\end{aligned}$$

□

7. Lineare gewöhnliche Differentialgleichungen mit konstanten Koeffizienten

!ACHTUNG!. Dieser Abschnitt ist noch nicht in latex aufgeschrieben und es ist unklar, ob und gegebenenfalls wann dies getan wird. Bitte benutzen Sie Ihre Vorlesungsmitschrift der Vorlesung vom 12.7..

Beispiele 7.1.

(1) Phasenportrait von $\varphi'(t) = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix} \varphi(t)$.

Lösungen: $\varphi(t) = \begin{pmatrix} e^{-2t} & 0 \\ 0 & e^{-t} \end{pmatrix} \varphi_0$, $\varphi_0 \in \mathbb{R}^2$.

Ein passendes Bild

(2) Phasenportrait von $\varphi'(t) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \varphi(t)$.

Lösungen: $\varphi(t) = \begin{pmatrix} e^{t} & 0 \\ 0 & e^{t} \end{pmatrix} \varphi_0$, $\varphi_0 \in \mathbb{R}^2$.

Ein passendes Bild

(3) Phasenportrait von $\varphi'(t) = \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix} \varphi(t)$.

Lösungen: $\varphi(t) = \begin{pmatrix} e^{-2t} & 0 \\ 0 & e^{t} \end{pmatrix} \varphi_0$, $\varphi_0 \in \mathbb{R}^2$.

Ein passendes Bild

(4) Phasenportrait von $\varphi'(t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \varphi(t)$.

Lösungen: $\varphi(t) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \varphi_0$, $\varphi_0 \in \mathbb{R}^2$.

Ein passendes Bild

(5) Phasenportrait von $\varphi'(t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \varphi(t)$.

Lösungen: $\varphi(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \varphi_0$, $\varphi_0 \in \mathbb{R}^2$.

Ein passendes Bild

8. Differentialgleichungen mit getrennten Variablen

Mi 24.7.

DEFINITION 8.1. Gegeben seien offene Intervalle I und J, und $g : I \to \mathbb{R}$ und $h : J \to \mathbb{R}$ stetig. Eine Differentialgleichung der Form

$$\varphi'(t) = g(t)h(\varphi(t))$$

nennt man eine gewöhnliche Differentialgleichung mit getrennten Variablen[24]. Kurzschreibweise

$$\varphi' = g(t)h(\varphi). \tag{8.2}$$

Der Satz 4.3 von Picard-Lindelöf ist **nicht** anwendbar, da

$$f(t, x) = g(t)h(x)$$

im allgemeinen[25] nicht Orts-Lipschitz-stetig ist.

[24] oder auch „mit getrennten Veränderlichen“, da eine Variable dasselbe wie eine Veränderliche ist.

[25] „im allgemeinen“ heißt hier: es gibt Funktionen g und h, so dass dies nicht erfüllt ist. Beispiel $g(t) = 1$, $h(x) = \sqrt[3]{x}$.

Schmutzige Rechnung (streng mathematisch nicht gerechtfertigt, aber in der Praxis oft erfolgreich)

$$\begin{aligned} & \frac{d\varphi}{dt} &&= g(t)h(\varphi) \\ \iff \quad & \frac{d\varphi}{h(\varphi)} &&= g(t)\,dt \\ \iff \quad & \int \frac{d\varphi}{h(\varphi)} &&= c + \int g(t)\,dt \end{aligned}$$

Ist nun H eine Stammfunktion von $1/h$ und G eine Stammfunktion von g, dann sollte es eine Konstante $c \in \mathbb{R}$ geben mit

$$H(\varphi(t)) = c + G(t),$$

also $\varphi(t) = H^{-1}(c + G(t))$.

Bestimme nun c so, dass Anfangsdaten erfüllt sind.

Jetzt eine exakte Herleitung:

Satz 8.3 (Trennung der Variablen)**.** Gegeben sei eine gewöhnliche Differentialgleichung mit getrennten Variablen wie in Definition 8.1, $h(x) \neq 0$ auf J. Sei $t_0 \in I$, $x_0 \in J$. Die Stammfunktion von $t \mapsto g(t)$ sei $G : I \to \mathbb{R}$ und die von $x \mapsto 1/h(x)$ sei $H : J \to \mathbb{R}$. Dann gibt es bis auf Restriktion genau eine stetig differenzierbare Lösung von

$$\varphi'(t) = g(t)h\big(\varphi(t)\big) \qquad \varphi(t_0) = x_0,$$

die auf einer offenen Umgebung von t_0 definiert ist, nämlich

$$\varphi(t) = H^{-1}(G(t) + H(x_0) - G(t_0)). \tag{8.4}$$

Beweis. Offensichtlich ist $t \mapsto 1/h(x)$ stetig. Also existiert H, ist stetig differenzierbar und monoton. Wir wenden Proposition 1.6 aus Kapitel 5 auf die Funktion $H : J \to \mathbb{R}$ an und erhalten wegen $H'(t) = 1/h(t) \neq 0$:

- $K := H(J)$ ist ein offenes Intervall,
- $H^{-1} : K \to J$ existiert,
- H^{-1} ist differenzierbar mit $(H^{-1})'(s) = 1/(1/h(H^{-1}(s))) = h(H^{-1}(s))$.

Also auch $H^{-1} \in C^1(K)$. Wegen $G \in C^1(I, \mathbb{R})$ ist die in Gleichung (8.4) definierte Funktion φ auf der offenen Menge

$$\{t \in I \mid G(t) + H(x_0) - G(t_0) \in K\} \ni t_0$$

in C^1. Man rechnet leicht nach, dass φ eine Lösung von (8.2) mit $\varphi(t_0) = x_0$ ist.

Nehmen wir, dass φ eine Lösung von (8.2) mit $\varphi(t_0) = x_0$ ist. Dann gilt

$$g(t) = \frac{\varphi'(t)}{h(\varphi(t))} = \frac{d}{dt}\left(H(\varphi(t))\right),$$

also ist $G(t) - H(\varphi(t))$ eine Konstante, nämlich $G(t_0) - H(x_0)$. Und dies ergibt (8.4).

Beispiel 8.5. $g : \mathbb{R} \to \mathbb{R}$, $g(t) = 1$. $h : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $h(x) = -x^2 \neq 0$. Die Differentialgleichung

$$\varphi'(t) = -\varphi(t)^2$$

besitzt unter anderem die Lösung $\varphi \equiv 0$. Jede andere Lösung wird nie Null (Eindeutigkeit in Picard-Lindelöf). Im Fall $\varphi \not\equiv 0$ können wir umformulieren in

$$\begin{aligned} & -\frac{\varphi'(t)}{\varphi(t)^2} = 1 \\ \iff & \frac{d}{dt}\frac{1}{\varphi(t)} = 1 \\ \iff & \frac{1}{\varphi(t)} = t + c \\ \iff & \varphi(t) = \frac{1}{t+c} \end{aligned}$$

Beim Anfangswert $\varphi(0) = \alpha < 0$ bekommen wir $c = 1/\alpha$ und das maximale Lösungs-Intervall ist $(-\infty, 1/|\alpha|)$. Die Lösung hört also bei $1/|\alpha|$ auf zu existieren. Wir haben gesehen, dass dieses Phänomen bei den linearen Gleichungen nie auftreten kann.

Beispiel 8.6. $\varphi'(t) = A(t)\varphi(t)$. NICHT GETIPPT...

Beispiel 8.7. $\varphi'(t) = \pi(\frac{1}{2}t^3 + \frac{1}{4}t)(4 + \varphi(t)^2)$. NICHT GETIPPT...

9. Mehr zum Fluss eines zeitabhängigen Vektorfelds

Sei wieder $U \subset \mathbb{R} \times \mathbb{R}^n$ eine offene Teilmenge und $f : U \to \mathbb{R}^n$ Orts-Lipschitz-stetig und stetig.

Wir definieren jetzt

$$\mathcal{D} \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n$$

durch $(\tilde{t}, s, y) \in \mathcal{D} \iff (s, y) \in U$ und es existiert eine Lösung $\varphi_{s,y} : I \to \mathbb{R}^n$ von

$$\varphi'(t) = f(t, \varphi(t)) \tag{9.1}$$

mit $\varphi_{s,y}(s) = y$ und $\tilde{t} \in I$.[26] Es folgt aus Zusatz 4.13:

$$\{(t,t,y) \mid (t,y) \in U\} \subset \mathring{\mathcal{D}}$$

Wir sehen später, dass $\mathcal{D}$ offen in $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^n$ ist, also $\mathring{\mathcal{D}} = \mathcal{D}$.

Wir definieren nun den Fluss Φ von f wie folgt:

$$\Phi : \mathcal{D} \to \mathbb{R}^n, \quad \Phi(t,s,y) := \varphi_{s,y}(t).$$

Dann gilt offensichtlich

$$\Phi(t,t,y) = y.$$

Lemma 9.2.

$$(t,s,y) \in \mathcal{D} \quad \Longleftrightarrow \quad (s,t,\Phi(t,s,y)) \in \mathcal{D}$$

und

$$\Phi(s,t,\Phi(t,s,y)) = y.$$

Beweis. Nach Definition von Φ gilt $\Phi(t,s,y) = \varphi_{s,y}(t)$. Daraus folgt

$$\varphi_{s,y} \equiv \varphi_{t,\Phi(t,s,y)}.$$

Wir erhalten dann an der Stelle s

$$y = \varphi_{s,y}(s) = \varphi_{t,\Phi(t,s,y)}(s) = \Phi(s,t,\Phi(t,s,y)).$$

□

So ähnlich zeigt man

$$\Phi(r,t,\Phi(t,s,y)) = \Phi(r,s,y),$$

[26] In anderen Worten $\tilde{t}$ liegt im zugehörigen maximalen Definitionsbereich.

falls alle Ausdrücke wohldefiniert sind.

Begründung:

$$\Phi(r,s,y) = \varphi s, y(r) = \varphi_{t,\Phi(t,s,y)}(r) = \Phi(r,t,\Phi(t,s,y))\,.$$

LEMMA 9.3. $\mathcal{D}$ ist offen.

Beweisskizze. [27] Sei $(t,s,y) \in \mathcal{D}$. Wir wollen eine offene Umgebung von (t,s,y) konstruieren, die in $\mathcal{D}$ enthalten ist. O.B.d.A. $t \geq s$. Wähle zu jedem $\tau \in [s;t]$ ein $\epsilon(\tau) > 0$ mit $B_{\epsilon(\tau)}((\tau,\tau,\Phi(\tau,s,y)) \subset \mathcal{D}$. Dies liefert eine offene Überdeckung $((\tau-\epsilon(\tau);\tau+\epsilon(\tau))_{\tau\in[s;t]}$ von $[s;t]$. Wähle mit Proposition 6.17 aus Kapitel 8 zu dieser Überdeckung eine Lebesguesche Zahl $\epsilon > 0$, und ein $m \in \mathbb{N}$, $m > 1/\epsilon$. Wir definieren

$$t_i := s + \frac{i}{m}(t-s).$$

Wir schreiben

$$\begin{cases} \varphi_1(s,y) := \Phi(t_1,s,y) \\ \varphi_i(y) := \Phi(t_i,t_{i-1},y) \qquad 2 \leq i \leq m-1 \\ \varphi_m(t,y) := \Phi(t,t_{m-1},y) \end{cases}$$

Wir können nun schreiben:

$$\begin{aligned} \Phi(\tilde{t},\tilde{s},\tilde{y}) &= \Phi(\tilde{t},t_{m-1},\Phi(t_{m-1},t_{m-2},\Phi(\ldots\Phi(t_1,\tilde{s},\tilde{y}) \\ &= \varphi_m(\tilde{t},\varphi_{m-1}(\varphi_{m-2}(\ldots\varphi_2(\varphi_1(\tilde{s},\tilde{y}))\ldots))) \end{aligned}$$

[27] Diese Beweisskizze wurde in der Vorlesung aus Zeitgründen übersprungen.

Wir betrachten beide Seiten der Gleichung als Funktion in $\tilde{t}$, $\tilde{s}$ und $\tilde{y}$, also $(\tilde{t}, \tilde{s}, \tilde{y}) \mapsto \mathbb{R}^n$. Dies ist eine „Verkettung“ von stetigen Funktionen, wobei die Funktionen nicht überall definiert sind. Ist $(\tilde{t}, \tilde{s}, \tilde{y})$ im Definitionsbereich der Verkettung, dann auch in $\mathcal{D}$. Im folgenden sei $D(h)$ der Definitionsbereich einer Funktion h. Der Definitionsbereich der Verkettung $(\tilde{s}, \tilde{y}) \mapsto \varphi_2(\varphi_1(\tilde{s}, \tilde{y}))$ ist

$$D(\varphi_1) \cap \big(\varphi_1\big)^{-1}\Big(D(\varphi_2)\Big),$$

also der Schnitt zweier offener Menge und somit wieder offen. Man zeigt dann induktiv, dass der Definitionsbereich der Verkettung

$$(\tilde{s}, \tilde{y}) \mapsto \varphi_{m-1}(\varphi_{m-2}(\dots \varphi_2(\varphi_1(\tilde{s}, \tilde{y}))\dots)))$$

die Menge

$$V := D(\varphi_1) \cap (\varphi_1)^{-1}\big(D(\varphi_2)\big) \cap (\varphi_2 \circ \varphi_1)^{-1}\big(D(\varphi_3)\big) \cap \dots (\varphi_{m-2} \circ \dots \circ \varphi_2 \circ \varphi_1)^{-1}\big(D(\varphi_{m-1})\big),$$

ist; und diese Menge ist als Schnitt offener Mengen ebenfalls offen. Der Definitionsbereich der Verkettung

$$(\tilde{t}, \tilde{s}, \tilde{y}) \mapsto \varphi_m(\tilde{t}, \varphi_{m-1}(\varphi_{m-2}(\dots \varphi_2(\varphi_1(\tilde{s}, \tilde{y}))\dots)))$$

ist dann

$$(\mathbb{R} \times V) \cap \big(\mathrm{id}_{\mathbb{R}} \times (\varphi_{m-1} \circ \dots \circ \varphi_2 \circ \varphi_1)\big)^{-1} D(\varphi_m),$$

also auch offen. □

PROPOSITION 9.4. Der Fluss $\Phi : \mathcal{D} \to \mathbb{R}^n$ ist stetig.

Beweisskizze. Wegen Satz 4.20 wissen wir bereits, zu jedem $(\tau, y) \in U$ gibt es eine Umgebung W von (τ, y) in U und ein $\epsilon_0 > 0$, so dass Φ auf $(\tau - \epsilon_0; \tau + \epsilon_0) \times W$ stetig ist. Man

beachten $\epsilon_0 = \epsilon_0(\tau, \epsilon)$ hängt von τ und ϵ ab. Zu der Überdeckung $\Big((\tau - \epsilon_0; \tau + \epsilon_0)\Big)_{\tau \in [t;s]}$ von $[t; s]$ wähle eine Lebesguesche Zahl $\epsilon > 0$, dann $m > 1/\epsilon$. Der restliche Beweis geht ähnlich, wie der Beweis der Offenheit von $\mathcal{D}$: wir schreiben lokal Φ als Verkettung vieler Abbildungen, bei denen wir bereits die Stetigkeit wissen. □

ZUSATZ 9.5. Ist $f \in C^k(U, \mathbb{R}^n)$, dann ist auch $\Phi \in C^k(\mathcal{D}, \mathbb{R}^n)$.

Anschaulich: ist $f \in C^k$, dann ist die Lösung $\varphi_{t_0,x_0}(t)$ ist nicht nur C^k in t sondern auch C^k in t_0 und x_0.[28]

Beweis. Dies zeigt man induktiv über k, indem man ähnlich wie in Bemerkung 4.16 argumentiert. □

KOROLLAR 9.6. Autonomer Fall: Sei U offen in $\mathbb{R}^n$, $f : U \to \mathbb{R}^n$, $\varphi'(t) = f(\varphi(t))$ und $\Phi : \mathcal{D} \to \mathbb{R}^n$ der zugehörige Fluss. Wir sagen f ist ein vollständiges Vektorfeld, falls $\mathcal{D} = \mathbb{R} \times \mathbb{R} \times U$. Dann ist $\Phi(t, s, y) = \Phi(t + c, s + c, y)$ und $\Phi(t, s, \cdot) : U \to U$ ist ein Homöomorphismus. (Denn $\Phi(s, t, \cdot) : U \to U$ ist die (stetige) Umkehrfunktion.)

Leicht zu prüfen. □

10. Einige Kommentare zum Langzeitverhalten von gew. Diffifferentialgleichungen

Fr. 26.7.

Wir haben mit dem Satz von Picard-Lindelöf gesehen, dass wir für kurze Zeiten Existenz und Eindeutigkeit von Lösungen von gewöhnlichen Differentialgleichungen mit gegebenen

[28] Aber die Formulierung im Zusatz ist noch etwas stärker!

Anfangswerten haben, vorausgesetzt, dass die Differentialgleichung in expliziter Form ist und die Gleichung stetig differenzierbar von den niedrigeren Ableitungen und der Zeit t abhängen. Gleiches gilt dann natürlich auch für die zugehörigen Flüsse.

Man interessiert sich nun aber auch für das langfristige Verhalten der Lösungen und der zugehörigen Flüsse. Wir haben bereits in Proposition 4.15 einige Ergebnisse zur Frage erhalten, wie lange Lösungen existieren. Wir betrachten nun den Fall, dass Lösungen φ für alle zukünftigen bzw. alle vergangenen Zeiten existieren, d.h. $t_{\max} = +\infty$ oder $t_{\min} = -\infty$. Wir betrachten der Einfachheit halber nur den autonomen, das heißt nicht explizit zeitabhängigen Fall.

Falls dann $\varphi(t)$ für $t \to \pm\infty$ gegen ein $x_{\pm\infty}$ konvergiert, und dieses[29] im Definitionsbereich der gewöhnlichen Differentialgleichung liegt, so ist existiert eine konstante Lösung $\varphi_\pm(t) := x_{\pm\infty}$. Dieser Fall wird im Rahmen der Stabilitätstheorie untersucht, über die wir im folgenden Abschnitt einen kleinen Überblick geben wollen.

Es kann aber auch vorkommen, dass φ periodisch ist, das heißt, es gibt ein kleinstes $L > 0$, die Periode, mit $\varphi(L + t) = \varphi(t)$ für alle t. Man fragt sich dann, ob Lösung in der Nähe ebenfalls periodisch sind. Andere Verhalten sind ebenfalls möglich, z.B. “quasiperiodisches Verhalten” oder “ergodisches Verhalten”, wobei man für letzteres manchmal auch das aufreißerischer klingende Wort “chaotisches Verhalten” nutzt. Die mathematischen Gebiete, die sich damit beschäfitgen nennen sich dann je nach Ausrichtung “dynamische Systeme”, “Chaos-Theorie”, “Ergodentheorie”, ...

[29]Bei Differentialgleichungen $\varphi^{(k)}(t) = f(\varphi(t), \dots, \varphi^{(k-1)}(t)$ der Ordnung k muss man hier natürlich überprüfen, ob $(x_{\pm\infty}, 0, \dots, 0) = \lim(\varphi(t), \varphi'(t), \dots \varphi^{(k-1)}(t))$ im Definitionsbereich der Funktion f liegt.

Eine reiche Quelle wichtiger und interessanter gewöhnlicher Differentialgleichungen ist die klassische Physik, hier insbesondere die Hamiltonsche Mechanik, die letztendlich auch die Himmelsmechanik umfasst. Wir wollen mal ein Beispiel geben, um anzudeuten, welche Art von Resultaten hier erwartet werden können.

Wenn man zum Beispiel das langfristige Verhalten unseres Sonnensystems betrachtet und die Entstehung des Lebens, die ja letztendlich einen über astronomische Zeiträume hinweg relativ konstanten Abstand zur Sonne benötigt, so erscheint es unabdingbar zu fordern, dass sich alle Planeten nahezu periodisch um die Sonne bewegen. Dieses Verhalten kann auch relativ leicht mathematisch hergeleitet werden, wenn nur ein Planet um die Sonne kreisen würde (Zwei-Körper-Problem). Es ist auch einfach zu zeigen, wenn man zwei Näherungen annimmt: zum einen nimmt man an, dass die Sonne auf grund der großen Masse sich nicht mehr bewegt und und zum anderen, dass sich die Planeten gegenseitig nicht anziehen. Wenn man diese Näherungen annimmt, dann hat man viele Erhaltungsgrößen, so dass man schließlich die Bewegung sehr gut beschreiben kann; man hat ein sogenanntes vollständig integrables dynamisches System. Die Lösungen solcher Systeme sind immer quasiperiodisch (sie sind nur dann periodisch, wenn die Verhältnisse der Umlaufzeiten der Planeten rationale Zahlen sind, sonst sind sie eben fast periodisch, was mit quasiperdidisch bezeichnet wird).

Diese beiden Näherungen sind aber leider in astronomischen Zeitskalen nicht mehr angebracht. Es handelt sich also um eine Störung eines vollständig integrablen Systems. Der Satz von Kolgomorow, Arnold und Moser aus den Jahren 1954-1963 besagt, dass fast alle Lösungen von derartigen Störungen von vollständig integrablen Systemen ebenfalls quasiperiodisch sind, vorausgesetzt, die Störungen sind nicht zu groß. Siehe zum Beispiel unter diesem Link für weitere Informationen:

https://de.wikipedia.org/wiki/Kolmogorow-Arnold-Moser-Theorem.

All diese Sätze werden wir aber leider im Rahmen des Analysis-Zyklus aus Zeitgründen nicht weiter behandeln können.

11. Stabilität und Verhalten in der Nähe von kritischen Punkten

In diesem Abschnitt wollen wir auf einige Beweise verzichten. In der Darstellung halten wir uns eng an [**23**], Abschnitt 4.5 und an [**35**], Kapitel VII. In diesen beiden Büchern findet man auch die fehlenden Beweise.

Im ganzen Abschnitt sei U eine offene Teilmenge von $\mathbb{R}^n$ und $v : U \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld.

Definition 11.1. $x_0 \in U$ heißt kritischer Punkt von v, falls $v(x_0) = 0$. Einen kritischen Punkt x_0 nennen wir stabil, falls für jede Umgebung $V \subset U$ von x_0 eine offene Umgebung $W \subset V$ von x_0 gibt, so dass jede Integralkurve φ von v mit $\varphi(0) \in W$ auf $[0, \infty)$ definiert werden kann, und so dass $\varphi([0, \infty)) \subset V$. Ein kritischer Punkt x_0 wird Attraktor genannt, falls er stabil ist und falls das obige W so gewählt werden kann, dass für alle Integralkurven φ von v mit $\varphi(0) \in W$ gilt $\lim_{t\to\infty} \varphi(t) = x_0$.

30

Beispiele 11.2. Wir setzen die Beispiele 7.1 fort.

[30]Man beachte, dass die hier verwendete Definition nicht ganz mit der in [**23**] verwendeten übereinstimmt. Wir haben Stabilität in die Definition des Attraktors hineingenommen, und sind damit in Übereinstimmung mit dem Großteil der fortgeschrittenen Literatur in diesem Gebiet. Der unten zitierte Beweis des Satzes von Poincaré-Ljapunow in [**23**] zeigt aber sogar unsere (stärkere) Version des Attraktors.

(1) In diesem Beispiel ist 0 ein Attraktor zum Vektorfeld

$$v : x \mapsto \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix} x$$

0 ist kein stabiler Punkt zum Vektorfeld $-v$.

(2) In diesem Beispiel ist 0 kein stabiler Punkt zum Vektorfeld

$$v : x \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x$$

0 ist Attraktor zum Vektorfeld $-v$.

(3) In diesem Beispiel ist 0 kein stabiler Punkt zum Vektorfeld

$$v : x \mapsto \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix} x$$

0 ist auch kein stabiler Punkt zum Vektorfeld $-v$.

(4) In diesem Beispiel ist 0 stabiler Punkt zum Vektorfeld

$$v : x \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x$$

0 ist auch stabiler Punkt zum Vektorfeld $-v$. Aber weder Attraktor zu v noch Attraktor zu $-v$.

(5) In diesem Beispiel ist 0 weder stabiler Punkt zum Vektorfeld

$$v : x \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x$$

noch ein stabiler Punkt zum Vektorfeld $-v$.

Allgemein gilt für $v(x) := Ax$: 0 ist Attraktor, genau dann, wenn jeder Eigenwert von A einen negativen Realteil hat. [31]

BEISPIEL 11.3. (aus [**23**])
Sei $v : \mathbb{R} \to \mathbb{R}$, $v(x) = -x^k$, $k \in \mathbb{N} \setminus \{0\}$. Der einzige kritische Punkt ist 0. Mit Trennung der Variablen sieht man, dass $\varphi' = v(\varphi)$ für $\varphi \neq 0$ äquivalent zu

$$\frac{d}{dt} \frac{1}{(k-1)\varphi(t)^{k-1}} = 1$$

ist, falls $k > 1$. Für $k = 1$ ähnlich mit Logarithmus. Mit Standard-Umformungen erhält man mit Anfangswert $\varphi(0) = x_1$

$$\varphi(t) = \begin{cases} x_1 e^{-t} & k = 1 \\ \frac{x_1}{\sqrt[k-1]{1+(k-1)x_1^{k-1}t)}} & k > 1 \end{cases}$$

Für $x_1 > 0$ oder k ungerade, existiert die Lösung für alle $t \in [0, \infty)$ und sie konvergiert gegen 0. Für k gerade und $x_1 < 0$ existiert die Lösung nur endlich lange. Deswegen ist 0 ein Attraktor, falls k ungerade, und 0 ist nicht einmal stabil, falls k gerade ist.

Ein passendes Bild

DEFINITION 11.4. Die Linearisierung von v in $x_0 \in U$ ist die Abbildung $D_{x_0}v : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto v'(x_0)x$.

Idee der folgenden Resultate: die Flusslinien (= Integralkurven) des Vektorfelds v sehen „ähnlich“ aus wie wie Flusslinien der Linearisierung. Das Phasenportrait in einer kleinen

[31] Hier wäre eigentlich ein Beweis nötig, zu dem wir keine Zeit mehr haben: Falls A diagonalisierbar ist, dann ist jede Lösung eine Linearkombination von Lösungen der Form $e^{t\lambda_0}\varphi_0$, mit $A\varphi_= \lambda_0\varphi_0$. Wenn A nicht diagonalisierbar ist, ist die Argumentation ähnlich.

Umgebung des kritischen Punkt ist eine „verbogene“ Version des Phasenportraits der Linearisierung.

THEOREM 11.5 (Poincaré-Ljapunow)**.** Sei x_0 ein kritischer Punkt von des C^1-Vektorfelds $v : U \to \mathbb{R}^n$. Für jeden komplexen Eigenwert λ der Matrix $v'(x_0)$ gelte $\mathbf{Re}\ \lambda < 0$. Dann ist x_0 ein Attraktor.

Beweis. Siehe [**23**], Abschnitt 4.5. □

DEFINITION 11.6. Eine Ljapunow-Funktion zum C^1-Vektorfeld $v : U \to \mathbb{R}^n$ und zum kritischen Punkt x_0 ist eine C^1-Funktion $L : U \to \mathbb{R}$, so dass [32]

- $L(x_0) = 0$
- $L(x) > 0$, falls $x \in U \setminus \{x_0\}$
- Für alle $x \in U$ gilt: $\partial_{x,v(x)} L \leq 0$. Oder es gilt für alle $x \in U$: $\partial_{x,v(x)} L \geq 0$

Beispiel: Ist E eine Erhaltungsgröße, die in x_0 ein absolutes Minimum hat, so ist $E - E(x_0)$ eine Ljapunow-Funktion.

Konsequenz: Ist φ eine Flusslinie von v, dann gilt

$$\frac{d}{dt} L(\varphi(t)) = L'(\varphi(t))\varphi'(t) = \partial_{\varphi(t),v(\varphi(t))} L \leq 0 \qquad \text{bzw. } \geq 0$$

THEOREM 11.7 (Ljapunow)**.** Sei v, x_0 und L wie oben.

(a) Falls $\forall x \in U : \partial_{x,v(x)} L \leq 0$, dann ist x_0 ein stabiler Punkt.
(b) Falls $\forall x \in U \setminus \{x_0\} : \partial_{x,v(x)} L < 0$, dann ist x_0 ein Attraktor
(c) Falls $\forall x \in U \setminus \{x_0\} : \partial_{x,v(x)} L > 0$, dann ist x_0 kein stabiler Punkt.

[32] Im folgenden ist $\partial_{x,v} f$ wieder die Richtungsableitung von f im Punkt x in Richtung v.

Beweis. Siehe [**23**], Abschnitt 4.5. □

Beispiel 11.8. Die Liénardsche Gleichung, siehe [**23**], Ende Abschnitt 4.5.

Theorem 11.9 (Linearisierungssatz von Grobman und Hartman, 1959–1960)**.** Sei x_0 ein kritischer Punkt eines stetig differenzierbaren Vektorfelds $v : U \to \mathbb{R}^n$. Wir nehmen an, dass für jeden Eigenwert λ von $v'(x_0)$ gilt: $\mathbf{Re}\,\lambda \neq 0$. Sei $D_{x_0}v : \mathbb{R}^n \to \mathbb{R}^n$ die Linearisierung von v in x_0. Dann gibt es eine offene Umgebung V von x_0 in U, eine offene Umgebung W von 0 und einen Homöomorphismus $h : V \to W$, $h(x_0) = 0$, so dass gilt

$$\begin{aligned} &\varphi : (a,b) \to V \text{ ist Flusslinie zu } v \\ \iff\ &h \circ \varphi : (a,b) \to W \text{ ist Flusslinie zu } D_{x_0}v \end{aligned}$$

Der Homöomorphismus h bildet Flusslinien auf Flusslinien ab. Die Flusslinien der Linearisierung sind die Abbildungen $t \mapsto e^{At}\varphi_0$ mit $\varphi_0 \in \mathbb{R}^n$, $A := v'(x_0)$.

Wir bekommen also

$$\varphi(t) = h^{-1}(e^{At}h(\varphi(0)))$$

vorausgesetzt wir sind „nahe genug“ bei x_0.

Wenn man das Theorem von Grobman und Hartman liest, so stellt sich die Frage, ob man nicht das Wort „Homöomorphismus“ durch „C^k-Diffeomorphismus“ ersetzen kann, sobald v genügend oft differenzierbar ist.

Man kann zeigen (siehe Abschnitt 2.9 von [**28**] für mehr Details):

- Ist $v \in C^2$, so gibt es einen C^1-Diffeomorphismus h mit den obigen Eigenschaften.
- Es gibt Vektorfelder $v \in C^\infty$, für die es keinen C^2-Diffeomorphismus h mit den obigen Eigenschaften gibt

ANHANG A

Mehr Details zu den Grundlagen der Logik

1. Das Russellsche Paradoxon

(Auch Russellsche Antinomie genannt. Vorläufer sind die Cantorsche Antinomie und das Burali-Forti-Paradoxon, siehe Wikipedia.)

Am Anfang der Vorlesung haben wir gesagt, dass unser Ziel ist, eine axiomatisch aufgebaute Mathematik zu erhalten. Diesem Anspruch sind wir bisher nicht gerecht geworden. Immer wieder wurde die Intuition genutzt, um Objekte zu definieren, anstatt wirklich saubere mathematische Definitionen hinzuschreiben. Das Vorgehen war bis weit ins 19. Jahrhundert ganz ähnlich. Das intuitive Verwenden des Mengenbegriffs führte dann aber zu dem Russellschen Paradoxon.

FRAGE 1.1. Gibt es die Menge aller Mengen?

Stand der Mathematik gegen Ende des 19. Jahrhunderts: Nach Cantors Beschreibung (Beschreibung 3.1) sollte es die Menge $\mathcal{M}$ aller Mengen geben. Das Elementzeichen $\in$ ist dann eine auf $\mathcal{E}\times\mathcal{M}$ definierte Aussageform, wobei $\mathcal{E}$ die Menge aller Elemente irgendeiner Menge bezeichne, also $\mathcal{E} = \bigcup \mathcal{M}$.

$$\begin{array}{rcl} \mathcal{E} \times \mathcal{M} & \longrightarrow & \{w, f\} \\ (x, M) & \mapsto & x \in M \end{array}$$

Russell bildete nun daraus

$$\mathcal{N} := \{x \in \mathcal{M} \mid x \notin x\}.$$

Gilt nun $\mathcal{N} \in \mathcal{N}$? Für alle Mengen x gilt

$$x \in \mathcal{N} \quad \Longleftrightarrow \quad x \notin x$$

und wenn wir $x := \mathcal{N}$ setzen, so erhalten wir

$$\mathcal{N} \in \mathcal{N} \quad \Longleftrightarrow \quad \mathcal{N} \notin \mathcal{N}.$$

Aussagen in der Form $A \leftrightarrow \neg A$ sind aber immer falsch, also haben wir einen Widerspruch erhalten.

Um die Mengenlehre und somit die Mathematik sauber auf Axiomen aufzubauen, müssen wir also viel genauer als bisher vorgehen und intuitive Argumente durch formal korrekte Beweise ersetzen.

Wie löst man dieses Paradoxon auf: Nicht jede „Zusammenfassung von bestimmten wohlunterschiedenen Objekten“ ist eine Menge. Die Axiome sollen regeln, wann solche eine Zusammenfassung eine Menge ist. Insbesondere ist die „Zusammenfassung“ $\mathcal{M}$ aller Mengen selbst keine Menge, sondern etwas größeres, das wir Klasse nennen wollen. Analog sind auch $\mathcal{N}$ und $\mathcal{E}$ keine Mengen, sondern Klassen.

2. Axiomatische Mengenlehre

Wir wollen nun skizzieren, wie man die Mengenlehre sauber auf Axiomen aufbauen kann. Man nimmt hierzu mehrere Axiomen an, all diese Axiome zusammen nennt man dann ein Axiomen-System. Historisch sind hier wichtige Fortschritte zu Ende zu 19. Jahrhunderts und zu Beginn des 20. Jahrhunderts gemacht worden. Nun haben verschiedene Mathematiker verschiedene Axiomensysteme entwickelt und benutzt. Das wohl bekannteste Axiomensystem heißt ZFC nach den Mathematikern Zermelo und Fraenkel und C für „Axiom of **C**hoice“ (= Auswahlaxiom). Wir stellen ein anderes vor (NBG), das sehr ähnliche Eigenschaften hat. Man kann zeigen, dass man mit ZFC und NBG dieselben Sätze über Mengen herleiten kann. Die Axiome von ZFC machen oft rein mengentheoretische Argumente einfacher, während die Axiome von NBG für das praktische Arbeiten oft besser geeignet sind.

Grundlegende Ideen:

- Man will gar nicht mehr erklären, was eine Menge ist, sondern fordert axiomatische Eigenschaften für Mengen.
- Man nimmt an, dass auch alle Elemente von Mengen ebenfalls Mengen sind. In anderen Worten: jede Menge ist ein Mengensystem über einer geeigneten Menge. In der Notation des letzten Abschnitts gilt also nun $\mathcal{M} = \mathcal{E}$. Die Zusammenfassung aller Regensburger Autos ist dann keine Menge. Der Satz „Es existiert kein Auto“ ist also in einer streng axiomatisch aufgebauten Vorlesung über Logik und Mengenlehre eine wahre Aussage, denn ein Auto ist ja keine Menge von Mengen.
- Trotz dieser Einschränkungen kann man immer noch natürliche, ganze, rationale und reelle Zahlen und vieles mehr als Mengen definieren. Die Zahl π ist also eine

Menge, deren Elemente wieder Mengen sind, deren Elemente wieder Mengen sind, ... Siehe dazu auch die Peano-Axiome in Abschnitt B in Kapitel „Zahlen“.

Die Axiome übernehme ich aus [**17**], siehe auch [**25**]. Zum Vergleich zu ZFC verweise ich auf [**13**].

Axiome 2.1 (Axiome der Mengenlehre nach von Neumann, Bernays, Gödel (NBG))**.**

- Es gibt zwei Sorten von Objekten, Mengen und Klassen. Außerdem gibt es noch die Beziehung $\in$, die besagt, wann eine Klasse das Element einer anderen Klasse ist.
- Jede Menge ist eine Klasse.
- Alle Elemente von Klassen sind Mengen.
- Extensionalität. Zwei Klassen sind genau dann gleich, wenn sie dieselben Elemente enthalten.
- Komprehension. Ist C eine auf allen Mengen definierte Aussagenform, die nach strikt festgelegten Regeln aufgebaut ist, die wir hier der Kürze halber nicht genauer diskutieren.[1] Dann ist

  $$\big\{x \text{ ist eine Menge} \,\big|\, C(x)\big\}$$

  eine Klasse.
- Die leere Klasse $\emptyset := \{x \text{ ist eine Menge} \mid x \neq x\}$ ist eine Menge.
- Aussonderungsmengenaxiom. Jede Teilklasse einer Menge ist eine Menge.
  Benutzt wurde hierbei die folgende Definition: Eine Klasse A ist eine Teilklasse einer Klasse B, wenn jedes Element von A ein Element von B ist.

[1] Siehe hierzu [**17**] und [**13**]. Gemeint ist eine Aussagenform 1. Stufe, in der nicht über Klassenvariable quantifiziert wird.

- Paarmengenaxiom. Sind A und B Mengen, so ist auch $\{A, B\}$ eine Menge.
- Vereinigungsaxiom. Ist A eine Menge, so ist auch

$$\bigcup A := \{x \mid \exists y \in A : \, x \in y\}$$

  eine Menge, die Vereinigungsmenge von A.
- Potenzmengenaxiom. Ist A eine Menge, so ist auch $\mathcal{P}(A) := \{x \mid x \subset A\}$ eine Menge, die Potenzmenge von A.
- Fundierungsaxiom. Ist $A \neq \emptyset$, so gibt es ein $x \in A$ mit $x \cap A = \emptyset$.
- Ersetzungsaxiom.Sind X und Y Klassen und ist F eine funktionale „Relationsklasse“[2] mit $D(F) = X$ und $B(F) \subset Y$. Ist die Teilklasse $A \subset X$ eine Menge, so ist auch $B(F|_A)$ eine Menge.
- Unendlichkeitsaxiom. Es gibt eine induktive Menge; eine Menge A heißt induktiv, wenn $\emptyset \in A$ und wenn für alle $x \in A$ auch $x \cup \{x\} \in A$ ist.
- Auswahlaxiom. Sei I eine nicht-leere Menge und $(M_i)_{i \in I}$ eine Familie von nicht-leeren Mengen. Dann ist $\prod_{i \in I} M_i$ nicht die leere Menge.

Bemerkungen 2.2.

(a) Besonders interessant ist das Auswahlaxiom. Es folgt bereits aus den Axiomen zuvor, dass $\prod_{i \in I} M_i$ eine Menge ist. Falls I endlich ist, so gilt offensichtlich $\prod_{i \in I} M_i \neq \emptyset$. Für unendliche Indexmengen kann man aber $\prod_{i \in I} M_i \neq \emptyset$ im allgemeinen nicht aus den anderen Axiomen herleiten. Wenn man es annimmt, „weiß“ man zwar, dass $\prod_{i \in I} M_i$ mindestens ein Element besitzt, es kann aber vorkommen, dass kein einziges Element von $\prod_{i \in I} M_i$ angegeben werden kann.

[2]Unter einer Relationsklasse verstehen wir eine Klasse, deren Elemente Paare sind. Man kann dann die üblichen Begriffe einer Relation definieren, z.B. „funktional“, Definitionsbereich, Bild, etc.

Es ist nun ganz lehrreich, zu untersuchen, welche Aussagen der Mathematik nur unter Hinzunahme des Auswahlaxioms gezeigt werden können, und für welche Aussagen, das Auswahlaxiom nötig ist. Angenommen wir hätten nur die Axiome bis zum Unendlichkeitsaxiom. Dann kann man zeigen, dass das Auswahlaxiom äquivalent zur Aussage „Jeder Vektorraum besitzt eine Basis“ ist.

Wenn wir das Auswahlaxiom annehmen, dann gibt es einen Vektorraum, der zwar eine Basis besitzt, aber zugleich kann keine Basis angegeben werden.

Das Auswahlaxiom ist auch äquivalent zur Aussage: „Sind X und Y Mengen, dann gibt es eine injektive Abbildung $f : X \longrightarrow Y$ oder es gibt eine injektive Abbildung $f : Y \longrightarrow X$“.

(b) Man kann aus den Axiomen der Mengenlehre nicht folgern, dass die Mengenlehre widerspruchsfrei ist! (Zweiter Gödelscher Unvollständigkeitssatz). Die Mathematik beruht auf der Annahme, dass die Axiome der Mengenlehre widerspruchsfrei sind.

(c) Es gibt Aussagen, von denen gezeigt werden kann, dass weder diese Aussage noch die Negation dieser Aussage beweisbar ist. Die Cantorsche Kontinuumshypothese ist solch eine Aussage:

„Sei M eine Menge, so dass es **keine** injektive Abbildung $M \longrightarrow \mathbb{N}$ gibt.
Dann gibt es eine injektive Abbildung $\mathbb{R} \longrightarrow M$.“

Solche Aussagen gibt es in jedem hinreichend großen logischen System (Erster Gödelscher Unvollständigkeitssatz).

Bemerkung 2.3. Das Auswahlaxiom ist „unabhängig“ von den anderen Axiomen der Mengenlehre. Aufgrund der Nicht-Konstruktivität und etwas ungewöhnlicher Konsequenzen wird im Normalfall explizit angegeben, wenn ein Beweis das Auswahlaxiom verwendet.

Aus den obigen Axiomen kann man nun Schritt für Schritt die ganze Mathematik heraus herleiten.

ANHANG B

Die Peano-Axiome

1. Die Axiome und erste Konsequenzen

Die natürlichen Zahlen wurden von Dedekind (1888) und Peano (1889) axiomatisiert. Siehe [**17**, Kapitel 3 und 4], [**20**], oder [**13**, Kapitel V] für mehr Details.

Axiome 1.1 („Axiome“ der natürlichen Zahlen (Peano-Axiome))**.** Gegeben sei

- eine Menge N,
- ein (ausgewähltes) Element in N, das wir 0 nennen,
- eine Abbildung $s : N \longrightarrow N$, $x \mapsto s(x)$, genannt die Nachfolger-Abbildung.

Wir sagen, dass $(N, 0, s)$ die Peano-Axiome erfüllt, falls gilt:

(P1) $0 \notin B(s)$

(P2) Die Abbildung $s : N \longrightarrow N$ ist injektiv.

(P3) (Induktionsaxiom) Erfüllt $T \subset N$ die Bedingungen $0 \in T$ und $s_{\#}(T) \subset T$, dann gilt bereits $T = N$.

Wenn $(N, 0, s)$ die Peano-Axiome erfüllt, sagt man $(N, 0, s)$ ist ein Modell der natürlichen Zahlen.

Proposition 1.2. Es gibt eine Menge N ein Element $0 \in N$ und eine Nachfolger-Abbildung $s : N \longrightarrow N$, so dass $(N, 0, s)$ die Peano-Axiome erfüllt.

Beweis folgt unten.

Proposition 1.3. Wenn $(\widetilde{N}, \widetilde{0}, \widetilde{s})$ und $(\widehat{N}, \widehat{0}, \widehat{s})$ die Peano-Axiome erfüllen, dann gibt es eine bijektive Abbildung $F : \widetilde{N} \longrightarrow \widehat{N}$, so dass $F(\widetilde{0}) = \widehat{0}$ und $F \circ \widetilde{s} = \widehat{s} \circ F$.

Man sagt dann oft: $(\widetilde{N}, \widetilde{0}, \widetilde{s})$ und $(\widehat{N}, \widehat{0}, \widehat{s})$ sind kanonisch isomorph.

Beweis später.

Die letzte Gleichung schreibt man am besten als Diagramm, ein sogenanntes kommutatives Diagramm

$$\begin{array}{ccc} \widetilde{N} & \xrightarrow{F} & \widehat{N} \\ \downarrow{\scriptstyle \widetilde{s}} & & \downarrow{\scriptstyle \widehat{s}} \\ \widetilde{N} & \xrightarrow{F} & \widehat{N} \end{array}$$

Alles, was man üblicherweise mit natürlichen Zahlen macht (Addition, Multiplikation, Teilbarkeit, Primzahlen, etc.) beruht letztendlich auf dem Tripel $(N, 0, s)$ und den Peano-Axiomen. Deswegen besagt Prop. 1.3, dass alle Eigenschaften von $(N, 0, s)$ die man aus den Peano-Axiomen herleiten kann, entweder in allen Modellen gelten oder in keinem. Beispiele: In jedem Modell gibt es unendlich viele Primzahlen, n ist eine Zahl mit 4 Teilern in $(\widetilde{N}, \widetilde{0}, \widetilde{s})$ genau dann wenn $F(n)$ eine Zahl mit 4 Teilern in $(\widehat{N}, \widehat{0}, \widehat{s})$ ist. Es kann uns also egal sein, welches Modell die natürlichen Zahlen beschreibt, wichtig sind allein die Peano-Axiome.

BEMERKUNG 1.4. Sind die Peano-Axiome weitere Axiome? Oder eine Definition? Oder Aussagen?
Es gibt hier zwei verschiedene Sichtweisen. Je nach Anwendung und Dozent wird die eine oder andere bevorzugt.

(a) Die mengentheoretische Sichtweise: Die Peano-Axiome sind eine **Definition**.
Proposition 1.2 besagt: Es gibt mindestens ein Tripel $(N, 0, s)$, das die Peano-Axiome erfüllt.
Und: Erfüllen $(\widetilde{N}, \widetilde{0}, \widetilde{s})$ und $(\widehat{N}, \widehat{0}, \widehat{s})$ die Peano-Axiome, so sind $(\widetilde{N}, \widetilde{0}, \widetilde{s})$ und $(\widehat{N}, \widehat{0}, \widehat{s})$ „im wesentlichen gleich“ (Proposition 1.3).

(b) Die axiomatische Sichtweise: wir wollen im mathematischen Teilgebiet „Theorie der natürlichen Zahlen“ annehmen, dass es eine Menge N, ein $0 \in N$ und eine Nachfolger-Abbildung gibt und nehmen die Peano-Axiome als **Axiome dieses Teilgebiets**. Proposition 1.2 besagt dann, dass es ein mengentheoretisches Modell gibt, das die Axiome erfüllt. Wie dieses Modell aber aussieht, ist für unsere weiteren Überlegungen irrelevant, da wir nur die Axiome für weitere Schritte nutzen.

Beweis von Proposition 1.2. Wiederholung: eine Menge A heißt induktiv, wenn $\emptyset \in A$ und wenn für alle $x \in A$ auch $x \cup \{x\} \in A$ ist.

Nach dem Unendlichkeitsaxiom gibt es eine induktive **Menge** A. Wir setzen $0 := \emptyset$. Wir definieren nun die Nachfolger-Abbildung

$$\widetilde{s} : A \longrightarrow A, \quad x \mapsto x \cup \{x\}.$$

Das Tripel $(A, 0, \widetilde{s})$ erfüllt offensichtlich (P1).

Für (P2) muss etwas gearbeitet werden, wir beweisen die Injektivität von $\widetilde{s} : A \longrightarrow A$ mit einem Widerspruchsbeweis.

Angenommen die Abbildung sei nicht injektiv. Dies bedeutet:

$$\exists x, y \in A : \; x \neq y \wedge \widetilde{s}(x) = \widetilde{s}(y) \tag{1.5}$$

Es folgt $x \cup \{x\} = y \cup \{y\}$ und hieraus folgen wiederum die Aussagen

$$x \in y \vee x \in \{y\} \tag{1.6}$$

und

$$y \in x \vee y \in \{x\}. \tag{1.7}$$

Nun ist aber $x \in \{y\}$ gleichbedeutend mit $x = y$ und dies haben wir oben ausgeschlossen. Somit erhalten wir aus (1.6) dann $x \in y$, und analog erhalten wir $y \in x$ aus (1.7).

Wir definieren nun $B := \{x, y\}$. Wegen $B \neq \emptyset$ ergibt das Fundierungsaxiom die Existenz eines $b \in B$ mit $b \cap B = \emptyset$. Im Fall $b = x$ erhalten wir $x \cap \{x, y\} \supset \{y\}$ also $b \cap B \neq \emptyset$. Im Fall $b = y$ zeigen wir $b \cap B \neq \emptyset$ analog. Wir haben einen Widerspruch erhalten, da wir gleichzeitig $b \cap B \neq \emptyset$ und $b \cap B = \emptyset$ erhalten haben. Also war eine Annahme (1.5) falsch. Wir haben mit einem Widerspruchsbeweis die Injektivität von $\widetilde{s}$ gezeigt.

Es verbleibt aber unklar, ob $(A, 0, \widetilde{s})$ auch (P3) erfüllt.

Wir setzen nun

$$N := \bigcap \{B \in \mathcal{P}(A) \mid B \text{ ist induktiv}\}.$$

Der Schnitt dieser induktiven Mengen ist wieder induktiv. Das Tripel $(N, 0, \widetilde{s}|_N)$ erfüllt offensichtlich (P1) und (P2). Erfüllt T die Voraussetzungen in (P3), so ist T induktiv und es gilt $T \in \mathcal{P}(N) \subset \mathcal{P}(A)$. Also folgt $N \subset T$, somit $T = N$. Wir erhalten (P3). □

Beispiel 1.8. Die Konstruktion liefert dann $0 = \emptyset$, $1 = s(0) = \{\emptyset\}$, $2 = s(1) = \{\emptyset, \{\emptyset\}\}$, $3 = s(2) = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$.

Proposition 1.9. Jede natürliche Zahl ungleich 0 ist der Nachfolger genau einer natürlichen Zahl.

Beweis (Direkter Beweis). Wir wissen bereits, dass jede natürliche Zahl der Nachfolger höchstens einer natürlichen Zahl ist. Zu zeigen bleibt also, dass jede natürlichen Zahl n ungleich 0 Nachfolger einer natürlichen Zahl ist. Wir definieren:

$$U := \{n \in N \mid n \text{ ist Nachfolger einer natürlichen Zahl}\}$$

und dann

$$T := U \cup \{0\}.$$

Die Menge T erfüllt die Eigenschaften im Peano-Axiom (P3): 0 ist in T, und wenn n in T enthalten ist, dann ist n eine natürliche Zahl, und somit ist $s(n)$ eine natürliche Zahl, die Nachfolger einer natürlichen Zahl ist. Somit ist $s(n)$ in T. Axiom 3 besagt also, dass T gleich N ist. Daraus folgt $U = N$ oder $U = N \setminus \{0\}$. Also ist jede Zahl ungleich 0 Nachfolger einer natürlichen Zahl. □Man kann den Beweis auch anders führen, als sogenannten Widerspruchsbeweis.

Beweis (Widerspruchsbeweis). **Wir nehmen an, die Aussage des Lemmas ist falsch, und wollen daraus einen Widerspruch herleiten. Wenn die Aussage des Lemmas falsch ist, dann können wir annehmen:** *Es gebe* eine natürliche Zahl n ungleich 0, die nicht Nachfolger einer natürlichen Zahl ist. Wir definieren $T := N \setminus \{n\}$. Die Menge T enthält 0, da $n \neq 0$. Ist $t \in T$, so ist t auch eine natürliche Zahl und somit ist auch der Nachfolger $s(t)$ eine natürliche Zahl. Da n kein Nachfolger einer natürlichen Zahl ist, folgt

$s(t) \neq n$ und somit $s(t) \in T$. Die Menge T erfüllt also die Eigenschaften in Peano-Axiom (P3) und somit gilt $T = N$. Zusammen mit $n \in N$ und $n \notin T$ ergibt sich ein Widerspruch. **Da die Existenz einer Zahl n mit den obigen Eigenschaften zu einem Widerpruch führen würde, wissen wir, dass es ein solches Gegenbeispiel nicht gibt, und die Aussage des Lemmas ist somit bewiesen.** □

2. Vollständige Induktion und rekursive Definition

In diesem Abschnitt nehmen wir an, dass $(N, 0, s)$ die Peano-Axiome erfüllt. Die Addition $+$ ist noch nicht definiert! Wir schreiben 1 für $s(0)$, 2 für $s(1)$ etc.. Die Notation $n+1$ ist im Sinne von $s(n)$ zu lesen.

Satz 2.1 (Vollständige Induktion)**.** Sei $A(\,\cdot\,)$ eine auf N definierte Aussageform. Wir setzen voraus, dass Induktionsanfang und Induktionsschritt erfüllt sind:
Induktionsanfang: $A(0)$ ist wahr.
Induktionsschritt: Für alle $n \in N$ gilt: $(A(n) \Longrightarrow A(n+1)))$.
Dann gilt für alle $n \in N$ die Aussage $A(n)$.

Im Induktionsschritt nennt man $A(n)$ die Induktionsvoraussetzung.

Beweis. Sei

$$T := \{n \in N \mid A(n)\}.$$

Auf Grund des Induktionsanfangs ist 0 in T. Der Induktionsschritt besagt: wenn $n \in T$, dann ist auch $n+1$ in T. Die Menge T erfüllt also die Eigenschaften in Peano-Axiom (P3) und somit gilt $T = N$. □

Viele Variation hiervon, z.B.

Satz 2.2 (Vollständige Induktion, Starke Version)**.** Sei $A(n)$ eine auf N definierte Aussage. Wir setzen voraus, dass Induktionssanfang und der modifizierte Induktionsschritt erfüllt sind:

Induktionsanfang: $A(0)$ ist wahr.

Modifizierter Induktionsschritt: Für alle $n \in N$ gilt: Aus $A(0) \wedge A(1) \wedge \ldots \wedge A(n)$ folgt $A(n+1)$).

Dann gilt für alle $n \in N$ die Aussage $A(n)$.

Beweis. Übungsblatt 4, Aufgabe 4 □

Satz 2.3 (Dedekindscher Rekursionssatz)**.** Sei M eine Menge, $a \in M$ und $g : M \times N \longrightarrow M$ eine Abbildung. Dann gibt es genau eine Abbildung $f : N \longrightarrow M$ so dass $f(0) = a$ und

$$\forall n \in N : \ f(s(n)) = g(f(n), n).$$

Natürlich kann man hier auch wieder $n+1$ statt $s(n)$ schreiben.

Wenn eine Abbildung auf diese Art und Weise definiert wird, nennen wir dies eine rekursive Definition.

Beweis. Wir beweisen zunächst durch vollständige Induktion die folgende Aussage $A(n)$, $n \in N$:

Es gibt eine eindeutige[1] Abbildung $f_n : \{0, 1, 2, \ldots, n\} \longrightarrow M$ mit den Eigenschaften[2]

$$f_n(0) = m \text{ und } \forall i \in \{0, 1, 2, \ldots, n-1\} : \ f_n(i+1) = g(f_n(i), i).$$

[1] „eindeutig“ bedeutet: es gibt so eine Funktion, und dies ist die einzige Funktion, die das erfüllt!

[2] Man muss sich an dieser Stelle eigentlich Gedanken machen, was hier mit $\{0, 1, 2 \ldots, n\}$ gemeint ist. Es ist die Menge K_n mit den Eigenschaften

Induktionsanfang: $n = 0$. Die Abbildung $f_0 : \{0\} \longrightarrow M$, $0 \mapsto m$ ist solch eine Abbildung. Und es ist offensichtlich die einzige.

Induktionsschritt:

Induktionsvoraussetzung: $A(n)$

Also wir haben eine Abbildung f_n wie oben.

Wir definieren dann

$$f_{n+1}(i) := \begin{cases} f_n(i) & \text{falls } i \in \{0, 1, 2, \ldots, n\}, \\ g(f_n(n), n) & \text{falls } i = n + 1. \end{cases}$$

Diese Abbildung erfüllt die in $A(n + 1)$ genannte Eigenschaft und man sieht leicht, dass es die einzige ist.

Wir setzen nun: $f(n) := f_n(n)$. Die so definierte Abbildung $f : N \longrightarrow M$ erfüllt die Eigenschaften des Satzes. □

Wir können nun Proposition 1.3 zeigen.

Beweis von Prop. 1.3. Wir definieren rekursiv $F(\widetilde{1}) = \widehat{1}$ und für $n \in N$: $F(\widetilde{s}(n)) = \widehat{s}(F(n))$. Dies ergibt eine Abbildung $F : \widetilde{N} \longrightarrow \widehat{N}$. Analog definiert man eine $G : \widehat{N} \longrightarrow \widetilde{N}$ durch $G(\widehat{1}) = \widetilde{1}$ und $G(\widehat{s}(n)) = \widetilde{s}(G(n))$. Man zeigt nun durch Induktion $G \circ F = \Delta_{\widetilde{N}}$ und $F \circ G = \Delta_{\widehat{N}}$, d.h. $G : \widehat{N} \longrightarrow \widetilde{N}$ ist die Umkehrfunktion von $F : \widetilde{N} \longrightarrow \widehat{N}$ und somit ist

(1) $0 \in K_n$,

(2) $s_{\#}(K_n \setminus \{n\}) \subset K_n$,

(3) Ist T eine Menge mit $0 \in T$ und $s_{\#}(T \setminus \{n\}) \subset T$, dann gilt $K_n \subset T$.

Dann ist zu zeigen dass $s(n) \notin K_n$ und $K_{n+1} = K_n \cup \{s(n)\}$. Man sieht also, dass K_n genau das ist, was wir uns unter $\{0, 1, 2, \ldots, n\}$ vorstellen.

$F : \widetilde{N} \longrightarrow \widehat{N}$ bijektiv. Die übrigen Eigenschaften dieser Abbildung folgen direkt aus der Definition von F. □

Aus Proposition 1.3 folgt: alles was wir für aus den Peano-Axiomen heraus für $(N, s, 0)$ zeigen, gilt auch für $(N', s', 0')$. Es ist also unerheblich, welches Modell wir nutzen. Wir nehmen nun eines her und schreiben ab sofort $\mathbb{N}$ an Stelle von N, $n \mapsto n + 1$ an Stelle von s und weiterhin 0. Wir nennen $\mathbb{N}$ die Menge der natürlichen Zahlen.

Beispiele 2.4.

(1) Addition: Sei $M = \mathbb{N}$, $g(m, n) := s(m)$, $a \in \mathbb{N}$. Wir erhalten eine Abbildung $\alpha_a : \mathbb{N} \longrightarrow \mathbb{N}$ mit $\alpha_a(0) = a$ und mit $\forall n \in \mathbb{N} : \alpha_a(s(n)) = s(a_a(n))$. Wir schreiben $a + n := \alpha_a(n)$. Für $n = 1$ stimmt dies mit der bisherigen Definition von $a + 1$ überein.

(2) Multiplikation: Sei $M = \mathbb{N}$, $i \in \mathbb{N}$, $g : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$, $(m, n) \mapsto m + i$. Wir erhalten eine Abbildung $\mu_i : \mathbb{N} \longrightarrow \mathbb{N}$ mit $\mu_i(0) = 0$ und mit $\forall n \in \mathbb{N} : \mu_i(s(n)) = \mu_i(n) + i$. Wir schreiben $i \cdot m := \mu_i(m)$.

(3) Potenzieren: Sei $M = \mathbb{N}$, $i \in \mathbb{N}$, $g : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$, $(m, n) \mapsto m \cdot i$. Wir erhalten eine Abbildung $p_i : N \longrightarrow \mathbb{N}$ mit $p_i(0) = 1$ und mit $\forall n \in N : p_i(s(n)) = p_i(n) \cdot i$. Wir schreiben $i^m := p_i(m)$.

(4) Fakultät: Sei $M = \mathbb{N}$, $g : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$, $(m, n) \mapsto (n + 1) \cdot m$. Wir definieren dadurch die Abbildung $! : \mathbb{N} \longrightarrow \mathbb{N}$ durch $0! := 1$ und $(m + 1)! := m! \cdot (m + 1)$. Also $n! = 1 \cdot 2 \cdot \cdots \cdot n$.

Satz 2.5. $(\mathbb{N}, +, \cdot)$ erfüllt die folgenden Eigenschaften:

(Aa) **Addition ist assoziativ.**

Für alle $x, y, z \in \mathbb{N}$ gilt

$$(x + y) + z = x + (y + z).$$

(An) **Addition hat neutrales Element.**

Es gibt ein Element $\tilde{0} \in \mathbb{N}$, so dass für alle $x \in \mathbb{N}$ gilt

$$x + \tilde{0} = \tilde{0} + x = x.$$

(Ak) **Addition ist kommutativ.**

Für alle $x, y \in \mathbb{N}$ gilt

$$x + y = y + x.$$

(Ma) **Multiplikation ist assoziativ.**

Für alle $x, y, z \in \mathbb{N}$ gilt

$$(x \cdot y) \cdot z = x \cdot (y \cdot z).$$

(Mn) **Multiplikation hat neutrales Element.**

Es gibt ein Element $\tilde{1} \in \mathbb{N}$, so dass für alle $x \in \mathbb{N}$ gilt

$$x \cdot \tilde{1} = \tilde{1} \cdot x = x.$$

(Mk) **Multiplikation ist kommutativ.**

Für alle $x, y \in \mathbb{N}$ gilt

$$x \cdot y = y \cdot x.$$

(AMd) **Addition und Multiplikation erfüllen das Distributivgesetz.**

Für alle $x, y, z \in \mathbb{N}$ gilt

$$x \cdot (y + z) = x \cdot y + x \cdot z$$

$$(y + z) \cdot x = y \cdot x + z \cdot x$$

Den Beweis kann man mit vollständiger Induktion durchführen, die wir im nächsten Abschnitt kennenlernen werden. Es ist eine gute Übung, einmal die Kommutativität der Addition durch vollständige Induktion oder direkt aus den Peano-Axiomen herzuleiten. Dies ist etwas mühsam, aber prinzipiell möglich.

Offensichtlich ist 0 die einzige Zahl, die an Stelle von $\tilde{0}$ die obige Eigenschaft erfüllt. Wir nennen $0 = \tilde{0}$ das neutrale Element der Addition. Analoges gilt für 1 und $\tilde{1}$, und man nennt $1 = \tilde{1}$ das neutrale Element der Multiplikation.

LEMMA 2.6 (Kürzungsregel)**.** Seien $n, m, k \in \mathbb{N}$ mit $n + k = m + k$. Dann gilt auch $m = n$.

In anderen Worten

$$\forall n, m, k \in \mathbb{N} : \; n + k = m + k \Longrightarrow m = n.$$

Der Beweis folgt durch Induktion nach k (wird nicht im Detail ausgeführt).

LEMMA 2.7. Jede von 0 verschiedene natürliche Zahl ist Nachfolger einer natürlichen Zahl.

Beweis. Angenommen $k \in \mathbb{N}$, $k \neq 0$ und $k \notin s_{\#}(\mathbb{N})$. Definiere $T := \mathbb{N} \setminus \{k\}$. Dann gilt $0 \in T$ und $s_{\#}(T) \subset T$ und somit erhalten wir den Widerspruch $T = \mathbb{N}$. □

LEMMA 2.8. Sind $n, m \in \mathbb{N}$. Es gelte $n + m = 0$. Dann gilt $n = m = 0$.

Beweis. (Widerspruchsbeweis) Wir nehmen an, dass $n \neq 0$ oder $m \neq 0$. Da der Fall $m \neq 0$ ganz analog zum Fall $n \neq 0$, genügt es den Fall $n \neq 0$ zu betrachten. [3] Dann gibt es eine Zahl $k \in \mathbb{N}$ mit $n = s(k)$. Es folgt $0 = s(k) + m = s(k + m)$, was der Tatsache widerspricht, dass 0 kein Nachfolger einer natürlichen Zahl ist. □

Hier gäbe es noch viel hinzuzufügen. Deswegen wird der Anhang evtl. hier noch erweitert.

3. Ordnung der natürlichen Zahlen

Wir definieren nun eine Relation:

$$\leq := \{(n, m) \in \mathbb{N} \times \mathbb{N} \mid \exists k \in \mathbb{N} : \ m = n + k\}.$$

Lemma 3.1. Diese Relation $\leq$ ist eine totale Ordnung auf $\mathbb{N}$. Es gilt:

(1) Reflexivität auf $\mathbb{N}$: Für alle m in $\mathbb{N}$ ist $m \leq m$ wahr.
(2) Antisymmetrie: Für alle n und m in $\mathbb{N}$ gilt

$$n \leq m \wedge m \leq n \Longrightarrow n = m.$$

(3) Transitivität: Für alle n, m und k in $\mathbb{N}$ gilt:

$$n \leq m \wedge m \leq k \Longrightarrow n \leq k.$$

(4) Totalität: Für alle n und m in $\mathbb{N}$ gilt

$$n \leq m \vee m \leq n$$

[3] Da Sätze in dieser Art oft vorkommen, sagt man „Ohne Beschränkung der Allgemeinheit (OBdA) gilt $n \neq 0$."

(5) Für alle $n \in \mathbb{N}$ gilt $n \leq n + 1$.

Beweis. Aussagen (1), (3) und (5) sind offensichtlich.

Zu (2) (Antisymmetrie): Es gelte $n \leq m$ und $m \leq n$. Dann gibt es $k_1, k_2 \in \mathbb{N}$ mit $m = n + k_1$ und $n = m + k_2$. Daraus folgt

$$n = m + k_2 = n + k_1 + k_2.$$

Mit der Kürzungsregel (Lemma 2.6) folgt $k_1 + k_2 = 0$ und mit Lemma 2.8 ergibt sich $k_1 = k_2 = 0$. Also $n = m$.

Zu (4) (Totalität): Zu $n \in \mathbb{N}$ definieren wir

$$T_n := \{m \in \mathbb{N} \mid m \leq n \vee n \leq m\}.$$

Man zeigt mit etwas Aufwand $0 \in T_n$ und $s_\#(T_n) \subset T_n$. Mit (P3) folgt $T_n = \mathbb{N}$, also die Behauptung. □

Definition 3.2. Sei R eine (partielle) Ordnungsrelation auf M. Ein Minimum (beziehungsweise Maximum) ist ein Element $m \in M$, so dass für alle $n \in M$ gilt: mRn (bzw. nRm).

Bemerkung. Wegen der Antisymmetrie gibt es höchstens ein Minimum.

Beispiele: Das offene Intervall $]0, 1[$ in $\mathbb{R}$ hat kein Minimum bezüglich $\leq$.

Die Menge $M := \{\{a\}, \{b\}, \{a, b\}\}$ trägt die Ordnungsrelation $\subset$. Es existiert kein Minimum in M.

Proposition 3.3 (Wohlordnung von $\mathbb{N}$)**.** Sei A eine nichtleere Teilmenge von $\mathbb{N}$, dann besitzt A ein Minimum.

Beweis. Wir nehmen an, A besäße kein Minimum. Definiere

$$\begin{aligned} T &:= \{n \in \mathbb{N} \mid \forall k \in A : k > n\} \\ &= \{n \in \mathbb{N} \mid \forall k \in \mathbb{N} : (k \leq n \Longrightarrow k \notin A)\} \\ &= \{n \in \mathbb{N} \mid \neg(\exists k \in \mathbb{N} : (k \leq n \wedge k \in A))\}. \end{aligned}$$

Offensichtlich gilt $0 \in T$: denn wenn $0 \notin T$ wäre, dann gäbe es ein $k \in \mathbb{N}$ mit $k \leq 0$ und $k \in A$, also $0 \in A$; und dann wäre 0 ein Minimum von A.

Wir zeigen nun

$$(n \in T) \Longrightarrow (n+1 \in T). \tag{3.4}$$

Daraus folgt dann mit (P3) die Aussage $T = \mathbb{N}$, also $A = \emptyset$.

Um (3.4) zu zeigen, nehmen wir an, es gebe ein $n \in T$ mit $n+1 \notin T$. Daraus folgt dann $n+1 \in A$, und aus $n \in T$ folgt mit der Totalität dann, dass $n+1$ ein Minimum von A ist. Dies ist ein Widerspruch zur obigen Annahme. □

4

[4]Alternativer Beweis: Wir nehmen an, A besitze kein Minimum. Zu zeigen ist, dass A die leere Menge ist. Wir zeigen induktiv die Aussage

$$P(n) \quad :\Longleftrightarrow \quad \{0, 1, 2, \ldots, n\} \cap A = \emptyset,$$

woraus die Aussage folgt.

Induktionsanfang: Angenommen 0 wäre in A. Dann ist 0 das Minimum. Da es aber kein Minimum in A gibt, folgt $0 \notin A$, also $P(0)$.

Induktionsschritt: Es gelte $P(n)$. Falls $n+1 \in A$, so ist $n+1$ ein Minimum von A. Da es aber kein Minimum gibt, gilt $n+1 \notin A$, und somit $P(n+1)$. □

BEMERKUNG 3.5. Sei R eine totale Ordnung auf einer Menge M. Wir sagen, R ist eine Wohlordnung oder (M, R) ist eine wohlgeordnete Menge, wenn jede nicht-leere Teilmenge A ein Minimum besitzt. Die letzte Proposition besagt also, dass $(\mathbb{N}, \leq)$ wohlgeordnet ist. Hingegen ist $(\mathbb{R}, \leq)$ nicht wohlgeordnet. Man kann aber zeigen:

Zu jeder Menge M gibt es eine Wohlordnung auf M.

Diese Aussage ist zum Auswahlaxiom äquivalent, wenn wir die übrigen Axiome der Mengenlehre annehmen.

ANHANG C

Konstruktion von $\mathbb{R}$ mit Hilfe von Cauchy-Folgen

In diesem Anhang stellen wir eine alternative Möglichkeit dar, ein Modell für die reellen Zahlen zu konstruieren, nämlich mit Hilfe von Cauchy-Folgen. Wenn man die Einführung von Cauchy-Folgen in diesem Anhang mit der Einführung im Hauptteil in Abschnitt 1.6 vergleicht, so sieht man, dass wir in diesem Anhang einige Eingeschaften explizit prüfen müssen, die im Hauptteil aus den im Abschnitt 6 behandelten Aussagen folgen. In Abschnitt 1.6 setzen wir voraus, dass wir bereits wissen, was $\mathbb{R}$ ist und dass jeder archimedisch geordnete Körper Teilkörper von $\mathbb{R}$ ist. Auch die Definition einer Cauchy-Folge muss leicht angepasst werden. [1]

1. Mehr zu Äquivalenzrelationen

Wir besprechen hier Äquivalenzrelationen noch etwas detailierter als zuvor. Derartige Sachverhalte wurden bereits in der Linearen Algebra I im Detail behandelt. Wir wiederholen sie in der Zentralübung am 26.11.2013, u.a. um sicherzustellen, dass auch diejenigen folgen können, die die Lineare Algebra I nicht hören.

[1] Dieser Teil wurde von meinem früheren Skript der Vorlesung im Wintersemester 2013/14 übernommen. Leider können kleinere übernahmebedingten Druckfehler nicht ganz ausgeschlossen werden. Auch werden manche Themen der Hauptvorlesung hier potentiell verdoppelt.

ÜBUNG 1.1. Seit $f : M \longrightarrow N$ eine Abbildung. Wir definieren

$$R_f := \{(x, y) \in M \times M \mid f(x) = f(y)\}.$$

Zeigen Sie, R ist eine Äquivalenzrelation.

Sei R eine Äquivalenzrelation auf M.

Für $x \in M$ definieren wir die Äquivalenzklasse von x als

$$[x] := \{y \in M \mid xRy\}.$$

Man sagt auch x ist ein Repräsentant von $[x]$ oder x repräsentiert $[x]$.

ÜBUNG 1.2. Zeigen Sie, dass für alle $x, y \in M$ gilt:

(a)

$$y \in [x] \Longleftrightarrow [x] = [y]$$

(b)

$$[x] = [y] \quad \veebar \quad [x] \cap [y] = \emptyset.$$

ÜBUNG 1.3. Sei R_f wie oben definiert. Bestimmen Sie die Äquivalenzklassen.

Ist R eine Äquivalenzrelation auf M, so definieren wir den Quotient von M nach R als

$$M/R := \{[x] \mid x \in M\}.$$

Die Abbildung $[\,\cdot\,] : M \to M/R$, $x \mapsto [x]$ nennt man die kanonische Projektion.

Übung 1.4. Sei eine Abbildung $F : M \longrightarrow X$ gegeben und R eine Äquivalenzrelation auf M. Zeigen Sie, es gibt genau dann eine Abbildung $\widehat{F} : M/R \longrightarrow X$, so dass $\widehat{F} \circ [\,\cdot\,] = F$, falls für alle $x, y \in M$ gilt

$$xRy \Longrightarrow F(x) = F(y).$$

Die Abbildung $\widehat{F} : M/R \longrightarrow X$ ist daraus eindeutig bestimmt.

$$\begin{array}{ccc} M & \xrightarrow{F} & X \\ {\scriptstyle [\,\cdot\,]}\downarrow & \nearrow_{\widehat{F}} & \\ M/R & & \end{array}$$

Gilt sogar

$$xRy \Longleftrightarrow F(x) = F(y),$$

dann wissen wir noch zusätzlich, dass $\widehat{F}$ injektiv ist.

Beispiel: Sei $f : M \longrightarrow N$, R_f wie oben. Dann gibt es eine eindeutige Abbildung $\widehat{f} : M/R_f \longrightarrow N$, so dass

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ {\scriptstyle [\,\cdot\,]}\downarrow & \nearrow_{\widehat{f}} & \\ M/R_f & & \end{array}$$

kommutiert. Außerdem ist $\widehat{f}$ injektiv. Wir erhalten eine bijektive Abbildung $\widehat{f} : M/R_f \longrightarrow B(f)$.

2. Folgen, Konvergenz und Cauchy-Folgen

Wenn wir die reellen Zahl über Cauchy-Folgen konstruieren wollen, so benötigen wir Begriffe wie Konvergenz, Cauchyfolgen und ähnliche für beliebige archimedisch geordnete Körper. Wir dürfen ja insbesondere nicht nutzen, dass jeder solche Körper ein Unterkörper von $\mathbb{R}$ ist. Deswegen müssen wir in großen Teilen unsere Einführung in Folgen und Reihen nochmals in etwas größerer Allgemeinheit wiederholen.

Wir schreiben im folgenden zumeist K für einen geordneten Körper $(K, +, \cdot, \leq)$. Und $|a| \in K_{\geq 0}$ sei (wie immer) der Betrag von $a \in K$.

DEFINITION 2.1. Eine K-wertige Folge $(a_i)_{i\in\mathbb{N}}$ heißt beschränkt[2] ,

$$
\begin{aligned}
&:\Longleftrightarrow \{a_i \mid i \in \mathbb{N}\} \text{ ist (nach oben und unten) beschränkt in } K \\
&\Longleftrightarrow \exists r_1, r_2 \in K : \forall i \in \mathbb{N} : r_1 \leq a_i \leq r_2 \\
&\Longleftrightarrow \exists r \in K : \forall i \in \mathbb{N} : |a_i| \leq r \\
&\Longleftrightarrow \{|a_i| \mid i \in \mathbb{N}\} \text{ ist beschränkt in } K
\end{aligned}
$$

BEMERKUNG 2.2. Falls K archimedisch ist, so gilt auch:

$$(a_i)_{i\in\mathbb{N}} \text{ ist beschränkt in } K \quad \Longleftrightarrow \quad \exists n \in \mathbb{N} : \forall i \in \mathbb{N} : |a_i| \leq n.$$

Die Implikation $\Longleftarrow$ ist klar, und $\Longrightarrow$ folgt direkt aus der archimedischen Eigenschaft

[2]Wenn man hier ganz exakt sein will, sollte man hier besser „beschränkt in K“ oder „K-beschränkt“ sagen, denn ob eine Folge beschränkt ist, hängt von K ab. Beispiel: Die Folge $(i)_{i\in\mathbb{N}}$ ist unbeschränkt in $\mathbb{Q}$, aber beschränkt im Körper der rationalen Funktionen mit rationalen Koeffizienten. Sind aber K_1 und K_2 archimedische geordnete Körper, so sieht man mit der unten stehenden Bemerkung, dass K_1-Beschränktheit und K_2-Beschränktheit die gleiche Bedeutung haben.

DEFINITION 2.3 (Konvergenz von Folgen). Eine K-wertige Folge $(a_i)_{i\in\mathbb{N}}$ konvergiert [3] gegen $a \in K$, falls gilt

$$\forall \epsilon \in K_{>0} : \exists i_0 \in \mathbb{N} : \forall i \in \mathbb{N} : (i \geq i_0 \Longrightarrow |a_i - a| \leq \epsilon).$$

Man nennt a den Grenzwert der Folge $(a_i)_{i\in\mathbb{N}}$ und schreibt $a = \lim_{i\to\infty} a_i$ oder $a_i \to a$ für $i \to \infty$. Wir sagen $(a_i)_{i\in\mathbb{N}}$ konvergiert, falls es ein derartiges $a \in K$ gibt. Folgen, die gegen 0 konvergieren, nennt man Nullfolgen. Falls eine Folge nicht konvergiert, so sagen wir dazu sie divergiert.

!ACHTUNG!. Wenn wir $a = \lim_{i\to\infty} a_i$ schreiben, so bedeutet dies immer:

- der Grenzwert existiert, und
- der Grenzert ist a.

BEMERKUNGEN 2.4.

(a) Falls eine Folge einen Grenzwert besitzt, so ist dieser eindeutig bestimmt. Seien a und a' zwei Grenzwerte von $(a_i)_{i\in\mathbb{N}}$. Zu einem gegebenen $\epsilon \in K_{>0}$ gilt also:

$$\exists i_0 \in \mathbb{N} : \forall i \in \mathbb{N} : (i \geq i_0 \Longrightarrow |a_i - a| \leq \epsilon).$$
$$\exists i'_0 \in \mathbb{N} : \forall i \in \mathbb{N} : (i \geq i'_0 \Longrightarrow |a_i - a'| \leq \epsilon).$$

[3] Wenn man hier ganz exakt sein will, sollte man hier besser „K-konvergiert“ sagen, denn ob eine Folge konvergiert, hängt von K ab. Beispiel: Die Folge $(1/i)_{i\in\mathbb{N}}$ $\mathbb{Q}$-konvergiert in $\mathbb{Q}$ gegen 0, aber nicht im Körper der rationalen Funktionen mit rationalen Koeffizienten. Sind aber K_1 und K_2 archimedische geordnete Körper, so sieht man mit einer der unten stehenden Bemerkungen, dass K_1-Konvergenz und K_2-Konvergenz die gleiche Bedeutung haben.

Wähle nun so ein i_0 und i_0'. Dann gilt für alle $i \geq \max\{i_0, i_0'\}$:

$$|a - a'| \leq |a - a_i| + |a_i - a'| \leq \epsilon + \epsilon = 2\epsilon.$$

Dies gilt für alle $\epsilon \in K_{>0}$. Angenommen wir haben $a \neq a'$, so gilt dies insbesondere für $\epsilon := |a - a'|/3 > 0$. Also folgt $3\epsilon \leq 2\epsilon$ und somit ergibt sich der Widerspruch $\epsilon \leq 0$. Die Annahme $a \neq a'$ war also falsch, d.h. es gilt $a = a'$.

(b) Falls $(a_i)_{i\in\mathbb{N}}$ konvergiert, so ist $(a_i)_{i\in\mathbb{N}}$ beschränkt. Um dies zu zeigen, wählen wir zu $\epsilon := 1$ ein passendes i_0. Es gilt somit für alle $i \in \mathbb{N}$ mit $i \geq i_0$:

$$|a_i| \leq |a_i - a| + |a| \leq |a| + 1.$$

Nun setzen wir

$$r := \max\{|a_1|, |a_2|, \ldots, |a_{i_0-1}|, |a| + 1\}.$$

Dann gilt für alle $i \in \mathbb{N}$: $|a_i| \leq r$. Somit ist $(a_i)_{i\in\mathbb{N}}$ beschränkt.

(c) Ist K archimedisch, so gilt

$$a = \lim_{i\to\infty} a_i \quad \Longleftrightarrow \quad \forall n \in \mathbb{N} : \exists i_0 \in \mathbb{N} : \forall i \in \mathbb{N} : (i \geq i_0 \Longrightarrow |a_i - a| \leq \frac{1}{n}).$$

Die Implikation $\Longrightarrow$ ist klar, und $\Longleftarrow$ folgt aus Lemma 2.5.

Lemma 2.5. Ist K archimedisch, dann gibt es für alle $\epsilon \in K_{>0}$ ein $n \in \mathbb{N}$ mit $\epsilon \geq (1/n)$.

Beweis. Zu $1/\epsilon \in K$ gibt es ein $n \in \mathbb{N}$ mit $1/\epsilon \leq n$, und dies ist äquivalent zu $\epsilon \geq (1/n)$. □

Beispiele 2.6. (a) Eine Folge $(a_i)_{i\in\mathbb{N}}$ heißt konstant, falls $a_1 = a_2 = a_3 = \ldots$. Konstante Folgen sind beschränkt und konvergieren. $a_1 = \lim_{i\to\infty} a_i$.

(b) Sei $K = \mathbb{Q}$. Dann ist $(\frac{1}{i})_{i\in\mathbb{N}}$ eine Nullfolge. (Nutze z.B. Bem. 2.4 (c) und setze $i_0 := n$).

(c) Die Folge $\left((-1)^i\right)_{i\in\mathbb{N}}$ hat 1 als obere und -1 als untere Schranke und ist somit beschränkt. Wir werden bald sehen, dass sie nicht konvergiert.

DEFINITION 2.7. Sei $(a_i)_{i\in\mathbb{N}}$ eine Folge von Elementen $a_i \in M$. Sei $f : \mathbb{N} \to \mathbb{N}$ eine Abbildung mit der Eigenschaft

$$\forall i, j \in \mathbb{N} : \; i < j \Longrightarrow f(i) < f(j)$$

(Man sagt zu dieser Eigenschaft: f ist streng monoton wachsend.) Dann nennt man $(a_{f(k)})_{k\in\mathbb{N}}$ eine Teilfolge von $(a_i)_{i\in\mathbb{N}}$.

BEISPIEL 2.8. Die Folge $(i)_{i\in\mathbb{N}}$ hat folgende Teilfolgen:[4]

- sich selbst,
- die Folge der ungeraden Zahlen $(1, 3, 5, 7, \ldots)$,
- die Folge der Primzahlen $(2, 3, 5, 7, \ldots)$.

Konvergiert $(a_i)_{i\in\mathbb{N}}$ gegen a, so konvergiert jede Teilfolge ebenfalls gegen a.

LEMMA 2.9. Die Folge $\left((-1)^i\right)_{i\in\mathbb{N}}$ aus Beispiele 2.6 (c) divergiert.

Beweis. Angenommen, die Folge $\left((-1)^i\right)_{i\in\mathbb{N}}$ konvergiert gegen ein $a \in K$. Dann konvergieren auch die Teilfolgen $\left((-1)^{2k}\right)_{k\in\mathbb{N}} = (1)_{k\in\mathbb{N}}$ und $\left((-1)^{2k+1}\right)_{k\in\mathbb{N}} = (-1)_{k\in\mathbb{N}}$ gegen a. Da diese Teilfolgen konstant sind, erhalten wir $a = 1$ und $a = -1$, was in K nicht möglich ist, denn es gilt ja $-1 < 0 < 1$. □

[4]Diese Folge hat natürlich noch **viel** mehr Teilfolgen. U.a. gibt es sogar Teilfolgen, die man gar nicht beschreiben kann. Wieso es nicht beschreibbare Teilfolgen gibt, kann erst später erklärt werden.

ÜBUNG 2.10. Seien $(a_i)_{i\in\mathbb{N}}$ und $(b_i)_{i\in\mathbb{N}}$ konvergente Folgen. Dann gilt:

$$\lim_{i\to\infty}(a_i + b_i) = \lim_{i\to\infty} a_i + \lim_{i\to\infty} b_i$$
$$\lim_{i\to\infty}(a_i - b_i) = \lim_{i\to\infty} a_i - \lim_{i\to\infty} b_i$$
$$\lim_{i\to\infty}(a_i \cdot b_i) = \lim_{i\to\infty} a_i \cdot \lim_{i\to\infty} b_i$$

Gilt zusätzlich: $\forall i \in \mathbb{N} : b_i \neq 0$, und ist $(b_i)_{i\in\mathbb{N}}$ keine Nullfolge, so gilt auch

$$\lim_{i\to\infty} \frac{a_i}{b_i} = \frac{\lim\limits_{i\to\infty} a_i}{\lim\limits_{i\to\infty} b_i}.$$

Die Beweise sind ähnlich wie in der folgenden Proposition.

DEFINITION 2.11. Eine K-wertige Folge $(a_i)_{i\in\mathbb{N}}$ heißt Cauchy-Folge,[5] falls

$$\forall \epsilon \in K_{>0} : \exists i_0 \in \mathbb{N} : \forall i, j \in \mathbb{N} : (i \geq i_0 \wedge j \geq i_0) \Longrightarrow |a_i - a_j| \leq \epsilon.$$

PROPOSITION 2.12.

(1) Jede konvergente Folge ist eine Cauchy-Folge.
(2) Jede Cauchy-Folge ist beschränkt.
(3) Seien $(a_i)_{i\in\mathbb{N}}$ und $(b_i)_{i\in\mathbb{N}}$ Cauchy-Folgen. Dann sind $(a_i + b_i)_{i\in\mathbb{N}}$, $(a_i - b_i)_{i\in\mathbb{N}}$ und $(a_i \cdot b_i)_{i\in\mathbb{N}}$ ebenfalls Cauchy-Folgen.

(4) Zusätzlich zu den Voraussetzungen in (3) gelte $\forall i \in \mathbb{N} : b_i \neq 0$, und die Folge $(b_i)_{i\in\mathbb{N}}$ sei keine Nullfolge. Dann ist auch $\left(\frac{a_i}{b_i}\right)_{i\in\mathbb{N}}$ eine Cauchy-Folge.

[5]Genau genommen müsste man hier wieder sagen: eine K-Cauchy-Folge, da die Definition zuächst von K abhängt. Die Bedeutung ist dann aber für alle archimedischen Körper dieselbe.

Um die Proposition zu zeigen, nutzen wir ein Lemma.

Lemma 2.13. Sei $A(\,\cdot\,)$ eine auf $K_{>0}$ definierte Aussageform, und $q \in K_{>0}$. Dann gilt

$$\forall \epsilon \in K_{>0} : A(\epsilon) \quad \Longleftrightarrow \quad \forall \epsilon \in K_{>0} : A(q\epsilon)$$

Das Lemma gibt es vielen Variationen. Wichtiger als die Aussage des Lemmas ist es, zu verstehen, wie man das Lemma (oder eine Variation davon!) kurz beweist.

Beweis des Lemmas.
„$\Longrightarrow$“: Es gelte

$$\forall \widehat{\epsilon} \in K_{>0} : A(\widehat{\epsilon}). \tag{2.14}$$

Für ein gegebenes $\epsilon > 0$ wollen wir nun $A(q\epsilon)$ zeigen. Wir wenden (2.14) für $\widehat{\epsilon} := q\epsilon$ an, und haben dann das gewünschte.

„$\Longleftarrow$“: Analog mit $\widehat{\epsilon} := q^{-1}\epsilon$ □

Beweis der Proposition.
(1): Es gelte $\lim_{i\to\infty} a_i = a$. Das heißt: für alle $\epsilon \in K_{>0}$ gibt es ein $i_0 \in \mathbb{N}$, so dass für alle natürlichen Zahlen $i \geq i_0$ gilt: $|a_i - a| \leq \epsilon$.

Für solch ein ϵ und ein passendes i_0 nehmen wir nun natürliche Zahlen $i \geq i_0$ und $j \geq i_0$ und rechnen nach:

$$|a_i - a_j| \leq |a_i - a| + |a - a_j| \leq \epsilon + \epsilon = 2\epsilon.$$

Wir haben nun also gezeigt:

$$\forall \epsilon \in K_{>0} : A(2\epsilon),$$

wobei $A(\cdot)$ die folgende auf $K_{>0}$ definierte Aussageform ist

$$A(\tilde{\epsilon}) \quad :\Longleftrightarrow \quad \exists i_0 \in \mathbb{N} : \forall i, j \in \mathbb{N} : \ (i \geq i_0 \wedge j \geq i_0) \Longrightarrow |a_i - a_j| \leq \tilde{\epsilon}$$

Nach dem obigen Lemma ist dies äquivalent zu

$$\forall \epsilon \in K_{>0} : \ A(\epsilon)$$

und dies ist gerade die Definition einer Cauchy-Folge.

(2): Ähnlich wie Bemerkung 2.4 (b).

(3): Angenommen $(a_i)_{i\in\mathbb{N}}$ und $(b_i)_{i\in\mathbb{N}}$ seien Cauchy-Folgen. Dies bedeutet, dass wir für jedes $\epsilon \in K_{>0}$ die folgenden Aussagen haben:

$$\exists i_0 \in \mathbb{N} : \forall i, j \in \mathbb{N} : \ (i \geq i_0 \wedge j \geq i_0) \Longrightarrow |a_i - a_j| \leq \epsilon$$

$$\exists j_0 \in \mathbb{N} : \forall i, j \in \mathbb{N} : \ (i \geq j_0 \wedge j \geq i_0) \Longrightarrow |b_i - b_j| \leq \epsilon$$

Wir wählen nun solch ein i_0 und solch ein j_0. Wir setzen $k_0 := \max\{i_0, j_0\}$. Dann gilt für dieses k_0:

$$\forall i, j \in \mathbb{N} : \ (i \geq k_0 \wedge j \geq k_0) \Longrightarrow |a_i - a_j| \leq \epsilon \wedge |b_i - b_j| \leq \epsilon.$$

Wir rechnen für $i, j \geq k_0$:

$$|(a_i + b_i) - (a_j + b_j)| = |(a_i - a_j) + (b_i - b_j)| \leq |a_i - a_j| + |b_i - b_j| \leq \epsilon + \epsilon = 2\epsilon.$$

Also ergibt sich insgesamt

$$\forall \epsilon \in K_{>0} : \exists k_0 \in \mathbb{N} : \forall i, j \in \mathbb{N} : \ (i \geq k_0 \wedge j \geq k_0) \Longrightarrow |(a_i + b_i) - (a_j + b_j)| \leq 2\epsilon.$$

Nach dem obigen Lemma können wir auch hier 2ϵ durch ϵ ersetzen. Dann steht hier gerade die Definition, dass $(a_i + b_i)_{i\in\mathbb{N}}$ eine Cauchy-Folge ist.

Der Beweis für $(a_i - b_i)_{i\in\mathbb{N}}$ ist völlig analog.

Für das Produkt $(a_i \cdot b_i)_{i\in\mathbb{N}}$ muss man etwas anders vorgehen. Zunächst nutzen wir die Tatsache, dass $(a_i)_{i\in\mathbb{N}}$ und $(b_i)_{i\in\mathbb{N}}$ beschränkt sind. Also gibt es ein $r \in K$ mit

$$\forall i \in \mathbb{N} : \ |a_i| \leq r \wedge |b_i| \leq r$$

Dann argumentieren wir wie bei der Summe, rechnen dann aber:

$$|a_i b_i - a_j b_j| = |a_i(b_i - b_j) + (a_i - a_j)b_j| \leq |a_i||b_i - b_j| + |a_i - a_j||b_j| \leq r\epsilon + \epsilon r = 2r\epsilon.$$

Nun argumentiert man wie bei der Summe, wobei man das Lemma mit $q := 2r$ nutzt.

(4): Es reicht zu zeigen: Sei $(b_i)_{i\in\mathbb{N}}$ eine Folge wie oben, dann ist $(\frac{1}{b_i})_{i\in\mathbb{N}}$ ebenfalls eine Cauchy-Folge. Die eigentliche Aussage folgt dann mit (3).

Wir zeigen zunächst durch Widerspruch, dass $(\frac{1}{b_i})_{i\in\mathbb{N}}$ beschränkt ist. Angenommen $(\frac{1}{b_i})_{i\in\mathbb{N}}$ sei nicht beschränkt, dann gilt

$$\forall r \in K : \ \exists n \in \mathbb{N} : \ \left|\frac{1}{b_n}\right| > r.$$

Dies impliziert

$$\forall r \in K_{>0} : \ \exists n \in \mathbb{N} : \ |b_n| < \frac{1}{r}$$

und dies ergibt

$$\forall \delta \in K_{>0} : \ \exists n \in \mathbb{N} : \ |b_n| < \delta, \tag{2.15}$$

wobei die folgende Variation des obigen Lemmas benutzt wurde

$$\forall r \in K_{>0} : \ A(1/r) \quad \Longleftrightarrow \quad \forall \delta \in K_{>0} : \ A(\delta).$$

Die Folge $(b_i)_{i\in\mathbb{N}}$ ist eine Cauchy-Folge. Zu einem gegebenen $\epsilon \in K_{>0}$ gibt es also ein $i_0 \in \mathbb{N}$, so dass gilt

$$\forall i, j \in \mathbb{N} : \; i \geq i_0 \wedge j \geq i_0 \Longrightarrow |b_i - b_j| \leq \epsilon \tag{2.16}$$

Wir wählen nun zu diesem i_0:

$$\delta := \frac{\min\{\epsilon, |b_1|, |b_2|, \ldots, |b_{i_0}|\}}{2}.$$

Mit (2.15) erhalten wir ein $n \in \mathbb{N}$ mit $|b_n| < \delta$. Auf Grund der Definition von δ gilt $n > i_0$. Mit (2.16) folgt:

$$\forall j \in \mathbb{N} : \; j \geq i_0 \Longrightarrow |b_n - b_j| \leq \epsilon$$

und mit der Rechnung

$$|b_j| \leq |b_j - b_n| + |b_n| \leq \epsilon + \frac{\epsilon}{2} \leq 2\epsilon$$

folgt

$$\forall \epsilon \in K_{>0} : \; \exists i_0 \in \mathbb{N} : \; \forall j \in \mathbb{N} : \; j \geq i_0 \Longrightarrow |b_j - 0| \leq 2\epsilon.$$

Dies ergibt unter Nutzung des Lemmas für $q = 2$, dass $(b_j)_{j\in\mathbb{N}}$ eine Nullfolge ist. Dies ist ein Widerspruch zur Annahme. Also haben wir die Beschränktheit gezeigt, d.h.

$$\exists r \in K : \; \forall n \in \mathbb{N} : \; \left|\frac{1}{b_n}\right| \leq r.$$

Nun sei wiederum $\epsilon > 0$ und i_0 wie oben. Wir rechnen dann

$$\left|\frac{1}{b_i} - \frac{1}{b_j}\right| \leq \left|\frac{b_j - b_i}{b_i b_j}\right| \leq \frac{|b_j - b_i|}{|b_i||b_j|} \leq r^2\epsilon$$

Unter Nutzung des Lemmas für $q = r^2$ erhalten wir, dass $(\frac{1}{b_i})_{i\in\mathbb{N}}$ eine Cauchy-Folge ist. □

Definition 2.17. Wir sagen K ist vollständig, falls jede K-wertige Cauchy-Folge in K konvergiert.

DEFINITION 2.18. Sei $A(\,\cdot\,)$ eine auf $\mathbb{N}$ definierte Aussageform.

$$\begin{aligned} & \text{Für fast alle } i \in \mathbb{N} \text{ gilt } A(i) \\ :\Longleftrightarrow\ & \text{Es gibt ein } i_0 \in \mathbb{N}, \text{ so dass für alle } i \in \mathbb{N}\text{: } i \geq i_0 \Longrightarrow A(i) \\ \Longleftrightarrow\ & \text{Die Menge } \{i \in \mathbb{N} \mid \neg A(i)\} \text{ ist endlich} \end{aligned}$$

Beispiel: $\lim_{i\to\infty} a_i = a \Longleftrightarrow$ Für alle $\epsilon \in K_{>0}$ gilt für fast alle $i \in \mathbb{N}$: $|a_i - a| \leq \epsilon \quad \not\Longleftrightarrow$ Für fast alle $i \in \mathbb{N}$ gilt für alle $\epsilon \in K_{>0}$: $|a_i - a| \leq \epsilon$

Notation: Da wir immer wieder Ausdrücke der Art

$$\forall i \in \mathbb{N} :\ i \geq i_0 \Longrightarrow A(i)$$

haben, schreiben wir hierfür kurz

$$\forall i \in \{i_0, i_0 + 1, \ldots\} :\ A(i).$$

LEMMA 2.19. Besitzt eine Cauchy-Folge eine konvergente Teilfolge, so ist die Cauchy-Folge bereits konvergent.

Beweis. Sei nun $(a_i)_{i\in\mathbb{N}}$ eine Cauchy-Folge, sei $f : \mathbb{N} \longrightarrow \mathbb{N}$ eine streng monotone wachsende Abbildung, und sei $(a_{f(n)})_{n\in\mathbb{N}}$ eine konvergente Teilfolge mit Grenzwert a.

Es gilt für ein zunächst fixiertes $\epsilon \in K_{>0}$:

$$\exists i_0 \in \mathbb{N} :\ \forall i, j \in \{i_0, i_0 + 1, \ldots\} :\ |a_i - a_j| \leq \epsilon. \tag{2.20}$$

und

$$\exists n_0 \in \mathbb{N} :\ \forall n \in \{n_0, n_0 + 1, \ldots\} :\ |a_{f(n)} - a| \leq \epsilon. \tag{2.21}$$

Für solche i_0 und n_0 setzen wir $k_0 := \max\{i_0, f(n_0)\}$. Da f streng monoton wachsend ist, können wir ein $n \in \{n_0, n_0+1, \ldots\}$ wählen mit $f(n) \geq k_0$. [6] Wir wählen dann $j := f(n)$ und haben somit für alle $i \in \{i_0, i_0+1, \ldots\}$ erhalten:

$$|a_i - a| \leq |a_i - a_{f(n)}| + |a_{f(n)} - a| \leq \epsilon + \epsilon \leq 2\epsilon.$$

Wir haben also insgesamt gezeigt:

$$\forall \epsilon \in K_{>0} : \exists i_0 \in \mathbb{N} : \forall i \in \{i_0, i_0+1, \ldots\} : |a_i - a| \leq 2\epsilon.$$

Und wenn wir Lemma 2.13 verwenden mit $q = 2$, erhalten wir $\lim_{i\to\infty} a_i = a$. □

Satz 2.22. Sei K ein geordneter Körper. Dann sind äquivalent:

(1) K erfüllt die Supremumseigenschaft
(2) K ist archimedisch und vollständig

Beweis. „(1) $\Longrightarrow$ (2)“: Angenommen K erfülle die Supremumseigenschaft. Wir haben bereits gesehen, dass dann K archimedisch ist.

Sei nun $(a_i)_{i\in\mathbb{N}}$ eine Cauchy-Folge. Wir definieren

$$\begin{aligned} M &:= \{x \in K \mid \text{Für fast alle } i \in \mathbb{N} \text{ gilt: } x \leq a_i\}, \\ &= \{x \in K \mid \exists i_0 \in \mathbb{N} : \forall i \in \{i_0, i_0+1, \ldots\} : x \leq a_i\}. \end{aligned}$$

Jede Cauchy-Folge ist beschränkt. Es gibt also r_1, r_2 mit $\forall i \in \mathbb{N} : r_1 \leq a_i \leq r_2$. Dann gilt $r_1 \in M$ und r_2 ist eine obere Schranke von M. Da M nicht-leer und nach oben beschränkt ist, existiert $a := \sup M \in K$.

[6] Man zeigt dazu zunächst $f(a+b) \geq f(a) + b$. Dann sieht man, dass es für $n := n_0 + (k_0 - f(n_0))$ erfüllt ist.

Wir wollen zeigen: $\lim_{i\to\infty} a_i = a$.

Sei $\epsilon \in K_{>0}$. Dann ist $a + \epsilon \notin M$, denn sonst wäre a keine obere Schranke von M. Also

$$\forall i_0 \in \mathbb{N} : \exists i \in \{i_0, i_0 + 1, \ldots\} : a + \epsilon > a_i. \tag{2.23}$$

Andererseits gilt: $x' \leq x \in M \implies x' \in M$. Wäre $a - \epsilon \notin M$ wahr, dann wäre auch $a - \epsilon$ eine obere Schranke, also a nicht das Supremum. Somit wissen wir $a - \epsilon \in M$. Wir erhalten

$$\exists k_0 \in \mathbb{N} : \forall k \in \{k_0, k_0 + 1, \ldots\} : a - \epsilon \leq a_k. \tag{2.24}$$

Sei nun $j_0 \in \mathbb{N}$ gegeben. Wir wählen ein k_0 wie in (2.24). Wende (2.23) mit $i_0 := \max\{k_0, j_0\}$ an. Dann existiert ein $i \in \{i_0, i_0 + 1, \ldots\}$ mit $a + \epsilon > a_i$ und $a - \epsilon \leq a_i$, also mit $|a_i - a| \leq \epsilon$. Wir haben gezeigt

$$\forall \epsilon \in K_{>0} : \forall j_0 \in \mathbb{N} : \exists i \in \{j_0, j_0 + 1, \ldots\} : |a_i - a| \leq \epsilon. \tag{2.25}$$

Wir definieren nun $f : \mathbb{N} \longrightarrow \mathbb{N}$ rekursiv. Wähle $f(1) := 1$. Ist $f(n)$ gewählt, so wenden wir (2.25) mit $\epsilon := 1/(n+1)$ und $j_0 := f(n) + 1$ an. Für das so erhaltene i setzen wir $f(n+1) := i$. Dann ist $(a_{f(n)})_{n\in\mathbb{N}}$ eine Teilfolge und $|a_{f(n)} - a| \leq 1/n$. Daraus folgt $a = \lim_{n\to\infty} a_{f(n)}$. Da $(a_i)_{i\in\mathbb{N}}$ eine Cauchy-Folge mit einer konvergenten Teilfolge ist, folgt mit Lemma 2.19 dann $\lim_{i\to\infty} a_i = a$.

„(2) $\implies$ (1)“: Angenommen K ist archimedisch und vollständig. Sei $\emptyset \neq M \subset K$, $r \in K$, $\forall x \in M : x \leq r$. Für jedes $q \in \mathbb{N}$ finden wir ein $p \in \mathbb{N}$ mit $qr \leq p$. Die Menge

$$A_q := \{p \in \mathbb{Z} \mid \frac{p}{q} \text{ ist obere Schranke von } M\}$$

ist somit nicht-leer. Zu einem $x \in M$ bestimme nun $s \in \mathbb{Z}$ mit $s \leq qx < s + 1$ (analog zu Aufgabe 4a) auf Übungsblatt 6). Dieses s ist untere Schranke von $A_q \subset \mathbb{Z}$. Jede nach

unten beschränkte nicht-leere Teilmenge von $\mathbb{Z}$ hat ein Minimum. [7] Somit existiert

$$p_q := \min A_q \in \mathbb{Z}.$$

Dann ist also $a_q := \frac{p_q}{q} \in \mathbb{Q}$ eine obere Schranke von M, wohingegen $\frac{p_q-1}{q}$ keine obere Schranke von M ist, d.h. es gibt ein $x_q \in M$ mit $a_q - (1/q) < x_q$. Es folgt für alle $q, i \in \mathbb{N}$: $a_q - (1/q) \leq a_i$ und dies wiederum besagt $|a_q - a_i| \leq \max\{1/q, 1/i\}$. Deswegen ist $(a_i)_{i\in\mathbb{N}}$ eine Cauchy-Folge, die auf Grund der Annahme gegen ein $a \in K$ konvergiert. Man sieht leicht, dass der Grenzwert von oberen Schranken wieder eine obere Schranke ist. Ebenso sind alle $a_i - (1/i)$ keine obere Schranke, und da K archimedisch ist, gibt es keine kleinere obere Schranke von M. Mit anderen Worten: a ist die kleinste obere Schranke von M, also $a = \sup M$. □

3. Existenz und Eindeutigkeit der reellen Zahlen

Wir wollen nun zeigen, dass es mindestens ein Modell der reellen Zahlen gibt (Existenzaussage), und dass je zwei Modelle (kanonisch) isomorph sind (Eindeutigkeitsaussage). Im Gegensatz zu Abschnitt 6.4 in Kapitel 2 wollen wir dies nicht mit Dedekindschen Schnitten, sondern mit Cauchy-Folgen durchführen. Wie so oft in der Mathematik ist es am besten mit der Eindeutigkeitsaussage anzufangen, denn der Beweis der Eindeutigkeit liefert entscheidende Ideen für den Beweis der Existenz. Die verwendete Technik heißt Vervollständigung und ist in leicht veränderter Form für viele Anwendungen der Mathematik (partielle Differentialgleichungen, Quantenmechanik, Allgemeine Relativitätstheorie,...) sehr wichtig.

[7] Denn ist A solche eine Teilmenge und s eine unter Schranke dann ist $\{p - s \mid p \in A\} \subset \mathbb{N}$ und hat deswegen nach Proposition 3.3 in Anhang 3 ein Minimum. Dann aber auch A.

Sei CF die Menge aller $\mathbb{Q}$-wertigen Cauchy-Folgen.

Angenommen $(K, +, \cdot, \leq)$ sei ein Modell der reellen Zahlen, d.h. $(K, +, \cdot, \leq)$ ist ein geordneter Körper mit Supremumseigenschaft. Wir betrachten wieder die injektive Abbildung $i_{\mathbb{Q}} : \mathbb{Q} \longrightarrow K$, die Addition, Multiplikation und die Ordnung erhält, siehe Lemma 5.7 in Kapitel 2. Für $(a_i)_{i\in\mathbb{N}} \in CF$ ist dann $(i_{\mathbb{Q}}(a_i))_{i\in\mathbb{N}}$ eine Cauchy-Folge in K. Da K die Supremumseigenschaft hat, existiert $f((a_i)_{i\in\mathbb{N}}) := \lim_{i\to\infty} i_{\mathbb{Q}}(a_i)$. Wir erhalten eine Abbildung $f : CF \longrightarrow K$.

Nun definieren wir eine Addition und Multiplikation auf CF:

$$\begin{aligned} + :&\quad CF \times CF \to CF \qquad (a_i)_{i\in\mathbb{N}} + (b_i)_{i\in\mathbb{N}} := (a_i + b_i)_{i\in\mathbb{N}} \\ \cdot :&\quad CF \times CF \to CF \qquad (a_i)_{i\in\mathbb{N}} \cdot (b_i)_{i\in\mathbb{N}} := (a_i \cdot b_i)_{i\in\mathbb{N}} \end{aligned}$$

Man sieht leicht, dass $(CF, +, \cdot)$ ein kommutativer Ring mit Eins ist. Hierbei ist $(0)_{i\in\mathbb{N}}$ das neutrale Element der Addition und $(1)_{i\in\mathbb{N}}$ das neutrale Element der Multiplikation. Es ist aber kein Körper, denn die Cauchy-Folge $(0, 1, 1, 1, \ldots)$ besitzt kein multiplikatives Inverses. Außerdem besagt Übung 2.10, dass f Addition und Multiplikation erhält.

Die Abbildung f ist nicht injektiv. Es gilt:

$$f((a_i)_{i\in\mathbb{N}}) = f((b_i)_{i\in\mathbb{N}}) \quad \Longleftrightarrow \quad (a_i - b_i)_{i\in\mathbb{N}} \text{ ist eine Nullfolge.}$$

Definition 3.1. Für $(a_i)_{i\in\mathbb{N}}, (b_i)_{i\in\mathbb{N}} \in CF$ definieren wir die Relation $\sim \subset CF \times CF$ durch

$$(a_i)_{i\in\mathbb{N}} \sim (b_i)_{i\in\mathbb{N}} \quad \Longleftrightarrow \quad (a_i - b_i)_{i\in\mathbb{N}} \text{ ist eine Nullfolge.}$$

Auf $\mathbb{R} := CF/\sim$ definieren wir Addition und Multiplikation wie folgt:

$$+ : \quad \mathbb{R} \times \mathbb{R} \to \mathbb{R} \qquad [(a_i)_{i\in\mathbb{N}}] + [(b_i)_{i\in\mathbb{N}}] := [(a_i + b_i)_{i\in\mathbb{N}}]$$

$$\cdot : \quad \mathbb{R} \times \mathbb{R} \to \mathbb{R} \qquad [(a_i)_{i\in\mathbb{N}}] \cdot [(b_i)_{i\in\mathbb{N}}] := [(a_i \cdot b_i)_{i\in\mathbb{N}}]$$

Die Ordnung definieren wir wie folgt:

$$[(a_i)_{i\in\mathbb{N}}] > [(b_i)_{i\in\mathbb{N}}] \quad :\Longleftrightarrow \quad \exists \epsilon \in \mathbb{Q}_{>0} : \text{ für fast alle } i \in \mathbb{N} \text{ gilt } a_i \geq b_i + \epsilon.$$

Man kann sich überlegen, dass hierdurch eine totale Ordnung auf $\mathbb{R}$ definiert wird.

Proposition 3.2. Die Addition und Multiplikation auf $\mathbb{R}$ sind wohldefiniert [8] und $(\mathbb{R}, +, \cdot, \leq)$ ist ein geordneter Körper. Die Abbildung $[\,\cdot\,] : CF \longrightarrow \mathbb{R}$, $(a_j)_{j\in\mathbb{N}} \mapsto [(a_j)_{j\in\mathbb{N}}]$ bewahrt die Addition, die Multiplikation und sendet die Eins von CF auf die Eins von $\mathbb{R}$. (Man nennt dies einen Homomorphismus von Ringen mit Eins.)

Beweisskizze. Der Beweis ist nun einfach, nur die Existenz eines multiplikativen Inversen ist etwas trickreich. Das Null-Element ist $[(0)_{i\in\mathbb{N}}] =: 0$.

Sei $[(b_i)_{i\in\mathbb{N}}] \neq 0$. Dann ist $(b_i)_{i\in\mathbb{N}}$ keine Nullfolge. Das heißt:

$$\exists \epsilon \in \mathbb{Q}_{>0} : \forall i_0 \in \mathbb{N} : \exists i \in \{i_0, i_0 + 1, \ldots\} : |b_i| \geq \epsilon$$

Wir [9] zeigen zunächst, dass $\{i \in \mathbb{N} \mid b_i = 0\}$ endlich ist. [10]

[8] „Wohldefiniert“ bedeutet hier: die Definition hängt nur von der Äquivalenzklasse ab und nicht von der Wahl des Repräsentanten. Da wir ja eine Abbildung definieren wollen, die einer Äquivalenzklasse etwas zurodnet, ist diese Eigenschaft das, was wir zeigen müssen, um zu sehen, dass diese Definitionen von + und · sinnvolle Definitionen sind.

[9] Beweis ab hier etwas anders als in der Vorlesung auf Grund einer Nachfrage

[10] Die leere Menge ist auch endlich, ist also hier nicht ausgeschlossen.

Angenommen $\{i \in \mathbb{N} \mid b_i = 0\}$ wäre unendlich groß, dann gibt es eine injektive Funktion $f : \mathbb{N} \longrightarrow \mathbb{N}$ mit $\forall n \in \mathbb{N} : \; b_{f(n)} = 0$. Dies bedeutet, dass $(b_i)_{i\in\mathbb{N}}$ eine Cauchy-Folge mit einer konvergenten Teilfolge ist. Wenn wir nun Lemma 2.19 anwenden, folgt daraus, dass $(b_i)_{i\in\mathbb{N}}$ eine Nullfolge ist. Das heißt wir erhalten den Widerspruch $[(b_i)_{i\in\mathbb{N}}] = 0$. Wir haben somit gesehen, dass $\{i \in \mathbb{N} \mid b_i = 0\}$ endlich ist.

Nun definieren wir

$$\tilde{b}_i := \begin{cases} b_i & \text{falls } b_i \neq 0 \\ 1 & \text{falls } b_i = 0 \end{cases}$$

Für fast alle $i \in \mathbb{N}$ gilt $b_i = \tilde{b}_i$, und somit bekommen wir auch $[(b_i)_{i\in\mathbb{N}}] = [(\tilde{b}_i)_{i\in\mathbb{N}}]$. Auf die Folge $(\tilde{b}_i)_{i\in\mathbb{N}}$ können wir nun Proposition 2.12 (4) anwenden. Wir sehen, dass

$$c_i := \frac{1}{\tilde{b}_i} = \begin{cases} (b_i)^{-1} & \text{falls } b_i \neq 0 \\ 1 & \text{falls } b_i = 0 \end{cases}$$

eine Cauchy-Folge ist.

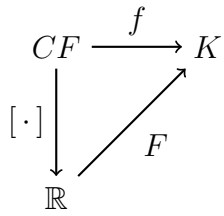
Es gilt dann für fast alle $i \in \mathbb{N}$: $c_i \cdot b_i = 1$, und deswegen gilt

$$[(c_i)_{i\in\mathbb{N}}] \cdot [(b_i)_{i\in\mathbb{N}}] = [(c_i)_{i\in\mathbb{N}}] \cdot [(\tilde{b}_i)_{i\in\mathbb{N}}] = [(1)_{i\in\mathbb{N}}].$$

□

Man beachte: die obige Definition, die obige Proposition und der Beweis benötigen die Existenz eines Modells nicht.

Beweis der Eindeutigkeit bis auf Isomorphie. Nehmen wir also wieder wie oben an, dass ein Modell $(K, +, \cdot, \leq)$ existiert. Dann gibt es genau eine Abbildung $F : \mathbb{R} \longrightarrow K$, so dass das folgende Diagramm kommutiert[11]



Alle Abbildungen hier erhalten Addition und Multiplikation und bilden die Eins auf die Eins ab. Außerdem ist F injektiv und erhält die Ordnung. Da R archimedisch geordnet ist, erhalten wir aus WS 2013/14 Aufgabe 4d) von Übungsblatt 6, dass $f : CF \longrightarrow K$ und somit $F : \mathbb{R} \longrightarrow K$ surjektiv ist. Also ist $F : CF \longrightarrow K$ ein Isomorphismus von geordneten Körpern und Satz 6.8 aus Kapitel 2 folgt im Fall $\widetilde{K} = \mathbb{R}$. Wenn nun $(K, +, \cdot \leq)$ und $(\widetilde{K}, \tilde{+}, \tilde{\cdot}, \widetilde{\leq})$ zwei Modelle der reellen Zahlen sind, so erhalten wir Isomorphismen $F : \mathbb{R} \longrightarrow K$ und $\widetilde{F} : \mathbb{R} \longrightarrow \widetilde{K}$. Dann ist auch $\widetilde{F} \circ F^{-1} : K \longrightarrow \widetilde{K}$ ein Isomorphismus und Satz 6.8 aus Kapitel 2 ist für alle $\widetilde{K}$ gezeigt. □

Wie bereits in Lemma 5.7 aus Kapitel 2 gesehen, gibt es nun eine injektive Abbildung $i_{\mathbb{Q}} : \mathbb{Q} \longrightarrow \mathbb{R}$, die Addition, Multiplikation und die Ordnung erhält. Es gilt hier $i_{\mathbb{Q}}(a) = [(a)_{i \in \mathbb{N}}]$, d.h. die rationalen Zahlen werden von konstanten Cauchy-Folgen repräsentiert.

Beweis Existenz. Wir zeigen, dass $(\mathbb{R}, +, \cdot, \leq)$ archimedisch und vollständig ist. Dann haben wir also ein Modell, d.h. Satz 6.7 aus Kapitel 2 ist gezeigt.

[11] „kommutiert“ bedeutet hier $F \circ [\,\cdot\,] = f$.

Zur archimedischen Eigenschaft: Sei ein Element $x \in \mathbb{R}$ gegeben. Zu zeigen ist: es gibt ein $n \in \mathbb{N}$ mit $x \leq i_{\mathbb{N}}(n)$. Es gibt nun eine $\mathbb{Q}$-wertige Cauchy-Folge $(a_i)_{i\in\mathbb{N}}$ mit $x = [(a_i)_{i\in\mathbb{N}}]$. Die Cauchy-Folge ist beschränkt, d.h. es gibt ein $r \in \mathbb{Q}$ mit $\forall i \in \mathbb{N} : |a_i| \leq r$. Da $\mathbb{Q}$ archimedisch ist, gibt es ein $n \in \mathbb{N}$ mit $r \leq n$. Man zeigt dann leicht, dass

$$x = [(a_i)_{i\in\mathbb{N}}] \leq [(n)_{i\in\mathbb{N}}] = i_{\mathbb{N}}(n).$$

Zur Vollständigkeit: Sei $(x_i)_{i\in\mathbb{N}}$ eine $\mathbb{R}$-wertige Cauchy-Folge. Falls es $a_i \in \mathbb{Q}$ gibt mit $x_i = i_{\mathbb{Q}}(a_i) = (a_i)_{n\in\mathbb{N}}$,[12] dann $(a_i)_{i\in\mathbb{N}} \in CF$ und es gilt dann

$$\lim_{i\in\mathbb{N}} x_i = [(a_i)_{i\in\mathbb{N}}].$$

Der allgemeine Fall ist ein bisschen aufwändiger und soll nur skizziert werden. Wir schreiben

$$x_i = [(a_{i,n})_{n\in\mathbb{N}}].$$

Man nutzt nun die Tatsache, dass sowohl $(x_i)_{i\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$ ist als auch dass für alle $i \in \mathbb{N}$ die Folge $(a_{i,n})_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{Q}$ ist, und konstruiert damit eine streng monoton wachsende Abbildung $f : \mathbb{N} \longrightarrow \mathbb{N}$,[13] so dass wiederum $(a_{i,f(i)})_{i\in\mathbb{N}}$ eine Cauchy-Folge ist. Man nennt solch eine Folge eine Diagonalfolge. Außerdem kann man nun

$$\lim_{i\in\mathbb{N}} x_i = [(a_{i,f(i)})_{i\in\mathbb{N}}]$$

zeigen. □

Ab jetzt identifizieren wir $\mathbb{Q}$ mit seinem Bild in $\mathbb{R}$ vermöge $i_{\mathbb{Q}}$.

---

[12] Hier ist kein Druckfehler. Gemeint ist die konstante Folge, die konstant a_i ist.

[13] Hier steckt etwas Konstruktionsarbeit, die wir (momentan) überspringen.

ANHANG Z

Überblick über algebraische Strukturen

(Aa) **Addition ist assoziativ.**

Für alle $x, y, z \in X$ gilt

$$(x + y) + z = x + (y + z).$$

(An) **Addition hat neutrales Element.**

Es gibt ein Element $0 \in X$, so dass für alle $x \in X$ gilt

$$x + 0 = 0 + x = x.$$

Man nennt 0 das neutrale Element der Addition.

(Ai) **Addition hat inverse Elemente.**

Zu jedem $x \in X$ gibt es ein $y \in X$, so dass

$$x + y = y + x = 0.$$

Man nennt y das Inverse von x bezüglich der Addition und schreibt normalerweise $-x$ anstelle von y.

(Ak) **Addition ist kommutativ.**

Für alle $x, y \in X$ gilt

$$x + y = y + x.$$

(Ma) **Multiplikation ist assoziativ.**

Für alle $x, y, z \in X$ gilt

$$(x \cdot y) \cdot z = x \cdot (y \cdot z).$$

(Mn) **Multiplikation hat neutrales Element.**

Es gibt ein Element $1 \in X$, so dass für alle $x \in X$ gilt

$$x \cdot 1 = 1 \cdot x = x.$$

Man nennt 1 das neutrale Element der Multiplikation.

(Mi) **Multiplikation hat inverse Elemente.**

Zu jedem $x \in X \setminus \{0\}$ gibt es ein $y \in X$, so dass

$$x \cdot y = y \cdot x = 1.$$

Man nennt y das Inverse von x bezüglich der Multiplikation und schreibt normalerweise x^{-1} anstelle von y.

(Mk) **Multiplikation ist kommutativ.**

Für alle $x, y \in X$ gilt

$$x \cdot y = y \cdot x.$$

(AMd) **Addition und Multiplikation erfüllen das Distributionsgesetz.**

Für alle $x, y, z \in X$ gilt

$$x \cdot (y + z) = x \cdot y + x \cdot z$$

$$(y + z) \cdot x = y \cdot x + z \cdot x$$

Literaturverzeichnis

[1] H. Amann, J. Escher, Analysis I, Birkhäuser

[2] H. Amann, J. Escher, Analysis II, Birkhäuser

[3] H. Amann, J. Escher, Analysis III, Birkhäuser

[4] H. Amann, Gewöhnliche Differentialgleichungen, de Gruyter

[5] G. Aumann, O. Haupt, Einführung in die reelle Analysis, de Gruyter

[6] M. Barner, F. Flohr, Analysis I, de Gruyter

[7] M. Barner, F. Flohr, Analysis II, de Gruyter

[8] C. Blatter, Analysis 1, Springer

[9] T. Bröcker, Analysis I, Spektrum

[10] T. Bröcker, Analysis II, BI Wissenschaftsverlag

[11] E. A. Coddington, N. Levinson, Theory of ordinary differential equations, McGraw Hill

[12] O. Deiser, Einführung in die Mengenlehre, Springer

[13] H.D. Ebbinghaus, Einführung in die Mengenlehre, BI Wissenschaftsverlag

[14] O. Forster, Analysis 1, Vieweg Studium

[15] O. Forster, Analysis 2, Vieweg Studium

[16] O. Forster, Analysis 3, Vieweg Studium

[17] U. Friedrichsdorf, A. Prestel, Mengenlehre für den Mathematiker, Vieweg Studium

[18] D. Grieser, Analysis I, Spriger Studium Mathematik

[19] D. Grieser, Vorlesungsskript Analysis I, Universität Oldenburg

[20] P. R. Halmos, Naive Mengenlehre, Vandenhoeck und Ruprecht

[21] S. Hildebrandt, Analysis I, Springer

[22] K. Königsberger, Analysis 1, Springer

[23] K. Königsberger, Analysis 2, Springer

[24] S. Lang, Real and functional analysis, Springer

[25] C. Löh, Vorlesungsskript Analysis I, Universität Regensburg

[26] C. Löh, Vorlesungsskript Analysis II, Universität Regensburg

[27] A. Peyerimhoff, Gewöhnliche Differentialgleichungen I und II, Akademische Verlagsgesellschaft Wiesbaden

[28] L. Perko, Differential equations and dynamical systems, Third edition, Springer

[29] B. v. Querenburg, Mengentheoretische Topologie, Springer

[30] W. Rudin, Analysis, Oldenbourg

[31] W. Rudin, Reelle und komplexe Analysis, Oldenbourg

[32] L. A. Steen, J. A. Seebach, Counterexamples in Topology

[33] W. Walter, Analysis 1, Springer

[34] W. Walter, Analysis 2, Springer

[35] W. Walter, Gewöhnliche Differentialgleichungen, Springer

Stichworte

C^1-Diffeomorphismus, 367
C^ℓ-Diffeomorphismus, 366
C^ℓ-Immersion, 385
n-dimensionale Sphäre, 372
p-Norm, 276
Äquivalenzklasse, 482
Äquivalenzrelation, 37
äquivalent, 10, 14
äquivalente Normen, 280
überabzählbar, 47, 147

Abbildung, 40
abgeschlossen, 203
abgeschlossene Hülle, 292
Ableitung, 215, 320
  der Umkehrfunktion, 220
Abschluss, 292
Absolutbetrag, 82
absolute Konvergenz, 160
abzählbar, 46
Addition, 473
  von Dedekindschen Schnitten, 97
Allquantor, 25
alternierende Reihe, 161
analytisch, 273
Anfangswert, 417
antisymmetrisch, 37
archimedisch geordneter Körper, 80
archimedisches Axiom, 80
Attraktor, 452
auf Y induzierte Topologie, 290
Aussage, 8
Aussageform, 11
aussagenlogische Formel, 12
aussagenlogische Verknüpfung, 10
Aussonderungsmengenaxiom, 460
Auswahlaxiom, 461
autonom, 407
Axiom
  archimedisches, 80
Axiome der Mengenlehre, 460
Axiome der reellen Zahlen, 6

Ball , 203

Banachraum, 284
beschränkt, 87, 121, 484
beschränkt im normierten Raum, 278
beschränkte Funktion, 276
Betragsfunktion, 82
  der komplexen Zahlen, 115
bijektiv, 44
Bild, 47
Bild der Relation, 40
Binomialkoeffizient, 68
Bolzano-Weierstraß
  Satz von, 140, 281

Cauchy-Folge, 141, 488
Cauchy-Produkt, 171

Dedekindscher Schnitt, 94
Definitheit der Norm, 276
Definitionsbereich, 40
Definitionsmenge, 19
Dezimal-Darstellung, 144
dicht, 294
Diffeomorphismus, 367
Differential, 320
Differentialgleichung
  autonome, 407
  explizite gewöhnliche, 412
  gewöhnliche, 406
  homogen, 433
  inhomogen, 433
  lineare, 433
  partielle, 408
Differenz
  symmetrische, 23
Differenzenquotient, 215
differenzierbar, 215, 320
disjunkt, 24
diskrete Topologie, 289
Divergenz
  von Folgen, 122, 485
Dreiecksungleichung der Norm, 276

Einbettung, 391
Einschränkung, 43
elliptisches Paraboloid, 353
endlich, 46
endliche Teilüberdeckung, 307
Entwicklungspunkt, 229
Erhaltungsgröße, 412
Ersetzungsaxiom., 461
erstes Integral, 412
erweiterten reellen Zahlen, 133
euklidischen Abstand, 202
Eulersche Zahl, 180
Existenzquantor, 25
explizite gewöhnliche Differentialgleichung, 412
explizite lineare gewöhnliche
    Differentialgleichung, 433

Exponentialfunktion, 173
Exponentialreihe, 172
Extensionalität, 460
Extremum
  lokales, 222, 399

Fakultät, 473
Familie, 49
fast alle, 152, 493
Fläche, 375
Fluss, 446
Flusslinie, 412
Folge
  M-wertige, 49
  in M, 49
  von Funktionen, 260
Folgen, 121
  beschränkte, 121
  monotone, 130
folgenkompakt, 209, 302
folgenstetig, 186, 297
Formel
  aussagenlogische, 12
Fundamentalsatz der Algebra, 117
Fundierungsaxiom, 461
Funktion
  beschränkte, 276
  reell-wertige, 41
Funktionalgleichung
  der Exponentialfunktion, 173
  der Logarithmusfunktion, 196

Gödelscher Unvollständigkeitssatz, erster, 462
Gödelscher Unvollständigkeitssatz, zweiter, 462
ganze Zahlen, 72
Gaußschen Zahlenebene, 113
Gebiet, 337
geometrische Reihe, 150
geordneter Körper, 77
gewöhnliche Differentialgleichung, 406
  autonome, 407
  explizite, 412
  lineare, 433
gewöhnliche Differentialgleichung mit
    getrennten Variablen, 442
gleich mächtig, 46
gleichmäßig stetig, 312
gleichmäßig stetig auf $[a, b]$, 244
Grad, 117
Grad des Polynoms, 82
Graph, 41
Grenzwert, 122
Grenzwert der Folge, 485
Grenzwert von Funktionen, 211

Häufungspunkt der Folge, 132
Häufungspunkt einer Menge, 208
harmonische alternierende Reihe, 163

harmonische Reihe, 150
Hauptsatz der Differential- und Integralrechnung
  Teil I, 245
  Teil II, 248
Hausdorff-Eigenschaft, 295
Hausdorffraum, 295
Heine-Borel-Eigenschaft, 306
Hesse-Abbildung, 339
Hesse-Matrix, 338
Hilbertraum, 284
homöomorph, 312
Homöomorphismus, 311
homogen, 433
Homogenität der Norm, 276
Homomorphismus von Ringen mit Eins, 498
hyperbolisches Paraboloid, 353
Hyperfläche, 375

Identität, 42
imaginär, 113
Imaginärteil, 113
Immersion, 385
implizit definierte, 371
indefinit, 351
Induktionsanfang, 57, 470
Induktionsaxiom, 54, 465
Induktionsschritt, 57, 470
Induktionsvoraussetzung, 57, 470
induktiv, 461, 467
induzierte Metrik, 202, 278
  auf Teilmenge, 202, 259
  von Norm, 278
induzierte Norm
  auf Unterraum, 277
induzierte Topologie
  von Teilmenge, 290
Infimum, 87
inhomogen, 433
injektiv, 44
Inklusion, 42
Innere, 292
inneren Kern, 292
Integralkurve, 412
Integration
  durch Substitution, 251
  partielle, 251
Intervalle, 133
Isometrie, 202
isomorph (als Ring), 74
Isomorphismus von Ringen, 74

Jacobi-Matrix, 329

k-mal stetig differenzierbar, 226
Körper, 76
  archimedisch geordneter, 80
  der komplexen Zahlen, 112

der rationalen Zahlen, 74
der reellen Zahlen, 84
geordneter, 77
Körper der rationalen Funktionen, 81
kanonische Projektion, 482
Klassen, 460
Klumpentopologie, 289
Koeffizienten, 117
kommutativer Ring, 74
kompakt, 307
Komplement, 23
komplexe Exponentialfunktion, 173
komplexe Konjugation, 115
komplexe Zahlen, 113
Komponente, 52
Komposition, 43
Komprehension, 460
Kontraktion, 285
Konvergenz
uneigentliche, 134
von Folgen, 122
Konvergenzradius, 157
konvergiert, 206
absolut, 160, 254
gegen ∞, 133
gegen unendlich, 133
gleichmäßig, 262
in $\overline{\mathbb{R}}$, 134
punktweise, 260
Kosinus, 176
Kreiszahl π, 199
kritischen Punkt, 356
kritischer Punkt, 399, 452

Lösung, 406
Lösungsintervall, 406
Lagrange-Multiplikatoren, 400
Lagrangesche Restglieddarstellung, 228
Lebesguesche Zahl, 314
leere Menge, 20
Leibniz-Regel, 161
Leitkoeffizient, 117
Limes, 122
Limes inferior, 137
Limes superior, 138
Linearisierung, 454
linksseitige Grenzwert oder Limes, 212
Lipschitz-Konstante, 416
Ljapunow-Funktion, 455
Logarithmus, 196
lokal Orts-Lipschitz-stetig, 416
lokal umkehrbar, 360
lokale C^ℓ-Parametrisierung, 391
Lokaler Umkehrsatz, 360
lokales Extremum, 222, 355, 399
lokales Maximum, 222, 355, 399
lokales Maximum oder Minimum, 355

lokales Minimum, 222, 355, 399
lokales striktes Maximum, 355
lokales striktes Minimum, 355
lokalkonstant, 336

mächtiger, 46
Majorante, 153
Majoranten-Kriterium, 152
maximale Lösungs-Intervall, 426
Maximum, 61, 477
  lokales, 355, 399
  lokales striktes, 355
Menge, 18
  der ganzen Zahlen, 72
  der komplexen Zahlen, 112
  der natürlichen Zahlen, 54, 473
  der rationalen Zahlen, 74
  der reellen Zahlen, 84
  leere, 20
Mengensystem, 30
Metrik, 201
metrischer Raum, 201, 257, 285
Minimum, 61, 477
  lokales, 355, 399
  lokales striktes, 355
Minkowski-Ungleichung, 277
Mittel
  arithmetisches, 83
  geometrisches, 83
Mittelwertsatz
  erster, 224
  zweiter, 224
Modell der natürlichen Zahlen, 465
Multi-Indizes, 348
Multiplikation, 473
  von Dedekindschen Schnitten, 100
Multiplizität der Nullstelle, 119

nach oben beschränkt, 87
nach unten beschränkt, 87
Nachfolger-Abbildung, 54, 465
Natürliche Zahlen, 4
Negation, 9
negativ definit, 351
negativ semi-definit, 351
nicht zusammenhängend, 299
Norm, 275
normierten Vektorraum, 276
Nullfolgen, 122, 485
Nullstelle, 117

obere Schranke, 87
Oberintegral, 235
Oberklasse, 94
offen, 203
offene Überdeckung, 306
offene Menge in $(X, \mathcal{O})$, 289
Ordnung, 37

partielle, 37
totale, 38
Ordnungsrelation, 37
Orts-Lipschitz-stetig, 416

Paar, (geordnetes), 32
Paarmengenaxiom, 461
Paraboloid
elliptisches, 353
hyperbolisches, 353
Parametrisierung
lokale ∼ einer Untermannigfaltigkeit, 391
Partialsumme, 148
partiell differenzierbar, 324
partielle Ableitung, 324
partielle Differentialgleichung, 408
Partielle Integration, 251
partielle Ordnung, 37
Partition, 233
Peano-Axiome, 53, 465
Permutation, 66
polynomiale Funktion, 117
positiv definit, 351
positiv semi-definit, 351
Potenzieren, 473
Potenzmenge, 30
Potenzmengenaxiom, 461
Potenzreihe, 156
Produkt
von Dedekindschen Schnitten, 100
Produkt, (kartesisches), 32
Produkt, (kartesisches,) von Mengenfamilien, 51
Produktregel, 218
Produktreihe, 169
Produktreihensatz, 169
Produkttopologie, 290
Produktzeichen, 55
Punkt
kritischer, 399
stationärer, 399

Quadrupel, 35
Quintupel, 35
Quotienten-Kriterium, 154
Quotientenregel, 219

Rand, 293
rationale Zahlen, 74
Realteil, 113
rechtsseitige Grenzwert oder Limes, 212
reell-analytisch, 273
reelle Exponentialfunktion, 173
Reelle Zahlen, 84
reflexiv, 37
Regel
von Leibniz, 161
regulärer Wert, 379
Reihe, 149

alternierende, 161
rein imaginär, 113
rekursive Definition, 471
Relation, 36
Relation, funktionale, 39
Repräsentant, 482
repräsentiert, 482
Restglied, 227
Restriktion, 43
Richtungsableitung, 324
Riemann-Integral, 236
Riemann-integrierbar, 236
Riemannscher Umordnungssatz, 164
Riemannsches Kriterium, 236
Ring, 74
der ganzen Zahlen, 72
kommutativer, 74
Ring mit Eins, 74
Russellsche Paradoxon, 457

Satz
über 1. Ableitung in Extrema, 222
über 2. Ableitung in Extrema, 226
vom Cauchy-Produkt, 171
von Bolzano-Weierstraß für endl.-dim. Vektorräume, 281
von Bolzano-Weierstraß, 141
von Cauchy-Hadamard, 158
von Dedekind, 107
von Rolle, 223
von Schröder-Bernstein, 46
von Taylor, 227
von Taylor (mehr-dimensional), 347
von Taylor in Multi-Index-Notation, 349
Schnitt, 22
Schnittmenge, 22
singulären Wert, 379
Sinus, 177
Spurtopologie, 290
stabil, 452
Stammfunktion, 247
stationären Punkt, 356
stationärer Punkt, 399
sternförmig, 301
stetig, 187, 202, 294
stetig differenzierbar, 225, 326
stetig in x_0, 294
streng monoton wachsend, 487
Substitution, 251
Summe
von Dedekindschen Schnitten, 97
Summenzeichen, 55
Supremum, 87
Supremums-Distanz, 262
Supremumseigenschaft, 88
Supremumsnorm, 276
surjektiv, 44

symmetrisch, 37

Tangentialraum, 396
Taylor, Satz von, 347
Taylor-Polynom, 229
Taylorreihe, 229
Teilüberdeckung
  endliche, 307
Teilfolge, 131, 487
Teilmenge, 20
  echte, 20
Teilsumme, 148
tertium non datur, 9
Topologie, 289
topologischer Raum, 289
total differenzierbar, 320
totale Ordnung, 38
transitiv, 37
Trennungszahl, 107
Treppenfunktion, 233
Tripel, 35

Umgebung, 203
Umkehrabbildung, 45
Umkehrsatz
  Lokaler, 360
Umkehrung, 45
Umordnung, 163
uneigentliche Konvergenz, 134
Uneigentliches Riemann-Integral, 253
unendlich, 46
Unendlichkeitsaxiom, 461
untere Schranke, 87
Unterintegral, 235
Unterklassen, 94
Untermannigfaltigkeit, 374
Untermannigfaltigkeitskarte, 375
Urbild, 48

Vektorfeld, 412
Vereinigung, 22
Vereinigungsaxiom, 461
Verkettung, 43
Verknüpfung
  aussagenlogische, 10
Vervollständigung, 496
vollständig, 142, 259, 492
Vollständige Induktion, 57, 470
vollständiges Vektorfeld, 449

Wahrheitstafel, 9
wegzusammenhängend, 301
Wert
  regulärer, 379
wohldefiniert, 80
wohlgeordnete Menge, 479
Wohlordnung, 479
Wohlordnung von $\mathbb{N}$, 61

Wronski-Determinante, 438
Wurzel-Kriterium, 155

Zerlegung, 233
Zielbereich, 41
zusammenhängend, 299

Symbole

$(e_1, \ldots, e_n)$, Standardbasis von $\mathbb{R}^n$, 280

$:=$, 4

$B(D, \mathbb{K})$, beschränkte Funktionen, 276

$C^0((a,b))$, 226

$C^\infty((a,b))$, 226

$C^k((a,b))$, 226

L_σ, 380

$\mathcal{S}_k$, 380

$\cap$, $\cup$, $\setminus$, Δ, 23

$\circ$, 44

$\deg(P)$, 117

d_{glm}, 262

$d_{\sup}$, 262

$\emptyset$, 20

$\exists$,$\forall$, 25

exp, 173

$\Longleftrightarrow$, $\Longrightarrow$, 15

inf, Infimum, 87

$\lim_{i\to\infty} a_i$, 485

$\lim_{j\to\infty} a_j$, 122

$\lim_{x\to x_0}$, Grenzwert von Funktionen, 211

lim inf, Limes inferior, 137

lim sup, Limes superior, 137

$\mathbb{N}$, 4

$\mathbb{N}_{>0}$, 4

$\mathbb{Q}$, 74

$\mathbb{R}$, 19, 84

$\mathbb{R}_{>0}$, 19

$\mathbb{Z}$, 72

$\mathcal{P}(M)$, 30

max, Maximum, 61

min, Minimum, 61

$\neg$, 9

$\prod_{j=1}^n$, 55

$\lceil r \rceil$, 129

$\lfloor r \rfloor$, 129

$\subset$, $\subsetneq$, 20

$\sum_{j=1}^n$, 55

sup, Supremum, 87

$\times$, 32

$\wedge$, $\vee$, $\veebar$, $\to$, $\leftarrow$, $\leftrightarrow$, 10

$\{\}$, 20

$a_i \to a$, 485

$a_j \to a$, 122

f', Ableitung von f, 215
$f^{\#}(N)$, Urbild von N unter f, 48
$f^{(k)}$, k-te Ableitung, 226
$f_{\#}(M)$, Bild von M unter f, 47
p-Norm, 276