

Oliver Deiser

Erste Hilfe in Analysis

für Caroline, Thalia und Larina

Inhalt

Vorwort	5
Die Themen des Buches	9
1. Kapitel Grundlegendes	13
1.1 Mengen	14
1.2 Umgang mit Mengen	16
1.3 Relationen	18
1.4 Funktionen	20
1.5 Visualisierung von Funktionen	22
1.6 Injektive, surjektive und bijektive Funktionen	24
1.7 Umgang mit Funktionen	26
1.8 Umkehrfunktionen und Einschränkungen	28
1.9 Bild und Urbild	30
1.10 Umgang mit Quantoren	32
1.11 Die vollständige Induktion	34
1.12 Das Prinzip vom kleinsten Element	36
2. Kapitel Die reellen und komplexen Zahlen	39
2.1 Irrationale Zahlen	40
2.2 Algebraische und transzendente Zahlen	42
2.3 Abzählbarkeit und Überabzählbarkeit	44
2.4 Die Körperaxiome	46
2.5 Die Anordnungsaxiome	48
2.6 Supremum und Infimum	50
2.7 Die Vollständigkeit	52
2.8 Die Dezimaldarstellung	54
2.9 Die Intervallschachtelung	56
2.10 Wurzeln und rationale Exponenten	58
2.11 Komplexe Zahlen	60

2.12 Umgang mit komplexen Zahlen	62
3. Kapitel Folgen und Grenzwerte	65
3.1 Folgen	66
3.2 Grenzwerte von Folgen	68
3.3 Monotone Folgen und Pendelfolgen	70
3.4 Die Limesregeln	72
3.5 Cauchy-Folgen	74
3.6 Teilfolgen	76
3.7 Häufungspunkte von Folgen	78
3.8 Der Satz von Bolzano-Weierstraß	80
3.9 Limes Superior und Inferior	82
3.10 Offene Epsilon-Umgebungen	84
3.11 Konvergenz in den komplexen Zahlen	86
3.12 Die Unendlichkeitssymbole	88
4. Kapitel Reihen	91
4.1 Unendliche Reihen	92
4.2 Folgen versus Reihen	94
4.3 Die geometrische Reihe	96
4.4 Dezimaldarstellungen als Reihen	98
4.5 Die harmonische Reihe	100
4.6 Das Cauchy-Kriterium für Reihen	102
4.7 Das Leibniz-Kriterium	104
4.8 Absolute und bedingte Konvergenz	106
4.9 Majorantenkriterium und Minorantenkriterium	108
4.10 Wurzelkriterium und Quotientenkriterium	110
4.11 Produkte von Reihen	112
4.12 Die Exponentialreihe	114
5. Kapitel Stetigkeit	117
5.1 Die Limesstetigkeit	118
5.2 Grenzwerte von Funktionen	120
5.3 Unstetigkeiten	122
5.4 Die Umgebungsstetigkeit	124
5.5 Die gleichmäßige Stetigkeit	126

5.6 Die Lipschitz-Stetigkeit	128
5.7 Stetige Fortsetzungen	130
5.8 Der Zwischenwertsatz	132
5.9 Der Extremwertsatz von Weierstraß	134
5.10 Die Stetigkeit der Umkehrfunktion	136
5.11 Punktweise und gleichmäßige Konvergenz	138
5.12 Potenzreihen	140
6. Kapitel Elementare Funktionen	143
6.1 Polynome	144
6.2 Rationale Funktionen	146
6.3 Die reelle Exponentialfunktion	148
6.4 Der natürliche Logarithmus	150
6.5 Die allgemeine Exponentialfunktion	152
6.6 Der allgemeine Logarithmus	154
6.7 Die komplexe Exponentialfunktion	156
6.8 Bilder der komplexen Exponentialfunktion	158
6.9 Sinus und Kosinus	160
6.10 Weitere trigonometrische Funktionen	162
6.11 Die Arkusfunktionen	164
6.12 Die Brücke zur Geometrie	166
7. Kapitel Differentiation	169
7.1 Geraden und ihre Darstellungen	170
7.2 Differenzen- und Differentialquotienten	172
7.3 Lineare Approximationen	174
7.4 Ableitungsregeln	176
7.5 Differenzierbarkeit und Stetigkeit	178
7.6 Höhere Ableitungen	180
7.7 Der Mittelwertsatz der Differentialrechnung	182
7.8 Ableitung und Monotonie	184
7.9 Lokale Extremwerte	186
7.10 Konvexität	188
7.11 Krümmungsverhalten	190
7.12 Die Taylor-Entwicklung	192

8. Kapitel Integration	195
8.1 Partitionen und Treppenfunktionen	196
8.2 Das Riemann-Integral	198
8.3 Das Darboux-Integral	200
8.4 Das Regelintegral	202
8.5 Eigenschaften des Integrals	204
8.6 Zum Umfang der integrierbaren Funktionen	206
8.7 Der Mittelwertsatz der Integralrechnung	208
8.8 Der Hauptsatz	210
8.9 Integrationsregeln	212
8.10 Uneigentliche Integrale	214
8.11 Der Vertauschungssatz	216
8.12 Integral und Flächeninhalt	218
Zum Studium der Mathematik	221
Anhänge	227
1. Junktoren	228
2. Quantoren	230
3. Axiome für die reellen Zahlen (2.4 – 2.7)	231
4. Epsilontik	232
5. Grenzwerte von Folgen und unendliche Summen	234
6. Reihenentwicklungen	235
7. Ableitungen	236
8. Integrierbare Funktionen	237
Literatur	238
Notationen	239
Index	242

Vorwort

Die Analysis gehört zusammen mit der linearen Algebra zu den Grundvorlesungen eines Mathematikstudiums an der Universität. Viele Studenten hören sie im ersten Semester. Bis man aber zur aus der Schule vertrauten Differential- und Integralrechnung kommt, vergehen Wochen, die mit Körperaxiomen, kleinsten oberen Schranken, Grenzwerten, Häufungspunkten und unzähligen kleinen Epsilons und Deltas angefüllt sind. Beweisen soll, wer rechnen geübt hat. Definieren muss, wer bislang mit einigermaßen genauen Beschreibungen zurechtkam. Für viele ist dieser methodische Neubeginn eine Herausforderung.

Dieses Buch möchte ein hilfreicher Begleittext für die Hörer einer Analysis-Vorlesung an der Universität sein (für das Fachstudium und das Lehramt an Gymnasien). Es will das Vorlesungsskript und die ausführliche und systematische Lehrbuchliteratur ergänzen. Im Gegensatz zu diesen Texten enthält es nur wenige Beweise und die Darstellung entspricht eher einem Garten als einem sich ausbreitenden Fluss. Zentrale Gegenstände der elementaren Analysis wurden ausgewählt, um sie möglichst übersichtlich, informativ und einprägsam in zweiseitigen Sektionen vorzustellen und zu diskutieren. Jede Sektion beginnt mit einer Definition oder einem Satz. Es folgen ein Diagramm, Erläuterungen, zahlreiche Beispiele und Gegenbeispiele, manchmal auch weitere Begriffe, Sätze und

Diagramme. Eingestreut sind zudem Bemerkungen zur Verwendung des behandelten Gegenstandes und zu möglichen Fehlerquellen, sowie Zusammenfassungen und Tabellen. So werden Hilfestellungen angeboten und der Überblick erleichtert, damit der Wald vor lauter Bäumen nicht aus den Augen gerät. Die mathematische Genauigkeit wird dabei nicht verwässert. Viele der angeführten Beispiele zeigen, wie wichtig die Beachtung der Voraussetzungen mathematischer Sätze ist.

Keineswegs tritt der Autor dafür ein, Systematik und vollständige fachsprachliche Beweise aus den Anfängervorlesungen und den zugehörigen Lehrbüchern zu verbannen – im Gegenteil. Aber die Erfahrung zeigt, dass viele Studienanfänger wichtige Definitionen und Sätze nicht in der Genauigkeit wiedergeben können, die für ein wissenschaftliches mathematisches Arbeiten notwendig ist, und dass sie diese Definitionen und Sätze oft auch nicht mit anschaulichen Vorstellungen und klärenden Beispielen verbinden können, die Verstehen, Einprägen und Anwenden erleichtern. Hier will der vorliegende Text eine helfende Hand reichen. Die Fähigkeit zum eigenständigen Erarbeiten von Anschauungen und Beispielen ist dabei Lernziel.

Verwendung des Buches

- (1) Zu vielen Themen einer im deutschsprachigen Raum üblichen Analysis-Vorlesung wird der Leser passende Sektionen in diesem Buch finden. Eine begleitende, regelmäßig wiederholende, vergleichende und evtl. auch vorgegreifende Lektüre bietet sich hier an.
- (2) Während der Vorlesung kann kontinuierlich überprüft werden, ob man fundamentale Definitionen und Sätze korrekt wiedergeben kann und ob man zugehörige Beispiele kennt. Die Struktur des Buches erleichtert diese Form der Selbstkontrolle und eine entsprechende Anpassung des eigenen Lernverhaltens.
- (3) Das Buch kann zur systematischen Vorbereitung auf eine schriftliche oder mündliche Prüfung benutzt werden.
- (4) In späteren Semestern kann das Buch zum Nachschlagen verwendet werden.

Was das Buch nicht sein will

Das Studium dieses Buches kann dem Leser helfen, seine Hausaufgaben leichter und erfolgreicher zu bearbeiten, und einige der vorgeführten Beispiele wird er vielleicht sogar als Übungsaufgaben wiederfinden. Angestrebt wurde mit dieser „Ersten Hilfe“ aber kein Text, dessen Hauptzweck es ist, sammelnd vorzuführen, wie man lösen kann, was man lösen muss. Ziel ist, ein vertieftes und tragfähiges Verständnis mathematischer Ideen zu vermitteln. Ein solches Verständnis trägt nach Meinung des Autors automatisch dazu bei, den vielfältigen Anforderungen eines Mathematikstudiums – einschließlich Übungen und Klausuren – selbstbewusster, erfolgreicher und freudvoller begegnen zu können.

Zum Inhalt des Buches

Im Anschluss an dieses Vorwort findet der Leser einen Überblick über die Themen der acht Kapitel des Haupttextes. Dem Haupttext folgen allgemeine Bemerkungen zum Mathematikstudium. In den Anhängen stellen wir Regeln über Junktoren und Quantoren zusammen, und es finden sich dort Tabellen zu den Axiomen für die reellen Zahlen, zur Epsilontik, zu Grenzwerten, Summen und Reihenentwicklungen, zu den Ableitungen der elementaren Funktionen und zu den Integralbegriffen. Das Buch endet mit einem Literaturverzeichnis, einem Überblick über Notationen und einem Index.

München, im März 2021

Oliver Deiser

Die Themen des Buches

Wo beginnt die Analysis als mathematische Disziplin? Viele würden sagen: beim Begriff des Grenzwerts, in seinen vielen verschiedenen Varianten und Spielarten. Stillschweigend wird dabei vorausgesetzt, dass die reellen Zahlen \mathbb{R} bekannt sind und dass Mengen und Funktionen sprachlich so weit zur Verfügung stehen, dass man mit der Untersuchung des Grenzwertbegriffs beginnen kann. Denn Mengen und Funktionen sind in allen mathematischen Disziplinen unentbehrlich. Und die reellen Zahlen bilden ebenfalls eine der Grundstrukturen der Mathematik, der man überall begegnet – auch in der Zahlentheorie, die oft die reellen Zahlen einsetzt, um die natürlichen Zahlen zu untersuchen.

Traditionell beginnt ein Mathematikstudium an der Universität mit den beiden Vorlesungen „Lineare Algebra“ und „Analysis“. Für die überall verwendeten Grundlagen ist keine eigene Vorlesung vorgesehen und sie werden deswegen in die beiden Vorlesungen integriert. Diese weisen dann oft längere Vorspanne auf, die sich mit Aussagen, Junktoren, Mengen, Relationen, Funktionen, vollständiger Induktion, Gruppen-, Ringen- und Körperaxiomen sowie dem Zahlensystem bis hin zu den komplexen Zahlen beschäftigen. Für die Analysis entspricht dieser Vorspann den beiden ersten Kapiteln dieses Buches. Die entsprechende Gestaltung ist aber in jeder Vorlesung anders, denn man kann diese Dinge sehr gründlich oder auch sehr kurz diskutieren. Der kürzeste Weg ist vielleicht, zu sagen, dass die reellen Zahlen \mathbb{R} als Punkte einer Linie und als Dezimalzahlen bekannt sind und dass \mathbb{R} durch die Existenz kleinster oberer Schranken ausgezeichnet ist. Zusammen mit einem naiven Funktionsbegriff gelangt man dann zügig zu den Grenzwerten. Bei einem sehr ausführlichen Weg dagegen wird die universelle mengentheoretische Sprache der Mathematik gründlich vorgestellt und einer axiomatischen Beschreibung der reellen Zahlen eine Konstruktion der reellen Zahlen mit Hilfe der rationalen Zahlen hinzugefügt. Beide Wege sind verbreitet, und daneben auch zahlreiche Mischformen und Kompromisse. Unsere Darstellung entspricht einem solchen Kompromiss. Wir behandeln Mengen naiv, Funktionen genauer, eine Axiomatisierung von \mathbb{R} im Überblick (ohne „Aus-schlachten“ der Axiome), eine Konstruktion von \mathbb{R} dagegen gar nicht und einen allgemeinen Grenzwertbegriff für metrische Räume deuten wir nur an.



Der Text ist in acht Kapitel gegliedert, die in je zwölf Sektionen unterteilt sind. Wir beginnen mit Grundlagen über Mengen, Relationen, Funktionen, Quantoren und Induktion. Wir beschränken uns dabei auf einen „naiven“ Mengenbegriff. Relationen und Funktionen führen wir dagegen exakt als gewisse Mengen geordneter Paare ein, d. h., wir identifizieren Relationen und Funktionen mit ihren Graphen. Der graphischen Visualisierung von Funktionen ist eine eigene Sektion gewidmet und vier weitere Sektionen besprechen grundlegende Dinge über Funktionen wie Injektivität, Einschränkung, Bild und Urbild. Es folgen eine Sektion über Quantorenregeln und zwei Sektionen über Induktion.

Im zweiten Kapitel betrachten wir die beiden Grundstrukturen der Analysis, die reellen Zahlen \mathbb{R} und die komplexen Zahlen \mathbb{C} . Ausgehend von der Existenz irrationaler Zahlen gelangen wir zu einer axiomatischen Charakterisierung der reellen Zahlen als vollständig angeordneter Körper. Im Zentrum steht das über größte obere und untere Schranken formulierte Vollständigkeitsaxiom und seine Anwendungen (Dezimaldarstellungen, Intervallschachtelung, Wurzeln). Die beiden letzten Sektionen behandeln komplexe Zahlen, wobei wir die komplexe Multiplikation von Anfang an geometrisch deuten.

Kapitel drei behandelt Folgen und ist von grundlegender Bedeutung für alles Weitere. Im Zentrum steht der Grenzwertbegriff, der in seinen vielen Spielarten die gesamte Analysis durchzieht. Nachdem wir den Grenzwertbegriff für Folgen eingeführt haben, untersuchen wir ihn an Folgen mit einer besonders einfachen Struktur. Neben den monotonen Folgen betrachten wir hier auch Pendelfolgen, deren Grenzwert man ebenfalls noch vor Augen sehen kann. Die Limesregeln eröffnen uns dann einen Kalkül für die Berechnung von Grenzwerten. Anschließend diskutieren wir Cauchy-Folgen und gehen dabei auch auf die Möglichkeit ein, die Vollständigkeit von \mathbb{R} mit ihrer Hilfe zu formulieren. Teilfolgen, Häufungspunkte, der Satz von Bolzano-Weierstraß sowie Limes Inferior und Superior bilden vier Sektionen, bei denen nicht wenige Anfänger ins Straucheln geraten. Neben vielen Beispielen illustrieren wir den Satz von Bolzano-Weierstraß durch einen anschaulichen Beweis. Anschließend formulieren wir unsere Grenzwertbegriffe mit offenen ε -Umgebungen, die uns neue suggestive Sprechweisen zur Verfügung stellen. Das Kapitel schließt mit Bemerkungen zur Konvergenz in \mathbb{C} und zur Verwendung der Unendlichkeitssymbole ∞ und $-\infty$, die wir bislang konsequent vermieden haben.

Das nächste Kapitel behandelt unendliche Reihen. Wir führen, wie heute üblich, unendliche Summen $x_0 + x_1 + x_2 + \dots + x_n \dots$ als Grenzwerte von Partialsummen ein, sodass diese Summen Grenzwerte von Folgen sind. Mit der geometrischen Reihe lernen wir eine der wichtigsten Reihen der Analysis kennen. Wir illustrieren ihren Einsatz bei der Untersuchung von Dezimaldarstellungen, später werden wir sie vor allem für Abschätzungen verwenden. Eine Diskussion der faszinierenden divergenten harmonischen Reihe darf nicht fehlen. Fünf Sektionen sind dann den Konvergenzkriterien für Reihen gewidmet. Danach besprechen wir Produkte von Reihen und insbesondere das Cauchy-Produkt. Die Exponentialreihe, die fortan die Rolle der Königin aller Reihen übernehmen wird, lernen wir in der letzten Sektion des Kapitels kennen.

Im fünften Kapitel besprechen wir den Begriff der Stetigkeit in vier Varianten: Limesstetigkeit, ε - δ -Stetigkeit, gleichmäßige Stetigkeit und Lipschitz-Stetigkeit. Gerade hier sind Diagramme wichtig, um abstrakte Anschauungen zu entwickeln. Im Umfeld der Limesstetigkeit definieren wir auch Grenzwerte für Funktionen (was wir prinzipiell auch bereits im Kapitel über Folgen hätten tun können). Nach einer knappen Diskussion über stetige Fortsetzungen stehen die Hauptsätze über stetige Funktionen im Vordergrund: der Zwischenwertsatz, der Extremwertsatz von Weierstraß über die Annahme von Minimum und Maximum, und der Satz über die Stetigkeit der Umkehrfunktion. Schließlich betrachten wir konvergente Funktionenfolgen und betonen, dass nur die gleichmäßige Konvergenz die Stetigkeit der Grenzfunktion sichert. Das Kapitel schließt mit einem Blick auf Potenzreihen und ihr Konvergenzverhalten.

Das sechste Kapitel ist den elementaren Funktionen gewidmet. Nach einer kurzen Diskussion von Polynomen und rationalen Funktionen stellen wir die Exponentialfunktion

an die Spitze und nutzen sie als Generator für andere Funktionen. Dabei gehen wir ausführlich auf die geometrischen Abbildungseigenschaften der komplexen Exponentialfunktion ein, sodass eine Definition von Sinus und Kosinus mit Hilfe dieser Funktion zwar überraschend, aber nicht unnatürlich ist, da der vertraute trigonometrische Gehalt klar zu sehen ist. In der letzten Sektion des Kapitels skizzieren wir, wie wir die geometrischen Eigenschaften der Exponentialfunktion auf rein analytischem Weg gewinnen können, sodass eine Brücke zur Geometrie sichtbar wird.

Das siebte Kapitel gibt einen knappen Überblick über die Grundlagen der Differentialrechnung. Wir beginnen mit einer ausführlichen Diskussion von Geraden, die die Definition von Differenzen-, Differentialquotienten und Tangenten vorbereitet. Unmittelbar nach der Definition der Ableitung rücken wir lineare Approximationen in den Vordergrund, die die Ableitung in einem neuen Licht erscheinen lassen und die spätere Taylor-Entwicklung vorbereiten. Nach den obligatorischen Ableitungsregeln betrachten wir ausführlich das Thema „Stetigkeit und Differenzierbarkeit“ und die höheren Ableitungen. Anschließend besprechen wir den Mittelwertsatz und erste Anwendungen, etwa die Bestimmung aller Funktionen mit $f' = f$. In den beiden sich anschließenden Sektionen über Ableitung, Monotonie und lokale Extrema ist vieles aus der Schule gut bekannt, Beispiele und Varianten vertiefen die Thematik. Zwei Sektionen sind dann dem Krümmungsbegriff gewidmet, und das Kapitel schließt mit dem Satz von Taylor und der Frage nach der Potenzreihenentwicklung einer reellen Funktion.

Im abschließenden achten Kapitel über Integration versuchen wir den verschiedenen verbreiteten Ansätzen, das Integral für Anfänger einzuführen, dadurch gerecht zu werden, dass wir sowohl das Riemann-, das Darboux- als auch das Regelintegral diskutieren. Daneben wird die elementare Integrationstheorie vorgestellt, vom Mittelwertsatz der Integralrechnung über den Hauptsatz und seine Anwendungen bis hin zur Vertauschbarkeit von Integration und Limesbildung. Wir schließen mit einer Betrachtung des Jordan-Inhalts von Punktmengen der Ebene, die die geometrische Bedeutung des Integrals noch einmal untermauert.

1. Kapitel

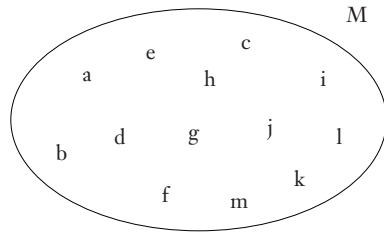
Grundlegendes

1.1 Mengen

Definition (Menge)

I M heißt eine *Menge*, falls ...

Unser Text beginnt mit einer „Definitions-lücke“. Mengen werden heute verwendet, um das Gebäude der Mathematik zu errichten – und bleiben undefiniert. Welche Mengen existieren, kann mit Hilfe eines Axiomensystems beschrieben werden. Man muss aber mit einer derartigen Axiomatik nicht in Berührung gekommen sein, um Mengen sicher anwenden zu können. Die Mengenlehre schenkt der modernen Mathematik ihre Sprache, und ein intuitives Grundverständnis reicht aus, um diese Sprache durch Nachahmung zu erlernen. Um die vorhandene Intuition zu schärfen, betrachten wir die Beschreibung des Mengenbegriffs von Georg Cantor aus dem Jahr 1895:



„Zusammenfassung zu einem Ganzen“

„Unter einer ‚Menge‘ verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die ‚Elemente‘ von M genannt werden) zu einem Ganzen.“

Cantor genügte seine Beschreibung zur Entwicklung einer Theorie, die sehr komplizierte unendliche Objekte untersuchen konnte.

Beispiele

- (1) \mathbb{N} = „die Menge aller natürlichen Zahlen (einschließlich der Null)“
- \mathbb{N}^* = „die Menge aller natürlichen Zahlen größer als Null“
- \mathbb{Z} = „die Menge aller ganzen Zahlen“
- \mathbb{Q} = „die Menge aller rationalen Zahlen“
- \mathbb{R} = „die Menge aller reellen Zahlen“
- (2) $[a, b]$ = „die Menge aller reellen Zahlen x mit $a \leq x$ und $x \leq b$ “
- $[a, b[$ = „die Menge aller reellen Zahlen x mit $a \leq x$ und $x < b$ “
- $]a, b]$ = „die Menge aller reellen Zahlen x mit $a < x$ und $x \leq b$ “
- $[a, +\infty[$ = „die Menge aller reellen Zahlen x mit $a \leq x$ “
- $] -\infty, b[$ = „die Menge aller reellen Zahlen x mit $x < b$ “, für reelle Zahlen $a \leq b$
- (3) \mathcal{I} = „die Menge aller Intervalle $]a, b[$ mit rationalen Endpunkten a, b “

Das letzte Beispiel zeigt, dass Mengen wieder Mengen als Elemente enthalten können. Ist m ein Element der Menge M , so schreiben wir $m \in M$. Ist m kein Element von M ,

so schreiben wir $m \in M$. Das Zeichen „ \in “ ist ein stilisiertes griechisches Epsilon „ ϵ “. Wir lesen „ $m \in M$ “ als

„ m ist ein Element von M “ oder „die Menge M besitzt das Element m “.

Beispiele

0 \in \mathbb{N} , $-1 \notin \mathbb{N}$, $0 \notin \mathbb{N}^*$, $2 \in \mathbb{Q}$, $-1/2 \in \mathbb{Q}$, $\sqrt{2} \in \mathbb{R}$.

Wie man sich Mengen und die Elementbeziehung vorstellt, bleibt jedem selbst überlassen – ob als Säcke, Punktwolken, Pfeildiagramme (vgl. 1.3), umzäunte Bereiche usw.

Prinzipiell können wir beliebige Zeichen links und rechts der Elementrelation verwenden. Oft bezeichnet man Grundobjekte wie Zahlen oder Punkte mit a, b, c, \dots und Mengen, die aus diesen Objekten gebildet werden, mit A, B, C, \dots . Diese Mengen können wir wieder zu Mengen zusammenfassen. Die Menge aus Beispiel (3) ist z. B. eine in der Analysis häufig betrachtete Menge von Mengen. Für derartige Mengensysteme bieten sich Zeichen wie $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$ an, man darf hier aber auch A, B, C, \dots schreiben.

Die wichtigste Beziehung zwischen Mengen ist die *Teilmengenbeziehung* oder *Inklusion*:

Definition (Teilmenge, Obermenge)

Seien M und N Mengen. Dann heißt N

- (a) *Teilmenge* von M , in Zeichen $N \subseteq M$, falls für alle $x \in N$ auch $x \in M$ gilt,
- (b) *echte Teilmenge* von M , in Zeichen $N \subset M$, falls $N \subseteq M$ und $N \neq M$,
- (c) *Obermenge* von M , in Zeichen $N \supseteq M$, falls $M \subseteq N$,
- (d) *echte Obermenge* von M , in Zeichen $N \supset M$, falls $M \subset N$.

Die Zeichen sind so gewählt, dass sie an $\leq, <, \geq$ und $>$ erinnern. Jedoch schreiben viele Mathematiker „ \subset “ für „Teilmenge“ und „ \subsetneq “ für „echte Teilmenge“ und begründen dies halb scherzhaft damit, dass die echte Inklusion selten vorkommt und man sich den Strich unter \subset sparen kann. Wir verwenden durchweg die Bezeichnungen der Definition.

Beispiele

- (1) Seien $a \leq b$ und $c \leq d$ reelle Zahlen. Dann gilt $[a, b] \subseteq [c, d]$ genau dann, wenn $a \geq c$ und $b \leq d$.
- (2) $\mathbb{N}^* \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

Häufig verwendet wird das folgende Prinzip, das in einer axiomatischen Formulierung der Mengenlehre ein Axiom ist:

Gilt $M \subseteq N$ und $N \subseteq M$, so gilt $N = M$. (*Extensionalitätsprinzip*)

Will man zeigen, dass zwei Mengen M und N gleich sind, so kann man also in zwei Schritten $M \subseteq N$ und $N \subseteq M$ zeigen. Dies ist oft einfacher als ein Nachweis von $M = N$ in einem Schritt.

1.2 Umgang mit Mengen

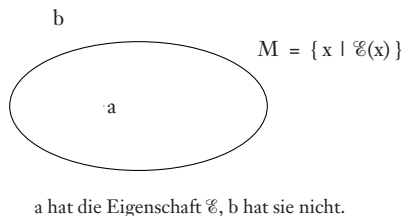
Definition (Mengenbildung über Eigenschaften, Mengenkompensation)

Ist M eine Menge und $\mathcal{E}(x)$ eine mathematische Eigenschaft, so schreiben wir

$$M = \{ x \mid \mathcal{E}(x) \},$$

falls für alle Objekte x gilt: $x \in M$ genau dann, wenn $\mathcal{E}(x)$. Wir lesen dann den Ausdruck $\{ x \mid \mathcal{E}(x) \}$ als „Menge aller x mit der Eigenschaft $\mathcal{E}(x)$ “.

Ist $M = \{ x \mid \mathcal{E}(x) \}$, so ist „ $x \in M$ “ nur eine andere Schreibweise für $\mathcal{E}(x)$. Was genau eine „mathematische Eigenschaft“ ist, braucht uns hier nicht zu kümmern. Klar ist, dass „ x ist eine rationale Zahl“ mathematisch und „ x ist eine Banane“ nicht mathematisch ist. Cantors Beschreibung lässt auch andere Eigenschaften zu, wir beschränken uns auf die Mathematik.



Die *Russell-Antinomie* „ $R = \{ M \mid M \notin M \}$ “ der Menge aller Mengen, die sich nicht selbst als Element enthalten, zeigt, dass nicht zu jeder Eigenschaft eine Menge gebildet werden kann. Für alle M würde $M \in R$ genau dann gelten, wenn $M \notin M$. Einsetzen von $M = R$ ergibt den Widerspruch: $R \in R$ genau dann, wenn $R \notin R$. Diesem bemerkenswerten und historisch äußerst bedeutsamen Widerspruch steht folgende (letztendlich axiomatisch begründete) Aussage gegenüber:

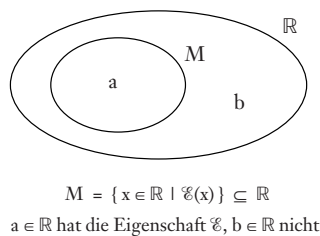
Ist $\mathcal{E}(x)$ eine in der Analysis betrachtete Eigenschaft, so existiert die Menge $M = \{ x \mid \mathcal{E}(x) \}$.

Solange man keine „wilden“ Konstruktionen durchführt, darf man die Kompensation $M = \{ x \mid \mathcal{E}(x) \}$ zur Definition einer Menge M bedenkenlos durchführen. Speziell gilt dies für *Aussonderungen* der Form

$$M = \{ x \in \mathbb{R} \mid \mathcal{E}(x) \} = \{ x \mid x \in \mathbb{R} \text{ und } \mathcal{E}(x) \}.$$

Wir reduzieren hier \mathbb{R} auf diejenigen x , für die $\mathcal{E}(x)$ gilt. Es entsteht eine Teilmenge von \mathbb{R} .

Die Kompensation erlaubt uns, Mengen und Operationen mit Mengen elegant und ordentlich einzuführen. Einige Beispiele sind:



Notation	Definition durch Kompensation	Name
\emptyset oder $\{ \}$	$\{ x \mid x \neq x \}$	leere Menge
$\{ a_1, \dots, a_n \}$	$\{ x \mid x = a_1 \text{ oder } \dots \text{ oder } x = a_n \}$	endliche Auflistung
$A \cup B$	$\{ x \mid x \in A \text{ oder } x \in B \}$	Vereinigung
$A \cap B$	$\{ x \mid x \in A \text{ und } x \in B \}$	Durchschnitt
$A - B$ oder $A \setminus B$	$\{ x \mid x \in A \text{ und } x \notin B \}$	Differenz

Unendliche Mengen werden oft mit Hilfe der „Pünktchenschreibweise“ definiert, etwa in $M = \{0, 2, 4, 6, \dots\}$. Diese Form ist vage, aber oft ist allen doch klar, was definiert wird (hier $M = \{n \in \mathbb{N} \mid n \text{ ist gerade}\}$), sodass wir diese gut lesbare Form nicht verbieten wollen.

In der Definition von $\{a_1, \dots, a_n\}$ steckt, dass es bei der endlichen Auflistung der Elemente weder auf die Reihenfolge der Elemente noch auf Wiederholungen ankommt:

Beispiel

■ Für alle a, b gilt: $\{a, b\} = \{b, a\} = \{a, a, b\} = \{a, b, a, b\} = \dots$

Sind etwa f, g und h Funktionen auf \mathbb{R} , so können wir diese Funktionen im Nullpunkt auswerten und die Menge $\{f(0), g(0), h(0)\}$ bilden. Diese Menge kann zum Beispiel von der Form $\{a, b, a\}$ sein, muss also nicht genau drei verschiedene Elemente besitzen. Allgemein hat die Menge $\{a, b, c\}$ mindestens ein und höchstens drei verschiedene Elemente.

Anfänger haben oft Schwierigkeiten mit der leeren Menge \emptyset . Diese Menge enthält keine Elemente, ist aber ein vollwertiges mathematisches Objekt und nicht etwa nichts. Zu unterscheiden ist sie von der Einermenge $\{0\}$, die nur die Null als Element enthält.

Beispiel

■ $\{\} = \{x \in \mathbb{R} \mid x^2 + 1 = 0\}$, $\{0\} = \{x \in \mathbb{R} \mid x^2 = 0\}$.

Reelle Polynome haben nur endliche viele Nullstellen, sodass die Mengen der Form $\{a_1, \dots, a_n\}$ hier eine wichtige Rolle spielen. Wollen wir, dass wir zu jedem Polynom dessen Nullstellenmenge bilden können, müssen wir $\{\}$ und $\{0\}$ zulassen. Dagegen sprechen letztendlich nur psychologische Argumente. In der Umgangssprache und auch etymologisch ist „Menge“ mit „viel“ und „ein Haufen“ belegt, etwa in: „Du hast ja eine Menge Geld“. Das sagt man nur im Spaß, wenn der Angesprochene einen leeren Geldbeutel hat, mathematisch wäre es korrekt. In den Anfangszeiten der Mengenlehre hat man die leere Menge auch gesondert betrachtet und manchmal als „uneigentliche Menge“ bezeichnet. Es braucht, wie bei der Null als Zahl, etwas Zeit, sich daran zu gewöhnen.

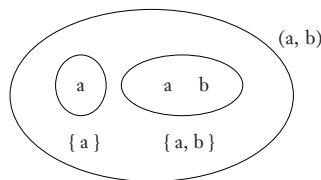
Neben ungeordneten Paarmengen $\{a, b\}$ werden auch *geordnete Paare* (a, b) gebraucht, etwa zur Angabe von Punkten der Ebene. Sie besitzen die folgende Koordinateneigenschaft:

Für alle a, b, c, d gilt $(a, b) = (c, d)$ genau dann, wenn $a = c$ und $b = d$.

Speziell ist also $(0, 1) \neq (1, 0)$, im Gegensatz zu $\{0, 1\} = \{1, 0\}$. Man kann geordnete Paare als Grundbegriff ansehen, aber wir können sie auch (trickreich!) als Mengen definieren, so dass die Koordinateneigenschaft gilt:

$$(a, b) = \{\{a\}, \{a, b\}\}.$$

(geordnetes Paar, Kuratowski-Paar)



Mit geordneten Paaren können wir nun das Kreuzprodukt zweier Mengen einführen:

$$A \times B = \{(a, b) \mid a \in A \text{ und } b \in B\}. \quad (\text{kartesisches Produkt, Kreuzprodukt})$$

Geordnete Tripel (a, b, c) können wir durch $(a, b, c) = ((a, b), c)$ definieren. Wir setzen dann $A \times B \times C = \{((a, b), c) \mid a \in A, b \in B, c \in C\}$. Analog ist $(a, b, c, d) = ((a, b, c), d)$ usw. Statt $A \times A$ schreiben wir auch A^2 , statt $A \times A \times A$ auch A^3 usw.

1.3 Relationen

Definition (Relation)

- (a) Eine Menge R heißt eine (zweistellige) *Relation*, falls jedes Element von R ein geordnetes Paar ist. Gilt $(x, y) \in R$, so schreiben wir auch $x R y$.
- (b) Gilt $R \subseteq M \times M$, so nennen wir R eine Relation *auf* der Menge M .

Relationen tauchen in der Mathematik in so vielen Formen auf, dass eine sehr allgemeine Definition angemessen ist. Obige Definition versucht gar nicht, die möglichen Beziehungen wie „größer“, „kleiner“, „äquivalent“, „umfassender“ oder „verwandt“, die zwischen zwei Objekten bestehen können, in ihrem Wesen zu beschreiben, sondern sie fängt alle möglichen Beziehungen durch eine einfache Struktureigenschaft ein. Dass x in einer bestimmten Relation R zu y steht, wird einfach durch $(x, y) \in R$ ausgedrückt.

Statt R wird oft ein Symbol wie \leq , $<$, \equiv , \sim verwendet. Dann wird die zu „ $(x, y) \in R$ “ gleichwertige Notation „ $x R y$ “ zur vertrauten Schreibweise

$$x \leq y, \quad x < y, \quad x \equiv y, \quad x \sim y.$$

Neu und oft schwierig für Anfänger ist, dass eine Relation als eine Menge aufgefasst wird, die beliebig bezeichnet werden kann.

Beispiele

- (1) Für alle $a, b \in \mathbb{Z}$ setzen wir:

$a \mid b$, falls a ein Teiler von b ist, d.h. falls ein $d \in \mathbb{Z}$ existiert mit $a d = b$.

Dann gilt z.B. $4 \mid 4$, $4 \mid -8$, $10 \mid 100$, $-3 \mid 9$, $0 \mid 0$, $0 \mid 4$.

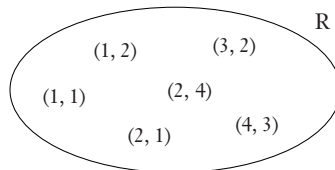
- (2) Sei $m \in \mathbb{N}^*$. Dann setzen wir für alle $a, b \in \mathbb{Z}$:

$a \equiv_m b$, falls $m \mid (a - b)$. (Kongruenz modulo m)

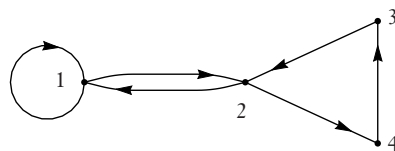
Es gilt $a \equiv_m b$, falls a und b bei Division durch m den gleichen Rest haben.

So ist etwa $2 \equiv_7 9 \equiv_7 -5$ und $\{a \in \mathbb{Z} \mid a \equiv_7 5\} = \{\dots, -9, -2, 5, 12, 19, \dots\}$.

Eine Relation R lässt sich manchmal durch Pfeildiagramme veranschaulichen. Wir verbinden alle x, y mit $x R y$ durch einen Pfeil von x nach y . Gilt $x R x$, so führt ein Pfeil von x zu sich selbst. Gilt $x R y$ und $y R x$, so sind x und y durch zwei Pfeile verbunden. Man kann dies graphisch vereinfachen, indem man x und y durch eine Linie verbindet. Die „ x - y -Straße“ ist dann in beiden Richtungen befahrbar.

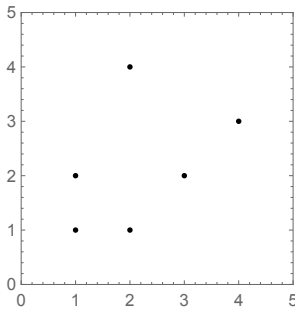


Eine Relation R auf $\{1, 2, 3, 4\}$. Es gilt
 $1 R 1, 1 R 2, 2 R 1, 2 R 4, 3 R 2, 4 R 3$.

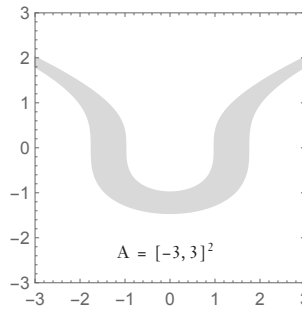


Die obige Relation R in Pfeildarstellung.
 In der diskreten Mathematik spricht man
 von einem *gerichteten Graphen*.

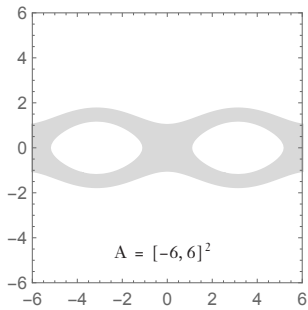
Für eine Relation R auf \mathbb{R} bietet sich eine andere Visualisierung an: Wir markieren alle Punkte (x, y) der Ebene, für die $x R y$ gilt. R erscheint dann als eine Punktwolke der Ebene.



noch einmal obige Relation R



$R = \{(x, y) \in A \mid 1 \leq x^2 - y^3 \leq 3\}$



$R = \{(x, y) \in A \mid 1/2 \leq \cos(x) + y^2 \leq 2\}$

Wir stellen Eigenschaften von Relationen tabellarisch zusammen.

Eine Relation R auf M heißt ...	falls ...
<i>reflexiv</i>	für alle $x \in M$ gilt $x R x$
<i>irreflexiv</i>	für kein $x \in M$ gilt $x R x$
<i>symmetrisch</i>	für alle $x, y \in M$ gilt: $x R y$ impliziert $y R x$
<i>antisymmetrisch</i>	für alle $x, y \in M$ gilt: $x R y$ und $y R x$ impliziert $x = y$
<i>transitiv</i>	für alle $x, y, z \in M$ gilt: $x R y$ und $y R z$ impliziert $x R z$

„Reflexiv“ bedeutet für Pfeildiagramme, dass von jedem x ein Pfeil zu x führt, sowie für Punktwolken, dass die Diagonale zur Punktwolke gehört. „Symmetrisch“ bedeutet für Pfeildiagramme, dass Pfeile zwischen zwei Punkten stets in beiden Richtungen verlaufen (und also durch Linien ersetzt werden können), sowie für Punktwolken, dass die Spiegelung an der Diagonalen $\{(x, x) \mid x \in \mathbb{R}\}$ die Punktwolke unverändert lässt.

Beispiele

- (1) Die Teilbarkeitsrelation \mid auf \mathbb{Z} ist reflexiv und transitiv. Sie ist nicht antisymmetrisch, da z. B. $-3 \mid 3$ und $3 \mid -3$, aber $3 \neq -3$ gilt.
- (2) Für jedes $m \in \mathbb{N}^*$ ist die Kongruenz \equiv_m modulo m reflexiv, symmetrisch und transitiv. (Relationen mit diesen drei Eigenschaften heißen *Äquivalenzrelationen*.)
- (3) Die Relationen \leq und \geq auf \mathbb{R} sind reflexiv, antisymmetrisch und transitiv, die Relationen $<$ und $>$ auf \mathbb{R} sind irreflexiv und transitiv. Die Gleichheit $=$ auf \mathbb{R} ist reflexiv, symmetrisch und transitiv.

Für alle reellen Zahlen $x \geq 0$ gilt zum Beispiel:

$$7/8 x \leq 2x < 1 + 2x \leq 1 + 2x + x^2 = (1+x)^2.$$

Die Abschätzungskette zeigt, dass $7/8 x < (1+x)^2$ für alle $x \geq 0$ gilt.

1.4 Funktionen

Definition (Funktion, Abbildung, Definitionsbereich, Wertebereich)

Eine Relation f heißt eine *Funktion* oder *Abbildung*, falls für alle $(x, y) \in f$ gilt:

Es gibt kein $y' \neq y$ mit $(x, y') \in f$. (Eindeutigkeitsbedingung)

Gilt $(x, y) \in f$, so schreiben wir $f(x) = y$. Wir nennen dann y den *Wert* von f an der *Stelle* x oder für das *Argument* x und x ein *Urbild* von y unter f . Wir setzen:

$\text{Def}(f) = \{ x \mid \text{es gibt ein } y \text{ mit } (x, y) \in f \}$, (Definitionsbereich von f)

$\text{Bild}(f) = \{ y \mid \text{es gibt ein } x \text{ mit } (x, y) \in f \}$. (Wertebereich oder Bild von f)

Wir sagen auch, dass f eine Funktion *auf* der Menge $\text{Def}(f)$ ist.

Der Funktionsbegriff ist abstrakt, allgemein und präzise. Wir müssen nicht erklären, was eine „eindeutige Zuordnung“ ist. Wir müssen die gefühlte Dynamik einer Funktion nicht einfangen. Eine Funktion ist eine „zweispaltige Tabelle“, in der links die Stellen und rechts die Funktionswerte stehen. Entscheidend ist die Eindeutigkeitsbedingung: Wir fordern, dass die Tabelle keine Einträge der Form (x, y) und (x, y') mit $y \neq y'$ enthält. Einträge der Form (x, y) und (x', y) mit $x \neq x'$ sind dagegen erlaubt, sie entsprechen der Gleichheit $f(x) = f(x')$ der Funktionswerte an zwei verschiedenen Stellen x und x' . Alle linken Einträge zusammengenommen ergeben den Definitionsbereich, alle rechten den Wertebereich der Funktion.

Stelle	Funktionswert
1	2
2	1
3	5
4	6
5	1

Eine Funktion f als Tabelle. Es gilt
 $\text{Def}(f) = \{ 1, \dots, 5 \}$, $\text{Bild}(f) = \{ 1, 2, 5, 6 \}$,
 $f(1) = 2$, $f(2) = 1$, $f(3) = 5$, $f(4) = 6$, $f(5) = 1$.

In Darstellungen, die auf eine Präzisierung des Funktionsbegriffs verzichten und diesen nur anschaulich erklären, heißt die Menge $\{ (x, y) \mid f(x) = y \}$ der *Graph* von f . Unsere Definition kann dann sehr griffig zum Ausdruck gebracht werden:

Wir identifizieren eine Funktion mit ihrem Graphen.

Obwohl Funktion und Graph für uns zusammenfallen, sprechen wir oft vom Graphen einer Funktion, wenn wir Visualisierungen im Auge haben.

Beispiele

- (1) $f = \{ (0, 1), (1, 1) \}$ ist die Funktion auf $\{ 0, 1 \}$ mit $f(0) = f(1) = 1$. Dagegen ist $R = \{ (1, 0), (1, 1) \}$ eine Relation mit $1 R 0$ und $1 R 1$, aber keine Funktion.
- (2) Setzen wir $f(n) = \text{„das größte } k \text{ mit } 2^k \mid n\text{“}$ für alle $n \in \mathbb{N}^*$, so erhalten wir eine Funktion auf \mathbb{N}^* . Es gilt z. B. $f(1) = f(3) = 0$, $f(2) = f(6) = 1$, $f(16) = f(48) = 4$.
- (3) Setzen wir $f(x) = x^2$ für alle $x \geq 0$, so erhalten wir die zweite Potenz f auf $[0, +\infty[$. Setzen wir $g(x) = x^2$ für alle $x \in \mathbb{R}$, so erhalten wir die zweite Potenz g auf \mathbb{R} .

Wir hatten eine Relation R eine Relation *auf* M genannt, falls $R \subseteq M \times M$. Die Sprechweise „ f ist eine Funktion auf M “ weicht hiervon ab. Bei Funktionen bedeutet „auf“ den Definitionsbereich. Als Relation ist eine Funktion f eine Relation auf $\text{Def}(f) \cup \text{Bild}(f)$.

Im Umgang mit Funktionen dominiert die vertraute Schreibweise

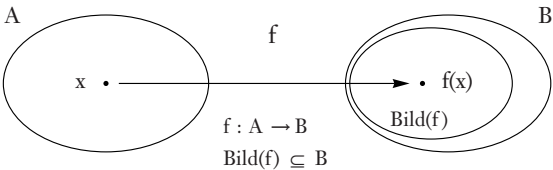
$$f(x) = y, \quad \text{gelesen: „} f \text{ von } x \text{ gleich } y\text{“,}$$

die Notation $(x, y) \in f$ tritt, wie $(x, y) \in R$ bei Relationen, zurück. Intuitiv können wir $f(x) = y$ so lesen, dass die Funktion f das Objekt x auf das Objekt y abbildet, x in y umformt oder in einem abstrakten Sinne umrechnet, dem x das y eindeutig zuordnet, etwa durch einen Pfeil usw. Wichtig ist, dass eine Funktion nicht durch einen Term definiert sein muss und dass wir im Allgemeinen ein $f(x)$ nicht einfach berechnen können.

Eine der wichtigsten Notationen für Funktionen verdient eine eigene Definition:

Definition ($f: A \rightarrow B$, Funktion von A nach B , Wertevorrat)

Sei f eine Funktion. Wir schreiben $f: A \rightarrow B$, falls $A = \text{Def}(f)$ und $\text{Bild}(f) \subseteq B$. Wir sagen dann, dass f eine Funktion *von* A *nach* B ist. B heißt ein *Wertevorrat* von f .



Gilt $f: A \rightarrow B$, so ist A der Definitionsbereich von f . Dagegen verlangen wir nicht, dass die Menge B der Wertebereich von f ist, sondern dass B den Wertebereich von f umfasst. So gilt für den Sinus auf den reellen Zahlen zum Beispiel

$$\sin: \mathbb{R} \rightarrow \mathbb{R} \text{ und } \sin: \mathbb{R} \rightarrow [-1, 1], \text{ nicht aber } \sin: \mathbb{R} \rightarrow [0, 1].$$

In $f: A \rightarrow B$ ist die Menge B der „ins Auge gefasste“ oder „prinzipiell zugelassene“ Wertevorrat von f . Oft kann eine Funktion $f: A \rightarrow B$ einfach definiert werden, die Bestimmung von $\text{Bild}(f)$ dagegen sehr schwierig sein.

In der Analysis werden Funktionen $f: A \rightarrow \mathbb{R}$ untersucht, deren Definitionsbereich A eine Teilmenge der reellen Zahlen \mathbb{R} ist, sog. *reelle Funktionen*. Typische Definitionsbereiche sind $[0, 1]$, $]0, +\infty[$, \mathbb{R} oder eine Menge wie $\mathbb{R}^* = \mathbb{R} - \{0\}$, die beispielsweise bei der Funktion f mit $f(x) = 1/x$ für alle $x \neq 0$ auftritt. Wir definieren für reelle Funktionen:

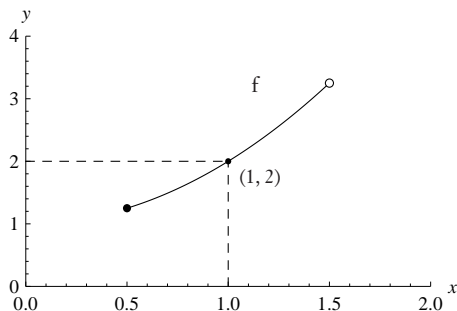
Ein $f: A \rightarrow \mathbb{R}$ heißt ...	falls für alle $x < y$ in A gilt:
<i>monoton steigend</i>	$f(x) \leq f(y)$
<i>streng monoton steigend</i>	$f(x) < f(y)$
<i>monoton fallend</i>	$f(x) \geq f(y)$
<i>streng monoton fallend</i>	$f(x) > f(y)$

Weiter heißt f (*streng*) *monoton*, falls f (*streng*) monoton fallend oder wachsend ist. Schließlich nennen wir f *beschränkt*, falls ein y existiert mit $-y \leq f(x) \leq y$ für alle $x \in A$.

1.5 Visualisierung von Funktionen

„Graphische Methode“

Sei $f: A \rightarrow B$ eine Funktion mit $A, B \subseteq \mathbb{R}$. Dann besteht die „graphische Methode“ in der Darstellung der Funktion, die entsteht, wenn wir jeden Punkt (x, y) der Ebene markieren, für den $f(x) = y$ gilt.



Graphische Darstellung der Funktion

$f: [1/2, 3/2[\rightarrow \mathbb{R}, f(x) = x^2 + 1$ für alle x .

Der schwarze (bzw. weiße) Punkt deutet an, dass $1/2$ (bzw. $3/2$) zum Definitionsbereich von f gehört (bzw. nicht gehört). Die gestrichelten Linien sind Hilfslinien, die deutlich machen, dass $f(1) = 2$ gilt.

Im Folgenden verzichten wir auf die hier noch angegebene Benennung der Achsen.

Das ist natürlich keine mathematische Definition im üblichen Sinne, sondern eher ein didaktisches Programm. Die graphische Methode wurde vor allem durch den an Fragen der Lehrerbildung interessierten Mathematiker Felix Klein zu Beginn des 20. Jahrhunderts propagiert. Sie wird heute in der Schule häufig verwendet und wir sind, wenn wir ein Universitätsstudium beginnen, mit den Graphen der Quadratwurzel, des Logarithmus und der trigonometrischen Funktionen gut vertraut. Mit Hilfe von Funktionsgraphen können wir viele wichtige analytische Eigenschaften erkennen und diskutieren:

Wachstumsverhalten, Nullstellen, Extremwerte, Krümmung, Wendepunkte, Oszillation, Sprungverhalten, eingeschlossene Fläche usw.

Eines der Ziele der Analysis ist es, diese anschaulichen Eigenschaften theoretisch und rechnerisch zu beherrschen. Funktionsgraphen haben oft einen hohen heuristischen Wert, aber sie besitzen keine Beweiskraft. Umgekehrt werden theoretisch fundierte Methoden eingesetzt, um die Graphen zu zeichnen.

Die graphische Methode fällt mit der Visualisierung von Relationen durch Punktwolken der Ebene zusammen. Eine reelle Funktion ist als relationale Punktwolke dadurch ausgezeichnet, dass sich über jedem Punkt der x -Achse höchstens ein markierter Punkt befindet. Dies ist gerade die Eindeigkeitseigenschaft einer Funktion. Der Definitionsbereich von f ist die Menge der Punkte der x -Achse, über denen sich ein markierter Punkt befindet. Der Wertebereich von f ist dagegen die Menge der y -Komponenten der markierten Punkte.

Nicht alle Funktionen lassen sich gut zeichnen. So lässt sich zum Beispiel die *Dirichlet'sche Sprungfunktion* $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = 0$, falls x rational, und $f(x) = 1$, falls x irrational, graphisch nur andeuten. Ähnliches gilt für stetige Funktionen auf \mathbb{R} , die in keinem Punkt differenzierbar sind. Auf diese Funktionen werden wir noch zurückkommen.

In den folgenden Diagrammen zeigen wir exemplarisch, wie sich eine rechnerische Manipulation einer Funktion anschaulich übersetzt. Es gelten die Entsprechungen:

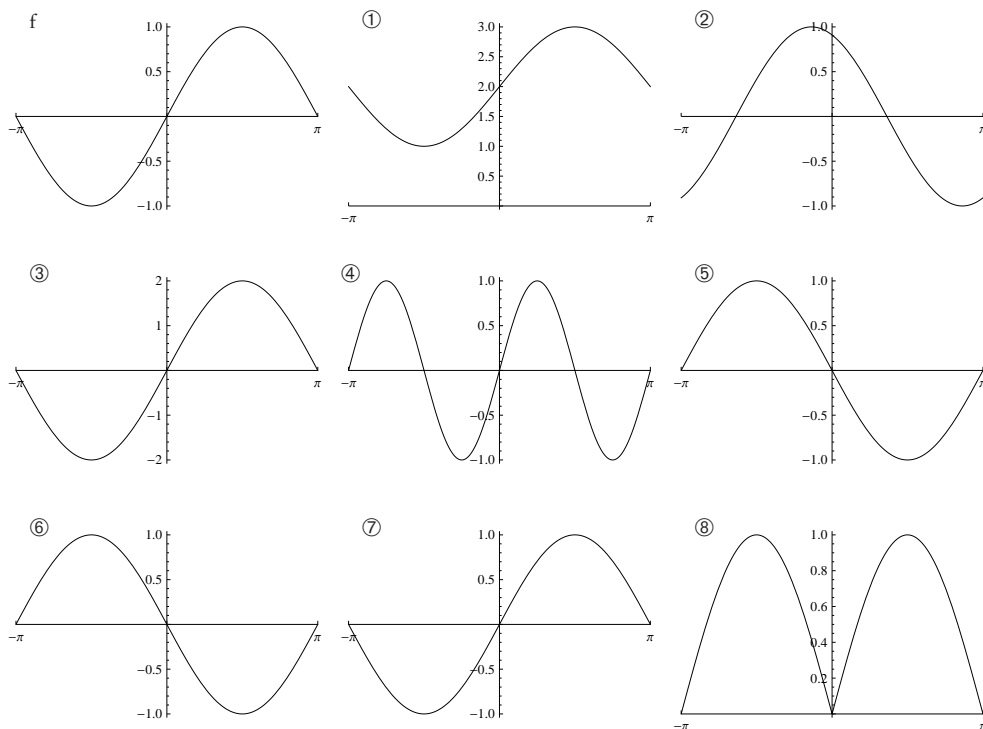
	<i>Transformation</i>	<i>geometrische Bedeutung</i>
①	$g(x) = f(x) + c$	Parallelverschiebung um c entlang der y-Achse
②	$g(x) = f(x + c)$	Parallelverschiebung um $-c$ entlang der x-Achse
③	$g(x) = c f(x)$	Streckung entlang der y-Achse um den Faktor c
④	$g(x) = f(cx)$	Stauchung entlang der x-Achse um den Faktor c
⑤	$g(x) = -f(x)$	Spiegelung an der x-Achse
⑥	$g(x) = f(-x)$	Spiegelung an der y-Achse
⑦	$g(x) = -f(-x)$	Punktspiegelung am Nullpunkt
⑧	$g(x) = f(x) $	Hochklappen des negativen Teils

Merkregel

Die Wirkung wird *invertiert*, wenn die Operation *in* der Funktion angewendet wird.

Die punktweise ausgeführte Operation $f(2x)$ liefert zum Beispiel eine Stauchung um den Faktor 2, die Operation $f(x + 1)$ eine Verschiebung um 1 nach links.

Die folgenden Diagramme zeigen die Wirkung der acht Transformationen der Tabelle für die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = \sin(x)$ für alle x , und die Konstante $c = 2$. Dargestellt sind die Funktionen jeweils im Intervall $[-\pi, \pi]$.

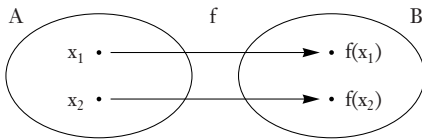


1.6 Injektive, surjektive und bijektive Funktionen

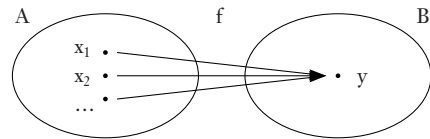
Definition (*injektiv, surjektiv, bijektiv*)

Sei $f : A \rightarrow B$. Dann heißt f

- (a) *injektiv*, falls für alle $x_1, x_2 \in A$ mit $x_1 \neq x_2$ gilt, dass $f(x_1) \neq f(x_2)$,
- (b) *surjektiv* (*auf B*), falls $\text{Bild}(f) = B$,
- (c) *bijektiv*, falls f injektiv und surjektiv auf B ist.



injektiv: Verschiedene Stellen haben verschiedene Bilder.

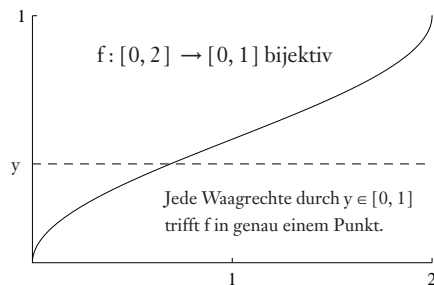


surjektiv: Jedes $y \in B$ hat mindestens ein Urbild.

Die Injektivität einer Funktion ist eine Eigenschaft von f . Für die beiden anderen Eigenschaften ist die Spezifikation eines Wertevorrats, also die Angabe von f in der Form $f : A \rightarrow B$ notwendig. Neben der Bezeichnung „injektiv“ ist im Deutschen auch „eindeutig“ gebräuchlich. Weiter spricht man auch substantivisch von Injektionen, Surjektionen und Bijektionen.

Die drei Begriffe haben eine sehr anschauliche Bedeutung. Injektivität bedeutet, dass kein Wert zweimal angenommen wird, Surjektivität bedeutet, dass jeder Wert des ins Auge gefassten Wertevorrats angenommen wird, und Bijektivität bedeutet, dass die Funktion f die Elemente von A und B vollständig paart, d. h., jedem Element von A wird genau ein Element von B zugeordnet und umgekehrt.

Für reelle Funktionen $f : A \rightarrow B$ existiert eine weitere anschauliche Interpretation: f ist injektiv, wenn jede Waagrechte der Ebene den Funktionsgraphen in höchstens einem Punkt schneidet. Und f ist surjektiv auf B , wenn jede Waagrechte der Ebene, die die Menge B der y -Achse trifft, den Funktionsgraphen in mindestens einem Punkt schneidet. Damit ist also eine Funktion $f : A \rightarrow B$ genau dann bijektiv, wenn jede Waagrechte durch B genau einen Punkt mit dem Graphen von f gemeinsam hat.



Beispiele

- (1) $\cos : \mathbb{R} \rightarrow \mathbb{R}$ ist weder injektiv noch surjektiv, $\cos : \mathbb{R} \rightarrow [-1, 1]$ ist surjektiv.
- (2) $f : \mathbb{R} \rightarrow [0, +\infty[$ mit $f(x) = x^2$ für alle $x \in \mathbb{R}$ ist surjektiv.
- (3) $g : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^3$ für alle $x \in \mathbb{R}$ ist bijektiv.
- (4) $\tan :]-\pi/2, \pi/2[\rightarrow \mathbb{R}$ und $\log :]0, +\infty[\rightarrow \mathbb{R}$ sind Bijektionen.

Sind M und N Mengen und existiert ein injektives $f: M \rightarrow N$, so hat die Menge M anschaulich höchstens so viele Elemente wie die Menge N , denn jedem $x \in M$ entspricht ein eindeutiges $y \in N$, nämlich $y = f(x)$. Man definiert deswegen

$$|M| \leq |N|, \text{ falls ein injektives } f: M \rightarrow N \text{ existiert, } (\text{Mächtigkeitsvergleich})$$

und sagt, dass die *Mächtigkeit* oder *Kardinalität* von M *kleinergleich* der *Mächtigkeit* oder *Kardinalität* von N ist. Analog definiert man

$$|M| = |N|, \text{ falls ein bijektives } f: M \rightarrow N \text{ existiert, } (\text{Gleichmächtigkeit})$$

und sagt, dass M und N *gleichmächtig* sind oder die *gleiche Kardinalität* haben.

Beispiel

■ Nach Beispiel (4) gilt $]-\pi/2, \pi/2[\mid = \mid \mathbb{R} \mid$ und $]0, +\infty[\mid = \mid \mathbb{R} \mid$.

Die Begriffsbildungen liefern eine Möglichkeit zur Definition der Endlichkeit und Unendlichkeit. Dabei genügt es, *einen* Begriff zu definieren, der andere ergibt sich dann durch Verneinung: „unendlich“ = „nicht endlich“ bzw. „endlich“ = „nicht unendlich“. Die folgende Tabelle stellt drei verschiedene äquivalente Möglichkeiten vor, dies zu tun.

M heißt ...	<i>falls ...</i>	<i>falls ...</i>	<i>falls ...</i>
<i>endlich</i>	es gibt ein $n \in \mathbb{N}$ mit $ M = \{0, \dots, n-1\} $		
<i>unendlich</i>		$ \mathbb{N} \leq M $, d.h., es gibt ein injektives $f: \mathbb{N} \rightarrow M$	es gibt eine echte Teilmenge N von M mit $ N = M $

Die erste Definition besagt, dass eine Menge M endlich ist, wenn wir sie mit natürlichen Zahlen $0, \dots, n-1$ durchzählen können ($n=0$ entspricht dem Fall $M=\emptyset$). Die zweite Definition nennt eine Menge M unendlich, wenn wir in ihr via

$$f(0), f(1), f(2), \dots, f(n), \dots$$

zählen können, d. h., M enthält eine „Kopie“ von \mathbb{N} . Bei der dritten Definition schließlich werden die unendlichen Mengen dadurch definiert, dass sie das klassische Prinzip

$$\text{„das Ganze ist größer als der Teil“}$$

verletzen. Diese auf Richard Dedekind zurückgehende Definition setzt als einzige die natürlichen Zahlen nicht voraus. Ihr genügt der Mengen- und Funktionsbegriff.

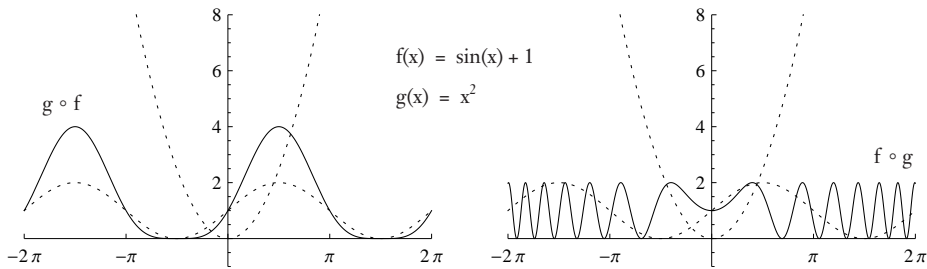
Nun wird der Leser vielleicht fragen, ob hier nicht viel Aufwand um recht wenig getrieben wird: Endlichkeit ist anschaulich klar und die nicht endlichen Mengen sind eben die unendlichen Mengen. Neben einer genauen mathematischen Definition der Unendlichkeit, die ja sicher philosophisch interessant ist, zeigt sich aber, dass im Reich des Unendlichen Größenunterschiede existieren, die durch den Mächtigkeitsvergleich mit Bijektionen ans Licht gebracht werden. Dieses spannende und für die Analysis bedeutsame Phänomen werden wir im zweiten Kapitel kennenlernen.

1.7 Umgang mit Funktionen

Definition (Verknüpfung, Komposition)

Seien $f : A \rightarrow B$ und $g : B \rightarrow C$ Funktionen. Dann ist die *Verknüpfung* oder *Komposition* $g \circ f : A \rightarrow C$ definiert durch

$$(g \circ f)(x) = g(f(x)) \quad \text{für alle } x \in A.$$



Die Verknüpfung $g \circ f$ lesen wir als „g nach f“ oder salopp als „g Kringel f“. Die anschauliche Dynamik ist:

Die Funktion $g \circ f$ schickt ein x aus A mit Hilfe von f zunächst nach B . Anschließend schickt sie $y = f(x)$ mit Hilfe von g nach C . Dann ist $g(y) = g(f(x))$ der Funktionswert von $g \circ f$ an der Stelle x .

Die Funktion f wird also zuerst ausgeführt. Dies erklärt die Sprechweise „g nach f“, die ja nicht der Lesereihenfolge entspricht.

Die Komposition $g \circ f$ kann nur dann definiert werden, wenn der Wertebereich von f eine Teilmenge des Definitionsbereichs der Funktion g ist. Ist g zum Beispiel die Quadratwurzelfunktion auf $[0, +\infty[$, so können wir nicht $g \circ f$ schreiben, wenn die Funktion f negative Werte annimmt, wie etwa die Logarithmusfunktion.

Sind $f : A \rightarrow B$, $g : B \rightarrow C$ und $h : C \rightarrow D$ Funktionen, so können wir die dreifachen Kompositionen $h \circ (g \circ f)$ und $(h \circ g) \circ f$ bilden. Für alle x aus A gilt aber

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

Die Komposition ist also assoziativ. Wir können die unangenehmen Klammern weglassen und einfach $h \circ g \circ f$ schreiben und zu $h(g(f(x)))$ auswerten.

Beispiel

Wir betrachten die Funktion $s : \mathbb{R} - \{1\} \rightarrow \mathbb{R}$ mit

$$s(x) = \frac{(\sin(x^2))^3 - 3}{x - 1} \quad \text{für alle } x \neq 1.$$

In $(\sin(x^2))^3$ ist eine dreifache Komposition der Form $h \circ g \circ f$ am Werk: f ist die zweite Potenz, g der Sinus und h die dritte Potenz. Wollen wir f differenzieren, werden wir unter anderem die Kettenregel heranziehen. Beim Nachdifferenzieren ist es dann wichtig, die Struktur der Funktion klar vor Augen zu haben.

Ein guter Blick für den „Aufbau“ einer Funktion ist generell hilfreich. Hierzu:

Definition (*punktweise Operationen für Funktionen*)

Seien $f, g : A \rightarrow \mathbb{R}$, $A \subseteq \mathbb{R}$, und sei $c \in \mathbb{R}$. Dann definieren wir die Funktionen cf , $f + g$, $f - g$, $f \cdot g$ und, falls $g(x) \neq 0$ für alle $x \in A$ gilt, f/g auf A , durch

$$\begin{aligned}(cf)(x) &= c f(x), & (f + c)(x) &= f(x) + c, \\(f + g)(x) &= f(x) + g(x), & (f - g)(x) &= f(x) - g(x), \\(f \cdot g)(x) &= f(x) \cdot g(x), & (f/g)(x) &= f(x)/g(x) \quad \text{für alle } x \in A.\end{aligned}$$

Weiter setzen wir $f^2 = f \cdot f$, $f^3 = f^2 \cdot f$ usw.

Der Definitionsbereich A wird durch diese Operationen nicht verändert, es gilt zum Beispiel $f + g : A \rightarrow \mathbb{R}$. In graphischen Darstellungen werden die Graphen der Funktionen punktweise skaliert, addiert, multipliziert, usw. (vgl. auch 1.5).

Beispiel

Sei $\text{id} : \mathbb{R} \rightarrow \mathbb{R}$ die Identität auf \mathbb{R} , d.h. es gilt $\text{id}(x) = x$ für alle x . Weiter sei $f : \mathbb{R} \rightarrow \mathbb{R}$ ein reelles Polynom,

$$f(x) = a_k x^k + \dots + a_1 x + a_0 \quad \text{für alle } x \in \mathbb{R},$$

für gewisse Koeffizienten $a_0, \dots, a_k \in \mathbb{R}$ (vgl. auch 6.1). Dann gilt

$$f = a_k \text{id}^k + \dots + a_1 \text{id} + a_0.$$

Alle Polynome lassen sich also aus der Identität durch Anwendung der punktweisen Operationen erzeugen. Diese Beobachtung ist an verschiedenen Stellen nützlich: Weiß man, dass die punktweisen Operationen die Stetigkeit erhalten, so folgt aus der Stetigkeit der Identität die Stetigkeit aller Polynome (vgl. 5.1). Ebenso folgt aus den Ableitungsregeln für die Addition und die Multiplikation eine Ableitungsregel für Polynome (vgl. 7.4).

Wir weisen noch auf eine Fehlerquelle hin, die auf unterschiedliche analytische und algebraische Konventionen zurückzuführen ist. In der Algebra wird die Komposition gerne als Multiplikation aufgefasst und man setzt dann konsequent für ein $f : A \rightarrow A$

$$f^2 = f \circ f, \quad f^3 = f \circ f^2 = f \circ f \circ f \quad \text{usw.}$$

Ist also $f : \mathbb{R} \rightarrow \mathbb{R}$ die vierte Potenz, so ist $f^2 = f \circ f$ die sechzehnte Potenz, denn es gilt

$$f^2(x) = f(f(x)) = f(x^4) = (x^4)^4 = x^{16} \quad \text{für alle } x \in \mathbb{R}.$$

Lesen wir dagegen f^2 als Multiplikation, d.h. $f^2 = f \cdot f$, so ist f^2 nur die achte Potenz, denn $f^2(x) = f(x) \cdot f(x) = x^4 \cdot x^4 = x^8$ für alle $x \in \mathbb{R}$. Die Bedeutung der Notation ist aber in der Regel aus dem Kontext heraus klar. Generell gilt, dass in der Analysis und der mathematischen Physik die punktweisen Operationen im Vordergrund stehen, sodass also \sin^2 nicht die Funktion $\sin \circ \sin$ ist, sondern $\sin^2(x) = (\sin(x))^2$ für alle x gilt. Eine Ausnahme bildet die Umkehrfunktion f^{-1} , die auch in der Analysis nicht die punktweise Invertierung $1/f$ bedeutet (vgl. auch die nächste Sektion).

1.8 Umkehrfunktionen und Einschränkungen

Definition (Umkehrfunktion)

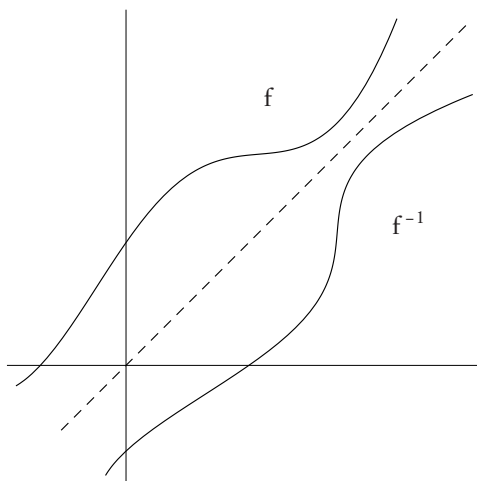
Sei $f: A \rightarrow B$ eine injektive Funktion. Dann setzen wir

$$f^{-1} = \{ (y, x) \mid (x, y) \in f \},$$

d. h., es gilt $f^{-1}(y) = x$ genau dann, wenn $f(x) = y$. Die Funktion f^{-1} heißt die *Umkehrfunktion* von f .

Die Umkehrung einer Funktion f beschreibt das „Rückgängigmachen“ der Bildung von $y = f(x)$: Mit f^{-1} kehren wir von y zurück zu x . In der graphischen Methode entspricht dieses Vorgehen der Spiegelung von f an der Hauptdiagonalen. Die Injektivität der Funktion f sorgt dafür, dass durch die Spiegelung wieder ein Graph entsteht, der die Eindeutigkeitsbedingung für Funktionen erfüllt.

Wichtige Eigenschaften für Umkehrfunktionen sind:



$$f(x) = y \text{ genau dann, wenn } f^{-1}(y) = x$$

$f^{-1} : \text{Bild}(f) \rightarrow \text{Def}(f)$ ist bijektiv.
$f^{-1} \circ f$ ist die Identität auf $\text{Def}(f)$ und $f \circ f^{-1}$ ist die Identität auf $\text{Bild}(f)$.
$(f^{-1})^{-1} = f$, d. h., die Umkehrfunktion der Umkehrfunktion von f ist f .

Die Schreibweise f^{-1} ist durch die Eigenschaft $f^{-1} \circ f = \text{id}$ (auf $\text{Def}(f)$) und die üblichen gruppentheoretischen Notationen für die Inversenbildung motiviert. Für die Analysis ist wichtig, dass f^{-1} nicht die punktweise Invertierung $1/f$ von f bezeichnet. Diese Verwechslung liegt nahe, da ja in einem Körper x^{-1} auch als $1/x$ geschrieben wird. Den Unterschied zwischen f^{-1} und $1/f$ kann man sich leicht klarmachen:

Beispiel

Sei $f: \mathbb{R} - \{0\} \rightarrow \mathbb{R}$ die Funktion mit $f(x) = 1/x$ für alle $x \neq 0$. Dann ist f injektiv und alle Funktionswerte sind von null verschieden. Also sind sowohl f^{-1} als auch $1/f$ definiert. Es gilt aber für alle $x \neq 0$:

$$f^{-1}(x) = 1/x = f(x), \quad (1/f)(x) = 1/f(x) = 1/(1/x) = x.$$

Also gilt $f^{-1} = f$, aber $1/f$ ist die Identität auf $\mathbb{R} - \{0\}$.

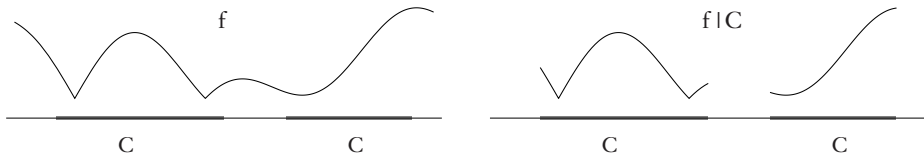
Gilt $y = f(x_1) = f(x_2)$ für $x_1 \neq x_2$, so ist f^{-1} als Funktion nicht definiert. Wir können aber für jede Funktion f die Menge $\{ (y, x) \mid (x, y) \in f \}$ definieren und als *Relation* untersu-

chen (diese Menge wird auch oft mit f^{-1} bezeichnet). Speziell ist hier die Menge aller x mit $f(x) = y$ von Bedeutung. Wir untersuchen sie in der folgenden Sektion genauer. Hier betrachten wir noch eine Operation, die in vielen analytisch interessanten Fällen eine gute „partielle Umkehrung“ ermöglicht:

Definition (*Einschränkung einer Funktion*)

Sei $f : A \rightarrow B$ eine Funktion und sei $C \subseteq A$. Dann ist die *Einschränkung* $f|_C : C \rightarrow B$ von f auf C definiert durch

$$(f|_C)(x) = f(x) \quad \text{für alle } x \in C.$$



Die Einschränkung einer Funktion verkleinert den Definitionsbereich, lässt die Wirkung der Funktion aber unangetastet. Diese Begriffsbildung bereitet Anfängern oft Schwierigkeiten, da sie Funktionen gerne als Terme auffassen und dann stillschweigend maximale Definitionsbereiche annehmen. Zur Illustration betrachten wir zwei Beispiele für nichtinjektive Funktionen, die durch eine geeignete Einschränkung zu wichtigen Umkehrfunktionen Anlass geben.

Beispiele

- (1) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ die zweite Potenz auf \mathbb{R} , d. h. $f(x) = x^2$ für alle $x \in \mathbb{R}$. Dann ist f nicht injektiv. Aber die Einschränkung $g = f|_{[0, +\infty[}$ ist injektiv. Die Umkehrfunktion g^{-1} von g ist die Quadratwurzelfunktion auf $[0, +\infty[$, es gilt

$$g^{-1}(x) = \sqrt{x} \quad \text{für alle } x \geq 0.$$

Da f jeden Wert höchstens zweimal annimmt, gilt

$$y = x^2 \quad \text{genau dann, wenn} \quad x = g^{-1}(y) = \sqrt{y} \quad \text{oder} \quad x = -g^{-1}(y) = -\sqrt{y}.$$

Dass man den rechten Ast der Parabel zur Invertierung auswählt, ist eine natürliche Konvention. Prinzipiell wäre auch der linke Ast geeignet.

- (2) Die Kosinusfunktion $\cos : \mathbb{R} \rightarrow \mathbb{R}$ nimmt jeden Wert unendlich oft an. Die Einschränkung $\cos_0 = \cos|_{[0, \pi]}$ des Kosinus auf seinen sogenannten Hauptzweig ist injektiv. Die Umkehrfunktion von \cos_0 ist als Arkuskosinus $\arccos : [-1, 1] \rightarrow [0, \pi]$ bekannt. Für alle $x \in [0, \pi]$ und alle y gilt:

$$y = \cos(x) \quad \text{genau dann, wenn} \quad x = \arccos(y).$$

Die Bedingung $x \in [0, \pi]$ ist hier wichtig. Im Allgemeinen erhalten wir die Stelle x nicht zurück, wenn wir auf den Wert $y = \cos(x)$ den Arkuskosinus anwenden. Es gilt zum Beispiel $\cos(2\pi) = 1$, aber $\arccos(1) = 0$. Allgemeiner gilt $\arccos(\cos(a2\pi)) = \arccos(1) = 0$ für alle $a \in \mathbb{Z}$.

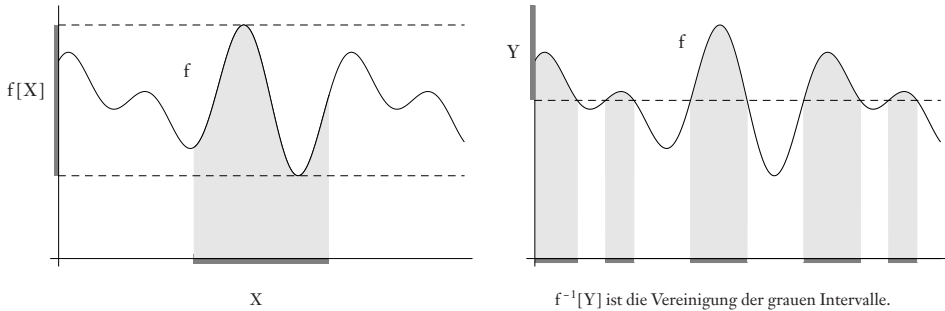
1.9 Bild und Urbild

Definition (Bild, Urbild)

Sei $f : A \rightarrow B$ eine Funktion, und seien $X \subseteq A$ und $Y \subseteq B$. Dann setzen wir:

$$f[X] = f(X) = \{f(x) \mid x \in X\}, \quad f^{-1}[Y] = f^{-1}(Y) = \{x \in A \mid f(x) \in Y\}.$$

Die Menge $f[X]$ heißt das *Bild* von X , die Menge $f^{-1}[Y]$ das *Urbild* von Y unter f .



Das Bild $f[X]$ ist also eine Menge von Funktionswerten und damit eine Teilmenge des Wertebereichs von f . Wir schicken alle $x \in X$ durch f und sammeln, was wir erhalten. Speziell gilt $f[A] = \text{Bild}(f)$. Das Urbild $f^{-1}[Y]$ ist dagegen eine Teilmenge von $A = \text{Def}(f)$. Wir sammeln alle Stellen $x \in A$, die durch Anwendung von f in Y landen. Wichtig ist:

Die Menge $f^{-1}[Y]$ ist immer definiert, die Injektivität von f wird nicht gefordert.

Die Notation $f[X]$ betont neben der gleichwertigen Notation $f(X)$, dass keine Funktionsauswertung der Form „ f an der Stelle X “ vorliegt. Wir verwenden im Folgenden $f[X]$.

Beispiele

(1) Sei $\text{id} : \mathbb{R} \rightarrow \mathbb{R}$ die Identität. Dann gilt für alle $X \subseteq \mathbb{R}$: $\text{id}[X] = \text{id}^{-1}[X] = X$.

(2) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ konstant gleich c . Dann gilt für alle $X, Y \subseteq \mathbb{R}$:

$$f[X] = \{c\}, \text{ falls } X \text{ nichtleer, und } f[X] = \emptyset \text{ sonst.}$$

$$f^{-1}[Y] = \mathbb{R}, \text{ falls } c \in Y, \text{ und } f^{-1}[Y] = \emptyset \text{ sonst.}$$

(3) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ die zweite Potenz, $f(x) = x^2$ für alle x . Dann gilt:

$$f[[0, 2]] = f[[-2, 2]] = [0, 4],$$

$$f^{-1}[[0, 2]] = [-\sqrt{2}, \sqrt{2}], \quad f^{-1}[[-\infty, 0]] = \{0\}.$$

(4) Sei $\cos : \mathbb{R} \rightarrow \mathbb{R}$ die Kosinusfunktion. Dann gilt

$$\cos[\mathbb{R}] = \text{Bild}(\cos) = \cos[[0, \pi]] = \cos[[0, 2\pi]] = [-1, 1],$$

$$\cos[\{a\pi \mid a \in \mathbb{Z}\}] = \{1, -1\},$$

$$\cos^{-1}[\{0\}] = \{x \in \mathbb{R} \mid \cos(x) = 0\} = \{\pi/2 + a\pi \mid a \in \mathbb{Z}\}.$$

Die Rechenregeln für die Bild- und Urbildoperation sind unterschiedlich. So gilt z. B.

$$f^{-1}[Y_1 \cap Y_2] = f^{-1}[Y_1] \cap f^{-1}[Y_2] \quad \text{für alle } Y_1, Y_2 \subseteq B,$$

während für das Bild im Allgemeinen lediglich gilt, dass

$$f[X_1 \cap X_2] \subseteq f[X_1] \cap f[X_2] \quad \text{für alle } X_1, X_2 \subseteq A.$$

Die Menge $f^{-1}[Y_1 \cap Y_2]$ besteht aus denjenigen Elementen von A , die f nach $Y_1 \cap Y_2$ transportiert, und dies sind genau diejenigen Elemente von A , die durch f sowohl nach Y_1 als auch nach Y_2 transportiert werden. Die ausführliche Argumentation liest sich wie folgt:

$$f^{-1}[Y_1 \cap Y_2] = \quad (\text{Definition des Urbilds})$$

$$\{x \in A \mid f(x) \in Y_1 \cap Y_2\} = \quad (\text{Definition des Durchschnitts})$$

$$\{x \in A \mid f(x) \in Y_1 \text{ und } f(x) \in Y_2\} = \quad (\text{Definition des Durchschnitts})$$

$$\{x \in A \mid f(x) \in Y_1\} \cap \{x \in A \mid f(x) \in Y_2\} = \quad (\text{Definition des Urbilds})$$

$$f^{-1}[Y_1] \cap f^{-1}[Y_2].$$

Dagegen können bei der Bild-Operation Überlappungen auftreten, die dazu führen, dass die Gleichheit verloren geht:

Beispiel

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ die zweite Potenz und seien $X_1 =]-\infty, 0]$, $X_2 = [0, +\infty[$. Dann gilt:

$$f[X_1 \cap X_2] = f[\{0\}] = \{0\} \subseteq [0, +\infty[= [0, +\infty[\cap [0, +\infty[= f[X_1] \cap f[X_2].$$

Erstellt man für die Inklusion $f[X_1 \cap X_2] \subseteq f[X_1] \cap f[X_2]$ eine ausführliche Argumentation, so findet man eine Implikation „ $\exists x (x \in M \wedge x \in N) \rightarrow \exists x x \in M \wedge \exists x x \in N$ “, die in der Rückrichtung nicht allgemeingültig ist (vgl. Anhang 2).

Wir stellen einige Rechenregeln für Bilder und Urbilder in einer Tabelle zusammen.

$f[X_1 \cap X_2] \subseteq f[X_1] \cap f[X_2]$ mit „ \subseteq “, falls f injektiv ist	$f^{-1}[Y_1 \cap Y_2] = f^{-1}[Y_1] \cap f^{-1}[Y_2]$
$f[X_1 \cup X_2] = f[X_1] \cup f[X_2]$	$f^{-1}[Y_1 \cup Y_2] = f^{-1}[Y_1] \cup f^{-1}[Y_2]$
$f[X_1 - X_2] \supseteq f[X_1] - f[X_2]$ mit „ \supseteq “, falls f injektiv ist	$f^{-1}[Y_1 - Y_2] = f^{-1}[Y_1] - f^{-1}[Y_2]$

Bilder und Urbilder tauchen in der Analysis häufig auf. Es gilt zum Beispiel (vgl. 5.9):

Das Bild eines Intervalls $[a, b]$ unter einer stetigen Funktion ist ein Intervall $[c, d]$.

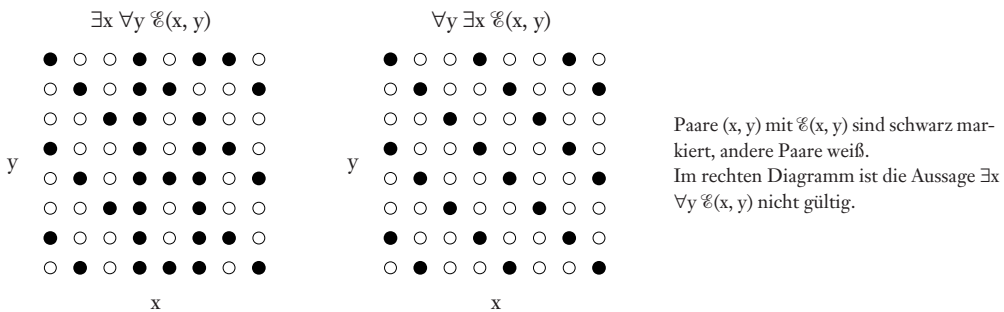
Und eine „höhere Formulierung“ der Stetigkeit einer Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ lautet, dass die Urbilder von offenen Intervallen unter f Vereinigungen von offenen Intervallen sind.

1.10 Umgang mit Quantoren

Satz (Quantorenregeln für den Allquantor \forall und den Existenzquantor \exists)

Für alle Eigenschaften \mathcal{E} und die Quantoren \forall „für alle“ und \exists „es gibt“ gilt:

- (a) $\neg \forall x \mathcal{E}(x)$ ist äquivalent zu $\exists x \neg \mathcal{E}(x)$,
 $\neg \exists x \mathcal{E}(x)$ ist äquivalent zu $\forall x \neg \mathcal{E}(x)$, (Verneinungsregeln)
- (b) $\forall x \forall y \mathcal{E}(x, y)$ ist äquivalent zu $\forall y \forall x \mathcal{E}(x, y)$,
 $\exists x \exists y \mathcal{E}(x, y)$ ist äquivalent zu $\exists y \exists x \mathcal{E}(x, y)$,
 $\exists x \forall y \mathcal{E}(x, y)$ impliziert $\forall y \exists x \mathcal{E}(x, y)$. (Vertauschungsregeln)



Wir verwenden hier das Zeichen „ \neg “ für die Negation „nicht“ (vgl. auch Anhänge 1 und 2). Die Ausdrücke werden exemplarisch wie folgt gelesen:

$\forall x \mathcal{E}(x)$	für alle x gilt $\mathcal{E}(x)$
$\exists x \mathcal{E}(x)$	es gibt ein x mit $\mathcal{E}(x)$
$\neg \forall x \mathcal{E}(x)$	nicht für alle x gilt $\mathcal{E}(x)$
$\forall x \forall y \mathcal{E}(x, y)$	für alle x und für alle y gilt $\mathcal{E}(x, y)$
$\forall x \exists y \mathcal{E}(x, y)$	für alle x gibt es ein y mit $\mathcal{E}(x, y)$

Die erste Zeile von Teil (a) des Satzes besagt also ausformuliert, dass

„nicht für alle x gilt $\mathcal{E}(x)$ “ äquivalent ist zu „es gibt ein x , für das $\mathcal{E}(x)$ nicht gilt“.

Ebenso besagt die zweite Zeile von Teil (a), dass

„es gibt kein x mit $\mathcal{E}(x)$ “ äquivalent ist zu „für alle x gilt non $\mathcal{E}(x)$ “.

Wenn es nicht an allen Tagen der Woche regnet, so gibt es einen Tag der Woche, an dem es nicht regnet – und umgekehrt. Und gibt es in einer Klasse kein Kind, das rote Haare hat, so haben alle Kinder der Klasse eine von rot verschiedene Haarfarbe – und umgekehrt.

Die Verneinungsregeln verallgemeinern sich für mehrere Quantoren wie folgt:

Die Aussage ...	ist äquivalent zu ...
$\neg \forall x \exists y \mathcal{E}(x, y)$	$\exists x \forall y \neg \mathcal{E}(x, y)$
$\neg \exists x \forall y \mathcal{E}(x, y)$	$\forall x \exists y \neg \mathcal{E}(x, y)$
$\neg \forall x \exists y \forall z \mathcal{E}(x, y, z)$	$\exists x \forall y \exists z \neg \mathcal{E}(x, y, z)$
$\neg \exists x \forall y \exists z \mathcal{E}(x, y, z)$	$\forall x \exists y \forall z \neg \mathcal{E}(x, y, z)$

Als Merkgregel:

Beim „Reinziehen“ einer Negation werden \forall und \exists ausgetauscht und die Negation landet vor der Eigenschaft, auf die Quantoren wirken.

Beispiel

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$. Dann bedeutet $\forall y \exists x f(x) = y$, dass f surjektiv ist (wobei die Variablen x und y über \mathbb{R} laufen), und $\neg \forall y \exists x f(x) = y$ ist äquivalent zu $\exists y \forall x f(x) \neq y$.

Das ist alles sehr hübsch, aber der Leser wird sich fragen, was dieses Logiktraining mit Analysis zu tun hat. Die Antwort ist: Die Verneinungsregeln werden in der Analysis tatsächlich häufig gebraucht. Begriffe wie

*Grenzwert einer Folge, Stetigkeit einer Funktion in einem Punkt,
gleichmäßige Konvergenz einer Funktionenfolge,
Riemann-Integrierbarkeit einer Funktion auf einem Intervall $[a, b]$*

besitzen Definitionen, bei denen sich mehrere All- und Existenzquantoren abwechseln. Eine übliche Formulierung der Stetigkeit einer Funktion in einem Punkt p hat zum Beispiel den Typ „ $\forall \epsilon \exists \delta$ “. Kennt man die Verneinungsregeln, so weiß man, dass die Unstetigkeit der Funktion im Punkt p vom Typ „ $\exists \epsilon \forall \delta$ “ ist.

Die beiden ersten Aussagen von (b) besagen, dass man Quantoren des gleichen Typs vertauschen darf. Das ist unproblematisch. Suggestiv und besser lesbar schreibt man oft

$$\forall x \forall y \mathcal{E}(x, y) \text{ als } \forall x, y \mathcal{E}(x, y) \quad \text{und} \quad \exists x \exists y \mathcal{E}(x, y) \text{ als } \exists x, y \mathcal{E}(x, y).$$

Eine große Fehlerquelle stellt dagegen die Vertauschung von verschiedenen Quantoren dar. Die dritte Vertauschungsregel besagt, dass der Übergang von $\exists \forall$ zu $\forall \exists$ erlaubt ist. Eine Eselsbrücke hierzu lautet:

Ein Allesfresser darf Alles Essen.

Der Übergang von $\forall \exists$ zu $\exists \forall$ führt dagegen oft zu falschen Aussagen. Das obige Diagramm führt den Unterschied zwischen $\exists x \forall y \mathcal{E}(x, y)$ und $\forall y \exists x \mathcal{E}(x, y)$ vor Augen. Wir betrachten die Relation $R = \{ (x, y) \mid \mathcal{E}(x, y) \}$ als Punktwolke. Anschaulich bedeutet $\exists x \forall y \mathcal{E}(x, y)$, dass es eine Senkrechte gibt, die schwarz gefärbt ist. Dagegen bedeutet $\forall y \exists x \mathcal{E}(x, y)$, dass alle Waagrechten mindestens einen schwarzen Punkt treffen. Die erste Aussage impliziert die zweite, aber die Umkehrung ist im Allgemeinen nicht richtig. „Einer ist für alles gut“ ist stärker als „Für alles ist irgendwer gut.“

1.11 Die vollständige Induktion

Satz (vollständige Induktion)

Sei $\mathcal{E}(n)$ eine Eigenschaft für natürliche Zahlen. Es gelte:

- (a) $\mathcal{E}(0)$.
- (b) Für alle $n \in \mathbb{N}$ gilt: $\mathcal{E}(n)$ impliziert $\mathcal{E}(n+1)$.

Dann gilt $\mathcal{E}(n)$ für alle natürlichen Zahlen.



Dieser Satz ist nicht nur in der Zahlentheorie, sondern in der gesamten Mathematik im Einsatz. Obwohl in der Analysis die reellen Zahlen und die komplexen Zahlen im Zentrum stehen, so werden doch häufig auch die natürlichen Zahlen eingesetzt, etwa in Potenzierungen $(x+y)^n$ oder in der Bildung der n -ten Ableitung $f^{(n)}$ einer Funktion f . Und dann ist die vollständige Induktion oft das Mittel der Wahl, um Aussagen wie zum Beispiel die binomische Formel $(x+y)^n = \sum_{k \leq n} \binom{n}{k} x^k y^{n-k}$ oder die Existenz der Ableitungen $\log^{(n)}$ des natürlichen Logarithmus $\log:]0, +\infty[\rightarrow \mathbb{R}$ für alle n zu beweisen.

Das Wort „Induktion“ steht im Gegensatz zur „Deduktion“, also einer Herleitung, eines Beweises. Induktiv argumentieren bedeutet, dass man von Einzelfällen auf das Allgemeine schließt. Wenn es an den ersten drei Urlaubstagen regnet, sagt man „induktiv“: „Hier regnet es immer, wir fahren wieder weg.“ Und wenn eine Eigenschaft $\mathcal{E}(n)$ für alle Zahlen $0, 1, 2, \dots, 100$ gilt, so liegt es nahe, dass sie für alle natürlichen Zahlen gilt. Diese endliche Induktion dient in der Mathematik dem Finden von Hypothesen – hier kann sie sehr wertvoll sein –, aber ein Beweis von „für alle n gilt $\mathcal{E}(n)$ “ ist auch dann nicht geführt, wenn die Eigenschaft \mathcal{E} für alle Zahlen $0, 1, 2, \dots, 10^{10}$ nachgewiesen wurde. Um zu betonen, dass man nicht von Einzelfällen auf das Allgemeine schließt, ist im deutschsprachigen Raum die Sprechweise „vollständige Induktion“ populär. Oft spricht man aber auch nur von „Induktion“, da ohnehin klar ist, dass in der Mathematik keine unvollständigen Beweise akzeptiert werden. Im Englischen ist „(mathematical) induction“ üblich. Der Ausdruck „complete induction“ bedeutet im Englischen dagegen eine starke Form der Induktion, die wir in der nächsten Sektion besprechen.

Der Leser wird in seinen Vorlesungen oder in der Literatur die vollständige Induktion häufig als „Beweisprinzip“ bezeichnet sehen. Dagegen ist nichts einzuwenden, da der obige Satz tatsächlich eine Methode zur Verfügung stellt, die in Anwendungen so empfunden wird wie etwa das Prinzip des indirekten Beweises. Dennoch gibt es Unterschiede: In der mengentheoretisch fundierten Mathematik ist die vollständige Induktion einfach ein beweisbarer Satz. Nur in der reinen Zahlentheorie betrachtet man die Induktion als Axiom.

Wir beschreiben nun, wie der Satz über die vollständige Induktion in der Praxis eingesetzt wird. Es liege also eine Behauptung

„Für alle $n \in \mathbb{N}$ gilt $\mathcal{E}(n)$.“

vor, die wir beweisen möchten. Ein Induktionsbeweis zerfällt in zwei Teile:

Teil 1: Induktionsanfang

Wir beweisen, dass die Null die Eigenschaft \mathcal{E} besitzt.

Teil 2: Induktionsschritt von n nach $n + 1$

Der zweite und oft viel schwierigere Teil des Beweises ist der Induktionsschritt. Dafür betrachten wir eine beliebige natürliche Zahl n , die für den Rest des Beweises nicht mehr verändert wird. Nun bekommen wir, ohne arbeiten zu müssen, ein Geschenk: Wir dürfen annehmen, dass n die Eigenschaft \mathcal{E} besitzt. Diese Gültigkeit von $\mathcal{E}(n)$ heißt *Induktionsvoraussetzung*, abgekürzt I. V. – man sollte sie als „Induktionsgeschenk“ auffassen. Oft notiert man die Induktionsvoraussetzung suggestiv als: „Sei also $\mathcal{E}(n)$ bereits bewiesen.“ Logisch korrekter ist eine schlichte Formulierung wie: „Es gelte also $\mathcal{E}(n)$ “ oder „Wir nehmen $\mathcal{E}(n)$ an.“ Nun müssen wir arbeiten und beweisen, dass $n + 1$ die Eigenschaft \mathcal{E} besitzt. Neben bereits bewiesenen einschlägigen Sätzen werden wir hierzu vor allem auch unser Geschenk einsetzen.

Nach Teil 1 und Teil 2 gelten die Voraussetzungen (a) und (b) des Satzes über die vollständige Induktion. Damit ist gezeigt, dass $\mathcal{E}(n)$ für alle n gilt.

Beispiel

Sei $x \in \mathbb{R}$ mit $x \geq -1$. Wir beweisen durch Induktion:

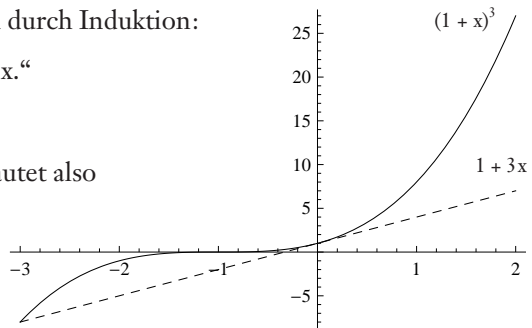
„Für alle $n \geq 0$ gilt $(1 + x)^n \geq 1 + nx$.“

(Bernoulli-Ungleichung)

Die betrachtete Eigenschaft $\mathcal{E}(n)$ lautet also

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

Die reelle Zahl x ist hier ein sog. Parameter der Eigenschaft \mathcal{E} .



Induktionsanfang $n = 0$:

Es gilt $(1 + x)^0 = 1 \geq 1 + 0x$. Also gilt $\mathcal{E}(0)$.

Induktionsschritt von n nach $n + 1$:

Sei n eine beliebige natürliche Zahl. Es gelte $(1 + x)^n \geq 1 + nx$ (dies ist die Induktionsvoraussetzung $\mathcal{E}(n)$). Dann gilt $\mathcal{E}(n + 1)$, denn

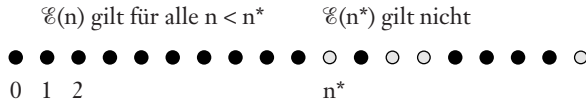
$$\begin{aligned} (1 + x)^{n+1} &= (1 + x)^n (1 + x) \geq_{\text{I.V. und } x \geq -1} (1 + nx)(1 + x) = \\ &= 1 + nx + x + nx^2 \geq_{nx^2 \geq 0} 1 + nx + x = 1 + (n + 1)x. \end{aligned}$$

Damit ist die Bernoulli-Ungleichung für x bewiesen.

1.12 Das Prinzip vom kleinsten Element

Satz (Prinzip vom kleinsten Element)

Sei $\mathcal{E}(n)$ eine Eigenschaft natürlicher Zahlen, und es gebe eine natürliche Zahl n , die die Eigenschaft $\mathcal{E}(n)$ nicht besitzt. Dann gibt es eine natürliche Zahl n^* derart, dass $\mathcal{E}(n^*)$ nicht gilt, aber $\mathcal{E}(n)$ für alle $n < n^*$ gilt.



Das Prinzip ist anschaulich klar: Wir zählen 0, 1, 2, 3, ... und warten so lange ab, bis wir zum ersten Mal eine natürliche Zahl n^* finden, für die $\mathcal{E}(n^*)$ nicht gilt. Dies ist nach Voraussetzung irgendwann der Fall. Der Satz wird dieser Anschauung folgend auch oft der *Satz vom kleinsten Ausreißer* oder *kleinsten Gegenbeispiel* genannt. Gibt es überhaupt ein Gegenbeispiel zur Eigenschaft $\mathcal{E}(n)$, so gibt es ein kleinstes Gegenbeispiel n^* . Dabei ist $n^* = 0$ durchaus möglich, d. h., bereits $\mathcal{E}(0)$ gilt nicht. Wie der weitere Verlauf der Gültigkeit der Eigenschaft ist, ist unbestimmt. Das kleinste Gegenbeispiel n^* kann das einzige Gegenbeispiel sein, aber es kann auch $\mathcal{E}(n^* + 1)$ gelten und $\mathcal{E}(n^* + 2)$ nicht gelten usw.

Eine aus der Sicht der Analysis interessante Anwendung des Prinzips ist der folgende „primzahlfreie“ Beweis der Irrationalität der Quadratwurzel aus 2.

Beispiel

„ $\sqrt{2}$ ist irrational, d. h., es gibt keine natürlichen Zahlen m und n mit $(m/n)^2 = 2$.“

Annahme nicht. Dann gibt es nach dem Prinzip vom kleinsten Element eine kleinste natürliche Zahl n^* , für welche ein $m^* \in \mathbb{N}$ existiert mit

$$(m^*/n^*)^2 = 2.$$

Dann ist $m^* \geq n^*$ und $m^* < 2n^*$ (da sonst $2n^{*2} = m^{*2} \geq 4n^{*2}$). Wir setzen nun:

$$n = m^* - n^* \in \mathbb{N} \text{ und } m = 2n^* - m^* \in \mathbb{N}.$$

Wegen $m^{*2} = 2n^{*2}$ ist

$$\begin{aligned} 2n^2 &= 2(m^* - n^*)^2 = 2m^{*2} - 4n^*m^* + 2n^{*2} = \\ &4n^{*2} - 4n^*m^* + m^{*2} = (2n^* - m^*)^2 = m^2, \end{aligned}$$

also gilt auch $(m/n)^2 = 2$. Nun ist aber

$$n = m^* - n^* < 2n^* - n^* = n^*,$$

im Widerspruch zur minimalen Wahl von n^* .

Einen weiteren Beweis für die Irrationalität von $\sqrt{2}$ werden wir im nächsten Abschnitt kennenlernen.

Das Prinzip vom kleinsten Element ist eng mit der vollständigen Induktion verwandt. Man kann die vollständige Induktion verwenden, um das Prinzip zu beweisen. Zum anderen führt das Prinzip selbst zu einer neuen Form der Induktion. Hierzu schreiben wir das Prinzip kompakt auf:

$$\exists n \neg \mathcal{E}(n) \rightarrow \exists n^* (\forall k < n^* \mathcal{E}(k) \wedge \neg \mathcal{E}(n^*)).$$

Eine Implikation $A \rightarrow B$ ist nach dem *Kontrapositionsgesetz* äquivalent zu $\neg B \rightarrow \neg A$ (vgl. Anhang 1). Wenden wir dies an, so erhalten wir

$$\neg \exists n^* (\forall k < n^* \mathcal{E}(k) \wedge \neg \mathcal{E}(n^*)) \rightarrow \neg \exists n \neg \mathcal{E}(n).$$

Die Verneinungsregeln für die Quantoren und die Äquivalenz von $\neg(A \wedge \neg B)$ und $A \rightarrow B$ und weiter von $\neg \neg A$ und A liefern nun:

$$\forall n^* (\forall k < n^* \mathcal{E}(k) \rightarrow \mathcal{E}(n^*)) \rightarrow \forall n \mathcal{E}(n).$$

Diese Aussage ist als „starke Induktion“ oder „Wertverlaufsinduktion“ bekannt. Besser lesbar aufgeschrieben lautet sie:

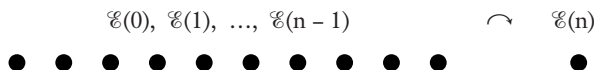
Satz (*starke Induktion, Wertverlaufsinduktion*)

Sei $\mathcal{E}(n)$ eine Eigenschaft für natürliche Zahlen. Für alle $n \in \mathbb{N}$ gelte:

(+) Gilt $\mathcal{E}(k)$ für alle $k < n$, so gilt $\mathcal{E}(n)$.

Dann gilt $\mathcal{E}(n)$ für alle natürlichen Zahlen.

Induktionsschritt



Induktionsvoraussetzung

Die rein logische Umformung zeigt, dass die starke Induktion eine „positive“ Form des Prinzips vom kleinsten Element ist. In der Praxis bedeutet der Satz, dass wir zum Beweis von $\mathcal{E}(n)$ nicht nur $\mathcal{E}(n-1)$ annehmen dürfen, sondern stärker $\mathcal{E}(k)$ für alle $k < n$. Unser Geschenk zum Beweis von $\mathcal{E}(n)$ lautet nun also: „Für alle $k < n$ gilt $\mathcal{E}(k)$.“ Ein Induktionsanfang entfällt hier, das gesamte Argument besteht im Schritt von „gültig für alle Zahlen kleiner als n “ zu „gültig für n “. (Im Fall $n = 0$ ist das Geschenk eine leere Schachtel.)

Schließlich notieren wir noch eine mengentheoretische Version des Prinzips:

Satz (*Wohlordnungssatz für \mathbb{N}*)

Sei A eine nichtleere Menge natürlicher Zahlen. Dann besitzt A ein kleinstes Element, d.h., es gibt ein $n^* \in A$ derart, dass $n \notin A$ für alle $n < n^*$ gilt.

Zum Kontrast betrachte der Leser die Menge $\{1, 1/2, 1/4, 1/8, \dots\}$ rationaler Zahlen. Diese Menge ist nichtleer, besitzt aber kein kleinstes Element. Gleiches gilt für \mathbb{Z} oder das offene Intervall $]0, 1[$ in \mathbb{R} .

2. Kapitel

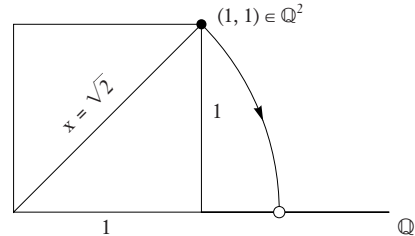
Die reellen und komplexen Zahlen

2.1 Irrationale Zahlen

Satz (Irrationalität der Quadratwurzel aus 2)

Sei d die Länge der Diagonalen des Einheitsquadrats. Dann ist d keine rationale Zahl, d. h., es gilt $d \neq n/m$ für alle natürlichen Zahlen n, m .

Die alten Griechen kannten bereits arithmetische und geometrische Beweise dieses Satzes. Das klassische Argument, das sich insbesondere bei Euklid findet, ist:



Ein arithmetischer Beweis

Wir nehmen an, dass $d = n/m$ für natürliche Zahlen n und m gilt. Durch Kürzen des Bruchs n/m dürfen wir weiter annehmen, dass n und m nicht beide gerade sind. Nach dem Satz des Pythagoras gilt $d^2 = 1 + 1$, also $n^2/m^2 = 2$. Also ist

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

und damit ist n^2 gerade. Dann ist aber auch n gerade, da das Quadrat einer ungeraden Zahl ungerade ist. Folglich ist n^2 durch 4 teilbar. Wegen $m^2 = n^2/2$ ist also m^2 durch 2 teilbar, also ist m^2 gerade. Folglich ist auch m gerade, *Widerspruch!*

Der Satz des Pythagoras liefert die entscheidende Gleichung $2m^2 = n^2$. Durch eine zahlentheoretische Argumentation mit geraden und ungeraden Zahlen können wir dann zeigen, dass n und m beide gerade sein müssen, was der Annahme der Teilerfremdheit von n und m widerspricht. (Der Leser vergleiche dazu auch den induktiven Beweis in 1.12.)

Für die Analysis ist das Ergebnis von fundamentaler Bedeutung. Sie benötigt ein Modell für ein Kontinuum, das alle in mathematischen Konstruktionen auftretenden Punkte einer Linie beschreiben kann. Der Satz zeigt:

Die rationalen Zahlen sind kein geeignetes Modell für ein Kontinuum.

Da zwischen zwei rationalen Zahlen p und q immer noch eine weitere rationale Zahl liegt, etwa $(p + q)/2$, besitzt \mathbb{Q} einen scheinbar kontinuierlichen Charakter. Und doch fängt die „Linie“ \mathbb{Q} , obwohl sie dicht mit Punkten bestückt ist, nicht alle Größen ein, die in der Mathematik auftreten. Übertragen wir wie im Diagramm oben die Diagonale des Einheitsquadrats auf die x -Achse, so treffen wir dort keine rationale Zahl an. Diese Lücke lässt sich funktional als „Fehlen von Nullstellen“ formulieren:

Beispiel

Sei $f: \mathbb{Q} \rightarrow \mathbb{Q}$ mit $f(q) = q^2 - 2$ für alle $q \in \mathbb{Q}$. Es gilt

$$f(0) = -2 < 0 \quad \text{und} \quad f(2) = 2 > 0.$$

Aber die Funktion f besitzt keine Nullstelle, da aus $f(q) = 0$ folgen würde, dass $q^2 = 2$, was für $q \in \mathbb{Q}$ nach dem Satz ausgeschlossen ist.

Erst wenn wir die reellen Zahlen zugrundelegen, können wir den allgemeinen Zwischenwertsatz beweisen, der für eine stetige Funktion die Existenz einer Nullstelle sichert, falls die Funktion sowohl negative als auch positive Werte annimmt (vgl. 5.8).

Es gibt viele weitere irrationale Zahlen. Mit $w = \sqrt{2}$ ist zum Beispiel auch

$$w^* = \sqrt{w} = \sqrt[4]{2}$$

irrational. Denn wäre w^* rational, so wäre auch $w^* \cdot w^* = w$ rational, da das Produkt zweier rationaler Zahlen rational ist. Induktiv zeigt die Überlegung, dass die Zahlen

$$\sqrt[2]{2}, \sqrt[4]{2}, \sqrt[8]{2}, \sqrt[16]{2}, \sqrt[32]{2}, \dots$$

irrational sind. Damit haben wir unendlich viele irrationale Zahlen gefunden. Den Griechen war neben der Irrationalität von $\sqrt{2}$ auch bekannt, dass die Zahlen

$$\sqrt{3}, \sqrt{5}, \sqrt{6}, \sqrt{7}, \sqrt{8}, \sqrt{10}, \sqrt{11}, \sqrt{12}, \sqrt{13}, \sqrt{14}, \sqrt{15}, \sqrt{17}$$

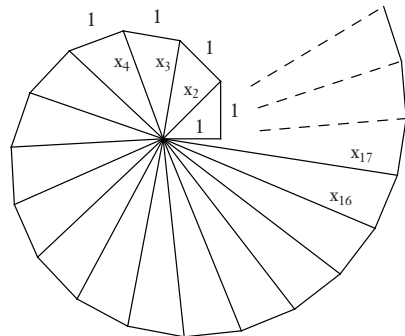
irrational sind. Warum sie bei 17 aufgehört haben, ist nicht restlos geklärt, aber die Liste legt die Vermutung nahe, dass die Quadratwurzel einer natürlichen Zahl n immer irrational ist, wenn n keine Quadratzahl 1, 4, 9, 16, 25, 36, ... ist. Von Gauß stammt ein Satz, der diese Vermutung – und viel mehr – beweist:

Satz (Satz von Gauß)

Seien $a_0, \dots, a_{k-1} \in \mathbb{Z}$ und sei $x \in \mathbb{R}$ mit

$$x^k + a_{k-1} x^{k-1} + \dots + a_1 x^1 + a_0 = 0.$$

Dann ist x eine ganze Zahl oder irrational.



Eine geometrische Erzeugung von $x_n = \sqrt{n}$ für $n = 2, 3, 4, \dots$. Ab $n = 18$ überlappt sich die Schnecke, was eine Erklärung für die historische Rolle von $\sqrt{17}$ sein könnte.

Beispiele

Die Gleichung $x^2 = 17$ hat keine ganzzahligen Lösungen. Also ist die Zahl $\sqrt{17}$, eine Lösung der Gleichung, irrational. Ebenso zeigt der Satz von Gauß für die Gleichung $x^3 = 3$, dass $\sqrt[3]{3}$ irrational ist.

Weiß man, dass die rationalen Zahlen durch endliche oder periodisch unendliche Dezimaldarstellungen charakterisiert sind (vgl. 2.8 und 4.4), so ist es einfach, spielerische Beispiele für irrationale Zahlen zu erzeugen:

Beispiele

Die unendlichen Dezimalbrüche

$$0,101001000100001\dots, \quad 0,010011000111\dots, \quad 0,13579111315171921\dots$$

sind nicht periodisch, stellen also irrationale Zahlen dar. Im Gegensatz zu $\sqrt{2}$ haben diese Zahlen aber keine offensichtliche geometrische oder analytische Bedeutung. Sie wurden konstruiert, um irrational zu sein.

2.2 Algebraische und transzendente Zahlen

Definition (algebraische und transzendente Zahlen, \mathbb{A})

Eine reelle Zahl x heißt *algebraisch*, falls rationale Zahlen a_0, \dots, a_k mit $a_k \neq 0$ existieren, sodass

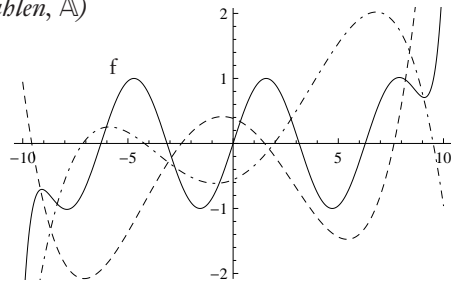
$$a_k x^k + a_{k-1} x^{k-1} + \dots + a_1 x + a_0 = 0.$$

Wir setzen

$$\mathbb{A} = \{x \in \mathbb{R} \mid x \text{ ist algebraisch}\}.$$

Eine reelle Zahl x heißt *transzendent*, falls x nicht algebraisch ist.

Es ist wichtig, nur rationale Koeffizienten a_0, \dots, a_k zuzulassen. Denn a ist eine Lösung von $x - a = 0$, sodass also jede reelle Zahl trivialerweise die Lösung einer algebraischen Gleichung mit reellen Koeffizienten ist. Dagegen erhalten wir die gleiche Menge \mathbb{A} , wenn wir nur ganzzahlige Koeffizienten betrachten. Multiplizieren wir nämlich die rationalen Koeffizienten a_0, \dots, a_k einer Gleichung mit ihrem gemeinsamen Hauptnenner, so erhalten wir eine neue Gleichung mit ganzzahligen Koeffizienten, die dieselben Lösungen wie die ursprüngliche Gleichung besitzt.



Die algebraischen Zahlen sind die Nullstellen von Polynomen mit rationalen Koeffizienten. Auch die Funktion f , die um die 0 herum wie der Sinus aussieht, ist ein solches Polynom (nämlich die Taylor-Approximation 21. Grades des Sinus im Nullpunkt, vgl. 7.12). Die erste positive Nullstelle p von h ist

$$p = 3,14159265360\dots, \text{ während}$$

$$\pi = 3,14159265358\dots$$

Beispiele

- (1) Die Gleichungen der Form $ax - b = 0$ mit ganzen Zahlen a und b , $a \neq 0$, zeigen, dass jede rationale Zahl algebraisch ist.
- (2) Die Gleichung $x^2 - 2 = 0$ zeigt, dass $\sqrt{2}$ und $-\sqrt{2}$ algebraisch sind. Damit ist insbesondere \mathbb{A} eine echte Obermenge von \mathbb{Q} .
- (3) Allgemeiner als (2) sind alle reellen Lösungen von Gleichungen zweiten Grades

$$x^2 + bx + c = 0$$

mit rationalen Koeffizienten b und c algebraisch. Nach der Lösungsformel für quadratische Gleichungen sind dies im Fall einer nichtnegativen Diskriminante $b^2 - 4c$ genau die Zahlen

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4c}}{2}.$$

- (4) Die Lösungen von

$$x^5 + 3x^4 - 2x^2 + x - 1 = 0$$

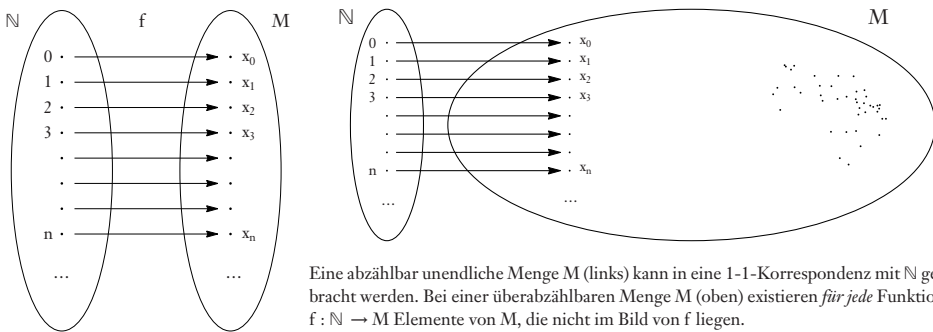
sind algebraisch. Das obige Diagramm zeigt den Graphen des zugehörigen Polynoms $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^5 + 3x^4 - 2x^2 + x - 1$ für alle x .

2.3 Abzählbarkeit und Überabzählbarkeit

Definition (*endlich, abzählbar, überabzählbar*)

Eine Menge M heißt

- (a) *endlich*, falls es eine Bijektion $f: \{0, \dots, n-1\} \rightarrow M$ für ein $n \in \mathbb{N}$ gibt,
- (b) *abzählbar unendlich*, falls es eine Bijektion $f: \mathbb{N} \rightarrow M$ gibt,
- (c) *abzählbar*, falls M endlich oder abzählbar unendlich ist,
- (d) *überabzählbar*, falls M nicht abzählbar ist.



Endliche Mengen lassen also eine Darstellung $M = \{x_0, \dots, x_{n-1}\}$ zu, abzählbaren Mengen eine Darstellung $M = \{x_0, \dots, x_n, \dots\}$ (vgl. 1.6). Die Elemente einer abzählbaren Menge können also mit den natürlichen Zahlen „vollständig durchgezählt“ werden.

Beispiele

(1) \mathbb{Z} ist abzählbar: $0, 1, -1, 2, -2, 3, -3, \dots$

(2) \mathbb{Q} ist abzählbar:

$$0/1, 1/1, -1/1, 2/1, -2/1, 1/2, -1/2, 3/1, -3/1, 1/3, -1/3, \\ 4/1, -4/1, 3/2, -3/2, 2/3, -2/3, 1/4, -1/4, \dots$$

Hier zählen wir alle gekürzten Brüche nach der Summe der Beträge von Zähler und Nenner auf.

(3) Die Menge aller algebraischen Gleichungen mit ganzzahligen Koeffizienten ist abzählbar:

$$x = 0, \quad x + 1 = 0, \quad x - 1 = 0, \quad x + 2 = 0, \quad x - 2 = 0, \quad 2x + 1 = 0, \\ -2x + 1 = 0, \quad x^2 = 0, \quad -x^2 = 0, \quad 2x - 1 = 0, \quad -2x - 2 = 0, \quad x^2 + 1 = 0, \quad \dots$$

Wir ordnen hier die Gleichungen nach der Summe $k + |a_0| + \dots + |a_k|$ ihres Grades und der Beträge ihrer Koeffizienten.

(4) \mathbb{A} ist abzählbar: Wir ersetzen in der Abzählung in (3) jede Gleichung durch ihre endlich vielen Nullstellen und streichen Wiederholungen.

Nun würde man vielleicht vermuten, dass auch die Menge der reellen Zahlen abzählbar ist. Unendlich ist scheinbar gleich unendlich, und damit gibt es scheinbar ebenso viele natürliche Zahlen wie rationale wie algebraische wie reelle Zahlen. Durch das sog. *Cantorsche Diagonalverfahren* lässt sich aber überraschend einfach zeigen, dass die reellen Zahlen nicht abzählbar sind. Hierzu betrachten wir eine beliebige Folge reeller Zahlen $x_0, x_1, \dots, x_n, \dots$ und schreiben in Dezimaldarstellung:

$$\begin{array}{rcccc} x_0 & = & \pm a_0, & d_{0,0} & d_{0,1} & d_{0,2} & \dots \\ x_1 & = & \pm a_1, & d_{1,0} & d_{1,1} & d_{1,2} & \dots \\ x_2 & = & \pm a_2, & d_{2,0} & d_{2,1} & d_{2,2} & \dots \\ & & & \dots & & d_{3,3} & \\ & & & & & & d_{4,4} \end{array}$$

Nun betrachten wir die Diagonale der Nachkommastellen des Diagramms und definieren für alle n :

$$b_n = \begin{cases} 6, & \text{falls } d_{n,n} = 7, \\ 7, & \text{falls } d_{n,n} \neq 7. \end{cases}$$

Nun setzen wir in Dezimaldarstellung

$$x^* = 0, b_0 b_1 b_2 \dots$$

Die Zahl x^* ist nach Konstruktion von allen x_0, x_1, x_2, \dots verschieden. Damit ist es also unmöglich, die reellen Zahlen abzuzählen. Wir haben bewiesen:

Satz (*Überabzählbarkeit der reellen Zahlen*)

Sind $x_0, x_1, \dots, x_n, \dots$ reelle Zahlen, so gibt es eine reelle Zahl x^* , die von allen x_n verschieden ist. Die reellen Zahlen sind also überabzählbar.

Die reellen Zahlen nehmen eine höhere Unendlichkeitsstufe ein als die Mengen \mathbb{N} , \mathbb{Q} und \mathbb{A} , die sich alle auf der Stufe der abzählbaren Unendlichkeit befinden. Weiter gilt:

\mathbb{R} entsteht nicht aus \mathbb{A} durch Hinzunahme von abzählbar vielen Zahlen.

Denn ist $\mathbb{A} = \{a_0, a_1, a_2, \dots\}$ und $X = \{x_0, x_1, x_2, \dots\}$ eine Menge von „neuen“ Zahlen, so ist die Menge

$$\mathbb{A} \cup X = \{a_0, x_0, a_1, x_1, a_2, x_2, \dots\}$$

abzählbar und damit von \mathbb{R} verschieden. \mathbb{R} wird also nicht dadurch erzeugt, indem wir zu \mathbb{A} eine Handvoll transzendenter Zahlen wie e , π , e^2 , π^2 , e^π , ... hinzunehmen. Die Menge $\mathbb{R} - \mathbb{A}$ der transzendenten Zahlen ist überabzählbar und damit sind „alle bis auf abzählbar viele“ reelle Zahlen transzendent – so schwer es auch ist, einzelne Zahlen als transzendent zu erkennen. Das Kontinuum ist umfassender, komplizierter und geheimnisvoller, als man meinen möchte.

2.4 Die Körperaxiome

Definition (Körper)

Sei K eine Menge, und seien $+, \cdot : K^2 \rightarrow K$. Weiter seien $0, 1 \in K$. Dann heißt $(K, +, \cdot)$ oder kurz K ein *Körper* mit *additiv neutralem Element* 0 und *multiplikativ neutralem Element* 1 , falls für alle $x, y, z \in K$ gilt:

- (K1) $x + (y + z) = (x + y) + z$, (Assoziativgesetz für $+$)
- (K2) $x + 0 = x$, (Neutralität von 0)
- (K3) es gibt ein x' mit $x + x' = 0$, (Inverse für $+$)
- (K4) $x + y = y + x$, (Kommutativgesetz für $+$)
- (K5) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$, (Assoziativgesetz für \cdot)
- (K6) $x \cdot 1 = x$, (Neutralität von 1)
- (K7) falls $x \neq 0$: es gibt ein x' mit $x \cdot x' = 1$, (Inverse für \cdot)
- (K8) $x \cdot y = y \cdot x$, (Kommutativgesetz für \cdot)
- (K9) $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$, (Distributivgesetz)
- (K10) $0 \neq 1$. (Verschiedenheit von 0 und 1)

Ein Körper ist also eine mit einer Addition, einer Multiplikation und zwei speziellen Elementen ausgestattete Menge, sodass die *Körperaxiome* (K1) – (K10) gelten. Wir schreiben wie üblich $x + y$ statt $+(x, y)$ und $x \cdot y$ oder xy statt $\cdot(x, y)$. Der Leser beachte, dass „ $+, \cdot : K^2 \rightarrow K$ “ beinhaltet, dass $x + y, x \cdot y \in K$ für alle $x, y \in K$ gilt. Die Menge K ist also abgeschlossen unter $+$ und \cdot .

+	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

\cdot	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

Beispiele

- (1) \mathbb{Q}, \mathbb{A} und \mathbb{R} bilden einen Körper, \mathbb{Z} dagegen nicht: Es gelten alle Axiome außer (K7): Für $2 \in \mathbb{Z}$ existiert kein $x' \in \mathbb{Z}$ mit $2 \cdot x' = 1$.
- (2) $\{0, 1\}$ bildet einen Körper mit der üblichen Multiplikation und der Addition $0 + 0 = 1 + 1 = 0$, $0 + 1 = 1 + 0 = 1$.
- (3) Ist p eine Primzahl, so bildet $\{0, \dots, p-1\}$ einen Körper, wenn wir modulo p rechnen. Man spricht vom *Restklassenkörper modulo p* . Sei zum Beispiel $p = 7$. Dann gilt $4 + 3 = 0$ und $5 \cdot 3 = 1$, da $4 + 3 = 7$ und $5 \cdot 3 = 15$ bei Division durch 7 den Rest 0 bzw. 1 ergeben. Damit gilt $-4 = 3$ und $5^{-1} = 3$ in diesem Körper. Beispiel (2) entspricht dem Spezialfall $p = 2$.

Additions- und Multiplikationstafel für den Restklassenkörper modulo 7 (Beispiel 3)

Der Körperbegriff lässt sich informal so beschreiben:

K ist ein Körper, falls die vier Grundrechenarten der Addition, Subtraktion, Multiplikation und Division erklärt sind und alle üblichen Rechengesetze gelten.

„Üblich“ ist mit Vorsicht zu lesen, da $1 + 1 = 0$ gelten kann, wie Beispiel (2) zeigt.

Bei dieser Beschreibung tauchen vier Operationen auf, während in der Körperdefinition nur von der Addition und Multiplikation die Rede ist. Die Subtraktion und die Division werden aber durch die Axiome (K3) und (K7) mitdefiniert. Hierzu vereinbaren wir:

Gilt $x + x' = 0$, so schreiben wir $-x$ für x' und nennen $-x$ das *additive Inverse* von x .

Gilt $x \cdot x' = 1$, so schreiben wir x^{-1} für x' und nennen x^{-1} das *multiplikative Inverse* von x .

Nun können wir eine Subtraktion und eine Division auf K einführen:

$$\begin{aligned} x - y &= x + -y \quad \text{für alle } x, y \in K, & (\text{Subtraktion in } K) \\ x / y &= x \cdot y^{-1} \quad \text{für alle } x, y \in K \text{ mit } y \neq 0. & (\text{Division in } K) \end{aligned}$$

Man kann nun die Gesetze für das Bruchrechnen beweisen, etwa

$$\frac{x_1}{y_1} + \frac{x_2}{y_2} = \frac{x_1 y_2 + x_2 y_1}{y_1 y_2} \quad \text{für alle } x_1, x_2, y_1, y_2 \in K \text{ mit } y_1, y_2 \neq 0.$$

Es ist bemerkenswert, dass die gesamte Schularithmetik der Grundrechenarten in zehn Axiomen eingefangen werden kann, die die beiden komplizierteren Operationen der Subtraktion und der Division nur implizit behandeln.

Wir diskutieren noch zwei wichtige spezielle Aspekte.

Die Sonderrolle der Null bei der Division

Für alle $x \in K$ gilt:

$$0 \cdot x = (0 + 0) x = 0 \cdot x + 0 \cdot x \quad \text{für alle } x \in K.$$

Dies zeigt, dass $0 \cdot x = 1$ für kein x gelten kann, da sonst $1 = 1 + 1$ und damit

$$0 = 1 - 1 = (1 + 1) - 1 = 1 + (1 - 1) = 1 + 0 = 1$$

gelten würde. Die Null kann also kein multiplikatives Inverses besitzen, wenn die anderen Axiome gelten sollen.

„Minus mal Minus gleich Plus“

Es gilt $(-1) \cdot (1 - 1) = (-1) \cdot 0 = 0$. Nach dem Distributivgesetz gilt also

$$(-1) \cdot 1 + (-1) \cdot (-1) = 0,$$

und damit

$$(-1) \cdot (-1) = -((-1) \cdot 1) = -(-1) = 1.$$

Allgemeiner gilt für alle $x, y \in K$

$$(-x) \cdot (-y) = (-1) \cdot x \cdot (-1) \cdot y = (-1) \cdot (-1) \cdot x \cdot y = 1 \cdot x \cdot y = x \cdot y.$$

2.5 Die Anordnungsaxiome

Definition (angeordneter Körper)

Sei $(K, +, \cdot)$ ein Körper, und sei $<$ eine Relation auf K . Dann heißt $(K, +, \cdot, <)$ oder kurz K ein *angeordneter Körper*, falls für alle $x, y, z \in K$ gilt:

$$(K11) \quad \text{nicht } x < x, \quad (\text{Irreflexivität})$$

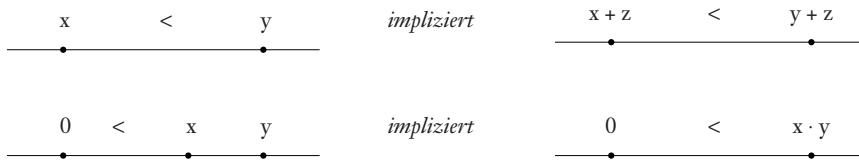
$$(K12) \quad x < y \text{ und } y < z \text{ impliziert } x < z, \quad (\text{Transitivität})$$

$$(K13) \quad x < y \text{ oder } x = y \text{ oder } y < x, \quad (\text{Vergleichbarkeit})$$

$$(K14) \quad x < y \text{ impliziert } x + z < y + z,$$

$$(K15) \quad 0 < x, y \text{ impliziert } 0 < x \cdot y.$$

Gilt $0 < x$, so heißt x ein *positives Element* von K . Gilt $x < 0$, so heißt x ein *negatives Element* von K . Weiter schreiben wir $x \leq y$, falls $x < y$ oder $x = y$.



Ein räumliches oder zeitliches Kontinuum ist vor allem durch ein „Links und Rechts“ bzw. „Früher oder Später“ ausgezeichnet. Diese Ordnungseigenschaften werden für die reellen Zahlen durch die Körperaxiome (K1) – (K10) noch nicht erfasst. Dies ändert sich durch die *Anordnungsaxiome* (K11) – (K15). Sie zerfallen in zwei Gruppen. Die ersten drei Axiome beschreiben grundlegende Eigenschaften einer auf einer Menge K vorhandenen linearen Ordnung. Die beiden anderen Axiome koppeln diese Ordnung an die Arithmetik des Körpers K .

Beispiele

- (1) \mathbb{Q} , \mathbb{A} und \mathbb{R} sind angeordnete Körper.
- (2) Die Restklassenkörper $(\{0, \dots, p-1\}, +, \cdot)$ modulo einer Primzahl p lassen sich nicht anordnen. Aus den Anordnungsaxiomen folgt, wie wir gleich sehen werden, dass $0 < 1$ gilt. Addition von 1 auf beiden Seiten liefert nach (K14), dass

$$1 = 0 + 1 < 1 + 1, \quad 1 + 1 < 1 + 1 + 1 \text{ usw.}$$

Induktiv folgt $0 < 1 + \dots + 1$, was in den Restklassenkörpern nicht gültig ist.

Erreichen wir in einem Körper K durch Aufsummieren der Eins nie die Null, so sagen wir, dass K die *Charakteristik* 0 besitzt. Die Überlegung in (2) zeigt, dass ein angeordneter Körper notwendig die Charakteristik 0 besitzt. Dass es umgekehrt Körper der Charakteristik 0 gibt, die sich nicht anordnen lassen, werden die komplexen Zahlen zeigen.

Wir diskutieren exemplarisch einige Folgerungen aus den Anordnungsaxiomen.

Die Summe positiver Zahlen ist positiv

Seien $x, y \in K$ mit $0 < x$

und $0 < y$. Dann gilt

$$\begin{array}{ccccccc} 0 & < & x & & y & \text{impliziert} & 0 & < & x + y \\ \cdot & & \cdot & & \cdot & & \cdot & & \cdot \end{array}$$

$$0 < y = 0 + y < x + y,$$

wobei wir die zweite Ungleichung aus $0 < x$ und (K14) erhalten.

$0 < 1$ und $-1 < 0$

Annahme, $0 < 1$ gilt nicht. Dann ist

$1 < 0$ nach (K10) und (K13). Somit ist

$$\begin{array}{ccc} -1 & & 0 & & 1 \\ \cdot & & \cdot & & \cdot \end{array}$$

$$0 = 1 - 1 < 0 - 1 = -1$$

nach (K14). Nach (K15) ist dann aber auch $0 < (-1)(-1) = 1$, *Widerspruch*.

Also gilt $0 < 1$. Addition von -1 auf beiden Seiten zeigt, dass $-1 < 0$.

Dass das Produkt zweier positiver Zahlen positiv ist, wird durch (K15) zum Axiom erhoben. Wir können nicht wie bei der Addition stärker fordern, dass die Multiplikation die Ordnung respektiert, dass also mit $x < y$ immer auch $xz < yz$ gilt. Denn wenn wir $0 < 1$ mit -1 multiplizieren, so erhalten wir links 0 und rechts -1 . Wegen $0 > -1$ ändert sich also das Ungleichheitszeichen. Allgemein gilt:

*Ungleichungen bleiben bei Addition und positiver Multiplikation erhalten.
Bei negativer Multiplikation wird aus „Kleiner“ ein „Größer“ und umgekehrt.*

In einem angeordneten Körper K definieren wir für alle $x \in K$:

$$|x| = x, \text{ falls } x \geq 0, \quad |x| = -x, \text{ falls } x < 0.$$

Das Körperelement $|x|$ heißt der *Betrag* von x . Es gilt $|x| = 0$ genau dann, wenn $x = 0$. Für alle x gilt $|x| \geq 0$. Weiter ist stets $|xy| = |x| |y|$. Ständig im Einsatz ist:

$$|x + y| \leq |x| + |y| \text{ für alle } x, y \in K. \quad (\text{Dreiecksungleichung})$$

Der Bezeichnung als „Dreiecksungleichung“ wird klar, wenn wir statt reeller Zahlen Vektoren x und y der Ebene betrachten und den Betrag durch die Länge der Vektoren ersetzen.

Beispiel

Eine häufige Anwendung der Dreiecksungleichung in der Analysis ist das „Einschieben der Null“ in eine Differenz. Sind $x, y \in \mathbb{R}$, so gilt für alle z :

$$\begin{aligned} |x - y| &= |x + 0 - y| = |x + (z - z) - y| = \\ &= |(x - z) + (z - y)| \leq |x - z| + |z - y|. \end{aligned}$$

Der Abstand zwischen x und y ist also beschränkt durch die Summe der Abstände von x und y zu einem Referenzobjekt z . Kann man also die Abstände $|x - z|$ und $|z - y| = |y - z|$ jeweils abschätzen durch ein $\varepsilon > 0$, so gilt also $|x - y| \leq 2\varepsilon$.

2.6 Supremum und Infimum

Definition (*obere Schranke, untere Schranke, Supremum, Infimum*)

Sei K ein angeordneter Körper, und seien $X \subseteq K$ und $s \in K$. Dann heißt s

- (a) eine *obere Schranke* von X , in Zeichen $X \leq s$, falls gilt:

$$x \leq s \quad \text{für alle } x \in X,$$

- (b) das *Supremum* von X , in Zeichen $s = \sup(X)$, falls gilt:

s ist die kleinste obere Schranke von X , d. h. $X \leq s$ und für alle $s' \geq X$ gilt $s \leq s'$,

- (c) eine *untere Schranke* von X , in Zeichen $s \leq X$, falls gilt:

$$s \leq x \quad \text{für alle } x \in X,$$

- (d) das *Infimum* von X , in Zeichen $s = \inf(X)$, falls gilt:

s ist die größte untere Schranke von X , d. h. $s \leq X$ und für alle $s' \leq X$ gilt $s' \leq s$.

Ein $X \subseteq K$ heißt *nach oben (unten) beschränkt*, falls ein s existiert mit $X \leq s$ ($s \leq X$).

Weiter heißt X *beschränkt*, falls X nach oben und unten beschränkt ist.



Mit dem Begriff des angeordneten Körpers haben wir den Unterschied zwischen den reellen Zahlen und den rationalen Zahlen noch nicht erfasst: \mathbb{R} und \mathbb{Q} sind durch die Axiome (K1) – (K15) nicht unterscheidbar. Die Begriffe „Supremum“ und „Infimum“ sind geeignet, dies zu leisten, und wir werden in der nächsten Sektion mit ihrer Hilfe die reellen Zahlen axiomatisch charakterisieren. Hier wollen wir diese beiden Schlüsselbegriffe für sich betrachten, da es Anfängern oft Schwierigkeiten bereitet, ihre Definitionen wiederzugeben. Ein sicherer Umgang mit Suprema und Infima ist in der Analysis sehr wichtig. Und die Mühe der Aneignung lohnt sich, denn viele Argumente lassen sich mit Suprema und Infima einfach und elegant führen.

Für die Analysis spielen die Begriffe der Definition vor allem für $K = \mathbb{Q}$ und $K = \mathbb{R}$ eine Rolle und wir konzentrieren uns im Folgenden auf diese Fälle. Dem Leser wird empfohlen, zunächst nur $K = \mathbb{R}$ zu betrachten und danach $K = \mathbb{Q}$ zur Kontrastbildung. Erst für die axiomatische Charakterisierung von \mathbb{R} in der nächsten Sektion wird die Definition für beliebige angeordnete Körper gebraucht.

Sei also $K = \mathbb{R}$, und sei $X \subseteq \mathbb{R}$ eine beliebige „lineare Punktmenge“. Diese Punktmenge kann zum Beispiel aus den Gliedern $x_0, x_1, \dots, x_n, \dots$ einer Folge bestehen oder aus Intervallen zusammengesetzt sein. Auch jede noch so „chaotische“ und „zerstäubte“ Teilmenge von \mathbb{R} ist zugelassen. Dann bedeutet $X \leq s$ für ein $s \in \mathbb{R}$, dass die reelle Zahl s rechts der Punktmenge liegt, wobei s auch der Menge X angehören kann (ist dies der Fall, so hat X

ein größtes Element und dieses ist gleich s). Stärker als $X \leq s$ bedeutet $s = \sup(X)$, dass der Punkt s rechts der Punktmenge liegt, aber jede Verkleinerung von s nicht mehr rechts der Punktmenge liegt, sondern in die Menge hinein- oder durch sie hindurchführt. Das Supremum von X ist anschaulich diejenige reelle Zahl, die X von rechts „berührt“, das „rechte Ende“ von X oder der „obere Grenzpunkt“ von X .

Beispiele für den Körper der reellen Zahlen

- (1) $[0, 1] \leq 2$, $1 \leq \{x > 2 \mid x \text{ ist rational}\}$, $\{1\} \leq 1$, $2 \leq \emptyset$.
- (2) $\sup([0, 1[) = \sup([0, 1]) = 1 = \inf(]1, 2[) = \inf([1, 2])$.
- (3) $\sup([0, 1] \cup \{2\}) = 2$, $\sup(\{1\}) = \inf(\{1\}) = 1$.
- (4) $\sup(\{\})$, $\inf(\mathbb{Q})$, $\sup(]0, +\infty[)$ existieren nicht, $0 = \inf(]0, +\infty[)$.
- (5) $1 = \sup(\{0,9; 0,99; 0,999; \dots\}) = \inf(\{1,1; 1,01; 1,001; \dots\})$.
- (6) $\sqrt{2} = \sup(\{x \in \mathbb{R} \mid x^2 < 2\}) = \inf(\{x \geq 0 \mid x^2 \geq 2\})$.
- (7) $\pi = \sup(\{3; 3,1; 3,14; 3,141; 3,1415; \dots\})$.

Beispiele für den Körper der rationalen Zahlen

(1) – (5) gelten in \mathbb{Q} genau wie in \mathbb{R} , aber

$$\sup(\{x \in \mathbb{R} \mid x^2 < 2\}), \inf(\{x \geq 0 \mid x^2 \geq 2\}), \sup(\{3; 3,1; 3,14; 3,141; \dots\})$$

existieren nicht in \mathbb{Q} .

Um zu beweisen, dass s das Supremum von X ist, kann man zeigen:

(S1) Sei $x \in X$. Dann gilt $x \leq s$.

(S2) Sei s' derart, dass $x \leq s'$ für alle $x \in X$. Dann gilt $s \leq s'$.

Zum Beweis von (S1) betrachtet man ein beliebiges $x \in X$ und zeigt $x \leq s$. Zum Beweis von (S2) nimmt man dagegen an, dass ein $s' \in K$ vorliegt, das größer oder gleich jedem Element x von X ist. Nun muss man zeigen, dass $s \leq s'$ gilt. Kurz: Man zeigt, dass s eine obere Schranke von X ist und dass s besser als jede andere obere Schranke von X ist.

Eine Variante dieses Beweismusters für „ $s = \sup(X)$ “, die sich sowohl für $K = \mathbb{Q}$ als auch für $K = \mathbb{R}$ eignet, lautet:

(S1) Sei $x \in X$. Dann gilt $x \leq s$.

(S2)* Sei $n \geq 1$. Dann gibt es ein $x \in X$ mit $s - 1/n \leq x$.

Analoge Beweismuster gelten für „ $s = \inf(X)$ “.

2.7 Die Vollständigkeit

Definition (vollständig angeordneter Körper)

Sei K ein angeordneter Körper. Dann heißt K (linear) *vollständig*, falls gilt:

(K16) Ist $X \subseteq K$ nichtleer und nach oben beschränkt, so existiert $\sup(X)$ in K .



Gibt es eine nichtleere und nach oben beschränkte Menge $X \subseteq K$, die kein Supremum in K besitzt, so heißt K *unvollständig*.



Unvollständigkeit: Jede obere Schranke von X lässt sich noch verkleinern.

In einem vollständig angeordneten Körper können wir also „ $s = \sup(X)$ “ immer bilden, abgesehen von der leeren Menge und Mengen, die gar keine obere Schranke besitzen. Automatisch existiert dann auch „ $s = \inf(X)$ “, falls X nichtleer und nach unten beschränkt ist, denn $\inf(X) = \sup(\{s \in K \mid s \leq X\})$.

Beispiele

- (1) \mathbb{R} ist (per Konstruktion oder per Axiom) vollständig.
- (2) \mathbb{Q} ist unvollständig: Die Menge $X = \{q \in \mathbb{Q} \mid q^2 < 2\}$ ist nichtleer und nach oben beschränkt (da z. B. $X \leq 2$), besitzt aber aufgrund der Irrationalität von $\sqrt{2}$ kein Supremum in \mathbb{Q} .
- (3) \mathbb{A} ist unvollständig: Die Menge

$$X = \{3; 3,1; 3,14; 3,141; 3,1415; \dots\}$$

ist nichtleer und nach oben beschränkt (da z. B. $X \leq 4$), besitzt aber aufgrund der Transzendenz von π kein Supremum in \mathbb{A} .

Die Aussage (K16) heißt auch das (lineare) *Vollständigkeitsaxiom*. Es ist das letzte der Axiome in unserer Zusammenstellung. Es gilt nämlich:

Satz (Existenz- und Eindeutigkeitssatz)

Es gibt bis auf Isomorphie genau einen vollständig angeordneten Körper.

Einen vollständig angeordneten Körper nennt man einen *Körper der reellen Zahlen*. Jeder solche Körper kann als „ \mathbb{R} “ dienen. In Einführungen in die Analysis wird oft auf einen Beweis des Existenz- und Eindeutigkeitssatzes verzichtet. Man nimmt \mathbb{R} als gegeben an und notiert:

\mathbb{R} ist mit den Grundrechenarten und einer Ordnung ausgestattet und erfüllt die Axiome (K1) – (K16). Alle weiteren Eigenschaften von \mathbb{R} lassen sich hierauf zurückführen und beruhen nicht auf speziellen Darstellungen oder Konstruktionen der reellen Zahlen.

Die Überabzählbarkeit von \mathbb{R} haben wir mit Hilfe von Dezimaldarstellungen bewiesen. Aber auch Dezimaldarstellungen können aus den Axiomen gewonnen werden (vgl. 2.8), sodass der Beweis doch auf den Axiomen beruht. Man kann alternativ auch direkt zeigen, dass ein Körper der reellen Zahlen überabzählbar sein muss. Das Vollständigkeitsaxiom zwingt uns, mit einer Menge \mathbb{R} zu arbeiten, die viel, viel, viel größer ist als \mathbb{N} , \mathbb{Q} oder \mathbb{A} .

Eine wichtige Konsequenz der Vollständigkeit von \mathbb{R} ist:

Satz (archimedische Anordnung von \mathbb{R})

■ Für alle $x, y > 0$ gibt es ein n mit $nx > y$.

Aus dem Satz folgt, dass $\inf_{n \geq 1} x/n = 0$ für alle $x > 0$ gilt und dass \mathbb{Q} *dicht* in \mathbb{R} ist, d. h., für alle reellen Zahlen $x < y$ existiert eine rationale Zahl q mit $x < q < y$.

Ein positives Element eines angeordneten Körpers, das kleiner als 1, $1/2$, $1/3$, $1/4$, ... ist, heißt *infinitesimal*. Historisch spielten diese Größen eine wichtige Rolle, man denke nur an die Bezeichnung „Infinitesimalrechnung“. Der im 19. Jahrhundert etablierte Aufbau der Analysis auf der Basis von (K1) – (K16) zeigte, dass man die Differential- und Integralrechnung auch ohne infinitesimale Größen entwickeln kann. Heute leben diese Größen in der sog. *Nonstandard-Analysis* weiter, und dort ist eine Infinitesimalrechnung im eigentlichen Sinne möglich.

Es gibt viele äquivalente Varianten des Axioms (K16). Oft wird es durch das sog. *metrische Vollständigkeitsaxiom* (K16)' ersetzt, das besagt, dass jede Cauchy-Folge konvergiert (vgl. 3.5). Man kann dann allerdings die archimedische Anordnung nicht mehr beweisen und muss sie als weiteres Axiom (K17) mit hinzunehmen.

Auch „supremumslose“ Versionen von (K16) sind möglich. Wie üblich sind für $a, b \in K$

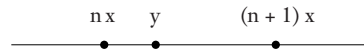
$$]a, b[, [a, b[,]a, b], [a, b]$$

definiert, etwa $[a, b[= \{x \in K \mid a \leq x < b\}$. Allgemein heißt ein $I \subseteq K$ ein *Intervall*, falls für alle $x < y$ in I gilt, dass $[x, y] \subseteq I$. Wir schreiben $X \leq Y$ für $X, Y \subseteq K$, falls $x < y$ für alle $x \in X$ und $y \in Y$. So ist z. B. $[-1, 0] \leq [0, 1]$. Es gilt nun:

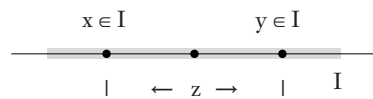
Satz (Varianten des Vollständigkeitsaxioms)

Sei K ein angeordneter Körper. Dann sind die folgenden Aussagen äquivalent:

- (a) (K16)
- (b) Jedes beschränkte Intervall $I \subseteq K$ hat die Form $]a, b[, [a, b[,]a, b]$ oder $[a, b]$.
- (c) Für alle nichtleeren Intervalle $I \leq J$ in K existiert ein s mit $I \leq s \leq J$.



Archimedische Anordnung: Die Vielfachen x , $2x, 3x, \dots, nx, \dots$ eines positiven x überspringen jedes positive y .

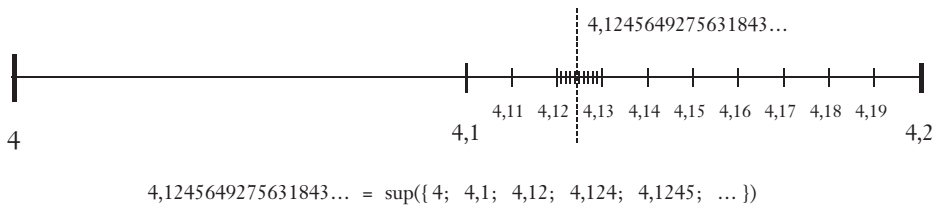


Ein Intervall I in K enthält mit je zwei Punkten $x < y$ auch alle Punkte z zwischen x und y . Dass es keine anderen beschränkten Intervalle als die Intervalle der vier Grundformen $]a, b[, [a, b[,]a, b], [a, b]$ gibt, ist äquivalent zur Vollständigkeit.

2.8 Die Dezimaldarstellung

Definition (Dezimaldarstellung)

- (a) Seien $m \in \mathbb{N}$ und $d_1, d_2, \dots, d_n, \dots$ eine Folge von Zahlen mit $d_i \in \{0, \dots, 9\}$ für alle $n \geq 1$. Dann nennen wir den Ausdruck $m, d_1 d_2 d_3 \dots$ einen *unendlichen Dezimalbruch* mit den *Dezimalziffern* $d_1, d_2, d_3 \dots$
- (b) Jedem Dezimalbruch $m, d_1 d_2 d_3 \dots$ ordnen wir eine reelle Zahl zu, die wir ebenfalls mit $m, d_1 d_2 d_3 \dots$ bezeichnen. Wir setzen hierzu:
- $$m, d_1 d_2 d_3 \dots = \sup(\{m, d_1 \dots d_n \mid n \geq 1\}),$$
- wobei die endlichen Dezimalbrüche $m, d_1 \dots d_n \in \mathbb{R}$ definiert sind durch
- $$m, d_1 \dots d_n = m + d_1/10 + d_2/100 + \dots + d_n/10^n.$$
- (c) Gilt $x = \pm m, d_1 d_2 d_3 \dots$ für eine reelle Zahl x , so nennen wir $\pm m, d_1 d_2 d_3 \dots$ eine *Dezimaldarstellung* oder *Dezimalbruchentwicklung* von x .



Strenger sollte man vielleicht $(m, d_1, d_2, d_3, \dots)$ für Dezimalbrüche und dann weiter $m, d_1 d_2 \dots$ für die durch den Dezimalbruch bezeichnete reelle Zahl schreiben. Die Identifizierung einer Folge mit einer reellen Zahl ist aber in der Regel ungefährlich und einfacher zu handhaben.

In den endlichen Dezimalbrüchen $m, d_1 \dots d_n$ verwenden wir gewisse endliche Summen, die uns auch in \mathbb{Q} und \mathbb{A} zur Verfügung stehen. Ein wesentlicher Einsatz des Vollständigkeitsaxioms erlaubt uns für \mathbb{R} die Definition von

$$m, d_1 d_2 d_3 \dots = \sup(\{m, d_1 \dots d_n \mid n \geq 1\}).$$

Die Menge des Supremums ist nichtleer und nach oben beschränkt durch $m + 1$. Also existiert das Supremum in \mathbb{R} . Wir haben damit unendliche Dezimalbrüche eingeführt, ohne unendliche Summen zu diskutieren. Dies ist aufgrund der Monotonie

$$m \leq m, d_1 \leq m, d_1 d_2 \leq \dots \leq m, d_1 \dots d_n \leq \dots$$

möglich. Kennt man unendliche Summen, so kann man zeigen, dass

$$m, d_1 d_2 d_3 \dots = a + d_1/10 + d_2/100 + \dots + d_n/10^n + \dots$$

Zur Definition und ersten Ergründung der Phänomene der Dezimaldarstellung genügt aber der Supremumsbegriff. Die Summenform diskutieren wir in 4.4.

Beispiele

$$\begin{array}{ll}
0 = 0,000\dots, & 1 = 1,000\dots = 0,999\dots, \\
1/20 = 0,05\dots = 0,04999\dots, & 1/7 = 0,142857142857142857\dots, \\
1/22 = 0,0454545\dots, & \sqrt{2} = 1,4142135356237309\dots, \\
\pi = 3,1415926535897932\dots, & e = 2,7182818284590452\dots
\end{array}$$

Während die Dezimaldarstellung von rationalen Zahlen eher in das Gebiet der elementaren Zahlentheorie fällt, ist es eine Aufgabe der Analysis, Methoden zur Darstellung von $\sqrt{2}$, π , e , $\log(2)$, $\sin(\pi/8)$ usw. zu finden. Dies gelingt unter anderem durch die sog. Potenzreihenentwicklung von Funktionen.

Das bekannte Phänomen der zweideutigen Darstellung lässt sich mit der obigen Supremumsdefinition klar erklären:

Die reelle Zahl $0,999\dots$ ist definiert als das Supremum der Menge

$$X = \{0; 0,9; 0,99; 0,999; \dots\}.$$

Dieses Supremum ist gleich 1. Also gilt $1 = 0,999\dots$

Zur Begründung, warum $\sup(X) = 1$ ist, beobachten wir, dass jeder endliche Dezimalbruch $0,9\dots9$ kleiner als 1 ist. Damit ist 1 eine obere Schranke von X und somit gilt sicher $\sup(X) \leq 1$. Andererseits ist $\sup(X) < 1$ unmöglich, da aufgrund der archimedischen Anordnung von \mathbb{R} die Differenzen $1/10^k = 0,0\dots01 = 1 - 0,9\dots9$ mit wachsender Zahl k der Nachkommastellen kleiner werden als jede positive Zahl. Die Überlegung zeigt auch, dass die Doppeldeutigkeit bereits in \mathbb{Q} auftritt. Die Menge $X \subseteq \mathbb{Q}$ besitzt auch im angeordneten Körper \mathbb{Q} das Supremum 1.

Insgesamt lässt sich zeigen:

Satz *(Ein- und Zweideutigkeit der Dezimaldarstellung)*

Jede irrationale Zahl besitzt genau eine und jede rationale Zahl genau eine oder genau zwei Dezimaldarstellungen. Die rationalen Zahlen sind durch periodische Darstellungen gekennzeichnet und zweideutige Darstellungen treten genau für die Zahlen der Form $n/(2^a 5^b)$, $n \neq 0$, $a, b \geq 0$ auf. Für diese Zahlen enden die beiden Darstellungen in der Periode 0 bzw. 9.

Ob man $0,05 = 0,0500\dots$ oder $0,04999\dots$ bevorzugt, ist Geschmackssache. Während viel für $0,05$ spricht, erbt $0,04999\dots$ von den eindeutigen Darstellungen die Eigenschaft, dass die Darstellung eine streng monoton wachsende Approximation mitliefert. Dagegen hat man bei $0 = 0,000\dots$ keine Wahl.

Wählt man statt $\{0, \dots, 9\}$ die Menge $\{0, \dots, b\}$ für ein $b \geq 1$ als Nachkommaziffern, so liefert der gleiche Ansatz die sog. *b-adischen Entwicklungen*.

Beispiel

Für $b = 2$ (Dualdarstellung, dyadische Entwicklung) gilt

$$1 = 0,111\dots = \sup(\{0; 0,1; 0,11; \dots\}) = \sup(\{0, 1/2, 1/2 + 1/4, \dots\}).$$

2.9 Die Intervallschachtelung

Satz (Prinzip der Intervallschachtelung)

Seien $I_n = [a_n, b_n]$ abgeschlossene Intervalle mit $a_n \leq b_n$ für alle $n \in \mathbb{N}$.

Weiter gelte $I_0 \supseteq I_1 \supseteq \dots \supseteq I_n \supseteq \dots$, d. h., es gelte

$$a_0 \leq a_1 \leq \dots \leq a_n \leq \dots \leq b_n \leq \dots \leq b_1 \leq b_0.$$

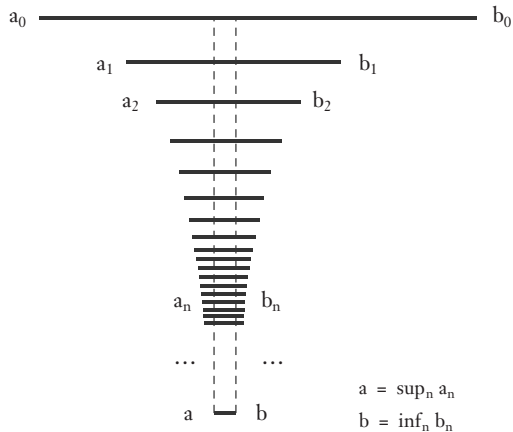
Sei dann

$$I = \bigcap_{n \in \mathbb{N}} I_n = \bigcap \{I_n \mid n \in \mathbb{N}\}.$$

Dann ist I ein Intervall der

Form $[a, b]$ mit $a \leq b$.

Insbesondere ist I nichtleer.



Das Prinzip ist sehr anschaulich und bereitet für sich genommen auch keine größeren Schwierigkeiten. Unsicherheiten bestehen aber zuweilen bei der Verwendung der Durchschnitte. Ist \mathcal{A} eine Menge von Mengen, so ist der „große Durchschnitt“ $\bigcap \mathcal{A}$ erklärt als

$$\bigcap \mathcal{A} = \bigcap_{A \in \mathcal{A}} A = \{x \mid \text{für alle } A \in \mathcal{A} \text{ gilt } x \in A\}.$$

Wir sammeln also alle x auf, die in jedem Element A von \mathcal{A} als Element enthalten sind. In der Situation des Satzes ist damit

$$I = \bigcap_{n \in \mathbb{N}} I_n = \bigcap \{I_n \mid n \in \mathbb{N}\} = \{x \mid \text{für alle } n \text{ gilt } x \in I_n\}$$

die Menge aller Punkte, die jedem Intervall I_n angehören. Es gilt

$$I \subseteq \dots \subseteq I_n \subseteq \dots \subseteq I_2 \subseteq I_1 \subseteq I_0,$$

das Intervall I ist also ein Teilintervall aller Intervalle I_n .

Der Satz besagt, dass das „Zusammenschrumpfen“ eines Intervalls $[a, b]$ nicht im Nichts endet, wenn jedes Schrumpfen ein abgeschlossenes Intervall hinterlässt. Wir betrachten hierzu einige Beispiele und Gegenbeispiele.

Beispiele

$$(1) \bigcap_{n \in \mathbb{N}} [0, 1 + 1/10^n] = [0, 1].$$

$$(2) \bigcap_{n \in \mathbb{N}} [x - 1/2^n, x + 1/2^n] = \{x\} \text{ für alle } x \in \mathbb{R}.$$

$$(3) \bigcap_{n \in \mathbb{N}}]0, 1/2^n[= \emptyset.$$

$$(4) \bigcap_{n \in \mathbb{N}} [n, +\infty[= \emptyset.$$

$$(5) \bigcap_{n \geq 1}]-1/n, 1/n[= \{0\}.$$

Die beiden ersten Beispiele zeigen, dass der Durchschnitt der betrachteten Intervalle von der Form $[a, b]$ mit $a < b$ sein kann, aber auch von der Form $[a, b] = \{a\} = \{b\}$ mit $a = b$. Im zweiten Fall gibt es genau einen Punkt, der in allen Intervallen als Element enthalten ist. Diese Eindeutigkeit taucht in der Praxis häufig auf, etwa dann, wenn die Länge des Intervalls I_{n+1} immer kleinergleich der Hälfte der Länge des Intervalls I_n ist. Allgemein besteht I genau dann aus einem einzigen Punkt, falls für alle $\varepsilon > 0$ ein n existiert, sodass die Länge von I_n (und folglich auch die Länge von I_m für alle $m > n$) kleiner ist als ε .

Die Beispiele (3) und (4) zeigen, dass die Aussage des Satzes im Allgemeinen nicht mehr richtig ist, wenn wir offene oder unbeschränkte abgeschlossene Intervalle verwenden. Die Intervalle

$$]0, 1[, \quad]0, 1/2[, \quad]0, 1/4[, \quad \dots$$

ziehen sich auf die leere Mengen zusammen, und auch die Intervalle

$$[0, +\infty[, \quad [1, +\infty[, \quad [2, +\infty[, \quad \dots$$

verlieren, obwohl sie alle unendlich lang sind, im Laufe der Verkleinerung ebenfalls jeden ihrer Punkte. Natürlich können, wie in Beispiel (5), auch offene geschachtelte Intervalle einen nichtleeren Schnitt besitzen. Aber für die allgemeine Gültigkeit der Aussage ist es wichtig, dass die Intervalle abgeschlossen und beschränkt sind. Intervalle dieses Typs sind auch als *kompakte Intervalle* bekannt. Sie tauchen in verschiedenen prominenten Sätzen der Analysis auf, etwa im Satz, dass jede stetige Funktion auf einem kompakten Intervall ihr Maximum annimmt. Später werden die kompakten Intervalle in der Analysis zu kompakten Mengen verallgemeinert und mit dem Auge der Topologie weiter untersucht.

Mit Hilfe von Suprema und Infima wird der Beweis des Prinzips der Intervallschachtelung sehr einfach. Es gilt

$$I = [\sup(\{a_n \mid n \in \mathbb{N}\}), \inf(\{b_n \mid n \in \mathbb{N}\})],$$

das Durchschnittsintervall ist also definiert durch das Supremum der linken Intervallgrenzen und das Infimum der rechten Intervallgrenzen. Umgekehrt kann man die Intervallschachtelung (zusammen mit der archimedischen Anordnung) als Axiom ansehen und dann das Vollständigkeitsaxiom beweisen. Das Prinzip ist damit für \mathbb{Q} und \mathbb{A} falsch, was zum Beispiel durch Intervalle in \mathbb{Q} bzw. \mathbb{A} , die sich in \mathbb{R} auf die Menge $\{\pi\}$ zusammenziehen, direkt belegt werden kann. So ist $\bigcap_n I_n = \emptyset$, falls

$$I_n = \{x \in \mathbb{A} \mid \pi - 1/2^n \leq x \leq \pi + 1/2^n\} \quad \text{für alle } n.$$

Die Mengen I_n sind keine Intervalle in \mathbb{R} , aber sie sind Intervalle in \mathbb{A} .

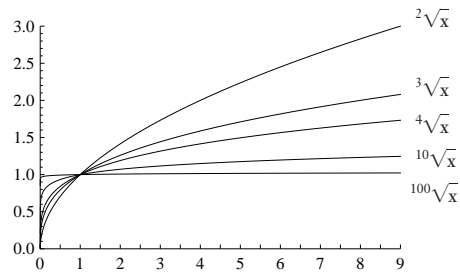
Die Intervallschachtelung lässt sich in der Analysis zum Beispiel verwenden, um den Satz von Bolzano-Weierstraß und den Zwischenwertsatz zu beweisen. Die Intervalle I_n werden dann oft rekursiv definiert. Man konstruiert I_1 mit Hilfe eines gegebenen Intervalls I_0 , I_2 mit Hilfe von I_1 , allgemein I_{n+1} mit Hilfe von I_n . Besonders beliebt ist dabei die fortgesetzte Halbierung. Hier zerlegt man I_n in zwei gleich lange Teilintervalle und wählt als I_{n+1} diejenige Hälfte, die eine gewisse gute Eigenschaft hat. Dann ist das eindeutige x im Durchschnitt aller Intervalle I_n oft ebenfalls ein guter Punkt, etwa ein Häufungspunkt einer Menge oder eine Nullstelle einer Funktion.

2.10 Wurzeln und rationale Exponenten

Definition (*n*-te Wurzel)

Für alle $n \geq 1$ und alle reellen Zahlen $x \geq 0$ heißt das eindeutige $y \geq 0$ mit $y^n = x$ die *n*-te Wurzel von x , in Zeichen $y = \sqrt[n]{x}$.

Für $n = 2$ heißt $\sqrt[2]{x} = \sqrt{x}$ auch die (positive) *Quadratwurzel* von x .



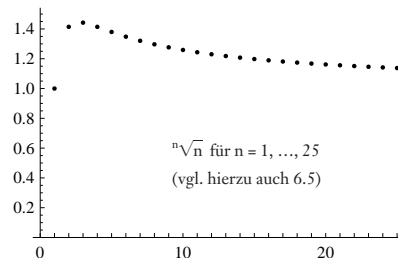
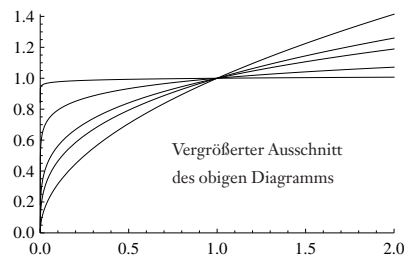
Beispiele

$$(1) \sqrt[n]{0} = 0, \sqrt[n]{1} = 1, \sqrt[n]{x^n} = x.$$

$$(2) \sqrt[2]{0,1} \sim 0,32, \sqrt[3]{0,1} \sim 0,46, \\ \sqrt[4]{0,1} \sim 0,56, \sqrt[10]{0,1} \sim 0,79.$$

$$(3) \sqrt[2]{1,1} \sim 1,05, \sqrt[3]{1,1} \sim 1,03, \\ \sqrt[4]{1,1} \sim 1,02, \sqrt[10]{1,1} \sim 1,01.$$

$$(4) \text{ Ist } w = \sqrt[3]{x}, \text{ so ist } \sqrt[4]{w} = \sqrt[12]{x} \\ \text{ und } w^2 = \sqrt[3]{x^2}.$$



Die obige Definition beruht auf:

Satz (Existenz von Wurzeln)

Für alle $n \geq 1$ und alle $x \geq 0$ existiert ein eindeutiges $y \geq 0$ mit $y^n = x$.

Die Existenz von Wurzeln für positive reelle Zahlen ist eine Konsequenz des Vollständigkeitsaxioms. Hierzu eine Überlegung:

Die auf die nichtnegative Halbachse eingeschränkte n-te Potenz $f: [0, +\infty[\rightarrow \mathbb{R}, f(x) = x^n$ für alle $x \geq 0$, ist streng monoton steigend und damit insbesondere injektiv. Also besitzt f eine Umkehrfunktion $g: [0, +\infty[\rightarrow \mathbb{R}$ und g ist dann offenbar die n-te Wurzelfunktion. Was ist hier überhaupt noch zu beweisen?

Wir spiegeln die n -te Potenz f auf $[0, +\infty[$ an der Winkelhalbierenden und erhalten die n -te Wurzel g . Das ist richtig, aber das Argument enthält eine Lücke. Es gilt $g: \text{Bild}(f) \rightarrow \mathbb{R}$, aber es ist nicht klar, dass $\text{Bild}(f)$ das Intervall $[0, +\infty[$ ist. Dies ist gerade zu zeigen! Richtig ist, dass der Zwischenwertsatz die Existenz von Wurzeln als Geschenk mitbringt. Wissen wir, dass eine auf einem Intervall definierte stetige Funktion, die zwei verschiedene Werte annimmt, auch jeden Wert zwischen diesen beiden Werten annimmt, so ist $\text{Bild}(f) = [0, +\infty[$, da f die Werte $0^n, 1^n, 2^n, 3^n, \dots$ annimmt. Damit ist $[0, +\infty[$ der Definitionsbereich von g . (Den Zwischenwertsatz besprechen wir in 5.8. Einen direkteren Beweis der Existenz von Wurzeln skizzieren wir in 3.4.)

Die Potenzfunktionen mit ungeraden Exponenten bilden \mathbb{R} bijektiv auf \mathbb{R} ab, und wir können so die dritte, fünfte, siebte etc. Wurzelfunktion nicht nur auf $[0, +\infty[$, sondern sogar auf ganz \mathbb{R} erklären und zum Beispiel $\sqrt[5]{-32} = -2$ schreiben. Im Allgemeinen werden in der Analysis aber die Wurzelfunktionen auf $[0, +\infty[$ oder sogar nur auf $]0, +\infty[$ betrachtet. Dies führt zu allgemeingültigen Rechenregeln. Wir können zum Beispiel für $w = \sqrt[5]{x}$ nicht $\sqrt[2]{w} = \sqrt[10]{x}$ rechnen, wenn x negativ ist.

Beim Umgang mit Wurzeln entdeckt man, dass das Wurzelziehen die Rechenregeln für die Exponentiation respektiert, wenn wir die n -te Wurzel als Exponent $1/n$ lesen. So gilt etwa $\sqrt{x} \cdot \sqrt{x} = x$, was $x^{1/2} \cdot x^{1/2} = x^{1/2 + 1/2} = x^1 = x$ entspricht. Wir definieren:

Definition (*Potenzfunktion für rationale Exponenten*)

Für alle reellen Zahlen $a > 0$ und alle $q = m/n \in \mathbb{Q}$ mit $m \in \mathbb{Z}$ und $n \geq 1$ setzen wir:

$$a^q = \sqrt[n]{a^m}.$$

Ist $m > 0$, so setzen wir weiter $0^q = 0$.

Diese Definition respektiert das Kürzen von Brüchen und ist also wohldefiniert. Es gelten die Rechengesetze der Exponentiation:

$$a^{-q} = 1/a^q, \quad a^q a^r = a^{q+r}, \quad (a^q)^r = a^{qr} \quad \text{für alle } a > 0 \text{ und alle } q, r \in \mathbb{Q}.$$

Zusammenfassend halten wir fest:

*In jedem Körper lässt sich eine Exponentiation für ganzzahlige Exponenten definieren.
Die Vollständigkeit von \mathbb{R} ermöglicht eine Exponentiation für rationale Exponenten.*

Neugierig fragen wir:

Kann man nicht auch beliebige reelle Exponenten zulassen?

Dies ist in der Tat möglich, und es stehen zwei verschiedene Wege hierzu offen:

Erste Möglichkeit: Stetige Fortsetzung

Definieren wir $f_a : \mathbb{Q} \rightarrow \mathbb{R}$ durch $f_a(q) = a^q$ für alle $q \in \mathbb{Q}$, so ist f_a streng monoton steigend für $a > 1$, konstant gleich 1 für $a = 1$ und streng monoton fallend für $0 < a < 1$. Dies motiviert die folgende Definition von a^x für alle $a > 0$ und $x \in \mathbb{R}$:

$$a^x = \begin{cases} \sup(\{a^q \mid q < x\}) & \text{falls } a > 1, \\ 1 & \text{falls } a = 1, \\ \inf(\{a^q \mid q < x\}) & \text{falls } 0 < a < 1. \end{cases}$$

Zweite Möglichkeit: Einsatz der Exponentialfunktion zur Basis e

Mit Hilfe der Exponentialfunktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$ und des Logarithmus $\log :]0, +\infty[\rightarrow \mathbb{R}$ zur Basis e definiert man $a^x = \exp(x \log(a))$ für alle $a > 0$ und $x \in \mathbb{R}$.

Beide Wege sind äquivalent. Wir besprechen sie in 5.7 und 6.5 noch genauer. Oft wird heute aufgrund der überragenden Bedeutung der Exponentialfunktion in der Mathematik und den Naturwissenschaften der zweite Weg beschritten.

2.11 Komplexe Zahlen

Definition (*komplexe Zahlen, Realteil, Imaginärteil, imaginäre Einheit*)

Wir setzen $\mathbb{C} = \mathbb{R}^2$ und nennen jedes Element von \mathbb{C} eine *komplexe Zahl*.

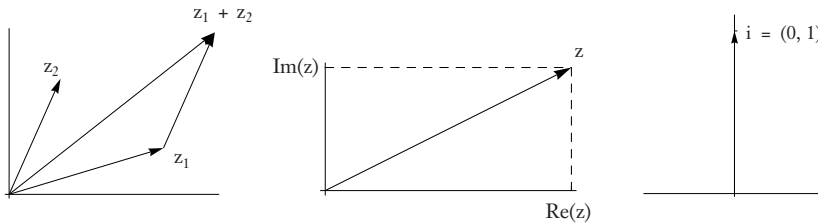
Für alle $(x_1, y_1), (x_2, y_2) \in \mathbb{C}$ definieren wir:

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

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2). \quad (\text{komplexe Multiplikation})$$

Ist $z = (x, y) \in \mathbb{C}$, so heißen $\operatorname{Re}(z) = x$ der *Real-* und $\operatorname{Im}(z) = y$ der *Imaginärteil* von z .

Weiter sei $i = (0, 1)$. Die komplexe Zahl i heißt die *imaginäre Einheit*.



Eine geometrische Deutung der komplexen Multiplikation entwickeln wir unten und in 12.2.

Die komplexen Zahlen sind aus historischen Gründen geheimnisumwittert. Das traditionelle Adjektiv „imaginär“ suggeriert, dass diese Zahlen im Gegensatz zu den reellen Zahlen nicht wirklich existieren, sondern Erfindungen unseres Geistes sind. Aus heutiger Sicht ist der Realitätsunterschied zwischen \mathbb{R} und $\mathbb{C} = \mathbb{R}^2$ verschwindend gering. Die komplexen Zahlen sind geordnete Paare reeller Zahlen, die wir als Punkte der Ebene vor uns sehen können. Sie lassen sich durch eine natürliche Fragestellung motivieren, die von einer Aufarbeitung der Geschichte der komplexen Zahlen unabhängig ist:

Kann man auf der Ebene \mathbb{R}^2 eine Addition und eine Multiplikation so definieren, dass $(\mathbb{R}^2, +, \cdot)$ ein Körper ist? Ist dies auch für $\mathbb{R}^3, \mathbb{R}^4, \mathbb{R}^5, \dots$ möglich?

Dass wir Vektoren der Ebene addieren können, ist gut bekannt, und die komplexe Addition ist nichts anderes als die übliche Vektoraddition. Sie erfüllt die ersten vier Körperaxiome mit additiv neutralem Element $0 = (0, 0)$. Die komplexe Multiplikation ist dagegen trickreicher, aber sie leistet, was wir wollen:

Satz (*Körperstruktur von \mathbb{C}*)

I $(\mathbb{C}, +, \cdot) = (\mathbb{R}^2, +, \cdot)$ ist ein Körper mit $0 = (0, 0)$ und $1 = (1, 0)$.

Die Antwort auf die Frage ist also „ja“ für \mathbb{R}^2 . Dagegen ist sie „nein“ für alle $n \geq 3$: Es ist nicht mehr möglich, auf $\mathbb{R}^3, \mathbb{R}^4, \dots$ eine Körperstruktur zu erklären. Die Ebene ist in diesem Sinn etwas besonderes!

Indem wir die x -Achse der Ebene mit \mathbb{R} identifizieren, können wir $\mathbb{R} \subseteq \mathbb{C}$ annehmen. Die komplexe Arithmetik ist dann eine Fortsetzung der reellen Arithmetik, da z. B.

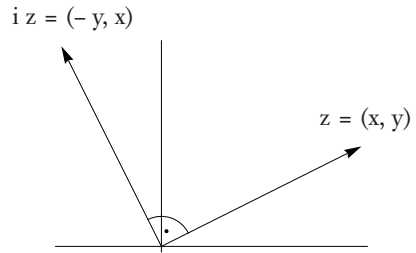
$$x y = (x, 0) \cdot (y, 0) = (xy - 0 \cdot 0, x \cdot 0 + 0 y) = (x y, 0) = x y.$$

Erst der algebraische Kalkül von \mathbb{C} macht die Vektoren der Ebene zu gefühlten Zahlen. Er wird von der imaginären Einheit dominiert:

Die Rolle der imaginären Einheit $i = (0, 1)$

Für alle $(x, y) \in \mathbb{C}$ gilt:

$$\begin{aligned} i \cdot (x, y) &= (0, 1) \cdot (x, y) = \\ (0x - 1y, 0y + 1x) &= (-y, x). \end{aligned}$$



Geometrisch gelesen bedeutet dies:

Die Multiplikation von (x, y) mit der imaginären Einheit i ist die Drehung des Vektors (x, y) um 90 Grad gegen den Uhrzeigersinn.

Speziell ist

$$i^2 = i \cdot i = i \cdot (0, 1) = (-1, 0) = -1 \in \mathbb{R}.$$

Weiter erhalten wir für alle $z = (x, y) \in \mathbb{C}$ die fundamentale Darstellung

$$z = (x, y) = (x, 0) + (0, y) = (x, 0) + i(y, 0) = x + iy = \operatorname{Re}(z) + i \operatorname{Im}(z).$$

Die komplexe Multiplikation lässt sich mit diesen Formeln rekonstruieren:

$$\begin{aligned} (x_1, y_1) (x_2, y_2) &= (x_1 + i y_1) (x_2 + i y_2) = \\ x_1 x_2 + i^2 y_1 y_2 + i (x_1 y_2 + x_2 y_1) &= (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2). \end{aligned}$$

Wegen $i^2 = -1$ gilt, wenn man die Wurzelschreibweise verwenden möchte, die Identität $i = \sqrt{-1}$. Sie ist nicht nur berühmt, sondern auch berüchtigt, da sie zu Fehlern wie

$$„-1 = i^2 = \sqrt{-1} \sqrt{-1} = \sqrt{(-1)(-1)} = \sqrt{1} = 1“$$

führen kann. Ohne jede Wurzelmythik sind aber i und $-i$ untadelige komplexe Lösungen der algebraischen Gleichung $z^2 + 1 = 0$. Allgemein gilt:

Satz (Fundamentalsatz der Algebra)

Sei $a_k z^k + \dots + a_1 z + a_0 = 0$ eine algebraische Gleichung k -ten Grades mit komplexen Koeffizienten. Dann gibt es komplexe Zahlen w_1, \dots, w_k , sodass

$$a_k z^k + \dots + a_1 z + a_0 = a_k (z - w_1) \cdot (z - w_2) \cdot \dots \cdot (z - w_k) \quad \text{für alle } z \in \mathbb{C}.$$

Die Gleichung hat also genau die komplexen Lösungen w_1, \dots, w_k .

Im Zahlensystem $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ beobachten wir also vier Erweiterungsschritte:

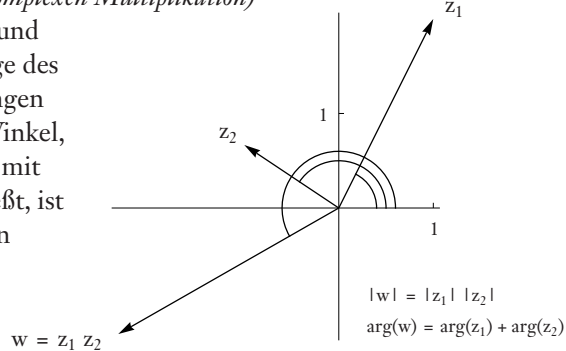
freie Subtraktion, freie Division, Vollständigkeit, Lösbarkeit von Gleichungen.

Beim Schritt von \mathbb{R} nach \mathbb{C} geht aber auch etwas verloren: Der Körper der komplexen Zahlen lässt sich nicht anordnen. Denn in jedem angeordneten Körper gilt $x^2 \geq 0$ für alle x , und dies ist in \mathbb{C} für $i = (0, 1)$ verletzt. In \mathbb{C} ist also ein „Kleiner“ und „Größer“ nicht mehr vorhanden. Das ist für Objekte, die wir Zahlen nennen, sicher gewöhnungsbedürftig.

2.12 Umgang mit komplexen Zahlen

Satz (geometrische Deutung der komplexen Multiplikation)

Seien z_1, z_2 komplexe Zahlen, und sei $w = z_1 z_2$. Dann ist die Länge des Vektors w das Produkt der Längen der Vektoren z_1 und z_2 . Der Winkel, den w im Gegenuhrzeigersinn mit der positiven x-Achse einschließt, ist die Summe der entsprechenden Winkel von z_1 und z_2 .



Kurz:

*Zwei komplexe Zahlen werden multipliziert,
indem man ihre Längen multipliziert und ihre Winkel addiert.*

Mit dieser geometrischen Multiplikationsregel haben wir der Vektoraddition eine ebenso anschauliche Deutung der komplexen Multiplikation an die Seite gestellt. Sie ist beim Umgang mit komplexen Zahlen äußerst hilfreich und erlaubt es, vergessene Formeln für komplexe Zahlen leicht rekonstruieren zu können. In einem systematischen Aufbau der Analysis wird diese Regel oft erst recht spät etabliert, da man dort Winkel nicht voraussetzen, sondern mit analytischen Methoden definieren möchte. Insofern erfordert die Verwendung der Multiplikationsregel an dieser Stelle eine „Winkelanleihe“ bei der Geometrie. Man darf sie, wenn man der rein analytischen Einführung geometrischer Grundbegriffe folgen möchte, vorerst nicht in Beweisen verwenden. Aber kennen und heuristisch verwenden darf man sie auch dann!

Bevor wir Beispiele für den Einsatz der Regel diskutieren, wollen wir die Begriffe „Länge“ und „Winkel“ noch etwas genauer betrachten.

Die *Länge* oder der *Betrag* $|z|$ einer komplexen Zahl $z = (x, y) \in \mathbb{R}^2$ ist die euklidische Länge von z :

$$|z| = \sqrt{x^2 + y^2}. \quad (\text{Länge oder Betrag von } z)$$

Es gilt also

$$|0| = 0, \quad |i| = 1, \quad |(1, 1)| = |1 + i| = \sqrt{2}, \quad \text{usw.}$$

Den *Winkel* oder das *Argument* $\arg(z)$ einer komplexen Zahl $z \neq 0$ übernehmen wir hier aus der Geometrie. Wir messen Winkel im Bogenmaß und vereinbaren $\arg(z) \in [0, 2\pi[$. Wir rechnen oft stillschweigend modulo 2π , sodass der Winkel $3/2\pi$ gleich dem Winkel $-\pi/2$ ist. Es gilt

$$\arg(1) = 0, \quad \arg(i) = \pi/2, \quad \arg(-1) = \pi, \quad \arg(-i) = 3/4\pi.$$

Der komplexen Zahl 0 wird in der Regel kein Winkel zugeordnet, aber sie bereitet auch keinerlei Schwierigkeiten bei der Multiplikation, da $z \cdot 0 = 0$ für alle z gilt. Wir können sie also hier außer Acht lassen oder $\arg(0) = 0$ setzen.

Beispiele

- (1) Oben hatten wir gesehen, dass iz einer Drehung von z um $\pi/2$ entspricht. Dies ist ein Spezialfall der allgemeinen Regel, denn es gilt $|i| = 1$ und $\arg(i) = \pi/2$.

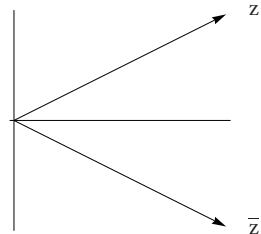
- (2) Für alle $z = (x, y) = x + iy$ heißt

$$\bar{z} = (x, -y) = x - iy = \operatorname{Re}(z) - i\operatorname{Im}(z)$$

die *komplex Konjugierte* von z . Geometrisch entspricht die Konjugation der Spiegelung des Vektors z an der x -Achse. Es gilt

$$z\bar{z} = |z|^2 \in [0, +\infty[. \text{ (Längenregel für } \mathbb{C} \text{)}$$

Denn es gilt $|z| = |\bar{z}|$, sodass $|z\bar{z}| = |z| |\bar{z}| = |z| |z| = |z|^2$. Weiter ist $\arg(\bar{z}) = 2\pi - \arg(z)$, also $\arg(z\bar{z}) = \arg(z) + \arg(\bar{z}) = 0$ modulo 2π .



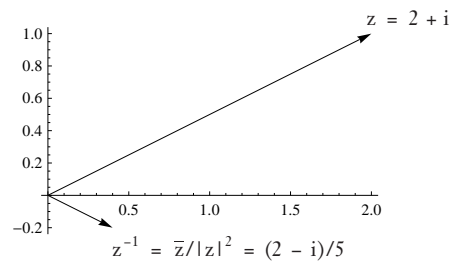
- (3) Für alle komplexen Zahlen $z \neq 0$ gilt

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2}. \text{ (Inversenregel für } \mathbb{C} \text{)}$$

Denn für $w = \bar{z}/|z|$ gilt

$$|w| = |\bar{z}|/|z| = 1/|z|,$$

$$\arg(w) = 2\pi - \arg(z).$$



Damit hat zw Länge 1 und Argument 0. Also ist $wz = 1$, d. h. $w = 1/z$.

- (4) Ähnlich wie in (2) und (3) kann man sehen, dass für alle $z \in \mathbb{C}$ gilt:

$$\operatorname{Re}(z) = (z + \bar{z})/2, \quad \operatorname{Im}(z) = (-i)(z - \bar{z})/2 = (z - \bar{z})/(2i).$$

- (5) Sei $w = (1, 1)/\sqrt{2}$. Dann hat w die Länge 1 und den Winkel $\pi/4$. Folglich ist $w^2 = i$ und also eine Lösung der Gleichung $z^2 - i = 0$.

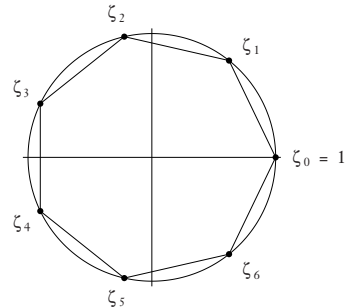
- (6) Sei $n \geq 1$ und seien $\zeta_0, \dots, \zeta_{n-1}$ die Vektoren des Einheitskreises mit den Winkeln $0, 1/n \cdot 2\pi, 2/n \cdot 2\pi, \dots, (n-1)/n \cdot 2\pi$. Dann gilt

$$\zeta_k^n = 1 \quad \text{für alle } 0 \leq k \leq n-1,$$

denn $|\zeta_k^n| = |\zeta_k|^n = 1^n = 1$ und modulo 2π ist

$$\arg(\zeta_k^n) = n \cdot \arg(\zeta_k) = n \cdot k/n \cdot 2\pi = k2\pi.$$

Die ζ_k lösen also die Gleichung $z^n = 1$. Da diese Gleichung höchstens n Lösungen hat, sind die ζ_k alle Lösungen der Gleichung. Man nennt $\zeta_0, \dots, \zeta_{n-1}$ die *n-ten Einheitswurzeln*. Geometrisch bilden sie ein gleichseitiges n -Eck und führen einen Spezialfall des Fundamentalsatzes vor Augen.



Die siebten Einheitswurzeln

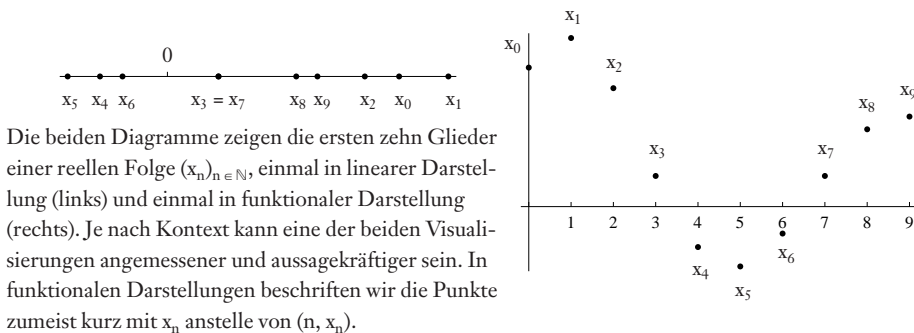
3. Kapitel

Folgen und Grenzwerte

3.1 Folgen

Definition (Folge, Notationen für Folgen)

- (a) Eine Funktion f auf \mathbb{N} nennen wir auch eine (unendliche) Folge (auf \mathbb{N}). Folgen notieren wir auch in den Folgennotationen
- $$(x_n)_{n \in \mathbb{N}}, (x_n)_{n \geq 0}, \langle x_n \mid n \in \mathbb{N} \rangle, (x_0, x_1, \dots, x_n, \dots), x_0, x_1, \dots, x_n, \dots$$
- Diese Ausdrücke bedeuten die Funktion f auf \mathbb{N} mit $f(n) = x_n$ für alle n .
- (b) Für alle n heißt $f(n) = x_n$ das n -te Glied oder das Glied der Folge mit Index n .
- (c) Eine Folge $(x_n)_{n \in \mathbb{N}}$ heißt eine Folge in einer Menge M , falls $x_n \in M$ für alle n . Eine Folge in \mathbb{R} nennen wir auch eine reellwertige Folge, eine Folge in \mathbb{C} auch eine komplexwertige Folge.



Beispiele

- (1) Für alle $c \in \mathbb{R}$ ist (c, c, c, \dots) eine Folge in \mathbb{R} . Diese Folge ist die Funktion $f: \mathbb{N} \rightarrow \mathbb{R}$ mit $f(n) = c$ für alle n . Sie heißt die *konstante Folge* mit Wert c .
- (2) $(0, 1, 2, 3, 4, \dots)$ ist eine Folge in \mathbb{N} , aber auch eine Folge in \mathbb{R} oder \mathbb{C} .
- (3) $(0, 1, 0, 1, 0, 1, 0, 1, \dots)$ ist eine Folge, die nur zwei verschiedene Werte annimmt. Dennoch hat diese Folge, wie jede Folge, unendlich viele Glieder x_n . Es gilt $x_n = 0$ für gerade und $x_n = 1$ für ungerade n .

Auch Varianten wie $(x_n)_{n \geq 1}$ sind gebräuchlich. Der Definitionsbereich ist hier nicht \mathbb{N} , sondern $\mathbb{N} - \{0\}$. So ist zum Beispiel $(1/n)_{n \geq 1}$ die Funktion f auf $\mathbb{N} - \{0\}$ mit

$$f(n) = 1/n \text{ für alle } n \geq 1.$$

Notieren wir diese Folge in der Form $(1, 1/2, 1/3, 1/4, \dots)$, so ist der Definitionsbereich unklar, denn es könnte auch die Funktion g auf \mathbb{N} mit $g(n) = 1/(n+1)$ für alle $n \geq 0$ gemeint sein. Dieser feine Unterschied ist aber in der Regel unproblematisch.

Beispiel

Ist $x = \pm k, d_1 d_2 d_3 \dots$ eine reelle Zahl in Dezimaldarstellung, so ist $(d_n)_{n \geq 1}$ die Folge der Dezimalziffern der Darstellung. Diese Folge ist eine Folge in $\{0, \dots, 9\}$.

Obwohl wir Folgen in M als Funktionen $f: \mathbb{N} \rightarrow M$ angeben können, so hat doch die Folgennotation $(x_n)_{n \in \mathbb{N}}$ viele Vorteile. Sie erlaubt es, Folgen, deren Glieder durch einen Term definiert sind, leicht anzugeben. Wir betrachten hierzu einige Beispiele.

Beispiele

(1) $(1/2^n)_{n \in \mathbb{N}} = (1, 1/2, 1/4, \dots)$, $((-1)^n/n)_{n \geq 1} = (-1, 1/2, -1/3, 1/4, \dots)$.

(2) Ist $g: \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion und $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} , so ist

$$(g(x_n))_{n \in \mathbb{N}} = (g(x_0), g(x_1), g(x_2), \dots).$$

Funktional gedacht ist $(g(x_n))_{n \in \mathbb{N}}$ die Verknüpfung $g \circ (x_n)_{n \in \mathbb{N}}$.

(3) $(\sin(n\pi))_{n \in \mathbb{N}} = (0, 0, 0, \dots)$.

Dass Folgen ihrer Natur nach Funktionen sind, hat aber ebenfalls Vorteile. Für eine Folge $(x_n)_{n \in \mathbb{N}}$ auf \mathbb{R} sind zum Beispiel, wie für jede Funktion $f: A \rightarrow \mathbb{R}$ mit $A \subseteq \mathbb{R}$, die Wachstumsbegriffe erklärt (vgl. 1.4):

Eine Folge $(x_n)_{n \in \mathbb{N}}$ heißt ...	falls ...
<i>monoton steigend</i>	$\forall n \ x_n \leq x_{n+1}$
<i>streng monoton fallend</i>	$\forall n \ x_n > x_{n+1}$
<i>nach oben beschränkt</i>	$\exists k \ \forall n \ x_n \leq k$
<i>beschränkt</i>	$\exists k \ \forall n \ x_n \leq k$

Beispiele

(1) $(0, 0, 1, 1, 2, 2, 3, 3, \dots)$ ist monoton steigend, nicht streng monoton steigend, unbeschränkt und nach unten beschränkt,

$(1, 1/2, 1/3, 1/4, \dots)$ ist streng monoton fallend und beschränkt,

$(0, 2, 1, 4, 3, 6, 5, \dots)$ ist weder monoton noch beschränkt.

(2) Ist $x = k, d_1 d_2 d_3 \dots$ in Dezimaldarstellung, so ist die Folge $(k, d_1 \dots d_n)_{n \geq 1}$ der n -ten Approximationen monoton steigend.

Für Folgen in \mathbb{C} gibt es mangels einer Ordnung in \mathbb{C} keine Monotoniebegriffe. Jedoch können wir die Beschränktheit über die komplexe Betragsfunktion definieren. Eine Folge $(z_n)_{n \in \mathbb{N}}$ heißt beschränkt, falls es ein $k \in \mathbb{N}$ gibt mit $|z_n| \leq k$ für alle n . Die Folgenglieder liegen dann alle in einer Kreisscheibe.

Beispiele

(1) $(i^n)_{n \in \mathbb{N}} = (1, i, -1, -i, 1, i, -1, -i, \dots)$ ist beschränkt.

(2) $(ni)_{n \in \mathbb{N}} = (0, i, 2i, 3i, \dots)$ ist unbeschränkt.

3.2 Grenzwerte von Folgen

Definition (Konvergenz von Folgen, Grenzwert, Limes, konvergent, divergent)

Eine Folge $(x_n)_{n \in \mathbb{N}}$ in \mathbb{R} konvergiert gegen ein $x \in \mathbb{R}$, falls gilt:

$$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 |x_n - x| < \varepsilon. \quad (\text{Konvergenzbedingung für } x)$$

Wir schreiben dann

$$\lim_{n \rightarrow \infty} x_n = x \quad \text{oder kurz} \quad \lim_n x_n = x, \quad (\text{Limesnotation})$$

und nennen x den *Grenzwert* oder *Limes* der Folge $(x_n)_{n \in \mathbb{N}}$.

Konvergiert die Folge nicht, so heißt sie *divergent*.

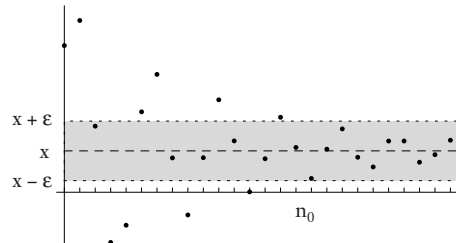
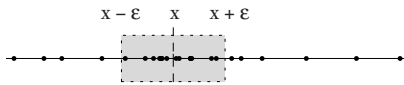


Diagramme zu „ $x = \lim_n x_n$ “. Für gegebenes $\varepsilon > 0$ liegen alle Folgenglieder ab einer Stelle n_0 im grauen Bereich. Im linken Diagramm ist dies ein ε -Intervall um x , im rechten ein ε -Streifen um x .

Der Grenzwertbegriff für Folgen spielt eine fundamentale Rolle in der Analysis, da er sich zur Definition vieler anderer Spielarten des Grenzwertbegriffs eignet. Es ist zum Beispiel heute üblich, die aus der Schule bekannten Grenzwerte für Funktionen, also etwa

$$\lim_{x \rightarrow 1, x \neq 1} \frac{x^2 - 1}{x - 1} = 2,$$

auf Grenzwerte von Folgen zurückzuführen (wie das geht, werden wir in 5.2 besprechen). Da die Ableitung einer Funktion in einem Punkt als Grenzwert einer Funktion eingeführt wird, beruht die Differentialrechnung auf dem Grenzwertbegriff für Folgen. Er spielt weiter eine herausragende Rolle in der Topologie und der Wahrscheinlichkeitstheorie.

Bemerkung

Das Unendlichkeitszeichen in der Limesnotation $\lim_{n \rightarrow \infty} x_n = x$ ist rein symbolisch. Die Konvergenzbedingung ist gerade dazu da, um die anschauliche Formulierung

„ x_n strebt gegen x , wenn n gegen unendlich strebt“

zu präzisieren. Die Situation ist insgesamt vergleichbar mit der symbolischen Intervallnotation $[0, +\infty[= \{x \in \mathbb{R} \mid x \geq 0\}$. In 3.12 werden wir Grenzwerte der Form $\lim_n x_n = \infty$ und $\lim_n x_n = -\infty$ einführen. Bis dahin ist $x \in \mathbb{R}$, falls $\lim_n x_n = x$.

Die Grenzwertdefinition ist ein erstes Beispiel für die vor allem von Karl Weierstraß im 19. Jahrhundert propagierte „Epsilonontik“, durch die die Grundbegriffe der Analysis präzisiert werden konnten. Diese Epsilonontik bereitet vielen Anfängern große Schwierigkeiten. Wir wollen deswegen im Detail betrachten, wie ein Nachweis der Konvergenzbedingung typischerweise verläuft. Danach betrachten wir Beispiele für Grenzwerte.

Wie beweist man $\lim_{n \rightarrow \infty} x_n = x$?

Wir beginnen mit „Sei $\varepsilon > 0$.“ (Dies steht so oft an der Tafel, dass „ $\varepsilon \leq 0$ “ als Insiderwitz gilt.) Zu diesem ε ist nun ein „guter“ Index n_0 zu finden, der in der Regel von ε abhängt. Gut heißt, dass ab der Stelle n_0 die Folgenglieder x_n von x höchstens den Abstand ε haben, d. h.:

$$(+)\text{ Für alle } n \geq n_0 \text{ gilt } |x_n - x| < \varepsilon.$$

Ein gutes n_0 findet man durch genaues Studium der Folge, nicht selten nach einigen Fehlversuchen und durch Rückwärtsargumentation. (Die Zahl n_0 ist nicht eindeutig bestimmt: Ist n_0 gut, so ist auch jedes $n_1 > n_0$ gut.) Ist n_0 festgelegt, so besteht der Rest des Beweises im Nachweis von (+). Gelingt dies, ist der Beweis fertig.

Viele Varianten sind möglich. Gleichwertig zur Konvergenzbedingung für x sind z. B.:

$$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 |x_n - x| \leq \varepsilon,$$

$$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 |x_n - x| < 2\varepsilon,$$

$$\forall \varepsilon > 0 \exists n_0 \forall n \geq 2n_0 |x_n - x| < \varepsilon.$$

Auch ein Nachweis einer dieser Aussagen zeigt also, dass $\lim_{n \rightarrow \infty} x_n = x$.

Will man dagegen zeigen, dass eine Folge $(x_n)_{n \in \mathbb{N}}$ divergiert, so ist zu zeigen, dass die Konvergenzbedingung für jedes x verletzt ist. Nach den Verneinungsregeln für die Quantoren ist also für jedes $x \in \mathbb{R}$ zu zeigen:

$$\exists \varepsilon > 0 \forall n_0 \exists n \geq n_0 |x_n - x| \geq \varepsilon. \quad (\text{Divergenzbedingung für } x)$$

Für ein beliebiges $x \in \mathbb{R}$ ist ein (in der Regel von x abhängiges) $\varepsilon > 0$ zu finden, sodass für unendlich viele n der Abstand $|x_n - x|$ zwischen x_n und x größergleich ε ist.

Beispiele

- (1) $\lim_n c = c$. Für alle $\varepsilon > 0$ ist $n_0 = 0$ ein guter Index, da für alle n gilt:

$$|x_n - c| = |c - c| = 0 < \varepsilon.$$

- (2) $\lim_{n \geq 1} (-1)^n/n = 0$. Für alle $\varepsilon > 0$ ist jedes n_0 mit $1/n_0 < \varepsilon$ gut, denn für $n \geq n_0$ gilt:

$$|x_n - 0| = |x_n| = |(-1)^n/n| = 1/n \leq 1/n_0 < \varepsilon.$$

- (3) $(0, 1, 0, 1, 0, 1, \dots)$ divergiert. Denn sei $x \in \mathbb{R}$. Dann gilt für alle n :

$$|x - x_n| \geq 1/2 \quad \text{oder} \quad |x - x_{n+1}| \geq 1/2.$$

Also verifiziert $\varepsilon = 1/2$ die Divergenzbedingung für x .

- (4) Seien $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$ Folgen mit

$$\lim_n (x_n)_{n \in \mathbb{N}} = 0 \quad \text{und} \quad |y_n| \leq |x_n| \quad \text{für alle } n.$$

Dann gilt $\lim_n y_n = 0$. Denn sei $\varepsilon > 0$ und n_0 gut für ε und die Folge $(x_n)_{n \in \mathbb{N}}$.

Dann ist n_0 auch gut für ε und die Folge $(y_n)_{n \in \mathbb{N}}$, denn für alle $n \geq n_0$ gilt:

$$|y_n - 0| = |y_n| \leq |x_n| = |x_n - 0| < \varepsilon.$$

3.3 Monotone Folgen und Pendelfolgen

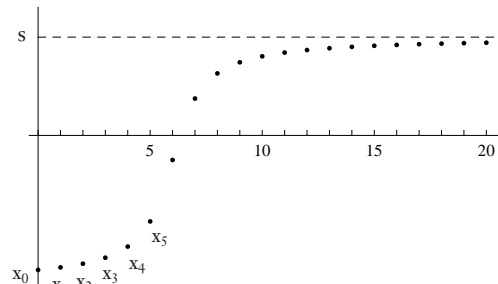
Satz (Konvergenz und Grenzwerte monotoner Folgen)

- (a) Sei $(x_n)_{n \in \mathbb{N}}$ monoton steigend und beschränkt in \mathbb{R} . Dann gilt

$$\lim_n x_n = \sup(\{x_n \mid n \in \mathbb{N}\}).$$

- (b) Analog gilt für eine monoton fallende und beschränkte Folge $(x_n)_{n \in \mathbb{N}}$, dass

$$\lim_n x_n = \inf(\{x_n \mid n \in \mathbb{N}\}).$$

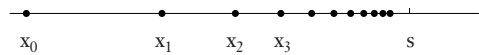


$$s = \sup_n x_n = \lim_n x_n$$

für eine monoton steigende und beschränkte Folge
in funktionaler und linearer Darstellung

Der Grenzwert einer monotonen beschränkten Folge ist also das Supremum oder Infimum der Folgenglieder.

Wir schreiben oft kurz $\sup_n x_n$ anstelle von $\sup(\{x_n \mid n \in \mathbb{N}\})$. Gleiches gilt für das Infimum.



Beispiele

- (1) Die Folge $(1/n)_{n \geq 1}$ ist monoton fallend und die Folge $(1 - 1/n)_{n \geq 1}$ ist monoton steigend. Folglich gilt

$$\lim_{n \geq 1} 1/n = \inf_n 1/n = 0,$$

$$\lim_n (1 - 1/n) = \sup_n (1 - 1/n) = 1.$$

- (2) Ist $x = k, d_1 d_2 d_3 \dots$ in Dezimaldarstellung mit $k \in \mathbb{N}$, so ist die Folge

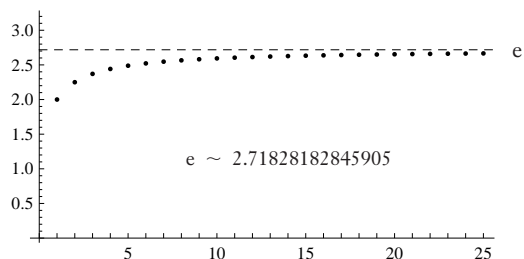
$$(k, d_1 \dots d_n)_{n \geq 1} = (k; k, d_1; k, d_1 d_2; \dots)$$

der n -ten Approximationen monoton steigend. Also gilt

$$x = \sup_{n \geq 1} k, d_1 \dots d_n \dots = \lim_{n \geq 1} k, d_1 \dots d_n.$$

- (3) Die Folge $((1 + 1/n)^n)_{n \geq 1}$ ist monoton steigend und beschränkt und damit konvergent. Man kann sie zur Definition der Eulerschen Zahl e verwenden:

$$e = \lim_{n \geq 1} (1 + 1/n)^n.$$



Ein häufig anzutreffender konvergenter nichtmonotoner Typ sind die „gedämpften Pendelfolgen“, deren Glieder sich hin und her bewegen und dabei immer mehr an Schwung verlieren:

Satz (Grenzwerte von Pendelfolgen)

(a) Sei $(x_n)_{n \in \mathbb{N}}$ eine rechtsstartende Pendelfolge in \mathbb{R} , d. h., es gelte

$$x_1 \leq x_3 \leq \dots \leq x_{2n+1} \leq \dots \leq x_{2n} \leq \dots \leq x_2 \leq x_0.$$

Weiter gelte

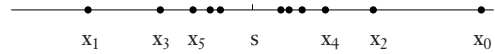
$$\forall \varepsilon > 0 \exists n \ x_{2n} - x_{2n+1} < \varepsilon.$$

(Konvergenzbedingung für Pendelfolgen)

$$\text{Dann gilt } \lim_n x_n = \sup_n x_{2n+1} = \inf_n x_{2n}.$$

(b) Analoges gilt für linksstartende Pendelfolgen

$$x_0 \leq x_2 \leq \dots \leq x_{2n} \leq \dots \leq x_{2n+1} \leq \dots \leq x_3 \leq x_1,$$



$s = \sup_n x_{2n+1} = \inf_n x_{2n} = \lim_n x_n$
in linearer Darstellung (Für eine funktionale Darstellung siehe das Diagramm zu Kettenbrüchen unten.)

Hier genügt es, zu jedem $\varepsilon > 0$ ein einziges gutes n zu finden!

Beispiele

(1) Die Folge $(1, -1, 1/2, -1/2, 1/3, -1/3, \dots)$ hat das Pendelverhalten

$$x_1 \leq x_3 \leq \dots \leq x_{2n+1} \leq \dots \leq x_{2n} \leq \dots \leq x_2 \leq x_0.$$

Weiter gilt die Konvergenzbedingung, da

$$x_{2n} - x_{2n+1} = 1/(n+1) - (-1/(n+1)) = 2/(n+1)$$

beliebig klein wird. Folglich ist $\lim_n x_n = \sup_{n \geq 1} -1/n = \inf_{n \geq 1} 1/n = 0$.

(2) Die sog. *Leibniz-Reihe* $(1, 1 - 1/3, 1 - 1/3 + 1/5, 1 - 1/3 + 1/5 - 1/7, \dots)$ ist eine rechtsstartende Pendelfolge und erfüllt die Konvergenzbedingung, da

$$x_{2n} - x_{2n+1} = 1/(4n+1) - 1/(4n+3) < 2/(4n+1)$$

beliebig klein wird. Also existiert $x = \lim_n x_n$. Es gilt $x = \pi/4$ (siehe auch 4.7).

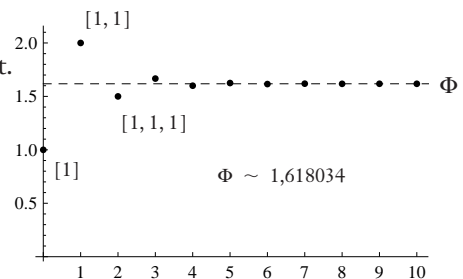
(3) Für positive natürliche Zahlen a_k sind die *endlichen Kettenbrüche* $[a_0, \dots, a_n]$ rekursiv definiert durch

$$[a_0] = a_0, \quad [a_0, \dots, a_{n+1}] = a_0 + \frac{1}{[a_1, \dots, a_{n+1}]}.$$

Für alle $(a_n)_{n \in \mathbb{N}}$ ist $([a_0, \dots, a_n])_{n \in \mathbb{N}}$ eine linksstartende Pendelfolge, die gegen eine irrationale Zahl konvergiert. Man setzt (*unendlicher Kettenbruch*):

$$[a_0, a_1, \dots] = \lim_n [a_0, \dots, a_n].$$

Der prominenteste unendliche Kettenbruch ist der Goldene Schnitt $\Phi = [1, 1, 1, \dots] = (1 + \sqrt{5})/2$.



3.4 Die Limesregeln

Satz (Limesregeln für punktweise Operationen)

Seien $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$ konvergente Folgen in \mathbb{R} , und sei $c \in \mathbb{R}$. Dann sind

$$(x_n + y_n)_{n \in \mathbb{N}}, \quad (c x_n)_{n \in \mathbb{N}},$$

$$(x_n - y_n)_{n \in \mathbb{N}}, \quad (x_n \cdot y_n)_{n \in \mathbb{N}}$$

konvergent und es gilt:

$$\lim_n (x_n + y_n) = \lim_n x_n + \lim_n y_n,$$

$$\lim_n (c x_n) = c \lim_n x_n,$$

$$\lim_n (x_n - y_n) = \lim_n x_n - \lim_n y_n,$$

$$\lim_n (x_n \cdot y_n) = \lim_n x_n \cdot \lim_n y_n.$$

Gilt $y_n \neq 0$ für alle n und $\lim_n y_n \neq 0$, so konvergiert $(x_n/y_n)_{n \in \mathbb{N}}$ mit

$$\lim_n \left(\frac{x_n}{y_n} \right) = \frac{\lim_n x_n}{\lim_n y_n}.$$

Für die *Summe zweier Folgen*

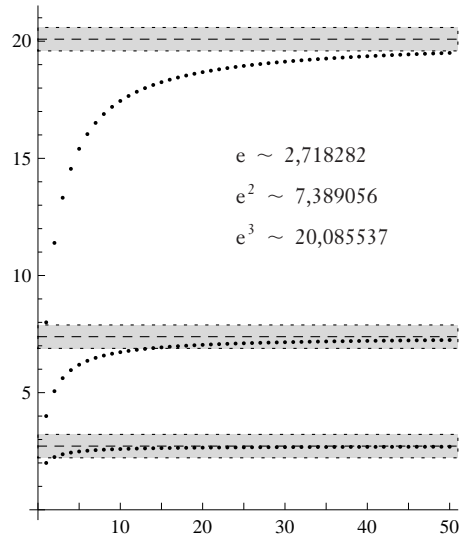
$$(x_n)_{n \in \mathbb{N}} + (y_n)_{n \in \mathbb{N}} = (x_n + y_n)_{n \in \mathbb{N}}$$

gilt also im Fall der Konvergenz: Der Grenzwert der Summe ist die Summe der Grenzwerte. Allgemeiner:

Für konvergente Folgen dürfen eine Limesbildung und eine punktweise arithmetische Operation in beliebiger Reihenfolge ausgeführt werden.

Die Vertauschbarkeit von Operationen ist eines der Grundmotive der Mathematik, dem wir noch öfter begegnen werden (etwa bei der Limesstetigkeit einer Funktion, vgl. 5.1).

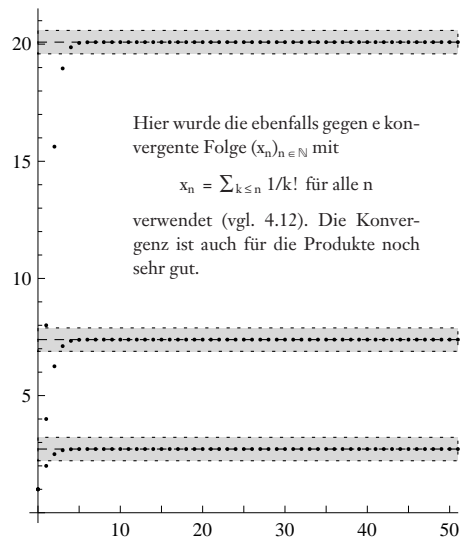
Die oft mühsame Bestimmung von Grenzwerten durch den Nachweis der Konvergenzbedingung wird durch die Limesregeln sehr erleichtert. Sie stellen uns einen „ ϵ -freien“ Kalkül zur Verfügung. Weiter erlauben sie einen eleganten Beweis der Existenz von Wurzeln in \mathbb{R} (vgl. Beispiel (6)).



Das Diagramm zeigt die je ersten 50 Glieder der drei Folgen $(x_n)_{n \geq 1}$, $(x_n^2)_{n \geq 1}$ und $(x_n^3)_{n \geq 1}$ mit

$$x_n = (1 + 1/n)^n \text{ für alle } n \geq 1.$$

Die erste Folge konvergiert gegen die Eulersche Zahl e , die beiden anderen gegen e^2 und e^3 (vgl. Beispiel (4)). Der ϵ -Streifen für $\epsilon = 1/2$ zeigt, wie in diesem Beispiel die Geschwindigkeit der Konvergenz durch die Produktbildung abnimmt.



Hier wurde die ebenfalls gegen e konvergente Folge $(x_n)_{n \in \mathbb{N}}$ mit

$$x_n = \sum_{k \leq n} 1/k! \text{ für alle } n$$

verwendet (vgl. 4.12). Die Konvergenz ist auch für die Produkte noch sehr gut.

Beispiele

(1) Es gilt

$$(1, 1/2, 1/3, \dots) - (1/(1 \cdot 2), 1/(2 \cdot 3), 1/(3 \cdot 4), \dots) = \\ (1 - 1/2, 1/2 - 1/6, 1/3 - 1/12, \dots) = (1/2, 1/3, 1/4, \dots),$$

d.h.

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

(2) Für alle $c \in \mathbb{R}$ gilt $\lim_n c = c$. Wegen $\lim_n 1/(n+1) = 0$ ist also

$$\lim_{n \geq 1} \left(c + \frac{1}{n+1}\right) = c + 0 = c.$$

Allgemeiner gilt: Ist $(x_n)_{n \in \mathbb{N}}$ eine gegen x konvergente Folge und $(y_n)_{n \in \mathbb{N}}$ eine Nullfolge, d.h., gilt $\lim_n y_n = 0$, so gilt

$$\lim_n (x_n + y_n) = \lim_n x_n + \lim_n y_n = x + 0 = x.$$

Wir können also eine konvergente Folge punktweise ohne Veränderung des Grenzwertes modifizieren, wenn wir dies in einer „im Limes unwesentlichen“ Art und Weise tun.

(3) Es gilt

$$(0, 1, 0, 1, 0, 1, \dots) + (1, 0, 1, 0, 1, 0, \dots) = (1, 1, 1, \dots).$$

Zwei divergente Folgen können also eine konvergente Summe haben. Aus der Existenz von $\lim_n x_n + y_n$ darf man im Allgemeinen nicht schließen, dass die Grenzwerte $\lim_n x_n$ und $\lim_n y_n$ existieren.

(4) Es gelte $\lim_n x_n = x$. Dann gilt

$$\lim_n (x_n^2) = \lim_n (x_n x_n) = (\lim_n x_n) \cdot (\lim_n x_n) = x \cdot x = x^2.$$

Induktiv zeigt man, dass allgemeiner $\lim_n (x_n^k) = x^k$ für alle $k \geq 1$ gilt.

(5) Es gilt

$$\lim_{n \geq 1} \frac{1 + 2n^2}{n^3} = \lim_{n \geq 1} \left(\frac{1}{n^3} + \frac{2n^2}{n^3}\right) = \\ \left(\lim_{n \geq 1} \frac{1}{n}\right)^3 + 2 \lim_{n \geq 1} \frac{1}{n} = 0^3 + 2 \cdot 0 = 0.$$

(6) Wir zeigen mit Hilfe der Limesregeln, dass $\sqrt{2}$ existiert. Nach dem Vollständigkeitsaxiom existiert $s = \sup(\{x \in \mathbb{R} \mid x^2 \leq 2\})$. Nun gilt aber

$$2 \leq \lim_{n \geq 1} (s + 1/n)^2 = (s + 0)^2 = s^2 = (s - 0)^2 = \lim_{n \geq 1} (s - 1/n)^2 \leq 2.$$

Damit ist $s^2 = 2$. Das Argument lässt sich leicht verallgemeinern, um die Existenz von n -ten Wurzeln zu zeigen.

3.5 Cauchy-Folgen

Definition (Cauchy-Folge)

Eine Folge $(x_n)_{n \in \mathbb{N}}$ in \mathbb{R} heißt eine *Cauchy-Folge*, falls gilt:

$$\forall \varepsilon > 0 \exists n_0 \forall n, m \geq n_0 |x_n - x_m| < \varepsilon.$$

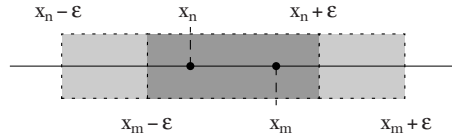
(Cauchy-Bedingung)

Die Cauchy-Bedingung stellt eine Bedingung an die Folgenglieder, ein Grenzwert der Folge ist nicht involviert.

Es gilt der fundamentale Satz:

Satz (Konvergenz von Cauchy-Folgen in \mathbb{R})

Jede Cauchy-Folge in \mathbb{R} konvergiert.



Zur Cauchy-Bedingung: Für gegebenes $\varepsilon > 0$ haben ab einem Index n_0 alle Glieder x_n, x_m einen Abstand kleiner als ε voneinander. Betrachten wir ein solches Paar x_n, x_m wie im Diagramm, so liegen alle x_k mit $k \geq n_0$ im dunkelgrauen Bereich, da sie sowohl von x_n als auch von x_m weniger als ε entfernt sind.

Umgekehrt ist jede konvergente Folge eine Cauchy-Folge. Denn ist $x = \lim_n x_n$, so gibt es für alle $\varepsilon > 0$ ein n_0 mit $|x - x_n| < \varepsilon/2$ für alle $n \geq n_0$. Dann gilt aber für alle $m, n \geq n_0$:

$$|x_n - x_m| = |x_n - x + x - x_m| \leq |x_n - x| + |x - x_m| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Dieses Vorgehen ist als „ $\varepsilon/2$ -Argument“ bekannt (vgl. auch das letzte Beispiel in 2.5).

Zusammenfassend gilt also:

Eine Folge in \mathbb{R} konvergiert genau dann, wenn sie eine Cauchy-Folge ist.

Beispiel

Wir betrachten die Folge

$$1, \quad 1 + \frac{1}{4}, \quad 1 + \frac{1}{4} + \frac{1}{9}, \quad 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16}, \quad \dots,$$

d. h., es gilt $x_n = \sum_{1 \leq k \leq n} 1/k^2$ für alle $n \geq 1$. Für alle $1 \leq n_0 \leq n \leq m$ gilt dann:

$$\begin{aligned} |x_m - x_n| &= \sum_{n < k \leq m} 1/k^2 < \sum_{n < k \leq m} (1/(k-1) - 1/k) = \\ 1/n - 1/(n+1) + 1/(n+1) - 1/(n+2) + \dots - 1/(m-1) + 1/(m-1) - 1/m &= \\ 1/n - 1/m &\leq 1/n + 1/m \leq 2/n_0, \end{aligned}$$

wobei wir $1/k^2 < 1/(k(k-1)) = 1/(k-1) - 1/k$ verwenden und eine sog. *Teleskopsumme* aufgelöst haben. Da für alle $\varepsilon > 0$ ein n_0 existiert mit $2/n_0 < \varepsilon$, ist die Folge eine Cauchy-Folge, und damit existiert $x = \lim_n x_n$. Mit weitergehenden Methoden (z. B. mit Fourier-Reihen) kann man den Grenzwert x als $\pi^2/6$ berechnen.

Man kann natürlich auch so argumentieren: Die Folge des Beispiels ist monoton steigend und die Teleskopsumme zeigt, dass sie beschränkt und folglich konvergent ist. Zum Nachweis der Konvergenz der Folge $(y_n)_{n \geq 1}$ mit $y_n = \sum_{1 \leq k \leq n} \sin(k)/k^2$ für alle $n \geq 1$ ist die Cauchy-Bedingung essentiell.

Die Konvergenz von Cauchy-Folgen wird oft als „metrisches Vollständigkeitsaxiom“ für die reellen Zahlen bezeichnet. Unser lineares Vollständigkeitsaxiom ist äquivalent zur metrischen Vollständigkeit zusammen mit der archimedischen Anordnung (vgl. auch 2. 7). Später betrachtet man in der Analysis sehr allgemeine metrische Räume und dann wird die metrische Version zur Formulierung der Vollständigkeit herangezogen. Supremum und Infimum stehen nicht mehr zur Verfügung.

Die Cauchy-Bedingung besagt, dass für jedes $\varepsilon > 0$ ab einer bestimmten Stelle der Abstand je zweier Folgenglieder kleiner als ε ist. Anschaulich bedeutet dies, dass sich die Glieder beliebig verdichten. Dabei ist aber Vorsicht geboten:

Es genügt nicht, dass der Abstand zwischen benachbarten Gliedern x_n und x_{n+1} beliebig klein wird.

Beispiel

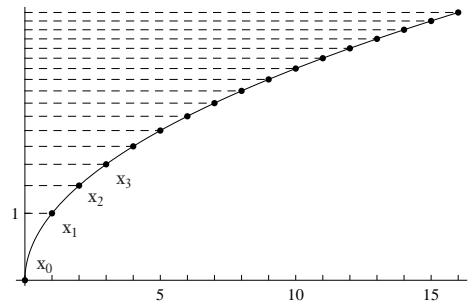
Für alle $n \geq 1$ gilt $\sqrt{1 + 1/n} < 1 + 1/n$ und damit

$$\begin{aligned}\sqrt{n+1} - \sqrt{n} &= \sqrt{n}(\sqrt{1 + 1/n} - 1) < \\ \sqrt{n}(1 + 1/n - 1) &= 1/\sqrt{n}.\end{aligned}$$

Also gilt $\lim_n (\sqrt{n+1} - \sqrt{n}) = 0$.

Setzen wir also $x_n = \sqrt{n}$ für alle n , so ist $(x_n)_{n \in \mathbb{N}}$ eine divergente Folge mit

$$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 |x_{n+1} - x_n| < \varepsilon.$$



Die Situation des Beispiels ist gar nicht so selten. Sie liegt zum Beispiel auch für die Folgen $(\log(n))_{n \geq 1}$ und für $(h_n)_{n \geq 1}$ mit

$$h_n = 1 + 1/2 + \dots + 1/n \quad \text{für alle } n \in \mathbb{N}$$

vor. Die Folge $(h_n)_{n \geq 1}$ ist die *harmonische Reihe*, die wir in 4.5 genauer besprechen werden.

Variante der Cauchy-Bedingung

Folgende gleichwertige Variante der Cauchy-Bedingung kommt mit nur einer Variablen oberhalb von n_0 aus:

$$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 |x_n - x_{n_0}| < \varepsilon. \quad (\text{Cauchy-Bedingung, II})$$

Dies folgt aus der ersten Version, wenn wir dort $m = n_0$ setzen. Umgekehrt ist es die Dreiecksungleichung, die uns von der scheinbar schwächeren zweiten Version zur ersten führt. Gilt nämlich die Version II und ist $\varepsilon > 0$ beliebig, so gibt es ein n_0 mit

$$\forall n \geq n_0 |x_n - x_{n_0}| < \varepsilon/2.$$

Für alle $n, m \geq n_0$ gilt dann

$$|x_n - x_m| \leq |x_n - x_{n_0}| + |x_{n_0} - x_m| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

3.6 Teilfolgen

Definition (*Teilfolge einer Folge*)

Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} , und sei $(i_n)_{n \in \mathbb{N}}$ eine streng monoton steigende Folge in \mathbb{N} . Dann heißt die Folge $(y_n)_{n \in \mathbb{N}}$ mit

$$y_n = x_{i_n} \text{ für alle } n$$

die durch die *Indexfolge* $(i_n)_{n \in \mathbb{N}}$ definierte *Teilfolge* von $(x_n)_{n \in \mathbb{N}}$.

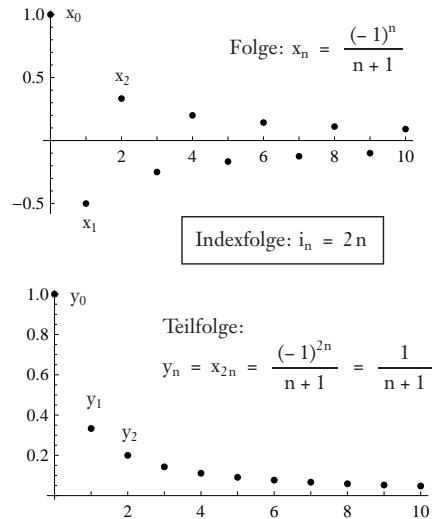
Folge: $x_0, x_1, x_2, x_3, \dots$

Indexfolge: $0, 2, 4, 6, \dots$

Teilfolge $(y_n)_{n \in \mathbb{N}}$: $x_0, x_2, x_4, x_6, \dots$

$y_0 = x_0, y_1 = x_1, y_2 = x_4, y_3 = x_6, \dots$

allgemein: $y_n = x_{i_n} = x_{2n}$ für alle n .



Anschaulich entsteht eine Teilfolge durch ein Aussieben der Folge, wobei unendlich viele Glieder den Siebprozess überleben. Das Primzahlsieb des Eratosthenes ist ein klassisches Beispiel für einen solchen Prozess. Die formale Beherrschung dieses sehr anschaulichen Begriffs bereitet Anfängern erfahrungsgemäß große Schwierigkeiten. Fragt man Studienanfänger nach einer Definition, so erhält man oft einen Zeichensalat aus x_n und i_n oder man erhält Beschreibungen wie „das erste Element der Teilfolge muss nicht das erste Element der Folge sein“ oder „es dürfen beliebig viele Elemente übersprungen werden“, die ja alle richtig sind, aber keine Definition des Begriffs darstellen.

Die Definition wird vielleicht bekömmlicher, wenn wir uns klarmachen, dass eine Verknüpfung von Funktionen am Werk ist:

Eine Folge $(x_n)_{n \in \mathbb{N}}$ ist eine Funktion f auf \mathbb{N} .
Eine Indexfolge $(i_n)_{n \in \mathbb{N}}$ wie in der Definition ist eine streng monoton steigende Funktion $g : \mathbb{N} \rightarrow \mathbb{N}$.
Die durch $(i_n)_{n \in \mathbb{N}}$ definierte Teilfolge $(y_n)_{n \in \mathbb{N}}$ von $(x_n)_{n \in \mathbb{N}}$ ist die Funktion $h = f \circ g$ auf \mathbb{N} , d. h., es gilt $h(n) = f(g(n)) = f(i_n) = x_{i_n} \text{ für alle } n \in \mathbb{N}.$ In Folgennotation ist h also die Folge $(x_{i_n})_{n \in \mathbb{N}}$.

Der Index in x_{i_n} ist immer als $x_{(i_n)}$ zu lesen, nie als $(x_i)_{i_n}$. Neben der Notation $(x_{i_n})_{n \in \mathbb{N}}$ ist gleichwertig auch $(x_{i(n)})_{n \in \mathbb{N}}$ gebräuchlich. Letztere Form vermeidet doppelte Indizes und betont den funktionalen Charakter der Indexfolge.

Beispiele

- (1) Die konstante Folge (c, c, c, \dots) besitzt nur eine Teilfolge, nämlich

$$(c, c, c, \dots).$$

Denn ist $(i_n)_{n \in \mathbb{N}}$ streng monoton steigend in \mathbb{N} , so gilt $x_{i_n} = c$ für alle n . Also ist die durch $(i_n)_{n \in \mathbb{N}}$ definierte Teilfolge von (c, c, c, \dots) konstant gleich c .

- (2) Ist $(x_n)_{n \in \mathbb{N}}$ monoton steigend, so ist auch jede Teilfolge $(x_{i_n})_{n \in \mathbb{N}}$ der Folge monoton steigend. Denn für alle $n \in \mathbb{N}$ gilt $i_n < i_{n+1}$, und damit ist auch $x_{i_n} \leq x_{i_{n+1}}$ aufgrund der Monotonie von $(x_n)_{n \in \mathbb{N}}$. Gleiches gilt für monoton fallende Folgen, und auch die strenge Monotonie vererbt sich auf Teilfolgen.

- (3) Wir betrachten die Folge $(n)_{n \in \mathbb{N}} = (0, 1, 2, 3, \dots)$. Dann gilt:

$(1, 3, 5, \dots)$ ist eine Teilfolge, definiert durch $(2n+1)_{n \in \mathbb{N}}$,

$(1, 0, 2, 3, 4, 5, \dots)$ und $(0, 0, 1, 2, 3, \dots)$ sind keine Teilfolgen.

Allgemein besitzt $(0, 1, 2, 3, 4, \dots)$ genau die streng monotonen Folgen in \mathbb{N} als Teilfolgen. Denn ist $(i_n)_{n \in \mathbb{N}}$ streng monoton in \mathbb{N} , so gilt

$$x_{i_n} = i_n \quad \text{für alle } n \in \mathbb{N},$$

da ja $x_k = k$ für alle $k \in \mathbb{N}$ gilt. Folglich ist die durch $(i_n)_{n \in \mathbb{N}}$ definierte Teilfolge von $(0, 1, 2, \dots)$ die Folge $(i_n)_{n \in \mathbb{N}}$ selbst. Umgekehrt erbt nach (2) jede Teilfolge von $(n)_{n \in \mathbb{N}}$ die strenge Monotonie dieser Folge.

- (4) Wir betrachten die Pendelfolge $(0, 1, 0, 1, 0, \dots)$. Die Indexfolgen $(2n)_{n \in \mathbb{N}}$ und $(2n+1)_{n \in \mathbb{N}}$ definieren die Teilfolgen $(0, 0, 0, \dots)$ und $(1, 1, 1, \dots)$. Daneben sind beispielsweise aber auch

$$(0, 0, 1, 1, 0, 0, 1, 1, \dots) \quad \text{und} \quad (0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, \dots)$$

Teilfolgen. Allgemein gilt:

Die Teilfolgen von $(0, 1, 0, 1, 0, 1, \dots)$ sind genau die Folgen in $\{0, 1\}$.

Der Wechsel zwischen 0 und 1 in einer Teilfolge muss nicht periodisch oder regelmäßig sein, auch „chaotische“ oder „zufällige“ 0-1-Folgen sind Teilfolgen.

- (5) Die Folge $(1/2^n)_{n \in \mathbb{N}}$ ist eine Teilfolge von $(1/(n+1))_{n \in \mathbb{N}}$. Sie wird durch die Indexfolge $(2^n - 1)_{n \in \mathbb{N}}$ definiert.
- (6) Sei $(x_n)_{n \in \mathbb{N}}$ eine Pendelfolge mit

$$x_0 \leq x_2 \leq \dots \leq x_{2n} \leq \dots \leq x_{2n+1} \leq \dots \leq x_3 \leq x_1.$$

Beispiel (4) zeigt, dass der Übergang zu einer Teilfolge im Allgemeinen die Pendelbewegung zerstört. Ist aber $(i_n)_{n \in \mathbb{N}}$ streng monoton in \mathbb{N} derart, dass i_{2n} gerade und i_{2n+1} ungerade für alle n ist, so ist $(x_{i_n})_{n \in \mathbb{N}}$ wieder eine Pendelfolge.

3.7 Häufungspunkte von Folgen

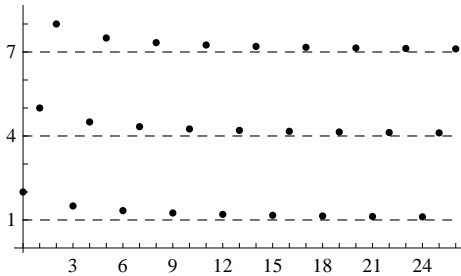
Definition (Häufungspunkt einer Folge)

Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} . Dann heißt ein x ein *Häufungspunkt* von $(x_n)_{n \in \mathbb{N}}$, falls es eine Teilfolge $(y_n)_{n \in \mathbb{N}}$ von $(x_n)_{n \in \mathbb{N}}$ gibt mit

$$x = \lim_{n \rightarrow \infty} y_n.$$

Die Häufungspunkte einer Folge sind also diejenigen Punkte, die wir durch eine Ausziehung der Folge „ansteuern“ können.

Wir bestimmen zunächst die Häufungspunkte der in den sechs Beispielen der vorhergehenden Sektion betrachteten Folgen.



$$x_n = r + 1 + \frac{1}{k+1} \text{ für } n = 3k + r, r \in \{0, 1, 2\}$$

Die Folge hat die Häufungspunkte 1, 4, und 7:

$$1 = \lim_{n \rightarrow \infty} x_{3n}, \quad 4 = \lim_{n \rightarrow \infty} x_{3n+1}, \quad 7 = \lim_{n \rightarrow \infty} x_{3n+2}.$$

Beispiele

- (1) Die Folge (c, c, c, \dots) besitzt genau den Häufungspunkt c .
- (2) Die Folge $(0, 1, 2, 3, \dots)$ besitzt keine Häufungspunkte.
- (3) Die Folge $(0, 1, 0, 1, 0, 1, \dots)$ besitzt genau die Häufungspunkte 0 und 1.
- (4) Die Folge $(1/n)_{n \geq 1}$ besitzt genau den Häufungspunkt 0.
- (5) Ist $(x_n)_{n \in \mathbb{N}}$ monoton steigend, so besitzt die Folge, falls sie beschränkt ist, genau den Häufungspunkt $x = \lim_{n \rightarrow \infty} x_n = \sup_{n \in \mathbb{N}} x_n$. Ist die Folge unbeschränkt, so besitzt sie keinen Häufungspunkt.
- (6) Ist $(x_n)_{n \in \mathbb{N}}$ eine Pendelfolge mit

$$x_0 \leq x_2 \leq \dots \leq x_{2n} \leq \dots \leq x_{2n+1} \leq \dots \leq x_3 \leq x_1,$$

so sind genau

$$a = \lim_{n \rightarrow \infty} x_{2n} = \sup_{n \in \mathbb{N}} x_{2n} \quad \text{und} \quad b = \lim_{n \rightarrow \infty} x_{2n+1} = \inf_{n \in \mathbb{N}} x_{2n+1}$$

die Häufungspunkte der Folge. Weiter gilt $a = b$ genau dann, wenn die Folge konvergiert, und in diesem Fall gilt $\lim_{n \rightarrow \infty} x_n = a = b$.

Die Beispiele zeigen, dass ein x ein Häufungspunkt einer Folge sein kann, obwohl sich die Glieder der Folge nicht räumlich um x herum häufen. Auch eine unendliche „Stapelung“ gilt als Häufungspunkt. Häufungen ohne Stapelungen werden wir im Begriff des Häufungspunktes einer Menge kennenlernen (vgl. 3.10).



Die Folge $0, 1, 0, 1, 0, 1, \dots$ in linearer Darstellung. Häufungspunkte sind 0 und 1.

Die Bestimmung der Häufungspunkte einer Folge ist im Allgemeinen eine komplizierte Angelegenheit. Im konvergenten Fall ist die Aufgabe einfach:

Satz (*Häufungspunkte konvergenter Folgen*)

Sei $(x_n)_{n \in \mathbb{N}}$ eine konvergente Folge in \mathbb{R} , und sei $x = \lim_n x_n$. Dann ist x der eindeutige Häufungspunkt von $(x_n)_{n \in \mathbb{N}}$.

Dagegen zeigen die obigen Beispiele, dass eine Folge auch gar keinen oder mehrere Häufungspunkte besitzen kann. Die auftretenden Mengen sind vielfältig, aber nicht beliebig:

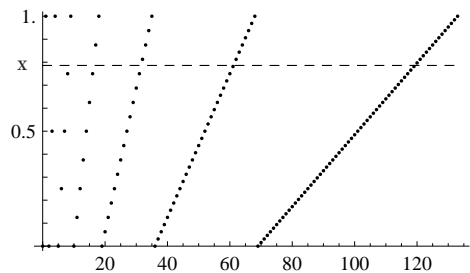
Beispiele

- (1) Die periodische Folge $(a_1, \dots, a_n, a_1, \dots, a_n, a_1, \dots, a_n, \dots)$ besitzt genau die Häufungspunkte a_1, \dots, a_n . Damit ist jede endliche Teilmenge von \mathbb{R} die Menge der Häufungspunkte einer Folge.

- (2) Jedes $x \in [0, 1]$ ist Häufungspunkt der Folge im Diagramm rechts, die $[0, 1]$ immer feiner durchläuft:

$0, 1, 0, 1/2, 1, 0, 1/4, 2/4, 3/4, 1,$
 $0, 1/8, 2/8, \dots, 6/8, 7/8, 1, \dots$

Ist $(x_n)_{n \in \mathbb{N}}$ eine Aufzählung von \mathbb{Q} , so ist sogar jede reelle Zahl ein Häufungspunkt der Folge. Eine Folge kann also überabzählbar viele Häufungspunkte besitzen.



- (3) Sind $1/2, 1/3, 1/4, \dots$ Häufungspunkte einer Folge, so ist auch 0 ein Häufungspunkt. Damit tritt nicht jede Teilmenge von \mathbb{R} als Häufungspunktmenge auf.

Der folgende Satz ist geeignet, die Aussagen dieser Beispiele zu beweisen:

Satz (*Charakterisierung der Häufungspunkte*)

Seien $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} und $x \in \mathbb{R}$. Dann ist x genau dann ein Häufungspunkt von $(x_n)_{n \in \mathbb{N}}$, falls gilt:

$$\forall \varepsilon > 0 \quad \forall n_0 \quad \exists n \geq n_0 \quad |x - x_n| < \varepsilon. \quad (\text{Häufungspunktbedingung für } x)$$

Eine äquivalente Formulierung der Häufungspunktbedingung für x lautet:

$$\text{Für alle } \varepsilon > 0 \text{ gibt es unendliche viele } n \text{ mit } |x_n - x| < \varepsilon.$$

Der Leser vergleiche dies noch einmal mit der Konvergenzbedingung für x :

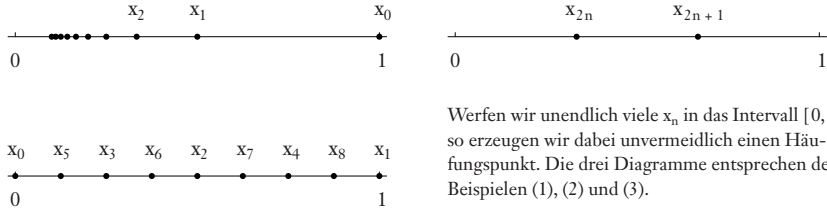
$$\text{Für alle } \varepsilon > 0 \text{ gibt es ein } n_0, \text{ sodass für alle } n \geq n_0 \text{ gilt, dass } |x_n - x| < \varepsilon.$$

Die Konvergenzbedingung für x ist stärker. Dort wird verlangt, dass es nur endlich viele Ausnahmen n zu „ $|x_n - x| < \varepsilon$ “ gibt. Für „ x ist Häufungspunkt“ wird lediglich verlangt, dass es unendlich viele Zeugen n für „ $|x_n - x| < \varepsilon$ “ gibt.

3.8 Der Satz von Bolzano-Weierstraß

Satz (*Satz von Bolzano-Weierstraß für Folgen*)

■ Jede beschränkte Folge in \mathbb{R} besitzt einen Häufungspunkt.



Eine Folge $(x_n)_{n \in \mathbb{N}}$ in \mathbb{R} ist beschränkt, falls die Menge der Folgenglieder eine untere und eine obere Schranke besitzt, d.h. falls es $a, b \in \mathbb{R}$ gibt mit

$$x_n \in [a, b] \quad \text{für alle } n \in \mathbb{N}.$$

Der Satz besagt, dass in diesem Fall immer eine konvergente Teilfolge der Folge existiert. Um ein Gefühl für die Aussage zu bekommen, möge der Leser das Intervall $[0, 1]$ auf ein Papier zeichnen und es mit beliebigen Punkten x_0, x_1, x_2, \dots füllen. Die Behauptung ist, dass eine derartige Füllung immer einen Häufungspunkt erzeugt, wobei auch wieder ein unendlich oft markierter Punkt als Häufungspunkt gilt. In der Tat bekommt man schnell den Eindruck, dass die Behauptung richtig ist. Eine Häufung scheint aufgrund der Enge des verfügbaren Platzes unvermeidbar.

Beispiele

- (1) Markieren wir in $[0, 1]$ die Punkte $1, 1/2, 1/4, \dots$, so erzeugen wir 0 als Häufungspunkt.
- (2) Springen wir in $[0, 1]$ zwischen $1/3$ und $2/3$ hin und her, so erzeugen wir die Häufungspunkte $1/3$ und $2/3$.
- (3) Nun versuchen wir Wiederholungen und ein Zulaufen auf einen Punkt zu vermeiden und unsere Punkte immer mit möglichst großem Abstand zu setzen. Wir markieren 0, 1, dann $1/2$, dann $1/4, 3/4$, dann $1/8, 3/8, 5/8, 7/8$, usw. So wird aber sogar jedes $x \in [0, 1]$ zu einem Häufungspunkt!

Die Situation ändert sich auch nicht, wenn wir anstelle des Intervalls $[0, 1]$ größere Intervalle wie $[0, 10], [0, 100], [0, 1000]$ usw. mit Punkten füllen. Sie ändert sich allerdings vollkommen, wenn wir das unbeschränkte Intervall $[0, +\infty[$ zur Verfügung haben. Wir können dann ins Unendliche fliehen und die Punkte $0, 1, 2, 3, \dots$ markieren, ohne dadurch einen Häufungspunkt zu erzeugen.

Der Satz von Bolzano-Weierstraß besitzt einen einfachen und anschaulichen Beweis, der von Anfängern aber oft als kompliziert empfunden wird, weil sie mit rekursiven Methoden und den doppelten Indizes der Teilfolgen nicht vertraut sind. Wir begnügen uns hier mit einer anschaulichen Formulierung der Beweisidee, die helfen will, die strengen

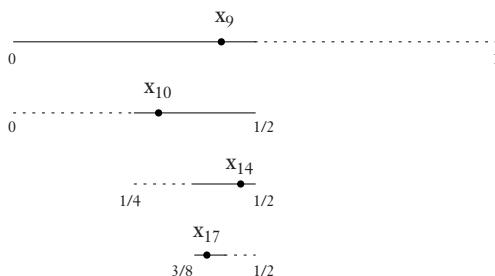
Beweise der Lehrbücher und Vorlesungen leichter lesbar zu machen. Der Einfachheit halber nehmen wir wie im obigen Gedankenexperiment an, die Folge $(x_n)_{n \in \mathbb{N}}$ befinde sich im Intervall $[0, 1]$. Ansonsten brauchen wir noch einen blauen Stift.

Anschauliche Darstellung des Beweises

Es liege also eine Folge $x_0, x_1, \dots, x_n, \dots$ in $[0, 1]$ vor. Wir teilen nun $[0, 1]$ in die zwei Hälften $[0, 1/2]$ und $[1/2, 1]$. Liegen unendlich viele Glieder der Folge in der linken Hälfte, so markieren wir dieses Intervall blau und versehen das erste Glied der Folge, das in diesem Intervall liegt, mit einem blauen Kreis. (Vielleicht ist es bereits x_0 , vielleicht ist es erst x_9 .) Andernfalls liegen unendlich viele Glieder der Folge in der rechten Hälfte. In diesem Fall markieren wir diese Hälfte und das erste Glied der Folge, das in dieser Hälfte liegt, blau.

(Es kann sein, dass sowohl im linken als auch im rechten Intervall unendlich viele Glieder liegen. In diesem Fall markieren wir trotzdem nur das linke Intervall blau. Was nicht sein kann, ist, dass sich links und rechts jeweils nur endlich viele Glieder befinden, denn es gibt insgesamt unendlich viele Glieder.)

Nun wiederholen wir die Halbierung mit dem ausgezeichneten blauen Intervall. So finden wir wieder ein Intervall – es hat nun die Länge $1/4$ –, in dem immer noch unendlich viele Glieder der Folge liegen. Wir markieren das Intervall blau und versehen das erste Folgenglied oberhalb des zuvor blau markierten Gliedes, das in unserem blauen Intervall der Länge $1/4$ liegt, mit einem blauen Kreis. (War es zuvor x_9 , so könnte es bereits x_{10} sein, vielleicht aber auch erst x_{42} .) Nun wiederholen wir dieses Verfahren unendlich oft. Wir erhalten geschachtelte blaue Intervallen, die sich aufgrund der fortgesetzten Halbierung auf einen Punkt x^* zusammenziehen. Weiter haben wir unendlich viele Folgenglieder mit einem blauen Kreis markiert. Sie bilden eine Teilfolge der Folge. Diese Teilfolge muss gegen x^* konvergieren, da für alle n das n -te blaue Glied nach Konstruktion im n -ten blauen Intervall liegt.



Jedes durchgezogene „blaue“ Intervall enthält unendlich viele, jedes gestrichelte Intervall endlich oder unendlich viele Glieder der Folge. In jedem durchgezogenen Intervall markieren wir ein neues Glied der Folge. So entsteht eine Teilfolge, die gegen den durch die Intervallschachtelung definierten Punkt konvergiert.

Eine genaue mathematische Formulierung kann diese anschauliche Darstellung des Beweises nicht ersetzen. Die Beweisidee ist vollständig vorhanden, aber die Form der Argumentation entspricht nicht der üblichen mathematischen Beweisführung. Der Anfänger ist aber aufgerufen, sich möglichst viele Beweise anschaulich klarzumachen. Beherrscht man dann noch die Übersetzung der Anschauung in die mathematische Fachsprache, so wird alles ganz einfach.

3.9 Limes Superior und Inferior

Definition (*Limes Superior und Limes Inferior*)

Sei $(x_n)_{n \in \mathbb{N}}$ eine beschränkte Folge in \mathbb{R} . Dann definieren wir:

$$\limsup_n x_n = \inf_{n \in \mathbb{N}} \sup_{m \geq n} x_m \quad (= \lim_n \sup_{m \geq n} x_m),$$

$$\liminf_n x_n = \sup_{n \in \mathbb{N}} \inf_{m \geq n} x_m \quad (= \lim_n \inf_{m \geq n} x_m).$$

Die reelle Zahl $\limsup_n x_n$ heißt der *Limes Superior* und die reelle Zahl $\liminf_n x_n$ der *Limes Inferior* der Folge $(x_n)_{n \in \mathbb{N}}$.

Die Begriffe wirken auf viele Anfänger aufgrund der lim-inf-sup-Kombination bedrohlich. Anschauliche Beschreibungen können bei der Befreundung helfen:

Wir betrachten zunächst das Supremum s_0 aller x_n , $n \geq 0$. Nun radieren wir x_0 weg und betrachten das Supremum s_1 der übrig gebliebenen x_n . Da wir nun weniger Punkte betrachten, gilt $s_0 \geq s_1$. Nun radieren wir x_1 weg und bilden das Supremum s_2 aller x_n mit $n \geq 2$. Dann gilt $s_0 \geq s_1 \geq s_2$. So fortfahrend erhalten wir eine monoton fallende Folge

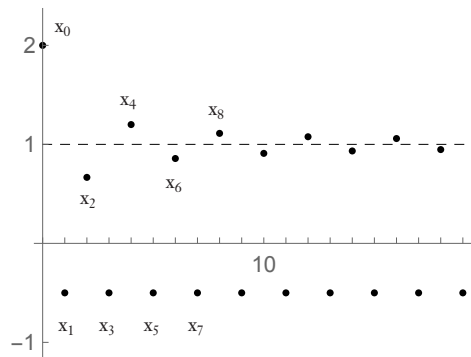
$$s_0 \geq s_1 \geq s_2 \geq \dots \geq s_n \geq \dots,$$

mit $s_n = \sup_{m \geq n} x_m$. Es gilt

$$s^* = \inf_n s_n = \lim_n s_n = \limsup_n x_n.$$

Analoges gilt für den Limes Inferior.

$$x_n = \begin{cases} 1 + (-1)^k/(n+1) & \text{für } n = 2k, \\ -1/2 & \text{für } n \text{ ungerade.} \end{cases}$$

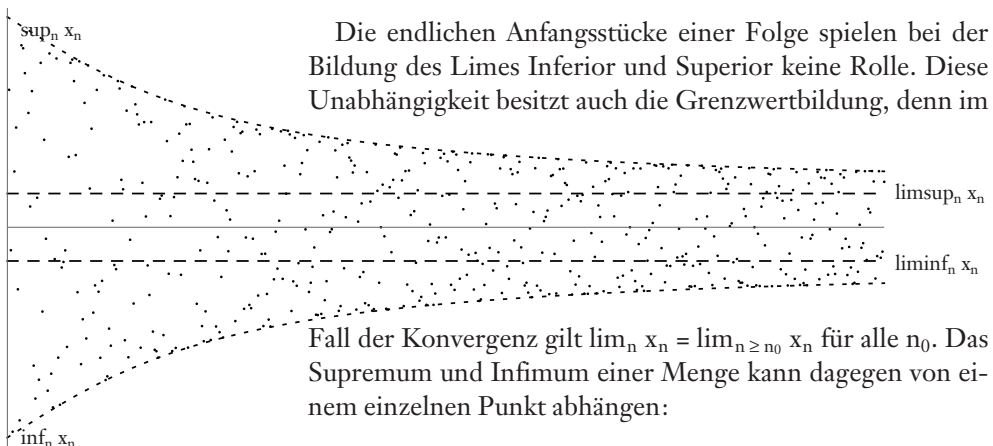


$$s_0 = x_0,$$

$$s_1 = s_2 = s_3 = s_4 = x_4,$$

$$s_5 = s_6 = s_7 = s_8 = x_8, \dots$$

$$\limsup_n x_n = \inf_n s_n = \inf_n x_{4n} = 1.$$



$$\sup([0, 1]) = 1 < 2 = \sup([0, 1] \cup \{2\}).$$

Beispiele

(1) Sei $(x_n)_{n \in \mathbb{N}} = (c, c, c, \dots)$ für ein $c \in \mathbb{R}$. Dann gilt

$$\liminf_n x_n = c = \limsup_n x_n.$$

(2) Sei $(x_n)_{n \in \mathbb{N}} = (0, 1, 0, 1, 0, 1, \dots)$. Dann gilt

$$\liminf_n x_n = 0 < 1 = \limsup_n x_n.$$

(3) Sei $(x_n)_{n \in \mathbb{N}} = (-1, 1, -1/2, 1/2, -1/4, 1/4, -1/8, 1/8, \dots)$. Dann gilt

$$\liminf_n x_n = 0 = \limsup_n x_n \quad \text{und} \quad 0 = \lim_n x_n.$$

(4) Für $(x_n)_{n \geq 1}$ wie im Diagramm oben rechts gilt

$$\liminf_n x_n = -1/2 < 1 = \limsup_n x_n.$$

Für alle Folgen $(x_n)_{n \in \mathbb{N}}$ gilt $\liminf_n x_n \leq \limsup_n x_n$. Die Gleichheit von Limes Inferior und Limes Superior tritt genau im Fall der Konvergenz der Folge ein:

Satz (*Charakterisierung des Grenzwerts mit liminf und limsup*)

Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} , und sei $x \in \mathbb{R}$. Dann sind äquivalent:

- (a) $\lim_n x_n = x$.
- (b) (x_n) ist beschränkt und $\liminf_n x_n = \limsup_n x_n = x$.

Allgemein gilt die folgende ansprechende Charakterisierung:

Satz (*Charakterisierung des Limes Inferior und Limes Superior*)

Sei $(x_n)_{n \in \mathbb{N}}$ eine beschränkte Folge in \mathbb{R} . Dann gilt:

- (a) $\limsup_n x_n$ ist der größte Häufungspunkt von $(x_n)_{n \in \mathbb{N}}$,
- (b) $\liminf_n x_n$ ist der kleinste Häufungspunkt von $(x_n)_{n \in \mathbb{N}}$.

Es folgt, dass eine beschränkte Folge mindestens einen Häufungspunkt besitzt. Der Charakterisierungssatz enthält also den Satz von Bolzano-Weierstraß (3.8). Stärker zeigt er, dass eine konvergente Folge genau einen Häufungspunkt und eine divergente beschränkte Folge mindestens zwei Häufungspunkte besitzt. Im divergenten Fall können noch weitere Häufungspunkte zwischen dem Limes Inferior und dem Limes Superior liegen (vgl. die Beispiele in 3.7).

Ausdehnung der Definition auf unbeschränkte Folgen

Der limsup kann allgemeiner für jede nach oben beschränkte Folge erklärt werden und der liminf für jede nach unten beschränkte Folge. So ist dann etwa der Limes Superior von $0, -1, 0, -2, 0, -3, \dots$ gleich 0 . Durch Verwendung der Symbole $+\infty$ und $-\infty$ können liminf und limsup sogar für jede Folge definiert werden. Die Definition bleibt gleich, wobei nun $+\infty$ und $-\infty$ bei der Bildung von Suprema und Infima zugelassen sind. Wir diskutieren dies in 3.12 genauer.

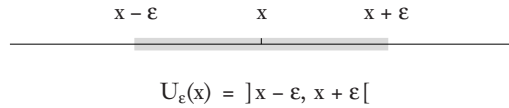
3.10 Offene Epsilon-Umgebungen

Definition (*offene ε -Umgebung*)

Für $x \in \mathbb{R}$ und $\varepsilon > 0$ setzen wir:

$$U_\varepsilon(x) = \{y \in \mathbb{R} \mid |y - x| < \varepsilon\}.$$

Die Menge $U_\varepsilon(x)$ heißt die (*offene*) ε -Umgebung des Punktes x .


Beispiele

$$(1) U_{1/2}(1) =]1/2, 3/2[, \quad U_n(0) =]-n, n[\quad \text{für alle } n \in \mathbb{N}.$$

$$(2)]a, b[= U_{(b-a)/2}((a+b)/2) \quad \text{für alle reellen Zahlen } a < b.$$

Auf den ersten Blick erscheint es vielleicht etwas übertrieben, für die Intervalle der Form $]x - \varepsilon, x + \varepsilon[$ einen eigenen Begriff und eine eigene Notation einzuführen. Doch der Schein trügt. Die Begriffe sind sehr gut geeignet, um Aussagen im Umfeld des Grenzwert- und Stetigkeitsbegriffs griffig formulieren zu können. Später bilden sie dann auch den Ausgangspunkt der topologischen Betrachtung der Analysis, die mehr Wert auf den Strukturunterschied zwischen $[a, b]$ und $]a, b[$ legt als auf die gemeinsame Länge $b - a$ der beiden Intervalle. Bevor wir einige Beispiele der „Umgebungssprache“ betrachten, wollen wir noch einige suggestive Sprechweisen präzisieren, die wir zuweilen schon verwendet haben. Damit soll betont werden, dass wir im Folgenden die exakte Ebene nicht verlassen.

Definition (*schließlich, immer wieder, unendlich oft*)

Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge. Dann gilt eine Eigenschaft \mathcal{E} für die Folge

schließlich, falls ein n_0 existiert, sodass $\mathcal{E}(x_n)$ für alle $n \geq n_0$ gilt,

immer wieder oder *unendlich oft*, falls $\mathcal{E}(x_n)$ für unendlich viele n gilt.



Schwarze Punkte markieren Zahlen n , für die $\mathcal{E}(x_n)$ gilt, weiße Punkte n , für die $\mathcal{E}(x_n)$ nicht gilt.

Beispiele

$(1/n)_{n \geq 1}$ ist schließlich kleiner als $1/2^{100}$.

$(0, 2, 1, 0, 3, 3, 3, \dots)$ ist schließlich konstant.

$(0, 2, 1, 0, 1, 2, 3, 4, 5, 6, \dots)$ ist schließlich streng monoton steigend.

$((-1)^n 1/2^n)_{n \in \mathbb{N}}$ ist immer wieder negativ.

$(0, 1, 0, 1, \dots)$ nimmt unendlich oft den Wert 1 an.

Eine Warnung:

Die Verneinung von „ \mathcal{E} gilt schließlich“ ist „ \mathcal{E} gilt unendlich oft nicht“, und nicht etwa „ \mathcal{E} gilt nur endlich oft“. Letztere Aussage ist die Verneinung von „ \mathcal{E} gilt unendlich oft“. Es kann zum Beispiel $\mathcal{E}(x_n)$ für alle geraden n gelten und für alle ungeraden n nicht gelten. Dann gilt \mathcal{E} zwar nicht schließlich, aber immerhin noch unendlich oft. Nach diesen logischen Vorbereitungen können wir nun zur Ernte der Früchte übergehen:

Umgebungsformulierung des Grenzwerts

Seien $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} und $x \in \mathbb{R}$. Dann sind äquivalent:

- (a) $\lim_n x_n = x$, d.h. $\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 \mid x - x_n \mid < \varepsilon$.
- (b) Für jede ε -Umgebung U von x gilt $x_n \in U$ schließlich.

Umgebungsformulierung von Häufungspunkten

Seien $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} und $x \in \mathbb{R}$. Dann sind äquivalent:

- (a) x ist ein Häufungspunkt von $(x_n)_{n \in \mathbb{N}}$, d.h., es gibt eine Teilfolge $(y_n)_{n \in \mathbb{N}}$ von $(x_n)_{n \in \mathbb{N}}$ mit $\lim_n y_n = x$.
- (b) Für jede ε -Umgebung U von x gilt $x_n \in U$ unendlich oft.

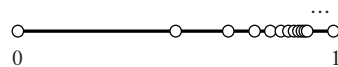
Ein x ist also der Grenzwert einer Folge, wenn jede ε -Umgebung von x schließlich alle Glieder einfängt. Und ein x ist ein Häufungspunkt, wenn die Folge jede ε -Umgebung von x immer wieder besucht. Entsprechend ist x kein Grenzwert, wenn es eine ε -Umgebung von x gibt, die die Folge immer wieder verlässt. Und ein x ist kein Häufungspunkt, wenn es eine ε -Umgebung gibt, die nur endlich oft besucht wird.

Die ε -Umgebungen eignen sich auch, um den Häufungspunktbegriff für Mengen einzuführen:

Definition (Häufungspunkte für Mengen)

Seien $P \subseteq \mathbb{R}$ und $x \in \mathbb{R}$. Dann heißt x ein *Häufungspunkt* von P , falls für alle $\varepsilon > 0$ die Menge $P \cap U_\varepsilon(x)$ unendlich ist.

$$P = [0, 1[- \{1 - 1/n \mid n \geq 1\}$$



Jedes $x \in [0, 1]$ ist Häufungspunkt von P .

Beispiel

0 ist ein Häufungspunkt von $]0, 1[$, $[0, 1]$, $\{1/2^n \mid n \in \mathbb{N}\}$, aber kein Häufungspunkt von $\{0, 1\}$. Dagegen ist 0 ein Häufungspunkt der Folge $0, 1, 0, 1, 0, 1, \dots$

Ein x kann also Häufungspunkt von $(x_n)_{n \in \mathbb{N}}$ sein, ohne ein Häufungspunkt der Menge $P = \{x_n \mid n \in \mathbb{N}\}$ zu sein. Nimmt die Folge aber jeden Wert höchstens endlich oft an, so sind die Häufungspunkte der Folge genau die Häufungspunkte von P .

Wie für Folgen gilt:

Satz von Bolzano-Weierstraß für Mengen

Jede unendliche beschränkte Teilmenge von \mathbb{R} besitzt einen Häufungspunkt.

3.11 Konvergenz in den komplexen Zahlen

Definition (Konvergenz in \mathbb{C})

Die Konvergenz einer Folge $(z_n)_{n \in \mathbb{N}}$ gegen ein $z \in \mathbb{C}$ wird wie im reellen Fall durch die folgende Konvergenzbedingung definiert:

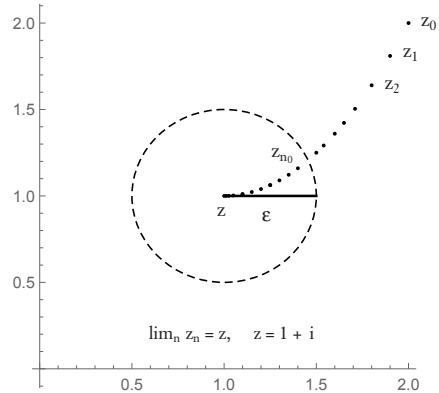
$$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 |z_n - z| < \varepsilon. \quad (\text{Konvergenzbedingung für } z)$$

Wir schreiben dann wieder $\lim_{n \rightarrow \infty} z_n = z$ oder $\lim_n z_n = z$.

Für alle komplexen Zahlen $w = (x, y)$ hatten wir definiert:

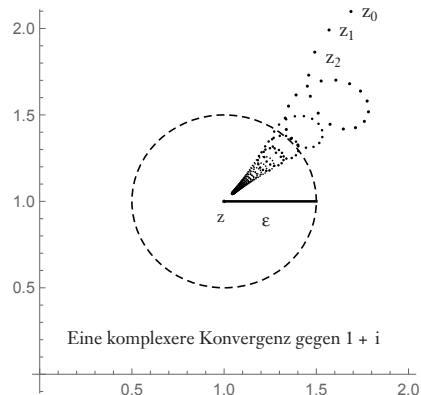
$$|w| = \sqrt{\operatorname{Re}(w)^2 + \operatorname{Im}(w)^2} = \sqrt{x^2 + y^2} \in \mathbb{R}.$$

In der Konvergenzbedingung für z ist also die Zahl $|z_n - z|$ reell und genauer der euklidische Abstand der Punkte z_n und z in der Ebene. Der Allquantor läuft wie früher nur über reelle Zahlen $\varepsilon > 0$, komplexe Zahlen tauchen nur innerhalb der Betragsstriche auf.



Beispiele

- (1) $\lim_n (n + i)/(n + 1) = 1$.
- (2) $\lim_n i^n/2^n = 0$.
- (3) $\lim_n (ni)$ und $\lim_n i^n$ existieren nicht.
- (4) Gilt $\lim_n x_n = x$ in \mathbb{R} , so gilt auch $\lim_n x_n = x$ in \mathbb{C} . Jede reelle Zahl x ist ja durch die Identifizierung von x und $(x, 0)$ auch eine komplexe Zahl, und deswegen ist eine Folge in \mathbb{R} auch eine Folge in \mathbb{C} .



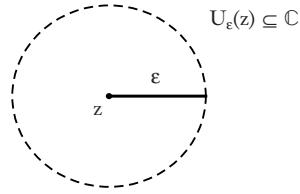
Ein Großteil unserer Überlegungen für Folgen in \mathbb{R} überträgt sich ohne Mühe und Modifikation auf Folgen in \mathbb{C} . Unterschiede ergeben sich nur durch die zweidimensionale Struktur und die fehlende Ordnung von \mathbb{C} .

- (a) Die Limesregeln für Folgen gelten auch für Folgen in \mathbb{C} .
- (b) Die Definition einer Cauchy-Folge in \mathbb{C} kann wörtlich übernommen werden. Die in \mathbb{C} konvergenten Folgen sind wieder genau die Cauchy-Folgen.
- (c) Häufungspunkte für Folgen in \mathbb{C} werden wie in \mathbb{R} über konvergente Teilfolgen definiert. Der Satz von Bolzano-Weierstraß gilt: Jede beschränkte Folge in \mathbb{C} besitzt einen Häufungspunkt. Dabei heißt eine Folge $(z_n)_{n \in \mathbb{N}}$ in \mathbb{C} beschränkt, falls es ein $x \in \mathbb{R}$ gibt mit $|z_n| < x$ für alle n .

- (d) Die Begriffe \liminf und \limsup stehen in \mathbb{C} nicht zur Verfügung.
- (e) Für $z \in \mathbb{C}$ und $\varepsilon > 0$ wird die offene ε -Umgebung von z definiert durch:

$$U_\varepsilon(z) = \{w \in \mathbb{C} \mid |w - z| < \varepsilon\},$$

also als die offene Kreisscheibe mit Mittelpunkt z und Radius ε . Die Umgebungsformulierungen von „Grenzwert“ und „Häufungspunkt“ können wörtlich übernommen werden. So gilt $\lim_n z_n = z$ in \mathbb{C} genau dann, wenn für jede offene ε -Umgebung von z die Folge schließlich in dieser Umgebung liegt.



Die Grenzwertbestimmung in den komplexen Zahlen lässt sich auf die Bestimmung von zwei reellen Grenzwerten zurückführen:

Satz (*koordinatenweise Konvergenz für Folgen in \mathbb{C}*)

Eine Folge $(z_n)_{n \in \mathbb{N}}$ in \mathbb{C} konvergiert genau dann gegen $z \in \mathbb{C}$, wenn die Folge $(\operatorname{Re}(z_n))_{n \in \mathbb{N}}$ gegen $\operatorname{Re}(z)$ und die Folge $(\operatorname{Im}(z_n))_{n \in \mathbb{N}}$ gegen $\operatorname{Im}(z)$ konvergieren.

Für alle komplexen Zahlen z gilt $z = (\operatorname{Re}(z), \operatorname{Im}(z))$. Im Konvergenzfall gilt also:

$$\lim_n z_n = \lim_n (\operatorname{Re}(z_n), \operatorname{Im}(z_n)) = (\lim_n \operatorname{Re}(z_n), \lim_n \operatorname{Im}(z_n)).$$

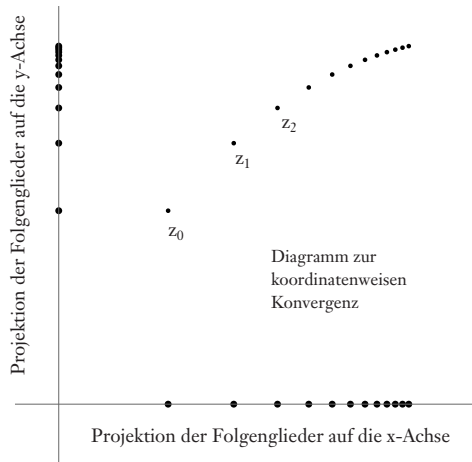
Diese Aussage ist ein weiteres Beispiel für eine Vertauschungsregel. Die Grenzwertbildung in \mathbb{C} respektiert den Aufbau $z = (x, y)$ einer komplexen Zahl. Man darf den Limes in das geordnete Paar reinziehen.

Die Diskussion zeigt, dass der durch die Konvergenzbedingung eingefangene Grenzwertbegriff für Folgen sehr allgemein ist und lediglich einen Abstandsbegriff benötigt. In der Analysis werden später sog. *metrische Räume* betrachtet, die eine Menge X mit einer Abstandsfunktion

$$d : X \times X \rightarrow \mathbb{R}$$

ausstatten. Die reelle Zahl $d(x, y)$ heißt dann der Abstand zwischen x und y in X . Die Konvergenz von Folgen $(x_n)_{n \in \mathbb{N}}$ in X wird nun durch die Konvergenzbedingung wie in \mathbb{R} und \mathbb{C} erklärt, wobei nun die Abstandsfunktion verwendet wird, da „Minus“ und „Betrag“ im Allgemeinen in X nicht mehr zur Verfügung stehen:

$$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 d(x_n, x) < \varepsilon. \quad (\text{metrische Konvergenzbedingung für } x)$$



3.12 Die Unendlichkeitssymbole

Definition (*unendliche Grenzwerte, Suprema und Infima in \mathbb{R}*)

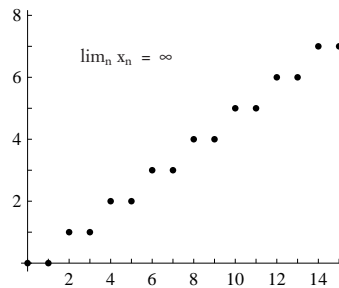
(a) Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} . Wir schreiben $\lim_{n \rightarrow \infty} x_n = \lim_n x_n = +\infty$, falls

$$\forall x \exists n_0 \forall n \geq n_0 \quad x_n \geq x. \quad (\text{uneigentliche Konvergenzbedingung f\"ur } +\infty)$$

Analog schreiben wir $\lim_{n \rightarrow \infty} x_n = \lim_n x_n = -\infty$, falls

$$\forall x \exists n_0 \forall n \geq n_0 \quad x_n \leq x. \quad (\text{uneigentliche Konvergenzbedingung f\"ur } -\infty)$$

Wir nennen dann $(x_n)_{n \in \mathbb{N}}$ *uneigentlich konvergent* oder auch *bestimmt divergent* gegen $+\infty$ bzw. $-\infty$.



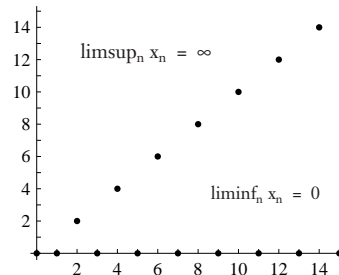
(b) Für $X \subseteq \mathbb{R}$ schreiben wir $\sup(X) = +\infty$, falls X nach oben unbeschränkt ist, d. h. falls für alle $y \in \mathbb{R}$ ein $x \in X$ existiert mit $y \leq x$. Analog schreiben wir $\inf(X) = -\infty$, falls X nach unten unbeschränkt ist.

(c) Wir definieren für jede (beschränkte oder unbeschränkte) Folge $(x_n)_{n \in \mathbb{N}}$:

$$\limsup_n x_n = \inf_{n \in \mathbb{N}} \sup_{m \geq n} x_m,$$

$$\liminf_n x_n = \sup_{n \in \mathbb{N}} \inf_{m \geq n} x_m,$$

wobei wir noch vereinbaren, dass $\inf(\{+\infty\}) = +\infty$ und $\sup(\{-\infty\}) = -\infty$.



Wir nehmen die Unendlichkeitssymbole in unseren Kalkül auf, indem wir vereinbaren:

$$-\infty < \infty, \quad -\infty < x, \quad x < \infty,$$

$$x + \infty = \infty, \quad x - \infty = -\infty, \quad x/\infty = x/(-\infty) = 0 \quad \text{für alle } x \in \mathbb{R},$$

$$x \cdot +\infty = +\infty, \quad x \cdot -\infty = -\infty \quad \text{für alle } x \in]0, \infty],$$

$$x \cdot \infty = -\infty, \quad x \cdot -\infty = \infty \quad \text{für alle } x \in [-\infty, 0[.$$

Oft nennt man die mit diesen Rechenregeln ausgestattete Menge

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$$

die *erweiterten reellen Zahlen*. $\overline{\mathbb{R}}$ ist im Gegensatz zu \mathbb{R} kein Körper, da die folgenden Ausdrücke nicht definiert sind:

$$+\infty - \infty, \quad -\infty + \infty, \quad +\infty \cdot 0, \quad -\infty \cdot 0. \quad (\text{undefinierte Ausdrücke in } \overline{\mathbb{R}})$$

Die Limesregeln gelten in erweiterter Form, etwa

$$\lim_n x_n = +\infty \text{ und } \lim_n y_n = y \text{ impliziert } \lim_n (x_n + y_n) = +\infty + y = +\infty,$$

$$\lim_n x_n = \infty \text{ und } \lim_n y_n = -\infty \text{ impliziert } \lim_n (x_n \cdot y_n) = \infty \cdot -\infty = -\infty.$$

Eine uneigentlich konvergente Folge ist nach wie vor divergent. Die Sprechweisen „uneigentlich konvergent“ und „bestimmt divergent“ sind sicher nicht ganz glücklich.

Es gilt $\limsup_n x_n = \infty$ genau dann, wenn $\sup_n x_n = \infty$. Dagegen ist $\limsup_n x_n = -\infty$ äquivalent zu $\lim_n x_n = -\infty$. Analoge Aussagen gelten für \liminf .

Beispiele

$$(1) \lim_n n^2 = +\infty, \lim_n -2^n = -\infty, \limsup_n (-1)^n n = \infty, \liminf_n (-1)^n n = -\infty.$$

$$(2) \text{ Die Folge } (x_n)_{n \in \mathbb{N}} = (1, 0, 3, 2, 5, 4, \dots) \text{ konvergiert uneigentlich gegen } +\infty. \\ \text{Es gilt } \lim_n x_n = \limsup_n x_n = \liminf_n x_n = \infty.$$

$$(3) \text{ Ist } r \geq 0 \text{ und } |x_n| \leq r \text{ unendlich oft, so konvergiert } (x_n)_{n \in \mathbb{N}} \text{ nicht uneigentlich.} \\ \text{Nach dem Satz von Bolzano-Weierstraß besitzt } (x_n)_{n \in \mathbb{N}} \text{ eine konvergente} \\ \text{Teilfolge. Denn die Teilfolge aller } x_n \text{ mit } |x_n| \leq r \text{ ist beschränkt in } \mathbb{R}, \text{ und eine} \\ \text{konvergente Teilfolge dieser Folge ist auch eine Teilfolge von } (x_n)_{n \in \mathbb{N}}.$$

Wir illustrieren noch, warum $+\infty - \infty$, $-\infty + \infty$, $+\infty \cdot 0$ und $-\infty \cdot 0$ undefiniert bleiben:

Beispiele

$$(1) \lim_n -n = -\infty, \quad \lim_n 2^n = +\infty, \quad \lim_n (-n + 2^n) = +\infty,$$

$$(2) \lim_n -n = -\infty, \quad \lim_n n = +\infty, \quad \lim_n (-n + n) = 0,$$

$$(3) \lim_n n = +\infty, \quad \lim_n 1/2^n = 0, \quad \lim_n (n \cdot 1/2^n) = 0,$$

$$(4) \lim_n 2^n = +\infty, \quad \lim_n 1/(n+1) = 0, \quad \lim_n (2^n \cdot 1/(n+1)) = +\infty.$$

In den komplexen Zahlen betrachtet man nur ein Unendlichkeitssymbol ∞ . Man definiert für eine Folge $(z_n)_{n \in \mathbb{N}}$ in \mathbb{C} :

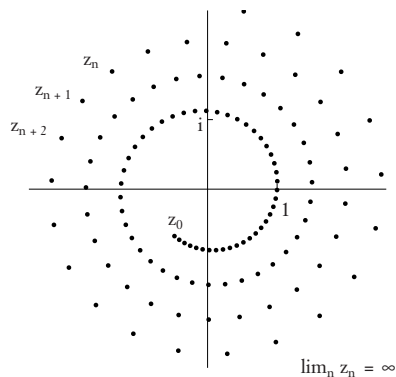
$$\lim_n z_n = \infty, \text{ falls } \forall x \geq 0 \exists n_0 \forall n \geq n_0 |z_n| \geq x.$$

Die Bedingung besagt, dass die Folgenglieder z_n schließlich außerhalb jeder beliebig großen Kreisscheibe um den Nullpunkt liegen.

Beispiele

$$\text{In } \mathbb{C} \text{ gilt } \lim_n (ni^n) = \infty,$$

$$\lim_n ((-1)^n n) = \infty, \quad \lim_n -n = \infty.$$



Die Folge $(-1)^n n$ konvergiert nicht uneigentlich in \mathbb{R} , und in \mathbb{R} ist $\lim_n -n = -\infty$. Die Notation für \mathbb{C} setzt also die für \mathbb{R} nicht fort. Oft verwendet man deswegen die Symbole $-\infty$ und $+\infty$ (statt ∞) für \mathbb{R} und ∞ für \mathbb{C} . Das ist genauer, wenn auch etwas schwerfälliger.

4. Kapitel

Reihen

4.1 Unendliche Reihen

Definition (*Partialsumme, unendliche Reihe, Doppelbedeutung von $\sum_{n \in \mathbb{N}} x_n$*)

Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} . Dann definieren wir für alle n die n -te *Partialsumme* s_n der Folge $(x_n)_{n \in \mathbb{N}}$ durch

$$s_n = x_0 + \dots + x_n = \sum_{k \leq n} x_k.$$

Die Folge $(s_n)_{n \in \mathbb{N}}$ heißt die durch die Folge $(x_n)_{n \in \mathbb{N}}$ *definierte (unendliche) Reihe* in \mathbb{R} . Die Folgenglieder x_n nennen wir auch die *Summanden* der Reihe.

(a) Konvergiert die Folge $(s_n)_{n \in \mathbb{N}}$ der Partialsummen, so setzen wir

$$\sum_{n \in \mathbb{N}} x_n = \lim_n s_n = \lim_n \sum_{k \leq n} x_k$$

und nennen $\sum_{n \in \mathbb{N}} x_n$ die (*unendliche*) *Summe* von $(x_n)_{n \in \mathbb{N}}$.

(b) Unabhängig von der Konvergenz von $(s_n)_{n \in \mathbb{N}}$ bezeichnen wir auch die Folge $(s_n)_{n \in \mathbb{N}}$ der Partialsummen selbst mit $\sum_{n \in \mathbb{N}} x_n$, d. h., wir setzen

$$\sum_{n \in \mathbb{N}} x_n = (s_n)_{n \in \mathbb{N}} = (\sum_{k \leq n} x_k)_{n \in \mathbb{N}}.$$

Statt $\sum_{n \in \mathbb{N}} x_n$ schreiben wir gleichwertig auch $\sum_{n=0}^{\infty} x_n$, $\sum_{n \geq 0} x_n$, $\sum_n x_n$ oder $x_0 + x_1 + x_2 + \dots + x_n + \dots$

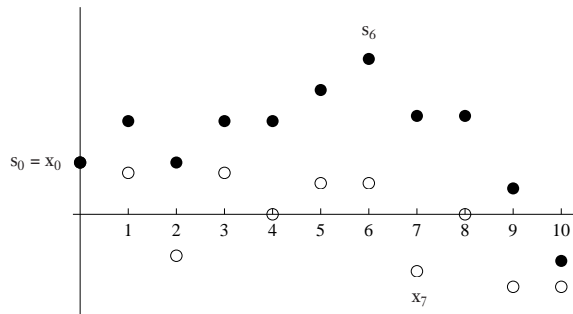
Den Ausgangspunkt bildet eine Folge $(x_n)_{n \in \mathbb{N}}$ in \mathbb{R} . Unser Ziel ist, alle x_n aufzusummieren, also die endliche Summenbildung

$$\sum_{k \leq n} x_k = x_0 + \dots + x_n$$

ins Unendliche zu erweitern. Wir können dann Aussagen wie

$$1/2 + 1/4 + 1/8 + \dots = 1$$

nicht nur an einem Tortendiagramm erläutern, sondern mit den Methoden der Analysis beweisen. Es ist naheliegend, für diese Aufgabe den entwickelten Grenzwertbegriff einzusetzen. Dieser Begriff steht uns aber nur für Folgen und noch nicht für Summen zur Verfügung. Wir ordnen deswegen der Folge $(x_n)_{n \in \mathbb{N}}$ eine Hilfsfolge $(s_n)_{n \in \mathbb{N}}$ zu.



Partialsummen $s_n = \sum_{1 \leq k \leq n} x_k$ (schwarze Punkte)
für gegebene Summanden x_n (weiße Punkte)

$$s_0 = x_0,$$

$$s_1 = x_0 + x_1,$$

$$s_2 = x_0 + x_1 + x_2,$$

...

$$s_n = x_0 + x_1 + \dots + x_n = \sum_{k \leq n} x_k.$$

Für unsere Hilfsfolge können wir nun die entwickelte Konvergenztheorie einsetzen. Konvergiert $(s_n)_{n \in \mathbb{N}}$, so setzen wir

$$\sum_n x_n = \lim_n s_n = \lim_n \sum_{k \leq n} x_k. \quad (\sum_n x_n \text{ als Grenzwert})$$

Statt der Sigma-Notation $\sum_{n \in \mathbb{N}} x_n$ können wir nun auch wieder informal

$$x_0 + x_1 + \dots + x_n + \dots$$

schreiben, ganz so, wie wir statt $(x_n)_{n \in \mathbb{N}}$ oft auch $x_0, x_1, \dots, x_n, \dots$ schreiben.

Zu den Schwierigkeiten, eine unendliche Summe als Grenzwert von Partialsummen zu begreifen, tritt nun noch die Doppelbedeutung der Notation $\sum_n x_n$. Wir setzen:

$$\sum_n x_n = (s_n)_{n \in \mathbb{N}}. \quad (\sum_n x_n \text{ als Folge der Partialsummen})$$

Wir fassen zusammen:

Was bedeutet „ $\sum_n x_n$ “?

1. Immer die Folge $(s_n)_{n \in \mathbb{N}}$ der Partialsummen von $(x_n)_{n \in \mathbb{N}}$.
2. Den Grenzwert von $(s_n)_{n \in \mathbb{N}}$, falls dieser existiert.

Die Doppelbedeutung führt zu Aussagen wie „ $\sum_{n \in \mathbb{N}} x_n$ ist der Limes von $\sum_{n \in \mathbb{N}} x_n$ “. Sie bewährt sich aber im Alltag. Die Formulierung „Sei $\sum_{n \in \mathbb{N}} x_n$ eine Reihe in \mathbb{R} “ ist einfacher und suggestiver als die Formulierung „Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} , und sei $(s_n)_{n \in \mathbb{N}}$ die Folge ihrer Partialsummen“.

Beispiele

- (1) $1/2 + 1/4 + 1/8 + \dots$ ist die Reihe $\sum_{n \geq 1} 1/2^n$. Die Reihe ist konvergent und es gilt $\sum_{n \geq 1} 1/2^n = 1$. Derartige sog. geometrische Reihen werden wir gleich noch genauer besprechen.

- (2) Die Reihe $\sum_n (-1)^n = 1 - 1 + 1 - 1 \pm \dots$ divergiert, denn hier gilt $s_n = 1$ für gerade und $s_n = 0$ für ungerade n . Der Streitfall

$$\text{„}\sum_n (-1)^n = (1 - 1) + (1 - 1) + (1 - 1) + \dots = 0\text{“}$$

$$\text{„}\sum_n (-1)^n = 1 + (-1 + 1) + (-1 + 1) + (-1 + 1) = 1\text{“}$$

wird also nicht durch den Kompromiss „ $\sum_n (-1)^n = 1/2$ “ gelöst. Die Reihe $\sum_n (-1)^n = (1, 0, 1, 0, 1, 0, \dots)$ ist divergent.

- (3) Die Reihe $x_0 + \dots + x_m + 0 + 0 + 0 + \dots$ konvergiert, da die Folge ihrer Partialsummen schließlich konstant gleich s_m ist. Es gilt $\sum_n x_n = s_m = \sum_{k \leq m} x_k$.

Auch die uneigentliche Konvergenz kann auf Reihen übertragen werden, sodass etwa $\sum_n x_n = +\infty$ bedeutet, dass $\lim_n s_n = +\infty$. Es gilt zum Beispiel $\sum_n n = +\infty$ und $\sum_n -n = -\infty$, während $\sum_n (-n)^n$ weder eigentlich noch uneigentlich existiert.

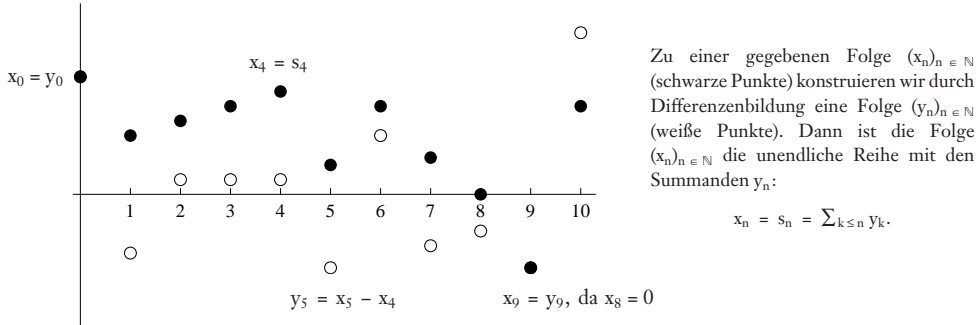
4.2 Folgen versus Reihen

Satz (Folgen als Reihen)

Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} . Weiter sei

$$y_0 = x_0 \text{ und } y_n = x_n - x_{n-1} \text{ für alle } n \geq 1.$$

Dann gilt $\sum_n y_n = (x_n)_{n \in \mathbb{N}}$ und im Fall der Konvergenz also $\sum_n y_n = \lim_n x_n$.



Bezeichnet $(s_n)_{n \in \mathbb{N}}$ die Folge der Partialsummen der Reihe $\sum_n y_n$, so gilt in der Tat

$$s_0 = y_0 = x_0,$$

$$s_1 = y_0 + y_1 = x_0 + (x_1 - x_0) = x_1,$$

$$s_2 = y_0 + y_1 + y_2 = x_0 + (x_1 - x_0) + (x_2 - x_1) = x_2,$$

...

$$s_n = x_0 + \sum_{1 \leq k \leq n} (x_k - x_{k-1}) = x_n,$$

...

Die Folgen $(x_n)_{n \in \mathbb{N}}$ und $(s_n)_{n \in \mathbb{N}}$ sind also gleich und damit ist die Folge $(x_n)_{n \in \mathbb{N}}$ identisch mit der Reihe $\sum_n y_n$. Die Folge $(y_n)_{n \in \mathbb{N}}$ beschreibt die Zuwächse der Folge $(x_n)_{n \in \mathbb{N}}$, wobei diese Zuwächse positiv, negativ oder gleich 0 sein können. Und durch Summieren der Zuwächse erhalten wir die Folgenglieder zurück. Ein Beispiel aus dem „unendlichen Alltag“:

Beispiel

Ist $(x_n)_{n \in \mathbb{N}}$ eine Folge von Kontoständen, so ist die im Satz definierte Folge $(y_n)_{n \in \mathbb{N}}$ die Folge der entsprechenden Einzahlungen und Abbuchungen. Weiter ist jeder Kontostand x_n die Summe aller Einzahlungen und Abbuchungen y_k mit $k \leq n$.

So simpel diese Überlegung ist, so stellt sie doch bereits eine diskrete Version des Hauptsatzes der Integral- und Differentialrechnung dar. Aus der Summe von Differenzen wird später ein Integral von Differentialquotienten (vgl. Kapitel 8).

Der Satz lässt sich manchmal anwenden, um Reihen zu berechnen:

Beispiel

Sei $x_n = n/(n+1)$ für alle n , und sei $(y_n)_{n \in \mathbb{N}}$ wie im Satz. Dann gilt $y_0 = x_0 = 0$ und

$$\begin{aligned} y_n &= x_n - x_{n-1} = \\ n/(n+1) - (n-1)/n &= \\ 1/(n(n+1)). \end{aligned}$$

für alle $n \geq 1$. Folglich ist

$$\begin{aligned} \sum_{n \geq 1} 1/(n(n+1)) &= \sum_n y_n = \\ \lim_n x_n &= \lim_n n/(n+1) = 1. \end{aligned}$$

Hier wurde das Pferd von hinten aufgezäumt. Wir haben die Summe

$$\begin{aligned} &1/2 + 1/6 + 1/12 + 1/20 + \\ &1/30 + \dots + 1/(n(n+1)) + \dots = 1 \end{aligned}$$

gefunden, indem wir ihre Partialsummen $(n/(n+1))_{n \geq 1}$ vorgegeben haben. Von der Reihe $\sum_{n \geq 1} 1/(n(n+1))$ wussten wir zunächst gar nichts.

Zusammenfassend halten wir fest:

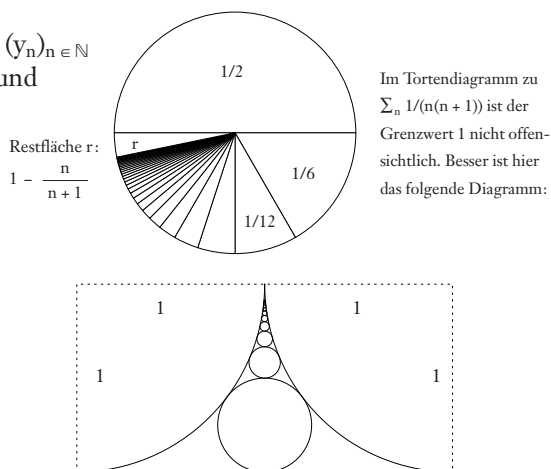
Jede Reihe ist (per Definition) eine Folge, und jede Folge ist als Reihe darstellbar.

Diese Beobachtung zeigt, dass die Beherrschung von Folgen zur Beherrschung von Reihen führt und umgekehrt. Die beiden Konzepte sind also prinzipiell gleichwertig. Dennoch brauchen wir in der Analysis beide Begriffe, denn es tauchen in natürlicher Weise sowohl Folgen als auch Reihen auf, und der Prozess der Übersetzung wäre auf Dauer zu mühsam.

Wir stellen die Notationen und Sprechweisen für Reihen tabellarisch zusammen.

	Folgen	Reihen
Notation	$(x_n)_{n \in \mathbb{N}}$	$\sum_n x_n$
Formal	$f: \mathbb{N} \rightarrow \mathbb{R}, f(n) = x_n$	$(s_n)_{n \in \mathbb{N}}, s_n = \sum_{k \leq n} x_k$
Informal	$x_0, x_1, x_2, \dots, x_n, \dots$	$x_0 + x_1 + x_2 + \dots + x_n + \dots$
Grenzwert	$\lim_n x_n$	$\sum_n x_n$

Der Begriff „Reihe“ wird in der Mathematik ganz anders gebraucht als in der Umgangssprache. Eine mathematische Reihe ist, als Zahl, eine unendliche Summe und, als Untersuchungsobjekt, eine Folge von Partialsummen. Umgangssprachlich hat „sich in einer Reihe aufstellen“, „Perlen aufreihen“, „Buchreihe“ mehr mit Ordnung als mit Summation zu tun. Auch eine „Messreihe“ bezeichnet eine Folge und keine Summe. Die Bedeutungen der Begriffe in der Mathematik sind wie immer durch die Definitionen festgelegt. Sie sind vielleicht nicht immer ganz glücklich, aber man gewöhnt sich in der Regel schnell daran.



4.3 Die geometrische Reihe

Satz (Konvergenz der geometrischen Reihe)

Sei $x \in]-1, 1[$. Dann konvergiert die sog. geometrische Reihe

$$\sum_n x^n = 1 + x^1 + x^2 + \dots + x^n + \dots$$

und es gilt:

$$\sum_n x^n = \frac{1}{1-x}, \quad \sum_{n \geq 1} x^n = \frac{x}{1-x}.$$

(Formeln für die geometrische Reihe)

Die geometrische Reihe gehört zu den wichtigsten Reihen der Analysis. Sie ist immer wieder im Einsatz, etwa beim Beweis der Konvergenzkriterien für Reihen oder bei der Bestimmung des Konvergenzradius von Potenzreihen. Um so erfreulicher ist, dass sie zu den recht wenigen Reihen gehört, deren Summe man leicht ausrechnen kann. Die Partialsummen

$$s_n = \sum_{k \leq n} x^k = x^0 + x^1 + \dots + x^n$$

lassen sich direkt bestimmen. Für alle reellen Zahlen x , $x \neq 1$, gilt:

$$s_n = \frac{1 - x^{n+1}}{1 - x}. \quad (\text{endliche geometrische Reihe})$$

Dies kann man durch vollständige Induktion oder durch Ausmultiplizieren des Produkts

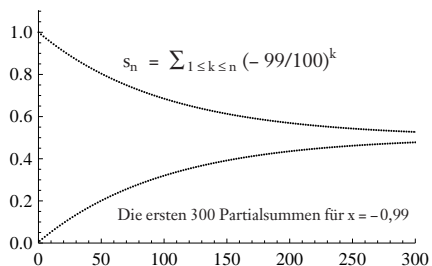
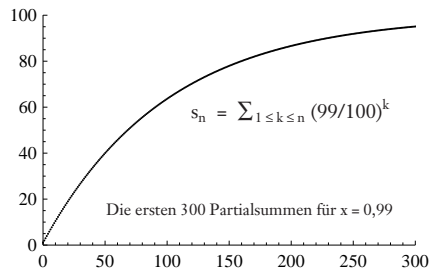
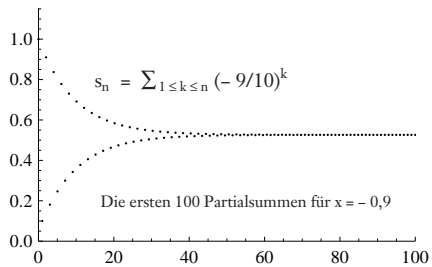
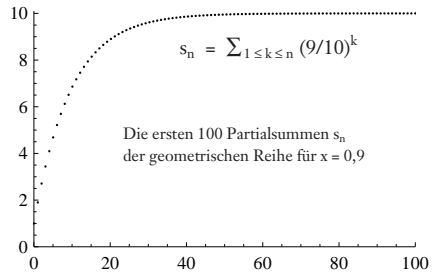
$$(1 + x + x^2 + \dots + x^n)(1 - x)$$

beweisen. Das Ausmultiplizieren hinterlässt eine Teleskopsumme, in der sich nur die beiden Terme 1 und $-x^{n+1}$ nicht gegenseitig weglöschen.

Nach der endlichen Summenformel gilt also für alle x mit $|x| < 1$:

$$\sum_n x^n = \lim_n s_n = \lim_n \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x},$$

wobei wir neben den Regeln für die Folgenarithmetik benutzen, dass $\lim_n x^n = 0$ für alle x mit $|x| < 1$. Damit ist also $\sum_n x^n$ bestimmt. Die Summe $\sum_{n \geq 1} x^n$ erhält man durch



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

oder alternativ durch eine analoge Überlegung, die auf der Berechnung

$$\sum_{1 \leq k \leq n} x^k = \frac{x - x^{n+1}}{1-x} \quad \text{für alle } x \neq 1$$

beruht.

Beispiele

$$(1) \sum_n (1/2)^n = 1/(1 - 1/2) = 2/1 = 2,$$

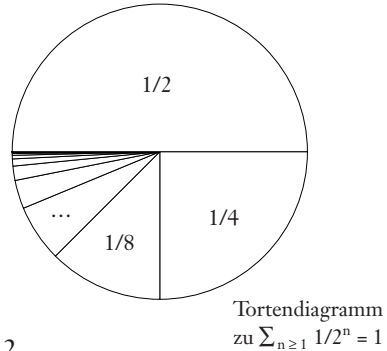
$$\sum_n (1/k)^n = k/(k-1) \quad \text{für alle } k \geq 2.$$

$$(2) \sum_n (-1/2)^n = 1/(1 + 1/2) = 2/3,$$

$$\sum_n (-1/k)^n = 1/(1 + 1/k) = k/(k+1), \quad k \geq 2.$$

$$(3) \sum_n q^n = 1/(1-q) = n/(n-m) \quad \text{für alle } q = m/n \in]-1, 1[.$$

$$(4) \sum_n x^k \text{ divergiert für alle } x \text{ mit } |x| \geq 1. \text{ Für alle } x \geq 1 \text{ gilt } \sum_n x^k = +\infty.$$



Die Menge aller x , für die die Reihe $\sum_n x^n$ konvergiert, bildet also ein Intervall. Dieses Verhalten ist typisch für die sog. *Potenzreihen* der Form $\sum_n a_n x^n$, die wir später besprechen werden (siehe 5.12). Die geometrische Reihe entspricht der besonders einfachen Potenzreihe mit den Koeffizienten $a_n = 1$ für alle n .

Unser Bestand an „berechenbaren“ Reihen lässt sich noch erweitern:

Beispiele

(1) Die Rechenregeln für die Exponentiation liefern für alle $x \in]-1, 1[$:

$$\sum_n x^{2n} = \sum_n (x^2)^n = \frac{1}{(1-x^2)}, \quad \sum_n x^{2n+1} = x \sum_n x^{2n} = \frac{x}{(1-x^2)}.$$

(2) Die Berechnung der Partialsummen der geometrischen Reihe unter Verwendung von Teleskopsummen lässt sich beispielsweise wie folgt variieren. Sei $x \neq 1$. Dann gilt für alle $n \geq 1$:

$$\begin{aligned} (x + 2x^2 + \dots + nx^n)(1-x)(1-x) &= \\ (x + x^2 + x^3 + \dots + x^n - nx^{n+1})(1-x) &= x - x^{n+1} - nx^{n+1} + nx^{n+2}. \end{aligned}$$

Folglich gilt für alle x mit $|x| < 1$:

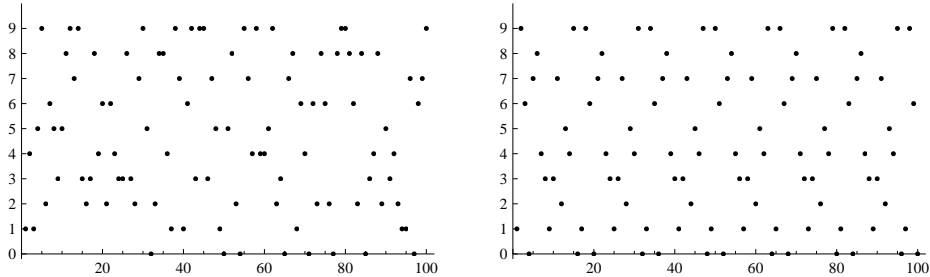
$$\sum_{n \geq 1} nx^n = \lim_n \frac{x - x^{n+1} - nx^{n+1} + nx^{n+2}}{(1-x)^2} = \frac{x}{(1-x)^2}.$$

Für $x = 1/2$ erhalten wir also $1/2 + 2/4 + 3/8 + 4/16 + 5/32 + \dots = 2$.

4.4 Dezimaldarstellungen als Reihen

Satz (Dezimaldarstellungen als Reihen)

■ Für $x = m, d_1 d_2 \dots d_n \dots \geq 0$ in Dezimaldarstellung gilt $x = m + \sum_{n \geq 1} d_n / 10^n$.



Die ersten 100 Dezimalziffern von $\pi = 3,1415\dots$ (links) und von $1/51 = 0,\overline{1960784313725490}$ (rechts)

Wir hatten bereits in Sektion 4.4 Dezimaldarstellungen über Suprema eingeführt. Für alle $m \in \mathbb{N}$ und Folgen $(d_n)_{n \geq 1}$ mit $d_n \in \{0, \dots, 9\}$ für alle $n \geq 1$ hatten wir definiert:

$$m, d_1 d_2 \dots d_n \dots = \sup_{n \geq 1} m, d_1 d_2 \dots d_n = \sup(\{m, d_1 d_2 \dots d_n \mid n \geq 1\}),$$

wobei die endlichen Dezimalbrüche $m, d_1 \dots d_n$ definiert sind durch:

$$m, d_1 \dots d_n = m + \sum_{k \leq n} d_k / 10^k = m + d_1 / 10 + d_2 / 10^2 + \dots + d_n / 10^n.$$

Es ist nun sehr natürlich, auch die unendlichen Dezimaldarstellungen als unendliche Summen zu betrachten. Die Aussage des Satzes ist dabei nicht schwer zu beweisen. Die Reihe $\sum_{n \geq 1} d_n / 10^n$ besitzt die Partialsummenfolge $(s_n)_{n \geq 1} = (0, d_1 \dots d_n)_{n \geq 1}$. Diese Folge ist monoton steigend und beschränkt. Daher gilt

$$\sum_{n \geq 1} d_n / 10^n = \lim_n s_n = \sup_n s_n = 0, d_1 d_2 \dots d_n \dots,$$

wobei das erste Gleichheitszeichen nach Definition der unendlichen Summe gilt, das zweite aufgrund der Monotonie der Folge $(s_n)_{n \in \mathbb{N}}$ und das dritte nach Definition der unendlichen Dezimaldarstellung. Folglich ist

$$m + \sum_{n \geq 1} d_n / 10^n = m + 0, d_1 d_2 \dots d_n \dots = m, d_1 d_2 \dots d_n \dots$$

Oft wird der Ausdruck „ $m + \sum_{n \geq 1} d_n / 10^n$ “ zur Definition von „ $m, d_1 d_2 \dots d_n \dots$ “ verwendet. Die Definition als Supremum benötigt den Begriff der unendlichen Reihe nicht und ist in diesem Sinne einfacher. Aber erst die unendlichen Reihen liefern das richtige Werkzeug zur Untersuchung der Dezimaldarstellung. Wir erinnern hierzu an einige Begriffe. Eine Dezimaldarstellung heißt *periodisch*, falls sie von der Form

$$m, a_1 \dots a_k \overline{b_1 \dots b_p} = m, a_1 \dots a_k b_1 \dots b_p b_1 \dots b_p b_1 \dots b_p \dots$$

ist, mit $k \geq 0, p \geq 1$. Der Ziffernblock $b_1 \dots b_p$ heißt eine *Periode* der Länge p der Darstellung.

Beispiel

12, 21, 1212, 2121, ... sind Perioden der Darstellung $0,121212\dots$

Man kann Perioden minimaler Länge bevorzugen (im Beispiel also 12 und 21 vor 1212 und 2121) und diejenige Periode auszeichnen, die zuerst einsetzt (im Beispiel also 12 und nicht 21). In diesem Sinne gibt es eine *kanonische Periode*. Man kann sie verkürzt die Periode nennen, wenn man sich der verschiedenen Zyklen bewusst ist.

Ist die Darstellung von der Form $m, \overline{b_1 \dots b_p}$ (also $k = 0$), so heißt sie *reinperiodisch*. Andernfalls heißt sie *gemischtperiodisch*. So ist zum Beispiel

$$1/7 = 0,\overline{142857} \text{ reinperiodisch und } 1/12 = 0,08\overline{3} \text{ gemischtperiodisch.}$$

Wegen $m, d_1 \dots d_n = m, d_1 \dots d_n \overline{0}$ ist jede endliche Darstellung gleich einer periodischen unendlichen Darstellung. Allgemeiner besitzen genau die rationalen Zahlen periodische Dezimaldarstellungen. Die geometrische Reihe zeigt, dass „periodisch“ die Rationalität impliziert (für die andere Implikation verwendet man Division mit Rest):

Einsatz der geometrischen Reihe

Wegen

$$m, a_1 \dots a_k \overline{b_1 \dots b_p} = m, a_1 \dots a_k + 1/10^k \cdot 0, \overline{b_1 \dots b_p}$$

können wir uns auf die reinperiodischen Darstellungen beschränken. Hier gilt

$$\begin{aligned} 0, \overline{b_1 \dots b_p} &= b_1/10 + \dots + b_p/10^p + b_1/10^{p+1} + \dots + b_p/10^{2p} + \dots = \\ &= (b_1/10 + \dots + b_p/10^p) + (b_1/10^p + \dots + b_p/10^{2p})/10^p + \dots \end{aligned}$$

Mit $q = b_1/10 + \dots + b_p/10^p \in \mathbb{Q}$ ist also

$$\begin{aligned} 0, \overline{b_1 \dots b_p} &= q + q/10^p + q/10^{2p} + \dots = q \sum_n (1/10^p)^n = \\ &= q/(1 - 1/10^p) = q \frac{10^p}{10^p - 1} \in \mathbb{Q}. \end{aligned}$$

Die gefundene Formel lässt sich als Umrechnungsformel für periodische Dezimaldarstellungen lesen:

$$0, \overline{b_1 \dots b_p} = b_1 \dots b_p / 9 \dots 9 \text{ (in Dezimalnotation),}$$

mit p Neunen im Nenner. Dieses Ergebnis wird manchmal auch als die *Regel von Robertson* bezeichnet. Man schreibt die gewünschte Periode in den Zähler und in den Nenner die Neuner-Zahl der entsprechenden Länge.

Beispiele

$$0, \overline{2} = 0,222\dots = 2/9, \quad 0, \overline{15} = 0,151515\dots = 15/99 = 5/33,$$

$$0, \overline{03210} = 0,032100321003210\dots = 3210/99999.$$

4.5 Die harmonische Reihe

Satz (Divergenz der harmonischen Reihe)

Für die sog. harmonische Reihe

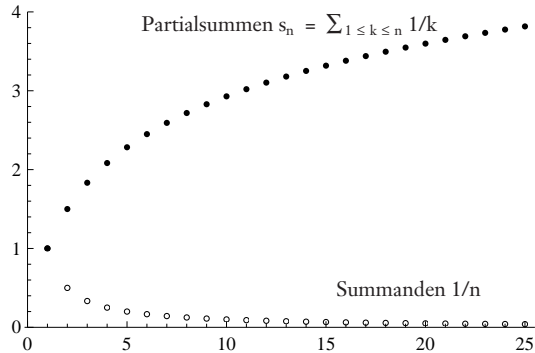
$$\sum_{n \geq 1} 1/n = 1 + 1/2 + 1/3 + 1/4 + 1/5 + \dots + 1/n + \dots$$

gilt: (a) $\sum_{n \geq 1} 1/n = +\infty$, (b) $\lim_{n \geq 1} 1/n = 0$.

Die harmonische Reihe ist ein Beispiel für eine divergente Reihe, deren Summanden eine Nullfolge bilden. Eine Berechnung der Partialsummen zeigt, dass die harmonische Reihe sehr langsam wächst. So gilt z. B.

$$1 + \dots + 1/10000 \sim 9.78761.$$

Dies scheint eher Konvergenz anzudeuten. Eine verblüffend einfache Argumentation zeigt jedoch, dass die harmonische Reihe divergiert. Das Argument gehört zu den Klassikern der Mathematik.



Die ersten 25 Partialsummen s_n der harmonischen Reihe.
Die Zahlen $s_n = \sum_{1 \leq k \leq n} 1/k$ heißen auch *harmonische Zahlen*.

Beweis der Divergenz der harmonischen Reihe

Die Idee ist, die Summanden der Reihe zu Blöcken der Länge 1, 2, 4, 8, 16, ... zusammenzufassen:

$$\begin{aligned} &1 + \\ &(1/2 + 1/3) + \\ &(1/4 + 1/5 + 1/6 + 1/7) + \\ &(1/8 + 1/9 + 1/10 + 1/11 + 1/12 + 1/13 + 1/14 + 1/15) + \\ &\dots \end{aligned}$$

Der erste Block ist größer als $1/2$. Der zweite Block hat zwei Summanden, die beide größer als $1/4$ sind, also ist dieser Block größer als $2/4 = 1/2$. Analog hat der dritte Block vier Summanden größer als $1/8$, also ist dieser Block größer als $4/8 = 1/2$. Allgemein hat der k -te Block 2^{k-1} Summanden größer als $1/2^k$ und ist damit größer als $2^{k-1}/2^k = 1/2$. Summieren wir also n Blöcke auf, so erhalten wir eine Partialsumme der harmonischen Reihe, die größer als $n/2$ ist. Da die Partialsummen monoton steigen, folgt $\sum_{n \geq 1} 1/n = +\infty$.

Die harmonische Reihe besitzt viele bemerkenswerte Eigenschaften. Auf drei von ihnen wollen wir noch etwas genauer eingehen.

Zusammenhang zum natürlichen Logarithmus

Sei \log der Logarithmus zur Basis e .

Dann existiert

$$C = \lim_n \left(\sum_{1 \leq k \leq n} 1/k - \log(n) \right).$$

C heißt *Euler-Mascheroni-Konstante*.

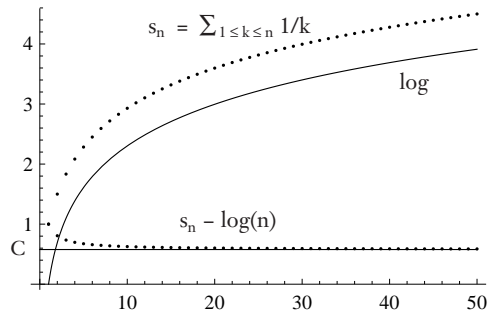
Es gilt

$$C = 0,57721566\dots$$

Die Konvergenzaussage wird falsch,

wenn wir einen Logarithmus zu

einer anderen Basis wählen. Insofern „steckt“ die Eulersche Zahl e in $\sum_{n \geq 1} 1/n$.



Exponenten größer als 1 im Nenner

Für $k = 2, 3, 4, \dots$ konvergiert

$$\sum_{n \geq 1} 1/n^k = 1 + 1/2^k + 1/3^k + \dots$$

Speziell gilt, wie schon erwähnt,

$$\sum_{n \geq 1} 1/n^2 = \pi^2/6 \sim 1,64493.$$

Allgemeiner ist $\sum_{n \geq 1} 1/n^x$ für

alle $x > 1$ konvergent, sodass

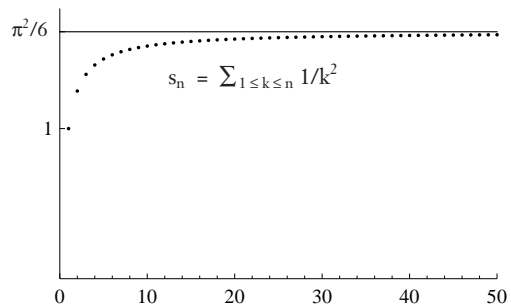
also $\sum_{n \geq 1} 1/n^{1,01}$ existiert.

Die Funktion

$$f :]1, \infty[\rightarrow \mathbb{R} \text{ mit } f(x) = \sum_{n \geq 1} 1/n^x \text{ für alle } x > 1$$

besitzt an der Stelle 1 einen Pol, der der Divergenz der harmonischen Reihe entspricht.

Die Fortsetzung dieser Funktion ins Komplexe führt zur Riemannschen Zetafunktion und dem bislang ungelösten Problem ihrer Nullstellen (Riemannsche Hypothese).



Die ersten Partialsummen von $1 + 1/4 + 1/9 + \dots$

Ausdünnungen der harmonischen Reihe

Man kann zeigen, dass die Reihe

$$\sum_{p \text{ prim}} 1/p = 1/2 + 1/3 + 1/5 + 1/7 + 1/11 + 1/13 + 1/17 + \dots$$

immer noch divergent ist. Es folgt, dass die Menge der Primzahlen unendlich sein muss, sonst wäre die Summe endlich. Dagegen konvergiert die Ausdünnung

$$\sum_{n \geq 1, n \text{ ist } 9\text{-frei}} 1/n = 1/1 + 1/2 + 1/3 + \dots + 1/8 + 1/10 + \dots + 1/18 + 1/20 + \dots,$$

wobei ein n 9-frei heißen soll, wenn die Dezimaldarstellung von n keine 9 enthält.

Umgekehrt bedeutet dies, dass die folgende Reihe divergiert:

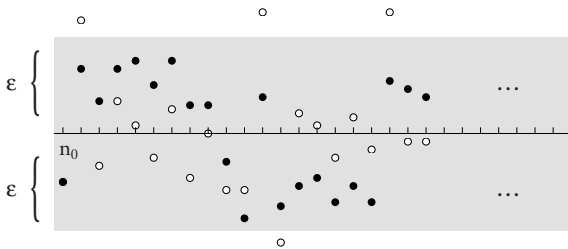
$$1/9 + 1/19 + 1/29 + \dots + 1/89 + 1/90 + 1/91 + \dots + 1/99 + 1/109 + \dots$$

4.6 Das Cauchy-Kriterium für Reihen

Satz (Cauchy-Kriterium der Konvergenz für Reihen)

Sei $\sum_n x_n$ eine Reihe in \mathbb{R} . Dann sind die folgenden Aussagen äquivalent:

- (a) $\sum_{n \in \mathbb{N}} x_n$ konvergiert.
- (b) $\forall \varepsilon > 0 \exists n_0 \forall n \geq m \geq n_0 \left| \sum_{m \leq k \leq n} x_k \right| < \varepsilon$. (Cauchy-Bedingung für Reihen)
- (c) $\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 \left| \sum_{n_0 \leq k \leq n} x_k \right| < \varepsilon$. (Cauchy-Bedingung für Reihen, II)



Zur Cauchy-Bedingung in der Form (c): Alle Summen $\sum_{n_0 \leq k \leq n} x_k$ (schwarze Punkte) der Summanden x_k (weiße Punkte) liegen im ε -Streifen um die x -Achse.

Die Reihe $\sum_n x_n$ ist per Definition die Folge $(s_n)_{n \in \mathbb{N}}$ ihrer Partialsummen, und diese Folge konvergiert wie jede Folge in \mathbb{R} genau dann, wenn sie eine Cauchy-Folge ist. Die Aussagen (b) und (c) des Satzes sind äquivalente Formulierungen der Aussage „ $(s_n)_{n \in \mathbb{N}}$ ist eine Cauchy-Folge“. Statt „ $< \varepsilon$ “ kann man wieder „ $\leq \varepsilon$ “ oder „ $< 2\varepsilon$ “ in die Bedingungen einsetzen, und die Summe kann man auch über „ $n_0 \leq k \leq 2n$ “ bzw. „ $m \leq k \leq 2n$ “ laufen lassen. Man darf aber nicht alles:

Warnung

Aus der Konvergenz von $\sum_n x_n$ folgt im Allgemeinen nicht, dass

$$\forall \varepsilon > 0 \exists n_0 \forall n \geq m \geq n_0 \sum_{m < k \leq n} |x_k| < \varepsilon.$$

Man darf also in der Regel die Betragsstriche nicht in die Summe reinziehen.

Beispiele hierfür werden wir in 4.7 und 4.8 kennenlernen.

Das Cauchy-Kriterium wird in der Praxis oft zum Nachweis der Konvergenz oder Divergenz einer Reihe eingesetzt. Oft genügen schwächere Versionen wie zum Beispiel:

Satz (Nullfolgenbedingung)

Sei $\sum_n x_n$ eine konvergente Reihe in \mathbb{R} . Dann ist die Folge $(x_n)_{n \in \mathbb{N}}$ der Summanden eine Nullfolge, d. h., es gilt $\lim_n x_n = 0$.

Setzen wir nämlich $n = m$ in der Cauchy-Bedingung, so erhalten wir

$$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 |x_n| < \varepsilon, \text{ d.h. } \lim_n x_n = 0.$$

Für obiges Diagramm heißt dies: Wir können durch Vergrößerung von n_0 erreichen, dass die Summen $\sum_{n_0 \leq k \leq m} x_k$ und die Summanden x_k im ε -Streifen liegen.

Beispiel

Gilt $|x_n| > 1/10$ für unendlich viele n , so ist $(x_n)_{n \in \mathbb{N}}$ keine Nullfolge, also divergiert die Reihe $\sum_n x_n$.

Das Cauchy-Kriterium für Reihen liefert stärkere Aussagen, die mit der Nullfolgenbedingung nicht mehr behandelt werden können:

Beispiel

Gilt $|x_n + \dots + x_{2^n}| > 1/10$ für unendlich viele n , so divergiert $\sum_n x_n$.

Die harmonische Reihe (vgl. 4.5) zeigt:

Die Reihe $\sum_n x_n$ kann divergieren, obwohl $\lim_n x_n = 0$ gilt!

Dass die Summanden eine Nullfolge bilden, ist also notwendig für die Konvergenz einer Reihe, aber nicht hinreichend. Es besteht die Möglichkeit der Divergenz bei verschwindenden Zuwächsen.

Schließlich halten wir noch fest:

Satz (*Abschneiden von Anfangsstücken*)

Sei $\sum_n x_n$ eine Reihe in \mathbb{R} und sei $m_0 \in \mathbb{N}$. Dann konvergiert $\sum_n x_n$ genau dann, wenn $\sum_{n \geq m_0} x_n$ konvergiert.

Auch dies folgt sofort aus dem Cauchy-Kriterium, denn oberhalb von m_0 sind die Werte $|\sum_{m \leq k \leq n} x_k|$ für beide Reihen gleich. Der Satz lässt sich noch durch eine Summenformel ergänzen. Im Fall der Konvergenz gilt für alle k :

$$\sum_n x_n = \sum_{n \leq k} x_n + \sum_{n > k} x_n. \quad (\text{endliche Abspaltungsregel})$$

Die Abspaltungsregel hat die Form eines Assoziativgesetzes:

$$x_0 + x_1 + \dots + x_n + \dots = (x_0 + \dots + x_k) + x_{k+1} + x_{k+2} + \dots$$

Allgemeiner ändern im Fall der Konvergenz auch unendlich viele Blockbildungen die Summe nicht:

$$x_0 + x_1 + \dots + x_n + \dots = (x_0 + \dots + x_{k_0}) + (x_{k_0+1} + \dots + x_{k_1}) + \dots$$

(endliche Zusammenfassungen)

Die divergente Reihe $\sum_n (-1)^n$ zeigt, dass derartige Blockbildungen Konvergenz (und verschiedene Summen) erzeugen können, wenn die Ausgangsreihe divergiert (vgl. 4.1.).

Wir fassen zusammen:

Liegt Konvergenz vor, so darf in einer unendlichen Reihe beliebig geklammert werden.

Das Thema „Kommutativgesetz im Unendlichen“ ist schwieriger. Wir diskutieren es in Sektion 4.8.

4.7 Das Leibniz-Kriterium

Satz (Leibniz-Kriterium, Konvergenzkriterium für alternierende Reihen)

Sei $(x_n)_{n \in \mathbb{N}}$ eine monoton fallende Nullfolge in \mathbb{R} , d. h., es gilt $x_n \geq x_{n+1}$ für alle n und $\lim_n x_n = 0$. Dann gilt für die Partialsummen s_n der Reihe $\sum_n (-1)^n x_n$:

$$(a) \quad s_1 \leq s_3 \leq \dots \leq s_{2n+1} \leq \dots \leq s_{2n} \leq \dots \leq s_2 \leq s_0,$$

$$(b) \quad \lim_n (s_{2n} - s_{2n+1}) = 0.$$

Folglich ist $\sum_n (-1)^n x_n$ konvergent und

$$\sum_n (-1)^n x_n = \sup_n s_{2n+1} = \inf_n s_{2n}.$$

Analoges gilt für $\sum_n (-1)^{n+1} x_n$.

Die Reihe $\sum_n (-1)^n x_n$ hat die Form

$$x_0 - x_1 + x_2 - x_3 + \dots + x_{2n} - x_{2n+1} + \dots,$$

wobei alle x_n größergleich null sind und monoton fallend gegen null konvergieren. Aufgrund der Monotonieeigenschaft ist die Folge der Partialsummen

$$s_0 = x_0, \quad s_1 = x_0 - x_1, \quad s_2 = x_0 - x_1 + x_2,$$

$$s_3 = x_0 - x_1 + x_2 - x_3, \quad \dots$$

eine rechtsstartende Pendelfolge wie in (a), und aufgrund der Konvergenz der Summanden x_n gegen 0 gilt

$$\lim_n (s_{2n} - s_{2n+1}) = \lim_n x_n = 0.$$

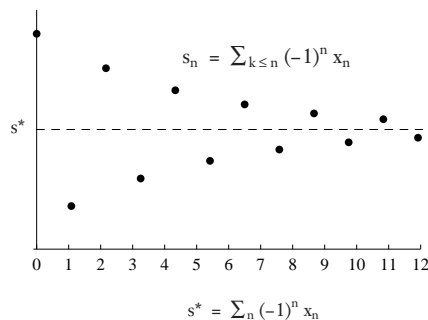
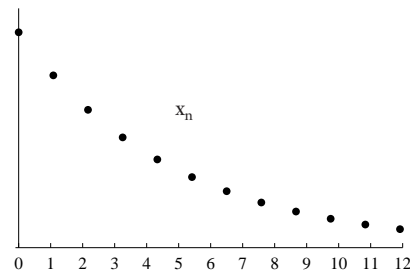
Damit ist $(s_n)_{n \in \mathbb{N}}$ eine Folge, die die Konvergenzbedingung für Pendelfolgen erfüllt (vgl. 3.3). Analoges gilt für die Reihe

$$\sum_n (-1)^{n+1} x_n = - \sum_n (-1)^n x_n,$$

deren Partialsummen eine linksstartende Pendelfolge bilden. Das Leibniz-Kriterium ist damit vollständig auf den Konvergenzsatz für Pendelfolgen zurückgeführt.

Eine Reihe $\sum_n y_n$, deren Summanden ständig das Vorzeichen wechseln, heißt *alternierend*. Das Leibniz-Kriterium besagt, dass eine alternierende Reihe konvergiert, wenn die Folge $(|y_n|)_{n \in \mathbb{N}}$ der Beträge ihrer Summanden eine monoton fallende Nullfolge ist. Man nennt das Kriterium deswegen oft auch das *Konvergenzkriterium für alternierende Reihen*.

Wir betrachten einige Beispiele für alternierende Reihen. Das Leibniz-Kriterium liefert die Konvergenz oft leicht, die Bestimmung der Grenzwerte ist dagegen meistens viel schwieriger.



Beispiele

- (1) Für alle $x \in]-1, 0[$ ist die geometrische Reihe $\sum_n x^n$ eine alternierende Reihe. Das Leibniz-Kriterium zeigt noch einmal die Konvergenz. Die Summe hatten wir bereits bestimmt, es gilt $\sum_n x^n = 1/(1-x)$.

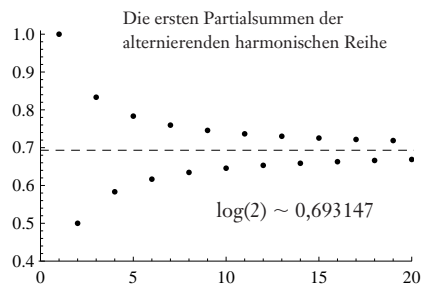
- (2) Sei $(a_n)_{n \in \mathbb{N}}$ eine um die Null pendelnde Folge der Form

$$a_1 \leq a_3 \leq \dots \leq a_{2n+1} \leq \dots \leq 0 \leq \dots \leq a_{2n} \leq \dots \leq a_2 \leq a_0.$$

Dann ist, für alle $x > 0$, die Reihe $\sum_n a_n x^n$ eine alternierende Reihe. Sie konvergiert nach dem Leibniz-Kriterium, falls $x \in]0, 1[$. Wir setzen hier $\lim_n a_n = 0$ nicht voraus. Die geometrische Reihe aus Beispiel (1) entspricht dem Fall $(a_n)_{n \in \mathbb{N}} = (1, -1, 1, -1, \dots)$.

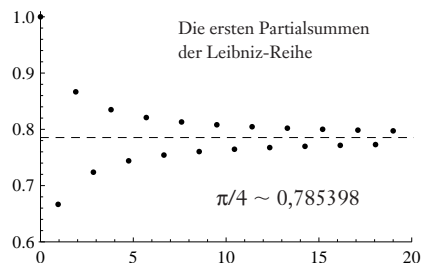
- (3) Die *alternierende harmonische Reihe* $\sum_{n \geq 1} (-1)^{n-1}/n$ konvergiert, da $(1/n)_{n \geq 1}$ monoton fallend gegen 0 konvergiert. Also ist die Reihe konvergent. Mit der Potenzreihenentwicklung der Logarithmusfunktion \log kann man die Summe identifizieren. Es gilt

$$1 - 1/2 + 1/3 - 1/4 + \dots = \log(2).$$



- (4) Die *Leibniz-Reihe* $\sum_n (-1)^n/(2n+1)$ konvergiert nach dem Satz. Die Potenzreihenentwicklung des Arkustangens und der Wert $\arctan(1) = \pi/4$ ermöglichen eine Bestimmung der Summe. Es gilt

$$1 - 1/3 + 1/5 - 1/7 + \dots = \pi/4.$$



Die alternierende harmonische Reihe und die Leibniz-Reihe besitzen nicht nur faszinierende Grenzwerte, sondern sie sind auch Beispiele für das folgende Phänomen:

$$\sum_n x_n \text{ kann konvergieren, während } \sum_n |x_n| \text{ divergiert.}$$

In der Tat gilt $\sum_{n \geq 1} 1/n = +\infty$ und $\sum_n 1/(2n+1) = +\infty$, während $\sum_{n \geq 1} (-1)^{n-1}/n$ und $\sum_n (-1)^n/(2n+1)$ konvergieren. Die merkwürdigen Eigenschaften derartiger Reihen werden wir in der nächsten Sektion diskutieren.

Alternierende Reihen treten häufiger auf, als man meinen möchte. Besitzt eine Reihe unendlich viele positive und unendlich viele negative Summanden, so können wir durch Zusammenfassen der positiven und negativen Blöcke eine alternierende Reihe erzeugen. Das Konvergenzverhalten wird durch diese Zusammenfassung nicht verändert. Im Allgemeinen erhalten wir aber keine Monotonie in den Summanden.

4.8 Absolute und bedingte Konvergenz

Definition (absolute und bedingte Konvergenz)

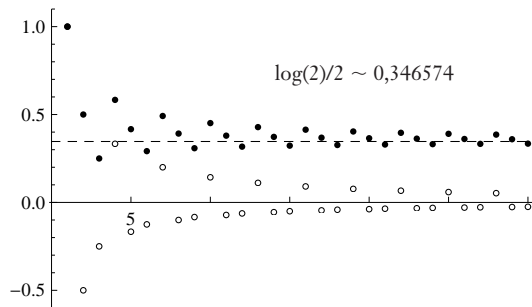
Eine Reihe $\sum_n x_n$ in \mathbb{R} heißt

- (a) *absolut konvergent*, falls $\sum_n |x_n|$ konvergiert,
- (b) *bedingt konvergent*, falls $\sum_n x_n$ konvergiert und $\sum_n |x_n|$ divergiert.

Aus dem Cauchy-Kriterium folgt wegen $|\sum_{m \leq k \leq n} x_k| \leq \sum_{m \leq k \leq n} |x_k|$, dass eine absolut konvergente Reihe $\sum_n x_n$ konvergent ist.

Beispiele

- (1) Für alle $x \in]-1, 1[$ ist die geometrische Reihe $\sum_n x^n$ absolut konvergent.
- (2) Die alternierende harmonische Reihe $\sum_{n \geq 1} (-1)^{n-1} / n$ und auch die Leibniz-Reihe $\sum_n (-1)^n / (2n+1)$ sind bedingt konvergent (vgl. 4.7).



Die ersten 30 Partialsummen (schwarze Punkte) und Summanden (weiße Punkte) der Umordnung

$$1 - 1/2 - 1/4 + 1/3 - 1/6 - 1/8 + 1/5 - 1/10 - 1/12 + \dots$$

der bedingt konvergenten alternierenden harmonischen Reihe.

Will man eine konvergente Reihe auf absolute oder bedingte Konvergenz überprüfen, so ist der folgende Satz hilfreich:

Satz (Kriterien für bedingte Konvergenz)

Sei $\sum_n x_n$ eine konvergente Reihe in \mathbb{R} . Dann sind äquivalent:

- (a) $\sum_n x_n$ konvergiert bedingt.
- (b) $\sum_n |x_n| = \infty$.
- (c) Die Summen über die negativen und positiven Summanden divergieren, d. h., es gilt $\sum_n \text{mit } x_n > 0 \ x_n = \infty$ und $\sum_n \text{mit } x_n < 0 \ x_n = -\infty$.

Als Merkregel gilt:

Absolute Konvergenz ist eine gute Eigenschaft.

Man kann mit absolut konvergenten Reihen sehr frei rechnen, ohne Fehler zu machen. Dies ist für bedingt konvergente Reihen ganz anders. Als Beispiel betrachten wir die konvergente alternierende harmonische Reihe:

$$\sum_{n \geq 1} (-1)^{n-1} / n = 1 - 1/2 + 1/3 - 1/4 + \dots = \log(2).$$

Die Summanden $1, -1/2, 1/3, -1/4, \dots$ dieser Reihe können wir anders anordnen:

Beispiele

$$(1) 1 - 1/2 - 1/4 + 1/3 - 1/6 - 1/8 + 1/5 - 1/10 - 1/12 + \dots = \log(2)/2.$$

Hier folgen zwei negative auf einen positiven Summanden. Die Dominanz der negativen Summanden führt zu einer kleineren Summe.

$$(2) 1 - 1/2 + 1/3 - 1/4 - 1/6 + 1/5 - 1/8 - 1/10 - 1/12 - 1/14 + 1/7 - \dots = -\infty$$

Hier folgen 1, 2, 4, 8, 16, ... negative Summanden auf einen positiven. Die Reihe divergiert bestimmt gegen $-\infty$.

Die Reihenfolge der Summation spielt also bei bedingt konvergenten Reihen eine Rolle. Für absolut konvergente Reihen kann dies nicht passieren. Zur genauen Formulierung der Ergebnisse nennen wir eine Reihe $\sum_n y_n$ eine *Umordnung* der Reihe $\sum_n x_n$, falls eine Bijektion $g: \mathbb{N} \rightarrow \mathbb{N}$ existiert mit $y_n = x_{g(n)}$ für alle n . Es gilt dann also

$$\sum_n y_n = x_{g(0)} + x_{g(1)} + x_{g(2)} + \dots,$$

d. h., die Funktion g gibt die Reihenfolge der Summation an. Dass $g: \mathbb{N} \rightarrow \mathbb{N}$ eine Bijektion ist, führt dazu, dass jedes x_n in der neuen Summe genau einmal erscheint. Es gilt der folgende Satz:

Satz (*Umordnungssatz*)

- (a) Ist $\sum_n x_n$ eine absolut konvergente Reihe, so ist auch jede Umordnung $\sum_n y_n$ von $\sum_n x_n$ absolut konvergent und es gilt $\sum_n x_n = \sum_n y_n$.
- (b) Ist $\sum_n x_n$ eine bedingt konvergente Reihe und $c \in \overline{\mathbb{R}} = [-\infty, +\infty]$, so existiert eine Umordnung $\sum_n y_n$ von $\sum_n x_n$ mit $\sum_n y_n = c$. Weiter existieren divergente Umordnungen von $\sum_n x_n$, die auch nicht uneigentlich konvergieren.

Kurzfassung:

*Das Kommutativgesetz gilt im Unendlichen für absolut konvergente Reihen.
Bei bedingter Konvergenz kann man jeden beliebigen Wert durch Umordnung erreichen.*

Bemerkung: Umordnungen von Folgen

Folgen sind stets immun gegen Umordnungen. Ist $x = \lim_n x_n$ und $g: \mathbb{N} \rightarrow \mathbb{N}$ bijektiv, so ist auch $x = \lim_n x_{g(n)}$. Denn in jeder Umgebung von x liegen fast alle x_n . Also liegen in jeder Umgebung von x auch fast alle $x_{g(n)}$.

Der Leser wird nun vielleicht fragen:

*Ist nicht jede Reihe auch eine Folge? Wie kann dann die
Umordnung für Reihen anders sein als für Folgen?*

Antwort: Die Reihe $\sum_n x_n$ ist die Partialsummenfolge $(s_n)_{n \in \mathbb{N}}$. Ist nun $\sum_n y_n$ eine Umordnung von $\sum_n x_n$, so ist im Allgemeinen die Partialsummenfolge $(t_n)_{n \in \mathbb{N}}$ von $\sum_n y_n$ keine Umordnung der Folge $(s_n)_{n \in \mathbb{N}}$. Sie kann ganz andere Glieder haben.

4.9 Majorantenkriterium und Minorantenkriterium

Satz (Majorantenkriterium für Konvergenz)

Sei $\sum_n x_n$ eine Reihe in \mathbb{R} . Weiter sei $\sum_n y_n$ eine konvergente Reihe mit Summanden in $[0, +\infty[$ mit der Eigenschaft

$$|x_n| \leq y_n \quad \text{für alle } n.$$

Dann konvergiert $\sum_n x_n$ absolut und es gilt $|\sum_n x_n| \leq \sum_n |x_n| \leq \sum_n y_n$.

Wir betrachten das Kriterium zuerst für den Fall, dass auch alle x_n größergleich null sind. Dann sind die Partialsummen $(s_n)_{n \in \mathbb{N}}$ von $\sum_n x_n$ und $(t_n)_{n \in \mathbb{N}}$ von $\sum_n y_n$ monoton steigend, und nach Voraussetzung existiert

$$t = \lim_n t_n = \sum_n y_n.$$

Wegen $s_n \leq t_n$ für alle n gilt

$$s_0 \leq s_1 \leq \dots \leq s_n \leq \dots \leq t,$$

also existiert

$$s = \sum_n x_n = \lim_n s_n \leq t.$$

Ist $\sum_n x_n$ eine Reihe mit $|x_n| < y_n$ für alle n , so zeigt das gerade geführte Argument, dass $\sum_n |x_n| \leq t$. Also ist die Reihe $\sum_n x_n$ absolut konvergent und damit konvergent.

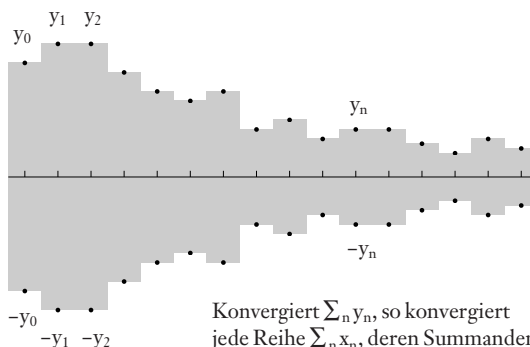
$$y_0 + y_1 + y_2 + \dots + y_n + \dots = y$$

$$\text{IV} \quad \text{IV} \quad \text{IV} \quad \dots \quad \text{IV}$$

$$x_0 + x_1 + x_2 + \dots + x_n + \dots$$

$$\text{IV} \quad \text{IV} \quad \text{IV} \quad \dots \quad \text{IV}$$

$$0 \quad 0 \quad 0 \quad \dots \quad 0 \quad \dots$$



Konvergiert $\sum_n y_n$, so konvergiert jede Reihe $\sum_n x_n$, deren Summanden x_n sich im grauen Bereich befinden.

Beispiele

- (1) Die Reihe $\sum_{n \geq 1} 1/n^2$ konvergiert, denn für alle $n \geq 1$ gilt

$$\frac{1}{n^2} \leq \frac{2}{n(n+1)}, \quad \text{da } n(n+1) = n^2 + n \leq 2n^2,$$

und diese Majorisierung ist konvergent, da $2 \sum_{n \geq 1} 1/(n(n+1)) = 2$ (vgl. 4.2).

- (2) Die konvergente Reihe $\sum_{n \geq 1} 1/n^2$ können wir nun selbst als Majorante verwenden. Ist $k \geq 2$, so ist $1/n^k \leq 1/n^2$ für alle $n \geq 1$ und damit ist die Reihe $\sum_{n \geq 1} 1/n^k$ konvergent.

- (3) Ist $(x_n)_{n \geq 1}$ eine (nicht notwendig konvergente) Folge in \mathbb{R} , so konvergiert die Reihe

$$\sum_{n \geq 1} \frac{\sin(x_n)}{n^2} = \frac{\sin(x_1)}{1} + \frac{\sin(x_2)}{4} + \frac{\sin(x_3)}{9} + \dots$$

absolut. Denn für alle $n \geq 1$ gilt $|\sin(x_n)/n^2| \leq 1/n^2$.

Oft findet man eine konvergente Reihe $\sum_n y_n$, die schließlich oberhalb der betrachteten Reihe $\sum_n x_n$ liegt, d. h., es gilt $|x_n| \leq y_n$ für alle n größergleich einem n_0 . In diesem Fall gilt die Konvergenzaussage des Majorantenkriteriums weiterhin, die Abschätzung ist dagegen anzupassen:

$$|\sum_n x_n| \leq \sum_n |x_n| \leq \sum_{n \geq n_0} y_n + \sum_{k < n_0} x_k.$$

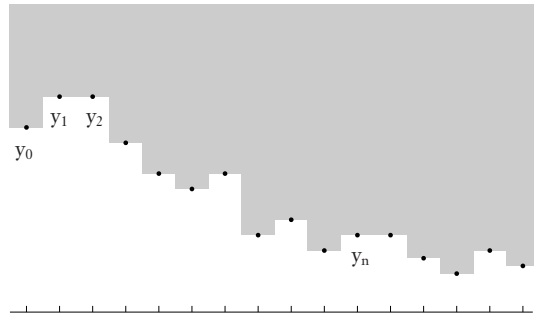
Das Majorantenkriterium ist die Mutter aller weiteren Konvergenzkriterien. Zwei wichtige Beispiele werden wir in der nächsten Sektion besprechen. Hier betrachten wir noch ein Kriterium zur Divergenz.

Satz (*Minorantenkriterium für Divergenz*)

Sei $\sum_n y_n$ eine divergente Reihe mit Summanden in $[0, +\infty[$. Weiter sei $\sum_n x_n$ eine Reihe mit Summanden in $[0, +\infty[$ mit der Eigenschaft

$$y_n \leq x_n \quad \text{für alle } n.$$

Dann divergiert $\sum_n x_n$.



Beispiel

Gleichmäßige progressive Ausdünnungen der harmonischen Reihe sind divergent. So gilt zum Beispiel

$$\sum_{n \geq 1} \frac{1}{6 + 4n} = \frac{1}{10} + \frac{1}{14} + \frac{1}{18} + \frac{1}{22} + \dots = +\infty.$$

Zum Beweis muss man nur beobachten, dass $1/10 \sum_{n \geq 1} 1/n$ eine divergente Minorante der Reihe ist.

Divergiert $\sum_n y_n$, so divergiert jede Reihe $\sum_n x_n$, deren Summanden x_n sich im grauen Bereich befinden.

Dem Leser ist vielleicht aufgefallen, dass die beiden Kriterien nicht vollkommen symmetrisch sind. Im Minorantenkriterium kommen keine Betragsstriche vor. Dies lässt sich in der Tat nicht erreichen:

Bemerkung

Ist $\sum_n y_n$ divergent und $\sum_n x_n$ eine Reihe in \mathbb{R} mit

$$y_n \leq |x_n| \quad \text{für alle } n,$$

so zeigt das Minorantenkriterium, dass die Reihe $\sum_n |x_n|$ divergiert. Dies heißt, dass $\sum_n x_n$ nicht absolut konvergiert, d. h., $\sum_n x_n$ divergiert oder konvergiert bedingt. Die alternierende harmonische Reihe $\sum_{n \geq 1} (-1)^{n-1}/n$ zeigt, dass der zweite Fall eintreten kann.

4.10 Wurzelkriterium und Quotientenkriterium

Satz (Wurzelkriterium)

Sei $\sum_n x_n$ eine Reihe in \mathbb{R} . Dann gilt:

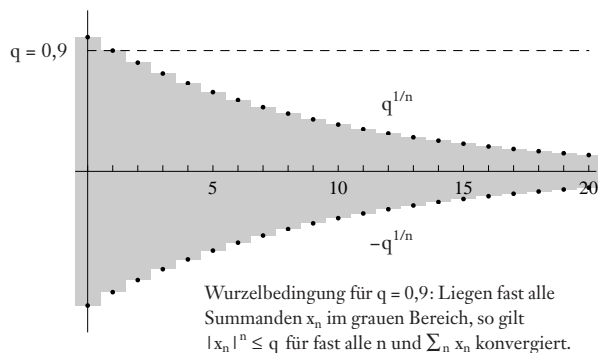
- (a) Gibt es ein $q \in]0, 1[$ und ein n_0 , sodass

$$\sqrt[n]{|x_n|} \leq q \quad \text{für alle } n \geq n_0, \quad (\text{Wurzelbedingung für } q)$$

so konvergiert $\sum_n x_n$ absolut.

- (b) Gilt $\sqrt[n]{|x_n|} \geq 1$ für unendlich viele n , so divergiert $\sum_n x_n$.

Das Wurzel- und das mit ihm eng verwandte Quotientenkriterium sind oftmals geeignet, um Reihen auf Konvergenz und Divergenz zu testen. Dabei sind die n -ten Wurzeln in der Wurzelbedingung zunächst eher abschreckend. Sie werden durch die folgende einfache Beobachtung sympathischer:



Die Aussage (a) ist eine Umformulierung der Majorisierung der Reihe $\sum_n |x_n|$ durch eine geometrische Reihe $\sum_n q^n$ (ab einem n_0).

Denn $|x_n| \leq q^n$ ist äquivalent zu $\sqrt[n]{|x_n|} \leq q$. Die Konvergenzaussage ist also eine Zusammenschau des Majorantenkriteriums und der Konvergenz der geometrischen Reihe. Die Divergenzaussage (b) ist zudem leicht einzusehen. Gilt nämlich $\sqrt[n]{|x_n|} \geq 1$ für unendlich viele n , so gilt $|x_n| \geq 1$ für unendlich viele n . Dann bilden die Summanden x_n keine Nullfolge, sodass $\sum_n x_n$ divergiert.

Beispiel

In der Reihe $1/2 + 1/5^2 + 1/2^3 + 1/5^4 + 1/2^5 + 1/5^6 + \dots$ ist n -te Wurzel des n -ten Summanden stets gleich $1/2$ oder $1/5$. Die Wurzelbedingung ist für $q = 1/2$ erfüllt und die Reihe also konvergent.

Oft existiert, im Gegensatz zu diesem Beispiel, $w = \lim_n \sqrt[n]{|x_n|}$. Gilt $w < 1$, so ist die Wurzelbedingung für alle $q \in]w, 1[$ erfüllt und die Reihe also konvergent. Gilt $w > 1$, so gilt Divergenz nach (b). Im Fall $w = 1$ kann man keine Aussage machen. Das Wurzelkriterium deckt nicht alle Fälle ab: Die Werte $\sqrt[n]{|x_n|}$ können sich von unten an die 1 anpirschen, und dann gilt weder die Voraussetzung von (a) noch die von (b). In solchen Fällen kann Konvergenz oder Divergenz vorliegen und man muss andere Methoden verwenden, um dies zu entscheiden. Wir diskutieren zwei Beispiele.

Beispiel

- (1) Für $\sum_{n \geq 1} (n/(n+1))^n$ gilt $\sqrt[n]{|x_n|} = n/(n+1)$ für alle $n \geq 1$, sodass weder (a) noch (b) greifen. Die Reihe divergiert, da die Summanden keine Nullfolge bilden:

$$\lim_{n \geq 1} (n/(n+1))^n = \lim_{n \geq 1} 1/(1+1/n)^n = 1/e.$$

- (2) Das Kriterium versagt auch für $\sum_n 1/n$ und für $\sum_n 1/n^2$, da

$$\lim_{n \geq 1} (1/n)^{1/n} = \lim_{n \geq 1} (1/n^2)^{1/n} = 1.$$

Eine handliche Abschwächung des Wurzelkriteriums ist, last but not least:

Satz (Quotientenkriterium)

Sei $\sum_n x_n$ eine Reihe in \mathbb{R} . Dann gilt:

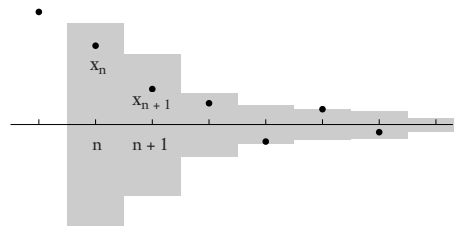
- (a) Gibt es ein $q \in]0, 1[$ und ein n_0 , sodass

$$|x_{n+1}/x_n| \leq q \quad \text{für alle } n \geq n_0,$$

so konvergiert $\sum_n x_n$ absolut.

- (b) Gibt es ein n_0 mit $|x_{n+1}/x_n| \geq 1$ für alle $n \geq n_0$, so divergiert $\sum_n x_n$.

(Quotientenbedingung für q)



Quotientenbedingung für $q=0,9$: Der graue Balken bei $n+1$ markiert das Intervall $[-q x_{n+1}, q x_{n+1}]$. Dieses Intervall hängt also nicht von n wie bei der Wurzelbedingung, sondern von x_n ab. Der Abfall der Summanden wird dadurch in der Regel beschleunigt.

Auch hier liegt eine geometrische Majorisierung vor, denn aus der Quotientenbedingung folgt „ $|x_n| \leq |x_0| q^n$ für alle $n \geq n_0$ “, wie eine Induktion nach n zeigt.

Beispiele

- (1) Die Reihe $\sum_{n \geq 1} n^2/2^n$ konvergiert, da für alle $n \geq 3$ gilt

$$|x_{n+1}/x_n| = (n+1)^2/2^{n+1} \cdot 2^n/n^2 = (1+1/n)^2/2 < (1+1/3)^2/2 = 8/9 < 1.$$

- (2) Für die Reihe $\sum_{n \geq 1} 1/n$ gilt

$$\lim_{n \geq 1} |x_{n+1}/x_n| = \lim_{n \geq 1} n/(n+1) = 1.$$

Das Quotientenkriterium ist nicht anwendbar, die Reihe divergiert.

- (3) Für die Reihe $\sum_{n \geq 1} 1/n^2$ gilt

$$\lim_{n \geq 1} |x_{n+1}/x_n| = \lim_{n \geq 1} (n/(n+1))^2 = 1.$$

Das Quotientenkriterium ist nicht anwendbar, die Reihe konvergiert.

- (4) Nach dem Wurzelkriterium konvergiert die obige Reihe

$$1/2 + 1/5^2 + 1/2^3 + 1/5^4 + 1/2^5 + 1/5^6 + \dots$$

Das Quotientenkriterium ist nicht anwendbar, da $|x_{n+1}/x_n|$ jeweils unendlich oft größer und kleiner als 1 ist. Dies zeigt auch, dass wir in (b) „für alle $n \geq n_0$ “ nicht durch „für unendliche viele $n \geq n_0$ “ wie im Wurzelkriterium ersetzen dürfen.

4.11 Produkte von Reihen

Satz (*Produkt absolut konvergenter Reihen*)

Seien $\sum_n x_n$ und $\sum_n y_n$ absolut konvergent. Dann gilt

$$\sum_n x_n \cdot \sum_n y_n = \sum_{n,m} x_n y_m,$$

wobei $\sum_{n,m} x_n y_m$ eine Reihe ist, die alle Produkte $x_n y_m$ in beliebiger Reihenfolge durchläuft.

Konvergente Folgen und Reihen teilen sich die Regeln für Summe und Skalierung, und für Folgen gilt eine Produktregel:

			...			
5						
	$x_0 y_4$	$x_1 y_4$	$x_2 y_4$	$x_3 y_4$	$x_4 y_4$	
4						
	$x_0 y_3$	$x_1 y_3$	$x_2 y_3$	$x_3 y_3$	$x_4 y_3$	
3						
	$x_0 y_2$	$x_1 y_2$	$x_2 y_2$	$x_3 y_2$	$x_4 y_2$...
2						
	$x_0 y_1$	$x_1 y_1$	$x_2 y_1$	$x_3 y_1$	$x_4 y_1$	
1						
	$x_0 y_0$	$x_1 y_0$	$x_2 y_0$	$x_3 y_0$	$x_4 y_0$	
0						
	0	1	2	3	4	5

	Folgen	Reihen
Summenregel	$\lim_n x_n + y_n = \lim_n x_n + \lim_n y_n$	$\sum_n x_n + y_n = \sum_n x_n + \sum_n y_n$
Skalierung	$\lim_n a x_n = a \lim_n x_n$	$\sum_n a x_n = a \sum_n x_n$
Produktregel	$\lim_n x_n y_n = \lim_n x_n \cdot \lim_n y_n$	$? = \sum_n x_n \cdot \sum_n y_n$

Der obige Satz gibt eine Antwort auf das Fragezeichen der Tabelle. Die Produktregel für Reihen ist dabei schwieriger als die Produktregel für Folgen und sie hat auch nicht dieselbe Form. Bereits im Endlichen gilt ja aufgrund des Distributivgesetzes, dass das Produkt zweier Summen die Summe aller Einzelprodukte ist:

$$\sum_{n \leq n_0} x_n \cdot \sum_{m \leq m_0} y_m = \sum_{n \leq n_0, m \leq m_0} x_n y_m.$$

Die Doppelsumme auf der rechten Seite hat genau $n_0 \cdot m_0$ viele Summanden. Aufgrund des Kommutativgesetzes ist es egal, in welcher Reihenfolge wir diese Summanden aufsummieren. Jeder hat ja beim Ausmultiplizieren so seine Vorlieben, aber diese Vorlieben beeinflussen das Ergebnis nicht. Übertragen wir nun die endliche Regel formal ins Unendliche, so erhalten wir eine Summenform wie im Satz:

$$\sum_n x_n \cdot \sum_m y_m = \sum_{n,m} x_n y_m.$$

Die Doppelsumme auf der rechten Seite hat $\mathbb{N} \times \mathbb{N}$ -viele Summanden. Wir können sie als einfache Summe schreiben, indem wir ein bijektives $b: \mathbb{N} \rightarrow \mathbb{N}^2$ fixieren, und definieren

$$\sum_{n,m} x_n y_m = \sum_n x_{n_1} y_{n_2}, \text{ wobei } b(n) = (n_1, n_2) \in \mathbb{N}^2 \text{ für alle } n.$$

Unsere Untersuchung der bedingten Konvergenz hat gezeigt, dass das unendliche Kommutativgesetz verletzt sein kann. Die Summe $\sum_{n,m} x_n y_m$ wird also im Allgemeinen von der Bijektion b abhängen. Der Satz besagt nun, dass die absolute Konvergenz wieder eine „gute“ Voraussetzung ist. Sind die beiden Reihen absolut konvergent, so können wir die

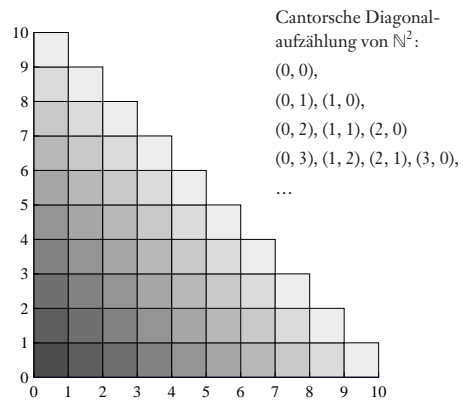
Paare $x_n y_m$ so aufsummieren, wie wir wollen. Wichtig ist nur, dass für jedes $(n, m) \in \mathbb{N}^2$ das Produkt $x_n y_m$ genau einmal in unserer Summe vorkommt.

Beispiel

Wir betrachten das Produkt zweier geometrischer Reihen. Seien also $x, y \in]-1, 1[$. Dann sind $\sum_n x^n$ und $\sum_n y^n$ absolut konvergent. Folglich gilt

$$\sum_{n,m} x^n y^m = \sum_n x^n \cdot \sum_n y^n = \frac{1}{(1-x)(1-y)}.$$

Eine besonders wichtige Anordnung der Summanden $x_n y_m$ folgt der Cantorschen Diagonalaufzählung der Menge \mathbb{N}^2 . Für die Summe $\sum_{n,m} x_n y_m$, die dieser Diagonalaufzählung entspricht, bietet sich eine endliche Blockbildung an, die alle Summanden auf den Diagonalen zusammenfasst. Die Elemente auf der n -ten Diagonalen sind durch die Bedingung, dass die Summe der Indizes konstant gleich n ist, ausgezeichnet. Der n -te Block hat also die Summe $d_n = \sum_{k \leq n} x_k y_{n-k}$. Unser Satz liefert:



Satz (Cauchy-Produkt oder Diagonalprodukt von Reihen)

Seien $\sum_n x_n$ und $\sum_n y_n$ absolut konvergente Reihen in \mathbb{R} , und für alle n sei

$$d_n = \sum_{k \leq n} x_k y_{n-k}.$$

Dann ist $\sum_n d_n$ absolut konvergent und es gilt $\sum_n d_n = \sum_n x_n \cdot \sum_n y_n$.

Das Cauchy-Produkt wurde vor allem aufgrund seiner hervorragenden Leistungen bei der Untersuchung der Exponentialreihe geadelt (vgl. 4.12 und dann vor allem 6.3). Es gibt aber auch hübsche kleinere Anwendungen:

Beispiel

Wir zeigen noch einmal, dass $\sum_{n \geq 1} n x^n = x/(1-x)^2$ für alle $x \in]-1, 1[$ (vgl. 4.3). Hierzu betrachten wir das Cauchy-Produkt der absolut konvergenten Reihen

$$\sum_n x^n \quad \text{und} \quad \sum_n x^{n+1}.$$

Auf der n -ten Diagonalen steht konstant $x^k x^{n-k+1} = x^{n+1}$, sodass

$$d_n = (n+1)x^{n+1} \quad \text{für alle } n.$$

Damit gilt

$$\sum_{n \geq 1} n x^n = \sum_n d_n = \sum_n x^n \cdot \sum_n x^{n+1} = \frac{1}{1-x} \cdot \frac{x}{1-x} = \frac{x}{(1-x)^2}.$$

4.12 Die Exponentialreihe

Definition (Exponentialreihe)

Für alle $x \in \mathbb{R}$ heißt die Reihe $\sum_n x^n/n!$ die *Exponentialreihe* für x .

Die Exponentialreihe wird oft als die wichtigste Reihe der Analysis bezeichnet, und dieser Rang wird ihr allenfalls von der geometrischen Reihe streitig gemacht. Vieles, was wir gemacht haben, haben wir auch deswegen gemacht, um die Exponentialreihe zu verstehen. Und vieles, was wir tun werden, wird auf der Exponentialreihe beruhen, etwa die Definitionen von x^y , \log , \sin und \cos . Sie ist mehr als ein roter Faden der Analysis. Sie ist ein dickes rotes Seil.

Das Quotientenkriterium zeigt, dass die Exponentialreihe für alle x absolut konvergiert. Denn für alle x und n gilt

$$\left| \frac{x^{n+1}/(n+1)!}{x^n/n!} \right| = \frac{|x|}{n+1}$$

und der rechte Term strebt gegen 0, falls n gegen unendlich strebt.

Wir haben diesmal nicht nur eine Reihe betrachtet, sondern wir haben für jede reelle Zahl x eine konvergente Reihe vorliegen. Dies gibt Anlass zur Definition einer Funktion:

$$\exp: \mathbb{R} \rightarrow \mathbb{R}, \quad \exp(x) = \sum_n x^n/n! \quad \text{für alle } x.$$

Die *Exponentialfunktion* $\exp: \mathbb{R} \rightarrow \mathbb{R}$ werden wir in 6.3 genauer untersuchen und dann an zahlreichen Stellen einsetzen.

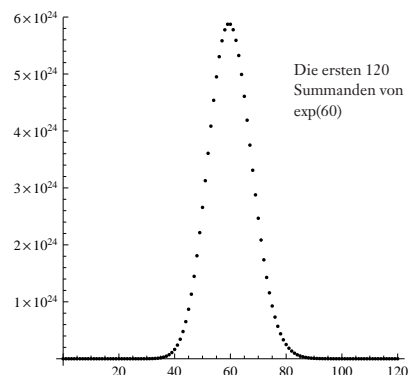
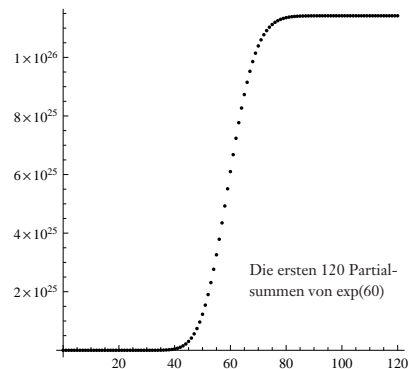
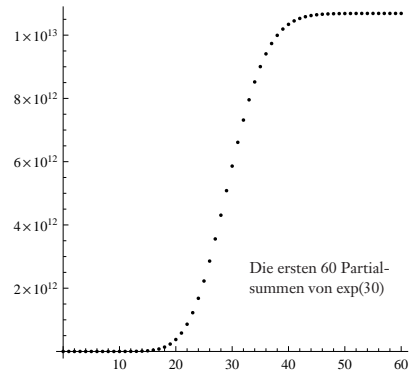
Die Exponentialreihe lässt sich auch zur Definition von e verwenden. Man setzt:

$$e = \exp(1) = \sum_n 1/n! = 1 + 1 + 1/2 + 1/6 + 1/24 + 1/120 + \dots \quad (\text{Eulersche Zahl})$$

Die ersten Partialsummen der Reihe $\sum_n 1/n!$ sind:

$$1; \quad 2; \quad 5/2 = 2,5; \quad 8/3 = 2,\overline{6}; \quad 65/24 = 2,708\overline{3}; \quad 163/60 = 2,71\overline{6}$$

Genauer gilt $e = 2,718281828459045\dots$. Die Eulersche Zahl ist transzendent (vgl. 2.2).



Die Exponentialreihe lässt sich auch in den komplexen Zahlen definieren. Hierzu übertragen wir die Theorie der Reihen nach \mathbb{C} . Eine komplexe Reihe $\sum_n z_n$ ist gegeben durch eine Folge $(z_n)_{n \in \mathbb{N}}$ in \mathbb{C} . Wir definieren sie wieder als Partialsummenfolge, d. h., es gilt

$$\sum_n z_n = (s_n)_{n \in \mathbb{N}} = (\sum_{k \leq n} z_k)_{n \in \mathbb{N}}. \quad (\sum_n z_n \text{ als Folge von Partialsummen})$$

Existiert der Grenzwert $\lim_n s_n = s \in \mathbb{C}$ der Partialsummen, so heißt die komplexe Zahl s die Summe der Reihe $\sum_n z_n$ und wir schreiben

$$\sum_n z_n = s. \quad (\sum_n z_n \text{ als Grenzwert})$$

Konvergiert $\sum_n z_n$, so gilt $\lim_n |z_n| = 0$.

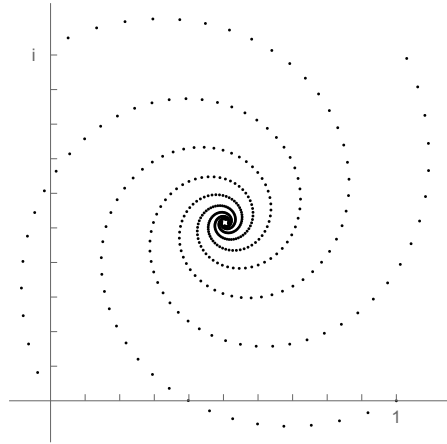
Beispiel

Für die geometrische Reihe gilt auch in \mathbb{C} :

$$\sum_{k \leq n} z^k = \frac{1 - z^{n+1}}{1 - z}, \quad \text{falls } z \neq 1,$$

$$\sum_n z^n = \frac{1}{1 - z}, \quad \text{falls } |z| < 1.$$

$\sum_{n \geq 1} z^n$ divergiert für $|z| \geq 1$, da dann die Beträge der Summanden wegen $|z^n| = |z|^n$ keine Nullfolge bilden.



Die komplexe geometrische Reihe für $z = 0,03 + i 0,99$.
Es gilt $1/(1 - z) \sim 0,504945 + i 0,515357$.

Ebenso gelten das Cauchy-Kriterium, das Majoranten- und Minorantenkriterium und die durch Majorisierung durch eine geometrische Reihe gewonnenen Kriterien, also das Wurzel- und Quotientenkriterium. In diesen Kriterien wird die reellwertige Betragsfunktion auf \mathbb{C} eingesetzt, sodass wir zum Beispiel die Wurzelbedingung

$${}^n\sqrt{|z_n|} \leq q \text{ für alle } n \geq n_0, \text{ mit einem } q < 1$$

übernehmen können, ohne uns über komplexe Wurzeln Gedanken zu machen. Auch nennen wir eine Reihe $\sum_n z_n$ in \mathbb{C} wieder absolut konvergent, wenn $\sum_n |z_n|$ konvergiert. Absolut konvergente Reihen in \mathbb{C} können beliebig umgeordnet werden und es gelten die Ergebnisse über die Produktbildung. Die Exponentialreihe in \mathbb{C} wird wie für \mathbb{R} definiert:

Definition (komplexe Exponentialreihe)

Für alle $z \in \mathbb{C}$ heißt die Reihe $\sum_n z^n/n!$ die *komplexe Exponentialreihe* für z .

Die komplexe Exponentialreihe konvergiert für alle $z \in \mathbb{C}$ gegen eine komplexe Zahl. Die zugehörige *komplexe Exponentialfunktion*

$$\exp : \mathbb{C} \rightarrow \mathbb{C}, \quad \exp(z) = \sum_n z^n/n! \quad \text{für alle } z \in \mathbb{C},$$

werden wir in 6.7 genauer betrachten. Zu den reellen Wachstumseigenschaften kommen ganz neuartige Dreieigenschaften hinzu, die dazu führen, dass sich die Exponentialfunktion zur Definition der trigonometrischen Funktionen und der Kreiszahl π eignet.

5. Kapitel

Stetigkeit

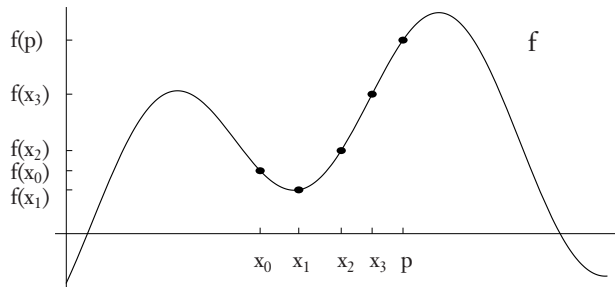
5.1 Die Limesstetigkeit

Definition (*Stetigkeit in einem Punkt und auf einer Menge*)

Sei $P \subseteq \mathbb{R}$, und sei $f: P \rightarrow \mathbb{R}$.

- (a) f heißt *stetig* in einem $p \in P$, falls für alle Folgen $(x_n)_{n \in \mathbb{N}}$ in P gilt:
 $\lim_n x_n = p$ impliziert $\lim_n f(x_n) = f(p)$.
- (b) f heißt *stetig auf* einer Menge $Q \subseteq P$, falls f stetig in allen $p \in Q$ ist.
- (c) f heißt *stetig*, falls f stetig auf P ist.

Wir setzen den Grenzwertbegriff für Folgen ein, um die Stetigkeit einer Funktion zu definieren. Das anschauliche Nichtspringen einer in einem Punkt p stetigen Funktion wird durch eine Folgenbedingung eingefangen. Wir verlangen, dass die Funktionswerte einer beliebigen Annäherung an p , die im Definitionsbereich der



Funktion zu erfolgen hat, gegen $f(p)$ konvergieren. Laufen wir im Definitionsbereich mit einer Folge $(x_n)_{n \in \mathbb{N}}$ auf den Punkt p zu, so läuft die Bildfolge $(f(x_n))_{n \in \mathbb{N}}$ im Wertebereich auf den Bildpunkt $f(p)$ zu. Im Stetigkeitsfall gilt also

$$\lim_n f(x_n) = f(\lim_n x_n).$$

Die Stetigkeit ist somit ein Beispiel für die Vertauschbarkeit von Operationen. Eine stetige Funktion respektiert Grenzwerte. Kurz:

Man darf einen Limes in eine stetige Funktion reinziehen.

Der Begriff der Stetigkeit ist ein Grundbegriff der Analysis. Er steht zwischen den beiden Säulen der Differenzierbarkeit und der Integrierbarkeit. Es wird sich zeigen:

Jede differenzierbare Funktion ist stetig und jede stetige Funktion ist integrierbar.

Der Stetigkeitsbegriff taucht auch außerhalb der reellen Analysis auf. In allgemeiner Form bildet er den Grundbegriff einer eigenen mathematischen Disziplin, nämlich der Topologie. (Die allgemeine Form deuten wir in 5.4 an.) Die Topologie untersucht stetige Abbildungen zwischen – oft sehr abstrakten – Räumen. Der Leser, der einen Teig in die Hand nimmt und ihn dehnt und knetet, ohne ihn zu zerreißen, hat, wenn er Ausgangs- und Endform vergleicht, ein „topologisches Beispiel“ für eine stetige Abbildung vor Augen. Eine Kugel kann, wie man in der Topologie zeigt, zum Beispiel nicht stetig in einen Autoreifen umgeformt werden.

Beispiele

- (1) Ist $c \in \mathbb{R}$, so ist die konstante Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = c$ für alle x stetig. Denn ist $\lim_n x_n = p$, so gilt

$$\lim_n f(x_n) = \lim_n c = c = f(p) (= f(\lim_n x_n)).$$

- (2) Die Identität $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x$ für alle x , ist stetig. Denn ist $\lim_n x_n = p$, so gilt

$$\lim_n f(x_n) = \lim_n x_n = p = f(p).$$

- (3) Sind $g, f: \mathbb{R} \rightarrow \mathbb{R}$ stetig, so auch $h = g \circ f$. Denn ist $\lim_n x_n = p$, so gilt

$$\lim_n h(x_n) = \lim_n g(f(x_n)) = g(\lim_n f(x_n)) = g(f(\lim_n x_n)) = g(f(p)).$$

Beim ersten Reinziehen des Limes verwenden wir, dass g stetig in $f(p) = \lim_n f(x_n)$ ist, beim zweiten, dass f stetig in $p = \lim_n x_n$ ist.

- (4) Sind $f, g: \mathbb{R} \rightarrow \mathbb{R}$ stetig und $a, b \in \mathbb{R}$, so ist auch $h = af + bg$ stetig. Denn ist $\lim_n x_n = p$, so gilt

$$\begin{aligned} \lim_n h(x_n) &= \lim_n (af(x_n) + bg(x_n)) = \\ &= a \lim_n f(x_n) + b \lim_n g(x_n) = af(p) + bg(p) = h(p). \end{aligned}$$

Dabei verwenden wir die Limesregeln für Folgen beim ersten Reinziehen des Limes, beim zweiten dann die Stetigkeit von f und g in p .

Die Beispiele zeigen, dass jede Polynomfunktion $f: \mathbb{R} \rightarrow \mathbb{R}$ stetig ist. Denn jede solche Funktion entsteht aus den konstanten Funktionen und der Identität durch wiederholte Bildung von Verknüpfungen und Linearkombinationen (vgl. 1.7). Beispiele für nichtstetige Funktionen besprechen wir in der folgenden Sektion. An dieser Stelle möchten wir noch auf das Thema „Definitionsbereich“ eingehen, das oft Schwierigkeiten bereitet, wenn die betrachtete Funktion nicht auf ganz \mathbb{R} definiert ist.

Beispiele

- (1) Sei $f:]-\infty, 0[\cup]0, +\infty[\rightarrow \mathbb{R}$ definiert durch $f(x) = -1$, falls $x < 0$, und $f(x) = 1$, falls $x > 0$. Dann ist die Funktion f stetig, obwohl sie an der Stelle 0 zu springen scheint. Die Funktion f ist aber im Nullpunkt gar nicht definiert, sodass hier die Stetigkeitsbedingung leer ist! Gleiches gilt für $g: \mathbb{R} - \{0\} \rightarrow \mathbb{R}$ mit $g(x) = 1/x$ für alle $x \neq 0$. Auch diese Funktion ist stetig.
- (2) Sei $P = \{p_1, \dots, p_n\}$ mit $p_1 < \dots < p_n$. Dann ist jede Funktion $f: P \rightarrow \mathbb{R}$ stetig. Denn gilt $\lim_n x_n = p_i$ in P , so ist die Folge $(x_n)_{n \in \mathbb{N}}$ schließlich konstant gleich p_i und also $\lim_n f(x_n) = f(p_i)$.

Bitte also abspeichern:

Für Stetigkeitsbetrachtungen ist die Angabe und Beachtung des Definitionsbereichs P der untersuchten Funktion unerlässlich.

5.2 Grenzwerte von Funktionen

Definition (Grenzwerte und Limesnotationen für Funktionen)

Sei $f: P \rightarrow \mathbb{R}$ eine Funktion. Weiter sei $p \in \mathbb{R}$ ein *Berührungspunkt* von P , d. h., es gebe eine Folge $(x_n)_{n \in \mathbb{N}}$ in P mit $\lim_n x_n = p$. Schließlich sei $y \in \mathbb{R}$. Dann schreiben wir

$\lim_{x \rightarrow p} f(x) = y$, falls für alle gegen p konvergenten Folgen

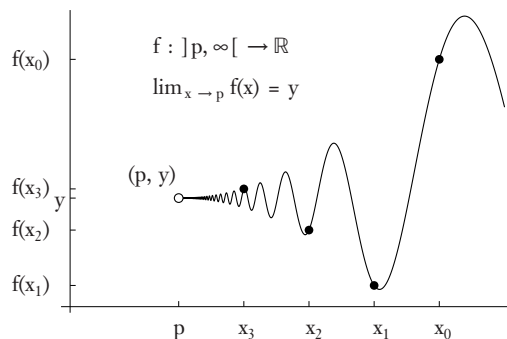
$(x_n)_{n \in \mathbb{N}}$ in P gilt, dass $\lim_n f(x_n) = y$.

(Grenzwertbedingung für Funktionen)

Wir nennen dann y den *Grenzwert* von f an der Stelle p und sagen, dass die Funktion f *gegen y strebt*, falls x in P *gegen p strebt*. Statt „ $\lim_{x \rightarrow p} f(x) = y$ “ schreiben wir auch „ $f(x) \rightarrow y$ für $x \rightarrow p$ “.

Damit haben wir Grenzwerte für Funktionen, die wir anschaulich begründet aus der Schule kennen, auf Grenzwerte für Folgen zurückgeführt – und damit exakt definiert.

In der Definition verlangen wir nicht, dass $p \in P = \text{Def}(f)$ gilt. Es muss aber eine Folge in P existieren, die gegen p konvergiert. Gilt $p \in P$, so ist dies immer der Fall, wie die konstante Folge $(p)_{n \in \mathbb{N}}$ zeigt. Ist $p \notin P$, so ist dies der Fall, wenn p ein Häufungspunkt von P ist, d. h. wenn in jeder ε -Umgebung von p unendlich viele Punkte von P liegen.



Beispiele

(1) 1 ist ein Berührungspunkt von $[0, 1/2] \cup \{1\}$ und von $[0, 1[$, nicht aber von $[0, 9/10]$.

(2) Sei $f: [0, 1[\rightarrow \mathbb{R}$ die Funktion mit $f(x) = x^2/2$ für alle $x \in [0, 1[$.

Dann gilt $\lim_{x \rightarrow 1} f(x) = 1/2$.

(3) Sei $g: [0, 1[\cup]1, 2[\rightarrow \mathbb{R}$ mit $g(x) = x^2$ für alle $x \in [0, 1[$ und $g(x) = 2$ für alle $x \in]1, 2[$. Dann existiert $\lim_{x \rightarrow 1} g(x)$ nicht, dagegen ist $\lim_{x \rightarrow 2} g(x) = 2$.

(4) Die Limesnotation können wir auch auf Terme anwenden, etwa in

$$\lim_{x \rightarrow 1} x^2 + 2 = 3.$$

Als Faustregel gilt, dass der Definitionsbereich der durch den Term definierten Funktion so groß wie möglich ist, für $x^2 + 2$ also \mathbb{R} , und in $\lim_{x \rightarrow 0} \sin(x)/x = 1$ also $\mathbb{R} - \{0\}$. Im Zweifel gibt man den Definitionsbereich besser an.

Der Leser wird sich nun vielleicht fragen, warum wir Grenzwerte von Funktionen im Kapitel über Stetigkeit besprechen. Wir hätten sie in der Tat früher einführen können, doch ihre Bedeutung für den Stetigkeitsbegriff beruht auf der folgenden eleganten und suggestiven Formulierung der Stetigkeit:

Stetigkeit als Grenzwerteigenschaft von Funktionen

Ein $f: P \rightarrow \mathbb{R}$ ist genau dann stetig in $p \in P$, wenn $\lim_{x \rightarrow p} f(x) = f(p)$.

Hier wird $p \in P$ vorausgesetzt. Der Fall $p \notin P$ spielt bei stetigen Fortsetzungen eine Rolle, die wir in 5.7 besprechen.

Der Grenzwertbegriff für Funktionen besitzt viele Spielarten. In der folgenden Tabelle ist $f: P \rightarrow \mathbb{R}$, $y \in [-\infty, +\infty]$ und es soll stets eine Folge geben, die unter den betrachteten Bedingungen (eigentlich oder uneigentlich) gegen p konvergiert.

$\lim_{x \rightarrow p, x \neq p} f(x) = y$	Für alle $(x_n)_{n \in \mathbb{N}}$ in $P - \{p\}$ mit $\lim_n x_n = p$ gilt $\lim_n f(x_n) = y$.
$\lim_{x \downarrow p} f(x) = y$	Für alle $(x_n)_{n \in \mathbb{N}}$ in $P \cap]p, \infty[$ mit $\lim_n x_n = p$ gilt $\lim_n f(x_n) = y$.
$\lim_{x \uparrow p} f(x) = y$	Für alle $(x_n)_{n \in \mathbb{N}}$ in $P \cap]-\infty, p[$ mit $\lim_n x_n = p$ gilt $\lim_n f(x_n) = y$.
$\lim_{x \rightarrow \infty} f(x) = y$	Für alle $(x_n)_{n \in \mathbb{N}}$ in P mit $\lim_n x_n = +\infty$ gilt $\lim_n f(x_n) = y$.
$\lim_{x \rightarrow -\infty} f(x) = y$	Für alle $(x_n)_{n \in \mathbb{N}}$ in P mit $\lim_n x_n = -\infty$ gilt $\lim_n f(x_n) = y$.

Die Ausdrücke „ $x \downarrow p$ “ und „ $x \uparrow p$ “ lesen wir als „ x strebt von oben gegen p “ bzw. als „ x strebt von unten gegen p “. Der entsprechende Grenzwert y heißt auch der *rechtsseitige* bzw. *linksseitige Grenzwert* von f an der Stelle p . Für diese beiden Konvergenztypen kann man auch nur monoton fallende bzw. monoton wachsende Folgen zulassen. Es gilt:

Stetigkeitsformulierung über links- und rechtsseitige Grenzwerte

Ein $f: P \rightarrow \mathbb{R}$ ist genau dann stetig in $p \in P$, falls gilt:

$$\lim_{x \uparrow p} f(x) = f(p) = \lim_{x \downarrow p} f(x),$$

wobei eine Limesbedingung zu streichen ist, wenn keine zugehörige Folge existiert (z. B. die rechte für $p = +\infty$ oder für $p = \sup(P)$, oder beide für $P = \{p\}$).

Beispiele

(1) Sei $f:]0, 1[\cup]1, +\infty[\rightarrow \mathbb{R}$ mit $f(x) = -1/x$ für alle x . Dann gilt

$$\lim_{x \uparrow 1} f(x) = -1 = \lim_{x \downarrow 1} f(x), \quad \lim_{x \rightarrow 0} f(x) = -\infty, \quad \lim_{x \uparrow \infty} f(x) = 0.$$

(2) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = 0$ für $x \leq 0$, $f(x) = 1$ für $x > 0$. Dann gilt

$$\lim_{x \uparrow 0} f(x) = 0, \quad \lim_{x \downarrow 0} f(x) = 1.$$

Warnung: Unterschiedliche Konventionen

In der Literatur sind in „ $\lim_{x \rightarrow p} f(x) = y$ “ oft nur Folgen in $P - \{p\}$ zugelassen, d. h., der Ausdruck bedeutet „ $\lim_{x \rightarrow p, x \neq p} f(x) = y$ “ in unserer Notation. Dies kann eine Fehlerquelle sein, denn ist $p \in P$ und existiert $\lim_{x \rightarrow p} f(x) = y$ in unserer Notation, so ist $y = \lim_n f(p) = f(p)$, d. h., f ist stetig in p . Dagegen folgt aus $\lim_{x \rightarrow p, x \neq p} f(x) = y$ nicht notwendig, dass $y = f(p)$.

5.3 Unstetigkeiten

Satz (Unstetigkeit in einem Punkt)

Sei $f:]a, b[\rightarrow \mathbb{R}$ eine Funktion mit $-\infty \leq a < b \leq +\infty$, und sei $p \in]a, b[$. Dann ist f genau dann unstetig im Punkt p , wenn einer der drei folgenden Fälle zutrifft:

Unstetigkeit der Art 1a

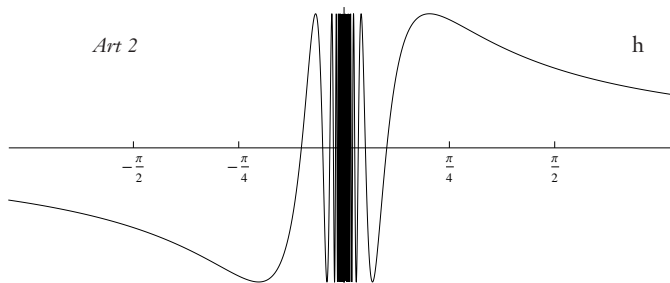
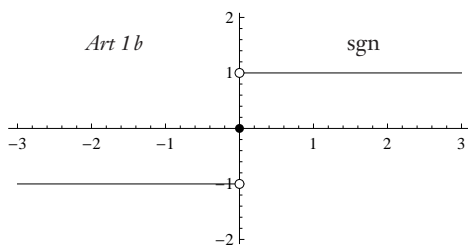
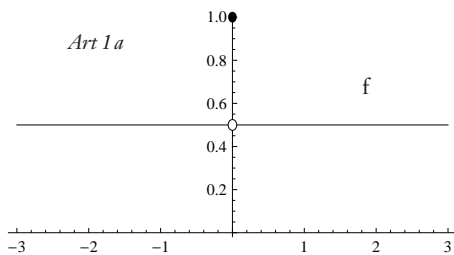
$\lim_{x \rightarrow p, x \neq p} f(x) = y$ existiert, aber es gilt $f(p) \neq y$.

Unstetigkeit der Art 1b

$\lim_{x \uparrow p} f(x) = y_1$ und $\lim_{x \downarrow p} f(x) = y_2$ existieren, aber es gilt $y_1 \neq y_2$.

Unstetigkeit der Art 2

$\lim_{x \uparrow p} f(x)$ existiert nicht oder $\lim_{x \downarrow p} f(x)$ existiert nicht.



Der Satz ist im Wesentlichen nur die Verneinung der Limesdefinition der Stetigkeit. Dennoch sind die Unstetigkeitsphänomene so vielfältig, dass sie eine eigene Betrachtung verdienen, und der Satz erlaubt eine Klassifikation des Typs der Unstetigkeit.

Wir betrachten zunächst die Art 1a. Hier existiert $\lim_{x \rightarrow p, x \neq p} f(x) = y$, aber die Funktion f verhält sich an der Stelle p selbst bockig. Sie tut dort nicht, was sie tun sollte.

Beispiel für Art 1a

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ die Funktion mit $f(x) = 1/2$ für alle $x \neq 0$ und $f(0) = 1$. Dann ist 0 eine Unstetigkeitsstelle der Art 1a, da für jede Folge $(x_n)_{n \in \mathbb{N}}$ in $\mathbb{R} - \{0\}$ mit $\lim_n x_n = 0$ gilt, dass $\lim_n f(x_n) = 1/2 \neq 1 = f(0)$. Ändern wir f an der Stelle 0 ab und setzen wir dort $1/2$ als neuen Wert ein, so erhalten wir eine im Nullpunkt stetige Funktion.

Bei einer Unstetigkeit der Art 1 b existieren zwar der links- und der rechtsseitige Grenzwert, aber sie sind verschieden. In diesem Fall führt dann auch keine Abänderung der Funktion an der Stelle p mehr zur Stetigkeit im Punkt p .

Beispiel für Art 1 b

Sei $\operatorname{sgn} : \mathbb{R} \rightarrow \mathbb{R}$ die Vorzeichenfunktion Signum mit

$$\operatorname{sgn}(x) = -1 \text{ für } x < 0, \quad \operatorname{sgn}(0) = 0, \quad \operatorname{sgn}(x) = 1 \text{ für } x > 0.$$

Dann hat sgn eine Unstetigkeitsstelle des Typs 1 b im Nullpunkt, da

$$\lim_{x \uparrow 0} \operatorname{sgn}(x) = -1 \neq 1 = \lim_{x \downarrow 0} \operatorname{sgn}(x).$$

Man kann die Funktion dieses Beispiels nicht als bockig bezeichnen, denn der Wert $\operatorname{sgn}(0) = 0$ ist der Mittelwert der Grenzwerte bei Annäherung von links bzw. von rechts. Die Funktion ist unstetig, aber sie tut ihr Bestes, um die Ansprüche von links und von rechts auszugleichen. Diese „vermittelnde Unstetigkeit“ taucht zum Beispiel in der Theorie der Fourier-Reihen auf.

Die Unstetigkeiten erster Art sind durch die Existenz der links- und rechtsseitigen Grenzwerte ausgezeichnet. Komplizierter sind die Unstetigkeiten zweiter Art, für die einer der beiden Grenzwerte nicht existiert:

Beispiel für Art 2

Sei $h : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$h(x) = \sin(1/x) \text{ für } x \neq 0, \quad h(0) = 0.$$

Dann ist 0 eine Unstetigkeitsstelle der zweiten Art. Für alle $c \in]-1, 1[$ existiert eine Nullfolge $(x_n)_{n \in \mathbb{N}}$ in $]0, +\infty[$ mit $\lim_n h(x_n) = c$. Damit existiert $\lim_{x \downarrow 0} h(x)$ nicht. Analoges gilt für die Annäherung von links.

Wir betrachten nun noch Funktionen, die sehr viele Unstetigkeitsstellen besitzen.

Beispiele für viele Unstetigkeitsstellen

- (1) Die *Dirichletsche Sprungfunktion* $f : \mathbb{R} \rightarrow \mathbb{R}$ ist definiert durch

$$f(x) = 1, \text{ falls } x \text{ rational, } f(x) = 0, \text{ falls } x \text{ irrational.}$$

Diese Funktion besitzt in allen $p \in \mathbb{R}$ eine Unstetigkeitsstelle der zweiten Art, da sowohl Folgen in \mathbb{Q} als auch Folgen in $\mathbb{R} - \mathbb{Q}$ existieren, die gegen p konvergieren.

- (2) Ist $g : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch $g(x) = x$, falls x rational, und $g(x) = -x$, falls x irrational, so ist f im Nullpunkt stetig und in allen anderen Punkten unstetig.
- (3) Sei $h : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch $h(x) = 0$, falls x irrational, $h(m/n) = 1/n$ für gekürzte Brüche m/n mit $n \geq 1$. Dann ist h genau dann stetig in einem Punkt p , wenn p irrational ist. Die Funktion h heißt auch die *Thomae-Funktion*.

Dagegen kann man zeigen, dass es keine Funktion gibt, die genau in allen rationalen Punkten stetig ist. Vieles ist möglich im Reich der Unstetigkeit, aber nicht alles!

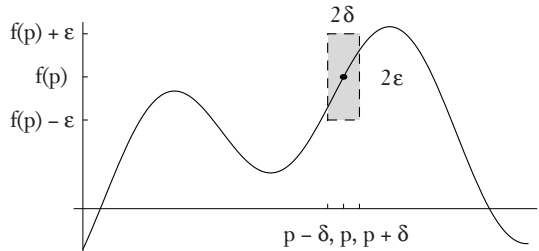
5.4 Die Umgebungsstetigkeit

Definition (Umgebungsstetigkeit oder ε - δ -Stetigkeit)

Sei $f: P \rightarrow \mathbb{R}$. Dann heißt f *umgebungsstetig* oder *ε - δ -stetig* in einem $p \in P$, falls gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in P (|x - p| < \delta \rightarrow |f(x) - f(p)| < \varepsilon). \quad (\varepsilon\text{-}\delta\text{-Bedingung})$$

Da ist sie wieder, die Epsilontik! Diesmal ist die Bedingung aber anders gebaut als die Konvergenzbedingung für Folgen. Nun sind zwei positive reelle Variablen ε und δ im Spiel, natürliche Zahlen sind dagegen nicht involviert. Die Variable ε betrifft Abstände von Funktionswerten, die Variable δ Abstände von Stellen. Wir halten also fest:



δ ist gut für ε :

Alle Punkte in $]p - \delta, p + \delta[$ landen in $]f(p) - \varepsilon, f(p) + \varepsilon[$.

Die ε -Variable lebt auf der y-Achse, die δ -Variable lebt auf der x-Achse.

Dem Allquantor über ε folgt ein Existenzquantor über δ . Wollen wir die ε - δ -Stetigkeit von f in p zeigen, so müssen wir also für ein beliebiges $\varepsilon > 0$ ein „gutes“ $\delta > 0$ finden, das im Allgemeinen von ε abhängen wird. Das kennen wir schon von der Konvergenzbedingung für Folgen, wo wir für jedes $\varepsilon > 0$ ein gutes n_0 finden mussten. Dort verlangten kleine ε nach großen n_0 . Die Faustregel lautet dagegen jetzt: Kleine ε lieben kleine δ . Unser δ muss derart sein, dass alle Punkte in $]p - \delta, p + \delta[$ durch f nach $]f(p) - \varepsilon, f(p) + \varepsilon[$ transportiert werden. Das lokale Wachstumsverhalten der Funktion f bestimmt die Größe von δ . Je stärker sich f in der Nähe von p verändert, desto kleiner ist δ zu wählen. Typische Werte für ein gegebenes ε sind $\delta = \varepsilon$, $\delta = \varepsilon/2$, $\delta = \varepsilon^2$ usw.

Beispiel

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = 10^{10}x$ für alle x . Wir zeigen die ε - δ -Stetigkeit von f im Punkt $p = 0$. Sei also $\varepsilon > 0$ beliebig. Wir setzen $\delta = \varepsilon/10^{10}$. Dieses δ ist gut für ε . Denn sei $x \in \mathbb{R}$ mit $|x - 0| = |x| < \delta$. Dann gilt

$$|f(x) - f(0)| = |10^{10}x - 0| = |10^{10}x| < 10^{10}\delta = \varepsilon.$$

Das Argument für die Limesstetigkeit lautet dagegen

$$\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} 10^{10}x = 0 = f(0).$$

Will man zeigen, dass eine Funktion f in p nicht ε - δ -stetig ist, so zeigt man:

$$\exists \varepsilon > 0 \forall \delta > 0 \exists x \in P (|x - p| < \delta \wedge |f(x) - f(p)| \geq \varepsilon). \quad (\text{negierte } \varepsilon\text{-}\delta\text{-Bedingung})$$

Neben den Verneinungsregeln für Quantoren aus 1.10 haben wir hier benutzt, dass eine Aussage $\neg(A \rightarrow B)$ äquivalent zu $A \wedge \neg B$ ist (vgl. Anhang 1).

Unsere beiden Stetigkeitsbegriffe sind, so verschieden sie sich anfühlen mögen, punktweise äquivalent:

Äquivalenz der Stetigkeitsbegriffe

Sei $f : P \rightarrow \mathbb{R}$ und sei $p \in P$. Dann ist f im Punkt p genau dann limesstetig, wenn f im Punkt p umgebungsstetig ist.

Dass wir zwei auf den ersten Blick recht verschiedene Formulierungen gefunden haben, die sich als äquivalent herausstellen, ist ein Indiz dafür, einen natürlichen und wichtigen mathematischen Begriff gefunden zu haben. Prinzipiell kann man sich immer aussuchen, mit welcher Version man arbeiten möchte. Dennoch ist es wichtig, beide Versionen zu kennen. Die ε - δ -Stetigkeit gibt, wie wir in den nächsten Sektionen sehen werden, Anlass zu den starken Formen der gleichmäßigen Stetigkeit und der Lipschitz-Stetigkeit, die in der Analysis als gute Voraussetzung oft eine zentrale Rolle spielen.

Die ε - δ -Bedingung lässt sich mit Umgebungen sehr ansprechend formulieren. Wir verwenden dabei ε -Umgebungen von $f(p)$ und δ -Umgebungen von p :

$$V_\varepsilon(f(p)) = \{y \in \mathbb{R} \mid |y - f(p)| < \varepsilon\},$$

$$U_\delta(p) = \{x \in \mathbb{R} \mid |x - p| < \delta\}.$$

Umgebungsformulierung der ε - δ -Stetigkeit

Für jede ε -Umgebung $V_\varepsilon(f(p))$ von $f(p)$ existiert eine δ -Umgebung $U_\delta(p)$ von p mit

$$f[U_\delta(p) \cap P] = \{f(x) \mid x \in U_\delta(p) \cap P\} \subseteq V_\varepsilon(p).$$

Kurz: Im Urbild jeder ε -Umgebung von $f(p)$ findet sich eine δ -Umgebung von p .

Diese Formulierung ist die Grundlage für den allgemeinen Stetigkeitsbegriff der Topologie. Sobald „Umgebung eines Punktes“ erklärt ist, kann die Stetigkeit einer Funktion erklärt werden.

Definition und Anschauung

Beide Formulierungen der Stetigkeit unterstützen die Anschauungen:

- (I) *Eine stetige Funktion macht keine Sprünge.*
- (II) *Die Funktionswerte $f(x)$ einer stetigen Funktion ändern sich wenig, wenn sich die Stellen x hinreichend wenig ändern.*

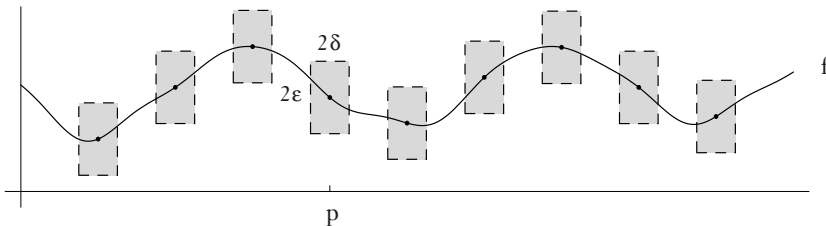
Dabei scheint die Limesstetigkeit die direktere Übersetzung von (I), die Umgebungsstetigkeit die direktere Übersetzung von (II) zu sein. Dagegen unterstützt der entwickelte Stetigkeitsbegriff keineswegs, dass die stetigen Funktionen diejenigen Funktionen sind, die wir mit dem Stift ohne abzusetzen zeichnen können. Sicher sind diese zeichenbaren Funktionen stetig, aber der Stetigkeitsbegriff lässt weit mehr stetige Funktionen zu. Es gibt zum Beispiel stetige Funktionen, die in keinem Punkt differenzierbar sind. Derartige Funktionen kann man nicht zeichnen. Ein Beispiel werden wir in Sektion 7.5 kennenlernen.

5.5 Die gleichmäßige Stetigkeit

Definition (gleichmäßige Stetigkeit)

Eine Funktion $f: P \rightarrow \mathbb{R}$ heißt *gleichmäßig stetig*, falls gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall p \in P \forall x \in P (|x - p| < \delta \rightarrow |f(x) - f(p)| < \varepsilon). \quad (\text{gleichmäßige } \varepsilon\text{-}\delta\text{-Bedingung})$$



Zum Vergleich der beiden ε - δ -Bedingungen schreiben wir noch einmal auf, was es heißt, dass eine Funktion $f: P \rightarrow \mathbb{R}$ ε - δ -stetig ist:

$$\forall p \in P \forall \varepsilon > 0 \exists \delta > 0 \forall x \in P (|x - p| < \delta \rightarrow |f(x) - f(p)| < \varepsilon).$$

Da wir zwei Allquantoren vertauschen dürfen, ist dies äquivalent zu:

$$\forall \varepsilon > 0 \forall p \in P \exists \delta > 0 \forall x \in P (|x - p| < \delta \rightarrow |f(x) - f(p)| < \varepsilon).$$

Nun ist der Unterschied zur gleichmäßigen Stetigkeit leicht zu sehen. In der gleichmäßigen Stetigkeit sind die Quantoren $\forall p \in P$ und $\exists \delta > 0$ vertauscht. Nach den Quantorenregeln (vgl. 1.10) ist diese Vertauschung in der Richtung von der gleichmäßigen zur normalen Form erlaubt. Mit anderen Worten:

Die gleichmäßige Stetigkeit ist eine Verstärkung der ε - δ -Stetigkeit.

In der gleichmäßigen Stetigkeit hängt δ immer noch von ε ab, aber die Abhängigkeit von p ist verschwunden. Ist $\varepsilon > 0$, so gibt es ein $\delta > 0$, das für alle $p \in P$ gut ist.

Beispiel

Sei $f: [0, 1] \rightarrow \mathbb{R}$ definiert durch $f(x) = x^2$ für alle x . Dann ist f gleichmäßig stetig. Denn sei $\varepsilon > 0$ beliebig. Sei $\delta = \varepsilon/2$. Wir zeigen, dass δ gut für ε ist. Seien also $x, p \in [0, 1]$ mit $|x - p| < \delta$. Dann gilt $x + p \leq 2$, da $x, p \in [0, 1]$. Also gilt

$$|f(x) - f(p)| = |x^2 - p^2| = |(x + p)(x - p)| \leq 2|x - p| < 2\delta = \varepsilon.$$

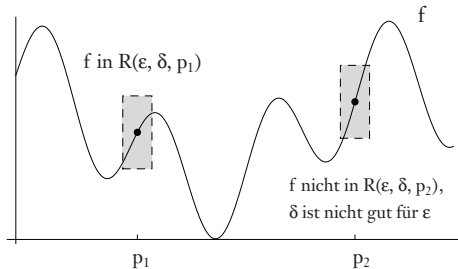
Die gleichmäßige Stetigkeit ist ihrer Natur nach ein globales Konzept. Die Stetigkeit einer Funktion $f: P \rightarrow \mathbb{R}$ hatten wir zunächst in einem Punkt $p \in P$ definiert, und dann hatten wir f stetig genannt, wenn f in jedem Punkt $p \in P$ stetig ist. Dagegen ergibt es keinen Sinn zu sagen, eine Funktion f sei im Punkt p gleichmäßig stetig. Die gleichmäßige Stetigkeit ist eine Eigenschaft, die alle Punkte des Definitionsbereichs von f betrifft.

Um uns den Begriff der gleichmäßigen Stetigkeit zu veranschaulichen, betrachten wir wieder die offenen Rechtecke der Form

$$R(\varepsilon, \delta, p) =]p - \delta, p + \delta[\times]f(p) - \varepsilon, f(p) + \varepsilon[$$

der Breite 2δ und der Höhe 2ε , die den Punkt $(p, f(p))$ in ihrem Zentrum haben. Die gleichmäßige Stetigkeit von f bedeutet, dass für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, sodass für alle $p \in P$ der Graph der Funktion im Streifen $]p - \delta, p + \delta[\times \mathbb{R}$ ganz im Rechteck $R(\varepsilon, \delta, p)$ liegt.

Die gleichmäßige Stetigkeit tritt recht häufig auf. Man kann mit Hilfe des Satzes von Bolzano-Weierstraß zeigen, dass jede stetige Funktion $f: [a, b] \rightarrow \mathbb{R}$ gleichmäßig stetig ist. Für offene Intervalle $]a, b[$ existieren Gegenbeispiele:



Beispiel

Sei $f:]0, 1[\rightarrow \mathbb{R}$ mit $f(x) = 1/x$. Dann ist f stetig, aber nicht gleichmäßig stetig. Denn sei $\varepsilon = 1$ und $\delta > 0$ beliebig. Wir betrachten nun ein beliebiges x derart, dass $x < \delta$ und $2x < 1$ gilt. Weiter sei dann $p = 2x$. Dann gilt $|x - p| = x < \delta$, aber

$$|f(x) - f(p)| = \frac{1}{x} - \frac{1}{2x} = \frac{1}{2x} = f(2x) = f(p) > 1 = \varepsilon.$$

Die Funktion f des Beispiels wächst zu schnell, um die Rechtecksbedingung erfüllen zu können, so klein δ auch gewählt wird. Neugierig fragen wir nun:

Gibt es Gegenbeispiele zur gleichmäßigen Stetigkeit mit beschränktem Wertebereich?

Die Antwort ist: Ja. Eine schnell oszillierende Funktion, wie sie uns bereits bei der Diskussion der Unstetigkeitsstellen zweiter Art begegnet ist, liefert ein solches Gegenbeispiel:

Beispiel

Sei $f:]0, 1[\rightarrow \mathbb{R}$ die Funktion mit $f(x) = \sin(1/x)$ für alle $x \in]0, 1[$ (vgl. 5.3). Dann ist f stetig, aber nicht gleichmäßig stetig. Die Frequenz der Funktion steigt bei Annäherung an die Null beliebig an, und die Funktion durchsticht, für jedes noch so kleine $\delta > 0$, die Rechtecke $R(1/2, \delta, p)$ in den Waagrechten, wenn p nahe genug an der Null gewählt wird.

Umgekehrt können auch unbeschränkte Funktionen gleichmäßig stetig sein. Die Identität auf \mathbb{R} ist ein Beispiel. Die Quadratfunktion oder auch das Produkt der Identität mit dem Sinus, also die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x \sin(x)$, zeigen, dass das Produkt zweier gleichmäßig stetiger Funktionen nicht mehr gleichmäßig stetig sein muss.

Die gleichmäßige Stetigkeit spielt in der Integrationstheorie eine wichtige Rolle im Beweis des Satzes, dass sich jede stetige, auf einem Intervall $[a, b]$ definierte Funktion gleichmäßig durch eine Treppenfunktion approximieren lässt und folglich integrierbar ist (vgl. 8.6).

5.6 Die Lipschitz-Stetigkeit

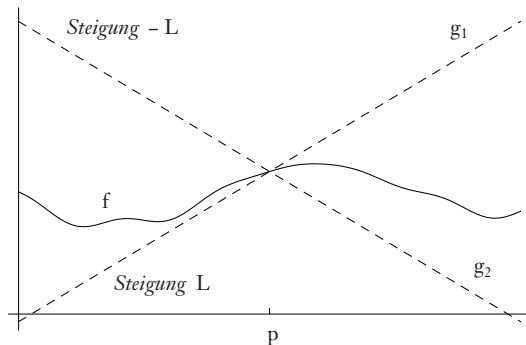
Definition (Lipschitz-Stetigkeit)

Eine Funktion $f : P \rightarrow \mathbb{R}$ heißt *Lipschitz-stetig* oder *dehnungsbeschränkt*, falls ein $L \geq 0$ existiert mit

$$\forall p \in P \quad \forall x \in P \quad |f(x) - f(p)| \leq L |x - p|. \quad (\text{Lipschitz-Bedingung für } L)$$

Die Zahl L heißt dann eine *Lipschitz-Konstante* für f .

In der Lipschitz-Bedingung sind die reellen Variablen ε und δ verschwunden, die Abschätzung der Abstände übernimmt eine Konstante L . Die Bedingung besagt, dass der Abstand je zweier Punkte p und x durch Anwendung der Funktion höchstens um den Faktor L gestreckt wird (wobei Lipschitz-Konstanten $L < 1$ einer Stauchung entsprechen). Diese Beschreibung erklärt auch den alternativen Namen der Dehnungsbeschränktheit.



Obwohl die Form der Lipschitz-Stetigkeit viel einfacher ist als die ε - δ -Stetigkeit mit ihren zahlreichen Quantorenwechseln, bereitet dieser Begriff vielen Anfängern zunächst Schwierigkeiten, da sie sich unter der Lipschitz-Bedingung nicht viel vorstellen können. Eine Variation der Rechteckmethode, mit der wir die gleichmäßige Stetigkeit anschaulich machen konnten, ist geeignet, um sich mit der Bedingung anzufreunden. Sei hierzu $p \in P$. Ist f Lipschitz-stetig mit der Konstanten L , so gilt für alle $x \neq p$:

$$\left| \frac{f(x) - f(p)}{x - p} \right| \leq L.$$

Die linke Seite kennen wir aus der Schule. Sie ist der Betrag des Differenzenquotienten von f für die Punkte x und p (in 7.1 werden wir ihn genauer besprechen). Die Bedingung besagt also, dass für alle $x \neq p$ die Gerade, die die Punkte $(p, f(p))$ und $(x, f(x))$ verbindet, eine Steigung im Intervall $[-L, L]$ besitzt. Damit gilt also die folgende anschauliche graphische Interpretation der Lipschitz-Stetigkeit mit der Konstanten L :

Sei $p \in P$ beliebig, und seien g_1 und g_2 die beiden Geraden der Ebene durch den Punkt $(p, f(p))$ mit der Steigung $-L$ bzw. L . Dann liegt der gesamte Funktionsgraph von f im durch die beiden Geraden eingeschlossenen waagrechten Bereich.

Genauer zerlegen die beiden Geraden für $L > 0$ die Ebene in vier Teile. Wir betrachten die beiden Teile, die die Waagrechte durch den Punkt $(p, f(p))$ enthalten. Ist $L = 0$, so ist diese Waagrechte identisch mit dem betrachteten Bereich. Allgemein gilt: Ist L sehr klein, so besitzt der betrachtete Bereich sehr spitze Winkel im Punkt $(p, f(p))$ und das Wachstum der Funktion ist dann sehr gering.

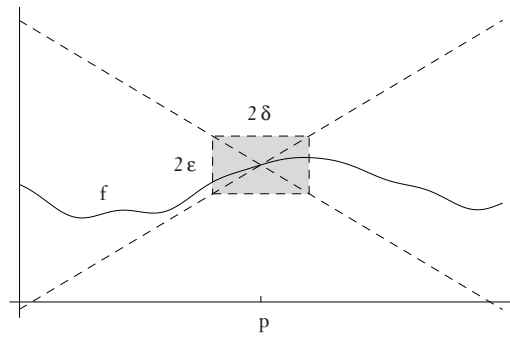
Wer es noch anschaulicher möchte, betrachte eine Taschenlampe mit dem Öffnungswinkel $2 \arctan(L)$. Die Lipschitz-Stetigkeit zur Konstanten L bedeutet dann: Legen wir die Taschenlampe parallel zur x -Achse irgendwo auf den Funktionsgraphen, so liegt der gesamte Funktionsgraph im nach vorne oder nach hinten gerichteten Lichtkegel.

Diese Überlegungen zeigen, dass man im Unterschied zur gleichmäßigen Stetigkeit auch wieder eine punktweise Form der Lipschitz-Stetigkeit betrachten könnte. Wir beschränken uns hier auf die globale Version, bei der eine Konstante für jedes Punktepaar p und x geeignet ist.

Nachdem wir uns nun mit dem Konzept angefreundet haben, ordnen wir es in die anderen Stetigkeitsbegriffe ein. Die Lipschitz-Stetigkeit ist sehr stark:

Lipschitz-Stetigkeit impliziert gleichmäßige Stetigkeit, aber nicht umgekehrt.

Die gültige Implikation kann man sich durch einen Vergleich der beiden Visualisierungen vor Augen führen: Liegt der Graph punktweise im durch zwei Geraden der Steigung L bzw. $-L$ eingeschlossenen Bereich, so erfüllt er auch die Rechtecksbedingung der gleichmäßigen Stetigkeit. Die andere Implikation gilt dagegen nicht:



Beispiel

Sei $f: [0, 1] \rightarrow \mathbb{R}$ definiert durch $f(x) = \sqrt{x}$ für alle x . Dann ist f gleichmäßig stetig. Dies folgt aus dem allgemeinen Satz oder durch direkten Nachweis: Sei $\varepsilon > 0$ beliebig. Aufgrund der Stetigkeit von f im Punkt 0 gibt es ein δ mit $|\sqrt{s} - \sqrt{0}| = \sqrt{s} < \varepsilon$ für alle $s \in [0, \delta]$. Dann ist δ aber auch für die gleichmäßige Stetigkeitsbedingung geeignet, denn für alle $p \leq x$ in $[0, 1]$ mit $x - p < \delta$ gilt

$$|f(x) - f(p)| = \sqrt{x} - \sqrt{p} \leq \sqrt{x - p} < \varepsilon,$$

wobei die erste Ungleichung durch Quadrieren aus $p \leq x$ folgt. Die Funktion f ist aber nicht Lipschitz-stetig. Denn die Wurzelfunktion beginnt im Nullpunkt mit einer unendlichen Steigung, und diese Steigung macht die Lipschitz-Stetigkeit unmöglich, da wir die Steigung unserer den Graphen einschließenden Geraden im Nullpunkt ebenfalls unendlich groß wählen müssten.

Dennoch gilt, dass die Lipschitz-Stetigkeit häufiger anzutreffen ist, als man vielleicht vermuten würde. Kennt man etwas Differentialrechnung, so kann man zeigen, dass jede stetig differenzierbare Funktion $f: [a, b] \rightarrow \mathbb{R}$ Lipschitz-stetig ist (siehe 7.7). Der Definitionsbereich hat hier die gute Form $[a, b]$, f ist differenzierbar und die Ableitung f' ist stetig. Dies ist zum Beispiel für den Sinus auf $[0, 2\pi]$ oder den Logarithmus auf $[1, 2]$ der Fall. Die Lipschitz-Stetigkeit taucht in der Analysis als „gute Voraussetzung“ an wichtigen Stellen auf, etwa im Existenz- und Eindeigkeitsatz der Theorie der Differentialgleichungen und im Banachschen Fixpunktsatz.

5.7 Stetige Fortsetzungen

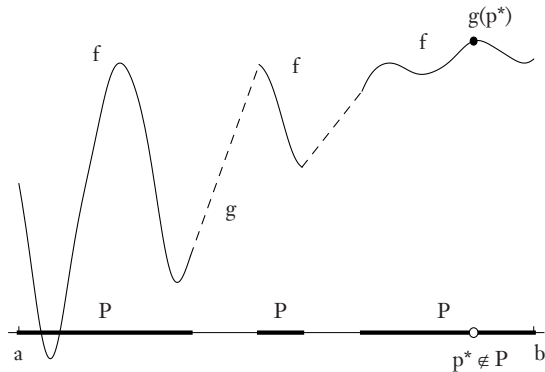
Definition (stetig fortsetzbar; stetige Fortsetzung)

Seien $f : P \rightarrow \mathbb{R}$ stetig und sei $A \subseteq \mathbb{R}$ eine Obermenge von P . Dann heißt f *stetig fortsetzbar* nach A , falls ein stetiges $g : A \rightarrow \mathbb{R}$ existiert mit $f = g|_P$.

Jedes solche g heißt dann eine *stetige Fortsetzung* von f nach A .

Graphisch lässt sich die Frage nach der Fortsetzbarkeit einer Funktion wie folgt beschreiben. Gegeben ist ein stetiges $f : P \rightarrow \mathbb{R}$. Nun betrachten wir einen umfassenderen Definitionsbereich A . Die Frage lautet:

Können wir den Graphen von f so ergänzen, dass ein stetiger Funktionsgraph auf A entsteht?



$$g : [a, b] \rightarrow \mathbb{R} \text{ stetig, } g(p) = f(p) \text{ für alle } p \in P$$

Wir müssen hierzu im Allgemeinen einfache Definitionslücken

schließen, Brücken zwischen Intervallgrenzen errichten, Funktionen, die auf Pulvermengen wie $P = \mathbb{Q}$ definiert sind, nach $A = \mathbb{R}$ ausdehnen usw. Die Aufgabe ist vielgestaltig und komplex. Wir betrachten im Folgenden lediglich einige wichtige Spezialfälle.

Aus der Schule kennen wir das Phänomen der hebbaren und nichthebbaren Definitionslücken. Diese tauchen vor allem dann auf, wenn wir eine Funktion durch einen Term erklären und dann nicht durch 0 teilen dürfen.

Beispiele

(1) Sei $f : \mathbb{R} - \{0\} \rightarrow \mathbb{R}$ definiert durch $f(x) = 1/x$ für alle $x \neq 0$. Dann besitzt die Funktion keine stetige Fortsetzung nach \mathbb{R} .

(2) Sei $f : \mathbb{R} - \{0\} \rightarrow \mathbb{R}$ definiert durch $f(x) = x \sin(1/x)$ für alle $x \neq 0$. Dann gilt

$$\lim_{x \rightarrow 0, x \neq 0} f(x) = 0, \quad \text{da}$$

$$|f(x)| = |x \sin(1/x)| \leq |x| \cdot 1 \leq |x| \quad \text{für alle } x \neq 0.$$

Die Funktion f besitzt also eine eindeutige stetige Fortsetzung g nach \mathbb{R} , die definiert ist durch

$$g(x) = \begin{cases} f(x) & \text{falls } x \neq 0, \\ 0 & \text{sonst.} \end{cases}$$

Eine weitere wichtige Technik der stetigen Fortsetzung ist die *lineare Interpolation*. Wir illustrieren sie an einem typischen Beispiel.

Beispiel

Sei $f: [0, 1] \cup [8, 9] \rightarrow \mathbb{R}$ eine stetige Funktion. Dann können wir f nach $[0, 9]$ stetig fortsetzen, indem wir die Punkte $(1, f(1))$ und $(8, f(8))$ der Ebene durch ein Geradenstück miteinander verbinden. Die so entstehende stetige Fortsetzung $g: [0, 9] \rightarrow \mathbb{R}$ von f lässt sich wie folgt darstellen:

$$g(x) = \begin{cases} f(x) & \text{falls } x \in [0, 1] \cup [8, 9], \\ f(1) + a(x-1) & \text{falls } x \in]1, 8[, \end{cases}$$

wobei $a = (f(8) - f(1))/(8 - 1)$ die Steigung der Geraden durch $(1, f(1))$ und $(8, f(8))$ ist.

Wir betrachten nun das schwierigere Problem der Fortsetzung einer stetigen Funktion $f: \mathbb{Q} \rightarrow \mathbb{R}$ nach \mathbb{R} . Eine einfache Überlegung zeigt, dass dies nicht für alle Funktionen möglich ist.

Beispiel

Die Funktion $f: \mathbb{Q} \rightarrow \mathbb{R}$ mit $f(x) = 0$ für $x < \sqrt{2}$ und $f(x) = 1$ für $x > \sqrt{2}$ ist stetig. Sie lässt sich aber nicht stetig nach \mathbb{R} fortsetzen, da

$$\lim_{x \uparrow \sqrt{2}, x \in \mathbb{Q}} f(x) = 0 \neq 1 = \lim_{x \downarrow \sqrt{2}, x \in \mathbb{Q}} f(x).$$

Das Beispiel zeigt, dass wir Sprünge vermeiden müssen, wenn die stetige Fortsetzung gelingen soll. Für monotone Funktionen genügt dies:

Satz (*Fortsetzungssatz für monotone Funktionen auf \mathbb{Q}*)

Sei $f: \mathbb{Q} \rightarrow \mathbb{R}$ stetig und monoton steigend. Der Wertebereich von f habe keine Lücken der Form $]f(p), f(q)[$, d. h. für alle $p < q$ in \mathbb{Q} gebe es ein $r \in \mathbb{Q}$ mit $f(p) < f(r) < f(q)$. Dann lässt sich f in eindeutiger Weise stetig nach \mathbb{R} fortsetzen. Ein analoger Satz gilt für monoton fallende Funktionen.

Die eindeutige stetige Fortsetzung $g: \mathbb{R} \rightarrow \mathbb{R}$ von f ist gegeben durch

$$g(x) = \lim_{q \uparrow x, q \in \mathbb{Q}} f(q) = \sup(\{f(q) \mid q \in \mathbb{Q}, q < x\}) \quad \text{für alle } x \in \mathbb{R}.$$

Der Satz lässt sich zur Definition der Exponentiation für reelle Exponenten anwenden (vgl. 2.10). Sei hierzu $a > 0$, und sei $f: \mathbb{Q} \rightarrow \mathbb{R}$ definiert durch

$$f(q) = a^q = \sqrt[n]{a^m} \quad \text{für alle } q = m/n, m \in \mathbb{Z}, n \geq 1.$$

Dann erfüllt f die Voraussetzungen des Satzes (mit „steigend“ für $a \geq 1$ und „fallend“ für $a < 1$). Also existiert eine stetige Fortsetzung g_a von f nach \mathbb{R} . Setzt man nun

$$a^x = g_a(x) \quad \text{für alle } x \in \mathbb{R} \text{ und alle } a > 0,$$

so ist die Exponentiation also für alle positiven Basen und alle reellen Exponenten erklärt. Die üblichen Rechenregeln gelten nach wie vor. Gleichwertig ist die Definition von a^x über die Exponentialfunktion e^x (vgl. 6.5).

5.8 Der Zwischenwertsatz

Satz (Zwischenwertsatz)

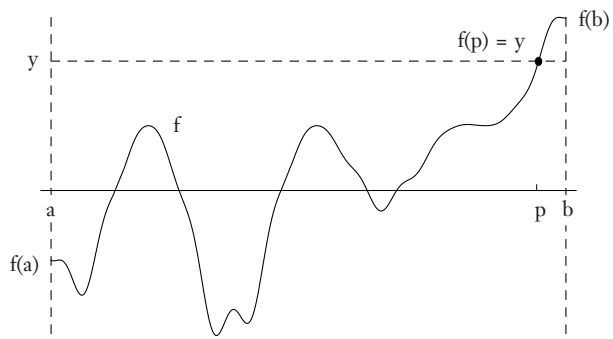
Seien $f : [a, b] \rightarrow \mathbb{R}$ stetig und $y \in \mathbb{R}$. Weiter liege y zwischen $f(a)$ und $f(b)$, d. h., es gelte

$f(a) \leq y \leq f(b)$, falls $f(a) \leq f(b)$, und

$f(b) \leq y \leq f(a)$, falls $f(b) \leq f(a)$.

Dann nimmt f den Wert y an, d. h., es existiert ein $p \in [a, b]$ mit $f(p) = y$.

Der Zwischenwertsatz ist sicher plausibel: Die Funktion beginnt bei $f(a)$ und endet bei $f(b)$. Da sie aufgrund der Stetigkeit nicht springen darf, muss sie auf dem Weg von a nach b jeden Wert zwischen $f(a)$ und $f(b)$ mindestens einmal annehmen. Der Satz muss aber natürlich bewiesen werden,



denn die Anschauung täuscht gerade bei der Stetigkeit oft. De facto ist es die Vollständigkeit der reellen Zahlen, die für die Gültigkeit des Zwischenwertsatzes verantwortlich ist.

Beispiele

- (1) Sei $f : [0, 2] \cap \mathbb{Q} \rightarrow \mathbb{R}$ definiert durch $f(q) = q^2 - 2$. Dann ist f stetig und $f(0) = -2$ und $f(2) = 2$. Aber die Funktion nimmt den Wert 0 zwischen $f(0)$ und $f(2)$ nicht an. Denn aus $f(p) = p^2 - 2 = 0$ würde $p^2 = 2$ folgen, was für keine rationale Zahl gelten kann (vgl. 2.1).
- (2) Sei $g : [0, 2] \cap \mathbb{A} \rightarrow \mathbb{R}$ definiert durch $g(x) = \cos(x)$, wobei \mathbb{A} die Menge der algebraischen Zahlen sei. Dann ist g stetig und $g(0) = 1$ und $g(2) < 0$. Aber die einzige Nullstelle des Kosinus im Intervall $[0, 2]$ ist die transzendente Zahl $\pi/2$ (vgl. 2.2). Also nimmt g den Wert 0 zwischen $g(0)$ und $g(2)$ nicht an.

Alle Beweise des Zwischenwertsatzes benutzen also notwendig die Vollständigkeit von \mathbb{R} , sei es bei der Bildung eines Supremums, der Bildung des Grenzwerts einer Cauchy-Folge, der Anwendung des Prinzips der Intervallschachtelung oder dergleichen mehr. Mit Hilfe von Suprema lässt sich ein p mit $f(p) = y$ im typischen Fall $f(a) < y < f(b)$ zum Beispiel wie folgt finden:

$$p = \sup(\{x \in [a, b] \mid f(x) < y\}).$$

Das Supremum existiert, da die Menge nach Voraussetzung den Punkt a enthält. Mit Hilfe der Stetigkeit von f weist man nun nach, dass in der Tat $f(p) = y$ gilt.

Beispiele

- (1) Sei $f: [0, 1] \cup [2, 3] \rightarrow \mathbb{R}$ die Funktion f mit $f(x) = 0$ für $x \in [0, 1]$ und $f(x) = 1$ für $x \in [2, 3]$. Dann ist f stetig und $f(0) = 0 < 1 = f(3)$, aber die Funktion nimmt den Wert $1/2$ zwischen 0 und 1 nicht an. Die Voraussetzung, dass f auf einem Intervall definiert ist, ist also wesentlich für die Gültigkeit des Zwischenwertsatzes.

- (2) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$f(x) = x^5 + 24x^4 - 7x^3 + x - 10 \quad \text{für alle } x \in \mathbb{R}.$$

Dann gilt $\lim_{x \rightarrow \infty} f(x) = +\infty$ und $\lim_{x \rightarrow -\infty} f(x) = -\infty$. Also gibt es ein $a < 0$ mit $f(a) < 0$ und ein $b > 0$ mit $f(b) > 0$. Da f auf $[a, b]$ stetig ist, nimmt f den Wert 0 zwischen $f(a)$ und $f(b)$ an. Allgemein zeigt die Überlegung, dass jedes reelle Polynom ungeraden Grades mindestens eine Nullstelle besitzt.

- (3) Sei $g: [0, +\infty[\rightarrow \mathbb{R}$ definiert durch

$$g(x) = x^2 \quad \text{für alle } x \geq 0.$$

Sei $y \geq 0$ beliebig. Dann gibt es ein b mit $b^2 \geq y$, etwa $b = \max(1, y)$. Die Funktion g ist stetig auf $[0, b]$ und y liegt zwischen $g(0)$ und $g(b)$. Also gibt es ein $p \in [0, b]$ mit $g(p) = p^2 = y$, d. h., es gilt

$$p = \sqrt{y}.$$

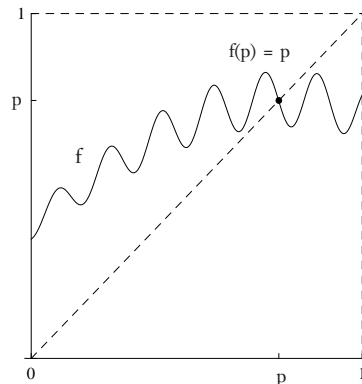
Allgemein zeigt die Überlegung, dass jedes $y \geq 0$ für jedes $n \geq 2$ eine n -te Wurzel besitzt. Der Zwischenwertsatz liefert also einen einfachen Beweis für die Existenz von Wurzeln (vgl. 2.10).

- (4) Sei $f: [0, 1] \rightarrow [0, 1]$ stetig. Dann besitzt f einen Fixpunkt, d. h., es gibt ein p mit $f(p) = p$. Sei hierzu $g: [0, 1] \rightarrow \mathbb{R}$ definiert durch

$$g(x) = f(x) - x \quad \text{für alle } x \in [0, 1].$$

Dann ist g stetig und 0 liegt zwischen $g(0) = f(0) \geq 0$ und $g(1) = f(1) - 1 \leq 0$. Also gibt es ein p mit $g(p) = 0$. Dann ist

$$f(p) = g(p) + p = p.$$



Das anschauliche Zusammentreffen von f mit der Diagonalen kann also auf den Zwischenwertsatz zurückgeführt werden. Analog gibt es für alle stetigen $f: [0, 1] \rightarrow [0, 1]$ ein p mit $f(p) = p^2$. Man betrachte hierzu g auf $[0, 1]$ mit $g(x) = f(x) - x^2$.

- (5) Sind $f, g: [a, b] \rightarrow \mathbb{R}$ stetig mit $f(a) < g(a)$, $g(b) < f(b)$, so gibt es ein p mit $f(p) = g(p)$. Man betrachte hierzu h auf $[a, b]$ mit $h(x) = f(x) - g(x)$.

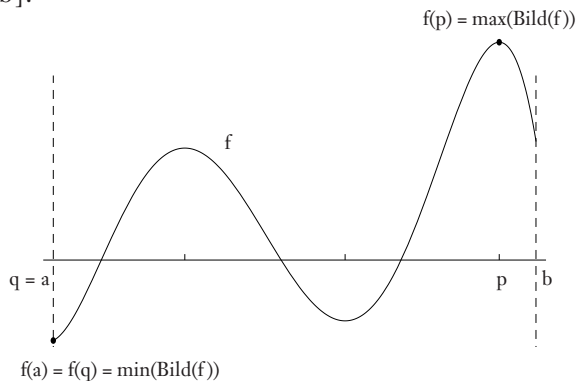
5.9 Der Extremwertsatz von Weierstraß

Satz (*Extremwertsatz, Annahme von Maximum und Minimum*)

Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig. Dann ist f beschränkt und es gibt $p, q \in [a, b]$ mit:

- (a) $f(p)$ ist das Maximum des Wertebereichs von f , d. h., es gilt
 $f(x) \leq f(p)$ für alle $x \in [a, b]$,
- (b) $f(q)$ ist das Minimum des Wertebereichs von f , d. h., es gilt
 $f(q) \leq f(x)$ für alle $x \in [a, b]$.

Der Extremwertsatz ist vielleicht ähnlich einleuchtend wie der Zwischenwertsatz. Eine stetige Funktion muss auf dem Weg von $f(a)$ nach $f(b)$ irgendwann einen maximalen und irgendwann einen minimalen Wert erreichen und annehmen, das kennen wir von jeder Bergwanderung. Auch hier gilt wieder, dass ein Beweis unerlässlich ist. Anschauungen ersetzen keine Beweise, und zudem basiert die Anschauung sehr stark auf einem „zeichenbaren Funktionsgraphen“, was den Stetigkeitsbegriff nicht voll einfängt.



Beweisskizze

Diesmal ist es der Satz von Bolzano-Weierstraß, der zum Beweis herangezogen wird, also erneut ein relativ starkes und abstraktes Geschütz. Man startet mit einer Folge $(f(x_n))_{n \in \mathbb{N}}$ im Wertebereich von f , die gegen das Supremum des Wertebereichs konvergiert, falls dieser nach oben beschränkt ist, und gegen $+\infty$ im anderen Fall. (Letzteres kann nicht passieren, aber das weiß man an dieser Stelle noch nicht). Nun wendet man den Satz von Bolzano-Weierstraß auf die Folge $(x_n)_{n \in \mathbb{N}}$ im Definitionsbereich an. Dies liefert einen Häufungspunkt p der Folge, und man zeigt nun mit Hilfe der Stetigkeit von f im Punkt p , dass die Funktion f im Punkt p wie gewünscht ihr Maximum annimmt. Eine analoge Argumentation oder ein Übergang zu $-f$ zeigt die Annahme des Minimums.

Eine stetige Funktion auf einem Intervall $[a, b]$ kann ihr Maximum und ihr Minimum mehrfach annehmen, man betrachte etwa den Kosinus auf dem Intervall $[0, 6\pi]$. Eine konstante Funktion nimmt sogar in jedem Punkt ihr Minimum und ihr Maximum an. Umgekehrt gilt: Ist das Minimum einer Funktion gleich ihrem Maximum, so ist die Funktion konstant.

Der Extremwertsatz ist für stetige Funktionen, die auf offenen oder halboffenen Intervallen definiert sind, im Allgemeinen nicht mehr gültig:

Beispiele

- (1) Die Funktion $f:]0, 1] \rightarrow \mathbb{R}$ mit $f(x) = 1/x$ nimmt ihr Minimum 1 im Punkt 1 an, aber ihr Wertebereich $[1, +\infty[$ ist nach oben unbeschränkt und hat kein Maximum.
- (2) Die Funktion $g:]0, 1[\rightarrow]0, 1[$ mit $f(x) = x$ hat den beschränkten Wertebereich $]0, 1[$, der kein Minimum und kein Maximum besitzt. Das Supremum des Wertebereichs ist 1, aber der Wert 1 wird nicht angenommen.

Der Zwischenwertsatz und der Extremwertsatz lassen sich sehr ansprechend zu einem einzigen Satz zusammenfassen:

Satz (*Wertebereich stetiger Funktionen*)

I Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig. Dann gibt es $c \leq d$ in \mathbb{R} mit $\text{Bild}(f) = [c, d]$.

Der Zwischenwertsatz sorgt dafür, dass das Bild von f ein Intervall ist, und der Extremwertsatz garantiert, dass die Randpunkte des Bildes angenommen werden und also das Bildintervall abgeschlossen ist.

Beschränkte abgeschlossene Intervalle nannten wir auch *kompakt* (vgl. 2.9). Damit kann man den Satz sehr griffig formulieren:

Stetige Funktionen bilden kompakte Intervalle auf kompakte Intervalle ab.

Allgemein gilt, dass stetige Funktionen Intervalle auf Intervalle abbilden. Das stetige Bild eines offenen Intervalls kann nun aber offen, abgeschlossen oder halboffen sein, wie die folgenden Beispiele zeigen.

Beispiele

- (1) Die Funktion $f:]0, 1[\rightarrow \mathbb{R}$ mit $f(x) = x$ hat das Bild $]0, 1[$.
- (2) Die Funktion $g:]0, 1[\rightarrow \mathbb{R}$ mit $g(x) = 1$ hat das Bild $\{1\} = [1, 1]$.
- (3) Die Funktion $h:]0, 1[\rightarrow \mathbb{R}$ mit $h(x) = |x - 1/2|$ hat das Bild $[0, 1/2[$.

Den kompakten Intervallen der Form $[a, b]$ kommt in der Analysis eine besondere Bedeutung zu. Beispiele sind:

<i>Prinzip der Intervallschachtelung</i>	Jede Intervallfolge $[a, b] \supseteq [a_1, b_1] \supseteq \dots$ besitzt einen nichtleeren Schnitt.
<i>Satz von Bolzano-Weierstraß</i>	Jede Folge in $[a, b]$ besitzt einen Häufungspunkt in $[a, b]$.
<i>Satz über die gleichmäßige Stetigkeit</i>	Jede stetige Funktion auf $[a, b]$ ist gleichmäßig stetig.
<i>Satz über den Wertebereich</i>	Jede stetige Funktion auf $[a, b]$ besitzt ein Intervall $[c, d]$ als Bild.

5.10 Die Stetigkeit der Umkehrfunktion

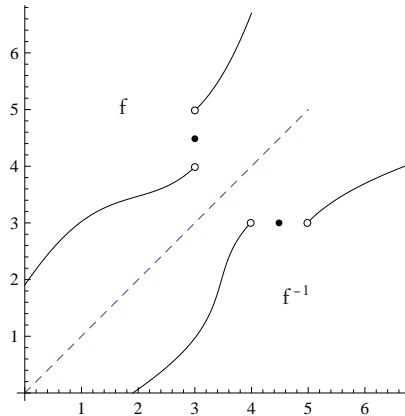
Satz (Stetigkeit der Umkehrfunktion)

Sei I ein Intervall, und sei $f: I \rightarrow \mathbb{R}$ streng monoton steigend. Dann ist die Umkehrfunktion f^{-1} von f streng monoton steigend und stetig. Eine analoge Aussage gilt für streng monoton fallende Funktionen.

Graphisch entsteht die Umkehrfunktion durch Spiegelung an der Hauptdiagonalen und es ist plausibel, dass durch diesen Spiegelungsprozess keine Sprünge entstehen. Der Satz besagt, dass dies tatsächlich auch so ist.

Bittet man die Hörer einer Anfängervorlesung, den Satz

„Die Umkehrfunktion einer streng monoton steigenden Funktion ist streng monoton ...“



zu ergänzen, so erhält man manchmal sogar eine Mehrheit für „streng monoton fallend“. Durch die Spiegelung an der Hauptdiagonalen scheint sich die Monotonie umzukehren. Dies ist aber nicht der Fall. Liegen $x_1 < x_2$ im Definitionsbereich von f , so gilt bei strenger Monotonie, dass

$$y_1 = f(x_1) < f(x_2) = y_2.$$

Ist nun g die Umkehrfunktion von f , so gilt also

$$g(y_1) = x_1 < x_2 = g(y_2).$$

Damit ist also auch g wieder streng monoton steigend. Analog ist die Umkehrfunktion einer streng monoton fallenden Funktion wieder streng monoton fallend. Graphisch kann man dies auch wie folgt einsehen: Gegeben sei der Graph von f in der üblichen Weise auf einem Blatt Papier. Nun drehen wir das Blatt Papier um 90 Grad gegen den Uhrzeigersinn. Dann haben wir bereits die Umkehrfunktion vor Augen, wobei wir im Gegensatz zur üblichen Darstellung die x -Achse von rechts ($-\infty$) nach links ($+\infty$) lesen müssen, die y -Achse wie üblich von unten ($-\infty$) nach oben ($+\infty$). Die Monotonie bleibt gleich.

Die Umkehrfunktion einer Funktion f existiert genau dann, wenn f injektiv ist. Ist f streng monoton, so gilt klarerweise $f(x_1) \neq f(x_2)$ für $x_1 \neq x_2$, d. h., f ist injektiv und somit existiert f^{-1} . Ist f eine stetige Funktion auf einem Intervall, so folgt aus der Injektivität von f umgekehrt die strenge Monotonie. Denn eine Konstellation $x_1 < x_2 < x_3$ und $f(x_1) < f(x_2)$, $f(x_2) > f(x_3)$ würde aufgrund des Zwischenwertsatzes nach sich ziehen, dass ein Wert zwischen $\max(f(x_1), f(x_3))$ und $f(x_2)$ mindestens zweimal angenommen wird. Analog ist $x_1 < x_2 < x_3$ und $f(x_1) > f(x_2)$, $f(x_1) < f(x_3)$ nicht mit der Injektivität vereinbar.

Der Satz erspart uns viele Stetigkeitsbeweise:

Beispiele

- (1) Die n -te Potenz $f: [0, +\infty[\rightarrow \mathbb{R}$ auf $[0, +\infty[$ mit $f(x) = x^n$ für alle $x \geq 0$, ist streng monoton steigend. Also ist ihre Umkehrfunktion, die n -te Wurzelfunktion $g: [0, +\infty[\rightarrow \mathbb{R}$ mit $g(x) = \sqrt[n]{x}$ für alle $x \geq 0$, streng monoton steigend und stetig.
- (2) Die Exponentialfunktion $\exp: \mathbb{R} \rightarrow \mathbb{R}$ ist streng monoton steigend. Also ist ihre Umkehrfunktion, der natürliche Logarithmus $\log:]0, +\infty[\rightarrow \mathbb{R}$, streng monoton steigend und stetig.

Dass der Definitionsbereich von f ein (beschränktes oder unbeschränktes) Intervall ist, ist notwendig für die Gültigkeit des Satzes. Eine Lücke im Definitionsbereich kann zu einem Sprung der Umkehrfunktion führen:

Beispiel

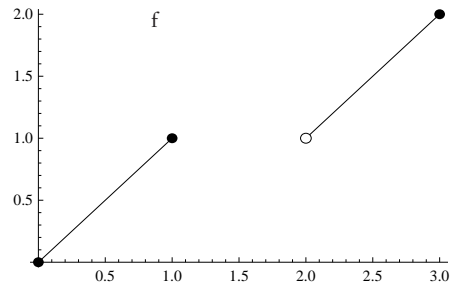
Sei $f: [0, 1] \cup]2, 3] \rightarrow \mathbb{R}$ definiert durch

$$f(x) = x \text{ für } x \in [0, 1],$$

$$f(x) = x - 1 \text{ für } x \in]2, 3].$$

Dann ist f streng monoton steigend (und stetig). Das Bild von f ist das Intervall $[0, 2]$, und für die Umkehrfunktion $g: [0, 2] \rightarrow \mathbb{R}$ gilt

$$\lim_{x \uparrow 1} g(x) = 1 \neq 2 = \lim_{x \downarrow 1} g(x).$$



Die Umkehrfunktion ist also im Punkt 1 unstetig.

Dem Leser ist vielleicht aufgefallen, dass im Satz nicht vorausgesetzt wird, dass die Funktion $f: I \rightarrow \mathbb{R}$ stetig ist. Das ist in der Tat kein Versäumnis. Für die Stetigkeit der Umkehrfunktion genügt die strenge Monotonie und die Intervallnatur des Definitionsbereichs. Die Funktion kann beliebig viele Sprünge machen. Sie führen zu Lücken im Definitionsbereich der Umkehrfunktion, die für die Stetigkeit irrelevant sind.

Beispiel

Sei $g: [0, 2] \rightarrow \mathbb{R}$ die Funktion des vorherigen Beispiels, d. h., es gilt

$$g(x) = \begin{cases} x & \text{falls } x \in [0, 1] \\ x + 1 & \text{falls } x \in]1, 2]. \end{cases}$$

Dann ist g streng monoton, aber nicht stetig. Die Umkehrfunktion von g ist die stetige Ausgangsfunktion $f: [0, 1] \cup]2, 3] \rightarrow \mathbb{R}$ des vorherigen Beispiels.

Ist eine auf einem Intervall I definierte streng monotone Funktion $f: I \rightarrow \mathbb{R}$ zusätzlich stetig, so ist der Definitionsbereich der Umkehrfunktion f^{-1} wieder ein Intervall.

5.11 Punktweise und gleichmäßige Konvergenz

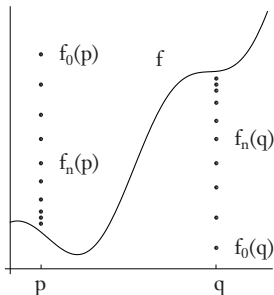
Definition (punktweise Konvergenz einer Funktionenfolge)

Sei $P \subseteq \mathbb{R}$, und seien $f_n : P \rightarrow \mathbb{R}$ für alle n . Es existiere

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) \quad \text{für alle } x \in P. \quad (\text{punktweise Konvergenzbedingung})$$

Dann heißt $f : P \rightarrow \mathbb{R}$ der *Grenzwert* der Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ und die Folge $(f_n)_{n \in \mathbb{N}}$ heißt (punktweise) *konvergent* gegen f . Wir schreiben dann

$$f = \lim_{n \rightarrow \infty} f_n \quad (\text{punktweise}).$$



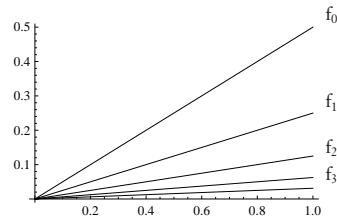
Wir haben unseren Grenzwertbegriff also noch einmal erweitert. Nun ist erklärt, was die Konvergenz von Funktionen $f_0, f_1, \dots, f_n, \dots$, gegen eine Funktion f bedeutet. Die Funktionen f_n haben alle denselben Definitionsbereich P , und wir verlangen, dass für jeden einzelnen Punkt p in P die Werte $f_0(p), f_1(p), \dots, f_n(p), \dots$ gegen $f(p)$ konvergieren.

Beispiele

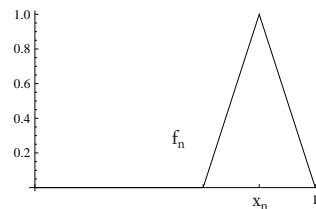
- (1) Seien $f_n : [0, 1] \rightarrow \mathbb{R}$ definiert durch

$$f_n(x) = (1/2^n) x \quad \text{für alle } x \text{ und } n.$$

Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ gegen die Nullfunktion auf $[0, 1]$.



- (2) Für alle n sei $f_n : [0, 1] \rightarrow \mathbb{R}$ eine Zackenfunktion wie im Diagramm rechts, mit Wert 1 an der Stelle x_n und Breite $2(1 - x_n)$ des Zackens. Es gelte $\lim_{n \rightarrow \infty} x_n = 1$. Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ gegen die Nullfunktion auf $[0, 1]$.

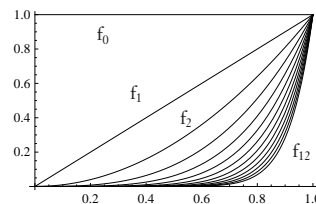


- (3) Seien $f_n : [0, 1] \rightarrow \mathbb{R}$ definiert durch

$$f_n(x) = x^n \quad \text{für alle } x \text{ und } n.$$

Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ gegen die Funktion $f : [0, 1] \rightarrow \mathbb{R}$ mit

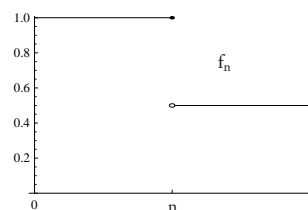
$$f(x) = 0 \quad \text{für } x \in [0, 1[\quad \text{und} \quad f(1) = 1.$$



- (4) Seien $f_n : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$f_n(x) = 1 \quad \text{für } |x| \leq n, \quad f_n(x) = 1/2 \quad \text{für } |x| > n.$$

Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ monoton steigend gegen die 1-Funktion auf \mathbb{R} .



(5) Für die Funktionen $f_n : [0, 1] \rightarrow \mathbb{R}$ aus Beispiel (3) gilt

$$\lim_n \lim_{x \uparrow 1} f_n(x) = \lim_n 1 = 1, \text{ aber } \lim_{x \uparrow 1} \lim_n f_n(x) = \lim_{x \uparrow 1} 0 = 0.$$

Das dritte Beispiel zeigt, dass die Stetigkeit beim Grenzübergang verloren gehen kann. Alle Funktionen f_n sind stetig und ihr Grenzwert f existiert, aber f hat eine Sprungstelle bei 1. Die Funktionen f_n nehmen alle den Wert $1/2$ an, der von der Grenzfunktion weit entfernt liegt. Dies motiviert die folgende Verstärkung der punktweisen Konvergenz:

Definition (*gleichmäßige Konvergenz einer Funktionenfolge*)

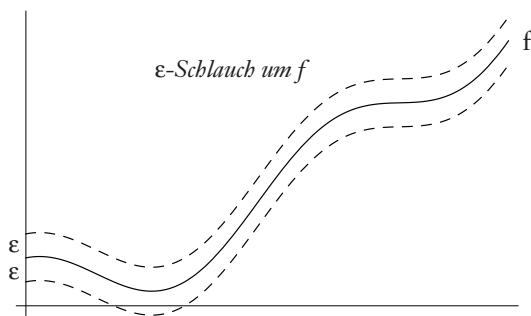
Sei $P \subseteq \mathbb{R}$, und seien $f_n : P \rightarrow \mathbb{R}$ für alle n und $f : P \rightarrow \mathbb{R}$. Es gelte

$$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 \forall x \in P \quad |f(x) - f_n(x)| < \varepsilon. \quad (\text{gleichmäßige Konvergenzbedingung})$$

Dann sagen wir, dass $(f_n)_{n \in \mathbb{N}}$ *gleichmäßig* gegen f konvergiert, und schreiben

$$f = \lim_{n \rightarrow \infty} f_n \quad (\text{gleichmäßig}).$$

Die gleichmäßige Konvergenz besagt anschaulich, dass sich alle Funktionen f_n schließlich in einem beliebig schmalen ε -Schlauch um f befinden. Der Allquantor über x steht hinter dem Existenzquantor für n_0 (diese Vertauschung kennen wir schon von der gleichmäßigen Stetigkeit aus 5.5). Rücken wir den Allquantor nach vorne, so erhalten



wir die schwächere punktweise Konvergenzbedingung. Die gleichmäßige Konvergenz impliziert also die punktweise Konvergenz. Dass die Umkehrung im Allgemeinen nicht gilt, zeigen die obigen Beispiele. Nur die Funktionenfolge des ersten Beispiels ist gleichmäßig konvergent, die Funktionenfolgen in (2), (3) und (4) konvergieren punktweise, aber nicht gleichmäßig. Dennoch taucht die gleichmäßige Konvergenz in der Analysis recht häufig auf. Ein wichtiges Beispiel liefern die Potenzreihen, die wir in der nächsten Sektion besprechen werden.

Der große Vorteil der gleichmäßigen Konvergenz ist der Erhalt der Stetigkeit:

Satz (*Stetigkeitssatz*)

Es gelte $f = \lim_n f_n$ (gleichmäßig) für Funktionen auf $P \subseteq \mathbb{R}$. Alle f_n seien stetig. Dann ist auch f stetig.

Das Konzept der gleichmäßigen Konvergenz unterstützt die Sicht, dass die stetigen Funktionen doch wieder nicht so wild sind, wie es manchmal scheint. Ein großer Satz von Weierstraß besagt, dass für jede stetige Funktion f auf einem Intervall $[a, b]$ eine Folge $(f_n)_{n \in \mathbb{N}}$ von Polynomen existiert, die gleichmäßig gegen f konvergiert. So zerknittert eine stetige Funktion f auf $[a, b]$ also auch sein mag: In jedem noch so schmalen ε -Schlauch um f findet sich ein „harmloses“ Polynom.

5.12 Potenzreihen

Definition (Potenzreihe, Konvergenzbereich)

Für $a_n, p, x \in \mathbb{R}$ heißt die Reihe $\sum_n a_n (x - p)^n$ die *Potenzreihe* mit *Koeffizienten* a_n und *Entwicklungspunkt* p im Punkt x . Weiter setzen wir

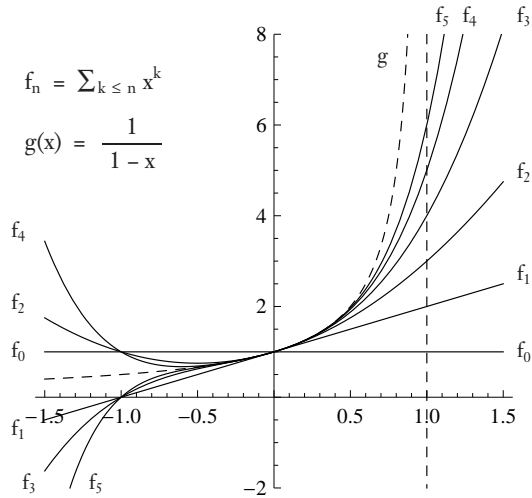
$$K = \{x \in \mathbb{R} \mid \text{die Reihe } \sum_n a_n (x - p)^n \text{ konvergiert}\}.$$

Die Menge K heißt der *Konvergenzbereich* der Potenzreihe $\sum_n a_n (x - p)^n$.

Um uns den neuen Begriffen zu nähern, nehmen wir zunächst $p = 0$ an. Dann hat eine Potenzreihe die Form $\sum_n a_n x^n$. Derartige mit einer Variablen x präsentierte Reihen sind uns schon begegnet:

Beispiele

- (1) Die geometrische Reihe $\sum_n x^n$ ist eine Potenzreihe mit den Koeffizienten $a_n = 1$ für alle n und dem Konvergenzbereich $K =]-1, 1[$. Für $x \in K$ gilt $\sum_n x^n = 1/(1 - x)$ (vgl. 4.3).
- (2) Die Exponentialreihe $\sum_n x^n/n!$ ist eine Potenzreihe mit den Koeffizienten $a_n = 1/n!$ für alle n . Es gilt $K = \mathbb{R}$ und $\sum_n x^n/n! = e^x$ für alle x (vgl. 4.12).



Jede Potenzreihe $\sum_n a_n x^n$ definiert eine Funktion auf ihrem Konvergenzbereich K , die oft ebenfalls mit $\sum_n a_n x^n$ bezeichnet wird. Damit steht „ $\sum_n a_n x^n$ “ für jedes $x \in \mathbb{R}$ für eine Reihe, für jedes $x \in K$ für eine reelle Zahl (die Summe der Reihe) und zudem jetzt auch noch für eine Funktion auf K .

Lassen wir beliebige Entwicklungspunkte p zu, so unterscheiden sich die Funktionen $\sum_n a_n (x - p)^n$ von $\sum_n a_n x^n$ nur durch eine Translation um p . Bei der Untersuchung des Konvergenzverhaltens von Potenzreihen können wir also ohne Verlust $p = 0$ annehmen.

Die große Frage, die sich auch wieder durch den Blick auf die Exponentialfunktion $\exp: \mathbb{R} \rightarrow \mathbb{R}$ mit $\exp(x) = \sum_n x^n/n!$ motivieren lässt, lautet:

Welche Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ lassen sich als Potenzreihe darstellen, d. h., wann gibt es Koeffizienten a_n mit $f(x) = \sum_n a_n x^n$ für alle x ? Wie berechnen wir die Koeffizienten?

In der reellen Differentialrechnung kann man diese Fragen ganz gut beantworten, in der komplexen Analysis (der Funktionentheorie) perfekt; in \mathbb{C} lässt sich jede differenzierbare Funktion als Potenzreihe darstellen, in \mathbb{R} nicht. Die Koeffizienten gewinnt man durch Differenzieren. Wir verweisen den Leser hierzu auf 7.12.

Die Konvergenzbereiche der geometrischen Reihe und der Exponentialreihe sind typisch. Der Konvergenzbereich K einer Potenzreihe $\sum_n a_n x^n$ ist immer ein symmetrisches Intervall um den Nullpunkt, also von der Form

$$]-R, R[,]-R, R], [-R, R[, [-R, R],$$

wobei der sog. *Konvergenzradius* R auch gleich $+\infty$ sein kann. Dieser lässt sich durch

$$R = \frac{1}{\limsup_n \sqrt[n]{|a_n|}} \quad (\text{Formel von Cauchy-Hadamard})$$

berechnen, wobei wir für diese Formel „ $1/0 = +\infty$ “ vereinbaren.

Alle vier Intervalltypen können als Konvergenzbereich auftreten. Die geometrische Reihe liefert ein Beispiel für den offenen Fall, die drei anderen Fälle werden durch die folgenden Beispiele belegt.

Beispiele

- (1) $\sum_{n \geq 1} x^n/n$ hat den Konvergenzbereich $[-1, 1[$. Der Punkt $x = 1$ entspricht der harmonischen Reihe und $x = -1$ der alternierenden harmonischen Reihe.
- (2) $\sum_{n \geq 1} (-1)^n/n x^n$ hat den Konvergenzbereich $] -1, 1]$. Nun entspricht $x = 1$ der alternierenden harmonischen Reihe und $x = -1$ der harmonischen Reihe. Die definierte Funktion ist die Spiegelung der Funktion in (1) an der y -Achse, da $\sum_{n \geq 1} (-1)^n/n x^n = \sum_{n \geq 1} (-x)^n/n$.
- (3) $\sum_{n \geq 1} x^n/n^2$ hat den Konvergenzbereich $[-1, 1]$.

Die durch die Potenzreihe in Beispiel (1) definierte Funktion lässt sich mit Methoden der Differential- und Integralrechnung bestimmen. Es gilt

$$\sum_{n \geq 1} x^n/n = -\log(1-x) \quad \text{für alle } x \in [-1, 1[, \text{ d.h.}$$

$$\sum_{n \geq 1} (-1)^{n-1}/n x^n = \log(1+x) \quad \text{für alle } x \in]-1, 1] \quad (\text{Logarithmus-Reihe})$$

mit dem Logarithmus zur Basis e . Das Einsetzen von $x = 1$ liefert

$$\log(2) = \sum_{n \geq 1} (-1)^{n-1}/n \cdot 1^n = 1 - 1/2 + 1/3 - 1/4 + \dots$$

Die Bestimmung der durch eine Potenzreihe dargestellten Funktion ist im Allgemeinen sehr schwierig. Man weiß zumindest, dass sie stetig ist:

Satz (Abelscher Grenzwertsatz)

I $\sum_n a_n x^n$ ist stetig auf ihrem Konvergenzbereich K .

Man kann durch Majorisierung mit einer geometrischen Reihe zeigen, dass $\sum_n a_n x^n$ für alle $r < R$ gleichmäßig auf $] -r, r[$ konvergiert, d.h., für alle $\varepsilon > 0$ existiert ein n_0 , so dass $|\sum_{n > n_0} a_n x^n| < \varepsilon$ für alle x mit $|x| < r$. Dies liefert die Stetigkeit der Potenzreihe auf allen Intervallen $] -r, r[, r < R$, und damit auf $] -R, R[$. Der Knackpunkt des Abelschen Satzes sind die Randpunkte $-R$ und R von K , falls diese zu K gehören.

6. Kapitel

Elementare Funktionen

6.1 Polynome

Definition (reelle und komplexe Polynome)

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion der Form

$$f(x) = a_k x^k + a_{k-1} x^{k-1} + \dots + a_1 x + a_0 \quad \text{für alle } x \in \mathbb{R},$$

mit $a_0, \dots, a_k \in \mathbb{R}$. Dann heißt f eine *reelle Polynomfunktion* oder kurz ein *reelles Polynom* mit den *Koeffizienten* a_0, \dots, a_k .

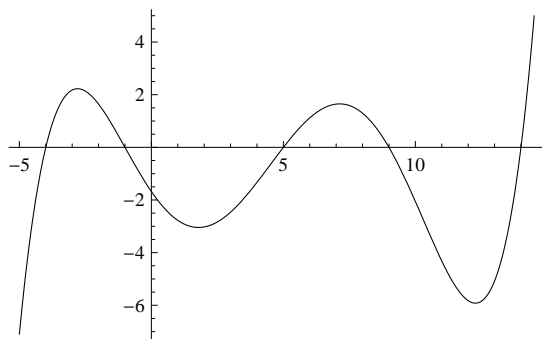
- (a) Ist $a_k \neq 0$, so heißt k der *Grad* und a_k der *Leitkoeffizient* von f .
Gilt $a_k = 1$, so heißt das Polynom *normiert*.
- (b) Gilt $f(x) = 0$ für alle x (d. h., sind alle Koeffizienten gleich 0), so heißt f das *Nullpolynom*. Wir ordnen ihm den symbolischen Grad $-\infty$ zu.

Analog sind komplexe Polynome $f: \mathbb{C} \rightarrow \mathbb{C}$ und die zugehörigen Sprechweisen definiert. Die Koeffizienten sind dann komplexe Zahlen.

Polynome sind die Grundfunktionen der Analysis. Sie sind relativ einfach beherrschbar und eignen sich, um andere Funktionen durch Approximation beherrschbar zu machen. Potenzreihen, deren Partialsummen Polynome sind, hatten wir schon kennengelernt. Im Satz von Taylor werden wir diesen Approximationsgedanken weiterverfolgen (vgl. 7.12).

Der Graph des Nullpolynoms ist die x -Achse. Der Graph eines

Polynoms vom Grad 0 ist eine zur x -Achse parallele, aber von der x -Achse verschiedene Gerade. Dem Grad 1 entsprechen die Geraden mit einer von null verschiedenen Steigung. Die Graphen der Polynome zweiten Grades sind Parabeln und ab dem Grad 3 zeigen sich die typischen polynomiellen Auf- und Abbewegungen.



$$f(x) = \frac{(x+4)(x+1)(x-5)(x-9)(x-14)}{1500}$$

Beispiele

- (1) „ $f(x) = 2x^3 - x^2 + x$ für alle $x \in \mathbb{R}$ “ definiert ein reelles Polynom zweiten Grades. Um die Sprechweise zu vereinfachen, identifiziert man in der Analysis Terme manchmal mit den durch sie definierten Funktionen (mit maximalem Definitionsbereich). Dann ist $2x^3 - x^2 + x$ ein Polynom.
- (2) $(x-4)(2x+5)(x-1)x$ ist ein reelles Polynom vierten Grades.
- (3) $z^3 + 2iz^2 - z + i$ ist ein komplexes Polynom dritten Grades.
- (4) $|x|$ und $1/x$ sind keine Polynome.

Die Summe und das Produkt zweier Polynome ist wieder ein Polynom. Die Division zweier Polynome führt dagegen zu den rationalen Funktionen, die wir in der nächsten Sektion besprechen.

Die algebraische Struktur der Polynome erlaubt es, für Polynome wie für ganze Zahlen eine Division mit Rest durchzuführen. Es gilt:

Satz (*Polynomdivision*)

Seien f, g (reelle oder komplexe) Polynome mit positivem Grad. Dann gibt es eindeutige Polynome q und r mit

$$f = qg + r \text{ und } \text{Grad}(r) < \text{Grad}(g).$$

Im Satz kann r das Nullpolynom sein. In diesem Fall heißt g ein *Teilerpolynom* oder kurz ein *Teiler* von f .

Beispiel

Sei $f: \mathbb{C} \rightarrow \mathbb{C}$ ein komplexes Polynom mit $\text{Grad}(f) \geq 1$, und sei $w \in \mathbb{C}$. Die Division von f durch $(z - w)$ liefert Polynome q und r mit

$$f = q(z - w) + r \text{ mit } \text{Grad}(r) < 1.$$

Speziell ist $f(w) = q(w)(w - w) + r(w) = r(w)$. Wegen $\text{Grad}(r) < 1$ ist r eine konstante Funktion. Ist also w eine Nullstelle von f , so ist r das Nullpolynom und es gilt

$$f = q(z - w). \quad (\text{Abspalten von Nullstellen})$$

Der Fundamentalsatz der Algebra liefert durch eine wiederholte Abspaltung von Nullstellen für jedes komplexe Polynom f eine Zerlegung in Linearfaktoren:

$$f = (z - w_1)(z - w_2) \dots (z - w_k), \quad k = \text{Grad}(f).$$

Die Nullstellen w_i können dabei mehrfach auftreten.

Die Macht des Imaginären über das Reelle ist groß und schenkt uns ein Zerlegungsergebnis für reelle Polynome. Ist $g: \mathbb{R} \rightarrow \mathbb{R}$ ein reelles Polynom vom Grad k , so sei $f: \mathbb{C} \rightarrow \mathbb{C}$ das komplexe Polynom, das entsteht, wenn wir in g die reelle Variable x durch die komplexe Variable z ersetzen. Ist w eine reelle Nullstelle von f , so ist w auch eine Nullstelle von g . Andernfalls ist \bar{w} eine weitere Nullstelle von f . Dann ist

$$(x - w) \cdot (x - \bar{w}) = x^2 - 2 \operatorname{Re}(w)x + w\bar{w}$$

ein reelles Polynom zweiten Grades. Diese Überlegung zeigt, dass ein reelles Polynom als Produkt von reellen Linearfaktoren und reellen Polynomen zweiten Grades geschrieben werden kann, die Nullstellen in \mathbb{C} , aber nicht in \mathbb{R} besitzen.

Aus der Zerlegung in Linearfaktoren folgt, dass ein Polynom k -ten Grades höchstens k Nullstellen besitzen kann. Eine hübsche Anwendung ist der *Identitätssatz* für Polynome: Stimmen Polynome f und g k -ten Grades auf $(k + 1)$ -vielen Punkten überein, so gilt $f = g$. Denn $f - g$ hat einen Grad kleinergleich k und $(k + 1)$ -viele Nullstellen. Also gilt $f - g = 0$ und damit $f = g$.

6.2 Rationale Funktionen

Definition (rationale Funktion)

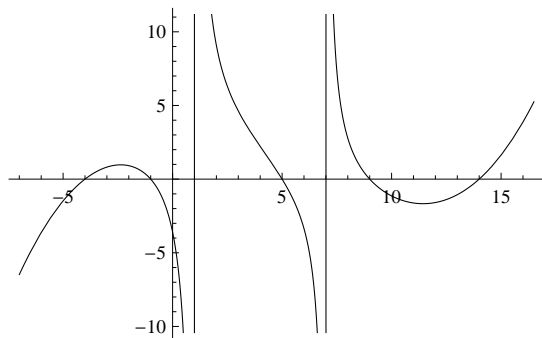
Sei $h : P \rightarrow \mathbb{R}$ eine Funktion der Form

$$h(x) = \frac{f(x)}{g(x)} \quad \text{für alle } x \in P, \quad P = \{x \in \mathbb{R} \mid g(x) \neq 0\},$$

mit Polynomen $f, g : \mathbb{R} \rightarrow \mathbb{R}$, g nicht das Nullpolynom. Dann heißt f eine *reelle rationale Funktion*.

Analog sind komplexe rationale Funktionen $h : P \rightarrow \mathbb{C}$ definiert. Die definierenden Polynome f und g sind hier komplexe Polynome.

Die rationalen Funktionen entstehen also durch Division zweier Polynome. Dabei werden die endlich vielen Nullstellen des Nennerpolynoms zu Definitionslücken. Hat der Nenner g den Grad k , so ist $h = f/g$ höchstens in k Punkten nicht definiert. In \mathbb{R} kann f/g überall definiert sein, in \mathbb{C} existiert immer mindestens eine Definitionslücke, wenn der Grad des Nenners positiv ist.



$$f(x) = \frac{(x+4)(x+1)(x-5)(x-9)(x-14)}{100(x-2)(x-7)}$$

Beispiele

- (1) Jedes Polynom f ist eine rationale Funktion, da $f = f/1$.
- (2) $1/z$ ist eine rationale Funktion mit Definitionsbereich $\mathbb{C} - \{0\}$.
- (3) $1/(x^2 + 1)$ ist eine rationale Funktion auf \mathbb{R} ,
 $1/(z^2 + 1)$ ist eine rationale Funktion auf $\mathbb{C} - \{i, -i\}$.
- (4) $(x^2 - 1)/(x - 1)$ ist eine rationale Funktion auf $\mathbb{R} - \{1\}$.
- (5) $x + 1$ ist eine rationale Funktion auf \mathbb{R} . Für alle $x \neq 1$ stimmt diese Funktion mit der Funktion in (4) überein, da $(x^2 - 1) = (x + 1)(x - 1)$.

Die beiden letzten Beispiele zeigen, dass der Definitionsbereich von f/g oft kleiner ist als nötig. Manche Definitionslücken sind stetig hebbar und es ist nur natürlich, die rationalen Funktionen zu bevorzugen, deren Definitionslücken echt, also nicht stetig hebbar sind (vgl. 5.7). Eine nicht stetig hebbare Definitionslücke heißt eine *Polstelle* von f/g . Da wir gemeinsame Nullstellen abspalten und wegekürzen können, können wir von f/g zu einer rationalen Funktion f^*/g^* übergehen, deren Polstellen genau die Nullstellen der Funktion g^* sind. Wir sagen dann, dass die rationale Funktion f^*/g^* einen *vollständigen Definitionsbereich* besitzt.

Allgemeiner heißt eine rationale Funktion f/g *gekürzt*, falls die Polynome f und g teilerfremd sind, d. h., es gibt keine Polynome q, f_1, g_1 mit $\text{Grad}(q) \geq 1$ und

$$f = q f_1, \quad g = q g_1.$$

Eine gekürzte rationale Funktion hat einen vollständigen Definitionsbereich, da man gemeinsame Linearfaktoren wegkürzen kann. In \mathbb{C} ist „gekürzt“ gleichwertig zu „vollständiger Definitionsbereich“, in \mathbb{R} ist $(x^2 + 1)/(x^2 + 1)$ nicht gekürzt, hat aber einen vollständigen Definitionsbereich. Eine gekürzte Darstellung f/g ist nur bis auf einen Faktor $c \neq 0$ eindeutig, da $f/g = (cf)/(cg)$ für alle $c \neq 0$ gilt. Als Funktionen sind f/g und $(cf)/(cg)$ identisch.

Ist f/g eine rationale Funktion mit vollständigem Definitionsbereich, so lässt sich ein n -facher Pol w der Funktion multiplikativ abspalten:

$$\frac{f}{g} = \frac{1}{(z-w)^n} \cdot \frac{f}{g^*}, \quad g^*(w) \neq 0.$$

Auch eine additive Abspaltung eines n -fachen Pols der Form

$$\frac{f}{g} = \frac{b_n}{(z-w)^n} + \dots + \frac{b_1}{z-w} + \frac{f^*}{g^*}, \quad g^*(w) \neq 0,$$

ist möglich und als *Partialbruchzerlegung* von f/g bekannt.

Beispiele

(1) Es gilt $(z(z+1))^{-1} = 1/z - 1/(z+1)$, also ist, wie wir schon wissen,

$$\sum_{n \geq 1} (n(n+1))^{-1} = \lim_n \sum_{1 \leq k \leq n} (1/k - 1/(k+1)) = \lim_n (1 - 1/(n+1)) = 1.$$

(2) Wegen $(z(z+1)(z+2))^{-1} = (1/2)/z - 1/(z+1) + (1/2)/(z+2)$ ist

$$\begin{aligned} \sum_{n \geq 1} (n(n+1)(n+2))^{-1} &= \lim_n \sum_{1 \leq k \leq n} ((1/2)/k - 1/(k+1) + (1/2)/(k+2)) = \\ &= \lim_n (1/4 - (1/2)/(n+1) + (1/2)/(n+2)) = 1/4. \end{aligned}$$

Die rationalen Funktionen besitzen eine bemerkenswerte algebraische Struktur:

Exkurs: Die rationalen Funktionen als Körper

Sei $K = \{ f/g \mid f/g \text{ ist eine gekürzte rationale Funktion auf } \mathbb{R} \}$. Wir nennen $f/g \in K$ *positiv*, falls die Leitkoeffizienten von f und g dasselbe Vorzeichen besitzen. So sind

$$x = x/1, \quad (x-3) = (x-3)/1, \quad (-x^2+1)/(-x) = (x^2-1)/x$$

positiv. Wir definieren nun für Elemente aus K :

$$f_1/g_1 + f_2/g_2 = \text{„die gekürzte Funktion } (f_1 g_2 + f_2 g_1)/(g_1 g_2)\text{“},$$

$$f_1/g_1 \cdot f_2/g_2 = \text{„die gekürzte Funktion } (f_1 f_2)/(g_1 g_2)\text{“},$$

$$f_1/g_1 < f_2/g_2, \quad \text{falls „die gekürzte Funktion } f_2/g_2 - f_1/g_1 \text{ ist positiv“}.$$

K ist ein nichtarchimedisch angeordneter Körper. Die Menge $\{ n1 \mid n \in \mathbb{N} \}$ ist nach oben beschränkt durch die Identität $x = x/1 \in K$.

6.3 Die reelle Exponentialfunktion

Definition (Exponentialfunktion)

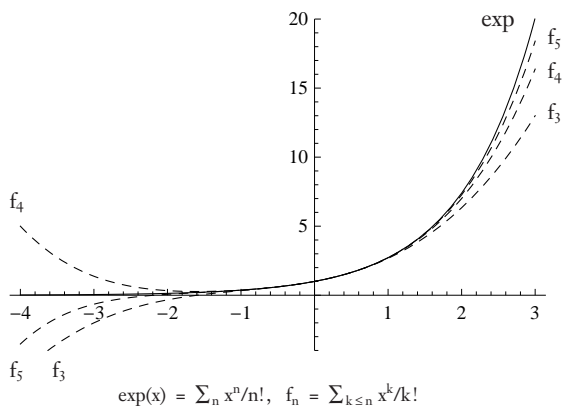
Wir definieren $\exp : \mathbb{R} \rightarrow \mathbb{R}$ durch

$$\exp(x) = \sum_n x^n/n! \quad \text{für alle } x \in \mathbb{R}. \quad (\text{Reihendefinition von } \exp)$$

Die Funktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$ heißt die (reelle) Exponentialfunktion.

Die Definition lässt sich durch gliedweises Differenzieren motivieren. Sie zielt darauf ab, dass sich die Exponentialfunktion beim Ableiten selbst reproduziert. Die Summanden $x^n/n!$ haben die Ableitung $n x^{(n-1)}/n! = x^{(n-1)}/(n-1)!$. Leiten wir also die Exponentialfunktion gliedweise ab, so erhalten wir

$$\begin{aligned} \exp'(x) &= \sum_{n \geq 1} x^{(n-1)}/(n-1)! \\ &= \sum_n x^n/n! = \exp(x). \end{aligned}$$



Dieses Vorgehen ist tatsächlich erlaubt (vgl. 7.6).

Wichtige Eigenschaften der Exponentialfunktion sind:

Stetigkeit	\exp ist stetig auf \mathbb{R}
Additionstheorem	$\exp(x+y) = \exp(x) \cdot \exp(y)$ für alle x, y
Monotonie	\exp ist streng monoton steigend
Wertebereich	$\text{Bild}(\exp) =]0, +\infty[$
Wachstum	$\lim_{x \rightarrow \infty} \exp(x)/x^k = +\infty$ für alle $k \geq 1$
Steigung im Nullpunkt	$\lim_{x \rightarrow 0, x \neq 0} (e^x - 1)/x = 1$
Restgliedabschätzung	$ \sum_{n \geq n_0} x^n/n! \leq 2 x ^{n_0}/n_0!, \text{ falls } n_0 \geq 2 x - 1$

Das Additionstheorem wird auch *Funktionalgleichung* genannt. Die Wachstumsaussage besagt, dass die Exponentialfunktion schneller wächst als jede Potenz x^k . Ein Merkspruch zur technischen Restgliedabschätzung lautet:

Der Fehler ist begrenzt durch das Doppelte des ersten nichtberechneten Gliedes.

Die Bedingung $n_0 \geq 2|x| - 1$ ist unterdrückt, was nur für $x \in [-1, 1]$ harmlos ist.

Elementare Werte der Exponentialfunktion sind $\exp(0) = 1$ und $\exp(1) = e$.

Warum gilt das Additionstheorem?

In Reihenform besagt das Additionstheorem, dass für alle x, y gilt:

$$\sum_n x^n/n! \cdot \sum_n y^n/n! = \sum_n (x+y)^n/n!$$

Dies zu beweisen sieht auf den ersten Blick hoffnungslos aus, aber es wird alles ganz einfach, wenn wir die linke Seite als Cauchy-Produkt berechnen, was wir aufgrund der absoluten Konvergenz der Reihen dürfen (vgl. 4.11):

$$\sum_n \frac{x^n}{n!} \cdot \sum_n \frac{y^n}{n!} = \sum_n \sum_{k \leq n} \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!} = \sum_n \frac{(x+y)^n}{n!},$$

wobei wir im letzten Schritt die binomische Formel verwendet haben:

$$(x+y)^n = \sum_{k \leq n} \frac{n!}{k!(n-k)!} x^k y^{n-k} \quad \text{für alle } n.$$

Warum ist das Additionstheorem wichtig?

Das Additionstheorem bringt die „exponentielle Natur“ der Exponentialfunktion ans Licht. Schreiben wir nämlich

$$e^x \text{ anstelle von } \exp(x), \quad (\textit{Exponentialschreibweise für exp})$$

so gelten alle üblichen Eigenschaften der Exponentiation und die bereits definierte Exponentiation e^q , $q \in \mathbb{Q}$, wird fortgesetzt. Beispielsweise gilt:

$$\begin{aligned} (1) \quad e^x \cdot e^y &= \exp(x) \cdot \exp(y) = \exp(x+y) = e^{x+y}, \\ (2) \quad (e^x)^3 &= \exp(x) \cdot \exp(x) \cdot \exp(x) = \exp(x+x+x) = \exp(3x) = e^{3x}, \\ (3) \quad e^x &= \exp(x) = \exp(x/2 + x/2) = \exp(x/2) \cdot \exp(x/2) = e^{x/2} \cdot e^{x/2}, \\ &\text{also } e^{x/2} = \sqrt{e^x} = (e^x)^{1/2}, \text{ usw.} \end{aligned}$$

Es gibt eine äquivalente Definition der Exponentialfunktion, die sich durch die Modellierung einer kontinuierlichen Verzinsung motivieren lässt. Erhalten wir 7 % Zinsen auf eine Geldeinheit, so haben wir, wenn wir die Zinsen monatlich ausgezahlt bekommen und gleich wieder mit 7 % anlegen, am Ende des Jahres $(1 + 0,07/12)^{12}$ Einheiten. Setzen wir 365 statt 12 ein, so entspricht $(1 + 0,07/365)^{365}$ einer täglichen Verzinsung. Ein Grenzübergang modelliert eine kontinuierliche Verzinsung:

$$\exp(x) = \lim_{n \geq 1} (1 + x/n)^n \quad \text{für alle } x \in \mathbb{R}, \quad (\textit{Limesdefinition von exp})$$

$$e = \exp(1) = \lim_{n \geq 1} (1 + 1/n)^n. \quad (\textit{Limesdefinition von e})$$

Zur Illustration skizzieren wir noch einen Beweis des Additionstheorems mit Hilfe der Limesdefinition. Für alle $x, y \in \mathbb{R}$ gilt:

$$\exp(x) \cdot \exp(y) = \lim_{n \geq 1} (1 + x/n)^n \cdot (1 + y/n)^n = \lim_{n \geq 1} (1 + (x+y + xy/n)/n)^n = \exp(x+y),$$

wobei wir im letzten Schritt verwenden, dass

$$\exp(x) = \lim_{n \geq 1} (1 + x_n/n)^n \quad \text{für alle Folgen } (x_n)_{n \geq 1} \text{ mit } \lim_{n \geq 1} x_n = x.$$

6.4 Der natürliche Logarithmus

Definition (natürlicher Logarithmus)

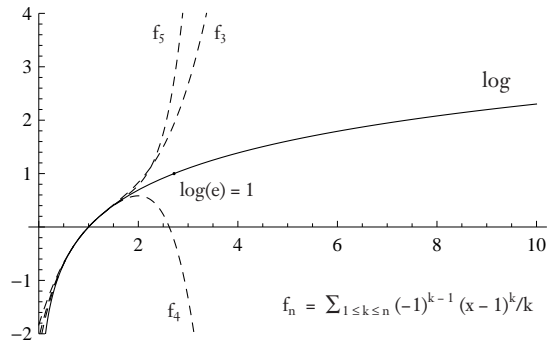
Die Umkehrfunktion der reellen Exponentialfunktion heißt der *natürliche Logarithmus*. In Zeichen schreiben wir

$$\log :]0, \mathbb{R}[\rightarrow \mathbb{R} \text{ oder } \ln :]0, \mathbb{R}[\rightarrow \mathbb{R}.$$

Die Bezeichnung „log“ ist heute weit verbreitet. Die traditionelle und oft auch in der Schule verwendete Bezeichnung „ln“ steht für „logarithmus naturalis“.

Die Logarithmusfunktion ist unser erstes Beispiel für den Einsatz der Exponentialfunktion als Generator für andere Funktionen. Die Umkehrfunktion existiert, da $\exp : \mathbb{R} \rightarrow \mathbb{R}$ monoton steigend und damit injektiv ist.

Aus $\exp(0) = 1$ und $\exp(1) = e$ folgt $\log(1) = 0$ und $\log(e) = 1$. Weitere Eigenschaften, die sich aus der Definition ergeben, sind:



Stetigkeit	\log ist stetig auf $]0, +\infty[$
Multiplikationstheorem	$\log(xy) = \log(x) + \log(y)$ für alle $x, y > 0$
Monotonie	\log ist streng monoton steigend
Vorzeichen	$\log(x) < 0$ für $x < 1$, $\log(1) = 0$, $\log(x) > 0$ für $x > 1$
Wachstum	$\lim_{x \rightarrow \infty} \log(x)/x^k = 0$ für alle $k \geq 1$
Steigung im Punkt 1	$\lim_{x \rightarrow 0, x \neq 0} \log(x+1)/x = 1$

Aus dem Multiplikationstheorem lassen sich weitere Regeln gewinnen:

Beispiele

- (1) $\log(x^2) = \log(xx) = \log(x) + \log(x) = 2\log(x)$,
- (2) $\log(1/x) = -\log(x)$, da $0 = \log(1) = \log(x \cdot 1/x) = \log(x) + \log(1/x)$,
- (3) $\log(\sqrt{x}) = \log(x)/2$, da $\log(x) = \log((\sqrt{x})^2) = 2\log(\sqrt{x})$ für alle $x > 0$.

Allgemeiner gilt $\log(x^y) = y\log(x)$ für alle $x > 0$ und $y \in \mathbb{Q}$, und allgemeiner auch für alle $y \in \mathbb{R}$, falls die Exponentiation x^y für alle $y \in \mathbb{R}$ bereits definiert ist (vgl. 2.10, 5.7).

Die Definition als Umkehrfunktion der Exponentialfunktion und die graphische Methode liefern uns ein gutes Bild über das Verhalten des Logarithmus:

Beispiele

- (1) Da \exp auf $[0, +\infty[$ schneller wächst als jede k -te Potenz, wächst der Logarithmus (Umkehrfunktion von \exp) langsamer als jede k -te Wurzel (Umkehrfunktion von x^k auf $[0, +\infty[$). Das ist gerade die Wachstumsaussage der Tabelle.
- (2) Die Exponentialfunktion fällt sehr schnell von 1 nach 0 auf dem Weg von 0 nach $-\infty$, da $\exp(-x) = 1/\exp(x)$ für $x \geq 0$. Also fällt der Logarithmus relativ langsam auf dem Weg von 1 nach 0 gegen $-\infty$ ab. De facto gilt

$$\lim_{x \downarrow 0} (x \log(x)) = \lim_{x \rightarrow \infty} \frac{\log(1/x)}{x} = - \lim_{x \rightarrow \infty} \frac{\log(x)}{x} = 0,$$

d. h., der Abfall des Logarithmus in $]0, 1[$ gegen $-\infty$ kann sich nicht einmal gegen einen linearen Abfall gegen 0 durchsetzen.

Eine Reihendarstellung steht uns zunächst nicht zur Verfügung (weshalb wir auch keine Restgliedabschätzung angeben können). Mit Methoden der Differentialrechnung lässt sich zeigen:

Reihendarstellung des Logarithmus

$$\log(x+1) = \sum_{n \geq 1} (-1)^{n-1} \frac{x^n}{n} \quad \text{für alle } x \text{ mit } -1 < x \leq 1, \text{ d. h.}$$

$$\log(x) = \sum_{n \geq 1} (-1)^{n-1} \frac{(x-1)^n}{n} \quad \text{für alle } x \text{ mit } 0 < x \leq 2. \quad (\text{Logarithmus-Reihe})$$

Die Auswertung der Logarithmus-Reihe am rechten Randpunkt liefert den Grenzwert der alternierenden harmonischen Reihe:

$$\log(2) = \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} = 1 - 1/2 + 1/3 - 1/4 + \dots$$

Die Logarithmus-Reihe für $\log(x)$ divergiert für alle $x > 2$. Dies zeigt, dass eine Potenzreihe auf einem beschränkten Intervall mit einer stetigen Funktion auf $]0, +\infty[$ übereinstimmen, aber außerhalb des Intervalls divergieren kann. Verantwortlich für dieses Phänomen ist nicht der Logarithmus, sondern die Intervallnatur des Konvergenzbereichs einer Potenzreihe. Auch eine Potenzreihendarstellung von $1/x$ im Punkt 1 zeigt dieses Verhalten.

Das Multiplikationstheorem liefert die Möglichkeit, mit der Logarithmus-Reihe $\log(x)$ für beliebige $x > 0$ zu berechnen. Denn ist n so groß gewählt, dass $x/2^n < 2$, so ist

$$\log(x) = \log(2^n \cdot x/2^n) = \log(2^n) + \log(x/2^n) = n \log(2) + \log(x/2^n),$$

und $\log(2)$ und $\log(x/2^n)$ lassen sich mit der Reihendarstellung berechnen.

6.5 Die allgemeine Exponentialfunktion

Definition (*Exponentialfunktion zu einer positiven Basis*)

Sei $a > 0$. Dann definieren wir:

$$\exp_a(x) = \exp(x \log(a)) \quad \text{für alle } x \in \mathbb{R}.$$

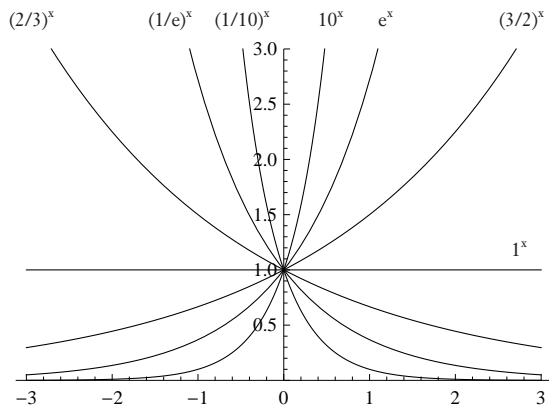
Die Funktion $\exp_a : \mathbb{R} \rightarrow \mathbb{R}$ heißt die *Exponentialfunktion zur Basis a*.

Wir schreiben auch a^x anstelle von $\exp_a(x)$.

Es gilt also:

*a^x verhält sich wie e^x ,
wobei der Exponent x
durch $\log(a)$ skaliert wird.*

Für $a > 1$ sieht der Graph von a^x aus wie der Graph von e^x . Für a zwischen 0 und 1 sieht er aus wie der Graph von $e^{-x} = (1/e)^x$, da dann $\log(a) < 0$. Schließlich ist 1^x die konstante Funktion 1, die die beiden Welten trennt.



Beispiele

$$(1) \quad x^x = e^{x \log(x)}, \quad x^{(x^2)} = e^{x^2 \log(x)}, \quad x^{(2^x)} = e^{2^x \log(x)} \quad \text{für alle } x > 0.$$

$$(2) \quad (a^2)^x = e^{x \log(a^2)} = e^{2x \log(a)} = a^{2x},$$

$$(a^x)^2 = e^{2 \log(a^x)} = e^{2x \log(a)} = a^{2x} \quad \text{für alle } a > 0 \text{ und alle } x \in \mathbb{R}.$$

$$(3) \quad \sqrt{2}^{\sqrt{2}} = e^{\sqrt{2} \log(\sqrt{2})} = e^{\sqrt{2} \log(2^{1/2})} = e^{\sqrt{2}/2 \log(2)}.$$

Die allgemeine Exponentiation erlaubt oft sehr elegante Berechnungen von Grenzwerten:

Beispiel

Dass $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$, lässt sich mit der binomischen Formel einsehen:

$$n = ((\sqrt[n]{n} - 1) + 1)^n \geq \frac{n!}{2! (n-2)!} (\sqrt[n]{n} - 1)^2 1^{n-2} = \frac{n(n-1)(\sqrt[n]{n} - 1)^2}{2},$$

also

$$\sqrt[n]{n} \leq \sqrt{2/(n-1)} + 1 \quad \text{für } n \geq 2.$$

Aber mit der allgemeinen Exponentiation geht es leichter:

$$\lim_n \sqrt[n]{n} = \lim_n n^{1/n} = \lim_n e^{1/n \log(n)} = e^0 = 1.$$

Rechenregeln für die Exponentiation

Für alle $a, b > 0$ und alle $x, y \in \mathbb{R}$ gilt:

- (a) $a^x \cdot a^y = e^{x \log(a)} \cdot e^{y \log(a)} = e^{(x+y) \log(a)} = a^{x+y},$
- (b) $(a^x)^y = (e^{x \log(a)})^y = e^{y \log(e^{x \log(a)})} = e^{y \cdot x \log(a)} = a^{x \cdot y},$
- (c) $a^x b^x = e^{x \log(a)} \cdot e^{x \log(b)} = e^{x (\log(a) + \log(b))} = e^{x \log(ab)} = (ab)^x.$

Die erste Regel zeigt, dass die Funktionen \exp_a das Additionstheorem erfüllen, d.h., es gilt $\exp_a(x+y) = \exp_a(x) \cdot \exp_a(y)$ für alle $x, y \in \mathbb{R}$. Im Reich der stetigen Funktionen gibt es keine weiteren nichttrivialen Beispiele:

Satz (*Charakterisierung der Funktionen mit Additionstheorem*)

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ eine stetig mit $f(x+y) = f(x) \cdot f(y)$ für alle $x, y \in \mathbb{R}$. Dann gilt:

- (a) Ist $f(1) = 0$, so ist f die Nullfunktion auf \mathbb{R} .
- (b) Ist $f(1) = a > 0$, so ist $f(x) = a^x$ für alle $x \in \mathbb{R}$.

Aus dem Satz folgt insbesondere, dass die zwei Wege der Einführung der allgemeinen Exponentiation äquivalent sind (vgl. 2.10 und 5.7). Denn beide Wege erzeugen stetige Exponentiationen auf \mathbb{R} zur Basis a , die das Additionstheorem erfüllen, und bei beiden Wegen gilt $a^1 = a$.

Im Charakterisierungssatz erscheint die Eulersche Zahl als Basis gleichberechtigt neben allen anderen Basen. Erst die Ableitung

$$(a^x)' = (\exp(x \log(a)))' = \log(a) \exp(x \log(a)) = \log(a) a^x$$

zeigt, dass die Basis e ausgezeichnet ist: Die Ableitung von $\exp = \exp_e$ ist wegen $\log(e) = 1$ einfacher als die von \exp_a für alle $a \neq e$.

Bislang hatten wir eine feste positive Basis a und einen variablen Exponenten x betrachtet. Wir können umgekehrt auch den Exponenten festhalten und die Basis laufen lassen:

Allgemeine Potenzfunktionen

Für $b \in \mathbb{R}$ sei $\text{pot}_b:]0, +\infty[\rightarrow \mathbb{R}$ definiert durch

$$\text{pot}_b(x) = x^b = e^{b \log(x)} \quad \text{für alle } x > 0.$$

Zu den Potenzfunktionen gehören viele vertraute Funktionen auf $]0, +\infty[$:

$$1, x, x^2, x^3, \dots, 1/x, 1/x^2, 1/x^3, \dots, \sqrt{x}, \sqrt[3]{x}, \sqrt[4]{x}, \dots$$

Es gilt $\text{pot}_b(x) \cdot \text{pot}_b(y) = \text{pot}_b(xy)$. Man kann zeigen, dass jede von der Nullfunktion verschiedene stetige Funktion f auf $]0, +\infty[$ mit $f(x) \cdot f(y) = f(xy)$ eine allgemeine Potenzfunktion ist. Genauer gilt dann $f = \text{pot}_b$ mit $b = \log(f(e))$.

Für $b > 0$ gilt $\lim_{x \downarrow 0} \text{pot}_b(x) = 0$, sodass pot_b für $b > 0$ durch $0 = 0^b$ stetig im Nullpunkt fortgesetzt werden kann.

6.6
Der allgemeine Logarithmus

Definition
(Logarithmus zu einer positiven Basis $a \neq 1$)

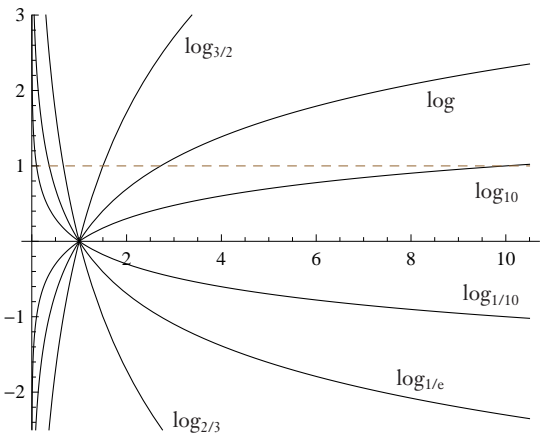
Sei $a > 0, a \neq 1$. Dann heit die Umkehrfunktion der Exponentialfunktion \exp_a der *Logarithmus zur Basis a* . In Zeichen schreiben wir

$\log_a :]0, +\infty[\rightarrow \mathbb{R}.$

Die Funktion \exp_a ist fr $a > 1$ streng monoton steigend und fr $a \in]0, 1[$ streng monoton fallend, sodass $\log_a = (\exp_a)^{-1}$ existiert. Die konstante Funktion \exp_1 lsst sich nicht umkehren, „ \log_1 “ ist nicht definiert.

Die Eigenschaften der Funktionen $\log_a :]0, +\infty[\rightarrow \mathbb{R}$ fr $a > 1$ sind denen des natrlichen Logarithmus hnlich, fr Basen a zwischen 0 und 1 ndern sich allerdings Monotonie und Vorzeichen.

Fr alle $a > 0$ gilt:



Stetigkeit	\log_a ist stetig auf $]0, +\infty[$
Multiplikationstheorem	$\log_a(xy) = \log_a(x) + \log_a(y)$ fr alle $x, y > 0$
Steigung im Punkt 1	$\lim_{x \rightarrow 0, x \neq 0} \log_a(x+1)/x = 1/\log(a)$
Wachstum	$\lim_{x \rightarrow +\infty} \log_a(x)/^k\sqrt{x} = 0$ fr alle $k \geq 1$

Fr $a > 1$ gilt:

Monotonie	\log_a ist streng monoton steigend
Vorzeichen	$\log_a(x) < 0$ fr $x < 1$, $\log_a(1) = 0$, $\log_a(x) > 0$ fr $x > 1$

Fr a mit $0 < a < 1$ gilt dagegen:

Monotonie	\log_a ist streng monoton fallend
Vorzeichen	$\log_a(x) > 0$ fr $x < 1$, $\log_a(1) = 0$, $\log_a(x) < 0$ fr $x > 1$

Die Grundregel für den Umgang mit den Logarithmusfunktionen lautet

$$y = a^x \text{ ist äquivalent zu } \log_a(y) = x, \text{ d. h.,} \\ \log_a \text{ holt den Exponenten aus } a^x.$$

Beispiele

- (1) $\log_a(a) = \log_a(a^1) = 1$, $\log_a(1) = \log_a(a^0) = 0$, $\log_a(1/a) = \log_a(a^{-1}) = -1$
nach der Grundregel. Letzteres folgt auch aus dem Multiplikationstheorem:
 $0 = \log_a(1) = \log_a(a \cdot 1/a) = \log_a(a) + \log_a(1/a) = 1 + \log_a(1/a).$
- (2) $\log_2(1) = 0$, $\log_2(2) = 1$, $\log_2(4) = 2$, $\log_2(1/16) = -4$,
 $\log_{10}(1) = 0$, $\log_{10}(10) = 1$, $\log_{10}(100) = 2$, $\log_{10}(1/1000) = -3.$

Die allgemeinen Logarithmusfunktionen besitzen, als Funktionenmenge betrachtet, eine Reihe von bemerkenswerten Eigenschaften. Im Folgenden ist x positiv und die Basen a und b sind positiv und von 1 verschieden.

Rechenregeln

- (1) $\log_a(x) = \frac{\log(x)}{\log(a)}$, da $x = e^{\log(x)} = e^{\log(x)/\log(a) \cdot \log(a)} = a^{\log(x)/\log(a)}.$
- (2) $\log_a(x) = \frac{\log(x)}{\log(a)} = \frac{\log_b(x) \log(b)}{\log_b(a) \log(b)} = \frac{\log_b(x)}{\log_b(a)}.$
- (3) $\log_a(x) = \frac{\log_{1/a}(x)}{\log_{1/a}(a)} = \frac{\log_{1/a}(x)}{-1} = -\log_{1/a}(x).$

Dies kann man auch aus $a^x = (1/a)^{-x}$ und der Grundregel gewinnen, denn \log_a holt aus $y = a^x$ das x und $\log_{1/a}$ holt aus $y = (1/a)^{-x}$ das $-x$.

- (4) $\log_a(b) = \frac{\log_b(b)}{\log_b(a)} = \frac{1}{\log_b(a)}.$

Es ist bemerkenswert, wie einfach sich die allgemeinen Logarithmen umrechnen und auf den natürlichen Logarithmus zurückführen lassen.

Beispiele

- (1) $\log_{10}(x) = \log(x)/\log(10).$
- (2) $\log_{10000}(x) = \log_{10}(x)/\log_{10}(10000) = \log_{10}(x)/4$, allgemein
 $\log_{a^c}(x) = \log_a(x)/\log_a(a^c) = \log_a(x)/c$ für alle $c \neq 0.$
- (3) $\log(2) = 1/\log_2(e)$, $\log_2(10) = 1/\log_{10}(2).$

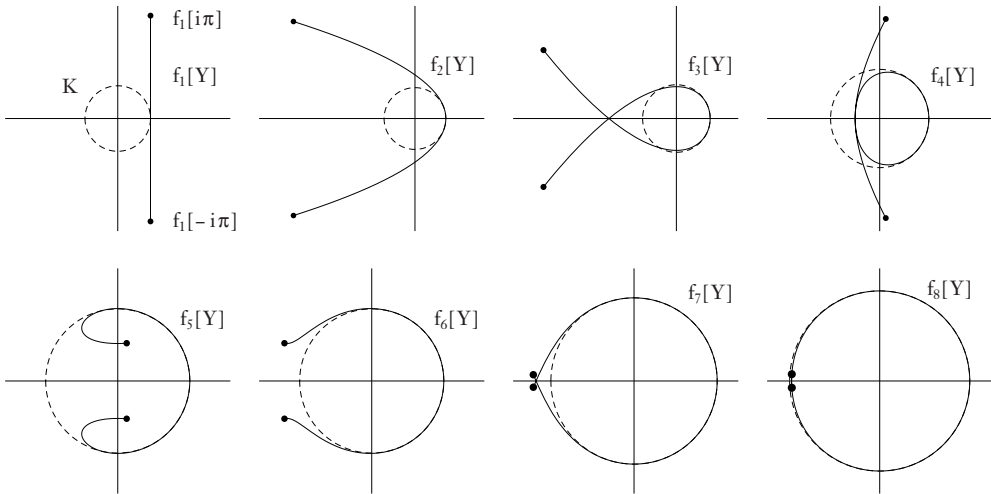
6.7 Die komplexe Exponentialfunktion

Definition (komplexe Exponentialfunktion)

Wir definieren die (komplexe) Exponentialfunktion $\exp : \mathbb{C} \rightarrow \mathbb{C}$ durch

$$\exp(z) = \sum_n z^n/n! \quad \text{für alle } z \in \mathbb{C}. \quad (\text{Reihendefinition von } \exp)$$

Wir schreiben auch e^z anstelle von $\exp(z)$.



Die Diagramme zeigen $K = \{z \in \mathbb{C} \mid |z| = 1\}$ und die Bildmengen $f_n[Y]$ für

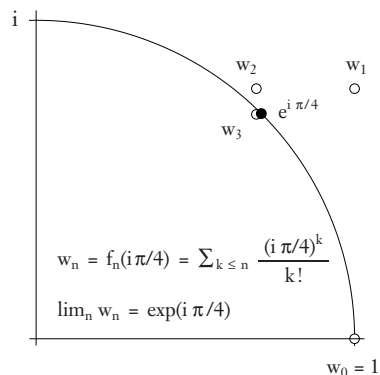
$$Y = \{iy \mid y \in [-\pi, \pi]\}, \quad n = 1, \dots, 8, \quad \text{wobei } f_n : \mathbb{C} \rightarrow \mathbb{C}, \quad f_n(z) = \sum_{1 \leq k \leq n} z^k/k!$$

Die komplexe Exponentialfunktion ist wie die reelle Exponentialfunktion definiert, und viele Resultate übertragen sich von \mathbb{R} nach \mathbb{C} : Die Exponentialfunktion ist stetig auf \mathbb{C} und das Cauchy-Produkt liefert das Additionstheorem:

$$e^{z+w} = e^z \cdot e^w \quad \text{für alle } z, w \in \mathbb{C}.$$

Speziell gilt für alle $x, y \in \mathbb{R}$:

$$e^{x+iy} = e^x e^{iy} \quad \text{für alle } x, y \in \mathbb{R}.$$



Um die Abbildungseigenschaften der komplexen Exponentialfunktion zu verstehen, genügt es also, die Werte e^x und e^{iy} mit $x, y \in \mathbb{R}$ zu untersuchen: Kennen wir die komplexe Exponentialfunktion auf der x - und der y -Achse, so kennen wir sie auf ganz \mathbb{C} . Das Verhalten auf der x -Achse ist uns gut bekannt, die komplexe Exponentialfunktion setzt ja die reelle Exponentialfunktion fort. Neuartig ist

$$e^{iy} = \sum_n i^n y^n/n! = 1 + iy - y^2/2 - iy^3/6 \pm \dots$$

In $\sum_n y^n/n!$ ist hier die zyklische Folge $1, i, -1, -i, 1, i, -1, -i, \dots$ eingewebt. Alle Summanden von e^{iy} befinden sich also auf dem Achsenkreuz der Ebene. Einen genaueren Einblick erlaubt uns aber erst die komplexe Konjugation. Für alle $z \in \mathbb{C}$ gilt nämlich

$$\overline{\exp(z)} = \overline{\sum_n z^n/n!} = \sum_n \overline{z^n/n!} = \sum_n \bar{z}^n/n! = \exp(\bar{z}),$$

und damit gilt für alle $y \in \mathbb{R}$:

$$e^{iy} \cdot \overline{e^{iy}} = e^{iy} \cdot e^{\overline{iy}} = e^{iy} \cdot e^{-iy} = e^0 = 1.$$

Wegen $z\bar{z} = |z|^2$ ist also $|e^{iy}| = 1$ für alle $y \in \mathbb{R}$. Die komplexe Exponentialfunktion bildet damit die y -Achse der Ebene stetig in den Einheitskreis $K = \{z \in \mathbb{C} \mid |z| = 1\}$ ab. Die erste Idee, die y -Achse stetig in K abzubilden, ist, sie aufzuwickeln. Man kann in der Tat zeigen, dass die komplexe Exponentialfunktion sich in dieser Weise verhält. Das genaue (wundervolle!) Ergebnis der Untersuchung lautet anschaulich formuliert:

Die komplexe Exponentialfunktion wickelt die y -Achse längentreu auf den Einheitskreis K auf. Es gilt $e^0 = (1, 0) \in K$. Entlang der positiven y -Achse verläuft die Aufwicklung gegen den Uhrzeigersinn, entlang der negativen y -Achse verläuft sie im Uhrzeigersinn.

Die y -Achse wird bei der Kreisaufwicklung nicht gedehnt und nicht gestaucht. Das Geradenstück $[0, 2\pi[\times \{i\}$ wird, beginnend im Punkt $(0, 1)$, gegen den Uhrzeigersinn zu K umgeformt. Ebenso wird $Y =]-\pi, \pi] \times \{i\}$ zu K umgeformt usw. Speziell erhalten wir:

Werte der Kreisaufwicklung

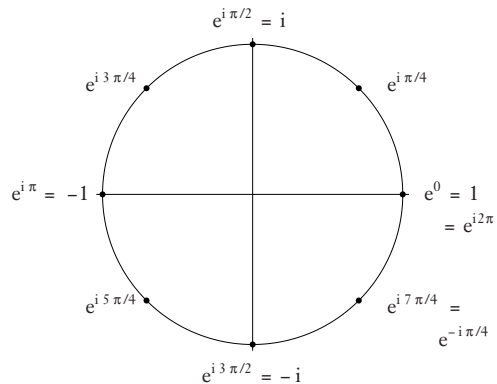
$$e^0 = 1 = (1, 0),$$

$$e^{i\pi/4} = (1, 1)/\sqrt{2} = (1+i)/\sqrt{2},$$

$$e^{i\pi/2} = i, \text{ und weiter:}$$

Eulersche Identität

$$e^{i\pi} + 1 = 0$$



Aus $e^{i2\pi} = 1$ und dem Additionstheorem ergibt sich weiter:

Längentreue Kreisaufwicklung der y -Achse durch e^z

Periodizität der komplexen Exponentialfunktion

$$e^z = e^{z + i2\pi} = e^{z + k2\pi} \text{ für alle } z \in \mathbb{C} \text{ und alle } k \in \mathbb{Z}.$$

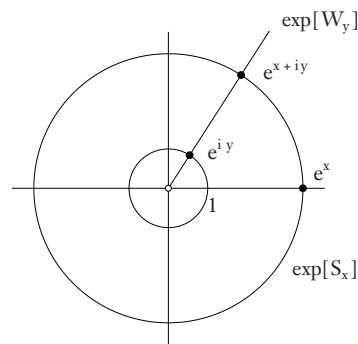
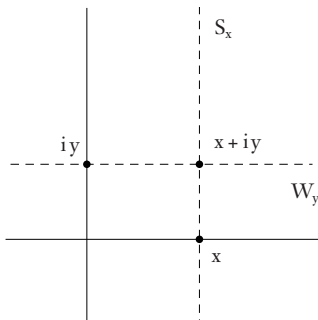
Wir haben in unserer Diskussion den Kreisumfang 2π aus der Geometrie übernommen. Heute wird die Zahl π in der Analysis oft analytisch eingeführt: Man definiert 2π als die Periode der komplexen Exponentialfunktion, d. h. als die kleinste positive reelle Zahl y mit $e^{iy} = 1$. Dass 2π dann tatsächlich der geometrische Umfang des Einheitskreises ist, folgt aus der Längentreue der Kreisaufwicklung. Wir diskutieren in der letzten Sektion dieses Kapitels, wie man diese Längentreue nachweisen kann.

6.8 Bilder der komplexen Exponentialfunktion

Satz (Abbildungseigenschaften der komplexen Exponentialfunktion)

Die komplexe Exponentialfunktion $\exp : \mathbb{C} \rightarrow \mathbb{C}$ bildet

- (a) jede Senkrechte $S_x = \{x\} \times \mathbb{R} = \{x + iy \mid y \in \mathbb{R}\}$ der Ebene surjektiv auf den Kreis $K_r = \{z \in \mathbb{C} \mid |z| = r\}$ mit Radius $r = e^x > 0$ ab. Dabei wird jedes halboffene Geradenstück von S_x der Länge 2π bijektiv auf K_r abgebildet.
- (b) jede Waagrechte $W_y = \mathbb{R} \times \{y\} = \{x + iy \mid x \in \mathbb{R}\}$ der Ebene bijektiv auf die Halbgerade $G = \{a e^{iy} \mid a > 0\}$ ab, die durch den Punkt e^{iy} des Einheitskreises verläuft, also mit der positiven x-Achse den im Bogenmaß gegen den Uhrzeigersinn gemessenen Winkel y einschließt.



Der Nullpunkt ist der einzige Punkt der Ebene, der nicht im Wertebereich von \exp liegt, d. h., es gilt $\text{Bild}(\exp) = \mathbb{C}^* = \mathbb{C} - \{0\}$. Ist $z \in \mathbb{C}$ und $e^z = w$, so werden genau die Punkte $z + k2\pi i$, $k \in \mathbb{Z}$, durch die Exponentialfunktion auf w abgebildet. Ist I ein halboffenes Intervall der Länge 2π , so gibt es für alle $z \neq 0$ eindeutige $r > 0$ und $\varphi \in I$ mit

$$z = r e^{i\varphi}. \quad (\text{Darstellung in Polarkoordinaten } r, \varphi \text{ bzgl. des Winkelintervalls } I)$$

Für Funktionen $f: \mathbb{C} \rightarrow \mathbb{C}$ steht uns eine einfache graphische Methode nicht mehr zur Verfügung. Die Visualisierung ist schwieriger, aber nicht unmöglich. Wir illustrieren am Beispiel der komplexen Exponentialfunktion, wie man mit der „Geometrie“ oder „Abbildungsdynamik“ von komplexen Funktionen vertraut werden kann.

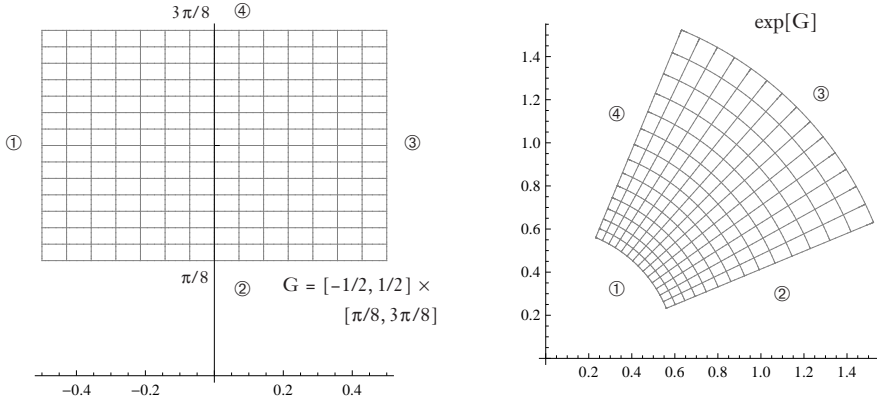
Der obige Satz folgt aus der in der letzten Sektion diskutierten Kreisaufwicklung der y-Achse und der Abspaltungsregel

$$e^{x+iy} = e^x e^{iy} \quad \text{für alle } x, y \in \mathbb{R}.$$

Die Wahl von $x = 0$ bei variablem y führt zu einer Aufwicklung der Senkrechten S_0 auf den Einheitskreis K , die Wahl von $x \neq 0$ zu einer Aufwicklung von S_x auf den Kreis mit Durchmesser $r = e^x$. Hier gilt $r < 1$ für $x < 0$ und $r > 1$ für $x > 0$. Analog können wir y festhalten und x als variabel ansehen, also Waagrechte W_y betrachten. Die Anwendung von \exp liefert eine vom Ursprung 0 ausgehende Halbgerade G , die durch den Punkt e^{iy} des Einheitskreises verläuft. Der Ursprung gehört dabei nicht zu G . Punkte auf G , die in der Einheitskreisscheibe liegen, entsprechen negativen x , und Punkte auf G außerhalb der Einheitskreisscheibe entsprechen positiven x .

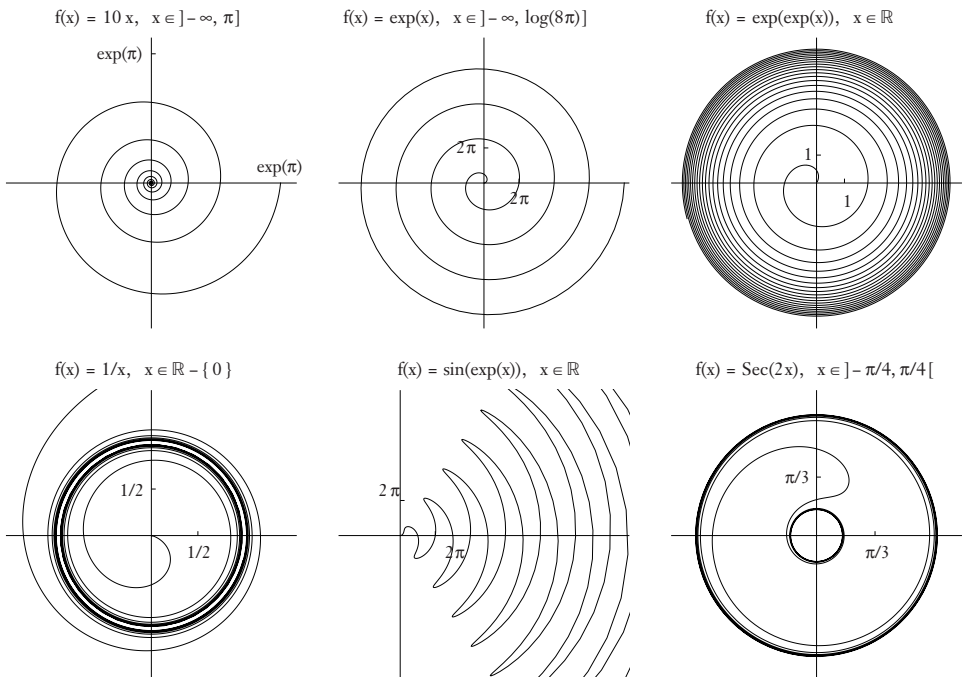
Anschaulich gilt: Verschieben wir eine Senkrechte in der Ebene, so sind die Bilder unter \exp sich aufblähende oder zusammenziehende Kreise. Verschieben wir eine Waagrechte in der Ebene, so sehen die Bilder unter \exp aus wie die Strahlen eines im Ursprung angebrachten Leuchtturms. Dies können wir auch so formulieren:

Ein Gitter der Ebene wird durch \exp in y -Richtung kreisförmig aufgefächert und in x -Richtung exponentiell gestreckt.



Beispiele

Die folgenden Diagramme visualisieren für reelle Funktionen $f : P \rightarrow \mathbb{R}$ das Bild des Graphen von f unter $\exp : \mathbb{C} \rightarrow \mathbb{C}$, d.h. die Menge $\exp[\{ (x, f(x)) \mid x \in P \}]$.



6.9 Sinus und Kosinus

Definition (Sinus, Kosinus)

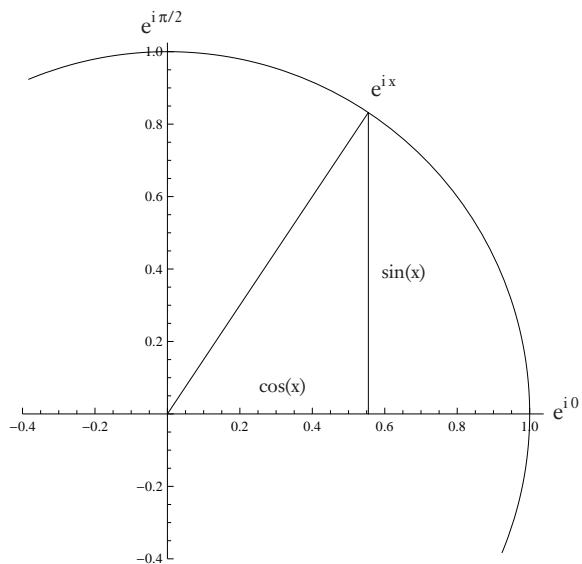
Wir definieren $\sin : \mathbb{R} \rightarrow \mathbb{R}$ und $\cos : \mathbb{R} \rightarrow \mathbb{R}$ durch:

$$\cos(x) = \operatorname{Re}(e^{ix}),$$

$$\sin(x) = \operatorname{Im}(e^{ix}) \quad \text{für alle } x \in \mathbb{R}.$$

Diese Funktionen heißen der *Kosinus* bzw. der *Sinus* auf \mathbb{R} .

Für Anfänger ist es oft überraschend, dass eine so komplizierte Funktion wie die komplexe Exponentialfunktion zur Definition der trigonometrischen Grundfunktionen herangezogen wird. Aber unsere Diskussion des Verhaltens der Exponentialfunktion in \mathbb{C} zeigt, dass wir in der Tat die guten alten Bekannten definieren, nur ohne die Verwendung von Dreiecken und Strahlensatz. Denn für alle reellen Zahlen x ist e^{ix} ein Punkt des Einheitskreises der Ebene. Gilt nun $e^{ix} = (a, b)$, so gilt nach Definition $\cos(x) = a$ und $\sin(x) = b$, und wir haben ein rechtwinkliges Dreieck mit den Punkten



$$0, \quad (0, \cos(x)), \quad e^{ix} = (\cos(x), \sin(x))$$

vorliegen, das im Nullpunkt den Winkel x besitzt. Die klassische geometrische Bedeutung von Kosinus und Sinus lässt sich also, die längentreue Aufwicklungseigenschaft der komplexen Exponentialfunktion vorausgesetzt, sofort aus der Definition über den Real- und Imaginärteil von e^{ix} gewinnen. Kosinus und Sinus sind die Längen der An- und Gegenkathete eines rechtwinkligen Dreiecks mit Winkel x und Hypotenuse 1. Der Strahlensatz liefert die üblichen Formeln „Ankathete durch Hypotenuse“ bzw. „Gegenkathete durch Hypotenuse“.

Die Verwendung der komplexen Exponentialfunktion legt nicht nur die klassische Bedeutung der beiden Funktionen im Reich der rechtwinkligen Dreiecke offen, sondern sie betont auch eine physikalische Interpretation:

Kosinus und Sinus sind die x- und y-Projektionen eines sich in der Zeit t gleichmäßig mit der Periode 2π auf dem Einheitskreis bewegenden Punktes.

Aus unserer Wertetabelle für die Kreisaufwicklung erhalten wir wichtige Werte für den Kosinus und Sinus:

Beispiele

- (1) $\cos(0) = \operatorname{Re}(e^{i0}) = \operatorname{Re}(1) = 1, \quad \sin(0) = \operatorname{Im}(e^{i0}) = \operatorname{Im}(1) = 0.$
- (2) $\cos(\pi/4) = \operatorname{Re}(e^{i\pi/4}) = \operatorname{Re}((1, 1)/\sqrt{2}) = 1/\sqrt{2},$
 $\sin(\pi/4) = \operatorname{Im}(e^{i\pi/4}) = \operatorname{Im}((1, 1)/\sqrt{2}) = 1/\sqrt{2}.$
- (3) $\cos(\pi/2) = \operatorname{Re}(e^{i\pi/2}) = \operatorname{Re}(i) = 0, \quad \sin(\pi/2) = \operatorname{Im}(e^{i\pi/2}) = \operatorname{Im}(i) = 1.$

Die erste positive Nullstelle des Kosinus kann zur analytischen Definition von π verwendet werden. Man setzt $\pi/2 =$ „das kleinste $x > 0$ mit $\cos x = 0$ “. Dies ist gleichwertig zur Definition von 2π als kleinstes $x > 0$ mit $e^{ix} = 1$ (vgl. hierzu 6.12).

Aus der Aufwicklungseigenschaft ergeben sich Nullstellen, Periodizität, Parität, Monotonieverhalten, lokale Extremwerte usw.

Die Formeln

$$e^{ix} = \cos(x) + i \sin(x) \quad (\text{Eulersche Formel}),$$

$$\cos(x) = \operatorname{Re}(e^{ix}), \quad \sin(x) = \operatorname{Im}(e^{ix})$$

erlauben bestechend kurze Beweise, insbesondere für die Additionstheoreme. So gilt

$$\cos(x+y) = \operatorname{Re}(e^{i(x+y)}) = \operatorname{Re}(e^{ix} e^{iy}) =$$

$$\operatorname{Re}((\cos(x) + i \sin(x))(\cos(y) + i \sin(y))) = \cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y).$$

Rechenregeln

$$(1) \cos(x)^2 + \sin(x)^2 = 1, \quad \sin(x + \pi/2) = \cos(x), \quad \cos(x + \pi/2) = -\sin(x).$$

$$(2) \sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y),$$

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y). \quad (\text{Additionstheoreme})$$

$$(3) \sin(x) - \sin(y) = 2 \sin((x-y)/2) \cos((x+y)/2),$$

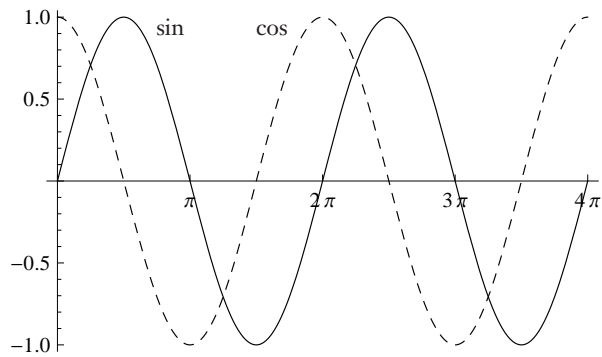
$$\cos(x) - \cos(y) = -2 \sin((x-y)/2) \sin((x+y)/2).$$

Aus $\operatorname{Re}(z) = (z + \bar{z})/2$, $\operatorname{Im}(z) = (z - \bar{z})/(2i)$ und $\bar{\bar{z}} = z$ gewinnen wir Reihendarstellungen, die sich auch zur effektiven Berechnung eignen:

Potenzreihenentwicklungen von sin und cos

$$\cos(x) = (\sum_n i^n x^n/n! + \sum_n (-i)^n x^n/n!)/2 = \sum_n (-1)^n x^{2n}/(2n)!,$$

$$\sin(x) = (\sum_n i^n x^n/n! - \sum_n (-i)^n x^n/n!)/(2i) = \sum_n (-1)^n x^{2n+1}/(2n+1)!.$$



6.10 Weitere trigonometrische Funktionen

Definition (*Tangens, Kotangens, Sekans, Kosekans*)

Wir setzen:

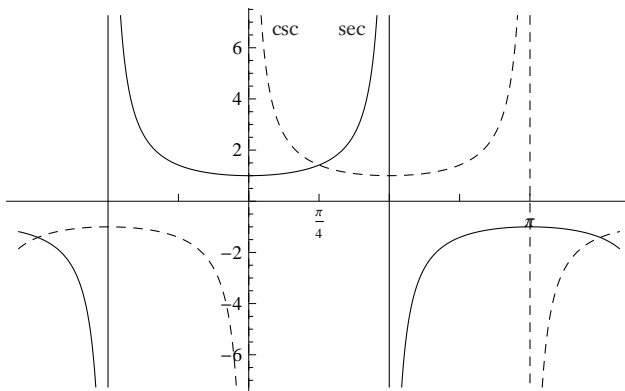
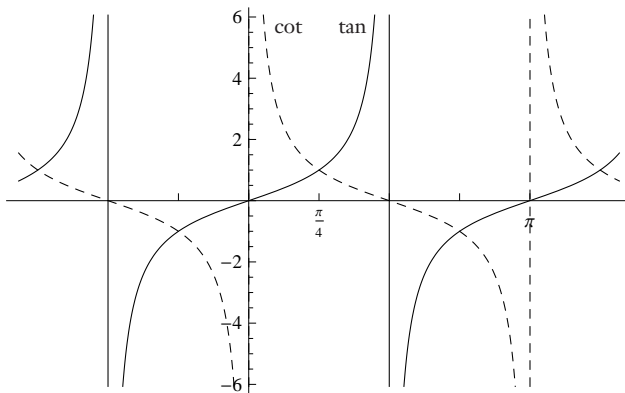
$$\tan(x) = \sin(x)/\cos(x) \quad \text{für alle } x \text{ mit } \cos(x) \neq 0,$$

$$\cot(x) = \cos(x)/\sin(x) \quad \text{für alle } x \text{ mit } \sin(x) \neq 0,$$

$$\sec(x) = 1/\cos(x) \quad \text{für alle } x \text{ mit } \cos(x) \neq 0,$$

$$\csc(x) = 1/\sin(x) \quad \text{für alle } x \text{ mit } \sin(x) \neq 0.$$

Diese Funktionen heißen *Tangens*, *Kotangens*, *Sekans* bzw. *Kosekans*.



Sind also

$$A = \{\pi/2 + k\pi \mid k \in \mathbb{Z}\}, \quad B = \{k\pi \mid k \in \mathbb{Z}\}$$

die Nullstellenmenge des Kosinus bzw. Sinus, so gilt

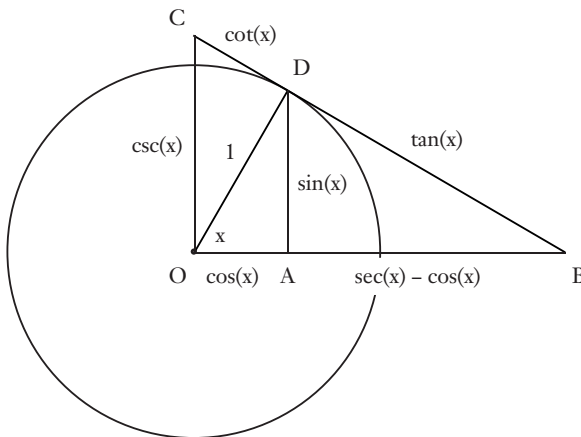
$$\tan : \mathbb{R} - A \rightarrow \mathbb{R}, \quad \cot : \mathbb{R} - B \rightarrow \mathbb{R}, \quad \sec : \mathbb{R} - A \rightarrow \mathbb{R}, \quad \csc : \mathbb{R} - B \rightarrow \mathbb{R}.$$

Die vier Funktionen sind stetig auf ihren Definitionsbereichen. Die Funktionsgraphen visualisieren viele weitere Eigenschaften (Monotonie, Periodizität, Verhalten an den Definitionslücken, Parität usw.), die sich aus den Eigenschaften des Sinus und Kosinus ergeben. Bemerkenswert ist, dass der Tangens das beschränkte Intervall $]-\pi/2, \pi/2[$ stetig und streng monoton steigend (und damit bijektiv) auf \mathbb{R} abbildet. Ebenso bildet der Kotangens das beschränkte Intervall $]0, \pi[$ stetig und streng monoton fallend auf \mathbb{R} ab. Damit ist zum Beispiel die Funktion $f:]0, 1[\rightarrow \mathbb{R}$ mit

$$f(x) = \tan(\pi x - \pi/2) \quad \text{für alle } x \in]0, 1[$$

eine stetige Bijektion zwischen $]0, 1[$ und \mathbb{R} .

Die vier Funktionen ließen sich prinzipiell mit Sinus und Kosinus darstellen. Dennoch ist es nützlich, eigene Bezeichnungen für sie einzuführen. Bereits $\cos(x)$ könnten wir ja immer durch $\sin(x + \pi/2)$ ersetzen. Da der Kosinus in geometrischen Kontexten häufig in Erscheinung tritt, würden wir ihn schnell vermissen, ein Ersetzen durch den Sinus wäre umständlich. Ebenso tauchen die vier anderen trigonometrischen Funktionen in natürlicher Weise auf. Betrachten wir einen Einheitskreis und das durch einen Winkel x zwischen 0 und 90 Grad gegebene Dreieck OAD, so sehen wir die Größen $\sin(x)$ und $\cos(x)$ vor uns. Zeichnen wir nun ein zweites Dreieck OBD wie im Diagramm unten, so können wir auch die vier anderen Größen $\tan(x)$, $\cot(x)$, $\sec(x)$ und $\csc(x)$ entdecken:

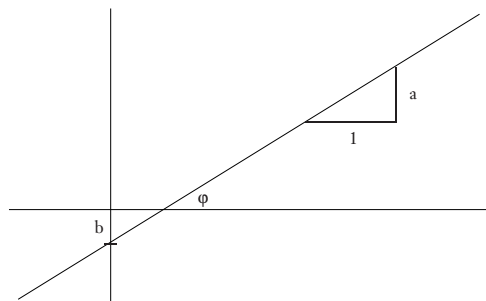


$$\begin{aligned}\cos(x) &= \overline{OA} \\ \sin(x) &= \overline{AD} \\ \tan(x) &= \overline{DB} \\ \cot(x) &= \overline{DC} \\ \sec(x) &= \overline{OB} \\ \csc(x) &= \overline{OC}\end{aligned}$$

In der Analysis spielt neben dem Sinus und dem Kosinus vor allem der Tangens eine wichtige Rolle. Er wird gebraucht, um Geraden zu untersuchen. Die Gerade g mit

$$g(x) = ax + b \quad \text{für alle } x$$

hat die Steigung a . Der Tangens kommt ins Spiel, wenn wir den Winkel φ betrachten, den die Gerade mit der x -Achse einschließt: Es gilt $\tan(\varphi) = a$ (vgl. 6.11).



6.11 Die Arkusfunktionen

Definition (Arkusfunktionen)

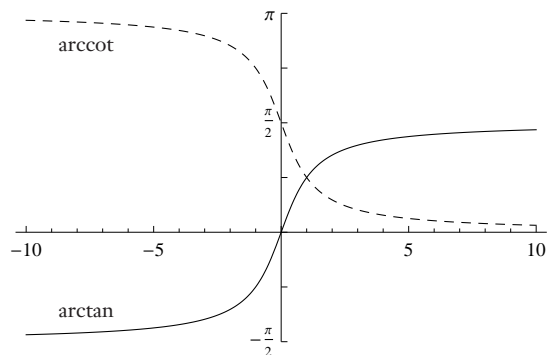
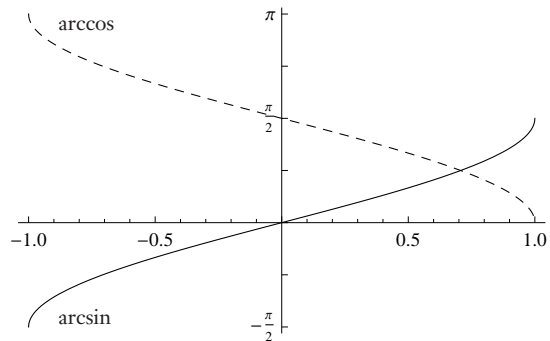
Der *Arkussinus* $\arcsin : [-1, 1] \rightarrow \mathbb{R}$, *Arkuskosinus* $\arccos : [-1, 1] \rightarrow \mathbb{R}$, *Arkustangens* $\arctan : \mathbb{R} \rightarrow \mathbb{R}$ und der *Arkuskotangens* $\operatorname{arccot} : \mathbb{R} \rightarrow \mathbb{R}$ werden definiert durch:

$$\arcsin = (\sin|[-\pi/2, \pi/2])^{-1}, \quad \arccos = (\cos|[0, \pi])^{-1},$$

$$\arctan = (\tan|]-\pi/2, \pi/2])^{-1}, \quad \operatorname{arccot} = (\cot|[0, \pi])^{-1}.$$

Die trigonometrischen Funktionen sind nicht injektiv und müssen deswegen geeignet eingeschränkt werden, bevor wir sie umkehren können. Die Intervalle, die wir zur Einschränkung verwenden können, sind nicht eindeutig bestimmt, der Sinus ist beispielsweise auch auf dem Intervall $[\pi/2, 3/2\pi]$ injektiv, und er nimmt dort ebenfalls alle Werte in $[-1, 1]$ an. Die in den Definitionen verwendeten Einschränkungen erscheinen aber aufgrund ihrer Nähe zur 0 als besonders natürlich. Als Merksatz gilt, dass die Einschränkungen der Funktionen, die ein „o“ enthalten, bei der 0 beginnen (wobei die 0 bei $\cot = \cos/\sin$ natürlich auszuschließen ist). Bei den anderen Funktionen liegt die Null in der Mitte des Intervalls und alle Intervalle haben die Länge π . Sie werden zum Wertebereich der Umkehrfunktionen. Die Umkehrfunktionen erben wie immer die Monotonieeigenschaften ihrer Herkunft.

Die Arkusfunktionen bereiten keine großen Schwierigkeiten. Die einzige Fehlerquelle besteht in der Vernachlässigung der Einschränkung:



Beispiele

$\sin(\arcsin(x)) = x$	für alle $x \in [-1, 1]$,
$\arcsin(\sin(x)) = x$	für alle $x \in [-\pi/2, \pi/2]$,
$\arcsin(\sin(x)) = \pi - x$	für alle $x \in [\pi/2, 3/2\pi]$,
$\arcsin(\sin(x)) = x - 2\pi$	für alle $x \in [3/2\pi, 5/2\pi]$ usw.

In der Analysis ist der Arkustangens die prominenteste Umkehrfunktion. Wir diskutieren zwei wichtige Anwendungen.

Winkelberechnung für Geraden

Der Winkel $\varphi \in]-\pi, \pi[$, den eine Gerade mit der Steigung $a \neq 0$ mit der positiven x -Achse einschließt, wird berechnet durch $\varphi = \arctan(a)$ (vgl. 6.10).

Polarkoordinaten

Unsere Analyse der komplexen Exponentialfunktion in 6.7 und 6.8 hat ergeben, dass wir jede komplexe Zahl $z \neq 0$ in der Form

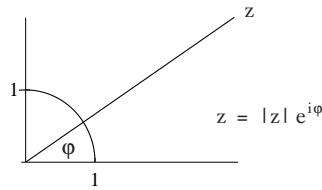
$$z = r e^{i\varphi} = r e^{i(\varphi + k2\pi)}, \quad k \in \mathbb{Z}, \quad (\text{Darstellung in Polarkoordinaten})$$

schreiben können. Wir nennen dann (r, φ) *Polarkoordinaten* der komplexen Zahl z . Eindeutigkeit erreichen wir, indem wir ein halboffenes Intervall der Länge 2π auszeichnen. Wir wählen hier $[0, 2\pi[$ (oft wird auch $]-\pi, \pi]$ verwendet). Für alle $z \in \mathbb{C}$ mit $z \neq 0$ sei also

$$\arg(z) = \text{„das eindeutige } \varphi \in [0, 2\pi[\text{ mit } z = |z| e^{i\varphi}\text{“}. \quad (\text{Argument von } z)$$

Zur Berechnung des Arguments wird der Arkustangens eingesetzt. Ist $x > 0$ und $y \geq 0$ für $z = (x, y)$, so gilt

$$\arg(z) = \arctan(y/x).$$



Liegt z in den anderen Quadranten, so sind Korrekturwinkel notwendig, da der Arkustangens ja nur Werte in $]-\pi/2, \pi/2[$ annimmt. Es ist günstig, beide Werte x und y einer Funktion zu übergeben, die die Fallunterscheidung durchführt. Wir definieren hierzu $\arctan_2 : \mathbb{R}^2 - \{0\} \rightarrow \mathbb{R}$:

$$\arctan_2(x, y) = \begin{cases} \arctan(y/x), & \text{falls } x > 0 \text{ und } y \geq 0, \\ \arctan(y/x) + 2\pi, & \text{falls } x > 0 \text{ und } y < 0, \\ \arctan(y/x) + \pi, & \text{falls } x < 0, \\ \pi/2, & \text{falls } x = 0 \text{ und } y > 0, \\ 3\pi/2, & \text{falls } x = 0 \text{ und } y < 0. \end{cases}$$

Es gilt dann:

$$\arg(z) = \arctan_2(x, y) \quad \text{für alle } z = (x, y) \in \mathbb{C} - \{0\}.$$

Also gibt

$$(\sqrt{x^2 + y^2}, \arctan_2(x, y)) \quad (\text{Umrechnungsformel von } x, y \text{ nach } r, \varphi)$$

die Umrechnung von kartesischen Koordinaten in Polarkoordinaten einer komplexen Zahl $z = (x, y) \neq 0$ an. Der Weg von Polarkoordinaten (r, φ) zu den kartesischen Koordinaten (x, y) wird mit Hilfe des Sinus und Kosinus berechnet. Gilt nämlich $z = r e^{i\varphi}$, so ist

$$x = r \cos(\varphi) \quad \text{und} \quad y = r \sin(\varphi). \quad (\text{Umrechnungsformel von } r, \varphi \text{ nach } x, y)$$

6.12 Die Brücke zur Geometrie

Satz (Definitionen der Zahl π)

Die folgenden Definitionen der Zahl π sind äquivalent:

Analytische Definition von π (über die erste Nullstelle des Kosinus)

$\pi/2 =$ „das kleinste $x > 0$ mit $\cos(x) = 0$ “.

Analytische Definition von π (über die Periode der komplexen Exponentialfunktion)

$2\pi =$ „das kleinste $y > 0$ mit $e^{iy} = 1$ “.

Geometrische Definition von π

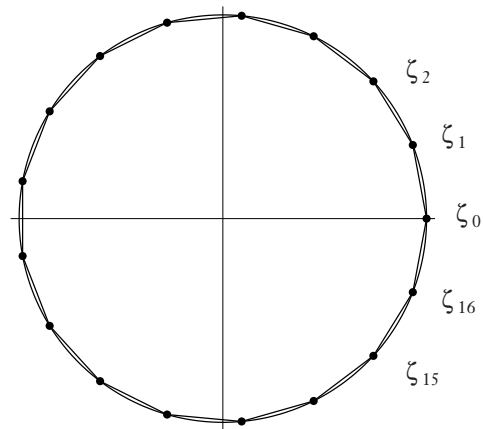
$2\pi = \lim_{n \geq 3} S_n$,

wobei S_n der Umfang eines in den Einheitskreis einbeschriebenen regelmäßigen n -Ecks ist.

Die geometrische Definition der Kreiszahl π spiegelt den seit Archimedes verwendeten Ansatz wider, den Kreis durch einfache geometrische Figuren zu approximieren. Dagegen basiert die analytische Definition von π auf der komplexen Exponentialfunktion $\exp: \mathbb{C} \rightarrow \mathbb{C}$,

$$\exp(z) = e^z = \sum_n z^n/n! \quad \text{für alle } z \in \mathbb{C},$$

deren Bezug zum Kreis zunächst unklar ist. Er ergibt sich, ohne jede Anleihe bei der Geometrie, in mehreren Schritten.



Für das 17-Eck gilt

$$2\pi - S_{17} = 2\pi - 34 \sin(\pi/17) = 0,0357017.$$

1. Schritt: Ortung von e^{ix} auf der Kreislinie

Für alle $x \in \mathbb{R}$ gilt $|e^{ix}| = 1$.

Dies ergibt sich (wie wir in 6.7 schon gesehen hatten) aus $\exp(-ix) = \overline{\exp(ix)}$, der allgemeinen Formel $z\bar{z} = |z|^2$ und dem Additionstheorem:

$$|e^{ix}|^2 = e^{ix} \cdot e^{-ix} = e^{ix - ix} = e^0 = 1.$$

Die Exponentialfunktion bildet also die imaginäre Achse stetig auf den Einheitskreis ab. Das ist auf viele Arten möglich und wir können „ $e^{ix} = 1$ für alle x “ noch nicht ausschließen. Für eine genauere Betrachtung verwenden wir die Sinus- und Kosinusfunktion auf \mathbb{R} :

$$e^{ix} = (\cos(x), \sin(x)) = \cos(x) + i\sin(x) \quad \text{für alle } x \in \mathbb{R}.$$

Nun können wir e^{ix} für kleine positive x genauer orten:

2. Schritt: Technische Analyse von e^{ix}

Auf dem Intervall $[0, 2]$ ist der Kosinus streng monoton fallend und der Sinus ist dort größergleich 0. Weiter gilt

$$1 - x^2/2 < \cos(x) < 1 - x^2/2 + x^4/24 \quad \text{für alle } x \in]0, 2].$$

Insbesondere ist $\cos(2) < 1 - 4/2 + 16/24 = -1/3 < 0$.

Wegen $\cos(0) = 1$ hat der Kosinus nach dem Zwischenwertsatz im Intervall $[0, 2]$ genau eine Nullstelle, die wir als $\pi/2$ bezeichnen. Da $e^{i\pi/2} = (0, \sin(\pi/2))$ ein Punkt des Einheitskreises mit $\sin(\pi/2) \geq 0$ ist, ist $\sin(\pi/2) = 1$ und damit

$$e^{i\pi/2} = (0, 1) = i \quad \text{und} \quad e^{i2\pi} = (e^{i\pi/2})^4 = i^4 = 1.$$

Damit ist π analytisch definiert. Die Untersuchung zeigt weiter, dass e^{ix} für $x \geq 0$ die Kreislinie gegen den Uhrzeigersinn mit der Periode 2π durchläuft: Man kann die Monotonieeigenschaften des Kosinus und Sinus aus der technischen Analyse gewinnen, und sie schließen ein kompliziertes Durchlaufen der Kreislinie mit tänzelnden Hin- und Rückbewegungen aus. Es ist aber keineswegs klar, dass 2π der Kreisumfang ist, wir kennen bislang nur die Abschätzung $0 < \pi/2 < 2$. Die Verbindung zu Archimedes liefert erst der nächste Schritt:

3. Schritt: n -te Einheitswurzeln und n -Eck

Für alle $n \geq 3$ spannen $\zeta_0 = e^{0 \cdot 2\pi i/n}, \dots, \zeta_{n-1} = e^{(n-1)2\pi i/n}$ ein regelmäßiges n -Eck mit dem Umfang $2n \sin(\pi/n)$ auf.

In der Tat gilt sogar für alle ganzen Zahlen k :

$$|e^{(k+1)2\pi i/n} - e^{k2\pi i/n}| = |e^{(k+1/2)2\pi i/n}| |e^{\pi i/n} - e^{-\pi i/n}| = |2 \operatorname{Im}(e^{i\pi/n})| = 2 \sin(\pi/n).$$

Der von k unabhängige Wert und $|e^{ix}| = 1$ zeigt, dass die Punkte $\zeta_0, \dots, \zeta_{n-1}$ ein regelmäßiges n -Eck aufspannen, und er liefert die Formel für den Umfang. Übrig bleibt:

4. Schritt: Grenzübergang

Es gilt $\lim_{n \geq 1} 2n \sin(\pi/n) = 2\pi$.

Zum Beweis verwendet man den Grenzwert $\lim_{x \rightarrow 0} \sin(x)/x = 1$ (den man zum Beispiel aus der Reihenentwicklung des Sinus gewinnen kann):

$$\lim_{n \geq 1} 2n \sin(\pi/n) = 2\pi \lim_{n \geq 1} \sin(\pi/n)/(\pi/n) = 2\pi \cdot 1 = 2\pi.$$

Damit ist die Äquivalenz der Definitionen von π etabliert. Führen wir den dritten und vierten Schritt statt 2π mit einem $x \in [0, 2\pi[$ durch, so erhalten wir, dass der durch 0 und e^{ix} definierte Bogen des Einheitskreises die Länge x hat. Damit ist die geometrische Bedeutung von $\sin(x)$ und $\cos(x)$ aufgezeigt. Weiter liefert die Formel $r_1 e^{ix_1} \cdot r_2 e^{ix_2} = r_1 r_2 e^{i(x_1 + x_2)}$ die geometrische Multiplikationsregel.

7. Kapitel

Differentiation

7.1 Geraden und ihre Darstellungen

Definition (Gerade)

Eine Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ heißt eine *Gerade*, falls $a, b \in \mathbb{R}$ existieren mit

$$g(x) = ax + b \quad \text{für alle } x \in \mathbb{R}.$$

Die reelle Zahl a heißt dann die *Steigung* von g .

Geraden gehören zu den wichtigsten Hilfsmitteln der Analysis. Die Grundidee der Differentialrechnung ist, eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ in einem Punkt p durch eine Gerade $g : \mathbb{R} \rightarrow \mathbb{R}$ zu approximieren. Bevor wir uns mit diesem Gedanken genauer auseinandersetzen, betrachten wir die durch Funktionen dargestellten Geraden noch etwas genauer.

Ist $g : \mathbb{R} \rightarrow \mathbb{R}$ eine Gerade mit $g(0) = 0$, so ist g eine *lineare Abbildung*, d. h., für alle $\alpha, \beta \in \mathbb{R}$ und alle $x_1, x_2 \in \mathbb{R}$ gilt

$$g(\alpha x_1 + \beta x_2) = \alpha g(x_1) + \beta g(x_2). \quad (\text{Linearität})$$

Beliebige Geraden sind *affine Abbildungen*, also Translationen von linearen Abbildungen. Ist nämlich $g : \mathbb{R} \rightarrow \mathbb{R}$ eine Gerade, so existiert eine eindeutige Gerade $h : \mathbb{R} \rightarrow \mathbb{R}$ mit $h(0) = 0$ und ein eindeutiges $b \in \mathbb{R}$ mit

$$g(x) = h(x) + b \quad \text{für alle } x \in \mathbb{R}.$$

Wir nennen dann h auch den *linearen Anteil* von h . Die Geraden g und h besitzen dieselbe Steigung und es gilt $b = g(0)$. Im Fall $b \neq 0$ ist die Gerade g keine lineare Abbildung, da dann $g(0) = b \neq 0$ gilt, während jede lineare Abbildung den Nullpunkt auf 0 abbildet.

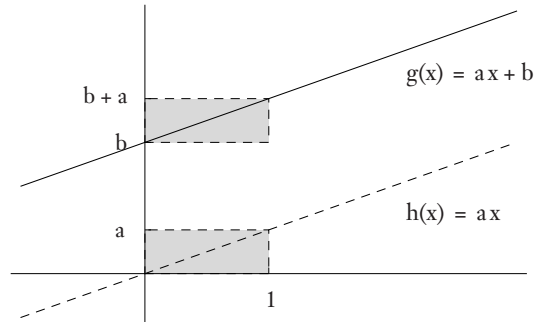
Sind zwei Punkte (x_1, y_1) und (x_2, y_2) der Ebene mit $x_1 \neq x_2$ gegeben, so gibt es genau eine Gerade g , die durch die beiden Punkte verläuft, d. h. für die gilt:

$$g(x_1) = y_1 \quad \text{und} \quad g(x_2) = y_2.$$

Bemerkung

Ist $x_1 = x_2$ und $y_1 \neq y_2$, so gibt es eine eindeutige Gerade durch die beiden Punkte im Sinne der Geometrie, nämlich die Senkrechte durch $x_1 = x_2$. Diese geometrische Gerade erlaubt aber keine funktionale Darstellung der Form $g(x) = ax + b$. In der Analysis wird dieser Fall durch uneigentliche Grenzwerte repräsentiert (vgl. etwa die Ableitung der Wurzelfunktion im Nullpunkt).

Es gibt zwei wichtige Möglichkeiten, die Gerade $g : \mathbb{R} \rightarrow \mathbb{R}$, die durch (x_1, y_1) und (x_2, y_2) mit $x_1 \neq x_2$ verläuft, darzustellen.



Eine Gerade g der Steigung a und ihr linearer Anteil h .

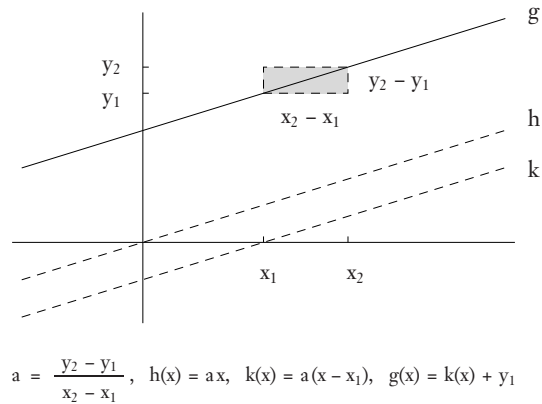
Merkregel: „1 nach rechts, a nach oben“,
„ x nach rechts, ax nach oben“.

Steigungsform oder Punkt-Richtungs-Darstellung

$$g(x) = \frac{y_2 - y_1}{x_2 - x_1} (x - x_1) + y_1.$$

Die Steigungsform kann man sich so klarmachen (und merken):

Der Wert $a = (y_2 - y_1)/(x_2 - x_1)$ ist die Steigung der gesuchten Geraden. Die Gerade $h(x) = ax$ der Steigung a durch den Nullpunkt verschieben wir nun um x_1 nach rechts und um y_1 nach oben. Die Verschiebung nach rechts liefert $a(x - x_1)$ und die Verschiebung nach oben $a(x - x_1) + y_1$.



Beispiel

Die Steigungsform der Geraden durch die Punkte (2, 5) und (4, 8) lautet

$$g(x) = (8 - 5)/(4 - 2) (x - 2) + 5 = 3/2 (x - 2) + 5.$$

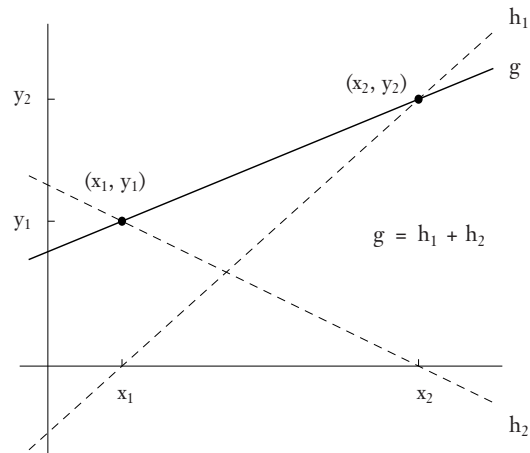
In „ $ax + b$ “-Form gilt also $g(x) = 3/2 x + 2$ für alle $x \in \mathbb{R}$.

Die Steigungsform behandelt die beiden Punkte x_1 und x_2 unsymmetrisch. Auf den Punkt (x_1, y_1) wird eine Gerade der richtigen Steigung aufgesetzt, sodass der Punkt (x_2, y_2) getroffen wird. Dieser Ansatz ist gerade für die Differentialrechnung besonders natürlich. Es gibt aber auch eine symmetrische Darstellung, bei der die gesuchte Gerade als Summe zweier Geraden h_1 und h_2 erscheint, die durch die Punkte $(x_1, 0)$ und (x_2, y_2) bzw. (x_1, y_1) und $(x_2, 0)$ verlaufen:

Lagrangesche Form oder Zweipunktdarstellung

$$g(x) = \frac{y_2}{x_2 - x_1} (x - x_1) + \frac{y_1}{x_1 - x_2} (x - x_2).$$

Dass die Lagrangesche Darstellung mit der Steigungsform übereinstimmt, kann man nachrechnen oder so einsehen: Beide Formen definieren Geraden, und Einsetzen von x_1 und x_2 zeigt, dass beide Geraden durch die Punkte (x_1, y_1) und (x_2, y_2) verlaufen.



7.2 Differenzen- und Differentialquotienten

Definition (*Differenzen- und Differentialquotient, Sekante, Tangente, differenzierbar*)

Sei $f: P \rightarrow \mathbb{R}$ eine Funktion, und sei $p \in P$.

(a) Für alle $x \in P$ mit $x \neq p$ heißt

$$a(p, x) = \frac{f(x) - f(p)}{x - p}$$

der *Differenzenquotient* von f für p und x . Die Gerade $s: \mathbb{R} \rightarrow \mathbb{R}$ mit

$$s(x) = f(p) + a(p, x)(x - p) \quad \text{für alle } x \in \mathbb{R}$$

heißt die *Sekante* von f durch $(p, f(p))$ und $(x, f(x))$.

(b) Existiert

$$f'(p) = \lim_{x \rightarrow p, x \in P, x \neq p} \frac{f(x) - f(p)}{x - p},$$

so heißt f *differenzierbar im Punkt* p und $f'(p)$ der *Differentialquotient* oder die *Ableitung* von f im Punkt p . Die Gerade $g: \mathbb{R} \rightarrow \mathbb{R}$ mit

$$g(x) = f(p) + f'(p)(x - p) \quad \text{für alle } x \in \mathbb{R}$$

heißt dann die *Tangente* an f im Punkt $(p, f(p))$.

Existiert $f'(p)$ für alle p , so heißt f *differenzierbar* und $f': P \rightarrow \mathbb{R}$ die *Ableitung* von f .

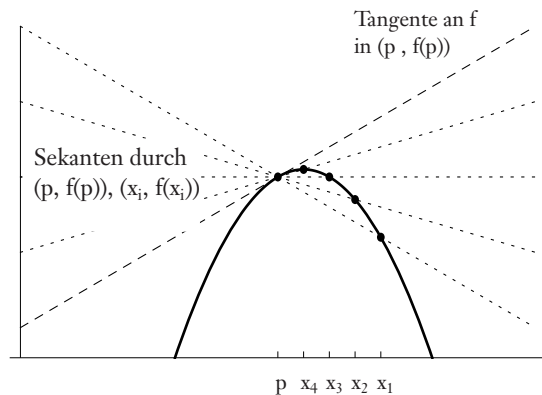
Neben f' sind auch Df , $df(x)/dx$ und $d/dx f(x)$ gebräuchlich, anstelle von $f'(p)$ auch $Df(p)$, $df/dx(p)$ und $df(x)/dx|_{x=p}$.

Der Differenzenquotient für p und x gibt die Steigung $a(p, x)$ der Geraden durch die Punkte $(p, f(p))$ und $(x, f(x))$ an. Als Steigung $f'(p)$ von f im Punkt p sehen wir, im Fall der Existenz, den Grenzwert von $a(p, x)$ für „ x gegen p “ an.

Viele Mathematiker betrachten den Differentialquotienten $a = f'(p)$ als einen geeigneten „Kode“ für den linearen Anteil $h: \mathbb{R} \rightarrow \mathbb{R}$, $h(x) = ax$, der Tangente von f im Punkt $(p, f(p))$. Die Ableitung wird dann sowohl als Zahl wie auch als lineare Abbildung betrachtet. In der mehrdimensionalen Analysis wird der zweite Aspekt noch deutlicher. Dort treten Matrizen an die Stelle der Zahl a .

Setzen wir $h = x - p$, so erhalten wir

$$f'(p) = \lim_{h \rightarrow 0, h \neq 0, p+h \in P} \frac{f(p+h) - f(p)}{h}. \quad (h\text{-Formulierung der Ableitung})$$



Wir notieren die Grenzwerte der Ableitung kurz als „ $\lim_{x \rightarrow p}$ “ und „ $\lim_{h \rightarrow 0}$ “.

Beispiele

(1) Ist $f(x) = c$ für alle x , so gilt $\lim_{x \rightarrow p} \frac{f(x) - f(p)}{x - p} = \lim_{x \rightarrow p} \frac{c - c}{x - p} = 0$.

(2) Ist $f(x) = x^n$, $n \geq 1$, für alle x , so gilt $f'(p) = np^{n-1}$ für alle $p \in \mathbb{R}$, denn

$$\begin{aligned} \lim_{x \rightarrow p} \frac{f(x) - f(p)}{x - p} &= \lim_{x \rightarrow p} \frac{x^n - p^n}{x - p} = \\ \lim_{x \rightarrow p} (x^{n-1}p^0 + x^{n-2}p^1 + \dots + x^1p^{n-2} + x^0p^{n-1}) &= np^{n-1}. \end{aligned}$$

Die Tangente an f im Punkt $(p, f(p))$ ist definiert durch $g(x) = np^{n-1}(x - p) + p^n$.

(3) Aus $\lim_{h \rightarrow 0} (e^h - 1)/h = 1$ und dem Additionstheorem folgt, dass für alle p gilt:

$$\lim_{h \rightarrow 0} \frac{e^{p+h} - e^p}{h} = e^p \lim_{h \rightarrow 0} \frac{e^h - e^0}{h} = e^p \cdot 1 = e^p.$$

Damit gilt also $\exp'(p) = \exp(p)$ für alle p , d.h. $\exp' = \exp$.

(4) Aus $\lim_{h \rightarrow 0} \sin(h)/h = 1$ und dem Additionstheorem folgt, dass für alle p gilt:

$$\begin{aligned} \lim_{h \rightarrow 0} (\sin(p+h) - \sin(p))/h &= \lim_{h \rightarrow 0} 2 \sin(h/2) \cos(p+h/2)/h = \\ \lim_{h \rightarrow 0} \sin(h/2)/(h/2) \cdot \lim_{h \rightarrow 0} \cos(p+h/2) &= 1 \cdot \cos(p) = \cos(p). \end{aligned}$$

Also gilt $\sin' = \cos$. Ähnlich zeigt man, dass $\cos' = -\sin$.

(5) Die Funktion f auf \mathbb{R} mit $f(x) = |x|$ ist im Punkt 0 nicht differenzierbar, da

$$\begin{aligned} \lim_{h \uparrow 0} (f(0+h) - f(0))/h &= \lim_{h \uparrow 0} |h|/h = -1, \\ \lim_{h \downarrow 0} (f(0+h) - f(0))/h &= \lim_{h \downarrow 0} h/h = 1. \end{aligned}$$

(6) Die Funktion f auf $[0, \infty[$ mit $f(x) = \sqrt{x}$ ist im Punkt 0 nicht differenzierbar, da

$$\lim_{h \rightarrow 0} (\sqrt{h+0} - 0)/h = \lim_{h \rightarrow 0} h^{-1/2} = +\infty.$$

(7) Sei $f: [0, 1] \cup \{2\} \rightarrow \mathbb{R}$. Dann ist f im Punkt 2 nicht differenzierbar, da keine Folge in $P - \{p\}$ existiert, die gegen p konvergiert.

Ein $f: P \rightarrow \mathbb{R}$ heißt *linksseitig* (bzw. *rechtsseitig*) *differenzierbar* in p , falls

$$\lim_{x \uparrow p} \frac{f(x) - f(p)}{x - p} \quad \text{bzw.} \quad \lim_{x \downarrow p} \frac{f(x) - f(p)}{x - p}$$

existieren. Diese Grenzwerte heißen dann die *links-* bzw. *rechtsseitige Ableitung* von f im Punkt p . Die Übereinstimmung der beiden Werte ist gleichwertig zur Differenzierbarkeit von f in p , vorausgesetzt, es gibt Folgen, die von unten bzw. von oben gegen p in P konvergieren. Nach (5) ist die Betragsfunktion im Punkt 0 links- und rechtsseitig differenzierbar.

7.3 Lineare Approximationen

Satz (*Approximationssatz*)

Sei $f: P \rightarrow \mathbb{R}$ eine Funktion, und sei $p \in P$ ein Häufungspunkt von P . Dann sind die folgenden Aussagen äquivalent:

- (a) f ist differenzierbar in p .
- (b) Es gibt ein $a \in \mathbb{R}$ und eine Funktion $r: P \rightarrow \mathbb{R}$ mit

$$f(x) = f(p) + a(x - p) + r(x) \text{ für alle } x \in P \text{ und } \lim_{x \rightarrow p} \frac{r(x)}{x - p} = 0.$$

- (c) Es gibt eine in p stetige Funktion $s: P \rightarrow \mathbb{R}$ mit

$$f(x) = f(p) + s(x)(x - p) \text{ für alle } x \in P.$$

Gelten (a), (b) und (c), so ist $f'(p) = a = s(p)$.

Der Approximationssatz besagt, dass sich eine in p differenzierbare Funktion in einer Umgebung von p „gut“ durch die Tangente

$$g(x) = f(p) + a(x - p)$$

an f im Punkt $(p, f(p))$ ersetzen lässt und dass diese Ersetzbarkeit umgekehrt auch die Differenzierbarkeit in p nach sich zieht. Was „gut“ genau heißt, wird durch die Funktionen r und s in (b) bzw. (c) zum Ausdruck gebracht.

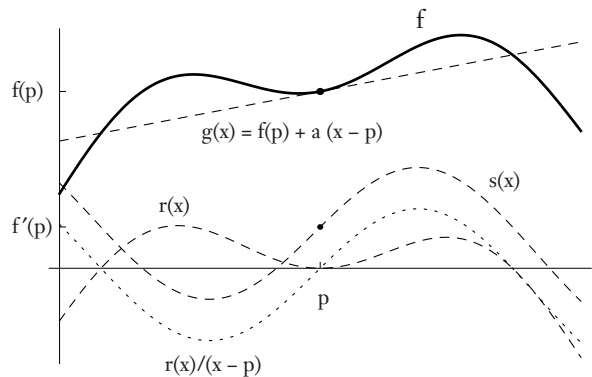
Wir betrachten zunächst die Darstellung in (b). Sie besagt:

f ist in der Nähe von p die Tangente g plus ein kleiner Rest r .

Die Kleinheit des Restes r wird durch $\lim_{x \rightarrow p} r(x)/(x - p) = 0$ präzisiert. Die Funktion r strebt nicht nur gegen 0, wenn x gegen p strebt, sondern sie strebt schneller gegen 0 als eine Gerade, so wie zum Beispiel x^2 für $x \rightarrow 0$ schneller gegen Null strebt als $x/100$. Man notiert die Darstellung auch suggestiv in der Form

$$f(x) = f(p) + a(x - p) + o(x - p) \text{ für } x \rightarrow p, \quad (\text{Klein-}o\text{-Notation})$$

wobei das sog. *Landau-Symbol* $o(x - p)$ für eine Funktion $r: P \rightarrow \mathbb{R}$ steht, für welche $\lim_{x \rightarrow p} r(x)/(x - p) = 0$ gilt. Analog steht $o((x - p)^k)$ für eine Funktion $r: P \rightarrow \mathbb{R}$ mit $\lim_{x \rightarrow p} r(x)/(x - p)^k = 0$. Derartige noch schneller gegen null konvergierende Reste werden wir bei der Taylor-Approximation kennenlernen (7.12).



Beispiel

Seien $f: \mathbb{R} \rightarrow \mathbb{R}$ die zweite Potenz und $p \in \mathbb{R}$. Wir definieren r durch:

$$f(x) = f(p) + 2p(x-p) + r(x) \quad \text{für alle } x \in \mathbb{R},$$

d. h., es gilt $r(x) = x^2 - p^2 - 2p(x-p)$ für alle $x \in \mathbb{R}$. Dann gilt

$$\begin{aligned} \lim_{x \rightarrow p} \frac{r(x)}{x-p} &= \lim_{x \rightarrow p} \left(\frac{x^2 - p^2}{x-p} - 2p \right) = \\ \lim_{x \rightarrow p} (x+p-2p) &= \lim_{x \rightarrow p} (x-p) = 0. \end{aligned}$$

Also gilt $f'(p) = 2p$.

Die Darstellung in (c) ist lediglich eine Umformulierung der Darstellung in (b). Ist

$$f(x) = f(p) + a(x-p) + r(x) \quad \text{für alle } x \in P, \quad \text{so gilt}$$

$$f(x) = f(p) + s(x)(x-p) \quad \text{für alle } x \in P, x \neq p, \quad \text{wenn wir}$$

$$s(x) = a + r(x)/(x-p) \quad \text{für alle } x \in P - \{p\}$$

setzen. Die Restbedingung „ $\lim_{x \rightarrow p} r(x)/(x-p) = 0$ “ ist äquivalent zur stetigen Fortsetzbarkeit von $s: P - \{p\} \rightarrow \mathbb{R}$ im Punkt p durch die Setzung „ $s(p) = a$ “.

Für die Funktion $s: P \rightarrow \mathbb{R}$ gilt

$$s(x) = \frac{f(x) - f(p)}{x-p} \quad \text{für alle } x \in P, x \neq p,$$

d. h., die Funktion s gibt auf $P - \{p\}$ die Differenzenquotienten von f in den Punkten p und x in der Variablen x wieder. f ist genau dann in p differenzierbar, wenn diese Differenzenquotientenfunktion im Punkt p stetig fortsetzbar ist.

Beispiele

- (1) Sei wieder $f: \mathbb{R} \rightarrow \mathbb{R}$ die zweite Potenz, und sei $p \in \mathbb{R}$. Dann gilt

$$f(x) = x^2 = p^2 + x^2 - p^2 = p^2 + (x+p)(x-p) = f(p) + s(x)(x-p)$$

mit $s(x) = x+p$ für alle $x \in \mathbb{R}$. $s: \mathbb{R} \rightarrow \mathbb{R}$ ist stetig in p , also gilt $f'(p) = s(p) = 2p$.

- (2) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ ein Polynom, und sei $f(p) = 0$. Dann gibt es ein Polynom s mit $f(x) = s(x)(x-p)$ für alle $x \in \mathbb{R}$ (Abspalten von Nullstellen). Dann gilt aber

$$f(x) = 0 + s(x)(x-p) = f(p) + s(x)(x-p) \quad \text{für alle } x \in \mathbb{R}.$$

Wegen der Stetigkeit von s gilt also $f'(p) = s(p)$.

- (3) Sei $s: \mathbb{R} \rightarrow \mathbb{R}$ stetig im Nullpunkt, und sei $f: \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$f(x) = x \cdot s(x) \quad \text{für alle } x \in \mathbb{R}.$$

Dann gilt $f(x) = f(0) + s(x)(x-0)$ für alle x . Also gilt $f'(0) = s(0)$.

7.4 Ableitungsregeln

Satz (Ableitungsregeln)

Unter der Voraussetzung der Existenz der linken Seite gilt:

$$(a) \quad (af + bg)'(p) = a f'(p) + b g'(p), \quad (\text{Linearität der Ableitung})$$

$$(b) \quad (fg)'(p) = f'(p) g(p) + f(p) g'(p), \quad (\text{Produktregel})$$

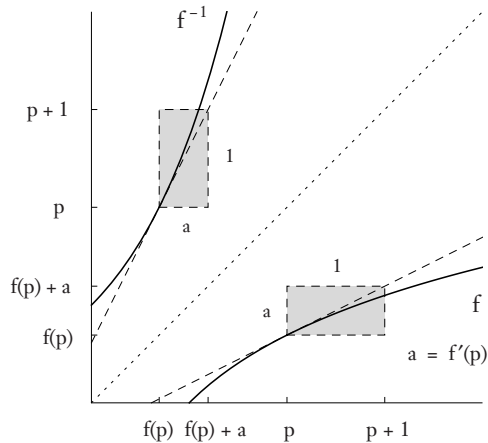
$$(c) \quad (f/g)'(p) = \frac{f'(p) g(p) - g'(p) f(p)}{g^2(p)}, \quad (\text{Quotientenregel})$$

$$(d) \quad (g \circ f)'(p) = g'(f(p)) \cdot f'(p), \quad (\text{Kettenregel})$$

$$(e) \quad (f^{-1})'(f(p)) = \frac{1}{f'(p)}. \quad (\text{Ableitung der Umkehrfunktion})$$

Die Ableitungsregeln bilden einen Kalkül für die Differentialrechnung. Die Berechnung von Differentialquotienten tritt in den Hintergrund, da die Ableitung durch Anwendung der Regeln aus wenigen bekannten elementaren Ableitungen gewonnen werden kann.

Wir haben die Regeln im Satz sehr kompakt angegeben. Ausführlich formuliert lautet die Kettenregel zum Beispiel:



Genaue Formulierung der Kettenregel

Seien $f: P \rightarrow \mathbb{R}$ und $g: Q \rightarrow \mathbb{R}$ Funktionen derart, dass die Verknüpfung $h = g \circ f$ existiert, d. h., für alle $x \in P$ ist $f(x) \in Q$. Weiter sei $p \in P$, und f sei differenzierbar in p und g sei differenzierbar in $f(p) \in Q$. Dann ist h differenzierbar in p und es gilt

$$h'(p) = g'(f(p)) \cdot f'(p).$$

Dass $g'(f(p))$ mit $f'(p)$ multipliziert werden muss, wird als *Nachdifferenzieren* bezeichnet. Dies wird gerne vergessen. Eine weitere Fehlerquelle ist, die Ableitung von g im Punkt p und nicht im Punkt $f(p)$ zu berechnen.

Bei Verwendung der graphischen Methode wird die Regel für die Ableitung der Umkehrfunktion f^{-1} sehr anschaulich. Hat f im Punkt $(p, f(p))$ die Steigung a , so hat die Umkehrfunktion von f im Punkt $(f(p), p)$ die Steigung $1/a$. Am klarsten sieht man dies vielleicht, wenn man die Spiegelung an der Winkelhalbierenden gar nicht durchführt, sondern den Graphen von f um 90 Grad gegen den Uhrzeigersinn dreht und die y -Achse (die neue x -Achse) von rechts nach links liest.

Beispiele

- (1) Für jedes Polynom
- $f: \mathbb{R} \rightarrow \mathbb{R}$
- ,
- $f(x) = \sum_{k \leq n} a_k x^k$
- , gilt

$$f'(p) = \sum_{1 \leq k \leq n} k a_k p^{k-1}.$$

- (2) Für den Spezialfall
- $f = g$
- liefert die Produktregel

$$(f^2)'(p) = (f \cdot f)'(p) = 2 f(p) f'(p).$$

Allgemein zeigt eine Induktion nach n , dass $(f^n)'(p) = n f(p)^{n-1} f'(p)$.

- (3) Auch die Kettenregel liefert die Formel in (2). Denn ist
- g
- die
- n
- te Potenz, so gilt
- $f^n = g \circ f$
- und
- $g'(f(p)) = n f(p)^{n-1}$
- . Also gilt nach der Kettenregel

$$(f^n)'(p) = (g \circ f)'(p) = g'(f(p)) \cdot f'(p) = n f(p)^{n-1} f'(p).$$

- (4) Für die Identität
- $\text{id}: \mathbb{R} \rightarrow \mathbb{R}$
- ,
- $\text{id}(x) = x$
- , gilt
- $\text{id}' = 1$
- . Nach (2) und (3) gilt also

$$(\text{id}^n)'(p) = n \text{id}(p)^{n-1} \text{id}'(p) = n p^{n-1} \quad \text{für alle } p \in \mathbb{R}.$$

Aber id^n ist die n -te Potenz auf \mathbb{R} und damit haben wir die Ableitungsregel für diese Funktion wiedergefunden.

- (5) Es gilt

$$d(x^3 + 2x^2 - 10)^2/dx(p) = 2(p^3 + 2p^2 - 10) \cdot (3p^2 + 4p),$$

$$de^{ax}/dx(p) = e^{ap} \cdot d(ax)/dx(p) = a e^{ap},$$

$$de^{x^2}/dx(p) = e^{p^2} \cdot dx^2/dx(p) = 2p e^{p^2},$$

$$da^x/dx(p) = d(e^{x \log(a)})/dx(p) = e^{p \log(a)} \cdot \log(a) = \log(a) a^p \quad \text{für } a > 0.$$

- (6) Die Wurzelfunktion
- $g:]0, +\infty[\rightarrow \mathbb{R}$
- ist die Umkehrfunktion der zweiten Potenz
- f
- auf
- $]0, +\infty[$
- . Es gilt
- $f'(p) \neq 0$
- für alle
- p
- . Also gilt für alle
- $q = f(p) = p^2$
- :

$$g'(q) = \frac{1}{f'(p)} = \frac{1}{2p} = \frac{1}{2\sqrt{q}}.$$

- (7)
- $\log:]0, +\infty[\rightarrow \mathbb{R}$
- ist die Umkehrfunktion von
- $\exp: \mathbb{R} \rightarrow \mathbb{R}$
- und es gilt
- $\exp'(p) = \exp(p) \neq 0$
- für alle
- p
- . Also gilt für alle
- $q = \exp(p) > 0$
- :

$$\log'(q) = \frac{1}{\exp'(p)} = \frac{1}{\exp(p)} = \frac{1}{\exp(\log(q))} = \frac{1}{q}.$$

- (8) Für alle
- $a \in \mathbb{R}$
- und
- $q > 0$
- gilt nach (7) und der Kettenregel:

$$\frac{dx^a}{dx}(q) = \frac{d(e^{a \log(x)})}{dx}(q) = e^{a \log(q)} \cdot \frac{d(a \log(x))}{dx}(q) = q^a \cdot \frac{a}{q} = a q^{a-1}.$$

Für $a = 1/2$ ergibt sich die Formel für die Ableitung der Quadratwurzel aus (6).

7.5 Differenzierbarkeit und Stetigkeit

Satz (Differenzierbarkeit impliziert Stetigkeit)

Sei $f: P \rightarrow \mathbb{R}$ differenzierbar in p .

Dann ist f stetig in p .

Der Satz lässt sich mit Hilfe des Approximationssatzes 7.3 leicht einsehen. Ist nämlich $f: P \rightarrow \mathbb{R}$ differenzierbar im Punkt p , so gilt

$$f(x) = f(p) + a(x - p) + r(x)$$

für alle $x \in P$, wobei $a = f'(p)$ und

$$\lim_{x \rightarrow p} r(x) = 0.$$

(Da sogar $\lim_{x \rightarrow p} r(x)/(x - p) = 0$.) Folglich ist f stetig in p , da

$$\lim_{x \rightarrow p} f(x) = \lim_{x \rightarrow p} (f(p) + a f'(p) + r(x)) = f(p) + a \cdot 0 + 0 = f(p).$$

Damit gilt:

Jede differenzierbare Funktion ist stetig.

Die Umkehrung ist, wie wir schon gesehen hatten, nicht richtig. Hier noch einmal einige Beispiele:

Beispiele

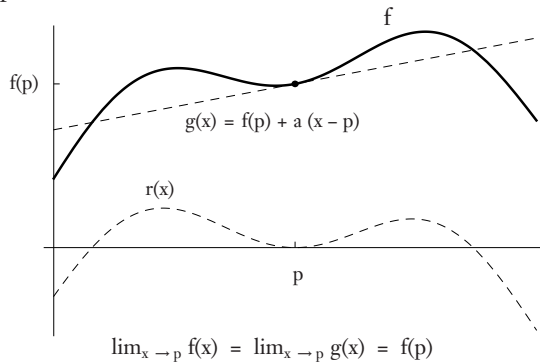
- (1) Die Betragsfunktion $\text{abs}: \mathbb{R} \rightarrow \mathbb{R}$, $\text{abs}(x) = |x|$ für alle x , ist stetig, aber im Nullpunkt nicht differenzierbar.
- (2) Die Wurzelfunktion $g: [0, +\infty[\rightarrow \mathbb{R}$ mit $g(x) = \sqrt{x}$ für alle $x \geq 0$ ist stetig, aber im Nullpunkt nicht differenzierbar. Das Gleiche gilt für die auf ganz \mathbb{R} definierte dritte Wurzel $h: \mathbb{R} \rightarrow \mathbb{R}$ mit

$$h(x) = \sqrt[3]{x} = \text{„das eindeutige } y \text{ mit } y^3 = x\text{“ für alle } x \in \mathbb{R}.$$

- (3) Sei $f: [-1, 1] \rightarrow \mathbb{R}$ definiert durch $f(x) = \sqrt{1 - x^2}$ für alle $x \in [-1, 1]$. Dann ist der Graph von f die obere Hälfte des Einheitskreises. Die Funktion f ist stetig, aber in den Punkten -1 und 1 nicht differenzierbar.

Im ersten Beispiel liegt ein „Knick“ vor, der verhindert, dass wir der Funktion eine lokale Steigung zuweisen können. Die Funktionen in den anderen Beispielen besitzen dagegen geometrische Tangenten an ihren Nullstellen, aber diese sind Senkrechte und damit existiert die Ableitung nach unserer Definition nicht.

Der Satz und die Beispiele lassen viele Fragen über den genaueren Zusammenhang der beiden Grundbegriffe der Differenzierbarkeit und der Stetigkeit offen. Wir wollen einige davon diskutieren. Zunächst fragen wir:



Ist die Ableitung f' einer differenzierbaren Funktion f stets wieder differenzierbar?

Die Antwort ist: Nein.

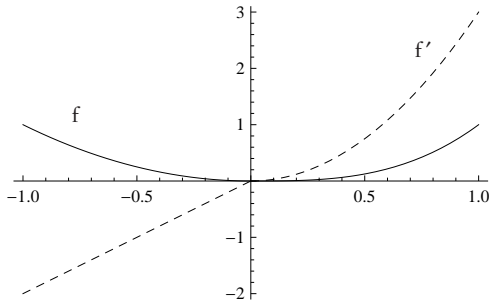
Beispiel

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2$ für $x < 0$ und $f(x) = x^3$ für $x \geq 0$. Dann gilt

$$f'(x) = \begin{cases} 2x, & \text{falls } x < 0, \\ 3x^2, & \text{falls } x \geq 0. \end{cases}$$

f' ist im Punkt 0 nicht differenzierbar.

Analoges gilt auch für $g: \mathbb{R} \rightarrow \mathbb{R}$ mit $g(x) = x|x|$ für alle x .



Bescheidener fragen wir:

Ist die Ableitung f' einer differenzierbaren Funktion f wenigstens immer stetig?

Die Antwort ist: Nein.

Beispiel

Die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2 \sin(1/x)$ für $x \neq 0$ und $f(0) = 0$ ist differenzierbar mit

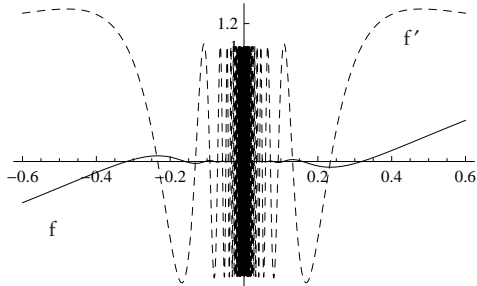
$$f'(p) = 2x \sin(1/x) - \cos(1/x)$$

für alle $x \neq 0$, und

$$f'(0) = \lim_{h \rightarrow 0} f(h)/h =$$

$$\lim_{h \rightarrow 0} h \sin(1/h) = 0.$$

f' hat also im Nullpunkt eine Unstetigkeitsstelle zweiter Art.



Man kann zeigen, dass eine Unstetigkeit der Ableitung immer zweiter Art ist, d. h., die Ableitung f' einer differenzierbaren Funktion kann keine einfachen Sprünge machen!

Schließlich fragen wir noch:

Ist eine stetige Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ wenigstens immer in einem Punkt differenzierbar?

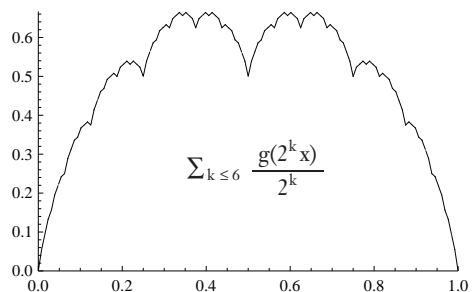
Auch dies ist zu verneinen. Erste Gegenbeispiele wurden von Bolzano und Weierstraß konstruiert, die folgende Funktion stammt von Takagi.

Beispiel

Sei $g: \mathbb{R} \rightarrow \mathbb{R}$ definiert durch $g(a + x) = |x|$ für alle $a \in \mathbb{Z}$ und $x \in [-1/2, 1/2[$. Dann definiert

$$f(x) = \sum_n g(2^n x)/2^n \text{ für alle } x \in \mathbb{R}$$

eine stetige Funktion auf \mathbb{R} , die in keinem Punkt differenzierbar ist.



7.6 Höhere Ableitungen

Definition (*höhere Ableitungen, n -fache stetige Differenzierbarkeit, glatt*)

Sei $f: P \rightarrow \mathbb{R}$ eine Funktion. Wir setzen $f^{(0)} = f$ und definieren rekursiv, solange die Ableitungen existieren,

$$f^{(n+1)} = f^{(n)'} \quad \text{für alle } n.$$

Die Funktion $f^{(n)}$ heißt, im Fall der Existenz, die n -te *Ableitung* von f oder die *Ableitung der Ordnung* n von f . Wir schreiben auch

$$D^n f = d^n f / dx^n = (d/dx)^n f = f^{(n)}.$$

Die Funktion f heißt

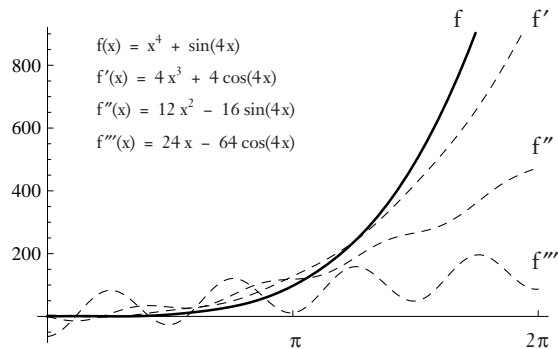
- (a) n -mal differenzierbar, falls $f^{(n)}$ existiert,
- (b) n -mal stetig differenzierbar oder eine \mathcal{C}^n -Funktion, falls $f^{(n)}$ stetig ist,
- (c) glatt oder eine \mathcal{C}^∞ -Funktion, falls $f^{(n)}$ für alle n existiert.

Es gilt also

$$f = f^{(0)}, \quad f' = f^{(1)}, \quad (f')' = f^{(2)}, \quad \dots$$

Für $n = 2$ ist auch f'' zur Bezeichnung von $f^{(2)}$ üblich. Die Funktion $f^{(0)}$ ist immer definiert und die Aussage, dass f eine \mathcal{C}^0 -Funktion ist, besagt einfach, dass f stetig ist. Die Funktion $f^{(1)}$ ist definiert, falls f differenzierbar ist, und f ist \mathcal{C}^1 , falls die Ableitung $f^{(1)}$ stetig ist.

Existiert $f^{(n+1)}$, so ist f eine \mathcal{C}^n -Funktion, da die Differenzierbarkeit von $f^{(n)}$ die Stetigkeit von $f^{(n)}$ impliziert. Wir betrachten einige Beispiele und ziehen dabei die Funktionen aus der letzten Sektion heran.



Beispiele

- (1) Die Betragsfunktion $\text{abs}: \mathbb{R} \rightarrow \mathbb{R}$ ist eine \mathcal{C}^0 -Funktion, aber $\text{abs}^{(1)}$ existiert nicht.
- (2) Für $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2 \sin(1/x)$ für $x \neq 0$ und $f(0) = 0$ existiert die Ableitung $f^{(1)}$, aber f ist keine \mathcal{C}^1 -Funktion.
- (3) Die Funktion $g: \mathbb{R} \rightarrow \mathbb{R}$ mit $g(x) = x^2$ für $x < 0$ und $g(x) = x^3$ für $x \geq 0$ ist \mathcal{C}^1 , aber $g^{(2)}$ existiert nicht, da g' einen Knick im Nullpunkt besitzt.
- (4) Auch die Funktion $h: [0, +\infty[\rightarrow \mathbb{R}$ mit $h(x) = x^{3/2}$ für alle $x \geq 0$ ist \mathcal{C}^1 , aber $h^{(2)}$ existiert nicht, da $h'(x) = 3/2 \sqrt{x}$ für alle $x \geq 0$ und also h' im Nullpunkt nicht differenzierbar ist.

Allgemeiner kann man für jedes n Funktionen f und g konstruieren, sodass gilt:

- (a) $f^{(n)}$ existiert, aber f ist keine \mathcal{C}^n -Funktion.
- (b) g ist eine \mathcal{C}^n -Funktion, aber $g^{(n+1)}$ existiert nicht.

Die erste und zweite Ableitung enthält viele Informationen über f . Wir werden sehen, dass wir mit ihrer Hilfe die lokalen Extremwerte und das Krümmungsverhalten von f untersuchen können (7.8 – 7.11). Die mehrfachen Ableitungen $f^{(n)}$ treten im Satz von Taylor und bei der Aufgabe, eine Funktion als Potenzreihe darzustellen, auf (7.12).

Glatte Funktionen lassen sich beliebig oft differenzieren. Sie haben keine „unterdrückten Oszillationen“ wie die Funktion in Beispiel (2), die in einer Ableitung irgendwann als Unstetigkeit zweiter Art erscheinen würden. Sie haben keine „verborgenen Knicke“ wie die Funktion in Beispiel (3), die beim wiederholten Differenzieren sichtbar werden würden. Und ihre Ableitungen zeigen keine geometrischen Tangenten der Steigung $+\infty$ wie die Funktion in Beispiel (4).

Beispiele

- (1) Jede Polynomfunktion f ist glatt. Wiederholtes Ableiten endet in der Nullfunktion. Ist f nicht die Nullfunktion (deren Grad gleich $-\infty$ ist), so gilt:

$$\text{Grad}(f) = \text{„das kleinste } n \text{ mit } f^{(n+1)} = 0\text{“}.$$

- (2) Für alle $n \geq 1$ und alle $1 \leq k \leq n$ gilt

$$\left(\frac{d}{dx}\right)^k x^n = n(n-1) \dots (n-k+1) x^{n-k} = \frac{n!}{(n-k)!} x^{n-k}.$$

$$\text{Speziell ist also } (d/dx)^n x^n = n!/0! x^0 = n!$$

- (3) Die Exponentialfunktion ist glatt. Es gilt $\exp^{(n)} = \exp$ für alle n .

- (4) Die Logarithmusfunktion $\log :]0, +\infty[\rightarrow \mathbb{R}$ ist glatt. Es gilt

$$\log^{(1)}(x) = x^{-1}, \quad \log^{(2)}(x) = -x^{-2}, \quad \log^{(3)}(x) = 2x^{-3}, \quad \log^{(4)}(x) = -6x^{-4},$$

allgemein

$$\log^{(n)}(x) = \frac{(-1)^{n-1} (n-1)!}{x^n} \quad \text{für alle } n \geq 1 \text{ und } x > 0.$$

- (5) Der Sinus ist glatt. Es gilt

$$\sin^{(2n)} = (-1)^n \sin, \quad \sin^{(2n+1)} = (-1)^n \cos \quad \text{für alle } n.$$

Die Folge der Ableitungen ist periodisch. Analoges gilt für den Kosinus.

- (6) Eine Potenzreihe $f = \sum_n a_n x^n$ mit Konvergenzradius R ist eine glatte Funktion auf $] -R, R[$. Für alle $x \in] -R, R[$ und alle $k \geq 0$ gilt:

$$f^{(k)}(x) = \sum_n a_n (d/dx)^k x^n = \sum_{n \geq k} a_n n(n-1) \dots (n-k+1) x^{n-k}.$$

(gliedweises Differenzieren)

7.7 Der Mittelwertsatz der Differentialrechnung

Satz (Mittelwertsatz der Differentialrechnung von Lagrange)

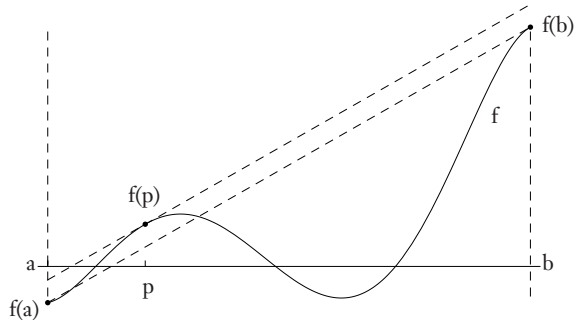
Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig und in allen $x \in]a, b[$ differenzierbar. Dann gibt es ein $p \in]a, b[$ mit

$$f'(p) = \frac{f(b) - f(a)}{b - a}.$$

Die Zahl $m = (f(b) - f(a))/(b - a)$ ist der Differenzenquotient von f für die Punkte a und b . Wir können sie als die mittlere Steigung, die f auf dem Weg von $f(a)$ nach $f(b)$ besitzt, auffassen. Der Mittelwertsatz besagt, dass ein Punkt im Inneren des Intervalls $[a, b]$ existiert,

in dem die lokale Steigung von f genau die globale mittlere Steigung von f ist.

Ist $f(a) = f(b)$, so gibt es ein p mit $f'(p) = 0$. Dieser Spezialfall ist als *Satz von Rolle* bekannt (benannt nach dem französischen Mathematiker Michel Rolle).



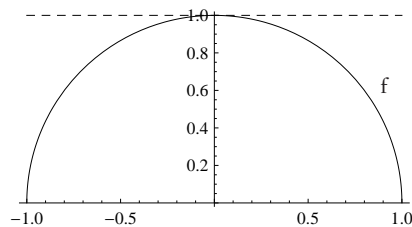
Beispiele

- (1) Sei $f: [-1, 1] \rightarrow \mathbb{R}$ mit

$$f(x) = \sqrt{1 - x^2} \text{ für alle } x \in [-1, 1].$$

Dann ist f stetig und in $] -1, 1 [$ differenzierbar. Nach dem Mittelwertsatz gibt es ein $p \in] -1, 1 [$ (nämlich $p = 0$) mit

$$f'(p) = \frac{f(-1) - f(1)}{1 - (-1)} = 0.$$



Das Beispiel illustriert die Bedingung „stetig in allen $p \in [a, b]$ und differenzierbar in allen $p \in]a, b[$ “. Wir würden etwas verlieren, wenn wir uns auf in $[a, b]$ differenzierbare Funktionen beschränken würden.

- (2) Auf die Stetigkeit in den Randpunkten kann nicht verzichtet werden. Denn sei $f: [0, 1] \rightarrow \mathbb{R}$ mit $f(x) = x$ für alle $x \in [0, 1[$ und $f(1) = 0$. Für diese Funktion gilt $(f(1) - f(0))/(1 - 0) = 0 \neq 1 = f'(x)$ für alle $x \in]0, 1[$.
- (3) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare nichtmonotone Funktion. Dann besitzt f' eine Nullstelle. Denn da f stetig, aber nicht monoton ist, gibt es $c < d$ in $[a, b]$ mit $f(c) = f(d)$ (vgl. die Beispiele zum Zwischenwertsatz). Nach dem Mittelwertsatz existiert also ein $p \in]c, d[$ mit

$$f'(p) = \frac{f(d) - f(c)}{d - c} = \frac{0}{d - c} = 0.$$

Der Mittelwertsatz der Differentialrechnung hat, wie der Zwischenwertsatz für stetige Funktionen, zahlreiche Anwendungen. Die Sätze, die man mit seiner Hilfe beweisen kann, umfassen den Satz über den Zusammenhang zwischen der Ableitung einer Funktion und ihrem Monotonieverhalten (vgl. die nächste Sektion), den Satz von Taylor (7.12) und den Hauptsatz der Differential- und Integralrechnung (8.8). Wir diskutieren hier noch zwei einfache, aber deswegen nicht weniger wichtige Anwendungen.

Anwendung 1: Bestimmung aller Funktionen f mit $f' = f$

Aus dem Mittelwertsatz folgt zunächst:

Ist I ein Intervall und $g : I \rightarrow \mathbb{R}$ differenzierbar mit $g' = 0$, so ist g konstant.

Denn für alle $a < b$ in I existiert ein p in $]a, b[$ mit

$$g'(p) = (g(b) - g(a))/(b - a).$$

Wegen $g'(p) = 0$ gilt also $g(a) = g(b)$. Da dies für beliebige $a < b$ in einem Intervall I gilt, kann also g nur einen einzigen Wert annehmen, d. h., g ist konstant. Nun folgt:

Gilt $f' = f$ für eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, so gibt es ein c mit $f = c \exp$.

Denn sei $g : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$g(x) = f(x) \cdot e^{-x} \quad \text{für alle } x \in \mathbb{R}.$$

Dann gilt nach der Produktregel

$$g'(x) = f'(x) \cdot e^{-x} - f(x) e^{-x} = f(x) \cdot e^{-x} - f(x) e^{-x} = 0 \quad \text{für alle } x \in \mathbb{R}.$$

Also gibt es ein c mit $g(x) = f(x) e^{-x} = c$ für alle $x \in \mathbb{R}$, d. h., es gilt

$$f(x) = c e^x \quad \text{für alle } x \in \mathbb{R}.$$

Speziell ist $\exp : \mathbb{R} \rightarrow \mathbb{R}$ durch „ $f' = f$ und $f(0) = 1$ “ eindeutig bestimmt.

Anwendung 2: Lipschitz-Stetigkeit ist gar nicht so selten

Wir hatten in 7.5 gesehen, dass die Differenzierbarkeit die Stetigkeit nach sich zieht. Aus dem Mittelwertsatz folgt nun, dass differenzierbare Funktionen in vielen Fällen sogar die starke Form der Lipschitz-Stetigkeit erfüllen. Sei nämlich $f : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar. Da f' stetig ist und stetige Funktionen auf $[a, b]$ beschränkt sind, gibt es ein $L \geq 0$ mit $|f'| \leq L$. Dann gilt für alle $x < y$ in $[a, b]$, dass

$$\left| \frac{f(x) - f(y)}{x - y} \right| \leq L,$$

da die linke Seite nach dem Mittelwertsatz ein Wert der Ableitung ist. Dies zeigt, dass

$$|f(x) - f(y)| \leq L |x - y| \quad \text{für alle } x, y \in [a, b],$$

d. h., f ist Lipschitz-stetig mit der Lipschitz-Konstanten L . Wir haben also gezeigt:

$\mathcal{C}^{(1)}$ -Funktionen auf kompakten Intervallen sind Lipschitz-stetig.

7.8 Ableitung und Monotonie

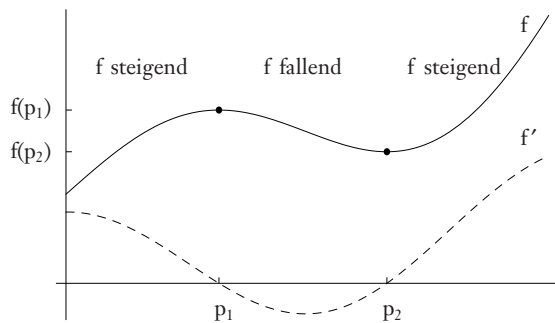
Satz (Ableitung und Monotonie)

Sei I ein Intervall und sei $f : I \rightarrow \mathbb{R}$ differenzierbar. Dann gilt:

- (a) $f' \geq 0$ genau dann, wenn f ist monoton steigend.
- (b) $f' \leq 0$ genau dann, wenn f ist monoton fallend.
- (c) $f' > 0$ impliziert f ist streng monoton steigend.
- (d) $f' < 0$ impliziert f ist streng monoton fallend.

Der Satz ist ein Paradebeispiel dafür, wie man mit Hilfe der Ableitung Informationen über die Funktion selbst gewinnen kann. Kennt man das Vorzeichenverhalten der Ableitung, so kennt man das Monotonieverhalten der Funktion.

Das Intervall I des Satzes kann offen, halboffen, abgeschlossen sein, und auch unbeschränkte Intervalle wie $]0, +\infty[$, $] -\infty, 0]$ oder $] -\infty, +\infty[$ sind zugelassen. Die Bedingungen an f sind punktweise zu lesen. „ $f' \leq 0$ “ bedeutet zum Beispiel, dass $f'(x) \leq 0$ für alle $x \in I$.



An den Stellen p_1 und p_2 ändert f die Monotonie und f' das Vorzeichen.

Beispiele

- (1) Es gilt $\exp'(x) = \exp(x) > 0$ für alle $x \in \mathbb{R}$. Also ist \exp streng monoton steigend. Weiter gilt $\operatorname{arccot}'(x) = -1/(1+x^2) < 0$ für alle $x \in \mathbb{R}$. Also ist arccot streng monoton fallend.
- (2) Für die zweite Potenz $f : \mathbb{R} \rightarrow \mathbb{R}$ gilt $f'(x) = 2x$. Nach dem Satz ist also $f|_{]-\infty, 0]}$ streng monoton fallend und $f|_{]0, +\infty[}$ streng monoton steigend.
- (3) Die dritte Potenz f auf \mathbb{R} ist streng monoton steigend mit $f'(x) = 3x^2 \geq 0$ für alle x . Das Beispiel zeigt, dass die umgekehrte Implikation in (c) im Allgemeinen nicht gilt. Analoges gilt für die Aussage (d), wie die Funktion $g = -f$ zeigt.
- (4) Sei $f : [0, 1] \cup [2, 3]$ definiert durch $f(x) = x$ für alle $x \in [0, 1]$ und $f(x) = x - 2$ für alle $x \in [2, 3]$. Dann ist f differenzierbar und es gilt $f'(x) = 1 > 0$ für alle x . Die Funktion ist aber nicht monoton steigend, da zum Beispiel $f(1) = 1 > 0 = f(2)$. Die Voraussetzung, dass f auf einem Intervall definiert ist, kann also nicht fallengelassen werden.

Im dritten und vierten Teil des Satzes erhalten wir, wie die Beispiele zeigen, keine Äquivalenzen. Die Bedingung „ $f' > 0$ “ bzw. „ $f' < 0$ “ ist stärker als die strenge Monotonie.

Es stellt sich die Frage, ob wir die Bedingung an die Ableitung geeignet abschwächen können, sodass Äquivalenzen entstehen und einfache Funktionen wie die dritte Potenz über ihre Ableitung als streng monoton steigend erkannt werden können. Dies ist in der Tat in einfacher Weise möglich. Wir führen hierzu für eine Funktion $f: I \rightarrow \mathbb{R}$ zwei (nur hier verwendete) Hilfsrelationen \leq^* und \geq^* ein:

$f \geq^* 0$, falls $f \geq 0$, und es gibt kein Teilintervall $[c, d]$ mit $c < d$ von I , sodass $f(x) = 0$ für alle $x \in [c, d]$ ist.

$f \leq^* 0$, falls $f \leq 0$, und es gibt kein Teilintervall $[c, d]$ mit $c < d$ von I , sodass $f(x) = 0$ für alle $x \in [c, d]$ ist.

Anschaulich gilt $f \geq^* 0$, wenn $f \geq 0$ ist und sich allenfalls punktweise auf der x -Achse aufhält. Analoges gilt für $f \leq^* 0$. Der Stern bedeutet also, dass die Funktion Nullstellen haben darf, aber keine Nullintervalle. Alle elementaren Funktionen mit Ausnahme der Nullfunktion haben diese Eigenschaft.

Die Relation „ $f \geq^* 0$ “ liegt im logischen Sinne zwischen „ $f > 0$ “ und „ $f \geq 0$ “:

$$f > 0 \text{ impliziert } f \geq^* 0, \quad f \geq^* 0 \text{ impliziert } f \geq 0.$$

Analoges gilt für „ $f \leq^* 0$ “. Nach diesen Vorbereitungen können wir nun den obigen Satz verbessern:

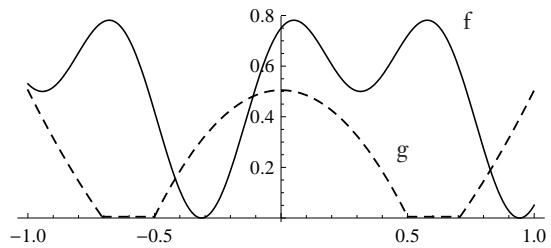
Satz (Ableitung und Monotonie, verbesserte Version)

Sei I ein Intervall und sei $f: I \rightarrow \mathbb{R}$ differenzierbar. Dann gilt:

- (a) $f' \geq 0$ genau dann, wenn f ist monoton steigend.
- (b) $f' \leq 0$ genau dann, wenn f ist monoton fallend.
- (c) $f' \geq^* 0$ genau dann, wenn f ist streng monoton steigend.
- (d) $f' \leq^* 0$ genau dann, wenn f ist streng monoton fallend.

Beispiele

- (1) Sei $n = 2m + 1$ ungerade, und sei $f: \mathbb{R} \rightarrow \mathbb{R}$ die n -te Potenz. Dann gilt $f'(x) = nx^{2m} \geq 0$ für alle x . Weiter gilt $f' \geq^* 0$, da f' nur im Nullpunkt gleich 0 ist. Also ist f streng monoton steigend.
- (2) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ konstant gleich c . Dann gilt $f'(x) = 0$ für alle x , aber es gilt weder $f' \leq^* 0$ noch $f' \geq^* 0$. f ist monoton (fallend und steigend), aber nicht streng monoton.



$f \geq 0$, $f \geq^* 0$, $g \geq 0$, nicht($g \geq^* 0$)

7.9 Lokale Extremwerte

Satz (Bedingungen für lokale Extrema)

Seien I ein offenes Intervall, $f: I \rightarrow \mathbb{R}$ differenzierbar und $p \in I$. Dann gilt (mit den Definitionen der Tabelle unten):

(a) *Notwendige Bedingung für ein lokales Extremum*

Besitzt f in p ein lokales Extremum, so gilt $f'(p) = 0$.

(b) *Hinreichende Bedingung für ein lokales Maximum*

Ist $]p - \varepsilon, p + \varepsilon[$ ein Teilintervall von I mit

$$f' \geq 0 \text{ in }]p - \varepsilon, p[\text{ und } f' \leq 0 \text{ in }]p, p + \varepsilon[,$$

so besitzt f in p ein lokales Maximum.

(c) *Hinreichende Bedingung für ein lokales Minimum*

Ist $]p - \varepsilon, p + \varepsilon[$ ein Teilintervall von I mit

$$f' \leq 0 \text{ in }]p - \varepsilon, p[\text{ und } f' \geq 0 \text{ in }]p, p + \varepsilon[,$$

so besitzt f in p ein lokales Minimum.

Gilt in (b) und (c) statt \leq und \geq stärker $<$ und $>$ (oder auch nur \leq^* und \geq^* wie in der letzten Sektion), so ist das lokale Extremum strikt.

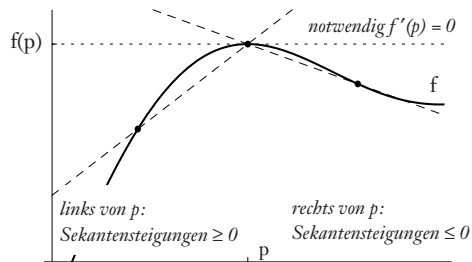


Diagramm zu (a): f besitzt in p ein lokales Maximum.

$f: P \rightarrow \mathbb{R}$ besitzt in $p \in P$ ein ...	falls gilt ...
lokales Maximum	$\exists \varepsilon > 0 \forall x \in U_\varepsilon(p) \cap P \ f(x) \leq f(p)$
lokales Minimum	$\exists \varepsilon > 0 \forall x \in U_\varepsilon(p) \cap P \ f(x) \geq f(p)$
striktes lokales Minimum	$\exists \varepsilon > 0 \forall x \in U_\varepsilon(p) \ (x \neq p \rightarrow f(x) > f(p))$
striktes lokales Maximum	$\exists \varepsilon > 0 \forall x \in U_\varepsilon(p) \ (x \neq p \rightarrow f(x) < f(p))$
(striktes) lokales Extremum	f besitzt in p ein (striktes) lokales Maximum oder ein (striktes) lokales Minimum

Hinweis: In 7.11 diskutieren wir eine weitere hinreichende Bedingung für lokale Extremwerte.

Der erste Teil des Satzes gibt eine anschauliche notwendige Bedingung für ein lokales Extremum an: Die Tangente einer differenzierbaren Funktion, die in p ein lokales Maximum oder Minimum besitzt, muss dort die Steigung 0 haben.

Beispiele

- (1) Sei f die dritte Potenz auf \mathbb{R} . Dann gilt $f'(0) = 0$, aber f besitzt kein lokales Extremum im Punkt 0. Die Bedingung (a) ist also nicht hinreichend.
- (2) Sei $f: [0, 1] \rightarrow \mathbb{R}$ definiert durch $f(x) = x$ für alle $x \in [0, 1]$. Dann ist f differenzierbar und es gilt $f'(x) = 1 \neq 0$ für alle $x \in [0, 1]$. f besitzt aber in 0 ein lokales Minimum und in 1 ein lokales Maximum. Die Voraussetzung, dass f auf einem offenen Intervall definiert ist, ist also wesentlich.

Die hinreichenden Bedingungen des Satzes fließen aus dem Satz über Ableitung und Monotonie (7.8). Denn ist $f' \geq 0$ in $]p - \varepsilon, p[$ und $f' \leq 0$ in $]p, p + \varepsilon[$, so ist f monoton steigend im linken und monoton fallend im rechten Intervall, und damit besitzt f als stetige Funktion ein lokales Maximum im Punkt p . Analog zeigt man die anderen Behauptungen. Insbesondere gilt die Merkregel:

Ein Vorzeichenwechsel der Ableitung zeigt ein striktes lokales Extremum an.

Auf den ersten Blick sieht es vielleicht so aus, als könnte ein lokales Extremum nur durch die Bedingungen in (b) und (c) entstehen. Dies ist aber nicht der Fall. Eine Funktion kann in p eine lokale Minimalstelle besitzen, aber in keinem Intervall $]p - \varepsilon, p[$ monoton fallen und in keinem Intervall $]p, p + \varepsilon[$ monoton steigen:

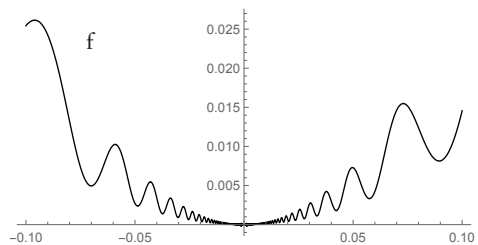
Beispiele

- (1) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2(2 + \sin(1/x))$ für $x \neq 0$ und $f(0) = 0$.

Wegen $-1 \leq \sin(1/x) \leq 1$ ist $f(x) > 0$ für $x \neq 0$ und damit ist 0 ein striktes (globales) Minimum von f . Es gilt

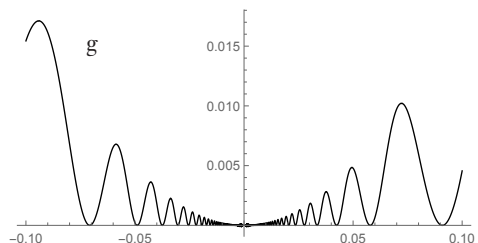
$$f'(x) = 2x(2 + \sin(1/x)) - \cos(1/x)$$

für $x \neq 0$ und $f'(0) = 0$. f' nimmt in allen Intervallen $] - \varepsilon, 0[$ und $] 0, \varepsilon[$ sowohl positive als auch negative Werte an und verletzt damit (c).



- (2) Sei $g: \mathbb{R} \rightarrow \mathbb{R}$ mit $g(x) = x^2(1 + \sin(1/x))$ für $x \neq 0$ und $g(0) = 0$.

Dann ist g wie in (1), aber das Minimum im Nullpunkt ist lokal und nicht strikt. Die Funktion zeigt auch, dass eine differenzierbare Funktion in der Umgebung eines nicht strikten lokalen Extremums nicht konstant sein muss.



7.10 Konvexität

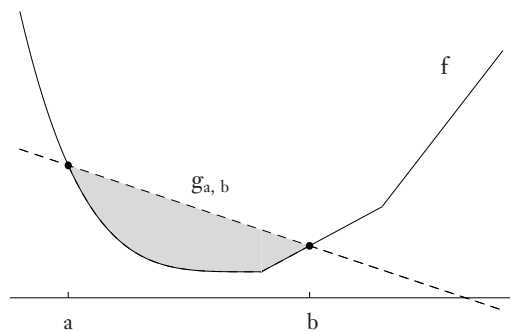
Definition (konvex, konkav)

Sei I ein Intervall und sei $f: I \rightarrow \mathbb{R}$. Für alle $a < b$ in I sei $g_{a,b}: \mathbb{R} \rightarrow \mathbb{R}$ die Sekante von f durch die Punkte $(a, f(a))$ und $(b, f(b))$. Dann heißt f

- (a) *konvex*, falls $f \leq g_{a,b}$ auf $]a, b[$,
- (b) *konkav*, falls $f \geq g_{a,b}$ auf $]a, b[$,
- (c) *streng konvex*, falls $f < g_{a,b}$ auf $]a, b[$,
- (d) *streng konkav*, falls $f > g_{a,b}$ auf $]a, b[$ für alle $a < b$ in I gilt.

Statt „konvex“ und „konkav“ spricht man auch von „linksgekrümmt“ bzw. „rechtsgekrümmt“. Die Konvexität bedeutet, dass f unterhalb aller Geradenstücke liegt, die zwei Punkte des Graphen miteinander verbinden. Bei der strengen Konvexität wird zusätzlich gefordert, dass f mit diesen Geradenstücken immer nur die Endpunkte gemeinsam hat.

Mit Hilfe der Differentialrechnung lassen sich die Regionen der Konkavität und Konvexität einer Funktion bestimmen. Wir diskutieren dies in der nächsten Sektion.



Beispiele

- (1) Jede Gerade ist konvex und konkav.
- (2) $\exp: \mathbb{R} \rightarrow \mathbb{R}$ ist streng konvex, $\log:]0, +\infty[\rightarrow \mathbb{R}$ ist streng konkav.
- (3) Die Funktion $f:]0, +\infty[\rightarrow \mathbb{R}$ mit $f(x) = 1/x$ für alle $x > 0$ ist konvex. f ist identisch mit ihrer Umkehrfunktion, die also ebenfalls konvex ist.
- (4) Die Funktion $f: [0, 1] \rightarrow \mathbb{R}$ mit $f(x) = 0$ für alle $x \in [0, 1[$ und $f(1) = 1$ ist nicht stetig, aber konvex. Die Funktion $g: [0, 1] \rightarrow \mathbb{R}$ mit $g(x) = -x^2$ für alle $x \in [0, 1[$ und $g(1) = -2$ ist nicht stetig und streng konkav.
- (5) Die zweite Potenz auf \mathbb{R} ist nichtmonoton und streng konvex.
- (6) Die dritte Potenz auf \mathbb{R} ist streng monoton, streng konkav auf $]-\infty, 0]$ und streng konvex auf $[0, +\infty[$.
- (7) Die Betragsfunktion $\text{abs}: \mathbb{R} \rightarrow \mathbb{R}$ ist nicht differenzierbar, aber konvex.

Das dritte Beispiel zeigt, dass sich die Eigenschaften „konvex“ und „konkav“ beim Übergang von f zu f^{-1} nicht immer austauschen.

Allgemeine geometrische Eigenschaften

- (a) Ist f stetig und streng monoton steigend, so tauschen sich „konvex“ und „konkav“ zwischen f und f^{-1} aus. Ist f stetig und streng monoton fallend, so bleiben „konvex“ und „konkav“ zwischen f und f^{-1} erhalten.
- (b) f ist genau dann (streng) konvex, wenn $-f$ (streng) konkav ist.
- (c) Sind f und g konvex (bzw. konkav) und sind $c, d \geq 0$, so ist auch $cf + dg$ konvex (bzw. konkav). Die strengen Versionen gelten für $c, d > 0$.

Obige Beispiele zeigen, dass eine konvexe oder konkave Funktion in den Randpunkten unstetig sein kann. Im Inneren des Intervalls erzwingen die Krümmungseigenschaften dagegen automatisch gute analytische Eigenschaften:

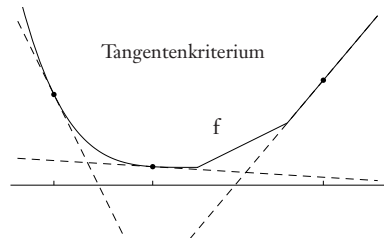
Stetigkeit und einseitige Differenzierbarkeit

Ist I ein offenes Intervall und f konvex oder konkav, so ist f stetig und links- und rechtsseitig differenzierbar.

Weiter gilt:

Tangentenkriterium

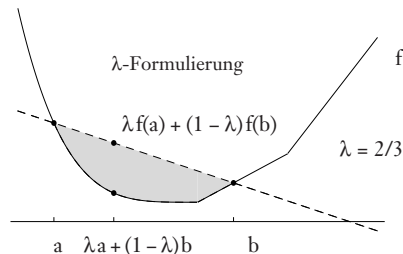
$f: I \rightarrow \mathbb{R}$ ist genau dann konvex, wenn für alle $p \in I$ eine Gerade g durch $(p, f(p))$ existiert mit $g \leq f$. Ist f differenzierbar, so gilt dies für die Tangente g an f in $(p, f(p))$.



λ -Formulierung

$f: I \rightarrow \mathbb{R}$ ist genau dann konvex, wenn für alle $a < b$ in I und $\lambda \in]0, 1[$ gilt:

$$f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b).$$



Die λ -Formulierung erlaubt eine Verallgemeinerung. Ist die Funktion f konvex, so gilt für alle x_1, \dots, x_n in I und alle $0 \leq \lambda_1, \dots, \lambda_n \leq 1$ mit $\lambda_1 + \dots + \lambda_n = 1$, dass

$$f(\lambda_1 x_1 + \dots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \dots + \lambda_n f(x_n). \quad (\text{Jensensche Ungleichung})$$

Beispiel

$\log:]0, \infty[\rightarrow \mathbb{R}$ ist konkav (und $-\log$ konvex), also gilt für alle $x_1, \dots, x_n > 0$:

$$\log\left(\frac{x_1}{n} + \dots + \frac{x_n}{n}\right) \geq \frac{\log(x_1)}{n} + \dots + \frac{\log(x_n)}{n}.$$

Die Anwendung der Exponentialfunktion liefert die Ungleichung zwischen arithmetischem und geometrischem Mittel:

$$\frac{x_1 + \dots + x_n}{n} \geq e^{\log(x_1)/n} \dots e^{\log(x_n)/n} = x_1^{1/n} \dots x_n^{1/n} = \sqrt[n]{x_1 \dots x_n}.$$

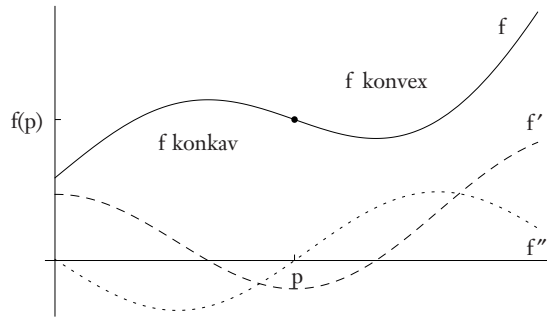
7.11 Krümmungsverhalten

Satz (Monotonie der Ableitung und Konvexität)

Sei I ein Intervall, und sei $f : I \rightarrow \mathbb{R}$ differenzierbar. Dann gilt:

- (a) f' ist monoton wachsend *genau dann, wenn* f ist konvex.
- (b) f' ist streng monoton wachsend *genau dann, wenn* f ist streng konvex.
- (c) f' ist monoton fallend *genau dann, wenn* f ist konkav.
- (d) f' ist streng monoton fallend *genau dann, wenn* f ist streng konkav.

Die Monotonie der ersten Ableitung spiegelt das Krümmungsverhalten der Funktion wider. Der Leser vergleiche diesen Satz mit dem analogen Satz über Ableitung und Monotonie in 7.8. Dieser besagte, dass das Vorzeichen der ersten Ableitung die Monotonie der Funktion widerspiegelt. Kombinieren wir beide Ergebnisse, so ergibt sich, dass das Vorzeichen der zweiten Ableitung das Krümmungsverhalten von f kodiert:



An der Stelle p ändert f die Krümmung,
 f' die Monotonie und f'' das Vorzeichen.

Ist $f : I \rightarrow \mathbb{R}$ zweimal differenzierbar, so gilt:

- (a) $f'' \geq 0$ *genau dann, wenn* f ist konvex.
- (b) $f'' > 0$ *impliziert* f ist streng konvex.
- (c) $f'' \leq 0$ *genau dann, wenn* f ist konkav.
- (d) $f'' < 0$ *impliziert* f ist streng konkav.

Setzt man in (b) und (d) die Relationen \geq^* und \leq^* statt $>$ bzw. $<$ ein (vgl. 7.8), so erhalten wir Äquivalenzen.

Beispiele

(1) Es gilt $\exp''(x) = \exp(x) > 0$ für alle $x \in \mathbb{R}$. Also ist \exp streng konvex auf \mathbb{R} .

(2) Für die vierte Potenz $f : \mathbb{R} \rightarrow \mathbb{R}$ gilt

$$f''(x) = 12x^2 \text{ für alle } x.$$

f'' ist nicht überall größer als die Nullfunktion, aber es gilt $f'' \geq^* 0$. Also ist f streng konvex nach der \geq^* -Version von (b).

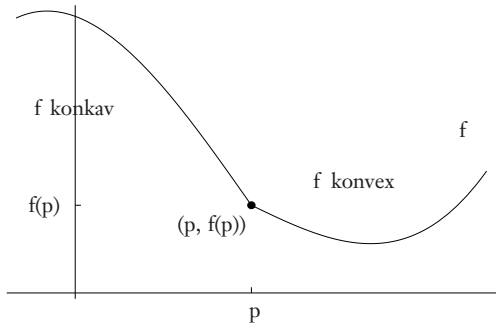
Wechselt die Ableitung f' auf I ihr Monotonieverhalten, so wechselt f das Krümmungsverhalten. Allgemein definieren wir für stetige Funktionen:

Definition (Wendepunkt)

Seien I ein Intervall und $f : I \rightarrow \mathbb{R}$ stetig. Dann besitzt f in einem $p \in I$ einen *Wendepunkt*, falls es ein $\varepsilon > 0$ mit $]p - \varepsilon, p + \varepsilon[\subseteq I$ gibt mit:

f ist konvex in $]p - \varepsilon, p[$ und
konkav in $]p, p + \varepsilon[$ *oder*

f ist konkav in $]p - \varepsilon, p[$ und
konvex in $]p, p + \varepsilon[$.



Die obigen Sätze liefern Möglichkeiten zur Bestimmung der Wendepunkte differenzierbarer Funktionen. Hat man das Monotonieverhalten von f' oder das Vorzeichenverhalten von f'' erkundet, so hat man auch die Wendepunkte von f bestimmt.

Besitzt eine zweimal differenzierbare Funktion in p einen Wendepunkt und ist ε wie in der Definition, so hat f'' in $]p - \varepsilon, p[$ und $]p, p + \varepsilon[$ verschiedene Vorzeichen. Ist f'' stetig, so gilt also $f''(p) = 0$. Damit erhalten wir:

Notwendige Bedingung für einen Wendepunkt

Ist $f : I \rightarrow \mathbb{R}$ zweimal stetig differenzierbar und besitzt f in p einen Wendepunkt, so gilt $f''(p) = 0$.

Beispiele

- (1) Für den Arkustangens $\arctan : \mathbb{R} \rightarrow \mathbb{R}$ gilt

$$\arctan'(x) = \frac{1}{1+x^2}, \quad \arctan''(x) = -\frac{2x}{(1+x^2)^2} \quad \text{für alle } x.$$

Damit ist \arctan auf $]-\infty, 0[$ streng konvex und auf $]0, +\infty[$ streng konkav. Der Nullpunkt ist der einzige Wendepunkt.

- (2) Für die vierte Potenz gilt $f''(0) = 0$. Die konvexe Funktion besitzt aber keinen Wendepunkt. Die notwendige Bedingung „ $f''(x) = 0$ “ ist also nicht hinreichend.

Wir notieren schließlich noch ein nützliches Kriterium für lokale Extremwerte, das die zweite Ableitung und damit die Krümmung von f heranzieht:

Zweite hinreichende Bedingung für lokale Extrema

Sei $f :]a, b[\rightarrow \mathbb{R}$ zweimal stetig differenzierbar und $f'(p) = 0$. Dann gilt:

Ist $f''(p) < 0$, so besitzt f in p ein striktes lokales Maximum. Ist $f''(p) > 0$, so besitzt f in p ein striktes lokales Minimum.

Das globale Minimum 0 der vierten Potenz lässt sich mit diesem Kriterium nicht ermitteln. Die Bedingung ist also nicht notwendig.

7.12 Die Taylor-Entwicklung

Satz (Satz von Taylor; polynomieller Approximationssatz, Taylor-Polynome)

Sei $f : I \rightarrow \mathbb{R}$ n -mal differenzierbar und $p \in I$. Dann gibt es ein $r : I \rightarrow \mathbb{R}$ mit

$$f(x) = T_p^n f(x) + r(x) \quad \text{für alle } x \in I, \quad \lim_{x \rightarrow p} \frac{r(x)}{(x-p)^n} = 0,$$

wobei $T_p^n f : \mathbb{R} \rightarrow \mathbb{R}$ mit

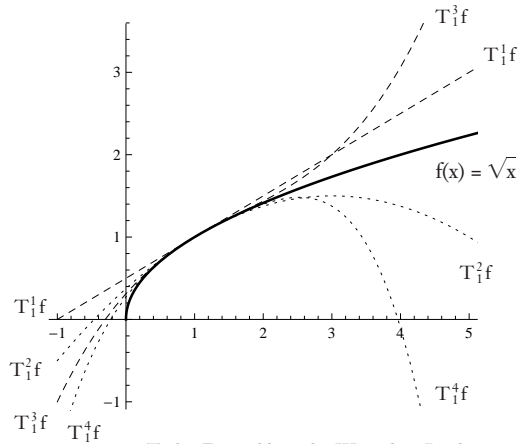
$$T_p^n f(x) = \sum_{k \leq n} \frac{f^{(k)}(p)}{k!} (x-p)^k$$

das *Taylor-Polynom n -ter Ordnung* von f im *Entwicklungspunkt* p ist.

$$T_p^0 f(x) = f(p),$$

$$T_p^1 f(x) = f(p) + f^{(1)}(p)(x-p),$$

$$T_p^2 f(x) = f(p) + f^{(1)}(p)(x-p) + \frac{f^{(2)}(p)}{2!} (x-p)^2.$$



Taylor-Entwicklung der Wurzel im Punkt 1

$$f^0(1) = 1, \quad f^{(1)}(1) = 1/2, \quad f^{(2)}(1) = -1/8,$$

$$f^{(3)}(1) = 1/16, \quad f^{(4)}(1) = -5/128.$$

Der Satz von Taylor ist im Fall $n = 1$ der Approximationssatz (vgl. 7.3):

$$f(x) = f(p) + f'(p)(x-p) + r(x), \quad \lim_{x \rightarrow p} r(x)/(x-p) = 0.$$

Ist diese Approximation erster Ordnung nicht gut genug, so kann man f durch Taylor-Polynome höheren Grades besser approximieren.

Das Restglied n -ter Ordnung

Über das *Restglied* $r : I \rightarrow \mathbb{R}$ der n -ten Taylor-Approximation kann man genauere Aussagen machen. Es gibt für alle $x \in I$ mit $x \neq p$ ein ξ zwischen p und x mit

$$r(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-p)^{n+1}. \quad (\text{Lagrangesche Form des Restglieds})$$

Mit Integralen kann man $r(x)$ genau berechnen. Wir geben die Formel hier an:

$$r(x) = \frac{1}{n!} \int_p^x (x-t)^n f^{(n+1)}(t) dt. \quad (\text{Integralform des Restglieds})$$

Kommt es nur auf die Approximation und nicht so sehr auf die Restfunktion an, so schreibt man oft mit Hilfe des Landau-Symbols (Klein-o-Notation):

$$f(x) = T_p^n f(x) + o((x-p)^n) \quad \text{für alle } x \in I.$$

Beispiele

- (1) Sei $f:]0, 2[\rightarrow \mathbb{R}$ mit $f(x) = \sqrt{x}$ für alle $x \in]0, 2[$. Dann gilt

$$f'(x) = x^{-1/2}/2, \quad f''(x) = -x^{-3/2}/4 \quad \text{für alle } x \in]0, 2[.$$

Also lautet die Taylor-Approximation zweiter Ordnung im Punkt $p = 1$:

$$f(x) = T_1^2 f(x) + o((x-1)^2) = 1 + (x-1)/2 - (x-1)^2/8 + o((x-1)^2).$$

- (2) Es gilt $\exp^{(n)}(0) = 1$ für alle n . Also gilt für alle x :

$$T_0^n \exp(x) = \sum_{k \leq n} \frac{f^{(k)}(0)}{k!} (x-0)^k = \sum_{k \leq n} \frac{x^k}{k!}.$$

Die Taylor-Polynome reproduzieren also die Exponentialreihe. Allgemein gilt:

- (3) Ist $f(x) = \sum_n a_n (x-p)^n$ eine Potenzreihe mit Konvergenzradius $R > 0$, so gilt $f^{(k)}(p) = k! a_k$ für alle k (gliedweises Differenzieren!), also $T_p^n f(x) = \sum_{k \leq n} a_k (x-p)^k$.

- (4) Die Taylor-Entwicklung lässt sich auch zur effektiven Umrechnung von Polynomdarstellungen verwenden. Ist

$$f(x) = 2(x-2)^3 + (x-2)^2 - (x-2) - 1 \quad \text{für alle } x, \text{ so gilt}$$

$$f^{(1)}(x) = 6(x-2)^2 + 2(x-2) - 1, \quad f^{(2)}(x) = 12(x-2) + 2, \quad f^{(3)}(x) = 12.$$

Wir betrachten $p = 1$ und berechnen

$$f^{(0)}(1)/0! = -1, \quad f^{(1)}(1)/1! = 3, \quad f^{(2)}(1)/2! = -5, \quad f^{(3)}(1)/3! = 2.$$

Das Lagrangesche Restglied ist 0, da $f^{(4)} = 0$. Also gilt

$$f(x) = T_1^3 f(x) = 2(x-1)^3 - 5(x-1)^2 + 3(x-1) - 1.$$

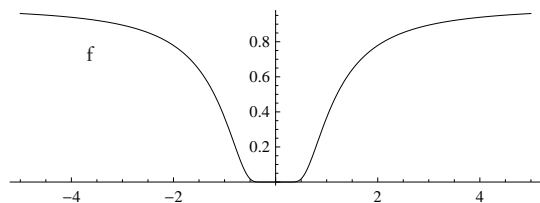
Es ist verführerisch, die Ordnung n der Approximation gegen unendlich gehen zu lassen, also zu fragen, ob für alle $x \in I$ gilt, dass

$$f(x) = T_p f(x) = \sum_n \frac{f^{(n)}(p)}{n!} (x-p)^n. \quad (\text{Versuch der Potenzreihenentwicklung})$$

Die Reihe $T_p f(x)$ heißt die *Taylor-Reihe* von f im Entwicklungspunkt p . Stimmt sie mit f überein, so ist f als Potenzreihe dargestellt. Dies gelingt zum Beispiel für \exp , \sin und \cos und allgemein, falls $\lim_n r_n(x) = 0$ für alle x gilt, wobei r_n das Restglied n -ter Ordnung ist. Die Taylor-Reihe kann aber auch nur im Entwicklungspunkt mit f übereinstimmen:

Beispiel

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ definiert durch $f(x) = e^{-1/x^2}$ für $x \neq 0$, $f(0) = 0$. Dann ist f glatt und $f^{(n)}(0) = 0$ für alle n . Also ist die Taylor-Reihe $T_0 f$ die Nullfunktion.



8. Kapitel

Integration

8.1 Partitionen und Treppenfunktionen

Definition (Partition, Zerlegungspunkte, Stützstellen)

Sei $[a, b]$ ein reelles Intervall, und seien $t_0, \dots, t_n, x_0, \dots, x_n$ reelle Zahlen mit

$$a = t_0 \leq x_0 \leq t_1 \leq x_1 \leq \dots \leq t_n \leq x_n \leq b.$$

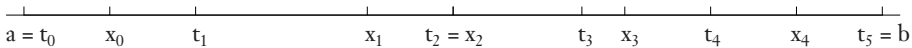
Dann heißt $p = (t_k, x_k)_{k \leq n}$ eine *Partition* von $[a, b]$ der *Länge* $n + 1$. Die Zahlen t_k heißen die *Zerlegungspunkte* und die Zahlen x_k die *Stützstellen* der Partition p . Wir setzen weiter $t_{n+1} = b$.

- (a) Ist $\delta > 0$ und gilt $t_{k+1} - t_k \leq \delta$ für alle $k \leq n$, so heißt p eine δ -*Partition* oder eine Partition der *Feinheit* δ . Die reelle Zahl

$$\delta(p) = \max_{k \leq n} (t_{k+1} - t_k)$$

heißt die (*minimale*) *Feinheit* von p .

- (b) Gilt $t_{k+1} - t_k = \frac{b-a}{n+1}$ für alle $k \leq n$, so heißt p *äquidistant*.



Eine Partition p von $[a, b]$ der Länge 5 mit Zerlegungspunkten t_k und Stützstellen x_k .

Es gilt $\delta(p) = t_2 - t_1$.

Im Gegensatz zur Differentiation analysieren wir in der Integrationstheorie eine Funktion $f: [a, b] \rightarrow \mathbb{R}$ nicht lokal, sondern als Ganzes. Partitionen werden uns dazu dienen, f durch einfache Funktionen zu approximieren. Bevor wir uns diesen Approximationen zuwenden, wollen wir Partitionen noch etwas genauer betrachten.

Sei also $[a, b]$ ein Intervall, und sei $p = (t_k, x_k)_{k \leq n}$ eine Partition von $[a, b]$ der Länge $n + 1$. Die Zerlegungspunkte t_k der Partition teilen das Intervall $[a, b]$ in $n + 1$ sich an den Randpunkten überschneidende abgeschlossene Teilintervalle auf:

$$[t_0, t_1], [t_1, t_2], \dots, [t_{n-1}, t_n], [t_n, t_{n+1}], \quad a = t_0, \quad b = t_{n+1}.$$

Dies motiviert, warum wir $n + 1$ und nicht n als Länge von p betrachten. Wir verlangen dabei nicht, dass t_k kleiner als t_{k+1} ist. Die Intervalle $[t_k, t_{k+1}]$ können aus genau einem Punkt bestehen. Es kann zum Beispiel auch $a = t_0 = t_1 = t_2$ und damit $\{a\} = [t_0, t_1] = [t_1, t_2]$ gelten.

Die Stützstellen x_k der Partition liegen in den Zerlegungsintervallen:

$$x_0 \in [t_0, t_1], \quad x_1 \in [t_1, t_2], \quad \dots, \quad x_n \in [t_n, t_{n+1}].$$

Auch sie können zusammenfallen. Es kann etwa $x_0 = t_1 = x_1$ gelten.

Eine Partition p hat immer die Feinheit $\delta(p)$. Ist $\delta \geq \delta(p)$, so ist p auch eine Partition der Feinheit δ . Ist $\delta < \delta(p)$, so ist p keine Partition der Feinheit δ .

Eine äquidistante Partition p von $[a, b]$ zerlegt $[a, b]$ in $n + 1$ gleichlange Teilintervalle. Es gilt $\delta(p) = (b - a)/(n + 1)$. Die Stützstellen können wieder beliebig in den Teilintervallen liegen.

Beispiele

- (1) Für ein Intervall $[a, b]$ und eine natürliche Zahl n sei

$$t_k = x_k = a + k \frac{b - a}{n + 1} \quad \text{für alle } k \leq n.$$

Dann ist $p_n = (t_k, x_k)_{k \leq n}$ eine äquidistante Partition von $[a, b]$ der Länge $n + 1$. Es gilt $\delta(p_n) = (b - a)/(n + 1)$ und damit $\lim_n \delta(p_n) = 0$.

- (2) Sei $[a, b]$ ein Intervall und seien

$$a = t_0 \leq t_1 \leq \dots \leq t_n \leq t_{n+1} = b.$$

Natürliche Möglichkeiten der Definition von Stützstellen x_k , $k \leq n$, sind:

$$\text{links: } x_k = t_k \quad \text{mittig: } x_k = \frac{t_k + t_{k+1}}{2} \quad \text{rechts: } x_k = t_{k+1}.$$

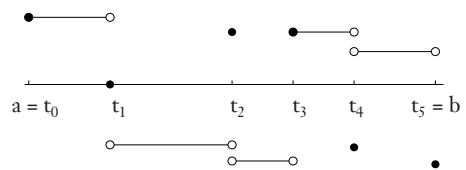
Für manche Betrachtungen genügen Zerlegungspunkte t_k eines Intervalls, Stützstellen werden nicht gebraucht. Wir nennen $p = (t_k)_{k \leq n}$ eine (*stützstellenfreie*) *Partition* von $[a, b]$ der Länge $n + 1$, falls $a = t_0 \leq \dots \leq t_n \leq b$ gilt. Wir setzen wieder $t_{n+1} = b$ und definieren die Feinheitbegriffe und die Äquidistanz wie früher. Diese einfacheren Partitionen eignen sich beispielsweise zur Definition von Treppenfunktionen:

Definition (Treppenfunktion)

Ein $f : [a, b] \rightarrow \mathbb{R}$ heißt eine *Treppenfunktion* (mit endlich vielen Stufen) auf $[a, b]$, falls es eine Partition $p = (t_k)_{k \leq n}$ von $[a, b]$ und $c_0, \dots, c_n \in \mathbb{R}$ gibt mit

$$f(x) = c_k \quad \text{für alle } k \leq n \text{ und } x \in]t_k, t_{k+1}[.$$

Die Werte von f an den Zerlegungspunkten t_k sind beliebig. Der Wertebereich von f ist also eine Teilmenge von $\{c_0, \dots, c_n, f(t_0), \dots, f(t_{n+1})\}$.



Beispiele

- (1) Sei $f : [0, 1] \rightarrow \mathbb{R}$ definiert durch $f(x) = 0$ für $x \in [0, 1/2[$, $f(1/2) = 1/2$ und $f(x) = 1$ für $x \in]1/2, 1]$. Dann ist f eine Treppenfunktion.
- (2) Sei $g : [0, 1] \rightarrow \mathbb{R}$ die Dirichletsche Sprungfunktion auf $[0, 1]$. Dann nimmt g nur die beiden Werte 0 und 1 an, aber g ist keine Treppenfunktion.
- (3) Sei $h : [0, 1] \rightarrow \mathbb{R}$ mit $h(x) = 1/2^n$ für alle n und alle $x \in]1/2^{n+1}, 1/2^n]$. Dann ist h keine Treppenfunktion, da h unendlich viele Werte annimmt.
- (4) Ist $f : [a, b] \rightarrow \mathbb{R}$ eine Treppenfunktion und stimmt $g : [a, b] \rightarrow \mathbb{R}$ bis auf endlich viele Stellen mit f überein, so ist auch g eine Treppenfunktion.

8.2 Das Riemann-Integral

Definition (Riemann-Summe und Riemann-Integral)

Sei $f: [a, b] \rightarrow \mathbb{R}$.

- (a) Für jede Partition $p = (t_k, x_k)_{k \leq n}$ von $[a, b]$ heißt die reelle Zahl

$$\sum_p f = \sum_{k \leq n} f(x_k) \cdot (t_{k+1} - t_k)$$

die *Riemann-Summe* von f bzgl. p .

- (b) f heißt (*Riemann-*)*integrierbar*, falls ein $c \in \mathbb{R}$ existiert mit:

Für alle $\varepsilon > 0$ existiert ein $\delta > 0$, sodass für alle δ -Partitionen p von $[a, b]$ gilt: $|\sum_p f - c| < \varepsilon$. (Integrierbarkeitsbedingung)

Die Zahl c heißt dann das (*bestimmte*) *Riemann-Integral* von f .

In Zeichen schreiben wir:

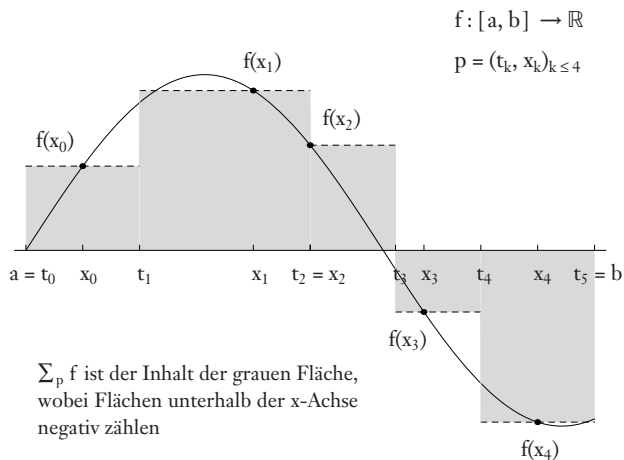
$$I(f) = \int_a^b f = \int_a^b f(x) dx = c.$$

In einer Riemann-Summe werten wir f in den Stützstellen einer Partition aus und berechnen den elementaren Flächeninhalt der entstehenden Figur. Dabei zählen Flächen unterhalb der x -Achse negativ. Wir sprechen deswegen auch von einer *vorzeichenbehafteten* oder *signierten Flächenmessung*. Mit Hilfe der Riemannschen Summen versuchen wir dann, die vom Graphen von f eingeschlossene signierte Fläche zu messen, indem wir immer feinere Partitionen zur approximativen Messung einsetzen. Die Integrierbarkeitsbedingung haben wir in einer ε - δ -Formulierung angegeben. Sie lässt sich gleichwertig auch wie folgt notieren:

Für alle Folgen $(p_n)_{n \in \mathbb{N}}$ von Partitionen von $[a, b]$ mit $\lim_n \delta(p_n) = 0$ gilt $\lim_n \sum_{p_n} f = c$. (Integrierbarkeitsbedingung, Folgenversion)

Eine suggestive Notation hierfür ist: $\int_a^b f = \lim_{\delta(p) \rightarrow 0} \sum_p f$.

Das Integralzeichen erinnert an das Summenzeichen \sum und suggeriert eine unendlich feine Summation. Die Definition des Integrals benötigt aber keine infinitesimalen Größen „ dx “.



Man erhält eine äquivalente Definition der Integrierbarkeit, wenn man nur äquidistante Partitionen zulässt:

$$\int_a^b f = \lim_{\delta(p) \rightarrow 0, p \text{ äquidistant}} \sum_p f = \lim_{\text{Länge}(p) \rightarrow \infty, p \text{ äquidistant}} \sum_p f.$$

Die letzte Form beruht auf der Beobachtung, dass für eine Folge $(p_n)_{n \in \mathbb{N}}$ äquidistanter Partitionen genau dann $\lim_n \delta(p_n) = 0$ gilt, wenn $\lim_n \text{Länge}(p_n) = +\infty$.

Weiß man, dass f integrierbar ist, so kann man eine beliebige Folge von Partitionen mit $\lim_n \delta(p_n) = 0$ zur Berechnung des Integrals verwenden. Wählen wir etwa äquidistante Partitionen mit linksseitigen Stützstellen $x_k = t_k$, so erhalten wir

$$\int_a^b f = \lim_n w_n \sum_{k \leq n} f(a + k w_n) \quad \text{mit } w_n = \frac{b-a}{n+1} \quad (\text{Berechnungsformel}).$$

Beispiele

- (1) Ist $f: [0, 1] \rightarrow \mathbb{R}$ konstant gleich c , so gilt für alle Partitionen $p = (t_k, x_k)_{k \leq n}$:

$$\sum_p f = \sum_{k \leq n} f(x_k) \cdot (t_{k+1} - t_k) = c \cdot \sum_{k \leq n} (t_{k+1} - t_k) = c \cdot 1 = c.$$

Also ist

$$\int_0^1 c \, dx = \int_0^1 f = \lim_{\delta(p) \rightarrow 0} \sum_p f = c.$$

- (2) Sei $g: [0, 1] \rightarrow \mathbb{R}$, $g(x) = x$ für alle x .

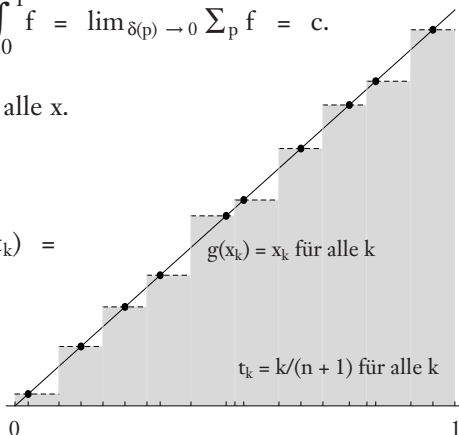
Dann gilt für alle äquidistanten

Partitionen $p = (t_k, x_k)_{k \leq n}$:

$$\sum_p g = \sum_{k \leq n} g(x_k) \cdot (t_{k+1} - t_k) =$$

$$\sum_{k \leq n} \frac{x_k}{n+1}.$$

Es gilt $t_k \leq x_k \leq t_{k+1}$,
und damit



$$\frac{1}{n+1} \sum_{k \leq n} \frac{k}{n+1} \leq \sum_p g \leq \frac{1}{n+1} \sum_{k \leq n} \frac{k+1}{n+1}.$$

Mit der Summenformel $\sum_{k \leq n} k = n(n+1)/2$ erhalten wir

$$\frac{n(n+1)}{2(n+1)^2} \leq \sum_p g \leq \frac{(n+1)(n+2)}{2(n+1)^2}.$$

Die linke und die rechte Seite konvergieren gegen $1/2$. Folglich ist

$$\int_0^1 x \, dx = \int_0^1 g = \lim_{\delta(p) \rightarrow 0, p \text{ äquidistant}} \sum_p g = 1/2.$$

8.3 Das Darboux-Integral

Definition (*Darboux-Summen, Ober- und Unterintegral, Darboux-Integral*)

Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt.

- (a) Für jede (stützstellenfreie) Partition $p = (t_k)_{k \leq n}$ von $[a, b]$ heißen

$$S_p f = \sum_{k \leq n} (t_{k+1} - t_k) \cdot \sup_{x \in [t_k, t_{k+1}]} f(x) \quad \text{und}$$

$$s_p f = \sum_{k \leq n} (t_{k+1} - t_k) \cdot \inf_{x \in [t_k, t_{k+1}]} f(x)$$

die *obere* bzw. *untere Darboux-Summe* von f bzgl. p .

- (b) $Sf = \inf_p S_p f$ und $sf = \sup_p s_p f$ heißen das *Ober-* bzw. *Unterintegral* von f .
Gilt $Sf = sf = c$, so heißt f *Darboux-integrierbar* und c das *Darboux-Integral* von f .

Auch das Darboux-Integral (gesprochen *darbú*) verwendet Partitionen, aber f wird nicht an Stützstellen ausgewertet, sondern von oben und unten approximiert. Wie für das Riemann-Integral kann man sich auf äquidistante Partitionen beschränken. Der Begriff der Feinheit wird nicht benötigt. Sf und sf sind immer definiert. Um zu zeigen, dass f Darboux-integrierbar ist, genügt es, für jedes $\varepsilon > 0$ Partitionen p_1 und p_2 anzugeben mit

$$S_{p_1} f - s_{p_2} f < \varepsilon.$$

(*Darbouxsche Integrierbarkeitsbedingung*)

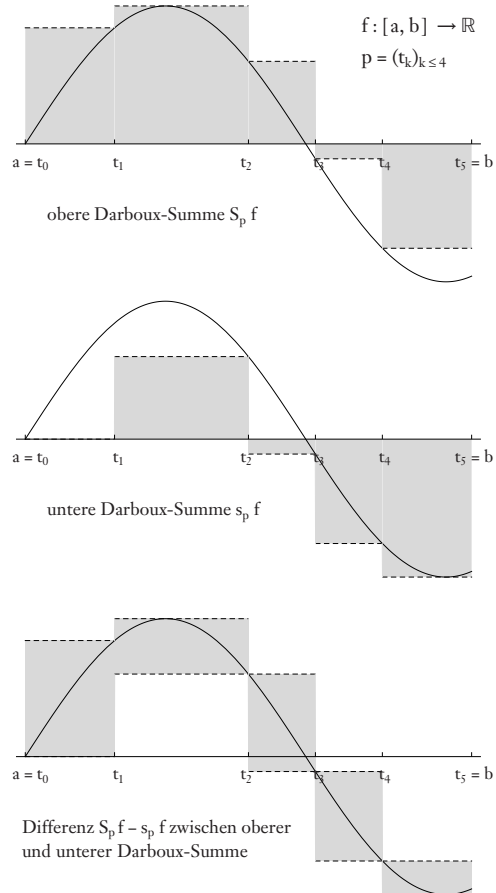
Denn aus $s_{p_2} f \leq sf \leq Sf \leq S_{p_1} f$ folgt $Sf - sf < \varepsilon$ für alle $\varepsilon > 0$, also $Sf = sf$.

Eine Riemann-Summe liegt zwischen den beiden Darboux-Summen:

$$s_p f \leq \sum_p f \leq S_p f,$$

wobei wir in $s_p f$ und $S_p f$ die Stützstellen von p ignorieren. Die Ungleichung lässt vermuten, dass die Darboux-Integrierbarkeit die Riemann-Integrierbarkeit impliziert. Dies ist richtig, und auch die Umkehrung gilt: Auch wenn f ihr Supremum und Infimum in einem

Intervall von p nicht annimmt, so können wir diese beiden Zahlen doch durch Funktionswerte beliebig genau approximieren und damit eine Darboux-Summe beliebig genau durch eine Riemann-Summe approximieren. Das Ergebnis lautet:



Äquivalenz von Riemann-Integral und Darboux-Integral

Eine Funktion $f: [a, b] \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn sie Darboux-integrierbar ist. In diesem Fall gilt

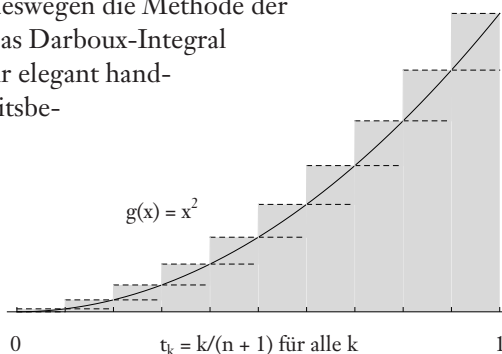
$$\int_a^b f = sf = Sf.$$

Eine Riemann-integrierbare Funktion ist automatisch beschränkt. Beim Darboux-Integral müssen wir die Beschränktheit voraussetzen, um die Infima und Suprema in den Teilintervallen der Partition bilden zu können.

Die beiden Ansätze sind äquivalent, aber konzeptionell verschieden. Die konkrete Berechnung von Riemann-Summen ist einfacher als die von Darboux-Summen, und Riemann-Summen sind deswegen die Methode der Wahl in numerischen Betrachtungen. Das Darboux-Integral lässt sich dagegen in der Theorie oft sehr elegant handhaben. Die Darboux'sche Integrierbarkeitsbedingung ist ein Beispiel hierfür.

Beispiel

Sei $g: [0, 1] \rightarrow \mathbb{R}$ mit $g(x) = x^2$ für alle x . Sei $n \in \mathbb{N}$, und sei p die äquidistante Partition von $[0, 1]$ der Länge $n + 1$. Dann gilt:



$$\begin{aligned} s_p g &= \sum_{k \leq n} (t_{k+1} - t_k) \cdot \inf_{x \in [t_k, t_{k+1}]} x^2 = \\ &= \sum_{k \leq n} \frac{1}{n+1} \cdot \frac{k^2}{(n+1)^2} = \frac{n(2n+1)}{6(n+1)^2}. \\ S_p g &= \sum_{k \leq n} (t_{k+1} - t_k) \cdot \sup_{x \in [t_k, t_{k+1}]} x^2 = \\ &= \sum_{k \leq n} \frac{1}{n+1} \cdot \frac{(k+1)^2}{(n+1)^2} = \frac{(n+2)(2n+3)}{6(n+1)^2}. \end{aligned}$$

Da g monoton steigt, sind die Infima und Suprema die Funktionswerte an den Randpunkten der Teilintervalle. Weiter haben wir die Summenformel $\sum_{k \leq n} k^2 = n(n+1)(2n+1)/6$ benutzt. Nun berechnen wir

$$\lim_n \frac{n(2n+1)}{(n+1)^2} = 2, \quad \lim_n \frac{(n+2)(2n+3)}{(n+1)^2} = 2.$$

Folglich gilt

$$sg = Sg = 1/6 \cdot 2 = 1/3, \text{ also } \int_0^1 x^2 dx = \int_0^1 g = 1/3.$$

Von den Mühen der Berechnung per Hand wird uns der Hauptsatz in 8.8 befreien!

8.4 Das Regelintegral

Definition (Regelfunktionen und Regelintegral)

1. Schritt: Integral für Treppenfunktionen

Sei $f: [a, b] \rightarrow \mathbb{R}$ eine Treppenfunktion, und seien $p = (t_k)_{k \leq n}$ eine Partition von $[a, b]$ und $c_0, \dots, c_n \in \mathbb{R}$ mit

$$f(x) = c_k \quad \text{für alle } x \in]t_k, t_{k+1}[.$$

Dann definieren wir das *Regelintegral* $R(f) \in \mathbb{R}$ von f durch

$$R(f) = \sum_{k \leq n} c_k (t_{k+1} - t_k).$$

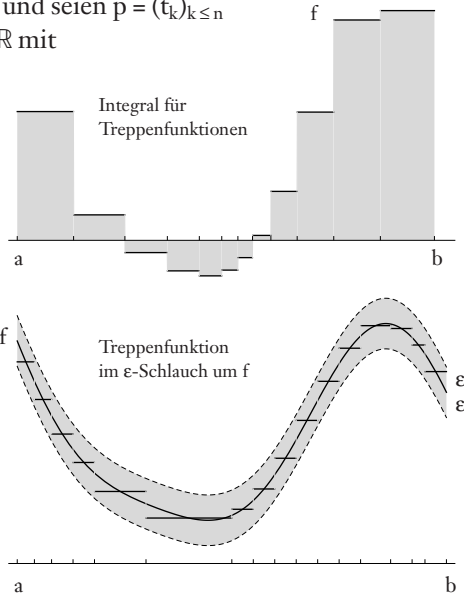
2. Schritt: Integral für Regelfunktionen

Ein $f: [a, b] \rightarrow \mathbb{R}$ heißt eine *Regelfunktion*, falls es eine Folge $(f_n)_{n \in \mathbb{N}}$ von Treppenfunktionen auf $[a, b]$ gibt mit

$$f = \lim_n f_n \quad (\text{gleichmäßig}).$$

Dann definieren wir das *Regelintegral* $R(f) \in \mathbb{R}$ von f durch

$$R(f) = \lim_n R(f_n).$$



Vom elementargeometrischen (signierten) Flächeninhalt einer Treppenfunktion führt uns eine natürliche Grenzwertbildung zu einer weiteren Variante des Integrals. Es ist für alle Funktionen f auf $[a, b]$ erklärt, die sich gleichmäßig durch Treppenfunktionen approximieren lassen, d. h. für die in jedem ϵ -Schlauch um f eine Treppenfunktion liegt. Dies ist für viele Funktionen der Fall, aber das Regelintegral ist im Gegensatz zum Darboux-Integral nicht mehr äquivalent zum Riemann-Integral: Mit dem Riemann-Integral kann man mehr Funktionen integrieren (siehe unten).

In der Definition werden spezielle Repräsentanten zur Definition von $R(f)$ benutzt. Eine Treppenfunktion lässt sich auf vielerlei Weisen mit Partitionen darstellen und eine Regelfunktion kann auf vielerlei Weisen gleichmäßig durch Treppenfunktionen approximiert werden. Man kann aber zeigen, dass $R(f)$ wohldefiniert ist, d. h., $R(f)$ hängt nicht von den gewählten Repräsentanten ab.

Eine von Treppenfunktionen unabhängige Charakterisierung lautet:

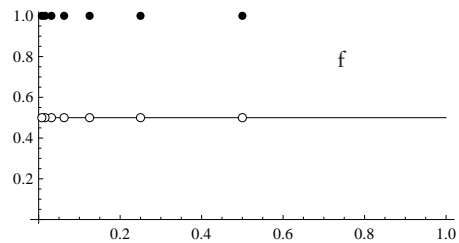
Charakterisierung der Regelfunktionen

Ein $f: [a, b] \rightarrow \mathbb{R}$ ist genau dann eine Regelfunktion, wenn f links- und rechtsseitige Grenzwerte besitzt, d. h. wenn gilt:

- (a) $\lim_{x \uparrow p} f(x)$ und $\lim_{x \downarrow p} f(x)$ existieren für alle $p \in]a, b[$,
- (b) $\lim_{x \downarrow a} f(x)$ und $\lim_{x \uparrow b} f(x)$ existieren.

Beispiele

- (1) Jede stetige Funktion f ist eine Regelfunktion. Die links- und rechtsseitigen Grenzwerte existieren (und stimmen überein).
- (2) Jede monotone Funktion auf $[a, b]$ ist eine Regelfunktion (und beschränkt durch $\max(|f(a)|, |f(b)|)$). Die Grenzwerte existieren, da jede beschränkte monotone Folge in \mathbb{R} konvergiert.
- (3) Sei $c \in [0, 1]$ irrational und sei $f: [0, 1] \rightarrow \mathbb{R}$ mit $f(x) = 0$ für $x \leq c$ und $f(x) = 1$ für $x > c$. Dann ist f eine Treppenfunktion. Im $1/2$ -Schlauch um f liegt keine Treppenfunktion mit einer äquidistanten Partition, da derartige Partitionen von $[0, 1]$ rationale Zerlegungspunkte besitzen. Das Beispiel zeigt, dass wir uns für das Regelintegral nicht auf äquidistante Partitionen beschränken können.
- (4) Sei $f: [0, 1] \rightarrow [0, 1]$ die Funktion mit $f(1/2^n) = 1$ für alle n und $f(x) = 1/2$ sonst. Dann ist f keine Regelfunktion, da $\lim_{x \downarrow 0} f(x)$ nicht existiert (oder weil im $1/4$ -Schlauch um f keine Treppenfunktion liegt).



Für „Regelintegral versus Riemann-Integral“ gilt:

Jede Regelfunktion ist Riemann-integrierbar und es gilt $R(f) = I(f) \quad (= \int_a^b f \, dx)$.

Die Umkehrung ist jedoch nicht richtig:

Beispiel

Die Funktion f in (4) ist Riemann-integrierbar mit $I(f) = 0$. Denn ein Intervall $[0, \varepsilon/2]$ trägt höchstens $\varepsilon/2$ zu einer Riemann-Summe $\sum_p f$ bei. In $[\varepsilon/2, 1]$ ist f nur endlich oft ungleich 0. Der Beitrag von $\sum_p f$ in $[\varepsilon/2, 1]$ ist also für hinreichend feine Partitionen kleiner als $\varepsilon/2$. Für diese Partitionen gilt also $\sum_p f < \varepsilon$.

Aus der Darboux'schen Integrierbarkeitsbedingung erhält man, dass eine Funktion $f: [a, b] \rightarrow \mathbb{R}$ genau dann Riemann-integrierbar ist, wenn gilt:

Für alle $\varepsilon > 0$ gibt es Treppenfunktionen g, h auf $[a, b]$ mit $g \leq f \leq h$ und $I(h) - I(g) < \varepsilon$.
(Integrierbarkeitsbedingung, II)

Riemann-integrierbare Funktionen lassen sich also zwischen zwei Treppenfunktionen einschließen, deren Integraldifferenz beliebig klein wird. Dies ist schwächer als die gleichmäßige Approximierbarkeit durch Treppenfunktionen!

Das Riemann-Integral ist also eine echte Fortsetzung des Regelintegrals. Das Regelintegral umfasst aber hinreichend viele Funktionen, sodass es durchaus eine Alternative zum Riemann-Integral darstellt, auch vor dem Hintergrund, dass beide Integrale in der höheren Analysis durch das sog. Lebesgue-Integral noch einmal erweitert werden.

8.5 Eigenschaften des Integrals

Satz (Eigenschaften des Regel- und des Riemann-Integrals)

Für alle integrierbaren $f, g : [a, b] \rightarrow \mathbb{R}$, $c, d \in \mathbb{R}$ und $s \leq t$ in $[a, b]$ gilt:

$$\int_a^b 1 \, dx = b - a, \quad (\text{Normiertheit})$$

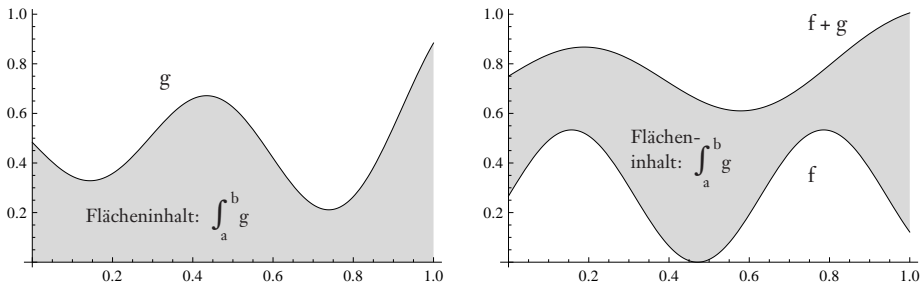
$$\int_a^b c f(x) + d g(x) \, dx = c \int_a^b f(x) \, dx + d \int_a^b g(x) \, dx, \quad (\text{Linearität})$$

$$\int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx, \quad \text{falls } f \leq g, \quad (\text{Monotonie})$$

$$\int_s^t f(x) \, dx \text{ existiert, d.h., } f|_{[s, t]} \text{ ist integrierbar,} \quad (\text{Einschränkung})$$

$$\int_a^b f(x) \, dx = \int_a^s f(x) \, dx + \int_s^b f(x) \, dx. \quad (\text{Aufspaltung})$$

Das betrachtete Integral ist dabei das Riemann- oder das Regelintegral.



Die Eigenschaften sind der Beginn eines Kalküls der Integralrechnung. Im Vergleich zur Differentiation fehlen Integrationsregeln für das Produkt, die Division und die Verknüpfung von Funktionen. Wir werden später den Kalkül der Integration noch durch die Produkt- und die Substitutionsregel erweitern, aber er bleibt unvollständig. Die weiteren Regeln gewinnen wir aus dem Hauptsatz durch Umkehrung von Differentiationsregeln. Auch sie liefern keine allgemeine Formel zur Berechnung des Integrals von fg und $g \circ f$.

Die Regeln des Satzes beinhalten die Integrierbarkeit der betrachteten Funktionen. So besagt zum Beispiel die Linearität, dass für alle integrierbaren Funktionen f und g auf $[a, b]$ die Linearkombination $cf + dg$ wieder integrierbar ist. Setzen wir für ein Intervall $[a, b]$

$$V = \{ f : [a, b] \rightarrow \mathbb{R} \mid f \text{ ist integrierbar} \},$$

so bildet V einen Untervektorraum des Vektorraumes aller reellen Funktionen auf $[a, b]$, und die Abbildung $I : V \rightarrow \mathbb{R}$ ist ein lineares Funktional auf V :

$$I(f) = \int_a^b f(x) \, dx \quad \text{für alle } f \in V.$$

Wir betrachten die Linearität des Integrals noch etwas genauer. Wir können sie zweiteilig schreiben als

$$\int_a^b c f(x) dx = c \int_a^b f(x) dx, \quad \int_a^b f(x) + g(x) dx = \int_a^b f(x) dx + \int_a^b g(x) dx.$$

Der erste Teil bedeutet: Strecken wir f um den Faktor c , so strecken wir auch die von f und der x -Achse eingeschlossene Fläche um diesen Faktor. Der zweite Teil ist überraschender: Setzen wir den Graphen der Funktion g auf den Graphen der Funktion f auf, so hat der Bereich zwischen f und $f + g$ dieselbe Fläche wie der von g und der x -Achse eingeschlossene Bereich. Mit Hilfe von Riemann-Summen kann man die Linearität des Riemann-Integrals aber schnell einsehen: Die endliche Summeneigenschaft $\sum_p (f + g) = \sum_p f + \sum_p g$ überträgt sich auf das Integral. Ähnliches gilt für das Regelintegral.

Beispiel

$$\begin{aligned} \int_0^1 4x^2 - 3x + 1/6 dx &= 4 \int_0^1 x^2 dx - 3 \int_0^1 x dx + 1/6 \int_0^1 1 dx = \\ &= 4 \cdot 1/3 - 3 \cdot 1/2 + 1/6 \cdot 1 = 0. \end{aligned}$$

Es ist nützlich, unsere Integralnotation noch zu erweitern, sodass die untere Integrationsgrenze auch größer als die obere Integrationsgrenze sein kann. Ist $f: [a, b] \rightarrow \mathbb{R}$ integrierbar, so definieren wir

$$\int_b^a f(x) dx = - \int_a^b f(x) dx. \quad (\text{Rückwärtsintegral})$$

Derartige Rückwärtsintegrale treten insbesondere beim Hauptsatz auf.

Beispiele

- (1) Sind $f, g: [a, b] \rightarrow \mathbb{R}$ integrierbar mit $f \leq g$, so gilt

$$\int_b^a f(x) dx = - \int_a^b f(x) dx \geq - \int_a^b g(x) dx = \int_b^a g(x) dx.$$

Die Monotonieregel dreht sich also bei Rückwärtsintegralen um.

- (2) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ auf allen kompakten Intervallen integrierbar. Dann gilt für alle $a, b, c \in \mathbb{R}$:

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx. \quad (\text{erweiterte Aufspaltung})$$

Denn gilt etwa $c < a < b$, so haben wir

$$\int_a^c f(x) dx + \int_c^b f(x) dx = - \int_c^a f(x) dx + \int_c^a f(x) dx + \int_a^b f(x) dx = \int_a^b f(x) dx.$$

Die anderen Fälle werden analog behandelt.

8.6 Zum Umfang der integrierbaren Funktionen

Satz (Umfang des Riemann- bzw. des Regelintegrals)

Sei $[a, b]$ ein Intervall. Dann gilt (für das Riemann- oder das Regelintegral):

Integrierbarkeit stückweise stetiger Funktionen

Jede stückweise stetige Funktion $f: [a, b] \rightarrow \mathbb{R}$ ist integrierbar, wobei f *stückweise stetig* heißt, wenn es eine Partition $p = (t_k)_{k \leq n}$ von $[a, b]$ gibt, sodass für alle $k \leq n$ gilt:

$$f|_{[t_k, t_{k+1}[} \text{ ist stetig und } \lim_{x \downarrow t_k} f(x), \lim_{x \uparrow t_{k+1}} f(x) \text{ existieren.}$$

Integrierbarkeit monotoner Funktionen

Jede monotone Funktion $f: [a, b] \rightarrow \mathbb{R}$ ist integrierbar.

Abgeschlossenheitseigenschaften

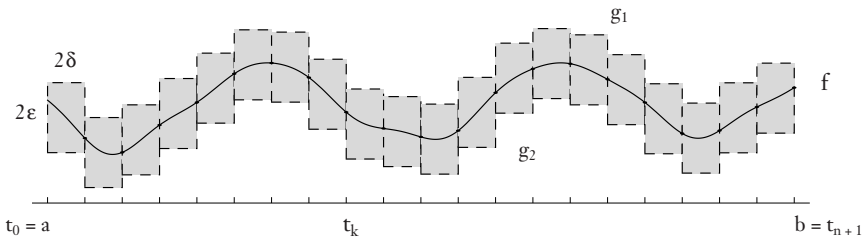
Sind $f, g: [a, b] \rightarrow \mathbb{R}$ integrierbar, so auch $f + g$, cf , fg und $|f|$.

Kompositionssatz

Ist $f: [a, b] \rightarrow [c, d]$ integrierbar und ist $g: [c, d] \rightarrow \mathbb{R}$ stetig, so ist $g \circ f: [a, b] \rightarrow \mathbb{R}$ integrierbar.

Abänderung an endlich vielen Stellen

Integrierbarkeit und Integral einer Funktion ändern sich nicht, wenn die Funktion an endlich vielen Stellen abgeändert wird.



Die gleichmäßige Stetigkeit einer stetigen Funktion $f: [a, b] \rightarrow \mathbb{R}$ liefert für alle $\varepsilon > 0$ eine äquidistante Partition von $[a, b]$ und Treppenfunktionen g_1 und g_2 , die f von oben und unten approximieren (vgl. 5.5). Der Inhalt $2\varepsilon(b-a)$ der Fläche zwischen g_1 und g_2 wird beliebig klein. Dies zeigt die Integrierbarkeit stetiger Funktionen.

Die stückweise stetigen und die monotonen Funktionen sind integrierbar und decken einen weiten Bereich ab. Jede stetige Funktion ist stückweise stetig. Wir halten aber fest, dass die stückweise Stetigkeit stärker ist als die Bedingung, dass f nur endlich viele Unstetigkeitsstellen besitzt. Die Funktion $f: [-1, 1] \rightarrow \mathbb{R}$ mit $f(x) = 1/x$ für alle $x \neq 0$ und $f(0) = 0$ hat nur eine Unstetigkeitsstelle, aber sie ist nicht stückweise stetig und als unbeschränkte Funktion auch nicht integrierbar.

Die Abgeschlossenheitseigenschaften des Satzes lassen sich durch Kombination noch vielfach erweitern. So sind zum Beispiel für eine integrierbare Funktion f auch

$$f^2 = f \cdot f, \quad f^3 = f^2 \cdot f, \quad \dots, \quad f^{n+1} = f^n \cdot f, \quad \dots$$

integrierbar. Ein weiteres Beispiel stellt die Zerlegung von einer Funktion in ihren Negativ- und Positivteil dar, die in der Integrationstheorie eine wichtige Rolle spielt (vgl. 8.12):

Beispiel

Für ein $f: [a, b] \rightarrow \mathbb{R}$ definieren wir den *Positivteil* $f^+: [a, b] \rightarrow \mathbb{R}$ und den *Negativteil* $f^-: [a, b] \rightarrow \mathbb{R}$ durch:

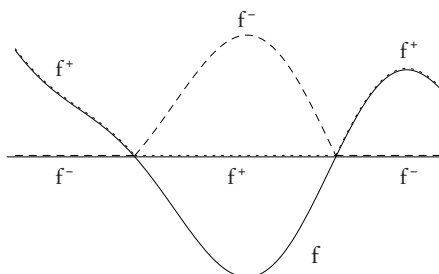
$$f^+(x) = \max(f(x), 0) \geq 0,$$

$$f^-(x) = \max(-f(x), 0) \geq 0$$

für alle $x \in [a, b]$. Dann gilt:

$$f = f^+ - f^-, \quad |f| = f^+ + f^-,$$

$$f^+ = (|f| + f)/2, \quad f^- = (|f| - f)/2.$$



Nach den Abgeschlossenheitseigenschaften ist also f genau dann integrierbar, wenn f^+ und f^- integrierbar sind. Nach Linearität des Integrals gilt dann

$$\int_a^b f = \int_a^b f^+ - \int_a^b f^-. \quad (\text{Zerlegung in Positiv- und Negativteil})$$

Wir vervollständigen unser Bild von der Menge der integrierbaren Funktionen nun noch durch einige Gegenbeispiele.

Beispiele

- (1) Jede unbeschränkte Funktion $f: [0, 1] \rightarrow \mathbb{R}$ ist nicht integrierbar.
- (2) Sei $f: [0, 1] \rightarrow \mathbb{R}$ die Dirichletsche Sprungfunktion auf $[0, 1]$. Dann ist f nicht integrierbar. Denn in jedem Intervall $[p_k, p_{k+1}]$ mit $p_k < p_{k+1}$ einer Partition nimmt f die Werte 0 und 1 an. Also ist f keine Regelfunktion und weiter auch nicht Riemann-integrierbar, da

$$S_a f = \sum_{k \leq n} (t_{k+1} - t_k) \cdot \sup_{x \in [t_k, t_{k+1}]} f(x) = \sum_{k \leq n} (t_{k+1} - t_k) 1 = 1,$$

$$s_p f = \sum_{k \leq n} (t_{k+1} - t_k) \cdot \inf_{x \in [t_k, t_{k+1}]} f(x) = \sum_{k \leq n} (t_{k+1} - t_k) 0 = 0.$$

- (3) Das Beispiel (2) zeigt auch, dass wir die Integrierbarkeit einer Funktion f durch Abänderung an abzählbar unendlich vielen Stellen zerstören können. Denn f ist eine Abänderung der Nullfunktion auf $[0, 1]$ an den abzählbar vielen rationalen Punkten in $[0, 1]$.
- (4) Sei $(q_n)_{n \in \mathbb{N}}$ eine Aufzählung aller rationalen Zahlen in $[0, 1]$ ohne Wiederholungen. Wir definieren $f, g: [0, 1] \rightarrow [0, 1]$ durch

$$f(q_n) = 1/2^n \text{ für alle } n, \quad f(x) = 0 \text{ für alle irrationalen } x,$$

$$g(x) = 1 \text{ für alle } x > 0, \quad g(0) = 0.$$

Dann sind f und g integrierbar, aber $g \circ f$ ist die nichtintegrierbare Dirichletsche Sprungfunktion aus (2). Im Kompositionssatz kann die Stetigkeit von g also nicht durch die Integrierbarkeit ersetzt werden. (Zum Beweis der Integrierbarkeit von f nutzt man, dass es für alle $\varepsilon > 0$ nur endlich viele Stellen gibt, an denen f größer als ε ist.)

8.7 Der Mittelwertsatz der Integralrechnung

Satz (*Mittelwertsatz der Integralrechnung für das Riemann- oder Regelintegral*)

Sei $f: [a, b] \rightarrow \mathbb{R}$ integrierbar und es gelte $s \leq f(x) \leq S$ für alle $x \in [a, b]$.

Einfacher Mittelwertsatz

Es gibt ein $m \in [s, S]$ mit

$$\int_a^b f = m(b - a).$$

Ist der Wertebereich von f ein Intervall, so gibt es ein $p \in [a, b]$ mit

$$\int_a^b f = f(p)(b - a). \quad (\text{Annahme des Mittelwerts})$$

Mittelwertsatz mit Gewichtsfunktion

Sei $g: [a, b] \rightarrow [0, +\infty[$ integrierbar. Dann existiert ein $m \in [s, S]$ mit

$$\int_a^b f g = m \int_a^b g.$$

Ist der Wertebereich von f ein Intervall, so gibt es ein $p \in [a, b]$ mit

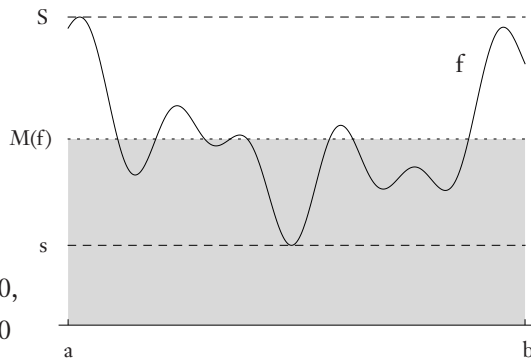
$$\int_a^b f g = f(p) \int_a^b g. \quad (\text{Annahme des gewichteten Mittelwerts})$$

Für eine integrierbare Funktion $f: [a, b] \rightarrow \mathbb{R}$, $a < b$, heißt die reelle Zahl

$$M(f) = \frac{1}{b - a} \int_a^b f$$

der *Mittelwert* von f . Das Rechteck

$$R = \begin{cases} [a, b] \times [0, M(f)] & \text{für } M(f) \geq 0, \\ [a, b] \times [M(f), 0] & \text{für } M(f) < 0 \end{cases}$$



hat dieselbe signierte Fläche, die f mit der x -Achse einschließt. Dies erklärt die Bezeichnung als „Mittelwert“. Der Mittelwertsatz besagt, dass $M(f)$ zwischen s und S liegt, falls alle Funktionswerte zwischen s und S liegen. Der Beweis ist einfach: Aufgrund der Monotonie des Integrals folgt aus $s \leq f(x) \leq S$ für alle x , dass

$$s(b - a) = \int_a^b s \leq \int_a^b f \leq \int_a^b S = S(b - a).$$

Also existiert ein $m \in [s, S]$ mit

$$m(b - a) = \int_a^b f.$$

Der Zusatz für Funktionen, deren Wertebereich ein Intervall ist (wie z. B. für stetige Funktionen), ergibt sich, wenn wir $s = \inf_{x \in [a, b]} f(x)$ und $S = \sup_{x \in [a, b]} f(x)$ setzen.

Die zweite Version des Mittelwertsatzes bringt eine nichtnegative sog. Gewichtsfunktion g ins Spiel. Die erste Version entspricht dem Spezialfall $g = 1$.

Beispiele

- (1) Sei $f: [a, b] \rightarrow \mathbb{R}$, $a < b$, stetig mit Integral $I(f) = 0$. Dann gibt es ein $p \in [a, b]$ mit $f(p)(b - a) = I(f) = 0$. Also besitzt f eine Nullstelle.
- (2) Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig, und sei $(t_k)_{k \leq n}$ eine Partition von $[a, b]$. Nach dem Mittelwertsatz gibt es $x_k \in [t_k, t_{k+1}]$ mit

$$\int_{t_k}^{t_{k+1}} f = f(x_k) (t_{k+1} - t_k) \quad \text{für alle } k \leq n \text{ (wobei wieder } t_{n+1} = b).$$

Ergänzen wir p um die Stützstellen x_k , so gilt

$$\int_a^b f = \sum_{k \leq n} \int_{t_k}^{t_{k+1}} f = \sum_{k \leq n} f(x_k) (t_{k+1} - t_k) = \sum_p f.$$

Das Integral über f ist also eine gewisse Riemann-Summe bzgl. p .

- (3) Seien $f, g: [-1, 1] \rightarrow \mathbb{R}$ mit $f(x) = g(x) = x$ für alle x . Dann gilt

$$\int_{-1}^1 f g = \int_{-1}^1 x^2 dx = 2/3 > 0, \text{ aber } m \cdot \int_{-1}^1 g = m \cdot 0 = 0 \text{ für alle } m.$$

Die Voraussetzung „ $g \geq 0$ “ ist also wesentlich.

Der Mittelwertsatz spielt eine tragende Rolle im Beweis des Hauptsatzes, den wir in der nächsten Sektion diskutieren werden. Das wesentliche Argument lässt sich schön isolieren und wir gestatten uns hier also einen Vorgriff:

Zusammenhang zwischen Integration und Differentiation

Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig. Wir setzen

$$F(x) = \int_a^x f(t) dt \quad \text{für alle } x \in [a, b].$$

Dann beschreibt $F: [a, b] \rightarrow \mathbb{R}$ das Flächenwachstum des Graphen von f auf dem Weg von a nach b . Ist nun $x \in [a, b]$, so gibt es nach dem Mittelwertsatz für alle h mit $x + h \in [a, b]$ ein p zwischen x und $x + h$ mit

$$\int_x^{x+h} f(t) dt = f(p) ((x + h) - x) = f(p) h, \quad \text{so dass}$$

$$\lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} f(t) dt = \lim_{p \rightarrow x} f(p) = f(x),$$

wobei wir zuletzt die Stetigkeit von f verwenden. Dies zeigt, dass die Flächenwachsfunktion F differenzierbar und eine *Stammfunktion* von f ist, d. h., es gilt $F' = f$.

8.8 Der Hauptsatz

Satz (Hauptsatz der Differential- und Integralrechnung)

Sei I ein Intervall, und sei $f: I \rightarrow \mathbb{R}$ (Riemann- oder Regel-) integrierbar.

Teil 1: Berechnung von Integralen mit Hilfe von Stammfunktionen

Sei $F: I \rightarrow \mathbb{R}$ eine Stammfunktion von f , d.h., F ist differenzierbar und es gilt $F' = f$. Dann gilt für alle Teilintervalle $[a, b]$ von I :

$$\int_a^b f = F(b) - F(a).$$

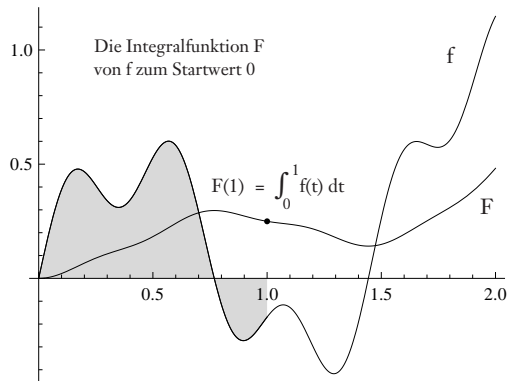
Teil 2: Existenz von Stammfunktionen

Sei $s \in I$ und sei $F: I \rightarrow \mathbb{R}$ definiert durch

$$F(x) = \int_s^x f(t) \, dt \quad \text{für alle } x \in I. \quad (\text{Integralfunktion von } f \text{ zum Startwert } s)$$

Dann ist F stetig und es gilt:

- (a) Ist f stetig, so ist F eine Stammfunktion von f .
- (b) Ist f eine Regelfunktion, so ist F links- und rechtsseitig differenzierbar und diese Ableitungen sind die links- bzw. rechtsseitigen Grenzwerte von f .



Der Hauptsatz besagt grob gesprochen:

Die Integration ist die Umkehrung der Differentiation.

Unser Wissen über das Differenzieren wird so zu einem Wissen über das Integrieren. Um das bestimmte Integral einer Funktion $f: [a, b] \rightarrow \mathbb{R}$ zu berechnen, müssen wir „lediglich“ eine Stammfunktion F von f finden. Dann gilt

$$\int_a^b f = F(b) - F(a).$$

Die Differenzen auf der rechten Seite tauchen so häufig auf, dass sich eine eigene Notation für sie etabliert hat. Wir setzen:

$$F \Big|_a^b = F(x) \Big|_a^b = [F(x)]_a^b = F(b) - F(a). \quad (\text{Auswertungsnotation})$$

Auf die Frage, wann f eine Stammfunktion besitzt und wie wir sie finden, gibt der Hauptsatz eine Teilantwort. Wir wählen einen Startwert s und betrachten die Integrale über f mit der festen unteren Grenze s und einer variablen oberen Grenze x . Dies definiert eine Funktion F auf I , die für stetige f eine Stammfunktion und für Regelfunktionen f eine „Quasistammfunktion“ von f ist.

Ist $f: I \rightarrow \mathbb{R}$ integrierbar und besitzt f eine Stammfunktion, so setzen wir

$$\int f = \int f(x) dx = \{F \mid F: I \rightarrow \mathbb{R}, F' = f\}. \quad (\text{unbestimmtes Integral von } f)$$

Die Menge $\int f$ heißt das *unbestimmte Integral* von f . Ist F eine Stammfunktion von f , so gilt

$$\int f = \{F + c \mid c \in \mathbb{R}\},$$

da sich zwei Funktionen h_1 und h_2 auf I mit derselben Ableitung nur um eine Konstante unterscheiden (da $(h_1 - h_2)' = 0$ ist $h_1 - h_2$ konstant, vgl. 7.7). In der Praxis arbeitet man oft mit einer beliebigen Stammfunktion anstatt mit einer Menge. Man schreibt etwa

$$\int \cos = \sin + c \quad \text{oder} \quad \int x^2 dx = x^3/3.$$

Solange man beachtet, dass eine Stammfunktion von f nur bis auf eine Konstante eindeutig bestimmt ist, ist diese Schreibweise ungefährlich.

Beispiele

$$(1) \int_1^e 1/x dx = \log(x) \Big|_1^e = \log(e) - \log(1) = 1.$$

$$(2) \int_0^1 \frac{1}{1+x^2} dx = \arctan(x) \Big|_0^1 = \arctan(1) - \arctan(0) = \pi/4.$$

$$(3) \int 0 = 0, \quad \int 1 = x, \quad \int x dx = x^2/2, \quad \int \tan = -\log(\cos).$$

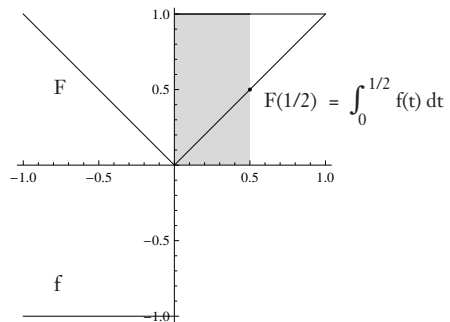
$$(4) \int_a^b f = \left(\int f \right) \Big|_a^b = \left(\int f \right)(b) - \left(\int f \right)(a).$$

(5) Ist f stetig differenzierbar, so ist f eine Stammfunktion von f' und für alle a, b gilt:

$$\int_a^b f'(x) dx = f \Big|_a^b = f(b) - f(a).$$

(6) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = -1$ für $x < 0$ und $f(x) = 1$ für $x \geq 0$. Dann besitzt f keine Stammfunktion. Die Integralfunktion F von f zum Startwert 0 ist die Betragsfunktion:

$$\int_0^x f(t) dt = |x| \quad \text{für alle } x.$$



(7) Sei $f: [0, 1] \rightarrow \mathbb{R}$ mit $f(x) = 0$ für $x \neq 1/2$, $f(1/2) = 1$. Dann ist die Integralfunktion F von f zum Startwert 0 die Nullfunktion. Es gilt $F' \neq f$, da $f(1/2) \neq 0$.

(8) Sei $f: [0, 1] \rightarrow \mathbb{R}$ mit $f(x) = x^2 \sin(1/x^2)$ für $x \neq 0$, $f(0) = 0$. Dann ist f differenzierbar mit $f'(x) = 2x \sin(1/x^2) - 2/x \cos(1/x^2)$ für $x \neq 0$ und $f'(0) = 0$. Als unbeschränkte Funktion ist f' nicht integrierbar. Eine Funktion, die eine Stammfunktion besitzt, muss also nicht integrierbar sein.

8.9 Integrationsregeln

Satz (*Integrationsregeln für das Riemann- oder Regelintegral*)

Partielle Integration

Seien $f, g : I \rightarrow \mathbb{R}$ stetig differenzierbar, I ein Intervall. Dann gilt:

$$\int f'(x) g(x) dx = fg - \int f(x) g'(x) dx \quad (\text{modulo einer Konstanten}),$$

$$\int_a^b f'(x) g(x) dx = fg \Big|_a^b - \int_a^b f(x) g'(x) dx \quad \text{für alle } a, b \in I.$$

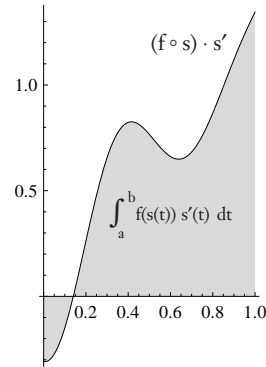
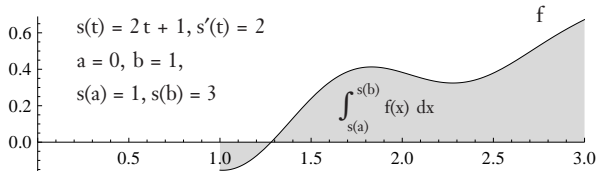
Substitutionsregel

Seien $f : J \rightarrow \mathbb{R}$ stetig und $s : I \rightarrow \mathbb{R}$ stetig differenzierbar, I, J Intervalle.

Weiter sei $f \circ s$ definiert, d.h., es gelte $s(x) \in J$ für alle $x \in I$. Dann gilt

$$\int f(s(t)) s'(t) dt = (\int f(x) dx) \circ s \quad (\text{modulo } c),$$

$$\int_a^b f(s(t)) s'(t) dt = \int_{s(a)}^{s(b)} f(x) dx \quad \text{für alle } a, b \in I.$$



Die rechte Fläche entsteht aus der linken durch Stauchung um den Faktor 2 in x-Richtung, Streckung um den Faktor 2 in y-Richtung und Translation um 1 nach links.

Die beiden Regeln ergänzen den Kalkül der Integralrechnung um zwei wichtige Elemente. Sie ergeben sich aus der Produktregel und der Kettenregel der Differentialrechnung durch Anwendung des Hauptsatzes. Wir diskutieren dies für die Substitutionsregel (die am meisten Schwierigkeiten bereitet) genauer.

Nach den Voraussetzungen der Substitutionsregel sind f und $g = (f \circ s) \cdot s'$ stetig. Nach dem Hauptsatz existieren also Stammfunktionen F und G von f bzw. g . Nach der Kettenregel gilt

$$(F \circ s)' = (F' \circ s) \cdot s' = (f \circ s) \cdot s' = g.$$

Also gilt $G = F \circ s$ (modulo einer Konstanten). Folglich ist

$$\int f(s(t)) s'(t) dt = \int g = G = F \circ s = \int f(x) dx \circ s \quad (\text{modulo } c)$$

und damit

$$\int_a^b f(s(t)) s'(t) dt = G \Big|_a^b = (F \circ s) \Big|_a^b = F \Big|_{s(a)}^{s(b)} = \int_{s(a)}^{s(b)} f(x) dx \quad \text{für } a, b \in I.$$

In Anwendungen geht man oft von $\int_c^d f(x) dx$ aus und führt eine geeignete Substitution „ $x = s(t)$ “ mit einer $\mathcal{C}^{(1)}$ -Funktion s durch. Für beliebige a, b mit $s(a) = c$, $s(b) = d$ gilt dann

$$\int_c^d f(x) dx = \int_a^b f(s(t)) s'(t) dt,$$

was man sich mittels $x = s(t)$, $dx = \frac{d(s(t))}{dt} \cdot dt = s'(t) dt$ merken kann.

Für injektive s ist $a = s^{-1}(c)$ und $b = s^{-1}(d)$. Die Injektivität von s ist aber nicht notwendig.

Beispiele

$$(1) \int \log(x) dx = \int \log(x) 1 dx = \log(x) \cdot x - \int x/x dx = x(\log(x) - 1).$$

$$(2) \int \arctan(x) dx = \int \arctan(x) 1 dx = \arctan(x) \cdot x - \int x/(1+x^2) dx = \\ \arctan(x) \cdot x - \int 1/2 (\log(1+x^2))' dx = \arctan(x) \cdot x - \log(1+x^2)/2.$$

$$(3) \int \cos^2(x) dx = \int \cos(x) \cos(x) dx = \int \cos(x) \sin'(x) dx = \\ \cos(x) \sin(x) - \int \cos'(x) \sin(x) dx = \cos(x) \sin(x) + \int \sin^2(x) dx.$$

$$\text{Addition von } \int \cos^2(x) \text{ liefert } 2 \int \cos^2(x) dx = \cos(x) \sin(x) + x.$$

$$(4) \int_a^b f(ct+d) dt = \frac{1}{c} \int_a^b f(ct+d) c dt = \frac{1}{c} \int_{ca+d}^{cb+d} f(x) dx \quad \text{für } c \neq 0.$$

Die Substitution ist „ $x = s(t) = ct + d$ “ mit $s'(t) = c$.

(5) Ist $s: [a, b] \rightarrow]0, +\infty[$ stetig differenzierbar, so gilt (mit $f(x) = 1/x$):

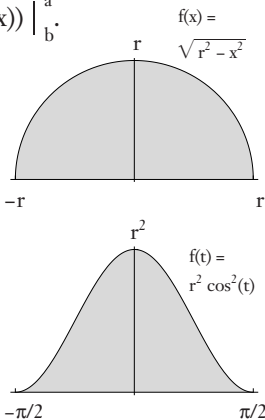
$$\int_a^b \frac{s'(t)}{s(t)} dt = \int_{s(a)}^{s(b)} \frac{1}{x} dx = \log(x) \Big|_{s(a)}^{s(b)} = \log(s(x)) \Big|_a^b.$$

(6) Wir berechnen die Fläche der oberen Hälfte des Kreises mit Radius r und Mittelpunkt 0:

$$\int_{-r}^r \sqrt{r^2 - x^2} dx = \int_{-\pi/2}^{\pi/2} \sqrt{r^2 - r^2 \sin^2(t)} r \cos(t) dt =$$

$$r^2 \int_{-\pi/2}^{\pi/2} \sqrt{1 - \sin^2(t)} \cos(t) dt = r^2 \int_{-\pi/2}^{\pi/2} \cos^2(t) dt =$$

$$r^2 \frac{\cos(x) \sin(x) + x}{2} \Big|_{-\pi/2}^{\pi/2} = r^2 (\pi/4 + \pi/4) = r^2 \pi/2.$$



Dabei haben wir „ $x = s(t) = r \sin(t)$ “ mit $s'(t) = r \cos(t)$ und (3) verwendet.

8.10 Uneigentliche Integrale

Definition (*Uneigentliches Integral für halboffene und offene Intervalle*)

Uneigentliches Integral für halboffene Intervalle

Sei $f : [a, b[\rightarrow \mathbb{R}$, $b \leq +\infty$. Dann setzen wir im Fall der Existenz

$$\int_a^b f = \lim_{c \uparrow b} \int_a^c f \in [-\infty, +\infty].$$

Analog setzen wir für $f :]a, b] \rightarrow \mathbb{R}$, $a \geq -\infty$ im Fall der Existenz

$$\int_a^b f = \lim_{d \downarrow a} \int_d^b f \in [-\infty, +\infty].$$

Uneigentliches Integral für offene Intervalle

Sei $f :]a, b[\rightarrow \mathbb{R}$ mit $-\infty \leq a < b \leq +\infty$. Dann setzen wir im Fall der Existenz

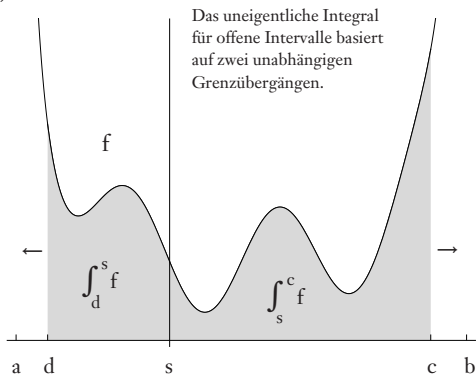
$$\int_a^b f = \int_a^s f + \int_s^b f \in [-\infty, +\infty],$$

mit $s \in]a, b[$ beliebig.

Existiert $r = \int_a^b f \in [-\infty, +\infty]$ für

eine auf einem halboffenen oder offenen Intervall definierte Funktion f , so heißt f *uneigentlich integrierbar* und r das *uneigentliche Integral* von f .

Das betrachtete Integral ist dabei das Riemann- oder das Regelintegral.



Die Formulierung „im Fall der Existenz“ beinhaltet sowohl die Existenz der Integrale auf allen beschränkten Teilintervallen $[c, d]$ des Definitionsbereichs als auch die der Grenzwerte, wobei letztere auch uneigentlich sein dürfen. Speziell muss f beschränkt auf allen Teilintervallen $[c, d]$ sein. Auch mit dem uneigentlichen Integral kann man also nicht über Polstellen hinwegintegrieren.

Für eine Funktion $f :]a, b[\rightarrow \mathbb{R}$ sind die folgenden drei Bedingungen notwendig und hinreichend für die uneigentliche Integrierbarkeit:

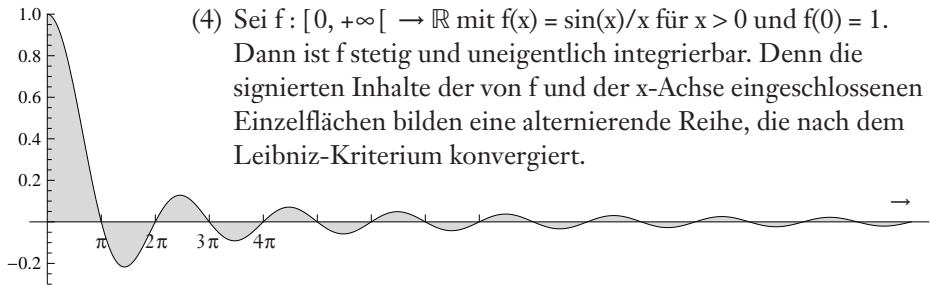
- (a) $f|_{[c, d]}$ ist integrierbar für alle $[c, d] \subseteq]a, b[$.
- (b) Es gibt ein $c \in]a, b[$, sodass $\int_a^c f$ und $\int_c^b f$ existieren.
- (c) $\int_a^c f + \int_c^b f$ ist nicht von der Form $-\infty + \infty$ oder $+\infty - \infty$.

Beispiele

$$(1) \int_0^1 \frac{1}{x} dx = \lim_{c \downarrow 0} \int_c^1 \frac{1}{x} dx = \lim_{c \downarrow 0} \log(x) \Big|_c^1 = +\infty.$$

$$(2) \int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{c \downarrow 0} \int_c^1 x^{-1/2} dx = \lim_{c \downarrow 0} 2x^{1/2} \Big|_c^1 = 2.$$

$$(3) \int_{-\infty}^{\infty} x dx \text{ existiert nicht, obwohl } \lim_{r \uparrow \infty} \int_{-r}^r x dx = 0.$$



- (4) Sei $f: [0, +\infty[\rightarrow \mathbb{R}$ mit $f(x) = \sin(x)/x$ für $x > 0$ und $f(0) = 1$. Dann ist f stetig und uneigentlich integrierbar. Denn die signierten Inhalte der von f und der x -Achse eingeschlossenen Einzelflächen bilden eine alternierende Reihe, die nach dem Leibniz-Kriterium konvergiert.

- (5) Sei $g: [-1, 1] \rightarrow \mathbb{R}$ definiert durch $g(x) = |x|^{-1/2}$ für $x \neq 0$ und $g(0) = 0$. Dann ist g unbeschränkt, also nicht integrierbar. Nach (2) schließt f die Fläche 4 mit der x -Achse ein. Dies kann man durch Beschneiden des Wertebereichs und

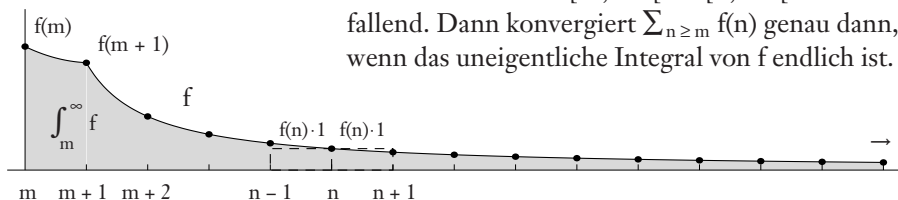
$$\lim_{s \uparrow \infty} \int_{-1}^1 \min(f(x), s) dx = 4$$

zum Ausdruck bringen. Dieser Grenzübergang ist aber kein Bestandteil des oben definierten uneigentlichen Integrals. Man könnte ihn als erneute Erweiterung des Integrals durchführen.

Uneigentliche Integrale führen zu einem neuen Konvergenzkriterium für Reihen:

Integralvergleichskriterium

Seien $m \in \mathbb{N}$ und $f: [m, +\infty[\rightarrow [0, +\infty[$ monoton fallend. Dann konvergiert $\sum_{n \geq m} f(n)$ genau dann, wenn das uneigentliche Integral von f endlich ist.



Die Endlichkeit des uneigentlichen Integrals ist äquivalent zur Endlichkeit von $\sum_{n \geq m} f(n) = \sum_{n \geq m} (f(n) \cdot 1)$.

Beispiele

$$(1) \text{ Die Reihe } \sum_{n \geq 1} 1/n \text{ divergiert, da } \int_1^{+\infty} 1/x dx = \lim_{c \rightarrow \infty} \log(c) = +\infty.$$

$$(2) \text{ Die Reihe } \sum_{n \geq 1} 1/n^2 \text{ konvergiert, da } \int_1^{+\infty} 1/x^2 dx = \lim_{c \rightarrow \infty} (1 - 1/c) = 1.$$

8.11 Der Vertauschungssatz

Satz (Vertauschungssatz für das Riemann- oder Regelintegral)

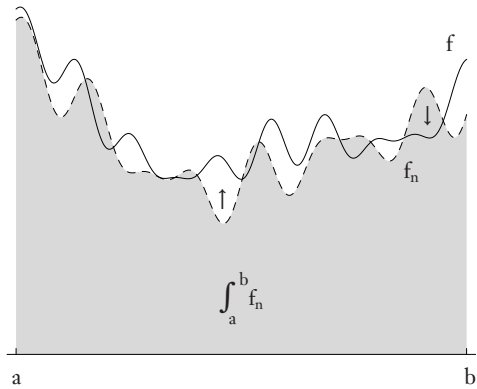
Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge integrierbarer Funktionen auf $[a, b]$, die punktweise gegen ein $f: [a, b] \rightarrow \mathbb{R}$ konvergiert. Dann gilt:

- (a) Ist $(f_n)_{n \in \mathbb{N}}$ beschränkt (d. h. gibt es ein $s \in [0, +\infty[$ mit $|f_n(x)| \leq s$ für alle n und alle $x \in [a, b]$) und f integrierbar, so gilt

$$\int_a^b f = \lim_n \int_a^b f_n.$$

(Vertauschungsregel für die Integration)

- (b) Gilt $\lim_n f_n = f$ (gleichmäßig), so ist $(f_n)_{n \in \mathbb{N}}$ beschränkt und f integrierbar. Folglich gilt die Vertauschungsregel in (a).



Der Satz ist ein weiteres Beispiel dafür, dass sich analytische Operationen unter „guten“ Bedingungen vertauschen lassen. Wir hatten schon die Limesstetigkeit

$$f(\lim_n x_n) = \lim_n f(x_n)$$

einer Funktion (vgl. 5.1) und die gliedweise Differenzierbarkeit

$$(\sum_n a_n x^n)' = \sum_n (a_n x^n)' = \sum_{n \geq 1} n a_n x^{n-1}$$

einer in $] -r, r[$ konvergenten Potenzreihe $\sum_n a_n x^n$ kennengelernt (vgl. 6.3, 7.6). Nun gilt

$$\int_a^b \lim_n f_n = \lim_n \int_a^b f_n,$$

vorausgesetzt, die Folge ist beschränkt und die Grenzfunktion $f = \lim_n f_n$ integrierbar. Beide Bedingungen gelten bei gleichmäßiger Konvergenz automatisch.

Legen wir das Riemann-Integral zugrunde und setzen wir statt des Integrals eine endliche Riemann-Summe mit einer Partition $p = (t_k, x_k)_{k \leq m}$ in die Vertauschungsregel ein, so erhalten wir

$$\sum_p \lim_n f_n = \lim_n \sum_p f_n, \text{ d. h.}$$

$$\sum_{k \leq m} \lim_n f_n(x_k) (t_{k+1} - t_k) = \lim_n \sum_{k \leq m} f_n(x_k) (t_{k+1} - t_k).$$

Der Satz besagt, dass sich diese endliche Summenregel in vielen Fällen auf das Integral überträgt. Analoges gilt für Treppenfunktionen und das Regelintegral.

Die Aussage (a) ist nicht einfach zu beweisen. Oft formuliert man den Satz deswegen nur für gleichmäßig konvergente Folgen, für die er einfach zu zeigen ist. Ein klareres Bild der Verhältnisse scheint aber die obige Formulierung zu liefern.

Beispiele

- (1) Sei $f: [-r, r] \rightarrow \mathbb{R}$, $f(x) = \sqrt{r^2 - x^2}$, die obere Hälfte des Einheitskreises. Weiter sei $(f_n)_{n \in \mathbb{N}}$ eine beschränkte Folge integrierbarer Funktionen auf $[-r, r]$, die punktweise gegen f konvergieren. Dann gilt

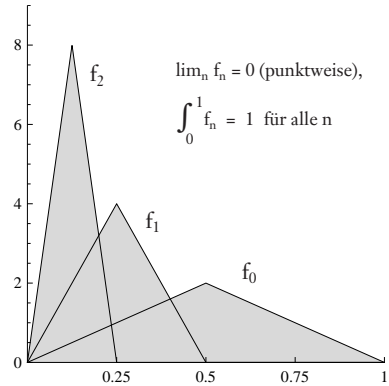
$$r^2 \pi/2 = \int_a^b f = \lim_n \int_a^b f_n.$$

Alle approximativen Verfahren zur Berechnung der Kreisfläche liefern also das gleiche Ergebnis (vgl. 8.9).

- (2) Für alle n sei $f_n: [0, 1] \rightarrow [0, +\infty[$ die stetige Zackenfunktion mit Zackenspitzen bei $(1/2^{n+1}, 2^{n+1})$ und Zackenbreiten 2^{-n} , d. h.,

$$\begin{aligned} f_n(x) &= 2^{2n+2} x && \text{in } [0, 1/2^{n+1}], \\ f_n(x) &= 2^{n+2} - 2^{2n+2} x && \text{in }]1/2^{n+1}, 1/2^n], \\ f_n(x) &= 0 && \text{in }]1/2^n, 1]. \end{aligned}$$

Dann schließt jedes f_n die Fläche 1 mit der x -Achse ein, und $(f_n)_{n \in \mathbb{N}}$ konvergiert punktweise gegen die Nullfunktion. Also ist



$$\int_a^b \lim_n f_n = \int_a^b 0 = 0 \neq 1 = \lim_n 1 = \lim_n \int_a^b f_n.$$

Auf die Beschränktheit der Folge kann also nicht verzichtet werden.

- (3) Seien $f_n: [0, +\infty[\rightarrow [0, 1]$, $n \geq 1$, Zackenfunktionen mit Zackenspitzen bei $(n/2, 1/n)$ und Zackenbreiten n . Dann gilt $\lim_n f_n = 0$ (gleichmäßig), aber das uneigentliche Integral von f_n ist gleich 1 für alle n . Die Vertauschungsregel gilt also nicht mehr für alle uneigentlichen Integrale.
- (4) Sei $(q_n)_{n \in \mathbb{N}}$ eine Aufzählung aller rationalen Zahlen in $[0, 1]$ (ohne Wiederholungen). Für alle n sei $f_n: [0, 1] \rightarrow [0, 1]$ definiert durch

$$f_n(q_k) = 1 \text{ für alle } k \leq n, \quad f_n(x) = 0 \text{ für alle } x \in [0, 1] - \{q_0, \dots, q_n\}.$$

Dann sind alle f_n integrierbar mit Integral 0, da f_n nur an endlich vielen Stellen von 0 verschieden ist. Die Folge $(f_n)_{n \in \mathbb{N}}$ konvergiert punktweise monoton steigend gegen die Dirichletsche Sprungfunktion auf $[0, 1]$. Die Integrierbarkeit der Grenzfunktion kann also bei punktwieser Konvergenz verloren gehen.

Situationen wie in Beispiel (2) sind unvermeidlich, aber der mögliche Verlust der Integrierbarkeit bei punktwieser beschränkter Konvergenz wie in Beispiel (4) ist eine Schwäche des Regel- und des Riemann-Integrals. In der höheren Analysis arbeitet man mit dem Lebesgue-Integral, das diese Schwäche nicht besitzt. Es setzt das Riemann-Integral echt fort. Die Dirichletsche Sprungfunktion ist Lebesgue-integrierbar mit Integral 0. Veränderungen an abzählbar vielen Stellen stören dieses Integral nicht.

8.12 Integral und Flächeninhalt

Definition (*Jordan-Inhalt*)

- (a) Ein $Q \subseteq \mathbb{R}^2$ heißt ein δ -Quadrat, falls es $\delta > 0$ und es $a, b \in \mathbb{Z}$ gibt mit

$$Q = [a\delta, (a+1)\delta] \times [b\delta, (b+1)\delta].$$

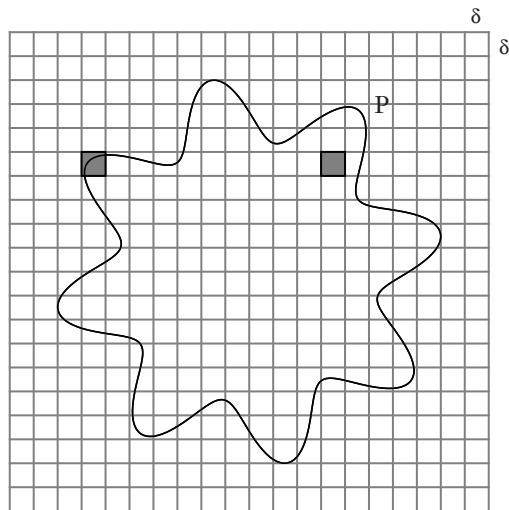
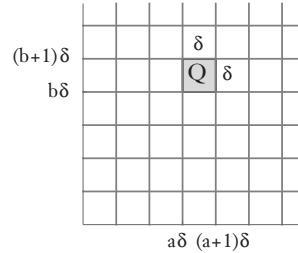
- (b) Sei $P \subseteq \mathbb{R}^2$ beschränkt, d. h., es gebe ein n mit $P \subseteq [-n, n]^2$. Dann heißen

$$J(P) = \inf_{\delta > 0} \delta^2 \cdot \text{„die Anzahl der } \delta\text{-Quadrate } Q \text{ mit } Q \cap P \neq \emptyset\text{“},$$

$$j(P) = \sup_{\delta > 0} \delta^2 \cdot \text{„die Anzahl der } \delta\text{-Quadrate } Q \text{ mit } Q \subseteq P\text{“}$$

der *äußere* bzw. *innere Jordan-Inhalt* von P .

- (c) Gilt $J(P) = j(P)$, so heißt P *Jordan-messbar* und $J(P)$ der *Jordan-Inhalt* von P .



Das linke δ -Quadrat trägt zu $J(P)$ bei, das rechte zu $j(P)$ und $J(P)$.

Die auf den ersten Blick vielleicht etwas technische Definition beruht auf der einfachen und anschaulichen Methode des Kästchenzählens, die wir von der Schule kennen. Ist eine Figur auf einem Millimeterpapier gegeben, so können wir den Flächeninhalt der Figur nach oben abschätzen, indem wir die Anzahl der Kästchen bestimmen, die mit der Figur mindestens einen Punkt gemeinsam haben. Analog können wir den Flächeninhalt der Figur nach unten abschätzen, indem wir die Anzahl der Kästchen bestimmen, die ganz innerhalb der Figur liegen. Wenn wir nun abstrahieren und beliebige kleine positive reelle Zahlen δ als Rastermaß und beliebige beschränkte Teilmengen P der Ebene als Figuren betrachten, so gelangen wir zur obigen Begriffsbildung.

Man darf vermuten, dass zwischen Inhalt und Integral ein enger Zusammenhang besteht. Das Integral haben wir ja bereits als eine „signierte Flächenmessung“ bezeichnet. Aber es gibt doch Unterschiede. Im Integral arbeiten wir mit Rechtecksflächen und nicht mit den Quadraten eines Gitters. Weiter tragen negative Bereiche negativ zum Integral bei, während beim Jordan-Inhalt alles positiv zählt. Der dritte und bedeutsamste Unterschied ist, dass der Jordan-Inhalt auch für Mengen P erklärt ist, die nicht zwischen einer Funktion und der x -Achse liegen, etwa für Kreise, Ellipsen, Kleeblätter, teichförmige Bereiche usw. Trotz dieser Unterschiede stellt sich heraus, dass Riemann-Integral und Jordan-Inhalt (und nicht etwa Regelintegral und Jordan-Inhalt!) gleichwertig sind. Die erste Erkenntnis hierzu ist:

Riemann-Integral für nichtnegative Funktionen und Jordan-Inhalt

Sei $f : [a, b] \rightarrow [0, +\infty[$ eine beschränkte Funktion, und sei

$$P_f = \{ (x, y) \in \mathbb{R}^2 \mid x \in [a, b], y \in [0, f(x)] \}$$

die von f und der x -Achse eingeschlossene Fläche. Dann ist f genau dann Riemann-integrierbar, wenn P_f Jordan-messbar ist. In diesem Fall gilt

$$\int_a^b f = J(P_f).$$

Damit ist der „geometrische Gehalt“ des Riemann-Integrals präzisiert: Das Riemann-Integral einer nichtnegativen Funktion ist nicht weniger und nicht mehr als der mit der Kästchenmethode berechnete Inhalt der von f definierten Fläche P_f . Prinzipiell könnte die Kästchenmethode ja auch dem engeren Regelintegral entsprechen oder sogar über das Riemann-Integral hinausgehen. Dies ist aber nicht der Fall.

Wie sieht es nun mit Funktionen aus, die negative Werte annehmen können? Der Jordan-Inhalt kann hier leicht mithalten, denn es gilt

$$\int_a^b f = \int_a^b f^+ - \int_a^b f^- = J(P_{f^+}) - J(P_{f^-})$$

mit dem Positivteil f^+ und dem Negativteil f^- von f (vgl. 8.6). Damit haben wir:

Der Jordan-Inhalt kann alles, was das Riemann-Integral kann.

Wie sieht es mit der Umkehrung aus? Kann man mit dem Jordan-Inhalt etwas messen, was man mit dem Riemann-Integral nicht messen kann? Die Antwort ist ja, und sie beruht auf dem dritten oben genannten Unterschied. Mit dem Riemann-Integral kann man zwar noch relativ bequem die Fläche eines Kreises mit Radius r berechnen, indem man die Fläche der oberen Hälfte des Kreises berechnet,

$$\int_{-r}^r \sqrt{r^2 - x^2} \, dx = r^2 \pi / 2 \quad (\text{vgl. 8.9})$$

und das Ergebnis verdoppelt. Aber die Berechnung von komplizierteren Flächen, die nicht durch Funktionen definiert werden, ist schwer bis unmöglich. Damit hat der Jordan-Inhalt zunächst einen theoretischen Vorsprung. Sobald man aber das Riemann-Integral auf die Dimension 2 erweitert hat, kann man diesen Vorsprung wettmachen. Ist nämlich $P \subseteq [-n, n]^2$ eine Jordan-messbare Menge, so ist die Funktion $1_P : [-n, n]^2 \rightarrow \mathbb{R}$ mit $1_P(x) = 1$ für $x \in P$ und $1_P(x) = 0$ für $x \notin P$ Riemann-integrierbar und das Riemann-Integral von 1_P ist gleich $J(P)$. Damit gilt:

Das zweidimensionale Riemann-Integral kann alles, was der Jordan-Inhalt kann.

Auch den Jordan-Inhalt kann man zu einem dreidimensionalen Jordan-Volumen erweitern. Das Jordan-Volumen kann dann wieder alles, was das zweidimensionale Riemann-Integral kann. Dies setzt sich fort und damit sind, im Hinblick auf höhere Dimensionen, das Riemann-Integral und der Jordan-Inhalt vollkommen äquivalent. Einen eleganten und effektiven Rechenkalkül besitzt nur das Integral.

Zum Studium der Mathematik

Vorbemerkung

Die folgenden Ausführungen erheben nicht den Anspruch eines allgemeinen Rezepts für einen erfolgreichen Start ins Mathematikstudium. Sie beruhen auf Lehrerfahrungen und persönlichen Einstellungen und wollen die fachliche Hilfestellung, die dieses Buch anstrebt, durch Hinweise zum Studium ergänzen.

Kennzeichen der Wissenschaft Mathematik ...

... ist der Beweis. Ein wissenschaftliches Studium der Mathematik spiegelt dieses Kennzeichen wider. Beweise werden oft mit Hilfe von Anschauungen gefunden und verstanden, aber nicht anschaulich geführt und notiert. Sie beruhen auf exakten Begriffsbildungen in Form von fachsprachlichen Definitionen. Das alles muss man mögen und noch besser lieben, wenn man Mathematik studieren möchte.

Vorbereitung und Studienbeginn

Viele Universitäten bieten mathematische Brückenkurse an. Man lernt dort das vergleichsweise hohe Tempo kennen, sammelt nützliches Wissen und kommt in Kontakt mit Dozenten und Tutoren, die über das Studium an der Universität informieren. Der Besuch eines solchen Kurses ist unbedingt zu empfehlen. Daneben bietet der Buchmarkt Texte für den Übergang zur Universität an. Wer an einem Brückenkurs nicht teilnehmen kann, möge den entsprechenden Dozenten per E-Mail um Lernmaterial oder Buchempfehlungen bitten.

Das Studium beginnt mit Einführungstagen, Gruppeneinteilung und ersten sozialen Kontakten. Danach bestimmen die mathematischen Grundvorlesungen und die wöchentlichen Übungsaufgaben den Studienalltag. Während jede einzelne Vorlesung überschaubar ist, häuft sich innerhalb von wenigen Wochen eine beachtliche Stoffmenge an. Wo man vielleicht Beispiele und Rechnungen erwarten würde, kommen sofort wieder neue Begriffe und Beweise. Viele der Übungsaufgaben beginnen mit „Zeigen Sie ...“ und man muss rasch lernen, damit umzugehen. Hierfür gibt es Tutorien, Kommilitonen, Skripte, Bücher – und den eigenen Kopf. Die beste Vorbereitung auf dieses „Zeigen Sie ...“ ist die Nachbereitung des Skripts in einer ruhigen Arbeitsatmosphäre. Die Beweise des Skripts sind in vielen Fällen Beispiel und Vorbild für die Beweise der Übungsaufgaben. Nachahmung ist erwünscht, auch stilistisch. Die angefertigten Lösungen der Übungsaufgaben sollen in der Sprache und Genauigkeit formuliert werden, die in der Vorlesung verwendet wird. Nach dem Finden der Lösung auf einem Schmierzettel steht noch das lesbare und strukturierende Aufschreiben bevor.

Um den Überblick zu behalten, kann man sich angewöhnen, jede Vorlesung in eigenen Worten zusammenzufassen: Was haben wir gemacht? Was war besonders wichtig? Ein Beispiel wäre: „Wir haben punktweise konvergente Funktionenfolgen betrachtet. Ein Beispiel zeigte, dass die Grenzfunktion einer punktweise konvergenten Folge stetiger Funktionen nicht mehr stetig sein muss. Daraufhin haben wir die punktweise Konvergenz zur gleichmäßigen Konvergenz verstärkt (Anschauung: ε -Schlauch). Nun konnten wir zeigen, dass die Stetigkeit erhalten bleibt (Stichwort: $\varepsilon/3$ -Argument).“

Lernverhalten

Ein punktweises und prüfungsadaptives Lernen, bei dem das Vergessen beim Antworten beginnt, funktioniert an der Universität nicht. Die Modulprüfungen, bei denen man auf sich allein gestellt ist, finden am Ende des Semesters statt, die wöchentlichen Übungsaufgaben kann man in Gruppen bearbeiten. Kontinuierliches und effizientes Arbeiten ist gefragt. Die Lernkurve ist entscheidend, nicht die kurzzeitige Leistungsspitze. Das Gras wächst nicht schneller, wenn man kurz vor der Klausur daran zieht. Sinnvoll ist es dagegen, die Graspflege – das eigene Lernverhalten – von Beginn an zu überprüfen und anzupassen. Befragungen zeigen, dass sehr viel Zeit in Vorlesungen und in Zusatzangeboten verbracht wird, während die Lernzeit „im stillen Kämmerlein“ sehr gering ausfällt. Und wöchentlich zehn Stunden für vier Übungsaufgaben und nur eine halbe Stunde für das Studieren von Skript und Lehrbüchern zu verwenden, ist in der Regel keine gute Zeiteinteilung.

Anschauung und Fachsprache

Um ein freudvolles Dasein als Mathematiker führen zu können, muss man eine zweifache Übersetzungs- oder Transformationsfertigkeit erwerben und einüben. Zum einen muss man abstrakte und in einer formalen Sprache formulierte Begriffe und Ergebnisse mit Anschauungen füllen können. Sonst bleiben sie formal und leblos. Zum anderen muss man eigene individuelle Anschauungen in die formale Sprache übersetzen können. Sonst bleibt der mathematische Gehalt der Anschauungen unscharf und die Kommunikation mit anderen Mathematikern wird behindert. Beide Fertigkeiten kann man trainieren, die erste durch Diagramme, Beispiele und ein Sprechen über Mathematik, die zweite durch das genaue Wiedergeben von Definitionen und Sätzen und durch eine Verschriftlichung von Mathematik, die stilistische Aspekte und Details der Notation beachtet. Vorlesungen und systematische Lehrbücher, die eine mathematische Theorie umfassend vorstellen wollen, sind nicht der Ort, jeden Begriff anschaulich zu erklären. Das geschieht letztendlich im eigenen Kopf. Je kreativer und individueller, desto besser.

Viele mathematische Begriffe können von verschiedenen Seiten aus betrachtet werden und zu einem Begriff können verschiedene kontextabhängige Auffassungen und Anschauungen nebeneinander existieren, die sich überschneiden können. Ein Beispiel hierfür ist die Ableitung einer Funktion f in einem Punkt. Man kann sie als Zahl sehen, mit der gerechnet werden kann, und die mit anderen Ableitungszahlen eine Funktion f' bildet. Man kann aber auch die Tangente an f im betrachteten Punkt in den Vordergrund rücken, die die Funktion in diesem Punkt linearisiert, und dann die Abweichung von f und dieser Tangente studieren. Ebenso können Matrizen sowohl mit linearen Gleichungssystemen assoziiert werden als auch mit linearen Abbildungen, mit algebraischen Ringen, mit Graphen und kombinatorischen oder stochastischen Fragestellungen.

Prüfungsvorbereitung

Zur Vorbereitung auf eine Klausur kann man – aufbauend auf kontinuierliche Arbeit während der Vorlesungszeit – wie folgt vorgehen:

- (a) Man wiederholt den gesamten Stoff der Vorlesung, wobei man sich auf die wesentlichen Themen konzentriert. Eine Bewertung in „sehr wichtig, wichtig, weniger wichtig“ vorzunehmen ist nicht leicht und in gewisser Weise bereits Teil der Prüfungsleistung. Dieses Buch möchte eine Hilfe hierzu sein.
- (b) Viele Klausuraufgaben sind modifizierte Übungsaufgaben. Es empfiehlt sich, alle Übungen noch einmal genau durchzugehen.
- (c) Von Kommilitonen, dem Tutor oder der Fachschaft kann man Klausuren vergangener Semester erhalten. Manchmal werden „reale Klausuraufgaben“ auch in den Tutorien diskutiert.
- (d) Oft darf man keine Hilfsmittel zur Klausur mitnehmen, zuweilen ist eine Seite mit Definitionen und Formeln erlaubt. Die Aufgaben bei der ersten Variante sind in der Regel etwas einfacher, und hier können auch Definitionen und Sätze abgefragt werden. Bei der zweiten Variante wird in der Regel die Anwendung von Begriffen und Ergebnissen etwas stärker betont.

In mündlichen Prüfungen sind eine Sicherheit in den Grundbegriffen sowie ein Überblick über den Verlauf und Aufbau der Vorlesung besonders wichtig. Das Durcharbeiten der Übungsaufgaben ist weniger bedeutsam, das Lernen mit Skript und Büchern dagegen sehr. Mündliche Prüfungen beginnen oft mit einer Definitionsfrage, an die sich Verständnis- und Einordnungsfragen anschließen. (Auch die Eingangsfrage „Worüber wollen Sie reden?“ kommt vor.) Beweise muss man nicht auswendig lernen, Beweisideen zu zentralen Sätzen sollte man aber angeben und diskutieren können.

Anhänge

1. Junktoren

Zeichen	Bedeutung	Name
\neg	nicht/non	Negation
\wedge	und	Konjunktion
\vee	oder	Disjunktion
\rightarrow	impliziert	Implikation
\leftrightarrow	genau dann, wenn	Äquivalenz

Die Tabelle zeigt die fünf wichtigsten Junktoren (logische Verknüpfungen) der Mathematik. Ihre Semantik kann durch sog. *Wahrheitstafeln* erläutert werden, mit „w“ für „wahr“ und „f“ für „falsch“:

\neg	A
f	w
w	f

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

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

A	\rightarrow	B
w	w	w
w	f	f
f	w	w
f	w	f

A	\leftrightarrow	B
w	w	w
w	f	f
f	f	w
f	w	f

So gilt eine Disjunktion $A \vee B$ („A oder B“) in der Mathematik als wahr, wenn mindestens eine der beiden Aussagen wahr ist. Und eine Implikation $A \rightarrow B$ („A impliziert B“) gilt in der Mathematik nur dann als falsch, wenn A wahr und B falsch ist.

Prinzipiell kann man sich auf die Negation und die Konjunktion beschränken, denn die anderen Junktoren lassen sich mit Hilfe von \neg und \wedge definieren:

$$A \vee B \text{ als } \neg(\neg A \wedge \neg B), \quad A \rightarrow B \text{ als } \neg A \vee B, \quad A \leftrightarrow B \text{ als } (A \rightarrow B) \wedge (B \rightarrow A).$$

Dass $A \rightarrow B$ äquivalent zu $\neg A \vee B$ ist, zeigt den „statischen“ Charakter der Implikation besonders deutlich. $A \rightarrow B$ drückt keine kausale Beziehung zwischen A und B aus.

Neben „A impliziert B“ sagen wir auch:

„Aus A folgt B“, „A ist hinreichend für B“, „B ist notwendig für A“.

Mathematische Sätze besagen, dass unter einer bestimmten Voraussetzung A eine Aussage B gilt. Es handelt sich also um Aussagen der Form $A \rightarrow B$. Um eine solche Implikation zu beweisen, nimmt man A an und versucht mit Hilfe dieser Annahme und bereits bewiesener Resultate durch logische Argumentation die Aussage B zu beweisen. Eine andere Beweisführung von $A \rightarrow B$, bei der man mit $\neg B$ als Annahme anfängt und $\neg A$ anstrebt, liefert das Kontrapositionsgesetz (siehe unten).

Häufig gebraucht werden die Verneinungsregeln für die Junktoren:

Die Aussage ...	ist äquivalent zu ...	Name der Regel
$\neg \neg A$	A	doppelte Verneinung
$\neg (A \wedge B)$ $\neg (A \vee B)$	$\neg A \vee \neg B$ $\neg A \wedge \neg B$	De Morgansche Regeln
$\neg (A \rightarrow B)$	$A \wedge \neg B$	Verneinung einer Implikation
$\neg (A \leftrightarrow B)$	entweder A oder B (exklusiv)	Verneinung einer Äquivalenz

Junktoren können wir mehrfach anwenden und zum Beispiel $(A \rightarrow B) \wedge (B \rightarrow (C \vee D))$ bilden. Um Klammern zu sparen, vereinbart man von stark nach schwach bindend:

$$\neg, \wedge, \vee, \rightarrow, \leftrightarrow \quad (\text{Bindungsstärke der Junktoren})$$

So ist etwa $A \vee \neg B \rightarrow \neg C \wedge D$ die Aussage $(A \vee (\neg B)) \rightarrow ((\neg C) \wedge D)$.

Zusammengesetzte Aussagen kann man wieder mit Wahrheitstafeln untersuchen. Wir illustrieren das Verfahren am Beispiel von

$$A \rightarrow B \leftrightarrow \neg B \rightarrow \neg A. \quad (\text{Kontrapositionsgesetz})$$

A	\rightarrow	B	\leftrightarrow	\neg	B	\rightarrow	\neg	A
w	w	w	w	f	w	w	f	w
w	f	f	w	w	f	f	f	w
f	w	w	w	f	w	w	w	f
f	w	f	w	w	f	w	w	f

Wahrheitstafel für das Kontrapositionsgesetz. Die Zahlen geben die Reihenfolge der Berechnung der Spalten an.

1 5 2 4 3

Tritt in der zuletzt berechneten Spalte nur der Wahrheitswert „w“ auf, so nennt man die betrachtete Aussage eine *Tautologie* oder *allgemeingültig*. Das Kontrapositionsgesetz ist also eine Tautologie. Es wird in der Mathematik in *indirekten Beweisen* verwendet: Anstelle von $A \rightarrow B$ kann man immer $\neg B \rightarrow \neg A$ zeigen, was einfacher sein kann.

Beispiele

Die folgenden Aussagen sind Tautologien:

$$(1) \neg \neg A \leftrightarrow A, \quad \neg (A \wedge B) \leftrightarrow \neg A \vee \neg B, \quad \neg (A \rightarrow B) \leftrightarrow A \wedge \neg B.$$

$$(2) A \vee \neg A \quad (\text{tertium non datur}), \quad A \wedge (A \rightarrow B) \rightarrow B \quad (\text{modus ponens}),$$

$$(B \rightarrow A) \wedge (\neg B \rightarrow A) \leftrightarrow A \quad (\text{Fallunterscheidung}).$$

2. Quantoren

Zeichen	Bedeutung	Name
\forall	für alle	Allquantor
\exists	es gibt ein	Existenzquantor

In der Mathematik werden vor allem der Allquantor „für alle, für jedes“ und der Existenzquantor „es gibt, es existiert ein“ – im Sinne von „es gibt mindestens ein“ – verwendet. So bedeuten

$$\exists x \cos(x) = 0, \quad \forall x \, x < x + 1, \quad \forall n \, \exists p \, (p \geq n \wedge p \text{ ist eine Primzahl}),$$

dass der Kosinus eine Nullstelle besitzt, dass eine reelle Zahl x stets kleiner als $x + 1$ ist und dass es unendlich viele Primzahlen gibt. Dass sich x auf die reellen und n und p auf die natürlichen Zahlen beziehen, ist in konkreten Fällen aus dem Kontext heraus klar. Zur Verdeutlichung kann man $\forall x \in \mathbb{R}, \forall n \in \mathbb{N}, \dots$ schreiben.

Für die Quantoren gelten die folgenden Verneinungs- und Vertauschungsregeln, die wir in 1.10 genauer besprechen:

$\neg \forall x A(x)$	\leftrightarrow	$\exists x \neg A(x)$
$\neg \exists x A(x)$	\leftrightarrow	$\forall x \neg A(x)$
$\forall x \forall y A(x, y)$	\leftrightarrow	$\forall y \forall x A(x, y)$
$\exists x \exists y A(x, y)$	\leftrightarrow	$\exists y \exists x A(x, y)$
$\exists x \forall y A(x, y)$	\rightarrow	$\forall y \exists x A(x, y)$

Nützlich sind zuweilen auch die folgenden Regeln, die angeben, wann man Quantoren in mit Junktoren zusammengesetzte Aussagen rein- und rausziehen darf:

$\forall x (A(x) \wedge B(x))$	\leftrightarrow	$\forall x A(x) \wedge \forall x B(x)$
$\exists x (A(x) \vee B(x))$	\leftrightarrow	$\exists x A(x) \vee \exists x B(x)$
$\forall x A(x) \vee \forall x B(x)$	\rightarrow	$\forall x (A(x) \vee B(x))$
$\exists x (A(x) \wedge B(x))$	\rightarrow	$\exists x A(x) \wedge \exists x B(x)$
$\forall x (A(x) \rightarrow B(x))$	\rightarrow	$\forall x A(x) \rightarrow \forall x B(x)$
$\exists x (A(x) \rightarrow B(x))$	\leftrightarrow	$\forall x A(x) \rightarrow \exists x B(x)$

3. Axiome für die reellen Zahlen (2.4 – 2.7)

(K1)	$x + (y + z) = (x + y) + z$	<i>Assoziativgesetz für +</i>
(K2)	$x + 0 = x$	<i>Neutralität von 0</i>
(K3)	$\exists x' \ x + x' = 0$	<i>Inverse für +</i>
(K4)	$x + y = y + x$	<i>Kommutativgesetz für +</i>
(K5)	$x \cdot (y \cdot z) = (x \cdot y) \cdot z$	<i>Assoziativgesetz für ·</i>
(K6)	$x \cdot 1 = x$	<i>Neutralität von 1</i>
(K7)	$x \neq 0 \rightarrow \exists x' \ x \cdot x' = 1$	<i>Inverse für ·</i>
(K8)	$x \cdot y = y \cdot x$	<i>Kommutativgesetz für ·</i>
(K9)	$x \cdot (y + z) = (x \cdot y) + (x \cdot z)$	<i>Distributivgesetz</i>
(K10)	$0 \neq 1$	<i>Verschiedenheit von 0 und 1</i>
(K11)	$\neg x < x$	<i>Irreflexivität</i>
(K12)	$x < y \wedge y < z \rightarrow x < z$	<i>Transitivität</i>
(K13)	$x < y \vee x = y \vee y < x$	<i>Vergleichbarkeit</i>
(K14)	$x < y \rightarrow x + z < y + z$	<i>Addition und Ordnung</i>
(K15)	$0 < x \wedge 0 < y \rightarrow 0 < x \cdot y$	<i>Multiplikation und Ordnung</i>
(K16)	Ist X nichtleer und nach oben beschränkt, so existiert $\sup(X)$.	<i>(lineare) Vollständigkeit</i>

In (K1) – (K15) sind x, y, z beliebige reelle Zahlen. In (K16) ist X eine beliebige Menge von reellen Zahlen. Das Axiom (K16) kann man durch die beiden folgenden Axiome ersetzen:

(K16)'	Ist $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge, so existiert $\lim_n x_n$.	<i>(metrische) Vollständigkeit</i>
(K17)	$\forall x > 0 \ \forall y > 0 \ \exists n \in \mathbb{N} \ n x > y$	<i>Archimedisches Axiom</i>

4. Epsilontik

Grenzwerte für Folgen (3.2, 3.5, 3.7, 3.12)

$\lim_n x_n = x$	$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 x_n - x < \varepsilon$
$(x_n)_{n \in \mathbb{N}}$ divergiert	$\forall x \exists \varepsilon > 0 \forall n_0 \exists n \geq n_0 x_n - x \geq \varepsilon$
$(x_n)_{n \in \mathbb{N}}$ ist eine Cauchy-Folge	$\forall \varepsilon > 0 \exists n_0 \forall n, m \geq n_0 x_n - x_m < \varepsilon$
x ist Häufungspunkt von $(x_n)_{n \in \mathbb{N}}$	$\forall \varepsilon > 0 \forall n_0 \exists n \geq n_0 x - x_n < \varepsilon$
$\lim_n x_n = \infty, \quad \lim_n x_n = -\infty$	$\forall x \exists n_0 \forall n \geq n_0 x_n \geq x, \quad \forall x \exists n_0 \forall n \geq n_0 x_n \leq x$

Häufungs- und Berührungpunkte von Mengen (3.10, 5.2)

x ist Häufungspunkt von P	$\forall \varepsilon > 0 \ P \cap U_\varepsilon(p)$ ist unendlich (mit $U_\varepsilon(p) =]p - \varepsilon, p + \varepsilon[$)
x ist Berührungpunkt von P	$\forall \varepsilon > 0 \ P \cap U_\varepsilon(p)$ ist nichtleer

Grenzwerte für Reihen (4.1, 4.6)

$\sum_n x_n = x$	$\lim_n \sum_{k \leq n} x_k = x$
Cauchy-Kriterium für Reihen	$\forall \varepsilon > 0 \exists n_0 \forall n \geq m \geq n_0 \sum_{m \leq k \leq n} x_k < \varepsilon$

Grenzwerte für Funktionen (5.2)

$\lim_{x \rightarrow p} f(x) = y,$ $f: P \rightarrow \mathbb{R}, p$ Berührungpunkt von P	$\forall (x_n)_{n \in \mathbb{N}} \text{ in } P (\lim_n x_n = p \rightarrow \lim_n f(x_n) = y)$
---	---

Stetigkeit von Funktionen (5.1, 5.4 – 5.6)

$f : P \rightarrow \mathbb{R}$ ist limesstetig in p	$\forall (x_n)_{n \in \mathbb{N}} \text{ in } P \text{ (} \lim_n x_n = p \rightarrow \lim_n f(x_n) = f(p) \text{)}$ oder gleichwertig $\lim_{x \rightarrow p} f(x) = f(p).$
$f : P \rightarrow \mathbb{R}$ ist ε - δ -stetig in p	$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in P \text{ (} x - p < \delta \rightarrow f(x) - f(p) < \varepsilon \text{)}$
$f : P \rightarrow \mathbb{R}$ ist gleichmäßig stetig	$\forall \varepsilon > 0 \exists \delta > 0 \forall p, x \in P \text{ (} x - p < \delta \rightarrow f(x) - f(p) < \varepsilon \text{)}$
$f : P \rightarrow \mathbb{R}$ ist Lipschitz-stetig	$\exists L \geq 0 \forall p, x \in P \text{ (} f(x) - f(p) \leq L x - p \text{)}$

Konvergenz von Funktionenfolgen (5.11)

$\lim_n f_n = f$ (punktweise) für $f, f_n : P \rightarrow \mathbb{R}$	$\forall x \in P \lim_n f_n(x) = f(x), \text{ d. h.}$ $\forall x \in P \forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 f_n(x) - f(x) < \varepsilon$
$\lim_n f_n = f$ (gleichmäßig) für $f, f_n : P \rightarrow \mathbb{R}$	$\forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 \forall x \in P f_n(x) - f(x) < \varepsilon$

Wert des Integrals (8.2 – 8.3)

$\int_a^b f(x) dx = c$	$\forall \varepsilon > 0 \exists \delta > 0 \forall p \text{ (} \delta(p) \leq \delta \rightarrow \sum_p f - c < \varepsilon \text{)}$ oder gleichwertig $\forall (p_n)_{n \in \mathbb{N}} \text{ (} \lim_n \delta(p_n) = 0 \rightarrow \lim_n \sum_{p_n} f = c \text{)}$ oder gleichwertig $\forall \varepsilon > 0 \exists p_1, p_2 \sum_{p_1} f - \sum_{p_2} f < \varepsilon$
------------------------	--

5. Grenzwerte von Folgen und unendliche Summen

$k, d_1 d_2 d_3 \dots d_n \dots = \lim_{n \geq 1} k, d_1 \dots d_n$	3.2
$[a_0, a_1, \dots, a_n, \dots] = \lim_n [a_0, \dots, a_n]$	3.3
$\lim_n (\sum_{1 \leq k \leq n} 1/k - \log(n)) = C$	4.5
$\lim_{n \geq 1} (1 + 1/n)^n = e$ $\lim_{n \geq 1} (1 + x_n/n)^n = e^x$, falls $\lim_n x_n = x$	3.2, 4.10, 6.3
$\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$	6.5
$\lim_{n \geq 1} 2n \sin(\pi/n) = 2\pi$	6.12

$\sum_n (-1)^n$ divergiert	4.1
$\sum_{n \geq 1} 1/(n(n+1)) = 1$ $\sum_{n \geq 1} 1/(n(n+1)(n+2)) = 1/4$	4.2, 4.9, 6.2
$\sum_n x^n = 1/(1-x)$ $\sum_{n \geq 1} x^n = x/(1-x)$ $\sum_{n \geq 1} n x^n = x/(1-x)^2$ $\sum_{n,m} x^n y^m = 1/((1-x)(1-y))$	4.3, 4.11, 4.12, 5.12
$0, d_1 d_2 \dots d_n \dots = \sum_{n \geq 1} d_n / 10^n$ $0, \overline{b_1 \dots b_p} = b_1 \dots b_p / 9 \dots 9$ (p Neunen)	4.4
$\sum_{n \geq 1} 1/n = \infty, \sum_{p \text{ prim}} 1/p = \infty$	4.5, 4.10, 8.10
$\sum_n 1/n^2 = \pi^2/6,$	3.5, 4.5, 4.9, 4.10, 8.10
$\sum_{n \geq 1} (-1)^{n-1}/n = \log(2)$	4.7, 4.8, 5.12, 6.4
$\sum_n (-1)^n/(2n+1) = \pi/4$	3.3, 4.7

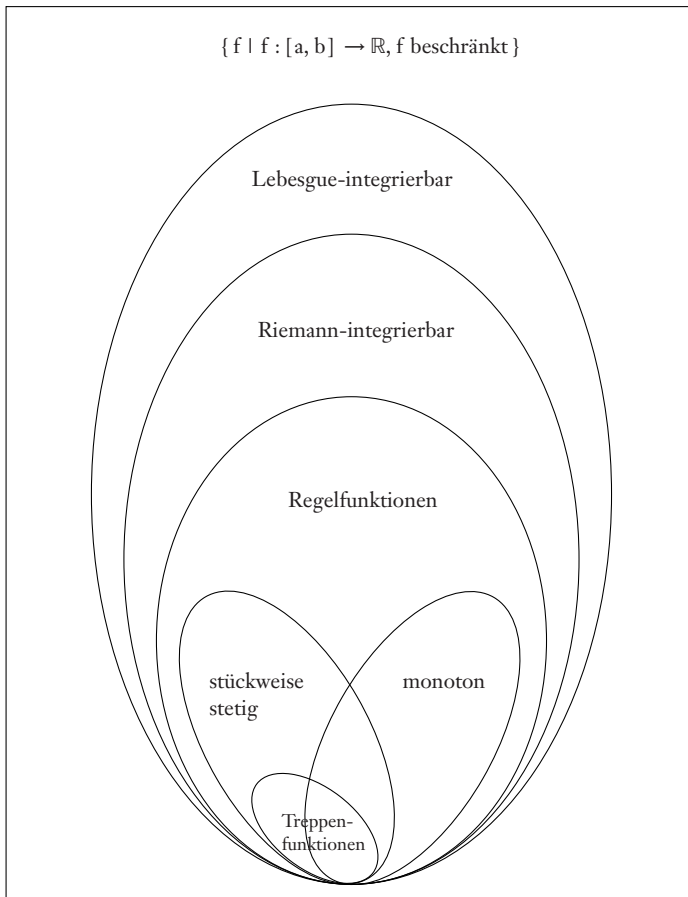
6. Reihenentwicklungen

Funktion und Reihendarstellung	gültig für	
$\frac{1}{1-x} = \sum_n x^n$	$ x < 1$	4.3, 4.12, 5.12
$\frac{x}{(1-x)^2} = \sum_{n \geq 1} n x^n$	$ x < 1$	4.3, 4.11
$\exp(x) = \sum_n \frac{x^n}{n!}$	$x \in \mathbb{R}$	4.12, 5.12, 6.3
$\exp(z) = \sum_n \frac{z^n}{n!}$	$z \in \mathbb{C}$	4.12, 6.7, 6.8
$\log(1+x) = \sum_{n \geq 1} (-1)^{n-1} \frac{x^n}{n}$	$x \in]-1, 1]$	5.12, 6.4
$\cos(x) = \sum_n (-1)^n \frac{x^{2n}}{(2n)!}$	$x \in \mathbb{R}$	6.9
$\sin(x) = \sum_n (-1)^n \frac{x^{2n+1}}{(2n+1)!}$	$x \in \mathbb{R}$	6.9
$f = T_p f(x) = \sum_n \frac{f^{(n)}(p)}{n!} (x-p)^n$	abhängig von f	7.12

7. Ableitungen

$d/dx \ a^x$	$\log(a) \ a^x$
$d/dx \ x^a$	$a \ x^{a-1}$
$d/dx \ \log_a(x)$	$\frac{1}{x \log(a)}$
$d/dx \ \log_x(a)$	$-\frac{\log_x(a)}{x \log(x)}$
$d/dx \ \sin(x)$	$\cos(x)$
$d/dx \ \cos(x)$	$-\sin(x)$
$d/dx \ \tan(x)$	$\frac{1}{\cos^2(x)}$
$d/dx \ \cot(x)$	$-\frac{1}{\sin^2(x)}$
$d/dx \ \sec(x)$	$\frac{\sin(x)}{\cos^2(x)}$
$d/dx \ \csc(x)$	$-\frac{\cos(x)}{\sin^2(x)}$
$d/dx \ \arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$
$d/dx \ \arccos(x)$	$-\frac{1}{\sqrt{1-x^2}}$
$d/dx \ \arctan(x)$	$\frac{1}{1+x^2}$
$d/dx \ \operatorname{arccot}(x)$	$-\frac{1}{1+x^2}$

8. Integrierbare Funktionen



Für ein reelles Intervall $[a, b]$ mit $a < b$ unterscheiden sich die verschiedenen Integralbegriffe durch die Menge der Funktionen, die integriert werden können (und nicht etwa durch die Integrale, falls diese existieren). Alle stückweise stetigen und alle monotonen Funktionen sind regelintegrierbar, alle Regelfunktionen sind Riemann-integrierbar und alle Riemann-integrierbaren Funktionen sind Lebesgue-integrierbar. Die Inklusionen sind echt: Es gibt ein Riemann-integrierbares f , das keine Regelfunktion ist, ein Lebesgue-integrierbares g , das nicht Riemann-integrierbar ist, und eine Funktion h , die nicht Lebesgue-integrierbar ist.

Literatur

- Amann, Herbert / Escher, Joachim** 2008 *Analysis I*. Birkhäuser, Basel, 3. Auflage 2008.
- Appell, Jürgen** 2008 *Analysis in Beispielen und Gegenbeispielen*. Springer, Berlin, 2008.
- Behrends, Ehrhard** 2011 *Analysis I*. Vieweg+Teubner, Wiesbaden, 5. Auflage 2011.
- Bornemann, Folkmar** 2008 *Konkrete Analysis für Studierende der Informatik*. Springer, Berlin, 2008.
- Bosch, Siegfried** 2009 *Lineare Algebra*. Springer, Berlin, 4. Auflage 2009.
- Deiser, Oliver** 2008 *Reelle Zahlen*. Springer, Berlin, 2. Auflage 2008.
- 2009 *Einführung in die Mengenlehre*. Springer, Berlin, 3. Auflage 2009.
 - 2010 *Grundbegriffe der wissenschaftlichen Mathematik*. Springer, Berlin, 2010.
 - 2012 *Erste Hilfe in Analysis*. Springer, Berlin, 2012. Erste Auflage dieses Buches.
 - 2013 *Analysis 1*. Springer, Berlin, 2. Auflage 2013.
 - 2015 *Analysis 2*. Springer, Berlin, 2. Auflage 2015.
- Deiser, Oliver / Lasser, Caroline** 2015 *Erste Hilfe in Linearer Algebra*. Springer, Berlin, 2015.
- Deiser, Oliver / Lasser, Caroline / Vogt, Elmar / Werner, Dirk** 2016 *12 x 12 Schlüsselkonzepte zur Mathematik*. Springer, Berlin, 2. Auflage 2016.
- Fischer, Gerd** 2010 *Lernbuch Lineare Algebra und Analytische Geometrie*. Vieweg+Teubner, Wiesbaden, 2010.
- Forster, Otto** 2011 *Analysis 1*. Vieweg+Teubner, Wiesbaden, 10. Auflage 2011.
- Freitag, Eberhard / Busam, Rolf** 2006 *Funktionentheorie 1*. Springer, Berlin, 4. Auflage 2006.
- Hairer, Ernst / Wanner, Gerhard** 2010 *Analysis in historischer Entwicklung*. Springer, Berlin, 2010.
- Heuser, Harro** 2009 *Analysis 1*. Vieweg+Teubner, Wiesbaden, 17. Auflage 2009.
- Hildebrandt, Stefan** 2006 *Analysis 1*. Springer, Berlin, 2. Auflage 2006.
- Königsberger, Konrad** 2004 *Analysis 1*. Springer, Berlin, 6. Auflage 2004.
- Reiss, Kristina / Schmieder, Gerald** 2007 *Basiswissen Zahlentheorie*. Springer, Berlin, 2. Auflage 2007.
- Remmert, Reinhold / Schumacher, Georg** 2002 *Funktionentheorie 1*. Springer, Berlin, 5. Auflage 2002.
- Rudin, Walter** 2009 *Analysis*. Oldenbourg Verlag, München, 4. Auflage 2009.
- Sonar, Thomas** 2011 *3000 Jahre Analysis. Geschichte, Kulturen, Menschen*. Springer, Berlin, 2011.
- Walter, Wolfgang** 2004 *Analysis 1*. Springer, Berlin, 7. Auflage 2004.
- Zorich, Vladimir** 2006 *Analysis I*. Springer, Berlin, 2006.

Notationen

Abkürzungen		$]a, b[$	14	f/g	27
I. A.: Induktionsanfang		$[a, +\infty[$	14	f^2	27
I. S.: Induktionsschritt		$] -\infty, b[$	14	f^3	27
I. V.: Induktionsvoraussetzung		$m \in M$	15	id	27
		$m \notin M$	15	f^{-1}	28
Griechisches Alphabet		\subseteq	15	$f C$	29
Alpha	A	α	15	$f[X]$	30
Beta	B	β	15	$f(X)$	30
Gamma	Γ	γ	15	$f^{-1}[Y]$	30
Delta	Δ	δ	16	$f^{-1}(Y)$	30
Epsilon	E	ε	16	\forall	32
Zeta	Z	ζ	16	\exists	32
Eta	H	η	16	\mathbb{A}	42
Theta	Θ	θ, ϑ	16	$x + y$	46
Jota	I	ι	16	$x \cdot y$	46
Kappa	K	κ	16	$x - y$	47
Lambda	Λ	λ	16	x / y	47
My	M	μ	17	$<$	48
Ny	N	ν	17	\leq	48
Xi	Ξ	ξ	17	$ x $	49
Omikron	O	o	17	$\sup(X)$	50
Pi	Π	π	17	$\inf(X)$	50
Rho	P	ρ	18	\mathbb{R}	52
Sigma	Σ	σ, ς	18	$m, d_1 d_2 d_3 \dots$	54
Tau	T	τ	18	$\bigcap \mathcal{A}$	56
Ypsilon	Y	υ	20	$\bigcap_{A \in \mathcal{A}} A$	56
Phi	Φ	φ	20	$\sqrt[n]{x}$	58
Chi	X	χ	20	a^q	59
Psi	Ψ	ψ	20	a^x	59
Omega	Ω	ω	21	\mathbb{C}	60
		\mathbb{R}^*	21	$\text{Re}(z)$	60
\mathbb{N}	14	$ M \leq N $	25	$\text{Im}(z)$	60
\mathbb{N}^*	14	$ M = N $	25	$i = (0, 1)$	60
\mathbb{Z}	14	$g \circ f$	26	$ z $	62
\mathbb{Q}	14	cf	27	$\arg(z)$	62
\mathbb{R}	14	$f + g$	27	\bar{z}	63
$[a, b]$	14	$f - g$	27	$\zeta_0, \dots, \zeta_{n-1}$	63
$[a, b[$	14	$f \cdot g$	27		

$(x_n)_{n \in \mathbb{N}}, (x_n)_{n \geq 0}, \langle x_n \mid n \in \mathbb{N} \rangle, (x_0, x_1, \dots, x_n, \dots), x_0, x_1, \dots, x_n, \dots$	66	e	149	$f^{(n)}$	180
$\lim_{n \rightarrow \infty} x_n = x$	68	$\log :]0, \mathbb{R}[\rightarrow \mathbb{R}$	150	$D^n f$	180
$\lim_n x_n = x$	68	$\ln :]0, \mathbb{R}[\rightarrow \mathbb{R}$	150	$d^n f/dx^n$	180
e	70	\exp_a	152	$(d/dx)^n f$	180
$[a_0, \dots, a_n]$	71	$\exp_a : \mathbb{R} \rightarrow \mathbb{R}$	152	\mathcal{C}^n	180
$[a_0, a_1, \dots]$	71	a^x	152	\mathcal{C}^∞	180
$(x_{i_n})_{n \in \mathbb{N}}$	76	$\text{pot}_b :]0, +\infty[\rightarrow \mathbb{R}$	153	$f \geq^* 0,$	185
$(x_{i(n)})_{n \in \mathbb{N}}$	76	$\log_a :]0, +\infty[\rightarrow \mathbb{R}$	154	$f \leq^* 0,$	185
$\limsup_n x_n$	82	$\exp : \mathbb{C} \rightarrow \mathbb{C}$	156	$T_p^n f$	192
$\liminf_n x_n$	82	\exp	156	T_p^n	192
$U_\varepsilon(x)$	84	$\sin : \mathbb{R} \rightarrow \mathbb{R}$	160	T_0^n	193
$\lim_{n \rightarrow \infty} z_n = z$	86	$\cos : \mathbb{R} \rightarrow \mathbb{R}$	160	T_p^n	193
$\lim_n z_n = z$	86	\tan	162	$T_p f(x)$	193
$U_\varepsilon(z)$	87	\cot	162	$\delta(p)$	196
$\overline{\mathbb{R}}$	88	\sec	162	$\sum_p f$	198
$\sum_{n \in \mathbb{N}} x_n$	92	\csc	162	$I(f)_b$	198
$\sum_{n=0}^\infty x_n$	92	\arcsin	164	$\int_a^b f$	198
$\sum_{n \geq 0} x_n$	92	\arccos	164	$\int_a^b f(x) dx$	198
$\sum_n x_n$	92	\arctan	164	$\lim_{\delta(p) \rightarrow 0} \sum_p f$	198
$m, d_1 d_2 \dots d_n \dots$	98	arccot	164	$\sum_p f$	199
$m, a_1 \dots a_k \overline{b_1 \dots b_p}$	98	\arcsin	164	$S_p f$	200
$C = 0,57721566$	101	\arccos	164	$s_p f$	200
$\sum_{n,m}$	112	\arctan	164	Sf	200
$\exp : \mathbb{R} \rightarrow \mathbb{R}$	114	arccot	164	sf	200
e	114	$\arg(z)$	165	$R(f)$	202
$e = 2,71\dots$	114	\arctan_2	165	$\int_a^b f(x) dx$	205
$\exp : \mathbb{C} \rightarrow \mathbb{C}$	115	π	166	$f^{[b]}$	207
$\lim_x \rightarrow_p f(x)$	120	$f'(p)$	172	f^-	207
$\lim_x \downarrow_p f(x)$	121	f'	172	$M(f)$	208
$\lim_x \uparrow_p f(x)$	121	Df	172	$F \Big _a^b$	210
$\lim_x \rightarrow_\infty f(x)$	121	$df(x)/dx$	172	$[F(x)]_a^b$	210
$\lim_x \rightarrow_{-\infty} f(x)$	121	$d/dx f(x)$	172	$\int_b^f f$	211
$\lim_x \rightarrow_{-\infty} f(x)$	121	$Df(p)$	172	$\int_b^f f$	214
$\lim_n \rightarrow_\infty f_n$	138f	$df/dx(p)$	172	$J(\mathbb{P})$	218
$\sum_n a_n (x-p)^n$	140	$df(x)/dx _{x=p}$	172	$j(P)$	218
$\exp : \mathbb{R} \rightarrow \mathbb{R}$	148	$a(p, x)$	172	\neg	228
\exp	148	$o(x-p)$	174	\wedge	228
e^x	149	$o((x-p)^k)$	174	\vee	228
		$f^{(0)}$	180	\rightarrow	228

\leftrightarrow	228
\forall	230
\exists	230

Index

A

Abbildung 20
 Abelscher Grenzwertsatz 141
 Ableitung 172, 180
 Ableitung der Umkehrfunktion 176
 Ableitung und Monotonie 184f
 Ableitungen 4, 236
 Ableitungsregeln 176
 absolut konvergent 106
 Abspalten von Nullstellen 145
 Abspaltungsregel 103
 Abstandsbegriff 87
 abzählbar 44
 abzählbar unendlich 44
 Addition 60
 Additionstheorem 148f, 153, 156
 Additionstheoreme 161
 additiv neutral 46
 affine Abbildungen 170
 algebraisch 42
 allgemeine Exponentialfunktion 3, 152
 allgemeingültig 229
 Allquantor 230
 alternierend 104
 alternierende harmonische Reihe 105
 Analytische Definition von π 166
 angeordneter Körper 48
 Annahme des Mittelwerts 208
 Annahme von Maximum und Minimum 134
 Anordnungsaxiome 48
 antisymmetrisch 19
 Approximationssatz 174, 192
 äquidistant 196
 Äquivalenz 228
 Äquivalenzrelationen. 19
 archimedisch 147
 archimedische Anordnung 53
 Archimedisches Axiom 231
 Argument 20, 62, 165
 Arkusfunktionen 164
 Arkuskosinus 164
 Arkuskotangens 164
 Arkussinus 164
 Arkustangens 164
 Assoziativgesetz 46
 auf 18, 20, 66, 118

Aufspaltung 204f
 Aus A folgt B 228
 Aussonderungen 16
 Auswertungsnotation 210
 Axiome für die reellen Zahlen 4, 231

B

b-adischen Entwicklungen 55
 Basis 152, 154
 bedingt konvergent 106
 Berechnung von Integralen 210
 Berechnungsformel 199
 Bernoulli-Ungleichung 35
 Berührungspunkt 120
 beschränkt 21, 50, 67
 bestimmt divergent 88
 Betrag 49, 62
 Betragsfunktion 178
 bijektiv 24
 Bild 20, 30
 Bindungsstärke der Junktoren 229
 binomische Formel 34, 149
 Blockbildung 103
 Bogenmaß 62
 Bruchrechnen 47

C

Cantorsche Diagonalverfahren 45
 Cauchy-Bedingung 74f
 Cauchy-Bedingung für Reihen 102
 Cauchy-Folge 53, 74
 Cauchy-Hadamard 141
 Cauchy-Kriterium 102
 Cauchy-Produkt 113, 149
 Charakteristik 48

D

Darboux-Integral 200
 Darboux-integrierbar 200
 Darboux'sche Integrierbarkeitsbedingung 200
 Darboux-Summe 200
 Darstellung in Polarkoordinaten 165
 Deduktion 34
 Definitionsbereich 20
 dehnungsbeschränkt 128
 Dezimalbruch 54

Dezimalbruchentwicklung 54
 Dezimaldarstellung 54, 98
 Dezimalziffern 54
 Diagonalaufzählung 113
 Diagonalprodukt 113
 Diagonalverfahren 45
 dicht 53
 Differentialquotient 172
 Differenzenquotient 128, 172
 differenzierbar 172, 180
 Dirichletsche Sprungfunktion 22, 123, 197,
 207, 217
 Disjunktion 228
 Diskriminante 42
 Distributivgesetz 46
 divergent 68
 Divergenz der harmonischen Reihe 100
 Divergenzbedingung 69
 Division 47
 Division mit Rest 145
 Doppelbedeutung 92
 Doppeldeutigkeit 55
 Doppelsumme 112
 Dreiecksungleichung 49
 Dualdarstellung 55
 Durchschnitt 56
 dyadische Entwicklung 55

E_____

echte Obermenge 15
 echte Teilmenge 15
 Eindeutigkeitsbedingung 20
 eingeschlossene Fläche 219
 eingeschlossene signierte Fläche 198
 Einheitswurzeln 63, 167
 Einschieben der Null 49
 Einschränkung 29, 204
 endlich 25, 44
 endliche geometrische Reihe 96
 endlichen Kettenbrüche 71
 Entwicklungspunkt 140, 192
 Epsilonontik 4, 232
 erweiterten reellen Zahlen 88
 euklidische Länge 62
 Euler-Mascheroni-Konstante 101
 Eulersche Formel 161
 Eulersche Identität 157
 Eulersche Zahl 114

Eulerschen Zahl 70
 Existenz von n -ten Wurzeln 73
 Existenz von Stammfunktionen 210
 Existenz von Wurzeln 58, 133
 Existenzquantor 230
 Exponentialfunktion 59, 114, 131, 148, 152,
 156
 Exponentialreihe 113ff
 Exponentialschreibweise für exp 149
 Exponentiation 59, 131
 Extensionalitätsprinzip 15
 Extremum 186
 Extremwertsatz 134

F_____

Fallunterscheidung 229
 Feinheit 196
 Fixpunkt 133
 Folge 66
 Folgen als Reihen 94
 Folgennotationen 66
 Formel von Cauchy-Hadamard 141
 Fortsetzung 130
 Fortsetzungssatz 131
 Fundamentalsatz der Algebra 61, 145
 Funktion 20
 Funktionalgleichung 148
 Funktionenfolge 138

G_____

gekürzt 147
 gemischtperiodisch 99
 Geometrische Definition von π 166
 geometrische Reihe 96, 99
 geordnete Paare 17
 geordnetes Paar 17
 Gerade 170
 glatt 180
 gleiche Kardinalität 25
 gleichmächtig 25
 Gleichmächtigkeit 25
 gleichmäßig stetig 126
 gleichmäßige Konvergenz 139
 gleichmäßige Stetigkeit 135
 Glied 66
 gliedweises Differenzieren 148, 181
 Grad 144
 Graph 20

Grenzwert 68, 120, 138

Grenzwertsatz 141

Grundrechenarten 47

H

harmonische Reihe 75, 100

harmonische Zahlen 100

Häufungspunkt 78, 85

Häufungspunktbedingung 79

Hauptsatz 94

Hauptsatz der Differential- und Integralrechnung 210

h-Formulierung der Ableitung 172

hinreichend 228

höhere Ableitungen 180

I

I.V. 35

Identität 27

Identitätssatz 145

imaginäre Einheit 60

Imaginärteil 60

immer wieder 84

Implikation 228

impliziert 228

in 66

Index 66

Indexfolge 76

indirekten Beweisen 229

Induktionf, 34

Induktionsanfang 35

Induktionsschritt 35

Induktionsvoraussetzung 35

Infimum 50

infinitesimal 53

injektiv 24

Inklusion 15

Integralform des Restglieds 192

Integralfunktion 210

Integralvergleichskriterium 215

Integration und Differentiation 209

Integrationsregeln 212

integrierbar 198, 214

Integrierbare Funktionen 4, 237

Integrierbarkeitsbedingung 198, 200, 203

Interpolation 130

Intervall 53

Intervallschachtelung 56, 135

Inverse 46

Inversenregel für \mathbb{C} 63

Irrationalität der Quadratwurzel aus 2 40

irreflexiv 19

Irreflexivität 48

J

Jensensche Ungleichung 189

Jordan-Inhalt 218

Jordan-messbar 218

Junktoren 228

K

kanonische Periode 99

Kardinalität 25

kartesisches Produkt 17

Kettenregel 176

Klein-o-Notation 174, 192

kleinstes Gegenbeispiel 36

Koeffizient 140, 144

Kommutativgesetz 46, 107

kompakt 57, 135

komplex Konjugierte 63

komplexe Exponentialfunktion 115, 156

komplexe Exponentialreihe 115

komplexe Zahl 60

Komposition 26

Kongruenz modulo m 18

Konjugation 63

Konjunktion 228

konkav 188

konstante Folge 66

Kontrapositionsgesetz 37, 229

konvergent 68, 138

Konvergenz 68

Konvergenz der geometrischen Reihe 96

Konvergenz in \mathbb{C} 86

Konvergenz von Cauchy-Folgen 74

Konvergenzbedingung 68, 86ff, 138f

Konvergenzbedingung für Pendelfolgen 71

Konvergenzbereich 140

Konvergenzkriterium für alternierende Reihen 104

Konvergenzradius 141

konvex 188

koordinatenweise Konvergenz 87

Körper 46, 60

Körper der reellen Zahlen 52

Körperaxiome 46

Kosekans 162

Kosinus 160

Kotangens 162

Kreisaufwicklung 157

Kreisumfang 157

Kreiszahl 166

Kreuzprodukt 17

Kuratowski-Paar 17

L_____

Lagrangesche Form 171

Lagrangesche Form des Restglieds 192

Landau-Symbol 174, 192

Länge 62, 196

Lebesgue-Integral 203, 217

Leibniz-Kriterium 104

Leibniz-Reihe 71, 105

Leitkoeffizient 144

Limes 68

Limes Inferior 82

Limes Superior 82

Limesdefinition von exp 149

Limesnotation 68

Limesnotationen 120

Limesregeln 72

linear) vollständig 52

lineare Abbildung 170

lineare Interpolation 130

linearen Anteil 170

Linearität 170, 176, 204

links- bzw. rechtsseitige Ableitung 173

linksgekrümmt 188

linksseitig (bzw. rechtsseitig) differenzierbar 173

linksseitige Grenzwert 121

Lipschitz-Bedingung für L 128

Lipschitz-Konstante 128

Lipschitz-stetig 128

Lipschitz-Stetigkeit 183

Logarithmen umrechnen 155

Logarithmus 101, 150, 154

Logarithmus-Reihe 141, 151

logische Verknüpfungen 228

lokale Extrema 186, 191

lokales Extremum 186

lokales Maximum 186

lokales Minimum 186

M_____

Mächtigkeit 25

Mächtigkeitsvergleich 25

Majorante 108

Majorantenkriterium 108

Menge 14

metrische Konvergenzbedingung 87

metrische Räume 87

metrische Vollständigkeitsaxiom 53

Minorantenkriterium 109

Minus mal Minus 47

Mittelwert 208

Mittelwertsatz der Differentialrechnung 182

Mittelwertsatz der Integralrechnung 208

modus ponens 229

monoton 21, 67

Monotonie 204

Monotonie der Ableitung und Konvexität 190

Multiplikation 60

Multiplikationsregel 62, 167

Multiplikationstheorem 150

multiplikativ neutral 46

N_____

nach oben beschränkt 67

Nachdifferenzieren 176

natürliche Logarithmus 150

Negation 228

negatives Element 48

Negativteil 207

nichtarchimedisch 147

Nonstandard-Analysis 53

normiert 144

Normiertheit 204

notwendig 228

n-te Glied 66

n-te Wurzel 58, 133

Nullfolge 73, 102

Nullfolgenbedingung 102

Nullpolynom 144

O_____

Ober- 200

obere Schranke 50

Obermenge 15

offene ε -Umgebung 84

o-Notation 174, 192

Ordnung 180, 192

P_____

Partialbruchzerlegung 147
 Partialsumme 92, 115
 Partielle Integration 212
 Partition 196
 Pendelfolge 71, 104
 Periode 98
 periodisch 98
 periodische Folge 79
 Polarkoordinaten 158, 165
 Polstelle 146
 Polynom 27, 144
 Polynomdivision 145
 positiv 147
 positives Element 48
 Positivteil 207
 Potenzfunktion 59
 Potenzfunktionen 153
 Potenzreihe 97, 140
 Potenzreihenentwicklung 161, 193
 Prinzip vom kleinsten Element 36
 Produkt 112
 Produktregel 112, 176
 Pünktchenschreibweise 17
 Punkt-Richtungs-Darstellung 171
 punktweise Konvergenz 138

Q_____

Quadratur des Kreises 43
 Quadratwurzel 58
 Quantoren 4, 230
 Quotientenbedingung 111
 Quotientenkriterium 111, 114
 Quotientenregel 176

R_____

rationale Exponenten 59
 rationale Funktion 146
 Realteil 60
 rechtsgekrümmt 188
 rechtsseitig) differenzierbar 173
 rechtsseitige 121
 rechtsseitige Ableitung 173
 reelle Exponenten 59, 131
 reelle Funktion 21
 reelle Polynomfunktion 144

reelles Polynom 144
 reflexiv 19
 Regel von Robertson 99
 Regelfunktion 202
 Regelintegral 202
 Reihe 92, 115
 Reihenentwicklungen 4, 235
 reinperiodisch 99
 Relation 18, 29
 Restglied 192
 Restgliedabschätzung 148
 Restklassenkörper 46
 Riemann-Integral 198, 219
 Riemann-Summe 198
 Rückwärtsintegral 205
 Russell-Antinomie 16

S_____

Satz vom kleinsten Ausreißer 36
 Satz von Bolzano-Weierstraß 80, 85, 134
 Satz von Gauß 41
 Satz von Rolle 182
 Satz von Taylor 192
 schließlich 84
 Sekans 162
 signiert 198
 Signum 123
 Sinus 160
 Skalierung 112
 Sonderrolle der Null 47
 Stammfunktion 209ff
 starke Induktion 37
 Steigung 165, 170
 Steigungsform 171
 Stelle 20
 stetig 118, 124
 stetig differenzierbar 180
 stetig fortsetzbar 130
 stetige Fortsetzung 130
 Stetigkeit der Umkehrfunktion 136
 Stetigkeitssatz 139
 strebt 120
 streng konkav 188
 streng konvex 188
 streng monoton 21, 67
 striktes lokales Maximum 186
 striktes lokales Minimum 186
 stückweise stetig 206

Stützstellen 196
 stützstellenfreie) Partition 197
 Substitutionsregel 212
 Subtraktion 47
 Summe 92, 115
 Summe zweier Folgen 72
 Summenregel 112
 Supremum 50
 surjektiv 24
 symmetrisch 19

T_____

Tangens 162f
 Tangente 172
 Tangentenkriterium 189
 Tautologie 229
 Taylor-Approximation 174
 Taylor-Entwicklung 3, 192
 Taylor-Polynom 192
 Taylor-Reihe 193
 Teiler 145
 Teilerpolynom 145
 Teilfolge 76
 Teilmenge 15
 Teilmengenbeziehung 15
 Teleskopsumme 74
 tertium non datur 229
 Thomae-Funktion 123
 transitiv 19
 Transitivität 48
 transzendent 42
 Treppenfunktion 197, 202
 Tripel 17

U_____

überabzählbar 44
 ε -Umgebung 84
 umgebungsstetig 124
 Umkehrfunktion 28
 Umordnung 107
 Umordnungssatz 107
 Umrechnungsformel 165
 unbestimmtes Integral 211
 uneigentlich integrierbar 214
 uneigentlich konvergent 88
 uneigentliche Konvergenzbedingung 88
 Uneigentliches Integral 214
 unendlich 25

unendlich oft 84
 unendliche Reihe 92
 unendlichen Dezimalbruch 54
 unendlicher Kettenbruch 71
 Unendlichkeitssymbole 88
 Unstetigkeit der Art 1a 122
 Unstetigkeit der Art 1b 122
 Unstetigkeit der Art 2 122
 Unstetigkeitsstelle zweiter Art 179
 untere Schranke 50
 Unterintegral 200
 unvollständig 52
 Urbild 20, 30

V_____

Vergleichbarkeit 48
 Verknüpfung 26
 Verneinungs- und Vertauschungsregeln 230
 Verneinungsregeln 32, 229
 Vertauschbarkeit 72, 118
 Vertauschungsregeln 32
 Vertauschungssatz 216
 vollständig 52
 vollständige Induktion 34
 vollständigen Definitionsbereich 146
 Vollständigkeit 231
 Vollständigkeitsaxiom 52
 von A nach B 21

W_____

Wachstumsbegriffe 67
 Wahrheitstafel 228
 Wendepunkt 191
 Wert 20
 Wertebereich 20
 Wertebereich stetiger Funktionen 135
 Wertevorrat 21
 Wertverlaufsinduktion 37
 Winkel 62
 Wohlordnungssatz 37
 Wurzel 58, 73, 133
 Wurzelbedingung 110
 Wurzelkriterium 110

Z_____

Zerlegung in Linearfaktoren 145
 Zerlegung in Positiv- und Negativteil 207
 Zerlegungspunkte 196

Zusammenfassung	14
Zweideutigkeit der Dezimaldarstellung	55
Zweipunktdarstellung	171
Zwischenwertsatz	132