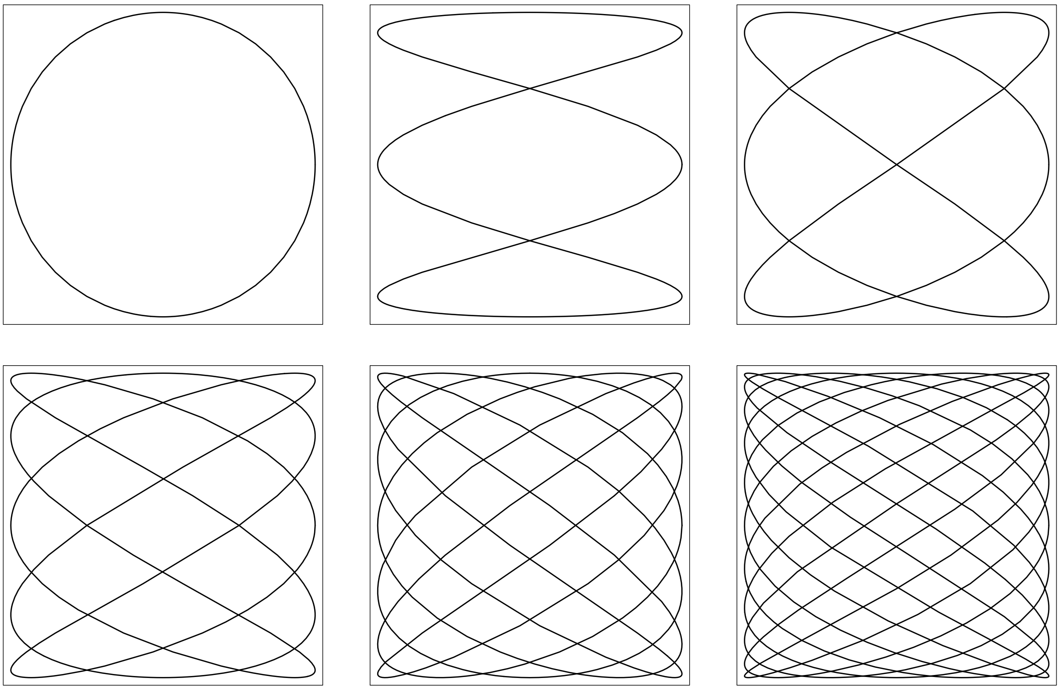
Analysis I

N.-P. Skoruppa

Vorlesungsskript — Universität Siegen, 2000



Version: Id

Die folgenden Seiten stellen eine nur leicht veränderte Version eines Skripts dar, welches ich 1990 frei nach Mitschriften zu Vorlesungen von F. Hirzebruch verfasste, die er in in Bonn in den Wintersemestern 1978/79 und 1989/90 gehalten hatte. Zu der zuletzt genannten Veranstaltung stellten mir verschiedene Studenten gelegentlich ihre jeweiligen Notizen zur Verfügung. Zu der ersten lag mir eine sehr schöne, noch handschriftlich angefertigte Vorlesungsausarbeitung von A. Küster vor, die dieser "mit spitzer Feder" niedergeschrieben hatte, wie F. Hirzebruch es nannte. Diese "spitze Feder", die punktgenau witzige Anmerkungen aus F. Hirzebruch's Vortrag aufzeichnete und zuweilen noch extrapolierte, habe ich hier erst gar nicht nachzuahmen versucht — obgleich ansonsten viele Details aus beiden Vorlesungen wortwörtlich übernommen wurden —, dagegen ist aber der methodische Aufbau der ursprünglichen Vorlesungen fast unverändert. Das ursprüngliche Skript von 1990 war unter dem Titel "Infinitesimalrechnung I" an der Universität Bonn verfügbar.

N.-P. Skoruppa, Siegen im September 2000

Das vorliegende Skript wurde auf PCs des Max-Planck-Instituts für Mathematik mit TEX gesetzt. Numerische Beispiele mit großen Zahlen waren dank PARI in Sekundenbruchteilen möglich, und die Graphen von Funktionen sind von MATHEMATICA gezeichnet worden. Die raffinierteren Aspekte des Layout wären ohne häufige Rücksprachen mit Ulrich Everling nicht zustande gekommen, und Rainer Jung verdanke ich einen wertvollen Hinweis, wie man von MATHEMATICA erstellte Graphen "glatter" reproduziern kann. Karin Deutler nahm die Mühe auf sich, ihren PC eine erste TEX -Rohform der Kapitel 1 bis 9 zu lehren. Das zwölfte Kapitel wurde von Hans-Olaf Herøy geschrieben (und in TEX gesetzt). Pia Bauer-Price, Rainer Jung, Hartmut Maennel brachten die Rohform der Kapitel 5,6,7 in eine TEX-lesbare Form. Verschiedene Mathematiker im Max-Planck-Institut für Mathematik und die Übungsleiter der Übungen zu den Vorlesungen halfen beim Korrektur lesen. Schließlich mußte das vorliegende Skript in Riesenauflage gedruckt und gebunden werden. Allen genannten und ungenannten Beteiligten möchte ich an dieser Stelle aufrichtig danken.

N-P. Skoruppa, Bonn, Februar 1990

INHALT

Im sechsten Kapitel wird die Parametrisierung $x \mapsto (\cos 2\pi x, \sin 2\pi x)$ des Einheitskreises betrachtet. Allgemeiner kann man Abbildungen der Gestalt $x \mapsto (\cos \frac{2\pi}{p} x, \sin \frac{2\pi}{q} x)$ studieren, wobei p und q irgendwelche positiven ganzen Zahlen bedeuten. Die dabei entstehenden Kurven heißen Lissajou Figuren. Das Titelblatt zeigt die Lissajou Figuren für $(p, q) = (1, 1)$, $(1, 3)$, $(2, 3)$, $(3, 5)$, $(5, 7)$, $(7, 11)$.

1 EINIGE HISTORISCHE BEMERKUNGEN ÜBER ZAHLSYSTEME

In diesem einführenden Kapitel werden die Zahlsysteme vorgestellt, denen man in der Infinitesimalrechnung begegnet. Insbesondere werden wir die für sie in der Mathematik heute allgemein gebräuchlichen Symbole vorstellen, wir werden andeuten, wie die einzelnen Zahlsysteme historisch auseinander hervorgegangen sind und wie ihre Entwicklung notwendig zur Infintesimalrechnung führen mußte. Nebenbei werden wir gleichzeitig schon einige der Schreibweisen kennenlernen, die man in der Mathematik heutzutage ständig benutzt.

Natürliche Zahlen

Mit $\mathbb{N}$ bezeichnet man die Menge der *natürlichen Zahlen*. Dies kann man symbolisch auch durch die Schreibweise

$$\mathbb{N}{:=}\{0, 1, 2, 3, \ldots\}.$$

ausdrücken. Das Zeichen ":=" besagt, daß das linke Symbol eine abkürzende Schreibweise für den Ausdruck auf der rechten sein soll. Der Ausdruck auf der rechten bezeichnet eine Menge, wobei die Elemente der Menge gerade die Objekte sind, die zwischen den geschweiften Klammern aufgeführt sind; da man nicht alle natürlichen Zahlen auflisten kann, behilft man sich mit "..." und appelliert damit an die Phantasie des Betrachters. Der im letzten Jahrhundert lebende Mathematiker Kronecker (1823–1891) hat gesagt, die natürlichen Zahlen habe der liebe Gott gemacht, der Rest sei Menschenwerk. (Die Zahl "0" ist möglicherweise auch schon als Menschenwerk zu betrachten.) Tatsächlich kann man alle anderen Zahlsysteme mengentheoretisch aus den natürlichen Zahlen heraus konstruieren und ihre Eigenschaften aus denen der natürlichen Zahlen ableiten. Mit natürlichen Zahlen kann man rechnen, man kann sie addieren und multiplizieren: die Menge der natülichen Zahlen ist *abgeschlossen unter den Operationen* "+" und "·".

Mengen

Ganze Zahlen

Allerdings kann man natürliche Zahlen nicht uneingeschränkt voneinander subtrahieren. Dies kann man erst im nächst größeren Bereich

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$$

der *ganzen Zahlen*. Hier kann nun uneingeschränkt addiert, subtrahiert und multipliziert werden, allerdings kann man nicht uneingeschränkt durch von 0 verschiedene Zahlen dividieren. Dies ist erst im wiederum nächst größeren Bereich $\mathbb{Q}$ der rationalen Zahlen möglich. Die Elemente von $\mathbb{Q}$ sind *Brüche* $\frac{p}{q}$, wobei p, q ganze Zahlen sind, allerdings q von 0 verschieden ist, und wobei

Rationale Zahlen

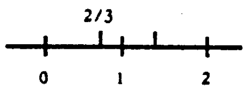
zwei Brüche $\frac{p}{q}$ und $\frac{p'}{q'}$ als gleich angesehen werden, falls die ganzen Zahlen pq' und $p'q$ gleich sind. Letzteres schreibt man auch symbolisch

$$\frac{p}{q} = \frac{p'}{q'} :\Longleftrightarrow pq' = p'q.$$

Es wird hier natürlich vorausgesetzt, daß man mit rationalen Zahlen rechnen kann, allerdings kann man—wie schon erwähnt—die rationalen Zahlen zusammen mit ihren Operationen "$\pm$", "$\cdot, \div$" mengentheoretisch aus den natürlichen Zahlen konstruieren. Nehmen wir einmal an, wir hätten schon die ganzen Zahlen konstruiert, so wäre eine typische Vorgehensweise zur Konstruktion der rationalen Zahlen, einen Bruch $\frac{p}{q}$ als Menge der *Paare* (tp, tq) zu *definieren*, wo t die Menge der ganzen, von 0 verschiedenen Zahlen durchläuft. Dabei wäre jetzt eigentlich noch der Begriff "Paar" zu erklären (mengentheoretisch ein Paar (a, b) wiederum definiert als die Menge $\{a, \{a, b\}\}$) etc..

Reelle Zahlen

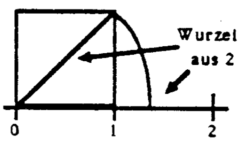
Das Hauptobjekt der Infinitesimalrechnung ist die Menge $\mathbb{R}$ der reellen Zahlen. Wählt man eine Gerade und darauf zwei Punkte, so kann man sich die reellen Zahlen als Punkte auf dieser Geraden vorstellen. Die beiden Punkte entsprechen den Zahlen 0 und 1, die natürliche Zahl n erhält man, indem man $n-$mal die Strecke von 0 nach 1 ausgehend von der 0 in Richtung der 1 abträgt, die negativen Zahlen, indem man in die entgegengesetzte Richtung abträgt, und die rationale Zahl $\frac{p}{q}$ — wobei wir q als positive ganze Zahl voraussetzen können — erhält man, indem man erst den p entsprechenden Punkt auffindet und dann die Strecke von 0 nach p in q gleiche Teile teilt; der dem Punkt 0 am nächsten gelegene Teilungspunkt ist dann die Zahl $\frac{p}{q}$.



Die Punkte der Geraden werden durch die rationalen Zahlen nicht ausgeschöpft. Es bleiben noch Punkte übrig: die irrationalen Zahlen, d.h. diejenigen reellen Zahlen, die nicht rational sind—in Symbolen: die Elemente der Menge $\mathbb{R} \setminus \mathbb{Q}$. Das es tatsächlich irrationale Zahlen gibt, war eine der großen Entdeckungen in der Mathematik. Eudoxos entdeckte die folgende Tatsache:

> Die Seite und die Diagonale eines Quadrates sind inkommensurabel.

Zwei Strecken heißen kommensurabel, falls man sie beide in Strecken aufteilen kann, von denen alle ein-und dieselbe Länge haben; andernfalls heißen sie inkommensurabel. Nimmt man ein Quadrat, dessen eine Seite gerade die Strecke von 0 nach 1 auf unserer Zahlengeraden ist, so besagt Eudoxos Satz also, daß die Länge der Diagonalen dieses Quadrats keine rationale Zahl ist. Die Länge dieser Diagonalen ist nach dem Satz des Pythagoras $\sqrt{2}$, wobei dies Symbol gerade die nichtnegative Zahl bedeutet, deren Quadrat 2 ergibt. In moderner Sprechweise besagt Eudoxos Entdeckung daher



Satz. *Die Zahl* $\sqrt{2}$ *ist irrational.*

Beweis. Der Beweis, den wir hier geben ist ein typisches Beispiel eines *indirekten Beweises*. Wir nehmen an, daß $\sqrt{2}$ rational ist und führen dies zum Widerspruch. Angenommen also, $\sqrt{2}$ sei rational. Dann gibt es ganze Zahlen p und $q \neq 0$, sodaß $\sqrt{2} = \frac{p}{q}$. Wir können dabei annehmen, daß p und q nicht beide zugleich gerade sind, denn andernfalls können wir p und q solange durch zwei teilen, bis mindestens eine der beiden Zahlen ungerade ist. Die Gleichung $\sqrt{2} = \frac{p}{q}$ ist gleichbedeutend mit $2q^2 = p^2$. Hieraus folgern wir, daß p gerade sein muß, denn hiernach ist jedenfalls p^2 gerade, und das Quadrat einer ungeraden Zahl wäre wieder ungerade. Also kann man $p = 2p'$ mit einer geeigneten ganzen Zahl p' schreiben. Setzen wir das in die letzte Gleichung ein und teilen wir beide Seiten der resultierenden Gleichung sogleich noch durch 2, so erhalten wir $q^2 = 2p'^2$. Wie eben folgt hieraus, daß q gerade ist. Da aber nicht beide, p und q, gerade sind, haben wir einen Widerspruch gefunden.

Zahldarstellungen

Wie notiert man reelle Zahlen. Für natürliche Zahlen hat man die Dezimalschreibweise

$$a_n a_{n-1} \dots a_3 a_2 a_1 a_0,$$

mit *Ziffern* a_k aus der Menge der Zahlen von 0 bis 9, die als Abkürzung für

$$a_n \cdot 10^n + \cdots + a_2 \cdot 100 + a_1 \cdot 10 + a_0$$

steht, und die schon in Indien vor Jahrhunderten verwendet wurde. Ganze Zahle erhält man, indem man gegebenenfalls ein Minus-Zeichen vor solche Ausdrücke setzt, und die rationalen Zahlen notiert man als Paare solcher Dezimalzahlen, eine über und eine unter dem Bruchstrich. Die Notierung von reellen Zahlen dagegen fordert eine völlig neue Begriffsbildung. Wie wir sahen, ist nicht jede reelle Zahl rational, d.h. nicht jeder Punkt auf der Zahlengeraden entspricht einem Bruch. Aber immerhin liegen die rationalen Zahlen *dicht* auf der Zahlengeraden, d.h. die Brüche kommen jedem Punkt auf der Zahlengeraden beliebig nahe. Eine reelle Zahl notiert man also am naheliegendsten, indem man eine Folge von rationalen Zahlen angibt, sodaß die Glieder dieser Folge der ins Auge gefaßten reellen Zahl immer näher kommen. Solch eine Approximation von reellen Zahlen durch rationale war im Grunde schon den Griechen bekannt, wenn auch eine logisch einwandfreie Fassung dieser "infiniten" Prozesse erst sehr viel später erfolgte. So hatte schon Archimedes benutzt, daß

$$\frac{265}{153} < \pi < \frac{1351}{780}$$

ist, wo π den Umfang des Kreises vom Radius 1 bezeichnet.

Wie notiert man also eine reelle, nicht notwendig rationale Zahl. Wir geben drei Beispiele; das erste ist ein Beispiel für die *Dezimaldarstellung* einer reellen Zahl, hier der Zahl π

$$
\begin{array}{l}
\pi = \\
3.1415926535897932384626433832795028841971693993751 05 \\
8209749445923078164062862089986280348253421170679821 4 \\
8086513282306647093844609550582231725359408128481117 4 \\
5028410270193852110555964462294895493038196442881097 5 \\
6659334461284756482337867831652712019091456485669234 6 \\
0348610454326648213393607260249141273724587006606315 5 \\
8817488152092096282925409171536436789259036001133053 0 \\
5488204665213841469519415116094330572703657595919530 9 \\
2186117381932611793105118548074462379962749567351885 7 \\
5272489122793818301194912983367336244065664308602139 4 \\
9463952247371907021798609437027705392171762931767523 8 \\
4674818467669405132000568127145263560827785771342757 7 \\
8960917363717872146844090122495343014654958537105079 2 \\
2796892589235420199561121290219608640344181598136297 7 \\
4771309960518707211349999998372978049951059731732816 0 \\
9631859502445945534690830264252230825334468503526193 1 \\
1881710100031378387528865875332083814206171776691473 0 \\
3598253490428755468731159562863882353787593751957781 8 \\
577805321712268066130019278766111959092164201 99\cdots
\end{array}
$$

Dies steht als Abkürzung für die Folge von rationalen Zahlen

$$
\begin{gathered}
3,\ 3+\frac{1}{10}=\frac{31}{10},\ 3+\frac{1}{10}+\frac{4}{100}=\frac{157}{50} \\
3+\frac{1}{10}+\frac{4}{100}+\frac{1}{1000}=\frac{3141}{1000}, \\
3+\frac{1}{10}+\frac{4}{100}+\frac{1}{1000}+\frac{5}{10\,000}=\frac{6283}{2000}.
\end{gathered}
$$

Der Nachteil dieser Darstellung ist, daß man nicht auf den ersten Blick sieht, wie es weiter geht — aber natürlich gibt es Algorithmen, um (theoretisch beliebig viele) weitere *Dezimalstellen* von π zu berechnen. In geschlossenerer Form steht die Folge von rationalen Zahlen

$$1, 1+\frac{1}{4}, 1+\frac{1}{4}+\frac{1}{9}, 1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}, \ldots,$$

die man noch prägnanter als

$$\sum_{n=1}^{\infty}\frac{1}{n^2}$$

schreiben kann. Hier wird der Algorithmus zur Berechnung der Folge von rationalen Zahlen sofort mitgeliefert. Aber natürlich haben solche einfachen Darstellungen auch Nachteile: man fragt sich nämlich sofort, welche ausgezeichnete reelle Zahl durch solch eine ausgezeichnet einfach gebaute Folge dargestellt wird. Die reelle Zahl, die hier dargestellt wird, ist $\frac{\pi^2}{6}$. Aber für das Beispiel

$$\sum_{n=1}^{\infty} \frac{1}{n^3}$$

ist die *Natur* der hier dargestellte Zahl bis heute noch unklar; wir werden unten noch einmal darauf zurückkommen und dies präzisieren.

Eine drittes Beispiel für die Darstellung reeller Zahlen ist der *unendliche Kettenbruch*

$$x = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}}}} \quad .$$

Hier ist der Algorithmus zur Berechnung der Folge von rationalen Zahlen sofort ersichtlich:

$$1, 1 + \frac{1}{1} = 2, 1 + \frac{1}{1+1} = \frac{3}{2}, 1 + \frac{1}{1 + \frac{1}{1+1}} = \frac{5}{3}, \ldots,$$

als auch die Natur der dargestellten Zahl: aufgrund der Struktur des Kettenbruchs muß offenbar gelten

$$x = 1 + \frac{1}{x},$$

d.h.

$$x^2 - x - 1 = 0,$$

also

$$x = \frac{1 + \sqrt{5}}{2}$$

(oder $= \frac{1-\sqrt{5}}{2}$, was aber sofort ausgeschloßen werden kann, da x offenbar positiv sein soll).

Wie schon das eben gegebene Beispiel des Kettenbruchs nahelegt, sind die Wurzeln von rationalen Zahlen in einem gewissen Sinne einfacher, den rationalen Zahlen näher verwandt als etwa π. Dies kann man leicht mittels der folgenden Beobachtung präzisieren? Die rationale Zahl r ist Lösung der *linearen* Gleichung

$$x - r = 0.$$

Die Zahl

$$\frac{1 + \sqrt{5}}{2}$$

ist Lösung der *quadratischen* Gleichung

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

In diesem Sinne sind die den rationalen Zahlen verwandtesten Zahlen diejenigen reellen Zahlen x, die Lösung einer Gleichung

$$x^n + r_{n-1}x^{n-1} + r_{n-2}x^{n-2} + \cdots + r_1 x + r_0 = 0,$$

Algebraische Zahlen

wo n eine positive ganze Zahl und die $r_{n-1}, r_{n-2}, \ldots$ rationale Zahlen sind. Solche Zahlen bezeichnet man als *(reell-)algebraische Zahlen*. Man kann zeigen daß der Bereich dieser Zahlen abgeschloßen ist unter Addition, Subtraktion, Multiplikation und Division, d.h. mit zwei algebraischen Zahlen ist auch wieder ihre Summe, Differenz, etc. eine algebraische Zahl. Tatsächlich ist nicht jede Zahl algebraisch; zum Beispiel ist die Zahl π nicht algebraisch.

Transzendente Zahlen

Nicht algebraische Zahlen nennt man *transzendent*. Die Transzendenz von π ist eine ganz und gar nicht offensichtliche Tatsache. Es wurde erst 1882 von Lindemann bewiesen, daß π transzendent ist. Die Transzendenz von π impliziert auch die Unmöglichkeit der Quadratur des Kreises: Es ist unmöglich lediglich mit Zirkel und Lineal die Seite eines Quadrats zu konstruieren, welches den gleichen Flächeninhalt wie der Kreis vom Radius 1 hat (die Koordinaten aller Punkte in einem kartesischen Koordinatensystem, die sich nur mit Zirkel und Lineal aus den Punkten mit Koordinaten $(0,0)$ und $(0,1)$ konstruieren laßen sind algebraische Zahlen). Beim Begriff "Transzendenz" gibt es heute noch viele ungelöste Fragen. Zum Beispiel weiß man nicht, ob die oben angegebene Zahl $\sum_{n=1}^{\infty} \frac{1}{n^3}$ transzendent ist oder nicht; man weiß — und dies erst seit relativ kurzem — lediglich, daß sie jedenfalls nicht rational ist (Apéry 1978).

Komplexe Zahlen

Betrachtet man algebraische Zahlen, so stößt man bald auf das Problem, daß solch einfache Gleichungen wie $x^2 + 1 = 0$ scheinbar keine Lösung besitzen. Natürlich besitzt diese Gleichung sehr wohl eine Lösung, allerdings nicht im Bereich der reellen Zahlen, sondern im größeren Bereich der komplexen Zahlen. Eine komplexe Zahl notiert man in der Form

$$a + ib,$$

wo a und b reelle Zahlen bedeuten. Man rechnet mit solchen Ausdrücken wie man es mit reellen Zahlen gewohnt ist, unter Benutzung der zusätzlichen Regel, daß stets eine Potenz i^2 durch die ganze Zahl -1 ersetzt werden kann. Das Symbol i bezeichnet man als *imaginäre Einheit*. Es ist gerade eine Lösung der Gleichung $x^2 + 1 = 0$; die andere Lösung ist $-i$ (d.h. das Symbol $0 + (-1)i$). Die fundamentale Aussage in diesem Zusammenhang ist, daß jede Gleichung

$$x^n + r_{n-1}x^{n-1} + r_{n-2}x^{n-2} + \cdots + r_1 x + r_0 = 0$$

(die *Koeffizienten* r_k sind irgendwelche vorgegebenen *komplexen* Zahlen) mindestens eine Lösung im Bereich der komplexen Zahlen besitzt (*Fundamentalsatz der Algebra*, Gauß, 1799). Als algebraische Zahl schlechthin (d.h. ohne den Zusatz *reell*) bezeichnet man jede Lösung einer solchen Gleichung mit rationalen Koeffizienten. Ein oft benütztes Symbol für die Gesamtheit der algebraischen Zahlen ist $\overline{\mathbb{Q}}$, das heute allgemein übliche Symbol für die Gesamtheit der komplexen Zahlen ist $\mathbb{C}$.

Wir fassen die in diesem ersten Kapitel angedeuteten Zusammenhänge noch einmal symbolisch in einem Diagramm zusammen:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \overline{\mathbb{Q}} \cap \mathbb{R} \subset \mathbb{R} \subset \mathbb{C} \supset \overline{\mathbb{Q}}$$

$$\sqrt{2} \in \overline{\mathbb{Q}} \setminus \mathbb{Q}, \quad \pi \in \mathbb{R} \setminus \overline{\mathbb{Q}}, \quad \sum_{n=1}^{\infty} \frac{1}{n^3} \notin \mathbb{Q}, \in \{?\}.$$

2 DIE AXIOMATIK DER REELLEN ZAHLEN

R hat zwei Verknüpfungen $+$ und $\cdot$, d.h. für zwei reelle Zahlen x und y ist $x+y$ ebenso wie $x \cdot y$ ein wohlbestimmtes Element von $\mathbb{R}$. Diese Verknüpfungen erfüllen eine Serie von Axiomen, von denen die ersten vier die folgenden sind:

Axiome für die Addition

(A1) Für alle $x, y, z \in \mathbb{R}$ gilt das *Assoziativgesetz* $(x+y)+z = x+(y+z)$.

(A2) Es gibt ein Element $0 \in \mathbb{R}$, sodaß $x+0 = 0+x = x$ für alle $x \in \mathbb{R}$ gilt.

(A3) Zu jedem $x \in \mathbb{R}$ existiert ein $x \in \mathbb{R}$ ein $x' \in \mathbb{R}$, sodaß $x+x' = x'+x = 0$

(A4) Für alle $x, y \in R$ gilt das *Kommutativgesetz* $x+y = y+x$.

Eine erste, einfache Folgerung aus diesen Axiomen ist die Eindeutigkeit des *additiven Neutralelements* 0; genauer:

Satz. *Ist $0' \in \mathbb{R}$, sodaß für jedes $x \in \mathbb{R}$ die Gleichung $0'+x = x+0' = x$ besteht, dann ist $0' = 0'$.*

Beweis. Die Voraussetung und Axiom (A2) angewandt auf $x = 0$ resp. $x = 0'$ liefert $0 = 0+0' = 0'$.

Ferner hat man

Satz. *Das zu jedem $x \in R$ nach (A3) existierende additives Inverse x' ist eindeutig bestimmt.*

Beweis. Es gelte $x+x'' = x''+x = 0$. Dann hat man

$$\begin{aligned} x'' &= x''+0 && \text{(nach (A2)} \\ &= x''+(x+x') && \text{(nach (A3))} \\ &= (x\text{"}+x)+x' && \text{(nach (A1))} \\ &= 0+x' && \text{(nach Voraussetzung)} \\ &= x' && \text{(nach (A2)).} \end{aligned}$$

Das zu einem x eindeutig bestimmte x' mit $x + x' = x' + x = 0$ wird mit $-x$ bezeichnet.

Satz. $-(-x) = x$

Beweis. Es gilt $x + (-x) = 0$. Also ist x das eindeutig bestimmte additive Inverse zu $(-x)$, d.h. $x = -(-x)$.

Satz. *Ist $a + x = b$ and $a + y = b$, so ist $x = y$.*

Beweis. Mit $a+x = b = a+y$ folgt $x = (-a)+(a + x) = (-a)+(a + y) = y$.

Dieser Satz besagt, daß die Gleichung $a + x = b$ zu fest gegebenen a, b von höchstens einem $x \in \mathbb{R}$ gelöst wird. Es ist klar, daß es mindestens eine Lösung, nämlich $x = b + (-a)$. Hierfür schreibt man auch $b - a$.

Es gibt noch eine Reihe weiterer wichtiger Folgerungen aus den Axiomen, wie etwa

$$a - (b - a) = (a - b) + c$$

(Man beweise zunächst $-(x + y) = (-x) + (-y)$, womit dann $-(b - c) = (-b) + (-(-c))$, wegen $-(-c) = c$ also $-(b - c) = (-b) + c$ folgt; setze dies in die behauptete Identität ein und wende (A1) an.) Alle diese Folgerungen sind nicht anderes als die vertrauten Rechenregeln für die Addition. Der wirklich bemerkenswerte Punkt ist hierbei, daß man alle diese Rechenregeln formal aus den Axiomen ableiten kann.

Die Axiome für die Multiplikation sind den Axiomen für die Addition völlig analog.

Axiome für die Multiplikation

(M1) Für alle $x, y, z \in \mathbb{R}$ gilt das Assoziativgesetz $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

(M2) Es gibt ein Element $e \in \mathbb{R}$, $e \neq 0$, sodaß $x \cdot e = e \cdot x = x$ für alle $x \in \mathbb{R}$ gilt.

(M3) Zu jedem $x \in \mathbb{R}$, $x \neq 0$ existiert ein $x \in \mathbb{R}$ ein $x' \in \mathbb{R}$, sodaß $x \cdot x' = x' \cdot x = e$

(M4) Für alle $x, y \in R$ gilt das Kommutativgesetz $x \cdot y = y \cdot x$.

Meist läßt man beim Rechnen den $\cdot$ weg, d.h. man schreibt also ab statt $a \cdot b$ etc.. Wie oben kann man beweisen

Satz. *Ist $e' \in \mathbb{R}$, sodaß für jedes $x \in \mathbb{R}$ die Gleichung $e' \cdot x = x \cdot e' = x$ besteht, dann ist $e' = e$.*

Satz. *Das zu jedem $x \in R$ nach (M3) existierende multiplikative Inverse x' ist eindeutig bestimmt.*

Das somit eindeutig bestimmte multiplikative Neutralelement e bezeichnet man üblicherweise mit 1, und für das zu gegebenem eindeutig bestimmte $a \neq 0$ multiplikative Inverse schreibt man a^{-1} oder $\frac{1}{a}$. Weiter folgt wie oben:

Satz. $\left(a^{-1}\right)^{-1} = a$.

Satz. *Zu $a \neq 0$ and b aus $\mathbb{R}$ existiert genau ein Element x mit $ax = b$.*

Dieses x ist gerade ba^{-1}. Man schreibt hierfür auch $\frac{b}{a}$.

Das letzte der für die Addition und Multiplikation geltende Axiom ist das *Distributivgesetz*:

Distributivgesetz

(D) $a(b+c) = ab + ac$ für alle $a, b, c \in R$.

Satz. *Für alle $x \in \mathbb{R}$ gilt $0 \cdot x = 0$.*

Beweis. Man hat $0 \cdot x = (0+0) \cdot x = 0 \cdot x + 0 \cdot x$ — die erste Gleichnung nach (A2), die zweite nach (D). Also ist $0 \cdot x$ eine Lösung der Differenzenaufgabe $0 \cdot x = y + 0x$. Da nach (A2) auch 0 eine ist, folgt nach einem oben ausgesprochenen Satz $0 = 0 \cdot x$.

Satz. $(-a) \cdot b = -(ab)$.

Beweis. Behauptet wird also, daß $(-a) \cdot b$ das additive Inverse zu ab ist d.h. $ab + (-a)b =$. Aber das ist wahr, denn $ab + (-a)b = (a + (-a))b = 0 \cdot b = 0$.

Satz. *$a(b-c) = ab - ac$ für all $a, b, c \in R$.*

Beweis. Behauptet wird, daß $a(b-c)$ Lösung der Differenzaufgaben $ac + y = ab$ ist. Tatsächlich ist $ac + (a(b-c)) = ac + ab + a(-c) = ab$.

Satz. *Sind $a_1, a_2 \neq 0$, so ist auch $a_1 \cdot a_2 \neq 0$. Außerdem ist*

$$\frac{b_1}{a_1} + \frac{b_2}{a_2} = \frac{b_1 a_2 + a_1 b_2}{a_1 \cdot a_2}.$$

Den Beweis lassen wir als"Übungsaufgabe". (Es genügt zu zeigen, daß $\frac{b_1}{d_1} + \frac{b_2}{a_2}$ eine Lösung der Divisionsaufgabe $(a_1 \cdot a_2) \cdot x = b_1 a_2 + a_1 b_2$ ist.)

Allgemeine Körper

Durch die bisher genannten Axiome

(A1) Assoziativität	(M1) Assoziativität
(A2) Neutrales 0	(M2) Neutrales $1 \neq 0$
(A3) Gegenelement	(M3) Gegenelement
(A4) Kommutativgesetz	(M4) Kommutativität

(D) Distributivität

sind die reellen Zahlen noch nicht eindeutig festgelegt; eine beliebige Menge M mit zwei Verknüpfungen $+$ und $\cdot$, die diese Axiome erfüllt, nennt man einen *Körper*. Tatsächlich gibt es viele Körper, die nichts mit dem gemein haben, was wir uns langläufig unter reellen Zahlen vorstellen. Man kann aus den angeführten Axiomen nicht einmal folgern, daß es mehr als zwei reelle Zahlen gibt, denn es gibt Körper mit endlich vielen Elementen, und sogar einen mit 2 Elementen.

Teilt man nämlich die natürlichen Zahlen in die zwei Klassen "ger" der geraden und "ung" der ungeraden Zahlen ein, addiert und multipliziert sie, wie ihre Elemente es tun würden, d.h.

ger+ger=ger	ger ·ger=ger
ger+ung=ung	ger ·ung=ger
ung+ger=ung	ung·ger=ger
ung+ung=ger	ung·ung=ung,

so entsteht eine Körper mit 2 Elementen, in dem ger der 0 und ung der 1 entspricht. Hier ist also insbesondere $1+1=0$, d.h. $1=-1$.

Ähnlich ist es mit den *Restklassen modulo* 3: Man unterteilt die natürlichen Zahlen in 3 Klassen danach, welchen Rest sie bei Division durch 3 lassen, z.B. gehört $17 = 5\cdot 3+2$ in die "2" genannte Klasse. Addition und Multiplikation der Klassen geschieht wieder repräsentantenweise: z.B. bestimmt man das Quadrat der Klasse 2 als Klasse 1 :

$$(3k+2)(3l+2) = 9kl + 6k + 6l + 4 = 3z + 1,$$

wobei

$$z = 3kl + 2(k+l) + 1$$

wie $k, l \in \mathbb{Z}$ ist. — So entsteht ein Körper mit 3 Elementen. Dagegen bilden die Restklassen modulo 4 keinen Körper, sondern nur einen *Ring*. Erfüllt eine Menge M mit zwei Verknüpfungen $+, \cdot$ die obigen Axiome mit Ausnahme von (M3), so nennt man sie einen kommutativen Ring. In einem Körper folgt z.B. aus $x\cdot y = 0$ stets $x = 0$ oder $y = 0$; in einem kommutativen Ring gilt das i.a. nicht mehr. (Gilt in einem Körper $x \cdot y = 0$, aber $y \neq 0$, so existiert das Inverse y^{-1}, und es ist $0 = (x\cdot y)\cdot y^{-1} = x$. Im Restklassenring mit vier Elementen ist $2 \neq 0$, aber $2\cdot 2 = 0$.)

Wie wir eben gesehen haben, benötigen wir also weitere Axiome, um die reellen Zahlen eindeutig zu charakterisieren. Ein Punkt, der die reellen Zahlen weiter unter allen möglichen Körpern hervorhebt, ist, daß sie *angeordnet* sind. Genauer gelten für die reellen Zahlen die sogenannten Anordnungsaxiome.

Anordnungsaxiome

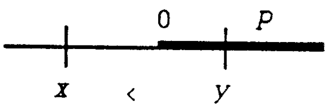
Es gibt eine Menge $P \subset \mathbb{R}$, sodaß gilt:

(O1) Von den drei Aussagen $x \in P$, $-x \in P$, $x = 0$ ist für jedes $x \in \mathbb{R}$ genau eine wahr.

(O2) Sind $x, y \in P$, so sind auch $x + y$ und $x \cdot y \in P$.

Die Menge P nennt man die Menge der positiven reellen Zahlen. Wir definieren nun

$$x > y :\Longleftrightarrow x - y \in P$$
$$x < y :\Longleftrightarrow y - x \in P.$$



Anschaulich sind auf der Zahlengeraden die positiven Zahlen diejenigen, die rechts von 0 liegen, und $x > y$ genau dann, wenn x rechts von y liegt. Insbesondere hat man in diesen Bezeichnungen

$$x > 0 \Longleftrightarrow x \in P, \qquad x < 0 \Longleftrightarrow -x \in P,$$

und von den drei Aussagen

$$x > 0,\ x < 0,\ x = 0$$

ist für jedes $x \in \mathbb{R}$ jeweils genau eine erfüllt.

Satz. *Für $x, y \in \mathbb{R}$ ist von den drei Aussagen $x > y,\ x < y,\ x = y$ stets genau eine erfüllt.*

Beweis. Man hat ja

$$x < y \Longleftrightarrow y - x \in P \Longleftrightarrow a \in P,$$
$$x > y \Longleftrightarrow x - y \in P \Longleftrightarrow a \notin P,$$
$$x = y \Longleftrightarrow y - x = 0 \Longleftrightarrow a = 0,$$

wobei $a = y - x$. Aus (O1) folgt nun die Behauptung.

Satz. *Für $x, y, z \in \mathbb{R}$ folgt aus $x < y$ und $y < z$ stets $x < z$.*

Beweis. Sind $y - x$ und $z - y$ in P, so ist nach (O2) auch $z - x = (z - y) + y - x$ in P.

Hier noch einige Regeln, die man entsprechend beweisen kann:

Satz. *Für $x, y, z \in \mathbb{R}$ gilt:*

$$x < y \Longleftrightarrow x + z < y + z$$
$$(x < y \wedge z > 0) \Rightarrow x \cdot z < y \cdot z.$$

Die zweite Aussage im letzten Satz ist übrigens äquivalent zu

$$0 < z \Rightarrow (x < y \Longleftrightarrow xz < yz).$$

Man nennt ein $x \in \mathbb{R}$ *negativ*, falls $-x \in P$ liegt. Damit gilt

Satz. *Für die Multiplikation von positiven und negativen Zahlen gelten die folgenden Regeln:*

pos·pos=pos
pos·neg=neg
neg·neg=pos.

Beweis. *Die erste Gleichung gilt nach (O2). Zur zweiten: Ist $x \in P$ und $-y \in P$, so ist $x \cdot (-y) = -xy \in P$ (nach (O2)), also xy negativ. Die dritte folgt analog.*

Hiermit kann man zum Beispiel zeigen:

Satz. *Für alle $x \in \mathbb{R}$ gilt:*

$$x > 0 \Rightarrow \frac{1}{x} > 0.$$

(Insbesondere ist $1 > 0$).

Beweis. Es ist zunächst $1 > 0$, denn nach dem ersten und dritten Fall des letzten Satzes ist ja $1 \cdot 1 > 0$; aber $1 = 1 \cdot 1$. Damit ist dann aber auch $\frac{1}{x} = 1 > 0$; da nach Voraussetzung $x > 0$ ist, muß nach dem letzten Satz also(1)$x > 0$gelten.

Eine weitere wichtige Folgerung des vorletzten Satzes ist der

Satz. *Für alle reellen Zahlen $a \neq 0$ gilt $a^2 > 0$.*

Satz. *Für alle reellen Zahlen x, y gilt*

$$0 < x < y \Rightarrow 0 < y^{-1} < x^{-1}.$$

(Hierbei steht $a < b < c$ zur Abkürzung für "$a < b$ und $b < c$".)

Beweis. Es ist

$$\frac{1}{x} - \frac{1}{y} = (y - x)(xy)^{-1}.$$

Aber $y - x > 0$ nach Voraussetzung, und wegen $x, y > 0$ und (Ord2) ist auch $xy > 0$. Nach obigem Satz ist also $(y - x)(xy)^{-1} > 0$.

Satz. *Für alle reellen Zahlen x, y gilt: $x, y \in P \Rightarrow \big(x < y \Longleftrightarrow x^2 < y^2\big)$*

Beweis. Es ist $y^2 - x^2 = (y - x)(y + x)$. Hierbei ist wegen $x, y \in P$ nach (O2) $y + x > 0$. Die Behauptung folgt nun leicht mit dem fünftletzten Satz.

Schließlich führen wir noch die folgende Bezeichnung ein:

$$x \leq y :\Longleftrightarrow (x < y \text{ oder } x = y).$$

(lies x *kleinergleich* y). Für diese Relation gelten als Folgerungen obiger Regeln zahllose ähnliche Regeln wie etwa

$$(x \leq y \text{ und } y \geq 0 \Rightarrow xz \leq yz,$$
$$x \geq 0 \quad y \leq 0 \Rightarrow \big(x \geq y \iff x^2 \leq y^2.\big)$$

Quadratwurzeln

Ist $a \in \mathbb{R}$, gibt es dann $x \in \mathbb{R}$, sodaß $x^2 = a$? Das ist ein ernsthaftes Problem, denn die Körper -und Ordnungsaxiome werden auch von der Menge $\mathbb{Q}$ der rationalen Zahlen erfüllt und seit Euklid wissen wir ja, daß die Gleichung $x^2 = 2$ in $\mathbb{Q}$ keine Lösung besitzt. Man kann also aus den bisher angeführten Axiomen nicht die Existenz von Quadratwurzeln folgern.

Dennoch kann man schon einiges über Quadratwurzeln folgern, wenn auch noch nicht ihre Existenz. Ist $a < 0$, dann gibt es keine Lösung von $x^2 = a$, denn Quadrate können nach einem oben aufgeführten Satz nicht negativ sein. Ist dagegen $a > 0$ so gibt es entweder keine Lösung von $x^2 = a$ oder aber genau zwei. Denn gibt es eine Lösung — etwa $x_1^2 = a$ — so ist auch $(-x_1)^2 = a$, und, da $x_1 \neq 0$, ist nach (O1) $-x_1 \neq x_1$; es gibt dann also mindestens zwei Lösungen. Ist $y^2 = a$, so folgt $0 = y^2 - x_1^2 = (y-x_1)(y+x_1)$, also nach einem oben bewiesenen Satz $y - x_1 = 0$ oder $y + x_1 = 0$, d.h. $y = x_1$ oder $y = -x_1$; es gibt also höchstens zwei Lösungen. Aus dem eben gegebenen Beweis folgt auch: Ist $a \geq 0$ und gibt es ein $x \leq 0$ mit $x^2 = a$, dann ist dieses x eindeutig bestimmt. Dieses x wird mit $\sqrt{a}$ bezeichnet.

Satz. *Seien $a \geq 0$ und $b > 0$, und $\sqrt{a}$, $\sqrt{b}$ mögen existieren. Dann existiert auch $\sqrt{ab}$ und es gilt $\sqrt{ab} = \sqrt{a} \cdot \sqrt{b}$.*

Beweis. Es ist

$$(\sqrt{a} \cdot \sqrt{b})^2 = \sqrt{a}^2 \cdot \sqrt{b}^2 = ab, \quad \sqrt{a} \cdot \sqrt{b} > 0..$$

Also existiert die Quadratwurzel von ab und ist gleich $\sqrt{a} \cdot \sqrt{b}$.

Absolutbetrag

Den *Absolutbetrag* einer reellen Zahl definiert man vermöge

$$|x| := \begin{cases} x & \text{falls} \quad x \in P \\ -x & \text{falls} \quad -x \in P \\ 0 & \text{falls} \quad x = 0 \end{cases}$$

Man beachte, daß die Fallunterscheidung hier nach (O1) vollständig ist: der Absolutbetrag ist also für jede reelle Zahl definiert. Man beachte, daß stets $|x| \geq 0$ gilt. Weitere einfache Folgerungen sind $|x| \geq x$ und $|x| = \sqrt{x^2}$ für alle reellen x. Letzteres wäre übrigens eine alternative Definition für den Absolutbetrag.

Satz. *Für alle reellen Zahlen x, y gilt:*

(R1) $|x| \geq 0,\ |x| = 0 \Longleftrightarrow x = 0$

(R2) $|x + y| \leq |x| + |y|$

(R3) $|x \cdot y| = |x| \cdot |y|\,.$

Beweis. Wir beweisen nur (R2), den Beweis der übrigen Aussagen lassen wir zur Übung Da beide Seiten der behaupteten Ungleichung nichtnegativ sind, genügt es $|x+y|^2 \leq (|x| + |y|)^2$ zu beweisen. Dies ist aber äquivalent zu $(x+y)^2 \leq x^2 + 2\,|x|\,|y| + y^2$ (da ja stets $|x|^2 = x^2$), und dies wiederum zu $xy \leq |xy|$, was wir oben schon als wahre Aussage notiert hatten.

Hiemit schließen wir für den Moment das Kapitel über die Axiome für die reellen Zahlen. Wie wir gesehen haben, charakterisieren die bisher aufgeführten Axiome die reellen Zahlen noch nicht ausreichend; die rationalen Zahlen erfüllen ebenso sämtliche der bisher aufgeführten Axiome. Wollen wir daher Lehrsätze für die reellen Zahlen ableiten, die nicht schon für die rationalen Zahlen gelten, so benötigen wir offenbar mehr Axiome. Tatsächlich stellt sich heraus, daß lediglich genau ein weiteres Axiom hinzugenommen werden muß, um die reellen Zahlen von allen anderen Körpern zu unterscheiden; dieses Axiom werden wir im Kapitel 5 behandeln.

3 ÜBER VOLLSTÄNDIGE INDUKTION

Natürliche, ganze und rationale Zahlen in $\mathbb{R}$

In dem von uns betrachteten Körper $\mathbb{R}$ gibt es die Elemente $0, 1$, also auch $1+1, 1+1+1, 1+1+1+1\ldots$ Diese Zahlen sind im Körper $\mathbb{R}$ auch alle verschieden, denn jedes Element dieser Serie ist größer als das vorhergehende. Wir nennen diese Elemente natürliche Zahlen. Des weiteren nennen wir die Menge der natürlichen Zahlen zusammen mit ihren nach (A3) in $\mathbb{R}$ existierenden additiven Inversen die Menge der ganzen Zahlen, und die Teilmenge aller reellen Zahlen, die eine Gleichung $px = q$ mit ganzen Zahlen p, q lösen, die Menge der rationalen Zahlen. Nimmt man also die Menge der reellen Zahlen als irgendwie gegeben an — wie es unser Standpunkt ist —, so hat man darin als Teilmengen die natürlichen, ganzen und rationalen Zahlen. Die grundlegenden Eigenschaften dieser Zahlbereiche kann man nun aus den für die reellen Zahlen geltenden Axiomen folgern (die allerdings noch nicht vollständig aufgelistet wurden). Es gibt allerdings auch den umgekehrten Standpunkt, daß man die Menge der natürlichen Zahlen mit geeigneten Axiomen eindeutig charakterisiert, und sodann aus der als gegeben angenommenen Menge der natürlichen Zahlen mittels mengentheoretischer Operationen eine Menge und darauf Operationen “+”, “·” konstruiert, und schließlich (aus den Axiomen für die natürlichen Zahlen heraus) beweist, daß diese Menge mit ebendiesen Operationen gerade alle Axiome erfüllt, die hier und im weiteren für die reellen Zahlen ausgesprochen werden. Wir werden hier allerdings weder den einen noch den anderen Weg konsequent gehen; d.h. wir werden weder die grundlegenden Tatsachen über natürliche Zahlen mittels der Axiome für die reellen Zahlen verifizieren, noch werden wir die reellen Zahlen aus den natürlichen konstruieren.

Induktionsprinzip

Eine der grundlegendsten Tatsachen über natürliche Zahlen ist die folgende. Sei $A(k)$ eine Aussage, die von der natürlichen Zahl k abhängt; z.B. könnte $A(k)$ die Aussage

$$\sum_{i=0}^{k} i^2 = \frac{k(k+1)(2k+1)}{6}$$

sein. In Worten: “die Summe der ersten k Quadratzahlen ist gleich …”. Für jede natürliche Zahl k hat man also eine Aussage, die richtig oder falsch sein kann. Wie kann man beweisen, daß gegebenenfalls $A(k)$ für alle $k \in \mathbb{N}$ richtig ist? Dies geschieht mit dem sogenannten Induktionsprinzip:

man zeigt:

(I1) $A(0)$ ist richtig

(I2) Ist $A(n)$ für irgendein n wahr, so ist auch $A(n+1)$ wahr.

Es ist unmittelbar einleuchtend, daß dann $A(k)$ für alle natürlichen Zahlen gilt: nach (I1) ist $A(0)$; nach dem *Induktionschritt* (I2) ist dann auch $A(1)$ wahr, also auch $A(2)$, also $A(3)$, etc..

Auf eine exakte Begründung des Induktionsprinzips verzichten wir hier. Ferner weisen wir nur daraufhin, daß es noch verschiedene Varianten des Induktionsprinzips gibt; z.B. kann der Induktionsanfang auch bei einer von 0 verschiedenen Zahl $k_0 > 0$ liegen (um zu beweisen daß eine Aussage für alle $k \geq k_0$ richtig ist). Schließlich ist das Induktionsprinzip noch äquivalent zu der Aussage, daß jede nichtleere Menge von natürlichen Zahlen ein kleinstes Element enthält (was man mittels der Betrachtungen aus dem nächsten Kapitel beweisen könnte). Wir verzichten hier auf eine Diskussion dieser Tatsachen und verweisen auf Bücher, in denen die sogenannten *Peano-Axiome* behandelt werden. Stattdessen geben wir ein Beispiel für die Anwendung des Induktionsprinzips.

Aus der Schule bekannt ist vermutlich die Formel

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

Von Gauss wird erzählt, daß er als Sechsjähriger die Zahlen von 1 bis 100 aufaddieren sollte — damit sein Lehrer Ruhe habe — ; Gauss war schnell fertig:

$$\sum_{i=1}^{100} i = \frac{100 \cdot 101}{2} = 5050.$$

Er schrieb die Zahlen von 1 bis 100 in eine Zeile, darunter nochmals in umgekehrter Reihenfolge, und bemerkte sodann, daß die Summe zweier übereinander stehender Zahlen stets 101 ist, sodaß das gefragte Ergebnis also $\frac{1}{2} \times 101 \times$(Anzahl der Zahlen in einer Zeile), also gerade die oben gegebene Zahl sein muß.

Wir beweisen

Behauptung.

$$\sum_{i=0}^{k} i^3 = \left(\frac{k(k+1)}{2}\right)^2.$$

Beweis. Der Induktionsanfang — z.B. für $k = 0$ — sei dem Leser überlassen. Induktionsschritt: zu zeigen ist für beliebiges k, daß

$$\sum_{i=0}^{k+1} i^3 = \left(\frac{(k+1)(k+2)}{2}\right)^2$$

aus

$$\sum_{i=0}^{k} i^3 = \left(\frac{k(k+1)}{2}\right)^2$$

folgt. Tatsächlich hat man

$$\begin{aligned}
\sum_{i=0}^{k+1} i^3 &= (k+1)^3 + \sum_{i=0}^{k} i^3 \\
&= (k+1)^3 + \left(\frac{k(k+1)}{2}\right)^2 \quad \text{(nach Ind.-vorausetzung)} \\
&= \left(k+1+\left(\frac{k}{2}\right)^2\right)(k+1)^2 \\
&= \left(\frac{(k+1)(k+2)}{2}\right)^2,
\end{aligned}$$

wie zu beweisen war.

4 GENERAL NONSENSE UND ELEMENTARE KOMBINATORIK

Naive Mengenlehre

EinigeBezeichnungen zur Mengenlehre: Seien $M, N, \ldots$ Mengen. Dann ist

$$M \cap N = \{x \mid x \in M \text{ und } x \in N\}$$

der Durchschnitt von M und N,

$$M \cup N = \{x \mid x \in M \text{ oder } x \in N\}$$

die Vereinigung von M und N,

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

die Differenzmenge von M und N. Ist klar, daß eine Menge A Teilmenge einer *Allmenge* X ist, schreibt man auch

$$A^{\mathrm{c}} = \{x \in X \mid x \notin A\}$$

für das Komplement von A in X. Der Beweis von Aussagen über Mengen geschieht am besten mittels eines naiven Gebrauchs einiger logischer Symbole. Anstelle der Wörter 'und','oder' und 'nicht' kann man die zwischen (bzw. vor) Aussagen stehenden Zeichen $\wedge$,$\vee$, $\neg$ (oder $\sim$) verwenden.

$$a \Longrightarrow b$$

steht für 'Wenn die Aussage a gilt, so gilt auch die Aussage b';

$$a \Longleftrightarrow b$$

steht für 'Die Aussage a ist äquivalent zur Aussage b', also

$$(a \Longleftrightarrow b) \Longleftrightarrow ((a \Rightarrow b) \wedge (b \Rightarrow a)).$$

Man schreibt abkürzend '$\forall x$' statt 'für alle x', und $\exists x$ für 'es gibt ein x'; also z.B.

$$\begin{aligned} N \subset M :&\Longleftrightarrow \text{ für jedes } x(x \in N \Rightarrow x \in M) \\ &\Longleftrightarrow \forall x(x \in N \Rightarrow x \in M). \end{aligned}$$

Es gibt zahlreiche Regeln für den Umgang mit diesen Symbolen. Zum Beispiel gilt

$$N \subset M \Longleftrightarrow N \cap M = N,$$

ferner die *de Morganschen Regeln*

$$(A \cup B)^{\mathrm{c}} = A^{\mathrm{c}} \cap B^{\mathrm{c}}$$
$$(A \cap B)^{\mathrm{c}} = A^{\mathrm{c}} \cup B^{\mathrm{c}}.$$

Ferner hat man die *Distributivgesetze*

$$M \cap (N_1 \cup N_2) = (M \cap N_1) \cup (M \cap N_2),$$
$$M \cup (N_1 \cap N_2) = (M \cup N_1) \cap (M \cup N_2),$$

und schließlich

$$\left(A^{\mathrm{c}}\right)^{\mathrm{c}} = A.$$

Letzteres gilt z.B., weil

$$x \in A \Longleftrightarrow \neg(\neg x \in A) \Longleftrightarrow x \in \left(A^{\mathrm{c}}\right)^{\mathrm{c}}.$$

Die *Potenzmenge* einer Menge, d.h. die Menge aller Teilmengen einer Menge, bilden zusammen mit den Verknüpfungen $\cap, \cup, \neg$ eine sogenannte *Boolesche Algebra*.

Abbildungen

Seien nun S und T zwei Mengen. Die Schreibweise

$$f: S \to T, x \mapsto f(x)$$

bedeutet: Jedem Element $x \in S$ wird (genau) ein Element $f(x) \in T$ zugeordnet. Man nennt dann f eine Abbildung von S nach T. S heißt *Definitionsbereich* (engl. Source), T *Wertebereich* (engl. Target) der Abbildung f. Beispiel: $S = T = \mathbb{R}$, $x \mapsto f(x) = x^2 + 1$ oder $x \mapsto g(x) = x + 2$. Als besonders ausgezeichnete Typen von Abbildungen erwähnen wir die *injektiven*, *surjektiven* und *bijektiven*:

$$\begin{aligned}
f \text{ injektiv } &:\Longleftrightarrow \forall x_1 \in S\, \forall x_2 \in S (x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)),\\
f \text{ surjektiv } &:\Longleftrightarrow \forall y \in T\, \exists x \in S (f(x) = y)\\
f \text{ bijektiv } &:\Longleftrightarrow (f \text{ injektiv und surjektiv}).
\end{aligned}$$

Beispiel:

$$f: \mathbb{R} \to \mathbb{R}, x \mapsto x + 1$$

ist sowohl injektiv als auch surjektiv: ist nämlich $x_1 + 1 = f(x_1) = f(x_2) = x_2 + 1$, so folgt $x_1 = x_2$; ferner ist $y \in R$ das Bild von $x = y - 1$ unter f, denn $f(x) = f(y-1) = y - 1 + 1 = y$.

Beispiel: $S = T = \mathbb{R}$, $g(x) = 2x^3 + 9$. Hier ist g auch injektiv und surjektiv.

$$2x_1^3 + 9 = 2x_2^3 + 9 \Rightarrow 0 = x_1^3 - x_2^3 = (x_1 - x_2)\left(x_1^2 + x_1x_2 + x_2^2\right).$$

Wäre $x_1 \neq x_2$, so ist $x_1^2 + x_1 x_2 + x_2^2 = 0$. Sei etwa $x_2 \neq 0$ (Andernfalls vertausche x_1 und x_2). Dann folgt $\frac{x_1^2}{x_2^2} + \frac{x_1}{x_2} + 1 = 0$. Die quadratische Gleichung $x^2 + x + 1$ hat aber keine reelle Lösung. Das ist ein Widerspruch. Also ist $x_1 = x_2$. Dies beweist die Injektivität von g. Die Surjektivität kann noch nicht gezeigt werden, denn es fehlt uns noch ein Satz über die Existenz von Wurzeln, zu dessen Beweis wir noch ein Axiom der reellen Zahlen benötigen.

Beispiel: $S = T = R,\ h(x) = x^2$. h ist weder injektiv noch surjektiv.

Ist eine Abbildung $f\colon S \to T$ bijektiv, so gibt es zu jedem $y \in T$ genau ein $x \in S$ mit $f(x) = y$. Dieses x bezeichnet man mit $f^{-1}(y)$. Damit erhält man eine neue Abbildung

$$f^{-1}\colon T \to S,$$

die *Umkehrabbildung* von f.

Beispiel: $S = T - R, -f(x) = x+1$. Die Umkehrabbildung f^{-1} ist $y \mapsto y-1$.

Zum Verständnis der alten Literatur geben wir noch ein kleines Wörterbuch:

neu	alt
$f : S \to T$ surjektiv	f ist Abbildung von S auf T
$f : S \to T$ injektiv	f ist eine eindeutige Abbildung von S in$T$
$f : S \to T$ bijektiv	f ist eine eindeutige Abb. von S auf T

Sind $f : S \to T$ und $g : T \to U$ zwei beliebige Abbildungen, so definiert man, die *Hintereinanderschaltung* (oder das *Kompositum* von f und g durch

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

Diese Kompositionsregel ist assoziativ:

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

Mit dieser Verknüpfung bilden die bijektiven Abbildungen einer Menge S eine Gruppe, d.h. es gelten — abgesehen von der Kommutativität — die gleichen Axiome für '$\circ$' wie für die Verknüpfung '+' bei den reellen Zahlen; das neutrale Element ist die Identität $Id\colon S \to S, x \mapsto x$, das Gegenelement von f ist f^{-1}. Diese Gruppe heiß die *Permutationsgruppe von* S.

Die Fakultät

Ist S eine endliche Menge mit n Elementen, so bezeichnet $n!$ (sprich *'enn-fakultät'*) die Anzahl der bijektiven Abbildungen von S auf S.

Man kann sich leicht überlegen, daß

$$n! = 1 \cdot 2 \cdots n,$$

ist. Für die ersten Werte hat man folgende Tabelle

n	$n!$	n	$n!$
1	1	13	6227020800
2	2	14	87178291200
3	6	15	1307674368000
4	24	16	20922789888000
5	120	17	355687428096000
6	720	18	6402373705728000
7	5040	19	121645100408832000
8	40320	20	2432902008176640000
9	362880	21	51090942171709440000
10	3628800	22	1124000727777607680000
11	39916800	23	25852016738884976640000
12	479001600	24	620448401733239439360000

Oft hilfreich ist der folgende

Satz. *S sei endliche Menge. Dann ist jede injektive Abbildung $f\colon S \to S$ bijektiv.*

Beweis. $f : S \to S$ sei injektiv. Dann hat

$$f(S) := \{y \in S \mid \exists x \in S(f(x) = y)\} \subset S$$

auch n Elemente. Es muß also $f(S) = S$ gelten, d.h. f muß surjektiv sein.

Kombinatorik

Wir beenden dieses Kapitel mit einigen Grundbegriffen aus der Kombinatorik.

Binomialkoeffizienten

Sei S eine n–elementige Menge, z.B. $S = \{1, 2, \ldots, n\}$. Dann setzt man

$$\binom{n}{k} := \text{die Anzahl der } k\text{-elementigen Teilmengen von } S$$

(sprich: 'enn-über-ka'). $\binom{n}{k}$ heißt *Binomialkoeffizient*. Ist etwa S die Menge $\{1, 2, 3, 4\}$, also $n = 4$ und etwa $k = 2$, dann sind die 2–elementigen Teilmengen

$$\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}$$

Also ist

$$\binom{4}{2} = 6.$$

Natürlich hängt der Wert von $\binom{n}{k}$ nur von der Anzahl der Elemente von S ab.

Satz.

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

Beweis: Sei $S = \{1, 2, ..., n-1, n\}$. Eine k–elementige Teilmenge enthält entweder das Element n oder aber nicht. In letzterem Fall ist sie eine k—elementige Teilmenge von $S - \{n\}$; davon gibt es $\binom{n-1}{n}$ viele. Im anderen Fall ist $A - \{n\}$ eine $(k-1)$–elementige Teilmenge von $S - \{n\}$, und davon gibt es genau $\binom{n-1}{k-1}$.

Am besten berechnet man die Binomialkoeffizienten mittels der Rekursion des vorstehenden Satzes, wobei man die Ergebnisse der einzelnen Rekursionsschritte wie im folgenden Schema, dem *Pascalschen Dreieck*, notiert:

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
1 8 28 56 70 56 28 8 1
1 9 36 84 126 126 84 36 9 1
1 10 45 120 210 252 210 120 45 10 1
1 11 55 165 330 462 462 330 165 55 11 1
1 12 66 220 495 792 924 792 495 220 66 12 1
1 13 78 286 715 1287 1716 1716 1287 715 286 78 13 1
1 14 91 364 1001 2002 3003 3432 3003 2002 1001 364 91 14 1
1 15 105 455 1365 3003 5005 6435 6435 5005 3003 1365 455 105 15 1
1 16 120 560 1820 4368 8008 11440 12870 11440 8008 4368 1820 560 120 16 1

(jede Zahl ist die Summe der beiden links und rechts darüberstehenden).

Bekannt war dieses Dreieck schon den Arabern im 13. Jahrhundert, weiter studiert haben es insbesonders Stiefel (1544) und Pascal (1659).

Die Berechnung der Binomialkoeffizienten kann allerdings auch nach der folgenden Formel erfolgen:

Satz.

$$\binom{n}{k} = \frac{n\cdot(n-1)\cdots(n-k+1)}{1\cdot 2\cdot 3\cdots k} = \frac{n!}{k!\,(n-k)!}$$

.

Der Beweis kann leicht mittels vollständiger Induktion geführt werden.

Die Bedeutung der Binomialkoeffizienten liegt für uns in folgendem Satz:

Satz.

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

Es ist zum Beispiel

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4.$$

Hierbei sind a, b etwa relle Zahlen. Der binomische Lehrsatz gilt aber auch in jedem kommutativen Ring! Man kann die Richtigkeit des Satz leicht durch vollständige Induktion beweisen — oder aber auch "sehen":

$$(a+b)^n = (a+b)(a+b)\cdots(a+b)$$

ist vollständig auszudistributieren; für jedes $k = 0, 1, 2, ...n$ hat man aus obigen n–vielen a's und n–vielen b'-s Produkte von n Faktoren zu bilden, k davon werden aus den a's genommen und $n-k$ von den b's. Die Auswahl der a's kann auf genau $\binom{n}{k}$ Weisen erfolgen, die Wahl der b's liegt dann schon fest, da aus jeder Klammer nur ein Summand als Faktor zu nehmen ist. Man hat also $\binom{n}{k}$ Produkte $a^k b^{n-k}$, und das ist genau der Koeffizient vor $a^k b^{n-k}$ in der behaupteten Formel.

Für $a = b = 1$ folgt aus dem binomischen Lehrsatz

$$\sum_{i=0}^{n} \binom{n}{i} = 2^n,$$

also nach der Definition der Binomialkoeffizienten:

Die Anzahl aller Teilmengen von $\{1, 2, ...n\}$ ist 2^n.

Schreibt man $\mathcal{P}(X)$ für die Potenzmenge von X und $|X|$ für die Anzahl der Elemente von X, so kann man die letztere Aussage auch schreiben als

$$\mathcal{P}(X) = 2^{|X|}.$$

Seien X, Y Mengen. Dann bezeichnet X^Y die Menge aller Abbildungen von Y in X.

Satz. *Sind X und Y endlich, so ist*

$$\left|X^Y\right| = |X|^{|Y|}$$

Den Beweis lassen wir als Übungsaufgabe. Insbesondere ist also 2^n = Anzahl aller Abbildungen $\{1, 2, ...n\} \to \{0, 1\}$. Die Abbildungen

$$\{1, 2, ..., n\} \to \{0, 1\}$$

entsprechen in eindeutiger Weise den Teilmengen von $\{1, 2, ..., n\}$, indem man einer Abbildung f einfach die Menge $f^{-1}(\{1\}) := \{k \mid f(k) = 1\}$ zuordnet. Der Satz ist in diesem Sinne also eine Verallgemeinerung des davor ausgesprochenen Resultats.

Satz. *Sei X eine Menge. Dann gibt es keine Bijektion zwischen X und $\mathcal{P}(X)$.*

Man bemerke insbesondere, daß $\mathcal{P}(\emptyset) \neq \emptyset$!

Beweis. Angenommen, es gibt eine bijektive Abbildung $g : X \to \mathcal{P}(X)$. Dann betrachte man

$$A := \{y \in X \mid \neg(y \in g(y))\}.$$

Wegen der Surjektivität von g gibt es dann $x_0 \in X$, s.d. $g(x_0) = A$. Es ist dann $x_0 \in A$ genau dann, wenn $\neg(x_0 \in g(x_0))$, wegen $A = g(x_0)$ hat man also

$$x_0 \in g(x_0) \Longleftrightarrow \neg(x_0 \in g(x_0)).$$

Das ist absurd.

Als Folgerung erhält man insbesondere.

$$n < 2^n$$

für jedes $n \in \mathbb{N}$.

5 DAS VOLLSTÄNDIGKEITSAXIOM

Bisher haben wir die Axiome eines angeordneten Körpers (Ai), (Mi), (D) und (Oj) ($i = 1, ..., 4$ und $j = 1, 2$) betrachtet. Allerdings werden diese auch von $\mathbb{Q}$ erfüllt, genügen also nicht, $\mathbb{R}$ vollständig zu charakterisieren. Es stellt sich heraus, daß genau ein weiteres Axiom ausreicht, um die reellen Zahlen von allen anderen angeordneten Körpern zu unterscheiden. Doch einige Bezeichnungen zuvor: sei S eine Teilmenge von $\mathbb{R}$.

Nach oben beschränkte Mengen

S heißt nach oben beschränkt, falls es ein $c \in \mathbb{R}$ gibt, sodaß alle $x \in S$ kleiner oder gleich c sind.

Oberee Schranke

Jedes solche c heißt obere Schranke von S.

Schließlich erklären wir noch:

Supremum

Eine Zahl s heißt kleinste obere Schranke oder Supremum von S , falls s obere Schranke von S ist, und für jede weitere obere Schranke c von S stets $s \leq c$ ist.

Nun können wir das letzte Axiom für die reellen Zahlen formulieren, das sogenannte Vollständigkeitsaxiom:

Vollständigkeitsaxiom

(VA) Jede nichtleere, nach oben beschränkte Teilmenge S von $\mathbb{R}$ besitzt eine kleinste obere Schranke.

Es ist leicht zu sehen, daß eine Menge höchstens ein Supremum besitzt, d.h. das Supremum einer Menge S ist eindeutig bestimmt: man schreibt dafür $\sup S$. Man beachte, daß jedes $s \in \mathbb{R}$ eine obere Schranke von $\emptyset$ ist; man setzt daher

$$\sup \emptyset = -\infty.$$

Schließlich setzt man für eine nicht nach oben beschränkte Menge S noch

$$\sup S = \infty.$$

Diese beiden letzten Setzungen haben hier keinen tieferen Sinn; allerdings erhalten viele Formeln, in denen das Supremum auftritt, damit eine naheliegende Interpretation auch in den Fällen, in denen das Supremum keine reelle Zahl bedeutet.

In $\mathbb{Q}$ gilt das Vollständigkeitsaxiom nicht. Beispiel: $M = \{x \in \mathbb{Q} \,|\, x^2 \leq 2\}$ ist in $\mathbb{Q}$ wie in $\mathbb{R}$ nach oben beschränkt, etwa durch 10^7. Man kann sich überlegen, daß ein Supremum s die Eigenschaft $s^2 = 2$ hat. Da $\sqrt{2}$ irrational ist, hat M in $\mathbb{Q}$ also kein Supremum; nach dem Vollständigkeitsaxiom besitzt M aber sehr wohl eine kleinste obere Schranke. Das Vollständigkeitsaxiom impliziert also insbesondere die Existenz der $\sqrt{2}$ und unterscheidet damit den Bereich der reellen Zahlen von dem der rationalen.

Wie schon erwähnt, kann man ganz allgemein zeigen, daß $\mathbb{R}$ tatsächlich mittels des Vollständigkeitsaxioms eindeutig bestimmt ist; genauer

Satz. *Gegeben seien zwei angeordnete Körper K (mit den Operationen $+, \cdot$ und dem Positivitätsbereich P) und K' (mit den Operationen $+', \cdot'$ und dem Positivitätsbereich P'), die das Vollständigkeitsaxiom erfüllen. Dann gibt es eine Bijektion*

$$\iota : K \to K',$$

sodaß für alle $x, y \in K$:

$$\iota(x + y) = \iota(x) +' \iota(y), \quad \iota(x \cdot y) = \iota(x) \cdot' \iota(y), \quad x \in P \Longleftrightarrow \iota(x) \in P'.$$

Wir werden diesen Satz hier nicht beweisen. Einen Beweis und weitere Erläuterungen findet man zum Beispiel in [M.Spivac: Calculus].

Ganz analog wie oben definiert man "*nach unten beschränkt*" und *größte untere Schranke* bzw. das Infimum einem Menge S — in Zeichen $\inf S$ —. Es gilt nun in $\mathbb{R}$ der

Satz. *Jede nichtleere, nach unten beschränkte Teilmenge S von $\mathbb{R}$ hat ein Infimum.*

.

Beweis. Sei c eine untere Schranke von S; die Menge $-S = \{-s \mid s \in S\}$ ist dann durch $-c$ nach oben beschränkt, besitzt also nach dem Axiom (VA) ein Supremum s. Dann ist aber $-s$ das Infimum von S; zunächst ist jedenfalls $-s$ untere Schranke von S, denn

$$x \in S \Rightarrow -x \in -S \Rightarrow -x \leq s \Rightarrow x \geq -s;$$

es ist aber auch die größte untere Schranke, denn ist c untere Schranke von S, so ist $-c$ obere Schranke von $-S$, also $-c \geq -s$, also $c \leq s$.

Wir wenden uns nun der grundlegendsten Folgerung aus dem (VA)–Axiom:

Satz des Archimedes

Satz des Archimedes. *Sei $d > 0$ und $M \in \mathbb{R}$. Dann gibt es ein $n \in \mathbb{N}$, sodaß $n \cdot d > M$.*

Man kann zeigen, daß dieser Satz noch nicht aus den Axiomen eines angeordneten Körpers folgt, d.h. es gibt angeordnete Körper, in denen der eben ausgesprochene Satz nicht gilt.

Beweis. Gäbe es nicht solch ein n wie im Satz, so wäre $S = \{n \cdot d \,|\, n \in \mathbb{N}\}$ durch M nach oben beschränkt. Sei $s = \sup S$. Dann ist $s - d < s$, also ist $s - d$ keine obere Schranke. Also gibt es ein Element $m \cdot d \in S$ s.d.

$$s - d < m \cdot d \leq s.$$

Dann folgt aber $(m+1)d > s$, also wäre s keine obere Schranke. Das ist ein Widerspruch.

Folgerung. *Zu jedem $\epsilon > 0$ gibt es ein $n \in \mathbb{N}\backslash\{0\}$, s.d. $\frac{1}{n} < \epsilon$.*

Wie klein $\epsilon > 0$ auch ist, die Kehrwerte der positiven natürlichen Zahlen werden kleiner, und je größer n desto kleiner $\frac{1}{n}$!

Beweis. Wegen

$$\epsilon > \frac{1}{n} \Longleftrightarrow n > \frac{1}{\epsilon}$$

ist klar, daß die Behauptung aus dem Satz des Archimedes mit $d = 1$ folgt.

Insbesondere folgern wir aus dem letzten Satz:

$$\forall \epsilon > 0 \, \exists n_0 \in \mathbb{N} \left(n > n_0 \Rightarrow \frac{1}{n} < \epsilon \right)$$

Das ist das erste Beispiel einer *Nullfolge*.

Folgen und Nullfolgen

Unter einer Folge $(a_n)_{n \in \mathbb{N}}$ von reellen Zahlen versteht man eine Abbildung $\mathbb{N} \to \mathbb{R}, n \mapsto a_n$. Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt Nullfolge, falls gilt:

$$\forall \epsilon > 0 \, \exists n_0 \in \mathbb{N} \, \forall n \in \mathbb{N} (n > n_0 \Rightarrow |a_n| < \epsilon).$$

Anschaulich verbirgt sich hinter dem Begriff der Nullfolge, daß sich die Glieder a_n der Folge mit wachsendem n der 0 immer mehr nähern.

Ist $A(n)$ eine Aussage über natürliche Zahlen n, so sagt man statt

$$\exists n_0 \in \mathbb{N} (n > n_0 \to A(n))$$

auch

“für fast alle n ist $A(n)$ wahr”.

Damit kann man die Definition der Nullfolge auch faßlicher in der folgenden Form ausprechen:

Für jedes noch so kleine $\epsilon > 0$ ist für fast alle n noch $|a_n| < \epsilon$

Beispiele für Nullfolgen sind die oben schon erwähnte Folge $n \mapsto \frac{1}{n}$. Ein weiteres Beispiel ist $n \mapsto \frac{1}{n^2}$: für jedes $n \geq 1$ ist $n^2 \geq n$, also $\frac{1}{n^2} \leq \frac{1}{n}$. Danach ist klar, daß mit $\left(\frac{1}{n}\right)_n$ auch $\left(\frac{1}{n^2}\right)_n$ eine Nullfolge ist. Diese Beispiele haben noch etwas Weiteres gemeinsam:

Monotonie

Eine Folge(a_n) heißt *monoton steigend (fallend,streng monoton steigend, streng monoton fallend)*, falls

$$a_n \ \leq \ (\geq, \ <, \ >) \ a_{n+1}$$

für alle $n \in \mathbb{N}$ gilt.

Nullfolgen brauchen keineswegs monoton zu fallen, z.B. ist

$$1, 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{3}, \frac{1}{9}, \frac{1}{4}, \frac{1}{16} ...$$

eine nicht monotone Nullfolge. Daß dies tatsächlich eine Nullfolge ist, folgt aus dem

Lemma. *Sind die Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ Nullfolgen, dann ist auch die Folge $a_0, b_0, a_1, b_1, a_2, b_2, ...$ eine Nullfolge.*

Beweis. Ist $|a_n| < \epsilon$ für $n > n_0$ und $|b_n| < \epsilon$ für $n > n_1$, so sind die Glieder $|c_n|$ der zu betrachtenden Folge kleiner als ϵ für $n > \max(n_0, n_1)$

Satz. *Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ Nullfolgen. Dann gilt:*

1. *die Folge $(a_n + b_n)_{n\in\mathbb{N}}$ ist eine Nullfolge;*
2. *für jedes $c \in \mathbb{R}$ ist $(ca_n)_{n\in\mathbb{N}}$ eine Nullfolge;*
3. *ist $(c_n)_{n\in\mathbb{N}}$ eine beschränkte Folge, so ist $(a_n \cdot c_n)_{n\in\mathbb{N}}$ eine Nullfolge, insbesondere ist*
4. *$(a_n \cdot b_n)_{n\in\mathbb{N}}$ eine Nullfolge.*

Hierbei haben wir noch den Begriff der monotonen Folge zu erklären:

Beschränkte Folgen

Eine Folge (a_n) heißt *beschränkt* falls ein $M \in \mathbb{R}$ existiert sodaß $a_n < M$ für all $n \in \mathbb{N}$ gilt.

Beweis. Es sei $\epsilon > 0$ gegeben. Dann gibt es n_0 und $m_0 \in \mathbb{N}$, sodaß $|a_n| < \epsilon$ für $n > n_0$ und $|b_n| < \epsilon$ für $n > m_0$. Also ist für $n > \max(n_0, m_0)$:

$$|a_n + b_n| \leq |a_n| + |b_n| < 2\epsilon.$$

Damit ist die Behauptung 1. bewiesen. Die Aussagen 2. und 3. folgen ganz analog jeweils mit den Ungleichungen:

$$|ca_n| = |c| \, |a_n| < |c| \cdot \epsilon$$
$$|a_n c_n| = |a_n| \, |c_n| < |a_n| \cdot M < M \cdot \epsilon,$$

und 4. folgt mit der Tatsache, daß Nullfolgen beschränkt sind.

Satz. *$(q^n)_{n\in\mathbb{N}}$ ist Nullfolge $\Longleftrightarrow |q| < 1$.*

Beweis. *Ist $|q| > 1$, so ist $|q| = 1 + h$ mit einem geeigneten $h > 0$. Also ist $|q^n| = |q|^n = (1+h)^n \geq 1 + n \cdot h$, letzteres mit einer einfachen Abschätzung mittels des binomischen Lehrsatzes. Aber $1 + h \cdot n$ wächst über alle Grenzen für $n \to \infty$, also ist $|q^n|$ unbeschränkt, also keine Nullfolge. Für $|q| = 1$ erhält man die Folgen $1, 1, ...$ und $1, -1, 1...$, jedenfalls keine Nullfolgen. Schließlich sei $|q| < 1$. Dann ist $\left|\frac{1}{q}\right| > 1$. Daher ist $\frac{1}{q} = 1 + h$ mit einem geeigneten $h > 0$. Für gegebenes $\epsilon > 0$ findet man dann wie oben ein n_0, sodaß für $n > n_0$ stets $\left|\frac{1}{q^n}\right| > 1 + n \cdot h > \frac{1}{\epsilon}$, also $|q^n| < \epsilon$ gilt. Also ist (q^n) tatsächlich eine Nullfolge.*

Konvergente Folgen

Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt konvergent gegen $a \in \mathbb{R}$ — in Zeichen $\lim_{n\to\infty} a_n = a$ — falls $(a_n - a)_{n\in\mathbb{N}}$ eine Nullfolge ist

Satz. *Konvergiert (a_n) gegen a und b, so ist $a = b$ (d.h. der Limes ist eindeutig bestimmt).*

Beweis. Es ist ja $|a - b| = |a - a_n - (b - a_n)| \leq |a - a_n| + |b - a_n| < 2\epsilon$ für genügend großes n. Also: $0 \leq |a - b| < \epsilon$ für jedes $\epsilon > 0$, was nur für $|a - b| = 0$ möglich ist.

Fundamental ist der folgende

Satz. *Jede monoton steigende (fallende), nach oben (unten) beschränkte Folge ist konvergent.*

Beweis. Wir betrachten den Fall einer monoton steigenden Folge. Die Menge $S = \{a_n \mid n \in \mathbb{N}\}$ ist nicht leer und nach oben beschränkt, hat also nach Axiom (VA) eine obere Grenze a. Wir behaupten, daß $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert: Für gegebenes $\epsilon > 0$ ist $a - \epsilon < a$, also $a - \epsilon$ keine obere Schranke. Also gibt es $n_0 \in \mathbb{N}$, sodaß $a - \epsilon < a_{n_0} \leq a$. Wegen der Monotonie ist für $n > n_0$ dann aber $a - \epsilon < a_n \leq a$, also $|a_n - a| < \epsilon$, und dies war zu zeigen. Den Fall einer monoton fallenden Folge behandelt man völlig analog, indem man im eben gegebenen Beweis sup durch inf ersetzt.

Wir geben einige Anwendungen des letzten Satzes. Jeder Dezimalbruch

$$a_{-e} \cdots a_{-1}.a_0 a_1 \cdots \qquad (e \in \mathbb{Z}, \mathrm{e} > 0, \forall n(a_n \in \{0,1,2,3,4,5,6,7,8,9\}))$$

stellt eine reelle Zahl dar: den Grenzwert x der Folge (S_n), wo

$$S_n = \sum_{k=-e}^{n} a_k \cdot 10^{-k}$$

ist. Daß dieser Grenzwert stets existiert, folgt mittels des letzten Satzes: Die Folge (S_n) ist monoton steigend und beschränkt. Letzteres folgt mit

$$S_n \leq \sum_{k=-e}^{0} a_k + \sum_{k=0}^{n} 9 \cdot 10^{-k} = \sum_{k=-e}^{n} a_k + \left(1 - 10^{-n}\right) \leq \sum_{k=-e}^{n} a_k + 1.$$

Umgekehrt kann gezeigt werden, daß man *jede* reelle Zahl so erhalten kann, d.h. daß sie als Dezimalbruch geschrieben werden kann.

Eine alte Aufgabe von Jakob dem Bernoulli (1654-1705) lautet

> Quaeritur: si creditor aliquis pecuniae summa foenori exponet, ea lege, ut singulis momentis pars proportionalis usurae amme sorti ammeretur; quantum ipsi finito anno debeatur.

Problem der stetigen Verzinsung

Also: Ein Kapital A wird nach einem Jahr mit 100 % verzinst. Nach einem Jahr ist also das Kapital $2A$. Schlägt man die Zinsen schon nach einem halben Jahr hinzu, so beträgt das Kapital nach einem Jahr

$$A \cdot \left(1 + \frac{1}{2}\right)\left(1 + \frac{1}{2}\right).$$

Schlägt man die Zinsen dritteljährlich hinzu, so ist das Kapital am Ende des Jahres

$$A\left(1 + \frac{1}{3}\right)\left(1 + \frac{1}{3}\right)\left(1 + \frac{1}{3}\right).$$

Teilt man das Jahr in n gleiche Teile, so liefert der Prozeß ein Endkapital von $A\left(1 + \frac{1}{n}\right)^n$. Was geschieht mit wachsendem n, d.h. bei stetiger Verzinsung? Die Antwort ist folgendermaßen: Es ist

$$\begin{aligned}\left(1 + \frac{1}{n}\right)^n &= \sum_{k=0}^{n} \frac{n(n-1)...(n-k+1)}{k!} \cdot \frac{1}{n^k} \\ &= \sum_{k=0}^{n} \frac{1\left(1 - \frac{1}{n}\right)...\left(1 - \frac{k-1}{n}\right)}{k!}.\end{aligned}$$

Aus dieser Darstellung liest man sofort ab, daß die Folge $\left(1 + \frac{1}{n}\right)^n$ monoton steigend ist. Diese Darstellung zeigt aber noch mehr, nämlich daß die Folge beschränkt ist:

$$\begin{aligned}&\sum_{k=0}^{n} \frac{1\left(1 - \frac{1}{n}\right)...\left(1 - \frac{k-1}{n}\right)}{k!} \\ \leq &\sum_{k=0}^{n} \frac{1}{k!} = \frac{1}{1} + \frac{1}{1} + \frac{1}{2} + \frac{1}{2 \cdot 3} + \frac{1}{2 \cdot 3 \cdot 4} + ... \frac{1}{n!} \\ &\leq 1 + 1 + \frac{1}{2} + \frac{1}{4} + ... + \frac{1}{2^{n-1}} \\ &= 1 + \frac{1 - \frac{1}{2}^n}{1 - \frac{1}{2}} < 3.\end{aligned}$$

Also ist die Folge nach dem letzten Satz konvergent. Man nennt den Limes die Eulerschesche Zahl — als Symbol e —, also

Eulersche Zahl e

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

Die ersten Dezimalstellen sind

$$\begin{aligned}\mathrm{e} = 2.&718281828459045235360287471352662497757247093 6\\ &9995957496696762772407663035354759457138217852 5\\ &1664274274663919320030599218174135966290435729 0\\ &0334295260595630738132328627943490763233829880 7\\ &5319525101901157383418793070215408914993488416 7\\ &5092447614606680822648001684774118537423454424 3\\ &7107539077744992069551702761838606261331384583 0\\ &0075204493382656029760673711320070932870912744 3\\ &7470472306969772093101416928368190255151086574 6\\ &3772111252389784425056953696770785449969967946 8\\ &6445490598793163688923009879312773617821542499 9\\ &2295763514822082698951936680331825288693984964 6\\ &5105820939239829488793320362509443117301238197 0\\ &6841614039701983767932068328237646480429531180 2\\ &3287825098194558153017567173613320698112509961 8\\ &1881593041690351598888519345807273866738589422 8\\ &7922849989208680582574927961048419844436346324 4\\ &9684875602336248270419786232090021609902353043 6\\ &9941849146314093431738143640546253152096183690 8\\ &8870701676839642437814059271456354906130310720 8\\ &5103837505101157477041718986106873969655212671 5\\ &\hfill 4688957035035396\ldots.\end{aligned}$$

Dies also die Antwort auf J. Bernoullis' Frage: bei stetiger Verzinsung mit Zinsfuß 100% hat sich nach einem Jahr das Grundkapital ver-e–facht.

Euler (1707-1783) kannte schon eine Kettenbruchentwicklung für e, die im Gegensatz zur Dezimalentwicklung eine sehr einfache Regelmässigkeit besitzt:

$$\mathrm{e} = 2 + \cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{4+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{6+\cfrac{1}{1+\cfrac{1}{1+\cdots}}}}}}}}}}$$

$$= [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10...].$$

Kettenbrüche

Hierzu einige allgemeine Bemerkungen über Kettenbrüche: seien b_0 eine irgendeine ganze Zahl und irgendwelche $b_1, b_2, b_3, \ldots > 0$ natürliche Zahlen, dann kann man dazu die Folge (z_n) der Kettenbrüche

$$z_n = [b_0, b_1 ... b_{n-1}] = b_0 + \cfrac{1}{b_1 + \cfrac{1}{b_2 + \cfrac{1}{b_3 + \cdots + \cfrac{1}{b_n}}}}$$

betrachten. Es ist $z_n \in \mathbb{Q}$; wir können also

$$z_n = \frac{p_n}{q_n}, \qquad \text{mit teilerfremden, positiven ganzen Zahlen } p_n, q_n$$

setzen.

Durch vollständige Induktion zeigt man leicht die Formeln

$$\begin{aligned} p_0 = 1, p_1 = b_0,\ p_{n+1} &= b_n p_n + p_{n-1} \\ q_0 = 0; q_1 = 1,\ q_{n+1} &= b_n q_n + q_{n-1}\ (n \geq 1). \end{aligned}$$

Weiter kann man zeigen, daß die Folge $(z_{2n+1})_{n\in\mathbb{N}}$ monoton steigend und $(z_{2n})_{n\in\mathbb{N}}$ monoton fallend ist. Außerdem gilt

$$(IV1) \qquad z_1 \leq z_3 \leq z_5 \leq \ldots \leq z_6 \leq z_4 \leq z_2$$

(es gilt hier sogar “$<$” statt “$\leq$”); insbesondere ist also die Folge $(z_{2n+1})_n$ monoton steigend und nach oben beschränkt, und die Folge $(z_{2n})_n$ ist monoton fallend und nach unten beschränkt. Also sind beide Folgen konvergent — etwa gegen z und z'. Weiter ist nach der oben gegebenen Rekursion für die p_n, q_n

$$p_{n+1}q_n - q_{n+1}p_n = p_{n-1}q_n - q_{n-1}p_n.$$

Hieraus folgt leicht durch vollständige Induktion

$$|p_{n+1}q_n - p_n q_{n+1}| = 1$$

für alle $n \in \mathbb{N}$. Dies bedeutet aber

$$\left| \frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} \right| = \frac{1}{q_n q_{n+1}},$$

woraus endlich

$$(IV2) \qquad \lim_{n\to\infty} |z_{n+1} - z_n| = 0.$$

folgt. Es ist also $z = z'$, und wir sehen daß die Folge (z_n) stets konvergent ist. Den Grenzwert z bezeichnet man mit

$$z = [b_0, b_1, b_2, b_3, \ldots],$$

und in diesem Sinn ist die oben gegebene Kettenbruchdarstellung für e zu verstehen. Die Folge (z_n) ist ein Beispiel für einen allgemeineren Begriff

Intervallschachtelung

> Eine Folge $(z_n)_{n\in\mathbb{N}}$ mit den Eigenschaften (IV1) und (IV2) heißt *Intervallschachtelung*.

Eine Intervallschachtelung ist (nach den gleichen Argumenten wie oben) offenbar stets konvergent.

Wir geben noch ein Beispiel für einen Kettenbruch, nämlich den schon in Kapitel 1 betrachteten:

$$[1, 1, 1, 1, 1, \ldots].$$

Hier ist also $b_0 = b_1 = \ldots = 1$, und damit

$$p_0 = 1, p_1 = 2, p_{n+1} = p_n + p_{n-1},$$
$$q_0 = 1, q_1 = 1 q_{n+1} = q_n + q_{n-1}.$$

Also ist $p_n = q_{n+1}$ und die ersten Glieder der Folge q_n sind

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89 \ldots$$

Fibonacci Zahlen

Dies sind die berühmtenFibonacci Zahlen (Fibonacci — alias Leonardo von Pisa, 1180-1250) Was ist der Limes der Folge (z_n)? Es gilt offenbar $1 + \frac{1}{z_{n-1}} = z_n$ für $n \in \mathbb{N}$. Setzt man also

$$z = \lim_{n\to\infty} z_n,$$

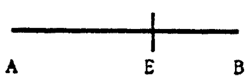
so folgt aus dem Hauptsatz für konvergente Folgen

$$1 + \frac{1}{z} = z,$$

Goldener Schnitt

da auch $\lim_{n\to\infty} z_{n-1} = z$. Also gilt Es folgt $z = \frac{1+\sqrt{5}}{2}$. Diese Zahl z hängt mit einem berühmten Problem bei Euklid zusammen, dem Problem des *goldenen Schnitts*:

> Man finde den Punkt auf der Strecke von A nach B, sodaß
>
> $$\frac{\overline{AB}}{\overline{AE}} = \frac{\overline{AE}}{\overline{BE}}.$$



Setzt man
$$u = \frac{\overline{AE}}{\overline{AB}},$$
so folgt
$$u^2 + u - 1 = 0,$$
d.h.
$$u = \frac{1}{z} = \frac{\sqrt{5}-1}{2}.$$
Diese Zahl hängt auch mit dem Problem der Konstruktion des regelmäßigen Fünfecks zusammen. Es ist nämlich gerade
$$\cos\frac{360^o}{5} = \frac{\sqrt{5}-1}{4} = \frac{1}{u}.$$
Als Übungsaufgabe vergleiche man den die ersten Näherungen $\frac{q_n}{2p_n}$ von $\frac{\sqrt{5}-1}{2}$ mit irgendwelchen anderswo gegebenen Werten.

Wir kehren zurück zur allgemeinen Theorie der konvergenten Folgen.

Satz. *Jede konvergente Folge ist beschränkt.*

Beweis. Sei $\lim_{n\to\infty} a_n = a$. Zu $\epsilon > 0$ gibt es ein n_0, sodaß für $n > n_0$ stets $|a_n - a| < \epsilon$, also $a - \epsilon < a_n < a + \epsilon$ gilt. Also liegen fast alle a_n in dieser $\epsilon-$ Umgebung von a. Eine Schranke für alle $|a_n|$ ist daher $|a_n| \le \max\left(|a+\epsilon|, |a-\epsilon|, |a_0|, |a_1|, \ldots, |a_{n_0}|\right)$.

Hauptsatz für konvergente Folgen

Satz. *Seien* $\lim_{n\to\infty} a_n = a$ *und* $\lim_{n\to\infty} b_n = b$. *Dann gilt:*

1. $\lim_{n\to\infty} a_n + b_n = a + b$
2. $\lim_{n\to\infty} a_n \cdot b_n = a \cdot b$,
3. $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$, *wenn alle* $b_n \neq 0$ *sind und* $b \neq 0$ *ist.*

Beweis. 1. Zu zeigen ist, daß $((a+b) - (a_n + b_n))$ eine Nullfolge ist; dies folgt aber mit
$$(a+b) - (a_n + b_n) = (a - a_n) + (b - b_n)$$
sofort aus dem Hauptsatz für Nullfolgen. 2. Hier schreibt man
$$ab - a_n b_n = a(b - b_n) + b_n(a - a_n).$$
Nun sind nach Voraussetzung $(b - b_n)$ und $(a - a_n)$ Nullfolgen, und es ist die Folge (b_n) als konvergente Folge beschränkt. Also ist nach dem Hauptsatz

über Nullfolgen $(ab - a_n b_n)$ eine Nullfolge. Ganz analog folgt 3. mittels der Darstellung

$$\frac{1}{b} - \frac{1}{b_n} = \frac{b_n - b}{b \cdot b_n}.$$

und der Tatsache daß $\frac{1}{b_n}$ eine beschränkte Folge ist. Letzteres folgt, weil ja $|b_n - b| < \frac{|b|}{2}$ für fast alle n, d.h. $|b_n| > \frac{|b|}{2}$, d.h. $\frac{1}{|b_n|} < \frac{2}{|b|}$ für fast alle n.

Teilfolgen

> Unter einer Teilfolge einer Folge $\{a_n\}_{n\in\mathbb{N}}$ versteht man eine Folge der Gestalt $\{a_{k_n}\}_{n\in\mathbb{N}}$, wo $\{k_n\}_{n\in\mathbb{N}}$ eine streng monoton steigende Folge von natürlichen Zahlen bedeutet.

Der folgende Satz ist fast offensichtlich.

Satz. *Jede Teilfolge einer gegen eine Zahl a konvergenten Folge ist konvergent, und zwar ebenfalls gegen a.*

Beweis. Zu gegebenen $\epsilon > 0$ ist für fast alle n stets $|a - a_n| < \epsilon$; also ist auch $|a - a_{k_n}| < \epsilon$, da ja $k_n \geq n$.

6 UNENDLICHE REIHEN

Wir beginnen mit einer Definition. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen.

Unendliche Reihen

Die Folge $(s_k)_{k\in\mathbb{N}}$ mit

$$s_k = \sum_{n=0}^{k} a_n$$

heißt *Partialsummenfolge* der *unendlichen Reihe*

$$\sum_{n=0}^{\infty} a_n.$$

Die a_n heißen die *Glieder der Reihe*.

Ist die unendliche Reihe $\sum_{n=0}^{\infty} a_n$ konvergent, d.h. ist die Partialsummenfolge $(s_k)_{k\in\mathbb{N}}$ konvergent, so bezeichnet man auch den *Wert der Reihe*, d.h. den Grenzwert der Partialsummenfolge, mit dem Symbol $\displaystyle\sum_{n=0}^{\infty} a_n$. Somit kann dieses Symbol zweierlei Bedeutung habe: eine Folge oder ihren Grenzwert; die jeweilige Bedeutung erschließt sich jeweils aus dem Zusammenhang.

Geometrische Reihe

Hier ist das vielleicht grundlegendste Beispiel einer unendlichen Reihe, die *geometrische Reihe*:

$$\sum_{n=0}^{\infty} q^n.$$

Satz. *Die vorstehende geometrische Reihe ist genau dann konvergent, wenn $|q| < 1$. Im Fall der Konvergenz ist ihr Wert $\frac{1}{1-q}$.*

Beweis. Ist s_k das k–te Glied der Partialsummenfolge der geometrischen Reihe, so hat man

$$s_k = 1 + q + \ldots + q^k = \frac{1-q^{k+1}}{1-q} \quad \text{für } q \neq 1$$
$$s_k = k+1 \quad \text{für } q = 1.$$

Die geometrische Reihe ist also genau dann konvergent, falls q^{k+1} konvergiert. Nach einem im letzten Kapitel bewiesenen Satz ist dies genau dann der Fall, falls $|q| < 1$ ist, und der Grenzwert der Partialsummenfolge hat dann den behaupteten Wert.

Als Illustration zu diesem Satz hat man etwa:

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 1 + \frac{1}{2} + \frac{1}{4} + \ldots = 2$$
$$\sum_{n=0}^{\infty} \left(-\frac{1}{2}\right)^n = 1 - \frac{1}{2} + \frac{1}{4} \pm \ldots = \frac{2}{3}.$$

Wir notieren nun einige einfache Eigenschaften der unendliche Reihen.

Satz. *Sei* $\sum_{n=0}^{\infty} a_n$ *konvergent und* $\sum_{n=0}^{\infty} b_n$ *konvergent. Dann ist auch* $\sum_{n=0}^{\infty} (a_n \pm b_n)$ *konvergent, und es gilt*

$$\sum_{n=0}^{\infty} (a_n \pm b_n) = \sum_{n=0}^{\infty} a_n \pm \sum_{n=0}^{\infty} b_n.$$

Beweis. Sei $s_k = \sum_{n=0}^{k} a_n$, $t_k = \sum_{n=0}^{k} b_n$, $s = \lim_{k\to\infty} s_k$ und $s = \lim_{k\to\infty} s_k$. Aus dem Assoziativ- und Kommutativgesetz der Addition folgt

$$s_k \pm t_k = \sum_{n=0}^{k} (a_n \pm b_n) =: r_k.$$

Der Hauptsatz für konvergente Folgen liefert die Konvergenz von $(r_k)_{k\in\mathbb{N}}$, und zwar gegen $s \pm t$. Dies ist die Behauptung.

Ganz analog zeigt man:

Satz. *Sei* $\sum_{n=0}^{\infty} a_n$ *konvergent und* $c \in \mathbb{R}$. *Dann ist auch* $\sum_{n=0}^{\infty} c \cdot a_n$ *konvergent und*

$$\sum_{n=0}^{\infty} c \cdot a_n = c \cdot \sum_{n=0}^{\infty} a_n.$$

Als weitere einfache Folgerungen aus der Theorie der Folgen kann man sich leicht die folgenden Tatsachen überlegen:

> Man darf endlich viele Glieder vertauschen, ohne daß sich am Konvergenzverhalten oder Limes etwas ändert

(denn die Partialsummenfolge bleibt bis auf endlich viele Glieder dieselbe).

> In einer Reihe darf man endlich viele Glieder weglassen, ohne daß sich am Konvergenzverhalten etwas ändert.

(Der Limes der neuen Reihe ist gleich dem Limes der alten Reihe minus der Summe der weggelassenen Glieder, sofern letzterer Limes existiert).

Ist $\sum_{n=0}^{\infty} a_n$ konvergent, so ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge

(denn $a_k = s_k - s_{k-1}$, also $\lim a_k = \lim s_k - \lim s_{k-1} = 0$). Man beachte aber, daß die Umkehrung des letzten Sachverhalts nicht gilt: die Folge mit den Gliedern $\frac{1}{n}$ ist eine Nullfolge, aber $\sum_{n=0}^{\infty} \frac{1}{n}$ ist divergent, wie wir unten sehen werden.

Wir kommen nun zu der eigentlichen Theorie der unendlichen Reihen. Den zentralen Punkt dieser Theorie bilden die *Konvergenzkriterien.* Wir beginnen mit einer Zusammenstellung solcher Konvergenzkriterien für Reihen mit nichtnegativen Gliedern.

Reihen mit nichtnegativen Gliedern

Satz. *Sei* $\sum_{n=0}^{\infty} a_n$ *eine Reihe mit nichtnegativen Gliedern, d.h. es gelte* $a_n \geq 0$ *für alle* $n \in \mathbb{N}$*. Dann ist* $\sum_{n=0}^{\infty} a_n$ *genau dann konvergent, wenn die Folge der Partialsummenfolge beschränkt ist.*

Beweis. Da die Folgenglieder nichtnegativ sind, ist die Folge der Partialsummen monoton steigend. Nach einem Satz des letzten Kapitels ist eine monoton steigende Folge aber genau dann konvergent, falls sie beschränkt ist.

Man kann in der Situation des letzten Satzes sogar noch etwas mehr sagen: man erinnere sich, daß der Limes einer nach oben beschränkten Folge gerade das Supremum der von allen Folgengliedern gebildeten Menge ist. Daher hat man für eine konvergente Reihe mit nichtnegativen Gliedern a_n die Formel:

$$\sum_{n=0}^{\infty} a_n = \sup \left\{ \sum_{n=0}^{k} a_n \middle| k \in \mathbb{N} \right\}$$

Harmonische Reihe

Als Illustration zum letzten Satz betrachten wir die harmonische Reihe:

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots.$$

Für die Partialsummen s_n hat man hier

$$\begin{aligned} s_{2^k} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \ldots + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^{k-1}+1} + \ldots + \frac{1}{2^k}\right) \\ &> 1 + \frac{1}{2} + \frac{2}{4} + \ldots + \frac{2^{k-1}}{2^k} = \frac{k+1}{2}. \end{aligned}$$

Also ist die Partialsummenfolge nicht nach oben beschränkt, mithin gilt

Satz. *Die harmonische Reihe* $\sum_{n=1}^{\infty} \frac{1}{n}$ *ist divergent.*

Satz (Majorantenkriterium). *Seien* $\sum_{n=0}^{\infty} a_n$ *und* $\sum_{n=0}^{\infty} b_n$ *unendliche Reihen mit nicht-negativen Gliedern, und es gelte* $a_n \leq b_n$ *für alle* $n \in \mathbb{N}$. *Dann gilt: ist* $\sum_{n=0}^{\infty} b_n$ *konvergent, so ist auch* $\sum_{n=0}^{\infty} a_n$ *konvergent, und es gilt*

$$\sum_{n=0}^{\infty} a_n \leq \sum_{n=0}^{\infty} b_n.$$

In der in diesem Satz beschriebenen Situation nennt man $\sum_{n=0}^{\infty} b_n$ eine *Majorante* der Reihe $\sum_{n=0}^{\infty} a_n$.

Beweis. Für die zugehörigen Partialsummenfolgen $(s_k)_{k\in\mathbb{N}}$ von $\sum_{n=0}^{\infty} a_n$ und $(t_k)_{k\in\mathbb{N}}$ von $\sum_{n=0}^{\infty} b_n$ gilt $s_k \leq t_k$ für alle $k \in \mathbb{N}$. Ist die aus den b_n gebildete Reihe konvergent, so ist dem oben angegebenen Satz die Folge der t_k beschränkt. Dann ist aber auch die Folge der s_k beschränkt, nach dem eben zitierten Satz ist die Reihe der a_n konvergent. Die Ungleichung folgt aus $s_k \leq t_k$ und der Formel im Anschluß an den Beweis ebendieses Satzes.

Eine Reihe $a_0 + a_1 + \ldots$ könnte man durch Umordnung ihrer Glieder, d.h. durch Anordnung ihrer Glieder in einer anderen Reihenfolge aufzusummieren versuchen. Noch anders ausgedrückt: statt $\sum_{n=0}^{\infty} a_n$ versucht man $\sum_{n=0}^{\infty} a_{\sigma(n)}$ zu berechnen, wo σ eine Bijektion der Menge der natürlichen Zahlen mit sich selbst bedeutet. Im Fall von Reihen mit nichtnegativen Gliedern hat man hierzu den

Satz (Umordnungssatz). *Sei* $\sum_{n=0}^{\infty} a_n$ *eine konvergente Reihe und es gelte* $a_n \geq 0$ *für alle* $n \in \mathbb{N}$. *Sei* $\sigma\colon \mathbb{N} \to \mathbb{N}$ *eine Bijektion. Dann ist* $\sum_{n=0}^{\infty} a_{\sigma(n)}$ *konvergent und*

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} a_{\sigma(n)}.$$

Beweis. Die Partialsummenfolgen $(s_k)_{k\in\mathbb{N}}$ und $(s'_k)_{k\in\mathbb{N}}$ der unendlichen Reihen $\sum_{n=0}^{\infty} a_n$ bzw. $\sum_{n=0}^{\infty} a_{\sigma(n)}$ sind monoton steigend. Setzen wir daher

$$M = (s_n | n \in \mathbb{N}), \qquad M' = (s'_n | n \in \mathbb{N}),$$

so ist nach einer oben gemachten Bemerkung

$$\lim_{n\to\infty} s_n = \sup M, \lim_{n\to\infty} s'_n = \sup M'.$$

Wir haben also zu zeigen, daß $\sup M' = \sup M$ gilt. Nun ist aber tatsächlich $s'_n \leq s_m$, wenn nur $m \geq \max(\sigma(k)|k \leq n)$ ist. Daher ist $\sup M' \leq \sup M$. Umgekehrt ist stets $s_n \leq s'_l$, wenn l so groß gewählt ist, daß zu jedem $k \leq n$ ein $k' \leq l$ existiert, sodaß $k = \sigma(k')$ ist. Also ist $\sup M \leq \sup M'$, womit der Satz bewiesen ist.

Im Fall einer Reihe mit sowohl unendliche vielen negativen als auch unendlich vielen positiven Gliedern ist die im Umordnungssatz ausgesprochene Tatsache völlig falsch. Wir kommen darauf unten noch einmal zurück.

Ganz natürlich schließt sich an den letzten Satz die Frage an, ob man eine gegebene Reihe durch ein Setzen von Klammern aufzusummieren versuchen darf, d.h. ob etwa

$$\sum_{n=0}^{\infty} a_n = \sum_{k=0}^{\infty} \sum_{n=n_k}^{n_{k+1}} a_n$$

für jede beliebige streng monotone Folge (n_k) gilt.

Satz. *Ist $\sum_{n=0}^{\infty} a_n$ konvergent, ist $(n_k)_{k\in\mathbb{N}}$ eine streng monotone Folge, so ist auch*

$$\sum_{k=0}^{\infty} b_k, \qquad b_k = \sum_{n=n_k}^{n_{k+1}} a_n$$

konvergent, und zwar gegen denselben Limes.

Man beachte, daß hier über die a_n nichts weiter vorausgesetzt wird. Insbesondere müssen sie nicht notwendig nichtnegativ sein.

Beweis. Die Partialsummenfolge von $\sum_{k=0}^{\infty} b_k$ ist eine Teilfolge der Partialsummenfolge von $\sum_{n=0}^{\infty} a_n$. Damit folgt die Behauptung aus einem entsprechenden Satz über Folgen.

In der Anwendung des eben bewiesenen Satzes muß man sehr vorsichtig sein. Die Reihe

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

ist eine konvergente Reihe, und zwar mit Grenzwert 0. Lassen wir aber die Klammern weg, so erhalten wir die Reihe

$$1 - 1 + 1 - 1 + 1 - 1 + \ldots.$$

Diese hat als Partialsummenfolge die Folge $1, 0, 1, 0 \ldots$ und ist demnach offenbar nicht konvergent ! (?) Daher

> Man darf in unendlichen Reihen Klammern setzen, aber nicht weglassen.

Die beiden folgenden Konvergenzkriterien sind vielleicht die bekanntesten.

Satz (Wurzelkriterium). *Für alle $n \in \mathbb{N}$ gelte $0 \leq a_n \leq q^n$ mit einer reellen Zahl $q < 1$. Dann ist $\sum_{n=0}^{\infty} a_n$ konvergent, und der Grenzwert ist $\leq \frac{1}{1-q}$.*

Man beachte, die Voraussetzung, daß q strikt kleiner als 1 ist. Für $q = 1$ ist die im Satz gemachte Aussage falsch. Der Satz heißt Wurzelkriterium, weil die Voraussetzungen des Satzes — also $a_n \leq q^n$ mit $q < 1$ — zumeist in Form

$$\sqrt[n]{a_n} \leq q < 1$$

ausgesprochen wird. Wir haben allerdings die Existenz der n–ten Wurzel einer nichtnegativen Zahl noch nicht bewiesen und benutzen der logischen Strenge wegen für den Augenblick vorsichtigerweise die gegebene Formulierung.

Beweis. Die Voraussetzung des Satzes besagt nichts anderes als daß die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ eine Majorante von $\sum_{n=0}^{\infty} a_n$ ist. Der Satz folgt daher aus dem oben bewiesenen Majorantenkriterium und und der ebenfalls oben bewiesenen Konvergenz der geometrischen Reihe im Fall $q < 1$.

Satz (Quotientenkriterium). *Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge von nichtnegativen Zahlen. Für alle $n \in \mathbb{N}$ gelte $\frac{a_{n+1}}{a_n} \leq q$ mit einer reellen Zahl $q < 1$. Dann ist $\sum_{n=0}^{\infty} a_n$ konvergent, und der Grenzwert ist $\leq \frac{a_0}{1-q}$.*

Beweis. Aus der Voraussetzung folgt leicht mittels Induktion

$$\frac{a_n}{a_0} \leq q^n.$$

Der Satz folgt daher aus dem Wurzelkriterium und mit $\sum \frac{a_n}{a_0} = \frac{1}{a_0} \sum a_n$.

Man beachte, daß auch hier — beim Quotientenkriterium — die Voraussetzung $q < 1$ ganz wesentlich ist und auf keinen Fall zu $q = 1$ abgeschwächt werden kann: Die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent und dabei ist stets $\frac{n}{n+1} < 1$.

Eine weitere, wichtige und nützliche Bemerkung ist, daß im Wurzel- und Quotientenkriterium die Voraussetzung “für alle n’ durch “für fast alle n” ersetzt werden dürfen. Wie oben festgestellt, ändert sich das Konvergenzverhalten einer Reihe ja nicht, falls man endlich viele Glieder wegläßt. (Die Abschätzungen für die Grenzwerte im Wurzel- und Quotientenkriterium bleiben bei der eben erwähnten Abschwächung der Voraussetzungen selbstverständlich nicht bestehen !)

Exponentialfunktion

Bevor wir weitere allgemeine Sätze über Reihen auflisten, wollen wir die bisher vorgestellte Theorie an einem der grundlegendsten Beispiele anwenden. eine reelle Zahl. Zu einer reellen Zahl x kann man die Reihe *Potenzreihe*

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

betrachten.

Satz. *Die Reihe* $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ *konvergiert für jedes* x.

Den Grenzwert der hier betrachteten Reihe bezeichnet man mit $\exp(x)$, also

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

Beweis. Wir beweisen hier die Konvergenz zunächst für $x \geq 0$. Für $x = 0$ ist die behauptete Konvergenz offenbar trivial. Sei also $x > 0$. Dann ist

$$\frac{x^{n+1}}{(n+1)!} \Big/ \frac{x^n}{n!} = \frac{x}{n+1}.$$

Nun gilt $\lim_{n\to\infty} \frac{x}{n+1} = 0$. Insbesondere liegen daher fast alle Glieder der Folge $\frac{x^{n+1}}{(n+1)!} / \frac{x^n}{n!}$ unterhalb — sagen wir $q = \frac{1}{2}$. Die behauptete Konvergenz folgt nun mit dem Quotientenkriterium. Zur Behandlung des Falls von negativen x müssen wir noch etwas weitere Theorie lernen.

Wir definieren:

Absolut konvergente Reihen

> Die Reihe $\sum_{n=0}^{\infty} a_n$ heißt *absolut konvergent*, falls $\sum_{n=0}^{\infty} |a_n|$ konvergiert.

Bevor wir zur allgemeinen Theorie der absolut konvergenten Reihen kommen, geben wir ein Beispiel:

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \dots.$$

Diese Reihe zeigt, daß die absolut konvergenten Reihen einen echten Teilbereich in der Gesamtheit aller konvergenten Reihen ausmachen: sie ist konvergent, aber nicht absolut konvergent. Letzteres ist gerade der Satz über die Divergenz der harmonischen Reihe, denn

$$\sum_{n=1}^{\infty} \left| (-1)^{n-1} \cdot \frac{1}{n} \right| = \sum_{n=1}^{\infty} \frac{1}{n}.$$

Zum Nachweis des ersteren betrachten wir die Partialsummenfolge der Reihe $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}$; aus

$$s_{2k} := \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{5} - \frac{1}{6}\right) + \dots + \left(\frac{1}{2k-1} - \frac{1}{2k}\right)$$

folgt

$$s_2 < s_4 < s_6 < \dots,$$

aus

$$s_{2k+1} := 1 - \left(\frac{1}{2} - \frac{1}{3}\right) - \left(\frac{1}{4} - \frac{1}{5}\right) - \ldots - \left(\frac{1}{2k} - \frac{1}{2k+1}\right)$$

folgt

$$\ldots < s_5 < s_3 < s_1.$$

Ferner ist stets

$$s_{2k} < s_{2l+1}.$$

Wegen

$$s_{2n+1} - s_{2n} = \frac{1}{2n+1}$$

ist schließlich

$$\lim_{n\to\infty} |s_n - s_{n-1}| = 0.$$

Die Partialsummenfolge bildet also eine Intervallschachtelung, und ist mithin konvergent.

Zurück zur allgemeinen Theorie: Sei $\sum_{n=0}^{\infty} a_n$ gegeben. Hieraus leiten wir zwei Reihen mit nichtnegativen Gliedern ab. Wir setzen für $b \in \mathbb{R}$

$$b^+ := \frac{|b| + b}{2} = \begin{cases} |b| & \text{für } b \geq 0 \\ 0 & \text{für } b \leq 0 \end{cases},$$

$$b^- := \frac{|b| - b}{2} = \begin{cases} 0 & \text{für } b \geq 0 \\ |b| & \text{für } b \leq 0 \end{cases}.$$

Die Reihen

$$\sum_{n=0}^{\infty} a_n^+, \qquad \sum_{n=0}^{\infty} a_n^-$$

haben dann offensichtlich nichtnegative Glieder. Hier ist ein Beispiel:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \ldots = \sum_{n=1}^{\infty} (-1)^{n-1} \cdot \frac{1}{n}$$

$$\sum_{n=1}^{\infty} \left((-1)^{n-1} \cdot \frac{1}{n}\right)^+ = 1 + 0 + \frac{1}{3} + 0 + \frac{1}{5} + \ldots$$

$$\sum_{n=1}^{\infty} \left((-1)^{n-1} \cdot \frac{1}{n}\right)^- = 0 + \frac{1}{2} + 0 + \frac{1}{4} + 0 + \ldots.$$

Für absolut konvergente Reihen gilt nun allgemein:

Satz. *Die Reihe* $\sum_{n=0}^{\infty} a_n$ *ist dann und nur dann absolut konvergent, wenn* $\sum_{n=0}^{\infty} a_n^+$ *und* $\sum_{n=0}^{\infty} a_n^-$ *konvergent sind.*

Beweis. "$\Rightarrow$" ist klar, da $0 \leq a_n^+ \leq |a_n|$ und $0 \leq a_n^- \leq |a_n|$, sodaß die behauptete Konvergenz mit dem Majorantenkriterium folgt. Die Umkehrung folgt wegen $a_n^+ + a_n^- = |a_n|$, wonach ja

$$\sum_{n=0}^{\infty} |a_n| = \sum_{n=0}^{\infty} a_n^+ + \sum_{n=0}^{\infty} a_n^-.$$

Satz. *Ist die Reihe* $\sum_{n=0}^{\infty} a_n$ *absolut konvergent, so ist sie auch konvergent.*

Beweis. Die Behauptung folgt wegen $a_n = a_n^+ - a_n^-$ mit dem letzten Satz und dem oben angeführten Satz über Summen konvergenter Reihen.

Man beachte, daß der Beweis des letzten Satzes insbesondere die Formel

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} a_n^+ - \sum_{n=0}^{\infty} a_n^-$$

liefert. Man kann die Idee zum Beweis der beiden letzten Sätze noch weiter variieren, um etwa zu zeigen: "Ist $\sum_{n=0}^{\infty} a_n$ konvergent und $\sum_{n=0}^{\infty} a_n^+$ konvergent, so ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent." (Dies folgt analog zum Beweis der letzten beiden Sätze, jetz aber mit $|a_n| = 2a_n^+ - a_n$), oder eine analoge Aussage mit mit $\sum a_n^-$ statt $\sum a_n^+$.

Wie wir soeben gesehen haben, impliziert die absolute Konvergenz insbesondere die gewöhnliche Konvergenz. Oben haben wir gezeigt, daß die Exponentialreihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ für jedes $x \in \mathbb{R}$ absolut konvergiert; also konvergiert sie für jedes $x \in \mathbb{R}$, womit der Beweis des Satz über die Konvergenz der Exponentialreihe vollständig ist.

Wir bleiben noch eine Weile beim Beispiel der Exponentialreihe. Wir betrachten nun für festes $x \geq 0$ die Folge $\left(\left(1 + \frac{x}{n}\right)^n\right)_{n \in \mathbb{N}}$. Für $x = 1$ sind wir ihr schon im letzten Kapitel im Beispiel der stetigen Verzinsung begegnet. Wie dort schreiben wir

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

Hieraus ersehen wir, daß die Folge der $\left(1 + \frac{x}{n}\right)^n$ monoton steigend ist, und — unter Benutzung des letzten Satzes — daß sie auch beschränkt ist:

$$\sum_{k=0}^{n} 1 \cdot \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \cdot \frac{x^k}{k!} \leq \sum_{k=0}^{n} \frac{x^k}{k!} \leq \exp(x).$$

Also ist die Folge der $\left(\left(1+\frac{x}{n}\right)^n\right)$ konvergent. Darüberhinaus zeigt die eben gemachte Abschätzung noch

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

Wir zeigen nun, daß auch die umgekehrte Ungleichung, d.h. also, daß hier sogar Gleichheit gilt. Wegen der Monotonie ist ja

$$\sum_{k=0}^{m}\left(1-\frac{1}{m}\right)\cdots\left(1-\frac{k-1}{m}\right)\cdot\frac{x^k}{k!}=\left(1+\frac{x}{m}\right)^m \le \left(1+\frac{x}{n}\right)^n$$

für $m \le n$. Halten wir nun m fest und lassen n gegen Unendlich streben, Offenbar konvergiert die linke Seite gegen $\sum_{k=0}^{m}\frac{x^k}{k!}$. Daher erhalten wir für die Grenzwerte

$$\sum_{k=0}^{m}\frac{x^k}{k!} \le \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n.$$

Dies gilt nun aber für jedes $m \in \mathbb{N}$. Also erhalten wir aus der letzten Ungleichung

$$\exp(x)=\lim_{m\to\infty}\sum_{k=0}^{m}\frac{x^k}{k!} \le \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n.$$

Damit haben wir den folgenden Satz bewiesen.

Satz. *Es gilt*

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \exp(x) = \sum_{n=0}^{\infty}\frac{x^n}{n!},$$

insbesondere

$$\mathrm{e} = \sum_{n=0}^{\infty}\frac{1}{n!}$$

Mit dem letzten Satz haben wir auch sofort eine Interpretation der Größe $\exp(x)$.

> Läßt man ein Kapital A bei stetiger Verzinsung zu $x\%$ genau ein Jahr liegen, so ist es zu $A\cdot\exp(x)$ gewachsen.

Dies folgt wie im letzten Kapitel: $A\left(1+\frac{x}{n}\right)^n$ ist das Kapital nach einem Jahr, wenn die Zinsen alle $\frac{1}{n}$ Jahre zugeschlagen werden, und stetige Verzinsung bedeutet, daß man n gegen unendlich streben läßt. Eine andere mögliche Deutung wäre: "Läßt man ein Kapital A bei stetiger Verzinsung zu 100% genau x Jahre liegen, so ist das Endkapital $A\cdot\exp(x)$. Wir überlassen es den geneigten StudentInnen, dies zu überprüfen.

Aufgrund der letzten Deutung wird man die folgende *Funktionalgleichung* erwarten:

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

denn das Endkapital bei stetiger Verzinsung eines Kapitals A zu 100% nach $x+y$ Jahren (also $A \cdot \exp(x+y)$) sollte doch das gleiche sein wie das Endkapital, welches man erhält, wenn man A erst x Jahre lang stetig zu 100% verzinst, und das dabei resultierende Endkapital nochmals y Jahre lang stetig zu 100% verzinst (also gleich $A \cdot \exp(x) \cdot \exp(y)$).

Die Funktionalgleichung ist tatsächlich wahr, wie wir unten sehen werden. Insbesondere hat man daher für natürliche Zahlen n

$$\exp(n) = \exp(1 + \ldots + 1) = \exp(1)^n = \mathrm{e}^n.$$

Es ist daher naheliegend und üblich

$$\mathrm{e}^x$$

statt $\exp(x)$ zu schreiben. Für jede reelle Zahl hat damit e^x eine wohlbestimmte Bedeutung (nämlich den Wert der unendlichen Reihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$). Die hierdurch erklärte Abbildung $x \mapsto \mathrm{e}^x$ nennt man *Exponentialfunktion*. Zum Beweis der Funktionalgleichung für die Exponentialfunktion benötigen wir noch einige Vorbereitungen aus der Theorie der unendlichen Reihen. Wir kehren daher zunächst wieder zur allgemeinen Theorie zurück.

Umordnungssatz

Satz. *Die Reihe $\sum_{n=0}^{\infty} a_n$ sei absolut konvergent, $\sigma : \mathbb{N} \to \mathbb{N}$ eine Bijektion. Dann ist auch $\sum_{n=0}^{\infty} a_{\sigma(n)}$ absolut konvergent und*

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} a_{\sigma(n)}.$$

Beweis. Es ist $a_n = a_n^+ - a_n^-$. Also ist auch $a_{\sigma(n)} = a_{\sigma(n)}^+ - a_{\sigma(n)}^-$. Nach Voraussetzung und nach dem Satz über die Vertauschung von Reihen mit nichtnegativen Gliedern ist $\sum_{n=0}^{\infty} a_{\sigma(n)}^+$ konvergent, und $\sum_{n=0}^{\infty} a_{\sigma(n)}^+ = \sum_{n=0}^{\infty} a_n^+$. Eine analoge Ausage gilt für $\sum_{n=0}^{\infty} a_{\sigma(n)}^-$. Nach dem Satz über die Summe (hier Differenz) konvergenter Reihen haben wir daher

$$\begin{aligned} \sum_{n=0}^{\infty} a_{\sigma(n)} &= \sum_{n=0}^{\infty} a_{\sigma(n)}^+ - \sum_{n=0}^{\infty} a_{\sigma(n)}^- \\ &= \sum_{n=0}^{\infty} a_n^+ - \sum_{n=0}^{\infty} a_n^- = \sum_{n=0}^{\infty} a_n \end{aligned} .$$

Der Aussage des letzten Satz über die Unveränderlichkeit des Grenzwerts bei Umordnung ist definitiv falsch für *bedingt konvergente Reihen*, d.h. nicht absolut konvergente Reihen. Für eine bedingt konvergente Reihe kann man zeigen, daß mittels geeigneter Umordnung jede beliebig vorgegebene Zahl als Grenzwert und sogar Divergenz erreicht werden kann.

Produkt unendlicher Reihen

Satz. *Die Reihen* $\sum_{n=0}^\infty a_n$ *und* $\sum_{n=0}^\infty b_n$ *seien absolut konvergent gegen a bzw. b. Sei* $\alpha\colon \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ *eine Bijektion. Dann ist die Reihe*

$$\sum_{n=0}^{\infty} a_{\alpha_1(n)} b_{\alpha_2(n)} \qquad (\ \alpha(n) = (\alpha_1(n), \alpha_2(n)\ )$$

absolut konvergent, und zwar gegen $a \cdot b$.

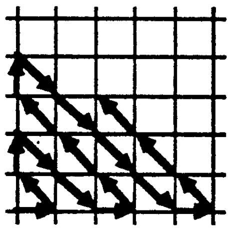
Der Wert der Reihe

$$\sum_{n=0}^{\infty} a_{\sigma(n)} b_{\sigma(n)}$$

ist insbesondere unabhängig von der speziellen Wahl von α, man bezeichnet ihn daher auch einfach mit

$$\sum_{m,n=0}^{\infty} a_m b_n.$$

Daß die Aussage des Satz nicht leer ist, d.h. daß überhaupt Bijektionen zwischen $\mathbb{N}$ und $\mathbb{N} \times \mathbb{N}$ bestehen, lassen wir als Übungsaufgabee Man kann eine solche Bijektionen etwa leicht aus folgender Skizze konstruieren:



Beweis. Es gilt

$$\sum_{n=0}^{k} \left| a_{\alpha_1(n)} b_{\alpha_2(n)} \right| \leq \sum_{n=0}^{m} |a_n| \sum_{n=0}^{m} |b_n| \,,$$

wobei m zu gegebenem k jeweils so groß gewählt ist, daß $\alpha_i(n) \leq m$ für alle $n \leq m$ gilt. Die rechte Seite dieser Ungleichung konvergiert nach Voraussetzung (monoton steigend) gegen

$$\sum_{n=0}^{\infty} |a_n| \sum_{n=0}^{\infty} |b_n| \,,$$

womit diese Zahl eine obere Schranke für die monoton steigende Partialsummenfolge der Reihe

$$\sum_{n=0}^{\infty} \left| a_{\alpha_1(n)} b_{\alpha_2(n)} \right|$$

darstellt. Mithin ist letztere Reihe konvergent, d.h. die Reihe

$$\sum_{n=0}^{\infty} a_{\alpha_1(n)} b_{\alpha_2(n)}$$

ist tatsächlich absolut konvergent. Nach dem Umordnungssatz für absolut konvergente Reihen ist der Grenzwert letzterer Reihe sogar unabhängig von der speziellen Wahl von α: Ist α' eine andere Bijektion, so erhält man ja die mit α' gebidete Reihe aus der mit α gebildeten durch Umordnung mit der Bijektion $\sigma: \mathbb{N} \to \mathbb{N}$, wobei $\sigma = \alpha^{-1} \circ \alpha'$. Zur Berechnung des Grenzwerts unserer Reihe können wir das α daher gemäß unseren eigenen Vorstellungen wählen. Wir wählen es so, daß

$$k \mapsto \sum_{n=0}^{k} a_n \sum_{n=0}^{k} b_n$$

eine Partialsummenfolge unserer Reihe ist; der Grenzwert unserer Reihe stimmt dann mit dem Grenzwert dieser Partialsummenfolge überein, und letztere ist offenbar gleich $a \cdot b$. Daß man tatsächlich solch ein α finden kann, wie wir es benutzt haben, lassen wir als Übungsaufgabe. Damit ist der Satz bewiesen.

Als Anwendung dieses Satzes können wir nun beweisen:

Satz. *Für* $x, y \in \mathbb{R}$ *gilt* $\exp(x+y) = \exp(x) \cdot \exp(y)$.

Beweis: Zunächst ist ja unter Beachtung der absoluten Konvergenz der Exponentialreihe und nach unserem letzten Satz

$$\begin{aligned} \exp(x) \cdot \exp(y) &= \sum_{r=0}^{\infty} \frac{x^r}{r!} \cdot \sum_{s=0}^{\infty} \frac{y^s}{s!} \\ &= \sum_{r,s=0}^{\infty} \frac{x^r}{r!} \cdot \frac{y^s}{s!}. \end{aligned}$$

Da man in konvergenten Reihen beliebig Klammern setzen darf, können wir für jedes $n \in \mathbb{N}$ diejenigen Paare (r, s) zusammenfassen, für die $r + s = n$ ist. Dann wird

$$\exp(x) \cdot \exp(y) == \sum_{n=0}^{\infty} c_n, \qquad c_n = \sum_{\substack{r,s \geq 0 \\ r+s=n}} \frac{x^r}{r!} \frac{y^s}{s!}.$$

Nun ist aber nach dem binomischen Lehrsatz

$$c_n = \frac{1}{n!} \sum_{\substack{r,s \geq 0 \\ r+s=n}} \frac{n!}{r!s!} x^r y^s = \frac{1}{n!}(x+y)^n.$$

Setzen wir dies in die letzte Identität ein, so erhalten wir

$$\exp(x) \cdot \exp(y) = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \exp(x+y),$$

und genau das wollten wir beweisen.

Die im Beweis mittels der dort durchgeführten Zusamenfassung der Produktreihe gebildete Reihe, die man offenbar stets zu je zwei absolut konvergenten Reihen bilden kann, nennt man *Cauchy-Produkt* der beiden Ausgangsreihen.

Aus der Funktionalgleichung erhält man insbesondere $\exp(x) \cdot \exp(-x) = \exp(0) = 1$. Da $\exp(x)$ für positives x positiv ist, wie man unmittelbar der definierenden Reihe entnimmt, ist daher stets $\exp(x) > 0$. Also können wir die Exponentialfunktion als Abbildung

$$\exp : \mathbb{R} \to \mathbb{R}_+ := \{r \mid r \in \mathbb{R} \wedge r > 0\}$$

auffassen. Man kann damit die Funktionalgleichung noch anders interpretieren: $\mathbb{R}$ und $\mathbb{R}_+$ sind Gruppen, die erste bezüglich der Addition, die zweite bezüglich der Multiplikation. Die Funktionalgleichung besagt nun nichts anderes, als daß exp einen Homomorphismus dieser Gruppen definiert.

Für $x_1, x_2 \in \mathbb{R}, x_1 < x_2$ ist $\exp(x_1) < \exp(x_2)$, denn man hat ja $\exp(x) = 1+$Reihe mit positiven Gliedern, d.h. $\exp(x) > 1$ für positive x, insbesondere also $\exp(x_2 - x_1) > 0$, und daher

$$\exp(x_2) = \exp(x_1) \cdot \exp(x_2 - x_1) > \exp(x_1).$$

Insbesondere ist also exp injektiv. Außerdem haben wir oben schon gesehen, daß $\exp(n) = (\mathrm{e})^n$. Die Folge $(\mathrm{e}^n)_{n\in\mathbb{N}}$ nimmt also beliebig große Werte an, die Folge $(\mathrm{e}^{-n})_{n\in\mathbb{N}}$ wiederum nimmt beliebig kleine Werte an. Es ist daher plausibel, daß exp auch surjektiv ist. Das ist auch tatsächlich wahr; allerdings reichen unsere Methoden noch nicht aus, dies zu beweisen. Den Beweis der Surjektivität werden wir im nächsten Kapitel als Folgerung unserer Untersuchungen an stetigen Funktionen geben.

Somit ist $\exp : \mathbb{R} \to \mathbb{R}_+$ ein Isomomorphismus, (= bijektiver Gruppenhomomorphismus). Insbesondere existiert die Umkehrabbildung, die sogenannte Logarithmusfunktion. Diese wird mit ln oder log bezeichnet (sprich “Logarithmus naturalis” oder einfach “Logarithmus”:

Der natürliche Logarithmus

$$\ln : \mathbb{R}_+ \to \mathbb{R}.$$

Der Logarithmus ist — als Umkehrfunktion eines Gruppenisomorphismus — wieder homomorph, d.h. es gilt für $a, b \in \mathbb{R}_+$:

$$\ln(a \cdot b) = \ln(a) + \ln(b), \qquad \ln(a/b) = \ln(a) - \ln(b).$$

Sei nun irgendeine Zahl $a > 0$ gegeben, $x \in \mathbb{R}$. Dann setzt man

$$a^x := \exp(x \cdot \ln a).$$

Offenbar definiert die Zuordnung

$$\mathbb{R} \to \mathbb{R}_+, \quad x \mapsto a^x$$

wieder einen Gruppenisomorphismus. Insbesondere hat man

$$a^{x+y} = a^x \cdot a^y.$$

Existenz n–ter Wurzeln

Als Folge hiervon erhält man $a^n = (a)^n$, d.h. a^n ist gleich der n-fachen Potenz von a. Dies wiederum zieht sofort die Existenz der n–ten Wurzel zu jeder positiven reellen Zahl y nach sich: in der Tat ist ja

$$\left(\exp\left(\frac{\log y}{n}\right)\right)^n = y.$$

Die Umkehrabbildung zu $x \mapsto a^x$ bezeichnen wir mit $\log_a$ ("Logarithmus zur Basis a); sie definiert also einen Gruppenisomorphismus $(\mathbb{R}_+, \cdot) \to (\mathbb{R}, +)$. Es gilt natürlich

$$\log_a y = \frac{\log y}{\log a},$$

wie man sich leicht überlegt.

Für $a = 10$ ist $\log_a$, als der "dekadische Logarithmus" oder "Logarithmus zur Basis 10" aus der Schule bekannt: wegen

$$a \cdot b = \exp(\ln(a \cdot b)) = \exp(\ln a + \ln b)$$

kann man sehr große Zahlen a und b näherungsweise multiplizieren, indem man $\log a$ und $\log b$ in einer Tabelle nachschlägt, die Ergebnisse addiert, und hierfür in der gleichen Tabelle die Zahl aufsucht, deren Logarithmus gerade diese Summe ist. Dieses Verfahren kennt man schon seit Jahrhunderten.

Der Mathematiker Bürgi, ein Meister der Potenzrechnung, hatte zu Beginn des 17. Jahrhunderts folgende geniale Idee: Die Potenzen x^n einer Zahl x sehr nahe bei 1 liefern eine sehr dichte Zahlenfolge. Bürgi stellte sich eine Tafel auf, mit deren Hilfe er sich wegen der bekannten Identität $x^k \cdot x^l = x^{k+l}$ für $k, l \in \mathbb{N}$ das Multiplizieren zu erleichtern gedachte: Er wählte die Zahl $x = 1.0001$ und tabellierte ihre Potenzen:

n	$(1.0001)^n$
1	1.0001000000000000
2	1.0002000100000000
3	1.0003000300010000
4	1.0004000600040001
5	1.0005001000100005
6	1.0006001500200015
7	1.0007002100350035
8	1.0008002800560070
9	1.0009003600840126
10	1.0010004501200210
11	1.0011005501650330
12	1.0012006602200495
13	1.0013007802860715
14	1.0014009103641001
15	1.0015010504551365
16	1.0016012005601820
17	1.0017013606802380
$\vdots$	$\vdots$
23027	9.9999977968106672
23028	10.000997796590348

Mit dieser Tabelle konnte er also Zahlen a, b zwischen 1 und 10 ziemlich genau untereinander multiplizieren. Suche a und b in der rechten Spalte auf, addiere zughörige n und n' aus der linken Spalte, suche $x^{n+n'}$ in der rechten. Hierbei kann es vorkommen, daß $n + n' > 23028$ ist. Dann sucht man $x^{n+n'-23028}$ nach. Dieser Wert ist ja nur um den Faktor $x^{23028} \approx 10$ kleiner als das gesuchte Ergebnis. Dieses Verfahren nutzt also in moderner Sprechweise den Logarithmus zur Basis $x = 1,0001$ zur Multiplikation großer (hier "sehr langer") Zahlen.

John Napier, ein Schotte (1550-1617), fand das zu ungenau. Er verlangte Besseres für die Seefahrt, nahm also als Basis $1 + 10^{-7} = 1,0000001$ und erstellte eine neue Tabelle (1614). Er entdeckte auch, wie man die Tafeln ineinander umrechnen kann. Briggs schlug dann vor, zur Basis 10 zu rechnen; der Vorteil sei: $\log_{10} 10 = 1$. Bürglis Idee, als Basis eine Zahl nahe bei 1 zu nehmen, führt übrigens zur Eulerschen Zahl e, denn

$$(1,0001)^{10000} \approx \mathrm{e}.$$

Kepler benutzte Logarithmentafeln für astronomische Berechnungen. Keplers Lehrer traute den Logarithmen jedoch nicht und rechnete lieber mit der alten Methode und der gut bekannten cos-Funktion, die wir unten besprechen werden. Man hat das Additionstheorem

$$\cos x \cdot \cos y = \frac{1}{2}(\cos(x+y) + \cos(x-y)).$$

Um $A \cdot B$ — etwa für A, B zwischen 0 und 1 — auszurechnen, ging man so vor: man sucht x, y auf, sodaß $\cos x = A$ und $\cos y = B$. Dann bildet man $x + y$ und $x - y$, addiert die zugehörigen Werte $\cos(x+y), \cos(x-y)$, teilt durch 2, etc.. Historisch ist also klar, daß auch der cos etwas mit exp zu tun haben muß, da auch der cos ja etwas über das Verhältnis der Multiplikation zur Addition sagt. Wir werden dies in einem Moment näher studieren.

Die Hyperbelfunktionen

Wir beginnen zunächst mit den begrifflich etwas einfacheren Hyperbelfunktionen:

$$\cosh x := \frac{1}{2}\left(\mathrm{e}^x + \mathrm{e}^{-x}\right)$$
$$\sinh x := \frac{1}{2}\left(\mathrm{e}^x - \mathrm{e}^{-x}\right)$$

(sprich "cosinus hyperbolicus" bzw. "sinus hyperbolicus"). Wir ordnen nun jedem $x \in \mathbb{R}$ das Paar

$$(\xi, \eta) = (\cosh x, \sinh x) = \left(\frac{1}{2}\left(e^x + e^{-x}\right), \frac{1}{2}\left(e^x - e^{-x}\right)\right) \in \mathbb{R}^2$$

zu. Was geschieht nun, wenn man x variiert, d.h. wenn man x die *Kurve* $\mathbb{R}$ durchlaufen läßt ? Dann wird wohl (ξ, η) auch eine Kurve — in der Ebene $\mathbb{R}^2$ — durchlaufen. Kann man diese beschreiben ? Setzt man $\xi = \cosh x$ und $\eta = \sinh x$, so ist

$$\xi + \eta = e^x > 0$$
$$\xi - \eta = e^{-x} > 0$$
$$\xi^2 - \eta^2 = 1$$

Wir haben also eine Abbildung

$$\mathbb{R} \to H := \left\{(\xi, \eta) \,|\, \xi^2 - \eta^2 = 1 \wedge \xi > 0\right\}$$
$$x \mapsto (\cosh x, \sinh x).$$

Satz. *Diese Abbildung ist bijektiv.*

Beweis. Sei $(\xi, \eta) \in H$, also insbesondere $\xi + \eta > 0$. Setze $x := \log(\xi + \eta)$, d.h. $\xi + \eta = \mathrm{e}^x$. Wegen $1 = \xi^2 - \eta^2 = (\xi + \eta)(\xi - \eta) = 1$ folgt dann $\xi - \eta = \mathrm{e}^{-x}$. Zusammengefaßt: $\xi = \cosh(x)$ und $\eta = \sinh(x)$. Dies beweist die Surjektivität. Eine nochmalige Durchsicht dieses Beweises zeigt aber auch, daß sich x eindeutig aus (ξ, η) berechnen läßt. Die ist aber gerade die behauptete Injektivität unserer Abbildung.

Damit ist die *Bildkurve* unserer Abbildung gerade eine (die "rechte") der beiden zusammenhängenden Teile der durch die Gleichung $\xi^2 - \eta^2 = 1$ beschriebenen Kurve im "(ξ, η)–Raum", einer *Hyperbel*. Mit Hilfe von cosh und sinh haben wir also eine *Parameterdarstellung* der Hyperbel gefunden, i.e. eine (bijektive) Abbildung $R \to H$.

Wie schon im Beweis des letzten Satz vorgeführt, ist es leicht, Eigenschaften der Hyperbelfunktionen aus entsprechenden Eigenschaften der Exponentialfunktion abzuleiten. Ein weiteres Beispiel hierzu ist:

Satz (Additionstheoreme). *Für alle x und y gilt*

$$\begin{aligned}\cosh(x+y) &= \cosh x \cdot \cosh y + \sinh x \cdot \sinh y\\ \sinh(x+y) &= \sinh x \cdot \cosh y + \cosh x \cdot \sinh y.\end{aligned}$$

Beweis. Der Beweis dieser Additionstheorem ist offensichtlich: zum Beispiel folgt die erste Identität wegen

$$\begin{aligned}&\cosh x \cdot \cosh y + \sinh x \cdot \sinh y\\ &= \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2} \cdot \frac{\mathrm{e}^y + \mathrm{e}^{-y}}{2} + \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2} \cdot \frac{\mathrm{e}^y - \mathrm{e}^{-y}}{2}\\ &= \ldots = \frac{\mathrm{e}^{(x+y)} + \mathrm{e}^{-(x+y)}}{2}\\ &= \cosh(x+y).\end{aligned}$$

Die zweite Identität folgt mit einer analogen Rechnung.

Unmittelbar aus der Definition der Hyperbelfunktionen liest man auch ab, daß cosh eine *gerade* und sinh eine *ungerade* Funktion ist, d.h. daß

$$\cosh(-x) = \cosh(x), \qquad \sinh(-x) = -\sinh(x)$$

gilt. Ferner erhält man mittels der Exponentialreihe auch sogleich eine Formel zur Berechnung der Hyperbelfunktionen, die *Potenzreihenentwicklung der Hyperbelfunktionen*. Aus

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \qquad e^{-x} = \sum_{n=0}^{\infty} (-x)^n/n! = \sum_{n=0}^{\infty} (-1)^n \frac{x^n}{n!},$$

folgt durch gliedweises Addieren dieser beiden Reihen:

$$\cosh(x) = \sum_{k=0}^{\infty} \frac{x^2 k}{(2k)!} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \ldots.$$

Diese Reihe konvergiert absolut für jedes $x \in \mathbb{R}$. Genauso erhält man:

$$\sinh(x) = \sum_{k=0}^{\infty} \frac{x^2 k + 1}{(2k+1)}! = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots.$$

Die Kreisfunktionen

Mittels der Exponentialfunktion konnten wir die Hyperbel $\xi^2 - \eta^2 = 1$ parametrisieren. Kann man etwas ähnliches für den Kreis durchführen ? Die Gleichung, die den Kreis beschreibt — etwa in der (ξ, η)–Ebene vom Radius 1 und mit Mittelpunkt in $(0, 0)$ — ist gegeben durch

$$\xi^2 + \eta^2 = 1.$$

Formal kommt man von der Hyperbel zum Kreis, indem man η durch $\frac{1}{i}\eta$ ersetzt, wo i die imaginäre Einheit bedeutet, also eine Lösung der Gleichung $x^2 = -1$. Da nun für jedes x stets $(\cosh(x))^2 + (\sinh(x))^2 = 1$ ist, erhalten wir durch formales Einsetzen, daß für jedes x auch

$$(\cosh(x))^2 + \left(\frac{1}{i} \cdot \sinh(x)\right)^2 = 1$$

Allerdings ist das keine Parametrisierung des Kreises, da ja $\frac{1}{i} \cdot \sinh(x)$ nicht reell ist. Ersetzen wir aber x durch ix, so wird man erwarten, daß die letzte Gleichung richtig bleibt, d.h. daß gilt

$$(\cosh(ix))^2 + \left(\frac{1}{i}\sinh(ix)\right)^2 = 1$$

welche Bedeutung $\cosh ix$ und $\sinh ix$ auch haben mögen. Die Interpretation dieser Symbolik liegt aber auf der Hand: ersetzt man formal in der oben gegebenen Reihenentwicklung von $\frac{1}{i}\sinh x$ die Zahl x durch ix, und beachtet man, daß in dieser Reihenentwicklung nur ungerade Potenzen von x auftreten, und daß

$$(ix)^{2k+1} = i \cdot i^{2k} x^{2k} = i \cdot (i^2)^k x^{2k} = i \cdot (-1)^k x^{2k},$$

so erhält man die Reihe

$$\cdot \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Dies ist für uns ein sinnvolles mathematisches Objekt. Da die sinh–Reihe für jedes x absolut konvergiert, ist diese Reihe genauso gut für jedes x absolut konvergent. Die folgende Definition ist daher für jede reelle Zahl einen Sinn:

$$\sin x := \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - + \cdots.$$

Führt man die eben getanen Überlegungen für cosh anstelle von sinh durch, so gelangt man zu

$$\cos x := \sum_{k=0}^{\infty} (-1)^{2k} \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + - \ldots$$

Die hier rechts stehende Reihe ist wieder für jedes x absolut konvergent. Aus der Definition ist sofort klar, daß der Sinus ungerade und der Cosinus eine ungerade Funktion ist, als Formel:

$$\cos(-x) = \cos(x), \qquad \sin(-x) = -\sin(x).$$

Ferner hat man

Satz (Additionstheorem). *Für alle x und y gilt*

$$\sin(x+y) = \sin(x)\cos(y) + \cos(x)\cdot\sin(y)$$
$$\cos(x+y) = \cos(x)\cdot\cos(y) - \sin(x)\cdot\sin(y).$$

Diesen Satz kann man prinzipiell und ohne (wirkliche) Schwierigkeiten unmittelbar aus den Reihendarstellungen unter Anwendung des Satzes über das Produkt von unendlichen Reihen beweisen. Es wird sich aber unten ein durchsichtiger und weniger künstlicher Beweis ergeben. Als unmittelbare Folge des letzten Satz kann man sich leicht überlegen:

Korollar.

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

Wir werden unten noch einmal darauf zurückkommen. Gemäß der im Korollar ausgesprochenen Behauptung definiert die Zuordnung

$$x \mapsto (\cos x, \sin x)$$

eine Abbildung

$$\mathbb{R} \to \left\{(\xi,\eta) \in \mathbb{R}^2 | \xi^2 + \eta^2 = 1\right\}.$$

Diese Abbildung ist zwar surjektiv, allerdings nicht injektiv, wie wir unten sehen werden

Komplexe Zahlen

Wesentlich durchsichtiger wird die Theorie der Kreisfunktionen bei Einbeziehung der komplexen Zahlen.Eine komplexe Zahl ist ein "Ausdruck" der Gestalt

$$a + ib,$$

wo a und b reelle Zahlen bedeuten. Wer durch die Wendung "Ausdruck" verunsichert ist, aber Sicherheit bei der naiven Mengenlehre findet, der mag $a+ib$ nur als eine andere Schreibweise für das Paar $(a,b) \in \mathbb{R}^2$ interpretieren. Die Zahl a heißt Realteil, die Zahl b Imaginärteil von $z = a+ib$; in Symbolen

$$a = \Re(z), \qquad b = \Im(z).$$

Auf der Menge der komplexen Zahlen definiert man eine Addition und eine Multiplikation vermöge

$$(a+ib) + (c+id) := (a+c) + i(b+d),$$
$$(a+ib)\cdot(c+id) := (ac-bd) + i(ad+bc).$$

Die komplexen Zahlen erfüllen zusammen mit diesen Operationen die Axiome eines Körpers [vgl. Kapitel 2]. Dies ist der Körper der komplexen Zahlen, und er wird mit $\mathbb{C}$ bezeichnet. Er enthält den Körper der reellen

Zahlen — dies sind gerade die komplexen Zahlen $a + ib$, wo $b = 0$. Die zu $z = a + ib$ *konjugiert* komplexe Zahl $\overline{z}$ ist definiert vermöge

$$\overline{z} = a - ib.$$

Die Zuordnung $z \mapsto \overline{z}$ definiert einen *Körperautomorphismus*, d.h. es gilt

$$\overline{(z + w)} = \overline{z} + \overline{w}, \qquad \overline{(z \cdot w)} = \overline{z} \cdot \overline{w}.$$

Darüberhinaus gilt unmittelbar nach der Definition

$$\overline{\overline{z}} = z,$$

d.h. die komplexe Konjugation ist eine *Involution*. Als *Betrag* der komplexen Zahl z bezeichnet man die Größe

$$|z| = \sqrt{a^2 + b^2} = \sqrt{z\overline{z}}.$$

Genaugenommen haben wir die Existenz der Quadratwurzel einer positiven reellen Zahl bislang noch nicht bewiesen, sodaß die eben gegebene Definition des Absolutbetrages im Rahmen eines streng logischen Aufbaus an dieser Stelle gar keinen Sinn ergäbe; wir greifen hier aber etwas vor und setzen für den Moment die Existenz von Wurzeln voraus. Der Beweis für ihre Existenz wird im nächsten Kapitel nachgetragen. Offenbar gilt

Satz. *Für alle komplexen Zahlen z ist*

$$|\Re(z)| \leq |z|, \qquad |\Im(z)| \leq |z|.$$

Etwas schwieriger ist der Nachweis der *Dreiecksungleichung*

Satz (Dreiecksungleichung). *Für alle komplexen Zahlen z, w gilt*

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

Beweis. Man hat ja

$$\begin{aligned} & |z + w|^2 \leq (|z| + |w|)^2 \\ \iff & (z + w)(\overline{z} + \overline{w}) \leq z\overline{z} + 2|z||w| + w\overline{w} \\ \iff & w\overline{z} + z\overline{w} \leq 2|z||w| \\ \iff & \frac{w\overline{z} + z\overline{w}}{2} \leq |z||w|. \end{aligned}$$

Die linke Seite der letzten Ungleichung ist aber gerade $\Re(w\overline{z})$, wie man leicht aus der einfachen Formel

$$\Re(y) = \frac{y + \overline{y}}{2}$$

abliest, und die rechte Seite ist gleich $|w\overline{z}|$, und

$$\Re(w\overline{z}) \leq |w\overline{z}|$$

ist ja nach dem letzten Satz wahr.

Folgen komplexer Zahlen

Wir betrachten jetzt unendliche Reihen von komplexen Zahlen.

Eine Folge $(a_n)_{N\in\mathbb{N}}$ von komplexen Zahlen a_n heißt konvergent, falls die beiden Folgen $(\Re(a_n))_{n\in\mathbb{N}}$ und $(\Im(a_n))_{n\in\mathbb{N}}$ konvergieren. Im Fall der Konvergenz setzt man

$$\lim_{n\to\infty} a_n := \lim_{n\to\infty} \Re(a_n) + i \lim_{n\to\infty} \Im(a_n).$$

Unter einer unendlichen Reihe von komplexen Zahlen verstehen wir wieder die Folge der Partialsummen, sodaß also die unendliche Reihe $\sum_{n=0}^{\infty} a_n$ ($a_n \in \mathbb{C}$) genau dann konvergiert, falls die beiden Reihen

$$\sum_{n=0}^{\infty} \Re(a_n), \qquad \sum_{n=0}^{\infty} \Im(a_n)$$

konvergieren. Im Fall der Konvergenz bezeichnen wir den Grenzwert wieder mit dem gleichen Symbol wie die unendliche Reihe, also

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} \Re(a_n) + i \cdot \sum_{n=0}^{\infty} \Im(a_n).$$

Man kann die Konvergenz einer Folge von komplexen Zahlen auch anders, anolag zum reellen Fall beschreiben.

Satz. *Sei $(a_n)_{N\in\mathbb{N}}$ eine Folge von komplexen Zahlen. Dann gilt*

$$\lim_{n\to\infty} a_n = a \iff \forall \epsilon > 0\, \exists n_0 \in \mathbb{N}\, \forall n \in \mathbb{N}\ (n \geq n_0 \implies |a - a_n| < \epsilon).$$

Beweis. Die Richtung “$\Longleftarrow$” ist eine Folge der Ungleichungen

$$|\Re(a - a_n)| \leq |a - a_n|\,, \qquad |\Im(a - a_n)| \leq |a - a_n|\,.$$

Die Richtung “$\Rightarrow$” folgt mit der Dreiecksungleichung, hier in der Form

$$|a - a_n| \leq |\Re(a - a_n)| + |\Im(a - a_n)|\,.$$

Die Details überlassen wir dem Leser.

Eine Reihe $\sum_{n=0}^{\infty} a_n$ von komplexen Zahlen a_n heißt absolut konvergent, falls $\sum_{n=0}^{\infty} \Re(a_n)$ und $\sum_{n=0}^{\infty} \Im(a_n)$ absolut konvergent sind.

Es ist klar, daß eine absolut konvergente Reihen komplexer Zahlen konvergent ist. Ferner gilt für absolut konvergente Reihen wortwörtlich wie im Fall der reellen unendlichen Reihen der Umordnungssatz und der Satz über die Produkte von unendlichen Reihen. Auch die Beweise dieser beiden Sätze können wortwörtlich übertragen werden, wobei man für den Beweis des Produktsatzes noch den folgenden Satz benötigt:

Satz. *Eine Reihe $\sum_{n=0}^{\infty} a_n$ von komplexen Zahlen a_n ist dann und nur dann absolut konvergent, wenn die Reihe $\sum_{n=0}^{\infty} |a_n|$ konvergiert.*

Beweis. Die Aussage folgt wie im Beweis des letzten Satzes mit den Ungleichungen

$$|\Re(a_n)|\,,\ |\Im(a_n)| \leq |a_n| \leq |\Re(a_n)| + |\Im(a_n)|\,.$$

Die Details überlassen wir wieder dem Leser.

Fortsetzung der Exponentialfunktion ins Komplexe

Nach diesen Vorbereitungen über Reihen komplexer Zahlen können wir nun alle oben betrachteten Funktionen via ihrer Reihendarstellungen für komplexe Argumente erklären. Wir setzen zunächst

$$\exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!} \qquad \text{für} \quad z \in \mathbb{C}.$$

Die hier auftretende unendliche Reihe von komplexen Zahlen ist absolut konvergent (sie konvergiert falls man z durch die reelle Zahl $|z|$ ersetzt), mithin ist sie konvergent, mithin ist der Grenzwert, den wir mit $\exp(z)$ oder auch e^z bezeichnen, für jedes komplexe z eine wohlbestimmte komplexe Zahl. Wie im reellen Fall kann man zeigen:

$$\forall z, w \in \mathbb{C}(\exp(z+w) = \exp(z) \cdot \exp(w)).$$

Insbesondere ist stets $\exp(w) \cdot \exp(-w) = \exp(0) = 1$, also $\exp(w) \neq 0$. Daher definiert die Zuordnung

$$z \mapsto \exp(z)$$

eine Abbildung

$$\mathbb{C} \to \mathbb{C}^* := \mathbb{C} \setminus \{0\};$$

nach der Funktionalgleichung ist diese Abbildung sogar ein Gruppenhomomorphismus. Wir setzen weiter für komplexes z:

$$\sin z = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!},$$
$$\cos z = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}.$$

Wieder sehen wir, daß diese Reihen absolut konvergent sind, da sie ja für reelle Argumente z konvergieren, sodaß wir damit für jede komplexe Zahl z einen wohldefinierten Wert $\sin z$ bzw. $\cos z$ erhalten. Mit

$$i^{4k} = 1, \quad i^{4k+1} = i, \quad i^{4k+2} = -1, \quad i^{4k+3} = -i \qquad (k \in \mathbb{Z})$$

erhält man unmittelbar aus den obigen Reihenentwicklungen den folgende

Eulersche Identität

Satz (Eulersche Identität). *Für alle komplexen Zahlen z ist*

$$\mathrm{e}^{iz} = \cos z + i \sin z.$$

Wie man sich leicht überlegt, kann man die Eulersche Identität auch formulieren als

$$\cos z = \frac{\mathrm{e}^{iz} + \mathrm{e}^{-iz}}{2}, \qquad \sin z = \frac{\mathrm{e}^{iz} - \mathrm{e}^{-iz}}{2i}.$$

Ist z reell, so besagt dies, daß $\cos x$ und $\sin x$ gerade gleich dem Realteil bzw. Imaginärteil von e^{ix} sind. Die Eulersche Identität ist eine Quelle vieler weiterer Identitäten: Die gesamten Additionstheoreme sämtlicher aus der Schule bekannten trigonometrischen Funktionen oder ihrer Umkehrfunktionen sind lediglich eine einfache Folgerung der Eulerschen Identität und der Funktionalgleichung der Exponentialfunktion. Um sich dies klar vor Augen zu führen, sollte man zur Übung die oben formulierten Additionstheoreme für sin und cos aus der Eulerschen Identität und dem Additionstheorem der Exponentialfunktion ableiten (man kann im Wesentlichen den oben gegebenen Beweis für die Additionstheoreme der hyperbolischen Funktionen abschreiben).

Die Exponentialfunktion definiert einen Gruppenhomomorphismus von der Gruppe der komplexen Zahlen (versehen mit der Addition) zur Gruppe $\mathbb{C}^*$ (versehen mit der Multiplikation). Dann definiert natürlich auch die Zuordnung

$$z \mapsto \mathrm{e}^{iz}$$

eine Gruppenhomomorphismus dieser Gruppen. Was ist das Bild der reellen Zahlen unter letzterem Homorphismus ? Es ist zunächst für reelle x

$$\begin{aligned} \overline{\exp(ix)} &= \overline{\cos x + i \sin x} = \cos x - i \sin x \\ &= \cos(-x) + i \sin(-x) = \exp(-ix), \end{aligned}$$

und daher

$$\left|\mathrm{e}^{ix}\right|^2 = \exp(ix)\overline{\exp(ix)} = \exp(ix)\exp(-ix) = 1.$$

Also definiert die Zuordnung

$$x \mapsto \mathrm{e}^{ix}$$

bei Beschränkung auf reelle x einen Gruppenhomomorphismus

$$\mathbb{R} \to \mathbb{S}^1 := \{z \in \mathbb{C} |\, |z| = 1\}.$$

(Man überzeuge sich, daß die Menge $\mathbb{S}^1$ bezüglich der Multiplikation komplexer Zahlen tatsächlich eine Gruppe ist.) Dies ist eine andere — allerdings wesentlich bedeutungsvollere — Sichtweise unserer oben betrachten Abbildung $x \mapsto (\cos x, \sin x)$. Was ist der Kern unseres Homomorphismus $x \mapsto \mathrm{e}^{ix}$? Es wird sich im nächsten Kapitel heraustellen — unsere bisherigen Methode reichen noch nicht, um dies hier schon einzusehen —, daß dieser Kern eine Untergruppe von $\mathbb{R}$ der Gestalt

$$d \cdot \mathbb{Z} = \{d \cdot z | z \in \mathbb{Z}\}$$

ist, wo d eine positive reelle Zahl ist. Offenbar ist d hierdurch eindeutig bestimmt. Angesichts des dargestellte Zusammenhangs mit der Exponentialfunktion und den trigonometrischen Funktionen ist klar, daß dieses d eine fundamentale Größe ist. Wir führen daher für diese d, genauer für $\frac{d}{2}$ ein eigenes Symbol ein:

$$\pi := \frac{d}{2},$$

oder prägnanter

Die Zahl π

> Der Kern des Gruppenhimomorphismus
>
> $$\mathbb{R} \to \mathbb{S}^1,\ x \mapsto \mathrm{e}^{2ix}$$
>
> besteht aus den ganzzahligen Vielfachen einer positiven reellen Zahl. Diese hierdurch eindeutig charakterisierte reelle Zahl wird mit π bezeichnet.

Weitere grundlegende Begriffe aus der Theorie der Folgen

Als Vorbereitung für das nächste Kapitel sind noch einige allgemeine Tatsachen aus der Theorie der Folgen nachzutragen. Der Begriff des Grenzwertes einer Folge $(a_n)_{n\in\mathbb{N}}$ von reellen Zahlen a_n ist für manche Betrachtungen etwas zu eng ausgelegt. Ein weniger enger Begriff ist der Begriff des Häufungspunktes einer Folge.

Häufungspunkt

> Eine reelle Zahl a heißt Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$, falls jede ϵ–Umgebung für unendlich viele n das Folgenglied a_n enthält.

Unter der ϵ–Umgebung einer reellen Zahl a verstehen wir hierbei die Menge

$$\{x \in \mathbb{R} | a - \epsilon < x < a + \epsilon\},$$

wobei wir stets voraussetzen, daß ϵ eine strikt positive Zahl ist. Die Sprechweise

"für unendlich viele n gilt die Eigenschaft $A(n)$"

steht hierbei als Abkürzung für diejenige Aussage, die in mathematischer Kurzschreibweise durch

$$\forall n_0 \in \mathbb{N} \exists n \in \mathbb{N}(n \geq n_o \wedge A(n))$$

wiedergegeben wird. Wir werden im Folgenden daneben noch die schon erklärte Sprechweise

"für fast alle n gilt $A(n)$"

benutzen, die für

$$\exists n_0 \in \mathbb{N} \forall n \in \mathbb{N}(n \geq n_0 \Longrightarrow A(n))$$

steht. Man erinnere sich, daß eine Zahl a Grenzwert der Folge $(a_n)_{n\in\mathbb{N}}$ genannt wird, falls jede ϵ–Unmgebung von a für fast alle n das Folgenglied a_n enthält. Danach ist klar, daß der Grenzwert einer Folge insbesondere ein Häufungspunkt dieser Folge ist. Die Umkehrung ist nicht richtig: Die Folge

$$1, -1, 1, -1, 1, -1, \ldots$$

hat offenbar zwei Häufungspunkte — $+1$ und -1 — dagegen hat sie keinen Grenzwert. Dies Beispiel zeigt auch, daß eine Folge durchaus mehrere Häufungspunkte haben kann, wogegen sie ja höchstens einen Grenzwert besitzt. Der genaue Zusammenhang zwischen dem Begriff Grenzwert und Häufungspunkt gibt der nächste Satz wieder.

Satz. *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen, a eine reelle Zahl. Dann sind die beiden folgenden Aussagen äquivalent:*

(i) *a ist Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$,*

(ii) *es gibt eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$, die gegen a konvergiert.*

Beweis. Da der Grenzwert einer Teilfolge insbesondere Häufungspunkt dieser Teilfolge, und der Häufungspunkt einer Teilfolge erst recht Häufungspunkt der gesamten Folge ist, so ist klar, daß (ii) die Aussage (i) impliziert. Sei umgekehrt die Aussage (i) wahr. Wir konstruieren dann folgendermäßen induktiv eine gegen a konvergente Teilfolge: sei a_{k_1} irgendein Folgenglied in der 1-Umgebung von a (diese enthält ja a_n für unendlich viele n). Es seien die Folgenglieder k_j für $j < n$ schon definiert. Dann gibt es in der $\frac{1}{n}$–Umgebung von a ein a_k für unendlich viele k, insbesondere gibt es in dieser Umgebung dann ein a_k mit $k > k_{n-1}$; wir setzen $k_n := k$. Nach Konstruktion ist klar, daß die Folge der k_n streng monoton steigt, und daß die Teilfolge $(a_{k_n})_{n\in\mathbb{N}}$ gegen a konvergiert.

Eine grundlegende Existenzaussage ist der

Satz (von Bolzano-Weierstrass). *Jede beschränkte Folge hat wenigstens einen Häufungspunkt.*

Beweis. Die Folge $(a_n)_{n\in\mathbb{N}}$ sei beschränkt, d.h. es gibt Zahlen $c < d$, sodaß $c \leq a_n \leq d$ für alle $n \in \mathbb{N}$ ist. Das Intervall $[c, d]$ teilen wir nun in zwei Hälften durch Einführung des Mittelpunkts $\frac{c+d}{2}$. Dann ist für unendlich viele n entweder $c \leq a_n \leq \frac{c+d}{2}$, oder aber jedenfalls $\frac{c+d}{2} \leq a_n \leq d$ für unendlich viele n. Trifft ersteres zu , so setzen wir $c_1 := c$, $d_1 = \frac{c+d}{2}$, trifft das zweite zu, so setzen wir $c_1 := \frac{c+d}{2}, d_1 := d$. Nun wiederholen wir diesen Prozeß für das Intervall $[c_1, d_1]$ an Stelle von $[c, d]$: Es gilt für unendlich viele n ist $c_1 \leq a_n \leq \frac{c_1+d_1}{2}$ oder für unendlich viele n ist $\frac{c_1+d_1}{2} \leq a_n \leq d$. Wir setzen $c_2 := c_1, d_2 := \frac{c_1+d_1}{2}$ im ersten Fall und sonst $c_2 = \frac{c_1+d_1}{2}, d_2 = d_1$. Auf diese Weise konstruieren wir induktiv eine Intervallschatelung, denn nach Konstruktion ist zum einen

$$c = c_0 \leq c_1 \leq c_2 \leq \ldots \leq d_2 \leq d_1 \leq d_0 = d_1,$$

und zum anderen

$$d_n - c_n = \frac{1}{2^n}(d - c).$$

Die beiden Folgen $(c_n)_{n\in\mathbb{N}}$ und $(d_n)_{n\in\mathbb{N}}$ konvergieren daher, und zwar gegen ein und denselben Grenzwert a. Aber dann ist a ein Häufungspunkt unserer Folge. Sei nämlich ein $\epsilon > 0$ gegeben. Wähle dann k so, daß$d_k - c_k = \frac{d-c}{2^k} < \epsilon$. Nach Konstruktion liegt für unendliche viele n das Folgenglied a_n im Intervall $[c_k, d_k]$. In diesem Intervall liegt aber auch der Grenzwert $a = \lim_{n\to\infty} c_n = \lim_{n\to\infty} d$, da $(c_n)_{n\in\mathbb{N}}$ aufsteigt und $(d_n)_{n\in\mathbb{N}}$ fällt. Folglich ist tatsächlich $|a - a_n| \leq \frac{d-c}{2k} < \epsilon$ für diese unendlich vielen n.

Satz. *Eine Folge ist dann und nur dann konvergent, falls sie beschränkt ist und genau einen Häufungspunkt besitzt.*

Beweis. Eine konvergente Folge ist beschränkt und ihr Grenzwert a ist ein Häufungspunkt. Da jeder Häufungspunkt Grenzwert einer Teilfolge ist, und da alle Teilfolgen einer konvergenten Folge gegen a konvergieren, kann es keine weiteren Häufungspunkte geben. Sei umgekehrt die Folge $(a_n)_{n\in\mathbb{N}}$ beschränkt. Dann hat sie einen Häugungspunt a. Wir nehmen an, daß die Folge $(a_n)_n$ nicht gegen a konvergiert; wir haben zu zeigen daß sie dann mindestens einen weiteren, von a verschiedenen Häufungspunkt besitzt. Da also $(a_n)_n$ nicht gegen a konvergiert, gibt es ein $\epsilon > 0$, sodaß $|a - a_m| \geq \epsilon$ für unendlich viele m zutrifft. Sei dann $(k_n)_{n\in\mathbb{N}}$ eine aus diesen unendlich vielen m gebildetet streng monotonen Folge. Mit $(a_n)_n$ ist auch die Teilfolge $(a_{k_n})_n$ beschränkt, also hat sie einen Häufungspunkt b, der dann auch Häufungspunkt von $(a_n)_n$ ist. Wir zeigen, daß $b \neq a$. Da $|a - a_{k_n}| \geq \epsilon$ ja für alle n, und $|b - a_{k_n}| < \frac{\epsilon}{2}$ für unendlich viele n gilt, gibt es mindesten ein n, für das beide Ungleichungen zutreffen. Mit diesem n ist dann

$$|a - b| = |a - a_{k_n} - (b - a_{k_n})| \geq |a - a_{k_n}| - |(b - a_{k_n})| \geq \epsilon - \frac{\epsilon}{2}.$$

Dies zeigt Behauptung.

Das folgende Konvergenzkriterium hat seine Bedeutung darin, daß es einem ein Mittel in die Hand gibt, die Konvergenz einer Folge zu zeigen, ohne den Grenzwert benennen zu müßen.

Satz (Cauchy Kriterium). *Eine Folge* $(a_n)_{n\in\mathbb{N}}$ *ist genau dann konvergent, falls gilt:*

$$\forall\epsilon > 0 \exists n_0 \in \mathbb{N} \forall m, n \in \mathbb{N}(m, n > n_0 \Rightarrow |a_n - a_m| < \epsilon)$$

Beweis. Die Folge der a_n sei konvergent gegen a. Zu gegebenem $\epsilon > 0$ kann man dann ein n_0 finden, sodaß

$$|a_n - a| < \frac{\epsilon}{2}$$

für alle $n \geq n_0$. Sind daher $n, m \geq n_0$, so hat man

$$|a_n - a_m| \leq |a_n - a| + |a - a_m| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Umgekehrt treffe das Cauchysche Konvergenzkriteriun zu. Dann ist $(a_n)_{n\in\mathbb{N}}$ jedenfalls beschränkt: zu $\epsilon = 17$ gibt es ein $n_0 \in \mathbb{N}$ sodaß für alle $n > n_0$ und alle $m > n_0$ dann $|a_n - a_m| < 17$ gilt; insbesondere liegt für alle $n > n_0$ das Folgenglied a_n in der 17–Umgebung von a_{n_0+1}, und dies impliziert sofort die Beschränktheit. Die Folge $(a_n)_{n\in\mathbb{N}}$ hat daher nach dem Satz von Bolzano-Weierstrass mindestens einen Häufungspunkt a. Dieser ist aber dann auch Grenzwert unserer Folge: zu gegebenem $\epsilon > 0$ gibt es ein n_1, sodaß $|a_m - a_n| < \frac{\epsilon}{2}$ für alle $n, m > n_1$, und zu n_0 gibt es ein $n_1 > n_0$, sodaß $|a - a_{n_1}| < \frac{\epsilon}{2}$ ist (ersteres nach dem Cauchyschen Konvergenzkriterium, letzteres, da a ja Häufungspunkt ist); damit ist für alle $n > n_0$

$$|a - a_n| = |a - a_{n_1} + a_{n_1} - a_n| \leq |a - a_{n_1}| + |a_{n_1} - a_n| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

und das war zu beweisen.

Es ist instruktiv, das Cauchysche Konvergenzkriteriums speziell für Reihen zu formulieren:

Satz. *Die Reihe* $\sum_{n=0}^{\infty} a_n$ *ist genau dann konvergent, wenn gilt*

$$\forall\epsilon > 0\, \exists n_0 \in \mathbb{N}\, \forall n, m \in \mathbb{N}(m > n > n_0 \Rightarrow |a_{n+1} + \ldots + a_m| < \epsilon).$$

Als Illustration der Nützlichkeit des Cauchy Kriteriums geben wir einen neuen Beweis der Tatsache, daß eine absolut konvergente Reihe auch konvergiert: in der Tat impliziert ja

$$||a_n| + \ldots + |a_m|| < \epsilon$$

via der Dreiecksungleichung

$$|a_n + \ldots a_m| < \epsilon,$$

sodaß, falls das Cauchy Kriterium auf die aus den $|a_n|$ gebildete Reihe zutrifft, es erst recht auf die aus den a_n gebildete zutrifft.

7 STETIGKEIT

Im zweiten Kapitel haben wir den Begriff der Abbildung erklärt:

$$f\colon M \to N,\ x \mapsto f(x)$$

steht für eine Zuordnung, die jedem Element der Menge M genau ein Element $f(x)$ der Menge N zuordnet. In den folgenden Kapiteln werden wir spezielle Abbildungen studieren:

Funktionen

> Unter einer *auf D definierten (reellwertigen) Funktion f* verstehen wir eine Abbildung $f\colon D \to B$, wo B eine Teilmenge von $\mathbb{R}$ ist.

Intervalle

Die Menge D heißt Definitionsbereich von f. Die Definitionsbereiche, die im Folgenden am häufigsten vorkommen werden, sind die *Intervalle*

$$\begin{aligned}(a,b) &:= \{x \in \mathbb{R} | a < x < b\},\\ [a,b] &:= \{x \in \mathbb{R} | a \le x \le b\},\\ (a,b] &:= \{x \in \mathbb{R} | a < x \le b\},\\ [a,b) &:= \{x \in \mathbb{R} | a \le x < b\}.\end{aligned}$$

Hierbei bedeuten a, b irgendwelche reellen Zahlen. Die aufgelisteten Intervalle bezeichnet man gelegentlich genauer als *offen, abgeschlossen, links halboffen* bzw. *rechts halboffen*. Daneben betrachtet man auch Intervalle der Gestalt

$$\begin{aligned}(a,+\infty) &:= \{x \in \mathbb{R} | a < x\},\\ (-\infty,a) &:= \{x \in \mathbb{R} | x < a\},\\ [a,+\infty) &:= \{x \in \mathbb{R} | a \le x\},\\ (-\infty,a] &:= \{x \in \mathbb{R} | x \le a\}.\end{aligned}$$

Die Symbole $\pm\infty$ sind hier nur eine Schreibweise ohne weitere Bedeutung. Spezielle Intervalle, die wir schon wiederholt benutzt haben, sind die ϵ–Umgebungen einer reellen Zahl a:

$$(a-\epsilon, a+\epsilon) = \{x \in \mathbb{R} |\, |a-x| < \epsilon\}.$$

Beispiele für Funktionen, die wir schon kennen, sind exp, sinh, cosh, sinh, und cosh; die sind jeweils Funktionen auf $\mathbb{R}$ (also $D = \mathbb{R}$). Als Beispiele für nicht auf ganz $\mathbb{R}$ definierte Funktionen haben wir schon log und $\log_a$ kennengelernt; hier ist $D = \mathbb{R}_+ = (0, \infty)$.

Im Grunde ungenau, aber der Bequemlichkeit halber doch üblich ist die Schreibweise $f(x)$ für eine Funktion f. Man sagt etwa "... die Funktion $\mathrm{e}^{\cos x}$ ...", und meint damit genauer "die durch die Zuordnung $x \mapsto \mathrm{e}^{\cos x}$ auf der Menge aller x, wo diese Zuordnung Sinn macht, definierte Funktion ...", oder man sagt "... die Funktion $f(x)/g(x)$ ..." und meint damit "... die durch die Zuordnung $x \mapsto f(x)/g(x)$ auf der Menge derjenigen reellen Zahlen, wo $f(x)$ und $g(x)$ erklärt sind und $g(x) \neq 0$ ist, definierte Funktion .... In einem anderen Zusammenhang kann natürlich — wenn x für irgendeine durch den Kontext festgelegte Zahl steht — das Symbol $\mathrm{e}^{\cos x}$ einfach die Zahl $\mathrm{e}^{\cos x}$ bedeuten, d.h. den Wert der eben erklärten Funktion für das Argument x.

Polynome

Weitere Beispiele für Funktionen sind die Polynome, die in vielerlei Hinsicht als Modell für Begriffsbildungen in der Theorie von Funktionen dienen. Ein *Polynom* (genauer: *ein Polynom mit reellen Koeffizienten*) ist ein symbolischer Ausdruck der Gestalt

$$a_0 + a_1X + ... + a_nX^n,$$

wo n irgendeine natürliche Zahl ist, und die a_k reelle Zahlen bedeuten. Was dabei X ist, lassen wir an dieser Stelle völlig offen. Hier ist es erstmal nur ein Symbol, eine Unbestimmte. Zwei solche Ausdrücke definieren das gleiche "Objekt", falls alle ihre Koeffizienten übereinstimmen. Wer sich durch diese Erklärung verunsichert fühlt und eine klare Definition mittels naiver Mengenlehre vorzieht, mag ein Polynom als eine Folge $(a_n)_n \in \mathbb{N}$ ansehen, wo für fast alle n das Folgenglied a_n gleich 0 ist, und die oben eingeführten "Ausdrücke" sind nur eine symbolische Schreibweise für solch eine spezielle Folge, wobei wir vereinbaren, daß man Symbole der Form "$0 \cdot X^i$" auch einfach weglassen kann. Die Menge aller Polynome mit reellen Koeffizienten wird üblicherweise mit $\mathbb{R}[X]$ bezeichnet. Das größte aller k, für das $a_k \neq 0$ ist, heißt *Grad des Polynoms* $a_0 + a_1X + ... + a_n \cdot X^n$. Dies ist wohldefiniert, d.h. solch ein maximales k existiert stets abgesehen von dem Fall, wo überhaupt kein Koeffizient von 0 verschieden ist: diesem *identisch verschwindenden Polynom* (oder auch *Nullpolynom*) ordnen wir hier keinen Grad zu.

Die Gesamtheit aller Polynome $R[X]$ wird zu einem Integritätsbereich, d.h. zu einem kommutativen Ring (mit 1) ohne Nullteiler, wenn man folgende Operationen einführt:

$$\begin{aligned}
&(a_0 + a_1X + ... + a_nX^n) + (b_0 + b_1X + ... + b_mX^m) \\
&:= (a_0 + b_0) + (a_1 + b_1)X + \cdots, \\
&(a_0 + a_1x + ...a_nX^n) \cdot (b_0 + b_1X + ... + b_mX^m) \\
&:= c_0 + c_1X + ...c_{n+m}X^{n+m},
\end{aligned}$$

wobei

$$c_k := \sum_{i=0}^{k} a_i \cdot b_{k-i}.$$

Die hier erklärte Multiplikation ist formal identisch mit dem im letzten Kapitel erklärten Cauchy-Produkt von unendlichen Reihen. Der eben erwähnte Begriff *Nullteiler* bedeutet, daß das Produkt von zwei Polynomen höchstens dann gleich dem Nullpolynom ist, wenn mindestens eines der Polynome selbst das Nullpolynom ist. Das dies tatsächlich der Fall ist, macht man sich leicht klar, indem man sich überlegt, daß sich bei Multiplikation die Grade addieren.

Ein Polynom $p = a_0 + a_1 \cdot X + ... + a_n \cdot X^n$ aus $\mathbb{R}[X]$ liefert offensichtlich eine Funktion $\mathbb{R} \to \mathbb{R}$, wenn wir jedem x in $\mathbb{R}$ die Zahl

$$p(x) := a_0 + a_1 x + ... + a_n x^n \in \mathbb{R},$$

zuordnen, die man erhält, indem man die Unbestimmte X durch die Zahl x ersetzt. Man beachte, daß "$p(x)$" hier lediglich eine abkürzende und naheliegende Schreibweise ist; ein Polynom an sich ist keine Abbildung. Eine Zahl $\alpha \in \mathbb{R}$ heißt Nullstelle des Polynoms $a_0 + a_1 \cdot X + ... + a_n \cdot X^n$, falls

Nullstellen von Polynomen

$$a_0 + a_1 \cdot \alpha + ... + a_n \cdot \alpha^n = 0$$

ist. Für Nullstellen gilt der

Satz. *Sei α Nullstelle des Polynoms p in $\mathbb{R}[X]$. Dann gibt es ein $q \in \mathbb{R}[X]$, sodaß*

$$p = (X - \alpha) \cdot g.$$

Jedes Polynom p mit α als Nullstelle kann man also durch das "einfachste" Polynom, welches α als Nullstelle hat — nämlich $X - \alpha$ —, im Ring $\mathbb{R}[X]$ teilen.

Beweis. Für das Polynom p — etwa $p = a_o + a_1 X + \cdots + a_n X^n$ — hat man unter Beachtung, daß α eine seiner Nullstellen ist

$$\begin{aligned} p &= (a_0 + a_1 \cdot X + ... + a_n X^n) - (a_0 + a_1 \alpha + ... + a_n \alpha^n) \\ &= a_1 (X - \alpha) + a_2 \big(X^2 - \alpha^2\big) + ... + a_n (X^n - \alpha^n). \end{aligned}$$

Setzt man hierin die Identitäten

$$X^k - \alpha^k = (X - \alpha)\big(X^{k-1} + \alpha x^{k-2} + a^2 X^{k-3} + \cdots + \alpha^{k-1}\big)$$

ein, so ist die Behauptung ersichtlich.

Als Folge des soeben bewiesenen Satzes erhält man

Satz. *Ein Polynom vom Grad n hat höchstens n verschiedene Nullstellen.*

Beweis. Sei p ein Polynom vom Grad n. Hat p keine Nullstelle, so ist nichts zu beweisen. Ist dagegen etwa α eine Nullstelle, so kann man schreiben $p = (X - \alpha)q$ mit einem geeigneten Polynom q. Wie man sich leicht überlegt, ist der Grad von q aber $n - 1$. Damit ist klar, daß die Behauptung sofort vermöge vollständiger Induktion über den Grad n folgt.

Satz. *Seien p, q zwei Polynome, die dieselbe Funktion definieren. Dann sind p und q als Polynome gleich, d.h. sämtliche Koeffizienten von p und q stimmen überein.*

Beweis. Das Differenzpolynom $p - q$ hat nach Voraussetzung jede Zahl α als Nullstelle. Also kann es nach dem letzten Satz keinen Grad n haben, d.h. es muß das Nullpolynom sein.

Solange der Grundbereich $\mathbb{R}$ ist, besagt der letzte Satzes, daß man zwischen dem Polynom p in $\mathbb{R}[X]$ und der zugehörigen, auf $\mathbb{R}$ definierten Funktion nicht wirklich zu unterscheiden braucht: man kann ein Polynom und die dazugehörige Abbildung miteinander *identifizieren*; d.h. die Zuordnung

$$\text{Polynom} \mapsto \text{Abbildung}$$

ist injektiv. Ist p ein Polynom, so sagt man daher einfach "...das Polynom $p(x)$ ...", und meint damit "die durch das Polynom p auf $\mathbb{R}$ definierte Abbilding $x \mapsto p(x)$ ...".

Ganz wesentlich war im Beweis des letzten Satz die Tatsache, daß es unendlich viele reelle Zahlen gibt. Tatsächlich ist die Situation bei Polynomen mit Koeffizienten in endlichen Körpern (z.B. $\mathbb{Z}/2\mathbb{Z}$) vollkommen anders: die Polynome

$$X,\ X^2,\ X^3, \ldots$$

definieren dort alle ein- und dieselbe Abbildiung

$$\mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}.$$

Der vorletzte Satz über die Teilbarkeit durch $X-$Nullstelle gilt dagegen sinngemäß in jedem Körper.

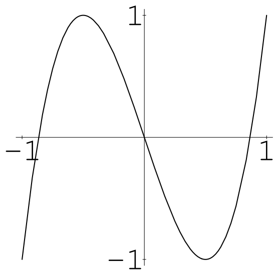
Rationale Funktionen

Mittels der Polynome gelangt man sofort zu einer weiteren Klasse von Funktionen: zu den rationalen Funktionen. Sind f und g Polynome (gemäß unserer eben gemachten Verabredung, Polynome und die durch sie definierten Abbildungen zu identifizieren, benutzen wir konsequenterweise von jetzt ab diejenigen Symbole für Polynome, die wir auch für Funktionen benutzen), und sei g nicht das Nullpolynom. Dann ist offenbar $f(x)/g(x)$ eine Funktion mit Definitionsbereich $\mathbb{R} \setminus \{\alpha \in \mathbb{R} | g(\alpha) = 0\}$. Funktionen von diesem Typus heißen rational.

Graphische Darstellungen von Funktionen sind jedem geläufig. Als *Graph* G_f einer auf D erklärten Funktion bezeichnet man die Menge

$$G_f = \{(x, f(x))|x \in D\}.$$

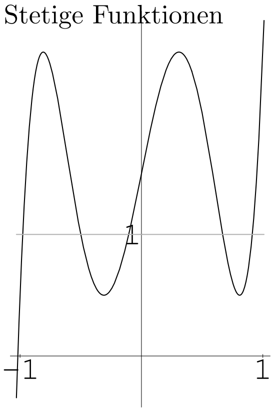
Dies ist eine Teilmenge des $\mathbb{R}^2$, und in einem kartesischen Koordinatensystem kann man zu konkret gegebenem f einige Punkte von G_f eintragen und diese durch irgendwelche passenden Geraden -oder Kurvenstücke verbinden. Links ist eine Skizze des Graphen des Polynoms $(4x^2-3)x$. Die zweite Skizze kann dagegen unmöglich den Graphen eines Polynoms $f(x)$ dritten Grades darstellen: $f(x)$ müßte etwa fünfmal den Wert 1 annehmen; dagegen hat $f(x)-1$ als Polynom dritten Grades höchstens drei Nullstellen, und drei ist zwei weniger als fünf. Tatsächlich zeigt die zweite Skizze den Graphen des Polynoms $(16x^4 - 20x^2 + 5)x + \frac{3}{2}$. Der Graph der Funktion



$$f(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q} \\ 0 & \text{für } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

läßt sich dagegen weniger gut darstellen. Der Punkt ist, daß die ersten beiden Beispiele stetige Funktionen wiedergaben, dagegen die soeben definierte Funktion nirgendwo stetig ist.

Stetige Funktionen



Diesen Begriff wollen wir nun präzisieren.

> Eine auf D erklärte Funktion f heißt stetig in einem Punkt $c \in D$, falls für jede Folge $(x_n)n \in \mathbb{N}$, sodaß $x_n \in D$ für alle $n \in \mathbb{N}$ und
> $$\lim_{n\to\infty} x_n = c$$
> gilt, auch stets
> $$\lim_{n\to\infty} f(x_n) = f(c)$$
> ist. Eine auf einer Menge D erklärte Funktion f heißt stetig auf D', falls $D' \subset D$ gilt, und sie in jedem $c \in D'$ stetig ist. Eine auf D erklärte Funktion f heißt stetig, falls sie in jedem $c \in D$ stetig ist.

Offenbar ist eine stetige Funktion auf jeder Teilmenge ihres Definitionsbereiches stetig. Wir betrachten einige Beispiele zum Begriff der Stetigkeit: Die Funktionen

$$\mathbb{R} \to \mathbb{R}, \quad x \mapsto x$$

(die *Identität auf* $\mathbb{R}$), ist stetig in jedem $c \in D = \mathbb{R}$, wie unmittelbar aus der Definition der Stetigkeit abzulesen ist. Die Funktion

$$f(x) = \begin{cases} 1 & \text{für } x \geq 0 \\ -1 & \text{für } x < 0 \end{cases}.$$

ist stetig an jeder Stelle $c \neq 0$ und unstetig (d.h. nicht stetig) an der Stelle $c = 0$. Zu letzterem: die Folge $\left((-)^n \frac{1}{n}\right)_{n\in\mathbb{N}}$ konvergiert gegen 0, aber die Folge der Bildwerte ist $1, -1, 1, -1...$, mithin nicht konvergent. Man beachte, daß in der Definition etwas für *jede* Folge — nicht etwa etwas für nur *eine* Folge — gefordert wird. Im letzten Beispiel gibt es durchaus Folgen, die gegen 0 konvergieren, und wo Folge die Bildwerte unter f dies auch tut (so etwa die Folge $\left(\frac{1}{n}\right)_n$. Ein weiteres einfaches Beispiel: die konstante Funktion a (d.h. die Abbildung $x \mapsto a$) ist überall stetig. Vielleicht nicht völlig offensichtlich ist, daß die unmittelbar vor der Definition der "Stetigkeit in einem Punkt" erklärte Funktion in keinem Punkt stetig ist.

Auf die Stetigkeit eines ganzen Bündels von Funktionen kann man mittels des folgenden Satzes schließen.

Satz. *Die auf D erklärten Funktionen f und g seien stetig in $c \in D$. Dann sind auch $f+g$, $f-g$, $f \cdot g$ stetig in c. Ferner ist auch f/g stetig in c, falls $g(x) \neq 0$ für $x \in D$. Hierbei bezeichnet $f+g$, $f \cdot g$ etc. die auf D erklärte Funktion $x \mapsto f(x) + g(x)$, bzw. $f(x) \cdot g(x)$ etc..*

Beweis. . Sei eine Folge $(x_n)_{n \in \mathbb{N}}$ von Punkten x_n aus D gegeben, sodaß $\lim_{n\to\infty} x_n = c$. Da f und g stetig in c sind, hat man $\lim_{n\to\infty} f(x_n) = f(c)$, $\lim_{n\to\infty} g(x_n) = g(c)$. Nach dem Hauptsatz über konvergente Folgen folgt

$$\lim_{n\to\infty} (f+g)(x_n) = \lim_{n\to\infty} f(x_n) + \lim_{n\to\infty} g(x_n) = f(c) + g(c).$$

Dies beweist die Stetigkeit von $f+g$ in c. Ganz analog impliziert der Haupsatz über konvergente Folgen die Behauptungen für $-$, "$\cdot$", "/" an Stelle von "+".

Nach dem eben bewiesenen Satz haben wir insbesondere

Satz. *Jedes Polynom definiert eine auf ganz $\mathbb{R}$ stetige Funktion. Jede rationale Funktion ist in jedem Punkt, wo sie definiert ist, stetig.*

Beweis. Ein Polynom $a_0 + a_1 x + ... + a_n x^n$ wird durch Multiplikation und Addition aus den konstanten Funktionen $a_0, ..., a_n$ und der Identität $x \mapsto x$ gebildet. Also folgt die Stetigkeit eines Polynoms aus dem letzten Satz. Damit ist dann aber auch einer rationale Funktion als Quotient zweier Polynome nach dem letzten Satz stetig.

Im Beweis von allgemeinen Aussagen über stetige Funktionen ist manchmal eine etwas andere Charakterisierung der Stetigkeit nützlich.

Satz. *Eine auf D erklärte Funktion f ist dann und nur dann stetig an der Stelle $c \in D$, wenn gilt:*

$$\forall \epsilon > 0 \, \exists \delta > 0 \, \forall x \in D (|x - c| < \delta \quad \Rightarrow \quad |f(x) - f(c)| < \epsilon).$$

In Worte gefaßt besagt dieser Satz also, f ist genau dann in $c \in D$ stetig, wenn zu jedem $\epsilon > 0$ ein $\delta > 0$ existiert, sodaß

$$f(\delta - \text{Umgebung von } c) \subset \epsilon - \text{Umgebung von } f(c)$$

ist. Wir wissen nach dem vorletzten Satz, daß die Funktion $f(x) = x^2$ in 0 stetig ist; doch überprüfen wir dies anhand des eben formulierten Kriteriums: sei $\epsilon > 0$ gegeben; wir suchen ein $\delta > 0$, sodaß $|f(x) - f(0)| = x^2 < \epsilon$, wenn nur $|x - 0| = |x| < \delta$ ist. Offenbar genügt es dafür, $\delta := \sqrt{\epsilon}$ zu wählen. Daß dies sogar das kleinste mögliche δ ist, spielt hierbei natürlich überhaupt keine Rolle. Zum Beweis des Satzes.

Beweis. "$\Leftarrow$": Sei $(x_n)_{n\in\mathbb{N}}$ gegeben, sodaß $x_n \in D$ $(n \in \mathbb{N})$ und sodaß $\lim_{n\to\infty} x_n = c$. Zu zeigen haben wir $\lim_{n\to\infty} f(x_n) = f(c)$, d.h. für jedes $\epsilon > 0$ ist $|f(x_n) - f(c)| < \epsilon$ für fast alle $n \in \mathbb{N}$. Sei also ein $\epsilon > 0$ gegeben. Nach Voraussetzung gibt es dazu ein $\delta > 0$, sodaß $|f(x) - f(c)| < \epsilon$ wenn nur $|x - c| < \delta$. Nun ist aber (x_n) konvergent gegen c, d.h. es ist $|x_n - c| < \delta$ für fast alle n; und damit dann auch $|f(x_n) - f(c)| < \epsilon$ für fast alle n.

"$\Rightarrow$" Bekanntlich ist eine Aussage "$a \Rightarrow b$" äquivalent zu ihrer Kontraposition "$\neg b \Rightarrow \neg a$". Es genügt daher, wenn wir aus der Negation der rechten Seite der Äquivalenz des Satzes auf die Unstetigkeit im Punkte c schließen können. Die Negation der rechten Seite lautet:

$$\exists \epsilon > 0 \,\forall \delta > 0 \,\exists x \in D(|x - c| < \delta \wedge |f(x) - f(c)| \geq \epsilon).$$

Sei $\epsilon > 0$ irgendeines der ϵ, deren Existenz hier behauptet wird. Zu jedem $\delta > 0$, insbesondere zu jedem $\frac{1}{n}$ $(n \in \mathbb{N})$ gibt es also ein $x_n \in D$, sodaß $|x_n - c| < \frac{1}{n}$ und $|f(x_n) - f(c)| \geq \epsilon$. Die Folge der x_n konvergiert somit gegen c, wogegen der Folge der $f(x_n)$ stets im Abstand $\geq \epsilon$ zu $f(c)$ bleibt. Danach kann f in c nicht stetig sein.

Wir kommen nun zu den ersten nicht offensichtlichen Eigenschaften stetiger Funktionen.

Satz. *Die auf D definierte Funktion f sei stetig an der Stelle $c \in D$, und es gelte $f(c) > 0$. Dann gibt es ein $\delta > 0$, sodaß $f(x) > 0$ für alle x aus $D \cap (a - \delta, a + \delta)$.*

Beweis. Wir wählen in dem Stetigkeitskriterium des letzten Satzes $\epsilon = f(c)$. Hierzu gibt es ein $\delta > 0$, sodaß für alle $x \in D$ mit $|x - c| < \delta$ die Ungleichung

$$|f(x) - f(c)| < \epsilon = f(c)$$

gilt. Diese Ungleichung ist aber äquivalent zu

$$0 < f(x) < 2f(c),$$

womit der Satz bewiesen ist.

Ganz analog gilt natürlich die Aussage: "Ist$f\colon D \to \mathbb{R}$ stetig an der Stelle $c \in D$ und $f(c) < 0$, dann gibt es ein $\delta > 0$ sodaß $f(x)$ für jedes $x \in D \cap (a-\delta, a+\delta)$ negativ ist."

Wir kommen nun zu den drei Haupsätzen der Theorie der stetigen Funktionen. Der erste lautet:

Stetige Funktionen sind auf abgeschlossenen Intervallen beschränkt

Satz. *Sei f eine auf dem abgeschlossenen Intervall $[a, b]$ stetige Funktion. Dann ist f beschränkt auf $[a, b]$, d.h. es gibt eine Zahl K, sodaß $|f(x)| < K$. für alle $a \leq x \leq b$ ist.*

Beweis. Wir beweisen die Kontraposition der Behauptung. Angenommen f ist nicht beschränkt. Dann gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in [a, b]$ sodaß $|f(x_n)| \geq n$. Die so gefundene Folge $(x_n)n \in \mathbb{N}$ liegt im Intervall $[a, b]$, ist daher beschränkt und hat somit nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge, etwa $\lim_{k\to\infty} x_{n_k} = c$. Wir werden gleich sehen, daß $a \leq c \leq b$ liegt, d.h. daß c im Definitionsbereich von f liegt. Aber $f(x_{n_k}) > n_k$, d.h. die Bildfolge der x_{n_k} unter f kann nicht gegen $f(c)$ konvergieren. Also ist f nicht stetig in c. Wir haben noch den Beweis nachzutragen, daß das hier konstruierte c tatsächlich die Ungleichung $a \leq c \leq b$ erfüllt. Dies ist offensichtlich ein Spezialfall des folgenden allgemeinen Sachverhalts.

Satz. *Sei $(a_n)_{n\in\mathbb{N}}$ eine gegen die Zahl α konvergente Folge. Sind alle $a_n \geq a$, so ist $\alpha \geq a$; sind alle $a_n \leq b$, so ist $\alpha \leq b$.*

Beweis. Angenommen $\alpha < a$. Dann gilt für fast alle Glieder der Folge $|a_n - \alpha| < \epsilon := a - \alpha$; aber diese Ungleichung impliziert $a_n < a$. Also sind fast alle a_n negativ — im Widerspruch zur Voraussetzung. Den Fall von $a_n \leq b$ behandelt man analog.

Wir betonen noch einmal, daß die Voraussetzung im ersten Hauptsatz, daß das betrachtete Intervall abgeschlossen ist, ganz wesentlich ist: die Funktion $f\colon= \frac{1}{x}$ ist auf dem *offenen* Intervall $(0, 1)$ stetig und unbeschränkt.

Wir wissen nun, daß für eine auf einem abgeschlossenen Intervall $[a, b]$ stetige Funktion die Menge

$$f[a, b] := \{f(x) | x \in [a, b]\}$$

beschränkt ist. Nach dem Vollständigkeitsaxiom (und einem Satz) besitzt $f[a, b]$ ein Supremum s und ein Infimum i. Kommen s und i als Funktionswerte vor ? Die Antwort auf diese Frage gibt der zweite Haupsatz über stetige Funktionen:

Stetige Funktionen nehmen auf abgeschlossenen Intervallen ihre Extremwerte an

Satz. *Sei f eine stetige Funktion auf dem abgeschlossenen Intervall $[a, b]$. Dann gibt es $x, y \in [a, b]$, sodaß*

$$f(x) = \sup f[a, b] \quad \textit{und} \quad f(y) = \inf f[a, b].$$

Beweis. Wir zeigen, daß das Supremum angenommen wird. Bezeichnet s das Supremum, so ist $s - \frac{1}{n}$ keine obere Schranke; also gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in [a, b]$ sodaß

$$s - \frac{1}{n} < f(x_n) \le s.$$

Sei $(x_{n_\nu})_{\nu \in \mathbb{N}}$ eine konvergente Teilfolge der Folge x_n, etwa mit Limes x. Nach dem oben bewiesenen Satz ist jedenfalls $x \in [a, b]$. Wegen der Stetigkeit von f konvergiert auch die Folge $(f(x_{n_\nu}))_{\nu \in \mathbb{N}}$ und zwar gegen $f(x)$. Wegen der Ungleichung für die $f(x_n)$ muß $f(x) = s$ gelten.

Wieder ist die Voraussetzung, daß das betrachtete Intervall abgeschlossen ist, ganz wesentlich: die stetige Funktion $f\colon (0,1) \to \mathbb{R}, x \mapsto x$ nimmt im (offenen) Intervall $(0,1)$ das Supremum 1 der Menge ihrer Funktionwerte nicht an.

Der dritte Haupsatz ist folgendermaßen:

Zwischenwertsatz

Satz (Zwischenwertsatz). *Sei f eine auf dem Intervall $[a, b]$ stetige Funktion, und es gelte, $f(a) < 0 < f(b)$. Dann gibt es ein $c \in [a, b]$, sodaß $f(c) = 0$.*

Beweis. Wir setzen

$$M := \{x \in [a, b] | f(x) \le 0\}.$$

Wegen $f(a) < 0$ ist dies jedenfalls eine nichtleere Menge. Als Teilmenge eines abgeschlossenen Intervalls ist sie nach oben beschränkt. Also existiert

$$c := \sup M.$$

Wir zeigen

$$f(c) = 0.$$

Es ist jedenfalls $c < b$, denn es gibt ja nach einem oben bewiesenen Satz ein $\delta > 0$, sodaß $f(x)$ noch für alle $b - \delta < x \le b$ positiv ist, wonach daher $c \le b - \delta$ gilt. Mit dem gleichen Argument folgt $c > a$. Also ist für ein hinreichend kleines $\epsilon > 0$ eine ganze ϵ–Umgebung von c in $[a, b]$ enthalten. Insbesondere gibt es daher zu jeder positiven ganzen Zahl n ein x_n^+ und ein x_n^- in $[a, b]$, sodaß

$$c - \frac{1}{n} \le x_n^- < c < x_n^+ \le c + \frac{1}{n}, \qquad f(x_n^-) < 0,\ f(x_n^+) > 0.$$

Aus der Stetigkeit von f in c folgt

$$\lim_{n \to \infty} f(x_n^\pm) = f(c);$$

wegen $f(x_n^-) < 0$ folgt mit einem oben bewiesenen Satz dann aber $f(c) \le 0$, wegen $f(x_n^+) > 0$ folgt mit demselben Argument $f(c) \ge 0$, mithin $f(c) = 0$. Genau das wollten wir zeigen.

Wir geben noch eine scheinbar allgemeinere Formulierung des Zwischenwertsatzes.

Satz. *Es sei f stetig auf $[a,b]$, und es gelte $f(a) \leq f(b)$. Dann gibt es zu jedem γ mit $f(a) \leq \gamma \leq f(b)$ ein $c \in [a,b]$, sodaß $f(c) = \gamma$.*

Beweis. Zu gegebenem γ wende man den vorhergehenden Satz auf die Funktion $x \mapsto f(x) - \gamma$ an.

Schließlich kann man sich noch eine zu der eben formulierten Version des Zwischenwertsatzes analoge Version mit $f(a) \geq f(b)$ an Stelle von $f(a) \leq f(b)$ zurechtlegen. "f sei stetig auf $[a,b]$ und $f(a) \geq f(b)$; dann gibt es zu jedem $f(a) \geq \gamma \geq f(b)$ ein $c \in [a,b]$, sodaß $f(c) = \gamma$."

Man kann die drei Hauptsätze über stetige Funktionen sehr handlich in einem zusammenfassen, nämlich zu der folgenden Aussage

Zusammenfasssung der Hauptsätze

Satz. *Sei f eine auf D stetige Funktion. Sei $[a,b] \subset D$. Dann ist $f([a,b])$ wieder ein abgeschlossenes Intervall.*

Beweis. Da f auf D stetig ist, ist f auch stetig auf $[a,b]$. Also existieren nach Haupsatz 1 $i = \inf f([a,b])$ und $s = \sup f([a,b])$; insbesondere ist

$$f([a,b]) \subset [i,s].$$

Nach Hauptsatz 2 existieren x und y in $[a,b]$, sodaß $i = f(x)$ und $y = f(y)$. Nach den diskutierten verschiedenen Versionen des Haupsatzes 3, des Zwischenwertsatzes, wird jeder Zwischenwert im Intervall $[f(x), f(y)]$ angenommen. Zusammengefaßt ist $f([a,b]) = [f(x), f(y)]$.

Umgekehrt impliziert der eben bewiesene Satz offenbar jeden einzelnen der angeführten drei Hauptsätze.

Wir kommen nun zu einigen Anwendungen der bisher entwickelten Theorie der stetigen Funktionen. Wir haben vom letzten Kapitel noch nachzutragen, daß die Exponentialfunktion stetig ist, und daß die Quadratwurzeln positiver reeller Zahlen existieren. Wir beginnen mit dem folgenden

Stetigkeit der Exponentialfunktion

Satz. *Die Exponentialfunktion ist auf ganz $\mathbb{R}$ stetig.*

Beweis. Wir nehmen an, wir hätten schon die Steigkeit der Exponentialfunktion im Punkt 0 gezeigt. Dann folgt sofort die Stetigkeit der Exponentialfunktion in jedem beliebigen Punkt c: Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge mit Limes c; da $(x_n - c)_{n\in\mathbb{N}}$ eine Nullfolge ist, folgt aus der Stetigkeit von exp in 0, daß $\exp(x_n - c)$ gegen 1 konvergiert; wegen

$$\exp(x_n) = \exp(c) \cdot \exp(x_n - c)$$

konvergiert dann aber $\exp(x_n)$ gegen $\exp(c)$.

Die Stetigkeit von exp in 0 ist ein Spezialfall des folgenden Satzes, der allerdings später noch wesentlich verallgemeinert wird.

Satz. *Die Reihe* $\sum_{n=0}^{\infty} a_n x^n$ *sei für jedes* $|x| < R$ *absolut konvergent. Es bezeichne* f *die auf* $x \in (-R, +R)$ *vermöge*

$$x \mapsto \sum_{i=0}^{\infty} a_n x^n.$$

erklärte Funktion. Dann ist f *an der Stelle* 0 *stetig.*

Beweis. Sei (h_n) eine Nullfolge, $|h_n| < R$. Ist dann $0 < r < R$, so ist $|h_n| < r$ für fast alle n, und für ebendiese n ist

$$\begin{aligned} |f(h_n) - f(0)| &= |h_n| \cdot \left| \sum_{k=1}^{\infty} a_k \cdot h_n^{k-1} \right| \\ &\leq |h_n| \cdot \sum_{k=1}^{\infty} |a_k| \, |\cdot h_n|^{k-1} \\ &\leq |h_n| \cdot \sum_{k=1}^{\infty} |a_k| \cdot r^{k-1} \\ &\leq |h_n| \cdot \text{Konstante.} \end{aligned}$$

Hieraus ist sofort ersichtlich, daß $f(h_n) - f(0)$ eine Nullfolge ist, d.h. daß $f(h_n)$ gegen $f(0)$ konvergiert, und das war zu zeigen.

Ganz analog wie die Stetigkeit der Exponentialfunktion kann man beweisen

Stetigkeit der Hyperbel- und Kreisfunktionen

Satz. *Die Funktionen* cosh, sinh, cos *und* sin *sind auf ganz* $\mathbb{R}$ *stetig.*

Beweis. Nach den Reihendarstellungen und dem vorangehenden Satz sind diese Funktionen jedenfalls stetig in 0. Die Stetigkeit in jedem anderen Punkt c folgt nun wie im Fall der Exponetialfunktion aus den jeweiligen Additionstheoremen. Im Fall des Sinus etwa hat man für jede gegen ein c konvergente Folge (x_n):

$$\begin{aligned} \lim_{n\to\infty} \sin x_n &= \lim_{n\to\infty} (\sin c \cdot \cos(x_n - c) + \cos c \sin(x_n - c)) \\ &= \sin c \cdot \lim_{n\to\infty} \cos(x_n - c) + \cos c \cdot \lim_{n\to\infty} \sin(x_n - c) \\ &= \sin c. \end{aligned}$$

Den Nachweis der Stetigkeit der anderen Funktionen überlassen wir dem Leser.

Als Folge der Stetigkeit der Exponentialfunktion können wir nun endlich beweisen:

Bijektivität der Exponentialfunktion

Satz. *Die Abbildung* $\exp : \mathbb{R} \to \mathbb{R}_+, x \mapsto e^x$ *ist bijektiv.*

Beweis. Die Injektivität hatten wir in Kapitel 6 schon bewiesen; sie war eine Folge der Tatsache, daß $e^{x_1} < \mathrm{e}^{x2}$ gilt, wenn nur $x_1 < x_2$ ist, was wiederum eine Folge der Funktionalgleichung war. Es bleibt die Surjektivität nachzuweisen. Sei dazu $\gamma \in \mathbb{R}_+$. Dann kann man ein n wählen, sodaß $e^{-n} < \gamma < e^n$, denn die Folge der $\mathrm{e}^n = (\mathrm{e})^n$ ist wegen $\mathrm{e} > 1$ nach oben unbeschränkt, die Folge der $\mathrm{e}^{-n} = (\frac{1}{\mathrm{e}})^n$ konvergiert wegen $\frac{1}{\mathrm{e}} < 1$ gegen 0. Da exp nach dem oben bewiesenen Satz stetig ist, insbesondere also auf $[\mathrm{e}^{-n}, \mathrm{e}^n]$ stetig ist, folgt nach dem Zwischenwertsatz die Existenz eines c mit $\exp(c) = \gamma$, und das war zu zeigen.

Die Exponentialfunktion ist stetig, und sie ist bijektiv. Es schließ sich naheliegenderweise die Frage an, was können wir für die Umkehrfunktion, den Logarithmus, schließen. Hier gilt nun allgemein der

Stetigkeit der Umkehrfunktion

Satz. *Sei* $f\colon [a,b] \to [i,s]$ *bijektiv und stetig. Dann ist die Umkehrabbildung*

$$f^{-1}\colon [i,s] \to [a,b]$$

ebenfalls stetig.

Beweis. Sei $c \in [i,s]$ und $(y_n)_{n\in\mathbb{N}}$ eine Folge in $[i,s]$, die gegen c konvergiert. Die Folge $\big(f^{-1}(y_n)\big)_{n\in\mathbb{N}}$ ist beschränkt und hat demnach einen Häufungspunkt α. Dies ist aber auch der einzige Häufungspunkt der Folge. Sei nämlich α_1 irgendein Häufungspunkt dieser Folge. Dann gibt es Teilfolgen $\big(f^{-1}(y_{k_n})\big)_{n\in\mathbb{N}}$ und $\big(f^{-1}(y_{l_n})\big)_{n\in\mathbb{N}}$, die gegen α_1 bzw. α konvergieren. Die sind aber beides Teilfolgen in $[a,b]$. Wegen der Stetigkeit von f ist daher

$$\begin{aligned} f(\alpha_1) &= \lim_{n\to\infty} f\big(f^{-1}(y_{k_n})\big) \\ &= \lim_{n\to\infty} y_{k_n} = \lim_{n\to\infty} y_n = c. \end{aligned}$$

Ganz analog folgt

$$f(\alpha) = c.$$

Aus der Injektivität von f folgt endlich

$$\alpha_1 = \alpha.$$

Als Folge mit genau einem Häufungspunkte ist $\big(f^{-1}(y_n)\big)_{n\in\mathbb{N}}$ konvergent, und zwar gegen α. Und da — wie wir gesehen haben — $f(\alpha) = c$, d.h. $f^{-1}(c) = \alpha$ gilt, haben wir die Stetigkeit von f^{-1} in c bewiesen.

Als Konsequenz des eben bewiesenen Satzes erhalten wir nun den folgenden:

Stetigkeit des Logarithmus

Satz.

$$\log\colon \mathbb{R}^+ \to \mathbb{R}$$

ist überall stetig.

Beweis. Sei $y \in \mathbb{R}_+$ und $x = \log y$, d.h. $e^x = y$. Da mit exp auch die Einschränkung

$$\underline{\exp} : [x-1, x+1] \to [y/\mathrm{e}, y\cdot\mathrm{e}], \quad x \mapsto \exp(x)$$

stetig und bijektiv ist, ist nach dem eben bewiesenen Satz dann aber auch die Umkehrfunktion

$$\big(\underline{\exp}\big)^{-1}\colon [y/\mathrm{e}, y\cdot\mathrm{e}] \to [x-1, x+1], \quad y \mapsto \log y$$

stetig. Hieraus folgt offensichtlich die Stetigkeit von

$$\log|_{[y/\mathrm{e},y\cdot\mathrm{e}]}\colon [y/\mathrm{e}, y\cdot\mathrm{e}] \to \mathbb{R}, \quad y \mapsto \log y.$$

Insbesondere ist die Einschränkung $\log|_{[y/\mathrm{e},y\cdot\mathrm{e}]}$ stetig in y. Hieraus folgt nun endlich die Stetigkeit von log in y mittels des folgenden kleinen Hilfsatzes:

Satz. *Sei f eine auf D erklärte Funktion, es sei $c \in E \subset D$, und die Einschränkung $f|_E$ von f auf E sei in c stetig. Es gebe ein $\delta > 0$, sodaß $(t-\delta, t+\delta) \cap D$ ganz in E enthalten ist. Dann ist auch f stetig an der Stelle c.*

Dieser Hilfssatzes mutet auf den ersten Blick womöglich etwas überflüssig an. Aber in der Tat kann man von der Stetigkeit der Einschränkung einer Funktion in einem Punkt durchaus nicht auf die Stetigkeit der uneingeschränkten Funktion schließen: so ist etwa — wie man sich leicht überlegt — für jede Funktion $f\colon \mathbb{R} \to \mathbb{R}$ die Einschränkung $f|_{\mathbb{Z}}$ stetig in jedem $t \in \mathbb{Z}$, wogegen natürlich nicht jede auf $\mathbb{R}$ erklärte Funktion auf $\mathbb{Z}$ stetig ist. Wir kommen nun zum Beweis des Hilfssatzes.

Beweis. Sei$(x_n)n \in \mathbb{N}$ eine Folge von Elementen von D mit $\lim_{n\to\infty} x_n = c$. Dann gilt für fast alle n, daß $|x_n - c| < \delta$, also sind nach Voraussetzung für fast alle n, etwa für $n \geq n_0$, die x_n in E gelegen. Nach der Stetigkeit von $f|_E$ in c konvergiert die Folge $(f|_E(x_n))_{n\geq n_0}$ gegen c, d.h. — wegen $f|_E(x_n) = f(x_n)$ $(n \geq n_0)$ — die Folge der $f(x_n)$ konvergiert gegen c, und das war zu zeigen.

Nachdem die Stetigkeit der Exponentialfunktion und des Logarithmus bewiesen ist, stellt sich die Frage nach abgeleiteten Funktionen wie $a^x = \mathrm{e}^{x\log a}$. Hier gilt der

Stetigkeit des Kompositums

Satz. *Seien f und g stetige Funktionen auf D bzw. E, und es gelte $D \supset g(E)$, sodaß $f \circ g$ auf E erklärt ist. Ist dann g stetig an der Stelle $c \in E$ und ist f stetig an der Stelle $g(c)(\in D)$, so ist $f \circ g$ stetig in c.*

Beweis. Sei nämlich $(x_n)_{n\in\mathbb{N}}$ irgendeine Folge, die gegen $c \in E$ konvergiert. Dann konvergiert die Folge $(g(x_n))_{n\in\mathbb{N}}$ gegen $g(c)$ — wegen der Stetigkeit von g in c — und somit konvergiert die Folge $(f(g(x_n)))_{n\in\mathbb{N}}$, d.h. die Folge $((f \circ g)(x_n))_{n\in\mathbb{N}}$, gegen $f(g(c)) = (f \circ g)(c)$ — wegen der Stetigkeit von f.

Als Folgerung erhalten wir

Stetigkeit von a^x

Satz. *Für eine vorgegebene positive reelle Zahl $a > 0$ ist die Funktion $x \mapsto a^x = e^{x\cdot\log a}$ auf ganz $\mathbb{R}$ stetig. Für eine vorgegebene reelle Zahl x ist die auf $\mathbb{R}_+$ erklärte Funktion $a \mapsto a^x$ auf ganz $\mathbb{R}_+$ stetig.*

Beweis. Die erste Funktion schreibt sich als $\exp \circ m_{\log a}$, wobei m_t für eine vorgegebene reelle Zahl t die Funktion $x \mapsto x \cdot t$ bedeutet. Da exp stetig ist, und $m_{\log a}$ (als Polynom) ebenfalls stetig ist, ist die Stetigkeit der ersten im Satz aufgeführten Funktion eine Folge des Satzes über das Kompositum stetiger Funktionen. Die Stetigkeit der zweiten Funktion folgt analog, indem man sie in der Form $\exp \circ m_x \circ \log$ schreibt.

Im letzten Kapitel hatten wir schon bemerkt, daß zu vorgegebener positiver ganzer Zahl n jede positive reelle Zahl y gleich der n–ten Potenz der Zahl $\exp\left(\frac{\log y}{n}\right)$ ist. (Allerdings fehlte uns dort noch der Nachweis der Existenz der Logarithmus Funktion, d.h. der Nachweis der Bijektivität der Exponentialfunktion.) Damit haben wir

Existenz von n–ten Wurzeln

Satz. *Zu jeder positiven reellen Zahl y und jeder positiven ganzen Zahl n gibt es eine (reelle) n–te Wurzel, d.h. eine reelle Zahl a mit $a^n = y$.*

Die Aussage des letzten Satzes kann man formulieren, indem man sagt, daß das Polynom $X^n - y$ stets eine (reelle) Nullstelle besitzt. Nun besitzt — wie wir wissen (z.B. $X^2 + 1$!) — nicht jedes Polynom eine reelle Nullstelle. Immerhin kann man aber (als eine weitere Anwendung des Zwischenwertsatzes) den folgenden Satz beweisen:

Noch einmal Nullstellen

Satz. *Jedes Polynom ungeraden Grades hat eine reelle Nullstelle.*

Beweis. Sei $f(x) = a_0 + a_1x + ... + a_nx^n$ ein Polynom ungeraden Grades n. Offenbar können wir $a_n > 0$ annehmen (sonst betrachte $-f$ statt f). Für $x \neq 0$ ist

$$f(x) = x^n \cdot \left(\frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + ... + a_n\right).$$

Man überlegt sich leicht die Existenz eines M, sodaß für alle x gilt

$$|x| \geq M \Rightarrow \left|\frac{a_0}{x^n} + ... + \frac{a_{n-1}}{x}\right| < \frac{a_n}{2}.$$

Für $|x| \geq M$ ist dann aber

$$\begin{aligned}\frac{f(x)}{x} &\geq x^{n-1} \cdot \left(a_n - \left|\frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + ... + \frac{a_{n-1}}{x}\right|\right) \\ &\geq M^{n-1} \cdot \frac{a_n}{2},\end{aligned}$$

wobei wir noch ausgenutzt haben, daß $n - 1$ gerade, mithin $x^{n-1} > 0$ ist. Danach haben wir insbesondere

$$\begin{aligned}&f(-M) < 0 \\ &f(+M) > 0.^{,}\end{aligned}$$

und nach dem Zwischenwertsatz hat f daher eine Nullstelle im Intervall $[-M, M]$.

8 DIFFERENZIERBARKEIT

Sei f eine auf dem offenen Intervall (a, b) erklärte Funktion.

Differenzenquotienten

Einen Ausdruck der Form

$$\frac{f(x_0 + h) - f(x_0)}{h},$$

wobei $x_0, x_0 + h \in (a, b)$ und $h \neq 0$ seien nennen wir *Differenzenquotient von f an der Stelle x_0.*.

Geometrisch gibt der Differenzenquotient die Steigung der Sekante des Graphen von f durch die Punkte $(x_0, f(x_0))$ und $(x_0 + h, f(x_0 + h))$ an.

Differenzierbare Funktion

f heißt differenzierbar in x_0, falls die Folge der Differenzenquotienten

$$\frac{f(x_0 + h_n) - f(x_0)}{h_n}.$$

für jede Nullfolge $(h_n)n \in \mathbb{N}$ mit $h_n \neq 0$ und $x_0 + h_n \in (a, b)$ konvergiert.

Ist f differenzierbar bei x_0, so ist der Grenzwert

$$\lim_{n \to \infty} \frac{f(x_0 + h_n) - f(x_0)}{h_n}$$

offenbar unabhängig von der speziellen Wahl der Nullfolge. Sind nämlich $(h_n)_{n \in \mathbb{N}}$ und $(k_n)_{n \in \mathbb{N}}$ Nullfolgen, die die Bedingungen in der Definition der Differenzierbarkeit erfüllen, so ist auch $h_1, k_1, h_2, k_2, h_3, k_3, ...$ eine solch zulässige Folge, konvergiert mithin gegen einen Grenzwert, gegen den dann auch jede Teilfolge, insbesondere die Teilfolg der h_n bzw. der k_n, konvergiert. Den gemeinsamen Limes aller zulässigen Teilfolgen bezeichnet man

Die Ableitung

als *Ableitung von f an der Stelle x_0* und benutzt dafür das Symbol

$$f'(x_0) \qquad \text{oder} \qquad \frac{df}{dx}(x_0).$$

Zur Erleichterung der Sprechweise vereinbaren wir eine auch sonst nützliche

Schreibweise:

> Sei g eine auf einer Teilmenge D der reellen Zahlen erklärte Funktion, sei x_0 ein *Häufungspunkt von D*, d.h. eine (nicht notwendig in D enthaltene) Zahl, sodaß jede ϵ–Umgebung von x_0 mindestens ein von x_0 verschiedenes Element enthält. Eine Zahl a heißt *Grenzwert von $g(x)$ bei Annäherung von x an x_0*, in Zeichen
> $$a = \lim_{x \to x_0} g(x),$$
> falls für jede Folge $(x_n)n \in \mathbb{N}$ in $D \setminus \{x_0\}$ mit Grenzwert x_0
> $$\lim_{n \to \infty} g(x_n) = a$$
> gilt.

Man überlegt sich leicht, daß der hier auftetende Grenzwert a, falls er existiert, eindeutig bestimmt ist (hierbei geht wesentlich ein, daß a ein Häufungspunkt ist !) Ferner überlegt man sich leicht Regeln für den Umgang mit diesem Symbol, wie etwa

$$\lim_{x \to x_0} (g(x) + h(x)) = \lim_{x \to x_0} g(x) + \lim_{x \to x_0} h(x),$$

falls die Grenzwerte auf der rechten Seite existieren, etc. Solche Regeln ergeben sich unmittelbar aus dem Hauptsatz für konvergente Folgen, und wir werden sie im Folgenden stillschweigend benutzen.

Offenbar besagt die Stetigkeit einer Funktion g an einer Stelle x_0 ihres Definitionsbereiches in der eben eingeführten Schreibweise nichts anderes als daß

$$\lim_{x \to x_0} g(x) = g(x_0),$$

und die Differenzierbarkeit besagt nichts anderes, als das

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Man kann dies noch etwas anders formulieren: die Differenzierbarkeit von f an der Stelle x_0 ist gleichbedeutend damit, daß man die *Differenzenquotientenfunktion*

$$(a, b) \setminus \{x_0\} \to \mathbb{R}, \quad x \mapsto \frac{f(x) - f(x_0)}{x - x_0} =: R(x),$$

zu einer im Punkt x_0 *stetigen* Funktion fortsetzen kann. In der Tat: ist f differenzierbar in x_0, so hat die Funktion

$$\underline{R}(x) := \begin{cases} R(x) & \text{für } x \in (a, b),\ x \neq x_0 \\ f'(x_0) & \text{für } x = x_0 \end{cases}$$

die Eigenschaft

$$\lim_{x \to x_0} \underline{R}(x) = \underline{R}(x_0),$$

ist also stetig in x_0. Kann man umgekehrt R zu einer in x_0 stetigen Funktion fortsetzen, so strebt $R(x)$ bei Annäherung von x an x_0 gegen einen Grenzwert, und das ist ja gerade die Definition der Differenzierbakeit von f in x_0.

Ist die Funktion f bei x_0 differenzierbar, so ist die eben betrachtete Funktion $\underline{R}$ stetig bei x_0. Nun ist aber offenbar

$$f(x) = f(x_0) + \underline{R}(x)(x - x_0).$$

Daher impliziert die Stetigkeit von $\underline{R}$ bei $x = x_0$ unmittelbar den

Satz. *Die auf dem offenen Intervall (a, b) erklärte Funktion f sei in $x_0 \in (a, b)$ differenzierbar. Dann ist f auch stetig in x_0.*

Beispiele differenzierbarer Funktionen und Regeln zur Berechnung der Ableitung

Wir betrachten Beispiele von differenzierbaren Funktionen und entwickeln damit gleichzeitig einige Differenzierbarkeitsregeln. Die konstante Funktion, d.h. die Funktion, die auf einem gegebenen Intervall nur einen einzigen Wert annimmt, hat offenbar den Wert 0. Die nächst einfache Funktion ist $f(x) = x^n$, die auf $(a, b) = (-\infty, \infty)$ erkärt ist. Hier ist

$$R(x) = \frac{x^n - x_0^n}{x - x_0} = x^{n-1} + x^{n-2}x_0 + ... + x_0^{n-1}$$

für $x \neq x_0$. Die rechte Seite dieser Identität kann man (als Polynom) offensichtlich als eine auf ganz $\mathbb{R}$ stetige Funktion auffassen, und zwar mit dem Wert nx_0^{n-1} für $x = x_0$. Also ist f bei x_0 differenzierbar, und es ist $f'(x_0) = nx_0^{n-1}$. Das spezielle Polynom $f(x) = x^n$ ist also in jedem Punkt differenzierbar; natürlich gilt dies für jedes Polynom, wie sich sofort aus dem folgenden Satz ergibt:

Satz. *Seien f und g auf dem Intervall (a, b) erklärte, in x_0 differenzierbare Funktionen. Dann sind auch $f + g$ und $f \cdot g$ an der Stelle x_0 differenzierbar, und für ihre Ableitungen gilt*

$$(f + g)'(x_0) = f'(x_0) + g'(x_0)$$
$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0).$$

Insbesondere ist $(c \cdot f)'(x_0) = cf'(x_0)$ für jede reelle Zahl c.

Die hier formulierte Regel $(fg)' = f'g + fg'$ ist die sogenannte *Produktregel.*

Beweis. Für $x \neq x_0$ ist

$$\frac{f(x) + g(x) - (f(x_0) + g(x_0))}{x - x_0}$$
$$= \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0}.$$

Nach Voraussetzung strebt die rechte Seite bei Annäherung von x an x_0 gegen $f'(x_0) + g'(x_0)$; dies gilt dann natürlich auch für die linke Seite. Ebenso folgt die Behauptung über Produkte, indem man

$$\begin{aligned}&\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0}\\ &= \frac{(f(x) - f(x_0)) \cdot g(x) + f(x_0) \cdot (g(x) - g(x_0))}{x - x_0}\end{aligned}$$

schreibt.

Ist nun $f(x)$ ein Polynom, etwa

$$f(x) = a_0 + a_1 x + ... + a_n x^n,$$

so ist nach dem letzten Satz mit den Funktionen x^n auch $f(x)$ an jeder Stelle x diefferenzierbar, und es ist $f'(x) = a_1 + 2a_2 x + 3a_3 x^2 + ... + n a_n x^{n-1}$.

Das nächste naheliegende Beispiel ist eine rationale Funktion. Hierzu beweisen wir zunächst allgemein den

Satz. *Sei f eine auf (a, b) erklärte, in $x_0 \in (a, b)$ differenzierbare Funktion. Es gelte $f(x) \neq 0$ für alle $x \in (a, b)$. Dann ist die Funktion $\frac{1}{f}$ in x_0 differenzierbar, und es gilt:*

$$\left(\frac{1}{f}\right)'(x_0) = -\frac{f'(x_0)}{f^2(x_0)}.$$

Beweis. Dies ergibt sich wieder wie im vorangehenden Satz, indem man nun

$$\frac{\frac{1}{f(x)} - \frac{1}{f(x_0)}}{x - x_0} = \frac{-\frac{f(x)-f(x_0)}{x-x_0}}{f(x) \cdot f(x_0)}$$

schreibt, und noch ausnutzt, daß f als im Punkt x_0 differenzierbare Funktion insbesondere stetig ist.

Die Voraussetzung “$f(x) \neq 0$ für alle $x \in (a, b)$” war nötig, um die Existenz der Funktion $\frac{1}{f}$ zu gewährleisten. Allerdings entspricht dies nicht der natürlichen Situation. Die natürliche Situation ist, daß f bei x_0 differenzierbar ist, und $f(x_0) \neq 0$ ist. Dann gibt es nämlich — da f dann bei x_0 stetig ist — nach einem Satz des letzten Kapitels ein offenes, in (a, b) enthaltenes Intervall I, welches den Punkt x_0 enthält, und sodaß $f(x) \neq 0$ für jedes $x \in I$ ist. Somit kann man die Funktion $1/f|_I$ bilden. Diese Funktion ist nun auch bei x_0 differenzierbar, und zwar mit der gleichen Ableitung wie im Satz. Dies folgt auch aus ebendiesem Satz, wobei man aber noch den folgenden einfachen Sachverhalt ausnutzen muß.

Satz. *Sei $a \le a' < b' \le b$, sei f eine auf (a,b) definierte Funktion, und sei $x_0 \in (a',b')$. Dann ist $f|_{(a',b')}$ genau dann in x_0 differenzierbar, wenn es f ist.*

Den offensichtlichen Beweis lassen wir als Übungsaufgabe.

Der vorletzte Satz ist ein Spezialfall der *Quotientenregel*:

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2},$$

Diese (und ihre präzise Formulierung) ergibt sich aus ebendiesem Satz und der Produktregel. Damit man sich die Reihenfolge besser merken kann:

$$A\left(\frac{\mathrm{Z}}{\mathrm{N}}\right) = \frac{\mathrm{N}A(\mathrm{Z}) - \mathrm{Z}A(\mathrm{N})}{\mathrm{N}^2} = \frac{\text{Netz} - \text{Zahn}}{\mathrm{N}^2}.$$

Anhand der letzten Regel und der Tatsache, daß Polynome differenzierbar sind, erkennen wir, daß rationale Funktionen überall, wo sie definiert sind, differenzierbar sind, und wir sind in der Lage, ihre Ableitungen zu berechnen.

Die bisher betrachteten Funktionen sind Beispiele für Funktionen, die in jedem Punkt ihres Definitionsbereichs differenzierbar sind.

> Sei f eine auf dem offenen Intervall (a,b) erklärte Funktion. Ist f in jedem Punkt von (a,b) differenzierbar, so heißt die Funktion
>
> $$(a,b) \to \mathbb{R}\ x \mapsto f'(x)$$
>
> die *Ableitung von* f, als Symbol f' oder $\frac{d}{dx}f$.

Höhere Ableitungen

Man kann den Prozeß des Bildens der Ableitung natürlich iterieren: Ist f' in jedem Punkt von (a,b) differenzierbar, so kann man $f'' := (f')'$ bilden, und allgemein kann man für eine natürliche Zahl n die *n–te Ableitung*

$$f^{(n)} := (f^{(n-1)})'$$

bilden, sofern die Ableitungen $f^{(k)}$ $(k < n)$ existieren und in jedem Punkt von (a,b) differenzierbar sind; in diesem Fall heißt f *n–mal differenzierbar*. Eine andere naheliegende Schreibweise für $f^{(n)}$ ist

$$\frac{d^n}{dx^n}f.$$

Die nächst einfache Funktion ist die Exponentialfunktion. Der Differenzenquotient ist

$$\frac{\mathrm{e}^{x_0+h} - \mathrm{e}^{x_0}}{h} = \mathrm{e}^{x_0}\frac{\mathrm{e}^h - 1}{h}.$$

Ableitung der Exponential-, Hyperbel- und Kreisfunktionen

Was geschieht nun mit $(e^h - 1)/h$, falls wir h gegen 0 streben lassen ? Setzen wir die Reihenentwicklung der Exponetialfunktion ein, so erhalten wir

$$\frac{e^h - 1}{h} = 1 + \frac{h}{2!} + \frac{h^2}{3!} + \cdots.$$

Auf der rechten Seite steht jetzt eine für jedes h absolut konvergente Reihe, und nach einem Satz aus dem letzten Kapitel ist sie stetig in 0. Dami finden wir

$$\lim_{h\to 0} \frac{e^h - 1}{h} = 1,$$

also den

Satz. *Die Exponentialfunktion ist überall differenzierbar, und es gilt*

$$\frac{d}{dx}\exp(x) = \exp(x).$$

Als Folgerung erhalten wir (unter Benutzung der angeführten Regeln):

$$\cosh' = \sinh \qquad \sinh' = \cosh\,.$$

Die Ableitungen der Kreisfunktionen sin und cos kann man ähnlich wie im Fall der Exponentialfunktion berechnen. Da $\cos(0) = 0$ ist, ist der Differenzenquotient bei $x_0 = 0$

$$\frac{\cos(h) - 1}{h} = -\frac{h}{2!} + \frac{h^3}{4!} - \cdots.$$

Die rechte Seite ist absolut konvergent für jedes h, und definiert somit eine bei 0 stetige Funktion; also existiert der Grenzwert, wenn h gegen 0 strebt, und ist gleich 0. Also $\cos'(0) = 0$. Analog erhält man $\sin'(0) = 1$. Damit können wir nun an einer beliebigen Stelle x_0 differenzieren: es ist

$$\frac{\cos(x_0 + h) - \cos x_0}{h} = \frac{\cos(x_0)(\cos h - 1) - \sin(x_0)\sin h}{h},$$

und daher

$$\lim_{h\to 0} \frac{\cos(x_0 + h) - \cos(x_0)}{h} = -\sin x_0.$$

Ganz ähnlich kann man den Sinus abhandeln, und man findet

$$\cos' = -\sin \qquad \sin' = \cos\,.$$

Es schließt sich die Frage nach den Differenzierbarkeitseigenschaften des Logarithmus (als Umkehrfunktion der differenzierbaren Exponentialfunktion) und der Funktionen a^x an. Hierzu beweisen wir die folgenden zwei Regeln, die Kettenregel und eine Regel für Umkehrfunktionen.

Kettenregel und Regel für Umkehrfunktionen.

Satz. *Sei f eine auf (a,b) und g eine auf (c,d) erklärte Funktion. Es gelte $f(a,b) \subset (c,d)$ (sodaß $g \circ f$ wohldefiniert ist). Schließlich sei f differenzierbar in x_0 und g sei differenzierbar in $f(x_0)$. Dann ist $g \circ f$ in x_0 differenzierbar, und es gilt*

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

Beweis. Es bezeichnen $R(x)$ und $S(y)$ die Differenzenquotientenfunktion von f und g bei x_0 bzw. $f(x_0)$, d.h.

$$\begin{matrix} f(x) = f(x_0) + R(x)(x - x_0) \\ g(y) = g(y_0) + S(y)(y - y_0) \end{matrix} .$$

Damit ist

$$g(fx) = g(f(x_0)) + S(f(x) \cdot R(x)(x - x_0)),$$

d.h.

$$T(x) := S(f(x)) \cdot R(x)$$

ist die Differenzenquotientenfunktion bei $(f \circ g)(x_0)$. Nach den Differenzierbarkeitsvorausetzungen über f und g existiert der Grenzwert von $T(x) = S(f(x))) \cdot R(x)$, wenn x gegen x_0 strebt, und ist gerade gleich dem behaupteten Wert.

Satz. *Sei $f : (a,b) \to (c,d)$ bijektiv und stetig. Es sei f in x_0 differenzierbar und $f'(x_0) \neq 0$. Dann ist $f^{-1} : (c,d) \to (a,b)$ im Punkt $y_0 := f(x_0)$ differenzierbar, es ist $\left(f^{-1}\right)'(y_0) = 1/f'(x_0)$.*

Man beachte, daß die Formel für die Ableitung von f^{-1} im Punkt y_0 eine einfache Folgerung der Kettenregel ist: es ist ja $x = (f^{-1} \circ f)(x)$, also

$$1 = \left(f^{-1}\right)'(f(x_0)) \cdot f'(x_0),$$

und dies ist gerade die behauptete Formel. Man beachte, daß dieser Schluß auch zeigt, daß die Voraussetzung $f'(x_0)$ eine notwendige Voraussetzung des Satzes ist, d.h. ist f differenzierbar in x_0 und ist f^{-1} differenzierbar in $y_0 = f(x_0)$ dann folgt $f'(x_0) \neq 0$. Man kann sich dies auch noch an dem Beispiel der Funktion $x \mapsto x^3$, $x_0 = 0$ illustrieren. Zum Beweis des Satzes ist also lediglich zu zeigen, daß f^{-1} überhaupt bei y_0 differenzierbar ist.

Beweis. Es ist

$$\begin{aligned} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} &= \frac{1}{\frac{y - y_0}{f^{-1}(y) - f^{-1}(y_0)}} \\ &= \frac{1}{\frac{f(f^{-1}(y)) - f(x_0)}{f^{-1}(y) - x_0}}. \end{aligned}$$

Die rechte Seite hat aber einen Grenzwert, wenn y gegen y_0 läuft: nach einem Satz des letzten Kapitels ist nämlich mit f auch f^{-1} stetig; daher läuft mit y gegen y_0 dann $f^{-1}(y)$ gegen x_0, und die Existenz des Grenzwerts der rechten Seite folgt aus der Differenzierbarkeit von f in x_0 (mit nichtverschwindender Ableitung !)

Die Formel für die Ableitung der Umkehrfunktion läßt sich folgendermaßen leicht merken: Schreibt man $y = f(x)$, so kann man die Ableitung von f als $\frac{dy}{dx}$ schreiben und die Ableitung von $x = f^{-1}$ als $\frac{dx}{dy}$.

Die Formel für die Ableitung besagt nun nichts anderes als

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}.$$

Ein Beispiel: $y = f(x) = x^2$ $(x > 0.)$, also $x = f^{-1}(y) = \sqrt{y}$. Es ist $\frac{dy}{dx} = 2x$, und so

$$(f^{-1})'(y) = \frac{dx}{dy} = 1/\frac{dx}{dy} = \frac{1}{2x} = \frac{1}{2\sqrt{y}}.$$

Ein weiteres Beispiel $y = \exp(x),\ x = \log y$, also

$$\log' y = 1/\frac{dx}{dy} = 1/\exp(x) = 1/y.$$

Hieraus, mit $\exp' = \exp$ und mittels der Kettenregel erhält man nun sofort den

Ableitung des Logaritmus und der Potenz

Satz. *Die Funktionen* $\log$ *$(x > 0)$, $x \mapsto a^x$ $(x \in \mathbb{R})$ und $a \mapsto a^x$ $(a > 0)$ sind in jedem Punkt ihres Definitionsbereiches differenzierbar, und es gilt*

$$\frac{d}{dx}\log x = \frac{1}{x}, \quad \frac{d}{dx}a^x = a^x \cdot \log a, \quad \frac{d}{da}a^x = xa^{x-1}.$$

Extremwerte, Satz von Rolle, Mittelwertsatz

Wir behandeln als nächstes die Hauptsätze der Theorie der differenzierbaren Funktionen.

> Die auf D definierte Funktion f hat bei $x_0 \in (a, b)$ ein *lokales Maximum*, falls
>
> $$\exists \delta > 0 \forall x \in (x_0 - \delta, x_0 + \delta) \cap D \quad (f(x) \leq f(x_0)).$$

Gilt die stärkere Bedingung “$f(x) < f(x_0)$” statt “$f(x) \leq f(x_0)$,” so spricht man von einem strikten lokalen Maximum an der Stelle x_0. Ganz entsprechend definiert man die Begriffe “lokales Minimum” und “striktes lokales Minimum”.

Satz. *Sei f eine auf (a,b) erklärte, in $x_0 \in (a,b)$ differenzierbare Funktion. Hat f ein lokales Maximum (Minimum) an der Stelle $x_0 \in (a,b)$, so ist $f'(x_0) = 0$.*

Beweis. Die Differenzierbarkeit ist gleichbedeutend mit der Stetigkeit der Funktion $\underline{R}$ in x_0, wo

$$\underline{R}(x) = \begin{cases} R(x) & \text{für } x \in (a,b),\ x \neq x_0 \\ f'(x_0) & \text{für } x = x_0 \end{cases}.$$

Hat f etwa ein lokales Maximum in x_0, so ist in einer δ-Umgebung von x_0

$$\begin{cases} R(x) \geq 0 & \text{für alle } x < x_0 \\ R(x) \leq 0 & \text{für alle } x > x_0. \end{cases}$$

Jede ϵ–Umgebung von x_0 enthält also sowohl Punkte, in denen $R(x) \geq 0$ gilt, als auch Punkte, in denen $R(x) \leq 0$ ist. Wegen der Stetigkeit von R in x_0 kann dann aber nach einem Satz des letzten Kapitels nur $R(x_0) = 0$ gelten.

Satz von Rolle. *Sei f eine auf dem abgeschlossenen Intervall $[a,b]$ erklärte, stetige Funktion, und f sei in (a,b) differenzierbar. Es gelte $f(a) = f(b) = 0$. Dann gibt es ein $x_0 \in (a,b)$ sodaß $f'(x_0) = 0$ gilt.*

Beweis. Da f stetig auf $[a,b]$ ist, nimmt f sein Supremum s und sein Infimum i an. Werden s und i in den Punkten a, b angenommen, so folgt $s = i = 0$, also $f(x) = 0$ für alle $x \in [a,b]$, also $f'(x) = 0$ für jedes $x \in (a,b)$. Werden dagegen s oder i in einem Punkt $x_0 \in (a,b)$ angenommen, so schließt man nach dem vorangehenden Satz $f(x_0) = 0$.

Satz (Mittelwertsatz). *Sei f eine auf dem abgeschlossenen Intervall $[a,b]$ erklärte, stetige Funktion, und f sei in (a,b) differenzierbar. Dann gibt es ein $x_0 \in (a,b)$, sodaß*

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

Beweis. Man wende den Satz von Rolle auf die Funktion

$$g(x) := f(x) - f(a) - \frac{f(b) - f(a)}{b - a} \cdot (x - a)$$

an.

Die geometrische Bedeutung des Mittelwertsatzes ist folgendermaßen: die Gleichung der Sekante an G_f durch die beiden Punkte $(x_0, f(x_0))$ und $(x_0 + h, f(x_0 + h))$ ist

$$y = \frac{f(x_0 + h) - f(x_0)}{h} \cdot (x - x_0) + f(x_0)$$

Lassen wir h gegen 0 streben, so geht die Sekantengleichung in die *Tangentenlgleichung* über:

$$y = f'(x_0)(x - x_0) + f(x_0).$$

Nun sind zwei verschiedene Geraden $y = \alpha x + \beta$ und $y = \alpha' \cdot x + \beta'$ genau dann parallel, wenn $\alpha = \alpha'$ ist. Demnach besagt der Mittelwertsatz, daß der Graph von f eine Tangente an einem über einem $x_0 \in (a, b)$ gelegenen Punkt besitzt, die parallel zur Sekante durch die Endpunkte des Graphen ist.

Als Anwendung des Mittelwertsatzes bewiesen wir den

Satz. *Sei f in (a, b) differenzierbar und $f'(x) > 0$ für alle $x \in (a, b)$. Dann ist f streng monoton wachsend.*

Beweis. Seien $\alpha, \beta \in (a, b), \alpha < \beta$. Nun ist $[\alpha, \beta] \subset (a, b)$ und die Einschränkung von f auf $[\alpha, \beta]$ erfüllt die Voraussetzungen des Mittelwertsatzes. Danach gibt es also ein $\gamma \in (\alpha, \beta)$ sodaß

$$\frac{f(\beta) - f(\alpha)}{\beta - \alpha} = f'(\gamma)(> 0).$$

Es folgt $f(\alpha) < f(\beta)$.

Ganz analog kann man zeigen, daß f streng monoton fallend ist, wenn $f'(x) < 0$ für alle $x \in (a, b)$ ist.

Der Satz über das Verschwinden der Ableitung in einem lokalen Extremwert läßt sich nicht umkehren: die Ableitung von $x \mapsto x^3$ in $x = 0$ verschwindet, obwohl dort kein lokaler Extremwert vorliegt.

Um aus dem Verschwinden der Ableitung auf ein lokales Extremum schließen zu können benötigt man also noch zusätzliche Informationen. Ein mögliches hinreichendes Kriterium bietet der folgende

Satz. *Sei f in (a, b) differenzierbar, $x_0 \in (a, b)$ und $f(x_0) = 0$. Außerdem gebe es ein $\delta > 0$ sodaß f' in $(x_0 - \delta, x_0 + \delta) \cap (a, b)$ von positiven zu negativen Werten übergeht, d.h. sodaß für alle $x \in (x_0 - \delta, x_0 + \delta) \cap (a, b)$ die folgenden Bedingungen gelten:*

$$\begin{cases} f'(x) > 0 & \text{für } x < x_0 \\ f'(x) < 0 & \text{für } x > x_0 \end{cases}.$$

Dann hat f in x_0 ein lokales striktes Maximum.

Beweis: Sei $x \in (a, b)$ aus der angegebenen δ-Umgebung von x_0. Für ein passendes γ zwischen x und x_0 ist dann

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\gamma),$$

also $f(x) < (x_0)$, falls $x < x_0$, denn dann ist ja $\gamma < x_0$ und somit $f'(\gamma) > 0$, als auch $f(x) < f(x_0)$, falls $x > x_0$, denn dann ist $\gamma > x_0$ und somit $f'(\gamma) > 0$.

Man legt sich leicht ein analoges Kriterium für ein lokales Minimum zurecht (indem man im Satz etwa f durch $-f$ ersetzt).

Diskussion der Kreisfunktionen

Als Anwendung unserer bisher entwickelten Theorie wollen wir die Kreisfunktionen genauer studieren. Wir hatten oben definiert:

$$\cos x := 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{k=0}^{\infty} \left(x^{2k}\right)(2k)!$$

$$\sin x := x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}$$

Schreiben wir die Reihendarstellung des Cosinus in der Form

$$\cos x = 1 - \frac{x^2}{2!}\left(1 - \frac{x^2}{3\cdot 4}\right) - \frac{x^6}{6!}\left(1 - \frac{x^2}{7\cdot 8}\right) - \frac{x^{10}}{10!}\left(1 - \frac{x^2}{11\cdot 12}\right) - \cdots,$$

so erkennen wir

$$\cos 2 < 1 - \frac{4}{2}\left(1 - \frac{4}{3\cdot 4}\right) = -\frac{1}{3}.$$

Aus $\cos(0) = 1$ und $\cos(2) < 0$ schließen wir nach dem Zwischenwertsatz auf die Existenz einer Nullstelle des Cosinus im Intervall (0,2). Sei ν das Infimum der Menge aller positiven Nullstellen des Cosinus; aus Stetigkeitsgründen ist dann $\nu > 0$ (wegen $\cos 0 = 1 > 0$ gibt es eine ganze ϵ–Umgebung der 0, auf der der Cosinus strikt positiv ist !) und auch $\cos\nu = 0$ (nach Definition von ν gibt es eine Folge von Nullstellen des Cosinus, die gegen ν konvergiert und der Cosinus ist stetig !). Also hat der Cosinus eine kleinste positive Nullestelle, die für den Augenblick mit ν bezeichnet wird. Es ist hierfür wegen

$$\mathrm{e}^{ix} = \cos x + i \sin x$$

und wegen

$$\left|\mathrm{e}^{ix}\right|^2 = \cos^2 x + \sin^2 x$$

(letzteres nur für reelle x !) offenbar $\mathrm{e}^{i\nu} = \pm i$.

Wegen

$$\sin x = x\left(1 - \frac{x^2}{3!}\right) + \frac{x^5}{5!}\left(1 - \frac{x^2}{6\cdot 7}\right) + \cdots,$$

ist der Sinus auf dem Intervall $(0, \sqrt{6})$ strikt positiv; da $0 < \nu < 2 < \sqrt{6}$ ist insbesondere $\sin\nu > 0$. Wie schließen damit $\mathrm{e}^{i\nu} = i$.

Offenbar ist insbesondere $\mathrm{e}^{4in\nu} = 1$ für jede ganze Zahl n. Sei y irgeneine relle Zahl, sodaß $\mathrm{e}^{iy} = 1$ gilt. Dann gibt es jedenfalls ein $n \in \mathbb{Z}$, sodaß

$$4n\nu \le y < 4(n+1)\nu.$$

Es folgt $0 \leq y - 4n\nu < 4\nu$ und $\mathrm{e}^{i(y-4n\nu)} = 1$, also $\mathrm{e}^{i(y-4n\nu)/4} = \pm i$, und so $0 \leq (y-4n\nu)/4 < \nu$ und $\cos(y-4n\nu)/4 = 0$. Da ja ν die kleinste positive Nullstelle des Cosinus ist, muß $0 = (y-4n\nu)/4$, d.h. $y = 4n\nu$ gelten. Damit haben wir den

Satz. *Der Kern des Homomorphismus* $\mathbb{R} \to \mathbb{S}^1, \quad x \to e^{ix}$ *ist die Untergruppe* $4\nu\mathbb{Z}$ *der additiven Gruppe der reellen Zahlen, wo* ν *die kleinste positive Nullstellle bezeichnet.*

Die Zahl π

Dies ist eine schon in Kapitel 6 angekündigte Behauptung. Dort hatten wir erklärt, daß man die hier auftretende Zahl 2ν mit π bezeichnet wird. Mit dieser Bezeichnung haben wir

Die kleinste positive Nullstelle des Cosinus ist die Zahl $\frac{\pi}{2}$.

Wie die oben angestellten Überlegungen zeigen, hätten wir auch die letzte Aussage als Definition von π nehmen können. Eine andere Charakterisierung ist:

2π ist die kleinste positive *Periode* des Cosinus, d.h. die kleinste positive Zahl, sodaß $\cos(x + 2\pi) = \cos x$ für alle x gilt.

In der Tat: Es ist jedenfalls 2π eine Periode des Cosinus, und ist P eine weitere Periode des Cosinus, so ist mit $\cos P = \cos 0 = 1$ dann auch $\mathrm{e}^{iP} = 1$, also $P \in 2\pi\mathbb{Z}$. Die hier erklärte Zahl π ist natürlich mit *dem* π identisch, in welchem Kontext man es auch immer kennengelernt haben mag.

Wie haben oben schon gesehen, daß $\mathrm{e}^{\pi/2} = i$ ist. Danach ist dann

$$\mathrm{e}^{x+\frac{\pi}{2}} = \mathrm{e}^x \cdot i$$

für alle x. Schreibt man die letzte Gleichung für den Real- und Imaginärteil getrennt auf, so lautet sie

$$\text{(G)} \qquad \cos(x + \frac{\pi}{2}) = -\sin x, \qquad \sin(x + \frac{\pi}{2}) = \cos x.$$

Der Cosinus ist in $(0, \frac{\pi}{2})$ strikt positiv, als gerade Funktion also auf dem Intervall $(-\frac{\pi}{2}, \frac{\pi}{2})$ strikt positiv. Wegen (G) ist daher der Sinus auf dem Intevall $(0, \pi)$ strikt positiv. Wegen $\cos' = -\sin$ folgt nach einem oben bewiesenen Satz

Satz. *Der Cosinus ist auf dem Intervall* $[0, \pi]$ *streng monoton fallend.*

Nach (G) haben wir dann sofort auch noch

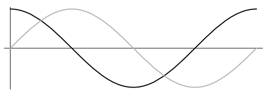
Satz. *Der Sinus ist auf dem Intervall* $[-\frac{\pi}{2}, -\frac{\pi}{2}]$ *streng monoton steigend.*

Aus (G) folgt noch — indem man x durch $x + \frac{\pi}{2}$ ersetzt —

$$\cos(x + \pi) = -\cos x,$$

d.h. der Cosinus ist schon durch seine Werte im Intervall $[0, \pi]$ festgelegt. Fassen wir zusammen:

> Der Cosinus ist stetig und differenzierbar. Er ist periodisch mit der Periode 2π. Von 0 nach π fällt er streng monoton vom Wert 1 zum Wert -1, von $[\pi, 2\pi]$ steigt er streng monoton vom Wert -1 nach 1.



Eine entsprechende Beschreibung überlegt man sich leicht für den Sinus (etwa mittels der Identiäten (G)). Wir erhalten damit auch eine sehr bildhafte Vorstellung von der Abbildung $x \to e^{ix}$: durchläuft x die reelle Zahlengerade, so durchläuft e^{ix} den Einheitskreis $\mathbb{S}^1$, und zwar so, daß der Punkt e^{ix} genau einmal im mathematisch positiven Sinn auf den Einheitskreis umläuft, wenn x ein volles Intervall der Länge 2π durchläuft.

Die Abbildungen

$$[0, \pi] \to [-1, +1], \quad x \mapsto \cos x,$$
$$[-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, +1], \quad x \mapsto \sin x$$

sind streng monoton fallend bzw. streng monoton steigend; also existieren hierzu die Umkehrfunktionen; sie werden mit

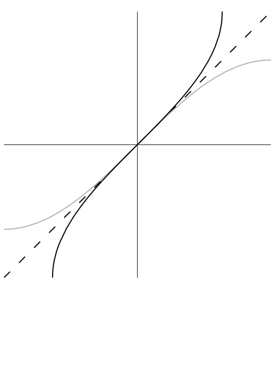
$$\arccos, \quad \text{bzw.} \quad \arcsin$$

bezeichnet. Aufgrund unserer bisher entwickelten Theorie wissen wir, daß diese Funktionen stetig sind, und im Innern ihres Definitionsbereichs (d.h. im Definitionsbereich ohne die Randpunkte) differenzierbar sind:

$$\arccos' y = \frac{-1}{\sqrt{1 - y^2}} \qquad \arcsin' y = \frac{1}{\sqrt{1 - y^2}}.$$

($x = \arcsin y$, $y = \sin x$, $\frac{dy}{dx} = \cos x$ also

$$\frac{dx}{dy} = \frac{1}{\cos x} = \frac{1}{\sqrt{1 - \sin^2 x}} = \frac{1}{\sqrt{1 - y^2}} \text{ !)}$$



Den Graphen von arcsin erhält man, indem man den Graphen von $\sin x$ an der Geraden $y = x$ spiegelt, d.h. indem man jeweils die Koordinaten der Punkte von $G_{\sin}$ vertauscht. Entsprechendes gilt natürlich auch für den Arcuscosinus, wie überhaupt für jede Umkehrfunktion.

Wichtig sind noch die Funktionen

$$\begin{cases} \tan x := \frac{\sin x}{\cos x} & \text{für } x \in \mathbb{R} \setminus \left(\frac{\pi}{2} + \pi\mathbb{Z}\right), \\ \cot x := \frac{\cos x}{\sin x} & \text{für } x \in \mathbb{R} \setminus \pi\mathbb{Z}. \end{cases}$$

Sie sind stetig und differenzierbar; es sind ungerade Funktionen, und sie haben die Periode π. Sie gehen auseinander vermöge

$$\tan(x + \frac{\pi}{2}) = -\cot(x)$$

hervor. Ihre Ableitungen sind

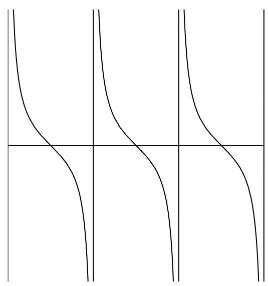
$$\tan' x = \frac{1}{\cos^2 x} = 1 + \tan^2 x, \quad \cot' x = \frac{-1}{\sin^2 x} = -(1 + \cot^2)x.$$

Insbesondere ist die Ableitung des Cotangens strikt negativ, und somit ist der Cotangens im jedem Intervall seines Definitionsbereichs streng monoton fallend; läuft x von 0 nach pi, so läuft $\cot y$ von $+\infty$ to $-\infty$.

Predigt an den Cotangens♣

Bei Dir ist alles Mahnen,
o Cotangens, vertan,
Du gleitest, hochgeboren,
hinab die schiefe Bahn.

Zwar bremst Du dazwischen Dein Sinken,
als packte Dich heilsame Reu,
doch kurz nur währt die Besinnung,
und hemmungslos fällst Du aufs neu.



Doch wenn Du bis minus Unendlich
gestürzt bist, Du haltloser Tor,
dann hebt Dich ein rettender Zauber
auf plus Unendlich empor.

Und wieder in lichten Höhen
erscheinst Du wunderbar,
doch gleicht Dein n–tes Leben
dem $(n-1)$–ten aufs Haar.

Du lernst nichts aus der Geschichte,
Du läufst im alten Trab,
unendlich oft wirst Du gehoben,
unendlich oft stürzt Du hinab.

♣ Dieser und die weiteren Verse in diesem Kapitel sind [H. Cremer: Carmina Mathematica, Verlag J.A.Mayer, Aachen] entnommen.

Den Tangens kann man ähnlich diskutieren:

Elegie um den Tangens

Aus verachteter Tiefe
quälst Du Dich aufwärts
in drückender Hitze
zum siebenten Himmel.

Doch was Du erspart hast,
das rauben bekanntlich
Inflationen und Kriege
alle π Jahre —
leicht läßt sich's beweisen.

Betrogener Tangens!
Bei so viel Tragik
bleibt mir die Spucke
weg und der Reim.

Die Umkehrfunktionen der Abbildungen

$$\begin{aligned} (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R},\ x \mapsto \tan x \\ (0, \pi) \to \mathbb{R},\ x \mapsto \cot x \end{aligned}$$

werden jeweils mit

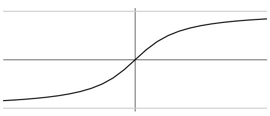
$$\arctan, \qquad \operatorname{arccot}$$

bezeichnet. Ihre Eigenschaften ergeben sich aus denen des Tangens und des Cotangens. Ihre Ableitungen sind

$$\arctan' y = \frac{1}{1+y^2} \qquad \operatorname{arccot}' y = \frac{-1}{1+y^2}.$$

Hier noch eine kurze Diskussion

Ode an die Arcustangensschlange



Du schleichst seit undenklichen Zeiten
so leis und so sanft heran,
Du stiegst in Ewigkeiten
kaum um ein δ an.
Nur langsam beginnst Du zu wachsen,
wie zum Beweis Deines Seins,
erreichst beim Schnittpunkt der Achsen
Deine höchste Steigung, die Eins.

Dann duckst Du Dich wieder zierlich
in stiller Bescheidenheit
und wandelst weiter manierlich
in die Unendlichkeit.

Hier stock ich im Lobgesange,
mir schwant, er wird mir vermiest:
Oh, Arcustangens-Schlange,
beißt Du nicht doch, Du Biest?!

9 TAYLOR-ENTWICKLUNG

In den vorangegangenen Kapiteln haben wir verschiedene Beispiele von Funktionen f gesehen, die sich als *Potenzreihe um einen Enwicklungspunkt* t schreiben lassen, d.h. eine Darstellung der Gestalt

$$f(x) = \sum_{\nu=0}^{\infty} a_\nu (x-t)^\nu$$

besitzen. Es liegt auf der Hand, zu fragen, welche Funktionen überhaupt sich so darstellen lassen. Um diese Frage zu beantworten wird man zum einen studieren, welche Eigenschaften einer Funktion sich daraus ergeben, daß sie eine Potenzreihendarstellung besitzt, zum anderen wird man nach hinreichenden Kriterien für die Existenz einer solchen Darstellung suchen. Ersteres werden wir im nächsten Kapitel diskutieren, das zweite ist der Inhalt dieses Kapitels.

Beim Studium des zweiten Problems wird man zunächst einmal zu gegebener Funktion Zahlen a_ν bestimmen wollen, die als Kandidaten einer Potenzreihenentwicklung um einen gegebenen Punkt t in Frage kommen. Hierzu wiederum wird man zunächst einfache Beispiele studieren. Die einfachsten Funktion, die eine Darstellung der gefragten Art besitzt, ist ein Polynom

$$f(x) = \sum_{\nu=0}^{n} a_\nu \cdot x^\nu.$$

Dieses hat per definitionem eine Potenzreihenentwicklung um $t = 0$. Es hat aber auch eine solche Entwicklung um jede beliebige andere vorgegebene Zahl t: man schreibe in der Entwicklung um 0 die Variable x als $(x - t) + t$ und beachte, daß jede einzelne der Potenzen $((x-t)+t)^\nu$ nach dem Binomischen Lehrsatz eine Entwicklung um t besitzt. Wir können also

$$f(x) = \sum_{\nu=0}^{\infty} b_\nu (x-t)^\nu$$

mit geeigneten Zahlen b_ν schreiben. Natürlich kann man die b_ν aus dem Binomischen Lehrsatz gewinnen, allerdings involviert diese Formel die Koeffizienten a_ν. Im Sinne unserer Fragestellung ist es aber effizienter nach einer Formel zu suchen, die die b_ν unmittelbar aus f bestimmt, ohne Bezug auf die a_ν. Denn ersetzen wir in solch einer Formel das Polynom f durch solche Funktionen, für die die Formel noch sinnvoll bleibt, so haben wir einen großen Schritt getan: wir haben Kandidaten für die Koeffizienten einer potentiellen Potenzreihenentwicklung.

Eine Formel der gesuchten Art für die b_ν gibt es in der Tat. Differenzieren wir nämlich $f(x)$ genau k–mal nach x, und nutzen wir aus, daß

$$\frac{d^k}{dx^k}(x-t)^\nu = \nu(\nu-1)(\nu-2)\cdots(\nu-k+1)(x-t)^{\nu-k}$$

gilt, so erhalten wir

$$\frac{d^k}{dx^k}f(x) = \sum_{\nu=0}^{n} b_\nu \cdot \nu(\nu-1)(\nu-2)\cdots(\nu-k+1)(x-t)^{\nu-k}.$$

Setzen wir hierin $x = t$, so folgt

$$f^{(k)}(t) = b_k \cdot k!$$

Damit haben wir bewiesen:

Satz. *Ist $f(x)$ ein Polynom und t eine gegebenen reelle Zahl, so gilt*

$$f(x) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(t)}{\nu!}(x-t)^\nu.$$

Zwei Beispiele: ist

$$f(x) = x^2 + 3x + 4,$$

so findet man nach dem eben bewiesenen Satz unmittelbar

$$f(x) = (x-t)^2 + (2t+3)(x-t) + \left(t^2 + 3t + 4\right).$$

Dies kann man natürlich auch mit dem binomischen Lehrsatz nachrechnen. Als zweites betrachten wir das Polynom x^n: Hier ist

$$\frac{1}{k!}\frac{d^k}{dx^k}x^n\Big|_{x=t} = \frac{n(n-1)\cdots(n-k+1)}{k!}t^{n-k} = \binom{n}{k}t^{n-k},$$

und daher nach dem letzten Satz

$$x^n = \sum_{\nu}^{n} \binom{n}{\nu} t^{n-\nu}(x-t)^\nu.$$

Dies ist der binomische Lehrsatz, den wir in den Beweis des Satzes hineingesteckt haben, und wie wir sehen, bekommen wir ihn also auch wieder zurück.

Sei jetzt f eine auf einem Intervall (a, b) erklärte, $(n-1)$-mal differenzierbare Funktion und $t \in (a, b)$.

Taylor-Polynom

Das Polynom

$$\sum_{\nu=0}^{n-1} \frac{f^{(\nu)}(t)}{\nu!} \cdot (x-t)^{\nu}$$

heißt das $(n-1)$–*te Taylorpolynom von* f *um* t.

Nach obigem Satz ist jedes Polynom vom Grad n gleich seinem $(n-1)$–ten Taylorpolynom um jeden beliebigen Punkt. Es ist naheliegend zu fragen, wie gut eine beliebig gegebene Funktion durch ihre Taylorpolynome approximiert wird, d.h. wie groß das *Restglied*

$$f(x) - ((n-1)\text{–tes Taylorpolynom um } t)(x)$$

ist. Hat man erst einmal eine geeignete Beschreibung dieses Restgliedes, so sind wir schon sehr nah bei ersten Antworten auf unsere uns leitende Fragestellung nach Kriterien für eine Potenzreihendarstellung. Man wird erwarten, daß das Restglied desto kleiner ist, je mehr f einem Polynom ähnelt. Eine charakteristische Eigenschaft von Polynomen ist es, daß ihre höheren Ableitungen identisch verschwinden. Also wird man erwarten, daß das Restglied desto kleiner ist, je “kleiner” die höheren Ableitungen von f sind, und allgemein sollte es eine Beschreibung des Restgliedes mittels höherer Ableitungen von f geben. Diese werden wir nun herleiten.

Dazu setzen wir voraus, daß f eine auf (a, b) mindestens n–mal differenzierbare Funktion ist und t irgendein beliebiger Punkt aus (a, b). Dann können wir schreiben

$$f(x) = f(t) + \frac{f'(t)}{1!}(x-t) + \cdots + \frac{f^{(n-1)}(t)}{(n-1)!}(x-t)^{n-1} + R_n(x, t),$$

mit einem durch eben diese Gleichung definierten *Fehler* $R_n(x, t)$. Wir wählen jetzt ein festes x in (a, b) und betrachten die Funktion

$$t \mapsto R_n(x, t) = f(x) - \left(f(t) + \frac{f'(t)}{1!}(x-t) + \cdots + \frac{f^{(n-1)}(t)}{(n-1)!}(x-t)^{n-1} \right).$$

Wir bezeichnen sie der Einfachheit halber mit $R(t)$. Da f ja n–mal differenzierbar sein soll, ist $R(t)$ in (a, b) mindestens einmal differenzierbar. Wir berechnen $R'(t)$: differenzieren wir den k–ten Term in der Klammer auf der rechten Seite der $R(t)$ definierenden Gleichung, so finden wir

$$\begin{aligned} &\frac{d}{dt} \frac{f^{(k)}(t)}{k!}(x-t)^k \\ &= \frac{d}{dt} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} \end{aligned}$$

für $k > 0$, und $= f'(t)$ für $k = 0$; demnach hebt sich der erste Summand der Ableitung des k-ten Terms jeweils gegen den zweiten Summanden der Ableitung des $(k+1)$–ten Terms auf, und wir finden

$$R'(t) = -\frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1}.$$

Nach dem Mittelwertsatz ist nun

$$R(t)(\ = R(t) - R(x)\ ) = R'(c) \cdot (t - x)$$

mit einem geeigneten c zwischen t und x. Setzen wir hierin unsere soeben abgeleitete Formel für $R'(t)$ ein, so sehen wir, daß wir den folgenden Satz bewiesen haben:

Cauchysches Restglied

Satz. *Es sei f eine auf dem Intervall (a, b) erklärte, n–mal differenzierbare Funktion, und es seien $x, t \in (a, b)$. Dann gib es ein c zwischen x und t, sodaß*

$$f(x) = f(t) + \frac{f'(t)}{1!}(x-t) + \cdots + \frac{f^{(n-1)}(t)}{(n-1)!}(x-t)^{n-1} \\ + \frac{f^{(n)}(c)}{(n-1)!}(x-c)^{n-1}(x-t).$$

Wir geben noch eine andere Beschreibung des Restgliedes. Hierzu benötigen wir eine etwas allgemeinere Formulierung des Mittelwertsatzes.

Cauchyscher Mittelwertsatz

Satz. *Seien f und g auf $[a, b]$ erklärte, stetige und in (a, b) differenzierbare Funktionen, und es sei $g(x) \neq 0$ für jedes $x \in (a, b)$. Dann gibt es ein $c \in (a, b)$, sodaß*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Man beachte, daß $g(b) - g(a) \neq 0$ (denn andernfalls gäbe es nach dem Satz von Rolle ein x mit $g'(x) = 0$).

Beweis. Die Funktion

$$f(x) - f(a) - \frac{f(b) - f(a)}{g(b) - g(a)} \cdot (g(x) - g(a)) =: h(x)$$

erfüllt die Voraussetzungen des Satzes von Rolle, d.h. es ist $h(b) = h(a) = 0$. Also gibt es nach ebendiesem Satz ein c mit $h'(c) = 0$, d.h.

$$0 = h'(c) = f'(c) - \frac{f(b) - f(a)}{f(b) - g(a)} \cdot g'(c),$$

und das ist gerade die Behauptung.

Wie angekündigt können wir mittels des verallgemeinerten Mittelwertsatzes noch eine andere Beschreibung des oben studierten Restgliedes $R(t)$ erhalten. Wenden wir ihn nämlich auf $f(t) = R(t)$ und $g(t) = (x-t)^n$ an, so finden wir

$$\frac{R(t)}{(x-t)^n} = \frac{R(x)-R(t)}{0-(x-t)^n} = \frac{R'(c)}{-n(x-c)^{n-1}}$$

mit einem geeigneten c zwischen x und t. Setzen wir hierin unsere oben berechnete Formel für $R'(t)$ ein, so erhalten wir den

Lagrangesches Restglied

Satz. *Es sei f eine auf dem Intervall (a,b) erklärte, n–mal differenzierbare Funktion, und es seien $x,t \in (a,b)$. Dann gib es ein c zwischen x und t, sodaß*

$$f(x) = f(t) + \frac{f'(t)}{1!}(x-t) + \ldots + \frac{f^{(n-1)}(t)}{(n-1)!}(x-t)^{n-1} + \frac{f^{(n)}(c)}{n!}(x-t)^n$$

ist.

Wir kehren nun zu unserer ursprünglichen Fragestellung nach der Möglichkeit einer Potenzreihenentwicklung zurück. Wir verabreden

Glatte Funktionen

> Eine auf dem Intervall (a,b) erklärte Funktion heißt *glatt* oder C^∞-Funktion, falls sie unendlich oft differenzierbar ist, d.h. falls die k–te Ableitung $f^{(k)}$ für jedes jedes $k \in \mathbb{N}$ existiert.

Ferner verabreden wir:

Taylor-Reihe

> Ist f eine auf dem Intervall (a,b) erklärte C^∞–Funktion, und ist $x_0 \in (a,b)$, so heißt die unendliche Reihe
>
> $$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} \cdot (x-x_0)^k$$
>
> *Taylor-Reihe von f um x_0*

Es ergeben sich sofort zwei Fragen: für welche x konvergiert diese Taylorreihe? Falls sie für ein x konvergiert, ist der Grenzwert dann $f(x)$? Wir werden gleich Beispiele dafür sehen, daß die Taylor-Reihe für alle $x \in (a,b)$ konvergiert, und zwar gegen $f(x)$. Im allgemeinen kann man aber zeigen, daß *jede beliebige* Reihe der Gestalt $\sum_{n=0}^{\infty} a_n x^n$ als Taylor-Reihe einer glatten Funktion erhalten werden kann, sodaß es insbesondere Funktionen gibt, deren Taylorreihe nur für $x = x_0$ konvergiert. Umgekehrt kann man Funktionen konstruieren, deren Taylorreihe für alle x konvergiert, aber nur für $x = x_0$ gegen $f(x)$.

Immerhin können wir mittels unserer Restgliedbeschreibungen ein hinreichendes Kriterium für die Darstellbarkeit einer Funktion durch ihre Taylor-Reihe aufstellen. Hierzu führen wir noch eine Sprechweise ein.

Gleichmäßig beschränkte Ableitungen

Ist f eine auf dem Intervall (a, b) erklärte C^∞–Funktion, so sagen wir *"die Ableitungen von f sind gleichmäßig beschränkt auf (a, b)"*, falls es eine Konstante M gibt, sodaß

$$\left|f^{(k)}(x)\right| \leq M$$

für alle $k \in \mathbb{N}$ und alle $x \in (a, b)$ gilt.

Kriterium der gleichmäßigen Beschränktheit.

Satz. *Sei f eine auf dem Intervall (a, b) erklärte C^∞-Funktion. Die Ableitungen von f seien im Intervall (a, b) gleichmäßig beschränkt. Dann konvergiert für jedes $x_0 \in (a, b)$ und für jedes $x \in (a, b)$ die Taylorreihe von f um x_0 gegen $f(x)$, d.h. für alle $x, x_0 \in (a, b)$ ist*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

Beweis. Es ist mit einem geeigneten (von x und x_0 abhängenden) c zwischen x und x_0 und einer von x, x_0 und n unabhängigen Konstanten M

$$\begin{aligned}\left|f(x) - \sum_{k=0}^{(n-1)} \frac{f^k(x_0)}{k!}(x - x_0)^k\right| &= \left|\frac{f^{(n)}(c)}{n!}(x - x_0)^n\right| \\ &\leq \frac{M}{n!}\left|x - x_0\right|^n =: a_n.\end{aligned}$$

Die erste Gleichung folgt aus der Lagrangeschen Restgliedformel, die Ungleichung aus der vorausgesetzten gleichmäßigen Beschränktheit der Ableitungen von f in (a, b). Wir lassen es als Übungsaufgabe, nachzuprüfen, daß die a_n eine Nullfolge bilden. Hiermit folgt dann unmittelbar die Behauptung des Satzes.

Wir illustrieren die Bedeutung der hier bewiesenen Sätze an einigen Beispielen.

Charakterisierung von cos, sin und exp mittels Differentialgleichungen

Sei f eine auf ganz $\mathbb{R}$ definierte C^∞–Funktion. Es gelte

$$f'' = -f.$$

Beispiele für solche f sind die Funktionen cos und sin. Wir behaupten, daß dies aber im Wesentlichen auch schon alle solche Funktionen sind, genauer: jedes f der betrachteten Art ist eine Linearkombination von sin und cos. In der Tat: zunächst bemerken wir, daß die Ableitungen von f auf jedem endlichen Intervall (a, b) gleichmäßig beschränkt sind — es ist ja $\left|f^{(k)}(x)\right| = |f(x)|$ oder $= |f'(x)|$, und f und f' sind stetig in $\mathbb{R}$, also insbesondere auf $[a, b]$ beschränkt. Nach dem Satz von der gleichmäßigen Beschränktheit ist daher für alle x und x_0

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \\ &= f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) \\ &- \frac{f(x_0)}{2!}(x - x_0)^2 - \frac{f'(x_0)}{3!}(x - x_0)^3 \\ &+ \frac{f(x_0)}{4!}(x - x_0)^4 + \frac{f'(x_0)}{5!}(x - x_0)^5 - \cdots . \end{aligned}$$

Vergleichen wir dies mit den Reihenentwicklungen von $\sin x$ und $\cos x$, so finden wir

$$f(x) = f(x_0)\cos(x - x_0) + f'(x_0)\sin(x - x_0),$$

und das haben wir behauptet. Man beachte, daß wir gleichzeitig noch verifiziert haben, daß die Reihenentwicklungen von $\sin x$ und $\cos x$ gerade ihre Taylor-Reihen um $x = 0$ sind. Dies ist kein Zufall: eine Erklärung dafür wird sich im nächsten Kapitel ergeben.

Mit der gleichen Methode wie eben kann man sich zur Übung die folgnede Aussage überlegen:

> Es sei f eine auf $\mathbb{R}$ definierte C^∞–Funktion, und es gelte $f' = f$. Dann gibt es eine Konstante c, sodaß $f = c \cdot \exp$.

Insbesondere wird man dabei noch finden, daß die die Exponentialfunktion definierende Reihe gerade ihre Taylor-Reihe um $x = 0$ ist.

Die Taylorreihe, sofern sie gegen die betrachtete Funktion konvergiert, ist zunächst einmal ein Mittel zur Berechnung dieser Funktion. Zur Illustration dieser Aussage berechnen wir e: nach der Lagrangeschen Restgliedformel hat man

$$e = e^1 = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} ... + \frac{1}{8!} + \frac{e^c}{q!}$$

mit einem c zwischen 0 und 1. Da die Exponentialfunktion monoton steigt, ist $\mathrm{e}^c < e$. Nun ist aber

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + ... < 1 + 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = 3,$$

und so

$$e - \left(1 + 1 + \frac{1}{2!} + \frac{1}{3!} ... + \frac{1}{8!}\right) < \frac{3}{9!} = 0,000008 \ldots .$$

Potenzreihen-darstellung von $\log x$

Als weiteres Beispiel betrachten wir $\log x$. Zunächst ist jedenfalls $\log x$ nach den Ergebnissen des letzten Kapitels eine für $x > 0$ erklärte C^∞-Funktion: dort wurde $\log' x = \frac{1}{x}$ bewiesen, und offenbar ist $\frac{1}{x}$ glatt; letzteres zeigt auch

$$\frac{d^k}{dx^k} \log x = (-1)^{k-1} \frac{(k-1)!}{x^k}.$$

Wir studieren damit nun die Taylor-Reihe von $\log(1+x)$ um $x = 0$: nach der Lagrangeschen- bzw. der Cauchyschen Restgliedformel haben wir zunächst für jedes $x > 0$

$$\log(1+x) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} x^k + R_n(x),$$

wobei

$$R_n(x) = \frac{(-1)^{n-1}}{n(1+c)^n} x^n$$

bzw. in der Cauchyschen Form

$$R_n(x) = \frac{(-1)^{n-1}}{(1+d)^n} (x-d)^{n-1} x$$

mit einem (von x abhängenden) c bzw. d zwischen 0 und x. Demnach hat man

$$x = 1 \Longrightarrow |R_n(x)| \leq \frac{1}{n}$$
$$|x| < 1 \Longrightarrow |R_n(x)| \leq C(x) \cdot |x|^n ,$$

letzteres mit einer geeigneten von x abhängenden, aber von n unabhängigen reelle Zahl $C(x)$. Hierbei ergibt sich die erste Aussage aus der Lagrangeschen Beschreibung des Restgliedes, und die zweite aus der Cauchyschen. In jedem Fall sehen wir so, daß bei vorgegebenem $x \in (-1, 1]$ die Folge der Restglieder $R_n(x)$ gegen 0 konvergiert. Wir haben damit den folgenden Satz bewiesen:

Satz. *Für alle $x \in (-1, 1]$ gilt*

$$\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k.$$

Man beachte hierbei den Spezialfall $x = 1$:

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \mp \cdots .$$

Der hier rechts stehenden Reihe, der *Leibniz Reihe*, sind wir früher schon als Beispiel für einer konvergente, aber nicht absolut konvergente Reihe begegnet; jetzt haben wir sogar ihren Wert berechnet.

Diese Reihe ist allerdings denkbar ungeeignet, um $\log 2$ zu berechnen, denn die einzelnen Glieder der Reihe konvergieren relativ langsam gegen 0. Eine bessere Methode geht folgendermaßen (nach James Gregory 1668[†]). Es ist für $|x| < 1$

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots,$$

und daher auch

$$\log(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots.$$

Subtrahieren wir die zweite dieser beiden Gleichungen von der ersten und benutzen wir noch die Regel $\log a - \log b = \log \frac{a}{b}$, so finden wir

$$\log\frac{1+x}{1-x} = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots\right).$$

Dies ist schon eine bessere Darstellung des Logarithmus als in der Taylorreihe im Satz: die Glieder der Reihe konvergieren etwas schneller gegen 0. Man kann diese Methode aber noch verfeinern. Für natürliche Zahlen p, q, $p > q > 0$ hat man insbesondere

$$\begin{aligned}\log p - \log q &= \log\frac{p}{q}\\ &= \log\frac{1+\frac{p-q}{p+q}}{1-\frac{p-q}{p+q}}\\ &= 2\left(\frac{p-q}{p+q} + \frac{1}{3}\left(\frac{p-q}{p+q}\right)^3 + \frac{1}{5}\left(\frac{p-q}{p+q}\right)^5 + \cdots\right)\end{aligned}$$

(man beachte hierzu, daß stets $0 < \frac{p-q}{p+q} < 1$ ist). Dies wenden wir an, um $\log 2$, $\log 3$ und $\log 5$ simultan zu berechnen. Dazu wählen wir in der letzten Formel

$$\begin{aligned}p &= 16, q = 15\\ p &= 25, q = 24\\ p &= 81, q = 80,\end{aligned}$$

und erhalten damit jeweils

$$\begin{aligned}4\log 2 \; - \log 3 \; - \log 5 &= 2\left(\tfrac{1}{31} + \tfrac{1}{3}\left(\tfrac{1}{31}\right)^3 + \cdots\right)\\ -3\log 2 \; - \log 3 + 2\log 5 &= 2\left(\tfrac{1}{49} + \tfrac{1}{3}\left(\tfrac{1}{49}\right)^3 + \cdots\right)\\ -4\log 2 + 4\log 3 \; - \log 5 &= 2\left(\tfrac{1}{161} + \tfrac{1}{3}\left(\tfrac{1}{162}\right)^3 + \cdots,\right)\end{aligned}$$

wobei wir $\log p - \log q$ sogleich unter Ausnutzung des Additionstheorems des Logarithmus in der Gestalt der linken Seiten dieser Gleichungen geschrieben

[†] vgl. [Gerhard Kovalewski: Die Klassischen Probleme der Analysis des Unendlichen. Leipzig 1921]

haben: dabei geht ein, daß die p, q gerade so gewählt sind, daß sie nur die Primteiler 2,3,5 enthalten. Wir haben nun ein Gleichungssystem in den drei "Unbekannten" $\log 2$,$\log 3$, $\log 5$. Durch Multiplikation der Gleichungen 1 bis 3 mit 7,5,3 respektive und anschließendem Aufaddieren der resultierenden Gleichungen erhält man

$$\begin{aligned}\log 2 = 14&\left(\frac{1}{31} + \frac{1}{3}\cdot\left(\frac{1}{31}\right)^3 + \cdots\right)\\ +10&\left(\frac{1}{49} + \frac{1}{3}\cdot\left(\frac{1}{49}\right)^3 + \cdots\right)\\ +6&\left(\frac{1}{161} + \frac{1}{3}\cdot\left(\frac{1}{161}\right)^3 + \cdots\right).\end{aligned}$$

Bricht man die Reihe nach dem zweiten Glied ab, so ergibt sich die Näherung

$$\log 2 \approx 0.693147076.$$

Diese ist tatsächlich auf 6 Stellen hinter dem Komma genau. Ähnlich gute Näherungen ergeben sich hierbei auch für $\log 3$ und $\log 5$.

Die Binomialreihe

Als letztes Beispiel verallgemeinern wir den Binomischen Lehrsatz. Dazu betrachten wir die für alle $x > -1$ erklärte Funktion

$$x \mapsto (1+x)^r = \exp(r \log(1+x)).$$

Ist $r \in \mathbb{N}$, so ist nach dem binomischen Lehrsatz

$$(1+x)^r = \sum_{k=0}^{r} \binom{r}{k}.$$

Hierbei ist

$$\binom{r}{k} = \frac{r(r-1)\cdots(r-k+1)}{k!}.$$

Offenbar hat die rechte Seite der letzten Gleichung für alle Zahlen r einen Sinn, und wir nehmen diese Gleichung als Definition von $\binom{r}{k}$ für beliebige r. Wir vereinbaren dabei noch

$$\binom{r}{0} = 1$$

(es ist fast immer richtig, ein Produkt mit 0 Faktoren, ein *leeres Produkt*, als 1 zu interpretieren.) Man beachte nun, daß $\binom{r}{k} = 0$ für natürliche $r < k$ gilt. Damit können wir in dem oben angeführten Binomischen Lehrsatz r durch das Symbol ∞ ersetzen, und es ist klar, daß man nun für nicht notwendig natürliche r erwartet, daß der Binomische Lehrsatz richtig bleibt, jedenfalls für gewisse x.

Satz. *Für jedes r und alle $|x| < 1$ gilt*

$$(1+x)^r = \sum_{k=0}^{\infty} \binom{r}{k}.$$

Beweis. Wir betrachten für festes $r \in \mathbb{R}$ die Funktion $f(x) := (1+x)^r$ auf dem Intervall $(-1, \infty)$. Die Ableitungen sind

$$\begin{aligned} f(x) &= (1+x)^r \\ f'(x) &= r(1+x)^{r-1} \\ f''(x) &= r(r-1)(1+x)^{r-2} \\ &\vdots \\ \frac{f^{(k)}(x)}{k!} &= \binom{r}{k}(1+x)^{r-k}. \end{aligned}$$

Die Taylorreihe von f bei $x = 0$ ist also genau die *Binomialreihe*, d.h. die Reihe auf der rechten Seite in der behaupteten Gleichung. Wir betrachten das Cauchysche Restglied $R_n(x)$ zum $(n-1)$–ten Taylorpolynom um $x = 0$:

$$R_n(x) = n\binom{r}{n}(1+c)^{r-n}(x-c)^{n-1}x$$

(c zwischen 0 und x.) Mit einer einfachen Abschätzung findet man

$$|R_n(x)| \leq A \cdot \left| n\binom{r}{n}x^n \right|$$

für $|x| < 1$ mit einer von n unabhängigen (allerdings von x abhängenden) Konstanten A. Die hier rechts stehende Größe ist aber das n–te Glied eine Nullfolge (etwa, weil die Folge der Quotienten von je zwei aufeinanderfolgenden Gliedern dieser Folge gegen $|x|$ konvergiert und $|x| < 1$ ist !) Damit folgt nun sofort die behauptete Identiät.

Abel hat noch mehr bewiesen, als wir im Satz ausgesprochen haben♣:

> "Alsdann ist
> $$1 + \frac{m}{1} \cdot a + \frac{m(m-1)}{1 \cdot 2} \cdot a^2 + \frac{m \cdot (m-1)(m-2)}{1 \cdot 2 \cdot 3} \cdot a^2 + \cdots = (1+a)^m.$$
>
> Dieser Ausdruck findet für jeden Werth von m statt, wenn der Zahlenwerth von a kleiner ist, als 1; ferner für jeden Werth von m, zwischen -1 und $+\infty$, wonn $a = 1$ ist, und für jeden positiven Werth von m, wenn $a = -1$ ist. Für andere Werthe von a und m ist das erste Glied eine divergente Reihe."

Wir geben abschließend noch drei theoretische Folgerungen unserer bisher entwickelten Theorie der Taylorreihen.

♣ Dies wurde im ersten Band von Crelles Journal [Journal Reine Angew. Math.] veröffentlicht

Satz. *Es sei f in (a,b) erklärt und $(n-1)$-mal differenzierbar, und außerdem sei $f^{(n-1)}$ noch in $x_0 \in (a,b)$ differenzierbar. Dann ist*

$$\lim_{x \to x_0} \frac{f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k}{(x-x_0)^n} = \frac{f^{(n)}(x_0)}{n!}.$$

Beweis. Wir beweisen die Behauptung des Satzes durch Induktion über n. Für $n = 1$ ist nichts zu beweisen. Angenommen, die Behauptung des Satzes ist richtig für $n-1$. Nach dem Cauchyschen Mittelwertsatz ist zunächst

$$\frac{f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k}{(x-x_0)^n}$$
$$= \frac{f'(c) - \sum_{k=1}^{n-1} \frac{f^{(k)}(x_0)}{(k-1)!}(c-x_0)^{k-1}}{n \cdot (c-x_0)^{n-1}},$$

wobei c zwischen x und x_0 liegt. Durchläuft hier x im ersten Ausdruck eine gegen x_0 konvergente Folge, so durchläuft die Zwischenstelle c eine ebenfalls gegen x_0 konvergente Folge; ferner ist hier die Summe auf der rechten Seite gerade das $(n-1)$-te Taylorpolynom von f' an einer Zwischenstelle c. Nach Induktionsannahme konvergiert daher die rechte Seite gegen

$$\frac{1}{n} \cdot \frac{(f')^{(n-1)}(x_0)}{(n-1)!},$$

d.h. gegen $f^{(n)}(x_0)/n!$, und das war zu zeigen.

Ein Extremwertkriterium

Satz. *Es sei f in (a,b) erklärt und $(n-1)$-mal differenzierbar, und außerdem sei $f^{(n-1)}$ noch in $x_0 \in (a,b)$ differenzierbar. Es gelte*

$$f'(x_0) = f''(x_0) = ... = f^{(n-1)}(x_0) = 0$$

und $f^{(n)}(x_0) \neq 0$. Dann gilt: Ist n gerade, so hat f in x_0 ein striktes lokales Minimum oder Maximum, je nachdem ob $f^{(n)}(x_0) > 0$ oder < 0 ist. Ist n ungerade, so hat f in x_0 weder ein lokales Minimum noch ein lokales Maximum.

Beweis. Nach dem letzten Satz und den Voraussetzungen über f konvergiert

$$\Delta(x) := \frac{f(x) - f(x_0)}{(x-x_0)^n}$$

gegen $f^{(n)}(x_0)/n!$, wenn x gegen x_0 strebt, d.h. $\Delta(x)$ kann zu einer in x_0 stetigen Funktion $\underline{\Delta}(x)$fortgesetzt werden, und es ist

$$f(x) - f(x_0) = \underline{\Delta}(x) \cdot (x-x_0)^n.$$

Es ist aber $\underline{\Delta}(x_0)(= f^n x_0/n!) > 0$ (bzw.< 0). Nach einem bekannten Satz über stetige Funktionen gibt es daher eine ganze ϵ–Umgebung von x_0, in der $\Delta(x)$ strikt positiv (bzw. strikt negativ) ist. Für ein x aus dieser Umgebung ist somit das Vorzeichen von $f(x) - f(x_0)$ gleich dem Vorzeichen von $(x-x_0)^n$ (bzw. -1 mal diesem Vorzeichen). Hieraus sind nun die im Satz aufgestellten Behauptungen unmittelbar abzulesen.

Regel von l'Hôpital

Satz. *Die Funktionen f, g seien $(n-1)$-mal differenzierbar in (a, b), es seien $f^{(n-1)}, g^{(n-1)}$ in $x_0 \in (a, b)$ differenzierbar. Gilt dann*

$$f(x_0) = \ldots = f^{(n-1)}(x_0) = g(x_0) = \ldots = g^{(n-1)}(x_0) = 0,$$

aber

$$g^{(n)}(x_0) \neq 0,$$

so ist

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{f^{(n)}(x_0)}{g^{(n)}(x_0)}.$$

Beweis. Nach unserem vorletzten Satz und wegen der Voraussetzungen ist

$$\lim_{x \to x_0} \frac{f(x)}{(x-x_0)^n} = \frac{f^{(n)}(x_0)}{n!},$$
$$\lim_{x \to x_0} \frac{g(x)}{(x-x_0)^n} = \frac{g^{(n)}(x_0)}{n!}.$$

Hieraus ist die behauptete Formel unmittelbar ersichtlich.

10 POTENZREIHEN

Eine Reihe der Gestalt

$$a_0 + a_1 x + a_2 x^2 \cdots = \sum_{k=0}^{\infty} a_k x^k,$$

wobei die a_k irgendwelche reellen Zahlen sind, nennt man *Potenzreihe*. Solch eine Reihe wird für gewisse Zahlen x konvergieren und für andere divergieren.

Satz. *Falls eine Potenzreihe für ein x_1 konvergiert, so ist sie für jedes x mit $|x| < |x_1|$ absolut konvergent.*

Beweis. Da die Reihe für x_1 konvergiert, bilden die Reihenglieder $a_k x_1^k$ eine Nullfolge, sind also insbesondere beschränkt — etwa $\left|a_k x_1^k\right| \leq M$ für alle k. Dann ist aber $\left|a_k x^k\right| \leq M \left|\frac{x}{x_1}\right|^k$, d.h. die geometrische Reihe $\sum_k M \cdot \left|\frac{x}{x_1}\right|^k$ ist eine Majorante für $\sum_k \left|a_k x^k\right|$.

Konvergenzradius

Demgemäß nennt man die Zahl

$$R := \sup\left\{r \mid r \geq 0 \wedge \sum |a_k|\, r^k \text{ ist konvergent}\right\}$$

Konvergenzradius der Potenzreihe $\sum_k a_k x^k$.

Satz. *Ist $|x| < R$, so ist $\sum_{k=0}^{\infty} a_k x^k$ absolut konvergent, ist $|x| > R$, so ist $\sum_{k=0}^{\infty} a_k x^k$ divergent.*

Beweis. Ist $|x| < R$ so gibt es — nach Definition von R — ein r, sodaß $|x| \leq r < R$ und $\sum_{k=0}^{\infty} |a_k|\, r^k$ konvergiert. Nach dem ersten Satz ist daher $\sum_{k=0}^{\infty} a_k x^k$ absolut konvergent. Wäre dagegen $\sum_{k=0}^{\infty} a_k x^k$ für ein $|x| > R$ konvergent, so wäre — wieder nach dem ersten Satz — die Reihe $\sum_{k=0}^{\infty} |a_k|\, r^k$ für jedes $R < r < |x|$ konvergent. Dies steht im Widerspruch zur Supremum-Eigenschaft von R.

Über das Verhalten in den Randpunkten $\pm R$ kann man keine allgemeine Aussage machen. Zum Beispiel ist für

$$\sum_{k=0}^{\infty} \frac{x^k}{k}$$

der Konvergenzradius $R = 1$; bei $x = 1$ ist die Reihe divergent, bei $x = -1$ konvergent. Man beachte ferner, daß auch die extremen Fälle $R = \infty$, d.h. die Potenzreihe konvergiert für alle x, oder auch $R = 0$, d.h. die Potenzreihe konvergiert nur für $x = 0$, auftreten können. Beispiele hierfür sind jeweils die Potenzreihen

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} \qquad \sum_{k=0}^{\infty} k!\, x^k.$$

Satz. *Die Potenzreihen* $\sum_{k=0}^{\infty} a_k x^k$ *und* $\sum_{k=0}^{\infty} k a_k x^{k-1}$ *haben den gleichen Konvergenzradius.*

Beweis. Die Konvergenzradien der beiden Reihen seien R und R'. Wir zeigen zunächst $R < R'$. Sei $|x| < R$. ist dann $|x| < |x_1| < R$, so ist $\sum_k a_k x_1^k$ konvergent, also $\left|a_k x_1^k\right| \leq M$ mit einer geeigneten Konstanten M, also ist $\sum_k kM \left|\frac{x}{x_1}\right|^k$ eine konvergente Majorante der Reihe $\sum_k k a_k x^{k-1}$, mithin $|x| \leq R'$. Zur umgekehrten Ungleichung $R' \leq R$: ist $|x| \leq R'$, so ist $\sum_k k \left|a_k x_1^{k-1}\right|$ konvergent für jedes $|x| < |x_1| < R$; es ist aber $\sum_k k \left|a_k x_1^{k-1}\right|$ eine Majorante für $\sum_k \left|a_k x_1^{k-1}\right|$, also ist $\sum_k \left|a_k x_1^k\right|$ konvergent, also ist erst recht $\sum_k k \left|a_k x^k\right|$ konvergent.

Der Satz ist eine Vorbereitung für den nächsten

Satz. *Der Konvergenzradius der Potenzreihe* $\sum_{k=0}^{\infty} a_k x^k$ *sei* $R > 0$. *Dann wird durch*

$$f(x) := \sum_{k=0}^{\infty} a_k x^k \qquad (|x| < R)$$

eine im Intervall $(-R, +R)$ *definierte, differenzierbare Funktion erklärt. Für die Ableitung gilt*

$$f'(x) = \sum_{k=0}^{\infty} k a_k x^{k-1}.$$

Beweis. Sei $|x| < R$ und h so klein, daß noch $|x| + |h| < R$ ist. Dann ist jedenfalls $|x + h| < R$ und wir können schreiben:

$$\begin{aligned} f(x+h) &= \sum_{k=0}^{\infty} a_k (x+h)^k \\ &= \sum_{k=0}^{\infty} \sum_{\kappa=0}^{k} a_k \binom{k}{\kappa} x^{k-\kappa} h^{\kappa}. \end{aligned}$$

Nun bleibt aber der hier rechts stehende Ausdruck wegen unserer Wahl von x und h sogar noch konvergent, wenn man die einzelne Summanden jeweils durch ihre Absolutbeträge ersetzt; denn mit der gleichen Rechnung wie eben ist

$$\sum_{k=0}^{\infty}\sum_{\kappa=0}^{k}\left|a_k\binom{k}{\kappa}x^{k-\kappa}h^{\kappa}\right| = \sum_{k=0}^{\infty}|a_k|\,(|x|+|h|)^k,$$

und es soll ja $|x|+|h| < R$ gelten. Nach einem allgemeinen Sachverhalt, den wir gleich noch als Satz nachstellen werden, folgt in dieser Situation, daß wir auf der rechten Seite unserer eben abgeleiteten Formel für $f(x+h)$ die Summationsreihenfolge vertauschen dürfen, d.h. daß

$$f(x+h) = \sum_{\kappa=0}^{\infty} f_{\kappa}(x)h^{\kappa}$$

mit

$$f_{\kappa}(x) = \sum_{k=\kappa}^{\infty}\binom{k}{\kappa}a_k x^{k-\kappa}$$

ist, und daß hierbei die Reihe der $f_{\kappa}(x)h^{\kappa}$ und die die $f_{\kappa}(x)$ definierenden Reihen absolut konvergent sind. Damit haben wir dann bei festgehaltenem x eine Darstellung von $f(x+h)$ als Potenzreihe in h. Insbesondere ist

$$\frac{f(x+h)-f(x)}{h} = \sum_{\kappa=1}^{\infty} f_{\kappa}(x)h^{\kappa-1},$$

und da nach einem früheren Satz die durch eine Potenzreihe dargestellte Funktion im Punkt 0 stetig ist, finden wir

$$\lim_{x\to 0}\frac{f(x+h)-f(x)}{h} = f_1(x).$$

Also ist f bei x differenzierbar, und die Ableitung ist

$$f_1(x) = \sum_{k=0}^{\infty} k a_k x^{k-1}.$$

Das war aber gerade zu beweisen.

Die im Beweis angesprochen allgemeine Tatsache ist der folgende

Vertauschung der Summationsreihenfolge bei absolut konvergenten Doppelreihen

Satz. *Die unendliche Doppelreihe*

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}|a_{m,n}| = \lim_{M\to\infty}\sum_{m=0}^{M}\left(\lim_{N\to\infty}\sum_{n=0}^{N}|a_{m,n}|\right)$$

sei konvergent. Dann sind die beiden Doppelreihen

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}a_{m,n} \qquad \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}a_{m,n}$$

konvergent, und ihre Werte sind gleich.

Der Beweis dieses Satzes kann mit ganz ähnlichen Überlegungen wie der Beweis des Satzes über das Produkt unendlicher Reihen im Kapitel 6 erbracht werden; wir lassen ihn als Übungsaufgabe.

Den vorletzten Satz kann man natürlich iterieren. Damit erhält man, daß eine durch eine Potenzreihe definierte Funktion im Innern des Konvergenzintervalls unendlich oft differenzierbar ist, also eine C^∞–Funktion darstellt, und die Ableitungen erhält man stets durch *gliedweises differenzieren.* Funktionen, die eine Potenzreihendarstellung besitzen, haben einen eigenen Namen.

Analytische Funktionen

Eine auf einem Intervall (a, b) erklärte Funktion f heißt analytisch im Punkt $x_0 \in (a, b)$, falls es eine Potenzreihe $\sum\limits_{k=0}^{\infty} a_k x^k$ mit positivem Konvergenzradius eine $\epsilon > 0$ gibt, sodaß

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$$

für alle $|x - x_0| < \epsilon$ gilt.

Eine auf einem offenen Intervall definierte Funktion heißt *analytisch*, falls sie in jedem Punkt des Intervalls analytisch ist. Wie wir im Beweis des vorletzten Satz gesehen haben, gilt

Satz. *Der Konvergenzradius der Potenzreihe $\sum_{k=0}^{\infty} a_k x^k$ sei $R > 0$. Dann wird durch*

$$f(x) := \sum_{k=0}^{\infty} a_k x^k \qquad (|x| < R)$$

eine im Intervall $(-R, +R)$ definierte, analytische Funktion erklärt.

Man beachte, daß keineswegs jede unendlich oft differenzierbare Funktion analytisch ist: $f(x) = \mathrm{e}^{-1/x^2}$ für $x > 0$ und $f(x) = 0$ für $x \leq 0$ ist unendlich oft differenzierbar, aber bei $x = 0$ nicht analytisch. Wichtig ist noch der folgende

Satz. *Ist f analytisch in x_0, d.h. gilt $f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$ für alle x aus einem x_0 enthaltenen offenen Intervall, so gilt*

$$a_k = \frac{f^{(k)}(0)}{k!}.$$

Beweis. Man differenziere $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ gliedweise und setze $x = x_0$. Dies liefert die behauptete Formel.

Insbesondere zeigt der eben bewiesene Satz, daß die *Potenzreihenentwicklung* einer analytischen Funktion um einen vorgegebenen Punkt eindeutig bestimmt ist: es ist gerade die Taylorentwicklung um diesen Punkt. Deshalb spricht man bei analytischen Funktionen auch von *der* Potenzreihenentwicklung um einen vorgegebenen Punkt.

Reihenentwicklung des Arcustangens

Als Anwendung der kleinen Theorie über Potenzreihen berechnen wir die Potenzreihenentwicklung des $\arctan x$ bei $x = 0$. Die Funktion

$$\tan x = \frac{\sin x}{\cos x}$$

ist definiert für $x \in \mathbb{R} \setminus \left(\frac{\pi}{2} + \pi\mathbb{Z}\right)$. Sie ist unendlich oft differenzierbar (sogar analytisch — was uns hier aber nicht interessiert) und periodisch: $\tan(\ x \ + \ \pi) = \tan(x)$ für alle x. Es ist leicht zu sehen, daß sie bei Einschränkung eine bijektive Abbildung $(-\frac{\pi}{2}, -\frac{\pi}{2}) \to \mathbb{R}$ definiert. Die Umkehrabbildung wird mit $\arctan y$ bezeichnet; es ist also zunächst einmal eine differenzierbare Funktion. Wir werden gleich sehen, daß sie bei $y = 0$ analytisch ist (sie ist sogar auf ganz $\mathbb{R}$ analytisch — was hier aber nicht weiter verfolgt wird.) Es gilt

$$\tan' x = 1 + \tan^2 x,$$

und daher

$$\arctan' y = \frac{1}{\tan'(\arctan y)} = \frac{1}{1 + y^2}.$$

Entwickeln wir $1/\left(1 + y^2\right)$ in die geometrische Reihe, so erhalten wir

$$\arctan' y = \sum_{k=0}^{\infty} (-1)^k y^{2k} \qquad (|y| < 1).$$

Hieraus erhalten wir durch Anwendung des oben angegebenen Satzes über das gliedweise Differenzieren von Potenzreihen:

$$\arctan y = c + \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1},$$

mit einer Konstanten c. Diese muß aber gleich 0 sein, wie man sieht, indem man in diese Gleichung $y = 0$ setzt (beachte: $\arctan 0 = 0$). Hierbei haben wir noch den folgenden einfachen, aber wichtigen Sachverhalt benutzt:

Satz. *Seien f, g zwei im Intervall (a, b) differenzierbare Funktionen, und es gelte $f' = g'$. Dann ist $f = g + c$ mit einer geeigneten Konstanten c.*

Beweis. Sei $h = f - g$. Sind x_1, x_2 beliebige Punkte in (a, b), so gilt nach dem Mittelwertsatz

$$h(x_1) - h(x_2) = h'(\xi)(x_1 - x_2)$$

mit einem geeigneten Zwischenwert ξ. Aber $h' = f' - g' = 0$, also $h(x_1) = h(x_2)$, und da x_1, x_2 beliebige Punkte sind, folgt, daß h konstant ist, und das wurde behauptet.

Somit erhalten wir den

Satz. *Für alle* $|y| < 1$ *ist*

$$\arctan y = x - \frac{x^3}{3} + \frac{x^5}{5} - + \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}.$$

11 GLEICHMÄSSIGE STETIGKEIT UND GLEICHMÄSSIGE KONVERGENZ

Als Vorbereitung für die Integralrechnung im nächsten Kapitel stellen wir hier zwei Sätze über stetige Funktionen und Folgen von stetigen Funktionen zusammen. Wir definieren zunächst

Gleichmäßige Stetigkeit

Eine auf M erklärte, reellwertige Funktion f heißt *gleichmäßig stetig*, falls gilt:

$$\forall \epsilon > 0 \exists \delta > 0 \forall x, y \in M(|x-y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon).$$

Man beachte den Unterschied zur gewöhnlichen Stetigkeit, der hier zunächst einmal in der anderen Reihenfolge der Quantoren augenfällig wird: eine auf M definierte, reellwertige Funktion f ist stetig, falls gilt:

$$\forall y \in M \forall \epsilon > 0 \exists \delta > 0 \forall xinM(|x-y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon).$$

Jedenfalls ist eine gleichmäßig stetige Funktion auch stetig. Die Umkehrung gilt im Allgemeinen nicht — es sei denn, man stellt gewisse Voraussetzungen an M, etwa, daß M ein abgeschlossenes Intervall ist.

Satz. *Sei f eine auf einem abgeschlossenen Intervall $[a, b]$ definierte, stetige Funktion. Dann ist f gleichmäßig stetig.*

Beweis. Wir nehmen an, daß f nicht gleichmäßig stetig wäre, d.h.

$$\exists \epsilon > 0 \forall \delta > 0 \exists x, y \in [a, b](|x-y| < \delta \land |f(x) - f(y)| \geq \epsilon).$$

Es gibt also ein $\epsilon > 0$, sodaß wir zu jeder ganzen Zahl $n > 0$ x_n und y_n im Intervall $[a, b]$ finden können, sodaß $|x_n - y_n| < \frac{1}{n}$, aber $|f(x_n) - f(y_n)| \geq \epsilon$. Die Folge $\{x_n\}$ ist als Zahlenfolge im Intervall $[a, b]$ beschränkt und hat somit eine konvergente Teilfoge — etwa $\{x_{k_n}\}$. Die Folge $\{y_{k_n}\}$ hat aus demselben Grund ebenfalls eine konvergente Teilfolge — etwa $\{y_{l_n}\}$. Die beiden Folgen $\{x_{l_n}\}$ und $\{y_{l_n}\}$ haben den gleichen Limes — etwa c, denn es ist ja

$$|x_{l_n} - y_{l_n}| < \frac{1}{l_n} \to 0.$$

Nun ist $c \in [a, b]$, da das Intervall ja abgeschlossen ist. Wegen der Stetigkeit von f ist daher

$$f(c) = \lim_n f(x_{l_n}) = \lim_n f(y_{l_n}).$$

Auf der anderen Seite ist aber

$$|f(x_{l_n}) - f(y_{l_n})| \geq \epsilon,$$

und dies ist absurd.

Punktweise Konvergenz

Eine Folge von auf einer Menge M definierten, reellwertigen Funktionen f_n heißt *punktweise konvergent* gegen die Funktion f, falls für jedes $x \in M$ die Folge $f_n(x)$ gegen $f(x)$ konvergiert.

Ein engerer Konvergenzbegriff für Folgen von Funktionen ist

Gleichmäßige Konvergenz

Eine Folge von auf einer Menge M definierten, reellwertigen Funktionen f_n heißt *gleichmäßig konvergent* gegen die Funktion f, falls gilt

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall x \in M \forall n \geq n_0 (|f_n(x) - f(x)| < \epsilon).$$

Offenbar ist eine gleichmäßig konvergente Folge von Funktionen insbesondere punktweise konvergent. Die Umkehrung gilt im Allgemeinen nicht. Auch die im nächsten Satz ausgesprochene Tasache ist für lediglich punktweise konveergente Folgen von Funktionen falsch.

Satz. *Die Folge von auf M definierten Funktionen f_n konvergiere auf M gleichmäßig gegen die Funktion f. Sind die f_n stetig, so ist auch f stetig.*

Beweis. Wir zeigen die Stetigkeit von f in einem $y \in M$. Zu gegebenem $\epsilon > 0$ gibt es wegen der gleichmässigen Konvergenz ein n_0 sodaß für alle

$$|f(x) - f_{n_0}(x)| < \frac{\epsilon}{3} \qquad \text{für alle } x \in M$$

. Wegen der Stetigleit von f_{n_0} gibt es ein $\delta > 0$, sodaß

$$|f_{n_0}(x) - f_{n_0}(y)| < \frac{\epsilon}{3} \qquad \text{für alle } |x - y| < \delta.$$

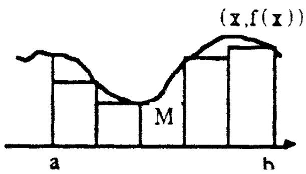
Dann gilt aber für alle $|x - y| < \delta$.

$$|f(x) - f(y)| \leq |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(y)| + |f_{n_0}(y) - f(y)| < \epsilon,$$

und das war zu zeigen.

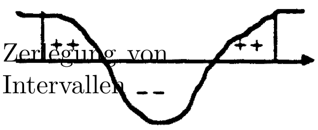
12 INTEGRAL

Ist eine Funktion $f\colon [a,b] \to \mathbb{R}$ gegeben, so daß $f(x) \geq 0$ für jedes $x \in [a,b]$ ist, kann man die Menge $M = \{(x,y) \,|\, x \in [a,b] \wedge 0 \leq y \leq f(x)\}$ betrachten und ihren Flächeninhalt zu bestimmen suchen, indem man M durch Rechtecke approximiert.



Natürlich kommt es auch vor, daß eine Funktion unter die Achse gerät. Dabei werden die "Flächeninhalte" der Stücke unterhalb der Achse negativ gezählt.

Einige Vorbereitungen: Intervalle sind im folgenden alle Mengen eines der folgenden Typen für $\alpha, \beta \in \mathbb{R}$ mit $\alpha \leq \beta$: $[\alpha,\beta], (\alpha,\beta], [\alpha,\beta), (\alpha,\beta)$ und $\emptyset$.



Zerlegung von Intervallen

Sei nun $a < b$. Dann ist eine Zerlegung ζ von $[a,b]$ eine Darstellung von $[a,b]$ als disjunkte Vereinigung von endlich vielen Intervallen I_k, $\zeta = \{I_1, \ldots I_q\}$. Wir verlangen also, daß $I_i \cap I_k = \emptyset$ falls $i \neq k$ und weiterhin, daß $\bigcup_{k=1}^{q} I_k = [a,b]$ ist.

Beispiel: Äquidistante Zerlegung. Man führe Punkte $x_k\colon = a + k\frac{b-a}{q}$ ein, für $k = 0, 1, \ldots q$, und setze dann $I_k\colon = [x_{k-1}, x_k)$ für $k = 1, \ldots q-1; I_q\colon = [x_{q-1}, x_q]$.

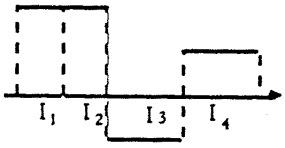
Zweites Beispiel: Man wähle irgendeine endliche Folge $a = x_0, x_1, \ldots x_n = b$ in $[a,b]$, so daß für $k = 1, 2, \ldots n,\ \ x_{k-1} < x_k$ ist. Dann ist

$$\zeta = \{(x_{k-1}, x_k), \{x_e\} \,|\, k = 1, \ldots n; e = 0, 1, \ldots n\}$$

eine Zerlegung von $[a,b]$.

Treppenfunktionen

Eine Treppenfunktion $t\colon [a,b] \to \mathbb{R}$ ist eine Abbildung, zu der es eine Zerlegung $\zeta = \{I_1, \ldots I_q\}$ von $[a,b]$ und Zahlen $c_1, \ldots c_k$ gibt, so daß $t(x) = c_k$ für jedes $x \in I_k$ ist, $k = 1, \ldots q$. Eine Treppenfunktion ist also "intervallweise konstant". Die zur Treppenfunktion gehörende Zerlegung ζ ist aber keineswegs eindeutig bestimmt.



Wir wollen nun stetige Funktionen durch Treppenfunktionen approximieren. Dazu betrachten wir den Vektorraum $V = \{g\colon [a,b] \to \mathbb{R} \,|\, g \text{ beschränkt }\}$. Er wird zum normierten Vektorraum, indem man die Norm

$$\|g\|\colon = \sup_{x \in [a,b]} |g(x)|$$

einführt. Die Abbildung $g \to \|g\|$ hat die üblichen Normeigenschaften:

$$\begin{aligned}
&(a).\ \|g\| \geq 0 \\
&(b).\ \|\lambda \cdot g\| = |\lambda| \cdot \|g\| \\
&(c).\ \|g + h\| \leq \|g\| + \|h\| \\
&(d).\ \|g\| = 0 \Leftrightarrow g = 0
\end{aligned}$$

für jedes g und h aus V, $\lambda \in \mathbb{R}$.

Satz. *Sei $f\colon [a, b] \to \mathbb{R}$ stetig. Dann gibt es zu jedem $\epsilon > 0$ eine Treppenfunktion t, so daß $\|f - t\| < \epsilon$, d.h. $\sup_{x \in [a,b]} |f(x) - t(x)| < \epsilon$.*

Bemerkung: Insbesondere gilt dann, daß

$$|f(x) - t(x)| < \epsilon \quad \text{für} \quad \forall x \in [a, b]. \tag{1}$$

Ist umgekehrt für jedes $x \in [a, b]$, $|f(x) - t(x)| < \epsilon$ so ist ϵ eine obere Schranke von $\{|f(x) - t(x)| \,|\, x \in [a, b]\}$, deren Supremum $\|f - t\| < \epsilon$ ist (eine - auf einer geschlossenen Menge definierte - stetige Funktion nimmt ihr Supremum an). Also ist die Behauptung des Satzes äquivalent zu (1).

Beweis. Wir wissen schon, daß f sogar gleichmäßig stetig ist, d.h.

$$\forall \epsilon > 0 \;\; \exists \delta > 0 \;\text{ so daß }\; \forall x, y \in [a, b] \;\; |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon. \tag{$*$}$$

Sei also $\epsilon > 0$ gegeben. Dann wählen wir eine Zerlegung ζ wie im zweiten Beispiel, so daß $|x_{k-1} - x_k| < \delta$ für $k = 1, 2, ...n$. Nun konstruieren wir wie folgt eine Treppenfunktion t: Setzen wir

$$\begin{aligned}
t(x) &:= f(x_{k-1}) \,\text{für}\, x \in (x_{k-1}, x_k) \\
t(x_k) &:= f(x_k).
\end{aligned}$$

Dann ist

$$|f(x) - t(x)| = \begin{cases} 0, & \text{falls } x = x_k \text{ für ein } k \\ |f(x) - f(x_{k-1})| < \epsilon, & \text{falls } x \in (x_{k-1}, x_k) \text{ für ein } k, \end{cases}$$

denn offenbar ist $|x - x_{k-1}|$ im letzteren Fall $< \delta$. Also ist $\|f - t\| < \epsilon$.

Stetige Funktionen sind gleichmäßig approximierbar durch Treppenfunktionen

Satz. *Sei $f\colon [a, b] \to \mathbb{R}$ stetig. Dann gibt es eine Folge von Treppenfunktionen t_n, die gleichmäßig gegen f konvergiert.*

Bemerkung: f_n konvergiert gleichmäßig gegen $f \Leftrightarrow \lim_{n\to\infty} \|f - f_n\| = 0$

Beweis. Setze $\epsilon := \frac{1}{n}$ und wähle Treppenfunktion t_n, so daß $\|f - t_n\| < \epsilon = \frac{1}{n}$ (dies ist möglich nach dem letzten Satz). Für jedes $x \in [a, b]$ ist also $|f(x) - t_n(x)| < \frac{1}{n}$. Dies bedeutet aber, daß die t_n gleichmäßig gegen

f konvergieren; ist $\epsilon > 0$ gegeben, wähle n_0 so, daß $\frac{1}{n_0} < \epsilon$. Dann ist für $n > n_0$, $|f(x) - t_n(x)| < \frac{1}{n} < \frac{1}{n_0} < \epsilon$, unabhängig von x.

Die Idee ist nun wie folgt: Wollen wir das Integral einer stetigen Funktion f erklären, werden wir diese durch Treppenfunktionen gleichmäßig approximieren, deren Integrale - die gleich definiert werden - das Integral von f approximieren.

Länge eines Intervalls

Dazu definieren wir die Länge $\lambda(I)$ des Intervalls I mit den Endpunkten α und β, $\alpha \leq \beta$ als $\lambda(I) := \beta - \alpha$.

Integral einer Treppenfunktion bzgl. einer Zerlegung

Ist t eine Treppenfunktion bezüglich der Zerlegung $\zeta = \{I_1, ...I_q\}$, so sei das Integral der Treppenfunktion t bezüglich der Zerlegung ζ

$$I_\zeta(t) := \sum_{k=1}^{q} c_k \cdot \lambda(I_k). \qquad \text{(wobei } t = c_k \text{ auf } I_k \text{ ist.)}$$

Da die Zerlegung ζ zur Treppenfunktion t nicht eindeutig bestimmt ist, könnte das Integral $I_\zeta(t)$ auch von ζ abhängen:

Satz. *Das Integral $I_\zeta(t)$ hängt nicht von ζ ab.*

Beweis. Eine Zerlegung ζ des Intervalls $[a, b]$ ist eine Zerlegung in disjunkte Teilintervalle. Diese Zerlegung kann verfeinert werden, indem man die Intervalle I_k der Zerlegung ζ weiter zerlegt: $I_k = \cup_{e=1}^{e_k} \tilde{I}_e^k$, dabei sind die $\tilde{I}_e^k$ ebenfalls disjunkte Intervalle. Dann ist offenbar $c_k \cdot \lambda(I_k) = \sum_{e=1}^{e_k} c_k \cdot \lambda(\tilde{I}_e^k)$. Ist also die Zerlegung ζ' durch Verfeinerung aus der Zerlegung ζ entstanden, so ist $I_\zeta(t) = I_{\zeta'}(t)$; ζ' ist natürlich auch eine Zerlegung zur Treppenfunktion t. Zum Beweis des Satzes ist nun lediglich noch zu überlegen, daß man zu je zwei vorgelegten Zerlegungen ζ_1 und ζ_2, bezüglich welcher t Treppenfunktion ist, eine beiden gemeinsame Verfeinerung finden kann: Ist $\zeta_1 = \left\{I_1^1, ...I_k^1\right\}$ und $\zeta_2 = \left\{I_1^2, ...I_e^2\right\}$,

$$\text{so kann man} \quad \zeta := \left\{I_i^1 \cap I_j^2 \,|\, i = 1, ...k;\ j = 1, ...e\right\} \quad \text{nehmen.}$$

Also ist $I_{\zeta_1}(t) = I_\zeta(t) = I_{\zeta_2}(t)$.

Integral einer Treppenfunktion

Nun können wir das Integral $I(t)$ für eine Treppenfunktion t definieren:

$$I(t) = I_\zeta(t) \quad , \qquad \text{wobei } \zeta \text{ irgendeine Zerlegung zu } t \text{ ist.}$$

Offenbar ist die Menge $V_{TF} = \{t : [a, b] \to \mathbb{R} \,|\, t \text{ Treppenfunktion}\}$ ein Vektorraum und es gilt:

Satz.

$$I(t_1 \pm t_2) = I(t_1) \pm I(t_2)$$
$$I(\lambda \cdot t) = \lambda \cdot I(t),$$

in Worten: I ist eine lineare Abbildung vom V_{TF} nach $\mathbb{R}$.

Beweis. Ist t eine Treppenfunktion bezüglich ζ, so gilt dies auch für $\lambda \cdot t$. Nimmt also t die Werte c_k an, so nimmt $\lambda \cdot t$ die Werte $\lambda \cdot c_k$ an. Also ist

$$I(\lambda \cdot t) = \sum_k \lambda \cdot c_k \cdot \lambda(I_k) = \lambda \cdot \sum_k c_k \cdot \lambda(I_k) = \lambda \cdot I(t).$$

Was die Summe $t_1 \pm t_2$ betrifft, wählen wir - wie im Beweis zum vorherigen Satz - eine gemeinsame Zerlegung ζ, bezüglich welcher t_1 und t_2 beide Treppenfunktionen sind und erhalten

$$I(t_1) \pm I(t_2) = \sum_k c_k \cdot \lambda(I_k) \pm \sum_k d_k \cdot \lambda(I_k) = \sum_k (c_k \pm d_k) \cdot \lambda(I_k) = I(t_1 \pm t_2),$$

wenn t_1 die Werte c_k und t_2 die Werte d_k annimmt.

Eine Abschätzung

Satz. $|I(t)| \leq (b-a) \cdot \|t\|$

Beweis. $|I(t)| \leq \sum_{k=1}^q |c_k| \lambda(I_k) \leq \max_k c_k \cdot \sum_{k=1}^q \lambda(I_k) = \|t\| \cdot (b-a)$, denn es ist $\|t\| := \sup_{x \in [a,b]} |t(x)| = \max_k |c_k|$, da die Funktion t nur die endlich vielen Werte c_k annimmt.

Satz. *Wenn die Folge $(t_n)_{n \in \mathbb{N}}$ von Treppenfunktionen t_n gleichmäßig gegen f konvergiert, dann konvergiert die Folge $(I(t_n))_{n \in \mathbb{N}}$.*

Bemerkung: Den Limes werden wir das Integral von f nennen, wenn f eine stetige Funktion ist. Wir brauchen aber noch eine Unabhängigkeitsaussage, um $\int_a^b f(x)\,dx$ definieren zu können.

Beweis. Entsprechend der Annahme gleichmäßiger Konvergenz gibt es zu jedem $\epsilon > 0$ ein n_0, so daß für $n > n_0$, $\|f - t_n\| < \epsilon$. Folglich ist

$$\text{für} \quad n, m > n_0, \ \ \|t_n - t_m\| \leq \|t_n - f\| + \|f - t_m\| < 2\epsilon.$$

Also ist

$$|I(t_m) - I(t_n)| = |I(t_n - t_m)| \leq \|t_m - t_n\| \cdot (b-a) < 2(b-a)\epsilon.$$

Da der Faktor $2(b-a)$ konstant ist, bildet die Folge der Integrale eine Cauchy-Folge, also konvergiert sie.

Satz. *Konvergieren $(t_n)_{n \in \mathbb{N}}$ und $(\tilde{t}_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f, so ist*

$$\lim_{n \to \infty} I(t_n) = \lim_{n \to \infty} I(\tilde{t}_n).$$

Beweis. Die Folge von Treppenfunktionen $t_1, \tilde{t}_1, t_2, \tilde{t}_2, ...$ konvergiert gleichmäßig gegen f, demnach konvergiert auch die Folge

$$I(t_1), I(\tilde{t}_1), I(t_2), I(\tilde{t}_2), ...$$

deren Teilfolgen denselben Limes haben. Die Behauptung folgt.

Definition des Integrals einer stetigen Funktion

Definieren wir nun das Integral einer stetigen Funktion $f\colon [a,b] \to \mathbb{R}$,

$$\int_a^b f(x)\,dx := \lim_{n\to\infty} I(t_n),$$

wobei $(t_n)_{n\in\mathbb{N}}$ eine Folge von Treppenfunktionen ist, die f gleichmäßig approximiert. Welche solche Folge man auswählt ist für den Wert des Integrals - nach dem letzten Satz - unerheblich.

Bemerkung: Auf die geschilderte Art kann man "$\int_a^b f(x)\,dx$" für jede Funktion $f\colon [a,b] \to \mathbb{R}$, die gleichmäßiger Limes einer Folge von Treppenfunktionen ist, definieren. Beispielsweise kann f dann selbst eine Treppenfunktion sein.

Einteilung eines Intervalls
Feinheit einer Einteilung

Von einer Einteilung E des Intervalls wollen wir immer dann sprechen, wenn wir eine endliche Folge $a = x_0 < x_1 < ... < x_q = b$ vorgelegt haben. Die Feinheit der Einteilung E ist per Definition $\varphi(E) := \max_{i=1,...q} |x_i - x_{i-1}|$.

Zur Einteilung E hatten wir bereits eine Zerlegung ζ angegeben (Zweites Beispiel). Zu dieser Zerlegung und einer vorgelegten Funktion f kann man durch beliebige Wahl von Punkten $\xi_k \in [x_{k-1}, x_k]$ eine Treppenfunktion definieren, und zwar durch

$$\begin{aligned} t(x) &= f(\xi_k) \qquad \text{für} \quad x \in (x_{k-1}, x_k) \\ t(x_k) &= f(x_k). \end{aligned}$$

Riemannsche Summe

Die Riemannsche Summe von f zur Einteilung E und den gewählten $\xi_k \in [x_{k-1}, x_k]$ ist per Definition

$$\sum_{k=1}^{q} f(\xi_k) \cdot (x_k - x_{k-1}) = I(t)$$

Riemannsche Untersumme

Diese Riemannsche Summe hängt also von E und den ξ_k ab. Als Spezialfall kann man die Riemannsche Untersumme von einer stetigen Funktion f zur Einteilung E erklären:

$$\sum_{k=1}^{q} m_k \cdot (x_k - x_{k-1}), \qquad \text{wobei } m_k := \text{ Infimum von } f \text{ auf } [x_{k-1}, x_k],$$

Riemannsche Obersumme

welches wegen der Stetigkeit von f bekanntlich an einer Stelle $\xi_k \in [x_{k-1}, x_k]$ angenommen wird. Analog erhält man mit den Maxima M_k von f in $[x_{k-1}, x_k]$ der Riemannschen Obersumme von f: $\sum_{k=1}^{q} M_k \cdot (x_k - x_{k-1})$.

Approximation durch Riemannsummen

Satz. *Es sei* $(E_n)_{n\in\mathbb{N}}$ *eine Folge von Einteilungen, für die* $\lim_{n\to\infty} \varphi(E_n) = 0$ *ist. Dann ist*

$$\lim_{n\to\infty} \sum_{k=1}^{q_n} f(\xi_k^{(n)}) \cdot (x_k^{(n)} - x_{k-1}^{(n)}) = \int_a^b f(x)\,dx$$

Bemerkung: Jede Folge von Riemannschen Summen zu den Einteilungen E_n konvergiert also gegen das Integral von f, sofern nur die Feinheiten der Einteilungen eine Nullfolge bilden.

Beweis. Führen wir die Treppenfunktionen t_n ein, durch

$$\begin{aligned} t_n(x) &= f(\xi_k^{(n)}) \qquad \text{falls} \quad x \in (x_{k-1}^{(n)}, x_k^{(n)}) \\ t_n(x_k^{(n)}) &= f(x_k^{(n)}). \end{aligned}$$

Wir wollen zeigen, daß die t_n gleichmäßig gegen f konvergieren, denn die $I(t_n)$ sind ja gerade die o.g. Riemannschen Summen. Zu $\epsilon > 0$ wähle man ein $\delta > 0$, so daß $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$, und man wähle ein n_0, so daß für $n > n_0$, $\varphi(E_n) < \delta$ ist. Dann ist $|f(x) - t_n(x)| < \epsilon$ für alle $x \in [a, b]$, sofern $n > n_0$ ist. Denn in $x = x_k^{(n)}$ ist die linke Seite gleich 0; ist aber $x \in (x_{k-1}^{(n)}, x_k^{(n)})$, so ist $t_n(x) = f(\xi_k^{(n)})$ und $\xi_k^{(n)} \in [x_{k-1}^{(n)}, x_k^{(n)}]$. Dann ist

$$|x - \xi_k^{(n)}| \le |x_k^{(n)} - x_{k-1}^{(n)}| \le \varphi(E_n) < \delta,$$

$$\text{wonach} \qquad |f(x) - t_n(x)| = |f(x) - f(\xi_k^{(n)})| < \epsilon \quad \text{ist.}$$

Beispiele

Es sei $f\colon [0,1] \to \mathbb{R},\ x \to x^2$. Wir wählen äquidistante Einteilungen

$$E_n\colon x_0^{(n)} = 0, x_1^{(n)} = \frac{1}{n}, \ldots, x_k^{(n)} = \frac{k}{n}, \ldots, x_n^{(n)} = 1.$$

Dann ist $\varphi(E_n) = \frac{1}{n}$, d.h. $\lim_{n\to\infty} \varphi(E_n) = 0$. Die Obersumme zu E_n ist $\sum_{k=1}^{n} \left(\frac{k}{n}\right)^2 \cdot \frac{1}{n} = (*)_n$, denn die Maxima von f auf den Intervallen $[x_{k-1}^{(n)}, x_k^{(n)}]$ werden am rechten Rand angenommen, da f monoton wächst.

$$(*)_n = \frac{1}{n} \cdot \frac{n(n+1)(2n+1)}{6} = \frac{(1+\frac{1}{n})(2+\frac{1}{n})}{6}.$$

Demnach gilt $\lim_{n\to\infty} (*)_n = \frac{1}{3} = \int_0^1 x^2\,dx$.

Zweites Beispiel: Es sei $f(x) := \frac{1}{x}$ auf dem Intervall $[a, b]$, wobei $0 < a < b$ ist. Immerhin ist $\frac{1}{x}\colon [a, b] \to \mathbb{R}$ gleichmäßig stetig, weil auf einem abgeschlossenen Intervall definiert und stetig. ($\frac{1}{x}\colon (0, \infty) \to \mathbb{R}$ ist nicht gleichmäßig stetig!)

Als n-te Einteilung wählen wir: $\quad a, ac_n, ac_n^2, \ldots ac_n^n, \quad$ wobei $\quad c_n := \sqrt[n]{\dfrac{b}{a}}.$

Die Einteilungen sind also geometrische Folgen, d.h. man hat konstante Quotienten aufeinanderfolgender Glieder. Die Riemannsche Obersumme ist

$$\sum_{k=1}^{n} \frac{1}{ac_n^{k-1}} \cdot (ac_n^k - ac_n^{k-1}) = n \cdot (\sqrt[n]{\frac{b}{a}} - 1),$$

da f monoton fällt. Was wissen wir über die Feinheit $\varphi(E_n)$? Die Intervallängen sind $ac_n^k - ac_n^{k-1} = ac_n^k \cdot (1 - \frac{1}{c_n})$ Da $c_n > 1$ ist, ist

$$\varphi(E_n) = ac_n^n - ac_n^{n-1} = b - \frac{b}{c_n} = b \cdot (1 - \frac{1}{c_n}),$$

es gilt also $\lim_{n\to\infty} \varphi(E_n) = 0$. Nach unseren Satz ist demnach

$$\int_a^b \frac{1}{x}\, dx = \lim_{n\to\infty} \frac{\left(\frac{b}{a}\right)^{\frac{1}{n}} - 1}{\frac{1}{n}} = \lim_{x\to 0} \frac{\left(\frac{b}{a}\right)^{x} - 1}{x} =$$

$$= \lim_{x\to 0} \frac{\ln(\frac{b}{a}) \cdot \left(\frac{b}{a}\right)^{x}}{1} \quad \text{(nach l'Hôpital)} \quad = \ln b - \ln a.$$

Hauptsatz der Differential- und Integralrechnung

Beweisen wir nun den Hauptsatz der Differential- und Integralrechnung, der etwas über die Verbindung von Integral und Differential aussagt:

Satz. *Es sei $F\colon [a,b] \to \mathbb{R}$ stetig, $F\colon (a,b) \to \mathbb{R}$ differenzierbar. Weiter sei $f\colon [a,b] \to \mathbb{R}$ stetig und $f = F'$ in (a,b). Dann ist $\int_a^b f(x)\,dx = F(b) - F(a)$*

Stammfunktion

Bemerkung: Man sagt auch, daß F Stammfunktion von f sei, wenn $F' = f$ ist. Ist F eine Stammfunktion von f, dann ist auch $(F + \text{Konstante})' = f$, demnach hat f mehrere Stammfunktionen (wenn - was noch gezeigt wird - f überhaupt eine Stammfunktion besitzt).

Beweis. Es sei eine Einteilung $a = x_0, x_1, \ldots, x_q = b$ gegeben. Dann ist nach dem Mittelwertsatz der Differenzialrechnung

$$F(x_k) - F(x_{k-1}) = f(\xi_k) \cdot (x_k - x_{k-1}), \qquad (*)$$

für einen Punkt ξ_k zwischen x_{k-1} und x_k. Also ist die zugehörige Riemannsche Summe $\sum_{k=1}^{q} f(\xi_k) \cdot (x_k - x_{k-1}) = F(b) - F(a)$, bei Wahl der ξ_k gemäß $(*)$. Nimmt man irgendeine Folge von Einteilungen E_n, für die die Feinheit $\varphi(E_n)$ gegen 0 konvergiert, und wählt man Riemannsche Summen für jedes E_n auf die geschilderte Weise, so gilt nach dem vorherigen Satz

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} [F(b) - F(a)] = F(b) - F(a).$$

Beispiele

Zu $f(x) = x^2$ ist $F(x) = \frac{x^3}{3}$ Stammfunktion , demnach ist $\int_a^b x^2\,dx = \frac{b^3}{3} - \frac{a^3}{3}$, ebenso folgt aus dem Hauptsatz $\int_a^b \frac{1}{x}\,dx = \ln b - \ln a$.

Flächeninhalt des Einheitskreises

Betrachten wir $f(x) = \sqrt{1-x^2}$ für $x \in [-1,1]$. Dort ist f stetig, und in $(-1,1)$ sogar differenzierbar. Am Rand ist f nicht differenzierbar, selbst wenn man die naheliegenden Limites nur einseitig bildete, denn die Kreislinie hat am Rand senkrechte Tangenten. Können wir eine Stammfunktion $F\colon [-1,1] \to \mathbb{R}$ (F stetig, in $(-1,1)$ differenzierbar und $F' = f$) zu $\sqrt{1-x^2}$ finden?

Versuchen wir es mit:

$$F(x) = \frac{1}{2} \cdot (\arcsin(x) \, + \, x\sqrt{1-x^2}).$$

Immerhin ist $F(x)$ stetig, und die Ableitung kennen wir auch,

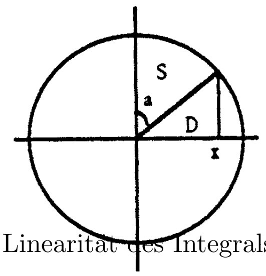
$$\text{denn} \qquad \arcsin'(x) = \frac{1}{\sqrt{1-x^2}}, \quad \text{d.h.}$$

$$F'(x) = \frac{1}{2} \cdot \frac{1}{\sqrt{1-x^2}} + \frac{1}{2} \cdot x \cdot \frac{-x}{\sqrt{1-x^2}} \frac{1}{2} \cdot \sqrt{1-x^2} = \sqrt{1-x^2}.$$

Man kann aber auch durch Nachdenken daraufkommen: $\frac{1}{2}x\sqrt{1-x^2}$ ist die Fläche des Dreiecks D. Kann auch $\arcsin(x)$ als Flächeninhalt gedeutet werden? Wir wissen, daß die Fläche des Sektors S gleich der halben Bogenlänge ist (d.h. eigentlich dürften wir es nicht wissen), also ist die Fläche des Sektors S gleich $\frac{a}{2}$. Da aber $\sin(a) = x$ ist, ist die Sektorfläche $\frac{1}{2}\arcsin(x)$, die Fläche von S und D zusammen deshalb gleich $F(x) = \frac{1}{2} \cdot (\arcsin(x) + x\sqrt{1-x^2})$.

Wir haben

$$F(1) - F(-1) = \int_{-1}^{1} \sqrt{1-x^2}\,dx = \frac{1}{2}\arcsin(1) - \frac{1}{2}\arcsin(-1) = \frac{\pi}{2}.$$



Der Flächeninhalt des ganzen Kreises ist demnach π.

Weitere allgemeine Eigenschaften des Integrals:

Linearität des Integrals

Satz.

$$\int_a^b (f_1 + f_2)(x)\,dx = \int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx$$

$$\int_a^b \alpha \cdot f(x)\,dx = \alpha \cdot \int_a^b f(x)\,dx$$

für stetige Funktionen f, f_1, f_2 und $\alpha \in \mathbb{R}$.

Bemerkung: Dieser Satz besagt, daß das Integral eine lineare Abbildung vom Vektoraum der stetigen Funktionen auf $[a,b]$ in den reellen Zahlen ist.

Beweis. Wir führen diesen Satz auf den entsprechenden Satz über Integrale von Treppenfunktionen zurück. Ist $f_1 = \lim_{n\to\infty} \hat{t}_n$ (gleichmäßig) und $f_2 = \lim_{n\to\infty} \tilde{t}_n$ (gleichmäßig), dann ist $f_1 + f_2 = \lim_{n\to\infty}(\hat{t}_n + \tilde{t}_n)$ (gleichmäßig), denn sind $(\|f_1 - \hat{t}_n\|)_{n\in\mathbb{N}}$ und $(\|f_2 - \tilde{t}_n\|)_{n\in\mathbb{N}}$ Nullfolgen, so gilt dies auch für $(\|f_1 + f_2 - \hat{t}_n - \tilde{t}_n\|)_{n\in\mathbb{N}}$,

$$\text{da} \quad \|f_1 + f_2 - \hat{t}_n - \tilde{t}_n\| \leq \|f_1 - \hat{t}_n\| + \|f_2 - \tilde{t}_n\| \quad \text{ist.}$$

Nach Definition ist demnach

$$\int_a^b f_1(x) + f_2(x)\,dx = \lim_{n\to\infty} I(\hat{t}_n + \tilde{t}_n) =$$
$$= \lim_{n\to\infty} I(\hat{t}_n) + \lim_{n\to\infty} I(\tilde{t}_n) = \int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx.$$

Genauso zeigt man die zweite Behauptung, wobei man verwendet, daß $\|\alpha t_n - \alpha f\| = |\alpha| \cdot \|t_n - f\|$ ist.

$|\int f| \leq \|f\| \cdot (b-a)$

Satz. *Es sei $f\colon [a,b] \to \mathbb{R}$ stetig. Dann ist*

$$|\int_a^b f(x)\,dx| \leq \|f\| \cdot (b-a)$$

Beweis. Approximiere f gleichmäßig durch Treppenfunktionen t_n, so daß die Funktionswerte von jedem t_n auch Funktionswerte von f sind; dann ist $\|t_n\| \leq \|f\|$ (vergleich die Definition der Norm), folglich gilt auch für jedes n: $I(t_n) \leq \|t_n\| \cdot (b-a) \leq \|f\| \cdot (b-a)$. Deshalb ist $\lim_{n\to\infty} I(t_n) = \int_a^b f(x)\,dx \leq \|f\| \cdot (b-a)$.

$f \geq 0 \Rightarrow \int f \geq 0$

Satz. *Sei f wie im letzten Satz und $f(x) \geq 0$ für $x \in [a,b]$. Dann ist $\int_a^b f(x)\,dx \geq 0$*

Beweis. Approximiere $\int_a^b f(x)\,dx$ durch Riemannsche Summen. Diese sind alle ≥ 0.

Monotonie des Integrals

Folgerung: Sind f_1 und f_2 stetig auf $[a,b]$ und ist $f_1 \leq f_2$ auf $[a,b]$, so ist $\int_a^b f_1(x)\,dx \leq \int_a^b f_2(x)\,dx$

Beweis. $f_2 - f_1$ ist stetig und ≥ 0 auf $[a,b]$, also ist

$$\int_a^b f_2(x)\,dx - \int_a^b f_1(x)\,dx = \int_a^b (f_2(x) - f_1(x))\,dx \geq 0.$$

$|\int f| \leq \int |f|$

Satz. *Es sei weiterhin f wie im vorletzten Satz. Dann gilt $|\int_a^b f(x)\,dx| \leq \int_a^b |f(x)|\,dx$.*

Bemerkung: $x \to |f(x)|$ ist als Kompositum stetiger Abbildungen stetig, demnach ist das Integral rechts wohlerklärt.

Beweis. Man kann den Satz als eine Art Dreiecksungleichung verstehen: Approximiert man das linke Integral mit Riemannsummen und benutzt für die Beträge die gewöhnliche Dreiecksungleichung, so erhält man eine Approximation mit Riemannsummen des rechten Integrals.

Ein zweiter Beweis: Es ist $f(x) \leq |f(x)|$ und $-f(x) \leq |f(x)|$. Wegen der Monotonie des Integrals ist $\pm \int_a^b f(x)\,dx = \int_a^b \pm f(x)\,dx \leq \int_a^b |f(x)|\,dx$, also ist $|\int_a^b f(x)\,dx| \leq \int_a^b |f(x)|\,dx$.

Satz. *Es sei f stetig auf $[a,b]$, und es sei m das Minimum, M das Maximum von f auf $[a,b]$. Dann ist*

$$(b-a)\cdot m \leq \int_a^b f(x)\,dx \leq (b-a)\cdot M.$$

Beweis. Betrachte die konstanten Funktionen $x \to m$ und $x \to M$. Ihre Integrale sind $\int_a^b m\,dx = m\cdot(b-a)$ und $\int_a^b M\,dx = M\cdot(b-a)$. Die Monotonie des Integrals impliziert die Behauptung.

Mittelwert

Der Mittelwert von f in $[a,b]$ ist per Definition

$$\frac{1}{b-a}\cdot\int_a^b f(x)\,dx.$$

Dieser Begriff ist eine Verallgemeinerung des

$$\text{„gewichteten Mittels:“} \quad \frac{w_1\cdot y_1 + w_2\cdot y_2 + \ldots + w_n\cdot y_n}{w_1 + w_2 + \ldots + w_n}$$

von n Zahlen y_i mit Gewichtsfaktoren w_i. Wählt man nämlich eine genügend feine Einteilung von $[a,b]$, so ist der Mittelwert von f “ungefähr” $\frac{1}{b-a}\cdot\sum_{k=1}^q f(\xi_k)\cdot(x_k - x_{k-1})$, da die Riemannsche Summe “ungefähr” das Integral und $b-a$ genau $\sum_{k=1}^q (x_k - x_{k-1})$ ist.

Umformuliert heißt der letzte Satz:

Satz. $m \leq \frac{1}{b-a}\cdot\int_a^b f(x)\,dx =$ *Mittelwert von f in $[a,b]$* $\leq M$.

Der Mittelwert von f wird aber nach dem Zwischenwertsatz angenommen, da f die Werte m und M annimmt. Demnach ergibt sich die

Mittelwertsatz der Integralrechnung

Folgerung: Ist $f\colon [a,b] \to \mathbb{R}$ stetig, so gibt es ein $\xi \in [a,b]$, so daß

$$f(\xi) = \frac{1}{b-a}\cdot\int_a^b f(x)\,dx$$

Bemerkung: Ist insbesondere $\int_a^b f(x)\,dx = 0$, so hat f eine Nullstelle.

Untersummen und Obersummen

Einige Bemerkungen über Untersummen und Obersummen: Zu jeder Einteilung $a = x_0 < x_1 < ... < x_q = b$ haben wir die Obersumme und Untersumme von f erklärt: $\sum_{k=1}^{q} M_k \cdot (x_k - x_{k-1})$ und $\sum_{k=1}^{q} m_k \cdot (x_k - x_{k-1})$. (Im Gegensatz zu den Riemannschen Summen sind Ober- und Untersumme zu einer Einteilung wohl erklärt!)

Nun ist die Untersumme kleiner oder gleich dem Integral. Dies sieht man wie folgt: Geht man von der gegebenen Einteilung zu einer neuen über, indem man einen Punkt u etwa zwischen x_k und x_{k+1} einfügt, so wird die Untersumme größer, denn das Minimum m_{k+1} von f auf $[x_k, x_{k+1}]$ ist nicht größer als das Minimum $m_{k+1}^{(1)}$ von f auf $[x_k, u]$ und auch nicht größer als das Minimum $m_{k+1}^{(2)}$ von f auf $[u, x_{k+1}]$, also ist der wesentliche Summand $m_{k+1} \cdot (x_{k+1} - x_k) \leq m_{k+1}^{(1)} \cdot (u - x_k) + m_{k+1}^{(2)} \cdot (x_{k+1} - u)$. Nimmt man eine Folge von Einteilungen E_n mit $\lim_{n\to\infty} \varphi(E_n) = 0$, so daß E_n eine Verfeinerung von E_{n-1} ist für jedes n, so bilden die Untesummen eine monoton steigende, die Obersummen eine monoton fallende Folge, und beide Folgen konvergieren gegen das Integral von f. Es folgt:

$$\text{Untersumme} \leq \int_a^b f(x)\,dx \leq \text{Obersumme}.$$

Zusammenhang mit dem Riemannschen Integral

Insbesondere ist

$$\begin{aligned}\int_a^b f(x)\,dx &= \text{obere Grenze aller Untersummen}\\ &\quad\ \text{bezüglich aller Einteilungen}\\ &= \text{untere Grenze aller Obersummen}\\ &\quad\ \text{bezüglich aller Einteilungen}\\ &=: \text{Riemannsche Integral von} f.\end{aligned}$$

Man beachte insbesondere, daß $m \cdot (b - a)$ und $M \cdot (b - a)$ die primitivsten denkbaren Unter- und Obersummen sind.

Wir wollen gleich zeigen, daß jede stetige Funktion eine Stammfunktion besitzt (vgl. Bemerkung zum Hauptsatz der Differential- und Integralrechnung). Dazu einige technische Vorbereitungen:

Es sei $f\colon [a, b] \to \mathbb{R}$ im folgenden stetig und $c, d \in [a, b]$. Man setzt

$$\int_c^d f(x)\,dx := \begin{cases} \int_c^d f|_{[c,d]}(x)\,dx, & \text{falls } c < d. \text{ Man schränke also nur}\\ & f \text{ auf } [c,d] \text{ ein und verwende die}\\ & \text{alten Definitionen.}\\ 0, & \text{falls } c = d\\ -\int_d^c f(x)\,dx, & \text{falls } d < c.\end{cases}$$

Zur Plausibilität dieser Definition: Wir hatten beim Rechteck die lotrechte Seite negativ gezählt, falls sich diese unterhalb der Achse befindet, also die "Orientierung" umgedreht worden ist. Dreht man nun die Orientierung der wagerechten Seite um, indem $d < c$ ist, so ist es naheliegend diese Seite auch negativ zu zählen.

$\int_c^e = \int_c^d + \int_d^e$

Satz. *Es sei* $f\colon [a,b] \to \mathbb{R}$ *stetig,* $c, d, e \in [a,b]$*. Dann ist*

$$\int_c^e f(x)\,dx = \int_c^d f(x)\,dx + \int_d^e f(x)\,dx.$$

Beweis. 1. Fall $c < d < e$: Wähle nur Einteilungen E_n des Intervalls $[c,e]$, die d als Teilungspunkt enthalten. Automatisch erhalten wir Einteilungen von $[c,d]$ und $[d,e]$. Außerdem wähle man die E_n so, daß $\lim_{n\to\infty} \varphi(E_n) = 0$. Bezüglich dieser Einteilung ist z.B. (n-te Untersumme von $f|_{[c,e]}$) = (n-te Untersumme von $f|_{[c,d]}$) + (n-te Untersumme von $f|_{[d,e]}$), woraus die Behauptung folgt.

2. Zwei der c, d, e sind gleich, z.B. $c = e$: Dann ist $\int_c^e f(x)\,dx = 0$ und die Behauptung folgt aus den Definitionen.

3. Etwa $d < c < e$: Dann ist

$$\int_d^c f(x)\,dx + \int_c^e f(x)\,dx = \int_d^e f(x)\,dx$$

nach Fall 1), also

$$\int_c^e f(x)\,dx = \int_d^e f(x)\,dx - \int_d^c f(x)\,dx,$$

aber per Definition ist $-\int_d^c f(x)\,dx = \int_c^d f(x)\,dx$. Analog zeigt man die verbleibenden Fälle.

Konstruktion einer Stammfunktion

Nun zur Konstruktion einer Stammfunktion zu f. Wir betrachten die Funktion

$$F(x) = \int_c^x f(t)\,dt \qquad \text{mit} \quad x \in [a,b].$$

(t nennt man auch gebundene Variable, diese kann beliebig umgetauft werden. Vgl. $\sum_{\nu=0}^n a_\nu = \sum_{\mu=0}^n a_\mu$).

Satz. *F ist differenzierbar in (a,b) und $F'(x) = f(x)$.*

Beweis.

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \cdot \left(\int_c^{x+h} f(t)\,dt - \int_c^x f(t)\,dt \right) = \frac{1}{h} \cdot \left(\int_x^{x+h} f(t)\,dt \right)$$

falls $h \neq 0$, nach dem letzten Satz. Sei nun etwa $h > 0$, dann ist $\frac{1}{h} \cdot \left(\int_x^{x+h} f(t)\,dt\right) = f(\xi_h)$ in einem Punkt $\xi_h \in [x, x+h]$ nach dem Mittelwertsatz der Integralrechnung.

Im Fall $h < 0$ ist $$\frac{1}{h} \cdot \left(\int_x^{x+h} f(t)\,dt\right) = \frac{1}{-h} \cdot \left(\int_{x+h}^{x} f(t)\,dt\right) = f(\xi_h)$$

in einem Punkt $\xi_h \in [x+h, x]$. (Man bemerke, daß hier $-h$ die Intervallänge ist.) Durchläuft also h eine Nullfolge, so durchlaufen die zugehörigen ξ_h eine gegen x konvergente Folge; wegen der Stetigkeit von f durchlaufen die $f(\xi_h)$ dann eine gegen $f(x)$ konvergente Folge. Wir haben demnach $\lim_{h\to 0} \frac{1}{h} \cdot \left(\int_x^{x+h} f(t)\,dt\right) = f(x)$.

Differenzierbar

Dieses Resultat läßt sich noch verbessern; doch zuvor einige Definitionen: g ist rechtsseitig differenzierbar an der Stelle a $:\Longleftrightarrow$

$$\lim_{\substack{h\to 0\\ h>0}} \frac{g(a+h) - g(a)}{h} \qquad \text{existiert.}$$

g ist linksseitig differenzierbar an der Stelle b $:\Longleftrightarrow$

$$\lim_{\substack{h\to 0\\ h<0}} \frac{g(b+h) - g(b)}{h} \qquad \text{existiert.}$$

Man bildet nur einseitige Limits, damit die Quotienten überhaupt definiert sind. $g\colon [a,b] \to \mathbb{R}$ heißt differenzierbar $:\Longleftrightarrow$ g ist differenzierbar in (a,b), rechtsseitig differenzierbar in a und linksseitig differenzierbar in b. Ist obendrein die Ableitung stetig, so heißt g stetig differenzierbar in $[a,b]$.

Dem Beweis des letzten Satzes entnimmt man dann

Hauptsatz: Die Existenz einer Stammfunktion einer stetigen Funktion

Satz. *F ist stetig differenzierbar in $[a,b]$ und $F'(x) = f(x)$ in $[a,b]$.*

Bemerkung: Demnach haben wir jetzt eingesehen, daß jede stetige Funktion eine Stammfunktion besitzt. Außerdem ist $\int_a^b f(x)\,dx = F(b) - F(a)$, da $\int_a^b f(x)\,dx = \int_c^b f(x)\,dx - \int_c^a f(x)\,dx$ ist.

$\int f(x)\,dx$ ist eine Schreibweise, die man gelegentlich in der Literatur findet. $\int f(x)\,dx$ bezeichnet dort *eine* Stammfunktion; $F(x) = \int f(x)\,dx$ besagt, daß F Stammfunktion zu f ist.

Die Methode, Integrale durch Stammfunktionen der Integranden zu berechnen, erlaubt uns Sätze der Differenzialrechnung in solchen der Integralrechnung umzuformulieren.

Wir wissen, daß für stetig differenzierbare Funktionen u und v auf $[a, b]$ gilt: $(u \cdot v)' = u' \cdot v + u \cdot v'$. Es ist also $u \cdot v$ Stammfunktion zu $u' \cdot v + u \cdot v'$, weshalb

Partielle Integration

$$\int_a^b u'(x) \cdot v(x)\, dx = u(b) \cdot v(b) - u(a) \cdot v(a) - \int_a^b u(x) \cdot v'(x)\, dx.$$

Beispiel

Schreiben wir $[F(x)]_a^b := F(b) - F(a)$, gilt

$$\int_0^{\frac{\pi}{2}} \sin^2(x)\, dx = [-\cos(x) \cdot \sin(x)]_0^{\frac{\pi}{2}} + \int_0^{\frac{\pi}{2}} \cos^2(x)\, dx,$$

(man setze $v = \sin(x);\ \ u = -\cos(x)$). Aber $\sin(0) = \cos(\frac{\pi}{2}) = 0$ und

$$\int_0^{\frac{\pi}{2}} \left(\sin^2(x) + \cos^2(x)\right) dx = \int_0^{\frac{\pi}{2}} dx = \frac{\pi}{2},$$

wonach

$$\int_0^{\frac{\pi}{2}} \sin^2(x)\, dx = \int_0^{\frac{\pi}{2}} \cos^2(x)\, dx = \frac{\pi}{4}.$$

Eine andere Anwendung der Regel der partiellen Integration: Dazu sei I ein offenes Intervall und f auf I n-mal stetig differenzierbar. Ferner seien $a, b \in I, n \geq 1$. Dann ist offenbar

$$f(b) = f(a) + \int_a^b f'(x)\, dx$$

Dies sieht schon wie die Taylorsche Formel aus, a spielt dabei die Rolle des Entwicklungspunktes, wenn $f(b)$ zu bestimmen ist. Entwickeln wir das Integral (Restglied) weiter, bekommen wir:

$$\int_a^b f'(x) \cdot (-1)\, dx = [f'(x) \cdot (b - x)]_a^b - \int_a^b f''(x) \cdot (b - x)\, dx$$

(falls wir $u = f'(x),\ v' = -1$ und $v = b - x$ setzen)

$$= -f'(a) \cdot (b - a) - \int_a^b f''(x)(b - x)\, dx.$$

Demnach ist

$$f(b) = f(a) + f'(a) \cdot (b - a) + \int_a^b f''(x) \cdot (b - x)\, dx.$$

Taylorsche Formel

Satz. *Sei f n-mal stetig differenzierbar auf dem Intervall I; $a, b \in I$. Dann ist*

Restglied in $\int$-form

$$f(b) = \sum_{\nu=0}^{n-1} f^{\nu}(a) \cdot \frac{(b-a)^{\nu}}{\nu!} + \int_a^b f^{(n)}(x) \cdot \frac{(b-x)^{n-1}}{(n-1)!}\, dx.$$

Beweis. Uns fehlt nur noch der Induktionsschritt. Dazu bestimmen wir

$$\begin{aligned}
&\int_a^b f^{(n)}(x) \cdot \frac{(b-x)^{n-1}}{(n-1)!}\, dx = \\
&(\text{setze } u = f^{(n)}(x),\ v' = \frac{(b-x)^{n-1}}{(n-1)!};\ v = -\frac{(b-x)^n}{n!}) \\
&= -[f^{(n)}(x) \cdot \frac{(b-x)^n}{n!}]_a^b + \int_a^b f^{(n+1)}(x) \cdot \frac{(b-x)^n}{n!}\, dx \\
&= f^{(n)}(a) \cdot \frac{(b-a)^n}{n!} + \int_a^b f^{(n+1)}(x) \cdot \frac{(b-x)^n}{n!}\, dx.
\end{aligned}$$

Die Induktionsannahme - Behauptung ist für n richtig - impliziert demnach die Richtigkeit der Behauptung für $n+1$.

Was auf differentialisch "Produktregel" heißt, heißt auf integralisch "Regel von der partiellen Integration". So bedeutet das differenzialische Wort "Kettenregel" im Integralischen

Transformations- und Substitutionsregel

Satz. *Es sei* $f\colon [a,b] \to \mathbb{R}$ *stetig und* $\varphi\colon [\alpha,\beta] \to [a,b] \subset \mathbb{R}$ *stetig differenzierbar. Dann ist*

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\, dx = \int_\alpha^\beta f(\varphi(u)) \cdot \varphi'(u)\, du$$

Bemerkung: Wie merkt man sich das? Setzen wir $x = \varphi(u)$, dann "ist" $dx = \varphi'(u) \cdot du$ (Kettenregel). Dabei geht der linke Integrand in den rechten Integrand über. Kennt man dann die rechten Grenzen, läuft also u von α nach β, so läuft $x = \varphi(u)$ von $\varphi(\alpha)$ nach $\varphi(\beta)$.

Beweis. Es sei $\nu \in [\alpha, \beta]$. Differenzieren wir

$$\int_\alpha^\nu f(\varphi(u)) \cdot \varphi'(u)\, du \quad \text{und} \quad \int_{\varphi(\alpha)}^{\varphi(\nu)} f(x)\, dx,$$

- die beide Funktionen von ν sind - nach ν, bekommen wir für den ersten Ausdruck $f(\varphi(\nu)) \cdot \varphi'(\nu)$. Für den zweiten Ausdruck schreiben wir $\phi(\nu) = \int_{\varphi(\alpha)}^{\varphi(\nu)} f(x)\, dx$. Offenbar ist ϕ ein Kompositum: $\phi(\nu) = \psi(\varphi(\nu))$, wobei $\psi(y) = \int_{\varphi(\alpha)}^{y} f(x)\, dx$.

Nach der Kettenregel ist $\quad \phi'(\nu) = \psi'(\varphi(\nu)) \cdot \varphi'(\nu) = f(\varphi(\nu)) \cdot \varphi'(\nu)$.

Da nun beide Integrale (als Funktion von ν) dieselbe Ableitung und denselben Wert für $\nu = \alpha$, nämlich 0, haben, sind sie gleich.

Fläche des Kreises

Eine Anwendung: In $\int_{-r}^{r} \sqrt{r^2 - x^2}\, dx$ setzen wir $x = r \cdot \sin(u)$ $(= \varphi(u))$, d.h. $dx = r \cdot \cos(u)\, du$, weshalb

$$\int_{-r}^{r} \sqrt{r^2 - x^2}\, dx = \int_{-\pi/2}^{\pi/2} \sqrt{r^2 - r^2 \cdot \sin^2(u)} \cdot r \cdot \cos(u)\, du$$

$$= r^2 \cdot \int_{-\pi/2}^{\pi/2} \cos^2(u)\, du = \frac{\pi \cdot r^2}{2} = \text{Fläche des Halbkreises vom Radius } r,$$

da $\int_{-\pi/2}^{0} \cos^2(u)\, du = \int_{0}^{-\pi/2} \cos^2(u)\, du$ ist, wie man durch Betrachtung der Funktion $\cos^2(u)$ erkennt, oder durch die Transformation $v = -u$ herleiten kann.

Versuchen wir nun $(*) = \int_0^x \sqrt{r^2 - \xi^2}\, d\xi$ zu bestimmen, d.h. eine Stammfunktion zu $\sqrt{r^2 - x^2}$ für $|x| \leq r$ zu finden. Wir setzen $\xi = r \cdot \sin(u)$, wonach $d\xi = r \cdot \cos(u)\, du$ ist, d.h.

$$(*) = \int_0^v \sqrt{r^2 - r^2 \cdot \sin^2(u)} \cdot r \cdot \cos(u)\, du = r^2 \cdot \int_0^v \cos^2(u)\, du,$$

wo $x = r \cdot \sin(v)$, oder umgeformt $v = \arcsin \frac{x}{r}$ ist. Es gilt $x \in [-r, r]$, demnach muß $u \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, und

$$\begin{aligned}
(*) &= r^2 \cdot \int_0^v \cos(u) \cdot \cos(u)\, du = r^2 \cdot [\sin(u)\cos(u)]_0^v + r^2 \cdot \int_0^v \sin^2(u)\, du \\
&= r^2 \cdot \sin(v)\cos(v) + r^2 \cdot \int_0^v (1 - \cos^2(u))\, du \\
&= r^2 \cdot \sin(v)\cos(v) + r^2 \cdot v - r^2 \cdot \int_0^v \cos^2(u)\, du.
\end{aligned}$$

Wir bekommen deshalb $(*) = \frac{1}{2} \cdot r^2 \cdot (\sin(v)\cos(v) + v)$, oder weil $v = \arcsin \frac{x}{r}$;

$$\int_0^x \sqrt{r^2 - \xi^2}\, d\xi = \frac{1}{2} \cdot x \cdot \sqrt{r^2 - x^2} + \frac{1}{2} \cdot r^2 \cdot \arcsin \frac{x}{r}.$$

13 ERGÄNZUNGEN ZUM INTEGRAL

In diesem abschließenden Kapitel stellen wir noch einige Tatsachen zusammen, die an sich nicht weiter zusammmenhängen, die aber bei der Behandlung des Integralbegriffs nicht unerwähnt bleiben sollten.

Wir beginnen mit einem Satz über die "Vertauschbarkeit von Integral und Limes".

Vertauschbarkeit von "$\int$ und lim"

Satz. *Gegeben sei eine Folge von Funktionen f_n, die auf dem Intervall $[a, b]$ stetig sind, und die auf $[a, b]$ gleichmäßig gegen f konvergieren. Dann ist f stetig, und es gilt*

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx$$

Beweis. Die Stetigkei von f wurde schon in Kapitel 11 nachgewiesen, sodaß wir lediglich noch die beuptete Identität zu beweisen haben. Sei dazu ein $\epsilon > 0$ gegeben. Dann gibt es ein n_0, sodaß für alle $n > n_0$ und alle $x \in [a, b]$ die Ungleichung

$$|f(x) - f_n(x)| < \epsilon$$

besteht. Ist daher $n > n_0$, so haben wir

$$\begin{aligned}\left|\int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx\right| &= \left|\int_a^b (f(x) - f_n(x))\,dx\right| \\ &\le \int_a^b |f(x) - f_n(x)|\;dx \le \int_a^b \epsilon\,dx \\ &= \epsilon\cdot(b-a).\end{aligned}$$

Da dies für jedes $\epsilon > 0$ gilt, ist die Behauptung bewiesen.

Die Voraussetzung, daß die Folge der f_n gleichmäßig konvergiert, ist nicht notwendig, aber auch nicht überflüssig. Sie ist nicht notwendig: vermöge

$$f_n(x) = \begin{cases} 2nx & \text{für } 0 \le x \le \frac{1}{2n}, \\ 2n\left(\frac{1}{n} - x\right) & \text{für } \frac{1}{2n} < x \le 1 \end{cases}$$

wird eine Folge von auf dem Intervall $[0, 1]$ stetigen Funktionen f_n erklärt, die punktweise gegen die identisch verschwindende Funktion konvergiert.

Die Konvergenz findet sicherlich nicht gleichmäßig statt, denn es ist ja stets $f_n(\frac{1}{2n}) = 1$, wonach es kein n_0 geben kann, sodaß etwa $|f_n(x)| < .01$ für alle $x \in [0,1]$ und alle $n \geq n_0$ gilt. Dennoch hat man hier

$$\int_0^1 f_n(x)\,dx = \frac{1}{2n},$$

d.h die Integrale der f_n konvergieren gegen 0, was ja auch der Wert des Integrals der Grenzfunktion ist.

Die Voraussetzung ist nicht überflüssig: hierzu variiert man das eben betrachtete Beispiel; wir setzen jetzt

$$g_n(x) = \begin{cases} 2n^2 x & \text{für } 0 \leq x \leq \frac{1}{2n}, \\ 2n^2\left(\frac{1}{n} - x\right) & \text{für } \frac{1}{2n} < x \leq 1 \end{cases}.$$

Dies definiert wieder eine Folge von stetigen Funktionen auf dem Intervall $[0,1]$, die punktweise und nicht gleichmäßig gegen die identisch verschwindende Funktion konvergiert; aber jetzt findet man

$$\int_0^1 g_n(x)\,dx = \frac{1}{2},$$

und diese Integrale konvergieren offenbar nicht gegen 0.

Mann kann die beiden Beispiele noch weiter variieren, indem man etwa in der Definition der g_n die *Steigung* der Geradenstücke $2n^2$ durch $2n^3$ ersetzt: dies ergibt eine Folge von gegen die Nullfunktion konvergierenden Funktionen, deren Integrale aber sogar über alle Grenzen wachsen.

Schließlich kann man auch Beispiele von glatten Funktionen konstruieren, die zeigen, daß die Vorausetzung der gleichmäßigen Konvergenz nicht überflüssig ist:

$$h_n(x) = 2nx\mathrm{e}^{-nx^2}$$

definiert eine Folge von auf ganz $\mathbb{R}$ glatten Funktionen, die punktweise gegen die Nullfunktion konvergiert; dabei ist aber

$$\int_0^1 h_n(x)\,dx = \int_0^1 2nx\mathrm{e}^{-nx^2}\,dx = -\mathrm{e}^{-nx^2}\Big|_O^1 = 1 - \mathrm{e}^{-n^2},$$

und dies sind die Glieder einer gegen 1, und nicht gegen 0 konvergenten Folge.

Als nächstes diskutieren wir eine Anwendung des Integralbegriffs in der Theorie der unendlichen Reihen, ein wichtiges weiteres Konvergenzkriterium.

Integralkriterium für unendliche Reihen

Satz. *Sei f eine auf dem Intervall $[1,+\infty)$ erklärte, stetige und monoton fallende Funktion, es gelte $f(x) \geq 0$ für alle $x \in [1,\infty)$. Dann sind die beiden folgenden Aussagen äquivalent:*

(i) *es gibt eine reelle Zahl a, sodaß* $\lim\limits_{n\to\infty} \displaystyle\int_0^n f(x)\,dx = a$,

(ii) *die unendliche Reihe* $\displaystyle\sum_{k=1}^{\infty} f(k)$ *ist konvergent.*

Beweis. Da f nirgenwo negative Werte annimmt, ist sowohl die Folge der $\int_0^n f(x)\,dx$ als auch die Partialsummenfolge der Reihe $\sum_{k=1}^{\infty} f(k)$ monoton steigend. Daher ist jede dieser beiden Folgen dann und nur dann konvergent, wenn sie nach oben beschränkt ist. Wegen der Monotonie von f und nach dem Mittelwertsatz der Integralrechnung hat man nun aber

$$f(k+1) \leq \int_k^{k+1} f(x)\,dx \leq f(k),$$

und daher

$$\sum_{k=1}^{n} f(k+1) \leq \int_1^{n+1} f(x)\,dx \leq \sum_{k=1}^{n} f(k).$$

Also ist die Folge der Partialsummen der Reihe in (ii) dann und nur dann beschränkt, wenn die Folge der Integrale in (i) beschränkt ist. Damit ist der Satz bewiesen.

Als Illustration zum Satz betrachten wir die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^s},$$

wobei s eine vorgegeben reelle Zahl bedeutet. Für $s = 1$ erkennen wir die harmonische Reihe, und diese divergiert. Für $s \neq 1$ haben wir

$$\int_1^N \frac{dx}{x^s} = \frac{x^{1-s}}{1-s}\bigg|_1^N = \frac{1}{1-s}\big(N^{1-s} - 1\big),$$

und die rechte Seite strebt offenbar genau dann mit wachsendem N gegen einen endlichen Grenzwert, falls $s > 1$ ist. Nach dem letzten Satz ist die Reihe $\sum_{n=1}^{\infty} n^{-s}$ daher genau dann konvergent, wenn $s > 1$ ist.

Berechnung von Bogenlängen

Eine andere Anwendung des Integralbegriffs ist die Berechnung der Bogenlänge einer Kurve. Wir beschränken uns hier auf solche Kurven, die man als Graph einer stetigen Funktion beschreiben kann. Sei dazu f eine auf dem Intervall $[a,b]$ erklärte, stetige Funktion. Zur Berechnung der Bogenlänge des Graphen

$$G_f = \{(x, f(x)) | a \leq x \leq b\}$$

betrachten wir eine Einteilung

$$E: \qquad a = x_0 < x_1 < x_2 < \ldots < x_q = b$$

des Intervalls $[a, b]$. Zu dieser Einteilung gehört der *Streckenzug* mit den Endpunkten

$$P_k := (x_k, f(x_k)),$$

d.h. der Streckenzug, der entsteht, wenn man P_1 mit P_2 verbindet, P_2 mit P_3 etc.. Wir stellen uns vor, daß dieser Streckenzug desto besser den Graphen G_f approximiert, desto kleiner der Feinheitsgrad $\varphi(E)$ ist; insbesondere stellen wir uns dabei vor, daß die Länge dieses Streckenzuges eine Approximation der gesuchten Bogenlänge ist. Aus der Elementargeometrie weiß man, daß der Abstand zweier Punkte (x, y) und (x', y') durch

$$\sqrt{(x - x')^2 + (y - y')^2}$$

gegeben ist. Demnach hat unser Streckenzug die Länge

$$\sum_{k=1}^{q} \sqrt{(x_k - x_{k-1})^2 + (f(x_k) - f(x_{k-1}))^2}$$
$$= (x_k - x_{k-1}) \cdot \sum_{k=1}^{q} \sqrt{1 + \left(\frac{f(x_k) - f(x_{k-1})^2}{x_k - x_{k-1}} \right)^2}.$$

Unter der Wurzel steht aber ein Differenzenquotient, sodaß wir nach dem Mittelwertsatz für die Länge des Streckenzuges auch

$$(x_k - x_{k-1}) \cdot \sum_{k=1}^{q} \sqrt{1 + f'(\xi_k)^2}$$

schreiben können, wo die $\xi_k \in (x_{k-1}, x_k)$ geeignete Zwischenstellen bezeichnen. Hierbei haben wir natürlich vorauszusetzen, daß f differenzierbar ist. Damit haben wir aber die Länge des zu E gehörenden Streckenzuges als eine Riemannsche Summe der Funktion

$$\sqrt{1 + f'(x)^2}$$

identifiziert. Nehmen wir jetzt noch an, daß die Ableitung von f im Intervall (a, b) stetig ist, und stetig auf $[a, b]$ fortgesetzt werden kann, so finden wir, daß für jede Folge von Einteilungen E_n die Länge der zugehörigen Streckenzüge gegen das Integral

$$\int_a^b \sqrt{1 + f'(t)^2}$$

konvergiert, wenn nur die Folge der Feinheiten der E_n gegen 0 konvergiert. Daher ist es plausibel, die folgende Definition zu treffen:

Definition der Bogenlänge

> Sei f eine auf dem Intervall $[a,b]$ stetige, im Intervall (a,b) differenzierbare Funktion, sodaß die Ableitung von f zu einer auf $[a,b]$ stetigen Funktion fortgesetzt werden kann. Als *Bogenlänge des Graphen von f* bezeichnet man dann das Integral
>
> $$\int_a^b \sqrt{1+f'(x)^2}\,dx.$$

Länge des Kreisbogens

Als Beispiel berechnen wir die Länge eines Kreisbogens, d.h. wir berechnen die Bogenlänge des Graphen der Funktion $y = \sqrt{1-x^2}$, etwa über dem Intervall $[0,t]$ mit $t < 1$. Dazu berechnen wir zunächst:

$$\frac{dy}{dx} = \frac{-x}{\sqrt{1-x^2}}.$$

Damit finden wir

$$\sqrt{1+\left(\frac{dy}{dx}\right)^2} = \frac{1}{\sqrt{1-x^2}}.$$

Eine Stammfunktion dieser Funktion ist $\arcsin x$, und so finden wir

> Die Bogenlänge des Graphen von $y = \sqrt{1-x^2}$ über dem Intervall $[0,t]$ ist $\arcsin t$.

Hierbei hatten wir in der hier entwickelten Theorie vorauszusetzen, daß t strikt kleiner als 1 ist, da ja $\frac{dy}{dx}$ sich nicht stetig nach 1 fortsetzen läßt (strebt x nach 1, so wächst die Ableitung von y über alle Grenzen). Allerdings hat $\arcsin t$ einen Grenzwert, wenn t gegen 1 geht, und sehen wir diesen als Bogenlänge des Viertelkreises an, so finden wir den Wert $\frac{\pi}{4}$ für ebendiese Bogenlänge.

Trapez- und Keplersche Faßregel

Abschließend wollen wir noch zwei Methoden besprechen, Integrale näherungsweise zu berechnen. Dazu betrachten wir eine auf einem abgeschlossenen Intervall stetige Funktion $f(x)$. Die Idee zur näherungsweisen Berechnung des Integrals von f ist, die Funktion f erst durch eine einfache, leicht intergierbare Funktion zu approximieren, und dann das Integral der approximierenden Funktion als gesuchten Näherungswert zu nehmen. Die einfachsten Funktionen hierzu sind vielleicht die linearen und quadratischen Polynome. Indem wir gegebenenfalls das abgeschlossene Intervall verschieben (d.h. die Integrationsvariable x durch $x + c$ mit einer geeigneten Konstanten c ersetzen), können wir im Folgenden annehmen, daß unser Intervall die Gestalt $[-h,h]$ hat.

Als lineare Funktion $l(x)$, die man als Approximation für f nehmen könnte, fällt einem zuerst die Funktion ein, die an den Intervallenden die gleichen Werte wie f annimmt:

$$l(x) = \frac{1}{2h}[f(h)(x+h) - f(-h)(x-h)].$$

Integriert man diese Funktion über dem Intervall $[-h, h]$, so findet man als Wert des Integrals

$$h[f(h) + f(-h)].$$

Dies ist die sogenannte *Trapezregel* zur (näherunsweisen) Berechnung von $\int_{-h}^{h} f(x)\, dx$.

Das naheliegendste quadratische Polynom $q(x)$ zur Approximation von f ist vielleicht dasjenige, welches an den Intervallenden und im Intervallmittelpunkt die gleichen Werte wie f annimmt:

$$q(x) = \frac{1}{2h^2}[f(h)x(x+h) - 2f(0)(x+h)(x-h) + f(-h)x(x-h)].$$

Für das Integral von q über dem Intervall $[-h, +h]$ findet man

$$\frac{h}{3} \cdot [f(h) + 4f(0) + f(-h)].$$

Dies wird als *Keplersche Faßregel* oder *Simpsonsche Regel* zur näherungsweisen Berechnung des Integrals von f bezeichnet.

Wir wollen nun den Fehler bei der Trapez- und Faßregel untersuchen. Dazu setzen wir für $0 \le t \le h$

$$\Delta_{\text{Tr.}}(t) = \int_{-t}^{+t} f(x)\, dx - t[f(t) + f(-t)]$$
$$\Delta_{\text{Faß.}}(t) = \int_{-t}^{+t} f(x)\, dx - \frac{t}{3} \cdot [f(t) + 4f(0) + f(-t)].$$

Zur Bestimmung von $\Delta_{\text{Tr.}}(h)$ nehmen wir an, daß f auf $[-h, +h]$ zweimal stetig differenzierbar ist, d.h. daß f' und f'' in $(-h, +h)$ existieren und auf $[-h, +h]$ stetig fortgesetzt werden können. Dann ist aber $\Delta_{\text{Tr.}}$ auf dem Intervall $[0, h]$ einmal stetig differenzierbar. Für die Ableitung findet man

$$\Delta'_{\text{Tr.}}(t) = -h[f'(t) - f'(-t)].$$

Wenden wir hierauf den Mittelwertsatz an, so haben wir

$$\Delta'_{\text{Tr.}}(t) = -2t^2 f''(\xi)$$

mit einem geeigneten $\xi \in (-t, +t)$, also jedenfalls

$$\left|\Delta'_{\text{Tr.}}(t)\right| \leq 2t^2 \cdot \|f''\|, \qquad \|f''\| := \sup_{x \in (-h,+h)} |f''(x)| \,.$$

Dann ist aber wegen

$$\Delta_{\text{Tr.}}(h) = \Delta_{\text{Tr.}}(h) - \Delta_{\text{Tr.}}(0) = \int_0^h \Delta'_{\text{Tr.}}(t)\, dt$$

schließlich

$$\left|\Delta_{\text{Tr.}}(h)\right| \leq \int_0^h \left|\Delta'_{\text{Tr.}}(t)\right| \, dt \leq \int_0^h 2t^2 \, dt \cdot \|f''\| = \frac{2}{3}h^3 \cdot \|f''\|$$

Zur Abschätzung von $\Delta_{\text{Faß.}}(h)$ setzen wir voraus, daß f in $[-h, +h]$ viermal stetig differenzierbar ist. Mit einer etwas längeren Rechnung findet man dann

$$\Delta'''_{\text{Faß.}}(t) = -\frac{h}{3}[f'''(h) - f'''(-h)],$$

nach dem Mittelwertsatz also wie oben

$$\left|\Delta'''_{\text{Faß.}}(t)\right| \leq \frac{2t^2}{3} \cdot \|f^{(4)}\|.$$

Damit findet man dann (wieder mit der gleichen Methode wie oben):

$$\left|\Delta''_{\text{Faß.}}(h)\right| \leq \frac{2}{9}h^3 \cdot \|f^{(4)}\|, \qquad \left|\Delta'_{\text{Faß.}}(h)\right| \leq \frac{1}{18}h^4 \cdot \|f^{(4)}\|,$$

$$\left|\Delta'_{\text{Faß.}}(h)\right| \leq \frac{1}{90}h^5 \cdot \|f^{(4)}\|$$

Damit haben wir den folgenden Satz bewiesen:

Fehler bei der Trapez- und Faßregel

Satz. *Sei f eine auf dem Intervall $[-h, h]$ erklärte, stetige Funktion. Dann gilt:*

(i) *Ist f zweimal differenzierbar in $(-h, h)$, sodaß sich die zweite Ableitung zu einer auf $[-h, +h]$ stetigen Funktion fortsetzen läßt, so ist*

$$\left| \int_{-h}^{+h} f(x)\, dx - h[f(h) + f(-h)] \right| \leq \frac{2}{3}h^3 \cdot \|f''\|.$$

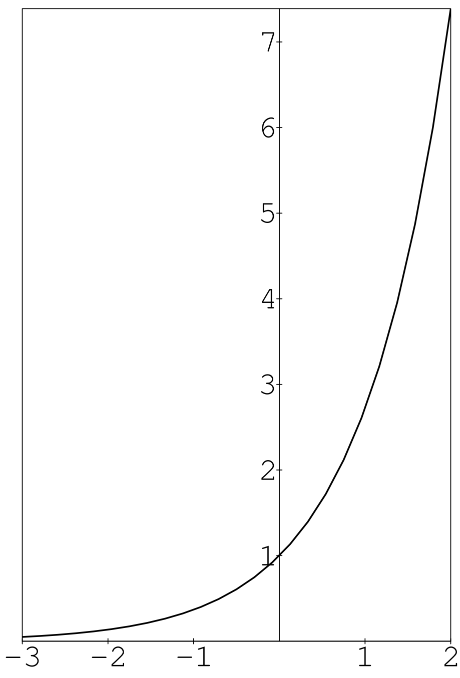
(ii) *Ist f viermal differenzierbar in $(-h, h)$, sodaß sich die vierte Ableitung zu einer auf $[-h, +h]$ stetigen Funktion fortsetzen läßt, so ist*

$$\left| \int_{-h}^{+h} f(x)\, dx - \frac{h}{3}[f(h) + 4f(0) + f(-h)] \right| \leq \frac{h^5}{90}\|f^{(4)}\|.$$

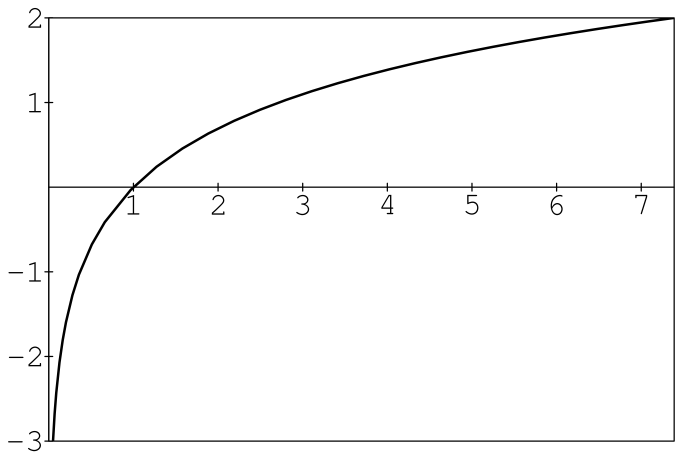
Hierbei ist jeweils $\|f^{(k)}\| = \sup_{x \in (-h,h)} \left|f^k(x)\right|$.

A DIE GRAPHEN EINIGER ELEMENTARER FUNKTIONEN

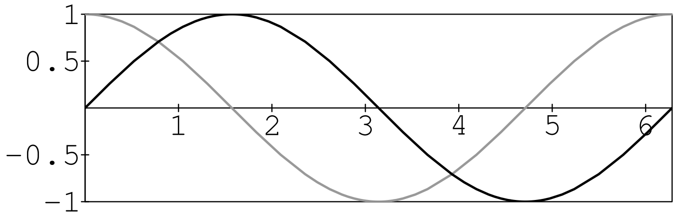
In diesem Anhang sind die Graphen der Funktionen skizziert, die in den vorangegangenen Kapiteln wiederholt als Beispiele besprochen wurden. Sind mehrere Graphen in einer einzigen Darstellung gegeben, so ist der Graph der im Text zuerst genannten Funktion hell, der Graph der im Text als nächstes genannten Funktion etwas dunkler etc..



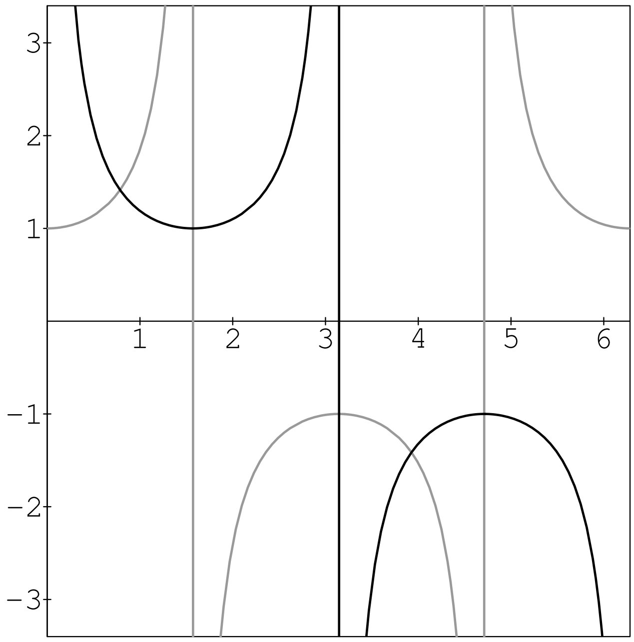
Die Exponentialfunktion e^x



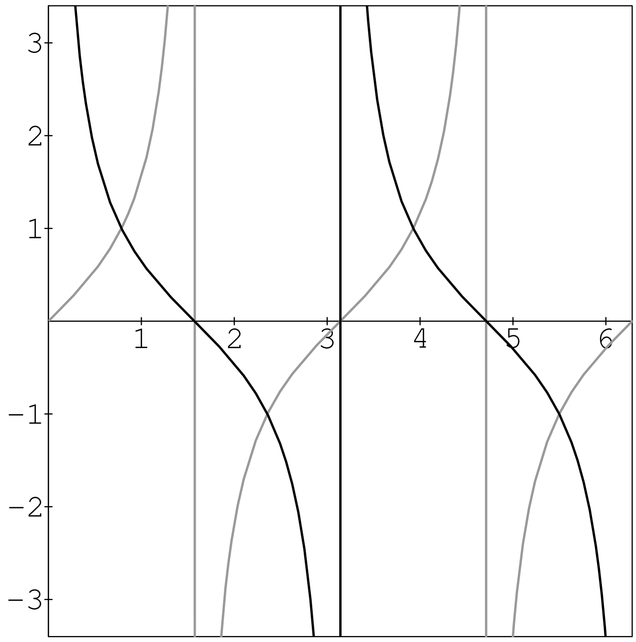
Der Logarithmus $\log x$



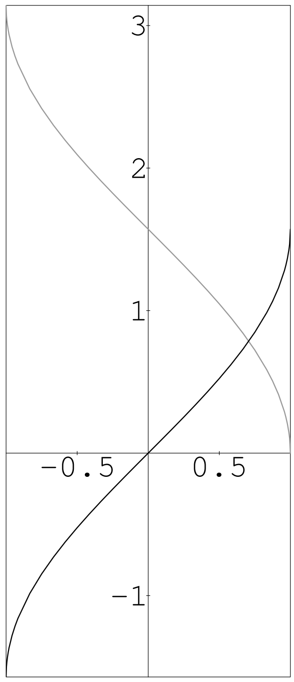
Cosinus $\cos x$ und Sinus $\sin x$



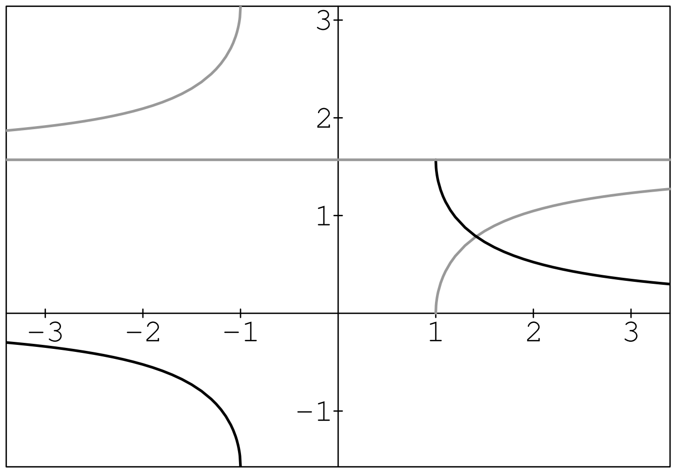
Secans $\sec x = \frac{1}{\cos x}$ und Cosecans $\csc x = \frac{1}{\sin x}$



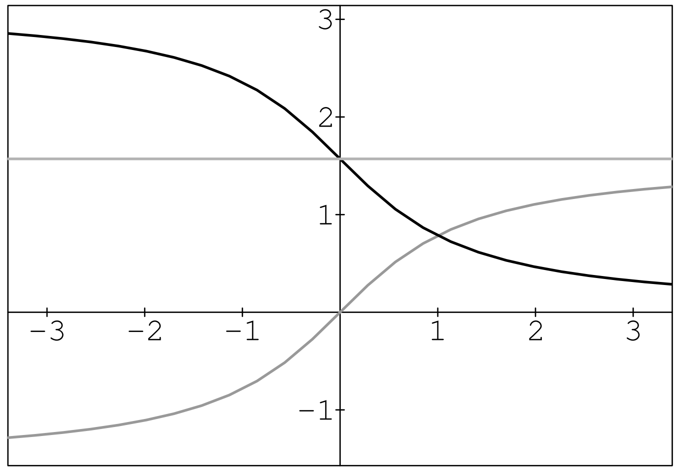
Tangens $\tan x = \frac{\sin x}{\cos x}$ und Cotangens $\cot x = \frac{\cos x}{\sin x}$



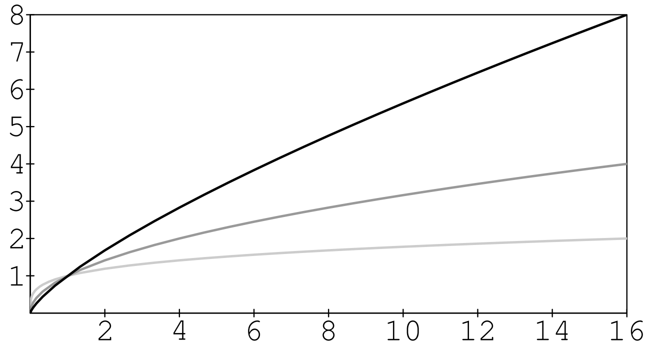
Arcuscosinus $\arccos x$ und $\arcsin x$



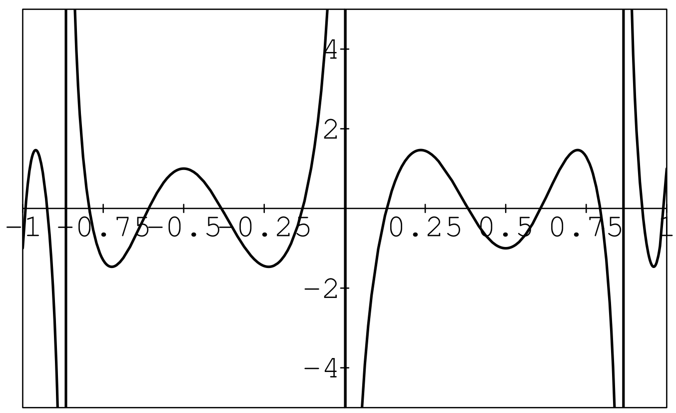
Arcussecans $\operatorname{arcsec} x$ und Arcuscosecans $\operatorname{arccsc} x$



Arcustangens $\arctan x$ und $\operatorname{arccot} x$



Die Funktionen $x^{\frac{1}{2}}$, $x^{\frac{2}{3}}$, und $x^{\frac{3}{4}}$



Eine rationale Funktion: $\frac{1-72x^2+840x^4-3584x^6+6912x^8-6144x^{10}+2048x^{12}}{-3x+4x^3}$

B AUFGABEN ZU DEN EINZELNEN KAPITELN

EINIGE HISTORISCHE BEMERKUNGEN ÜBER ZAHLSYSTEME

1. Zeigen Sie: $\sqrt{3}$ ist irrational.

2. Für welche natürlichen Zahlen n ist $\sqrt{n}$ irrational?

3. Für welche ganzen Zahlen p, q sind die Nullstellen der Gleichung $x^2 + px + q = 0$ irrational.

4. Was ist der Wert des unendlichen Kettenbruchs

$$4 + \cfrac{1}{2 + \cfrac{1}{8 + \cfrac{1}{2 + \cfrac{1}{8 + \cfrac{1}{2 + \cdots}}}}} \quad ?$$

Berechnen Sie die ersten Konvergenten des Kettenbruchs aus der vorstehenden Aufgabe. Berechnen Sie den Wert von $2\sqrt{5}$ mit einem Taschenrechner. Vergleichen Sie.

5. Was ist der Wert des unendlichen Kettenbruchs

$$n + \cfrac{1}{n + \cfrac{1}{n + \cfrac{1}{n + \cfrac{1}{n + \cfrac{1}{n + \cdots}}}}}$$

für eine gegebene natürliche Zahl n ?

6. Was ist der Wert der unendlichen Reihe $1 + \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \frac{1}{81} + \cdots$?

(Subtrahiert man von der Reihe 1 und multipliziert das Ergebnis mit 3, so erhält man ...?)

7. Schreiben Sie $\frac{61111}{49500}$ als Dezimalzahl.

8. Man veranschauliche die Addition und Multiplikation von komplexen Zahlen, indem man die komplexe Zahl $a + ib$ mit dem Punkt mit Koordinaten (a, b) in der Zahlenebene identifiziert.

DIE AXIOMATIK DER RELLEN ZAHLEN

9. Man zeige, daß die Menge der komplexen Zahlen mit Absolutbetrag 1 unter Multiplikation abgeschlossen ist.

10. Es bezeichne M eine Menge und $\mathcal{P}(M)$ die Menge aller Teilmengen von M. Für zwei Elemente A, B aus $\mathcal{P}(M)$ setzen wir

$$A + B := (A \setminus B) \cup (B \setminus A) \quad \text{und} \quad A \cdot B := A \cap B.$$

Zeige: Die Menge $\mathcal{P}(M)$ zusammen mit den Operationen "+" und "·" bildet einen kommutativen Ring. Für welche Mengen M erhält man sogar einen Körper ?

11. (i) Eine natürliche Zahl ist genau dann durch 3 teilbar, wenn es ihre Quersumme ist. (Hinweis: $10 = 3 \cdot 3 + 1$.)

(ii) Finden Sie einen ähnlichen Test wie in der vorstehenden Aufgabe für Teilbarkeit durch die Zahl 11.

12. Seien a, b reelle Zahlen, es gelte $0 \le a \le b$. Zeige:

$$0 \le a \le \sqrt{ab} \le \frac{a+b}{2} \le b.$$

Wann kann hier das vorletzte "$\le$" durch "$=$" ersetzt werden ?

13. Bestimme die reellen Zahlen x für die gilt:

(i) $\frac{1}{x} + \frac{1}{1-x} > 0$

(ii) $\frac{x-1}{x+1} > 0$

(iii) $x^2 + px + q > 0 \quad (p, q \in \mathbb{R})$

(iv) $|x - 1| + |x + 1| < 4$

(v) $|x - 1| \cdot |x + 1| < 4$.

14. Zeige:

$$\max(x, y) = \frac{x + y + |x - y|}{2}, \quad \min(x, y) = \frac{x + y - |x - y|}{2}.$$

(Die Symbole $\max(x, y)$ und $\min(x, y)$ bezeichnen jeweils die größere bzw. die kleinere der beiden reellen Zahlen x, y.)

15. Finde analoge Formeln wie in der vorhergehenden Übungsaufgabe für $\max(x, y, z)$ und $\min(x, y, z)$.

16. Schreibe die folgenden Ausdrücke mit mindestens einem Absolutbetragstrich weniger:

(i) $\left|\sqrt{2}+\sqrt{3}-\left|\sqrt{5}-\sqrt{7}\right|\right|$

(ii) $||a+b|-|a|-|b||$

(iii) $\left|x^2-2xy+y^2\right|$

(iv) $||a+b|+|c|-|a+b+c||$

(v) $\left|\sqrt{2}+\sqrt{3}-\sqrt{5}+\sqrt{7}\right|$.

17. Zeigen Sie: $||x|-|y|| \leq |x-y|$ für alle $x, y \in \mathbb{R}$.

ÜBER VOLLSTÄNDIGE INDUKTION

18. Finde und beweise (durch vollständige Induktion) eine Formel für

(i) $\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\cdots+\frac{1}{n(n+1)}$,

(Hinweis: $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$.)

(ii) $\frac{3}{1^2\cdot 2^2}+\frac{5}{2^2\cdot 3^2}+\cdots+\frac{2n+1}{n^2(n+1)^2}$.

19. Beweise die Bernoullische Ungleichung:

$$(1+h)^n \geq 1+hn \quad \text{fuer alle } n \in \mathbb{N} \text{ und } h \in \mathbb{R} \text{ mit } h \geq -1.$$

Wann gilt hier "=" ?

20. Berechne die ersten Konvergenten $\frac{p_n}{q_n}$ von

$$\cfrac{4}{1+\cfrac{1^2}{2+\cfrac{3^2}{2+\cfrac{5^2}{2+\cfrac{7^2}{2+\cdots}}}}} .$$

Finde eine Rekursionsformel für p_n und q_n.

21. Finde und beweise Formeln für

(i) $1+3+5+\cdots+(2n+1)$

(ii) $1+9+25+\cdots+(2n+1)^2$.

22. Zeigen Sie

(i) Es gibt genau eine Zahlenfolge B_n ($n \in N$) mit der folgenden Eigenschaft: Es ist $B_0 = 1$ und für jede natürliche Zahl $n \geq 1$ und jede reelle Zahl x ist

$$B_n(x+1) - B_n(x) = nx^{n-1},$$

wenn man $B_n(x)$ durch

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} B_k x^{n-k}$$

erklärt.

(ii) Finden Sie eine Formel für

$$1 + 2^n + 3^n + 4^n + \cdots + k^n$$

zu jeder natürlichen Zahl n.

23. Zeigen Sie durch vollständige Induktion:

$$(a^n + a^{n-1}b + a^{n-2}b^2 + \cdots + ab^{n-1} + b^n)(a-b) = a^{n+1} - b^{n+1}.$$

GENERAL NONSENSE UND ELEMENTARE KOMBINATORIK

24. Seien S, T endliche Mengen. Bestimmen Sie die Anzahl der injektiven Abbildungen von S nach T.

25. Seien A_i ($1 \leq i \leq n$) endliche Mengen. Zeigen Sie:

$$\left| \bigcup_{1 \leq i \leq n} A_i \right| = - \sum_{\emptyset \neq I \subset \{1,\ldots,n\}} (-1)^{|I|} \left| \bigcap_{i \in I} A_i \right|.$$

26. Es bezeichne R einen kommutativen Ring und M eine beliebige Menge. Sind $f, g \in R^M$, so erklären wir die *Summe* $f+g$ und das *Produkt* $f \cdot g$, indem wir

$$(f+g)(x) = f(x) + g(x), \quad (f \cdot g)(x) = f(x) \cdot g(x) \qquad (x \in M)$$

setzen. Zeigen Sie: R^M wird vermöge dieser Operationen zu einem kommutativen Ring.

27. Für beliebige Abbildungen $f\colon S \to T$ und Teilmengen $X, Y \subset S$ gilt im allgemeinen *nicht* $f(X \cap Y) = f(X) \cap f(Y)$. Finden Sie ein Gegenbeispiel. Was gilt mit '$\cup$' an Stelle von '$\cap$' ?

28. (i) Bestimmen Sie die Anzahl der surjektiven Abbildungen einer endlichen Menge S auf eine endliche Menge T.

(ii) Bestimmen Sie die Anzahl der Zerlegungen einer endlichen Menge in eine vorgegebene Anzahl von nichtleeren, paarweise disjunkten Teilmengen. Hinweis: Die ersten Anzahlen von Zerlegungen sind:

$$\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 3 & & 1 & & & & \\
 & & & 1 & & 7 & & 6 & & 1 & & & \\
 & & 1 & & 15 & & 25 & & 10 & & 1 & & \\
 & 1 & & 31 & & 90 & & 65 & & 15 & & 1 & \\
1 & & 63 & & 301 & & 350 & & 140 & & 21 & & 1
\end{array}$$

29. Sei $f: S \to T$ eine Abbildung. Zeigen Sie, daß sich f als $f = i \circ s$ mit einer surjektiven Abbildung $s: S \to X$ und einer injektiven Abbildung $i: X \to T$ schreiben läßt.

DAS VOLLSTÄNDIGKEITS AXIOM

30. Bestimmen Sie das Supremum und das Infimum der folgenden Mengen, sofern sie existieren:

(i) $\left\{\frac{1}{n} \mid n \in \mathbb{N} \setminus \{0\}\right\}$

(ii) $\left\{\frac{1}{n} \mid n \in \mathbb{Z} \setminus \{0\}\right\}$

(iii) $\left\{x \in \mathbb{R} \mid x^2 + x - 1 < 0\right\}$

(iv) $\left\{x \in \mathbb{R} \mid x^2 + x - 1 > 0\right\}$

(v) $\left\{\frac{1}{n} + (-1)^n \mid n \in \mathbb{N} \setminus \{0\}\right\}$.

31. Seien A, B nichtleere, nach oben beschränkte Teilmengen von $\mathbb{R}$. Dann ist auch $A + B$ nichtleer und nach oben beschränkt, und es gilt

$$\sup(A + B) = \sup(A) + \sup(B).$$

(Hier bezeichnet $A + B$ die Menge aller reellen Zahlen, die man in der Gestalt $x + y$ mit $x \in A$, $y \in B$ schreiben kann. Hinweis: Die Ungleichung '$\sup(A + B) \leq \sup(A) + \sup(B)$' ist einfach einzusehen; für die umgekehrte Ungleichung genügt es $\sup(A+B) \geq \sup(A)+\sup(B)-\varepsilon$ für jedes $\varepsilon > 0$ zu beweisen—warum?)

32. Zeigen Sie, daß $\frac{n!}{n^n}$ eine Nullfolge ist.

33. Für eine von 0 verschiedene natürliche Zahl n bezeichne $\alpha(n)$ die Anzahl der Primteiler von n, also

n	1	2	3	4	5	6	7	8	9	10	...
$\alpha(n)$	1	1	1	2	1	2	1	3	2	2	...

Zeigen Sie, daß $\frac{\alpha(n)}{n}$ eine Nullfolge ist. (Hinweis: vergleiche mit Anwesenheitsaufgabe 9.)

34. Sei A eine nichtleere, nach unten beschränkte Teilmenge der reellen Zahlen. Zeigen Sie

$$\sup\{x \mid x \le A\} = \inf A.$$

(Hierbei ist '$x \le A$' eine abkürzende Schreibweise für '$x \le y$ für alle $y \in A$'.)

35. Seien A, B nichtleere Teilmengen von $\mathbb{R}$, und es gelte $A \le B$ (d.h. $x \le y$ für alle $x \in A, y \in B$). Zeigen Sie, daß A nach oben und B nach unten beschränkt ist, und daß $\sup A \le \inf B$ gilt.

36. Jede nach unten beschränkte, monoton fallende Folge von reellen Zahlen ist konvergent.

37. Für eine natürliche Zahl n sei $\nu(n) = \sup\{\nu \in \mathbb{N} \mid 2^\nu \le n\}$. Zeigen Sie, daß $\frac{\nu(n)}{n}$ eine Nullfolge ist.

38. Zeigen Sie, daß sich jede reelle Zahl als *Dualzahl* schreiben läßt; genauer: zu jeder reellen Zahl x gibt es eine ganze Zahl e und eine Folge $(z_i)_{e \le i < \infty}$ mit $z_i \in \{0,1\}$ für alle $e \le i < \infty$, sodaß

$$x = \operatorname{sign}(x) \lim_{n\to\infty} \sum_{i=e}^{n} z_i 2^{-i}.$$

Zeigen Sie, daß für jedes von 0 veschiedene x die Folge der z_i so gewählt werden kann, daß $z_e \ne 0$ und $z_i = 0$ für unendlich viele i ist, und daß es zu jedem x auch *höchstens eine* solche Darstellung als Dualzahl gibt.

39. Konstruieren Sie eine surjektive Abbildung

$$f\colon \mathcal{P}(N) \longrightarrow \{x \in R | 0 \le x < 1\}$$

(Hinweis: vorstehende Aufgabe). Beschreiben Sie eine Teilmenge von $\mathcal{P}(\mathbb{N})$, sodaß die Einschränkung Ihrer Abbildung auf diese Teilmenge bijektiv wird.

40. Beweisen Sie die folgenden Identitäten:

(i) $\lim\limits_{n\to\infty} \sqrt[n]{a} = 1 \quad (a > 0)$

(ii) $\lim\limits_{n\to\infty} \sqrt[n]{a^n + b^n} = \max(a, b) \quad (a, b > 0)$

(iii) $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$

(iv) $\lim\limits_{n\to\infty} \sqrt[n]{17n^{17} + 16n^{16} + \cdots + 2n^2 + n} = 1.$

41. Finden Sie alle konvergenten Teilfolgen der Folge

$$1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, \ldots.$$

42. Ist a_0 eine ganze Zahl, $a_1, a_2, \ldots$ eine Folge von positiven natürlichen Zahlen, so setzt man

$$[a_0, a_1, \ldots, a_n] := a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ldots + \cfrac{1}{a_n}}}}.$$

Zeigen Sie:

(i) Definiert man p_n, q_n durch

$$\begin{array}{llll} p_0 = a_0, & p_1 = a_0 a_1 + 1, & p_n = a_n p_{n-1} + p_{n-2} & (n \geq 2) \\ q_0 = 1, & q_1 = a_1, & q_n = a_n q_{n-1} + q_{n-2} & (n \geq 2), \end{array}$$

so gilt
$$\frac{p_n}{q_n} = [a_0, \ldots, a_n].$$

(ii) Es ist $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$.

(iii) Mit $x_n := \frac{p_{2n}}{q_{2n}}$ und $y_n := \frac{p_{2n+1}}{q_{2n+1}}$ wird eine Intervallschachtelung $(x_n|y_n)_{n\in\mathbb{N}}$ definiert. (Insbesondere existiert stets der Grenzwert $[a_0, a_1, a_2, \ldots] := \lim\limits_{n\to\infty} [a_o, a_1, \ldots, a_n]$.)

43. Zeigen Sie: Jede irrationale reelle Zahl x besitzt eine eindeutige *Darstellung als unendlicher Kettenbruch;* genauer: es gibt eine und nur eine Folge ganzer Zahlen $(a_n)_{n\in\mathbb{N}}$ mit positiven a_n für $n \geq 1$, sodaß $x = [a_0, a_1, a_2, \ldots]$ gilt.

(Hinweis: Betrachten Sie die Folge $x_0 = x,\ x_n = [x_n] + \frac{1}{x_{n+1}}$ —hierbei steht das Symbol $[y]$ für die größte ganze Zahl, die noch kleiner oder gleich y ist.)

44. (i) Zeigen Sie, daß sich jede rationale Zahl in einen *endlichen Kettenbruch* entwickeln läßt, d.h. daß sich jede rationale Zahl als

$[a_0, a_1, \ldots, a_n]$ mit geeigneten ganzen Zahlen a_r, $a_r > 0$ für $r \geq 1$, schreiben läßt.

(ii) Entwickeln Sie $-691/2730$ in einen Kettenbruch. Ist Ihre Darstellung eindeutig ?

45. Für welche x existiert $\lim_{n\to\infty} (-1)^n \binom{x}{n} n^{1+x}$? Berechnen Sie den Grenzwert für $x \in \mathbb{Z}$.

(Hinweis: Man kann etwa so vorgehen: zunächst untersucht man für solche x, wo die entsprechende Folge monoton ist (hierbei darf ohne Beweis die für *beliebige* reelle $x \geq 1$ und $h \geq 0$ geltende Bernoullische Ungleichung "$(1+h)^x \geq 1+hx$" angewandt werden); die Untersuchung für beliebige x führt man auf die schon untersuchten Fälle zurück, indem man $\binom{x}{n} = \frac{x}{n}\binom{x-1}{n-1}$ ausnutzt.)

UNENDLICHE REIHEN

46. Sei $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$ und $\lim_{n\to\infty} a_n = 0$. Dann konvergiert die Reihe $\sum_{n=0}^{\infty} (-1)^n a_n$.

(Hinweis: Denken Sie an Intervallschachtelungen !)

47. Zeigen Sie, daß $\sum_{n=0}^{\infty} \binom{x}{n}$ für jede reelle Zahl $x > -1$ konvergiert und sonst divergiert. Für welche Zahlen x ist die Reihe absolut konvergent ? Raten Sie den Wert der Reihe für $x = 1/2$ —Sie müssen Ihre Vermutung allerdings nicht beweisen.

(Hinweis: Kann man die vorstehende Aufgabe anwenden ?)

48. Entscheiden Sie, welche der folgenden Reihen konvergent bzw. divergent ist.

$$\sum_{n=0}^{\infty} \frac{n^2}{n!} \qquad \sum_{n=0}^{\infty} \frac{n^2}{n^3+1} \qquad \sum_{n=0}^{\infty} \frac{n^k + a_k n^{k-1} + \cdots + a_1 n + a_0}{n^l + b_l n^{l-1} + \cdots + b_1 n + b_0}$$

$$\sum_{n=0}^{\infty} \frac{n^n}{(2.7)^n n!} \qquad \sum_{n=0}^{\infty} \frac{n^n}{(2.8)^n n!}.$$

49. Für welche Zahlen x konvergiert $\displaystyle\sum_{n=1}^{\infty} \frac{x^n n!}{n^n}$?

50. Für welche x und k konvergiert $\displaystyle\sum_{n=1}^{\infty} n^k x^n$?

51. Streichen Sie in der harmonischen Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ alle Terme $\frac{1}{n}$, wo die Dezimaldarstellung von n die Ziffer 0 enthält. Ist die verbleibende unendliche Reihe divergent oder konvergent ?

52. Man konstruiere eine Folge (a_n) von positiven reellen Zahlen mit den folgenden Eigenschaften: Die Folge $\left(\frac{a_{n+1}}{a_n}\right)$ ist nach oben unbeschränkt und $\sum_n a_n$ konvergiert.

53. Es bezeichne F_n die n-te Fibonacci-Zahl, d.h. es ist $F_n = F_{n-1} + F_{n-2}$ $(n \geq 2)$, $F_0 = F_1 = 1$. Für welche x konvergiert $\sum_{n=0}^{\infty} F_n x^n$?

54. Sei $\sum_{n=0}^{\infty} a_n$ eine konvergente, aber nicht absolut konvergente Reihe. Zeigen Sie: Zu jeder reellen Zahl x existiert eine Bijektion $\sigma : \mathbb{N} \to \mathbb{N}$, sodaß $\sum_{n=0}^{\infty} a_{\sigma(n)} = x$.

55. Zeigen Sie, daß $\displaystyle\sum_{n=0}^{\infty} \frac{B_n}{n!} x^n = \frac{x}{\mathrm{e}^x - 1}$ für hinreichend kleines $|x|$.

(Die B_n bezeichnen hierbei die durch "$(B+1)^n = B_n$" erklärten Bernoulli-Zahlen — vgl. Aufgabe 22.)

56. Zeigen Sie, daß die Folge der Zahlen $\exp\left(\sum_{\ell=1}^{N} \frac{1}{\ell}\right)/N$ gegen einen endlichen Grenzwert konvergiert.

57. (i) Man zeige für alle reellen Zahlen x und $y \neq 0$ die Identitäten

$$\sum_{\ell=0}^{n-1} \sin(x + \ell y) = \sin\left(x + \frac{n-1}{2} y\right) \cdot \frac{\sin\left(\frac{n}{2} y\right)}{\sin\left(\frac{y}{2}\right)}.$$

(ii) Beweisen Sie eine analoge Formel für $\displaystyle\sum_{\ell=0}^{n-1} \cos(x + \ell y)$.

(Hinweis: Benutzen Sie die Eulersche Gleichung und betrachten Sie $\displaystyle\sum_{\ell=0}^{n-1} \mathrm{e}^{i(x+\ell y)}$.)

58. Zeigen Sie, daß es zu jeder positiven reellen Zahl a und zu jeder ganzen Zahl $n \geq 1$ genau n verschiedene komplexe Zahlen b_i $(1 \leq i \leq n)$ mit der Eigenschaft $b_i^n = a$ gibt.

STETIGKEIT

59. Sei

$$f(x) = \begin{cases} x \sin(\frac{1}{x}) & \text{für } x \in \mathbb{R} - \{0\} \\ 0 & \text{für } x = 0. \end{cases}$$

Zeigen Sie, daß f in 0 stetig ist.

60. (i) Zeigen Sie: Ist f eine auf dem Intervall $[a, b]$ stetige Funktion, so gibt es eine auf ganz $\mathbb{R}$ definierte, stetige Funktion F, sodaß $F(x) = f(x)$ für alle x in $[a, b]$ gilt.

(ii) Zeigen Sie, daß die vorstehende Aussage falsch wird, wenn man $[a, b]$ durch (a, b) ersetzt.

61. Zu einer positiven reellen Zahl a und einer natürlichen Zahl $n \neq 0$ definieren wir eine Funktion $f\colon (0, \infty) \to (0, \infty)$ vermöge

$$f(x) = \frac{(n-1)x^n + a}{nx^{n-1}}.$$

Zeigen Sie, daß für jede positive reelle Zahl x_0 die Folge

$$x_0, f(x_0), f(f(x_0)), f(f(f(x_0))), \cdots$$

konvergiert. Wogegen ? Berechnen Sie die ersten Folgenglieder im Fall $a = 5, n = 2, x_0 = 2$.

62. Zeigen Sie: Ist f eine auf ganz $\mathbb{R}$ definierte, stetige Funktion mit den Eigenschaften $f(x+y) = f(x) \cdot f(y)$, $f(1) = a > 0$, so ist $f(x) = a^x$.

(Hinweis: Sie dürfen ohne Beweis benutzen, daß $x \mapsto a^x$ stetig ist. Zeigen Sie zunächst $f(x) = a^x$ für rationale x.)

63. Sei $f\colon [a, b] \to [a, b]$ eine Funktion mit der folgenden Eigenschaft: es gibt eine reelle Zahl $0 < c < 1$, sodaß für alle x, y in $[a, b]$ die Ungleichung $|f(x) - f(y)| < c|x - y|$ erfüllt ist.

(i) Zeigen Sie, daß f stetig ist.

(ii) Zeigen Sie, daß eine reelle Zahl x in $[a, b]$ existiert, sodaß $f(x) = x$ gilt.

(Hinweis: Betrachten Sie für irgendein x_0 in $[a, b]$ die Folge

$$f(x_0), f(f(x_0)), f(f(f(x_0))), \ldots .)$$

64. Zeigen Sie, daß für jede reelle Zahl a und jede positive reelle Zahl ϵ das Intervall $(a-\epsilon, a+\epsilon)$ sowohl unendlich viele rationale als auch unendlich viele irrationale Zahlen enthält.

65. (i) Seien f, g zwei stetige Funktionen auf dem Intervall $[a, b]$ mit der Eigenschaft $f(a) \geq g(a)$ und $f(b) \leq g(b)$. Zeigen Sie, daß sich die Graphen von f und g in mindestens einem Punkt schneiden.

(ii) Sei $f\colon [0, 1] \to [0, 1]$ eine stetige Funktion. Zeigen Sie, daß f mindestens einen Fixpunkt hat, d.h. daß $f(x) = x$ für mindestens ein x in $[0, 1]$ gilt.

66. Zu jedem normierten Polynom $p(X)$ von geradem Grad, existiert ein c_0, sodaß die Gleichung $p(x) = c$ mindestens eine Lösung für $c \geq c_0$ und keine Lösung für $c < c_0$ hat.

67. Zeigen Sie, daß es zu je $n+1$ vorgegebenen Zahlenpaaren x_k, a_k $(0 \leq k \leq n)$ genau ein Polynom vom Grad n gibt, sodaß $p(x_k) = a_k$ für alle k gilt.

68. Zeigen Sie: es gibt keine auf $\mathbb{R}$ stetige Funktion, die jede reelle Zahl genau 2-mal als Wert annimmt.

(Hinweis: Wäre f eine solche Funktion und $f(a) = f(b)$, so hätte man entweder $f(x) \geq f(a) = f(b) \geq f(y)$ für alle x in und alle y außerhalb $[a, b]$, oder aber das Gleiche mit '$\leq$' statt '$\geq$'.)

69. Finden Sie eine auf $\mathbb{R}$ stetige Funktion, die jede reelle Zahl genau 3-mal als Wert annimmt.

70. Finden Sie eine auf $\mathbb{R}$ definierte Funktion, die

(i) in keinem Punkt stetig ist, wogegen ihr Absolutbetrag in jedem Punkt stetig ist.

(ii) in einer vorgegebenen reellen Zahl a stetig und in jedem anderen Punkt unstetig ist.

(ii) in jedem irrationalen Punkt stetig, in jedem rationalen Punkt unstetig ist.

71. Skizzieren Sie die Graphen der folgenden auf D definierten Funktionen:

(i) $f(x) = a^x \qquad (a > 0, D = \mathbb{R})$

(ii) $g(x) = x^r \qquad (r$ fest vorgegenben, $D = (0, +\infty)$)

(iii) $\Gamma(x) = \lim\limits_{n\to\infty} \dfrac{(-1)^n n^{x-1}}{\binom{-x}{n}} \qquad (D = \mathbb{R} - \{0, -1, -2, -3, \ldots\})$

(zur Existenz der Grenzwerte siehe oben; sie können ohne Beweis benutzen, daß $\Gamma(x)$ stetig ist).

72. Wieviele Schnittpunkte können die Graphen $G_{f_i} = \{(x, f_i(x)) \mid x \in \mathbb{R}\}$ $(i = 1, 2)$ zweier polynomialer Funktion f_1, f_2 höchstens haben ?

DIFFERENZIER-BARKEIT

73. Zeigen Sie, daß die Funktion

$$f(x) = \begin{cases} \mathrm{e}^{-1/x} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$$

unendlich oft differenzierbar ist.

74. Man konstruiere zu vorgegebenen reellen Zahlen $\epsilon > 0$ und a, M eine auf $\mathbb{R}$ unendlich oft differenzierbare Funktion $f(x)$, sodaß stets $f(x) > 0$ für x in und $f(x) = 0$ für x außerhalb $(a-\epsilon, a+\epsilon)$ ist, und zusätzlich noch $\sup_{x\in\mathbb{R}} f(x) = M$ gilt.

(Benutzen Sie die vorstehende Aufgabe.)

75. (i) Seien f, g n-mal differenzierbare Funktionen. Beweisen Sie

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$$

(hierbei bezeichnet $f^{(k)}$ die k-te Ableitung von f).

(ii) Berechnen Sie die Formel, die man erhält, wenn man in der vorstehenden Gleichung $f(x) = \mathrm{e}^{ax}$, $g(x) = \mathrm{e}^{bx}$ $(a, b \in \mathbb{R})$ und $x = 0$ setzt.

(iii) Finden Sie eine Formel, die die n-te Ableitung eines Produkts $(f_1 \cdots f_m)$ durch die Ableitungen $f_h^{(k)}$ $(1 \le h \le m,\ 0 \le k \le n)$ ausdrückt.

76. (i) Beweisen Sie: zu jedem Polynom $p(X)$ gibt es ein Polynom $\widetilde{p}(X)$, sodaß

$$\sum_{k=0}^{n} \binom{n}{k} p(k) = 2^n \widetilde{p}(n)$$

für alle natürlichen Zahlen n.

(ii) Berechnen Sie $\widetilde{p}(X)$ für den Fall $p(X) = x^3$.

(Hinweis: Es genügt, den Fall $p(X) = x^h$ zu behandeln; betrachten Sie hierzu den Ausdruck $\frac{d^h}{dx^h}(\mathrm{e}^x + 1)^n$—via Induktion sieht man leicht, daß er einerseits in der Form $\sum_{\ell=0}^{h} q_\ell(n)\mathrm{e}^{\ell x}(\mathrm{e}^x + 1)^{n-\ell}$ mit geeigneten Polynomen $q_\ell(X)$ geschrieben werden kann, andererseits kann man ihn aber auch mit dem binomischen Lehrsatz berechnen.)

77. Finden Sie mindestens eine Funktion $f(x)$, sodaß

(i) $f'(x) = \frac{1}{x \log x}$

(ii) $f'(x) = \mathrm{e}^{\mathrm{e}^{\mathrm{e}^x}} \mathrm{e}^{\mathrm{e}^x} \mathrm{e}^x$

(iii) $f'(x) = f(x)x^r$.

78. Sei $\mathbb{R}(\cot)$ die Menge aller Funktionen f, die sich in der Gestalt $f(x) = r(\cot(x))$ schreiben lassen, wo r eine rationale Funktion ist. Zeigen Sie, daß $\mathbb{R}(\cot)$ unter Differentiation abgeschlossen ist.

79. Sei $p(X)$ ein normiertes Polynom vom Grad r mit reellen Nullstellen x_i $(1 \leq i \leq r)$. Zeigen Sie:

$$\prod_{\substack{1 \leq i,j \leq r \\ i \neq j}} (x_i - x_j) = \prod_{i=1}^{r} p'(x_i).$$

80. Berechnen Sie $\frac{f'(x)}{f(x)}$ für

(i) $f(x) = \prod_{r=1}^{n} (1 + \mathrm{e}^{rx})$

(ii) $f(x) = \prod_{r=1}^{n} (1 + \mathrm{e}^{rx})^{-1}$

(iii) $f(x) = \binom{x}{n}$.

81. Zeigen Sie, daß $\lim\limits_{\substack{x \to 0 \\ x > 0}} x^r \mathrm{e}^{-1/x} = 0$ für jede reelle Zahl r.

82. Sei f eine auf dem Intervall $(-\epsilon, 1 + \epsilon)$ zweimal differenzierbare Funktion; es gelte $f(0) = 0$, $f(1) = 1$, $f'(0) = f'(1) = 0$. Zeigen Sie, daß es ein x im Intervall $(0,1)$ gibt, sodaß $|f''(x)| \geq 4$.

(Ein Teilchen, welches in einer Zeiteinheit die Strecke 1 zurücklegt, erleidet zu irgendeinem Zeitpunkt mindestens die Beschleunigung 4 ! Hinweis: Ist $f(\frac{1}{2}) \geq \frac{1}{2}$, so gibt es ein x in $(0, \frac{1}{2})$ mit $f''(x) = 4$; ist $f(1) - f(\frac{1}{2}) \geq \frac{1}{2}$, so gibt ein x in $(\frac{1}{2}, 1)$ mit $f''(x) = -4$.)

83. Zeigen Sie: für alle $r \geq 1$ und $h > -1$ gilt $(1+h)^r \geq 1 + hr$.

(Hinweis: Untersuchen Sie $h \mapsto (1+h)^r - (1+hr)$ auf Monotonie.)

84. (i) Sei f eine auf $(a, +\infty)$ differenzierbare Funktion, sodaß sowohl $\lim\limits_{x \to \infty} f(x)$ als auch $\lim\limits_{x \to \infty} f'(x)$ existieren. Zeige: $\lim\limits_{x \to \infty} f'(x) = 0$.

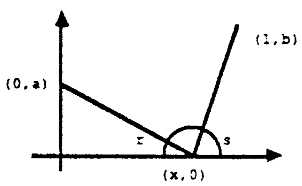
(ii) Finden Sie eine auf $(0, +\infty)$ differenzierbare Funktion, sodaß der Grenzwert $\lim\limits_{x \to \infty} f(x)$ existiert, aber nicht $\lim\limits_{x \to \infty} f'(x)$.

85. Sei f eine auf dem Intervall $(-a, +a)$ $(n+1)$–mal stetig differenzierbare Funktion und $p(X)$ ein Polynom vom Grad n, sodaß

$$|f(x) - p(x)| \leq M \left| x^{n+1} \right| \qquad \text{für alle } x \in (-a, +a) \qquad (M \text{ konstant})$$

gilt. Zeigen Sie, daß dann $p(X)$ das n–te Taylorpolynom von f um den Punkt 0 ist.

86. Ein Teilchen bewegt sich auf einer Geraden vom Punkt $(0, a)$ zum Punkt $(x, 0)$ und dann auf einer Geraden von $(x, 0)$ zum Punkt $(1, b)$ $(a, b > 0)$. Zeigen Sie, daß die zurückgelegte Strecke minimal ist, falls die Winkel r und s gleich sind.



87. (i) Sei $f\colon (a,b) \to (c,d)$ eine bijektive, differenzierbare Abbildung, und sei F eine Funktion auf (a,b), sodaß $F' = f$. Dann ist die Funktion

$$G(x) := xf^{-1}(x) - F(f^{-1}(x))$$

differenzierbar, und es gilt $G' = f^{-1}$.

(ii) Finden Sie Funktionen $G(x)$, sodaß

$$\text{(a)}\quad G'(x) = \log(x) \qquad \text{(b)}\quad G'(x) = \arcsin(x)$$
$$\text{(c)}\quad G'(x) = \arctan(x).$$

88. (i) Sei $f\colon (0,1) \to (0,1)$ eine bijektive, monoton steigende Funktion, sodaß $f^{-1} = f$. Dann gilt $f(x) = x$.

(ii) Konstruieren Sie unendlich viele bijektive, monoton fallende, unendlich oft differenzierbare Funktionen $f\colon (0,1) \to (0,1)$ mit der Eigenschaft $f^{-1} = f$.

89. Sei $f(x)$ eine auf dem Intervall $[a,b]$ stetige, im Intervall (a,b) differenzierbare Funktion, sodaß $|f'(x)| \leq M$ für alle $x \in (a,b)$ gilt. Zeigen Sie: $|f(b) - f(a)| \leq M(b-a)$.

90. (i) Sei $a_1 < \ldots < a_n$. Finden Sie das Minimum der Funktion $f(x) = \sum_{i=1}^{n} (x - a_i)^2$.

(ii) Finden Sie das Minimum der Funktion $f(x) = \sum_{i=1}^{n} |x - a_i|$.

91. Für reelle Zahlen p, q bezeichne V den Vektorraum aller auf $\mathbb{R}$ unendlich oft differenzierbaren Funktionen f, die der Gleichung

$$f'' + pf' + qf = 0$$

genügen. Zeigen Sie, daß die Abbildung

$$V \longrightarrow \mathbb{R}^2, \quad f \mapsto (f(0), f'(0))$$

eine Bijektion definiert (sie ist dann sogar ein *Isomorphismus von Vektorräumen*).

Hinweis: Zum Nachweis der Surjektivität betrachte man Funktionen der Gestalt $\mathrm{e}^{\lambda t}, t\mathrm{e}^{\lambda t}$, bzw. — falls $\lambda \notin \mathbb{R}$ — die Real-und Imaginärteile dieser Funktionen. Zum Nachweis der Injektivität genügt es zu zeigen, daß nur die identisch verschwindende Funktion in V der Bedingung $f(0) = f'(0) = 0$ genügt. Warum ? Zum Nachweis des letzteren betrachte man zu gegebener reeller Zahl c die Funktion $F_c(t) := \mathrm{e}^{-ct}(f(t)^2 + f'(t)^2)$; es ist $F_c(0) = 0$, $F_c(t) \geq 0$; für ein geeignetes c ist aber $F_c'(t) \leq 0$, und dies ist nur möglich, falls F_c identisch verschwindet.

92. Beweisen Sie das Additionstheorem des arcsin:

$$\arcsin x + \arcsin y = \arcsin(x\sqrt{1-y^2} + y\sqrt{1-x^2}).$$

93. Es ist $\frac{\pi}{2}$ die kleinste positive Nullstelle von $\cos(x)$.

(i) Zeigen Sie, daß $\tan(x) := \frac{\sin(x)}{\cos(x)}$ bei Einschränkung eine bijektive differenzierbare Funktion $(-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ definiert. Die Umkehrabbildung wird mit arctan bezeichnet.

(ii) Zeigen Sie, daß $\frac{\pi}{4} = \arctan(1)$.

(iii) Beweisen Sie das Additionstheorem für den Arcustangens:

$$\arctan x + \arctan y = \arctan \frac{x+y}{1-xy} \qquad (0 \le x \le 1, y < 1).$$

Hinweis: Beweisen Sie zunächst mittels der Eulerschen Formel das äquivalente Additiontheorem für den Tangens.

(iv) Zeigen Sie die *Formel von John Machin*:

$$\frac{\pi}{4} = 4 \arctan \frac{1}{5} - \arctan \frac{1}{239}.$$

94. (i) Zeigen Sie die folgende Formel für alle $m, p, q \ge 1,\ m^2 + 1 = pq$ gültige Formel:

$$\arctan \frac{1}{m} = \arctan \frac{1}{m+p} + \arctan \frac{1}{m+q}.$$

Hinweis: Diese Formel geht im wesentlichen auf Lewis Carroll zurück.

(ii) Zeigen Sie die Formel

$$\frac{\pi}{4} = 2 \arctan \frac{1}{3} + \arctan \frac{1}{7}.$$

(Eine Formel von Euler zur Berechnung von π.)

95. (i) Seien r_k reelle Zahlen, n_k ganze Zahlen $(1 \le k \le n)$. Man beweise, daß die beiden folgenden Aussagen äquivalent sind:

(a) $\displaystyle\sum_{k=1}^{n} n_k \arctan r_k = \pi \ell \quad$ mit einem $\ell \in \mathbb{Z}$

.

(b) $\displaystyle\prod_{k=1}^{n}(1+ir_k)^{n_k} \in \mathbb{R}$.

Hinweis: $x = \ell\pi$ $(\ell \in \mathbb{Z}) \leftrightarrow \mathrm{e}^{ix} = \pm 1$.

(ii) Beweisen Sie mit (i) nochmals die Formeln:

$$\frac{\pi}{4} = 4\arctan\frac{1}{5} - \arctan\frac{1}{239}, \qquad \frac{\pi}{4} = 2\arctan\frac{1}{3} + \arctan\frac{1}{7}.$$

96. (i) Zeigen Sie, daß die Exponentialfunktion transzendent über dem Körper der rationalen Funktionen ist, d.h. daß es keine rationalen Funktionen $f_0, f_1, \ldots, f_n$ gibt, sodaß

$$\mathrm{e}^{nx} + \mathrm{e}^{(n-1)x} f_{n-1}(x) + \ldots + \mathrm{e}^{x} f_1(x) + f_0(x) = 0$$

für alle x gilt, für die die f_k definiert sind.

(ii) Zeigen Sie das gleiche für den Sinus.

TAYLOR-ENTWICKLUNG

97. Man berechne die Taylorentwicklung von $\log\left(\frac{1-x}{1+x}\right)$ um $x = 0$.

98. Setzt man in der Machinschen Formel

$$\frac{\pi}{4} = 4\arctan\frac{1}{5} - \arctan\frac{1}{239}.$$

die Taylorentwicklung

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - + \cdots \qquad (-1 < x \leq 1)$$

ein, so erhält man eine unendliche Reihe, die gegen π konvergiert und deren n-te Partialsumme s_n eine effektiv berechenbare, rationale Zahl ist. Finden Sie eine Fehlerabschätzung für $|\pi - s_n|$, und bestimmen Sie anhand Ihrer Abschätzung ein n_0, sodaß die ersten 100 Stellen der Dezimalentwicklung von s_{n_0} mit den ersten 100 Stellen der Dezimalentwicklung von π übereinstimmen.

99. (i) Zeigen Sie $\pi = 48\arctan\frac{1}{18} + 32\arctan\frac{1}{57} - 20\arctan\frac{1}{239}$.

(ii) Setzt man in diese Formel die Taylorentwicklung des arctan um $x = 0$ ein, so erhält man eine Darstellung von π als unendliche Reihe. Bestimmen Sie ein n, sodaß die n-te Partialsumme dieser Reihe sich von π um weniger als 10^{-100} unterscheidet. Vergleichen Sie dieses Ergebnis mit der entsprechenden Abschätzung für die Machinsche Formel (siehe oben).

(iii) Berechnen Sie die ersten 6 Stellen von π.

100. (i) Zeigen Sie, daß die Zuordnung $x \mapsto \tanh x := \frac{\sinh x}{\cosh x}$ eine bijektive, unendlich oft differenzierbare Abbildung $\mathbb{R} \to (-1,1)$ definiert. Die Umkehrfunktion wird mit arctanh bezeichnet.

(ii) Berechnen Sie die Taylorentwicklung von $\operatorname{arctanh} x$ bei $x = 0$, und zeigen Sie, daß sie für jedes x aus $(-1,1)$ gegen $\operatorname{arctanh} x$ konvergiert.

Hinweis: Man kann die Funktion $\operatorname{arctanh} x$ mittels des Logarithmus ausdrücken.

101. Berechnen Sie die Taylorreihen der folgenden Funktionen um $x = 0$, bestimmen Sie den Konvergenzradius dieser Taylorreihen, und zeigen Sie, daß diese für $|x| <$ Konvergenzradius gegen die jeweilige Funktion konvergieren:

(i) $\arctan x$ (ii) $\operatorname{arccot} x$

(ii) $\operatorname{arcsinh} x$ (iv) $\operatorname{arcsec}(x^2+2)$.

($\operatorname{arcsec} x$ ist die Umkehrfunktion der Einschränkung von $\sec x = \frac{1}{\cos x}$ auf das Intervall $(0, \frac{\pi}{2})$.)

102. Beweisen Sie die folgenden Aussage: Seien f, g in einem offenen Intervall I definierte, differenzierbare Funktionen, und sei $a \in \mathbb{R} \cup \{\pm\infty\}$ Randpunkt von I (d.h. es ist $I = (b,a)$ oder $I = (a,b)$ mit einem geeigneten b). Es gelte

$$\lim_{\substack{x \to a \\ x \in I}} f(x) = 0, \qquad \lim_{\substack{x \to a \\ x \in I}} g(x) = 0,$$

und $\lim\limits_{\substack{x \to a \\ x \in I}} \dfrac{f'(x)}{g'(x)}$ existiere. Dann existiert auch $\lim\limits_{\substack{x \to a \\ x \in I}} \dfrac{f(x)}{g(x)}$, und es gilt

$$\lim_{\substack{x \to a \\ x \in I}} \frac{f(x)}{g(x)} = \lim_{\substack{x \to a \\ x \in I}} \frac{f'(x)}{g'(x)}.$$

(Dies ist eine weitere von zahlreichen Varianten von *L'Hôpital's Regel*. Hinweis: benutzen Sie den Cauchyschen Mittelwertsatz.)

103. Berechnen Sie die folgenden Grenzwerte:

(i) $\lim\limits_{x \to 0} \dfrac{x}{\tan x}$

(ii) $\lim\limits_{x \to 0} \dfrac{\cos^2 x - 1}{\tan x}$

(Hinweis: Regel von L'Hôpital.)

104. Berechnen Sie die folgenden Grenzwerte:

(i) $\lim_{x \to \frac{\pi}{4}} (\cot x)^{\tan 2x}$

(ii) $\lim_{x \to 0} (1 + \sin x)^{\frac{1}{\sin x}}$

(Hinweis: Regel von L'Hôpital.)

POTENZREIHEN

105. Berechnen Sie die Konvergenzradien der folgenden Potenzreihen:

(i) $\sum_{n=1}^{\infty} n^\alpha x^n \ (\alpha \in \mathbb{R})$

(ii) $\sum_{n=0}^{\infty} \sin(n) x^n$

(iii) $\sum_{n=1}^{\infty} \log(n) x^n$

(iv) $\sum_{n=0}^{\infty} \frac{x^n}{(4+(-1)^n)^{3n}}$.

106. Entwickeln Sie $\sqrt{x^2 - x + 1}$ in eine Potenzreihe um $x = 1$.

107. (i) Zeigen Sie, daß $\sum_{n=0}^{\infty} F_n x^n$ ($F_n = n$-te Fibonacci-Zahl) eine rationale Funktion ist.

(ii) Entwickeln Sie diese rationale Funktion in eine Taylorreihe um $x = 0$, indem Sie sie als Summe von geometrischen Reihen schreiben.

(iii) Folgern Sie eine explizite Formel für F_n.

108. (i) Für eine nach oben beschränkte Zahlenfolge $\{a_n\}$ bezeichnet man als *Limes superior* — in Zeichen $\limsup_{n\to\infty} a_n$ — die Zahl $\sup M$, wo M die Menge der Häufungspunkte der Folge $\{a_n\}$ bedeutet. Zeigen Sie, daß diese Definition sinnvoll ist, d.h. daß die Menge M tatsächlich nach oben beschränkt ist (wir benutzen hier die Konvention $\sup \emptyset = -\infty$). Für eine nach oben unbeschränkte Folge setzt man noch $\limsup_{n\to\infty} = \infty$.

(ii) Zeigen Sie für den Konvergenzradius einer gegebenen Potenzreihe $\sum_n a_n x^n$ die *Formel von Hadamard*:

$$\sup \left\{ r \geq 0 \mid \sum |a_n| r^n \text{ ist konvergent} \right\} = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}}.$$

(Hierbei ist die rechte Seite gegebenenfalls im Sinne von $\frac{1}{0} = \infty$ und $\frac{1}{\infty} = 0$ zu interpretieren.)

109. Sei (a_n) eine nach oben beschränkte Zahlenfolge. Zeigen Sie

$$\limsup_{n\to\infty} \{a_k | k \geq n\} = \limsup_{n\to\infty} a_n.$$

110. (i) Bestimmen Sie den Konvergenzradius R der Reihe

$$1 + \sum_{k=1}^{\infty} \left(\frac{1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2k-1)}{2^k \cdot k!} \right)^2 t^k.$$

(ii) Zeigen Sie, daß die durch diese Potenzreihe in dem Intervall $(-R, R)$ definierte Funktion f der Differentialgleichung

$$f'' + \frac{1-2t}{t(1-t)} f' - \frac{1}{4t(1-t)} f = 0$$

genügt, und daß jede weitere bei $t = 0$ analytische Lösung dieser Differentialgleichung von der Gestalt $c \cdot f$ mit einer geeigneten Konstanten c ist.

GLEICHMÄSSIGE STETIGKEIT UND GLEICHMÄSSIGE KONVERGENZ

111. Ein abgeschlossenes Intervall I und darin Punkte x_k $(0 \leq k \leq n)$ seien vorgegeben. Konstruieren Sie eine auf diesem Intervall punktweise konvergente Folge von stetigen Funktionen, sodaß die Grenzfunktion auf $I \setminus \{x_0, x_1, \ldots, x_n\}$ stetig und in den Punkten x_k unstetig ist.

112. Es bezeichne $I \subset \mathbb{R}$ ein abgeschlossenes Intervall und $\mathcal{C}(I)$ die Menge der stetigen Funktionen auf I. Zeigen Sie nacheinander die folgenden Behauptungen:

(i) Zu jedem $f \in \mathcal{C}(I)$ existiert eine Folge von auf I stückweise linearen, stetigen Funktion l_n $(n = 1, 2, 3, \ldots)$, die auf I gleichmäßig gegen f konvergiert. — Hierbei heißt eine stetige Funktion l auf I stückweise linear, falls es eine Zerlegung $I = \cup_{k=1}^{r} I_k$ von I in paarweise disjunkte Teilintervalle I_k gibt, sodaß zu jedem k Zahlen a, b existieren, sodaß $l(x) = ax + b$ für alle $x \in I_k$ ist.

(Hinweis: Jedes Element in $\mathcal{C}(I)$ ist gleichmäßig stetig auf I.)

(ii) Zu jeder auf I stetigen, stückweise linearen Funktion l kann man Zahlen a_k und x_k $(0 \leq k \leq n)$ finden, sodaß

$$l(x) = \sum_{k=0}^{n} a_k |x - x_k| \qquad \text{für alle } x \in I.$$

(iii) Zu jedem $\epsilon > 0$ existiert ein Polynom $p(x)$, sodaß

$$||x| - p(x)| < \epsilon \quad \text{für alle } x \in I.$$

(Hinweis: Man kann die Binomialreihe und die Formel $|x| = \sqrt{x^2}$ benutzen.)

(iv) Zu jedem $f \in \mathcal{C}(I)$ existiert eine Folge von Polynomen p_n ($n = 1, 2, 3, \ldots$), die auf I gleichmäßig gegen f konvergiert.

113. Zeigen Sie, daß für auf einer Menge M definierte Funktionen f, f_n ($n = 0, 1, 2, \ldots$) die folgenden beiden Aussagen äquivalent sind:

(A) Die Folge f_n konvergiert auf M gleichmäßig gegen f.

(B) Es gibt ein n_0, sodaß die Funktionen $f_n - f$ für $n \geq n_0$ beschränkt sind, und die Folge $\sup |f - f_n|(M)$ ($n \geq n_0$) ist eine Nullfolge.

114. Zeigen Sie, daß eine Reihe von Funktionen $\sum_n f_n(x)$ auf einer Menge M gleichmäßig konvergiert, falls es eine konvergente Reihe $\sum_n c_n$ gibt, sodaß $|f_n(x)| \leq c_n$ für alle n und alle $x \in M$ gilt. (Da die Grenzfunktion einer gleichmäßig konvergenten Folge von stetigen Funktionen wieder stetig ist, impliziert die eben bewiesene Tatsache insbesondere, daß eine Potenzreihe im Innern ihres Konvergenzkreises eine stetige Funktion darstellt.)

INTEGRAL

115. Es sei $f(x) = 1$ für $x \in I \cap \mathbb{Q}$ und $f(x) = 0$ für $x \in I \setminus \mathbb{Q}$, wobei $I = [0, 1]$. Zeigen Sie:

(i) Es gibt eine Folge von Treppenfunktionen t_n, die auf I punktweise gegen f konvergiert.

(ii) Es gibt dagegen keine Folge von Treppenfunktionen, die auf I gleichmäßig gegen f konvergiert.

116. Berechnen Sie $\int_0^a \cos x \, dx$ mittels Treppenfunktionen, d.h. finden Sie eine geeignete Folge von Treppenfunktionen t_n, die auf dem Intervall $[0, a]$ gleichmäßig gegen $\cos x$ konvergiert, und berechnen Sie direkt und ohne Benutzung des 'Hauptsatzes' den Grenzwert der Folge $\int_0^a t_n(x) \, dx$.

117. Sei f eine auf einem abgeschlossenen Intervall I definierte, streng monotone Funktion. Dann gibt es eine Folge von Treppenfunktionen, die auf I gleichmäßig gegen f konvergiert.

118. Sei f eine auf dem Intervall $[0, a]$ definierte, streng monoton steigende Funktion, $f(0) = 0,\ f(a) = b$. Dann gilt

$$\int_0^a f(x)\,dx + \int_0^b f^{-1}(y)\,dy = a \cdot b.$$

(Hinweis: Veranschaulichen Sie sich die Behauptung an einer Skizze!)