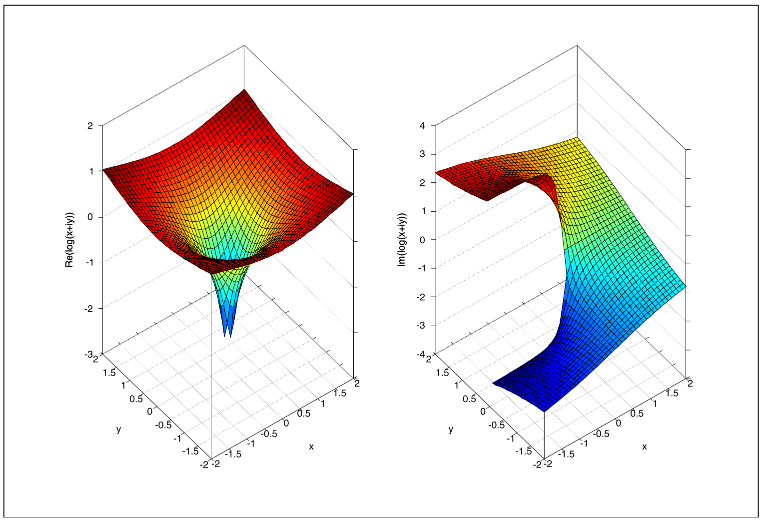
Analysis (BA-INF 022)

Dr. Michael Welter



MATHEMATISCHES INSTITUT

RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN

3. Mai 2023

Inhaltsverzeichnis

Vorwort

Die erste Version des Skripts in geTEXter Form wurde im Sommersemester 2009 von Herrn Christian Kipping erstellt, dem ich für seine sehr gute Arbeit an dieser Stelle von ganzem Herzen danke. Ich danke auch der Fachschaft Informatik, die damals die aus Studienbeitragsmitteln finanzierte Erstellung eines Skripts angeregt hatte. Seitdem hat sich das Skript enorm weiterentwickelt.

Die Vorlesung *Analysis* schließt im Bachelorstudiengang Informatik an die Vorlesung *Logik und diskrete Strukturen* an. Die dort behandelten Inhalte werden hier natürlich stillschweigend vorausgesetzt. Einige grundlegende Definitionen und Ergebnisse sind im Anhang zusammengefasst. Es sei noch darauf hingewiesen, dass die benutzten Notationen leicht von denen in der *Logik und diskrete Strukturen* abweichen können. So bedeutet hier z.B. $A \subset B$, dass A eine Teilmenge von B ist; dabei ist der Fall $A = B$ nicht ausgeschlossen.

Für Fehlermeldungen, Anregungen und Kommentare bin ich jederzeit dankbar. Bitte senden Sie diese an welter@uni-bonn.de.

1. Die reellen und komplexen Zahlen

Wenn man ein Gebäude errichten will, so braucht man ein solides Fundament, das dieses Gebäude trägt. Ebenso ist es auch in der Mathematik. In der Mathematik muss man alle neuen Aussagen beweisen, das heißt aus bereits als richtig erkannten Aussagen ableiten. Aber irgendwo muss man anfangen. Gewisse Aussagen muss man einfach als richtig hinnehmen. Dies sind sogenannte *Axiome*.

Unser Fundament der Analysis sind die *reellen Zahlen*. Wir nehmen diese als gegeben hin. Man sagt auch, dass wir die reellen Zahlen axiomatisch beschreiben. Das heißt, wir setzen voraus, dass es eine Menge von Objekten („Zahlen") gibt, in denen gewisse Aussagen („Axiome") gelten. Wir bezeichnen diese Menge dann als die Menge der reellen Zahlen $\mathbb{R}$.

Es gibt auch noch andere Zugänge zu den reellen Zahlen. Bei diesen Zugängen wählt man „kleineres"Axiomensystem und leitet dann die Existenz der reellen Zahlen hieraus ab, indem man eine Menge mit den gewünschten Eigenschaften konstruiert.

Wir wählen aber den axiomatischen Zugang.

Die reellen Zahlen lassen sich durch drei Gruppen von Axiomen charakterisieren:

(i) die Körperaxiome,

(ii) die Anordnungsaxiome,

(iii) das Vollständigkeitsaxiom.

Während die ersten beiden Gruppen von Axiomen auch für die rationalen Zahlen gelten, stellt das Vollständigkeitsaxiom den grundlegenden Unterschied zwischen diesen und den reellen Zahlen dar. Ihm werden wir später einen eigenen Abschnitt widmen.

Wir betrachten die dem Leser sicherlich bekannten Mengen der natürlichen Zahlen $\mathbb{N} := \{1, 2, 3, \ldots\}$, der ganzen Zahlen $\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ und der rationalen Zahlen $\mathbb{Q} := \{\frac{p}{q} | p, q \in \mathbb{Z}, q \neq 0\}$ als Teilmengen der reellen Zahlen. In einem kurzen Exkurs werden wir klären, wieso man das machen kann.

Ebenfalls als bekannt setzen wir die in Anhang A zusammengefassten Grundlagen, wie etwa das Prinzip der vollständigen Induktion, mengentheoretische Schreibweisen und algebraische Grundstrukturen, als bekannt voraus.

Wir werden an der ein oder anderen Stelle ohne Beweis auf Ergebnisse der parallel stattfindenden Linearen Algebra zurückgreifen und auch manchmal ohne Definition die Sprache der Linearen Algebra benutzen; etwa wenn wir sagen, dass eine Menge von Funktionen einen Vektorraum bilden.

Nun wollen wir uns die reellen Zahlen genauer anschauen.

1.1. Die reellen Zahlen als archimedisch angeordneter Körper

Zunächst einmal wollen wir mit den reellen Zahlen natürlich rechnen können. Die üblichen Rechenoperationen (Addition, Subtraktion, Multiplikation und Division) sollen also auf den reellen Zahlen erklärt sein. Wir erinnern an die Definition der algebraischen Struktur *Körper*.

Definition 1 (Körper)**:** Es sei K eine Menge und $+ : K \times K \to K$ und $\cdot : K \times K \to K$ zwei Abbildungen. Ein Tripel $(K, +, \cdot)$ heißt Körper (mit Nullelement 0_K und Einselement 1_K), wenn die folgenden Bedingungen erfüllt sind:

- *Abgeschlossenheit:*
$$\begin{aligned} &\forall a, b \in K: \; a + b \in K, \\ &\forall a, b \in K \setminus \{0_K\}: \; a \cdot b \in K \setminus \{0_K\}. \end{aligned}$$

- *Assoziativität:*
$$\begin{aligned} &\forall a, b, c \in K: \; (a + b) + c = a + (b + c), \\ &\forall a, b, c \in K: \; (a \cdot b) \cdot c = a \cdot (b \cdot c). \end{aligned}$$

- *Kommutativität:*
$$\forall a, b \in K : a + b = b + a \text{ und } a \cdot b = b \cdot a$$

- Es gibt in K eindeutig bestimmte, *neutrale Elemente* $0_K \neq 1_K$ mit
$$\begin{aligned} &\forall a \in K \quad : a + 0_K = 0_K + a = a, \\ &\forall a \in K \setminus \{0_K\} \; : a \cdot 1_K = 1_K \cdot a = a. \end{aligned}$$

- Es gibt eindeutig bestimmte *inverse Elemente*, dass heißt
$$\begin{aligned} &\forall a \in K \; \exists -a \in K: \; a + (-a) = 0_K, \\ &\forall a \in K \setminus \{0_K\} \; \exists a^{-1} \in K: \; a \cdot a^{-1} = 1_K. \end{aligned}$$

- *Distributivität:*

$$\forall a, b, c \in K: \ a \cdot (b + c) = (a \cdot b) + (a \cdot c).$$

Axiom 1: $(\mathbb{R}, +, \cdot)$ ist ein Körper.

BEMERKUNG: Natürlich bezeichnen wir das neutrale Element der Addition wie üblich mit 0 und das der Multiplikation mit 1.

Problem 1 (Bruchrechnung)**:** *(i) Es seien $a, b \in \mathbb{R}$ und $b \neq 0$. Zeigen Sie, dass die Gleichung $bx = a$ eindeutig lösbar ist.*

(ii) Wir definieren für $a, b \in \mathbb{R}, b \neq 0$,

$$\frac{a}{b} := a \cdot b^{-1},$$

wobei b^{-1} natürlich das multiplikative Inverse von b bezeichnet. Beweisen Sie die folgenden Rechenregeln der Bruchrechnung ($a, b, c, d \in \mathbb{R}, b \neq 0, d \neq 0$):

- $\frac{a}{b} = \frac{c}{d}$ *genau dann, wenn* $ad = bc$*,*
- $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$*,*
- $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$*,*
- $$\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc},$$
  wobei auch $c \neq 0$ sei.

Hinweis: Es gilt stets $(ab)^{-1} = b^{-1}a^{-1}$.

Bemerkung: In dieser Aufgabe geht es darum, dass Sie sich bei jedem Schritt klar machen, welche Eigenschaften eines Körpers in die Argumentation einfließen.

Unsere nächste Forderung an die reellen Zahlen soll sein, dass man die reellen Zahlen *anordnen* kann. Um diese Forderung präzise formulieren zu können, benötigen wir folgende Definition.

Definition 2 (Ordnungsrelation)**:** Es sei $(K, +, \cdot)$ ein Körper mit Nullelement 0_K. Eine Relation $\leq$ auf K heißt *Ordnungsrelation*, wenn für alle $a, b, c \in K$ die folgenden Eigenschaften gelten:

(A1) $a \leq b \vee b \leq a$ (Vergleichbarkeit)

(A2) $(a \leq b \wedge b \leq a) \Rightarrow a = b$ (Identitätseigenschaft)

(A3) $(a \leq b \wedge b \leq c) \Rightarrow a \leq c$ (Transitivität)

(A4) $a \leq b \Rightarrow a + c \leq b + c$ (Monotonie)

(A5) $(0_K \leq a \wedge 0_K \leq b) \Rightarrow 0_K \leq a \cdot b$ (Monotonie)

Ein Element $a \in K$ heißt *positiv*, falls $0_K \leq a$ und $a \neq 0_K$ gilt, und a heißt *negativ*, falls $a \leq 0_K$ und $a \neq 0_K$ gilt.

BEMERKUNG: Aus der Ordnungsrelation $\leq$ lassen sich die anderen Ordnungsbeziehungen wie $<$, $\geq$ oder $>$ ableiten. Man setzt

$$\begin{aligned} a < b &:\Leftrightarrow\ a \leq b\ \wedge\ a \neq b \\ a > b &:\Leftrightarrow\ b < a \\ a \geq b &:\Leftrightarrow\ b \leq a \\ a \leq b \leq c &:\Leftrightarrow\ a \leq b\ \wedge\ b \leq c \\ a \leq b < c &:\Leftrightarrow\ a \leq b\ \wedge b < c \\ &\ldots . \end{aligned}$$

Definition 3 (angeordneter Körper)**:** Wir sagen, dass ein Körper K *angeordnet* werden kann, wenn auf K eine Ordnungsrelation $\leq$ definiert werden kann. Das Tupel $(K, \leq)$ wird dann als ein *angeordneter Körper* bezeichnet.

Axiom 2: Der Körper $\mathbb{R}$ kann angeordnet werden.

In den folgenden Hilfssätzen fassen wir einige Rechenregeln für den Umgang mit der Ordnungsrelation zusammen. Wir formulieren diese Aussagen nur für die reellen Zahlen, obwohl sie natürlich in beliebigen angeordneten Körpern gelten.

Lemma 4: Für jede reelle Zahl a gilt:

(i) $a \geq 0 \Leftrightarrow -a \leq 0$

(ii) $a \neq 0 \Rightarrow a \cdot a > 0$

(iii) $a \cdot a \geq 0$

(iv) $a > 0 \Rightarrow a^{-1} > 0$

BEWEIS:

(i) Aus $a \geq 0$ folgt mit $(A4)$

$$0 = a + (-a) \geq 0 + (-a) = -a.$$

Damit haben wir die eine Richtung der Äquivalenz gezeigt. Mit $0 \geq -a$ und der Eigenschaft $(A4)$ erhält man aber auch

$$a = 0 + a \geq -a + a = 0.$$

(ii) Ist $a \neq 0$, so ist $a \cdot a \neq 0$, da $\mathbb{R}$ ein Körper ist. Falls $a \geq 0$ ist, so folgt aus Eigenschaft (A5) $a \cdot a \geq 0$. Da $a = 0$ ausgeschlossen wurde, gilt $a \cdot a > 0$. Falls $a \leq 0$ ists, so folgt wegen (i) die Ungleichung $(-a) \geq 0$. Ebenfalls wieder wegen (A5) erhalten wir $(-a) \cdot (-a) \geq 0$ und damit auch $a \cdot a > 0$.

(iii) Für $a = 0$ ist auch $a \cdot a = 0$. Ist aber $a \neq 0$, so haben wir in (ii) gesehen, dass $a \cdot a > 0$ ist. In jedem Fall haben wir also $a \cdot a \geq 0$.

(iv) Es ist $a^{-1} \cdot a^{-1} > 0$. Multipliziert man diese Gleichung mit a, so erhält man die gewünschte Ungleichung.

□

BEMERKUNG: Ist a eine beliebige reelle Zahl, so schreiben wir a^2 anstelle von $a \cdot a$.

Lemma 5: Für beliebige reelle Zahlen a, b, c gelten die folgenden Aussagen:

(i) Es trifft genau eine der Aussagen zu:

a) $a = b$

b) $a < b$

c) $a > b$

(ii) $(a \leq b \wedge b < c) \Rightarrow a < c$

(iii) $(0 < a \wedge 0 < b) \Rightarrow 0 < a \cdot b$

(iv) $a \leq b \wedge c \geq 0 \Rightarrow a \cdot c \leq b \cdot c$

(v) $a \leq b \wedge c \leq 0 \Rightarrow a \cdot c \geq b \cdot c$

Problem 2: *Beweisen Sie das vorstehende Lemma.*

Nach Lemma 4 ist $1 = 1^2 \geq 0$ und somit $1 > 0$. Addieren wir auf beide Seiten dieser Ungleichung eine 1, so erhalten wir $1 + 1 > 1$. Addieren wir nochmals eine 1 auf beide Seiten der Ungleichung, so haben wir $1 + 1 + 1 > 1 + 1$. Wir können diesen Prozess beliebig oft wiederholen und erhalten also

Satz 6: $\mathbb{R}$ hat unendlich viele Elemente.

BEMERKUNG: Diese Argumentation lässt sich auf jeden angeordneten Körper übertragen. Endliche Körper, wie z.B. $\mathbb{Z}_p$, können folglich nicht angeordnet werden.

Lemma 7 (Monotonie der Quadratfunktion)**:** Sind a, b reelle Zahlen und ist $a \geq 0$, so gilt

$$a < b \Rightarrow a^2 < b^2.$$

BEWEIS: Für $a = 0$ folgt die Behauptung schon aus Lemma 4. Es sei also $0 < a$ und $a < b$ vorausgesetzt. Dann ist $a \cdot a \leq a \cdot b$ nach Lemma 5 (iv). Aufgrund der Transitivität ergibt sich aus $0 < a$ und $a < b$ auch $0 < b$ und somit ergibt sich wieder aus Lemma 5 (iv) aus $a < b$ auch die Ungleichung $a \cdot b \leq b \cdot b$.
Wäre hier $ab = bb$, so gälte $(a - b) \cdot b = 0$, also $a = b$ oder $b = 0$, was jedoch beides ausgeschlossen wurde. Also haben wir $a \cdot a \leq a \cdot b$ und $a \cdot b < b \cdot b$. Zusammen ergibt das $a \cdot a < b \cdot b$. □

Wir nehmen das vorstehende Lemma zum Anlass, um Potenzen von reellen Zahlen mit ganzzahligen Exponenten zu definieren.

Definition 8 (Potenzen mit ganzzahligen Exponenten)**:** Für eine beliebige reelle Zahl a definieren wir rekursiv:

$$a^0 := 1$$
$$a^{n+1} := a^n \cdot a$$

für $n \in \mathbb{N}_0$. Für $n \in \mathbb{N}$ und $a \neq 0$ setzen wir außerdem

$$a^{-n} := (a^n)^{-1},$$

a^{-n} ist also das multiplikative Inverse von a^n.

BEMERKUNG: Wir haben in der vorstehenden Definition Potenzen a^n mit Exponenten $n \in \mathbb{Z}$ definiert. Dies legt die Frage nahe, wie man sinnvollerweise Potenzen a^r mit Exponenten $r \in \mathbb{Q}$ oder Potenzen a^x mit Exponenten $x \in \mathbb{R}$ definieren sollte. Sinnvoll ist eine solche Erweiterung der Definition sicherlich nur dann, wenn die erweiterte Definition für Exponenten $n \in \mathbb{Z}$ das selbe Ergebnis liefert wie die vorstehende Definition.

Problem 3 (Binomischer Lehrsatz)**:** *Zeigen Sie, dass für alle $a, b \in \mathbb{R}$ gilt:*

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

***Erinnerung (aus Logik und diskrete Strukturen)**:*
Für $n, k \in \mathbb{N}_0$ und $k \leq n$ definieren wir den Binomialkoeffizienten

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

wobei $n! := 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n$ die n-te Fakultät *bezeichnet. Es gelten die folgenden Regeln:*

(i) $\binom{m}{k} = \binom{m}{m-k}$.

(ii) $\binom{m+1}{k+1} = \binom{m}{k+1} + \binom{m}{k}$.

(iii) $\binom{m}{k} \in \mathbb{N}$.

Problem 4 (Verallgemeinerte 3. Binomische Formel)**:** *Zeigen Sie, dass für alle $q \in \mathbb{R}$ gilt:*

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

Folgern Sie hieraus für alle $a, b \in \mathbb{R}$

$$a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^{n-1-k} b^k.$$

Satz 9 (Bernoulli-Ungleichung)**:** Für jede reelle Zahl $x > -1$ und jedes $n \in \mathbb{N}_0$ gilt

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

BEWEIS: Wir beweisen die Bernoulli-Ungleichung durch vollständige Induktion nach n. Es sei x eine reelle Zahl mit $x > -1$.

Die *Induktionsverankerung* nehmen wir bei $n = 0$ vor. Es gilt $(1+x)^0 = 1 = 1+0\cdot x$.

Nun folgt der *Induktionsschritt*. Als *Induktionsannahme* gehen wir davon aus, dass die in der Bernoulli-Ungleichung behauptet Ungleichung $(1+x)^n \geq 1+n\cdot x$ für ein $n \in \mathbb{N}_0$ gilt. Wir müssen nun zeigen, dass die Ungleichung auch für $n+1$ gilt. Wir haben

$$\begin{aligned}(1+x)^{n+1} &= (1+x)^n(1+x)\\ &\geq (1+n\cdot x)(1+x)\\ &= 1+n\cdot x+x+(n\cdot x)\cdot x\\ &= 1+(n+1)\cdot x+n\cdot x^2\\ &\geq 1+(n+1)\cdot x.\end{aligned}$$

Beim Übergang von der ersten zur zweiten Zeile haben wir in der vorstehenden Ungleichungskette die Induktionsannahme benutzt.

□

Definition 10 (Betrag)**:** Für beliebiges $a \in \mathbb{R}$ setzen wir

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

$|a|$ heißt der *Betrag von* a.

Auf die selbe Weise lässt sich in jedem angeordneten Körper eine Betragsfunktion definieren und der folgende Satz wäre dann auch gültig.

Satz 11: Für alle $a, b \in \mathbb{R}$ gilt:

(i) $|a| \geq 0$ und $(|a| = 0 \Leftrightarrow a = 0)$ (Definitheit)

(ii) $|a\cdot b| = |a|\cdot|b|$ (Homogenität)

(iii) $|a+b| \leq |a|+|b|$ (Dreiecksungleichung)

(iv) $||a|-|b|| \leq |a-b|$

Problem 5: *Beweisen Sie die Teile (i) und (ii).*

BEWEIS: Bevor wir die beiden verbleibenden Teile (iii) und (iv) beweisen, halten wir noch zwei offensichtliche Beobachtungen zur Betragsfunktion fest. Zum einen gilt für jedes Element a in einem angeordneten Körper sowohl $|a| \geq a$ als auch $|a| \geq -a$. Zum anderen gilt für jedes Element a in einem angeordneten Körper $|a| = |-a|$.

(iii) Nach der eben gemachten Vorbemerkung gilt also sowohl $|a| + |b| \geq a + b$ als auch $|a| + |b| \geq (-a) + (-b) = -(a+b)$. Dann muss $|a| + |b| \geq |a+b|$ sein, da $|a+b|$ entweder gleich $a+b$ oder gleich $-(a+b)$ ist.

(iv) Mit (iii) erhalten wir $|a| = |a-b+b| \leq |a-b| + |b|$ und somit $|a| - |b| \leq |a-b|$. Ebenso gilt $|b| = |b-a+a| \leq |b-a| + |a|$, also $-(|a| - |b|) = |b| - |a| \leq |a-b|$. Aus diesen beiden Ungleichungen folgt mit dem gleichen Argument wie bei Teil (iii) die Behauptung.

$\square$

Definition 12 (Normierter Körper / Norm)**:** Ein Körper K, auf welchem eine Abbildung

$$\begin{aligned} |\cdot| : K &\to \mathbb{R} \\ x &\mapsto |x| \end{aligned}$$

definiert ist, so dass die Eigenschaften (i), (ii), (iii) aus Satz 11 erfüllt sind, heißt *normierter (oder bewerteter) Körper* und die Abbildung $|\cdot|$ wird *Norm* genannt.

BEMERKUNG: Satz 11 zeigt, dass jeder angeordnete Körper ein normierter Körper ist. Jedoch gibt es auch normierte Körper, die nicht angeordnet werden können. Wir werden bald mit den komplexen Zahlen einen solchen Körper konstruieren.

Problem 6 (Normen auf $\mathbb{R}$?)**:** *Prüfen Sie für die folgenden beiden Abbildungen $|\cdot|_i : \mathbb{R} \to \mathbb{R}$, $(i = 1, 2)$, ob diese eine Norm auf $\mathbb{R}$ sind, wenn wir für $a \in \mathbb{R}$ setzen:*

(i)

$$|a|_1 := \begin{cases} 0 & , \text{ falls } a = 0 \\ 1 & , \text{ falls } a \neq 0 \end{cases}.$$

(ii)

$$|a|_2 := \frac{|a|}{1 + |a|},$$

wobei $|\cdot|$ den gewöhnlichen Betrag auf $\mathbb{R}$ bezeichnet.

Exkurs: $\mathbb{N}, \mathbb{Z}$ und $\mathbb{Q}$

In diesem Exkurs wollen wir uns kurz überlegen, wie man die Mengen $\mathbb{N}, \mathbb{Z}$ und $\mathbb{Q}$ in den reellen Zahlen wiederfinden kann.

Definition 13: Wir nennen eine Teilmenge $A \subset \mathbb{R}$ eine *induktive Menge*, falls sie die folgenden beiden Eigenschaften hat:

(i) $1 \in A$,

(ii) Liegt $a \in A$, so ist auch $a + 1 \in A$

Als wir uns überlegt haben, dass $\mathbb{R}$ unendliche viele Elemente hat, haben wir dazu die Zahlen $1, 1+1, 1+1+1, \ldots$ betrachtet. Wir wollen nun wie üblich $2 := 1+1, 3 := 1+1+1, \ldots$ schreiben. Die Menge aller Zahlen, die man durch wiederholtes Addieren der 1 erhält, ist offensichtlich induktiv. Wir wollen diese Menge für den Moment mit N bezeichnen. N ist bei weitem nicht die einzige induktive Teilmenge der reellen Zahlen. Auch $\mathbb{R}$ selber ist induktiv. Und auch $\mathbb{R}_+ := \{x \in \mathbb{R} | x > 0\}$ ist induktiv. Offensichtlich gilt aber für jede induktive Teilmenge $A \subset \mathbb{R}$, dass $N \subset A$ ist.

Wir definieren nun

Definition 14:

$$\mathbb{N} := \bigcap \{A \subset \mathbb{R} | A \text{ ist eine induktive Menge}\}.$$

Problem 7: *Überlegen Sie sich, dass* $\mathbb{N} = N$ *ist.*

Definition 15: Wir nennen $x \in \mathbb{R}$ eine ganze Zahl, wenn $x \in \mathbb{N}$ oder $-x \in \mathbb{N}$ oder $x = 0$ gilt. Die Menge aller ganzen Zahlen bezeichnen wir mit $\mathbb{Z}$. Weiter setzen wir

$$\mathbb{Q} := \left\{ x \in \mathbb{R} | \text{ Es gibt } a, b \in \mathbb{Z}, b \neq 0, \text{ so dass } x = a \cdot b^{-1} =: \frac{a}{b} \right\}.$$

Die Elemente von $\mathbb{Q}$ nennen wir die rationalen Zahlen. Eine reelle Zahl, die nicht rational ist, heißt *irrationale Zahl*.

Ende des Exkurses

1.2. Die komplexen Zahlen

Da in angeordneten Körpern stets $x^2 \geq 0$ ist, hat die Gleichung $x^2 = -1$ in $\mathbb{R}$ sicherlich keine Lösung. Wir wollen nun einen Körper konstruieren, in dem auch diese Gleichung eine Lösung hat und der die reellen Zahlen als Teilmenge enthält.

Hierzu definieren wir auf $\mathbb{R}^2$ zwei Verknüpfungen $+$ und $\cdot$ durch

$$\begin{aligned}(a,b)+(x,y) &:= (a+x, b+y)\\ (a,b)\cdot(x,y) &:= (ax-by, bx+ay).\end{aligned}$$

Satz 16: $(\mathbb{R}^2, +, \cdot)$ ist ein Körper. Wir nennen diesen Körper den *Körper der komplexen Zahlen*.

Es gilt

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

Wir setzen $i := (0,1)$ und identifizieren eine reelle Zahl $x \in \mathbb{R}$ mit dem Tupel $(x,0) \in \mathbb{R}^2$. Dann lässt sich die vorstehende Gleichung als

$$i^2 = -1$$

schreiben. In dem gerade konstruierten Körper besitzt also die negativen Zahl -1 eine Wurzel; genauer natürlich zwei, denn es ist auch $(-i)^2 = -1$.

Außerdem gilt für $\lambda \in \mathbb{R}$ und $(a,b) \in \mathbb{R}^2$

$$\lambda(a,b) = (\lambda,0)(a,b) = (\lambda a, \lambda b) \text{ und } (a,b)\lambda = (a,b)(\lambda,0) = (a\lambda, b\lambda) = (\lambda a, \lambda b).$$

Somit ist

$$a + ib = (a,0) + (0,b) = (a,b).$$

Dies ist die übliche Art komplexe Zahlen zu notieren. Wir setzen

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

Wenn man $i^2 = -1$ beachtet, kann man in $\mathbb{C}$ genauso rechnen, wie man es von den reellen Zahlen gewohnt ist, also etwa

$$(1+i)(2-i) = 1\cdot 2 - 1\cdot i + 2\cdot i - i^2 = 2 - i + 2i - (-1) = 3 + i.$$

Ist $z = a + ib \in \mathbb{C}$ eine komplexe Zahl, so bezeichnet man a als den *Realteil von* z und b als den *Imaginärteil von* z. Wir schreiben auch $\mathrm{Re}(z)$ für den Realteil und $\mathrm{Im}(z)$ für den Imaginärteil von z.

Zu $z := a + ib$ definieren wir das Konjugierte von z durch

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

Für zwei beliebige komplexe Zahlen z_1 und z_2 gilt

$$\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2} \text{ und } \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}.$$

Außerdem ist $z\overline{z} = a^2 + b^2$, also eine reelle Zahl. Dies ist zum Beispiel nützlich, wenn man das multiplikative Inverse einer komplexen Zahl berechnen möchte. Ist $z = a + ib \neq 0$, so ist

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

Problem 8: *Bestimmen Sie reelle Zahlen a und b, so dass die folgenden Gleichungen erfüllt sind (für jede Gleichung ergeben sich natürlich andere a und b):*

(i) $(3 + i) + (2 - i) = a + bi$,

(ii) $(3 + i)(2 - i) = a + bi$,

(iii) $\overline{3 + i} = a + bi$,

(iv) $\frac{1}{3+i} = a + bi$,

(v) $\frac{2-i}{3+i} = a + bi$.

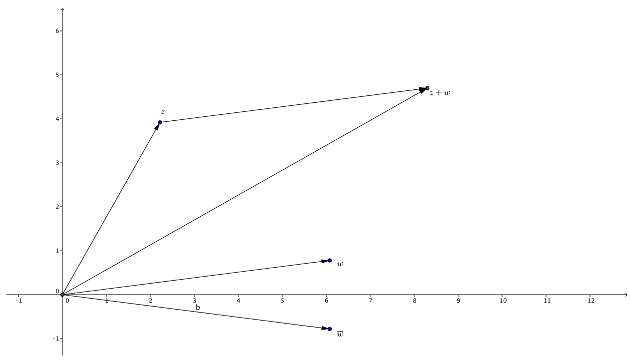


Abbildung 1.1.: Geometrische Bedeutung der komplexen Addition und der komplexen Konjugation

$x^2 = -1$ ist nicht die einzige Gleichung, die in den komplexen Zahlen lösbar ist und in den reellen nicht.

Wir wollen uns kurz anschauen, dass jede quadratische Gleichung

$$z^2 + \alpha z + \beta = 0$$

mit komplexen Koeffizienten $\alpha, \beta \in \mathbb{C}$ eine komplexe Lösung z besitzt.

Zunächst betrachten wir Gleichungen vom Typ

$$z^2 = \gamma$$

mit $\gamma \in \mathbb{C}$. Wenn wir $z = x + iy$ und $\gamma = a + ib$ schreiben, dann können wir $z^2 = \gamma$ auch als

$$(x + iy)^2 = a + ib$$

schreiben beziehungsweise, wenn wir die linke Seite ausmultiplizieren, als

$$x^2 - y^2 + 2ixy = a + ib.$$

Da zwei komplexe Zahlen genau dann gleich sind, wenn Real- und Imaginärteile übereinstimmen, muss also

$$x^2 - y^2 = a \text{ und } 2xy = b$$

sein. Ist $b = 0$ und $a \geq 0$, so muss $y = 0$ und $x = \pm\sqrt{a}$ sein. Ist $b = 0$ und $a \leq 0$, so muss $x = 0$ und $y = \pm\sqrt{-a}$ sein.

Ist $b \neq 0$, so sind auch x und y ungleich Null. Aus der zweiten Gleichung erhalten wir dann, dass $y = \frac{b}{2x}$ sein muss. Wenn wir dies in die erste Gleichung einsetzen, so erhalten wir

$$x^2 - \left(\frac{b}{2x}\right)^2 = a$$

beziehungsweise, wenn wir mit x^2 durchmultiplizieren und alles auf eine Seite holen,

$$x^4 - \frac{b^2}{4} - ax^2 = 0.$$

Dies ist eine Gleichung mit reellen Koeffizienten, deren reelle Lösung wir suchen. Wir ergänzen quadratisch und erhalten

$$0 = x^4 - ax^2 - \frac{b^2}{4} = x^4 - 2\frac{a}{2}x^2 + \left(\frac{a}{2}\right)^2 - \left(\frac{a}{2}\right)^2 - \frac{b^2}{4} = \left(x^2 - \frac{a}{2}\right)^2 - \frac{a^2}{4} - \frac{b^2}{4}.$$

Wir erhalten also

$$\left(x^2 - \frac{a}{2}\right)^2 = \frac{a^2 + b^2}{4}.$$

Somit ist

$$x^2 = \frac{a}{2} \pm \frac{\sqrt{a^2 + b^2}}{2}.$$

Da das Quadrat einer reellen Zahl nicht negativ sein kann, haben wir

$$x = \pm\sqrt{\frac{a}{2} + \frac{\sqrt{a^2 + b^2}}{2}}.$$

Wir haben also im Fall $b \neq 0$ zwei Lösungen gefunden.

Wenden wir uns nun der allgemeinen quadratischen Gleichung

$$z^2 + \alpha z + \beta = 0$$

zu. Wie oben können wir auch hier quadratisch ergänzen und erhalten

$$0 = z^2 + \alpha z + \beta = z^2 + 2\frac{\alpha}{2}z + \left(\frac{\alpha}{2}\right)^2 - \left(\frac{\alpha}{2}\right)^2 + \beta = \left(z + \frac{\alpha}{2}\right)^2 - \left(\frac{\alpha}{2}\right)^2 + \beta$$

beziehungsweise

$$\left(z+\frac{\alpha}{2}\right)^2 = \left(\frac{\alpha}{2}\right)^2 - \beta.$$

Mit $w := z + \frac{\alpha}{2}$ und $\gamma := \left(\frac{\alpha}{2}\right)^2 - \beta$ lässt sich dies als $w^2 = \gamma$ schreiben und wir haben somit eine Gleichung vom zuvor behandelten Typ.

Problem 9: *Bestimmen Sie die beiden komplexen Lösungen* $z_1 = x_1 + iy_1$ *und* $z_2 = x_2 + iy_2$ *mit* $x_1, x_2, y_1, y_2 \in \mathbb{R}$ *der Gleichung*

$$z^2 + (-1+i)z - (4+8i) = 0.$$

Bemerkung: Von den Lösungen sind die Real- und Imaginärteile explizit anzugeben.

Wir sehen also, dass jede quadratische Gleichung mit komplexen Koeffizienten eine komplexe Lösung besitzt. Allgemeiner gilt aber noch mehr.

Satz 17 (Fundamentalsatz der Algebra)**:** Es seien $n \in \mathbb{N}, a_0, a_1, \ldots, a_n \in \mathbb{C}$ und $a_n \neq 0$. Dann gibt es ein $\alpha \in \mathbb{C}$, so dass

$$a_n\alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_1\alpha + a_0 = 0$$

ist.

Einen Beweis dieses Satzes geben wir auf Seite 253. Wir wollen uns aber noch überlegen, dass ein Polynom über den komplexen Zahlen vollständig in lineare Faktoren zerfällt.

Lemma 18: Es seien $n \in \mathbb{N}, a_0, a_1, \ldots, a_n \in \mathbb{C}$ mit $a_n \neq 0$ und $\alpha \in \mathbb{C}$, so dass

$$a_n\alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_1\alpha + a_0 = 0$$

ist. Dann existieren $b_0, b_1, \ldots, b_{n-1} \in \mathbb{C}$ mit $b_{n-1} \neq 0$, so dass für alle $z \in \mathbb{C}$ gilt:

$$a_nz^n + a_{n-1}z^{n-1} + \ldots + a_1z + a_0 = (z-\alpha)(b_{n-1}z^{n-1} + \ldots + b_1z + b_0).$$

Beweis: Wir multiplizieren zunächst die rechte Seite der angestrebten Gleichung aus und erhalten

$$\begin{aligned} & (z-\alpha)(b_{n-1}z^{n-1} + \ldots + b_1z + b_0) \\ = \; & b_{n-1}z^n + (b_{n-2} - \alpha b_{n-1})z^{n-1} + (b_{n-3} - \alpha b_{n-2})z^{n-2} + \ldots + (b_0 - \alpha b_1)z - \alpha b_0 \end{aligned}$$

Setzen wir nun nacheinander

$$\begin{array}{rcl} b_{n-1} & := & a_n(\neq 0) \\ b_{n-2} & := & a_{n-1} + \alpha b_{n-1} \\ b_{n-3} & := & a_{n-2} + \alpha b_{n-2} \\ \vdots & & \vdots \\ b_0 & := & a_1 + \alpha b_1, \end{array}$$

so haben wir

$$(z-\alpha)(b_{n-1}z^{n-1} + \ldots + b_1 z + b_0) = a_n z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \ldots + a_1 z - \alpha b_0$$

für alle $z \in \mathbb{C}$. Es bleibt also zu zeigen, dass $-\alpha b_0 = a_0$ ist. Hierzu setzen wir in diese Gleichung nun α für z ein. Die linke Seite der Gleichung wird wegen des Faktors $(z-\alpha)$ zu Null. Es ist also

$$a_n\alpha^n + a_{n-1}\alpha^{n-1} + a_{n-2}\alpha^{n-2} + \ldots + a_1\alpha - \alpha b_0 = 0.$$

Nach Voraussetzung ist auch

$$a_n\alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_1\alpha + a_0 = 0.$$

Ziehen wir nun die zweite von der ersten Gleichung ab, so ergibt sich $0 = -a_0 - \alpha b_0$, also $-\alpha b_0 = a_0$. Damit ist der Satz bewiesen. $\square$

Mittels vollständiger Induktion folgt nun sofort aus dem Fundamentalsatz:

Korollar 19: Es seien $n \in \mathbb{N}, a_0, a_1, \ldots, a_n \in \mathbb{C}$ mit $a_n \neq 0$. Dann existieren komplexe Zahlen $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$, so dass für alle $z \in \mathbb{C}$ gilt:

$$a_n z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0 = a_n(z-\alpha_1)(z-\alpha_2)\cdot \ldots \cdot(z-\alpha_n).$$

Die komplexen Zahlen haben aber nicht nur Vorteile gegenüber den reellen Zahlen. Wie eingangs erwähnt gilt in jedem angeordneten Körper $x^2 \geq 0$ für alle Elemente x.

Lemma 20: Der Körper $\mathbb{C}$ kann nicht angeordnet werden.

Obwohl die komplexen Zahlen nicht angeordnet werden können, kann man aber auf ihnen einen Betrag definieren, so dass $\mathbb{C}$ zu einem normierten Körper wird.

Satz 21 (Betrag komplexer Zahlen)**:** Die Menge der komplexen Zahlen $\mathbb{C}$ ist zusammen mit der Abbildung $|\cdot| : \mathbb{C} \to \mathbb{R}$

$$z \mapsto |z| := \sqrt{z \cdot \overline{z}} = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}$$

ein normierter Körper.

Der vorstehende Satz versetzt uns nun in die Lage, auch in den komplexen Zahlen Analysis betreiben zu können. Damit ist das Spielfeld bereitet und wir können uns nun den wirklichen Objekten der Analysis, nämlich Folgen, Reihen und Funktionen zuwenden.

1.3. Die Vollständigkeit der reellen Zahlen

Auch die rationalen Zahlen $\mathbb{Q}$ sind ein angeordneter Körper. Man stellt sich also unweigerlich die Frage, was die reellen Zahlen von den rationalen Zahlen unterscheidet. Wir erinnern dazu an ein Ergebnis, welches dem Leser hoffentlich bekannt ist.

Satz 22: Es gibt keine rationale Zahl x derart, dass $x^2 = 2$ ist.

BEMERKUNG: Auf `http://www.cut-the-knot.org/proofs/sq_root.shtml` finden sich über 20 Beweise für die Irrationalität von $\sqrt{2}$!

BEWEIS: Wir nehmen an, dass es eine derartige rationale Zahl x doch gibt. Dann gibt es teilerfremde, ganze Zahlen u, v derart, dass $x = \frac{u}{v}$ ist. Also haben wir

$$u^2 = 2v^2.$$

Da 2 also u^2 teilt, muss 2 auch schon u teilen. Es gibt also ein w derart, dass $u = 2w$ ist. Setzen wir dies in die vorstehende Gleichung ein, so erhalten wir $4w^2 = 2v^2$ und nachdem wir durch 2 gekürzt haben

$$2w^2 = v^2.$$

Hieraus folgt aber wie zuvor, dass v durch 2 teilbar ist, was einen Widerspruch zur Teilerfremdheit von u und v darstellt. $\square$

Aus der Schule weiß man, dass es eine reelle Zahl $\sqrt{2}$ gibt, für die $(\sqrt{2})^2 = 2$ gilt. Dies ist also ein Unterschied zwischen den rationalen und den reellen Zahlen. Wir wollen uns nun klarmachen, worauf dieser Unterschied beruht. Dazu führen wir zunächst ein paar

Begriffe ein. Danach werden wir die Existenz der reellen Zahl $\sqrt{2}$ (und noch mehr) formal beweisen.

Definition 23 (Supremum, Infimum, Maximum, Minimum, Schranke)**:** Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper und $A \subset K$.

a) Die Teilmenge A heißt *nach oben beschränkt*, wenn es ein $S \in K$ gibt, so dass $x \leq S$ für alle $x \in A$ gilt.
Die Zahl S heißt dann *eine obere Schranke von* A.

b) Die Teilmenge A heißt *nach unten beschränkt*, wenn es ein $s \in K$ gibt, so dass $x \geq s$ für alle $x \in A$ gilt.
Die Zahl s heißt dann *eine untere Schranke von* A.

c) Die Zahl $S \in K$ heißt *kleinste obere Schranke* oder *Supremum von* A, falls gilt:

- S ist eine obere Schranke von A.
- Ist S' eine obere Schranke von A, so gilt $S \leq S'$.

Wir schreiben $\sup(A)$ für das Supremum von A.

d) Die Zahl s heißt *größte untere Schranke* oder *Infimum von* A, falls gilt:

- s ist eine untere Schranke von A.
- Ist s' eine untere Schranke von A, so gilt $s' \leq s$.

Wir schreiben $\inf(A)$ für das Infimum von A.

e) Gilt $\sup(A) \in A$, so nennt man $\sup(A)$ auch das *Maximum von* A und schreibt $\max(A)$ für $\sup(A)$.

f) Gilt $\inf(A) \in A$, so nennt man $\inf(A)$ auch das *Minimum von* A und schreibt $\min(A)$ für $\inf(A)$.

Beispiel: Ist $A := \{x \in \mathbb{R} | 0 \leq x < 7,5\}$, so gilt

$$\begin{aligned} \inf(A) &= \min(A) = 0 \\ \sup(A) &= 7,5. \end{aligned}$$

Das Maximum von A existiert nicht.

Das folgende Lemma gibt eine nützliche alternative Charakterisierung für das Supremum an. Den Beweis des Lemmas überlassen wir dem Leser.

Lemma 24: Es sei K ein angeordneter Körper und $A \subset K$ eine Teilmenge von K. Eine Zahl $S \in K$ ist genau dann das Supremum von A, wenn die beiden folgenden Bedingungen erfüllt sind:

(i) für alle $x \in A$ gilt $x \leq S$ und

(ii) zu jedem $\varepsilon > 0$ existiert ein $x \in A$ mit $S - \varepsilon < x \leq S$.

Problem 10: *Beweisen Sie das vorstehende Lemma.*

BEMERKUNG: Dies ist das erste Mal, dass in dieser Vorlesung der Ausdruck *zu jedem* $\varepsilon > 0$ vorkommt. Dieser Ausdruck wird noch sehr oft vorkommen. Eine Mathematikerin denkt sich bei diesem Ausdruck ε automatisch als eine sehr kleine, positive reelle Zahl. Dies ist auch ein guter Moment, um über Konventionen in der Mathematik zu sprechen. Es gibt einige solcher Konventionen. Natürliche Zahlen werden gerne mit n, m oder k, l bezeichnet, reelle Zahlen mit x oder y und komplexe Zahlen mit z oder w. Solche Konventionen erhöhen die Lesbarkeit von mathematischen Texten. Formal sagen

$$\forall \varepsilon > 0 \exists n \in \mathbb{N} : \frac{1}{2^n} < \varepsilon$$

und

$$\forall n > 0 \exists \varepsilon \in \mathbb{N} : \frac{1}{2^\varepsilon} < n$$

natürlich das gleiche aus, aber keine ausgebildete Mathematikerin würde letzteres schreiben. Und Sie sollten das auch nicht tun! Wir schließen diese Bemerkung mit dem angeblich kürzesten Mathematikerwitz der Welt: „Sei $\varepsilon < 0$."

Definition 25: Ein angeordneter Körper K heißt *vollständig*, wenn jede nichtleere, nach oben beschränkte Teilmenge $A \subset K$ ein Supremum besitzt.

Axiom 3: Der Körper $\mathbb{R}$ der reellen Zahlen ist vollständig.

Als erste Konsequenz der Vollständigkeit erhalten wir das sogenannte *Archimedische Prinzip*.

Satz 26 (Archimedisches Prinzip)**:** Zu jedem $x \in \mathbb{R}$ existiert ein $n \in \mathbb{N}$ mit $x < n$; die Menge der natürlichen Zahlen ist also nicht nach oben beschränkt.

BEWEIS: Wir beweisen dies durch einen Widerspruchsbeweis und nehmen dazu an, dass die Menge der natürlichen Zahlen $\mathbb{N}$ eine nach oben beschränkte Teilmenge der reellen Zahlen ist. Natürlich ist die Menge der natürlichen Zahlen nichtleer. Also existiert nach dem Vollständigkeitsaxiom 3 das Supremum von $\mathbb{N}$, setze $S := \sup(\mathbb{N})$. Da $S - 1$ keine obere Schranke für $\mathbb{N}$ sein kann, muss es eine natürliche Zahl n mit $n > S - 1$ geben. Mit n ist auch $n + 1$ eine natürliche Zahl und es ist $n + 1 > S = \sup(\mathbb{N})$, was der Definition des Supremums widerspricht. □

BEMERKUNG: Natürlich bleibt das Archimedische Prinzip korrekt, wenn man $\mathbb{R}$ durch $\mathbb{Q}$ ersetzt. Das Archimedische Prinzip ist also nicht äquivalent zum Vollständigkeitsaxiom.

Das Archimedische Prinzip gibt uns das folgende Resultat über das Wachstum von Potenzen.

Satz 27 (Wachstum von Potenzen)**:** Es sei b eine reelle Zahl.

(i) Ist $b > 1$, so existiert zu jedem $M \in \mathbb{R}$ ein $n \in \mathbb{N}$, so dass
$$b^n > M$$
gilt.

(ii) Ist $0 < b < 1$, so existiert zu jedem $\varepsilon > 0$ ein $n \in \mathbb{N}$, so dass
$$b^n < \varepsilon$$
gilt.

BEWEIS:

(i) Es sei also M eine beliebige reelle Zahl und $b > 1$. Setzen wir $x := b - 1$, so ist $x > 0$. Aufgrund der Bernoulli-Ungleichung gilt für jedes $n \in \mathbb{N}$
$$b^n = (1 + x)^n \geq 1 + nx.$$
Nun benötigen wir also ein n derart, dass $1 + nx > M$ ist. Nach dem Archimedischen Prinzip gibt es ein $n \in \mathbb{N}$ mit $n > \frac{M-1}{x}$, woraus auch $b^n > M$ folgt.

(ii) Für den zweiten Teil des Satzes brauchen wir eine Vorüberlegungen: Es gilt $(b^{-1})^n = (b^n)^{-1}$. Dies beweisen wir mit vollständiger Induktion nach n. Der Fall $n = 1$ ist klar. Der Induktionsschritt erschließt sich wie folgt:
$$b^{n+1} \cdot \left(b^{-1}\right)^{n+1} \overset{\text{Def}}{=} b^n \cdot b \cdot b^{-1} \left(b^{-1}\right)^n = b^n \cdot \left(b^{-1}\right)^n \overset{\text{Ind.Vor}}{=} 1.$$

Es sei nun $0 < b < 1$. Da $b > 0$ ist, ist auch $b^{-1} > 0$. Multipliziert man nun beide Seiten der Ungleichung $b < 1$ mit b^{-1}, so ergibt sich $1 < b^{-1}$. Nach (i) existiert zu $M = \varepsilon^{-1}$ ein n, so dass

$$\left(b^{-1}\right)^n > \varepsilon^{-1}$$

gilt. Multipliziert man beide Seiten der Ungleichung mit $b^n\varepsilon$ und berücksichtigt, dass nach Vorüberlegung $\left(b^{-1}\right)^n = \left(b^n\right)^{-1}$ ist, so ergibt sich die Behauptung.

□

Wir wollen nun zum Abschluss dieses Abschnitts wieder auf die Frage nach der Existenz reeller Lösungen der Gleichung $x^2 = 2$ oder allgemeiner $x^k = a$ zurückkommen und deren Existenz zeigen.

Satz 28: Es existiert genau eine reelle Zahl $b > 0$ mit $b^2 = 2$.

BEWEIS: *Eindeutigkeit:* Sind b_1, b_2 reelle Zahlen mit $b_1^2 = b_2^2 = 2$, so ist $0 = b_1^2 - b_2^2 = (b_1 - b_2)(b_1 + b_2)$. Es ist also entweder $b_1 = b_2$ oder $b_1 = -b_2$, was zeigt, dass es höchstens eine positive Zahl b wie im Satz geben kann.

Existenz: Aufgrund des Vollständigkeitsaxioms 3 existiert das Supremum der Menge

$$A := \{x \in \mathbb{R}_+ | x^2 \leq 2\},$$

wobei wir $\mathbb{R}_+ := \{x \in \mathbb{R} \,|\, x > 0\}$ setzen. Dieses Supremum sei mit b bezeichnet und wir werden zeigen, dass $b^2 = 2$ ist.

Die Idee ist dabei, dass wir die beiden Fälle $b^2 < 2$ und $b^2 > 2$ ausschließen. Dies machen wir, in dem wir eine Zahl x konstruieren, deren Quadrat zwischen b^2 und 2 liegt. Hierzu machen wir einmal den Ansatz $x := b + 1/n$ und im zweiten Fall $x := b - 1/n$. n ist dabei geeignet zu wählen.

Sei also angenommen, dass $b^2 < 2$ ist. Es soll nun

$$\left(b + \frac{1}{n}\right)^2 \overset{!}{<} 2$$

gelten. Wenn wir die linke Seite ausmultiplizieren, so erhalten wir

$$\left(b + \frac{1}{n}\right)^2 = b^2 + \frac{2b}{n} + \frac{1}{n^2} \leq b^2 + \frac{2b}{n} + \frac{1}{n}.$$

Es reicht also aus, ein n zu finden, so dass

$$b^2 + \frac{2b}{n} + \frac{1}{n} < 2$$

ist beziehungsweise
$$\frac{2b+1}{2-b^2} < n.$$
Da $b^2 < 2$ sein soll, ist $2 - b^2 > 0$. Nach dem Archimedischen Prinzip gibt es tatsächlich eine natürliche Zahl n mit
$$\frac{2b+1}{2-b^2} < n$$
und somit haben wir
$$\left(b + \frac{1}{n}\right)^2 = b^2 + \frac{2b}{n} + \frac{1}{n^2} \leq b^2 + \frac{2b}{n} + \frac{1}{n} < 2.$$
Damit ist aber $b + \frac{1}{n} \in A$ und $b + \frac{1}{n} > b = \sup(A)$, was nicht sein kann.

Wäre andererseits $b^2 > 2$, so gäbe es wieder nach dem Archimedischen Prinzip eine natürliche Zahl n mit
$$\max\left\{\frac{1}{b}, \frac{2b}{b^2-2}\right\} < n.$$
Hieraus folgte dann
$$\left(b - \frac{1}{n}\right)^2 = b^2 - \frac{2b}{n} + \frac{1}{n^2} > b^2 - \frac{2b}{n} > 2.$$
Damit wäre $b - \frac{1}{n}$ eine obere Schranke von A und kleiner als $b = \sup(A)$, was wieder nicht sein kann.

Es muss also tatsächlich $b^2 = 2$ gelten. □

BEMERKUNG: Die folgende Teilmenge der rationalen Zahlen
$$A := \{x \in \mathbb{Q} | x^2 \leq 2\}$$
ist nichtleer und nach oben beschränkt ist. Sie besitzt aber in $\mathbb{Q}$ kein Supremum, da für dieses wie im vorstehenden Beweis gezeigt $\sup(A)^2 = 2$ gelten müsste, was aber Satz 22 widersprechen würde. Also ist der Körper der rationalen Zahlen $\mathbb{Q}$ nicht vollständig.

Allgemeiner kann man zeigen

Satz 29: Zu jedem $a \in \mathbb{R}_+$ und $k \in \mathbb{N}$ existiert genau eine reelle Zahl $b > 0$ mit $b^k = a$.

BEMERKUNG: Der Beweis verläuft sehr ähnlich zu dem Beweis von Satz 28. Man kann
$$b := \sup\{x \in \mathbb{R}_+ | x^k \leq a\}$$
wählen. Dann schätzt man $(b + \frac{1}{n})^k$ und $(b - \frac{1}{n})^k$ mit Hilfe des Binomischen Lehrsatzes und der Bernoulli-Ungleichung ab.

Definition 30 (Quadratwurzeln und allgemeine Wurzeln positiver reeller Zahlen)**:** Es sei $a \in \mathbb{R}_+$ und $k \in \mathbb{N}$. Die eindeutig bestimmte positive, reelle Zahl b mit der Eigenschaft, dass $b^k = a$ ist, heißt die *k-te Wurzel von a* und wir schreiben $\sqrt[k]{a}$ für b. Ist $k = 2$, so nennen wir b auch die *(Quadrat-)Wurzel von a* und schreiben $\sqrt{a}$ statt $\sqrt[2]{a}$.

Mit Hilfe der allgemeinen Wurzeln können wir die Definition von Potenzen für positive reelle Basen auf Potenzen mit rationalen Exponenten ausdehnen.

Definition 31 (Potenzen mit rationalen Exponenten)**:** Es sei $a \in \mathbb{R}_+$ und $r = \frac{p}{q} \in \mathbb{Q}$ mit $p, q \in \mathbb{Z}$ und $q > 0$. Dann setzen wir

$$a^r := (\sqrt[q]{a})^p.$$

Wir halten noch kurz einige Rechenregeln fest. Wir sparen uns den Beweis, da wir später mit Hilfe der Exponentialfunktion und des Logarithmus noch allgemeinere Potenzen mit reellen Exponenten definieren werden. Bei dieser Gelegenheit werden wir dann diese Rechenregeln in allgemeinerer Form beweisen.

Satz 32: Für beliebige $a, b \in \mathbb{R}_+$ und $r, s \in \mathbb{Q}$ gilt

(i) $(a^r)^s = a^{rs}$,

(ii) $a^r a^s = a^{r+s}$,

(iii) $a^r b^r = (ab)^r$,

(iv) $a \neq 0 \Rightarrow (\frac{1}{a})^r = \frac{1}{a^r} = a^{-r}$.

Eine schöne Anwendung des Archimedischen Prinzips ist der folgende Satz, der zeigt, dass die rationalen Zahlen nicht so dünn gestreut sind wie die ganzen Zahlen.

Satz 33: Sind $x, y \in \mathbb{R}$ mit $x < y$, dann liegen zwischen x und y unendlich viele rationale Zahlen. Ebenso liegen zwischen x und y unendlich viele irrationale Zahlen.

BEMERKUNG: Für die Tatsache, dass zwischen zwei reellen Zahlen stets eine rationale Zahl zu finden ist, sagt man auch, dass $\mathbb{Q}$ *dicht in $\mathbb{R}$ liegt.*

BEWEIS: Wir konstruieren zunächst ein $\alpha = \frac{a}{b} \in \mathbb{Q}$ mit $x < \alpha < y$. Nach dem Archimedischen Prinzip existiert eine natürliche Zahl b mit $b > \frac{1}{y-x}$. Wir setzen $a := [xb] + 1$, wobei $[xb]$ die kleinste ganze Zahl bezeichnet, die kleiner oder gleich xb ist; die sogenannte *Gauß-Klammer* von xb. Dann gilt $xb = xb - 1 + 1 < [xb] + 1 \leq xb + 1$ und somit

$$x = \frac{xb}{b} < \frac{[xb]+1}{b} \leq \frac{xb+1}{b} = x + \frac{1}{b} < x + (y-x) = y.$$

Diese Konstruktion lässt sich nun beliebig oft wiederholen, d.h. wir können auch zwischen x und α eine rationale Zahl finden und ebenso zwischen α und y. Somit lassen sich also beliebig viele rationale Zahlen zwischen x und y finden.

Sind $\alpha_1, \alpha_2 \in \mathbb{Q}$ mit $x < \alpha_1 < \alpha_2 < y$, so ist $\beta := \alpha_1 + \frac{\alpha_2 - \alpha_1}{\sqrt{2}}$ irrational und es gilt $\alpha_1 < \beta < \alpha_2$. Wegen Satz 33 gibt es unendlich viele derartige irrationale Zahlen.
$\square$

Problem 11: *Zeigen Sie, dass das β im vorstehenden Beweis irrational ist.*

2. Folgen und Reihen

2.1. Funktionen

Ein Hauptgegenstand der Analysis ist die Untersuchung von Funktionen, die auf einer Teilmenge der reellen Zahlen oder allgemeiner auf einer Teilmenge eines höherdimensionalen Raum $\mathbb{R}^d$ definiert sind und reelle Werte annehmen.

Bevor wir richtig anfangen, wollen wir ein paar grundlegende Begriffe klären.

Definition 34: Sind A und B Mengen, so nennt man eine Zuordnungsvorschrift, die jedem $a \in A$ genau ein $b \in B$ zuordnet, eine *Funktion (oder Abbildung) von A nach B*. Nennen wir diese Funktion f, so bezeichnen wir das dem Element $a \in A$ eindeutig zugeordnete Element $b \in B$ mit $f(a)$ und schreiben für die Funktion

$$\begin{aligned} f : A &\to B \\ a &\mapsto f(a). \end{aligned}$$

Die Menge A bezeichnet man als den Definitionsbereich der Funktion f und die Menge B als ihren Wertebereich.
Zwei Funktionen $f : A \to B$ und $g : X \to Y$ sind genau dann gleich, wenn $A = X, B = Y$ ist und $f(a) = g(a)$ für alle $a \in A$ gilt.

BEMERKUNG: Eine Funktion ist also erst dann komplett definiert, wenn man sowohl Definitions- als auch Wertebereich und die zugehörige Zuordnungsvorschrift angibt. Die beiden Funktionen $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ und $f : \mathbb{Q} \to \mathbb{R}, x \mapsto x^2$ sind nach der vorstehenden Definition nicht gleich. Häufig spart man sich aber dennoch einige Angaben und macht implizite Annahmen. Geben wir etwa bei einer Zuordnungsvorschrift auf den reellen Zahlen keinen Definitionsbereich an, so nehmen wir als Definitionsbereich die größtmögliche Teilmenge der reellen Zahlen, auf denen der Ausdruck $f(x)$ definiert ist. Wenn wir also zum Beispiel von der Funktion $f(x) = 1/x$ sprechen, so nehmen wir implizit an, dass der Definitionsbereich die Menge aller reellen Zahlen ungleich der Null ist.

Beispiele:

1) Sind $a_0, a_1, \ldots, a_n$ reelle Konstanten, so nennt man eine Funktion der Art $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) := \sum_{k=0}^{n} a_k x^k = a_n x^n + \ldots + a_1 x + a_0$ eine reelle Polynomfunktion.

2) Sind f und g zwei Polynomfunktionen, so nennt man eine Funktion h mit $h(x) = f(x)/g(x)$ eine rationale Funktion. Sie ist offensichtlich für alle x definiert, für die $g(x) \neq 0$ ist.

Der Wertebereich einer Funktion kann mehr Elemente enthalten, als von der Funktion wirklich angenommen werden. Deshalb unterscheiden wir ihn vom Bild der Funktion.

Definition 35: Ist $f : A \to B$ eine Funktion und $X \subset A$, so bezeichnen wir mit

$$f(X) := \{b \in B | \exists x \in X : f(x) = b\} \subset B$$

das *Bild von X unter der Funktion f*. $f(A)$ nennen wir *das Bild der Funktion f*. Die Menge

$$\{(a, f(a)) : a \in A\}$$

nennen wir *den Graphen von f*.

Nun noch drei wichtige Eigenschaften von Funktionen, die mit den eben eingeführten Begriffen zusammenhängen.

Definition 36: Es sei $f : A \to B$ eine Funktion.

(i) Wir sagen, dass f *surjektiv* ist (oder A surjektiv auf B abbildet), wenn $f(A) = B$ ist; d.h. Wertebereich und Bild der Funktion f stimmen überein. f ist also genau dann surjektiv, wenn

$$\forall b \in B \exists a \in A : f(a) = b$$

gilt.

(ii) Wir sagen, dass f *injektiv* ist, wenn

$$\forall a, a' \in A : f(a) = f(a') \Longrightarrow a = a'$$

gilt.

(iii) Wir sagen, dass f *bijektiv* ist (oder A bijektiv auf B abbildet), wenn f injektiv und surjektiv ist.

Problem 12: *Entscheiden Sie von den folgenden Abbildungen, ob sie injektiv, surjektiv oder bijektiv sind:*

(i) $f : \mathbb{N} \to \mathbb{N}, n \mapsto n + 1$.

(ii) $g : \mathbb{N} \to \{-1, 1\}, n \mapsto (-1)^n$.

(iii) $h : \mathbb{N} \to \mathbb{N}_0, n \mapsto n - 1$.

(iv) $j : \mathbb{N} \to \mathbb{N}, n \mapsto n^2$.

Exkurs: Visualisierung von Funktionen mit Matlab/Octave

Mittels Matlab und Octave lassen sich reellwertige Funktionen, die von einer oder zwei reellen Veränderlichen abhängen, sehr gut visualisieren.

Das folgende Skript plottet den Graphen der Funktion $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^3 - 3x^2 + x + 1$ für x zwischen -6 und 6.

```
figure(1)
clf()
x=-6:0.1:6
y=x.^3-3*x.^2+x+1;
plot(x,y)
grid on
xlabel('x')
ylabel('f(x)')
hold on
plot([-6 6],[0 0],'k')
plot([0 0],[-400 200],'k')
```

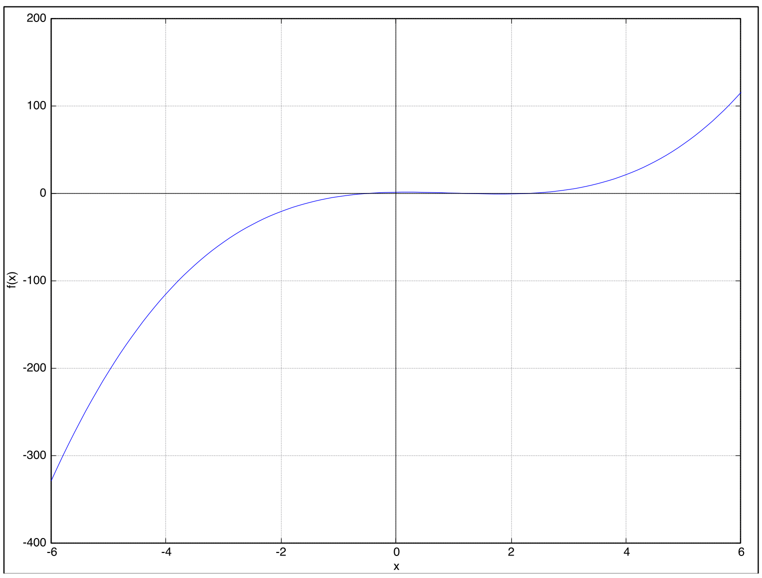


Abbildung 2.1.: Plot des Graphen der Funktion $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^3 - 3x^2 + x + 1$ für x zwischen -6 und 6

Das folgende Skript plottet den Graphen der Funktion $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 + y^2$ für x und y zwischen -5 und 5.

```
figure(1)
clf()
[x,y]=meshgrid(-5:0.1:5,-5:0.1:5);
f=x.^2+y.^2;
surf(x,y,f)
grid on
```

Ein paar Worte zur Erläuterung: Matlab bzw. Octave können sehr gut mit Arrays und Matrizen rechnen. `x=-6:0.1:6` legt einen Array mit Einträgen von -6 bis 6 mit einer Schrittweite von $0,1$ an, also $-6; -5,9; -5,8; ...; 5,8; 5,9; 6$. Wenn man nun einen Array `x` hat, so ist `x.^3` ein Array, in dem jeder Eintrag die dritte Potenz des entsprechenden Eintrags des Arrays `x` ist. Die Operation „hoch 3“ wird also elementweise durchgeführt. Deshalb steht auch ein Punkt vor dem `^`. Der Punkt sagt, dass die nachfolgende Operation elementweise durchgeführt werden soll. Hätte man hier den Punkt weggelassen, so hätte Matlab bzw. Octave versucht, das Matrixprodukt `x·x·x` zu berechnen, was aus Formatgründen nicht möglich ist und somit zu einer Fehlermeldung geführt hätte. Sind `x` und `y` zweigleich lange Arrays, so zeichnet `plot(x,y)` die Punkte, die man erhält, wenn man den i-ten Eintrag in `x` als die x-Koordinate des i-ten Punktes nimmt und den i-ten Eintrag in `y` als die y-Koordinate des i-ten Punktes. Die Punkte werden mit Geradestücken verbunden. Matlab bzw. Octave zeichnet also einen Polygonzug.

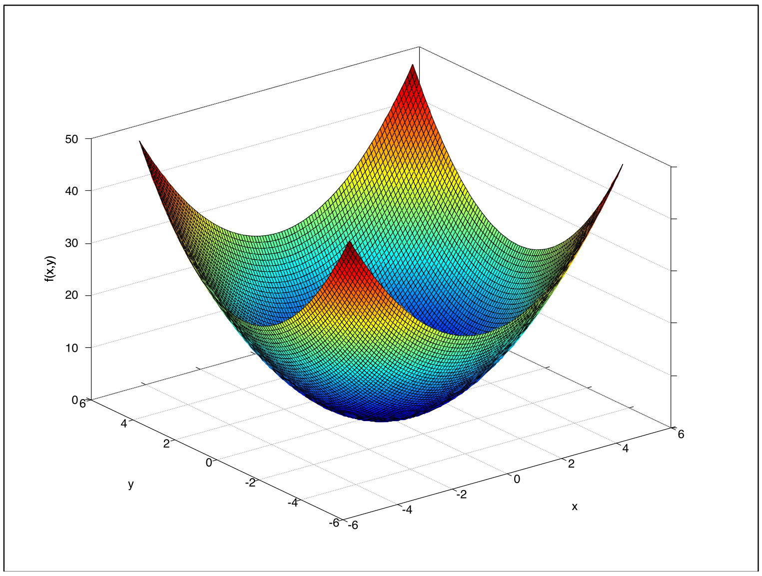


Abbildung 2.2.: Plot des Graphen der Funktion $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 + y^2$ auf $[-5, 5] \times [-5, 5]$

Ende des Exkurses

2.2. Folgen

Die ersten Funktionen, die wir studieren wollen, sind solche, deren Definitionsbereich im Wesentlichen gleich den natürlichen Zahlen ist. Wenn man eine Funktion $f : \mathbb{N} \to \mathbb{R}$ hat, dann kann man die Funktionswerte $f(1), f(2), f(3), \ldots$ als eine Aufzählung der Werte interpretieren. Man spricht hier von einer Folge. Historisch hat sich hier eine Schreibweise durchgesetzt, die von der bei Funktionen abweicht.

Definition 37 (Folge)**:** Es sei M eine Menge. Eine *Folge in* M ist eine Abbildung

$$\begin{aligned} a : \mathbb{N} &\to M \\ n &\mapsto a(n). \end{aligned}$$

Statt $a(n)$ schreibt man aber meist a_n und für diese Abbildung dann kürzer $(a_n)_{n\in\mathbb{N}}$, $(a_n)_{n\geq 1}$ oder auch nur (a_n). Das Element a_n bezeichnet man als das n-te Folgenglied der Folge (a_n).

BEMERKUNG: Folgen müssen nicht immer bei $n = 1$ beginnen. Manchmal beginnen sie auch bei $n = 0$, sind also Funktionen $a : \mathbb{N}_0 \to M$. Manchmal beginnen sie auch erst später oder noch früher und man schreibt dann zum Beispiel $(a_n)_{n\geq 7}$ oder $(a_n)_{n\geq -3}$. Dies ist völlig unproblematisch, da durch eine geeignete Indexverschiebung, die Situation aus der vorstehenden Definition hergestellt werden kann. Beispielsweise nimmt die Folge $\left(\frac{1}{n-2}\right)_{n\geq 3}$ die selben Werte wie die Folge $\left(\frac{1}{n}\right)_{n\geq 1}$.

Beispiele:

1) Es sei $a \in \mathbb{C}$ eine fest gewählte Zahl und $a_n := a$ für alle $n \in \mathbb{N}$. Dies ist also die Folge $a, a, a, a, a, \ldots$. Eine solche Folge heißt *konstant*.

2) $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ ist die Folge $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$

3) $(q^n)_{n\in\mathbb{N}}$ mit $q \in \mathbb{C}$ fest. Wähle für q zum Beispiel:

   - $q = \frac{1}{2}: \quad \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$
   - $q = -1: \quad -1, 1, -1, 1, -1, 1, \ldots$
   - $q = i: \quad i, -1, -i, 1, i, \ldots$
   - $q = 2: \quad 2, 4, 8, 16, 32, \ldots$

4) $\left(\left(\frac{1}{n}, \frac{n+1}{n}\right)\right)_{n\in\mathbb{N}}$ ist eine Folge in $\mathbb{R}^2$. Dazu kommt später mehr.

Für Mathematiker steht meist die Konvergenzfrage im Mittelpunkt des Interesses. Hiermit werden wir uns ab dem nächsten Abschnitt beschäftigen. Für Informatiker sind aber auch unbeschränkte Folgen interessant, etwa bei Laufzeitabschätzungen von Algorithmen. Ein Beispiel hierzu geben wir gleich in einem Exkurs, wenn wir uns AVL-Bäume anschauen.

Wir wollen noch eine Notation angeben, die wahrscheinlich jedem Informatiker geläufig ist.

Definition 38 (Groß-O Notation)**:** Es seien (a_n) und (b_n) zwei Folgen positiver reeller Zahlen. Die Folge (b_n) ist $O(a_n)$ (gesprochen: b_n ist Groß-O von a_n), wenn es eine positive reelle Konstante C gibt, so dass für alle $n \in \mathbb{N}$

$$b_n \leq C a_n$$

gilt; d.h. also in Quantorenschreibweise

$$\exists C > 0 \forall n \in \mathbb{N} : b_n \leq C a_n.$$

Exkurs: Visualisierung von Folgen mit Hilfe von Matlab oder Octave

Das folgende Matlab-Skript plottet einige der in den Beispielen behandelten Folgen:

```
clear all
figure(1)
clf()
n=1:100;
an=ones(1,100)*0.5;
bn=1./n;
cn=0.5.^n;
dn=(-1).^n;
plot(n,an,'.',n,bn,'.',n,cn,'.',n,dn,'.');
axis([-0.5 100 -1.5 1.5])
legend('a_n','b_n','c_n','d_n')
xlabel('n')
```

```
clear all
figure(2)
clf()
n=1:100;
an=1./n;
bn=(n+1)./n;
subplot(1,2,1)
plot(an,bn,'.');
```

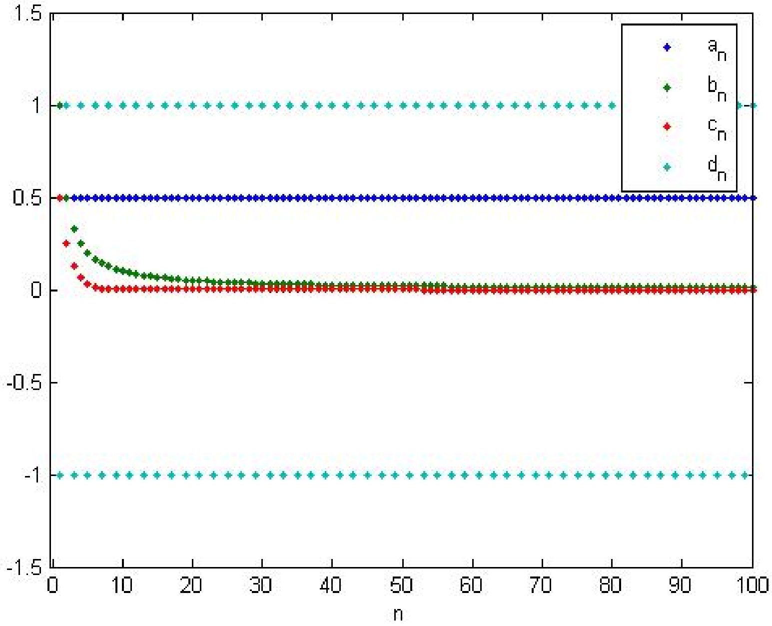


Abbildung 2.3.: Visualisierung einiger Folgen mit Matlab

```
xlabel('x')
ylabel('y')
subplot(1,2,2)
plot3(n,an,bn,'.');
grid on
xlabel('n')
ylabel('x')
zlabel('y')
```

Ende des Exkurses

Exkurs: Ein Beispiel für das Auftreten der Fibonacci-Folge in der Informatik

In dem Rechenbuch *Liber Abacci* von Leonardo von Pisa, genannt Fibonacci, aus dem 13. Jahrhundert findet sich die folgende Aufgabe: „Jemand setzt ein Paar Kaninchen in einen Garten, der auf allen Seiten von einer Mauer umgeben ist, um herauszufinden, wie viele Kaninchen innerhalb eines Jahre geboren werden. Wenn angenommen wird, dass jeden Monat jedes Paar ein weiteres Paar erzeugt, und dass Kaninchen zwei Monate nach ihrer Geburt geschlechtsreif sind, wie viele Paare Kaninchen werden dann jedes Jahr geboren?“

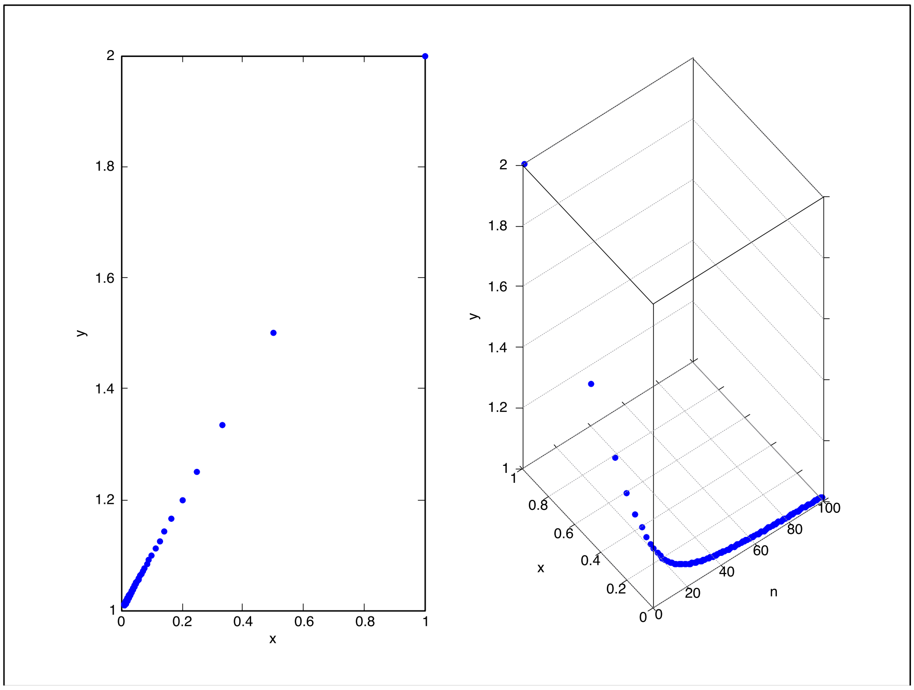


Abbildung 2.4.: Visualisierungen der Folge $(\frac{1}{n}, (\frac{n+1}{n}))$ im $\mathbb{R}^2$

Definition 39 (Fibonacci-Folge)**:** Die Fibonacci-Folge ist wie folgt definiert:

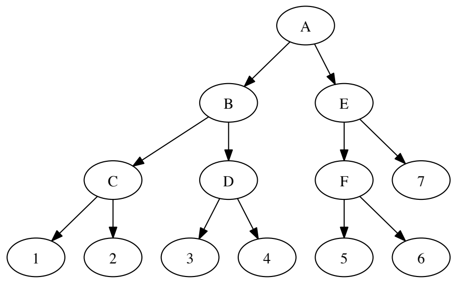
$$F_{n+1} = F_n + F_{n-1} \quad \text{für } n \geq 2, \quad F_1 = F_2 = 1.$$

Die Fibonacci-Zahlen finden sich in der Natur: Viele Pflanzen weisen in ihrem Bauplan Spiralen auf, etwa Ananas oder Sonnenblumen, deren Anzahlen durch Fibonacci-Zahlen gegeben sind.

Aber haben die Fibonacci-Zahlen auch etwas mit Informatik zu tun?

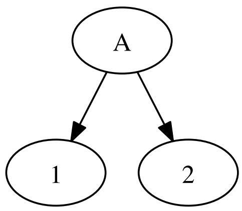
Definition 40 (AVL-Baum)**:** Es sei t ein binärer Suchbaum. Dann heißt t höhenbalanciert oder ein AVL-Baum, falls für jeden Knoten u des Baums gilt, dass sich die Höhe des linken Teilbaums von der Höhe des rechten Teilbaums um höchstens 1 unterscheidet.

Beispiel für einen AVL-Baum der Höhe 3:

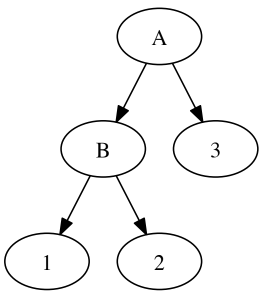


Wir wollen uns nun überlegen, welchen Zusammenhang es zwischen der Höhe eines AVL-Baums und der Mindestanzahl seiner Blätter gibt.

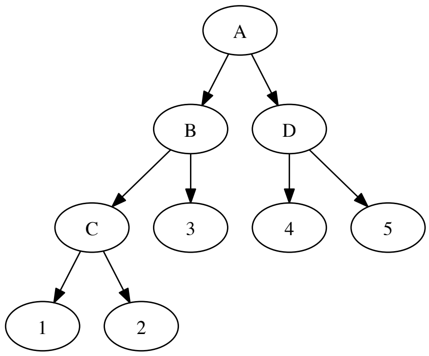
Ein AVL-Baum der Höhe 1 hat 2 Blätter :



Ein AVL-Baum der Höhe 2 hat mindestens 3 Blätter:



Ein AVL-Baum der Höhe 3 hat mindestens 5 Blätter:



Man überlegt sich leicht, dass ein AVL-Baum der Höhe h mindestens F_{h+2} Blätter hat.

In Aufgabe 19 wird gezeigt werden, dass

$$F_h \approx 1,170 \cdot 1,618^h$$

ist, woraus man nun folgern kann, dass ein AVL-Baum mit N Blättern höchstens eine Höhe von $1,44 \ldots \log_2 N + 1$ hat. Die Laufzeit für eine Suche in einem AVL-Baum mit N gespeicherten Informationen ist also $O(\log_2 N)$.

Ende des Exkurses

2.3. Der Grenzwert reeller Zahlenfolgen

Wir kommen nun zu einem der zentralen Begriffe der Analysis.

Definition 41 (Grenzwert einer Folge)**:** Es sei (a_n) eine reelle Zahlenfolge. Eine reelle Zahl a heißt ein *Grenzwert der Folge* (a_n), wenn zu jeder (noch so kleinen) reellen Zahl $\varepsilon > 0$ eine natürliche Zahl n_0 existiert, so dass für alle natürlichen Zahlen $n \geq n_0$

$$|a_n - a| < \varepsilon$$

gilt, d.h. für alle bis auf eventuell endlich viele Folgeglieder a_n gilt

$$a - \varepsilon < a_n < a + \varepsilon.$$

Man sagt dann auch, dass die Folge (a_n) gegen a *konvergiert*.

BEMERKUNG: (i) In Quantorenschreibweise bedeutet dies:

$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |a_n - a| < \varepsilon.$$

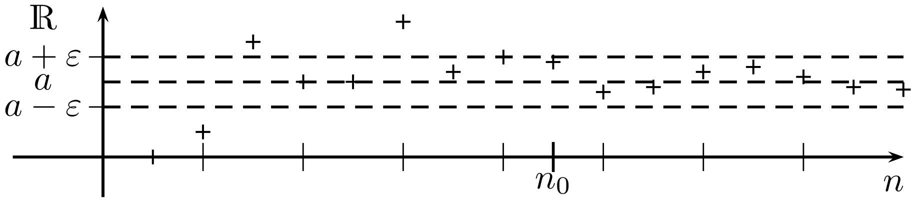


Abbildung 2.5.: Konvergente Folge in $\mathbb{R}$ mit Grenzwert a

(ii) Im Folgenden ist ε stets, auch wenn dies nicht explizit vermerkt wird, eine reelle, positive Zahl.

Definition 42 (Konvergenz)**:** Eine reelle Folge (a_n) heißt *konvergent*, wenn sie einen Grenzwert besitzt. Ist (a_n) nicht konvergent, so heißt (a_n) *divergent*.

BEMERKUNG: Die Quantorenschreibweise für die Aussage (a_n) *ist konvergent* lautet also

$$\exists a \in \mathbb{R} \forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |a_n - a| < \varepsilon$$

Somit lautet die Quantorenschreibweise für die Aussage (a_n) *ist divergent* :

$$\begin{aligned} &\neg (\exists a \in \mathbb{R} \forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |a_n - a| < \varepsilon) \\ =& (\forall a \in \mathbb{R} \exists \varepsilon > 0 \forall n_0 \in \mathbb{N} \exists n \geq n_0 : |a_n - a| \geq \varepsilon) \end{aligned}$$

Bevor wir uns Beispiele konvergenter und divergenter Folgen anschauen, wollen wir uns überlegen, dass der Grenzwert einer konvergenten Folge eindeutig bestimmt ist.

Satz 43: Sind a und a' Grenzwerte einer konvergenten Folge (a_n), so ist $a = a'$; das heißt der Grenzwert einer konvergenten Folge ist eindeutig bestimmt. Wir schreiben

$$\lim_{n \to \infty} a_n = a.$$

BEWEIS: Wir führen einen Widerspruchsbeweis.

Angenommen, die Folge (a_n) konvergiere gleichzeitig gegen a und a', und es wäre $a \neq a'$. Dann ist $|a - a'| > 0$. Wir wählen nun ein ε mit $0 < \varepsilon \leq \frac{1}{2}|a - a'|$. Laut Definition des Grenzwerts existieren natürliche Zahlen n_1 und n_2 derart, dass

$$\begin{aligned} \forall n \geq n_1 &: |a_n - a| < \varepsilon \\ \forall n \geq n_2 &: |a_n - a'| < \varepsilon \end{aligned}$$

gilt. Für $n \geq \max(n_1, n_2)$ erhält man dann unter Anwendung der Dreiecksungleichung

$$\begin{aligned} |a - a'| &= |a - a_n + a_n - a'| \\ &\leq |a - a_n| + |a_n - a'| \\ &< 2\varepsilon \\ &\leq |a - a'|. \end{aligned}$$

Dies ist offensichtlich ein Widerspruch. □

BEMERKUNG: Der vorstehende Beweis enthält einige für Beweisführungen in der Analysis sehr typische Elemente. Zum einen wird die Dreiecksungleichung angewendet. Um die Dreiecksgleichungen aber anwenden zu können, muss zunächst der Ausdruck der untersucht werden soll (hier: $|a - a'|$) *aufgebläht* werden, indem ein Term abgezogen und gleich wieder draufaddiert wird. Nach Anwendung der Dreiecksungleichungen entstehen dann zwei Ausdrücke, die mit Hilfe der Ausgangsvoraussetzungen abgeschätzt werden können.

Beispiele:

1) Die Folge $a_n = \frac{1}{n}$ konvergiert gegen 0; wir nennen (a_n) deshalb auch eine *Nullfolge*. Es sei $\varepsilon > 0$ beliebig. Nach dem Archimedischen Axiom existiert ein $n_0 \in \mathbb{N}$, so dass $1 < n_0\varepsilon$ ist. Für alle $n \geq n_0$ gilt also

$$|a_n - 0| = \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon.$$

2) Die Folge (b_n) mit

$$b_n := \begin{cases} 1 & \text{, falls } n \text{ gerade} \\ -1 & \text{, falls } n \text{ ungerade} \end{cases}$$

ist divergent. Denn, wenn wir annehmen, dass $b \in \mathbb{R}$ Grenzwert der Folge (b_n) ist, so gäbe es zu $\varepsilon := 1$ eine natürliche Zahl n_0, so dass für alle $n \geq n_0$

$$|b - b_n| < 1$$

ist. Für alle $n \geq n_0$ haben wir damit

$$\begin{aligned} 2 &= |b_{n+1} - b_n| \\ &= |b_{n+1} - b + b - b_n| \\ &\leq |b_{n+1} - b| + |b - b_n| \\ &< 2, \end{aligned}$$

was nicht möglich ist. Somit ist die Annahme nicht haltbar.

3) Es sei $c_n = q^n$ mit festem $q \in \mathbb{R}$.

- Für $q = 1$ haben wir die konstante Einsfolge und für $q = -1$ die Folge aus dem vorstehenden Beispiel.

- Ist $|q| < 1$, so ist $\lim_{n\to\infty} q^n = 0$.
  Es sei $\varepsilon > 0$ beliebig. Nach Satz 27 (ii) mit $b := |q| \in \mathbb{R}$ existiert ein $n_0 \in \mathbb{N}$, so dass $b^{n_0} < \varepsilon$ ist. Für alle $n \geq n_0$ gilt also
  $$|c_n - 0| = |q^n| = |q|^n \leq |q|^{n_0} = b^{n_0} < \varepsilon.$$

- Ist $|q| > 1$, so ist (q^n) divergent.
  Es sei $c \in \mathbb{R}$ beliebig. Nach Satz 27 (i) existiert ein $n_0 \in \mathbb{N}$, so dass
  $$|q|^n > 1 + |c|$$
  ist. Für alle $n \geq n_0$ gilt also
  $$|c_n - c| = |q^n - c| \geq |q|^n - |c| > 1 + |c| - |c| = 1.$$

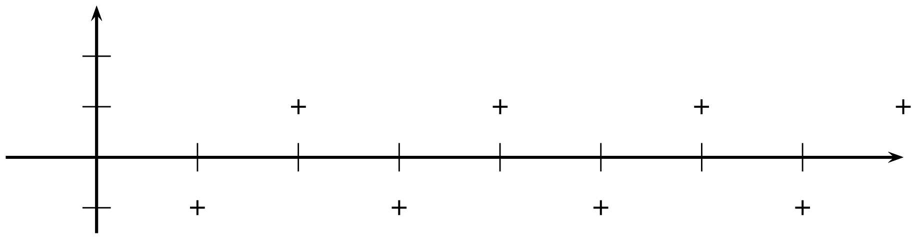


Abbildung 2.6.: Die divergente Folge (b_n)

Problem 13: *Zeigen Sie, dass eine reelle Zahlenfolge (a_n) genau dann gegen eine reelle Zahl a konvergiert, wenn $(a_n - a)$ eine Nullfolge ist, das heißt gegen 0 konvergiert.*

Problem 14: *Zwei Folgen (a_n) und (b_n) heißen äquivalent, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt, so dass für alle $n \geq n_0$*
$$|a_n - b_n| < \varepsilon$$
gilt. Zeigen Sie: Ist (a_n) konvergent, dann ist auch (b_n) konvergent und konvergiert gegen den selben Grenzwert.

Eine Anwendung des Grenzwertbegriffs bei Folgen in der Informatik ist die sogenannte *klein-o Notation*, die Ihnen sicherlich geläufig ist.

Definition 44 (klein-o Notation)**:** Es seien (a_n) und (b_n) zwei Folgen positiver reeller Zahlen. Die Folge (b_n) ist $o(a_n)$ (gesprochen: b_n ist klein-o von a_n), wenn
$$\lim_{n\to\infty} \frac{b_n}{a_n} = 0$$
ist.

Definition 45 (Beschränktheit von Folgen)**:** Es sei (a_n) eine reelle Zahlenfolge.

(i) Die Folge (a_n) heißt *nach oben beschränkt*, wenn es eine Konstante $K \in \mathbb{R}$ gibt, so dass $a_n \leq K$ für alle $n \in \mathbb{N}$ ist.

(ii) Die Folge (a_n) heißt *nach unten beschränkt*, wenn es eine Konstante $k \in \mathbb{R}$ gibt, so dass $a_n \geq k$ für alle $n \in \mathbb{N}$ ist.

(iii) Die Folge (a_n) heißt *beschränkt*, wenn es eine reelle Konstante $M > 0$ gibt, so dass für alle $n \in \mathbb{N}$

$$|a_n| \leq M$$

gilt.

Problem 15: *Finden Sie jeweils ein Beispiel für eine reelle Zahlenfolge, die*

(i) nach oben, aber nicht nach unten beschränkt ist.

(ii) die sowohl nach oben als auch nach unten nicht beschränkt ist.

(iii) die beschränkt ist, aber nicht konvergent.

Problem 16: *Zeigen Sie, dass eine reelle Zahlenfolge (a_n) genau dann beschränkt ist, wenn sie nach oben und nach unten beschränkt ist.*

Satz 46: Jede konvergente reelle Zahlenfolge (a_n) ist beschränkt.

BEWEIS: Es bezeichne a den Grenzwert der Folge (a_n). Dann existiert $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$

$$|a_n - a| < 1$$

gilt. Wir wenden nochmal den Kniff an, dass wir einen Ausdruck durch gleichzeitiges addieren und subtrahieren einer Zahl aufblähen, um dann die Dreiecksungleichung anwenden zu können. Wir haben für alle $n \geq n_0$ die Abschätzung

$$\begin{aligned} |a_n| &= |a_n - a + a| \\ &\leq |a_n - a| + |a| \\ &< 1 + |a|. \end{aligned}$$

Setze $M := \max\left(\{|a_1|, |a_2|, \ldots, |a_{n_0-1}|, 1 + |a|\}\right)$, dann folgt $|a_n| \leq M$ für alle $n \in \mathbb{N}$. $\square$

BEMERKUNG: Die Umkehrung gilt im Allgemeinen nicht. Wir haben oben gesehen, dass die durch

$$b_n := \begin{cases} 1, & n \text{ gerade} \\ -1, & n \text{ ungerade} \end{cases}$$

definierte Folge (b_n) divergent ist. Sie ist aber offensichtlich beschränkt, denn es gilt $|b_n| = 1$ für alle $n \in \mathbb{N}$.

Das Berechnen von Grenzwerten mittels der Definition ist recht mühsam. Es gibt einige Sätze, die uns die Berechnung von Grenzwerten erleichtern.

Satz 47 (Rechenregeln für Grenzwerte)**:** Es seien (a_n) und (b_n) konvergente Folgen in $\mathbb{R}$ mit $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Dann gilt:

(i) Die Folge $(a_n + b_n)$ konvergiert gegen den Grenzwert $a + b$, und die Folge $(a_n - b_n)$ ist konvergent mit Grenzwert $a - b$.

(ii) Die Folge $(a_n b_n)$ ist konvergent mit Grenzwert ab; insbesondere ist die Folge (λa_n) für festes $\lambda \in \mathbb{R}$ konvergent gegen λa.

(iii) Ist $a \neq 0$, so existiert ein $n_0 \in \mathbb{N}$, so dass $a_n \neq 0$ für alle $n \geq n_0$ ist. Die Folge $\left(\frac{1}{a_n}\right)_{n\geq n_0}$ ist konvergent mit Grenzwert $\frac{1}{a}$.

(iv) Die Folge $(|a_n|)$ ist konvergent mit Grenzwert $|a|$.

BEMERKUNG: Aus der Konvergenz der Folge $(|a_n|)$ kann man nicht auf die Konvergenz der Folge (a_n) schließen, wie etwa das Beispiel $a_n := (-1)^n$ zeigt. Die Umkehrung von (iv) gilt also nicht.

Problem 17: *Beweisen oder widerlegen Sie die folgenden Aussagen für reelle Zahlenfolgen:*

(i) Summe, Differenz, Produkt und Quotient zweier divergenter Folgen ist ebenfalls divergent (getrennt für alle vier Operationen).

(ii) Die beiden Folgen (a_n) und (b_n) sind genau dann konvergent, wenn die beiden Folgen $(a_n + b_n)$ und $(a_n - b_n)$ konvergent sind.

(iii) Sind die beiden Folgen (a_n) und (b_n) konvergent und gilt $b_n \neq 0$ für alle $n \in \mathbb{N}$, dann ist die Folge $(\frac{a_n}{b_n})$ konvergent.

BEWEIS:

(i) Wir führen hier nur den Beweis für die Folge $(a_n + b_n)$. Der Beweis für $(a_n - b_n)$ geht vollkommen analog.

Wir haben zu zeigen, dass wir zu jedem $\varepsilon > 0$, das uns vorgegeben wird, ein $n_0 \in \mathbb{N}$ finden können, so dass für alle $n \geq n_0$

$$|a_n + b_n - (a + b)| < \varepsilon$$

ist.

Es sei $\varepsilon > 0$ beliebig, aber fest vorgegeben. Nach der Definition des Grenwerts angewendet auf die Folgen (a_n) und (b_n) existieren natürliche Zahlen n_1 und n_2, so dass für alle $n \geq n_1$

$$|a_n - a| < \varepsilon$$

und für alle $n \geq n_2$

$$|b_n - b| < \varepsilon$$

gilt. Für alle $n \geq \max(\{n_1, n_2\}) =: n_0$ folgt

$$|a_n + b_n - (a + b)| \leq |a_n - a| + |b_n - b| < 2\varepsilon.$$

Die Konstante 2 ist unabhängig von ε (und natürlich auch von anderen von ε abhängenden Größen wie n_0, n_1) und kann durch Skalierung eliminiert werden. Wir werden dies an diesem Beispiel einmal vorführen. Ist $\varepsilon' > 0$ beliebig klein vorgegeben, so wählen wir oben $\varepsilon := \varepsilon'/2$. Zu diesem ε existieren wie zuvor gesehen $n_0, n_1 \in \mathbb{N}$, so dass für alle $n \geq \max(\{n_0, n_1\})$

$$|a_n + b_n - (a + b)| < 2\varepsilon = \varepsilon'.$$

gilt.

(ii) Wir haben zu zeigen, dass wir zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ finden können, so dass für alle $n \geq n_0$

$$|a_n b_n - ab| < \varepsilon$$

ist.

Es sei wieder $\varepsilon > 0$ beliebig, aber fest vorgegeben. Wie in Teil (i) existieren natürliche Zahlen n_1 und n_2, so dass für alle $n \geq n_1$

$$|a_n - a| < \varepsilon$$

und für alle $n \geq n_2$

$$|b_n - b| < \varepsilon$$

gilt. Wir müssen nun $|a_n b_n - ab|$ abschätzen und zeigen, dass dies klein wird. Hierbei hilft uns unser Kniff, den Ausdruck durch Addition und Subtraktion der gleichen Zahl (Aber welcher?[1]) aufzublähen. Es ist

$$a_n b_n - ab = a_n b_n - a b_n + a b_n - ab = (a_n - a) b_n + a(b_n - b).$$

[1] Lassen Sie den Kopf nicht hängen, wenn Sie da nicht selber drauf gekommen sind. In solchen Kniffen steckt viel mathematische Erfahrung. Die kommt nicht über Nacht. Aber je mehr Mathematik -in unserem Fall Analysis- man sieht und je mehr man sich auch mit den Beweisen beschäftigt, desto größer wird der Erfahrungsschatz und desto mehr wird man in die Lage versetzt, selber auf solche Kniffe zu kommen.

Da jede konvergente Folge beschränkt ist, existiert ein $B > 0$ derart, dass $|b_n| \leq B$ für alle $n \in \mathbb{N}$. Also folgt für alle $n \geq \max(\{n_1, n_2\})$

$$\begin{aligned} |a_n b_n - ab| \leq& |a_n - a||b_n| + |a||b_n - b| \\ \leq& \varepsilon B + |a|\varepsilon \\ =& (B + |a|) \cdot \varepsilon. \end{aligned}$$

Da die Konstante $(B + |a|)$ unabhängig von ε ist, haben wir die Behauptung bewiesen.

Den Zusatz erhält man, wenn man anstelle der Folge (b_n) die konstante Folge (λ) wählt.

□

Problem 18: *Beweisen Sie die Teile (iii) und (iv) den vorstehenden Satz.*

In den Übungen werden Sie beweisen:

Problem 19: *Zeigen Sie mittels vollständiger Induktion folgende explizite Darstellung der Fibonacci-Folge:*

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$$

Beweisen Sie schließlich mit der gewonnenen Darstellung, dass

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \frac{1+\sqrt{5}}{2} (\approx 1,6180).$$

Dieser Grenzwert ist der berühmte goldene Schnitt.

Beispiel: Wir wollen mit Hilfe der obigen Rechenregeln den Grenzwert der Folge (a_n) mit

$$a_n = \frac{n+7}{n^2+3n+5}$$

bestimmen und somit insbesondere zeigen, dass die Folge überhaupt konvergent ist. Zähler und Nenner sind für sich genommen divergente Folgen, d.h. wir können die Rechenregeln nicht sofort auf sie anwenden. Die Idee ist nun, dass man in Zähler und Nenner jeweils die größte auftretende Potenz von n herauszieht.

$$a_n = \frac{n+7}{n^2+3n+5} = \frac{n\left(1+\frac{7}{n}\right)}{n^2\left(1+\frac{3}{n}+\frac{5}{n^2}\right)}.$$

Da $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ eine Nullfolge ist, ist auch $\left(\frac{1}{n^2}\right)_{n\in\mathbb{N}}$ eine Nullfolge. Und wir haben

$$a_n = \frac{n+7}{n^2+3n+5} = \frac{n\left(1+\frac{7}{n}\right)}{n^2\left(1+\frac{3}{n}+\frac{5}{n^2}\right)} = \underbrace{\frac{1}{n}}_{\to 0} \cdot \frac{\overbrace{1+\frac{7}{n}}^{\to 1}}{\underbrace{1+\frac{3}{n}+\frac{5}{n^2}}_{\to 1}}.$$

Also ist $\lim_{n\to\infty} a_n = 0$.

Problem 20: *Untersuchen Sie die Folgen $(a_n), (b_n)$ und (c_n) auf Konvergenz bzw. Divergenz und bestimmen Sie im Konvergenzfall den Grenzwert für $n \to \infty$, wobei*

$$a_n := \frac{31n + 2n^2 - n^3}{(5n-2)^3},\ b_n := \sqrt{n+1} - \sqrt{n} \text{ und } c_n := \left(\left|\frac{2-2i}{1+4i}\right|\right)^n.$$

Tipp zu (b_n): Erweitern Sie geeignet, um die 3. binomische Formel ins Spiel zu bringen.

Auch die folgende Aufgabe werden Sie sich in den Übungen anschauen:

Problem 21: *(i) Es sei $(a_n)_{n\in\mathbb{N}}$ eine konvergente reelle Zahlenfolge mit $a_n \geq 0$ für alle $n \in \mathbb{N}$. Es sei $a := \lim_{n\to\infty} a_n$. Zeigen Sie, dass dann gilt*

$$\lim_{n\to\infty} \sqrt{a_n} = \sqrt{a}.$$

Tipp: Die 3. binomische Formel könnte hilfreich sein. Betrachten Sie den Fall $a = 0$ gesondert.

(ii) Berechnen Sie den Grenzwert (falls er existiert!) der Folge $(a_n)_{n\in\mathbb{N}}$ mit

$$a_n := \frac{\sqrt{n^2-1}}{n}.$$

Wir wollen noch eine Klasse von divergenten Folgen in den reellen Zahlen besonders hervorheben.

Definition 48 (Bestimmte Divergenz)**:** Es sei (a_n) eine Folge in $\mathbb{R}$.

(i) Wir sagen, dass (a_n) *bestimmt gegen $+\infty$ divergiert* (in Zeichen: $\lim_{n\to\infty} a_n = +\infty$), wenn es zu jeder Konstante $M > 0$ ein $n_0 \in \mathbb{N}$ gibt, so dass $a_n > M$ für alle $n \geq n_0$ ist.

(ii) Weiter nennen wir (a_n) *bestimmt divergent gegen $-\infty$* (in Zeichen: $\lim_{n\to\infty} a_n = -\infty$), wenn die Folge $(-a_n)$ bestimmt gegen $+\infty$ divergiert.

Problem 22: *(i) Geben Sie jeweils ein Beispiel für eine bestimmt gegen $+\infty$ bzw. $-\infty$ divergierende reelle Zahlenfolge.*

(ii) Geben Sie ein Beispiel für eine reelle Zahlenfolge, die unbeschränkt und divergent, aber nicht bestimmt divergent ist.

(iii) Zeigen Sie: Ist die Folge (a_n) bestimmt divergent gegen $+\infty$ oder $-\infty$, so existiert ein $n_0 \in \mathbb{N}$, so dass $a_n \neq 0$ für alle $n \geq n_0$ ist und die Folge $(1/a_n)_{n\geq n_0}$ ist eine Nullfolge.

(iv) Gilt in (iii) auch die Umkehrung? Beweis oder Gegenbeispiel.

2.4. Das Schachtelungsprinzip

Die Ergebnisse in diesem und dem nächsten Abschnitt benutzen ganz stark die Ordnungsrelation auf den reellen Zahlen. Diese Ergebnisse lassen sich nicht ohne Weiteres auf zum Beispiel komplexe Folgen übertragen.

Satz 49: Sind (a_n) und (b_n) reelle, konvergente Zahlenfolgen mit $a_n \leq b_n$ für alle $n \in \mathbb{N}$, so gilt

$$a := \lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n =: b$$

BEWEIS: Wir beweisen den Satz indirekt, d.h. wir zeigen, dass wenn $a \leq b$ nicht gilt, auch die Voraussetzung im Satz, dass $a_n \leq b_n$ für alle $n \in \mathbb{N}$ ist, nicht gilt. Es gelte also $a > b$. Dann können wir $\varepsilon := \frac{a-b}{2}(> 0)$ setzen und es existieren natürliche Zahlen n_0 und n_1, so dass für alle $n \geq n_0$

$$|a_n - a| < \varepsilon$$

und für alle $n \geq n_1$

$$|b_n - b| < \varepsilon$$

gilt. Für alle $n \geq \max(n_0, n_1)$ gilt also

$$a_n > a - \varepsilon = a - \frac{a-b}{2} = \frac{a+b}{2} = b + \frac{a-b}{2} = b + \varepsilon > b_n.$$

$\square$

BEMERKUNG: Wir betrachten nun die beiden Folgen (a_n) und (b_n) mit $a_n := \frac{1}{n}$ und $b_n := 0$. Die Folge $(\frac{1}{n})$ ist größer als Null für alle n, also haben wir für alle $n \in \mathbb{N}$

$$a_n > b_n.$$

Es gilt aber

$$\lim_{n\to\infty} \frac{1}{n} = 0 = \lim_{n\to\infty} 0.$$

Also folgt aus $a_n < b_n$ für alle n im Allgemeinen auch nur

$$\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$$

und nicht etwa die strikte Ungleichung.

Satz 50 (Schachtelungsprinzip)**:** Es seien $(a_n), (b_n), (c_n)$ reelle Zahlenfolgen mit $a_n \leq b_n \leq c_n$ für alle $n \in \mathbb{N}$. Weiterhin seien (a_n) und (c_n) konvergent mit

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = a.$$

Dann ist auch die Folge (b_n) konvergent und es gilt

$$\lim_{n\to\infty} b_n = a.$$

BEWEIS: Wir müssen zeigen, dass es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt, so dass für alle $n \geq n_0$

$$|b_n - a| < \varepsilon$$

gilt.

Es sei $\varepsilon > 0$ beliebig, dann existiert ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$

$$|a_n - a| < \varepsilon \text{ und } |c_n - a| < \varepsilon$$

gilt.[2] Wegen

$$a - \varepsilon < a_n \leq b_n \leq c_n < a + \varepsilon$$

ist die Behauptung $|b_n - a| < \varepsilon$ bewiesen. □

Beispiel: Wir wollen $b_n = \frac{n}{2^n}$ auf Konvergenz untersuchen.
Für $n > 3$ gilt $2^n \geq n^2$, also

$$0 < \frac{n}{2^n} \leq \frac{n}{n^2} = \frac{1}{n} \to 0.$$

Wählen wir also $a_n := 0$ und $c_n := \frac{1}{n}$ für alle $n \in \mathbb{N}$, so ergibt sich aus dem Schachtelungsprinzip

$$\lim_{n\to\infty} \frac{n}{2^n} = 0.$$

Problem 23: *(i) Folgern Sie aus dem Satz über das Wachstum von Potenzen (Satz 27), dass die Folge* $(\sqrt[n]{2})$ *gegen den Grenzwert 1 konvergiert.*

(ii) Bestimmen Sie mit Hilfe des Schachtelungsprinzips den Grenzwert der Folge $(\sqrt[n]{3^n + 5^n})$.

[2] Genauer ist es natürlich wieder so, dass ein $n_1 \in \mathbb{N}$ existiert, so dass $|a_n - a| < \varepsilon$ für alle $n \geq n_1$ gilt, und ein $n_2 \in \mathbb{N}$, so dass $|c_n - a| < \varepsilon$ für alle $n \geq n_2$ gilt. Ist n dann größer als die größere der beiden Zahlen n_1 und n_2, so gelten für n beide Aussagen gleichzeitig. Wir nehmen also $n_0 := \max\{n_1, n_2\}$. Wenn man aber etwas versierter im Umgang mit Folgen ist, dann sollten einem solche Feinheiten klar sein und dann kann man das auch mal kürzer fassen.

2.5. Der Satz von der monotonen Konvergenz

Definition 51 (Monotonie)**:** Eine reelle Zahlenfolge (a_n) heißt

- *monoton wachsend*, wenn $a_n \leq a_{n+1}$ für alle $n \in \mathbb{N}$ gilt,
- *monoton fallend*, wenn $a_n \geq a_{n+1}$ für alle $n \in \mathbb{N}$ gilt,
- *streng monoton wachsend*, wenn $a_n < a_{n+1}$ für alle $n \in \mathbb{N}$ gilt,
- *streng monoton fallend*, wenn $a_n > a_{n+1}$für alle $n \in \mathbb{N}$ gilt.

Die Folge heißt *monoton*, wenn sie monoton wachsend oder monoton fallend ist.

Lemma 52: Ist (a_n) eine monoton wachsende, konvergente reelle Zahlenfolge mit Grenzwert a, so gilt für alle $n \in \mathbb{N}$

$$a_n \leq a.$$

BEMERKUNG: Es gilt natürlich auch die folgende Aussage: Ist (a_n) eine monoton fallende, konvergente reelle Zahlenfolge mit Grenzwert a, so gilt für alle $n \in \mathbb{N}$

$$a \leq a_n.$$

Der Beweis verläuft analog oder man wendet bei einer monoton fallenden Folge (a_n) das Lemma auf die Folge $(-a_n)$ an, die monoton wachsend ist.

BEWEIS: Wir führen einen Widerspruchsbeweis und nehmen an, dass (a_n) eine monoton wachsende, konvergente reelle Zahlenfolge mit Grenzwert a ist und dass aber nicht für alle $n \in \mathbb{N}$

$$a_n \leq a$$

gilt. Dann existiert also ein $n_0 \in \mathbb{N}$ mit $a_{n_0} > a$. Da die Folge monoton wachsend ist, gilt dann für alle $n \geq n_0$

$$a_n \geq a_{n_0} > a.$$

Für $\varepsilon := \frac{1}{2}(a_{n_0} - a) > 0$ und alle $n \geq n_0$ gilt also

$$|a_n - a| = a_n - a \geq a_{n_0} - a > \frac{1}{2}(a_{n_0} - a) = \varepsilon.$$

Dies widerspricht der Tatsache, dass a der Grenzwert der Folge ist. Also muss die eingangs gemachte Annahme falsch sein. $\square$

Satz 53 (Satz von der monotonen Konvergenz)**:** Jede beschränkte, monotone reelle Zahlenfolge (a_n) konvergiert.

BEWEIS: Wir beschränken uns hier auf den Fall, dass (a_n) eine monoton wachsende, beschränkte Folge ist.

Da die Menge

$$A := \{a_n | n \in \mathbb{N}\}$$

nichtleer und beschränkt ist, besitzt sie ein Supremum.

Wir werden zeigen, dass $a := \sup A$ der Grenzwert der Folge ist.

Es sei $\varepsilon > 0$ beliebig. Da a das Supremum der Menge aller Folgenglieder ist, existiert zu diesem ε ein $N \in \mathbb{N}$ mit

$$a - \varepsilon < a_N \leq a.$$

Für alle $n \geq N$ gilt also

$$a - \varepsilon < a_N \leq a_n \leq a.$$

Das linke $\leq$ gilt wegen der Monotonie der Folge, das rechte $\leq$ da a eine obere Schranke der Menge A ist.

Also gilt für alle diese n

$$|a_n - a| = a - a_n < \varepsilon.$$

$\square$

Problem 24: *Beweisen oder widerlegen Sie die folgenden Aussagen:*

(i) Eine Folge konvergiert, wenn sie monoton und beschränkt ist.

(ii) Eine konvergente Folge ist monoton und beschränkt.

(iii) Wenn eine Folge nicht monoton ist, kann sie nicht konvergieren.

(iv) Wenn eine Folge nicht beschränkt ist, kann sie nicht konvergieren.

Wir hatten schon früher (Satz 29) gezeigt, dass jede positive reelle Zahl eine Wurzel besitzt. Wir wollen nun noch einen konstruktiven zweiten Beweis dieses Satzes liefern.

Satz 54: Es sei $a \in \mathbb{R}_+ := \{x \in \mathbb{R} \,|\, x > 0\}$. Die Folge (x_n) sei rekursiv durch

$$\begin{aligned} x_1 &:= a \\ x_{n+1} &:= \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \quad \text{für } n \geq 1 \end{aligned}$$

definiert. Dann ist die Folge (x_n) konvergent und für ihren Grenzwert x gilt $x > 0$ und $x^2 = a$.

BEMERKUNG: Man sieht im Beweis des Satzes, dass man jede positive Zahl x_1 als Startwert nehmen kann.

BEWEIS: Induktiv sieht man, dass x_n für alle $n \geq 0$ positiv ist. Aus der induktiven Definition der Folge leitet man für alle $n \geq 2$ die folgenden beiden Aussagen ab:

(i) $x_n^2 - a \geq 0$,

(ii) $x_n \geq x_{n+1}$.

Denn zum einen gilt für $n \geq 2$

$$\begin{aligned} x_n^2 - a =& \frac{1}{4}\left(x_{n-1} + \frac{a}{x_{n-1}}\right)^2 - a \\ =& \frac{1}{4}\left(x_{n-1}^2 + 2a + \frac{a^2}{x_{n-1}^2}\right) - \frac{4a}{4} \\ =& \frac{1}{4}\left(x_{n-1}^2 - 2a + \frac{a^2}{x_{n-1}^2}\right) \\ =& \frac{1}{4}\left(x_{n-1} - \frac{a}{x_{n-1}}\right)^2 \geq 0. \end{aligned}$$

Und zum anderen hat man dann

$$\begin{aligned} x_n - x_{n+1} =& x_n - \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \\ =& \frac{1}{2}\left(x_n - \frac{a}{x_n}\right) \\ =& \frac{1}{2x_n}\left(x_n^2 - a\right) \geq 0. \end{aligned}$$

Die Folge $(x_n)_{n\geq 2}$ ist also monoton fallend und für alle $n \geq 2$ gilt $0 < x_n \leq x_2$. Somit konvergiert die Folge nach Satz 53. Bezeichne nun x den Grenzwert der Folge (x_n). Gemäß Definition der Folge gilt die Beziehung

$$x_n x_{n+1} = \frac{1}{2}\left(x_n^2 + a\right).$$

Da mit x_n auch x_{n+1} gegen x konvergiert, erhalten wir

$$x^2 = \frac{1}{2}\left(x^2 + a\right)$$

also $x^2 = a$.

□

BEMERKUNG: Das Vorgehen beim Beweis des vorstehenden Satzes ist sehr typisch für das Vorgehen bei der Bestimmung des Grenzwertes einer rekursiv definierten Folge. Meistens zeigt man zunächst unter zur Hilfenahme des Satzes von der monotonen Konvergenz, dass die rekursiv definierte Folge überhaupt konvergiert, also einen Grenzwert besitzt. Den Grenzwert selber bestimmt man anschließend, indem man folgert, dass der Grenzwert einer Fixpunktgleichung genügen muss.

Der folgende Satz ist eine Verallgemeinerung von Satz 54. Wir werden den Satz hier nicht beweisen.

Satz 55: Es sei $a \in \mathbb{R}_+ := \{x \in \mathbb{R} \,|\, x > 0\}$ und $k \in \mathbb{N}$ mit $k > 1$. Die Folge (x_n) sei rekursiv durch

$$\begin{aligned} x_1 &:= a \\ x_{n+1} &:= \frac{1}{k}\left((k-1)x_n + \frac{a}{x_n^{k-1}}\right) \quad \text{für } n \geq 1 \end{aligned}$$

definiert. Dann ist die Folge konvergent in $\mathbb{R}$ und für ihren Grenzwert x gilt $x > 0$ und $x^k = a$.

Exkurs: Zwei Matlab-Funktionen zur Wurzelberechnung

Wir stellen hier zwei Beispiele vor, wie man die obigen Folgen in Matlab implementieren und zur Berechnung reeller Wurzeln benutzen kann.

Bei der ersten Implementierung wird eine feste Anzahl von Folgenglieder berechnet.

```
function erg    = wurzel1(a,N)
if a<0
    error('a muss groesser als Null sein');
end
if nargin==1
    N=10;
elseif nargin~=2
    error('Es werden 1 oder 2 Argumente erwartet')
```

```
end
x(1)=a;
for n=1:N
    x(n+1)=0.5*(x(n)+a/x(n));
end
erg=x(N+1);
```

Auf der anderen Seite liefert uns der Beweis des obigen Satzes eine Fehlerabschätzung. Es ist nämlich $x_n^2 \geq a$ (bzw. $\frac{1}{x_n^2} \leq \frac{1}{a}$) für alle n, also auch $\sqrt{a} \leq x_n$ und $\frac{a}{x_n} \leq \sqrt{a}$. Es gilt also für alle n

$$\frac{a}{x_n} \leq \sqrt{a} \leq x_n.$$

Bei der folgenden Implementierung wird solange gerechnet, bis der Fehler kleiner als eine vorgeschriebene Größe ist. Neben der Näherung für die Wurzel liefert die Funktion auch die Anzahl der berechneten Folgenglieder zurück.

```
function [erg n]  = wurzel2(a,e)
if a<0
    error('a muss groesser als Null sein');
end
if nargin==1
    e=0.0001;
elseif nargin~=2
    error('Es werden 1 oder 2 Argumente erwartet')
end
x(1)=a;
x(2)=0.5*(1+a);
n=2;
while(x(n)-a/x(n)>e)
    x(n+1)=0.5*(x(n)+a/x(n));
    n=n+1;
end
erg=x(n);
```

Ende des Exkurses

2.6. Intervallschachtelungen und der Satz von Bolzano-Weierstraß

Zunächst führen wir ein paar abkürzende Bezeichnungsweisen für typische Teilmengen der reellen Zahlen ein, die immer wieder vorkommen; sogenannte Intervalle. Wir werden den Begriff des Intervalls später allgemeiner definieren.

Es seien $a, b \in \mathbb{R}$. Wir nennen die Menge

$$[a, b] := \{x \in \mathbb{R} | a \leq x \leq b\}$$

das abgeschlossene Intervall von a bis b, die Menge

$$]a,b] := \{x \in \mathbb{R} | a < x \leq b\}$$

das linksoffene Intervall von a bis b, die Menge

$$[a,b[:= \{x \in \mathbb{R} | a \leq x < b\}$$

das rechtsoffene Intervall von a bis b und die Menge

$$]a,b[:= \{x \in \mathbb{R} | a < x < b\}$$

das offene Intervall von a bis b. Ist $a > b$, so sind die vorstehenden Mengen offensichtlich alle leer. Für den späteren Gebrauch führen wir weiter noch die folgenden uneigentlichen Intervalle ein. Es sei

$$[a,+\infty[:= \{x \in \mathbb{R} | x \geq a\},$$
$$]a,+\infty[:= \{x \in \mathbb{R} | x > a\},$$
$$]-\infty,a] := \{x \in \mathbb{R} | x \leq a\}.$$

und

$$]-\infty,a[:= \{x \in \mathbb{R} | x < a\}.$$

Satz 56 (Intervallschachtelung)**:** Es seien (x_n) und (y_n) reellen Folgen derart, dass die folgenden drei Bedingungen erfüllt sind:

a) $\forall n \in \mathbb{N}: \; I_n := [x_n, y_n] \neq \emptyset$

b) $\forall n \in \mathbb{N}: I_n \supset I_{n+1}$ und

c) $\lim\limits_{n\to\infty} y_n - x_n = 0$

Dann existiert genau ein $a \in \mathbb{R}$ mit

$$a \in [x_n, y_n]$$

für alle $n \in \mathbb{N}$.

BEWEIS: Existenz des a:

Aus den Eigenschaften a) und b) folgt, dass die Folge (x_n) monoton wächst und die Folge (y_n) monoton fällt. Außerdem gilt für alle $n \in \mathbb{N}$

$$x_1 \leq x_n \leq y_n \leq y_1.$$

Damit sind die Folgen (x_n) und (y_n) beschränkt und somit konvergent. Es existieren also Zahlen $a, b \in \mathbb{R}$ mit $\lim_{n\to\infty} x_n = a$ und $\lim_{n\to\infty} y_n = b$. Aus der Eigenschaft c) und den Rechenregeln für konvergente Folgen ergibt sich

$$0 = \lim_{n\to\infty} y_n - x_n = \lim_{n\to\infty} y_n - \lim_{n\to\infty} x_n = b - a$$

beziehungsweise
$$a = \lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = b.$$
Nach Lemma 52 gilt für alle $n \in \mathbb{N}$
$$x_n \le a \le y_n$$
und somit
$$\lim_{n\to\infty} x_n = a = \lim_{n\to\infty} y_n,$$
womit der Satz vollständig bewiesen ist. □

Definition 57 (Teilfolge)**:** Es sei (a_n) eine Folge und (n_k) eine Folge natürlicher Zahlen mit $n_1 < n_2 < \ldots$, dann heißt $(a_{n_k})_{k\in\mathbb{N}} = (a_{n_1}, a_{n_2}, a_{n_3}, \ldots)$ eine *Teilfolge der Folge* (a_n).

Beispiel: Ist (a_n) eine Folge, so können wir zum Beispiel die Teilfolge der Folgenglieder mit geradem Index betrachten. Dies wäre dann (a_{2n}).

Eine wichtige Anwendung der Intervallschachtelung ist der folgende Satz, den wir noch des öfteren anwenden werden.

Satz 58 (Satz von Bolzano-Weierstraß)**:** Jede beschränkte Folge (a_n) reeller Zahlen besitzt eine konvergente Teilfolge.

BEWEIS: Sei also (a_n) eine beschränkte Folge reeller Zahlen. Dann gibt es aufgrund der Beschränktheit Konstanten $A, B \in \mathbb{R}$, so dass $A \le a_n \le B$ für alle $n \in \mathbb{N}$ ist. Wir betrachten das Intervall
$$[A, B] := \{x \in \mathbb{R} | \; A \le x \le B\}$$
und konstruieren induktiv eine Folge von Intervallen $([A_k, B_k])_{k\in\mathbb{N}}$ und eine Folge natürlicher Zahlen (n_k), so dass für alle $k \in \mathbb{N}$ die folgenden Bedingungen erfüllt sind:

1) in $I_k := [A_k, B_k]$ liegen unendlich viele Folgenglieder von (a_n),

2) $[A_k, B_k] \subseteq [A_{k-1}, B_{k-1}]$,

3) $B_k - A_k = \frac{1}{2}(B_{k-1} - A_{k-1})$,

4) $n_k > n_{k-1}$,

5) $a_{n_k} \in [A_k, B_k]$.

Induktionsverankerung $(k = 1)$*:* Zunächst setzen wir

$$A_1 := A, \qquad B_1 := B, \qquad n_1 := 1.$$

Induktionsschritt $(k \to k+1)$*:* Sind $[A_k, B_k]$ und n_k bereits konstruiert, so ist $M_k := \frac{A_k+B_k}{2}$ die Mitte des Intervalls. Wegen 1) müssen in $[A_k, M_k]$ oder in $[M_k, B_k]$ unendlich viele Folgeglieder liegen. Ist mit $[A_{k+1}, B_{k+1}]$ dieses Intervall bezeichnet[3], so sind die Eigenschaften 1) bis 3) erfüllt.

Da $[A_{k+1}, B_{k+1}]$ unendlich viele Folgeglieder enthält, gibt es ein $n > n_k$ mit $a_n \in [A_{k+1}, B_{k+1}]$. Es sei n_{k+1} das kleinste solche n. So sieht man, dass die Eigenschaften 4) und 5) erfüllt sind.

Nach Satz 56 gibt es genau ein a, das in allen Intervallen I_k liegt. Wegen

$$|a - a_{n_k}| \leq B_k - A_k \leq \left(\frac{1}{2}\right)^{k-1} (B_1 - A_1)$$

konvergiert die Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ gegen dieses a, womit der Satz bewiesen ist.

□

2.7. Cauchy-Folgen

Definition 59 (Cauchy-Folge)**:** Eine Folge (a_n) in $\mathbb{R}$ heißt *Cauchy-Folge*, wenn es zu jedem $\varepsilon > 0$ eine Zahl $n_0 \in \mathbb{N}$ gibt, so dass $|a_n - a_m| < \varepsilon$ für alle natürlichen Zahlen $m, n \geq n_0$ gilt.

BEMERKUNG: In Quantorenschreibweise lautet die Cauchy-Eigenschaft also:

$$\forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N} \ \forall m, n \geq n_0 : \ |a_n - a_m| < \varepsilon.$$

Dies sieht sehr ähnlich zur Definition des Grenzwertes einer Folge aus.

Lemma 60: Jede Cauchy-Folge (a_n) in $\mathbb{R}$ ist beschränkt.

BEWEIS: Aufgrund der Cauchy-Eigenschaft der Folge (a_n) existiert ein $n_0 \in \mathbb{N}$, so dass für alle $m, n \geq n_0$

$$|a_n - a_m| < 1$$

[3] Liegen sowohl in $[A_k, M_k]$ als auch in $[M_k, B_k]$ unendlich viele Folgeglieder, so können wir eines beliebig auswählen.

gilt. Damit erhalten wir für alle $n \geq n_0$ die Abschätzung

$$\begin{aligned} |a_n| &= |a_n - a_{n_0} + a_{n_0}| \\ &\leq |a_n - a_{n_0}| + |a_{n_0}| \\ &< 1 + |a_{n_0}|. \end{aligned}$$

Setze $M := \max(\{|a_1|, |a_2|, \ldots, |a_{n_0-1}|, 1 + |a_{n_0}|\})$, dann folgt $|a_n| \leq M$ für alle $n \in \mathbb{N}$. □

Satz 61: Jede konvergente Folge in $\mathbb{R}$ ist eine Cauchy-Folge.

BEWEIS: Es sei also (a_n) eine konvergente Folge, deren Grenzwert mit a bezeichnet sei. Wir müssen zeigen, dass (a_n) der Cauchy-Eigenschaft genügt, also, dass es zu jedem $\varepsilon > 0$ ein geeignetes $n_0 \in \mathbb{N}$ gibt, so dass für alle $n, m \geq n_0$

$$|a_n - a_m| < \varepsilon$$

gilt.

Es sei $\varepsilon > 0$ beliebig, aber fest vorgegeben. Da die Folge (a_n) konvergent ist, existiert nach der Definition des Grenzwerts ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$

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

gilt. Sind nun $n, m \geq n_0$, so gilt

$$\begin{aligned} |a_n - a_m| &= |a_n - a + a - a_m| \\ &\leq |a_n - a| + |a - a_m| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

□

Satz 62: Jede Cauchy-Folge in $\mathbb{R}$ besitzt einen Grenzwert in $\mathbb{R}$. Mit anderen Worten: Jede Cauchy-Folge in $\mathbb{R}$ ist konvergent.

Problem 25: *Finden Sie ein Beispiel für eine Cauchy-Folge in $\mathbb{Q}$, die keinen Grenzwert in $\mathbb{Q}$ besitzt.*

BEWEIS: Es sei (a_n) eine Cauchy-Folge. Wir werden zeigen, dass die Folge konvergent ist. Das Problem ist, dass wir keinen natürlichen Kandidaten für den Grenzwert der Folge habe. Wir können also nicht direkt mit ε starten und loslegen.

Nach Lemma 60 ist die Folge (a_n) aber beschränkt. Somit besitzt sie nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge (a_{n_k}). Der Grenzwert dieser Teilfolge sei mit a bezeichnet. Wir haben nun zu zeigen, dass die Folge (a_n) gegen a konvergiert.

Sei $\varepsilon > 0$ beliebig.

Da die Teilfolge (a_{n_k}) gegen a konvergiert, existiert ein $N_1 \in \mathbb{N}$, so dass für alle $n_k \geq N_1$
$$|a_{n_k} - a| < \frac{\varepsilon}{2}$$
gilt.

Aufgrund der Cauchy-Eigenschaft der Folge (a_n) existiert ein $N_2 \in \mathbb{N}$, so dass für alle $n, m \geq N_2$
$$|a_n - a_m| < \frac{\varepsilon}{2}$$
gilt.

Wir setzen nun $n_0 := \max\{N_1, N_2\}$ und fixieren ein $n_k \geq n_0$. Dann gilt für alle $n \geq n_0$
$$|a_n - a| = |a_n - a_{n_k} + a_{n_k} - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \varepsilon.$$
□

Wir sehen also, dass in den reellen Zahlen der Begriff der Cauchy-Folge äquivalent zu dem Begriff der konvergenten Folge ist. Die Cauchy-Eigenschaft hat aber den Vorteil, dass in ihr der Grenzwert nicht vorkommt. Man kann also eventuell für eine Folge die Cauchy-Eigenschaft nachweisen, auch wenn man keine Idee hat, was ihr Grenzwert ist. Wir werden diesen Vorteil unter Anderem im kommenden Abschnitt über Reihen ausnutzen.

2.8. Konvergenz komplexer Zahlenfolgen und Konvergenz in $\mathbb{R}^r$ und $\mathbb{C}^r$

Wir haben früher gesehen, dass wir auch auf den komplexen Zahlen einen Betrag definieren können. Damit können wir die Definition des Grenzwertes praktisch wortwörtlich auf komplexe Zahlenfolgen übertragen.

Definition 63 (Grenzwert einer komplexen Zahlenfolge)**:** Es sei (a_n) eine komplexen Zahlenfolge. Eine komplexe Zahl a heißt *Grenzwert der Folge* (a_n), wenn zu jeder (noch so kleinen) reellen Zahl $\varepsilon > 0$ eine natürliche Zahl n_0 existiert, so dass für alle natürlichen Zahlen $n \geq n_0$

$$|a_n - a| < \varepsilon$$

gilt.

BEMERKUNG: Sätze, die nur unter Anwendung dieser Definition bewiesen wurden, bleiben auch für komplexe Zahlenfolgen gültig; etwa die Rechenregeln in Satz 47. Sätze, die ausnutzen, dass die reellen Zahlen ein angeordneter Körper sind, lassen sich natürlich nicht auf komplexe Zahlenfolgen übertragen; etwa das Schachtelungsprinzip.

Aufgrund der fehlenden Anordbarkeit der komplexen Zahlen macht es natürlich keinen Sinn, davon zu sprechen, dass eine Folge komplexer Zahlen nach oben (oder nach unten beschränkt) ist. Es ist aber sinnvoll, von beschränkten Folgen zu reden.

Definition 64: Ist (a_n) eine Folge komplexer Zahlen, so nennen wir (a_n) *beschränkt*, wenn es eine positive (reelle!) Konstante M gibt, so dass für alle $n \in \mathbb{N}$

$$|a_n| \leq M$$

gilt.

Da sich viele Ergebnisse über den komplexen Zahlen ebenso formulieren lassen, wie über den reellen Zahlen, arbeiten wir künftig mit der folgenden Konvention.

Konvention: Mit $\mathbb{K}$ bezeichnen wir den Körper der reellen Zahlen $\mathbb{R}$ oder den Körper der komplexen Zahlen $\mathbb{C}$, innerhalb eines Satzes oder einer Definition aber stets denselben Körper.

Definition 65 (Konvergenz von Folgen in $\mathbb{K}^r$)**:** Es sei $r \in \mathbb{N}$ beliebig. Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}^r$, so ist jedes Folgenglied ein r-Tupel in $\mathbb{K}^r$, also

$$a_n = (a_n^{(1)}, a_n^{(2)}, \ldots, a_n^{(r)}).$$

Wir sagen, dass die Folge (a_n) in $\mathbb{K}^r$ genau dann gegen $a = (a^{(1)}, a^{(2)}, \ldots, a^{(r)}) \in \mathbb{K}^r$ konvergiert, wenn für jede Komponentenfolge $(a_n^{(i)})_{n\in\mathbb{N}}, i = 1, \ldots, r$, gilt

$$\lim_{n\to\infty} a_n^{(i)} = a^{(i)}.$$

Beispiel: Die Folge $\left(\left(\frac{1}{n}, \frac{n}{n+1}\right)\right)_{n\in\mathbb{N}}$ im $\mathbb{R}^2$ konvergiert gegen $(0,1)$, da $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ eine Nullfolge ist und die Folge $\left(\frac{n}{n+1}\right)_{n\in\mathbb{N}}$ gegen 1 konvergiert.

Jede Folge in $\mathbb{C}$ kann auch als Folge in $\mathbb{R}^2$ auffassen. Wir haben also zwei unterschiedliche Konvergenzbegriffe, die wir anwenden können. Der folgende Hilfssatz zeigt, dass die beiden Konvergenzbegriffe aber zusammenfallen.

Lemma 66: Ist (z_n) eine Folge in $\mathbb{C}$ und $z \in \mathbb{C}$, so gilt

$$\lim_{n\to\infty} z_n = z \Leftrightarrow \begin{cases} \lim_{n\to\infty} \mathrm{Re}(z_n) = \mathrm{Re}(z) \\ \lim_{n\to\infty} \mathrm{Im}(z_n) = \mathrm{Im}(z). \end{cases}$$

Problem 26: *Beweisen Sie das vorstehenden Lemma.*

Auch der Begriff der Cauchy-Folge lässt sich nahezu wortwörtlich ins Komplexe übertragen.

Definition 67 (Cauchy-Folge in $\mathbb{C}$)**:** Eine Folge (z_n) in $\mathbb{C}$ heißt *Cauchy-Folge*, wenn es zu jedem $\varepsilon > 0$ eine Zahl $n_0 \in \mathbb{N}$ gibt, so dass $|z_n - z_m| < \varepsilon$ für alle natürlichen Zahlen $m, n \geq n_0$ gilt.

Satz 68: Eine Folge (z_n) in $\mathbb{C}$ ist genau dann eine Cauchy-Folge, wenn die beiden Folgen $(\mathrm{Re}(z_n))$ und $(\mathrm{Im}(z_n))$ Cauchy-Folgen (in $\mathbb{R}$) sind.

Problem 27: *Beweisen Sie den vorstehenden Satz.*

Korollar 69: Eine Folge (z_n) in $\mathbb{C}$ ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.

Exkurs: Die Mandelbrotmenge und Juliamengen

Für festes $c \in \mathbb{C}$ betrachten wir die Folge (z_n) komplexer Zahlen, die durch $z_1 := 0$ und

$$z_{n+1} := z_n^2 + c$$

für alle $n \in \mathbb{N}$ definiert ist. Man sieht leicht, dass sich die Folge für unterschiedliche Werte von c sehr unterschiedlich verhält. Es gibt Werte für den Parameter c, so dass die zugehörige Folge (z_n) konvergiert. Es gibt Werte, für die die Folge zwar beschränkt ist, aber divergiert. Und es gibt schließlich auch Werte für den Parameter c, für die die zugehörige Folge (z_n) unbeschränkt ist. Die *Mandelbrotmenge* ist die Menge aller Parameter c, für die die zugehörige Folge (z_n) beschränkt ist.

Problem 28: *Zeigen Sie: Gibt es ein N, so dass $|z_N| > 2$ und $|z_N| > |c|$ ist, so existiert ein (reelles!) $q > 1$, so dass $|z_{N+m}| \geq q^m |z_N|$ für alle $m \in \mathbb{N}$ ist. Die Folge (z_n) ist dann also unbeschränkt und c gehört nicht zur Mandelbrotmenge. Dies ist insbesondere für $|c| > 2$ der Fall.*

Das folgende Matlab-Skript plottet die Mandelbrotmenge. Wir nutzen in dieser Implementierung die Stärken von Matlab (resp. Octave), nämlich dass das Programm sehr gut und schnell mit komplexen Zahlen und Matrizen rechnen kann.

```
clear all
xmin=-2;
xmax=1;
ymin=-1.5;
ymax=1.5;
xpixel=1000;
ypixel=1000;
hx=(xmax-xmin)/xpixel;
hy=(ymax-ymin)/ypixel;
[ cx , cy ] = meshgrid (xmin:hx:xmax,ymin:hy:ymax);
c=cx+i*cy;
[m,n]=size(cx)
z=zeros(m,n);
M=zeros(m,n);
for k=1:200
    z=z.^2+c;
end
valid=(abs(z)<=2);
invalid=not(valid);
M(invalid)=40;
figure(4);
clf()
image(M);
```

Wir schauen uns nun nur einen Ausschnitt der Menge an, indem wir das Skript wie folgt abändern.

```
xmin=-1.78;
xmax=-1.74;
ymin=-0.03;
ymax=0.03;
xpixel=1000;
ypixel=1000;
```

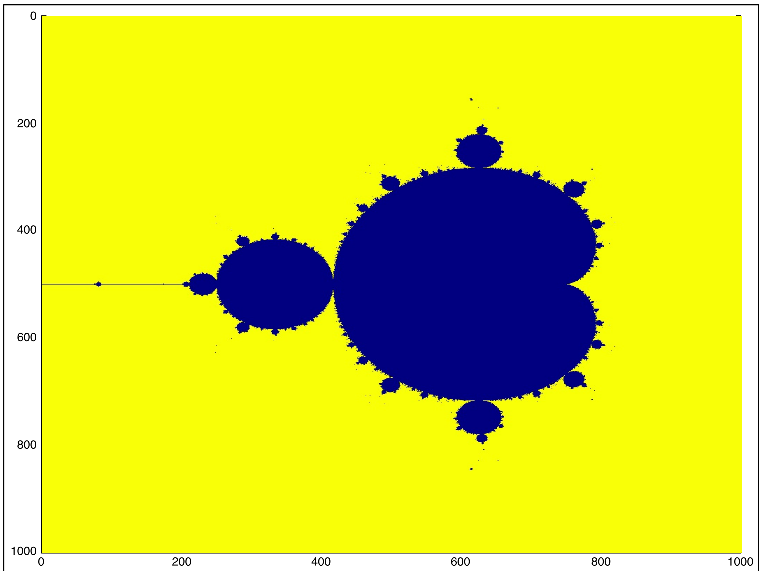


Abbildung 2.7.: Die Mandelbrotmenge

Der Ausschnitt weist eine große Ähnlichkeit mit der ganzen Mandelbrotmenge auf. Eine Menge, die selbstähnliche Teilmengen enthält, nennen wir ein *Fraktal*. Auch andere interessante Strukturen sind zu entdecken, wenn man Teile der Mandelbrotmenge stark vergrößert.

```
xmin=-0.75;
xmax=-0.74;
ymin=-0.09;
ymax=-0.085;
```

In diesem Ausschnitt des Randes der Mandelbrotmenge erkennt man zum Beispiel Seepferdchen-ähnliche Strukturen.

Ab den 1980er Jahren bekam die Mandelbrotmenge auch außerhalb von Mathematikerkreisen eine enorme für mathematische Themen eher unübliche Bekanntheit, was sicherlich zu einem erheblichen Teil an ästhetisch anspruchsvollen Darstellungen der Menge lag, die nun mit Hilfe von Computer erzeugt werden konnten. Hierbei werden üblicherweise die Punkte die nicht zur Mandelbrotmenge gehören je nach der Geschwindigkeit, mit der die Folge divergiert, in unterschiedlichen Farben eingefärbt. Ein Beispiel für eine solche Darstellung liefert das folgende Skript

```
clear all
xmin=-2;xmax=1;
ymin=-1.5;ymax=1.5;
```

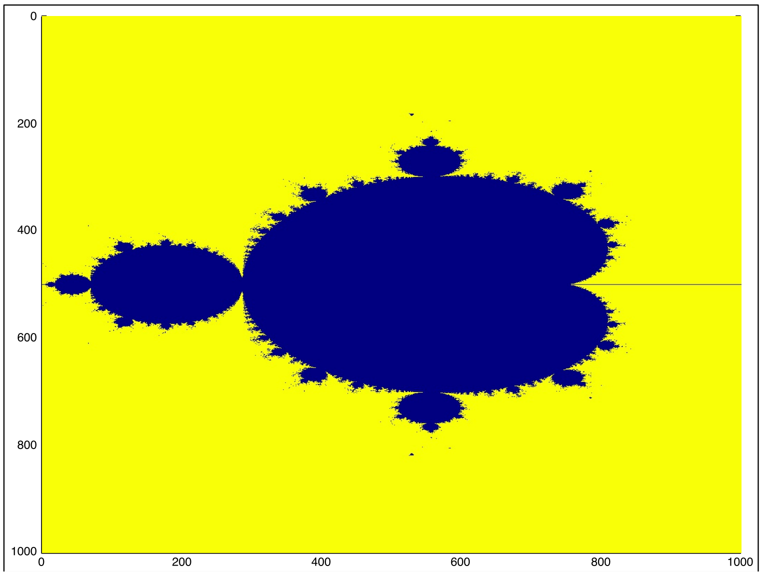


Abbildung 2.8.: Eine zur Mandelbrotmenge ähnliche Teilmenge der Mandelbrotmenge

```
xpixel=1000;ypixel=1000;
hx=(xmax-xmin)/xpixel;hy=(ymax-ymin)/ypixel;
[ cx , cy ] = meshgrid (xmin:hx:xmax,ymin:hy:ymax);
c=cx+i*cy;
[m,n]=size(cx)
z=zeros(m,n);
M=zeros(m,n);
for k=1:100
    invalid=not(abs(z)<=2);
    M(invalid)=M(invalid)+0.5;
    z=z.^2+c;
end
figure(4);
clf()
image(M);
```

Für aufwendigere Darstellungen der Mandelbrotmenge und anderer Fraktale kann man das freie Computerprogramm `Xaos (http://matek.hu/xaos/doku.php)` nutzen. Auch gibt es einen Zusammenhang zwischen Fraktalen und der sogenannten Chaostheorie, die alleine schon durch ihren Namen die Phantasie der Leute anspricht.

Die Mandelbrotmenge ist eng verwandt mit einer anderen Klasse von Fraktalen, den sogenannten Juliamengen. Die Juliamenge $J(c)$ zum Parameter c ist die Menge aller Startwerte z_1, für die, die durch

$$z_{n+1} := z_n^2 + c$$

für alle $n \in \mathbb{N}$ definierte Folge (z_n) beschränkt ist. Es gibt also überabzählbar viele Juliamengen, nämlich zu jedem Parameter c eine. Man kann zeigen, dass die Juliamenge $J(c)$ genau dann nichtleer ist, wenn c in der Mandelbrotmenge liegt.

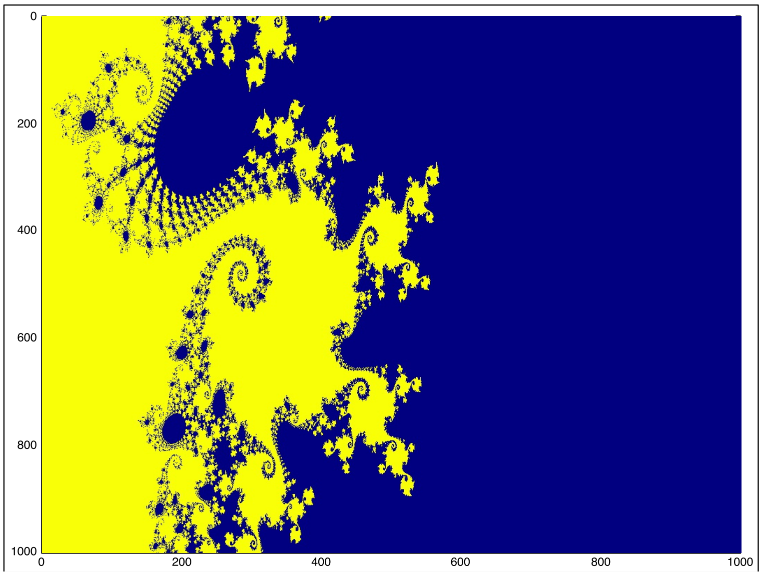


Abbildung 2.9.: Ein Ausschnitt der Mandelbrotmenge

Das folgende Skript plottet mehrere Juliamengen; wieder mit Farbverläufen, die die Geschwindigkeit der Divergenz für nicht zur Juliamenge gehörende Punkte wiedergeben.

```
clear all
xmin=-2;
xmax=2;
ymin=-2;
ymax=2;
xpixel=400;
ypixel=400;
hx=(xmax-xmin)/xpixel;
hy=(ymax-ymin)/ypixel;
[ zx , zy ] = meshgrid (xmin:hx:xmax,ymin:hy:ymax);
z=zx+i*zy;
figure(4);
clf()
[m,n]=size(zx)
M=zeros(m,n);
c=0;
for k=1:200
    invalid=not(abs(z)<=2);
    M(invalid)=M(invalid)+0.1;
    z=z.^2+c;
end
subplot(2,2,1)
image(M);
z=zx+i*zy;
c=-1;
M=zeros(m,n);
for k=1:200
    invalid=not(abs(z)<=2);
    M(invalid)=M(invalid)+0.1;
    z=z.^2+c;
end
subplot(2,2,2)
```

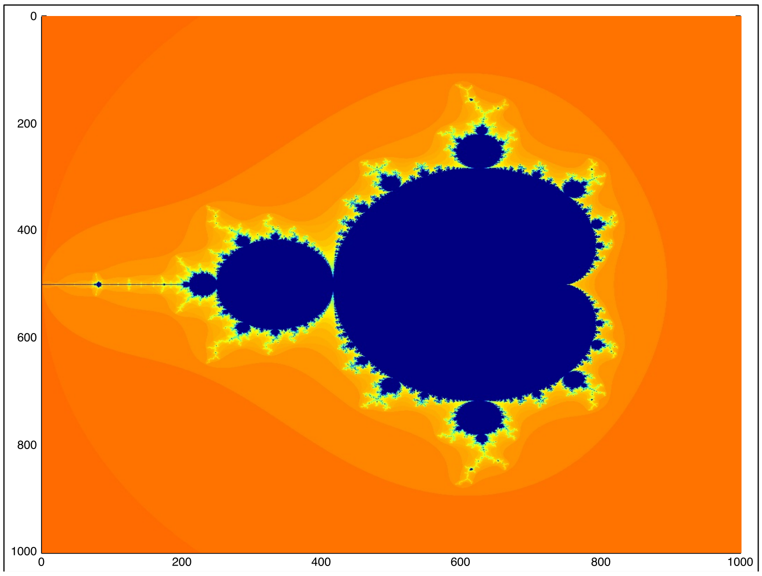


Abbildung 2.10.: Die Mandelbrotmenge (mit Farbverläufen)

```
image(M);
z=zx+i*zy;
c=-0.7;
M=zeros(m,n);
for k=1:200
    invalid=not(abs(z)<=2);
    M(invalid)=M(invalid)+0.1;
    z=z.^2+c;
end
subplot(2,2,3)
image(M);
z=zx+i*zy;
c=0.64*i;
M=zeros(m,n);
for k=1:200
    invalid=not(abs(z)<=2);
    M(invalid)=M(invalid)+0.1;
    z=z.^2+c;
end
subplot(2,2,4)
image(M);
```

Der Musiker Jonathan Coulton hat sogar ein Lied über die Mandelbrotmenge geschrieben, das man unter `http://www.jonathancoulton.com/wiki/Mandelbrot_Set` finden kann. Es gibt verschiedene Versionen des Liedes. In einigen wird fälschlicherweise nicht die Konstruktionsvorschrift für die Mandelbrotmenge, sondern die Juliamenge beschrieben. Zwei schöne Youtube-Videos zu dem Song finden Sie unter https://www.youtube.com/watch?v=ZDU40eUcTj0 und https://www.youtube.com/watch?v=ES-yKOYaXq0.

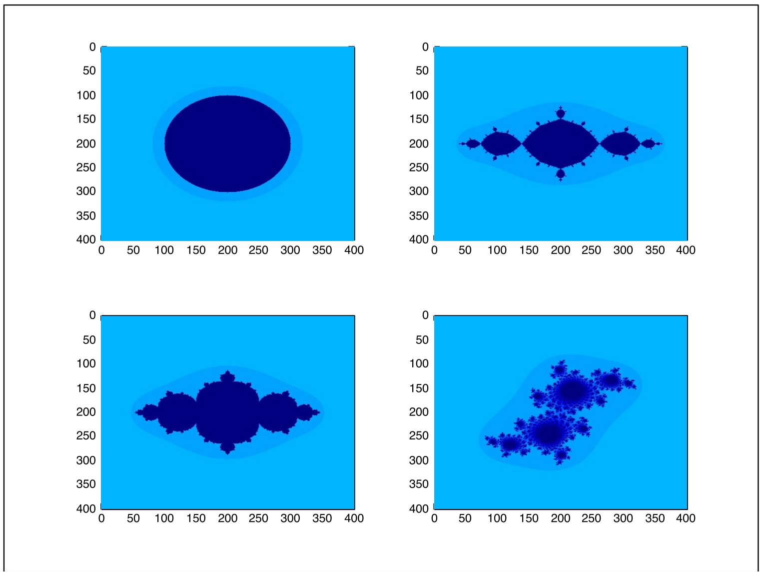


Abbildung 2.11.: Juliamengen zu $c = 0, c = -1, c = -0.7$ und $c = 0.64i$

In der Literatur werden auch Verallgemeinerungen der hier behandelten Mandelbrotmenge und Juliamengen untersucht, bei denen man die Rekursion

$$z_{n+1} := z_n^2 + c$$

durch eine andere Rekursion der Form

$$z_{n+1} := f(z_n, c)$$

mit einer geeigneten Funktion f ersetzt.

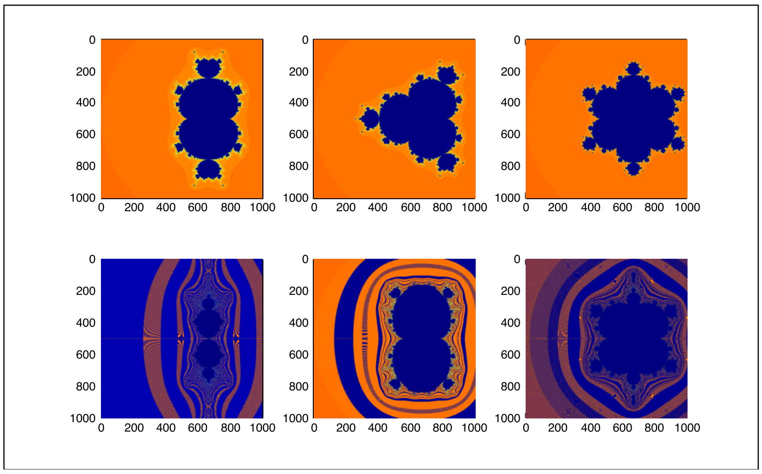


Abbildung 2.12.: Mandelbrotmengen zu $f(z,c) = z^3 + c, f(z,c) = z^4 + c, f(z,c) = z^7 + c, f(z,c) = z^3 + z + c, f(z,c) = z^5 + z^3 + c, f(z,c) = z^7 + z^5 + c.$

Mathematiker haben viele Jahre lang Fraktale untersucht ohne dabei konkrete Anwendungen im Kopf zu haben. Fraktale erscheinen wie andere Bereiche der Mathematik auch gerne wie eine nette mathematische Spielerei. Aber in der Tat hat man in den 1980er und 1990er Jahren auch praktische Anwendungen von Fraktalen oder Idee dahinter, nämlich aus einfachen Regeln komplexe Strukturen aufzubauen, die Selbstähnlichkeiten aufweisen, gefunden. So gibt es zum Beispiel verlustbehaftete Kompressionsverfahren für digitale Bilddateien, bei denen Selbstähnlichkeiten in den Bildern ausgenutzt werden. Man spricht hier *fraktaler Bildkompression*.

Ende des Exkurses

2.9. Reihen

Kann man einer Summe mit unendlich vielen Summanden einen Wert zuweisen? Ist

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

Oder

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

Oder macht dies alles überhaupt keinen Sinn?

Zunächst müssen wir erst einmal definieren, was wir unter einer Summe mit unendlich vielen Summenden verstehen. Und das kleine Beispiel zeigt schon, dass wir festlegen müssen, wie wir die Summenden aufsummieren.

Wir betrachten dazu eine Folge von ganz bestimmten endlichen Summen.

Definition 70 (Reihe)**:** Es sei (a_k) eine Folge in $\mathbb{K}$. Für $n \in \mathbb{N}$ setzen wir

$$S_n := \sum_{k=1}^{n} a_k.$$

Die Folge $(S_n)_{n\in\mathbb{N}}$ nennt man dann eine *Reihe* und S_n ist die n*-te Partialsumme (oder Teilsumme) der Reihe*. Man schreibt

$$\sum_{n=1}^{\infty} a_n$$

für die Folge $(S_n)_{n\in\mathbb{N}}$, unabhängig davon, ob die Folge konvergiert oder nicht. Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt *konvergent*, wenn die Folge $(S_n)_{n\in\mathbb{N}}$ konvergiert. Andernfalls heißt die Reihe *divergent*. Ist die Reihe konvergent, so wird auch ihr Grenzwert mit $\sum_{n=1}^{\infty} a_n$ bezeichnet. Ist die Reihe divergent, so wird dem Symbol $\sum_{n=1}^{\infty} a_n$ kein Zahlenwert zugeordnet.

BEMERKUNG: Das Symbol $\sum_{n=1}^{\infty} a_n$ hat also eine Doppelbedeutung. Es kann sowohl die Folge der Partialsummen als auch im Konvergenzfall den Grenzwert dieser Folge bezeichnen.

Problem 29: *Es seien (a_k) eine Folge in $\mathbb{K}$ und $n_0 \in \mathbb{N}$. Zeigen Sie, dass die Reihe $\sum_{n=1}^{\infty} a_n$ genau dann konvergiert, wenn die Reihe $\sum_{n=n_0}^{\infty} a_n$ konvergiert.*

Satz 71: Konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$, so ist die Folge (a_n) eine Nullfolge, das heißt

$$\lim_{n\to\infty} a_n = 0.$$

BEWEIS: Es sei $S_n := \sum_{k=1}^{n} a_k$ und (S_n) eine konvergente Reihe. Dann ist (S_n) eine Cauchy-Folge. Ist nun $\varepsilon > 0$ eine beliebige positive reelle Zahl, so existiert ein $n_0 \in \mathbb{N}$, so dass $|S_n - S_m| < \varepsilon$ für alle $n, m \geq n_0$ ist. Mit $m = n - 1$ ergibt dies

$$|a_n| = |S_n - S_{n-1}| < \varepsilon.$$

Alternativer Beweis:
Es sei $S_n := \sum_{k=1}^{n} a_k$. Nach Voraussetzung ist die Folge (S_n) konvergent. Sei $S := \lim_{n\to\infty} S_n$. Es ist

$$S_n - S_{n-1} = \sum_{k=1}^{n} a_k - \sum_{k=1}^{n-1} a_k = a_n$$

und somit aufgrund der Rechenregeln für Grenzwerte

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (S_n - S_{n-1}) = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = S - S = 0.$$

□

Beispiele:

1) *Geometrische Reihe:* Es sei $q \in \mathbb{C}$ fixiert. Die Reihe

$$\sum_{n=0}^{\infty} q^n = 1 + q + q^2 + q^3 + \ldots$$

heißt *geometrische Reihe*. Ist $|q| \geq 1$, so ist $(q^n)_{n\in\mathbb{N}}$ keine Nullfolge, also divergiert die Reihe.
Ist jedoch $|q| < 1$, so gilt (siehe Problem 4)

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

Daraus folgt also für $|q| < 1$

$$\sum_{k=0}^{\infty} q^k = \lim_{n\to\infty} S_n = \frac{1}{1-q}.$$

2) *Harmonische Reihe:* Die Reihe

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

heißt *harmonische Reihe*. Diese Reihe divergiert. Um dies zu beweisen, nehmen wir an, dass (S_n) konvergent ist und es sei $S := \lim_{n\to\infty} S_n$. Dann ist auch $\lim_{n\to\infty} S_{2n} = S$ und

$$\lim_{n\to\infty} (S_{2n} - S_n) = \lim_{n\to\infty} S_{2n} - \lim_{n\to\infty} S_n = S - S = 0.$$

Für alle natürlichen Zahlen n gilt aber

$$S_{2n} - S_n = \sum_{k=n+1}^{2n} \frac{1}{k} \geq n \cdot \frac{1}{2n} = \frac{1}{2}.$$

Wir erhalten einen Widerspruch, und somit muss unsere Annahme, dass die harmonische Reihe konvergiert, falsch gewesen sein.
Insbesondere zeigt dieses Beispiel, dass Satz 71 nur eine notwendige und keine hinreichende Bedingung für Konvergenz ist.

3) Unser drittes Beispiel ist die Reihe

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

Wir setzen wieder $S_n := \sum_{k=1}^{n} \frac{1}{k^2}$. Die Folge ist also offensichtlich monoton wachsend. Die Beschränkheit der Folge sieht man wie folgt

$$0 \leq \sum_{k=1}^{n} \frac{1}{k^2} = 1 + \sum_{k=2}^{n} \frac{1}{k^2} < 1 + \sum_{k=2}^{n} \frac{1}{k(k-1)} = 1 + \sum_{k=2}^{n} \left(\frac{1}{k-1} - \frac{1}{k} \right) = 1 + 1 - \frac{1}{n} < 2$$

Die Folge ist also beschränkt und monoton wachsend und damit konvergent.

Problem 30: *Untersuchen Sie die folgenden Reihen auf Konvergenz bzw. Divergenz:*

(i) $\sum_{n=1}^{\infty} (-1)^n \frac{n-1}{n}$,

(ii) $\sum_{n=0}^{\infty} \left(\frac{2}{5}\right)^n$,

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

Exkurs: Matlab-Experiment zur harmonischen Reihe

Die harmonische Reihe divergiert sehr langsam. Wir verdeutlichen dies anhand einiger Berechnungen mit Matlab. Die Partialsummen kann man mit der folgenden Matlabfunktion berechnen.

```
function erg=harmonische_reihe(N)
erg=0;
for i=1:N
    erg=erg+1/i;
end
```

Das Skript

```
clear all
format long
i=10;
while (i<1000000000000)
    fprintf('n=%d  Summe bis n=%f \n',i,harmonische_reihe(i));
    i=i*10;
end
```

hat die folgende Ausgabe

```
n=10  Summe bis n=2.928968
n=100  Summe bis n=5.187378
n=1000  Summe bis n=7.485471
n=10000  Summe bis n=9.787606
n=100000  Summe bis n=12.090146
n=1000000  Summe bis n=14.392727
n=10000000  Summe bis n=16.695311
n=100000000  Summe bis n=18.997896
n=1000000000  Summe bis n=21.300482
n=1.000000e+010  Summe bis n=22.064778
n=1.000000e+011  Summe bis n=22.064778
```

Man könnte angesichts der letzten beiden Zeilen zu dem Schluss kommen, dass die zugehörige Reihe konvergiert. Wir haben aber eben bewiesen, dass die Reihe nicht konvergiert. Mehr noch, wir haben gezeigt, dass sich die beiden letzten Werte um mehr als $\frac{1}{2}$ unterscheiden müssten. Der Computer liefert hier also fehlerhafte Werte, die wir aufgrund unseres mathematischen Wissens als fehlerhaft erkannt haben. Aber auch wenn unser Computer an dieser Stelle richtig gerechnet hätte, so wächst die harmonische Reihe doch so langsam an, dass man sich nie sicher wäre, ob sie nicht vielleicht doch konvergiert.

Um die harmonische Reihe zu zeichnen, ist die folgende Implementierung der harmonischen Reihe vorteilhafter.

```
function erg=harmonische_reihe1(N)
for i=1:length(N)
    erg(i)=0;
    for n=1:N(i)
```

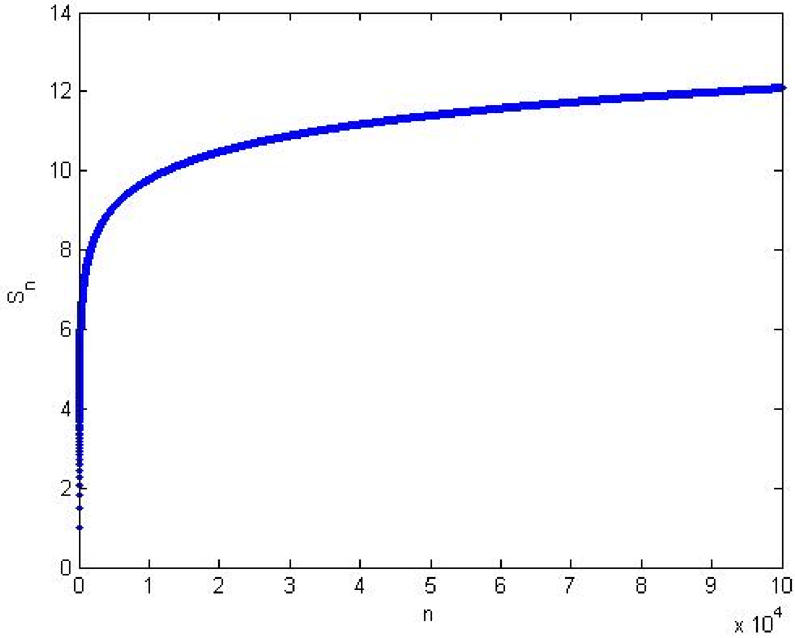


Abbildung 2.13.: Partialsummen der harmonischen Reihe

```
            erg(i)=erg(i)+1/n;
        end
end
```

Diese Funktion kann man dann wie im folgenden Beispiel in einem Skript benutzen.

```
clear all
figure(1)
clf()
n=1:100000;
plot(n,harmonische_reihe1(n),'.')
xlabel('n')
ylabel('S_n')
```

Ende des Exkurses

Satz 72 (Linearkombinationen konvergenter Reihen)**:** Sind

$$\sum_{n=1}^{\infty} a_n \text{ und } \sum_{n=1}^{\infty} b_n$$

zwei konvergente Reihen in $\mathbb{K}$ und ist $\lambda \in \mathbb{K}$, so konvergiert auch die Reihe

$$\sum_{n=1}^{\infty} (a_n + \lambda b_n)$$

und es gilt
$$\sum_{n=1}^{\infty}(a_n + \lambda b_n) = \sum_{n=1}^{\infty} a_n + \lambda \sum_{n=1}^{\infty} b_n.$$

BEWEIS: Die Partialsummen von $\sum a_n$ beziehungsweise $\sum b_n$ seien mit S_n beziehungsweise T_n bezeichnet. Dann gilt für alle natürlichen Zahlen n
$$S_n + \lambda T_n = \sum_{k=1}^{n}(a_k + \lambda b_k)$$
und die Konvergenz der Reihe $\sum(a_n + \lambda b_n)$ und die Gleichung im Satz folgen wegen der Konvergenz der Folgen (S_n) und (T_n) sofort aus den Rechenregeln für Grenzwerte. □

Satz 73 (Leibniz-Kriterium)**:** Ist (a_k) eine monotone Nullfolge reeller Zahlen, so konvergiert die Reihe
$$\sum_{k=1}^{\infty}(-1)^k a_k.$$

Problem 31: *Finden Sie eine Nullfolge (a_n) derart, dass die Reihe $\sum_{n=1}^{\infty}(-1)^n a_n$ divergiert. Die Monotonie der Folge muss im Leibniz-Kriterium also wirklich vorausgesetzt werden.*

Beispiel: Die sogenannte *alternierende harmonische Reihe*
$$\sum_{k=1}^{\infty}(-1)^k \frac{1}{k}$$
ist konvergent.

BEWEIS: Wir betrachten hier nur den Fall, dass (a_k) monoton fallend ist. Da die Folge (a_k) zudem eine Nullfolge sein soll, ist sie nach Lemma 52 nach unten durch Null beschränkt ist.

Es sei wieder
$$S_n := \sum_{k=1}^{n}(-1)^k a_k.$$

Wir zeigen nun, dass sowohl (S_{2n}) als auch (S_{2n+1}) konvergieren und dass die Grenzwerte übereinstimmen. Es gilt

$$\begin{aligned} S_{2n+2} - S_{2n} &= a_{2n+2} - a_{2n+1} \leq 0 \text{ und} \\ S_{2n+3} - S_{2n+1} &= -a_{2n+3} + a_{2n+2} \geq 0 \end{aligned}$$

Die Folgen (S_{2n}) und (S_{2n+1}) sind also monoton. Außerdem gilt

$$S_{2n+1} - S_{2n} = -a_{2n+1} \leq 0.$$

Hieraus folgt

$$S_2 \geq S_{2n} \geq S_{2n+1} \geq S_1$$

und somit sind beide Folgen beschränkt. Es existieren also reelle Zahlen s und $\tilde{s}$ mit

$$\begin{aligned} \lim_{n\to\infty} S_{2n} &= s \text{ und} \\ \lim_{n\to\infty} S_{2n+1} &= \tilde{s}. \end{aligned}$$

Wir wollen nun zeigen, dass $s = \tilde{s}$ gilt.

$$\begin{aligned} s - \tilde{s} &= \lim_{n\to\infty} S_{2n} - \lim_{n\to\infty} S_{2n+1} \\ &= \lim_{n\to\infty} (S_{2n} - S_{2n+1}) \\ &= \lim_{n\to\infty} a_{2n+1} = 0 \end{aligned}$$

Aus der Konvergenz der beiden Teilfolgen (S_{2n}) und (S_{2n+1}) gegen den selben Grenzwert s folgt, dass auch die Folge (S_n) gegen s konvergiert. Ist nämlich $\varepsilon > 0$ vorgegeben, so existieren natürliche Zahlen n_0, n_1, so dass

$$\begin{aligned} \forall n \geq n_0 &: |S_{2n} - s| < \varepsilon \\ \forall n \geq n_1 &: |S_{2n+1} - s| < \varepsilon \end{aligned}$$

gilt. Für alle $n \geq \max\{2n_0, 2n_1 + 1\}$ gilt also $|S_n - s| < \varepsilon$. □

Problem 32: *Untersuchen Sie die folgenden Reihen auf Konvergenz bzw. Divergenz:*

(i) $\displaystyle\sum_{n=1}^{\infty} \left((-1)^n \frac{1}{n} + \frac{1}{n} \right)$,

(ii) $\displaystyle\sum_{n=1}^{\infty} \left((-1)^n \frac{1}{n} + \frac{1}{n^2} \right)$,

(iii) $\displaystyle\sum_{n=1}^{\infty} \left((-1)^n \frac{1}{\sqrt{n}} + \frac{1}{n^2} \right)$,

(iv) $\displaystyle\sum_{n=1}^{\infty} \left(\frac{1}{\sqrt{n}} + (-1)^n \frac{1}{n^2} \right)$,

Exkurs: Die Kochsche Schneeflocke

Wir werden nun über wiederholtes Anwenden einer einfachen Konstruktionsvorschrift ein kompliziertes Objekt konstruieren.

Wir starten mit einer Strecke fester Länge ℓ. Wir teilen diese Strecke in drei gleich lange Teile, entfernen den mittleren der drei Teile und setzen an dieser Stelle zwei Strecken der Länge $\ell/3$ ein, die ein gleichseitiges Dreieck bilden. Wir haben jetzt einen Streckenzug mit vier Strecken der Länge $\ell/3$. Mit jeder der vier Teilstrecken wiederholen wir diese Konstruktion. Wir erhalten dann einen Streckenzug, der aus 16 Strecken besteht, von denen jede die Länge $\ell/9$ hat. Diese Konstruktionsvorschrift können wir nun wieder auf jede der Teilstrecken des Streckenzuges anwenden. In der folgenden Abbildung sieht man, was man nach 6 Wiederholungen der Konstruktionsvorschrift erhält.

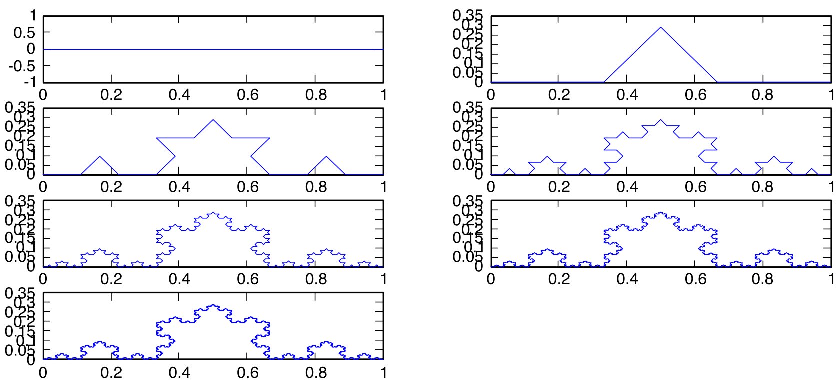


Abbildung 2.14.: 6fache Wiederholung der Konstruktionsvorschrift

Die Kochsche Schneeflocke erhält man nun, wenn man nicht mit einer Strecke, sondern mit einem gleichseitigen Dreieck startet und auf jede der drei Seiten und die dann entstehenden Seiten die obige Konstruktionsvorschrift unendlich oft anwendet. Wie die Mandelbrotmenge ist auch die Kochsche Schneeflocke ein Fraktal. Die Kochsche Schneeflocke ist ein geometrisches Objekt das man nicht zeichnen kann. Man kann nur Näherungen zeichnen, indem man die Konstruktion nach einer festen Anzahl von Schritten abbricht. Das folgende Skript erledigt dies.

```
clear all
x=[0 1 2 0];
y=[0 sqrt(3) 0 0];
figure(1)
clf()

N=8

subplot(N+1,3,1)
   plot(x, y)
   axis ([-1 2.5 -1 2.5],'square')
```

```
for n=1:N
for i=1:(length(x)-1)
    v= x(i+1)-x(i);
    w= y(i+1)-y(i);

    xneu(4*i-3)=x(i);
    yneu(4*i-3)=y(i);
    xneu(4*i-2)=x(i)+1/3*v;
    yneu(4*i-2)=y(i)+1/3*w;
    xneu(4*i)=x(i)+2/3*v;
    yneu(4*i)=y(i)+2/3*w;

    xneu(4*i-1)=xneu(4*i-2)+(cos(pi/3)*v-sin(pi/3)*w)/3;
    yneu(4*i-1)=yneu(4*i-2)+(sin(pi/3)*v+cos(pi/3)*w)/3;
 end
 x=[xneu 0];
 y=[yneu 0];

subplot(N+1,3,n+1)
  plot(x, y)
 axis ([-1 2.5 -1 2.5],'square')
 end
```

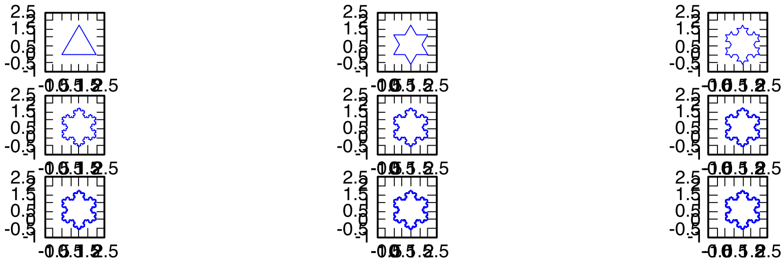


Abbildung 2.15.: Die ersten Näherungen der Kochschen Schneeflocke

Was hat jetzt die Kochsche Schneeflocke mit Folgen und Reihen zu tun? Wir können uns fragen, welchen Umfang und welchen Flächeninhalt die Kochsche Schneeflocke hat, wenn wir mit einem gleichseitigen Dreieck mit Seitenlänge ℓ starten. Wir bezeichnen mit u_n die Länge der Strecke, die die n-te Näherung der Kochschen Schneeflocke umrandet, und mit f_n deren Flächeninhalt. Dabei soll n-te Näherung bedeuten, dass man die obige Konstruktionsvorschrift n-mal auf die auftretenden Strecken angewendet hat. Das Dreieck ist also die 0-te Näherung und somit $u_0 = 3\ell$ und $f_0 = \sqrt{3}\ell^2/4$. In jedem Konstruktionsschritt wird jede Teilstrecke durch 4 Streckenstücke ersetzt, deren Länge $1/3$ der Länge der Teilstrecke ist. Also haben wir

$$u_n = \frac{4}{3}u_{n-1}$$

und somit

$$u_n = \left(\frac{4}{3}\right)^n 3\ell.$$

Das heißt aber, dass man dem Rand der Kochschen Schneeflocke wegen $\lim_{n\to\infty} u_n = +\infty$ keine sinnvolle Länge zuordnen kann.

Wie sieht es mit dem Flächeninhalt der n-ten Näherung aus? Die n-te Näherung besteht aus $4^n \cdot 3$ Teilstrecken. Der Fläche der n-ten Näherung ist gleich der Fläche der $n-1$-ten Näherung plus den Flächen der im n-ten Konstruktionsschritt neu angesetzten gleichseitigen Dreiecke. Da die $n-1$-te Näherung $4^{n-1} \cdot 3$ Teilstrecke hat, werden auch $4^{n-1} \cdot 3$ gleichseitige Dreiecke angesetzt. Jedes von diesen hat eine Seitenlänge von $\ell/3^n$, also einen Flächeninhalt von $\frac{1}{2}\frac{\ell}{3^n}\sqrt{\frac{3}{4}}\frac{\ell}{3^n} = \frac{\sqrt{3}\ell^2}{4}\left(\frac{1}{9}\right)^n$. Damit haben wir

$$f_n = \frac{\sqrt{3}}{4}\ell^2 + \sum_{k=1}^{n} 4^{k-1} 3 \frac{\sqrt{3}\ell^2}{4}\left(\frac{1}{9}\right)^k = \frac{\sqrt{3}}{4}\ell^2\left(1 + \sum_{k=1}^{n}\frac{1}{3}\left(\frac{4}{9}\right)^{k-1}\right).$$

Da die geometrische Reihe über $4/9$ konvergiert, ist der Flächeninhalt der Kochschen Schneeflocke gleich

$$\frac{\sqrt{3}}{4}\ell^2\left(1 + \frac{1}{3}\sum_{n=n}^{\infty}\left(\frac{4}{9}\right)^n\right) = \frac{\sqrt{3}\ell^2}{4}\left(1 + \frac{1}{3}\frac{1}{1-\frac{4}{9}}\right) = \frac{2\sqrt{3}}{5}\ell^2.$$

Die Kochsche Schneeflocke besitzt also einen endlichen Flächeninhalt, wird aber von einem unendlich langen Rand umschlossen!

Ende des Exkurses

2.10. Absolute Konvergenz und einige Konvergenzkriterien

Definition 74 (Absolute Konvergenz)**:** Eine Reihe $\sum_{k=1}^{\infty} a_k$ heißt *absolut konvergent*, wenn die Reihe

$$\sum_{k=1}^{\infty} |a_k|$$

konvergent ist.

Beispiel: Die alternierende harmonische Reihe $\sum(-1)^n\frac{1}{n}$ ist konvergent, aber nicht absolut konvergent.

Satz 75 (Dreiecksungleichung für absolut konvergente Reihen)**:** Eine absolut konvergente Reihe $\sum_{k=1}^{\infty} a_k$ ist auch konvergent (gemäß Definition 70) und für die Grenzwerte gilt die Dreiecksungleichung

$$\left|\sum_{k=1}^{\infty} a_k\right| \leq \sum_{k=1}^{\infty} |a_k|.$$

Beweis: Da $\sum_{k=1}^{\infty} |a_k|$ konvergiert, ist $S_n := \sum_{k=1}^{n} |a_k|$ eine Cauchy-Folge. Also existiert zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass für alle $m \geq n \geq n_0$

$$\left|\sum_{k=n+1}^{m} |a_k|\right| = |S_m - S_n| < \varepsilon$$

gilt. Auf Grund der Dreiecksungleichung folgt

$$\left|\sum_{k=n+1}^{m} a_k\right| \leq \sum_{k=n+1}^{m} |a_k| < \varepsilon.$$

Also ist auch $\sigma_n = \sum_{k=1}^{n} a_k$ eine Cauchy-Folge und die Reihe $\sum_{k=1}^{\infty} a_k$ konvergiert. Da (σ_n) konvergiert, konveriert auch $(|\sigma_n|)$ und es gilt

$$\lim_{n\to\infty} |\sigma_n| = \left|\sum_{k=1}^{\infty} a_k\right|.$$

Abermalige Anwendung der Dreiecksungleichung liefert uns

$$|\sigma_n| = \left|\sum_{k=1}^{n} a_k\right| \leq \sum_{k=1}^{n} |a_k| = S_n$$

und daraus folgt

$$\left|\sum_{k=1}^{\infty} a_k\right| = \lim_{n\to\infty} |\sigma_n| \leq \lim_{n\to\infty} S_n = \sum_{k=1}^{\infty} |a_k|.$$

□

Satz 76 (Majoranten- und Minorantenkriterium)**:** Es seien (a_n) und (b_n) zwei Folgen mit $|a_n| \leq |b_n|$ für alle $n \in \mathbb{N}$.

(i) *Majorantenkriterium:* Ist $\sum_{k=1}^{\infty} b_k$ eine absolut konvergente Reihe, so ist auch die Reihe $\sum_{k=1}^{\infty} a_k$ absolut konvergent.

(ii) *Minorantenkriterium:* Ist die Reihe $\sum_{k=1}^{\infty} |a_k|$ divergent, so divergiert auch die Reihe $\sum_{k=1}^{\infty} |b_k|$.

BEWEIS:

(i) Die Folgen (S_n) und (T_n) mit

$$S_n := \sum_{k=1}^{n} |a_k| \text{ und } T_n := \sum_{k=1}^{n} |b_k|$$

sind offensichtlich monoton wachsend. Aufgrund der Voraussetzung $|a_n| \leq |b_n|$ für alle $n \in \mathbb{N}$ haben wir

$$S_n \leq T_n.$$

Da (T_n) eine monotone, konvergente Folge ist, gilt wegen Lemma 52

$$T_n \leq \lim_{n\to\infty} T_n = \sum_{k=1}^{\infty} |b_k|.$$

Die Folge (S_n) ist also auch beschränkt und somit konvergent.

Alternativer Beweis: Es sei $\varepsilon > 0$ beliebig. Da die Reihe $\sum_{k=1}^{\infty} b_k$ absolut konvergiert, existiert ein $n_0 \in \mathbb{N}$, so dass für alle $m \geq n \geq n_0$

$$\sum_{k=n+1}^{m} |b_k| < \varepsilon$$

ist. Auf Grund der Vorraussetzung $|a_n| \leq |b_n|$ folgt

$$\sum_{k=n+1}^{m} |a_k| < \varepsilon.$$

Also ist $(\sum_{k=1}^{n} |a_k|)_{n\in\mathbb{N}}$ eine Cauchy-Folge und insbesondere konvergent, da $\mathbb{R}$ vollständig ist.

(ii) Die Behauptung folgt sofort aus (i).

□

Beispiel: Wir wissen, dass die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ konvergiert. Für $k \geq 2$ und $n \geq 1$ gilt $\frac{1}{n^k} \leq \frac{1}{n^2}$. Also sind für alle $k \geq 2$ die Reihen

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

konvergent.

Problem 33: *(i) Für eine Folge positiver reeller Zahlen (a_n) bezeichne $\Theta(a_n)$ die Menge der Folgen positiver Zahlen (b_n) mit der Eigenschaft, dass positive reelle Konstanten c_1, c_2 und ein $n_0 \in \mathbb{N}$ existieren, so dass für alle $n \geq n_0$*

$$c_1 a_n \leq b_n \leq c_2 a_n$$

gilt.
Zeigen Sie: Ist $(b_n) \in \Theta(a_n)$, so sind die beiden Reihen $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ entweder beide absolut konvergent oder beide divergent.

(ii) Zeigen Sie, dass $(b_n) \in \Theta(a_n)$ ist, falls der Grenzwert $\lim_{n\to\infty} \frac{a_n}{b_n}$ existiert und ungleich 0 ist.

(iii) Untersuchen Sie die Reihe $\displaystyle\sum_{n=1}^{\infty} \frac{n^2+n}{n^4+11n^2+3}$ auf Konvergenz bzw. Divergenz.

Wir halten das Ergebnis der vorstehenden Aufgabe in einem Satz fest.

Satz 77 (Vergleichskriterium)**:** Sind (a_n) und (b_n) zwei reelle Zahlenfolgen mit

$$\lim_{n\to\infty} \left|\frac{a_n}{b_n}\right| > 0,$$

so konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$ genau dann absolut, wenn die Reihe $\sum_{n=1}^{\infty} b_n$ absolut konvergiert.

Problem 34: *(i) Es sei $\sum_{n=1}^{\infty} a_n$ eine absolut konvergente Reihe und $(c_n)_{n\in\mathbb{N}}$ eine beschränkte Folge reeller Zahlen. Zeigen Sie, dass die Reihe $\sum_{n=1}^{\infty} c_n a_n$ absolut konvergiert.*

(ii) Geben Sie ein Beispiel einer konvergenten Reihe $\sum_{n=1}^{\infty} a_n$ und einer konvergenten Folge $(c_n)_{n\in\mathbb{N}}$ an, so dass die Reihe $\sum_{n=1}^{\infty} c_n a_n$ divergiert.

Wir kommen nun zu zwei sehr wichtigen Konvergenzkriterien für Reihen.

Satz 78 (Quotientenkriterium)**:** Es sei (a_n) eine Folge in $\mathbb{K}$ mit $a_n \neq 0$ für alle $n \geq n_0$. Es gebe eine reelle Zahl θ mit $0 < \theta < 1$, so dass für alle $n \geq n_0$ die Ungleichung

$$\left|\frac{a_{n+1}}{a_n}\right| \leq \theta$$

erfüllt ist. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$ absolut.

BEWEIS: Es sei $\left|\frac{a_{n+1}}{a_n}\right| \leq \theta$ für alle $n \in \mathbb{N}$, dann gilt

$$|a_{n+1}| \leq \theta|a_n| \leq \theta^2|a_{n-1}| \leq \ldots \leq \theta^n|a_1|$$

für alle $n \in \mathbb{N}$. Durch Anwenden der geometrischen Reihe folgt

$$\sum_{n=0}^{\infty} |a_1|\theta^n = \frac{|a_1|}{1-\theta}.$$

Dies ist eine konvergente Majorante für $\sum_{k=1}^{\infty} |a_k|$, woraus die Behauptung folgt. $\square$

Beispiele:

1) Wir betrachten die Reihe $\sum_{n=1}^{\infty} a_n$ mit
$$a_n := \frac{n^2}{2^n}$$
und wenden darauf das Quotientenkriterium an. Es ist
$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{\frac{(n+1)^2}{2^{n+1}}}{\frac{n^2}{2^n}} = \frac{(n+1)^2}{2^{n+1}} \cdot \frac{2^n}{n^2} = \frac{1}{2} \cdot \left(\frac{n+1}{n}\right)^2 = \frac{1}{2}\left(1 + \frac{1}{n}\right)^2$$
Für $n \geq 3$ folgt
$$\left|\frac{a_{n+1}}{a_n}\right| \leq \frac{1}{2}\left(\frac{4}{3}\right)^2 = \frac{8}{9}.$$
In diesem Fall ist das Quotientenkriterium mit $\theta = \frac{8}{9}$ anwendbar.

2) Wir betrachten abermals die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$. Von dieser wissen wir, dass sie divergiert. Für den Quotienten gilt
$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{n+1} \cdot \frac{n}{1} < 1.$$
Jedoch gibt es kein $\theta < 1$ mit $\frac{n}{n+1} \leq \theta < 1$. Dieses Beispiel zeigt, dass im Quotientenkriterium die Voraussetzung $\left|\frac{a_{n+1}}{a_n}\right| \leq \theta < 1$ nicht zu $\left|\frac{a_{n+1}}{a_n}\right| < 1$ abgeschwächt werden kann.

3) Wie wir wissen konvergiert die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$. Auch hier gilt
$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 1.$$
Es existiert also kein θ wie in Satz 78. Das Quotientenkriterium bietet also nur ein hinreichendes jedoch kein notwendiges Kriterium für die absolute Konvergenz einer Reihe.

Häufig kann man die folgende etwas leichter zu handhabende Variante des Quotientenkriteriums anwenden.

Korollar 79: Existiert in der Situation des vorstehenden Satzes der Grenzwert
$$\theta := \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|,$$
so gilt:

- Ist $\theta < 1$, so konvergiert die Reihe $\sum a_n$ absolut.
- Ist $\theta > 1$, so divergiert die Reihe $\sum a_n$.
- Im Fall $\theta = 1$ ist keine Aussage möglich.

Problem 35: *Beweisen Sie das vorstehende Korollar.*

Satz 80 (Wurzelkriterium)**:** Es sei (a_n) eine Folge in $\mathbb{K}$. Es gebe eine reelle Zahl θ mit $0 < \theta < 1$ und ein $n_0 \in \mathbb{N}$, so dass

$$\sqrt[n]{|a_n|} \leq \theta$$

für alle $n \geq n_0$ gilt. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$ absolut.

BEWEIS: Die Bedingung $\sqrt[n]{|a_n|} \leq \theta$ ist äquivalent zu $|a_n| \leq \theta^n$. Wieder dient die geometrische Reihe als Majorante. Also gilt

$$\sum_{n=1}^{\infty} |a_n| \leq \sum_{n=1}^{\infty} \theta^n.$$

□

Beispiel: Wir wissen, dass die Reihe $\sum_{n=1}^{\infty} n^{-k}$ für $k \geq 2$ absolut konvergiert. Wir wollen zeigen, dass das Wurzelkriterium hier jedoch nicht anwendbar ist.
Setze $b_n := \sqrt[n]{n} - 1$. Wir werden zeigen, dass dies eine Nullfolge ist.

$$n = \left(\sqrt[n]{n}\right)^n = (1 + b_n)^n \geq 1 + \binom{n}{2} b_n^2$$

Das $\geq$-Zeichen ergibt sich durch Anwenden der allgemeinen binomischen Formel und weglassen der gemischten Terme. Hiermit erhält man also

$$b_n^2 \leq \frac{n-1}{\binom{n}{2}} = \frac{n-1}{\frac{n(n-1)}{2}} = \frac{2}{n} \xrightarrow[n\to\infty]{} 0.$$

Weiter folgt

$$\sqrt[n]{\frac{1}{n}} = \frac{1}{\sqrt[n]{n}} \xrightarrow[n\to\infty]{} 1$$

und für festes $k \geq 2$ also

$$\sqrt[n]{\frac{1}{n^k}} = \left(\frac{1}{\sqrt[n]{n}}\right)^k \xrightarrow[n\to\infty]{} 1.$$

Hieraus sieht man, dass auch das Wurzelkriterium nur ein hinreichendes, aber kein notwendiges Kriterium für die absolute Konvergenz einer Reihe ist.

Wir halten noch die folgende Erkenntnis aus dem vorstehenden Beispiel fest.

Lemma 81: Es ist
$$\lim_{n\to\infty} \sqrt[n]{n} = 1.$$

Wir notieren auch für das Wurzelkriterium eine etwas einfachere Variante als Korollar.

Korollar 82: Existiert in der Situation des vorstehenden Satzes der Grenzwert
$$\theta := \lim_{n\to\infty} \sqrt[n]{|a_n|},$$
so gilt:

- Ist $\theta < 1$, so konvergiert die Reihe $\sum a_n$ absolut.
- Ist $\theta > 1$, so divergiert die Reihe $\sum a_n$.
- Im Fall $\theta = 1$ ist keine Aussage möglich.

Problem 36: *Beweisen Sie das vorstehende Korollar.*

Beispiel: Die Reihe
$$\sum_{n=0}^{\infty} \frac{n}{(-2)^n}$$
konvergiert absolut, denn es ist
$$\sqrt[n]{\left|\frac{n}{(-2)^n}\right|} = \sqrt[n]{\frac{n}{2^n}} = \frac{\sqrt[n]{n}}{\sqrt[n]{2^n}} \xrightarrow[n\to\infty]{} \frac{1}{2}.$$

Problem 37: *Untersuchen Sie die Reihe*
$$\sum_{n=1}^{\infty} \frac{2+(-1)^n}{2^n}$$
sowohl mit dem Quotienten- als auch mit dem Wurzelkriterium.

Exkurs: Der Limes superior und der Limes inferior

Für $\emptyset \neq X \subset \mathbb{R}$ haben wir bereits zuvor definiert, was wir unter dem Supremum und dem Infimum von X verstehen. Falls X nicht nach oben beschränkt ist, dann existiert $\sup X$

nicht. Wir schreiben in diesem Fall $\sup X := \infty$. Falls X nicht nach unten beschränkt ist, schreiben wir $\inf X := -\infty$. Für eine reelle Folge $(a_n)_{n\in\mathbb{N}}$ definieren wir nun

den *Limes superior* als $\overline{\lim}\, a_n = \limsup_{n\to\infty} a_n := \lim_{n\to\infty} (\sup\{a_k | k \geq n\}) \in \mathbb{R} \cup \{\infty\}$

und

den *Limes inferior* als $\underline{\lim}\, a_n = \liminf_{n\to\infty} a_n := \lim_{n\to\infty} (\inf\{a_k | k \geq n\}) \in \mathbb{R} \cup \{-\infty\}$.

Problem 38: *Zeigen Sie, dass eine Folge (a_n) genau dann gegen $a \in \mathbb{R}$ konvergiert, wenn*

$$\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = a.$$

Mit dieser Begriffsbildung lassen sich das Wurzel- und das Quotientenkriterium wie folgt formulieren:

Problem 39: *(i) Ist $(a_k)_{k\in\mathbb{N}}$ eine Folge komplexer Zahlen mit*

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} < 1,$$

dann konvergiert die Reihe $\sum_{k=1}^{\infty} a_k$ absolut. Ist der vorstehende $\limsup$ größer 1, so divergiert die Reihe.

(ii) Ist $(a_k)_{k\in\mathbb{N}}$ eine Folge komplexer Zahlen mit $a_k \neq 0$ für alle $k \in \mathbb{N}$ und

$$\limsup_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| < 1,$$

dann konvergiert die Reihe $\sum_{k=1}^{\infty} a_k$ absolut. Ist

$$\liminf_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| > 1,$$

so divergiert die Reihe.

Ende des Exkurses

Exkurs: Der Riemannsche Umordnungssatz

Wir wollen die Bedeutung der absoluten Konvergenz einer Reihe noch etwas besser herausarbeiten. Dazu führen wir zunächst einen neuen Begriff ein.

Definition 83: Es sei $\sigma : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Dann nennt man die Reihe $\sum_{n=1}^{\infty} a_{\sigma(n)}$ eine *Umordnung* der Reihe $\sum_{n=1}^{\infty} a_n$.

Nun stellt sich die Frage, ob jede Umordnung einer konvergenten Reihe auch wieder konvergent ist und ob sie gegen den gleichen Grenzwert konvergiert. Wir wollen uns überlegen, dass dies für konvergente, aber nicht absolut konvergente Reihen nicht gilt. Hier gilt der folgende faszinierende Satz.

Satz 84 (Riemannscher Umordnungssatz)**:** Es sei $\sum_{n=1}^{\infty} a_n$ eine konvergente, aber nicht absolut konvergente Reihe. Dann existiert zu jeder reellen Zahl S eine bijektive Abbildung $\sigma : \mathbb{N} \to \mathbb{N}$ derart, dass die Reihe $\sum_{n=1}^{\infty} a_{\sigma(n)}$ konvergiert und den Grenzwert S hat.

Wie werden den Satz nicht beweisen, wollen aber die Idee hinter dem Beweis darstellen. Für eine reelle Zahl a setzen wir

$$a^+ := \begin{cases} a & \text{, falls } a \geq 0, \\ 0, & \text{sonst,} \end{cases}$$

und

$$a^- := \begin{cases} 0 & \text{, falls } a \geq 0, \\ -a, & \text{sonst.} \end{cases}$$

Für jede reelle Zahl a gilt dann

- $a^+ - a^- = a,$
- $a^+ + a^- = |a|,$
- $0 \leq a^+, a^- \leq |a|.$

Lemma 85: Ist $\sum_{n=1}^{\infty} a_n$ eine konvergente, aber nicht absolut konvergente Reihe, so divergieren die beiden Reihen $\sum_{n=1}^{\infty} a_n^+$ und $\sum_{n=1}^{\infty} a_n^-$; genauer divergieren die beiden Reihen bestimmt gegen $+\infty$.

Problem 40: *Beweisen Sie das vorstehende Lemma.*

Problem 41: *Beweise oder widerlege: Konvergiert eine der beiden Reihen $\sum_{n=1}^{\infty} a_n^+$ und $\sum_{n=1}^{\infty} a_n^-$, so ist $\sum_{n=1}^{\infty} a_n$ eine absolut konvergente Reihe.*

Die Idee zum Beweis des Riemannschen Umordnungssatzes ist nun die folgende: Man summiert solange Elemente der Folge (a_n^+) bis der Wert der Summe den Wert S übersteigt. Dann subtrahiert man solange Elemente der Folge (a_n^-) bis der Wert der Summe den Wert S unterschreitet. Dann summiert man wieder solange Elemente der Folge (a_n^+) bis der Wert der Summe den Wert S übersteigt u.s.w.

Bei absolut konvergenten Reihen kann einem dies nicht passieren. Dort gilt der folgende Satz, den wir aber nicht beweisen wollen.

Satz 86: Ist $\sum_{n=1}^{\infty} a_n$ eine absolut konvergente Reihe, so konvergiert auch jede Umordnung $\sum_{n=1}^{\infty} a_{\sigma(n)}$ absolut und hat den selben Grenzwert.

Ende des Exkurses

2.11. Dezimalbrüche und b-adische Darstellung reeller Zahlen

Definition 87: Es sei $k \in \mathbb{N}_0$. Eine Reihe der Form

$$\pm \sum_{n=-k}^{\infty} a_n 10^{-n}$$

mit $a_n \in \{0, 1, 2, \dots, 9\}$ für $n \in \mathbb{Z}, n \geq -k$, heißt *Dezimalbruch*. Üblicherweise schreibt man hierfür $\pm a_{-k} \dots a_{-1} a_0, a_1 a_2 a_3 \dots$.

BEMERKUNG: 10 ist natürlich selber ein Dezimalbruch mit $a_{-1} = 1$ und $a_0 = 0$. In der den Dezimalbruch definierenden Reihe in der vorstehenden Definition ist 10 als ein Zeichen für die Zahl $1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1$ zu lesen.

Problem 42: *Überlegen Sie sich, dass sowohl der Dezimalbruch* $1,00000000\dots$ *als auch der Dezimalbruch* $0,99999\dots$ *gegen die Zahl 1 konvergieren.*

Definition 88: Ein Dezimalbruch $x = a_{-k} \dots a_{-1} a_0, a_1 a_2 a_3 \dots$ heißt *periodisch*, wenn es $M, N \in \mathbb{N}$ gibt, so dass $a_{n+N} = a_n$ für alle $n \geq M$ gilt.

Satz 89: Jeder Dezimalbruch konvergiert, das heißt jeder Dezimalbruch stellt eine reelle Zahl dar.

Problem 43: *Man beweise den Satz, indem man eine konvergente Majorante angibt.*

Man stellt sich jetzt natürlich die Frage, ob die Umkehrung dieses Satzes auch gilt, also ob sich jede reelle Zahl durch einen Dezimalbruch darstellen lässt.

Definition 90: Für reelles x bezeichne $[x]$ die größte ganze Zahl, die kleiner oder gleich x ist. Man nennt $[x]$ die *Gauß-Klammer* von x.

Satz 91: Sei $x \in [0,1[$ beliebig. Wir definieren die Folge (a_n) rekursiv durch

$$a_1 := [10x] \text{ und } a_n := \left[10^n \left(x - \sum_{k=1}^{n-1} a_k 10^{-k}\right)\right].$$

Dann konvergiert der Dezimalbruch $0, a_1 a_2 a_3 \ldots$ gegen x.

BEWEIS: Zunächst halten wir fest, dass für jede reelle Zahl a gilt:

$$a - 1 < [a] \leq a.$$

Nun schauen wir uns eine Partialsumme des Dezimalbruchs an und schätzen diese nach oben und nach unten ab. Wir haben

$$\begin{aligned}\sum_{k=1}^{n} a_k 10^{-k} &= \sum_{k=1}^{n-1} a_k 10^{-k} + a_n 10^{-n} \\ &\leq \sum_{k=1}^{n-1} a_k + 10^n \left(x - \sum_{k=1}^{n-1} a_k 10^{-k}\right) 10^{-n} \\ &= x.\end{aligned}$$

Andererseits ergibt sich

$$\begin{aligned}\sum_{k=1}^{n} a_k 10^{-k} &= \sum_{k=1}^{n-1} a_k 10^{-k} + a_n 10^{-n} \\ &\geq \sum_{k=1}^{n-1} a_k + \left(10^n \left(x - \sum_{k=1}^{n-1} a_k 10^{-k}\right) - 1\right) 10^{-n} \\ &= x - 10^{-n}.\end{aligned}$$

Aus dem Schachtelungsprinzip ergibt sich nun die Behauptung des Satzes. □

Korollar 92: Jede reelle Zahl lässt sich als Dezimalbruch darstellen.

BEMERKUNG: Die Dezimalbruchdarstellung einer reellen Zahl ist aber, wie wir zuvor gesehen haben, im Allgemeinen nicht eindeutig. Wenn man Dezimalbrüche mit Periode 9 nicht zulässt und diese immer aufrundet, d.h. als Dezimalbruch mit Periode 0 darstellt, dann ist die Dezimalbruchdarstellung einer reellen Zahl eindeutig.

Problem 44: *Zeigen Sie: Sind $a = 0,a_1a_2a_3\ldots$ und $b = 0,b_1b_2b_3\ldots$ zwei verschiedene Dezimalbruchdarstellungen der selben reellen Zahl x, so gibt es eine natürliche Zahl N, so dass*

$$\begin{aligned} a_n &= b_n, \text{ für } n = 1,\ldots,N-1 \\ a_N &= b_N + 1 \\ a_n = 0 \text{ und } b_n &= 9, \text{ für alle } n > N, \end{aligned}$$

wobei man eventuell die Rollen von a und b vertauschen muss.

Korollar 93: Zu jeder reellen Zahl x gibt es eine Folge (a_n) in $\mathbb{Q}$ mit

$$\lim_{n\to\infty} a_n = x.$$

Allgemeiner gilt:

Satz 94: Ist b eine natürliche Zahl größer als 1, so lässt sich jede reelle Zahl x als Reihe der Form

$$\pm \sum_{n=-k}^{\infty} a_n b^{-n}$$

mit $a_n \in \{0, 1, 2, \ldots, b-1\}$ für $n \in \mathbb{Z}, n \geq -k$, schreiben. Man nennt diese Reihe dann, die *b-adische Darstellung von x*.

Satz 95: Für $p, q \in \mathbb{N}$ mit $p < q$ seien die Folgen $(a_n), (r_n)$ in $\mathbb{N}_0$ rekursiv definiert durch

$$r_1 := p \text{ und } 10r_n = a_n q + r_{n+1} \text{ mit } 0 \leq r_{n+1} < q.$$

Dann konvergiert der Dezimalbruch $0,a_1a_2a_3\ldots$ gegen $\frac{p}{q}$.

BEWEIS: Zunächst überlegen wir uns, dass die Definition der Folgen (a_n) und (r_n) sinnvoll ist und die beiden Folgen eindeutig durch die obigen Vorschriften gegeben sind. Ist r_n für ein $n \in \mathbb{N}$ bereits bekannt, so existieren zu r_n und q nach dem Satz über die Division mit Rest eindeutig bestimmte Zahlen a_n, r_{n+1} mit

$$10r_n = a_n q + r_{n+1} \text{ und } 0 \leq r_{n+1} < q.$$

Also sind für alle $n \in \mathbb{N}$ die Folgenglieder a_n und r_n eindeutig festgelegt.

Wir zeigen nun per vollständiger Induktion, dass für alle natürlichen Zahlen n

$$\sum_{k=1}^{n} a_k 10^{-k} = \frac{p}{q} - \frac{r_{n+1}}{10^n q}$$

gilt, woraus dann die Behauptung sofort folgt.

Für $n = 1$ ist nach Definition der Folgen (a_n) und (r_n)

$$a_1 q = 10 r_1 - r_2 = 10p - r_2,$$

woraus nach Division durch $10q$

$$a_1 10^{-1} = \frac{p}{q} - \frac{r_2}{10^1 q}$$

folgt. Also ist die Induktionsverankerung gelungen.

Sei nun die Behauptung für n bereits bewiesen. Wir müssen zeigen, dass dann auch

$$\sum_{k=1}^{n+1} a_k 10^{-k} = \frac{p}{q} - \frac{r_{n+2}}{10^{n+1} q}$$

gilt. Es ist

$$a_{n+1} 10^{-(n+1)} - \frac{r_{n+1}}{10^n q} = \frac{1}{10^{n+1} q} (a_{n+1} q - 10 r_{n+1}) = -\frac{r_{n+2}}{10^{n+1} q}.$$

Unter Anwendung der Induktionsvoraussetzung erhalten wir damit

$$\begin{aligned}\sum_{k=1}^{n+1} a_k 10^{-k} &= \sum_{k=1}^{n} a_k 10^{-k} + a_{n+1} 10^{-(n+1)} \\ &= \frac{p}{q} - \frac{r_{n+1}}{10^n q} + a_{n+1} 10^{-(n+1)} = \frac{p}{q} - \frac{r_{n+2}}{10^{n+1} q}.\end{aligned}$$

□

Satz 96: Eine reelle Zahl x besitzt genau dann eine periodische Dezimalbruchdarstellung, wenn $x \in \mathbb{Q}$ gilt.

BEWEIS: Ist x eine rationale Zahl, dann kann man eine Dezimalbruchdarstellung mittels Satz 95 für x gewinnen. Diese ist sicherlich periodisch, da die Folge (r_n) nur endlich viele verschiedene Werte annimmt und r_n eindeutig die Werte von a_n und r_{n+1} bestimmt. Ist also $r_n = r_{n+N}$, so sind auch $a_n = a_{n+N}$ und $r_{n+1} = r_{n+N+1}$.

Besitzt anderseits x eine periodische Dezimalbruchdarstellung, so existieren also natürliche Zahlen M, N, so dass $a_{n+N} = a_n$ für alle $n \geq M$ gilt. Es sei $k_0 \in \mathbb{N}$ derart, dass $k_0 N \geq M$ ist. Dann gilt

$$\begin{aligned}\sum_{n=k_0 N}^{\infty} a_n 10^{-n} &= \sum_{k=k_0}^{\infty} \sum_{\ell=0}^{N-1} a_{kN+\ell} 10^{-(kN+\ell)} = \sum_{k=k_0}^{\infty} \sum_{\ell=0}^{N-1} a_{k_0 N+\ell} 10^{-(kN+\ell)} \\ &= \sum_{k=k_0}^{\infty} (10^{-N})^k \sum_{\ell=0}^{N-1} a_{k_0 N+\ell} 10^{-\ell}.\end{aligned}$$

Dies ist aber aufgrund der Formel für die geometrische Reihe eine rationale Zahl, denn es ist

$$\begin{aligned}\sum_{k=k_0}^{\infty}(10^{-N})^k &= \sum_{k=0}^{\infty}(10^{-N})^k - \sum_{k=0}^{k_0-1}(10^{-N})^k \\ &= \frac{1}{1-10^{-N}} - \frac{1-10^{-Nk_0}}{1-10^{-N}} = \frac{10^{-Nk_0}}{1-10^{-N}}.\end{aligned}$$

□

Exkurs: Abzählbarkeit und Überabzählbarkeit

Wir haben früher gezeigt, dass die rationalen Zahlen $\mathbb{Q}$ dicht in $\mathbb{R}$ liegen.

Definition 97 (Abzählbarkeit)**:** Eine Menge A heißt *abzählbar*, wenn es eine Folge $(x_n)_{n\in\mathbb{N}}$ gibt, so dass

$$A = \{x_n | n \in \mathbb{N}\}.$$

Mit anderen Worten: Eine Menge A ist genau dann abzählbar, wenn es eine surjektive Abbildung $f : \mathbb{N} \to A$ gibt.

Beispiel: Die ganzen Zahlen $\mathbb{Z}$ sind abzählbar. Wir wählen $x_1 = 0, x_2 = 1, x_3 = -1, x_4 = 2, x_5 = -2, \ldots$ bzw. $x_{2k} = k$ und $x_{2k-1} = -(k-1)$ für alle $k \in \mathbb{N}$.

Lemma 98: Die Mengen $\mathbb{N}^n, n \in \mathbb{N}$, sind abzählbar.

BEWEIS: Für $n = 2$ geht man wie folgt vor: Man zählt alle Tupel $(x, y) \in \mathbb{N}^2$ der Reihe nach mit steigender Komponenten Summe $x+y$ auf. Bei gleicher Komponentensumme zählt man die Tupel mit steigendem x auf. Also:

$$(1,1),(1,2),(2,1),(1,3),(2,2),(3,1),\ldots.$$

Für $n > 2$ folgt die Abzählbarkeit von $\mathbb{N}^n$ dann induktiv.

□

Satz 99: $\mathbb{Q}$ ist abzählbar.

BEWEIS: Die Abbildung $\mathbb{N}_0^3 \to \mathbb{Q}, (i,j,k) \mapsto \frac{i-j}{1+k}$ ist surjektiv. $\square$

Satz 100: Das Intervall $I \; :=]0,1[$ ist nicht abzählbar und damit sind auch die reellen Zahlen $\mathbb{R}$ nicht abzählbar.

BEWEIS: Wir müssen zeigen, dass zu jeder Folge (x_n) in I mindestens ein $a \in I$ existiert mit $a \notin \{x_n | n \in \mathbb{N}\}$. Die Dezimalbruchentwicklung von x_n sei $0, x_{n,1}x_{n,2}x_{n,3}\ldots$. Wir setzen nun

$$a_n := \begin{cases} 2 & , \text{ falls } x_{n,n} = 1 \\ 1 & , \text{ falls } x_{n,n} \neq 1 \end{cases}$$

und a sei der Dezimalbruch $0, a_1a_2a_3a_4\ldots$. Offensichtlich ist $a \in I$. a unterscheidet sich aber von jedem x_n spätestens an der n-ten Nachkommastelle.

$\square$

Ende des Exkurses

2.12. Das Cauchy-Produkt

Definition 101 (Cauchy-Produkt)**:** Es seien (a_n) und (b_n) Folgen in $\mathbb{K}$. Dann heißt die Reihe

$$\sum_{k=1}^{\infty} \left(\sum_{j=1}^{k} a_j b_{k-j+1} \right)$$

das *Cauchy-Produkt* der Reihen $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$.

BEMERKUNG: Beginnen die Reihen bei $n = 0$, so lautet das Cauchy-Produkt von $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$

$$\sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} a_j b_{k-j} \right).$$

Satz 102: Sind $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ absolut konvergente Reihen, so ist auch das Cauchy-Produkt absolut konvergent und es gilt

$$\sum_{n=1}^{\infty} a_n \cdot \sum_{m=1}^{\infty} b_m = \sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} a_n b_m \right) = \sum_{m=1}^{\infty} \left(\sum_{n=1}^{\infty} a_n b_m \right) = \sum_{k=1}^{\infty} \left(\sum_{j=1}^{k} a_j b_{k-j+1} \right).$$

BEWEIS: Dieser Satz wird hier nicht bewiesen. Der Beweis ist sehr technisch und findet sich zum Beispiel in [For08a]. □

2.13. Potenzreihen

In der Schule haben Sie in Analysis wahrscheinlich vor allem Funktionen untersucht. Auch wir wollen bald hierzu kommen. Es ist aber ziemlich langweilig über Funktionen zu reden, wenn man außer Polynomfunktionen und rationalen Funktionen keine interessanten Funktionen kennt. Viele interessante Funktionen kann man über Reihen einführen, etwa über Potenzreihen, die wir im Folgenden behandeln wollen.

Definition 103 (Potenzreihe)**:** Es sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{K}$ und $x_0 \in \mathbb{K}$. Die Reihe

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + a_3(x-x_0)^3 + a_4(x-x_0)^4 + \dots$$

heißt *Potenzreihe in der Variaben x mit Entwicklungspunkt x_0 und Koeffizienten(folge) a_n*.

Die Fragestellung, die uns bei Potenzreihen zunächst interessiert, ist: Für welche Werte $x \in \mathbb{K}$ konvergiert die Potenzreihe

$$P(x) := \sum_{n=0}^{\infty} a_n(x-x_0)^n.$$

Für diese x macht es nämlich Sinn, die Funktion $x \mapsto P(x)$ zu betrachten und weiter zu untersuchen. Wir werden später sehen, dass wir durch Potenzreihen viele interessante Funktionen definieren können.

Wir schauen uns zunächst aber ein paar Beispiele zum Konvergenzverhalten an.

Beispiele:

1) Es sei die Potenzreihe $\sum_{n=0}^{\infty} x^n$ mit Entwicklungspunkt $x_0 = 0$ gegeben. Dies ist eine geometrische Reihe. Wir wissen, dass diese Reihe für $|x| < 1$ konvergiert und für $|x| \geq 1$ divergiert.

2) Nun betrachten wir die Reihe

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

Wir wenden das Quotientenkriterium an und erhalten

$$\left|\frac{(n+1)x^{n+1}}{nx^n}\right| = \frac{n+1}{n}|x| \xrightarrow[n\to\infty]{} |x|.$$

Hieraus folgt, dass die Potenzreihe für $|x| < 1$ konvergiert. Ist $|x| \geq 1$, so ist die Folge (nx^n) keine Nullfolge und folglich divergiert die Potenzreihe.

3) Die Potenzreihe

$$\sum_{n=0}^{\infty} n(x-1)^n$$

konvergiert ebenso für $|x-1| < 1$ und divergiert für $|x-1| \geq 1$.

Wir sehen in diesen Beispielen, dass der Konvergenzbereich der betrachteten Potenzreihen immer kreisförmig ist – mal mit Rand, mal ohne Rand. Das nächste Lemma zeigt, dass das kein Zufall ist. Der Konvergenzbereich einer Potenzreihe ist stets kreisförmig.

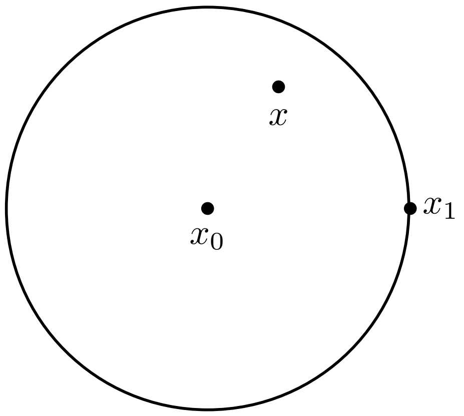


Abbildung 2.16.: Die Situation in Lemma 104

Lemma 104: Konvergiert die Potenzreihe

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n$$

in einem Punkt $x_1 \neq x_0$, so konvergiert sie auch in jedem Punkt $x \in \mathbb{K}$ mit $|x-x_0| < |x_1-x_0|$ absolut.

BEWEIS: Da die Reihe $\sum_{n=0}^{\infty} a_n(x_1-x_0)^n$ konvergiert, ist $(a_n(x_1-x_0)^n)_{n\in\mathbb{N}_0}$ eine Nullfolge. Es existiert also ein $M \in \mathbb{R}_+$ mit

$$|a_n(x_1-x_0)^n| \leq M$$

für alle $n \in \mathbb{N}_0$. Also haben wir für $x \in \mathbb{K}$ mit $|x - x_0| < |x_1 - x_0|$

$$\begin{aligned}|a_n(x - x_0)^n| &= \left|a_n(x_1 - x_0)^n \frac{(x - x_0)^n}{(x_1 - x_0)^n}\right| \\ &= |a_n(x_1 - x_0)^n| \cdot \left|\frac{(x - x_0)^n}{(x_1 - x_0)^n}\right| \\ &\leq M q^n\end{aligned}$$

mit $q := \left|\frac{(x-x_0)}{(x_1-x_0)}\right| < 1$. Nach dem Majorantenkriterium konvergiert $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ für die entsprechenden x absolut. □

Aus dem vorstehenden Lemma ergibt sich der folgende Satz, der das Konvergenzverhalten von Potenzreihen weitgehend beschreibt.

Satz 105: Es sei $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ wieder eine Potenzreihe. Dann tritt genau einer der drei folgenden Fälle ein:

(i) Die Potenzreihe konvergiert für alle $x \in \mathbb{K}$ absolut.

(ii) Die Potenzreihe divergiert für alle $x \in \mathbb{K} \backslash \{x_0\}$.

(iii) Es gibt genau ein $R \in \mathbb{R}_+$, so dass die Potenzreihe für alle $x \in \mathbb{K}$ mit $|x - x_0| < R$ absolut konvergiert und für $|x - x_0| > R$ divergiert.

Definition 106 (Konvergenzradius)**:** Konvergiert die Potenzreihe $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ für alle x mit $|x - x_0| < R$ und divergiert für $|x - x_0| > R$, so nennt man R den *Konvergenzradius* von $\sum_{n=0}^{\infty} a_n(x - x_0)^n$. Konvergiert die Potenzreihe für alle $x \in \mathbb{K}$, so sagt man, dass der Konvergenzradius Unendlich ist $(R = \infty)$. Divergiert die Potenzreihe für alle $x \in \mathbb{K} \backslash \{x_0\}$, so sagt man, dass der Konvergenzradius Null ist.

BEMERKUNG: (i) Den Konvergenzradius kann man meist mittels des Quotienten- oder des Wurzelkriteriums bestimmen.

(ii) Auf Grund von Lemma 104 und Satz 105 ist der Konvergenzradius wohldefiniert.

Hat eine Potenzreihe einen Konvergenzradius $R \in \mathbb{R}_+$ (also ungleich 0 oder ∞), so kann man über das Verhalten der Reihe auf dem Konvergenzrand $|x - x_0| = R$ keine allgemeingültige Aussage treffen, wie die folgenden Beispiele zeigen:

Beispiele:

1) Die geometrische Reihe $\sum_{n=1}^{\infty} x^n$ divergiert für alle $x \in \mathbb{C}$ mit $|x| = 1$, da dann die Folge (x^n) keine Nullfolge ist.

2) Die Reihe $\sum_{n=1}^{\infty} \frac{x^n}{n}$ konvergiert in $x = -1$ auf Grund des Leibniz-Kriteriums. Für $x = 1$ entspricht sie der harmonischen Reihe und divergiert demnach.

3) Die Potenzreihe $\sum_{n=1}^{\infty} \frac{x^n}{n^2}$ konvergiert für alle $|x| = 1$. Für diese x gilt nämlich:

$$\left| \sum_{n=1}^{\infty} \frac{x^n}{n^2} \right| \leq \sum_{n=1}^{\infty} \left| \frac{x^n}{n^2} \right| = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Somit ist diese Reihe in den Punkten $|x| = 1$ absolut konvergent.

Problem 45: *Bestimmen Sie die Konvergenzradien der folgenden Potenzreihen.*

(i)

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} x^n,$$

(ii)

$$\sum_{n=1}^{\infty} \frac{(x-1)^n}{n},$$

(iii)

$$\sum_{n=1}^{\infty} \frac{n^3}{7^n} x^{2n}.$$

2.14. Die Exponentialfunktion und die trigonometrischen Funktionen

Satz 107: Die Potenzreihe

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

konvergiert für alle $x \in \mathbb{C}$ absolut.

BEWEIS: Es sei $x \in \mathbb{C}$ so gilt

$$\left| \frac{x^{n+1}}{(n+1)!} \cdot \frac{n!}{x^n} \right| = \frac{|x|}{n+1} \xrightarrow[n\to\infty]{} 0.$$

Das Quotientenkriterium ist also für alle x erfüllt. Der Konvergenzradius ist damit unendlich. □

Definition 108 (Exponentialfunktion)**:** Die durch

$$\begin{aligned} \exp : \mathbb{K} &\to \mathbb{K} \\ x &\mapsto \exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!} \end{aligned}$$

definierte Funktion heißt (komplexe bzw. reelle) *Exponentialfunktion*.

Konvention: Bei Funktionen (bzw. Potenzreihen) im Komplexen schreibt man meist z statt x (und z_0 statt x_0).

Satz 109 (Funktionalgleichung der Exponentialfunktion)**:** Für alle $x, y \in \mathbb{K}$ gilt die Funktionalgleichung

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

BEWEIS: Der Beweis verwendet das Cauchy-Produkt aus Definition 101 beziehungsweise aus Satz 102.

$$\begin{aligned} \exp(x) \cdot \exp(y) &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \cdot \sum_{m=0}^{\infty} \frac{y^m}{m!} \\ &= \sum_{k=0}^{\infty} \sum_{j=0}^{k} \left(\frac{x^j}{j!} \cdot \frac{y^{k-j}}{(k-j)!} \right) \\ &= \sum_{k=0}^{\infty} \left(\frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} x^j y^{k-j} \right) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} (x+y)^k \\ &= \exp(x+y) \end{aligned}$$

□

Definition 110 (Euler'sche Zahl)**:** Die Zahl

$$e := \exp(1) = \sum_{k=0}^{\infty} \frac{1}{k!}$$

heißt *Euler'sche Zahl*.

Satz 111: Es gilt:

(i) $\exp(0) = 1$,

(ii) $z \in \mathbb{C} \Rightarrow \exp(z) \neq 0$,

(iii) $z \in \mathbb{C} \Rightarrow \exp(-z) = \frac{1}{\exp(z)}$,

(iv) $x \in \mathbb{R} \Rightarrow \exp(x) > 0$; ist $x > 0$, so gilt sogar $\exp(x) > 1$,

(v) $r \in \mathbb{Q} \Rightarrow \exp(r) = e^r$.

BEWEIS:

(i) Ist klar!

(ii) Angenommen es existiert ein $z \in \mathbb{C}$ mit $\exp(z) = 0$, dann gilt

$$1 = \exp(0) = \exp(z - z) = \exp(z)\exp(-z) = 0\exp(-z) = 0.$$

Widerspruch!

(iii) Folgt aus dem Beweis von Teil (ii), da $1 = \exp(z)\exp(-z)$ für alle $z \in \mathbb{C}$.

(iv) Für $x \geq 0$ gilt $\exp(x) = 1 + \sum_{k=1}^{\infty} \frac{x^k}{k!} \geq 1$. Daraus folgt $\exp(-x) = \frac{1}{\exp(x)} > 0$.

(v) Für $r = \frac{p}{q}$ mit $p, q \in \mathbb{N}$ gilt

$$(\exp(r))^q = \exp(r \cdot q) = \exp(p) = \exp(1)^p = e^p.$$

Also ist $\exp(r) = \sqrt[q]{e^p} = e^r$. Ist $r < 0$, so haben wir aufgrund von (iii) und der bekannten Rechengesetze für Potenzen

$$\exp(r) = \frac{1}{\exp(-r)} = \frac{1}{e^{-r}} = e^r.$$

$\square$

Satz 112 (Restgliedabschätzung der Exponentialfunktion)**:** Es sei $n \in \mathbb{N}_0$. Dann gilt für alle reellen Zahlen x mit $|x| \leq 1 + \frac{n}{2}$

$$\left| \exp(x) - \sum_{k=0}^{n} \frac{x^k}{k!} \right| \leq 2\frac{|x|^{n+1}}{(n+1)!}.$$

BEMERKUNG: Wir werden diesen Satz im Fall $n = 0$ später benutzen, um die Stetigkeit der Exponentialfunktion nachzuweisen. Die Bedeutung solcher Restgliedabschätzungen in der Numerik wird in einem Exkurs thematisiert.

BEWEIS:Es ist

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

Nun ist $(n+1+k)! = (n+1)!(n+2) \cdot \ldots \cdot (n+1+k) \geq (n+1)!(n+2)^k$ für $k \geq 0$, wenn man leere Produkte zu 1 festsetzt. Wir haben also

$$\left|\exp(x) - \sum_{k=0}^{n} \frac{x^k}{k!}\right| \leq \frac{|x|^{n+1}}{(n+1)!} \sum_{k=0}^{\infty} |x|^k \frac{(n+1)!}{(n+1+k)!} \leq \frac{|x|^{n+1}}{(n+1)!} \sum_{k=0}^{\infty} \left(\frac{|x|}{(n+2)}\right)^k.$$

Ist $\frac{|x|}{(n+2)} \leq \frac{1}{2}$, so konvergiert die geometrische Reihe ganz rechts und ihr Wert lässt sich nach oben durch 2 abschätzen, womit der Satz bewiesen ist. □

Eine hübsche Anwendung der Restgliedabschätzung der Exponentialfunktion ist das folgende Korollar.

Korollar 113: Die Eulersche Zahl e ist irrational.

BEWEIS: Wir nehmen an, dass es ganze Zahlen p, q mit $q > 0$ gibt, so dass $e = p/q$ ist. Dann ist für alle $n \in \mathbb{N}$

$$n!q\left(e - \sum_{k=0}^{n} \frac{1}{k!}\right) \in \mathbb{Z} \setminus \{0\}.$$

Also haben wir, wenn wir die Restgliedabschätzung anwenden,

$$0 < \left|n!q\left(e - \sum_{k=0}^{n} \frac{1}{k!}\right)\right| \leq n!q\frac{2}{(n+1)!} = \frac{2q}{n+1}.$$

Wählen wir nun $n \geq 2q$, so erhalten wir einen Widerspruch, da jede ganze Zahl ungleich Null betraglich größer oder gleich 1 ist.

□

Exkurs: Ein analytisches Irrationalitätskriterium

Man kann den vorstehenden Beweis verallgemeinern und auf ein allgemeines Irrationalitätskriterium zurückführen.

Satz 114 (Irrationalitätskriterium)**:** Es seien (p_n) und (q_n) Folgen ganzer Zahlen und $\alpha \in \mathbb{R}$. Gelten dann die beiden Bedingungen

- $\alpha q_n - p_n \neq 0$ für alle hinreichend großen n,
- $\lim_{n\to\infty} |\alpha q_n - p_n| = 0$,

so ist $\alpha \notin \mathbb{Q}$.

BEWEIS: Wir nehmen an, dass die beiden Bedingungen im Satz erfüllt sind und dass dennoch α rational ist. Dann existieren also ganze Zahlen p, q, so dass $\alpha = \frac{p}{q}$ ist. Ist n genügend groß, so gilt aufgrund der beiden Bedingungen

$$0 < |\alpha q_n - p_n| < \frac{1}{q}.$$

Hieraus folgt aber für die ganze Zahl $q\alpha q_n - qp_n$

$$0 < |q\alpha q_n - qp_n| < 1,$$

was nicht möglich ist. □

Beispiele:

1) Irrationalität von e: Wir setzen $p_n := n! \sum_{k=0}^{n} \frac{1}{k!}$ und $q_n := n!$. Dass damit die Bedingungen des Irrationaltätskriteriums erfüllt sind, zeigen die Abschätzung aus dem Beweis des Korollars 113.

2) Irrationalität von $\sqrt{2}$: Wir definieren die Folgen (p_n) und (q_n) rekursiv, indem wir setzen $p_1 = q_1 = 1$ und $q_{n+1} := 2p_n q_n$ und $p_{n+1} := 2q_n^2 + p_n^2$. Hiermit gilt, dass für alle $n \in \mathbb{N}$

$$(\sqrt{2}q_{n+1} - p_{n+1}) = -(\sqrt{2}q_n - p_n)^2$$

ist, woraus sofort folgt, dass die erste Bedingung im obigen Irrationalitätskriterium erfüllt ist. Induktiv folgert man weiter, dass

$$|\sqrt{2}q_n - p_n| < \frac{1}{2^n}$$

ist, woraus dann die zweite Bedingung folgt.

Ende des Exkurses

Exkurs: Numerische Approximation der Exponentialfunktion

Der vorstehende Satz gibt uns eine Möglichkeit, wie man Werte der Exponentialfunktion näherungsweise mit Hilfe des Computers berechnen kann und er gibt uns darüber hinaus eine Abschätzung für den Fehler, den wir dabei möglicherweise begehen. Das folgende Listing zeigt eine einfache Implementierung in Matlab.

```
function erg=exp_reihe(a,N)
if nargin==1
    N=50;
elseif nargin~=2
    error('Es werden 1 oder 2 Argumente erwartet')
end
for i=1:length(a)
  erg(i)=sum(1./factorial(0:N).*a(i).^(0:N));
end
```

Das folgende Skript zeichnet die Graphen der Exponentialfunktion und einiger Näherungspolynome in einer Umgebung des Nullpunktes.

```
clear all
x=-3:0.2:3;
figure(1)
clf()
plot(x,exp(x),x,exp_reihe(x,1),x,exp_reihe(x,2),x,exp_reihe(x,3),x,exp_reihe(x
    ,5))
legend('exp','n=1','n=2','n=3','n=5')
```

Ende des Exkurses

Satz 115: Für $x \in \mathbb{R}$ gilt

$$|\exp(ix)| = 1.$$

BEWEIS: Es sei (z_n) eine gegen z_0 konvergente Folge in $\mathbb{C}$, so erhalten wir mit Lemma 66, dass $\overline{z_n}$ gegen $\overline{z_0}$ konvergiert. Für $z \in \mathbb{C}$ gilt somit

$$\overline{\sum_{n=0}^{m} \frac{z^n}{n!}} = \sum_{n=0}^{m} \frac{\overline{z^n}}{n!} = \sum_{n=0}^{m} \frac{\overline{z}^n}{n!}.$$

Daraus folgt

$$\overline{\exp(z)} = \exp(\overline{z}),$$

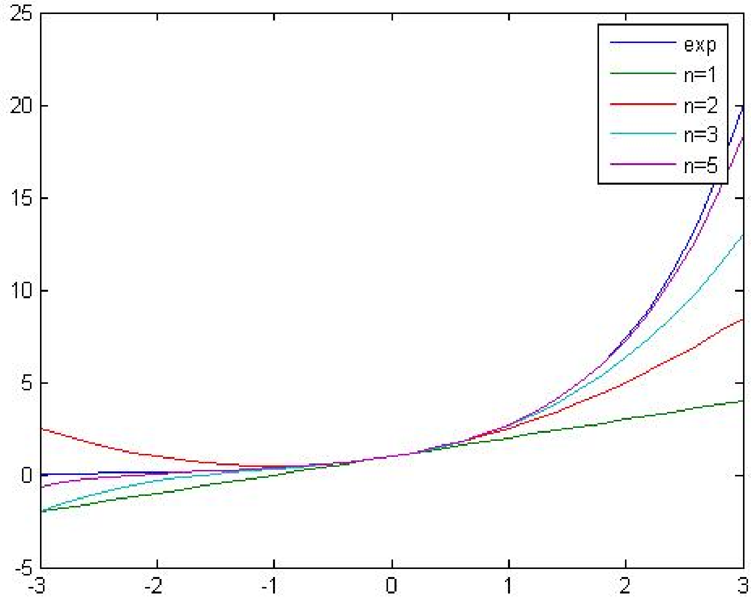


Abbildung 2.17.: Approximation der Exponentialfunktion durch Näherungspolynome

und hieraus erhalten wir für reelles x

$$|\exp(ix)|^2 = \exp(ix)\overline{\exp(ix)} = \exp(ix)\exp(-ix) = \exp(ix - ix) = \exp(0) = 1.$$

$\square$

Definition 116 (Sinus und Kosinus)**:** Für $x \in \mathbb{R}$ definieren wir die Funktionen $\sin, \cos : \mathbb{R} \to \mathbb{R}$ durch

$$\begin{aligned} \sin(x) &= \mathrm{Im}(\exp(ix)) \\ \cos(x) &= \mathrm{Re}(\exp(ix)). \end{aligned}$$

Für $x \in \mathbb{R}$ gilt also die *Euler'sche Formel*

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

Wir können jetzt auch den Real- und Imaginärteil der komplexen Exponentialfunktion mittels reeller Funktionen ausdrücken, denn es ist

$$\exp(x + iy) = e^x(\cos(y) + i\sin(y)) = e^x\cos(x) + ie^x\sin(y).$$

Bisher sind $\cos$ und $\sin$ für uns allerdings kaum mehr als Namen. Wir werden diese beiden Funktionen später noch eingehender studieren.

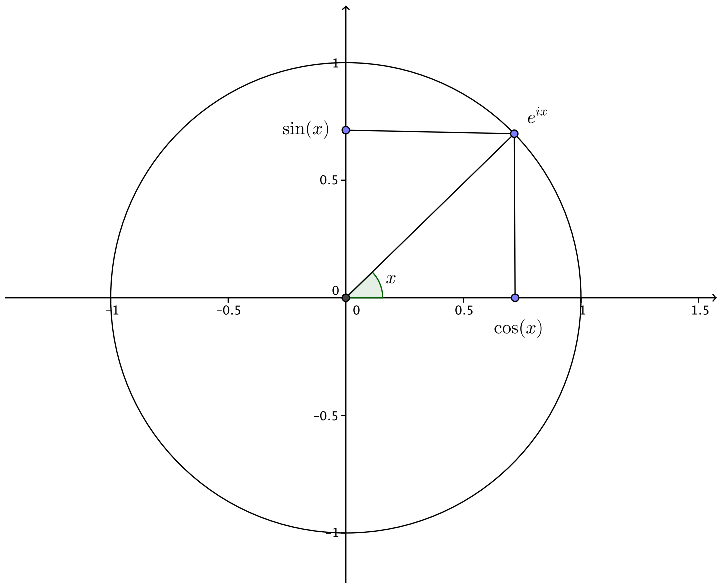


Abbildung 2.18.: Die Definition von Sinus und Cosinus für reelles x

```
clear all
figure(1)
clf()
[x,y]=meshgrid(-4:0.1:10,-4:0.1:10);
f=exp(x).*cos(y);
surf(x,y,f)
grid on
xlabel('x')
ylabel('y')
zlabel('Re(exp(x+iy))')
```

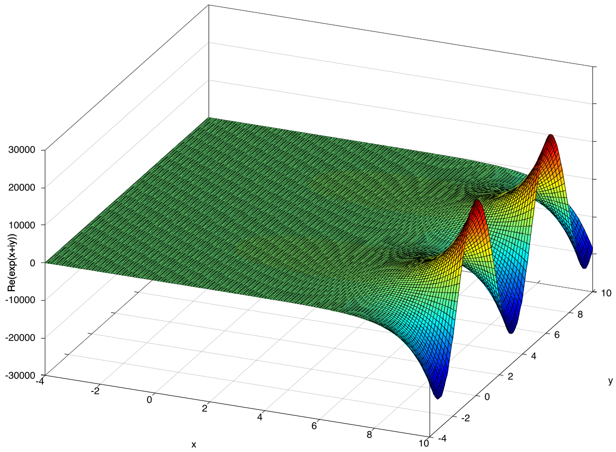


Abbildung 2.19.: Realteil der komplexen Exponentialfunktion

Lemma 117: Die beiden Potenzreihen

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

und

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}$$

konvergieren absolut für alle $z \in \mathbb{C}$.

BEWEIS: Wir beweisen die Aussage nur für die erste Reihe. Der Beweis für die andere Reihe geht analog. Zum Beweis wenden wir das Quotientenkriterium auf die Reihe $\sum a_n$ mit $a_n := \frac{(-1)^n}{(2n+1)!} z^{2n+1}$ an. Es ist

$$\left| \frac{(-1)^{n+1} z^{2(n+1)+1}}{(2(n+1)+1)!} \cdot \frac{(2n+1)!}{(-1)^n z^{2n+1}} \right| = \frac{|z|^2}{(2n+3)(2n+2)} \xrightarrow[n\to\infty]{} 0,$$

und somit konvergiert die Reihe für alle $z \in \mathbb{C}$. $\square$

Lemma 118: Für alle $x \in \mathbb{R}$ gilt

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \text{ und}$$
$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$$

BEWEIS: Ist $n \in \mathbb{N}$ gerade, so können wir $n = 2m$ schreiben, anderenfalls $n = 2m+1$. Da wir wissen, dass die auftretenden Reihen konvergent sind, haben wir also

$$\begin{aligned}\exp(ix) = \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} &= \sum_{m=0}^{\infty} i^{2m} \frac{x^{2m}}{(2m)!} + \sum_{m=0}^{\infty} i^{2m+1} \frac{x^{2m+1}}{(2m+1)!} \\ &= \sum_{m=0}^{\infty} i^{2m} \frac{x^{2m}}{(2m)!} + i \sum_{m=0}^{\infty} i^{2m} \frac{x^{2m+1}}{(2m+1)!}.\end{aligned}$$

Nun ist $i^{2m} = (i^2)^m = (-1)^m$ und somit

$$\exp(ix) = \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m}}{(2m)!} + i \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m+1}}{(2m+1)!}.$$

Also haben wir

$$\begin{aligned}\mathrm{Re}(\exp(ix)) &= \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m}}{(2m)!} \\ \mathrm{Im}(\exp(ix)) &= \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m+1}}{(2m+1)!}.\end{aligned}$$

□

Da die beiden Potenzreihen im vorstehenden Lemma für alle komplexen Argumente konvergieren, können wir diese Potenzreihen nutzen, um die bisher nur auf den reellen Zahlen definierten Funktionen Sinus und Cosinus auf die ganze komplexe Ebene *fortzusetzen*.

Definition 119 (Sinus und Kosinus)**:** Für alle $z \in \mathbb{C}$ seien die Funktionen $\sin, \cos : \mathbb{C} \to \mathbb{C}$ durch

$$\begin{aligned}\sin(z) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} \quad \text{und} \\ \cos(z) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}\end{aligned}$$

definiert.

BEMERKUNG: Lemma 118 zeigt, dass diese Definition mit der ursprünglichen Definition von $\sin$ und $\cos$ für reelle x verträglich ist.

Satz 120: Für alle $z \in \mathbb{C}$ gilt

$$\sin(z) = \frac{1}{2i}\left(\exp(iz) - \exp(-iz)\right)$$
$$\cos(z) = \frac{1}{2}\left(\exp(iz) + \exp(-iz)\right).$$

Problem 46: *Beweisen Sie den vorstehenden Satz.*

Satz 121: Für alle $z, w \in \mathbb{C}$ gilt

(i) $\cos(z+w) = \cos(z)\cos(w) - \sin(z)\sin(w)$
$\sin(z+w) = \sin(z)\cos(w) + \cos(z)\sin(w)$

(ii) $\sin^2(z) + \cos^2(z) = 1$

(iii) $\cos(z) = \cos(-z)$, $\sin(z) = -\sin(-z)$

(iv) $\cos(0) = 1$, $\sin(0) = 0$

Problem 47: *Beweisen Sie den vorstehenden Satz.*

```
clear all
figure(1)
clf()
[x,y]=meshgrid(-8:0.1:8,-8:0.1:8);
f=sin(x).*cosh(y);
surf(x,y,f)
grid on
title('Realteil der komplexen Sinus-Funktion')
xlabel('x')
ylabel('y')
zlabel('Re(sin(x+iy))')
```

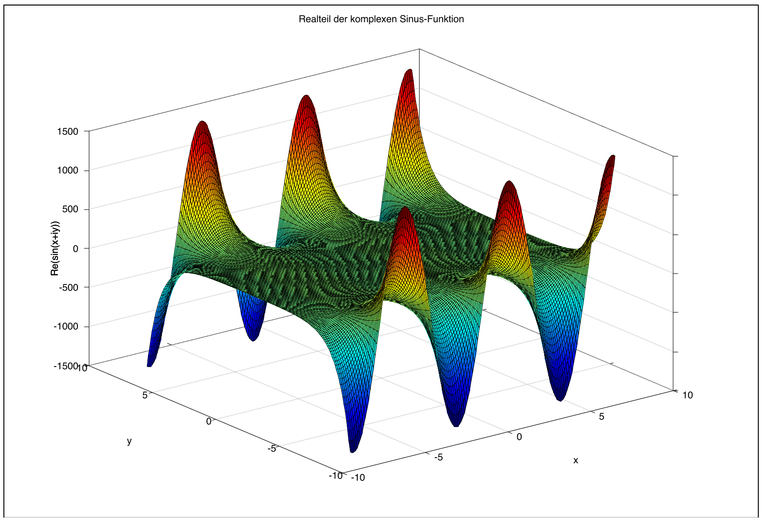


Abbildung 2.20.: Realteil der komplexen Sinus-Funktion

3. Stetigkeit

3.1. Stetige reellwertige Funktionen

Wir haben zuvor gesehen, dass wenn (a_n) gegen a konvergiert, auch

- $|a_n|$ gegen $|a|$ konvergiert und
- $\sqrt{a_n}$ gegen $\sqrt{a}$ konvergiert.

Man kann dies auch als eine Eigenschaft der beiden Funktionen $|\cdot|$ und $\sqrt{\cdot}$ ansehen. Diese Eigenschaft heißt *Stetigkeit*.

Stetigkeit ist eine sehr wichtige Eigenschaft reellwertiger Funktionen. Wir werden sie gleich für Funktionen auf $\mathbb{R}^r$ definieren. In den folgenden Kapitel werden wir aber fast ausschließlich Funktionen betrachten, die von einer reellen Variablen abhängen.

Definition 122 (Stetigkeit reellwertiger Funktionen)**:** Es seien $r \in \mathbb{N}$, $D \subset \mathbb{R}^r$, $a \in D$ und $f : D \to \mathbb{R}$. Wir sagen f *ist stetig in* a, wenn für jede Folge (a_n) in D mit $\lim_{n\to\infty} a_n = a$

$$\begin{aligned} \lim_{n\to\infty} f(a_n) &= f(a), \text{ also} \\ \lim_{n\to\infty} f(a_n) &= f\left(\lim_{n\to\infty} a_n\right) \end{aligned}$$

gilt. Die Funktion $f : D \to \mathbb{R}$ heißt *stetig (auf D)*, wenn f in jedem $a \in D$ stetig ist.

Beispiele:

1) Konstante Funktionen, also Funktionen der Art $\text{konst}_c : \mathbb{R} \to \mathbb{R}, x \mapsto c$ für ein festes $c \in \mathbb{R}$, sind stetig.

2) Die identische Abbildung $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}, x \mapsto x$ ist stetig.

3) Sind $a_0, \ldots, a_n \in \mathbb{R}$, so ist die Polynomfunktion $P : \mathbb{R} \to \mathbb{R}, x \mapsto \sum_{k=0}^{n} a_k x^k$ stetig.

4) Die Funktion $|\cdot| : \mathbb{R} \to \mathbb{R}, \quad x \mapsto |x|$ ist aufgrund von Satz 47 stetig.

5) Die Funktion $f : \mathbb{R} \to \mathbb{R}$, die definiert ist durch

$$f(x) := \begin{cases} 1, & \text{falls } x \geq 0, \\ -1, & \text{falls } x < 0, \end{cases}$$

ist nicht stetig in 0, denn es ist

$$\lim_{n\to\infty} f\left(-\frac{1}{n}\right) = -1 \neq 1 = f(0).$$

6) Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}\ \ (x, y) \mapsto x + y$ ist aufgrund von Satz 47 stetig.

7) Die Funktionen $Proj_k : \mathbb{R}^r \to \mathbb{R},\ (x_1, \ldots, x_k, \ldots, x_r) \mapsto x_k$ für $k = 1, \ldots, r$ sind stetig. $Proj_k$ heißt *Projektion auf die k-te Komponente*.

8) Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ ist stetig. Nach der Restgliedabschätzung in Satz 112 ist $|\exp(x) - 1| \leq 2|x|$ für alle x mit $|x| \leq 1$. Es sei a eine beliebige reelle Zahl und (a_n) eine konvergente Folge mit Grenzwert a. Dann ist $(a_n - a)$ eine Nullfolge. Aus der gerade zitierten Abschätzung folgt also

$$\lim_{n\to\infty} \exp(a_n - a) = 1.$$

Aus der Funktionalgleichung der Exponentialfunktion folgt

$$\exp(a_n) = \exp(a_n - a + a) = \exp(a_n - a)\exp(a)$$

und somit

$$\lim_{n\to\infty} \exp(a_n) = \exp(a).$$

9) Wir werden später sehen, dass durch Potenzreihen gegebenen Funktionen im Inneren des Konvergenzkreises stetig sind; insbesondere also auch die trigonometrischen Funktionen $\sin$ und $\cos$. Wir sparen uns deshalb an dieser Stelle explizite Nachweise der Stetigkeit dieser Funktionen.

Problem 48: *Bestimmen Sie für die folgenden Funktionen $f : D \to \mathbb{R}, x \mapsto f(x)$ alle $x \in D$, in denen f stetig ist (mit Beweis!).*

(i) $D := \mathbb{R}$ *und*

$$f(x) := \begin{cases} 1 & \text{, falls } x \leq 0, \\ 0 & \text{, falls } x > 0. \end{cases}$$

(ii) $D := \mathbb{R}$ *und*

$$f(x) := \begin{cases} 1 & \text{, falls } x \in \mathbb{Q}, \\ 0 & \text{, falls } x \notin \mathbb{Q}. \end{cases}$$

(iii) $D := \mathbb{R}$ *und*

$$f(x) := \begin{cases} 1 & \text{, falls } x \leq \sqrt{2}, \\ 0 & \text{, falls } x > \sqrt{2}. \end{cases}$$

(iv) $f(x)$ *wie in (iii), aber* $D := \mathbb{Q}$.

Problem 49: *Ist die durch*

$$f(x,y) := \begin{cases} \frac{xy}{x^2+y^2} & \text{, für } (x,y) \neq (0,0), \\ 0 & \text{, für } (x,y) = (0,0) \end{cases}$$

definierte Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ *in* $(0,0)$ *stetig?*

Satz 123: Es seien $r \in \mathbb{N}$ und $D \subset \mathbb{R}^r$. Weiter seien $f, g : D \to \mathbb{R}$ in $a \in D$ stetige Funktionen und $\lambda \in \mathbb{R}$. Dann sind auch die Funktionen

$$\begin{aligned} f+g &: D \to \mathbb{R} & x &\mapsto f(x)+g(x) \\ \lambda f &: D \to \mathbb{R} & x &\mapsto \lambda \cdot f(x) \\ fg &: D \to \mathbb{R} & x &\mapsto f(x) \cdot g(x) \end{aligned}$$

in a stetig. Gilt außerdem $g(a) \neq 0$, so ist auch

$$fg^{-1} : D' \to \mathbb{R} \qquad x \mapsto \frac{f(x)}{g(x)}$$

in a stetig, wobei D' durch $D' := \{x \in D | g(x) \neq 0\}$ definiert ist.

BEWEIS: Der Satz folgt direkt aus den Grenzwertrechenregeln für Folgen, siehe Satz 47. □

Satz 124: Es seien $r \in \mathbb{N}$, $D \subset \mathbb{R}^r$ und $E \subset \mathbb{R}$. Sind nun $f : D \to \mathbb{R}$ und $g : E \to \mathbb{R}$ Funktionen mit $f(D) \subset E$ und ist f in $a \in D$ und g in $b := f(a) \in E$ stetig, dann ist die Funktion

$$g \circ f : D \to \mathbb{R} \qquad x \mapsto g(f(x))$$

in a stetig.

BEWEIS: Es sei (x_n) eine Folge in D mit $\lim_{n\to\infty} x_n = a$. Da f in a stetig ist, ist $f(x_n)$ eine Folge in E mit $\lim_{n\to\infty} f(x_n) = f(a) = b$. Da g in b stetig ist, gilt also

$$\lim_{n\to\infty} (g \circ f)(x_n) = \lim_{n\to\infty} g(f(x_n)) = g(\lim_{n\to\infty} f(x_n)) = g(f(a)) = (g \circ f)(a).$$

□

Beispiel: Es sei $f : D \to \mathbb{R}$ stetig, dann ist auch

$$\begin{aligned} |f| : D &\to \mathbb{R} \\ x &\mapsto |f(x)| \end{aligned}$$

stetig, denn es gilt $|f| = |\cdot| \circ f$.

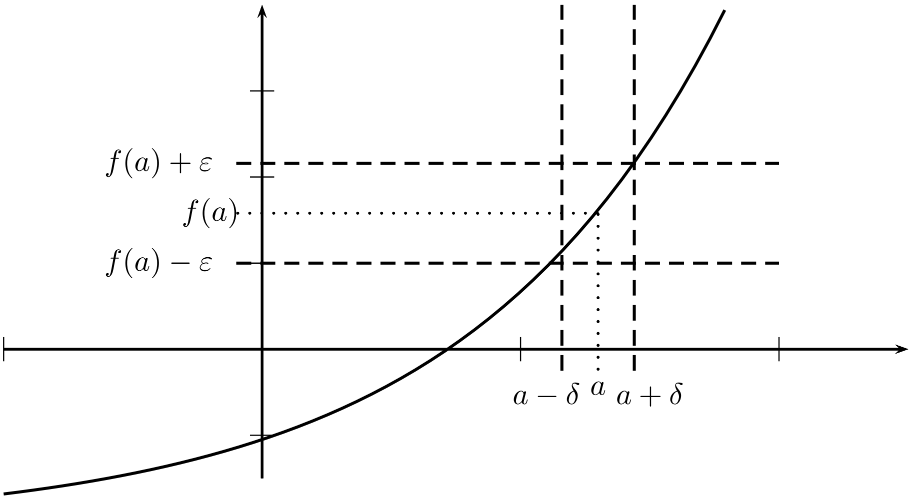


Abbildung 3.1.: ε-δ Definition der Stetigkeit

Unsere bisherige Definition der Stetigkeit bezeichnet man auch als Folgenstetigkeit. Wir wollen jetzt zumindest im Eindimensionalen noch eine äquivalente Charakterisierung kennenlernen.

Satz 125 (ε-δ-Stetigkeit)**:** Es sei $D \subset \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion, so sind die folgenden zwei Aussagen äquivalent.

(i) Die Funktion f ist stetig in $a \in D$.

(ii) Für alle $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für alle $x \in D$ mit $|x - a| < \delta$

$$|f(x) - f(a)| < \varepsilon$$

folgt. In Quantorenschreibweise heißt das

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in D : |x - a| < \delta \;\Rightarrow |f(x) - f(a)| < \varepsilon.$$

Beweis: $(i) \Rightarrow (ii)$: Angenommen (ii) gilt nicht, das heißt

$$\exists \varepsilon > 0 \; \forall \delta > 0 \; \exists x \in D : |x - a| < \delta \text{ und } |f(x) - f(a)| \geq \varepsilon.$$

Für jedes $n \in \mathbb{N}$ existiert zu $\delta = \frac{1}{n}$ ein $x_n \in D$ mit $|x_n - a| < \delta = \frac{1}{n}$ und $|f(x_n) - f(a)| \geq \varepsilon$. Es gilt jedoch $\lim_{n\to\infty} x_n = a$. Da f stetig in a ist, gilt $\lim_{n\to\infty} f(x_n) = f(a)$, was $|f(x_n) - f(a)| \geq \varepsilon$ widerspricht.

$(ii) \Rightarrow (i)$: Es sei $\varepsilon > 0$ beliebig, so gibt es nach (ii) ein $\delta > 0$, so dass für alle $x \in D$ mit $|x - a| < \delta$ auch

$$|f(x) - f(a)| < \varepsilon$$

gilt. Es sei nun (x_n) eine konvergente Folge in D mit Grenzwert a. Dann existiert zu jedem $\delta > 0$ ein $n_0 \in \mathbb{N}$, so dass $|x_n - a| < \delta$ für alle $n \geq n_0$ gilt. Daraus folgt

$$|f(x_n) - f(a)| < \varepsilon$$

für alle $n \geq n_0$, das heißt

$$\lim_{n\to\infty} f(x_n) = f(a).$$

□

Zum Abschluss dieses Kapitels wollen wir noch ein paar nützliche neue Schreibweisen für Grenzwerte bei reellwertigen Funktionen einführen.

Definition 126 (Berührpunkt)**:** Es seien $r \in \mathbb{N}$ und $D \subset \mathbb{R}^r$. Ein Punkt $a \in \mathbb{R}^r$ heißt *Berührpunkt von* D, wenn es mindestens eine Folge (a_n) in D gibt, derart, dass $\lim_{n\to\infty} a_n = a$ gilt.

Beispiel: Ist D ein Intervall der Gestalt $]a, b[$, so ist jedes $x \in [a, b]$ ein Berührpunkt von D.

Definition 127 (Grenzwerte bei Funktionen)**:** Es seien $D \subset \mathbb{R}^r$, $f : D \to \mathbb{R}$ eine Funktion, $a \in \mathbb{R}^r$ ein Berührpunkt von D und $b \in \mathbb{R} \cup \{\pm\infty\}$.

a) Gilt für jede Folge (a_n) in $D \setminus \{a\}$ mit $\lim_{n\to\infty} a_n = a$

$$\lim_{n\to\infty} f(a_n) = b,$$

so schreiben wir

$$\lim_{x\to a} f(x) = b.$$

b) Speziell im Fall $r = 1$ definieren wir weiter: Gilt für jede Folge (a_n) in D mit $a_n > a$ für alle n und $\lim_{n\to\infty} a_n = a$

$$\lim_{n\to\infty} f(a_n) = b$$

gilt, so schreiben wir

$$\lim_{x \searrow a} f(x) = b.$$

Gilt für jede Folge (a_n) in D mit $a_n < a$ für alle n und $\lim_{n\to\infty} a_n = a$

$$\lim_{n\to\infty} f(a_n) = b$$

gilt, so schreiben wir

$$\lim_{x \nearrow a} f(x) = b.$$

c) Bei $r = 1$ definieren wir: Gilt für jede Folge (a_n) in D mit $\lim_{n\to\infty} a_n = +\infty$

$$\lim_{n\to\infty} f(a_n) = b$$

gilt, dann schreiben wir

$$\lim_{x\to\infty} f(x) = b.$$

Gilt für jede Folge (a_n) in D mit $\lim_{n\to\infty} a_n = -\infty$

$$\lim_{n\to\infty} f(a_n) = b$$

gilt, dann schreiben wir

$$\lim_{x\to-\infty} f(x) = b.$$

Problem 50: *Es sei I ein nicht-leeres Intervall. Beweisen oder widerlegen Sie die folgenden Aussagen:*

(i) Eine Funktion $f : I \to \mathbb{R}$ ist genau dann stetig in $x_0 \in I$, wenn

$$\lim_{x\to x_0} f(x)$$

existiert und gleich $f(x_0)$ ist.

(ii) Eine Funktion $f : I \to \mathbb{R}$ ist genau dann stetig in $x_0 \in I$, wenn die Grenzwerte $\lim_{x\searrow x_0} f(x)$ und $\lim_{x\nearrow x_0} f(x)$ existieren und gleich $f(x_0)$ sind.

Definition 128 (Stetige Fortsetzung)**:** Es seien $r \in \mathbb{N}$, $D \subset \mathbb{R}^r$, $a \notin D$ ein Berührpunkt von D und $f : D \to \mathbb{R}$ eine stetige Funktion. Wir sagen f *ist stetig fortsetzbar in* a, wenn es ein $b \in \mathbb{R}$ gibt, so dass

$$\lim_{x \to a} f(x) = b$$

ist. Die Funktion $\tilde{f} : D \cup \{a\} \to \mathbb{R}$, die definiert ist durch

$$\begin{aligned} \tilde{f}(x) &:= f(x) \text{ für alle } x \in D \\ \tilde{f}(a) &:= b, \end{aligned}$$

heißt *(stetige) Fortsetzung von f in a.*

Beispiel: Das folgende Lemma zeigt, dass die Funktion $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ mit

$$f(x) = \frac{\exp(x) - 1}{x}$$

durch die Festsetzung $f(0) := 1$ in 0 stetig fortsetzbar ist.

Lemma 129: Es gilt

$$\lim_{x \to 0} \frac{\exp(x) - 1}{x} = 1.$$

BEWEIS: Die Exponentialfunktion ist gegeben durch die Potenzreihe

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

Es ist also

$$\frac{\exp(x) - 1}{x} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{k!}$$

und somit

$$\begin{aligned} \left| \frac{\exp(x) - 1}{x} - 1 \right| &\leq \sum_{k=2}^{\infty} \frac{|x|^{k-1}}{k!} \\ &= \sum_{k=1}^{\infty} \frac{|x|^k}{(k+1)!} = |x| \sum_{k=1}^{\infty} \frac{|x|^{k-1}}{(k+1)!} \\ &= |x| \sum_{k=0}^{\infty} \frac{|x|^k}{(k+2)!} \\ &\leq |x| \sum_{k=0}^{\infty} \frac{|x|^k}{k!} = |x| \exp(|x|). \end{aligned}$$

Da sowohl der Absolutbetrag als auch die Exponentialfunktion stetig sind, folgt hieraus die Behauptung. □

3.2. Sätze über stetige Funktionen einer Veränderlichen

Satz 130 (Zwischenwertsatz, 1. Version)**:** Es sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion und $f(a) \cdot f(b) < 0$. Dann existiert (mindestens) ein $p \in]a, b[$ mit $f(p) = 0$.

BEWEIS: Wir konstruieren ein geeignetes p mittels einer Intervallschachtelung. Die Voraussetzung $f(a) \cdot f(b) < 0$ besagt, dass die beiden Zahlen $f(a)$ und $f(b)$ nicht das selbe Vorzeichen haben können. Eine der beiden Zahlen ist also negativ, die andere positiv. Ohne Einschränkung können wir dazu annehmen, dass $f(a) < 0 < f(b)$.

Wir setzen $a_1 := a$ und $b_1 := b$. Angenommen $a_1, \ldots, a_{n-1}, b_1, \ldots, b_{n-1}$ sind bereits konstruiert, so setzen wir

$$a_n := \begin{cases} a_{n-1}, & \text{falls } f\left(\frac{a_{n-1}+b_{n-1}}{2}\right) \geq 0 \\ \frac{a_{n-1}+b_{n-1}}{2}, & \text{sonst} \end{cases}$$

$$b_n := \begin{cases} \frac{a_{n-1}+b_{n-1}}{2}, & \text{falls } f\left(\frac{a_{n-1}+b_{n-1}}{2}\right) \geq 0 \\ b_{n-1}, & \text{sonst.} \end{cases}$$

Nach Konstruktion gilt $f(a_n) < 0 \leq f(b_n)$ für alle $n \in \mathbb{N}$. Weiter gelten für alle $n \in \mathbb{N}$ die folgenden drei Aussagen

$$\begin{aligned} a_n &< b_n, \\ [a_n, b_n] &\subset [a_{n-1}, b_{n-1}] \quad \text{und} \\ b_n - a_n &= 2^{-(n-1)}(b_1 - a_1). \end{aligned}$$

Nach dem Satz 56 über die Intervallschachtelung existiert genau ein p mit $p \in [a_n, b_n]$ für alle $n \in \mathbb{N}$. Wir zeigen nun, dass unser so konstruiertes p die geforderte Eigenschaft besitzt. Nach Konstruktion gilt

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = p.$$

Und da f stetig ist, gilt somit

$$0 \geq \lim_{n\to\infty} f(a_n) = f(p) = \lim_{n\to\infty} f(b_n) \geq 0,$$

woraus $f(p) = 0$ folgt. Da nach Voraussetzung $f(a) \cdot f(b) \neq 0$ ist, liegt p tatsächlich in dem Intervall $]a, b[$. □

Problem 51: *Wie viele Nullstellen besitzt die Polynomfunktion*

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^5 - 3x^4 - \frac{11}{3}x^3 + 10x^2 + x - 3$$

im Intervall $[-2, 2]$*?*
Hinweis: Eine Polynomfunktion vom Grad n *hat maximal* n *Nullstellen. Dies folgt aus dem Hauptsatz der Algebra mittels Polynomdivision.*

Korollar 131 (Zwischenwertsatz, 2. Version)**:** Es sei $f : [a,b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) < f(b)$. Dann existiert zu jedem $c \in [f(a), f(b)]$ ein $p \in [a,b]$ mit $f(p) = c$.

BEWEIS: Wende die erste Version des Zwischenwertsatzes auf $g(x) := f(x) - c$ an. □

Exkurs: Nullstellenberechnung mittels des Zwischenwertsatzes

Das folgende Skript nutzt die im Beweis des Zwischenwertsatzes benutzte Intervallschachtelung, um eine Nullstelle einer vorgegebenen Funktion f zu berechnen. Man kann dies zum Beispiel zur Berechnung von reellen Wurzeln $\sqrt{a}$ nutzen, indem man sie auf $f(x) = x^2 - a$ anwendet. Diese Mehode ist allerdings wesentlich langsamer als die früher vorgestellte Methode mit Hilfe der Folge aus Satz 54.

```
  function erg = zws(f,a,b)
%
% zws(f,a,b) - berechnet eine Nullstelle von f zwischen a und b mittels
    Intervallhalbierungen, wie im Beweis des Zwischenwertsatzes.
% f kann hier bei eine Matlab-Funktion sein, die in der Form @Funktionsname
    übergeben werden muss. Oder
% f ist eine selbstdefinierte "anonyme" Funktion.
% Beispiele
% > f = @(x) x^2-2
% > zws(f,0,2)
% > zws(@sin, 1, 4)
%
 if(f(a)*f(b)>0)
  error('Es muss f(a)*f(b) kleiner oder gleich 0 sein.')
end
x=min(a,b);
y=max(a,b);
while(y-x>100*eps)
  c=(x+y)/2;
  if(f(x)*f(c)>0)
    x=c;
  else
    y=c;
  end
end
erg=(x+y)/2;
```

Ein Beispielskript, das die gerade definierte Funktion benutzt, könnte wie folgt aussehen:

```
clear all
```

```
% Wie definiert man eine anonyme Funktion?
f=@(x) sin(x.^2).*exp(x)-cos(x);
% Wie plottet man eine Funktion?
x=-1:0.1:3;
figure(1)
clf()
plot(x,f(x),[-1 3],[0 0])
xlabel('x')
ylabel('f(x)')
% Wie berechnet man eine Nullstelle?
zws(f,0,1)
```

Ende des Exkurses

Definition 132 (Verallgemeinertes Intervall)**:** Eine Teilmenge $I \subset \mathbb{R}$ heißt *(verallgemeinertes) Intervall*, wenn für alle $a, b, c \in \mathbb{R}$ mit $a < b < c$ gilt: Wenn $a, c \in I$ sind, dann ist auch $b \in I$.

BEMERKUNG: Man überlegt sich leicht, dass die früher eingeführten und als Intervalle bezeichneten Mengen $[a, b]$, $]a, b[$, $]a, b]$, und $[a, b[$ verallgemeinerte Intervalle sind. Aber auch die eingeführten uneigentlichen Intervalle, wie etwa $]a, +\infty[$, sind verallgemeinerte Intervalle.

Korollar 133: Ist $I \subset \mathbb{R}$ ein verallgemeinertes Intervall und $f : I \to \mathbb{R}$ stetig, so ist das Bild $f(I)$ ein verallgemeinertes Intervall.

BEWEIS: Es seien $a, b, c \in \mathbb{R}$ Punkte mit $a < b < c$ und $a, c \in f(I)$. Nun müssen wir zeigen, dass auch $b \in f(I)$ gilt. Es seien $x_1, x_2 \in I$ mit $f(x_1) = a$ und $f(x_2) = c$, also gilt

$$f(x_1) < b < f(x_2).$$

Ist $x_1 < x_2$, so existiert nach Korollar 131 ein $p \in [x_1, x_2]$ mit $f(p) = b$, also gilt $b \in f(I)$. Ist $x_1 > x_2$, so wendet man Korollar 131 auf $-f$ und $-b$ an, denn es ist

$$-f(x_2) < -b < -f(x_1).$$

□

Definition 134: Es sei $D \subset \mathbb{R}^r$ und $f : D \to \mathbb{R}$ eine Funktion.

a) Die Funktion f heißt *nach oben beschränkt*, wenn es ein $M \in \mathbb{R}$ gibt mit

$$f(x) \leq M \text{ für alle } x \in D.$$

b) Die Funktion f heißt *nach unten beschränkt*, wenn es ein $m \in \mathbb{R}$ gibt mit

$$f(x) \geq m \text{ für alle } x \in D.$$

c) Die Funktion f heißt *beschränkt*, wenn es $m, M \in \mathbb{R}$ gibt mit

$$m \leq f(x) \leq M \text{ für alle } x \in D.$$

d) Ein Punkt $x_0 \in D$ heißt *Maximalstelle von* f, wenn für alle $x \in D$

$$f(x) \leq f(x_0)$$

gilt. Man sagt dann auch, dass f in x_0 ein *Maximum* annimmt.

e) Ein Punkt $x_0 \in D$ heißt *Minimalstelle von* f, wenn für alle $x \in D$

$$f(x_0) \leq f(x)$$

gilt. Man sagt dann auch, dass f in x_0 ein *Minimum* annimmt.

f) Ein Punkt $x_0 \in D$ heißt *Extremstelle von* f, wenn x_0 eine Maximal- oder Minimalstelle ist. Man sagt dann auch, dass f in x_0 ein *Extremum* annimmt.

Satz 135: Jede stetige Funktion $f : [a, b] \to \mathbb{R}$ ist beschränkt und hat in $[a, b]$ (mindestens) eine Maximalstelle und eine Minimalstelle.

BEMERKUNG: Der Satz ist im Allgemeinen für

$$f : \begin{array}{l}]a, b[\to \mathbb{R} \\ ]a, b] \to \mathbb{R} \\ [a, b[\to \mathbb{R} \end{array}$$

falsch. Man betrachte etwa die Funktion $f :]0, a] \to \mathbb{R}$, welche durch $x \mapsto \frac{1}{x}$ definiert ist.

BEWEIS: Wir nehmen an, dass f nicht nach oben beschränkt ist. (Nach unten beschränkt geht analog.) Dann existiert zu jedem $n \in \mathbb{N}$ ein $y_n \in f([a, b])$ mit $y_n > n$. Zu jedem y_n existiert ein $x_n \in [a, b]$ mit $f(x_n) = y_n$. Die Folge (x_n) liegt in $[a, b]$ und ist somit beschränkt. Nach dem Satz 58 von Bolzano-Weierstraß besitzt (x_n) eine konvergente Teilfolge (x_{n_k}). Wir setzen $x_0 := \lim_{k\to\infty} x_{n_k}$, dann ist $x_0 \in [a, b]$, da alle $x_{n_k} \in [a, b]$ sind, siehe Schachtelungssatz.
Da f stetig ist, gilt

$$\lim_{k\to\infty} f(x_{n_k}) = f(x_0).$$

Nach Annahme ist die Folge $f(x_{n_k})$ jedoch unbeschränkt, also insbesondere divergent. Dies ist ein Widerspruch, damit ist gezeigt, dass f beschränkt ist.

Es sei nun

$$M := \sup\{f(x)|x \in [a,b]\}.$$

Wir müssen zeigen, dass $M \in f([a,b])$ ist. Es existiert zu jedem $\varepsilon > 0$ ein $y_\varepsilon \in f([a,b])$ mit

$$M - \varepsilon < y_\varepsilon \leq M.$$

Betrachten wir etwa $\varepsilon = \frac{1}{n}$ für $n \in \mathbb{N}$, so sehen wir, dass es zu jedem n ein $y_n \in f([a,b])$ mit

$$M - \frac{1}{n} < y_n \leq M.$$

gibt. Da $y_n \in f([a,b])$ ist, existiert zu jedem y_n ein $x_n \in [a,b]$ mit $f(x_n) = y_n$ und wir haben

$$M - \frac{1}{n} < f(x_n) \leq M.$$

Daraus folgt, wenn wir n gegen Unendlich laufen lassen, nach dem Schachtelungsprinzip

$$\lim_{n\to\infty} f(x_n) = M.$$

Die Folge (x_n) ist wieder beschränkt und wie oben folgt, dass es eine konvergente Teilfolge (x_{n_k}) gibt. Wir setzen wieder $x_0 := \lim_{k\to\infty} x_{n_k}$. Wegen der Stetigkeit von f gilt wieder

$$M = \lim_{k\to\infty} f(x_{n_k}) = f(x_0).$$

Der Beweis für die Minimalstelle erfolgt analog. □

Die Aussage aus Satz 135 lässt sich auch wie folgt formulieren:

Korollar 136: Das Bild eines abgeschlossenen Intervalls unter einer stetigen Funktion ist wieder ein abgeschlossenes Intervall, d.h. ist $f : [a,b] \to \mathbb{R}$ eine stetige Funktion, so existieren $c, d \in \mathbb{R}, c \leq d$, mit $f([a,b]) = [c,d]$.

Problem 52: *Beweise oder widerlege: Das Bild eines offenen Intervalls unter einer stetigen Funktion ist wieder ein offenes Intervall, d.h. ist* $f :]a,b[\to \mathbb{R}$ *eine stetige Funktion, so existieren* $c, d \in \mathbb{R}, c < d$, *mit* $f(]a,b[) =]c,d[$.

3.3. Funktionenfolgen und Stetigkeit

Nun wollen wir uns mit Folgen beschäftigen, deren Folgenglieder Funktionen sind. Man spricht dann von *Funktionenfolgen*. Wir wollen einen Konvergenzbegriff erarbeiten, d.h. wir wollen definieren, was wir darunter verstehen wollen, dass eine Funktionenfolge gegen eine Funktion konvergiert. Dabei soll der Konvergenzbegriff so gewählt sein, dass sich gewünschte Eigenschaften der Funktionen der Funktionenfolge auf die Grenzfunktion übertragen. Denken Sie hier etwa an Potenzreihen. Wir betrachten die Funktionenfolge (f_n), die gegeben ist durch $f_n(x) := \sum_{k=0}^{n} a_k(x-x_0)^k$. Dies sind alles Polynomfunktionen und somit recht einfach zu untersuchen. Z.B. sind alle f_n stetig. Falls der folgende Grenzwert existiert, haben wir $\sum_{n=0}^{i} nftya_n(x-x_0)^n = \lim_{n\to\infty} f_n(x)$. Es wäre doch schön, wenn man wüsste, dass sich die Stetigkeit auf die Grenzfunktion überträgt.

Naheliegend ist also der folgende Ansatz.

Definition 137 (Punktweise Konvergenz)**:** Es sei $I \subset \mathbb{R}$ und $f_n : I \to \mathbb{R}$ mit $n \in \mathbb{N}$ sei eine Folge von Funktionen. Die Folge (f_n) heißt *punktweise konvergent* gegen eine Funktion $f : I \to \mathbb{R}$, falls für alle $x \in I$

$$\lim_{n\to\infty} f_n(x) = f(x)$$

gilt.

Beispiel: Wir betrachten die Funktionenfolge (f_n) mit $f_n : [0,1] \to \mathbb{R}, x \mapsto x^n$. Für $x \in [0,1]$ gilt

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0, & x \in [0,1[\\ 1, & x = 1. \end{cases}$$

Die Funktionenfolge (f_n) konvergiert also punktweise gegen eine Funktion $f : [0,1] \to \mathbb{R}$ mit

$$f(x) := \begin{cases} 0, & x \in [0,1[\\ 1, & x = 1. \end{cases}$$

Im vorstehenden Beispiel haben wir eine Funktionenfolge stetiger Funktionen, die gegen eine unstetige Funktion konvergiert. Die Stetigkeit hat sich also nicht auf die Grenzfunktion übertragen. Die punktweise Stetigkeit ist für unsere Zwecke der falsche Konvergenzbegriff.

Deshalb wollen wir nun einen anderen Konvergenzbegriff einführen. Wir lassen uns dabei von der Definition der Konvergenz bei Zahlenfolgen inspirieren: $\forall \epsilon \exists n \in \mathbb{N} \forall n \geq n_0 : |a_n - a| < \varepsilon$. Wenn nun die (a_n) keine Zahlen, sondern Funktionen sein sollen, dann benötigen wir eine Art von Betragsfunktion auf der Menge der Funktionen.

Definition 138 (Supremumsnorm)**:** Es sei $I \subset \mathbb{R}$ und $f : I \to \mathbb{R}$ eine beschränkte Funktion. Wir definieren die *Supremumsnorm von f auf I* durch

$$\|f\|_I := \sup\{|f(x)|\,|x \in I\}.$$

BEMERKUNG: Ist $I := [a, b]$ und f stetig, so gilt

$$\|f\|_I = \max\{|f(x)|\,|x \in I\}.$$

Man spricht dann auch von der Maximumsnorm.

Die Funktion $\|.\|_I$ hat viele Eigenschaften, die die Betragsfunktion auch hat.

Lemma 139: Es sei $I \subset \mathbb{R}, \alpha \in \mathbb{R}$ und $f, g : I \to \mathbb{R}$ beschränkte Funktionen. Dann gelten:

(i) $\|f\|_I \geq 0$,

(ii) $\|f\|_I = 0$ gilt genau dann, wenn $f(x) = 0$ für alle $x \in I$ ist,

(iii) $\|\alpha \cdot f\|_I = |\alpha| \cdot \|f\|_I$,

(iv) $\|f \cdot g\|_I \leq \|f\|_I \cdot \|g\|_I$,

(v) $\|f + g\|_I \leq \|f\|_I + \|g\|_I$.

Problem 53: *Beweisen Sie dieses Lemma und geben Sie ein Beispiel an, in dem $\|f \cdot g\|_I < \|f\|_I \cdot \|g\|_I$ ist.*

Der folgende Satz liefert eine äquivalente Charakterisierung der gleichmäßigen Konvergenz.

Definition 140: Es sei $I \subset \mathbb{R}$, $f : I \to \mathbb{R}$ eine Funktion und (f_n) eine Folge von Funktionen $f_n : I \to \mathbb{R}$. Die Funktionenfolge (f_n) heißt *gleichmäßig konvergent* gegen die Funktion f, falls

$$\lim_{n\to\infty} \|f_n - f\|_I = 0$$

ist.

Problem 54: *Zeigen Sie: Gibt es ein $\varepsilon_0 > 0$ und eine Folge (x_n) in I, so dass $|f_n(x_n) - f(x_n)| \geq \varepsilon_0$ für unendlich viele n gilt, so kann (f_n) nicht gleichmäßig gegen f konvergieren.*

Problem 55: *Es sei $I \subset \mathbb{R}$, $f : I \to \mathbb{R}$ eine Funktion und (f_n) eine Folge von Funktionen $f_n : I \to \mathbb{R}$. Zeigen Sie*

(i) Die Folge (f_n) konvergiert genau dann punktweise gegen f, falls
$$\forall x \in I\ \forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N :\ |f_n(x) - f(x)| < \varepsilon.$$

(ii) Die Folge (f_n) konvergiert genau dann gleichmäßig gegen f, falls
$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N\ \forall x \in I :\ |f_n(x) - f(x)| < \varepsilon$$

In der vorstehenden Aufgabe erkennt man sofort, dass die Eigenschaft der gleichmäßigen Konvergenz stärker ist als die der punktweisen Konvergenz. Insbesondere konvergiert jede Funktionenfolge, die gleichmäßig gegen eine Grenzfunktion konvergiert, auch punktweise gegen diese Grenzfunktion. Die punktweise Konvergenz ist also nicht unnütz. Sie kann uns helfen, die Grenzfunktion zu finden.

Beispiel: Wir betrachten noch einmal die Funktionenfolge (f_n) mit $f_n : [0,1] \to \mathbb{R}, x \mapsto x^n$ und ihre punktweise Grenzfunktion
$$f(x) := \lim_{n\to\infty} f_n(x) = \begin{cases} 0, & x \in [0,1[\\ 1, & x = 1. \end{cases}$$
Wir wollen zeigen, dass (f_n) nicht gleichmäßig gegen f konvergiert. Es ist $f_n(1) - f(1) = 0$. Somit ist für festes n
$$\begin{aligned} \|f_n - f\|_{[0,1]} &= \sup\{|f_n(x) - f(x)| \mid x \in [0,1]\} \\ &= \sup\{|f_n(x) - f(x)| \mid x \in [0,1[\} \\ &= \sup\{|x^n| \mid x \in [0,1[\} = 1. \end{aligned}$$

Dann ist aber auch
$$\lim_{n\to\infty} \|f_n - f\|_I = 1$$
und (f_n) konvergiert nicht gleichmäßig gegen f. (f_n) kann auch nicht gegen eine andere Funktion gleichmäßig konvergieren, da die Folge gegen diese Funktion dann auch punktweise konvergieren müsste. Die vorstehende Rechnung zeigt, dass
$$\|f_n - f\|_{[0,1[} = \|f_n - f\|_{[0,1]}$$
ist. Es reicht also nicht, wenn wir nur die 1 aus dem Definitionsbereich entfernen. Auch auf dem Intervall $[0,1[$ konvergiert die Funktionenfolge nicht gleichmäßig gegen die Grenzfunktion.
Wenn wir die Funktionenfolge allerdings auf einer abgeschlossenen echten Teilmenge des Definitionsbereichs der Form $[0,q]$ mit einem $0 < q < 1$ betrachten, so konvergiert die Funktionen dort gleichmäßig gegen die Grenzfunktion f, die dort gleich der Nullfunktion ist. Es ist dann nämlich
$$\|f_n - f\|_{[0,q]} = \sup\{|f_n(x) - f(x)| \mid x \in [0,q]\} = \sup\{|x^n| \mid x \in [0,q]\} = q^n$$
und somit
$$\lim_{n\to\infty} \|f_n - f\|_{[0,q]} = 0.$$

Problem 56: *Die Funktionen $f_n : \mathbb{R} \to \mathbb{R}$ seien definiert durch*

$$f_n(x) := \frac{x^{2n}}{1 + x^{2n}}.$$

(i) *Gegen welche Funktion $f : \mathbb{R} \to \mathbb{R}$ konvergiert die Funktionenfolge (f_n) punktweise?*

(ii) *Zeigen Sie, dass (f_n) auf $\mathbb{R}$ nicht gleichmäßig gegen f konvergiert.*

(iii) *Es sei $0 < q < 1$. Zeigen Sie, dass die Funktionenfolge (f_n) gleichmäßig konvergiert, wenn wir den Definitionsbereich der Funktionen auf die Menge $\{x \in \mathbb{R} | |x| \leq q\}$ einschränken.*

(iv) *Es sei $q > 1$. Zeigen Sie, dass die Funktionenfolge (f_n) gleichmäßig konvergiert, wenn wir den Definitionsbereich der Funktionen auf die Menge $\{x \in \mathbb{R} | |x| \geq q\}$ einschränken.*

Satz 141: Konvergiert eine Folge stetiger Funktionen (f_n) gleichmäßig gegen eine Funktion f, dann ist f ebenfalls stetig.

BEWEIS: Wir müssen zeigen, dass f in jedem beliebigen x' des Definitionsbereichs stetig ist. Dazu benutzen wir die ε-δ-Stetigkeit, d.h. wir zeigen

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x : \; |x - x'| < \delta \Rightarrow |f(x) - f(x')| < \varepsilon.$$

Es sei $\varepsilon > 0$ beliebig vorausgesetzt. Da (f_n) gleichmäßig konvergiert, existiert ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ und x, x'

$$|f_n(x) - f(x)| < \varepsilon \text{ und}$$
$$|f(x') - f_n(x')| < \varepsilon$$

gilt. Da die Funktion f_n stetig in x' ist, existiert zu ε ein $\delta > 0$, so dass für alle x mit $|x - x'| < \delta$

$$|f_n(x) - f_n(x')| < \varepsilon$$

gilt. Damit folgt

$$\begin{aligned} |f(x) - f(x')| &= |f(x) - f_n(x) + f_n(x) - f_n(x') + f_n(x') - f(x')| \\ &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(x')| + |f_n(x') - f(x')| \\ &\leq 3\varepsilon \end{aligned}$$

□

Satz 142: Es sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe mit Konvergenzradius $R>0$. Außerdem sei $0<r<R$. Dann konvergiert die Funktionenfolge

$$\begin{aligned} f_n : [x_0-r, x_0+r] &\to \mathbb{R} \\ x &\mapsto \sum_{k=0}^{n} a_k(x-x_0)^k \end{aligned}$$

gleichmäßig gegen

$$\begin{aligned} f : [x_0-r, x_0+r] &\to \mathbb{R} \\ x &\mapsto \sum_{k=0}^{\infty} a_k(x-x_0)^k. \end{aligned}$$

BEWEIS: Wir argumentieren zunächst wie im Beweis von Lemma 104. Ist $\rho < R$, so konvergiert die Reihe $\sum_{k=0}^{\infty} a_k\rho^k$. Also ist $(a_k\rho^k)$ eine Nullfolge und es existiert ein $C \in \mathbb{R}_+$, so dass für alle $k \in \mathbb{N}_0$

$$|a_k\rho^k| \le C$$

gilt. Für x mit $|x-x_0| \le r$ und $r<\rho<R$ gilt also

$$\left|a_k(x-x_0)^k\right| = \left|a_k\rho^k \frac{(x-x_0)^k}{\rho^k}\right| \le C\theta^k$$

mit $\theta := \frac{r}{\rho} \in]0,1[$. Da die geometrische Reihe $\sum_{k=0}^{\infty} C\theta^k$ konvergiert, existiert zu jedem $\varepsilon>0$ ein $N \in \mathbb{N}$, so dass

$$\sum_{k=N+1}^{\infty} C\theta^k < \varepsilon$$

gilt. Für $n \ge N$ gilt also

$$\begin{aligned} |f_n(x)-f(x)| &= \left|\sum_{k=n+1}^{\infty} a_k(x-x_0)^k\right| \\ &\le \sum_{k=n+1}^{\infty} C\theta^k < \varepsilon. \end{aligned}$$

□

Korollar 143: Es sei $P(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe mit Konvergenzradius $R>0$. Dann ist die Funktion $P :]x_0-R, x_0+R[\to \mathbb{R}, x \mapsto P(x)$ stetig.

BEMERKUNG: Nach dem vorstehenden Korollar sind insbesondere die trigonometrischen Funktionen $\sin : \mathbb{R} \to \mathbb{R}$ und $\cos : \mathbb{R} \to \mathbb{R}$ stetig.

Lemma 144: Es gilt

$$\lim_{x\to 0} \frac{\sin x}{x} = 1$$

und

$$\lim_{x\to 0} \frac{\cos x - 1}{x} = 0.$$

BEWEIS: Für $x \neq 0$ ist

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

Rechts steht eine Potenzreihe mit Entwicklungspunkt 0 und Konvergenzradius Unendlich. Diese Potenzreihe konvergiert natürlich in 0 und nimmt dort den Wert 1 an. Da die Potenzreihe eine stetige Funktion darstellt, folgt die Behauptung sofort. □

Problem 57: *Beweisen Sie den zweiten Grenzwert in Lemma 144.*

Satz 145 (Identitätssatz für Potenzreihen)**:** Es sei $P(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit positivem Konvergenzradius. Außerdem sei (α_k) eine Nullfolge mit $\alpha_k \neq 0$ für alle $k \in \mathbb{N}$. Gilt dann

$$P(\alpha_k) = 0$$

für alle hinreichend großen k (das heißt $\exists k_0 \forall k \geq k_0 \ldots$), so gilt

$$a_n = 0$$

für alle $n \in \mathbb{N}_0$.

BEWEIS: Der Beweis erfolgt mit Induktion.

Induktionsverankerung (n = 0):

$$a_0 = P(0) = \lim_{n\to\infty} P(\alpha_n) = \lim_{n\to\infty} 0 = 0$$

Induktionsschritt ($n \to n+1$): Auf Grund der Induktionsverankerung wissen wir, dass $a_0 = \ldots = a_n = 0$ gilt. Setze

$$Q(z) = \frac{P(z)}{z^{n+1}} = a_{n+1} + a_{n+2} z + \ldots .$$

Damit folgt weiter

$$a_{n+1} = Q(0) = \lim_{k\to\infty} Q(\alpha_k) = \lim_{k\to\infty} \frac{P(\alpha_k)}{(\alpha_k)^{n+1}} = 0.$$

□

Korollar 146 (Alternative Version des Identitätssatzes für Potenzreihen)**:** Sind $P(z) = \sum_{n=0}^{\infty} a_n z^n$ und $Q(z) = \sum_{n=0}^{\infty} b_n z^n$ zwei Potenzreihen mit positivem Konvergenzradius. Außerdem sei die Folge (α_k) wie in Satz 145. Gilt dann auch noch $P(\alpha_k) = Q(\alpha_k)$ für alle $k \geq k_0$, so folgt für alle $n \in \mathbb{N}_0$

$$a_n = b_n.$$

BEWEIS: Wir erhalten das Korollar, wenn wir den Identitätssatz für Potenzreihen auf die Reihe

$$R(z) := P(z) - Q(z) = \sum_{n=0}^{\infty} (a_n - b_n) z^n$$

anwenden. □

Der Identitätssatz ist auch die Grundlage für den sogenannten *Potenzreihenansatz*, mit dem sich viele Probleme lösen lassen. Wir werden später sehen, wie man hiermit sogenannte Differentialgleichungen lösen kann.

Beispiel: Wir suchen eine Funktion f, die der Funktionalgleichung

$$f(x) + f(x)^2 = 0$$

zumindest in der Nähe von 0 genügt. Wir nehmen an, dass sich die Funktion f als Potenzreihe mit Entwicklungspunkt 0 darstellen lässt, also dass in der Nähe von 0

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

gilt. Wenn wir diese Potenzreihe in die Funktionalgleichung einsetzen, so erhalten wird

$$0 = \sum_{n=0}^{\infty} a_n x^n + \left(\sum_{n=0}^{\infty} a_n x^n \right)^2.$$

Das Quadrat der Potenzreihe können wir mit dem Satz vom Cauchy-Produkt, Satz 102, ausrechnen. Es ist

$$\left(\sum_{n=0}^{\infty} a_n x^n \right)^2 = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k a_{n-k} x^n.$$

Wir erhalten also, wenn wir dies in die obige Gleichung einsetzen,

$$0 = \sum_{n=0}^{\infty} \left(a_n + \sum_{k=0}^{n} a_k a_{n-k} \right) x^n.$$

In der vorstehenden Gleichung steht rechts eine Potenzreihe. Diese ist nur dann gleich der Nullfunktion, wenn alle Koeffizienten gleich Null sind. Es muss also für alle $n \in \mathbb{N}_0$

$$a_n + \sum_{k=0}^{n} a_k a_{n-k} = 0$$

gelten. Für $n = 0$ erhalten wir $0 = a_0 + a_0^2 = a_0(1 - a_0)$; es muss also $a_0 = 0$ oder $a_0 = -1$ gelten. Für $n = 1$ erhalten wir $0 = a_1 + a_0 a_1 + a_1 a_0$. Ist $a_0 = 0$, so ist also auch $a_1 = 0$. Für $a_0 = -1$ ergibt sich aber auch $0 = a_1 - 2a_1 = -a_1$, also $a_1 = 0$. Induktiv zeigt man sehr leicht, dass sowohl für $a_0 = 0$ als auch für $a_0 = -1$ stets $a_n = 0$ für alle $n \geq 1$ gelten muss. Die einzigen beiden Lösungsfunktionen, die sich als Potenzreihen schreiben lassen, sind also die beiden konstanten Funktionen $f = 0$ und $f = -1$.

Wir schauen uns noch ein Beispiel an, wo man die Potenzreihe einer Funktion mittels *Koeffizientenvergleich* bestimmt.

Beispiel: Wir wollen schauen, ob man die Funktion f, die in der Nähe von Null durch

$$f(x) = \frac{1 + x^2}{1 + x}$$

gegeben ist, als eine Potenzreihe mit Entwicklungspunkt 0 darstellen kann. Wir machen den Ansatz

$$f(x) = \frac{1 + x^2}{1 + x} = \sum_{n=0}^{\infty} a_n x^n.$$

Multiplikation mit $1 + x$ liefert uns

$$1 + x^2 = (1 + x) \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} a_n x^{n+1} = \sum_{n=0}^{\infty} a_n x^n + \sum_{n=1}^{\infty} a_{n-1} x^n.$$

Wir haben also

$$1 + x^2 = a_0 + \sum_{n=1}^{\infty} (a_n + a_{n-1}) x^n.$$

Auf beiden Seiten der Gleichung steht eine Potenzreihe. Da die beiden Potenzreihen gleich sein sollen, müssen die Koeffizientenfolgen übereinstimmen. Wir erhalten also

$$\begin{aligned} 1 &= a_0 \\ 0 &= a_1 + a_0 \\ 1 &= a_2 + a_1 \\ 0 &= a_n + a_{n-1} \text{ für alle } n \geq 3. \end{aligned}$$

Wir erhalten also $a_0 = 1, a_1 = -1, a_2 = 2$ und $a_n = -a_{n-1}$ für $n \geq 3$. Also ist

$$f(x) = 1 - x + \sum_{n=2}^{\infty} 2 \cdot (-1)^n x^n.$$

Wir müssen aber zugeben, dass man dies auch schneller aus der geometrischen Reihe ableiten kann. Es ist nämlich für $|x| < 1$

$$\frac{1}{1+x} = \frac{1}{1-(-x)} = \sum_{n=0}^{\infty}(-x)^n = \sum_{n=0}^{\infty}(-1)^n x^n.$$

Also ist

$$\begin{aligned} \frac{1+x^2}{1+x} &= (1+x^2)\sum_{n=0}^{\infty}(-1)^n x^n = \sum_{n=0}^{\infty}(-1)^n x^n + \sum_{n=0}^{\infty}(-1)^n x^{n+2} \\ &= \sum_{n=0}^{\infty}(-1)^n x^n + \sum_{n=2}^{\infty}(-1)^n x^n = 1 - x + 2 \cdot \sum_{n=2}^{\infty}(-1)^n x^n. \end{aligned}$$

Exkurs: Polynome und Polynomfunktionen

Ein Polynom über einem Körper K ist ein formaler Ausdruck der Form $a_0 + a_1 X + a_2 X^2 + \ldots$, wobei die Koeffizienten $a_0, a_1, a_2, \ldots \in K$ sind und alle Koeffizienten bis auf endlich viele gleich Null sind. Per Definition sind zwei Polynome gleich, wenn sie die gleichen Koeffizienten haben. Die Menge der Polynome über einem Körper K ist damit identisch zu der Menge aller Folgen in K, für die alle bis auf endlich viele Folgenglieder gleich Null sind.

Sind alle Koeffizienten mit einem Index größer als n gleich Null, so schreiben wir $a_0 + a_1 X + \ldots + a_n X^n$ für das Polynom.

Einem Polynom $a_0 + a_1 X + \ldots + a_n X^n$ über dem Körper K kann man nun eine Funktion $P : K \to K, x \mapsto a_0 + a_1 x + \ldots + a_n x^n$ zuordnen. Man nennt eine solche Funktion eine Polynomfunktion. Sind P und Q zwei Polynomfunktionen über einem Körper K, so sind sie per definitionem gleich, wenn $P(x) = Q(x)$ für alle $x \in K$ gilt.

Die beiden Gleichheitsbegriffe für Polynome und für Polynomfunktionen fallen nicht unbedingt zusammen.

Betrachtet man beispielsweise den endlichen Körper $\mathbb{Z}_2$ und das Polynom X^2+X. Für die zugehörige Polynomfunktion P gilt $P(0) = 0$ und $P(1) = 0$. Die Polynomfunktion P ist also die Nullfunktion, aber das Polynom ist nicht das Nullpolynom, da nicht alle Koeffizienten gleich Null sind.

In dem Körper der reellen Zahlen $\mathbb{R}$ und dem Körper der komplexen Zahlen $\mathbb{C}$ fallen die beiden Gleichheitsbegriffe aber zusammen, weshalb man, wenn man in diesem Körpern arbeitet, die Begriffe Polynom und Polynomfunktion häufig synonym benutzt und nicht unterscheidet.

Korollar 147: Für Polynomfunktionen $f, g : \mathbb{K} \to \mathbb{K}$ mit

$$f(z) = a_0 + a_1 z + \ldots + a_n z^n$$
$$g(z) = b_0 + b_1 z + \ldots + b_m z^m$$

mit $a_0, \ldots, a_n, b_0 \ldots b_n \in \mathbb{K}$ sind die folgenden zwei Aussagen äquivalent.

(i) Für alle $z \in \mathbb{K}$ gilt $f(z) = g(z)$.

(ii) Es gilt $m = n$ und für alle $k \in \{0, \ldots n\}$ gilt $a_k = b_k$.

BEWEIS: Von (ii) nach (i) ist klar. Die Gegenrichtung von (i) nach (ii) folgt aus dem Identitätssatz. □

Ende des Exkurses

3.4. Umkehrfunktionen monotoner Funktionen

Definition 148 (Umkehrfunktion)**:** Es sei $I \subset \mathbb{R}$ und $f : I \to \mathbb{R}$ eine Funktion. Dann heißt $f^{-1} : f(I) \to I$ eine *Umkehrfunktion von* f, wenn für alle $x \in I$

$$(f^{-1} \circ f)(x) = x$$

gilt.

BEMERKUNG: Man verwechsle die Umkehrfunktion f^{-1} nicht mit der Funktion $x \mapsto \frac{1}{f(x)}$.

Lemma 149: Es sei $I \subset \mathbb{R}$ und $f : I \to \mathbb{R}$ eine Funktion. Dann sind äquivalent:

(i) Die Funktion f besitzt eine Umkehrfunktion.

(ii) Die Funktion $f : I \to f(I)$ ist injektiv und surjektiv.

(iii) Es gibt eine eindeutig definierte Funktion $f^{-1} : f(I) \to I$, so dass für alle $x \in I$

$$(f^{-1} \circ f)(x) = x$$

und für alle $y \in f(I)$
$$(f \circ f^{-1})(y) = y$$
gilt.

BEWEIS:

(i)⇒(ii): Die Surjektivität von $f : I \to f(I)$ ist offensichtlich. Wir zeigen nun, dass die Funktion auch injektiv ist. Es seien $x_1, x_2 \in I$ mit $f(x_1) = f(x_2)$. Wenn wir auf diese Gleichung f^{-1} anwenden, dann erhalten wir
$$x_1 = f^{-1}(f(x_1)) = f^{-1}(f(x_2)) = x_2.$$

(ii)⇒(iii): Da f injektiv ist, existiert zu jedem $y \in f(I)$ ein eindeutig bestimmtes $x \in I$ mit $f(x) = y$. Ist $g : f(I) \to I$ eine Funktion derart, dass $f(g(y)) = y$ ist, so muss also $g(y) = x$ sein, wenn $f(x) = y$ ist. Die einzige Möglichkeit, die Umkehrfunktion $f^{-1} : f(I) \to I$ zu definieren, ist also, dass wir $f^{-1}(y) := x$ setzen, falls $f(x) = y$. Man rechnet leicht nach, dass diese Funktion die beiden Eigenschaften aus (iii) erfüllt.

(iii)⇒(i): Trivial.

□

BEMERKUNG: Angesichts von (iii) im vorstehenden Satz könnte man sich eventuell fragen, ob man in Definition 148 anstelle von
$$\forall x \in I : (f^{-1} \circ f)(x) = x$$
auch
$$\forall y \in f(I) : (f \circ f^{-1})(y) = y$$
fordern könnte. Dies geht aber nicht. Wenn wir uns etwa die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$ anschauen, so ist $f(\mathbb{R}) = \mathbb{R}_{\geq 0}$ und für die Funktion $g : \mathbb{R}_{\geq 0} \to \mathbb{R}$ mit $g(y) = \sqrt{y}$ gilt für alle $y \in \mathbb{R}_{\geq 0}$
$$(f \circ g)(y) = f(\sqrt{y}) = (\sqrt{y})^2 = y.$$
Es ist aber $(g \circ f)(-1) = \sqrt{(-1)^2} = 1$. g ist also nicht die Umkehrfunktion von f. Da f nicht injektiv ist, besitzt f keine Umkehrfunktion.

Definition 150 (Monotone Funktionen)**:** Es sei $D \subset \mathbb{R}$. Die Funktion $f : D \to \mathbb{R}$ heißt

(i) *monoton wachsend* genau dann, wenn für alle $x, x' \in D$ mit $x < x'$ die Ungleichung $f(x) \leq f(x')$ gilt. Gilt sogar stets die strikte Ungleichung $f(x) < f(x')$, so nennen wir f *streng monoton wachsend*.

(ii) *monoton fallend* genau dann, wenn für alle $x, x' \in D$ mit $x < x'$ die Ungleichung $f(x) \geq f(x')$ gilt. Gilt sogar stets die strikte Ungleichung $f(x) > f(x')$, so nennen wir f *streng monoton fallend.*

Problem 58: *Es sei I ein nicht-leeres Intervall. Beweisen oder widerlegen Sie die folgenden Aussagen:*

(i) Jede stetige Funktion $f : I \to \mathbb{R}$ ist beschränkt.

(ii) Besitzt eine stetige Funktion $f : I \to \mathbb{R}$ eine Nullstelle, so existieren $x_1, x_2 \in I$ mit $f(x_1) < 0$ und $f(x_2) > 0$.

(iii) Ist $f : I \to \mathbb{R}$ stetig und streng monoton wachsend, so existiert zu jedem $y_0 \in f(I)$ genau ein $x_0 \in I$ mit $f(x_0) = y_0$.

(iv) Ist $f : I \to \mathbb{R}$ stetig und monoton wachsend, so existiert zu jedem $y_0 \in f(I)$ genau ein $x_0 \in I$ mit $f(x_0) = y_0$.

Satz 151: Es sei $f : [a, b] \to \mathbb{R}$ eine stetige, streng monoton wachsende Funktion. Setze $A := f(a)$, $B := f(b)$. Dann bildet f das Intervall $[a, b]$ bijektiv auf $[A, B]$ ab und es existiert die Umkehrfunktion $f^{-1} : [A, B] \to [a, b]$. f^{-1} ist ebenfalls stetig und streng monoton wachsend.

BEMERKUNG: Man kann streng monoton wachsend durch streng monoton fallend ersetzen. Dann ist auch $f^{-1} : [B, A] \to [a, b]$ streng monoton fallend.

BEWEIS: Aus $a < x < b$ folgt wegen der Monotonie $f(a) < f(x) < f(b)$. Also gilt $f([a, b]) \subseteq [A, B]$. Nach dem Zwischenwertsatz 131 wird jeder Wert in $[A, B]$ von f angenommen, das heißt f bildet $[a, b]$ surjektiv auf $[A, B]$ ab. Da f streng monoton wachsend ist, ist f injektiv, denn aus $x < x'$ folgt $f(x) < f(x')$. Somit bildet f das Intervall $[a, b]$ bijektiv auf $[A, B]$ ab. Folglich existiert die Umkehrfunktion

$$f^{-1} : [A, B] \to [a, b].$$

Nun zeigen wir, dass die Umkehrfunktion f^{-1} streng monoton wachsend ist. Es seien $y, y' \in [A, B]$ mit $y < y'$, dann existieren $x, x' \in [a, b]$ mit $f(x) = y$ und $f(x') = y'$ beziehungsweise $f^{-1}(y) = x$ und $f^{-1}(y') = x'$. Wäre $x \geq x'$, so auch $y \geq y'$ wegen der Monotonie von f. Also gilt

$$f^{-1}(y) = x < x' = f^{-1}(y').$$

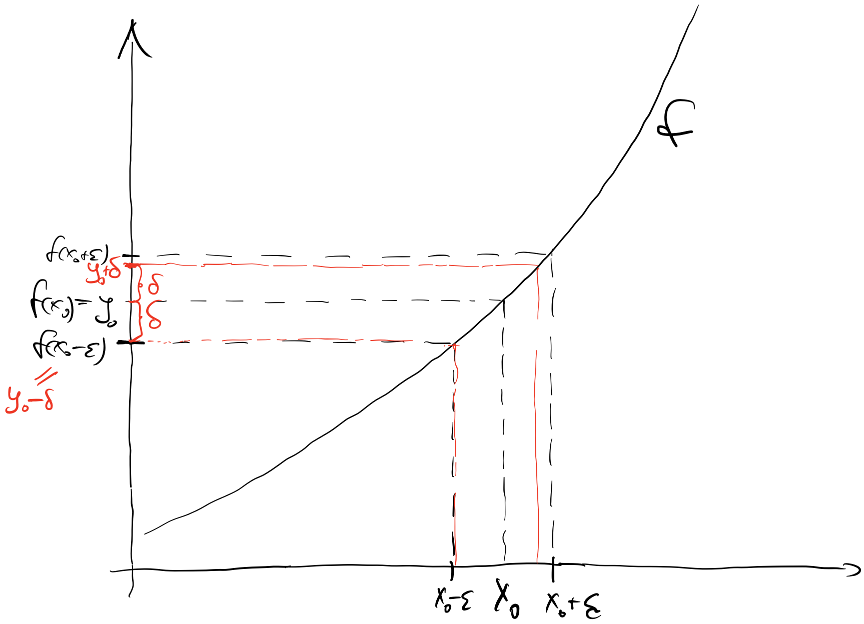


Abbildung 3.2.: Skizze zum Beweis der Stetigkeit der Umkehrfunktion von f

Zuletzt zeigen wir noch die Stetigkeit von f^{-1}.
Wenn wir die Definition der $\varepsilon-\delta-$Stetigkeit benutzen, so müssen wir zeigen:

$$\forall y_0 \in [A,B] \forall \varepsilon > 0 \exists \delta > 0 \forall y \in [A,B] : |y - y_0| < \delta \Rightarrow |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon.$$

Seien also $y_0 \in [A,B]$ und $\varepsilon > 0$ beliebig. Zu y_0 existiert ein $x_0 \in [a,b]$ mit $f(x_0) = y_0$. Wegen der Monotonie haben wir

$$f(x_0-\varepsilon) < f(x_0) < f(x_0+\varepsilon).$$

Wir setzen nun δ auf die kleinere der beiden Zahlen $|f(x_0) - f(x_0-\varepsilon)|$ und $|f(x_0) - f(x_0+\varepsilon)|$, also

$$\delta := \min\{|f(x_0) - f(x_0-\varepsilon)|, |f(x_0) - f(x_0+\varepsilon)|\}.$$

Dann ist $]y_0-\delta, y_0+\delta[=]f(x_0)-\delta, f(x_0)+\delta[\subset]f(x_0-\varepsilon), f(x_0+\varepsilon)[$. Ist nun $y \in]y_0-\delta, y_0+\delta[$ beliebig, also $|y-y_0| < \delta$, so haben wir also $f(x_0-\varepsilon) < y < f(x_0+\varepsilon)$ und es existiert aufgrund des Zwischenwertsatzes ein $x \in]x_0-\varepsilon, x_0+\varepsilon[$ mit $f(x) = y$. Hiermit erhalten wir

$$|f^{-1}(y) - f^{-1}(y_0)| = |f^{-1}(f(x)) - f^{-1}(f(x_0))| = |x - x_0| < \varepsilon.$$

Alternativ: Wenn man die Definition der Folgenstetigkeit benutzt, dann sieht der Beweis in etwa wie folgt aus:
Es sei $y \in [A,B]$ und y_n eine Folge in $[A,B]$ mit $\lim_{n\to\infty} y_n = y$. Nun werden wir $\lim_{n\to\infty} f^{-1}(y_n) = f^{-1}(y)$ beweisen. Angenommen dies gilt nicht, so existiert ein $\varepsilon > 0$, so dass für alle $n_0 \in \mathbb{N}$ ein $n \geq n_0$ existiert mit $|f^{-1}(y_n) - f^{-1}(y)| \geq \varepsilon$. Es gibt also ein $\varepsilon > 0$ und eine Teilfolge (y_{n_k}) mit

$$|f^{-1}(y_{n_k}) - f^{-1}(y)| \geq \varepsilon.$$

Die Folge $(f^{-1}(y_{n_k}))$ liegt in $[a,b]$, somit ist sie beschränkt. Nach dem Satz von Bolzano-Weierstraß existiert eine konvergente Teilfolge $f^{-1}\left(y_{n_{k_i}}\right)$ mit Grenzwert $c \in [a,b]$. Da f stetig ist, gilt

$$\begin{aligned} f(c) &= \lim_{i\to\infty} f\left(f^{-1}\left(y_{n_{k_i}}\right)\right) \\ &= \lim_{i\to\infty} y_{n_{k_i}} \\ &= y. \end{aligned}$$

Daraus folgt

$$c = f^{-1}(f(c)) = f^{-1}(y),$$

also gilt $\lim_{i\to\infty} f^{-1}\left(y_{n_{k_i}}\right) = f^{-1}(y)$, was aber unserer Annahme widerspricht. □

Nun ist der vorstehende Satz nur für Funktionen definiert, die auf einem abgeschlossenen Intervall definiert sind. Aus dem Satz lässt sich aber folgendes Korollar ableiten.

Korollar 152: Es sei I ein verallgemeinertes Intervall und $f : I \to \mathbb{R}$ eine stetige, streng monoton wachsende Funktion. Dann bildet f das Intervall I bijektiv auf $f(I)$ ab, was wieder ein verallgemeinertes Intervall ist, und es existiert die Umkehrfunktion $f^{-1} : f(I) \to I$. f^{-1} ist ebenfalls stetig und streng monoton wachsend.

Problem 59: *Beweisen Sie das Korollar.*

Beispiel: Man kann den vorstehenden Satz bzw. das Korollar nutzen, um die Wurzel-Funktion einzuführen.[1] Die Funktion $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, x \mapsto x^2$ ist offensichtlich eine bijektive, stetige, streng monoton wachsende Funktion. Sie besitzt also eine ebensolche Umkehrfunktion.

Problem 60: *Es sei $f : \mathbb{R} \to \mathbb{R}$ eine Funktion. Die Funktionen $f^2, f^3, \ldots$ sind punktweise definiert über $f^2(x) = (f(x))^2, f^3(x) = (f(x))^3, \ldots$. Beweisen oder widerlegen Sie die folgenden Aussagen:*

(i) Ist f stetig, so sind auch f^2 und f^3 stetig.

(ii) Ist f^2 stetig, so ist auch f stetig.

(iii) Ist f^3 stetig, so ist auch f stetig.

[1] Dies ist mindestens die dritte Art, wie wir in diesem Skript die Wurzel definieren.

3.5. Die Logarithmusfunktion und die allgemeinen Potenzfunktionen

Nun kann man diesen Satz auf die Exponentialfunktion anwenden.

Satz 153 (Logarithmus)**:** Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ ist streng monoton wachsend und bildet $\mathbb{R}$ auf $\mathbb{R}_+$ ab. Die Umkehrfunktion

$$\log : \mathbb{R}_+ \to \mathbb{R}$$

ist also stetig, streng monoton wachsend und heißt der (natürliche) *Logarithmus*. Für alle $x, y \in \mathbb{R}_+$ gilt die Funktionalgleichung

$$\log(xy) = \log(x) + \log(y).$$

BEWEIS: Wir müssen zeigen, dass

- die Funktion $\exp$ streng monoton wachsend ist,
- $\exp$ die reellen Zahlen $\mathbb{R}$ auf $\mathbb{R}_+$ abbildet,
- die Funktionalgleichung gilt.

Für $x > 0$ gilt

$$\exp(x) = 1 + x + \sum_{n=2}^{\infty} \frac{x^n}{n!} \geq 1 + x > 1$$

und somit folgt für $x, x' \in \mathbb{R}$ mit $x < x'$

$$\exp(x') = \exp(x' - x + x) = \underbrace{\exp(x' - x)}_{>1} \exp(x) > \exp(x).$$

Also ist die Funktion $\exp$ streng monoton wachsend.
Um zu zeigen, dass $\exp$ den $\mathbb{R}$ auf $\mathbb{R}_+$ abbildet, betrachten wir zunächst die Grenzwerte

$$\lim_{x\to\infty} \exp(x) \quad \text{und}$$
$$\lim_{x\to-\infty} \exp(x).$$

Wegen $\exp(x) \geq 1 + x$ gilt $\lim_{x\to\infty} \exp(x) = +\infty$. Außerdem gilt weiter

$$0 < \exp(-x) = \frac{1}{\exp(x)} \leq \frac{1}{1+x},$$

woraus sofort $\lim_{x\to-\infty}\exp(x) = 0$ folgt. Da $\exp(x) > 0$ für alle $x \in \mathbb{R}$ gilt und da $\exp$ stetig ist, folgt

$$\exp(\mathbb{R}) = \mathbb{R}_+.$$

Es seien nun $x, y \in \mathbb{R}_+$ beliebige Zahlen. Dann existieren $\alpha, \beta \in \mathbb{R}$ mit

$$x = \exp(\alpha) \qquad \text{und} \qquad y = \exp(\beta)$$

beziehungsweise $\log(x) = \alpha$ und $\log(y) = \beta$. Mit dieser Festlegung gilt nun

$$\begin{aligned}\log(xy) &= \log(\exp(\alpha)\exp(\beta))\\ &= \log(\exp(\alpha+\beta))\\ &= \alpha+\beta\\ &= \log(x)+\log(y).\end{aligned}$$

□

Bisher wurden nur Potenzen a^r für $a \in \mathbb{R}_+$ und $r \in \mathbb{Q}$ definiert. Im Folgenden wird dem Ausdruck a^x für $x \in \mathbb{R}$ ein Sinn gegeben. Natürlich soll diese Definition für $x \in \mathbb{Q}$ mit der alten Definition übereinstimmen. Finden wir eine stetige Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(r) = a^r$ für alle $r \in \mathbb{Q}$, dann bietet sich diese als Definition für $f(x) = a^x$ an. Das nächste Lemma zeigt, dass es höchstens eine solche stetige Funktion geben kann.

Lemma 154: Sind $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ zwei stetige Funktionen, die auf $\mathbb{Q}$ übereinstimmen (das heißt $f_1(r) = f_2(r)$ für alle $r \in \mathbb{Q}$), so sind sie gleich.

BEWEIS: Zu jedem $x \in \mathbb{R}$ gibt es nach Korollar 93 eine Folge $(r_n)_{n\in\mathbb{N}}$ in $\mathbb{Q}$ mit $\lim_{n\to\infty} r_n = x$. Daraus folgt dann wegen der Stetigkeit der beiden Funktionen

$$\begin{aligned}f_1(x) &= \lim_{n\to\infty} f_1(r_n)\\ &= \lim_{n\to\infty} f_2(r_n) = f_2(x).\end{aligned}$$

□

Satz 155: Es sei $a > 0$. Wir definieren die Funktion

$$f_a : \mathbb{R} \to \mathbb{R}$$

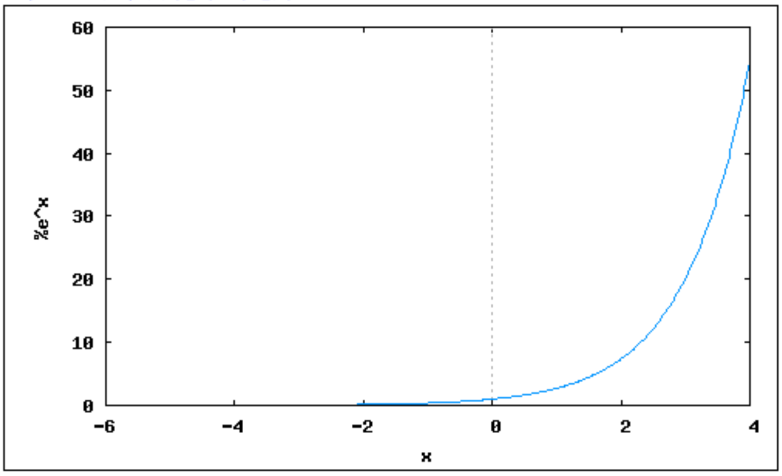
durch

$$f_a(x) := \exp(x\log(a)).$$

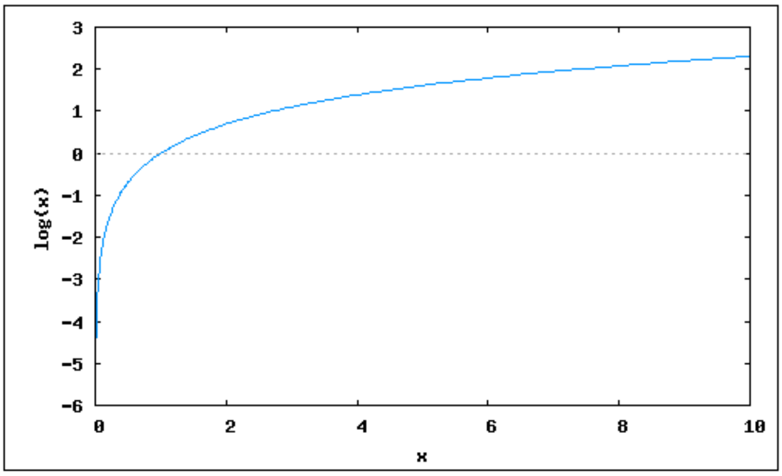
Dann gilt für alle $r \in \mathbb{Q}$ die Gleichung

$$f_a(r) = a^r.$$

```
(%i22) wxplot2d(exp(x),[x,-6,4]);
(%t22)
```




```
(%o22)
(%i23) wxplot2d(log(x),[x,0,10]);
plot2d: expression evaluates to non-numeric value somewhere in plotting range.
(%t23)
```




```
(%o23)
```

Abbildung 3.3.: Die Exponentialfunktion und die Logarithmusfunktion im Maximaplot

BEWEIS: Mit Induktion zeigt man, dass die Funktion $f_a(n)$ für $n \in \mathbb{N}$, der bekannten Potenzfunktion entspricht.

Induktionsverankerung:

$$f_a(1) = \exp(1 \cdot \log(a)) = \exp(\log(x)) = a = a^1$$

Induktionsschritt:

$$\begin{aligned} f_a(n+1) &= \exp((n+1)\log(a)) \\ &= \exp(n\log(a))\exp(\log(a)) \\ &= f_a(n)f_a(1) = a^n a^1 = a^{n+1} \end{aligned}$$

Damit ist die Aussage für $n \in \mathbb{N}$ bewiesen. Weiter gilt

$$\begin{aligned} f_a(-n) &= \exp(-n \log(a)) \\ &= \frac{1}{\exp(n \log(a))} \\ &= \frac{1}{f_a(n)} \\ &= \frac{1}{a^n} \\ &= a^{-n}. \end{aligned}$$

Also gilt $f_a(n) = a^n$ sogar für alle $n \in \mathbb{Z}$. Es sei nun $r = \frac{p}{q}$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ so gilt

$$\begin{aligned} (f_a(r))^q &= (\exp(r \log(a)))^q \\ &= \exp(qr \log(a)) \\ &= \exp(p \log(a)) \\ &= f_a(p) \\ &= a^p \end{aligned}$$

Also ist $f_a(r) = \sqrt[q]{a^p} = a^r$. □

Wir können nun die in Satz 155 definierte Funktion f_a nutzen, um die Definition von Potenzen von rationalen Exponenten auf beliebige reelle Exponenten auszuweiten.

Definition 156 (Potenz)**:** Es sei $a > 0$ eine reelle Zahl. Für beliebige $x \in \mathbb{R}$ definieren wir

$$a^x := \exp(x \log(a)).$$

BEMERKUNG: Die Eulersche Zahl e ist definiert durch $e := \exp(1)$. Insbesondere gilt also für alle reellen Zahlen x

$$e^x = \exp(x \log(e)) = \exp(x).$$

Satz 157: Für alle $x, y \in \mathbb{R}$ und $a, b \in \mathbb{R}_+$ gilt

(i) $a^{x+y} = a^x a^y$,

(ii) $a^{-x} = \frac{1}{a^x}$,

(iii) $(a^x)^y = a^{xy}$,

(iv) $(ab)^x = a^x b^x$.

Problem 61: *Beweisen Sie den vorstehenden Satz.*

Definition 158: Für positive reelle Zahlen x und a definieren wir den *Logarithmus von x zur Basis a* durch

$$\log_a x := \frac{\log x}{\log a}.$$

Problem 62: *Es sei a eine positive reelle Zahl.*

(i) Was ist $\log_a 1$?

(ii) Zeigen Sie, dass für alle $x, y \in \mathbb{R}_{>0}$

$$\log_a(xy) = \log_a x + \log_a y$$

gilt.

(iii) Zeigen Sie, dass für alle $x \in \mathbb{R}$

$$\log_a(a^x) = x$$

und für alle $x \in \mathbb{R}_{>0}$

$$a^{\log_a x} = x$$

ist.

Problem 63: *Bestimmen Sie alle reellen Lösungen der Gleichung*

$$4^x - 5 \cdot 2^x + 6 = 0.$$

3.6. Einige Grenzwerte

Im Beweis von Satz 153 haben wir gesehen, dass

$$\lim_{x\to\infty} e^x = +\infty$$

und

$$\lim_{x\to-\infty} e^x = 0$$

ist. Im folgenden Lemma schauen wir uns einige weitere interessante Grenzwerte an.

Lemma 159:

(i) Für alle $a > 0$ gilt

$$\lim_{n\to\infty} \sqrt[n]{a} = 1.$$

(ii) Für alle $k \in \mathbb{N}$ gilt

$$\lim_{x\to\infty} \frac{e^x}{x^k} = \infty \qquad \text{und} \qquad \lim_{x\to\infty} x^k e^{-x} = 0.$$

Die Exponentialfunktion wächst also schneller als jede Polynomfunktion.

(iii)

$$\lim_{x\to\infty} \log(x) = \infty \qquad \text{und} \qquad \lim_{x\searrow 0} \log(x) = -\infty$$

(iv) Für alle $\alpha > 0$ gilt

$$\lim_{x\to\infty} \frac{\log(x)}{x^\alpha} = 0 \qquad \text{und} \qquad \lim_{x\searrow 0} x^\alpha \log(x) = 0.$$

Der Logarithmus wächst langsamer als jede Polynomfunktion.

(v) Für alle $\alpha > 0$ gilt

$$\lim_{x\to\infty} x^{-\alpha} = 0 \qquad \text{und} \qquad \lim_{x\searrow 0} x^\alpha = 0.$$

Aufgrund des zweiten Grenzwertes setzt man häufig $0^\alpha := 0$.

BEMERKUNG: Im Beispiel zum Wurzelkriterium haben wir bereits

$$\lim_{n\to\infty} \sqrt[n]{n} = 1$$

bewiesen. Dies lässt sich auch aus Punkt (iv) des vorstehenden Satzes und der Stetigkeit der Exponentialfunktion folgern, denn es ist

$$\sqrt[n]{n} = n^{\frac{1}{n}} = \exp\left(\frac{\log n}{n}\right)$$

BEWEIS:

(i) Da die Exponentialfunktion stetig ist, gilt

$$\begin{aligned}\lim_{n\to\infty} \sqrt[n]{a} &= \lim_{n\to\infty} a^{\frac{1}{n}} \\ &= \lim_{n\to\infty} \exp\left(\frac{1}{n}\log(a)\right) \\ &= \exp(0) = 1\end{aligned}$$

(ii) Für beliebige $k \in \mathbb{N}$ und $x > 0$ gilt

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \geq \frac{x^{k+1}}{(k+1)!}.$$

Woraus sofort

$$\frac{e^x}{x^k} \geq \frac{x}{(k+1)} \xrightarrow[x\to\infty]{} \infty$$

folgt und somit erhalten wir die erste Behauptung. Die zweite Behauptung folgt durch Bilden des Kehrwerts.

(iii) Wegen der strengen Monotonie ergibt sich für beliebige $x, S \in \mathbb{R}_+$ mit $x > e^S$ die Ungleichung $\log(x) > S$. Also wächst der Logarithmus über jede Schranke hinaus. Betrachten wir weiter

$$\lim_{x\searrow 0} \log(x) = -\lim_{x\searrow 0} \log\left(\frac{1}{x}\right)$$

und setzen $y := \frac{1}{x}$, so ist $x \searrow 0$ gleichbedeutend mit $y \to \infty$, also gilt

$$\lim_{x\searrow 0} \log(x) = -\lim_{y\to\infty} \log(y) = -\infty.$$

(iv) Wegen (iii) ist $x \to \infty$ gleichbedeutend mit $y \to \infty$ für $y = \alpha \log(x)$. Nun folgt weiter

$$\lim_{x\to\infty} \frac{\log(x)}{x^\alpha} = \lim_{x\to\infty} \frac{\log(x)}{\exp(\alpha\log(x))} = \lim_{y\to\infty} \frac{y}{\alpha}\frac{1}{\exp(y)} \overset{\text{(ii)}}{=} 0.$$

Aus

$$x^\alpha \log(x) = \frac{-\log\left(\frac{1}{x}\right)}{\left(\frac{1}{x}\right)^\alpha}$$

ergibt sich die zweite Behautpung.

(v) Diese Behauptung folgt aus Aussage (iv), da

$$0 \leq x^{-\alpha} \leq \frac{\log(x)}{x^\alpha}$$

für alle $x \geq e$ und

$$0 \leq x^\alpha \leq x^\alpha |\log(x)|$$

für alle $x \leq \frac{1}{e}$ gilt.

$\square$

Der folgende Satz liefert eine alternative Darstellung für die Exponentialfunktion und zeigt, dass man die Exponentialfunktion nicht nur mittels einer Reihe einführen kann.

Satz 160: Es gilt die Identität

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

BEWEIS: Setze $y := e^h - 1$ was gleichbedeutend zu $h := \log(y+1)$ ist. Nun gilt wegen Lemma 129

$$\lim_{y\to 0}\frac{y}{\log(1+y)} = \lim_{h\to 0}\frac{e^h-1}{h} = 1.$$

Für festes $x \in \mathbb{R}, x \neq 0,$ erhalten wir also

$$\lim_{y\to 0}\frac{xy}{\log(1+xy)} = 1$$

und somit auch $\lim_{y\to 0}\frac{\log(1+xy)}{xy} = 1$. Wegen $a\log(b) = \log(b^a)$ erhalten wir

$$\frac{\log(1+xy)}{y} = \log\left((1+xy)^{\frac{1}{y}}\right).$$

Hieraus ergibt sich

$$1 = \lim_{y\to 0}\frac{\log(1+xy)}{xy} = \frac{1}{x}\lim_{y\to 0}\log\left((1+xy)^{\frac{1}{y}}\right)$$

beziehungsweise

$$x = \lim_{y\to 0}\log\left((1+xy)^{\frac{1}{y}}\right).$$

Aufgrund der Stetigkeit der Exponentialfunktion erhalten wir weiter

$$\begin{aligned}\exp(x) &= \exp\left(\lim_{y\to 0}\log\left((1+xy)^{\frac{1}{y}}\right)\right)\\ &= \lim_{y\to 0}\exp\left(\log\left((1+xy)^{\frac{1}{y}}\right)\right)\\ &= \lim_{y\to 0}(1+xy)^{\frac{1}{y}}.\end{aligned}$$

Wählen wir für y nun die Nullfolge $\frac{1}{n}$, so ergibt sich die Behauptung

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

□

BEMERKUNG: Man kann im letzten Schritt des Beweises anstelle der Folge $(1/n)$ natürlich auch jede andere Nullfolge nehmen. Es gilt also etwa auch

$$\exp(x) = \lim_{n\to\infty}\left(1-\frac{x}{n^2}\right)^{-n^2}.$$

Der vorstehende Satz gibt uns also eine zweite Möglichkeit Werte der Exponentialfunktion näherungsweise zu bestimmen. Es zeigt sich aber, dass die Partialsummen der Exponentialreihe viel schneller konvergieren als die oben betrachte Folge.

```
clear all;
n=1:100;
x=1;
f=(1+x./n).^n;
for i=n
g(i)=sum(x.^(0:i)./factorial(0:i));
end
plot(n,f,'r.',n,g,'g.');
```

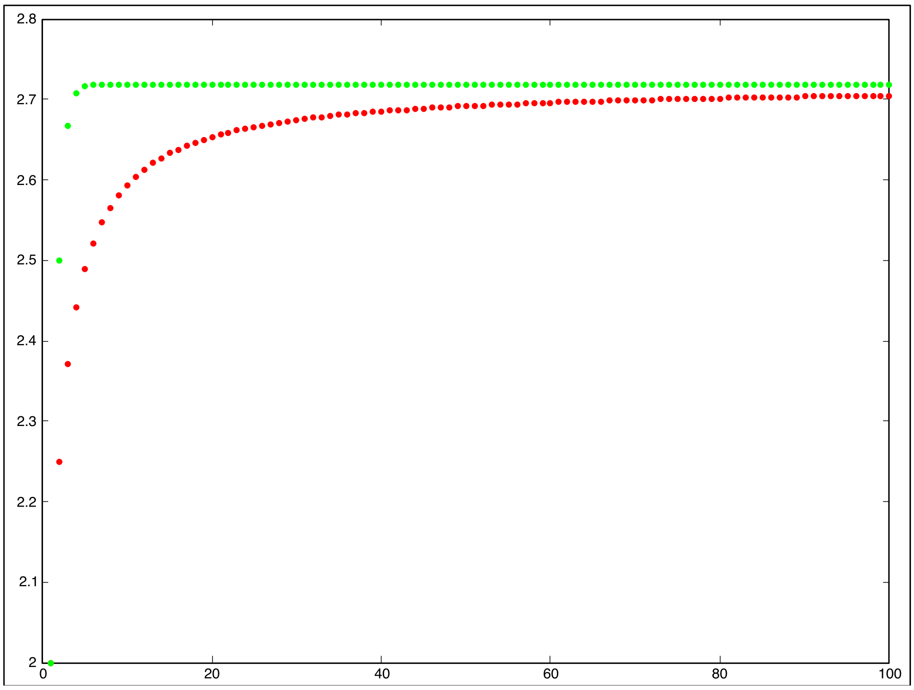


Abbildung 3.4.: Approximation der Zahl e mittels Partialsummen der Exponentialreihe (in grün) und mittels der Folge $(1+1/n)^n$ (in rot)

3.7. Asymptoten

Wenn sich der Graph einer Funktion f für $x \to \infty$ oder $x \to -\infty$ einer festen Gerade immer mehr annähert, so nennt man die Gerade eine Asymptote der Funktion f. Wir wollen dies nun genauer definieren.

Definition 161 ((schräge und horizontale) Asymptote)**:** Es seien $m, a \in \mathbb{R}$ und $f : D \to \mathbb{R}$. Wir sagen, dass die Gerade $y = mx + a$ eine *Asymptote* der Funktion f ist, wenn einer der folgenden Fälle vorliegt:

(i) D enthält ein nach oben unbeschränktes Intervall und es ist

$$\lim_{x\to\infty} (f(x) - mx - a) = 0,$$

(ii) D enthält ein nach unten unbeschränktes Intervall und es ist

$$\lim_{x\to-\infty} (f(x) - mx - a) = 0.$$

Ist $m = 0$, so sagen wir, dass a eine *horizontale Asymptote von* f ist. Ist $m \neq 0$ so spricht man von einer *schrägen Asymptote*.

Beispiele:

1) Es ist $\lim_{x\to\infty} \frac{x^2+3x+2}{2x^2-4} = \frac{1}{2}$ und somit

$$\lim_{x\to\infty} \left(\frac{x^2+3x+2}{2x^2-4} - \frac{1}{2} \right) = 0.$$

Also ist $y = \frac{1}{2}$ eine horizontale Asymptote der Funktion $\frac{x^2+3x+2}{2x^2-4}$.

2) Wir wollen die Asymptoten der Funktion $f(x) = \frac{x}{|x|}$ bestimmen. Es ist

$$\lim_{x\to\infty} \frac{x}{|x|} = 1$$

und

$$\lim_{x\to-\infty} \frac{x}{|x|} = -1.$$

f hat also die beiden horizontalen Asymptoten $y = 1$ und $y = -1$.

3) Wir wollen die Asymptoten der Funktion

$$f(x) = \frac{x^2+3x+3}{x+1}$$

bestimmen. Mittels Polynomdivision erhält man, dass $\frac{x^2+3x+3}{x+1} = x + 2 + \frac{1}{x+1}$ ist. Also ist

$$\lim_{x\to\infty} \left(\frac{x^2+3x+3}{x+1} - x - 2 \right) = \lim_{x\to\infty} \frac{1}{x+1} = 0$$

und ebenso

$$\lim_{x\to-\infty} \left(\frac{x^2+3x+3}{x+1} - x - 2 \right) = \lim_{x\to-\infty} \frac{1}{x+1} = 0.$$

$y = x + 2$ ist also sowohl für $x \to \infty$ also auch für $x \to -\infty$ eine schräge Asymptote der Funktion f.

Es gibt noch einen weiteren Typ von Asymptoten.

Definition 162 (vertikale Asymptote)**:** Es sei $f : D \to \mathbb{R}$ eine Funktion und $a \notin D$ sei ein Berührpunkt von D. Dann heißt die Gerade $x = a$ eine vertikale Asymptote von f, wenn mindestens eine der folgenden Bedingungen zutrifft:

$$\lim_{x \nearrow a} f(x) = +\infty, \lim_{x \nearrow a} f(x) = -\infty, \lim_{x \searrow a} f(x) = +\infty, \lim_{x \searrow a} f(x) = -\infty.$$

Beispiele:

1) Da

$$\lim_{x \searrow 0} \frac{1}{x} = +\infty$$

ist, ist 0 eine vertikale Asymptote der Funktion $\frac{1}{x}$. Es ist auch

$$\lim_{x \nearrow 0} \frac{1}{x} = -\infty.$$

2) Wir wollen die vertikalen Asymptoten von

$$\frac{2x+4}{x^2+3x+2}$$

bestimmen. Für alle x, für die der Nenner ungleich Null ist, ist die Funktion stetig. Dort können keine vertikalen Asymptoten vorliegen. Wir müssen also schauen, wo der Nenner gleich Null ist. Es ist $x^2 + 3x + 2 = (x+1)(x+2)$. Also kommen nur $x = -1$ und $x = -2$ in Betracht. Es ist

$$\lim_{x \nearrow -1} \frac{2x+4}{x^2+3x+2} = \lim_{x \nearrow -1} \frac{2(x+2)}{(x+1)(x+2)} = \lim_{x \nearrow -1} \frac{2}{(x+1)} = -\infty.$$

Also ist $x = -1$ eine vertikale Asymptote. Weiter ist

$$\lim_{x \nearrow -2} \frac{2x+4}{x^2+3x+2} = \lim_{x \nearrow -2} \frac{2}{(x+1)} = -2 = \lim_{x \searrow -2} \frac{2}{(x+1)} = \lim_{x \searrow -2} \frac{2x+4}{x^2+3x+2}.$$

Die Funktion lässt sich also stetig in $x = -2$ fortsetzen und es liegt dort keine vertikale Asymptote vor.

3) Die Funktion

$$f(x) = \frac{x^2+3x+3}{x+1}$$

hat offensichtlich die vertikale Asymptote $x = -1$, da -1 keine Nullstelle des Zählers ist.

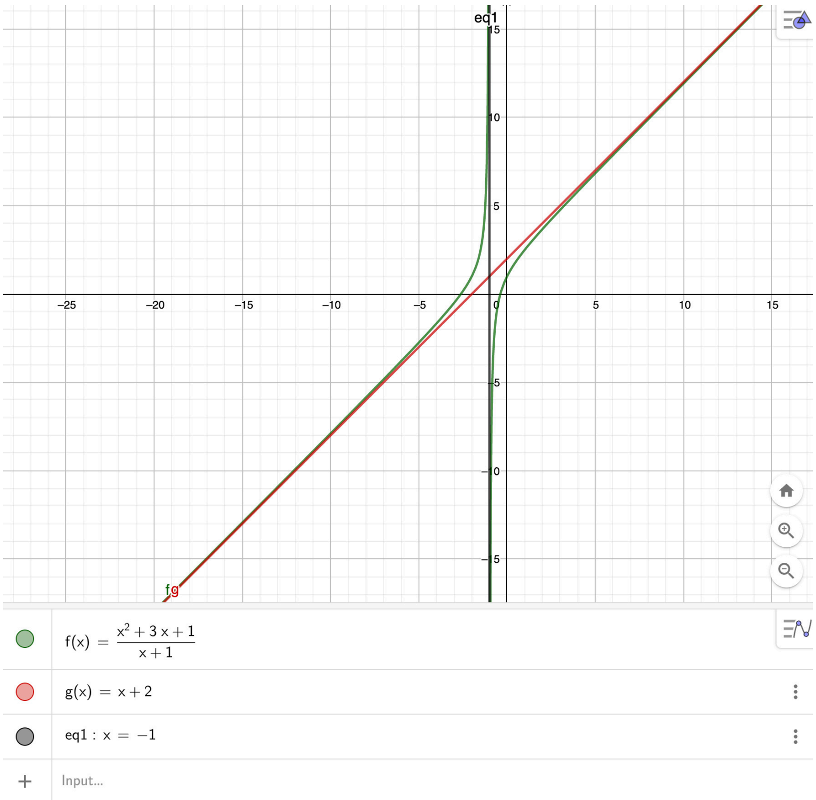


Abbildung 3.5.: GeoGebra-Plot der Funktion $f(x) = \frac{x^2+3x+3}{x+1}$ mit schräger Asymptote $y = x+2$ und vertikaler Asymptote $x = -1$.

4. Differential- und Integralrechnung

4.1. Differentialrechnung reellwertiger Funktionen einer Veränderlichen

Generalvoraussetzung: In diesem Kapitel ist $I \subset \mathbb{R}$ stets ein Intervall, dass mehr als einen Punkt enthält.

Definition 163 (Differenzierbarkeit)**:** Eine Funktion $f : I \to \mathbb{R}$ heißt *differenzierbar in* $x_0 \in I$, falls ein $a \in \mathbb{R}$ und eine Funktion $r_{x_0} : I \to \mathbb{R}$ mit

$$\lim_{x \to x_0} \frac{r_{x_0}(x)}{x - x_0} = 0$$

existieren, so dass für alle $x \in I$

$$f(x) = f(x_0) + a(x - x_0) + r_{x_0}(x)$$

gilt. Der Wert a heißt die *Ableitung von f an der Stelle* x_0, welchen wir mit $f'(x_0)$ bezeichnen wollen.
Die Funktion f heißt *differenzierbar* (auf ganz I), falls f in allen Punkten $x \in I$ differenzierbar ist. Die Funktion $f' : I \to \mathbb{R}$ mit $x \mapsto f'(x)$ heißt dann die *Ableitung von* f.

BEMERKUNG: Wegen

$$r_{x_0}(x) = \frac{r_{x_0}(x)}{x - x_0}(x - x_0)$$

sieht man sofort, dass $\lim_{x \to x_0} r_{x_0}(x) = 0$. Die Bedingung

$$\lim_{x \to x_0} \frac{r_{x_0}(x)}{x - x_0} = 0$$

sagt aber sogar aus, dass $r_{x_0}(x)$ für x gegen x_0 schneller gegen 0 geht als $x - x_0$.

BEMERKUNG: Statt $f'(x)$ schreibt man auch $\frac{df}{dx}(x)$ oder $\frac{d}{dx}(f(x))$. In der Physik wird auch gerne die Schreibweise $\dot{v}$ statt v' benutzt.

Geometrische Interpretation: Ist x nahe bei x_0, so kann $r(x)$ vernachlässigt werden und f wird annähernd durch die Gerade

$$f(x_0) + a \cdot (x - x_0)$$

beschrieben. Die Ableitung $f'(x_0)$ ist die Steigung dieser Geraden. Wir nennen diese Gerade auch die Tangente an die Funktion f im Punkt x_0.

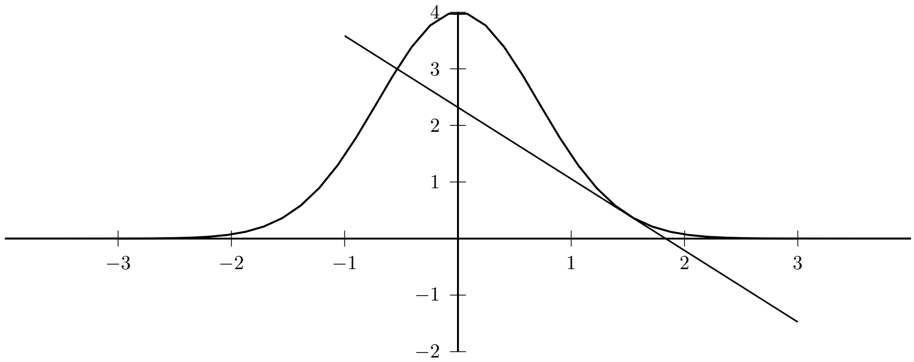


Abbildung 4.1.: Eine Funktion mit eingezeichneter Tangente in einem Punkt.

Beispiele:

1) Es sei f eine konstante Funktion, also $f(x) = c$ für alle $x \in \mathbb{R}$. Da gilt für alle $x, x_0 \in \mathbb{R}$
$$f(x) = f(x_0) + 0(x - x_0).$$
Also ist $f'(x_0) = 0$ für alle $x_0 \in \mathbb{R}$.

2) Es sei $f(x) = x^n$ mit $n \in \mathbb{N}$ und $x_0 \in \mathbb{R}$ beliebig. Es ist
$$x^n - x_0^n = (x - x_0) \sum_{k=0}^{n-1} x^k x_0^{n-k-1}$$
und
$$\lim_{x \to x_0} \sum_{k=0}^{n-1} x^k x_0^{n-k-1} = n x_0^{n-1}.$$
Setzt man
$$r(x) := r_{x_0}(x) := (x - x_0) \sum_{k=0}^{n-1} x^k x_0^{n-k-1} - n x_0^{n-1}(x - x_0),$$
so gilt $\lim_{x \to x_0} \frac{r(x)}{x - x_0} = 0$ und $x^n = x_0^n + n x_0^{n-1}(x - x_0) + r(x)$ für alle $x \in \mathbb{R}$. Also ist $f'(x_0) = n x_0^{n-1}$.

Satz 164: Jede in einem Punkt x_0 differenzierbare Funktion ist in x_0 auch stetig.

Problem 64: *Beweisen Sie den vorstehenden Satz.*

BEMERKUNG: Die Umkehrung gilt jedoch nicht. Die Funktion $|\cdot| : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$ ist in 0 stetig, aber nicht differenzierbar. Dies sieht man leicht mit folgendem Satz.

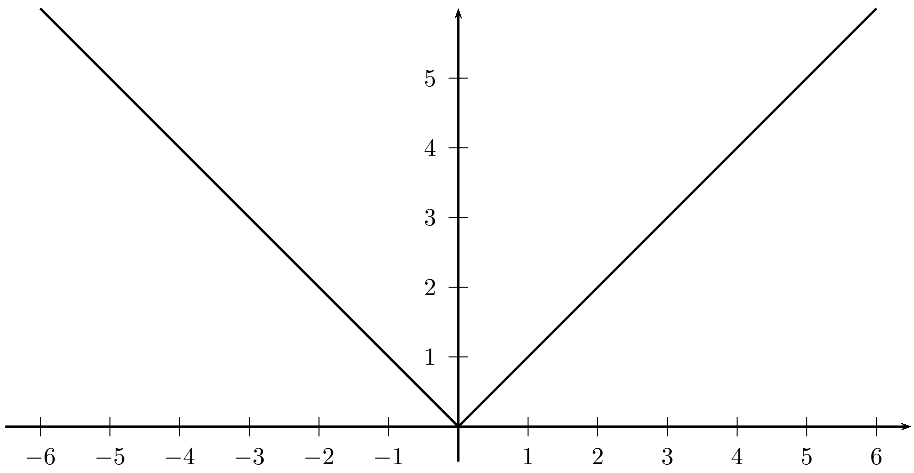


Abbildung 4.2.: Betragsfunktion

Satz 165: Es sei $f : I \to \mathbb{R}$ eine Funktion, dann sind folgende Aussagen äquivalent:

(i) Die Funktion f ist differenzierbar in $x_0 \in I$ mit $f'(x_0) = a$.

(ii) Der Grenzwert

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert und ist gleich a.

(iii) Der Grenzwert

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

existiert und ist gleich a.

Problem 65: *Beweisen Sie den vorstehenden Satz.*

BEMERKUNG: Da der Differenzenquotient

$$\frac{f(x) - f(x_0)}{x - x_0}$$

nur für $x \in I \setminus \{x_0\}$ erklärt ist, ist bei den Grenzwerten im vorstehenden Satz $x \neq x_0$ bzw. $h \neq 0$ vorauszusetzen.

Beispiel (Ableitung der Exponentialfunktion)**:** Es sei $x_0 \in \mathbb{R}$ beliebig. Aufgrund des Additionstheorems der Exponentialfunktion gilt

$$\exp(x_0 + h) - \exp(x_0) = \exp(x_0)(\exp(h) - 1).$$

Aus Lemma 129 kennen wir die Aussage $\lim_{h\to 0} \frac{e^h - 1}{h} = 1$. Hiermit erhalten wir

$$\lim_{h\to 0} \frac{e^{x_0+h} - e^{x_0}}{h} = e^{x_0}.$$

Es gilt also $\exp(x_0) = \exp'(x_0)$. Dies nennt man auch die *Differentialgleichung der Exponentialfunktion*.

Beispiel (Ableitung der trigonometrischen Funktionen)**:** Es sei $x_0 \in \mathbb{R}$ beliebig. Aufgrund der Additionstheoreme von Sinus und Cosinus gilt

$$\sin(x_0 + h) - \sin(x_0) = \sin(x_0)(\cos(h) - 1) + \cos(x_0)\sin(h)$$

und

$$\cos(x_0 + h) - \cos(x_0) = \cos(x_0)(\cos(h) - 1) - \sin(x_0)\sin(h)$$

Aus Lemma 144 kennen wir die Aussagen $\lim_{h\to 0} \frac{\cos(h)-1}{h} = 0$ und $\lim_{h\to 0} \frac{\sin h}{h} = 1$. Hiermit erhalten wir

$$\sin'(x_0) = \cos(x_0) \text{ und } \cos'(x_0) = -\sin(x_0).$$

Problem 66: *Die Funktion $f : \mathbb{R} \to \mathbb{R}$ sei definiert durch*

$$f(x) := \begin{cases} 0 & , \text{ falls } x \leq 0, \\ x^2 & , \text{ falls } x > 0, \end{cases}.$$

Zeigen Sie, dass f auf ganz $\mathbb{R}$ differenzierbar ist. Zeigen Sie, dass die Ableitung f' nicht auf ganz $\mathbb{R}$ differenzierbar ist.

Problem 67: *Es sei $g : [-1, 1] \to \mathbb{R}$ eine Funktion. Beweisen Sie die folgenden Aussagen und bestimme $f'(0)$ bzw. $h'(0)$.*

(i) Ist g stetig in 0, so ist $f(x) := x \cdot g(x)$ differenzierbar in 0.

(ii) Ist g beschränkt, so ist $h(x) := x^2 \cdot g(x)$ differenzierbar in 0.

Exkurs: Numerische Berechnung der Ableitung

Die im letzten Satz gegebene alternative Charakterisierung der Differenzierbarkeit ermöglicht auch eine näherungsweise Berechnung der Ableitung einer Funktion f an einer vorgegebenen Stelle mit Hilfe des Computers.

```
function erg = ableitung(f,a,h)
%
% ableitung(f,a,b) - berechnet die Ableitung von f an der Stelle a mittels Differenzenquotient
% f kann hier bei eine Matlab-Funktion sein, die in der Form
% @Funktionsname übergeben werden muss, oder
% f ist eine selbstdefinierte "anonyme" Funktion.

if nargin == 2
    h=10^(-5);
end
erg=(f(a+h)-f(a))/h;
```

Numerisch stabiler ist die folgende Implementierung. Dabei wird ausgenutzt, dass auch

$$f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0-h)}{2h}$$

ist, wenn f differenzierbar ist.

```
function erg = ableitung2(f,a,h)
  %
  % ableitung2(f,a,b) - berechnet die Ableitung von f an der Stelle a mittels Differenzenquotient
  % f kann hier bei eine Matlab-Funktion sein, die in der Form
  % @Funktionsname übergeben werden muss, oder
  % f ist eine selbstdefinierte "anonyme" Funktion.

if nargin == 2
    h=10^(-5);
end
erg=(f(a+h)-f(a-h))/(2*h);
```

Zum Abschluss geben wir noch ein kleines Anwendungsbeispiel für diese Funktionen.

```
clear all;
x=0:0.1:7;
figure(1)
clf()
plot(x,sin(x),x,ableitung(@sin,x))
legend('sin(x)','Ableitung der Sinus-Funktion')
xlabel('x')
```

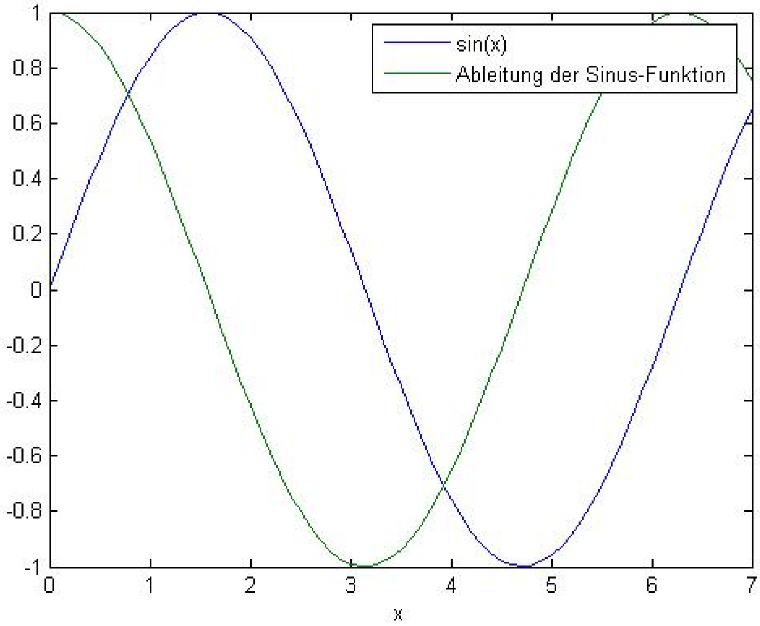


Abbildung 4.3.: Ausgabe des Beispielskripts.

Ende des Exkurses

Satz 166 (Ableitungsregeln)**:** Es seien $f, g : I \to \mathbb{R}$ in $x_0 \in I$ differenzierbar und $\lambda \in \mathbb{R}$ beliebig. Dann sind die Funktionen

$$f + g,\ \lambda \cdot f \text{ und } f \cdot g$$

in x_0 differenzierbar mit

(i) $(f + g)'(x_0) = f'(x_0) + g'(x_0)$, (Linearität)

(ii) $(\lambda \cdot f)'(x_0) = \lambda \cdot f'(x_0)$, (Linearität)

(iii) $(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0)$. (Produktregel)

(iv) Ist $g(x_0) \neq 0$, so ist auch $\frac{f}{g}$ in x_0 differenzierbar mit

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}.$$ (Quotientenregel)

BEWEIS:

(i) Es existieren r_1 und r_2 wie in Definition 163, so dass

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r_1(x)$$
$$g(x) = g(x_0) + g'(x_0)(x - x_0) + r_2(x)$$

gilt. Somit gilt

$$(f+g)(x) = f(x_0) + g(x_0) + (f'(x_0) + g'(x_0))(x - x_0) + r_1(x) + r_2(x)$$

und

$$\lim_{x \to x_0} \frac{r_1(x) + r_2(x)}{x - x_0} = 0.$$

(ii) Folgt direkt aus (iii), da $(\lambda)' = 0$ ist.

(iii) Um die Produktregel herzuleiten, setzen wir $h := x - x_0$. Haben wir wie bei (i)

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r_1(x)$$
$$g(x) = g(x_0) + g'(x_0)(x - x_0) + r_2(x),$$

so ergibt sich

$$\begin{aligned} f(x) \cdot g(x) &= (f(x_0) + f'(x_0)h + r_1(x_0 + h))\,(g(x_0) + g'(x_0)h + r_2(x_0 + h)) \\ &= f(x_0)g(x_0) + (f'(x_0)g(x_0) + f(x_0)g'(x_0)) \cdot h \\ &\quad + r_1(x_0 + h)r_2(x_0 + h) \\ &\quad + r_1(x_0 + h)(g(x_0) + g'(x_0)h) \\ &\quad + r_2(x_0 + h)(f(x_0) + f'(x_0)h) \end{aligned}$$

Setzen wir $r(h) := r_1(x_0+h)r_2(x_0+h) + r_1(x_0+h)(g(x_0)+g'(x_0)h) + r_2(x_0+h)(f(x_0)+f'(x_0)h)$, so gilt wegen $\frac{r_1(x_0+h)}{h} \to 0$ und $\frac{r_2(x_0+h)}{h} \to 0$ auch $\frac{r(h)}{h} \to 0$ für $h \to 0$. Damit ist die Produktregel bewiesen und die zweite Linearitätsbehauptung folgt daraus.

(iv) Zunächst wird ein Spezialfall der Quotientenregel bewiesen.

$$\begin{aligned} \left(\frac{1}{g}\right)'(x_0) &= \lim_{h \to 0} \frac{1}{h} \left(\frac{1}{g}(x_0 + h) - \frac{1}{g}(x_0) \right) \\ &= \lim_{h \to 0} \frac{1}{h} \left(\frac{g(x_0) - g(x_0 + h)}{g(x_0 + h)g(x_0)} \right) \\ &= \lim_{h \to 0} \frac{1}{g(x_0 + h)g(x_0)} \cdot \lim_{h \to 0} -\frac{g(x_0 + h) - g(x_0)}{h} \\ &= \frac{1}{(g(x_0))^2} \cdot (-g'(x_0)) \end{aligned}$$

Auf Grund der Produktregel folgt nun die Quotientenregel

$$\begin{aligned}\left(\frac{f}{g}\right)'(x_0) &= \left(f \cdot \frac{1}{g}\right)'(x_0) \\ &= f'(x_0)\frac{1}{g(x_0)} + f(x_0)\left(\frac{1}{g}\right)'(x_0) \\ &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}.\end{aligned}$$

□

Satz 167 (Kettenregel)**:** Die Hintereinanderausführung zweier differenzierbarer Funktionen $f : I \to J$ und $g : J \to \mathbb{R}$ ist differenzierbar und es gilt

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

BEWEIS: Es gelte

$$f(x_0 + h) = f(x_0) + f'(x_0)h + r_1(h)$$

mit $\frac{r_1(h)}{h} \xrightarrow[h\to\infty]{} 0$ und für $y_0 := f(x_0)$ analog

$$g(y_0 + \tilde{h}) = g(y_0) + g'(y_0)\tilde{h} + r_2(\tilde{h})$$

mit $\frac{r_2(h)}{h} \xrightarrow[h\to\infty]{} 0$. Nun erhält man

$$\begin{aligned}(g \circ f)(x_0 + h) &= g(f(x_0 + h)) \\ &= g(\underbrace{f(x_0)}_{=y_0} + \underbrace{f'(x_0)h + r_1(h)}_{=:\tilde{h}(h)}) \\ &= g(y_0) + g'(y_0)\tilde{h} + r_2(\tilde{h}(h)).\end{aligned}$$

Außerdem es gilt $\lim_{h\to 0} \tilde{h}(h) = 0$ und $\lim_{h\to 0} \frac{\tilde{h}(h)}{h} = f'(x_0)$. Daraus folgt

$$\begin{aligned}\lim_{h\to 0} \frac{r_2(\tilde{h}(h))}{h} &= \lim_{h\to 0} \frac{r_2(\tilde{h}(h))\tilde{h}(h)}{h\tilde{h}(h)} \\ &= \lim_{h\to 0} \frac{\tilde{h}(h)}{h} \cdot \lim_{h\to 0} \frac{r_2(\tilde{h}(h))}{\tilde{h}(h)} \\ &= f'(x_0) \cdot 0 = 0.\end{aligned}$$

Zusammen ergibt dies

$$(g \circ f)(x_0 + h) = g(f(x_0)) + g'(f(x_0)) \cdot f'(x_0) \cdot h + r(h)$$

mit $r(h) = g'(f(x_0))r_1(h) + r_2(\tilde{h}(h))$ und $\frac{r(h)}{h} \xrightarrow[h\to 0]{} 0.$ □

Problem 68: *Bestimmen Sie für die folgenden Funktionen jeweils die Ableitung.*

(i)

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto e^x(x^2 + 1),$$

(ii)

$$g : \mathbb{R} \to \mathbb{R}, x \mapsto \frac{e^x}{x^2 + 1},$$

(iii)

$$h : \mathbb{R} \to \mathbb{R}, x \mapsto \sin(x^2 + 1).$$

Satz 168 (Ableitung der Umkehrfunktion)**:** Es sei $f : I \to J$ eine bijektive Funktion. Ist f außerdem in $y_0 \in I$ differenzierbar mit $f'(y_0) \neq 0$, so ist $f^{-1} : J \to I$ in $x_0 := f(y_0)$ differenzierbar mit

$$\left(f^{-1}\right)'(x_0) = \frac{1}{f'(y_0)} = \frac{1}{f'(f^{-1}(x_0))}.$$

BEMERKUNG: Die Voraussetzung $f'(y_0) \neq 0$ ist wichtig. Betrachten wir etwa die Funktion $f : [0, 1] \to [0, 1], x \mapsto x^2$. f ist differenzierbar und bildet das Intervall $[0, 1]$ bijektiv auf $[0, 1]$ ab. Die Umkehrabbildung f^{-1} existiert also und bildet ebenfalls das Intervall $[0, 1]$ bijektiv auf $[0, 1]$ ab. Wir kennen die Umkehrfunktion schon, es ist nämlich die Wurzelfunktion und diese ist in 0 nicht differenzierbar, wie man sieht, wenn man den Differenzenquotienten betrachtet:

$$\frac{f^{-1}(x) - f^{-1}(0)}{x - 0} = \frac{\sqrt{x} - 0}{x - 0} = \frac{\sqrt{x}}{\sqrt{x}\sqrt{x}} = \frac{1}{\sqrt{x}}.$$

Der Grenzwert für x gegen 0 existiert offensichtlich nicht, da der Ausdruck unbeschränkt ist. Da $f'(x) = 2x \neq 0$ für alle $x \neq 0$ ist, sagt der Satz, dass die Wurzelfunktion für alle $x > 0$ differenzierbar ist und dass

$$(\sqrt{x})' = \left(f^{-1}\right)'(x) = \frac{1}{f'(f^{-1}(x))} = \frac{1}{2f^{-1}(x)} = \frac{1}{2\sqrt{x}}.$$

BEWEIS: Da wegen Satz 151 f^{-1} in einer Umgebung von y_0 stetig ist, gilt $\lim_{x \to x_0} f^{-1}(x) = f^{-1}(x_0)$. Setze $y := f^{-1}(x)$ so gilt $y \to y_0$ für $x \to x_0$. Dann gilt

$$\lim_{x \to x_0} \frac{f^{-1}(x) - f^{-1}(x_0)}{x - x_0} = \lim_{y \to y_0} \frac{y - y_0}{f(y) - f(y_0)} = \frac{1}{f'(y_0)} = \frac{1}{f'(f^{-1}(x_0))}.$$

□

Problem 69: *Warum ist der folgende „Beweis“ des Satzes nicht vollständig?*
Für alle $x_0 \in J$ gilt $x_0 = f(f^{-1}(x_0))$. Leitet man diese Gleichung auf beiden Seiten ab, so erhält man nach Kettenregel $1 = f'(f^{-1}(x_0))(f^{-1}(x_0))'$, woraus die Behauptung folgt.

Beispiel: Betrachte die Exponentialfunktion $\exp$ mit der Umkehrfunktion $\exp^{-1}(x) = \log(x)$ für $x > 0$. Es gilt also

$$\log'(x) = \frac{1}{\exp'(\log(x))} = \frac{1}{\exp(\log(x))} = \frac{1}{x}.$$

Beispiel: Nun berechnen wir die Ableitung der Wurzelfunktion noch auf einem anderen Weg, nämlich mit Hilfe der Exponential- und der Logarithmusfunktion. Es sei

$$f(x) = \sqrt{x} = x^{\frac{1}{2}} = \exp\left(\frac{1}{2}\log(x)\right).$$

Dann gilt nach Kettenregel

$$f'(x) = \left(\frac{1}{2}\log(x)\right)' \exp\left(\frac{1}{2}\log(x)\right) = \frac{1}{2x}\exp\left(\frac{1}{2}\log(x)\right) = \frac{x^{\frac{1}{2}}}{2x} = \frac{1}{2}x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}}.$$

Ebenso zeigt man für beliebiges $\alpha \in \mathbb{R}$, dass für $x > 0$

$$(x^\alpha)' = \alpha x^{\alpha-1}$$

ist.

$f(x)$	$f'(x)$	Definitionsbereich
$x^n, n \in \mathbb{N}$	nx^{n-1}	$x \in \mathbb{R}$
$\exp(x)$	$\exp(x)$	$x \in \mathbb{R}$
$\log(x)$	$\frac{1}{x}$	$x > 0$
$x^\alpha, \alpha \in \mathbb{R}$	$\alpha x^{\alpha-1}$	$x > 0$
$\sin(x)$	$\cos(x)$	$x \in \mathbb{R}$
$\cos(x)$	$-\sin(x)$	$x \in \mathbb{R}$

Abbildung 4.4.: Einige wichtige Ableitungen

Definition 169 (Raum der stetig differenzierbaren Funktion)**:** Es sei $f : I \to \mathbb{R}$ eine Funktion. Wir setzen $f^{(0)} := f$. Ist $f^{(n-1)} : I \to \mathbb{R}$ differenzierbar, so heißt

$$\begin{aligned} f^{(n)} &: I \to \mathbb{R} \\ x &\mapsto \left(f^{(n-1)}\right)'(x) \end{aligned}$$

die *n-te Ableitung von f*. Die Funktion f heißt dann *n-mal differenzierbar*. Ist die n-te Ableitung stetig, so heißt f n-mal stetig differenzierbar. Mit

$$C^n(I) := \{f : I \to \mathbb{R} \mid f\ n\text{-mal stetig differenzierbar}\}$$

bezeichnet man die Menge der n-mal stetig differenzierbaren Funktionen. Setze außerdem

$$C^\infty(I) := \{f : I \to \mathbb{R} \mid f \text{ ist beliebig oft stetig differenzierbar}\} = \bigcap_{n\in\mathbb{N}} C^n(I).$$

Die gerade definierten Mengen sind Vektorräume.

4.2. Implizite Differentiation und logarithmische Ableitung

Es kann vorkommen, dass eine Funktion nicht explizit gegeben ist, sondern nur implizit durch eine Gleichung gegeben ist. Aber auch in solchen Fällen kann man häufig die Ableitung der Funktion berechnen. Wir wollen uns dies in diesem Abschnitt an einigen Beispielen anschauen.

Wir betrachten zunächst alle Punkte $(x, y) \in \mathbb{R}^2$ die der Gleichung

$$x^2 + y^2 = 9$$

genügen. Die Gleichung beschreibt bekanntermaßen einen Kreis um den Nullpunkt mit Radius 3. Welche Steigung hat die Tangente an diesen Kreis im Punkt $(2, -\sqrt{5})$?

Ein Lösungsweg sieht wie folgt aus: Wir betrachten y als Funktion, die von x abhängt, also $y = y(x)$. Da einem x eventuell zwei y-Werte auf dem Kreis zugeordnet sind, müssen wir zwischen dem oberen und dem unteren Halbkreis unterscheiden. Der Punkt $(2, -\sqrt{5})$ liegt auf dem unteren Halbkreis, also ist

$$y(x) = -\sqrt{9 - x^2}.$$

Diese Funktion könnten wir nun ableiten und so die Antwort auf die Frage finden. Wir wollen das nicht weiter ausführen.

Wir können die Ableitung von $y(x)$ an der Stelle $x = 2$ aber auch berechnen, ohne die Gleichung explizit nach $y(x)$ aufzulösen - was bei vielen Problemen auch gar nicht möglich ist. Wir leiten dazu die obige Gleichung auf beiden Seiten ab, wir betrachten also

$$\frac{d}{dx}(x^2 + y(x)^2) = \frac{d}{dx}9.$$

Wenn wir auf der linken Seite die Kettenregel beachten, erhalten wir

$$2x + 2y(x)y'(x) = 0$$

beziehungsweise

$$y'(x) = -\frac{2x}{2y(x)} = -\frac{x}{y}.$$

Die Steigung der Tangenten an den Kreis im Punkt $(2, -\sqrt{5})$ beträgt also

$$y'(2) = -\frac{2}{-\sqrt{5}} = \frac{2}{\sqrt{5}}.$$

Man spricht hier vom *impliziten ableiten* oder *impliziten differenzieren.*

Schauen wir uns ein anderes Beispiel an. Unsere Funktion $y = y(x)$ sei implizit gegeben durch die Gleichung

$$yx^2 + e^y = x^3.$$

Wie können wir die Ableitung von $y(x)$ berechnen. Wieder leiten wir die Gleichung auf beiden Seiten nach x ab und erhalten

$$\frac{d}{dx}(yx^2 + e^y) = \frac{d}{dx}x^3.$$

Dies ergibt

$$y'x^2 + 2yx + e^y y' = 3x^2$$

beziehungsweise

$$y'(x) = \frac{3x^2 - 2yx}{x^2 + e^y}.$$

Als letztes Beispiel wollen wir die Ableitung der Funktion $f(x) = x^x$ berechnen. Wir können dies natürlich tun, indem wir $f(x) = \exp(x \log x)$ schreiben und nun die Kettenregel und die Produktregel anwenden. Etwas einfacher geht es, wenn wir die Gleichung $f(x) = x^x$ zunächst logarithmieren und dann implizit ableiten. Wir haben

$$\log f(x) = x \log x.$$

Wenn wir beide Seiten nach x ableiten, erhalten wir

$$\frac{f'(x)}{f(x)} = \log x + x \cdot \frac{1}{x} = 1 + \log x.$$

Man sagt hierzu auch, dass man die *logarithmische Ableitung von f* bildet. Wenn wir nun mit $f(x)$ multiplizieren und später die Definition von f einsetzen, erhalten wir

$$f'(x) = f(x)(1 + \log x) = x^x(1 + \log x).$$

Exkurs: Differentialrechnung mit Computeralgebrasystemen

Beim Berechnen von Ableitungen sind Computeralgebrasysteme wie Mathematica, Maple oder das von uns benutzte Maxima ein sehr leistungsstarkes Hilfsmittel.

Selbst komplizierteste verschachtelte Funktionen lassen sich in sekundenschnelle ableiten. Das folgende Beispiel zeigt, wie man die Ableitung der Funktion $x^2 \exp(x^2)$ mit Maxima berechnen kann.

```
(%i7) diff(exp(x^2)*x^2,x,1);
```

(%o7) $2\,x^3\,\%e^{x^2} + 2\,x\,\%e^{x^2}$

Auch das berechnen der zweiten oder fünften Ableitung geht sehr schnell.

(%i16) `diff(exp(x^2)*x^2,x,2);`

(%o16) $4\,x^4\,\%e^{x^2}+10\,x^2\,\%e^{x^2}+2\,\%e^{x^2}$

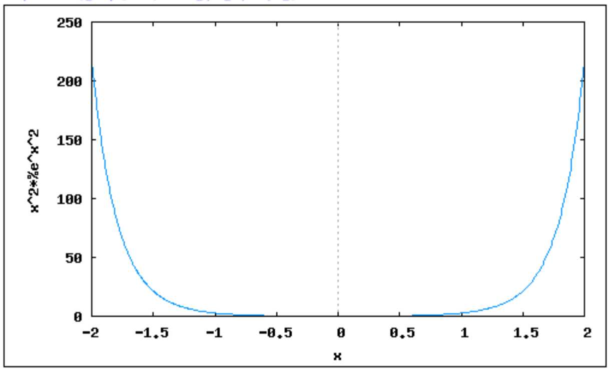
(%i17) `diff(exp(x^2)*x^2,x,5);`

(%o17) $32\,x^7\,\%e^{x^2}+320\,x^5\,\%e^{x^2}+760\,x^3\,\%e^{x^2}+360\,x\,\%e^{x^2}$

Wir lassen noch schnell die Funktion und die Ableitung plotten.

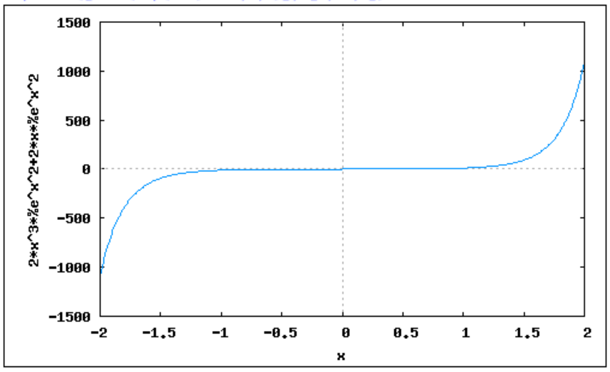
(%i9) `wxplot2d([exp(x^2)*x^2], [x,-2,2])$`

(%t9)



(%i10) `wxplot2d([diff(exp(x^2)*x^2,x,1)], [x,-2,2])$`

(%t10)



Ende des Exkurses

4.3. Sätze über differenzierbare Funktionen

Satz 170: Es sei $a < b$, $I :=]a, b[$ und $f : I \to \mathbb{R}$ differenzierbar. Ist $x_m \in I$ eine *Extremstelle von* f, so ist

$$f'(x_m) = 0.$$

BEWEIS: Angenommen x_m sei ein Maximum. Für Minima geht der Beweis analog. So gilt für alle $x \in I$ die Ungleichung $f(x) \leq f(x_m)$. Für alle $x \in I$ mit $x < x_m$ gilt

$$\frac{f(x) - f(x_m)}{x - x_m} \geq 0$$

und damit

$$\lim_{x \nearrow x_m} \frac{f(x) - f(x_m)}{x - x_m} \geq 0.$$

Für alle $x \in I$ mit $x > x_m$ gilt analog dazu

$$\frac{f(x) - f(x_m)}{x - x_m} \leq 0$$

und damit

$$\lim_{x \searrow x_m} \frac{f(x) - f(x_m)}{x - x_m} \leq 0.$$

Da der Grenzwert $\lim_{x \to x_m} \frac{f(x)-f(x_m)}{x-x_m}$ nach Vorraussetzung existiert, müssen die beiden einseitigen Grenzwerte gleich sein. Womit bewiesen ist, dass die Ableitung den Wert Null hat. □

Satz 171 (Satz von Rolle)**:** Es sei $a < b$ außerdem sei die Funktion $f : [a, b] \to \mathbb{R}$ stetig und auf $]a, b[$ differenzierbar. Ist außerdem $f(a) = f(b)$, so existiert ein $\xi \in]a, b[$ mit $f'(\xi) = 0$.

BEWEIS: Für konstante Funktionen ist dies klar. Es sei f also nicht konstant, so nimmt f nach Satz 135 auf $[a, b]$ ihr Maximum und ihr Minimum an. Ist x_m eine Minimalstelle und x_M eine Maximalstelle, so ist

$$f(x_m) \neq f(x_M),$$

da f nicht konstant ist. Also liegt x_m oder x_M in $]a, b[$ und die Behauptung folgt aus Satz 170. □

Wir kommen nun zu einem sehr wichtigen und zentralen Satz der Analysis. Aus diesem Satz werden wir im Folgenden zum Beispiel ein Kriterium ableiten, mit dem differenzierbare Funktionen auf Extrema untersucht werden können. Auch bei der Untersuchung der trigonometrischen Funktionen $\sin$ und $\cos$ wird uns der folgende Satz wertvolle Dienste leisten.

Satz 172 (Satz von Taylor)**:** Es sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine $(n + 1)$-mal differenzierbare Funktion. Ist $x_0 \in I$, dann existiert für alle $x \in I$ ein $\xi \in]x, x_0[$ (bzw. $\xi \in]x_0, x[$) mit

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n + 1)!}(x - x_0)^{n+1}.$$

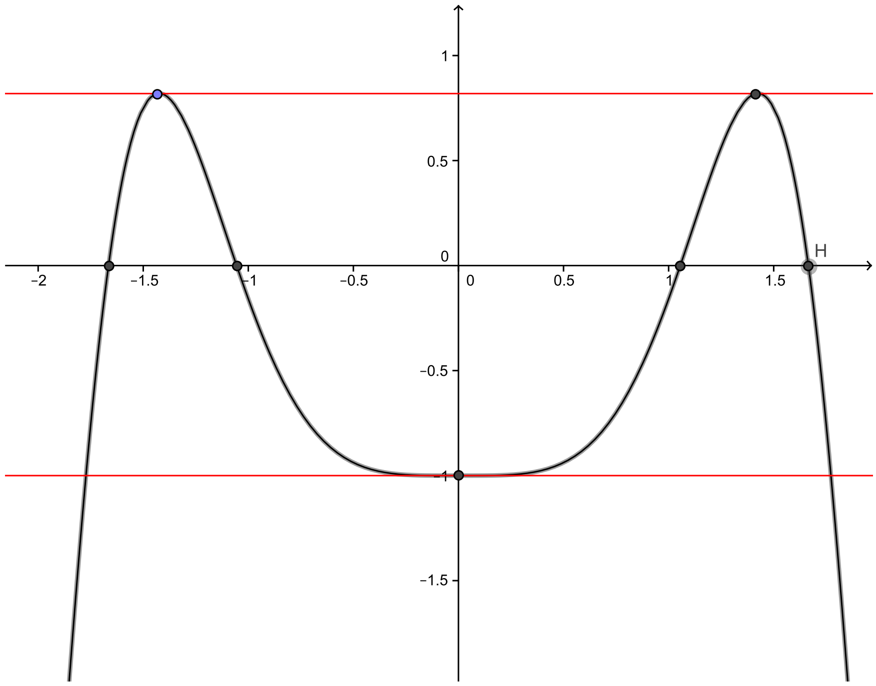


Abbildung 4.5.: Geometrische Deutung des Satzes von Rolle

BEMERKUNG: Je öfter eine Funktion also stetig differenzierbar ist, desto „ähnlicher“ wird sie einer Potenzreihe. Dennoch ist nicht jede C^∞-Funktion durch eine Potenzreihe gegeben. Ein Beispiel hierfür lautet

$$\begin{aligned} f &: \mathbb{R} \to \mathbb{R} \\ x &\mapsto \begin{cases} e^{-\frac{1}{x^2}}, & x > 0 \\ 0, & x \leq 0. \end{cases} \end{aligned}$$

Dass $f \in C^\infty(\mathbb{R})$ ist, folgt nach einiger Rechnung aus Lemma 159. Ließe sich f als Potenzreihe schreiben, so müssten alle Koeffizienten nach dem Identitätssatz gleich Null sein.

BEWEIS: Es sei $x < x_0$. Man betrachte die Hilfsfunktion

$$H(t) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k - \frac{\rho}{(n+1)!}(x-t)^{n+1},$$

wobei wir ρ so wählen, dass $H(x_0) = 0$ ist. Außerdem ist $H(x) = 0$. Die Hilfsfunktion ist auf $[x, x_0]$ stetig und auf $]x, x_0[$ differenzierbar, da f $(n+1)$-mal differenzierbar ist. Nach dem Satz von Rolle 171 existiert ein $\xi \in]x, x_0[$ mit $H'(\xi) = 0$.

Schauen wir uns die Ableitung (nach der Variablen t) der Funktion H an:

$$\begin{aligned} H'(t) &= -f'(t) - \sum_{k=1}^{n} \left(\frac{f^{(k+1)}(t)}{k!}(x-t)^k - \frac{f^{(k)}(t)}{k!} k\,(x-t)^{k-1} \right) \\ &\quad + \frac{\rho}{(n+1)!}(n+1)(x-t)^n \\ &= -\sum_{k=0}^{n} \frac{f^{(k+1)}(t)}{k!}(x-t)^k + \sum_{k=0}^{n-1} \frac{f^{(k+1)}(t)}{k!}\,(x-t)^k + \frac{\rho}{n!}(x-t)^n \\ &= -\frac{f^{(n+1)}(t)}{n!}(x-t)^n + \frac{\rho}{n!}(x-t)^n. \end{aligned}$$

Aus $H'(\xi) = 0$ und $\xi \neq x$ folgt also $\rho = f^{(n+1)}(\xi)$. □

BEMERKUNG: Im Beweis sieht man, dass es ausreicht, wenn f in $[x, x_0]$ (bzw. $[x_0, x]$) n-mal stetig differenzierbar und in $]x, x_0[$ (bzw. $]x_0, x[$) $(n+1)$-mal differenzierbar ist. Die $(n+1)$-Ableitungen in x und x_0 werden nicht benötigt.

Definition 173 (Taylorpolynom)**:** Es sei $f : I \to \mathbb{R}$ n-mal differenzierbar und $x_0 \in I$. Die Polynomfunktion

$$T_{f,n}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

heißt *Taylorpolynom vom Grad* n *der Funktion* f *um den Entwicklungspunkt* x_0.

Beispiel: Wir wollen mit Hilfe des Satzes von Taylor $\sqrt{2}$ näherungsweise berechnen. Dazu betrachten wir die Funktion $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ mit $f(x) = \sqrt{x+1} = (x+1)^{1/2}$ und setzen $x_0 = 0$. Der Satz von Taylor sagt, dass es ein $\xi \in]0, 1[$ gibt mit

$$\sqrt{2} = f(1) = f(0) + \frac{f^{(1)}(0)}{1!} + \frac{f^{(2)}(0)}{2!} + \frac{f^{(3)}(0)}{3!} + \frac{f^{(4)}(\xi)}{4!}.$$

Es sind

$$\begin{aligned} f'(x) &= \frac{1}{2}(x+1)^{-\frac{1}{2}} \\ f''(x) &= -\frac{1}{4}(x+1)^{-\frac{3}{2}} \\ f'''(x) &= \frac{3}{8}(x+1)^{-\frac{5}{2}} \\ f^{(4)}(x) &= -\frac{15}{16}(x+1)^{-\frac{7}{2}}. \end{aligned}$$

Also haben wir

$$\sqrt{2} = f(1) \approx 1 + \frac{1}{2} - \frac{1}{8} + \frac{1}{16} = 1,4375$$

und der Fehler den wir bei dieser Abschätzung machen ist kleiner oder gleich

$$\max_{\xi\in]0,1[}\left|\frac{f^{(4)}(\xi)}{4!}\right| \leq \frac{15}{4!\cdot 16} = 0,0390625.$$

Wir erhalten bessere Abschätzungen, wenn wir Werte approximieren, die näher am Entwicklungspunkt Null liegen.

Exkurs: Ableiten mit GeoGebra

Wir wollen die vorstehenden Rechnungen bei der näherungsweisen Berechnung von $\sqrt{2}$ auf Grundlage des Satzes von Taylor nun mit dem Computeralgebrasystem GeoGebra nachvollziehen.

Die Funktion wird einfach durch Eingabe von `f(x)=sqrt(x+1)` definiert. Möchte man die Ableitung berechnen, so kann man `f′` eingeben. Alternativ kann man auch `Ableitung(f)` eingeben. Auf die zweite Ableitung kann man mit `f″` oder `Ableitung(f,2)` zugreifen.

Die obige Näherung von $\sqrt{2}$ kann man also mittels `f(0)+f′(0)+f″(0)/2!+f″′(0)/3!` berechnen lassen.

Den maximalen Fehler, den man hierbei macht, berechnet man mittels `max(abs(f″″(x)/4!),0,1)`. Hier wird also das Maximum von $\left|f^{(4)}(x)/4!\right|$ für x zwischen 0 und 1 bestimmt.

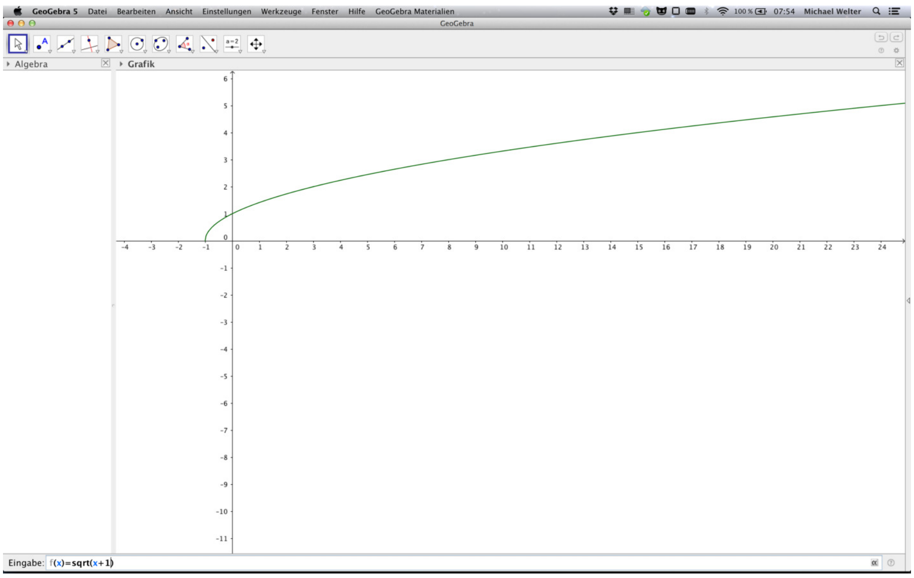


Abbildung 4.6.: Definition der Funktion $f(x) = \sqrt{x+1}$ in GeoGebra

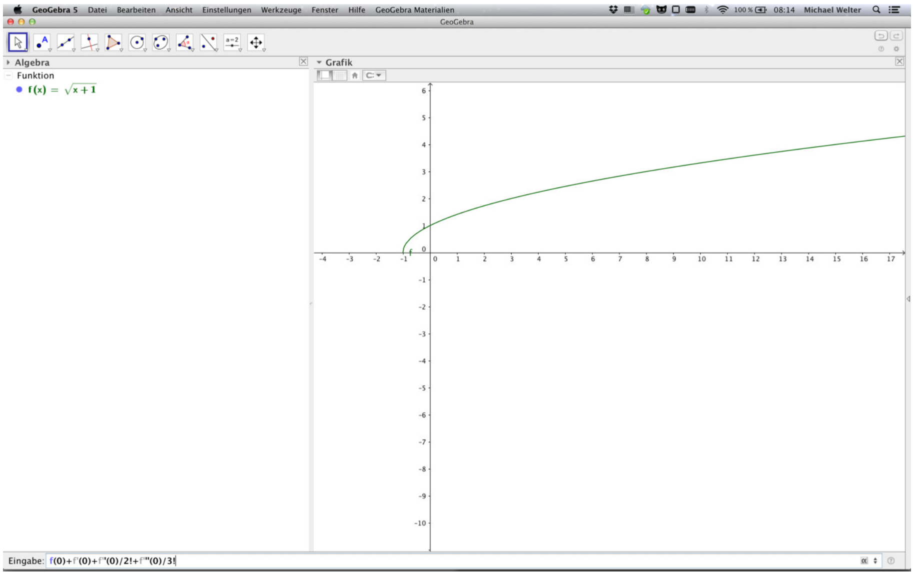


Abbildung 4.7.: Berechnung der Näherung von $\sqrt{2}$ mittels des Satzes von Taylor in GeoGebra

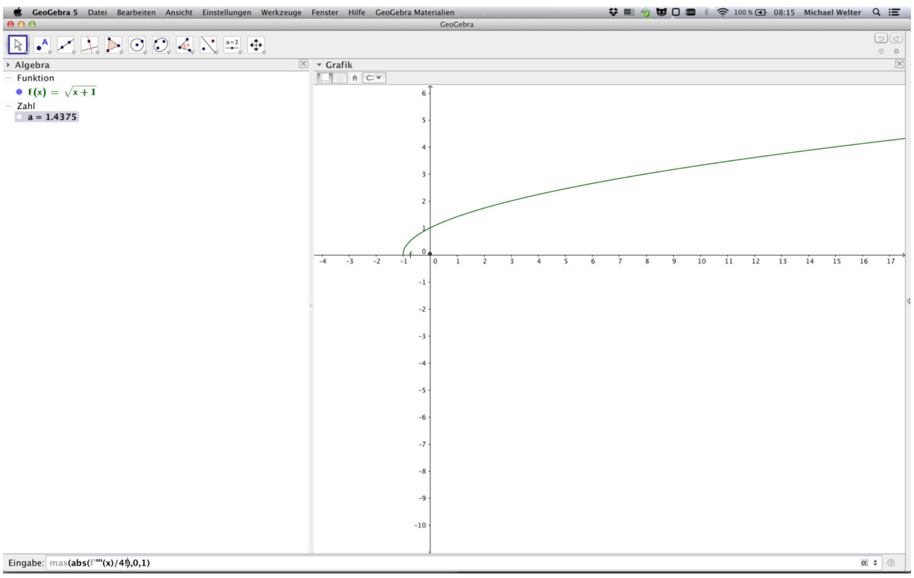


Abbildung 4.8.: Berechnung des Fehlers bei der Approximation von $\sqrt{2}$ in GeoGebra

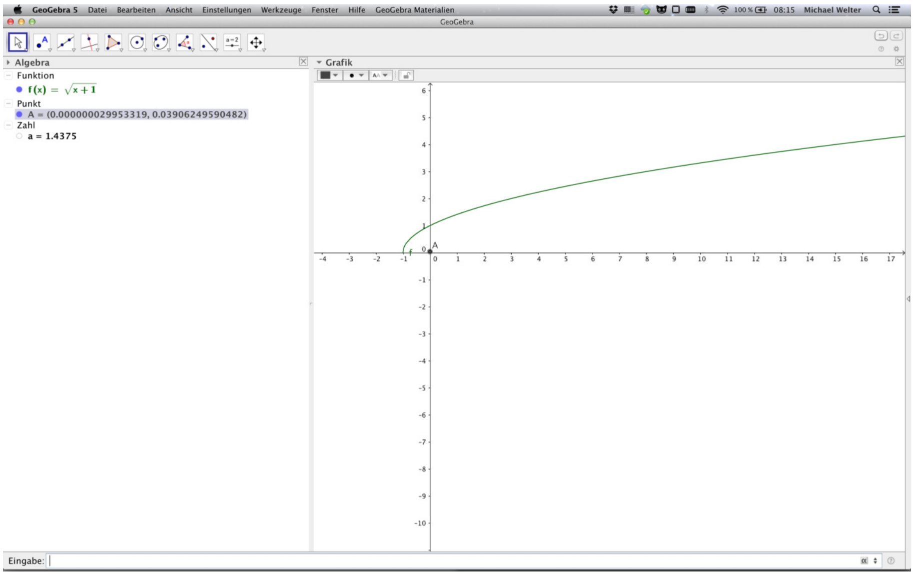


Abbildung 4.9.: Ausgabe des Approximationsfehlers in GeoGebra

Ende des Exkurses

Wir haben einige wichtige Funktionen, wie etwa $\exp, \sin$ und $\cos$ über Potenzreihen eingeführt. Der Satz von Taylor zeigt, dass Funktionen umso mehr Potenzreihen ähneln, je öfter sie abgeleitet werden dürfen. Dies heißt aber nicht, dass jede beliebig oft differenzierbare Funktion durch eine Potenzreihe dargestellt werden kann. Wir wollen uns dies genauer anschauen.

Definition 174 (Taylorreihe)**:** Es sei $f : I \to \mathbb{R}$ beliebig oft differenzierbar und $x_0 \in I$. Die Potenzreihe

$$T_f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

heißt die *Taylorreihe von f mit Entwicklungspunkt* x_0.

BEMERKUNG: Die Reihe T_f muss nicht notwendigerweise gegen f konvergieren, das heißt es kann $x \in I$ geben, für die $T_f(x)$ konvergiert, aber $T_f(x) \neq f(x)$ gilt. Wir betrachten hierzu wieder die Funktion

$$f : \mathbb{R} \to \mathbb{R}$$
$$x \mapsto \begin{cases} e^{-\frac{1}{x^2}}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

aus obigem Beispiel. Es ist $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}$. Die Taylorreihe von f ist also die Nullfunktion und diese stimmt für alle positiven x nicht mit der Funktion f überein.

Satz 175: Es sei $f : I \to \mathbb{R}$ beliebig oft differenzierbar und $x_0 \in I$. Existiert ein $\delta > 0$, so dass $D := \{x \in \mathbb{R} |\ |x - x_0| < \delta\} \subset I$ ist, und gilt für alle $x \in D$

$$\lim_{n\to\infty} r_n(x, x_0) = 0,$$

wobei $r_n(x, x_0) := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$ das Restglied im Satz von Taylor ist, so lässt sich f in einer Umgebung von x_0 in eine Potenzreihe entwickeln und es gilt dort $f(x) = T_f(x)$.

Beispiele:

1) Wir können den vorstehenden Satz nun benutzen, um die Potenzreihendarstellung des Logarithmus um den Entwicklungspunkt 1 zu berechnen. Wir setzen also $f(x) := \log(x)$. Wenn man sich die ersten Ableitung des Logarithmus anschaut, sieht man schnell, dass für alle $k \in \mathbb{N}$

$$f^{(k)}(x) = (-1)^{k+1}\frac{(k-1)!}{x^k}$$

ist. Dies lässt sich natürlich induktiv beweisen. Also ist $f^{(k)}(1) = (-1)^{k+1}(k - 1)!$ für $k \geq 1$ und somit wegen $\log(1) = 0$

$$T_f(x) = \sum_{k=1}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k = \sum_{k=0}^{\infty}(-1)^{k+1}\frac{(x-1)^k}{k}.$$

Für das Restglied erhalten wir nach dem Satz von Taylor, dass ein ξ zwischen 1 und x (oder x und 1) existiert, so dass

$$r_n(x, 1) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-1)^{n+1} = (-1)^{n+2}\frac{(x-1)^{n+1}}{(n+1)\xi^{n+1}}.$$

Dies geht sicherlich in einer Umgebung von 1 gegen Null, wo $|x - 1| < |\xi|$ gilt.

2) Der Weg die Potenzreihenentwicklung einer gegebenen Funktion über die Ableitungen der Funktion zu bestimmen, ist oft mühsam. Einfacher ist es häufig, neue Reihenentwicklungen aus bekannten Reihen herzuleiten. Wir werden uns in Abschnitt 4.17 alternative Wege anschauen, wie man aus bekannten Reihenentwicklungen über Ableiten und Integrieren neue Reihen herleiten kann. Hier wollen wir jetzt nur ein paar einfache Beispiele geben, in denen wir in bekannte Reihen das x durch eine Transformation ersetzen. Starten wir etwa mit der für $|x| < 1$ gültigen Formel

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k,$$

so erhalten wir, wenn wir x durch x^d ersetzen

$$\frac{1}{1-x^d} = \sum_{k=0}^{\infty} x^{dk},$$

wenn wir x durch $-x$ ersetzen

$$\frac{1}{1+x} = \sum_{k=0}^{\infty}(-1)^k x^k,$$

oder wenn wir x durch $1-x$ ersetzen -solange $|1-x|<1$ ist-

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

Starten wir mit der Exponentialreihe, so können wir etwa

$$e^{x^2} = \sum_{n=0}^{\infty}\frac{1}{n!}x^{2n}$$

herleiten.

Definition 176 (Lokales Minimum bzw. lokales Maximum)**:** Es sei $f:]a,b[\to \mathbb{R}$ eine Funktion.

(i) Man sagt, dass f in $x_0 \in]a,b[$ ein *lokales Maximum* hat, wenn ein $\varepsilon > 0$ existiert, so dass

$$f(x_0) \geq f(x)$$

für alle $x \in]x_0-\varepsilon, x_0+\varepsilon[$ gilt. Tritt in dieser Forderung Gleichheit nur für $x = x_0$ ein, so spricht man von einem *isolierten, lokalen Maximum*.

(ii) Man sagt, dass f in $x_0 \in]a,b[$ ein *lokales Minimum* hat, wenn ein $\varepsilon > 0$ existiert, so dass

$$f(x_0) \leq f(x)$$

für alle $x \in]x_0-\varepsilon, x_0+\varepsilon[$ gilt. Tritt in dieser Forderung Gleichheit nur für $x = x_0$ ein, so spricht man von einem *isolierten, lokalen Minimum*.

BEMERKUNG: Ist f differenzierbar, so gilt nach Satz 170 $f'(x_0) = 0$ für lokale Extrema. Wähle $I =]x_0-\varepsilon, x_0+\varepsilon[$.

Satz 177: Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f \in C^{n+1}(I)$. Ist $x_0 \in I$ derart, dass

$$f'(x_0) = f^{(2)}(x_0) = \ldots = f^{(n)}(x_0) = 0$$

und $f^{(n+1)}(x_0) \neq 0$ gilt, so folgt:

(i) Ist n gerade, so ist x_0 kein Extremum von f.

(ii) Ist n ungerade, so ist x_0 ein lokales Minimum, falls $f^{(n+1)}(x_0) > 0$ gilt und ein lokales Maximum, falls $f^{(n+1)}(x_0) < 0$ gilt.

BEWEIS: Angenommen es sei $f^{(n+1)}(x_0) > 0$. Der andere Fall geht analog. Da $f^{(n+1)}$ stetig ist, existiert ein $\varepsilon > 0$, so dass $f^{(n+1)}(x) > 0$ für alle $x \in]x_0 - \varepsilon, x_0 + \varepsilon[$ ist. Nach dem Satz von Taylor 172 existiert zu jedem x ein ξ zwischen x und x_0, so dass

$$f(x) = f(x_0) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

gilt. Ist n gerade und damit $n+1$ ungerade, so wechselt der Ausdruck $(x - x_0)^{n+1}$ je nachdem, ob x größer oder kleiner x_0 ist, das Vorzeichen. Also kann x_0 kein Extremum sein.
Ist n ungerade, so ist $(x - x_0)^{n+1} \geq 0$ für alle $x \in]x_0 - \varepsilon, x_0 + \varepsilon[$. Also gilt $f(x) \geq f(x_0)$ für alle $x \in]x_0 - \varepsilon, x_0 + \varepsilon[$ und x_0 ist ein lokales Minimum. □

Satz 178 (Mittelwertsatz der Differentialrechnung)**:** Es sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, welche auf $]a, b[$ differenzierbar ist. Dann existiert ein $\xi \in]a, b[$ mit

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

BEWEIS: Der Beweis folgt sofort aus dem Satz 172 von Taylor mit $n = 0$. (Beachte die Bemerkung nach dem Beweis von Satz 172.) □

Problem 70: *(i) Es seien $a, b \in \mathbb{R}, a < b$ und $c \in]a, b[$. Die Funktion $f :]a, b[\to \mathbb{R}$ sei stetig auf $]a, b[$ und differenzierbar auf $]a, b[\setminus \{c\}$. Zeigen Sie, dass f in c differenzierbar ist, wenn*

$$\lim_{x \searrow c} f'(x) = \lim_{x \nearrow c} f'(x)$$

gilt.
Tipp: Mittelwertsatz der Differentialrechnung.

(ii) Ist die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) := \begin{cases} x^2 + 1 & , x \neq 0 \\ 0 & , x = 0 \end{cases}$$

in 0 differenzierbar?

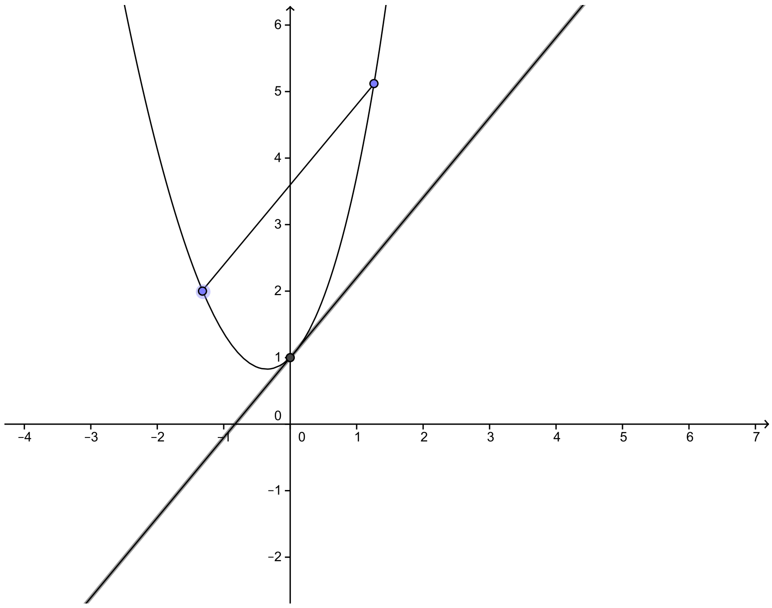


Abbildung 4.10.: Geometrische Deutung des Mittelwertsatzes der Differentialrechnung

(iii) Ist die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) := \begin{cases} x^2 \sin \frac{1}{x^2} & , x \neq 0 \\ 0 & , x = 0 \end{cases}$$

in 0 differenzierbar?

Satz 179: Ist $f :]a, b[\to \mathbb{R}$ differenzierbar, so gilt:

(i) Ist $f'(x) = 0$ für alle $x \in]a, b[$, so ist f auf $]a, b[$ konstant.

(ii) Ist $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$) für alle $x \in]a, b[$, so ist f auf diesem Intervall monoton wachsend (monoton fallend).

(iii) Ist $f'(x) > 0$ (bzw. $f'(x) < 0$) für alle $x \in]a, b[$, so ist f auf diesem Intervall streng monoton wachsend (streng monoton fallend).

BEWEIS:

(i) Es gelte also $f'(x) = 0$ für alle $x \in]a, b[$. Es seien $c, d \in]a, b[$ mit $c < d$ beliebig. Nach dem Mittelwertsatz 178 gibt es ein $\xi \in]c, d[$, so dass

$$\frac{f(d) - f(c)}{d - c} = f'(\xi)$$

gilt. Da $f'(\xi) = 0$ ist, muss also $f(c) = f(d)$ sein, was wegen der Beliebigkeit von c und d die Behauptung beweist.

(ii) Es gelte also $f'(x) \geq 0$ für alle $x \in]a, b[$. Wieder seien $c, d \in]a, b[$ mit $c < d$ beliebig gewählt. Auf Grund des Mittelwertsatzes existiert ein $\xi \in]c, d[$, so dass

$$f'(\xi) = \frac{f(d) - f(c)}{d - c}$$

gilt. Da $f'(\xi) \geq 0$ ist, muss also $f(d) \geq f(c)$ sein.

□

Problem 71: *Beweisen Sie Teil (iii) des vorstehenden Satzes.*

Problem 72: *Wir betrachten die Funktion $f : \mathbb{R} \to \mathbb{R}, x \mapsto e^{-x^2}$.*

(i) Bestimmen Sie die lokalen Extrema von f.

(ii) Wie ist das Monotonieverhalten von f?

(iii) Bestimmen Sie

$$\lim_{x \to +\infty} f(x) \quad \textit{und} \quad \lim_{x \to -\infty} f(x).$$

(iv) Skizzieren Sie den Graphen von f.

Satz 180 (Verallgemeinerter Mittelwertsatz)**:** Es seien $a < b$ Punkte und $f, g : [a, b] \to \mathbb{R}$ stetige Funktionen, welche auf $]a, b[$ differenzierbar sind. Außerdem sei $g'(x) \neq 0$ für alle $x \in]a, b[$. Dann existiert ein $\xi \in]a, b[$ mit

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

BEMERKUNG: Für $g(x) = x$ folgt der Mittelwertsatz 178.

Problem 73: *Beweisen Sie den vorstehenden Satz.*
Tipp: Man betrachte die Hilfsfunktion $H(x) := f(x) - \frac{f(b)-f(a)}{g(b)-g(a)}(g(x) - g(a))$ und wende auf diese den Satz von Rolle an.

4.4. Untersuchung der trigonometrischen Funktion

Lemma 181: Für alle $x \in]0, 2[$ gilt $\sin(x) > 0$.

BEWEIS: Es sei $x \in]0, 2[$ beliebig. Wir wenden den Satz von Taylor (Satz 172) mit $n = 2$ und $x_0 = 0$ an und erhalten, dass ein $\xi \in]0, 2[$ mit

$$\begin{aligned}\sin(x) &= \sin(0) + \frac{\sin'(0)}{1!}x + \frac{\sin^{(2)}(0)}{2!}x^2 + \frac{\sin^{(3)}(\xi)}{3!}x^3 \\ &= 0 + x + 0 \cdot x^2 - \frac{\cos(\xi)}{3!}x^3\end{aligned}$$

existiert. Da für alle x die Ungleichung $|\cos(x)| \leq 1$ gilt, erhält man für alle $x \in]0, 2[$

$$\left|\frac{\cos(\xi)}{3!}x^3\right| \leq \frac{|x|^3}{3!} \overset{x<2}{\leq} \frac{4}{6}|x| = \frac{2}{3}|x|.$$

Somit findet man auf $]0, 2[$ für den Sinus die Abschätzung

$$\sin(x) \geq x - \frac{2}{3}x = \frac{1}{3}x > 0.$$

□

Lemma 182: Es gilt die Ungleichung

$$\cos(2) \leq -\frac{1}{3}.$$

BEWEIS: Es sei $x \in [0, 2]$. Auf Grund des Satzes von Taylor erhält man für $x_0 = 0$ und $n = 3$ wieder

$$\cos(x) = 1 + 0 - \frac{x^2}{2} + 0 + R(x)$$

mit

$$R(x) = \left|\frac{\cos(\xi)}{4!}x^4\right| \leq \frac{|x|^4}{4!} \leq \frac{16}{24} = \frac{2}{3}$$

und $\xi \in]0, 2[$. Daraus folgt

$$\cos(2) \leq 1 - \frac{4}{2} + \frac{2}{3}.$$

□

Satz 183: Die Funktion $\cos : [0,2] \to \mathbb{R}$ besitzt genau eine Nullstelle.

BEWEIS: Aus Satz 179 und Lemma 181 folgt, dass $\cos$ auf $[0,2]$ streng monoton fallend ist. Also hat $\cos$ auf diesem Intervall höchstens eine Nullstelle.
Da $\cos(0) = 1$ und $\cos(2) \leq -\frac{1}{3}$ gilt, hat $\cos$ nach dem Zwischenwertsatz 130 mindestens eine Nullstelle in $[0,2]$. $\square$

Definition 184 (Die Zahl π)**:** Die eindeutige bestimmte Nullstelle der Funktion $\cos : [0,2] \to \mathbb{R}$ wird mit $\frac{\pi}{2}$ bezeichnet.

Satz 185: (i) Die Funktionen $\sin, \cos : \mathbb{R} \to \mathbb{R}$ sind *periodisch mit Periode* 2π, das heißt für alle $x \in \mathbb{R}$ gilt

$$\begin{aligned}\cos(x+2\pi) &= \cos(x)\\ \sin(x+2\pi) &= \sin(x).\end{aligned}$$

(ii) Für alle $x \in \mathbb{R}$ gilt

$$\begin{aligned}\cos(x+\pi) &= -\cos(x)\\ \sin(x+\pi) &= -\sin(x)\\ \cos\left(x+\frac{\pi}{2}\right) &= -\sin(x)\\ \sin\left(x+\frac{\pi}{2}\right) &= \cos(x).\end{aligned}$$

(iii) Folgende Wertetabelle gilt:

x	0	$\frac{\pi}{2}$	π	$\frac{3}{2}\pi$	2π
$\cos(x)$	1	0	-1	0	1
$\sin(x)$	0	1	0	-1	0

BEWEIS: Nach Definition ist $\cos\left(\frac{\pi}{2}\right) = 0$. Außerdem gilt nach Satz 121 $\cos^2\left(\frac{\pi}{2}\right) + \sin^2\left(\frac{\pi}{2}\right) = 1$ und damit folgt $\sin^2\left(\frac{\pi}{2}\right) = 1$. Wegen Lemma 181 erhält man $\sin\left(\frac{\pi}{2}\right) = 1$. Außerdem folgt

$$\begin{aligned}\cos(\pi) &= \cos\left(\frac{\pi}{2}+\frac{\pi}{2}\right) = \cos^2\left(\frac{\pi}{2}\right) - \sin^2\left(\frac{\pi}{2}\right) = -1\\ \sin(\pi) &= \sin\left(\frac{\pi}{2}+\frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2}\right)\sin\left(\frac{\pi}{2}\right) + \sin\left(\frac{\pi}{2}\right)\cos\left(\frac{\pi}{2}\right) = 0.\end{aligned}$$

Betrachte nun

$$\begin{aligned}\cos(x+\pi) &= \cos(x)\cos(\pi) - \sin(x)\sin(\pi) = -\cos(x)\\ \sin(x+\pi) &= \cos(x)\sin(\pi) + \sin(x)\cos(\pi) = -\sin(x),\end{aligned}$$

damit folgt nun

$$\begin{aligned}\cos(x+2\pi) &= \cos((x+\pi)+\pi) = -\cos(x+\pi) = \cos(x)\\ \sin(x+2\pi) &= \sin((x+\pi)+\pi) = -\sin(x+\pi) = \sin(x).\end{aligned}$$

Außerdem gilt

$$\begin{aligned}\cos\left(x+\frac{\pi}{2}\right) &= \cos(x)\cos\left(\frac{\pi}{2}\right) - \sin(x)\sin\left(\frac{\pi}{2}\right) = -\sin(x)\\ \sin\left(x+\frac{\pi}{2}\right) &= \cos(x)\sin\left(\frac{\pi}{2}\right) - \sin(x)\cos\left(\frac{\pi}{2}\right) = \cos(x).\end{aligned}$$

$\square$

Satz 186: Es sei $x \in \mathbb{R}$.

(i) Es gilt $\cos(x) = 0$ genau dann, wenn $x \in \left\{\frac{\pi}{2} + k\pi \,|\, k \in \mathbb{Z}\right\}$.

(ii) Es gilt $\sin(x) = 0$ genau dann, wenn $x \in \{k\pi \,|\, k \in \mathbb{Z}\}$.

BEWEIS: Wegen $\cos\left(x+\frac{\pi}{2}\right) = -\sin(x)$ reicht es Aussage (ii) zu zeigen.
Hinrichtung: Für $x \in \left[0, \frac{\pi}{2}\right[$ gilt $\cos(x) > 0$. Wegen der Symmetrie $(\cos(-x) = \cos(x))$, gilt dies also für alle $x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. Wegen $\sin\left(x+\frac{\pi}{2}\right) = \cos(x)$ gilt $\sin(x) > 0$ für alle $x \in]0,\pi[$. Auf Grund von $\sin(x+\pi) = -\sin(x)$, besitzt $\sin$ also in $[0, 2\pi]$ nur die Nullstellen 0, π und 2π.
Es sei nun $y \in \mathbb{R}$ mit $\sin(y) = 0$. Es existiert ein $m \in \mathbb{Z}$, so dass $0 \leq y - 2\pi m < 2\pi$ gilt. Wegen der Periodizität gilt

$$\sin(y - 2\pi m) = \sin(y) = 0.$$

Also ist $y - 2\pi m \in \{0, \pi\}$ und y von der Gestalt $k\pi$ mit $k \in \mathbb{Z}$.

Rückrichtung: Ist klar. $\square$

Satz 187 (Polarkoordinaten)**:** (i) Die Funktion $\exp : \mathbb{C} \to \mathbb{C}$ hat Periode $2\pi i$.

(ii) Zu jeder komplexen Zahl $z \in \mathbb{C}\backslash\{0\}$ existiert eine eindeutig bestimmte Zahl $\varphi \in]-\pi, \pi]$, so dass

$$z = |z|e^{i\varphi}$$

gilt. Diese Darstellung trägt die Bezeichnung *Polarkoordinatendarstellung der komplexen Zahl z*. φ bezeichnet man als das *Argument von z*.

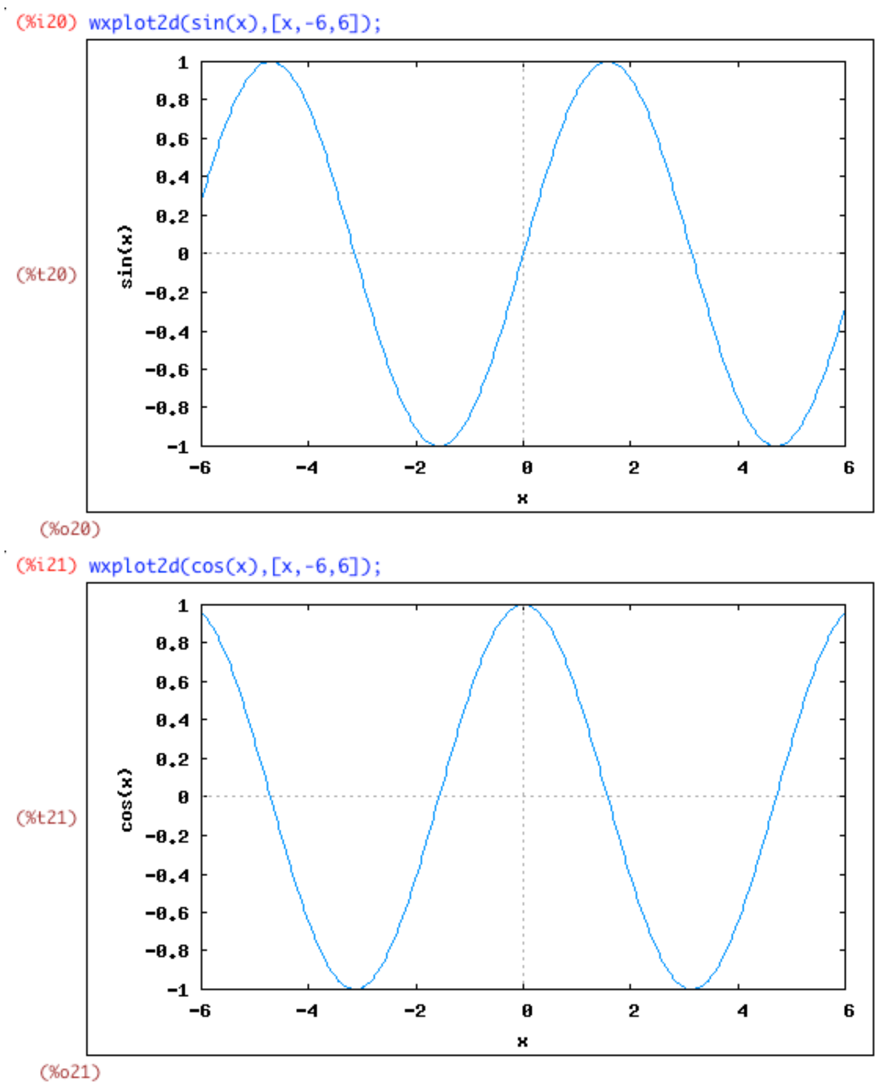


Abbildung 4.11.: Die Sinusfunktion und die Cosinusfunktion im Maximaplot

BEMERKUNG: Die Polarkoordinatendarstellung erlaubt es uns, eine geometrische Interpretation der komplexen Multiplikation zu geben. Das Argument φ beschreibt den Winkel zwischen der x-Achse und der Strecke vom Nullpunkt zur komplexen Zahl z, wenn wir diese als Punkt im $\mathbb{R}^2$ auffassen. Der Betrag $|z|$ ist der Abstand von z zum Nullpunkt. Multipliziert man zwei komplexe Zahlen z und w, so werden die Beträge multipliziert und die Argumente addiert.

BEWEIS:

(i) Es gilt

$$\begin{aligned} \exp(z + 2\pi i) &= \exp(z)\exp(2\pi i) \\ &= \exp(z)(\cos(2\pi) + i\sin(2\pi)) \\ &= \exp(z). \end{aligned}$$

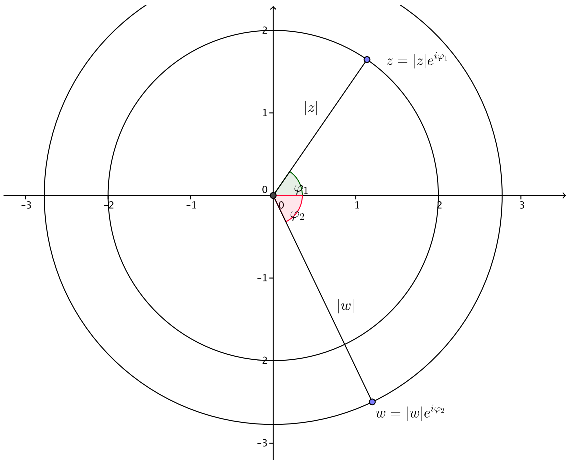


Abbildung 4.12.: Situation bei der komplexen Multiplikation

(ii) Es sei $z \in \mathbb{C}\backslash\{0\}$ beliebig und setze $\omega = \frac{z}{|z|}$. Dann gilt $|\omega| = 1$. Es sei weiter $\omega = a + ib$ mit $a, b \in \mathbb{R}$. Wegen $1 = |\omega|^2 = a^2 + b^2$ ist $|a| \leq 1$. Also existiert genau ein $\alpha \in [0, \pi]$ mit $\cos(\alpha) = a$. Auf Grund von $\cos^2(\alpha) + \sin^2(\alpha) = 1$ gilt

$$\sin(\alpha) = \pm\sqrt{1 - \cos^2(\alpha)} = \pm\sqrt{1 - a^2} = \pm b.$$

Ist $\sin(\alpha) = b$, so setze $\varphi := \alpha$, ist $\sin(\alpha) = -b$, so setze $\varphi := -\alpha$. In beiden Fällen gilt

$$e^{i\varphi} = \cos(\varphi) + i\sin(\varphi) = a + ib = \omega = \frac{z}{|z|}.$$

□

Nun können wir ganz einfach alle Lösungen von Gleichungen der Art

$$z^n = \gamma$$

mit $n \in \mathbb{N}$ und $\gamma \in \mathbb{C} \setminus \{0\}$ angeben. Dazu schreiben wir zunächst γ in Polarkoordinaten. Ist $\gamma = re^{i\varphi}$, so lösen die n komplexen Zahlen

$$z_k := \sqrt[n]{r}e^{\frac{\varphi}{n}i + \frac{2\pi i}{n}k}$$

mit $k \in \{0, \ldots, n-1\}$ die Gleichung $z^n = \gamma$.

Die n paarweise verschiedenen Lösungen der Gleichung $z^n = 1$ nennt man die *n-ten Einheitswurzeln*.

Problem 74: *Zeigen Sie, dass die z_k mit $k \in \{0, \ldots, n-1\}$ paarweise verschieden sind.*

Problem 75: *Für alle $x \in \mathbb{R}$ mit $\cos x \neq 0$ definieren wir die* Tangensfunktion $\tan$ *durch*

$$\tan x := \frac{\sin x}{\cos x}.$$

(i) Was ist der Definitionsbereich der Tangensfunktion?

(ii) Zeigen Sie: Für alle x aus dem Definitionsbereich der Tangensfunktion gilt

a) $\tan(x + \pi) = \tan x$,

b) $\tan(-x) = -\tan x$,

c) $(\tan x)' = \frac{1}{\cos^2 x} = 1 + \tan^2 x$.

(iii) Zeigen Sie, dass die Funktionen $\sin : [-\pi/2, \pi/2] \to [-1, 1]$, $\cos : [0, \pi] \to [-1, 1]$ und $\tan :]-\pi/2, \pi/2[\to \mathbb{R}$ bijektiv sind und folglich Umkehrfunktionen besitzen. Diese heißen Arkus-Funktionen *und werden mit $\arcsin, \arccos$ und $\arctan$ bezeichnet.*

(iv) Bestimmen Sie für jede der Arkus-Funktionen, wo sie differenzierbar ist und berechnen Sie die Ableitungen der Arkus-Funktionen.

Exkurs: Die komplexe Logarithmusfunktion und komplexe Potenzen

Definition 188 (Hauptwert des Arguments einer komplexen Zahl)**:** Die zu $z \in \mathbb{C} \setminus \{0\}$ eindeutig bestimmt Zahl $\varphi \in]-\pi, \pi]$ mit $z = |z| e^{i\varphi}$ nennen wir den *Hauptwert des Arguments von z* und für schreiben $\arg(z)$ für φ.

Definition 189 (Die komplexe Logarithmusfunktion)**:** Für $z \in \mathbb{C} \setminus \{0\}$ setzen wir

$$\log(z) := \log(|z|) + i \arg(z),$$

wobei der $\log$ auf der rechten Seite die reelle Logarithmusfunktion und $\arg(z)$ den Hauptzweig des Logarithmus bezeichnet.

BEMERKUNG: Ohne die von uns getroffene Festlegung, welchen Wert des Arguments wir hier nehmen wollen, wäre die komplexe Logarithmusfunktion nicht eindeutig definiert.

```
clear all
figure(1)
clf()
[x,y]=meshgrid(-2:0.1:2,-2:0.1:2);
f=log(sqrt(x.^2+y.^2));
subplot(1,2,1)
surf(x,y,f)
grid on
xlabel('x')
ylabel('y')
zlabel('Re(log(x+iy))')
[x,y]=meshgrid(-2:0.1:2,0.1:0.1:2);
f=arg(x+i*y);
subplot(1,2,2)
surf(x,y,f)
grid on
xlabel('x')
ylabel('y')
zlabel('Im(log(x+iy))')
hold on
[x,y]=meshgrid(-2:0.1:2,-2:0.1:-0.1);
f=arg(x+i*y);
surf(x,y,f)
[x,y]=meshgrid(0.1:0.1:2,-2:0.1:2);
f=arg(x+i*y);
surf(x,y,f)
```

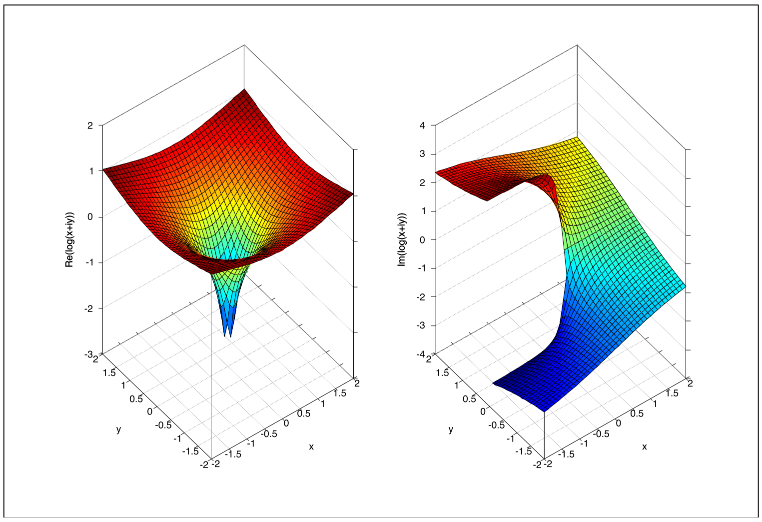


Abbildung 4.13.: Real- und Imaginärteil der komplexen Logarithmusfunktion

Problem 76: *Zeigen Sie, dass die komplexe Logarithmusfunktion für reelle* $z \in]-\infty, 0[$ *nicht stetig ist.*

Beispiel: Wegen $e^{i\pi} = -1$ ist also zum Beispiel $\log(-1) = i\pi$. Wir erinnern daran, dass der reelle Logarithmus von -1 nicht definiert ist.

Satz 190:

Für alle $z \in \mathbb{C}\backslash\{0\}$ gilt
$$\exp(\log(z)) = z.$$

Für alle $z \in \mathbb{C}$ mit $\mathrm{Im}(z) \in]-\pi, \pi]$ gilt
$$\log(\exp(z)) = z.$$

Für $z, w \in \mathbb{C} \setminus \{0\}$ gilt
$$\log(zw) = \log(z) + \log(w) + 2\pi\beta i$$
mit
$$\beta = \begin{cases} -1 & \text{, falls } \arg(z) + \arg(w) > \pi \\ 0 & \text{, falls } -\pi < \arg(z) + \arg(w) \leq \pi \\ 1 & \text{, falls } \arg(z) + \arg(w) \leq -\pi \end{cases}$$

Problem 77: *Beweisen Sie den vorstehenden Satz.*

Wir können nun auch unsere Definition von allgemeinen Potenzen ins Komplexe ausdehnen.

Definition 191: Für $a \in \mathbb{C} \setminus \{0\}$ und $z \in \mathbb{C}$ setzen wir
$$a^z := \exp(z \log(a))$$
und für $z \in \mathbb{C} \setminus \{0\}$
$$0^z := 0.$$

BEMERKUNG: Weiterhin ist 0^0 nicht definiert. In verschiedenen Situationen kann es aber aufgrund von Konvergenzüberlegungen durchaus Sinn machen, 0^0 einen Wert zuzuweisen. Hier ergeben sich aber in verschiedenen Zusammenhängen durchaus unterschiedliche (in dem jeweiligen Zusammenhang!) sinnvolle Definitionen.

Beispiel: Es ist
$$\log(i) = \log(|i|) + i\arg(i) = \log(1) + i\frac{\pi}{2} = i\frac{\pi}{2}$$
und somit
$$i^i = \exp(i\log(i)) = \exp\left(i \cdot i\frac{\pi}{2}\right) = e^{-\frac{\pi}{2}}.$$

Ende des Exkurses

4.5. Der Satz von de l'Hospital

Es folgt ein Satz zur Berechnung von Grenzwerten, welcher auf dem Mittelwertsatz aufbaut.

Satz 192 (Satz von de l'Hospital)**:** Es sei I ein offenes, nichtleeres Intervall und $f, g : I \to \mathbb{R}$ seien differenzierbare Funktionen. Außerdem sei a ein Randpunkt von I, wobei $a = \pm\infty$ zugelassen ist. Für alle $x \in I$, sei $g(x) \neq 0$ und $g'(x) \neq 0$ und es gelte

$$\lim_{x\to a} f(x) = 0 = \lim_{x\to a} g(x)$$

oder

$$\lim_{x\to a} f(x) = \infty = \lim_{x\to a} g(x).$$

Wenn der Grenzwert

$$\lim_{x\to a} \frac{f'(x)}{g'(x)}$$

existiert, dann existiert auch der Grenzwert

$$\lim_{x\to a} \frac{f(x)}{g(x)}$$

und es gilt

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

BEWEIS: Der Beweis erfolgt hier nur für $a \in \mathbb{R}$ und $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$. Ohne Einschränkung sei a der linke Randpunkt von I. Wähle ein $b > a$ mit $b \in I$. Die Funktionen

$$\begin{aligned} F : [a,b] &\to \mathbb{R} \\ x &\mapsto \begin{cases} 0, & \text{falls } x = a \\ f(x), & \text{sonst,} \end{cases} \\ G : [a,b] &\to \mathbb{R} \\ x &\mapsto \begin{cases} 0, & \text{falls } x = a \\ g(x), & \text{sonst,} \end{cases} \end{aligned}$$

sind stetig auf $[a,b]$ und differenzierbar auf $]a,b[$. Nach dem verallgemeinerten Mittelwertsatz 180 existiert ein $\xi_b \in]a,b[$ mit

$$\frac{f'(\xi_b)}{g'(\xi_b)} = \frac{F'(\xi_b)}{G'(\xi_b)} = \frac{F(b) - F(a)}{G(b) - G(a)} = \frac{f(b)}{g(b)}.$$

Offensichtlich gilt $\lim_{b\to a} \xi_b = a$. Da $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ nach Vorraussetzung existiert, gilt

$$\lim_{\xi_b\to a} \frac{f'(\xi_b)}{g'(\xi_b)} = \lim_{b\to a} \frac{f'(b)}{g'(b)}$$

und somit auch

$$\lim_{b\to a} \frac{f(b)}{g(b)} = \lim_{b\to a} \frac{f'(b)}{g'(b)}.$$

□

Da es sich bei dem Punkt a im vorstehenden Satz um einen Randpunkt des Intervalls handelt, sind die Grenzwerte im Satz einseitige Grenzwerte, wenn a eine reelle Zahl ist. Der Satz ist aber natürlich auch anwendbar, wenn wir es mit einer Definitionslücke zu tun haben. Wir halten dies kurz fest.

Korollar 193: Es seien $\alpha, a, \beta \in \mathbb{R}$ mit $\alpha < a < \beta$, $I :=]\alpha, \beta[\setminus \{a\}$ und $f, g : I \to \mathbb{R}$ seien differenzierbare Funktionen. Für alle $x \in I$, sei $g(x) \neq 0$ und $g'(x) \neq 0$ und es gelte

$$\lim_{x\to a} f(x) = 0 = \lim_{x\to a} g(x)$$

oder

$$\lim_{x\to a} f(x) = \infty = \lim_{x\to a} g(x).$$

Wenn der Grenzwert

$$\lim_{x\to a} \frac{f'(x)}{g'(x)}$$

existiert, dann existiert auch der Grenzwert

$$\lim_{x\to a} \frac{f(x)}{g(x)}$$

und es gilt

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

BEWEIS: Man wende den Satz von de l'Hospital einmal mit $I =]\alpha, a[$ und einmal mit $I =]a, \beta[$ an. □

Beispiele:

1) Betrachte den Grenzwert

$$\lim_{x\to\infty} \frac{\log(x)}{\sqrt{x}}$$

auf dem offenen Intervall $\mathbb{R}_+$. Zunächst müssen wir prüfen, ob der Satz von de l'Hospital anwendbar ist, d.h. ob die Voraussetzungen erfüllt sind. Es ist $\lim_{x\to\infty}\log(x) = +\infty = \lim_{x\to\infty}\sqrt{x}$. Für $x > 0$ ist sowohl der Nenner $\sqrt{x} \neq 0$ als auch dessen Ableitung $\frac{1}{2\sqrt{x}} \neq 0$. Somit ist der Satz anwendbar und es gilt

$$\frac{\log'(x)}{(\sqrt{x})'} = \frac{\frac{1}{x}}{\frac{1}{2\sqrt{x}}} \xrightarrow[x\to\infty]{} 0$$

Also folgt $\lim_{x\to\infty}\frac{\log(x)}{\sqrt{x}} = 0$.

2) Betrachte nun den Grenzwert

$$\lim_{x\to 0}\frac{e^x - \sin(x) - 1}{x^2}$$

auf $\mathbb{R}_+$. Wieder müssen wir zunächst die Voraussetzungen prüfen. Es ist $\lim_{x\to 0} e^x - \sin(x) - 1 = 0 = \lim_{x\to 0} x^2$. Für $x \neq 0$ ist sowohl $x^2 \neq 0$ als auch $2x \neq 0$. Der Satz von de l'Hospital ist also anwendbar und durch Differenzieren von Zähler und Nenner erhält man

$$\frac{(e^x - \sin(x) - 1)'}{(x^2)'} = \frac{e^x - \cos(x)}{2x}.$$

Von dem rechten Ausdruck kann man den Grenzwert jedoch noch nicht so einfach ablesen, weil $\lim_{x\to 0} e^x - \cos(x) = \lim_{x\to 0} 2x = 0$ gilt. Darum wendet man den Satz von de l'Hospital nun auf $\lim_{x\to 0}\frac{e^x-\cos(x)}{2x}$ an. Die Voraussetzungen sind weiterhin erfüllt und es ist

$$\frac{(e^x - \cos(x))'}{(2x)'} = \frac{e^x + \sin(x)}{2} \xrightarrow[x\to 0]{} \frac{1}{2}.$$

Damit erhält man also durch zweimaliges Anwenden des Satzes von de l'Hospital

$$\lim_{x\to 0}\frac{e^x - \sin(x) - 1}{x^2} = \frac{1}{2}.$$

3) Als letztes Beispiel betrachten wir

$$\lim_{x\to 0}\frac{1}{\sin x} - \frac{1}{x}.$$

Hier müssen wir den Ausdruck im Grenzwert erst auf die richtige Form bringen. Der Satz von de l'Hospital behandelt die Fälle „$\frac{0}{0}$“ und „$\frac{\infty}{\infty}$“, aber nicht den Fall „$\infty - \infty$“. Es ist aber

$$\frac{1}{\sin x} - \frac{1}{x} = \frac{x - \sin x}{x\sin x}$$

und $\lim_{x\to 0} x - \sin x = 0 = \lim_{x \sin x}$. Außerdem ist in einer Umgebung um Null, die Null nicht enthält, etwa für $x \in]-1, 0[\cup]0, 1[$, sowohl der Nenner $x\sin x \neq 0$ als auch dessen Ableitung $\sin x + x\cos x \neq 0$. Wir können also den Satz von de l'Hospital anwenden und erhalten

$$\frac{(x - \sin x)'}{(x\sin x)'} = \frac{1 - \cos x}{\sin x + x\cos x}.$$

Es ist allerdings wiederum $\lim_{x\to 0} 1 - \cos x = 0 = \lim_{x\to 0} \sin x + x\cos x$. Wir müssen also noch einmal ableiten. In einer Umgebung um Null, die Null nicht enthält, sind der Nenner $\sin x + x\cos x$ und dessen Ableitung nicht Null und wir haben

$$\frac{(1-\cos x)'}{(\sin x + x\cos x)'} = \frac{\sin x}{\cos x + \cos x - x\sin x} = \frac{\sin x}{2\cos x - x\sin x}.$$

Da $\cos 0 = 1$ ist, verschwindet der Nenner in der Nähe von Null nicht und wir haben

$$\lim_{x\to 0} \frac{\sin x}{2\cos x - x\sin x} = 0.$$

Zweimaliges Anwenden des Satzes von de l'Hospital ergibt dann

$$\lim_{x\to 0} \frac{1}{\sin x} - \frac{1}{x} = 0.$$

Problem 78: *Wir setzen $f(x) = x + \sin x\cos x$ und $g(x) = (x + \sin x\cos x)e^{\sin x}$. Untersuchen Sie, ob die folgenden Grenzwerte existieren und berechnen Sie sie gegebenfalls:*

(i) $\lim_{x\to\infty} \frac{f(x)}{g(x)}$,

(ii) $\lim_{x\to\infty} \frac{f'(x)}{g'(x)}$.

Warum darf die Regel von de l'Hospital nicht angewendet werden?

4.6. Das Newton-Verfahren

Wir wollen noch kurz eine Anwendung der Differentialrechnung aus der Numerik skizzieren, das so genannte *Newton-Verfahren* zum Lösen von Gleichungen. Ist f eine differenzierbare Funktion, so interessiert man sich oft, für welche x die Gleichung $f(x) = 0$ erfüllt ist. Die Idee des Verfahrens ist:

Es sei x_0 mit $f'(x_0) \neq 0$ gegeben. Dann kann man f bei x_0 durch die Gerade

$$g(x) := f(x_0) + f'(x_0)(x - x_0)$$

annähern. Ist nun x_1 die Nullstelle dieser Geraden, also

$$x_1 := x_0 - \frac{f(x_0)}{f'(x_0)},$$

so hofft man, dass x_1 näher an einer Nullstelle von f liegt.

Gilt wieder $f'(x_1) \neq 0$, so kann man dieses Vorgehen wiederholen. Man definiert also die Folge (x_n) induktiv durch

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)},$$

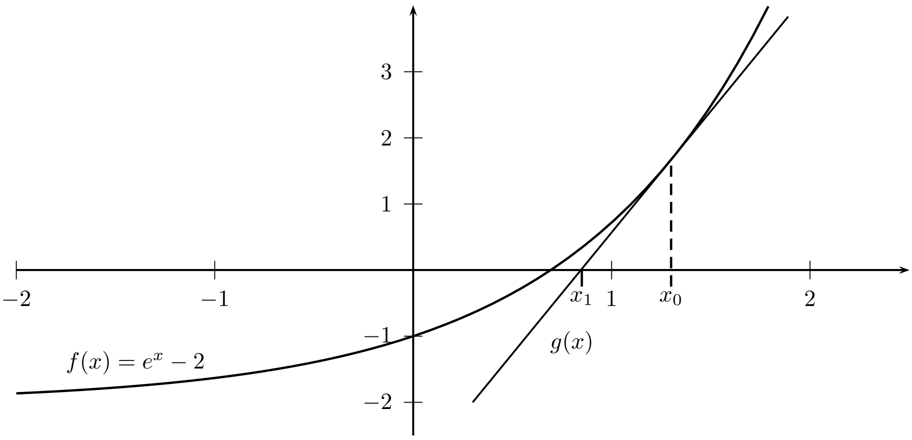


Abbildung 4.14.: Idee des Newton-Verfahrens

wobei $f'(x_n) \neq 0$ für alle n gelten muss. Falls dieses Verfahren gegen einen Grenzwert $\overline{x}$ konvergiert, so gilt

$$\overline{x} = \overline{x} - \frac{f(\overline{x})}{f'(\overline{x})},$$

falls auch f' stetig ist und $f'(\overline{x}) \neq 0$ gilt. Insbesondere folgt aus dieser Gleichung $f(\overline{x}) = 0$. Im Allgemeinen braucht das Verfahren aber nicht zu konvergieren. Man kann aber zeigen, dass es stets konvergent ist, wenn zum Beispiel $f : [a, b] \to \mathbb{R}$ zweimal differenzierbar mit $f'(x_0) \neq 0$ und $f''(x) \geq 0$ (oder $f''(x) \leq 0$) für alle $x \in [a, b]$ ist und außerdem noch $f(a) > 0$, $f(b) < 0$ und $x_0 \in [a, b]$ gilt.

Beispiel: Die Folgen in den Sätzen 56 und 55 sind Beispiele für das Newton-Verfahren. Wir wollen hier noch ein konkretes Beispiel geben und die Anwendung des gerade diskutierten Newton-Verfahren zur Berechnung von $\sqrt[3]{2}$ erklären.

Nach Definition ist $\sqrt[3]{2}$ die Lösung der Gleichung $x^3 - 2 = 0$. Setze $f(x) = x^3 - 2$ dann gilt

$$\begin{aligned} f'(x) &= 3x^2 \neq 0 \text{ für } x > 0, \\ f^{(2)}(x) &= 6x > 0 \text{ für } x > 0, \\ f(1) &= -1 \text{ und} \\ f(2) &= 6. \end{aligned}$$

Wähle $x_0 = 2$, dann konvergiert

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^3 - 2}{3x_n^2}$$

gegen den gesuchten Wert.

Eine Implementierung des Newton-Verfahrens in Matlab könnte wie folgt aussehen.

```
function [erg a]= newton(f,start,N)
if nargin==2
  N=100;
end

a(1)=start;
for i=2:N
  a(i)=a(i-1)-f(a(i-1))/ableitung2(f,a(i-1));
end
erg=a(N);
```

Wir wenden diese Funktion nun auf das obige Beispiel an.

```
f= @(x) x^3-2;
newton(f,3,100)
```

4.7. Extrema reellwertiger Funktionen mehrerer Veränderlicher

In diesem Kapitel werden Funktionen $f : D \to \mathbb{R}$ betrachtet, wobei $D \subset \mathbb{R}^n$ ist. Immer wenn das Format der betrachteten Vektoren eine Rolle spielt, identifizieren wir $v \in \mathbb{R}^n$ mit einer $n \times 1$-Matrix, d.h. wir fassen den Vektor als Spaltenvektor auf. v^t ist dann ein Zeilenvektor.

Definition 194 (Norm auf $\mathbb{R}^n$)**:** Für $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ definiere durch

$$\|v\|_2 := \sqrt{\sum_{i=1}^{n} v_i^2}$$

eine Norm von $\mathbb{R}^n$.

BEMERKUNG: Eine komplexe Zahl $z = a + ib$ kann man mit dem Vektor $v = (a, b) \in \mathbb{R}^2$ identifizieren. Es ist

$$\|v\|_2 = |z|.$$

Definition 195 (Lokales Maximum bzw. lokales Minimum)**:** Es sei $D \subset \mathbb{R}^n$ und $f : D \to \mathbb{R}$ eine Funktion. Ein Punkt $a \in D$ heißt *lokales Maximum (bzw. lokales Minimum) von* f, wenn es ein $\varepsilon > 0$ gibt, so dass

$$U_\varepsilon(a) := \{x \in \mathbb{R}^n | \|a - x\|_2 < \varepsilon\} \subset D$$

und für alle $x \in U_\varepsilon(a)$ die Ungleichung

$$f(x) \leq f(a) \qquad \text{(bzw. } f(x) \geq f(a)\text{)}$$

gilt.

BEMERKUNG: Für $n = 1$ stimmt diese Definition mit der Definition 176 überein. Auch die Begriffe isoliertes, lokales Maximum beziehungsweise Minimum und globales Maximum beziehungsweise Minimum übertragen sich analog.

Definition 196 (Offene Menge)**:** Es sei $D \subset \mathbb{R}^n$. Ein Punkt $a \in D$ heißt *innerer Punkt von* D, wenn es ein $\varepsilon > 0$ gibt, so dass

$$U_\varepsilon(a) := \{x \in \mathbb{R}^n | \|a - x\|_2 < \varepsilon\} \subset D$$

ist. Eine Menge D heißt genau dann *offen*, wenn alle $a \in D$ innere Punkte sind.

Beispiel: Als nächstes werden Funktionen mehrer Veränderlicher auf Extrema untersucht. Man betrachte die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, die gegeben ist durch $f(x, y) := x^2 + y^2$. Angenommen der Punkt (x_0, y_0) ist ein Extremum von f ist. Betrachte die beiden Funktionen *einer* Veränderlichen $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ mit $f_1(x) = x^2 + y_0^2$ und $f_2(y) = x_0^2 + y^2$. Diese haben die Ableitungen

$$f_1'(x) = 2x$$
$$f_2'(y) = 2y$$

Ist (x_0, y_0) ein Extremum von f, so ist x_0 ein Extremum von f_1 und y_0 ein solches von f_2. Also gilt $f_1'(x_0) = 2x_0 = 0$ und $f_2'(y_0) = 2y_0 = 0$ nach Satz 170. Setzt man dies gleich Null, so sieht man dass $(0, 0)$ der einzige Kandidat für ein Extremum von f ist.

Definition 197 (Partielle Ableitung)**:** Es sei $D \subset \mathbb{R}^n$ offen und $a = (a_1, \ldots, a_n) \in D$. Die Funktion

$$f : D \to \mathbb{R}$$
$$(x_1, \ldots, x_n) = x \mapsto f(x_1, \ldots, x_n)$$

heißt in a *partiell differenzierbar bzgl.* x_i, falls die durch

$$f_i(x_i) := f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$$

definierte Funktion einer Veränderlichen in a_i differenzierbar ist. Für die partielle Ableitung nach der i-ten Komponente in dem Punkt a schreibt man

$$\frac{\partial f}{\partial x_i}(a) := f_i'(a_i).$$

Ist f in allen Punkten von D nach x_i partiell differenzierbar, so heißt die Funktion

$$\frac{\partial f}{\partial x_i} : D \to \mathbb{R}$$

die *partielle Ableitung von f nach x_i*. Die Funktion f heißt auf D *partiell differenzierbar*, falls $\frac{\partial f}{\partial x_i}(a)$ für alle $a \in D$ und alle i existiert. Sind außerdem noch alle partielle Ableitungen stetig, so heißt f *stetig partiell differenzierbar*.

BEMERKUNG: Wenn wir also eine Funktion f in der Variablen $x_1, x_2, \dots, x_n$ nach der Variablen x_i partiell ableiten wollen, so behandeln wir die Funktion f so, als wenn sie nur von der einen Variablen x_i abhängt und als wenn die anderen Variablen $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ Konstanten wären.

Beispiel: Es sei $f : \mathbb{R}^3 \to \mathbb{R}$ gegeben durch $f(x, y, z) := xy^2 + x^3y - xz + yz$. Dann ist

$$\begin{aligned}
\frac{\partial f}{\partial x}(x, y, z) &= y^2 + 3x^2y - z, \\
\frac{\partial f}{\partial y}(x, y, z) &= 2xy + x^3 + z, \\
\frac{\partial f}{\partial z}(x, y, z) &= -x + y.
\end{aligned}$$

Definition 198 (Gradient)**:** Es sei $f : \mathbb{R}^n \to \mathbb{R}$ eine partiell differenzierbare Funktion. Dann heißt der Vektor

$$\nabla f(a) := \operatorname{grad}(f)(a) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}$$

der *Gradient von f an der Stelle a*.

Geometrische Interpretation: Wir haben gesehen, dass bei einer Funktion $f : I \to \mathbb{R},\ I \subset \mathbb{R}$, die Ableitung $f'(x_0)$ die Steigung der Tangente an den Graphen von f im Punkt x_0 ist. Ist f nun eine Funktion, die auf einer Teilmenge $D \subset \mathbb{R}^2$ definiert ist, also $f : D \to \mathbb{R}$, so besitzt der Graph von f zu jedem Punkt $a := (x_0, y_0)$ eine Ebene $T_a(x, y)$, die den Graphen von f in dem Punkt a schneidet und die Steigung des Graphen in dem

Punkt a beschreibt. Wir nennen diese Ebene die *Tangentialebene an f im Punkt a*. Die Tangentialebene ist gegeben durch

$$T_a(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)\cdot(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)\cdot(y-y_0).$$

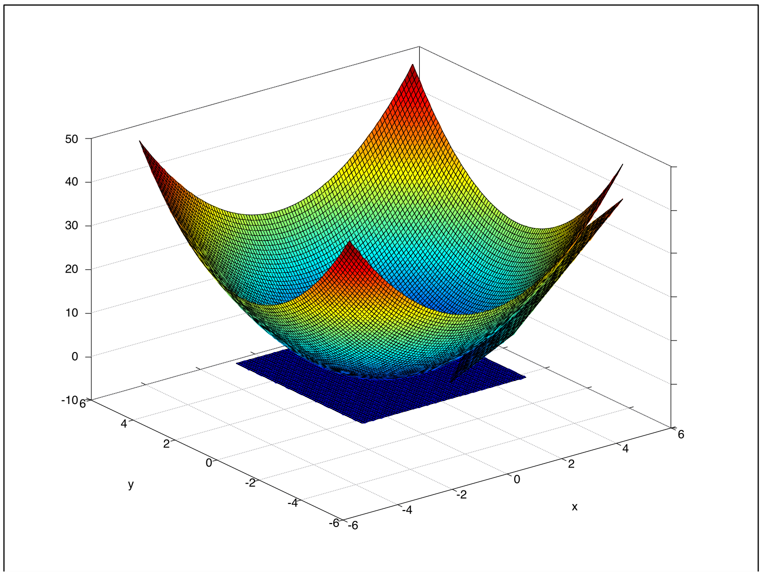


Abbildung 4.15.: Tangentialebenen an den Graphen der Funktion $f(x,y) = x^2 + y^2$ in den Punkten $(0,0)$ und $(3,-3)$

Dies lässt sich noch weiter auf Funktionen beliebig vieler Veränderlicher verallgemeinern.

Betrachten wir eine Funktion $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$, so existiert zu jedem Punkt $a \in D$ eine Hyperebene $T_a(x)$, die den Graphen von f in dem Punkt a schneidet und den Graphen von f in einer Umgebung des Punktes a näherungsweise beschreibt. Mit dem Standardskalarprodukt $\langle x,y\rangle := \sum_{\nu=1}^{n} x_\nu y_\nu$ auf dem $\mathbb{R}^n$ lässt sich diese Tangentialebene an f in a als

$$T_a(x) = f(a) + \langle \mathrm{grad}(f)(a), x-a\rangle$$

schreiben.

Der Plot der Tangentialebenen an den Graphen der Funktion $f(x,y) = x^2 + y^2$ in den Punkten $(0,0)$ und $(3,-3)$ wurde mit dem folgenden Skript realisiert:

```
figure(1)
clf()
[x,y]=meshgrid(-5:0.1:5,-5:0.1:5);
f=x.^2+y.^2;
surf(x,y,f)
hold on
%Tangentialebene in (3,-3)
[x,y]=meshgrid(1:0.1:5,-5:0.1:-1);
t=(3)^2+(-3)^2+6*(x-3)-6*(y+3);
```

```
surf(x,y,t)
%Tangentialebene in (0,0)
[x,y]=meshgrid(-3:0.1:3,-3:0.1:3);
t=0*x+0*y;
surf(x,y,t)
xlabel('x')
ylabel('y')
hold off
```

Satz 199: Es sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ partiell differenzierbar. Hat f in einem (inneren) Punkt $a \in D$ ein lokales Extremum, so gilt

$$\operatorname{grad}(f)(a) = 0 \in \mathbb{R}^n.$$

BEWEIS: Klar. □

Definition 200: Es sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ eine partiell differenzierbare Funktion. Ein Punkt $a \in D$ heißt *kritischer Punkt von* f, wenn $\operatorname{grad}(f)(a) = 0 \in \mathbb{R}^n$ ist.

Beispiel: Betrachte die Funktion

$$\begin{aligned} f : \mathbb{R}^4 &\to \mathbb{R} \\ x &\mapsto x_1 x_2^2 + x_3 x_1^2 - x_4 x_2. \end{aligned}$$

Die partiellen Ableitungen lauten:

$$\frac{\partial f}{\partial x_1}(x) = x_2^2 + 2x_3x_1 \qquad \frac{\partial f}{\partial x_2}(x) = 2x_1x_2 - x_4 \qquad \frac{\partial f}{\partial x_3}(x) = x_1^2 \qquad \frac{\partial f}{\partial x_4}(x) = -x_2.$$

Nullsetzen des Gradienten

$$0 \overset{!}{=} \operatorname{grad}(f)(x) = \begin{pmatrix} x_2^2 + 2x_3x_1 \\ 2x_1x_2 - x_4 \\ x_1^2 \\ -x_2 \end{pmatrix}$$

liefert $x_1 = x_2 = x_4 = 0$. Die kritischen Punkte von f sind also alle Punkte $(0,0,x_3,0)^t$ mit $x_3 \in \mathbb{R}$ beliebig.

Um zu sehen, ob in einem kritischen Punkt tatsächlich ein Maximum oder ein Minimum vorliegt, hat man in der Theorie einer Veränderlichen die nächste Ableitung betrachtet. Dieses Konzept wird nun für Funktionen mehrer Veränderliche eingeführt.

Definition 201: Es sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ stetig partiell differenzierbar. Sind alle partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ partiell differenzierbar, so heißt f *zweimal partiell differenzierbar*.
Induktiv: Ist f k-mal stetig partiell differenzierbar und sind alle partiellen Ableitungen k-ter Ordnung

$$\underbrace{\frac{\partial}{\partial x_{i_k}}\left(\frac{\partial}{\partial x_{i_{k-1}}}\cdots\left(\frac{\partial f}{\partial x_{i_1}}(a)\right)\right)}_{k-\text{mal}} \qquad i_1,\ldots,i_k \in \{1,\ldots,n\}$$

partiell differenzierbar, so heißt f $(k+1)$-mal partiell differenzierbar. Statt

$$\frac{\partial}{\partial x_{i_k}}\left(\frac{\partial}{\partial x_{i_{k-1}}}\cdots\left(\frac{\partial f}{\partial x_{i_1}}(a)\right)\right)$$

schreibt man

$$\frac{\partial^k f}{\partial x_{i_k}\partial x_{i_{k-1}}\cdots\partial x_{i_1}}(a).$$

Satz 202 (von Schwarz)**:** Ist f zweimal stetig partiell differenzierbar, so gilt für alle inneren Punkte $a \in D$ und alle $i, j \in \{1,\ldots,n\}$

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$$

BEMERKUNG: Auf Grund dieses Satzes schreibt man zum Beispiel statt $\frac{\partial^3}{\partial x_1 \partial x_2 \partial x_1}$ oftmals $\frac{\partial^3}{\partial x_2 \partial x_1^2}$.

BEWEIS: In [Sch94] findet man einen Beweis, der im wesentlichen nur den hier behandelten Mittelwertsatz der Differentialrechnung benutzt. □

Bei der Untersuchung der kritischen Punkte von Funktionen einer Veränderlicher hat sich der Satz von Taylor als hilfreich erwiesen. Wir wollen hier keine Version des Satzes von Taylor für Funktionen mehrerer Veränderlicher behandeln. Um das Vorgehen plausibel zu machen, reicht es aber auch aus, wenn wir den Satz von Taylor für Funktionen einer Veränderlichen wiederholt anwenden.

Nehmen wir also an, dass (x_0, y_0) ein kritischer Punkt einer Funktion $f : D \to \mathbb{R}$ mit

$A \subset \mathbb{R}^2$ ist. Nun will man f in der Nähe von (x_0, y_0) untersuchen. Es seien $h, k \in \mathbb{R}$ "klein". Mit THO wird im Folgenden *Terme höherer Ordnung* abgekürzt. Diese können später für kleine h, k vernachlässigt werden. Der Satz von Taylor angewendet auf die Veränderliche x liefert

$$\begin{aligned} f(x_0 + h, y_0 + k) = f(x_0, y_0 + k) &+ \frac{\partial f}{\partial x}(x_0, y_0 + k) \cdot h \\ &+ \frac{1}{2} \cdot \frac{\partial^2 f}{\partial x^2}(x_0, y_0 + k) \cdot h^2 + \mathrm{THO}\,. \end{aligned}$$

Wendet man den Satz von Taylor nun auf die zweite Komponente an, erhällt man

$$\begin{aligned} f(x_0, y_0 + k) &= f(x_0, y_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot k + \frac{1}{2} \cdot \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \cdot k^2 + \mathrm{THO} \\ \frac{\partial f}{\partial x}(x_0, y_0 + k) &= \frac{\partial f}{\partial x}(x_0, y_0) + \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) \cdot k + \mathrm{THO} \\ \frac{\partial^2 f}{\partial x^2}(x_0, y_0 + k) &= \frac{\partial^2 f}{\partial x^2}(x_0, y_0) + \mathrm{THO}\,. \end{aligned}$$

Damit folgt

$$\begin{aligned} f(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \cdot h + \frac{\partial f}{\partial y}(x_0, y_0) \cdot k + \frac{1}{2} \cdot \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \cdot h^2 \\ + \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) \cdot h \cdot k + \frac{1}{2} \cdot \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \cdot k^2 + \mathrm{THO}\,. \end{aligned}$$

Da (x_0, y_0) ein kritischer Punkt ist, ist

$$\mathrm{grad}(f)(x_0, y_0) = \begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix} = 0$$

und wir haben

$$\begin{aligned} f(x_0 + h, y_0 + k) = f(x_0, y_0) &+ \frac{1}{2} \cdot \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \cdot h^2 \\ &+ \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) \cdot h \cdot k + \frac{1}{2} \cdot \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \cdot k^2 + \mathrm{THO}\,. \end{aligned}$$

Dies wird übersichtlicher, wenn man die aus der linearen Algebra bekannten Matrixschreibweise benutzt. Setzt man

$$H_f(x_0, y_0) := \begin{pmatrix} \frac{\partial^2}{\partial x^2} f(x_0, y_0) & \frac{\partial^2}{\partial x \partial y} f(x_0, y_0) \\ \frac{\partial^2 f}{\partial y \partial x} f(x_0, y_0) & \frac{\partial^2}{\partial y^2} f(x_0, y_0) \end{pmatrix}$$

und

$$v := \begin{pmatrix} h \\ k \end{pmatrix},$$

so gilt

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{1}{2} v^t \cdot H_f(x_0, y_0) \cdot v + \mathrm{THO}\,.$$

Man kann also an

$$v^t \cdot H_f(x_0, y_0) \cdot v$$

ablesen, ob man ein Maximum, Minimum oder gar kein Extremum vorliegt.

Dieses Vorgehen lässt sich auf Funktionen in mehr als zwei Veränderliche verallgemeinern. Dies ist nicht wirklich schwerer, aber wie man sich leicht ausmalen kann, noch unübersichtlicher.

Definition 203 (Hesse-Matrix)**:** Es sei $D \subset \mathbb{R}^n$ offen, $a \in D$ und $f : D \to \mathbb{R}$ zweimal stetig partiell differenzierbar. Dann heißt die Matrix

$$H_f(a) := \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a)\right)_{1 \leq i,j \leq n} \in M(n \times n, \mathbb{R})$$

die *Hesse-Matrix von f an der Stelle a*.

BEMERKUNG: Wegen des Satzes 202 von Schwarz ist die $H_f(a)$ eine symmetrische Matrix.

Beispiel: Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $x \mapsto x_1^2 + x_2^2$. Die partiellen Ableitungen lauten

$$\frac{\partial f}{\partial x_1}(x) = 2x_1$$
$$\frac{\partial f}{\partial x_2}(x) = 2x_2.$$

Dann lautet die Hesse-Matrix im Punkt 0

$$H_f(0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(0) & \frac{\partial^2 f}{\partial x_2^2}(0) \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Exkurs: Einige Ergebnisse aus der Linearen Algebra über symmetrische Matrizen

Die Beweise der Sätze in diesem Exkurs findet man zum Beispiel in [Fis03].

Definition 204 (Definite Matrix)**:** Es sei $M \in M(n \times n, \mathbb{R})$ eine symmetrische Matrix.

(i) Die Matrix M heißt *positiv definit*, falls für alle $v \in \mathbb{R}^n \setminus \{0\}$

$$\mathbb{R} \ni v^t M v = (v_1, \ldots, v_n) M \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} > 0$$

gilt.

(ii) Die Matrix M heißt *positiv semidefinit*, falls für alle $v \in \mathbb{R}^n \setminus \{0\}$

$$v^t M v \geq 0$$

gilt.

(iii) Die Matrix M heißt *negativ (semi)definit*, falls $-1 \cdot M$ positiv (semi)definit ist.

(iv) Die Matrix M heißt *indefinit*, falls es $v, w \in \mathbb{R}^n$ mit $v^t M v > 0$ und $w^t M w < 0$ gibt.

Beispiel: Ist M eine Diagonalmatrix, etwa

$$M = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

dann gilt

$$v^t M v = v^t \cdot \begin{pmatrix} \lambda_1 v_1 \\ \vdots \\ \lambda_n v_n \end{pmatrix} = \sum_{i=1}^{n} \lambda_i v_i^2.$$

Sind $e_1, \ldots, e_n$ die Vektoren der Standardbasis des $\mathbb{R}^n$, so ist $Me_i = \lambda_i e_i$. Die Matrix M ist also genau dann positiv definit, wenn die Einträge $\lambda_1, \ldots, \lambda_n$ positiv sind.

Definition 205 (Eigenwert, Eigenvektor)**:** Es sei $A \in M(n \times n, \mathbb{K})$. Eine Zahl $\lambda \in \mathbb{K}$ heißt *Eigenwert der Matrix* A, wenn es einen Vektor $v \in \mathbb{K}^n$ mit $v \neq 0$ gibt, so dass

$$Av = \lambda v$$

ist. Der Vektor v ist dann *ein Eigenvektor von* A *zum Eigenwert* λ.

BEMERKUNG: Man kann zeigen, dass eine symmetrische Matrix $M \in M(n \times n, \mathbb{R})$ stets n linear unabhängige Eigenvektoren besitzt und alle Eigenwerte reell sind.

Satz 206: Es sei $M \in M(n \times n, \mathbb{R})$ symmetrisch. Dann gilt

(i) Die Matrix M ist positiv definit genau dann, wenn alle Eigenwerte von M positiv sind.

(ii) Die Matrix M ist positiv semidefinit genau dann, wenn alle Eigenwerte von M nicht negativ sind.

(iii) Die Matrix M ist negativ definit genau dann, wenn alle Eigenwerte von M negativ sind.

(iv) Die Matrix M ist negativ semidefinit genau dann, wenn alle Eigenwerte von M nicht positiv sind.

Für 2×2 Matrizen gilt:

Satz 207: Es sei

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix} \in M(2 \times 2, \mathbb{R}).$$

Dann gilt:

(i) Die Matrix A ist positiv definit genau dann, wenn $ad - b^2 > 0$ und $a > 0$ gilt.

(ii) Die Matrix A ist negativ definit genau dann, wenn $ad - b^2 > 0$ und $a < 0$ gilt.

(iii) Die Matrix A ist indefinit genau dann, wenn $ad - b^2 < 0$ gilt.

Ende des Exkurses

Nun können wir das Hauptresultat des Kapitels formulieren und anwenden.

Satz 208: Es sei $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}$ zweimal stetig differenzierbar und $a \in D$ ein kritischer Punkt von f, das heißt

$$\operatorname{grad}(f)(a) = 0.$$

Dann gilt:

(i) Ist die Hesse-Matrix $H_f(a)$ positiv definit, so hat f in a ein isoliertes, lokales Minimum.

(ii) Ist $H_f(a)$ negativ definit, so hat f in a ein isoliertes, lokales Maximum.

(iii) Ist $H_f(a)$ indefinit, so hat f in a kein lokales Extremum. Man sagt in diesem Fall, dass a ein Sattelpunkt von f ist.

Korollar 209: Es sei $D \subset \mathbb{R}^2$ offen, $f : D \to \mathbb{R}$ zweimal stetig differenzierbar und $a \in D$ ein kritischer Punkt von f. Es sei

$$H_f(a) = \begin{pmatrix} \alpha & \beta \\ \beta & \delta \end{pmatrix} \in M(2 \times 2, \mathbb{R})$$

die Hesse-Matrix von f in a. Dann gilt:

(i) Ist $\alpha\delta - \beta^2 > 0$ und $\alpha > 0$, so hat f in a ein isoliertes, lokales Minimum.

(ii) Ist $\alpha\delta - \beta^2 > 0$ und $\alpha < 0$, so hat f in a ein isoliertes, lokales Maximum.

(iii) Ist $\alpha\delta - \beta^2 < 0$, so hat f in a kein lokales Extremum. Man sagt in diesem Fall, dass a ein Sattelpunkt von f ist.

Beispiele:

1) Man betrachte die Funktion

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto x^2 - y^2 \end{aligned}$$

mit den partiellen Ableitungen

$$\begin{aligned} 0 &\stackrel{!}{=} \frac{\partial f}{\partial x}(x, y) = 2x \\ 0 &\stackrel{!}{=} \frac{\partial f}{\partial y}(x, y) = -2y. \end{aligned}$$

Also ist $(0, 0)$ der einzige kritische Punkt von f. Die Hessematrix an diesem Punkt lautet

$$H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Auf Grund der dritten Aussage von Satz 207 ist $H_f(0, 0)$ indefinit und f hat somit in $(0, 0)$ kein lokales Extremum.

2) Betrachte nun die Funktion $f(x, y) = x^2 + y^2 - e^{xy}$ und ihre partiellen Ableitungen

$$\begin{aligned} &\begin{cases} 0 \stackrel{!}{=} \frac{\partial f}{\partial x}(x, y) = 2x - ye^{xy} \mid \cdot x & \Rightarrow 2x^2 - xye^{xy} = 0 \\ 0 \stackrel{!}{=} \frac{\partial f}{\partial y}(x, y) = 2y - xe^{xy} \mid \cdot y & \Rightarrow 2y^2 - xye^{xy} = 0 \end{cases} \\ \Rightarrow &(x - y)(x + y) = x^2 - y^2 = 0 \end{aligned}$$

1. Fall ($x = y$): So erhält man aus der ersten Gleichung

$$0 = 2x - xe^{x^2} = x\left(2 - e^{x^2}\right).$$

Also ist $x = y = 0$ oder $2 - e^{x^2} = 0 \Leftrightarrow e^{x^2} = 2 \Leftrightarrow x^2 = \log(2)$. Für diesen Fall erhalten wir die kritischen Punkte:

$$(0,0), \quad \left(\sqrt{\log(2)}, \sqrt{\log(2)}\right), \quad \left(-\sqrt{\log(2)}, -\sqrt{\log(2)}\right).$$

2. Fall ($x = -y$): Die zweite Gleichung wird zu

$$0 = 2y + ye^{-y^2} = y\left(2 + e^{-y^2}\right).$$

Also ist $y = 0$ oder $2 + e^{-y^2} = 0 \Leftrightarrow e^{-y^2} = -2$. Jedoch hat die Gleichung $e^{-y^2} = -2$ in $\mathbb{R}$ keine Lösung. Also erhällt man im zweiten Fall keine neuen, kritischen Punkte.
Die Hesse-Matrix lautet

$$H_f(x,y) = \begin{pmatrix} 2 - y^2e^{xy} & -e^{xy} - yxe^{xy} \\ -e^{xy} - xye^{xy} & 2 - x^2e^{xy} \end{pmatrix}.$$

Setzt man die kritischen Punkte ein und wendet Satz 207 an, so folgt

$$\det(H_f(0,0)) = 3 \text{ und } \frac{\partial^2 f}{\partial x^2}(0,0) = 2 > 0$$

und

$$\begin{aligned}\det\left(H_f\left(\sqrt{\log(2)}, \sqrt{\log(2)}\right)\right) &= \det\left(H_f\left(-\sqrt{\log(2)}, -\sqrt{\log(2)}\right)\right) \\ &= -16\log(2) < 0.\end{aligned}$$

Also ist $(0,0)$ ein Minimum und die anderen beiden Punkte sind keine Extrema.

Problem 79 (Methode der kleinsten Quadrate; Ausgleichsgerade)**:** *(i) Bestimmen Sie zu n gegebenen Messwertpaaren $(x_1, y_1), \ldots, (x_n, y_n)$ diejenige $a, b \in \mathbb{R}$, für die die Funktion F mit*

$$F(a,b) = \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

minimal wird. Man bezeichnet dann die Gerade $y = a + bx$ als Ausgleichsgerade *zu den Messwertpaaren $(x_1, y_1), \ldots, (x_n, y_n)$.*

(ii) Ein Fallschirmspringer öffnet seinen Fallschirm und misst mit Hilfe eines Höhenmessers zu verschiedenen Zeitpunkten nach dem Öffnen des Schirms seine Höhe über dem Erdboden. Die Messung ergab folgende Werte:

Fallzeit t in Sekunden	*5*	*10*	*15*	*20*	*25*
Höhe h in Metern	*230*	*210*	*190*	*170*	*140*

Bestimmen Sie die Ausgleichsgerade der Form $f(x) = a + bx$. Welche Höhe würden Sie nach dem Modell nach einer Fallzeit von 30 Sekunden erwarten? Nach wieviel Sekunden erreicht der Springer nach dem Modell den Erdboden?

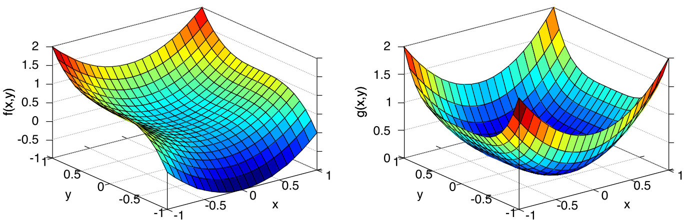


Abbildung 4.16.: Graphen der Funktion $f(x,y) = x^2 + y^3$ und $g(x,y) = x^2 + y^4$ in einer Umgebung des Ursprungs.

Problem 80: *Zeigen Sie anhand der Funktionen $f, g : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) := x^2 + y^3$ und $g(x,y) := x^2 + y^4$, dass man aus einer positiv semidefiniten Hessematrix nicht auf das Vorliegen eines Extremums schließen kann.*

4.8. Stammfunktionen

In diesem Abschnit sei $I \subset \mathbb{R}$ stets ein offenes, nicht leeres Intervall.

Definition 210 (Stammfunktion)**:** Es sei $f : I \to \mathbb{R}$ eine Funktion. Eine differenzierbare Funktion $F : I \to \mathbb{R}$ heißt *Stammfunktion von f* oder *unbestimmtes Integral von f*, wenn für alle $x \in I$

$$F'(x) = f(x)$$

gilt. Mit $\int f(x)dx$ bezeichnen wir eine (**nicht eindeutig bestimmte**) Stammfunktion von f.

BEMERKUNG: Wenn wir etwas schreiben wie $\int f(x)\,dx = F(x)$, so ist dies als „eine Stammfunktion von f ist gegeben durch die Funktion F“ bzw. „F ist eine Stammfunktion von f“zu lesen. Dies ist keine herkömmliche Gleichheit und damit ist diese Schreibweise nicht ganz unproblematisch. Man sieht zum Beispiel schnell, dass sowohl

$$\int 1\, dx = x$$

als auch

$$\int 1\, dx = x + 42$$

gilt. Offensichtlich kann man diese beiden „Gleichungen“ nun aber nicht gleich setzen und $x = x + 42$ folgern.

Der folgende Satz zeigt, dass sich zwei Stammfunktion einer gegebenen Funktion f allerdings höchstens durch eine additive Konstante unterscheiden.

Satz 211: Sind F und G zwei Stammfunktionen von f auf I, so existiert eine Konstante $c \in \mathbb{R}$, so dass für alle $x \in I$

$$G(x) = F(x) + c$$

gilt.

BEWEIS: Setze $\varphi(x) = F(x) - G(x)$. Für alle $x \in I$ gilt dann

$$\varphi'(x) = F'(x) - G'(x) = f(x) - f(x) = 0.$$

Also ist φ auf I konstant. □

BEMERKUNG: Gilt $\int f(x)\, dx = F(x)$, so gilt also auch $\int f(x)\, dx = F(x) + c$ für alle $c \in \mathbb{R}$. Und aus den Beziehungen $\int f(x)\, dx = F(x)$ und $\int f(x) = G(x)$ folgt, dass es ein $c \in \mathbb{R}$ gibt, so dass $F(x) = G(x) + c$ ist.

Aus jeder Differentationsregel erhält man eine Regel zum Bilden der Stammfunktion.

Beispiele:

1) $\displaystyle\int c\, dx = cx$ auf ganz $\mathbb{R}$,

2) $\displaystyle\int x^\alpha\, dx = \frac{x^{\alpha+1}}{\alpha+1}$ auf $\begin{cases} \mathbb{R} & \text{für } \alpha \in \mathbb{N}, \\ ]-\infty, 0[\wedge]0, \infty[& \text{für } \alpha = -2, -3, \ldots, \\ ]0, \infty[& \text{für } \alpha \in \mathbb{R}\backslash\{-1\}, \end{cases}$

3) $\displaystyle\int x^{-1}\, dx = \log(|x|)$ auf $]-\infty, 0[$ und auf $]0, \infty[$,

4) $\displaystyle\int e^x\, dx = e^x$ auf $\mathbb{R}$,

5) $\displaystyle\int \cos(x)\, dx = \sin(x)$ und $\displaystyle\int \sin(x)\, dx = -\cos(x)$ auf $\mathbb{R}$.

Satz 212: Es seien $f, g : I \to \mathbb{R}$ und F eine Stammfunktion von f. Dann gilt

(i) Sind $a, b \in \mathbb{R}$ und besitzt g eine Stammfunktion, so gilt

$$\int af(x) + bg(x)\, dx = a \int f(x)\, dx + b \int g(x)\, dx.$$

(ii) (Partielle Integration) Ist g differenzierbar und besitzt Fg' eine Stammfunktion, so gilt

$$\int f(x)g(x)\, dx = F(x)g(x) - \int F(x)g'(x)\, dx.$$

(iii) (Substitionsregel) Ist g differenzierbar, so gilt

$$\int f(g(x))g'(x)\, dx = F(g(x)).$$

BEWEIS: Die Aussagen folgen sofort aus den entsprechenden Differentationsregeln. Aussage (i) folgt aus der Linearität der Ableitung und Aussage (ii) aus der Produktregel. Aussage (iii) erhält man auf Grund der Kettenregel. Wegen

$$(F(g(x)))' = F'(g(x))g'(x) = f(g(x))g'(x).$$

ist also $F(g(x))$ eine Stammfunktion von $f(g(x))g'(x)$. □

Beispiele:

1) Ist $g(x) \neq 0$ für alle $x \in I$ und g differenzierbar auf I. Dann gilt

$$\int \frac{g'(x)}{g(x)}\, dx = \log(|g(x)|).$$

2) Wir haben in Aufgabe 75 die Tangensfunktion eingeführt. Mithilfe des ersten Beispiel können wir eine Stammfunktion dieser Funktion bestimmen. Es ist $\tan(x) = \sin(x)/\cos(x)$ und $\cos'(x) = -\sin(x)$. Somit haben wir für alle $x \in \mathbb{R}$ mit $\cos(x) \neq 0$

$$\int \tan(x)\, dx = \int \frac{\sin(x)}{\cos(x)}\, dx = -\int \frac{(\cos(x))'}{\cos(x)}\, dx = -\log(|\cos(x)|).$$

3) Die Stammfunktion von $\cos^2$.

$$\int \cos^2(x)\, dx = \sin(x)\cos(x) - \int -\sin^2(x)\, dx \quad \text{und}$$
$$\int \sin^2(x)\, dx = \int 1 - \cos^2(x)\, dx = x - \int \cos^2(x)\, dx$$

Zusammen ergibt dies

$$2\int \cos^2(x)\, dx = \sin(x)\cos(x) + x$$

beziehungsweise

$$\int \cos^2(x)\, dx = \frac{1}{2}\left(\sin(x)\cos(x) + x\right).$$

4) *Partialbruchzerlegung:* Mit Hilfe der Partialbruchzerlegung lassen sich Stammfunktionen zu rationalen Funktionen $\frac{Q(x)}{P(x)}$ bestimmen, wobei P und Q Polynome in x sind.
Wir betrachten zunächst als Beispiel $\frac{1}{P(x)}$, wobei P die Polynomfunktion $P(x) = 1 - x^2 = (1-x)(1+x)$ sei. Nun werden Zahlen $A_1, A_2 \in \mathbb{R}$ gesucht, so dass die Gleichung

$$\frac{1}{1-x^2} = \frac{A_1}{1-x} + \frac{A_2}{1+x}$$

erfüllt ist. Die beiden Brüche $A_1/(1-x)$ und $A_2/(1+x)$ werden *Partialbrüche* genannt. Es gilt

$$\begin{aligned}\frac{1}{1-x^2} &= \frac{A_1}{1-x} + \frac{A_2}{1+x}\\ &= \frac{A_1 + A_1x + A_2 - A_2x}{1-x^2}\\ &= \frac{(A_1+A_2) + (A_1-A_2)x}{1-x^2}.\end{aligned}$$

Damit die Zähler übereinstimmen, müssen die Koeffizienten vor den verschiedenen Potenzen von x übereinstimmen. Man erhält so das Gleichungssystem

$$\begin{aligned}A_1 + A_2 &= 1\\ A_1 - A_2 &= 0\end{aligned}$$

Es ist also $A_1 = A_2 = \frac{1}{2}$. Die Stammfunktion von $\frac{1}{1-x^2}$ lässt sich nun wie folgt berechnen:

$$\begin{aligned}\int \frac{1}{1-x^2}\, dx &= \int \frac{\frac{1}{2}}{1-x}\, dx + \int \frac{\frac{1}{2}}{1+x}\, dx\\ &= \frac{1}{2}\left(-\int \frac{1}{x-1}\, dx + \int \frac{1}{x+1}\, dx\right)\\ &= \frac{1}{2}\left(-\log(|x-1|) + \log(|x+1|)\right)\\ &= \frac{1}{2}\log\left(\left|\frac{x+1}{x-1}\right|\right).\end{aligned}$$

Es sei nun P eine beliebige Polynomfunktion vom Grad n mit reellen Koeffizienten, also $P(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_0$ mit festen $a_0, a_1, \ldots, a_n \in \mathbb{R}$. Nach dem Fundamentalsatz der Algebra (Satz 17; einen Beweis findet man auf Seite 253) besitzt

P genau n nicht notwendigerweise paarweise verschiedene Nullstellen $\alpha_1, \dots, \alpha_n$ in den komplexen Zahlen. Von diesen n Nullstellen seien genau r viele reell. Ohne Einschränkung der Allgemeinheit seien dies die Nullstellen $\alpha_1, \dots, \alpha_r$. Die Nullstellen $\alpha_{r+1}, \dots, \alpha_n$ liegen also in $\mathbb{C} \setminus \mathbb{R}$. Ist $\alpha \in \mathbb{C}$ Nullstelle eines Polynoms P mit reellen Koeffizienten, so ist auch die komplex-konjugierte Zahl $\overline{\alpha}$ eine Nullstelle des Polynoms P, denn es ist $P(\overline{\alpha}) = \overline{P(\alpha)} = 0$.

Die $n-r$ nicht-reellen Nullstellen sind also (sagen wir mal s viele) Paare von komplex-konjugierten Zahlen. Wir können diese dann schreiben als

$$\alpha_{r+1}, \dots, \alpha_{r+s}, \overline{\alpha_{r+1}}, \dots, \overline{\alpha_{r+s}}.$$

Damit lässt sich P in der folgenden Form schreiben

$$P(x) = a_n \prod_{\rho=1}^{r} (x - \alpha_\rho) \prod_{\sigma=1}^{s} (x - \alpha_{r+\sigma})(x - \overline{\alpha_{r+\sigma}}).$$

Schauen wir uns ein Produkt der Form $(x - \alpha)(x - \overline{\alpha})$ an. Es ist $(x - \alpha)(x - \overline{\alpha}) = x^2 - (\alpha + \overline{\alpha})x + \alpha\overline{\alpha} = x^2 - 2\operatorname{Re}(\alpha)x + |\alpha|^2$. Folglich existieren (nicht notwendigerweise paarweise verschiedene) reelle Zahlen $b_1, \dots, b_s$ und $c_1, \dots, c_s$, so dass man das Polynom P schreiben kann als

$$P(x) = a_n \prod_{\rho=1}^{r} (x - \alpha_\rho) \prod_{\sigma=1}^{s} (x^2 + b_\sigma x + c_\sigma).$$

Wir können an dieser Stelle also festhalten, dass jedes Polynom mit reellen Koeffizienten über den reellen Zahlen in ein Produkt aus Polynomen vom Grad 1 oder 2 zerfällt.

Nun können wir das Verfahren der Partialbruchzerlegung allgemein beschreiben.

Gegeben sei also eine rationale Funktion $\frac{Q(x)}{P(x)}$. Ist der Grad von Q größer oder gleich dem von P, so kann man mittels Polynomdivision Polynome R und Q_1 berechnen, so dass

$$\frac{Q(x)}{P(x)} = R(x) + \frac{Q_1(x)}{P(x)}$$

gilt und der Grad von Q_1 echt kleiner als der von P ist. Wir nehmen deshalb im weiteren an, dass der Grad von Q echt kleiner als der von P ist und wir eine Stammfunktion von Q/P suchen.

a) Ist α eine reelle Nullstelle von P und $(x-\alpha)^k$ die höchste Potenz, die als Faktor in der obigen Produktdarstellung von P auftaucht, so ordnen wir diesem Faktor die k Partialbrüche

$$\frac{A_1}{x - \alpha}, \frac{A_2}{(x - \alpha)^2}, \dots, \frac{A_k}{(x - \alpha)^k}$$

zu.

b) Ist x^2+bx+c ein Faktor vom Grad 2 in der obigen Produktdarstellung von P und $(x^2+bx+c)^l$ die höchste Potenz, die als Faktor in der obigen Produktdarstellung von P auftaucht, so ordnen wir diesem Faktor die l Partialbrüche

$$\frac{B_1x+C_1}{x^2+bx+c}, \frac{B_2x+C_2}{(x^2+bx+c)^2}, \ldots, \frac{B_lx+C_l}{(x^2+bx+c)^l}$$

zu.

c) Wir führen die Schritte a) und b) für alle verschiedenen Faktoren von P durch. Dabei sind die Unbestimmten A_k, B_l, C_l natürlich jeweils unterschiedlich zu indizieren.

d) Wir setzen Q/P gleich der Summe aller Partialbrüche, die sich in den vorstehenden Schritten ergeben haben, und wir addieren die Summe der Partialbrüche auf, so dass wir einen Bruch mit dem Nenner P erhalten. Im Zähler dieses Bruches tauchen die Unbestimmten A_k, B_l, C_l auf. Wir sortieren den Zähler dieses Bruches nach aufsteigenden Potenzen von x, ebenso verfahren wir mit Q. Wir setzen die Koeffizienten gleicher Potenzen von x gleich und erhalten ein lineares Gleichungssystem, aus dem wir die Unbestimmten A_k, B_l, C_l berechnen können.

Schauen wir uns ein größeres Beispiel an.

Wir wollen eine Stammfunktion von

$$\frac{x^7-4x^6+9x^5-18x^4+23x^3-19x^2+15x-4}{x^5-4x^4+6x^3-6x^2+5x-2}$$

bestimmen.

Mittels Polynomdivison sehen wir, dass

$$\begin{aligned} &x^7-4x^6+9x^5-18x^4+23x^3-19x^2+15x-4 \\ &\quad = (x^5-4x^4+6x^3-6x^2+5x-2)(x^2+3)+(x^2+2) \end{aligned}$$

beziehungsweise

$$\begin{aligned} &\frac{x^7-4x^6+9x^5-18x^4+23x^3-19x^2+15x-4}{x^5-4x^4+6x^3-6x^2+5x-2} \\ &\quad = (x^2+3)+\frac{x^2+2}{x^5-4x^4+6x^3-6x^2+5x-2} \end{aligned}$$

ist. Als nächstes suchen wir Nullstellen von $x^5-4x^4+6x^3-6x^2+5x-2$. Sollte es ganzzahlige Nullstellen geben, so müssen diese den Koeffizienten a_0, in unserem Beispiel also -2, teilen. Als Kandidaten für ganzzahlige Nullstellen kommen also $-2, -1, 1, 2$ in Frage. Man rechnet schnell nach, dass sowohl 1 als auch 2 Nullstellen des Polynoms sind. Eine Polynomdivison durch $x-1$ ergibt

$$x^5-4x^4+6x^3-6x^2+5x-2 = (x^4-3x^3+3x^2-3x+2)(x-1).$$

Da 1 auch Nullstelle von $x^4 - 3x^3 + 3x^2 - 3x + 2$ ist, können wir noch einen Faktor $x - 1$ ausdividieren und erhalten

$$x^5 - 4x^4 + 6x^3 - 6x^2 + 5x - 2 = (x^3 - 2x^2 + x - 2)(x-1)^2.$$

Nun dividieren wir einen Faktor $x - 2$ aus, so dass wir

$$x^5 - 4x^4 + 6x^3 - 6x^2 + 5x - 2 = (x^2 + 1)(x-2)(x-1)^2$$

erhalten. $x^2 + 1$ hat offensichtlich keine reelle Nullstellen. Wir haben also die Faktorisierung von P in reelle Polynome vom Grad 1 oder 2 gefunden. Es ist

$$\frac{x^2+2}{x^5 - 4x^4 + 6x^3 - 6x^2 + 5x - 2} = \frac{x^2+2}{(x^2+1)(x-2)(x-1)^2}.$$

Nun bestimmen wir die Partialbrüche.

Zum Faktor $x - 2$ gehört der Partialbruch

$$\frac{A_1}{x-2},$$

zum Faktor $(x-1)^2$ gehören die Partialbrüche

$$\frac{A_2}{x-1}, \frac{A_3}{(x-1)^2}$$

und zum Faktor $x^2 + 1$ gehört der Partialbruch

$$\frac{Bx+C}{x^2+1}.$$

Wir haben also

$$\frac{x^2+2}{(x^2+1)(x-2)(x-1)^2} = \frac{A_1}{x-2} + \frac{A_2}{x-1} + \frac{A_3}{(x-1)^2} + \frac{Bx+C}{x^2+1}$$

Für die rechte Seite der Gleichung gilt

$$\begin{aligned} & \frac{A_1}{x-2} + \frac{A_2}{x-1} + \frac{A_3}{(x-1)^2} + \frac{Bx+C}{x^2+1} \\ = \quad & \frac{A_1(x-1)^2(x^2+1)}{(x-2)(x-1)^2(x^2+1)} + \frac{A_2(x-2)(x-1)(x^2+1)}{(x-1)(x-2)(x-1)(x^2+1)} \\ & + \frac{A_3(x-2)(x^2+1)}{(x-1)^2(x-2)(x^2+1)} + \frac{(Bx+C)(x-2)(x-1)^2}{(x^2+1)(x-2)(x-1)^2}. \end{aligned}$$

Multiplizieren wir nun die Zähler aus, addieren die Brüche und sortieren nach den Potenzen von x, so erhalten wir

$$\begin{aligned}
& \frac{A_1}{x-2}+\frac{A_2}{x-1}+\frac{A_3}{(x-1)^2}+\frac{Bx+C}{x^2+1} \\
= \quad & \frac{(A_1+A_2+B)x^4+(-2A_1-3A_2+A_3-4B+C)x^3}{(x^2+1)(x-2)(x-1)^2} \\
& +\frac{(2A_1+3A_2-2A_3+5B-4C)x^2+(-2A_1-3A_2+A_3-2B+5C)x}{(x^2+1)(x-2)(x-1)^2} \\
& +\frac{A_1+2A_2-2A_3-2C}{(x^2+1)(x-2)(x-1)^2}
\end{aligned}$$

Vergleichen wir den Zähler dieses Bruches mit x^2+2, so erhalten wir mittels Koeffizientenvergleich das folgende lineare Gleichungssystem, aus dem sich A_1, A_2, A_3, B, C bestimmen lassen.

$$\begin{aligned}
A_1+A_2+B &= 0 \\
-2A_1-3A_2+A_3-4B+C &= 0 \\
2A_1+3A_2-2A_3+5B-4C &= 1 \\
-2A_1-3A_2+A_3-2B+5C &= 0 \\
A_1+2A_2-2A_3-2C &= 2.
\end{aligned}$$

Wir erhalten $A_1=\frac{6}{5}, A_2=-1, A_3=-\frac{3}{2}, B=-\frac{1}{5}, C=\frac{1}{10}$. Wir haben also

$$\begin{aligned}
& \frac{x^7-4x^6+9x^5-18x^4+23x^3-19x^2+15x-4}{x^5-4x^4+6x^3-6x^2+5x-2} \\
= \quad & x^2+3+\frac{\frac{6}{5}}{x-2}-\frac{1}{x-1}-\frac{\frac{3}{2}}{(x-1)^2}+\frac{-\frac{1}{5}x+\frac{1}{10}}{x^2+1} \\
= \quad & x^2+3+\frac{6}{5}\cdot\frac{1}{x-2}-\frac{1}{x-1}-\frac{3}{2}\cdot\frac{1}{(x-1)^2}-\frac{1}{5}\cdot\frac{x}{x^2+1}+\frac{1}{10}\cdot\frac{1}{x^2+1}
\end{aligned}$$

und somit

$$\begin{aligned}
& \int \frac{x^7-4x^6+9x^5-18x^4+23x^3-19x^2+15x-4}{x^5-4x^4+6x^3-6x^2+5x-2}dx \\
= \quad & \int x^2+3\ dx+\frac{6}{5}\int\frac{1}{x-2}dx-\int\frac{1}{x-1}dx-\frac{3}{2}\int\frac{1}{(x-1)^2}dx \\
& -\frac{1}{5}\int\frac{x}{x^2+1}dx+\frac{1}{10}\int\frac{1}{x^2+1}dx.
\end{aligned}$$

Die Stammfunktionen in der letzten Zeile können wir leicht berechnen. Wir haben, wie man durch Ableiten leicht nachrechnen,

$$\begin{aligned}\int x^2+3\,dx &= \frac{1}{3}x^3+3x,\\ \int \frac{1}{x-2}dx &= \log|x-2|,\\ \int \frac{1}{x-1}dx &= \log|x-1|,\\ \int \frac{1}{(x-1)^2}dx &= -\frac{1}{x-1},\\ \int \frac{x}{x^2+1}dx &= \frac{1}{2}\int \frac{2x}{x^2+1}dx = \frac{1}{2}\log|x^2+1|,\\ \int \frac{1}{x^2+1}dx &= \arctan(x),\end{aligned}$$

wobei wir bei der letzten Stammfunktion Erkenntnisse aus Problem 75 benutzt haben. Zusammenfassend haben wir also erhalten

$$\begin{aligned}&\int \frac{x^7-4x^6+9x^5-18x^4+23x^3-19x^2+15x-4}{x^5-4x^4+6x^3-6x^2+5x-2}dx\\ = \;&\frac{1}{3}x^3+3x+\frac{6}{5}\log|x-2|-\log|x-1|+\frac{3}{2}\cdot\frac{1}{x-1}-\frac{1}{10}\log|x^2+1|+\frac{1}{10}\arctan(x).\end{aligned}$$

5) Wir wollen

$$\int x^3\sin(x^2+1)dx$$

bestimmen. Im Prinzip verlangt dieses Integral nach einer partiellen Integration, da wir das Produkt zweier Funktionen haben und der Integrand „einfacher" wird, wenn wir x^3 ableiten und $\sin$ durch seine Stammfunktion ersetzen. Dabei stört uns aber das x^2+1, das im Sinus steht. Dieses substituieren wir deshalb zunächst weg. Wir setzen also $y := x^2+1$. Nun noch ein kleiner Trick zur Substitutionsregel. $\frac{dy}{dx}$ ist eine andere Schreibweise für die Ableitung y', wenn y eine Funktion in x ist. Wir tun nun so, als ob $\frac{dy}{dx}$ tatsächlich ein Bruch von dy und dx wäre[1]. Es ist $\frac{dy}{dx}=2x$, also $dx=\frac{dy}{2x}$. Dies setzen wir in unser Integral ein und wir müssen nun alle auftretenden Ausdrücke in x durch Ausdrücke in y ersetzen, bis wir schließlich ein Integral haben, welches nur noch Ausdrücke in y enthält. Wir erhalten

$$\int x^3\sin(x^2+1)dx=\int x^3\sin(y)\frac{dy}{2x}=\frac{1}{2}\int x^2\sin(y)dy=\frac{1}{2}\int(y-1)\sin(y)dy,$$

wobei wir ausgenutzt haben, dass $x^2=y-1$ ist. Das Integral rechts lösen wir nun per partieller Integration mit $g(y)=y-1, g'(y)'=1, f(y)'=\sin(y), F(y)=-\cos(y)$. Wir haben

$$\int(y-1)\sin(y)dy=-(y-1)\cos y+\int\cos y dy=-(y-1)\cos y+\sin y.$$

[1] Man kann durchaus zeigen, dass die hier genutzen Schreibweisen mathematisch Sinn machen und korrekt sind. Das würde unseren Rahmen aber sprengen.

Wenn wir nun resubstituieren, erhalten wir

$$\begin{aligned}\int x^3 \sin(x^2+1)dx &= \frac{1}{2}\int (y-1)\sin(y)dy = -\frac{1}{2}(y-1)\cos y + \frac{1}{2}\sin y \\ &= -\frac{1}{2}x^2\cos(x^2+1) + \frac{1}{2}\sin(x^2+1).\end{aligned}$$

Exkurs: Partialbruchzerlegung und Stammfunktionen mit Computeralgebrasystemen

Wir wollen hier kurz zeigen, wie man die im obigen Beispiel behandelte Partialbruchzerlegung Differentialgleichung mit Hilfe von Maxima durchführen kann. Außerdem zeigt das folgende Beispiel auch, wie man mit Maxima Stammfunktionen berechnen lässt.

```
(%i1) partfrac((x^7-4*x^6+9*x^5-18*x^4+23*x^3-19*x^2+15*x-4)/(x^5-4*x^4+6*x^3-6*x^2+5*x-2), x);
```

(%o1) $-\frac{2x-1}{10(x^2+1)}+x^2-\frac{1}{x-1}-\frac{3}{2(x-1)^2}+\frac{6}{5(x-2)}+3$

```
(%i2) integrate((x^7-4*x^6+9*x^5-18*x^4+23*x^3-19*x^2+15*x-4)/(x^5-4*x^4+6*x^3-6*x^2+5*x-2), x);
```

(%o2) $-\frac{\log(x^2+1)}{10}+\frac{\operatorname{atan}(x)}{10}-\log(x-1)+\frac{6\log(x-2)}{5}+\frac{x^3+9x}{3}+\frac{3}{2x-2}$

Ende des Exkurses

4.9. Differentialgleichungen erster Ordnung

Definition 213 (Differentialgleichung erster Ordnung)**:** Es sei $D \subset \mathbb{R}^2$ offen und $f : D \to \mathbb{R}$ stetig. Die Gleichung

$$y'(x) = f(x, y(x))$$

wird *gewöhnliche Differentialgleichung (Dgl.) erster Ordnung* genannt.
Eine Lösung ist eine auf einem Intervall I differenzierbare Funktion φ, deren Graph in D liegt und die für alle $x \in I$ die gegebene Differentialgleichung erfüllt. Das heißt für alle $x \in I$ gilt

$$\varphi'(x) = f(x, \varphi(x)).$$

BEMERKUNG: Für diese Differentialgleichung schreibt man oft kürzer $y' = f(x, y)$. Besonders einfach ist die Situation, wenn die Differentialgleichung *trennbar* ist, das heisst, wenn es stetige Funktionen $h : I \to \mathbb{R}$ und $g : J \to \mathbb{R}$ mit $f(x, y(x)) = h(x) \cdot g(y(x))$ gibt. Ist nämlich y eine Lösung von

$$y'(x) = h(x) \cdot g(y(x))$$

und $g(t) \neq 0$ für alle $t \in J$, so gilt

$$\frac{y'(x)}{g(y(x))} = h(x).$$

Nun bezeichne G eine Stammfunktion von $\frac{1}{g(t)}$, so ist

$$(G(y(x)))' = G'(y(x)) \cdot y'(x) = \frac{1}{g(y(x))} y'(x),$$

also die linke Seite der Gleichung $\frac{y'(x)}{g(y(x))} = h(x)$. Da $g(t) \neq 0$ vorausgesetzt wurde, gilt $G'(t) > 0$ oder $G'(t) < 0$ für alle $t \in J$. Deshalb ist G streng monoton und somit existiert die Umkehrfunktion G^{-1}. Ist H eine Stammfunktion von h, so sind also alle Lösungen von $y'(x) = h(x) \cdot g(y(x))$ von der Form

$$y(x) = G^{-1}(H(x)),$$

denn es gilt

$$\begin{aligned} y'(x) =& (G^{-1}(H(x)))' \\ =& \frac{1}{G'(G^{-1}(H(x)))} H'(x) \\ =& g(y(x)) \cdot h(x). \end{aligned}$$

Beispiel: Betrachte die Differentialgleichung $y'(x) = x \cdot y^2(x)$. Wähle $g(t) = t^2$. Diese Funktion ist ist für $t \neq 0$ positiv. Eine Stammfunktion G von $\frac{1}{g(t)} = \frac{1}{t^2}$ lautet $G(t) = -\frac{1}{t} + c_1$ mit $c_1 \in \mathbb{R}$. Eine Stammfunktion zu $h(x) = x$ lautet $H(x) = \frac{1}{2}x^2 + c_2$ mit $c_2 \in \mathbb{R}$. Wegen $G(y(x)) = H(x)$ gilt also

$$-\frac{1}{y(x)} = \frac{1}{2}x^2 + c$$

für ein $c \in \mathbb{R}$. Daraus folgt

$$y(x) = \frac{1}{-\frac{1}{2}x^2 - c} = \frac{2}{-2c - x^2}.$$

Aber auch $y(x) = 0$ für alle $x \in \mathbb{R}$ löst die Differentialgleichung. Es gibt unendlich viele Lösungen.

Definition 214 (Anfangswertproblem)**:** Die Differentialgleichung

$$y'(x) = f(x, y(x))$$

zusammen mit der zusätzlichen Bedingung

$$y(x_0) = y_0$$

heißt *Anfangswertproblem (Awp)*. Eine Lösung $\varphi \in C^1(I)$ des Anfangswertproblems ist eine Lösung φ der Differentialgleichung mit $\varphi(x_0) = y_0$.

Beispiel: Betrachte das Anfangswertproblem

$$\begin{aligned} y'(x) =& xy^2(x) \\ y(0) =& 1. \end{aligned}$$

Die allgemeine Lösung ist wie schon gesehen die Funktion

$$y(x) = \frac{2}{-2c - x^2}.$$

Aus $1 = y(0) = \frac{2}{-2c}$ folgt $c = -1$ und somit ist

$$y(x) = \frac{2}{2 - x^2}$$

die Lösung des Anfangswertproblems. Diese Lösung ist eindeutig, denn nach obiger Betrachtung muss y von der Gestalt $\frac{2}{-2c-x^2}$ sein.

Beispiel: Wir wollen uns noch ein weiteres Beispiel anschauen und hierbei noch einmal auf einen Trick zurückgreifen, den wir schon bei der Substitionsregel im Kapitel über Stammfunktionen benutzt haben. Wir suchen die Lösung des Anfangswertproblems $y' = \cos(x)(y+1)$ mit $y(0) = 1$. Es ist $y' = \frac{dy}{dx}$. Nun tun wir wieder so, als wenn dies ein normaler Bruch wäre und bringen alle Ausdrücke mit y auf die linke Seite der Gleichung und alle Ausdrücke mit x auf die rechte Seite, also $\frac{dy}{y+1} = \cos(x)dx$. Schnell noch ein $\int$ davor geschrieben und wir erhalten

$$\int \frac{dy}{y+1} = \int \cos(x)dx.$$

Da $y(0) = 1$ positiv und y stetig ist, können wir annehmen, dass $y(x)$ in einer Umgebung von 0 positiv ist. Es existiert eine Konstante c, so dass

$$\log(y+1) = \sin(x) + c$$

ist, also

$$y(x) = \exp(\sin(x) + c) - 1.$$

Um c zu bestimmen, nutzen wir nun die Anfangswertbedingung aus und erhalten

$$1 = y(0) = \exp(c) - 1,$$

also $c = \log 2$.

Definition 215 (Lineare Differentialgleichung erster Ordnung)**:** Sind $a, g : I \to \mathbb{R}$ Funktionen auf einem Intervall $I \subset \mathbb{R}$, so heißt

$$y'(x) + a(x)y(x) = g(x)$$

lineare Differentialgleichung erster Ordnung. Die Funktion a heißt Koeffizientenfunktion und g heißt Inhomogenität. Ist $g = 0$ so heißt die Differentialgleichung *homogen*, sonst *inhomogen*.

Satz 216: Es sei y_p eine Lösung der Differentialgleichung

$$y'(x) + a(x)y(x) = g(x).$$

Die Lösungsmenge dieser Gleichung ist gleich der Menge

$$\mathcal{L} = \{y_p + y_h \,|\, y_h \text{ löst } \; y_h'(x) + a(x)y_h(x) = 0 \,\}.$$

BEWEIS: Es sei y eine beliebige Lösung von $y'(x) + a(x)y(x) = g(x)$. Es ist zu zeigen, dass $y \in \mathcal{L}$ liegt. Es gilt

$$(y(x) - y_p(x))' + a(x)(y(x) - y_p(x)) = \underbrace{y'(x) + a(x)y(x)}_{g(x)} - \underbrace{(y_p'(x) + a(x)y_p(x))}_{g(x)} = 0.$$

Also löst $y - y_p$ die homogene Differentialgleichung. □

Satz 217: Wir betrachten die homogenen Differentialgleichung

$$y'(x) + a(x)y(x) = 0.$$

Besitzt a eine Stammfunktion, so ist die allgemeine Lösung dieser Differentialgleichung gegeben durch

$$y_h(x) = Ke^{-\int a(x)\,dx}.$$

mit $K \in \mathbb{R}$.

BEWEIS: Es gilt $\frac{y'(x)}{y(x)} = -a(x)$. Also folgt

$$\log(|y|) = \int \frac{y'(x)}{y(x)}\,dx = -\int a(x)\,dx$$

beziehungsweise

$$|y| = e^{-\int a(x)\,dx}.$$

Damit gilt $y(x) \neq 0$ für alle x, also gilt

$$y(x) = \operatorname{sign}(y(x)) \cdot |y(x)|,$$

wobei die Signumfunktion sign durch

$$\operatorname{sign}(y(x)) := \begin{cases} 1, & y(x) > 0 \\ -1, & y(x) < 0 \end{cases}$$

definiert ist. □

Um eine spezielle Lösung der inhomogenen Differentialgleichung

$$y'(x) = +a(x)y(x) = g(x)$$

zu finden, nutzt man die sogenannte *Variation der Konstanten*. Die Funktion $y_h(x) = K \cdot e^{-\int a(x)\,dx}$ löst die Gleichung $y'(x) + a(x)y(x) = 0$. Man fasst K nun als Funktion von x auf und hofft, dass

$$y_p(x) = K(x)e^{-A(x)}$$

die inhomogene Differentialgleichung löst, wobei A eine Stammfunktion von a sei. Es gilt

$$\begin{aligned} y_p'(x) =& K'(x)e^{-A(x)} - K(x)e^{-A(x)}a(x) \\ =& K'(x)e^{-A(x)} - a(x)y_p(x). \end{aligned}$$

Damit folgt

$$\begin{aligned} g(x) &= y_p'(x) + a(x)y_p(x) \\ &= K'(x)e^{-A(x)} \end{aligned}$$

also gilt

$$K'(x) = g(x)e^{A(x)}$$

und somit

$$K(x) = \int g(x)e^{A(x)}\,dx.$$

Beispiel: Wir betrachten das Anfangswertproblem

$$\begin{aligned} y'(x) + 2y(x) &= e^{4x} + 1 \\ y(0) &= 1. \end{aligned}$$

Dann lautet die allgemeine Lösung der homogenen Gleichung

$$y_h(x) = K \cdot e^{-2x},$$

wobei $K \in \mathbb{R}$ beliebig ist. Variation der Konstanten ergibt

$$\begin{aligned} K(x) =& \int e^{2x}(e^{4x} + 1)\,dx \\ =& \frac{1}{6}e^{6x} + \frac{1}{2}e^{2x} \end{aligned}$$

und somit

$$\begin{aligned} y_p(x) =& K(x) \cdot e^{-2x} \\ =& \left(\frac{1}{6}e^{6x} + \frac{1}{2}e^{2x}\right) \cdot e^{-2x} \\ =& \frac{1}{6}e^{4x} + \frac{1}{2}. \end{aligned}$$

Für $K \in \mathbb{R}$ löst dann also

$$y(x) = y_p(x) + y_h(x) = \frac{1}{6}e^{4x} + \frac{1}{2} + K \cdot e^{-2x}$$

die Differentialgleichung. Wegen

$$1 = y(0) = \frac{1}{6} + \frac{1}{2} + K$$

muss also

$$K = 1 - \frac{2}{3} = \frac{1}{3}$$

sein und somit lautet die Lösung des Anfangswertproblems

$$y(x) = \frac{1}{6}e^{4x} + \frac{1}{2} + \frac{1}{3}e^{-2x}.$$

Differentialgleichungen sind ein weites Feld. Es gibt viele verschiedene Lösungsansätze je nach Typ der betrachteten Differentialgleichung. Auch gibt es allgemeine Sätze, die klären, unter welchen Bedingungen Lösungen auf jeden Fall existieren und auch wann diese Lösungen eindeutig sind. Wir verweisen hier auf die weiterführende Literatur, wie etwa [Wal00]

Exkurs: Numerisches Lösen von Differentialgleichungen

Wir können nun einige wenige Typen von Differentialgleichungen explizit lösen. Es gibt auch Differentialgleichungen die man gar nicht explizit lösen kann. In Fällen, wo man die Differentialgleichung nicht explizit lösen kann, kann man aber versuchen, die Lösung der Differentialgleichung – genauer eines Anfangswertproblems – näherungsweise numerisch zu bestimmen. Wir wollen dies zunächst an der Differentialgleichung

$$y' = y^2 + x$$

mit der Anfangswertbedingung $y(1) = 2$ demonstrieren. Diese Gleichung ist nichtlinear und auch nicht trennbar. Dazu ersetzt man die Ableitung y' auf der linken Seite der Differentialgleichung durch den Differenzenquotienten

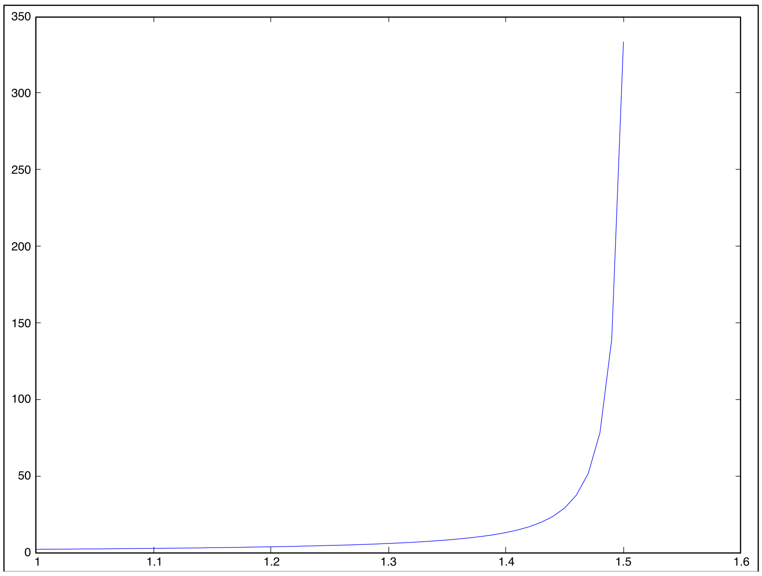
$$\frac{y(x+h) - y(x)}{h}$$

und erhält dann

$$\frac{y(x+h) - y(x)}{h} \approx y(x)^2 + x \text{ bzw. } y(x+h) \approx y(x) + h(y(x)^2 + x).$$

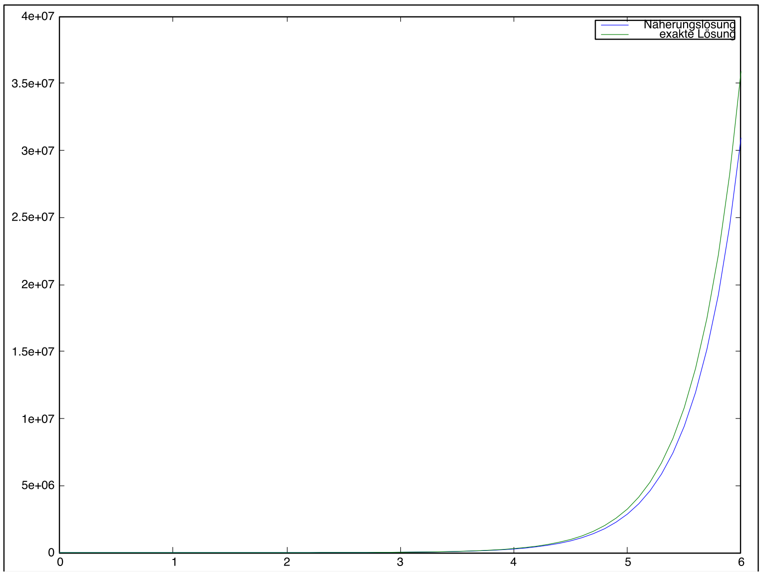
Kennen wir nun y zum Zeitpunkt x, so können wir auch y zum Zeitpunkt $x + h$ (näherungsweise) berechnen. Dann können wir y zum Zeitpunkt $x + h + h$ berechnen u.s.w. Das folgende Matlab-Skript löst näherungsweise die Differentialgleichung $y' = y^2 + x$ für x zwischen 1 und 1,5 zum Anfangswert $y(1) = 2$ und plottet den Graphen der Näherungslösung.

```
% Numerische Lösung der Differentialgleichung y'=y^2+x mit y(1)=2
%
clear all
Xend=1.5;
h=0.01;
y(1)=2;
x(1)=1;
for i=1:((Xend-1)/h)
y(i+1)=y(i)+h*(y(i)^2+x(i));
x(i+1)=x(i)+h;
end
figure(1)
clf()
plot(x,y)
```



Hier noch ein Skript, welches die Differentialgleichung $y' = c * y$ mit $y(0) = 20$ und $c = 2.4$ numerisch löst und sowohl die numerische als auch die exakte Lösung plottet.

```
clear all;
N=6;
h=0.1;
y(1)=20;
c=2.4;
y(2)=y(1)*(1+c*h);
t=0:h:N;
for i=2:length(t)-1
y(i+1)=y(i-1)+c*y(i)*2*h;
end
figure(3)
clf()
f=20*exp(c*t);
plot(t,y,t,f)
legend('Näherungslösung','exakte Lösung')
```



Ende des Exkurses

4.10. Das Integral von Treppenfunktionen

In diesem und den folgenden Abschnitten seien stets $a, b \in \mathbb{R}$ mit $a < b$.

Motivation: Mit Hilfe des Integralbegriffes ordnet man einer nichtnegativen Funktion f den Flächeninhalt $F_a^b(f)$ unterhalb des Graphen zu.

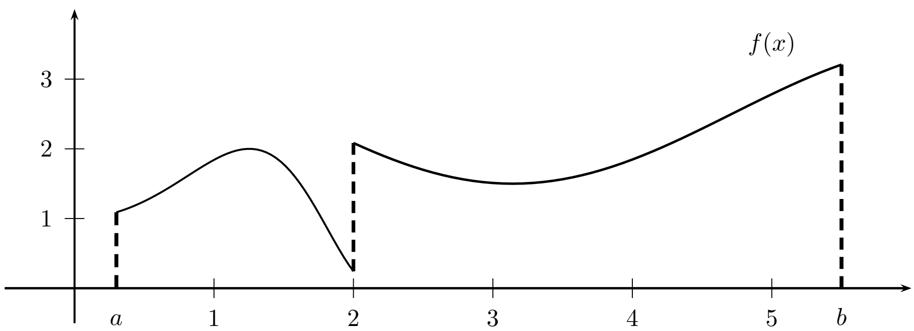


Abbildung 4.17.: Eine nichtnegative stückweise stetige Funktion

An den Integralbegriff werden bezüglich des Flächeninhalts einige Erwartungen gestellt:

a) Er soll mit bekannten Flächeninhalten übereinstimmen. Zum Beispiel soll der Flächeninhalt unter eine konstante Funktion auf dem Intervall $[a, b]$ dem eines Rechtecks entsprechen. Es soll also

$$F_a^b(f) = c(b - a)$$

gelten.

b) Unterteilt man die Fläche an einer beliebigen Stelle t, welche zwischen a und b liegt, so soll

$$F_a^b(f) = F_a^t(f) + F_t^b(f)$$

gelten.

c) Gilt $f(x) \leq g(x)$ für alle $x \in [a, b]$, so soll $F_a^b(f) \leq F_a^b(g)$ gelten.

Aus a) und b) wird klar, welchen Flächeninhalt $F_a^b(f)$ man der Fläche unter einer Treppenfunktion wie in der folgenden Abbildung zuordnet.

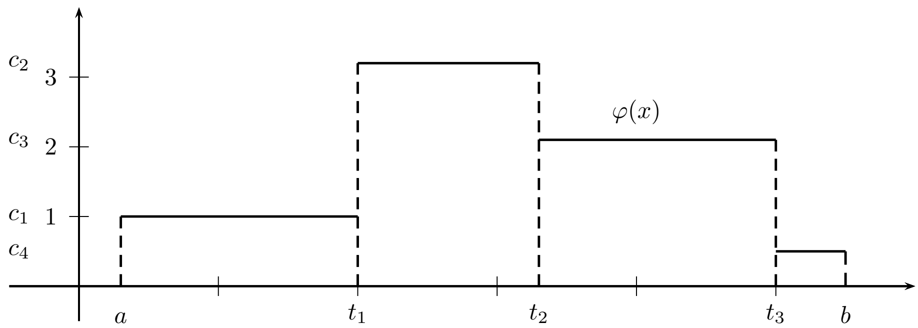


Abbildung 4.18.: Flächeninhaltsbestimmung bei einer Treppenfunktion

Der entsprechende Flächeninhalt ist nämlich

$$F_a^b(\varphi) = c_1(t_1 - a) + c_2(t_2 - t_1) + c_3(t_2 - t_3) + c_4(b - t_3).$$

Hat man eine beliebige Funktion f durch zwei Treppenfunktionen φ und ψ eingeschachtelt, so gilt wegen c)

$$F_a^b(\varphi) \leq F_a^b(f) \leq F_a^b(\psi).$$

Dies kann man nutzen um $F_a^b(f)$ einen Wert zuzuordnen, indem man f durch Treppenfunktionen immer weiter annähert.

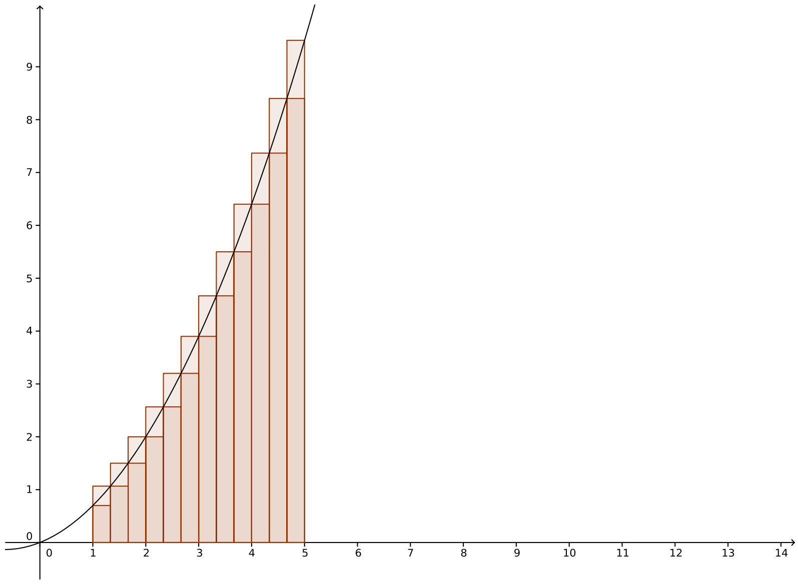


Abbildung 4.19.: Approximation der Fläche unterhalb eines Funktionsgraphen mittels Treppenfunktionen

Definition 218 (Treppenfunktion)**:** (i) Eine endliche Menge $Z := \{t_0, t_1, \ldots, t_n\}$ heißt *Unterteilung* oder *Zerlegung* von $[a, b]$, falls

$$a = t_0 < t_1 < \ldots < t_n = b$$

gilt.

(ii) Die Funktion $\varphi : [a, b] \to \mathbb{R}$ heißt *Treppenfunktion*, wenn es eine Zerlegung $Z := \{t_0, \ldots, t_n\}$ von $[a, b]$ gibt, so dass

$$\forall k \in \{1, \ldots, n\} \; \exists c_k \in \mathbb{R} \; \forall x \in]t_{k-1}, t_k[: \; \varphi(x) = c_k$$

gilt. Die Werte $\varphi(t_k)$ an den Sprungstellen t_k sind beliebig. Mit $T([a, b])$ wird die Menge aller Treppenfunktionen auf $[a, b]$ bezeichnet.

Lemma 219: Es sei $\varphi \in T([a, b])$ eine Treppenfunktion bezüglich einer Zerlegung Z. Dann ist φ auch eine Treppenfunktion bezüglich jeder feineren Zerlegung $Z' := \{t'_0, \ldots, t'_n\}$ mit $Z \subset Z'$.

Korollar 220: Es seien $\varphi, \psi \in T([a, b])$. Die Funktion φ ist eine Treppenfunktion bezüglich der Zerlegung Z und die Funktion ψ bezüglich der Zerlegung Z'. Dann sind φ und ψ Treppenfunktionen bezüglich der Zerlegung $Z \cup Z'$.

Korollar 221: Die Menge $T([a, b])$ ist ein Untervektorraum des Vektorraum $\mathrm{Abb}([a, b], \mathbb{R})$, welcher alle Abbildungen von $[a, b]$ nach $\mathbb{R}$ beinhaltet.

Definition 222 (Fläche unter einer Treppenfunktion)**:** Es sei $\varphi \in T([a, b])$ eine Treppenfunktion und $Z := \{t_0, \dots, t_n\}$ eine Zerlegung derart, dass

$$\varphi(x) = c_k$$

für alle $x \in]t_{k-1}, t_k[$ gilt. Dann ist der (orientierte) Flächeninhalt unter der Treppenfunktion definiert durch

$$F_a^b(\varphi, Z) := \sum_{k=1}^{n} c_k(t_k - t_{k-1}).$$

Satz 223: Es seien Z und Z' zwei Zerlegungen von $[a, b]$. Ist φ eine Treppenfunktion sowohl bezüglich Z als auch bezüglich Z', so gilt

$$F_a^b(\varphi, Z) = F_a^b(\varphi, Z')$$

BEWEIS: *1. Schritt:* Für $t \in [a, b] \setminus Z$ gilt

$$F_a^b(\varphi, Z) = F_a^b(\varphi, Z \cup \{t\}),$$

denn es gibt ein i, so dass $t_{i-1} < t < t_i$ gilt und es ist

$$c_i(t_i - t_{i-1}) = c_i(t - t_{i-1}) + c_i(t_i - t),$$

woraus $F_a^b(\varphi, Z) = F_a^b(\varphi, Z \cup \{t\})$ sofort folgt.

2. Schritt: Mittels vollständiger Induktion und dem ersten Schritt folgt nun für $Z \subset Z'$

$$F_a^b(\varphi, Z) = F_a^b(\varphi, Z').$$

3. Schritt: Sind nun Z und Z' Zerlegungen wie sie im Satz gefordert wurden, so gilt auf Grund des 2. Schrittes

$$F_a^b(\varphi, Z) = F_a^b(\varphi, Z \cup Z') = F_a^b(\varphi, Z').$$

□

Definition 224: Für $\varphi \in T([a,b])$ setzt man

$$\int_a^b \varphi(x)\, dx := F_a^b(\varphi) := F_a^b(\varphi, Z),$$

wobei Z irgendeine Zerlegung ist, bezüglich der φ eine Treppenfunktion ist.

Satz 225: Für $\varphi, \psi \in T([a,b])$ und $\lambda \in \mathbb{R}$ gilt

(i)

$$\int_a^b (\lambda\varphi(x) + \psi(x))\, dx = \lambda \int_a^b \varphi(x)\, dx + \int_a^b \psi(x)\, dx$$

(ii) Ist $\varphi \leq \psi$, das heißt für alle $x \in [a,b]$ gilt $\varphi(x) \leq \psi(x)$, dann gilt

$$\int_a^b \varphi(x)\, dx \leq \int_a^b \psi(x)\, dx.$$

BEWEIS: Man kann φ und ψ bezüglich der gleichen Zerlegung betrachten. Die obigen Aussagen sind dann Aussagen über endliche Summen, die man sofort einsieht. □

Satz 226 (Standardabschätzung für das Integral einer Treppenfunktion)**:** Für $\varphi \in T([a,b])$ gilt

$$\left| \int_a^b \varphi(x)\, dx \right| \leq \|\varphi\|_{[a,b]}(b-a).$$

4.11. Regelfunktionen

Definition 227: Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt *Regelfunktion*, wenn es eine Folge (φ_n) in $T([a,b])$ gibt, die gleichmäßig gegen f konvergiert.

BEMERKUNG: Es sei daran erinnert, dass nach Definition 140 die gleichmäßige Konvergenz von (φ_n) gegen f gleichbedeutend mit

$$\lim_{n\to\infty} \|\varphi_n - f\| = 0$$

ist, wenn $\|\cdot\|$ die Supremumsnorm ist.

Wir wollen nun beweisen, dass jede auf einem abgeschlossenen Intervall stetige Funktion dort eine Regelfunktion ist. Hierzu führen wir zunächst einen neuen Stetigkeitsbegriff ein.

Definition 228: Es sei $D \subset \mathbb{R}$. Eine Funktion $f : D \to \mathbb{R}$ heißt *gleichmäßig stetig auf* D, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x, x' \in D$ mit $|x - x'| < \delta$ gilt, dass $|f(x) - f(x')| < \varepsilon$ ist. In Quantorenschreibweise heißt das

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, x' \in D : |x - x'| < \delta \Rightarrow |f(x) - f(x')| < \varepsilon.$$

BEMERKUNG: Jede auf D gleichmäßig stetige Funktion ist insbesondere auf ganz D stetig, denn aus Satz 125 folgt, dass f genau dann auf ganz D stetig ist, wenn

$$\forall x' \in D \forall \varepsilon > 0 \exists \delta > 0 \forall x \in D : |x - x'| < \delta \Rightarrow |f(x) - f(x')| < \varepsilon.$$

Da hier das δ von x' abhängen kann, ist dies offensichtlich eine schwächere Forderung an die Funktion f als die gleichmäßige Stetigkeit auf D.

Satz 229: Es sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Dann ist f auf $[a, b]$ gleichmäßig stetig. D.h. jede auf einem abgeschlossenen Intervall stetige Funktion ist dort auch gleichmäßig stetig.

BEWEIS: Wir nehmen an, dass f auf $[a, b]$ nicht gleichmäßig stetig ist. Dann existiert also ein ε_0, so dass es zu jedem $\delta > 0$ reelle Zahlen $x, x' \in [a, b]$ gibt, für die $|x - x'| < \delta$ und $|f(x) - f(x')| \geq \varepsilon_0$ ist.

Zu jeder natürlichen Zahl n gibt es dann also mit $\delta = \frac{1}{n}$ Zahlen $x_n, x'_n \in [a, b]$, für die $|x_n - x'_n| < \frac{1}{n}$ und $|f(x_n) - f(x'_n)| \geq \varepsilon_0$ ist. Insbesondere ist also die Folge $(x_n - x'_n)$ eine Nullfolge.

Die Folgen (x_n) und (x'_n) müssen nicht konvergieren, sie sind aber auf jeden Fall beschränkt, da $a \leq x_n, x'_n \leq b$ für alle $n \in \mathbb{N}$ ist. Nach dem Satz von Bolzano-Weierstraß existieren also konvergente Teilfolgen (x_{n_k}) und (x'_{n_k}). Da die Folge $(x_n - x'_n)$ eine Nullfolge ist, müssen diese beiden Teilfolgen gegen denselben Grenzwert konvergieren. Diese sei mit x_0 bezeichnet.

Wir haben nun auf Grund der Stetigkeit von f

$$0 = f(x_0) - f(x_0) = f(\lim_{k\to\infty} x_{n_k}) - f(\lim_{k\to\infty} x'_{n_k}) = \lim_{k\to\infty} f(x_{n_k}) - \lim_{k\to\infty} f(x_{n_k})$$
$$= \lim_{k\to\infty} f(x_{n_k}) - f(x_{n_k}),$$

was aber der Abschätzung $|f(x_n) - f(x'_n)| \geq \varepsilon_0$ widerspricht. Also ist die Annahme, dass f nicht gleichmäßig stetig ist, nicht haltbar.

□

Beispiel: Wir zeigen nun, dass die Funktion $f :]0,1] \to \mathbb{R}$, die gegeben ist durch $f(x) := \frac{1}{x}$, nicht gleichmäßig stetig auf $]0,1]$ ist. Wir betrachten $|f(x) - f(x')|$. Es ist

$$|f(x) - f(x')| = \left|\frac{1}{x} - \frac{1}{x'}\right| = \left|\frac{x' - x}{xx'}\right|.$$

Sei $\delta > 0$ beliebig. Wir wählen ein x mit $0 < x < \min\{\frac{1}{2}, \delta\}$ und setzen $x' := 2x$. Damit erhalten wir

$$|f(x) - f(x')| = \left|\frac{x}{2x^2}\right| = \frac{1}{2x} > 1.$$

Zu $\varepsilon := 1$ gibt es also für alle $\delta > 0$ Zahlen x und x', so dass $|x - x'| < \delta$, aber $|f(x) - f(x')| > \varepsilon$ ist; also ist f nicht gleichmäßig stetig.

Satz 230: Jede stetige Funktion $f : [a,b] \to \mathbb{R}$ ist eine Regelfunktion.

BEWEIS: Wir müssen eine Folge von Treppenfunktionen konstruieren, die gleichmäßig gegen f konvergiert.

Sei also $\varepsilon > 0$ vorgegeben. Da f auf $[a,b]$ gleichmäßig stetig ist, existiert zu diesem ε ein $\delta > 0$, so dass $|f(x) - f(x')| < \varepsilon$ für alle $x, x' \in [a,b]$ mit $|x - x'| < \delta$ ist.

Für alle $n \in \mathbb{N}$ und $k = 0, \ldots, n$ setzen wir

$$t_k^{(n)} := a + k \cdot \frac{b-a}{n}$$

und definieren für $x \in [t_{k-1}^{(n)}, t_k^{(n)}[$ die Treppenfunktionen φ_n durch

$$\varphi_n(x) := f(t_{k-1}^{(n)})$$

und

$$\varphi_n(b) := f(b).$$

Nach dem Archimedischen Prinzip existiert ein $n_0 \in \mathbb{N}$ so, dass $(b-a)/n_0 < \delta$ ist. Sei nun $n \geq n_0$ und $x \in [a,b]$. Ist $x = b$, so ist $f(x) = \varphi_n(x)$. Sei nun $x \neq b$. Es gibt dann ein $k \in \{1, 2, \ldots, n\}$, so dass $x \in [t_{k-1}^{(n)}, t_k^{(n)}[$ und haben wir dann

$$|f(x) - \varphi_n(x)| = |f(x) - f(t_{k-1}^{(n)})| < \varepsilon,$$

da $|x - t_{k-1}^{(n)}| \leq (b-a)/n \leq (b-a)/n_0 < \delta$ ist. Also konvergiert die Folge (φ_n) gleichmäßig gegen f. □

Definition 231: Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt *stückweise stetig (auf* $[a,b]$*)*, wenn es eine Zerlegung $Z := \{t_0, t_1, \ldots, t_n\}$ von $[a,b]$ gibt, so dass die Funktionen $f_i :]t_{i-1}, t_i[\to \mathbb{R}, x \mapsto f(x)$ für alle $i \in \{1, \ldots, n\}$ stetig und stetig fortsetzbar in t_{i-1} und t_i sind.

Satz 232: Jede stückweise stetige Funktion $f : [a,b] \to \mathbb{R}$ ist eine Regelfunktion.

Problem 81: *Beweisen Sie den vorstehenden Satz.*

4.12. Das Integral von Regelfunktionen

Satz 233: (i) Konvergiert eine Folge (φ_n) von Treppenfunktionen in $T([a,b])$ gleichmäßig gegen eine Funktion $f : [a,b] \to \mathbb{R}$, so konvergiert auch die Folge der Integrale

$$\left(\int_a^b \varphi_n(x)\, dx \right).$$

(ii) Konvergieren zwei Folgen (φ_n) und (ψ_n) von Treppenfunktionen in $T([a,b])$ gleichmäßig gegen dieselbe Funktion $f : [a,b] \to \mathbb{R}$, so ist

$$\lim_{n\to\infty} \int_a^b \varphi_n(x)\, dx = \lim_{n\to\infty} \int_a^b \psi_n(x)\, dx.$$

BEWEIS:

(i) Wir zeigen, dass die Folge der Integrale eine Cauchy-Folge ist. Sei also $\varepsilon > 0$ beliebig, aber fest vorgegeben.

Da (φ_n) gleichmäßig gegen f konvergiert, existiert ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$

$$\|\varphi_n - f\|_{[a,b]} < \varepsilon$$

ist.

Für $m, n \geq n_0$ ist also

$$\|\varphi_n - \varphi_m\|_{[a,b]} = \|\varphi_n - f + f - \varphi_m\|_{[a,b]} \leq \|\varphi_n - f\|_{[a,b]} + \|\varphi_m - f\|_{[a,b]} < 2\varepsilon.$$

Aus der Standardabschätzung für das Integral einer Treppenfunktion folgt hiermit

$$\left|\int_a^b \varphi_n(x)\,dx - \int_a^b \varphi_m(x)\,dx\right| = \left|\int_a^b \varphi_n(x) - \varphi_m(x)\,dx\right| \leq \|\varphi_n - \varphi_m\|_{[a,b]}(b-a) < 2(b-a)\varepsilon.$$

(ii) Wir zeigen, dass die Folge

$$\left(\int_a^b \varphi_n(x)\,dx - \int_a^b \psi_n(x)\,dx\right)$$

eine Nullfolge ist.

Sei also $\varepsilon > 0$ beliebig, aber fest vorgegeben.

Da (φ_n) und (ψ_n) gleichmäßig gegen f konvergieren, existiert ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$

$$\|\varphi_n - f\|_{[a,b]} < \varepsilon$$

und

$$\|\psi_n - f\|_{[a,b]} < \varepsilon$$

ist.

Für $n \geq n_0$ ist also

$$\|\varphi_n - \psi_n\|_{[a,b]} = \|\varphi_n - f + f - \psi_n\|_{[a,b]} \leq \|\varphi_n - f\|_{[a,b]} + \|\psi_n - f\|_{[a,b]} < 2\varepsilon.$$

Damit haben wir nun

$$\left|\int_a^b \varphi_n(x)\,dx - \int_a^b \psi_n(x)\,dx\right| = \left|\int_a^b \varphi_n(x) - \psi_n(x)\,dx\right| \leq \|\varphi_n - \psi_n\|_{[a,b]}(b-a) < 2(b-a)\varepsilon.$$

$\square$

Definition 234: Es sei $f : [a, b] \to \mathbb{R}$ eine Regelfunktion und (φ_n) eine Folge von Treppenfunktionen, die gleichmäßig gegen f konvergiert. Wir setzen

$$\int_a^b f(x)\,dx := \lim_{n\to\infty} \int_a^b \varphi_n(x)\,dx$$

und nennen dies das *(bestimmte) Integral über $f(x)$ von a nach b.* Weiter setzen wir

$$\int_a^a f(x)\,dx := 0$$

und

$$\int_b^a f(x)\, dx := -\int_a^b f(x)\, dx.$$

Beispiel: Wir wollen $\int_0^1 x\, dx$ berechnen. Wir gehen dazu analog zum Beweis, dass jede stetige Funktion eine Regelfunktion ist, vor.

Für $n \in \mathbb{N}$ setzen wir $\varphi_n(x) := k/n$ für $x \in [k/n, (k+1)/n[$, wobei $k \in \{0, 1, \dots, n-1\}$, und $\varphi_n(1) = 1$. Dann ist

$$\int_0^1 \varphi_n(x)\, dx = \sum_{k=0}^{n-1} \frac{k}{n} \cdot \frac{1}{n} = \frac{1}{n^2} \sum_{k=1}^{n-1} k = \frac{1}{n^2} \cdot \frac{(n-1)n}{2}.$$

Hiermit erhalten wir

$$\int_0^1 x\, dx = \lim_{n\to\infty} \int_0^1 \varphi_n(x)\, dx = \frac{1}{2}.$$

Satz 235 (Linearität und Monotonie des Integrals)**:** Sind $f, g : [a, b] \to \mathbb{R}$ Regelfunktionen und $\lambda \in \mathbb{R}$ so gilt:

(i) Die Funktion $f + g$ ist eine Regelfunktion und es gilt

$$\int_a^b f(x) + g(x)\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx.$$

(ii) Die Funktion $\lambda \cdot f$ ist eine Regelfunktion und es gilt

$$\int_a^b \lambda f(x)\, dx = \lambda \int_a^b f(x)\, dx.$$

(iii) Ist $f \leq g$, so gilt

$$\int_a^b f(x)\, dx \leq \int_a^b g(x)\, dx.$$

Problem 82: *Beweisen Sie diesen Satz.*

BEMERKUNG: Die Eigenschaften (i) und (ii) beweisen die Linearität der Integralfunktion und die Eigenschaft (iii) zeigt ihre Monotonie.

Satz 236 (Standardabschätzung)**:** Ist $f : [a,b] \to \mathbb{R}$ eine Regelfunktion, so gilt

$$\left| \int_a^b f(x)\, dx \right| \leq (b-a)\|f\|_{[a,b]}.$$

BEWEIS: Sei (φ_n) eine Folge von Treppenfunktionen, die gleichmäßig gegen f konvergiert. Aus der Dreiecksungleichung für die Supremumsnorm folgert man analog zur normalen Betragsfunktion, dass

$$\left| \|\varphi_n\|_{[a,b]} - \|f\|_{[a,b]} \right| \leq \|\varphi_n - f\|_{[a,b]}$$

ist. Wegen $\lim_{n\to\infty} \|\varphi_n - f\|_{[a,b]} = 0$ können wir hieraus $\lim_{n\to\infty} \|\varphi_n\|_{[a,b]} = \|f\|_{[a,b]}$ folgern.

Aufgrund der Rechenregeln für Grenzwert gilt

$$\left| \lim_{n\to\infty} \int_a^b \varphi_n(x)\, dx \right| = \lim_{n\to\infty} \left| \int_a^b \varphi_n(x)\, dx \right|.$$

Mit der Standardabschätzung für das Integral von Treppenfunktionen können wir also folgern

$$\begin{aligned} \left| \int_a^b f(x)\, dx \right| &= \left| \lim_{n\to\infty} \int_a^b \varphi_n(x)\, dx \right| \\ &= \lim_{n\to\infty} \left| \int_a^b \varphi_n(x)\, dx \right| \\ &\leq \lim_{n\to\infty} (b-a)\|\varphi_n\|_{[a,b]} = (b-a)\|f\|_{[a,b]}. \end{aligned}$$

□

Satz 237: Es sei $f : [a,b] \to \mathbb{R}$ eine Funktion.

(i) Ist $t \in]a,b[$, so ist f genau dann eine Regelfunktion, wenn f auf $[a,t]$ und auf $[t,b]$ eine Regelfunktion ist. Dann gilt

$$\int_a^b f(x)\, dx = \int_a^t f(x)\, dx + \int_t^b f(x)\, dx.$$

(ii) Sind $x_1, \ldots, x_n \in [a,b]$ und ist $g : [a,b] \to \mathbb{R}$, dass $f(x) = g(x)$ für alle $x \in [a,b] \backslash \{x_1, \ldots, x_n\}$ gilt, so ist g genau dann eine Regelfunktion, wenn f eine Regelfunktion ist und es gilt

$$\int_a^b g(x)\, dx = \int_a^b f(x)\, dx.$$

Problem 83: *Beweisen Sie den vorstehenden Satz.*

Problem 84: *Es sei $g : [a, b] \to \mathbb{R}$ eine stetige Funktion mit $g \geq 0$ und*

$$\int_a^b g(x)\,\mathrm{d}x = 0.$$

Zeigen Sie, dass dann $g(x) = 0$ für alle $x \in [a, b]$ ist.

Wir haben bisher in den Beweisen und Beispielen äquidistante Zerlegungen benutzt. Eine größere Flexibilität bei der Berechnung von Integralen erhält man natürlich, wenn man auch Zerlegungen betrachtet, bei denen die Punkte nicht gleichmäßig über das Interval verteilt sind. Wir wollen hierzu ein paar Bezeichnungen einführen und ein allgemeines Ergebnis angeben, das sich mit den Ideen der zuvor bewiesenen Resultate leicht zeigen lässt.

Definition 238 (Feinheit der Zerlegung)**:** Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $Z := \{t_0, t_1, \ldots, t_n\}$ eine Zerlegung von $[a, b]$. Die Zahl

$$\mu(Z) := \max_{j=1}^{n} t_j - t_{j-1}$$

bezeichnet man als die *Feinheit der Zerlegung* Z.

Definition 239 (Riemannsche Summen; Riemann-Folge)**:** Es sei $f : [a, b] \to \mathbb{R}$ eine Funktion und Z eine Zerlegung von $[a, b]$. Weiter seien $\xi_i \in [t_{i-1}, t_i]$ für $i = 1, 2, \ldots, n$ und $\underline{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)$. Dann nennt man

$$R_f(Z, \underline{\xi}) := \sum_{i=1}^{n} f(\xi_i)(t_i - t_{i-1})$$

die Riemannsche Summe der Funktion f bezüglich der Zerlegung Z und der Stützstellen $\underline{\xi}$. Ist (Z_n) eine Folge von Zerlegungen des Intervals $[a, b]$ mit $\lim_{n\to\infty} \mu(Z_n) = 0$ und $(\underline{\xi}_n)$ eine zugehörige Stützstellenfolge, so heißt die Folge der Riemannschen Summen $(R_f(Z_n, \underline{\xi}_n))$ *eine Riemann-Folge von f.*

BEMERKUNG: Man beachte, dass jede Riemannsche Summe das Integral einer Treppenfunktion darstellt und umgekehrt jedes Integral einer Treppenfunktion als Riemannsche Summe geschrieben werden kann. Man kann den Integralbegriff auch über Riemannsche Summen anstelle von Treppenfunktionen einführen. Häufig wird dabei ein etwas allgemeinerer Integralbegriff als das von uns benutzte Regelintegral eingeführt, das sogenannte Riemann-Integral. Die Klasse der integrierbaren Funktionen ist beim Riemann-Integral etwas größer als beim Regelintegral. In der Praxis ist der Unterschied aber unbedeutend. Eine signifikant größere Klasse an integrierbaren Funktionen erhält man mittels des sogenannten Lebesgue-Integrals. Dieses würde aber den Rahmen dieses Buches übersteigen.

Satz 240 (Berechnung von Integralen mittels Riemannschen Summen)**:** Es sei $f : [a,b] \to \mathbb{R}$ eine stetige Funktion. Weiter sei (Z_n) eine Folge von Zerlegungen des Intervalls $[a,b]$ mit

$$\lim_{n\to\infty} \mu(Z_n) = 0$$

und $(\underline{\xi}_n)$ eine zugehörige Stützstellenfolge. Dann gilt

$$\lim_{n\to\infty} R_f(Z_n, \underline{\xi}_n) = \int_a^b f(x)\, dx.$$

Problem 85: *Beweisen Sie diesen Satz. Zeigen Sie dazu, dass die zur Riemann-Folge gehörenden Treppenfunktionen gleichmäßig gegen f konvergieren, indem Sie ähnlich wie im Beweis von 230 vorgehen.*

Beispiel: Als Anwendung dieses Satzes wollen wir das Integral

$$\int_a^b x^\alpha\, dx$$

für $\alpha \neq -1$ und $0 < a < b$ berechnen.

Wir wählen hier eine sogenannte geometrische Zerlegung des Intervals $[a,b]$. Für $n \in \mathbb{N}$ wählen wir die Punkte der Zerlegung Z_n durch $t_i^{(n)} := aq^i$ für $i = 0, \ldots, n$, wobei $q := q(n) := \sqrt[n]{\frac{b}{a}} > 1$ gesetzt sei. Für $i = 1, \ldots, n$ setzen wir weiter $\xi_i^{(n)} := aq^{i-1}$.

Es ist

$$\mu(Z_n) = \max_{i=1}^{n}\{a(q^i - q^{i-1})\} \leq b(q-1).$$

Aus $\lim_{n\to\infty} q(n) = 1$ folgt nun $\lim_{n\to\infty} \mu(Z_n) = 0$.

Es ist

$$\begin{aligned} R_f(Z_n, \underline{\xi}_n) &= \sum_{i=1}^{n} (aq^{i-1})^\alpha (aq^i - aq^{i-1}) = a^{\alpha+1} \sum_{i=1}^{n} (q^{i-1})^\alpha q^{i-1}(q-1) \\ &= a^{\alpha+1}(q-1) \sum_{i=1}^{n} q^{(i-1)(\alpha+1)} = a^{\alpha+1}(q-1) \sum_{i=0}^{n-1} (q^{\alpha+1})^i \\ &= a^{\alpha+1}(q-1) \frac{q^{(\alpha+1)n} - 1}{q^{\alpha+1} - 1}. \end{aligned}$$

Nun ist $q^{(\alpha+1)n} = \left(\frac{b}{a}\right)^{\alpha+1}$ und somit $a^{\alpha+1}(q^{(\alpha+1)n} - 1) = b^{\alpha+1} - a^{\alpha+1}$. Wir haben also

$$R_f(Z_n, \underline{\xi}_n) = (b^{\alpha+1} - a^{\alpha+1}) \frac{q-1}{q^{\alpha+1} - 1}.$$

Wenn wir jetzt n gegen Unendlich laufen lassen, dann dürfen wir nicht vergessen, dass q von n abhängt und dass $\lim_{n\to\infty} q(n) = 1$ ist. Wir müssen also noch

$$\lim_{q\to 1} \frac{q-1}{q^{\alpha+1}-1}$$

bestimmen. Hierzu substituieren wir $t := \log(q)$, also $q = e^t$. Dann ist

$$\lim_{q\to 1} \frac{q-1}{q^{\alpha+1}-1} = \lim_{t\to 0} \frac{e^t-1}{e^{(\alpha+1)t}-1} = \lim_{t\to 0} \frac{e^t-1}{(\alpha+1)t} \lim_{t\to 0} \frac{(\alpha+1)t}{e^{(\alpha+1)t}-1}.$$

Wir haben bereits gesehen, dass

$$\lim_{t\to 0} \frac{e^t-1}{t} = 1$$

ist. Also ist

$$\lim_{q\to 1} \frac{q-1}{q^{\alpha+1}-1} = \frac{1}{\alpha+1} \lim_{t\to 0} \frac{e^t-1}{t} \lim_{t\to 0} \frac{(\alpha+1)t}{e^{(\alpha+1)t}-1} = \frac{1}{\alpha+1}$$

und wir haben

$$\int_a^b x^\alpha \, dx = \lim_{n\to\infty} R_f(Z_n, \underline{\xi}_n) = (b^{\alpha+1} - a^{\alpha+1}) \frac{1}{\alpha+1}.$$

Exkurs: Numerische Berechnung von Integralen

Der vorstehende Satz liefert viele Möglichkeiten, wie man Integrale stetiger Funktionen numerisch berechnen kann. Hier eine ganz einfache Implementierung, die die Approximation

$$\int_a^b f(x) \, dx \approx \frac{b-a}{N} \sum_{k=1}^{N} f(x_k)$$

nutzt. Diese Formel nennt man die *summierte Rechteckregel*.

```
function erg = integral(f,a,b,N)
  %
  % integral(f,a,b) - berechnet das Integral von f zwschen a und b
  % f kann hier bei eine Matlab-Funktion sein, die in der Form
  % @Funktionsname übergeben werden muss, oder
  % f ist eine selbstdefinierte "anonyme" Funktion.

if nargin == 3
    N=100;
end

for k=0:N
    t(k+1)=a+k*(b-a)/N;
end

erg=sum(f(t).*(b-a)/N);
```

Bessere Ergebnisse liefert im Allgemeinen die *summierte Trapezregel*

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2N} f(a) + \frac{b-a}{N} \sum_{k=2}^{N} f(x_k) + \frac{b-a}{2N} f(b).$$

Es gibt aber noch wesentlich aufwendigere Verfahren.

```
function erg = integral(f,a,b,N)
k=1:(N+1);
x=a+(k-1)*(b-a)/N;
erg=0;
for k=2:N
    erg=erg+f(x(k));
end
erg=(b-a)/N*erg+(b-a)/(2*N)*(f(a)+f(b));
```

Ende des Exkurses

4.13. Der Hauptsatz der Differential- und Integralrechnung

Satz 241 (Mittelwertsatz der Integralrechnung)**:** Es seien $f : [a, b] \to \mathbb{R}$ eine stetige Funktionen und $g : [a, b] \to \mathbb{R}$ eine Regelfunktion mit $g \geq 0$. Dann gibt es ein $\zeta \in [a, b]$ mit

$$\int_a^b f(x)g(x)\,dx = f(\zeta) \int_a^b g(x)\,dx.$$

Insbesondere gilt für $g = 1$

$$\int_a^b f(x)\,dx = f(\zeta)(b-a).$$

BEWEIS: Da f stetig ist, nimmt f auf $[a, b]$ Minimum und Maximum an. Wir setzen

$$m := \min\{f(x) | x \in [a, b]\}$$
$$M := \max\{f(x) | x \in [a, b]\}.$$

Da $g \geq 0$ vorrausgesetzt wurde, gilt

$$m \cdot g \leq f \cdot g \leq M \cdot g$$

und nach Satz 235 gilt

$$m \int_a^b g(x)\,dx \leq \int_a^b f(x)g(x)\,dx \leq M \int_a^b g(x)\,dx.$$

Gilt $\int_a^b g(x)\,dx = 0$, so ist auch $\int_a^b f(x)g(x)\,dx = 0$ und die Aussage des Satzes ist trivial. Nehmen wir nun also an, dass $\int_a^b g(x)\,dx \neq 0$, also insbesondere $\int_a^b g(x)\,dx > 0$, so ist

$$\mu = \frac{\int_a^b f(x)g(x)\,dx}{\int_a^b g(x)\,dx} \in [m, M].$$

Nach dem Zwischenwertsatz existiert also ein $\zeta \in [a, b]$ mit $f(\zeta) = \mu$. □

Satz 242 (Hauptsatz der Differential- und Integralrechnung)**:** Es sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Für $x \in [a, b]$ definiere $F : [a, b] \to \mathbb{R}$ durch

$$F(x) := \int_a^x f(t)\,dt.$$

Dann gilt:

(i) Die Funktion F ist eine Stammfunktion von f, das heißt für alle $x \in [a, b]$ gilt $F'(x) = f(x)$.

(ii) Ist G eine beliebige Stammfunktion von f, so gilt

$$\int_a^b f(x)\,dx = G(b) - G(a) =: G(x)|_a^b.$$

Beweis: (ii) folgt aus (i). Ist die Funktion F aus (i) nämlich eine Stammfunktion von f und G eine weitere Stammfunktion von f, so existiert eine Konstante $c \in \mathbb{R}$, so dass $F(x) = G(x) + c$ für alle $x \in [a, b]$ gilt. Also folgt wegen $F(a) = 0$

$$\int_a^b f(x)\,dx = F(b) = F(b) - F(a) = G(b) + c - G(a) - c = G(b) - G(a).$$

(i) Es sei $h \neq 0$. Betrachte den Differenzenquotientent

$$\begin{aligned}\frac{F(x+h) - F(x)}{h} &= \frac{1}{h}\left(\int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt\right)\\ &= \frac{1}{h}\int_x^{x+h} f(t)\,dt.\end{aligned}$$

Nach dem Mittelwertsatz der Integralrechnung 241 existiert ein $\zeta_h \in]x, x+h[$, so dass

$$\int_x^{x+h} f(t)\,dt = f(\zeta_h)(x + h - x) = f(\zeta_h) \cdot h$$

gilt. Außerdem gilt

$$\lim_{h\to 0} \zeta_h = x.$$

Da f stetig ist, erhält man nun

$$\lim_{h\to 0} f(\zeta_h) = f(x).$$

Damit folgt

$$\lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = f(x).$$

□

Exkurs: Das SIR-Modell in der mathematischen Epidemiologie

In diesem Exkurs schauen wir uns ein System von Differentialgleichungen an.

Wir betrachten eine Population, die sich aus drei Teilpopulationen zusammensetzt:

- $S(t)$ bezeichne die Anzahl der gesunden, potentiell infizierbaren Personen in der Population zum Zeitpunkt t. [2]

- $I(t)$ bezeichne die Anzahl der infizierten Personen in der Population zum Zeitpunkt t. Wir gehen davon aus, dass diese ansteckend sind.

- $R(t)$ bezeichne die Anzahl der geheilten und immunen Personen in der Population zum Zeitpunkt t. [3]

Wir modellieren nur das Infektionsgeschehen. Änderungen der Größen der Populationen durch Geburt und Tod sollen vernachlässigt werden.

Das einfache SIR-Modell sieht dann wie folgt aus:

$$\begin{aligned} S'(t) &= -\beta S(t)I(t) \\ I'(t) &= \beta S(t)I(t) - \gamma I(t) \\ R'(t) &= \gamma I(t) \end{aligned}$$

[2] Wir wollen im folgenden die Teilpopulation der gesunden, potentiell infizierbaren Personen ebenfalls mit S bezeichnen; und die anderen beiden Teilpopulationen entsprechend mit I und R.

[3] Man erhält das gleiche Modell, wenn man davon ausgeht, dass die Krankheit tödlich verläuft und $R(t)$ den Anteil der an der Krankheit Verstorbenen bezeichnet.

mit Anfangswerten

$$S(0) = S_0, I(0) = I_0 \text{ und } R(0) = R_0,$$

wobei β und γ positive Konstanten sind. β bezeichnet man als Infektionsrate. γ ist die Genesungsrate.

Man sieht sofort, dass $S'(t)+I'(t)+R'(t) = 0$ ist für alle Zeiten t. Also ist $S(t)+I(t)+R(t)$ für alle Zeiten t konstant, etwa

$$S(t) + I(t) + R(t) = N.$$

Man kann zeigen, dass $1/\gamma$ der mittleren Infektionsdauer entspricht.

Nun wollen wir das Differentialgleichungssystem normalisieren. Dazu teilen wir die obigen Anzahlen $S(t), I(t)$ und $R(t)$ jeweils durch N, d.h. wir gehen von Anzahlen zu relativen Häufigkeiten über. Zum anderen wollen wir Zeit skalieren gemäß der mittleren Infektionsdauer. Wir setzen:

$$\begin{aligned} u(t) &:= \frac{S(t/\gamma)}{N}, \\ v(t) &:= \frac{I(t/\gamma)}{N}, \\ w(t) &:= \frac{R(t/\gamma)}{N} \end{aligned}$$

und

$$u_0 = S_0/N, v_0 = I_0/N \text{ und } w_0 = R_0/N.$$

Das obige Differentialgleichungssystem transformiert sich dann zu

$$\begin{aligned} u'(t) &= -\mathcal{R}u(t)v(t) \\ v'(t) &= \mathcal{R}u(t)v(t) - v(t) = (\mathcal{R}u(t) - 1)v(t) \\ w'(t) &= v(t) \end{aligned}$$

mit den Anfangswerten

$$u(0) = u_0, v(0) = v_0 \text{ und } w(0) = w_0,$$

wobei

$$\mathcal{R} = \frac{\beta N}{\gamma}$$

ist. Diese Größe gibt also an, wie viele Individuen ein Infizierter während der mittleren Infektionsdauer in einer Population mit N infizierbaren Individuen infiziert.

Für $t \geq 0$ ist nach dem Hauptsatz der Differential- und Integralrechnung

$$w(t) = w_0 + \int_0^t v(s)ds.$$

Man kann sich also auf u und v konzentrieren. Offensichtlich kann man $\mathcal{R} > 0$, $u(t), v(t) \geq 0$ und $u(t) + v(t) \leq 1$ für alle Zeiten t annehmen.

Nach der ersten Differentialgleichung ist $u'(t) \leq 0$ und somit u monoton fallend. Da u monoton fallend und nach unten beschränkt ist, existiert $\lim_{t\to\infty} u(t)$.

Es ist

$$\frac{v'(t)}{u'(t)} = -1 + \frac{1}{\mathcal{R}u(t)}.$$

Multipliziert man dies mit $u'(t)$ und integriert, so erhält man

$$v(t) - v_0 = u_0 - u(t) + \mathcal{R}^{-1}(\log u(t) - \log u_0).$$

Wir können nun v in Abhängigkeit von u schreiben:

$$v(u) := c - u + \mathcal{R}^{-1} \log u$$

mit $c := v_0 + u_0 - \mathcal{R}^{-1} \log u_0$.

u und v streben für große Zeiten t gegen einen nahezu konstanten Zustand.

v nimmt also sein Maximum für $u = \mathcal{R}^{-1}$ an. Ist aber $\mathcal{R} \leq 1$, so kann dieser Wert von u nicht angenommen werden. Wir sehen in der zweiten Differentialgleichung, dass in diesem Fall auch $v(t)$ monoton fallend ist.

Ist $\mathcal{R} > 1$ so ist der maximale Wert an Infizierten gleich

$$v_{max} = -\mathcal{R}^{-1}(1 + \log \mathcal{R}) + u_0 + v_0 - \mathcal{R}^{-1} \log u_0.$$

Man sagt, dass sich das Differentialgleichungssystem in einem *Gleichgewichtszustand* befindet, wenn

$$u'(t) = v'(t) = w'(t) = 0$$

gilt. Wir müssen uns beim einfachen SIR-Modell also das Gleichungssystem

$$\begin{aligned} 0 &= -\mathcal{R}u(t)v(t) \\ 0 &= \mathcal{R}u(t)v(t) - v(t) = v(t)(\mathcal{R}u(t) - 1) \\ 0 &= v(t) \end{aligned}$$

anschauen. Es muss offensichtlich $v(t) = 0$ gelten, d.h. es gibt keine Infizierten. Für u gibt es aber zwei Möglichkeiten; es kann $u(t) = 0$ sein oder $u(t) = \mathcal{R}^{-1}$.

Mit dem folgenden Octave/Matlab-Skript kann man die Lösungen des Differentialgleichungssystems näherungsweise lösen:

```
clear all
v0=0.001;
u(1)=1-v0;
v(1)=v0;
w(1)=0;
```

```
h=0.1;
Tend=50;
 R=0.9

for  i=1:(Tend/h)
   u(i+1)=u(i)-h*R*u(i)*v(i);
   v(i+1)=v(i)+h*R*u(i)*v(i)-h*v(i);
   w(i+1)=w(i)+h*v(i);
end
t=0:h:Tend;
figure(45)
clf()
plot(t,u,t,v,t,w)
```

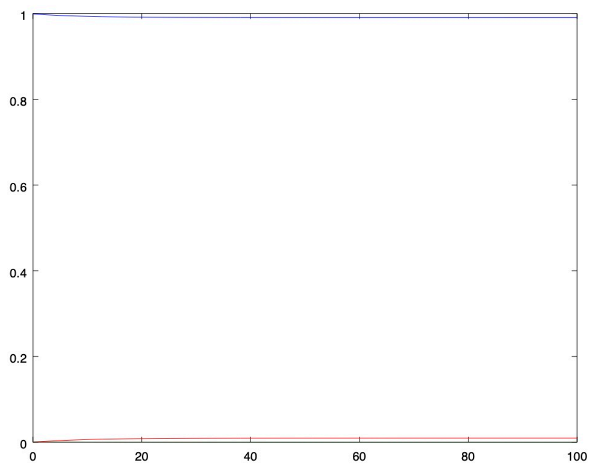


Abbildung 4.20.: SIR-Modell mit $\mathcal{R} = 0.9$ und einem Anteil an Infizierten zum Zeitpunkt $t = 0$ von 0,1$\%$

Unter `http://mysite.science.uottawa.ca/rsmith43/Zombies.pdf` finden Sie einen Artikel von Munz u.a., die ähnliche Modelle für eine Zombie-Apokalypse aufgestellt haben. Sie sollten einiges wieder erkennen. Am Ende des Artikels finden Sie auch Matlab-Programme, die Sie sich ruhig einmal anschauen sollten.

Ende des Exkurses

4.14. Weitere Rechenregeln für Integrale

Die Sätze in diesem Abschnitt ergeben sich aus den bereits behandelten Rechenregeln für unbestimmte Integrale und dem Hauptsatz der Differential- und Integralrechnung.

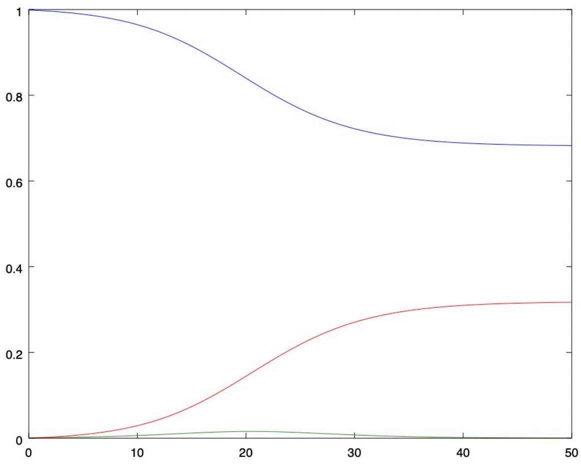


Abbildung 4.21.: SIR-Modell mit $\mathcal{R} = 1.2$ und einem Anteil an Infizierten zum Zeitpunkt $t = 0$ von 0,1$\%$

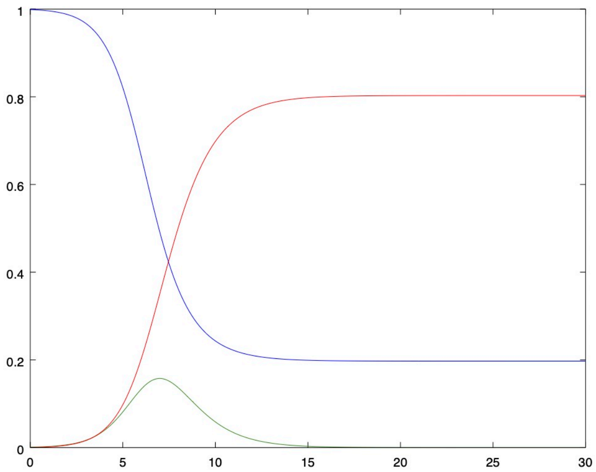


Abbildung 4.22.: SIR-Modell mit $\mathcal{R} = 2$ und einem Anteil an Infizierten zum Zeitpunkt $t = 0$ von 0,1$\%$

Satz 243 (Produktregel; partielle Integration)**:** Es seien $u, v : I \to \mathbb{R}$ stetig differenzierbar, dann ist auch uv stetig differenzierbar, und es gilt

$$\int_a^b u(x)v'(x)\,\mathbf{d}x = u(x)v(x)|_a^b - \int_a^b u'(x)v(x)\,\mathbf{d}x.$$

BEWEIS: Diese Aussage folgt direkt aus dem Satz 212 und dem Hauptsatz der Integral- und Differentialrechnung 242. □

Satz 244 (Substitutionsregel)**:** Es seien $f : I \to \mathbb{R}$ eine Regelfunktion und $g : [a, b] \to I$ eine stetig differenzierbare Funktion. Dann gilt

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(t)\,dt.$$

BEMERKUNG: Dies kann man sich mit folgender, symbolischer Schreibweise merken. Ist $t(x) = g(x)$ so ist $\frac{dt}{dx}(x) = g'(x)$ und damit $dt = g'(x)\,dx$. Läuft x von a nach b, so läuft t von $g(a)$ nach $g(b)$.

BEWEIS: Diese Aussage folgt ebenfalls wieder direkt aus dem Satz 212 und dem Hauptsatz der Integral- und Differentialrechnung 242. ☐

Beispiele:

1) Betrachte das Integral

$$\int_a^b \frac{x}{1+x^2}\,dx.$$

Es ist schon bekannt, dass $\log|g(x)|$ die Stammfunktion von $\frac{g'(x)}{g(x)}$ ist. Setze nun $g(x) = 1 + x^2$ dann gilt $g'(x) = 2x$. Also folgt

$$\int_a^b \frac{x}{1+x^2}\,dx = \frac{1}{2}\int_a^b \frac{2x}{1+x^2}\,dx = \frac{1}{2}\left(\log(1+b^2) - \log(1+a^2)\right).$$

2) Mit Hilfe des Integrals

$$\int_{-r}^{r} \sqrt{r^2 - x^2}\,dx$$

lässt sich der Flächeninhalt eines Halbkreises mit Radius r bestimmen. Verwende die Substitutionsregel und setze dazu $x := r\sin(t)$. Dann ist $dx = r\cos(t)\,dt$ und wenn x von r nach $-r$ läuft, muss $\sin(t)$ von -1 nach 1 laufen. Damit läuft t von $-\frac{\pi}{2}$ nach $\frac{\pi}{2}$ und es gilt

$$\begin{aligned}\int_{-r}^{r} \sqrt{r^2 - x^2}\,dx =& \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{r^2 - r^2\sin^2(t)}r\cos(t)\,dt\\ =&r^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(t)\,dt,\end{aligned}$$

da $1-\sin^2(t)=\cos^2(t)$ ist. Eine Stammfunktion von $\cos^2$ ist schon aus dem Beispiel von Seite 194 bekannt. Damit erhält man

$$\begin{aligned} r^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos^2(t)\,dt =& r^2\frac{1}{2}\left(\cos(t)\sin(t)\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}}+t\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\right)\\ =& r^2\frac{1}{2}\left(\cos\left(\frac{\pi}{2}\right)\sin\left(\frac{\pi}{2}\right)+\frac{\pi}{2}\right)-r^2\frac{1}{2}\left(\cos\left(-\frac{\pi}{2}\right)\sin\left(-\frac{\pi}{2}\right)-\frac{\pi}{2}\right)\\ =& r^2\frac{\pi}{2} \end{aligned}$$

Also ist der Flächeninhalt eines Kreises mit Radius r gleich $2r^2\frac{\pi}{2}=\pi r^2$.

3) Es gilt

$$\int\frac{1}{1+x^2}\,dx=\arctan(x).$$

Nun wird das Integral

$$\int_0^1\frac{1}{1+x+x^2}\,dx$$

berechnet. Der Nenner dieses Integrals lässt sich umformen in

$$\begin{aligned} 1+x+x^2 =& \left(x+\frac{1}{2}\right)^2+\frac{3}{4}\\ =& \frac{3}{4}\left(\left(\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)\right)^2+1\right). \end{aligned}$$

Nun wendet man wieder Substitutionsregel an und setzt $t=\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)$. Damit erhält man $dt=\frac{2}{\sqrt{3}}\,dx$ beziehungsweise $dx=\frac{\sqrt{3}}{2}\,dt$ und somit folgt

$$\begin{aligned} \int_0^1\frac{1}{1+x+x^2}\,dx =& \frac{4}{3}\int_0^1\frac{1}{\left(\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)\right)^2+1}\,dx\\ =& \frac{4}{3}\int_{\frac{1}{\sqrt{3}}}^{\frac{3}{\sqrt{3}}=\sqrt{3}}\frac{1}{t^2+1}\cdot\frac{\sqrt{3}}{2}dt\\ =& \frac{2}{\sqrt{3}}\arctan(t)\Big|_{\frac{1}{\sqrt{3}}}^{\sqrt{3}} \end{aligned}$$

Exkurs: Ein alternativer Zugang zum Logarithmus und zur Exponentialfunktion

Wir haben in diesem Buch zunächst die Exponentialfunktion über ihre Potenzreihe definiert. Dies war schon zu einem sehr frühen Zeitpunkt möglich, nachdem wir uns mit

Folgen und Reihen beschäftigt hatten. Nachdem wir dann stetige Funktionen betrachtet haben, haben wir die Logarithmusfunktion als Umkehrfunktion der Exponentialfunktion eingeführt.

Hier werden wir einen alternativen Zugang skizzieren, den man auch auch sehr häufig in der Literatur findet.

Für $x > 0$ definieren wir die Funktion $L :]0, \infty] \to \mathbb{R}$ durch

$$L(x) := \int_1^x \frac{1}{t} dt.$$

Aus dem Hauptsatz der Differential- und Integralrechnung (Satz 242) ergibt sich sofort, dass die Funktion L differenzierbar (und damit auch stetig) ist und dass

$$L'(x) = \frac{1}{x}.$$

Da offensichtlich $L(1) = 0$ ist, stimmt diese Funktion also mit der von uns bereits früher definierten Logarithmusfunktion überein.

Da die Ableitung $L'(x)$ offensichtlich für alle $x > 0$ positiv ist, ist die Funktion L streng monoton wachsend und besitzt somit eine differenzierbare Umkehrfunktion E.

Wie sieht der Wertebereich von L aus?

Mit der oben berechneten Ableitung von L (oder mittels der Substitutionsregel (Satz 244)) zeigt man leicht, dass für alle $x, y > 0$ und $\alpha \in \mathbb{R}$

$$L(x^\alpha) = \alpha L(x)$$

und

$$L(xy) = L(x) + L(y)$$

gilt. Man definiere etwa eine Hilfsfunktion $H(x) = L(x^\alpha) - \alpha L(x)$ bzw. $H(x) = L(xy) - L(x) - L(y)$ und zeige dass diese identisch verschwindet, indem man in beiden Fällen $H'(x) = 0$ für alle $x > 0$ und $H(1) = 0$ nachrechnet.

Aufgrund der Monotonie von L und $L(1) = 0$, ist $L(1/2) < 0$ und $L(2) > 0$. Damit erhalten wir

$$\lim_{n\to\infty} L(1/2^n) = \lim_{n\to\infty} nL(1/2) = -\infty$$

und

$$\lim_{n\to\infty} L(2^n) = \lim_{n\to\infty} nL(2) = +\infty.$$

Wir haben also $L(\mathbb{R}_+) = \mathbb{R}$ und damit $E : \mathbb{R} \to \mathbb{R}_+$.

Sollte man die Exponentialfunktion und den Logarithmus unabhängig voneinander, d.h. die eine Funktion nicht als Umkehrfunktion der anderen, eingeführt haben, so kann man leicht diesen Zusammenhang nachweisen, indem man

$$L(e^x) = x$$

nachweist. Hierzu betrachte man die Hilfsfunktion $H(x) = L(e^x) - x$ und gehe wie oben vor.

Ende des Exkurses

4.15. Uneigentliche Integrale

Bisher ist das Integral nur für Funktionen definiert worden, die auf einem abgeschlossenen Intervall $[a, b]$ beschränkt sind. In diesem Kapitel werden wir die Definition ausweiten und drei Fälle sogenannter uneigentlicher Integrale behandeln.

I. Eine Integrationsgrenze ist unendlich.

Definition 245: Es sei $f : [a, \infty[\to \mathbb{R}$ eine Funktion, die auf jedem Intervall $[a, R]$, $a < R < \infty$, eine Regelfunktion ist. Dann setze

$$\int_a^\infty f(x)\,dx := \lim_{R\to\infty} \int_a^R f(x)\,dx,$$

falls dieser Grenzwert existiert. Analoge Definition gelte für $\int_{-\infty}^b f(x)\,dx$ mit $f :]-\infty, b] \to \mathbb{R}$.

Beispiel: Das Integral

$$\int_1^\infty \frac{1}{x^s}\,dx$$

existiert genau für alle $s > 1$. Es gilt nämlich

$$\int_1^R \frac{1}{x^s}\,dx = \frac{1}{-s+1} x^{-s+1}\Big|_1^R = \frac{1}{1-s}\left(\frac{1}{R^{s-1}} - 1\right).$$

Grenzübergang liefert dann

$$\int_1^\infty \frac{1}{x^s}\,dx = \lim_{R\to\infty} \frac{1}{1-s}\left(\frac{1}{R^{s-1}} - 1\right) = \frac{1}{s-1}$$

falls $s > 1$ ist.

Problem 86: *Zeigen Sie, dass das Integral im vorstehenden Beispiel für $s = 1$ und $s < 1$ nicht existiert.*

II. Der Integrand ist an einer Grenze nicht definiert.

Definition 246: Es sei $f :]a, b] \to \mathbb{R}$ eine Funktion, die auf jedem Intervall $[a + \varepsilon, b]$ mit $0 < \varepsilon < b - a$ eine Regelfunktion ist. Dann setzt man

$$\int_a^b f(x)\,dx = \lim_{\varepsilon \searrow 0} \int_{a+\varepsilon}^b f(x)\,dx,$$

falls dieser Grenzwert existiert. Analog definiert man $\int_a^b f(x)\,dx$ für $f : [a, b[\to \mathbb{R}$.

Beispiel: Das Integral

$$\int_0^1 \frac{1}{x^s}\,dx$$

existiert genau für $s < 1$. Es gilt

$$\int_\varepsilon^1 \frac{1}{x^s}\,dx = \frac{1}{1-s}\left(1 - \varepsilon^{1-s}\right)$$

und damit folgt für $s < 1$

$$\int_0^1 \frac{1}{x^s}\,dx = \frac{1}{1-s}.$$

III. Beide Integrationsgrenzen sind kritisch

Definition 247: Es sei $f :]a, b[\to \mathbb{R}$ mit $a \in \mathbb{R} \cup \{-\infty\}$ und $b \in \mathbb{R} \cup \{\infty\}$ eine Funktion, die auf jedem Intervall $[\alpha, \beta] \subset]a, b[$ eine Regelfunktion ist. Es sei $c \in]a, b[$. Dann definiert man

$$\int_a^b f(x)\,dx := \int_a^c f(x)\,dx + \int_c^b f(x)\,dx,$$

falls die beiden rechten Integrale existieren.

BEMERKUNG: Man muss zeigen, dass diese Definition unabhängig von der Wahl von c ist.

Beispiele:

1) Nach den obigen Beispielen existiert das Integral

$$\int_0^\infty \frac{dx}{x^s}$$

für kein $s \in \mathbb{R}$.

2) Untersuche das uneigentliche Integral

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2}.$$

Es gilt

$$\int_{-R}^{0} \frac{dx}{1+x^2} = \arctan(x)|_{-R}^{0} = 0 - \arctan(-R) \xrightarrow[R\to\infty]{} -\left(-\frac{\pi}{2}\right)$$

und

$$\int_{0}^{R} \frac{dx}{1+x^2} = \arctan(x)|_{0}^{R} = \arctan(R) \xrightarrow[R\to\infty]{} \frac{\pi}{2}$$

womit insgesamt

$$\int_{-\infty}^{\infty} \frac{ds}{1+x^2} = \pi$$

gilt.

3) Wir betrachten bei festen $x > 0$

$$\int_{0}^{\infty} t^{x-1}e^{-t}dt.$$

Der Integrand ist für $t = 0$ nicht definiert. Da sich die Exponentialfunktion gegen jede Polynomfunktion durchsetzt, ist $\lim_{t\to\infty} t^{x+1}e^{-t} = 0$. Für große t hat man also $t^{x-1}e^{-t} \leq t^{-2}$ und folglich existiert das Integral $\int_1^\infty t^{x-1}e^{-t}dt$. Für alle $t \geq 0$ ist $t^{x-1}e^{-t} \leq t^{x-1} = \frac{1}{t^{1-x}}$. Also existiert auch $\int_0^1 t^{x-1}e^{-t}dt$. Die durch

$$\Gamma(x) := \int_{0}^{\infty} t^{x-1}e^{-t}dt.$$

definierte Funktion $\Gamma : \mathbb{R}_+ \to \mathbb{R}$ heißt *Gamma-Funktion*. Aus der folgenden Aufgabe folgt, dass $\Gamma(n+1) = n!$ ist, und damit stellt die Gamma-Funktion eine Fortsetzung der Fakultät-Funktion dar.

Problem 87: *Für die Gamma-Funktion zeige man*

(i) $\Gamma(1) = 1$.

(ii) Für alle $x > 0$ *gilt* $\Gamma(x+1) = x\Gamma(x)$.
Tipp: Partielle Integration.

(iii) Wie lässt sich die Funktionalgleichung aus Teil (ii) nutzen, um Γ *auf* $\mathbb{R}\backslash\{0, -1, -2, -3, \ldots\}$ *fortzusetzen?*

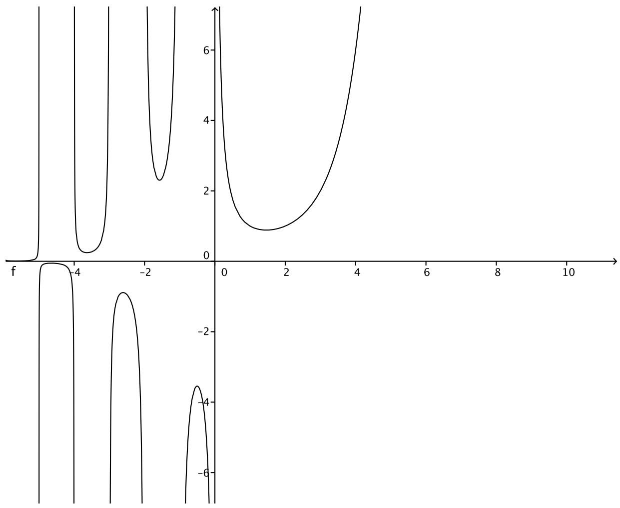


Abbildung 4.23.: Graph der Gamma-Funktion

Satz 248 (Integralkriterium für Reihen)**:** Es sei $f : [1, \infty[\to \mathbb{R}_+ \cup \{0\}$ eine monoton fallende Funktion. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} f(n)$ genau dann, wenn das uneigentliche Integral $\int_1^{\infty} f(x)\, dx$ existiert.

BEWEIS: Da f nach Vorraussetzung monoton fallend ist, gilt

$$f(n) \leq f(x) \leq f(n-1)$$

für $n - 1 \leq x \leq n$. Daraus folgt

$$f(n) \leq \int_{n-1}^{n} f(x)\, dx \leq f(n-1).$$

Aufsummieren ergibt dann

$$\sum_{n=2}^{N} f(n) \leq \int_1^N f(x)\, dx \leq \sum_{n=1}^{N-1} f(n).$$

Existiert der Ausdruck $\int_1^{\infty} f(x)\, dx$, so ist $\sum_{n=2}^{N} f(n)$ beschränkt und konvergiert somit. Konvergiert andererseits die Reihe $\sum_{n=1}^{\infty} f(n)$, so ist $\int_1^R f(x)\, dx$ für $\mathbb{R} \to \infty$ monoton wachsend und beschränkt, woraus die Konvergenz folgt. $\square$

Beispiel (Riemannsche ζ-Funktion)**:** Es wurde gezeigt, dass $\int_1^{\infty} \frac{dx}{x^s}$ für $s > 1$ existiert, also existiert auch die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^s}$ für $s > 1$. Die Funktion $\zeta :]1, \infty[\to \mathbb{R}$ mit

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}$$

heißt *Riemannsche ζ-Funktion*. Die Riemannsche ζ-Funktion spielt unter anderem eine ganz wesentliche Rolle in der modernen Zahlentheorie, wenn es um Untersuchungen von Primzahlen geht. Den Zusammenhang zwischen der Riemannschen ζ-Funktion und den Primzahlen erkannte als erstes Euler, der folgende alternative Darstellung der Funktion fand,

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1},$$

wobei sich das Produkt über alle Primzahlen p erstreckt[4].

Die Riemannsche ζ-Funktion spielt eine wichtige Rolle bei vielen Beweisen des sogenannten Primzahlsatzes. Schon Gauß hatte vermutet, dass die Anzahl $\pi(x)$ aller Primzahlen, die kleiner oder gleich einer vorgegebenen Zahl x sind, also

$$\pi(x) := \#\{p \text{ prim} \mid p \leq x\},$$

für große x gut durch den Ausdruck $x/\log(x)$ beschrieben werden kann. Es dauerte aber noch etwa 100 Jahre bis unabhängig voneinander Hadamard und de La Vallée Poussin zeigen konnten, dass

$$\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\log(x)}} = 1$$

ist.

In diesen Beweisen benutzt man aber nicht die oben angegebene Funktion einer reellen Veränderlichen, sondern man setzt die Riemannsche ζ-Funktion ins Komplexe fort. Die obige Reihe konvergiert für alle $s \in \mathbb{C}$ mit $\mathrm{Re}(s) > 1$. Man kann aber sogar zeigen, dass man die durch diese Reihe gegebene Funktion auf ganz $\mathbb{C} \setminus \{1\}$ fortsetzen kann. Man interessiert sich dann für möglichst große Gebiete von $\mathbb{C}$, wo die Riemannsche ζ-Funktion keine Nullstellen hat.

Die Nullstellen der Riemannschen ζ-Funktion führen uns dann zu einer der berühmtesten Vermutungen in der Mathematik, der sogenannten *Riemannschen Vermutung*. Man kann zeigen, dass $\zeta(-2n) = 0$ für alle $n \in \mathbb{N}$. Diese Nullstellen nennt man auch die trivialen Nullstellen der Riemannschen ζ-Funktion. Für alle weiteren Nullstellen der Riemannschen ζ-Funktion weiß man, dass sie im sogenannten kritischen Streifen

$$\{z \in \mathbb{C} \mid 0 < \mathrm{Re}(z) < 1\}$$

liegen. Die Riemannsche Vermutung besagt nun, dass alle nichttrivialen Nullstellen der Riemannschen ζ-Funktion sogar auf der kritischen Gerade

$$\{z \in \mathbb{C} \mid \ \mathrm{Re}(z) = \frac{1}{2}\}$$

liegen.

[4] Mit unendlichen Produkten haben wir uns nicht beschäftigt. Man kann Ihnen auf ähnliche Art einen Sinn geben, wie wir das bei Reihen gemacht haben.

4.16. Volumen und Länge eines Graphen

Die Einführung des Integrals war motiviert durch den Wunsch einen vernünftigen Flächeninhaltsbegriff zu haben. In diesem Kapitel werden wir uns überlegen, dass man mit Hilfe des Integrals auch Volumen gewisser dreidimensionaler Körper und die Länge des Graphen von reellwertigen Funktionen definieren kann.

Wir wollen zunächst eine vernünftige Definition für das Volumen eines Körpers im dreidimensionalen Raum herleiten, wenn wir den Flächeninhalt aller Querschnittsflächen, die etwa orthogonal zur x-Achse sind, kennen. Ein Körper der Höhe h, dessen Querschnitts-

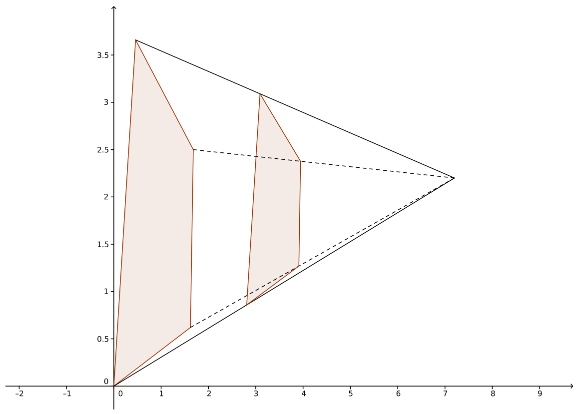


Abbildung 4.24.: Ein dreidimensionaler Körper mit hervorgehobener Querschnittsfläche

flächen auf jeder Höhe den selben Flächeninhalt F haben, hat das Volumen hF. Haben wir nun einen Körper, bei dem der Flächeninhalt der Querschnittsfläche in der x-Richtung durch eine stetige Funktion $F : [a, b] \to \mathbb{R}$ gegeben ist, so können wir das Volumen des Körpers durch

$$\sum_{i=0}^{n-1} (x_{i+1} - x_i) F(x_i)$$

approximieren, wenn $\{x_0, x_1, \dots, x_n\}$ eine Zerlegung des Intervalls $[a, b]$ ist. Je größer n ist, desto besser ist die Annäherung des gesuchten Volumens. Bei der Summe $\sum_{i=0}^{n-1}(x_{i+1} - x_i)F(x_i)$ handelt es sich um eine Riemannsche Summe der Funktion F.

Definition 249: Ist in einem dreidimensionalen Körper der Flächeninhalt der Querschnittsfläche in der x-Richtung durch eine stetige Funktion $F : [a, b] \to \mathbb{R}$ gegeben, so definieren wir das Volumen V dieses Körpers durch

$$V := \int_a^b F(x) dx.$$

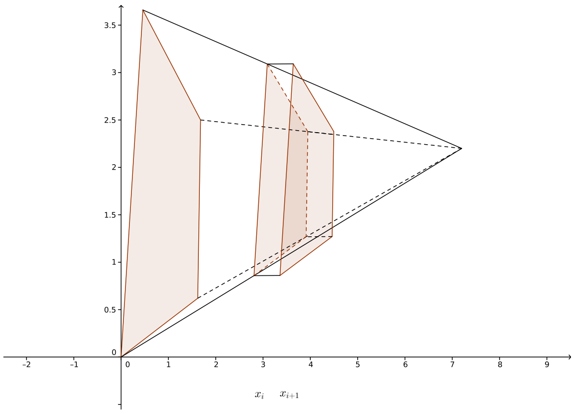


Abbildung 4.25.: Approximation des Volumens eines dreidimensionaler Körper

Satz 250 (Volumen eines Rotationskörpers)**:** Entsteht ein Körper im dreidimensionalen Raum durch Drehung einer Kurve $y = f(x)$, $a \leq x \leq b$, um die x-Achse, so ist das Volumen V des Rotationskörpers gegeben durch

$$V = \pi \int_a^b f(x)^2 dx.$$

Problem 88: *Beweisen Sie den vorstehenden Satz.*

Nun wollen wir eine vernünftige Definition für die Länge des Graphen einer Funktion $f : [a, b] \to \mathbb{R}$ herleiten. Hierzu sei $\{x_0, x_1, \dots, x_n\}$ wieder eine Zerlegung des Intervalls $[a, b]$ und wir verbinden die Punkte $(x_i, f(x_i)), (i = 0, 1, \dots, n)$, mit einem Polygonzug. Das Geradenstück zwischen den Punkten $(x_i, f(x_i))$ und $(x_{i+1}, f(x_{i+1}))$ hat dann nach Pythagoras die Länge $\sqrt{(x_{i+1} - x_i)^2 + (f(x_{i+1}) - f(x_i))^2}$. D.h. die Länge des Graphen der Funktion f wird durch

$$\sum_{i=0}^{n-1} \sqrt{(x_{i+1} - x_i)^2 + (f(x_{i+1}) - f(x_i))^2} = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \sqrt{1 + \left(\frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \right)^2}$$

approximiert. Nach dem Mittelwertsatz der Differentialrechnung existieren $\xi_i \in]x_i, x_{i+1}[$, so dann

$$f'(\xi_i) = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}$$

ist, und die obige Summe lässt sich schreiben

$$\sum_{i=0}^{n-1} \sqrt{(x_{i+1} - x_i)^2 + (f(x_{i+1}) - f(x_i))^2} = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \sqrt{1 + (f'(\xi_i))^2}.$$

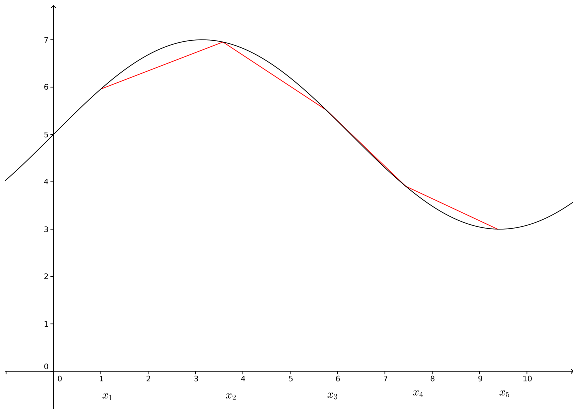


Abbildung 4.26.: Graph einer Funktion mit einem die Länge des Graphen approximierenden Polygonzug

Dies ist eine Riemannsche Summe der Funktion $\sqrt{1+(f'(x))^2}$.

Definition 251: Die Länge L des Graphen einer stetig differenzierbaren Funktion $f : [a,b] \to \mathbb{R}$ ist gegeben durch

$$L := \int_a^b \sqrt{1+(f'(x))^2}\, dx.$$

4.17. Funktionenfolgen II: Integrierbarkeit und Differenzierbarkeit

Satz 252: Es seien $f_n : [a,b] \to \mathbb{R}$ mit $n \in \mathbb{N}$ stetige Funtionen, die gleichmäßig gegen eine Funktion $f : [a,b] \to \mathbb{R}$ konvergieren. Dann gilt

$$\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b f(x)\, dx.$$

BEMERKUNG: Die folgende Aufgabe zeigt, dass dieser Satz in dieser Form leider wirklich nur für bestimmte Integrale gültig ist.

Problem 89: *Die Funktionen* $f_n : \mathbb{R}_{\geq 0} \to \mathbb{R}, n \in \mathbb{N}$, *seien durch* $f_n(x) := \frac{x}{n^2} e^{-x/n}$ *definiert. Zeigen Sie:*

(i) *Die Funktionenfolge* (f_n) *konvergiert gleichmäßig gegen die Nullfunktion.*

(ii) *Es gilt*

$$\lim_{n\to\infty} \int_0^\infty f_n(x) dx = 1.$$

BEWEIS: Zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$, so dass für alle $x \in [a, b]$ die Ungleichung $|f_n(x) - f(x)| < \varepsilon$ gilt. Insbesondere gilt also

$$\|f_n - f\|_{[a,b]} < \varepsilon.$$

Mit Hilfe der Standardabschätzung für Integrale folgt

$$\begin{aligned} \left| \int_a^b f_n(x)\, dx - \int_a^b f(x)\, dx \right| &\leq (b-a)\|f_n - f\|_{[a,b]} \\ &< (b-a)\varepsilon. \end{aligned}$$

Da man ε beliebig klein wählen kann, folgt die Behauptung des Satzes. □

Beispiel: Wir leiten nun aus der geometrischen Reihe die Potenzreihendarstellung des Logarithmus her. Es sei $x \in]-1, 1[$. Wegen der Sätze 142 und 252 gilt

$$\int_0^x \sum_{k=0}^\infty t^k\, dt = \int_0^x \lim_{n\to\infty} \sum_{k=0}^n t^k\, dt = \lim_{n\to\infty} \int_0^x \sum_{k=0}^n t^k\, dt = \sum_{k=0}^\infty \frac{1}{k+1} t^{k+1} \Big|_0^x = \sum_{k=1}^\infty \frac{x^k}{k}.$$

Andererseits gilt

$$\sum_{k=0}^\infty t^k = \frac{1}{1-t}.$$

Eine Stammfunktion von $\frac{1}{1-t}$ ist gegeben durch

$$\int_0^t \frac{1}{1-x}\, dx = -\log(1-t).$$

Hieraus erhält man die Potenzreihendarstellung

$$-\log(1-x) = \sum_{k=1}^\infty \frac{x^k}{k}$$

beziehungsweise die Potenzreihendarstellung des Logarithmus um den Entwicklungspunkt 1

$$\log(x) = \sum_{k=1}^\infty (-1)^{k+1} \frac{(x-1)^k}{k}.$$

Satz 253: Es seien $f_n : [a,b] \to \mathbb{R}$ mit $n \in \mathbb{N}$ stetig differenzierbare Funtionen, die punktweise gegen eine Funktion $f : [a,b] \to \mathbb{R}$ konvergieren. Die Folge der Ableitungen $f'_n : [a,b] \to \mathbb{R}$ konvergiere gleichmäßig. Dann ist die Grenzfunktion f differenzierbar und es gilt für alle $x \in [a,b]$

$$f'(x) = \lim_{n\to\infty} f'_n(x).$$

BEWEIS: Wir definieren die Funktion $g : [a,b] \to \mathbb{R}$ durch

$$g(x) := \lim_{n\to\infty} f'_n(x).$$

für alle $x \in [a,b]$. Es ist zu zeigen, dass $f'(x) = g(x)$ für alle $x \in [a,b]$ ist.

Nach Satz 141 ist die Funktion g stetig. Nach dem Hauptsatz der Differential- und Integralrechnung gilt für alle $x \in [a,b]$

$$f_n(x) = f_n(a) + \int_a^x f'_n(t)dt.$$

Nach Satz 252 konvergiert $\int_a^x f'_n(t)dt$ gegen $\int_a^x g(t)dt$, wenn n gegen unendlich geht. Also erhält man

$$f(x) = f(a) + \int_a^x g(t)dt.$$

Leitet man dies nun ab, so erhält man wiederum aus dem Hauptsatz der Differential- und Integralrechnung $f'(x) = g(x)$.

□

Kombiniert man diesen Satz mit Satz 142, so erhält man (im Wesentlichen)

Satz 254: Es sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann hat auch die Potenzreihe

$$\sum_{n=1}^{\infty} n a_n x^{n-1}$$

den Konvergenzradius R und die Funktion

$$f :]-R,R[\to \mathbb{R}$$
$$x \mapsto \sum_{n=0}^{\infty} a_n x^n$$

ist differenzierbar mit Ableitung

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

Potenzreihen dürfen also gliedweise differenziert werden.

Problem 90: *Zeigen Sie, dass die Potenzreihe $\sum_{n=1}^{\infty} n a_n x^{n-1}$ den selben Konvergenzradius wie die Reihe $\sum_{n=0}^{\infty} a_n x^n$ hat.*

Beispiele:

1)

$$\begin{aligned}
\sin'(x) &= \left(\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \right)' \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} (2n+1) x^{2n} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \\
&= \cos(x)
\end{aligned}$$

2)

$$\begin{aligned}
\cos'(x) &= \left(\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \right)' \\
&= \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!} 2n x^{2n-1} \\
&= \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n-1)!} x^{2n-1} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+1)!} x^{2n+1} \\
&= -\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \\
&= -\sin(x)
\end{aligned}$$

Korollar 255: Durch Potenzreihen definierte Funktionen sind beliebig oft stetig differenzierbar. Insbesondere sind also $\exp, \sin, \cos \in C^{\infty}(\mathbb{R})$.

BEWEIS: Der Beweis erfolgt durch beliebig ofte Anwendung von Satz 254. □

Beispiel: Ausgehend von der für $|x| < 1$ gültigen Formel

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$$

wollen wir nun weitere Potenzreihenentwicklungen herleiten. Wenn wir beide Seiten der Gleichung ableiten, erhalten wir

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

Erneutes Ableiten liefert

$$\frac{2}{(1-x)^3} = \sum_{k=1}^{\infty} (k+1)kx^{k-1} = \sum_{k=0}^{\infty} (k+2)(k+1)x^k.$$

Allgemein erhält man für $n \geq 0$

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

beziehungsweise

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

Exkurs: Lösen von Differentialgleichungen mit Potenzreihen

Man kann sich die Tatsache, dass Potenzreihen gliedweise abgeleitet werden dürfen, zunutze machen, um Differentialgleichungen zu lösen. Die Methode ist im Allgemeinen recht aufwendig und eigentlich nur das letzte Mittel, wenn andere Methoden nicht zur Verfügung stehen. Natürlich können einem entsprechende Computerprogramme die Arbeit auch hier erleichtern. Wir wollen die Methode an der folgenden nicht-linearen Differentialgleichung erster Ordnung demonstrieren.

Gesucht sei die Lösung des Anfangswertproblems

$$y' = y^2 + x$$

mit $y(1) = 2$.

Wir nehmen nun an, dass die Lösung durch eine Potenzreihe um den Entwicklungspunkt 1 gegeben ist. Wir machen also den Ansatz

$$y(x) = \sum_{n=0}^{\infty} a_n (x-1)^n.$$

Wegen der Anfangswertbedingung haben wir $a_0 = y(1) = 2$. Wir kennen also schon den ersten Koeffizienten.

Da wir Potenzreihen gliedweise ableiten dürfen, haben wir

$$y'(x) = \sum_{n=1}^{\infty} n a_n (x-1)^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} (x-1)^n.$$

Das Cauchy-Produkt liefert uns

$$y^2(x) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k a_{n-k} (x-1)^n.$$

Hiermit erhalten wir für die für die rechte Seite der Differentialgleichung

$$\begin{aligned} y^2(x) + x &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k a_{n-k} (x-1)^n + x - 1 + 1 \\ &= a_0^2 + 1 + (2a_0 a_1 + 1)(x-1) + \sum_{n=2}^{\infty} \sum_{k=0}^{n} a_k a_{n-k} (x-1)^n. \end{aligned}$$

Aufgrund der Differentialgleichung $y' = y^2 + x$ haben wir also

$$\sum_{n=0}^{\infty} (n+1) a_{n+1} (x-1)^n = a_0^2 + 1 + (2a_0 a_1 + 1)(x-1) + \sum_{n=2}^{\infty} \sum_{k=0}^{n} a_k a_{n-k} (x-1)^n.$$

Nach dem Identitätssatz für Potenzreihen müssen die Koeffizienten der beiden Reihen übereinstimmen, d.h. man hat

$$\begin{aligned} a_1 &= a_0^2 + 1 = 5 \\ 2a_2 &= 2a_0 a_1 + 1 = 21, \text{ also } a_2 = \frac{21}{2} \\ 3a_3 &= 2a_2 a_0 + a_1^2 = 42 + 25 = 67, \text{ also } a_3 = \frac{67}{3} \end{aligned}$$

und allgemein

$$(n+1) a_{n+1} = \sum_{k=0}^{n} a_k a_{n-k} \text{ für alle } n \geq 3.$$

So kann man jetzt sukzessive die weiteren Koeffizienten berechnen. Dies kann ziemlich mühsam sein. Wenn man Glück hat, dann erkennt man ein Bildungsgesetz. Oder man hat zumindest den Reihenanfang, den man als Approximation der Lösung ansehen kann.

Ende des Exkurses

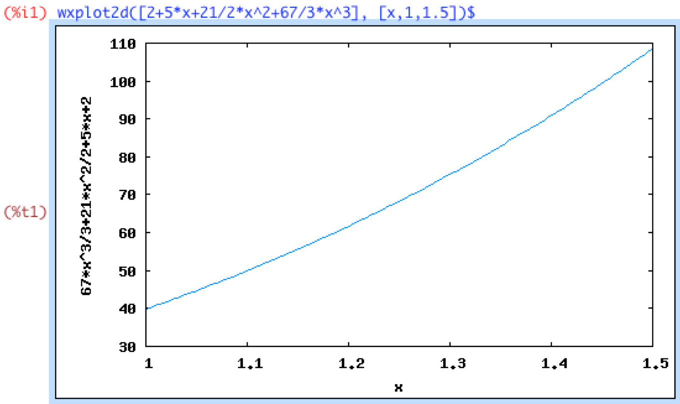


Abbildung 4.27.: Die ersten Glieder der Potenzreihe im Maxima-Plot

4.18. Die Stirlingsche Formel und das Wallissche Produkt

In den meisten Büchern zur Analysis wird die Stirlingsche Formel, die wie in diesem Abschnitt herleiten wollen, mit Hilfe der Gamma-Funktion hergeleitet. Wir nehmen hier einen etwas anderen Weg, da das hier benutze Hilfsmittel des Satz von der partiellen Summation bzw. der Eulerschen Summenformel auch für sich genommen sehr hübsch ist und bei anderen Fragestellungen eingesetzt werden kann. Der folgende Satz etwa stellt eine Verbindung zwischen einem diskreten Objekt (einer endlichen Summe) und einem analytischen Objekt (einem Integral) her. Der Name „Satz von der partiellen Summation“ erinnert natürlich an die partielle Integration und in der Tat weisen die auftretenden Formeln auch eine sehr starke Ähnlichkeit auf, wenn man die Summenfunktion $A(t)$ als Stammfunktion der Folge a_n interpretiert.

Satz 256 (Satz über die partielle Summation)**:** Es seien (a_n) eine Folge reeller Zahlen, (t_n) eine streng monoton wachsende, unbeschränkte Folge reeller Zahlen und $A(t)$ die Summe über diejenigen a_n, deren Indizes n der Bedingung $t_n \le t$ genügen. Ist dann $g : [t_1, \infty[\to \mathbb{R}$ stetig differenzierbar, so gilt für alle reellen $x \ge t_1$

$$\sum_{\substack{n \\ t_n \le x}} a_n g(t_n) = A(x)g(x) - \int_{t_1}^{x} A(t)g'(t)dt.$$

BEWEIS: Es sei $x \geq t_1$. Dann existiert ein N, so dass $t_N \leq x < t_{N+1}$ ist. Für alle t mit $t_n \leq t < t_{n+1}$ ist stets $A(t) = A(t_n)$ ist. Damit erhalten wir nun

$$\begin{aligned}
\int_{t_1}^{x} A(t)g'(t)dt &= \sum_{n=1}^{N-1} \int_{t_n}^{t_{n+1}} A(t)g'(t)dt + \int_{t_N}^{x} A(t)g'(t)dt \\
&= \sum_{n=1}^{N-1} \int_{t_n}^{t_{n+1}} A(t_n)g'(t)dt + \int_{t_N}^{x} A(t_N)g'(t)dt \\
&= \sum_{n=1}^{N-1} A(t_n) \int_{t_n}^{t_{n+1}} g'(t)dt + A(t_N) \int_{t_N}^{x} g'(t)dt \\
&= \sum_{n=1}^{N-1} A(t_n)\left(g(t_{n+1}) - g(t_n)\right) + A(t_N)\left(g(x) - g(t_N)\right) \\
&= \sum_{n=2}^{N} A(t_{n-1})g(t_n) - \sum_{n=1}^{N} A(t_n)g(t_n) + A(t_N)g(x) \\
&= \sum_{n=2}^{N} (A(t_{n-1}) - A(t_n))g(t_n) - A(t_1)g(t_1) + A(t_N)g(x)
\end{aligned}$$

Nun nutzen wir aus, dass $A(t_n) - A(t_{n-1}) = a_n$ für $n \geq 2$ ist. Außerdem ist $A(t_N) = A(x)$ und $A(t_1) = a_1$. Wir haben also

$$\int_{t_1}^{x} A(t)g'(t)dt = -\sum_{n=2}^{N} a_n g(t_n) - a_1 g(t_1) + A(x)g(x),$$

woraus die behauptete Gleichung sofort folgt. □

Korollar 257 (Eulersche Summenformel)**:** Ist $g : [0, \infty[\to \mathbb{R}$ eine stetig differenzierbare Funktion, so gilt

$$\sum_{n=1}^{N} g(n) = \int_{1}^{N} g(t)dt + \frac{1}{2}(g(1) + g(N)) + \int_{1}^{N} (t - [t] - \frac{1}{2})g'(t)dt,$$

wobei $[t]$ die Gauss-Klammer von t bezeichnet.

BEWEIS: Man wendet den Satz über die partielle Summation mit $a_n := 1$ und $t_n := n$ für alle $n \in \mathbb{N}$. Man erhält dann für alle $x \geq 1$

$$\begin{aligned}
\sum_{n \leq x} g(n) &= [x]g(x) - \int_{1}^{x} [t]g'(t)dt \\
&= [x]g(x) - \int_{1}^{x} (t - \frac{1}{2})g'(t)dt + \int_{1}^{x} (t - [t] - \frac{1}{2})g'(t)dt.
\end{aligned}$$

Wendet man nun auf das erste Integral die Regel der partielle Integration an, so erhält man für $x := N \in \mathbb{N}$ die Behauptung. $\square$

Problem 91: *Führen Sie die im Beweis nicht näher ausgeführte partielle Integration aus und vervollständigen Sie den Beweis.*

Problem 92: *Folgern Sie aus der Eulerschen Summenformel, dass der Grenzwert*

$$\gamma := \lim_{N\to\infty} \left(\sum_{n=1}^{N} \frac{1}{n} - \log N \right)$$

existiert und $\gamma > \frac{1}{2}$ ist. Man bezeichnet γ auch als die Eulersche Konstante*.*

Korollar 258 (Stirlingsche Formel; einfache Version)**:** Der folgende Grenzwert existiert

$$\lim_{N\to\infty} \frac{N!e^N}{N^N\sqrt{N}}.$$

Für große N gilt also näherungsweise

$$\log N! \approx N \log N - N + \frac{1}{2} \log N.$$

Beweis: Mit $g(x) := \log x$ erhält man aus der Eulerschen Summenformel

$$\log N! = \sum_{n=1}^{N} \log n = \int_1^N \log t + \frac{1}{2} \log N + \int_1^N (t - [t] - \frac{1}{2})\frac{1}{t} dt$$

und somit aus dem Hauptsatz der Differential- und Integralrechnung

$$\log N! = N \log N - N + \frac{1}{2} \log N + \int_1^N (t - [t] - \frac{1}{2})\frac{1}{t} dt.$$

Wenn wir nun zeigen können, dass das Integral $\int_1^\infty (t - [t] - \frac{1}{2})\frac{1}{t} dt$ existiert, so liefert uns die Anwendung der Exponentialfunktion auf diese Gleichung die Existenz des Grenzwertes im Korollar.

Nun ist nach dem Hauptsatz der Differential- und Integralrechnung

$$\begin{aligned}\int_n^{n+1} (t - [t] - \frac{1}{2})\frac{1}{t} dt = \int_n^{n+1} (t - n - \frac{1}{2})\frac{1}{t} dt &= \int_n^{n+1} 1 - \frac{n + 1/2}{t} dt \\ &= n + 1 - n - (n + \frac{1}{2})(\log(n+1) - \log n) \\ &= 1 - (n + \frac{1}{2}) \log(1 + \frac{1}{n}).\end{aligned}$$

Wegen der Potenzreihenentwicklung des Logarithmus

$$-\log(1-x)=\sum_{k=1}^{\infty}\frac{x^k}{k}$$

ist

$$-\log(1+\frac{1}{n})=\sum_{k=1}^{\infty}\frac{\left(-\frac{1}{n}\right)^k}{k}=-\frac{1}{n}+\frac{1}{2n^2}-\frac{1}{3n^3}+\dots.$$

Also

$$\begin{aligned}-n\log(1-\frac{1}{n})&=-1+\frac{1}{2n}-\frac{1}{3n^2}\pm\dots,\\ -\frac{1}{2}\log(1-\frac{1}{n})&=-\frac{1}{2n}+\frac{1}{4n^2}-\frac{1}{6n^3}\pm\dots.\end{aligned}$$

Zusammen ergibt dies für große n

$$\int_n^{n+1}(t-[t]-\frac{1}{2})\frac{1}{t}dt=1-(n+\frac{1}{2})\log(1+\frac{1}{n})=-\frac{1}{12n^2}+O(n^-3).$$

Da die Reihe $\sum_{n=1}^{\infty}n^{-2}$ konvergiert, existiert somit auch das Integral $c:=\int_1^{\infty}(t-[t]-\frac{1}{2})\frac{1}{t}dt$ und das Korollar ist bewiesen. □

Für viele Anwendungen reicht diese Form der Stirlingschen Formel. Wir wollen aber den Grenzwert genau berechnen. Hierzu benötigen wir ein paar Vorarbeiten.

Lemma 259: Für

$$I_m:=\int\sin^m x\,dx$$

gilt die Rekursionsformel

$$I_m=-\frac{1}{m}\cos x\sin^{m-1}x+\frac{m-1}{m}I_{m-2}.$$

Problem 93: *Beweisen Sie das vorstehende Lemma.*

Korollar 260: Für

$$A_m:=\int_0^{\frac{\pi}{2}}\sin^m x\,dx$$

gilt $A_0=\pi/2, A_1=1$ und die Rekursionsformel

$$A_m=\frac{m-1}{m}A_{m-2}.$$

Problem 94: *Beweisen Sie das vorstehende Korollar.*

Satz 261 (Wallissches Produkt)**:** Es gilt

$$\lim_{N\to\infty} \prod_{n=1}^{N} \frac{4n^2}{4n^2-1} = \frac{\pi}{2}.$$

BEWEIS: Es ist mit den Bezeichnungen des letzten Korollars

$$\begin{aligned}\prod_{n=1}^{N} \frac{4n^2}{4n^2-1} &= \prod_{n=1}^{N} \frac{2n}{2n-1}\frac{2n}{2n+1} \\ &= \prod_{n=1}^{N} \frac{A_{2n-2}}{A_{2n}}\frac{A_{2n+1}}{A_{2n-1}} \\ &= \frac{A_0}{A_{2N}}\frac{A_{2N+1}}{A_1} = \frac{\pi}{2}\frac{A_{2N+1}}{A_{2N}}.\end{aligned}$$

Da $0 \le \sin x \le 1$ für alle $0 \le x \le \pi/2$ ist, gilt $A_k \ge A_{k+1}$ für alle $k \in \mathbb{N}$. Also ist

$$\frac{A_{2N+2}}{A_{2N}} \le \frac{A_{2N+1}}{A_{2N}} \le 1.$$

Wegen des vorstehenden Korollars ist aber

$$\lim_{N\to\infty} \frac{A_{2N+2}}{A_{2N}} = 1,$$

womit die Behauptung des Satzes aus dem Schachtelungssatz folgt. □

Problem 95: *Folgern Sie aus dem Wallisschen Produkt, dass auch*

$$\lim_{N\to\infty} \prod_{n=1}^{N} \frac{(2n)^2}{(2n-1)^2} \cdot \frac{1}{2N} = \frac{\pi}{2}$$

ist.

Korollar 262 (Stirlingsche Formel)**:** Es gilt

$$\lim_{n\to\infty} \frac{N!e^N}{N^N\sqrt{N}} = \sqrt{2\pi}.$$

BEWEIS: Es sei
$$b_N := \frac{N!e^N}{N^N\sqrt{N}} \text{ und } b := \lim_{N\to\infty} b_N.$$

Dann ist
$$\lim_{N\to\infty} \frac{b_{2N}}{b_N^2} = \frac{b}{b^2} = \frac{1}{b}.$$

Es ist aber
$$\frac{b_{2N}}{b_N^2} = \frac{(2N)!e^{2N}}{(2N)^{2N}\sqrt{2N}} \left(\frac{N^N\sqrt{N}}{N!e^N}\right)^2 = \frac{1}{\sqrt{2}}\sqrt{N}\binom{2N}{N}\frac{1}{2^{2N}}.$$

Man überlegt sich leicht, dass
$$\binom{2N}{N}\frac{1}{2^{2N}} = \prod_{n=1}^{N} \frac{2n-1}{2n}$$

ist. Aus dem Wallischen Produkt (beachten Sie die Aufgabe dazu!) folgt also, dass
$$\lim_{N\to\infty} \sqrt{N}\binom{2N}{N}\frac{1}{2^{2N}} = \frac{1}{\sqrt{\pi}}$$

ist. Somit haben wir
$$\frac{1}{b} = \lim_{N\to\infty} \frac{b_{2N}}{b_N^2} = \lim_{N\to\infty} \frac{1}{\sqrt{2}}\sqrt{N}\binom{2N}{N}\frac{1}{2^{2N}} = \frac{1}{\sqrt{2\pi}}$$

beziehungsweise $b = \sqrt{2\pi}$. □

Problem 96: *Zeigen Sie, dass*
$$\binom{2N}{N}\frac{1}{2^{2N}} = \prod_{n=1}^{N} \frac{2n-1}{2n}$$

ist.

4.19. Stetige, reellwertige Funktionen mehrerer Veränderlicher und der Fundamentalsatz der Algebra

Wir wollen uns zunächst überlegen, dass ein ähnliches Ergebnis wie Satz 135 auch für stetige, reellwertige Funktionen mehrerer Veränderlicher gilt. Das angestrebte Ergebnis soll also von der Form sein: *Ist K eine* geeignete *Teilmenge des $\mathbb{R}^r$, so ist jede stetige Funktion $f : K \to \mathbb{R}$ beschränkt und hat in K (mindestens) eine Maximalstelle und eine Minimalstelle.*

Aber was heißt hier geeignet?

Um uns klar zu machen, welche Eigenschaften K haben muss, schauen wir uns den Beweis von Satz 51 an. Zum einen wird dort ausgenutzt, dass das Intervall $[a, b]$ beschränkt ist. Dies erlaubt es uns nämlich, den Satz von Bolzano-Weierstraß anzuwenden.

Zum anderen haben wir ausgenutzt, dass der Grenzwert der Folge (x_{n_k}) in $[a, b]$ liegen muss, wenn alle Folgenglieder in dem Intervall liegen. An dieser Stelle würde der Beweis auch scheitern, wenn wir lediglich ein halboffenes Intervall, etwa $]a, b]$, oder ein offenes Intervall $]a, b[$ betrachten würden.

Definition 263: Es sei $A \subset \mathbb{R}^r$.

(i) Die Menge A heißt *beschränkt*, wenn es eine Konstante M gibt, so dass für alle $x = (x_1, \ldots, x_r) \in B$

$$\max_i |x_i| \leq M$$

gilt.

(ii) Die Menge A heißt *abgeschlossen*, wenn für jede Folge (x_n) in A, die (in $\mathbb{R}^r$) gegen einen Grenzwert x_0 konvergiert, auch $x_0 \in A$ ist.

(iii) Die Menge A heißt *kompakt*, wenn sie beschränkt und abgeschlossen ist.

Da im Beweis von Satz 135 der Satz von Bolzano-Weierstraß (Satz 58) benutzt worden ist, überlegen wir uns zunächst, dass ein analoges Ergebnis auch für Folgen im $\mathbb{R}^r$ gilt.

Definition 264: Eine Folge (a_n) im $\mathbb{R}^r$ heißt *beschränkt*, wenn die Menge $\{a_n | n \in \mathbb{N}\}$ beschränkt ist.

Satz 265: [Satz von Bolzano-Weierstraß im $\mathbb{R}^r$] Jede beschränkte Folge (a_n) im $\mathbb{R}^r$ besitzt eine konvergente Teilfolge.

BEWEIS: Wir überlegen uns den Beweis nur für Folgen im $\mathbb{R}^2$. Aber die Argumentation lässt sich offensichtlich auf den $\mathbb{R}^r$ übertragen. Sei also (a_n) eine beschränkte Folge im $\mathbb{R}^2$. Für jedes n ist a_n ein Tupel im $\mathbb{R}^2$, etwa $a_n = (a_n^{(1)}, a_n^{(2)})$. Die Folge $(a_n^{(1)})$ ist eine beschränkte Folge reeller Zahlen. Nach dem Satz von Bolzano-Weierstraß (Satz 58) besitzt diese Folge also eine konvergente Teilfolge. Diese sei mit $(a_{n_k}^{(1)})$ bezeichnet. Die Teilfolge $(a_{n_k}^{(2)})$ der Folge $(a_n^{(2)})$ ist natürlich auch beschränkt. Sie besitzt also – wieder nach Satz 58 – auch eine konvergente Teilfolge. Wir haben also jetzt eine Folge von Indizes gefunden, so dass sowohl die Teilfolge in der 1. Komponente als auch die Teilfolge in der 2. Komponente konvergent ist. Damit ist der Satz für $r = 2$ bewiesen. □

Nun haben wir alles zusammen, um folgende Verallgemeinerung von Satz 135 zu beweisen. Der Beweis überträgt sich nun fast wortwörtlich.

Satz 266: Es sei $K \subset \mathbb{R}^r$ eine kompakte Menge. Jede stetige Funktion $f : K \to \mathbb{R}$ ist beschränkt und hat in K (mindestens) eine Maximalstelle und eine Minimalstelle.

BEWEIS: Wir nehmen an, dass f nicht nach oben beschränkt ist. (Nach unten beschränkt geht analog.) Dann existiert zu jedem $n \in \mathbb{N}$ ein $y_n \in f(K)$ mit $y_n > n$. Zu jedem y_n existiert ein $x_n \in K$ mit $f(x_n) = y_n$. Die Folge (x_n) liegt in K und ist somit beschränkt. Nach dem Satz 265 von Bolzano-Weierstraß im $\mathbb{R}^r$ besitzt (x_n) eine konvergente Teilfolge (x_{n_k}). Wir setzen $x_0 := \lim_{k\to\infty} x_{n_k}$, dann ist $x_0 \in K$, da alle $x_{n_k} \in K$ sind und K kompakt, also insbesondere abgeschlossen, ist.
Da f stetig ist, gilt

$$\lim_{k\to\infty} f(x_{n_k}) = f(x_0).$$

Nach Annahme ist die Folge $f(x_{n_k})$ jedoch unbeschränkt, also insbesondere divergent. Dies ist ein Widerspruch, damit ist gezeigt, dass f beschränkt ist.

Es sei nun

$$M := \sup f(x)|x \in K.$$

Wir müssen zeigen, dass $M \in f(K)$ ist. Es existiert zu jedem $\varepsilon > 0$ ein $y_\varepsilon \in f(K)$ mit

$$M - \varepsilon \leq y_\varepsilon < M.$$

Betrachten wir etwa $\varepsilon = \frac{1}{n}$ für $n \in \mathbb{N}$, so sehen wir, dass es zu jedem n ein x_n gibt mit

$$M - \frac{1}{n} \leq f(x_n) < M.$$

Daraus folgt

$$\lim_{n\to\infty} f(x_n) = M.$$

Die Folge (x_n) ist wieder beschränkt und wie oben folgt, dass es eine konvergente Teilfolge (x_{n_k}) gibt. Wir setzen wieder $x_0 := \lim_{k\to\infty} x_{n_k}$. Wegen der Stetigkeit von f gilt wieder

$$M = \lim_{k\to\infty} f(x_{n_k}) = f(x_0).$$

Der Beweis für die Minimalstelle erfolgt analog. □

Da wir die komplexen Zahlen $\mathbb{C}$ mit dem $\mathbb{R}^2$ identifizieren können, haben wir insbesondere das folgende Korollar.

Korollar 267: Es sei $K \subset \mathbb{C}$ eine kompakte Menge. Jede stetige Funktion $f : K \to \mathbb{R}$ ist beschränkt und hat in K (mindestens) eine Maximalstelle und eine Minimalstelle.

Nun haben wir alles zusammen, um den Fundamentalsatz der Algebra beweisen zu können. Die Idee zu der hier präsentierten einfachen Beweisvariante dieses Satzes stammt von de Oliveira[5].

Satz 268 (Fundamentalsatz der Algebra)**:** Es seien $n \in \mathbb{N}, a_0, a_1, \ldots, a_n \in \mathbb{C}$ und $a_n \neq 0$. Dann gibt es ein $z \in \mathbb{C}$, so dass

$$a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 = 0$$

ist.

BEWEIS: Wir setzen $P(z) := a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0$ betrachten die stetige Funktion

$$|P| : \mathbb{C} \to \mathbb{R}, z \mapsto |P(z)|.$$

Wegen

$$\begin{aligned} |P(z)| &\geq |a_n z^n| - |a_{n-1} z^{n-1} + \ldots + a_1 z + a_0| \\ &\geq |a_n||z|^n - |a_{n-1}||z|^{n-1} - \ldots - |a_1||z| - |a_0| \\ &= |a_n||z|^n (1 - |a_{n-1}||z|^{-1} - \ldots - |a_1||z|^{-(n-1)} - |a_0||z|^{-n}). \end{aligned}$$

gilt

$$\lim_{|z| \to \infty} |P(z)| = \infty.$$

Die stetige Funktion $|P|$ nimmt auf den kompakten Mengen $K_R := \{z \in \mathbb{C} |\ |z| \leq R\}$ ein lokales Minimum an. Wegen $\lim_{|z|\to\infty} |P(z)| = \infty$ muss dieses Minimum aber ein globales Minimum sein, wenn R nur hinreichend groß gewählt wird. Ohne Einschränkung der Allgemeinheit können wir annehmen, dass dieses globale Minimum in $z_0 = 0$ angenommen wird (Betrachte ansonsten $P(z + z_0)$ anstelle von $P(z)$!). Wir wollen nun zeigen, dass $P(0) = 0$ ist.

Für alle $z \in \mathbb{C}$ gelte also

$$|P(z)|^2 \geq |P(0)|^2.$$

[5] Oswaldo Rio Branco de Oliveira, The Fundamental Theorem of Algebra: An Elementary and Direct Proof, The Mathematical Intelligencer July 2011, Volume 33, Issue 2, pp 1-2.

Weiter existiert ein Polynom Q mit $Q(0) \neq 0$ und eine natürliche Zahl k, so dass wir P schreiben können als $P(z) = P(0) + z^k Q(z)$ für alle $z \in \mathbb{C}$. Dann ist

$$\begin{aligned}
0 &\leq |P(z)|^2 - |P(0)|^2 \\
&= |P(0) + z^k Q(z)|^2 - |P(0)|^2 \\
&= \overline{(P(0) + z^k Q(z))}(P(0) + z^k Q(z)) - |P(0)|^2 \\
&= \overline{P(0)} z^k Q(z) + P(0)\overline{z^k Q(z)} + |z^k Q(z)|^2 \\
&= 2\,\mathrm{Re}\left(\overline{P(0)} z^k Q(z)\right) + |z|^{2k}|Q(z)|^2.
\end{aligned}$$

Nun können wir jedes z als $z = |z|e^{i\varphi}$ schreiben. Wenn wir dies einsetzen, erhalten wir

$$0 \leq 2|z|^k \,\mathrm{Re}\left(\overline{P(0)} e^{ik\varphi} Q(z)\right) + |z|^{2k}|Q(z)|^2.$$

Dividieren wir dies durch $|z|^k$ und betrachten wir dann den Grenzübergang $|z| \to 0$ bei festem φ, so erhalten wir aufgrund der Stetigkeit der rechten Seite der vorstehenden Ungleichung

$$0 \leq 2\,\mathrm{Re}\left(\overline{P(0)} e^{ik\varphi} Q(0)\right).$$

Diese Ungleichung gilt für alle $\varphi \in \mathbb{R}$. Wir wählen nun sukzessive $\varphi = 0, \frac{\pi}{k}$ und $\varphi = \frac{\pi}{2k}, \frac{3\pi}{2k}$; also $e^{ik\varphi} = \pm 1$ und $e^{ik\varphi} = \pm i$, um zu schließen, dass sowohl $\pm\,\mathrm{Re}(\overline{P(0)}Q(0)) \geq 0$ als auch $\pm\,\mathrm{Im}(\overline{P(0)}Q(0)) \geq 0$ gelten muss. Folglich ist $\overline{P(0)}Q(0) = 0$. Da $Q(0) \neq 0$ ist, muss somit $P(0) = 0$ sein.

$\square$

5. Ausblicke

5.1. Differentialrechnung im $\mathbb{R}^r$

Wir haben uns bereits in Kapitel 4.7 mit Extrema bei Funktionen mehrerer Veränderlicher beschäftigt. Dabei hatten wir aber keine eigenständige Differentialrechnung im $\mathbb{R}^r$ entwickelt, sondern alle Begriffe und Ergebnisse aus der Differentialrechnung im Eindimensionalen abgeleitet.

Hier werden wir kurz skizzieren, wie man analog zur Differentialrechnung, die wir in $\mathbb{R}$ eingeführt haben, eine solche im $\mathbb{R}^r$ einführt.

Wir benutzen die in Kapitel 4.7 eingeführten Bezeichnungen und Konventionen. Da wir Funktionen betrachten wollen, deren Wertebereich im $\mathbb{R}^m$ liegt, müssen wir früher gemachte Definition von Grenzwerten von Funktionen verallgemeinern. Ist $D \subset \mathbb{R}^r$, $f : D \to \mathbb{R}^m$ eine Funktion, $a \in \mathbb{R}^r$ ein Berührpunkt von D und $b \in \mathbb{R}^m$, so schreiben wir

$$\lim_{x \to a} f(x) = b,$$

falls für jede Folge (x_n) in D mit $\lim_{n\to\infty} x_n = a$

$$\lim_{n \to \infty} f(x_n) = b$$

gilt.

Generalvoraussetzung: In diesem Kapitel sei $D \subset \mathbb{R}^r$ stets eine offene, nichtleere Teilmenge des $\mathbb{R}^r$.

Definition 269 (Differenzierbarkeit im $\mathbb{R}^r$)**:** Eine Funktion $f : D \to \mathbb{R}^p$ heißt *differenzierbar in* $x_0 \in D$, falls ein $A \in M(p \times r, \mathbb{R})$ und eine Funktion $r_{x_0} : D \to \mathbb{R}^p$ mit

$$\lim_{x \to x_0} \frac{1}{\|x - x_0\|_2} r_{x_0}(x) = 0 \in \mathbb{R}^p$$

existieren, so dass für alle $x \in D$

$$f(x) = f(x_0) + A(x - x_0) + r_{x_0}(x)$$

gilt. Der Matrix A heißt die *(totale) Ableitung von f an der Stelle* x_0, welchen wir mit $f'(x_0)$ bezeichnen wollen.
Die Funktion f heißt *(total) differenzierbar* (auf ganz D), falls f in allen Punkten $x \in D$ differenzierbar ist.

Eine Funktion $f : D \to \mathbb{R}^p$ wie sie in der vorstehenden Definition betrachtet wird, setzt sich aus Komponenten Funktionen zusammen, d.h. es gibt Funktionen $f_1, \dots, f_p : D \to \mathbb{R}$, so dass man die Funktion f schreiben kann als

$$x \mapsto \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_p(x) \end{pmatrix} = f(x).$$

Die obige Definition des Differenzierbarkeitsbegriffs lässt sich sowohl auf f als auch auf die Komponentenfunktionen anwenden. Der nächste Satz zeigt, dass beides zum gleichen Ergebnis führt.

Satz 270: Es sei $D \subset \mathbb{R}^r$ eine offene Menge und $f_1, \dots, f_p : D \to \mathbb{R}$ Funktionen. Die Funktion $f : D \to \mathbb{R}^p$ mit

$$x \mapsto \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_p(x) \end{pmatrix} =: f(x)$$

ist genau dann differenzierbar in $x_0 \in D$, wenn alle Komponentenfunktionen $f_1, \dots, f_p$ in x_0 differenzierbar sind. In diesem Fall gilt

$$f'(x_0) = \begin{pmatrix} f_1'(x_0) \\ f_2'(x_0) \\ \vdots \\ f_p'(x_0) \end{pmatrix}.$$

Der folgende Satz stellt einen Zusammenhang zum früher eingeführten Begriff der partiellen Ableitungen her und ist ein praktisches Kriterium, mit dem man prüfen kann, ob eine Funktion differenzierbar ist.

Satz 271: Es sei $D \subset \mathbb{R}^r$ eine offene Menge und $f = (f_1, \dots, f_m) : D \to \mathbb{R}^p$ eine Funktion. f ist in $x_0 \in D$ differenzierbar, wenn alle pr partiellen Ableitungen

$$\frac{\partial f_\pi}{\partial x_\rho}(x) \text{ für } \pi = 1, \dots, p, \;\; \rho = 1, \dots, r$$

in einer Umgebung von x_0 existieren und in x_0 stetig sind. In diesem Fall gilt

$$f'(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_r}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \cdots & \frac{\partial f_2}{\partial x_r}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(x_0) & \frac{\partial f_p}{\partial x_2}(x_0) & \cdots & \frac{\partial f_p}{\partial x_r}(x_0) \end{pmatrix}$$

Da die Definition der Differenzierbarkeit bei Funktionen mehrerer Veränderlicher eine natürliche Verallgemeinerung der Definition im Eindimensionalen darstellt, verwundert es nicht, dass sich die üblichen Rechenregeln übertragen lassen. Die wichtigste Regel lautet

Satz 272 (Kettenregel im $\mathbb{R}^n$)**:** Es seien $D \subset \mathbb{R}^r$ und $E \subset \mathbb{R}^p$. Die Hintereinanderausführung zweier differenzierbarer Funktionen $f : D \to E$ und $g : E \to \mathbb{R}^q$ ist differenzierbar und es gilt für alle $x_0 \in D$

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

BEMERKUNG: Auf der rechten Seite der Kettenregel steht das Produkt einer $q \times p$- und einer $p \times r$-Matrix.

Wir wollen uns den wichtigen Spezialfall $r = q = 1$ der Kettenregel etwas genauer anschauen. Die Komponentenfunktionen von f seien mit $f_1, \dots, f_p$ bezeichnet. Dann ist

$$f'(x_0) = \begin{pmatrix} f_1'(x_0) \\ f_2'(x_0) \\ \vdots \\ f_p'(x_0) \end{pmatrix}$$

und

$$g'(f(x_0)) = \left(\frac{\partial g}{\partial x_1}(f(x_0)), \dots, \frac{\partial g}{\partial x_p}(f(x_0)) \right).$$

Also haben wir

$$(g \circ f)'(x_0) = \sum_{\pi=1}^{p} \frac{\partial g}{\partial x_\pi}(f(x_0)) f_\pi'(x_0).$$

Nun sei $D \subset \mathbb{R}^2$ und $F : D \to \mathbb{R}$. Weiter seien $x, x_0 \in D$ derart, dass für alle $t \in [0, 1]$ auch $x_0 + t(x - x_0)$ in D liegt, d.h. die Verbindungsstrecke von x_0 und x liegt in D. Es sei noch $(h_1, h_2) := x - x_0$ gesetzt. Wir betrachten nun die durch $\varphi(t) := F(x_0 + t(x - x_0))$ definierte Funktion $\varphi : [0, 1] \to \mathbb{R}$. Ist F $(N + 1)$-mal stetig partiell differenzierbar, so ist φ $(N + 1)$-mal differenzierbar auf $[0, 1]$ und es existiert nach dem Satz von Taylor (Satz 172) ein $\xi \in]0, 1[$, so dass

$$\varphi(1) = \sum_{n=0}^{N} \frac{\varphi^{(n)}(0)}{n!}(1 - 0)^n + \frac{\varphi^{(N+1)}(\xi)}{(N + 1)!}(1 - 0)^{N+1}.$$

Wir wollen nun die Ableitungen von φ berechnen. Nach dem gerade betrachteten Spezialfall der Kettenregel ist

$$\varphi'(t) = h_1 \frac{\partial F}{\partial x_1}(x_0 + t(x - x_0)) + h_2 \frac{\partial F}{\partial x_2}(x_0 + t(x - x_0)).$$

Erneutes Differenzieren liefert

$$\begin{aligned}\varphi''(t) &= h_1(h_1\frac{\partial^2 F}{\partial x_1\partial x_1}(x_0+t(x-x_0))+h_2\frac{\partial^2 F}{\partial x_2\partial x_1}(x_0+t(x-x_0)))\\ &\quad +h_2(h_1\frac{\partial^2 F}{\partial x_1\partial x_2}(x_0+t(x-x_0))+h_2\frac{\partial^2 F}{\partial x_2\partial x_2}(x_0+t(x-x_0)))\\ &= h_1h_1\frac{\partial^2 F}{\partial x_1\partial x_1}(x_0+t(x-x_0))+2h_1h_2\frac{\partial^2 F}{\partial x_2\partial x_1}(x_0+t(x-x_0))\\ &\quad +h_2h_2\frac{\partial^2 F}{\partial x_2\partial x_2}(x_0+t(x-x_0)),\end{aligned}$$

wobei die letzte Gleichheit aufgrund des Satzes von Schwarz (Satz 202) gilt. Schließlich rechnet man leicht nach, dass aufgrund des Satzes von Schwarz (Satz 202)

$$\varphi^{(n)}(t)=\sum_{\nu=0}^{n}\binom{n}{\nu}h_1^{n-\nu}h_2^{\nu}\frac{\partial^n F}{\partial x_2^{\nu}\partial x_1^{n-\nu}}(x_0+t(x-x_0))$$

ist. Wir führen nun eine abkürzende Schreibweise ein, die aufgrund des Satzes von Schwarz und des binomischen Lehrsatzes gerechtfertigt ist, wenn man mit den Ableitungsoperatoren $\frac{\partial}{\partial x_1}$ und $\frac{\partial}{\partial x_2}$ ganz formal rechnet. Wir setzen

$$\left(h_1\frac{\partial}{\partial x_1}+h_2\frac{\partial}{\partial x_2}\right)^n:=\sum_{\nu=0}^{n}\binom{n}{\nu}h_1^{n-\nu}h_2^{\nu}\frac{\partial^n}{\partial x_2^{\nu}\partial x_1^{n-\nu}}.$$

Dann haben wir also

$$\varphi^{(n)}(t)=\left(h_1\frac{\partial}{\partial x_1}+h_2\frac{\partial}{\partial x_2}\right)^n F(x_0+t(x-x_0))$$

Mit diesen Schreibweisen haben wir gezeigt:

Satz 273 (Satz von Taylor im $\mathbb{R}^2$)**:** Es seien $D\subset\mathbb{R}^2$ und $F:D\to\mathbb{R}$ $(N+1)$-mal stetig partiell differenzierbar. Weiter seien $x,x_0\in D$ derart, dass die Verbindungsstrecke von x_0 und x in D liegt. Setzen wir $(h_1,h_2):=x-x_0$, so existiert ein $\xi\in]0,1[$ mit

$$F(x)=F(x_0)+\sum_{n=1}^{N}\frac{\left(h_1\frac{\partial}{\partial x_1}+h_2\frac{\partial}{\partial x_2}\right)^n F(x_0)}{n!}+\frac{\left(h_1\frac{\partial}{\partial x_1}+h_2\frac{\partial}{\partial x_2}\right)^{N+1} F(x_0+\xi(x-x_0))}{(N+1)!}.$$

Allgemeiner gilt

Satz 274 (Satz von Taylor im $\mathbb{R}^r$)**:** Es seien $D\subset\mathbb{R}^r$ und $F:D\to\mathbb{R}$ $(N+1)$-mal stetig partiell differenzierbar. Weiter seien $x,x_0\in D$ derart, dass die Verbindungsstrecke von x_0 und x in D liegt. Setzen wir $(h_1,\dots,h_r):=x-x_0$, so existiert ein $\xi\in]0,1[$ mit

$$F(x)=F(x_0)+\sum_{n=1}^{N}\frac{\left(\sum_{\rho=1}^{r}h_\rho\frac{\partial}{\partial x_\rho}\right)^n F(x_0)}{n!}+\frac{\left(\sum_{\rho=1}^{r}h_\rho\frac{\partial}{\partial x_\rho}\right)^{N+1} F(x_0+\xi(x-x_0))}{(N+1)!}.$$

Wir wollen noch eine Anwendung des Satzes skizzieren. Wir wollen den numerischen Wert von $1,05^{1,02}$ näherungsweise bestimmen. Dazu wenden wir den Satz von Taylor auf $F(x_1, x_2) := x_1^{x_2} = \exp(x_2 \log x_1)$ mit $N = 2$ und $x_0 = (1,1)$ an. Es ist $h_1 = 0.05$ und $h_2 = 0,02$. Als partielle Ableitungen erhalten wir

$$
\begin{aligned}
\frac{\partial}{\partial x_1} F(x_1,x_2) &= \exp(x_2 \log x_1)\tfrac{x_2}{x_1} && (= 1 \text{ für } (x_1,x_2) = (1,1)) \\
\frac{\partial}{\partial x_2} F(x_1,x_2) &= \exp(x_2 \log x_1) \log x_1 && (= 0 \text{ für } (x_1,x_2) = (1,1)) \\
\frac{\partial^2}{\partial x_1^2} F(x_1,x_2) &= \exp(x_2 \log x_1)\left(\tfrac{x_2}{x_1}\right)^2 - \exp(x_2 \log x_1)\tfrac{x_2}{x_1^2} && (= 0 \text{ für } (x_1,x_2) = (1,1)) \\
\frac{\partial^2}{\partial x_2 \partial x_1} F(x_1,x_2) &= \exp(x_2 \log x_1)\tfrac{x_2}{x_1} \log x_1 + \exp(x_2 \log x_1)\tfrac{1}{x_1} && (= 1 \text{ für } (x_1,x_2) = (1,1)) \\
\frac{\partial^2}{\partial x_2^2} F(x_1,x_2) &= \exp(x_2 \log x_1)(\log x_1)^2 && (= 0 \text{ für } (x_1,x_2) = (1,1)).
\end{aligned}
$$

Aus dem Satz von Taylor ergibt sich nun

$$
\begin{aligned}
1,05^{1,02} &= f(1+0,05, 1+0,02) \approx f(1,1) + 0,05 \cdot \frac{\partial}{\partial x_1} F(1,1) + 0,05 \cdot 0,02 \cdot \frac{\partial^2}{\partial x_2 \partial x_1} F(1,1) \\
&= 1 + 0,05 + 0,0010 = 1,051
\end{aligned}
$$

Um die Güte dieser Abschätzung zu überprüfen, müssten wir nun das Restglied im Satz von Taylor abschätzen. Man kann zeigen, dass der Fehler dieser Abschätzung $< 0,0001$ ist.

Wir beschließen diesen Abschnitt mit einer Übertragung des in Kapitel 4.6 besprochenen Newton-Verfahrens ins Mehrdimensionale. Die Newton-Folge (x_n) war damals induktiv durch

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$$

definiert. Ist nun $D \subset \mathbb{R}^n$ und $f : D \to \mathbb{R}^n$, so übertragt sich dies ins Mehrdimensionale zu

$$x_{n+1} := x_n - f'(x_n)^{-1} f(x_n).$$

Hier bezeichnet $f'(x_n)^{-1}$ natürlich die zu $f'(x_n)$ inverse Matrix, d.h. wir müssen nun mindestens voraussetzen, dass die $n \times n$-Matrix $f'(x_n)$ invertierbar sein muss.

Das folgende Matlab-Skript berechnet nun ausgehend vom Startwert $(1,1)$ eine Näherungslösung für eine Nullstelle der Funktion

$$
\begin{aligned}
&f : \mathbb{R}^2 \to \mathbb{R}^2 \\
&(x,y) \mapsto (\sin(x)\cos(y), x^2 + y^2 - 3),
\end{aligned}
$$

d.h. eine Näherungslösung für das nichtlineare Gleichungssystem

$$\begin{aligned} \sin(x)\cos(y) &= 0 \\ x^2 + y^2 - 3 &= 0. \end{aligned}$$

```
%newtonmehrdim.m
x=[1 1]'
for i=1:5
  A=[cos(x(1))*cos(x(2)) -sin(x(1))*sin(x(2)); 2*x(1) 2*x(2)]
  y=[sin(x(1))*cos(x(2))  x(1)^2+x(2)^2-3]'
  x=x-A\y %A\y berechnet die Lösung des Gleichungssystems Ax=y, also A^-1y
end
[sin(x(1))*cos(x(2))  x(1)^2+x(2)^2-3]' %f(x)
```

Das Skript liefert nach 5 Schritten:

```
x =

   0.72979
   1.57080

ans =

   4.0825e-17
   0.0000e+00
```

5.2. Extrema mit Nebenbedingungen: Lagrange-Multiplikatoren

Bei Anwendungen stellt sich häufig das Problem, dass man nicht die lokalen Extrema einer Funktion bestimmen muss, sondern, dass dabei eine oder mehrere Nebenbedingungen erfüllt sein sollen. Wir verdeutlichen dies an einem einfachen Beispiel.

Wir wollen unter allen Rechtecken eines fest vorgegebenen Umfangs, das mit dem größten Flächeninhalt bestimmen.

Bezeichnen wir die Seitenlängen des Rechtecks mit a und b, so wird der Flächeninhalt durch die Funktion $F(a, b) = ab$ beschrieben. Diese Funktion hat offensichtlich gar keine Extrema. Nun soll aber der Umfang der betrachteten Rechtecke fest vorgegeben sein. Es soll etwa $U(a, b) = 2a + 2b = u$ sein, wobei u fest ist. Die Aufgabe lautet also nun: Bestimmen Sie die Extrema der Funktion $F(a, b) = ab$ unter der Nebenbedingung $U(a, b) = 2a + 2b = u$. Diese Aufgabe kann man mit Hilfe der Differentialrechnung in einer Veränderlichen lösen. Aus der Nebenbedingung ergibt sich, dass $b = u/2 - a$ sein muss. Setzt man dies in F ein, so sieht man, dass man nun die Extrema der Funktion $f(a) = a(u/2 - a) = au/2 - a^2$ bestimmen muss. Die Ableitung dieser Funktion hat eine Nullstelle bei $a = u/4$. Da die zweite Ableitung an dieser Stelle negativ ist, nimmt die Funktion für $a = b = u/4$ ein Maximum an.

Nicht immer lassen sich solche Probleme aber auf die Bestimmung von Extrema von Funktionen einer Veränderlichen zurückführen. Der folgende Satz liefert ein tiefliegendes Ergebnis, das aber häufig zum Ziel führt. Wir werden ihn nicht beweisen. Es gibt auch noch allgemeinere Versionen dieses Satzes für den Fall, dass man mehr als nur eine Nebenbedingung berücksichtigen muss.

Satz 275: Es sei $U \subset \mathbb{R}^n$ offen und $f, F : U \to \mathbb{R}$ seien stetig differenzierbare Funktionen. Es gelte $\operatorname{grad}(F)(x) \neq 0$ für alle $x \in U$. Ist dann x_0 ein lokales Extremum von f auf

$$S := \{x | F(x) = 0\},$$

so existiert ein $\lambda \in \mathbb{R}$ mit

$$\operatorname{grad}(f)(x_0) = \lambda \operatorname{grad}(F)(x_0).$$

λ wird *Lagrange-Multiplikator* genannt.

Zunächst zeigen wir nun, wie man das obige Problem mit diesem Satz löst. Wir setzen $f(a, b) := ab$ und $F(a, b) := 2a + 2b - u$. Falls es ein lokales Extremum unter der Nebenbedingung $F(a, b) = 0$ gibt, so existiert eine reelle Zahl λ mit $\operatorname{grad}(f)(a, b) = \lambda \operatorname{grad}(F)(a, b)$, also

$$\begin{aligned} b = \frac{\partial f}{\partial a}(a, b) &= \lambda \frac{\partial F}{\partial a}(a, b) = \lambda 2 \\ a = \frac{\partial f}{\partial b}(a, b) &= \lambda \frac{\partial F}{\partial b}(a, b) = \lambda 2. \end{aligned}$$

Es muss also $a = b$ sein. Aus der Nebenbedingung $F(a, b) = 0$ errechnet man dann die Lösung $a = b = u/4$, die wir zuvor auch schon gefunden hatten.

Schauen wir uns ein etwas kompliziertes Beispiel in diese Richtung an. Wir wollen eine Getränkedose mit möglichst großem Volumen bei einer vorgegebenen Oberfläche von 625cm^2 konstruieren. Wir stellen uns die Dose als einen Zylinder mit Radius r und Höhe h vor. Die Nebenbedingung lautet dann

$$F(r, h) = 2\pi rh + 2\pi r^2 - 625cm^2 = 0.$$

Die zu maximierende Funktion lautet

$$f(r, h) = \pi r^2 h.$$

Falls es ein lokales Extremum unter der Nebenbedingung $F(r, h) = 0$ gibt, so existiert eine reelle Zahl λ mit $\operatorname{grad}(f)(r, h) = \lambda \operatorname{grad}(F)(r, h)$, also

$$\begin{aligned} 2\pi rh = \frac{\partial f}{\partial r}(r, h) &= \lambda \frac{\partial F}{\partial r}(r, h) = \lambda(2\pi h + 4\pi r) \\ \pi r^2 = \frac{\partial f}{\partial h}(r, h) &= \lambda \frac{\partial F}{\partial h}(r, h) = \lambda 2\pi r. \end{aligned}$$

Aufgrund der 2. Gleichung muss $r = 0$ oder $r = 2\lambda$ sein. $r = 0$ können wir verwerfen, da dann das Volumen der Dose gleich 0 ist. Wir ersetzen nun $r = 2\lambda$ in der ersten Gleichung und erhalten

$$0 = \lambda(2\pi h + 4\pi r) - 2\pi r h = 2\lambda\pi(4\lambda - h).$$

Also muss $\lambda = 0$ oder $h = 4\lambda$ sein. $\lambda = 0$ impliziert aber $r = 0$. Wir setzen nun $r = 2\lambda$ und $h = 4\lambda$ in die Nebenbedingung $2\pi r h + 2\pi r^2 - 625cm^2 = 0$ ein und erhalten nach kurzer Rechnung, dass $\lambda^2 = \frac{625}{24\pi}$ sein muss. Da r und h positiv sein sollten, erhalten wir also $\lambda = \sqrt{\frac{625cm^2}{24\pi}}$. Es ergibt sich also ein maximales Volumen von

$$\pi r^2 h = \pi 4\lambda^2 4\lambda = 16\pi \frac{625}{24\pi}\sqrt{\frac{625cm^2}{24\pi}} = \frac{2}{3}625cm^2\sqrt{\frac{625}{24\pi}}cm \approx 1199.6325cm^3$$

Als weitere Anwendung dieses Satzes wollen wir berechnen, welcher Punkt (x, y) der Hyperbel

$$H = \{(x, y)|x^2 + 8xy + 7y^2 = 225\}$$

den kürzesten Abstand zum Nullpunkt hat.

Wir wollen also die Extrema der Abstandsfunktion $f(x, y) = x^2 + y^2$ unter der Nebenbedingung $F(x, y) = x^2 + 8xy + 7y^2 - 225 = 0$ bestimmen.

Falls ein lokales Extremum unter der Nebenbedingung $F(a, b) = 0$ existiert, so gibt es eine reelle Zahl λ mit $\operatorname{grad}(f)(x, y) = \lambda \operatorname{grad}(F)(x, y)$, also

$$\begin{aligned} 2x = \frac{\partial f}{\partial x}(x, y) &= \lambda\frac{\partial F}{\partial x}(x, y) = \lambda(2x + 8y) \\ 2y = \frac{\partial f}{\partial y}(x, y) &= \lambda\frac{\partial F}{\partial y}(x, y) = \lambda(8x + 14y). \end{aligned}$$

Diese beiden Gleichungen ergeben ein lineares Gleichungssystem für die Unbestimmten x, y mit dem Parameter λ. Dieses lineare Gleichungssystem hat genau dann eine nichttriviale Lösung, wenn die Determinate der Koeffizientenmatrix verschwindet. Es ist

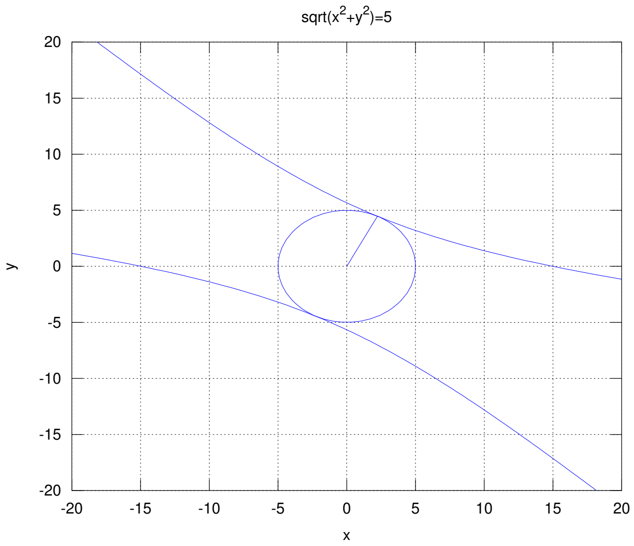
$$\det\begin{pmatrix} 2\lambda - 2 & 8\lambda \\ 8\lambda & 14\lambda - 2 \end{pmatrix} = (2\lambda - 2)(14\lambda - 2) - 64\lambda = -4(\lambda + 1)(9\lambda - 1).$$

Wir erhalten also, dass $\lambda = -1$ oder $\lambda = 1/9$ sein muss.

Für $\lambda = -1$ erhalten wir, dass $x = -2y$ gelten muss. Aus der Nebenbedingung $F(x, y) = 0$ ergibt sich dann, dass $-5y^2 = 225$ sein muss. In diesem Fall gibt es also keine Lösung, die auf der Hyperbel liegt. Für $\lambda = 1/9$ erhalten wir, dass $x = y/2$ bzw. $y = 2x$ gelten muss. Aus der Nebenbedingung $F(x, y) = 0$ ergibt sich hiermit, dass $x = \pm\sqrt{\frac{225}{45}} = \pm\sqrt{5}$ sein muss und somit $f(x, y) = x^2 + y^2 = 5x^2 = 25$. Die Punkte mit dem kleinsten Abstand zum Nullpunkt sind also die Punkte $(\sqrt{5}, \sqrt{20})$ und $(-\sqrt{5}, -\sqrt{20})$, und der kürzeste Abstand von der Hyperbel zum Nullpunkt ist 5.

```
ezplot('x^2+8*x*y+7*y^2=225',[-20 20 -20 20])
hold on
```

```
plot([0 sqrt(5)],[0 2*sqrt(5)])
ezplot('sqrt(x^2+y^2)=5',[-20 20 -20 20])
grid on
```



5.3. Fourierreihen

In diesem Kapitel geht es um gute Darstellungen von 2π-periodischen Funktionen. Die Ergebnisse lassen sich auf Funktionen mit beliebiger Periode übertragen. Hat nämlich $g : \mathbb{R} \to \mathbb{R}$ die Periode p, so hat $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = g\left(\frac{p}{2\pi}x\right)$ die Periode 2π, denn es gilt

$$f(x + 2\pi) = g\left(\frac{p}{2\pi}(x + 2\pi)\right) = g\left(\frac{p}{2\pi}x + p\right) = g\left(\frac{p}{2\pi}x\right) = f(x).$$

Definition 276: Es seien $f, g : \mathbb{R} \to \mathbb{R}$ 2π-periodische, integrierbare Funktionen. So setzt man

$$\langle f, g \rangle := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)\, dx.$$

Lemma 277: Für alle $\lambda \in \mathbb{R}$ und $f_1, f_2, f, g : \mathbb{R} \to \mathbb{R}$ 2π-periodischen, integrierbaren Funktionen gilt:

(i) $\langle \lambda f_1 + f_2, g \rangle = \lambda \langle f_1, g \rangle + \langle f_2, g \rangle$

(ii) $\langle f, g \rangle = \langle g, f \rangle$

(iii) $\|f\|^2 := \langle f, f \rangle = \dfrac{1}{\pi} \displaystyle\int_{-\pi}^{\pi} (f(x))^2\, dx \geq 0$. Ist f stetig, so gilt sogar $\langle f, f \rangle = 0 \Leftrightarrow f = 0$.

Problem 97: *Beweisen Sie das vorstehende Lemma.*

BEMERKUNG: Die Funktion $\langle \cdot, \cdot \rangle$ ist eine semi-positiv definite, symmetrische Bilinearform auf dem Vektorraum der 2π-periodischen, integrierbaren Funktionen. Diese Bilinearform hat bis auf $\langle f, f \rangle = 0 \Leftrightarrow f = 0$ alle Eigenschaften eines Skalarproduktes. Die durch dieses Skalarprodukt in (iii) definierte Norm der Funktion f ist nicht die Supremumsnorm, die wir früher betrachtet haben. Im aktuellen Kapitel meinen wir mit $\|\cdot\|$ immer die im vorstehenden Lemma definierte Norm und niemals die Supremumsnorm.

Lemma 278: Für alle $n, m \in \mathbb{N}$ gelten folgende *Orthogonalitätsbedingungen*:

$$\langle \cos(nx), \sin(mx) \rangle = 0$$

$$\langle \cos(nx), \cos(mx) \rangle = \langle \sin(nx), \sin(mx) \rangle = \begin{cases} 0, & \text{falls } n \neq m \\ 1, & \text{falls } n = m. \end{cases}$$

BEWEIS: Auf Grund der Additionstheoreme gilt

$$\begin{aligned} \cos(\alpha + \beta) &= \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \\ \cos(\alpha - \beta) &= \cos(\alpha)\cos(-\beta) - \sin(\alpha)\sin(-\beta) \\ &= \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta). \end{aligned}$$

Durch Addition erhält man

$$\cos(\alpha+\beta)+\cos(\alpha-\beta)=2\cos(\alpha)\cos(\beta).$$

Womit

$$\cos(mx)\cos(nx)=\frac{1}{2}\left(\cos((m-n)x)+\cos((m+n)x)\right)$$

und für $n \neq m$

$$\int_{-\pi}^{\pi}\cos(mx)\cos(nx)\,dx=0$$

folgt. Für $n=m$ gilt

$$\int_{-\pi}^{\pi}\cos^2(mx)\,dx=\frac{1}{m}\int_{-m\pi}^{m\pi}\cos^2(t)\,dt=\frac{1}{m}\frac{1}{2}\left(\sin(t)\cos(t)+t\right)\Big|_{-m\pi}^{m\pi}=\pi.$$

□

Problem 98: *Beweisen die restlichen Aussagen im vorstehenden Satz.*

Definition 279 (Fourierreihe)**:** Es sei f eine 2π-periodische, integrierbare Funktion. Dann heißen die Zahlen

$$a_k := \langle f, \cos(k\,\cdot)\rangle \text{ für } k=0,1,2,\ldots$$

und

$$b_k := \langle f, \sin(k\,\cdot)\rangle \text{ für } k=1,2,\ldots$$

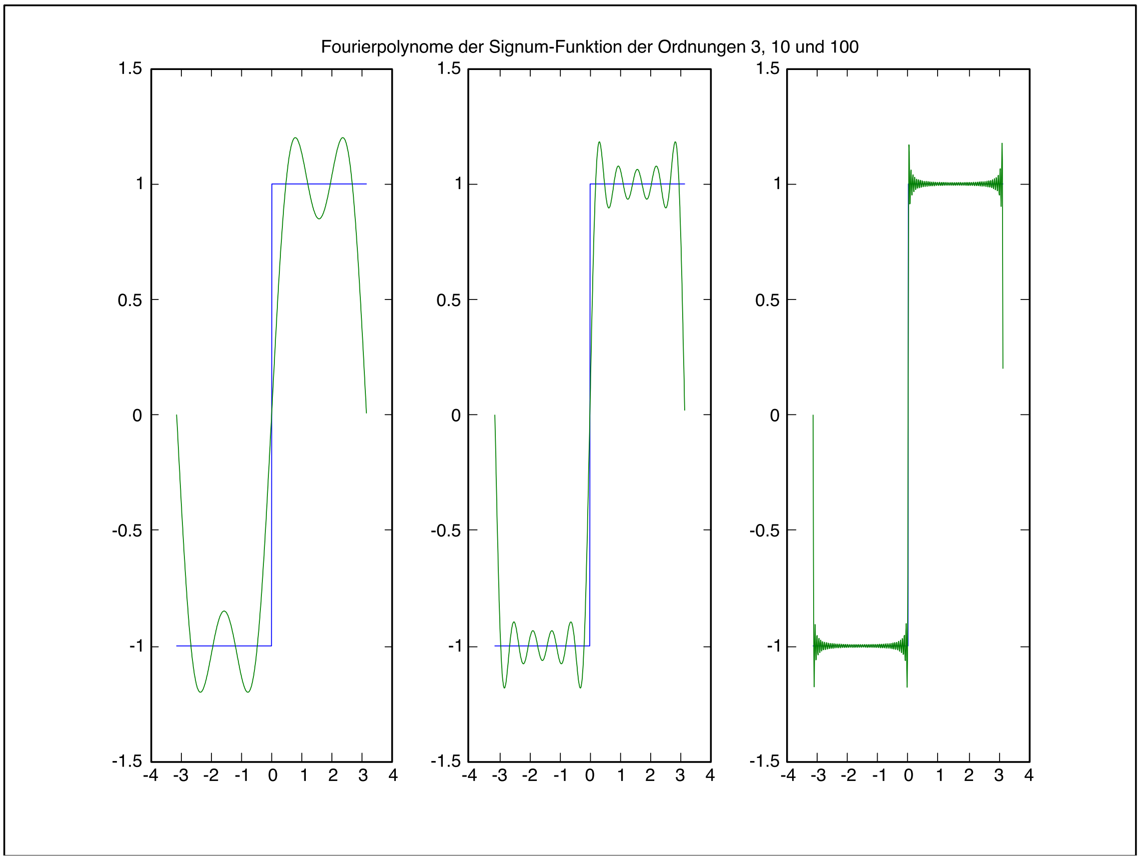
die *Fourierkoeffizienten von* f. Die Funktion

$$S_n(f)(x) := S_n(x) := \frac{a_0}{2}+\sum_{k=1}^{n} a_k\cos(kx)+b_k\sin(kx)$$

heißt *Fourierpolynom der Ordnung* n *von* f und die Reihe

$$S(f)(x) := S(x) := \frac{a_0}{2}+\sum_{k=1}^{\infty} a_k\cos(kx)+b_k\sin(kx)$$

heißt *Fourierreihe* von f.



Satz 280 (Besselsche Ungleichung)**:** Es sei f eine π-periodische, integrierbare Funktion mit Fourierkoeffizienten a_k und b_k. Dann gilt

$$\|f\|^2 \geq \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k^2 + b_k^2\right).$$

BEWEIS: Auf Grund der Linearität gilt

$$\begin{aligned} \|f - S_n(f)\|^2 &= \langle f - S_n(f), f - S_n(f) \rangle \\ &= \langle f, f \rangle - 2 \langle f, S_n(f) \rangle + \langle S_n(f), S_n(f) \rangle . \end{aligned}$$

Ausschreiben des Fourierpolynoms ergibt

$$\begin{aligned}
\langle f, S_n\rangle &= \frac{a_0}{2}\langle f, 1\rangle + \sum_{k=1}^{n}\left(a_k\langle f, \cos(kx)\rangle + b_k\langle f, \sin(kx)\rangle\right)\\
&= \frac{a_0^2}{2} + \sum_{k=1}^{n}(a_k^2 + b_k^2),\\
\langle S_n(f), S_n(f)\rangle &= \frac{a_0^2}{4}\langle 1, 1\rangle + \sum_{k=1}^{n} a_k^2\langle \cos(kx), \cos(kx)\rangle + \sum_{k=1}^{n} b_k^2\langle \sin(kx), \sin(kx)\rangle\\
&= \frac{a_0^2}{2} + \sum_{k=1}^{n}(a_k^2 + b_k^2).
\end{aligned}$$

Zusammen ergibt dies

$$0 \le \|f - S_n\|^2 = \|f\|^2 - \left(\frac{1}{2}a_0^2 + \sum_{k=1}^{n}(a_k^2 + b_k^2)\right),$$

woraus die Besselsche Ungleichung sofort folgt. □

Satz 281 (Parsevalsche Gleichung)**:** Gilt $\lim_{n\to\infty}\|f - S_n(f)\| = 0$, so folgt

$$\|f\|^2 = \frac{1}{2}a_0^2 + \sum_{k=1}^{\infty}(a_k^2 + b_k^2).$$

BEWEIS: Klar. □

Beispiele:

1) Es sei $0 < x < 2\pi$. Für die 2π-periodische Funktion f mit

$$f(x) = \frac{\pi - x}{2}$$

ergeben sich die Fourierkoeffiziente $a_k = 0$ und $b_k = \frac{1}{k}$ für alle k. Außerdem gilt

$$\begin{aligned}
\|f\|^2 =& \frac{1}{\pi} \int_0^{2\pi} \frac{(\pi - x)^2}{4} \, dx \\
=& \frac{1}{4\pi} \int_0^{2\pi} \pi^2 - 2\pi x + x^2 \, dx \\
=& \frac{1}{4\pi} \left(\pi^2 x - \pi x^2 + \frac{1}{3} x^3 \right) \Big|_0^{2\pi} \\
=& \frac{1}{4\pi} \left(2\pi^3 - 4\pi^3 + \frac{8}{3}\pi^3 \right) \\
=& \left(\frac{1}{2} - 1 + \frac{2}{3} \right) \pi^2 \\
=& \frac{1}{6}\pi^2.
\end{aligned}$$

Außerdem gilt $\lim_{n\to\infty} \|f - S_n(f)\|^2 = 0$, was hier nicht bewiesen wird und somit folgt

$$\frac{\pi^2}{6} = \|f\|^2 = \sum_{k=1}^{\infty} b_k^2 = \sum_{k=1}^{\infty} \frac{1}{k^2}.$$

2) Betrachte nun

$$f(x) = \begin{cases} 1, & 0 < x < \pi \\ 0, & x = 0 \\ -1, & -\pi < x < 0 \end{cases}$$

und setze f 2π-periodisch fort. Man erhält nun $\|f\|^2 = 2$ und

$$\begin{aligned}
\forall k :\ & a_k = 0 \\
k = 1, 3, 5, \ldots :\ & b_k = \frac{4}{k\pi} \\
k = 2, 4, \ldots :\ & b_k = 0.
\end{aligned}$$

Damit folgt

$$\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}.$$

Definition 282 (Konvergenz im quadratischen Mittel)**:** Es seien f und f_n für alle $n \in \mathbb{N}$ 2π-periodische, integrierbare Funktionen. Man sagt, die Folge (f_n) *konvergiert im quadratischen Mittel gegen* f, wenn

$$\lim_{n\to\infty} \|f_n - f\| = 0$$

gilt.

BEMERKUNG: Die vorstehende Definition sieht genauso aus, wie die Definition der gleichmäßigen Konvergenz von Funktionenfolgen. Es wird aber hier eine andere Norm benutzt. Deshalb haben wir es hier mit einem anderen Konvergenzbegriff zu tun. Konvergiert die Folge (f_n) gleichmäßig gegen f, so konvergiert sie auch im quadratischen Mittel gegen f. Jedoch folgt aus der Konvergenz von (f_n) im quadratischen Mittel gegen f im Allgemeinen nicht einmal die punktweise Konvergenz gegen f, da sich der Wert des Integrals nicht ändert, wenn man den Integranden in einem Punkt abändert.

Satz 283: Es sei f eine 2π-periodische, integrierbare Funktion, so konvergiert die Folge der Fourierpolynome $S_n(f)$ im quadratischen Mittel gegen f.

BEWEIS: Ein Beweis für diesen Satz findet sich zum Beispiel in [For08a]. □

Exkurs: Diskrete Fourier-Transformation

Man kennt f nur an den Stützstellen $x_i = \frac{2\pi}{N} i$ für $i = 0, \ldots, N-1$. Nun kann man die Fourierkoeffizienten approximieren:

$$\begin{aligned}
a_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx)\, dx \\
&\approx \frac{1}{\pi} \sum_{i=0}^{N-1} f(x_i) \cos(kx_i) \frac{2\pi}{N} \\
&= \frac{2}{N} \sum_{i=0}^{N-1} f(x_i) \cos(kx_i) =: \widehat{a}_k \\
a_0 &= \int_0^{2\pi} f(x)\, dx \\
&\approx \frac{1}{\pi} \sum_{i=0}^{N-1} f(x_i) \frac{2\pi}{N} \\
&= \frac{2}{N} \sum_{i=0}^{N-1} f(x_i) =: \widehat{a}_0 \\
b_k &\approx \frac{2}{N} \sum_{i=0}^{N-1} f(x_i) \sin(kx_i) =: \widehat{b}_k .
\end{aligned}$$

Im folgenden bezeichnet $[\cdot]$ die Gaußklammer. Nun gelten für $i = 0, \ldots, N-1$ folgende Gleichheiten. Ist N ungerade so gilt

$$f(x_i) = \frac{\widehat{a}_0}{2} + \sum_{k=1}^{\left[\frac{N}{2}\right]} \left(\widehat{a}_k \cos(kx_i) + \widehat{b}_k \sin(kx_i) \right),$$

ist andernfalls N gerade so erhält man

$$f(x_i) = \frac{\widehat{a}_0}{2} + \sum_{k=1}^{\frac{N}{2}-1} \left(\widehat{a}_k \cos(kx_i) + \widehat{b}_k \sin(kx_i) \right) + \frac{\widehat{a}_{\frac{N}{2}}}{2} \cos\left(\frac{N}{2} x_i \right).$$

Diese Gleichungen sind exakt und keine Approximationen. Einen Beweis für die Gleichungen findet sich in [Sto05].
Gegeben seien N Werte $f(x_0), \ldots, f(x_{N-1})$. Von diesen berechne man die N Fourierkoeffizienten $\widehat{a}_0, \ldots, \widehat{a}_{[\frac{N}{2}]}$, $\widehat{b}_1, \ldots, \widehat{b}_{[\frac{N}{2}]}$, wobei $\widehat{b}_{[\frac{N}{2}]} = 0$ gilt, falls N gerade ist. Da die Fourierkoeffizienten unterschiedlich wichtige Informationen tragen, kann man einige "vergessen", wenn man Daten reduzieren will. Die Fouriertransformation ist heutzutage ein wichtiges Hilfsmittel zum Beispiel bei der Bildbearbeitung, Tonaufnahmen und beim Scannen. Jedoch gibt es heute noch schnellere Fouriertransformationsverfahren als das der diskreten Fouriertransformation.

Im folgenden Beispiel sieht man zunächst, dass durch die Fouriertransformation keine Informationen verloren gehen. Anschließend werden immer mehr Koeffizienten gelöscht, d.h. es geht immer mehr Information verloren. In dem Beispiel sind die Koeffizienten, die gelöscht werden, allerdings nicht nach ihrem Informationsgehalt ausgewählt worden, sondern es werden einfach die höchsten Koeffizienten gelöscht. Man könnte eventuell für die selben Anzahlen von gelöschten Koeffizienten bessere Ergebnisse erzielen, wenn man die zu löschenden Koeffizienten gezielter auswählen würde.

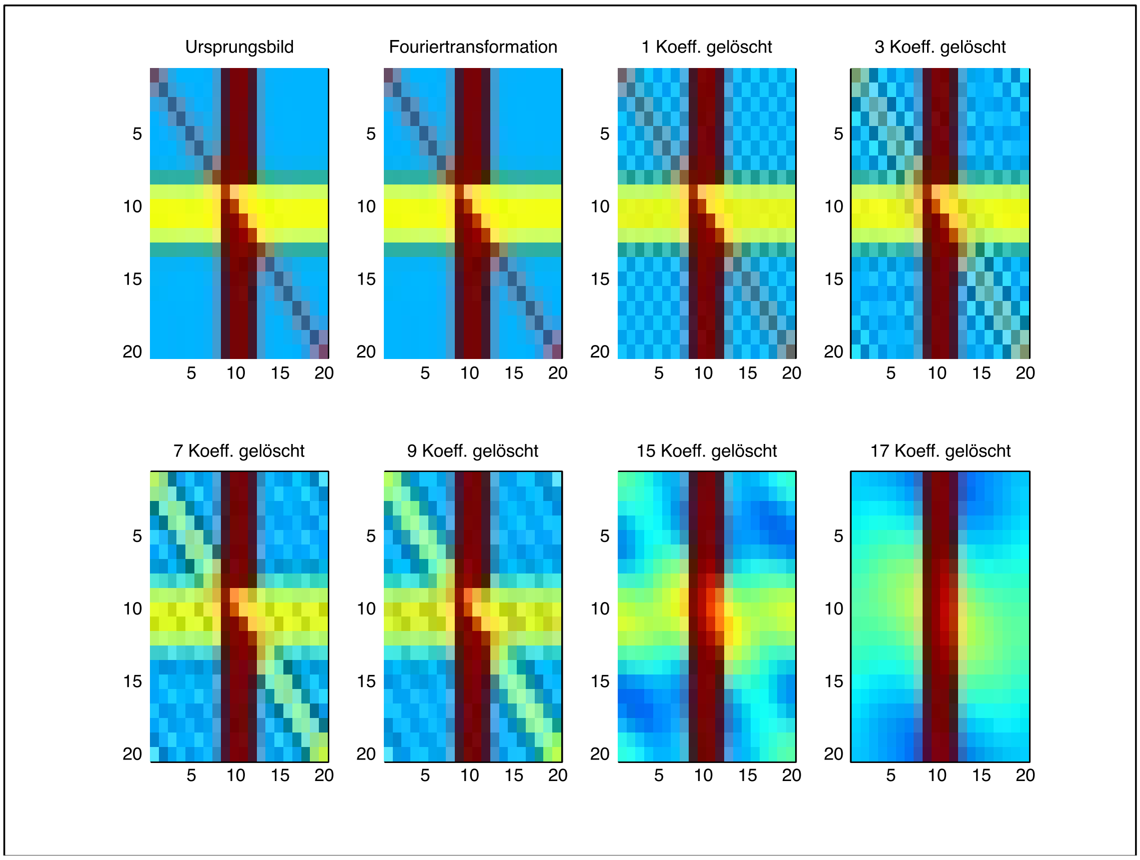
Wir betrachten hier ein Bild mit 20×20 Bildpunkten. Wir betrachten nun die Werte einer jeden Spalte als Funktionswerte $f(x_0), \ldots, f(x_{19})$ und berechnen zu diesen die Fourierkoeffizienten $\widehat{a}_0, \ldots, \widehat{a}_{10}$ und $\widehat{b}_1, \ldots, \widehat{b}_{10}$. Eigentlich müssten wir $\widehat{b}_{10}$ nicht berechnen, da bekannt ist, dass dieser gleich 0 ist. Wir nutzen dieses Wissen in der folgenden Implementierung nicht aus und berechnen 21 Fourierkoeffizienten, obwohl 20 ausreichen würden. Wenn wir dies für jede der 20 Spalten machen, haben wir also 20×20 Fourierkoeffizienten zu berechnen. Die Anzahl der zu speichernden Informationen hat sich also nicht verringert. Wenn wir aber zu jeder Spalte weniger als 20 Koeffizienten speichern, so lässt sich der Speicherbedarf verringern. Damit geht natürlich ein Informationsverlust einher, aber dieser ist unter Umständen besser hinzunehmen als der Informationsverlust, den man hätte, wenn man einen entsprechenden Prozentsatz des Bildes nicht speichern würde, also etwa durch weisse Pixel ersetzen würden.

```
clear all
N=20;
for k=1:N
   for l=1:N
      y(k,l)=20;
   end
end
for k=1:N
       y(9,k)=40;
       y(10,k)=40;
       y(11,k)=40;
       y(12,k)=40;
       y(k,9)=70;
       y(k,10)=70;
```

```
      y(k,11)=70;
      y(k,12)=70;
      y(k,k)=55;
end

figure(1)
clf();
subplot(2,4,1)
image(y);
title('Ursprungsbild');

x=2*pi/N*(0:(N-1));
k=1:floor(N/2);
c=cos(k'*x);
s=sin(k'*x);
a0=2/N*sum(y);
a=2/N*(c*y);
b=2/N*(s*y);

if mod(N,2)==0
   a(N/2,:)=a(N/2,:)/2;
end

M=11
for k=M:floor(N/2)
   a(k,:)=0*ones(1,N);
   b(k,:)=0*ones(1,N);
end

temp=a'*c+b'*s;
temp=temp';
z2=1/2*ones(N,1)*a0+temp;

subplot(2,4,2)
image(z2)
title('Fouriertransformation')

M=10
for k=M:floor(N/2)
   a(k,:)=0*ones(1,N);
   b(k,:)=0*ones(1,N);
end

temp=a'*c+b'*s;
temp=temp';
z=1/2*ones(N,1)*a0+temp;

subplot(2,4,3)
image(z)
title('1 Koeff. gelöscht')

M=9
for k=M:floor(N/2)
   a(k,:)=0*ones(1,N);
   b(k,:)=0*ones(1,N);
end

```

```
temp=a'*c+b'*s;
temp=temp';
z=1/2*ones(N,1)*a0+temp;

subplot(2,4,4)
image(z)
title('3 Koeff. gelöscht')

M=7
for k=M:floor(N/2)
   a(k,:)=0*ones(1,N);
   b(k,:)=0*ones(1,N);
end

temp=a'*c+b'*s;
temp=temp';
z=1/2*ones(N,1)*a0+temp;

subplot(2,4,5)
image(z)
title('7 Koeff. gelöscht')

M=6
for k=M:floor(N/2)
   a(k,:)=0*ones(1,N);
   b(k,:)=0*ones(1,N);
end

temp=a'*c+b'*s;
temp=temp';
z=1/2*ones(N,1)*a0+temp;

subplot(2,4,6)
image(z)
title('9 Koeff. gelöscht')

M=3
for k=M:floor(N/2)
   a(k,:)=0*ones(1,N);
   b(k,:)=0*ones(1,N);
end

temp=a'*c+b'*s;
temp=temp';
z=1/2*ones(N,1)*a0+temp;

subplot(2,4,7)
image(z)
title('15 Koeff. gelöscht')

M=2
for k=M:floor(N/2)
   a(k,:)=0*ones(1,N);
   b(k,:)=0*ones(1,N);
end

temp=a'*c+b'*s;
```

```
temp=temp';
z=1/2*ones(N,1)*a0+temp;

subplot(2,4,8)
image(z)
title('17 Koeff. gelöscht')
```



Ende des Exkurses

5.4. Weitere Arten von Integralen

Satz 284 (Parameterabhängige Integrale)**:** Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $I \subset \mathbb{R}$ ein Intervall. Die Funktion

$$\begin{aligned} f : [a,b] \times I &\to \mathbb{R} \\ (x,y) &\mapsto f(x,y) \end{aligned}$$

sei stetig und die Funtion $\phi : I \to \mathbb{R}$ sei definiert durch

$$\phi(y) := \int_a^b f(x,y)dx.$$

Dann gelten die folgenden Aussagen:

(i) ϕ ist stetig.

(ii) Ist f nach y stetig partiell differenzierbar, so ist ϕ stetig differenzierbar und es ist

$$\phi'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y)dx.$$

BEWEIS:

(i) Es sei $c \in I$ beliebig und (y_n) eine Folge in I mit $\lim_{n\to\infty} y_n = c$. Setzt man

$$F_n(x) := f(x, y_n),$$

so kann man zeigen (was wir hier nicht tun, da wir die Konzepte Kompaktheit und gleichmäßige Stetigkeit im $\mathbb{R}^2$ nicht behandelt haben), dass die Funktionenfolge $(F_n(x))$ gleichmäßig gegen die Funktion F mit $F(x) = f(x, c)$ konvergiert. Dann gilt aber wegen Satz 252

$$\int_a^b F(x)dx = \lim_{n\to\infty} \int_a^b F_n(x)dx$$

beziehungsweise

$$\int_a^b f(x, c)dx = \lim_{n\to\infty} \int_a^b f(x, y_n)dx,$$

was gleichbedeutend ist mit

$$\phi(x) = \lim_{n\to\infty} \phi(y_n).$$

Also ist die Funktion ϕ stetig.

(ii) Es seien (y_n) und c wie zuvor. Man kann zeigen, dass die durch

$$G_n(x) := \frac{f(x, y_n) - f(x, c)}{y_n - c}$$

definierte Funktionenfolge (G_n) gleichmäßig gegen die Funktion G mit

$$G(x) = \frac{\partial f}{\partial y}(x, c)$$

konvergiert. Also gilt wieder nach Satz 252

$$\int_a^b G(x)dx = \lim_{n\to\infty} \int_a^b G_n(x)dx.$$

Nun ist aber

$$\int_a^b G_n(x)dx = \int_a^b \frac{f(x,y_n)-f(x,c)}{y_n-c}dx = \frac{\phi(y_n)-\phi(c)}{y_n-c}$$

und somit

$$\lim_{n\to\infty}\int_a^b G_n(x)dx = \lim_{n\to\infty}\frac{\phi(y_n)-\phi(c)}{y_n-c} = \phi'(c).$$

Andererseits ist

$$\int_a^b G(x)dx = \int_a^b \frac{\partial f}{\partial y}(x,c)dx,$$

womit

$$\phi'(c) = \int_a^b \frac{\partial f}{\partial y}(x,c)dx$$

gezeigt ist. Da $\frac{\partial f}{\partial y}$ stetig ist, ist nach Teil (i) auch ϕ' stetig.

□

Ist nun $f : [a,b]\times[c,d]\to\mathbb{R}$ eine stetige Funktion, so ist nach dem vorstehenden Satz

$$\int_a^b f(x,y)dx$$

eine stetige Funktion in der Variablen y, also insbesondere integrierbar. Es macht also Sinn, dass mehrfache Integral

$$\int_c^d\left(\int_a^b f(x,y)dx\right)dy$$

zu betrachten. Ebenso macht es Sinn, das mehrfache Integral

$$\int_a^b\left(\int_c^c f(x,y)dy\right)dx$$

zu betrachten. Der folgende Satz, den wir nicht beweisen werden, sagt nun, dass unter den gemachten Voraussetzungen die beiden Ausdrücke übereinstimmen.

Satz 285: Es seien $a,b,c,d\in\mathbb{R}$ mit $a<b, c<d$ und $f : [a,b]\times[c,d]\to\mathbb{R}$ eine stetige Funktion. Dann gilt

$$\int_a^b\left(\int_c^d f(x,y)dy\right)dx = \int_c^d\left(\int_a^b f(x,y)dx\right)dy.$$

Definition 286 (Mehrfaches Integral über Rechteckbereiche)**:** Es seien $a, b, c, d \in \mathbb{R}$ mit $a < b, c < d$ und $f : [a, b] \times [c, d] \to \mathbb{R}$ eine beschränkte Funktion. Wir sagen, dass f über $R := \{(x, y) \in \mathbb{R}^2 | a \leq x \leq b, c \leq y \leq d\}$ integrierbar ist, wenn die beiden mehrfachen Integrale

$$\int_a^b \left(\int_c^c f(x, y)dy \right) dx \text{ und } \int_c^d \left(\int_a^b f(x, y)dx \right) dy.$$

existieren und übereinstimmen. Wir setzen dann

$$\int_R f(x, y)d(x, y) := \int_a^b \int_c^d f(x, y)dydx := \int_a^b \left(\int_c^d f(x, y)dy \right) dx.$$

BEMERKUNG: So wie das einfache Riemann-Integral

$$\int_a^b f(x)dx$$

den Flächeninhalt unterhalb des Funktionsgraphen von f zwischen den Grenzen a und b beschreibt, so beschreibt das Integral

$$\int_R f(x, y)d(x, y)$$

das Volumen des Raums unter der Fläche $f(x, y)$ und über dem Rechteck R.

Beispiel: Das Volumen unter der Paraboloidfläche $f(x, y) = x^2 + y^2$ über dem Rechteck $[0, 1] \times [0, 1]$ ergibt sich zu

$$\begin{aligned} \int_0^1 \int_0^1 x^2 + y^2 dydx &= \int_0^1 \left(\int_0^1 x^2 + y^2 dy \right) dx = \int_0^1 (x^2 y + \frac{1}{3} y^3)|_0^1 dx \\ &= \int_0^1 x^2 + \frac{1}{3} dx = (\frac{1}{3} x^3 + \frac{1}{3} x)|_0^1 \\ &= \frac{2}{3}. \end{aligned}$$

Definition 287 (Indikatorfunktion)**:** Ist $D \subset \mathbb{R}^2$, so heißt die Funktion $1_D : \mathbb{R}^2 \to \mathbb{R}$ mit

$$1_D(x, y) := \begin{cases} 1 & , \text{ falls } (x, y) \in D \\ 0 & , \text{ falls } (x, y) \notin D \end{cases}$$

die *Indikatorfunktion der Menge* D.

Mit Hilfe der Indikatorfunktion können wir nun Integrale von Funktionen über beliebigen beschränkten Teilmengen des $\mathbb{R}^2$ definieren.

Definition 288: Es sei $D \subset \mathbb{R}^2$ derart, dass es $a, b, c, d \in \mathbb{R}$ gibt, so dass $D \subset R := \{(x, y) \in \mathbb{R}^2 | a \leq x \leq b, c \leq y \leq d\}$ ist. Ist $f : D \to \mathbb{R}$ eine beschränkte Funktion, so sagen wir, dass f über D integrierbar ist, wenn die Funktion $f \cdot 1_D$ über R integrierbar ist. Wir setzen dann

$$\int_D f(x, y) d(x, y) := \int_R f(x, y) 1_D(x, y) d(x, y).$$

BEMERKUNG: Zur Berechnung des Integrals über D setzen wir die Funktion f also außerhalb von D auf einem Rechteckbereich mit 0 fort und berechnen das Integral über R.

Beispiel: Das Volumen V eines Zylinders mit Radius r und Höhe h lässt sich wie folgt berechnen. Die Grundfläche D des Zylinders ist ein Kreis mit Radius r. D lässt sich in das Rechteck $R := [-r, r] \times [-r, r]$ einbetten. Für alle $(x, y) \in D$ gilt $x^2 + y^2 \leq r^2$. Bei festem $x \in [-r, r]$ muss für y also $-\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2}$ gelten. Ist also $(x, y) \in R$ mit $y > \sqrt{r^2 - x^2}$ oder $y < -\sqrt{r^2 - x^2}$, so ist $1_D(x, y) = 0$. Wir haben also

$$\begin{aligned} V &= \int_D h d(x, y) = \int_R h 1_D(x, y) d(x, y) \\ &= \int_{-r}^{r} \int_{-\sqrt{r^2 - x^2}}^{\sqrt{r^2 - x^2}} h dy dx \\ &= h \int_{-r}^{r} y|_{-\sqrt{r^2 - x^2}}^{\sqrt{r^2 - x^2}} dx \\ &= 2h \int_{-r}^{r} \sqrt{r^2 - x^2} dx = 2hr^2 \frac{\pi}{2} \\ &= r^2 \pi h. \end{aligned}$$

Dabei haben wir ausgenutzt, dass wir das Integral $\int_{-r}^{r} \sqrt{r^2 - x^2} dx$ bereits auf Seite 229 ausgerechnet haben, als wir den Flächeninhalt eines Kreises berechnet haben.

5.5. Mehr über Differentialgleichungen

In diesem Kapitel wollen wir uns mit Systemen von Differentialgleichungen und mit Differentialgleichungen höherer Ordnung beschäftigen, also mit Differentialgleichungen, in denen nicht nur die erste Ableitung, sondern auch höhere Ableitungen vorkommen.

Zunächst müssen wir dafür aber unseren Differenzierbarkeitsbegriff auf Funktionen $f : I \to \mathbb{R}^n$ oder $f : I \to \mathbb{C}^n$ erweitern.

Definition 289: Es sei $I \subset \mathbb{R}$ ein Intervall. Eine komplexwertige Funktion $f : I \to \mathbb{C}$ heißt *(reell-)differenzierbar in* $x_0 \in I$ genau dann, wenn die beiden reellwertigen Funktionen $\operatorname{Re} f : I \to \mathbb{R}$ und $\operatorname{Im} f : I \to \mathbb{R}$ in x_0 differenzierbar sind. Wir setzen dann

$$f'(x_0) := (\operatorname{Re} f)'(x_0) + i(\operatorname{Im} f)'(x_0)$$

Wir erinnern an die Konvention, dass $\mathbb{K}$ wieder den Körper der reellen Zahlen oder den Körper der komplexen Zahlen bezeichnet; innerhalb eines Satzes oder einer Definition aber wieder stets den selben Körper.

Definition 290: Es sei $I \subset \mathbb{R}$ ein Intervall und $f_1, \dots, f_n : I \to \mathbb{K}$ Funktionen. Die Funktion $f : I \to \mathbb{K}^n$ mit

$$x \mapsto \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix} =: f(x)$$

heißt *differenzierbar in* $x_0 \in I$ genau dann, wenn alle Komponentenfunktionen $f_1, \dots, f_n$ differenzierbar sind. Wir setzen dann

$$f'(x_0) := \begin{pmatrix} f_1'(x_0) \\ f_2'(x_0) \\ \vdots \\ f_n'(x_0) \end{pmatrix}.$$

Beispiel: Es sei $v \in \mathbb{K}^n$ und $\lambda \in \mathbb{K}$. Wir definieren die Funktion $f : \mathbb{R} \to \mathbb{R}$ durch

$$f(x) := ve^{\lambda x} = \begin{pmatrix} v_1 e^{\lambda x} \\ v_2 e^{\lambda x} \\ \vdots \\ v_n e^{\lambda x} \end{pmatrix}.$$

Ist $\lambda = \lambda_1 + i\lambda_2$ mit reellen λ_1, λ_2, so erhalten wir für $i = 1, \dots, n$

$$\begin{aligned}
(v_i e^{\lambda x})' &= (v_i e^{(\lambda_1 + i\lambda_2)x})' \\
&= (v_i e^{\lambda_1 x}(\cos(\lambda_2 x) + i\sin(\lambda_2 x)))' \\
&= \lambda_1 v_i e^{\lambda_1 x}(\cos(\lambda_2 x) + i\sin(\lambda_2 x)) + v_i e^{\lambda_1 x}(-\lambda_2 \sin(\lambda_2 x) + i\lambda_2 \cos(\lambda_2 x)) \\
&= \lambda_1 v_i e^{(\lambda_1 + i\lambda_2)x} + \lambda_2 v_i e^{\lambda_1 x}(i^2 \sin(\lambda_2 x) + i\cos(\lambda_2 x)) \\
&= (\lambda_1 + i\lambda_2) v_i e^{(\lambda_1 + i\lambda_2)x} \\
&= \lambda v_i e^{\lambda x}.
\end{aligned}$$

Wir erhalten also insgesamt $f'(x) = \lambda v e^{\lambda x}$. Es ist also $f'(x) = \lambda f(x)$, genauso wie wir es für $\lambda, v \in \mathbb{R}$ schon viel früher gesehen hatten.

Problem 99: *Es seien $f, g : I \to \mathbb{K}^n$ zwei differenzierbare Funktionen. Zeigen Sie, dass $(f(x) + g(x))' = f'(x) + g'(x)$ für alle $x \in I$ gilt. Ist A eine $n \times n$ Matrix über $\mathbb{K}$, so gilt außerdem $(Af(x))' = Af'(x)$ für alle $x \in I$.*

Kommen wir damit nun zu Systemen von Differentialgleichungen. Der einfachste Fall sind n entkoppelte, lineare, homogene Differentialgleichungen, wie wir sie früher bereits betrachtet hatten, also

$$\begin{array}{rcl} y_1' & = & a_1 y_1 \\ y_2' & = & a_2 y_2 \\ \vdots & & \vdots \\ y_n' & = & a_n y_n. \end{array}$$

Die allgemeine Lösung dieses Systems ist gegeben durch

$$y_k(x) = c_k e^{a_k x} \text{ für } k = 1, \dots, n.$$

Setzen wir nun

$$y(x) := \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix},$$

so lässt sich das System auch schreiben als

$$y' = Ay,$$

wobei A die $n \times n$ Diagonalmatrix mit den Einträgen $a_1, \dots, a_n$ auf der Hauptdiagonale ist.

Als nächstes wollen wir uns das System $y' = Ay$ für den Fall anschauen, dass A eine obere Dreiecksmatrix ist.

Also

$$A := \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,n} \\ 0 & a_{2,2} & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & a_{n-1,n} \\ 0 & \cdots & \cdots & 0 & a_{n,n} \end{pmatrix}$$

Die letzte Gleichung des Systems lautet dann

$$y_n' = a_{n,n} y_n.$$

Die allgemeine Lösung dieser linearen Differentialgleichung lautet

$$y_n(x) = c_n e^{a_{n,n} x}.$$

Dann lautet die vorletzte Gleichung

$$y_{n-1}' = a_{n-1,n-1} y_{n-1} + a_{n-1,n} y_n = a_{n-1,n-1} y_{n-1} + a_{n-1,n-1} c_n e^{a_{n,n} x}.$$

Dies ist eine inhomogene, lineare Differentialgleichung, die wir lösen können.

Sind die Lösungen $y_n, y_{n-1}, \ldots, y_{k+1}$ bekannt, so lässt sich y_k als Lösung der inhomogenen, linearen Differentialgleichung

$$y_k' = a_{k,k} + \sum_{l=k+1}^{n} a_{k,l} y_l$$

berechnen.

Satz 291: Es sei A eine $n \times n$ Matrix über $\mathbb{C}$ und S eine invertierbare $n \times n$ Matrix über $\mathbb{C}$. Eine Funktion $\phi : \mathbb{R} \to \mathbb{C}^n$ ist genau dann Lösung des Differentialgleichungssystems $y' = Ay$, wenn die Funktion $\psi := S^{-1}\phi : \mathbb{R} \to \mathbb{C}^n$ eine Lösung von $z' = (S^{-1}AS)z$ ist.

BEWEIS: Gilt $\phi' = A\phi$, so ist

$$(S^{-1}AS)\psi = (S^{-1}AS)S^{-1}\phi = S^{-1}A\phi = S^{-1}\phi' = (S^{-1}\phi)' = \psi'.$$

Andererseits ist $\psi = S^{-1}\phi$ gleichbedeutend mit $\phi = S\psi$. Gilt also $\psi' = (S^{-1}AS)\psi$, so haben wir

$$S^{-1}\phi' = \psi' = (S^{-1}AS)\psi = S^{-1}A\phi.$$

Mulitplikation der vorstehenden Gleichung von links mit S liefert dann $\phi' = A\phi$. $\square$

Aus der Linearen Algebra ist bekannt, dass es zu A genau dann eine intertierbare Matrix S derart gibt, dass $S^{-1}AS$ eine Diagonalmatrix ist, wenn es eine Basis von Eigenvektoren von A gibt. Man kann aber sogar zeigen, dass es zu jeder komplexen $n \times n$ Matrix A eine invertierbare $n \times n$ Matrix S über $\mathbb{C}$ gibt, so dass $S^{-1}AS$ eine obere Dreiecksmatrix ist. Nach dem vorstehenden Satz können wir also jedes System $y' = Ay$ lösen, indem wir $z' = (S^{-1}AS)z$ lösen.

Satz 292: Es sei A eine $n \times n$ Matrix über $\mathbb{C}$. Ist v ein Eigenvektor von A zum Eigenwert λ, so ist die Funktion ϕ, die definiert ist durch

$$\phi(x) := ve^{\lambda x}$$

eine Lösung von $y' = Ay$.
Ist A eine reelle Matrix, so sind reelle Lösungen gegeben durch $\psi_1 := \operatorname{Re}\phi$ und $\psi_2 := \operatorname{Im}\phi$.

BEWEIS: Nach obigem Beispiel ist $\phi' = \lambda\phi$. Andererseits haben wir auch

$$A\phi = Ave^{\lambda x} = \lambda ve^{\lambda x} = \lambda\phi,$$

da $Av = \lambda v$ ist. □

Ist A reell und sind λ und v wie im vorstehenden Satz, so ist auch $\bar{\lambda}$ ein Eigenwert von A und $\bar{v}$ ein zugehöriger Eigenvektor.

Sind $\lambda = \alpha + i\beta$ und $v = a + ib$, so erhält man

$$\operatorname{Re}(ve^{\lambda x}) = (a\cos\beta x - b\sin\beta x)e^{\alpha x}$$

und

$$\operatorname{Im}(ve^{\lambda x}) = (b\cos\beta x + a\sin\beta x)e^{\alpha x}.$$

Satz 293: Ist A eine $n \times n$ Matrik über dem Körper $\mathbb{K}$, so ist der Lösungraum L_H des Differentialgleichungssystems $y' = Ay$ ein n-dimensionaler Untervektorraum des Vektorraums aller Abbildungen $f : \mathbb{R} \to \mathbb{K}^n$. Es gibt also stets n linear unabhängige Lösungen, die eine Basis des Lösungsraums L_H darstellen. Ist $b : I \to \mathbb{K}^n$ eine Funktion und ψ eine Lösung des inhomogenen Systems

$$y' = Ay + b(x),$$

so gilt für den Lösungsraum L_I des inhomogenen Systems

$$L_I = \{\psi + \phi | \phi \in \mathcal{L}_H\}.$$

BEWEIS: Einen Beweis findet man in [Wal00]. □

Definition 294: Es sei $b : I \to \mathbb{R}$ eine Funktion und $a_0, \ldots, a_{n-1}$ seien reelle Zahlen. Eine Differentialgleichung der Art

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1y' + a_0y = b$$

heißt *lineare Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten.*

Setze

$$\begin{array}{rcl} y_0 & := & y \\ y_1 & := & y_0' = y' \\ y_2 & := & y_1' = y'' \\ \vdots & & \vdots \\ y_{n-1} & := & y_{n-2}' = y^{(n-1)}. \end{array}$$

Dann ist y genau dann eine Lösung der Differentialgleichung

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = b,$$

wenn $y_0, y_1, \ldots, y_{n-1}$ das System

$$z' = Az + \bar{b}$$

lösen, wobei

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

und

$$\bar{b} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(x) \end{pmatrix}.$$

Definition 295: Das Polynom

$$P(X) := X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0$$

heißt *charakteristisches Polynom der linearen Dgl.*

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = 0.$$

Satz 296: Es sei P das charakteristische Polynom der Differentialgleichung

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = 0.$$

Ist λ eine Nullstelle von P der Vielfachheit m, das heißt es existiert eine Polynom Q mit $Q(\lambda) \neq 0$ und $P(X) = (X - \lambda)^m Q(X)$, so sind die Funktionen

$$e^{\lambda x}, xe^{\lambda x}, \ldots, x^{m-1}e^{\lambda x}$$

linear unabhängige Lösungen der Differentialgleichung.
Ist λ komplex, so erhält man reelle Lösungen, indem man den Real- und den Imaginärteil der komplexen Lösungen betrachtet.

BEWEIS: Einen Beweis findet man in [Wal00]. □

Beispiel: Wir wollen den vorstehenden Satz auf die Differentialgleichung

$$y'' = -ky$$

anwenden, wobei k eine positive reelle Konstante sein soll. Das zugehörige charakteristische Polynom lautet $P(X) = X^2 + k$. Dieses hat offensichtlich die beiden konjugierten, komplexen Nullstellen $i\sqrt{k}$ und $-i\sqrt{k}$. Somit haben wir also die beiden komplexwertigen Lösungen $e^{i\sqrt{k}}$ und $e^{-i\sqrt{k}}$. Da wir an reellen Lösungen interessiert sind, müssen wir uns Real- und Imaginärteil anschauen und erhalten

$$e^{i\sqrt{k}x} = \cos(\sqrt{k}x) + i\sin(\sqrt{k}x),$$

das heißt die beiden Funktionen $\cos(\sqrt{k}x)$ und $\sin(\sqrt{k}x)$ sind linear unabhängig Lösungen der Differentialgleichung.

Exkurs: Lösen von Differentialgleichungen mit Computeralgebrasystemen

Wir wollen hier kurz zeigen, wie man die im obigen Beispiel behandelte Differentialgleichung mit Hilfe von Maxima lösen kann.

(%i2) `'diff(y,x,2)+k*y=0;`

(%o2) $\frac{d^2}{d\,x^2}\,y+k\,y=0$

(%i3) `ode2(%,y,x);`

Is k positive, negative, or zero? `positive;`

(%o3) $y=\%k1\,\sin\left(\sqrt{k}\,x\right)+\%k2\,\cos\left(\sqrt{k}\,x\right)$

Ende des Exkurses

A. Einiges, was bekannt sein sollte...

In diesem Anhang werden kurz einige Grundlagen dargestellt, die beim Leser als bekannt vorausgesetzt werden.

A.1. Mengen und Abbildungen

Eine *Menge* ist eine Zusammenfassung von verschiedenen Objekten zu einem Ganzen. Die einzelnen Objekte werden *Elemente* genannt.

Endliche Mengen, das heißt Mengen, die nur endlich viele verschiedene Elemente enthalten, können explizit durch die Angabe aller Elemente gegeben werden. Häufig wählt man auch eine beschreibende Darstellung, beispielsweise

$$\begin{aligned} A &= \{1, 2, 3, 4, 5, 6\} \\ &= \{x \mid x \text{ ist eine positive ganze Zahl mit } x \leq 6\} \end{aligned}$$

(dies ist zu lesen als: „die Menge aller x mit der Eigenschaft: x ist eine positive ganze Zahl mit $x \leq 6$ “).

$x \in A$ bedeutet, dass x ein Element der Menge A ist, anderenfalls schreiben wir $x \notin A$.

$\emptyset$ bezeichnet die leere Menge, also die Menge, die kein Element enthält.

$A \subset B$ bedeutet, dass A eine Teilmenge von B ist, das heißt alle Elemente, die in A sind, sind auch in B. Hier ist insbesondere auch der Fall $A = B$ möglich.

Sind A und B Mengen, so definieren wir die *Vereinigung von A und B*

$$A \cup B := \{x \mid x \in A \text{ oder } x \in B\}$$

(das heißt $A \cup B$ ist die Menge aller x, die in A oder in B liegen (oder in beiden!)) und den *Durchschnitt von A und B*

$$A \cap B := \{x \mid x \in A \text{ und } x \in B\}$$

(das heißt $A \cap B$ ist die Menge aller x, die in A und in B liegen). Schließlich ist die *Differenz von A und B* definiert als $A \setminus B := \{x \mid x \in A \text{ und } x \notin B\}$, also die Menge aller x, die in A, aber nicht in B liegen.

Gilt $A \cap B = \emptyset$, so heißen A und B disjunkt.

Ist $A \subset \Omega$, so ist $\overline{A}^{\Omega} := \{x \in \Omega | x \notin A\}$, das sog. *Komplement von* A. Ist die Grundmenge Ω offensichtlich, so schreibt man häufig $\overline{A}$ statt $\overline{A}^{\Omega}$.

Für $A, B, C \subset \Omega$ gelten die folgenden Regeln

- Kommutativität: $A \cap B = B \cap A$ und $A \cup B = B \cup A$,
- Assoziativität: $(A \cap B) \cap C = A \cap (B \cap C)$ und $(A \cup B) \cup C = A \cup (B \cup C)$,
- Distributivität: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ und $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
- die De Morganschen Regeln: $\overline{A \cap B} = \overline{A} \cup \overline{B}$ und $\overline{A \cup B} = \overline{A} \cap \overline{B}$,
- $A \setminus B = A \cap \overline{B}$.

Sind A und B Mengen, so ist das *Mengenprodukt von* A *und* B wie folgt definiert:

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

dabei bezeichnet (a, b) das geordnete Tupel der Elemente a und b. Zwei Tupel (x, y) und (a, b) sind genau dann gleich, wenn $x = a$ und $y = b$ sind.

Induktiv definieren wir das mehrfache Produkt einer Menge mit sich selber. Es ist

$$\begin{aligned} &A^2 := A \times A, \\ &A^n := A^{n-1} \times A \text{ für alle natürlichen Zahlen } n \geq 3. \end{aligned}$$

Man kann A^n mit der Menge aller n-Tupel $(x_1, \ldots, x_n)$ mit $x_i \in A$ für alle $i = 1, \ldots, n$ identifizieren.

Eine Teilmenge $R \subset A \times B$ bezeichnet man auch als *Relation auf* A *und* B. Statt $(x, y) \in R$ schreibt man aber üblicherweise xRy (etwa $3 < 4$ für die Relation $<$).

A.2. Die natürlichen Zahlen, vollständige Induktion und das Wohlordnungsprinzip

Mit $\mathbb{N}$ bezeichnen wir die Menge der natürlichen Zahlen $1, 2, 3, 4, \ldots$ und mit $\mathbb{N}_0$ die Menge $\mathbb{N} \cup \{0\}$. Die Menge der ganzen Zahlen ist mit $\mathbb{Z}$ bezeichnet.

Axiom 4 (Vollständige Induktion)**:** Ist A eine Teilmenge der natürlichen Zahlen, also $A \subset \mathbb{N}$, und gelten die folgenden beiden Eigenschaften

(IV) $1 \in A$,

(IS) $n \in A \Rightarrow n+1 \in A$,

dann gilt: $A = \mathbb{N}$.

Einige typische Anwendungen der vollständigen Induktion finden sich in den Problemen in diesem und im nächsten Abschnitt.

Problem 100 (Binomialkoeffizienten)**:** *Es seien* $n, k \in \mathbb{N}_0$ *und* $k \leq n$. *Zeigen Sie, dass für die* Binomialkoeffizienten

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

wobei $n! := 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n$ *die* n*-te* Fakultät *bezeichnet, gilt:*

(i) $\binom{m}{k} = \binom{m}{m-k}$.

(ii) $\binom{m+1}{k+1} = \binom{m}{k+1} + \binom{m}{k}$.

(iii) $\binom{m}{k} \in \mathbb{N}$.

Problem 101: *Beweisen Sie mittels vollständiger Induktion:*

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

(ii) $3 \cdot \sum_{k=1}^{n+1} (2k-1)^2 = 4n^3 + 12n^2 + 11n + 3$

Man kann zeigen, dass das Axiom der vollständigen Induktion äquivalent zum folgenden Axiom ist:

Axiom 5 (Wohlordnungsprinzip)**:** Jede nicht-leere Teilmenge A der natürlichen Zahlen besitzt ein kleinstes Element, also

$$A \subset \mathbb{N}, A \neq \emptyset \Rightarrow \exists a \in A \forall b \in A : a \leq b.$$

Mit Hilfe des Wohlordnungsprinzips kann man zum Beispiel den folgenden Satz beweisen.

Satz 297 (Division mit Rest)**:** Es seien $a, m \in \mathbb{Z}, m > 1$. Dann gibt es eindeutig bestimmte ganze Zahlen $q, r \in \mathbb{Z}$ mit

$$a = qm + r \text{ und } 0 \leq r < m.$$

A.3. Gruppen, Ringe und Körper

Ist G eine nichtleere Menge, so bezeichnen wir eine Abbildung $\odot : G \times G \to G$ als *Verknüpfung auf* G.

Definition 298 (Gruppe)**:** Ist G eine Menge und $\odot : G \times G \to G$ eine Verknüpfung auf G, so sagen wir, dass das Tupel $(G, \odot)$ eine *Gruppe* ist, wenn die folgenden Bedingungen erfüllt sind:

(A) $\forall a, b, c \in G : (a \odot b) \odot c = a \odot (b \odot c)$, (Assoziativität)

(N) $\exists e \in G \forall a \in G : a \odot e = e \odot a = a$, (Existenz eines neutralen Elements)

(I) $\forall a \in G \exists a' \in G : a \odot a' = a' \odot a = e$. (Existenz von inversen Elementen)

Gilt darüber hinaus noch

(K) $\forall a, b \in G : a \odot b = b \odot a$, (Kommutativität)

so nennen wir die Gruppe *abelsch* oder *kommutativ*.

Oft sagt man einfach, dass eine Menge G eine Gruppe ist und unterschlägt dabei die Verknüpfung. Schreibt man die Verknüpfung mit dem Zeichen $+$, so schreibt man für das neutrale Element üblicherweise 0 statt e und für das zu a inverse Element $-a$ statt a'. Schreibt man die Verknüpfung mit dem Zeichen $\cdot$, so schreibt man für das neutrale Element üblicherweise 1 statt e und für das zu a inverse Element a^{-1} statt a'.

Beispiele:

1) $(\mathbb{N}, +)$ ist keine Gruppe, da zum Beispiel (N) nicht erfüllt ist.

2) $(\mathbb{N}_0, +)$ ist auch keine Gruppe, da es zum Beispiel zu 2 kein Inverses gibt.

3) $(\mathbb{Z}, +)$ ist eine abelsche Gruppe; ebenso $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$.

4) Die Menge der rationalen Zahlen mit der üblichen Multiplikation $(\mathbb{Q}, \cdot)$ ist keine Gruppe, denn es existiert kein Inverses zur 0.

5) $(\mathbb{Q} \setminus \{0\}, \cdot)$ und $(\mathbb{R} \setminus \{0\}, \cdot)$ sind abelsche Gruppen.

6) Bezeichnen wir mit $[a]_m$ den Rest r aus Satz 297 (Division mit Rest), den a bei Division durch m lässt, so können wir auf der Menge $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z} := \{0, 1, 2, \ldots, m-1\}$ eine Addition und eine Multiplikation wie folgt definieren:

$$\begin{aligned} \forall a, b \in \mathbb{Z}_m : \quad & a + b := [a + b]_m \\ \forall a, b \in \mathbb{Z}_m : \quad & a \cdot b := [a \cdot b]_m . \end{aligned}$$

$(Z_m, +)$ ist dann für alle $m > 1$ eine abelsche Gruppe. $(Z_m \setminus \{0\}, \cdot)$ ist genau dann eine abelsche Gruppe, wenn m eine Primzahl ist; dabei heißt eine natürliche Zahl m Primzahl, wenn sie genau zwei Teiler hat, nämlich 1 und sich selber.

Definition 299 (Ring)**:** Ist R eine Menge und sind $+ : R \times R \to R$ („Addition“) und $\cdot : R \times R \to R$ („Multiplikation“) Verknüpfungen auf R, so sagen wir, dass das Tripel $(R, +, \cdot)$ ein *Ring* ist, wenn die folgenden Bedingungen erfüllt sind:

(G+) $(R, +)$ ist eine abelsche Gruppe.

(A*) $\forall a, b, c \in R : (a \cdot b) \cdot c = a \cdot (b \cdot c)$, (Assoziativität der Multiplikation)

(D) $\forall a, b, c \in R : a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ und $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$,(Distributivgesetze)

Gilt darüber hinaus

(K*) $\forall a, b \in G : a \cdot b = b \cdot a$, (Kommutativität der Multiplikation)

so nennen wir den Ring *kommutativ*. Besitzt der Ring R ein neutrales Element bezüglich der Multiplikation, so nennen wir R einen *Ring mit Eins* und jedes $a \in R$, welches ein Inverses bezüglich der Multiplikation besitzt, heißt *invertierbar* oder *Einheit von* R.

Konventionen: Man lässt den Verknüpfungspunkt · bei der Multiplikation häufig weg und schreibt einfach ab statt $a \cdot b$. Es gilt *Punktrechnung vor Strichrechnung*. Dann lauten die Distributivgesetze

$$\forall a, b, c \in R : a(b + c) = ab + ac \text{ und } (a + b)c = ac + bc.$$

Definition 300 (Körper)**:** Wir nennen einen kommutativen Ring mit Eins $(K, +, \cdot)$ einen *Körper*, wenn das neutrale Element der Addition 0_K ungleich dem neutralen Element der Multiplikation 1_K ist und jedes $a \in K \setminus \{0_K\}$ invertierbar ist.

Problem 102: *Wir betrachten die folgende Teilmenge der reellen Zahlen* $\mathbb{R}$*:*

$$K := \{a + b\sqrt{2} | a, b \in \mathbb{Q}\}.$$

Zeigen Sie, dass $(K, +, \cdot)$ *ein Körper ist, wobei* $+$ *und* $\cdot$ *die übliche Addition bzw. Multiplikation auf den reellen Zahlen bedeutet.*
Bemerkung: In dieser Aufgabe darf ohne Beweis benutzt werden, dass $\sqrt{2}$ *irrational ist. Wo wird die Irrationalität von* $\sqrt{2}$ *benutzt?*

B. Wichtige Sätze, Beispiele und Gegenbeispiele

- Der Körper der reellen Zahlen $\mathbb{R}$ ist ein vollständiger, archimedisch angeordneter Körper.

- Der Körper der komplexen Zahlen $\mathbb{C}$ kann nicht angeordnet werden.

- Die Menge
$$\{x \in \mathbb{Q} |\, x^2 \leq 2\}$$
ist beschränkt. Sie besitzt aber kein Supremum in $\mathbb{Q}$. Der Körper der rationalen Zahlen $\mathbb{Q}$ ist nicht vollständig.

- Rechenregeln für Grenzwerte: Es seien (a_n) und (b_n) konvergente Folgen in $\mathbb{R}$ mit $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Dann gilt:

  (i) Die Folge $(a_n \pm b_n)$ ist konvergent mit Grenzwert $a \pm b$.

  (ii) Die Folge $(a_n b_n)$ ist konvergent mit Grenzwert ab; insbesondere ist die Folge (λa_n) für festes $\lambda \in \mathbb{R}$ konvergent gegen λa.

  (iii) Ist $a \neq 0$, so existiert ein $n_0 \in \mathbb{N}$, so dass $a_n \neq 0$ für alle $n \geq n_0$ ist. Die Folge $\left(\frac{1}{a_n}\right)_{n \geq n_0}$ ist konvergent mit Grenzwert $\frac{1}{a}$.

  (iv) Die Folge $(|a_n|)$ ist konvergent mit Grenzwert $|a|$.

- Schachtelungsprinzip: Es seien $(a_n), (b_n), (c_n)$ reelle Zahlenfolgen mit $a_n \leq b_n \leq c_n$ für alle $n \in \mathbb{N}$. Weiterhin seien (a_n) und (c_n) konvergent mit
$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = a.$$
Dann ist auch die Folge (b_n) konvergent und es gilt
$$\lim_{n\to\infty} b_n = a.$$

- Satz von der monotonen Konvergenz: Jede beschränkte, monotone reelle Zahlenfolge (a_n) konvergiert.

- Satz von Bolzano-Weierstraß: Jede beschränkte Folge (a_n) reeller Zahlen besitzt eine konvergente Teilfolge.

- Die geometrische Reihe
$$\sum_{n=0}^{\infty} q^n$$
konvergiert genau für die $q \in \mathbb{C}$, für die $|q| < 1$ ist. Für diese q ist
$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}.$$

- Die Folge $(\frac{1}{n})$ ist eine Nullfolge. Die harmonische Reihe
$$\sum_{n=1}^{\infty} \frac{1}{n}$$
ist aber divergent.

- Leibniz-Kriterium: Ist (a_k) eine monotone Nullfolge reeller Zahlen, so konvergiert die Reihe
$$\sum_{k=1}^{\infty} (-1)^k a_k.$$

- Die alternierende harmonische Reihe
$$\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$$
konvergiert; sie konvergiert aber nicht absolut.

- Majorantenkriterium: Es seien (a_n) und (b_n) zwei Folgen mit $|a_n| \leq |b_n|$ für alle $n \in \mathbb{N}$. Ist dann $\sum_{k=1}^{\infty} b_k$ eine absolut konvergente Reihe, so ist auch die Reihe $\sum_{k=1}^{\infty} a_k$ absolut konvergent.

- Quotientenkriterium: Es sei (a_n) eine Nullfolge in $\mathbb{K}$ mit $a_n \neq 0$ für alle $n \geq n_0$. Es gebe eine reelle Zahl θ mit $0 < \theta < 1$, so dass für alle $n \geq n_0$ die Ungleichung
$$\left| \frac{a_{n+1}}{a_n} \right| \leq \theta$$
erfüllt ist. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$ absolut.

- Wurzelkriterium: Es sei (a_n) eine Folge in $\mathbb{K}$. Es gebe eine reelle Zahl θ mit $0 < \theta < 1$ und ein $n_0 \in \mathbb{N}$, so dass
$$\sqrt[n]{|a_n|} \leq \theta$$
für alle $n \geq n_0$ gilt. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$ absolut.

- Es sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ wieder eine Potenzreihe. Dann tritt genau einer der drei folgenden Fälle ein:

  (i) Die Potenzreihe konvergiert für alle $x \in \mathbb{K}$ absolut.

  (ii) Die Potenzreihe divergiert für alle $x \in \mathbb{K}\backslash\{x_0\}$.

  (iii) Es gibt genau ein $R \in \mathbb{R}_+$, so dass die Potenzreihe für alle $x \in \mathbb{K}$ mit $|x-x_0| < R$ absolut konvergiert und für $|x-x_0| > R$ divergiert.

- Die drei Potenzreihen
$$\sum_{n=1}^{\infty} x^n, \sum_{n=1}^{\infty} \frac{1}{n} x^n, \sum_{n=1}^{\infty} \frac{1}{n^2} x^n$$
  konvergieren alle für $|x| < 1$ und divergieren für $|x| > 1$; d. h. der Konvergenzradius ist in jedem der drei Fälle gleich Eins. Die Reihe $\sum_{n=1}^{\infty} x^n$ divergiert für alle x mit $|x| = 1$. Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2} x^n$ konvergiert für alle x mit $|x| = 1$. Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n} x^n$ ist für einige x mit $|x| = 1$ konvergent, etwa für $x = -1$, und für einige x mit $|x| = 1$ divergent, etwa für $x = 1$.

- Die Exponentialreihe
$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$
  konvergiert für alle $x \in \mathbb{C}$ absolut.

- Die Funktion $f : \mathbb{R} \to \mathbb{R}$, die definiert ist durch
$$f(x) := \begin{cases} 1, & \text{falls } x \geq 0, \\ -1, & \text{falls } x < 0, \end{cases}$$
  ist nicht stetig in 0.

- Die Funktion $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ mit
$$f(x) = \frac{\exp(x) - 1}{x}$$
  ist in 0 stetig fortsetzbar.

- Zwischenwertsatz: Es sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) < f(b)$. Dann existiert zu jedem $c \in [f(a), f(b)]$ ein $p \in [a, b]$ mit $f(p) = c$.

- Jede stetige Funktion $f : [a, b] \to \mathbb{R}$ ist beschränkt und hat in $[a, b]$ (mindestens) eine Maximalstelle und eine Minimalstelle.

- Konvergiert eine Folge stetiger Funktionen (f_n) gleichmäßig gegen eine Funktion f. Dann ist f ebenfalls stetig.

- Es sei $f_n : [0,1] \to \mathbb{R}$ mit $x \mapsto x^n$. Für $x \in [0,1]$ gilt

$$f(x) := \lim_{n\to\infty} f_n(x) = \begin{cases} 0, & x \in [0,1[\\ 1, & x = 1. \end{cases}$$

  Die Folge stetiger Funktionen (f_n) konvergiert also punktweise gegen eine unstetige Funktion f.

- Umkehrfunktionen stetiger Funktionen: Es sei $f : [a,b] \to \mathbb{R}$ eine stetige, streng monoton wachsende Funktion. Setze $A := f(a)$, $B := f(b)$. Dann bildet f das Intervall $[a,b]$ bijektiv auf $[A,B]$ ab. Die Umkehrfunktion $f^{-1} : [A,B] \to \mathbb{R}$ ist ebenfalls stetig und streng monoton wachsend.

- Wichtige Grenzwerte:

  (i) Für alle $a > 0$ gilt

  $$\lim_{n\to\infty} \sqrt[n]{a} = 1.$$

  (ii) Für alle $k \in \mathbb{N}$ gilt

  $$\lim_{x\to\infty} \frac{e^x}{x^k} = \infty \qquad \text{und} \qquad \lim_{x\to\infty} x^k e^{-x} = 0.$$

  Die Exponentialfunktion wächst also schneller als jede Polynomfunktion.

  (iii)

  $$\lim_{x\to\infty} \log(x) = \infty \qquad \text{und} \qquad \lim_{x\searrow 0} \log(x) = -\infty$$

  (iv) Für alle $\alpha > 0$ gilt

  $$\lim_{x\to\infty} \frac{\log(x)}{x^\alpha} = 0 \qquad \text{und} \qquad \lim_{x\searrow 0} x^\alpha \log(x) = 0.$$

  Der Logarithmus wächst langsamer als jede Polynomfunktion.

  (v) Für alle $\alpha > 0$ gilt

  $$\lim_{x\to\infty} x^{-\alpha} = 0 \qquad \text{und} \qquad \lim_{x\searrow 0} x^\alpha = 0.$$

  Aufgrund des zweiten Grenzwertes setzt man häufig $0^\alpha := 0$.

- Jede in einem Punkt x_0 differenzierbare Funktion ist dort stetig.

- Die Betragsfunktion $x \mapsto |x|$ ist in $x = 0$ stetig, aber nicht differenzierbar.

- Ableitungsregeln: Es seien $f, g : I \to \mathbb{R}$ in $x_0 \in I$ differenzierbar und $\lambda \in \mathbb{R}$ beliebig. Dann sind die Funktionen

$$f + g,\ \lambda \cdot f \text{ und } f \cdot g$$

in x_0 differenzierbar mit

(i) $(f + g)'(x_0) = f'(x_0) + g'(x_0)$, (Linearität)

(ii) $(\lambda \cdot f)'(x_0) = \lambda \cdot f'(x_0)$, (Linearität)

(iii) $(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0)$. (Produktregel)

(iv) Ist $g(x_0) \neq 0$, so ist auch $\frac{f}{g}$ in x_0 differenzierbar mit

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}. \qquad \text{(Quotientenregel)}$$

- Kettenregel: Die Hintereinanderausführung zweier differenzierbarer Funktionen $f : I \to J$ und $g : J \to \mathbb{R}$ ist differenzierbar und es gilt

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

- Ableitung der Umkehrfunktion: Es sei $f : I \to J$ eine bijektive Funktion. Ist f außerdem in $y_0 \in I$ differenzierbar mit $f'(y_0) \neq 0$, so ist $f^{-1} : J \to I$ in $x_0 =: f(y_0)$ differenzierbar mit

$$\left(f^{-1}\right)'(x_0) = \frac{1}{f'(y_0)} = \frac{1}{f'(f^{-1}(x_0))}.$$

- Es sei $a < b$, $I :=]a, b[$ und $f : I \to \mathbb{R}$ differenzierbar. Ist $x_m \in I$ eine *Extremstelle von* f, so ist

$$f'(x_m) = 0.$$

- Satz von Rolle: Es sei $a < b$ außerdem sei die Funktion $f : [a, b] \to \mathbb{R}$ stetig und auf $]a, b[$ differenzierbar. Ist außerdem $f(a) = f(b)$, so existiert ein $\xi \in]a, b[$ mit $f'(\xi) = 0$.

- Satz von Taylor: Es sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine $(n + 1)$-mal differenzierbare Funktion. Sind $x, x_0 \in I$, dann existiert für alle $x \in I$ ein $\xi \in]x, x_0[$ (bzw. $\xi \in]x_0, x[$) mit

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n + 1)!}(x - x_0)^{n+1}.$$

- Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f \in C^{n+1}(I)$. Ist $x_0 \in I$ derart, dass

$$f'(x_0) = f^{(2)}(x_0) = \ldots = f^{(n)}(x_0) = 0$$

und $f^{(n+1)}(x_0) \neq 0$ gilt, so folgt:

(i) Ist n gerade, so ist x_0 kein Extremum von f.

(ii) Ist n ungerade, so ist x_0 ein lokales Minimum, falls $f^{(n+1)}(x_0) > 0$ gilt und ein lokales Maximum, falls $f^{(n+1)}(x_0) < 0$ gilt.

- Mittelwertsatz der Differentialrechnung: Es sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, welche auf $]a, b[$ differenzierbar ist. Dann existiert ein $\xi \in]a, b[$ mit

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

- Ist $f :]a, b[\to \mathbb{R}$ differenzierbar, so gilt:

(i) Ist $f'(x) = 0$ für alle $x \in]a, b[$, so ist f auf $]a, b[$ konstant.

(ii) Ist $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$) für alle $x \in]a, b[$, so ist f auf diesem Intervall monoton wachsend (monoton fallend).

(iii) Ist $f'(x) > 0$ (bzw. $f'(x) < 0$) für alle $x \in]a, b[$, so ist f auf diesem Intervall streng monoton wachsend (streng monoton fallend).

- Satz von de l'Hospital: Es sei I ein offenes Intervall und $f, g : I \to \mathbb{R}$ seien differenzierbare Funktionen. Außerdem sei a ein Randpunkt von I, wobei $a = \pm\infty$ zugelassen ist. Für alle $x \in I$, sei $g(x) \neq 0$ und $g'(x) \neq 0$ und es gilt

$$\lim_{x \to a} f(x) = 0 = \lim_{x \to a} g(x)$$

oder

$$\lim_{x \to a} f(x) = \infty = \lim_{x \to a} g(x).$$

Dann folgt

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)},$$

falls der rechte Grenzwert existiert.

- Es sei $A \subset \mathbb{R}^2$ offen, $f : A \to \mathbb{R}$ zweimal stetig differenzierbar und $a \in A$ ein kritischer Punkt von f. Es sei

$$H(f)(a) = \begin{pmatrix} \alpha & \beta \\ \beta & \delta \end{pmatrix} \in M(2 \times 2, \mathbb{R})$$

die Hesse-Matrix von f in a. Dann gilt:

(i) Ist $\alpha\delta - \beta^2 > 0$ und $\alpha > 0$, so hat f in a ein isoliertes, lokales Minimum.

(ii) Ist $\alpha\delta - \beta^2 > 0$ und $\alpha < 0$, so hat f in a ein isoliertes, lokales Maximum.

(iii) Ist $\alpha\delta - \beta^2 < 0$, so hat f in a kein lokales Extremum. Man sagt in diesem Fall, dass a ein Sattelpunkt von f ist.

- Mittelwertsatz der Integralrechnung: Es seien $f, g : [a, b] \to \mathbb{R}$ stetige Funktionen und $g \geq 0$. Dann gibt es ein $\zeta \in [a, b]$ mit

$$\int_a^b f(x)g(x)\, dx = f(\zeta) \int_a^b g(x)\, dx.$$

Insbesondere gilt für $g = 1$

$$\int_a^b f(x)\, dx = f(\zeta)(b - a).$$

- Hauptsatz der Differential- und Integralrechnung: Es sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Für $x \in [a, b]$ definiere $F : [a, b] \to \mathbb{R}$ durch

$$F(x) := \int_a^x f(t)\, dt.$$

Dann gilt:

(i) Die Funktion F ist eine Stammfunktion von f, das heißt für alle $x \in [a, b]$ gilt $F'(x) = f(x)$.

(ii) Ist G eine beliebige Stammfunktion von f, so gilt

$$\int_a^b f(x)\, dx = G(b) - G(a) =: G(x)|_a^b.$$

- Produktregel; partielle Integration: Es seien $u, v : I \to \mathbb{R}$ stetig differenzierbar, dann ist auch uv stetig differenzierbar, und es gilt

$$\int_a^b u(x)v'(x)\, dx = u(x)v(x)|_a^b - \int_a^b u'(x)v(x)\, dx.$$

- Substitionsregel: Es seien $f : I \to \mathbb{R}$ eine integrierbare Funktion und $g : [a, b] \to I$ eine stetig differenzierbare Funktion. Dann gilt

$$\int_a^b f(g(x))g'(x)\, dx = \int_{g(a)}^{g(b)} f(t)\, dt.$$

- Integralkriterium für Reihen: Es sei $f : [1, \infty[\to \mathbb{R}_+ \cup \{0\}$ eine monoton fallende Funktion. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} f(n)$ genau dann, wenn das uneigentliche Integral $\int_1^{\infty} f(x)\, dx$ existiert.

- Es seien $f_n : [a, b] \to \mathbb{R}$ mit $n \in \mathbb{N}$ stetige Funtionen, die gleichmäßig gegen eine Funktion $f : [a, b] \to \mathbb{R}$ konvergieren. Dann gilt

$$\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b f(x)\, dx.$$

- Es seien $f_n : [a, b] \to \mathbb{R}$ mit $n \in \mathbb{N}$ stetig differenzierbare Funtionen, die punktweise gegen eine Funktion $f : [a, b] \to \mathbb{R}$ konvergieren. Die Folge der Ableitungen $f'_n : [a, b] \to \mathbb{R}$ konvergiere gleichmäßig. Dann ist die Grenzfunktion f differenzierbar und es gilt für alle $x \in [a, b]$

$$f'(x) = \lim_{n\to\infty} f'_n(x).$$

- Stirlingsche Formel: Es gilt

$$\lim_{n\to\infty} \frac{N!e^N}{N^N\sqrt{N}} = \sqrt{2\pi}.$$

C. Computerprogramme, die in diesem Skript benutzt werden

C.1. Das Computeralgebrasystem Maxima

Ein gutes Tutorial zu Maxima finden Sie unter `http://www.csulb.edu/~woollett/`. Ansonsten ist die Seite des Maxima-Projekts `http://maxima.sourceforge.net` ein guter Startpunkt für weitere Dokumentationen und Tutorials.

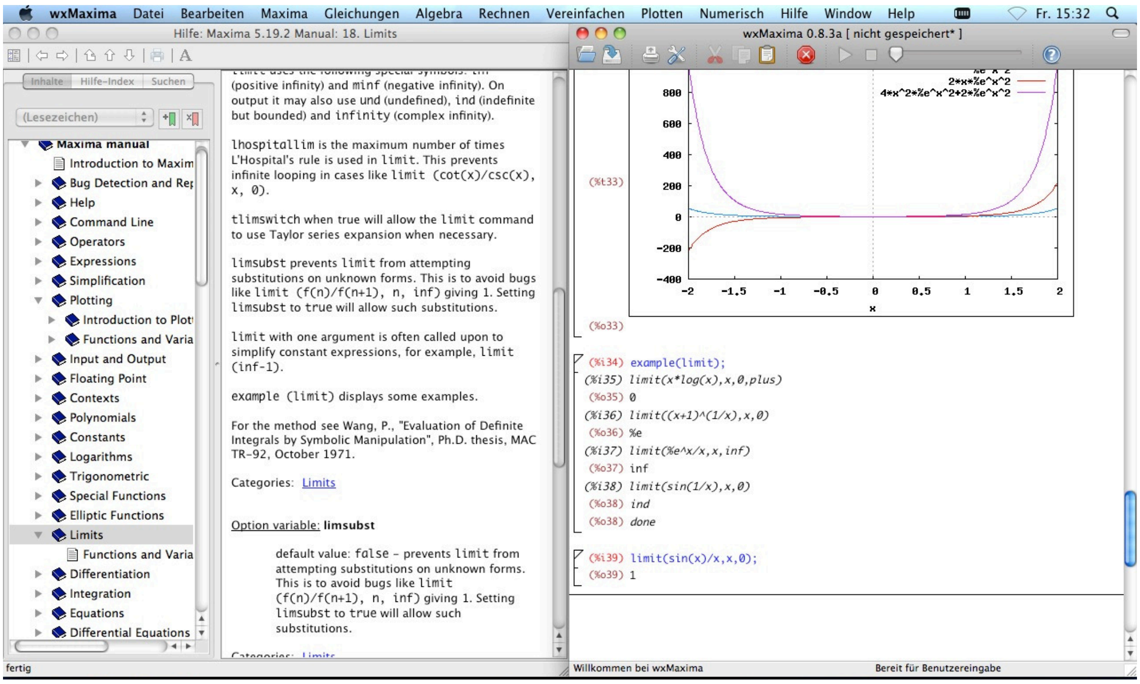


Abbildung C.1.: wxMaxima unter MacOS X

C.2. Das Computeralgebrasystem GeoGebra

Materialien zu GeoGebra finden Sie auf `https://www.geogebra.org`.

C.3. Matlab und Alternativen

Neben Computeralgebrasystemen gibt es auch noch numerische Mathematikpakete, die nicht für symbolisches Rechnen gedacht sind, sondern für numerische Berechnungen. Es handelt sich hierbei also eigentlich nur um Programmiersprachen, die speziell auf mathematische Anwendungen hin konzipiert sind. Wir benutzen hier die Programmiersprache des Pakets Matlab.

Es gibt kostenlose Alternativen zu MATLAB:

- Scilab: http://www.scilab.org
- Freemat: http://freemat.sf.net
- Octave: http://www.gnu.org/software/octave/

Alle drei Pakete gibt es unter Windows, Linux und MacOS X.

Die grundlegenden Elemente der hier beschriebenen Programmiersprache sind in allen Programmen vorhanden. Die meisten Programmbeispiele sollten also auf allen Systemen laufen. Keines der Programme wurde allerdings unter Scilab getestet. Das leistungsstärkste der drei Pakete scheint uns Octave zu sein. Seit der Version 3.8.0 (Erscheiniungsdatum 31.12.2013) kommt Octave mit einer (zur Zeit noch experiementellen) Benutzeroberfläche, die an Matlab erinnert. Auch FreeMat besitzt eine Oberfläche, die sich stark an Matlab orientiert; allerdings ist der Funktionsumfang (noch) recht beschränkt. Die von Scilab benutzte Programmiersprache scheint die meisten Abweichungen zur Matlab-Sprache zu besitzen.

Wer keine Software installieren möchte, kann die Matlab-Sprache auch online ausprobieren. Es gibt verschiedene Anbieter, etwa `http://octave-online.net`.

Es gibt viele Einführungen und Tutorials im Internet. Viele Materialien sind auf den oben genannten Seiten verlinkt. Ein sehr umfangreiches Tutorials finden Sie unter

`http://www.tutorialspoint.com/matlab/index.htm`.

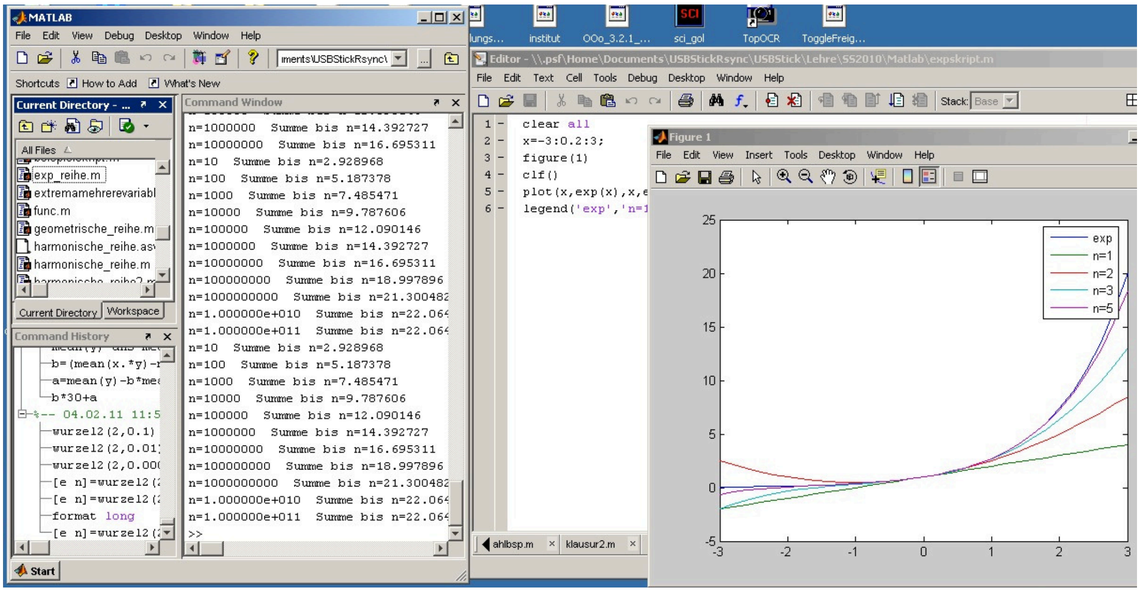


Abbildung C.2.: Matlab unter Windows XP

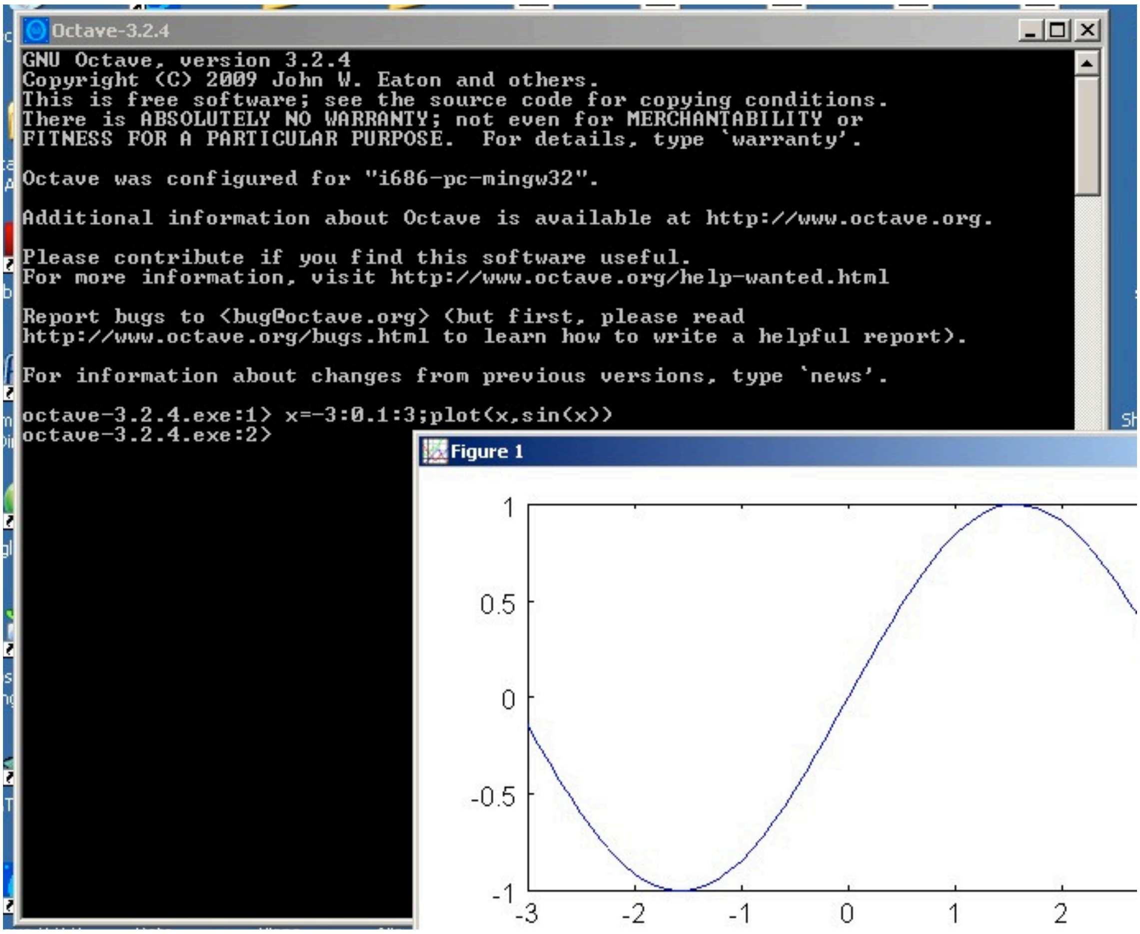


Abbildung C.3.: Octave unter Windows XP

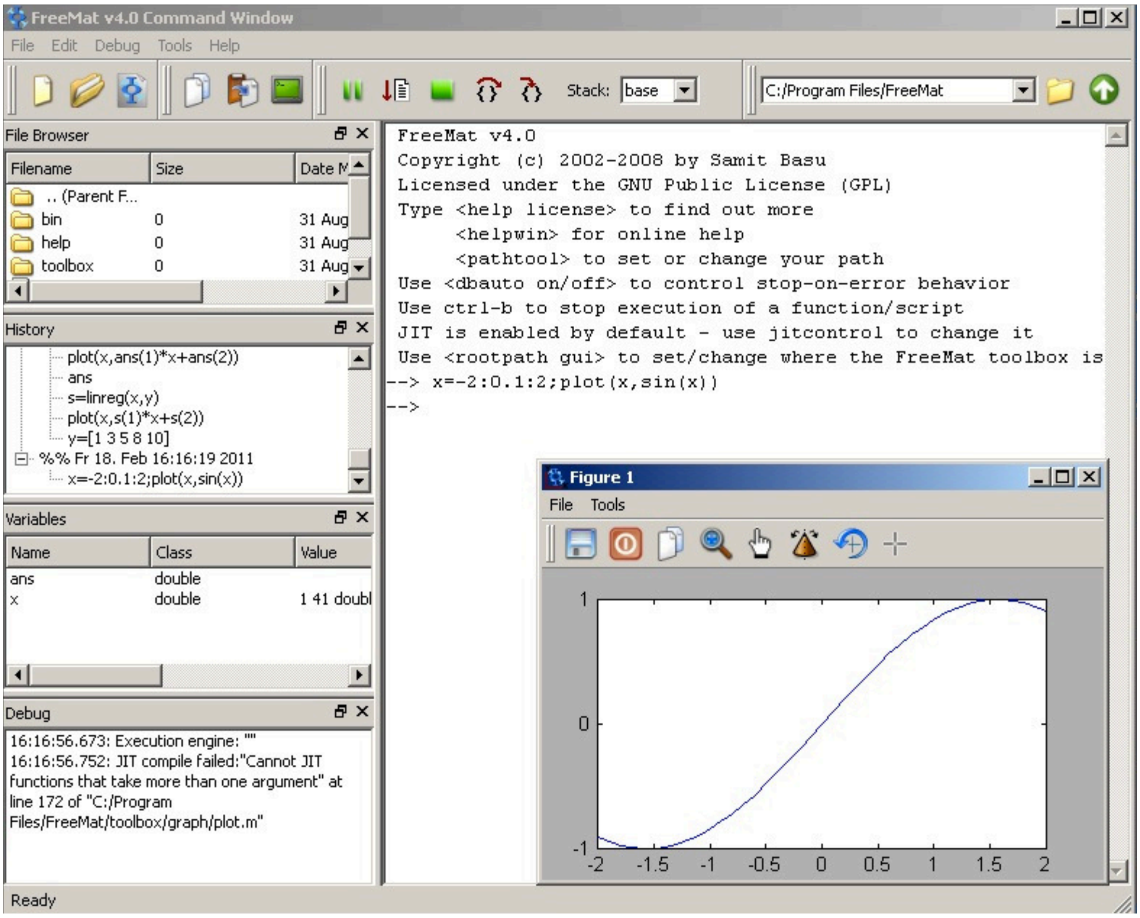


Abbildung C.4.: FreeMat unter Windows XP

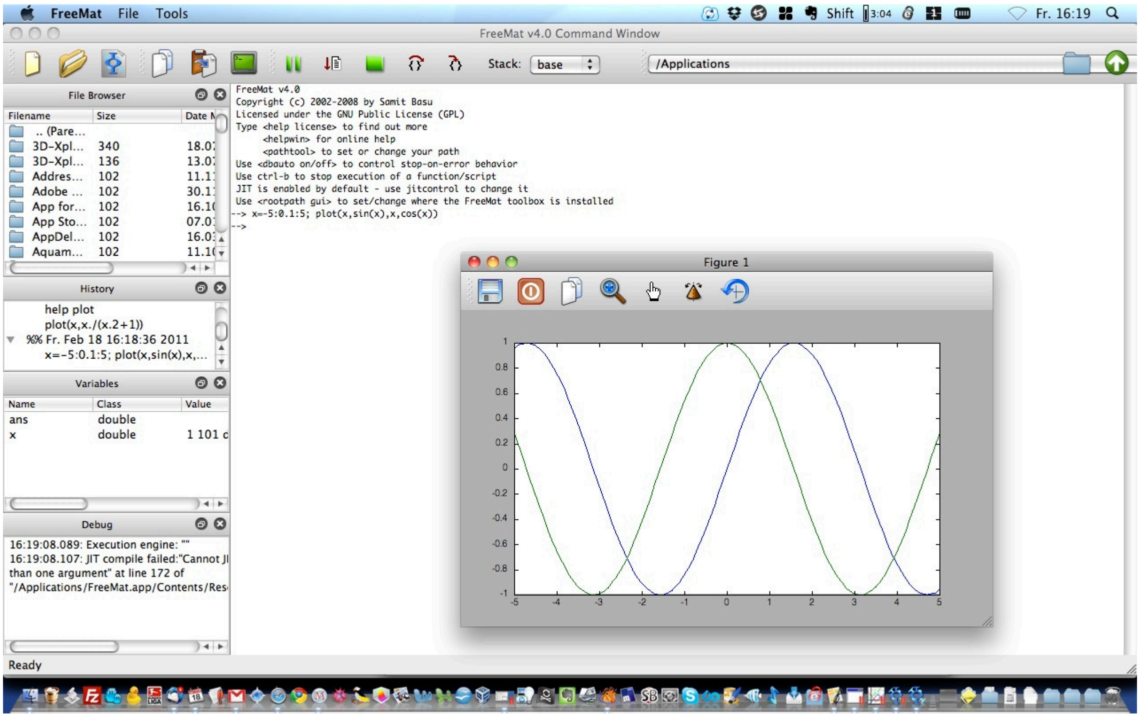


Abbildung C.5.: FreeMat unter MacOS X

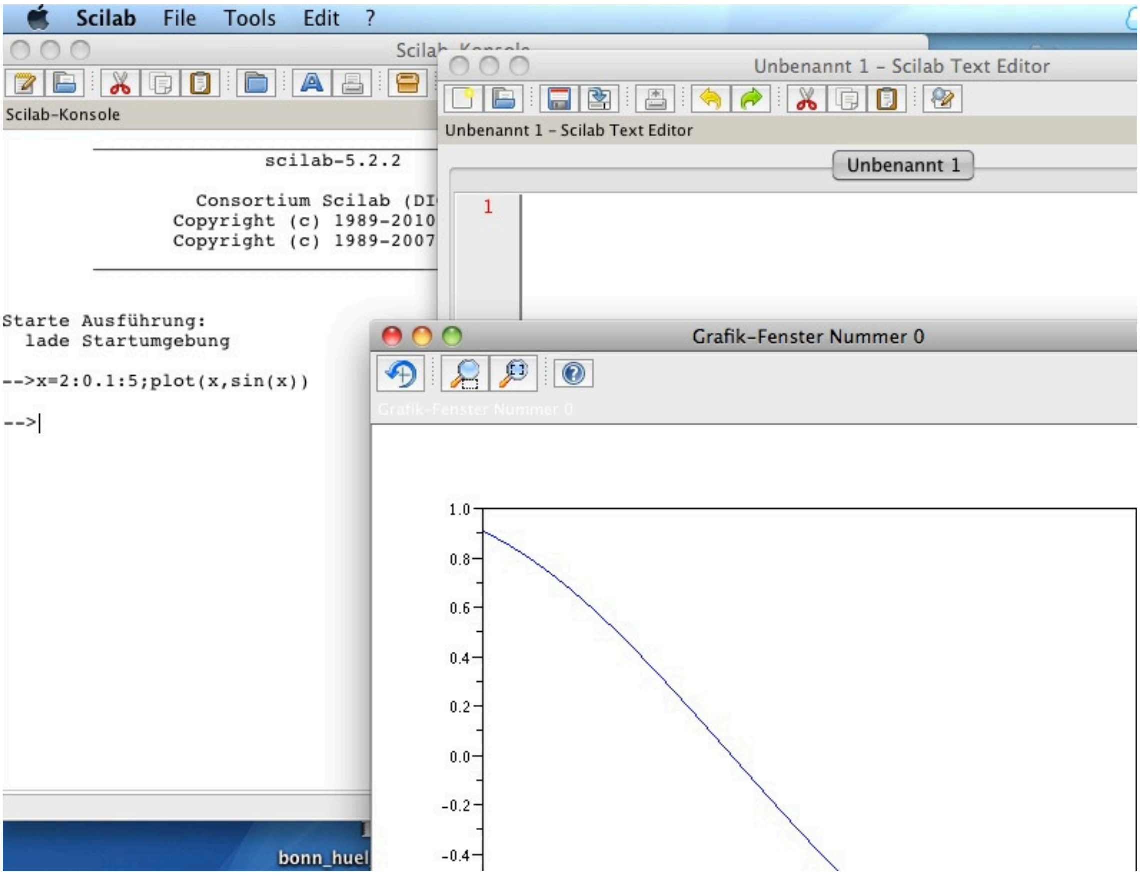


Abbildung C.6.: SciLab unter MacOS X

Literaturverzeichnis

[Fis03] Gerd Fischer. *Lineare Algebra*. Vieweg Verlag, 2003.

[For08a] Otto Forster. *Analysis 1*. Vieweg Verlag, 9. edition, 2008.

[For08b] Otto Forster. *Analysis 2*. Vieweg & Teubner Verlag, 8. edition, 2008.

[Har06] Peter Hartmann. *Mathematik für Informatiker*. Vieweg & Teubner Verlag, 2006.

[Has13] George B. Thomas & Maurice D. Weir & Joel Hass. *Basisbuch Analysis*. Pearson, 2013.

[Ost09] Michael Oberguggenberger & Alexander Ostermann. *Analysis für Informatiker*. Springer-Verlag, 2. edition, 2009.

[Sch94] Gerald Schmieder. *Analysis*. Vieweg-Verlag, 1. edition, 1994.

[Sto05] Josef Stoer. *Numerische Mathematik 1. Eine Einführung – unter Berücksichtigung von Vorlesungen von F. L. Bauer.* Berlin: Springer, 9th ed. edition, 2005.

[Wal00] Wolfgang Walter. *Gewöhnliche Differentialgleichungen. Eine Einführung. 7., neu bearb. und erweiterte Aufl.* Springer-Lehrbuch. Berlin: Springer., 2000.

Sachregister