

Mathematik 1
für Bau- und Umweltingenieurwesen
Vorlesung und Übungen
TU Wien
Wintersemester 2020/21

Reinhard Winkler,
mit zahlreichen Übungsaufgaben von
Gabriel Maresch

18. September 2020

Sehr geehrte Hörerinnen und Hörer!

Das vorliegende Skriptum dient als wichtigstes Begleitmaterial zu Vorlesung und Übung „Mathematik 1 für Bau- und Umweltingenieurwesen“ an der TU Wien im Wintersemester 2020/21. Gegenüber der Vorgängerversion vom Wintersemester 2018/19 wurden zahlreiche kleinere Fehler korrigiert und ein paar geringfügige Umstellungen vorgenommen, um den gedanklichen Aufbau der Vorlesung zu verbessern.

Wir befinden uns zur Zeit mitten in der Corona-Krise mit den damit verbundenen Komplikationen. Es ist nicht garantiert, dass die Lehrveranstaltung über die gesamte Zeit in der Form abgehalten werden kann, wie dies zur Zeit (Anfang September) geplant ist, nämlich wenigstens teilweise als Präsenzlehre. Bitte lesen Sie diese Präambel unter den entsprechenden Vorbehalten.

Neben den präzisen Formulierungen der mathematischen Inhalte (Definitionen, Sätze etc.) sind im Skriptum auch ausführliche Erklärungen der wichtigsten Ideen enthalten. Wenn Sie hin und wieder die Vorlesung versäumen, sollte es mit Hilfe des vorliegenden Skriptums also problemlos möglich sein, Versäumtes nachzulernen. Überdies ist erstmals geplant, die Vorlesung auch aufzuzeichnen, so dass Sie diese auch im Nachhinein sehen und hören können. Trotzdem ist ein – im Rahmen der Corona-bedingten Einschränkungen – möglichst regelmäßiger Besuch der Vorlesung sehr zu empfehlen. Denn so besteht die Möglichkeit zu Zwischenfragen und damit zur raschen Klärung von Verständnisproblemen. Außerdem lassen sich im mündlichen Vortrag an der Tafel Gewichtungen, Intuitionen und zahlreiche andere wichtige Aspekte mathematischer Inhalte wesentlich besser vermitteln als mit bedrucktem Papier. Wann immer sinnvoll, gibt es in der Vorlesung auch Hinweise, in welcher Form die Inhalte geprüft werden.

Als begleitende Pflichtlehrveranstaltung zur Vorlesung gibt es auch Übungen, die überwiegend mit Aufgabenmaterial aus dem Skriptum bestritten werden. Einen beträchtlichen Teil davon hat mein Kollege Gabriel Maresch beige-steuert. Aus Zeitgründen, aber nicht ausschließlich deshalb, kann nur ein relativ kleiner Teil der mehr als 300 Übungsaufgaben in der Lehrveranstaltung „Übung“ behandelt werden. Dennoch empfehle ich, sich in einem gewissen Umfang auch mit den dort nicht behandelten Aufgaben zu beschäftigen. Dazu sind einige Erläuterungen am Platz.

Als Orientierungshilfe für den Umgang mit den Aufgaben haben wir eine Zuordnung zu drei Kategorien T (Test), P (Prüfung) und E (Ergänzung) vorgenommen. Für die Vorbereitung auf Übungstests empfehlen wir vor allem die Beschäftigung mit den T-, für die Prüfung darüber hinaus mit den P-Aufgaben. Von den E-Aufgaben sind einige etwas anspruchsvoller. Lassen Sie sich nicht entmutigen, wenn Sie manche davon nicht lösen können. Wenn Sie sich aber bei jeder E-Aufgabe wenigstens klar machen, worin die Aufgabenstellung besteht, so wird das Ihr Verständnis ganz wesentlich vertiefen und daher eine sehr sinnvolle Vorbereitung auf die Prüfung sein.

Abgesehen vom vorliegenden Skriptum werde ich auf elektronischem Wege auch eine Sammlung von Anwendungen des Stoffes aus Mathematik im Bauingenieurwesen, die Kolleginnen und Kollegen aus Ihrer Fakultät zur Verfügung gestellt haben, ausgeben. Für das meiste daraus wird Stoff aus Mathematik 2 oder gar aus Mathematik 3 erforderlich sein. Prüfungsstoff zu den Mathematik-Lehrveranstaltungen sind diese Anwendungsbeispiele nicht, sondern Motivation, damit Sie frühzeitig sehen, wie Mathematik in Ihrem eigentlichen Fach wirksam wird und warum sie deshalb als Grundlage unverzichtbar ist.

Ergänzend noch einige Bemerkungen zur Vorbereitung auf die Prüfung: Im Skriptum sind auch die meisten Beweise enthalten. Für Anwender in den Ingenieurwissenschaften kommt es nicht darauf an, dass man all diese Beweise aktiv beherrscht. Um bei Anwendungen zu den richtigen Methoden greifen zu können, ist es aber trotzdem von entscheidender Bedeutung, die wichtigsten Ideen wenigstens in einem intuitiven Sinn erfasst zu haben. Sehr zielführend dabei ist es, beim Lernen sehr wohl auch die Beweise zu studieren, wenn auch nicht unbedingt mit dem Anspruch, sie frei reproduzieren zu können. Entsprechend kommen längere Beweise bei mir nicht als Prüfungsaufgaben. Einfache Schlussweisen, wie sie als Teile von Beweisen häufig vorkommen, sollten Ihnen aber geläufig sein. Als erfolgsversprechende Strategie beim Lernen empfehle ich deshalb, sich zu bemühen, wenigstens lesend möglichst den gesamten Stoff aus den Unterlagen zu verstehen. „Möglichst“ heißt dabei, dass Sie sich an einzelnen unklaren Stellen nicht ewig aufhalten sollen, sondern zunächst weitergehen dürfen. Vielleicht klärt sich vieles beim nächsten Durchlesen. Wenn für Sie jedoch zu viele Beweise unklar bleiben, stellt sich die Frage, ob Sie für die Prüfung schon hinreichend gut vorbereitet sind. Einige weitere Hinweise zur Prüfung generell wie auch Prüfungsangaben aus den vergangenen Jahren finden Sie auf meiner homepage unter:

<http://dmg.tuwien.ac.at/winkler/pruefungen/>

Bitte stellen Sie sich darauf ein, dass ein universitäres Ingenieursstudium einer Vollzeitbeschäftigung entspricht und beträchtliche Anforderungen stellt, sowohl intellektuell als auch an Ihren persönlichen Einsatz. Genies, die so eine Aufgabe nebenbei bewältigen, sind äußerst selten. Auch die Mathematik als Grundlagenfach im ersten Studienjahr verlangt Ihnen einen großen Arbeitsaufwand ab – weit mehr als Sie vermutlich in der Schule für dieses Fach leisten mussten. Nehmen Sie es also bitte nicht auf die leichte Schulter. Dafür ist es eine umso größere Genugtuung, die Herausforderungen eines anspruchsvollen Studiums zu meistern. Es ist meine Aufgabe als Vortragender in der Vorlesung, mein Bestes zu geben, um Ihnen bei der Aneignung des umfangreichen Stoffes behilflich zu sein und Sie dazu zu motivieren, gleichfalls Ihr Bestes zu geben. Mit den besten Vorsätzen dazu wünsche ich Ihnen viel Erfolg in den Mathematiklehrveranstaltungen und in Ihrem gesamten Studium.

Reinhard Winkler, im September 2020

Inhaltsverzeichnis

1	Vorbemerkungen und Grundbegriffe	9
1.1	Einordnung der Mathematik in Studium und Welt	9
1.1.1	Ingenieurwissenschaft	9
1.1.2	Naturwissenschaft	10
1.1.3	Mathematik	10
1.1.4	Philosophie	12
1.2	Die Sprache der Mathematik:	
	Logik und Mengen	13
1.2.1	Mathematische Sprache und zweiwertige Logik	13
1.2.2	Junktoren und Quantoren	15
1.2.3	Mengen und mengentheoretische Operationen	24
1.2.4	Kartesische Produkte und Relationen	27
1.2.5	Der Funktionsbegriff	29
1.3	Die Zahlenbereiche	32
1.3.1	Die Zahlenmengen $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ im Überblick	33
1.3.2	Veranschaulichung von \mathbb{R} als Zahlengerade	35
1.3.3	Axiomatik am Beispiel von \mathbb{R}	36
1.3.4	Die Supremumseigenschaft als Konsequenz der Vollständigkeit	39
1.3.5	Die archimedische Eigenschaft	40
1.3.6	Betrag, Potenzen und Wurzeln	41
1.3.7	Das Induktionsprinzip	43
1.3.8	Anwendung des Induktionsprinzips:	
	Rekursionen und Induktionsbeweise	45
1.3.9	Teilbarkeit und Primfaktorzerlegung in \mathbb{N} und \mathbb{Z}	49
1.3.10	Zahlendarstellungen	52
1.3.11	Die komplexe Zahlenebene	55
1.4	Elementare Kombinatorik	62
1.4.1	Einfachste Anzahlformeln	62
1.4.2	Anzahlen von Funktionen und Permutationen; Faktorielle	63
1.4.3	Teilmengen, Kombinationen und Binomialkoeffizienten	65
1.4.4	Binomischer und polynomischer Lehrsatz	66

1.5	Vektoren im n -dimensionalen Raum \mathbb{R}^n	68
1.5.1	Koordinatisierung und Vektorraumoperationen	69
1.5.2	Eigenschaften der Vektorraumoperationen	70
1.5.3	Die kanonische Basis	71
1.5.4	Linearität am Beispiel der Drehung	72
1.5.5	Inneres (Skalar-) Produkt und Längenmessung	74
1.5.6	Zwei wichtige Ungleichungen	77
1.5.7	Das äußere (Vektor-) Produkt im \mathbb{R}^3	79
1.5.8	Interessante Teilmengen in \mathbb{R}^2 und \mathbb{R}^3	81
1.5.9	Interessante Teilmengen in beliebigen metrischen Räumen	85
2	Folgen und Reihen	89
2.1	Reelle Folgen	89
2.1.1	Einfache Beispiele und Eigenschaften reeller Folgen	89
2.1.2	Der Grenzwert einer Folge	92
2.1.3	Konvergenzregeln (Grenzwertsätze)	95
2.1.4	Einige Beispiele konvergenter Folgen	99
2.1.5	Häufungspunkte	101
2.1.6	Obere und untere Limiten	103
2.1.7	Die wichtigsten Sätze aus der Theorie der Folgen	105
2.1.8	Rekursive Folgen und Iterationsfolgen	111
2.2	Reihen	112
2.2.1	Der Wert einer unendlichen Reihe	112
2.2.2	Einige wichtige Beispiele konvergenter und divergenter Reihen	115
2.2.3	Wurzel- und Quotientenkriterium	117
2.2.4	Besonderheiten absolut konvergenter Reihen	119
2.2.5	Alternierende Reihen	122
3	Funktionen und Stetigkeit	125
3.1	Reelle Funktionen	125
3.1.1	Graphische Darstellung und einfache Eigenschaften	125
3.1.2	Grenzwert von Funktionen und Stetigkeit	129
3.1.3	Vererbungseigenschaften stetiger Funktionen	131
3.1.4	Beispiele zu Stetigkeit und Unstetigkeit	132
3.1.5	Konsequenzen der Vollständigkeit von \mathbb{R} für stetige Funktionen	134
3.1.6	Fixpunktiteration	137
3.1.7	Stetigkeit und gleichmäßige Konvergenz	139
3.2	Polynome und rationale Funktionen (Elemente der Algebra)	142
3.2.1	Definition, Auswertung, Polynomdivision	142
3.2.2	Nullstellen und Eindeutigkeitssatz	144
3.2.3	Fundamentalsatz der Algebra	146
3.2.4	Nullstellenbestimmung	149

3.2.5	Approximation und Interpolation durch Polynome	151
3.2.6	Partialbruchzerlegung	153
3.3	Weitere wichtige Beispiele	
	stetiger Funktionen	156
3.3.1	Wurzeln und Potenzen mit rationalem Exponenten	156
3.3.2	Exponentialfunktion	157
3.3.3	Logarithmen, Potenzfunktionen, Zusammenfassung	159
3.3.4	Trigonometrische Funktionen	163
4	Differentialrechnung	167
4.1	Die Ableitung einer reellen Funktion	167
4.1.1	Motivation: Tangente, Momentangeschwindigkeit	167
4.1.2	Differentialquotient und lineare Approximation	169
4.1.3	Ableitungsregeln	172
4.1.4	Die Ableitung von Umkehrfunktionen, insbesondere von Wurzeln	177
4.2	Taylorapproximation und Potenzreihen	178
4.2.1	Monotonie und erste Ableitung, Anfang	179
4.2.2	Der Mittelwertsatz	180
4.2.3	Stetige Differenzierbarkeit und höhere Ableitungen	182
4.2.4	Der Satz von Taylor	183
4.2.5	Die Exponentialreihe und die Eulersche Zahl e	186
4.2.6	Verallgemeinerung: Potenzreihen und ihr Konvergenzbereich in \mathbb{C}	188
4.2.7	Wichtige Eigenschaften von Potenzreihen	189
4.2.8	Die Regel von de l'Hospital	191
4.3	Wichtige Beispiele differenzierbarer Funktionen	193
4.3.1	Logarithmus, Fortsetzung	193
4.3.2	Allgemeine Potenzen und binomische Reihe	195
4.3.3	Die Differenzierbarkeit der trigonometrischen Funktionen	196
4.3.4	Die Arcusfunktionen	199
4.3.5	Die Hyperbel- und Areafunktionen	201
4.4	Anwendungen der Differentialrechnung	203
4.4.1	Monotonie und erste Ableitung, Fortsetzung	203
4.4.2	Konvexität, Krümmung und zweite Ableitung	205
4.4.3	Extremwertbestimmung und höhere Ableitungen	208
4.4.4	Fixpunktiteration, Dynamik und Chaos	210
4.4.5	Das Newton-Verfahren	212
5	Integralrechnung	215
5.1	Das Riemannintegral	215
5.1.1	Ober- und Untersummen	216
5.1.2	Riemannsummen	218
5.1.3	Einfache Eigenschaften	219

5.1.4	Das Riemannintegral stetiger Funktionen	222
5.2	Der Hauptsatz und seine Anwendungen	224
5.2.1	Zwei Versionen des Hauptsatzes	224
5.2.2	Die Berechnung von Integralen mittels Stammfunktionen	226
5.2.3	Integrationsregeln	228
5.2.4	Beispiele zur Integration	230
5.3	Das Integral bezüglich eines Maßes	234
5.3.1	Der Begriff des Maßes	234
5.3.2	Vertiefende Bemerkungen zum Lebesgueschen Maß	236
5.3.3	Neuinterpretation des Riemannintegrals als Lebesgueintegral	240
5.3.4	Einige Sätze aus der Lebesgueschen Theorie	245
5.4	Einige Anwendungen und ausgewählte Themen	246
5.4.1	Uneigentliche Integrale	247
5.4.2	Unendliche Reihen und Integrale	249
5.4.3	Numerische Integration	251
5.4.4	Kurven und ihre Länge	253

Kapitel 1

Vorbemerkungen und Grundbegriffe

Während die späteren Kapitel 2 (Folgen und Reihen), 3 (Stetige Funktionen), 4 (Differentialrechnung) und 5 (Integralrechnung) jeweils ein bestimmtes Thema im Auge haben, stellt das erste Kapitel mancherlei Grundlegendes zusammen. In den fünf Abschnitten des ersten Kapitels geht es um: die Einordnung der Mathematik in einen größeren Zusammenhang (1.1), die vor allem logisch und mengentheoretisch bedingten Besonderheiten der Sprache der Mathematik (1.2), die fünf klassischen Zahlenbereiche (1.3), elementare Kombinatorik (1.4) und Vektoren und Geometrie (1.5).

1.1 Einordnung der Mathematik in Studium und Welt

Um die Rolle der Mathematik im Rahmen eines universitären Ingenieurstudiums, aber auch in noch größerem Kontext realistisch einzuschätzen, lohnt es, sich mit Gegenstand und Methode wie auch mit dem Verhältnis von Mathematik (1.1.3), Natur- (1.1.2) und Ingenieurwissenschaft (1.1.1) zueinander auseinanderzusetzen. Diese Auseinandersetzung wiederum ist der Philosophie (1.1.4), insbesondere ihren Teilgebieten Wissenschafts- und Erkenntnistheorie zuzuordnen. Ohne in die Tiefe zu gehen, sollen hier einige wichtige Charakteristika dieser Bereiche schlagwortartig erwähnt werden.

1.1.1 Ingenieurwissenschaft

Inhalt in Kurzfassung: Einige wichtige Merkmale der Ingenieurwissenschaften in Schlagworten.

Ziel: Gestaltung der Wirklichkeit, Problemlösung

Methode: empirisch, experimentell

Instrumente: Verständnis der Naturgesetze, Einsicht in die Bedürfnisse des Menschen

1.1.2 Naturwissenschaft

Inhalt in Kurzfassung: Einige wichtige Merkmale der Naturwissenschaften in Schlagworten.

Ziel: Verständnis der Naturgesetze

Methode: empirisch, experimentell

Instrumente: u.a. Sprache und Begriffe der Mathematik; Ergebnisse der Mathematik sind nur in dem Maße auf die Wirklichkeit anwendbar, in dem die mathematischen Modelle die Wirklichkeit adäquat beschreiben

1.1.3 Mathematik

Inhalt in Kurzfassung: Einige wichtige Merkmale der Mathematik.

Ziel: Präzisierung der Begriffe, in denen wir über die Wirklichkeit nachdenken können; Erforschung von Denkmöglichkeiten und Denknötigkeiten

Methode: axiomatisch, d.h. logisch-deduktiv; entsprechend sind die Ergebnisse der Mathematik in Bezug auf die mathematischen Begriffe (nicht unbedingt auf ihre nur ungefähren Entsprechungen in der Realität), denknötig, scharf und sicher.

Instrumente: reines Denken, Bleistift, Papier, Papierkorb

Einige Erläuterungen zur axiomatischen Methode sind am Platz. Sie wurde in der griechischen Antike eingeführt. Vor allem die *Elemente* des Euklid aus dem 3. Jahrhundert v. Chr. wurden für Jahrtausende zum exemplarischen Lehrbuch für diese Methode.

Dabei werden in **Axiomen** jene Grundsätze formuliert, für die man keinen Beweis verlangt. Man kann das so wie in der Antike interpretieren, nämlich dass die Axiome offensichtliche und unumstößliche Wahrheiten darstellen. Heutzutage sieht man in Axiomen eher implizite Definitionen, das heißt Vereinbarungen, worüber man spricht; nämlich über jene Dinge oder Systeme von Dingen, die sich so verhalten wie in den Axiomen beschrieben. Wir werden zum Beispiel ein Axiomensystem für die reellen Zahlen kennenlernen und, wenn auch nicht konsequent, so doch exemplarisch damit arbeiten.

Wichtig für die axiomatische Methode sind auch **Definitionen**, die in der Mathematik einem besonders hohen Anspruch an Präzision genügen müssen.

Man kann Definitionen als Abkürzungen für umständlichere Formulierungen auffassen. Insofern sind sie nicht prinzipiell notwendig, de facto, vor allem aus Gründen der leichten Fasslichkeit einer Theorie, aber schon.

Zwei Beispiele von Definitionen, anhand derer wir auch gleich eine gebräuchliche Notation einführen können: Indem wir etwa schreiben $e := \sum_{n=0}^{\infty} \frac{1}{n!}$, legen wir fest, dass das (einfache) Symbol e für jene reelle Zahl (die Eulersche Zahl, Basis des natürlichen Logarithmus) verwendet wird, die durch die unendliche Summe rechts eindeutig (aber mit mehr Schreibaufwand) definiert wird. Indem wir $:=$ mit Doppelpunkt schreiben statt $=$, wollen wir sagen, dass es sich hierbei nicht um eine zu beweisende Behauptung (ein Theorem) handelt, sondern um eine Definition, die uns, sofern wir das Symbol e noch nicht dauerhaft für etwas anderes verbraucht haben, frei steht. Ist diese Definition einmal gegeben, darf das Symbol e (wenigstens im gegebenen Kontext) nur mehr in diesem Sinn verwendet werden. Ähnlich führt man Kurzsprech- oder -schreibweisen ein wie etwa $a|b : \leftrightarrow \exists t \in \mathbb{Z} : b = at$ (Teilbarkeit; wir werden diese Formelsprache bald zu lesen lernen). Von nun an steht die Zeichenkette $a|b$ als Abkürzung für die Aussage rechts des Äquivalenzdoppelpfeils \leftrightarrow . Der Doppelpunkt vor dem Doppelpfeil deutet an, dass die beiden Aussagen links und rechts davon definitionsgemäß äquivalent sind. Ohne den Doppelpunkt würde es sich um eine Behauptung handeln, die, bevor sie als wahr anerkannt wird, erst bewiesen werden müsste. Doch zurück zur axiomatischen Methode.

Von den Axiomen und Definitionen ausgehend leitet man auf rein logischem Wege neue Aussagen her, die man dann **Sätze** oder **Theoreme** der durch das Axiomensystem und seine Sprache gegebenen **Theorie** nennt. Sehr einfache Theoreme nennt man manchmal auch **Propositionen** (Aussagen). Ein Satz, den man eher als Vorstufe zu einem interessanteren Ergebnis betrachtet, heißt auch **Hilfssatz** oder **Lemma**. Sätze, die recht unmittelbar aus einem (vielleicht wichtigeren) Theorem folgen, nennt man **Folgerungen** oder **Korollare**. Argumentationsketten, die zwingend zu Propositionen, Lemmata, Theoremen und Korollaren führen, heißen deren **Beweise**. Worin genau das Zwingende von mathematischen Beweisen besteht, das sie etwa von Schlüssen im Alltag (wie z.B. *Der Oktober hat angefangen, da wird es in den nächsten Wochen wohl kälter werden*) unterscheidet, lässt sich mit den Mitteln der mathematischen Logik exakt bestimmen und analysieren. Genaueres würde hier aber zu weit führen.

Für uns wichtig sind einige Bemerkungen zur mathematischen Praxis, womit gemeint ist: zur Art und Weise, wie Mathematiker ihre Wissenschaft erleben und betreiben. Für Mathematiker sind Beweise von interessanten Theoremen das, wonach sie suchen und womit sie Anerkennung in der Fachwelt erwerben. Der Weg dorthin ist aber vor allem von Intuitionen und Ideen geprägt, die zunächst durchaus vage sein können. Das übergeordnete Ziel in der Mathematik wie generell in der grundlagenorientierten Wissenschaft besteht darin, Zusammenhänge besser zu verstehen. Und das wiederum bedeutet, Vorstellungen und Begriffe von den Dingen, die einen interessieren, in sinnvolle Beziehung zueinander zu bringen.

Auch bei der Vorbereitung auf die Mathematikprüfung empfiehlt es sich, dies im Auge zu haben. Komplizierte Beweise müssen dazu nicht auswendig

beherrscht werden. Indem man sich aber bemüht, die Beweise wenigstens lesend zu verstehen, wird man automatisch wichtige Ideen mitnehmen. Dabei ist es nicht tragisch, wenn man hin und wieder einzelne Schritte nicht nachvollziehen kann. Im Großen und Ganzen sollte man aber das Gefühl haben, klar zu wissen, wovon die Rede ist und warum welche Überlegungen angestellt werden.

Weil es um das Verständnis vor allem im Großen geht, nimmt man sich in der Mathematik trotz aller grundsätzlich anzustrebenden Präzision und Strenge mancherlei Freiheiten in Notation etc., ohne jedes Mal extra darauf hinzuweisen. Auch wir werden das so halten.

Abschließend noch eine Bemerkung zur mathematischen Formelsprache. Diese wurde vor allem deshalb entwickelt, weil sie im Erkenntnisprozess eine praktische und wertvolle Hilfe ist, nicht als Selbstzweck. Praktisch ist sie, weil sie Abkürzungen ermöglicht und damit Schreibarbeit sparen hilft. Sie ist aber auch darüber hinaus wertvoll, weil sie es ermöglicht, inhaltliche Zusammenhänge deutlicher sichtbar machen. Wie viel Formalismus man in einem konkreten Zusammenhang (z.B. im vorliegenden Skriptum in Hinblick auf eine mathematische Einführungslehrveranstaltung einer technischen Studienrichtung) verwendet, ist deshalb eine pragmatische Frage, deren Antwort davon abhängt, wie groß die genannten Vorteile sind in Vergleich mit der Hürde, die mathematische Formelsprache für Anfänger darstellt. Für welche Dosis auch immer man sich schließlich entscheidet – es ist sehr wahrscheinlich, dass man damit angesichts großer individueller Unterschiede den Bedürfnissen nicht aller Beteiligten in gleicher Weise gerecht wird. Trotzdem sind redliche Bemühungen um ein sinnvolles Maß Grundvoraussetzung.

1.1.4 Philosophie

Inhalt in Kurzfassung: Einige wichtige Merkmale der Philosophie in Schlagworten.

Ziel: Identifikation und Diskussion der für den Menschen insgesamt wichtigen Fragen. Das hat wiederum entscheidende Rückwirkungen auf die Ziele in den Ingenieurwissenschaften, siehe 1.1.1. Aufgrund des universellen Anspruchs der Philosophie darf man aber nur selten abschließende Antworten erwarten.

Methode: offen (anything goes), d.h. der Logik im strengen Sinn der Mathematik zwar verpflichtet (ihre Gesetze dürfen nicht verletzt werden), es sind aber auch andere Wege der Erkenntnis zugelassen. Insbesondere können Ergebnisse aus sämtlichen Einzelwissenschaften (nicht nur mathematisch geprägten) relevant werden.

Instrumente: reines Denken, Bleistift, Papier

1.2 Die Sprache der Mathematik: Logik und Mengen

Vor allem im Laufe des 20. Jahrhunderts hat sich die Erkenntnis durchgesetzt, dass die bereits angesprochene methodische Strenge der Mathematik nur dadurch zu gewährleisten ist, dass man sich einer besonders klaren Sprache bedient. Die nötige Klarheit muss auf zwei Ebenen erreicht werden, von denen sich die eine mit Syntax/Grammatik, die andere mit Semantik/Bedeutung natürlicher Sprachen vergleichen lässt. Im Fall der Mathematik wird, formalistisch betrachtet, die Rolle der Syntax von der (formalen) Logik übernommen, die Rolle der Semantik von den Mengen (beides ist Teil der mathematischen Logik). Im Hinblick auf ingenieurwissenschaftliche Anwendungen liegt uns dieser formalistische Standpunkt zwar nicht per se am Herzen. Bis zu einem gewissen Punkt müssen diese Grundlagen aber durchdrungen werden, um auch kompliziertere mathematische Inhalte bewältigen zu können. Der Logik zuzuordnen sind die unmittelbar folgenden ersten beiden Unterabschnitte, die sich mit der klassischen zweiwertigen Logik (1.2.1) sowie mit Junktoren und Quantoren (1.2.2) beschäftigen. Die drei daran anschließenden betreffen grundlegende mengentheoretische Konstruktionen: unmittelbar an die Logik anschließende (1.2.3), kartesische Produkte und Relationen (1.2.4) und den Funktionsbegriff (1.2.5).

1.2.1 Mathematische Sprache und zweiwertige Logik

Inhalt in Kurzfassung: Die gesamte Mathematik ist davon geprägt, dass sie nur solche Aussagen macht, die entweder wahr oder falsch sind, Zwischenstufen sind nicht möglich. Entsprechend präzise müssen mathematische Begriffe, Formulierungen und Beweise sein. Somit kommt einer Sprache, die das ermöglicht, eine überragende Bedeutung zu.

Wie jede Wissenschaft macht die Mathematik **Aussagen**. Die Gültigkeit mathematischer Aussagen ist (im Gegensatz zu den meisten Bereichen des Alltags, aber auch zu manchen anderen Wissenschaften) keine graduelle, sondern eine strikte. Der Mathematik wird daher die **zweiwertige (klassische) Logik** zugrunde gelegt, wonach jede korrekt formulierte Aussage entweder *wahr* ist oder *falsch*, nichts drittes (*tertium non datur*).¹ Es gibt also nur **zwei Wahrheitswerte**. (Sehr wohl sind graduelle Zwischenstufen zwischen wahr und falsch bei „metamathematischen“ Aussagen, also Aussagen *über* die Mathematik möglich, z.B.: „Die heutige Mathematikvorlesung war recht interessant.“)

Es sei betont, dass die Mathematik mit der Wahrscheinlichkeitstheorie (siehe Mathematik 2) und mit der mehrwertigen Logik auch Teildisziplinen bereitstellt, die sehr wohl auch den methodisch strengen Umgang mit Aussagen ermöglicht, die nicht sicher wahr oder sicher falsch sind, sondern deren Gültigkeitsstatus von

¹Das bedeutet allerdings noch lange nicht, dass man immer entscheiden kann, ob eine bestimmte Aussage wahr oder falsch ist. So wie in jeder lebendigen Wissenschaft gibt es auch in der Mathematik viele offene Fragen. Von einigen kann man sogar beweisen, dass sie nicht entscheidbar sind – jedenfalls nicht in einer Weise, wie man das ursprünglich gehofft hat.

anderer Art ist, möglicherweise z.B. in irgendeiner Weise irgendwo zwischen 0 und 1 liegt. Doch wollen wir der zu entwickelnden Mathematik im Weiteren stets die klassische, zweiwertige Logik zugrunde legen.

Im einfachsten Fall bringt eine (mathematische oder auch außermathematische) Aussage zum Ausdruck, dass ein bestimmtes Objekt eine bestimmte Eigenschaft hat. Zum Beispiel: *3 ist eine Primzahl*. Diese Aussage ist wahr, etwa im Gegensatz zur gleichermaßen sinnvollen und korrekt gebildeten, aber falschen Aussage: *4 ist eine Primzahl*. In Begriffen der Grammatik bestehen solche Aussagen (Sätze) nur aus einem Subjekt (nämlich 3 bzw. 4; in der Mathematik spricht man informell allerdings eher von *Objekten* statt von Subjekten) und einem **Prädikat** (nämlich Primzahl zu sein).

Weitere sehr einfach gebaute Aussagen sind **Gleichungen** wie $2 + 2 = 4$ (wahr) oder $2 + 2 = 5$ (falsch). Obwohl darin jeweils mehrere Objekte vorkommen, lassen sich Gleichungen, aber auch Ungleichungen wie $4 < 5$ (wahre Aussage), gleichfalls als Aussagen interpretieren, die lediglich aus einem Subjekt und einem Prädikat bestehen. Im Beispiel $4 < 5$ kann man als Subjekt nämlich das geordnete Paar $(4, 5)$ auffassen und als Prädikat die Kleinerbeziehung $<$.

Die Analogie zwischen mathematischer und natürlicher Sprache ist auch in folgendem Punkt erhellend. Und zwar geht es um die Unterscheidung zwischen **Bezeichnetem** (mathematischem Objekt) und **Bezeichnendem** (sprachlichem Symbol). Was gemeint ist, wird sofort klar angesichts einer Gleichung wie $2+2=4$. Das Gleichheitssymbol $=$ besagt, dass es sich links und rechts davon um identische Objekte handelt, die in jedem beliebigen Kontext wechselseitig austauschbar sind. Die schriftliche Aneinanderreihung der Symbole 2, + und nochmals 2 auf der linken Seite ist allerdings nicht identisch mit dem Symbol 4 auf der rechten Seite. Daraus schließen wir, dass sich das Gleichheitssymbol $=$ auf die bezeichneten Objekte bezieht und nicht auf die bezeichnenden Symbole. Die Unterscheidung zwischen Symbol und Objekt ist also unerlässlich. Mathematische Aussagen beziehen sich (analog wie übrigens auch Aussagen in der natürlichen Sprache) in der Regel² auf die bezeichneten mathematischen Objekte und nicht auf die bezeichnenden Symbole, mit denen diese bezeichnet werden.³

Diese Erörterungen mögen hier etwas abgehoben wirken. Viele Irrtümer und Missverständnisse von Studienanfängern (und nicht nur solchen) beruhen aber gerade auf diesbezüglichen Verwechslungen. Beispiele dazu werden uns noch begegnen, z.B. bei der Unterscheidung zwischen **Zahlen** und ihren **Zifferndarstellungen**.

Die bisher behandelten mathematischen Aussagen sind elementar oder auch atomar in dem Sinn, dass sie sich nicht auf einfachere Bestandteile zurückführen lassen. Die Mathematik wäre aber ärmlich, müsste sie sich mit solchen Atomformeln begnügen. Dann könnte sie nicht einmal so einfache Gesetze formulieren wie z.B.:

²Ausnahmen zu dieser Regel kommen vor, besonders in der Informatik, in der mathematischen Logik, in der Algebra etc., müssen dann aber ausdrücklich als solche gekennzeichnet werden.

³Wer mag, darf sich an der Assoziation zu Magrittes berühmtem Bild einer Pfeife mit dem Text *Ceci n'est pas une pipe* erfreuen.

Kommutativgesetz (erste Fassung): Das Ergebnis einer Addition hängt nicht von der Reihenfolge der Summanden ab.

Damit ist offenbar gemeint, dass $x + y = y + x$ gilt, egal welche Zahlen für x und y eingesetzt werden. Die Symbole x und y stehen also für beliebige Werte aus einer bestimmten Menge. Man nennt x und y in diesem Zusammenhang (also ausnahmsweise in ihrer Eigenschaft als Symbole) **Variable**. Wir erlauben uns in der Mathematik, Formeln zu bilden, in denen Variable vorkommen, wobei der Wahrheitswert der entstehenden Formel davon abhängen kann, was für diese Variablen eingesetzt wird. Beispielsweise ist die Formel $x < y$ per se weder wahr noch falsch. Erst wenn x und y spezifiziert werden, ergibt sich eine wahre (z.B. für $x = 4$ und $y = 5$) oder falsche (z.B. für $x = 5$ und $y = 4$) Aussage. Gelegentlich spricht man dabei statt von einer Aussage auch von einer **Aussageform**.

Wir kehren nochmals zum Kommutativgesetz zurück und finden eine konzisere Fassung als oben:

Kommutativgesetz (zweite Fassung): Für alle Zahlen x und y gilt $x + y = y + x$.

Die Wendung *Für alle ...* spielt in der Mathematik eine derart wichtige Rolle, dass man dafür ein eigenes Symbol, den sogenannten **Allquantor** \forall eingeführt hat. Mit seiner Hilfe lässt sich noch kürzer schreiben:

Kommutativgesetz (dritte Fassung): $\forall x, y : x + y = y + x$.

Um zu würdigen, dass es sich dabei nicht nur um eine Abkürzung handelt, sondern dass dadurch wesentliche Elemente der Mathematik deutlich werden, lohnt es, etwas weiter auszuholen. Davor noch eine Übung:

Übungsaufgabe 1. (P) Formulieren Sie in natürlicher Sprache (d.h. ohne Verwendung mathematischer Formeln) Aussagen, die

1. Gegenstand der Mathematik sind.
2. normalerweise nicht Gegenstand der Mathematik sind, die aber geeignet wären, Gegenstand der Mathematik zu sein.
3. nicht geeignet sind, Gegenstand der Mathematik zu sein.

1.2.2 Junktoren und Quantoren

Inhalt in Kurzfassung: Aus jeder Aussage kann durch Negation (\neg , einstelliger Junktor) eine Aussage (in diesem Fall mit entgegengesetztem Wahrheitswert) gebildet werden, ähnlich aus je zwei Aussagen durch Verbindung durch einen der zweistelligen Junktoren \wedge („und“), \vee („oder“), \rightarrow (Implikation) und \leftrightarrow (Äquivalenz). Wenn man Variable verwendet, die eventuell unendlich viele verschiedene Werte annehmen können, kann man mit Hilfe von Allquantor \forall und Existenzquantor \exists sogar beliebig viele Aussagen mit gleicher Struktur zu einer einzigen verbinden. In jedem dieser Fälle ergibt sich der Wahrheitswert der neuen Aussage nach formalen Regeln eindeutig aus den Wahrheitswerten ihrer Bestandteile. Bei allgemeingültigen Aussagen wie Tautologien ergibt sich für jede mögliche

Belegung der Bestandteile der Wahrheitswert „wahr“.

Junktoren und **Quantoren** dienen dazu, aus gegebenen Aussagen kompliziertere zu bilden. Wichtig ist dabei, dass sich der Wahrheitswert der neuen Aussage allein aus den Wahrheitswerten der ursprünglichen eindeutig ergibt, also sonst von nichts abhängt. Allgemein gebräuchlich sind fünf Junktoren und zwei Quantoren. Ein Junktor, die Negation \neg , ist **einstellig**, weil er auf eine Aussage angewendet wird. Die vier Junktoren \wedge (Konjunktion), \vee (Disjunktion), \rightarrow (Implikation) und \leftrightarrow (Äquivalenz) sind **zweistellig**, weil sie auf zwei Aussagen angewendet werden. Die beiden Quantoren \forall (Allquantor) und \exists (Existenzquantor) beziehen sich auf beliebig, typischerweise unendlich viele Einzelaussagen. Die genaue Bedeutung, zunächst der Negation:

Negation, nicht, \neg : Der Wahrheitswert der Aussage $\neg A$ ist verschieden von jenem der Aussage A . Beispiel einer wahren Aussage mit Negation ist $\neg(5 < 4)$.⁴ Negationen schreibt man oft auch als $3 \neq 4$ für $\neg(3 = 4)$, $-1 \notin \mathbb{N}$ für $\neg(-1 \in \mathbb{N})$, ... Die Abhängigkeit des Wahrheitswertes der Aussage $\neg A$ kann einprägsam in Form einer (in diesem Fall eines einstelligen Junktors sehr einfachen) **Wahrheitstafel** festgehalten werden:

A	$\neg A$
w	f
f	w

Es folgen die vier zweistelligen Junktoren:

Konjunktion, und, \wedge : Die Aussage $A \wedge B$ ist genau dann wahr, wenn sowohl A als auch B wahr sind. Beispiel einer wahren Aussage mit Konjunktion: $(4 < 5) \wedge (5 < 6)$.

Disjunktion, oder, \vee : Die Aussage $A \vee B$ ist genau dann wahr, wenn wenigstens eine der Aussagen A und B wahr ist. Beispiel einer wahren Aussage (nicht nur) mit Disjunktion: *Für alle reellen Zahlen a, b gilt $(a \leq b) \vee (b \leq a)$.*

Implikation, wenn-dann, \rightarrow : Die Aussage $A \rightarrow B$ ist nur dann falsch, wenn A (das Vorderglied der Implikation) wahr und B (das Hinterglied der Implikation) falsch ist. Beispiel einer wahren Aussage (nicht nur) mit Implikation: *Für alle reellen Zahlen a gilt $(a < 4) \rightarrow (a < 5)$.* Man beachte, dass eine Implikation $A \rightarrow B$ laut Vereinbarung jedenfalls dann wahr ist, wenn B wahr ist, und auch dann, wenn A falsch ist. Zum Beispiel ist die Implikation $2 < 1 \rightarrow 1 < 2$ wahr. Das mag für den Anfänger ungewohnt sein, ist in der Mathematik aber zweckmäßig. Vielleicht wird es dann verständlich, wenn man sich als Merkregel einprägt: Die Implikation $A \rightarrow B$

⁴Wir verwenden Klammern in der üblichen Weise, d.h. um klarzustellen, welche Teilformeln zuerst zu lesen sind.

bedeutet, dass B mindestens so wahr ist wie A . Oder: Gilt A , dann erst recht B ; gilt A nicht, dann ist uns für B alles recht.

Äquivalenz, genau-dann-wenn, \leftrightarrow : Die Aussage $A \leftrightarrow B$ ist genau dann wahr, wenn A und B denselben Wahrheitswert haben. Beispiel einer wahren Aussage (nicht nur) mit Äquivalenz: *Für alle a gilt $(a < 4) \leftrightarrow (4 > a)$.*

Die Wahrheitstafel für die vier behandelten zweistelligen Junktoren haben also die folgende Gestalt:

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

Und nun zu den beiden Quantoren:

Allquantor, für alle, \forall : Kommt in einer Formel eine Variable, z.B. x , vor, so hängt der Wahrheitswert von $A(x)$ davon ab, was für x eingesetzt wird. Die Aussage $\forall x : A(x)$ ist genau dann wahr, wenn für alle (im gegebenen Kontext sinnvollen) Werte a , die für die Variable x eingesetzt werden können, die Aussage $A(a)$ wahr ist. Die Beispiele zur Disjunktion, Implikation und Äquivalenz lassen sich auch mit Allquantor schreiben, etwa zur Äquivalenz: $\forall x : (x < 4) \leftrightarrow (4 > x)$. Man beachte, dass so eine Allaussage auch als Konjunktion unendlich vieler Aussagen interpretiert werden kann, im Beispiel (wenn wir den Quantor auf die natürlichen Zahlen $0, 1, 2, \dots$ beziehen) als $((0 < 4) \leftrightarrow (4 > 0)) \wedge ((1 < 4) \leftrightarrow (4 > 1)) \wedge \dots \wedge ((5 < 4) \leftrightarrow (4 > 5)) \wedge \dots$

Existenzquantor, es gibt, \exists : Die Aussage $\exists x : A(x)$ ist genau dann wahr, wenn unter allen (im gegebenen Kontext sinnvollen) Werten a , die für die Variable x eingesetzt werden können, wenigstens einer existiert, für den die Aussage $A(a)$ wahr ist. Beispiel einer wahren Aussage mit All- und Existenzquantor, bezogen auf natürliche Zahlen: $\forall x \exists y : y > x$ (zu jeder natürlichen Zahl x gibt es eine größere Zahl y). Falsch wäre hingegen: $\exists y \forall x : y > x$ (es gibt eine natürliche Zahl y , die größer ist als alle x). Analog zum Allquantor lässt sich der Existenzquantor auch als unendliche Disjunktion interpretieren. Man beachte, dass der Existenzquantor auch bei Negationen des Allquantors auftritt: Wenn etwas nicht für alle gilt, so gibt es zumindest ein Gegenbeispiel, genauer: Die Formel $(\neg \forall x : A(x)) \leftrightarrow (\exists x : \neg A(x))$ ist, wie man sagt, **allgemeingültig**, weil sie unabhängig von der inhaltlichen Bedeutung von $A(x)$ immer wahr ist. Ebenfalls allgemeingültig ist die dazu duale Formel: $(\neg \exists x : A(x)) \leftrightarrow (\forall x : \neg A(x))$ (Übung). Erwähnt sei auch die oft praktische Notation $\exists!$ für „es gibt genau ein ...“. Wenn man das Gleichheitszeichen $=$ zur Verfügung hat, lässt sich $\exists! x : A(x)$ definieren als Abkürzung für $(\exists x : A(x)) \wedge (\forall x, y : (A(x) \wedge A(y)) \rightarrow x = y)$.

Wollte man auch für All- und Existenzquantor den Wahrheitswerteverlauf in Tabellenform angeben, so bräuchte man in der Regel unendlich viele Spalten für die Aussagen $A(x)$, für jedes in Frage kommende x eine. Nur wenn in jeder x -Spalte der Wert „w“ stünde, wäre das auch in der Spalte für die Gesamtformel $\forall x : A(x)$ der Fall, während die Spalte für $\exists x : A(x)$ ein „w“ bekäme, sobald in wenigstens einer x -Spalte ein „w“ steht.

Damit sind im Wesentlichen alle für die Mathematik relevanten Möglichkeiten beschrieben, wie aus einfachen Aussagen komplexere gebildet werden können. Will man mathematische Aussagen (wenigstens in ihrer formalen Bedeutung) verstehen, muss man also vor allem diese Sprachelemente beherrschen. Formeln, in denen keine Quantoren, sondern nur Junktoren vorkommen, heißen auch **aussagenlogische Formeln**.

Besonders wichtig sind in der Mathematik Gleichheit $=$ und Äquivalenz \leftrightarrow . Eine fundamentale Beziehung, die Gleichheit mit Äquivalenz verbindet, lässt sich in der Formel

$$x = y \rightarrow (A(x) \leftrightarrow A(y))$$

fassen, wobei A eine Aussagen ist, die von einer Variablen x abhängt, d.h. die eine Eigenschaft zum Ausdruck bringt, die ein Objekt, das man für x einsetzen kann, haben kann oder nicht. Verbal bedeutet die Formel: Sind x und y identisch, so gleichen sie sich hinsichtlich jeder beliebigen Eigenschaft A . In jeder beliebigen Aussage, die für x wahr ist, kann man an beliebigen Stellen, an denen x auftritt, x durch y ersetzen (und umgekehrt) und erhält dabei wieder eine wahre Aussage.

Sehr ähnlich wie mit der Gleichheit (Identität) von Objekten kann man mit der Äquivalenz von Aussagen verfahren. Gilt nämlich $A \leftrightarrow B$ (man sagt dann, A und B sind äquivalent), so kann man (ähnlich wie bei Gleichungen $a = b$ zwischen zwei Objekten, zum Beispiel Zahlen, a und b) die Aussage A beliebig durch die Aussage B ersetzen und vice versa. Zum Beispiel sind für alle reellen Zahlen a und b die Aussagen $a - b = 0$ (Aussage A) und $a = b$ (Aussage B) äquivalent. Ersetzt man die eine durch die andere, spricht man von einer **Äquivalenzumformung**.

Wir werden die Symbole für die Junktoren und Quantoren keineswegs immer einsetzen, nur weil das möglich ist, sondern eher selten. Sehr oft werden wir eine verbale Ausdrucksweise bevorzugen. Indem man sich die Symbole trotzdem einprägt, schärft man aber den Blick für die Struktur mathematischer Aussagen. Das erweist sich als sehr hilfreich.

Um in der Notation Klammern zu sparen, verwenden wir folgende allgemein üblichen Bindungsregeln für Junktoren: Am stärksten bindet die Negation \neg , dann Konjunktion \wedge gleich stark wie Disjunktion \vee (zwischen den beiden müssen also jedenfalls Klammern gesetzt werden), am schwächsten \rightarrow und \leftrightarrow (zwischen diesen beiden sind auch Klammern üblich). Zum Beispiel ist $\neg A \wedge B \rightarrow B$ als $((\neg A) \wedge B) \rightarrow B$ zu lesen.

In der mathematischen Argumentation spielen allgemeingültige Aussagen – sie wurden im Zusammenhang mit dem Existenzquantor bereits definiert – ei-

ne besondere Rolle. Aussagenlogische Formeln, die allgemeingültig sind, heißen auch **Tautologien**. Sie sind also für beliebige Wahrheitswerte der Teilaussagen selber immer wahr. Um zu überprüfen, ob es sich bei einer gegebenen aussagenlogischen Formel um eine Tautologie handelt, genügt es, alle Fälle möglicher Wahrheitswertbelegungen der vorkommenden Formeln zu untersuchen. Davon gibt es 2, 4, 8, ...; je nachdem, ob eine, zwei, drei oder mehr Formeln A, B, C, \dots darin vorkommen. Praktisch sind dabei Wahrheitstabellen, wie wir sie schon bei der Definition der Junktoren verwendet haben.

Als Beispiel folgt die Wahrheitstabelle für eine der beiden de Morganschen Regeln, siehe auch Übungsaufgabe 4. Der tautologische Charakter der aussagenlogischen Formel $\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$ kommt darin zum Ausdruck, dass in der entsprechenden Spalte nur der Wahrheitswert w auftritt, unabhängig davon, welche Wahrheitswerte die Formeln A und B haben:

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

Es sei darauf hingewiesen, dass ein sorgloser Umgang mit formaler Logik mancherlei verblüffende Paradoxa generieren kann. Ein typisches Beispiel wären die folgenden beiden Aussagen zweier Personen A und B :

A : „ B sagt die Wahrheit.“

B : „ A lügt.“

Fragen wir nach dem Wahrheitsgehalt dieser Aussagen, können wir zwei Fälle unterscheiden. 1.Fall, A sagt die Wahrheit. Dann sagt auch B die Wahrheit also lügt A , Widerspruch zur Annahme. Also ist der 1.Fall nicht möglich. 2.Fall, A lügt, dann kann aber A nicht lügen, ebenfalls Widerspruch. Dieser Konstellation kann also keine sinnvolle Bedeutung gegeben werden – etwa so wie wenn wir sinnlose Laute aneinanderreihen. Eine Analyse zeigt, dass die innere Widersprüchlichkeit obiger Behauptungen damit zusammenhängt, dass Aussagen etwas über sich selbst behaupten, hier wechselseitig. Eine besonders raffinierte und mathematisch höchst relevante Anwendung solcher Phänomene steckt in dem berühmten Unvollständigkeitssatz von Kurt Gödel (1906-1978). Eine Vertiefung dieses Themas würde hier aber viel zu weit führen und muss Spezialvorlesungen über mathematische Logik für Fachmathematiker überlassen bleiben. Die folgenden Übungsaufgaben enthalten jedoch keine derartigen Feinheiten und können mit Hilfe des hier Behandelten bearbeitet werden.

Übungsaufgabe 2. (T) Drei Personen A , B und C machen die folgenden Aussagen:

A : B und C sagen die Wahrheit.

B: *A* sagt die Wahrheit.

C: *A* lügt und *B* sagt die Wahrheit.

Finden Sie mittels Aussagenlogik heraus, wer lügt und wer die Wahrheit sagt.

Übungsaufgabe 3. (T) Für ein Verbrechen gibt es 3 Verdächtige **A**, **B** und **C**, sowie die folgenden Ermittlungsergebnisse:

- Wenn sich **B** oder **C** als Täter herausstellen, dann ist **A** unschuldig.
- Wenn **A** oder **C** unschuldig sind, dann muss **B** der Täter sein.
- Wenn **C** schuldig ist, dann ist **A** sein Mittäter.

Finden Sie mittels Aussagenlogik heraus, wer der oder die Täter sind.

Übungsaufgabe 4. (T) Begründen Sie, warum die folgenden zusammengesetzten Aussagen Tautologien sind, d.h. immer wahr, unabhängig davon, welche Wahrheitswerte die vorkommenden Einzelaussagen *A*, *B*, *C* haben.

1. $\neg(A \wedge \neg A)$ (Satz vom Widerspruch)
2. $A \vee \neg A$ (tertium non datur)
3. $(A \rightarrow B) \leftrightarrow (\neg A \vee B)$ (erste Umschreibung der Implikation)
4. $(A \rightarrow B) \leftrightarrow \neg(A \wedge \neg B)$ (zweite Umschreibung der Implikation)
5. $(A \leftrightarrow B) \leftrightarrow (A \wedge B) \vee (\neg A \wedge \neg B)$ (Umschreibung der Äquivalenz)
6. $(A \rightarrow B) \leftrightarrow (\neg B \rightarrow \neg A)$ (Kontraposition)
7. $A \wedge (B \vee C) \leftrightarrow (A \wedge B) \vee (A \wedge C)$ (erstes Distributivgesetz)
8. $A \vee (B \wedge C) \leftrightarrow (A \vee B) \wedge (A \vee C)$ (zweites Distributivgesetz)
9. $\neg(A \wedge B) \leftrightarrow \neg A \vee \neg B$ (erste de Morgansche Regel)
10. $\neg(A \vee B) \leftrightarrow \neg A \wedge \neg B$ (zweite de Morgansche Regel)

Kommen in einer Formel auch Quantoren mit Variablen vor, so können die Variablen eventuell unendlich viele Werte durchlaufen, was unendlich viele Fälle bedeutet. Will man entscheiden, ob die Formel allgemeingültig ist, so steht keine allgemeine Methode wie die der Wahrheitswertetabellen zur Verfügung. In einfachen Fällen führen aber inhaltliche Überlegungen zum Ziel.

Zum Beispiel ist die Formel

$$(\forall x : A(x)) \vee (\forall x : B(x)) \rightarrow (\forall x : (A(x) \vee B(x)))$$

allgemeingültig. Begründung: Als Implikation könnte die Formel nur falsch sein, wenn das Vorderglied $(\forall x : A(x)) \vee (\forall x : B(x))$ der Implikation wahr und das Hinterglied $\forall x : (A(x) \vee B(x))$ falsch wäre. Ist das Vorderglied wahr, so muss

(erster Fall) für alle x die Aussage $A(x)$ wahr sein oder (zweiter Fall) für alle x die Aussage $B(x)$. In beiden Fällen ist aber auch $A(x) \vee B(x)$ für alle x wahr, also $\forall x : (A(x) \vee B(x))$, das Hinterglied. Also ist die Implikation insgesamt allgemeingültig. Die umgekehrte Implikation

$$(\forall x : (A(x) \vee B(x))) \rightarrow (\forall x : A(x)) \vee (\forall x : B(x))$$

ist jedoch *nicht* allgemeingültig. Um das zu einzusehen, beziehe man die Variable x zum Beispiel auf sämtliche natürliche Zahlen. $A(x)$ stehe für die Aussage „ x ist gerade“, $B(x)$ für die Aussage „ x ist ungerade“. Weil jede natürliche Zahl x gerade oder ungerade ist, ist das Vorderglied wahr. Weil aber weder alle natürlichen Zahlen x gerade, noch alle natürlichen Zahlen x ungerade sind, ist das Hinterglied der Implikation und somit die gesamte Implikation falsch.

Die hier nur an einem Beispiel erläuterte Methode des logischen Schließens lässt sich in einem ganz bestimmten Sinn vervollständigen, worüber der ebenfalls von Kurt Gödel erstmals bewiesene Vollständigkeitssatz Auskunft gibt. Etwas verkürzt könnte man sagen: Die Methode mathematischen Beweisens mit rein logischen Mitteln lässt sich vollständig beschreiben und leistet alles, was man sich von ihr sinnvollerweise nur wünschen kann.

In den folgenden Übungen bedeutet die Schreibweise $A(x, y)$, dass die Aussage von zwei Variablen abhängt, wie zum Beispiel $x < y$.

Übungsaufgabe 5. (T) Überlegen Sie, welche der folgenden Formeln allgemeingültig sind und begründen Sie Ihre Antwort. Im negativen Fall bedeutet das das Finden eines Beispiels, d.h. die Angabe konkreter A, B, x, y, \dots

1. $\neg(\forall x : A(x)) \leftrightarrow \exists x : \neg A(x)$ (erste de Morgansche Regel für Quantoren)
2. $\neg(\exists x : A(x)) \leftrightarrow \forall x : \neg A(x)$ (zweite de Morgansche Regel für Quantoren)
3. $(\forall x : A(x)) \vee (\forall x : B(x)) \leftrightarrow (\forall x : (A(x) \vee B(x)))$
4. $(\exists x : A(x)) \vee (\exists x : B(x)) \leftrightarrow (\exists x : (A(x) \vee B(x)))$
5. $(\forall x : A(x)) \wedge (\forall x : B(x)) \leftrightarrow (\forall x : (A(x) \wedge B(x)))$
6. $(\exists x : A(x)) \wedge (\exists x : B(x)) \leftrightarrow (\exists x : (A(x) \wedge B(x)))$
7. $(\forall x \exists y : A(x, y)) \rightarrow (\forall y \exists x : A(x, y))$
8. $(\forall x \exists y : A(x, y)) \rightarrow (\exists y \forall x : A(x, y))$
9. $(\exists x \forall y : A(x, y)) \rightarrow (\forall y \exists x : A(x, y))$

Übungsaufgabe 6. (T) $A(x)$ stehe für „Person x schweigt“, $B(x)$ für „Person x stimmt zu“. Schreiben Sie mittels logischer Symbole und Quantoren:

1. Wer schweigt, stimmt zu.
2. Es gibt jemanden der schweigt, aber trotzdem widerspricht.

3. Jeder der zustimmt, schweigt.

4. Niemand der spricht, stimmt zu.

Entscheiden Sie welche der Aussagen (2), (3) und (4) der Aussage (1) widersprechen, sowie welche der Aussagen (1), (2) und (3) zur Aussage (4) äquivalent sind.

Auf spätere Abschnitte (1.2.3 und 1.3.1) vorgehend verwenden wir in den folgenden Übungsaufgaben schon die aus der Schule vertraute Notation $x \in \mathbb{N}$ für „ x ist eine natürliche Zahl“ bzw., wörtlich, „ x ist ein Element der Menge \mathbb{N} “ u.ä. Zur Erinnerung: \mathbb{N} besteht aus den natürlichen Zahlen $0, 1, 2, 3, \dots$, bei der Menge \mathbb{Z} der ganzen Zahlen kommen auch die negativen Werte $-1, -2, -3, \dots$ hinzu, bei \mathbb{Q} (Menge der rationalen Zahlen) die Brüche ganzer Zahlen (außer Nenner 0) wie $\frac{2}{3}, -\frac{3}{2}, \frac{15}{4}$. Die noch viel umfassendere Menge \mathbb{R} der reellen Zahlen enthält alle Zahlenwerte, die Punkten auf der Zahlengeraden entsprechen wie $\pi, \sqrt{17}$ u.ä. Die Menge \mathbb{C} der komplexen Zahlen schließlich besteht aus allen $a + ib$, wobei $a, b \in \mathbb{R}$ und $i^2 = -1$. Wenn diese Kurzbeschreibung nicht reicht, finden sich ausführlichere Informationen in Abschnitt 1.3.

Übungsaufgabe 7. (T) Sind die folgenden Formeln wahr oder falsch? Geben Sie jeweils eine Begründung oder ein Gegenbeispiel (was genau bedeutet das?) an.

1. $\exists x \in \mathbb{N} \exists y \in \mathbb{N} : x > y$

2. $\exists x \in \mathbb{N} \forall y \in \mathbb{N} : x > y$

3. $\forall x \in \mathbb{N} \exists y \in \mathbb{N} : x > y$

4. $\forall x \in \mathbb{N} \forall y \in \mathbb{N} : x > y$

Was passiert wenn sie in den Formeln \mathbb{N} durch \mathbb{Z} ersetzen?

Übungsaufgabe 8. (T) Sind die folgenden Formeln wahr oder falsch? Geben Sie jeweils eine Begründung oder ein Gegenbeispiel an.

1. $\forall x \in \mathbb{R} \exists y \in \mathbb{R} : x = y^2$

2. $\forall x \in \mathbb{R} \exists y \in \mathbb{R} : x^2 = y$

3. $\forall x \in \mathbb{R} \exists y \in \mathbb{R} : y^2 < x$

4. $\exists x \in \mathbb{R} \forall y \in \mathbb{R} : x < y^2$

Was passiert, wenn Sie in den ersten beiden Formeln \mathbb{R} durch \mathbb{C} bzw. wenn Sie in den letzten beiden Formeln \mathbb{R} durch \mathbb{Q} ersetzen?

Besonders häufig kommt in der Mathematik die Negation von Formeln mit einem oder mehreren Quantoren vor. Verwendet man mehrmals die de Morgan'schen Regeln aus Übungsaufgabe 5, so sieht man, dass zum Beispiel $\neg \forall x \exists y \forall z :$

$A(x, y, z)$ äquivalent ist zu $\exists x \forall y \exists z : \neg A(x, y, z)$. Anders ausgedrückt: Die Formel

$$(\neg \forall x \exists y \forall z : A(x, y, z)) \leftrightarrow (\exists x \forall y \exists z : \neg A(x, y, z))$$

ist allgemeingültig.

Im Anschluss an diesen Themenkreis noch eine Bemerkung und Konvention: Kommen in einer Formel Variable vor, ohne dass sie vorher bei einem Quantor stehen (man sagt: von einem Quantor **gebunden** werden), so nennt man sie **freie Variable**. Kommt in einer Formel eine freie Variable vor, so handelt es sich, wie schon früher erwähnt, nicht um eine Aussage, sondern bestenfalls um eine Aussageform, die für sich genommen keinen Wahrheitswert hat. Um sie zu einer Aussage zu machen, bieten sich drei Möglichkeiten an: Indem man für die freie Variable einen speziellen Wert einsetzt, indem man sie durch einen Existenzquantor bindet, oder indem man sie durch einen Allquantor bindet. Wird diesbezüglich nichts explizit dazu gesagt, geht man meist von der Konvention aus, dass ein Allquantor dazu zu denken ist. Wird also beispielsweise die Formel $x + y = y + x$ als Aussage behauptet, ist meistens $\forall a, b : x + y = y + x$ gemeint oder, genauer $\forall x \forall y : x + y = y + x$. Ist etwas anderes gemeint, muss dies ausdrücklich gekennzeichnet werden. Darauf zu vergessen ist ein äußerst häufiger Anfängerfehler.

Übungsaufgabe 9. (T) Für positive natürliche Zahlen $m, n \in \mathbb{N}$ ⁵ gibt es den Begriff der Teilbarkeit (vgl. auch Abschnitt 1.3.9). Wir schreiben $m|n$, wenn die natürliche Zahl n die natürliche Zahl m teilt, d.h. wenn es eine natürliche Zahl k gibt mit $n = mk$. Betrachten Sie für $n, m \in \mathbb{N}$ die folgenden Aussagen:

1. Dann und nur dann gilt $m|n$, wenn der Bruch $\frac{n}{m}$ eine natürliche Zahl ist.
2. Zu jeder natürlichen Zahl gibt es mindestens einen Teiler.
3. Wenn $m|n$, dann auch $m|n^2$
4. Wenn $2|mn$, dann $2|m$ oder $2|n$.

Schreiben Sie diese Aussagen als Formeln, d.h. mit Quantoren und Implikationen an und überzeugen Sie sich davon, dass sie wahr sind.

Übungsaufgabe 10. (T) Finden Sie die Negationen der Aussagen aus Übungsaufgabe 9 zuerst in Alltagssprache und dann als mathematische Formeln. Achtung: Da die ursprünglichen Aussagen wahr sind, sind ihre Negationen falsch. In dieser Aufgabe sollen Sie diese (falschen Aussagen) natürlich nicht beweisen, sondern lediglich formulieren.

⁵Etwas vorgreifend darf in dieser Aufgaben diese Notation verwendet werden, nämlich dafür, dass n und m natürliche Zahlen, also Elemente der Menge \mathbb{N} aller natürlichen Zahlen sind.

1.2.3 Mengen und mengentheoretische Operationen

Inhalt in Kurzfassung: Eine Menge ist vollständig dadurch gegeben, welche Elemente sie enthält. Logische Junktoren finden ihre Entsprechungen in mengentheoretischen Operationen und Beziehungen wie Durchschnitt, Vereinigung, Komplement, Teilmenge, Gleichheit. Den logischen Quantoren entsprechen großer Durchschnitt und große Vereinigung.

Die Mengenlehre spielt in der modernen Mathematik nicht nur *eine* fundamentale Rolle, sondern mindestens drei. Erstens wird mit „Mengenlehre“ ein Teilgebiet der mathematischen Logik bezeichnet, die ihrerseits ein Teilgebiet der Mathematik ist, das sich besonders mit Grundlagen beschäftigt. Der ursprüngliche Gegenstand der Mengenlehre, der sich im Laufe der Zeit natürlich aufgefächert hat, ist die Größe von Mengen, insbesondere die verschiedenen Unendlichkeiten. Zweitens liefert die Mengenlehre so etwas wie eine große vereinheitlichte Theorie für die gesamte Mathematik – etwas wonach beispielsweise die Physik trotz wichtiger Erfolge immer noch auf der Suche ist. Der vereinheitlichende Charakter der Mengenlehre besteht darin, dass sich (fast) alle Objekte der Mathematik als Mengen auffassen lassen. Das verleiht der Mathematik eine besondere methodische Einheitlichkeit und Denkökonomie. In Verbindung damit steht die dritte und für uns weitaus wichtigste Rolle der Mengenlehre, nämlich als Sprache. Diese Sprache wollen wir uns nun bis zu einer gewissen Geläufigkeit aneignen.

Obwohl der Begriff der **Menge** erst in der zweiten Hälfte des 19. Jahrhunderts durch das Werk Georg Cantors (1845-1918) Eingang in die mathematische Terminologie fand, handelt es sich dabei nach allgemeiner heutiger Auffassung in Verbindung mit der Elementschäftsbeziehung \in um den fundamentalen Begriff der Mathematik schlechthin. Dabei müssen wir uns hier keineswegs mit allen Aspekten beschäftigen. Das für uns Wichtige lässt sich recht knapp zusammenfassen.

Als Versuch einer Definition formulierte Cantor:

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.

Ohne auf die logischen Probleme dieser Definition näher einzugehen⁶ genügt für uns, dass es – ganz im Sinne der zweiwertigen Logik – immer nur darauf ankommt, ob für gegebenes m und M die Elementsbeziehung gilt oder nicht, symbolisch $m \in M$ bzw. $m \notin M$. Diesem Paradigma entsprechend ist jede Menge allein durch ihre Elemente bestimmt, anders formuliert: Haben die Mengen

⁶Dem englischen Philosophen und Logiker Bertrand Russell fiel auf, dass die Menge R aller Mengen, die sich selbst nicht als Element enthalten, zum Widerspruch $R \in R \leftrightarrow R \notin R$ führt. Als Konsequenz ließ man die Bildung von Mengen nicht mehr in dieser absolut freizügigen Weise zu. Man formulierte ausdrücklich, welche Konstruktionen erlaubt sind. Beispiel (Vereinigungsmengenaxiom): *Zu jeder Menge von Mengen gibt es die Vereinigungsmenge.* Je nach Zählweise knapp zehn derartige Aussagen wurden zu Axiomen erhoben, auf deren Basis Mengenlehre und in weiterer Folge der größte Teil der modernen Mathematik betrieben wird.

A und B dieselben Elemente, so sind sie gleich, und natürlich haben umgekehrt gleiche Mengen dieselben Elemente. Formal:

$$(\forall x : (x \in A) \leftrightarrow (x \in B)) \leftrightarrow A = B.$$

Gebräuchlich sind Schreibweisen wie beispielsweise: $\{1, 2, 3\}$ für die Menge, die genau die drei Elemente 1, 2 und 3 enthält; $\{1, 2, 3, \dots\}$ für die (unendliche) Menge aller positiven ganzen Zahlen; $\{x : 1 < x < 2\}$ für die Menge aller (reellen) Zahlen, die zwischen 1 und 2 liegen etc.

Übungsaufgabe 11. (P) Erklären Sie: $\{1, 2\} = \{1, 2, 1\}$ aber $\{1, 2\} \neq \{\{1\}, \{2\}\}$.

Cantor verlieh seinem Mengenbegriff Tiefe, weil in seinen Untersuchungen (zu sogenannten Fourierreihen) auch unendliche Mengen auftraten und er erkannte, dass man auch dort von kleineren und größeren Mengen sprechen kann. Mengenlehre bedeutet gemäß den historischen Wurzeln also vor allem die Untersuchung der verschiedenen Unendlichkeiten. Diesen Aspekt werden wir am Ende von 1.2.5 bei der Unterscheidung abzählbarer und überabzählbarer Mengen nochmals kurz ins Auge fassen.

Zunächst halten wir naiv die Schreibweise $|M|$ für die Anzahl der Elemente einer Menge M fest, sofern M endlich ist, z.B. $|\{4, 5, 7\}| = 3$. Eine besondere Rolle spielt die leere Menge $\{\}$, für die man auch \emptyset schreibt. Sie ist dadurch gekennzeichnet, dass sie kein Element enthält. Für sie gilt daher $|\emptyset| = 0$.

Im Laufe des 20. Jahrhunderts kristallisierte sich der Mengenbegriff als Grundbegriff für die Mathematik schlechthin heraus, weil man lernte, (fast) alle interessanten mathematischen Objekte als Mengen zu deuten. Wie schon einleitend erwähnt, gibt das der Mathematik eine Vereinheitlichung, wie sie beispielsweise der Physik noch abgeht. Neben ästhetischen Qualitäten hat das auch große prinzipielle Vorteile, die für uns allerdings kaum eine Rolle spielen.

Für uns ist vor allem die Sprache der Mengenlehre nützlich. Denn sehr häufig erweist es sich als vorteilhaft, wenn man von einer abstrakten Eigenschaft E zu einer (als vergleichsweise konkretes Objekt vorstellbaren) Menge übergehen kann, nämlich zur Menge M aller Elemente mit der Eigenschaft E . Man schreibt

$$M = \{x : E(x)\} \quad \text{oder} \quad M = \{x \mid E(x)\}.$$

Bei der Eigenschaft $E(x)$ kann es sich auch um eine Gleichung, Ungleichung oder gar ein System von Gleichungen und/oder Ungleichungen handeln, in denen vielleicht neben der Variablen x auch noch weitere Variablen, etwa y, z, \dots vorkommen. In diesem Fall schreibt man $E(x, y, z)$ statt $E(x)$, und M ist die Menge aller Lösungen dieses Systems, die sogenannte **Lösungsmenge**.

Doch interessiert uns zunächst das Verhältnis zwischen Logik und Mengen. Dabei stoßen wir auf mengentheoretische Übersetzungen der logischen Junktoren und Quantoren.

Konjunktion \wedge und (Durch-) Schnitt \cap : Der sogenannte Durchschnitt oder auch die Schnittmenge $M \cap N$ der Mengen M und N besteht aus jenen Elementen, die sowohl in M als auch in N liegen, formal:

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

Ist der Schnitt $M \cap N = \emptyset$ leer, so nennt man die Mengen M und N **disjunkt**.

Disjunktion \vee und Vereinigung \cup : Die sogenannte Vereinigung $M \cup N$ besteht aus jenen Elementen, die in wenigstens einer der Mengen M oder N liegen, formal:

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

Negation \neg , Komplement und Differenz \setminus : Unter dem Komplement M^C der Menge M versteht man $M^C := \{x : \neg x \in M\}$ (absolutes Komplement). Meist (streng genommen muss man das sogar immer tun) bezieht man sich allerdings auf eine umfassende Menge X , in der alle betrachteten x liegen, also eigentlich $M^C = \{x : x \in X \wedge \neg x \in M\} = \{x \in X : x \notin M\}$ (relatives Komplement), ein Spezialfall der für beliebige Mengen M, N definierten **mengentheoretischen Differenz**

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

Gelegentlich von Interesse ist auch die **symmetrische Differenz**

$$M \triangle N := (M \setminus N) \cup (N \setminus M).$$

Implikation \rightarrow und Teilmengenbeziehung \subseteq : In diesem Fall ist die interessante Analogie zwischen Junktoren und Menge von formal etwas anderer Art als zuvor. Und zwar nennen wir M eine **Teilmenge** von N , N eine **Obermenge** von M , symbolisch $M \subseteq N$ bzw. $N \supseteq M$, wenn für alle x die Implikation $x \in M \rightarrow x \in N$ wahr ist, das heißt wenn jedes Element von M auch ein Element von N ist.

Äquivalenz \leftrightarrow und Gleichheit $=$: Bereits weiter oben haben wir gesehen: $M = N$ genau dann, wenn die Äquivalenz $x \in M \leftrightarrow x \in N$ für alle x wahr ist.

Allquantor \forall und großer Durchschnitt \bigcap : Angenommen für jedes $i \in I$ (I ist eine sogenannte Indexmenge und darf auch unendlich sein) ist eine Menge M_i gegeben. Dann lässt sich der Durchschnitt D all dieser Mengen bilden, symbolisch:

$$D = \bigcap_{i \in I} M_i := \{x \mid \forall i \in I : x \in M_i\}$$

D besteht also aus jenen x , die in jedem einzelnen M_i als Element enthalten sind. In der Mengentheorie ist auch die alternative Schreibweise $\bigcap S := \{x \mid \exists ! M \in S : x \in M\}$ für den Durchschnitt aller Mengen M , die selbst Element einer gegebenen Menge S von Mengen sind. Für obiges D wäre also $S := \{M_i : i \in I\}$ zu setzen.

Existenzquantor und große Vereinigung \bigcup : Dual zum Durchschnitt: Angenommen für jedes $i \in I$ (Indexmenge) ist eine Menge M_i gegeben. Dann lässt sich die Vereinigung V all dieser Mengen bilden, symbolisch:

$$V = \bigcup_{i \in I} M_i := \{x \mid \exists i \in I : x \in M_i\}$$

V besteht also aus jenen x , die in wenigstens einem der M_i als Element enthalten sind. Analog zum Durchschnitt ist in der Mengentheorie auch die alternative Schreibweise $V = \bigcup S := \{x \mid \forall M \in S : x \in M\}$ gebräuchlich; und zwar für die Vereinigung aller Mengen M , die selbst Element einer gegebenen Menge S von Mengen sind. Für obiges V wäre wieder $S := \{M_i : i \in I\}$ zu setzen.

Übungsaufgabe 12. (T) Untersuchen Sie, ob die nachfolgenden Gleichungen für alle Mengen A, B, C, \dots, A_i ($i \in I$), B_j ($j \in J$) gelten. Stets ist eine Begründung zu geben; lautet die Antwort nein, dann in Form eines Gegenbeispiels.

1. $A \cap B = B \cap A$
2. $A \setminus B = B \setminus A$
3. $A \triangle B = B \triangle A$
4. $A \cup (B \cap C) = (A \cup B) \cap C$
5. $A \cup (B \cap C) = (A \cup B) \cap C$
6. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
7. $(\bigcup_{i \in I} A_i) \cap B = \bigcup_{i \in I} (A_i \cap B)$
8. $(\bigcup_{i \in I} A_i) \cap (\bigcup_{j \in J} B_j) = \bigcup_{i \in I} \bigcup_{j \in J} (A_i \cap B_j)$

Hinweis: Oft lassen sich die Mengengleichungen in logische Beziehungen zwischen Aussagen der Form $x \in A$, $x \in B$ etc. umformulieren, deren Allgemeingültigkeit dann zu überprüfen ist.

1.2.4 Kartesische Produkte und Relationen

Inhalt in Kurzfassung: Weil die meisten interessanten mathematischen Aussagen von mehr als nur einem Objekt abhängen (z.B. $a < b$), ist es wichtig, neben Einzelobjekten a, b, \dots auch von Paaren (a, b) , Tripeln (a, b, c) etc. sprechen zu können. Die Menge aller Paare, die aus den Elementen zweier Mengen A, B gebildet werden können, ist das kartesische Produkt $A \times B$. Die Teilmengen von $A \times B$ heißen auch Relationen. Z.B. kann die Relation $<$ auf der Menge \mathbb{N} aller natürlichen Zahlen mit der Menge all jener Paare (a, b) mit $a \in A = \mathbb{N}$ und $b \in B = \mathbb{N}$ identifiziert werden, für die $a < b$ gilt.

Weiter oben im Zusammenhang mit Prädikaten sind bereits geordnete Paare aufgetaucht. Ganz allgemein kommt es bei einem **geordneten Paar** (a, b) nur auf die beiden **Komponenten** a und b und ihre Reihenfolge an. Die entscheidende Eigenschaft lässt sich in der Formel

$$(a, b) = (c, d) \leftrightarrow ((a = c) \wedge (b = d))$$

ausdrücken. Man beachte, dass das geordnete Paar (a, b) nicht mit der Zweiermenge $\{a, b\}$ verwechselt werden darf, weil ja nach Definition der Gleichheit von Mengen immer $\{a, b\} = \{b, a\}$ gilt, während $(a, b) = (b, a)$ nur für $a = b$ zutrifft.⁷

Mittels geordneter Paare lässt sich der Begriff des **kartesischen Produktes**

$$A \times B := \{(a, b) : a \in A, b \in B\}$$

bilden. Dabei handelt es sich also um die Menge aller geordneten Paare (a, b) , deren erste Komponente a ein Element von A und deren zweite Komponente b ein Element von B ist. Die Bezeichnung bezieht sich auf *Cartesius*, die lateinische Namensform für René Descartes (1596-1650). Sie findet sich auch im Begriff des kartesischen Koordinatensystems wieder, der natürlichen Veranschaulichung eines kartesischen Produktes.

Übungsaufgabe 13. (*P*) Erklären Sie die Mengengleichheiten

$$A \times \left(\bigcup_{i \in I} B_i \right) = \bigcup_{i \in I} (A_i \times B) \quad \text{und} \quad A \times \left(\bigcap_{i \in I} B_i \right) = \bigcap_{i \in I} (A_i \times B).$$

Man kann natürlich auch **geordnete Tripel, Quadrupel, ..., n -tupel** definieren durch $(a, b, c) := ((a, b), c)$ und allgemein (rekursiv, wir werden Rekursionen in 1.3.8 noch systematischer behandeln)

$$(a_1, a_2, \dots, a_n, a_{n+1}) := ((a_1, a_2, \dots, a_n), a_{n+1}).$$

Entsprechend können kartesische Produkte mehrerer Mengen $A_i, i = 1, 2, \dots, n$, definiert werden als

$$A_1 \times \dots \times A_n := \{(a_1, \dots, a_n) : a_i \in A_i, i = 1, \dots, n\}.$$

Im Fall $A_1 = \dots = A_n = A$ schreibt man für $A_1 \times \dots \times A_n = A \times \dots \times A$ auch A^n .

Wenn wir uns nochmals an zweistellige Prädikate erinnern, zum Beispiel an die Kleinerbeziehung $<$, so können wir diese übersetzen in die Menge aller Paare (a, b) , für die $a < b$ gilt. Beziehen wir uns beispielsweise auf natürliche oder reelle Zahlen, so wird damit $<$ selbst zu einer Menge (das entspricht dem bereits erwähnten Paradigma, alle Objekte der Mathematik als Mengen aufzufassen),

⁷Es ist möglich, geordnete Paare selbst so als Mengen zu definieren, dass die oben formulierte entscheidende Eigenschaft erfüllt ist, beispielsweise als $(a, b) := \{\{a\}, \{a, b\}\}$. Für uns ist das aber nicht weiter von Bedeutung.

und zwar zur Teilmenge eines kartesischen Produktes (im Beispiel $\mathbb{N} \times \mathbb{N}$ oder $\mathbb{R} \times \mathbb{R}$).

Generell nennt man jede Menge R der Gestalt $R \subseteq A \times B$ eine **Relation** zwischen den Mengen A und B , im Falle $A = B$ auch eine Relation auf A . $R^{(-1)} = \{(b, a) : (a, b) \in R\}$ heißt die **inverse Relation** (manchmal auch **Umkehrrelation**) von R , Beispiel: $>$ ist die inverse Relation von $<$.

Übungsaufgabe 14. (P) Auf einer beliebigen Menge \mathcal{M} von Mengen, kann man die Teilmengenbeziehung \subseteq auch als Relation, d.h. als Menge R von Paaren auffassen. Tun Sie das für die Menge aller Teilmengen von $\{1, 2, 3\}$, indem Sie alle Elemente von R angeben. Hinweis: Erstellen Sie eine passende 8×8 Tabelle.

1.2.5 Der Funktionsbegriff

Inhalt in Kurzfassung: Funktionen (oder Abbildungen) f sind solche Relationen $f \subseteq A \times B$, wo es zu jedem $a \in A$ nur ein $b \in B$ in Relation $(a, b) \in f$ gibt. Man schreibt $b = f(a)$ für dieses eindeutige b und $f : A \rightarrow B$, wenn es zu jedem $a \in A$ auch tatsächlich ein $b \in B$ mit $f(a) = b$ gibt. An den Begriff der Funktion schließen mehrere Komplexe wichtiger Begriffe an: Injektivität etc., Verkettung etc., Kardinalität etc.

Von besonderem Interesse sind bestimmte spezielle Relationen $R \subseteq A \times B$, für die man andere Schreibweisen bevorzugt. Und zwar nennt man $f \subseteq A \times B$ eine **Abbildung** oder **Funktion**⁸ von A nach B und schreibt $f : A \rightarrow B$, wenn es zu jedem $a \in A$ genau ein $b \in B$ gibt mit $(a, b) \in f$. Dieses b nennt man das **Bild** oder den (**Funktions-**) **Wert** von f bei a (an der Stelle a). Man schreibt auch $b = f(a)$ oder $f : a \mapsto b$. A nennt man **Definitionsmenge/-bereich** von f , B **Zielfmenge/-bereich** oder **Wertevorrat**. Jene $b \in B$, die tatsächlich als Bilder von einem oder mehreren Elementen $a \in A$ auftreten, fasst man zur **Bildmenge** $f(A) := \{f(a) : a \in A\} \subseteq B$ zusammen. Oft nennt man $f(A)$ auch die **Wertemenge** und B den **Wertevorrat** von f . Man beachte, dass durch f selbst (nämlich als Menge gewisser geordneter Paare) A und $f(A)$ eindeutig bestimmt sind, nicht aber B . Denn jede Obermenge von $f(A)$ ist als Wertevorrat zulässig. Ist $f(A) = B$, so nennt man $f : A \rightarrow B$ **surjektiv**. Weil dieser Begriff von B abhängt, sagt man der Deutlichkeit halber oft auch **surjektiv auf B** . Haben verschiedene Elemente $a_1 \neq a_2 \in A$ stets verschiedene Bilder (oder gleichbedeutend: folgt aus $f(a_1) = f(a_2)$ stets $a_1 = a_2$), so heißt f **injektiv**. Ist $f : A \rightarrow B$ sowohl surjektiv als auch injektiv, so nennt man f **bijektiv**. Eine bijektive Abbildung nennt man auch eine **Bijektion**. Unter sämtlichen Abbildungen f ist genau für die bijektiven auch die inverse Relation (Umkehrrelation) $f^{(-1)}$ von f wieder eine Abbildung (Funktion) $f^{(-1)} : B \rightarrow A$, genannt die **inverse Abbildung (Funktion)**, **Umkehrabbildung** oder **Umkehrfunktion** von f .

⁸Formal bedeuten beide Worte **Abbildung** und **Funktion** dasselbe. In der reellen Analysis, der große Teile der Vorlesung zuzuordnen sind, werden wir häufiger von Funktionen sprechen, in der Linearen Algebra in der Vorlesung Mathematik 2 hingegen wird viel von (linearen) Abbildungen die Rede sein.

Übungsaufgabe 15. (T) Durch welche der folgenden Vorschriften wird eine Funktion definiert? Falls eine Funktion zwischen den angegebenen Mengen vorliegt, entscheiden Sie, ob diese injektiv, surjektiv oder bijektiv ist.

1. $f : \{\text{Studierende der TU Wien}\} \rightarrow \mathbb{N}$ und $f(x) = \text{Matrikelnummer von } x$.
2. $f : \{\text{Studierende der TU Wien}\} \rightarrow \{A, B, AB, 0\}$ und $f(x) = \text{Blutgruppe von } x$.
3. $f : \{\text{Noten in der Partitur von Beethovens 9. Symphonie}\} \rightarrow \{d, e, f, g, a, b, c\}$ und $f(x) = \text{Tonname von } x$. Hinweis: „Freude“
4. $f : \{\text{Einwohner von Österreich}\} \rightarrow \{\text{Vornamen}\}$ und $f(x) = \text{Vorname der Tochter von } x$.
5. $f : \{\text{Einwohner von Österreich}\} \rightarrow \{\text{Vornamen}\}$ und $f(x) = \text{Vorname des Vaters von } x$.
6. $f : \{\text{Elemente des Periodensystems}\} \rightarrow \mathbb{N}$ und $f(x) = \text{Ordnungszahl von } x$.
7. $f : \{\text{Elemente des Periodensystems}\} \rightarrow \mathbb{N}$ und $f(x) = \text{Massenzahl von } x$.

Beispiele von Funktionen werden uns sehr bald und dann fast ständig begleiten. (Ein typisches Beispiel ist etwa $f(x) := x^2$ als weder injektive noch surjektive reelle Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$.)

Oft betrachtet man die Wirkung einer Funktion $f : D \rightarrow X$ auf einer Teilmenge $T \subseteq D$ des Definitionsbereichs D . Genauer: Die Menge $f_T := \{(x, f(x)) : x \in T\}$ ist dann wieder eine Funktion, nämlich $f_T : T \rightarrow X$. Man nennt f_T die **Einschränkung** von f auf T , f eine **Fortsetzung** von f_T auf D . Meist verzichtet man auf den Index T in f_T und bezeichnet, ungenau, Einschränkung und Fortsetzung mit demselben Symbol.

Auf einer beliebigen Menge M gibt es die sogenannte **identische Funktion** oder **Identität** $\text{Id}_M : M \rightarrow M$, wo jedes $m \in M$ sich selbst als Bild zugeordnet wird, also $m \mapsto \text{Id}_M(m) := m$. Von besonderer Wichtigkeit ist das Konzept der **Verkettung**, **Komposition**, **Verknüpfung** oder **Hintereinanderausführung** (oder manchmal auch etwas missverständlich **Multiplikation**) $g \circ f : A \rightarrow C$ zweier Funktionen $f : A \rightarrow B$ und $g : B \rightarrow C$. Sie ist definiert durch $(g \circ f)(a) := g(f(a))$ für alle $a \in A$. Ein Element a wird also zuerst f unterworfen, das Resultat $f(a) \in B$ schließlich g . Man sieht sehr schnell:

1. $f \circ \text{Id}_A = f$ und $\text{Id}_B \circ f = f$
2. $(h \circ g) \circ f = h \circ (g \circ f)$ für $h : C \rightarrow D$
3. $f^{(-1)} \circ f = \text{Id}_A$ und $f \circ f^{(-1)} = \text{Id}_B$ für bijektives f

Übungsaufgabe 16. (T) Sei $A = \{1, 2, 3\}$. Wie viele Abbildungen $f : A \rightarrow A$ gibt es? Welche davon sind injektiv, surjektiv, bijektiv? Geben Sie die Umkehrabbildungen $f^{(-1)}$ aller bijektiven $f : A \rightarrow A$ an.

Sei außerdem $B := \{1, 2, 3, 4\}$. Wieviele injektive, surjektive, bijektive $g : A \rightarrow B$ und $h : B \rightarrow A$ gibt es dann?

Übungsaufgabe 17. (P) Seien $f : A \rightarrow B$ und $g : B \rightarrow C$ irgendwelche Abbildungen und $h := g \circ f$, also $h(a) := g(f(a))$. Zeigen Sie:

1. Sind f und g injektiv, so auch h .
2. Sind f und g surjektiv, so auch h .
3. Sind f und g bijektiv, so auch h .

Übungsaufgabe 18. (P) Seien wieder $f : A \rightarrow B$ und $g : B \rightarrow C$ irgendwelche Abbildungen und $h := g \circ f$.

1. Sei h injektiv. Folgt daraus auch die Injektivität von f oder von g ?
2. Sei h surjektiv. Folgt daraus auch die Surjektivität von f oder von g ?
3. Sei h bijektiv. Folgt daraus auch die Bijektivität von f oder von g ?

Begründen Sie jeweils Ihre Antwort (im negativen Fall ist ein Gegenbeispiel anzugeben).

Auch die für die Mathematik so wichtigen Rechenoperationen lassen sich als spezielle Beispiele von Funktionen auffassen. Zum Beispiel ist die Addition $+$ eine Funktion, die jedem Zahlenpaar (a, b) seine Summe $a + b$ zuordnet, also $+$: $(a, b) \mapsto a + b$. Als Definitionsbereich ist die Menge aller Paare aus dem zugrunde gelegten Zahlenbereich zu nehmen, im umfassendsten Fall also z.B. $+$: $\mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$. Auf die Zahlenbereiche kommen wir in Abschnitt 1.3 aber noch viel ausführlicher zu sprechen.

Übungsaufgabe 19. (T) Viele aus der Schule bekannte Verknüpfungen und Operationen (bei Bedarf können Sie auch im Index nachsehen) sind assoziativ, d.h. $a \star (b \star c) = (a \star b) \star c$. Dabei bezeichnet \star die fragliche Verknüpfung. Eine weitere interessante Eigenschaft ist die Kommutativität, d.h. die Gültigkeit von $a \star b = b \star a$. Untersuchen Sie folgende Beispiele auf Assoziativität und Kommutativität:

1. Exponentiation: $a \star b := a^b$ mit $a, b \in \mathbb{N}$.
2. Größter gemeinsamer Teiler: $n \star m := \text{ggT}(n, m)$ mit $n, m \in \mathbb{N}$.
3. Kreuz- oder Vektorprodukt: $\mathbf{x} \star \mathbf{y} := \mathbf{x} \times \mathbf{y}$, für Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ (Vorgriff auf 1.5.7).

Geben Sie jeweils ein Gegenbeispiel an, falls eine Eigenschaft verletzt ist.

Seien $f : A \rightarrow B$ und A, B endliche Mengen. Ist f injektiv, so beobachten wir $|A| \leq |B|$; ist f surjektiv oder gar bijektiv, so gilt analog $|A| \geq |B|$ bzw. $|A| = |B|$. Cantor nutzte genau diese Eigenschaften injektiver und bijektiver Abbildungen, um auch für unendliche Mengen Größenvergleiche anzustellen, indem er definierte: A ist kleiner oder gleich B (die **Mächtigkeit** oder **Kardinalität** von B ist mindestens so groß wie die von A), symbolisch $|A| \leq |B|$, falls es ein

injektives $f : A \rightarrow B$ gibt. Gibt es sogar ein bijektives $f : A \rightarrow B$, so nennt man A und B gleich groß (von gleicher Mächtigkeit) und schreibt $|A| = |B|$. (Unendliche) Mengen, die gleich groß sind wie die Menge $\mathbb{N} = \{0, 1, 2, \dots\}$ heißen **abzählbar** (das ist die kleinste mögliche Größe einer unendlichen Menge), solche die strikt größer sind als \mathbb{N} heißen **überabzählbar**. Explizit bedeutet die Abzählbarkeit einer unendlichen Menge A , dass man alle ihre Elemente der Reihe nach als sogenannte Folge (Abzählung) a_0, a_1, a_2, \dots auflisten kann. Denn nimmt man kein Element $a \in A$ dabei doppelt, so ist die Zuordnung $n \mapsto a_n$ eine Bijektion zwischen $\mathbb{N} = \{0, 1, 2, \dots\}$ und $A = \{a_0, a_1, a_2, \dots\}$ wie gefordert.

Als Geburtsstunde der Mengenlehre wird allgemein jener Moment (vermutlich im Jahr 1873) angesehen, als Cantor bemerkte, dass zwar $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$ gilt (dass also jede der Mengen \mathbb{N} der natürlichen, \mathbb{Z} der ganzen und \mathbb{Q} der rationalen Zahlen abzählbar ist), dass aber die Menge \mathbb{R} der reellen Zahlen überabzählbar ist, siehe Abschnitt 1.3, insbesondere Übungsaufgabe 64. Allgemein schreibt man $|A| < |B|$ wenn es eine injektive, aber keine bijektive Abbildung $f : A \rightarrow B$ gibt. Insbesondere gilt also $|\mathbb{N}| < |\mathbb{R}|$. Wir verzichten auf die Vertiefung dieses interessanten Themas. Wer will ist jedoch eingeladen, sich mit der folgenden Übungsaufgabe zu beschäftigen, die allerdings auch Vorgriffe auf enthält.

Nun ist es höchste Zeit, die hier bereits auftretenden Zahlenmengen systematisch zu behandeln.

1.3 Die Zahlenbereiche

Wie schon früher erwähnt, werden in der modernen reinen Mathematik alle Objekte aus Mengen aufgebaut. Man beginnt mit der Zahl 0 als leerer Menge und kann daraus nacheinander die Mengen \mathbb{N} der natürlichen, \mathbb{Z} der ganzen, \mathbb{Q} der rationalen, \mathbb{R} der reellen und \mathbb{C} der komplexen Zahlen konstruieren. Aus Sicht der Ingenieurwissenschaften ist diese Vorgehensweise aber weniger interessant. Statt dessen geben wir zunächst einen Überblick über die genannten Zahlenbereiche (1.3.1). Sodann wenden wir uns direkt dem System \mathbb{R} der reellen Zahlen zu, das für uns die wichtigste Rolle spielt. Am besten stellt man sich \mathbb{R} als Zahlengerade vor (1.3.2), zu der aber noch die arithmetischen Operationen kommen. Insgesamt lässt sich \mathbb{R} sehr gut axiomatisch als vollständig angeordneter Körper beschreiben, woraus sich die elementaren Rechenregeln ableiten lassen (1.3.3). Die sogenannte Vollständigkeit ist hauptverantwortlich für die Supremumseigenschaft (1.3.4), für die archimedische Eigenschaft (1.3.5) und für die Existenz beliebiger Wurzeln positiver reeller Zahlen (1.3.6). Die Mengen \mathbb{N} , \mathbb{Z} und \mathbb{Q} finden sich als Teilmengen von \mathbb{R} wieder. Besonderes Augenmerk verdienen bei \mathbb{N} die eng verwandten Themen Induktion (1.3.7), Anwendung derselben in Form von Rekursionen und Induktionsbeweisen (1.3.8), bei \mathbb{N} und \mathbb{Z} Teilbarkeit und Primfaktorzerlegung (1.3.9), außerdem Zahlendarstellungen generell (1.3.10) und bei \mathbb{C} die komplexe Zahlenebene (1.3.11).

1.3.1 Die Zahlenmengen $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ im Überblick

Inhalt in Kurzfassung: Ein kurzer Überblick zeigt, aus welchen Bedürfnissen heraus die Zahlenmengen \mathbb{N} (endliche Kardinalitäten), \mathbb{Z} (Subtraktion), \mathbb{Q} (Division), \mathbb{R} (Zahlengerade ohne Lücken) und \mathbb{C} ($\sqrt{-1}$) zustande kommen.

Die Menge $\mathbb{N} = \{0, 1, 2, \dots\}$ der **natürlichen Zahlen** besteht aus jenen Zahlen, die Mächtigkeiten endlicher Mengen angeben. Auf \mathbb{N} sind die beiden Rechenoperationen **Addition** und **Multiplikation** uneingeschränkt ausführbar, genauer: Zu allen $a, b \in \mathbb{N}$ gibt es in \mathbb{N} die **Summe**, für die wir $a + b$, und das **Produkt**, für das wir $a \cdot b$ oder auch nur ab schreiben. Begrifflich sind also $+$ und \cdot , wenn wir sie nur auf natürliche Zahlen anwenden, Abbildungen von $\mathbb{N} \times \mathbb{N}$ nach \mathbb{N} . Diese beiden Operationen entsprechen den mengentheoretischen Operationen \cup und \times , genauer: Sind A, B endliche Mengen mit $|A| = a$ und $|B| = b$, so gilt, sofern A und B disjunkt sind, $|A \cup B| = a + b$, und (auch im nicht disjunkten Fall) $|A \times B| = ab$. Außerdem schreibt man $a \leq b$, wenn es Mengen A und B mit $|A| = a$, $|B| = b$ und $A \subseteq B$ gibt oder, äquivalent, mit $|A| \leq |B|$ (im Sinne des Größenvergleichs von Mengen mittels injektiver Abbildungen). Ist $a \leq b$ und $a \neq b$ (d.h. $\neg a = b$), so schreibt man $a < b$ oder auch $b > a$. Daraus ergeben sich mancherlei vertraute Rechenregeln, auf die wir an späterer Stelle noch ausdrücklich zu sprechen kommen. Unter anderem macht man sich klar: Gibt es zu $a, b \in \mathbb{N}$ eine Lösung $c \in \mathbb{N}$ der Gleichung $a + x = b$, d.h. ein c mit $a + c = b$, so ist diese Lösung c eindeutig bestimmt. Man schreibt dafür auch $c = b - a$ und nennt c das Resultat der **Subtraktion** oder auch die **Differenz** von b und a .

Ist jedoch $b < a$, so gibt es in \mathbb{N} keine Differenz $b - a$. Um sicherzustellen, dass Differenzen immer existieren, erweitert man \mathbb{N} zur Menge $\mathbb{Z} = \{\dots - 3, -2, -1, 0, 1, 2, 3, \dots\}$ der **ganzen Zahlen**. Auf ganz \mathbb{Z} ist der **Betrag** definiert durch $|n| = |-n| = n$ für alle $n \in \mathbb{N}$. In der gewohnten Weise sind in \mathbb{Z} Addition, Multiplikation und Subtraktion uneingeschränkt ausführbar, außerdem ist auch auf \mathbb{Z} Größenvergleich möglich, z.B. $-5 < -3$. Überlegungen analog zu denen, die uns von \mathbb{N} zu \mathbb{Z} geführt haben, lassen sich auf \mathbb{Z} anwenden, wenn man von Gleichungen $ax = b$ ausgeht: Sind a und b ganze Zahlen, so gibt es eine Lösung $c \in \mathbb{Z}$ definitionsgemäß nur dann, wenn a ein **Teiler** von b ist (man sagt in diesem Fall auch: a teilt b), symbolisch $a|b$. In diesem Fall ist für $a \neq 0$ die Lösung eindeutig bestimmt. Man schreibt dafür $c = \frac{b}{a}$ (manchmal auch $c = b : a$ oder $c = b/a$) und nennt c das Resultat der **Division** oder auch den **Quotienten** oder **Bruch** von b durch a mit **Zähler** b und **Nenner** a . Ist a kein Teiler von b , so gibt es in \mathbb{Z} keine Lösung, was zur Erweiterung von \mathbb{Z} auf \mathbb{Q} führt.

Die Menge $\mathbb{Q} = \{\frac{b}{a} : a, b \in \mathbb{Z}, a \neq 0\}$ der **rationalen Zahlen** besteht aus allen Brüchen, wo Zähler und Nenner ganzzahlig sind, der Nenner allerdings nicht 0. Bekanntlich können verschiedene Brüche dieselbe Zahl darstellen, z.B. $\frac{2}{3} = \frac{4}{6}$. Stets gibt es eine **gekürzte** Darstellung, wo die Beträge von Zähler und Nenner nicht mehr durch Kürzen verkleinert werden können. In der bekannten Weise sind auf \mathbb{Q} alle Grundrechnungsarten Addition, Multiplikation, Subtrak-

tion und Division (außer durch 0) uneingeschränkt ausführbar. Außerdem lassen sich auch beliebige rationale Zahlen der Größe nach vergleichen.

Ein mögliches Ungenügen an \mathbb{Q} ist von anderer Art als bei \mathbb{N} und \mathbb{Z} . Und zwar stellt sich heraus, dass es in \mathbb{Q} Löcher gibt in folgendem Sinn: Angenommen man sucht eine Zahl $x > 0$ mit $x^2 = 2$ (die Wurzel aus 2, symbolisch $\sqrt{2}$ hätte diese Eigenschaft), so lässt sich schnell einsehen, dass so ein x nicht in \mathbb{Q} liegen kann.

Der klassische Beweis macht von elementaren Rechenregeln Gebrauch: So ein x hätte eine gekürzte Darstellung $x = \frac{p}{q}$ als Bruch natürlicher Zahlen. Aus $x^2 = \frac{p^2}{q^2} = 2$ folgt $p^2 = 2q^2$. Also ist p^2 gerade, was nur für gerades $p = 2k$, $k \in \mathbb{N}$, möglich wäre, also $4k^2 = p^2 = 2q^2$. Kürzen durch 2 zeigt, dass auch $q^2 = 2k^2$ gerade ist, folglich auch $q = 2l$. Die Beziehung $\frac{p}{q} = \frac{2k}{2l} = \frac{k}{l}$ liefert einen Widerspruch zur Gekürztheit der Darstellung $\frac{p}{q}$. Also kann x nicht in \mathbb{Q} liegen.

Übungsaufgabe 20. (P) In dieser Argumentation wurde davon Gebrauch gemacht, dass das Produkt zweier ungerader Zahlen wieder ungerade ist. Geben Sie an, wo genau dies verwendet wurde, und beweisen sie es.

Ein x mit $x^2 = 2$ wollen wir aber unbedingt zur Verfügung haben, nämlich z.B. als Länge der Diagonale eines Quadrats mit Seitenlänge 1 (Pythagoras!).

Übungsaufgabe 21. (E) Formulieren und beweisen Sie den **Satz von Pythagoras**. (Schulstoff!)

Deshalb ergänzt man die Menge \mathbb{Q} der rationalen Zahlen um alle sogenannten **irrationalen** Zahlen (da gehören z.B. auch die Zahlen e und π oder $\ln 2$ dazu) zur Menge \mathbb{R} der sogenannten **reellen Zahlen**. Man stellt sich \mathbb{R} am besten als sogenannte Zahlengerade vor, wo jeder Punkt genau einer reellen Zahl entspricht und umgekehrt. Wir kommen etwas später darauf zurück, wie man \mathbb{R} und damit auch die Mengen \mathbb{N} , \mathbb{Z} und \mathbb{Q} sehr genau beschreiben kann.

Übungsaufgabe 22. (E) Zeigen Sie, dass auch $\sqrt{3}$ irrational ist. Hinweis: Orientieren Sie sich am Beweis für die Irrationalität von $\sqrt{2}$, achten Sie aber sorgfältig auf die Unterschiede und darauf, was sie in Ihrer Argumentation verwenden.

Wir haben die Zahl $\sqrt{2}$ benutzt, um die Erweiterung von \mathbb{Q} zu \mathbb{R} zu motivieren. Das sollte aber keinesfalls das Missverständnis suggerieren, dass man in \mathbb{R} nun beliebige Wurzeln ziehen kann. Denn Quadrate reeller Zahlen sind immer nichtnegativ, also haben negative Zahlen keine reellen Quadratwurzeln. Trotzdem kann man sich solche wünschen, also etwa $\sqrt{-1}$. Tatsächlich kann dieser Wunsch sogar erfüllt werden. Wenn man eine solche Quadratwurzel von -1 wie üblich mit dem Symbol i (für **imaginäre Einheit**) bezeichnet, erweist es sich als sinnvoll, die Menge $\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$ der **komplexen Zahlen** zu bilden. Diese Menge \mathbb{C} leistet vielerlei auf einmal: Sie enthält alle reellen Zahlen $r = r + i0$ sowie die imaginäre Einheit $i = 0 + i1$, und alle vier Grundrechenarten sind in \mathbb{C} gewohnter Weise (d.h. außer Division durch 0) ausführbar. Darüber

hinaus lässt sich sehr viel Interessantes über \mathbb{C} sagen, was aber erst etwas später getan werden soll.

1.3.2 Veranschaulichung von \mathbb{R} als Zahlengerade

Inhalt in Kurzfassung: Die geometrische Veranschaulichung von \mathbb{R} als Menge der Punkte auf der Zahlengeraden bewährt sich besonders in Bezug auf die Kleinerrelation $<$, die sich als „links von“ interpretieren lässt. Damit lassen sich auch abgeschlossene, offene und halboffene Intervalle sehr gut beschreiben.

Bevor wir in 1.3.3 ein Axiomensystem angeben, welches \mathbb{R} präzise und sogar in einem gewissen Sinne vollständig beschreibt, wollen wir mit der anschaulichen Deutung beginnen. Und zwar verstehen wir die Menge \mathbb{R} aller reellen Zahlen als die Menge der Punkte auf einer (nach beiden Richtungen unbeschränkten) Geraden, der meist waagrecht gezeichneten **Zahlengeraden**. Liegt ein Punkt a links vom Punkt b , so schreiben wir $a < b$ oder $b > a$ bzw. $a \leq b$ oder $b \geq a$, wenn wir auch $a = b$ zulassen wollen. Der Punkt 0 teilt die Zahlengerade in einen positiven Teil $\mathbb{R}^+ := \{x \in \mathbb{R} : x > 0\}$ und einen negativen $\mathbb{R}^- := \{x \in \mathbb{R} : x < 0\}$. Außerdem schreiben wir $\mathbb{R}_0^+ := \{x \in \mathbb{R} : x \geq 0\}$ und $\mathbb{R}_0^- := \{x \in \mathbb{R} : x \leq 0\}$.

Besonders wichtige Teilmengen von \mathbb{R} sind die sogenannten **Intervalle**, bestehend aus allen Punkten x zwischen den **Randpunkten** a und b . Für $a \leq b$ schreiben wir

$$\begin{aligned} [a, b] &:= \{x \in \mathbb{R} : a \leq x \leq b\} \text{ (abgeschlossenes Intervall),} \\ (a, b) &:= \{x \in \mathbb{R} : a < x < b\} \text{ (offenes Intervall),} \\ (a, b] &:= \{x \in \mathbb{R} : a < x \leq b\} \text{ (links halboffenes Intervall) und} \\ [a, b) &:= \{x \in \mathbb{R} : a \leq x < b\} \text{ (rechts halboffenes Intervall).} \end{aligned}$$

Übungsaufgabe 23. (T) Wählen Sie ein konkretes $x \in \mathbb{R}$ und konkrete Zahlen $\varepsilon_1 > \varepsilon_2 > 0$. Zeichnen Sie die Mengen $\{y \in \mathbb{R} : |x - y| < \varepsilon_1\}$ und $(x - \varepsilon_2, x + \varepsilon_2)$ auf der Zahlengeraden ein.

Übungsaufgabe 24. (T) Gegeben seien drei nichtleere Intervalle $A := (a_1, b_1]$, $B = (a_2, b_2)$ und $C = [a_3, b_3]$ mit konkreten Randpunkten $a_1 < b_1, a_2 < b_2, a_3 < b_3 \in \mathbb{R}$ vor. Bestimmen Sie die Mengen

$$A^c, \quad B^c, \quad C^c, \quad A \cap (B \cup C), \quad A \cup (B \cap C), \quad A \setminus C, \quad B \triangle C, \quad ((A \setminus B) \cap C) \cup A$$

und schreiben Sie diese als Intervall bzw. als Vereinigung von Intervallen an. Diskutieren Sie, welche Möglichkeiten denkbar sind, je nach Lage der Randpunkte der Intervalle zueinander. (Zur Erinnerung: A^c ist das Komplement von A , hier relativ in Bezug auf \mathbb{R} ; $B \triangle C$ ist die symmetrische Differenz von B und C .)

Übungsaufgabe 25. (T) Wählen Sie ein konkretes $a > 0$ und zeichnen Sie die beiden Mengen $\{x : x < a\}$ und $\{x : x^2 < a^2\}$ auf der Zahlengeraden ein. Ist eine der beiden Mengen in der anderen enthalten? Wiederholen Sie die Vorgehensweise mit $-a$ anstelle von a .

1.3.3 Axiomatik am Beispiel von \mathbb{R}

Inhalt in Kurzfassung: Das System \mathbb{R} der reellen Zahlen, genauer der vollständig angeordnete Körper \mathbb{R} , lässt sich durch eine überschaubare Liste von Eigenschaften (Axiomen) nicht nur gut beschreiben, sondern (in einem streng präzisierbaren Sinn) auch charakterisieren. Dabei treten drei Strukturen in Verbindung miteinander: die additive Gruppe auf \mathbb{R} , die multiplikative Gruppe auf $\mathbb{R} \setminus \{0\}$ und die durch \leq gegebene Ordnungsstruktur auf \mathbb{R} . Axiome an jede dieser drei Strukturen für sich zusammen mit solchen, die zwischen je zwei dieser Strukturen gewisse Verträglichkeiten zum Ausdruck bringen, tragen die gesamte interessante Information über \mathbb{R} .

Das System der reellen Zahlen lässt sich durch eine überschaubare Liste von Daten und Forderungen (Axiomen) sehr gut, in einem präzisierbaren Sinne sogar vollständig beschreiben als sogenannter **vollständig angeordneter Körper**. Man kann rein mengentheoretisch ein solches System auf verschiedene Arten konstruieren. Wir bevorzugen jedoch den axiomatischen Zugang, weil die Axiome durchwegs allgemein vertraute Gesetze sind. Es kommt nicht darauf an, die Liste dieser Axiome auswendig zu lernen. Bemerkenswert ist vor allem, dass eine derartige Liste ausreicht, um als Axiomensystem für \mathbb{R} zu fungieren. Die Mengen \mathbb{N} , \mathbb{Z} und \mathbb{Q} werden wir dann als Teilmengen von \mathbb{R} identifizieren können.

Die Daten sind die Menge \mathbb{R} und darauf die Addition $+$, die Multiplikation \cdot und die Ordnungsrelation \leq . Die Forderungen (Axiome) sind die folgenden.

\mathbb{R} bildet mit der Addition $+$ eine sogenannte **abelsche Gruppe**. Das wiederum beinhaltet:

1. **Assoziativgesetz:** $(a + b) + c = a + (b + c)$ für alle $a, b, c \in \mathbb{R}$.
2. Existenz eines **neutralen Elements**, in diesem Fall $0 \in \mathbb{R}$ mit $a + 0 = 0 + a$ für alle $a \in \mathbb{R}$.
3. Zu jedem $a \in \mathbb{R}$ existiert ein **inverses Element**, nämlich $-a$, mit $a + (-a) = (-a) + a = 0$ (neutrales Element).
4. Die ersten drei Forderungen definieren eine **Gruppe**, abelsch wird sie durch das (früher schon einmal erwähnte) **Kommutativgesetz:** $a + b = b + a$ für alle $a, b \in \mathbb{R}$.

Dieselben Gesetze gelten für die Addition auch auf \mathbb{Z} , \mathbb{Q} und \mathbb{C} . In \mathbb{N} gelten mangels inverser Elemente aber nur 1., 2., und 4.

Auch die Multiplikation \cdot ist assoziativ, und eingeschränkt auf $\mathbb{R} \setminus \{0\}$ liegt ebenfalls eine abelsche Gruppe vor. Das neutrale Element bezüglich \cdot ist $1 \in \mathbb{R} \setminus \{0\}$, das Inverse zu $a \neq 0$ bezeichnen wir mit a^{-1} . Analog zu oben wird also gefordert (für alle a, b, c):

1. $(ab)c = a(bc)$

2. $a1 = 1a = a$
3. $aa^{-1} = a^{-1}a = 1$, sofern $a \neq 0$
4. $ab = ba$

Addition und Multiplikation sind durch das **Distributivgesetz** miteinander verbunden: $a(b+c) = ab+ac$ für alle $a, b, c \in \mathbb{R}$.⁹

Die bisherigen Gesetze gelten für die Addition und Multiplikation auch auf \mathbb{Q} und \mathbb{C} – man spricht von einem **Körper**. \mathbb{Z} hingegen ist mangels multiplikativer Inverser kein Körper (und \mathbb{N} schon gar nicht).

Viele vertraute Rechenregeln gelten in allen Körpern, weil sie sich allein aus den bisher genannten Axiomen ableiten lassen. Dazu einige einfache Beispiele, bei denen man sich jeweils klar mache, dass an jeder Stelle eines der Axiome für einen Körper oder bereits zuvor Bewiesenes verwendet, sonst aber nichts benötigt wird.

Erstes Beispiel: $-(-a) = a$ für alle $a \in \mathbb{R}$: Für das Inverse $-(-a)$ von $-a$ gilt nach Definition $(-a) + (-(-a)) = 0$. Addieren wir auf beiden Seiten der Gleichung a von links, erhalten wir auf der linken Seite $a + ((-a) + (-(-a))) = (a + (-a)) + (-(-a)) = 0 + (-(-a)) = -(-a) + 0 = -(-a)$, auf der rechten Seite $a + 0 = a$. Also stimmen beide Seiten überein: $-(-a) = a$.

Zweites Beispiel: $0a = a0 = 0$ für alle $a \in \mathbb{R}$: $0a = (0+0)a = 0a + 0a$, also liefert Addition des additiven Inversen $-0a$ von $0a$ auf beiden Seiten $0 = 0a + (-0a) = (0a + 0a) + (-0a) = 0a + (0a + (-0a)) = 0a + 0 = 0a = a0$, die Behauptung.

Drittes Beispiel: $(-a)(-b) = ab$. Nach dem Distributivgesetz und der eben bewiesenen Beziehung ist $ab + a(-b) = a(b + (-b)) = a0 = 0a = 0$. Addition des additiven Inversen $-ab$ liefert $a(-b) = -ab$, wegen der Kommutativität analog $(-a)b = -ab$, folglich $(-a)(-b) = -a(-b) = -(-ab) = ab$.

Übungsaufgabe 26. (E) Zeigen Sie mit Hilfe der Axiome für \mathbb{R} , dass dort (wie in jedem Körper) die folgenden beiden Kürzungsregeln gelten:

1. Aus $a+x = a+y$ folgt $x = y$ (für alle $x, y \in \mathbb{R}$).
2. Ist $a \neq 0$, so folgt $x = y$ auch aus $ax = ay$ (für alle $x, y \in \mathbb{R}$).

Wie man schon allein aufgrund der Darstellung von \mathbb{R} als Zahlengerade, wo ja kleinere Zahlen links und größere rechts liegen, erkennt, ist über die arithmetische (d.h. durch die Rechenoperationen gegebene) Struktur hinausgehend auch die **Ordnungsrelation** \leq wesentlich. Man verlangt (jeweils für alle $a, b, c \in \mathbb{R}$):

1. **Reflexivität:** $a \leq a$.
2. **Antisymmetrie:** Gilt gleichzeitig $a \leq b$ und $b \leq a$, so ist $a = b$.

⁹Man beachte die Bindungsregel *Punkt geht vor Strich*, wonach der Ausdruck $ab+ac$ im Sinne der Klammerung $(a \cdot b) + (a \cdot c)$ zu verstehen ist.

3. **Transitivität:** Aus $a \leq b$ und $b \leq c$ folgt $a \leq c$.
4. **Trichotomie:** Für $a, b \in \mathbb{R}$ gilt stets genau eine der drei Möglichkeiten $a < b$, $a = b$ oder $a > b$.

Die Ordnungsrelation ist mit Addition und Multiplikation verträglich im Sinn der beiden ebenfalls als Axiome geforderten **Monotoniegesetze**:

1. Aus $a, b, c \in \mathbb{R}$ und $a \leq b$ folgt $a + c \leq b + c$.
2. Aus $a, b, c \in \mathbb{R}$, $c \geq 0$ und $a \leq b$ folgt $ac \leq bc$.

Aus diesen Forderungen lassen sich gemeinsam mit den vorangegangenen weitere einfache Rechenregeln ableiten wie z.B.: $a \geq 0$ genau dann wenn $-a \leq a$; $a \leq b$ genau dann wenn $-b \leq -a$; für $0 < a, b$ ist $a \leq b$ genau dann wenn $b^{-1} \leq a^{-1}$.

Alle bisher ausgesprochenen Gesetze (Axiome) samt ihren Folgerungen gelten entsprechend auch für \mathbb{Q} (man spricht von einem **angeordneten Körper**),

Übungsaufgabe 27. (P) Leiten Sie die obigen drei Äquivalenzen für $a, b \in \mathbb{R}$ aus den definierenden Eigenschaften eines angeordneten Körpers ab:

1. $a \geq 0$ genau dann wenn $-a \leq a$.
2. $a \leq b$ genau dann wenn $-b \leq -a$.
3. Für $0 < a, b$ ist $a \leq b$ genau dann wenn $b^{-1} \leq a^{-1}$.

Das wichtigste Beispiel eines Körpers, der nicht angeordnet ist, ist der Körper \mathbb{C} der komplexen Zahlen. Dies wird später aus Übungsaufgabe 77 hervorgehen.

Übungsaufgabe 28. (T) In dieser Aufgabe darf aus der Schule Bekanntes (Eigenschaften involvierter Funktionen wie Wurzeln, Winkelfunktionen und Logarithmus, ungefähre numerischer Wert der Kreiszahl π) verwendet werden, nicht jedoch elektronische Rechenhilfen. Entscheiden Sie, welche der beiden Zahlen größer ist:

1. $\sqrt{7} + \sqrt{11}$ oder $\sqrt{8} + \sqrt{10}$
2. 42 oder $\sqrt{17^2 + 25^2}$
3. $\log_{10} 1000$ oder $\log_{10} 500 + \log_{10} 500$
4. $\sin 3,14 + \sin 3,15$ oder 0
5. $\cos 3,14 - \cos 3,15$ oder 0
6. $\sin(1)$ oder $\cos(1)$

Der anhand der bisherigen Eigenschaften noch nicht erkennbare, aber entscheidende Unterschied gegenüber \mathbb{Q} ist die **Vollständigkeit** von \mathbb{R} . Sie lässt sich wie folgt fassen:

Angenommen, \mathbb{R} wird in zwei nichtleere Mengen A und B mit $A < B$ zerlegt. Das soll bedeuten: $A, B \subseteq \mathbb{R}$, $A \cup B = \mathbb{R}$, $A \neq \emptyset \neq B$ und $a < b$ für alle $a \in A$ und $b \in B$. Dann garantiert die Vollständigkeit, dass es einen Grenzpunkt zwischen A und B gibt, also ein $x \in \mathbb{R}$ (x kann zu A oder auch zu B gehören) mit $a \leq x \leq b$ für alle $a \in A$ und $b \in B$. Man kann sich x als Schnittpunkt vorstellen, an dem die Zahlengerade bei der Zerlegung in die Mengen A und B zerschnitten wird.

Man beachte, dass \mathbb{Q} tatsächlich nicht vollständig in diesem Sinn ist. Man wähle dazu irgendeine irrationale Zahl r wie etwa $\sqrt{2}$, π oder e und nehme als A und B die Mengen aller $q \in \mathbb{Q}$ mit $q < r$ bzw. mit $q > r$. Der Grenzpunkt („Schnittpunkt“) zwischen A und B kann nur r selbst sein, r ist aber in \mathbb{Q} nicht als Element vorhanden.

Somit beschreibt die gegebene Axiomatik \mathbb{R} hinreichend genau, um \mathbb{R} von den anderen Zahlenbereichen \mathbb{N} , \mathbb{Z} , \mathbb{Q} und \mathbb{C} zu unterscheiden. Doch nicht nur das: Man kann beweisen, dass jedes System, das die genannten Bedingungen erfüllt, dieselbe Struktur wie \mathbb{R} hat, also – wie man sagt – zu \mathbb{R} **isomorph** ist. In diesem Sinn ist also die Struktur von \mathbb{R} als **vollständig angeordneter Körper** eindeutig bestimmt, und wir dürfen für alles Weitere von diesen Eigenschaften ausgehen.

Es fällt leicht, die Mengen \mathbb{N} , \mathbb{Z} und \mathbb{Q} als Teilmengen von \mathbb{R} zu identifizieren: \mathbb{N} enthält die (eindeutig bestimmten) neutralen Elemente bezüglich $+$ und \cdot , also 0 und 1, sowie alle Zahlen, die sich daraus durch iterierte Addition von 1 bilden lassen; sonst nichts. Bei \mathbb{Z} kommen die additiven Inversen $-n$ aller $n \in \mathbb{N}$ dazu, bei \mathbb{Q} alle Quotienten $\frac{a}{b} := ab^{-1}$ mit $a, b \in \mathbb{Z}$ und $b \neq 0$. Noch formaler kann man definieren

$$\mathbb{N} := \bigcap \{T : T \subseteq \mathbb{R} \wedge 0 \in T \wedge (\forall t : t \in T \rightarrow t + 1 \in T)\},$$

\mathbb{Z} und \mathbb{Q} entsprechend.

Übungsaufgabe 29. (E) Ähnlich zu obiger Definition der Menge \mathbb{N} sind auch \mathbb{Z} und \mathbb{Q} als Teilmengen von \mathbb{R} formal anzuschreiben. Hinweis: Die Abgeschlossenheit von \mathbb{N} bezüglich Addition von 1 ($t \mapsto t+1$) ist durch stärkere Abgeschlossenheitseigenschaften zu ersetzen.

1.3.4 Die Supremumseigenschaft als Konsequenz der Vollständigkeit

Inhalt in Kurzfassung: Die für die Analysis entscheidende Eigenschaft, die \mathbb{R} gegenüber \mathbb{Q} auszeichnet, ist die Vollständigkeit. In 1.3.3 wurde sie über die Existenz eines Schnittpunktes definiert, wo immer man die Zahlengerade in einen linken und einen rechten Teil zerlegt. Äquivalent dazu ist auch die Existenz von

Infimum (größte untere Schranke) und Supremum (kleinste obere Schranke) einer beliebigen nichtleeren beschränkten Teilmenge von \mathbb{R} .

Ist $M \subseteq \mathbb{R}$ und $m \leq r$ für alle $m \in M$, so heißt r eine **obere Schranke** von M . Gibt es unter allen oberen Schranken eine kleinste s (die also $s \leq r$ für alle oberen Schranken r erfüllt), so nennt man s das **Supremum** von M , symbolisch $s := \sup M$. Ist $s \in M$, nennt man s auch das **Maximum** von M und schreibt $s = \max M$. Eine Menge M , zu der es wenigstens eine obere Schranke gibt, heißt **nach oben beschränkt**. Die **Supremumseigenschaft** besagt nun, dass jede nach oben beschränkte nichtleere Teilmenge $M \subseteq \mathbb{R}$ auch ein Supremum hat. Der Beweis ist leicht: Man teilt \mathbb{R} in zwei Mengen A und B auf, wobei B aus allen oberen Schranken von M besteht, A aus dem Rest. Man sieht schnell, dass die Definition der Vollständigkeit anwendbar ist, es also ein $x \in \mathbb{R}$ gibt mit $a \leq x \leq b$ für alle $a \in A$ und $b \in B$. Offenbar ist $x = \sup M$.

Ganz symmetrisch zu oberen Schranken definiert man **untere Schranken**, **Infimum** \inf , **Minimum** \min , **nach unten beschränkt**, und die **Infimumseigenschaft**, die aus Symmetriegründen ebenfalls aus der Vollständigkeit folgt. Eine Menge die sowohl nach oben als auch nach unten beschränkt ist, heißt schlicht **beschränkt**.

Übungsaufgabe 30. (E) Beweisen Sie die Infimumseigenschaft.

Übungsaufgabe 31. (T) Bestimmen Sie für folgenden Teilmengen von \mathbb{R} jeweils alle oberen Schranken und geben Sie, sofern vorhanden, das Supremum an.

1. $\{1 + \frac{1}{n} : n \in \mathbb{N}, n \geq 1\}$
2. $\{1 - \frac{1}{n} : n \in \mathbb{N}, n \geq 1\}$
3. $\{\frac{m}{n} : m, n \in \mathbb{N}, 0 < n < m\}$
4. $\{x \in \mathbb{Q} : x^2 \leq 3\}$

Übungsaufgabe 32. (T) Wie Aufgabe 31, aber mit unteren Schranken und Infimum.

1.3.5 Die archimedische Eigenschaft

Inhalt in Kurzfassung: Jede positive Quantität übersteigt, wenn sie hinreichend oft addiert wird, jede vorgegebene Schranke. Dieser Sachverhalt wird mit der sogenannten archimedischen Eigenschaft begrifflich gefasst.

Eine weitere wichtige Konsequenz der Vollständigkeit ist die **archimedische Eigenschaft**. Sie besagt Folgendes: Ist $\varepsilon > 0$ irgendeine positive reelle Zahl (z.B. $\varepsilon = 1$), so ist die Menge $M := \{n\varepsilon : n \in \mathbb{N}\}$ aller Vielfachen von ε nach oben unbeschränkt in \mathbb{R} , d.h. M hat keine obere Schranke r , mit anderen Worten: Es gibt keine Zahl $r \in \mathbb{R}$ mit $n\varepsilon < r$ für alle $n \in \mathbb{N}$. Man kann die archimedische

Eigenschaft auch so formulieren: Mit jedem noch so kleinen positiven ε lässt sich jede vorgegebene Größe r übertreffen, wenn man ε nur hinreichend oft addiert.

Die Beweis für die archimedische Eigenschaft mit Hilfe der Vollständigkeit lautet wie folgt: Wäre die archimedische Eigenschaft verletzt, so gäbe es eine obere Schranke r von $M = \{n\varepsilon : n \in \mathbb{N}\}$. Wegen der Supremumseigenschaft gäbe es also auch $s := \sup M$. Wir betrachten $s' := s - \varepsilon < s$. Weil s die kleinste obere Schranke von M ist, kann s' keine mehr sein, also $s' < n\varepsilon \in M$ für ein geeignetes $n \in \mathbb{N}$. Es gilt aber auch $(n+1)\varepsilon \in M$ und wegen $s' < n\varepsilon$ (Monotoniegesetz für $+$ anwenden) $s = s' + \varepsilon < n\varepsilon + \varepsilon = (n+1)\varepsilon \in M$. Also ist s doch keine obere Schranke von M , Widerspruch.

Übungsaufgabe 33. (P) Überlegen Sie sich, dass die Menge $\mathbb{Z} \subseteq \mathbb{R}$ in \mathbb{R} sowohl nach oben als auch nach unten unbeschränkt ist. Dabei sollen Sie nur bereits Bewiesenes verwenden.

1.3.6 Betrag, Potenzen und Wurzeln

Inhalt in Kurzfassung: Für reelle Zahlen werden in üblicher Weise Betrag, Potenzen und Wurzeln definiert.

Wegen der Trichotomie der Ordnungsrelation trifft für jede reelle Zahl r genau eine der folgenden drei Möglichkeiten zu:

1. $r > 0$, in diesem Fall heißt r **positiv**.
2. $r = 0$
3. $-r > 0$, in diesem Fall heißt r **negativ**.

Unter dem **Betrag** oder auch **Absolutbetrag** $|r|$ von r versteht man jenen der beiden Werte r und $-r$, der nichtnegativ ist. Sehr häufig verwendet man die **Dreiecksungleichung** $|x + y| \leq |x| + |y|$, hier zunächst nur für reelle Zahlen. Man sieht sie unmittelbar ein, wenn man die verschiedenen Möglichkeiten für positives/negatives x bzw. y durchspielt.

Übungsaufgabe 34. (P) Begründen Sie die Dreiecksungleichung $|x + y| \leq |x| + |y|$ für beliebige reelle Zahlen x, y durch geeignete Fallunterscheidung.

Wie wir aus Übungsaufgabe 77 wissen, ist das **Quadrat** $q = r^2$ einer reellen Zahl r nie negativ. Umgekehrt tritt, wie wir später aus dem Zwischenwertsatz (seinerseits eine Folgerung aus der Vollständigkeit von \mathbb{R}) werden schließen können, jede nichtnegative reelle Zahl als Quadrat auf, genauer: Zu jedem $q \geq 0$ gibt es ein r mit $r^2 = q$. Wegen $(-r)^2 = r^2 = q$ hat $-r$ dieselbe Eigenschaft. Den nichtnegativen der beiden Werte nennt man die **Quadratwurzel** von q , symbolisch $r = \sqrt{q}$. Auch all diese Eigenschaften lassen sich aus den Axiomen für \mathbb{R} ableiten. In Zukunft werden wir das bei elementaren Regeln nur mehr in Ausnahmefällen ausdrücklich betonen.

Übungsaufgabe 35. (T)

1. Berechnen Sie alle Lösungen der quadratischen Gleichung $x^2 - 4x + 3 = 0$:
2. Machen Sie die Probe für Ihre Lösungen aus 1. Dokumentieren Sie dabei alle Ihre Rechenschritte.
3. Geben Sie die Lösungsmenge der quadratischen Ungleichung $-x^2 + 4x - 3 < 0$ in Intervallschreibweise an.
4. Fertigen Sie eine saubere und aussagekräftige Skizze an, welche Ihre Resultate illustriert.

Übungsaufgabe 36. (T) Wie Aufgabe 35, aber mit $x^2 + 2x - 3 = 0$.

Übungsaufgabe 37. (T) Lösen Sie Gleichungen im Bereich der reellen Zahlen und fertigen Sie jeweils eine Skizze an.

1. $|x|^2 - 2x - 3 = 0$
2. $x^2 - 2|x| - 3 = 0$
3. $|x^2 - 2x - 3| = 0$

Übungsaufgabe 38. (T) Bestimmen Sie die Lösungsmenge über \mathbb{R} , d.h. die Menge aller reellen x , welche folgende Ungleichung erfüllen:

1. $\frac{3}{|x-2|} < 2 + 3x$
2. $\sqrt{|x+2|} \leq |x+1|$
3. $\frac{|1+x|}{1+|1+x|} \leq \frac{1}{2} + \frac{|x|}{1+|x|}$

Ganz ähnlich wie Quadrate definiert man höhere **Potenzen** $r^n := r \cdot r \cdot \dots \cdot r$ als Produkt von n Faktoren (n heißt in diesem Zusammenhang der **Exponent**), die alle den Wert r haben. Etwas eleganter kann man auch rekursiv definieren (siehe 1.3.8) $r^0 := 1$ und $r^{n+1} := r^n r$. Entsprechend heißt $r \geq 0$ die **n -te Wurzel** von q (symbolisch $r = \sqrt[n]{q}$), falls $r^n = q$, wobei im Fall $q > 0$ auch $r > 0$ sein soll. Man beachte den Unterschied zwischen geradem bzw. ungeradem Exponenten n : Für ungerades n ist r schon durch die Forderung $r^n = q$ eindeutig bestimmt, und zwar für jedes $q \in \mathbb{R}$. Für gerades n und $q > 0$ gibt es zwei Lösungen für r , die sich um das Vorzeichen unterscheiden, weshalb man erst durch die Forderung $r > 0$ Eindeutigkeit erzwingt. Für gerades n und $q < 0$ gibt es gar keine Lösung, weshalb die Wurzel nur für $q \geq 0$ definiert ist.

Exemplarisch für zahlreiche bekannte Rechenregeln für Potenzen und Wurzeln überlege man sich die folgenden.

Übungsaufgabe 39. (P) Begründen Sie die folgenden Rechenregeln für Potenzen und Wurzeln reeller Zahlen ($0 \leq a, b \in \mathbb{R}$, $m, n \in \mathbb{N}$):

1. $a^{m+n} = a^m a^n$
2. $(ab)^n = a^n b^n$

$$3. (a^m)^n = a^{mn}$$

$$4. \sqrt[n]{a} \sqrt[n]{b} = \sqrt[n]{ab}$$

$$5. \sqrt[m]{\sqrt[n]{a}} = \sqrt[mn]{a}$$

Nützlich sind überdies die Symbole ∞ für **unendlich** und $-\infty$, die allerdings keine reellen Zahlen darstellen. Man vereinbart $-\infty < r < \infty$, $\infty + r = \infty$ für alle $r \in \mathbb{R}$ und ähnliche selbsterklärende Konventionen.

1.3.7 Das Induktionsprinzip

Inhalt in Kurzfassung: Die natürlichen Zahlen \mathbb{N} haben im Gegensatz zu den anderen prominenten Zahlenbereichen die Eigenschaft, dass man, indem man von 0 ausgeht und immer wieder von der jeweils aktuellen Zahl n zum Nachfolger $n + 1$ weitergeht, jede vorgegebene natürliche Zahl erreichen kann. Dieses Prinzip („Dominoeffekt“ heißt Induktionsprinzip. Die meisten Sätze, die sich auf beliebige natürliche Zahlen beziehen, können mit dem Induktionsprinzip bewiesen werden.

Innerhalb \mathbb{R} haben wir die Menge \mathbb{N} definiert, indem wir von 0 ausgingen und beliebige Additionen von 1 zuließen. Das bedeutet, dass jede Menge $T \subseteq \mathbb{N}$, die 0 enthält und mit jeder Zahl $n \in \mathbb{N}$ auch die Zahl $n + 1$ (den **Nachfolger** von n), bereits ganz \mathbb{N} sein muss. Dieses an sich sehr einfache Prinzip erweist sich als unglaublich wichtig in der Mathematik und heißt **Induktionsprinzip**. Als Formel lässt es sich schreiben als:

$$\forall T \subseteq \mathbb{N} : (0 \in T \wedge (\forall n : n \in T \rightarrow n + 1 \in T)) \rightarrow T = \mathbb{N}$$

Die Bedeutung des Induktionsprinzips liegt darin, dass mit seiner Hilfe viele Aussagen $A(n)$ über alle (also unendlich viele) $n \in \mathbb{N}$ in nur endlich vielen Schritten bewiesen werden können. Dazu betrachtet man die Menge T aller $n \in \mathbb{N}$, die die zu beweisende Aussage erfüllen. Zu beweisen ist demnach $T = \mathbb{N}$. Nach dem Induktionsprinzip genügt es dafür $0 \in T$ und für alle $n \in \mathbb{N}$ die Implikation $n \in T \rightarrow n + 1 \in T$ nachzuweisen. Dafür wiederum genügt der Beweis folgender beiden Aussagen:

1. $A(0)$, d.h. $0 \in T$: Die für alle n zu beweisende Aussage gilt also zunächst für $n = 0$ (**Induktionsanfang** oder auch **Induktionsstart**). Das zu zeigen, ist oft sehr leicht.
2. $A(n) \rightarrow A(n + 1)$ für beliebiges $n \in \mathbb{N}$, d.h.: Von einem beliebigen n vererbt sich die Aussage $A(n)$ (äquivalent $n \in T$), die sogenannte **Induktionsannahme** oder **Induktionsvoraussetzung**, auf $n + 1$, d.h. auf die sogenannte **Induktionsbehauptung** $A(n + 1)$ (äquivalent $n + 1 \in T$). Die Vererbung selbst, d.h. die Implikation $n \in T \rightarrow n + 1 \in T$ nennt man den **Induktionsschritt**. Weil dabei $n \in \mathbb{N}$ beliebig anzunehmen ist, wird dabei insgesamt die im Induktionsprinzip auftretende Teilformel $\forall n : n \in T \rightarrow n + 1 \in T$ bewiesen.

Gemäß dem Induktionsprinzip ist damit $T = \mathbb{N}$ bewiesen, das heißt die Menge aller n , für die $A(n)$ gilt, stimmt mit \mathbb{N} überein, mit anderen Worten: Die zu beweisende Aussage $A(n)$ gilt für alle $n \in \mathbb{N}$. In der Praxis spricht man nicht immer von der Menge T , sondern nur von den n .

Insgesamt hat man also erstens, als Induktionsanfang, $A(0)$ zu überprüfen und zweitens, als Induktionsschritt, dass für beliebiges $n \in \mathbb{N}$ aus der Induktionsannahme $A(n)$ die Induktionsbehauptung $A(n+1)$ folgt.

Eine einfache Variante besteht darin, dass man beim Induktionsanfang statt $0 \in T$, z.B. $1 \in T$ zeigt. Dann folgt nur, dass alle natürlichen Zahlen n ab $n = 1$ die Aussage erfüllen. Entsprechendes gilt, wenn man $n_0 \in T$ für einen anderen Startwert $n_0 \in \mathbb{N}$ nachweist.

Ein interessantes Beispiel eines Induktionsbeweises liefert den wichtigen **Satz vom kleinsten Element**:

Satz 1.3.7.1. *Jede nichtleere Menge M natürlicher Zahlen hat ein kleinstes Element.*

Beweis. Wir gehen indirekt vor, indem wir annehmen $M \subseteq \mathbb{N}$ habe kein kleinstes Element und daraus $M = \emptyset$ folgern. Wir betrachten dazu die Menge $T := \mathbb{N} \setminus M$ aller $n \in \mathbb{N}$ mit der Eigenschaft $\{0, \dots, n\} \cap M = \emptyset$, für die also keine Zahl bis n in M liegt. Mit Hilfe des Induktionsprinzips wollen wir $T = \mathbb{N}$ zeigen, woraus $M = \emptyset$ folgt, wie gewünscht. Nach dem Induktionsprinzip sind also zwei Aussagen zu beweisen, erstens $0 \in T$ und zweitens die Implikation $n \in T \rightarrow n+1 \in T$ für beliebiges $n \in \mathbb{N}$:

$0 \in T$: Wäre $0 \notin T$, so folgte $\{0\} \cap M \neq \emptyset$, also $0 \in M$. Dann hätte M doch ein kleinstes Element, nämlich 0, Widerspruch.

Aus $n \in T$ folgt $n+1 \in T$: Andernfalls wäre $\{0, \dots, n\} \cap M = \emptyset$ und $\{0, \dots, n, n+1\} \cap M \neq \emptyset$, was nur für $n+1 \in M$ möglich ist. Dann wäre aber $n+1$ das kleinste Element von M , Widerspruch zur indirekten Annahme. \square

Satz 1.3.7.1 rechtfertigt eine modifizierte, etwas stärkere Variante des Induktionsprinzips, die wir **Ordnungsinduktion** nennen. Zur Unterscheidung nennen wir die bisher behandelte und auch im Beweis von Satz ?? verwendete Variante **Nachfolgerinduktion**, weil im Induktionsschritt der Übergang von n auf den Nachfolger $n+1$ der entscheidende Schritt war. Bei der Ordnungsinduktion verwendet man eine stärkere Induktionsannahme, nämlich dass nicht nur $n \in T$, sondern $k \in T$ für alle $k = 0, 1, \dots, n$ gilt. Wegen Satz 1.3.7.1 folgt tatsächlich auch so die zu beweisende Aussage $T = \mathbb{N}$. Wäre nämlich $T \neq \mathbb{N}$, so wäre $M := \mathbb{N} \setminus T \neq \emptyset$, und der Satz würde ein kleinstes Element m von M liefern. Für $n := m-1$ wäre dann $k \in T$ für alle $k = 0, 1, \dots, n$, trotzdem aber $n+1 \notin T$, entgegen der Voraussetzung bei Ordnungsinduktion.

Wir werden sowohl solche Anwendungsbeispiele kennen lernen, wo die Nachfolgerinduktion reicht, als auch solche, wo die Ordnungsinduktion (oder, äquivalent, der Satz vom kleinsten Element) wesentlich praktischer ist. Für beide Versionen ist auch die Bezeichnung **vollständige Induktion** gebräuchlich.

Übungsaufgabe 40. (P) Man kann das Induktionsprinzip auch als **Dominoeffekt** verstehen: Fällt der erste Stein und ist sichergestellt, dass jeder fallende Stein seinen Nachfolger mitreißt, so fällt jeder Stein. Begründen Sie das mit Hilfe des Induktionsprinzips. Was ist dabei die Menge T ? Und welche Annahme über die Aufstellung der Dominosteine muss man dabei machen?

Und nun eine erste Übung zur Anwendung des Induktionsprinzips:

Übungsaufgabe 41. (T) Stellen Sie sich eine quadratische Fläche mit einem Schachbrettmuster mit 2^n dm Seitenlänge vor. Jedes Feld des Musters ist ein Quadrat von 1 dm Seitenlänge. Die Gesamtfläche soll mit L-förmigen Platten gepflastert werden, von denen jede genau drei Felder des Schachbrettmusters überdeckt. Die langen Seiten einer jeden solchen Platte sind also jeweils 2 dm lang, die kurzen 1 dm. Innerhalb der Fläche steht auf einem der quadratischen 1×1 -Felder ein Pfeiler (dieser Bereich muss also nicht gepflastert werden). Zeigen Sie mit vollständiger Induktion, dass für alle $n \in \mathbb{N}$ mit $n \geq 1$ so eine Pflasterung möglich ist.

Hinweis: Zerlegen Sie beim Induktionsschritt das Quadrat in vier kleinere Quadrate und verwenden Sie eine L-förmige Platte in der Mitte, um auf jene drei der Quadrate, in denen nicht der Pfeiler steht, die Induktionsannahme anwenden zu können.

1.3.8 Anwendung des Induktionsprinzips: Rekursionen und Induktionsbeweise

Inhalt in Kurzfassung: Das Induktionsprinzip ermöglicht die rekursive Definition von Folgen (z.B. durch Angabe eines Anfangsgliedes und einer Rekursionsgleichung) und Beweise interessanter Eigenschaften solcher Folgen, insbesondere von Formeln, die für alle $n \in \mathbb{N}$ gelten.

Eine wichtige Konsequenz des Induktionsprinzips ist die Möglichkeit **rekursiver Definitionen**. Ein bereits bekanntes Beispiel: Wir können die Potenzen r^n für $r \in \mathbb{R}$ und $n \in \mathbb{N}$ durch $r^0 := 1$ und $r^{n+1} := r^n r$ (für alle $n \in \mathbb{N}$) definieren. Bezeichne dazu T die Menge aller $n \in \mathbb{N}$, bis zu denen die Potenzen r^n dadurch eindeutig definiert sind. Dann ist $0 \in T$ wegen $r^0 = 1$ (der Induktionsanfang ist also erfüllt), und r^{n+1} errechnet sich über $r^{n+1} = r^n r$ eindeutig aus r^n und r selbst. Somit gilt $n \in T \rightarrow n+1 \in T$ für alle n , also funktioniert auch der Induktionsschritt.

Ganz allgemein lässt sich offenbar sagen: Legt man für eine Funktion $f : \mathbb{N} \rightarrow X$ (hier ist X irgendeine Menge, in der die zu definierenden Werte von f liegen sollen) den Wert $f(0) \in X$ fest und erklärt man, wie sich für beliebiges $n \in \mathbb{N}$ aus den Werten $f(0), \dots, f(n) \in X$ der Wert $f(n+1) \in X$ ergibt, so ist dadurch f eindeutig festgelegt. Für dieses Prinzip ist auch die Bezeichnung **Rekursionssatz** gebräuchlich.

Eine Funktion mit Definitionsbereich \mathbb{N} nennt man auch eine (unendliche) **Folge**. Die Funktionswerte $f(n)$, $n \in \mathbb{N}$, einer Folge heißen die **Glieder** der

Folge. (Folgen mit reellen Gliedern werden in Abschnitt 2.1 im Zentrum stehen.) Die gesamte Folge notiert man auch als $f = (f(n))_{n \in \mathbb{N}}$, wobei man oft die Indexschreibweise $f_n := f(n)$, also $f = (f_n)_{n \in \mathbb{N}}$ bevorzugt. Steht die Funktion weniger im Vordergrund als die Folgenglieder, wählt man meist andere Buchstaben für die Folgenglieder, z.B. a_n für f_n . Oft beginnt der Definitionsbereich von Folgen nicht mit 0, sondern erst mit 1, hin und wieder auch noch später. In der Regel werden wir für Folgenglieder nicht den Buchstaben f , sondern vorzugsweise a, b, c, x, y, \dots verwenden.

Eine Variante sind **mehrgliedrige Rekursionen**, wie etwa die zweigliedrig definierte **Fibonacci-Folge**. Man beginnt mit den beiden Anfangsgliedern $a_0 := 0$ und $a_1 := 1$ (manchmal auch $a_0 := 1$ und $a_1 := 1$) und $a_{n+1} := a_{n-1} + a_n$ für $n \geq 1$. Die ersten Glieder lauten dann 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \dots , siehe auch Übungsaufgabe 142.

Geht man von einer Folge $(a_n)_{n \in \mathbb{N}}$ mit $a_n \in \mathbb{R}$ oder auch $a_n \in \mathbb{C}$ (d.h. $X = \mathbb{R}$ oder $X = \mathbb{C}$) aus, kann man die Folge der sogenannten **Partialsommen** $s_n := a_0 + a_1 + \dots + a_n$ rekursiv definieren durch $s_0 := a_0$ und $s_{n+1} := s_n + a_{n+1}$. Für s_n schreibt man auch

$$s_n = \sum_{k=0}^n a_k.$$

Fangen die a_n erst mit $n = 1$ oder später an, so setzt man $a_0 = 0$ und schreibt entsprechend $s_n = \sum_{k=1}^n a_k$ etc. Die Folge der Partialsommen nennt man auch die aus den Gliedern a_n gebildete **Reihe**. Das ganze Kapitel 2 wird Folgen und Reihen gewidmet sein. An dieser Stelle untersuchen wir die Summen, die sich aus arithmetischen und geometrischen Folgen ergeben:

Bei der **arithmetischen Reihe** ist $a_n = n$. Mittels Nachfolgerinduktion lässt sich die Formel

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

beweisen. In der Notation von früher ist also

$$T = \left\{ n \in \mathbb{N} : \sum_{k=1}^n k = \frac{n(n+1)}{2} \right\}$$

Und zwar stellt man fest, dass zunächst für $n = 0$ beide Seiten $s_0 = \sum_{k=1}^0 k = 0$ und $\frac{0(0+1)}{2} = 0$ den Wert 0 haben. Also ist $0 \in T$ (Induktionsanfang). Nimmt man ihre Richtigkeit der Formel (also $n \in T$) für ein beliebiges, festes n an (Induktionsannahme, IA), so folgt (Induktionsschritt) auch

$$\begin{aligned} s_{n+1} &= s_n + (n+1) \stackrel{\text{IA}}{=} \frac{n(n+1)}{2} + (n+1) = \\ &= (n+1) \left(\frac{n}{2} + 1 \right) = (n+1) \left(\frac{n+2}{2} \right) = \frac{(n+1)(n+2)}{2}, \end{aligned}$$

die Formel für $n+1$ statt n (also $n+1 \in T$, Induktionsbehauptung). Also gilt für beliebiges $n \in \mathbb{N}$ die Implikation $n \in T \rightarrow n+1 \in T$. Wegen der Beliebigkeit

von $n \in \mathbb{N}$ ist somit neben $0 \in T$ auch die zweite der beiden Prämissen im Induktionsprinzip erfüllt: $\forall n : n \in T \rightarrow n + 1 \in T$. Aus dem Induktionsprinzip folgt $T = \mathbb{N}$, also gilt die zu beweisende Formel für alle $n \in \mathbb{N}$. In Zukunft werden wir Induktionsbeweise nicht mehr so ausführlich darstellen. Insbesondere werden wir die Menge T aller $n \in \mathbb{N}$, für die die zu beweisende Aussage gilt, nur mehr in Ausnahmefällen explizit anführen.

Übungsaufgabe 42. (T) Schreiben Sie die folgenden Terme mithilfe des Summensymbols \sum an und erklären Sie Ihre Vorgehensweise:

1. $1 + 3 + 5 + \dots + 99$
2. $4 + 8 + 12 + \dots + 100$
3. $1 + 4 + 9 + 16 + 25 + \dots + 10^4$
4. $1 - 2 + 3 - 4 + 5 - 6 + \dots + 101$

Übungsaufgabe 43. (T)

1. Schreiben Sie folgende Behauptung als Formel (möglichst mit Summensymbol) an: „Die Summe aller geraden Zahlen von 2 bis $2n$ kann berechnet werden, indem man die Anzahl der geraden Zahlen von 2 bis $2n$ mit $n + 1$ multipliziert.“
2. Führen Sie den Induktionsstart des Induktionsbeweises für die Aussage aus 1. durch.
3. Führen Sie den Induktionsschritt des Induktionsbeweises durch.
4. Markieren Sie jene Stelle Ihres Induktionsbeweises, an der Sie die Induktionsvoraussetzung verwendet haben.

Übungsaufgabe 44. (T) Wie Übungsaufgabe 43, aber mit: „Die Summe aller ungeraden Zahlen von 1 bis $2n - 1$ kann auch berechnet werden, indem man die Anzahl der ungeraden Zahlen von 1 bis $2n - 1$ mit n multipliziert.“

Bei der **geometrischen Reihe** ist $a_n = q^n$ für $n = 0, 1, 2, \dots$ und ein festes $q \in \mathbb{R}$ (oder auch $q \in \mathbb{C}$). Dann ist

$$qs_n = q(q^0 + q^1 + \dots + q^{n-1} + q^n) = q^1 + q^2 + \dots + q^n + q^{n+1} = s_n + q^{n+1} - 1,$$

also $(q - 1)s_n = q^{n+1} - 1$. Für $q \neq 1$ liefert das die Formel

$$s_n = \sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q}.$$

Will man die informelle Schreibweise mit den drei Punkten \dots (für „und so weiter“) vermeiden, geht man mit Induktion vor:

Übungsaufgabe 45. (T) Beweisen Sie die Summenformel für die geometrische Reihe mit Induktion. Kennzeichnen Sie dabei sorgfältig Induktionsanfang, -annahme, -schritt und -behauptung.

Übungsaufgabe 46. (T) Berechnen Sie für allgemeines $n \in \mathbb{N}$ folgende Summen:

$$\begin{array}{llll} 1. \sum_{k=0}^n 2^k & 2. \sum_{k=1}^n (-3)^k & 3. \sum_{k=2}^n (-1)^k & 4. \sum_{k=n}^{2n} 2^{-k} \\ 5. \sum_{k=n}^{2n} (-2)^k & & & \end{array}$$

Übungsaufgabe 47. (T) Beweisen Sie mit vollständiger Induktion für allgemeines $n \in \mathbb{N}$

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

Übungsaufgabe 48. (T) Beweisen Sie mit vollständiger Induktion für allgemeines $n \in \mathbb{N}$

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

Hinweis: Beachten Sie, dass auf der linken Seite beide Summationsgrenzen von n abhängen.

Übungsaufgabe 49. (T) Zeigen Sie für $n = 2, 3, \dots$ mittels vollständiger Induktion:

$$\sum_{k=0}^{n-1} (n+k)(n-k) = \frac{n(n+1)(4n-1)}{6}.$$

Analog zu Summen mehrerer Summanden lassen sich auch Produkte $p_n = \prod_{k=1}^n a_n$ mehrerer Faktoren rekursiv definieren durch $p_0 := 1$ und $p_{n+1} := p_n a_{n+1}$.

Übungsaufgabe 50. (T) Zeigen Sie mit vollständiger Induktion, dass für alle $n = 2, 3, \dots$ gilt

$$\prod_{k=2}^n \left(1 - \frac{1}{k}\right) = \frac{1}{n}.$$

Übungsaufgabe 51. (T) Beweisen Sie mittels vollständiger Induktion, dass für jedes $n \in \mathbb{N}$ die natürliche Zahl $n^3 + 2n$ durch 3 teilbar ist.

Hinweis: Schreiben Sie zuerst die zu beweisende Aussage als Formel an.

Übungsaufgabe 52. (T) Zeigen Sie mit vollständiger Induktion, dass für natürliche Zahlen $n \in \mathbb{N}$ stets $n < 2^n$ gilt.

Übungsaufgabe 53. (T) Zeigen Sie mit vollständiger Induktion, dass für hinreichend große natürliche Zahlen $n \in \mathbb{N}$ (wie groß?) stets $n^2 < 2^n$ gilt.

Übungsaufgabe 54. (T) Zeigen Sie mit vollständiger Induktion: Für beliebige Konstanten $a, b > 0$ und $n \in \mathbb{N}$ groß genug gilt $n^2 \geq an + b$. Beschreiben Sie graphisch und verbal, wie sich diese Ungleichung interpretieren läßt!

Hinweis: $n \geq a + b + 1$ ist eine gute Wahl.

Übungsaufgabe 55. (T) Die Anzahl der Teile, in die man eine Kreisfläche durch n Geraden (gerade Schnitte) zerlegen kann, kann die Zahl $\frac{n^2+n+2}{2}$ nie übersteigen.

Hinweis: Überlegen Sie zuerst, wie viele alte Schnitte ein neuer Schnitt höchstens kreuzen kann. Wie viele neue Teile entstehen daher maximal?

1.3.9 Teilbarkeit und Primfaktorzerlegung in \mathbb{N} und \mathbb{Z}

Inhalt in Kurzfassung: Jede positive natürliche Zahl, lässt sich in ein Produkt von Primzahlen zerlegen. Das ist nicht schwer einzusehen. Aber erst dass diese Zerlegung sogar (bis auf die Reihenfolge der Faktoren) eindeutig ist, gibt den Primzahlen ihre große Bedeutung. Zum Beispiel lassen sich anhand der Primfaktorzerlegung größter gemeinsamer Teiler und kleinstes gemeinsames Vielfaches besser verstehen.

Für zwei ganze Zahlen a, b schreibt man $a|b$ (a teilt b , a ist ein **Teiler** von b , b ist ein **Vielfaches** von a), wenn es eine ganze Zahl c gibt mit $b = ac$. Eine positive ganze Zahl $p \geq 2$ heißt **Primzahl**, wenn 1 und p die einzigen positiven Teiler von p sind (1 ist keine Primzahl!). Die Menge der Primzahlen bezeichnen wir mit \mathbb{P} . Es gilt also $\mathbb{P} = \{2, 3, 5, 7, 11, 13, 17, 19, \dots\}$.

Mit Hilfe des Satzes vom kleinsten Element beweist man, dass sich jede natürliche Zahl $n \geq 1$ als ein Produkt von Primzahlen, ihren **Primfaktoren** darstellen lässt. Diese Darstellung heißt **Primfaktorzerlegung**. Dabei interpretiert man 1 als das leere Produkt (siehe die Definition von Produkten mit 0 Faktoren).

Zum Beweis: Tatsächlich gilt die Aussage für $n = 1$ (Produkt von 0 Faktoren). Wäre sie für gewisse n verletzt, so gäbe es nach dem Satz vom kleinsten Element 1.3.7.1 auch ein kleinstes $n > 1$ ohne Primfaktorzerlegung. Wäre n eine Primzahl, so Produkt von nur einem Primfaktor, hätte also sehr wohl eine Primfaktorzerlegung, Widerspruch. Wenn n aber keine Primzahl ist, so wäre $n = ab$ mit $1 < a, b < n$. Da n minimal ohne Primfaktorzerlegung gewählt war, haben a und b Primfaktorzerlegungen, die sich aber zu einer ihres Produktes $n = ab$ zusammensetzen lassen, ebenfalls Widerspruch. Es kann also kein n ohne Primfaktorzerlegung geben.

Sucht man nach den ersten Primzahlen, so identifiziert man unterhalb von $n = 50$ sehr schnell die Primzahlen 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43 und 47. Diese Reihe reißt nie ab. Es gibt nämlich unendlich viele Primzahlen:

Übungsaufgabe 56. (E) Beweisen Sie dies, indem Sie sich überlegen, dass zu jeder endlichen Menge von Primzahlen p_1, \dots, p_k die Zahl $n := p_1 \cdot \dots \cdot p_k + 1$ durch keines der p_i teilbar ist, daher ...

Einer der wichtigsten Sätze der Mathematik ist der von der **Eindeutigkeit der Primfaktorzerlegung**:

Satz 1.3.9.1. *Je zwei Primfaktorzerlegungen einer natürlichen Zahl $n \geq 1$ unterscheiden sich nur durch die Reihenfolge der Primfaktoren, nicht aber hinsichtlich der Häufigkeit, mit der jeder Faktor auftritt.*¹⁰

Ein einfaches Beispiel: Wir können die Zahl $n = 12$ zwar auf drei Arten in Primfaktoren zerlegen, nämlich $12 = 2 \cdot 2 \cdot 3 = 2 \cdot 3 \cdot 2 = 3 \cdot 2 \cdot 2$. Doch kommt in jeder Zerlegung der Primfaktor 2 doppelt vor, 3 einmal und alle anderen Primzahlen gar nicht. Nach dem Satz gilt das nicht nur für die Zahl 12, sondern für alle natürlichen Zahlen. Der Beweis braucht weit weniger als eine Seite und gelingt z.B. wie folgt:

Beweis der Eindeutigkeit der Primfaktorzerlegung: Wir gehen indirekt vor. Wäre die Eindeutigkeitsaussage falsch, so gäbe es nach dem Satz vom kleinsten Element (1.3.7.1) ein minimales $n > 0$ mit zwei (nicht nur bis auf die Reihenfolge der Faktoren) verschiedenen Primfaktorzerlegungen, also

$$n = p_1 \cdot \dots \cdot p_r = q_1 \cdot \dots \cdot q_s.$$

Hilfsbehauptung: Alle p_i sind von allen q_j verschieden. Denn wäre $p_i = q_j$, so könnte man die Gleichung durch diesen Wert kürzen und hätte zwei verschiedene Zerlegungen für die natürliche Zahl $n_0 := \frac{n}{p_i} < n$, im Widerspruch zur Minimalität von n als Gegenbeispiel. Also gilt tatsächlich diese Hilfsbehauptung. Insbesondere ist also $p_1 \neq q_1$. Sei o.B.d.A.¹¹ $p_1 < q_1$ (im Fall $q_1 > p_1$ symmetrisch argumentieren), dann hätte die Zahl

$$\begin{aligned} n' &:= (q_1 - p_1) \cdot q_2 \cdot q_3 \cdot \dots \cdot q_s = n - p_1 \cdot (q_2 \cdot q_3 \cdot \dots \cdot q_s) = \\ &= p_1 \cdot (p_2 \cdot \dots \cdot p_r - q_2 \cdot \dots \cdot q_s) = p_1 \cdot m \end{aligned}$$

(mit $m := p_2 \cdot \dots \cdot p_r - q_2 \cdot \dots \cdot q_s$) zwei verschiedene Zerlegungen: Die Zerlegung $n' = (q_1 - p_1) \cdot q_2 \cdot q_3 \cdot \dots \cdot q_s$ enthält den Primfaktor p_1 nicht. Denn dieser kommt weder unter den q_j vor, noch kann p_1 den ersten Faktor $q_1 - p_1$ teilen, weil es dann auch die größere Primzahl q_1 teilen müsste, entgegen der Hilfsbehauptung. Die andere Zerlegung ergibt sich aus $n' = p_1 m$, indem man eine Primfaktorzerlegung von m einsetzt, und enthält offenbar den Faktor p_1 . Wegen $n' < n$ stünde auch das im Widerspruch zur Minimalität von n . \square

Aus der Eindeutigkeit der Primfaktorzerlegung von natürlichen Zahlen lassen sich mehrere interessante Aussagen über Teilbarkeit, größten gemeinsamen Teiler und kleinstes gemeinsames Vielfaches ablesen bzw. gewinnen:

¹⁰Man beachte, dass analoge Aussagen gelten, wenn man nicht nur in \mathbb{N} arbeitet, sondern auch negative ganze Zahlen zulässt. Notwendige Modifikationen im Folgenden bestehen z.B. darin, dass bei Teilbarkeitsaussagen stets sowohl der positive wie auch der negative Wert zugelassen werden müssen. Im Zusammenhang mit Polynomen wird es wichtig sein, diese Art von Verallgemeinerung im Auge zu haben.

¹¹Diese Abkürzung steht für „ohne Beschränkung der Allgemeinheit“. Man verwendet sie, um anzudeuten, dass zwar andere Fälle möglich sind, diese aber genauso behandelt werden können wie der explizit ausgeführte Fall.

1. Jede natürliche Zahl $n \geq 1$ bestimmt eindeutig die Vielfachheiten $e_p(n) \in \mathbb{N}$, $p \in \mathbb{P}$, mit der die Primzahlen p als Faktoren in der Zerlegung

$$n = \prod_{p \in \mathbb{P}} p^{e_p(n)}$$

von n vorkommen. Dabei gilt $e_p(n) > 0$ nur für endlich viele $p \in \mathbb{P}$. Umgekehrt definiert jede Familie von Vielfachheiten $e_p(n)$, $p \in \mathbb{P}$, von denen nur endlich viele von 0 verschieden sind, eine eindeutige natürliche Zahl n , für die obige Formel gilt.

2. Die Vielfachheiten eines Produktes ergeben sich als Summe der einzelnen Vielfachheiten:

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

für alle $m, n \in \mathbb{N}$ mit $m, n \geq 1$ und alle $p \in \mathbb{P}$. Entsprechendes gilt auch für Produkte von mehr als zwei Faktoren.

3. Für zwei natürliche Zahlen $m, n \geq 1$ gilt Teilbarkeit $m|n$ genau dann, wenn für alle $p \in \mathbb{P}$ die Ungleichung $e_p(m) \leq e_p(n)$ gilt.
4. Eine Primzahl p , die das Produkt ab zweier natürlicher Zahlen a und b teilt, muss mindestens einen der beiden Faktoren a oder b teilen. (Nur die Primzahlen sowie, aus trivialen Gründen, die Zahlen 0 und 1 haben diese Eigenschaft.)

5. Zu je zwei natürlichen Zahlen

$$m = \prod_{p \in \mathbb{P}} p^{e_p(m)} \quad \text{und} \quad n = \prod_{p \in \mathbb{P}} p^{e_p(n)}$$

mit $m, n \geq 1$ betrachten wir die Zahl

$$t := \prod_{p \in \mathbb{P}} p^{\min\{e_p(m), e_p(n)\}}.$$

Wegen $e_p(v) = \min\{e_p(m), e_p(n)\} \leq e_p(m), e_p(n)$ für alle $p \in \mathbb{P}$ ist diese Zahl t ein gemeinsamer Teiler sowohl von m als auch von n , also $t|m$ und $t|n$. Darüber hinaus hat t unter allen gemeinsamen Teilern von m und n die größt möglichen Vielfachheiten $e_p(t)$. Das bedeutet, dass für jeden anderen gemeinsamen Teiler $t' \in \mathbb{N}$ von m und n stets $t'|t$ gilt. Deshalb heißt t auch der **größte gemeinsame Teiler** von m und n , symbolisch $t = \text{ggT}(m, n)$. Ganz analog gibt es einen größten gemeinsamen Teiler $\text{ggT}(n_1, \dots, n_k)$ auch von mehr als nur zwei natürlichen Zahlen n_1, \dots, n_k .

6. Eine analoge Konstruktion, lediglich mit \max statt \min liefert das kleinste gemeinsame Vielfache:

Zu je zwei natürlichen Zahlen

$$m = \prod_{p \in \mathbb{P}} p^{e_p(m)} \quad \text{und} \quad n = \prod_{p \in \mathbb{P}} p^{e_p(n)}$$

mit $m, n \geq 1$ betrachten wir die Zahl

$$v := \prod_{p \in \mathbb{P}} p^{\max\{e_p(m), e_p(n)\}}.$$

Wegen $e_p(t) = \max\{e_p(m), e_p(n)\} \geq e_p(m), e_p(n)$ für alle $p \in \mathbb{P}$ ist diese Zahl v ein gemeinsames Vielfaches sowohl von m als auch von n , also $m|v$ und $n|v$. Darüber hinaus hat v unter allen gemeinsamen Vielfachen von m und n die kleinst möglichen Vielfachheiten $e_p(v)$. Das bedeutet, dass für jedes andere gemeinsame Vielfache $v' \in \mathbb{N}$ von m und n stets $v|v'$ gilt. Deshalb heißt v auch das **kleinste gemeinsame Vielfache** von m und n , symbolisch $v = \text{kgV}(m, n)$. Ganz analog gibt es ein kleinstes gemeinsames Vielfaches $\text{kgV}(n_1, \dots, n_k)$ auch von mehr als nur zwei natürlichen Zahlen n_1, \dots, n_k .

Übungsaufgabe 57. (E) Rekapitulieren Sie alle obigen Behauptungen und ergänzen Sie jene Erklärungen die nur knapp ausgeführt worden sind. Was ist zu bedenken, wenn man auch $n = 0$ zulässt? (Man beachte, dass 0 Vielfaches aller natürlichen Zahlen ist, Teiler aber nur von sich selbst.) Ist es möglich bzw. sinnvoll ggT und kgV auch von unendlich vielen natürlichen Zahlen zu bilden?

Übungsaufgabe 58. (T) Ein einfaches Getriebe besteht aus zwei Zahnrädern mit N_a bzw. N_b Zähnen. Berechnen Sie nach wie vielen Umdrehungen ein Zahn des ersten Rades wieder in dieselbe Stelle des zweiten Rades greift.

1. $N_a = 50, N_b = 28$
2. $N_a = 52, N_b = 28$
3. $N_a = 53, N_b = 11$

Übungsaufgabe 59. (T) Wie Aufgabe 58, nur seien nun die beiden Zahnräder über eine Kette mit $N_k = 110$ Gliedern verbunden. Berechnen Sie nach wie vielen Umdrehungen ein Kettenglied wieder in denselben Zahn des ersten Rades greift.

1.3.10 Zahlendarstellungen

Inhalt in Kurzfassung: Wie schon an früherer Stelle betont, sind mathematische Objekte zu unterscheiden von ihrer Darstellung. Dass bei Zahlen solche Verwechslungen immer wieder auftreten, liegt nicht zuletzt daran, dass es sehr praktische Darstellungen wie vor allem die dekadische zur Basis 10 gibt, die sich durch weitgehende Eindeutigkeit auszeichnet. Obwohl das eigentlich alles elementarer Schulstoff ist, lohnt es, manch Grundlegendes zu wiederholen und sorgfältig zu durchdenken.

Der üblichen dekadischen Zahlendarstellung liegt der Satz zugrunde, dass es bei vorgegebener **Basis** $b \in \mathbb{N}$ mit $b \geq 2$ zu jedem $n \in \mathbb{N}$ eine Summendarstellung

$$n = \sum_{i=0}^k a_i b^i$$

mit $a_i \in \{0, 1, \dots, b-1\}$ gibt. Wählt man k minimal (d.h. $a_k \neq 0$ sofern nicht $n = 0$), so sind auch die a_i eindeutig bestimmt und heißen die **Ziffern** von n bezüglich der Basis b . Ist $b = 10$, so spricht man von **dekadischer** oder **Dezimaldarstellung**, für $b = 2$ von **binärer**.

Dekadisches Beispiel: $2018 = 2 \cdot 10^3 + 0 \cdot 10^2 + 1 \cdot 10^1 + 8 \cdot 10^0$

Der praktische Wert solcher Darstellungen besteht vor allem darin, dass damit die wichtigsten Rechenoperationen effizient durchgeführt werden können. Zur Übung soll das am Beispiel der Addition durchdacht werden.

Übungsaufgabe 60. (E) Beschreiben Sie möglichst genau den Additionsalgorithmus für natürliche Zahlen, wie sie ihn in der Volksschule gelernt haben. Tun Sie es, weil es etwas weniger Aufwand erfordert, in Bezug auf die binäre (und nicht die dekadische) Darstellung. Die Aufgabe besteht also im Folgenden:

Seien

$$m = \sum_{i=0}^k a_i 2^i, \quad n = \sum_{i=0}^k b_i 2^i, \quad m+n = \sum_{i=0}^{k+1} c_i 2^i$$

mit $a_i, b_i, c_i \in \{0, 1\}$. Beschreiben Sie das Verfahren, mit Hilfe dessen man die c_i aus den a_i und b_i gewinnt.

Reelle Zahlen r lassen sich eindeutig darstellen als $r = n + x$ mit $n \in \mathbb{Z}$ und $0 \leq x < 1$. Man schreibt dafür auch $n = [r]$ und $x = \{r\}$ (nicht mit der einelementigen Menge mit Element r verwechseln!) und nennt n den **ganzzahligen** und x den **gebrochenen Anteil** von r .

Für $n \geq 0$ schreibt man mit Kommastrich $r = a_k a_{k-1} \dots a_1 a_0, x_1 x_2 x_3 \dots$. Dabei sind die a_i die Ziffern von n wie oben und die Ziffern $x_i \in \{0, 1, \dots, b-1\}$ so gewählt, dass $x = \sum_{i=1}^{\infty} x_i 10^{-i}$.

Die exakte Bedeutung so einer unendlichen Summe (Reihe) werden wir im nächsten Kapitel ausführlich besprechen. Dekadisches Beispiel:

$$3,14159 \dots = 3 \cdot 10^0 + 1 \cdot 10^{-1} + 4 \cdot 10^{-2} + 1 \cdot 10^{-3} + 5 \cdot 10^{-4} + 9 \cdot 10^{-5} + \dots$$

Negative Zahlen kennzeichnet man bekanntlich durch das vorangestellte Symbol $-$.

Zur Wiederholung von früher: Sehr wichtig ist es, die Darstellung einer Zahl nicht mit der Zahl selbst zu verwechseln. Schließlich können verschiedene Zeichenketten dasselbe Objekt bezeichnen, z.B.: 4, $2+2$, vier, four, die Anzahl der Himmelsrichtungen, ... Ein anderes Beispiel (über das mathematische Laien oft hitzig diskutieren) ist die Gleichung $1 = 1,0000 \dots = 0,9999 \dots$, deren Grund wir im Zusammenhang mit dem Grenzwert von Reihen besser verstehen werden.

Auf der Darstellung ganzer Zahlen aufbauend ist auch die Darstellung einer rationalen Zahl $q = \frac{a}{b}$ als Bruch ganzer Zahlen a, b von großer Bedeutung. Stets gibt es auch eine **gekürzte Darstellung**, in der a und b teilerfremd sind. Man erhält sie, indem man sowohl ursprünglichen Zähler a als auch Nenner durch ihren größten gemeinsamen Teiler dividiert. Fordert man zusätzlich $b > 0$, so wird die gekürzte Darstellung sogar eindeutig.

Übungsaufgabe 61. (T) Wandeln Sie von dekadischer in binäre Darstellung um bzw. vice versa.

1. Dekadisch: 9, 10, 42, 259
2. Binär: 10, 111, 1000, 11001100

Übungsaufgabe 62. (T) Schätzen Sie ab, wie viele Stellen die angegebenen Zahlen in Dezimaldarstellung haben:

$$1. \frac{10^{(3^2)} 10^{-4} 10^6}{100^3 (10^3)^{-2}}, \quad 2. \frac{2^{(10^3)} 5^3 10^5}{(100^{-3})^{-2} (30^2)^2}, \quad 3. \frac{2^{(2^4)} 5^3 10^{-5}}{(10000^5)^{-2} 2^8}.$$

Hinweis: $2^{10} = 1024 \approx 10^3$

Übungsaufgabe 63. (T) Wie Übungsaufgabe 62, aber mit den Stellen in der Binärdarstellung.

Zum Abschluss die bereits am Ende von Abschnitt 1.2 angekündigte Übungsaufgabe zur Abzählbarkeit von \mathbb{Z} und \mathbb{Q} sowie zur Überabzählbarkeit von \mathbb{R} .

Übungsaufgabe 64. (E) In dieser Aufgabe greifen wir nochmals voraus auf die aus der Schule wohlbekannten Zahlenmengen \mathbb{N} , \mathbb{Z} , \mathbb{Q} und \mathbb{R} (siehe auch 1.3.1). Zeigen Sie:

1. Die Menge \mathbb{Z} aller ganzen Zahlen ist abzählbar. Hinweis: Beginnen Sie mit $a_0 = 0$, $a_1 = 1$, $a_2 = -1$ etc.
2. $\mathbb{N} \times \mathbb{N}$ ist abzählbar. Hinweis: Beginnen Sie bei der Abzählung mit den Paaren $(0, 0)$, $(0, 1)$, $(1, 0)$, $(0, 2)$, $(1, 1)$, $(2, 0)$, \dots . Beschreiben Sie sorgfältig wie Sie fortsetzen.
3. Für jede abzählbare Menge A ist auch $A \times A$ abzählbar. Hinweis: Verwenden Sie Aussage 2.
4. Die Menge \mathbb{Q} aller rationalen Zahlen ist abzählbar. Hinweis: Jede rationale Zahl lässt sich als Bruch (Paar) ganzer Zahlen deuten, und verwenden Sie vorangegangene Teile.
5. Die Menge M aller unendlichen 0-1-Folgen ist überabzählbar. Hinweis: Formal ist M die Menge aller Abbildungen $f : \mathbb{N} \rightarrow \{0, 1\}$, $n \mapsto a_n$, für die wir aber gewohnheitsmäßig lieber $f = \mathbf{a} = (a_0, a_1, \dots)$ schreiben, siehe auch Anfang von Kapitel 2. Nehmen Sie an, es gäbe eine Abzählung von M , d.h. eine Folge von Folgen $\mathbf{a}^{(k)} = (a_0^{(k)}, a_1^{(k)}, \dots)$, $k \in \mathbb{N}$, die ganz M durchläuft. Man erhält einen Widerspruch, indem man eine 0-1-Folge (b_0, b_1, \dots) konstruiert, die sicher nicht in der Abzählung vorkommt, weil sie sich im nullten Glied von der nullten Folge unterscheidet (d.h. $b_0 \neq a_0^{(0)}$), im ersten Glied von der ersten (d.h. $b_1 \neq a_1^{(1)}$) und so weiter.
6. Die Menge \mathbb{R} ist überabzählbar. Hinweis: Gäbe es eine Abzählung von \mathbb{R} , so auch eine der Teilmenge aller reellen Zahlen, in deren unendlicher Dezimalentwicklung nur die Ziffern 0 und 1 vorkommen. Das ist aber wegen Aussage 5 unmöglich.

7. Man beweise $|\mathbb{R}| = |\mathbb{C}|$. Hinweis: Die Schwierigkeit besteht darin, eine injektive Abbildung $f: \mathbb{C} \rightarrow \mathbb{R}$, also von der „zweidimensionalen“ Menge \mathbb{C} in die (intuitiv kleinere) „eindimensionale“ Menge zu konstruieren. Die Grundidee besteht darin, aus zwei 0-1-Folgen (die Real- und Imaginärteil a bzw. b einer komplexen Zahl $a + ib$ entsprechen), eine einzige zu machen, so dass dieser Vorgang auch umgekehrt werden kann. Dies ist möglich, indem man je zwei Folgen (a_1, a_2, \dots) und b_1, b_2, \dots die gemischte Folge $(a_1, b_1, a_2, b_2, \dots)$ zuordnet.

1.3.11 Die komplexe Zahlenebene

Inhalt in Kurzfassung: Komplexe Zahlen lassen sich als Punkte der sogenannten Gaußschen Zahlenebene auffassen. Die x -Koordinate heißt Real, die y -Koordinate heißt Imaginärteil. Geometrisch entsprechen der Addition und Subtraktion komplexer Zahlen die (in Abschnitt 1.5 zu behandelnden) entsprechenden Vektoroperationen. Für Multiplikation und Division ist die Polardarstellung komplexer Zahlen geeigneter. Ein Punkt der Ebene wird dabei durch zwei andere geometrische Größen codiert: durch seinen Abstand vom Koordinatenursprung (Betrag) und durch den Winkel, den seine Richtung mit der x -Achse einschließt (Argument). Damit wird auch das Potenzieren komplexer Zahlen durchsichtiger und man erkennt, dass jede komplexe Zahl $\neq 0$ genau n verschiedene n -te Wurzeln hat. Der für das Verständnis all dessen gleichfalls sehr nützlichen komplexen Konjugation entspricht geometrisch eine Spiegelung der Ebene um die x -Achse.

Auf die komplexen Zahlen stößt man notgedrungen, wenn man nach einer Erweiterung von \mathbb{R} sucht, in der erstens die vier Grundrechnungsarten (Addition, Subtraktion, Multiplikation und Division) nach den gewohnten Regeln ausführbar sind (genauer: die Erweiterung soll einen Körper bilden, siehe Axiomatik von \mathbb{R}) und zweitens auch die Wurzeln negativer Zahlen enthalten sind. Es genügt sogar, lediglich von einer sogenannten **imaginären Einheit** auszugehen, die man traditionell mit i bezeichnet und von der man als wesentliche Eigenschaft $i^2 = -1$ voraussetzt. Sind a und b irgendwelche reelle Zahlen, so sollen also auch das Produkt ib und die Summe $a + ib$ in der Erweiterung liegen. Wunderbarerweise reicht das, wie wir gleich sehen werden, schon aus. Man nennt a den **Real-** und b den **Imaginärteil** der komplexen Zahl $z = a + ib$, schreibt dafür auch $a = \Re z$ und $b = \Im z$, so dass stets $z = \Re z + i \Im z$ gilt, und spricht auch von der **Koordinatendarstellung** von z . Offenbar sind zwei komplexe Zahlen $a_1 + ib_1$ und $a_2 + ib_2$ genau dann gleich, wenn $a_1 = a_2$ und $b_1 = b_2$ gilt.¹² Die Zahl $\bar{z} := a - ib$ heißt die zu z (komplex) **Konjugierte**. Die Menge \mathbb{C} der **komplexen Zahlen** ist dann gegeben durch

$$\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\},$$

¹²Will man auf mengentheoretisch festem Grund stehen, könnte man $a + ib$ als alternative und suggestive Schreibweise für das Paar $(a, b) \in \mathbb{R}^2$ auffassen, was aber formale Komplikationen aufwirft, die wir hier vermeiden wollen.

enthält alle reellen Zahlen a in der Gestalt $a + i0$ (weshalb die Teilmengenbeziehung $\mathbb{R} \subseteq \mathbb{C}$ besteht) und auch die imaginäre Einheit $i = 0 + i1$. Die Addition muss nach der Regel

$$(a + ib) + (c + id) = (a + c) + i(b + d)$$

erfolgen, die Multiplikation gemäß

$$(a + ib)(c + id) = ac + aid + ibc + ibid = ac + iad + ibc + i^2bd = (ac - bd) + i(ad + bc).$$

Man rechnet für beide Operationen leicht die Assoziativ- und Kommutativgesetze nach, ebenso das Distributivgesetz.

Übungsaufgabe 65. (P) Rechnen Sie für komplexe Zahlen nach, indem Sie die entsprechenden Gesetze für reelle Zahlen verwenden:

1. das Kommutativgesetz für die Addition
2. das Assoziativgesetz für die Addition
3. das Kommutativgesetz für die Multiplikation
4. das Assoziativgesetz für die Multiplikation
5. das Distributivgesetz

Außerdem findet man auf den ersten Blick die neutralen Elemente $0 = 0 + i0$ (additiv) und $1 = 1 + i0$ (multiplikativ) sowie zu jedem $a + ib$ ein additiv Inverses, nämlich $-(a + ib) = -a + i(-b)$. Nicht auf den ersten Blick offensichtlich ist das multiplikativ Inverse z^{-1} einer komplexen Zahl $z = a + ib$. Man findet es aber schnell, wenn man unbekümmert drauf losrechnet und mit der Konjugierten erweitert:

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

Auch wenn diese Umformung mangels Rechtfertigung der Regeln für die Division (Brüche!) noch auf Sand gebaut war, so erweist sich doch das Ergebnis, nämlich die komplexe Zahl mit Realteil $\frac{a}{a^2 + b^2}$ und Imaginärteil $\frac{-b}{a^2 + b^2}$ tatsächlich als das multiplikativ Inverse zu $z = a + ib$ (Übung: nachrechnen). Unsere Menge \mathbb{C} mit den Operationen, wie sie oben definiert wurden, erfüllt alle Gesetze eines Körpers.

Übungsaufgabe 66. (E) Überprüfen Sie, dass \mathbb{C} mit den oben definierten Operationen tatsächlich einen Körper bildet. Gehen Sie dazu schrittweise vor:

1. Rekapitulieren Sie aus 1.3.3, welche Forderungen in der Definition eines Körpers enthalten sind.
2. Prüfen Sie, welche dieser Forderungen bereits nachgeprüft worden sind. Berufen Sie sich dabei auf das Skriptum.

3. Kontrollieren Sie jene Punkte, die noch offen sind.

Man beachte, dass wir bei unseren Definitionen gar keine andere Wahl hatten, wollten wir wieder einen Körper bekommen, der i enthält. In diesem Sinn sind also auch die komplexen Zahlen durch relativ bescheidene Minimalanforderungen eindeutig bestimmt und tragen keinerlei definitonische Willkür in sich.

Die Zahl $a^2 + b^2$ im Nenner des oben berechneten Inversen z^{-1} von z ist, abgesehen vom Fall $z = 0$, reell und positiv. Deshalb können wir die Wurzel ziehen und durch

$$|z| = |a + ib| := \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$$

den **Betrag** $|z|$ von z definieren. Das führt uns zur geometrischen Interpretation von \mathbb{C} .

Wir fassen \mathbb{C} als Ebene \mathbb{R}^2 auf, wo $z = a + ib$ dem Punkt mit den Koordinaten a und b entspricht (**Gauß'sche Zahlenebene**). Manchmal denken wir uns $a + ib$ auch als Vektor (Pfeil) vom Koordinatenursprung zum Punkt (a, b) . Dabei entspricht die x -Achse den reellen Zahlen und heißt deshalb auch die **reelle Achse**. Die y -Achse nennen wir die **imaginäre Achse**.

Die Addition komplexer Zahlen entspricht der üblichen Vektoraddition gemäß der Parallelogrammregel. Dabei stellt man sich Vektoren als Pfeile vor, die man aneinander hängt (Schaft an Spitze). Führt man diese Operation mit zwei Vektoren in unterschiedlicher Reihenfolge aus, entsteht ein Parallelogramm.

Übungsaufgabe 67. (T) Berechnen Sie die Summe der komplexen Zahlen $2 + 3i$ und $1 - i$ und illustrieren Sie Ihre Rechnung graphisch.

Der oben definierte Betrag $|z|$ einer komplexen Zahl $z = a + ib$ stimmt nach Pythagoras mit dem Abstand des Punktes (a, b) vom Koordinatenursprung überein, lässt sich also als die Länge der Strecke von 0 nach z deuten.

Es stellt sich die Frage, ob auch die Multiplikation eine geometrische Interpretation zulässt. Die Antwort lautet ja. Dazu empfiehlt es sich, zur sogenannten **trigonometrischen** oder **Polardarstellung** komplexer Zahlen überzugehen. Wir verwenden, einer exakteren Behandlung im Rahmen der Differentialrechnung vorgreifend, die Winkelfunktionen Cosinus und Sinus, symbolisch \cos und \sin in naiver Weise. Und zwar nutzen wir aus, dass ein Punkt der Ebene eindeutig gekennzeichnet ist, wenn wir wissen, in welche Richtung und wie weit wir, beim Ursprung beginnend, gehen müssen. Die Richtung ist gegeben durch einen Winkel α , der mit der positiven reellen Achse eingeschlossen wird, der Abstand von 0 ist eine nichtnegative Zahl $r = |z|$. Für $z = a + ib$ gilt für den Realteil $a = r \cos \alpha$, für den Imaginärteil $b = r \sin \alpha$. Wir können somit schreiben:

$$z = a + ib = r(\cos \alpha + i \sin \alpha)$$

Weil z in dieser Weise durch r und α bestimmt ist, schreiben wir etwas kürzer auch $z = [r, \alpha]$ oder – warum, wird erst durch die Differentialrechnung klar werden – in Exponentialschreibweise

$$z = re^{i\alpha}.$$

Dabei ist r weiterhin der Betrag von z , α heißt auch das **Argument** von z . Wir untersuchen zunächst das Produkt zweier komplexer Zahlen

$$z_1 = [1, \alpha] = \cos \alpha + i \sin \alpha$$

und

$$z_2 = [1, \beta] = \cos \beta + i \sin \beta$$

vom Betrag 1. Ihr Produkt ist eine komplexe Zahl $z_3 = z_1 z_2$ mit Realteil

$$\cos \alpha \cos \beta - \sin \alpha \sin \beta$$

und Imaginärteil

$$\sin \alpha \cos \beta + \cos \alpha \sin \beta.$$

Die **Additionstheoreme** für Cosinus und Sinus besagen:

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$$

und

$$\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta.$$

Übungsaufgabe 68. (E) *Beweisen Sie das Additionstheorem*

1. für den Cosinus

2. für den Sinus

auf elementargeometrischem Weg.

In 1.5.4 werden wir auch noch einen Beweis mittels Vektorrechnung führen.

Wir fassen all das zusammen zu

$$(\cos \alpha + i \sin \alpha) \cdot (\cos \beta + i \sin \beta) = \cos(\alpha + \beta) + i \sin(\alpha + \beta)$$

oder, wenn wir zu komplexen Zahlen mit beliebigen Beträgen r, s übergehen,

$$r(\cos \alpha + i \sin \alpha) \cdot s(\cos \beta + i \sin \beta) = rs(\cos(\alpha + \beta) + i \sin(\alpha + \beta)).$$

Nun zeigt sich der Vorteil der Polardarstellung, in der diese Beziehung die sehr übersichtliche Gestalt

$$[r, \alpha] \cdot [s, \beta] = [rs, \alpha + \beta]$$

annimmt. Also: Komplexe Zahlen werden multipliziert, indem man die Beträge multipliziert und die Winkel (Argumente) addiert. Multiplikation mit der komplexen Zahl $[r, \alpha]$ bedeutet also eine Streckung um den Faktor r , verknüpft mit einer Drehung um den Winkel α . Später werden wir die Beziehung $e^{i\alpha} = \cos \alpha + i \sin \alpha$, kennenlernen, womit sich auch die Exponentialschreibweise

$$(re^{i\alpha})(se^{i\beta}) = (rs)e^{i(\alpha+\beta)}$$

ergibt.

Auch die Division als Umkehrung der Multiplikation ergibt sich in Polardarstellung, für $s \neq 0$, sofort als

$$\frac{[r, \alpha]}{[s, \beta]} = \left[\frac{r}{s}, \alpha - \beta \right].$$

Durch Iteration der Multiplikation erhalten wir die Regel fürs **Potenzieren**:

$$z^n = [r, \alpha]^n = [r^n, n\alpha].$$

Für $r = 1$ lässt sich das auch als sogenannte **Moivresche Formel** schreiben:

$$(\cos \alpha + i \sin \alpha)^n = \cos n\alpha + i \sin n\alpha$$

Übungsaufgabe 69. (T) Berechnen Sie Real- und Imaginärteil der Summe

$$\sum_{k=0}^n \cos k\alpha + i \sin k\alpha.$$

Hinweis: Interpretieren Sie die Summe als geometrische Reihe mit $q := e^{i\alpha} = \cos \alpha + i \sin \alpha$ und verwenden Sie die Moivresche Formel.

Offenbar lässt sich auch die Operation des Potenzierens umkehren, was bedeutet, dass in \mathbb{C} stets auch n -te Wurzeln existieren. Denn für vorgegebenes $z = [r, \alpha]$ ist $w := [\sqrt[n]{r}, \frac{\alpha}{n}]$ eine komplexen Zahl mit $w^n = [r, \alpha] = z$. Wir haben also eine n -te Wurzel w von z gefunden. Dabei ist allerdings zu beachten, dass dieses w nicht die einzige n -te Wurzel ist. Wohl ist der Betrag $\sqrt[n]{r}$ einer n -ten Wurzel von $z = [r, \alpha]$ eindeutig bestimmt, nicht jedoch das Argument. Das liegt daran, dass zwei Argumente α und β zum selben z führen, wenn sie sich genau um ein ganzzahliges Vielfaches des vollen Kreisumfangs 2π unterscheiden, wenn sie also, wie man sagt, **kongruent modulo 2π** sind, formal: wenn es ein $k \in \mathbb{Z}$ gibt mit $\alpha - \beta = 2k\pi$. Angenommen, es gibt neben w eine weitere n -te Wurzel w' von $z \neq 0$. Dann ist der Quotient $\zeta := \frac{w}{w'}$ wegen $\zeta^n = \frac{w^n}{w'^n} = \frac{z}{z} = 1$ eine n -te Wurzel von 1, eine sogenannte n -te **Einheitswurzel**¹³. Als n -te Wurzeln von z kommen also genau die Zahlen ζw mit n -ten Einheitswurzeln ζ in Frage. Über letztere verschafft man sich schnell einen Überblick, indem man $\zeta = [r_0, \beta]$, also $1 = [1, 0] = \zeta^n = [r_0^n, n\beta]$ ansetzt, $r_0 = 1$ abliest und nur noch die Lösungen von $n\beta = 2k\pi$, also $\beta_k := k \frac{2\pi}{n}$ mit $k \in \mathbb{Z}$ aufsucht. Für $k = 0, 1, \dots, n-1$ erhält man verschiedene n -te Einheitswurzeln $\zeta_k = [1, \beta_k] = [1, k \frac{2\pi}{n}]$, ab $k = n$ wiederholt sich aber alles zyklisch. Man beachte, dass $\zeta_k = \zeta_1^k$ gilt. In der komplexen Ebene bilden die n -ten Einheitswurzeln ein regelmäßiges n -Eck, dessen Ecken auf dem Einheitskreis liegen, eine davon im Punkt $1 = 1 + i0 = (1, 0)$. Wir fassen zusammen:

¹³Die Bezeichnung mit dem griechischen Buchstaben ζ , gesprochen Zeta, hat sich für Einheitswurzeln eingebürgert.

Satz 1.3.11.1. Sei $n \in \mathbb{N}$ positiv und $z = [r, \alpha]$ eine komplexe Zahl mit $r > 0$. Dann gibt es genau n verschiedene n -te Wurzeln von z . Eine davon ist $w = w_0 := [\sqrt[n]{r}, \frac{\alpha}{n}]$. Sämtliche n -te Wurzeln sind gegeben durch

$$w_k := \zeta^k w = [\sqrt[n]{r}, \frac{\alpha}{n} + k \frac{2\pi}{n}], \quad k = 0, 1, \dots, n-1,$$

wobei ζ die n -te Einheitswurzel $\zeta = [1, \frac{2\pi}{n}]$ bezeichnet.

Auch die w_k bilden in der komplexen Ebene ein regelmäßiges n -Eck mit Mittelpunkt in 0, allerdings i.A. ohne Eckpunkt auf der reellen Achse.

Für einige Winkel ist es nützlich, die Werte von Sinus und Cosinus stets parat zu haben:

$$\begin{aligned} \sin 0 = \cos \frac{\pi}{2} = 0, \quad \sin \frac{\pi}{6} = \cos \frac{\pi}{3} = \frac{1}{2}, \\ \sin \frac{\pi}{4} = \cos \frac{\pi}{4} = \frac{\sqrt{2}}{2}, \quad \sin \frac{\pi}{3} = \cos \frac{\pi}{6} = \frac{\sqrt{3}}{2} \end{aligned}$$

Übungsaufgabe 70. (E) Begründen Sie die angegebenen Werte von Sinus und Cosinus mit elementargeometrischen Überlegungen. Gehen Sie in folgenden Schritten vor:

1. Zeigen Sie, dass die Winkelsumme eines Dreiecks stets gleich $\pi = 180^\circ$ ist.
2. Welche Winkel hat folglich ein gleichseitiges Dreieck?
3. Verwenden Sie Ihr Resultat, um $\sin \frac{\pi}{6} = \cos \frac{\pi}{3}$ (warum stimmen diese beiden Werte überein?) zu bestimmen.
4. Verwenden Sie den Satz von Pythagoras für die fehlenden Werte.

Übungsaufgabe 71. (T) Berechnen Sie sämtliche n -te Wurzeln der komplexen Zahl z und stellen Sie diese graphisch dar für: 1. $z = 1$, $n = 6$ 2. $z = -2$, $n = 4$ 3. $z = \frac{-1+i}{\sqrt{2}}$, $n = 3$ 4. $z = 3 + 4i$, $n = 5$ Stellen Sie Real- und Imaginärteil der Ergebnisse, sofern möglich (das ist nicht immer der Fall), mittels reeller Wurzeln ohne Winkelfunktionen dar.

Abschließend noch eine für später sehr nützliche Bemerkung über die Rolle der komplexen Konjugation $z = a + ib \mapsto \bar{z} = a - ib$ (in Koordinatendarstellung) bzw. $z = [r, \alpha] \mapsto [r, -\alpha]$ (in Polardarstellung). Geometrisch lässt sie sich deuten als Spiegelung an der reellen Achse. Konjugation lässt genau die reelle Achse fest, d.h. $\bar{z} = z$ genau dann, wenn $z \in \mathbb{R}$. Aus der Koordinatendarstellung liest man unmittelbar $\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2$ ab (weil sich nur das Vorzeichen der Imaginärteile ändert), aus der Polardarstellung $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$ und, für $z_2 \neq 0$, $\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}$ (weil sich nur das Vorzeichen der Winkel ändert). Konjugation lässt also reelle Zahlen fest und verträgt sich auf sehr geschmeidige Weise mit den Rechenoperationen.¹⁴

¹⁴In der Algebra nennt man die Konjugation deshalb einen \mathbb{R} -Automorphismus von \mathbb{C} .

Man kann diese Tatsache auch so interpretieren: Die imaginäre Einheit i und ihre Konjugierte $-i$ spielen in Bezug auf \mathbb{R} genau dieselbe Rolle. Das überrascht nicht, weil wegen $(-i)^2 = -1$ ja auch $-i$ eine Quadratwurzel von -1 ist, also statt i als Ausgangspunkt unserer Überlegungen zu den komplexen Zahlen hätte genommen werden können. Dass wir in der komplexen Ebene den Punkt $(0, 1)$ mit i und nicht mit $-i$ identifizieren, ist also reine Konvention.

Übungsaufgabe 72. (T) Berechnen Sie die komplexe Zahl z

$$1. z = \frac{1+2i}{1+3i} \qquad 2. z = |1 \pm 7i|$$

$$3. z = \frac{2}{(3+i)^2} + \frac{2}{(3-i)^2}$$

und schreiben Sie das Ergebnis in der Standardform $\Re z + i\Im z$ an. Skizzieren Sie Ihr Ergebnis in der Gauß'schen Zahlenebene und geben Sie jeweils auch die konjugiert komplexe Zahl an.

Übungsaufgabe 73. (T) Berechnen Sie unter Benutzung der Darstellung in Polarform

$$1. (-3 + 3i)^{10} \qquad 2. (1 - \sqrt{3}i)^9 \qquad 3. \sqrt[4]{2 - 2i}$$

$$4. \sqrt[3]{i + \sqrt{3}}$$

und schreiben Sie das Ergebnis in der Standardform $\Re z + i\Im z$ an. Vereinfachen Sie dabei so weit wie möglich (nicht immer wird das in befriedigender Weise gelingen).

Übungsaufgabe 74. (T) Es sei $z = \frac{1-i}{1+i}$.

1. Zeichnen Sie z in der Gauß'schen Zahlenebene ein.

2. Schreiben Sie z in Polardarstellung an.

3. Berechnen Sie Real- und Imaginärteil von z^{99} .

Übungsaufgabe 75. (T) Wie Übungsaufgabe 74 aber mit $z = \frac{1+i}{1-i}$.

Übungsaufgabe 76. (T) Bestimmen Sie die Lösungsmengen der folgenden Gleichungen bzw. Ungleichungen und stellen Sie sie graphisch dar:

$$1. \Im z = -1, \qquad 2. \Re(-iz) = 1, \qquad 3. \Re(z(1+i)) = 1,$$

$$4. \Re z^2 \geq 1, \qquad 5. \bar{z}^2 = z^2, \qquad 6. |z+1+i| = 1,$$

$$7. |\Re z - \Im z| \leq 1, \qquad 8. (\Re z)(\Im z) \leq 1.$$

Übungsaufgabe 77. (E) Überlegen Sie sich allein mit Hilfe der Axiome für einen angeordneten Körper (insbesondere der Monotoniegesetze), dass in \mathbb{R} (wie in jedem angeordneten Körper) für alle $a \in \mathbb{R}$ stets $a^2 \geq 0$ gilt, Quadrate also nie negativ sein können. Schließen Sie daraus auch $0 < 1$ und damit weiter, dass \mathbb{C} mit dem Element i , welches $i^2 = -1$ erfüllt, nicht zu einem angeordneten Körper gemacht werden kann. Hinweis: Verwenden Sie, dass für alle $a \in \mathbb{R}$ stets $a \geq 0$ oder $-a \geq 0$ gilt, das Monotoniegesetz für die Multiplikation und die etwas weiter oben bewiesene Rechenregel $(-a)(-b) = ab$ für den Spezialfall $b = a$.

1.4 Elementare Kombinatorik

Die Hauptaufgabe der Kombinatorik ist die Bestimmung der Anzahl der Elemente von Mengen (insbesondere endlichen), die auf unterschiedliche Weise zustande kommen können. Die wichtigsten der elementaren Beispiele wollen wir in diesem Abschnitt behandeln: einfache Anzahlformeln, insbesondere das Inklusions-Exklusionsprinzip in 1.4.1, Permutationen und Faktorielle in 1.4.2, Kombinationen als Anzahlen von Teilmengen in 1.4.3 und den binomischen und polynomischen Lehrsatz als Anwendung in 1.4.4.

1.4.1 Einfachste Anzahlformeln

Inhalt in Kurzfassung: Disjunkten Vereinigungen entspricht bei den Kardinalitäten (Anzahlen) die Summe, kartesischen Produkten das Produkt natürlicher Zahlen. Die Kardinalität einer Vereinigung mehrerer beliebiger (d.h. i.a. nicht disjunkter) endlicher Mengen kann mit dem Inklusions-Exklusionsprinzip berechnet werden.

Im Zusammenhang mit den Rechenoperationen auf \mathbb{N} haben wir Addition und Multiplikation motiviert über ihre Bedeutung bei mengentheoretischer Vereinigung und kartesischem Produkt zweier Mengen. Die offensichtlichen Verallgemeinerungen auf n statt nur zwei Mengen lauten

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{i=1}^n |A_i|, \quad \text{sofern die Mengen } A_i \text{ paarweise disjunkt sind,}$$

und

$$|A_1 \times A_2 \times \dots \times A_n| = \prod_{i=1}^n |A_i|.$$

Diese Zusammenhänge können verwendet werden, um auch von Summen oder Produkten unendlicher Mächtigkeiten, sogenannter Kardinalzahlen zu sprechen. Wir können das hier nicht vertiefen und wollen uns statt dessen überlegen, was bei möglicherweise nicht disjunkten Vereinigungen zu tun ist. Für nur zwei Mengen macht man sich sehr schnell klar:

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Für drei Mengen ist die Situation ebenfalls noch recht überschaubar:

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

Übungsaufgabe 78. (T) Bestimmen Sie, wie viele natürliche Zahlen von 1 bis 600 durch mindestens eine der Zahlen

(a) 2, 3 oder 12

(b) 3, 5 oder 8

(c) 6, 8 oder 14

teilbar sind.

Übungsaufgabe 79. (P) Formulieren und begründen Sie eine analoge Formel für die Vereinigung $A \cup B \cup C \cup D$ von vier Mengen.

Auch für eine beliebige endliche Anzahl n von Mengen A_1, \dots, A_n lässt sich die Kardinalität der Vereinigung durch eine entsprechende Formel ausdrücken. Dabei kommen alle Durchschnitte, die aus den A_i gebildet werden, vor. Das sind die Mengen $A_I = \bigcap_{i \in I} A_i$ für beliebige nichtleere (Index-)Teilmengen $I \subseteq \{1, 2, \dots, n\}$. Die Formel nennt man auch **Inklusions-Exklusionsprinzip**. Sie lautet:

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{I: \emptyset \neq I \subseteq \{1, 2, \dots, n\}} (-1)^{|I|+1} |A_I|$$

Die Schreibweise der Summe rechts bedeutet, dass für jede nichtleere Teilmenge I von $\{1, 2, \dots, n\}$ mit ungerader Kardinalität $|I|$ der Summand $|A_I|$ mit positivem Vorzeichen versehen wird, sonst mit negativem.

Der Beweis erfolgt am besten mittels Induktion nach n . Beim Schritt von n auf $n+1$ geht man von der Beziehung

$$\bigcup_{i=1}^{n+1} A_i = \left(\bigcup_{i=1}^n A_i \right) \cup A_{n+1}$$

aus und verwendet zunächst die Formel für die Vereinigung von nur zwei Mengen:

$$\left| \bigcup_{i=1}^{n+1} A_i \right| = \left| \bigcup_{i=1}^n A_i \right| + |A_{n+1}| - \left| \bigcup_{i=1}^n (A_i \cap A_{n+1}) \right|.$$

Auf die beiden darin auftretenden Vereinigungen von n Mengen lässt sich die Induktionsannahme anwenden, wodurch man nach kurzer Analyse der Vorzeichen sehr schnell die behauptete Formel erkennt. Das ist eine gute Übung für die Arbeit mit Indizes:

Übungsaufgabe 80. (E) Führen Sie die noch fehlenden Rechnungen im Beweis des Inklusions-Exklusionsprinzips aus.

1.4.2 Anzahlen von Funktionen und Permutationen; Faktorielle

Inhalt in Kurzfassung: So wie Addition von Vereinigungen und Multiplikation von kartesischen Produkten, so lässt sich das Potenzieren kombinatorisch von der Anzahlbestimmung von Funktionen von einer Menge in eine andere gegebene ableiten. Beschränkt man sich auf injektive oder gar auf bijektive Funktionen (analog: auf Permutationen einer Menge), so stößt man notgedrungen sehr schnell auf Faktorielle.

Seien A und B endliche Mengen mit $m := |A|$ und $n := |B|$ Elementen. Sei F die Menge aller Funktionen $f : A \rightarrow B$. Wir fragen nach der Anzahl $|F|$. Für jedes Element $a \in A$ gibt es n Möglichkeiten, es auf ein $b = f(a) \in B$ abzubilden. Außerdem können wir unsere Wahl für jedes $a \in A$ unabhängig von den anderen treffen, also multiplizieren sich die Möglichkeiten auf: n^2 für $m = 2$, n^3 für $m = 3$ etc. Allgemein gilt daher $|F| = n^m$ für die Menge aller Funktionen von einer m - in eine n -elementige Menge. (Man spricht auch von **Variationen mit Wiederholung**.) Deshalb schreibt man für die Menge F oft auch B^A . Ist $A = \{a_1, \dots, a_m\}$, so kann man ein $f \in F$ als m -tupel $(f(a_1), \dots, f(a_m))$ codieren, also als Element des kartesischen Produktes $B \times \dots \times B$ (m -mal). Das passt wegen $|B| = n$ auch bestens mit $n^m = n \cdot \dots \cdot n$ (m -mal) zusammen.

Wir schränken uns nun auf injektive Abbildungen $f : A \rightarrow B$ ein. Interessant ist nur der Fall $m = |A| \leq |B| = n$, weil es andernfalls gar keine solchen f gibt. Ist wieder $A = \{a_1, \dots, a_m\}$ und $|B| = n$, so haben wir n Möglichkeiten für $f(a_1) \in B$. Für $f(a_2)$ stehen uns jetzt aber nur mehr $n - 1$ Möglichkeiten offen, nämlich alle $b \in B \setminus \{f(a_1)\}$, für $f(a_3)$ nur mehr $n - 2$ Möglichkeiten etc. Somit gibt es insgesamt $n \cdot (n - 1) \cdot \dots \cdot (n - m + 1)$ injektive Abbildungen von A nach B oder, in anderer Sprechweise, m -tupel aus paarweise verschiedenen Elementen von B . (Man spricht auch von **Variationen ohne Wiederholung**). Mit der Notation $n! = 1 \cdot 2 \cdot \dots \cdot n$ lässt sich diese Anzahl offenbar auch schreiben als $\frac{n!}{(n-m)!}$. Im Fall $m = n$ sind injektive Abbildungen von A nach B sogar bijektiv. Nach denselben Überlegungen wie bisher entsprechen diese jenen n -tupeln, in denen jedes von n Elementen (jene von B) genau einmal auftritt, und es gibt genau $n!$ davon. An dieser Anzahl ändert sich nichts, wenn $A = B = M$ mit $|M| = n$ ist. Also zählt $n!$ auch die bijektiven Abbildungen einer Menge M mit n Elementen auf sich selbst, die sogenannten **Permutationen** von M . Nochmals anders formuliert: Es gibt $n!$ mögliche Anordnungen von n Elementen. Überdies legt die im Fall $m = n$ resultierende Gleichheit $\frac{n!}{(n-m)!} = n!$ die Definition $0! = (n - m)! = 1$ nahe. Das passt auch mit unserer Interpretation zusammen: Es gibt genau eine Möglichkeit, 0 Symbole anzuschreiben, nämlich gar nichts hinzuschreiben.

Führt man das Symbol $n!$ (gesprochen als n **Faktorielle** oder n **Fakultät**) zunächst ohne kombinatorische Bezugnahme rein arithmetisch ein, so definiert man daher rekursiv: $0! = 1$ und $(n + 1)! = n!(n + 1)$. Als weitere Werte ergeben sich $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, $6! = 720$, $7! = 5040$, $8! = 40320$, $9! = 362880$, $10! = 3628800, \dots$; eine äußerst rasch wachsende Zahlenfolge!

Übungsaufgabe 81. (P) Welche Endziffern (in der Dezimaldarstellung) von $n!$ können auftreten?

Übungsaufgabe 82. (T) Zeigen Sie mit vollständiger Induktion, dass $n! > 2^n$ für alle $n = 4, 5, 6, \dots$

Ohne Beweis geben wir die oft sehr nützliche **Stirlingsche Formel** an, welche das Wachstum von $n!$ beschreibt:

$$n! \sim \frac{n^n \sqrt{2\pi n}}{e^n}$$

Das Symbol \sim soll bedeuten, dass der prozentuelle Unterschied zwischen den Größen rechts und links beliebig nahe bei 0 liegt, sofern nur n genügend groß wird. Man beachte, dass in dieser Formel (wie auch in unzähligen anderen vergleichbaren Fällen) die Kreiszahl $\pi = 3,14159\dots$ vorkommt, obwohl ein Zusammenhang mit ihrer geometrischen Bedeutung (Verhältnis von Umfang zu Durchmesser eines Kreises) überhaupt nicht offensichtlich ist.

1.4.3 Teilmengen, Kombinationen und Binomialkoeffizienten

Inhalt in Kurzfassung: Kombinatorisch treten die sogenannten Binomialkoeffizienten auf, wenn man die k -elementigen Teilmengen einer n -elementigen Menge zählt. Ihre arithmetische Bedeutung wird erst später, im binomischen Lehrsatz sichtbar werden.

Eine typische kombinatorische Fragestellung lautet: Wie viele Teilmengen hat eine Menge M mit n Elementen? Antwort: 2^n . Begründung: Für jedes der n Elemente gibt es zwei Möglichkeiten (nennen wir sie 1 und 0): es in eine Teilmenge T aufzunehmen oder nicht. Wollen wir T insgesamt festlegen, müssen wir für jedes der n Elemente von M eine solche 0-1-Entscheidung treffen, d.h.: Jedem $T \subseteq M$ entspricht genau eine 0-1-Folge der Länge n , also ein Element von $\{0, 1\} \times \dots \times \{0, 1\}$ (n -mal), und vice versa. Also gibt es $2 \cdot 2 \cdot \dots \cdot 2 = 2^n$ Teilmengen $T \subseteq M$. Man beachte, dass hier implizit eine Bijektion zwischen der Menge der Teilmengen von M und der Menge aller 0-1-Folgen der Länge n verwendet wurde. Weil es eine solche Bijektion gibt, haben die Mengen gleich viele Elemente.

Etwas feinere Überlegungen braucht es, um die Anzahl der k -elementigen Teilmengen einer n -elementigen Menge M zu bestimmen. (Man spricht auch von **Kombinationen (ohne Wiederholung)**.) Wir bezeichnen diese Anzahl mit $\binom{n}{k}$. Wir setzen hier $k \leq n$ voraus, weil ja andernfalls trivialerweise $\binom{n}{k} = 0$ ist. Wir betrachten die Abbildung f , die jeder der $n!$ Anordnungen $o = (m_1, m_2, \dots, m_k, m_{k+1}, \dots, m_n)$ der n Elemente von M jene k -elementige Teilmenge zuordnet, die aus den ersten k Elementen in dieser Anordnung besteht, also $f(o) = \{m_1, m_2, \dots, m_k\}$. Diese Abbildung $f: O(M) \rightarrow \mathcal{T}(k, M)$ von der Menge $O(M)$ aller Anordnungen der Elemente von M in die Menge $\mathcal{T}(k, M)$ aller k -elementigen Teilmengen von M ist f surjektiv, nicht jedoch injektiv. Gehen wir von zwei Anordnungen $o, o' \in O(M)$ aus, so gilt $f(o) = f(o')$ genau dann, wenn o' aus o dadurch hervorgeht, dass man die ersten k Elemente m_1, \dots, m_k gemäß einer geeigneten Permutation π untereinander vertauscht (es gibt $k!$ solche π) und ebenso durch eine Permutation ρ die verbleibenden $n - k$ Elemente m_{k+1}, \dots, m_n untereinander (es gibt $(n - k)!$ solche ρ). Es gibt also $k!(n - k)!$ solche Paare (π, ρ) , und jedes solche Paar entspricht genau einem o' mit $f(o') = f(o)$. Unter f werden also jeweils $k!(n - k)!$ verschiedene Elemente von O auf dasselbe Element von $\mathcal{T}(k, M)$ abgebildet, und genau die Elemente aus $\mathcal{T}(k, M)$ werden von f getroffen. Folglich gilt $|O| = k!(n - k)! |\mathcal{T}(k, M)|$.

Wegen $|O| = n!$ folgt daraus

$$\binom{n}{k} = |\mathcal{T}(k, M)| = \frac{n!}{k!(n-k)!}.$$

Diese Zahlen heißen, weil sie im Binomischen Lehrsatz auftreten, auch **Binomialkoeffizienten**.

Oft ist es praktisch, die Größenordnung von Binomialkoeffizienten abzuschätzen. Hält man zum Beispiel k fest, stellt sich heraus, dass sich $\binom{n}{k}$ etwa so verhält wie ein bestimmter Bruchteil der k -ten Potenz n^k :

Übungsaufgabe 83. (P) Sei $k \in \mathbb{N}$ fest vorgegeben. Finden Sie eine positive reelle Zahl c_k , von der Sie zeigen können: Es gibt ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ gilt:

$$c_k n^k \leq \binom{n}{k} \leq n^k.$$

Hinweis: Stünde an dieser Stelle schon der Begriff des Grenzwertes zur Verfügung ließe sich stärker zeigen:

$$\lim_{n \rightarrow \infty} \binom{n}{k} n^{-k} = \frac{1}{k!}$$

Diesen Hinweis sollen Sie in Ihrem Beweis hier allerdings noch nicht verwenden. Er kann aber zur Kontrolle dienen.

1.4.4 Binomischer und polynomialer Lehrsatz

Inhalt in Kurzfassung: Die n -te Potenz einer Summe $a+b$ lässt sich laut binomischem Lehrsatz als Summe von Ausdrücken der Form $a^{n-k}b^k$ mit $k = 0, \dots, n$ darstellen, deren Vielfachheiten durch die Binomialkoeffizienten $\binom{n}{k}$ gegeben sind, die wir als Anzahlen k -elementiger Teilmengen einer n -elementigen Menge kennen gelernt haben. Die Verallgemeinerung auf mehr als zwei Summanden nennt man poly- oder multinomischen Lehrsatz.

Häufig verwendet man folgende Regeln für Binomialkoeffizienten ($0 \leq k \leq n \in \mathbb{N}$):

$$\binom{n}{k} = \binom{n}{n-k}, \quad \text{speziell: } \binom{n}{0} = \binom{n}{n} = 1 \quad \text{und} \quad \binom{n}{1} = \binom{n}{n-1} = n.$$

Besonders wichtig ist die Rekursionsformel

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

Diese Formel lässt sich leicht nachrechnen.

Übungsaufgabe 84. (T) Überprüfen Sie diese Formel mit Hilfe der allgemeinen Formel $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ für Binomialkoeffizienten.

Reizvoller ist die kombinatorische Deutung: Entstehe die $n + 1$ -elementige Menge $M' = M \cup \{m'\}$, indem man zu einer n -elementigen Menge M noch ein weiteres Element $m' \notin M$ hinzufügt. Unter den $k + 1$ -elementigen Teilmengen T' von M' gibt es zwei Sorten; die erste Sorte mit $m' \in T'$, die zweite mit $m' \notin T'$. Zur ersten Sorte: Entfernt man m' aus T' , so erhält man eine k -elementige Teilmenge T von M , jede genau einmal. Davon gibt es $\binom{n}{k}$, also gibt es auch so viele T' der ersten Sorte. Von der zweiten Sorte sind genau die $k + 1$ -elementigen Teilmengen von M , wovon es $\binom{n}{k+1}$ gibt. Die Gesamtzahl $\binom{n+1}{k+1}$ der T' ist also tatsächlich die Summe dieser beiden Anzahlen $\binom{n}{k}$ und $\binom{n}{k+1}$.

Mit Hilfe der bisherigen Regeln lässt sich auf sehr einprägsame Weise das aus den Binomialkoeffizienten bestehende **Pascalsche Dreieck** aufbauen. Der Binomialkoeffizient $\binom{n}{k}$ steht dabei in der n -ten Zeile und k -ten Spalte (Numerierung beginnt bei 0), d.h. die Spitze des Dreiecks bildet $\binom{0}{0} = 1$, symmetrisch schräg darunter stehen ebenfalls die zwei Einsen, links $\binom{1}{0} = 1$ und rechts $\binom{1}{1} = 1$, in der nächsten Reihe $\binom{2}{0} = 1$, $\binom{2}{1} = 2$ und $\binom{2}{2} = 1$ etc. Außen stehen wegen $\binom{n}{0} = \binom{n}{n} = 1$ stets Einsen, und die Eintragungen dazwischen ergeben sich nach obiger Rekursionsformel als Summe der Eintragungen unmittelbar schräg darüber.

Übungsaufgabe 85. (T) Berechnen Sie das Pascalsche Dreieck bis zur Zeile $n = 10$.

Eine ihrer wichtigsten Anwendungen haben die Binomialkoeffizienten und das Pascalsche Dreieck im **Binomischen Lehrsatz (Binomische Formel)**:

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$$

Diese Formel gilt für alle $n \in \mathbb{N}$ und $a, b \in \mathbb{R}$ (oder auch $a, b \in \mathbb{C}$). Der Beweis kann z.B. mittels Induktion nach n erfolgen, wobei beim Induktionsschritt die Rekursionsformel $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$ von weiter oben die entscheidende Rolle spielt. Für $n = 0$ ist alles trivial, weil links wie rechts nur 1 steht. Im Induktionsschritt setzt man in $(a + b)^{n+1} = (a + b)^n(a + b)$ für den ersten Faktor die Induktionsannahme $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$ ein, und multipliziert mit a und mit b . Die beiden resultierenden Summen können nach einer geeigneten Indexverschiebung zu einer einzigen zusammengefasst werden, die der rechten Seite der Binomischen Formel für $n + 1$ statt n entspricht. Die einzelnen Rechenschritte sind wieder eine gute Übung.

Übungsaufgabe 86. (T) Führen Sie den angedeuteten Induktionsbeweis der Binomischen Formel vollständig aus.

Übungsaufgabe 87. (T) Zeigen Sie direkt mithilfe des Binomischen Lehrsatzes:

$$1. \sum_{k=0}^n \binom{n}{k} = 2^n, \text{ für } n \geq 0$$

2. $\sum_{k=0}^n (-1)^k \binom{n}{k} = 0$, für $n \geq 0$
3. $(1+x)^n \geq 1 + \frac{n^2}{4}x^2$, für $n \geq 2$ und $x \geq 0$.

Die Verallgemeinerung des Binomischen Lehrsatzes von zwei auf $m \in \mathbb{N}$ Summanden ist der **Poly-** oder auch **Multinomische Lehrsatz** :

$$\left(\sum_{i=1}^m a_i \right)^n = \sum_{(k_1, \dots, k_m) \in I} \frac{n!}{\prod_{i=1}^m k_i!} \prod_{i=1}^m a_i^{k_i}$$

Dabei erstreckt sich die Summe über die Indexmenge I , bestehend aus allen m -tupeln natürlicher Zahlen k_1, \dots, k_m mit $\sum_{i=1}^m k_i = n$.

Übungsaufgabe 88. (E) Für $m = 1$ ist der polynomische Lehrsatz trivial, für $m = 2$ geht er in den binomischen über. Beweisen Sie den Polynomischen Lehrsatz in seiner allgemeinen Form mittels Induktion nach m .

Ein alternativer Beweis gelingt durch naheliegende Übertragung der Ideen beim Binomischen Lehrsatz ohne grundsätzlich Neues: Zuerst erkennt man auf ganz ähnliche Weise wie bei der kombinatorischen Deutung der Binomialkoeffizienten, dass die in der Formel auftretenden Terme $\frac{n!}{\prod_{i=1}^m k_i!}$ zählen, auf wie viele Arten man $n = \sum_{i=1}^m k_i$ Objekte auf m wohlunterschiedene Behälter, mit Fassungsvermögen k_1, k_2, \dots, k_m aufteilen kann. Sodann überlegt man sich, dass dabei dasselbe kombinatorische Problem vorliegt, wie wenn man fragt, wie oft das Produkt $\prod_{i=1}^m a_i^{k_i}$ beim Ausmultiplizieren der n -ten Potenz der Summe der a_i auftritt.

1.5 Vektoren im n -dimensionalen Raum \mathbb{R}^n

Im letzten Abschnitt dieses Kapitels widmen wir uns der Vektorrechnung. Im Gegensatz zum umfangreichen Kapitel über Lineare Algebra im zweiten Semester betreiben wir aber keine abstrakte algebraische Strukturtheorie. Das Ziel dieses Abschnitts besteht darin, die Sprache der Vektorrechnung zu nutzen, um geometrische Gegebenheiten besser beschreiben zu können. Im Wesentlichen denkt man dabei an ein kartesisches Koordinatensystem in der zweidimensionalen Ebene, im dreidimensionalen oder auch im n -dimensionalen Raum \mathbb{R}^n . Damit lassen sich die Vektorraumoperationen definieren (1.5.1). Die wichtigsten Rechenregeln sind Inhalt von 1.5.2. Indem man sich auf eine Basis bezieht, wobei wir hier nur die kanonische betrachten (1.5.3), lassen sich zum Beispiel Drehungen als Beispiele linearer Abbildungen beschreiben (1.5.4). Als Nebenprodukt erhält man die Additionstheoreme für Sinus und Cosinus. Darüber hinaus behandeln wir das (Standard-) Skalarprodukt (1.5.5), die euklidische Länge/Norm/Metrik (1.5.6) und das Vektorprodukt im \mathbb{R}^3 (1.5.7). Schließlich beschreiben wir einfache geometrische Figuren (1.5.8) und führen topologischen Begriffe wie z.B. die der offenen und abgeschlossenen Menge ein (1.5.9).

1.5.1 Koordinatisierung und Vektorraumoperationen

Inhalt in Kurzfassung: Der Vektorraum \mathbb{R}^2 entsteht formal als kartesisches Produkt, indem man alle geordneten Paare (a, b) mit $a, b \in \mathbb{R}$ bildet und komponentenweise Addition und Multiplikation mit einem Skalar definiert. Geometrisch entspricht dem die Vektoraddition nach der Parallelogrammregel, bzw. Verlängerung oder Verkürzung um einen skalaren Faktor. Jeder zweidimensionale Vektor kann auch als Punkt der Ebene interpretiert werden. Entsprechendes gilt für \mathbb{R}^3 als dreidimensionaler Anschauungsraum. Weniger anschaulich aber formal völlig analog lassen sich auch die n -dimensionalen Vektorräume \mathbb{R}^n definieren.

Auf der (reellen) Zahlengeraden werden Punkte mit Zahlen $x \in \mathbb{R}$ identifiziert. Im Falle der (komplexen) Zahlenebene \mathbb{C} entsprechen den Punkten Zahlenpaare $\mathbf{x} = (x_1, x_2)$. Weil zwei Zahlen nötig sind, nennt man die Ebene zweidimensional. Als Menge kann \mathbb{C} mit dem kartesischen Produkt $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ gleichgesetzt werden. Normalerweise werden wir annehmen, dass die Koordinatenachsen im rechten Winkel aufeinander stehen, in welchem Fall man von einem kartesischen Koordinatensystem spricht.

Entsprechend können die Punkte des dreidimensionalen Raumes durch Zahlentripel $\mathbf{x} = (x_1, x_2, x_3)$, die aus drei Komponenten $x_1, x_2, x_3 \in \mathbb{R}$ bestehen, angegeben werden, und wir schreiben \mathbb{R}^3 für den dreidimensionalen Raum. Analog interpretieren wir

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_i \in \mathbb{R}\}$$

mit beliebigem $n \in \mathbb{N}$ als n -dimensionalen Raum.

Die Elemente $\mathbf{x} = (x_1, x_2, \dots, x_n)$ von \mathbb{R}^n sind n -tupel, die wir je nach Kontext ebenfalls **Punkte** nennen oder auch **Vektoren** mit den **Koordinaten** x_1, \dots, x_n . Oft entsteht ein übersichtlicheres Schriftbild, wenn man die Komponenten der n -tupel nicht waagrecht (Zeilenvektor), sondern senkrecht als Spaltenvektor schreibt, also:

$$\mathbf{x} = (x_1, x_2, \dots, x_n) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Will man die fett gedruckten Symbole $\mathbf{x}, \mathbf{y}, \dots$ vermeiden, bezeichnet man Vektoren oft auch mit \vec{x}, \vec{y}, \dots

Beim Matrizenkalkül, den wir in Mathematik 2 im Kapitel über Lineare Algebra kennen lernen werden, ist zwischen Zeilen- und Spaltenvektoren zu unterscheiden. Jetzt ist das aber noch nicht notwendig, weil es nur um zwei verschiedene Schreibweisen für dieselben Objekte, nämlich Elemente des \mathbb{R}^n geht. Wir werden deshalb meist zur jeweils gerade sympathischeren Darstellung greifen.

Beim Wort *Vektor* schwingt die Idee einer durch einen Pfeil repräsentierten Bewegung mit, die vom Ursprung (Nullvektor) des Koordinatensystems

$(0, 0, \dots, 0)$ nach (x_1, x_2, \dots, x_n) führt. Oft denkt man aber auch nur an den Punkt (x_1, x_2, \dots, x_n) . Beide Vorstellungen (Vektor und Punkt) werden durch dieselben mathematischen Objekte, nämlich eben durch die n -tupel repräsentiert. Steht die Interpretation als Punkt im Vordergrund, insbesondere für die Dimensionen $n = 2$ und $n = 3$, so schreibt man manchmal auch $(x|y)$ oder (x/y) bzw. $(x|y|z)$ oder $(x/y/z)$ statt (x, y) bzw. (x, y, z) und A, B, C, \dots statt $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$.

Die Interpretation als Vektor legt es nahe, nicht nur am Ursprung anzusetzen, sondern an einem beliebigen Punkt. Das führt zur **Vektoraddition**:

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

Ganz ähnlich kann man Vektoren auch mit **Skalaren** (Zahlen) $r \in \mathbb{R}$ multiplizieren:

$$r\mathbf{x} = r \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} rx_1 \\ rx_2 \\ \vdots \\ rx_n \end{pmatrix}$$

Wir wollen diese Multiplikation **Skalarmultiplikation** (Skalar mal Vektor ergibt Vektor) nennen. Man lasse sich durch die (etwas missverständliche) Wortwahl nicht verwirren: Für $n > 1$ ist diese Skalarmultiplikation zu unterscheiden von der Multiplikation zweier Skalare (also der gewöhnlichen Multiplikation zweier Zahlen), außerdem von zwei später noch zu behandelnden Operationen von Vektoren, dem sogenannten Skalarprodukt (Vektor mal Vektor ergibt Zahl) und dem Vektorprodukt im \mathbb{R}^3 (Vektor mal Vektor ergibt Vektor).

Die geometrischen Interpretationen der **Vektorraumoperationen** (Addition und Skalarmultiplikation) für $n = 1, 2, 3$ liegen auf der Hand: Die Addition folgt, so wie die Addition komplexer Zahlen in der Ebene als 2-dimensionaler Spezialfall (siehe 1.3.11), der Parallelogrammregel. Multiplikation mit dem Skalar r entspricht der Streckung des Pfeils um den Faktor r .

Übungsaufgabe 89. (T) Geben Sie sich zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ vor, deren Koordinaten alle von 0 verschieden sind, berechnen Sie die Summe $\mathbf{z} = \mathbf{x} + \mathbf{y}$ und illustrieren Sie Ihre Rechnung mit einer sorgfältigen Skizze.

1.5.2 Eigenschaften der Vektorraumoperationen

Inhalt in Kurzfassung: Vertraute Gesetze (Assoziativgesetz u.ä.) vererben sich von den reellen Zahlen wegen der komponentenweisen Definition für Vektoren auch auf Vektorräume. Deshalb ist jeder Vektorraum auch eine abelsche Gruppe mit zusätzlicher Skalarmultiplikation, die gleichfalls die üblichen Gesetze erfüllt.

Folgende Eigenschaften der Vektorraumoperationen sind besonders hervorzuheben:

Die Elemente (Vektoren) von \mathbb{R}^n bilden bezüglich $+$ eine abelsche Gruppe. Das heißt, zur Erinnerung, explizit: Es gelten Kommutativgesetz $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ und Assoziativgesetz $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ (für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$). Überdies gibt es ein neutrales Element, den sogenannten **Nullvektor**

$$\mathbf{0} := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n,$$

weil $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$ für alle $\mathbf{x} \in \mathbb{R}^n$ gilt. Schließlich gibt es zu jedem $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ ein (eindeutig bestimmtes) Inverses $-\mathbf{x}$, nämlich

$$-\mathbf{x} := \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix},$$

mit $\mathbf{x} - \mathbf{x} = \mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

Für die Skalarmultiplikation gelten außerdem die beiden Distributivgesetze $(r+s)\mathbf{x} = r\mathbf{x} + s\mathbf{x}$ und $r(\mathbf{x} + \mathbf{y}) = r\mathbf{x} + r\mathbf{y}$, ein Assoziativgesetz, nämlich $(rs)\mathbf{x} = r(s\mathbf{x})$, und $1\mathbf{x} = \mathbf{x}$ (der Skalar 1 ist also auch bezüglich der Skalarmultiplikation eine Art neutrales Element), jeweils für $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ und $r, s \in \mathbb{R}$.

Übungsaufgabe 90. (P) Überprüfen Sie die vier angegebenen Rechenregeln: zwei Distributivgesetze, Assoziativgesetz und 1 als neutrales Element.

Man könnte weitere geläufige Rechenregeln angeben. Die meisten davon erweisen sich aber als Folgerungen der obigen. Deshalb wählt man die angegebenen Regeln zur Definition von Vektorräumen als abstrakten Strukturen. Dieser Gesichtspunkt wird im zweiten Semester im Kapitel über Lineare Algebra ausführlicher behandelt werden. Hier beziehen wir uns ausschließlich auf die Räume \mathbb{R}^n , wie wir sie im vorangegangenen Unterabschnitt eingeführt haben. Wir nennen diese Räume auch die n -dimensionalen **Standardvektorräume**.

1.5.3 Die kanonische Basis

Inhalt in Kurzfassung: Jene Vektoren, in denen eine Komponente = 1 ist, die übrigen = 0, heißen kanonische Basisvektoren. Mit ihnen kann jeder beliebige Vektor eindeutig als Linearkombination dargestellt werden.

Eine besondere Rolle spielen im Standardvektorraum \mathbb{R}^n die sogenannten **kanonischen Einheitsvektoren** $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$. Der Vektor $\mathbf{e}_i = (0, 0, \dots, 0, 1, 0, \dots, 0)$ heißt i -ter Einheitsvektor. Er hat 1 als i -te Koordinate, alle anderen Koordinaten von \mathbf{e}_i sind 0. Gemeinsam bilden die \mathbf{e}_i die **kanonische Basis** $B = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$.

Offenbar lässt sich jeder beliebige Vektor $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ in eindeutiger Weise als **Linearkombination** der \mathbf{e}_i (also als Summe von Vielfachen der \mathbf{e}_i) schreiben, nämlich

$$\mathbf{x} = (x_1, x_2, \dots, x_n) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \sum_{i=1}^n x_i \mathbf{e}_i.$$

Man sagt, dass die \mathbf{e}_i den ganzen Raum \mathbb{R}^n erzeugen oder aufspannen und, weil die Darstellung eindeutig ist, dass sie linear unabhängig sind. Ganz allgemein charakterisieren die beiden Eigenschaften „erzeugen“ und „linear unabhängig“ eine **Basis**. Man könnte auch andere Basen verwenden. Abgesehen von der folgenden Übungsaufgabe werden wir dem aber erst im Kapitel über Lineare Algebra systematisch nachgehen.

Übungsaufgabe 91. (E) Zeigen Sie, dass auch die Vektoren $\mathbf{e}'_1 := \mathbf{e}_1$ und $\mathbf{e}'_i := \mathbf{e}_{i-1} + \mathbf{e}_i$ für $i = 2, \dots, n$ eine Basis von \mathbb{R}^n bilden, d.h.: Zu jedem Vektor $\mathbf{x} \in \mathbb{R}^n$ gibt es eindeutige Skalare x'_1, \dots, x'_n mit $\mathbf{x} = \sum_{i=1}^n x'_i \mathbf{e}'_i$.

1.5.4 Linearität am Beispiel der Drehung

Inhalt in Kurzfassung: Lineare Abbildungen sind solche, die mit den Vektorraumoperationen verträglich sind. Das läuft darauf hinaus, dass sie die Koeffizienten von Linearkombinationen unverändert lassen. Als Konsequenz gibt es zwischen zwei Vektorräumen genau eine lineare Abbildung, die auf den Elementen einer Basis vorgegebene Werte hat. Somit enthält eine Tabelle mit den Koordinaten dieser Werte (eine sogenannte Matrix) die gesamte Information über die lineare Abbildung. Wie sich das sehr effektiv ausnützen lässt, wird anhand von Drehungen der Ebene illustriert. Als Ernte fallen dabei auch die Additionstheoreme für Cosinus und Sinus heraus.

Zentral in der Linearen Algebra ist der Begriff der **linearen Abbildung**. Dabei handelt es sich um eine Abbildung f zwischen zwei Vektorräumen, die mit den Vektorraumoperationen (Linearkombinationen) verträglich ist in dem Sinn, dass stets $f(r\mathbf{x} + s\mathbf{y}) = rf(\mathbf{x}) + sf(\mathbf{y})$ gilt. Weniger abstrakt formuliert: Das Bild des r -fachen eines Vektors ist das r -fache des Bildes, als Formel: $f(r\mathbf{x}) = rf(\mathbf{x})$; und das Bild einer Summe ist die Summe der Bilder, als Formel: $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$.¹⁵

Hier soll der anschaulich-geometrische Gehalt dieser Definition am Beispiel der Drehung $d_\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ illustriert werden, welche die Ebene \mathbb{R}^2 um den

¹⁵Um Missverständnissen vorzubeugen: Im Zusammenhang mit Polynomen und reellen Funktionen bezeichnet man oft jede Funktion f der Bauart $f(x) = kx + d$ mit $k, d \in \mathbb{R}$ als linear, siehe auch 3.2. Ist so ein f aber auch im hier definierten Sinn linear, so folgt $d = f(0) = f(0 + 0) = f(0) + f(0) = 2d$, also $d = 0$. Diesen Fall nennt man auch den **homogenen**, und zwar im Unterschied zum **inhomogenen** mit $d \neq 0$.

Winkel α in die sogenannte mathematisch positive Richtung, also gegen den Uhrzeigersinn dreht. Mittelpunkt der Drehung ist der Koordinatenursprung.

Die Anschauung lehrt, dass d_α tatsächlich linear im Sinne obiger Definition ist. Und zwar geht eine Parallelogramm (das ja die Addition zweier Vektoren veranschaulicht) durch eine Drehung wieder in ein Parallelogramm über, ebenso geht ein um einen Faktor gestreckter Vektor in den um denselben Faktor gestreckten gedrehten Vektor über. Als Formel notiert bedeutet das für ein beliebiges $\mathbf{x} = (x_1, x_2) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$

$$d_\alpha(\mathbf{x}) = d_\alpha(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = x_1d_\alpha(\mathbf{e}_1) + x_2d_\alpha(\mathbf{e}_2).$$

Also ist d_α durch die Linearität auf ganz \mathbb{R}^2 festgelegt, sofern wir nur seine Wirkung auf die Basisvektoren \mathbf{e}_1 und \mathbf{e}_2 kennen. Die ist aber klar, denn offenbar gilt

$$d_\alpha(\mathbf{e}_1) = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \quad \text{und} \quad d_\alpha(\mathbf{e}_2) = \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}.$$

Aufgrund der Linearität genügen also diese Daten, um die Abbildung d_α eindeutig festzulegen. Die Darstellung

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

als sogenannte 2×2 -Matrix mit den beiden Vektoren $d_\alpha(\mathbf{e}_1)$ und $d_\alpha(\mathbf{e}_2)$ als Spalten enthält somit alle relevante Information über die lineare Abbildung d_α . Entsprechendes gilt für eine beliebige lineare Abbildung f . Die Spalten stellen die Bilder der (kanonischen) Basisvektoren unter f dar und bestimmen f vermittels Linearität eindeutig. In der Linearen Algebra wird diese Beobachtung wesentlich vertieft werden.

Doch zurück zu den Drehungen. In unseren Überlegungen können wir den Winkel α durch einen anderen Winkel β ersetzen und die Zusammensetzung (Hintereinanderausführung) $d_\beta \circ d_\alpha = d_{\alpha+\beta}$, also die Drehung um den Winkel $\alpha + \beta$, betrachten.

Für die Bildvektoren $\mathbf{y}_i := d_{\alpha+\beta}(\mathbf{e}_i) = d_\beta(\mathbf{x}_i)$ mit $\mathbf{x}_i := d_\alpha(\mathbf{e}_i)$, $i = 1, 2$, erhalten wir einerseits

$$\mathbf{y}_1 = \begin{pmatrix} \cos(\alpha + \beta) \\ \sin(\alpha + \beta) \end{pmatrix} \quad \text{und} \quad \mathbf{y}_2 = \begin{pmatrix} -\sin(\alpha + \beta) \\ \cos(\alpha + \beta) \end{pmatrix},$$

andererseits durch iteriertes Ausrechnen vermittels der Formeln für d_α und d_β :

$$\begin{aligned} \mathbf{y}_1 &= d_\beta(d_\alpha(\mathbf{e}_1)) = d_\beta \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} = d_\beta(\cos \alpha \mathbf{e}_1 + \sin \alpha \mathbf{e}_2) = \\ &= (\cos \alpha) d_\beta(\mathbf{e}_1) + (\sin \alpha) d_\beta(\mathbf{e}_2) = \cos \alpha \begin{pmatrix} \cos \beta \\ \sin \beta \end{pmatrix} + \sin \alpha \begin{pmatrix} -\sin \beta \\ \cos \beta \end{pmatrix} = \\ &= (\cos \alpha \cos \beta - \sin \alpha \sin \beta) \mathbf{e}_1 + (\cos \alpha \sin \beta + \sin \alpha \cos \beta) \mathbf{e}_2 \end{aligned}$$

und ganz analog

$$\mathbf{y}_2 = d_\beta(d_\alpha(\mathbf{e}_2)) = -(\sin \alpha \cos \beta + \cos \alpha \sin \beta)\mathbf{e}_1 + (\cos \alpha \cos \beta - \sin \alpha \sin \beta)\mathbf{e}_2$$

(diese Rechenschritte entsprechen eigentlich der Berechnung eines Matrizenproduktes, das wir im Kapitel über Lineare Algebra allgemein besprechen werden), also

$$\mathbf{y}_1 = d_\beta(d_\alpha(\mathbf{e}_1)) = \begin{pmatrix} \cos \alpha \cos \beta - \sin \alpha \sin \beta \\ \sin \alpha \cos \beta + \cos \alpha \sin \beta \end{pmatrix}$$

und

$$\mathbf{y}_2 = d_\beta(d_\alpha(\mathbf{e}_2)) = \begin{pmatrix} -\sin \alpha \cos \beta - \cos \alpha \sin \beta \\ \cos \alpha \cos \beta - \sin \alpha \sin \beta \end{pmatrix}.$$

Vergleich der Koordinaten in den beiden gewonnenen Darstellungen von \mathbf{y}_1 und \mathbf{y}_2 zeigt uns die **Additionstheoreme** für Cosinus und Sinus, die wir im Zusammenhang mit den Polarkoordinaten komplexer Zahlen schon verwendet haben:

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$$

und

$$\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta.$$

Übungsaufgabe 92. (P) Überprüfen Sie die Additionstheoreme rechnerisch, aber ohne Taschenrechner für die Winkel $\alpha = \beta = \frac{\pi}{4}$.

Übungsaufgabe 93. (T) Drehen Sie das Dreieck mit den Eckpunkten $(-1/1)$, $(0/2)$ und $(1/2)$ um 45° im Uhrzeigersinn (mathematisch negativ).

Übungsaufgabe 94. (P) Durch die Gleichung $5x^2 + 6xy + 5y^2 = 1$ wird eine Ellipse festgelegt. Bestimmen Sie die Gleichung der um 45° gegen den Uhrzeigersinn (mathematisch positiv) gedrehten Ellipse indem Sie die in Abschnitt 1.5.4 untersuchte Transformation $x \mapsto x \cos \alpha + y \sin \alpha$ und $y \mapsto -x \sin \alpha + y \cos \alpha$ durchführen.

1.5.5 Inneres (Skalar-) Produkt und Längenmessung

Inhalt in Kurzfassung: Das innere Produkt oder auch Skalarprodukt zweier Vektoren ist eine bestimmte Zahl. Diese ergibt sich linear aus beiden Komponenten. Auf den kanonischen Basisvektoren nimmt das Skalarprodukt den Wert 0 (für verschiedene) oder 1 (für zweimal denselben kanonischen Basisvektor) an. Durch diese definierenden Eigenschaften ist das Skalarprodukt sogar eindeutig bestimmt. Nahegelegt wird seine Definition durch die geometrische Deutung: Länge des einen Vektors mal Länge der Projektion des anderen auf den ersten. Das Skalarprodukt eines Vektors mit sich selbst liefert somit das Quadrat seiner Länge. Damit wird mit Hilfe des Skalarproduktes auch Längenmessung möglich. Zwei Vektoren stehen genau dann normal aufeinander (sie heißen dann orthogonal), wenn ihr Skalarprodukt = 0 ist. Mit Hilfe der Cosinusfunktion und des Skalarproduktes lassen sich auch beliebige Winkel messen.

Man kann das (gewöhnliche) **Skalarprodukt** oder auch **innere Produkt** $\mathbf{x}\mathbf{y}$ der Vektoren $\mathbf{x} = (x_1, x_2, \dots, x_n)$ und $\mathbf{y} = (y_1, y_2, \dots, y_n)$ im Standardvektorraum \mathbb{R}^n durch die Definition

$$\mathbf{x}\mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \sum_{i=1}^n x_i y_i$$

einführen.

Alternativ wollen wir hier von der geometrischen Interpretation ausgehen, der Einfachheit halber zunächst für $n = 2$, also für Vektoren in der Ebene. Und zwar lässt sich $\mathbf{x}\mathbf{y}$ interpretieren als das Produkt der Länge von \mathbf{x} , für die wir $\|\mathbf{x}\|$ schreiben, mit der Länge der Orthogonalprojektion von \mathbf{y} auf \mathbf{x} . Ist α der Winkel zwischen den beiden Vektoren, so bedeutet das

$$\mathbf{x}\mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \alpha.$$

Gewisse Spezialfälle sind besonders wichtig: Für $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$ ist $\mathbf{x}\mathbf{y} = 0$ genau dann, wenn $\cos \alpha = 0$, also wenn \mathbf{x} und \mathbf{y} **orthogonal** oder auch **normal** aufeinander stehen. Der Nullvektor steht definitionsgemäß auf jeden anderen Vektor normal. Der andere Extremfall, nämlich $\cos \alpha = 1$ liegt vor, wenn $\mathbf{x} = \mathbf{y}$. Dann ist $\mathbf{x}\mathbf{x} = \|\mathbf{x}\|^2$ bzw. $\|\mathbf{x}\| = \sqrt{\mathbf{x}\mathbf{x}}$.

Vertauschen wir die Rollen von \mathbf{x} und \mathbf{y} , ist der Winkel in die umgekehrte Richtung zu durchlaufen, also α durch $-\alpha$ zu ersetzen. Wegen $\cos(-\alpha) = \cos \alpha$ folgt daraus

$$(i) \quad \mathbf{x}\mathbf{y} = \mathbf{y}\mathbf{x}.$$

Klarerweise gilt außerdem

$$(ii) \quad (r\mathbf{x})\mathbf{y} = r(\mathbf{x}\mathbf{y})$$

und, wie die geometrische Anschauung lehrt,

$$(iii) \quad \mathbf{x}(\mathbf{y}_1 + \mathbf{y}_2) = \mathbf{x}\mathbf{y}_1 + \mathbf{x}\mathbf{y}_2.$$

Man beachte, dass (ii) und (iii) analog sind zu den Linearitätsbedingungen von vorher, genauer: Die Abbildungen $\mathbf{x} \mapsto \mathbf{x}\mathbf{y}$ ist für festes \mathbf{y} ebenso linear wie die Abbildungen $\mathbf{y} \mapsto \mathbf{x}\mathbf{y}$ für festes \mathbf{x} . Weil das Skalarprodukt von zwei Komponenten abhängt, spricht man deshalb auch von **Bilinearität**. Aufgrund dieser Eigenschaft lassen sich beliebige Werte des Skalarproduktes sehr leicht ausrechnen, wenn man sie für die Basisvektoren \mathbf{e}_i kennt. Und diese ergeben sich in offensichtlicher Weise aus der geometrischen Anschauung: $\mathbf{e}_1\mathbf{e}_1 = \mathbf{e}_2\mathbf{e}_2 = 1$ (Länge 1) und $\mathbf{e}_1\mathbf{e}_2 = \mathbf{e}_2\mathbf{e}_1 = 0$ (orthogonal).

Für beliebige Vektoren

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 \quad \text{und} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2$$

aus \mathbb{R}^2 ergibt sich zunächst unter Verwendung von (iii)

$$(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \cdot (y_1\mathbf{e}_1 + y_2\mathbf{e}_2) \stackrel{\text{(iii)}}{=} (x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \cdot (y_1\mathbf{e}_1) + (x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \cdot (y_2\mathbf{e}_2).$$

Für den ersten dieser beiden Summanden erhält man

$$\begin{aligned} (x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \cdot (y_1\mathbf{e}_1) &\stackrel{\text{(i)}}{=} (y_1\mathbf{e}_1) \cdot (x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \stackrel{\text{(ii)}}{=} y_1(\mathbf{e}_1 \cdot (x_1\mathbf{e}_1 + x_2\mathbf{e}_2)) \stackrel{\text{(iii)}}{=} \\ &= y_1(\mathbf{e}_1 \cdot (x_1\mathbf{e}_1) + \mathbf{e}_1 \cdot (x_2\mathbf{e}_2)) \stackrel{\text{(i)}}{=} y_1((x_1\mathbf{e}_1) \cdot \mathbf{e}_1 + (x_2\mathbf{e}_2) \cdot \mathbf{e}_1) \stackrel{\text{(ii)}}{=} \\ &= y_1x_1(\mathbf{e}_1\mathbf{e}_1) + y_1x_2(\mathbf{e}_2\mathbf{e}_1) = x_1y_1 \cdot 1 + y_1x_2 \cdot 0 = x_1y_1. \end{aligned}$$

Für den zweiten Summanden erhält man analog

$$(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \cdot y_2\mathbf{e}_2 = x_2y_2.$$

Zusammengefasst liefert das tatsächlich

$$\mathbf{xy} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1y_1 + x_2y_2,$$

was mit der eingangs erwähnten Definition des Skalarproduktes im Fall $n = 2$ übereinstimmt. Auch für beliebiges $n \in \mathbb{N}$ führen ganz analoge (wenn auch nicht immer so anschauliche) Überlegungen zu der entsprechenden Beziehung, so dass wir auch dort die eingangs gegebene Definition übernehmen können. Die sich daraus ergebende Länge

$$\|(x_1, x_2, \dots, x_n)\| := \sqrt{\sum_{i=1}^n x_i^2}$$

heißt auch die **euklidische Länge** oder **Norm** des Vektors $\mathbf{x} = (x_1, x_2, \dots, x_n)$. Man beachte den Spezialfall $n = 1$, wo mit $\mathbf{x} = x_1 = x$, $\mathbf{y} = y_1 = y$ das Skalarprodukt $\mathbf{xy} = xy$ mit dem gewöhnlichen Produkt und die Norm $\|\mathbf{x}\| = \sqrt{xx} = |x|$ mit dem Betrag zusammenfällt.

Die Bedingungen (i), (ii) und (iii) wurden in der obigen Rechnung mehrmals verwendet. Wichtig ist auch die bereits weiter oben implizit aufgetretene Beobachtung

$$\text{(iv) } \mathbf{xx} \geq 0,$$

wobei Gleichheit nur für $\mathbf{x} = \mathbf{0}$ gilt. Aus (i) bis (iv) allein lassen sich viele weitere vertraut anmutende Rechenregeln wie etwa $\mathbf{x0} = 0$ ableiten.

Übungsaufgabe 95. (P) Leiten Sie das Gesetz $\mathbf{x0} = 0$ alleine mit Hilfe der Gesetze (i) bis (iii) für Skalarprodukte her.

Übungsaufgabe 96. (P) Zeigen Sie für das rechtwinkelige Dreieck mit den Eckpunkten $A = (0/0/0)$, $B = (b/0/0)$, $C = (b/c/0)$ ($b, c > 0$) durch Nachrechnen, dass der Cosinus des von $\mathbf{x} := \overline{AB}$ (dem Vektor von A nach B) und

$\mathbf{y} := \overline{AC}$ (dem Vektor von A nach C) eingeschlossenen Winkels α tatsächlich durch

$$\frac{\mathbf{xy}}{\sqrt{\mathbf{xx}}\sqrt{\mathbf{yy}}}$$

gegeben ist.

Übungsaufgabe 97. (T) Berechnen Sie alle Winkel (Angabe des Cosinus genügt) sowie den Flächeninhalt des Dreiecks mit den Eckpunkten A, B, C :

$$A = (1/0/0) \quad B = (0/1/0) \quad C = (0/0/1).$$

1.5.6 Zwei wichtige Ungleichungen

Inhalt in Kurzfassung: Die Cauchy-Schwarzsche Ungleichung $|\mathbf{xy}| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$ gilt ebenso wie die Dreiecksungleichung $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ in beliebigen Vektorräumen mit Skalarprodukt. Definiert man damit einen Abstand (eine sogenannte Metrik) d durch $d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$, so nimmt die Dreiecksungleichung die Gestalt $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ an.

Weil alle bisherigen Überlegungen zum Skalarprodukt im Wesentlichen nur auf den Eigenschaften (i) bis (iv) aus 1.5.5 beruhen, verwendet man diese zur Definition eines allgemeinen Skalarproduktes in einem Vektorraum. Zwar brauchen wir diesen allgemeineren Begriff jetzt noch nicht unbedingt. Jedoch spielen die Eigenschaften (i) bis (iv) eine wichtige Rolle bei der Herleitung der **Cauchy-Schwarzschen Ungleichung**:

$$|\mathbf{xy}| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

Über die geometrische Definition des Skalarproduktes als $\mathbf{xy} = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cos \alpha$ ist sie klar, weil ja $|\cos \alpha| \leq 1$ für alle Winkel α gilt. Bemerkenswert ist, dass diese wichtige Ungleichung alleine aus den Eigenschaften (i) bis (iv) folgt. Davon kann man oft Gebrauch machen. Deshalb bringen wir auch diesen Beweis:

Es genügt, den nichttrivialen Fall $\mathbf{y} \neq \mathbf{0}$ zu untersuchen (für $\mathbf{y} = \mathbf{0}$ sind beide Seiten der Ungleichung $= 0$). Man betrachtet den Vektor $\mathbf{z} = \mathbf{x} + r\mathbf{y}$, eine Linearkombination von \mathbf{x} und \mathbf{y} , wobei für den Skalar r zweckmäßigerweise $r := -\frac{\mathbf{xy}}{\|\mathbf{y}\|^2}$ gesetzt wird. Gesetz (iv) aus 1.5.5 garantiert $0 \leq \|\mathbf{z}\|^2 = (\mathbf{x} + r\mathbf{y})(\mathbf{x} + r\mathbf{y})$. Den Term rechts multipliziert man mittels (i)-(iii) aus und erhält dafür nach kurzer Rechnung $0 \leq \|\mathbf{x}\|^2 - \frac{(\mathbf{xy})^2}{\|\mathbf{y}\|^2}$. Bringt man den negativen Term auf die linke Seite der Ungleichung und multipliziert mit $\|\mathbf{y}\|^2 > 0$, so erhält man $(\mathbf{xy})^2 \leq \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$ und nach Ziehen der Quadratwurzel auf beiden Seiten die behauptete Cauchy-Schwarzsche Ungleichung.

Rekapituliert man den Beweis, so kann man die recht nützliche Erkenntnis gewinnen, dass statt \leq sogar Gleichheit $=$ gesetzt werden kann, wenn $\mathbf{z} = \mathbf{0}$ gilt, was bedeutet, dass \mathbf{x} ein Vielfaches von \mathbf{y} ist.

Übungsaufgabe 98. (P) Deuten Sie für gegebene Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ die linke und die rechte Seite in der Cauchy-Schwarzschen Ungleichung jeweils als Fläche und begründen Sie geometrisch, warum zwischen ihnen die Beziehung \leq gilt.

Eine Folgerung aus der Ungleichung von Cauchy-Schwarz ist die mindestens ebenso wichtige **Dreiecksungleichung**, die ebenfalls in beliebigen Vektorräumen mit einem Skalarprodukt gilt:

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

Der Name rührt daher, dass man die Vektoren \mathbf{x} , \mathbf{y} und $\mathbf{x} + \mathbf{y}$ als die Seiten eines Dreiecks interpretieren kann. Dann besagt die Ungleichung: Geht man die Seite $\mathbf{x} + \mathbf{y}$ auf direktem Wege entlang, wird der Weg nie länger sein, als über den Umweg (zuerst \mathbf{x} und dann \mathbf{y}).

Der Beweis der Dreiecksungleichung mittels Cauchy-Schwarz ist kurz:

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y})(\mathbf{x} + \mathbf{y}) = \|\mathbf{x}\|^2 + 2\mathbf{x}\mathbf{y} + \|\mathbf{y}\|^2 \leq \\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

Ziehen der Quadratwurzel auf beiden Seiten liefert die Dreiecksungleichung.

Der Abstand (die Distanz) $d(\mathbf{x}, \mathbf{y})$ zweier Punkte \mathbf{x} und \mathbf{y} ist definiert durch

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|.$$

(Liegt das gewöhnliche Skalarprodukt zugrunde, spricht man vom **euklidischen Abstand**. Die Dreiecksungleichung hat dann die Gestalt $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$, d.h. der direkte Weg von \mathbf{x} nach \mathbf{z} ist nie länger als der Umweg über die Zwischenstation \mathbf{y} . Beweis:

$$d(\mathbf{x}, \mathbf{z}) = \|\mathbf{x} - \mathbf{z}\| = \|\mathbf{x} - \mathbf{y} + \mathbf{y} - \mathbf{z}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\| = d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}),$$

Klarerweise gilt auch stets $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$, $d(\mathbf{x}, \mathbf{y}) \geq 0$, $d(\mathbf{x}, \mathbf{y}) = 0$ nur für $\mathbf{x} = \mathbf{y}$. Diese Eigenschaften machen d zu einer sogenannten **Metrik**, in diesem Fall spricht man von der **euklidischen Metrik**. Im Spezialfall $n = 1$, d.h. $\mathbf{x} = x$, $\mathbf{y} = y \in \mathbb{R}$ ist $d(\mathbf{x}, \mathbf{y}) = d(x, y) = |x - y|$.

Allgemein nennt man eine Menge X zusammen mit einer Metrik d auf X , einen **metrischen Raum**, wenn $d : X \times X \rightarrow \mathbb{R}$ die entsprechenden Eigenschaften hat: $d(x, y) \geq 0$ (**Nichtnegativität**), $d(x, y) = 0$ nur für $x = y$ (**Positivität**), $d(x, y) = d(y, x)$ (**Symmetrie**) und $d(x, z) \leq d(x, y) + d(y, z)$ (**Dreiecksungleichung**) für alle $x, y, z \in X$.

Nicht sehr überraschend kann man auch bei der Dreiecksungleichung analysieren, wann Gleichheit gilt. Das ist genau dann der Fall, wenn $\mathbf{x}\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|$ gilt, und das gemäß unseren Überlegungen zu Cauchy-Schwarz wiederum genau dann, wenn einer der beiden Vektoren ein Vielfaches des anderen ist und zusätzlich (das Vorzeichen links könnte auch negativ sein) in dieselbe Richtung weist.

Noch eine weitere Erkenntnis können wir aus demselben Beweis ziehen, nämlich den **Satz des Pythagoras**: Die Gleichheit $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ gilt genau dann, wenn $\mathbf{x}\mathbf{y} = 0$, wenn also \mathbf{x} und \mathbf{y} einen rechten Winkel einschließen.

Diese Beziehungen ermöglichen es, geometrisch motivierte Begriffe wie Orthogonalität, Länge, Winkel etc. und auch den Satz von Pythagoras auf abstrakte Räume zu übertragen, sofern dort nur ein Skalarprodukt mit den Eigenschaften (i)-(iv) aus 1.5.5 definiert ist.

1.5.7 Das äußere (Vektor-) Produkt im \mathbb{R}^3

Inhalt in Kurzfassung: Das äußere Produkt, genannt auch Vektorprodukt von zwei dreidimensionalen Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ ist wieder ein Vektor in \mathbb{R}^3 und wird mit $\mathbf{x} \times \mathbf{y}$ bezeichnet. Dieser Vektor ist durch folgende drei Bestimmungsstücke festgelegt: (i) $\mathbf{x} \times \mathbf{y}$ steht normal sowohl auf \mathbf{x} als auch auf \mathbf{y} ; (ii) die Länge von $\mathbf{x} \times \mathbf{y}$ ist gleich der Maßzahl der Fläche des von \mathbf{x} und \mathbf{y} in \mathbb{R}^3 aufgespannten Parallelogramms; (iii) die Orientierung von $\mathbf{x} \times \mathbf{y}$ ergibt sich aus der „Rechtsschraubenregel“ oder abstrakt, mit Hilfe des Begriffs der Determinante aus der Linearen Algebra (Mathematik 2) aus der Forderung $\det(\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y}) \geq 0$ an die Determinante.

Im \mathbb{R}^3 spielt neben den bisher behandelten Operationen mit Vektoren (Addition, Skalarmultiplikation, Skalarprodukt) noch eine weitere wichtige Rolle: das sogenannte **äußere** oder auch **Vektor-** oder **Kreuzprodukt** zweier Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, für das man auch $\mathbf{x} \times \mathbf{y}$ schreibt. Es liefert wieder einen Vektor im \mathbb{R}^3 , kann daher als Abbildung $\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ (Achtung, hier bezeichnet das erste Symbol \times das Vektorprodukt, das zweite das kartesische Produkt zweier Mengen), $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \times \mathbf{y}$, geschrieben werden.

Der Vektor $\mathbf{x} \times \mathbf{y}$ ist geometrisch durch folgende Eigenschaften eindeutig festgelegt:

1. Das Vektorprodukt $\mathbf{x} \times \mathbf{y}$ steht normal sowohl auf \mathbf{x} als auch auf \mathbf{y} .
2. Die (eindimensionale Maßzahl der) Länge $\|\mathbf{x} \times \mathbf{y}\|$ des Vektorproduktes $\mathbf{x} \times \mathbf{y}$ ist gleich der (zweidimensionalen Maßzahl der) Fläche des von \mathbf{x} und \mathbf{y} aufgespannten Parallelogramms.
3. Die Orientierung von $\mathbf{x} \times \mathbf{y}$ ist so gewählt, dass das Vorzeichen der durch die Vektoren \mathbf{x}, \mathbf{y} und $\mathbf{x} \times \mathbf{y}$ (in dieser Reihenfolge) gebildeten Determinante (siehe Kapitel Lineare Algebra in Mathematik 2) nicht negativ ist.

Man stellt sich das Vektorprodukt geometrisch also am besten so vor: Zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ im Anschauungsraum spannen ein Parallelogramm auf. Der Vektor $\mathbf{x} \times \mathbf{y}$ steht normal auf das Parallelogramm und seine Länge ist zahlenmäßig gleich der Fläche des Parallelogramms. Damit ist $\mathbf{x} \times \mathbf{y}$ bis auf die Orientierung eindeutig bestimmt. Diese dreht sich um, wenn \mathbf{x} und \mathbf{y} vertauscht werden. Im realen dreidimensionalen Raum kann man die Orientierung des Vektorproduktes über die Orientierung der Koordinatenachsen festlegen. Oft wird das über eine sogenannte Rechtsschraubenregel getan. Dabei stellt man sich vor, dass eine Schraube durch die Drehung des Vektors \mathbf{x} zum Vektor \mathbf{y} sich in die Richtung von $\mathbf{z} := \mathbf{x} \times \mathbf{y}$ bewegt, so wie das der Fall sein sollte, wenn man $\mathbf{x} = \mathbf{e}_1, \mathbf{y} = \mathbf{e}_2$ und $\mathbf{z} = \mathbf{e}_3$ (Einheitsvektor in Richtung von x -, y - und z -Achse) setzt.

Es zeigt sich, dass für $\mathbf{x} = (x_1, x_2, x_3)$ und $\mathbf{y} = (y_1, y_2, y_3)$ die Definition

$$\mathbf{x} \times \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$$

genau die behaupteten Eigenschaften hat. Man beachte, dass das Vektorprodukt linear von beiden Komponenten abhängt. Damit sind folgende Beziehungen gemeint:

$$\mathbf{x} \times (\mathbf{y}_1 + \mathbf{y}_2) = \mathbf{x} \times \mathbf{y}_1 + \mathbf{x} \times \mathbf{y}_2 \quad \text{und} \quad \mathbf{x} \times (r\mathbf{y}) = r(\mathbf{x} \times \mathbf{y})$$

und analog

$$(\mathbf{x}_1 + \mathbf{x}_2) \times \mathbf{y} = \mathbf{x}_1 \times \mathbf{y} + \mathbf{x}_2 \times \mathbf{y} \quad \text{und} \quad (r\mathbf{x}) \times \mathbf{y} = r(\mathbf{x} \times \mathbf{y})$$

für alle $\mathbf{x}, \mathbf{x}_i, \mathbf{y}, \mathbf{y}_i \in \mathbb{R}^3$ und $r \in \mathbb{R}$. Überdies gilt $\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$, $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$ und $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$.

Übungsaufgabe 99. (P) Überprüfen Sie die obigen Beziehungen, beginnend mit der Linearität.

Die Normalitätseigenschaft des Vektorproduktes lässt sich leicht überprüfen:

Übungsaufgabe 100. (P) Rechnen Sie nach, dass der Vektor $\mathbf{x} \times \mathbf{y}$ sowohl auf \mathbf{x} als auch auf \mathbf{y} normal steht.

Für die anderen beiden hilft es sehr, wenn man schon den Begriff der Determinante zur Verfügung hat, was aber erst im zweiten Semester in der Linearen Algebra der Fall sein wird. Ein bereits an dieser Stelle verständliche Deutung ergibt sich aus der folgenden Übungsaufgabe:

Übungsaufgabe 101. (P) Verwenden Sie die Formel $\text{Volumen} = \text{Grundfläche} \times \text{Höhe}$ um nachzuweisen, dass das Volumen des Parallelepipeds („verzerrter Würfel“) mit Seiten $\mathbf{x}, \mathbf{y}, \mathbf{z}$ durch das sogenannte Spatprodukt $|\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z})|$ gegeben ist. Wie kann man daraus durch eine geometrische Überlegung folgern, dass

$$|\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z})| = |\mathbf{y} \cdot (\mathbf{z} \times \mathbf{x})| = |\mathbf{z} \cdot (\mathbf{x} \times \mathbf{y})|$$

gilt? (Bemerkung: Diese Gleichungen würden auch nach Weglassen der Absolutbeträge gelten, jedoch falsch werden, wenn man dann an einer Stelle die Reihenfolge im Vektorprodukt vertauschte.)

Etwas konkreter lässt sich die zweite Eigenschaft in folgendem Spezialfällen überprüfen:

Übungsaufgabe 102. (P) Zeigen Sie für das rechtwinkelige Dreieck $A = (0/0/0)$, $B = (b/0/0)$, $C = (b/c/0)$ durch Nachrechnen, dass der Flächeninhalt tatsächlich durch den Betrag des Vektorproduktes $\frac{1}{2} \overline{AB} \times \overline{AC}$ gegeben ist.

Übungsaufgabe 103. (T) Geben Sie sich zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ vor, die weder normal noch parallel zueinander stehen und auch nicht parallel zu einer der drei Koordinatenebenen sind. Bestimmen Sie $\mathbf{x} \times \mathbf{y}$ rechnerisch und fertigen Sie eine Skizze (zum Beispiel im Schrägriss) an.

Übungsaufgabe 104. (E) In Anwendungen tritt gelegentlich das doppelte Vektorprodukt $\mathbf{x} \times (\mathbf{y} \times \mathbf{z})$ auf. Dieses Produkt ist offensichtlich der Nullvektor, falls \mathbf{y} und \mathbf{z} parallel sind. Wenn \mathbf{y} und \mathbf{z} nicht parallel sind, so bilden \mathbf{y}, \mathbf{z} und $\mathbf{y} \times \mathbf{z}$ eine Basis (siehe 1.5.3). Die Vektoren \mathbf{y}, \mathbf{z} spannen eine Ebene auf und $\mathbf{y} \times \mathbf{z}$ steht normal auf diese Ebene. Somit können wir schreiben

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) = \alpha_1 \mathbf{y} + \alpha_2 \mathbf{z} + \alpha_3 (\mathbf{y} \times \mathbf{z})$$

mit noch zu bestimmenden Koeffizienten $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. Der Vektor $\mathbf{x} \times (\mathbf{y} \times \mathbf{z})$ steht normal auf $\mathbf{y} \times \mathbf{z}$ und liegt daher in der von \mathbf{y} und \mathbf{z} aufgespannten Ebene. Der Koeffizient α_3 beim dritten Basisvektor $\mathbf{y} \times \mathbf{z}$ ist also $= 0$. Wir bilden nun das Skalarprodukt mit \mathbf{x}

$$0 = \mathbf{x}(\mathbf{x} \times (\mathbf{y} \times \mathbf{z})) = \alpha_1 \mathbf{x}\mathbf{y} + \alpha_2 \mathbf{x}\mathbf{z}$$

und erhalten $\alpha_1 = \lambda \mathbf{x}\mathbf{z}, \alpha_2 = -\lambda \mathbf{x}\mathbf{y}$ mit nun nur mehr einem Parameter λ :

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) = \lambda ((\mathbf{x}\mathbf{z})\mathbf{y} - (\mathbf{x}\mathbf{y})\mathbf{z}).$$

Berechnen Sie durch direktes Einsetzen in die Definition des Vektorprodukts den Wert des Parameters λ in dieser Formel.

1.5.8 Interessante Teilmengen in \mathbb{R}^2 und \mathbb{R}^3

Inhalt in Kurzfassung: Mit Hilfe der bisher erarbeiteten Begriffe finden wir verschiedene Darstellungen (Koordinatenform, Parameterform, Normalvektorform) zahlreicher geometrischer Gebilde. Zuerst behandeln wir Objekte in der Ebene (Gerade, Rechteck, Kreis), sodann Objekte im dreidimensionalen Raum (Ebene, Quader, Zylinder, Kegel, Kugel). Praktische Parameterdarstellungen ergeben sich z.B. durch die an das entsprechende Objekt angepasste Zylinder- oder Kugelkoordinaten.

Mit Hilfe der Vektorrechnung lassen sich viele wichtige geometrische Gebilde hervorragend beschreiben. Hier einige wichtige Beispiele:

Geraden g im \mathbb{R}^n lassen sich am anschaulichsten in **Parameterform** darstellen:

$$g = \{\mathbf{x}_0 + r\mathbf{x}_1 : r \in \mathbb{R}\} \quad \text{mit} \quad \mathbf{x}_0, \mathbf{x}_1 \in \mathbb{R}^n$$

Dabei ist \mathbf{x}_0 irgendein Punkt auf der Geraden (Ortsvektor) und $\mathbf{x}_1 \neq \mathbf{0}$ (Richtungsvektor) ein Vektor, der in die Richtung von g weist. Wenn der Parameter r ganz \mathbb{R} durchläuft, durchläuft $\mathbf{x}_0 + r\mathbf{x}_1$ die gesamte Gerade $g \subseteq \mathbb{R}^n$.

In der Ebene \mathbb{R}^2 lassen sich Geraden auch in der **(Koordinaten-) Form**

$$g = \{(x, y) : ax + by = c\}$$

darstellen, wobei a und b nicht beide 0 sein dürfen. Ist $a = 0$ und $b \neq 0$, so ist g parallel zur x -Achse, im umgekehrten Fall zur y -Achse. Ist $b \neq 0$ lässt sich die Gleichung $ax + by = c$ umformen zu $y = -\frac{a}{b}x + \frac{c}{b}$. Die Steigung k dieser

Geraden ist also gegeben durch $k = -\frac{a}{b}$. Aus der Schule sollte bekannt sein, wie man Parameter- und Koordinatenform für Geraden ineinander überführen kann.

Man kann die Koordinatenform $ax + by = c$ auch umschreiben zu

$$g = \{\mathbf{x} : \mathbf{n}\mathbf{x} = c\}.$$

Darin ist $\mathbf{n}\mathbf{x}$ das Skalarprodukt des festen Vektors $\mathbf{n} = (a, b)$ mit dem variablen Vektor $\mathbf{x} = (x, y)$. Die Forderung, dass dieses Skalarprodukt für alle $\mathbf{x} \in g$ den konstanten Wert c liefert, bedeutet für je zwei $\mathbf{x}_1, \mathbf{x}_2 \in g$ dass $\mathbf{n}(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{n}\mathbf{x}_1 - \mathbf{n}\mathbf{x}_2 = c - c = 0$ gilt, dass also die Verbindungslinie zwischen \mathbf{x}_1 und \mathbf{x}_2 normal steht auf \mathbf{n} . Mit anderen Worten: Der Vektor \mathbf{n} steht normal auf g , weshalb man auch von einer **Normalvektorform** für g spricht. Liegt \mathbf{x} nicht auf g , so ist $\mathbf{n}\mathbf{x} - c \neq 0$, wobei der Wert auf der linken Seite proportional zum Normalabstand d von \mathbf{x} auf g ist. Die Proportionalität wird (eventuell bis auf das Vorzeichen) zur Gleichheit, wenn der Normalvektor \mathbf{n} auf $\|\mathbf{n}\| = 1$ normiert ist. Normierung lässt sich erreichen, indem man \mathbf{n} und c durch $\mathbf{n}' := \frac{\mathbf{n}}{\|\mathbf{n}\|}$ bzw. $c' := \frac{c}{\|\mathbf{n}\|}$ ersetzt. Die resultierende Formel

$$d = \frac{|\mathbf{n}\mathbf{x} - c|}{\|\mathbf{n}\|}$$

heißt **Hessesche Abstandsformel**.

In der Ebene \mathbb{R}^2 lassen sich achsenparallele **Rechtecke** R sehr einfach mittels Koordinaten anschreiben:

$$R = \{(x, y) \in \mathbb{R}^2 : a_1 \leq x \leq a_2, b_1 \leq y \leq b_2\}$$

mit geeigneten Begrenzungskordinaten $a_1 < a_2$ und $b_1 < b_2$.

Ein **Kreis** k in der Ebene besteht aus jenen $\mathbf{x} = (x, y) \in \mathbb{R}^2$, deren Abstand zu einem Mittelpunkt $\mathbf{x}_0 = (x_0, y_0)$ den festen Abstand $d(\mathbf{x}, \mathbf{x}_0) = \|\mathbf{x} - \mathbf{x}_0\| = r_0 \geq 0$ (Radius) haben, also

$$k = \{\mathbf{x} = (x, y) \in \mathbb{R}^2 : \|\mathbf{x} - \mathbf{x}_0\| = r_0\} = \{(x, y) \in \mathbb{R}^2 : (x - x_0)^2 + (y - y_0)^2 = r_0^2\}.$$

Dem Kreis besonders gut angepasst sind Polarkoordinaten, wie wir sie schon bei den komplexen Zahlen kennengelernt haben. Dabei stellt man die Koordinaten eines Punktes (x, y) dar als $x = r \cos \alpha$ und $y = r \sin \alpha$, wobei $r = \sqrt{x^2 + y^2} \geq 0$ der Abstand des Punktes vom Ursprung ist und $\alpha \in [0, 2\pi)$ jener Winkel, um den die x -Achse um den Ursprung bis zum Punkt (x, y) zu drehen ist. Bei vorgegebenem Radius $r_0 \geq 0$ und Mittelpunkt im Ursprung nimmt die Kreisgleichung die besonders einfache Gestalt $r = r_0$ an. Dabei ist r_0 als feste Zahl zu verstehen, nämlich der Radius des Kreises, und r als eine der beiden Variablen in der Polardarstellung eines Punktes

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \alpha \\ r \sin \alpha \end{pmatrix} \in \mathbb{R}^2.$$

Der Kreis k hat daher auch die Parameterdarstellung

$$k = \left\{ \begin{pmatrix} r \cos \alpha \\ r \sin \alpha \end{pmatrix} : 0 \leq \alpha < 2\pi, r = r_0 \right\}.$$

Fassen wir \mathbb{R}^2 als komplexe Zahlenebene auf, so stehen mit Real- und Imaginärteil noch zusätzliche symbolische Mittel zur Verfügung, um verschiedene geometrische Gebilde zu beschreiben, vgl. dazu Übungsaufgabe 76.

Was über Geraden $g \subseteq \mathbb{R}^2$ gesagt wurde, gilt analog für **Ebenen** $\varepsilon \subseteq \mathbb{R}^3$. Die Parameterdarstellung hat die Gestalt

$$\varepsilon = \{ \mathbf{x}_0 + r\mathbf{x}_1 + s\mathbf{x}_2 : r, s \in \mathbb{R} \},$$

wobei $\mathbf{x}_0 \in \varepsilon$ irgendein Ortsvektor auf ε ist und $\mathbf{x}_1, \mathbf{x}_2$ zwei nicht parallele Richtungsvektoren, die folglich ε aufspannen. Es gilt auch eine Koordinatendarstellung

$$\varepsilon = \{ (x, y, z) \in \mathbb{R}^3 : ax + by + cz = d \}$$

mit $a, b, c, d \in \mathbb{R}$ derart, dass $\mathbf{n} := (a, b, c) \neq \mathbf{0}$ ein Normalvektor auf ε ist. Auch eine Hessesche Normalabstandsmessung zwischen Punkt und Ebene ist genauso möglich wie bei $n = 2$ zwischen Punkt und Gerade.

Übungsaufgabe 105. (T) Gegeben ist das Quadrat Q mit den Eckpunkten $(1/1)$, $(1/-1)$, $(-1/-1)$ und $(-1/1)$ sowie die Gerade g_α durch die Punkte $(0/0)$ und $(\cos \alpha / \sin \alpha)$.

Berechnen Sie die Abstände aller Eckpunkte von der Geraden g_α sowie die (orthogonale) Projektion des Quadrates Q auf die Gerade g_α .

Analog zu Rechtecken in der Ebene lassen sich im dreidimensionalen Raum \mathbb{R}^3 achsenparallele **Quader** Q mittels Koordinaten anschreiben:

$$Q = \{ (x, y, z) \in \mathbb{R}^3 : a_1 \leq x \leq a_2, b_1 \leq y \leq b_2, c_1 \leq z \leq c_2 \}$$

mit geeigneten Begrenzungskordinaten $a_1 < a_2$, $b_1 < b_2$ und $c_1 < c_2$.

In den folgenden Übungsaufgaben kommen die geometrischen Interpretationen von Skalar- und Vektorprodukt zum Einsatz. Sie dürfen verwenden, dass die Formel $\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \alpha$ auch für Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ im Raum, nicht nur in der Ebene, gilt, wenn α der von ihnen eingeschlossene Winkel ist.

Übungsaufgabe 106. (T) Berechnen Sie alle Winkel (Angabe des Cosinus genügt) sowie den Flächeninhalt des Dreiecks mit Eckpunkten A, B, C . Geben Sie außerdem die Gleichung jener Ebene an, welche A, B, C enthält:

$$A = (0/1/1) \quad B = (-1/0/1) \quad C = (1/1/0).$$

Übungsaufgabe 107. (T) Berechnen Sie den Winkel unter dem sich die beiden Ebenen $\varepsilon_1 : 2x + z = 2$ und $\varepsilon_2 : 3y + z = 2$ schneiden. (Unter dem Winkel zwischen zwei Ebenen im \mathbb{R}^3 versteht man den Winkel ihrer Normalvektoren.)

Übungsaufgabe 108. (T) Bestimmen Sie die Oberfläche und das Volumen des Tetraeders mit Eckpunkten: $A = (3/1/-1)$, $B = (1/0/1)$, $C = (-5/3/1)$ und $D = (-4/1/2)$. Hinweis: Das Volumen eines Kegels oder einer Pyramide ist ein Drittel des Volumens des Zylinders oder des Prismas mit derselben Grundfläche und Höhe. (Warum?)

Übungsaufgabe 109. (T) Bestimmen Sie für das Tetraeder aus Aufgabe 108 jeweils den Winkel, den zwei Seitenflächen miteinander einschließen.

Eine Kombination von Kreis und Quader stellt der achsenparallele **Zylinder** $Z \subseteq \mathbb{R}^3$ dar. Der Einfachheit halber wollen wir annehmen, dass der Basiskreis von Z in der Grundebene liegt, wo also für die dritte Koordinate $z = 0$ gilt. Der Mittelpunkt des Kreises sei $(x_0, y_0, 0)$, der Radius r_0 , die Höhe des Zylinders h_0 . Dann gilt:

$$Z = \{(x, y, z) : (x - x_0)^2 + (y - y_0)^2 \leq r_0^2, 0 \leq z \leq h_0\}.$$

Eine alternative Beschreibung lautet

$$Z = \left\{ \begin{pmatrix} x = r \cos \alpha \\ y = r \sin \alpha \\ z \end{pmatrix} : 0 \leq \alpha < 2\pi, 0 \leq r \leq r_0, 0 \leq z \leq h \right\}.$$

Ersetzt man r_0^2 durch $(r_0 - \frac{r_0}{h_0}z)^2$, so wird aus dem Zylinder ein **Kegel**. Doch zurück zum Zylinder: Übernimmt man die Polarkoordinaten vom Kreis und ergänzt sie durch die z -Koordinate, erhält man die **Zylinderkoordinaten** (r, α, z) , wobei weiterhin $x = r \cos \alpha$ und $y = r \sin \alpha$ gilt. Die Menge Z ist dann durch die beiden einfachen Beziehungen $r \leq r_0$ und $0 \leq z \leq h_0$ beschrieben.

Die **Vollkugel** $K \subseteq \mathbb{R}^3$ mit Mittelpunkt (x_0, y_0, z_0) und Radius $r_0 \geq 0$ ist gegeben durch

$$K = \{(x, y, z) \in \mathbb{R}^3 : (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 \leq r_0^2\}.$$

Man beachte, dass man auch kürzer schreiben kann

$$K = \{\mathbf{x} \in \mathbb{R}^3 : d(\mathbf{x}, \mathbf{x}_0) \leq r_0\},$$

wenn man den euklidischen Abstand d benutzt. Der Kugel besonders gut angepasst sind **Kugelkoordinaten** r, α, β mit $r \geq 0$, $0 \leq \alpha < 2\pi$ und $-\frac{\pi}{2} \leq \beta \leq \frac{\pi}{2}$. Mit ihnen lässt sich jeder Punkt $(x, y, z) \in \mathbb{R}^3$ in der Form $x = r \cos \alpha \cos \beta$, $y = r \sin \alpha \cos \beta$ und $z = r \sin \beta$ darstellen. Also:

$$K = \left\{ \begin{pmatrix} x = r \cos \alpha \cos \beta \\ y = r \sin \alpha \cos \beta \\ z = r \sin \beta \end{pmatrix} : 0 \leq \alpha < 2\pi, -\frac{\pi}{2} \leq \beta \leq \frac{\pi}{2}, 0 \leq r \leq r_0 \right\}.$$

Wieder ist r der Abstand zum Koordinatenursprung, α spielt dieselbe Rolle wie bei Kreis und Zylinder (auf einem Globus deutbar als geographische Länge), β ist der Winkel in der Vertikalen (geographische Breite). K ist dann – analog zum Kreis in der Ebene – schlicht durch die Beziehung $r \leq r_0$ beschrieben.

Übungsaufgabe 110. (P) Welche geometrischen Objekte werden durch

$$x_1(\varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} \quad \text{bzw.} \quad x_2(\varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ \varphi \end{pmatrix}$$

mit $\varphi \in [0, 2\pi]$ und festem $r > 0$ beschrieben?

Übungsaufgabe 111. (T) Geben Sie die Gleichung(en) des Kreises mit Mittelpunkt $M = (0/0/1)$ und Radius $r = 2$, welcher sich in der Ebene $z = 1$ befindet, jeweils in kartesischen Koordinaten, in Zylinderkoordinaten und in Kugelkoordinaten an.

Übungsaufgabe 112. (E) Nehmen Sie an, die Erde wäre eine ideale Kugel mit Radius $R = 6,371 \times 10^6$ m. Der Koordinatenursprung liege im Erdmittelpunkt, die x -Achse gehe durch den Schnittpunkt von Äquator und Nullmeridian, die y -Achse durch den Schnittpunkt von Äquator und Meridian 90° östliche Länge, die z -Achse durch den Nordpol. Mit \mathbf{x} sei der Radiusvektor vom Erdmittelpunkt zum Hörsaal Ihrer Übungsgruppe an der TU Wien bezeichnet. Geben Sie \mathbf{x} möglichst genau in Kugelkoordinaten und in kartesischen Koordinaten an.

1.5.9 Interessante Teilmengen in beliebigen metrischen Räumen

Inhalt in Kurzfassung: Wir präzisieren geometrisch anschauliche Begriffe wie Randpunkt, innerer Punkt und äußerer Punkt einer Menge im \mathbb{R}^n , um damit topologische Grundbegriffe wie z.B. Umgebung, offene, abgeschlossene Menge etc. präzise zu definieren.

Obwohl die folgenden (topologischen) Begriffe mit Hilfe einer beliebigen Metrik erklärt werden können, genügt es für unsere Zwecke, an den euklidischen Abstand d zu denken.

Ist $\mathbf{x}_0 \in \mathbb{R}^n$ und $\varepsilon > 0$ so nennt man die Menge

$$K(\mathbf{x}_0, \varepsilon) := \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{x}_0) < \varepsilon\}$$

die **offene ε -Kugel um \mathbf{x}_0** . Ist $M \subseteq \mathbb{R}^n$ irgendeine Teilmenge von \mathbb{R}^n , so unterscheiden wir in Bezug auf M drei Arten von Punkten $\mathbf{x} \in \mathbb{R}^n$. Gibt es eine ε -Kugel um \mathbf{x} , die ganz in M liegt, so heißt \mathbf{x} **innerer Punkt** von M . In diesem Fall heißt M auch eine **Umgebung** von \mathbf{x} . Gibt es eine ε -Kugel um \mathbf{x} , die ganz außerhalb M liegt (d.h. disjunkt zu M ist), so heißt \mathbf{x} **äußerer Punkt** von M . Ist \mathbf{x} weder innerer noch äußerer Punkt von M , so enthält jede ε -Kugel um \mathbf{x} sowohl Punkte aus M als auch Punkte, die nicht zu M gehören. In diesem Fall heißt \mathbf{x} ein **Randpunkt** von M . Innere Punkte von M gehören immer zu M , äußere Punkte nie, für Randpunkte ist beides möglich. Gibt man zu einer Menge M alle Randpunkte dazu, erhält man ihren **Abschluss**, den man auch mit \overline{M} bezeichnet. Entfernt man alle Randpunkte, so erhält man das **Innere** von M , das man meist mit M° bezeichnet. Klarerweise gilt $M^\circ \subseteq M \subseteq \overline{M}$.

Eine Menge $M \subseteq \mathbb{R}^n$ heißt **offen**, wenn jeder Punkt $\mathbf{x} \in M$ innerer Punkt von M ist. Das ist genau dann der Fall, wenn M keinen seiner Randpunkte enthält. (Offene ε -Kugeln sind offen auch in diesem Sinn.) Enthält M hingegen alle seine Randpunkte, dann heißt M **abgeschlossen**. Man beachte, dass eine Menge weder offen noch abgeschlossen sein muss, denn sie kann ja manche Randpunkte enthalten und andere nicht.

Diese Sprechweisen passen im Fall $n = 1$ auch mit jenen für Intervalle zusammen. Sei also $a < b \in \mathbb{R}$, so hat jedes der Intervalle $[a, b]$, (a, b) , $(a, b]$ und (a, b) genau zwei Randpunkte, nämlich a und b . Die Menge der inneren Punkte ist jeweils (a, b) , die Menge der äußeren Punkte ist $(-\infty, a) \cup (b, \infty)$. Beide Randpunkte enthält $[a, b]$, also ist das abgeschlossene Intervall $[a, b]$ auch im topologischen Sinn abgeschlossen. Aus symmetrischem Grund ist das Intervall (a, b) auch topologisch offen. Die sogenannten halboffenen Intervalle $[a, b)$ und $(a, b]$ sind weder offen noch abgeschlossen.

Übungsaufgabe 113. (T) Geben Sie die Menge aller inneren Punkte, aller Randpunkte und aller äußeren Punkte der Menge $A, B \subseteq \mathbb{R}$ bzw. der Menge $C \subseteq \mathbb{R}^2$ an.

1. $A := \bigcup_{k \in \mathbb{Z}} (k - \frac{1}{4}, k + \frac{1}{4}) \subseteq \mathbb{R}$
2. $B := \{\frac{1}{n} : n \in \mathbb{N}, n \geq 1\} \cup (\frac{1}{2}, 2) \subseteq \mathbb{R}$
3. $C := K(\mathbf{x}_0, 1) \cup K(\mathbf{x}_1, 1) \subseteq \mathbb{R}^2$ mit $\mathbf{x}_0 = (-1, 0)$ und $\mathbf{x}_1 = (1, 0)$.
4. Genau wie 3., nur mit $\mathbf{x}_0 = (0, 0)$.

Übungsaufgabe 114. (E) Begründen Sie folgende Aussagen über Teilmengen von \mathbb{R}^2 und illustrieren Sie die Situation anhand einer Skizze.

1. Jede Kugel $K(\mathbf{x}, \varepsilon)$ ist offen.
2. Der Abschluss \overline{M} einer Menge M ist selbst abgeschlossen, und zwar die kleinste abgeschlossene Menge, die M als Teilmenge enthält.
3. Das Innere M° einer Menge M ist selbst offen, und zwar die größte offene Teilmenge von M .

Ist X ein metrischer Raum mit Metrik d , so heißt eine Teilmenge $M \subseteq X$ **beschränkt**, wenn es ein $x \in X$ und ein $r \in \mathbb{R}$ gibt mit $d(x, y) < r$ für alle $y \in M$, anders formuliert: Wenn M in einer Kugel von endlichem Radius r enthalten ist. Für $X = \mathbb{R}^n$ mit der euklidischen Metrik ist M genau dann beschränkt, wenn es eine (endliche) reelle Zahl $r \in \mathbb{R}$ gibt mit $\|x\| \leq r$ für alle $x \in M$.

$M \subseteq \mathbb{R}^n$ heißt **kompakt**, wenn M sowohl abgeschlossen als auch beschränkt ist.¹⁶

¹⁶Als Definition wird oft eine andere, zur hier gegebenen äquivalente Eigenschaft genommen, die vor allem für Verallgemeinerungen besser geeignet ist. Diese sogenannte Überdeckungseigenschaft lautet: Wann immer M in der Vereinigung (möglicherweise unendlich vieler) offener Mengen enthalten ist, gibt es endlich viele dieser offenen Mengen, deren Vereinigung immer noch eine Obermenge von M ist. Für unsere Zwecke ist diese, dem Anfänger wahrscheinlich komplizierter anmutende Definition aber nicht notwendig.

Ist $n = 1$, also $M \subseteq \mathbb{R}$, so heißt M **zusammenhängend**, wenn M ein Intervall ist (offen, abgeschlossen oder halboffen) oder von einem der Typen \emptyset , (a, ∞) , $[a, \infty)$, $(-\infty, b)$, $(-\infty, b]$ mit $a, b \in \mathbb{R}$ oder ganz \mathbb{R} . Für $n > 1$ muss die korrekte Definition zusammenhängender Mengen komplizierter gefasst werden. Und zwar ist M *nicht* zusammenhängend, wenn es in zwei nichtleere Teilmengen zerlegt werden kann, von denen keine Randpunkte der anderen enthält.¹⁷

¹⁷Auch diese Definition lässt sich auf beliebige topologische Räume übertragen.

Kapitel 2

Folgen und Reihen

Wir wissen bereits: Folgen sind Funktionen mit Definitionsbereich \mathbb{N} (oder einer Teilmenge davon wie z.B. \mathbb{N}^+) und Werten in einer zunächst beliebigen Menge. Oft schreiben wir $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ oder auch $\mathbf{a} = (a_n)$ für die Folge $n \mapsto a_n$ oder, weniger formal, a_0, a_1, a_2, \dots . Die a_n heißen auch Folgenglieder. In diesem Kapitel interessieren uns besonders (wenn auch nicht ausschließlich) reelle Folgen, wo also $a_n \in \mathbb{R}$, siehe 2.1. Im Zusammenhang mit Folgen treffen wir erstmals auf den vielleicht wichtigsten Begriff der Mathematik überhaupt: den Begriff des Grenzwertes. Besonderes Augenmerk verdienen Reihen (2.2). Sie sind selbst wieder Folgen, nämlich Folgen von Partialsummen, die sich durch Addition der Glieder einer Folge ergeben.

2.1 Reelle Folgen

Wir beginnen mit einfachen Beispielen und Eigenschaften von Folgen (2.1.1). Sodann kommen wir zum für die gesamte Mathematik fundamentalen Begriff des Grenzwertes und damit verwandten (2.1.2). Rechenregeln dazu folgen in 2.1.3, ergänzt durch weitere Beispiele (2.1.4). Eine für die Theorie sehr wichtige Abschwächung des Grenzwertes ist der Begriff des Häufungspunktes (2.1.5) mit den für reelle Folgen sehr nützlichen Spezialfällen des oberen und unteren Grenzwertes oder Limes (2.1.6). Die für die weiteren Kapitel fundamentalen allgemeinen, eher theoretischen Resultate folgen in 2.1.7. Abschließend werden noch Folgen behandelt, die auf besondere Weise, nämlich durch Rekursion erzeugt werden (2.1.8).

2.1.1 Einfache Beispiele und Eigenschaften reeller Folgen

Inhalt in Kurzfassung: Den Einstieg bilden sehr einfach gebaute Folgen wie konstante, arithmetische und geometrische. Die Bernoullische Ungleichung wird bewiesen und auf die geometrische Folge angewandt. Die genannten einfachen Beispiele geben Anlass zu allgemeineren Begriffsbildungen. So definieren wir

u.a. was man unter beschränkten, monotonen, alternierenden und periodischen Folgen versteht.

Generell ist eine Folge bekanntlich eine Abbildung $\mathbf{a} : \mathbb{N} \rightarrow X$, $n \mapsto a_n$, die auf der Menge \mathbb{N} definiert ist. Man schreibt auch $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$. Manchmal nimmt man als Definitionsmenge statt \mathbb{N} auch \mathbb{Z} oder eine Teilmenge von \mathbb{N} wie \mathbb{N}^+ oder eine von der Gestalt $\{0, \dots, N\}$ mit $N \in \mathbb{N}$. In letzterem Fall spricht man von endlichen Folgen. Wir betrachten als Standardfall aber den Definitionsbereich \mathbb{N} . Liegen die Werte der Folge in $X = \mathbb{R}$, so sprechen wir von einer **reellen Folge**.

Sehr einfache Beispiele von Folgen $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ sind die **konstanten Folgen**, für die es ein c gibt mit $a_n = c$, weiters in \mathbb{R} die **arithmetischen Folgen**, für die es ein d (Schrittlänge) gibt mit $a_n = a_0 + nd$, jeweils für alle $n \in \mathbb{N}$. Ist $d > 0$, so werden die Folgenglieder **unbeschränkt** immer größer, und man spricht von einer **(streng) monoton wachsenden Folge**, entsprechend für $d < 0$ von einer **(streng) monoton fallenden Folge**, siehe auch Definition 2.1.1.1.

Von besonderer Wichtigkeit sind **geometrische** oder **Exponentialfolgen**, wo $a_n = a_0 q^n$ mit einer festen Basis $q \in \mathbb{R}$. Für $a_0 = 1$, also $a_n = q^n$, untersuchen wir die verschiedenen Fälle für q .

Im Fall von $q > 1$ ist $q = 1 + \varepsilon$ mit $\varepsilon > 0$. Klarerweise ist die Folge in diesem Fall streng monoton wachsend. Sie ist aber auch (nach oben) unbeschränkt. Denn aus dem binomischen Lehrsatz folgt bei $n \geq 1$ die sogenannte **Bernoullische Ungleichung**

$$q^n = (1 + \varepsilon)^n = \sum_{k=0}^n \binom{n}{k} \varepsilon^k \geq \sum_{k=0}^1 \binom{n}{k} \varepsilon^k = \binom{n}{0} \varepsilon^0 + \binom{n}{1} \varepsilon^1 = 1 + n\varepsilon,$$

die in dieser Form offenbar auch für $n = 0$ gilt. Wegen der archimedischen Eigenschaft übersteigt $1 + n\varepsilon$ für hinreichend großes n jede beliebig vorgegebene reelle Zahl, also ist auch die Folge der q^n unbeschränkt.

Für $q = 1$ liegt die konstante Folge mit Wert $a_n = c = 1$ vor, für $q = 0$ ist die Folge ab $n = 1$ konstant mit Wert $a_n = 0$. Für $0 < q < 1$ werden die Folgenglieder mit wachsendem n immer kleiner und nähern sich 0 an. Mit Hilfe des Grenzwertbegriffs werden wir das gleich noch präzisieren. Für $q < 0$ ist die Folge der q^n alternierend, wobei dann die Folge der $(-q)^n = (-1)^n q^n$ zu einem der bereits behandelten Fälle gehört.

Wir kommen nun zu den Definitionen einiger elementarer Eigenschaften, die gewisse Folgen von Interesse haben. Zunächst zu Beschränktheit und verwandten Begriffen:

Definition 2.1.1.1. Eine Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ in einem metrischen Raum X heißt beschränkt, wenn die Menge $\{a_n : n \in \mathbb{N}\}$ ihrer Glieder beschränkt ist. Im Fall reeller Folgen ($X = \mathbb{R}$) unterscheidet man entsprechend **nach (oben/unten) beschränkt**, wenn die Menge $M(\mathbf{a}) := \{a_n : n \in \mathbb{N}\}$ der Folgenglieder die jeweilige Eigenschaft hat, wenn es also ein $r_1 \in \mathbb{R}$ gibt mit $a_n \leq r_1$ für alle

$n \in \mathbb{N}$ (nach oben beschränkt) bzw. wenn es ein $r_0 \in \mathbb{R}$ gibt mit $a_n \geq r_0$ für alle $n \in \mathbb{N}$ (nach unten beschränkt). Beschränkt ist die reelle Folge also genau dann, wenn es sowohl ein r_1 als auch ein r_0 mit den genannten Eigenschaften gibt oder, äquivalent dazu, wenn es ein $r \in \mathbb{R}$ mit $|a_n| \leq r$ für alle $n \in \mathbb{N}$ gibt. Im nach oben/unten beschränkten Fall nennt man das Supremum/Infimum der Menge $M(\mathbf{a})$ auch **Supremum/Infimum** der Folge \mathbf{a} und schreibt dafür $\sup \mathbf{a}$ bzw. $\inf \mathbf{a}$ oder, häufiger, $\sup_{n \in \mathbb{N}} a_n$ bzw. $\inf_{n \in \mathbb{N}} a_n$.

Für die Theorie der reellen Folgen sind die Begriffe rund um Monotonie von ähnlich großer Wichtigkeit:

Definition 2.1.1.2. Eine reelle Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ heißt **monoton wachsend**, wenn $a_n \leq a_{n+1}$, bzw. **streng monoton wachsend**, wenn $a_n < a_{n+1}$, bzw. **monoton fallend**, wenn $a_n \geq a_{n+1}$, bzw. **streng monoton fallend**, wenn $a_n > a_{n+1}$ für alle $n \in \mathbb{N}$ gilt. Beide, monoton wachsende und monoton fallende Folgen heißen auch schlicht **monoton**, streng monoton wachsende und streng monoton fallende Folgen heißen auch **streng monoton**. Statt streng verwendet man auch das Wort **strikt**.

Allgemeiner heißt eine Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ in einem metrischen Raum X mit Metrik d beschränkt, wenn die Menge es einen Punkt

Offenbar folgt aus dieser Definition für beliebige $n_1 < n_2$ stets $a_{n_1} \leq a_{n_2}$ (monoton wachsend), bzw. $a_{n_1} < a_{n_2}$ (streng monoton wachsend), bzw. $a_{n_1} \geq a_{n_2}$ (monoton fallend) bzw. $a_{n_1} > a_{n_2}$ (streng monoton fallend).

Gelegentlich spielen auch solche Folgen eine wichtige Rolle, deren Glieder abwechselndes (alternierendes) Vorzeichen haben:

Definition 2.1.1.3. Die reelle Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ heißt **alternierend**, wenn aufeinanderfolgende Glieder wechselndes Vorzeichen haben, wenn also für alle $n \in \mathbb{N}$ gilt $a_n a_{n+1} \leq 0$ bzw. $a_n a_{n+1} < 0$ (**streng** oder **strikt alternierend**).

Als letzte dieser recht elementaren Eigenschaften von (diesmal beliebigen, nicht nur reellen) Folgen, definieren wir Periodizität:

Definition 2.1.1.4. Die Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ heißt **periodisch**, wenn es eine **Periode**, d.h. ein $p \in \mathbb{N}^+$, mit $a_{n+p} = a_n$ für alle $n \in \mathbb{N}$ gibt.

Übungsaufgabe 115. (T) Untersuchen Sie die Folge $a_n = \frac{n!}{n^n}$ auf die Eigenschaften aus den Definitionen 2.1.1.1 bis 2.1.1.4.

Übungsaufgabe 116. (T) Wie Übungsaufgabe 115, aber für die Folgen

$$a_n = \sin\left(\frac{1}{n}\pi\right), \quad b_n = \sin\left(n\frac{\pi}{2}\right), \quad c_n = n \cos\left(\frac{1}{n}\pi\right).$$

Übungsaufgabe 117. (P) Fertigen Sie eine 5×5 Tabelle für die fünf Eigenschaften beschränkt, monoton wachsend, monoton fallend, alternierend und periodisch an. Jedes der 25 Felder entspricht einem Paar (E_1, E_2) von Eigenschaften aus dieser 5-elementigen Menge, wenn die Zeile E_1 und die Spalte E_2 entspricht. Uns interessieren nur die 10 Felder unter der Diagonale (warum?). Für jedes Feld (E_1, E_2) stellen sich vier Fragen:

1. Gibt es reelle Folgen mit beiden Eigenschaften E_1 und E_2 ?
2. Gibt es reelle Folgen mit Eigenschaft E_1 aber ohne Eigenschaft E_2 ?
3. Gibt es reelle Folgen mit Eigenschaft E_2 aber ohne Eigenschaft E_1 ?
4. Gibt es reelle Folgen, die weder die Eigenschaft E_1 noch die Eigenschaft E_2 haben?

Lautet die Antwort zu einer solchen Frage ja, so ist so eine Folge anzugeben; bei nein ist dies zu begründen.

Behandeln Sie alle oder wenigstens einige interessante der 10 Felder (Kombinationen) in der beschriebenen Weise.

2.1.2 Der Grenzwert einer Folge

Inhalt in Kurzfassung: Der Grenzwert x einer Folge mit den Gliedern a_n ist durch die Eigenschaft gekennzeichnet, dass der Abstand zwischen Grenzwert und Folgengliedern in dem Sinn beliebig klein wird, dass zu jedem (beliebig kleinen) $\varepsilon > 0$ ein (hinreichend großer) Index n_0 existiert, ab dem sämtliche Folgen Glieder a_n die Beziehung $|a_n - x| < \varepsilon$ erfüllen. Eine Folge heißt konvergent, wenn es einen Grenzwert gibt, sonst divergent. Die Grundidee dieses Begriffs ist nicht nur auf \mathbb{R} sondern z.B. auf beliebige metrische Räume anwendbar. Der Grenzwertbegriff ist das Um und Auf großer Teile der Mathematik und wird sich mit geringfügigen Variationen auch durch die meisten weiteren Kapitel der Vorlesung ziehen. Eng verwandt mit dem Begriff des Grenzwertes einer Folge ist jener der Cauchyfolge.

Wir gehen aus vom Beispiel $a_n = \frac{1}{n}$ für $n \geq 1$, haben es also mit der Folge $a_1 = 1$, $a_2 = \frac{1}{2} = 0,5$, $a_3 = \frac{1}{3} = 0,333\dots$, $a_4 = \frac{1}{4} = 0,25$ etc. zu tun. Anscheinend nähern sich die a_n dem Wert 0 an, dem sogenannten **Grenzwert** der Folge. Doch was genau soll das bedeuten? Wir stehen hier vor der Aufgabe, das intuitiv recht überzeugende Konzept der *beliebigen Annäherung* mit einem scharfen Begriff zu fassen. Es ist keine Übertreibung, wenn man behauptet, dass das Ringen um diesen Begriff Jahrtausende gedauert hat. Denn schon Eudoxos im 4. und Archimedes, der gemeinhin als der größte Mathematiker der Antike gilt, im 3. Jahrhundert vor unserer Zeitrechnung waren ganz nahe am modernen Grenzwertbegriff, ebenso wie fast zwei Jahrtausende später Gottfried Wilhelm Leibniz (1646-1716). Doch erst etwa seit Augustin-Louis Cauchy (1789-1857) herrscht Einigkeit über die exakte Definition des Grenzwertes einer Folge.

Definition 2.1.2.1. Sei $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ eine reelle Folge. Die Zahl $x \in \mathbb{R}$ heißt **Grenzwert** von \mathbf{a} , symbolisch

$$x = \lim_{n \rightarrow \infty} a_n$$

oder auch

$$a_n \rightarrow x \quad \text{für } n \rightarrow \infty,$$

wenn es zu jedem $\varepsilon > 0$ einen Index $n_0 = n_0(\varepsilon) \in \mathbb{N}$ gibt derart, dass für alle $n \geq n_0$ gilt: $|a_n - x| < \varepsilon$. Die gesamte Definition für $x = \lim_{n \rightarrow \infty} a_n$ lässt sich konzis als logische Formel schreiben:

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |a_n - x| < \varepsilon$$

Existiert ein Grenzwert, so heißt die Folge **konvergent**. Man sagt, sie **konvergiert** oder **strebt** gegen ihren Grenzwert. Ist dieser Grenzwert 0, so spricht man von einer **Nullfolge**. Eine Folge, die nicht konvergiert, heißt **divergente Folge**, sie **divergiert**.

Übungsaufgabe 118. (T) Schreiben Sie auch die Divergenzaussage als Formel mit Quantoren an, ohne das Negationssymbol \neg zu verwenden. Unterscheiden Sie dabei zwei verschiedene Aussagen:

1. Die Folge der a_n konvergiert nicht gegen x .
2. Die Folge der a_n divergiert.

Für den späteren Gebrauch (vor allem in Mathematik 2) bringen wir die Definition des Grenzwertes allgemeiner auch für Folgen in beliebigen metrischen Räumen X (statt \mathbb{R}) mit Metrik d (siehe Definition in 1.5.9):

Definition 2.1.2.2. Sei $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ eine Folge in einem metrischen Raum X . Der Punkt $x \in X$ heißt **Grenzwert** von \mathbf{a} , symbolisch

$$x = \lim_{n \rightarrow \infty} a_n \quad \text{oder auch} \quad a_n \rightarrow x \quad \text{für} \quad n \rightarrow \infty,$$

wenn es zu jeder Umgebung U von x einen Index $n_0 = n_0(\varepsilon) \in \mathbb{N}$ gibt derart, dass $a_n \in U$ für alle $n \geq n_0$.

Alle darauf aufbauenden Begriffe und Sprechweisen können vom reellen Fall übernommen werden. So heißt eine Folge auch in einem beliebigen metrischen Raum **konvergent**, wenn sie einen Grenzwert hat, und man sagt, die Folge **konvergiert** oder **strebt** gegen ihren Grenzwert. Wieder heißt eine Folge, die nicht konvergiert, eine **divergente Folge**, sie **divergiert**.

Praktisch ist im Zusammenhang mit Grenzwerten auch folgende Sprechweise: Eine Aussage gilt **für fast alle** $n \in \mathbb{N}$, wenn es nur endlich viele (sofern überhaupt) Ausnahmen gibt. Dann lässt sich die Grenzwertbeziehung $x = \lim_{n \rightarrow \infty} a_n$ umformulieren zu: Für jedes $\varepsilon > 0$ bzw. für jede Umgebung U von x liegen fast alle Folgenglieder im Intervall $(x - \varepsilon, x + \varepsilon)$ bzw. in U .

Übungsaufgabe 119. (P) Argumentieren Sie sorgfältig, warum diese Formulierung tatsächlich äquivalent dazu ist, dass die a_n im Sinne von Definition 2.1.2.1 gegen x konvergieren.

Klarerweise sind konstante Folgen $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ mit $a_n = c \in \mathbb{R}$ für alle $n \in \mathbb{N}$ konvergent mit Grenzwert $x = c$. Denn zu beliebig vorgegebenem $\varepsilon > 0$ können wir $n_0 = 0$ setzen, so dass für alle $n \geq n_0 = 0$ die geforderte Beziehung

$|a_n - x| = |x - x| = 0 < \varepsilon$ gilt. Weniger trivial sind konvergente Folgen, die nicht konstant sind. Eines der einfachsten aber auch wichtigsten Beispiele dieser Art ist die schon eingangs zur Motivation betrachtete Folge mit den für alle $n \geq 1$ definierten Gliedern $a_n = \frac{1}{n}$. Sie hat den Grenzwert $x = 0$, ist also eine Nullfolge. Um das streng nach der Definition des Grenzwertes nachzuprüfen, muss diesmal (was für Grenzwertuntersuchungen typisch ist) zu jedem vorgegebenem $\varepsilon > 0$ ein spezifisches, von ε abhängiges $n_0 = n_0(\varepsilon)$ angegeben werden. Das gelingt, wenn man irgendein $n_0(\varepsilon) > \frac{1}{\varepsilon}$ nimmt, was ja immer möglich ist. Denn dann erfüllen alle $n \geq n_0(\varepsilon)$ gleichfalls die Beziehung $|a_n - x| = |\frac{1}{n} - 0| = \frac{1}{n} < \frac{1}{n_0(\varepsilon)} < \varepsilon$. Also gilt tatsächlich $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$.

Zur Übung ein sehr ähnliches Beispiel:

Übungsaufgabe 120. (T) Zeigen Sie direkt anhand von Definition 2.1.2.1, dass die Folge

$$x_n = \frac{1-n}{1+n}$$

gegen den Grenzwert $x = -1$ konvergiert. Bestimmen Sie für $\varepsilon = 1$, $\varepsilon = \frac{1}{10}$ und $\varepsilon = \frac{1}{100}$ jeweils die kleinste Zahl $n_\varepsilon \in \mathbb{N}$, so dass $|x_n - x| < \varepsilon$ für alle $n \geq n_\varepsilon$ gilt.

Ganz ähnlich gilt $\lim_{n \rightarrow \infty} a_n = 0$ auch für die weiter oben betrachteten Folgen der Bauart $a_n = q^n$ mit $|q| < 1$ (hier sind auch komplexe q zugelassen). Das ist leicht einzusehen: Wegen $|q^n| = |q|^n$ dürfen wir der Einfachheit halber $q > 1$ voraussetzen. Sei $Q := q^{-1} > 1$. Dann gilt $Q = 1 + \delta$ mit einem $\delta > 0$, nach der Bernoullischen Ungleichung aus 2.1.1 daher $Q^n \geq 1 + n\delta$. Um $\lim_{n \rightarrow \infty} q^n = 0$ zu beweisen, haben wir uns ein beliebiges $\varepsilon > 0$ vorzugeben. Wegen $\delta > 0$ und der archimedischen Eigenschaft werden die Q^n mit wachsendem n beliebig groß, genauer: Es gibt ein $n_0 \in \mathbb{N}$ mit $Q^n \geq 1 + n\delta > \varepsilon^{-1}$ für alle $n \geq n_0$. Für diese n gilt dann aber auch $0 < q^n = (Q^n)^{-1} = \varepsilon$, was zu zeigen war.

Sehr ähnlich dem Begriff der Konvergenz einer Folge ist jener der **Cauchyfolge**:

Definition 2.1.2.3. Eine reelle Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ heißt Cauchyfolge, wenn es zu beliebig (klein) vorgegebenem $\varepsilon > 0$ einen (hinreichend großen) Index $n_0 \in \mathbb{N}$ gibt, ab dem je zwei Glieder (d.h. alle $n_1, n_2 \geq n_0$) einen Abstand $|a_{n_1} - a_{n_2}|$ kleiner ε haben. Als Formel:

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n_1, n_2 \geq n_0 : |a_{n_1} - a_{n_2}| < \varepsilon.$$

Zur Erläuterung: Der wesentliche Unterschied zur Konvergenz besteht also darin, dass es dort um den Abstand $|a_n - x|$ der Folgeglieder zu einem (in der Formel explizit auftretenden) Grenzwert x geht, bei der Cauchyfolge hingegen um den Abstand $|a_{n_1} - a_{n_2}|$ von Folgengliedern zueinander. Es ist nicht schwer zu sehen, dass jede konvergente Folge eine Cauchyfolge ist. Dass in \mathbb{R} auch die Umkehrung gilt, ist weniger selbstverständlich und liegt an der Vollständigkeit

der reellen Zahlen. In 2.1.7 werden wir darauf noch eingehen.

Folgende Beobachtungen scheinen erwähnenswert:

1. Um uns zu vergewissern, dass es bei der symbolischen Schreibweise $\lim_{n \rightarrow \infty} a_n$ nicht zu Mehrdeutigkeiten kommen kann, halten wir fest, dass eine Folge höchstens einen Grenzwert haben kann. Das ist leicht zu sehen, siehe 2.1.7
2. Ändert man nur endlich viele Glieder einer Folge ab, lässt man nur endlich viele weg oder fügt man nur endlich viele hinzu, so ändert das nichts an Konvergenzverhalten und gegebenenfalls am Grenzwert. (Dies wird sofort klar, wenn man an die Formulierung des Grenzwertes mit *fast alle* denkt.) Entsprechend wollen wir den Grenzwert gelegentlich auch dann als solchen gelten lassen, wenn endlich viele Glieder der Folge gar nicht definiert sind (eventuell weil für gewisse n ein Nenner 0 werden kann). Analoges gilt für Häufungspunkte von Folgen, siehe 2.1.5.
3. Jede **Teilfolge** a_{n_0}, a_{n_1}, \dots ($n_0 < n_1 < n_2 < \dots$) einer konvergenten Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ ist konvergent gegen denselben Grenzwert. Das ist klar, wenn man an die Formulierung mit „fast alle“ für „höchstens endlich viele Ausnahmen“ denkt. Denn wenn es nur endlich viele n gibt, für die $|a_n - \alpha| < \varepsilon$ nicht gilt, dann erst recht, wenn man nur mehr gewisse, nämlich die n_k , in Betracht zieht.
4. Jede konvergente Folge ist beschränkt. Ist nämlich $x = \lim_{n \rightarrow \infty} a_n$ (die a_n und x dürfen in einem beliebigen metrischen Raum X mit Metrik d liegen), so gibt es ein $n_0 \in \mathbb{N}$ mit $d(a_n, x) < 1$ für alle $n \geq n_0$. Sei r die größte der endlich vielen reellen Zahlen $d(a_0, x) + 1, d(a_1, x) + 1, \dots, d(a_{n_0-1}, x) + 1$, so folgt $d(a_n, x) < r$ für alle $n \in \mathbb{N}$. Also liegen sämtliche Folgenglieder in der Kugel um x mit Radius r , bilden also eine beschränkte Menge. Somit ist die Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ beschränkt.

2.1.3 Konvergenzregeln (Grenzwertsätze)

Inhalt in Kurzfassung: Wir konzentrieren uns nun auf reelle und komplexe Folgen und studieren vor allem das Verhalten der Grundrechnungsarten in Bezug auf den Grenzwert, außerdem der Betrags- und der Wurzelfunktionen. Es erweist sich, dass sie sich auf angenehme Weise mit dem Grenzwert vertragen, z.B.: Der Grenzwert einer Summe $a_n + b_n$ ist gleich der Summe der Grenzwerte der a_n und der b_n , der Grenzwert von k -ten Wurzeln $\sqrt[k]{a_n}$ ist die k -te Wurzel des Grenzwertes der $a_n \geq 0$ etc. In anderer Terminologie, die erst im nächsten Kapitel ausführlich behandelt werden wird: Sowohl die Grundrechnungsarten als auch die genannten Funktionen sind stetig.

Der Begriff des Grenzwertes spielt sehr gut mit den Rechenoperationen Addition etc. zusammen. Die nachfolgenden Konvergenzregeln (manchmal auch **Grenzwertsätze** genannt) bringen das zum Ausdruck.

Satz 2.1.3.1. Seien $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ und $\mathbf{b} = (b_n)_{n \in \mathbb{N}}$ konvergente reelle Folgen mit $\lim_{n \rightarrow \infty} a_n = x$ und $\lim_{n \rightarrow \infty} b_n = y$. Dann sind auch die nachfolgend angegebenen Folgen $\mathbf{c} = (c_n)_{n \in \mathbb{N}}$ konvergent gegen die angegebene Zahl $z = \lim_{n \rightarrow \infty} c_n$.

1. $c_n = a_n + b_n$, $z = x + y$, also:

$$\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n$$

2. $c_n = a_n - b_n$, $z = x - y$, also:

$$\lim_{n \rightarrow \infty} (a_n - b_n) = \lim_{n \rightarrow \infty} a_n - \lim_{n \rightarrow \infty} b_n$$

3. $c_n = a_n b_n$, $z = xy$, also:

$$\lim_{n \rightarrow \infty} a_n b_n = \lim_{n \rightarrow \infty} a_n \cdot \lim_{n \rightarrow \infty} b_n$$

Zusatz: Ist $x = 0$, so genügt es sogar, wenn die b_n beschränkt sind, sie müssen nicht konvergieren.

4. Hier sei $b_n \neq 0$ für alle n und auch $y \neq 0$ vorausgesetzt, $c_n = \frac{a_n}{b_n}$, $z = \frac{x}{y}$, also:

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}$$

5. $c_n = |a_n|$, $z = |x|$, also:

$$\lim_{n \rightarrow \infty} |a_n| = \left| \lim_{n \rightarrow \infty} a_n \right|$$

6. Hier sei $a_n \geq 0$ vorausgesetzt, $c_n = \sqrt{a_n}$ und $z = \sqrt{x}$, also:

$$\lim_{n \rightarrow \infty} \sqrt{a_n} = \sqrt{\lim_{n \rightarrow \infty} a_n}$$

7. $c_n = \sqrt[k]{a_n}$ (wobei für gerades k auch hier $a_n \geq 0$ vorausgesetzt sei), also:

$$\lim_{n \rightarrow \infty} \sqrt[k]{a_n} = \sqrt[k]{\lim_{n \rightarrow \infty} a_n}.$$

8. Stimmen die Grenzwerte $x = \lim_{n \rightarrow \infty} a_n = y = \lim_{n \rightarrow \infty} b_n$ überein, und gilt $a_n \leq c_n \leq b_n$ für alle $n \in \mathbb{N}$, so auch $\lim_{n \rightarrow \infty} c_n = x = y$.

Beweis. 1. Sei $\varepsilon > 0$. Wegen $\lim_{n \rightarrow \infty} a_n = x$ gibt es ein n_1 mit $|a_n - x| < \frac{\varepsilon}{2}$ für alle $n \geq n_1$. Analog gibt es ein n_2 mit $|b_n - y| < \frac{\varepsilon}{2}$ für alle $n \geq n_2$. Wir setzen $n_0 := \max\{n_1, n_2\}$. Dann gilt für alle $n \geq n_0$: $|c_n - z| = |(a_n + b_n) - (x + y)| = |(a_n - x) + (b_n - y)| \leq |a_n - x| + |b_n - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

2. Man ersetze in 1. b_n durch $-b_n$ und y durch $-y$.

3. Hier formt man die abzuschätzende Differenz zweckmäßig wie folgt um:
 $a_n b_n - xy = a_n b_n - a_n y + a_n y - xy = a_n(b_n - y) + (a_n - x)y$. Weil die a_n konvergieren, sind sie auch beschränkt, etwa durch $|a_n| \leq a$ mit $a > 0$. Wegen der Konvergenz der b_n gibt es ein n_1 mit $|b_n - y| < \frac{\varepsilon}{2|a|+1}$ für alle $n \geq n_1$, wegen jener der a_n ein n_2 mit $|a_n - x| < \frac{\varepsilon}{2|y|+1}$ für alle $n \geq n_2$. Für alle $n \geq n_0 := \max\{n_1, n_2\}$ rechnet man mit Obigem

$$|c_n - z| = |a_n b_n - xy| = |a_n(b_n - y) + (a_n - x)y| \leq |a_n| \cdot |b_n - y| + |a_n - x| \cdot |y| < \varepsilon$$

nach. Den Zusatz beweist man sehr ähnlich (Übung).

4. Zuerst betrachten wir den Spezialfall $a_n = 1$ für alle n , also $x = 1$. Die abzuschätzende Differenz ist dann $\frac{1}{b_n} - \frac{1}{y} = \frac{y - b_n}{y b_n}$. Ähnlich wie in 3. sieht man, dass dieser Wert für hinreichend große n betragsmäßig kleiner als ε wird. Somit ist dieser Spezialfall erledigt. Der allgemeine Fall folgt dann wegen $\frac{a_n}{b_n} = a_n \frac{1}{b_n}$ aus 3.
5. Wegen $||a_n| - |x|| \leq |a_n - x|$ kann jedes $n_0(\varepsilon)$ für die a_n auch für die $|a_n|$ genommen werden.
6. Sei zunächst $x = 0$. Zu vorgegebenem $\varepsilon > 0$ ist auch $\varepsilon^2 > 0$. Also gilt $0 \leq a_n < \varepsilon^2$ und somit auch $0 \leq \sqrt{a_n} < \varepsilon$ für fast alle n .

Sei nun $x \neq 0$. Abzuschätzen ist die Differenz

$$\sqrt{a_n} - \sqrt{x} = \frac{(\sqrt{a_n} - \sqrt{x})(\sqrt{a_n} + \sqrt{x})}{\sqrt{a_n} + \sqrt{x}} = \frac{a_n - x}{\sqrt{a_n} + \sqrt{x}} = (a_n - x)q_n$$

mit $q_n = \frac{1}{\sqrt{a_n} + \sqrt{x}}$. Nach Voraussetzung gilt $a_n - x \rightarrow 0$. Wegen $x > 0$ und $a_n \rightarrow x$ müssen die q_n ab einem geeigneten n_0 beschränkt sein. Wegen 3. folgt daraus $\sqrt{a_n} - \sqrt{x} = (a_n - x)q_n \rightarrow 0$. Somit gilt $\sqrt{a_n} \rightarrow \sqrt{x}$.

7. Ähnlich wie für Quadratwurzeln, Übung.
8. Sei $\varepsilon > 0$ vorgegeben. Nach Voraussetzung gibt es $n_1, n_2 \in \mathbb{N}$ mit $|a_n - x| < \varepsilon$ für alle $n \geq n_1$ und $|b_n - x| < \varepsilon$ für alle $n \geq n_2$. Sei nun $n \geq n_0 := \max\{n_1, n_2\}$. Ist $x \leq c_n$, also $x \leq c_n \leq b_n$, so folgt $|c_n - x| \leq |b_n - x| < \varepsilon$. Ist hingegen $c_n < x$, also $a_n \leq c_n < x$, so folgt ebenfalls $|c_n - x| \leq |a_n - x| < \varepsilon$, was zu zeigen war. □

Übungsaufgabe 121. (E) Im Beweis von Satz 2.1.3.1 sind vereinzelt Übungen übrig geblieben. Tragen Sie diese Argumente nach:

1. Beweisen Sie den Zusatz in Aussage 3.
2. Im Beweis von Aussage 6 wurde behauptet, dass die q_n beschränkt sind. Begründen Sie das sorgfältig.

3. Beweisen Sie Aussage 7 für $k = 3$. Anleitung:

$$a_n - x = (\sqrt[3]{a_n} - \sqrt[3]{x})(\sqrt[3]{a_n^2} + \sqrt[3]{a_n x} + \sqrt[3]{x^2}) = a_n - x.$$

4. Beweisen Sie Aussage 7 für allgemeines k . Anleitung: Suchen Sie nach einer Verallgemeinerung der Anleitung für den Fall $k = 3$.

Übungsaufgabe 122. (P) Gibt es reelle Folgen mit Gliedern a_n bzw. b_n , wo zwar die Summen $a_n + b_n$ und Produkte $a_n b_n$ konvergieren, nicht aber die a_n und die b_n selbst. Wenn ja, geben Sie ein Beispiel, wenn nein, begründen Sie dies.

Übungsaufgabe 123. (P) Finden Sie konkrete Beispiele für folgende Situationen:

1. Zwei Folgen mit Gliedern a_n bzw. b_n , welche gegen 4 konvergieren und zusätzlich $a_n > b_n > 4$ für alle n erfüllen.
2. Drei verschiedene Folgen, welche divergieren, obwohl unendlich viele Folgenglieder gleich 0 sind.
3. Eine divergente Folge mit Gliedern a_n , für welche $|a_n|$ gegen 2 konvergiert.

Übungsaufgabe 124. (P)

Welche der folgenden Aussagen sind wahr? Begründen Sie Ihre Antwort bzw. geben Sie ein Gegenbeispiel, wenn die Aussage falsch ist.

1. In jedem (noch so kleinen) Intervall $(-\varepsilon, \varepsilon)$ mit $\varepsilon > 0$ gibt es eine divergente Folge.
2. Sind zwei Folgen divergent, so ist ihre Differenz auch stets divergent.
3. Ist (a_n) eine beschränkte Folge, so konvergiert $(\frac{a_n}{n^2})$.
4. Ist (a_n) eine divergente Folge, so divergiert auch (a_{2n}) .
5. Ist (a_n) eine konvergente Folge, so konvergiert auch (a_{2n}) .

Übungsaufgabe 125. (T) In welchem Schritt der folgenden Rechnung, die offenbar ein falsches Ergebnis liefert, liegt der Fehler?

$$1 = \lim_{n \rightarrow \infty} 1 = \lim_{n \rightarrow \infty} \frac{1}{n} \cdot n = \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \lim_{n \rightarrow \infty} n = 0 \cdot \lim_{n \rightarrow \infty} n = 0$$

Übungsaufgabe 126. (E) Es sei die reelle Folge a_n konvergent mit Grenzwert a . Betrachten Sie die Folge

$$b_n = \frac{a_1 + a_2 + \dots + a_n}{n}.$$

Es ist also b_n der Mittelwert der ersten n Folgenglieder von (a_n) . Zeigen Sie, dass auch die Folge der b_n gegen den Grenzwert a konvergiert.

Hinweis: Die Differenz $|b_n - a|$ ist abzuschätzen. Zerlegen Sie diese in drei Summanden I, II und III und untersuchen Sie deren Konvergenzverhalten.

$$b_n - a = \underbrace{\frac{(a_1 - a) + \dots + (a_{n(\varepsilon)} - a)}{n}}_I + \underbrace{\frac{(a_{n(\varepsilon)+1} - a) + \dots + (a_n - a)}{n - n(\varepsilon)}}_II \underbrace{\frac{n}{n}}_III$$

2.1.4 Einige Beispiele konvergenter Folgen

Inhalt in Kurzfassung: Mit Hilfe der aus 2.1.3 verfügbaren Grenzwertsätze wollen wir uns nun mit einigen interessanten Beispielen konvergenter Folgen befassen. Ausgehend vom archetypischen Beispiel einer nicht konstanten aber konvergenten Folge, nämlich jener mit den Gliedern $a_n = \frac{1}{n}$, die wir schon als Nullfolge ausgewiesen haben, erlauben die Grenzwertsätze, den Grenzwert beliebiger gebrochen rationaler Folgen unmittelbar abzulesen. Weitere interessante konvergente Folgen sind jene mit den Gliedern $a_n = \sqrt[n]{n}$ (Grenzwert 1) und $a_n = \left(1 + \frac{1}{n}\right)^n$ (Grenzwert e , Eulersche Zahl, Beweis später mittels Differentialrechnung).

Wir beginnen mit einem Beispiel wie etwa $a_n = \frac{3n-4}{5n+2}$. Bekanntlich kann man den Grenzwert mit Hilfe der Konvergenzregeln (Grenzwertsätze, siehe Satz 2.1.3.1) berechnen:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{3n-4}{5n+2} &= \lim_{n \rightarrow \infty} \frac{3 - \frac{4}{n}}{5 + \frac{2}{n}} = \frac{\lim_{n \rightarrow \infty} (3 - \frac{4}{n})}{\lim_{n \rightarrow \infty} (5 + \frac{2}{n})} = \frac{\lim_{n \rightarrow \infty} 3 - \lim_{n \rightarrow \infty} \frac{4}{n}}{\lim_{n \rightarrow \infty} 5 + \lim_{n \rightarrow \infty} \frac{2}{n}} = \\ &= \frac{3 - \lim_{n \rightarrow \infty} 4 \cdot \lim_{n \rightarrow \infty} \frac{1}{n}}{5 + \lim_{n \rightarrow \infty} 2 \cdot \lim_{n \rightarrow \infty} \frac{1}{n}} = \frac{3 - 4 \cdot 0}{5 + 2 \cdot 0} = \frac{3}{5} \end{aligned}$$

Abgesehen von den erwähnten Regeln haben wir als einzige Grenzwertaussagen die an früherer Stelle diskutierten, mehr oder weniger trivialen Beziehungen $\lim_{n \rightarrow \infty} c = c$ für konstante Folgen sowie $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$ verwendet.

In dieser Rechnung waren für den Grenzwert der Folge offenbar allein die Koeffizienten bei der höchsten Potenz in Zähler und Nenner entscheidend. Wir wollen dieses Beispiel auf beliebige **gebrochen rationale Folgen** verallgemeinern. Definitionsgemäß haben dann die Glieder a_n die Bauart

$$a_n = \frac{\alpha_k n^k + \alpha_{k-1} n^{k-1} + \dots + \alpha_1 n + \alpha_0}{\beta_l n^l + \beta_{l-1} n^{l-1} + \dots + \beta_1 n + \beta_0}$$

mit gewissen $k, l \in \mathbb{N}$ und $\alpha_i, \beta_j \in \mathbb{R}$, wobei wir $\alpha_k \neq 0 \neq \beta_l$ voraussetzen dürfen. Wir unterscheiden die drei Fälle $k < l$, $k = l$ und $k > l$. Im Fall $k < l$ dividieren wir Zähler und Nenner durch n^l und erhalten $a_n = \frac{Z(n)}{N(n)}$ mit $Z(n) = \alpha_k n^{k-l} + \alpha_{k-1} n^{k-1-l} + \dots + \alpha_1 n^{1-l} + \alpha_0 n^{-l}$ und $N(n) = \beta_l + \beta_{l-1} n^{-1} + \dots + \beta_1 n^{1-l} + \beta_0 n^{-l}$. Aufgrund der entsprechenden Konvergenzregeln konvergiert jeder der Summanden im Zähler gegen $\lim_{n \rightarrow \infty} \alpha_i n^{i-l} = \alpha_i (\lim_{n \rightarrow \infty} \frac{1}{n})^{i-l} = 0$, also auch deren Summe $Z(n)$, während $\lim_{n \rightarrow \infty} N(n) = \beta_l \neq 0$. Insgesamt gilt

also $\lim_{n \rightarrow \infty} a_n = \frac{Z(n)}{N(n)} = \frac{\lim_{n \rightarrow \infty} Z(n)}{\lim_{n \rightarrow \infty} N(n)} = \frac{0}{\beta_l} = 0$. Ganz ähnlich überlegt man sich für den Fall $k = l$ die Konvergenz $\lim_{n \rightarrow \infty} a_n = \frac{\alpha_k}{\beta_l} = \frac{\alpha_k}{\beta_k}$. Ist hingegen $k > l$, so dominiert der Zähler gegenüber dem Nenner, und es herrscht Divergenz gegen ∞ oder $-\infty$, abhängig davon, ob $\alpha_k \beta_l > 0$ oder $\alpha_k \beta_l < 0$.

Übungsaufgabe 127. (T) Wie Übungsaufgabe 120, aber mit $x_n = \frac{2n^3}{n^3+n}$ und dem Grenzwert $x = 2$.

Von Interesse ist auch die spezielle Folge $a_n := \sqrt[n]{n}$. Interessant daran ist, dass für $n \rightarrow \infty$ das n unter der Wurzel vergrößernd wirkt, jenes über dem Wurzelsymbol jedoch dämpfend. Es zeigt sich $\lim_{n \rightarrow \infty} a_n = 1$, dass also der dämpfende Einfluss stärker ist. Zum Beweis: Sei $\varepsilon > 0$ beliebig. Wegen der archimedischen Eigenschaft wird $\frac{n+1}{2}\varepsilon^2$ für große n beliebig groß, genauer: Es gibt ein $n_0 \in \mathbb{N}$ mit $\frac{n+1}{2}\varepsilon^2 \geq 1$ für alle $n \geq n_0 \geq 2$. Mit dem binomischen Lehrsatz folgt

$$(1 + \varepsilon)^n = 1 + n\varepsilon + \frac{n(n+1)}{2}\varepsilon^2 + \dots \geq \frac{n(n+1)}{2}\varepsilon^2 \geq n,$$

also nach Ziehen der n -ten Wurzel $1 + \varepsilon \geq \sqrt[n]{n} \geq 1$, woraus die Behauptung folgt.

Ein anderes prominentes Beispiel einer Folge, wo wachsende und dämpfende Einflüsse gegeneinander abzuwägen sind, ist die Folge $a_n = (1 + \frac{1}{n})^n$. Es zeigt sich, dass diese Folge gegen eine irrationale Zahl zwischen 2 und 3 konvergiert. Es handelt sich dabei um die sogenannte **Eulersche Zahl** $e = 2,71828\dots$. Die genaueren Zusammenhänge werden wir im Zuge der Differentialrechnung besser verstehen. Deshalb verzichten wir an dieser Stelle auf einen Beweis.

Übungsaufgabe 128. (T) Untersuchen Sie die Folgen auf Konvergenz und bestimmen Sie gegebenenfalls den Grenzwert:

$$\begin{aligned} 1. a_n &= (-1)^n \cdot (-n+2)^3 \\ 2. b_n &= \frac{2n^2+n}{n-1} - \frac{2n^2}{n+2} & 3. c_n &= \frac{(1+(-1)^n)n+n^2}{(n+1)^2} \\ 4. d_n &= \frac{((-1)^n \cdot n+4)n}{n^2+1} \end{aligned}$$

Übungsaufgabe 129. (T) Berechnen Sie die Grenzwerte, falls vorhanden:

$$\begin{aligned} 1. a_n &= \sqrt{n+3} - \sqrt{n+1} \\ 2. b_n &= \sqrt{n+1}(\sqrt{n+2} - \sqrt{n}) & 3. c_n &= n(\sqrt{n+1} - \sqrt{n}) \\ 4. d_n &= \frac{\sqrt{n}\sin(n)}{n+1} & 5. e_n &= \frac{(3n+1)^2 + \sin^2(\sqrt{n})}{n^2} \\ 6. f_n &= \sqrt[n]{(-\frac{1}{2})^n + 2(\frac{1}{2})^n} \end{aligned}$$

Übungsaufgabe 130. (T) Berechnen Sie die Grenzwerte, falls vorhanden:

$$1. a_n = n^2 \left(\sqrt{1 + \frac{1}{n^2}} - 1 \right) \quad 2. b_n = n \left(\frac{1}{\sqrt{n+2}} - \frac{1}{\sqrt{n}} \right)$$

Übungsaufgabe 131. (T) Verwenden Sie die Beziehung $\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x$, um die Grenzwerte, falls vorhanden, zu berechnen:

$$a_n = \frac{(n+1)^{2n+1} n}{n^{2n+1}(n+1)}, \quad b_n = \left(1 + \frac{2}{n+1}\right)^{n^2+n}, \quad c_n = \left(1 + \frac{2}{n^2}\right)^{2n}$$

2.1.5 Häufungspunkte

Inhalt in Kurzfassung: Der Begriff des Häufungspunktes x einer Folge mit den Gliedern a_n ist eine Abschwächung des Grenzwertbegriffs. Und zwar verlangt man, dass die charakteristische Ungleichung $|a_n - x| < \varepsilon$ für beliebige $\varepsilon > 0$ wenigstens für unendlich viele und nicht unbedingt für fast alle n gelten muss. So wie der Grenzwert, ist dieser Begriff nicht nur in \mathbb{R} , sondern in beliebigen metrischen Räumen sinnvoll. Eine Variante des Häufungspunktes einer Folge ist der Häufungspunkt einer Menge in \mathbb{R} oder allgemeiner in einem metrischen Raum).

Für $x \in \mathbb{R}$ ist die Eigenschaft, Häufungspunkt einer Folge zu sein, schwächer als jene, Grenzwert zu sein. Im Gegensatz zum Grenzwert, wo in jeder Umgebung U von x ab einem gewissen Index alle Folgenglieder liegen müssen, genügt es für einen Häufungspunkt, dass immer wieder (also unendlich viele) Folgenglieder in U liegen, eventuell aber auch unendliche viele andere Folgenglieder außerhalb von U . Um Analogie wie Unterschied zum Grenzwert deutlich zu machen, geben wir die Definition nach demselben Muster, d.h. in klassischer ε -Formulierung:

Definition 2.1.5.1. Die Zahl $x \in \mathbb{R}$ heißt **Häufungspunkt** der reellen Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$, wenn in jeder Umgebung von x unendlich viele Glieder von \mathbf{a} liegen, mit anderen Worten: wenn es zu jedem $\varepsilon > 0$ und zu jedem $n_0 \in \mathbb{N}$ ein $n \geq n_0$ gibt mit $|a_n - x| < \varepsilon$, als logische Formel:

$$\forall \varepsilon > 0 \forall n_0 \in \mathbb{N} \exists n \geq n_0 : |a_n - x| < \varepsilon$$

Zur Erinnerung die entsprechende Formel für den Grenzwert $\lim_{n \rightarrow \infty} a_n = x$:

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |a_n - x| < \varepsilon$$

Der Vergleich der beiden Formeln zeigt, dass lediglich die letzten beiden Quantoren durch den jeweils anderen ersetzt wurden. Klarerweise ist die Forderung für den Grenzwert die stärkere. Denn da müssen ab einem gewissen Index *alle* Folgenglieder in der ε -Umgebung von x liegen, insbesondere also unendlich viele; beim Häufungspunkt dürfen immer wieder Glieder weit entfernt sein von x , solange nur immer wieder andere Glieder in die Nähe kommen.

Ein sehr einfaches Beispiel, das den Unterschied zwischen Grenzwert und Häufungspunkt zeigt, ist die alternierende Folge $a_n = (-1)^n$. Sie hat zwei Häufungspunkte, nämlich 1 und -1 , aber keinen Grenzwert. Wir sehen also erstmals anhand von zwei wichtigen Begriffen, dass die quantorenlogische Struktur wirklich entscheidend ist. Immerhin lässt sich bei Vorliegen eines Häufungspunktes

aus den a_n aber eine Teilfolge auswählen, die gegen den Häufungspunkt konvergiert. Auf diesen Sachverhalt werden wir in 2.1.7 noch zurückkommen, weil es in der Weiterentwicklung der Theorie eine wichtige Rolle spielt.

Folgen können sehr viele Häufungspunkte haben. Zur Illustration einige Übungen dazu:

Übungsaufgabe 132. (P) Für eine reelle Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ bezeichne $H(\mathbf{a})$ die Menge der Häufungspunkte von \mathbf{a} . Geben Sie ein \mathbf{a} an mit:

1. $H(\mathbf{a}) = \emptyset$
2. $H(\mathbf{a}) = \{1, 2, 3\}$
3. $H(\mathbf{a}) = \{1, 2, 3, \dots, n\}$
4. $H(\mathbf{a}) = \mathbb{Z}$
5. $H(\mathbf{a}) = [0, 1]$
6. $H(\mathbf{a}) = \mathbb{R}$

Nach dieser Übungsaufgabe könnte man auf den ersten Blick glauben, dass alle Teilmengen von \mathbb{R} als Menge $H(\mathbf{a})$ der Häufungspunkte einer geeigneten Folge \mathbf{a} auftreten können. Das ist aber nicht der Fall, weil $H(\mathbf{a})$ immer abgeschlossen sein muss. Das ist aber die einzige Einschränkung, wie die folgende, etwas anspruchsvollere Übungsaufgabe zeigt.

Übungsaufgabe 133. (E) Zeigen Sie:

1. Die Menge $H(\mathbf{a})$ aller Häufungspunkte einer reellen Folge \mathbf{a} ist abgeschlossen.
2. Ist umgekehrt $A \subseteq \mathbb{R}$ irgendeine abgeschlossene Teilmenge von \mathbb{R} , so gibt es eine reelle Folge \mathbf{a} mit $H(\mathbf{a}) = A$.

So wie zum Grenzwert formulieren wir auch zum Häufungspunkt noch die allgemeinere Definition für Folgen in beliebigen metrischen Räumen:

Definition 2.1.5.2. Sei X ein metrischer Raum mit Metrik d . Der Punkt $x \in X$ heißt **Häufungspunkt** der Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$, $a_n \in X$, wenn in jeder Umgebung U von x unendlich viele Glieder von \mathbf{a} liegen, mit anderen Worten: wenn es zu jeder Umgebung U von x und zu jedem $n_0 \in \mathbb{N}$ ein $n \geq n_0$ gibt mit $a_n \in U$.

Der Vollständigkeit halber bringen wir auch noch die Definition des Häufungspunktes einer Menge (statt einer Folge). Sie ist jener für Folgen sehr ähnlich ist, bezieht sich aber direkt auf die Elemente der Menge statt auf die Folgenindizes.

Definition 2.1.5.3. Sei X ein metrischer Raum mit einer Metrik d . Dann heißt ein Punkt $x \in X$ **Häufungspunkt** der Menge $M \subseteq X$, wenn in jeder Umgebung U von x unendlich viele Elemente von M liegen, mit anderen Worten: wenn es zu jedem $\varepsilon > 0$ ein $y \neq x$ gibt mit $y \in M$ und $d(x, y) < \varepsilon$, als logische Formel:

$$\forall \varepsilon > 0 \exists y \in M \setminus \{x\} : d(x, y) < \varepsilon$$

2.1.6 Obere und untere Limiten

Inhalt in Kurzfassung: Bei reellen Folgen spielen spezielle Häufungspunkte eine nützliche Rolle, nämlich oberer und unterer Grenzwert, genannt auch Limes superior und Limes inferior, symbolisch \limsup und \liminf . Der obere Grenzwert x einer reellen Folge mit den Gliedern a_n zeichnet sich dadurch aus, dass für jedes $\varepsilon > 0$ (wenn überhaupt) nur endliche viele a_n größer als $x + \varepsilon$ sind, aber unendlich viele a_n größer als $x - \varepsilon$ (symmetrisch der untere Grenzwert). Jeder Grenzwert ist gleichzeitig sowohl oberer als auch unterer Grenzwert. Nach oben/unten unbeschränkten Folgen ordnet man den oberen/unteren Grenzwert $\infty/-\infty$ zu. Gibt es für jedes $\alpha \in \mathbb{R}$ nur endlich viele n mit $a_n > \alpha$, so definiert man $\limsup_{n \rightarrow \infty} a_n = -\infty$, entsprechend $\liminf_{n \rightarrow \infty} a_n = \infty$, wenn für jedes $\alpha \in \mathbb{R}$ nur endlich oft $a_n < \alpha$ gilt.

Um oberen und unteren Limes zu illustrieren, betrachten wir zunächst als Beispiel die Folge \mathbf{a} mit den Gliedern $a_n = (-1)^n(1 + \frac{1}{n})$. Man sieht schnell: Die Glieder mit geradem Index nähern sich von oben der Zahl 1 an, die mit ungeradem Index von unten der Zahl -1 . Beide Werte sind Häufungspunkte von \mathbf{a} und, wovon man sich schnell überzeugt, sogar die einzigen. Somit ist 1 der größte Häufungspunkt der nach oben beschränkten Folge \mathbf{a} , der sogenannte **obere Grenzwert**, genannt auch **oberer Limes** oder **Limes superior**, symbolisch $\limsup_{n \rightarrow \infty} a_n$. Entsprechend heißt -1 als kleinster Häufungspunkt der nach unten beschränkten Folge \mathbf{a} der **untere Grenzwert**, genannt auch **unterer Limes** oder **Limes inferior**, symbolisch $\liminf_{n \rightarrow \infty} a_n$.

Es erweist sich als praktisch, auch unbeschränkte Folgen mit oberen und unteren Limiten zu versehen, die dann den Wert $\limsup_{n \rightarrow \infty} a_n = \infty$ (im nach oben unbeschränkten Fall) bzw. $\liminf_{n \rightarrow \infty} a_n = -\infty$ (im nach unten unbeschränkten Fall) zuzuweisen. Es folgt eine ausführliche Diskussion und Definition dieser Begriffe.

Definition 2.1.6.1. Sei $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ eine reelle Folge. Wir definieren den **oberen Grenzwert** $\limsup_{n \rightarrow \infty} a_n$, genannt auch **oberer Limes** oder **Limes superior**, von \mathbf{a} durch folgende Fallunterscheidung:

- 1.Fall: Sei \mathbf{a} nach oben beschränkt und gebe es ein $x \in \mathbb{R}$ mit folgender Eigenschaft: Zu jedem $\varepsilon > 0$ gibt es unendlich viele $n \in \mathbb{N}$ mit $a_n > x - \varepsilon$, gleichzeitig aber nur endlich viele mit $a_n \geq x + \varepsilon$. Dann setzen wir $\limsup_{n \rightarrow \infty} a_n := x$.
- 2.Fall: Sei \mathbf{a} nach oben beschränkt, es gebe aber kein x wie im 1.Fall. Dann setzen wir $\limsup_{n \rightarrow \infty} a_n := \infty$.
- 3.Fall: Sei \mathbf{a} nicht nach oben beschränkt. Dann setzen wir $\limsup_{n \rightarrow \infty} a_n := -\infty$.

Symmetrisch dazu definieren wir den **unteren Grenzwert**, genannt auch **unterer Limes** oder **Limes inferior**, symbolisch $\liminf_{n \rightarrow \infty} a_n$ von \mathbf{a} durch folgende Fallunterscheidung:

- 1.Fall: Sei \mathbf{a} nach unten beschränkt und gebe es ein $x \in \mathbb{R}$ mit folgender Eigenschaft: Zu jedem $\varepsilon > 0$ gibt es unendlich viele $n \in \mathbb{N}$ mit $a_n < x + \varepsilon$, gleichzeitig aber nur endlich viele mit $a_n \leq x - \varepsilon$. Dann setzen wir $\liminf_{n \rightarrow \infty} a_n := x$.
- 2.Fall: Sei \mathbf{a} nach unten beschränkt, es gebe aber kein x wie im 1.Fall. Dann setzen wir $\liminf_{n \rightarrow \infty} a_n := \infty$.
- 3.Fall: Sei \mathbf{a} nicht nach unten beschränkt. Dann setzen wir $\limsup_{n \rightarrow \infty} a_n := -\infty$.

Bemerkung: Man könnte \liminf auch kürzer mittels \limsup definieren durch $\liminf_{n \rightarrow \infty} a_n := -\limsup_{n \rightarrow \infty} (-a_n)$.

Übungsaufgabe 134. (P) Begründen Sie, warum

$$\liminf_{n \rightarrow \infty} a_n := -\limsup_{n \rightarrow \infty} (-a_n)$$

tatsächlich zum selben Ergebnis führt wie Definition 2.1.6.1.

Proposition 2.1.6.2. Sei $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ eine reelle Folge. Dann gilt:

1. Ist $x = \limsup_{n \rightarrow \infty} a_n \in \mathbb{R}$ eine reelle Zahl, so ist x ein Häufungspunkt von \mathbf{a} , und zwar der größte.
2. Ist $x = \liminf_{n \rightarrow \infty} a_n \in \mathbb{R}$ eine reelle Zahl, so ist x ein Häufungspunkt von \mathbf{a} , und zwar der kleinste.
3. Sowohl $\liminf_{n \rightarrow \infty} a_n$ als auch $\limsup_{n \rightarrow \infty} (-a_n)$ sind eindeutig bestimmt.

Beweis. 1. Sei $\varepsilon > 0$. Nach Definition von $x = \limsup_{n \rightarrow \infty} a_n \in \mathbb{R}$ gibt es unendlich viele n mit $a_n > x - \varepsilon$. Davon erfüllen aber nur endlich viele $a_n \geq x + \varepsilon$. Also muss $x - \varepsilon < x < x + \varepsilon$ für unendlich viele n gelten. Da dies für jedes $\varepsilon > 0$ gilt, ist x Häufungspunkt von \mathbf{a} . Es gibt aber sicher keinen größeren Häufungspunkt $x' > x$ von \mathbf{a} . Nehmen wir zum Beweis indirekt an, ein solches x' liege vor. Wir setzen $\varepsilon := \frac{x' - x}{2}$. Nach Definition eines Häufungspunktes müsste $x' - \varepsilon < a_n < x' + \varepsilon$, insbesondere also $a_n > x' - \varepsilon = x + \varepsilon$ für unendlich viele n gelten. Das widerspricht aber der Definition von $x = \limsup_{n \rightarrow \infty} a_n$.

2. Symmetrisch zu 1.

3. Definition 2.1.6.1 ist im 2.Fall und im 3.Fall offensichtlich eindeutig. Im 1.Fall führt sie zu einer reellen Zahl, die der größte bzw. kleinste Häufungspunkt von \mathbf{a} ist, wegen der Totalordnung von \mathbb{R} also ebenfalls eindeutig bestimmt ist.

□

Bemerkung: Die dritte Aussage von Proposition 2.1.6.2 inkludiert noch nicht, dass auch im 3. Fall $\limsup_{n \rightarrow \infty} a_n$ und $\liminf_{n \rightarrow \infty} a_n$ existieren. Das werden wir erst in 2.1.7.4 beweisen.

Für die Symbole $-\infty$ und ∞ und alle $r \in \mathbb{R}$ vereinbaren wir wie gewohnt die Schreibweise $-\infty < r < \infty$. Man beachte aber, dass es sich bei ∞ und $-\infty$ tatsächlich nur um Symbole handelt, nicht um reelle Zahlen. In 2.1.7 werden wir sehen, dass auf diese Weise \limsup und \liminf (im Gegensatz zu \lim) für alle reellen Folgen definiert sind. Im Fall $\liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n = \infty$ schreiben wir auch $\lim_{n \rightarrow \infty} a_n = \infty$ und sprechen von **Divergenz gegen ∞** , manchmal auch von **uneigentlicher Konvergenz**; entsprechend für $-\infty$ statt ∞ .

Übungsaufgabe 135. (P) Geben Sie eine Folge mit paarweise verschiedenen Gliedern a_n an, für die gilt:

$$\inf\{a_n : n \in \mathbb{N}\} < \liminf_{n \rightarrow \infty} a_n < \limsup_{n \rightarrow \infty} a_n < \sup\{a_n : n \in \mathbb{N}\}.$$

Sehr leicht beweist man Rechenregeln, für die man salopp schreibt: $\infty + c = \infty$, $\infty + \infty = \infty$, $c \cdot \infty = \infty$ (für $0 < c \in \mathbb{R}$), $\frac{c}{\infty} = 0$, (für $0 \neq c \in \mathbb{R}$) etc. Die erste davon soll bedeuten: Divergieren die a_n gegen ∞ und konvergieren die b_n gegen $c \in \mathbb{R}$, so divergiert $a_n + b_n$ gegen ∞ etc. Man mache sich klar, dass aber beispielsweise über $\infty - \infty$, $\frac{\infty}{\infty}$ und $\frac{c}{0}$ keine allgemeingültigen Aussagen gemacht werden können. Man spricht in diesem Zusammenhang von **unbestimmten Formen**.

Übungsaufgabe 136. (P) Sei $r \in \mathbb{R}$ eine vorgegebene reelle Zahl.

1. Geben Sie zwei Folgen mit Gliedern $a_n \rightarrow \infty$ bzw. $b_n \rightarrow \infty$ an, so dass für die Differenzen $d_n := a_n - b_n$ gilt:

$$(i) d_n \rightarrow -\infty \quad (ii) d_n \rightarrow r \quad (iii) d_n \rightarrow \infty$$

2. Analog für die Quotienten $q_n := \frac{a_n}{b_n}$ statt der Differenzen d_n .
3. Wie Teilaufgabe 1, nur mit $a_n \rightarrow 0$ und $b_n \rightarrow 0$.
4. Wie Teilaufgabe 3, nur mit den Quotienten q_n aus Teil 2 statt der Differenzen d_n .

2.1.7 Die wichtigsten Sätze aus der Theorie der Folgen

Inhalt in Kurzfassung: In diesem Abschnitt werden die für die Theorie reeller Folgen wichtigsten Tatsachen zusammengestellt und bewiesen. Unterschieden wird dabei hinsichtlich der Rolle der Vollständigkeit von \mathbb{R} , eine für die reelle Analysis entscheidende Eigenschaft. Zuerst kommen Aussagen, die unabhängig von der Vollständigkeit gelten, dann jene, die von ihr abhängen. Die wichtigsten Beispiele für letzteres lauten: Jede monotone und beschränkte Folge konvergiert.

Jede beschränkte Folge hat einen Häufungspunkt und daher eine konvergente Teilfolge (Bolzano-Weierstraß). Jede Cauchyfolge in \mathbb{R} konvergiert. Ebenfalls eine Konsequenz der Vollständigkeit ist die uneingeschränkte Existenz von oberem und unterem Grenzwert (eigentlich oder uneigentlich) von reellen Folgen.

Wir wollen nun die wichtigsten Zusammenhänge zwischen den verschiedenen Eigenschaften von reellen Folgen beweisen und möglichst übersichtlich zusammenstellen. Bei einigen davon spielt die Vollständigkeit von \mathbb{R} eine entscheidende Rolle. Um dies deutlicher sichtbar zu machen, unterscheiden wir nach diesem Gesichtspunkt. Wir beginnen zunächst mit (typischerweise eher elementaren) Aussagen, wo die Vollständigkeit keine Rolle spielt und die deshalb typischerweise auch in viel allgemeinerem Kontext gelten (Satz 2.1.7.1). Erst danach arbeiten wir jene wichtigen Sätze (??) heraus, die für \mathbb{R} spezifisch sind oder für die bei allfälligen Verallgemeinerungen (etwa auf beliebige metrische Räume) zumindest eine der Vollständigkeit von \mathbb{R} entsprechende Eigenschaft vorausgesetzt werden muss. Sodann ziehen wir Folgerungen für \limsup und \liminf (Satz 2.1.7.4).

Zur besseren Übersicht wiederholen wir auch einige Tatsachen, die schon an früherer Stelle bemerkt worden sind.

Satz 2.1.7.1. *Die folgenden Aussagen über reelle Folgen gelten auch für Folgen in beliebigen metrischen Räumen. Dazu bezeichne $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ eine reelle Folge oder allgemeiner eine Folge in einem metrischen Raum X mit einer Metrik d . In \mathbb{R} kann wie üblich die Metrik $d(x, y) = |x - y|$ genommen werden.*

1. *Konvergiert eine Folge, so ist ihr Grenzwert auch ein Häufungspunkt.*
2. *Hat eine Folge mehr als einen Häufungspunkt, so divergiert sie.*
3. *Der Grenzwert einer konvergenten Folge ist eindeutig bestimmt.*
4. *Jede Teilfolge a_{n_0}, a_{n_1}, \dots ($n_0 < n_1 < n_2 < \dots$) einer konvergenten Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ ist konvergent gegen denselben Grenzwert $\lim_{k \rightarrow \infty} a_{n_k} = \lim_{n \rightarrow \infty} a_n$.*
5. *Jede konvergente Folge ist eine Cauchyfolge.*
6. *Jede Cauchyfolge ist beschränkt.*
7. *Zu jedem Häufungspunkt einer Folge gibt es eine konvergente Teilfolge.*
8. *Ist x ein Häufungspunkt der Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$, $a_n \in A$, und ist $A \subseteq X$ abgeschlossen, so folgt $x \in A$.*

Beweis. Zur Erinnerung aus 1.5.9: Ist in den nachfolgenden Beweisen der einzelnen Aussagen von Umgebungen U eines Punktes $x \in X$ die Rede, so sind damit Mengen gemeint, zu denen es ein $\varepsilon > 0$ gibt derart, dass alle $y \in X$ mit $d(x, y) < \varepsilon$ in U liegen. Im Fall $X = \mathbb{R}$ denkt man sich dabei am einfachsten das offene Intervall $(x - \varepsilon, x + \varepsilon)$ um x mit „Radius“ ε , also mit Länge 2ε .

1. Konvergiere \mathbf{a} gegen x , sei U eine Umgebung von x und $n_0 \in \mathbb{N}$ beliebig. Zu zeigen ist, dass es ein $n \in \mathbb{N}$ gibt mit $n \geq n_0$ und $a_n \in U$. Weil \mathbf{a} gegen x konvergiert, gibt es ein $n_1 \in \mathbb{N}$, so dass $a_n \in U$ für alle $n \geq n_1$. Sei $n := \max\{n_0, n_1\}$ das Maximum von n_0 und n_1 . Dieses n hat die gewünschte Eigenschaft $a_n \in U$.
2. Seien x_1 und x_2 seien Häufungspunkte, x_1 sogar Grenzwert der konvergenten Folge \mathbf{a} . Es genügt, die indirekte Annahme $x_1 \neq x_2$ zu einem Widerspruch zu führen. Tatsächlich folgt aus dieser Annahme $d := d(x_1, x_2) > 0$, somit auch $\varepsilon := \frac{d}{2} > 0$. Nach Definition des Grenzwertes müsste es dann ein n_1 geben mit $d(a_n, x_1) < \varepsilon$ für alle $n \geq n_1$. Nach Definition des Häufungspunktes gibt es außerdem ein $n \geq n_1$ mit $d(a_n, x_2) < \varepsilon$. Dann folgt laut Dreiecksungleichung der Widerspruch

$$d = d(x_1, x_2) \leq d(x_1, a_n) + d(a_n, x_2) < \varepsilon + \varepsilon = \frac{d}{2} + \frac{d}{2} = d.$$

3. Sind x_1 und x_2 Grenzwerte der Folge \mathbf{a} , so nach dem ersten Teil auch Häufungspunkte. Konvergiert \mathbf{a} , so ist das nach der zweiten Aussage aber nur für $x_1 = x_2$ möglich.
4. Wir denken an die Formulierung mit „fast alle“ für „höchstens endlich viele Ausnahmen“: Konvergiert die Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ sämtlicher Glieder a_n , $n \in \mathbb{N}$, gegen x , so gibt es zu einer beliebigen Umgebung U von x höchstens endlich viele n mit $x \notin U$. Daran kann sich aber nichts ändern, wenn man gewisse n aus der Betrachtung streicht. Also liegen auch fast alle Glieder der Teilfolge in U . Weil dies für jede Umgebung U von x gilt, bedeutet dies, dass x auch Grenzwert der Teilfolge ist.
5. Sei wieder x Grenzwert von \mathbf{a} . Um zu zeigen, dass \mathbf{a} eine Cauchyfolge ist, sei $\varepsilon > 0$ vorgegeben. Nach Definition des Grenzwertes gibt es einen Index $n_0 \in \mathbb{N}$ mit $d(a_n, x) < \frac{\varepsilon}{2}$ für alle $n \geq n_0$. Dann gilt für den Abstand von je zwei beliebigen Folgengliedern mit Indizes $n_1, n_2 \geq n_0$ wegen der Dreiecksungleichung tatsächlich

$$d(a_{n_1}, a_{n_2}) \leq d(a_{n_1}, x) + d(x, a_{n_2}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

6. Sei \mathbf{a} eine Cauchyfolge. Dann gibt es ein n_0 mit $d(a_{n_1}, a_{n_2}) < 1$ für alle $n_1, n_2 \geq n_0$. Sei r die größte der Zahlen $d(a_0, a_{n_0}) + 1, d(a_1, a_{n_0}) + 1, \dots, d(a_{n_0-1}, a_{n_0}) + 1$. Dann gilt $d(a_n, a_{n_0}) < r$ für alle $n \in \mathbb{N}$, also liegen alle Folgenglieder in der (beschränkten) Kugel um a_{n_0} mit Radius r , bilden also eine beschränkte Folge.
7. Sei x ein Häufungspunkt der Folge \mathbf{a} . Beginnend mit einem $k_1 \in \mathbb{N}$ mit $d(a_{k_1}, x) < 1$ liefert eine iterierte Anwendung der Definition des Häufungspunktes eine unendliche Folge aufsteigender Indizes $k_1 < k_2 < \dots$ mit $d(a_{k_n}, x) < \frac{1}{n+1}$. Somit bilden die $a_{k_1}, a_{k_2} \dots$ eine Teilfolge von \mathbf{a} , die gegen x konvergiert.

8. Weil A abgeschlossen ist, genügt es zu zeigen, dass in jeder Umgebung U von x Folgenglieder und damit Elemente von A liegen. Das folgt aber unmittelbar daraus, dass x Häufungspunkt der $a_n \in A$ ist.

□

Da die Aussagen aus Satz 2.1.7.1 für Folgen in beliebigen metrischen Räumen gelten, hätten wir statt $X = \mathbb{R}$ z.B. auch $X = \mathbb{Q}$ nehmen können. Das ist anders, sobald die Vollständigkeit von \mathbb{R} verwendet wird, wie im nachfolgenden Satz.

Satz 2.1.7.2. *Die folgenden Aussagen gelten für reelle Folgen (jedoch nicht allgemein für Folgen in beliebigen metrischen Räumen).*

1. *Jede beschränkte und monotone reelle Folge konvergiert. (**Satz von der Konvergenz durch Monotonie und Beschränktheit**) Folglich ist eine monotone reelle Folge genau dann beschränkt, wenn sie konvergiert.*
2. *Jede beschränkte reelle Folge hat einen Häufungspunkt in \mathbb{R} . (Analog hat jede unendliche beschränkte Teilmenge von \mathbb{R} einen Häufungspunkt.) Insbesondere hat jede beschränkte reelle Folge eine konvergente Teilfolge. (**Satz von Bolzano-Weierstraß**)*
3. *Jede reelle Cauchyfolge konvergiert. Folglich sind in \mathbb{R} genau die Cauchyfolgen die konvergenten. (**Cauchy Kriterium**)*

Beweis. 1. Wir wissen bereits, dass jede konvergente Folge beschränkt ist. Zu zeigen ist daher nur mehr die erste Aussage, dass nämlich jede monotone und beschränkte Folge konvergiert. Wir führen den Beweis für *monoton wachsend*, der andere Fall ist ganz analog. Wegen der Supremumseigenschaft gibt es das Supremum s aller a_n . Dieses s erweist sich schnell als Grenzwert. Denn für beliebiges $\varepsilon > 0$ ist $s - \varepsilon$ nicht mehr obere Schranke der a_n . Also gibt es ein n_0 mit $a_{n_0} > s - \varepsilon$. Dann muss wegen der Monotonie aber für alle $n \geq n_0$ gelten: $s - \varepsilon < a_{n_0} \leq a_n \leq s$, folglich $|a_n - s| < \varepsilon$.

2. Ist die Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ beschränkt, so gibt es Zahlen $x_0 \leq y_0 \in \mathbb{R}$ mit $x_0 \leq a_n \leq y_0$ für alle $n \in \mathbb{N}$. Wir wählen die Folge von Intervallen $I_k = [x_k, y_k]$ mit $I_0 = [x_0, y_0] \supseteq I_1 \supseteq I_2 \dots \supseteq I_k \supseteq I_{k+1} \supseteq \dots$ derart, dass jedes I_{k+1} die linke oder die rechte Hälfte von I_k ist und $a_n \in I_{k+1}$ für unendlich viele n . Dass dies in jedem Schritt möglich ist, folgt unmittelbar mittels Induktion nach k : In I_0 liegen sogar alle Folgenglieder, insbesondere also unendlich viele; und wenn I_k unendlich viele Folgenglieder enthält, so muss das auch für wenigstens eine der beiden Hälften von I_k , welche dann als I_{k+1} genommen werden kann, der Fall sein. Die Längen der I_k halbieren sich in jedem Schritt, das heißt $y_k - x_k = \frac{y_0 - x_0}{2^k} \rightarrow 0$ für $k \rightarrow \infty$. Wir erhalten eine monoton wachsende Folge $x_0 \leq x_1 \leq x_2 \leq \dots$ und eine monoton fallende Folge $y_0 \geq y_1 \geq y_2 \geq \dots$. Beide Folgen sind beschränkt (die der x_k nach oben durch die y_k , die der y_k nach unten durch die x_k), also existieren laut erstem Teil dieses Satzes (Satz von der Konvergenz

durch Monotonie und Beschränktheit) die Grenzwerte $x = \lim_{k \rightarrow \infty} x_k$ und $y = \lim_{k \rightarrow \infty} y_k$. Wegen $x - y = \lim_{k \rightarrow \infty} x_k - \lim_{k \rightarrow \infty} y_k = \lim_{k \rightarrow \infty} (x_k - y_k) = 0$ folgt $x = y$. Um zu zeigen, dass x Häufungspunkt der a_n ist, sei $\varepsilon > 0$. Für hinreichend großes k ist $x - \varepsilon < x_k < y_k < x + \varepsilon$. Laut Konstruktion gilt $a_n \in I_k = [x_k, y_k]$ für unendlich viele n . Alle diese n erfüllen aber erst recht $|a_n - x| < \varepsilon$. Weil $\varepsilon > 0$ beliebig war, ist x ein Häufungspunkt von \mathbf{a} .

3. Dass jede konvergente Folge eine Cauchyfolge ist, wissen wir schon aus Satz 2.1.7.1, Aussage 5. Sei daher umgekehrt \mathbf{a} eine Cauchyfolge. Nochmals wegen Satz 2.1.7.1, diesmal Aussage 6, ist \mathbf{a} beschränkt. Nach der soeben bewiesenen Aussage 2 (Satz von Bolzano-Weierstraß) hat \mathbf{a} daher einen Häufungspunkt x . Wir zeigen nun, dass x sogar Grenzwert ist. Sei dazu $\varepsilon > 0$ vorgegeben. Weil \mathbf{a} eine Cauchyfolge ist, gibt es ein $n_0 \in \mathbb{N}$ mit der Eigenschaft, dass $|a_{n_1} - a_{n_2}| < \frac{\varepsilon}{2}$ für alle $n_1, n_2 \geq n_0$. Weil x ein Häufungspunkt von \mathbf{a} ist, gibt es ein $n_3 \geq n_0$ mit $|a_{n_3} - x| < \frac{\varepsilon}{2}$. Für alle $n \geq n_0$ gilt somit $|a_n - x| \leq |a_n - a_{n_3}| + |a_{n_3} - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. \square

Will man die Ergebnisse aus Satz 2.1.7.2 auch für Folgen in allgemeineren metrischen Räumen beweisen, braucht man typischerweise einen entsprechenden Vollständigkeitsbegriff für metrische Räume. Einen solchen erhält man, indem man für einen **vollständigen metrischen Raum** *per definitionem* fordert, dass alle Cauchyfolgen konvergieren – eine Eigenschaft, die wir für den metrischen Raum $X = \mathbb{R}$ mit der Metrik $d(x, y) = |x - y|$ gleich beweisen werden. Die höherdimensionalen euklidischen Räume \mathbb{R}^n haben diese Eigenschaft:

Satz 2.1.7.3. *Die Räume \mathbb{R}^n , $n \in \mathbb{N}$, sind mit der euklidischen Metrik vollständige metrische Räume.*

Beweis. (Skizze) Zu zeigen ist, dass jede Cauchyfolge in \mathbb{R}^n konvergiert. Für den euklidischen Abstand d_n von zwei Punkten $\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_2) \in \mathbb{R}^n$ gilt offensichtlich $d(\mathbf{x}, \mathbf{y}) \geq |x_i - y_i| = d_1(x_i, y_i)$ für jede Komponenten $i = 1, \dots, n$. Hält man sich die Definition einer Cauchyfolge vor Augen, so zeigt dies: Liegt eine Cauchyfolge in \mathbb{R}^n vor, so bilden die i -ten Komponenten der Folgenglieder eine Cauchyfolge in \mathbb{R} . In \mathbb{R} konvergiert aber jede Cauchyfolge. Somit gibt es ein $\mathbf{x} \in \mathbb{R}^n$, dessen Komponenten die Grenzwerte dieser Komponentenfolgen sind. Daraus folgt sehr schnell (Übung), dass die ursprünglich vorgegebene Folge in \mathbb{R}^n gegen dieses \mathbf{x} konvergiert. \square

Übungsaufgabe 137. (P) *Ergänzen Sie die Details im letzten, nicht vollständig ausgeführten Schritt im Beweis von Satz 2.1.7.3.*

Auch die bereits früher angekündigte Existenz von Limes superior und Limes inferior für beliebige reelle Folgen ist eine Konsequenz der Vollständigkeit von \mathbb{R} :

Satz 2.1.7.4. *Sei $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ eine beliebige Folge in \mathbb{R} . Dann gilt:*

1. Der obere Grenzwert $\limsup_{n \rightarrow \infty} a_n$ existiert, entweder im eigentlichen Sinn (d.h. als reelle Zahlen) oder im uneigentlichen Sinn (als ∞ oder $-\infty$). Nach Proposition 2.1.6.2, Aussage 3, ist er sogar eindeutig bestimmt. Dasselbe gilt für $\liminf_{n \rightarrow \infty} a_n$.
2. Die Folge \mathbf{a} konvergiert genau dann, wenn $\limsup_{n \rightarrow \infty} a_n$ und $\liminf_{n \rightarrow \infty} a_n$ im eigentlichen Sinn, d.h. als reelle Zahlen, existieren und übereinstimmen.

Beweis. 1. Aus Symmetriegründen genügt der Beweis für \limsup . Drei Fälle sind zu unterscheiden.

Fall 1: Gibt es zu jedem $x \in \mathbb{R}$ nur endliche viele $n \in \mathbb{N}$ mit $a_n > x$, so ist definitionsgemäß $\limsup_{n \rightarrow \infty} a_n = -\infty$.

Fall 2: Ist \mathbf{a} nicht nach oben beschränkt, so ist $\limsup_{n \rightarrow \infty} a_n = \infty$.

Fall 3: Liege weder Fall 1 noch Fall 2 vor. Dann gibt es eine reelle Zahl x_0 mit $a_n > x_0$ für unendlich viele n (weil nicht Fall 1). Andererseits ist \mathbf{a} nach oben beschränkt (weil nicht Fall 2). Sei M die Menge aller $x \in \mathbb{R}$ mit $a_n > x$ für unendlich viele n . Wegen $x_0 \in M$ ist M nicht leer. Weil wir nicht in Fall 2 sind, ist M nach oben beschränkt. Nach dem Supremumsprinzip gibt es daher $s := \sup M \in \mathbb{R}$. Dieses s erfüllt die Definition von $\limsup_{n \rightarrow \infty} a_n$. Um das zu beweisen geben wir uns ein beliebiges $\varepsilon > 0$ vor und beweisen, dass $a_n > s - \varepsilon$ für unendlich viele n (Behauptung 1), $a_n > s + \varepsilon$ aber nur für endlich viele n (Behauptung 2) gilt.

Beweis von Behauptung 1: Weil s die kleinste obere Schranke von M ist, gibt es ein $m \in M$ mit $m > s - \varepsilon$. Nach Definition von M gilt $a_n > m > s - \varepsilon$ für unendlich viele n .

Beweis von Behauptung 2: Weil s obere Schranke von M ist, folgt $s + \varepsilon \notin M$. Also gibt es nur endlich viele n mit $a_n > s + \varepsilon$.

2. Konvergiert \mathbf{a} gegen $x \in \mathbb{R}$, so erfüllt x als Grenzwert auch die Definition von $\limsup_{n \rightarrow \infty} a_n$ und $\liminf_{n \rightarrow \infty} a_n$. Seien umgekehrt beide Definitionen für $x \in \mathbb{R}$ erfüllt. Wir haben $x = \lim_{n \rightarrow \infty} a_n$ zu zeigen. Sei dazu $\varepsilon > 0$ beliebig. Dann gelten die Ungleichungen $a_n \geq x + \varepsilon$ und $a_n \leq x - \varepsilon$ jeweils nur für endlich viele n , die erste wegen $x = \limsup_{n \rightarrow \infty} a_n$, die zweite wegen $x = \liminf_{n \rightarrow \infty} a_n$. Umgekehrt bedeutet das aber: $x - \varepsilon < a_n < x + \varepsilon$ für fast alle n , was zu zeigen war. □

Übungsaufgabe 138. (E)

1. Im Beweis von Satz 2.1.7.4, Fall 3, wurde $s = \sup M = \limsup_{n \rightarrow \infty} a_n$ behauptet. Begründen Sie das sorgfältig.
2. Im Beweis von Satz 2.1.7.4 wurde aus Symmetriegründen auf den Beweis der \liminf -Aussage verzichtet. Beweisen Sie diese.

2.1.8 Rekursive Folgen und Iterationsfolgen

Inhalt in Kurzfassung: Rekursive Folgen entstehen, wenn man, mit einem bestimmten Startelement beginnend, jedes weitere Folgenglied mit Hilfe der vorangegangenen Glieder definiert. Der Spezialfall der Iterationsfolge liegt vor, wenn diese Definition immer nur auf das unmittelbare Vorgängerglied zugreift, und zwar unter Verwendung einer gleichbleibenden Transformation T , d.h. $a_{n+1} := T(a_n)$. Im Zusammenhang gewinnen periodische Punkte, Fixpunkte und Spinnwebdiagramme an Interesse.

Rekursive Folgen haben wir bereits im Zusammenhang mit dem Rekursionsatz kennen gelernt. Sie sind so definiert, dass zumindest der Startwert (das nullte Folgenglied a_0) gegeben ist und darüber hinaus erklärt wird, wie man jeweils das nächste aus den bisherigen erhält. Hängt jedes Folgenglied nur vom unmittelbar vorangehenden ab, liegt eine eingliedrige Rekursion vor: $a_{n+1} = T(a_n)$ für alle $n \in \mathbb{N}$ mit einer Funktion $T : D \rightarrow D$ auf einem geeigneten Definitionsbereich D .¹ Definieren wir die Iterationen von T durch $T^{(0)} := \text{Id}_D$ (Identität auf D) und $T^{(n+1)} := T \circ T^{(n)}$, so gilt offenbar $a_n = T^{(n)}(a_0)$, weshalb wir auch von **Iterationsfolgen**² sprechen.

Ist $a_p = a_0$ und $p > 0$, so ist die Iterationsfolge **a** periodisch mit (nicht notwendig kleinster) Periodenlänge p . Man nennt so ein $a = a_0$ einen **periodischen Punkt** von T . Ist $p = 1$, so ist a sogar ein **Fixpunkt** von T .

Im Spezialfall reeller Folgen ist $D \subseteq \mathbb{R}$ und T eine reelle Funktion. Häufig nähern sich die Folgenglieder einem bestimmten Wert an. Der Mechanismus lässt sich sehr gut durch sogenannte **Spinnwebdiagramme** illustrieren.

Erstes Beispiel: Ist $D = [0, \infty)$ und $T(x) := \sqrt{x}$, so kann ein beliebiges positives $a_0 > 0$ gewählt werden, stets werden sich die a_n mit wachsendem n dem Wert 1 annähern, der ein Fixpunkt ist: $T(1) = 1$. Startet man mit $a_0 > 1$, erfolgt die Annäherung von rechts, bei $a_0 < 1$ von links.

Übungsaufgabe 139. (P) Testen Sie anhand verschiedener Startwerte (wenigstens mit einem $a_0 < 1$ und einem $a_0 > 1$) das oben beschriebene Verhalten der Iterationsfolge für $T : x \mapsto \sqrt{x}$ und versuchen Sie es zu begründen.

Zweites Beispiel: Auch bei $D = \mathbb{R}$ und $T(x) := 1 - \frac{x}{2}$ erfolgt bei beliebigem Startwert Annäherung an den Fixpunkt $\frac{2}{3}$, in diesem Fall jedoch alternierend von rechts und links.

Übungsaufgabe 140. (P) Analog Aufgabe 139 mit $T(x) := 1 - \frac{x}{2}$.

Übungsaufgabe 141. (P) Analog Aufgabe 139, aber mit der Transformation $T(x) = \sqrt{x - \frac{1}{4}}$ und den Startwerten $a_0 = \frac{1}{2}$ bzw. $a_0 = 2$.

¹Funktionen, wo Definitionsbereich und Zielbereich übereinstimmen, nennt man oft auch **Transformationen**, deshalb in diesem Zusammenhang meist die Bezeichnung T .

²Trotz derselben Notation $T^{(n)}$ darf die hier damit bezeichnete n -te Iteration nicht mit der n -ten Ableitungen verwechselt werden, wie sie später in 4.2.3 definiert wird.

Ein tieferes Verständnis dieses Phänomens ergibt sich aus Kombination zweier sehr grundlegender Begriffe, die eng miteinander verwandt sind. Beim ersten handelt es sich um den an dieser Stelle bereits bekannten Grenzwert von Folgen. Der zweite ist die Stetigkeit reeller Funktionen. Sie wird im Zentrum des nächsten Kapitels stehen.

An dieser Stelle befassen wir uns vorerst aber in einer Übungsaufgabe mit einem prominenten Beispiel einer zweigliedrigen Rekursion.

Übungsaufgabe 142. (E) *Eingangs war von eingliedrigen Rekursionen die Rede. Ein berühmtes Beispiel einer zweigliedrigen Rekursion führt zu den sogenannten Fibonacci-Zahlen³. Sie beginnen mit den beiden Werten $F_0 = F_1 := 1$ und folgen der zweigliedrigen Rekursion $F_{n+1} := F_{n-1} + F_n$. Berechnen Sie einige Glieder der Folge, bis Sie eine wohlbegründete Vermutung über die Größenordnung des Wachstums der Glieder F_n aufstellen können. Zur Auswahl stehen zum Beispiel lineares Wachstum $F_n \approx kn + d$ (k und d geeignete reelle Zahlen), quadratisches Wachstum $F_n \approx n^2$, Wachstum mit einer höheren k -ten Potenz $F_n \approx n^k$, exponentielles Wachstum $F_n \approx a^n$ ($a \in \mathbb{R}$ geeignet), superexponentielles Wachstum wie z.B. $F_n \approx a^{n^2}$, $F_n \approx n!$, $F_n \approx n^n$ etc. Versuchen Sie Ihre Vermutung so weit wie möglich zu präzisieren und zu begründen. (Hier wird keine perfekte Antwort erwartet, sondern dass Sie sich eigenständige und sinnvolle Gedanken gemacht haben.)*

2.2 Reihen

Eine Reihe ergibt sich durch Summation der Glieder einer Folge. Formal sind Reihen selbst wieder Folgen – nämlich Folgen von Partialsummen. Umgekehrt kann jede Folge als Reihe interpretiert werden, wo nämlich die Differenzen aufeinanderfolgender Glieder aufsummiert werden. Grundsätzlich müsste man zwischen Folgen und Reihen also gar nicht unterscheiden. Trotzdem treten unter dem Gesichtspunkt des Summierens ganz typische Fragen und Phänomene auf, deretwegen die Theorie der Reihen einen eigenständigen und sehr wichtigen Platz in der Analysis einnimmt. Nach den grundlegenden Definitionen in 2.2.1 lernen wir in 2.2.2 einige Beispiele kennen. 2.2.3 bringt wichtige Konvergenzaussagen, 2.2.4 Spezifika absoluter Konvergenz und 2.2.5 Wissenswertes über alternierende Reihen.

2.2.1 Der Wert einer unendlichen Reihe

Inhalt in Kurzfassung: Die Partialsummen s_n einer Reihe entstehen durch fortgesetzte Addition der Glieder a_n , also $s_0 := a_0$ und $s_{n+1} := s_n + a_{n+1}$. Konvergiert die Folge der s_n gegen ein $\alpha \in \mathbb{R}$, so heißt α der Wert der Reihe $\sum_{n=0}^{\infty} a_n$ und die Reihe konvergent. Mit dieser Definition wird die Theorie der Folgen (angewendet auf die s_n) auch auf Reihen anwendbar. Zusätzliche Aspekte kommen

³nach dem italienischen Mathematiker Leonardo Fibonacci (um 1170 bis nach 1240), der damit die Population sich von Generation zu Generation vermehrender Kaninchen beschrieb

zum Beispiel durch den stärkeren Begriff der absoluten Konvergenz (Konvergenz der Reihe mit den Gliedern $|a_n|$) ins Spiel, welches im Vergleichskriterium (konvergente Majoranten, divergente Minoranten) auftritt. Die Glieder einer konvergenten Reihe bilden stets eine Nullfolge.

Wir gehen wieder aus von einer reellen Folge $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$, interessieren uns diesmal aber nicht nur für die gegebenen Folgenglieder a_n , sondern auch für deren sogenannte **Partial-** oder **Teilsommen**

$$s_n := \sum_{k=0}^n a_k.$$

Beginnen die a_n nicht mit dem Index $n = 0$ sondern mit einem anderen n_0 , so beginnt auch die Summation entsprechend mit $k = n_0$ statt mit $k = 0$, und auch die nachfolgenden Überlegungen sind entsprechend anzupassen. Meist ist $n_0 = 0$ oder $n_0 = 1$.

Übungsaufgabe 143. (T) Finden Sie möglichst einfache Ausdrücke für die Teilsommen der Folgen

1. $a_n = n$
2. $b_n = (-1)^n$
3. $c_n = \frac{(-1)^n}{2^n}$
4. $d_n = \frac{1}{n(n+1)}$
5. $e_n = \frac{1+(-2)^n}{2^n}$

Hinweis: Werfen Sie für 5. einen Blick auf Übungsaufgabe 146.

Wir wollen das Konvergenzverhalten einer Reihe mit den Gliedern a_n in Beziehung setzen. Sind beispielsweise die $a_n \geq 0$ nichtnegativ, so ist die Folge der s_n monoton wachsend. Für die s_n sind dann Beschränktheit und Konvergenz äquivalent. Konvergieren die s_n , so nennt man ihren Grenzwert den **Wert der unendlichen Reihe** mit den **Gliedern** oder **Summanden** a_n und schreibt:

$$s = \sum_{n=0}^{\infty} a_n := \lim_{n \rightarrow \infty} s_n.$$

Man sagt, die Reihe **konvergiert** gegen s , und die Reihe heißt **konvergent**. Man sagt, die Reihe **divergiert** bzw. ist **divergent**, wenn die Folge der s_n divergiert.

Auf die Folge $\mathbf{s} = (s_n)_{n \in \mathbb{N}}$ der Partialsummen können nun alle Begriffe und Überlegungen, die wir schon von Folgen kennen, angewandt werden, z.B. gilt:

$$\sum_{n=0}^{\infty} (a_n + b_n) = \sum_{n=0}^{\infty} a_n + \sum_{n=0}^{\infty} b_n$$

und

$$\sum_{n=0}^{\infty} c a_n = c \sum_{n=0}^{\infty} a_n,$$

sofern die Reihen mit den Gliedern a_n und b_n konvergieren.

Übungsaufgabe 144. (P) Beweisen Sie diese beiden Formeln mit Hilfe geeigneter Grenzwertsätze.

Die Reihe $\sum_{n=0}^{\infty} a_n$ heißt **absolut konvergent**, wenn sogar die Reihe mit den Gliedern $|a_n|$ konvergiert, symbolisch

$$\sum_{n=0}^{\infty} |a_n| < \infty$$

Die Reihe heißt **bedingt konvergent**, wenn nur $\sum_{n=0}^{\infty} a_n$ konvergiert, nicht aber $\sum_{n=0}^{\infty} |a_n|$. Das Wort *sogar* in diesem Satz erklärt sich durch folgenden Satz.

Satz 2.2.1.1. *Jede absolut konvergente Reihe ist konvergent.*

Dieser Satz ist eine unmittelbare Konsequenz des nun folgenden, in zwei äquivalenten Formulierungen präsentierten **Vergleichskriteriums**,⁴ wenn man dort $b_n := |a_n|$ setzt.

Satz 2.2.1.2. *Angenommen, über die Reihen mit den Gliedern a_n bzw. b_n sei bekannt, dass $|a_n| \leq b_n$ für alle $n \in \mathbb{N}$ gilt. Dann gelten die beiden folgenden (zueinander offenbar äquivalenten) Implikationen.*

1. *Konvergiert die Reihe $\sum_{n=0}^{\infty} b_n$, so auch die Reihe $\sum_{n=0}^{\infty} a_n$, und zwar sowohl im gewöhnlichen wie auch im absoluten Sinn. Man nennt die Reihe der b_n auch eine **konvergente Majorante** für die Reihe der a_n (**Majorantenkriterium**).*
2. *Divergiert die Reihe $\sum_{n=0}^{\infty} a_n$, so auch die Reihe $\sum_{n=0}^{\infty} b_n$. Man nennt die Reihe der a_n auch eine **divergente Minorante** für die Reihe der b_n (**Minorantenkriterium**).*

Beweis. Es genügt, das Majorantenkriterium zu beweisen. Sei dazu wie bisher $s_n := \sum_{k=0}^n a_k$. Nach dem Cauchyschen Konvergenzkriterium 2.1.7.2 genügt es, wenn wir zu beliebig vorgegebenem $\varepsilon > 0$ einen Index $n_0 \in \mathbb{N}$ finden mit $|s_{n_2} - s_{n_1}| < \varepsilon$ für alle $n_1, n_2 \geq n_0$. Wir bezeichnen mit $S_n := \sum_{k=0}^n b_k$ die Partialsummen der Reihe der b_n . Nach Voraussetzung konvergieren die S_n , bilden also eine Cauchy-Folge. Definitionsgemäß gibt es daher ein n_0 mit $|S_{n_2} - S_{n_1}| < \varepsilon$ für alle $n_1, n_2 \geq n_0$. Somit sind wir mit dem Beweis fertig, wenn wir $|s_{n_2} - s_{n_1}| \leq |S_{n_2} - S_{n_1}|$ zeigen können. Und tatsächlich ergibt sich das (o.B.d.A., d.h. *ohne Beschränkung der Allgemeinheit*, für $n_1 < n_2$) unmittelbar aus der Dreiecksungleichung:

$$|s_{n_2} - s_{n_1}| = \left| \sum_{k=n_1+1}^{n_2} a_k \right| \leq \sum_{k=n_1+1}^{n_2} |a_k| \leq \sum_{k=n_1+1}^{n_2} b_k = |S_{n_2} - S_{n_1}| < \varepsilon$$

□

⁴Die Bezeichnung *Kriterium* ist insofern irreführend, als nur hinreichende bzw. notwendige Bedingungen für Konvergenz gegeben werden aber keine äquivalenten. Dennoch hat sie sich eingebürgert.

In jedem Fall bilden die Glieder einer konvergenten Reihe eine Nullfolge:

Satz 2.2.1.3. *Konvergiert die Reihe $\sum_{n=0}^{\infty} a_n$, so folgt $\lim_{n \rightarrow \infty} a_n = 0$.*

Beweis. Mit den bisherigen Bezeichnungen gilt $a_n = s_n - s_{n-1}$ für $n \geq 1$, folglich

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} (s_n - s_{n-1}) = \lim_{n \rightarrow \infty} s_n - \lim_{n \rightarrow \infty} s_{n-1} = s - s = 0.$$

□

Die Umkehrung gilt nicht, wie wir bald am Beispiel der harmonischen Reihe sehen werden.

Es sei bemerkt, dass alles Bisherige, sofern die Ordnungsrelation \leq nicht im Spiel war, nicht nur für reelle, sondern genauso für komplexe Folgen und somit Reihen sinnvoll war und gilt. Sehr nützlich ist außerdem die Beobachtung, dass Konvergenz bzw. Divergenz einer Reihe stets erhalten bleibt, wenn man nur endlich viele Glieder verändert, wegnimmt oder hinzufügt. Der Wert der Reihe kann sich dabei aber sehr wohl ändern – im Gegensatz zum Grenzwert von Folgen.

Übungsaufgabe 145. (E) *Begründen Sie die Aussagen des letzten Absatzes sorgfältig.*

Übungsaufgabe 146. (T) *Wir betrachten die Reihe mit den Gliedern $a_n = \frac{1}{n^2 + n}$.*

1. *Überprüfen Sie, dass für $n \geq 1$ gilt: $a_n = \frac{1}{n^2 + n} = \frac{1}{n} - \frac{1}{n+1}$.*
2. *Berechnen Sie $s_n = \sum_{k=1}^n a_k$.*
3. *Berechnen Sie $a_n := s_n - s_{n-1}$ mit s_n aus Teil 2.*
4. *Berechnen Sie für a_n und s_n aus Teil 2 bzw. 3 die Grenzwerte $\lim_{n \rightarrow \infty} s_n$ und $\lim_{n \rightarrow \infty} a_n$.*

Übungsaufgabe 147. (T) *Untersuchen Sie die Reihen $\sum_{n=1}^{\infty} a_n, \sum_{n=1}^{\infty} b_n, \dots$ aus Aufgabe 143 auf Konvergenz. Bestimmen Sie gegebenenfalls den Grenzwert.*

2.2.2 Einige wichtige Beispiele konvergenter und divergenter Reihen

Inhalt in Kurzfassung: Wir untersuchen geometrische, harmonische und hyperharmonische Reihen auf Konvergenz .

Zu Beginn behandeln wir die **geometrische Reihe** mit den Gliedern $a_n := q^n$, $n \in \mathbb{N}$, mit einer festen Basis $q \neq 1$. In 1.3.8, nämlich im Zusammenhang mit Induktion und Rekursion, haben wir für ihre Partialsummen die Formel

$$s_n = \sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q}$$

bewiesen. Für $|q| > 1$ wird der Zähler mit wachsendem n betragsmäßig beliebig groß, also auch s_n . Somit herrscht Divergenz. Ist aber $|q| < 1$, so gilt $q^n \rightarrow 0$ (siehe 2.1.1). Folglich gilt für diesen Fall die äußerst wichtige **Formel für die unendliche geometrische Reihe**:

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$$

Als Anwendung zeigen wir, dass $x := 0,999\dots = 1$, wobei die Gleichheit wirklich eine exakte ist. Tatsächlich liefert die Formel für die geometrische Reihe

$$x = 0,999\dots = \sum_{n=1}^{\infty} \frac{9}{10^n} = \frac{9}{10} \cdot \sum_{n=0}^{\infty} \left(\frac{1}{10}\right)^n = \frac{9}{10} \cdot \frac{1}{1 - \frac{1}{10}} = \frac{9}{10} \cdot \frac{10}{9} = 1.$$

Die sogenannte **harmonische Reihe** hat die Glieder $a_n = \frac{1}{n}$, $n \geq 1$, die gegen 0 konvergieren. Dennoch ist die harmonische Reihe divergent. Man sieht dies sehr leicht, indem man die Summanden zu endlichen Teilsommen

$$t_n := \frac{1}{2^n + 1} + \frac{1}{2^n + 2} + \dots + \frac{1}{2^{n+1}}$$

zusammenfasst. Die Summe t_n besteht aus 2^n Summanden, deren letzter, $\frac{1}{2^{n+1}}$, am kleinsten ist, also $t_n \geq 2^n \cdot \frac{1}{2^{n+1}} = \frac{1}{2}$. Somit gilt

$$s_{2^{n+1}} = 1 + t_0 + t_1 + \dots + t_n \geq 1 + \frac{n+1}{2} \rightarrow \infty$$

für $n \rightarrow \infty$. Die Partialsummen der harmonischen Reihe sind also unbeschränkt. Somit divergiert die harmonische Reihe.

Übungsaufgabe 148. (P) Betrachten Sie die drei endlichen Reihen

$$\begin{aligned} & \frac{1}{4} + \frac{1}{5} + \dots + \frac{1}{N_1}, \\ & \frac{1}{16} + \frac{1}{17} + \dots + \frac{1}{N_2}, \\ & \frac{1}{1000} + \frac{1}{1001} + \dots + \frac{1}{N_3}. \end{aligned}$$

Wie weit müssen Sie summieren, d.h. wie groß müssen Sie N_1, N_2 und N_3 wählen, damit obige Summen jeweils sicherlich den Wert $\frac{1}{2}$ übersteigen? Schließen Sie aus dem gerade angewandten Verfahren, dass die harmonische Reihe divergiert und geben Sie ein N_4 an, sodass $\sum_{n=1}^{N_4} \frac{1}{n} > 10^{10}$.

Es ist bemerkenswert, dass eine scheinbar nur geringfügige Modifikation der harmonischen Reihe, nämlich zur **hyperharmonischen Reihe** $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ mit einem Exponenten $\alpha = 1 + \varepsilon > 1$ reicht, um Konvergenz zu erzwingen. (Wir greifen insofern auf 3.3.2 vor, als wir schon hier beliebige reelle Exponenten α zulassen.)

Ein Beweis lässt sich mit Hilfsmitteln aus der Integralrechnung (siehe 5.4.2), die wir an dieser Stelle noch nicht zur Verfügung haben, noch leichter führen. Ein elementarer Zugang folgt gleich als Übungsaufgabe.

Übungsaufgabe 149. (E) Zeigen Sie, dass die hyperharmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ für $\alpha > 1$ konvergiert, für $\alpha < 1$ divergiert. Hinweis für $\alpha > 1$: Fassen Sie zunächst ähnlich wie beim Beweis der Konvergenz der harmonischen Reihe die Glieder für die Indizes $2^n + 1$ bis 2^{n+1} zusammen und schätzen Sie deren Summe durch ein geeignetes q^n ab. Damit ist eine geometrische Reihe als konvergente Majorante gefunden.

Übungsaufgabe 150. (T) Untersuchen Sie die Konvergenz der folgenden Reihen und geben Sie - falls existent - den Grenzwert an.

$$1. \sum_{m=1}^{\infty} 4 \left(-\frac{1}{4}\right)^m \quad 2. \sum_{n=0}^{\infty} \frac{1}{4 \cdot (-1)^n} \quad 3. \sum_{k=2}^{\infty} \frac{1}{4 \cdot 2^k} \quad 4. \sum_{i=0}^{\infty} \frac{1}{4} 2^i.$$

Übungsaufgabe 151. (T) Bestimmen Sie Real- und Imaginärteil von $\sum_{n=1}^{\infty} \frac{e^{in\psi}}{2^n}$.

Übungsaufgabe 152. (T) Untersuchen Sie die folgenden Reihen auf Konvergenz bzw. Divergenz, indem Sie eine geeignete Majorante bzw. Minorante angeben:

$$1. \sum_{n=1}^{\infty} \frac{\sin n}{2^n} \quad 2. \sum_{n=1}^{\infty} \frac{n}{1+n^2} \quad 3. \sum_{n=0}^{\infty} \left(1 + \frac{1}{n}\right)^n \quad 4. \sum_{n=1}^{\infty} \frac{(-1)^n}{1+2^n}.$$

Übungsaufgabe 153. (T) Wie Übungsaufgabe 152.

$$1. \sum_{n=0}^{\infty} \frac{2n-4}{n^4+2} \quad 2. \sum_{n=3}^{\infty} \frac{24n}{n^2-4}$$

$$3. \sum_{n=1}^{\infty} \frac{2n + \cos(4n)}{4n^2}$$

Übungsaufgabe 154. (T) Untersuchen Sie mittels Majoranten- und Minorantenkriterium auf Konvergenz bzw. Divergenz.

$$\sum_{n=1}^{\infty} \frac{\sqrt[3]{n}}{\sqrt{n^3 - \sqrt{n}}} \quad \text{und} \quad \sum_{n=1}^{\infty} \frac{1}{\sum_{k=1}^n \frac{1}{\sqrt{k}}}$$

2.2.3 Wurzel- und Quotientenkriterium

Inhalt in Kurzfassung: Wurzel- und Quotientenkriterium geben einerseits hinreichende, andererseits notwendige Bedingungen für die Konvergenz bzw. Divergenz von Reihen. Sie ergeben sich ziemlich unmittelbar aus einer Anwendung der Vergleichskriterien auf die geometrische Reihe. Eine erste, aber bereits sehr wichtige Anwendung des Quotientenkriteriums ist die uneingeschränkte absolute Konvergenz der Exponentialreihe.

Häufig dienen geometrische Reihen als konvergente Majoranten. Gegeben eine Reihe mit Gliedern a_n , hat man dazu lediglich ein q mit $|q| < 1$ und $|a_n| \leq q^n$ oder, äquivalent, $\sqrt[n]{|a_n|} \leq |q| < 1$ zu finden. Die resultierende Bedingung heißt **Wurzelkriterium** und lautet:

Satz 2.2.3.1. *Angenommen, es gibt ein $q < 1$, so dass für alle $n \in \mathbb{N}$ (auch endlich viele Ausnahmen sind zugelassen) $\sqrt[n]{|a_n|} \leq q$ gilt. Dann konvergiert die Reihe $\sum_{n=0}^{\infty} a_n$ absolut. Insbesondere ist Konvergenz garantiert, sofern der Grenzwert $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}$ existiert und < 1 ist.*

Man beachte, dass die Voraussetzung nicht zu $q \leq 1$ abgeschwächt werden kann. Denn dann würde beispielsweise die harmonische Reihe die Voraussetzungen erfüllen, obwohl sie divergiert.

Weil der Umgang mit Wurzeln hohen Grades unangenehm sein kann, ist in vielen Fällen das **Quotientenkriterium** leichter anwendbar:

Satz 2.2.3.2. *Angenommen, es gibt ein $q < 1$, so dass für alle $n \in \mathbb{N}$ (auch endlich viele Ausnahmen sind zugelassen) $\frac{|a_{n+1}|}{|a_n|} \leq q$ gilt. Dann konvergiert die Reihe $\sum_{n=0}^{\infty} a_n$ absolut. Insbesondere ist Konvergenz garantiert, sofern der Grenzwert $\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|}$ existiert und < 1 ist.*

Beweis. Die Voraussetzung lässt sich zu $|a_{n+1}| \leq q|a_n|$ umschreiben. Mittels Induktion folgt $|a_n| \leq q^n a_0$, also ist die Reihe mit den Gliedern $q^n a_0$ als a_0 -fache der geometrischen Reihe mit Basis $q < 1$ eine konvergente Majorante. \square

Das prominenteste Anwendungsbeispiel des Quotientenkriteriums ist die **Exponentialreihe**

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

mit irgendeiner festen Zahl x . Die Glieder a_n dieser Reihe sind also $a_n = \frac{x^n}{n!}$. Die Quotienten aufeinanderfolgender Glieder konvergieren gegen

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n \rightarrow \infty} \frac{x^{n+1}}{(n+1)!} \frac{n!}{x^n} = \lim_{n \rightarrow \infty} \frac{x}{n+1} = 0.$$

Laut Quotientenkriterium folgt daraus die Konvergenz der Exponentialreihe für beliebiges $x \in \mathbb{R}$ und sogar $x \in \mathbb{C}$. Wir werden später sehen, dass $\exp(x) = e^x$ gilt mit der **Eulerschen Zahl**

$$e = \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} + \frac{1}{720} + \frac{1}{5040} + \dots \approx 2,718281828459\dots$$

Übungsaufgabe 155. (P) Berechnen Sie die Zahl $e = \sum_{n=0}^{\infty} \frac{1}{n!}$ (**Eulersche Zahl**) näherungsweise (ohne Taschenrechner oder Computer), indem Sie eine rationale Zahl $q = \sum_{n=0}^N \frac{1}{n!}$ angeben, für die Sie $|e - q| < 10^{-2}$ beweisen können.

Hinweis: Der Reihenrest $r_N = \sum_{n=N+1}^{\infty} \frac{1}{n!}$ kann durch Vergleich mit einer geometrischen Reihe abgeschätzt werden.

Übungsaufgabe 156. (T) Untersuchen Sie die Konvergenz der folgenden Reihen mit Hilfe eines geeigneten Kriteriums:

$$\begin{array}{lll}
1. \sum_{n=1}^{\infty} \frac{1}{n} \left(2 + \frac{1}{n}\right)^n & 2. \sum_{n=1}^{\infty} \frac{n^4}{e^{n^2}} & 3. \sum_{n=1}^{\infty} \frac{n!}{n^n} \\
4. \sum_{n=1}^{\infty} \left(\frac{n}{n+1}\right)^{n^2} & &
\end{array}$$

Hinweis: Verwenden Sie $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$ bzw. $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$.

2.2.4 Besonderheiten absolut konvergenter Reihen

Inhalt in Kurzfassung: Die wichtigsten Besonderheiten absolut konvergenter Reihen kommen im Riemannsches Umordnungssatz zum Ausdruck (die Glieder dürfen ohne Einfluss auf den Wert der Reihe vertauscht werden, im extremen Gegensatz zu bedingt konvergenten Reihen) und in der Möglichkeit, Produkte zweier Reihen als Cauchyprodukte zu berechnen.

Unter den konvergenten Reihen spielen die absolut konvergenten eine besonders wichtige Rolle. Einen wesentlichen Grund dafür kann man darin sehen, dass ihr Wert nicht von der Reihenfolge der Glieder abhängt. Sehr eindrucksvoll kommt das im **Riemannsches Umordnungssatz** zum Ausdruck.

Satz 2.2.4.1. Sei $\sum_{n=0}^{\infty} a_n$ eine konvergente Reihe mit Wert s . Je nach absoluter oder bedingter Konvergenz liegen zwei konträre Situationen vor:

1. Ist die Reihe absolut konvergent gegen $s \in \mathbb{R}$, so konvergiert auch jede Reihe, die durch Umordnung der Glieder a_n zustande kommt, gegen denselben Wert s . Genauer gesagt: Ist $\pi : \mathbb{N} \rightarrow \mathbb{N}$ bijektiv, also irgendeine Permutation der Indexmenge \mathbb{N} , so gilt

$$\sum_{n=0}^{\infty} a_{\pi(n)} = \sum_{n=0}^{\infty} a_n = s.$$

2. Ist die Reihe nur bedingt konvergent, so kann durch geeignete Umordnung der Glieder jeder beliebige andere Wert, wie auch Divergenz erzielt werden. Genauer gesagt: Ist $t \in \mathbb{R} \cup \{\infty, -\infty\}$ beliebig vorgegeben, so gibt es ein bijektives $\pi : \mathbb{N} \rightarrow \mathbb{N}$, d.h. eine Permutation der Indexmenge \mathbb{N} mit

$$\sum_{n=0}^{\infty} a_{\pi(n)} = t.$$

Beweis. 1. Wir bezeichnen die Partialsummen der umgeordneten Reihe mit $s'_n := \sum_{k=0}^n a_{\pi(k)}$ und führen den Beweis, indem wir ein beliebiges $\varepsilon > 0$ vorgeben und zeigen, dass es ein $n_0 \in \mathbb{N}$ gibt derart, dass alle $n \geq n_0$ die Ungleichung $|s_n - s'_n| < \varepsilon$ erfüllen, woraus

$$s - \varepsilon \leq \liminf_{n \rightarrow \infty} s'_n \leq \limsup_{n \rightarrow \infty} s'_n \leq s + \varepsilon.$$

folgt. Weil darin $\varepsilon > 0$ beliebig ist, beweist das $\lim_{n \rightarrow \infty} s'_n = s$, wie behauptet.

Will man tatsächlich $|s_n - s'_n| < \varepsilon$ garantieren, liegt es nahe, n so groß zu wählen, dass in beiden Partialsummen nur Glieder mit kleinem Beitrag fehlen. Und das ist möglich, weil es wegen der absoluten Konvergenz der gegebenen Reihe ein n_1 gibt mit $\sum_{n=n_1}^{\infty} |a_n| < \frac{\varepsilon}{2}$. Weil π surjektiv ist, gibt es ein hinreichend großes $n_0 \geq n_1$ derart, dass unter den Indizes $\pi(0), \pi(1), \dots, \pi(n_0)$ alle Zahlen $0, 1, \dots, n_1$ vorkommen. Für $n \geq n_0$ unterscheiden sich die Partialsummen s_n und s'_n daher nur mehr um Summanden $a_{n'}$ mit $n' > n_1$, also:

$$|s_n - s'_n| \leq 2 \sum_{n=n_1}^{\infty} |a_n| < \varepsilon.$$

2. Ist $\sum_{n=0}^{\infty} a_n$ hingegen bedingt konvergent, müssen die positiven Summanden für sich eine divergente Reihe bilden, genauso die negativen. Ist $t \in \mathbb{R}$ vorgegeben, kann man die a_n in folgender Weise zu einer Reihe mit Gliedern $a_{\pi(n)}$ und Wert t umordnen: Zunächst nimmt man aus a_0, a_1, \dots der Reihe nach so lange positive Glieder $a_{\pi(0)}, a_{\pi(1)}, \dots, a_{\pi(k_0)} \geq 0$, bis ihre Summe erstmals $s'_{k_0} > t$ (Notation wie im Beweis der ersten Aussage) erfüllt. Sodann wählt man von den verbliebenen, wieder der Reihe nach, negative Glieder $a_{\pi(k_0+1)}, \dots, a_{\pi(k_1)} < 0$ aus, bis erstmals $s'_{k_1} < t$ gilt. Dieses Verfahren setzt man fort: Man wählt wieder positive Glieder bis $s'_{k_2} > t$, dann negative bis $s'_{k_3} < t$ etc. Auf diese Weise werden alle Glieder aufgebraucht, mit anderen Worten: Es entsteht eine Bijektion $\pi: \mathbb{N} \rightarrow \mathbb{N}$. Die Partialsummen der resultierenden Reihe schwanken um t , mit lokalen Maxima $s_{k_{2n}}$ und lokalen Minima $s_{k_{2n+1}}$. Weil wir als k_i immer den ersten möglichen Wert nehmen, ist stets $t - a_{k_{2n+1}} \leq s'_{k_{2n+1}} < t < s'_{k_{2n}} \leq t + a_{k_{2n}}$. Da die Reihe der a_n konvergiert, bilden die a_n eine Nullfolge, also werden die Oszillationen der s'_n um t beliebig klein, und es gilt $\lim_{n \rightarrow \infty} s'_n = t$, wie behauptet.

Für $t = \infty$ geht man ähnlich vor, verlangt aber z.B. $s'_{2k} > k$ und $s'_{2k+1} < k$. Auf diese Weise ist Divergenz $s'_n \rightarrow \infty$ für $n \rightarrow \infty$ garantiert; entsprechend für $t = -\infty$.

□

Diese Unveränderlichkeit des Wertes absolut konvergenter Reihen bei Umordnung der Glieder hat eine technische Konsequenz, deren Bedeutung wir später im Zusammenhang mit Potenzreihen erst richtig ermessen können. Und zwar geht es um das Produkt zweier unendlicher Reihen mit Gliedern a_n bzw. b_n . Das Distributivgesetz ist zunächst für nur zwei Summanden definiert, liefert aber sehr schnell auch die entsprechende Formel für beliebige Anzahlen $n_1, n_2 \in \mathbb{N}$ endlich vieler Summanden:

$$s_{n_1, n_2} := \left(\sum_{k=0}^{n_1} a_k \right) \cdot \left(\sum_{l=0}^{n_2} b_l \right) = (a_0 + a_1 + \dots + a_{n_1}) \cdot (b_0 + b_1 + \dots + b_{n_2}) = \sum_{k=0}^{n_1} \sum_{l=0}^{n_2} a_k b_l,$$

wobei es wegen des Kommutativgesetzes auf die Reihenfolge der endlich vielen Summanden $a_k b_l$ nicht ankommt. Wachsen die oberen Summationsgrenzen n_1 und n_2 , so kommen nach und nach alle Summanden $a_k b_l$ mit $k, l \in \mathbb{N}$ ins Spiel. Um eine natürliche Ordnung der Summanden festzulegen, fasst man zunächst für jedes $n \in \mathbb{N}$ jene mit $k + l = n$ zusammen. Man bildet also die endlichen Summen $c_n := \sum_{k=0}^n a_k b_{n-k}$. Die aus diesen Gliedern gebildete Reihe

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \sum_{k=0}^n a_k b_{n-k}$$

heißt das **Cauchyprodukt** der Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$. Sofern die gegebenen Reihen absolut konvergieren, zeigen ähnliche Überlegungen wie beim Riemannschen Umordnungssatz: Für hinreichend große n_1, n_2 unterscheidet sich die Summe s_{n_1, n_2} einerseits nur mehr beliebig wenig vom Produkt der beiden gegebenen Reihen, andererseits auch nur beliebig wenig von der Partialsumme $s_{n_1+n_2} := \sum_{n=0}^{n_1+n_2} c_n$ des Cauchyproduktes. Folglich konvergiert das Cauchyprodukt gegen das Produkt der gegebenen Reihen. Diese Überlegungen führen zu folgendem Satz:

Satz 2.2.4.2. *Das Cauchyprodukt zweier absolut konvergenter Reihen mit den Gliedern a_n bzw. b_n konvergiert selbst absolut, und zwar gegen das Produkt der Werte der beiden Reihen,⁵ als Formel:*

$$\left(\sum_{n=0}^{\infty} a_n \right) \cdot \left(\sum_{n=0}^{\infty} b_n \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right)$$

Beweis. Aus psychologischen Gründen denken wir zunächst nur an den Fall positiver Glieder $a_n, b_n \geq 0$. Zur Abkürzung schreiben wir $A := \sum_{n=0}^{\infty} a_n$, $B := \sum_{n=0}^{\infty} b_n$ und $C := \sum_{n=0}^{\infty} c_n$ mit $c_n := \sum_{k=0}^n a_k b_{n-k}$. Es ist also $AB = C$ zu beweisen (wobei sich die Konvergenz von C erst dabei ergeben wird). Weiters schreiben wir $A_N := \sum_{n=0}^N a_n$, $B_N := \sum_{n=0}^N b_n$ und $C_N := \sum_{n=0}^N c_n$ für alle $N \in \mathbb{N}$. Nach der Definition unendlicher Reihen gilt $\lim_{N \rightarrow \infty} A_N = A$ und $\lim_{N \rightarrow \infty} B_N = B$ und $\lim_{N \rightarrow \infty} C_N = C$. Wenn es uns gelingt $\lim_{N \rightarrow \infty} D_N = 0$ für die Differenzen $D_N := A_N B_N - C_N = 0$ zu beweisen, folgt daraus die Konvergenz $\lim_{N \rightarrow \infty} C_N = C$ samt Gleichung $AB = C$.

Ausmultiplizieren der beiden Summen A_N und B_N zeigt, dass das Produkt $A_N B_N$ die Summe all jener Produkte $a_k b_l$ von Summanden der beiden gegebenen Reihen ist, für die $0 \leq k, l \leq N$ gilt, während bei C_N die Bedingung $k + l \leq N$ gilt. Somit ist die Differenz d_N gerade die Summe jener $a_k b_l$ mit $0 \leq k, l \leq N$ und $k + l > N$. Letzteres bedeutet, dass in jedem Summanden $a_k b_l$ von d_N eine der beiden Möglichkeiten $k \geq M$ oder $l \geq M$ zutreffen muss, wobei mit M die Zahl $\frac{N}{2}$ (falls N gerade ist) oder $\frac{N-1}{2}$ (falls N ungerade ist) bezeichnet sei. Damit folgt

$$d_N \leq (A_N - A_M) B_N + A_N (B_N - B_M).$$

⁵Der Satz gilt mit etwas feinsinnigerem Beweis übrigens auch dann, wenn von den beiden gegebenen Reihen nur eine absolut konvergiert.

Wegen der Positivität der Glieder a_n und b_n gilt sicher $A_N \leq A$ und $B_N \leq B$ (Beschränktheit der A_N und B_N), außerdem wegen der Konvergenz beider Reihen und $M \rightarrow \infty$ für $N \rightarrow \infty$ auch

$$\lim_{N \rightarrow \infty} (A_N - A_M)B_N \leq \lim_{N \rightarrow \infty} B(A - A_M) = 0$$

und analog $\lim_{N \rightarrow \infty} A_N(B_N - B_M) = 0$. Daraus folgt insgesamt tatsächlich

$$\lim_{n \rightarrow \infty} d_N = \lim_{N \rightarrow \infty} d_N \leq (A_N - A_M)B_N + A_N(B_N - B_M) = 0.$$

Für den allgemeinen Fall, dass nämlich die Summanden a_n und b_n irgendwelche, möglicherweise negative oder auch komplexe Zahlen sind, die jedoch absolut konvergente Reihen bilden, überzeuge man sich, dass im behandelten Fall schlicht an geeigneten Stellen Absolutbeträge einzufügen und sind. Wegen der vorausgesetzten absoluten Konvergenz von A und B ändert das jedoch nichts an der Gültigkeit der Argumentation insgesamt. \square

Übungsaufgabe 157. (P) Überprüfen Sie Aussage von Satz 2.2.4.2 anhand des Beispiels $a_n = \frac{1}{2^n}$ und $b_n = \frac{1}{3^n}$.

Übungsaufgabe 158. (P) Wie Übungsaufgabe 157 mit $a_n = \frac{2^n}{n!}$ und $b_n = \frac{(-3)^n}{n!}$.

Übungsaufgabe 159. (E) Finden Sie ein Beispiel, das zeigt, dass die Voraussetzung der absoluten Konvergenz in Satz 2.2.4.2 nicht einfach weggelassen werden kann.

Hinweis: Da die Reihen der a_n und b_n bedingt konvergieren sollen, müssen Sie vor allem die Vorzeichen geschickt wählen. Vielleicht gelingt es, dies so zu arrangieren, dass die Summanden in $c_n = \sum_{k=0}^n a_k b_{n-k}$ für festes n gleiches Vorzeichen haben und die c_n mit wachsendem n nicht gegen 0 konvergieren. Dann konvergiert das Cauchyprodukt überhaupt nicht.

2.2.5 Alternierende Reihen

Inhalt in Kurzfassung: Jede alternierende Reihe (d.h. mit Gliedern abwechselnden Vorzeichens), deren Glieder betragsmäßig eine monotone Nullfolge bilden, konvergiert (nach Leibniz).

Bei sogenannten **Leibnizreihen** spielt absolute Konvergenz keine Rolle. Trotzdem konvergieren sie. Das einfachste Beispiel dieser Art ist die bedingt konvergente Reihe $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$. (An späterer Stelle wird sich mit etwas Theorie ergeben, dass diese Reihe den Wert $-\ln 2$ hat.)

Satz 2.2.5.1. Angenommen, die positiven Glieder $a_0 \geq a_1 \geq \dots \geq 0$ bilden eine monoton fallende Nullfolge. Dann konvergiert die aus ihnen gebildete alternierende Reihe

$$\sum_{n=0}^{\infty} (-1)^n a_n.$$

Beweis. Die Partialsummen $s_n := \sum_{k=0}^n (-1)^k a_k$ erfüllen offenbar

$$s_1 \leq s_3 \leq s_5 \leq \dots \leq s_4 \leq s_2 \leq s_0.$$

Aufgrund von Monotonie und Beschränktheit existieren die Grenzwerte

$$s^- := \lim_{n \rightarrow \infty} s_{2n+1} \leq s^+ := \lim_{n \rightarrow \infty} s_{2n}.$$

Wegen

$$s^+ - s^- = \lim_{n \rightarrow \infty} s_{2n} - s_{2n+1} = \lim_{n \rightarrow \infty} a_{2n+1} = 0$$

müssen diese sogar denselben Wert

$$s := s^+ = s^- = \lim_{n \rightarrow \infty} s_n = \sum_{n=0}^{\infty} (-1)^n a_n$$

haben. □

Übungsaufgabe 160. (E) Betrachten Sie die Partialsummen einer konvergenten alternierenden Reihe $\sum a_k$. Gehen Sie nochmals die Argumentation im Beweis von Satz 2.2.5.1 durch und begründen Sie damit die folgende Abschätzung

$$\left| \sum_{k=1}^{n-1} a_k - \sum_{k=1}^{\infty} a_k \right| \leq |a_n|.$$

Übungsaufgabe 161. (T) Berechnen Sie die Partialsummen $s_n := \sum_{k=1}^n (-1)^k \frac{1}{k}$ für $n = 1, 2, \dots, 10$ exakt als Bruch und in Dezimaldarstellung auf 2 Stellen gerundet.

Übungsaufgabe 162. (T) Bei welchen der folgenden Reihen darf das Leibniz'sche Kriterium 2.2.5.1 verwendet werden? Welche Genauigkeit ist mit 10 Summanden möglich?

1. $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}, \quad x \in \mathbb{R}$
2. $\sum_{n=1}^{\infty} (-1)^n \frac{\sin(n)}{n}$
3. $\sum_{n=1}^{\infty} \frac{\sin(n)}{n}$
4. $\sum_{n=1}^{\infty} (-1)^n \left(e - \left(1 + \frac{1}{n}\right)^n \right)$
5. $\sum_{n=1}^{\infty} \left(\frac{xn}{n+1} \right)^n, \quad x \in \mathbb{R}.$

Hinweis: Sie dürfen verwenden, dass die Folge der $a_n := \left(1 + \frac{1}{n}\right)^n$ monoton gegen $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$ konvergiert.

Übungsaufgabe 163. (T) Untersuchen Sie die folgenden Reihen auf Konvergenz bzw. Divergenz mit einem Kriterium Ihrer Wahl.

1.
$$\sum_{n=1}^{\infty} (-1)^n \sin\left(\frac{1}{n}\right)$$

2.
$$\sum_{n=1}^{\infty} (-1)^n \sin\left(-\frac{1}{n}\right)$$

3.
$$\sum_{n=1}^{\infty} \sin\left(\frac{(-1)^n}{n}\right)$$

Übungsaufgabe 164. (T) Untersuchen Sie die folgenden Reihen auf Konvergenz bzw. Divergenz mit einem Kriterium Ihrer Wahl.

1.
$$\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}$$

2.
$$\sum_{n=1}^{\infty} \frac{n!}{(2n)!}$$

3.
$$\sum_{n=1}^{\infty} \frac{n^n}{3^{(n^2)}}$$

4.
$$\sum_{n=1}^{\infty} a_n, \quad \text{mit } a_n = \begin{cases} \frac{2}{n}, & n \text{ ungerade} \\ -\frac{1}{n-1}, & n \text{ gerade} \end{cases}$$

Kapitel 3

Funktionen und Stetigkeit

In diesem Kapitel geht es um reelle Funktionen, vor allem unter dem Gesichtspunkt von Stetigkeit und Unstetigkeit. Viele Vertiefungen werden erst im nachfolgenden Kapitel 4 über Differentialrechnung folgen. Hier behandeln wir in 3.1 allgemeine Eigenschaften reeller Funktionen. In 3.2 geht es um Polynome und rationale Funktionen. Dabei spielen auch algebraische Aspekte eine wichtige Rolle. Weitere wichtige Beispiele reeller Funktionen sind schließlich Gegenstand von 3.3.

3.1 Reelle Funktionen

Unter einer reellen Funktion f verstehen wir eine solche, wo Definitionsbereich und Wertebereich in \mathbb{R} enthalten sind, also $f : \mathbb{R} \supseteq D \rightarrow \mathbb{R}$. Wir beginnen mit der graphischen Darstellung und einfachen Eigenschaften (3.1.1). Der Grenzwertbegriff im Kontext von Funktionen führt uns in 3.1.2 zum Begriff der Stetigkeit. Vererbungseigenschaften stetiger Funktionen behandeln wir in 3.1.3, illustrative Beispiele folgen in 3.1.4. Wie schon bei Folgenkonvergenz beruhen viele wichtige theoretische Ergebnisse auf der Vollständigkeit von \mathbb{R} (siehe 3.1.5). In 3.1.6 geht es um eine interessante Verbindung von Stetigkeit mit rekursiven Folgen, die durch Iteration zustande kommen. Den Abschluss bildet das Konzept der gleichmäßigen Konvergenz (siehe 3.1.7), das für das Verständnis gewisser auf den ersten Blick überraschender Phänomene hilfreich ist.

3.1.1 Graphische Darstellung und einfache Eigenschaften

Inhalt in Kurzfassung: Formal sind reelle Funktionen gewisse Teilmengen der Ebene \mathbb{R}^2 . Das ermöglicht Veranschaulichungen der Funktionen selbst wie auch mit ihnen zusammenhängender Begriffe wie Nullstellen, Polstellen, Beschränktheit, Monotonie, Parität (gerade, ungerade), Periodizität, Asymptoten, Extremwerte und -stellen.

Nach Definition sind reelle Funktionen $f : \mathbb{R} \supseteq D \rightarrow \mathbb{R}$ gewisse Teilmengen der Ebene \mathbb{R}^2 , nämlich $f = \{(x, f(x)) : x \in D\} \subseteq \mathbb{R}^2$. Entsprechend lassen sie sich auf offensichtliche Weise graphisch darstellen. Oft wird die Menge $f \subseteq \mathbb{R}^2$ deshalb auch als **Graph von f** bezeichnet. Formal ist die Unterscheidung zwischen f und dem Graph von f hinfällig. Wir werden deshalb nur dann vom Graph von f sprechen, wenn wir wirklich an die graphische Darstellung denken. Ähnlich wie bei Folgen vereinbaren wir: Ist f ohne Angabe von D nur durch einen Term in einer Variablen (meist x) gegeben, so wählen wir als Definitionsbereich D die Menge aller Werte der Variablen, für die dieser Term sinnvoll ist.

Wir behandeln nun einige wichtige Begriffe samt ihrer anschaulichen Bedeutung.

Eine Zahl $x_0 \in D$ heißt **Nullstelle** von f , wenn $f(x_0) = 0$. Allgemeiner sprechen wir für $f(x_0) = y_0$ von einer y_0 -Stelle x_0 von f . Nullstellen zeigen sich als Schnittpunkte des Graphen von f mit der x -Achse. Beispiel: $x_0 = \pi$ ist eine Nullstelle der Funktion $f = \sin$.

Analog zu Mengen und Folgen zeichnet man auch unter reellen Funktionen gewisse als (nach oben/unten) **beschränkt** aus; und zwar jene, für die die Bildmenge $f(D) = \{f(x) : x \in D\}$ als Teilmenge von \mathbb{R} die entsprechende Eigenschaft hat. Beispiele: \sin, \cos sind auf ganz \mathbb{R} beschränkt (nach unten durch -1 , nach oben durch 1), $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$, ist nur nach unten beschränkt (im Beispiel mit Infimum 0), nicht aber nach oben.

Eine **Polstelle** $x_0 \in \mathbb{R}$ von f zeigt sich anschaulich daran, dass f in jeder Umgebung von x_0 unbeschränkt ist. Eine genaue Definition erfordert Begriffe, die hier nicht zur Verfügung stehen. Für uns genügt: Eine Polstelle liegt nicht im Definitionsbereich D von f und entsteht z.B. dann, wenn durch einen Nenner mit Nullstellen dividiert wird. Typisches Beispiel einer Polstelle ist $x_0 = 0$ für $f(x) := \frac{1}{x}$ ($x \neq 0$).

Analog zu Folgen nennt man auch reelle Funktionen (**streng**) **monoton wachsend** bzw. **fallend**, wenn für alle $x_1 < x_2 \in D$ gilt: $f(x_1) \leq f(x_2)$ (monoton wachsend), $f(x_1) < f(x_2)$ (streng monoton wachsend), $f(x_1) \geq f(x_2)$ (monoton fallend) bzw. $f(x_1) > f(x_2)$ (streng monoton fallend). Eine Funktion der Gestalt $x \mapsto kx + d$ mit $k, d \in \mathbb{R}$ ist für $k > 0$ streng monoton wachsend, für $k < 0$ streng monoton fallend und für $k = 0$ konstant und somit sowohl monoton wachsend als auch monoton fallend, nicht jedoch im strengen Sinn.

Eine Funktion f heißt **gerade** bzw. **ungerade**, wenn der Definitionsbereich D symmetrisch um 0 ist (d.h. $x \in D$ genau dann, wenn $-x \in D$) und wenn $f(-x) = f(x)$ bzw. $f(-x) = -f(x)$ für alle $x \in D$ gilt. Der Graph gerader Funktionen ist symmetrisch bezüglich der y -Achse, der ungerader Funktionen ist symmetrisch bezüglich des Koordinatenursprungs. Typische Beispiele gerader bzw. ungerader Funktionen sind die Potenzfunktionen $x \mapsto x^k$ mit geradem bzw. ungeradem k .

Eine Funktion f heißt **periodisch** mit **Periode** p , wenn $f(x + p) = f(x)$ für alle $x \in D$ gilt. Wie die Bezeichnung suggeriert, erkennt man Periodizität daran, dass sich bei Fortschreiten entlang der x -Achse im Abstand von p alles wiederholt. Die klassischen Beispiele periodischer Funktionen sind die trigono-

metrischen \cos , \sin (mit kleinster Periode 2π) und \tan , \cot (mit kleinster Periode π). Wir werden sie in 3.3.4 und 4.3.3 noch genauer behandeln.

Eine (eventuell inhomogene) lineare Funktion l , d.h. $l : x \mapsto kx + d$ mit festen Zahlen $k, d \in \mathbb{R}$ heißt **Asymptote** von f , wenn x sich für $x \rightarrow \infty$ beliebig an l anschmiegt, d.h. wenn $\lim_{x \rightarrow \infty} (f(x) - l(x)) = 0$ und/oder $\lim_{x \rightarrow -\infty} (f(x) - l(x)) = 0$. Das Symbol $\lim_{x \rightarrow \infty}$ ist analog zu verstehen wie $\lim_{n \rightarrow \infty}$ bei Folgen, mit dem einzigen Unterschied, dass $x \in \mathbb{R}$, während $n \in \mathbb{N}$. Beispiel: Die Funktion $f : x \mapsto x + \frac{1}{x}$ hat die Asymptote $l : x \mapsto x$. (Genaue Definition des Grenzwerts einer Funktion folgt in 3.1.2.1.)

Gibt es ein $a \in D$ mit $f(a) \geq f(x)$ für alle $x \in D$, so heißt $f(a)$ (**globales Maximum**) von f und a **Stelle eines Maximums**. Vom **lokalen Maximum** $f(a)$ von f bzw. von der **Stelle a eines lokalen Maximums** spricht man dann, wenn es eine Umgebung $U = (a - \varepsilon, a + \varepsilon)$ von a mit $\varepsilon > 0$ gibt, so dass die Einschränkung von f auf $D \cap U$, d.h. die reelle Funktion $\{(x, f(x)) : x \in D \cap U\}$ ein (globales) Maximum in a hat. Entsprechend (nämlich über $f(a) \leq f(x)$ statt $f(a) \geq f(x)$) definiert man **globales/lokales Minimum** und **Stelle eines globalen/lokalen Minimums**. Maxima und Minima bzw. Stellen von Maxima und Minima fasst man in den Begriffen der **Extrema** bzw. **Extremstellen** zusammen. Graphisch zeigen sich Maxima und Minima als höchst bzw. tiefst gelegene Punkte des Funktionsgraphen.

Typische Beispiele: Die Funktion \cos nimmt an allen Stellen a der Gestalt $a = 2k\pi$ mit $k \in \mathbb{Z}$ das (globale) Maximum 1 an, an allen Stellen der Gestalt $a = (2k+1)\pi$ mit $k \in \mathbb{Z}$ das (globale) Minimum -1 . Im Gegensatz dazu hat die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^3 - 12x$, an der Stelle $a = -2$ ein lokales Maximum $f(-2) = 16$, das nicht global ist; an der Stelle $a = 2$ ein (ebenfalls nicht globales) lokales Minimum $f(2) = -16$. Die Differentialrechnung wird uns Methoden in die Hand geben, sehr systematisch nach Extrema zu suchen.

Übungsaufgabe 165. (P) Wählen Sie eine konkrete reelle Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$. Skizzieren Sie sodann die beiden Mengen

$$A := \{(x, f(x)) : x \in \mathbb{R}\} \subseteq \mathbb{R}^2 \quad \text{und} \quad B := \{(f(x), x) : x \in \mathbb{R}\} \subseteq \mathbb{R}^2.$$

Erklären Sie, wie A und B geometrisch zusammenhängen und finden Sie heraus, welche Eigenschaft von f dafür verantwortlich ist, ob B ebenfalls der Graph einer Funktion ist.

Wenn in Ihrem konkreten Beispiel B der Graph einer Funktion ist, finden Sie abschließend eine andere konkrete Funktion f , sodass B nun nicht mehr der Graph einer Funktion ist. Wenn in Ihrem konkreten Beispiel B nicht der Graph einer Funktion war, finden Sie eine andere konkrete Funktion f , sodass B nun doch Graph einer Funktion ist.

Übungsaufgabe 166. (T) Zeichnen Sie die folgenden Mengen in der x - y -Ebene ein und entscheiden Sie, ob es sich dabei um den Graphen einer Funktion $y = f(x)$ handelt. Falls ja, geben Sie die Abbildungsvorschrift, sowie Definitionsbereich und Wertebereich an.

1. $\{(x, y) : x^2 + y^2 = 1\}$,
 $\{(x, y) : x^2 + y^2 = 1 \text{ und } x \geq 0\}$,
 $\{(x, y) : x^2 + y^2 = 1 \text{ und } y \geq 0\}$.
2. $\{(x, 0) : x \in \mathbb{R}\}$,
 $\{(0, y) : y \in \mathbb{R}\}$,
 $\{(x, x) : x \in \mathbb{R}\}$.
3. $\{(x^2, x) : x \in \mathbb{R}\}$,
 $\{(x, x^2) : x \in \mathbb{R}\}$.

Übungsaufgabe 167. (T) Gegeben ist die reelle Funktion

$$f(x) = (x + 1)^2 - 4.$$

Geben Sie einen möglichst großen Definitionsbereich $D \subseteq \mathbb{R}$ an, so dass f auf D injektiv ist. Berechnen Sie die Umkehrfunktion $f^{(-1)}$ und skizzieren Sie sowohl den Graphen von f als auch den Graphen von $f^{(-1)}$.

Übungsaufgabe 168. (T) Gegeben ist die reelle Funktion $f(x) := \frac{x-1}{x+1}$ für $x \neq -1$.

1. Geben Sie einen möglichst großen Definitionsbereich $D \subseteq \mathbb{R}$, sowie einen Wertebereich $W \subseteq \mathbb{R}$ an, sodass $f : D \rightarrow W$ bijektiv ist.
2. Berechnen Sie die Umkehrfunktion und geben Sie diese in der Form $f^{(-1)}(x) = \dots$ an.
3. Geben Sie $f^{(44)}(x) = \underbrace{f \circ f \circ \dots \circ f(x)}_{44 \text{ mal}}$ und $f^{(45)}(x) = \underbrace{f \circ f \circ \dots \circ f(x)}_{45 \text{ mal}}$ an.

Hinweis: In Teilaufgabe 3 soll natürlich nicht 44 oder 45 mal die Funktionsvorschrift eingesetzt werden. Berechnen Sie statt dessen zuerst $f^{(-2)}$ und verwenden Sie dann $f^{(44)} = f^{(42)} \circ f^{(2)} = f^{(40)} \circ f^{(2)} \circ f^{(2)} = \dots$

Übungsaufgabe 169. (T) Gegeben seien die beiden reellen Funktionen f und g durch $f(x) := x^2 - 1$ und $g(x) := \sin x$.

1. Berechnen Sie $f(g(x))$ und $g(f(x))$ und skizzieren Sie die Funktionsgraphen dieser Funktionen.
2. Geben Sie möglichst große Definitionsbereiche $D_1, D_2 \subset \mathbb{R}$ an, sodass f auf D_1 und g auf D_2 injektiv ist. Bestimmen Sie dort die Umkehrfunktionen $f^{(-1)}$ und $g^{(-1)}$.

Übungsaufgabe 170. (P) Es sei f eine gerade und g eine ungerade Funktion. Überprüfen Sie, ob die Funktionen 1. $t(x) := f(g(x))$, 2. $u(x) := g(f(x))$, 3. $v(x) := f(x) \cdot g(x)$ und 4. $w(x) := f(x) + g(x)$ stets gerade, stets ungerade oder im allgemeinen keines von beidem sind.

Übungsaufgabe 171. (P) Es sei $f: \mathbb{R} \rightarrow \mathbb{R}^+$ eine streng monoton wachsende Funktion. Bestimmen Sie das Monotonieverhalten von

1. $f_1(x) = -f(x)$, 2. $f_2(x) = \frac{1}{f(x)}$, 3. $f_3(x) = f^{(-1)}(x)$, 4. $f_4(x) = \sqrt{f(x)}$
anhand der Definition von Monotonie. Überlegen Sie sich abschließend, was passiert, wenn Sie „streng monoton“ durch „monoton“ bzw. „wachsend“ durch „fallend“ ersetzen.

Übungsaufgabe 172. (T) Untersuchen Sie folgende reelle Funktionen jeweils auf Injektivität, Surjektivität, Umkehrbarkeit, Monotonie, Periodizität und Parität (gerade oder ungerade) und skizzieren Sie die Funktionsgraphen. Geben Sie jeweils auch Definitions- und Wertebereich an.

1. $f_1(x) = x^2$, 2. $f_2(x) = \sqrt{x}$, 3. $f_3(x) = |x|$, 4. $f_4(x) = \frac{1}{x}$.

Übungsaufgabe 173. (T) Untersuchen Sie folgende reelle Funktionen auf Monotonie und Umkehrbarkeit. Geben Sie den maximalen Definitionsbereich an und argumentieren Sie für die Monotonieeigenschaften nicht nur anschaulich, sondern rechnen Sie die jeweilige Definition nach. Skizzieren Sie anschließend die Funktionsgraphen. Falls möglich, geben Sie auch die Umkehrfunktion (samt ihrem Definitions- und Wertebereich) an.

1. $f_1(x) = \frac{1}{\sqrt{x}}, x \neq 0$ 2. $f_2(x) = \frac{1}{\sqrt[3]{x}}, x \neq 0$
3. $f_3(x) = x|x|$ 4. $f_4(x) = 1$ 5. $f_5(x) = x$
6. $f_6(x) = \operatorname{sgn}(x) = \begin{cases} 1 & \text{für } x > 0 \\ 0 & \text{für } x = 0 \\ -1 & \text{für } x < 0 \end{cases}$

Übungsaufgabe 174. (T) Analog zu Übungsaufgabe 173 (In dieser Aufgabe dürfen Sie Ihr Schulwissen über die Exponentialfunktion \exp und den Logarithmus \ln zur Basis e , siehe auch 3.3.2 und 3.3.3, einsetzen.):

1. $f_1(x) = \cos x$ 2. $f_2(x) = \exp x$ 3. $f_3(x) = \exp(-|x|)$
4. $f_4(x) = \ln x$ 5. $f_5(x) = |\ln x|$ 6. $f_6(x) = \ln |x|$

3.1.2 Grenzwert von Funktionen und Stetigkeit

Inhalt in Kurzfassung: Der Begriff des Grenzwertes einer reellen Funktion $f(x)$ für $x \rightarrow x_0$ kann in weitgehender Analogie zum Folgentrenzwert definiert werden (ε - δ -Definition). Funktionen, für die dieser Grenzwert nicht nur existiert, sondern auch mit $f(x_0)$ übereinstimmt, heißen stetig in x_0 . Die Analogie zum Folgentrenzwert äußert sich im Kriterium der Folgenstetigkeit. Häufig ist es von Interesse, auch rechts- und/oder linksseitige Grenzwerte von Funktionen zu betrachten.

Wir haben bereits die Symbolik $\lim_{x \rightarrow \infty} f(x)$ verwendet. Ihre Bedeutung liegt auf der Hand, weil sich der Folgentrenzwert direkt übertragen lässt. Man hat lediglich $n \in \mathbb{N}$ durch $x \in \mathbb{R}$ zu ersetzen. Wir wollen auch Grenzwerte der Form $\lim_{x \rightarrow x_0} f(x)$ mit $x_0 \in \mathbb{R}$ betrachten.

Definition 3.1.2.1. Sei $f : D \rightarrow \mathbb{R}$ eine reelle Funktion und x_0 Häufungspunkt von D , d.h. in jeder Umgebung von x_0 liegen Punkte $x \in D$ mit $x \neq x_0$. Wir nennen $y_0 \in \mathbb{R}$ den **Grenzwert von f für $x \rightarrow x_0$** und schreiben

$$\lim_{x \rightarrow x_0} f(x) = y_0,$$

wenn folgendes gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D \cap (x_0 - \delta, x_0 + \delta) \setminus \{x_0\} : |f(x) - y_0| < \varepsilon.$$

Zieht man statt der Werte $x \in (x_0 - \delta, x_0 + \delta)$ nur x -Werte mit $x < x_0$ heran, also $x \in (x_0 - \delta, x_0)$, so spricht man von **linksseitigem Grenzwert** $\lim_{x \rightarrow x_0^-} f(x)$, analog für $x > x_0$ von **rechtsseitigem Grenzwert** $\lim_{x \rightarrow x_0^+} f(x)$.

Wir schreiben $\lim_{x \rightarrow x_0} f(x) = \infty$, wenn es für alle $r \in \mathbb{R}$ ein $\delta > 0$ gibt, mit $f(x) > r$ für alle $x \in D \cap (x_0 - \delta, x_0 + \delta)$ gilt, analog $\lim_{x \rightarrow x_0} f(x) = -\infty$ mit $f(x) < r$ (statt $f(x) > r$). Ersetzt man in der definierenden Formel für $\lim_{x \rightarrow x_0} f(x) = y_0$ den Bestandteil $\exists \delta > 0 \forall x \in D \cap (x_0 - \delta, x_0 + \delta)$ durch $\exists x_1 \in \mathbb{R} \forall x \in D \cap (x_1, \infty)$, so erhält man die Definition für $\lim_{x \rightarrow \infty} f(x) = y_0$; ersetzt man darin wiederum (x_1, ∞) durch $(-\infty, x_1)$, so erhält man die Definition für $\lim_{x \rightarrow -\infty} f(x) = y_0$. Sinngemäß sind auch die Beziehungen $\lim_{x \rightarrow \pm\infty} f(x) = \pm\infty$ (4 Vorzeichenvarianten) definiert.

Die exakte Definition der letztgenannten Varianten des Grenzwertes erfolgt im gleichen Geist und soll als Übung behandelt werden:

Übungsaufgabe 175. (E) Wie würden Sie $\lim_{x \rightarrow \infty} f(x) = \infty$, $\lim_{x \rightarrow \infty} f(x) = -\infty$, $\lim_{x \rightarrow -\infty} f(x) = \infty$ und $\lim_{x \rightarrow -\infty} f(x) = -\infty$ präzise definieren?

Man beachte, dass die Formel für den Grenzwert einer Funktion dieselbe logische Struktur hat wie beim Grenzwert für Folgen. Das kann schon allein deshalb nicht verwundern, weil ja etwas sehr Ähnliches zum Ausdruck kommt, nämlich beliebige Annäherung an den Grenzwert. Der Unterschied besteht darin, dass $n \rightarrow \infty$ durch $x \rightarrow x_0$ zu ersetzen ist. Von besonderem Interesse ist der stetige Fall, wo $y_0 = f(x_0)$. Weil die folgende Stetigkeitsdefinition auf der vorangegangenen Definition mit ε und δ fußt, spricht man oft auch vom **ε - δ -Kriterium**:

Definition 3.1.2.2. Die reelle Funktion $f : D \rightarrow \mathbb{R}$ heißt **(lokal) stetig in $x_0 \in D$** , falls $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ gilt; f heißt schlechthin oder **(global) stetig**, wenn f lokal stetig in x_0 für alle $x_0 \in D$ ist. Explizit bedeutet Stetigkeit von f in x_0 :

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in (x_0 - \delta, x_0 + \delta) \cap D : |f(x) - f(x_0)| < \varepsilon.$$

Gilt $\lim_{x \rightarrow x_0^-} f(x) = f(x_0)$, also $\lim_{x \rightarrow x_0^-} f(x) = f(\lim_{x \rightarrow x_0} x)$, so spricht man von **linksseitiger Stetigkeit**, bei $\lim_{x \rightarrow x_0^+} f(x) = f(x_0)$ von **rechtsseitiger Stetigkeit** der Funktion f in x_0 .

Einfachste Beispiele stetiger Funktionen sind die konstanten Funktionen f , wo es also ein $c \in \mathbb{R}$ gibt mit $f(x) = c$ für alle $x \in D$. In der Stetigkeitsdefinition kann man zu jedem $\varepsilon > 0$ sogar ein beliebiges $\delta > 0$ nehmen. Nur geringfügig schwieriger ist das im Fall der Identität $f = \text{Id}_D$ mit $f(x) = x$ für alle $x \in D$. Hier setzt man zu vorgegebenem $\varepsilon > 0$ am einfachsten $\delta := \varepsilon$.

Übungsaufgabe 176. (T) Zeigen Sie anhand der Definition der Stetigkeit, dass $f(x) = (x - 1)^2$ auf ganz \mathbb{R} stetig ist. Geben Sie sich ein konkretes $x_0 \in \mathbb{R}$ und $\varepsilon > 0$ vor und berechnen Sie ein passendes $\delta > 0$, sodass die ε - δ Charakterisierung der Stetigkeit an x_0 erfüllt ist.

Eine verbale Beschreibung der Stetigkeit von f an der Stelle x_0 könnte lauten: Nähert sich x beliebig an x_0 an, so nähert sich $f(x)$ beliebig an $f(x_0)$ an. Durchläuft x eine Folge mit Gliedern a_n , die gegen x_0 konvergiert, so konvergiert insbesondere die Folge der Funktionswerte $f(a_n)$ gegen $f(x_0)$. Tatsächlich lässt sich diese Beobachtung zu einem Kriterium verschärfen, das mit dem Schlagwort **Folgenstetigkeit** verbunden ist.

Satz 3.1.2.3. Eine reelle Funktion $f : D \rightarrow \mathbb{R}$ ist genau dann stetig in $x_0 \in D$, wenn für alle Folgen $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ ($a_n \in D$ für alle $n \in \mathbb{N}$) mit $\lim_{n \rightarrow \infty} a_n = x_0$ auch $\lim_{n \rightarrow \infty} f(a_n) = f(x_0)$ gilt.

Beweis. Recht einfach ist der Beweis, dass aus dem ε - δ -Kriterium die Folgenstetigkeit folgt: Ist $\varepsilon > 0$, so gibt es ein $\delta > 0$ wie im ε - δ -Kriterium, also: $|f(x) - f(x_0)| < \varepsilon$ für alle $x \in D$ mit $|x - x_0| < \delta$. Weil die a_n gegen x_0 konvergieren, gibt es ein $n_0 \in \mathbb{N}$ mit $|a_n - x_0| < \delta$ für alle $n \geq n_0$, für welche dann auch $|f(a_n) - f(x_0)| < \varepsilon$ gilt, was zu zeigen war.

Um umgekehrt zu zeigen, dass aus der Folgenstetigkeit das ε - δ -Kriterium folgt, gehen wir indirekt vor. Das bedeutet, dass wir annehmen, die Behauptung wäre falsch. Dann gäbe es ein $\varepsilon > 0$ derart, dass es für alle $\delta > 0$, insbesondere also für alle $\delta = \frac{1}{n}$ mit $n = 1, 2, 3, \dots$ ein $a_n \in D$ gibt mit $|a_n - x_0| < \frac{1}{n}$ aber $|f(a_n) - f(x_0)| \geq \varepsilon$. Damit wäre aber eine Folge mit $\lim_{n \rightarrow \infty} a_n = x_0$ gefunden, deren Werte $f(a_n)$ für $n \rightarrow \infty$ sicher nicht gegen $f(x_0)$ konvergieren, Widerspruch zur vorausgesetzten Folgenstetigkeit. \square

3.1.3 Vererbungseigenschaften stetiger Funktionen

Inhalt in Kurzfassung: Stetigkeit vererbt sich auf Summe, Differenz, Produkt, Quotienten (sofern definiert) und Verkettung stetiger Funktionen, sowohl lokal als auch global.

Mit Hilfe der Folgenstetigkeit lassen sich Grenzwertsätze für Folgen aus 2.1.3.1 fast unmittelbar auf Funktionen übertragen:

Satz 3.1.3.1. Seien die reellen Funktionen $f, g : D \rightarrow \mathbb{R}$ stetig in $x_0 \in D$, dann sind auch die Funktionen $f + g, f - g, fg : D \rightarrow \mathbb{R}$, definiert durch $(f + g)(x) := f(x) + g(x)$, $(f - g)(x) := f(x) - g(x)$, $(fg)(x) := f(x)g(x)$, stetig in x_0 . Das

gilt auch für die Funktion $\frac{f}{g}$, $\frac{f}{g}(x) := \frac{f(x)}{g(x)}$, sofern die zusätzliche Voraussetzung $g(x) \neq 0$ für alle $x \in D$ erfüllt ist.

Beweis. Wir beweisen den Satz für die Summe, für die anderen Fälle ist die Argumentation ganz analog. Seien $a_n \in D$ mit $\lim_{n \rightarrow \infty} a_n = x_0$, so gilt wegen der Stetigkeit von f und g in x_0 auch $\lim_{n \rightarrow \infty} f(a_n) = f(x_0)$ und $\lim_{n \rightarrow \infty} g(a_n) = g(x_0)$. Weil der Grenzwert einer Summe mit der Summe der Grenzwerte übereinstimmt (siehe Satz 2.1.3.1, Teil 1), folgt daraus $\lim_{n \rightarrow \infty} (f+g)(a_n) = (f+g)(x_0)$. Damit ist $f+g$ folgenstetig in x_0 und somit wegen Satz 3.1.2.3 stetig in x_0 . \square

Übungsaufgabe 177. (E) Zeigen Sie, dass analog zu Folgen auch für Funktionen gilt:

Ist f (wenigstens in der Nähe von x_0) beschränkt und $\lim_{x \rightarrow x_0} g(x) = 0$, so folgt $\lim_{x \rightarrow x_0} f(x)g(x) = 0$.

Für die Verkettung von Funktionen muss man nur geringfügig anders argumentieren: Aus $a_n \rightarrow x_0$ für $n \rightarrow \infty$ folgt, wenn f in x_0 stetig ist, $f(a_n) \rightarrow f(x_0)$ und daraus, wenn g stetig in $f(x_0)$ ist, auch $(g \circ f)(a_n) = g(f(a_n)) \rightarrow g(f(x_0)) = (g \circ f)(x_0)$. Also:

Satz 3.1.3.2. Ist $f : D_f \rightarrow \mathbb{R}$ stetig in $x_0 \in D_f$ und $g : D_g \rightarrow \mathbb{R}$ mit $f(D_f) \subseteq D_g$ stetig in $f(x_0) \in D_g$, so ist auch $g \circ f : D_f \rightarrow \mathbb{R}$ stetig in x_0 . Folglich sind Verkettung stetiger Funktionen wieder stetig.

3.1.4 Beispiele zu Stetigkeit und Unstetigkeit

Inhalt in Kurzfassung: Mit Hilfe des Vorangegangenen erhält man leicht vielfältige Beispiele stetiger Funktionen. Im Gegensatz dazu braucht es zur Konstruktion unstetiger Funktionen neue Ideen. Mittels Definition durch Fallunterscheidung beispielsweise kann man die Signumfunktion definieren, deren Unstetigkeit an der Stelle 0 dem Typus Sprungstelle entspricht. Einem anderen Typus entsprechen Oszillationsstellen. Schließlich stellt die Dirichletsche Sprungfunktion ein Beispiel einer Funktion dar, die sogar an jeder Stelle unstetig ist. An typischen Polstellen hingegen ist eine Funktion gar nicht definiert, also weder stetig noch unstetig.

Aus stetigen Funktionen können, wie wir gesehen haben, durch Bildung von Summen, Differenzen, Produkten, Quotienten und Kompositionen nur wieder stetige Funktionen entstehen. Gehen wir von den konstanten Funktionen und der identischen Funktion mit $\text{Id}(x) := x$ (die ja alle stetig sind) aus, erhalten wir mittels Summen, Differenzen und Produkte sämtliche **Polynomfunktionen**, genannt auch **ganzzrationale Funktionen**. Das sind Funktionen f zu denen es ein $n \in \mathbb{N}$ und Koeffizienten $a_0, \dots, a_n \in \mathbb{R}$ gibt mit

$$f(x) = \sum_{i=0}^n a_i x^i$$

für alle $x \in \mathbb{R}$ (oder aus dem Definitionsbereich von f). Ist $a_n \neq 0$, so heißt n der Grad von f . Sind f und g Polynomfunktionen, so nennt man die Funktion $r := \frac{f}{g}$ eine **gebrochen rationale Funktion**. Man beachte: Als Definitionsbereich für Polynomfunktionen kann immer ganz \mathbb{R} genommen werden. Bei gebrochen rationalen Funktionen muss man die Nullstellen des Nenners ausnehmen.

Wir werden bald sehen, dass auch die trigonometrischen Funktionen, die Exponentialfunktion und der Logarithmus stetig sind. Wie sehen dann überhaupt unstetige Funktionen aus?

Die einfachste Form einer Unstetigkeit liegt dann vor, wenn eine Funktion an einem Punkt speziell definiert ist. So ist beispielsweise die Funktion f mit $f(x) := 0$ für alle $x \neq 0$ und $f(0) := 1$ unstetig bei 0, sonst überall stetig. In diesem Fall – genauer: wenn $\lim_{x \rightarrow x_0} f(x)$ existiert, aber von $f(x_0)$ verschieden ist – spricht man auch von einer **hebbaren Unstetigkeit** bei x_0 , weil ja eine geeignete Umdefinition von f nur an dieser einen Stelle die Unstetigkeit aufheben kann.

Die nächst einfache Form von Unstetigkeiten sind **Sprungstellen**, wo die einseitigen Grenzwerte $\lim_{x \rightarrow x_0^-} f(x)$ und $\lim_{x \rightarrow x_0^+} f(x)$ existieren, aber nicht übereinstimmen. Die Differenz $\lim_{x \rightarrow x_0^+} f(x) - \lim_{x \rightarrow x_0^-} f(x)$ nennt man die **Sprunghöhe** von f bei x_0 . Ein nützliches Beispiel ist die **Signumfunktion** sgn mit $\operatorname{sgn}(x) := -1$ für $x < 0$, $\operatorname{sgn}(0) := 0$ und $\operatorname{sgn}(x) := 1$ für $x > 0$. Egal wie $\operatorname{sgn}(0)$ undefiniert würde, die Unstetigkeit bei 0 kann nicht aufgehoben werden.

Es ist aber auch denkbar, dass nicht einmal einseitige Grenzwerte existieren. Das Standardbeispiel lautet $f(x) := \sin \frac{1}{x}$ für $x \neq 0$. Unabhängig davon, wie $f(0)$ definiert wird, die Funktion kann an dieser Stelle schon deshalb nicht rechts- oder linksseitig stetig sein, weil nicht einmal die einseitigen Grenzwerte bei 0 existieren. Man spricht von **Oszillation**.

Als extremstes Gegenbeispiel zur Stetigkeit dient uns die überall unstetige sogenannte **Dirichletsche Sprungfunktion**.¹ Sie nimmt an allen rationalen Stellen den Wert 1, an allen irrationalen den Wert 0 an. Im Zusammenhang mit ihr führen wir als allgemeine Notation ein: Für $f : D \rightarrow \mathbb{R}$ und $A \subseteq D$ schreiben wir auch $f = \mathbf{1}_A$, wenn $f(x) = 1$ für alle $x \in A$ und $f(x) = 0$ für alle $x \in D \setminus A$. Die Funktion $\mathbf{1}_A$ heißt die **charakteristische Funktion** oder auch **Indikatorfunktion** der Menge A auf D . Die Dirichletsche Sprungfunktion f ist also die charakteristische oder Indikatorfunktion der Menge \mathbb{Q} , $f = \mathbf{1}_{\mathbb{Q}} : \mathbb{R} \rightarrow \{0, 1\}$.

Übungsaufgabe 178. (E) Erklären Sie, warum die Dirichletsche Sprungfunktion an jedem $x \in \mathbb{R}$ unstetig ist.

Übungsaufgabe 179. (E) Zeigen Sie, dass die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} x, & \text{für } x \in \mathbb{Q} \\ 0, & \text{für } x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases}$$

ausschließlich an der Stelle 0 stetig ist.

¹nach dem deutschen Mathematiker Johann Peter Gustav Lejeune Dirichlet (1805-1859)

Eine besondere Erwähnung verdienen **Polstellen** wie die Stelle 0 für die Funktion $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \frac{1}{x}$. Weil 0 nicht im Definitionsbereich liegt, kann dort streng genommen weder von Stetigkeit noch von Unstetigkeit gesprochen werden. Trotzdem werden im Zusammenhang mit Unstetigkeiten oft auch Polstellen behandelt.

Übungsaufgabe 180. (T) *Existieren folgende Grenzwerte? Begründen Sie Ihre Antworten anhand der Definition des Grenzwertes für Funktionen.*

1. $\lim_{x \rightarrow 0^+} (\operatorname{sgn}(x) + 1),$
2. $\lim_{x \rightarrow 0^-} (\operatorname{sgn}(x) + 1),$
3. $\lim_{x \rightarrow 0} (\operatorname{sgn}(x) + 1),$
4. $\lim_{x \rightarrow 0^+} \frac{1}{x},$
5. $\lim_{x \rightarrow 0^-} \frac{1}{x},$
6. $\lim_{x \rightarrow 0} \frac{1}{x}.$

Übungsaufgabe 181. (T) *Skizzieren Sie folgende Funktionen und argumentieren Sie, wo die Funktionen stetig bzw. unstetig sind.*

$$1. f(x) = \begin{cases} 2x^2, & \text{für } x < 1, \\ 3, & \text{für } x = 1, \\ \frac{1}{3-x}, & \text{für } x > 1. \end{cases}$$

$$2. f(x) = \operatorname{sgn}((-x-1)(-x+1)(-x+3)).$$

Übungsaufgabe 182. (T) *Gegeben ist die Funktion*

$$f_{a,b}(x) = \begin{cases} x^2, & \text{für } x < -1, \\ ax + b, & \text{für } -1 \leq x < 1, \\ \frac{1}{2bx-1}, & \text{für } x \geq 1. \end{cases}$$

Bestimmen Sie jene Parameter $a, b \in \mathbb{R}$ für die $f_{a,b}$ auf ganz \mathbb{R} stetig ist.

Übungsaufgabe 183. (T) *Wählen Sie ein $\alpha \in \mathbb{R}$ so, dass die Funktion*

$$f(x) = \frac{x^2 + 2x + \alpha}{x - 3}$$

auf ganz \mathbb{R} stetig fortsetzbar ist und begründen Sie Ihre Vorgehensweise.

3.1.5 Konsequenzen der Vollständigkeit von \mathbb{R} für stetige Funktionen

Inhalt in Kurzfassung: So wie über Folgen gibt es auch über Funktionen zahlreiche wichtige Ergebnisse, die entscheidend von der Vollständigkeit von \mathbb{R} abhängen, darüber hinaus auch von der Stetigkeit. Wir behandeln hier die folgenden: Zwischenwertsatz für stetige Funktionen, Satz von der stetigen Umkehrfunktion und Satz vom Maximum.

Stetige Funktionen haben einige sehr wichtige Eigenschaften, die anschaulich ziemlich klar erscheinen und deren tieferer Grund in der Vollständigkeit von \mathbb{R} liegt. Die erste davon wird durch den **Zwischenwertsatz** beschrieben: Eine stetige Funktion auf einem Intervall nimmt zwischen je zwei Werten $f(a) < f(b)$ jeden Zwischenwert $y_0 \in [f(a), f(b)]$ an.

Satz 3.1.5.1. *Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ an. Insbesondere hat f eine Nullstelle in $[a, b]$, sofern $f(a)$ und $f(b)$ unterschiedliches Vorzeichen haben.*

Die Idee besteht in der Vorstellung, dass sich die Werte $f(x)$, wenn man bei $x = a$ beginnt und x sich kontinuierlich in Richtung b bewegt, entsprechend kontinuierlich (Stetigkeit!) von $f(a)$ nach $f(b)$ bewegen, den Wert y_0 irgendwann passieren und wegen der Vollständigkeit nicht überspringen können. Eine Exaktifizierung dieser Idee lautet wie folgt.

Beweis. Sei o.B.d.A. $f(a) < y_0 < f(b)$. Wir wollen ein $x_0 \in [a, b]$ mit $f(x_0) = y_0$ finden. Wir zerlegen \mathbb{R} in zwei Teile A und B . Dabei besteht A aus allen a' mit der Eigenschaft, dass es kein $x \in [a, b]$ gibt mit $x < a'$ und $f(x) > y_0$. $B := \mathbb{R} \setminus A$ bestehe aus dem Rest. Z.B. ist $a \in A$ und $b \in B$. Man überzeugt sich schnell, dass $a < b$ für alle $a \in A$ und $b \in B$ gilt, dass diese Zerlegung also die in der Definition der Vollständigkeit von \mathbb{R} geforderten Eigenschaften hat. Also gibt es ein x_0 mit $a' \leq x_0 \leq b'$ für alle $a' \in A$ und $b' \in B$, insbesondere ist $x_0 \in [a, b]$. Wegen der Stetigkeit von f existiert der Grenzwert $\lim_{x \rightarrow x_0} f(x) = f(x_0)$. Laut Konstruktion gilt $f(x) \leq y_0$ für alle $x < x_0$, was nur für $f(x_0) \leq y_0$ möglich ist. Ähnlich gilt $f(x) \geq y_0$ für Werte von $x > x_0$ beliebig nahe an x_0 . Das ist wiederum nur möglich, wenn $f(x_0) \geq y_0$ gilt. Beides zusammen bedeutet aber $f(x_0) = y_0$. \square

Eine anderer, populärer Beweis des Zwischenwertsatzes erfolgt mittels Intervallschachtelung. Die Idee besteht darin, das Intervall $[a, b]$ immer wieder zu halbieren und in jedem Schritt jene Hälfte zu nehmen, wo die Funktionswerte an den beide Randpunkten weder beide über noch beide unter y_0 liegen. Die so entstehende Folge von Intervallen zieht sich auf einen Punkt $x_0 \in [a, b]$ zusammen, an dem f aus Stetigkeitsgründen den Wert $f(x_0) = y_0$ annehmen muss.

Klarerweise gilt die Zwischenwertsatz für unstetige Funktionen im Allgemeinen nicht. Zum Beispiel könnte ein Zwischenwert y_0 ja an einer Sprungstelle übersprungen werden.

Übungsaufgabe 184. (P) *Eine an beiden Enden eingespannte Saite werde durch eine stetige Funktion $g : [0, 2] \rightarrow \mathbb{R}$ mit $g(0) = g(2) = 0$ dargestellt.*

1. *Zeigen Sie, dass es immer zwei Punkte $x_1, x_2 \in [0, 2]$ mit $|x_1 - x_2| = 1$ gibt, die gleich weit ausgelenkt werden, i.e. für die $g(x_1) = g(x_2)$ gilt. Hinweis: Wenden Sie den Zwischenwertsatz auf eine geeignete Hilfsfunktion $h(x)$ an.*
2. *Illustrieren Sie diesen Sachverhalt speziell für die Funktionen $g(x) = \sin(\frac{\pi}{2}x)$ und $g(x) = \sin(\pi(\frac{x}{2})^2)$.*

3. Geben Sie ein Beispiel einer stetigen Funktion $g : [0, 2] \rightarrow \mathbb{R}$ an, so dass Sie (außer $x_1 = 0$ und $x_2 = 2$) keine zwei Punkte x_1, x_2 mit $|x_1 - x_2| > 1$ und $g(x_1) = g(x_2)$ finden können.

Sei nun $f : [a, b] \rightarrow \mathbb{R}$ nicht nur stetig, sondern auch streng monoton, o.B.d.A. monoton wachsend, insbesondere also injektiv. Sei $f(a) = c$ und $f(b) = d$. Die Zwischenwerteigenschaft besagt dann, dass $f : [a, b] \rightarrow [c, d]$ surjektiv, also sogar bijektiv ist. Somit gibt es eine Umkehrfunktion $f^{(-1)} : [c, d] \rightarrow [a, b]$, die ebenfalls streng monoton (in diesem Fall wachsend) und stetig ist.

Übungsaufgabe 185. (E) Begründen Sie, dass $f^{(-1)}$ diese beiden Eigenschaften hat. Illustrieren Sie insbesondere anhand einer Skizze die ε - δ -Stetigkeit von $f^{(-1)}$.

Die gleichen Überlegungen lassen sich auf Funktionen anwenden, die nicht nur auf einem abgeschlossenen Intervall definiert sind. Es gilt also der **Satz von der stetigen Umkehrfunktion**:

Satz 3.1.5.2. Sei $f : D \rightarrow \mathbb{R}$ injektiv mit Umkehrfunktion. Dann ist $f^{(-1)} : f(D) \rightarrow D$ sicher dann stetig, wenn D kompakt (abgeschlossen und beschränkt) oder zusammenhängend ist. Im zusammenhängenden Fall ist f auch streng monoton.

Beispiel: Die Potenzfunktionen $f : x \mapsto x^n$ mit $n = 1, 2, \dots$ lassen sich als monotone Funktionen auf $[0, \infty)$ (für gerades n) oder sogar auf ganz \mathbb{R} (für ungerades n) zu den Wurzelfunktionen $x \mapsto x^{\frac{1}{n}} = \sqrt[n]{x}$ umkehren. Analoges werden wir mit der Exponentialfunktion tun und auf ihre Umkehrfunktion, den Logarithmus stoßen, sowie mit den trigonometrischen Funktionen auf geeigneten Intervallen, wo Monotonie herrscht.

Übungsaufgabe 186. (P) Führen Sie durch schrittweises Zerlegen der Funktion f in Ihre elementaren Bestandteile die Stetigkeit von f auf die Stetigkeit der Exponentialfunktion $x \mapsto e^x$ (diese dürfen Sie hier verwenden) und auf die Stetigkeit der identischen Abbildung $x \mapsto x$ zurück. Geben Sie einen möglichst großen Definitionsbereich für f an.

$$f(x) = \frac{|\ln(x^2 + 4)|}{\sqrt{1 - x^2}} + \ln\left((\ln(x) + e^x)^2\right).$$

Auf abgeschlossenen Intervallen haben stetige Funktionen neben der Zwischenwerteigenschaft noch weitere interessante Eigenschaften.

Hilfssatz 3.1.5.3. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann ist f beschränkt.

Beweis. Diesmal arbeiten wir mit dem Satz von Bolzano-Weierstraß, der ja auch auf der Vollständigkeit von \mathbb{R} beruht. Aus Symmetriegründen genügt es, einen Widerspruch aus der Annahme abzuleiten, f wäre nach oben unbeschränkt. In diesem Fall gäbe es nämlich zu jedem $n \in \mathbb{N}$ ein $a_n \in [a, b]$ mit $n \leq f(a_n)$. Nach dem Satz von Bolzano-Weierstraß gibt es eine Teilfolge der a_n , etwa a_{n_1}, a_{n_2}, \dots

mit $n_1 < n_2 < \dots$, die gegen einen Grenzwert $x_0 \in [a, b]$ konvergiert. Wegen der Folgenstetigkeit müssten die unbeschränkten $f(a_{n_k}) \geq n_k$ gleichzeitig gegen $f(x_0) \in \mathbb{R}$ konvergieren, was aber unmöglich ist. \square

Auch hier sieht man schnell, dass weder die Abgeschlossenheit des Intervalls $[a, b]$ noch die Stetigkeit von f als Voraussetzung weggelassen werden können. Für die Abgeschlossenheit dient als Beispiel die Funktion $f : (0, 1] \rightarrow \mathbb{R}$, $x \mapsto \frac{1}{x}$, für die Stetigkeit dieselbe Funktion auf $(0, 1]$, wenn wir sie zusätzlich auch noch an der Stelle $x = 0$ irgendwie definieren, etwa $f(0) := 0$.

Wegen der Beschränktheit gibt es zu einem f wie in Hilfssatz 3.1.5.3 auch Infimum und Supremum. Diese erweisen sich sogar als Minimum bzw. Maximum. Als Verschärfung von 3.1.5.3 gilt nämlich der **Satz vom Maximum**:

Satz 3.1.5.4. *Sei $f : [a, b] \rightarrow \mathbb{R}$, $a \leq b$, stetig. Dann nimmt f auf $[a, b]$ sowohl ein Minimum als auch ein Maximum an. (Auch hier genügt es, statt eines abgeschlossenen Intervalls einen beliebigen kompakten Definitionsbereich von f voranzusetzen.)*

Beweis. Der Beweis läuft ähnlich wie zuvor mittels Bolzano-Weierstraß. Allerdings gehen wir vom Supremum $s := \sup f([a, b])$ von f aus, das wegen der im Hilfssatz 3.1.5.3 bereits bewiesenen Beschränktheit (und weil der Wertebereich nicht leer ist) existiert. Zu jedem $n = 1, 2, \dots$ gibt es ein $a_n \in [a, b]$ mit $s - \frac{1}{n} < f(a_n) \leq s$. Wieder wählen wir aus den a_n eine konvergente Teilfolge a_{n_1}, a_{n_2}, \dots aus, an deren Grenzwert x_0 für den Funktionswert aus Stetigkeitsgründen $f(x_0) = \lim_{k \rightarrow \infty} f(a_{n_k}) = s$ gelten muss. Also ist s sogar Maximum. Analog zeigt man, dass f auf $[a, b]$ auch ein Minimum annimmt. \square

Übungsaufgabe 187. *(E) Der Beweis von Satz 3.1.5.4 wurde nur für das Maximum geführt. Führen Sie ihn auch für das Minimum, indem Sie $-f$ statt f betrachten und die Aussage über das Maximum verwenden.*

Auch hier zeigen einfache Beispiele, dass man die Voraussetzungen nicht abschwächen kann:

Übungsaufgabe 188. *(P) Finden Sie von den bereits behandelten Beispielen verschiedene reelle Funktionen f und g auf geeigneten Definitionsbereichen (diese sind unbedingt anzugeben!) mit folgenden Eigenschaften:*

1. f ist stetig, nimmt aber weder Maximum noch Minimum an. (Wie muss der Definitionsbereich jedenfalls beschaffen sein?)
2. f ist auf einem abgeschlossenen Intervall $[a, b]$ definiert, nimmt aber weder Maximum noch Minimum an. (Wie muss f jedenfalls beschaffen sein?)

3.1.6 Fixpunktiteration

Inhalt in Kurzfassung: Wir kombinieren das erstmals in 2.1.8 angeschnittene Thema Iterationsfolgen $a_{n+1} := T(a_n)$ mit dem aktuellen Hauptthema Stetigkeit, indem wir für $T : D \rightarrow D$ mit $D \subseteq \mathbb{R}$ Stetigkeit voraussetzen. Ist D

abgeschlossen, so gilt für sogenannte Kontraktionen T das Kontraktionsprinzip (der Banachsche Fixpunktsatz), wonach bei beliebigem Startpunkt a_0 Konvergenz gegen den einzigen Fixpunkt von T eintritt.

Wir schließen an 2.1.8 an, wo wir Iterationsfolgen $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ betrachtet haben, die, beginnend mit einem Startwert a_0 , durch eine Rekursion $a_{n+1} = T(a_n)$ mit einer Funktion $T : D \rightarrow D$ zustande kommen. Nun konzentrieren wir uns auf reelle Funktionen T , also $D \subseteq \mathbb{R}$, die überdies stetig sind. Noch mehr als in 2.1.8 spielen Fixpunkte bei stetigem T eine besondere Rolle, weil nur sie als Grenzwerte $x = \lim_{n \rightarrow \infty} a_n$ von Iterationsfolgen in Frage kommen. Für den sehr einfachen Beweis dieser Aussage verwenden wir die Folgenstetigkeit von T (Satz 3.1.2.3):

$$T(x) = T\left(\lim_{n \rightarrow \infty} a_n\right) = \lim_{n \rightarrow \infty} T(a_n) = \lim_{n \rightarrow \infty} a_{n+1} = x.$$

Ist beispielsweise $D = [a, b]$, so muss T mindestens einen Fixpunkt haben. Denn die Funktion $f(x) := T(x) - x$ erfüllt $f(a) = T(a) - a \geq 0$ und $f(b) = T(b) - b \leq 0$, muss also nach dem Zwischenwertsatz 3.1.5.1 eine Nullstelle x_0 haben. Folglich gilt $T(x_0) = x_0$, und x_0 ist ein Fixpunkt von T .

Bei Konvergenzuntersuchungen ist das eine große Hilfe. Allerdings ist dadurch nicht gesichert, dass bei Iteration tatsächlich Konvergenz eintritt. Das kann oft mit dem **Fixpunktprinzip** garantiert werden. Die hinreichende Voraussetzung besteht darin, dass T eine sogenannte **Kontraktion** ist. Das soll bedeuten, dass es eine Zahl $\lambda < 1$ gibt (eine sogenannte **Kontraktionskonstante**) mit $|T(x_1) - T(x_2)| \leq \lambda|x_1 - x_2|$ für alle $x_1, x_2 \in D$. Insbesondere ist so ein T stetig, weil in der Stetigkeitsdefinition $\delta = \varepsilon$ gesetzt werden kann. Man beachte, dass eine Kontraktion nur einen Fixpunkt haben kann, weil für Fixpunkte $x_1 = T(x_1), x_2 = T(x_2)$ die Beziehung $|x_1 - x_2| = |T(x_1) - T(x_2)| \leq \lambda|x_1 - x_2|$ bei $\lambda < 1$ nur für $|x_1 - x_2| = 0$, also $x_1 = x_2$ möglich ist.

Nun zur Konvergenz der Iterationsfolge: Startet man mit irgendeinem $a_0 \in D$, so folgt dann $|a_2 - a_1| = |T(a_1) - T(a_0)| \leq \lambda|a_1 - a_0|$, analog $|a_3 - a_2| \leq \lambda|a_2 - a_1| \leq \lambda^2|a_1 - a_0|$ und allgemein $|a_{n+1} - a_n| \leq \lambda^n|a_1 - a_0|$ (Induktion nach n). Die Abstände verschiedener Glieder werden also kleiner, mindestens so schnell wie die Glieder einer konvergenten geometrischen Reihe; als Formel für $n_1 < n_2$:

$$|a_{n_2} - a_{n_1}| \leq \sum_{k=n_1}^{n_2-1} |a_{k+1} - a_k| \leq \sum_{k=n_1}^{n_2-1} \lambda^k |a_1 - a_0| \leq |a_1 - a_0| \lambda^{n_1} \sum_{n=0}^{\infty} \lambda^n = \frac{|a_1 - a_0|}{1 - \lambda} \lambda^{n_1}$$

Der Faktor $\frac{|a_1 - a_0|}{1 - \lambda}$ auf der rechten Seite ist eine feste Zahl. Der letzte Faktor λ^{n_1} konvergiert für $n_1 \rightarrow \infty$ gegen 0. Also gibt es zu jedem $\varepsilon > 0$ ein n_0 derart, dass $|a_{n_2} - a_{n_1}| < \varepsilon$ für alle $n \geq n_0$ gilt. Es liegt also eine Cauchyfolge vor, weshalb die a_n gegen ein $x \in \mathbb{R}$ konvergieren (Satz 2.1.7.2). Ist D abgeschlossen, so muss dieser Grenzwert wieder in D liegen. Wegen der Stetigkeit von T kommt nach obigen Überlegungen nur der eindeutige Fixpunkt von T in Frage. Damit ist das sogenannte **Kontraktionsprinzip** (der **Banachsche Fixpunktsatz**) für reelle Folgen bewiesen:

Satz 3.1.6.1. Sei $D \subseteq \mathbb{R}$ abgeschlossen und $T : D \rightarrow D$ eine Kontraktion. Dann hat T genau einen Fixpunkt x . Für jeden Startwert $a_0 \in D$ konvergiert die durch $a_{n+1} := T(a_n)$ definierte rekursive Folge gegen diesen Fixpunkt x .

Beispiele von Kontraktionen sind Funktionen $T : \mathbb{R} \rightarrow \mathbb{R}$ mit $T(x) := kx + d$, sofern $|k| < 1$. Der Parameter $|k|$ ist dann die Kontraktionskonstante λ . Die Fixpunktgleichung $x = T(x) = kx + d$ hat die eindeutige Lösung $x = \frac{d}{1-k}$. Die Beispiele zu rekursiven Folgen an früherer Stelle erscheinen nun in einem neuen Licht. Es sei auch an die dort erwähnten Spinnwebdiagramme erinnert.

Übungsaufgabe 189. (T) Untersuchen Sie die rekursiv definierte Folge auf Konvergenz:

$$a_0 := 2, \quad \text{und} \quad a_{n+1} := \frac{3}{4 - a_n} \quad \text{für } n \geq 1.$$

Übungsaufgabe 190. (T) Wie Übungsaufgabe 189 aber mit den Startwerten $a_0 = 0$ bzw. $a_0 = 5$.

Aus den bisherigen Überlegungen ist ersichtlich, dass die Fixpunktiteration umso schneller konvergiert, je kleiner die Kontraktionskonstante λ ist. Wir werden diesen Gedanken nochmals in 4.4.5 beim Newton-Verfahren aufgreifen, wo λ bei Annäherung an den Fixpunkt sogar gegen 0 konvergiert.

3.1.7 Stetigkeit und gleichmäßige Konvergenz

Inhalt in Kurzfassung: Für eine Folge von Funktionen $f_n : D \rightarrow \mathbb{R}$, für die der Grenzwert $f(x) := \lim_{n \rightarrow \infty} f_n(x)$ für alle $x \in D$ existiert, liegt es nahe, die resultierende Funktion $f : D \rightarrow \mathbb{R}$ als (punktweisen) Grenzwert oder Grenzfunktion der Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ zu definieren. Dabei kann das überraschende Phänomen auftreten, dass alle f_n stetig sind, f aber nicht. (Beispiel: $f_n(x) := x^n$ auf $D := [0, 1]$ an der Stelle 1.) Analysiert man die Situation, stößt man auf den stärkeren Begriff der gleichmäßigen Konvergenz einer Funktionenfolge auf D , der sehr wohl garantiert, dass die Grenzfunktionen stetiger f_n selbst wieder stetig ist.

Nun betrachten wir nicht mehr einzelne Funktionen, sondern Folgen reeller Funktionen $f_n : D \rightarrow \mathbb{R}$, z.B. $D = [0, 1]$ und $f_n(x) := x^n$, $n \in \mathbb{N}$. Für jedes $x \in D$ mit $x < 1$ herrscht die Konvergenz $\lim_{n \rightarrow \infty} f_n(x) = 0$, lediglich für $x = 1$ konvergieren die $f_n(x) = f_n(1) = 1$ gegen 1. Es liegt nahe, die Funktion $f : [0, 1] \rightarrow \mathbb{R}$, $f(x) := 0$ für $x < 1$ und $f(1) := 1$ als **(punktweise) Grenzfunktion** oder als **(punktweisen) Limes** $f = \lim_{n \rightarrow \infty} f_n$ zu bezeichnen. Überraschend mag anmuten, dass f an der Stelle $x = 1$ nicht stetig ist, obwohl alle f_n stetig sind. Dieses Phänomen kann nur auftreten, wenn die Konvergenz nicht gleichmäßig ist. Die Definition lautet:

Definition 3.1.7.1. Man sagt, die Funktionen $f_n : D \rightarrow \mathbb{R}$, $n \in \mathbb{N}$, **konvergieren gleichmäßig** gegen $f : D \rightarrow \mathbb{R}$, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt derart, dass für alle $n \geq n_0$ und für alle $x \in D$ gilt: $|f_n(x) - f(x)| < \varepsilon$.

Gleichmäßige Konvergenz gegen f auf D lässt sich veranschaulichen mit Hilfe eines ε -Schlauchs $S_\varepsilon = \{(x, y) : x \in D, f(x) - \varepsilon < y < f(x) + \varepsilon\}$ um f . Wie eng dieser Schlauch auch gelegt, d.h. wie klein $\varepsilon > 0$ auch gewählt wird, ab einem gewissen Index n_0 liegen sämtliche f_n mit $n \geq n_0$ ganz in S_ε . Beim Beispiel mit $f_n(x) = x^n$ ist das nicht der Fall, weil bei $\varepsilon < 1$ diese Schlauchbedingung von x -Werten, die hinreichend nahe bei 1 liegen, verletzt wird.

Der logische Unterschied zwischen punktweiser und gleichmäßiger Konvergenz zeigt sich wieder einmal an der Rolle der logischen Quantoren. Die Formeln für punktweise bzw. gleichmäßige Konvergenz lauten:

$$\begin{aligned} \forall \varepsilon > 0 \quad \forall x \in D \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : |f_n(x) - f(x)| < \varepsilon \\ \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall x \in D \quad \forall n \geq n_0 : |f_n(x) - f(x)| < \varepsilon \end{aligned}$$

Wichtig ist, dass der Begriff der gleichmäßigen Konvergenz immer nur in Bezug auf einen Definitionsbereich D sinnvoll ist.

Übungsaufgabe 191. (P) Wir betrachten wieder die Funktionen $f_n(x) := x^n$, $n \in \mathbb{N}$, mit der Grenzfunktion $f(x) = 0$ für $x < 1$ und $f(1) = 1$, diesmal aber auf unterschiedlichen Definitionsbereichen $D \subseteq [0, 1]$. Begründen Sie:

1. Ist $D = [0, a]$ mit $a < 1$, so ist die Konvergenz gleichmäßig.
2. Auf keinem der Intervalle $D := [1 - \varepsilon, 1]$ mit $\varepsilon > 0$ ist die Konvergenz gleichmäßig.
3. Die Konvergenz ist genau dann gleichmäßig auf D , wenn 1 kein Häufungspunkt von D ist, d.h. wenn es keine Folge \mathbf{a} mit Folgengliedern $a_n \in D$ und $a_n \neq 1$ für alle n gibt, so dass $\lim_{n \rightarrow \infty} a_n = 1$.

Die allgemein interessante Aussage über Stetigkeit und gleichmäßige Konvergenz lautet nun:

Satz 3.1.7.2. Angenommen, die reellen Funktionen $f_n : D \rightarrow \mathbb{R}$ sind stetig und konvergieren gleichmäßig gegen $f : D \rightarrow \mathbb{R}$. Dann ist auch f stetig.

Beweis. Sei $x_0 \in D$. Für den Beweis, dass f in x_0 stetig ist, sei $\varepsilon > 0$ und n_0 so gewählt, dass für alle $x \in D$ und alle $n \geq n_0$ die Ungleichung $|f_n(x) - f(x)| < \frac{\varepsilon}{3}$ gilt (gleichmäßige Konvergenz). Weiters gibt es wegen der Stetigkeit von f_{n_0} in x_0 ein δ derart, dass $|f_{n_0}(x) - f_{n_0}(x_0)| < \frac{\varepsilon}{3}$ für alle $x \in D$ mit $|x - x_0| < \delta$ gilt. Laut Dreiecksungleichung folgt für solche x

$$|f(x) - f(x_0)| \leq |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(x_0)| + |f_{n_0}(x_0) - f(x_0)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

was die Stetigkeit von f bei x_0 und, weil $x_0 \in D$ beliebig war, von f beweist. \square

Gleichmäßige Konvergenz garantiert also, dass der Grenzwert stetiger Funktionen (für $n \rightarrow \infty$) wieder stetig ist. Weil Stetigkeit selbst über Grenzwerteigenschaften definiert ist (nämlich für $x \rightarrow x_0$), können wir die Aussage des Satzes als Formel auch so schreiben:

$$\lim_{x \rightarrow x_0} \lim_{n \rightarrow \infty} f_n(x) = \lim_{x \rightarrow x_0} f(x) = f(x_0) = \lim_{n \rightarrow \infty} f_n(x_0) = \lim_{n \rightarrow \infty} \lim_{x \rightarrow x_0} f_n(x).$$

Satz 3.1.7.2 lässt sich also auch so lesen: Bei gleichmäßiger Konvergenz dürfen zwei bestimmte Grenzübergänge vertauscht werden, ohne das Ergebnis zu verändern. Es gibt in der Mathematik viele verwandte Aussagen, wobei teilweise andere Ausprägungen des Grenzwertbegriffs wie Differentiation und Integration ins Spiel kommen. Ein paar davon werden wir später noch kennen lernen.

Übungsaufgabe 192. (P) Für $n \in \mathbb{N}$ sei

$$g_n(x) = \begin{cases} \frac{2}{2nx + \frac{1}{2nx}}, & \text{für } x > 0 \\ 0, & \text{für } x = 0. \end{cases}$$

1. Zeigen Sie, dass g_n auf $[0, \infty)$ für jedes $n \in \mathbb{N}$ eine stetige Funktion mit Maximum 1 ist. Verwenden Sie dazu, dass $x \mapsto x + \frac{1}{x}$ genau für $x = 1$ ein Minimum besitzt.
2. Skizzieren Sie die Funktionsgraphen von g_n für $n = 1, 2, 3$.
3. Bestätigen Sie, dass die Funktionenfolge der g_n auf $[0, \infty)$ punktweise, aber nicht gleichmäßig gegen die Funktion $g(x) = 0$ (Nullfunktion) konvergiert. Die g_n konvergieren jedoch gleichmäßig auf jedem Teilintervall der Form $[a, \infty)$, mit $a > 0$.

Übungsaufgabe 193. (P) Betrachten Sie die reelle Funktion $f_{x_0}(x) = \frac{1}{1+(x-x_0)^2}$.

1. Setzen Sie $x_0 = 0$ und argumentieren Sie, warum f_0 stetig und beschränkt ist sowie ein (globales) Maximum in $x = 0$ besitzt. Fertigen Sie eine Skizze an.
2. Überlegen Sie sich, wie Sie aus dem Funktionsgraphen von f_0 sofort und ohne weitere Rechnung die Funktionsgraphen von f_{x_0} für beliebiges x_0 gewinnen können.
3. Betrachten Sie die Funktionenfolge $(f_n)_{n=0}^{\infty}$ (d.h. der Parameter x_0 durchläuft nun der Reihe nach alle natürlichen Zahlen). Bestimmen Sie das Konvergenzverhalten dieser Funktionenfolge (punktweise und/oder gleichmäßig) und berechnen Sie die Grenzfunktion.

Übungsaufgabe 194. (P) Finden Sie jeweils ein (vom bisher untersuchten verschiedenes) Beispiel von (paarweise verschiedenen) stetigen Funktionen $f_n : [0, 1] \rightarrow \mathbb{R}$, die eine punktweise Grenzfunktion f haben, und für die zusätzlich gilt:

1. Die Konvergenz ist gleichmäßig auf $[0, 1]$. (Folglich ist f stetig.)
2. Die Konvergenz ist nicht gleichmäßig auf $[0, 1]$, f ist aber trotzdem stetig.
3. Die Konvergenz ist nicht gleichmäßig auf $[0, 1]$, und f ist nicht stetig.

Sie müssen die f_n und f nicht unbedingt durch eine Formel angeben. Es genügt, wenn Sie Ihre Ideen mit Hilfe qualitativer Skizzen erläutern.

3.2 Polynome und rationale Funktionen (Elemente der Algebra)

Polynome sind, grob gesprochen, jene Funktionen, die sich allein aus Addition, Subtraktion und Multiplikation ergeben. In mancherlei Hinsicht ähneln sie den ganzen Zahlen. Nimmt man auch noch die Division hinzu, gelangt man zu den gebrochen rationalen Funktionen. Das Studium von Polynomen, insbesondere die Nullstellenbestimmung, bildet ein Herzstück jenes Teilgebietes der Mathematik, das man Algebra nennt. Wichtig sind Polynome, weil man einerseits mit ihnen und gebrochen rationalen Funktionen in vielerlei Hinsicht besser umgehen kann als mit beliebigen Funktionen, andererseits, weil Funktionen aus wesentlich größeren Klassen durch Polynome (und gebrochen rationale Funktionen) angenähert werden können. Diese Anwendungen werden in der Differentialrechnung (Kapitel 4) deutlich werden. In diesem Abschnitt beschäftigen wir uns zunächst in 3.2.1 und 3.2.2 mit relativ elementaren Aspekten. Einer der wichtigsten Sätze der Mathematik ist der sogenannte Fundamentalsatz der Algebra, dem zufolge jedes nichtkonstante komplexe Polynom mindestens eine Nullstelle hat, siehe 3.2.3. Viele Fragen über Polynome laufen auf Nullstellenbestimmung hinaus, mit der wir uns in 3.2.4 beschäftigen. Erste Ergebnisse darüber, wie weit Polynome andere Funktionen annähern/simulieren können, finden sich in 3.2.5. Abschließend behandeln wir in 3.2.6 die Partialbruchzerlegung gebrochen rationaler Funktionen.

3.2.1 Definition, Auswertung, Polynomdivision

Inhalt in Kurzfassung: Im Zusammenhang mit Polynomen und gebrochen rationalen Funktionen werden einige elementare Begriffe wiederholt, die aus der Schule weitgehend bekannt sein dürften: Koeffizienten und Grad eines Polynoms; konstante, lineare, quadratische und kubische Polynome;

Die Menge der **Polynomfunktionen** oder auch **ganzrationalen Funktionen** (etwas schlampig werden wir oft auch schlicht von **Polynomen** sprechen) entsteht, wenn man ausgehend von den konstanten reellen Funktionen $c_r : x \mapsto r$, $r \in \mathbb{R}$, und der Identität $\text{Id}_{\mathbb{R}} : x \mapsto x$ sämtliche Summen, und Produkte bildet. Wir bezeichnen diese Menge mit $\mathbb{R}[x]$. Man macht sich leicht klar, dass jedes Polynom $f \in \mathbb{R}[x]$ von der Gestalt

$$f : x \mapsto \sum_{i=0}^n a_i x^i$$

mit einem $n \in \mathbb{N}$ und sogenannten **Koeffizienten** $a_i \in \mathbb{R}$ (bzw. $a_i \in \mathbb{C}$ für $f \in \mathbb{C}[x]$) ist. Ist $a_n \neq 0$, so heißt n der **Grad**² von f , symbolisch $n = \text{grad} f$. Polynome vom Grad 0 heißen **konstant**, solche von Grad ≤ 1 heißen auch

²Wir werden später sehen, dass sowohl Grad als auch Koeffizienten durch die Funktion eindeutig bestimmt sind. Nur deshalb ist diese Definition überhaupt sinnvoll.

linear (i.A. inhomogen, also nicht im Sinne der Linearitätsdefinition in Kapitel 1), solche vom Grad 2 **quadratisch**, von Grad 3 **kubisch**. Die Nullfunktion ist auch ein Polynom (nämlich mit $n = 0$ und $a_n = a_0 = 0$), das **Nullpolynom**, welches wir mit 0 bezeichnen. Nach obiger Definition hätte das Nullpolynom keinen Grad, weil ja auch $a_0 = 0$ ist. Es erweist sich als praktisch, ihr den Grad $-\infty$ zuzuordnen.

Man kann genauso komplexe Koeffizienten a_i zulassen. So erhält man die Menge $\mathbb{C}[x]$ der komplexen Polynome. Auch die nachfolgenden Untersuchungen lassen sich gleichermaßen auf diesen Fall übertragen. Vorläufig denken wir aber an reelle Polynome.

Die elementarste Aufgabe besteht darin, den Wert eines Polynoms $f : x \mapsto \sum_{i=0}^n a_i x^i$ an einer bestimmten Stelle x_0 auszuwerten. Dies erfolgt effizient mit dem **Hornerschema**, dem die Darstellung

$$f(x) = a_0 + x(a_1 + x(a_2 + \dots + x(a_{n-1} + xa_n) \dots))$$

zugrunde liegt. Setzt man für x speziell x_0 ein, hat man, mit a_n beginnend, abwechselnd Multiplikationen (jeweils Zwischenergebnis mal x_0) und ebenso viele, nämlich n Additionen (jeweils Zwischenergebnis plus ein a_i) durchzuführen. Das ist weniger als bei einer naiven Berechnung aller Potenzen x_0^i , Multiplikation mit den a_i und Summation der Zwischenergebnisse $a_i x_0^i$.

Lässt man neben Addition, Subtraktion und Multiplikation auch Division zu, so erhält man die Menge $\mathbb{R}(x)$ der sogenannten (**gebrochen**) **rationalen Funktionen**, von denen jede die Gestalt $f(x) = \frac{f_1(x)}{f_2(x)}$ mit $f_1, f_2 \in \mathbb{R}[x]$ hat. Aus dem Definitionsbereich von f sind die Nullstellen von f_2 auszunehmen. Auf dem verbleibenden Definitionsbereich ist die Funktion aber stetig (siehe Satz 3.1.3.1).

Auch die Verkettung ist innerhalb von $\mathbb{R}[x]$ bzw. $\mathbb{R}(x)$ möglich:

Übungsaufgabe 195. (P) Begründen Sie: Sind f, g Polynome bzw. gebrochen rationale Funktionen, dann auch die Verkettung $g \circ f$. Was lässt sich im Fall von Polynomen über den Grad sagen?

Übungsaufgabe 196. (T) Geben Sie eine rationale Funktion an, die

1. genau bei $x = 1$ und $x = 0$ Nullstellen hat, die bei $x = -1$ gegen $-\infty$ divergiert und für alle reellen Zahlen außer -1 definiert ist.
2. genau bei $x = 1, x = 0$ und $x = 2$ Nullstellen hat, die bei $x = -3$ gegen $+\infty$ divergiert und für alle reellen Zahlen außer -3 definiert ist.

Skizzieren Sie Ihre Funktion. Können Sie noch eine zweite solche Funktion angeben?

Eine wichtige Rolle wird für uns die Analogie zwischen $\mathbb{R}[x]$ und \mathbb{Z} wie auch zwischen $\mathbb{R}(x)$ und \mathbb{Q} spielen. Beispielsweise ist in $\mathbb{R}[x]$ Division nur in besonderen Fällen möglich. Ist $f = \frac{f_1}{f_2}$ mit Polynomen $f_1, f_2 \in \mathbb{R}[x]$ selbst wieder eine Polynomfunktion (und als solche auf ganz \mathbb{R} definiert), so nennen wir f_2 einen

Teiler von f_1 . In jedem Fall ist aber, so wie in \mathbb{Z} , **Division mit Rest** möglich: Zu je zwei Polynomen f_1 und f_2 gibt es Polynome q und r mit $f_1 = qf_2 + r$ derart, dass $\text{grad}(r) < \text{grad}(f_2)$. Im Falle von Teilbarkeit $f_2|f_1$ ist $r = 0$.

Das Verfahren der Polynomdivision wird schnell klar an einem Beispiel: Sei $f_1(x) = 2x^3 + 2x^2 - 3x + 5$, $f_2(x) = x^2 - x + 2$. Um die Gleichung $f_1 = qf_2 + r$ in der behaupteten Weise zu erfüllen muss q offenbar linear, d.h. von der Form $q(x) = ax + b$ sein und r ebenfalls höchstens linear, also $r(x) = cx + d$ ($a, b, c, d \in \mathbb{R}$). Ausmultiplizieren führt zur Gleichung

$$(qf_2+r)(x) = (ax+b)(x^2-x+2)+cx+d = ax^3+(-a+b)x^2+(2a-b+c)x+(2b+d).$$

Koeffizientenvergleich mit dem Polynom $f(x) = 2x^3 + 2x^2 - 3x + 5$, beginnend mit dem Glied x^3 , führt zu den Gleichungen $a = 2$, $-a + b = 2$, $2a - b + c = -3$ und $2b + d = 5$. Diese haben die Lösung $a = 2$, $b = 4$, $c = -3$ und $d = -3$. Also bilden die Polynome $q(x) = 2x + 4$ und $r(x) = -3x - 3$ die Lösung der Aufgabe. Man macht sich leicht klar, dass sich dasselbe Verfahren allgemein auf beliebige Polynome anwenden lässt. Zweckmäßigerweise notiert man die Polynomdivision etwas sparsamer, ähnlich der Division ganzer Zahlen mit Rest, wie sie schon in der Volksschule gelehrt wird.

Übungsaufgabe 197. (E) Überlegen Sie sich ein praktisches Schema für die händische Polynomdivision und illustrieren Sie es am obigen Zahlenbeispiel. (Vielleicht ist Ihnen ein solches schon aus der Schule in Erinnerung.)

Interessant ist der Fall, wo $f_2(x) = x - \alpha$ mit einem $\alpha \in \mathbb{R}$. Dann hat der Rest $r = r_0$ in $f(x) := f_1(x) = q(x)(x - \alpha) + r_0$ den Grad 0, ist also eine Konstante. Dividiert man auch $q_0 := q$ mit Rest durch $f_2(x) = x - \alpha$, erhält man analog $q_0(x) = q_1(x)(x - \alpha) + r_1$ usw. Setzt man ein, so bedeutet das

$$f(x) = r_0 + (x - \alpha)q(x) = r_0 + (x - \alpha)(r_1 + (x - \alpha)(r_2 + \dots)) = \sum_{i=0}^n r_i(x - \alpha)^i.$$

Man spricht von der Entwicklung von f um die Stelle α . Schon aus dem Ergebnis der ersten Polynomdivision liest man $f(\alpha) = q(\alpha)(\alpha - \alpha) + r_0 = r_0$ ab.

3.2.2 Nullstellen und Eindeutigkeitssatz

Inhalt in Kurzfassung: Hat ein Polynom f eine Nullstelle α , so lässt sich (wie man mittels Polynomdivision einsieht) ein Linearfaktor abspalten zu $f(x) = q(x)(x - \alpha)$. Weil sich ein Polynom höchstens in so viele Linearfaktoren zerlegen lässt, wie sein Grad beträgt, schließt man daraus, dass es auch nicht mehr Nullstellen haben kann.

Polynome vom Grad 0 (also konstante Polynome $\neq 0$) haben keine Nullstelle, Polynome vom Grad 1 genau eine. Denn für $f(x) = a_1x + a_0$ mit $a_1 \neq 0$ lässt sich die Nullstellengleichung $f(x) = 0$ eindeutig lösen, und wir erhalten die Nullstelle $\alpha = -\frac{a_0}{a_1}$ lösen. Wenn sich ein beliebiges Polynom als Produkt anderer Polynome

schreiben lässt, so erbt es alle Nullstellen dieser Polynome. Speziell bedeutet dies: Ist $f = f_1 g$ das Produkt eines linearen Polynoms f_1 (eines sogenannten **Linearfaktors** von f) und eines beliebigen anderen Polynoms g , so bedeutet dies, dass auch f zumindest eine Nullstelle hat, nämlich die von f_1 . Wir wollen uns nun überlegen, dass auch eine Art Umkehrung gilt: Hat f eine Nullstelle, dann auch einen Linearfaktor mit dieser Nullstelle.

Dazu nehmen wir an, dass $\alpha \in \mathbb{R}$ (analog mit $\alpha \in \mathbb{C}$) eine Nullstelle des Polynoms f ist, dass also $f(\alpha) = 0$ gilt. Wie bei der Entwicklung von f um α dividieren wir f durch das Polynom $x - \alpha$ und erhalten $f(x) = q_1(x)(x - \alpha) + r$ mit einem geeigneten Polynom q_1 und $r \in \mathbb{R}$. Setzen wir $x = \alpha$, erhalten wir $0 = f(\alpha) = q_1(\alpha)(\alpha - \alpha) + r = r$, also $f(x) = q_1(x)(x - \alpha)$. Das lineare Polynom $x - \alpha$ teilt also f .

Angenommen, f hat mehrere Nullstellen $\alpha = \alpha_1, \alpha_2, \dots, \alpha_k$, die wir als voneinander verschieden annehmen. Für $i = 2, 3, \dots, k$ ist dann in $0 = f(\alpha_i) = q_1(\alpha_i)(\alpha_i - \alpha)$ der letzte Faktor von 0 verschieden, folglich gilt $q_1(\alpha_i) = 0$. Das Polynom q_1 hat also mit der möglichen Ausnahme von $\alpha = \alpha_1$ alle Nullstellen von f selbst als Nullstelle. Wenden wir unsere frühere Überlegung auf q_1 statt auf f an, zeigt das $q_1(x) = q_2(x)(x - \alpha_2)$ mit einem geeigneten q_2 etc. Setzen wir der Reihe nach ein, erhalten wir

$$f(x) = (x - \alpha_1)q_1(x) = (x - \alpha_1)(x - \alpha_2)q_2(x) = (x - \alpha_1) \dots (x - \alpha_k)q_k(x)$$

mit geeigneten Polynomen q_i absteigenden Grades. Das Polynom auf der rechten Seite der Gleichungskette hat mindestens den Grad k . Somit erfüllt der Grad n von f die Bedingung $k \leq n$, also:

Satz 3.2.2.1. *Ein Polynom $f(x) = \sum_{i=0}^n a_i x^i$ vom Grad n hat höchstens n Nullstellen.*

Lässt sich sogar

$$f(x) = a_n(x - \alpha_1)^{e_1}(x - \alpha_2)^{e_2} \dots (x - \alpha_k)^{e_k}$$

mit $e_i \in \mathbb{N}$ schreiben – in diesem Fall muss $\sum_{i=1}^k e_i = n$, der Grad von f , sein –, so sagt man: Das Polynom f **zerfällt in Linearfaktoren**. Für paarweise verschiedene α_i sind die Exponenten e_i dann die Vielfachheiten der Nullstellen α_i . Generell versteht man unter der **Vielfachheit** einer Nullstelle α von f den höchsten Exponenten $e \in \mathbb{N}$, so dass $(x - \alpha)^e$ das Polynom f teilt.

Eine wichtige Konsequenz des Bisherigen ist die Eindeutigkeit der Koeffizienten von Polynomfunktionen:

Satz 3.2.2.2. *Stimmen zwei Polynome $f_1(x) = \sum_{i=0}^n a_i x^i$ und $f_2(x) = \sum_{i=0}^n b_i x^i$ vom Grad $\leq n$ als Funktionen an $n+1$ Stellen überein, gilt also $f_1(x) = f_2(x)$ für $x = x_0, x_1, \dots, x_n$ (paarweise verschieden), so stimmen auch die Koeffizienten $a_i = b_i$ für $i = 0, 1, \dots, n$ überein. Insbesondere ist der (über die Koeffizienten definierte) Grad einer Polynomfunktion f eindeutig bestimmt ist.*

Beweis. Denn andernfalls wäre $f_1 - f_2$ ein Polynom $\neq 0$ vom Grad $\leq n$ mit mehr als n , nämlich mindestens mit den $n+1$ Nullstellen x_0, x_1, \dots, x_n , was nach Satz 3.2.2.1 unmöglich ist. \square

Übungsaufgabe 198. (P) Geben Sie drei verschiedene reelle Polynome mit den Nullstellen $0, 1, 2, 3, 4$ an und überlegen Sie sich, wie man alle reellen Polynome mit diesen Nullstellen beschreiben kann. Hinweis: Es gibt ein, in einem naheliegenden Sinn minimales, Polynom mit dieser Eigenschaft. Alle anderen sind Vielfache davon.

3.2.3 Fundamentalsatz der Algebra

Inhalt in Kurzfassung: Einer der großen Sätze der Mathematik ist der sogenannte Fundamentalsatz der Algebra. Er besagt, dass jedes komplexe Polynom in Linearfaktoren zerfällt, insbesondere also (bei Zählung der Vielfachheiten) so viele Nullstellen hat, wie sein Grad beträgt. Für reelle Polynome gilt eine ähnliche Zerlegung in höchstens quadratische Faktoren. Die Zerlegung eines Polynoms in irreduzible Faktoren ist im Wesentlichen eindeutig.

Die bisherigen Überlegungen und Ergebnisse gelten genauso, wenn man statt mit reellen mit komplexen (oder auch mit rationalen) Polynomen arbeitet. Davon werden wir nun Gebrauch machen. Wie weithin üblich, werden wir komplexe Variablen oft mit z bezeichnen.

Die Bedeutung der komplexen Zahlen beruht zu einem guten Teil auf dem sogenannten **Fundamentalsatz der Algebra (erste Version)**:

Satz 3.2.3.1. Jedes komplexe Polynom $f \in \mathbb{C}[x]$ vom Grad $n \geq 1$ hat mindestens eine komplexe Nullstelle.

Beweis. Idee: Man findet eine Stelle α , wo der Betrag des Funktionswerts $|f(\alpha)|$ minimal ist. Wäre $f(\alpha) \neq 0$, so zeigt man, dass durch geringfügige Verschiebung von α zu einem geeigneten α' die Ungleichung $|f(\alpha')| < |f(\alpha)|$ erzwungen werden könnte, im Widerspruch zur Minimalität von $|f|$ in α . Doch nun zur Umsetzung dieser Idee, die Anleihen bei der Analysis von reellwertigen Funktionen in der Ebene aus Mathematik 2 nimmt, allerdings in völliger Analogie zu aus dem Eindimensionalen bereits Bekanntem.

Ausführung: Sei $f : z \mapsto \sum_{i=0}^n a_i z^i$ mit $a_i \in \mathbb{C}$, $a_n \neq 0$. Für komplexe Zahlen z mit großem Betrag $|z|$, dominiert in

$$f(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_0 = a_n z^n \left(1 + \frac{a_{n-1}}{a_n z} + \frac{a_{n-2}}{a_n z^2} + \dots + \frac{a_0}{a_n z^{n-1}} \right)$$

betragsmäßig der höchste Term $a_n z^n$ gegenüber den anderen $a_i z^i$. Weil $|a_n z^n| = |a_n| \cdot |z|^n$ sehr groß wird, gilt das auch für $|f(z)|$. Kleine $|f(z)|$ (etwa in der Größenordnung von $|a_0| = |f(0)|$) können daher nur im Inneren einer beschränkten und abgeschlossenen Kreisscheibe K in der komplexen Zahlenebene rund um 0 auftreten. Die Funktion $z \mapsto |z|$ ist stetig, somit auch $z \mapsto |f(z)|$. Auf \mathbb{C} gilt ganz analog wie auf \mathbb{R} ein Satz vom Maximum: Jede stetige reellwertige Funktion nimmt auf jeder nichtleeren, beschränkten und abgeschlossenen (kompakten) Teilmenge von \mathbb{C} sowohl Maximum als auch Minimum an. Sei also $\alpha \in K$ eine Stelle des Minimums von $|f|$ auf K . Wir beweisen nun indirekt, dass $f(\alpha) = 0$ gilt.

Wir entwickeln f um α , also $f(z) = \sum_{j=0}^n b_j(z - \alpha)^j$ mit gewissen $b_j \in \mathbb{C}$, und wählen $k \in \{1, 2, \dots, n\}$ so, dass (man beachte $b_0 = f(\alpha)$)

$$f(z) = f(\alpha) + (z - \alpha)^k \left(b_k + (z - \alpha) \sum_{j=k+1}^n b_j(z - \alpha)^{j-k-1} \right)$$

mit $b_k \neq 0$ gilt. Wäre $b_0 = f(\alpha) \neq 0$, so ließe sich mit einem geeigneten α' sehr nahe bei α der Widerspruch $|f(\alpha')| < |f(\alpha)|$ dazu ableiten, dass α Minimalstelle für $|f|$ ist. Das sieht man so ein:

In der obigen Entwicklung von $f(z)$ um α wird für $z \rightarrow \alpha$ der zweite Summand in der großen Klammer beliebig klein. Also dominiert in dieser Klammer $b_k \neq 0$. Somit muss nur noch erzwungen werden, dass die komplexe Zahl $f(z) - f(\alpha) \approx (z - \alpha)^k b_k$ (aufgefasst als Vektor von 0 weg) annähernd in entgegengesetzte Richtung weist wie $b_0 = f(\alpha)$, also $(z - \alpha)^k \approx -\varepsilon \frac{f(\alpha)}{b_k}$ mit einem kleinen $\varepsilon > 0$. Weil die komplexe Zahl $-\varepsilon \frac{f(\alpha)}{b_k}$ eine k -te Wurzel $w \in \mathbb{C}$ hat, lässt sich auch ein entsprechendes $\alpha' = \alpha + w$ nahe bei α mit $|f(\alpha')| < |f(\alpha)|$ finden. \square

Anders formuliert besagt der Fundamentalsatz, dass sich aus jedem komplexen Polynom f vom Grad $n \geq 1$ ein Linearfaktor $x - \alpha$ mit $\alpha = \alpha_1 \in \mathbb{C}$ abspalten lässt, d.h. $f(x) = q_1(x)(x - \alpha_1)$. Ist q_1 nicht schon eine Konstante, so garantiert wieder der Fundamentalsatz einen Linearfaktor $x - \alpha_2$ auch von q_1 etc. Wir schließen also auf die **zweite Version des Fundamentalsatzes**:

Satz 3.2.3.2. *Jedes komplexe Polynom $f(x) = \sum_{i=0}^n a_i x^i$ vom Grad $n \geq 1$ zerfällt in ein Produkt von Linearfaktoren:*

$$f(x) = a_n(x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_n).$$

Die $\alpha_1, \alpha_2, \dots, \alpha_n$ sind die Nullstellen von f .

Wir können auch noch eine interessante Folgerung für reelle Polynome ziehen. Sie beruht auf der früheren Beobachtung (gegen Ende von 1.3.11), dass die Bildung der konjugiert komplexen Zahl \bar{z} von $z \in \mathbb{C}$ mit den Grundrechnungsarten verträglich ist: $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$, $\overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2$, $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$, $\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}$ ($z_2 \neq 0$). Beachtet man auch noch $\bar{\bar{r}} = r$ für $r \in \mathbb{R}$, so folgt daraus für Polynome $f(x) = \sum_{i=0}^n a_i x^i$ mit reellen Koeffizienten $a_i \in \mathbb{R}$ und für $z \in \mathbb{C}$:

$$f(\bar{z}) = \sum_{i=0}^n a_i \bar{z}^i = \overline{\sum_{i=0}^n a_i z^i} = \overline{f(z)}.$$

Ist nun α eine komplexe Nullstelle von f , so folgt daraus, dass auch $\bar{\alpha}$ eine Nullstelle von f ist:

$$f(\bar{\alpha}) = \overline{f(\alpha)} = \bar{0} = 0.$$

Mit jeder komplexen Nullstelle $\alpha = a + ib$ ($a, b \in \mathbb{R}$) eines reellen Polynoms f ist also auch $\bar{\alpha} = a - ib$ eine Nullstelle von f . Daher ist f teilbar durch das Produkt

$$p(x) = (x - \alpha)(x - \bar{\alpha}) = (x - a - ib)(x - a + ib) = (x - a)^2 + b^2.$$

Das ist ein quadratisches Polynom, das reell **irreduzibel** ist, d.h. das nicht als Produkt reeller Polynome kleineren Grades dargestellt werden kann. Denn diese Polynome müssten Grad 1 und deshalb eine reelle Nullstelle haben; p hat aber keine. Also: Jedes reelle Polynom positiven Grades ohne reelle Nullstelle hat wenigstens einen quadratischen, reell irreduziblen Faktor. Aus nichtkonstanten reellen Polynomen lassen sich also immer lineare oder quadratische reelle Faktoren abspalten. Iteration solcher Abspaltungen zeigt:

Satz 3.2.3.3. *Jedes nicht konstante reelle Polynom lässt sich als ein Produkt irreduzibler reeller Polynome der Grade 1 und/oder 2 darstellen.*

Wir können diese Einsichten noch weiter treiben, um ein Analogon zur eindeutigen Primfaktorzerlegung in \mathbb{Z} zu erhalten. Die Rolle der Primzahlen übernehmen nun die irreduziblen Polynome. So wie in \mathbb{Z} z.B. 2 und -2 bezüglich Teilbarkeit dieselbe Rolle spielen, tun das bei den Polynomen zum Beispiel $x - 3$ und $2x - 6$. Wenn wir uns entscheiden müssen, ziehen wir $x - 3$ vor, weil es **normiert** ist in dem Sinn, dass der führende Koeffizient $= 1$ ist, was durch Division stets erreicht werden kann. Somit lässt sich jedes Polynom f als Produkt einer Konstanten (dem führenden Koeffizienten von f) und sonst lauter irreduziblen normierten Faktoren schreiben. Welche irreduzible normierte Faktoren auftreten, ist durch die Nullstellen von f eindeutig bestimmt. Auch die Vielfachheit, mit der sie auftreten ist eindeutig. Denn wir können durch jeden irreduziblen Faktor einmal dividieren und mit dem verbleibenden Polynom weiterarbeiten: Seine Nullstellen bestimmen jene irreduziblen Faktoren, die in f mehr als einmal auftreten etc. Somit gilt:

Satz 3.2.3.4. *Jedes Polynom $f \in \mathbb{R}[x]$ (analog: $f \in \mathbb{C}[x]$) vom Grad $n \geq 1$ lässt sich als ein Produkt $f = a_n f_1 f_2 \dots f_k$ des höchsten Koeffizienten a_n von f und normierter irreduzibler Polynome $f_k \in \mathbb{R}[x]$ (bzw. $f_k \in \mathbb{C}[x]$) zerlegen. Diese Darstellung ist eindeutig in folgendem Sinn: Ist $f = b_n g_1 g_2 \dots g_l$ irgendeine weitere Produktdarstellung von f mit einer Konstanten b_n und normierten irreduziblen Polynomen $g_i \in \mathbb{R}[x]$ (bzw. $g_i \in \mathbb{C}[x]$), so gilt $a_n = b_n$, $k = l$, und es gibt eine Permutation π der Indexmenge $\{1, 2, \dots, k\}$ derart, dass $g_i = f_{\pi(i)}$ für $i = 1, 2, \dots, k$ gilt.*

Satz 3.2.3.4 gilt analog für rationale Polynome (also mit $\mathbb{Q}[x]$ statt $\mathbb{R}[x]$ bzw. $\mathbb{C}[x]$) und sogar für Polynome über beliebigen Körpern. Der Beweis müsste aber anders erfolgen, was hier zu weit geführt hätte. In jedem Fall gibt die eindeutige Faktorisierung Anlass zu ähnlichen Begriffen wie in der Teilbarkeitslehre ganzer Zahlen, wie etwa größter gemeinsamer Teiler, kleinstes gemeinsames Vielfaches etc.

Übungsaufgabe 199. (T) Zerlegen Sie das Polynom f in irreduzible komplexe und irreduzible reelle Faktoren.

1. $f(x) = x^3 - 1$
2. $f(x) = x^5 - 1$
3. $f(x) = x^4 + x^3 + x^2 + x + 1$
4. $f(x) = x^n - 1, \quad n \in \mathbb{N}$

3.2.4 Nullstellenbestimmung

Inhalt in Kurzfassung: Der klassische Gegenstand der Algebra ist das Lösung von algebraischen Gleichungen, also die Bestimmung von Nullstellen von Polynomen. Der Fundamentalsatz der Algebra aus 3.2.3 garantiert die Existenz von komplexen Lösungen. Nun geht es um deren explizite Angabe. Allgemeine Lösungsformeln gibt es für Polynome bis zum Grad 4, wir begnügen uns aber mit dem quadratischen Fall. Darüber hinaus interessant ist ein Satz, der im Fall rationaler Koeffizienten die Suche nach rationalen Lösungen auf eine endliche Menge einschränkt.

Der Fundamentalsatz der Algebra garantiert, dass Polynome f stets komplexe Nullstellen haben. Damit ist aber noch nicht erklärt, wie man diese Nullstellen auch findet. Besonders praktisch wären natürlich explizite Formeln für die Nullstelle(n) von f , wo man nur die Koeffizienten von f einzusetzen braucht.

Fast trivial ist es, eine Lösungsformel für lineare Polynome $f(x) = kx + d$ zu bekommen. Ist $k \neq 0$, f also vom Grad 1, so formt man $f(x) = kx + d = 0$ um zu $kx = -d$ und zur expliziten Lösungsformel $x = -\frac{d}{k}$.

Quadratische Gleichungen $ax^2 + bx + c = 0$ mit $a \neq 0$ dividiert man am besten zunächst durch a und erhält die Normalform $f(x) := x^2 + px + q = 0$. Bekanntlich führt Ergänzung auf ein sogenanntes vollständiges Quadrat zum Ziel: $x^2 + px + q = (x + \frac{p}{2})^2 - \frac{p^2}{4} + q$, was zur Gleichung $(x + \frac{p}{2})^2 = \frac{p^2}{4} - q$ führt, welche offenbar die beiden (auch für reelles f möglicherweise komplexen) Lösungen

$$x = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$$

hat. Hier fließt ein, dass jede komplexe Zahl (hier die Zahl $\frac{p^2}{4} - q$) komplexe Quadratwurzeln hat. (Dividiert man anfangs nicht durch a , kann man ähnlich vorgehen und erhält eine etwas kompliziertere aber gleichwertige Formel, die manchmal die *große Lösungsformel* genannt wird. Sie hat aber keinen zusätzlichen praktischen Wert, sondern ist nur schwerer zu merken.)

Übungsaufgabe 200. (T) Lösen Sie die folgenden quadratischen Gleichungen und Ungleichungen:

1. $2x^2 + 6x - 8 = 0$ und $2x^2 + 6x - 8 < 0$.
2. $x^2 - 2|x| - 3 = 0$ und $x^2 - 2|x| - 3 \geq 0$.

Fertigen Sie jeweils eine aussagekräftige Skizze an, aus der man Nullstellen und Vorzeichen der entsprechenden Terme ablesen kann.

Auch für kubische und sogar für Polynome vierten Grades gibt es analoge explizite Lösungsformeln, die von italienischen Mathematikern des 16. Jahrhunderts gefunden wurden. Diese Formeln sind jedoch schon einigermaßen kompliziert und werden deshalb in der Praxis nur selten eingesetzt. Versuche, auch Formeln für Gleichungen noch höheren Grades herzuleiten, scheiterten. Um 1830 stellte sich heraus, dass solche Formeln (natürlich hat man zu spezifizieren, welche Art von Formeln genau man damit meint) nicht existieren.³ Man hat sich deshalb auf andere Methoden zu besinnen. Diese werden wir im Zuge der Differentialrechnung vor allem im Näherungsverfahren von Newton kennenlernen.

Im Fall rationaler bzw. ganzzahliger Koeffizienten a_i von

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0$$

ist die folgende Beobachtung oft von Nutzen. Sind die $a_i = \frac{p_i}{q_i} \in \mathbb{Q}$ Brüche mit $p_i, q_i \in \mathbb{Z}$, so lässt sich die Gleichung $f(x) = 0$ durch Multiplikation mit einem gemeinsamen Vielfachen der q_i auf den Fall ganzer Koeffizienten zurückführen. Ist nun $\alpha = \frac{p}{q}$ mit $p, q \in \mathbb{Z}$ eine rationale Nullstelle von f in gekürzter Darstellung, d.h. mit teilerfremden p, q , dann folgt aus

$$0 = f(\alpha) = f\left(\frac{p}{q}\right) = a_n \left(\frac{p}{q}\right)^n + a_{n-1} \left(\frac{p}{q}\right)^{n-1} + \dots + a_2 \left(\frac{p}{q}\right)^2 + a_1 \frac{p}{q} + a_0$$

nach Multiplikation mit q^n die Beziehung

$$0 = a_n p^n + a_{n-1} p^{n-1} q + \dots + a_2 p^2 q^{n-2} + a_1 p q^{n-1} + a_0 q^n.$$

Alle $a_i p^i q^{n-i}$ mit $i \geq 1$ sind durch p teilbar, folglich muss p auch den verbleibenden letzten Summanden $a_0 q^n$ teilen. Weil q und (wegen der Eindeutigkeit der Primfaktorzerlegung) deshalb auch q^n zu p teilerfremd ist, muss p ein Teiler von a_0 sein. Ganz analog erweist sich q als ein Teiler von a_n . Damit haben wir bewiesen:

Satz 3.2.4.1. *Sei $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0$ mit ganzzahligen Koeffizienten $a_i \in \mathbb{Z}$ und α eine Nullstelle von f mit der Darstellung $\frac{p}{q}$ eine gekürzter Bruch ganzer Zahlen p, q . Dann ist p ein Teiler von a_0 und q ein Teiler von a_n .*

Man beachte, dass es nur endlich viele Teiler p von a_0 und endlich viele Teiler q von a_n gibt. Also lassen sich durch Einsetzen nach endlich vielen Schritten sämtliche rationalen Nullstellen von f finden. Sind die a_i teilerfremd (was ja immer durch Kürzen garantiert werden kann), so sind weniger Möglichkeiten zu überprüfen als andernfalls.

Beispiel: Für $f(x) := 3x^3 + 10x^2 - 23x + 10$ kommen nur rationale Nullstellen in Frage, wo bei gekürzter Darstellung $\alpha = \frac{p}{q}$ der Zähler p ein Teiler von $a_0 = 10$ und der Nenner q ein Teiler von $a_n = a_3 = 3$ ist, also $p \in \{\pm 1, \pm 2, \pm 5, \pm 10\}$ und $q \in \{\pm 1, \pm 3\}$. Unter den sich daraus ergebenden 16 Möglichkeiten identifiziert man (relativ) rasch die drei rationalen Nullstellen $1, \frac{2}{3}$ und -5 von f .

³Die beiden jung verstorbenen Mathematiker Niels Henrik Abel (1802-1829) und Évariste Galois (1811-1832) sind in diesem Zusammenhang an erster Stelle zu nennen.

Übungsaufgabe 201. (T) Bestimmen Sie alle komplexen Nullstellen von f und faktorisieren Sie f einmal in reelle und einmal in komplexe irreduzible Faktoren:

$$f(x) = x^6 + x^5 - 2x^4 - x^2 - x + 2$$

Übungsaufgabe 202. (T)

1. Sei $p(x) = x^3 + x^2 + x + 1$ und $q(x) = x^6 + x^4 + x^2 + 1$. Zerlegen Sie p und q in möglichst viele Faktoren mit reellen Koeffizienten.
2. Bestimmen Sie alle Lösungen der Gleichungen $p(x) = 0$ und $q(x) = 0$ für $x \in \mathbb{C}$.

3.2.5 Approximation und Interpolation durch Polynome

Inhalt in Kurzfassung: Polynome können auf einem abgeschlossenen Intervall $[a, b]$ eine beliebig vorgegebene stetige Funktion beliebig gut approximieren (Satz von Stone-Weierstraß, den wir allerdings nicht beweisen). Elementar ist hingegen einzusehen, dass es stets sogenannte Interpolationspolynome im folgenden Sinn gibt: Sind $n + 1$ verschiedene Stellen x_i und zugeordnete Werte y_i vorgegeben, so gibt es ein eindeutiges Polynom p vom Grad $\leq n$ mit $p(x_i) = y_i$ für alle i . Dieses Polynom kann mit der Lagrange'schen Interpolationsformel oder, algorithmisch effektiver, mit dem Newtonschen Interpolationsverfahren ermittelt werden.

Wegen der Übersichtlichkeit der Theorie der Polynome ist es oft von Interesse, Funktionen f allgemeinerer Natur durch Polynome p zu ersetzen. Zu unterscheiden sind **Approximation** und **Interpolation**. Bei der Approximation sucht man nach einem p mit kleiner Differenz $|p(x) - f(x)|$. Das wichtigste Ergebnis in diese Richtung ist der **Approximationssatz von Weierstraß**, der hier ohne Beweis erwähnt sei:

Satz 3.2.5.1. Zu jeder stetigen reellen Funktion $f : [a, b] \rightarrow \mathbb{R}$ und jedem $\varepsilon > 0$ gibt es ein Polynom $p : [a, b] \rightarrow \mathbb{R}$ mit $|p(x) - f(x)| < \varepsilon$ für alle $x \in [a, b]$.

Dieser Satz besagt insbesondere, dass es zu jedem solchen f eine Folge von Polynomen p_n gibt, die auf $[a, b]$ gleichmäßig gegen f konvergiert. Man braucht nur zu jedem $n = 1, 2, \dots$ ein Polynom p_n mit $|p_n(x) - f(x)| < \frac{1}{n}$ für alle $x \in [a, b]$ zu nehmen, was laut Satz von Stone-Weierstraß ja möglich ist. Allerdings ist damit zu rechnen, dass die Grade der p_n gegen ∞ streben. Das ist sogar immer der Fall, wenn f nicht schon selbst ein Polynom ist.

Im Gegensatz zur Approximation geht es beim **Interpolationsproblem** darum, vorgegebene Werte $p(x_i)$ an gewissen Stellen x_i genau zu treffen. Allerdings ist das nur an endlich vielen Stellen möglich. Die Situation lässt sich sehr genau durch den folgenden Satz (**Lagrangesche Interpolationsformel**) beschreiben.

Satz 3.2.5.2. *Sind die $n + 1$ verschiedenen Stellen $x_0, x_1, \dots, x_n \in \mathbb{R}$ (analog für \mathbb{C}) sowie Werte $y_0, y_1, \dots, y_n \in \mathbb{R}$ vorgegeben, dann gibt es genau ein reelles Polynom p vom Grad $\leq n$, für das $p(x_i) = y_i$ für $i = 0, 1, \dots, n$ gilt.*

Beweis. Leicht zu beweisen ist die Eindeutigkeitsaussage. Denn angenommen, es gibt zwei Polynome p_1 und p_2 mit den geforderten Eigenschaften, so ist die Differenz $p_1 - p_2$ ein Polynom vom Grad $\leq n$ mit $n + 1$ verschiedenen Nullstellen x_0, x_1, \dots, x_n . Das ist aber nur dann möglich, wenn $p_1 - p_2 = 0$, also $p_1 = p_2$.

Nun zur Existenz eines p wie in der Behauptung. Das Polynom $q(x) := \prod_{i=0}^n (x - x_i)$ hat Grad $n + 1$ und jeden der Linearfaktoren $x - x_i$ als Teiler. Folglich sind die

$$q_i(x) := \prod_{j: 0 \leq j \leq n, j \neq i} (x - x_j)$$

$i = 0, 1, \dots, n$, selbst Polynome vom Grad n . Die Nullstellen von q_i sind genau die x_j mit $j \neq i$, während $q_i(x_i) \neq 0$. Somit lässt sich das Polynom

$$p(x) := \sum_{i=0}^n y_i \frac{q_i(x)}{q_i(x_i)}$$

vom Grad $\leq n$ bilden, welches offenbar $p(x_i) = y_i$ für $i = 0, 1, \dots, n$ erfüllt. \square

Der Formel für das **Interpolationspolynom** p aus dem obigen Beweis sieht man leicht an, dass p die gewünschten Eigenschaften hat. Die explizite Berechnung der Koeffizienten von p ist aber verhältnismäßig aufwendig. Für die Praxis besser geeignet ist das **Newtonsche Interpolationsverfahren**, das wegen der Eindeutigkeit des Interpolationspolynoms zum selben Ergebnis führt. Man generiert dabei nach und nach Polynome $p_0, p_1, p_2, \dots, p_n$, von denen jedes p_k einen Grad $\leq k$ hat und Interpolation $p_k(x_i) = y_i$ für $i = 0, 1, \dots, k$ leistet. Man beginnt mit dem konstanten Polynom $p_0(x) := y_0$ und konstruiert rekursiv $p_{k+1}(x) := p_k(x) + \lambda_k r_k(x)$ mit $r_k(x) := (x - x_0)(x - x_1) \dots (x - x_k)$, wobei λ_k gerade so gewählt wird, dass $p_{k+1}(x_{k+1}) = y_{k+1}$ gilt. Weil $r_k(x_{k+1})$ sicher nicht 0 ist (die x_i sind ja paarweise verschieden), gibt es ein eindeutiges λ_k mit der gewünschten Eigenschaft. Für $i = 0, 1, \dots, k$ ist hingegen $r_k(x_i) = 0$, also $p_{k+1}(x_i) = p_k(x_i) + \lambda_k r_k(x_i) = p_k(x_i) = y_i$. Somit ist $p := p_n$ tatsächlich das gesuchte Interpolationspolynom.

Übungsaufgabe 203. (T) Gegeben ist die Interpolationsaufgabe $f(0) = 1$, $f(\frac{1}{2}) = 2$ und $f(\frac{3}{4}) = 4$.

1. Finden Sie irgendein Interpolationspolynom für so ein f .
2. Finden Sie ein anderes Polynom mit denselben Werten an den angegebenen Stellen.
3. Beschreiben Sie die Menge aller Interpolationspolynome und setzen Sie Ihre Erkenntnis in Beziehung zur Eindeutigkeitsaussage von Satz 3.2.5.2.

3.2.6 Partialbruchzerlegung

Inhalt in Kurzfassung: Für gebrochen rationale Funktionen gibt es neben der Darstellung als gekürzter Bruch noch eine zweite Normalform, nämlich als Summe eines Polynoms und Brüchen spezieller Gestalt, wo insbesondere der Zählergrad stets kleiner ist als der Nennergrad. Die resultierende Darstellung nennt man Partialbruchzerlegung. Sie ist u.a. für die Integration gebrochen rationaler Funktionen von praktischem Nutzen. Zuerst wollen wir das Wichtigste für den komplexen Fall besprechen, dann für den reellen.

Die Partialbruchzerlegung ist die Darstellung einer beliebigen gebrochen rationalen Funktion als Summe von Brüchen etwas einfacherer Bauart. Insbesondere ist diese Darstellung nützlich bei Integrationsaufgaben, siehe 5.2.3. Auch für Kurvendiskussionen ist die Partialbruchzerlegung nützlich, weil sie auf den ersten Blick Polstellen erkennen lässt. Es folgt eine Beschreibung der praktischen Durchführung.

Sei $r(x) = \frac{p(x)}{q(x)}$ eine gebrochen rationale Funktion. Ist der Grad von p nicht von Beginn an kleiner als der von q , so lässt sich mittels Polynomdivision eine Darstellung der Gestalt $r(x) = p_0(x) + \frac{p_1(x)}{q(x)}$ mit Polynomen p_0 und p_1 erzielen, wobei der Grad von p_1 kleiner ist als der von q . Weiters setzen wir p_1 und q als teilerfremd voraus. (Andernfalls kürzen wir durch diesen gemeinsamen Teiler.)

Sei daher ab nun die Darstellung $r(x) = \frac{p(x)}{q(x)}$ gekürzt und $\text{grad}(p) < \text{grad}(q)$. Außerdem dürfen wir der Einfachheit halber q als normiert annehmen, d.h. der Koeffizient bei der höchsten x -Potenz sei 1. Nach dem Fundamentalsatz lässt sich $q(x) = \prod_{i=1}^k (x - \alpha_i)^{e_i}$ in komplexe Linearfaktoren zerlegen mit paarweise verschiedenen Nullstellen α_i und mit Vielfachheiten $e_i \in \mathbb{N}$, $e_i \geq 1$. Man kann zeigen, dass sich $r(x)$ als Summe von Funktionen der Gestalt $\frac{c_{i,j}}{(x - \alpha_i)^j}$ mit geeigneten Zahlen $c_{i,j} \in \mathbb{C}$ ($i = 1, \dots, k; j = 1, \dots, e_i$) darstellen lässt.

Ist r nicht komplex, sondern reell, und ist man an einer rein reellen Darstellung interessiert, so zerlegt man q in reell irreduzible Faktoren, die aber nun auch quadratisch sein können:

$$q(x) = \prod_{i=1}^k (x - \alpha_i)^{e_i} \prod_{j=1}^l (x^2 + \beta_j x + \gamma_j)^{f_j}$$

mit reell irreduziblen quadratischen Polynomen $x^2 + \beta_j x + \gamma_j$ ($j = 1, \dots, l; \alpha_j, \beta_j, \gamma_j \in \mathbb{R}; e_i, f_j \in \mathbb{N}; e_i, f_j \geq 1$). Die Partialbruchzerlegung von r setzt sich nun aus zwei Arten von Summanden zusammen: Jenen der Gestalt $\frac{c_{i,j}}{(x - \alpha_i)^j}$ mit geeigneten Zahlen $c_{i,j} \in \mathbb{R}$ ($i = 1, \dots, k; j = 1, \dots, e_i$), wie sie analog schon im komplexen Fall aufgetreten sind, und aus jenen der Gestalt $\frac{a_{i,j}x + b_{i,j}}{(x^2 + \beta_i x + \gamma_i)^j}$ mit geeigneten Zahlen $a_{i,j}, b_{i,j} \in \mathbb{R}$ ($i = 1, \dots, l; j = 1, \dots, f_i$).

Der Beweis dieses **Satzes von der Partialbruchzerlegung** ist etwas technisch, und führte etwas zu weit, ohne interessante Einsichten zu bringen. Wir begnügen uns daher mit der Beschreibung, wie man für gegebenes f die Partialbruchzerlegung erhält.

Gegeben sei eine gebrochene rationale Funktion $f(x) = \frac{p(x)}{q(x)}$. Dann sind folgende Schritte durchzuführen, von denen die ersten drei auch in anderer Reihenfolge möglich sind:

1. Falls der Grad von p nicht kleiner ist als der von q , liefert Polynomdivision eine Darstellung von f als Summe eines Polynoms und eines Bruches mit Zählergrad $<$ Nennergrad. Mit diesem Bruch kann nun weitergearbeitet werden.
2. Faktorisierung von q in (reell) irreduzible Faktoren. In der Praxis kann das schwierig sein. In Übungsaufgaben ist die Faktorisierung oft schon vorgegeben oder relativ leicht zu ermitteln.
3. Überprüfung, ob einer der irreduziblen Faktoren des Nenners q auch ein Teiler des Zählers p ist. In diesem Fall ist durch solche gemeinsame Faktoren zu kürzen, bis resultierende Zähler und Nenner teilerfremd sind. Tatsächlich steht mit dem euklidischen Algorithmus eine effektive Methode zur Verfügung, um den größten gemeinsamen Teiler von zwei Polynomen (wie übrigens auch von zwei ganzen Zahlen) zu ermitteln.⁴ In Übungsaufgaben ist die gegebene Darstellung meist bereits eine gekürzte. (Das gilt übrigens auch in der Praxis, weil bei komplizierten Zahlenwerten gemeinsame Teiler unwahrscheinlich sind.)
4. Partialbruchzerlegung mit unbestimmten Werten, oft mit A, B, C, \dots bezeichnet, die die Rolle der α_i, β_i und γ_i spielen, ansetzen. Die gesamte verbleibende Aufgabe besteht in der Ermittlung von A, B, C, \dots
5. Die entstandene Gleichung für die beiden gebrochen rationalen Funktionen (f in der ursprünglich gegebenen Form links und die Partialbruchzerlegung mit den Unbekannten A, B, C, \dots rechts) durch Multiplikation mit dem kleinsten gemeinsamen Vielfachen der Nennerpolynome multiplizieren. Dadurch entsteht eine Gleichung zwischen zwei Polynomen, rechts mit den zu bestimmenden noch unbekanntem Koeffizienten A, B, C, \dots
6. Der Eindeutigkeitssatz für Polynome besagt, dass die erhaltene Gleichung nur dann gilt, wenn entsprechende Koeffizienten links und rechts übereinstimmen. Das liefert ein System mehrerer linearer Gleichungen in ebenso vielen Variablen A, B, C, \dots , welches laut Satz von der Partialbruchzerlegung eine Lösung hat. (Genauere Analyse zeigt, dass die Lösung für A, B, C, \dots sogar eindeutig ist.)

⁴Eine kurze Beschreibung des euklidischen Algorithmus zur Berechnung des größten gemeinsamen Teilers: Sind a und b zwei ganze Zahlen oder auch zwei Polynome, dann führt man Division mit Rest durch und erhält $b = aq_0 + r_0$ mit einem Rest r_0 , der dem Betrag bzw. dem Grad nach kleiner ist als a . Das gleiche macht man nun mit a und r_0 , also $a = r_0q_1 + r_2$, sodann $r_0 = r_1q_2 + r_3$ etc. bis $r_n = r_{n+1}q_{n+2}$ (also bis der Rest $r_{n+2} = 0$ wegfällt). Dann ist, wie man leicht nachprüft, r_{n+1} der gesuchte größte gemeinsame Teiler von a und b . Denn jeder gemeinsame Teiler von a und b ist auch ein Teiler von r_0 , folglich von r_1 etc. und schließlich von r_{n+1} . Umgekehrt ist r_{n+1} ein Teiler von r_n , folglich von r_{n-1} etc., folglich von r_0 , folglich von a und schließlich auch von b .

Ein einfaches Beispiel zur Illustration, vor allem zur Lösung des auftretenden Gleichungssystems, ist

$$f(x) = \frac{x^2 + 2x - 3}{(x-2)(x^2+1)} = \frac{A}{x-2} + \frac{Bx+C}{x^2+1}.$$

Die rechte Seite entspricht dem allgemeinen Ansatz. Wäre im Nenner links zum Beispiel $(x-2)^2$ statt $(x-2)$ aufgetreten, müsste im Ansatz rechts auch ein Summand der Gestalt $\frac{D}{(x-2)^2}$ auftreten. Ähnlich zöge $(x^2+1)^2$ statt (x^2+1) links einen zusätzlichen Summanden $\frac{E x+F}{(x^2+1)^2}$ auf der rechten Seite nach sich. In der einfachen Situation hier multiplizieren wir mit dem gemeinsamen Nenner $(x-2)(x^2+1)$ und erhalten die Gleichung

$$x^2 + 2x - 3 = A(x^2 + 1) + (Bx + C)(x - 2) = (A + B)x^2 + (-2B + C)x + (A - 2C).$$

Weil die quadratischen Polynome links und rechts nur dann übereinstimmen können, wenn das auch alle Koeffizienten tun, folgen die drei Gleichungen $1 = A + B$, $2 = -2B + C$ und $-3 = A - 2C$. Sie haben die eindeutige gemeinsame Lösung $A = 1$, $B = 0$, $C = 2$, die man zum Beispiel durch Elimination erhält. (Die systematische Lösung linearer Gleichungssysteme ist Gegenstand der Linearen Algebra, siehe Mathematik 2. Für die einfachen Beispiele hier sollte Schulwissen ausreichen.)

Alternative Methode (die sogenannte Polmethode) zur Berechnung der Koeffizienten A, B, C, \dots , illustriert an diesem Beispiel: Bei linearen Nennerpolynomen (im Beispiel $x-2$) kann man den zugehörigen Koeffizienten (im Beispiel A) auch sehr schnell ermitteln, indem man in der Polynomgleichung für x die Nullstelle des Polynoms (im Beispiel $x=2$) einsetzt. Dann bleibt nämlich nur mehr $2^2 + 2 \cdot 2 - 3 = A(2^2 + 1)$, also $5 = A \cdot 5$ stehen, woraus unmittelbar $A = 1$ abgelesen werden kann. Für die irreduziblen quadratischen Polynome (im Beispiel x^2+1) müsste man dann aber komplexe Nullstellen (im Beispiel $x=i$) einsetzen. Im Beispiel erhält man $i^2 + 2i - 3 = (Bi + C)(i - 2)$, also $-4 + 2i = (-B - 2C) + (-2B + C)i$, was nach Vergleich von Real- und Imaginärteil zu den beiden Gleichungen $-4 = -B - 2C$ und $2 = -2B + C$ führt. Die erste zeigt $B = 4 - 2C$, was eingesetzt in die zweite $2 = -2(4 - 2C) + C = 5C - 8$, folglich $C = 2$ und $B = 4 - 2C = 4 - 2 \cdot 2 = 0$ führt, wie oben behauptet. Kommen im Nenner höhere Potenzen irreduzibler Polynome vor, muss man in mehreren Schritten vorgehen. Besser als das Auswendiglernen vollständiger Rezepte ist das eigenständige und verständige Durchrechnen einiger Übungsaufgaben.

Übungsaufgabe 204. (T) Bestimmen Sie mit einer Methode Ihrer Wahl die Partialbruchzerlegung von

1. $\frac{2x^3 + 4x + 6}{(x-1)^2(x^2+3)}$
2. $\frac{3 + 2x + 5x^2 + 2x^3}{1 + 2x + 2x^2 + 2x^3 + x^4}$
3. $\frac{x^3 - 5x^2 + 7x + 3}{x^4 - 4x^3 + 5x^2 - 4x + 4}$

3.3 Weitere wichtige Beispiele stetiger Funktionen

Es ist hoch an der Zeit, wichtige Beispiele reeller Funktionen abseits von Polynom- und gebrochen rationalen Funktionen zu besprechen: Wurzel- und Potenzfunktionen mit rationalem Exponenten (3.3.1), Exponentialfunktion (3.3.2), Logarithmus und Potenzfunktionen mit beliebigem reellen Exponenten (3.3.3) und trigonometrische Funktionen (3.3.4).

3.3.1 Wurzeln und Potenzen mit rationalem Exponenten

Inhalt in Kurzfassung: Die Definition der Wurzel- und, allgemeiner, der Potenzfunktionen mit rationalem Exponenten ergibt sich zwangsläufig, wenn man vertraute Rechenregeln wie $(xy)^n = x^n y^n$, $x^m x^n = x^{m+n}$ und $(x^n)^m = x^{nm}$ auch für die Verallgemeinerungen fordert.

Die **Potenzfunktionen** $\text{pot}_n : x \mapsto x^n$ mit $n \in \mathbb{Z}$ sind stetig und (sofern $n \neq 0$) wenigstens für $x > 0$ streng monoton; wachsend für $n > 0$ und fallend für $n < 0$. Also gibt es auf \mathbb{R}^+ auch eine stetige Umkehrfunktion (siehe Satz 3.1.5.2). Für $n > 0$ nennt man sie die n -te **Wurzelfunktion**. Die übliche Notation ist $x \mapsto \sqrt[n]{x}$, die für ungerades n sogar auf ganz \mathbb{R} , für gerades n für $x \geq 0$ definiert ist.

Die Rechenregel $(xy)^n = x^n y^n$ für $n \in \mathbb{Z}$ lässt sich als Funktionalgleichung $\text{pot}_n(xy) = \text{pot}_n(x)\text{pot}_n(y)$ anschreiben. Diese Regel vererbt sich offenbar auch auf die Umkehrfunktion. Also liegt es nahe, auch die Wurzelfunktionen als Potenzfunktionen pot_α mit einem geeigneten α aufzufassen. Aber mit welchem α ? Die naheliegende Antwort ergibt sich aus der zunächst für ganzzahlige Exponenten n, m gültigen Rechenregel $(x^n)^m = x^{nm}$, d.h. $\text{pot}_m \circ \text{pot}_n = \text{pot}_{nm}$. Es sollte demnach $\text{pot}_{\frac{1}{n}} \circ \text{pot}_n = \text{pot}_{\frac{1}{n} \cdot n} = \text{pot}_1 = \text{Id}$ gelten, $\text{pot}_{\frac{1}{n}}$ sollte also die Umkehrfunktion der n -ten Potenz pot_n sein, also $\text{pot}_{\frac{1}{n}}(x) := \sqrt[n]{x}$. Für beliebige rationale Exponenten $\alpha = \frac{m}{n}$, $m, n \in \mathbb{Z}$, $n \neq 0$, definiert man aus ähnlichen Gründen

$$\text{pot}_\alpha(x) = x^\alpha = x^{\frac{m}{n}} := \sqrt[n]{x^m}.$$

Man überzeugt sich durch etwas langwierige aber einfache Überlegungen und Rechnungen, dass diese Definition nicht von der Darstellung von $\alpha = \frac{m}{n}$ als Bruch abhängt (z.B. erhält man für $m = 2$ und $n = 3$, also $\alpha = \frac{2}{3}$ denselben Wert x^α wie für $m = 4$ und $n = 6$).

Übungsaufgabe 205. (E) Warum genau ist das ganz allgemein, d.h. für beliebige $m, n \in \mathbb{Z}$, $n \neq 0$, so? Begründen Sie sorgfältig!

Somit sind Potenzfunktionen $\text{pot}_\alpha(x) = x^\alpha$ für alle $\alpha \in \mathbb{Q}$ und $x > 0$ erklärt. Außerdem überzeugt man sich leicht davon, dass die üblichen, oben erwähnten Rechenregeln wie $x^\alpha x^\beta = x^{\alpha+\beta}$, d.h. $\text{pot}_\alpha \text{pot}_\beta = \text{pot}_{\alpha+\beta}$ etc. auch für rationale Exponenten gelten.

Übungsaufgabe 206. (P) Begründen Sie für $x, y \in \mathbb{R}$ und $\alpha, \beta \in \mathbb{Q}$ die Rechenregeln:

$$1. (xy)^\alpha = x^\alpha y^\alpha \quad 2. (x^\alpha)^\beta = x^{\alpha\beta} \quad 3. x^{\alpha+\beta} = x^\alpha x^\beta$$

Verwenden Sie in Ihrer Argumentation nur die entsprechenden Regeln mit ganzzahligen Exponenten und die Definition von Potenzen mit rationalen Exponenten.

Was aber ist zu tun, wenn auch irrationale α zugelassen werden? Dazu ist es nützlich, nicht die Basis, sondern den Exponenten als Variable aufzufassen. So gelangt man zur Exponentialfunktion.

3.3.2 Exponentialfunktion

Inhalt in Kurzfassung: Aus den bisherigen Konstruktion ergeben sich auf sehr natürliche (wenn auch etwas theoretische) Weise die Exponentialfunktionen \exp_a zu einer beliebigen positiven Basis a . Höhepunkt ist ein zusammenfassender Existenz- und Eindeigkeitssatz. Er besagt im Wesentlichen, dass alle Konstruktionsschritte tatsächlich zu Funktionen mit den gewünschten Eigenschaften führen, noch dazu auf eindeutige Weise.

Durch den vorangegangenen Abschnitt ist a^x definiert für $a > 0$ und $x \in \mathbb{Q}$. Wir wollen die Definition auf beliebige $x \in \mathbb{R}$ ausdehnen. Die naheliegende Vorgangsweise besteht darin, irrationale x durch rationale x' zu approximieren und darauf zu hoffen, dass die bereits definierten Werte $a^{x'}$ als Approximationen für den zu definierenden Wert a^x verwendet werden können. Mit anderen Worten: Man sucht als Funktion die stetige Fortsetzung bereits definierter Werte zu einer global definierten Funktion $x \mapsto a^x$ und hofft sogar auf Eindeutigkeit. Glücklicherweise ist all dies möglich. Zur Erläuterung folgt ein kurzer theoretischer Exkurs.

Als entscheidende Eigenschaft erweist sich die gleichmäßige Stetigkeit der Funktion $x \mapsto a^x$ auf jedem Intervall $D := \mathbb{Q} \cap [a, b]$ (sogenannte lokal gleichmäßige Stetigkeit). Dieser Begriff ist eine Verschärfung der gewöhnlichen Stetigkeit dahingehend, dass es nicht nur für jedes einzelne $x \in D$ zu jedem $\varepsilon > 0$ ein entsprechendes δ im Sinne der Stetigkeitsdefinition gibt, sondern dass zu jedem $\varepsilon > 0$ ein δ existiert, das für alle $x \in D$ gleichermaßen funktioniert. Der feine Unterschied wird wieder an der Reihenfolge der logischen Quantoren sichtbar. Bei der Stetigkeit beginnt die Formel mit $\forall \varepsilon > 0 \forall x \in D \exists \delta > 0 \dots$, bei der gleichmäßigen Stetigkeit mit $\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D \dots$. Mit einigem rechnerischen Aufwand zeigt man, dass bei $x \mapsto a^x$ tatsächlich die behauptete lokal gleichmäßige Stetigkeit vorliegt. Gleichmäßig stetige wie auch „lokal gleichmäßig stetige“ Funktionen haben, wie man relativ leicht überprüft, die Eigenschaft, dass sie Cauchyfolgen auf Cauchyfolgen abbilden.

Übungsaufgabe 207. (E) Zeigen Sie: Bilden die $x_n, n \in \mathbb{N}$, eine Cauchyfolge in $D \subseteq \mathbb{R}$ und ist f gleichmäßig stetig auf D , dann bilden auch die $f(x_n), n \in \mathbb{N}$, eine Cauchyfolge.

Konvergieren also rationale α_n für $n \rightarrow \infty$ gegen x , so bilden sie eine Cauchyfolge, die in die Cauchyfolge a^{α_n} übergeht. Eine solche konvergiert in \mathbb{R} (Satz 2.1.7.2, Aussage 3). Den Grenzwert $\lim_{x' \rightarrow x, x' \in \mathbb{Q}} a^{x'}$ nennen wir a^x . Wieder überzeugt man sich leicht davon, dass dieser Wert a^x nicht von der speziellen Wahl der α_n (es gibt ja auch andere rationale Folgen, die gegen x konvergieren) abhängt und dass sich die gleichmäßige Stetigkeit von D auf die resultierende Funktion $x \mapsto a^x$ auf ganz $[a, b]$ vererbt. Weil all dies auf jedem Intervall $[a, b]$ möglich ist, ist eine Funktion $x \mapsto a^x$ auf ganz \mathbb{R} definiert. Relativ leicht prüft man nach, dass die Rechenregel $a^x a^y = a^{x+y}$ auch auf dem zu ganz \mathbb{R} erweiterten Definitionsbereich gilt, ebenso $(a^x)^y = a^{xy}$ und $a^x b^x = (ab)^x$.

Für die Zusammenfassung unserer Überlegungen formulieren wir etwas um und ergänzen weitere wichtige Beobachtungen (**Existenz- und Eindeutigkeitssatz für die Exponentialfunktion**):

Satz 3.3.2.1. *Für jedes $a > 0$ gibt es genau eine auf ganz \mathbb{R} definierte Funktion (die wir mit \exp_a bezeichnen und die wir die **Exponentialfunktion zur Basis a** nennen) welche folgenden Bedingungen genügt:*

1. $\exp_a(x + y) = \exp_a(x) \exp_a(y)$ für alle $x, y \in \mathbb{R}$ (Funktionalgleichung der Exponentialfunktion)
2. \exp_a ist stetig an wenigstens einem Punkt $x_0 \in \mathbb{R}$.
3. $\exp_a(1) = a$ (Normierung)

Für diese Funktion gilt $\exp_a(x) = a^x$ für alle $x \in \mathbb{R}$. Sie ist stetig auf ganz \mathbb{R} . Für $a \neq 1$ nimmt sie genau sämtliche positive Zahlen $y > 0$ als Werte an. In diesem Fall ist \exp_a streng monoton; wachsend für $a > 1$, fallend für $a < 1$.

Beweis. (Skizze) Die Überlegungen in 3.3.1 sind von Rechenregeln für Exponenten ausgegangen, die im Wesentlichen die Funktionalgleichung aus Bedingung 1 zum Ausdruck bringen. Sie haben gezeigt, dass diese zusammen mit Bedingung 3 die Werte $\exp_a(x)$ für alle $x \in \mathbb{Q}$ eindeutig festlegen. Die Fortsetzung von \mathbb{Q} auf ganz \mathbb{R} folgt, sofern die Funktion stetig ist. Tatsächlich folgt Stetigkeit an einem beliebigen Punkt x_1 aus Kombination der Bedingungen 1 und 2:

$$\begin{aligned} \lim_{x \rightarrow x_1} \exp_a(x) &= \lim_{x \rightarrow x_1} \exp_a((x_1 - x_0) + (x + x_0 - x_1)) \stackrel{1.}{=} \\ &= \lim_{x \rightarrow x_1} (\exp_a(x_1 - x_0)) \cdot \left(\lim_{x \rightarrow x_1} \exp_a(x + x_0 - x_1) \right) = \\ &= \exp_a(x_1 - x_0) \cdot \lim_{x \rightarrow x_0} \exp_a(x) \stackrel{1.,2.}{=} \frac{\exp_a(x_1)}{\exp_a(x_0)} \exp_a(x_0) = \exp_a(x_1). \end{aligned}$$

□

Übungsaufgabe 208. (E) Führen Sie in aller Ausführlichkeit den oben nur skizzierten Beweis von Satz 3.3.2.1. Formulieren Sie dabei auch sorgfältig alle Hilfsbehauptungen und bauen Sie Ihren Beweis entsprechend auf.

In Satz 3.3.2.1 kommen die charakterisierenden Eigenschaften der Exponentialfunktion sehr klar zum Ausdruck. Wir werden sehen, dass eine bestimmte Basis a eine besondere Rolle spielt, nämlich die Basis $a = e := \sum_{n=0}^{\infty} \frac{1}{n!} \approx 2,7 \dots$ (**Eulersche Zahl**). Warum das so ist, wird aber erst im Rahmen der Differentialrechnung klar werden.

Übungsaufgabe 209. (P) Begründen Sie, warum die Gleichung $x = e^{-x}$ eine Lösung im Intervall $[0, 1]$ besitzt und berechnen Sie eine solche Lösung näherungsweise. Sie können dazu die Stetigkeit der Exponentialfunktion als gegeben betrachten.

3.3.3 Logarithmen, Potenzfunktionen, Zusammenfassung

Inhalt in Kurzfassung: Der Logarithmus \log_a zur Basis a lässt sich mittlerweile ohne weitere Schwierigkeiten als Umkehrfunktion der Exponentialfunktion \exp_a definieren. Eine besondere Rolle, die allerdings erst im Zusammenhang mit der Differentialrechnung klar werden wird, spielt dabei die Basis $a = e$ (Eulersche Zahl) mit der Exponentialfunktion $\exp := \exp_e$ und dem natürlichen Logarithmus $\ln := \log_e$. Potenzfunktionen p_α mit beliebigem reellen Exponenten α ergeben sich als Verkettung $p_\alpha = \exp \circ \ln_\alpha \circ \ln$ mit der linearen Funktion $\ln_\alpha : x \mapsto \alpha x$. Sowohl für Logarithmen als auch für Potenz- und für (homogene) lineare Funktionen gelten analoge Existenz- und Eindeutigkeitssätze wie für die Exponentialfunktionen.

Sei $a > 0$ und $a \neq 1$. Als stetige und streng monotone Funktion (Satz 3.3.2.1) hat $\exp_a : \mathbb{R} \rightarrow \mathbb{R}^+ = (0, \infty)$ eine gleichfalls stetige und streng monotone Umkehrfunktion, den Logarithmus zur Basis a , symbolisch $\log_a : \mathbb{R}^+ \rightarrow \mathbb{R}$. Die Funktionalgleichung der Exponentialfunktion spiegelt sich für den Logarithmus in der Gestalt $\log_a(xy) = \log_a(x) + \log_a(y)$ wider. Es gilt auch ein ganz analoger Existenz- und Eindeutigkeitssatz wie bei der Exponentialfunktion, den wir etwas später in Verbindung mit den analogen Sätzen für lineare Funktionen \ln_k und für Potenzfunktionen pot_α beweisen werden. Viele weitere interessante Eigenschaften des Logarithmus werden sich im Rahmen der Differentialrechnung zeigen (4.3.1). Dort wird sich auch klären, warum der Logarithmus \log_e zur Basis e (Eulersche Zahl, wie bereits erwähnt ist $e = \sum_{n=0}^{\infty} \frac{1}{n!} \approx 2,7 \dots$) eine besondere Rolle spielt. Man nennt ihn auch den **natürlichen Logarithmus**, symbolisch $\log_e =: \ln$ für **logarithmus naturalis**.

Mit Hilfe des Logarithmus lassen sich, unter voraussetzender Verwendung gewohnter Rechenregeln, auch die Potenzfunktionen pot_α für beliebiges $\alpha \in \mathbb{R}$ auf angenehme Art anschreiben: $\text{pot}_\alpha(x) = x^\alpha = (e^{\ln x})^\alpha = e^{\alpha \ln x} = \exp(\alpha \ln(x))$. Weil diese Rechenregeln aber noch nicht uneingeschränkt zur Verfügung stehen, nehmen wir die resultierende Beziehung

$$x^\alpha = \text{pot}_\alpha(x) := \exp(\alpha \ln(x))$$

als Definition für pot_α für beliebige $\alpha \in \mathbb{R}$. Somit ist $\text{pot}_\alpha = \exp \circ \ln_\alpha \circ \ln$ die Komposition der stetigen Funktionen $\ln : x \mapsto \ln x$, $\ln_\alpha : y \mapsto \alpha y$ und $\exp : z \mapsto e^z$. Weil diese drei Funktionen stetig sind, ist es auch pot_α .

Überdies ergibt sich jetzt mit Hilfe der Funktionalgleichungen für den Logarithmus und für die Exponentialfunktion ganz allgemein auch jene für Potenzfunktionen:

$$\begin{aligned} \text{pot}_\alpha(xy) &= e^{\alpha \ln(xy)} = e^{\alpha(\ln(x)+\ln(y))} = e^{\alpha \ln(x)+\alpha \ln(y)} = e^{\alpha \ln(x)} e^{\alpha \ln(y)} = \\ &= \text{pot}_\alpha(x) \text{pot}_\alpha(y) \end{aligned}$$

In der obigen Definition der Potenzfunktion $\text{pot}_\alpha : x \mapsto x^\alpha$ sind Exponentialfunktion und natürlicher Logarithmus \ln zur speziellen Basis e verwendet worden. Man hätte in dieser Definition aber genauso gut auch eine andere Basis $a \neq e$ nehmen können.

Übungsaufgabe 210. (E) Beweisen Sie dies, indem Sie für beliebige Basen $a, b, x > 0$, $a, b \neq 1$, die Gleichung

$$\exp_a(\alpha \log_a(x)) = \exp_b(\alpha \log_b(x))$$

zeigen.

Eine nützliche Folgerung besteht in der vertrauten Rechenregel $(x^\alpha)^\beta = x^{\alpha\beta}$ aus 3.3.1, die wir nun für beliebige $\alpha, \beta \in \mathbb{R}$ und nicht nur wie dort für rationale zeigen können:

$$(x^\alpha)^\beta = (\exp(\alpha \ln(x)))^\beta = \exp(\beta \ln(\exp(\alpha \ln(x)))) = \exp(\beta\alpha \ln(x)) = x^{\alpha\beta}$$

Damit können wir nun auch leicht zwischen Logarithmen zu verschiedenen Basen a und b umrechnen. Dazu gehen wir aus von $x = a^{\log_a(x)} = (e^{\ln(a)})^{\log_a(x)} = e^{\ln(a) \log_a(x)}$, analog $x = e^{\ln(b) \log_b(x)}$, also $\ln(a) \log_a(x) = \ln(x) = \ln(b) \log_b(x)$, folglich

$$\log_b(x) = \alpha \log_a(x) \quad \text{mit} \quad \alpha = \frac{\ln(a)}{\ln(b)},$$

und daher speziell für $a = e$:

$$\log_b(x) = \frac{1}{\ln(b)} \ln(x).$$

Noch eine wichtige Rechenregel ergibt sich unmittelbar aus der Definition von pot_α : $\ln(x^\alpha) = \ln(\exp(\alpha \ln(x))) = \alpha \ln(x)$. Analog gilt

$$\log_a(x^\alpha) = \alpha \log_a(x)$$

für beliebiges $a > 0$, $a \neq 1$.

Übungsaufgabe 211. (T) Beweisen Sie diese Regel für beliebiges $a > 0$, $a \neq 1$.

Nun zu den bereits angekündigten **Existenz- und Eindeutigkeitssätzen**. Ganz analoge Aussagen wie Satz 3.3.2.1 über die Exponentialfunktion gelten auch für **lineare, Logarithmus- und für Potenzfunktionen**. Um die Analogie hervorzuheben, fassen wir im folgenden Satz alle vier Typen zusammen, inklusive der bereits in Satz 3.3.2.1 behandelten Exponentialfunktionen nun in etwas veränderter Formulierung.

Satz 3.3.3.1. Die Exponentialfunktionen $\exp_a, x \mapsto a^x$ mit $a > 0$, die lineare Funktionen $\text{lin}_k, x \mapsto kx$, mit $k \in \mathbb{R}$, die Logarithmusfunktionen $\log_a, x \mapsto \log_a(x)$, mit $a > 0$ und $a \neq 1$ sowie die Potenzfunktionen $\text{pot}_\alpha, x \mapsto x^\alpha$ mit $\alpha \in \mathbb{R}$, sind durch folgende Eigenschaften charakterisiert.

1. Für jedes $a > 0$ gibt es genau eine auf ganz \mathbb{R} definierte Funktion f , welche folgenden Bedingungen genügt:

- (a) $f(x + y) = f(x)f(y)$ für alle $x, y \in \mathbb{R}$
- (b) f ist stetig an wenigstens einem Punkt $x_0 \in \mathbb{R}$.
- (c) $f(1) = a$

Bei dieser Funktion f handelt es sich um die **Exponentialfunktion** zur Basis a : $f(x) = \exp_a(x) = a^x$ für alle $x \in \mathbb{R}$.

2. Für jedes $k \in \mathbb{R}$ gibt es genau eine auf ganz \mathbb{R} definierte Funktion f , welche folgenden Bedingungen genügt:

- (a) $f(x + y) = f(x) + f(y)$ für alle $x, y \in \mathbb{R}$
- (b) f ist stetig an wenigstens einem Punkt $x_0 \in \mathbb{R}$.
- (c) $f(1) = k$

Bei dieser Funktion f handelt es sich um die (homogene) **lineare Funktion** mit Steigung k : $f(x) = kx$ für alle $x \in \mathbb{R}$.

3. Für jedes $a > 0, a \neq 1$ gibt es genau eine auf \mathbb{R}^+ definierte Funktion f , welche folgenden Bedingungen genügt:

- (a) $f(xy) = f(x) + f(y)$ für alle $x, y \in \mathbb{R}^+$
- (b) f ist stetig an wenigstens einem Punkt $x_0 \in \mathbb{R}^+$.
- (c) $f(a) = 1$

Bei dieser Funktion f handelt es sich um den **Logarithmus** zur Basis a : $f(x) = \log_a(x)$ für alle $x \in \mathbb{R}^+$.

4. Für jedes $\alpha \in \mathbb{R}$ gibt es genau eine auf \mathbb{R}^+ definierte Funktion f , welche folgenden Bedingungen genügt:

- (a) $f(xy) = f(x)f(y)$ für alle $x, y \in \mathbb{R}^+$
- (b) f ist stetig an wenigstens einem Punkt $x_0 \in \mathbb{R}^+$.
- (c) $f(e) = e^\alpha$ (Äquivalent kann $f(b) = b^\alpha$ für ein beliebiges $b > 0, b \neq 1$ gefordert werden.)

Bei dieser Funktion f handelt es sich um die **Potenzfunktion** mit Exponenten α : $f(x) = \text{pot}_\alpha(x) = x^\alpha$ mit für alle $x \in \mathbb{R}^+$.

Beweis. Zunächst beobachten wir, dass die Funktionen $\exp_a : \mathbb{R} \rightarrow \mathbb{R}^+$ für 1., $\ln_k : \mathbb{R} \rightarrow \mathbb{R}$ für 2., $\log_a : \mathbb{R}^+ \rightarrow \mathbb{R}$ für 3. und $\text{pot}_\alpha : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ für 4. die Bedingungen a), b) und c) erfüllen. Es gibt also Funktionen mit den geforderten Eigenschaften. Damit ist für jede der vier Behauptungen nur mehr die Eindeutigkeitsaussage zu beweisen. Zu diesem Zwecke besteht unsere Beweisstrategie darin, die erste Aussage aus Satz 3.3.2.1 zu übernehmen und die anderen drei auf jeweils sehr ähnliche Weise auf vorangegangene zurückzuführen.

1. Folgt direkt aus Satz 3.3.2.1.

2. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ irgendeine Funktion, die a), b) und c) aus Behauptung 2 erfüllt. Mit diesem f definieren wir die Funktion $g := \exp \circ f : \mathbb{R} \rightarrow \mathbb{R}^+$, womit umgekehrt $f = \ln \circ g$ gewährleistet ist. Man sieht sehr schnell, dass g die Bedingungen a) und b) an die Funktion f aus Teil 1 erfüllt:

a) $g(x+y) = e^{f(x+y)} = e^{f(x)+f(y)} = e^{f(x)}e^{f(y)} = g(x)g(y)$ für alle $x, y \in \mathbb{R}$.

b) f ist stetig in $x_0 \in \mathbb{R}$, \exp ist global stetig, daher (Satz 3.1.3.2) ist $g = \exp \circ f$ stetig in x_0 .

c) $g(1) = e^{f(1)} = e^k =: a$.

Aus Teil 1 folgt damit $g = \exp_a = \exp_{e^k}$, woraus $f(x) = \ln(g(x)) = \ln((e^k)^x) = \ln(e^{kx}) = kx$ für alle $x \in \mathbb{R}$ folgt, was zu zeigen war.

3. Sei $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ irgendeine Funktion, die a), b) und c) aus Behauptung 3 erfüllt. Mit diesem f definieren wir die Funktion $g := f \circ \exp : \mathbb{R} \rightarrow \mathbb{R}$, womit umgekehrt $f = g \circ \ln$ gewährleistet ist. Diesmal werden wir zeigen, dass die Funktion g die Bedingungen a), b) und c) an die Funktion f aus Teil 2 erfüllt:

a) $g(x+y) = f(e^{x+y}) = f(e^x e^y) = f(e^x) + f(e^y) = g(x) + g(y)$ für alle $x, y \in \mathbb{R}$.

b) f ist stetig in $x_0 \in \mathbb{R}^+$, \exp ist global stetig, daher (wieder Satz 3.1.3.2) ist $g = f \circ \exp$ stetig in $\ln(x_0)$.

Aus Teil 2 folgt damit $g = \ln_k$ mit $k = g(1)$, also $f(x) = g(\ln(x)) = k \ln(x) = \frac{1}{\ln(a')} \ln(x) = \log_{a'}(x)$ für alle $x \in \mathbb{R}^+$, sofern man $k := \frac{1}{\ln(a')}$ setzt. Laut Voraussetzung a) an f gilt daher speziell ($x = a$ setzen) $\log_{a'}(a) = f(a) = 1$, was wegen der Injektivität der Logarithmen nur für $a' = a$ und somit $f = \log_a$ möglich ist.

4. Sei $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ irgendeine Funktion, die a), b) und c) aus Behauptung 4 erfüllt. Mit diesem f definieren wir die Funktion $g := f \circ \exp f : \mathbb{R} \rightarrow \mathbb{R}^+$, womit umgekehrt $f = g \circ \ln$ gewährleistet ist. Diese Funktion g erfüllt die Bedingungen a) und b) aus Teil 1:

a) $g(x+y) = f(e^{x+y}) = f(e^x e^y) = f(e^x)f(e^y)$ für alle $x, y \in \mathbb{R}$.

b) f ist stetig in $x_0 \in \mathbb{R}$, \exp ist global stetig, daher (Satz 3.1.3.2) ist $g = f \circ \exp$ stetig in $\ln(x_0)$.

Somit folgt aus Teil 1 $g = \exp_a$ mit $a := g(1)$ und somit $f(x) = g(\ln(x)) = a^{\ln(x)} = e^{\ln(a) \ln(x)} = x^{\ln(a)}$ für alle $x \in \mathbb{R}^+$, also $f = \text{pot}_\alpha$ mit $\alpha = \ln(a)$ folgt (was auch mit Voraussetzung c) kompatibel ist).

□

3.3.4 Trigonometrische Funktionen

Inhalt in Kurzfassung: Die wichtigsten Eigenschaften der trigonometrischen Funktionen Cosinus, Sinus, Tangens und Cotangens werden auf elementargeometrischem Wege hergeleitet. Auf jeweils geeigneten Definitionsbereichen werden auch die Umkehrfunktionen, die sogenannten Arcusfunktionen definiert. Später wird die Differentialrechnung wesentliche Vertiefungen des Verständnisses und Exaktifizierungen ermöglichen, siehe 4.3.3 und 4.3.4.

Im Gegensatz zu Exponentialfunktion und Logarithmus können wir für die trigonometrischen Funktionen Sinus, Cosinus, Tangens und Cotangens an dieser Stelle noch keine überzeugende rigorose Theorie begründen. Wohl wäre es möglich, z.B. den Cosinus schon jetzt über die sogenannte Potenzreihe $\cos x := \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$ zu definieren. Wir kennen bereits das Quotientenkriterium, mit dem man leicht nachprüft, dass diese Reihe für alle $x \in \mathbb{R}$ (sogar für alle $x \in \mathbb{C}$) konvergiert. Doch fehlen uns noch die Grundlagen, um den Zusammenhang zur üblichen geometrischen Interpretation zu verstehen. Wir wollen deshalb vorläufig nur ein paar wichtige Eigenschaften aus der elementaren Anschauung ablesen, an denen wir die trigonometrischen Funktionen später auch in anderem Gewand wiedererkennen werden. Auch für die Definition lassen wir uns davon leiten.

Wir beginnen mit der Zeichenebene \mathbb{R}^2 , die wir mit einem kartesischen (rechtwinkligen) Koordinatensystem, bestehend aus x -Achse $\{(x, 0) : x \in \mathbb{R}\}$ und y -Achse $\{(0, y) : y \in \mathbb{R}\}$ versehen. Um den Ursprung mit den Koordinaten $(0, 0)$ als Mittelpunkt legen wir einen Kreis K mit dem Radius 1 (**Einheitskreis**), also $K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Wir bewegen uns entlang der Kreislinie, beginnend beim Punkt $(1, 0)$ entgegen dem Uhrzeigersinn. Überstrichene Winkel messen wir im **Bogenmaß**, d.h. als Länge α des zurückgelegten Kreisbogens. (Das Nichtrigorose unseres Zugangs tritt an dieser Stelle auf, weil der Begriff der Bogenlänge nur in einer naiven Weise zur Verfügung steht.) Die **Kreiszahl** π ist (hier eben nur geometrisch-anschaulich, nicht rigoros) so definiert, dass wir genau bei $\alpha = 2\pi$ die erste volle Umrundung am Kreis K durchgeführt haben. Die Zahl π entspricht also einem Winkel von 180° . Offensichtlich ist $\pi \leq 4$ (halber Umfang des k umschriebenen Quadrats), etwas feinsinnigere Überlegungen zeigen $\pi \geq 3$ oder, mit deutlich mehr Aufwand, $3,1 \leq \pi \leq 3,2$. Mit Hilfe der Differentialrechnung werden wir effektivere Möglichkeiten zur Berechnung von π kennen lernen.

Für jedes α erreichen wir, nachdem wir auf K die Bogenlänge α zurückgelegt haben, einen Punkt $\phi(\alpha)$ dessen x -Koordinate wir mit $\cos \alpha$ und dessen y -Koordinate wir mit $\sin \alpha$ bezeichnen, also $\phi(\alpha) = (\cos \alpha, \sin \alpha)$. Da wir die Kreislinie in beide Richtungen beliebig weit entlang gehen, d.h. beliebig oft durchlaufen können, sind dadurch zwei reelle Funktionen

$$\text{Cosinus} : \alpha \mapsto \cos \alpha \quad \text{und} \quad \text{Sinus} : \alpha \mapsto \sin \alpha$$

definiert. Man kann auch formulieren: In einem rechtwinkligen Dreieck mit Hypotenuse (längste Seite) der Länge 1 ist der Cosinus eines spitzen Winkels (d.h. $\alpha < \frac{\pi}{2}$) die Länge der anliegenden Kathete = Ankathete, der Sinus die

der Gegenkathete, oder kurz für beliebige Hypotenusen: Cosinus = Ankathete durch Hypotenuse, Sinus = Gegenkathete durch Hypotenuse. Beide Funktionen sind periodisch mit kleinster Periode 2π . Es gilt also $\cos(\alpha + 2\pi) = \cos(\alpha)$ und $\sin(\alpha + 2\pi) = \sin(\alpha)$, außerdem $\cos(\alpha) = \sin(\alpha + \frac{\pi}{2})$. Die Wertemenge beider Funktionen ist $[-1, 1]$. An den Stellen $\alpha = k\frac{\pi}{2}$, $k \in \mathbb{Z}$, nehmen Cosinus und Sinus immer nur einen der Werte $-1, 0, 1$ an. Welchen, hängt davon ab, welchen Wert der Rest $r \in \{0, 1, 2, 3\}$ in $k = 4n + r$ mit $n \in \mathbb{Z}$ (Division von k durch 4 mit Rest r) hat.

Wir wiederholen aus 1.3.11: Weil die Koordinaten von $\phi(\alpha)$ die Kreisgleichung erfüllen, gilt $\cos^2 \alpha + \sin^2 \alpha = 1$. Bei $\alpha = \frac{\pi}{4}$ überlegt man sich damit rasch $\cos \frac{\pi}{4} = \sin \frac{\pi}{4} = \frac{\sqrt{2}}{2}$. Das Dreieck, das von den Punkten $(0, 0)$, $(1, 0)$ und $\phi(60^\circ) = \phi(\frac{\pi}{3})$ gebildet wird, ist ein gleichseitiges. Daraus wiederum folgt $\cos \frac{\pi}{3} = \frac{1}{2}$ und (Kreisgleichung) $\sin \frac{\pi}{3} = \sqrt{1 - \cos^2 \frac{\pi}{3}} = \frac{\sqrt{3}}{2}$. Aus Symmetriegründen schließen wir daraus $\cos \frac{\pi}{6} = \frac{\sqrt{3}}{2}$ und $\sin \frac{\pi}{6} = \frac{1}{2}$. Es ist klar, dass man aus den bereits bekannten Funktionswerten von \cos und \sin an den Stellen $\alpha = 0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}$ auch leicht die Werte an allen Stellen der Form $\alpha + \frac{k\pi}{2}$ mit $k \in \mathbb{Z}$ erhält.

Eine sehr wichtige Rolle spielen die Additionstheoreme

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$$

und

$$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta),$$

siehe Übungsaufgabe 68. Verwendet haben wir sie bereits in 1.3.11 bei der Polardarstellung komplexer Zahlen und ihrer Multiplikation, bewiesen in 1.5.4 im Zusammenhang mit linearen Abbildungen, indem wir die Komposition der Drehungen um die Winkel α und β studierten.

Cosinus und Sinus sind stetig, weil die Veränderung von x - und y -Koordinate am Einheitskreis sicher durch die zurückgelegte Bogenlänge nach oben beschränkt ist, als Formel: $|\cos \alpha - \cos \beta| \leq |\alpha - \beta|$ und $|\sin \alpha - \sin \beta| \leq |\alpha - \beta|$. In der Stetigkeitsdefinition kann also stets $\delta = \varepsilon$ gesetzt werden.

Mit Hilfe von Cosinus und Sinus definiert man die Funktionen **Tangens** als $\tan := \frac{\sin}{\cos}$ und **Cotangens** als $\cot := \frac{\cos}{\sin}$. Für die Definitionsbereiche müssen beim Tangens alle $\alpha = \frac{\pi}{2} + k\pi$ (Nullstellen von \cos), beim Cotangens alle $\alpha = k\pi$ (Nullstellen von \sin), jeweils mit $k \in \mathbb{Z}$, ausgenommen werden. Man überlegt sich geometrisch, dass man $\tan \alpha$ auch als y -Koordinate jenes Punktes erhält, in dem sich die Gerade g durch $(0, 0)$ und $\phi(\alpha)$ mit der Gerade $x = 1$ schneidet. Schneidet man g mit $y = 1$, so erhält man einen Punkt mit x -Koordinate $\cot \alpha$. Rein geometrische Merkgelien: Tangens = Gegenkathete durch Ankathete, Cotangens = Ankathete durch Gegenkathete.

Übungsaufgabe 212. (E) *Tangens und Cotangens sind mittels Sinus und Cosinus definiert. Wegen $\sin^2 + \cos^2 = 1$ definieren \sin und \cos einander bis aufs Vorzeichen. Folglich müssen auch \tan^2 und \cot^2 sowohl durch \sin^2 als auch durch \cos^2 eindeutig bestimmt sein. Diese Zusammenhänge lassen sich auch*

umkehren. Ihre Aufgabe besteht darin, diese Zusammenhänge vollständig zu erfassen, indem Sie

1. \sin^2 , \cos^2 und \tan^2 als Funktionen von \cot^2
2. \sin^2 , \cos^2 und \cot^2 als Funktionen von \tan^2
3. \sin^2 , \tan^2 und \cot^2 als Funktionen von \cos^2
4. \cos^2 , \tan^2 und \cot^2 als Funktionen von \sin^2

ausdrücken.

Die trigonometrischen Funktionen sind nicht auf ihren gesamten Definitionsbereichen bijektiv, sehr wohl aber, wenn man sie auf gewisse Teilbereiche, wo sie streng monoton sind, einschränkt. Dort lassen sich (bijektive) Umkehrfunktionen definieren, die sogenannten **Arcusfunktionen**:

Arcus cosinus: $\arccos := \cos^{(-1)} : [-1, 1] \rightarrow [0, \pi]$

Arcus sinus: $\arcsin := \sin^{(-1)} : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$

Arcus tangens: $\arctan := \tan^{(-1)} : \mathbb{R} \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$

Arcus cotangens: $\operatorname{arccot} := \cot^{(-1)} : \mathbb{R} \rightarrow (0, \pi)$

Übungsaufgabe 213. (P) Jede harmonische Schwingung der Form

$$A \sin(\omega t + \delta)$$

mit Amplitude A , Kreisfrequenz ω und Phasenfaktor δ ist auch in der Form

$$A \sin(\omega t + \delta) = a \cos(\omega t) + b \sin(\omega t)$$

darstellbar.

1. Geben Sie den Zusammenhang zwischen den Parametern A, δ einerseits und a, b andererseits explizit an.
2. Wandeln Sie $\sin(t + \frac{\pi}{4})$ bzw. $\sin t + \cos t$ in die jeweils andere Darstellung um.

Weitere wichtige Eigenschaften der trigonometrischen Funktionen, wie zum Beispiel den Zusammenhang mit der Exponentialfunktion über den Umweg der komplexen Zahlen, werden wir mit Hilfe der Differentialrechnung erschließen.

Kapitel 4

Differentialrechnung

Stetigkeit haben wir kennengelernt als eine Eigenschaft von Funktionen, die sich auf das lokale Änderungsverhalten der Funktionswerte bezieht. Genauer lässt sich sagen: Gehen wir von einer Stelle x_0 aus und ändern wir sie um $\Delta x = x - x_0$ geringfügig ab bis zu x , so bedeutet das für den Funktionswert $f(x) = f(x_0) + r(x)$ eine Änderung, die sich im Restterm $\Delta f = r(x)$ niederschlägt. Bei stetigen Funktionen folgt aus $x \rightarrow x_0$ die Konvergenz $f(x) \rightarrow f(x_0)$ der Funktionswerte, also $r(x) \rightarrow 0$. Bei differenzierbaren Funktionen kann man noch genauere Aussagen darüber machen, mit welcher Geschwindigkeit die Konvergenz $r(x) \rightarrow 0$ erfolgt, wenn $x \rightarrow x_0$. Und zwar herrscht annähernd lineare Konvergenz. In 4.1 behandeln wir den grundlegenden Begriff der Ableitung einer Funktion und wie sie sich berechnen lässt. Iteriert man das Ableiten, stößt man auf höhere Ableitungen, Taylor- und Potenzreihen (4.2). Wichtige Beispiele differenzierbarer Funktionen folgen in 4.3, Anwendungen der Differentialrechnung in 4.4.

4.1 Die Ableitung einer reellen Funktion

Historisch hatte die Differentialrechnung im physikalischen Zugang Newtons und im geometrischen Zugang Leibniz' ihren Ursprung (4.1.1). Der oben skizzierten Idee folgend, lässt sich die Ableitung nicht nur als Differentialquotient definieren, sondern auch mittels linearer Approximation (4.1.2). Der resultierende Begriff der Ableitung einer reellen Funktion ist der grundlegende Begriff dieses Abschnitts. In 4.1.3 finden sich zahlreiche Regeln zur Berechnung der Ableitung. In 4.1.4 wird die Ableitung von Umkehrfunktionen in Beziehung gesetzt zur Ableitung der ursprünglichen Funktion.

4.1.1 Motivation: Tangente, Momentangeschwindigkeit

Inhalt in Kurzfassung: Der historischen Parallelität von Leibniz und Newton sowie ihren beiden Zugängen zur Differentialrechnung entsprechend motivieren

wir den später zu exaktifizierenden Begriff der Ableitung auf zweifache Weise: über das Tangentenproblem (Leibniz) und über die Momentangeschwindigkeit (Newton). Sehr schnell erkennt man, dass es sich rein mathematisch um dasselbe Problem handelt.

Als Begründer der Infinitesimalrechnung (= Differential- und Integralrechnung) gelten Isaac Newton (1643-1727) und Gottfried Wilhelm Leibniz (1646-1716). Newton ging vom physikalischen Problem der Momentangeschwindigkeit aus, Leibniz vom geometrischen Tangentenproblem. Beides führt zum selben Begriff, nämlich zu dem der Ableitung einer Funktion.

Bekanntlich ist die Geschwindigkeit v definiert als Quotient aus zurückgelegtem Weg durch verstrichene Zeit. Ist t unsere Variable für die Zeit und bezeichnet $s(t)$ den zum Zeitpunkt t zurückgelegten Weg, so gilt also $v = \frac{s(t_1)-s(t_0)}{t_1-t_0}$ (**Differenzenquotient**). Dadurch wird aber nur die Durchschnittsgeschwindigkeit zwischen den Zeitpunkten t_0 und t_1 gemessen. Weil sich die Geschwindigkeit ja während dieser Zeitspanne ändern kann, kennen wir damit nicht die noch zu definierende Momentangeschwindigkeit $v(t)$ zu einem Zeitpunkt t . Die Hoffnung besteht darin, dass bei Annäherung von t_1 an t_0 der Differenzenquotient einem Wert zustrebt, den wir dann als $v(t_0)$ nehmen können.

Ganz ähnlich ist die Situation beim Tangentenproblem. Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ eine reelle Funktion, so interessiert uns die Steigung der Tangente an den Graphen in einem Punkt $(x_0, f(x_0))$. Die Steigung einer linearen Funktion l ist durch $\frac{l(x_1)-l(x_0)}{x_1-x_0}$ definiert und lässt sich als durchschnittliche Änderungsrate der Funktion interpretieren. Sie hängt bei linearem l nicht von der speziellen Wahl der Punkte x_0 und x_1 ab, bei nichtlinearem f aber sehr wohl. Für jedes Punktepaar $x_0 \neq x_1$ lässt sich eine lineare Funktion durch $(x_0, f(x_0))$ und $(x_1, f(x_1))$ legen (Interpolationspolynom vom Grad 1, siehe 3.2.5), die Sekante des Graphen durch diese Punkte. Ihre Steigung ist durch den Differenzenquotienten $\frac{f(x_1)-f(x_0)}{x_1-x_0}$ gegeben. Nähert sich x_1 dem Punkt x_0 an, so besteht Hoffnung, dass sich die Steigung der Sekante an die gesuchte Steigung der Tangente annähert. Bis auf die Bezeichnungsweise (Variablenbezeichnung x statt t , Funktionsbezeichnung f statt s) sind wir exakt beim selben mathematischen Problem angelangt wie zuvor, nämlich den Grenzwert von Differenzenquotienten zu bestimmen.

Übungsaufgabe 214. (E) Galileo Galilei beschrieb als erster das Fallgesetz, wonach ein knapp über der Erdoberfläche ruhender und ab dem Zeitpunkt $t = 0$ zu Boden fallender Körper, gäbe es keinen Luftwiderstand, bevor er den Boden erreicht nach t Zeiteinheiten eine Fallhöhe von ct^2 Längeneinheiten zurücklegen würde. Dabei hängt die Konstante $c = \frac{g}{2}$, $g =$ Erdbeschleunigung, nicht vom fallenden Körper ab. Setzen Sie anhand dieser Situation die Newtonsche Frage nach der Momentangeschwindigkeit durch eine geeignete graphische Illustration des Problems mit der Leibnizschen Frage nach der Tangentensteigung in Beziehung.

4.1.2 Differentialquotient und lineare Approximation

Inhalt in Kurzfassung: Die angekündigte Exaktifizierung des Begriffs der Ableitung (und damit auch gleich der Stammfunktion) einer Funktion erfolgt mit Hilfe ihres Differentialquotienten, der als Grenzwert des Differenzenquotienten definiert wird. Es ist ein sehr fruchtbarer Gesichtspunkt, die Ableitung auch als lineare Approximation von Funktionen aufzufassen, wie es schon in Leibniz' Zugang über das Tangentenproblem augenscheinlich war. Auch einseitige Ableitungen sind noch Gegenstand dieses Unterabschnitts.

Obige Überlegungen legen die folgende, für dieses Kapitel und weit darüber hinaus grundlegende Definition nahe.

Definition 4.1.2.1. Sei $f : D \rightarrow \mathbb{R}$ eine reelle Funktion und $x_0 \in D$ innerer Punkt von D . Existiert der Grenzwert (der sogenannte **Differentialquotient**)

$$f'(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

so heißt f **differenzierbar** in x_0 . Ist D offen und f differenzierbar in x für alle $x \in D$, so heißt f differenzierbar auf D . Die Funktion $f' : D \rightarrow \mathbb{R}$, $x \mapsto f'(x)$ heißt dann die **Ableitung** von f , f heißt **Stammfunktion** von f' .

Die einfachsten Beispiele: Konstante Funktionen $f : x \mapsto c$, $c \in \mathbb{R}$ fest, sind differenzierbar, und ihre Ableitung ist wegen

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{c - c}{x - x_0} = 0$$

die Nullfunktion. Die Funktion $f : x \mapsto x$ ist ebenfalls differenzierbar mit konstanter Ableitung, jedoch mit Wert 1:

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{x - x_0}{x - x_0} = 1.$$

Ein drittes, etwas weniger triviales Beispiel ist die Funktion $f : x \mapsto \frac{1}{x}$. Auch sie erweist sich für alle $x_0 \in \mathbb{R} \setminus \{0\}$ als differenzierbar:

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{\frac{1}{x} - \frac{1}{x_0}}{x - x_0} = \lim_{x \rightarrow x_0} \frac{x_0 - x}{xx_0(x - x_0)} = \lim_{x \rightarrow x_0} -\frac{1}{xx_0} = -\frac{1}{x_0^2}.$$

Wir dürfen also die Ableitungsregeln $c' = 0$, $x' = 1$ und $\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$ festhalten.

Übungsaufgabe 215. (T) Berechnen Sie analog, d.h. durch direkten Rückgriff auf die Definition, f' für:

1. $f(x) = x^2$
2. $f(x) = x^n$ für allgemeines $n \in \mathbb{N}$
3. $f(x) = \sqrt{x}$ (nur für $x > 0$)

4. $f(x) = \sqrt[n]{x}$ für allgemeines $n \in \mathbb{N}$, $n \geq 1$, und $x > 0$

Schreiben wir für eine in x_0 differenzierbare Funktion f die Situation in der Form

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x)$$

an und denken wir uns $r(x)$ als „sehr kleinen“ Rest, so schließen sich daran erhellende Überlegungen an. Diese Formel bringt zum Ausdruck, dass sich der Funktionswert $f(x)$ an einer anderen (benachbarten) Stelle x ungefähr ergibt, indem man von $f(x_0)$ ausgehend einen Summanden, nämlich $f'(x_0)(x - x_0)$ addiert, der proportional ist zum Anstieg $f'(x_0)$ von f in x_0 und zum Abstand $x - x_0$. Damit hat man $f(x)$ angenähert – bis auf einen Fehler oder Rest $r(x)$, von dem man hofft, dass er klein wird, wenn x sich x_0 annähert. Der Unterschied zur Stetigkeit besteht darin, dass $r(x)$ für $x \rightarrow x_0$ nicht nur gegen 0 konvergiert, sondern sogar dann, wenn man durch $x - x_0$ dividiert. Denn wegen $\frac{r(x)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0)$ ergibt sich im Fall von Differenzierbarkeit von f in x_0 für den Fehler $r(x)$ die Beziehung

$$\lim_{x \rightarrow x_0} \frac{r(x)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) = f'(x_0) - f'(x_0) = 0.$$

Vergleich mit der Stetigkeit (wo $\lim_{x \rightarrow x_0} r(x) = 0$ mit $r(x) = f(x) - f(x_0)$ gilt) zeigt: Subtrahieren wir von f die lineare Funktion $x \mapsto f(x_0) + f'(x_0)(x - x_0)$, so garantiert Differenzierbarkeit von f für den Rest $r(x)$ viel schnellere Konvergenz gegen 0, als das bei bloßer Stetigkeit der Fall ist.

Anders formuliert: Bei der Stetigkeitsdefinition wird f durch die konstante Funktion mit Wert $f(x_0)$ approximiert, bei Differenzierbarkeit durch die (etwas kompliziertere) lineare Funktion $x \mapsto f(x_0) + f'(x_0)(x - x_0)$. Die Belohnung für den größeren Aufwand ist die bessere Approximation.

Im Vergleich mit Stetigkeit haben wir Differenzierbarkeit folglich als noch feinere Approximierbarkeitseigenschaft identifiziert. Folgender Satz sollte also nicht überraschen:

Satz 4.1.2.2. *Ist die reelle Funktion $f : D \rightarrow \mathbb{R}$ im Punkt $x_0 \in D$ differenzierbar, so ist f in x_0 auch stetig.*

Beweis. Aus der Differenzierbarkeitsbedingung

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x)$$

mit $\lim_{x \rightarrow x_0} \frac{r(x)}{x - x_0} = 0$ folgt wegen $r(x) = (x - x_0) \frac{r(x)}{x - x_0}$ auch $r(x) \rightarrow 0$ für $x \rightarrow x_0$, und damit erst recht $\lim_{x \rightarrow x_0} f(x) = f(x_0)$. \square

Eine schärfere Aussage, die noch deutlicher den Gesichtspunkt der linearen Approximation ins Zentrum rückt, lautet:

Satz 4.1.2.3. *Eine reelle Funktion $f : D \rightarrow \mathbb{R}$ ist im inneren Punkt x_0 von D genau dann differenzierbar, wenn es eine lineare Approximation $l : x \mapsto kx$ ($k \in \mathbb{R}$) von f gibt in dem Sinn, dass*

$$f(x) = f(x_0) + l(x - x_0) + r(x) = f(x_0) + k(x - x_0) + r(x)$$

gilt mit $\lim_{x \rightarrow x_0} \frac{r(x)}{x-x_0} = 0$. In diesem Fall ist $k \in \mathbb{R}$ durch diese Bedingung eindeutig bestimmt als $k = f'(x_0)$.

Eine ungenaue, aber oft sehr praktische Schreibweise verwendet $\Delta x := x - x_0$ (der griechische Buchstabe Δ , sprich Delta, wird oft für kleine, meist positive Differenzen verwendet):

$$f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x$$

mit einem Fehler viel kleiner als Δx , sobald Δx hinreichend klein wird.

Manchmal ist es von Interesse, auch Randpunkte x_0 von D zuzulassen und die sogenannten **einseitigen Ableitungen**

$$f'(x_0^-) = f'_-(x_0) := \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \quad (\text{linksseitige Ableitung})$$

bzw.

$$f'(x_0^+) = f'_+(x_0) := \lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \quad (\text{rechtsseitige Ableitung})$$

zu definieren. Die Funktion $f : x \rightarrow |x|$ beispielsweise hat an der Stelle 0 die linksseitige Ableitung $f'(0^-) = -1$ und die rechtsseitige Ableitung $f'(0^+) = 1$. Weil diese beiden Werte verschieden sind, existiert $f'(0)$ nicht. Also ist f in 0 nicht differenzierbar. Anschaulich zeigt sich das daran, dass der Funktionsgraph von f bei $x = 0$ ein Eck hat, das auch bei beliebiger Vergrößerung des Bildes nie verschwindet und nie zu einer glatten Kurve mit Tangente wird.

Ist für eine Funktion $f : [a, b] \rightarrow \mathbb{R}$, $a < b$, von der Ableitung bei a oder b die Rede, so ist stets $f'(a^+)$ bzw. $f'(b^-)$ gemeint.

Übungsaufgabe 216. (T) Wie verhält sich die Funktion $f : [0, \infty) \rightarrow \mathbb{R}$, $f(x) := \sqrt{x}$, hinsichtlich Differenzierbarkeit an der Stelle $x = 0$?

Übungsaufgabe 217. (T) Wählen Sie die reellen Parameter a, b und c so, dass die auf ganz \mathbb{R} definierte Funktion

$$f(x) = \begin{cases} 4x & \text{falls } x \leq 0, \\ ax^2 + bx + c & \text{falls } 0 < x \leq 1, \\ 3 - 2x & \text{falls } x > 1. \end{cases}$$

1. stetig und differenzierbar ist.
2. stetig, aber nicht differenzierbar ist.
3. in 0 und 1 unstetig ist.

4.1.3 Ableitungsregeln

Inhalt in Kurzfassung: Wir wollen nun Rechenregeln herleiten, die es erlauben, die Ableitung einer Funktion zu berechnen, die sich in unterschiedlicher Weise aus anderen Funktionen aufbaut, deren Ableitung bereits bekannt ist. Und zwar behandeln wir Summe, Produkt, Verkettung und Quotienten zweier Funktionen.

Der Übersicht halber stellen wir zunächst ohne weitere Kommentare und Spezifikationen die wichtigsten Ableitungsregeln zusammen. Präzisierungen, Erläuterungen und Beweise werden im Laufe dieses und späterer Unterabschnitte folgen.

konstante Funktionen: $c' = 0$ (c konstante Funktion)

Identität: $x' = 1$ (x steht für die identische Funktion $x \mapsto x$)

Inversion: $\text{pot}'_{-1} = -\text{pot}_{-2}$, d.h. $(\frac{1}{x})' = -\frac{1}{x^2}$

einfache Potenzen: $\text{pot}'_n = n \text{pot}_{n-1}$ (zunächst für $n \in \mathbb{N}$)

Linearität: $(f + g)' = f' + g'$ und $(cf)' = cf'$ (für konstantes c)

Ableitung von Polynomen: $(\sum_{i=0}^N a_n x^n)' = \sum_{n=0}^N n a_n x^{n-1} = \sum_{n=0}^{N-1} (n + 1) a_{n+1} x^n$

Produktregel: $(fg)' = fg' + f'g$

Kettenregel: $(g \circ f)' = (g' \circ f)f'$

Quotientenregel: $(\frac{f}{g})' = \frac{f'g - fg'}{g^2}$

Umkehrfunktionen: $(f^{-1})' = \frac{1}{f' \circ f^{-1}}$

Exponentialfunktion: $\exp' = \exp$

Logarithmus: $\ln' = \text{pot}_{-1}$, d.h. $\ln'(x) = \frac{1}{x}$

allgemeine Potenzfunktionen: $\text{pot}'_\alpha = \alpha \text{pot}_{\alpha-1}$, d.h. $(x^\alpha)' = \alpha x^{\alpha-1}$ für beliebiges $\alpha \in \mathbb{R} \setminus \{-1\}$

Sinus, Cosinus: $\sin' = \cos$, $\cos' = -\sin$

Ähnlich wie gewöhnliche Grenzwerte verträgt sich auch Differenzierbarkeit mit den Grundrechnungsarten, allerdings sind einige der resultierenden Formeln (Ableitungsregeln) etwas komplizierter. Unproblematisch ist die Linearität: $(f + g)' = f' + g'$ und $(cf)' = cf'$ für eine Konstante $c \in \mathbb{R}$. Man kann sie ablesen aus:

$$\begin{aligned} (f + g)(x_0 + \Delta x) &= f(x_0 + \Delta x) + g(x_0 + \Delta x) \approx \\ &\approx f(x_0) + f'(x_0)\Delta x + g(x_0) + g'(x_0)\Delta x = \\ &= (f + g)(x_0) + (f'(x_0) + g'(x_0))\Delta x \end{aligned}$$

und

$$(cf)(x_0 + \Delta x) = cf(x_0 + \Delta x) \approx c(f(x_0) + f'(x_0)\Delta x) = (cf)(x_0) + cf'(x_0)\Delta x.$$

Eine präzisere Formulierung lautet (**Linearität des Differenzierens**) :

Proposition 4.1.3.1. *Seien die Funktionen $f, g : D \rightarrow \mathbb{R}$ im inneren Punkt x_0 von D differenzierbar. Dann gilt:*

1. $f + g : D \rightarrow \mathbb{R}$ ist in x_0 differenzierbar, und es gilt $(f + g)'(x_0) = f'(x_0) + g'(x_0)$. (**Summenregel**)
2. Für beliebiges $c \in \mathbb{R}$ ist $cf : D \rightarrow \mathbb{R}$ in x_0 differenzierbar, und es gilt $(cf)'(x_0) = cf'(x_0)$. (**Homogenität der Ableitung**)

Beweis. 1. Man berechnet direkt den Differentialquotienten:

$$\begin{aligned} (f + g)'(x_0) &= \lim_{x \rightarrow x_0} \frac{(f + g)(x) - (f + g)(x_0)}{x - x_0} = \\ &= \lim_{x \rightarrow x_0} \frac{f(x) + g(x) - f(x_0) - g(x_0)}{x - x_0} = \\ &= \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} + \lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} = \\ &= f'(x_0) + g'(x_0) \end{aligned}$$

2. Analog:

$$\begin{aligned} (cf)'(x_0) &= \lim_{x \rightarrow x_0} \frac{(cf)(x) - (cf)(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{cf(x) - cf(x_0)}{x - x_0} = \\ &= c \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \\ &= cf'(x_0) \end{aligned}$$

□

Differentiation, aufgefasst als Funktion $f \mapsto f'$, ist also selbst linear. Damit gilt natürlich auch die entsprechende Aussage für Differenzen, insbesondere $(f - g)' = f' - g'$, weil sich ja $f - g = f + (-1)g$ als eine iterierte Anwendung von Addition und Multiplikation mit der Konstanten -1 schreiben lässt. Nur unwesentlich komplizierter ist die **Produktregel**. Die Δ -Schreibweise bringt uns auf die richtige Spur:

$$\begin{aligned} fg(x_0 + \Delta x) &= f(x_0 + \Delta x)g(x_0 + \Delta x) \approx \\ &\approx (f(x_0) + f'(x_0)\Delta x)(g(x_0) + g'(x_0)\Delta x) = \\ &= fg(x_0) + (f'(x_0)g(x_0) + f(x_0)g'(x_0))\Delta x + f'(x_0)g'(x_0)(\Delta x)^2. \end{aligned}$$

Weil Δx ja für $\Delta x \rightarrow 0$ interessiert, ist $(\Delta x)^2$ viel kleiner als Δx , kann also vernachlässigt werden. Somit deutet obige approximative Gleichung hin auf $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$. Tatsächlich gilt:

Proposition 4.1.3.2. *Seien die Funktionen $f, g : D \rightarrow \mathbb{R}$ im inneren Punkt x_0 von D differenzierbar. Dann ist auch $fg : D \rightarrow \mathbb{R}$ in x_0 differenzierbar, und es gilt*

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

Sind f und g auf ganz D differenzierbar, dann auch fg , und es gilt

$$(fg)' = f'g + fg'.$$

Beweis. Im Zähler des Differentialquotienten formt man (ähnlich wie bei der entsprechenden Regel für den Grenzwert einer Produktfolge)

$$\begin{aligned} f(x)g(x) - f(x_0)g(x_0) &= f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0) = \\ &= (f(x) - f(x_0))g(x) + f(x_0)(g(x) - g(x_0)) \end{aligned}$$

um. Daraus folgt

$$\begin{aligned} (fg)'(x_0) &= \lim_{x \rightarrow x_0} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} = \\ &= \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} g(x) + \lim_{x \rightarrow x_0} f(x_0) \frac{g(x) - g(x_0)}{x - x_0} = \\ &= f'(x_0)g(x_0) + f(x_0)g'(x_0). \end{aligned}$$

Man beachte, dass wegen der Differenzierbarkeit von g in x_0 und Satz 4.1.2.2 auch die in obiger Rechnung verwendete Stetigkeitsbeziehung $\lim_{x \rightarrow x_0} g(x) = g(x_0)$ gilt. \square

Aus den bisherigen Ableitungsregeln schließen wir: Alle Polynomfunktionen sind, weil sie sich durch Addition, Subtraktion und Multiplikation aus den differenzierbaren Funktionen $x \mapsto x$ und $x \mapsto c$ (Konstante mit $c \in \mathbb{R}$) ergeben, differenzierbar. Zur Berechnung der Ableitungen beginnen wir mit den Potenzfunktionen $x \mapsto x^n$. Aus der Produktregel ergibt sich die Rekursionsformel $(x^{n+1})' = (x^n x)' = (x^n)'x + x^n x' = (x^n)'x + x^n$. Einsetzen für $n = 1, 2, \dots$ liefert $(x^2)' = (x^1)'x + x^1 = 2x$, $(x^3)' = (x^2)'x + x^2 = 2xx + x^2 = 3x^2$ und allgemein, mittels Induktion, $(x^n)' = nx^{n-1}$.

Übungsaufgabe 218. (T) *Führen Sie diesen Induktionsbeweis vollständig durch.*

Für beliebige Polynome folgt dann aus der Linearität der Differentiation die Ableitungsregel

$$\left(\sum_{i=0}^n a_i x^i \right)' = \sum_{i=0}^n i a_i x^{i-1} = \sum_{i=0}^{n-1} (i+1) a_{i+1} x^i.$$

Besonders wichtig und auch charakteristisch ist die Ableitungsregel für die Verkettung differenzierbarer Funktionen. Wieder führt die approximative Schreibweise zur richtigen Vermutung. Denn

$$\begin{aligned} (g \circ f)(x_0 + \Delta x) &= g(f(x_0 + \Delta x)) \approx g(f(x_0) + f'(x_0)\Delta x) \approx \\ &\approx g(f(x_0)) + g'(f(x_0))f'(x_0)\Delta x \end{aligned}$$

suggeriert $(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$ (in diesem Zusammenhang heißen $f'(x_0)$ **innere**, $g'(f(x_0))$ **äußere Ableitung** der Verkettung), und tatsächlich lautet die **Kettenregel**:

Proposition 4.1.3.3. *Sei die Funktion $f : D_f \rightarrow \mathbb{R}$ im inneren Punkt x_0 von D_f differenzierbar, ebenso die Funktionen $g : D_g \rightarrow \mathbb{R}$ im inneren Punkt $y_0 := f(x_0)$ von D_g . Weiters sei $f(D_f) \subseteq D_g$, so dass die Verkettung $g \circ f : D_f \rightarrow \mathbb{R}$, $x \mapsto g(f(x))$, definiert ist. Dann ist $g \circ f$ in x_0 differenzierbar, und es gilt*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Sind f und g auf ganz D_f bzw. D_g differenzierbar, so also auch $g \circ f$ auf D_f , und es gilt $(g \circ f)' = f' \cdot (g' \circ f)$.

Beweis. Wir schreiben die Voraussetzungen an in der Form

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + r_f(x), \quad \text{und} \\ g(y) &= g(y_0) + g'(y_0)(y - y_0) + r_g(y) \end{aligned}$$

mit $\lim_{x \rightarrow x_0} \frac{r_f(x)}{x - x_0} = 0$ und $\lim_{y \rightarrow y_0} \frac{r_g(y)}{y - y_0} = 0$. Zu zeigen haben wir

$$(g \circ f)(x) = g(f(x_0)) + f'(x_0)g'(f(x_0))(x - x_0) + r(x)$$

mit $\lim_{x \rightarrow x_0} \frac{r(x)}{x - x_0} = 0$. Wir schreiben $y = f(x)$, setzen

$$\begin{aligned} (g \circ f)(x) - g(f(x_0)) &= g(y) - g(y_0) = g'(y_0)(y - y_0) + r_g(y) = \\ &= \left(g'(y_0) \frac{y - y_0}{x - x_0} + \frac{r_g(y)}{y - y_0} \cdot \frac{y - y_0}{x - x_0} \right) (x - x_0) \end{aligned}$$

ein in

$$r(x) = (g \circ f)(x) - g(f(x_0)) - f'(x_0)g'(f(x_0))(x - x_0)$$

und erhalten

$$\frac{r(x)}{x - x_0} = g'(y_0) \frac{y - y_0}{x - x_0} + \frac{r_g(y)}{y - y_0} \cdot \frac{y - y_0}{x - x_0} - f'(x_0)g'(f(x_0)).$$

Weil f als differenzierbare Funktion stetig ist, folgt aus $x \rightarrow x_0$ auch $y = f(x) \rightarrow y_0 := f(x_0)$. Somit schließen wir aus $\lim_{x \rightarrow x_0} \frac{y - y_0}{x - x_0} = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$ und $\lim_{y \rightarrow y_0} \frac{r_g(y)}{y - y_0} = 0$ (siehe weiter oben) tatsächlich

$$\lim_{x \rightarrow x_0} \frac{r(x)}{x - x_0} = g'(y_0)f'(x_0) + 0 \cdot f'(x_0) - f'(x_0)g'(f(x_0)) = 0.$$

□

Übungsaufgabe 219. (E) Sei $f : D \rightarrow D$ differenzierbar mit (als bekannt anzunehmender) Ableitung f' . Wir betrachten die Iterationen (Achtung, an dieser Stelle nicht mit den höheren Ableitungen von f verwechseln!) $f^{(n)}$, die rekursiv definiert sind durch $f^{(0)} := Id_D$ (Identität auf D) und $f^{(n+1)} := f \circ f^{(n)}$.

1. Begründen Sie, warum alle $f^{(n)}$ differenzierbar sind. Hinweis: Vollständige Induktion.
2. Stellen Sie eine Formel für die Ableitung von $f^{(n)}$ auf und beweisen Sie diese mittels Induktion nach n .
3. Angenommen $x_0 \in D$ ist ein Fixpunkt von f . Wie vereinfacht sich dann die Formel aus Teil 2 an der Stelle x_0 ?

Wenden wir die Kettenregel auf die Funktion $g \circ f$ mit $g : y \mapsto \frac{1}{y}$ (die Ableitung $g'(y) = -\frac{1}{y^2}$ von g kennen wir schon) an, so erhalten wir

$$\left(\frac{1}{f}\right)'(x_0) = (g \circ f)'(x_0) = g'(f(x_0))f'(x_0) = -\frac{f'(x_0)}{f(x_0)^2}.$$

In Kombination mit der Produktregel wenden wir diese Ableitungsregel auf Quotienten an:

$$\begin{aligned} \left(\frac{f}{g}\right)'(x_0) &= \left(f \cdot \frac{1}{g}\right)'(x_0) = f'(x_0)\frac{1}{g(x_0)} + f(x_0)\left(\frac{1}{g}\right)'(x_0) = \\ &= \frac{f'(x_0)}{g(x_0)} - \frac{f(x_0)g'(x_0)}{g(x_0)^2} = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}. \end{aligned}$$

Hieraus liest man die wohlbekannte **Quotientenregel**

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$

ab, genauer:

Proposition 4.1.3.4. Seien die Funktionen $f, g : D \rightarrow \mathbb{R}$ im inneren Punkt x_0 von D differenzierbar, außerdem $g(x_0) \neq 0$. Dann ist $\frac{f}{g}$ in einer geeigneten Umgebung von x_0 definiert und in x_0 differenzierbar mit:

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

In Kurzschreibweise: $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$.

Übungsaufgabe 220. (T) Beweisen Sie für $f = f_1 \cdot f_2 \cdot \dots \cdot f_n$ die Ableitungsregel $f' = f \cdot \sum_{i=1}^n \frac{f'_i}{f_i}$ mittels Induktion nach n .

Welche Funktionen wir auch immer mittels Grundrechnungsarten (Addition, Subtraktion, Multiplikation und Division) sowie durch Verkettung aufbauen: Mit unseren Differentiationsregeln können wir alle Ableitungen bilden, wenn wir sie für die Ausgangsfunktionen kennen. Bisher sind wir von den konstanten Funktionen und der Funktion $x \mapsto x$ ausgegangen und auf diesem Wege zu den Polynomfunktionen gekommen. Dafür haben Summenregel, Homogenität

der Ableitung und Produktregel ausreicht. Mit der Quotientenregel können wir nun auch gebrochen rationale Funktionen behandeln. Weil Kompositionen gebrochen rationaler Funktionen wieder gebrochen rational sind, bringt die Kettenregel an dieser Stelle nichts Neues. Sehr wohl von neuer Qualität (wenngleich man Zusammenhänge mit der Kettenregel ausmachen kann) ist die Ableitung von Umkehrfunktionen.

4.1.4 Die Ableitung von Umkehrfunktionen, insbesondere von Wurzeln

Inhalt in Kurzfassung: Die Ableitung einer Umkehrfunktion gewinnt man, grob gesprochen, als Kehrwert der Ableitung der ursprünglichen Funktion. Diese Einsicht hat weitreichende Anwendungen (auf Exponential-, Potenz- u.a. Funktionen) und lässt sich anschaulich (Spiegelung des Graphen um die erste Mediane) ebenso leicht plausibel machen wie formal (Kehrwert des Differentialquotienten bilden).

Wir wollen nun annehmen, $f : D \rightarrow f(D) \subseteq \mathbb{R}$ sei in x differenzierbar und habe eine Umkehrfunktion $f^{(-1)} : f(D) \rightarrow D$. Man ist versucht, die Gleichung $(f^{(-1)} \circ f)(x) = x$ unter Verwendung der Kettenregel auf beiden Seiten abzuleiten und auf $(f^{(-1)})'(f(x))f'(x) = 1$, also $(f^{(-1)})'(f(x)) = \frac{1}{f'(x)}$ bzw., wenn man $y = f(x)$ und $x = f^{(-1)}(y)$ setzt, auf $(f^{(-1)})'(y) = \frac{1}{f'(f^{(-1)}(y))}$ zu schließen.

Tatsächlich gilt diese Formel auch. Allerdings haben wir, indem wir die Kettenregel verwendet haben, vorausgesetzt, dass $f^{(-1)}$ differenzierbar ist, wenigstens im Punkt $f(x)$. Das wollen wir aber aus der Differenzierbarkeit von f erst herleiten. Die Überlegungen dazu lauten wie folgt.

Wir setzen voraus, dass f in x_0 differenzierbar ist mit $f'(x_0) \neq 0$ und in einer Umgebung von x_0 stetig und umkehrbar. Wir wollen uns überlegen, dass unter diesen Voraussetzungen die Umkehrfunktion $f^{(-1)}$ in $f(x_0)$ differenzierbar ist, nach obigen Überlegungen dann notwendig mit $(f^{(-1)})'(f(x_0)) = \frac{1}{f'(x_0)}$. Wir wissen bereits aus Satz 3.1.5.2, dass $f^{(-1)}$ wieder stetig ist (wenigstens in einer geeigneten Umgebung von $f(x_0)$). Somit sind für $y = f(x)$ und $y_0 = f(x_0)$ (also auch $x = f^{(-1)}(y)$ und $x_0 = f^{(-1)}(y_0)$) die Bedingungen $x \rightarrow x_0$ und $y \rightarrow y_0$ äquivalent. Deshalb gilt

$$(f^{(-1)})'(y_0) = \lim_{y \rightarrow y_0} \frac{f^{(-1)}(y) - f^{(-1)}(y_0)}{y - y_0} = \lim_{x \rightarrow x_0} \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{f'(x_0)},$$

also:

Satz 4.1.4.1. *Sei die reelle Funktion $f : D \rightarrow f(D)$ in $x_0 \in D \subseteq \mathbb{R}$ differenzierbar mit $f'(x) \neq 0$, außerdem in einer Umgebung U von x_0 stetig und umkehrbar. Dann ist die (lokale) Umkehrfunktion $f^{(-1)} : f(U) \rightarrow U$ in $f(x_0)$ differenzierbar, und es gilt $(f^{(-1)})'(f(x_0)) = \frac{1}{f'(x_0)}$. Ist D offen (meist ist D auch zusammenhängend) und f auf ganz D differenzierbar und umkehrbar mit $f'(x) \neq 0$ für alle $x \in D$, so gilt $(f^{(-1)})' = \frac{1}{f' \circ f^{(-1)}}$.*

Wir können diese Regel auf die Potenzfunktionen $f(x) := x^n$ für $n = 1, 2, \dots$ mit $D := \mathbb{R}^+$ anwenden. Offenbar sind die Voraussetzungen erfüllt. Die Umkehrfunktionen sind die Wurzelfunktionen $p_{\frac{1}{n}} : x \mapsto \sqrt[n]{x}$. Dann haben wir $f'(x) = nx^{n-1}$ zu setzen. In der Formel für die Ableitung bekommen wir mit $y = f(x) = x^n$ und $x = \sqrt[n]{y}$ die Regel $(y^{\frac{1}{n}})' = (\sqrt[n]{y})' = \frac{1}{n(\sqrt[n]{y})^{n-1}} = \frac{y^{\frac{1}{n}-1}}{n}$. Einprägsamer ist sie wohl in der Gestalt

$$(x^\alpha)' = \alpha x^{\alpha-1}$$

mit $\alpha = \frac{1}{n}$. Wir werden später sehen, dass diese Regel sogar für beliebige $\alpha \in \mathbb{R}$ gilt.

Übungsaufgabe 221. (T) Berechnen Sie für $f(x) = \sqrt{x + \sqrt{x + \sqrt{x}}}$ die erste Ableitung nach den bekannten Ableitungsregeln und erklären Sie deren Verwendung. Wo ist die erste Ableitung jeweils definiert?

Übungsaufgabe 222. (T) Die Schwingungsdauer τ eines Pendels mit der Länge l berechnet sich zu $\tau = 2\pi\sqrt{\frac{l}{g}}$. Dabei bezeichnet g die Fallbeschleunigung (welche in Wien, wenn Sie am Ufer der Donau etwa 160m über N.N stehen, ca. 9.7888 m/s^2 beträgt). Auf wie viel % genau kann man die Schwingungsdauer τ angeben, wenn man die Pendellänge l auf 0.1% genau bestimmt?

Hinweis: Linearisieren Sie die Funktion $\tau(l)$.

4.2 Taylorapproximation und Potenzreihen

Geht man von Polynomen aus und erweitert man die endliche Summe, durch die sie dargestellt werden, zu unendlichen Reihen der Bauart

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

oder (bei Entwicklung um einen Punkt $x_0 \neq 0$)

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n,$$

so landet man beim Begriff der Potenzreihe. Im Fall der Konvergenz zeigt sich, dass die Koeffizienten a_n aufs Engste mit den höheren Ableitungen von f zusammenhängen, weshalb Potenzreihen auch als sogenannte Taylorreihen aufgefasst werden können. Ihre große Bedeutung erhalten sie dadurch, dass die meisten wichtigen Funktionen sich als Potenzreihen darstellen lassen.

Wir beginnen mit Vorbereitungen, und zwar wird in 4.2.1 eine Beziehung zwischen Monotonie und Extrema von Funktionen mit deren Ableitung in eine Beziehung gesetzt. Diese spielt eine wesentliche Rolle auch bei dem für die

weitere Theorie grundlegenden Mittelwertsatz (4.2.2), dessen Fortsetzung unter Zuhilfenahme höherer Ableitungen (4.2.3) der Satz von Taylor (4.2.4) ist. Von einer anderen Seite kommend, nämlich auf der Suche nach einer Funktion $\neq 0$, die ihre eigene Ableitung ist, stößt man auf die Exponentialreihe und die Eulersche Zahl e (4.2.5). Sie ist das archetypische Beispiel für Potenzreihen, die anschließend allgemeiner untersucht werden, vor allem in Hinblick auf ihren Konvergenzbereich (4.2.6) und ihre äußerst angenehmen Differenzierbarkeitseigenschaften (4.2.7). Unter dem Gesichtspunkt von Potenzreihen wird auch die bei Prüflingen (mehr als bei Prüfern) beliebte Regel von de l'Hospital (4.2.8) besser verständlich.

4.2.1 Monotonie und erste Ableitung, Anfang

Inhalt in Kurzfassung: Das Vorzeichen der Ableitung f' spiegelt Monotonieverhalten wieder. Im Zentrum stehen jene Überlegungen, die zeigen, dass eine lokale Extremstelle x_0 einer differenzierbaren Funktion f im Inneren des Definitionsbereichs stets $f'(x_0) = 0$ erfüllt. Verfeinerte Überlegungen werden in 4.4.1 folgen.

Angenommen $f : D \rightarrow \mathbb{R}$ ist in $x_0 \in D$ (innerer Punkt von D) differenzierbar, d.h. der Grenzwert

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Nach Definition des Grenzwertes gibt es dann zu jedem $\varepsilon > 0$ ein $\delta > 0$ derart, dass

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| < \varepsilon$$

für alle $x \in (x_0 - \delta, x_0 + \delta)$. Ist $f'(x_0) \neq 0$ können wir $\varepsilon > 0$ so klein wählen, dass der Differenzenquotient $\frac{f(x) - f(x_0)}{x - x_0}$ für alle $x \in (x_0 - \delta, x_0 + \delta)$ dasselbe Vorzeichen wie $f'(x_0)$ hat. Im Fall $f'(x_0) > 0$ bedeutet das $f(x) > f(x_0)$ für alle $x \in (x_0, x_0 + \delta)$ und $f(x) < f(x_0)$ für alle $x \in (x_0 - \delta, x_0)$. In diesem Sinne ist also f im Punkt x_0 strikt monoton wachsend, entsprechend fallend für $f'(x_0) < 0$. An der Stelle x_0 liegt somit sicher kein Extremum von f , sofern x_0 innerer Punkt des Definitionsbereichs ist und $f'(x_0) \neq 0$. Man beachte, dass dies nicht unbedingt bedeutet, dass f in einer ganzen Umgebung um x_0 monoton ist. Später wird es uns nicht schwer fallen, Beispiele dazu zu konstruieren. (Diese müssen unstetige Ableitungen haben.) Unsere Überlegungen zeigen aber:

Satz 4.2.1.1. *Hat die reelle Funktion $f : D \rightarrow \mathbb{R}$ in einem inneren Punkt x_0 von D ein lokales Extremum, und ist f differenzierbar in x_0 , so gilt $f'(x_0) = 0$.*

Die notwendige Bedingung $f'(x_0) = 0$ ist jedoch nicht hinreichend, wie das Beispiel $f(x) := x^3$ im Punkt $x_0 = 0$ zeigt. Wegen $f'(x) = 3x^2$ ist $f'(x_0) = f'(0) = 0$, in x_0 liegt aber weder ein lokales Minimum noch ein lokales Maximum von f .

Außerdem beachte man, dass an Randpunkten x_0 des Definitionsbereichs Extrema sehr wohl auch dann auftreten können, wenn für die (dann nur einseitige) Ableitung $f'(x_0) \neq 0$ gilt. Ein sehr einfaches Beispiel ist die Funktion $f(x) := x$ auf $D := [0, 1]$. Im Randpunkt $x_0 = 0$ nimmt f das Minimum an, im Randpunkt $x_0 = 1$ das Maximum, während $f'(0) = f'(1) = 1 \neq 0$.

Die hier begonnenen Untersuchungen der Monotonie mit Hilfe der Differentialrechnung werden wir (unter Zuhilfenahme des Mittelwertsatzes und seiner Folgerungen) in 4.4.1 fortsetzen.

4.2.2 Der Mittelwertsatz

Inhalt in Kurzfassung: Der Mittelwertsatz taucht beim Aufbau der weiteren Theorie ständig auf, auch in Situationen, wo man es auf den ersten Blick nicht vermutet. Seine Aussage ist anschaulich sehr überzeugend: Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und im Inneren von $[a, b]$ differenzierbar. Zieht man die Verbindungslinie zwischen den Endpunkten Punkten $(a, f(a))$ und $(b, f(b))$ des Graphen von f , so gibt es an einem geeigneten Punkt $(x_0, f(x_0))$ eine zu dieser Linie parallele Tangente.

Der für die Entwicklung der Theorie grundlegende **Mittelwertsatz der Differentialrechnung** lautet:

Satz 4.2.2.1. Für $a < b \in \mathbb{R}$ sei die Funktion $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und im Inneren (a, b) differenzierbar. Dann gibt es ein $x_0 \in (a, b)$ mit $f'(x_0) = \frac{f(b)-f(a)}{b-a}$ bzw., äquivalent, mit $f(b) = f(a) + f'(x_0)(b - a)$.

Anschaulich bedeutet das: Zieht man die Verbindungslinie zwischen den Punkten $(a, f(a))$ und $(b, f(b))$, so gibt es zwischen a und b einen Punkt, wo die Tangente an f parallel zu dieser Linie ist. Für einen strengen Beweis betrachtet man zuerst den Spezialfall $f(a) = f(b)$, auf den sich der **Satz von Rolle** bezieht:

Proposition 4.2.2.2. Für $a < b \in \mathbb{R}$ sei die Funktion $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und im Inneren (a, b) sogar differenzierbar. Gilt außerdem $f(a) = f(b)$, so gibt es ein $x_0 \in (a, b)$ mit $f'(x_0) = 0$.

Beweis. Ist f eine konstante Funktion, so ist $f' = 0$ die Nullfunktion, und wir können jedes $x_0 \in (a, b)$ nehmen. Interessanter ist der Fall, dass f nicht konstant ist. Als stetige Funktion nimmt f auf $[a, b]$ ein Minimum und ein Maximum an (Satz 3.1.5.4). Wenigstens eines der beiden ist von $f(a) = f(b)$ verschieden. Es gibt also eine Stelle $x_0 \in [a, b]$ eines lokalen Extremums mit $x_0 \neq a, b$. Folglich liegt x_0 im Inneren (a, b) , wo f differenzierbar ist. An so einer Stelle muss, wie wir aus Satz 4.2.1.1 wissen, $f'(x_0) = 0$ gelten. \square

Den Mittelwertsatz beweist man, indem man ihn auf den Satz von Rolle zurückführt. Dazu betrachten wir die Funktion

$$g(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Man rechnet unmittelbar $g(b) = f(a) = g(a)$ nach. Außerdem erbt g von f die Stetigkeit auf $[a, b]$ und die Differenzierbarkeit auf (a, b) , erfüllt also die Voraussetzungen des Satzes von Rolle. Folglich gibt es ein $x_0 \in (a, b)$ mit $g'(x_0) = 0$. Wegen $f(x) = g(x) + \frac{f(b)-f(a)}{b-a}(x-a)$ folgt daraus $f'(x_0) = g'(x_0) + \frac{f(b)-f(a)}{b-a} = \frac{f(b)-f(a)}{b-a}$, wie im Mittelwertsatz 4.2.2.1 behauptet.

Als interessante Folgerung erhalten wir, dass Stammfunktionen auf zusammenhängenden Teilmengen von \mathbb{R} (Intervallen) bis auf eine additive Konstante eindeutig bestimmt sind.

Satz 4.2.2.3. *Sei $D \subseteq \mathbb{R}$ zusammenhängend, und seien F_1 und F_2 Stammfunktionen von $f : D \rightarrow \mathbb{R}$, so gibt es ein $c \in \mathbb{R}$ mit $F_2 = F_1 + c$, d.h. $F_2(x) = F_1(x) + c$ für alle $x \in D$.*

Beweis. Die Differenz $F := F_2 - F_1$ hat die Ableitung $F' = (F_2 - F_1)' = F_2' - F_1' = f - f = 0$. Nach dem Mittelwertsatz 4.2.2.1 gibt es zu je zwei Punkten $a < b$ in D eine Zwischenstelle $x \in (a, b) \subseteq D$ (D ist zusammenhängend!) mit $F(b) = F(a) + F'(x)(b-a)$, wegen $F' = 0$ also $F(b) = F(a)$. Folglich ist F konstant. Mit dem Wert $c := F(a) = F(b)$ von F (für beliebige $a, b \in D$) gilt dann die Behauptung. \square

Übungsaufgabe 223. (P) *Beweisen Sie die Ungleichung*

$$|\sin b - \sin a| \leq |b - a|$$

für alle reellen Zahlen a und b indem Sie den Mittelwertsatz auf die Funktion $\sin x$ anwenden. Sie dürfen dabei schon verwenden, dass $\sin' = \cos$.

Ganz ähnlich wie den Mittelwertsatz beweist man den **Zwischenwertsatz für Ableitungen**:

Satz 4.2.2.4. *Ist f auf $[a, b]$ differenzierbar, so nimmt f' jeden Wert zwischen $f'(a)$ und $f'(b)$ an.*

Ist die Ableitung f' stetig, so ist diese Aussage bereits durch den Zwischenwertsatz 3.1.5.1 für stetige Funktionen garantiert. Der Zwischenwertsatz für Ableitungen lässt sich aber auch ohne diese Voraussetzung beweisen. Die Idee ist die folgende: Ist $f'(a) < \alpha < f'(b)$, so betrachtet man die Funktion $g(x) := f(x) - \alpha x$. Als stetige Funktion muss sie auf $[a, b]$ ein Extremum annehmen, das (warum?) an einer Stelle x_0 im Inneren von $[a, b]$ liegt. Dort muss $f'(x_0) = \alpha$ gelten.

Übungsaufgabe 224. (E) *Arbeiten Sie diese Ideen zu einem vollständigen Beweis von Satz 4.2.2.4 aus.*

Die Arbeit in dieser Übungsaufgabe rührte daher, dass wir auch ohne die Stetigkeitsvoraussetzung an f' durchkommen wollten. Wir wollen uns nun mit den Möglichkeiten beschäftigen, die sich eröffnen, wenn man immer stärkere Voraussetzungen in diese Richtung zur Verfügung hat.

4.2.3 Stetige Differenzierbarkeit und höhere Ableitungen

Inhalt in Kurzfassung: Ist eine Funktion f differenzierbar, so kann man sich fragen, ob ihre Ableitung selbst auch wieder interessante Eigenschaften wie Stetigkeit oder gar Differenzierbarkeit aufweist. Im ersten Fall nennt man f stetig differenzierbar, im zweiten lässt sich die sogenannte zweite Ableitung $f'' := (f')'$ bilden. Iteration dieses Prozesses führt zu den höheren Ableitungen von f . Rekursive Definition: $f^{(0)} := f$ (0-te Ableitung), $f^{(n+1)} := (f^{(n)})'$ (Definition der $n + 1$ -ten Ableitung von f als Ableitung der n -ten Ableitung von f).

Wir wollen nun annehmen, dass die reelle Funktion $f : D \rightarrow \mathbb{R}$ differenzierbar ist, und ihre Ableitung $f' : D \rightarrow \mathbb{R}$ ins Auge fassen. Ist f' stetig, so heißt f **stetig differenzierbar**.

Es ist auf den ersten Blick gar nicht leicht, Beispiele von Funktionen zu finden, die differenzierbar, aber nicht stetig differenzierbar sind. Das Standardbeispiel ist die Funktion $f(x) := x^2 \sin \frac{1}{x}$ für $x \neq 0$ und $f(0) := 0$. Wir greifen vor und verwenden die Differenzierbarkeit des Sinus samt Ableitungsregel $\sin' = \cos$. Aus Produkt- und Kettenregel berechnet man damit $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$ für $x \neq 0$. Für $x \rightarrow 0$ konvergiert der erste Summand dieses Ausdrucks gegen 0, der zweite jedoch oszilliert zwischen den Werten -1 und 1 . Also existiert kein Grenzwert $\lim_{x \rightarrow 0} f'(x)$, obwohl f auch an der Stelle 0 differenzierbar ist mit Ableitung

$$f'(0) = \lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \rightarrow 0} \frac{f(x)}{x} = \lim_{x \rightarrow 0} x \sin \frac{1}{x} = 0.$$

Für stetig differenzierbares f stellt sich die Frage, ob sogar die Ableitung f' differenzierbar ist etc. Man definiert daher die höheren Ableitungen rekursiv durch $f^{(0)} := f$ und

$$f^{(n+1)} := (f^{(n)})' \quad (n + 1\text{-te Ableitung}),$$

sofern die n -te Ableitung von f existiert und selbst differenzierbar ist.¹ Die Funktion f heißt **n -mal differenzierbar**, wenn $f^{(n)}$ existiert; sie heißt **n -mal stetig differenzierbar**, wenn $f^{(n)}$ sogar stetig ist. Ist f sogar n -mal differenzierbar für jedes $n \in \mathbb{N}$, so heißt f **unendlich oft differenzierbar** (manchmal auch **glatt**).

Die 0-te Ableitung von f ist also f selbst, die erste ist die (gewöhnliche) Ableitung f' von f , die zweite Ableitung f'' ist die Ableitung der Ableitung etc.

Beispiele unendlich oft differenzierbarer Funktionen sind Polynome und gebrochen rationale Funktionen. Gehen wir von einem Polynom

$$f(x) = \sum_{k=0}^n a_k x^k$$

¹Trotz derselben Notation darf die n -te Ableitung nicht mit der n -fachen Iteration von f aus 2.1.8 verwechselt werden.

aus, so sehen wir: $f(0) = a_0$, $f'(0) = a_1$, $f''(0) = 2a_2$, $f'''(0) = 6a_3$ und allgemein $f^{(k)}(0) = k!a_k$ für $k = 0, 1, \dots, n$. Man beachte auch, dass $f^{(n)} = n!a_n$ konstant und $f^{(k)}$ für $k > n$ die Nullfunktion ist. Die Koeffizienten $a_k = \frac{f^{(k)}(0)}{k!}$ lassen sich also auch über die höheren Ableitungen von f an der Stelle 0 bestimmen. Analoges gilt bei Entwicklung von f an einer anderen Stelle x_0 :

$$f(x) = \sum_{k=0}^n b_k (x - x_0)^k \quad \text{mit} \quad b_k = \frac{f^{(k)}(x_0)}{k!}.$$

Diesen Ansatz wollen wir nun für beliebige hinreichend oft differenzierbare Funktionen f , die keine Polynome sein müssen, weiter verfolgen.

Zuvor aber noch eine Übungsaufgabe:

Übungsaufgabe 225. (E) Für $m, n \in \mathbb{N}$ und $x \neq 0$ sei $f_{m,n}(x) := x^m \sin\left(\frac{1}{x^n}\right)$, außerdem $f_{m,n}(0) := 0$.

1. Begründen Sie, warum alle $f_{m,n}$ auf $\mathbb{R} \setminus \{0\}$ unendlich oft differenzierbar (und folglich auch stetig) sind und berechnen Sie $f'_{m,n}(x)$ für $x \neq 0$.
2. Für welche Paare (m, n) ist f stetig?
3. Für welche Paare (m, n) ist f differenzierbar?
4. Für welche Paare (m, n) ist f stetig differenzierbar?
5. Gegeben irgendein Paar (m, n) . Wie oft ist $f_{m,n}$ differenzierbar, wie oft stetig differenzierbar?

Sie dürfen in dieser Aufgabe bereits die Ableitungsregeln $\sin' = \cos$ und $\cos' = -\sin$ verwenden.

4.2.4 Der Satz von Taylor

Inhalt in Kurzfassung: So wie die Differenzierbarkeit einer Funktion f an einer Stelle x_0 in der Nähe von x_0 eine gute Approximation von f durch eine lineare Funktionen ermöglicht, so erlaubt n -fache Differenzierbarkeit eine noch viel bessere Approximation von f durch jenes Polynom n -ten Grades, das in x_0 in seinen Ableitungen mit f übereinstimmt. Man nennt dieses Polynom das Taylorpolynom n -ten Grades von f in x_0 . Der Satz von Taylor macht eine Aussage über die Approximationsgüte, die von der Größe der $n + 1$ -ten Ableitung abhängt. Technisch gesehen ist der Satz von Taylor die natürliche Verallgemeinerung des Mittelwertsatzes auf höhere Ableitungen. Funktioniert Taylorapproximation für alle n , so ergibt sich im Grenzwert eine Approximation von f durch eine unendliche Reihe, die sogenannte Taylorreihe. In vielen interessanten Fällen stimmt, wie wir später sehen werden, die Taylorreihe sogar exakt mit f überein.

Definition 4.2.4.1. Ist die n -mal differenzierbare reelle Funktion $f : D \rightarrow \mathbb{R}$ vorgegeben und $x_0 \in D$, so heißt

$$p_{f,x_0,n}(x) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

das **Taylorpolynom n -ten Grades** oder auch das **n -te Taylorpolynom** von f an der Stelle x_0 .

Offenbar ist das Polynom $p_{f,x_0,n}$ so definiert, dass an der Stelle x_0 seine 0-te Ableitung (= sein Funktionswert), seine erste Ableitung und seine höheren Ableitung bis zur n -ten mit jenen von f übereinstimmen, also $p_{f,x_0,n}^{(k)}(x_0) = f^{(k)}(x_0)$ für $k = 0, 1, \dots, n$. Die erste Ableitung einer Funktion entspricht ihrer linearen Approximation, während das Taylorpolynom eine Approximation höheren Grades ist. Das berechtigt zur Hoffnung, dass $p_{f,x_0,n}$ für wachsende n die Funktion f immer besser approximiert. Bedenkt man, dass $p_{f,x_0,n}^{(n+1)}$ die Nullfunktion ist, so darf man erwarten, dass die Differenz $f - p_{f,x_0,n}$ davon abhängt, wie stark f von dieser Eigenschaft abweicht, sprich von der Größe der $n+1$ -ten Ableitung $f^{(n+1)}$. Betrachten wir zunächst den sehr einfachen Fall $n = 0$. Das 0-te Taylorpolynom $p_{f,x_0,0}$ ist die konstante Funktion mit dem Wert $f(x_0)$. Der Mittelwertsatz besagt, dass die Abweichung $f(x) - p_0(x) = f(x) - f(x_0)$ gleich $f'(\xi)(x - x_0)$ mit einer Zwischenstelle ξ ist. Das ist tatsächlich eine Aussage von der vermuteten Art. Von derselben Struktur, lediglich mit höheren Ableitungen, ist der **Satz von Taylor**, der den Mittelwertsatz als Spezialfall mit $n = 0$ enthält.

Satz 4.2.4.2. Die reelle Funktion $f : D \rightarrow \mathbb{R}$ sei stetig auf dem abgeschlossenen Intervall $[x_0, x]$, $x_0 < x \in D$, und $n + 1$ -mal differenzierbar auf dem offenen Intervall (x_0, x) . Wir bezeichnen mit

$$r_{f,x_0,n}(x) := f(x) - p_{f,x_0,n}(x)$$

den Fehler (genannt auch **Restglied**), den man an der Stelle $x \neq x_0 \in D$ macht, wenn man f durch das n -te Taylorpolynom $p_{f,x_0,n}$ von f an der Stelle x_0 ersetzt.

Dann gibt es dann eine Stelle ξ zwischen x_0 und x mit

$$r_{f,x_0,n}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}.$$

(Für $x < x_0$ gilt dieselbe Aussage, wenn man die Intervalle $[x, x_0]$ und (x, x_0) statt $[x_0, x]$ bzw. (x_0, x) verwendet.)

Beweis. Wir bestimmen die Zahl λ so, dass für die Funktion

$$F : t \mapsto p_{f,x_0,n}(t) + \lambda(t - x_0)^{n+1}$$

die Beziehung $F(x) = f(x)$ gilt, also $\lambda = \frac{f(x) - p_{f,x_0,n}(x)}{(x - x_0)^{n+1}}$, und betrachten die Differenz

$$g(t) := F(t) - f(t).$$

Die Funktion g ist stetig auf $[x_0, x]$ und $n + 1$ -mal differenzierbar auf (x_0, x) mit $g(x_0) = g'(x_0) = g''(x_0) = \dots = g^{(n)}(x_0) = 0$ und $g(x) = 0$. Nach dem Satz von Rolle angewendet auf g gibt es eine Zwischenstelle ξ_1 zwischen x_0 und x mit $g'(\xi_1) = 0$. Nun wenden wir den Satz von Rolle auf g' an und erhalten eine Zwischenstelle ξ_2 zwischen x_0 und ξ_1 mit $g''(\xi_2) = 0$. Auf diese Weise setzen wir fort und erhalten schließlich eine Zwischenstelle $\xi := \xi_{n+1}$ mit $g^{(n+1)}(\xi) = 0$. Man beachte, dass $g^{(n+1)}(t) = \lambda(n+1)! - f^{(n+1)}(t)$. Setzen wir ξ für t ein, so folgt $\lambda = \frac{f^{(n+1)}(\xi)}{(n+1)!}$ und somit

$$r_{f,x_0,n}(x) = f(x) - p_{f,x_0,n}(x) = F(x) - p_{f,x_0,n}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1},$$

die Behauptung. \square

Sind Schranken für die $n + 1$ -te Ableitung $f^{(n+1)}$ bekannt, so gibt der Satz von Taylor Auskunft darüber, wie gut die Funktion f durch ihr n -tes Taylorpolynom $p_{f,x_0,n}$ approximiert wird.

Besonders angenehm wäre es, wenn f unendlich oft differenzierbar wäre und wenn man zeigen könnte, dass der Fehlerterm

$$r_{f,x_0,n}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$$

für $n \rightarrow \infty$ gegen 0 konvergiert. Denn das bedeutete

$$f(x) = \lim_{n \rightarrow \infty} p_{f,x_0,n}(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n.$$

Die Reihe rechts nennt man die **Taylorreihe** von f an der **Entwicklungs-** oder **Anschlussstelle** x_0 .

Übungsaufgabe 226. (T) Berechnen Sie für folgende Funktionen das Taylorpolynom dritten Grades mit dem angegebenen Entwicklungspunkt x_0

1. $x^3 + 3x^2 + 5$, $x_0 = 0$.
2. $x^3 + 3x^2 + 5$, $x_0 = 1$.

Sehr häufig tritt die Entwicklungsstelle $x_0 = 0$ auf. In diesem Fall hat die Taylorreihe die Gestalt als sogenannte **Maclaurinreihe**

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

Wegen des stark wachsenden Nenners $n!$, der jedenfalls gegenüber dem Term $(x-x_0)^n$ dominiert, stehen die Chancen nicht schlecht, dass der Fehlerterm $r_{f,x_0,n}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$ für $n \rightarrow \infty$ gegen 0 strebt und somit die Funktion f durch ihre Taylorreihe dargestellt wird. Garantieren können wir das z.B., wenn die Ableitungen $f^{(n)}$ für $n \rightarrow \infty$ beschränkt bleiben. Dem einfachsten und gleichzeitig wichtigsten Beispiel, wo das der Fall ist, wollen wir uns nun zuwenden.

4.2.5 Die Exponentialreihe und die Eulersche Zahl e

Inhalt in Kurzfassung: Wachstumsprozesse beispielsweise, wo der Zuwachs proportional zum aktuellen Bestand einer Population ist, legen es nahe, nach einer Funktion f mit $f' = f$ zu suchen. Nimmt man $f(0) = 1$ an, so ergibt sich daraus eindeutig die Taylorreihe von f als Exponentialreihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. Mit diversen Konvergenzaussagen über Reihen überprüft man, dass diese Reihe nicht nur eine auf ganz \mathbb{C} definierte Funktion darstellt, sondern dass diese Funktion auf \mathbb{R} auch sämtliche Voraussetzungen des Eindeutigkeitssatzes für Exponentialreihen erfüllt. Folglich handelt es sich bei f um eine der Funktionen \exp_a , wobei sich als Basis $a = f(1)$ die Eulersche Zahl e ergibt. Somit ist die Exponentialreihe als Darstellung der Exponentialfunktion zur Basis e identifiziert und auch gleich eine natürliche Fortsetzung ins Komplexe gefunden. Ohne allzu große Schwierigkeiten überprüft man, dass auch tatsächlich $\exp' = \exp$ gilt, wie ursprünglich für die Funktion f angenommen.

Sehr häufig treten Wachstumsprozesse auf, wo der momentane Zuwachs proportional zum vorhandenen Bestand ist. Ein typisches Beispiel ist das Bevölkerungswachstum bei gleichbleibender Fortpflanzungsrate pro Individuum. Bezeichnen wir mit $f(t)$ die Bevölkerung zum Zeitpunkt t , so sollte also annähernd $f'(t) = cf(t)$ mit einer Konstanten $c \in \mathbb{R}$ gelten. Herrscht positives Wachstum, so ist $c > 0$, und wir können die Einheiten so wählen, dass $c = 1$ und $f(0) = 1$ gilt. Gesucht ist also eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(0) = 1$, die ihre eigene Ableitung ist: $f' = f$.

Durch diese Vorgaben ist die Taylorreihe von f bei $x_0 = 0$ aber schon eindeutig bestimmt, weil ja dann $f = f' = f'' = \dots = f^{(n)} = \dots$, insbesondere $f^{(n)}(0) = f(0) = 1$ für alle $n \in \mathbb{N}$ gilt. Es muss dann gelten

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Die Reihe rechts ist die **Exponentialreihe**. Wir haben bereits bewiesen, dass sie für alle $x \in \mathbb{R}$ und sogar (mit dem gleichen Beweis mittels Quotientenkriterium) für alle $x \in \mathbb{C}$ absolut konvergiert. Also können wir bei Bedarf

$$f : \mathbb{C} \rightarrow \mathbb{C}, \quad x \mapsto \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

auch als Funktion auf ganz \mathbb{C} auffassen. Zunächst bleiben wir aber im Reellen.

Wegen der absoluten Konvergenz können wir das Produkt zweier Funktionswerte von f als Cauchyprodukt berechnen:

$$f(x)f(y) = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) \cdot \left(\sum_{n=0}^{\infty} \frac{y^n}{n!} \right) = \sum_{n=0}^{\infty} c_n$$

mit

$$c_n = \sum_{k=0}^n \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!} = \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} = \frac{(x+y)^n}{n!}$$

(Binomischer Lehrsatz, siehe 1.4.4). Also können wir zu

$$f(x+y) = f(x)f(y)$$

zusammenfassen. Die Funktion f erfüllt somit dieselbe Funktionalgleichung wie die Exponentialfunktionen aus 3.3.2. Nach dem Eindeutigkeitssatz 3.3.2.1 für die Exponentialfunktion muss $f = \exp = \exp_a$ für $a = \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!} = e$ gelten, sofern wir die Stetigkeit von f an wenigstens einem Punkt x_0 beweisen können. Wir fassen $x_0 = 0$ ins Auge und schreiben zu diesem Zweck

$$f(x) = 1 + x\left(1 + \frac{x}{2!} + \frac{x^2}{3!} + \dots\right) = 1 + x \sum_{n=0}^{\infty} \frac{x^n}{(n+1)!}.$$

Für $|x| \leq 1$ ist wegen $\left|\frac{x^n}{(n+1)!}\right| \leq \frac{1}{n!}$ die unendliche Reihe $\sum_{n=0}^{\infty} \frac{x^n}{(n+1)!}$ betragsmäßig beschränkt durch $\sum_{n=0}^{\infty} \frac{1}{n!} = e$, also folgt

$$\lim_{x \rightarrow 0} f(x) = 1 + \lim_{x \rightarrow 0} x \sum_{n=0}^{\infty} \frac{x^n}{(n+1)!} = 1 = f(0),$$

was die Stetigkeit von f bei $x_0 = 0$ beweist. Also ist tatsächlich

$$e^x = \exp_e(x) = \exp(x) = f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

für alle $x \in \mathbb{R}$. Für die Eulersche Zahl ergibt sich der Wert $e = 2,718281828459045 \dots$

Für \exp können wir jetzt sogar Differenzierbarkeit und die Ableitungsregel

$$\exp' = \exp \quad \text{d.h.} \quad (e^x)' = e^x$$

beweisen. Dazu setzen wir in der Definition des Differentialquotienten $h = x - x_0$, also

$$\exp'(x_0) = \lim_{x \rightarrow x_0} \frac{\exp(x) - \exp(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{\exp(x_0 + h) - \exp(x_0)}{h}.$$

Für den Zähler gilt wegen der Funktionalgleichung $\exp(x_0 + h) = \exp(x_0) \exp(h)$. Somit können wir weiter schreiben

$$\exp'(x_0) = \exp(x_0) \lim_{h \rightarrow 0} \frac{\exp(h) - 1}{h}.$$

Der letzte Ausdruck

$$\frac{\exp(h) - 1}{h} = \frac{1}{h} \left(h + \frac{h^2}{2} + \frac{h^3}{6} + \dots \right) = 1 + \frac{h}{2!} + \frac{h^2}{3!} + \dots = 1 + h \sum_{n=0}^{\infty} \frac{h^n}{(n+2)!}$$

konvergiert für $h \rightarrow 0$ gegen 1 (Argument wie weiter oben bei der Stetigkeit), also gilt wirklich $\exp'(x_0) = \exp(x_0)$.

Übungsaufgabe 227. (T) In der obigen Rechnung wurde verwendet, dass $\sum_{n=0}^{\infty} \frac{h^n}{(n+2)!}$ für h aus einem Intervall um 0 beschränkt ist. Begründen Sie das für $h \in [-1, 1]$.

Aus der Beziehung $a^x = e^{(\ln a)x}$ liefert die Kettenregel auch die Ableitung von Exponentialfunktionen mit einer Basis $a \neq e$, nämlich

$$(a^x)' = \left(e^{(\ln a)x} \right)' = \ln a e^{(\ln a)x} = (\ln a)a^x.$$

4.2.6 Verallgemeinerung: Potenzreihen und ihr Konvergenzbereich in \mathbb{C}

Inhalt in Kurzfassung: Lässt man in der Exponentialreihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ statt der Faktoren $\frac{1}{n!}$ beliebige Koeffizienten a_n und statt x^n Ausdrücke $(x - x_0)^n$ mit einem festen x_0 (der sogenannten Entwicklungs- oder Anschlussstelle) zu, so stößt man auf den Begriff der Potenzreihe. Die erste wichtige Einsicht: Der Konvergenzbereich hat (mit eventueller Ausnahme von Randpunkten) in \mathbb{C} stets Kreis- in \mathbb{R} folglich Intervallgestalt, woraus sich der Begriff des Konvergenzradius r ergibt. Auch die Extremfälle $r = 0$ (Konvergenz nur für $x = x_0$) und $r = \infty$ (Konvergenz überall, wie bei der Exponentialreihe) sind möglich. Meist interessieren wir uns nur für das Verhalten auf \mathbb{R} . Eine Ausnahme wird vor allem die Exponentialfunktion sein, wenn wir nochmals auf die trigonometrischen Funktionen zu sprechen kommen werden.

Zahlreiche Phänomene, die wir an der Exponentialreihe beobachten konnten, gelten viel allgemeiner. Der geeignete Rahmen sind sogenannte **Potenzreihen**, die wir zweckmäßig gleich als komplexe Funktionen auffassen. Es handelt sich um Reihen der Form

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n.$$

Dabei heißt x_0 die **Anschluss-** oder **Entwicklungsstelle**, die Zahlen $a_n \in \mathbb{C}$ heißen die **Koeffizienten** der Potenzreihe. Eine Potenzreihe ist also der (zunächst punktweise) Grenzwert der (stetigen) Polynome $\sum_{k=0}^n a_k (x - x_0)^k$ für $n \rightarrow \infty$.

Die wichtigste Frage ist die nach der Konvergenz. Vom Wurzelkriterium wissen wir, dass es darauf ankommt, ob ab einem gewissen n_0 , also für alle $n \geq n_0$

$$\sqrt[n]{|a_n (x - x_0)^n|} = |x - x_0| \sqrt[n]{|a_n|} \leq q < 1$$

garantiert werden kann. Die entscheidende Rolle dabei spielt der obere Grenzwert $s := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}$ der n -ten Wurzeln der Beträge der a_n . Gilt die Ungleichung $|x - x_0| < s^{-1}$, so herrscht Konvergenz, ist hingegen $|x - x_0| > s^{-1}$, so divergiert die Reihe. Der **Konvergenzbereich** K , das ist die Menge aller $x \in \mathbb{C}$, für die die Reihe konvergiert, hat also die Gestalt einer Kreisscheibe in der komplexen Ebene mit Mittelpunkt x_0 und Radius $r := s^{-1}$ bzw. einer zusammenhängenden Menge in \mathbb{R} . Man nennt r den **Konvergenzradius** der

Reihe. Die Punkte auf dem Kreisrand können zu K gehören oder auch nicht. Im Extremfall kann auch $r = \infty$ sein (wie zum Beispiel bei der Exponentialreihe, also bei $a_n = \frac{1}{n!}$) oder $r = 0$ (wie etwa bei $a_n = n!$). Im ersten Fall konvergiert die Reihe überall, im zweiten nur im Punkt x_0 . Einen positiven aber endlichen Konvergenzradius $r = 1$ hat die geometrische Reihe $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$, die also ebenfalls als Potenzreihe aufgefasst werden kann, und zwar mit den Koeffizienten $a_n = 1$.

Interessiert man sich nur für das Verhalten im Reellen, ist der Konvergenzbereich K einer Potenzreihe ein Intervall (Schnitte eines Kreises mit der reellen Achse), das sogenannte **Konvergenzintervall** der Reihe, wobei in den Extremfällen auch $K = \mathbb{R}$ oder $K = \{x_0\}$ möglich ist.

Übungsaufgabe 228. (T) Welche der Reihen lassen sich als Potenzreihen deuten? Geben Sie für jene, wo das der Fall ist, Entwicklungsstelle und Konvergenzbereich an.

$$1. \sum_{n=0}^{\infty} \left(\frac{x}{3} - 2\right)^n \qquad 2. \sum_{n=0}^{\infty} \left(x - \frac{1}{n}\right)^n$$

$$3. \sum_{n=0}^{\infty} \frac{(3x-1)^n}{n!}$$

Übungsaufgabe 229. (T) Berechnen Sie die Konvergenzradien, und untersuchen Sie das Konvergenzverhalten an den Randpunkten des Konvergenzbereichs.

$$1. \sum_{n=1}^{\infty} n^{\ln(n)/n} x^n \qquad 2. \sum_{n=1}^{\infty} \left(1 + \frac{1}{8} + \frac{1}{27} + \dots + \frac{1}{n^3}\right) x^n \qquad 3. \sum \frac{x^{n^2}}{2^n}$$

Übungsaufgabe 230. (E) Für $|x| < 1$ gilt bekanntlich $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$. Mit anderen Worten: die Funktion $\frac{1}{1-x}$ wird durch die Potenzreihe $\sum_{n=0}^{\infty} x^n$ dargestellt. Berechnen Sie die Potenzreihendarstellung der Funktion

$$f(x) = \frac{x}{x^2 - x - 1},$$

Hinweis: Berechnen Sie die Partialbruchzerlegung von $f(x)$ und verwenden Sie die Potenzreihendarstellung der geometrischen Reihe.

4.2.7 Wichtige Eigenschaften von Potenzreihen

Inhalt in Kurzfassung: Funktionen, die als Potenzreihen dargestellt werden können, haben auf dem Inneren ihres Konvergenzbereichs so ziemlich alle Eigenschaften, die man sich wünschen kann. Sie sind dort als Reihen sogar absolut konvergent und unendlich oft differenzierbar. Ihre Ableitung erhält man durch gliedweises Differenzieren der ursprünglichen Potenzreihe wieder als Potenzreihe.

Angenommen die Potenzreihe $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ hat einen positiven Konvergenzradius $r > 0$. Sei K' die Kreisscheibe in der komplexen Zahlenebene mit

Mittelpunkt x_0 und Radius $r' < r$. Mit etwas mühsamen, aber nicht sehr schwierigen Rechnungen kann man sich davon überzeugen, dass die Konvergenz auf K' dann sogar gleichmäßig erfolgt. Also ist die durch die Potenzreihe dargestellte Funktion als gleichmäßiger Grenzwert stetiger Polynome wieder stetig. Noch feinsinnigere Überlegungen zeigen, dass Potenzreihen auch differenzierbar sind etc. Ohne auf Beweise näher einzugehen, fassen wir die wichtigsten Tatsachen zusammen.

Satz 4.2.7.1. *Gegeben sei die Potenzreihe*

$$f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$$

mit $a_n \in \mathbb{R}$ (eventuell auch $a_n \in \mathbb{C}$). Ihr Konvergenzbereich sei $K \subseteq \mathbb{C}$ (f aufgefasst als komplexe Funktion) bzw. $I \subseteq \mathbb{R}$ (wenn der Definitionsbereich von f in \mathbb{R} enthalten ist), $r > 0$ der Konvergenzradius.

1. Für jedes x mit $|x - x_0| < r$ konvergiert die Potenzreihe absolut. Insbesondere stellen Cauchyprodukte von zwei Potenzreihen im Inneren des gemeinsamen Konvergenzbereichs das Produkt der dargestellten Funktionen dar (vgl. Satz 2.2.4.2).
2. Ist $r' < r$, so konvergiert die Potenzreihe sogar gleichmäßig auf der Menge aller $x \in \mathbb{C}$ mit $|x - x_0| \leq r'$.
3. Die Funktion f ist auf ganz K bzw. I stetig und auf dem Inneren von K unendlich oft differenzierbar (sogar im komplexen Sinn, der analog zum reellen Fall über Differentialquotienten definiert ist).
4. Die Ableitung f' von f (und damit auch alle höheren Ableitungen) hat ebenfalls eine Darstellung als Potenzreihe um x_0 mit demselben Konvergenzradius r , die man durch gliedweises Differenzieren erhält, d.h.:

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} (x - x_0)^n.$$

5. Die Potenzreihe ist ihre eigene Taylorreihe um x_0 , das heißt $a_n = \frac{f^{(n)}(x_0)}{n!}$.

Übungsaufgabe 231. (P) Die reellen Zahlen $a < b$ seien vorgegeben. Geben Sie konkrete Potenzreihen an, deren Konvergenzbereich in \mathbb{R} genau folgende Mengen sind:

$$[a, b], \quad [a, b), \quad (a, b], \quad (a, b), \quad \{a\}, \quad \mathbb{R}.$$

Das Verhältnis von Potenzreihen und Taylorreihen ist also das folgende: Ist eine Funktion f an einem Punkt x_0 unendlich oft differenzierbar, so ist die Taylorreihe von f in x_0 definiert als $\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}$. Diese Reihe ist eine Potenzreihe. Es ist aber nicht garantiert, dass diese Potenzreihe konvergiert, geschweige

denn gegen f . Im Inneren des Konvergenzbereichs einer beliebigen Potenzreihe jedoch ist die durch sie dargestellte Funktion unendlich oft differenzierbar, und ihre Taylorreihe ist die gegebene Potenzreihe.

Das klassische Beispiel zur Illustration möglicher Schwierigkeiten ist die Funktion $f(x) := e^{-\frac{1}{x^2}}$ für $x \neq 0$ und $f(0) := 0$. Man zeigt mit Induktion nach n durch simples Differenzieren, dass für $x \neq 0$ die n -te Ableitung von der Bauart $f^{(n)} = r_n f$ mit einer gebrochen rationalen Funktion r_n ist. Für gebrochen rationale Funktionen r gilt (Variablensubstitution $y = \frac{1}{x}$) $\lim_{x \rightarrow 0} r(x)f(x) = \lim_{y \rightarrow \pm\infty} \frac{r(\frac{1}{y})}{e^{y^2}} = 0$, weil der exponentielle Term in Nenner den höchstens polynomialen im Zähler dominiert. Nochmals Induktion nach n zeigt $f^{(n)}(0) = 0$ für alle n : Nicht nur f selbst ist stetig bei 0 mit Wert 0 (Induktionsanfang), sondern auch der Differentialquotient

$$f^{(n+1)}(0) = \lim_{x \rightarrow 0} \frac{f^{(n)}(x) - 0}{x - 0} = \lim_{x \rightarrow 0} \frac{r_n(x)}{x} f(x) = 0$$

ist von derselben Bauart (mit $\frac{r_n(x)}{x}$ statt $r_n(x)$) und mit Wert 0 (Induktionsschritt). Somit ist f auch an der Stelle $x_0 = 0$ unendlich oft differenzierbar, und alle Ableitungen nehmen den Wert 0 an. Die Taylorreihe von f an der Stelle x_0 ist somit die Nullreihe und stellt daher die Nullfunktion dar, nicht aber die Funktion f , von der wir ausgegangen sind. Die Funktion f hat also keine Potenzreihendarstellung bei $x_0 = 0$, obwohl sie unendlich oft differenzierbar ist.

Übungsaufgabe 232. (E) Die Funktion $f(x) := e^{-\frac{1}{x^2}}$ für $x \neq 0$ und $f(0) := 0$ ist, wie wir gesehen haben in $x = 0$ unendlich oft differenzierbar aber nicht in eine Potenzreihe entwickelbar. Gilt dasselbe für:

1. $f(x) := e^{-\frac{1}{x^3}}$ für $x \neq 0$ und $f(0) := 0$?
2. $f(x) := e^{-\frac{1}{x^4}}$ für $x \neq 0$ und $f(0) := 0$?

Doch zurück zu weniger ungewöhnlichen Situationen.

4.2.8 Die Regel von de l'Hospital

Inhalt in Kurzfassung: Gilt $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0$ für differenzierbare Funktionen f und g , so besagt die Regel von de l'Hospital in einer typischen Variante $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$.

Ein sehr nützliches Instrument zur Berechnung von Grenzwerten der Form $\frac{0}{0}$ oder $\frac{\infty}{\infty}$ ist die **Regel von de l'Hospital**.

Wir gehen zunächst aus von der Situation $f(a) := \lim_{x \rightarrow a^+} f(x) = 0$ und $g(a) := \lim_{x \rightarrow a^+} g(x) = 0$ für $a \in \mathbb{R}$ aus. Sofern die auftretenden Nenner $\neq 0$ sind, gilt in diesem Fall

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{\frac{f(x) - f(a)}{x - a}}{\frac{g(x) - g(a)}{x - a}}.$$

Für $x \rightarrow a$ streben Zähler und Nenner des Doppelbruchs rechts gegen $f'(a)$ bzw. $g'(a)$, also gilt $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}$. Mit Hilfe einer Verallgemeinerung des Mittelwertsatzes kann man noch stärkere Aussagen verwandter Art beweisen, die wir ohne weitere Beweisdetails in folgendem Satz zusammenfassen.

Satz 4.2.8.1. *Auf dem Definitionsbereich $D = (a, b)$ mit $-\infty < a < b < \infty$ seien die Funktionen f und g differenzierbar. Es gebe ein $\varepsilon > 0$ so, dass die Ableitung g' im kleinen Bereich $(a, a + \varepsilon)$ rechts von a keine Nullstelle habe. Gilt entweder $\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^+} g(x) = 0$ oder $\lim_{x \rightarrow a^+} g(x) \in \{-\infty, \infty\}$, so folgt*

$$\lim_{x \rightarrow a^+} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a^+} \frac{f'(x)}{g'(x)},$$

sofern der rechte Grenzwert als reelle Zahl oder auch nur als ∞ oder $-\infty$ existiert. Die entsprechende Aussage gilt für $\lim_{x \rightarrow b^-}$, außerdem sinngemäß für $a = -\infty$ bzw. $b = \infty$.

Ein typisches Anwendungsbeispiel, wo die Regel von de l'Hospital sogar zweimal zum Zug kommt, ist

$$\lim_{x \rightarrow 0} \frac{\exp(x) - x - 1}{x^2} = \lim_{x \rightarrow 0} \frac{\exp(x) - 1}{2x} = \lim_{x \rightarrow 0} \frac{\exp(x)}{2} = \frac{1}{2}.$$

Für noch interessantere Anwendungen müssen wir unseren Fundus an differenzierbaren Funktionen erweitern. Doch zunächst noch einige Übungsaufgaben.

Übungsaufgabe 233. (E) Die Funktionen f und g in Satz 4.2.8.1 mögen Potenzreihendarstellungen an der Stelle a haben. Weiters gelte $f(a) = g(a) = 0$, und g sei nicht konstant 0. Beweisen Sie die Regel von de l'Hospital direkt unter diesen Voraussetzungen, ohne Satz 4.2.8.1 zu verwenden.

Übungsaufgabe 234. (T) Berechnen Sie den Grenzwert mit der Regel von de l'Hospital und mittels Entwicklung von Zähler und Nenner in Potenzreihen. Vergleichen Sie den Rechenaufwand.

$$\lim_{x \rightarrow 0} \frac{x(e^x - e^{-x})}{e^{x^2} - e^{-x^2}}$$

Übungsaufgabe 235. (T) Berechnen Sie die folgenden Grenzwerte für alle positiven reellen Zahlen α :

$$\lim_{x \rightarrow \infty} \frac{x}{\exp(\alpha x)} \quad \text{und} \quad \lim_{x \rightarrow \infty} \frac{\ln x}{x^\alpha}.$$

Schließen Sie daraus auf die Grenzwerte von

$$\lim_{x \rightarrow \infty} \frac{x^\beta}{\exp(\alpha x)} \quad \text{und} \quad \lim_{x \rightarrow \infty} \frac{(\ln x)^\beta}{x^\alpha},$$

für alle positiven reellen Zahlen α und β .

Übungsaufgabe 236. (P) Sei

$$f(x) = \frac{x^2 \cos \frac{1}{x}}{\sin x}.$$

1. Warum lässt sich der Grenzwert $\lim_{x \rightarrow 0} f(x)$ nicht mit der Regel von de l'Hospital berechnen? Hinweis: Verwenden Sie an geeigneter Stelle die beiden Folgen $a_k = \frac{1}{k\pi}$ und $b_k = \frac{2}{(2k+1)\pi}$.

2. Berechnen Sie den Grenzwert mit einer anderen Methode.

Hinweis: $f(x) = \frac{x}{\sin x} x \cos\left(\frac{1}{x}\right)$

4.3 Wichtige Beispiele differenzierbarer Funktionen

Im vorangegangenen Abschnitt diente die Exponentialfunktion als Ausgangspunkt und repräsentatives Beispiel für die Theorie der Potenzreihen. Nun sollen weitere wichtige Funktionen behandelt werden. Bemerkenswerterweise stammen sie alle in der einen oder anderen Weise von der Exponentialfunktion ab. Ihre besondere Rolle in der Mathematik kommt dadurch eindrucksvoll zum Ausdruck. Konkret werden wir den Logarithmus (4.3.1), Potenzfunktionen mit beliebigen reellen Exponenten (4.3.2), die trigonometrischen (4.3.3), die Arcus- (4.3.4) und die Areefunktionen (4.3.5) unter dem Gesichtspunkt der Differentialrechnung behandeln.

4.3.1 Logarithmus, Fortsetzung

Inhalt in Kurzfassung: Der Logarithmus \log_a zur Basis $a > 0$, $a \neq 1$, ist als die Umkehrfunktion von \exp_a , der Exponentialfunktion zur Basis a , definiert. Für $a = e$ spricht man auch von natürlichen Logarithmus und schreibt \ln . Er genügt der Ableitungsregel $\ln' x = \frac{1}{x}$ und hat eine einfache Potenzreihendarstellung (logarithmische Reihe) mit Konvergenzradius 1. Vermittels der Beziehung $\log_a x = \frac{\ln x}{\ln a}$ lassen sich Eigenschaften des natürlichen Logarithmus \ln auch leicht auf \log_a zu einer anderen Basis a übertragen.

Wie wir bereits wissen, ist der Logarithmus \log_a zur Basis $a > 0$, $a \neq 1$, als Umkehrfunktion der Exponentialfunktion \exp_a zur Basis a definiert. In Hinblick auf die Differentialrechnung empfiehlt es sich, mit der Basis $a = e$ zu beginnen. Zur Erinnerung aus 3.3.3: Man schreibt $\ln := \log_e$ und nennt den Logarithmus zur Basis e auch den **natürlichen Logarithmus**. Als Umkehrfunktion der differenzierbaren Funktion \exp ist \ln selbst differenzierbar mit Ableitung (Formel für die Ableitung einer Umkehrfunktion aus Satz 4.1.4.1)

$$\ln'(x) = \frac{1}{\exp'(\ln(x))} = \frac{1}{\exp(\ln(x))} = \frac{1}{x}.$$

Übungsaufgabe 237. (T) Berechnen Sie die erste Ableitung von $f(x) := \ln(\ln(\ln(x)))$. Wo ist f definiert?

Jetzt sind wir in der Lage, die wichtige und früher bereits erwähnte Formel

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$$

sehr schnell mit Hilfe der Regel von de l'Hospital herzuleiten. Und zwar ersetzen wir die Variable n durch die reelle Variable $x > 0$, formen um zu $(1 + \frac{1}{x})^x = e^{x \ln(1 + \frac{1}{x})}$ und setzen dann noch $x = \frac{1}{y}$. Anwendung der Regel von de l'Hospital liefert

$$\lim_{x \rightarrow \infty} x \ln\left(1 + \frac{1}{x}\right) = \lim_{y \rightarrow 0} \frac{\ln(1+y)}{y} = \lim_{y \rightarrow 0} \frac{\frac{1}{1+y}}{1} = 1.$$

Unter Verwendung der Stetigkeit der Exponentialfunktion fassen wir zusammen zu

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e^{\lim_{x \rightarrow \infty} x \ln(1 + \frac{1}{x})} = e^1 = e,$$

was behauptet wurde.

Aus der Ableitungsregel für den Logarithmus lässt sich für $-1 < x \leq 1$ auch die Darstellung von \ln als **logarithmische Reihe**

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}$$

gewinnen. Wir verwenden dabei, dass eine Funktion auf einem Intervall durch ihre Ableitung und einen Funktionswert (hier $\ln 1 = 0$) eindeutig gegeben ist. Deshalb genügt es, obige Reihe abzuleiten und unter Verwendung der Formel für die geometrische Reihe

$$\left(\sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}\right)' = \sum_{n=1}^{\infty} (-1)^{n+1} x^{n-1} = \sum_{n=0}^{\infty} (-x)^n = \frac{1}{1+x} = \ln'(1+x)$$

abzulesen. Nach Satz 4.2.2.3 müssen die beiden Funktionen $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}$ und $\ln(1+x)$ auf $(-1, 1)$ also tatsächlich übereinstimmen. Bemerkenswert ist die Ausdehnung auf den Fall $x = 1$: Der **Abelsche² Grenzwertsatz**, den wir hier nicht beweisen, besagt, dass eine Potenzreihe, sofern sie an einem Randpunkt ihres Konvergenzbereichs konvergiert, an diesem Punkt gegen die stetige Fortsetzung der auf dem Inneren des Konvergenzbereichs dargestellten Funktion (wo ja immer Stetigkeit herrscht) konvergiert. Bei der logarithmischen Reihe ist das für $x = 1$ der Fall. Daraus folgt:

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots$$

²nach dem norwegischen Mathematiker Niels Henrik Abel (1802-1829), siehe auch 3.2.4.

Übungsaufgabe 238. (P) Gewinnen Sie durch Anwendung der Rechenregeln für den Logarithmus die Potenzreihendarstellung der Funktion $\ln \sqrt{\frac{1+x}{1-x}}$ aus der logarithmischen Reihe um den Entwicklungspunkt $x = 0$.

Logarithmen \log_a zu einer Basis $a \neq e$ ($a > 0$, $a \neq 1$) lassen sich leicht auf den natürlichen Logarithmus zurückführen. Denn aus

$$e^{\ln x} = x = a^{\log_a x} = (e^{\ln a})^{\log_a x} = e^{(\ln a) \log_a x}$$

lesen wir $\ln x = (\ln a) \log_a x$ ab, also

$$\log_a x = \frac{\ln x}{\ln a}.$$

Logarithmen zu verschiedenen Basen unterscheiden sich also nur durch multiplikative Konstante. Die Ableitung von \log_a ist gegeben durch

$$\log'_a x = \frac{1}{(\ln a)x}.$$

4.3.2 Allgemeine Potenzen und binomische Reihe

Inhalt in Kurzfassung: Durch Verkettung von Logarithmus, Multiplikation mit einer Konstanten α und Exponentialfunktion erhält man die Potenzfunktion $p_\alpha(x) = x^\alpha$, auch für irrationales α . Aus dieser Darstellung und der Kettenregel ergibt sich die Ableitungsregel $(x^\alpha)' = \alpha x^{\alpha-1}$, nun für beliebiges reelles α . Auch für jede Potenzfunktion gibt es eine Potenzreihendarstellung mit Konvergenzradius 1. Und zwar treten darin geeignet verallgemeinerte Binomialkoeffizienten $\binom{\alpha}{n}$ auf, weshalb man von der Binomialreihe spricht.

Die Potenzfunktionen $p_\alpha : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, $x \mapsto x^\alpha$, lassen sich darstellen als

$$p_\alpha(x) = x^\alpha = (e^{\ln x})^\alpha = e^{\alpha \ln x},$$

d.h. als Komposition $p_\alpha = \exp \circ m_\alpha \circ \ln$ mit der Multiplikationsfunktion $m_\alpha : x \mapsto \alpha x$. Alle drei Funktionen sind differenzierbar, nach der Kettenregel also auch ihre Komposition, und es gilt für die Ableitung

$$(x^\alpha)' = p'_\alpha(x) = \exp'(m_\alpha(\ln x)) \cdot m'_\alpha(\ln x) \cdot \ln'(x)$$

Nach Berechnung der drei Faktoren

$$\begin{aligned} \exp'(m_\alpha(\ln x)) &= \exp(\alpha \ln x) = x^\alpha, \\ m'_\alpha(\ln x) &= \alpha \quad \text{und} \\ \ln'(x) &= \frac{1}{x} \end{aligned}$$

ergibt sich die vertraute Differentiationsregel

$$(x^\alpha)' = p'_\alpha(x) = x^\alpha \frac{\alpha}{x} = \alpha x^{\alpha-1},$$

die erst jetzt für beliebige reelle, insbesondere auch irrationale Exponenten α bewiesen ist.

Übungsaufgabe 239. (T) Wo ist die Funktion $f(x) := x^x$ definiert? Berechnen Sie f' , untersuchen Sie das Verhalten von f insbesondere in der Nähe von $x = 0$ und fertigen Sie eine Skizze an.

Auch für Potenzfunktionen gibt es eine nützliche Potenzreihendarstellung, nämlich für $|x| < 1$ durch die **binomische Reihe**:

$$f(x) := (1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n.$$

Der Witz besteht darin, dass α nicht mehr ganzzahlig sein muss. Der hier auftretende **verallgemeinerte Binomialkoeffizient** ist definiert durch

$$\binom{\alpha}{n} := \frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-n+1)}{n!}.$$

(Man beachte, dass diese Definition für $0 \leq n \leq \alpha \in \mathbb{N}$ mit der bisher gegebenen übereinstimmt.) Durch iteriertes Differenzieren von f erhält man $f'(x) = \alpha(1+x)^{\alpha-1}$, $f''(x) = \alpha(\alpha-1)(1+x)^{\alpha-2}$ etc. und sieht sehr schnell, dass die binomische Reihe die Taylorreihe von f an der Stelle $x_0 = 0$ ist. Dass die Taylorreihe auch tatsächlich f darstellt, erreicht man in diesem Fall nicht durch Abschätzung des Restgliedes $r_{f,x_0,n}(x)$, wie es in unserer Version des Satzes 4.2.4.2 von Taylor beschrieben wurde. Ein Beweis ergibt sich aus der Beobachtung, dass die binomische Reihe positiven Konvergenzradius $r = 1$ hat und in diesem Bereich deshalb eine Funktion g darstellt. Wir wollen $g(x) = (1+x)^\alpha$ zeigen. Dazu rechnet man zunächst die Beziehung $(1+x)g'(x) = \alpha g(x)$ nach (Übung), woraus

$$\ln(g(x))' = \frac{g'(x)}{g(x)} = \frac{\alpha}{1+x}$$

folgt. Sämtliche Stammfunktionen F der rechten Seite haben die Form $F(x) = \alpha \ln(1+x) + c$ mit $c \in \mathbb{R}$ (siehe auch Satz 4.2.2.3), also $\ln(g(x)) = \alpha \ln(1+x) + c$ oder $g(x) = e^{\alpha \ln(1+x)} e^c = (1+x)^\alpha e^c$. Einsetzen des speziellen Wertes $x = 0$ ergibt $1 = g(0) = e^c$, also $g(x) = (1+x)^\alpha$, wie behauptet.

Übungsaufgabe 240. Zeigen Sie, dass die binomische Reihe tatsächlich Konvergenzradius $r = 1$ hat und daher für $|x| < 1$ tatsächlich eine Funktion g mit dieser Potenzreihe darstellt. Prüfen Sie für dieses g die in obiger Überlegung nur behauptete Beziehung $(1+x)g'(x) = \alpha g(x)$ nach.

4.3.3 Die Differenzierbarkeit der trigonometrischen Funktionen

Inhalt in Kurzfassung: Über den Umweg der komplexen Zahlen erweisen sich auch die trigonometrischen Funktionen als eng mit der Exponentialfunktion

verbunden. Und zwar ist $\cos \alpha$ für reelles α der Realteil und $\sin \alpha$ der Imaginärteil der komplexen Zahl $\exp(i\alpha) = e^{i\alpha}$. Diese Zusammenhänge ermöglichen mathematisch strenge Definitionen von \cos und \sin als Potenzreihen, die sogar auf ganz \mathbb{C} konvergieren und die Ableitungsregeln $\cos' = -\sin$, $\sin' = \cos$ erfüllen. Als Quotienten dieser beiden Funktionen sind nunmehr auch Tangens und Cotangens rigoros definiert.

Mit Hilfe von Potenzreihen können wir unsere Überlegungen zu den trigonometrischen Funktionen aus 3.3.4 auf eine feste Basis stellen. Wir gehen nochmals aus von der Exponentialfunktion. Und zwar betrachten wir sie nun als komplexe Funktion. Dabei werden wir einen verblüffenden Zusammenhang mit den trigonometrischen Funktionen feststellen. Und zwar betrachten wir die Funktion $\phi: \mathbb{R} \rightarrow \mathbb{C}$, $\phi: x \mapsto \exp(ix)$, $x \in \mathbb{R}$. Dabei setzen wir $\phi(x) = f(x) + ig(x)$ mit Realteil $f(x)$ und Imaginärteil $g(x)$, wodurch die reellen Funktionen $f, g: \mathbb{R} \rightarrow \mathbb{R}$ definiert sind. In der Exponentialreihe für $\exp(ix)$ wird der Realteil von den Gliedern mit geradem Index $n = 2k$ gebildet, der Imaginärteil von denen mit ungeradem $n = 2k + 1$ (jeweils $k \in \mathbb{N}$), also

$$\exp(ix) = \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Daraus folgen für die Funktionen f und g die Potenzreihendarstellungen

$$f(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

und

$$g(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}.$$

Wir sammeln nun einige Informationen über f und g . Zunächst liest man

$$e^{-ix} = f(x) - ig(x) = \overline{e^{ix}}$$

ab, also

$$|\phi(x)|^2 = |e^{ix}|^2 = e^{ix} \overline{e^{ix}} = e^{ix} e^{-ix} = e^{ix-ix} = e^0 = 1.$$

Somit liegt der Wertebereich von ϕ auf dem Einheitskreis, also $f(x)^2 + g(x)^2 = 1$. Weiters zeigt (gliedweise) Differentiation der Potenzreihen für f und g , dass

$$f'(x) = -g(x) \quad \text{und} \quad g'(x) = f(x).$$

Für kleine Δx bedeutet das $f(x + \Delta x) \approx f(x) + \Delta x f'(x) = f(x) - \Delta x g(x)$ sowie $g(x + \Delta x) \approx g(x) + \Delta x g'(x) = g(x) + \Delta x f(x)$. Setzen wir Real- und Imaginärteil von $\phi(x)$ zusammen und verwenden wir $i\phi(x) = i(f(x) + ig(x)) = -g(x) + if(x)$, bedeutet das

$$\begin{aligned} \phi(x + \Delta x) &= f(x + \Delta x) + ig(x + \Delta x) \approx \\ &\approx f(x) - \Delta x g(x) + ig(x) + i\Delta x f(x) = \phi(x) + i\Delta x \phi(x). \end{aligned}$$

Interpretieren wir $\phi(x)$ als Bewegung auf dem komplexen Einheitskreis in Abhängigkeit vom Zeitparameter x , so bewegen wir uns also in Δx Zeiteinheiten um einen Weg der Länge $|i\Delta x\phi(x)| = \Delta x$ weiter, d.h. mit Geschwindigkeit 1. Multiplikation mit i bedeutet in der komplexen Ebene eine Drehung um einen rechten Winkel nach links. Also erfolgt die Bewegung auf dem Einheitskreis gegen den Uhrzeigersinn. Offenbar ist $\phi(0) = 1$, also startet die Bewegung im Zeitpunkt $x = 0$ beim Punkt 1.

Vergleichen wir mit unserer (damals, am Ende von Kapitel 3, noch nicht völlig rigorosen) Definition der trigonometrischen Funktionen \cos und \sin , so beobachten wir völlige Übereinstimmung, wobei $f = \cos$ und $g = \sin$ zu setzen ist. Diese nicht rigoros, sondern teilweise über die Anschauung gewonnene Einsicht motiviert uns, **Cosinus** und **Sinus** nun streng durch die gewonnenen Reihen zu definieren:

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

und

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

Man beachte, dass diese Definitionen auch für komplexes x sinnvoll sind. Unsere bisherigen Überlegungen zeigen den Zusammenhang

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

für alle $x \in \mathbb{R}$ und allgemeiner, wenn wir komplexe Argumente $a+ib$ mit Realteil a und Imaginärteil b zulassen,

$$\exp(a+ib) = e^{a+ib} = e^a e^{ib} = r(\cos b + i \sin b) \quad \text{mit } r := e^a.$$

Außerdem lesen wir aus den weiter oben abgeleiteten Beziehungen zwischen den Potenzreihen die Ableitungsregeln

$$\cos' x = -\sin x \quad \text{und} \quad \sin' x = \cos x$$

ab. Die Kreiszahl π lässt sich nun streng definieren, beispielsweise als das Doppelte der kleinsten positiven Nullstelle des Cosinus. Auch die Additionstheoreme für \cos und \sin , die uns erstmals in 1.5.4 begegnet sind, kann man nun auf eine neue und einfache Art herleiten. Für reelle x, y braucht man nur in

$$\begin{aligned} \cos(x+y) + i \sin(x+y) &= e^{i(x+y)} = e^{ix+iy} = e^{ix} e^{iy} = \\ &= (\cos x + i \sin x)(\cos y + i \sin y) = \\ &= \cos x \cos y - \sin x \sin y + i(\sin x \cos y + \cos x \sin y) \end{aligned}$$

Real- und Imaginärteil zu vergleichen.

Man beachte auch die oft nützlichen Beziehungen

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \text{und} \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}.$$

Mit \cos und \sin sind nun auch **Tangens** und **Cotangens** durch $\tan x := \frac{\sin x}{\cos x}$ und $\cot x := \frac{\cos x}{\sin x}$ auf ihren jeweiligen Definitionsbereichen streng definiert. Die geometrischen Interpretationen stehen natürlich weiterhin in gewohnter Weise zur Verfügung. Die Ableitung \tan' ergibt sich aus der Quotientenregel:

$$\tan' x = \left(\frac{\sin x}{\cos x} \right)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = 1 + \tan^2 x$$

Für den letzten Ausdruck rechts wurde der Bruch im mittleren dritten in zwei Summanden aufgeteilt. Ähnlich berechnet man \cot' :

Übungsaufgabe 241. (T) Berechnen Sie \cot' .

Übungsaufgabe 242. (T) Berechnen Sie für folgende Funktionen das Taylorpolynom dritten Grades mit dem angegebenen Entwicklungspunkt x_0 .

1. $\sin(x) - x$, $x_0 = 0$
2. $\sin(x) - x$, $x_0 = \frac{\pi}{2}$
3. $\sin(x) + x$, $x_0 = \pi$

Übungsaufgabe 243. (T) Zeigen Sie mit vollständiger Induktion, dass die n -te Ableitung von $\sin x e^x$ durch

$$\begin{cases} (-4)^k \sin x e^x & \text{wenn } n = 4k \\ (-4)^k (\cos x + \sin x) e^x & \text{wenn } n = 4k + 1 \\ (-4)^k 2 \cos x e^x & \text{wenn } n = 4k + 2 \\ (-4)^k 2 (\cos x - \sin x) e^x & \text{wenn } n = 4k + 3 \end{cases}$$

gegeben ist.

Übungsaufgabe 244. (T) Berechnen Sie jeweils das Differential von $f(x)$, d.h. die lineare Approximation, an der angegebenen Stelle x_0 und berechnen Sie damit näherungsweise den angegebenen Wert.

1. $f(x) = e^x$, $x_0 = 0$, $e^{0.1} = ?$
2. $f(x) = \sin(x)$, $x_0 = 0$, $\sin(5^\circ) = ?$
3. $f(x) = \sqrt{x}$, $x_0 = 1$, $\sqrt{1.1} = ?$
4. $f(x) = x^{20}$, $x_0 = 1$, $(1.02)^{20} = ?$

4.3.4 Die Arcusfunktionen

Inhalt in Kurzfassung: Die Definition der Arcusfunktionen als Umkehrfunktionen der trigonometrischen auf jeweils geeignetem Definitionsbereich ist bereits aus 3.3.4 bekannt. Nun werden für diese Funktionen auch die Ableitungsregeln hergeleitet.

Wir wissen bereits aus 3.3.4: Die Arcusfunktionen **Arcus cosinus**, **Arcus sinus**, **Arcus tangens** und **Arcus cotangens**, symbolisch \arccos , \arcsin ,

arctan und arccot, sind definiert als Umkehrfunktionen der lokalen trigonometrischen Funktionen

$$\begin{aligned} \cos &: [0, \pi] \rightarrow [-1, 1] & \arccos &: [-1, 1] \rightarrow [0, \pi] \\ \sin &: \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow [-1, 1] & \arcsin &: [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ \tan &: \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \rightarrow \mathbb{R} & \arctan &: \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \\ \cot &: (0, \pi) \rightarrow \mathbb{R} & \operatorname{arccot} &: (0, \pi) \rightarrow \mathbb{R} \end{aligned}$$

Vor allem aus Sicht der Integralrechnung (siehe nächstes Kapitel), sind die Ableitungsregeln von Interesse. Unter Verwendung der Beziehung

$$\cos(\arcsin x) = \sqrt{1 - \sin^2(\arcsin x)} = \sqrt{1 - x^2}$$

und der Formel aus Satz 4.1.4.1 für die Ableitung von Umkehrfunktionen berechnen wir

$$\arcsin' x = \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - x^2}}$$

für $-1 < x < 1$ (den Definitionsbereich von arcsin). Ganz ähnlich erhält man auf demselben Bereich

$$\arccos' x = -\frac{1}{\sqrt{1 - x^2}}.$$

Am häufigsten kommt die Ableitung des arctan vor: Mit Hilfe der Ableitungsregel $\tan' x = 1 + \tan^2$ aus 4.3.3 erhalten wir wieder nach der Differentiationsregel für Umkehrfunktionen

$$\arctan' x = \frac{1}{\tan'(\arctan x)} = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2},$$

diesmal für alle $x \in \mathbb{R}$; ganz analog ergibt sich $\operatorname{arccot}' x = -\frac{1}{1+x^2}$.

Übungsaufgabe 245. (T) Beweisen Sie die Ableitungsregel $\operatorname{arccot}' x = -\frac{1}{1+x^2}$ für den Arcus cotangens.

Interessant ist, dass aus der Ableitung von arctan zusammen mit der Formel für die geometrische Reihe eine Reihendarstellung für π gewonnen werden kann. Aus

$$\arctan' x = \frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-x^2)^n = \sum_{n=0}^{\infty} (-1)^n x^{2n} = 1 - x^2 + x^4 - x^6 + \dots$$

zusammen mit $\arctan 0 = 0$ folgt nämlich (wie bei der logarithmischen Reihe wegen der Eindeutigkeit von Stammfunktionen bis auf eine additive Konstante)

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}.$$

Bei dieser **Arcustangensreihe** handelt es sich für $0 \leq x \leq 1$ um eine Leibnizreihe. Nach dem Abelschen Grenzwertsatz (siehe auch 4.3.1 im Zusammenhang mit der logarithmischen Reihe) gilt die Reihendarstellung auch noch für den Randpunkt $x = 1$ des Konvergenzbereichs. Wegen $\arctan 1 = \frac{\pi}{4}$ bedeutet das:

$$\pi = 4 \arctan 1 = 4 \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = 4 \left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots \right)$$

Übungsaufgabe 246. (P) Finden Sie eine analoge Reihendarstellung für π , indem Sie statt wie oben $\tan \frac{\pi}{4} = 1$ nun die Beziehung $\tan \frac{\pi}{6} = \frac{1}{\sqrt{3}}$ begründen und verwenden.

Übungsaufgabe 247. (P) Zeigen Sie die Gültigkeit der trigonometrischen Formel

$$\arctan x + \arctan y = \arctan \left(\frac{x+y}{1-xy} \right) \quad \text{für } xy < 1.$$

Hinweis: Rechnen Sie durch explizites Differenzieren nach, dass die linke und die rechte Seite dieser Gleichung dieselbe Ableitung haben. Was ist noch zu tun?

4.3.5 Die Hyperbel- und Areafunktionen

Inhalt in Kurzfassung: Ähnlich wie die trigonometrischen Funktionen (sogar leichter, weil kein Umweg ins Komplexe nötig ist) lassen sich mit Hilfe der Exponentialfunktion die Hyperbelfunktionen definieren. Ihre Umkehrfunktionen heißen Areafunktionen und lassen sich ganz analog zu den Arcusfunktionen aus 4.3.4 behandeln.

Aus der Exponentialfunktion \exp haben wir den Logarithmus (als Umkehrfunktion) gewonnen, die Potenzfunktionen (als Verknüpfung von Logarithmus, Multiplikation und Exponentiation), Cosinus und Sinus (als Real- bzw. Imaginärteil der komplexen Exponentialfunktion), Tangens und Cotangens als Quotienten von Sinus und Cosinus und die Arcusfunktionen als lokale Umkehrfunktionen von Cosinus, Sinus, Tangens und Cotangens. Auch die vier **Hyperbelfunktionen** sind mittels \exp definiert. Zunächst **Cosinus hyperbolicus** \cosh und **Sinus hyperbolicus** \sinh durch

$$\cosh x := \frac{e^x + e^{-x}}{2} \quad \text{und} \quad \sinh x := \frac{e^x - e^{-x}}{2}.$$

Analog zur Beziehung $\cos^2 + \sin^2 = 1$ der trigonometrischen Funktionen, die zeigt, dass alle Punkte $(x, y) = (\cos t, \sin t)$ mit $t \in \mathbb{R}$ die Gleichung $x^2 + y^2 = 1$ erfüllen und deshalb in \mathbb{R}^2 auf dem Einheitskreis liegen, gilt für die hyperbolischen Funktionen

$$\cosh^2 - \sinh^2 = 1.$$

Übungsaufgabe 248. (T) Leiten Sie diese Beziehung aus den Definitionen von \cosh und \sinh her.

Alle Punkte $(x, y) = (\cosh t, \sinh t)$ mit $t \in \mathbb{R}$ erfüllen daher die Gleichung $x^2 - y^2 = 1$ und liegen folglich in \mathbb{R}^2 auf der durch diese Gleichung beschriebenen (gleichseitigen) Hyperbel. Das erklärt die Bezeichnung *Hyperbelfunktionen*.

Direkt aus den Definitionen liest man die Ableitungsregeln

$$\cosh' = \sinh \quad \text{und} \quad \sinh' = \cosh$$

ab. Die Berechnung der Reihendarstellungen für \cosh und \sinh ist vergleichsweise einfach:

Übungsaufgabe 249. (T) *Ermitteln Sie Reihendarstellungen für \cosh und \sinh .*

Die offensichtliche Analogie zu den trigonometrischen Funktionen fortsetzend definiert man auch **Tangens hyperbolicus** \tanh und **Cotangens hyperbolicus** \coth durch

$$\tanh := \frac{\sinh}{\cosh} \quad \text{und} \quad \coth := \frac{\cosh}{\sinh}.$$

Für \coth muss 0 aus dem Definitionsbereich herausgenommen werden. Hingegen sind \cosh , \sinh und \tanh auf ganz \mathbb{R} definiert. Die Ableitungen sind

$$\tanh' = \frac{1}{\cosh^2} = 1 - \tanh^2 \quad \text{und} \quad \coth' = -\frac{1}{\sinh^2} = 1 - \coth^2.$$

Übungsaufgabe 250. (T) *Rechnen Sie diese Ableitungsregeln nach.*

In ähnlicher Weise wie bei den trigonometrischen und den Arcusfunktionen betrachtet man auch für die Hyperbelfunktionen ihre lokalen Umkehrfunktionen, die sogenannten **Areafunktionen: Area cosinus hyperbolicus, Area sinus hyperbolicus, Area tangens hyperbolicus** und **Area cotangens hyperbolicus**, symbolisch Arcosh , Arsinh , Artanh und Arcoth . Zur besseren Übersicht wieder eine Tabelle mit den Definitionsbereichen, wo Bijektivität herrscht:

$$\begin{array}{ll} \cosh : [0, \infty) \rightarrow [1, \infty) & \text{Arcosh} : [1, \infty) \rightarrow [0, \infty) \\ \sinh : \mathbb{R} \rightarrow \mathbb{R} & \text{Arsinh} : \mathbb{R} \rightarrow \mathbb{R} \\ \tanh : \mathbb{R} \rightarrow (-1, 1) & \text{Artanh} : (-1, 1) \rightarrow \mathbb{R} \\ \coth : \mathbb{R} \setminus \{0\} \rightarrow (-\infty, -1) \cup (1, \infty) & \text{Arcoth} : (-\infty, -1) \cup (1, \infty) \rightarrow \mathbb{R} \setminus \{0\} \end{array}$$

Die Ableitungen auf den jeweiligen Definitionsbereichen erhält man analog zu jenen der Arcusfunktionen in 4.3.4 mit Hilfe von Satz 4.1.4.1 und der Beziehung $\cosh^2 - \sinh^2 = 1$:

$$\text{Arcosh}'x = \frac{1}{\sqrt{x^2 - 1}} \quad \text{und} \quad \text{Arsinh}'x = \frac{1}{\sqrt{x^2 + 1}}$$

sowie

$$\text{Artanh}'x = \frac{1}{1 - x^2} \quad \text{und} \quad \text{Arcoth}'x = \frac{1}{1 - x^2}.$$

Der Unterschied zwischen den letzten beiden Ableitungen liegt im Definitionsbereich: $|x| < 1$ für Arctanh und $|x| > 1$ für Arcoth .

Übungsaufgabe 251. (P) Erklären Sie ausführlich, wie man die Formeln für die Ableitungen der Areafunktionen erhält.

4.4 Anwendungen der Differentialrechnung

In diesem letzten Abschnitt zur Differentialrechnung beschäftigen wir uns mit einigen wichtigen innermathematischen Anwendungen. Wir beginnen mit Untersuchungen des Zusammenhangs zwischen Differentialrechnung einerseits und Monotonie andererseits (4.4.1). Große Analogien dazu finden sich in der Beziehung zwischen Konvexität, zweiter Ableitung und Krümmung (4.4.2). Daran schließen Extremwertuntersuchungen unter Zuhilfenahme auch höherer Ableitungen an (4.4.3). Andere Gesichtspunkte treten auf, wenn man Iterationen betrachtet, Schlagwort Dynamik und Chaos (4.4.4). Von besonderem Interesse ist das Newtonverfahren zur approximativen Berechnung von Nullstellen stetig differenzierbarer Funktionen (= Lösung auch nichtlinearer Gleichungen). Dabei handelt es sich um einen Spezialfall einer Iteration, den wir in 4.4.5 genauer untersuchen.

4.4.1 Monotonie und erste Ableitung, Fortsetzung

Inhalt in Kurzfassung: Die Ergebnisse aus 4.2.1 lassen sich durch Einsatz des Mittelwertsatzes verstärken zu Aussagen wie: Gilt $f'(x) \geq 0$ für alle x auf einem ganzen Intervall, so ist f dort monoton wachsend (analog fallend bei \leq und streng monoton wachsend/fallend bei $>$ bzw. $<$). Kurz gehen wir auch auf eine Art Umkehrung ein: Kann man aus Monotonie auf Stetigkeit oder gar Differenzierbarkeit schließen?

Wir schließen an 4.2.1 an, werden uns nun aber Ergebnisse zunutze machen, die uns dort noch nicht zur Verfügung standen, vor allem den Mittelwertsatz 4.2.2.1.

Bei einer monoton wachsenden Funktion f sind Differenzenquotienten immer ≥ 0 , als deren Grenzwert daher auch die Ableitung $f' \geq 0$ (sofern existent). Analog gilt $f' \leq 0$ für monoton fallendes differenzierbares f . Dass auch ein umgekehrter Schluss möglich ist, zeigt der Mittelwertsatz. Ihm zufolge gilt (o.B.d.A. sei $a < b$)

$$f(b) = f(a) + f'(x)(b - a)$$

mit einer geeigneten Zwischenstelle $x \in (a, b)$. Ist $f'(x) \geq 0$ für alle $x \in (a, b)$, so folgt $f(a) \leq f(b)$. Weil dieses Argument statt für $a < b$ auch für je zwei beliebige a', b' mit $a \leq a' < b' \leq b$ zutrifft schließen wir, dass f in diesem Fall monoton wachsend auf $[a, b]$ ist. Diese sowie entsprechende Aussagen bei anderer Ungleichung fassen wir zusammen zu:

Satz 4.4.1.1. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und differenzierbar auf (a, b) , dann gelten folgende Implikationen:

1. Gilt $f'(x) \geq 0$ für alle $x \in (a, b)$, so ist f auf $[a, b]$ monoton wachsend.

2. Gilt $f'(x) > 0$ für alle $x \in (a, b)$, so ist f auf $[a, b]$ streng monoton wachsend.
3. Gilt $f'(x) \leq 0$ für alle $x \in (a, b)$, so ist f auf $[a, b]$ monoton fallend.
4. Gilt $f'(x) < 0$ für alle $x \in (a, b)$, so ist f auf $[a, b]$ streng monoton fallend.

Zum Vergleich: Aus 4.2.1 wissen wir bereits, dass im Fall von $f'(x_0) \neq 0$ die Funktion f im Punkt x_0 monoton sein muss in dem Sinn, dass es eine Umgebung von U gibt derart, dass für $x \in U$ stets $f(x) < f(x_0)$ bei $x < x_0$ und $f(x) > f(x_0)$ bei $x > x_0$ gilt. Es muss aber keine ganze Umgebung von x_0 geben, auf der f monoton ist. Als Beispiel kann die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f : x \mapsto x + 2x^2 \sin \frac{1}{x}$ für $x \neq 0$ und $f : (0) = 0$ dienen. Ähnlich wie schon in 4.2.3 zeigt man $f'(0) = 1$, während f' in jeder Umgebung von 0 zwischen den Werten -1 und 3 oszilliert. Letzteres verträgt sich nicht mit Monotonie.

Man könnte auch die Frage stellen, ob aus der Monotonie einer reellen Funktion f auf einem Intervall auch ohne zusätzliche Voraussetzung gewisse Schlüsse über Stetigkeit und Differenzierbarkeit gezogen werden können. Offenbar können monotone Funktionen durchaus Unstetigkeiten (nämlich Sprungstellen, aber nur solche) aufweisen. Wir überlegen weiter: An jeder Sprungstelle wird ein ganzes Intervall aus potentiellen Funktionswerten übersprungen. Jedes solche Sprungintervall enthält rationale Zahlen. Die (offenen) Sprungintervalle sind paarweise disjunkt. Also kann es nicht mehr Sprungintervalle und somit Sprungstellen geben als rationale Zahlen. Und das sind nur abzählbar viele. Also: Eine monotone Funktion kann nur abzählbar viele Unstetigkeitsstellen haben, muss also (sobald sie wenigstens auf einem Intervall definiert ist) vorwiegend (nämlich überabzählbar viele) Stetigkeitsstellen haben.

Übungsaufgabe 252. (E) Sei $f : [a, b] \rightarrow \mathbb{R}$ monoton und $a < x_0 < b$. Zeigen Sie, dass in x_0 sowohl der linksseitige Grenzwert $\lim_{x \rightarrow x_0^-} f(x)$ als auch der rechtsseitige Grenzwert $\lim_{x \rightarrow x_0^+} f(x)$ von f existieren. Ist f in x_0 unstetig, so liegt also eine Sprungstelle vor.

Anleitung: Betrachten Sie im monoton wachsenden Fall $\sup_{x < x_0} f(x)$ und $\inf_{x > x_0} f(x)$.

Auch zur analogen Frage über Differenzierbarkeit einer monotonen Funktion f lässt sich Interessantes sagen: Die Menge der Punkte x , wo $f'(x)$ nicht existiert, kann zwar sehr wohl überabzählbar sein, in einem etwas weiteren Sinn aber trotzdem nicht sehr groß. Diese Menge muss nämlich eine sogenannte Lebesgue-Nullmenge sein (siehe 5.3.1). Der Beweis sprengt unsere Möglichkeiten allerdings bei Weitem.

Übungsaufgabe 253. (T) Gegeben sei die reelle Funktion

$$f(x) = (x^2 - 1)e^x.$$

Fertigen Sie eine möglichst aussagekräftige Skizze des Graphen dieser Funktion an. Dazu ist es u.a. hilfreich, sich folgendes zu überlegen:

- Wo ist f überhaupt definiert?
- Wie viele Nullstellen hat f ?
- Hat der Graph von f Sprünge oder Knicke?
- Hat f Polstellen und wenn ja, wie verhält sich f in der Nähe eines Pols?
- Wie verhält sich $f(x)$ für $x \rightarrow \infty$? Hat f sogar Asymptoten und wenn ja, wie sehen diese aus?
- Ist f (stückweise) monoton?
- Hat f lokale bzw. globale Extremstellen?
- Wie sieht der Wertebereich von f aus, werden manche Wert gar nicht oder mehrfach angenommen?

Übungsaufgabe 254. (T) Wie Übungsaufgabe 253, aber mit $f(x) = \ln(1 + x^2) - \frac{3x}{5} + 1$.

Übungsaufgabe 255. (T) Wie Übungsaufgabe 253, aber mit $f(x) = \ln(\cos x)$.

Übungsaufgabe 256. (T) Wie Übungsaufgabe 253, aber mit $f(x) = |x \ln x|$.

4.4.2 Konvexität, Krümmung und zweite Ableitung

Inhalt in Kurzfassung: Eine Funktion heißt konvex, wenn sie sich – grob gesprochen – nach oben hin krümmt, konkav im umgekehrten Fall. In auffälliger Analogie zu den Ergebnissen aus 4.4.1 folgt aus $f''(x) \geq 0$ für alle x auf einem ganzen Intervall die Konvexität von f (analog Konkavität bei \leq und strenge Konvexität/Konkavität bei $>$ bzw. $<$). Kurz und ohne strenge Beweise wird auch der Begriff der Krümmung angeschnitten.

Ähnlich wie die erste Ableitung einer Funktion eng mit ihrem Monotonieverhalten zusammenhängt, gibt die zweite Ableitung über Konvexität bzw. Konkavität Auskunft. Zunächst die Definition dieser Begriffe.

Definition 4.4.2.1. Eine reelle Funktion $f : D \rightarrow \mathbb{R}$ mit zusammenhängendem Definitionsbereich D heißt **konvex**, wenn für je zwei Punkte $a < b \in D$ und alle $t \in (0, 1)$ stets

$$f((1-t)a + tb) \leq (1-t)f(a) + tf(b).$$

gilt. Kann \leq durch $<$, \geq bzw. $>$ ersetzt werden, so heißt f **streng konvex**, **konkav** bzw. **streng konkav**.

Um Missverständnisse zu vermeiden, sei hier auch erwähnt, was man unter einer **konvexen Menge** versteht. Im \mathbb{R}^n ist das eine Teilmenge $K \subseteq \mathbb{R}^n$ mit der Eigenschaft, dass mit je zwei Punkten $a, b \in \mathbb{R}^n$ auch die gesamte Verbindungslinie, das ist die Menge aller **Konvexkombinationen** $x = (1-t)a + tb$

mit $t \in [0, 1]$ in K enthalten ist. Für $n = 1$, d.h. auf der Zahlengeraden \mathbb{R} sind genau die zusammenhängenden Mengen konvex. In \mathbb{R}^2 sind beispielsweise volle Dreiecke, Quadrate oder Kreisscheiben konvex, nicht aber Kreislinien (die immerhin noch zusammenhängend sind).³

Der Bezug zwischen Konvexität einer Menge und einer Funktion besteht darin, dass der Bereich über dem Graphen einer konvexen Funktion eine konvexe Teilmenge der Ebene bildet, genauer:

Übungsaufgabe 257. (E) *Beweisen Sie folgende Aussage: Ist $f : D \rightarrow \mathbb{R}$ eine konvexe Funktion auf dem zusammenhängenden Bereich D , dann ist $K := \{(x, y) : x \in D, f(x) \leq y\}$ eine konvexe Teilmenge der Ebene.*

Doch zurück zur Konvexität einer reellen Funktion f . Man überlegt sich, was diese Eigenschaft für eine Funktion anschaulich bedeutet, nämlich dass die geradlinige Verbindung (Sekante) zwischen zwei Punkten auf dem Graphen von f nirgends unterhalb des Graphen liegt. Bei Konkavität gilt Entsprechendes mit „oberhalb“ statt „unterhalb“. Typische Beispiele konvexer Funktionen sind also solche, die nach oben hin gekrümmt sind wie die **Exponentialfunktionen** \exp_a oder die **Potenzfunktionen** $x \mapsto x^\alpha$ mit $\alpha > 1$. Konkav sind beispielsweise **Logarithmen** \log_a und die **Potenzfunktionen** p_α mit $0 < \alpha < 1$.

In Analogie zu Satz 4.4.1.1 über erste Ableitung und Monotonie, lässt sich über zweite Ableitung und Konvexität genauer sagen:

Satz 4.4.2.2. *Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und zweimal differenzierbar auf (a, b) , dann gelten folgende Implikationen:*

1. *Gilt $f''(x) \geq 0$ für alle $x \in (a, b)$, so ist f auf $[a, b]$ konvex.*
2. *Gilt $f''(x) > 0$ für alle $x \in (a, b)$, so ist f auf $[a, b]$ streng konvex.*
3. *Gilt $f''(x) \leq 0$ für alle $x \in (a, b)$, so ist f auf $[a, b]$ konkav.*
4. *Gilt $f''(x) < 0$ für alle $x \in (a, b)$, so ist f auf $[a, b]$ streng konkav.*

Beweis. (Idee) Für den Beweis (z.B. der ersten Behauptung) kann man verwenden, dass nach dem früheren Resultat bei $f'' \geq 0$ die Ableitung f' monoton wachsend ist. Damit lässt sich mit Hilfe des Mittelwertsatzes zeigen, dass die Steigung von Sekanten von links nach rechts zunehmen muss. Mit etwas Rechenaufwand lässt sich, der Anschauung folgend, daraus auf die Konvexität von f schließen. Wir verzichten auf Details. \square

Entsprechend ist für **Wendepunkte** x die Beziehung $f''(x) = 0$ typisch. Was genau ein Wendepunkteiner Funktion f ist lässt sich unterschiedlich definieren. Anschaulich denken wir an eine Stelle, wo eine Bereich von Konvexität

³Am einfachsten lässt sich der Zusammenhang einer Kreislinie damit begründen, dass sie stetiges Bild (nämlich z.B. unter der Abbildung $t \mapsto (\cos(t), \sin(t))$) der zusammenhängenden Menge \mathbb{R} ist. Und stetige Abbildungen bilden zusammenhängende Mengen stets wieder auf solche ab. Der Beweis ist leicht, setzt aber eine sorgfältige Behandlung der am Ende von 1.5.9 angedeuteten Begriffe voraus.

links von x in einen von Konkavität rechts von x übergeht oder umgekehrt, wo also f'' das Vorzeichen wechselt. Man könnte aber auch an Funktionen interessiert sein, für die f'' nicht einmal existiert (ähnlich wie wir Extrema auch für nicht differenzierbare Funktionen definieren), sehr wohl aber eine Tangente t an f in einem Punkt x , wo $f < t$ auf der einen Seite von x und $f > t$ auf der anderen gilt oder umgekehrt. Wir wollen das hier jedoch nicht vertiefen.

Ähnlich wie Monotonie allein schon gewisse Stetigkeitseigenschaften impliziert, haben Konvexität und Konkavität sogar noch stärkere Implikationen dieser Art: Konvexe Funktionen $f : D \rightarrow \mathbb{R}$ auf offenem D sind stetig, und sogar die Differenzierbarkeit von f kann nur an höchstens abzählbar vielen Punkten von D verletzt sein.

Offenbar hat die zweite Ableitung f'' mit der **Krümmung** des Graphen von f zu tun. Man kommt aber zu keinem sinnvollen Begriff, wenn man versucht, die Krümmung als Zahlenwert alleine über f'' zu definieren. Man halte sich das Beispiel $f(x) = x^2$ vor Augen, wo der Graph eine Parabel mit Scheitelpunkt $(0, 0)$ ist. Diese Funktion f hat konstante zweite Ableitung $f''(x) = 2$, während das, was man als Krümmung des Graphen zu bezeichnen geneigt ist, variiert, nämlich beim Scheitel in $x = 0$ am größten ist und für $|x| \rightarrow \infty$ immer geringer wird. Doch welchen präzisen Begriff von Krümmung hat man dabei vor Augen? Naheliegender ist es, einen Kreis an den Graphen zu legen, der diesen nicht nur im Sinne einer Tangente berührt, sondern sich sogar im Sinne eines Taylorpolynoms zweiten Grades anschmiegt (**Schmiegunskreis**). Je kleiner der Radius dieses Kreises, desto größer die Krümmung. Also definiert man die Krümmung κ als den Kehrwert des Radius so eines Schmiegunskreises. Man kann zeigen, dass dieser Wert tatsächlich nicht von f'' alleine abhängt, sondern auch von f' (je größer f' , desto kleiner die Krümmung – man denke an die quadratische Parabel). Als korrekte Formel für die Krümmung $\kappa = \kappa(f, x_0)$ des Graphen von f im Punkt $(x_0, f(x_0))$ erweist sich

$$\kappa = \frac{|f''(x_0)|}{(1 + f'(x_0)^2)^{\frac{3}{2}}}.$$

Übungsaufgabe 258. (T) Berechnen Sie die Krümmung κ der Parabel $f(x) := x^2$ in einem beliebigen Punkt (x_0, x_0^2) .

Übungsaufgabe 259. (E) Beschreiben Sie einen oberen Halbkreis mit Radius r als Funktion f und nehmen Sie ein beliebiges x_0 aus dem Inneren des Definitionsbereichs von f an.

1. Welchen Wert sollte κ aus der obigen Formel annehmen, wenn die Formel wirklich beschreibt, was sie beschreiben soll?
2. Überprüfen Sie das!

Übungsaufgabe 260. (T) Berechnen Sie für die Ellipse mit Gleichung $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ (d.h. mit Halbachsen $a, b > 0$) die Krümmung an den Punkten $(\pm a, 0)$ und $(0, \pm b)$ und interpretieren Sie Ihr Ergebnis. (Stichwort: Schmiegunskreise.)

4.4.3 Extremwertbestimmung und höhere Ableitungen

Inhalt in Kurzfassung: Bisher haben wir das Verschwinden der Ableitung in einem Punkt x_0 , also $f'(x_0) = 0$, (in Verbindung mit anderen Voraussetzungen) als notwendige Bedingung für das Vorliegen eines Extremums von f in x_0 kennen gelernt. Unter Zuhilfenahme höherer Ableitungen erhält man genauere Aussagen, insbesondere auch hinreichende Bedingungen.

Schon früher haben wir gesehen, dass z.B. für die Funktion $f(x) := x^3$ bei $x_0 = 0$ zwar die Ableitung den Wert $f'(0) = 0$ hat, trotzdem aber kein Extremum vorliegt. Das gleiche gilt für alle Potenzfunktionen $p_n : x \mapsto x^n$ mit ungeradem $n \geq 3$, während bei geradem n sehr wohl $x_0 = 0$ eine Extremstelle ist. Dieses Phänomen wollen wir besser verstehen. Der Satz von Taylor ermöglicht eine befriedigende Analyse. Wir müssen auch den Fall im Auge haben, dass an der Stelle x_0 möglicherweise nicht nur die erste, sondern mehrere Ableitungen bis hin zur n -ten den Wert $f^{(k)}(x_0) = 0$ haben, $k = 1, 2, \dots, n$. Sei also die $n + 1$ -te Ableitung die erste mit $f^{(n+1)}(x_0) \neq 0$.

Wir wollen den Fall $f^{(n+1)}(x_0) > 0$ ausführlich besprechen, analog den Fall $f^{(n+1)}(x_0) < 0$. In diesem Fall muss $f^{(n)}$ wenigstens in einer Umgebung von x_0 definiert und an der Stelle x_0 strikt monoton sein in dem Sinn, dass es ein $\delta > 0$ gibt mit $f^{(n)}(x) < f^{(n)}(x_0) = 0$ für $x_0 - \delta < x < x_0$ und $f^{(n)}(x) > f^{(n)}(x_0) = 0$ für $x_0 < x < x_0 + \delta$. Der Satz von Taylor besagt

$$f(x) = f(x_0) + \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n$$

(es gilt ja $f^{(k)}(x_0) = 0$ für $k = 1, 2, \dots, n - 1$) mit einem ξ zwischen x_0 und x . Wir unterscheiden zwei Fälle, n gerade oder ungerade. Für gerades n ist der Potenzausdruck $(x - x_0)^n$ immer positiv, während $f^{(n)}(\xi)$ bei $\xi = x_0$ das Vorzeichen wechselt. Also bekommen wir $f(x) > f(x_0)$ für $x > x_0$ und $f(x) < f(x_0)$ für $x < x_0$. Folglich ist x_0 keine Extremstelle. Für ungerades n ist die Situation umgekehrt. Sowohl $(x - x_0)^n$ als auch $f^{(n)}$ wechseln bei x_0 das Vorzeichen. In jedem Fall gilt daher $f(x) > f(x_0)$ für $x \neq x_0$ mit $x_0 - \delta < x < x_0 + \delta$. Folglich ist x_0 Stelle eines lokalen Minimums von f , entsprechend eines Maximums, sofern $f^{(n+1)}(x_0) < 0$. Damit haben wir den folgenden Satz bewiesen.

Satz 4.4.3.1. *Sei die reelle Funktion $f : D \rightarrow \mathbb{R}$ gegeben und $x_0 \in D$ ein innerer Punkt von D . In x_0 existiere die $n + 1$ -te Ableitung von f (sogar in einer Umgebung von x_0 folglich die n -te). Wir nehmen an, dass die erste bis n -te Ableitung von f in x_0 verschwinde: $f'(x_0) = f''(x_0) = \dots = f^{(n)}(x_0) = 0$, die $n + 1$ -te jedoch nicht: $f^{(n+1)}(x_0) \neq 0$. Dann gilt die folgende Fallunterscheidung:*

1. *Ist n gerade, so hat f in x_0 kein lokales Extremum.*
2. *Ist n ungerade und $f^{(n+1)}(x_0) > 0$, so hat f in x_0 ein lokales Minimum.*
3. *Ist n ungerade und $f^{(n+1)}(x_0) < 0$, so hat f in x_0 ein lokales Maximum.*

Man beachte, dass für das Einstiegsbeispiel $f(x) := x^3$ speziell $n = 2$ zu setzen ist, um aus dem Satz die interessante Aussage (nämlich dass $x_0 := 0$ keine Extremstelle ist) abzulesen.

Übungsaufgabe 261. (E) Sei $n \in \mathbb{N}$ beliebig vorgegeben.

1. Sei $n = 2k$ gerade. Geben Sie ein f und x_0 an, so dass alle Voraussetzungen und somit auch Aussage 1 in Satz 4.4.3.1 gelten.
2. Sei $n = 2k + 1$ ungerade. Geben Sie ein f und x_0 an, so dass alle Voraussetzungen und Aussage 2 in Satz 4.4.3.1 gelten.
3. Sei $n = 2k + 1$ ungerade. Geben Sie ein f und x_0 an, so dass alle Voraussetzungen und Aussage 3 in Satz 4.4.3.1 gelten.

Übungsaufgabe 262. (E) Verwenden Sie die Funktion $f(x) := e^{-\frac{1}{x^2}}$ für $x \neq 0$ und $f(0) := 0$ aus 4.2.7, um zu illustrieren, dass es unendlich oft differenzierbare Funktionen gibt, deren Extremwertverhalten nicht mit Satz 4.4.3.1 allein geklärt werden kann.

Übungsaufgabe 263. (P) Wieviele lokale Extremstellen kann ein Polynom vom Grad n , das auf ganz \mathbb{R} definiert wird, haben? Was ist die maximale Zahl, und was ist sonst noch möglich? Illustrieren Sie Ihre Überlegungen auch durch Skizzen.

Übungsaufgabe 264. (P) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ von der Gestalt $f(x) = \prod_{i=1}^n (x - \alpha_i)$ mit $\alpha_1 < \alpha_2 < \dots < \alpha_n$. Wie viele Nullstellen hat die Ableitung f' ?

Übungsaufgabe 265. (T) Ein Himmelskörper der Masse m (Sie dürfen $m = 1$ setzen), der sich im Abstand $r > 0$ von der Sonne befindet, besitzt das sogenannte „effektive Potential“

$$V(r) = -G \frac{mM}{r} + \frac{L^2}{2r^2m} - \frac{ML^2G}{r^3mc^2},$$

wobei $M, L > 0$ feste Parameter sind ($M =$ Sonnenmasse, $L =$ Drehimpuls), G die Gravitationskonstante und c die Lichtgeschwindigkeit bezeichnen.

Erläuterung für physikalisch Interessierte: Die ersten beiden Summanden in der Formel für $V(r)$ entsprechen der Newtonschen Gravitation, der dritte ist eine relativistische Korrektur.

1. Bestimmen Sie das asymptotische Verhalten (was genau könnte damit gemeint sein?) und skizzieren Sie $V(r)$.
2. Bestimmen Sie die Minima von $V(r)$, indem Sie $V'(r) = 0$ setzen und Ihre Resultate aus 1. verwenden.

4.4.4 Fixpunktiteration, Dynamik und Chaos

Inhalt in Kurzfassung: Wir setzen die Überlegungen aus 3.1.6 über die Konvergenz von Iterationsfolgen mit $a_{n+1} = T(a_n)$ fort, nun unter dem Gesichtspunkt der Differenzierbarkeit von T . Die Kontraktionseigenschaft hängt nun eng mit der Bedingung $|T'| < 1$ zusammen. Abschließend kommen wir auch kurz auf die Möglichkeit chaotischen, d.h. extrem komplizierten Verhaltens, selbst bei relativ einfachen Transformationen T , zu sprechen.

Wir kommen nochmals auf die Fixpunktiteration aus 3.1.6 zurück. Sei also $T : D \rightarrow D$ mit $D \subseteq \mathbb{R}$. Wir interessieren uns für Folgen mit Gliedern $a_n = T^{(n)}(a_0)$, die also mit einem gewissen Startwert $a_0 \in D$ der Rekursion $a_{n+1} := T(a_n)$ für alle $n \in \mathbb{N}$ genügen. Der Einfachheit halber setzen wir voraus, dass D abgeschlossen ist. Wir wollen der in 3.1.6 thematisierten Frage nachgehen, ob Iteration von T Konvergenz gegen einen Fixpunkt bewirkt, nun unter dem Gesichtspunkt der Differenzierbarkeit. Tatsächlich gilt der folgende Satz:

Satz 4.4.4.1. *Sei $D \subseteq \mathbb{R}$ abgeschlossen, $T : D \rightarrow D$ differenzierbar in $x_0 \in D$ mit $|T'(x_0)| < 1$ und x_0 ein Fixpunkt von T .⁴ Dann gilt:*

1. *Es gibt eine Umgebung U von x_0 derart, dass für alle $x \in U$ die Iterationsfolge mit Startwert x gegen x_0 konvergiert.*
2. *Ist D zusammenhängend und T auf ganz D differenzierbar mit $|T'(x)| \leq \lambda$ für alle $x \in D$ mit einem $\lambda < 1$, dann ist T eine Kontraktion mit λ als Kontraktionskonstante. Folglich gibt es genau einen Fixpunkt von T , und für jedes $a_0 \in D$ konvergiert die Iterationsfolge $a_{n+1} := T^{(n)}(a_0)$ gegen diesen Fixpunkt.*

Beweis. 1. Aus der Definition der Ableitung mittels Differenzenquotient folgt wegen $|T'(x_0)| < 1$, dass es ein $\lambda < 1$ und eine Intervallumgebung $U = (x_0 - \delta, x_0 + \delta) \subseteq U$, $\delta > 0$, von x_0 gibt mit folgender Eigenschaft: Für alle $x \in U$ gilt die Ungleichung

$$\left| \frac{T(x) - x_0}{x - x_0} \right| = \left| \frac{T(x) - T(x_0)}{x - x_0} \right| \leq \lambda < 1$$

und somit $|T(x) - x_0| \leq \lambda|x - x_0| < |x - x_0|$. Diese Beziehung zeigt auch $T(U) \subseteq U$. Also liegen mit einem Startwert $a_0 \in U$ auch alle Glieder $a_n := T^{(n)}(a_0)$ der Iterationsfolge wieder in U und erfüllen $|a_n - x_0| \leq \lambda^n|a_0 - x_0|$. Das sieht man mittels Induktion mit Induktionsanfang $|a_0 - x_0| = \lambda^0|a_0 - x_0|$ und Induktionsschritt

$$|a_{n+1} - x_0| = |T(a_n) - x_0| \leq \lambda|a_n - x_0| \leq \lambda \cdot \lambda^n|x - x_0| = \lambda^{n+1}|x - x_0|.$$

Wegen $\lambda < 1$ folgt daraus $\lim_{n \rightarrow \infty} T^{(n)}(x) = \lim_{n \rightarrow \infty} a_n = x_0$.

⁴Ist D zusammenhängend und x_0 ein Randpunkt von D , so genügt auch einseitige Differenzierbarkeit von T in x_0 .

2. Nach dem Banachschen Fixpunktsatz (Kontraktionsprinzip, Satz 3.1.6.1) genügt es, für $x, x' \in D$ die Ungleichung $|T(x) - T(x')| \leq \lambda|x - x'|$ zu beweisen. Diese Ungleichung ist für $x = x'$ trivial, und sonst ergibt sie sich unmittelbar aus dem Mittelwertsatz 4.2.2.1, wonach es eine Zwischenstelle ξ zwischen x und x' gibt mit $T(x) - T(x') = T'(\xi)(x - x')$, also

$$|T(x) - T(x')| = |T'(\xi)(x - x')| \leq |T'(\xi)| \cdot |x - x'| \leq \lambda|x - x'|.$$

□

Der Beweis von Satz 4.4.4.1 zeigt, dass die Konvergenz der Iterationsfolge sehr schnell erfolgt, wenn $\lambda > 0$ sehr klein ist.

Übungsaufgabe 266. (E) Rekapitulieren Sie den Beweis von Satz 4.4.4.1, um die Fußnote betreffend einseitige Differenzierbarkeit von T an Randpunkten von D zu bestätigen.

Sind die genannten Konvergenzvoraussetzungen nicht erfüllt, so kann die **Dynamik** von Iterationsfolgen (das heißt ihr Verhalten im Laufe der Zeit, sprich für $n \rightarrow \infty$) extrem unübersichtlich sein, woraus sich das damit verbundene populäre Schlagwort **Chaos** erklärt. Ein berühmtes Beispiel ist das folgende.

Wir betrachten die Funktion $T_c(x) := cx(1 - x)$ auf dem Einheitsintervall $[0, 1]$ für jene Werte von c , für die $T_c : [0, 1] \rightarrow [0, 1]$ gesichert ist, also für $c \in [0, 4]$. Für kleine Werte von c lässt sich das Verhalten der Iterationsfolgen $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$, $x_{n+1} := T_c(x_n)$, in Abhängigkeit vom Anfangswert x_0 noch sehr gut überblicken. Wenn sich c aber dem Wert 4 annähert, wird die Situation immer verwickelter. Zur Übung kann man sich mit folgenden Fragen befassen.

- Welche Fixpunkte hat T_c ? (Es gibt nie mehr als 2 Fixpunkte.)
- Für welche Anfangswerte konvergiert die Iterationsfolge gegen welchen Fixpunkt?
- Gibt es abgesehen von Fixpunkten periodische Punkte?
- Für welche Anfangswerte landet man irgendwann in einem periodischen Punkt?
- Gibt es Iterationsfolgen, die sich irgendwann periodischen Iterationsfolgen annähern?
- Gibt es Iterationsfolgen, die keinem der obigen Typen zuzuordnen sind?

Je nach Wert von c sind diese Fragen unterschiedlich schwierig. In manchen Fällen werden Sie vermutlich nicht alle Fragen beantworten können. Sinnvoll ist die Unterscheidung im Sinne der folgenden Übungsaufgabe:

Übungsaufgabe 267. Untersuchen Sie obige Fragen für c aus folgenden Bereichen:

1. (T) $0 \leq c \leq 1$. (Dieser Fall sollte sehr einfach sein.)
2. (T) $1 < c \leq 2$. (Immer noch einfach.)
3. (P) $2 < c \leq 1 + \sqrt{3}$. (Schon anspruchsvoller, aber lösbar. Welche Rolle spielt dabei $1 + \sqrt{3}$?)
4. (E) $1 + \sqrt{3} < \lambda \leq 3$. (Ähnliches Ergebnis wie für $2 < c \leq 1 + \sqrt{3}$, die Analyse ist aber um einiges verwickelter.)
5. (E) $3 < c \leq 4$. (Wer diesen Fall vollständig versteht, kann berühmt werden. Versuchen Sie wenigstens herauszubekommen, was sich gegenüber den bisherigen Fällen verändert.)

4.4.5 Das Newton-Verfahren

Inhalt in Kurzfassung: Das Newtonverfahren dient der Approximation von Nullstellen einer Funktion f (d.h. der Lösung auch sehr allgemeiner, nichtlinearer Gleichungen) entsprechend dem Paradigma der Differentialrechnung: Man ersetzt f durch die lineare Approximation (Taylorpolynom ersten Grades) von f in der Nähe einer Nullstelle. Die leicht zu bestimmende Nullstelle dieser linearen Approximation verwendet man als Ausgangspunkt für die nächste Approximation der gesuchten Nullstelle von f . Unter ziemlich schwachen Bedingungen lässt sich rasche Konvergenz des Verfahrens garantieren. Als Anwendungsbeispiel wird die rasche Approximation von Quadratwurzeln ausschließlich mit Hilfe von Grundrechnungsarten behandelt.

Eine für die Praxis besonders wichtige Iteration ist das **Newtonverfahren** zur näherungsweise Bestimmung von Nullstellen von Funktionen. Im Zusammenhang mit Polynomen wurde es bereits erwähnt. Es handelt sich dabei um eine Methode, die sich aber auf eine viel größere Klasse von Funktionen anwenden lässt. Die Grundidee besteht darin, dass erstens differenzierbare Funktionen durch lineare Funktionen (Tangenten an den Funktionsgraphen) gut approximiert werden können und zweitens die Nullstellen von linearen Funktionen sehr leicht zu bestimmen sind.

Sei die reelle Funktion $f : D \rightarrow \mathbb{R}$ also differenzierbar und $x_0 \in D$. Dann ist das erste Taylorpolynom $p = p_{f, x_0, 1}$ von f bei x_0 gegeben durch $p(x) = f(x_0) + f'(x_0)(x - x_0)$ (lineare Approximation von f). Die Nullstelle von p findet man durch Lösung der Gleichung $0 = f(x_0) + f'(x_0)(x - x_0)$ nach x . Die Lösung x_1 ist gegeben durch $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$. Ist p im relevanten Bereich eine gute Approximation von f , so hat man guten Grund zur Annahme, dass die Nullstelle x_1 von p in der Nähe einer Nullstelle von f liegt. Iteration der Vorgangsweise generiert eine rekursive Folge, die durch $x_{n+1} := T(x_n)$ mit

$$T(x) := x - \frac{f(x)}{f'(x)}$$

definiert ist. Spricht man vom Newtonverfahren mit Anfangspunkt x_0 , so ist diese Folge gemeint. Natürlich sind Probleme denkbar. Zum Beispiel könnte ein Folgenglied aus dem Definitionsbereich D von f fallen oder die Ableitung im Nenner 0 werden. Unter geeigneten, nicht sehr starken Voraussetzungen, konvergiert das Newton-Verfahren aber ganz hervorragend: Jede Nullstelle ξ von f mit $f'(\xi) \neq 0$ ist wegen $T(\xi) = \xi - \frac{f(\xi)}{f'(\xi)} = \xi$ ein Fixpunkt von T . Ist beispielsweise f zweimal stetig differenzierbar, so ist T (abgesehen von Nullstellen von f') selbst wenigstens einmal stetig differenzierbar mit der Ableitung

$$T'(x) = 1 - \frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2}.$$

Wegen $f(\xi) = 0$ ist auch $T'(\xi) = 0$. Geben wir uns eine beliebig kleine Kontraktionskonstante $\lambda > 0$ vor, so folgt aufgrund der Stetigkeit von T' für alle x aus einer hinreichend kleinen Umgebung von ξ die Ungleichung $|T'(x)| < \lambda$. Nach unseren Überlegungen aus 4.4.4 konvergiert das Newtonverfahren in diesem Bereich daher extrem rasch. Man spricht von quadratischer Konvergenz des Verfahrens.

Satz 4.4.5.1. *Die zweimal differenzierbare Funktion $f : D \rightarrow \mathbb{R}$ möge in $\xi \in D$ eine Nullstelle haben, wobei ξ innerer Punkt von D sei. Außerdem sei $f'(\xi) \neq 0$. Dann gibt es eine Umgebung $U = (\xi - \delta, \xi + \delta) \subseteq D$, $\delta > 0$, von ξ derart, dass für jeden Startwert $x_0 \in U$ das Newtonverfahren (bei stetigem f'' sogar sehr schnell) gegen ξ konvergiert.*

Als Beispiel wollen wir die Quadratwurzel \sqrt{a} eines gegebenen $a > 0$ berechnen. Ein naiver Zugang wäre **Intervallschachtelung**. D.h. man beginnt mit zwei Approximationen $x_0 < y_0$ für \sqrt{a} mit $x_0^2 < a < y_0^2$, betrachtet das arithmetische Mittel $m_0 := \frac{x_0 + y_0}{2}$ und testet, ob $m_0^2 < a$ oder $m_0^2 > a$ gilt. Im ersten Fall arbeitet man mit $x_1 := m_0$ und $y_1 := y_0$ weiter, im zweiten mit $x_1 := x_0$ und $y_1 := m_0$ etc.

Übungsaufgabe 268. (T) Berechnen Sie $\sqrt{2}$ auf diese Weise auf zwei Nachkommastellen genau.

Das Newton-Verfahren ist wesentlich effektiver. Denn a ist die Nullstelle von $f(x) = x^2 - a$. Also hat man nur einen geeigneten Startwert x_0 zu wählen (in diesem Fall ist jedes $x_0 > 0$ geeignet) und nach der Rekursion

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - a}{2x_n} = \frac{x_n}{2} + \frac{a}{2x_n},$$

also $T(x) = \frac{x}{2} + \frac{a}{2x}$ vorzugehen. Nehmen wir als Beispiel $a = 2$ und $x_0 = 1$, so liefert die Rekursion $x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}$ die Werte $x_1 = \frac{3}{2} = 1,5$, $x_2 = \frac{17}{12} = 1,41\bar{6}$, $x_3 = \frac{577}{408} = 1,414215686 \dots$ etc., also ziemlich rasche Konvergenz gegen $\sqrt{2} = 1,414213562 \dots$. Man beachte, dass sich die einzelnen Rechenschritte bei dieser Berechnung der Quadratwurzel allein mit Grundrechnungsarten durchführen lassen.

Übungsaufgabe 269. (T) Berechnen Sie $\sqrt[3]{2}$ mittels Newtonverfahren und ohne Taschenrechner auf zwei Nachkommastellen genau.

Übungsaufgabe 270. (T) Gesucht sind Lösungen der Gleichung $3x = e^x$.

1. Begründen Sie, warum es im Intervall $[0, 1]$ genau eine Lösung gibt.
2. Berechnen Sie diese Lösung ohne Taschenrechner auf zumindest eine Nachkommastelle genau.

Hinweis: Verwenden Sie Intervallschachtelung und nützen Sie aus, dass die Exponentialreihe sehr schnell konvergiert, d.h. dass Sie nur wenige Terme addieren müssen, um die geforderte Genauigkeit zu erreichen.

3. Suchen Sie die Lösung mit Hilfe des Newtonverfahrens.

Kapitel 5

Integralrechnung

Den Ausgangspunkt der Integralrechnung bildet das Anliegen, Flächen zu berechnen, und zwar auch solche, die durch krumme Kurven begrenzt sind. Die wesentliche Schwierigkeit hat man schon dann bewältigt, wenn man den Fall beherrscht, dass in einem Rechteck nur eine der vier Seiten durch eine krumme Linie ersetzt wird. Lässt sich diese Linie als Graph einer Funktion deuten, so ist man beim Begriff des Riemannintegrals angelangt, dem der erste Abschnitt des Kapitels (5.1) gewidmet ist. Eine der wichtigsten Erkenntnisse besteht in dem engen Zusammenhang zur Differentialrechnung. Er drückt sich im Hauptsatz der Differential- und Integralrechnung (5.2) aus, aus dem gemeinsam mit bereits bekannten Ableitungsregeln sich die meisten Integrationsregeln als Folgerung ergeben. Für die mehrdimensionale Integration wie auch für die Wahrscheinlichkeitstheorie ist es von Interesse, das Riemannintegral unter einem neuen Gesichtspunkt zu sehen, der weiter reichende Verallgemeinerungen ermöglicht. Und zwar handelt es sich dabei um die Integration bezüglich eines Maßes. Sie ist der Gegenstand von 5.3. Der letzte Abschnitt (5.4) schließlich bringt weitere Aspekte und Anwendungen der Integralrechnung.

5.1 Das Riemannintegral

Die Grundaufgabe des Riemannintegrals lässt sich deuten als Berechnung einer Fläche, die in der Ebene zwischen x -Achse, den senkrechten Geraden $x = a$, $x = b$ und dem Graphen einer reellen Funktion $f : [a, b] \rightarrow \mathbb{R}$ liegt, wobei Flächenteile unterhalb der x -Achse (also mit negativen y -Werten) mit negativem Vorzeichen versehen sind. Die Grundidee besteht darin, das Intervall $[a, b]$ zu unterteilen und auf jedem der Teilstücke die entstehenden schmalen Streifen, deren krumme Grenze hoffentlich nur mehr sehr kurz ist, durch schmale Rechtecksflächen, die leicht berechenbar sind, zu ersetzen. Wird die Unterteilung hinreichend fein, so hofft man, sollte auch der Approximationsfehler beliebig klein gemacht werden können. Diese Grundidee lässt sich mit Ober- und Untersummen (5.1.1) wie auch mit Riemannsummen (5.1.2) präzisieren. Nach der Herleitung einfa-

cher Eigenschaften (5.1.3) zeigen wir, dass die Konzepte jedenfalls im Fall eines stetigen f zum Erfolg führen (5.1.4).

5.1.1 Ober- und Untersummen

Inhalt in Kurzfassung: Als ersten Zugang zum (Riemann-)Integral wählen wir jenen über Ober- und Untersummen. Für eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ definiert jede Zerlegung Z eine Obersumme $O(f, Z)$ und eine Untersumme $U(f, Z)$. Für $f \geq 0$ sind das obere bzw. untere Approximationen der Fläche zwischen der x -Achse, dem Graphen von f und den Senkrechten $x = a$ und $x = b$. Nähern sich diese beiden Approximationen für feine Zerlegungen Z einem gemeinsamen Wert an, so definiert man diesen als das Riemann-Integral $\int_a^b f(x) dx$ (auch bestimmtes Integral) des Integranden f über dem Integrationsbereich $[a, b]$.

Unter einer **Zerlegung** von $[a, b]$, $a \leq b$, verstehen wir eine endliche Teilmenge $Z = \{a = x_0 \leq x_1 \leq \dots \leq x_n = b\}$ von $[a, b]$, die sowohl a als auch b enthält. Die größte der Differenzen $x_i - x_{i-1}$, $i = 1, 2, \dots, n$, nennen wir die **Feinheit** von Z und schreiben dafür

$$\phi(Z) := \max\{x_i - x_{i-1} : 1 \leq i \leq n\}.$$

(Im nicht sehr interessanten Fall $a = b$ und $n = 0$ setzen wir $\phi(Z) := 0$.) Sei $a \leq b$ und $f : [a, b] \rightarrow \mathbb{R}$ beschränkt. Die zugehörige **Obersumme** ist definiert durch

$$O(f, Z) := \sum_{i=1}^n (x_i - x_{i-1}) \sup_{x_{i-1} \leq x \leq x_i} f(x),$$

die zugehörige **Untersumme** durch

$$U(f, Z) := \sum_{i=1}^n (x_i - x_{i-1}) \inf_{x_{i-1} \leq x \leq x_i} f(x).$$

Klarerweise gilt $U(f, Z) \leq O(f, Z)$ (und, im Fall $a = b$ und $n = 0$, $U(f, Z) = O(f, Z) = 0$). Wir interessieren uns für das Verhalten von Ober- und Untersummen, wenn $\phi(Z)$ gegen 0 konvergiert. Dafür ist es zweckmäßig, den Begriff der **Verfeinerung** Z_2 einer Zerlegung Z_1 zu definieren, der schlicht $Z_1 \subseteq Z_2$ bedeutet. Man überlegt sich sehr leicht, dass in diesem Fall stets $U(f, Z_1) \leq U(f, Z_2) \leq O(f, Z_2) \leq O(f, Z_1)$ gilt.

Übungsaufgabe 271. (P) Begründen Sie sorgfältig, warum für Zerlegungen $Z_1 \subseteq Z_2$ und beschränktes f tatsächlich stets $U(f, Z_1) \leq U(f, Z_2) \leq O(f, Z_2) \leq O(f, Z_1)$ gilt.

Bei Verfeinerungen rücken Ober- und Untersumme einander also näher. Bedenkt man noch, dass es zu je zwei beliebigen Zerlegungen Z_1 und Z_2 (die nicht

unbedingt Verfeinerungen voneinander sein müssen) stets eine gemeinsame Verfeinerung $Z := Z_1 \cup Z_2$ gibt, so schließen wir auch in diesem Fall

$$U(f, Z_1) \leq U(f, Z) \leq O(f, Z) \leq O(f, Z_2).$$

Das bedeutet, dass alle Untersummen untere Schranken für alle Obersummen sind und, umgekehrt, alle Obersummen obere Schranken für alle Untersummen. Die interessanten Größen sind daher das Supremum aller Untersummen und das Infimum aller Obersummen. Man nennt sie das **untere Riemannintegral** (kurz auch **R-Integral**) bzw. das **obere Riemannintegral** von f über $[a, b]$. Wir schreiben dafür

$$\int_a^b f(x) dx := \sup_Z U(f, Z) \quad \text{bzw.} \quad \overline{\int}_a^b f(x) dx := \inf_Z O(f, Z).$$

Aus unseren bisherigen Überlegungen folgt $\int_a^b f(x) dx \leq \overline{\int}_a^b f(x) dx$. Die entscheidende Frage ist, ob Gleichheit erzielt werden kann. Sie ist Anlass für die folgende, für dieses Kapitel grundlegende Definition.

Definition 5.1.1.1. Sei $a \leq b$. Die beschränkte reelle Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt **Riemann-integrierbar (R-integrierbar)**, sofern

$$\int_a^b f(x) dx = \overline{\int}_a^b f(x) dx.$$

In diesem Fall heißt der gemeinsame Wert von oberem und unterem Riemann-Integral das **Riemannintegral (R-Integral)** oder auch **bestimmtes Integral** von f über $[a, b]$ und wird mit

$$\int_a^b f(x) dx$$

bezeichnet. Im Fall $a > b$ setzt man $\int_a^b f(x) dx := -\int_b^a f(x) dx$. Die Funktion f heißt in diesem Zusammenhang auch **Integrand**, das Intervall $[a, b]$ heißt **Integrationsbereich**.

Statt der Integrationsvariablen x und dx kann natürlich auch ein anderes Symbol verwendet werden, z.B. t und dt . Feste Konvention ist die Wahl des Buchstabens d (für **Differential**).

Übungsaufgabe 272. (P) Sei $f(x) := x$ auf $[a, b] := [0, 1]$ und $Z(n) := \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1 - \frac{1}{n}, 1\}$. Berechnen Sie die Ober- und Untersummen $O(f, Z(n))$ und $U(f, Z(n))$ und mit deren Hilfe das Integral $\int_0^1 x dx$.

Wir werden sehen, dass zum Beispiel alle stetigen Funktionen R-integrierbar sind. Um die Probleme, die dabei auftreten können, zu verstehen, ist es wichtig, auch ein Beispiel einer (beschränkten) Funktion kennen zu lernen, die nicht R-integrierbar ist. Das klassische Beispiel ist die Dirichletsche Sprungfunktion $\mathbf{1}_{\mathbb{Q}}$,

die wir schon in 3.1.4 kennengelernt haben, zum Beispiel auf der Definitionsmenge $[0, 1]$. Zu jeder Zerlegung Z von $[0, 1]$, egal wie fein, gilt für Unter- und Obersumme $U(f, Z) = 0$ und $O(f, Z) = 1$. Also ist f nicht R-integrierbar. Man beachte, dass dieses f an jeder Stelle von $[0, 1]$ unstetig ist. Tatsächlich werden wir noch ein Kriterium kennen lernen, welches zeigt, dass R-Integrierbarkeit davon abhängt, dass die Menge der Unstetigkeitsstellen eine sogenannte Lebesgue-Nullmenge, also im maßtheoretischen Sinn klein ist.

5.1.2 Riemannsummen

Inhalt in Kurzfassung: Wenn man in den Unter- und Obersummen aus 5.1.1 die auftretenden Infima bzw. Suprema durch irgendwelche Funktionswerte im entsprechenden Teilintervall ersetzt, erhält man sogenannte Riemannsummen. Ihre Werte liegen offenbar zwischen entsprechender Unter- und Obersumme. Das Hauptergebnis dieses Unterabschnitts enthält mehrere äquivalente Beschreibungen des Riemannintegrals.

Oft ist es zweckmäßig, wenn man in den Obersummen und Untersummen das Supremum bzw. Infimum durch spezielle Funktionswerte ersetzt. Um das zu präzisieren, führen wir den Begriff einer **Belegung** B zu einer Zerlegung $Z = \{a = x_0 < x_1 < \dots < x_n = b\}$ von $[a, b]$ ein. Darunter verstehen wir ein n -tupel $(\xi_1, \xi_2, \dots, \xi_n)$ mit $x_{i-1} \leq \xi_i \leq x_i$ für $i = 1, 2, \dots, n$. Die damit definierte Summe

$$S(f, Z, B) := \sum_{i=1}^n (x_i - x_{i-1}) f(\xi_i)$$

nennen wir die **Riemannsumme** von f zur Zerlegung Z und Belegung B . Klarerweise gilt $U(f, Z) \leq S(f, Z, B) \leq O(f, Z)$ für jede Belegung B zu Z . Riemannsummen approximieren das Integral $\int_a^b f(x) dx$ also nicht schlechter als Ober- und Untersumme. Tatsächlich führen die Zugänge über Ober- und Untersummen einerseits und Riemannsummen andererseits zum selben Begriff. Wir fassen dies in folgendem Satz zusammen:

Satz 5.1.2.1. *Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion und $\alpha \in \mathbb{R}$. Dann sind die folgenden Aussagen äquivalent.*

1. *Die Funktion f ist Riemann-integrierbar mit Wert $\alpha = \int_a^b f(x) dx$. (Explizit bedeutet das: Beide, das Supremum aller Untersummen $U(f, Z)$ und das Infimum aller Obersummen $O(f, Z)$, wobei Z alle Zerlegungen von $[a, b]$ durchläuft, stimmen mit α überein.)*
2. *Für jedes $\varepsilon > 0$ gibt es eine Zerlegung Z von $[a, b]$ mit $U(f, Z) \leq \alpha \leq O(f, Z)$ und $O(f, Z) - U(f, Z) < \varepsilon$.*
3. *Es gibt eine Folge von Zerlegungen $Z(n)$, $n \in \mathbb{N}$, von $[a, b]$ mit*

$$\lim_{n \rightarrow \infty} U(f, Z(n)) = \lim_{n \rightarrow \infty} O(f, Z(n)) = \alpha.$$

4. Für beliebiges $\varepsilon > 0$ gibt es ein $\delta > 0$ derart, dass für alle Zerlegungen Z mit $\phi(Z) < \delta$ sowohl $O(f, Z) - U(f, Z) < \varepsilon$ als auch $|S(f, Z, B) - \alpha| < \varepsilon$ für alle Belegungen B zu Z gilt.

Beweis. In 5.1.1 haben wir uns bereits überlegt, dass für je zwei Zerlegungen Z_1, Z_2 von $[a, b]$ stets $U(f, Z_1) \leq O(f, Z_2)$ gilt. Folglich gilt auch $\sup_Z U(f, Z) \leq \inf_Z O(f, Z)$, wobei Supremum und Infimum über sämtliche Zerlegungen Z von $[a, b]$ gebildet wird. Mit dieser Erkenntnis führen wir nun einen zyklischen Äquivalenzbeweis.

1 \Rightarrow 2: Ist die erste Bedingung erfüllt, so gibt es Zerlegungen Z_1 und Z_2 mit $U(f, Z_1) > \alpha - \frac{\varepsilon}{2}$ und $O(f, Z_2) < \alpha + \frac{\varepsilon}{2}$. Mit $Z := Z_1 \cup Z_2$ gilt laut bereits Bewiesenen $\alpha - \frac{\varepsilon}{2} < U(f, Z_1) \leq U(f, Z) \leq O(f, Z) \leq O(f, Z_2) < \alpha + \frac{\varepsilon}{2}$, woraus $O(f, Z) - U(f, Z) < \varepsilon$ folgt.

2 \Rightarrow 3: Zu jedem $n \in \mathbb{N}$ wählen wir $Z(n)$ mit $O(f, Z(n)) - U(f, Z(n)) < \frac{1}{n+1}$, was laut Bedingung 2 möglich ist. Diese Zerlegungen $Z(n)$ haben die gewünschte Eigenschaft.

3 \Rightarrow 4: Laut Voraussetzung finden wir ein n derart, dass für $Z := Z(n)$ die Bedingung $O(f, Z) - U(f, Z) < \frac{\varepsilon}{2}$ erfüllt ist. Sei $Z = \{a = x_0 \leq x_1 \leq \dots \leq x_n = b\}$. Dann gibt es ein positives $\delta < \frac{1}{2}(\min_{1 \leq i \leq n} |x_i - x_{i-1}|)$ derart, dass außerdem $n\delta \sup_{x \in [a, b]} |f(x)| < \frac{\varepsilon}{4}$. Für so ein δ kann man nach etwas Rechnung die gewünschten Eigenschaften nachweisen.

4 \Rightarrow 1: Weil laut erstem Absatz Untersummen nie größer als Obersummen sind, genügt es, irgendeine Zerlegung Z mit $O(f, Z) - U(f, Z) < \varepsilon$ zu finden. Laut Voraussetzung gibt es ein $\delta > 0$, so dass jedes Z mit $\phi(Z) < \delta$ diese Eigenschaft hat. Man kann also zum Beispiel $Z := \{a = x_0, x_1, \dots, x_n = b\}$ mit $n > \frac{b-a}{\delta}$ nehmen. \square

Übungsaufgabe 273. (P) Berechnen Sie das Riemann-Integral $\int_1^2 (2-x) dx$ mittels äquidistanter Riemann-Summen.

Übungsaufgabe 274. (P) Berechnen Sie das Riemann-Integral $\int_0^1 e^x dx$ mittels äquidistanter Riemann-Summen. Verwenden Sie dazu die Summenformel für die geometrische Reihe.

5.1.3 Einfache Eigenschaften

Inhalt in Kurzfassung: Monotone Funktionen sind stets Riemann-integrierbar. Integrale über angrenzende Intervalle können gestückelt werden und liefern als Wert die Summe der Einzelintegrale. Das Integral ist verträglich mit Addition und Skalarmultiplikation (Linearität), mit Ungleichungen (Monotonie) und mit Beträgen (der Betrag eines Integrals ist kleiner gleich dem Integral des Betrags).

Die einfachsten Beispiele R-integrierbarer Funktionen sind die konstanten. Denn für $f : [a, b] \rightarrow \mathbb{R}, x \mapsto c, c \in \mathbb{R}$, nehmen alle Ober-, Unter- und Riemannsummen den Wert $c(b-a)$ an, der somit der Wert des R-Integrals $\int_a^b c dx$ ist. Mit etwas mehr Rechenaufwand lässt sich zum Beispiel auch die R-Integrierbarkeit

der Funktion $f : x \mapsto x$ auf jedem beschränkten Intervall beweisen, z.B. auf $[0, 1]$ mit $\int_0^b f(x) dx = \int_0^1 x dx = \frac{b^2}{2}$, vgl. Übungsaufgabe 272. Wir verzichten aber auf diesen Aufwand und fassen eine größere Klasse von Funktionen ins Auge:

Satz 5.1.3.1. *Jede monotone Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist Riemann-integrierbar.*

Beweis. (Skizze) Es genügt, den monoton wachsenden Fall zu behandeln, der monoton fallende ist analog. Sei ε vorgegeben. Wir müssen eine Zerlegung Z finden mit $O(f, Z) - U(f, Z) < \varepsilon$. Wir wissen aus 4.4.1, dass alle Unstetigkeitsstellen einer monotonen Funktion Sprungstellen sind. Die Sprunghöhen können sich maximal auf $d := f(b) - f(a)$ aufaddieren. Also kann es nur endlich viele Sprungstellen s_1, \dots, s_k mit Sprunghöhe $\geq \frac{\varepsilon}{2(b-a)}$ geben. Um jedes dieser s_i legen wir jeweils ein sehr kleines abgeschlossenes Teilintervall J_i der Länge $< \delta$ mit $k\delta d < \frac{\varepsilon}{2}$. Die durch die J_i noch nicht überdeckten Teile von $[a, b]$ können wir in endlich viele Teilintervalle $[x_{j-1}, x_j]$ zerlegen mit $f(x_j) - f(x_{j-1}) < \frac{\varepsilon}{2(b-a)}$. Für die resultierende Zerlegung Z gilt tatsächlich $O(f, Z) - U(f, Z) < \varepsilon$. \square

Übungsaufgabe 275. (E) Prüfen Sie die im Beweis von Satz 5.1.3.1 behauptete Ungleichung $O(f, Z) - U(f, Z) < \varepsilon$ nach.

Sehr leicht (Übung) beweist man:

Proposition 5.1.3.2. *Sei $f : [a, c] \rightarrow \mathbb{R}$, und seien die Einschränkungen von f sowohl auf $[a, b]$ als auch auf $[b, c]$ R-integrierbar. Dann ist f auch auf $[a, c]$ R-integrierbar mit*

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx$$

Beweis. (Idee) Man hat lediglich geeignete Zerlegungen von $[a, b]$ und $[b, c]$ zusammenzusetzen, zunächst unter der Voraussetzung $a \leq b \leq c$. Für andere Konstellationen ist die allgemein übliche (und für Riemann-Integrale sehr bewährte) Konvention $\int_a^b f(x) dx = -\int_b^a f(x) dx$ zu beachten. \square

Übungsaufgabe 276. (P) Erklären Sie den Beweis von Proposition 5.1.3.2 für $a \leq b \leq c$ ausführlich.

Kombiniert man Satz 5.1.3.1 über die Integrierbarkeit monotoner Funktionen mit der durch Proposition 5.1.3.2 aufgezeigten Möglichkeit, Intervalle zusammenzustückeln, so bekommt man schon eine recht große Klasse Riemann-integrierbarer Funktionen. Von theoretischem Interesse ist der Umstand, dass die Riemann-integrierbaren Funktionen auf $[a, b]$ einen Vektorraum bilden, auf dem das Integral als Abbildung $f \mapsto \int_a^b f(x) dx$ linear ist.

Proposition 5.1.3.3. *Angenommen, die Funktionen $f, g : [a, b] \rightarrow \mathbb{R}$ sind R-integrierbar und $c \in \mathbb{R}$. Dann sind die Funktionen cf und $f + g$ ebenfalls R-integrierbar auf $[a, b]$, und es gilt:*

$$\int_a^b cf(x) dx = c \int_a^b f(x) dx \quad \text{und} \quad \int_a^b (f+g)(x) dx = \int_a^b f(x) dx + \int_a^b g(x) dx.$$

Beweis. (Skizze) Für jede Zerlegung $Z = \{a = x_0 \leq x_1 \leq \dots \leq x_{n-1} \leq x_n = b\}$ von $[a, b]$ gilt offenbar $cU(f, Z) = U(cf, Z) \leq O(cf, Z) = cO(f, Z)$. Weil wegen der R-Integrierbarkeit von f durch geeignete Wahl der Zerlegung Z Untersumme $U(f, Z)$ und Obersumme $O(f, Z)$ für f beliebig nahe aneinander gebracht werden können, gilt das auch für cf statt f . (Präzisierung: Übung)

Bei der Behandlung von $f + g$ geht man ähnlich vor, beginnt aber mit der folgenden Überlegung: Für jedes $x \in [x_{i-1}, x_i]$ ist

$$(f + g)(x) = f(x) + g(x) \leq \sup_{x_{i-1} \leq x \leq x_i} f(x) + \sup_{x_{i-1} \leq x \leq x_i} g(x),$$

also auch

$$\sup_{x_{i-1} \leq x \leq x_i} (f + g)(x) \leq \sup_{x_{i-1} \leq x \leq x_i} f(x) + \sup_{x_{i-1} \leq x \leq x_i} g(x).$$

Summation über alle $i = 1, 2, \dots, n$ zeigt $O(f + g, Z) \leq O(f, Z) + O(g, Z)$. Analog zeigt man $U(f, Z) + U(g, Z) \leq U(f + g, Z)$, insgesamt also

$$U(f, Z) + U(g, Z) \leq U(f + g, Z) \leq O(f + g, Z) \leq O(f, Z) + O(g, Z).$$

Weil außerdem

$$U(f, Z) + U(g, Z) \leq \int_a^b f(x) dx + \int_a^b g(x) dx \leq O(f, Z) + O(g, Z)$$

gilt, genügt es, für beliebiges $\varepsilon > 0$ eine Zerlegung Z mit $O(f + g, Z) - U(f + g, Z) < \varepsilon$ zu finden. Tatsächlich, zu vorgegebenem ε können wir wegen Satz 5.1.2.1, Aussage 4, ein δ finden, so dass für alle Zerlegungen Z von $[a, b]$ mit Feinheit $\phi(Z) < \delta$ sowohl $O(f, Z) - U(f, Z) < \frac{\varepsilon}{2}$ als auch $O(g, Z) - U(g, Z) < \frac{\varepsilon}{2}$ gilt. So ein Z wählen wir und schließen damit

$$O(f + g, Z) - U(f + g, Z) \leq O(f, Z) + O(g, Z) - U(f, Z) - U(g, Z) < \varepsilon.$$

□

Übungsaufgabe 277. (E) Im Beweis von Proposition 5.1.3.3 wurde $f + g$ vollständig behandelt, cf nur skizzenhaft. Führen Sie die Argumentation auch für cf vollständig aus.

Weitere wichtige Beobachtungen betreffen die Verträglichkeit von Integral mit Ungleichungen (**Integralungleichungen**).

Proposition 5.1.3.4. Die Funktionen $f, g : [a, b] \rightarrow \mathbb{R}$ seien R-integrierbar.

1. Aus $f \leq g$, d.h. $f(x) \leq g(x)$ für alle $x \in [a, b]$, folgt $\int_a^b f(x) dx \leq \int_a^b g(x) dx$. (Monotonie oder auch Positivität des Integrals)
2. $|\int_a^b f(x) dx| \leq \int_a^b |f(x)| dx$ (vgl. Dreiecksungleichung für Summen)

Beweis. 1. (Idee) Die entsprechenden Beziehungen gelten für alle Ober- und Untersummen, daher auch für die Integrale.

2. Die R-Integrierbarkeit von f überträgt sich, wie man durch Abschätzung der Differenz von Ober- und Untersummen leicht einsieht (Übung), auch auf die beiden nichtnegativen Funktionen $f^+(x) := \max\{f(x), 0\}$ und $f^-(x) := \max\{-f(x), 0\}$. Mit ihnen gilt $f = f^+ - f^-$ und $|f(x)| = f^+(x) + f^-(x)$. Proposition 5.1.3.3 und die bereits bewiesene erste Aussage in Verbindung mit der Dreiecksungleichung implizieren daher

$$\begin{aligned} \left| \int_a^b f(x) dx \right| &= \left| \int_a^b f^+(x) - f^-(x) dx \right| = \left| \int_a^b f^+(x) dx - \int_a^b f^-(x) dx \right| \leq \\ &\leq \left| \int_a^b f^+(x) dx \right| + \left| \int_a^b f^-(x) dx \right| = \\ &= \int_a^b f^+(x) dx + \int_a^b f^-(x) dx = \\ &= \int_a^b f^+(x) + f^-(x) dx = \int_a^b |f(x)| dx. \end{aligned}$$

□

Übungsaufgabe 278. (E) Zeigen Sie die im Beweis von Proposition 5.1.3.4 behauptete R-Integrierbarkeit von f^+ (analog für f^-). Hinweis: Zeigen Sie zunächst $O(f^+, Z) - U(f^+, Z) \leq O(f, Z) - U(f, Z)$ für jede Zerlegung Z .

Übungsaufgabe 279. (E) Wann gilt in Aussage 2 von Proposition 5.1.3.4 nicht nur die Ungleichung mit \leq , sondern sogar Gleichheit?

5.1.4 Das Riemannintegral stetiger Funktionen

Inhalt in Kurzfassung: Jede stetige Funktion ist Riemann-integrierbar. Im Beweis ist entscheidend, dass stetige Funktionen auf abgeschlossenen Intervallen (besser: auf kompakten Mengen) sogar gleichmäßig stetig sind.

Mit den stückweise monotonen Funktionen haben wir bereits eine große Klasse R-integrierbarer Funktionen kennen gelernt. Die wichtigste Klasse steht aber noch aus, nämlich die der stetigen Funktionen $f : [a, b] \rightarrow \mathbb{R}$. Auch sie erweisen sich als R-integrierbar. Die ε - δ -Definition der Stetigkeit legt dies auch nahe, weil ja hinreichend kleines δ garantiert, dass die Funktionswerte auf Intervallen der Länge $< \delta$ um weniger als ε variieren. In dieser Schlussweise verbirgt sich aber eine Tücke, die von ähnlicher Art ist, wie an früherer Stelle der Unterschied zwischen punktweiser und gleichmäßiger Konvergenz. Denn das δ aus der Stetigkeitsdefinition bezieht sich auf nur eine Stelle x_0 . Es wäre also denkbar, dass wir für verschiedene Stellen (davon gibt es ja unendlich viele) immer kleinere δ wählen müssen und deshalb kein gemeinsames $\delta > 0$ für das ganze Intervall $[a, b]$

und folglich auch kein endliches Z finden, für das Ober- und Untersumme einander beliebig nahe kommen. Wir brauchen deshalb die folgende Verschärfung der Stetigkeit.

Definition 5.1.4.1. Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt **gleichmäßig stetig** auf D , wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt derart, dass für alle $x_1, x_2 \in D$ mit $|x_1 - x_2| < \delta$ auch $|f(x_1) - f(x_2)| < \varepsilon$ gilt.

Klarerweise ist jede gleichmäßig stetige Funktion auch stetig. Die Umkehrung gilt jedoch nicht. Zum Beispiel ist $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^2$, wie wir wissen, auf ganz \mathbb{R} stetig. Gleichmäßige Stetigkeit liegt aber nicht vor, weil sogar jedes $\varepsilon > 0$ ein Gegenbeispiel ist. Denn für jedes $\delta > 0$ laufen die Funktionswerte $f(x) = x^2$ und $f(x + \delta) = x^2 + 2x\delta + \delta^2$ beliebig weit auseinander, sofern nur x hinreichend groß ist.

Der logische Unterschied der beiden Begriffe wird wieder einmal deutlich, wenn man Quantoren verwendet. Dann lautet Stetigkeit auf D

$$\forall \varepsilon > 0 \forall x \in D \exists \delta > 0 \forall x' \in D : |x - x'| < \delta \rightarrow |f(x) - f(x')| < \varepsilon,$$

während gleichmäßige Stetigkeit auf D

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D \forall x' \in D : |x - x'| < \delta \rightarrow |f(x) - f(x')| < \varepsilon$$

bedeutet. Die Quantoren $\forall x \in D$ und $\exists \delta > 0$ wurden also vertauscht.

Ein anderes Beispiel, anhand dessen sich sehr gut der Unterschied zwischen Stetigkeit und gleichmäßiger Stetigkeit illustrieren lässt, ist die Funktion $f(x) := \frac{1}{x}$. Sie ist Gegenstand der folgenden Übungsaufgabe.

Übungsaufgabe 280. (P) Wir betrachten die Funktion $f(x) := \frac{1}{x}$ auf dem Definitionsbereich $D := (0, \infty)$.

1. Zeigen Sie, dass f auf D nicht gleichmäßig stetig ist.
2. Zeigen Sie, dass f auf $[1, \infty) \subseteq D$ sehr wohl gleichmäßig stetig ist.
3. Worauf kommt es bei einer Teilmenge $D_0 \subseteq D$ an, damit f auf D_0 gleichmäßig stetig ist?

Glücklicherweise fällt der Unterschied zwischen Stetigkeit und gleichmäßiger Stetigkeit auf kompakten Mengen weg.

Proposition 5.1.4.2. Jede auf einer kompakten (d.h. abgeschlossenen und beschränkten) Menge $D \subseteq \mathbb{R}$ definierte stetige Funktion $f : D \rightarrow \mathbb{R}$ ist sogar gleichmäßig stetig.

Beweis. Wir gehen indirekt vor und nehmen an, die stetige Funktion sei nicht gleichmäßig stetig. Dann gibt es ein $\varepsilon > 0$ derart, dass sich zu jedem $\delta_n = \frac{1}{n}$ mit $n = 1, 2, \dots$ ein Paar von x -Werten x_n und x'_n in $[a, b]$ finden lässt mit $|x_n - x'_n| < \frac{1}{n}$ aber $|f(x_n) - f(x'_n)| \geq \varepsilon$. Nach dem Satz von Bolzano-Weierstraß haben die x_n einen Häufungspunkt $x_0 \in [a, b]$. In jeder Umgebung von x_0 liegen Paare von

x -Werten, deren Funktionswerte um mindestens ε voneinander entfernt sind. Wenigstens einer davon muss dann von $f(x_0)$ mindestens $\frac{\varepsilon}{2}$ entfernt sein. Das widerspricht aber der Stetigkeit von f in x_0 . Dieser Widerspruch beweist die gleichmäßige Stetigkeit von f auf $[a, b]$. \square

Nun können wir relativ leicht beweisen:

Satz 5.1.4.3. *Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$, $a < b$, ist R -integrierbar.*

Beweis. Weil f wegen Proposition 5.1.4.2 sogar gleichmäßig stetig auf $[a, b]$ ist, gibt es zu vorgegebenem $\varepsilon > 0$ ein $\delta > 0$ derart, dass $|f(x) - f(x')| < \frac{\varepsilon}{b-a}$ für alle $x, x' \in [a, b]$ mit $|x - x'| < \delta$. Wir wählen eine Zerlegung $Z = \{a = x_0 \leq x_1 \leq \dots \leq x_n = b\}$ mit $\phi(Z) < \delta$. Dann gilt $\sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x) < \frac{\varepsilon}{b-a}$ für alle $i = 1, 2, \dots, n$. Summation über alle i liefert die gewünschte Beziehung

$$O(f, Z) - U(f, Z) < \sum_{i=1}^n (x_i - x_{i-1}) \frac{\varepsilon}{b-a} = \varepsilon.$$

\square

5.2 Der Hauptsatz und seine Anwendungen

Wie sein Name suggeriert, bildet der Hauptsatz der Differential- und Integralrechnung das Bindeglied zwischen diesen beiden Themenbereichen, die somit zur Einheit jenes so wichtigen Gebietes der Mathematik verschmelzen, das man auch als Infinitesimalrechnung bezeichnet (5.2.1). Neben der zentralen Bedeutung für die Theorie ist der Hauptsatz nämlich auch die Grundlage für die praktische Berechnung vieler Integrale (5.2.2), insbesondere jener, die mit den Integrationsregeln aus 5.2.3 zugänglich sind. Wir schließen den Abschnitt in 5.2.4 mit einigen Beispielen dazu.

5.2.1 Zwei Versionen des Hauptsatzes

Inhalt in Kurzfassung: Der Hauptsatz der Differential- und Integralrechnung zeigt, inwiefern Differenzieren und Integrieren zueinander inverse Operationen sind, analog zu Subtraktion und Addition. Damit verschmelzen beide Gebiete zur sogenannten Infinitesimalrechnung. Neben seiner Bedeutung für die Theorie dient der Hauptsatz, wie wir später sehen werden, auch zur expliziten Berechnung von Integralen. Dazu sind auch noch Überlegungen zur Eindeutigkeit von Stammfunktionen am Platz.

Der **Hauptsatz der Differential- und Integralrechnung** stellt die Verbindung zwischen diesen beiden Gebieten her. Und zwar geht es um die Beziehung

$$\int_a^b f(x) dx = F(b) - F(a),$$

wobei F eine Stammfunktion von f ist, also $F' = f$ erfüllt (an den Randpunkten a und b eventuell als einseitige Ableitung). Zwei Formulierungen sind von Interesse, die eine für vorgegebenes f (Voraussetzung: f stetig), die andere für vorgegebenes F (Voraussetzung: Riemann-integrierbare Ableitung F'). Die präzisen Formulierungen:

Satz 5.2.1.1. (Erste Version des Hauptsatzes) Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig ($a < b$). Dann ist die Funktion

$$F : [a, b] \rightarrow \mathbb{R}, \quad x \mapsto \int_a^x f(t) dt,$$

differenzierbar, und es gilt $F' = f$ (an den Randpunkten a, b als einseitige Ableitung). Wegen $F(a) = 0$ gilt insbesondere auch

$$\int_a^b f(x) dx = F(b) - F(a).$$

Beweis. Für ein beliebiges $x_0 \in [a, b]$ untersuchen wir den Differentialquotienten

$$F'(x_0) = \lim_{x \rightarrow x_0} \frac{F(x) - F(x_0)}{x - x_0}.$$

Die Differenz im Zähler erfüllt $F(x) - F(x_0) = \int_{x_0}^x f(t) dt$. Weil f stetig ist, gibt es zu beliebig vorgegebenem $\varepsilon > 0$ ein $\delta > 0$ derart, dass für alle $x \in [a, b]$ mit $|x - x_0| < \delta$ die Ungleichung $f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon$ garantiert ist. Für solche x gilt folglich $(f(x_0) - \varepsilon)(x - x_0) < F(x) - F(x_0) < (f(x_0) + \varepsilon)(x - x_0)$ und somit $f(x_0) - \varepsilon < \frac{F(x) - F(x_0)}{x - x_0} < f(x_0) + \varepsilon$. Weil $\varepsilon > 0$ beliebig war, existiert der Grenzwert für $x \rightarrow x_0$ und hat den Wert $F'(x_0) = f(x_0)$. \square

Satz 5.2.1.2. (Zweite Version des Hauptsatzes) Angenommen die Funktion $F : [a, b] \rightarrow \mathbb{R}$ habe eine R-integrierbare Ableitung $f := F'$ (in den Randpunkten a, b als einseitige Ableitung), so gilt

$$\int_a^b f(x) dx = F(b) - F(a).$$

Beweis. Wir gehen von einer Zerlegung $Z = \{a = x_0 < x_1 < \dots < x_n = b\}$ aus. Nach dem Mittelwertsatz gibt es zu jedem $i = 1, 2, \dots, n$ eine Zwischenstelle $\xi_i \in (x_{i-1}, x_i)$ mit $f(\xi_i) = F'(\xi_i) = \frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}}$. Somit hat die Riemannsumme von f für Z und zur Belegung $B := \{\xi_1, \dots, \xi_n\}$ den Wert

$$\begin{aligned} S(F, Z, B) &= \sum_{i=1}^n F(\xi_i)(x_i - x_{i-1}) = \sum_{i=1}^n (F(x_i) - F(x_{i-1})) = \\ &= F(x_n) - F(x_0) = F(b) - F(a). \end{aligned}$$

Weil f nach Voraussetzung R-integrierbar ist, herrscht für $\phi(Z) \rightarrow 0$ Konvergenz der Riemannsummen $S(f, Z, B)$ gegen $\int_a^b f(x) dx$, also gilt Nach Satz 5.1.2.1 die behauptete Identität. \square

Übungsaufgabe 281. (P) Sei f stetig mit einer Stammfunktion F . Außerdem seien $a(x)$ und $b(x)$ differenzierbare Funktionen. Wir definieren die Funktion

$$G(x) := \int_{a(x)}^{b(x)} f(t) dt.$$

1. Nehmen Sie zunächst a als konstante Funktion an. Begründen Sie, warum G differenzierbar ist und berechnen Sie die Ableitung G' . Anleitung: Fassen Sie G als Verkettung $G = F \circ b$ auf.
2. Analog mit konstantem b aber nicht konstantem a .
3. Behandeln Sie nun den allgemeinen Fall, wo weder a noch b konstant sein müssen.

Übungsaufgabe 282. (T) Verwenden Sie den Hauptsatz der Differential- und Integralrechnung, um zu entscheiden ob F eine Stammfunktion von f ist:

1. $f(x) = -\sin(x)$, $F(x) = \int_x^0 \sin(t) dt$.
2. $f(x) = x^3$, $F(x) = \int_0^{2x} 3t^2 dt$.
3. $f(x) = e^x$, $F(x) = 2 \int_1^{x/2} e^{2t} dt$.

5.2.2 Die Berechnung von Integralen mittels Stammfunktionen

Inhalt in Kurzfassung: Weil Stammfunktionen, grob gesprochen, bis auf additive Konstante eindeutig bestimmt sind, gelingt die Berechnung eines (bestimmten) Integrals sehr leicht, wenn eine beliebige Stammfunktion des Integranden vorliegt. Das wird an einem sehr einfachen Beispiel erläutert und vorgeführt.

Von einer Anwendung des Mittelwertsatzes 4.2.2.3 wissen wir bereits, dass auf zusammenhängenden Mengen D eindeutig sind bis auf eine additive Konstante. Diese Aussage gilt nicht mehr, wenn D nicht zusammenhängend ist. Beispiel:

Sei $f : D \rightarrow \mathbb{R}$ mit $D = \mathbb{R} \setminus \{0\}$ und $f(x) := \frac{1}{x}$. Dann erhält man sämtliche Stammfunktionen von f , indem man für jedes Paar $c_-, c_+ \in \mathbb{R}$ die Funktion F_{c_+, c_-} definiert durch $F_{c_+, c_-}(x) := \ln|x| + c_-$ für $x < 0$ und $F_{c_+, c_-}(x) := \ln x + c_+ = \ln|x| + c_+$ für $x > 0$. Beweis: Jede dieser Funktionen ist, wie man sofort nachrechnet, Stammfunktion von f (für $x < 0$ ist nach der Kettenregel $(\ln|x|)' = (\ln(-x))' = \frac{1}{-x}(-1) = \frac{1}{x}$). Weil die Teile $x < 0$ und $x > 0$ für sich genommen zusammenhängend sind, ist für jeden von ihnen der Satz über die Eindeutigkeit von Stammfunktionen bis auf eine additive Konstante anwendbar. Somit gibt es umgekehrt für jede Stammfunktion F von f Zahlen c_-, c_+ mit $F = F_{c_-, c_+}$.

Ist die Funktion f gegeben, so hat sich für die Stammfunktion(en) die Schreibweise $\int f(x) dx + c$ oder, noch kürzer $\int f$, und die Bezeichnung **unbestimmtes Integral** eingebürgert; im Gegensatz zum bestimmten (Riemann-)Integral $\int_a^b f(x) dx$, in dem immer auch der Integrationsbereich angegeben werden muss. Eigentlich ist mit $\int f(x) dx + c$ die Menge aller Stammfunktionen von f gemeint. Ist der Definitionsbereich zusammenhängend und F eine Stammfunktion von f , so ist also genau genommen

$$\int f = \int f(x) dx + c = \{F + c : c \in \mathbb{R}\}.$$

Bei nicht zusammenhängendem Definitionsbereich ist, wie oben am Beispiel von $f(x) = \frac{1}{x}$ auf $\mathbb{R} \setminus \{0\}$ diskutiert wurde, die Menge entsprechend komplizierter. Wir werden die additive Konstante nicht immer anschreiben und z.B. kurz $\int x^2 dx = \frac{x^3}{3}$ notieren.

Übungsaufgabe 283. (P) Sei $f : D \rightarrow \mathbb{R}$ stetig und D von der Form $D = \bigcup_{i=1}^n (a_i, b_i)$ mit $a_1 < b_1 \leq a_2 < b_2 \leq \dots \leq a_{n-1} < b_{n-1} \leq a_n \leq b_n$. Weiters sei F eine Stammfunktion von f .

1. Beschreiben Sie das unbestimmte Integral $\int f dx$ als Menge aller Stammfunktionen von f . Hinweis: Für jeden zusammenhängenden Teilbereich von D ist eine eigene additive Konstante zu nehmen.
2. Für welche Intervalle $[a, b]$ ist das bestimmte Integral $\int_a^b f(x) dx$ definiert, und wie lässt es sich aus den gegebenen Daten berechnen?

Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, so gibt es nach der ersten Version des Hauptsatzes 5.2.1.1 eine (differenzierbare) Stammfunktion F von f . Nach Satz 4.2.2.3 unterscheidet sich jede andere Stammfunktion von f von F nur durch eine additive Konstante, ist daher ebenfalls differenzierbar mit (R-integrierbarer) Ableitung f . Nach der zweiten Version des Hauptsatzes 5.2.1.2 gilt also nicht nur für eine spezielle, sondern für jede Stammfunktion F von f die dort behauptete Beziehung. Zusammenfassend haben wir daher:

Satz 5.2.2.1. Ist F irgendeine Stammfunktion der stetigen Funktion $f : [a, b] \rightarrow \mathbb{R}$, so ist

$$\int_a^b f(x) dx = F(b) - F(a).$$

Eine übliche Notation bei konkreten Berechnungen, hier durchgeführt für das einfache Beispiel $f(x) = x$, $F(x) = \frac{x^2}{2}$, $a = 1$ und $b = 2$, lautet:

$$\int_1^2 x dx = \frac{x^2}{2} \Big|_{x=1}^2 = \frac{2^2}{2} - \frac{1^2}{2} = \frac{4-1}{2} = \frac{3}{2}.$$

Bestimmte Integrale stetiger Funktionen auf Intervallen lassen sich also sicher dann berechnen, wenn man eine Stammfunktion des Integranden f findet,

also f , wie man sagt, **integriert**. Dabei hat man offenbar den Vorgang des Differenzierens umzukehren. Grundsätzlich existieren Integrale zwar eher als Ableitungen. Denn jede differenzierbare Funktion ist stetig und somit erst recht \mathbb{R} -integrierbar, nicht aber umgekehrt. Die explizite Berechnung von Stammfunktionen ist in der Praxis aber meist schwieriger als die der Ableitung. Man bedenke, dass die Ableitungsregeln insofern vollständig sind, als sie uns sagen, wie wir – ausgehend von elementaren Funktionen wie Polynomen, \exp , \ln , trigonometrischen Funktionen etc. – deren Summen, Differenzen, Produkte, Quotienten und Verkettungen behandeln müssen. Für derart aufgebaute Funktionen liefern diese Regeln stets ein Ergebnis. Für die Berechnung von Stammfunktionen ist dies leider nicht der Fall. Immerhin können sehr viele einfach gebaute Funktionen mit Hilfe weniger Integrationsregeln integriert werden.

Übungsaufgabe 284. (P) Die **Signumfunktion** sgn ist definiert durch

$$\operatorname{sgn}(x) := -1 \quad \text{für } x < 0, \quad \operatorname{sgn}(0) := 0, \quad \text{und} \quad \operatorname{sgn}(x) := 1 \quad \text{für } x > 0.$$

1. Hat sgn eine Stammfunktion? Hinweis: Sie dürfen Übungsaufgabe 224 verwenden.
2. Verallgemeinern Sie Ihre Antwort auf beliebige Funktionen mit Sprungstellen.
3. Finden Sie eine Funktion mit einer Unstetigkeit anderer Art, die sich in Hinblick auf die Existenz von Stammfunktionen anders verhält als sgn .

5.2.3 Integrationsregeln

Inhalt in Kurzfassung: Mit Hilfe des Hauptsatzes lassen sich Differentiationsregeln zu Integrationsregeln umkehren, teilweise aber nur mit eingeschränktem Anwendungsbereich. Unmittelbar umkehrbar sind die bereits bekannten Ableitungsregeln für diverse elementare Funktionen (Polynome, \exp , trigonometrische Funktionen, Potenzfunktionen etc.). Außerdem vererbt sich die Linearität des Differenzierens unmittelbar auf das Integrieren. Nur eingeschränkt zielführend sind hingegen partielle Integration (Umkehrung der Produktregel) und Substitutionsregel (Umkehrung der Kettenregel).

Die verfügbaren Standardmethoden bestehen in der Verwendung der Stammfunktionen elementarer Funktionen, der Linearität, der partiellen Integration (entspricht der Produktregel beim Differenzieren) und der Substitutionsregel (entspricht der Kettenregel beim Differenzieren). Der Beweis dieser Regeln ergibt sich jeweils mehr oder weniger unmittelbar aus den entsprechenden Differentiationsregeln.

Um die Notation im Folgenden übersichtlich zu halten, beziehen wir die Regeln in diesem Abschnitt stillschweigend auf den jeweiligen Definitionsbereich. Meist wird nur eine Stammfunktion angegeben, d.h. ohne Berücksichtigung der additiven Konstante(n).

Stammfunktionen elementarer Funktionen:

- Potenzen: $\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1}$ für $\alpha \in \mathbb{R} \setminus \{-1\}$ und $\int x^{-1} dx = \int \frac{1}{x} dx = \ln|x|$
- Exponentialfunktion und Logarithmus: $\int \exp(x) dx = \exp(x)$,
 $\int \ln(x) dx = x \ln(x) - x$
- Trigonometrische Funktionen: $\int \cos x dx = \sin x$, $\int \sin x dx = -\cos x$,
 $\int \tan x dx = -\ln|\cos x|$, $\int \cot x dx = -\ln|\sin x|$
- Hyperbelfunktionen: $\int \cosh x dx = \sinh x$, $\int \sinh x dx = \cosh x$,
 $\int \tanh x dx = \ln(\cosh x)$, $\int \coth x dx = \ln|\sinh x|$.
- Weitere nützliche Stammfunktionen, die sich direkt aus Differentiationsregeln ergeben: $\int \frac{1}{1+x^2} dx = \arctan x$ oder $\int \frac{1}{1+x^2} dx = -\operatorname{arccot} x$,
 $\int \frac{1}{\sqrt{1-x^2}} dx = \arcsin x$ oder $\int \frac{1}{\sqrt{1-x^2}} dx = -\arccos x$,
 $\int \frac{1}{\sqrt{1+x^2}} dx = \operatorname{Arsinh} x = \ln(x + \sqrt{x^2 + 1})$, $\int \frac{1}{\sqrt{x^2-1}} dx = \operatorname{Arcosh} x = \ln(x + \sqrt{x^2 - 1})$.

Es ist nicht notwendig, sich alle Regeln auswendig zu merken. Sehr wohl einprägen sollte man sich die Stammfunktionen von Potenzen, exp, sin, cos und $\frac{1}{1+x^2}$.

Linearität: Die Linearität des unbestimmten Integrals kann man sowohl aus der Linearität des Differenzierens wie auch aus der des bestimmten Integrals und dem Hauptsatz folgern:

$$\int (f + g) = \int f + \int g \quad \text{und} \quad \int (cf) = c \int f \quad \text{für } c \in \mathbb{R}$$

Partielle Integration: Aus der Produktregel fürs Differenzieren folgt die Integrationsregel

$$\int fg = Fg - \int (Fg'),$$

wobei F eine Stammfunktion von f bezeichne. Der Beweis ergibt sich durch Differenzieren auf beiden Seiten. Links erhält man nach Definition des unbestimmten Integrals fg , rechts mit Hilfe von Linearität und Produktregel ebenfalls:

$$\left(Fg - \int Fg' \right)' = (Fg)' - Fg' = F'g + Fg' - Fg' = F'g = fg.$$

Ein einfaches Beispiel: Im unbestimmten Integral $\int xe^x dx$ wählen wir $f(x) = e^x$, also $F(x) = e^x$ und $g(x) = x$, also $g'(x) = 1$. Somit ist

$$\int xe^x dx = e^x x - \int e^x dx = e^x x - e^x = e^x(x - 1).$$

Partielle Integration führt nicht immer zum Ziel, selbst wenn man eine Stammfunktion F von f kennt. Denn auf der rechten Seite taucht wieder ein

Integral auf, das eventuell leichter zu berechnen ist als das ursprüngliche, nicht aber in allen Fällen; deshalb das Wort *partiell* in der Bezeichnung der Regel.

Substitutionsregel: So wie die partielle Integration der Produktregel fürs Differenzieren entspricht, hat auch die Kettenregel eine Art Umkehrung, die aber ebenfalls nicht immer zum Ziel führt. Mit einer Stammfunktion F von f (jeweils als Funktionen in x) lässt sich das allgemeine Schema wie folgt notieren:

$$\int f(x) dx = F(x) = F(g(t)) = \int f(g(t))g'(t) dt$$

Darin besagt die erste Gleichheit, dass F eine Stammfunktion von f ist; die zweite bringt die Variablensubstitution $x = g(t)$ mit Hilfe einer (geschickt zu wählenden) Funktion g zum Ausdruck; und die dritte entspricht der Kettenregel, der zufolge der Integrand rechts die Ableitung von F nach t ist. Man liest daraus ab, dass neben der Substitution $x = g(t)$ auch $dx = g'(t)dt$ gesetzt werden muss, was man oft durch den Einschub eines Kästchens in folgender Weise notiert:

$$\int f(x) dx = \left| \begin{array}{l} x = g(t) \\ dx = g'(t)dt \end{array} \right| = \int f(g(t))g'(t) dt$$

Am besten versteht man die Methode anhand eines Beispiels, etwa des unbestimmten Integrals $\int e^{\sqrt{x}} dx$. Wir setzen $x = t^2 =: g(t)$, also $dx = g'(t) dt = 2t dt$, und erhalten

$$\int e^{\sqrt{x}} dx = \left| \begin{array}{l} x = t^2 \\ dx = 2t dt \end{array} \right| = \int e^t 2t dt = 2(te^t - e^t) = 2(\sqrt{x}e^{\sqrt{x}} - e^{\sqrt{x}}).$$

In der dritten Gleichheit wurde das weiter oben behandelte Beispiel zur partiellen Integration verwendet. Geht man von einem bestimmten Integral aus, so kann man sich die Rücksubstitution ersparen, wenn man die Integrationsgrenzen mittransformiert. Am Ende von 5.2.4 werden wir ein typisches Beispiel dazu behandeln.

5.2.4 Beispiele zur Integration

Inhalt in Kurzfassung: Wir behandeln ausführlich die Integration gebrochener rationaler Funktionen mittels Partialbruchzerlegung. Als Beispiele zur partiellen Integration wählen wir die Integranden \ln und \cos^2 . Die Substitutionsregel schließlich illustrieren wir anhand der Berechnung der Fläche eines Kreises.

Mit Hilfe der in 5.2.3 angegebenen Integrationsregeln lassen sich Polynomfunktionen in offensichtlicher Weise integrieren:

$$\int \sum_{k=0}^n a_k x^k dx = \sum_{k=0}^n a_k \int x^k dx = \sum_{k=0}^n \frac{a_k}{k+1} x^{k+1}.$$

Nicht so offensichtlich, aber dennoch gemäß einem festen Schema lassen sich auch **gebroschen rationale Funktionen** behandeln, obwohl sie durch Division

zustande kommen, für welche keine allgemeine Integrationsregel zur Verfügung steht. Und zwar stützt man sich auf die **Partialbruchzerlegung**. Offenbar reicht es, Polynome und Funktionen folgender vier Typen zu integrieren:

1. $f_1(x) = \frac{1}{x-\alpha}$
2. $f_2(x) = \frac{1}{(x-\alpha)^k}$ mit $k \geq 2$
3. $f_3(x) = \frac{Ax+B}{x^2+\beta x+\gamma}$
4. $f_4(x) = \frac{Ax+B}{(x^2+\beta x+\gamma)^k}$ mit $k \geq 2$

Dabei sind $A, B, \alpha, \beta, \gamma$ reelle Zahlen und das Polynom $x^2 + \beta x + \gamma$ im Nenner von f_3 und f_4 reell irreduzibel.

Sehr einfach ist die Aufgabe bei den ersten beiden Typen:

$$\int f_1(x) dx = \int \frac{1}{x-\alpha} dx = \ln|x-\alpha|$$

und

$$\begin{aligned} \int f_2(x) dx &= \int \frac{1}{(x-\alpha)^k} dx = \int (x-\alpha)^{-k} dx = \frac{(x-\alpha)^{-k+1}}{-k+1} = \\ &= -\frac{1}{(k-1)(x-\alpha)^{k-1}}. \end{aligned}$$

Bei f_3 und f_4 kann man durch geeignete lineare Substitutionen erreichen, dass der Nenner in einer neuen Variablen y ein Vielfaches von $y^2 + 1$ wird. Im ersten Schritt bringt man mittels der Substitution $x = t + \frac{\beta}{2}$ das lineare Glied zum Verschwinden, landet also beim Typus $t^2 + \gamma'$ im Nenner mit (wegen der Irreduzibilität) positivem γ' . Sodann setzt man $t = \sqrt{\gamma'}y$, also $t^2 + \gamma' = \gamma'y^2 + \gamma' = \gamma'(y^2 + 1)$. Wir dürfen uns also auf den Spezialfall $\beta = 0$ und $\gamma = 1$ beschränken. Spalten wir ein f_3 dieser Form noch gemäß dem Zähler $Ax + B$ auf, so ist dieser Fall abzuhandeln mit Hilfe der beiden unbestimmten Integrale

$$\int \frac{x}{1+x^2} dx = \frac{\ln(1+x^2)}{2}$$

und

$$\int \frac{1}{1+x^2} dx = \arctan(x).$$

(Deshalb die rechnerische Bedeutung des Arcustangens!)

Es bleibt der Typ f_4 , für den der eine der beiden Summanden auch keine großen Probleme macht:

$$\int \frac{x}{(1+x^2)^k} dx = -\frac{1}{2(k-1)(1+x^2)^{k-1}}.$$

Der einzige etwas mühsamere Fall betrifft das unbestimmte Integral

$$I_k := \int \frac{1}{(1+x^2)^k} dx.$$

Den Fall $k = 1$ haben wir bereits bei f_3 behandelt. Wegen der (durch Differenzieren nachzurechnenden) Rekursionsformel

$$I_{k+1} = \frac{x}{2k(1+x^2)^k} + \left(1 - \frac{1}{2k}\right)I_k$$

lässt sich der Fall $k > 1$ iterativ auf den bereits behandelten zurückführen.

Übungsaufgabe 285. (T) Berechnen Sie die Integrale mittels Partialbruchzerlegung:

$$\begin{array}{ll} 1. \int \frac{dx}{x(x^2-1)} & 2. \int \frac{dy}{y(y^2+1)} \\ 3. \int \frac{z^2+z+1}{z(z^2+1)} dz & \end{array}$$

In der Liste der Stammfunktionen $F(x)$ elementarer Funktionen $f(x)$ kam auch $f(x) = \ln x$ mit $F(x) = x \ln x - x$ vor. So leicht dies auch durch Ableiten zu verifizieren ist, so lehrreich ist es zu sehen, dass man auch mittels partieller Integration zu diesem Ergebnis kommt. Man interpretiert $\ln x$ als $\ln(x) = f(x)g(x)$ mit $f(x) = 1$ und $g(x) = \ln(x)$, also $F(x) = x$ und $g'(x) = \frac{1}{x}$. So erhält man

$$\begin{aligned} \int \ln(x) dx &= \int fg = Fg - \int (Fg') = x \ln(x) - \int \frac{x}{x} dx = \\ &= x \ln(x) - \int 1 = x \ln(x) - x. \end{aligned}$$

Ein anderes typisches Beispiel für partielle Integration, diesmal mit $f = g = \cos$, $F = \sin$ und $g' = -\sin$ ist

$$\begin{aligned} \int \cos^2 &= Fg - \int Fg' = \sin \cos + \int \sin^2 = \sin \cos + \int (1 - \cos^2) = \\ &= \sin \cos + \int 1 - \int \cos^2. \end{aligned}$$

Wir bringen $\int \cos^2$ von rechts auf die linke Seite und erhalten nach Division durch 2:

$$\int \cos^2(x) dx = \frac{1}{2}(\sin(x) \cos(x) + x).$$

Ähnlich verfährt man mit Integranden wie \sin^2 oder $\sin \cdot \cos$, die in der Theorie der Fourierreihen häufig auftreten.

Übungsaufgabe 286. (T) Berechnen Sie die Integrale mit Hilfe partieller Integration:

1. $\int (\ln x)^2 dx$
2. $\int y^2 \sin y dy$
3. $\int z \cos^2 z dz$

Übungsaufgabe 287. (T) Wie die vorige Aufgabe, aber mit

1. $\int x^2 e^x dx$
2. $\int x \ln x dx$
3. $\int \frac{\ln x}{x} dx$

Zum Abschluss berechnen wir als Anwendungsbeispiel für die Substitutionsregel sowie die mit ihr einhergehenden Transformationen von Intervallgrenzen die Kreisfläche. Sie setzt sich aus vier Viertelkreisen zusammen. Nehmen wir zunächst den Radius 1. Dann ist die Fläche V des Viertelkreises gegeben durch $V = \int_0^1 \sqrt{1-x^2} dx$, also:

$$V = \int_0^1 \sqrt{1-x^2} dx = \left| \begin{array}{l} x = \sin t \\ dx = \cos t dt \end{array} \right| = \int_0^{\frac{\pi}{2}} \sqrt{1-\sin^2 t} \cos t dt = \int_0^{\frac{\pi}{2}} \cos^2 t dt.$$

Weiter oben haben wir eine Stammfunktion F des Integranden $f(t) := \cos^2 t$ bereits berechnet als $F(t) = \frac{1}{2}(\sin(t) \cos(t) + t)$. Setzen wir ein, so erhalten wir für die Fläche V des Viertelkreises

$$V = F\left(\frac{\pi}{2}\right) - F(0) = \frac{1}{2}\left(\sin\left(\frac{\pi}{2}\right) \cos\left(\frac{\pi}{2}\right) + \frac{\pi}{2}\right) - \frac{1}{2}(\sin(0) \cos(0) + 0) = \frac{\pi}{4}.$$

Die Fläche des gesamten Kreises ist somit π . Den Fall eines allgemeinen Radius r führt man (wieder für den Viertelkreis) auf das bereits behandelte Integral zurück und erhält für die Fläche $A(r)$ des Kreises mit Radius r die bekannte Formel

$$\begin{aligned} A(r) &= 4 \int_0^r \sqrt{r^2 - x^2} dx = \left| \begin{array}{l} x = rt \\ dx = r dt \end{array} \right| = 4 \int_0^1 \sqrt{r^2 - r^2 t^2} r dt = \\ &= 4 \int_0^1 r \sqrt{1-t^2} r dt = 4r^2 \int_0^1 \sqrt{1-t^2} dt = r^2 \pi. \end{aligned}$$

Übungsaufgabe 288. (P) Ermitteln Sie mit Hilfe der Integralrechnung die Fläche der Ellipse mit Halbachsen $a, b > 0$.

Übungsaufgabe 289. (T) Berechnen Sie die Integrale mit Hilfe geeigneter Substitutionen:

1. $\int x^2 \sqrt{x-2} dx$
2. $\int \tan y dy$
3. $\int \sin(\ln z) dz$

Übungsaufgabe 290. (T) Wie die vorige Aufgabe, aber mit

1. $\int x^3 \sqrt{x-1} dx$
2. $\int \sin x \cos^3 x dx$
3. $\int \frac{1}{\sqrt{1+e^x}} dx$

5.3 Das Integral bezüglich eines Maßes

Der Riemannsche Integralbegriff, der zu Beginn dieses Kapitels behandelt wurde, ist nach Bernhard Riemann (1826-1866) benannt. Zu Beginn des 20. Jahrhunderts entwickelte Henri Lebesgue (1875-1941) eine Integrationstheorie, die der älteren in mancherlei Hinsicht überlegen ist. Einer ihrer Vorteile besteht in der großen Verallgemeinerbarkeit. So ist der Lebesguesche Zugang auch Grundlage für die moderne Stochastik (= Wahrscheinlichkeitstheorie und Statistik). Fundamental ist der äußerst naheliegende Begriff des Maßes (5.3.1). In Hinblick auf die vielfältigen Anwendungsmöglichkeiten werden in 5.3.2 die wesentlichen Grundideen, die zum Lebesgueschen Maß führen, dargestellt und in 5.3.3 der enge Zusammenhang zwischen dem auf diesem Maßbegriff aufbauenden Lebesgueschen Integral und dem bereits bekannten Riemannintegral skizziert. Charakteristische Ergebnisse, weitgehend ohne strenge Beweise schließen mit 5.3.4 den Abschnitt ab. Ziel ist vor allem ein intuitives Verständnis, mit Hilfe dessen allfällige Berührungspunkte mit der Lebesgueschen Theorie abgebaut werden können. Das ist unter anderem (aber keineswegs nur) für ein adäquates Verständnis von Wahrscheinlichkeitstheorie und Statistik Voraussetzung. Der Abschnitt hat vergleichsweise stärker theoretischen Charakter und wird in der Vorlesung eher überblicksmäßig behandelt, mit Fokus auf den grundsätzlichen Anliegen, die hinter der Theorie stehen.

5.3.1 Der Begriff des Maßes

Inhalt in Kurzfassung: Beim Begriff „Maß“ denken wir vor allem an Längen-, Flächen- und Volumsmaße, aber auch an Wahrscheinlichkeitsmaße (siehe vor allem Mathematik 2). Begrifflich ist ein Maß μ eine Abbildung, die gewissen Teilmengen A einer gemeinsamen Obermenge X (für uns zunächst am wichtigsten: $X = \mathbb{R}$) als ihre Maßzahl $\mu(A)$ eine nichtnegative reelle Zahl oder ∞ zuordnet. Neben $\mu(\emptyset) = 0$ lautet die wichtigste Forderung an ein Maß: Die Maßzahl paarweise disjunkter Vereinigungen zweier, endlich vieler oder auch abzählbar unendlich vieler Mengen ist die (endliche oder unendliche) Summe der einzelnen Maßzahlen. Vom Definitionsbereich eines Maßes fordert man, dass es sich um eine sogenannte σ -Algebra (enthält \emptyset und X und ist abgeschlossen bezüglich Komplementbildung sowie abzählbaren Durchschnitten und Vereinigungen) handelt. Besonders wichtig ist das Lebesguemaß λ auf $X = \mathbb{R}$, das jedem Intervall $[a, b]$ seine Länge $\lambda([a, b]) := b - a$ zuordnet. Es ist auch für sehr komplizierte Mengen definiert, wie zum Beispiel für die Cantormenge. Sie ist das Paradebeispiel einer überabzählbaren Menge vom Maß 0.

Ob wir Längen, Flächen, Volumina oder andere Quantitäten wie Wahrscheinlichkeiten messen, immer wieder erweist sich der Begriff des Maßes als geeigneter begrifflicher Rahmen: Man weist gewissen Mengen Zahlenwerte (ihr Maß) zu derart, dass der Vereinigung disjunkter Mengen die Summe der einzelnen Maße entspricht. Man beachte die Analogie zur Anzahlbestimmung von Vereinigungsmengen. Eine Schwierigkeit, die auf den ersten Blick nicht unbedingt

auffällt, besteht darin, dass nicht unbedingt allen (beliebig komplizierten) Mengen in sinnvoller Weise ein Maß zugeordnet werden kann. Diese Schwierigkeit beschäftigt vor allem theoretische Mathematiker. Hier brauchen uns die damit verbundenen Probleme wenig zu kümmern. Zwecks mathematischer Korrektheit brauchen wir lediglich die folgende Definition.

Definition 5.3.1.1. Sei X eine Menge und \mathcal{A} eine Menge von Teilmengen von X , die wir die (bezüglich \mathcal{A}) **messbaren Mengen** nennen. Wir nennen \mathcal{A} eine **σ -Algebra** (sprich: Sigma-Algebra) auf X , wenn folgende Bedingungen erfüllt sind:

1. $\emptyset, X \in \mathcal{A}$, die leere und die gesamte Menge sind also messbar.
2. Für jedes (messbare) $A \in \mathcal{A}$ liegt auch $X \setminus A$ in \mathcal{A} (ist also messbar).
3. Für eine Folge von (messbaren) Mengen $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, liegt auch die Vereinigung $\bigcup_{n \in \mathbb{N}} A_n$ in \mathcal{A} (ist also messbar).

Diese Definition garantiert, dass alle üblichen mengentheoretischen Operationen innerhalb der σ -Algebra \mathcal{A} der messbaren Mengen ausführbar sind, und zwar sowohl für endlich wie auch für abzählbar unendlich viele Mengen. Zum Beispiel ist wegen

$$\bigcap_{n \in \mathbb{N}} A_n = X \setminus \bigcup_{n \in \mathbb{N}} (X \setminus A_n)$$

auch der Durchschnitt einer (abzählbaren) Folge messbarer Mengen A_n wieder messbar.

Ein einfaches Beispiel einer σ -Algebra auf einer Menge X ist die sogenannte **Potenzmenge** von X , das ist die Menge aller Teilmengen von X . Leider müssen wir uns in vielen interessanten Fällen, so auch beim Lebesgueschen Maß, dem für uns an dieser Stelle weitaus wichtigsten Beispiel, mit kleineren σ -Algebren begnügen. Die wichtigste Rolle von σ -Algebren besteht darin, als Definitionsbereich von Maßen zu fungieren.

Definition 5.3.1.2. Sei \mathcal{A} eine σ -Algebra. Eine Funktion $\mu : \mathcal{A} \rightarrow [0, \infty]$ heißt **Maß**, wenn $\mu(\emptyset) = 0$ und die sogenannte **σ -Additivität** erfüllt ist:

Sind die messbaren Mengen A_n , $n \in \mathbb{N}$, paarweise disjunkt, d.h. $A_m \cap A_n = \emptyset$ für alle $m \neq n$, so gilt

$$\mu \left(\bigcup_{n \in \mathbb{N}} A_n \right) = \sum_{n \in \mathbb{N}} \mu(A_n).$$

Man beachte, dass aus der σ -Additivität auch die **endliche Additivität** folgt. Setzt man nämlich $A_n = \emptyset$ für alle $n = n_0, n_0 + 1, \dots$, so nimmt σ -Additivität die Gestalt

$$\mu(A_0 \cup A_1 \cup \dots \cup A_{n_0-1}) = \mu(A_0) + \mu(A_1) + \dots + \mu(A_{n_0-1})$$

an. Diese Bedingung besagt schlicht, dass sich Maßzahlen aufaddieren, wenn wir nichtüberlappende Mengen zu ihrer Vereinigung zusammenfassen. Das ist eigentlich eine Selbstverständlichkeit. Denn genau diese Situation liegt bei Längen-, Flächen- und Volumsmessungen genauso vor wie bei Wahrscheinlichkeiten und so ziemlich allem, wofür wir das Wort *messen* verwenden.

Übungsaufgabe 291. (P) Ein einfaches, aber trotzdem wichtiges Beispiel eines Maßes ist das **Zählmaß** μ . Es ist auf einer beliebigen Menge X definiert durch $\mu(A) := |A|$, wenn $A \subseteq X$ eine endliche Teilmenge von X ist und durch $\mu(A) := \infty$ für alle unendlichen $A \subseteq X$. Die Algebra \mathcal{A} ist also die Potenzmenge (die Menge aller Teilmengen) von X . Überprüfen Sie, dass es sich beim Zählmaß tatsächlich um ein Maß handelt.

In diesem Kapitel denken wir meist an das gewöhnliche Längenmaß auf der Zahlengeraden \mathbb{R} , das wir mit λ bezeichnen. Einzelne Punkte $x \in \mathbb{R}$ haben die Länge 0, also $\lambda(\{x\}) = 0$ für alle $x \in \mathbb{R}$. Generell nennt man Mengen M vom Maß 0 **Nullmengen**. Wegen der σ -Additivität sind endliche wie auch abzählbar unendliche Teilmengen von \mathbb{R} stets Nullmengen bezüglich λ . Bemerkenswert ist, dass es sogar überabzählbare Nullmengen gibt. Das Standardbeispiel ist die sogenannte **Cantormenge**. Wir beginnen mit $M_0 := [0, 1]$. Wir erwarten $\lambda(M_0) = \lambda([0, 1]) = 1$, weil das Einheitsintervall $[0, 1]$ ja die Länge 1 hat. Nehmen wir das mittlere Drittel heraus (nicht die Randpunkte), so bleibt die Menge $M_1 := [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ übrig, die das Maß $\lambda(M_1) = \frac{2}{3}$ hat. Aus jedem der beiden Teilintervalle von M_1 entfernen wir wieder das mittlere Drittel und erhalten eine Menge M_2 vom Maß $\lambda(M_2) = (\frac{2}{3})^2$. Auf diese Weise fahren wir fort mit Mengen M_n vom Maß $\lambda(M_n) = (\frac{2}{3})^n$. Die als Schnittmenge

$$M := \bigcap_{n \in \mathbb{N}} M_n$$

definierte Cantormenge ist in allen M_n enthalten, hat also kleineres Maß als jedes M_n . Es folgt $0 \leq \lambda(M) \leq \lambda(M_n) = (\frac{2}{3})^n \rightarrow 0$ für $n \rightarrow \infty$, was nur für $\lambda(M) = 0$ möglich ist. Trotzdem ist M überabzählbar. Wie man sich leicht überlegt, liegen in M nämlich alle Zahlen x der Bauart $x = \sum_{n=1}^{\infty} \frac{2a_n}{3^n}$ mit $a_n \in \{0, 1\}$ (Zifferndarstellung zur Basis 3). Jedem solchen x kann $f(x) := \sum_{n=1}^{\infty} \frac{a_n}{2^n} \in [0, 1]$ (Binärdarstellung) zugeordnet werden. Die so definierte Funktion $f : M \rightarrow [0, 1]$ ist offenbar surjektiv. Denn jede Zahl in $[0, 1]$ lässt sich in der beschriebenen Weise binär darstellen, tritt daher als Bild unter f auf. Weil $[0, 1]$ überabzählbar ist, muss demnach auch M überabzählbar sein.

5.3.2 Vertiefende Bemerkungen zum Lebesgueschen Maß

Inhalt in Kurzfassung: Dieser Unterabschnitt dient der Vertiefung des theoretischen Verständnisses der Maßtheorie und ihrer grundsätzlichen Anliegen. Als zentrales Ergebnis geben wir ohne Beweis einen Satz an, durch den das n -dimensionale Lebesguesche Maß λ_n eindeutig charakterisiert wird. Außerdem

konstruieren wir eine Menge, die nicht messbar ist und erwähnen kurz das berühmte Paradoxon von Banach-Tarski („aus 1 mach 2“).

Im vorangegangenen Unterabschnitt sind wir naiv vorgegangen. Gemäß unserer intuitiven Vorstellung einer Längenmessung in \mathbb{R} , haben wir gewissen Mengen A ihr Maß $\lambda(A)$ zugewiesen, wobei alles auf der Festlegung $\lambda([a, b]) = b - a$ für $a \leq b$ und auf der σ -Additivität von λ fußte. Es ist aber keinesfalls trivial, dass ein solches Maß λ überhaupt existiert. Tatsächlich ist der Satz über die Existenz des Lebesgueschen Maßes ein tiefliegender, dessen Beweis beträchtlichen Aufwand erfordert, den wir hier nicht treiben werden. Das Hauptergebnis gilt nicht nur für das eindimensionale Längenmaß $\lambda = \lambda_1$ auf \mathbb{R} , sondern ganz analog für das zweidimensionale Flächenmaß λ_2 auf \mathbb{R}^2 , das dreidimensionale Volumensmaß λ_3 auf \mathbb{R}^3 und ganz allgemein für das n -dimensionale Lebesguemaß λ_n auf \mathbb{R}^n . Weil wir uns in Mathematik 2 intensiv auch mit dem zwei- und dreidimensionalen Fall beschäftigen werden, folgt nun der allgemeine, alle Dimensionen erfassende Satz.

Satz 5.3.2.1. *Auf der Menge \mathbb{R}^n , $n \in \mathbb{N}$, gibt es eine σ -Algebra \mathcal{A}_n und ein auf \mathcal{A}_n definiertes Maß λ_n mit folgenden beiden Eigenschaften:*

Großer Definitionsbereich des Maßes: \mathcal{A}_n enthält alle Intervalle bzw. achsenparallelen Rechtecke bzw. achsenparallelen Quader bzw. alle verallgemeinerten achsenparallelen n -dimensionale Quader.¹

Die Werte des Maßes entsprechen den erwarteten: Für Intervalle $I = [a, b]$ gilt $\lambda(I) = b - a$. Für n -dimensionale achsenparallele Quader $Q = I_1 \times I_2 \times \dots \times I_n$, $I_k = [a_k, b_k]$, gilt

$$\lambda_n(Q) = \lambda_n(I_1 \times I_2 \times \dots \times I_n) = \lambda(I_1) \cdot \dots \cdot \lambda(I_n) = (b_1 - a_1) \cdot \dots \cdot (b_n - a_n).$$

(Intervalle erhalten als Maß also ihre **Länge**, Rechtecke ihre **Fläche**, 3-dimensionale achsenparallele Quader ihr **Volumen** im üblichen Sinn, n -dimensionale achsenparallele Quader ihr Volumen in einem entsprechend verallgemeinerten Sinn.)

Man kann von \mathcal{A}_n und λ_n überdies folgende Bedingungen fordern, wodurch sogar Eindeutigkeit hergestellt wird:

Vollständigkeit des Maßes: Wann immer $A \subseteq N$ gilt mit einer Nullmenge $N \in \mathcal{A}_n$, also $\lambda(N) = 0$, dann ist auch $A \in \mathcal{A}_n$ messbar und eine Nullmenge.

¹Dabei verstehen wir unter einem n -dimensionalen achsenparallelen Quader eine Menge $Q \subseteq \mathbb{R}^n$ der Form $Q = I_1 \times I_2 \times \dots \times I_n$ mit Intervallen $I_k = [a_k, b_k]$. Folglich enthält \mathcal{A}_n auch alle abzählbaren Vereinigungen von Intervallen, Rechtecken, Quadern etc., somit alle offenen und alle abgeschlossenen Mengen, somit auch deren abzählbare Vereinigungen und Schnitte etc. Diese Bedingung garantiert also, dass sehr viele und sogar ziemlich komplizierte Mengen messbar sind.)

Regularität des Maßes: Zu jeder beschränkten und messbaren Menge $A \in \mathcal{A}_n$ und zu jedem $\varepsilon > 0$ gibt es eine Menge E , die endliche Vereinigung von Intervallen (bzw. achsenparallelen Rechtecken oder Quadern im üblichen oder auch verallgemeinerten Sinn) ist und A in folgendem Sinn approximiert: Es gibt eine Folge von offenen Intervallen (bzw. achsenparallelen Rechtecken, Quadern im üblichen oder im verallgemeinerten Sinn) I_0, I_1, \dots mit $\sum_{k \in \mathbb{N}} \lambda(I_k) < \varepsilon$ und $A \triangle E \subseteq \bigcup_{k \in \mathbb{N}} I_k$.

Die Elemente dieser eindeutig bestimmten σ -Algebra \mathcal{A}_n heißen die (n -dimensionalen) **(Lebesgue-)messbaren Mengen**. Das (ebenfalls eindeutige) Maß λ_n heißt das (n -dimensionale) **Lebesguesche Maß** oder **Lebesguemaß**. Das Maß λ_n ist überdies

bewegungsinvariant: Für jedes (messbare) $A \in \mathcal{A}_n$ und jede **Bewegung** $\beta : \mathbb{R}^n \rightarrow \mathbb{R}^n$ des \mathbb{R}^n (d.h. jede beliebige Verkettung von Translationen, Drehungen und Spiegelungen) gilt $\lambda_n(\beta(A)) = \lambda_n(A)$.

Der einfache Zusammenhang zwischen dem ein- und dem höherdimensionalen Lebesguemaß, der in der Formel

$$\lambda_n(I_1 \times \dots \times I_n) = \lambda(I_1) \cdot \dots \cdot \lambda(I_n)$$

zum Ausdruck kommt, macht λ_n zum n -fachen Produktmaß von λ (bzw. zur Vervollständigung desselben). Man schreibt dafür auch $\lambda_n = \lambda \otimes \lambda \otimes \dots \otimes \lambda$ (n -mal). Wir werden davon in Mathematik 2 reichlich Gebrauch machen, etwa beim Satz von Fubini über die Berechnung höherdimensionaler Integrale durch Iteration eindimensionaler oder in der Stochastik bei Produkträumen zwecks Modellierung unabhängiger Zufallsgrößen.

Doch zurück zum Lebesgueschen Maß: So natürlich und einleuchtend vor allem die ersten drei Eigenschaften von λ_n aus Satz 5.3.2.1 auch erscheinen, so mache man sich klar, wie sehr alles an den reellen Zahlen liegt. Angenommen, wir hätten nur die rationalen Zahlen zur Verfügung, wollten an der σ -Additivität festhalten und gleichzeitig auch rationalen Intervallen, also Mengen $[a, b] \cap \mathbb{Q}$ als Maß die Länge $b - a$ zuordnen. Dann wären einpunktige wie auch abzählbare Mengen Nullmengen. Aber auch die rationalen Intervalle wären abzählbar, was sich nicht mit einer positiven Länge vereinbaren lässt. \mathbb{Q} wäre also für die Zwecke der Maß- und Integrationstheorie ungeeignet. Wieder erweist sich \mathbb{R} als überlegen und in vielerlei Hinsicht unverzichtbar. Man beachte in diesem Zusammenhang insbesondere die Wichtigkeit der Unterscheidung zwischen abzählbaren und überabzählbaren Mengen.

Auch wenn hier nicht auf die Konstruktion von \mathcal{A}_n und λ_n sowie auf den Beweis von Satzes 5.3.2.1 eingegangen werden kann, sollte man Folgendes mitnehmen: Es führt zu keinen Widersprüchen, wenn man mit λ in der Weise operiert, wie wir es weiter oben z.B. im Zusammenhang mit der Cantormenge getan haben. Wir werden sehen, dass darauf eine Integrationstheorie begründet werden kann, die mit dem Riemannsches Integral zusammenpasst, uns aber zusätzliche Möglichkeiten eröffnet.

Übungsaufgabe 292. (E) Sei $f : [a, b] \rightarrow [0, \infty)$ und sei $M := \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq b, 0 \leq y \leq f(x)\}$.

1. Skizzieren Sie die Menge M .
2. Argumentieren Sie anschaulich, warum

$$\lambda_2(M) = \int_a^b f(x) dx$$

gelten sollte.

3. Bestätigen Sie mittels Rückgriff auf weiter oben angegebene Eigenschaften des (in diesem Fall zweidimensionalen) Lebesguemaßes λ_2 und auf die Definition des R -Integrals über Ober- und Untersummen, dass die Gleichung aus dem zweiten Teil der Aufgabe tatsächlich gilt, wenn f R -integrierbar ist.

Man muss einigen begrifflichen und beweistechnischen Aufwand treiben, um das Lebesguemaß vermittle Satz 5.3.2.1 eindeutig definieren zu können. Wie schon angedeutet, ist das deshalb notwendig, weil nicht alle Teilmengen von \mathbb{R} (oder auch \mathbb{R}^n) messbar sind. Die interessierte Leserin und der interessierte Leser werden sich fragen, wie denn eine nicht messbare Menge aussehen mag. Es folgt eine Skizze der Konstruktion mit anschließender Übungsaufgabe.

Die klassische Konstruktion einer nicht messbaren Teilmenge A von \mathbb{R} läuft wie folgt. Man denkt sich das Einheitsintervall $[0, 1]$ zu einem Kreis S aufgewickelt, wo die Punkte 0 und 1 wieder zusammenfallen. Nun zerlegt man S in viele Klassen K . Und zwar kommt ein beliebiges $x \in [0, 1]$ mit all jenen $y \in [0, 1]$ in dieselbe Klasse, für die die Differenz $x - y$ rational ist. Nimmt man aus jedem K genau ein Element $a_K \in K$ heraus, so erweist sich die Menge A all dieser a_K als nicht messbar, weil die Annahme einer Längenmaßzahl $\lambda(A)$ zu einem Widerspruch führt. Um das einzusehen, sind folgende Überlegungen durchzuführen:

1. Die Menge $[0, 1]$ zerfällt tatsächlich in lauter Klassen der beschriebenen Art; und zwar so, dass jedes $x \in [0, 1]$ genau einer Klasse angehört. (Man hat sich vor allem zu überlegen, dass wenn y in dieselbe Klasse wie x und z kommt, auch x und z in derselben Klasse liegen müssen.)
2. Wegen $\emptyset \subseteq A \subseteq [0, 1]$ muss für die (hypothetische) Maßzahl $\lambda(A)$ die Ungleichung $0 = \lambda(\emptyset) \leq \lambda(A) \leq \lambda([0, 1]) = 1$ gelten. Wir werden zwei mögliche Fälle unterscheiden: $\lambda(A) = 0$ (Fall 1) oder $0 < \lambda(A) \leq 1$ (Fall 2).
3. Wenn man A um eine rationale Zahl $q \in \mathbb{Q}$ verschiebt (auf dem Kreis dreht), entsteht eine Menge $A_q = A + q$, die wegen der Translationsinvarianz von λ gleiches Maß hat wie A , also $\lambda(A + q) = \lambda(A)$.

4. Jedes $x \in S$ liegt in genau einem A_q . Deshalb ist

$$S = \bigcup_{q \in \mathbb{Q} \cap [0,1)} A_q$$

eine abzählbare disjunkte Vereinigung.

5. Im Fall 1 ($\lambda(A_q) = \lambda(A) = 0$) folgt daraus der Widerspruch

$$1 = \lambda(S) = \lambda\left(\bigcup_{q \in \mathbb{Q} \cap [0,1)} A_q\right) = \sum_{q \in \mathbb{Q} \cap [0,1)} \lambda(A_q) = \sum_{q \in \mathbb{Q} \cap [0,1)} 0 = 0.$$

6. Im Fall 2 ($\lambda(A_q) = \lambda(A) > 0$) folgt daraus der Widerspruch

$$1 = \lambda(S) = \lambda\left(\bigcup_{q \in \mathbb{Q} \cap [0,1)} A_q\right) = \sum_{q \in \mathbb{Q} \cap [0,1)} \lambda(A_q) = \sum_{q \in \mathbb{Q} \cap [0,1)} \lambda(A) = \infty.$$

Übungsaufgabe 293. (E) *Behandeln Sie die oben skizzierten sechs Schritte zur Konstruktion einer nicht messbaren Menge A ausführlich.*

Dass es Mengen gibt, die nicht messbar sind, lässt sich noch an viel spektakuläreren Phänomenen erkennen wie dem **Paradoxon von Banach-Tarski**: Mit mengentheoretischen Mitteln kann man zeigen, dass die volle dreidimensionale Einheitskugel in fünf Mengen zerlegt werden kann, die, im Raum bewegt und geschickt wieder zusammengesetzt, zwei volle Kugeln derselben Größe ergeben. Das Gesamtvolumen wurde also durch Bewegungen verdoppelt. Auf den ersten Blick scheint das der Bewegungsinvarianz des dreidimensionalen Lebesguemaßes λ_3 zu widersprechen. Das ist aber nicht der Fall. Man hat lediglich die Konsequenz zu ziehen, dass die Teilmengen, in die die Vollkugel zerlegt worden ist, nicht messbar, also geradezu unvorstellbar kompliziert sind. Denn nur für messbare Mengen gilt die Additivität des Maßes.

5.3.3 Neuinterpretation des Riemannintegrals als Lebesgueintegral

Inhalt in Kurzfassung: Eine etwas veränderte Sichtweise lässt uns das Riemann-Integral auch als Integral bezüglich des Lebesguemaßes verstehen, als sogenanntes Lebesgue-Integral, dessen Definition wir nun geben werden. Wann immer das Riemann-Integral für einen gegebenen Integranden existiert, so auch das Lebesgue-Integral – darüber hinaus aber auch in allgemeineren Fällen. In den meisten Situationen, die für den Ingenieur wichtig sind, ist der Unterschied der beiden Begriffe nicht relevant. Es ist aber wünschenswert, sich nicht verwirren zu lassen, wenn man in der Literatur auf die etwas andere Notation stößt. Außerdem werden wir in Mathematik 2 in der Stochastik großen Nutzen aus dem

allgemeineren Begriff ziehen.

Wir wollen nochmals den Begriff des R-Integrals unter einem geringfügig veränderten Gesichtspunkt rekapitulieren. Und zwar gehen wir aus vom Konzept der Untersumme

$$U(f, Z) = \sum_{i=1}^n \left(\inf_{x \in [x_{i-1}, x_i]} f(x) \right) (x_i - x_{i-1})$$

zu einer Zerlegung $Z = \{a = x_0 \leq x_1 \leq x_2 \leq \dots \leq x_n = b\}$ für eine Funktion $f : [a, b] \rightarrow \mathbb{R}$, von der wir zunächst $f \geq 0$ annehmen wollen. Die Summe rechts stimmt mit dem Integral $\int_a^b f_Z(x) dx$ überein, wenn wir die Funktion f_Z als stückweise konstant ansetzen mit Werten $f_Z(x) = c_i := \inf_{x \in [x_{i-1}, x_i]} f(x)$ für $x_{i-1} \leq x < x_i$ (und $f_Z(x_n) = f_Z(b) := c_n$). Insbesondere gilt $f_Z \leq f$ und somit auch

$$\int_a^b f_Z(x) dx \leq \int_a^b f(x) dx,$$

sofern das R-Integral rechts existiert. In diesem Fall kann bei hinreichend feiner Zerlegung Z das Integral links beliebig nahe an jenes rechts angenähert werden. Also lässt sich auch schreiben

$$\int_a^b f(x) dx = \sup_Z \int_a^b f_Z(x) dx,$$

wobei Z alle Zerlegungen von $[a, b]$ durchläuft.

Wir wollen das Supremum rechts sogar auf eine größere Klasse von Funktionen ausdehnen als die der f_Z , wie sie in der oben beschriebenen Weise aus Zerlegungen Z von $[a, b]$ entstehen. Dafür beobachten wir zunächst, dass sich die Funktion Z auch in der Form

$$f_Z(x) = \sum_{i=1}^n c_i \mathbf{1}_{A_i}$$

mit $A_i = [x_{i-1}, x_i)$ für $i = 1, \dots, n-1$ und $A_n = [x_{n-1}, x_n]$ schreiben lässt, und das Integral von f_Z als

$$\int_a^b f_Z(x) dx = \sum_{i=1}^n c_i \lambda(A_i).$$

Damit sind wir an dem Punkt angelangt, wo sich in natürlicher Weise ein modifizierter Integralbegriff anbietet. Die Modifikation gegenüber dem R-Integral besteht darin, dass wir nicht nur Intervalle, sondern beliebige Mengen A_i , für die $\lambda(A_i)$ definiert ist, also beliebige messbare Mengen A_i zulassen. Statt $\int_a^b f(x) dx$ schreiben wir für diesen neu zu definierenden Integralbegriff $\int_{[a,b]} f d\lambda$. Die korrekten Definitionen lauten wie folgt.

Sei $X \subseteq \mathbb{R}$ eine messbare Menge. Dann heißt jede Linearkombination

$$t = \sum_{i=1}^n c_i \mathbf{1}_{A_i}$$

von charakteristischen Funktionen $\mathbf{1}_{A_i}$ messbarer Teilmengen $A_1, \dots, A_n \subseteq X$ mit $c_i \geq 0$ eine (positive) **Treppenfunktion** auf X . Das Integral von t ist definiert durch

$$\int t \, d\lambda = \int_X t \, d\lambda := \sum_{i=1}^n c_i \lambda(A_i).$$

Dieser Wert kann, sofern A_i unendliches Maß hat, auch ∞ sein. Er ist aber (wie man sich relativ leicht überlegen kann) unabhängig von der speziellen Darstellung von t . (Beispiel: Für $X = [0, 2]$ und $n = 2$ wird mit $A_1 = [0, 1]$, $A_2 = (1, 2]$, $c_1 = 1$ und $c_2 = 2$ dieselbe Funktion t dargestellt wie mit $A_1 = [0, 2]$, $A_2 = (1, 2]$, $c_1 = 1$ und $c_2 = 1$. Trotzdem liefert die obige Formel für $\int t \, d\lambda$ in beiden Fällen den gleichen Wert 3.)

Ist das Integral wie oben für positive Treppenfunktionen t definiert, so können wir für ein beliebiges $f : X \rightarrow \mathbb{R}$ mit $f \geq 0$ definieren:

$$\int_X f \, d\lambda = \int f \, d\lambda := \sup_{t: t \leq f} \int t \, d\lambda,$$

wobei sich das Supremum über alle positiven Treppenfunktionen t mit $t \leq f$ erstreckt. Auch dieser Wert kann unendlich sein, sogar bei $\lambda(X) < \infty$, sofern f unbeschränkt ist (z.B. $X = (0, 1]$ und $f(x) := \frac{1}{x}$). Ist die Bedingung $f \geq 0$ nicht erfüllt, stellt man $f = f^+ - f^-$ als Differenz des Positivteils $f^+(x) := \max\{f(x), 0\}$ und des Negativteils $f^-(x) := \max\{-f(x), 0\}$ dar und definiert $I^+(f) := \int f^+ \, d\lambda$ sowie $I^-(f) := \int f^- \, d\lambda$ und damit

$$I(f) := \int f \, d\lambda := I^+(f) - I^-(f) = \int f^+ \, d\lambda - \int f^- \, d\lambda.$$

Zu einem Problem kommt es nur, wenn $I^+(f) = I^-(f) = \infty$. In allen anderen Fällen ist $I(f) = \int f \, d\lambda$ definiert, sei es als endlicher Wert (wenn $I^+(f), I^-(f) < \infty$), als $\infty = \infty - c$ (wenn $I^+(f) = \infty$ und $I^-(f) = c \in \mathbb{R}$) oder als $-\infty = c - \infty$ (wenn $I^+(f) = c \in \mathbb{R}$ und $I^-(f) = \infty$).

Hat man es mit komplexwertigen Funktionen zu tun, d.h. mit Funktionen der Gestalt $x \mapsto f(x) + ig(x)$, wobei die Funktionen f und g reellwertig sind, so setzt man

$$\int f + ig \, d\lambda := \int f \, d\lambda + i \int g \, d\lambda,$$

vorausgesetzt Real- und Imaginärteil sind beide endlich. Damit ist das **Lebesguesche Integral** $\int f \, d\lambda$ für alle f definiert, für die das sinnvoll möglich ist.

Bei der R-Integrierbarkeit einer Funktion f war Voraussetzung, dass Ober- und Untersumme beliebig nahe gebracht werden können. Beim Lebesgue-Integral war bisher noch von überhaupt keiner vergleichbaren Einschränkung die Rede.

Für eine befriedigende Integrationstheorie erweist sich eine solche allerdings sehr wohl als notwendig, auch wenn sie viel weniger restriktiv ist als im Riemannschen Fall. Und zwar geht es um die sogenannte **Messbarkeit** der Funktion f . Um f nämlich im Sinne des Lebesgueschen Integrals gut durch Treppenfunktionen approximieren zu können, muss es möglich sein, für beliebig vorgegebene $a < b$ das Maß der Menge jener x zu bestimmen, wo $a < f(x) < b$ gilt. Eine Funktion $f : X \rightarrow \mathbb{R}$ heißt also **messbar**, wenn für jedes Intervall $I = (a, b)$ (ob offen oder abgeschlossen läuft auf dasselbe hinaus) das Urbild $f^{-1}(I) = \{x \in X : f(x) \in I\}$ messbar ist, d.h. in \mathcal{A} liegt. Relativ leicht macht man sich klar, dass stetige Funktionen stets messbar sind, darüber hinaus aber viele weitere. Ja, man kann sagen: Alle reellen Funktionen, die einem naiverweise in den Sinn kommen, sind messbar.

Die Lebesguesche Integrationstheorie, zu der im nachfolgenden Unterabschnitt die wichtigsten Tatsachen zusammengestellt werden sollen, bezieht sich durchwegs auf Integrale messbarer Funktionen, wo überdies das Integral im obigen Sinn als Wert aus $\mathbb{R} \cup \{-\infty, \infty\}$ definiert ist. Ausgeschlossen sind also von den messbaren f nur jene mit $\int f^+ d\lambda = \int f^- d\lambda = \infty$. Genau dann, wenn beides, $I^+(f)$ und $I^-(f)$, endliche reelle Zahlen sind, ist auch $I(|f|) := \int |f| d\lambda < \infty$, und man nennt f **integrierbar** oder, wenn man den Unterschied zur R-Integrierbarkeit betonen will, **Lebesgue-** oder kurz **L-integrierbar**.

Wichtig ist für uns, dass alle Riemannintegrale auch als Lebesgueintegrale interpretiert werden können. Es gilt nämlich der folgende, nicht sehr schwierig zu beweisende Satz:

Satz 5.3.3.1. *Ist die Funktion $f : [a, b] \rightarrow \mathbb{R}$ mit $a < b$ R-integrierbar, so ist sie auch im Lebesgueschen Sinn integrierbar, und es gilt*

$$\int_a^b f(x) dx = \int_{[a,b]} f d\lambda.$$

Übungsaufgabe 294. *(E) Begründen Sie, dass in Satz 5.3.3.1 jedenfalls \leq gilt. Anleitung: Jede Riemann-Untersumme lässt sich als Lebesgueintegral einer Treppenfunktion $\leq f$ interpretieren. (Für den Beweis der Ungleichung \geq empfiehlt es sich, Hilfsmittel aus 5.3.4 zu verwenden, die jetzt noch nicht zur Verfügung stehen.)*

Bemerkenswert ist, dass das Lebesguesche Maß auch das adäquate Instrument ist, um Riemann-Integrierbarkeit besser zu verstehen. Es gilt nämlich das folgende **Kriterium für R-Integrierbarkeit**:

Satz 5.3.3.2. *Eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann R-integrierbar wenn die Menge der Unstetigkeitsstellen von f bezüglich des Lebesgueschen Maßes eine Nullmenge ist.²*

²Nebenbei sei bemerkt, dass es auch einen entsprechenden Satz gibt, der Messbarkeit von $f : D \rightarrow \mathbb{R}$, $D \subseteq \mathbb{R}$, über Stetigkeitseigenschaften von f charakterisiert, nämlich den Satz von *Lusin*. Ihm zufolge lautet die zur Messbarkeit äquivalente Bedingung: Zu jedem $\varepsilon > 0$ gibt es eine messbare Ausnahmemenge A mit $\lambda(A) < \varepsilon$ derart, dass die Einschränkung von f auf $D \setminus A$ stetig ist.

Dass wir mit dem Lebesgueschen Integralbegriff auch wirklich etwas dazugewinnen, zeigt z.B. die Dirichletsche Sprungfunktion $f = \mathbf{1}_A$ mit $A := \mathbb{Q} \cap [0, 1]$ auf $[0, 1]$. Am Ende von 5.1.1 haben wir uns klar gemacht, dass sie nicht R-integrierbar ist. Weil sie nur die Werte 0 und 1 annimmt, ist sie aber eine Treppenfunktion im Lebesgueschen Sinn. Die Menge A ist abzählbar, also messbar vom Maß 0. Entsprechend ist ihr Komplement auf $[0, 1]$, nämlich die Menge $B := [0, 1] \setminus \mathbb{Q}$, ebenfalls messbar vom Maß $\lambda(B) = 1$ (wegen $1 = \lambda([0, 1]) = \lambda(A) + \lambda(B) = 0 + \lambda(B)$). Folglich ist $f = \mathbf{1}_A$ messbar (nur die messbaren Mengen \emptyset , A und B und $[0, 1]$ treten als Urbilder unter f auf) mit Integral

$$\int_{[0,1]} \mathbf{1}_A d\lambda = 1 \cdot \lambda(A) + 0 \cdot \lambda(B) = 1 \cdot 0 + 0 \cdot 1 = 0.$$

Integrierbarkeit im Lebesgueschen Sinn entspricht in vielerlei Hinsicht der absoluten Konvergenz unendlicher Reihen. Das wird offensichtlich, wenn man das Zählmaß μ auf \mathbb{N} zugrunde legt. Zur Erinnerung: Für alle Teilmengen $A \subseteq X := \mathbb{N}$ ist $\mu(A) := |A|$ die Mächtigkeit der Menge A , wobei $\mu(A) = \infty$ zu setzen ist, sofern A unendlich ist. Fassen wir eine Folge von reellen Gliedern a_n als Funktion $f : \mathbb{N} \rightarrow \mathbb{R}$, $n \mapsto a_n$ auf, so gilt in unserer Notation

$$\sum_{n=0}^{\infty} a_n = \int_{\mathbb{N}} f d\mu,$$

wobei das Integral genau dann eine reelle Zahl, f somit integrierbar bezüglich des Zählmaßes ist, wenn die Reihe absolut konvergiert.

Was die Grundidee betrifft, kann man sich den Unterschied zwischen Riemannschem und Lebesgueschem Integral auch folgendermaßen einprägen. Beim Riemannschem Integral unterteilt man den Definitionsbereich in kleine Teilintervalle $A_i = [x_{i-1}, x_i]$. Auf jedem dieser Intervalle ersetzt man f durch eine konstante Funktion nahe f (bei Untersummen $\leq f$, bei Obersummen $\geq f$, bei Riemannsummen zu einer Belegung durch irgendeinen Funktionswert y_i dazwischen) und berechnet das Integral der Approximation als Summe $\sum y_i(x_i - x_{i-1}) = \sum y_i \lambda(A_i)$ von Rechtecksflächen. Beim Lebesgueschen Integral hingegen beginnt man mit einer Unterteilung der y -Achse und sucht zu je zwei benachbarten Teilungspunkten $y_j < y_{j+1}$ die Menge B_j , bestehend aus allen x auf der x -Achse, wo $y_j \leq f(x) < y_{j+1}$. Als Approximation für das Integral von f nimmt man aber wieder die analog gebaute Summe, nämlich $\sum_j y_j \lambda(B_j)$.

Die Lebesguesche Sichtweise aufs Integral lässt sich immer dann anwenden, wenn ein Maß zur Verfügung steht. Insbesondere ist das der Fall bei höherdimensionalen Integralen, wo der Definitionsbereich keine Teilmenge von \mathbb{R} ist, sondern zum Beispiel von \mathbb{R}^2 (Flächenmaß) oder \mathbb{R}^3 (Volumsmaß), aber auch wenn die Trägermenge die eines Wahrscheinlichkeitsraumes ist. In Mathematik 2 werden diese Anwendungen wichtig sein.

Übungsaufgabe 295. (E) Als Einstimmung auf die Stochastik (Wahrscheinlichkeitstheorie und Statistik) betrachte man die Menge aller Paare (a, b) mit $a, b \in \{1, 2, 3, 4, 5, 6\}$, die wir als Ergebnis eines Wurfes mit zwei (voneinander

unterscheidbaren) Würfeln W_a und W_b interpretieren. Jeder dieser Versuchsausgänge möge gleiche Wahrscheinlichkeit $\frac{1}{36}$ haben. Wir interessieren uns für die möglichen Augensummen $a + b \in X := \{2, 3, \dots, 12\}$ und die Wahrscheinlichkeiten dafür. Definieren Sie ein Maß μ , das jeder Teilmenge $T \subseteq \{2, 3, \dots, 12\}$ die Wahrscheinlichkeit zuordnet, bei einem Wurf mit den beiden Würfeln eine Augensumme zu erzielen, die in T liegt.

5.3.4 Einige Sätze aus der Lebesgueschen Theorie

Inhalt in Kurzfassung: Man kann etwas verkürzt sagen, dass der Lebesguesche Integralbegriff alle Annehmlichkeiten des Riemannsches mit diesem teilt, darüber hinaus aber noch weitere Eigenschaften hat, die aus Sicht der Theorie große Vorteile bieten. Das ist auch der Grund für den Siegeszug des Lebesgueschen Integrals in der Mathematik seit über 100 Jahren und soll hier in aller Kürze angedeutet werden.

Ähnlich wie für das Riemann-Integral kann man auch für das Lebesguesche Integral Linearität beweisen: Das Integral einer Linearkombination ist die entsprechende Linearkombination der Integrale, als Formel

$$\int f + g \, d\lambda = \int f \, d\lambda + \int g \, d\lambda \quad \text{und} \quad \int cf \, d\lambda = c \int f \, d\lambda \quad \text{für } c \in \mathbb{R}.$$

Gleichfalls nicht schwierig zu beweisen sind Monotonie

$$f \leq g \text{ impliziert } \int f \, d\lambda \leq \int g \, d\lambda$$

und die Betragsungleichung

$$\left| \int f \, d\lambda \right| \leq \int |f| \, d\lambda.$$

Auf den ersten Blick ist der Unterschied der Bedingungen für R-Integrierbarkeit und für Messbarkeit (die im Falle beschränkter Funktionen auf Intervallen $[a, b]$ dasselbe bedeutet wie Lebesgue-Integrierbarkeit) vielleicht nicht so deutlich abschätzbar. Aus theoretischer Sicht ist er aber enorm. Das spiegelt sich vor allem in Ergebnissen wider, die sich auf Grenzwerte beziehen. So ist der punktweise Grenzwert messbarer Funktionen stets wieder messbar. Der punktweise Grenzwert von R-integrierbaren Funktionen muss aber nicht mehr Riemann-integrierbar sein.

Beispiel: Weil \mathbb{Q} abzählbar ist, lässt sich $\mathbb{Q} \cap [0, 1] = \{q_n : n \in \mathbb{N}\}$ schreiben. Die $q_n \in \mathbb{Q}$ bilden also eine Abzählung der Menge $A := \mathbb{Q} \cap [0, 1]$. Mit der Bezeichnung $A_n := \{q_0, q_2, \dots, q_n\}$, $n \in \mathbb{N}$, sei $f_n := \mathbf{1}_{A_n}$ auf $[0, 1]$. Jedes f_n hat nur endlich viele (hebbare) Unstetigkeitsstellen, nämlich die q_i mit $i = 0, 1, \dots, n$. Also sind die f_n alle R-integrierbar. Ihr Grenzwert hingegen ist die Dirichletsche Sprungfunktion $\mathbf{1}_A$, also nicht R-integrierbar.

In der Lebesgueschen Theorie kann diese Schwierigkeit nicht auftreten. Es gilt darüber hinaus unter relativ schwachen Bedingungen die Vertauschbarkeit

$$\lim_{n \rightarrow \infty} \int f_n d\lambda = \int \lim_{n \rightarrow \infty} f_n d\lambda$$

von Grenzwert $n \rightarrow \infty$ und Integration. Hinreichend ist zum Beispiel, dass die Konvergenz der f_n gegen f monoton erfolgt (**Satz von der monotonen Konvergenz**) oder, nach dem noch wichtigeren Lebesgueschen **Satz von der dominierten Konvergenz**, dass es eine integrierbare Funktion g gibt mit $|f_n| \leq g$ für alle $n \in \mathbb{N}$. Im Riemannschen Fall dagegen hat man die viel stärkere gleichmäßige Konvergenz der f_n vorauszusetzen, damit auch für die Grenzfunktion f R-Integrierbarkeit garantiert werden kann.

Verwandt damit ist die folgende sehr bemerkenswerte Tatsache: Für jede reelle Zahl $p \geq 1$ kann man aus der Menge aller Lebesgue-messbaren Funktionen f jene auszeichnen, für die die sogenannte **p -Norm** $\|f\|_p := \int |f|^p d\lambda$ einen endlichen Wert annimmt. Fasst man f und g mit dieser Eigenschaft als äquivalent auf ($f \sim g$), wenn sie sich nur auf einer Ausnahmemenge A vom Maß $\lambda(A) = 0$ unterscheiden, als identisch auf, so erhält man als Menge der Äquivalenzklassen die sogenannten L_p -Räume. Bezüglich der Metrik $d(f, g) := \|f - g\|$ handelt es sich dabei sogar um vollständige metrische Räume, die in der Theorie eine herausragenden Rolle spielen. Von besonderem Interesse ist der Fall $p = 2$, für den mit $f \cdot g := \int fg d\lambda$ (im noch wichtigeren Fall komplexwertiger Funktionen $f \cdot g := \int f\bar{g} d\lambda$) sogar ein Skalarprodukt vorliegt. Man spricht dann von einem Hilbertraum, in dessen Rahmen sich erst die Theorie der sogenannten Fourierreihen gut verstehen lässt.

Insgesamt machen diese, teilweise nur angedeuteten Eigenschaften des Lebesgueschen Integralbegriffs die Menge aller im Lebesgueschen Sinn integrierbaren Funktionen zu einem mathematisch viel besser handhabbaren Objekt als die Menge der Riemann-integrierbaren. Für uns wird vor allem die Allgemeinheit dieses Begriffs noch sehr wertvolle Dienste leisten, nämlich, wie schon erwähnt, in der Stochastik, dem letzten Kapitel in Mathematik 2.

Dennoch gibt es auch nicht messbare Funktionen. Eine solche zu konstruieren ist Gegenstand der folgenden Übungsaufgabe.

Übungsaufgabe 296. (P) *Geben Sie unter Zuhilfenahme von Übungsaufgabe 293 eine nicht messbare Funktion an.*

Zum Abschluss des Kapitels (und der Vorlesung Mathematik 1) dürfen wir uns aber wieder elementaren Themen zuwenden.

5.4 Einige Anwendungen und ausgewählte Themen

In diesem letzten Abschnitt versammeln wir einige verstreute Themen, in denen die Integralrechnung eine wichtige Rolle spielt: zuerst, als unmittelbare Ausdehnung des Begriffs des (eigentlichen) R-Integrals, uneigentliche Integrale (5.4.1),

die approximative Berechnung von Summen mit Integralen wie etwa in der Eulerschen Summenformel (5.4.2), numerische Integration (5.4.3) sowie Kurven und ihre Länge (5.4.4).

5.4.1 Uneigentliche Integrale

Inhalt in Kurzfassung: Uneigentliche Integrale entstehen als Grenzwerte von Riemann-Integralen, wenn die Integrationsgrenzen gegen $\pm\infty$ oder gegen eine Stelle streben, in deren Umgebung der Integrand unbeschränkt ist. Unter Umständen haben solche Integrale dennoch einen endlichen Wert. Im Rahmen der Lebesgueschen Theorie lassen sich all diese Situationen einheitlich behandeln. Interessante Anwendungsbeispiele sind die Potenzfunktionen.

Bei der Definition des R-Integrals wurde vorausgesetzt, dass sowohl Integrand als auch Integrationsbereich beschränkt sind. Beim Lebesgueschen Integral fallen diese Einschränkungen weg. Somit sind, um typische Beispiele zu geben, die R-Integrale

$$I(\alpha, a, b) := \int_a^b f_\alpha(x) dx \quad \text{mit} \quad f_\alpha(x) := x^\alpha$$

und $\alpha \in \mathbb{R}$ nur für $b < \infty$ und, sofern $\alpha < 0$, nur für $a > 0$ definiert. Für $\alpha \neq -1$ und $0 < a < b < \infty$ liefern die bekannten Integrationsregeln

$$I(\alpha, a, b) = \int_a^b x^\alpha dx = \frac{1}{\alpha + 1} (b^{\alpha+1} - a^{\alpha+1}).$$

Nehmen wir $\alpha < -1$ an, so konvergiert $b^{\alpha+1}$ für $b \rightarrow \infty$ gegen 0, also

$$I(\alpha, a, \infty) = \int_a^\infty x^\alpha dx := \lim_{b \rightarrow \infty} \int_a^b x^\alpha dx = \int_{[a, \infty)} f_\alpha d\lambda = -\frac{a^{\alpha+1}}{\alpha + 1} = \frac{a^{\alpha+1}}{-\alpha - 1}.$$

Für diesen Wert ist auch die Schreibweise $\int_{[a, \infty)} x^\alpha d\lambda(x)$ üblich, die das Integral im Lebesgueschen Sinn meint, im Term $d\lambda(x)$ aber trotzdem die Integrationsvariable x sichtbar macht.

Ähnlich sieht es für $-1 < \alpha < 0$ aus, wenn $b \in \mathbb{R}$ fest bleibt und der Grenzwert $a \rightarrow 0$ und folglich $a^{\alpha+1} \rightarrow 0$ betrachtet wird:

$$I(\alpha, 0, b) = \int_0^b x^\alpha dx := \lim_{a \rightarrow 0^+} \int_a^b x^\alpha dx = \int_{(0, b]} f_\alpha d\lambda = \int_{(0, b]} x^\alpha d\lambda(x) = \frac{b^{\alpha+1}}{\alpha + 1}.$$

Übungsaufgabe 297. (P) Überlegen Sie sich ein Beispiel einer Funktion $f : [0, \infty) \rightarrow \mathbb{R}$, die ausschließlich positive Werte annimmt und für die trotzdem

$$\int_0^\infty f d\lambda < \infty$$

gilt.

Der Wert des Lebesgueschen Integrals über unbeschränkten Integranden und/oder Integrationsbereichen stimmt in diesen Fällen mit dem entsprechenden Grenzwert Riemannscher Integrale überein, wenn die Intervallgrenzen sich $\pm\infty$ oder beispielsweise einer Polstelle annähern. Man spricht in diesem Zusammenhang auch von **uneigentlichen (Riemann)-Integralen**.

Eine Besonderheit verdient Erklärung. Wir wollen sie am Beispiel der Funktion

$$f(x) := \frac{\sin x}{x} \quad \text{für } x \neq 0$$

illustrieren, die wir vermittels $f(0) := 1$ zu einer globalen reellen Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ stetig fortsetzen (Regel von de l'Hospital) und für $x \geq 0$ betrachten. Zwischen 0 und π sind die Funktionswerte positiv, zwischen π und 2π negativ, und so weiter alternierend. Wegen des wachsenden Nenners ist stets $|f(x + \pi)| < |f(x)|$. Folglich werden die Integrale $I_k := \int_{k\pi}^{(k+1)\pi} f(x) dx$, $k \in \mathbb{N}$, mit wachsendem k betragsmäßig immer kleiner, allerdings mit alternierendem Vorzeichen. Analog zu Leibnizreihen folgt daraus die Konvergenz von $\int_0^b f(x) dx$ für $b \rightarrow \infty$ gegen einen endlichen Wert. Das Lebesguesche Integral $\int_{[0, \infty)} f d\lambda$ ist jedoch NICHT definiert. So wie für die alternierende harmonische Reihe kann man sich nämlich leicht klar machen, dass Positivteil $\sum_{k \in \mathbb{N}} I_{2k}$ und Negativteil $-\sum_{k \in \mathbb{N}} I_{2k+1}$ des Integrals beide den Wert ∞ haben, das Integral als Differenz $\infty - \infty$ daher nicht sinnvoll definiert werden kann. Man kann dieses Phänomen auch so interpretieren: Das Lebesguesche Integral hängt abgesehen vom Integranden f nur vom Maß der Teilmengen des Integrationsbereichs ab, nicht aber von ihrer Anordnung auf der Zahlengerade. Bei der Bildung des Grenzwertes $b \rightarrow \infty$ hingegen spielt die Ordnung sehr wohl eine wichtige Rolle und bewirkt im Beispiel mit $f(x) = \frac{\sin x}{x}$, dass positive und negative Anteile gerade so gerecht abwechseln, dass sie einander im Wesentlichen die Waage halten. (Man vergleiche die Situation mit Leibnizreihen bzw. mit dem Riemannschen Umordnungssatz für absolut bzw. für bedingt konvergente Reihen.)

Übungsaufgabe 298. (T) Berechnen Sie das uneigentliche Integral

$$\int_0^\infty x e^{-x^2} dx.$$

Übungsaufgabe 299. (T) Weisen Sie nach, dass

$$\int_0^1 \frac{1}{(1-x)^\alpha} dx$$

genau für $\alpha < 1$ einen endlichen Wert besitzt.

Übungsaufgabe 300. (T) Benutzen Sie ein Vergleichsargument, um zu zeigen, dass die Integrale

$$\int_0^\infty e^{-x^2} dx \quad \text{und} \quad \int_1^\infty \frac{\sin(x)}{x^2} dx$$

endlich sind.

Übungsaufgabe 301. (P) Zeigen Sie, dass $\lim_{\substack{s \rightarrow 0^+ \\ t \rightarrow \infty}} \int_s^t \frac{\sin(x)}{x} dx$ einen endlichen Grenzwert besitzt. Anleitung: Formen Sie $\int_s^t \frac{\sin(x)}{x} dx$ mittels partieller Integration so um, dass ersichtlich wird, warum

$$\left| \int_s^t \frac{\sin(x)}{x} dx \right| \leq \frac{1}{s} + \frac{1}{t} + \int_s^t \frac{1}{x^2} dx = \frac{2}{s}.$$

5.4.2 Unendliche Reihen und Integrale

Inhalt in Kurzfassung: Die Definition des Riemannintegrals führt das Integral auf endliche Summen zurück. Der daraus resultierende Zusammenhang zwischen Summe und Integral ermöglicht auch den umgekehrten Weg, nämlich endliche oder unendliche Summen mit Hilfe von Integralen wenigstens näherungsweise zu berechnen. Das soll an einigen Beispiel illustriert werden.

Die soeben angeklungenen Vergleiche zwischen Integral und unendlichen Reihen lassen sich – wenig überraschend – auch auf eine strenge Grundlage stellen. Wieder wollen wir von den Funktionen $f_\alpha(x) := x^\alpha$ mit $\alpha < 0$ ausgehen, die Integrale $\int_a^b f_\alpha(x) dx$ mit den Reihen $\sum_{n=a+1}^b f_\alpha(n)$ vergleichen und den Grenzwert $b \rightarrow \infty$ untersuchen. Die Summe $\sum_{n=a+1}^b f_\alpha(n)$ lässt sich auch als Integral $\int_a^b g_\alpha(x) dx$ interpretieren, wobei g_α auf den Intervallen $[n, n+1)$ konstant ist mit Wert $f_\alpha(n+1)$. Wegen der fallenden Monotonie von f_α ist $g_\alpha \leq f_\alpha$, folglich $\sum_{n=a+1}^b f_\alpha(n) \leq \int_a^b f_\alpha(x) dx$ und im Grenzwert

$$\sum_{n=a+1}^{\infty} f_\alpha(n) \leq \int_a^{\infty} f_\alpha(x) dx.$$

Der allgemeine Sachverhalt lässt sich als **Vergleichskriterium** formulieren:

Satz 5.4.2.1. Sei f eine auf $[N, \infty)$, $N \in \mathbb{Z}$, definierte nichtnegative und monoton fallende Funktion, dann gilt:

$$\sum_{n=N+1}^{\infty} f(n) \leq \int_N^{\infty} f(x) dx \leq \sum_{n=N}^{\infty} f(n)$$

Insbesondere konvergiert die Reihe mit den Gliedern $f(n)$ genau dann, wenn $\int_N^{\infty} f(x) dx$ einen endlichen Wert hat.

Anwendungsbeispiel: Im vorangegangenen Abschnitt haben wir gesehen, dass für $\alpha < -1$ das Integral für $b \rightarrow \infty$ gegen einen endlichen Wert konvergiert. Also folgt die schon im Kapitel über Reihen behauptete Konvergenz der unendlichen Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} = 1 + \sum_{n=2}^{\infty} \frac{1}{n^\alpha} \leq 1 + \int_1^{\infty} \frac{dx}{x^\alpha} = 1 + \frac{1}{\alpha-1} < \infty \quad \text{für } \alpha > 1.$$

Indem man die untere Summationsgrenze von $a + 1$ auf a verändert, kann man auch eine Ungleichung in die umgekehrte Richtung bekommen:

$$\int_a^\infty f_\alpha(x) dx \leq \sum_{n=a}^\infty f_\alpha(n)$$

Im divergenten Fall (etwa für $-1 \leq \alpha$) kann man auf diese Weise auch die Divergenz von Reihen zeigen, wenn man diese Information für das entsprechende Integral bereits hat.

Übungsaufgabe 302. (T) Argumentieren Sie wie oben, um zu begründen, dass die Reihe $\sum_{n=1}^\infty n^\alpha$ genau für $\alpha < -1$ konvergiert.

Es liegt auf der Hand, wie in der beschriebenen Weise auch allgemeinere **Integralkriterien** für die Konvergenz und Divergenz von Reihen erhalten werden können. Wir verzichten hier auf Details.

Übungsaufgabe 303. (T) Entscheiden Sie, ob die unendliche Reihe

$$\sum_{n=2}^\infty \frac{1}{n \ln(n)}$$

konvergiert oder divergiert.

Außerdem sieht man, dass man im Fall monotoner Integranden unendliche Reihen durch Integrale approximieren kann. Verfeinerungen unserer Überlegungen führen zu Formeln, die es eventuell ermöglichen, den Approximationsfehler abzuschätzen. Als wichtigstes Beispiel dieser Art sei die **Eulersche Summenformel** erwähnt, wo Information über die Ableitung des Integranden entscheidend ist. Es gibt Verfeinerungen, wo auch höhere Ableitungen einfließen.

Übungsaufgabe 304. (P) Berechnen Sie für $N = 10, 100, 1000, \dots, 10^k$ die endliche Summe $s_N := \sum_{n=1}^N \frac{1}{n}$ unter Verwendung geeigneter Integrale und ohne Computersoftware näherungsweise. Geben Sie möglichst gute Fehlerabschätzungen an.

Übungsaufgabe 305. (E) In 1.4.2 wurde die Stirlingsche Näherungsformel für $n!$ angegeben. Leiten Sie eine einfachere Variante davon her, indem Sie

$$\ln n! = \sum_{k=1}^n \ln k$$

verwenden und diese Summe mit ähnlichen Methoden approximieren wie in Übungsaufgabe 304.

Ein Zwischenresümee: Sowohl beim Riemannschen wie auch beim Lebesgueschen Zugang wurde das Integral mittels (zunächst endlicher) Summen definiert, und zwar im Wesentlichen als ihr Grenzwert. Über den Hauptsatz der

Differential- und Integralrechnung als wichtigstem Instrument haben wir Integrale in einer Weise zu berechnen gelernt, die es ermöglicht, auch wieder Rückschlüsse auf Summen bzw. deren Grenzwert, also auf unendliche Reihen zu machen. Nun gehen wir nochmals wie zu Beginn den umgekehrten Weg, nämlich von der Summe zum Integral.

5.4.3 Numerische Integration

Inhalt in Kurzfassung: Integration mit Hilfe des Hauptsatzes der Differential- und Integralrechnung stößt u.a. dann auf Grenzen, wenn die Stammfunktionen eines gegebenen Integranden nicht explizit dargestellt werden können. Der Ausweg besteht meist in sogenannten numerischen Methoden. Ohne Hauptsatz muss man dann meist in der einen oder anderen Weise auf Riemannsummen oder ähnliche Approximationen des exakten Wertes des zu berechnenden Integrals zurückgreifen. Typischerweise muss man dabei höhere Genauigkeit durch größeren Rechenaufwand erkaufen und hat dabei Kosten und Nutzen gegeneinander abzuwägen.

Die Integrationsmethoden aus Abschnitt 5.2 fußen auf dem Hauptsatz der Differential- und Integralrechnung und der Hoffnung, eine Stammfunktion eines gegebenen Integranden f zu finden. Ist f z.B. stetig, so ist die Existenz einer Stammfunktion theoretisch gesichert. Für die praktische Berechnung hätte man freilich gerne einfache Formelausdrücke. Solche zu bekommen war das Ziel von Integrationsverfahren wie partieller Integration oder Substitutionsregel. Wie schon an früherer Stelle erwähnt, führen diese Methoden leider nicht immer zum Ziel. Es gibt sogar sehr einfach gebaute Funktionen, von denen man zeigen kann, dass sich ihre Stammfunktion (die laut Theorie existiert) überhaupt nicht als einfache Formel schreiben lässt, die nur aus den elementaren Funktionen, die wir schon kennen gelernt haben (gebrochen rationale, exp, Logarithmus, trigonometrische etc.) aufgebaut ist. Prominente Beispiele sind die Funktionen $x \mapsto \frac{\sin x}{x}$ (die weiter oben schon aufgetaucht ist) oder $x \mapsto e^{-x^2}$ (die in der Wahrscheinlichkeitstheorie im Zusammenhang mit der Normalverteilung eine extrem wichtige Rolle spielt). Je komplizierter Integranden werden, desto eher ist damit zu rechnen, dass man an der Ermittlung einer expliziten Formel für die Stammfunktion scheitert. Man kann sogar so weit gehen, zu sagen, dass jene Funktionen, die elementar darstellbare Stammfunktionen haben, die seltene Ausnahme bilden. In der Praxis ist diese Sicht heutzutage angemessener denn je, weil der Computer viel kompliziertere Funktionen behandelbar macht, als das früher, in der Zeit händischen Rechnens, möglich war. Deshalb verliert die Kunst des trickreichen Integrierens (im Sinne des Ermittlens einer Stammfunktion) im Vergleich zu anderen Methoden immer mehr an Gewicht. Welche Möglichkeiten eröffnet der Computer?

Vor allem lässt er viel aufwendigere numerische Integrationen zu. Dabei wird in konkreten Anwendungen auf exakte Darstellungen von Stammfunktionen verzichtet. Man begnügt sich mit numerischen (zahlenmäßigen) Näherungen bestimmter Integrale, für die man hinreichende Genauigkeit garantieren kann.

Beispielsweise kann man dabei die Definition des R-Integrals als Grenzwert von Riemannsummen nachvollziehen. Man geht von

$$\int_a^b f(x) dx \approx \sum_i c_i f(x_i)$$

aus und überlegt sich, wie die Belegung (x_1, \dots, x_n) sowie die Gewichte c_i (sie ersetzen die Differenzen $x_i - x_{i-1}$ in Riemannsummen) gewählt werden sollen, damit man für eine interessante Klasse von Funktionen f gute Näherungen bekommt. So eine Klasse ist meistens bestimmt durch Regularitätsbedingungen an f wie z.B. mehrmalige Differenzierbarkeit eventuell mit Schranken für gewisse höhere Ableitungen. Ähnlich wie im Satz von Taylor kann man dann beispielsweise erreichen, dass alle Polynome bis zu einem gewissen Grad n exakt integriert werden, und für alle anderen Funktionen der maximale Fehler durch die $n + 1$ -te Ableitung beschränkt werden kann. Eine besonders wichtige Rolle spielen numerische Methoden bei höherdimensionalen Integralen. Weil die dort eingesetzten Methoden zu weit führten, begnügen wir uns hier mit einem einfachen, eindimensionalen Beispiel, das in eine etwas andere Richtung zielt.

Man kann zum Beispiel hoffen, dass f zwischen je zwei benachbarten Zerlegungspunkten $a = x_0 < x_1 < \dots < x_n = b$ annähernd linear verläuft. Kennt man die Werte $f(x_i)$, und nimmt man **Äquidistanz** an, d.h. $x_i - x_{i-1} = \frac{b-a}{n}$, so nähert man $\int_a^b f(x) dx$ durch das Integral über diese stückweise lineare Funktion l an und erhält die sogenannte **Sehnenformel** oder auch **Trapezregel**:

$$\int_a^b f(x) dx \approx \int_a^b l(x) dx = (b-a) \left(\frac{f(a) + f(b)}{2n} + \frac{1}{n} \sum_{i=1}^{n-1} f(x_i) \right).$$

Es liegt nahe, diesen Ansatz zu verfeinern, indem man nicht durch stückweise lineare, sondern stückweise quadratische Polynome oder auch Polynome höheren Grades (sogenannte **Splinefunktionen**) interpoliert, deren Integrale man leicht berechnen kann. Kurzes Nachdenken über die Möglichkeiten zeigt schnell das Grundproblem, das in der numerischen Mathematik fast immer im Zentrum steht: Höhere Genauigkeit braucht (bei gleicher Methode) höheren Aufwand. Der Gewinn an Genauigkeit muss also immer gegen die Kosten des höheren Aufwandes abgewogen werden. Es gibt keine pauschale Regel, wo der optimale Kompromiss liegt, weil das von der konkreten Anwendungssituation abhängt. Damit eine Methode vielseitig einsetzbar ist, sollte sie mit möglichst scharfen Aussagen über die Qualitätsparameter verbunden sein. Solche herzuleiten ist das tägliche Brot numerischer Mathematiker. Anwender sollten die entsprechenden mathematischen Aussagen verstehen und ihren Wert je nach Anwendung richtig einschätzen können.

Übungsaufgabe 306. (P) *Vergleichen und diskutieren Sie am Beispiel des bestimmten Integrals $\int_{-1}^1 f(x) dx$ mit $f(x) := \frac{1}{1+x^2}$ verschiedene der oben ange-deuteten Methoden:*

1. *Geben Sie den exakten Wert des Integrals auf zwei Nachkommastellen an.*

2. Berechnen Sie Ihnen geeignet erscheinende Ober-, Unter- und Riemannsummen zu äquidistanten Zerlegungen des Integrationsintervalls in $n = 1, 2, 3, 4, 5, 6$ Teile.
3. Wenden Sie auf dieselben Zerlegungen die Trapezregel an.
4. Ersetzen Sie $f(x)$ durch das quadratische Taylorpolynom t_2 von f mit Entwicklungsstelle $x_0 = 0$ und berechnen Sie das Integral $\int_{-1}^1 t_2(x) dx$.
5. Ersetzen Sie $f(x)$ durch das quadratische Interpolationspolynom für f an den Stützstellen $-\pi, 0, \pi$.
6. Genauso, nur mit variierenden Stützstellen $-a, 0, a$ ($0 < a < 1$).

Übungsaufgabe 307. (E) Der Graph der Funktion $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2}$ wird auch als Gauß'sche Glockenkurve bezeichnet. Berechnen Sie näherungsweise die Fläche unter der Gauß'schen Glockenkurve indem Sie folgende Schritte ausführen:

1. Geben Sie sich eine Genauigkeit ε vor, z.B. $\varepsilon = 0,01$ und suchen Sie einen Wert $M > 0$ sodass $f(x) < \varepsilon$ für $x \notin [-M, M]$.
2. Überlegen Sie, wann die Ungleichung $e^{-x^2} < e^{-|x|}$ gilt und schätzen Sie damit die Fläche der Glockenkurve außerhalb des Intervalls $[-M, M]$ ab.
3. Ersetzen Sie auf dem Intervall $[-M, M]$ die Funktion $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2}$ durch ihr Taylorpolynom $T_{f,n,0}$. Bestimmen Sie den Grad n so groß, dass $T_{f,n,0}$ auf dem Intervall $[-M, M]$ nicht negativ wird und führen Sie eine Fehlerabschätzung $|f - T_{f,n,0}|$ auf $[-M, M]$ durch!
4. Berechnen Sie mit den konkreten Werten M und n , die Sie erhalten haben, das Integral

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2} dx \approx \int_{-M}^M T_{f,n,0}(x) dx$$

und geben Sie eine Abschätzung des dabei gemachten Fehlers an.

5.4.4 Kurven und ihre Länge

Inhalt in Kurzfassung: Will man den Begriff einer Kurve exaktifizieren, kommt man nicht umhin, Parametrisierungen zu verwenden. Begrifflich setzt man eine Kurve sogar mit ihrer Parametrisierung gleich. Dabei handelt es sich um eine Abbildung von einem Intervall (Parameterintervall) in die Ebene (oder in einen höherdimensionalen Raum \mathbb{R}^n). Von dieser Abbildung setzt man Stetigkeit voraus, d.h. Stetigkeit der Komponentenfunktionen, die ja reelle Funktionen sind. Liegt eine Parametrisierung vor, kann man mit Hilfe von Zerlegungen des Parameterintervalls die Längen der approximierenden Polygonzüge verwenden, um das Supremum all dieser Längen als Länge der Kurve (Bogenlänge) zu definieren. Ist die Parametrisierung stetig differenzierbar, so erhält man eine Darstellung der Bogenlänge als Integral. Angewendet auf die durch $t \mapsto (\cos t, \sin t)$,

$t \in [0, 2\pi]$, parametrisierte Kreislinie erhält man damit die Formel $u = 2r\pi$ für den Umfang u eines Kreises mit Radius r . Das ist jetzt keine elementargeometrische Definition der Zahl π mehr, sondern ein beweisbarer Satz. Die darin auftretende Zahl π ist dabei als über die Cosinusreihe (nämlich als das Doppelte der kleinsten positiven Nullstelle) definiert zu verstehen.

Wir wollen bei der oft exemplarisch zitierten Frage nach der Länge der Küste Großbritanniens beginnen. Wir stellen uns vor, die Küste entlang zu gehen und in gewissen Abständen Markierungen zu machen, deren geradlinigen Abstände voneinander wir addieren. Die Summe betrachten wir als Approximation der Gesamtlänge der Küste. Da diese sehr zerklüftet ist, misst man umso größere Werte, je kleinere Buchten man berücksichtigt. A priori ist es denkbar, dass die erhaltenen Werte unbeschränkt größer werden.

Ein mathematisch streng definierbares Objekt dieser Art ist die sogenannte **Kochsche** oder auch **Schneeflockenkurve**. Sie kommt wie folgt zustande: Man beginnt mit einem Geradensegment s_0 der Länge $L(s_0) = 1$ in der Ebene. Im ersten Schritt teilen wir s_0 in drei gleiche Teile. Über dem mittleren Drittel d als Basis errichten wir ein gleichseitiges Dreieck und ersetzen d durch die beiden anderen Seiten dieses Dreiecks. Das Resultat s_1 ist ein Polygonzug aus vier Strecken, jeweils der Länge $\frac{1}{3}$. Die Gesamtlänge von s_1 ist also gegeben durch $L(s_1) = \frac{4}{3}$. Im zweiten Schritt verfahren wir mit jeder der vier Strecken so wie im ersten Schritt mit s_0 . Auf diese Weise entsteht ein Polygonzug s_2 , bestehend aus 16 Geradenstücken, mit Gesamtlänge $L(s_2) = \frac{16}{9} = (\frac{4}{3})^2$. So kann man fortfahren und eine Folge von immer feiner gegliederten Polygonzügen s_n der Längen $L(s_n) = (\frac{4}{3})^n$ generieren. Die Grenzkurve $s = \lim_{n \rightarrow \infty} s_n$ (dieser hier etwas unpräzise Begriff lässt sich auf natürliche Weise, die der Anschauung sehr gut entspricht, exaktifizieren) ist die besagte Kochsche Kurve. Weil die Längen $L(s_n) = (\frac{4}{3})^n$ der approximierenden Kurven (Polygonzüge) s_n für $n \rightarrow \infty$ gegen ∞ streben, liegt es nahe, auch s die Länge $L(s) = \infty$ zuzuordnen.

Übungsaufgabe 308. (E) Skizzieren Sie die Kochsche Kurve und illustrieren Sie an diesem Beispiel die Problematik unendlicher Kurvenlänge.

Man kann sogar Kurven definieren, die zum Beispiel das ganze Einheitsquadrat ausfüllen, sogenannte **Peanokurven**.³

Wir wollen realistische Voraussetzungen an Kurven finden, wo solche Phänomene nicht auftreten können. Wir hoffen, dass dann sehr wohl obere Schranken für die Messwerte auch bei beliebiger Verfeinerung der Approximation existieren. Dann wäre das Supremum aller Messwerte ein sinnvoller Kandidat, um die Länge der Kurve zu definieren. Wir wollen diesen Ansatz präzisieren.

Zunächst definieren wir eine Kurve in der Ebene \mathbb{R}^2 (analog im drei- oder auch höherdimensionalen Raum \mathbb{R}^n) als eine Funktion $f : [a, b] \rightarrow \mathbb{R}^2$. Den Definitionsbereich $[a, b]$ können wir uns beispielsweise als ein Zeitintervall vorstellen,

³Sie sich nach dem italienischen Mathematiker Giuseppe Peano (1858-1932) benannt. Nach ihm sind auch die berühmten Peanoaxiome benannt, eines der ersten Axiomensystem für die natürlichen Zahlen.

innerhalb dessen jedem Zeitpunkt t durch f ein Punkt $f(t) = (x(t), y(t))$ zugeordnet wird. Die beiden reellen (eindimensionalen) Funktionen $x, y : [a, b] \rightarrow \mathbb{R}$ setzen wir als stetig voraus, wenn wir von einer Kurve sprechen. Sie beschreiben die Koordinaten der Punkte in der Ebene, die während des Zeitintervalls $[a, b]$ mittels $f : t \mapsto (x(t), y(t))$ durchlaufen werden. Die Menge all dieser Punkte $(x(t), y(t))$, $a \leq t \leq b$, also das Bild von $[a, b]$ unter f , entspricht unserer intuitiven Vorstellung einer Kurve. Begrifflich jedoch ist diese Vorstellung ungenau, weil dabei die Abhängigkeit von t , die sogenannte **Parametrisierung** der Kurve unterschlagen wird. Erst durch die Parametrisierung werden, wie wir gleich sehen werden, Differential- und Integralrechnung verfügbar.

Um die Länge der Kurve begrifflich zu fassen, betrachten wir – wie schon bei der Definition des R-Integrals – Zerlegungen $Z = \{a = t_0 \leq t_1 \leq \dots \leq t_n = b\}$ von $[a, b]$. Die Gesamtlänge $L(f)$ der Kurve muss mindestens so lang sein wie der Polygonzug, der $f(t_0)$ mit $f(t_1)$ geradlinig verbindet, $f(t_1)$ mit $f(t_2)$ etc. Die Länge $L(f, Z)$ des zu Z gehörigen Polynomzugs ist nach Pythagoras (euklidischer Abstand) gegeben durch

$$L(f, Z) := \sum_{i=1}^n \sqrt{(x(t_i) - x(t_{i-1}))^2 + (y(t_i) - y(t_{i-1}))^2}.$$

Für jedes Z verlangen wir $L(f) \geq L(f, Z)$ und definieren deshalb die **Bogenlänge** $L(f)$ von f als

$$L(f) := \sup_Z L(f, Z),$$

wobei Z alle Zerlegungen des Intervalls $[a, b]$ durchläuft.

Die Beziehung zur Integralrechnung lässt sich herstellen, indem wir f , d.h. die reellen Funktionen $x : t \mapsto x(t)$ und $y : t \mapsto y(t)$ als differenzierbar voraussetzen und auf jedem der Teilintervalle $[t_{i-1}, t_i]$ den Mittelwertsatz bemühen. Er besagt, dass es Zwischenstellen $\alpha_i, \beta_i \in (t_{i-1}, t_i)$ gibt mit $x(t_i) - x(t_{i-1}) = x'(\alpha_i)(t_i - t_{i-1})$ und $y(t_i) - y(t_{i-1}) = y'(\beta_i)(t_i - t_{i-1})$. Der resultierende Wert

$$\begin{aligned} L(f, Z) &= \sum_{i=1}^n \sqrt{(x(t_i) - x(t_{i-1}))^2 + (y(t_i) - y(t_{i-1}))^2} = \\ &= \sum_{i=1}^n (t_i - t_{i-1}) \sqrt{x'(\alpha_i)^2 + y'(\beta_i)^2} \end{aligned}$$

erinnert an eine Riemannsumme für die Funktion $g(t) := \sqrt{x'(t)^2 + y'(t)^2}$. Wäre $\alpha_i = \beta_i$, so hätten wir es tatsächlich mit einer solchen zu tun, und zwar zur Belegung $B = (\alpha_1, \alpha_2, \dots, \alpha_n)$ zur Zerlegung Z . Dann könnten wir direkt argumentieren: Sind die Ableitungen x' und y' stetig, so auch g . Also ist g R-integrierbar, und die $L(f, Z)$ konvergieren für $\phi(Z) \rightarrow 0$ gegen das Integral

$$\int_a^b \sqrt{x'(t)^2 + y'(t)^2} dt.$$

Weil im Allgemeinen $\alpha_i \neq \beta_i$ gilt, ist diese Schlussweise so noch nicht zwingend. Mit etwas mehr technischem Aufwand, der keine grundsätzlich neuen Ideen braucht, lässt sich aber tatsächlich beweisen:

Satz 5.4.4.1. *Für die Funktion $f : [a, b] \rightarrow \mathbb{R}^2$, $t \mapsto (x(t), y(t))$ seien die Komponentenfunktionen $x : [a, b] \rightarrow \mathbb{R}$ und $y : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar (an den Randpunkten als einseitige Ableitung). Dann ist die Länge $L(f)$ der durch f gegebenen Kurve gegeben durch*

$$L(f) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2} dt.$$

Ein analoger Satz gilt auch für \mathbb{R}^n mit $n > 2$ statt \mathbb{R}^2 , also für Kurven in höherdimensionalen Räumen.

Als Spezialfall behandeln wir nun die Kreislinie mit Radius r . Hier ist $f : [0, 2\pi] \rightarrow \mathbb{R}^2$ mit $t \mapsto (x(t), y(t))$, $x(t) := r \cos t$ und $y(t) := r \sin t$, also $x'(t) = -r \sin t$, $y'(t) = r \cos t$ und $\sqrt{x'(t)^2 + y'(t)^2} = \sqrt{r^2 \sin^2 t + r^2 \cos^2 t} = r$. Unsere Formel für die Bogenlänge liefert daher die bekannte Formel $u = \int_0^{2\pi} \sqrt{x'(t)^2 + y'(t)^2} = \int_0^{2\pi} r dt = 2r\pi$ für den Umfang u eines Kreises mit Radius r .

Der Vektor $f'(t) := (x'(t), y'(t))$ lässt sich als Geschwindigkeitsvektor deuten. Denn er gibt die Richtung und vermittelt seiner Länge auch die Geschwindigkeit der Bewegung an. Bei der Kreisbewegung erhalten wir als (euklidische) Länge

$$\|f'(t)\| = \sqrt{x'(t)^2 + y'(t)^2} = r.$$

Auf dem Einheitskreis in der komplexen Zahlenebene wird durch $t \mapsto e^{it} = \cos t + i \sin t$ daher eine Bewegung mit Geschwindigkeit 1 beschrieben, so wie wir das in 4.3.3 bereits diskutiert haben. Erst jetzt jedoch haben wir einen exakten Begriff der Bogenlänge zur Verfügung. Die einfache Formel zur Berechnung der Bogenlänge von Kurven mit stetig differenzierbaren Parametrisierungen verdanken wir der Differential- und Integralrechnung.

Abschließend sei noch der Fall eines Funktionsgraphen erwähnt, wo $x(t) = t$, also $x'(t) = 1$ und somit

$$L(f) = \int_a^b \sqrt{1 + y'(t)^2} dt.$$

gilt.

Übungsaufgabe 309. (T) Berechnen Sie die Länge einer Parabel, die durch den Funktionsgraphen von $f(x) = x^2$ gegeben ist, im Bereich $0 \leq x \leq a$. Hinweis: Im auftretenden Integral bietet sich die Substitution $x = \sinh t$ an.

Übungsaufgabe 310. (T) Berechnen Sie die Bogenlänge des Funktionsgraphen der Funktion

$$f(x) = \operatorname{Arcosh} \frac{1}{x} - \sqrt{1 - x^2}, \quad \text{für } x \in \left[\frac{1}{e}, 1\right].$$

Übungsaufgabe 311. (T) Berechnen Sie die Bogenlänge der Kurve mit

$$x(t) = t \cos(t), y(t) = t \sin(t), \quad \text{für } t \in [0, \pi].$$

Übungsaufgabe 312. (P) Welche Integrale treten bei der Berechnung der Bogenlänge einer Ellipse auf, welche bei einer Hyperbel? (Sie müssen diese Integrale nicht ausrechnen.)

Index

- Abbildung, 29
abelsche Gruppe, 36
Abelscher Grenzwertsatz, 194
abgeschlossen, 86
abgeschlossenes Intervall, 35
Ableitung, 169, 179
Ableitung der Umkehrfunktion, 177
Ableitungsregeln, 172
Abschluss, 85
absolut konvergente Reihe, 114
Absolutbetrag, 41
abzählbar, 32
Addition, 33
Additionstheoreme, 58, 74, 198
allgemeine Potenz, 159, 195
allgemeingültig, 17
Allquantor, 17, 26
alternierend, 91
alternierende Reihe, 122
angeordneter Körper, 38
Anschlussstelle, 185, 188
Antisymmetrie, 37
Anzahlformeln, 62
Approximationssatz von Weierstraß, 151
Äquidistanz, 252
Äquivalenz, 17, 26
Äquivalenzumformung, 18
archimedische Eigenschaft, 40
Arcus cosinus, 165, 199
Arcus cotangens, 165, 199
Arcus sinus, 165, 199
Arcus tangens, 165, 199
Arcusfunktionen, 165, 199
Arcustangensreihe, 201
Area cosinus hyperbolicus, 202
Area cotangens hyperbolicus, 202
Area sinus hyperbolicus, 202
Area tangens hyperbolicus, 202
Areafunktionen, 201
Argument einer komplexen Zahl, 58
arithmetische Folge, 90
arithmetische Reihe, 46
Assoziativgesetz, 36
Asymptote, 127
Aussage, 13
Aussageform, 15
äußere Ableitung, 175
äußerer Punkt, 85
äußeres Produkt, 79
Axiom, 10
Axiomatik, 36

Banachscher Fixpunktsatz, 138
Basis, 52, 72
bedingt konvergente Reihe, 114
Belegung, 218
Bernoullische Ungleichung, 90
beschränkt, 90
beschränkt, 40, 86
beschränkte Funktion, 126
bestimmtes Integral, 217, 227
Betrag, 33, 41, 57
Bewegung, 238
Bewegungsinvarianz, 238
Beweis, 11
Bijektion, 29
bijektiv, 29
Bild, 29
Bildmenge, 29
bilinear, 75
Binomialkoeffizient, 66, 196
Binomische Formel, 67
binomische Reihe, 195, 196
Binomischer Lehrsatz, 67

- binär, 53
- Bogenlänge, 253, 255
- Bogenmaß, 163
- Bruch, 33

- Cantormenge, 236
- Cauchy-Schwarzsche Ungleichung, 77
- Cauchyfolge, 94
- Cauchyprodukt, 121
- Chaos, 211
- charakteristische Funktion, 133, 242
- Cosinus, 163, 198
- Cosinus hyperbolicus, 201
- Cotangens, 164, 199
- Cotangens hyperbolicus, 202

- Definition, 10
- Definitionsbereich, 29
- Definitionsmenge, 29
- dekadisch, 53
- Dezimaldarstellung, 53
- Differential, 217
- Differentialquotient, 169
- Differentialrechnung, 167
- Differenz, 26, 33
- Differenzenquotient, 168
- differenzierbar, 169
- Dirichletsche Sprungfunktion, 133, 217, 244, 245
- disjunkt, 26
- Disjunktion, 16, 26
- Distributivgesetz, 37
- divergent, 93
- divergente Reihe, 113
- Divergenz, 93
- Divergenz gegen ∞ , 105
- Division, 33
- Division mit Rest, 144
- Dominoeffekt, 45
- Dreiecksungleichung, 41, 78
- Durchschnitt, 25, 26
- Dynamik, 211

- Ebene, 83
- Eindeutigkeit der Exponentialfunktion, 158
- Eindeutigkeit der Exponentialfunktionen, 161
- Eindeutigkeit der linearen Funktionen, 161
- Eindeutigkeit der Logarithmusfunktionen, 161
- Eindeutigkeit der Potenzfunktionen, 161
- Eindeutigkeit der Primfaktorzerlegung, 50
- Einheitskreis, 163
- Einheitswurzel, 59
- Einschränkung, 30
- einseitige Ableitung, 171
- einseitiger Grenzwert, 130
- einstelliger Junktor, 16
- endliche Additivität, 235
- Entwicklungsstelle, 185, 188
- ε - δ -Kriterium, 130
- euklidische Länge, 76
- euklidische Metrik, 78
- euklidischer Abstand, 78
- Eulersche Summenformel, 250
- Eulersche Zahl, 100, 118, 159
- Existenz der Exponentialfunktion, 158
- Existenz der Exponentialfunktionen, 161
- Existenz der linearen Funktionen, 161
- Existenz der Logarithmusfunktionen, 161
- Existenz der Potenzfunktionen, 161
- Existenz und Eindeutigkeit der Exponentialfunktion, 158
- Existenz und Eindeutigkeit der Exponentialfunktionen, 161
- Existenz und Eindeutigkeit der linearen Funktionen, 161
- Existenz und Eindeutigkeit der Logarithmusfunktionen, 161
- Existenz und Eindeutigkeit der Potenzfunktionen, 161
- Existenzquantor, 17, 27
- Exponent, 42
- Exponentialfolge, 90
- Exponentialfunktion, 157
- Exponentialreihe, 118, 186
- Extrema, 179
- Extremum, 127
- Extremwertbestimmung, 208

- Faktorielle, 64
- Fakultät, 64
- fast alle, 93
- Feinheit, 216
- Fibonacci-Folge, 46
- Fixpunktiteration, 137
- Fixpunktprinzip, 138
- Fläche, 237
- Flächenberechnung, 215
- Folge, 45
- Folglied, 46
- Folgenstetigkeit, 131
- Folgerung, 11
- Formel für die geometrische Reihe, 116
- Fortsetzung, 30
- freie Variable, 23
- Fundamentalsatz der Algebra, 146, 147
- Fundamentalsatz der Zahlentheorie, 50
- Funktion, 29
- Funktionswert, 29

- ganze Zahlen, 33
- ganzrationale Funktion, 132, 142
- ganzzahliger Anteil, 53
- Gauß'sche Zahlenebene, 57
- gebrochen rationale Folge, 99
- gebrochen rationale Funktion, 133, 143, 230
- gebrochener Anteil, 53
- gebundene Variable, 23
- gekürzter Bruch, 53
- gekürzt, 33
- geometrische Folge, 90
- geometrische Reihe, 47, 115
- geordnetes Paar, 28
- Gerade, 81
- gerade Funktion, 126
- glatt, 182
- Gleichheit, 26
- gleichmäßig konvergent, 139
- gleichmäßig stetig, 223
- gleichmäßige Konvergenz, 139
- Gleichung, 14
- Glied einer Folge, 46
- Glied einer Reihe, 113
- gliedweises Differenzieren, 190
- global stetig, 130
- globales Extremum, 127
- globales Maximum, 127
- globales Minimum, 127
- Grad, 142
- Graph einer Funktion, 126
- Grenzfunktion, 139
- Grenzwert einer Folge, 92
- Grenzwert einer Funktion, 130
- Grenzwertsätze, 95
- große Vereinigung, 27
- großer Durchschnitt, 26
- Gruppe, 36
- größter gemeinsamer Teiler, 51, 148

- halboffenes Intervall, 35
- harmonische Reihe, 116
- Häufungspunkt einer Folge, 101, 102
- Häufungspunkt einer Menge, 102
- Hauptsatz der Differential- und Integralrechnung, 224
- hebbare Unstetigkeit, 133
- Hessesche Abstandsformel, 82
- Hilfssatz, 11
- Hintereinanderausführung, 30
- homogen, 72
- Homogenität der Ableitung, 173
- Hornerschema, 143
- Hyperbelfunktionen, 201
- hyperharmonische Reihe, 116
- höhere Ableitung, 182

- identische Funktion, 30
- Identität, 30
- imaginäre Achse, 57
- imaginäre Einheit, 34, 55
- Imaginärteil, 55
- Implikation, 16, 26
- Indikatorfunktion, 133
- Induktionsanfang, 43
- Induktionsannahme, 43
- Induktionsbehauptung, 43
- Induktionsprinzip, 43
- Induktionsschritt, 43
- Induktionsstart, 43
- Induktionsvoraussetzung, 43

- Infimum, 40
- Infimum einer Folge, 91
- Infimumseigenschaft, 40
- Infinitesimalrechnung, 168
- Ingenieurwissenschaft, 9
- inhomogen, 72
- injektiv, 29
- Inklusions-Exklusionsprinzip, 63
- innere Ableitung, 175
- innerer Punkt, 85
- Inneres, 85
- inneres Produkt, 75
- Integrabilitätskriterium, 243
- Integralkriterium, 250
- Integralrechnung, 215
- Integralungleichungen, 221
- Integrand, 217
- Integration, 228
- Integrationsbereich, 217
- Integrationsregeln, 228
- integrierbar, 243
- Intergration, 226
- Interpolationspolynom, 152
- Intervall, 35
- Intervallschachtelung, 213
- inverse Abbildung, 29
- inverse Funktion, 29
- inverse Relation, 29
- inverses Element, 36
- irrationale Zahl, 34
- isomorph, 39
- Iterationsfolge, 111

- Junktor, 16

- kanonische Basis, 71
- kanonischer Einheitsvektor, 71
- Kardinalität, 31
- kartesisches Produkt, 28
- Kegel, 84
- Kettenregel, 175
- klassische Logik, 13
- kleinstes gemeinsames Vielfaches, 52, 148
- Kochsche Kurve, 254
- Koeffizient, 142, 188
- Kombination, 65
- Kombinatorik, 62
- Kommutativgesetz, 15, 36
- kompakt, 86
- Komplement, 26
- komplex konjugiert, 55
- komplexe Zahlen, 34, 55
- Komponente, 28
- Komposition, 30
- kongruent, 59
- Konjugation, 25
- Konjugierte, 55
- Konjunktion, 16
- konkav, 205
- konstante Folge, 90
- konstantes Polynom, 142
- Kontraktion, 138
- Kontraktionskonstante, 138
- Kontraktionsprinzip, 138
- konvergente Folge, 93
- konvergente Reihe, 113
- Konvergenz, 93
- Konvergenzbereich, 188
- Konvergenzintervall, 189
- Konvergenzradius, 188
- Konvergenzregeln, 95
- konvexe Funktion, 205
- konvexe Menge, 205
- Konvexität und zweite Ableitung, 205
- Konvexkombination, 206
- Koordinate, 69
- Koordinatendarstellung, 55
- Koordinatenform, 81
- Korollar, 11
- Kreis, 82
- Kreiszahl π , 163
- Kreuzprodukt, 79
- Krümmung, 207
- kubisch, 143
- Kugel, 84
- Kugelkoordinaten, 84
- Körper, 37
- L-integrierbar, 243
- Lagrangesche Interpolationsformel, 151
- Länge, 237

- Lebesgue-integrierbar, 243
- Lebesgue-messbare Menge, 238
- Lebesgue-Nullmenge, 218
- Lebesguemaß, 238
- Lebesguesches Integral, 242
- Lebesguesches Maß, 238
- Leibnizreihe, 122
- Lemma, 11
- Limes inferior, 103, 109
- Limes superior, 103, 109
- lineare Abbildung, 72
- lineare Approximation, 169, 170
- lineares Polynom, 143
- Linearfaktor, 145
- Linearität des Differenzierens, 173
- Linearität des Integrals, 229
- Linearkombination, 72
- links halboffenes Intervall, 35
- linksseitige Ableitung, 171
- linksseitiger Grenzwert, 130
- logarithmische Reihe, 194
- Logarithmus, 159, 193
- logarithmus naturalis, 159
- lokal stetig, 130
- lokales Extremum, 127
- lokales Maximum, 127
- lokales Minimum, 127
- Lösungsmenge, 25

- Maclaurinreihe, 185
- Majorantenkriterium, 114
- Mathematik, 10
- Maximum, 40, 127
- Maß, 235
- mehrgliedrige Rekursion, 46
- Menge, 24
- mengentheoretische Differenz, 26
- mengentheoretische Operationen, 24
- messbare Funktion, 243
- messbare Menge, 235, 238
- Metrik, 78
- metrischer Raum, 78
- Minimum, 40, 127
- Minorantenkriterium, 114
- Mittelwertsatz, 180

- Mittelwertsatz der Differentialrechnung, 180
- Moivresche Formel, 59
- monoton, 91, 126
- monoton fallend, 91, 126
- monoton wachsend, 91, 126
- monotone Funktion, 126, 220
- Monotonie, 179
- Monotonie und Ableitung, 203
- Monotoniegesetze, 38
- Multinomischer Lehrsatz, 68
- Multiplikation, 30, 33
- Mächtigkeit, 31

- n -mal differenzierbar, 182
- n -mal stetig differenzierbar, 182
- nach oben beschränkt, 90
- nach unten beschränkt, 90
- Nachfolger, 43
- Nachfolgerinduktion, 44
- Naturwissenschaft, 10
- natürliche Zahlen, 33
- natürlicher Logarithmus, 193
- Negation, 16, 26
- negativ, 41
- Nenner, 33
- neutrales Element, 36
- Newtonsches Interpolationsverfahren, 152
- Newtonverfahren, 212
- Norm, 76
- normal, 75
- Normalvektorform, 82
- normiertes Polynom, 148
- n -tupel, 28
- Nullfolge, 93
- Nullmenge, 218, 236
- Nullpolynom, 143
- Nullstelle, 126
- Nullstellen von Polynomen, 149
- Nullvektor, 71
- numerische Integration, 251

- o.B.d.A., „ohne Beschränkung der Allgemeinheit“, 50
- obere Schranke, 40
- oberer Grenzwert, 103

- oberer Limes, 103
- oberes R-Integral, 217
- oberes Riemannintegral, 217
- Obermenge, 26
- Obersumme, 216
- offen, 86
- offene Kugel, 85
- offenes Intervall, 35
- Ordnungsinduktion, 44
- Ordnungsrelation, 37
- orthogonal, 75
- Oszillation, 133

- Paradoxon von Banach-Tarski, 240
- Parameterform, 81
- Parametrisierung, 255
- Partialbruchzerlegung, 231
- Partialsomme, 46, 113
- partielle Integration, 229
- Pascalsches Dreieck, 67
- Periode, 91, 126
- periodisch, 91, 126
- periodischer Punkt, 111
- Permutation, 64
- p -Norm, 246
- Polardarstellung, 57
- Polstelle, 126, 134
- Polynom, 132, 142
- Polynomfunktion, 132, 142
- Polynomischer Lehrsatz, 68
- positiv, 41
- Potenz, 42, 156, 159, 195
- Potenzen komplexer Zahlen, 59
- Potenzfunktion, 156
- Potenzmenge, 235
- Potenzreihe, 178, 188
- Primfaktor, 49
- Primfaktorzerlegung, 49
- Primzahl, 49
- Produkt, 33
- Produktregel, 173
- Proposition, 11
- Prädikat, 14
- Punkt, 69
- punktweise Grenzfunktion, 139
- punktweiser Limes, 139

- Pythagoräischer Lehrsatz, 34

- Quader, 83
- Quadrat, 41
- quadratisch, 143
- Quadratwurzel, 41
- Quadrupel, 28
- Quantor, 16
- Quotient, 33
- Quotientenkriterium, 118
- Quotientenregel, 176

- R-Integral, 217
- Randpunkt, 35, 85
- rationale Funktion, 142
- rationale Zahlen, 33
- Realteil, 55
- Rechteck, 82
- rechts halboffenes Intervall, 35
- rechtsseitige Ableitung, 171
- rechtsseitiger Grenzwert, 130
- reelle Achse, 57
- reelle Folge, 90
- reelle Funktion, 125
- reelle Zahlen, 34
- Reflexivität, 37
- Regel von de l'Hospital, 191
- Regularität eines Maßes, 238
- Reihe, 46, 112
- Rekursion, 45
- Rekursionssatz, 45
- rekursive Definition, 45
- rekursive Folge, 111
- Relation, 29
- Restglied, 184
- Riemannintegral, 215, 217
- Riemannscher Umordnungssatz, 119
- Riemannsumme, 218

- Satz, 11
- Satz des Pythagoras, 78
- Satz vom kleinsten Element, 44
- Satz vom Maximum, 137
- Satz von der dominierten Konvergenz, 246

- Satz von der Konvergenz durch Monotonie und Beschränktheit, 108
 Satz von der monotonen Konvergenz, 246
 Satz von der stetigen Umkehrfunktion, 136
 Satz von Pythagoras, 34
 Satz von Rolle, 180
 Satz von Taylor, 184
 Schmiegunskreis, 207
 Schneeflockenkurve, 254
 Schnitt, 25
 Schranke, 40
 Sehnenformel, 252
 σ -Additivität, 235
 σ -Algebra, 235
 Signumfunktion, 133
 Sinus, 163, 198
 Sinus hyperbolicus, 201
 Skalar, 70
 Skalarmultiplikation, 70
 Skalarprodukt, 75
 Spinnwebdiagramm, 111, 139
 Splinefunktion, 252
 Sprunghöhe, 133
 Sprungstelle, 133
 Stammfunktion, 169
 Standardvektorraum, 71
 stetig, 130
 stetig differenzierbar, 182
 streng alternierend, 91
 streng konkav, 205
 streng konvex, 205
 streng monoton, 91, 126
 streng monoton fallend, 91, 126
 streng monoton wachsend, 91, 126
 strikt alternierend, 91
 strikt monoton fallend, 91
 strikt monoton wachsend, 91
 Substitutionsregel, 230
 Subtraktion, 33
 Summand, 113
 Summe, 33
 Summenregel für die Ableitung, 173
 Supremum, 40
 Supremum einer Folge, 91
 Supremumseigenschaft, 40
 surjektiv, 29
 symmetrische Differenz, 26
 Tangens, 164, 199
 Tangens hyperbolicus, 202
 Tautologie, 19
 Taylorapproximation, 178
 Taylorpolynom, 184
 Taylorreihe, 185
 Teiler, 33, 49
 Teilfolge, 95, 106
 Teilmenge, 26
 Teilsumme, 113
 Theorem, 11
 Theorie, 11
 Transformation, 111
 Transitivität, 38
 Trapezregel, 252
 Treppenfunktion, 242
 Trichotomie, 38
 trigonometrische Darstellung, 57
 trigonometrische Funktionen, 163, 196
 Tripel, 28
 überabzählbar, 32
 Umfang des Kreises, 256
 Umgebung, 85
 Umkehrabbildung, 29
 Umkehrfunktion, 29, 177
 Umkehrrelation, 29
 unbestimmte Form, 105
 unbestimmtes Integral, 227
 uneigentliche Konvergenz, 105
 uneigentliches R-Integral, 247
 unendlich, 43
 unendlich oft differenzierbar, 182
 unendliche Reihe, 112
 ungerade Funktion, 126
 Unstetigkeit, 132
 untere Schranke, 40
 unterer Grenzwert, 103
 unterer Limes, 103
 unteres R-Integral, 217
 unteres Riemannintegral, 217

- Untersumme, 216
- Variable, 15
- Variationen mit Wiederholung, 64
- Vektor, 69
- Vektoraddition, 70
- Vektorprodukt, 79
- Vektorraumoperation, 70
- verallgemeinerter Binomialkoeffizient, 196
- Vereinigung, 26, 27
- Verfeinerung, 216
- Vergleichskriterium, 114, 249
- Verkettung, 30
- Verknüpfung, 30
- Vielfaches, 49
- Vielfachheit, 145
- Vollkugel, 84
- vollständiger metrischer Raum, 109
- vollständig angeordneter Körper, 36, 39
- vollständige Induktion, 44
- Vollständigkeit eines Maßes, 237
- Vollständigkeit von \mathbb{R} , 39, 105, 134
- Volumen, 237

- Wahrheitstafel, 16
- Wahrheitswert, 13
- Wert, 29
- Wert einer unendlichen Reihe, 113
- Wertemenge, 29
- Wertevorrat, 29
- Wurzel, 42, 156
- Wurzelfunktion, 156
- Wurzelkriterium, 117

- Zahlengerade, 35
- zerfallendes Polynom, 145
- Zerlegung, 216
- Zielbereich, 29
- Zielmenge, 29
- Ziffern, 53
- zusammenhängend, 87
- zweistelliger Junktor, 16
- zweiwertige Logik, 13
- Zwischenwertsatz, 135
- Zwischenwertsatz für Ableitungen, 181
- Zylinder, 84
- Zylinderkoordinaten, 84
- Zähler, 33
- Zählmaß, 236