Von den reellen Zahlen zum Satz von Taylor

Klaus-R. Löffler

Inhaltsverzeichnis

1 Notationen und Sprechweisen

1.1 Mathematische Grundbegriffe

Aussagen Eine (in unserem Fall mathematische) Aussage ist ein sprachliches Gebilde (Text, Formel oder beides gemischt), dem eindeutig genau einer der Wahrheitswerte *wahr* oder *falsch* zugeordnet werden kann[1]. Dabei ist nicht entscheidend, ob diese Zuordnung von einer Person vorgenommen werden kann, sondern nur, dass sie existiert[2]. Beispiele:

- wahr: 3 ist eine ungerade Zahl, $7 \cdot 11 \cdot 13 \cdot 713 = 713713$, Andreas Gryphius starb 1666
- falsch: 3 ist die kleinste Primzahl, $3 + 5 = 35$, Konrad Adenauer war erster deutscher Bundespräsident
- Keine Aussagen in diesem Sinne sind: Der Mond nietet null, 32 ist eine ziemlich große Zahl, die sommerliche Hitze 2018 war eine Strafe des Himmels

Verknüpfungen von Aussagen Durch Verbinden mit *und* (mathematisches Zeichen: $\wedge$) bzw. *oder* (mathematisches Zeichen: $\vee$) werden zwei Aussagen zu einer neuen Aussage verknüpft. Wenn A und B zwei Aussagen sind, ist die Aussage

- $A \wedge B$ (gelesen: A und B) genau dann wahr[3], wenn A und B beide wahr sind,
- $A \vee B$ (gelesen: A oder B) genau dann falsch, wenn A und B beide falsch sind

Im Gegensatz zum gelegentlichen umgangssprachlichen Gebrauch ist dieses „oder“ nicht ausschließend, die Aussage $A \vee B$ ist also auch wahr, wenn A und B beide wahr sind.
Sind A und B Aussagen, so ist die Aussage $A \Rightarrow B$ (gesprochen: Aus A folgt B) wahr, wenn A falsch ist oder wenn B wahr ist, also insbesondere, wenn die Richtigkeit von A (zu zeigen durch logisches Schließen) die Richtigkeit von B nach sich zieht.

Aussageformen Sprachliche Gebilde mit einem oder mehreren Platzhaltern, aus denen durch Einsetzen geeigneter Werte aus einem Einsetzungsbereich Aussagen werden, heißen *Aussageformen.* Beispiele:

$A(x)$: „Die Zahl x ist größer als 41“. Einsetzungsbereich ist zum Beispiel die Menge der ganzen Zahlen.
Dann sind $A(42)$ und $A(4711)$ wahre, $A(40)$ und $A(-112)$ falsche Aussagen.

$B(x)$: „$x \cdot 0 = 0,001$“. Einsetzungsbereich ist z.B. die Menge der rationalen Zahlen.
Dann entsteht für jede Einsetzung von x eine falsche Aussage.

[1] Statt „die Aussage ... ist wahr“ formuliert man oft kürzer „die Aussage ... gilt“.
[2] Behauptet zum Beispiel ein Matrose, eine Körpergröße von 177,7 cm zu haben, ist dies auch dann eine Aussage, wenn er zum Zeitpunkt dieser Behauptung als Schiffbrüchiger auf eine Insel ohne Geräte zur Längenmessung verschlagen worden ist.
[3] In diesem Text werden die Begriffe *wahr* und *richtig* synonym verwendet.

$C(x)$: „$(x+3)^2 = x^2 + 6x + 9$“. Einsetzungsbereich ist z.B. die Menge der reellen Zahlen. Dann entsteht für jede Einsetzung von x eine wahre Aussage. Aussageformen, die für jede zulässige Einsetzung eine wahre Aussage liefern, werden auch als Formeln bezeichnet.

Mengen Eine Menge - meistens bezeichnet mit einem Buchstaben wie $A, B, \ldots M \ldots$ - ist eine Zusammenfassung von unterscheidbaren Objekten. Diese Objekte werden als *Elemente* der Menge bezeichnet. Ist a ein Element der Menge M wird dies notiert als „$a \in M$“[4]. Eine Menge kann durch Aufzählung ihrer Elemente beschrieben werden; dabei erfolgt die Zusammenfassung der Elemente zu einer Menge durch geschweifte Klammern. Dabei ist die so definierte Menge unabhängig davon, in welcher Reichenfolge die Elemente aufgezählt werden[5]. Definiert man eine Menge M durch

$$M := \{2, 5, 77, 2018\}, \text{so sind } 5 \in M, 77 \in M \text{ und } 2018 \in M \text{ wahre Aussagen,}$$

dagegen ist die Aussage $6 \in M$ falsch. Dass ein Objekt a nicht Element einer Menge M ist, wird in der Form $a \notin M$ notiert.
Häufig benutze Standardmengen sind

- Die Menge der *natürlichen* Zahlen[6]: $\mathbb{N} = \{0, 1, 2, 3, 4, 5, 6, \ldots\}$,
- Die Menge der *ganzen* Zahlen: $\mathbb{Z} = \{0, -1, 1, -2, 2, -3, 3, -4, \ldots\}$,

Weitere Standardsymbole für Zahlenmengen sind $\mathbb{Q}$ (Menge der *rationalen* Zahlen und $\mathbb{R}$ (Menge der *reellen* Zahlen.
Anstatt eine Menge durch Aufzählung ihrer Elemente zu definieren, kann man sie auch durch eine charakterisierende Eigenschaft ihrer Elemente beschreiben. Ist z.B. die Menge M definiert durch $M := \{2, 4, 6, 8\}$, kann man M auch als die Menge der geraden ganzen Zahlen zwischen 1 und 9 beschreiben: $M = \{x \in \mathbb{Z} \mid 1 < x < 9 \wedge x \text{ ist gerade}\}$.
Ist allgemeiner $A()$ eine Aussageform mit einem Einsetzungsbereich E, so kann man die Menge M aller Elemente x von E bilden, deren Einsetzen in $A()$ eine wahre Aussage $A(x)$ ergibt.: $M = \{x \in E \mid A(x)\}$. Beispiele:

* $E = \mathbb{N}, A(x) : x^2 < 77$ ergibt $M = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$,
* $E = \mathbb{Z}, A(x) : x^2 = 123,21$ ergibt $M = \{-111, 111\}$,
* $E = \mathbb{Q}, A(x) : x^2 = 2$ ergibt $M = \{\}$, die leere Menge.

Wenn eine Menge M unendlich viele Elemente enthält, wird dies in der Form $\#(M) = \infty$ notiert. Andernfalls bezeichnet $\#(M)$ die Anzahl der Elemente von M.

Verknüpfungen von und Relationen zwischen Mengen Sind A und B Mengen, so versteht man unter dem *Durchschnitt* (auch: *Schnittmenge*) von A und B (in Zeichen: $A \cap B$) die Menge der gemeinsamen Elemente von A und B. Die *Vereinigungsmenge* von A und B (in Zeichen: $A \cup B$) enthält alle Elemente von A und alle Elemente von B. Die Elemente der Vereinigungsmenge sind also diejenigen, die Element von mindestens einer der beiden Mengen A, B sind.
Als *Potenzmenge* von M, wird die Menge bezeichnet, deren Elemente genau die Teilmengen von M sind. Die Potenzmenge einer Menge M wird notiert als 2^M[7].

$$A \cap B = \{x \mid x \in A \wedge x \in B\}, \quad A \cup B = \{x \mid x \in A \vee x \in B\}.$$

[4] Übliche Sprechweisen hierfür sind auch *a gehört zu M* oder *a liegt in M*.

[5] Aus Gründen der Übersichtlichkeit ordnet man oft der Größe nach unterscheidbare Elemente bei der Aufzählung auf- oder absteigend an.

[6] In der deutschen Literatur wird häufig die Zahl 0 nicht den natürlichen Zahlen zugeordnet. Falls man nicht ausdrücklich - wie oben im Text - festgelegt hat, dass die Null eingeschlossen ist, spricht man zur Vermeidung von Missverständnissen von der Menge der *nicht-negativen ganzen Zahlen*.

[7] Wie durch Induktion leicht zu zeigen ist, gilt für jede endliche Menge M die Gleichung $\#(M) = 2^{\#(M)}$.

Die *Restmenge* (oder Differenzmenge) $A \backslash B$ (gesprochen: „A ohne B") ist die Menge aller Objekte, die Element von A, aber kein Element von B sind: $A \backslash B = \{x \mid x \in A \wedge x \notin B\}$.
A ist *Teilmenge* von B (in Zeichen $A \subset B$), wenn alle Elemente von A (auch) Elemente von B sind[8].

Quantoren Ist $A(x)$ eine Aussageform und ergibt für eine Menge M das Einsetzen eines Elements x aus M immer eine wahre Aussage, lässt sich dies mit *Allquantor* notieren:

$$\bigwedge_{x \in M} A(x) \quad (\text{Für alle } x \text{ aus } M \text{ gilt } A(x)) .$$

Gibt es unter den Elementen der Menge M mindestens eines, das beim Einsetzen in die Aussageform A() zu einer wahren Aussage führt, kann man das mit dem *Existenzquantor* beschreiben:

$$\bigvee_{x \in M} A(x) \quad (\text{Es existiert ein } x \text{ aus } M, \text{ für das } A(x) \text{ gilt}).$$

Es ist zu beachten, dass bei der Formulierung „Es existiert ein ..." immer gemeint ist „Es existiert <u>mindestens ein</u> ...". Folgende Formulierungen sind also äquivalent:

$$\bigvee_{x \in M} A(x) \quad \text{und} \quad \{x \mid A(x)\} \neq \{\}$$

Beispiele zur Anwendung von Quantoren:

$$\bigwedge_{n \in \mathbb{N}} \bigvee_{k \in \mathbb{N}} k > n : \qquad \text{Zu jeder natürlichen Zahl gibt es eine größere.}$$
$$\bigvee_{n \in \mathbb{N}} \bigwedge_{k \in \mathbb{N}} k \geq n : \qquad \text{Es gibt eine kleinste natürliche Zahl, (nämlich 0).}$$
$$\bigwedge_{r \in \mathbb{N}} \bigvee_{p \in \mathbb{Z}} \bigvee_{q \in \mathbb{N}} r = \frac{p}{q} : \qquad \text{Jede rationale Zahl lässt sich darstellen als Quotient einer}$$

ganzen und einer natürlichen Zahl.

1.2 Verwendete Sprechweisen

Wie fast überall in der Sprache ist auch in mathematischen Texten die Bedeutung vieler Zeichen- und Wortfolgen kontextabhängig. Beispiele

$x \in M$: Je nach Kontext wird eine Aussage über ein bestimmtes x gemacht oder x ist eine gebundene Variable.
$x \in M$: x ist ein Element der Menge M.
Für alle $x \in M$ gilt $A(x)$: Für alle Elemente x der Menge M ist die Aussage $A(x)$ richtig.

Eine meint je nach Kontext <u>alle</u> oder <u>mindestens eine</u>
Eine zweistellige Primzahl ist ungerade: <u>Jede</u> zweistellige Primzahl ist ungerade.
Die Menge $\{33, 43, 53, 63\}$ enthält eine zweistellige Primzahl: Die Menge $\{33, 43, 53, 63\}$ enthält <u>mindestens eine</u> zweistellige Primzahl[9] .

[8] Dies schließt die Möglichkeit ein, dass A und B in allen Elementen übereinstimmen, also gleich sind ($A = B$).
[9] Wenn ein einem Satz der Form *Die Menge ... enthält ein ...* gemeint ist, dass es ein und nur ein Element mit der genannten Eigenschaft gibt, wird die Formulierung *Die Menge ... enthält <u>genau ein</u> ...* bevorzugt verwendet.

enthält

Die Menge $\mathbb{R}$ enthält die Menge $\mathbb{Q}$: $\mathbb{Q}$ ist eine Teilmenge von $\mathbb{R}$ $(\mathbb{Q} \subset \mathbb{R})$.
Die Potenzmenge von $\mathbb{R}$ enthält die Menge $\mathbb{Q}$: $\mathbb{Q}$ ist ein Element von $2^{\mathbb{R}}$ $(\mathbb{Q} \in 2^{\mathbb{R}})$.

verschieden

Wenn die Rede z.B. von n $(n \in \mathbb{N})$ Zahlen ist, kann gemeint sein, dass die Zahlen beliebig sind oder dass keine zwei gleichen dabei sind. Aus Deutlichkeitsgründen werden daher die Formulierungen *n nicht notwendig verschiedene Zahlen* bzw. *n paarweise verschiedene Zahlen* verwendet.

2 Die reellen Zahlen

Wir gehen von der uns vertrauten Menge $\mathbb{Q}$ der rationalen Zahlen [10] aus, unter denen sich die noch vertrauteren natürlichen Zahlen[11] befinden, und in der zwei Verknüpfungen, genannt Addition und Multiplikation („plus“mit Zeichen „+“und „mal“mit Zeichen „·“) erklärt sind. Die durch Addition verknüpften Zahlen werden dabei bekanntlich als Summanden, die bei der Multiplikation verknüpften Zahlen als Faktoren bezeichnet. Dabei setzen wir voraus, daß die folgenden Eigenschaften erfüllt sind:

(0) Addiert oder multipliziert man zwei rationale Zahlen, so erhält man als Ergebnis wieder eine rationale Zahl.

Man kann diese Eigenschaft auch so formulieren: Die Menge $\mathbb{Q}$ ist gegenüber der Addition und gegenüber der Multiplikation abgeschlossen.
Für die Addition gelten außerdem die folgenden Regeln:

(1) Bei drei Summanden darf man beliebig zwei benachbarte zu einer Teilsumme zusammenfassen: $(a+b)+c = a+(b+c)$. Diese Regel ist als Mittel der rechnerischen Vereinfachung beim Kopfrechnen aus Beispielen wie 128 + 796 + 204 vertraut, wo es einfacher ist, zuerst die Summe aus zweitem und drittem Summanden zu bilden.

(2) Beim Vertauschen von Summanden bleibt die Summe die gleiche: $a + b = b + a$

(3a) Die Addition mit dem Summanden 0 liefert als Ergebnis den anderen Summanden: $a + 0 = a$ und $0 + a = a$. Man formuliert diesen Sachverhalt auch so: Die Zahl 0 ist bezüglich der Addition neutral. Zwar weiß man, dass es keine andere Zahl gibt, die diese Eigenschaft der Null hat, aber dies ist auch eine unmittelbare Folgerung aus der Neutralitätseigenschaft. Denn betrachtet man eine (möglicherweise von 0 verschiedene Zahl) n, die ebenfalls neutral bezüglich der Addition ist, so liefert die Addition $0 + n$ als Ergebnis einerseits n, weil 0 neutral ist, andererseits 0, weil n neutral ist. Somit ist $n = 0$, es gibt also kein weiteres bezüglich der Addition neutrales Element in $\mathbb{Q}$.

(3b) Zu jeder rationalen Zahl a gibt es eine Zahl b, für welche die Addition von a und b die Summe 0 ergibt. Zwar ist uns bekannt, dass es nicht mehrere solcher - als Gegenzahl von a bezeichneten - rationalen Zahlen gibt, es ist aber auch eine unmittelbare Folgerung aus den bisher aufgezählten Eigenschaften. Um dies zu beweisen, überzeugen wir uns, dass aus $a + b = 0$ und $a + p = 0$ folgt $b = p$. Dabei ist p eine (zunächst möglicherweise von b verschiedene Gegenzahl von a). Die spontane Idee zur Führung des Beweises mag zunächst in der korrekten Gleichsetzung $a + b = a + p$ bestehen, um anschließend durch Subtraktion von a auf beiden Seiten zur gewünschten Behauptung zu gelangen. Doch von einer Subtraktion war bisher noch gar nicht die Rede, und

[10] Eine Zahl wird als *rational* bezeichnet, wenn sie sich als Quotient einer ganzen und einer von null verschiedenen natürlichen Zahl erhalten lässt.

[11] In dieser Darstellung ist die Menge der natürlichen Zahlen $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.

erinnert man sich nun, dass Subtrahieren Addieren der Gegenzahl bedeutet, dann wird klar, dass mit dem Subtrahieren auch die Eindeutigkeit der Gegenzahl vorausgesetzt wird, also gerade das, was man hier beweisen will. Wählen wir also einen neuen Ansatz und verwenden hierbei, dass b und p Gegenzahlen zu a sind:

$$b = b + 0 = b + (a + p) = (b + a) + p = (a + b) + p = 0 + p = p, \quad \text{also} \quad b = p.$$

Man beachte, dass hierbei sowohl von der Eigenschaft der 0 als neutralem Element, den Eigenschaften von b und p als Gegenzahlen zu a, sowie dem Assoziativgesetz Gebrauch gemacht wurde.

Man fasst die Eigenschaften (0) bis (3b) in der folgenden Formulierung zusammen: Die Menge der rationalen Zahlen bildet bezüglich der Addition eine kommutative Gruppe - oder kürzer: $(\mathbb{Q}, +)$ ist eine kommutative Gruppe.
In fast analoger Weise lassen sich die Eigenschaften der rationalen Zahlen bezüglich der Multiplikation darstellen.

(1) Die Multiplikation ist assoziativ: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

(2) Die Multiplikation ist kommutativ: $a \cdot b = b \cdot a$

(3a) Es gibt ein neutrales Element, nämlich die Zahl 1, sodass für alle rationalen Zahlen a die Gleichungen $1 \cdot a = a$ und $a \cdot 1 = a$ gelten.

Schließlich sind Addition und Multiplikation noch durch das Distributivgesetz verbunden:

(4) $a \cdot (b + c) = a \cdot c + b \cdot c$. Auch von diesem Gesetz machen wir beim praktischen Rechnen häufig Gebrauch, zum Beispiel wenn wir im Kopf die Aufgabe $5 \cdot 37$ als $5 \cdot 7 + 5 \cdot 30$ lösen.

Anscheinend vergessen wurde zunächst Regel (3b), die analog zu den Eigenschaften der Addition heißen müsste: Zu jeder rationalen Zahl a gibt es eine Zahl b (die bekanntlich als Kehrwert von a bezeichnet wird), sodass gilt $a \cdot b = 1$ und $b \cdot a = 1$. Dem steht entgegen, dass wir gelernt haben $0 \cdot b = 0$, sodass entweder gelten müsste $0 = 1$ oder aber 0 keine Kehrzahl hat. Allerdings: Wäre 0 = 1, dann wäre 0 neutrales Element, also $0 \cdot b = b$ und somit $b = 0$; die Menge $\mathbb{Q}$ bestünde nur aus dem einzigen Element 0, was den Widerspruch beseitigen, aber gleichzeitig die Betrachtung der Menge ziemlich uninteressant machen würde. Aber wissen wir überhaupt, dass die Gleichung $0 \cdot b = 0$ richtig ist? Prüfen wir es nach: Bezeichnen wir die Gegenzahl von $0 \cdot b$ mit p, so gilt:

$$0 = 0 \cdot b + p = (0 + 0) \cdot b + p = (0 \cdot b + 0 \cdot b) + p = 0 \cdot b + (0 \cdot b + p) = 0 \cdot b + 0 = 0 \cdot b,$$

woraus sich ergibt, dass wirklich die Multiplikation von 0 mit irgendeiner rationalen Zahl aufgrund der vorausgesetzten Regeln stets 0 ergibt. Die Zahl 0 wird aufgrund dieser Eigenschaft als *absorbierendes* Element der Multiplikation bezeichnet; es kann also keine Kehrzahl zu 0 geben.
Andererseits kann aber nie das Produkt zweier von 0 verschiedener Zahlen a und b den Wert 0 haben, wenn man voraussetzt, dass jede von 0 verschiedene Zahl einen Kehrwert hat. Denn sind a und b von 0 verschiedene Zahlen und p ist der Kehrwert von b, so folgt aus $a \cdot b = 0$ die Gleichung $(a \cdot b) \cdot p = 0 \cdot p$, also $a \cdot (b \cdot p) = 0$ und somit $a \cdot 1 = 0$, mithin $a = 0$, was ja zuvor ausgeschlossen war. Ein Produkt zweier rationaler Zahlen hat also genau dann den Wert null, wenn mindestens einer der Faktoren diesen Wert hat.
Nicht nur die Menge $\mathbb{Q}$, sondern auch die Menge der rationalen Zahlen ohne 0 (Bezeichnung $\mathbb{Q}^*$) ist abgeschlossen bezüglich der Multiplikation. Hier gelten nun die analogen Regeln zur Addition in $\mathbb{Q}$, insbesondere

(3b) Zu jeder von 0 verschiedenen rationalen Zahl a gibt es eine von 0 verschiedene rationale Zahl b (die Kehrzahl oder der Kehrwert) von a, sodass gilt: $a \cdot b = b \cdot a = 1$.

Mit den Eigenschaften (0) bis (3b) für $(\mathbb{Q}, +)$ bzw. $(\mathbb{Q}^*, \cdot)$ und dem Distributivgesetz (4) bildet die Menge der rationalen Zahlen bezüglich Addition und Multiplikation einen Körper.
Oder in kürzerer Formulierung: $(\mathbb{Q}, +, \cdot)$ ist ein Körper.
Fasst man abstrahierend die Gruppeneigenschaften von $(\mathbb{Q}, +)$ bzw. $(\mathbb{Q}^*, \cdot)$ zusammen, indem man den Platzhalter M für die Menge $\mathbb{Q}$ bzw $\mathbb{Q}^*$ und den Platzhalter $\circ$ für die Verknüpfungszeichen „+“ bzw. „$\cdot$ “verwendet, so hat man:

(0) Die Menge M ist bezüglich $\circ$ abgeschlossen: Für alle Elemente a, b aus M gibt es ein eindeutig bestimmtes c aus M, sodass gilt $a \circ b = c$.

(1) (1) Für alle a, b, c aus M gilt $(a \circ b) \circ c = a \circ (b \circ c)$.

(2) (2) Für alle a, b aus M gilt $a \circ b = b \circ a$.

(3a) Es gibt ein Element n in M, sodass für alle a aus M gilt: $n \circ a = a \circ n = a$. n heißt neutrales Element; erinnert die Verknüpfung an eine Addition, wird es als Nullelement oder gelegentlich als o bezeichnet, erinnert die Verknüpfung an eine Multiplikation, bezeichnet man das neutrale Element häufig als Einselement bzw. mit e.

Es kann nicht mehrere bezüglich $\circ$ neutrale Elemente geben.

(3b) Zu jedem Element a aus der Menge M gibt es ein Element b aus M, sodass die Gleichung $a \circ b = n$ erfüllt ist. b heißt *zu a inverses Element*. Es kann nicht mehrere zu a inverse Elemente geben. Das zu a inverse Element wird als a^{-1} notiert. Ist die Verknüpfung eine Addition, ist es üblich, das zu a inverse Element als *Gegenzahl* von a (in Zeichen: $-a$) zu bezeichnen. Ist die Verknüpfung $\circ$ eine Multiplikation, wird das zu a inverse Element manchmal als *Kehrwert* von a bezeichnet. Ist speziell das neutrale Element mit 1 bezeichnet, notiert man den Kehrwert von a als $\frac{1}{a}$.

In der kommutativen Gruppe $(M, \circ)$ ist auf natürliche Weise eine Umkehroperation $\bullet$ zu definieren, indem man für a, b aus M definiert: $a \bullet b = a \circ b^{-1}$.
Die Umkehroperation zur Addition wird als Subtraktion, die Umkehroperation zur Multiplikation als Division bezeichnet. Abkürzend schreibt man speziell für die Umkehrungen von Addition und Multiplikation

- anstelle von $a + (-b)$ kürzer $a - b$; das Zeichen „$-$“wird also als Operationszeichen für das Bilden der Gegenzahl und als Operationszeichen für die Umkehrung der Addition benutzt,
- anstelle von $a \cdot \frac{1}{b}$ kürzer $\frac{a}{b}$.

Übungsaufgaben: Beweise nur unter Verwendung der Rechenregeln (0)..(4) und ggf. bereits hergeleiteter Resultate die Richtigkeit der folgenden Formeln:

(i) $-(-a) = a$

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

(iii) $a \cdot (b - c) = a \cdot b - a \cdot c$

(iv) $\frac{a \cdot b}{a \cdot c} = \frac{b}{c}$

(v) $(-1) \cdot a = -a$

(vi) $-\frac{1}{a} = \frac{-1}{a}$

Die rationalen Zahlen bilden aber nicht nur einen Körper (bezüglich + und ·), sie sind auch als Menge *vollständig geordnet*[12].
Eine Zahl x, für welche $0 < x$ gilt, wird als positiv bezeichnet, eine Zahl x, für die $x < 0$ gilt, heißt negativ. Die Menge der rationalen Zahlen setzt sich also aus den positiven Zahlen, der Zahl 0 und den negativen Zahlen zusammen. Anstelle von $a < b$ verwendet man gleichwertig die Schreibweise $b > a$.
Die folgenden Zusammenhänge zwischen den Verknüpfungen + und · und der Ordnung in $\mathbb{Q}$ werden vorausgesetzt:

(M+) Für beliebiges c folgt aus $a > b$ stets $a + c > b + c$ (Monotoniegesetz).

(I) Aus $c < 0$ und $a < b$ folgt $a \cdot c > b \cdot c$ (Inversionsgesetz).

Hieraus folgt beispielsweise, dass die Zahl 1 positiv ist, denn aus $1 < 0$ folgt nach dem Inversionsgesetz $1 \cdot 1 > 1 \cdot 0$, also $1 > 0$ mit offensichtlichem Widerspruch.

Übungsaufgaben : Beweise

(1) das Monotoniegesetz der Multiplikation: Aus $c > 0$ und $a < b$ folgt $a \cdot c < b \cdot c$.

(2) ... , dass jede negative Zahl einen negativen Kehrwert hat.

(3) ... , dass jede negative Zahl eine positive Gegenzahl hat.

(4) ... , dass für jede von 0 verschiedene Zahl x die Zahl x^2 positiv ist.

Es erweist sich häufig als praktisch, in einer geordneten Menge M die Fälle $a < b$ und $a = b$ in der Bezeichnung $a \leq b$ zusammenfassen. In einer mit $\leq$ geordneten Menge M gelten dann die folgenden Eigenschaften:

(1) Für jedes x aus M gilt $x \leq x$ (Reflexivität),

(2) für x, y aus M folgt aus $x \leq y$ und $y \leq x$ stets $x = y$ (Antisymmetrie),

(3) für x, y, z aus M folgt aus $x \leq y$ und $y \leq z$ stets $x \leq z$ (Transitivität).

Wenn M eine geordnete Menge ist, a ein Element von M, und T eine Teilmenge von M, dann heißt s *eine obere Schranke* T, wenn für alle Elemente x aus T gilt $x \leq s$. Analog ist s eine *untere Schranke* von T, wenn $s \leq x$ für alle x aus T gilt.
Die Menge T heißt nach *oben beschränkt* bzw. *nach unten beschränkt*, wenn sie eine obere bzw. untere Schranke hat. Sie heißt *beschränkt*, wenn sie sowohl eine obere wie eine untere Schranke hat.
Wenn s eine obere Schranke von T ist, und t ist größer als s, dann ist auch t eine obere Schranke von T. Der einfache Beweis dieser Aussage wird als Übung empfohlen.

Beispiele:

(1) Die Menge $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ der natürlichen Zahlen ist nach unten beschränkt.

(2) Die Menge $\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots\}$ ist nach unten und nach oben beschränkt. Untere Schranken sind $\frac{1}{2}$ und jede kleinere Zahl, obere Schranke ist 1 und jede größere Zahl.

(3) Die Menge aller positiven rationalen Zahlen x, für welche $x^2 \leq 2$ gilt, ist beschränkt. Denn eine untere Schranke ist zum Beispiel 0, eine obere Schranke zum Beispiel 2.

In allen drei Beispielen gibt es unter allen unteren Schranken eine kleinste, nämlich 0 bzw. $\frac{1}{2}$ bzw. 0. Anders sieht es bei den oberen Schranken aus.:

[12] Das bedeutet, dass je zwei rationale Zahlen a und b in genau einer der drei folgenden Beziehungen stehen: $a < b$ oder $a = b$ oder $b < a$, wobei folgende Transitivitätseigenschaft erfüllt ist: Aus $a < b$ und $b < c$ folgt $a < c$.

- Die Menge $\mathbb{N}$ hat überhaupt keine obere Schranke, was hier nicht bewiesen, sondern nur postuliert werden soll - und sich später aus den Axiomen für die reellen Zahlen ergibt.
- Unter den oberen Schranken der Menge $\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots\}$ gibt es eine kleinste, nämlich 1.
- Unter den oberen Schranken der Menge aller positiven rationalen Zahlen x, für welche $x^2 \leq 2$ gilt, gibt es in der Menge der rationalen Zahlen keine kleinste.

Diese letzte Aussage wird nachfolgend bewiesen, indem gezeigt wird: Wenn s eine obere Schranke der betrachteten Menge ist und man setzt $t := \frac{1}{2}(s + \frac{2}{s})$, dann ist t eine kleinere obere Schranke. Dazu werden die folgenden vier Aussagen (0)...(3) betrachtet:

$$(0) \quad s^2 \neq 2, \qquad (1) \quad t > 0, \qquad (2) \quad t^2 \geq 2, \qquad (3) \quad t < s.$$

Zu (0): Es gehört zu den Standardbeweisen, die bereits in der Schule geführt werden, zu zeigen, dass keine rationale Zahl das Quadrat 2 hat[13].

Zu (1): Summen und Quotienten positiver Zahlen sind positiv, also ist t positiv.

Zu (2): Da s^2 verschieden von 2 ist, ist $s^2 - 2 \neq 0$ und daher $(s^2 - 2)^2 > 0$.

$$t^2 = (\frac{1}{2}(s + \frac{2}{s}))^2 = \frac{1}{4}(s^2 + 4s + \frac{4}{s^2}) = \frac{1}{4}(s^2 - 4s + \frac{4}{s^2} + 8s) = \frac{1}{4}(s - \frac{2}{s})^2 + 2s > 2$$

Zu (3) Aus $s^2 > 2$ folgt $\frac{2}{s} < s$, also $s + \frac{2}{s} < 2s$ und mithin $\frac{1}{2}(s + \frac{2}{s}) < s$.

Es gibt also nichtleere Teilmengen von $\mathbb{Q}$, die zwar nach oben beschränkt sind, aber keine kleinste obere Schranke haben. Das Postulieren einer solchen Schranke führt zur Ausdehnung der rationalen auf die reellen Zahlen[14]. Von der Menge $\mathbb{R}$ der reellen Zahlen setzen wir voraus, dass sie einen vollständig geordneten Körper bilden, der die Menge der natürlichen Zahlen umfasst und in dem jede nach oben beschränkte nichtleere Menge eine kleinste obere Schranke hat. Als Körper, der die natürlichen Zahlen umfasst, schließt $\mathbb{R}$ auch die ganzen Zahlen ein (die negativen Zahlen als Gegenzahlen zu den positiven natürlichen Zahlen) und somit auch alle rationalen als Ergebnisse von Quotienten aus ganzen Zahlen. Die kleinste obere Schranke einer Teilmenge T von $\mathbb{R}$ wird *Supremum* von T (in Zeichen sup T) genannt. Ist das Supremum einer Menge T gleichzeitig Element von T, wird es als Maximum von T (in Zeichen: max T) bezeichnet. Analog werden Infimum einer Menge und Minimum einer Menge erklärt.
Äquivalent zur Aussage, dass die Menge $\mathbb{N}$ der natürlichen Zahlen nach oben nicht beschränkt ist, formulieren wir das folgende *archimedische Axiom*: Zu jeder positiven reellen Zahl ε gibt es eine natürliche Zahl n, deren Kehrwert kleiner als ε ist: $\frac{1}{n} < \varepsilon$.
Außerdem fordern wir noch, dass die rationalen Zahlen in $\mathbb{R}$ dicht liegen, dass es also zu jeder reellen Zahl r und zu jeder noch so kleinen reellen Zahl ε eine rationale Zahl q mit $r < q < r + \varepsilon$ gibt.
Reelle Zahlen, die nicht in $\mathbb{Q}$ liegen, werden als irrationale Zahlen bezeichnet. Als Zeichen für die Menge der positiven reellen Zahlen wird nachfolgend gelegentlich die Notation $\mathbb{R}_+$ verwendet.

[13] Gäbe es natürliche Zahlen p, q mit der Eigenschaft $(\frac{p}{q})^2 = 2$, so dürfte man zunächst voraussetzen, dass nicht p und q beide gerade sind. Dann muss p^2 und mithin auch p wegen $p^2 = 2q^2$ gerade, also q ungerade sein. Da p gerade ist, ist p^2 Vielfaches von 4, also ist auch q^2 und somit q gerade, was den gewünschten Widerspruch liefert.

[14] Ein konstruktiver Aufbau der reellen Zahlen, der auf den den rationalen Zahlen aufbaut, wird in einer anderen Abhandlung dargestellt.

2.0.1 Intervalle und Umgebungen

Für zwei reelle Zahlen a, b (mit $a \leq b$) versteht man unter dem *abgeschlossenen Intervall* von a bis b (in Zeichen $[a; b]$) die Menge aller reellen Zahlen, die nicht kleiner als a und gleichzeitig nicht größer als b sind: $[a; b] := \{x | a \leq x \leq b\}$.
Ähnlich definiert man *offene Intervalle*: $]a; b[:= \{x | a < x < b\}$ und *halboffene Intervalle*, nämlich so: $]a; b] := \{x | a < x \leq b\}$ bzw. $[a; b[:= \{x | a \leq x < b\}$.
Man bezeichnet eine Teilmenge T von $\mathbb{R}$ als *Umgebung* einer reellen Zahl a, wenn es ein offenes Intervall gibt, das a als Element enthält und Teilmenge von T ist; dann haben diese Umgebungen die folgenden Eigenschaften:
(0) Für jedes a aus $\mathbb{R}$ ist $\mathbb{R}$ eine Umgebung von a .

(1) Für jedes a aus $\mathbb{R}$ und jede Umgebung T von a ist a Element von T.

(2) Sind S und T zwei Umgebungen von a , dann ist auch die Schnittmenge von S und T eine Umgebung von a .

(3) Ist T eine Umgebung von a und S ist eine Obermenge von T, dann ist auch S eine Umgebung von a .

(4) Wenn T eine Umgebung von a ist, dann gibt es eine Umgebung S von a, die Teilmenge von T ist, und wobei T Umgebung eines jeden Elements von S ist.

(5) Sind a und b verschiedene reelle Zahlen, dann gibt es eine Umgebung T von a und eine Umgebung S von b, die disjunkt sind.

Bemerkung: Ist allgemein - für eine beliebige Menge M anstelle von $\mathbb{R}$ - für jedes a aus M ein System von Umgebungen genannten Teilmengen von M definiert, sodass die Eigenschaften (0)...(4) erfüllt sind, spricht man von einem *topologischen Raum*. Gilt in einem topologischen Raum auch (5), wird der Raum als *Hausdorffraum*[15] bezeichnet.
Die Elemente von R werden auch als Punkte des Raums R bezeichnet.
Für eine Teilmenge T von R und einen Punkt a aus R definiert man:

- a heißt *Berührpunkt* von T, wenn jede Umgebung von a mindestens ein Element von T enthält.
- a heißt *Häufungspunkt* von T, wenn jede Umgebung von a mindestens ein von a verschiedenes Element von T enthält.
- a heißt *Randpunkt* von T, wenn jede Umgebung von a mindestens ein Element von T und mindestens ein nicht in T liegendes Element von R enthält.
- a heißt *isolierter Punkt* von T, wenn es eine Umgebung von a gibt, deren einziges in T liegendes Element a ist.

Eine Teilmenge T eines topologischen Raums R heißt

- *offen*, wenn sie Umgebung eines jeden der in ihr enthaltenen Elemente ist.
- *abgeschlossen*, wenn sie alle ihre Häufungspunkte enthält.
- *diskret*, wenn sie ausschließlich aus isolierten Punkten besteht.

Hinweis: Die Begriffe abgeschlossen und offen stellen im Gegensatz zur Umgangssprache keine Gegensätze dar. Es gibt Mengen, die weder abgeschlossen noch offen sind, andererseits kann eine Menge sowohl offen wie auch abgeschlossen sein.

[15] nach dem Mathematiker *Felix Hausdorff*, gest. 1942 .

- Die Menge aller Berührpunkte von T heißt abgeschlossene Hülle von T (Notation: $\overline{T}$).
- Die Menge aller Häufungspunkte von T wird als Ableitung von T bezeichnet; Notation: T'.
- Die Menge aller Randpunkte von T heißt Rand von T mit der Notation ∂T
- Die Menge aller Punkte von T, die nicht in R liegen, heißt Komplement von T, notiert als $\tilde{T}$.
- Die Menge aller Punkte von T, für die T eine Umgebung ist, heißt Inneres von T oder offener Kern von T, notiert als T^0.

Bemerkung: Von der Hülle einer Menge spricht man dann, wenn folgende Eigenschaften vorliegen:

1) Jede Menge ist Teilmenge ihrer Hülle ($T \subset \overline{T}$)
2) Die Hüllenbildung ist idempotent, d.h. mehrfaches Bilden der Hülle führt nach der ersten Hüllenbildung nicht zu einer erweiterten Menge ($\overline{T} = \overline{\overline{T}}$)

Neben der oben definierten abgeschlossenen Hülle gibt es z.B. in der Geometrie die konvexe Hülle[16] und in der linearen Algebra die lineare Hülle[17].
Unmittelbar aus den Definitionen folgt für jede Teilmenge T eines topologischen Raums:

- T^0 ist Teilmenge von T,
- T ist Teilmenge von $\overline{T}$,
- T' ist Teilmenge von $\overline{T}$,
- ∂T ist Teilmenge von $\overline{T}$,
- $\partial(\tilde{T}) = \partial T$.

Die folgenden Gleichungen bedürfen eines (als Übung empfohlenen Nachweises):

- $T \cap T = \overline{T}$
- $T \cap \partial T = \overline{T}$
- $\widetilde{T^0} = \overline{\tilde{T}}$

Ebenso beweisbedürftig ist die (nur scheinbar selbstverständliche) Aussage: Der offene Kern von T ist offen, die abgeschlossene Hülle von T ist abgeschlossen.

2.0.2 Relationen

Unter dem kartesischen Produkt zweier Mengen A und B versteht man die Menge aller Paare (a, b), für die a Element von A und b Element von B ist. Das kartesische Produkt von A und B wird als $A \times B$ notiert:

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

Beispiele:

a) $\{p, q, r\} \times \{u, v\} = \{(p, u), (p, v), (q, u), (q, v), (r, u), (r, v)\}$

[16] Die *konvexe Hülle* einer Menge M von Punkten der Ebene ist die Gesamtheit aller Punkte, die auf einer der Verbindungsstrecken von zwei Punkten aus M liegen.

[17] Ist M eine Menge von Vektoren eines linearen Raums V, dann ist die *lineare Hülle* von M (in Zeichen oft als $< M >$ oder $\mathrm{span}(M)$ notiert) die Menge aller Linearkombinationen von Vektoren aus M.

b) $\{\} \times \{1, 2, 3, 4\} = \{\}$

c) $\mathbb{N} \times \{0, -7)\} = \{(0, 0), (0, -7), (1, 0), (1, -7), (2, 0), (2, -7), \ldots\}$

Eine Teilmenge T eines kartesischen Produkts $A \times B$ definiert eine *Relation*. Dabei stehen genau dann zwei Elemente a, b (mit $a \in A$ und $b \in B$) in dieser Relation, wenn (a, b) Element von T ist. Ist speziell $A = B$ spricht man auch von einer Relation *in* A.
Eine Relation T in A heißt

- *reflexiv*, wenn für jedes a aus A die Menge T das Element (a, a) enthält,
- *irreflexiv*, wenn die Menge T für kein a aus A das Element (a, a) enthält,
- *symmetrisch*, wenn T zugleich mit einem Paar (a, b) immer auch das Paar (b, a) enthält,
- *asymmetrisch*, wenn T nie gleichzeitig mit dem Paar (a, b) auch das Paar (b, a) enhält,
- *antisymmetrisch*, wenn aus (a, b) aus T und (b, a) aus T folgt, dass $a = b$ gilt,
- *transitiv*, wenn zusammen mit (a, b) und (b, c) stets (a, c) in T liegt.

Eine Relation, die antisymmetrisch und transitiv ist, wird als *Ordnungsrelation* bezeichnet, wenn sie (sogar) asymmetrisch ist, heißt die Ordnungsrelation *streng*; eine Relation, die reflexiv und transitiv ist, nennt man *Äquivalenzrelation*.
Beispiele für Relationen in der Menge der ganzen Zahlen $\mathbb{Z}$:

a) Mit $T = \{(a, b) \mid a < b\}$ definiert man eine (nämlich die vertraute) strenge Ordnungsrelation,

b) mit $T = \{(a, b) \mid (b - a)$ ist ohne Rest durch 7 teilbar$\}$ legt man eine Äquivalenzrelation fest.

Eine Äquivalenzrelation T in einer Menge A zerlegt diese Menge in sogenannte Äquivalenzklassen. Betrachtet man für jedes a aus A die Menge $\{x \mid (a, x) \in T\}$, dann sind je zwei dieser Mengen entweder elementfremd oder gleich.

3 Funktionen, Grenzwert, Stetigkeit

Eine Relation F heißt *funktional*, wenn aus $(a, b) \in F$ und $(a, c) \in F$ stets $b = c$ folgt. Diese Eigenschaft einer Relation wird auch als *Rechtseindeutigkeit* bezeichnet.
Sind A und B Mengen, und ist die funktionale Relation F eine Teilmenge von $A \times B$, dann wird durch F eine Funktion f (genauer : eine Klasse von Funktionen) festgelegt. Ist hierbei D die Menge aller Elemente $a \in A$, für welche ein $b \in B$ mit $(a, b) \in F$ existiert, dann heißt D *Definitionsbereich* (oder *Argumentmenge*) der Funktion. Da die Relation funktional ist, gibt es zu jedem $a \in D$ genau ein derartiges b. Man bezeichnet dieses b als Wert von f an der Stelle a und drückt diesen Sachverhalt durch die Schreibweise $f(a) = b$ aus. Die Menge W aller $b \in B$, für die ein $a \in A$ mit $f(a) = b$ existiert, wird als *Wertemenge* von f bezeichnet.
Beispiele:

(1) $F = \{(1, 2), (2, 3), (3, 4), (4, 5), (5, 5)\} : D = \{1, 2, 3, 4, 5\}, W = \{2, 3, 4, 5\}$

(2) $F = \{(1, 2), (2, 2), (3, 2), (4, 2), (5, 2)\} : D = \{1, 2, 3, 4, 5\}, W = \{2\}$

(3) $F = \{(-1, 1), (0, 0), (1, 1), (2, 4), (5, 25)\} : D = \{-1, 0, 1, 2, 5\}, W = \{0, 1, 4, 25\}$

(4) $F = \{(1, 1), (w, w), (pp, pp)\} : D = \{1, w, pp\}, W = \{1, w, pp\}$

(5) $F = \{(2,3),(3,5),(7,11),(17,19)\} : D = \{2,3,7,17\}, W = \{3,5,11,19\}$

Anstatt die funktionale Relation explizit vorzugeben, kann man in vielen Fällen den Definitionsbereich und die Vorschrift angeben, mit deren Hilfe zu einer Stelle der Argumentmenge der zugehörige Funktionswert ermittelt werden kann; in den obigen Beispielen so:

(1) $f(x) = x + 1$; (2) $f(x) = 2$; (3) $f(x) = x^2$; (4) $f(x) = x$;
(5) $f(x)$ ist in der geordneten Folge der Primzahlen der Nachfolger von x.

Notationsvereinbarung:
Nachfolgend werden Funktionen wie in den Beispielen (1)-(4) häufig durch einen Funktionsterm beschrieben, der nur eine echte Teilmenge der reellen Zahlen als Einsetzungen für die Variable zulässt; diese Menge wird gelegentlich als Einsetzungsbereich oder auch als Definitionsbereich $D(T)$ des Terms T bezeichnet, wie in folgenden Beispielen:

$$D\left(\frac{x+1}{x-2}\right) = \mathbb{R} \setminus \{2\}, \quad D\left(\sqrt{4-x^2}\right) = [-2;2], \quad D\left(\log(1-x^2)\right) =]-1;1[$$

Um hier eine Verwechslung mit dem Definitionsbereich einer Funktion zu vermeiden, wird nachfolgend beim Definitionsbereich von Funktionen der alternative Begriff *Argumentmenge* (dann mit der Mengenbezeichnung A) bevorzugt. Die Aussage *Die Funktion f hat die Argumentmenge A und als Werte reelle Zahlen* wird nachfolgend in der Form $f : A \to \mathbb{R}$ notiert.
Eine Funktion f wird als *konstant* bezeichnet, wenn die Wertemenge nur ein einziges Element enthält, wie im Beispiel (2). Nachfolgend werden (fast) nur Funktionen bestrachtet, deren Argumentmenge und Wertemenge beide Teilmengen der reellen Zahlen sind. Eine solche Funktion mit Argumentmenge A und Wertemenge W heißt

- nach oben beschränkt bzw. nach unten beschränkt bzw. beschränkt, wenn W diese Eigenschaft hat.,
- *injektiv*, wenn sie an verschiedenen Stellen stets auch verschiedene Werte annimmt, wenn also für $a, b \in A$ aus $f(a) = f(b)$ stets $a = b$ folgt.
- *isoton* (oder *monoton zunehmend*), wenn für alle $a, b \in A$ mit $a < b$ stets $f(a) \leq f(b)$ gilt,
- *streng isoton* (oder *streng monoton zunehmend*), wenn sie isoton und injektiv ist,
- *antition* (oder *monoton abnehmend*), wenn für alle $a, b \in A$ mit $a < b$ stets $f(a) \geq f(b)$ gilt,
- *streng antiton* (oder *streng monoton abnehmend*), wenn sie antiton und injektiv ist.
- *monoton*, wenn sie antiton oder isoton ist.

Unter den obigen Beispielen sind (1), und (2) isoton, (5) streng isoton, (4) und (5) injektiv.

3.1 Der Funktionenraum $\mathbb{R}^A$

Für eine Menge A reeller Zahlen wird nachfolgend mit $\mathbb{R}^A$ die Menge aller Funktionen mit der Argumenmenge A bezeichnet, deren Werte reelle Zahlen sind[18]:

$$\mathbb{R}^A := \{f \mid f : A \to \mathbb{R}\}$$

[18] Allgemeiner bezeichnet man für Mengen A, B mit B^A die Menge aller Abbildungen der Menge A in die Menge B. In diesem Sinne erklärt sich auch die Bezeichnung 2^M für die Potenzmenge einer Menge M. Diese Potenzmenge ist isomorph zur Menge aller Abbildungen von M in eine 2-elementige Menge.

Mit der Voraussetung

$$A \subset \mathbb{R}, \quad f : A \to \mathbb{R}, \quad g : A \to \mathbb{R}$$

definiert man die Funktion $f + g$ mit der Argumentmenge A durch

$$\bigwedge_{x \in A} (f + g)(x) := f(x) + g(x).$$

Mit den Rechenregeln in der kommutativen Gruppe $(\mathbb{R}, +)$ rechnet man leicht nach, dass mit der so definierten Addition $(\mathbb{R}^A, +)$ eine kommutative Gruppe ist; neutrales Element ist die Funktion, die an jeder Stelle $x \in A$ den Wert 0 hat.
Auf analoge Weise definiert man[19] das Produkt rf der reellen Zahl r mit der Funktion f, das Produkt $f \cdot g$ (oder einfach fg) von zwei Funktionen f und g, sowie für $n \in \mathbb{N}$ die n-te Potenz f^n einer Funktion f. Ebenso lässt sich für die Funktionen f und g die Quotientenfunktion $\frac{f}{g}$ erklären, sofern die Funktion g an keiner Stelle den Wert 0 annimmt.

3.1.1 Lineare Funktionen

Definition Eine Funktion f heißt *linear*[20], wenn es reelle Zahlen m und c mit der Eigenschaft gibt, dass sich für jedes x aus der Argumentmenge von f der Funktionswert $f(x)$ als $mx+c$ berechnen lässt.

An dieser Stelle ist der folgende Hinweis wichtig: In weiten Teilen der Mathematik werden die Begriffe Abbildung und Funktion als Synonyme gebraucht, wobei der Funktionsbegriff bevorzugt bei Abbildungen zwischen Mengen reeller Zahlen (oder - hier nicht behandelt - komplexer Zahlen) verwendet wird.
In der linearen Algebra heißt eine Abbildung f zwischen zwei Vektorräumen *linear*, wenn sie additiv (d.h. $f(x + y) = f(x) + f(y)$) und homogen (d.h. $f(rx) = r \cdot f(x)$) ist. In diesem Sinne wäre die einzige lineare Funktion die mit einem Funktionsterm der Form mx, also mit $c = 0$. Der hier in weiterem Sinne verwendete Begriff der linearen Funktion bezieht eine Begründung aus dem Graphen der linearen Funktion; bei der Darstellung in einem kartesischen Koordinatensystem liegen alle Punkte des Graphen nämlich auf einer Geraden (lateinisch: linea).

3.2 Grenzwert einer Funktion

Ergänzend zu dem Abschnitt über Umgebungen ist es zur späteren Vermeidung von Fallunterscheidungen zweckmäßig, den topologischen Raum $\mathbb{R}$ um die uneigentlichen Elemente ∞ und $-\infty$ zu erweitern. In die geordnete Menge der reellen Zahlen werden diese uneigentlichen Elemente eingebettet durch die für jede reelle Zahl x geltende Anordnung $-\infty < x < \infty$.
Die Menge $\overline{\mathbb{R}}$ (definiert als $\mathbb{R} \cup \{-\infty, \infty\}$) wird zu einem topologischen Raum, wenn man ergänzend zu den Umgebungen reeller Zahlen die Umgebungen von $-\infty$ bzw. ∞ wie folgt erklärt:
Eine Menge U reeller Zahlen ist eine Umgebung

- von $-\infty$, wenn es eine reelle Zahl a gibt mit $\{x \mid x < a\} \subset U$
- von ∞, wenn es eine reelle Zahl a gibt mit $\{x \mid x > a\} \subset U$

Die Intervallschreibweise wird für $a \in \mathbb{R}$ um die folgenden Notationen erweitert:

$$]-\infty; a[= \{x \mid x < a\}, \quad]a; \infty[= \{x \mid x > a\}, \quad]-\infty; \infty[= \mathbb{R}$$

[19]... für Funktionen aus $\mathbb{R}^A$...
[20]lat. linea : die (gerade) Linie

Definition Ist f eine Funktion mit Argumentmenge A und a ein Häufungspunkt von A, dann bezeichnet man die reelle Zahl (oder den uneigentlichen Wert) r als Grenzwert von f bei a, wenn es zu jeder Umgebung V von r eine Umgebung U von a gibt, für die folgendes gilt:
Für jedes Element $x \in A \setminus \{a\}$, das in U liegt, liegt der Wert $f(x)$ in V.

Gemäß dieser Definition ist es unwesentlich, ob a zur Argumentmenge gehört und - wenn ja - welchen Funktionswert f an der Stelle a hat. Da jede Umgebung einer reellen Zahl z als Teilmenge ein offenes Intervall um z enthält, kann man den definierenden Teil für reelle Grenzwerte an reellen Stellen folgendermaßen formulieren:

... wenn es zu jedem offenen Intervall V mit $r \in V$ ein offenes Intervall U mit $a \in U$ gibt, für die folgendes gilt: Für jedes Element $x \in (A \setminus \{a\})$, das in U liegt, liegt der Wert $f(x)$ in V.

Berücksichtigt man, dass jedes offene Intervall um eine reelle Zahl z als Teilmenge ein offenes Intervall mit Mittelpunkt z hat, kann man den definierenden Teil schließlich auf folgende Weise formulieren:

... wenn es zu jeder positiven Zahl ε eine positive Zahl δ gibt, sodass
für jedes Element $x \in]a-\delta, a+\delta[\setminus\{a\}$ der Funktionswert $f(x)$ in $]r-\varepsilon, r+\varepsilon[$ liegt.

Verwendet man die letzte Charakterisierung, so kommt man schließlich zu der häufig verwendeten Definition:

Definition: Ist f eine Funktion mit Definitionsbereich A und a ein Häufungspunkt von A, dann bezeichnet man die reelle Zahl r als Grenzwert von f bei a, wenn es zu jeder positiven Zahl ε eine positive Zahl δ gibt, sodass für jedes $x \in (A \setminus \{a\}$ mit $|x-a| < \delta$ gilt $|f(x) - r| < \varepsilon$.

Oder in der wesentlich kompakteren und daher übersichtlicheren mathematischen Kurznotation unter der Voraussetzung $A \subset \mathbb{R}, \quad f : A \to \mathbb{R}, \quad a \in A', \quad r \in \mathbb{R}$:

$$\lim_{x\to a} f(x) = r :\Leftrightarrow \bigwedge_{\varepsilon\in\mathbb{R}_+} \bigvee_{\delta\in\mathbb{R}_+} \bigwedge_{x\in A} (0 < |x-a| < \delta \Rightarrow |f(x) - r| < \varepsilon).$$

3.2.1 Grenzwertcharakterisierung durch Folgen

Unter den Voraussetzungen

$$A \subset \mathbb{R}, \quad f : A \to \mathbb{R}, \quad c \in A', \quad r \in \mathbb{R}$$

sind die folgenden Aussagen sind äquivalent:

(a) $\lim_{x\to a} f(x) = r$

(b) Jede unendliche Folge (c_n) mit Gliedern in $A\setminus\{c\}$, die gegen c konvergiert, hat eine gegen r konvergente Bildfolge.

Aus (a) folgt (b)

Gegeben sei eine in A gegen c konvergente Folge (c_n). Zum Nachweis von $\lim_{n\to\infty} f(c_n) = r$ sei ein positives ε vorgegeben. Gesucht wird eine Nummer n_0 mit der Eigenschaft, dass für alle Nummern n ab n_0 gilt $|f(c_n) - r| < \varepsilon$.

Da f bei c gegen r konvergiert, gibt es eine positive Zahl δ mit

$$(1) \qquad \bigwedge_{x\in A\setminus\{c\}} (0 < |x-c| < \delta \Rightarrow |f(x) - r| < \varepsilon).$$

Außerdem gibt es wegen der vorausgesetzten Konvergenz von (c_n) gegen c eine Nummer n_0, die folgendes leistet:

(2) $$\bigwedge_{n\in\mathbb{N}} (n \geq n_0 \Rightarrow |c_n - c| < \delta).$$

Aus (1) und (2) folgt, dass die Nummer n_0 aus (2) die gewünschte Eigenschaft hat.

Aus non (a) folgt non (b)

Da f als nicht konvergent gegen r bei c vorausgesetzt wird, existiert positive Zahl ε mit folgender Eigenschaft

(3) $$\bigwedge_{\delta\in\mathbb{R}_+} \bigvee_{x\in A} (|x - c| < \delta \wedge |f(x) - f(c)| \geq \varepsilon).$$

Wählt man also für jede positive ganze Zahl n den Wert $\frac{1}{n}$ als δ, so liefert (3) die Existenz einer Folge (c_n), bei der für jedes n gilt: $(|c_n - c| < \frac{1}{n} \wedge |f(c_n) - r| \geq \varepsilon$.

Da die Folge $(|c_n - c|)$ eine Nullfolge als Majorante hat, konvergiert (c_n) gegen c, ohne dass die Bildfolge den Grenzwert r hat; die Existenz einer solchen Folge war zu zeigen.

3.2.2 Grenzwertregeln

Bei den nachfolgenden Grenzwert-Regeln wird kein Beweis geführt, da die Beweise sich fast wörtlich durch Übertragung und Anpassung aus den entsprechenden an späterer Stelle gegebenen Stetigkeitsbeweisen ergeben. Unter den Voraussetzungen

$$A \subset \mathbb{R}, \quad a \in A', \quad \lambda, \mu, r \in \mathbb{R}, \quad n \in \mathbb{N}, \quad f, g : A \to \mathbb{R}, \quad \lim_{x\to a} f(x) = \lambda, \quad \lim_{x\to a} g(x) = \mu$$

lassen sich die folgenden Grenzwertsätze herleiten:

(1) $\lim_{x\to a} (rf)(x) = r\lambda$ (Faktorregel)

(2) $\lim_{x\to a} (f + g)(x) = \lambda + \mu$ (Summenregel)

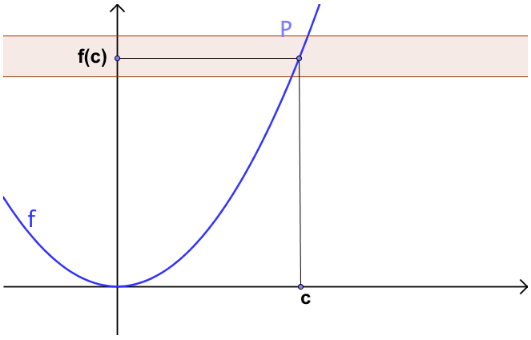
(3) $\lim_{x\to a} (f \cdot g)(x) = \lambda \cdot \mu$ (Produktregel)

(4) $\lim_{x\to a} (f^n)(x) = \lambda^n$ (Potenzregel)

(5) $\lim_{x\to a} (\frac{f}{g})(x) = \frac{\lambda}{\mu}$, falls $\mu \neq 0$ (Quotientenregel)

3.3 Stetigkeit einer Funktion

Bei der Anwendung des Begriffs der Stetigkeit (oder Kontinuität) auf Funktionen mag die Vorstellung der zeichnerischen Darstellung eine wichtige Rolle gespielt haben: Eine Funktion, deren Graphen man ohne Absetzen des Stiftes zeichnen kann, entspricht dieser Vorstellung. Anschaulich gesprochen: Legt man einen beliebig schmalen Horizontalstreifen um einen Punkt P mit den Koordinaten $(c|f(c))$ des Graphen einer Funktion f, so sollen alle Punkte $(x|f(x))$ in diesem Horizontalstreifen liegen, wenn x nahe genug bei c liegt.



Die Funktion f wird also als stetig an der Stelle c bezeichnet, wenn es zu jedem Horizontalstreifen um die Gerade mit der Gleichung $y = f(c)$ einen Vertikalstreifen um die Gerade mit der Gleichung $x = c$ mit folgender Eigenschaft gibt: Der Teil des Graphen von f der ganz im Vertikalstreifen liegt, wird auch vollständig vom Horizontalstreifen überdeckt.
Wenn wir die Streifen so legen, dass die angegebenen Geraden jeweils die Mittelparallelen im Streifen sind, und wenn wir die halbe Breite des Horizontalstreifens mit ε und die halbe Breite des Vertikalstreifens mit δ bezeichnen, so ergibt sich die folgende Charakterisierung dafür, dass f an der Stelle c stetig ist:
Zu jeder positiven reellen Zahl ε gibt es eine positive reelle Zahl δ mit folgender Eigenschaft: Für jedes x aus der Argumentmenge A von f, das von c um weniger als δ entfernt ist, unterscheidet sich $f(x)$ von $f(c)$ um weniger als ε. Oder kürzer in logischer Notation

$$\bigwedge_{\varepsilon\in\mathbb{R}_+} \bigvee_{\delta\in\mathbb{R}_+} \bigwedge_{x\in A} (|x-c| < \delta \Longrightarrow |f(x) - f(c)| < \varepsilon)$$

Die Definition lässt unmittelbar einige Folgerungen zu

- Ist c isolierte Stelle der Argumentmenge von f, ist f bei c stetig

  Denn wenn c isoliert in A ist, gibt es eine positive Zahl δ, für die das offene Intervall $]c-\delta, c+\delta[$ außer c kein Element von A enthält. Mit diesem δ wird die Definition der Stetigkeit bei c für beliebiges ε erfüllt.

- Ist c Häufungspunkt von A, ist f genau dann an der Stelle $c \in A$ stetig, wenn der Grenzwert von f an der Stelle c existiert und mit $f(c)$ übereinstimmt.

- Ist f an der Stelle c stetig und ist $f(c)$ positiv, dann gibt es ein offenes Intervall um c, über dem f nur größere Werte als $\frac{f(c)}{2}$ annimmt, also insbesondere positiv ist.

  Zum Nachweis braucht man nur die Stetigkeitsdefinition auf $\varepsilon = \frac{f(c)}{2}$ anzuwenden. Der diesen Sachverhalt angebende Satz wird gelegentlich als *Positivitätssatz für stetige Funktionen* bezeichnet.

3.3.1 Stetigkeitscharakterisierung durch Folgen

Bei den ersten der nachfolgenden Stetigkeits-Regeln wird der Beweis unter Rückgriff auf die Definition geführt. Alternativ ergeben sich die Regeln unmittelbar aus den entsprechenden Konvergenzsätzen für reelle Zahlenfolgen, wenn man - wie bei den weiter unten angegebenen Regeln - die Äquivalenz folgender Aussagen benutzt:

(a) Die auf der Menge A definierte Funktion f ist an der Stelle c stetig.

(b) Jede unendliche Folge (c_n) mit Gliedern in A, die gegen c konvergiert, hat eine gegen $f(c)$ konvergente Bildfolge.

Aus (a) folgt (b)

Gegeben sei eine in A gegen c konvergente Folge (c_n). Zum Nachweis von $\lim_{n\to\infty} f(c_n) = f(c)$ sei ein positives ε vorgegeben. Gesucht wird eine Nummer n_0 mit der Eigenschaft, dass für alle Nummern n ab n_0 gilt $|f(c_n) - f(c)| < \varepsilon$.

Da f bei c stetig ist, gibt es eine positive Zahl δ mit

$$(4) \qquad \bigwedge_{x\in A} (|x - c| < \delta \Rightarrow |f(x) - f(c)| < \varepsilon).$$

Außerdem gibt es wegen der vorausgesetzten Konvergenz eine Nummer n_0, die folgendes leistet:

$$(5) \qquad \bigwedge_{n\in N} (n \geq n_0 \Rightarrow |c_n - c| < \delta).$$

Aus (1) und (2) folgt, dass die Nummer n_0 aus (2) die gewünschte Eigenschaft hat.

Aus non (a) folgt non (b)

Da f als unstetig bei c vorausgesetzt wird, existiert positive Zahl ε mit folgender Eigenschaft

$$(6) \qquad \bigwedge_{\delta\in\mathbb{R}_+} \bigvee_{x\in A} (|x - c| < \delta \wedge |f(x) - f(c)| \geq \varepsilon).$$

Wählt man also für jede positive ganze Zahl n den Wert $\frac{1}{n}$ als δ, so liefert (3) die Existenz einer Folge (c_n), bei der für jedes n gilt: $(|c_n - c| < \frac{1}{n} \wedge |f(c_n) - f(c)| \geq \varepsilon$.

Da die Folge $(|c_n - c|)$ eine Nullfolge als Majorante hat , konvergiert (c_n) gegen c, ohne dass die Bildfolge den Grenzwert $f(c)$ hat; die Existenz einer solchen Folge war zu zeigen.

3.4 Regeln zur Stetigkeit an einer Stelle c

Nachfolgend seien f und g stets Funktionen mit Argumentmenge A, stetig an einer Stelle $c \in A$.

3.4.1 Faktorregel

Es sei r eine reelle Zahl; dann ist die Funktion rf (definiert durch $\bigwedge_{x\in A}(rf)(x) := r\cdot f(x))$ bei c stetig. Zu beweisen ist nur etwas für den Fall $r \neq 0$, da die Funktion rf sonst konstant (mit Wert 0) und daher stetig ist.
Es sei $\varepsilon > 0$ vorgegeben. Gemäß der Stetigkeitsdefinition sei δ zu f für $\frac{\varepsilon}{|r|}$ passend. Dann gilt für alle Stellen x aus $A\cap]c - \delta, c + \delta[$

$$|(rf)(x) - (rf)(c)| = |r \cdot (f(x) - f(c)| = |r| \cdot |f(x) - f(c)| < |r| \cdot \frac{\varepsilon}{|r|} = \varepsilon.$$

3.4.2 Summenregel

Die Summenfunktion $f + g$ (definiert durch $\bigwedge_{x\in A}(f + g)(x) := f(x) + g(x))$ ist bei c stetig.
Zum Beweis sei $\varepsilon > 0$ vorgegeben. Gemäß der Stetigkeitsdefinition sei δ_f zu f für $\frac{\varepsilon}{2}$ passend; entsprechend sei δ_g zu g für $\frac{\varepsilon}{2}$ passend. Mit $\delta := min(\delta_f, \delta_g)$ gilt für alle Stellen x aus $A\cap]c - \delta, c + \delta[$

$$|f(x) - f(c)| < \frac{\varepsilon}{2} \quad \wedge \quad |g(x) - g(c)| < \frac{\varepsilon}{2},$$

also

$$\begin{aligned} |(f + g)(x) - (f + g)(c)| = |f(x) + g(x) - f(c) - g(c)| &= |f(x) - f(c) + g(x) - g(c)| \\ &\leq |f(x) - f(c)| + |g(x) - g(c)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

3.4.3 Produktregel

Die Produktfunktion fg (definiert durch $\bigwedge_{x\in A}(fg)(x) := f(x)g(x))$ ist bei c stetig.
Zum Beweis sei $\varepsilon > 0$ vorgegeben. Man setze $s := \max(\varepsilon + |g(c)|, |f(c)|)$. Gemäß der Stetigkeitsdefinition sei δ_f zu f für $\frac{\varepsilon}{2s}$ passend; entsprechend sei δ_g zu g für $\frac{\varepsilon}{2s}$ passend. Mit $\delta := min(\delta_f, \delta_g)$ gilt für alle Stellen x aus $A\cap]c-\delta, c+\delta[$

$$|f(x) - f(c)| < \frac{\varepsilon}{2s} \quad \wedge \quad |g(x) - g(c)| < \frac{\varepsilon}{2s},$$

also

$$\begin{aligned}
|(fg)(x) - (fg)(c)| &= |f(x)g(x) - f(c)g(x) + f(c)g(x) - f(c)g(c)| \\
&= |(f(x) - f(c)) \cdot g(x) + f(c) \cdot (g(x) - g(c))| \\
&\leq |f(x) - f(c)| \cdot |g(x)| + |f(c)| \cdot |g(x) - g(c)| \\
= |f(x) - f(c)| \cdot |g(x) - g(c) + g(c)| &+ |f(c)| \cdot |g(x) - g(c)| \\
\leq |f(x) - f(c)| \cdot s + s \cdot |g(x) - g(c)| &< \frac{\varepsilon}{2s} \cdot s + s \cdot \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

3.4.4 Potenzregel

Aus der Produktregel folgt durch vollständige Induktion die Potenzregel:
Ist f eine an einer Stelle c stetige Funktion und n eine positive ganze Zahl, dann ist auch die Funktion f^n (mit $f^n(x) := (f(x))^n$) bei c stetig.

3.4.5 Kehrwertregel

Unter der zusätlichen Voraussetzung, dass f keine Nullstelle hat, ist die Kehrwertfunktion $\frac{1}{f}$ (definiert durch $\bigwedge_{x\in A} \frac{1}{f}(x) := \frac{1}{f(x)}$) stetig an der Stelle c.
Zum Beweis wird vorausgesetzt, dass $f(c)$ positiv ist (andernfalls verläuft der Beweis analog). Vorgegeben sei ein positives ε.
Nach der Stetigkeitsdefinition gibt es ein positives δ_1, so dass für alle Stellen x der Argumentmenge A mit $|x - c| < \delta_1$ gilt $|f(x) - f(c)| < \frac{f^2(c)}{2} \cdot \varepsilon$.
Und nach dem oben angegebenen Positivitätssatz gibt es eine positive Zahl δ_2, für die $f(x) > \frac{f(c)}{2}$ gilt, wenn die Argumentstelle x im Intervall $]c - \delta_2, c + \delta_2[$ liegt.
Mit $\delta := \min(\delta_1, \delta_2)$ gilt daher für alle $x \in A\cap]c - \delta, c + \delta[$:

$$|\frac{1}{f}(x) - \frac{1}{f}(c)| = |\frac{1}{f(x)} - \frac{1}{f(c)} = |\frac{f(c) - f(x)}{f(x) \cdot f(c)}| \leq |\frac{f(c) - f(x)}{\frac{f(c)}{2} \cdot f(c)}| < |\frac{\frac{f^2(c)}{2} \cdot \varepsilon}{\frac{f(c)}{2} \cdot f(c)}| = \varepsilon$$

3.4.6 Quotientenregel

Wenn die Funktion g keine Nullstelle hat, ist die Funktion $\frac{f}{g}$ (definiert durch $\bigwedge_{x\in A} \frac{f}{g}(x) := \frac{f(x)}{g(x)}$ an der Stellei c stetig.
Wegen $\frac{f}{g} = f \cdot \frac{1}{g}$ folgt das unmittelbar aus Kehrwert- und Produktregel.

3.4.7 Betragsregel

Die Funktion abs f, definiert durch $\bigwedge_{x\in A}$ (abs $f)(x) := |f(x)|$ ist bei c stetig. Das folgt, da nach der Differenzenungleichung

$$|(\text{abs}f)(x) - (\text{abs}f)(c)| = |\ |f(x)| - |g(x)|\ | \leq |f(x) - g(x)|$$

gilt, ohne Konstruktion eines neuen δ zum vorgegebenen ε aus der Stetigkeitsdefinition.

3.4.8 Verkettungsregel

Unter der Voraussetzung

$$\begin{aligned} A, B \subset \mathbb{R}; c \in A \\ f : A \to \mathbb{R}, \text{stetig an der Stelle } c; f(A) \subset B \\ g : B \to \mathbb{R}, \text{stetig an der Stelle } f(c) \end{aligned}$$

ist die durch Verkettung entstehende Funktion $g \circ f$ stetig an der Stelle $f(c)$.
Denn wenn (a_n) eine gegen c konvergierende Folge in A ist, dann konvergiert $(f(a_n))$ gegen $f(c)$ und somit $(g(f(a_n))$ gegen $g(f(c))$.

3.4.9 Einschränkungsregel

Unter der Voraussetzung

$$c \in B \subset A \subset \mathbb{R}; f : A \to \mathbb{R}, \text{stetig an der Stelle } c; g : B \to \mathbb{R}; \bigwedge_{x \in B} g(x) = f(x)$$

ist auch die Funktion g an der Stelle c stetig.
g wird auch als *Einschränkung von f auf B* bezeichnet, in Zeichen : $g = f|_B$.
Zum Beweis ist lediglich zu beachten, dass jede gegen c konvergente Folge mit Gliedern in B auch eine gegen c konvergente Folge mit Gliedern in A ist.

3.4.10 Ausdehnungsregel

Unter der Voraussetzung

$$\begin{aligned} A, B \subset \mathbb{R}; c \in A \\ f : A \to \mathbb{R}, \text{stetig an der Stelle } c; \\ g : B \to \mathbb{R}; \bigwedge_{x \in B \cap A} g(x) = f(x) \end{aligned}$$

ist die auf $A \cup B$ definierte Funktion h mit

$$\bigwedge_{x \in A} h(x) = f(x) \quad \wedge \quad \bigwedge_{x \in B \setminus A} h(x) = g(x)$$

genau dann stetig bei c, wenn c entweder kein Häufungspunkt von B ist oder g an der Stelle c den Grenzwert $f(c)$ hat.
Denn wenn c kein Häufungspunkt von B ist, dann liegen von einer in $A \cup B$ gegen c konvergierenden Folge von einer Nummer an alle Folgenglieder in A; die Bildfolge konvergiert also gegen $f(c)$.

4 Stetige Funktionen auf abgeschlossenen Intervallen

4.1 Definition der stetigen Funktion

Eine auf einer Argumentmenge $A \subseteq \mathbb{R}$ definierte Funktion f heißt stetig, wenn sie an jeder Stelle $x \in A$ stetig ist.
Von den in diesem Abschnitt betrachteten Funktionen wird jeweils außer der Stetigkeit vorausgesetzt, dass ihre Argumentmenge A ein abgeschlossenes Intervall $[a; b]$ ist.

4.2 Der Zwischenwertsatz für stetige Funktionen

Wesentlich wird nachfolgend von den Eigenschaften der reellen Zahlen mehrfach die folgende verwendet: Jede nach oben beschränkte nicht-leere Menge reeller Zahlen hat ein Supremum (also eine kleinste obere Schranke).

4.2.1 Nullstellensatz

Ist f eine auf $[a;b]$ definierte stetige Funktion mit $f(a) < 0 < f(b)$, dann gibt es ein $\xi \in]a;b[$ mit $f(\xi) = 0$.
Zum Beweis sei $M = \{x \in [a;b] | \bigwedge_{x \in [a;x]} f(x) < 0\}$. Die Menge M ist nach oben durch b beschränkt und wegen $f(a) < 0$ nicht leer, hat also ein Supremum ξ. Da M Teilmenge des abgeschlossenen Interevalls $[a;b]$ ist, liegt auch ξ in $[a;b]$ und f ist dort stetig.
Wäre $f(\xi) < 0$, gäbe es ein Intervall $]\xi - \delta; \xi + \delta[$, über dem $f(x)$ negativ wäre, also wäre ξ nicht die kleinste obere Schranke von M. Entsprechend würde aus $f(\xi) > 0$ folgen, dass $f(x)$ in einem ganzen Intervall um ξ positiv ist, es gäbe also Elemente $x < \xi$, die nicht in M liegen. Somit folgt, wie zu beweisen war, $f(\xi) = 0$.

4.2.2 Zwischenwertsatz

Gilt für die auf einer das abgeschlossene Intervall $[a;b]$ umfassenden Argumentmenge definierte Funktion f für die reelle Zahl λ die Ungleichung $f(a) \leq \lambda \leq f(b)$, dann gibt es im Intervall $[a;b]$ eine Stelle ξ mit $f(\xi) = \lambda$.
Ein Beweis muss nur für den Fall $f(a) < \lambda < f(b)$ geführt werden. Dazu definiert man auf A die Funktion h durch die Vorschrift $\bigwedge_{x \in A} h(x) = f(x) - c$. Die Funktion h ist stetig und erfüllt die Voraussetzungen des Nullstellensatzes, hat also eine Nullstelle $\xi \in [a;b]$.

4.2.3 Fixpunktsatz

Wenn die stetige Funktion f ein abgeschlossenes Intervall $[a;b]$ in sich abbildet, dann hat sie mindestens einen Fixpunkt ξ, es gibt also mindestens eine Stelle $\xi \in [a;b]$ mit $f(\xi) = \xi$.
Zum Beweis betrachte man die auf $[a;b]$ durch $\bigwedge_{x \in [a;b]} h(x) = f(x) - x$ definierte Hilfsfunktion h. Diese ist stetig und erfüllt wegen $h(a) = f(a) - a < 0$ und $h(b) = f(b) - b > 0$ die Voraussetzungen des Nullstellensatzes, hat also eine Nullstelle ξ. Aus $h(\xi) = 0$ folgt $f(\xi) = \lambda$, was zu zeigen war.

4.2.4 Erläuterung zur gleichmäßigen Stetigkeit

Gibt man bei einer stetigen Funktion eine positive Zahl ε vor, so gibt es für jede Stelle c der Argumentmenge A nach der Stetigkeitsdefinition eine positive Zahl δ, so dass für alle Argementstellen x aus $]c - \delta; c + \delta[$ gilt $|f(x) - f(c) < \varepsilon$. Allerdings ist die Größe eines solchen geeigneten δ nicht unabhängig von der jeweiligen Argumentstelle c.
Es ist eine besondere Eigenschaft der auf abgeschlossenen Intervallen definierten stetigen Funktionen, dass für sie zu vorgegebenem ε immer ein Wert von δ existiert, die im Sinne der Stetigkeitsdefinition zu allen Stellen der Argumentmenge passt. Diese Eigenschaft einer stetigen Funktion wird als *gleichmäßige Stetigkeit* bezeichnet. Es gilt also der folgende Satz:

4.2.5 Satz zur gleichmäßigen Stetigkeit

Hat die stetige Funktion f als Argumentmenge ein abgeschlossenes Intervall $[a;b]$, dann ist sie gleichmäßig stetig; es gilt also

(7) $$\bigwedge_{\varepsilon \in \mathbb{R}_+} \bigvee_{\delta \in \mathbb{R}_+} \bigwedge_{c,x \in A} (|x - c| < \delta \Rightarrow |f(x) - f(c)| < \varepsilon)$$

Zum Beweis wird die Negation der Aussage (4) zum Widerspruch geführt. Es wird also angenommen, es gäbe ein positives ε, bei dem man für jedes positive δ Argumentstellen c, x finden kann, die sich zwar um weniger als δ unterscheiden, deren Funktionswerte aber eine Differenz von mindestens ε haben. Wählt man nun speziell für δ einen Stammbruch, so hat man die Annahme

$$(8) \qquad \bigvee_{\varepsilon \in R_+} \bigwedge_{n \in N^*} \bigvee_{c_n, x_n \in A} (|x_n - c_n| < \frac{1}{n} \wedge |f(x_n) - f(c_n)| \geq \varepsilon)$$

Jede beschränkte Folge reeller Zahlen hat eine konvergente Teilfolge, eine solche Teilfolge von (c_n) sei mit (d_n) bezeichnet; die indexgleiche Teilfolge von (x_n) sei (y_n).
Da alle Folgenglieder in der Argumentmenge A, also einem abgeschlossenen Intervall liegen, ist auch der Grenzwert $\lim_{n \to \infty} d_n (=: d)$ Element von A. An der Stelle d kann aber keine Stetigkeit vorliegen, da (y_n) und (d_n) beide gegen d konvergieren, aber $(f(y_n) - f(d_n))$ keine Nullfolge ist.

4.3 Satz vom Maximum und Minimum

4.3.1 Beschränktheit stetiger Funktionen

Jede auf einem abgeschlossenen Intervall $[a; b]$ stetige Funktion ist dort beschränkt, d.h. es gibt ein abgeschlossenes Intervall, das alle Werte von f enthält. Oder anschaulich gesprochen: Der Graph von f wird von einem geeigneten Rechteck vollständig überdeckt.
Es genügt, die Beschränktheit nach oben zu zeigen.
Zum Beweis definiere man die Menge M so: $M = \{x \in [a; b] \mid f|_{[a;x]} \text{ nach oben beschränkt}\}$. Die nach oben durch b beschränkte Menge M enthält a, ist also nicht leer, und hat somit ein Supremum $s \in [a; b]$. Da f bei s stetig ist, gibt es nach der Stetigkeitsdefinition (hier angewendet für $\varepsilon = 1$) eine positive Zahl δ (o.B.d.A. $\delta < b - a$) mit der Eigenschaft

$$\bigwedge_{x \in [a;b] \cap]s - \delta; s + \delta[} f(x) < f(s) + 1$$

Da f somit über dem Intervall $[a; b] \cap [s - \frac{\delta}{2}; s + \frac{\delta}{2}]$ beschränkt ist, kann s nicht kleiner als b sein; also ist $s = b$. Und da f über $]s - \frac{\delta}{2}; s]$ und (nach Definition von s) über $[a; s - \frac{\delta}{2}]$beschränkt ist, folgt die Beschränktheit von f für ganz $[a; b]$.

4.3.2 Satz vom Maximum

Oben wurde gezeigt, dass jede auf einem abgeschlossenen Intervall stetige Funktion f nach oben beschränkt ist; der nachfolgende Beweis zeigt, dass das Supremum s der Funktion in der Wertemenge $f([a; b])$ liegt, also eine Stelle $c \in [a; b]$ mit $f(c) = s$ existiert. In jeder beschränkten, nicht leeren Menge reeller Zahlen gibt es eine Folge (y_n), die gegen das Supremum konvergiert. Es sei also (y_n) eine gegen s konvergente Folge, (a_n) eine passende Urbildfolge in $[a; b]$, also eine Folge, in der für jeden Index n gilt $f(a_n) = y_n$. Da jede beschränkte reelle Folge eine konvergente Teilfolge hat, gibt es zu (a_n) eine konvergente Teilfolge (c_n); ihr Grenzwert sei c. Da $[a; b]$ abgeschlossen ist, liegt der Grenzwert c in $[a; b]$. Dann ist $(f(c_n))$ als Teilfolge von (y_n) ebenfalls gegen s konvergiert, andererseits aufgrund der Stetigkeit von f konvergent gegen $f(c)$. Aufgrund der Eindeutigkeit des Grenzwerts ist $f(c) = s$; also ist $f(c)$ Maximum der Bildmenge.
Analog ergibt sich die Existenz des Minimums von f.

4.4 Zusammenfassung von Zwischenwertsatz und Maximumssatz

Die beiden in der Überschritt genannten Sätze lassen sich folgendermaßen zusammenfassen:

Die Bildmenge einer auf einem abgeschlossenen Intervall definierten
und dort stetigen Funktion ist ein abgeschlossenes Intervall.

4.5 Integrierbarkeit stetiger Funktionen

Als Folgerung aus den vorhergehenden Abschnitten ergibt sich:
Jede auf einem abgeschlossenen Intervall $[a;b]$ definierte Funktion ist dort (Riemann-) integrierbar.
Denn zu vorgegebenem positiven ε gibt es aufgrund der gleichmäßigen Stetigkeit eine positive ganze Zahl n mit folgender Eigenschaft:

$$\bigwedge_{x\in[a;b]} \bigwedge_{y\in[a;b]} (|x-y| < \frac{b-a}{n} \Rightarrow |f(x)-f(y)| < \frac{\varepsilon}{b-a})$$

Bei der Zerlegung $Z = (a = x_0, x_1, x_2, \ldots, x_n = b)$ des Intervalls $[a,b]$ in n Teilintervalle gleicher Länge gilt dann für die zugehörigen Ober- und Untersummen O_Z bzw. U_Z:

$$\begin{aligned} O_Z - U_Z \quad &= \quad \sum_{i=1}^{n} \max(f|_{[x_{i-1};x_i]}) \cdot \frac{b-a}{n} - \sum_{i=1}^{n} \min(f|_{[x_{i-1};x_i]}) \cdot \frac{b-a}{n} \\ &= \frac{b-a}{n} \cdot \sum_{i=1}^{n} (\max(f|_{[x_{i-1};x_i]}) - \min(f|_{[x_{i-1};x_i]}) < \frac{b-a}{n} \cdot \sum_{i=1}^{n} \frac{\varepsilon}{b-a} = \varepsilon \end{aligned}$$

Nach dem Riemannschen Kriterium für Integrierbarkeit ist f mithin integrierbar.

4.6 Stetigkeit der Umkehrfunktion

Eine auf einem abgeschlossenen Intervall definierte stetige Funktion ist genau dann umkehrbar, wenn sie streng monoton ist, denn

1. Eine streng monotone Funktion ist stets umkehrbar.

2. Wenn umgekehrt eine auf einem Intervall $[a;b]$ definierte stetige Funktion nicht streng monoton ist, ist sie nicht injektiv. Denn wenn sie zum Beispiel nicht streng isoton ist, gibt es im Definitionsintervall $[a;b]$ Stellen x_1, x_2, x_3 mit

   $$x_1 < x_2 < x_3 \quad \wedge \quad f(x_1) < f(x_2) \quad \wedge \quad f(x_2) > f(x_3).$$

   Dann werden aber alle Werte aus dem offenen Intervall $]\max(f(x_1), f(x_3)); f(x_2)[$ nach dem Zwischenwertsatz sowohl über $]x_1; x_2[$ als auch über $]x_2; x_3[$ angenommen.

Die Umkehrfuntion einer über einem abgeschlossenen Intervall definierten injektiven stetigen Funktion ist wieder stetig. Denn allgemeiner gilt:
Die Umkehrfunktion zu einer streng monotonen Funktion ist immer stetig.

5 Beispiele stetiger Funktionen

In diesem Abschnitt wird zur Abkürzung die Argumentmenge einer Funktion f mit A_f bezeichnet.

5.1 Rationale Funktionen

(1) Jede konstante Funktion ($k \in \mathbb{R}; \bigwedge_{x\in A_f} f(x) = k$) sowie die identische Funktion ($\bigwedge_{x\in A_f} f(x) = x$) sind stetig, denn $\lim_{x\to c} k = k$ und $\lim_{x\to c} x = c$

(2) Jede ganzrationale Funktion, also jede Funktion, deren Funktionsterm sich mit geeignetem $n \in \mathbb{N}$ und geeigneten $a_1, a_2, a_3, \ldots, a_n \in \mathbb{R}$ in der Form $\sum_{i=0}^{n} a_i x^i$ darstellen lässt, ist stetig, wie mit Hilfe von Produkt- und Summenregel aus (1) folgt.

(3) Jede gebrochen-rationale Funktion (Quotient zweier ganz-rationaler Funktionen) ist stetig, wie aus der Quotientenregel für Stetigkeit folgt.

Hinweis: Bei der Bildung des Quotienten von Funktionen fallen die Nullstellen der Nennerfunktion aus der Argumentmenge heraus. Dies führt aber nicht zur Unstetigkeit an diesen Stellen (gelegentlich wurde früher in der Schule von hebbaren bzw. nicht hebbaren Unstetigkeitsstellen gesprochen), vielmehr ist die Quotientenfunktion überall stetig, da die Frage nach der Stetigkeit nur Stellen aus der Argumentmenge betrifft. Der mathematische Stetigkeitsbegriff stimmt nicht in allen Fällen mit der anschaulichen Vorstellung einer ohne Absetzen des Stifts zu zeichnenden Graphen überein.

5.2 Wurzelfunktion

Die auf der Menge der nichtnegativen reellen Zahlen oder einer Teilmenge A definierte Wurzelfunktion ($\bigwedge_{x\in A} f(x) = \sqrt{x}$) ist stetig, wie unmittelbar aus dem Wurzelsatz für Grenzwerte folgt.

5.3 Die elementarene Funktionen

Die elementaren Funktionen sin, cos, tan, cot, sinh, cosh, exp und log sind stetig. Der Nachweis wird hier nicht gegeben.

5.4 Stückweise definierte Funktionen

Funktionen, deren Graphenverlauf scheinbar oder wirklich nicht durch einen geschlossenen Ausdruck anzugeben ist, können stückweise definiert werden. Dabei werden zunächst Funktionen $f_1, f_2, \dots, f_n$ über ihren jeweiligen Argumentmengen $A_1, A_2, \dots, A_n$ definiert und dann zu einer Funktion f mit $A_f = \bigcup_{i=1}^{n} A_i$ zusammengesetzt.
Vorausgesetzt wird dabei die Kompatibilität der Funktionswerte:

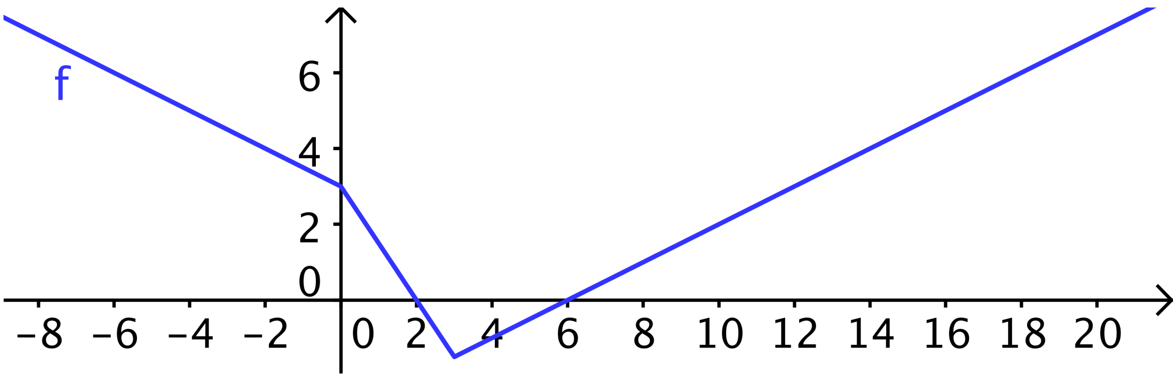
$$\bigwedge_{i,j\in\{1,2,3,\dots,n\}} \bigwedge_{x\in A_i\cap A_j} f_i(x) = f_j(x)$$

Da an den in der Argumentmenge liegenden Häufungspunkten der Argumentmenge das Vorliegen von Stetigkeit äquivalent dazu ist, dass an der betreffenden Stelle ein Grenzwert der Funktion existiert und mit dem Funktionswert dort übereinstimmt, ist für die Vererbung der Stetigkeit von den Teilfunktionen f_i auf f die folgende Bedingung notwendig und hinreichend

$$\bigwedge_{i,j\in\{1,2,3,\dots,n\}} \bigwedge_{c\in A_i'} (c \in A_j \Rightarrow \lim_{x\to c} f_i(x) = f_j(c))$$

5.4.1 Beispiele

(1) Intervallweises Zusammensetzen ganzrationaler Funktionen



Im dargestellten Beispiel sind drei verschiedene, über $]-\infty;-2]$ bzw. $[-2;3]$ bzw. $[3;\infty[$ definierte lineare Funktionen zu einer Funktion f zusammengesetzt:

Unter Verwendung der Betragsfunktion lässt sich der Funktionsterm geschlossen als $|x-3|-|\frac{x}{2}+1|$ darstellen.

(2) Schließen von Definitionslücken

a) ... bei gebrochen-rationalen Funktionen

Die als Quotient ganzrationalen Funktionen entstehende Funktion g mit $g(x) = \frac{x^2-3x+2}{x-1}$ hat aufgrund der Nullstelle der Nennerfunktion die Argumentmenge $A = \mathbb{R} \setminus \{1\}$.

Da für $x \neq 1$ aber die Identität $\frac{x^2-3x+2}{x-1} = x-2$ gilt, kann g an der Stelle 1 stetig ergänzt werden.

Die erweiterte Funktion f ist nun auf der Menge aller reellen Zahlen definiert und ist durch $f(x) = \frac{x^2-3x+2}{x-1}$ für $x \neq 1, f(x) = -1$ oder kürzer durch $\bigwedge_{x\in\mathbb{R}} f(x) = x-2$ anzugeben.

b) ... bei Beispielen mit der Sinusfunktion

i. $f(x) = \frac{\sin(x)}{x}$

Die durch Quotientenbildung entstehende Funktion ist zunächst an der Stelle 0 nicht definiert. Wie man bei geometrischer Definition der Sinusfunktion mit Hilfe einer Flächeninhaltsabschätzung zeigt, hat die Funktion f an der Stelle 0 den Grenzwert 1. Mit der Definition $f(0) := 1$ dehnt man also f auf ganz $\mathbb{R}$ stetig aus.

ii. $f(x) = x \cdot \sin(\frac{1}{x})$

Die Funktion f ist auf $\mathbb{R}^*$ definiert. Die Lücke bei 0 lässt sich aber stetig beheben, da die Sinusfunktion beschränkt ist und das Produkt einer an einer Stelle c gegen 0 konvergenten Funktion mit einer in einer Umgebung von c beschränkten Funktion bei c den Grenzwert 0 hat. Mit der Definition $f(0) := 0$ dehnt man also f auf ganz $\mathbb{R}$ stetig aus.

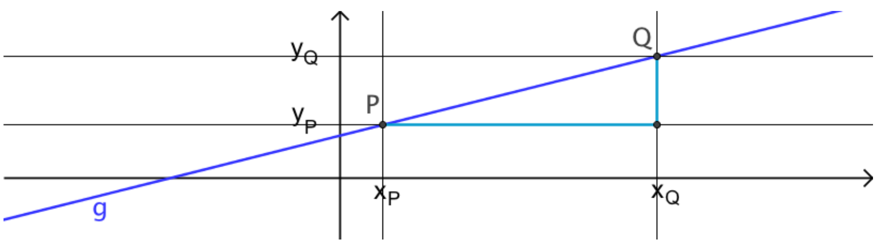
iii. Gegenbeispiel: $f(x) = \sin(\frac{1}{x})$

Da die Funktion f in jeder Umgebung von 0 alle Werte des Intervalls $[-1;1]$ annimmt, hat sie keinen Grenzwert an der Stelle 0. Eine stetige Ausdehnung auf die Menge aller reellen Zahlen ist also hier nicht möglich.

6 Differenzierbarkeit

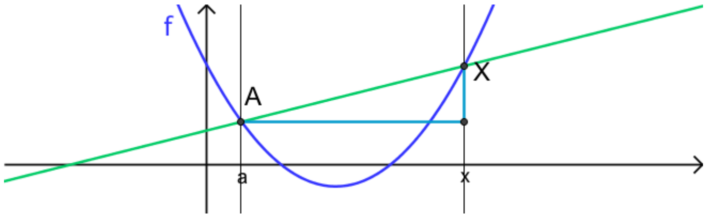
6.1 Hinführung

Die Steigung m einer Geraden g gibt an, wie steil bzw. flach eine Gerade verläuft und ob sie steigt oder fällt. Bei einer Ursprungsgeraden ist das die zweite Koordinate des Punktes $P(1|y_P)$ auf der Geraden. Allgemein hat die Gerade g durch zwei nicht auf einer Parallelen zur y-Achse liegenden Punkte P und Q die Steigung $m_g = \frac{y_Q-y_P}{x_Q-x_P}$.



Möchte man bei einem nicht geradelinigen Kurvenverlauf die Steigung beschreiben, so kann man das näherungsweise mithilfe von Sekanten tun: Um einen Näherungswert für die Steigung des Graphen einer Funktion f an einer Stelle a zu erhalten, legt man eine Sekante durch den Kurvenpunkt $A(a|f(a))$ und einen davon verschiedenen Kurvenpunkt $X(x|f(x))$ und berechnet die Steigung der Strecke AX.

$$m_{AX} = \frac{f(x) - f(a)}{x - a}$$



Allerdings entspricht die so ermittelte Durchschnittssteigung des Graphen von f über dem Intervall $[a; x]$ (bzw. $[x; a]$, falls x kleiner als a ist) anschaulich nur dann näherungsweise der Steigung des Graphen von f bei a, wenn $|x - a|$ klein genug ist. So ergibt sich im skizzierten Beispiel als Steigung von AX eine positive Zahl, während ganz offensichtlich der Graph von f durch den Punkt A nicht steigt, sondern fällt.
Man möchte die Steigung an der Stelle a zu einer vorgegebenen Fehlerschranke ε durch einen Wert beschreiben, der sich auch bei noch größerer Annäherung der Stelle a durch die Stelle x von der zugehörigen Durchschnittssteigung um weniger als ε unterscheidet.
Diese Forderung erfüllt - falls er existiert - der Grenzwert der Durchschnittsteigungsfunktion g_a an der Stelle a:

$$g_a : A_f \setminus \{a\} \longrightarrow \mathbb{R} \text{ mit } g_a(x) = \frac{f(x) - f(a)}{x - a}$$

Dies führt zu folgender Definition der Differenzierbarkeit:

6.2 Definition der Differenzierbarkeit

Eine Funktion f heißt *differenzierbar* an der Stelle c ihrer Argumengmenge A, wenn die durch $g(x) = \frac{f(x)-f(c)}{x-c}$ über $A \setminus \{c\}$ definierte Funktion g an der Stelle c einen Grenzwert hat; also formaler:

$$c \in A \subset \mathbb{R}; \quad f : A \longrightarrow \mathbb{R}$$
$$f \text{ differenzierbar bei } c :\Longleftrightarrow \bigvee_{\lambda \in \mathbb{R}} \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = \lambda$$

Falls dieser Grenzwert λ existiert, wird er als *Ableitung von f an der Stelle c* bezeichnet und als $f'(c)$ notiert. Damit wird auf der Menge aller Stellen von A, an denen f differenzierbar ist, eine Funktion f' definiert.

6.3 Folgerungen

6.3.1 Tangentengleichung

Die Gerade mit der Steigung m, die durch den Punkt P verläuft, hat bekanntlich die Gleichung $y = y_P + m \cdot (x - x_P)$.
Die im Punkte P mit den Koordinaten $(c|f(c))$ an den Graphen von f angelegte Tangente hat die Steigung $f'(c)$ und daher die Gleichung $y = f(c) + f'(c) \cdot (x - c)$.

6.3.2 Zusammenhang zwischen Stetigkeit und Differenzierbarkeit

(1) Aus Differenzierbarkeit folgt Stetigkeit

Zu zeigen ist, dass aus $\lim_{x\to c}\left(\frac{f(x)-f(c)}{x-c} - f'(c)\right) = 0$ folgt $\lim_{x\to c}(f(x) - f(c)) = 0$. Dies ergibt wegen

$$f(x) - f(c) = \left(\frac{f(x) - f(c)}{x - c} - f'(c)\right)(x - c) + f'(x) \cdot (x - c) \longrightarrow 0 \quad (\text{für } x \to c)$$

(2) Aus Stetigkeit folgt <u>nicht</u> Differenzierbarkeit

Dies wird durch Beispiel einer Funktion gezeigt, die an der Stelle 0 zwar stetig, aber nicht differenzierbar ist:

Durch

$$\bigwedge_{x\in\mathbb{R}^*} f(x) := x \cdot \sin(\frac{1}{x}) \quad \wedge \quad f(0) = 0$$

wird eine auf ganz $\mathbb{R}$ stetige Funktion definiert, da das Produkt einer bei 0 gegen 0 konvergenten Funktion mit einer in einer Umgebung von 0 beschränkten Funktion bei 0 den Grenzwert 0 hat. Aber die Durchschnittssteigung über dem Intervall $[0; x]$, also $\sin(\frac{1}{x})$ nimmt in jeder Umgebung von 0 die Werte 0 und 1 an (nämlich bei $x = \frac{n}{\pi}$ bzw. $x = \frac{2}{\pi(1+4n)}; n \in \mathbb{N}$), hat also an der Stelle 0 keinen Grenzwert.

6.3.3 Lokaler Wachstumssatz

Ist eine Funktion f mit Argumentmenge A an einer Stelle $c \in A$ differenzierbar, und ist die Ableitung dort positiv, so gibt es ein offenes Intervall I um die Stelle c, für das folgendes gilt:

$$\bigwedge_{x,y\in I\cap A} (x < c < y \Rightarrow f(x) < f(c) < f(y))$$

Denn nach dem Positivitätssatz für Grenzwerte folgt aus $\lim_{x\to c} \frac{f(x)-f(c)}{x-c} > 0$, dass es ein offenes Intervall I um c gibt mit

$$\bigwedge_{x\in I\cap A\setminus\{c\}} \frac{f(x) - f(c)}{x - c} > 0.$$

Für alle $x \in I \cap A \setminus \{c\}$ sind daher $f(x) - f(c)$ und $x - c$ vorzeichengleich. Daher gilt

$$\bigwedge_{x,y\in I\cap A} (x < c \Rightarrow f(x) < f(c)) \quad \wedge \quad (y > c \Rightarrow f(y) > f(c)).$$

6.3.4 Notwendige Bedingung für ein lokales Maximum

Eine Funktion f mit der Argumentmenge A hat genau dann an einer Stelle $c \in A$ ein lokales Maximum, wenn es ein offenes Intervall I um c gibt, in dem die Funktion an keiner Stelle einen größeren Wert als $f(c)$ annimmt:

$$f \text{ hat bei } c \text{ ein lokales Maximum } :\Leftrightarrow \bigvee_{\delta\in\mathbb{R}_+} \bigwedge_{x\in A} (|x - c| < \delta \Rightarrow f(x) \leq f(c))$$

Wäre die Ableitung von f an einer Stelle c mit einem lokalen Maximum verschieden von 0, so müsste c ein Randpunkt der Argumentmenge sein, da sonst nach dem lokalen Wachstumssatz in jedem offenen

Intervall um c auch größere Werte als $f(c)$ angenommen würden. Daher gilt die folgende notwendige Bedingung für das Vorliegen eines lokalen Maximums:
Ist c innerer Punkt der Argumentmenge A einer bei c differenzierbaren Funktion f, und hat f bei c ein lokales Maximum, so gilt $f'(c) = 0$.

6.4 Ableitungsregeln

Zum Nachweis der Differenzierbarkeit einer Funktion f an einer Stelle c und zur Berechnung der Ableitung $f'(c)$ kann nach folgendem Schema verfahren werden:

1. Notieren des Terms $\frac{f(x)-f(c)}{x-c}$ für die mittlere Steigung über dem Intervall $[c; x]$ (bzw. $[x; c]$), des sog. *Differenzenquotienten.* Da der Nenner und (bei Stetigkeit von f an der Stelle c) der Zähler beide bei c den Grenzwert 0 haben, lässt sich ohne weitere Umformung noch keine Aussage über die Differenzierbarkeit machen.

2. Algebraische Umformung des Differenzenquotienten zu einem oder mehreren Operanden, für die jeweils ein Grenzwert an der Stelle c zu erkennen ist.

Anstelle des oben angegebenen Differenzenquotienten $\frac{f(x)-f(c)}{x-c}$ kann man auch den durch Substitution mit $x = c+h$ entstehenden Ausdruck $\frac{f(c+h)-f(c)}{h}$ betrachten. Er hat genau dann bei 0 einen Grenzwert, wenn die Funktion f bei c differenzierbar ist, und dieser Grenzwert ist die Ableitung $f'(c)$.

6.4.1 Ableitungsbeispiele

- Eine konstante Funktion f ist an jeder Stelle c differenzierbar; $f'(c) = 0$.

  Denn wenn k die reelle Zahl ist, die von f an jeder Stelle als Wert angenommen wird, so gilt:

$$\frac{f(x)-f(c)}{x-c} = \frac{k-k}{x-c} = 0 \to 0 \quad \text{für } x \to c.$$

- Die identische Funktion f, die an jeder Stelle x den Wert x annimmt ist an jeder Stelle c differenzierbar; $f'(c) = 1$, denn

$$\frac{f(x)-f(c)}{x-c} = \frac{x-c}{x-c} = 1 \to 1 \quad \text{für } x \to c.$$

- Die Kehrwertfunktion f, die an jeder Stelle $x \in \mathbb{R}^*$ den Wert $\frac{1}{x}$ annimmt, ist an jeder Stelle $c \in \mathbb{R}^*$ differenzierbar; $f'(c) = -\frac{1}{c^2}$, denn

$$\frac{f(x)-f(c)}{x-c} = \frac{\frac{1}{x}-\frac{1}{c}}{x-c} = \frac{c-x}{xc(x-c)} = \frac{-1}{xc} \to -\frac{1}{c^2} \quad \text{für } x \to c.$$

- Die Wurzelfunktion f, die an jeder Stelle $x \in \mathbb{R}_+ \cup \{0\}$ den Wert $\sqrt{x}$ annimmt, ist an jeder Stelle $c \in \mathbb{R}_+$ differenzierbar; $f'(c) = \frac{1}{2\cdot\sqrt{c}}$, denn

$$\frac{f(x)-f(c)}{x-c} = \frac{\sqrt{x}-\sqrt{c}}{x-c} = \frac{x-c}{(\sqrt{x}+\sqrt{c})(x-c)} = \frac{1}{\sqrt{x}+\sqrt{c}} \to \frac{1}{2\cdot\sqrt{c}} \quad \text{für } x \to c.$$

6.4.2 Ableitung einer Integralfunktion

Ist f eine auf einem Intervall $[a;b]$ definierte, integrierbare Funktion, dann erklärt man bekanntlich die zugehörige Integralfunktion F durch

$$F:[a;b]\longrightarrow\mathbb{R};\ \bigwedge_{x\in[a;b]} F(x):=\int_a^x f(t)\mathrm{d}t.$$

Die Integralfunktion ist zwar überall stetig, aber nicht notwendigerweise differenzierbar.
Ist allerdings der Integrand f an einer Stelle c stetig, dann ist F bei c differenzierbar und hat dort die Ableitung $f(c)$, wie nachfolgend gezeigt wird.

$$\left|\frac{F(x)-F(c)}{x-c}-f(c)\right| = \left|\frac{\int_a^x f(t)\mathrm{d}t-\int_a^c f(t)\mathrm{d}t}{x-c}-\frac{\int_c^x f(c)\mathrm{d}t}{x-c}\right| = \left|\frac{\int_c^x f(t)\mathrm{d}t-\int_c^x f(c)}{x-c}\right|$$
$$= \left|\frac{\int_c^x (f(t)-f(c))\mathrm{d}t}{x-c}\right| \le \left|\frac{\int_c^x |f(t)-f(c)|\mathrm{d}t}{x-c}\right| \le \frac{|x-c|\cdot\sup\{|f(t)-f(c)|\mid t\in[x;c]\cup[c;x]\}}{|x-c|}$$
$$= \sup\{|f(t)-f(c)|\mid t\in[x;c]\cup[c;x]\}\to 0 \quad \text{für } x\to c.$$

Beispiel: Der natürliche Logarithmus $\ln(x)$ einer positiven, reellen Zahl x kann auf folgende Weise definiert werden

$$\bigwedge_{x\in\mathbb{R}_+} \ln(x):=\int_1^x \frac{1}{t}\,\mathrm{d}t.$$

Als Integralfunktion mit stetigem Integranden ist ln überall differenzierbar; $\ln'(x)=\frac{1}{x}$.

6.4.3 Faktorregel

Ist f eine an einer Stelle c differenzierbare Funktion, k eine reelle Zahl, dann ist $k\cdot f$ bei c differenzierbar; $(k\cdot f)'(c)=k\cdot f'(c)$, denn

$$\frac{(k\cdot f)(x)-(k\cdot f)(c)}{x-c}=\frac{k\cdot f(x)-k\cdot f(c)}{x-c}=k\cdot\frac{f(x)-f(c)}{x-c}\to k\cdot f'(c) \quad \text{für } x\to c.$$

6.4.4 Summenregel

Sind die Funktionen f und g bei c differenzierbar, dann ist auch die Summe $f+g$ bei c differenzierbar; $(f+g)'(c)=f'(c)+g'(c)$, denn

$$\frac{(f+g)(x)-(f+g)(c)}{x-c}=\frac{f(x)+g(x)-(f(c)+g(c))}{x-c}=\frac{f(x)-f(c)+g(x)-g(c)}{x-c}$$
$$=\frac{f(x)-f(c)}{x-c}+\frac{g(x)-g(c)}{x-c}\to f'(c)+g'(c) \quad \text{für } x\to c.$$

6.4.5 Produktregel

Sind die Funktionen f und g bei c differenzierbar, dann ist auch das Produkt $f\cdot g$ bei c differenzierbar; $(f\cdot g)'(c)=f'(c)\cdot g(c)+f(c)\cdot g'(c)$, denn

$$\frac{(f\cdot g)(x)-(f\cdot g)(c)}{x-c}=\frac{f(x)\cdot g(x)-f(c)\cdot g(c))}{x-c}=\frac{(f(x)-f(c))\cdot g(x)+f(c)\cdot(g(x)-g(c))}{x-c}$$
$$=\frac{f(x)-f(c)}{x-c}\cdot g(x)+f(c)\cdot\frac{g(x)-g(c)}{x-c}\to f'(c)\cdot g(c)+f(c)\cdot g'(c) \quad \text{für } x\to c.$$

6.4.6 Potenzregel

Aus der Produktregel ergibt sich durch vollständige Induktion: Ist n eine positive ganze Zahl, und ist die Funktionen f bei c differenzierbar, dann ist auch die Funktion f^n an der Stelle c differenzierbar; $(f^n)'(c) = n \cdot f^{n-1}(c) \cdot f'(c)$.

6.4.7 Kettenregel

Gehört für eine durch Verkettung entstandene Funktion $g \circ f$ die Stelle c zum Differenzierbarkeitsbereich von f, die Stelle $f(c)$ zum Differenzierbarkeitsbereich von g, dann ist die Funktion $g \circ f$ bei c differenzierbar; $(g \circ f)'(c) = g'(f(c)) \cdot f'(c)$.
Der Nachweis wird zunächst für den Fall $f'(c) \neq 0$ geführt; in diesem Fall sind nach dem lokalen Wachstumssatz in einer geeigneten Umgebung von c alle Funktionswerte $f(x)$ für $x \neq c$ verschieden von $f(c)$.

$$\begin{aligned}\frac{(g \circ f)(x) - (g \circ f)(c)}{x - c} = \frac{g(f(x)) - g(f(c))}{x - c} &= \frac{g(f(x)) - g(f(c))}{f(x) - f(c)} \cdot \frac{g(x) - g(c)}{x - c} \\ &\to g'(f(c)) \cdot f'(c) \quad \text{für } x \to c.\end{aligned}$$

Zu zeigen ist noch, dass $g \circ f$ auch im Falle $f'(c) = 0$ bei c differenzierbar ist und dort die Ableitung 0 hat.
Aufgrund der Differenzierbarkeitsvoraussetzungen für g bei $f(c)$ und für f bei c haben die folgenden beiden Funktionen γ und ϕ bei 0 den Grenzwert 0:

$$\gamma(t) := \frac{g(f(c) + t) - g(f(c))}{t} - g'(f(c)); \quad \phi(h) := \frac{f(c + h) - f(c)}{h}$$

Umformung ergibt

$$g(f(c) + t) = g(f(c)) + (\gamma(t) + g'(f(c))) \cdot t; \quad f(c + h) = f(c) + h \cdot \phi(h)$$

und somit

$$g(f(c + h)) = g(f(c) + h \cdot \phi(h)) = g(f(c)) + (\gamma(h \cdot \phi(h)) + g'(f(c))) \cdot h \cdot \phi(h).$$

Daraus folgt

$$\begin{aligned}\frac{g(f(c + h)) - g(f(c))}{h} &= \frac{(\gamma(h \cdot \phi(h)) + g'(f(c))) \cdot h \cdot \phi(h)}{h} \\ &= \gamma(h \cdot \phi(h)) + g'(f(c))) \cdot \phi(h) \to 0 \quad \text{für } h \to 0.\end{aligned}$$

6.4.8 Reziprokenregel

Ist die Funktion f an der Stelle c differenzierbar, und ist $f(c) \neq 0$, dann ist auch $\frac{1}{f}$ bei c differenzierbar; $(\frac{1}{f})'(c) = -\frac{f'}{f^2}(c)$, denn da f stetig ist, sind wegen $f(c) \neq 0$ alle Funktionswerte $f(x)$ in einer geeigneten Umgebung von c verschieden von null. Bezeichnet man die Kehrwertfunktion mit g, so ist $\frac{1}{f} = g \circ f$. Nach der Kettenregel ist $\frac{1}{f}$ also bei c differenzierbar:

$$(\frac{1}{f})'(c) = (g \circ f)'(c) = g'(f(c)) \cdot f'(c) = -\frac{1}{f^2(c)} \cdot f'(c)$$

6.4.9 Quotientenregel

Sind die Funktionen f und g an der Stelle c differenzierbar, und ist $g(c) \neq 0$ (und damit $g(x) \neq 0$ in einer geeigneten Umgebung von c), dann ist $\frac{f}{g} (= f \cdot \frac{1}{g})$ nach der Produktregel bei c differenzierbar.
Man erhält

$$(\tfrac{f}{g})'(c) = (f \cdot \tfrac{1}{g})'(c) = f'(c) \cdot (\tfrac{1}{g})(c) + f(c) \cdot (\tfrac{1}{g})'(c) = \tfrac{f'(c)}{g(c)} + f(c) \cdot \tfrac{-g'(c)}{g^2(c)} = \tfrac{f'(c) \cdot g(c) - f(c) \cdot g'(c)}{g^2(c)}$$

6.4.10 Ableitung der Umkehrfunktion

Ist die Funktion f streng monoton über einem Intervall um c und differenzierbar bei c mit $f'(c) \neq 0$, dann ist die Umkehrfunktion f^{-1} differenzierbar bei $f(c)$, und es gilt $(f^{-1})'(f(c)) = \frac{1}{f'(c)}$.

$$\frac{f^{-1}(f(x)) - f^{-1}(f(c))}{f(x) - f(c)} = \frac{x - c}{f(x) - f(c)} = \frac{1}{\frac{f(x)-f(c)}{x-c}} \to \frac{1}{f'(c)} \quad \text{für } x \to c.$$

Ergänzende Bemerkungen:

- Die Voraussetzung $f'(c) \neq 0$ ist nicht bereits durch die strenge Monotonie von f gesichert, wie z.B. die Potenzfunktion dritten Grades (für $c = 0$) zeigt.
- Da f bei c stetig ist, gilt $\lim_{x\to c} f(x) = f(c)$; nach dem lokalen Wachstumssatz sind für x aus einer geeigneten Umgebung von $c \quad (x \neq c)$ die Funktionswerte $f(x)$ verschieden von $f(c)$. Wenn also $\frac{f^{-1}(f(x))-f^{-1}(f(c))}{f(x)-f(c)}$ an der Stelle $f(c)$ einen Grenzwert hat, dann folgt daraus die behauptete Differenzierbarkeit bei $f(c)$, und der Grenzwert ist die gesuchte Ableitung.
- Setzt man bereits voraus, dass die Umkehrfunktion f^{-1} bei $f(c)$ differenzierbar ist, dann ist die Ableitung mithilfe der Kettenregel und der Ableitung der identischen Funktion einfach zu berechnen:

$$\bigwedge_x (f^{-1} \circ f)(x) = x \Longrightarrow (f^{-1})'(f(c)) \cdot f'(c) = 1 \qquad \text{und daher} \quad (f^{-1})'(f(c)) = \frac{1}{f'(c)}.$$

6.5 Tabelle der wichtigsten Ableitungen

Funktion	f	$k \cdot f$	$f + g$	$f \cdot g$	$\frac{f}{g}$	$g \circ f$	f^{-1}
Ableitung	f'	$k \cdot f'$	$f' + g'$	$f' \cdot g + f \cdot g'$	$\frac{f' \cdot g - f \cdot g'}{g^2}$	$(g' \circ f) \cdot f'$	$\frac{1}{f'}$

Außerdem gilt für spezielle Funktionen mit

$$p_n(x) := x^n, \quad r(x) = \frac{1}{x}, \quad w(x) := \sqrt{x}; \quad \exp(x) := e^x :$$

Funktion	p_n	r	w	exp	ln	sin	cos	tan	sinh	cosh
Ableitung	$n \cdot p_{n-1}$	$-\frac{r'}{r^2}$	$\frac{1}{2 \cdot w}$	exp	r	cos	$-\sin$	$\frac{1}{\tan^2}$	cosh	sinh

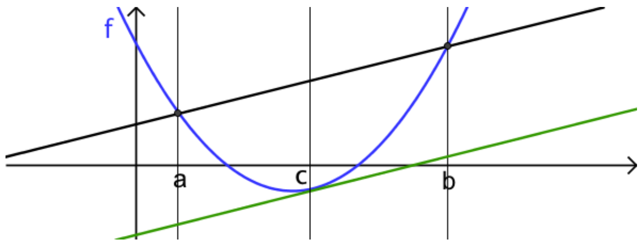
7 Eigenschaften differenzierbarer Funktionen

Eine Funktion heißt differenzierbar, wenn sie an jeder Stelle ihrer Argumentmenge differenzierbar ist.

7.1 Mittelwertsatz der Differentialrechnung

7.1.1 Mittelwertsatz

Ist f eine auf einem abgeschlossenen Intervall $[a; b]$ stetige und im Inneren des Intervalls differenzierbare Funktion, dann gibt es im Inneren des Intervalls eine Stelle c , an der die zugehörige Tangente parallel zur Sekante durch die Punkte $(a|f(a))$ und $(b|f(b))$ verläuft.



$$f : [a;b] \longrightarrow \mathbb{R}, \text{stetig, auf }]a;b[\text{ differenzierbar} \Longrightarrow \bigvee_{c\in]a;b[} \frac{f(b)-f(a)}{b-a} = f'(c)$$

Der Beweis wird im Anschluss an eine Reduktion auf einen Spezialfall durchgeführt.

1. Reduktion auf den Spezialfall $f(a) = f(b) = 0$

   Definiert man die Hilfsfunktion h durch

   $$\bigwedge_{x\in[a;b]} h(x) := f(x) - f(a) \cdot \frac{f(b)-f(a)}{b-a} \cdot (x-a),$$

   also als Differenz von f und der linearen Funktion, deren Graph die betrachtete Sekante ist, dann ist h differenzierbar und hat an den Stellen a und b den Wert 0. Gilt also der Mittelwertsatz in der reduzierten Form, dann gibt es im Intervall $]a;b[$ eine Stelle c mit $h'(c) = 0$; wegen $h'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$ ergibt sich daraus die allgemeine Beahuptung des Mittelwertsatzes.

2. Nachweis der reduzierten Behauptung

   Satz von Rolle: Als stetige Funktion auf einem abgeschlossenen Intervall hat f dort Minimum und Maximum. Werden beide am Rand angenommen ist die Funktion wegen $f(a) = f(b)$ konstant, hat also (z.B.) an der Stelle $c = \frac{a+b}{2}$ die Ableitung 0.

   Wird aber Minimum oder Maximum an einer Stelle c aus $]a;b[$ angenommen, so folgt aus dem lokalen Wachstumssatz, dass $f'(c) = 0$ gelten muss.

3. Modifikation der Behauptung

   Durch Multiplikation der Gleichung in (1) mit $b-a$ und Addition von $f(a)$ erhält die Aussage des Mittelwertsatzes die Form $\bigvee_{c\in]a;b[} f(b) = f(a) + f'(c) \cdot (b-a)$.

Die (anschaulich klare) Folgerung, dass eine Funktion, deren Ableitungsfunktion auf einem Intervall I den konstanten Wert 0 hat, selber konstant ist, ergibt sich aus dem Mittelwertsatz. Denn an je zwei verschiedenen Stellen $x, y \in I$ wird der gleiche Funktionswert angenommen, da es nach dem Mittelwertsatz zwischen x und y eine Stelle ξ gibt mit

$$f(y) = f(x) + f'(\xi) \cdot (y-x) = f(x) + 0 \cdot (y-x) = f(x).$$

7.1.2 Verallgemeinerter Mittelwertsatz

Während der oben angegebene Mittelwertsatz von der Anschauung her naheliegend ist, bietet sich eine anschauliche Deutung bei seiner Verallgemeinerung nicht so unmittelbar an. Es gilt nämlich allgemeiner:

$$(f, g : [a;b] \longrightarrow \mathbb{R}, \text{stetig, auf }]a;b[\text{ diffb.} \wedge \bigwedge_{x\in]a;b[} g'(x) \neq 0) \Longrightarrow \bigvee_{c\in]a;b[} \frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}.$$

Als Spezialfall ergibt sich daraus mit $\bigwedge_{x\in[a;b]} g(x) := x$ der oben angegebene Mittelwertsatz.
Die Verallgemeinerung lässt sich nicht einfach durch zweimaliges Anwenden des Mittelwertsatzes (nämlich auf f und auf g) gewinnen, da die für jede der Funktionen im Sinne des Satzes existierende Stelle c im allgemeinen nicht die gleiche ist.
Zum Nachweis des verallgemeinerten Mittelwersatzes betrachte man unter den angegebenen Voraussetzungen für f und g die Hilfsfunktion h:

$$\bigwedge_{x\in[a;b]} h(x) := \begin{vmatrix} g(x) & g(b) & g(a) \\ f(x) & f(b) & f(a) \\ 0 & 1 & 1 \end{vmatrix}.$$

Da h als Linearkombination von f und g differenzierbar im Inneren und stetig am Rand von $[a;b]$ ist, sind wegen $h(a) = h(b) = 0$ die Voraussetzungen des Satzes von Rolle erf"llt. Es gibt also im Intervall $]a;b[$ eine Stelle c mit

$$h'(c) = 0; \quad \text{also} \quad f'(c)\cdot(g(b)-g(a)) - g'(c)\cdot(f(b)-f(a)) = 0.$$

Da g aufgrund der Voraussetzung streng monoton ist, darf durch $g(b) - g(a)$ dividiert werden, so dass sich die behauptete Gleichung ergibt.

7.2 Folgerungen aus dem Mittelwertsatz

7.2.1 $f' = o \quad \Rightarrow \quad f$ ist konstant

Gibt es für eine auf $]a;b[$ differenzierbare und bei a, b stetige Funktion f Stellen c_1, c_2, an den unterschiedliche Funktionswerte angenommen werden, gibt es nach dem Mittelwertsatz auch eine Stelle c zwischen c_1 und c_2, an denen die Ableitung verschieden von null ist. Nur eine konstante Funktion hat also als Ableitung die Nullfunktion.

7.2.2 Globaler Wachstumssatz

Hat eine auf einem abgeschlossenen Intervall $[a;b]$ stetige, im Inneren des Intervalls differenzierbare Funktion an allen Stellen aus $]a;b[$ eine positive Ableitung, dann ist sie streng isoton.
Zum Nachweis ist zu zeigen:

$$\bigwedge_{x,y\in[a;b]} (x < y \Rightarrow f(x) < f(y))$$

Aus $a \leq x < y \leq b$ folgt nach dem Mittelwertsatz die Existenz einer Stelle $c \in]x;y[$ mit

$$f(y) = f(x) + f'(c)(y-x) > f(x).$$

Bemerkung: Umgekehrt lässt sich aus der strengen Isotonie einer differenzierbaren Funktion über einem Intervall nicht folgern, dass die Ableitung der Funktion überall positiv ist, wie z.B. das Gegenbeispiel der Potenzfunktion 3. Grades zeigt.

7.2.3 Fixpunktsatz für differenzierbare Funktionen

Ist f eine stetige, auf einem abgeschlossenen Intervall $[a;b]$ definierte Funktion, deren Funktionswerte alle in $[a;b]$ liegen, dann hat f nach dem entsprechenden Satz über stetige Funktionen (mindestens) einen Fixpunkt. Setzt man zusätzlich voraus, dass f differenzierbar ist, und eine reelle Zahl q mit $0 \leq q < 1$ existiert, durch welche die Ableitung an allen Stellen betraglich nach oben beschränkt ist, so gilt zusätzlich zur reinen Existenzaussage über einen Fixpunkt:

(1) Die Funktion f hat genau einen Fixpunkt c.

(2) Definiert man eine Folge (a_n) durch

$$a_0 := a \quad \wedge \quad \bigwedge_{n\in\mathbb{N}} a_{n+1} := f(a_n),$$

dann konvergiert die Folge (a_n) gegen den Fixpunkt c.

Zu (1): Es muss nur noch gezeigt werden, dass nicht zwei verschiedene Fixpunkte c und d existieren können. Aus der Annahme, dass c und d $(c, d \in [a;b]; c \neq d)$ Fixpunkte sind, folgt nach dem Mittelwertsatz, dass im Intervall mit den Rändern c und d eine Stelle ξ existiert mit

$$|d-c| = |f(d) - f(c)| = |f'(\xi)| \cdot |d-c| \leq q \cdot |d-c| < |d-c|,$$

woraus durch Widerspruch die Eindeutigkeit des Fixpunkts folgt.

Zu (2): Nach dem Mittelwertsatz gibt es für jede natürliche Zahl n mit $a_n \neq c$ eine Stelle ξ im Intervall zwischen a_{n+1} und c, für die folgendes gilt:

$$(|a_{n+1} - c| = |f(a_n) - f(c)| = |f'(\xi)| \cdot |a_n - c| \leq q \cdot |a_n - c|)$$

Die so erhaltene Ungleichung $|a_{n+1} - c| \leq q \cdot |a_n - c|$ gilt wegen $f(c) = c$ auch für den Fall $a_n = c$. Durch vollständige Induktion zeigt man

$$|a_0 - c| \leq |a - c| \wedge \bigwedge_{n\in\mathbb{N}} |a_{n+1} - c| \leq q \cdot |a_n - c| \Longrightarrow \bigwedge_{n\in\mathbb{N}} |a_n - c| \leq q^n \cdot |a - c|.$$

Da (q^n) wegen $|q| < 1|$ eine Nullfolge ist, ist nach dem Majorantensatz auch $(a_n - c)$ eine Nullfolge, was äquivalent zu $\lim_{n\to\infty} a_n = c$ ist.

Bemerkung: Aufgrund der Eindeutigkeit des Grenzwerts folgt aus (2) noch einmal die in (1) bewiesene Eindeutigkeit des Fixpunktes.

7.3 Zwischenwertsatz für Ableitungen

Ist eine Funktion f auf einem Intervall $[a;b]$ differenzierbar mit $f'(a) < f'(b)$, und ist λ eine reelle Zahl aus dem offfenen Intervall $]f(a); f(b)[$, dann gibt es in $]a;b[$ eine Stelle ξ mit $f'(\xi) = \lambda$ [21] .

1. Reduktion

   Es genügt, den Beweis für den Spezialfall $\lambda = 0$ zu führen, denn für die durch

   $$\bigwedge_{x\in[a;b]} g(x) := f(x) - \lambda x$$

   definierte Funktion g folgt dann wegen $g'(a) = f'(a) - \lambda < 0 \ \wedge \ g'(b) = f'(b) - \lambda x > 0$

   die Existenz einer Stelle $\xi \in]a;b[$ mit $g'(\xi) = 0$, also $f'(\xi) = \lambda$.

[21] Dieser Satz wird auch als *Zwischenwertsatz von Darboux* nach dem französischen Mathematiker Gaston Darboux (gest. 1917) bezeichnet.

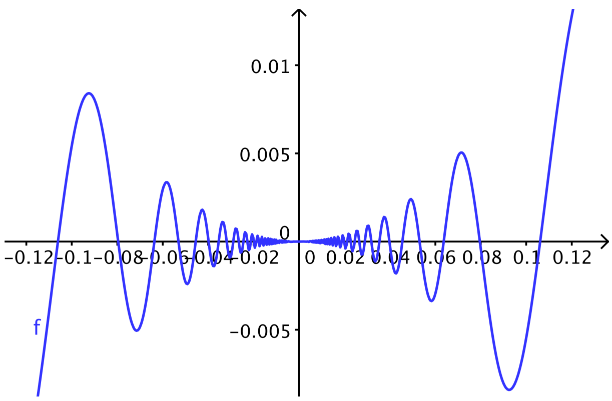
2. Beweis der auf den Spezialfall $\lambda = 0$ reduzierten Behauptung

   Nach dem Satz vom Maximum und Minimum stetiger Funktionen auf abgeschlossenen Intervallen gibt es eine Stelle $\xi \in [a; b]$, an der f ein Minimum annimmt. Dabei kann ξ keiner der Werte a oder b sein, da nach dem lokalen Wachstumssatz f wegen $f'(a) < 0$ rechts von a kleinere Werte als $f(a)$ annimmt; analog folgt $\xi \neq b$.

   Mithin liegt ξ im offenen Intervall $]a; b[$; aufgrund der notwendigen Bedingung für Extrema an inneren Punkten der Argumentmenge gilt also $f'(\xi) = 0$.

3. Abgrenzung zum Zwischenwertsatz für stetige Funktionen durch ein Gegenbeispiel

   - Da sich jede auf einem Intervall $[a; b]$ definierte stetige Funktion f als Ableitung der durch
     $$\bigwedge_{x\in[a;b]} F(x) := \int_a^x f(t)\mathrm{dt}$$
     definierten Integralfunktion F erhalten lässt, ist damit auch ein (weiterer) Beweis des Zwischenwertsatzes für stetige Funktionen gegeben.
   - Umgekehrt braucht aber die Ableitung einer differenzierbaren Funktion keineswegs stetig zu sein, wie das Beispiel der folgenden auf ganz $\mathbb{R}$ definierten Funktion f zeigt:
     $$\bigwedge_{x\in\mathbb{R}^*} f(x) := x^2 \cdot \sin(\frac{1}{x}); \quad f(0) := 0$$
     Die Funktion ist bei 0 differenzierbar mit Ableitung 0, denn
     $$\frac{f(x) - f(0)}{x - 0} = \frac{x^2 \cdot \sin(\frac{1}{x}) - 0}{x} = x \cdot \sin(\frac{1}{x}) \to 0 \quad (x \to 0).$$
     Für jede von 0 verschiedene Stelle x liefern die Ableitungsregeln
     $$f'(x) = 2x \cdot \sin(\frac{1}{x}) + x^2 \cdot \cos(\frac{1}{x}) \cdot \frac{-1}{x^2} = 2x \cdot \sin(\frac{1}{x}) - \cos(\frac{1}{x}).$$
     An der Stelle 0 hat $f'(0)$ nicht den Grenzwert 0 (und auch keinen anderen Grenzwert), denn $\lim_{x\to 0} 2x \cdot \sin(\frac{1}{x}) = 0$, wohingegen es zu $\cos(\frac{1}{x})$ in jedem Intervall um 0 Stellen gibt, an denen der Wert 1 angenommen wird; man betrachte dazu die Stelle $x = \frac{1}{2n\pi}; n \in N^*$.



4. Folgerung zur Art der möglichen Unstetigkeit bei Ableitungen

   Die einzig mögliche Art der Unstetigkeit ist oszillatorisch, da sowohl eine Sprungstelle wie auch ein uneigentlicher Grenzwert ∞ durch den oben angegebenen Zwischenwertsatz für Ableitungen ausgeschlossen werden.

8 Vom Satz von Rolle zum Satz von Taylor

8.1 Der verallgemeinerte Satz von Rolle

Stellt man sich einen mit Langhölzern beladenen Wagen vor, der im Tal startend und in einem anderen Tal seine Fahrt beendend über einen Berg fährt, so ist anschaulich klar, dass zu irgendeinem Zeitpunkt der Fahrt die Richtung der Langhölzer horizontal ist. Die entsprechende Eigenschaft findet man beim Graphen einer Funktion, vorausgesetzt sie erfüllt gewisse Voraussetzungen. Dies ist Inhalt des weiter oben bewiesenen Satzes von Rolle[22].
Dieser Satz wird nachfolgend verallgemeinert:

Voraussetzung: Es sei n eine natürliche Zahl und f eine auf einem abgeschlossenen Intervall $[a;b]$ stetige Funktion, die bei a n-mal und im Inneren des Intervalls $(n+1)$-mal differenzierbar ist; weiterhin gelte: $f(a) = f'(a) = f''(a) = ... = f^{(n)}(a) = f(b) = 0$.

Behauptung: Es gibt im Intervall $]a;b[$ eine Stelle c mit $f^{(n+1)}(c) = 0$.

Beweis (durch Induktion): Die Aussage des Satzes ist für $n = 0$ offensichtlich richtig, denn dann stimmt der Satz mit dem Satz von Rolle überein.

Der erste Teil des Induktionsbeweises ist damit abgeschlossen.
Im zweiten Teil des Induktionsbeweises wird nun vorausgesetzt, dass n eine natürliche Zahl ist, und die Aussage des verallgemeinerten Satzes von Rolle für n richtig ist. Hiermit ist dann zu zeigen:

Wenn f eine auf einem abgeschlossenen Intervall $[a;b]$ stetige Funktion ist, die an der Stelle a $(n+1)$-mal und im Inneren des Intervalls $(n+2)$-mal differenzierbar ist,

wobei $f(a) = f'(a) = f''(a) = ... = f^{(n)}(a) = f^{(n+1)}(a) = f(b) = 0$ gilt, dann gibt es im Intervall $]a;b[$ eine Stelle c mit $f^{(n+2)}(c) = 0$.

Nach Induktionsannahme gibt es wegen $f(a) = f'(a) = f''(a) = ... = f^{(n)}(a) = f(b) = 0$ im Intervall $]a;b[$ eine Stelle d mit $f^{(n+1)}(d) = 0$.

Somit gilt $f^{(n+1)}(a) = f^{(n+1)}(d) = 0$. Nach dem Satz von Rolle gibt es daher im Intervall $]a;d[$ eine Stelle c mit $(f^{(n+1)})'(c) = 0$.

Wegen $(f^{(n+1)})'(c) = f^{(n+2)}(c)$ und $a < c < d < b$ ist c eine Stelle im offenen Intervall $]a;b[$, für die $f^{(n+2)}(c) = 0$ gilt.

Damit ist der Induktionsbeweis abgeschlossen.

8.2 Der Satz von Taylor

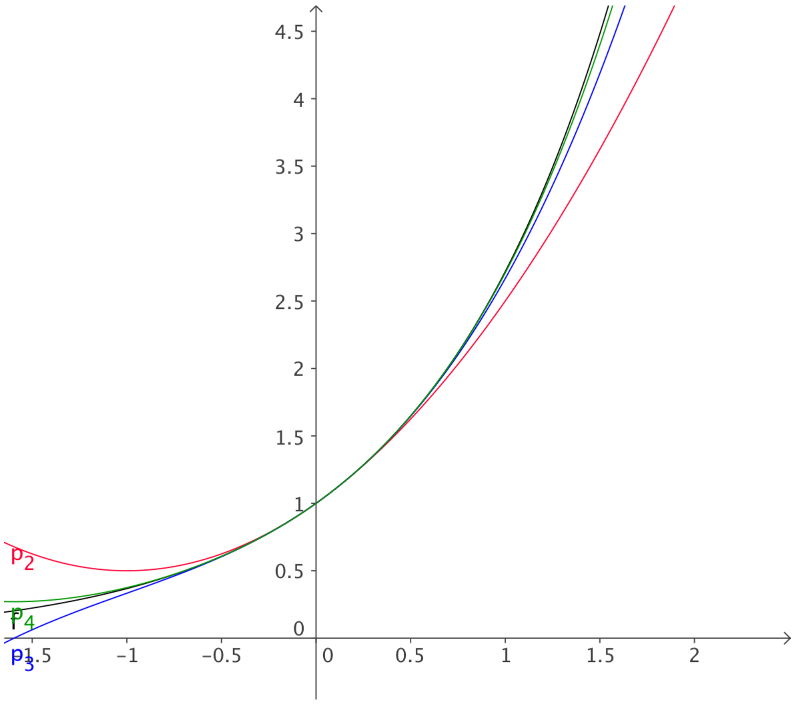
8.2.1 Schmiegepolynome

Der Funktionsterm einer ganzrationalen Funktion p vom Grade n ($n \in \mathbb{N}$) erlaubt eine Darstellung der Form $\sum_{i=0}^{n} a_i x^i$ mit reellen Koeffizienten $a_0, a_1, a_2, \ldots, a_n$.
Als k-te Ableitung ($k \in \{0, 1, 2, \ldots, n\}$) ergibt sich

$$p^{(k)}(x) = \sum_{i=0}^{n-k} \frac{(i+k)!}{i!} \, a_{i+k} \, x^i \ , \quad \text{also speziell für x} = 0 \quad p^{(k)}(0) = k! \cdot a_k \quad .$$

[22] Ist f eine auf einem Intervall $[a;b]$ stetige, im Inneren des Intervalls differenzierbare Funktion mit $f(a) = f(b) = 0$, so gibt es im offenen Intervall $]a;b[$ eine Stelle c, sodass $f'(c) = 0$ gilt.

Unter dem (an der Stelle 0 entwickelten) n-ten Schmiegepolynom $p_n(x)$ einer bei 0 n-mal differenzierbaren Funktion f versteht man den Funktionsterm einer ganzrationalen Funktion n-ten Grades p_n, welche für $i = 0, 1, 2, ..., n$ die Gleichung $p_n^{(i)}(0) = f^{(i)}(0)$ und daher $p^{(k)}(0) = k! \cdot a_k$ erfüllt. Dieses Schmiegeplolynom hat somit die Gleichung $p_n(x) = \sum_{i=1}^{n} \frac{f^{(i)}(0)}{i!} x^i$.
Die Betrachtung von Beispielen solcher Schmiegepolynome, z.B. für die elementaren Funktionen exp, sin, cos, sinh, cosh zeigt, dass sich für größere Werte von n die Graphen der Schmiegepolynome in einer Umgebung von 0 recht gut an den Graphen von f anschmiegen, sodass man die leicht zu berechnenden Funktionswerte des Taylorpolynoms als Ersatzwerte für die nicht so einfach zu berechnenden Funktionswerte von f verwenden kann. Die Skizze zeigt den Graphen der natürlichen Exponentialfunktion sowie die Graphen ihrer Schmiegepolynome p_2, p_3 und p_4.



Allerdings reicht dieser Eindruck einer guten Approximation keineswegs; wenn etwa die konkrete Aufgabe gestellt wird, die Eulersche Zahl e (also exp(1)) mit einem Fehler von weniger als 10^{-6} zu berechnen, ist noch nicht sicher, ob es genügt, das Taylorpolynom $p_n(x)$ für ein geeignetes n an der Stelle 1 zu berechnen und - wenn ja - welcher Wert von n für die gewünschte Genauigkeit ausreichend ist. Man braucht also noch eine Abschätzung für die Differenz $|f(1) - p_n(1)|$.
Die Möglichkeit einer solchen Abschätzung liefert der Satz von Taylor.

Dabei wird folgendes vorausgesetzt:

Die auf dem abgeschlossenen Intervall $[0; b]$ definierte Funktion f ist bei 0 n-mal und im Inneren des Definitionsintervalls $(n+1)$-mal differenzierbar und bei b stetig.

Dann sagt der Satz von Taylor folgendes aus:

Es gibt ein c aus $]0; b[$, für das gilt: $f(b) - p_n(b) = \frac{f^{(n+1)}(c)}{(n+1)!} \cdot b^{n+1}$.

Zum Beweis betrachte man die durch

$$h(x) = f(x) - p_n(x) - (f(b) - p_n(b)) \cdot \frac{x^{n+1}}{b^{n+1}}$$

definierte Hilfsfunktion h.

Diese erfüllt die Voraussetzungen des verallgemeinerten Satzes von Rolle, denn:

- Da zu f nur noch ganzrationale - also beliebig oft differenzierbare - Funktionen addiert werden, übertragen sich nach der Summenregel der Differentialrechnung die Differenzierbarkeits- und Stetigkeitsvoraussetzungen von f auf h.
- Für $i = 0, 1, 2, ..., n$ gilt $h^{(i)}(0) = 0$, denn nach der Definition des Taylorpolynoms ist für alle diese $f^{(i)}(0) = p_n^{(i)}(0)$. Und die i-te Ableitung der Potenzfunktion $(n+1)$-ten Grades enthält für i von 0 bis n den Faktor x^{n+1-i}, nimmt also an der Stelle $x = 0$ den Wert null an.
- $h(b) = f(b) - p_n(b) - (f(b) - p_n(b)) \cdot \frac{b^{n+1}}{b^{n+1}} = 0.$

Es gibt daher eine Stelle c im Intervall $]0; c[$ mit $h^{(n+1)}(c) = 0$.

Das bedeutet: $f^{(n+1)}(c) - p_n^{(n+1)}(c) - (f(b) - p_n(b)) \cdot \frac{(n+1)!}{b^{n+1}} = 0$

Da die n-te Ableitung einer ganzrationalen Funktion n-ten Grades eine Konstante ist, folgt

$$p_n^{(n+1)}(c) = 0; \qquad f^{(n+1)}(c) = (f(b) - p_n(b)) \cdot \frac{(n+1)!}{b^{n+1}}$$

Division durch $(n+1)!$ und Multiplikation mit b^{n+1} ergibt - nach Umordnen - die behauptete Gleichung.

9 Folgerungen, Anwendungen und Gegenbeispiele

9.1 Jede ganzrationale Funktion ist ihr eigenes Taylorpolynom

Zum Beweis betrachte man die Differenz zwischen Funktionswert und Wert des Taylorpolynoms an einer Stelle x; diese lässt sich nach dem Satz von Taylor als Vielfaches der $(n+1)$-ten Ableitung von f an einer Stelle c darstellen; die $(n+1)$-te Ableitung einer ganzrationalen Funktion n-ten Grades ist aber 0, woraus die Behauptung folgt.

9.2 Der binomische Satz gilt:

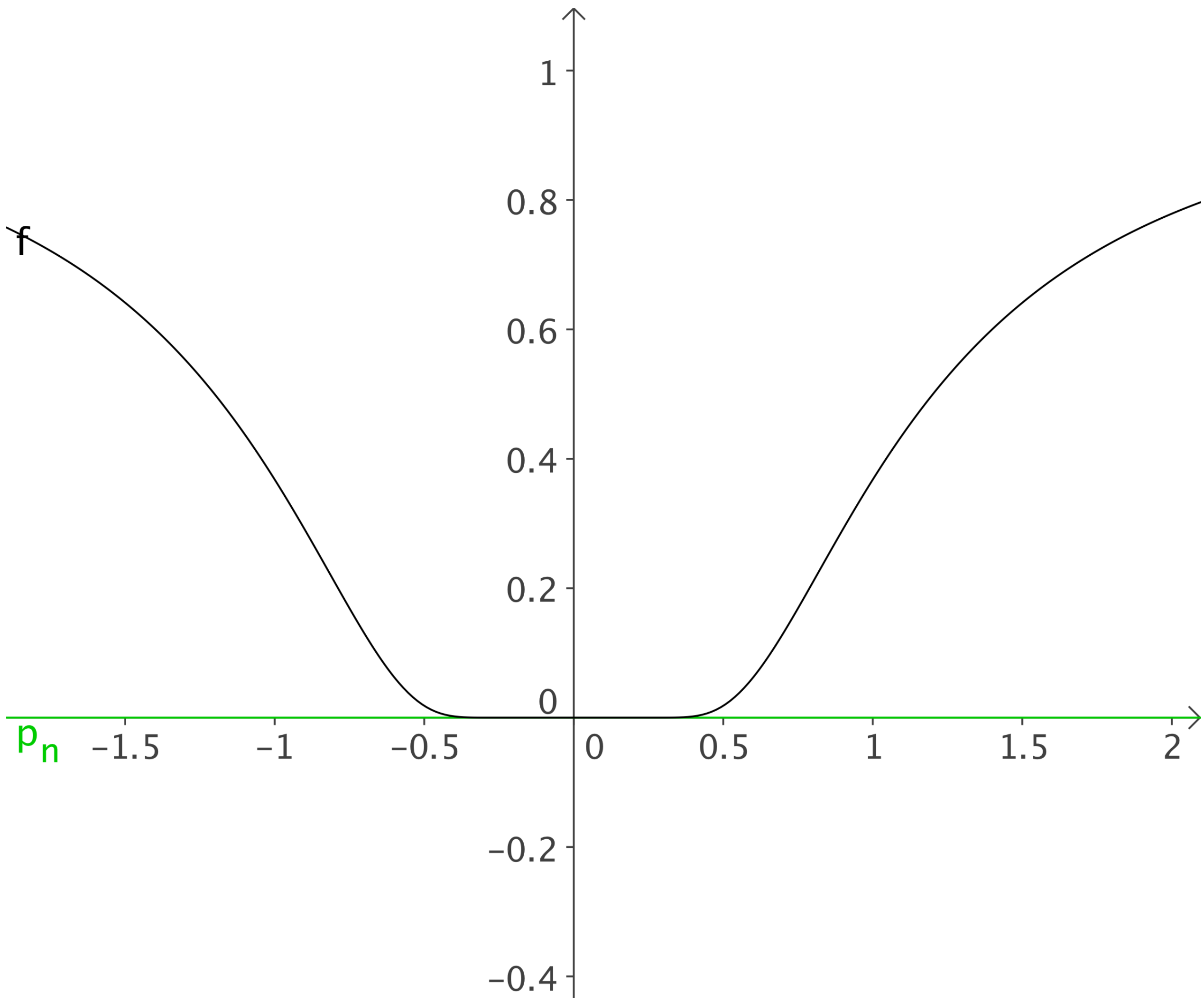
$(a + b)^n = \sum_{i=1}^{n} \binom{n}{i} a^i \cdot b^{n-i}$
Zum Beweis bestimme man das n-te Taylorpolynom $p_n(x)$ der Funktion f mit $f(x) = (a + x)^n$. Nach der ersten angegebenen Folgerung sind die Werte $f(x)$ und $p_n(x)$ für jedes x gleich. Wenn man diese Gleichung aufschreibt, steht (mit $x = b$) der binomische Satz da.

9.3 Approximierbarkeit einer speziellen Funktionenklasse

Nach einer bekannten Übungsaufgabe zur Analysis wird für jede reelle Zahl b und jedes reelle k der Wert des Ausdrucks $\frac{b^{n+1}}{(n+1)!} \cdot k$ beliebig klein, wenn man n hinreichend groß wählt. Hieraus folgt unmittelbar, dass sich für jede Funktion f die auf der Menge der reellen Zahlen beliebig oft differenzierbar ist und gleichmäßig beschränkte Ableitungen hat, für jede Stelle b die Differenz zwischen $f(b)$ und $p_n(b)$ beliebig klein machen lässt.
Dabei bedeutet gleichmäßige Beschränktheit der Ableitungen von f, dass eine reelle Zahl k mit der Eigenschaft existiert, dass für alle Stellen x und für alle Nummern n gilt $|f^{(n)}(x)| < k$.

9.4 Beispiel einer Funktion, die nicht durch ihr bei 0 entwickeltes Taylorpolynom approximiert wird

Die Vermutung, dass sich mithilfe der an der Stelle 0 entwickelten Schmiegepolynome Funktionswerte einer jeden unendlich oft differenzierbaren Funktion beliebig genau berechnen lassen, widerlegt das folgende Gegenbeispiel:
Die durch $f(0) = 0$ und $f(x) = \exp^{-\frac{1}{x^2}}$ für $x \neq 0$ definierte Funktion f ist überall unendlich oft differenzierbar, wobei für jede natürliche Zahl n gilt: $f^{(n)}(0) = 0$; daher ist $p_n(x) = 0$ für jedes x und jedes n .



10 Die Integralform des Satzes von Taylor

10.1 Satz von Lagrange

Voraussetzung: Es sei n eine natürliche Zahl und u, v auf einem abgeschlossenen Intervall $[a; b]$ (n+1)-mal differenzierbare Funktionen.

Behauptung:

a) $$\left[\sum_{i=0}^{n} (-1)^i u^{(n-i)} v^{(i)}\right]' = u^{(n+1)} v + (-1)^n u v^{(n+1)}$$

b) $$\left[\sum_{i=0}^{n}(-1)^i u^{(n-i)}v^{(i)}\right]_a^b = \int_a^b u^{(n+1)}(x)\cdot v(x)\mathrm{dx} + (-1)^n\int_a^b u(x)v^{(n+1)}(x)\mathrm{dx}$$

Beweis :

$$\begin{aligned}\left[\sum_{i=0}^{n}(-1)^i u^{(n-i)}v^{(i)}\right]' &= \sum_{i=0}^{n}(-1)^i u^{(n-i+1)}v^{(i)} + \sum_{i=0}^{n}(-1)^i u^{(n-i)}v^{(i+1)}\\ &= u^{(n+1)}v + \sum_{i=1}^{n}(-1)^i u^{(n-i+1)}v^{(i)} + \sum_{i=0}^{n-1}(-1)^i u^{(n-i)}v^{(i+1)} + (-1)^n uv^{(n+1)}\\ &= u^{(n+1)}v + \sum_{i=1}^{n}(-1)^i u^{(n-i+1)}v^{(i)} + \sum_{i=1}^{n}(-1)^{i-1} u^{(n-i+1)}v^{i} + (-1)^n uv^{(n+1)}\end{aligned}$$

Da die beiden Summen in der Mitte des letzten Ausdrucks entgegengesetzt gleich sind, folgt Teil a der Behauptung. Und aus dem Hauptsatz der Differential- und Integralrechnung ergibt sich damit die Richtigkeit von Teil b der Behauptung.

10.2 Die Integralform des Restglieds

Ist nun f eine über $[a,b]$ (n+1)-mal stetig differenzierbare Funktion und $u := f, v(x) := \frac{(b-x)^n}{n!}$, so gilt $(-1)^i v^{(i)} = \frac{(b-x)^{n-i}}{(n-i)!}$ für $i \in \{0,1,2,\ldots,n\}$. Für die Funktion v hat man:

$$(-1)^n v^{(n)}(x) \equiv 1, \quad v^{(n+1)}(x) \equiv 0, \quad v^{(i)}(b) = 0 \ (i = 1,2,3,\ldots,n-1)$$

Durch Einsetzen in Teil b des Satzes von Lagrange ergibt sich daraus

$$\begin{aligned}\left[\sum_{i=0}^{n} f^{(n-i)}(x)\frac{(b-x)^{n-i}}{(n-i)!}\right]_a^b &= \int_a^b f^{(n+1)}(x)\cdot\frac{(b-x)^n}{n!}\,\mathrm{dx}, \quad \text{also}\\ f(b) - \sum_{i=0}^{n} f^{(n-i)}(a)\,\frac{(b-x)^{n-i}}{(n-i)!} &= \int_a^b f^{(n+1)}(x)\cdot\frac{(b-x)^n}{n!}\,\mathrm{dx}\\ \text{und somit}\quad f(b) &= \sum_{i=0}^{n} f^{(i)}(a)\,\frac{(b-x)^i}{i!} + \int_a^b f^{(n+1)}(x)\cdot\frac{(b-x)^n}{n!}\,\mathrm{dx}\ .\end{aligned}$$

$\sum_{i=0}^{n}\frac{f^{(i)}(a)}{i!}(b-x)^i$ ist aber gerade der Wert des n-ten Schmiegepolynoms zu f, entwickelt an der Stelle a für das Argument b:

$$f(b) = p_n(b) + \int_a^b f^{(n+1)}(x)\cdot\frac{(b-x)^n}{n!}\,\mathrm{dx}, \quad \text{für variables } b:$$

$$f(x) = p_n(x) + \int_a^x f^{(n+1)}(t)\cdot\frac{(x-t)^n}{n!}\,\mathrm{dt} \qquad .$$

Als Restglied r_n $(= f(x) - p_n(x))$ ergibt sich damit: $r_n = \int_a^x f^{(n+1)}(t)\cdot\frac{(x-t)^n}{n!}\,\mathrm{dt}$.

10.3 Andere Restgliedformen

Auf die Integralform des Restglieds

$$r_n = \int_a^x f^{(n+1)}(t) \cdot \frac{(x-t)^n}{n!}\ \mathrm{dt} = \int_a^x f^{(n+1)}(t) \cdot \frac{(x-t)^{n-k}}{n!} \cdot (x-t)^k\ \mathrm{dt}$$

in der zuletzt dargestellten Form darf der Mittelwertsatz der Integralrechnung angewendet werden, da der Faktor $(x-t)^k$ im Integrationsintervall keinen Vorzeichenwechsel hat. Es gibt also (in Abhängigkeit von k) eine Stelle c im offenen Intervall $]a, x[$ mit der Eigenschaft

$$\begin{aligned}\int_a^x f^{(n+1)}(t) \cdot \frac{(x-t)^{n-k}}{n!} \cdot (x-t)^k\ \mathrm{dt} &= f^{(n+1)}(c) \cdot \frac{(x-c)^{n-k}}{n!} \cdot \int_a^x (x-t)^k\ \mathrm{dt} \\ &= f^{(n+1)}(c) \cdot \frac{(x-c)^{n-k}}{n!} \cdot \frac{(x-a)^{k+1}}{k+1}\end{aligned}$$

Je nach Wahl von k erhält man jetzt die verschiedenen Restgliedformen:

$$\begin{aligned}k = 0: \qquad r_n(x) &= f^{(n+1)}(c)\ \frac{(x-c)^n}{n!}(x-a) \quad \text{(Cauchy)} \\ k = n: \qquad r_n(x) &= f^{(n+1)}(c)\ \frac{(x-a)^{n+1}}{(n+1)!} \quad \text{(Lagrange)}\end{aligned}$$

Hinweis: Die Herleitung des Restglieds nach Lagrange unter Verwendung des verallgemeinerten Satzes von Rolle kommt, wie die Herleitung weiter oben zeigte, mit reduzierten Voraussetzungen aus: Die Stetigkeit der (n+1)-ten Ableitung von f an den Randstellen des betrachteten Intervalls $[a, b]$ braucht nicht vorausgesetzt zu werden.