Analysis für Informatiker und Ingenieure

SS 2022 - Vorlesung: 4-stündig, Übungen: 2-st., Tutorien: 2-st.

Übungsleiter: Lars von der Heide

Math-Lab: Matthias Rost

Tutoren: Anne Becker, Theo Nonnenmacher, Antonio Brkic

Hörerzahl: ca. 300

Literatur:

K. Kiyek, F. Schwarz, Mathematik für Informatiker 1

u.v.m.

Inhaltsverzeichnis

Kapitel I: Grundlegende arithmetische Techniken

§ 1 Das Prinzip der vollständigen Induktion

Dies ist eine Methode, um Aussagen über natürliche Zahlen ($\mathbb{N}$) zu beweisen. Sei dazu $A(n)$ $(n \in \mathbb{N})$ eine solche Aussage. Es möge gelten:

(1) $A(1)$ ist richtig (Induktionsanfang)

(2) Aus der Richtigkeit von $A(n)$ folgt die Richtigkeit von $A(n+1)$.
$\Big($kurz: $A(n) \Rightarrow A(n+1)\Big)$ (Induktionsschritt)

Dann gilt: $A(n)$ ist richtig $\forall\, n \in \mathbb{N}$.
(Dabei kürzen wir mit dem Zeichen $\forall$ die Sprechweise "für alle" ab.)

Symbolischer Vergleich mit dem Fall von Dominosteinen:

(1) besagt, dass ein erster Stein fallen muss;

(2) besagt, dass das Umfallen eines Steines den Fall des nächsten bewirkt.

Wenn **beide** Prinzipien greifen, fällt der ganze aufgebaute Domino-Parcours.

Beispiele:

(i) Behauptung: Die Summe der ersten n natürlichen Zahlen ist gleich

$$\frac{n(n+1)}{2}$$

(Das ist der "kleine Satz von Gauß"), also

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2} \quad \forall\, n \in \mathbb{N}. \quad \Big(\text{Das ist } A(n).\Big)$$

(1) $A(1): \quad 1 = \dfrac{1 \cdot (1+1)}{2}$. Das stimmt (war also nur zu verifizieren).

(2) Gelte also $A(n)$ (Induktionshypothese). Wir zeigen $A(n+1)$, also

$$1{+}2{+}3{+}\ldots{+}n{+}(n{+}1) = \frac{(n+1)(n+1+1)}{2} \quad \text{(Induktionsbehauptung)}$$

Dazu: $1{+}2{+}\ldots{+}n{+}(n{+}1) \underbrace{=}_{A(n)} \dfrac{n(n+1)}{2} + (n{+}1) = (n{+}1)\cdot(\dfrac{n}{2}{+}1) =$

$(n+1) \cdot \dfrac{n+2}{2}$ q.e.d.

Für Formeln dieser Art benutzen wir auch das Summenzeichen $\sum$ (großes griechisches S):

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

(ii) Wir zeigen die sog. **geometrische Summenformel:**

$$q \in \mathbb{R} \text{ oder } \mathbb{C},\ q \neq 1 \ \Rightarrow\ \sum_{k=0}^{n} q^k = \frac{q^{n+1}-1}{q-1} \quad \forall\, n \in \mathbb{N}$$

(wobei wir $q^0 := 1$ setzen).

(1) $A(1):\ 1+q = \dfrac{q^2-1}{q-1}$ stimmt nach der 3. Binomischen Formel

(2) Gelte also $A(n)$ $\Rightarrow \displaystyle\sum_{k=0}^{n+1} q^k = \sum_{k=0}^{n} q^k + q^{n+1} \underbrace{=}_{A(n)} \frac{q^{n+1}-1}{q-1} + q^{n+1} =$
$\dfrac{q^{n+1}-1+q^{n+2}-q^{n+1}}{q-1} = \dfrac{q^{n+2}-1}{q-1}$, also $A(n+1)$.

(iii) Behauptung: $2^n \geq n \ \ \forall\, n \in \mathbb{N}$

$n=1:\ 2^1 = 2 \geq 1$ stimmt.

$n \to n+1:\ 2^{n+1} = 2 \cdot 2^n \underbrace{\geq}_{A(n)} 2n = n+n \geq n+1$ q.e.d.

Bemerkung:

Der Induktionsanfang kann verschoben werden. Sei $n_0 \in \mathbb{N}$ oder auch in $\mathbb{Z}$ fest. Für die Aussage $A(n)$ gelte:

(1) $A(n_0)$ ist richtig

(2) $A(n)$ richtig ($n \geq n_0$) $\Rightarrow$ $A(n+1)$ richtig

Dann ist $A(n)$ richtig $\forall\, n \in \mathbb{Z},\ n \geq n_0$.

Beispiel:

$$2^n \geq n^2 \quad \forall\, n \geq 4, \ \text{ denn}$$

$n=4:\ 2^4 \geq 4^2$ stimmt

$n \to n+1:\ 2^{n+1} - (n+1)^2 = 2 \cdot 2^n - (n+1)^2 \underbrace{\geq}_{A(n)} 2 \cdot n^2 - (n+1)^2 =$

$n^2 - 2n - 1 = (n-1)^2 - 2 \underbrace{\geq}_{n \geq 4} 3^2 - 2 = 7 \geq 0.$

§ 2 Das Rechnen mit reellen Zahlen

Wir gehen in dieser Vorlesung mit einer gewissen Grundvertrautheit der Hörerschaft im Umgang mit den reellen Zahlen $(\mathbb{R})$ aus. Um Irritationen zu vermeiden, seien an dieser Stelle die wichtigsten Grundeigenschaften zusammengefasst:

A. Die Grundrechenarten

Es gibt im Zahlsystem $\mathbb{R}$ der reellen Zahlen zwei zweistellige Operationen, genannt plus $(+)$ und mal $(\,\cdot\,)$, die folgende Eigenschaften erfüllen:

1. Abgeschlossenheit: Beide Operationen führen nicht aus dem Bereich der reellen Zahlen hinaus:

   $$a, b \in \mathbb{R} \Rightarrow a + b \in \mathbb{R} \quad \text{und} \quad a \cdot b \in \mathbb{R}$$

2. Beide Operationen sind kommutativ, d.h.

   $$a + b = b + a \quad \text{und} \quad a \cdot b = b \cdot a \ \ \forall\, a, b \in \mathbb{R}$$

3. Beide Operationen sind assoziativ, d.h.

   $$(a + b) + c = a + (b + c) \quad \text{und} \quad (a \cdot b) \cdot c = a \cdot (b \cdot c) \ \ \forall\, a, b, c \in \mathbb{R}$$

   (Deshalb lässt man diese Klammern i.a. weg und - das ist ebenso Konvention - schreibt auch den Malpunkt nicht.)

4. Beide Operationen besitzen ein neutrales Element 0 bzw. $1\ (\neq 0)$:

   $$a + 0 = a \quad \text{und} \quad a \cdot 1 = a \ \ \forall\, a \in \mathbb{R}$$

5. Beide Operationen sind umkehrbar in folgendem Sinn:

   $$\forall\, a \in \mathbb{R} \ \exists\, -a \in \mathbb{R} \quad \text{mit} \quad a + (-a) = 0,$$

   und dann definiert man die Differenz

   $$a - b := a + (-b) \quad \text{für} \quad a, b \in \mathbb{R}.$$

   $$\forall\, a \in \mathbb{R},\ a \neq 0 \ \exists\, \frac{1}{a}\ (= a^{-1}) \quad \text{mit} \quad a \cdot \frac{1}{a} = 1,$$

   und dann definiert man den Quotienten

   $$\frac{a}{b} := a \cdot \frac{1}{b},\ (a, b \in \mathbb{R},\ b \neq 0.)$$

   Das Symbol $\exists$ steht für "es gibt" oder "es existiert".

6. Die Operationen sind verlinkt über das Distributivgesetz:

   $$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \ \ \forall\, a, b, c \in \mathbb{R},$$

   wobei man für die rechte Seite kurz $ab + ac$ schreibt ("Punkt-vor-Strich-Konvention").

Aus diesen Grundregeln lassen sich die anderen üblichen Rechenregeln herleiten, wie z.B. die Bruchrechenregeln

$$\frac{a}{b}+\frac{c}{d} = \frac{ad+bc}{bd}\ (b,d \neq 0),\quad \frac{a}{b}\cdot\frac{c}{d} = \frac{ac}{bd}\ (b,d \neq 0) \quad \text{und} \quad \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc},\quad (b,c,d \neq 0),$$

die Vorzeichenregeln

$$(-a)\cdot b = a\cdot(-b) = -(ab) \quad \text{und} \quad (-a)\cdot(-b) = ab.$$

Man kann kürzen und erweitern:

$$\frac{ac}{bc} = \frac{a}{b} \ \ \forall\, a,b,c \in \mathbb{R},\ b,c \neq 0,$$

und es gilt die häufig benutzte Regel

$$a\cdot b = 0 \iff a = 0 \quad \text{oder} \quad b = 0$$

Ungleichungen:

Die reellen Zahlen sind **angeordnet** in dem Sinn, dass es eine "Kleiner-Relation" $<$ gibt mit folgenden Eigenschaften:

1. Für $a,b \in \mathbb{R}$ gilt genau eine der Eigenschaften
$$a < b, \quad b < a, \quad a = b \quad \text{("Trichotonie")}$$
2. Für $a,b,c \in \mathbb{R}$ gilt:
$$a < b \quad \text{und} \quad b < c \ \Rightarrow\ a < c \quad \text{("Transitivität")}$$
3. Für $a,b,c \in \mathbb{R}$ gilt:
$$a < b \ \Rightarrow\ a+c < b+c \quad \text{("Monotonie bzgl. +")}$$
4. Für $a,b,c \in \mathbb{R}$ gilt:
$$a < b,\ 0 < c \ \Rightarrow\ ac < bc \quad \text{("Monotonie bzgl. } \cdot \text{")}$$

Daraus leitet man dann weitere Bezeichnungen und Rechenregeln ab:

(i) Statt $a < b$ schreibt man auch $b > a$ (b ist größer als a).

$a \leq b$ bzw. $b \geq a$ steht für $a < b$ oder $a = b$ (kleiner gleich bzw. größer gleich);

a positiv heißt $a > 0$; a negativ steht für $a < 0$.

(ii) $a > 0,\ b > 0 \ \Rightarrow\ ab > 0$

$a > 0,\ b < 0$ oder $a < 0,\ b > 0 \ \Rightarrow\ ab < 0$

$a < 0,\ b < 0 \ \Rightarrow\ ab > 0$

(iii) $a \neq 0 \ \Rightarrow\ a^2 > 0$ (folgt aus (ii) denn $a^2 = a\cdot a = (-a)\cdot(-a)$)

Insbesondere ist $1 = 1^2 > 0$

(iv) $a < b,\ c < 0 \ \Rightarrow\ ac > bc$ (Achtung: beliebte Fehlerquelle)

Schließlich führt man noch den **(Absolut-)Betrag** einer Zahl ein:

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

Für $|a|$ gelten dann die folgenden Rechenregeln:

(i) $|a| = |-a|$ stets; $|a| \geq 0$ stets und $|a| = 0 \iff a = 0$

(ii) $|ab| = |a| \cdot |b| \;\; \forall\, a, b \in \mathbb{R}$

(iii) $|a + b| \leq |a| + |b| \;\; \forall\, a, b \in \mathbb{R}$ und
$|a + b| \geq |a| - |b|$ sowie $|a + b| \geq |b| - |a|$ stets

Als geometrisches Modell führt man gerne den **Zahlenstrahl** ein. Das ist eine orientierte skalierte Gerade, meist horizontal gedacht.

Ein ausgezeichneter Punkt ist der Nullpunkt 0.

Alle Zahlen rechts von 0 sind positiv, alle links von 0 sind negativ.

a und $-a$ liegen spiegelbildlich bzgl. 0.

$|a|$ ist der Abstand von a zu 0, $|a - b|$ ist der Abstand zwischen a und b .

Benutzt man die Betragsfunktion, so hat man es streng genommen mit Fallunterscheidungen zu tun. Das ist nicht immer ganz einfach. Wir illustrieren dies an einem

Bespiel:

Welche reellen Zahlen $x \in \mathbb{R}$ erfüllen die Ungleichung

$$(*) \;\; |x - 1| + |x + 1| < 4 \;?$$

Zum Auflösen unterscheiden wir $x \leq -1$, $-1 < x < 1$ und $x \geq 1$:

- Fall 1: $x \leq -1$: Dann ist (*) äquivalent zu

  $$-x + 1 - x - 1 < 4 \iff -2x < 4 \iff x > -2$$

  Das sind dann insgesamt die Zahlen $x \in \mathbb{R}$ mit $-2 < x \leq -1$.

- Fall 2: $-1 < x < 1$: Dann ist (*) äquivalent zu

  $$-x + 1 + x + 1 < 4 \iff 2 < 4$$

  Das stimmt immer. Also erfüllen alle x mit $-1 < x < 1$ die Ungleichung (*).

- Fall 3: $x \geq 1$: Dann ist (*) äquivalent zu

  $$x - 1 + x + 1 < 4 \iff 2x < 4 \iff x < 2$$

  Das sind dann insgesamt die Zahlen $x \in \mathbb{R}$ mit $1 \leq x < 2$.

Zusammengefasst ergibt sich:

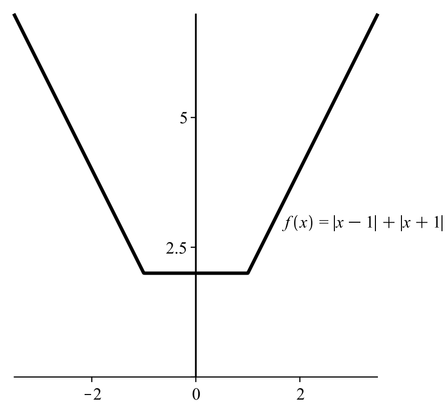
Es sind genau die Zahlen $x \in \mathbb{R}$ mit

$$-2 < x < 2,$$

welche diese Ungleichung erfüllen, also lösen.

Diese Menge nennen wir das (offene) Intervall $(-2, 2)$.

Zur Illustration noch der Graph der Funktion $f(x) = |x-1| + |x+1|$:



Bemerkung:

Gelegentlich erweitert man die Menge der reellen Zahlen zur Menge der $\mathbb{C}$ **komplexen** Zahlen. Hier führt man eine "Zahl" i ein mit der Eigenschaft

$$i^2 = -1.$$

die sog. **imaginäre Einheit.**

Eine komplexe Zahl z ist dann ein Gebilde der Bauart

$$z = x + iy \quad \text{mit} \quad x, y \in \mathbb{R}.$$

Wir nennen dann x den **Realteil** von z, i.Z. $x = \operatorname{Re} z$.

y heißt der **Imaginärteil** von z, i.Z. $y = \operatorname{Im} z$.

Das Rechnen in $\mathbb{C}$ erfolgt dann wie üblich mit der Konvention $i^2 = -1$, also:

$$z = x_1 + iy_1,\ w = x_2 + iy_2 \ \ (x_j, y_j \in \mathbb{R},\ j = 1, 2) \ \Rightarrow$$
$$z \pm w = (x_1 \pm x_2) + i(y_1 \pm y_2)$$
$$z \cdot w = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

Beim Dividieren behilft man sich mit einem "Trick":

$$\frac{1}{x+iy} = \frac{x-iy}{(x+iy)(x-iy)} = \frac{x}{x^2+y^2} + i\,\frac{-y}{x^2+y^2},$$

d.h. man erweitert mit

$$\bar{z} := x - iy,$$

der sogenannten **konjugiert komplexen Zahl**, und beachtet die dritte Binomische Formel $(a+b)(a-b) = a^2 - b^2$ und $i^2 = -1$.

Die Rechenregeln der Grundrechenarten in $\mathbb{R}$ bleiben komplett erhalten.

Eine Anordnung wie in $\mathbb{R}$ gibt es nicht (und kann es auch nicht geben: denn dann wäre $1 = 1^2 > 0$, aber auch $-1 = i^2 > 0$, was mit der Trichotonie nicht vereinbar wäre).

Eine geometrische Veranschaulichung analog zum Zahlenstrahl ist durch die "Gaußsche Zahlenebene" gegeben. Dabei identifiziert man die Zahl

$$z = x + iy$$

mit dem Punkt

$$(x, y) \in \mathbb{R}^2.$$

Die Addition entspricht dann der vektoriellen Addition im $\mathbb{R}^2$;

$-z$ liegt spiegelbildlich zu z bzgl. des Ursprungs;

$\bar{z}$ liegt spiegelbildlich zu z bzgl. der reellen Achse.

Schließlich definiert man noch den Betrag einer komplexen Zahl

$$|z| := \sqrt{x^2 + y^2} \; (= \sqrt{z\bar{z}}).$$

Die Rechenregeln (i)-(iii) für den reellen Betrag übertragen sich auch auf den komplexen Betrag.

Geometrisch ist wieder $|z|$ der Abstand von z zum Nullpunkt,

$|z - w|$ der Abstand zwischen den Zahlen z und w.

Wir werden in der Analysis die komplexen Zahlen eher selten einsetzen.

C. Vollständigkeit

Theoretisch könnte man sich mit dem System $\mathbb{Q}$ der rationalen Zahlen begnügen: x rational heißt dabei: x ist von der Form

$$x = \frac{p}{q} \quad \text{mit} \quad p \in \mathbb{Z},\ q \in \mathbb{N}.$$

Die Grundrechenarten und die Anordnung in A. und B. wären davon nicht betroffen. Das daraus resultierende Problem besteht aber darin, dass es Lücken im Zahlenstrahl gibt, wenn man nur rationale Zahlen zulässt. Das populärste Beispiel ist etwa der folgende

Satz:

Es gibt keine rationale Zahl x mit $x^2 = 2$ (m.a.W.: $\sqrt{2}$ ist irrational.)

Der Beweis hierfür gehört zur mathematischen Allgemeinbildung. Deswegen führen wir ihn auch:

Wir nehmen an, es gebe eine rationale Zahl x mit $x^2 = 2$, etwa

$$x = \frac{p}{q} \quad \text{mit} \quad p \in \mathbb{Z},\ q \in \mathbb{N}.$$

Wir nehmen an, x sei in durchgekürzter Form gegeben ("o.B.d.A.")

$$
\begin{aligned}
x^2 &= 2 \\
\Rightarrow \ \frac{p^2}{q^2} &= 2 \\
\Rightarrow \ p^2 &= 2q^2 \\
\Rightarrow \ p^2 & \quad \text{ist gerade} \\
\Rightarrow \ p & \quad \text{ist gerade}
\end{aligned}
$$

(denn das Quadrat einer ungeraden Zahl ist immer ungerade), also etwa

$$
\begin{aligned}
p = 2r & \quad \text{mit} \quad r \in \mathbb{Z} \\
\Rightarrow \ (2r)^2 &= 2q^2 \\
\Rightarrow \ q^2 &= 2r^2 \\
\Rightarrow \ q^2 & \quad \text{ist gerade} \\
\underbrace{\Rightarrow}_{\text{wie oben}} \ q & \quad \text{ist gerade.}
\end{aligned}
$$

Also kann man in $\frac{p}{q}$ doch kürzen (mit 2) im Widerspruch zu unserer Annahme. Dies ist nur dann kein Widerspruch, wenn man die Annahme verwirft, dass es solche Zahlen p und q gibt.

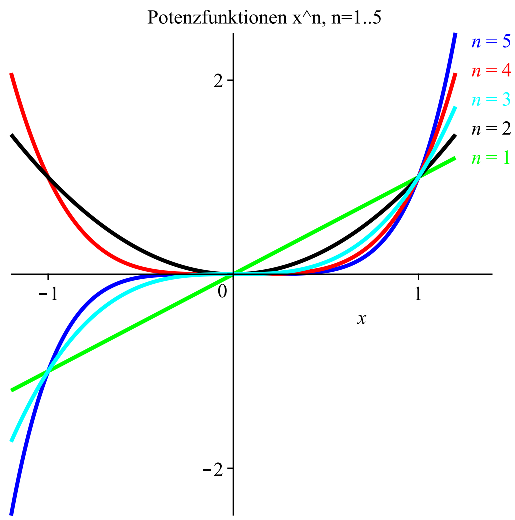
Wie man das Zahlsystem vervollständigt so, dass keine Lücken mehr im Zahlenstrahl vorhanden sind, werden wir später sehen.

§ 3 Potenzen, Wurzeln, Polynome und rationale Funktionen

A. Potenzen

Das sind die Funktionen $y = f(x) = x^n,\ n \in \mathbb{N}$ fest.

Wir skizzieren zunächst die Graphen für $n = 1, \ldots, 5$:



Man sieht:

Die Graphen sind

- **achsensymmetrisch** bzgl. der $y-$Achse, wenn n gerade ist und
- **punktsymmetrisch** bzgl. 0, wenn n ungerade ist.

Wir wiederholen die Potenzrechenregeln:

$$x^m x^n = x^{m+n}; \quad \frac{x^m}{x^n} = x^{m-n};$$

$$\left(x^m\right)^n = x^{mn}$$

$$(xy)^m = x^m y^m; \quad \left(\frac{x}{y}\right)^m = \frac{x^m}{y^m}$$

($x, y \in \mathbb{R}$ oder auch $\mathbb{C}$. Auftretende Nenner seien $\neq 0$; $m, n \in \mathbb{N}$.)

Die Rechenregeln gelten auch für $m, n \in \mathbb{Z}$, wenn man setzt:

$$x^0 := 1; \quad x^{-n} := \frac{1}{x^n} \quad \text{für} \quad n \in \mathbb{N}.$$

Frage: Gibt es auch eine Umformungsmöglichkeit für $(x+y)^n$ $(n \in \mathbb{N}_0)$?

$n=0:\ (x+y)^0 = 1$

$n=1:\ (x+y)^1 = 1 \cdot x + 1 \cdot y$

$n=2:\ (x+y)^2 = 1 \cdot x^2 + 2 \cdot xy + 1 \cdot y^2$

$n=3:\ (x+y)^3 = (x^2+2xy+y^2)(x+y) = 1 \cdot x^3 + 3 \cdot x^2y + 3 \cdot xy^2 + 1 \cdot y^3$

$n=4:\ (x+y)^4 = \ldots = 1 \cdot x^4 + 4 \cdot x^3y^1 + 6 \cdot x^2y^2 + 4x^1y^3 + 1 \cdot y^4$

usw.

Offenbar entsteht eine Summe aus Produkten $c \cdot x^k y^l$, wobei die Summe der Exponenten $= n$ ist: $k+l=n$ und die Vorfaktoren c in folgendem Schema angeordnet werden können:

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
 & & & & & \ldots & & & & &
\end{array}
$$

Das ist das sog. **Pascalsche Dreieck**. Die auftretenden Zahlen

$$
\begin{array}{ccccccc}
 & & & a_{00} & & & \\
 & & a_{10} & & a_{11} & & \\
 & a_{20} & & a_{21} & & a_{22} & \\
a_{30} & & a_{31} & & a_{32} & & a_{33} \\
 & & & \ldots & & &
\end{array}
$$

unterliegen offenbar folgendem Bildungsgesetz:

$$a_{n,0} = a_{n,n} = 1; \quad a_{n+1,k} = a_{n,k-1} + a_{n,k},\ k = 1, \ldots, n$$

Satz:

Das Bildungsgesetz im Pascalschen Dreieck leisten die Zahlen

$$\binom{n}{k} := \frac{n(n-1)(n-2)\ldots(n-k+1)}{1 \cdot 2 \cdot \ldots \cdot k} \quad \text{bei} \quad \binom{n}{0} := 1,$$

die sogenannten **Binomialkoeffizienten** n über k.

Beweis:

$$a_{n,k-1} + a_{n,k} = \frac{n(n-1)\cdot(n-k)}{1 \cdot 2 \cdot \ldots \cdot (k-1)} + \frac{n(n-1)(n-2)\ldots(n-k+1)}{1 \cdot 2 \cdot \ldots \cdot k} =$$

$$\frac{n(n-1)\ldots(n-k)(k+n-k+1)}{1 \cdot 2 \cdot \ldots \cdot k} = \frac{(n+1)n(n-1)(n-2)\ldots(n-k)}{1 \cdot 2 \cdot \ldots \cdot k} =$$

$$\binom{n+1}{k} \quad \text{q.e.d.}$$

Bemerkungen und Beispiele:

(i) Der Nenner $1 \cdot 2 \cdot k =: k!$ heißt $k-$ **Fakultät**.

Diese Zahlen werden sehr schnell sehr groß, z.B. $7! = 5040$

Der TR-Overflow liegt zwischen $n = 69$ und $n = 70$.

(ii) Die Formel für $\binom{n}{k}$ ist leicht zu merken: Zähler und Nenner bestehen aus jeweils k Faktoren, im Nenner aufsteigend, beginnend mit 1, im Zähler abfallend, beginnend mit n, also z.B.

$$\binom{7}{3} = \frac{7 \cdot 6 \cdot 5}{1 \cdot 2 \cdot 3} = 35$$

(iii) Erweitert man in $\binom{n}{k}$ mit $(n-k)!$, so sieht man, dass $\binom{n}{k}$ komplett mit Fakultäten geschrieben werden kann:

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

woraus man die Symmetrie im Pascalschen Dreieck ablesen kann:

$$\binom{n}{k} = \binom{n}{n-k}$$

(iv) $\binom{n}{k}$ gibt auch an, wieviele Möglichkeiten man hat, aus n Objekten k ($\leq n$) verschiedene auszuwählen. Daher kommt auch der englische Name für $\binom{n}{k}$: (from) n choose k, z.B.

- Anzahl Skatrunden (zu je 3 Personen) aus 8 Personen:

$$\binom{8}{3} = \frac{8 \cdot 7 \cdot 6}{1 \cdot 2 \cdot 3} = 56$$

- Klassiker (Lotto): 6 aus 49 Zahlen sind auszuwählen. Die Anzahl der Varianten ist dann

$$\binom{49}{6} = \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}{6!} = 13983816 \quad \text{Möglichkeiten.}$$

(v) Die Rechenregel für $(x+y)^n$ lautet also

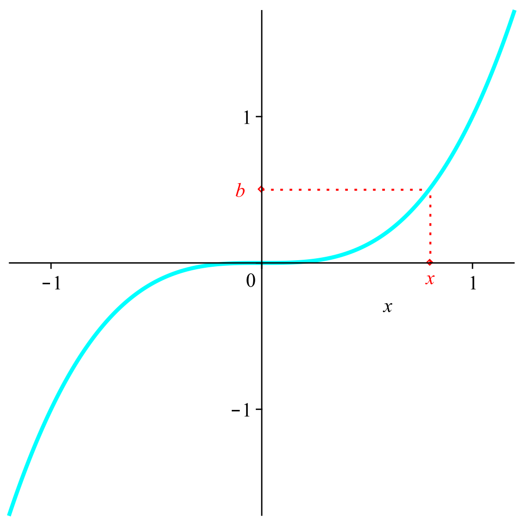
$$(x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \ldots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n = \sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k.$$

Das ist der sogenannte **Binomische Lehrsatz.**

B. Wurzeln

Wir wollen nun Gleichungen der Form $x^n = b$ ($n \in \mathbb{N}$, $b \in \mathbb{R}$ fest) lösen. Aus den Graphen aus A. ist ersichtlich:

(a) Ist n ungerade, so ist liest man an folgender Skizze das Ergebnis ab:



$x^n = b$ ist für **jedes** b **eindeutig** lösbar, durch die sog. $n-$**te Wurzel** aus b:

$$x = \sqrt[n]{b}$$

Die Existenz liegt dabei an der Stetigkeit von x^n: keine Lücken im Bildbereich.

Die Eindeutigkeit liegt an der strengen Monotonie $\to$ Begriffe später).

Beispiel:

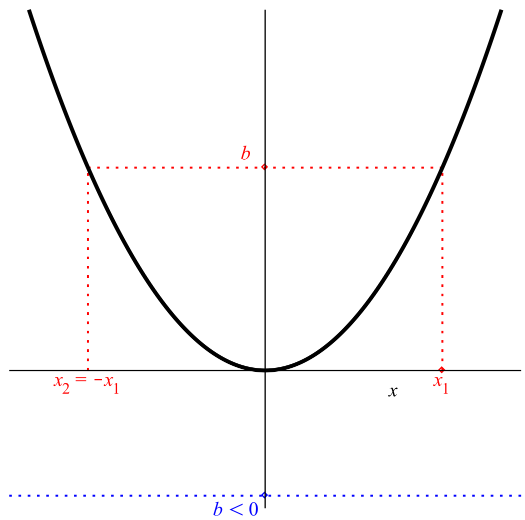
$$x^3 = 27 \iff x = \sqrt[3]{27} = 3$$

Es ist dann

$$\left(\sqrt[n]{b}\right)^n = b \ \ \forall\, b \in \mathbb{R} \quad \text{und} \quad \sqrt[n]{x^n} = x \ \ \forall\, x \in \mathbb{R}:$$

x^n und $\sqrt[n]{x}$ sind Umkehrfunktionen zueinander.

(b) Ist n gerade, so betrachten wir für die Gleichung $x^n = b$ die folgende Skizze:



Dies zeigt: Die Gleichung $x^n = b$ besitzt

- im Falle $b < 0$ keine Lösung,
- im Falle $b = 0$ die "doppelte" Lösung $x = 0$ und
- im Falle $b > 0$ zwei Lösungen, eine positive und eine negative, die auch symmetrisch zur $y-$Achse liegen.

  Die positive heißt dann wieder $n-$te Wurzel aus b, $(\sqrt[n]{b})$, die zweite ist dann $-\sqrt[n]{b}$. Das festzulegen, ist reine Konvention. Hier gilt
  $$\left(\sqrt[n]{b}\right)^n = b \;\; \forall\, b \geq 0 \quad \text{und} \quad \sqrt[n]{x^n} = |x| \;\; \forall\, x \in \mathbb{R}$$

Also könnte man korrekterweise sagen:

- $\sqrt[4]{16} = 2$
- Die Gleichung $x^6 = 64$ wird gelöst von $x_1 = \sqrt[6]{64} = 2$ und $x_2 = -x_1 = -2$.
- $\sqrt[3]{-8} = -2$
- $\sqrt[4]{-8}$ existiert nicht (im Reellen).

Die Potenzgesetze bleiben gültig, wenn man setzt:

$$\sqrt[n]{x} := x^{\frac{1}{n}},\; n \in \mathbb{N},\; x \in \mathbb{R} \quad \text{bzw.} \quad x \geq 0$$

C. Polynome

oder auch ganzrationale Funktionen genannt sind Funktionen vom Typ

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$

mit gegebenen Zahlen $a_0, \ldots, a_n$ **(”Koeffizienten”)**; n heißt der **Grad** von $p(x)$, wenn $a_n \neq 0$ ist.

Oft treten polynomiale Gleichungen $p(x) = 0$, (also Nullstellen von $p(x)$) auf. Einige Anmerkungen dazu:

(i) $n = 1: \; a_1 x + a_0 = 0 \;\Rightarrow\; x = -\dfrac{a_0}{a_1}$ (trivial)

(ii) $n = 2: \; a_2 x^2 + a_1 x + a_0 = 0 \;\Rightarrow\; x_{1/2} = \dfrac{1}{2a_2}\Big(- a_1 \pm \sqrt{a_1^2 - 4a_0 a_2}\Big)$

(”Mitternachtsformel”) mit den verschiedenen Interpretationen je nach Vorzeichen des Radikanden.

(iii) $n = 3, 4$: Es gibt ”Mitternachtsformeln” . Diese sind schon lange bekannt (Cardano, Tartaglia, Ferrari im 16. Jh.) , sind allerdings sehr kompliziert und werden kaum benutzt

(iv) $n \geq 5$: Es gibt keine allgemeine Lösungsformel (Galois, Abel, Anfang des 19. Jh.).

In Spezialfällen ist u.U. eine Lösung auf dem Papier möglich.

(v) Gelegentlich gelingt eine Reduktion durch Substitution, z.B.

- $x^4 - 4x^2 + 3 = 0$:

  Die Substitution $x^2 = u$ liefert als Hilfsgleichung

  $$u^2 - 4u + 3 = 0 \quad \text{mit den Lösungen} \quad u_1 = 3, \; u_2 = 1.$$

  Resubstitution $x^2 = 3$ bzw. $x^2 = 1$ liefert dann die 4 Lösungen

  $$x_{1/2} = \pm\sqrt{3}, \; x_{3/4} = \pm 1.$$

- $x^{12} - x^6 - 2 = 0$:

  Die Substitution $x^6 = u$ liefert die Hilfsgleichung

  $$u^2 - u - 2 = 0 \quad \text{mit den Lösungen} \quad u_1 = 2, \; u_2 = -1.$$

  Resubstitution führt dann bei $u_1 = 2$ zu den Lösungen

  $x_{1/2} = \pm\sqrt[6]{2}$, bei $u_1 = -1$ zu keinen weiteren (reellen) Lösungen.

(vi) Hat man in $p(x) = 0$ eine Lösung x_1 gefunden (Raten, Spicken, etc.), so kann man bekanntlich einen ”Linearfaktor” abdividieren (Polynomdivision: Algorithmus wie beim schriftlichen Dividieren ganzer Zahlen), d.h. es ist

$$p(x) = (x - x_1)\tilde{p}(x),$$

wobei $\tilde{p}(x)$ ein Polynom vom Grad $n-1$ ist. Die Lösungen von $p(x)=0$ sind dann x_1 und die Lösungen von $\tilde{p}(x)=0$.

Beispiel: $x^3-5x+2=0:\ \ x_1=2$ geraten

$$x^3-5x+2 \underbrace{=}_{\text{Div.}} (x-2)(x^2+2x-1) \;\Rightarrow\; x_{2/3}=\frac{-2\pm\sqrt{8}}{2}=-1\pm\sqrt{2}.$$

(vii) Tipp zum Raten: Liegt $p(x)$ vor in der Form

$$p(x)=1\cdot x^n+\ldots+a_0$$

mit <u>ganzzahligen</u> Koeffizienten $a_0,\ \ldots,\ a_{n-1}$, so kommen als ganzzahlige Lösungen von $p(x)=0$ <u>nur</u> die Teiler von a_0 in Frage. Das ist eine Konsequenz aus dem sogenannten Wurzelsatz von Vieta aus der Algebra.

Beispiel: $x^5+x^4-9x^3-9x^2+20x+20=0:$

$a_0=20$ besitzt die Teiler $\pm 1,\ \pm 2,\ \pm 4,\ \pm 5,\ \pm 10,\ \pm 20$.

Durchchecken liefert die drei Lösungen $x_1=-1,\ x_2=2,\ x_3=-2$. Polynomdivision (entweder dreimal mit je einem Linearfaktor oder einmal durch das Produkt $(x+1)(x-2)(x+2)=x^3+x^2-4x-4$ liefert

$$p(x)=(x+1)(x-2)(x+2)(x^2-5)$$

mit den weiteren Lösungen $x_{4/5}=\pm\sqrt{5}$.

(viii) Wenn alles nichts hilft, bleiben nur noch numerische Näherungsverfahren (z.B. Newton-Methode $\rightarrow$ später)

D. Rationale Funktionen

Das sind Funktionen vom Typ $f(x)=\dfrac{p(x)}{q(x)}$, wobei $p(x)$ und $q(x)$

Polynome sind. Hier ist zu beachten, dass der Definitionsbereich i.a. eingeschränkt ist auf $\mathbb{R}$ ohne die Nullstellen von $q(x)$.

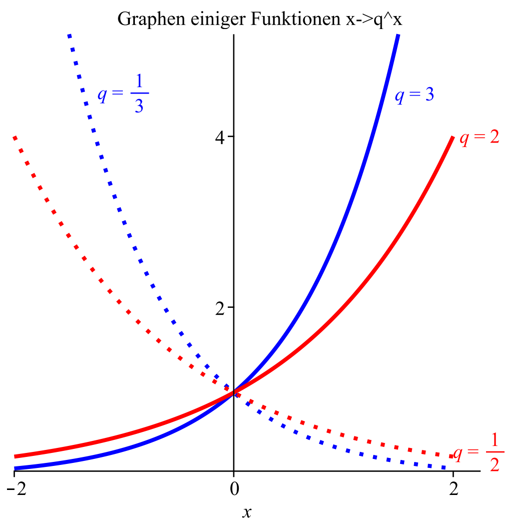
§ 4 Exponentialfunktionen

Dies sind die Funktionen vom Typ

$$f(x) = q^x$$

mit einer festen **Basiszahl** $q > 0, \ q \neq 1$.

Wir skizzieren die Graphen für $q = 2, \ 3, \ \frac{1}{2}, \ \frac{1}{3}$:



Offensichtlich gilt: $f(x)$ ist

- streng monoton wachsend, wenn $q > 1$ gilt und
- streng monoton fallend für $0 < q < 1$.

Diese Funktionen beschreiben im Falle $q > 1$ Prozesse mit konstanten Wachstumsraten etwa beim

- Zinseszins,
- BIP-Wachstum (im Idealfall) oder
- Populationsentwicklungen, etc.

bzw. im Fall $0 < q < 1$ Prozesse mit konstanten Zerfallsraten, etwa beim

- radioaktiven Zerfall oder bei
- degressiver Abschreibung etc.

Eine jährliche Zunahme um jeweils $3\ \%$ wird dann durch die Funktion 1.03^x beschrieben (x demnach in Jahren gemessen),

eine jährliche Abnahme um $3\ \%$ durch die Funktion 0.97^x.

Ein interessantes Phänomen ist, dass die Zeiträume, in denen sich bei exponentiellen Prozessen die Funktionswerte von q^x verdoppeln (bei $q > 1$) bzw. halbieren (bei $q < 1$), **konstant** sind:

$$q > 1 : \quad q^{x+\Delta x} = 2q^x \iff q^x \cdot q^{\Delta x} = 2q^x \iff q^{\Delta x} = 2.$$

Δx hängt also nicht von x ab. Wie man Δx berechnet, sehen wir in § 5.

Man kennt diesen Sachverhalt, wenn man bei einer Populationsentwicklung vom Verdoppelungszeitraum spricht oder bei radioaktiven Substanzen von der Halbwertszeit.

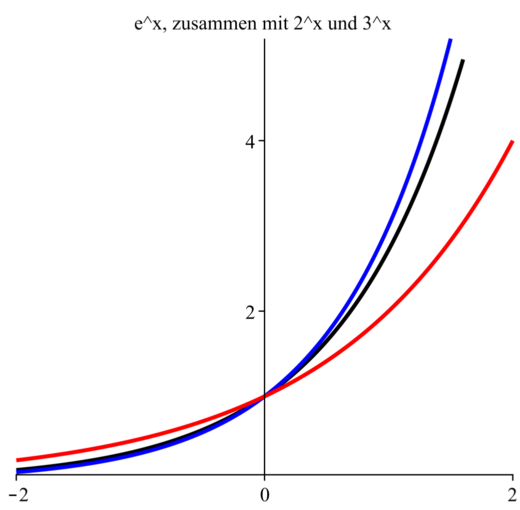
<u>Zur Zahl e:</u>

Wir werden sehen, dass man nur eine Basiszahl $q > 0$ benötigt, um alle Exponentialfunktionen beschreiben zu können. Lange Zeit war es üblich, die Basiszahl $q = 10$ zu wählen, was durch die Verwendung des Dezimalsystems gerechtfertigt werden kann. Genauso könnte man sich am Binärsystem orientieren und als Basiszahl $q = 2$ wählen. Allgemein üblich ist heute die Verwendung der **"Eulerschen Zahl"** e, die man über einen Grenzwert definiert:

$$e := \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = 2,71828\ldots$$

Zum Begriff des Grenzwerts kommen wir später.

Wir zeichnen nun noch den Graphen von e^x ein:



§ 5 Logarithmen

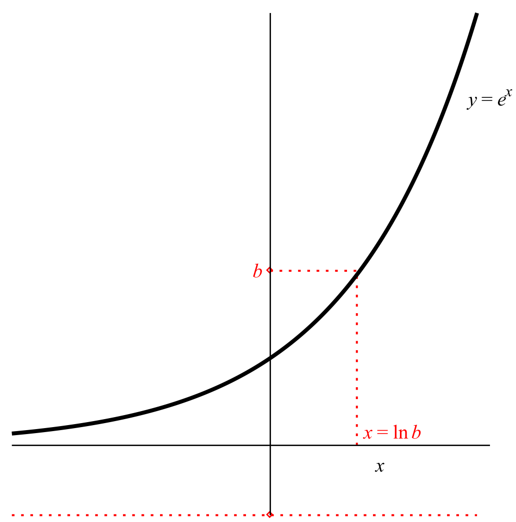
Hier stellt sich weder die Umkehrfrage: Wir wollen für gegebenes $b \in \mathbb{R}$ die Gleichung

$$e^x = b$$

lösen.

Mit Blick auf den Graphen von e^x sehen wir: Diese Gleichung besitzt

- im Falle $b \leq 0$ keine Lösung und
- im Falle $b > 0$ genau eine Lösung, die wir analog zu Wurzeln zeichnerisch gewinnen können (s.Skizze):



Diese eindeutige Lösung nennen wir den **(natürlichen) Logarithmus** von b :

$$x = \ln b.$$

Die Wortherkunft ist dabei ziemlich nichtssagend: es stecken die griechischen Worte 'logos' (Verhältnis, Berechnung) und' arithmos' (Zahl) drin.

Es ist also

$$\ln(e^x) = x \quad \forall \;\; x \in \mathbb{R} \quad \text{und} \quad e^{\ln b} = b \quad \forall \;\; b > 0$$

Akademischer formuliert: Die Funktionen

$$e^x : \mathbb{R} \to (0, \infty)$$

und

$$\ln x : (0, \infty) \to R$$

sind Umkehrfunktionen zueinander.

Grundsätzlich gilt alles bisher Gesagte auch für jede der Funktionen

$$q^x : \mathbb{R} \to (0, \infty), \ (q > 0, \ q \neq 1) :$$

$$q^x = b \iff x = \log_q b$$

(Logarithmus zur Basis q). Es genügt allerdings, mit einer Basis, etwa e zu arbeiten:

$$q^x = \left(e^{\ln q}\right)^x \underbrace{=}_{\text{Rechenregeln}} e^{x \ln q}$$

(wobei wir stillschweigend voraussetzen, dass die Potenzregeln auch für reelle Exponenten gelten). Es ist also

$$q^x = b \iff e^{x \ln q} = b \iff x \ln q = \ln b \iff x = \frac{\ln b}{\ln q},$$

also

$$\log_q b = \frac{\ln b}{\ln q}.$$

Mit den Logarithmen ist es also wie vielfach im Leben: Kennt man einen, so kennt man alle.

Die Rechenregeln für Logarithmen ergeben sich aus den Potenzgesetzen:

Setzt man etwa $\ln x = u, \ \ln y = v,$ also $x = e^u, \ y = e^v,$ so folgt:

$x \cdot y = e^u \cdot e^v = e^{u+v} \ \Rightarrow \ u + v = \ln(xy)$

$\dfrac{x}{y} = \dfrac{e^u}{e^v} = e^{u-v} \ \Rightarrow \ u - v = \ln\left(\dfrac{x}{y}\right)$

$x^\alpha = \left(e^{lnx}\right)^\alpha = e^{\alpha \ln x} \ \Rightarrow \ln(x^\alpha) = \alpha \ \ln x$

Diese Rechnungen kann man mit jeder anderen Basis durchführen.

Fazit:

Für jeden Logarithmus $\log$ gilt:

(i) $\log(xy) = \log x + \log y \quad \forall \ x, y > 0$

(ii) $\log \dfrac{x}{y} = \log x - \log y \quad \forall \ x, y > 0, \ y \neq 1$

(iii) $\log\left(x^\alpha\right) = \alpha \ \log x \quad \forall \ x > 0, \ \alpha \in \mathbb{R}.$

Beachte: Es gibt keine elementare Umformungsmöglichkeit für $\log(x \pm y)$!!

Beispiele:

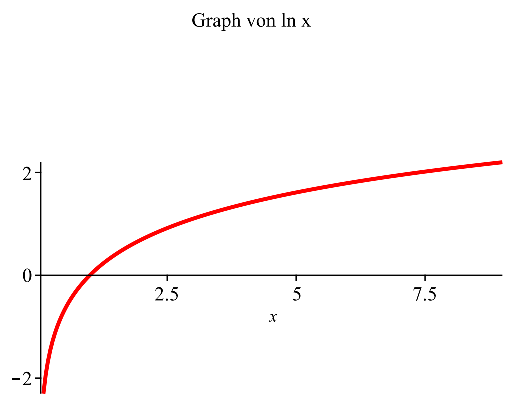
(i) $2^x = 64 \iff x = \dfrac{\ln 64}{\ln 2} = 6$ und

$2^x = 7 \iff x = \dfrac{\ln 7}{\ln 2} = 2.81$

(ii) $3^x + 5 = 17 \iff 3^x = 12 \iff x = \dfrac{\ln 12}{\ln 3} = 2.26$

(iii) $4 \cdot 7^x = 3 \iff 7^x = \dfrac{3}{4} \iff x = \dfrac{\ln(3/4)}{\ln 7} = -0.15$

(iv) Graph von $\ln x$



(v) Verdoppelungszeitraum bei 2 % jährlichem Wachstum:

$$1.02^{\Delta x} = 2 \iff \Delta x = \frac{\ln 2}{\ln 1.02} = 35 \text{ (Jahre)}$$

(vi) Halbwertszeit bei 10 % jährlichem Schwund:

$$0.9^{\Delta x} = 0.5 \iff \Delta x = \frac{\ln 0.5}{\ln 0.9} = 6.6 \text{ (Jahre)}$$

(vii) $2^{3x-7} = 3 \iff 3x - 7 = \dfrac{\ln 3}{\ln 2} = 1.58 \iff x = \dfrac{1.58 + 7}{3} = 2.86$

(viii) $2^x \cdot 3^{7-2x} = 6 \iff x \ln 2 + (7-2x) \ln 3 = \ln 6 \iff$

$$(\ln 2 - 2\ln 3)x = \ln 6 - 7\ln 3 \iff x = \frac{\ln 6 - 7\ln 3}{\ln 2 - 2\ln 3} = 3.92$$

(alternativ: $2^x \cdot 3^{7-2x} = 6 \iff \dfrac{2^x}{3^{2x}} = \dfrac{6}{3^7} \iff \dfrac{2^x}{9^x} = \dfrac{6}{3^7} \iff$

$\left(\dfrac{2}{9}\right)^x = \dfrac{6}{3^7} \iff x = \dfrac{\ln 6/3^7}{\ln 2/9} = 3.92$)

(ix) $2^x + 3^x = 7:$ leider nichts machbar

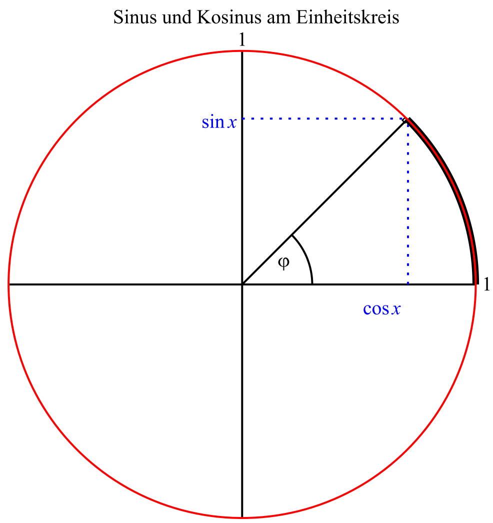
(x) $\ln x + \ln(8x) - \ln\left(\dfrac{1}{x}\right) = 0 \iff \ln\left(\dfrac{x \cdot 8x}{1/x}\right) = 0 \iff$

$$\ln(8x^3) = 0 \iff 8x^3 = 1 \iff x = \frac{1}{2}$$

§ 6 Die Kreisfunktionen

Wir definieren die trigonometrischen Funktionen anhand des Einheitskreises

$$EK = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \; x^2 + y^2 = 1 \right\} \left(= \left\{ z \in \mathbb{C} : \; |z| = 1 \right\} \right)$$



Sei dazu $P = \begin{pmatrix} x \\ y \end{pmatrix} \in$ EK und φ der Winkel zur positiven $x-$Achse.

Die $x-$Koordinate von P $(= \text{Re } z)$ nennen wir den **Kosinus** von φ,

die $y-$Koordinate von P $(= \text{Im } z)$ den **Sinus** von φ :

$$x = \cos\varphi, \; y = \sin\varphi$$

und definieren noch den **Tangens** über

$$\tan\varphi := \frac{\sin\varphi}{\cos\varphi}, \quad \text{falls} \quad \cos\varphi \neq 0.$$

Den Winkel φ können wir im Gradmaß messen oder (mittlerweile eher üblich) im Bogenmaß x. (= Länge des Bogens von 1 nach P.)

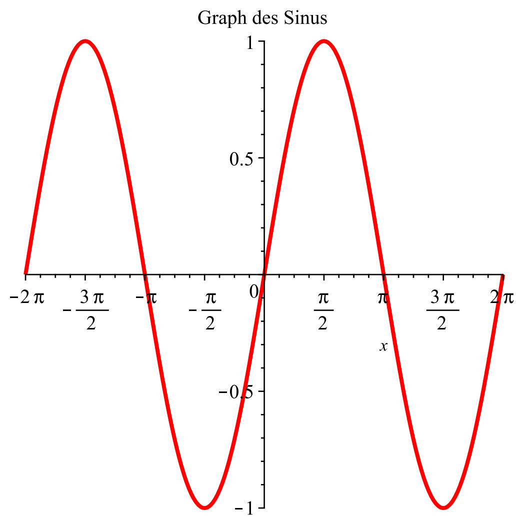
Der Zusammenhang erschließt sich über den Dreisatz:

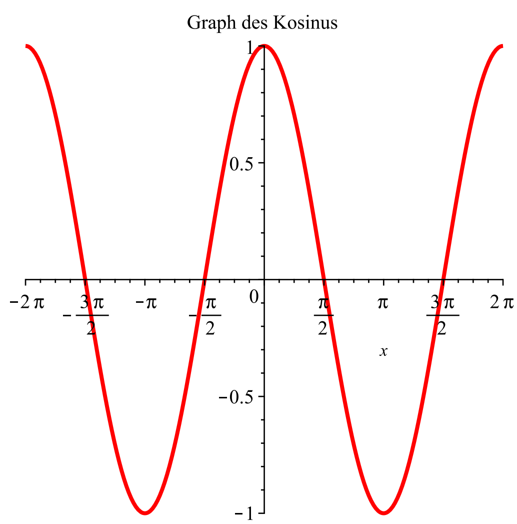
$$\frac{\varphi}{360^0} = \frac{x}{2\pi}$$

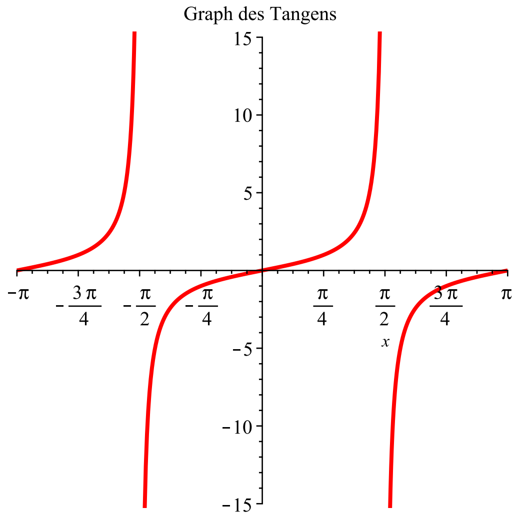
Kleine Wertetabelle:

x (Bogenmaß)	0	$\frac{\pi}{2}$	π	$\frac{3}{2}\,\pi$	$\frac{\pi}{4}$	$\frac{5}{4}\,\pi$	$-\frac{\pi}{4}$	$\frac{\pi}{6}$
φ (Gradmaß)	0	90^0	180^0	270^0	45^0	225^0	-45^0	30^0
$\cos x$	1	0	-1	0	$\frac{1}{2}\,\sqrt{2}$	$-\frac{1}{2}\,\sqrt{2}$	$\frac{1}{2}\,\sqrt{2}$	$\frac{1}{2}\,\sqrt{3}$
$\sin x$	0	1	0	-1	$\frac{1}{2}\,\sqrt{2}$	$\frac{1}{2}\,\sqrt{2}$	$-\frac{1}{2}\,\sqrt{2}$	$\frac{1}{2}$
$\tan x$	0	$-$	0	$-$	1	-1	-1	$\frac{1}{\sqrt{3}}$

Graphen:







Eigenschaften: (aus dem Bild oder den Graphen ablesbar)

(0) Sinus und Kosinus sind $2\pi-$periodisch; der Tangens ist $\pi-$periodisch:

$$\sin(x+2\pi) = \sin x,\ \cos(x+2\pi) = \cos x;\ \tan(x+\pi) = \tan x$$

(i) $\sin(-x) = -\sin x$ ("ungerade Funktion")
$\cos(-x) = \cos x$ ("gerade Funktion")

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

(iii) $-1 \le \cos x \le 1,\ -1 \le \sin x \le 1$ stets

(iv) $x \ge 0 \Rightarrow \sin x \le x$

(v) Nullstellen von $\sin x:\ 0,\ \pm\pi,\ \pm 2\pi, \dots$

Nullstellen von $\cos x:\ \pm\frac{\pi}{2},\ \pm\frac{3}{2}\pi, \dots$

Definitionslücken von $\tan x:\ \pm\frac{\pi}{2},\ \pm\frac{3}{2}\pi, \dots$

Erwähnt seien noch die beiden Additionstheoreme

$$\cos(x+y) = \cos x \cos y - \sin x \sin y \quad \text{und}$$

$$\sin(x+y) = \sin x \cos y + \cos x \sin y$$

Gleichungen in $\sin x,\ \cos x,\ \tan x$ **:**

- Die Gleichung $\sin x = b$

  (a) ist im Falle $b > 1$ oder $b < -1$ unlösbar (in $\mathbb{R}$).

  (b) besitzt im Falle $-1 \leq b \leq 1$ unendlich viele Lösungen.

  Genau eine davon fällt in das Intervall $\left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$.

  Diese nennt man dann den **Arcussinus** von b :

  $$x_1 = \arcsin b$$

  und ist so auch in den Taschenrechnern implementiert.

  Es gibt dann im Intervall $[0, 2\pi]$ noch eine zweite Lösung, nämlich

  $$x_2 = \pi - x_1.$$

  Alle anderen Lösungen ergeben sich wegen der Periodizität zu

  $$x_1 + 2k\pi \quad \text{bzw.} \quad x_2 + 2k\pi,\ k \in \mathbb{Z}.$$

- Die Gleichung $\cos x = b$ diskutieren wir in den Übungen (Definition von $\arccos b$ so, dass $\arccos 0 = 1$ gilt)

- Die Gleichung $\tan x = b$ besitzt für jedes $b \in \mathbb{R}$ unendlich viele Lösungen.

  Genau eine davon fällt in das Intervall $\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right)$. Diese nennen wir dann **Arcustangens** von b :

  $$x = \arctan b.$$

  Alle anderen Lösungen sind dann

  $$\arctan b + k\pi,\ k \in \mathbb{Z}.$$

Kapitel II Der Grenzwertbegriff

§ 1 Zahlenfolgen

Rein formal sind Zahlenfolgen Funktionen, deren Definitionsbereich

$$\mathbb{N} \quad \text{oder} \quad \mathbb{N}_0 \quad \text{oder} \quad \{m, m+1, m+2, \ldots\} \ \ (m \in \mathbb{Z} \quad \text{fest}) \text{ ist.}$$

Wir bezeichnen sie in der Regel mit

$$(a_n)_{n\in\mathbb{N}}, \ (a_n)_{n=m}^{\infty}, \ (a_n)_{n\geq m} \quad \text{o.ä.}$$

Man kann Zahlenfolgen auf verschiedene Arten definieren:

- aufzählend, z.B. $(a_n) = (1, 2, 4, 8, 16, 32, \ldots)$: wird oft gemacht, ist aber nicht präzise
- als Formelvorschrift, z.B. obige Folge: $a_n = 2^n, \ n \in \mathbb{N}$
- rekursiv, z.B. Fibonacci-Folge:

  $$F_0 = 0, \ F_1 = 1, \ F_n := F_{n-1} + F_{n-2} \quad \text{für} \quad n \geq 2,$$

  also

  $$F_2 = 1, \ F_3 = 2, \ F_4 = 3, \ F_5 = 5, \ F_6 = 8, \ F_7 = 13, \ \ldots$$

Wir wollen in diesem einleitenden Abschnitt salopp einige Beispiele anschauen im Hinblick auf ihr Verhalten für große Werte von n $(n \to \infty)$

(i) Was haben die Folgen

$$n, \ n \ln n, \ 2^n, \ n^2, \ 1.1^n, \ n^n, \ n\sqrt{n}, \ n! \quad \text{und die Fibonacci-Folge } F_n$$

gemeinsam? Man erkennt wohl mit wenig Mühe:

- alle sind wachsend
- alle sind unbeschränkt, " $\to \infty$, allerdings
- "mit "unterschiedlicher Geschwindigkeit".

Man kann sich ein Ranking vorstellen: geordnet nach aufsteigender Geschwindigkeit wäre dies etwa

$$n \prec n \ln n \prec n^2 \prec 1.1^n \prec F_n \prec 2^n \prec n! \prec n^n$$

(ii) Was haben die Folgen

$$\frac{1}{n},\ \frac{1}{2^n},\ \left(\frac{3}{4}\right)^n,\ \frac{1}{\ln n}$$

gemeinsam?

- alle sind monoton fallend
- "gehen alle gegen 0",
- auch wieder mit unterschiedlicher Geschwindigkeit.
- Ranking? Nach der Vorschrift "schnellste" zuerst, würde das wohl so aussehen:
$$\frac{1}{2^n} \prec \left(\frac{3}{4}\right)^n \prec \frac{1}{n} \prec \frac{1}{\ln n}$$
- Wie passt $\left(-\frac{1}{2}\right)^n$ hier herein?

  $\rightarrow$ geht auch gegen 0, aber nicht monoton wie die anderen

(iii) Wie verhalten sich die nachstehenden Folgen für große n ?

$\dfrac{n}{n+1} \quad \rightarrow 1?$

$2 - \dfrac{1}{n} \quad \rightarrow 2?$

$\dfrac{n^2+3n+5}{3n^2-8n+17} \quad \rightarrow \dfrac{1}{3}?$

(iv) Wie ist das Verhalten von $a_n = (-1)^n$? $\rightarrow$ divergent

(v) Wie messen wir am besten das Ranking bei den Folgen aus (i)?
Am besten, man setzt je zwei aufeinanderfolgende ins Verhältnis und beobachtet, (wie?) dass dieser Quotient jeweils gegen 0 geht, also

$$\frac{n}{n \ln n} \rightarrow 0,\ \frac{n \ln n}{n^2} \rightarrow 0,\ \frac{n^2}{1.1^n} \rightarrow 0 \quad \text{usw.}$$

§ 2 Grenzwerte von Zahlenfolgen

Definition:

Gegeben sei eine Folge (a_n) reeller oder komplexer Zahlen. Wir sagen:
(a_n) **konvergiert** gegen $a \in \mathbb{R}$ (oder $\mathbb{C}$), falls gilt:

$$\text{Zu jedem } \varepsilon > 0 \text{ gibt es ein } N \text{ mit } |a_n - a| < \varepsilon \quad \forall\ n > N.$$

Wir schreiben dann

$$\lim_{n\to\infty} a_n = a \quad \text{oder} \quad a_n \to a,\ n \to \infty,$$

und wir nennen a den **Grenzwert** oder **Limes** von (a_n).
Ist speziell $\lim\limits_{n\to\infty} a_n = 0$, so sprechen wir von einer **Nullfolge.**
Nicht konvergente Folgen nennen wir auch **divergent.**

Bemerkungen und Beispiele:

(i) In $\mathbb{R}$ bedeutet $|a_n - a| < \varepsilon$ auch

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

d.h. ab einem Index N liegen **alle** Folgenglieder a_n im Intervall

$$(a - \varepsilon, a + \varepsilon)$$

bei vorgegebener Fehlerschranke $\varepsilon > 0$.

Man kann das auch so formulieren: Außerhalb dieses Intervalls liegen nur endlich viele dieser Folgenglieder a_n. Deshalb ist auch die folgende Sprechweise geläufig:

Fast alle Folgenglieder liegen in dem Intervall $(a - \varepsilon, a + \varepsilon)$.

In $\mathbb{C}$ bedeutet $|a_n - a| < \varepsilon$: die Folgenglieder a_n liegen ab dem Index N in der offenen Kreisscheibe um a mit Radius ε.

(ii) Die Definition präzisiert die gebräuchliche, aber etwas unsaubere Sprechweise " a_n nähert sich immer mehr dem Wert a an."

Dabei liegt die Betonung auf <u>zu jedem</u> $\varepsilon > 0$. Es liegt also eine gewisse Dynamik zugrunde. Das dazu "passende" N wird in der Regel von der Vorgabe ε abhängen: $N = N(\varepsilon)$. I.a. dürfen wir erwarten, dass mit kleiner werdendem ε die Schranke N immer größer gewählt werden muss.

Hat man ein passendes $N = N(\varepsilon)$ gefunden, so leistet jede Zahl größer als N das Gewünschte auch. Diese sind dann zwar nicht optimal, aber das wird in der Definition ja auch nicht gefordert.

(iii) Beispiel: Wir behaupten:

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

Beweis:

Sei $\varepsilon > 0$ vorgegeben. Wähle dann $N := \frac{1}{\varepsilon}$.

Dann gilt für $n > N$: $\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \varepsilon$.

(iv) Beispiel: Wir behaupten:

$$\lim_{n\to\infty} \left(\frac{1}{2}\right)^n = 0.$$

Beweis:

Sei $\varepsilon > 0$ gegeben. Wähle $N := \frac{\ln \varepsilon}{\ln \frac{1}{2}}$

$n > N \;\Rightarrow\; |a_n - a| = \left|\left(\frac{1}{2}\right)^n - 0\right| = \left(\frac{1}{2}\right)^n < \varepsilon \iff n \ln \frac{1}{2} < \ln \varepsilon \iff$

$$n > \frac{\ln \varepsilon}{\ln \frac{1}{2}}.$$

(Beachte dabei: $\ln \frac{1}{2} < 0$!!)

Nebenbei: Gelegentlich findet man in der Literatur, dass dieses N eine natürliche Zahl sein soll. Das ist überflüssig, denn man kann ja statt $\frac{\ln \varepsilon}{\ln \frac{1}{2}}$ auch die nächst größere natürliche Zahl wählen, notiert gerne in der Form

$$\left[\frac{\ln \varepsilon}{\ln \frac{1}{2}}\right] + 1,$$

wobei $[x]$ die größte ganze Zahl $\leq x$ bedeute (Gauß-Klammer,"floor"-Funktion, entspricht dem Abrunden auf die nächst kleinere ganze Zahl).

(v) Verallgemeinerung von (iv):

Ist $z \in \mathbb{C}$ mit $|z| < 1$, so ist

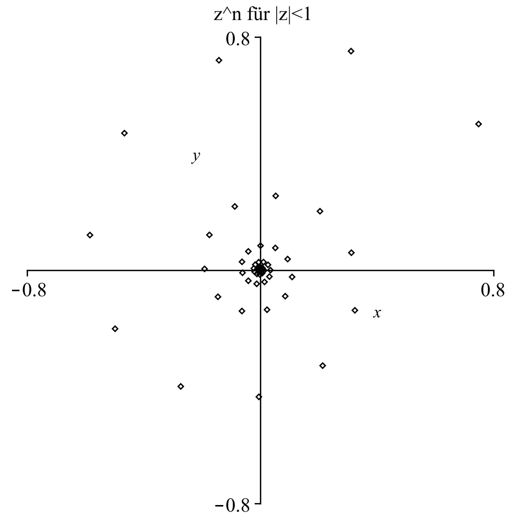
$$\lim_{n\to\infty} z^n = 0,$$

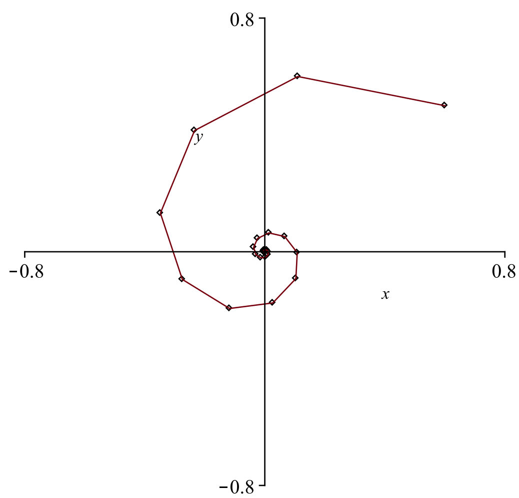
denn: $|z^n - 0| = |z|^n < \varepsilon$, falls $n > \frac{\ln \varepsilon}{\ln |z|} =: N$.

Also ist z.B. $\left(\frac{1}{2}\right)^n \to 0$, $\left(\frac{2022}{2023}\right)^n \to 0$ (beide sogar in monoton fallender Weise), aber auch

$\left(-\frac{1}{2}\right)^n \to 0$ (in alternierender Weise) oder auch

$\left(\frac{i}{2}\right)^n \to 0$, $\left(\frac{3+2i}{4}\right)^n \to 0$, etc. (in "drehender Weise").





(v) Beispiel: Wir behaupten:

$$\frac{n-1}{3n} \to \frac{1}{3},$$

Beweis dazu:

$$\left|\frac{n-1}{3n} - \frac{1}{3}\right| = \left|\frac{n-1-n}{3n}\right| = \frac{1}{3n} < \varepsilon \text{ für } n > N := \frac{1}{3\varepsilon}.$$

(vi) Für die Frage, **ob** eine Folge konvergiert oder nicht, ist es nicht entscheidend, ob die Schranke N optimal gewählt wurde oder nicht. Oft kann man die Ungleichung $|a_n - a| < \varepsilon$ gar nicht oder nur sehr schwer nach n auflösen. Betrachte etwa als Beispiel

$$a_n = \frac{1}{n^3 + 5n^2 + n + 42}.$$

Dann ist a_n Nullfolge, denn $|a_n - 0| \leq \dfrac{1}{n} < \varepsilon$ für $n > N := \dfrac{1}{\varepsilon}$, wenngleich dieses N sicherlich nicht optimal ist.

Man sieht hier sehr schön, wie in der Analysis oft gearbeitet wird: Man darf zu Beweiszwecken "abschätzen", um einfacher zum Resultat zu kommen. Man hat hierbei relativ viel Freiheit, das zu tun. Im Beispiel hätte auch die Abschätzung

$$|a_n| \le \frac{1}{n^3}$$

gereicht, was zu $N = \dfrac{1}{\sqrt[3]{\varepsilon}}$ geführt hätte. Zu viel des Guten wäre die Abschätzung

$$|a_n| \le \frac{1}{42}$$

gewesen, was zwar stimmt, aber nicht zielführend ist.

(vii) Zu einfachen Abschätzungen kommt man, wenn man im Zähler negative Terme oder im Nenner positive Terme einfach weglässt. Ist dies nicht so einfach möglich, kann man ein wenig in die Trickkiste greifen. Betrachte dazu das folgende

Beispiel: Wir behaupten:

$$\lim_{n\to\infty} \frac{n^2+n}{n^3-7} = 0$$

Beweis dazu:

$$\left|\frac{n^2+n}{n^3-7}\right| \le \frac{n^2+n^2}{n^3-\frac{1}{2}\,n^3} = \frac{4}{n} < \varepsilon, \quad \text{für } \; n > \frac{4}{\varepsilon} =: N.$$

(viii) Der Grenzwert einer Zahlenfolge ist, wenn er überhaupt existiert, eindeutig bestimmt.

(ix) Konvergente Folgen (a_n) sind beschränkt, d.h. es gibt $K > 0$ mit

$$|a_n| \le K \;\; \forall\, n \in \mathbb{N}.$$

Dies sieht man, wenn man in der Definition ein konkretes ε wählt, etwa $\varepsilon = 42 > 0$. Dann gilt gibt es ein $N \in \mathbb{N}$ mit $|a_n - a| < 42 \quad \forall \;\, n > N$. Dann ist aber für diese n

$$|a_n| = |a_n - a + a| \le |a_n - a| + |a| \le 42 + |a|,$$

und man kann

$$K := \max\{\ 42 + |a|, , |a_1|, |a_2|,\ \ldots,\ |a_N|\ \}$$

setzen.

Die Aussage ist ein Argument dafür, dass Folgen wie

$$n,\ n^2,\ \sqrt{n},\ \ln n$$

etc. alle divergieren, da sie unbeschränkt sind.

(x) Ist eine Folge (a_n) konvergent, so gilt

$a_{n+1} - a_n \to 0$, auch $a_{n+2} - a_n \to 0$, $\ldots$, $a_{n+p} - a_n \to 0$ für jedes p .

Das ist die sog. "Cauchy-Bedingung" (Augustin Cauchy, 1789 - 1857).

Zusatz: Fordert man

$$\forall\, \varepsilon > 0 \;\; \exists\, N \quad \text{mit } |a_{n+p} - a_n| < \varepsilon \quad \forall \;\; n > N, \;\; \forall \;\; p \in \mathbb{N},$$

also eine leicht verschärfte Cauchy-Bedingung, so ist diese auch hinreichend für Konvergenz.

Dies ist insofern bemerkenswert, als dies ein Instrument darstellt, um Konvergenz zu zeigen **ohne Kenntnis des Grenzwerts.**

Als Anwendung zeigen wir, dass die Folge $(-1)^n$ divergiert:

$$|a_{n+1} - a_n| = \left(\left|(-1)^{n+1} - (-1)^n\right| = 2 \not\to 0\right),$$

Allgemeiner gilt sogar:

Ist $z \in \mathbb{C}$, $|z| \geq 1$, $z \neq 1$, so ist z^n divergent.

Beweis dazu:

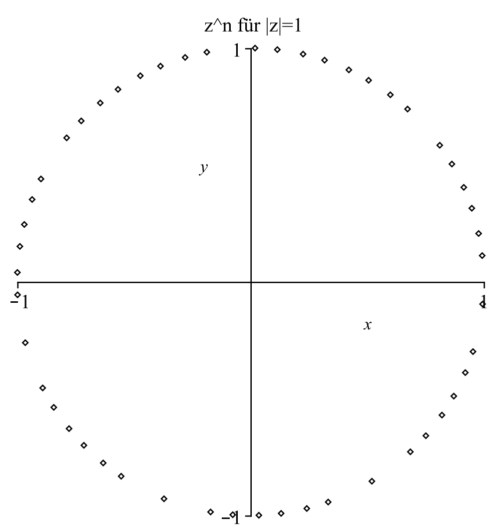
$$\left|z^{n+1} - z^n\right| = |z|^n \cdot |z-1| \geq |z-1| \not\to 0.$$

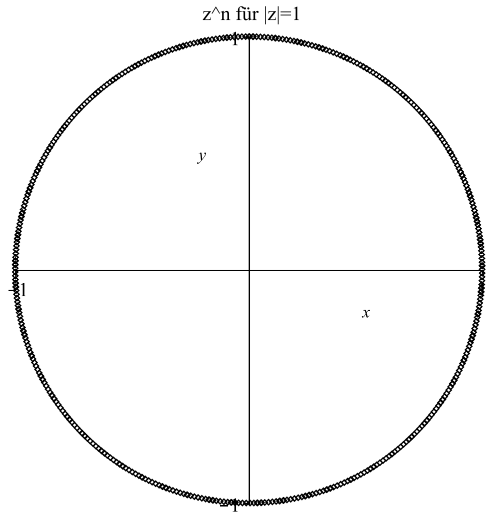
Also sind

$$\left(\frac{11}{10}\right)^n, \; \left(-\frac{2023}{2022}\right)^n, \; (1+i)^n, \; \left(\frac{3}{5} + \frac{4}{5}i\right)^n$$

allesamt divergent (mit unterschiedlichem Charakter).

Die letzte dieser Folgen sei in der Gaußschen Zahlenenbene veranschaulicht:





(xi) Ist eine Folge (a_n) monoton wachsend, d.h.

$$a_n \le a_{n+1} \quad \forall \ n \qquad (\text{Notation:} \quad (a_n)\uparrow),$$

so ist sie konvergent, wenn sie nach oben beschränkt ist.

Das ist der **Satz über monotone Konvergenz.**

Ebenso konvergiert eine monoton fallende Folge (a_n), d.h.

$$a_n \ge a_{n+1} \quad \forall \ n \ \ (\text{Notation:} \quad (a_n)\downarrow) \ ,$$

wenn sie nach unten beschränkt ist.

Das ist genau das Vollständigkeitsmerkmal der reellen Zahlen, was in Abschnitt I.1.C. noch gefehlt hat. Begnügt man sich nur mit den rationalen Zahlen, so so stimmt dies nämlich nicht mehr. Beispielsweise ist die Folge

$$(a_n) = (1,\ 1.4,\ 1.41,\ 1.414,\ 1.4142,\ \ldots)$$

(also die Folge der Dezimalbruchentwicklung von $\sqrt{2}$ monoton wachsend, $\in \mathbb{Q}$, und nach oben beschränkt (durch $\sqrt{2}$), aber ihr Grenzwert, (der in $\mathbb{R}$ existiert), liegt nicht mehr in $\mathbb{Q}$.

(xii) Ein weiteres Beispiel ist die Definition der Eulerschen Zahl e.

$a_n = \left(1 + \dfrac{1}{n}\right)^n$ ist monoton wachsend, denn:

$$\frac{a_{n+1}}{a_n} = \frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{\left(1+\frac{1}{n}\right)^n} = \left(1+\frac{1}{n}\right)\left(\frac{1+\frac{1}{n+1}}{1+\frac{1}{n}}\right)^{n+1} =$$

$$\left(1+\frac{1}{n}\right)\left(1+\frac{1+\frac{1}{n+1}-1-\frac{1}{n}}{1+\frac{1}{n}}\right)^{n+1} = \left(1+\frac{1}{n}\right)\left(1+\frac{\frac{1}{n+1}-\frac{1}{n}}{1+\frac{1}{n}}\right)^{n+1} \underbrace{\ge}_{\text{Bernoulli}}$$

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

Ganz analog kann man zeigen, dass $b_n=\left(1+\frac{1}{n}\right)^{n+1}$ monoton fällt. Damit ist a_n aber nach oben beschränkt (durch jedes b_n, also z.B. durch $b_1=2^2=4$) und daher konvergent.

(xiii) Die Konstruktion von e in Beispiel (xii) ist verallgemeinerbar zum sog. **Intervallschachtelungsprinzip:**

Sind $(a_n),(b_n)$ reelle Zahlenfolgen mit

$$(a_n)\uparrow,\ (b_n)\downarrow,\ a_n\le b_n \quad \text{und} \quad b_n-a_n\to 0,$$

so konvergieren (a_n) und (b_n) gegen denselben Wert.

(xiv) Nützlich und eng verwandt mit dem Intervallschachtelungsprinzip ist der sog. **Einschnürungssatz,** manchmal auch **Quetschlemma** oder **Sandwich-Theorem** genannt:

Gilt

$$a_n\le c_n\le b_n,\ \ a_n\to s \quad \text{und} \quad b_n\to s,$$

so gilt auch

$$c_n\to s.$$

Als Beispiel zeigen wir:

$$\lim_{n\to\infty}\sqrt[n]{n}=1.$$

Dazu die folgende Überlegung:

Ist $h>0$, so ist $(1+h)^n\ge\binom{n}{2}h^2$ (nach dem Binomischen Lehrsatz).

Wir setzen $h=\sqrt[n]{n}-1$ ein und erhalten

$$\underbrace{\left(\sqrt[n]{n}\right)^n}_{=n}\ge\underbrace{\binom{n}{2}}_{=\frac{n(n-1)}{2}}\left(\sqrt[n]{n}-1\right)^2$$

$$\Rightarrow \underbrace{0}_{\to 0}\le\sqrt[n]{n}-1\le\underbrace{\sqrt{\frac{2}{n-1}}}_{\to 0}$$

Das Ergebnis werden wir mit besseren Hilfsmitteln (Regel von l'Hospital) später eleganter zeigen können.

Rechenregeln für Grenzwerte:

- $a_n \to 0, (b_n)$ beschränkt $\Rightarrow a_n \cdot b_n \to 0$
  (denn: $|b_n| \le K, \;\; |a_n| < \dfrac{\varepsilon}{K} \;\; \forall \; n > N \;\Rightarrow\; |a_n b_n| < \varepsilon \;\; \forall \; n > N$)
- $a_n \to a, \; c \in \mathbb{C}$ beliebig $\Rightarrow ca_n \to ca$
  (denn: $a_n - a \to 0, \; b_n \equiv c \underbrace{\Rightarrow}_{\text{1.Regel}} ca_n - ca \to 0$)
- $a_n \to a, \; b_n \to b \;\Rightarrow\; a_n \pm b_n \to a \pm b$
  (denn: Dreiecksungleichung: $|a_n + b_n - (a+b)| \le |a_n - a| + |b_n - b|$)
- $a_n \to a, \; b_n \to b \;\Rightarrow\; a_n \cdot b_n \to a \cdot b$
  (denn: $a_n b_n - ab = \underbrace{a_n}_{\text{beschr.}} \underbrace{(b_n - b)}_{\to 0} + b \underbrace{(a_n - a)}_{\to 0} \to 0$)
- $a_n \to a, \; b_n \to b \neq 0 \;\Rightarrow\; \dfrac{a_n}{b_n} \to \dfrac{a}{b}$
  (denn: $\dfrac{1}{b_n} - \dfrac{1}{b} = \dfrac{b - b_n}{b b_n} \to 0$)
  (eigentlich müsste man noch zeigen: $\dfrac{1}{b_n}$ ist beschränkt)

Beispiele:

Bestimme die Grenzwerte der Folgen

$$\frac{(-1)^n + \sin(n^5 - 42)}{n}, \quad \frac{6n^2 - n + 3}{4n^2 + 1} \quad \text{und} \quad \frac{3^n + 2^n}{(-1)^n + 3^{n+1}}$$

(i) Wir behaupten

$$\frac{(-1)^n + \sin(n^5 - 42)}{n} \to 0.$$

Das ist aber klar, denn $\dfrac{1}{n} \to 0$ und $|(-1)^n + \sin(n^5 - 42)| \le 1 + 1 = 2$

(ii) $$\frac{6n^2 - n + 3}{4n^2 + 1} = \frac{6 - \frac{1}{n} + \frac{3}{n^2}}{4 + \frac{1}{n^2}} \to \frac{6}{4} = \frac{3}{2}$$

(iii) $$\frac{3^n + 2^n}{(-1)^n + 3^{n+1}} = \frac{1 + \left(\frac{2}{3}\right)^n}{\left(-\frac{1}{3}\right)^n + 3} \to \frac{1}{3}$$

Abschließend noch der wichtige Vergleich zwischen Potenzfunktionen und Exponentialfunktionen (in Form von Folgen):

Sei dazu $h > 0$ und $k \in \mathbb{N}$. Wir betrachten $a_n = \dfrac{n^k}{(1+h)^n}$.

$$\Rightarrow \; 0 \le a_n \underbrace{\le}_{\text{Binomi, } n \ge k+1} \frac{n^k}{\binom{n}{k+1} h^{k+1}} = \frac{1}{h^{k+1}} \frac{n^k}{n^{k+1} + \ldots} \to 0$$

Auch dieses Ergebnis werden wir mit verbessertem Instrumentarium einfacher bekommen.

Also ist z.B.

$$\lim_{n\to\infty} \frac{n^2}{2^n} = 0, \quad \lim_{n\to\infty} \frac{n^{10}}{3^n} = 0, \quad \lim_{n\to\infty} \frac{n^{2022}}{1.001^n} = 0.$$

Etwas salopp kann man sagen: Die Exponentialfunktionen wachsen schneller als jede Potenz. Dies überträgt sich sofort auf den Vergleich zwischen Wurzeln und Logarithmen. Da dies die jeweiligen Umkehroperationen sind, wächst also der Logarithmus langsamer als jede Wurzel:

$$\lim_{n\to\infty} \frac{\ln n}{n^\alpha} = 0 \quad \forall \ \alpha > 0.$$

Bemerkungen:

(i) Bei Ausdrücken der Form $\frac{\infty}{\infty}$, $\infty-\infty$, $0\cdot\infty$, $\frac{0}{0}$, 1^∞ kommt man meist nicht ohne arithmetische Umformungen oder andere Überlegungen aus, denn es kann im wahrsten Sinne des Wortes "alles" passieren, z.B.

- $\infty - \infty$:

  $(n+1)^2 - (2n+n^2) = 1 \to 1$

  $(n+1)^2 - n^2 = 2n+1 \to \infty$

  $\ln(n+1) - \ln n = \ln\left(\frac{n+1}{n}\right) = \ln\left(1+\frac{1}{n}\right) \to 0$ ("Stetigkeit des ln")

  $\sqrt{n+1} - \sqrt{n} = \frac{1}{\sqrt{n+1}+\sqrt{n}} \to 0$

- $0 \cdot \infty$:

  $\frac{1}{n} \cdot n^2 = n \to \infty$

  $\frac{1}{n^2} \cdot n = \frac{1}{n} \to 0$

  $\left(\sqrt{n+1} - \sqrt{n}\right) \cdot \sqrt{n} = \frac{\sqrt{n}}{\sqrt{n+1}+\sqrt{n}} = \frac{1}{\sqrt{1+\frac{1}{n}}+1} \to \frac{1}{2}$

- 1^∞ :

  $\left(1+\frac{1}{n}\right)^n \to e$

  $\left(1+\frac{1}{n}\right)^{n^2} \underbrace{\geq}_{\text{Bernoulli}} 1 + \frac{n^2}{n} = 1 + n \to \infty$

(ii) Landausche O-o-Symbolik: (Edmund Landau,1877 - 1938)

Für Folgen $(a_n), (b_n)$ aus $\mathbb{R}$ schreibt man oft:

- $a_n = O(b_n)$, falls $\left|\frac{a_n}{b_n}\right|$ beschränkt ist bzw.
- $a_n = o(b_n)$, falls $\left|\frac{a_n}{b_n}\right| \to 0$ gilt und
- $a_n \sim b_n$, falls $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = 1$ gilt. (Hier sagt man dann auch: (a_n) und (b_n) verhalten sich **asymptotisch gleich**).

z.B.

$$3n^2 + 6n - 4 = O(n^2); \;\; n^7 = o(7^n); \;\; n^3 \sim n^3 + n^2; \; n \ln n = o(n^2).$$

Ein sehr schönes Ergebnis ist die Asymptotik von Stirling:

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

was wir hier nicht zeigen können. Man liest daraus ab:

$$n! = o(n^n)$$

Bei nicht konvergenten Folgen führt man oft Ersatzbegriffe ein:

- Gegeben (a_n) . Streicht man einige Folgenglieder weg (endlich viele oder unendlich viele, aber so, dass noch eine Folge "übrig bleibt"),so spricht man von einer **Teilfolge.**

  Hat diese einen Grenzwert s , so nennt man s einen **Häufungswert** von (a_n).

  Zum Beispiel hat die Folge $a_n = (-1)^n$ die beiden Häufungswerte

  $$1 \; (= \lim_{n\to\infty} a_{2n}) \quad \text{und} \quad -1 \; (= \lim_{n\to\infty} a_{2n+1}).$$

- Satz von **Bolzano-Weierstraß**: Jede beschränkte Folge besitzt mindestens einen Häufungswert.

  Unter den Häufungswerten gibt es dann einen größten, den man den **limes superior** nennt ($\limsup a_n$) und auch einen kleinsten, den **limes inferior** ($\liminf a_n$). Also ist z.B.

  $$\liminf(-1)^n = -1 \quad \text{und} \quad \limsup(-1)^n = 1.$$

- Bei unbeschränkten Folgen können mehrere Fälle eintreten, die wir an Beispielen erörtern:

  (i) $a_n = n^2$:

  Man sagt, diese Folge sei **bestimmt divergent** gegen ∞, d.h.

  $$\forall\, K > 0 \;\; \exists\, N \quad \text{mit} \quad a_n \geq K \;\; \forall\, n \geq N$$

  und setzt in diesem Fall $\limsup a_n = \liminf a_n := +\infty$

  (ii) $a_n = \Big(1 + (-1)^n\Big) n^2$:

  Diese besitzt einen endlichen Häufungswert $0 = \lim\limits_{n\to\infty} a_{2n+1}$ und eine nach oben unbeschränkte Teilfolge a_{2n}. Sinnvollerweise setzt man hier $\liminf a_n = 0$ und $\limsup a_n = +\infty$

  (iii) $a_n = (-1)^n \cdot n^2$:

  Diese Folge besitzt eine nach oben unbeschränkte Teilfolge a_{2n} und eine nach unten unbeschränkte Teilfolge a_{2n+1} . Hier sollte man $\limsup a_n := +\infty$ und $\liminf a_n := -\infty$ setzen.

  Ergänzt man die Begriffsbildung durch solche Erweiterungen, so kann man sagen: Nicht jede Folge besitzt einen Limes, aber jede Folge besitzt einen Limes superior und einen Limes inferior.

§ 3 Funktionsgrenzwerte und Stetigkeit

Wir studieren das Verhalten einer Funktion $f(x)$ an einer Stelle a, indem wir untersuchen, wie sich die Funktionswerte $f(x_n)$ einer Folge verhalten, wenn $x_n \to a$ gilt. Dazu zunächst zwei Beispiele:

(i) Sei $f(x) = x^2$, und $a = 2$.

Ist (x_n) irgendeine Folge mit $x_n \to 2$, so gilt wegen der Grenzwertregeln für Folgen $f(x_n) = x_n^2 \to 2^2 = 4$. Wir werden sagen:

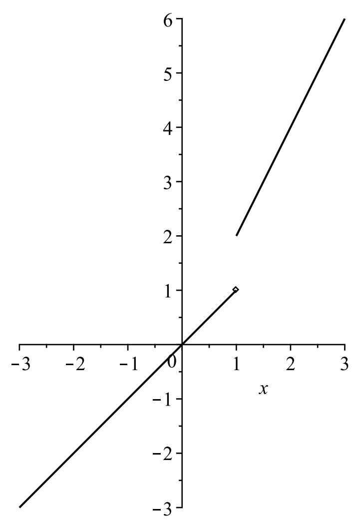
$$\lim_{x \to 2} f(x) = 4.$$

Allgemeiner gilt: Ist $a \in \mathbb{R}$ beliebig, und ist (x_n) eine Folge mit $x_n \to a$, so ist nach den Grenzwertregeln $f(x_n) \to a^2$. Wir schreiben daher wieder:

$$\lim_{x \to a} x^2 = a^2.$$

Da auch $f(a) = a^2$ ist, werden wir f **stetig** an jeder Stelle a nennen.

(ii) Sei $f(x) = \begin{cases} x & x \leq 1 \\ 2x & x > 1 \end{cases}$ und $a = 1$.



Wählen wir $x_n = 1 - \dfrac{1}{n}$, so ist $x_n \to 1$ und $f(x_n) = 1 - \dfrac{1}{n} \to 1$;

Wählen wir aber $x_n = 1 + \dfrac{1}{n}$, so ist auch $x_n \to 1$, aber diesmal ist $f(x_n) = 2\left(1 + \dfrac{1}{n}\right) \to 2$.

Wir sagen:

$$\lim_{x \to 1} f(x) \quad \text{existiert nicht.}$$

Ist indes $x_n \to 1$, $x_n < 1$, so gilt: $f(x_n) \to 1$. Wir werden sinnvollerweise von einem **linksseitigen Grenzwert** reden:

$$\lim_{x \to 1-} f(x) = 1.$$

Ebenso: $x_n \to 1,\ x_n > 1 \Rightarrow f(x_n) = 2x_n \to 2$. Sinnvollerweise sollten wir von einem **rechtsseitigen** Grenzwert reden:

$$\lim_{x\to 1+} f(x) = 2$$

Wegen $f(1) = 1$ werden wir f an der Stelle $a = 1$ **linksseitig stetig** nennen, rechtsseitig stetig aber nicht.

Definition:

Gegeben sei eine Funktion $f : D \to \mathbb{R}$ und eine Stelle $a \in D$.

Wir sagen:

f ist **konvergent** für x gegen a gegen den Wert s,

wenn gilt:

Für jede Folge x_n aus D mit $x_n \to a,\ x_n \neq a$ gilt $f(x_n) \to s$.

Wir notieren diese Situation in der Form

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

Existiert $s = \lim_{x\to a} f(x)$ und ist zusätzlich $s = f(a)$, so nennt man f **stetig** an der Stelle a .

Analog definiert man Grenzwerte für $x \to \pm\infty$:

$$\lim_{x\to\pm\infty} f(x) = s, \text{falls } f(x_n) \to s \text{ für alle Folgen } x_n \text{ mit } x_n \to \pm\infty$$

und auch **einseitige** Grenzwerte

$$\lim_{x\to a-} f(x),\ \lim_{x\to a+} f(x)$$

(es werden nur Testfolgen x_1 mit $x_n < a$ bzw. $x_n > a$ zugelassen.)

Schließlich reden von **linksseitiger** bzw. **rechtsseitiger Stetigkeit** bei a, falls

$$\lim_{x\to a-} f(x) = f(a) \quad \text{bzw.} \quad \lim_{x\to a+} f(x) = f(a) \quad \text{gilt.}$$

Bemerkung:

Das ist die sogenannte diskrete Version des Begriffs des Funktionswerts. In manchen Fällen, ist es praktischer, die sogenannte kontinuierliche Version zu wählen:

$$\lim_{x\to a} f(x) = s \iff \forall\ \varepsilon > 0\ \exists\, \delta > 0 \text{ mit : } x \in D,\ |x-a| < \delta\ \Rightarrow\ |f(x)-s| < \varepsilon.$$

Beweis der Äquivalenz:

- Gelte das kontinuierliche Kriterium:

  Sei (x_n) eine Folge aus D mit $x_n \to a$, und sei $\varepsilon > 0$ gegeben .

  $$\Rightarrow\ \exists\, N \quad \text{mit} \quad |x_n - a| < \delta \quad \forall\ n > N$$

  $$\underbrace{\Rightarrow}_{\text{kont. Krit.}} \ |f(x_n) - s| < \varepsilon \quad \forall\, n > N$$

  $$\Rightarrow f(x_n) \to s.$$

- Gelte das kont. Kriterium nicht, d.h. $\exists\ \varepsilon_0 > 0$ mit:

  Zu jedem $\delta > 0\ \exists\, x \in D$ mit $|x - a| < \delta$, aber $|f(x) - s| \geq \varepsilon_0$.

  Speziell für $\delta = \dfrac{1}{n} > 0$ $(n \in \mathbb{N})$ gibt es $x = x_n \in D$ mit

  $|x_n - a| < \dfrac{1}{n}$, aber $|f(x_n) - s| \geq \varepsilon_0 \ \Rightarrow\ x_n \to a$, aber $f(x_n) \not\to s$.

Weitere Beispiele:

(i) $f(x) = x^3,\ a = 2$:

$$\lim_{x\to 2} x^3 = 8\ (\ = f(2) : \ f \quad \text{ist stetig bei } 2\ )$$

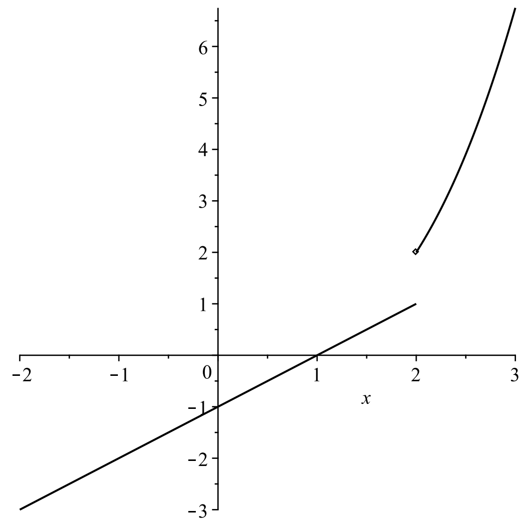
(ii) $f(x) = x^3,\ a \in \mathbb{R}$ beliebig:

$$\lim_{x\to a} x^3 = a^3 : \ f \quad \text{ist überall stetig.}$$

(iii) $f(x) = \begin{cases} x - 1 & x < 2 \\ x^2 - 2 & x \geq 2 \end{cases}$ bei $a = 2$:

$$\lim_{x\to 2-} f(x) = 1,\ \lim_{x\to 2+} f(x) = 2;\ \lim_{x\to 2} f(x) \quad \text{existiert nicht.}$$

Wegen $f(2) = 2$ ist f rechtsseitig stetig bei 2, nicht aber linksseitg stetig.



Das sind die gängigsten Unstetigkeitsstellen, sogenannte **Sprungstellen.**

(iv) $f(x) = \begin{cases} x & x \neq 3 \\ 4 & x = 3 \end{cases}$ bei $a = 3$

$\lim\limits_{x\to 3} = 3$ existiert zwar, aber wegen $f(3) = 4$ ist f nicht stetig bei $a = 3$.

Diese Unstetigkeit ist aber "hebbar", wenn man einfach $f(3) := 3$ umdefiniert.

(v) $f(x) = \begin{cases} \frac{x^2-1}{x-1} & x \neq 1 \\ \alpha & x = 1 \end{cases}$

Kann man α so wählen, dass f stetig wird bei $a = 1$?

Klar kann man das:

$$x \neq 1 \;\Rightarrow\; \frac{x^2-1}{x-1} = \frac{(x+1)(x-1)}{x-1} = x + 1 \to 2,\, x \to 1.$$

Wähle also sinnvollerweise $\alpha = 2$.

(vi) $f(x) = \dfrac{1}{x^2}$.

Kann man f stetig bei 0 ergänzen ? Nein, denn:

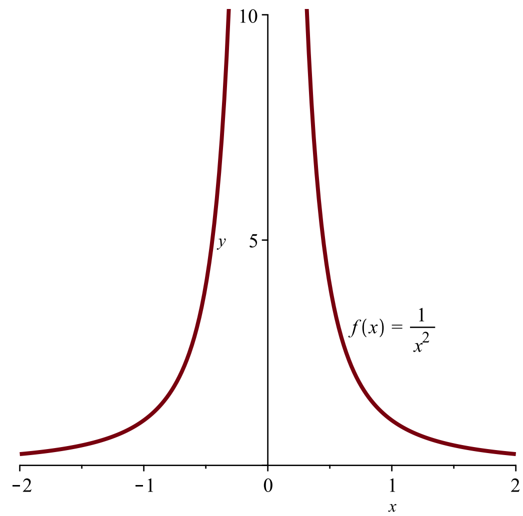
Ist x_n irgendeine Nullfolge, so ist $f(x_n)$ nach oben unbeschränkt

(z.B. ist $\dfrac{1}{n} \to 0$, aber $f\left(\dfrac{1}{n}\right) = n^2 \to \infty$) $\Rightarrow \lim\limits_{x\to 0} \dfrac{1}{x^2}$ existiert nicht.

Manchmal schreibt man hier noch

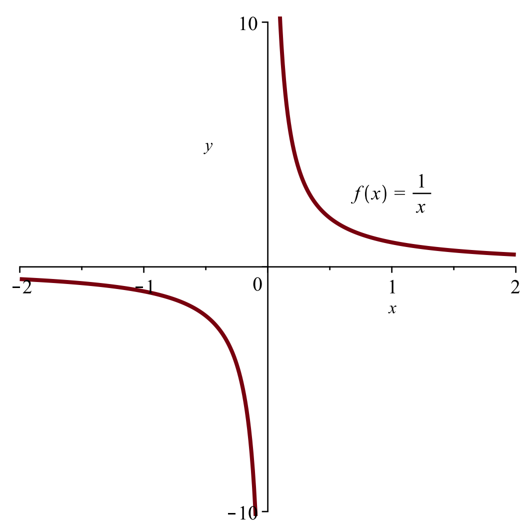
$$\lim_{x\to 0} \frac{1}{x^2} = \infty.$$

0 ist ein sogenannter **Pol.**



(vii) $f(x) = \dfrac{1}{x}$.

Hier ist 0 wieder eine Polstelle, etwas anders als in Beispiel (vi):



Hier würde man sagen:

$$\lim_{x\to 0+} \frac{1}{x} = \infty, \quad \lim_{x\to 0-} \frac{1}{x} = -\infty.$$

(viii) $f(x) = \sin\left(\dfrac{1}{x}\right)$.

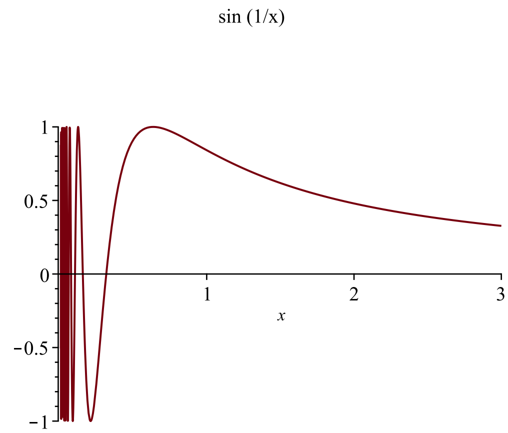
Kann man hier f bei 0 stetig ergänzen ? (Ein Pol liegt hier nicht vor, denn der Sinus ist beschränkt (durch 1)).

- $x_n = \dfrac{1}{n\pi} \to 0,\ f(x_n) = \sin(n\pi) = 0 \to 0$
- $x_n = \dfrac{1}{\frac{\pi}{2} + 2n\pi} \to 0,\ f(x_n) = \sin\dfrac{\pi}{2} = 1 \to 1$

- $w \in [-1,1],\ x_n = \dfrac{1}{\arcsin(2n\pi + w)} \to 0,\ f(x_n) = w \to w.$

Hier gibt es keine einseitigen Grenzwerte. Es handelt sich auch nicht um einen Pol. 0 ist eine sog. **Oszillationsstelle.**

Man könnte alle $y-$Werte im Intervall $[-1,1]$ als Häufungswerte bezeichnen.

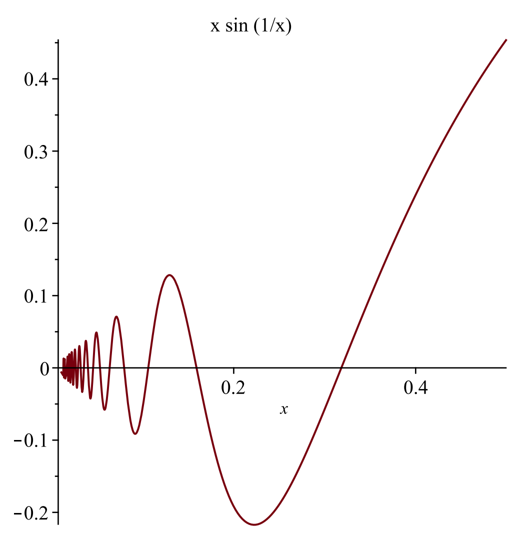


(ix) $f(x) = x \sin \dfrac{1}{x}$.

Hier existiert $\lim\limits_{x\to 0} f(x)$ und ist $=0$. Zum Beweis wenden wir hier zur Abwechslung mal das kontinuierliche Kriterium an:

Sei also $\varepsilon > 0$. Setze $\delta := \varepsilon\ (>0)$

$$x \in \mathbb{R},\ |x-0| < \delta \ \Rightarrow\ |f(x) - 0| = |x \sin \frac{1}{x}| \leq |x| < \varepsilon.$$



(x) Rechenregeln für Stetigkeit folgen unmittelbar aus den Rechenregeln für Folgengrenzwerte:

Summe, Differenz, Produkt und Quotient stetiger Funktionen sind wieder stetig (letztere falls auftretende Nenner $\neq 0$ sind). Auch die Hintereinanderausführung (Komposition) $f \circ g$ zweier stetiger Funktionen ist wieder stetig

(Beweis dazu: $x_n \to a \underbrace{\Rightarrow}_{g \text{ stetig}} g(x_n) \to g(a) \underbrace{\Rightarrow}_{f \text{ stetig}} f\Big(g(x_n)\Big) \to f\Big(g(a)\Big)$.)

(xi) Polynome, a^x, e^x, $\ln x$, $\cos x$, $\sin x$ etc. sind alle stetig auf ihrem gesamten Definitionsbereich.

Für Polynome sieht man das für x^n per Induktion und den Rest über die Rechenregeln. Für die anderen Funktionen ist das etwas technisch. Wir lassen diese Beweise aus.

Rationale Funktionen

$$\frac{p(x)}{q(x)},\ p, q \quad \text{Polynome}$$

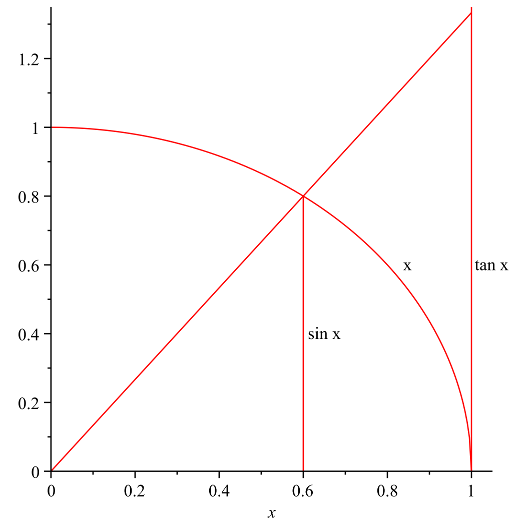
sind wegen der Rechenregeln überall stetig bis auf die Nullstellen des Nenners. Ausnahmen können auftreten wie in Beispiel (v). Dann muss allerdings auch das Zählerpolynom an dieser Stelle verschwinden.

(xi) Kann man $\dfrac{\sin x}{x}$ bei $a = 0$ stetig ergänzen?

TR: $\dfrac{\sin 0.01}{0.01} = 0.9999833$. Wir vermuten stark:

$$\lim_{x \to 0} \frac{\sin x}{x} = 1.$$

Betrachte dazu folgende Skizze:



Misst man x im Bogenmaß, so gilt offenbar für $0 < x < \frac{\pi}{2}$

$$\sin x \leq x \leq \tan x = \frac{\sin x}{\cos x}$$

$$\Rightarrow \ \cos x \leq \frac{\sin x}{x} \leq 1.$$

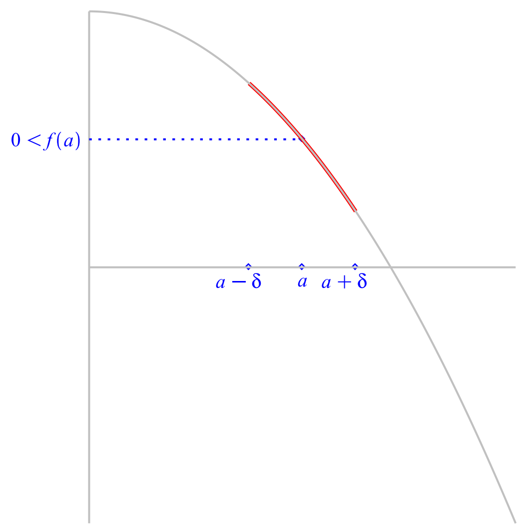
Für $x \to 0$ ist aber auch $\cos x \to \cos 0 = 1$.

§ 4 Eigenschaften stetiger Funktionen

Wir diskutieren im folgenden, was die Stetigkeit einer Funktion $f(x)$ an ein er Stelle a oder auf einem ganzen Intervall I ausmacht, was also die stetigen Funktionen abgrenzt gegenüber irgendwelchen anderen Funktionen.

- Stetige Funktionen wechseln nicht abrupt ihr Vorzeichen, genauer:

  Ist $f : I \to \mathbb{R}$ stetig bei $a \in I$ und gilt $\boxed{f(a) > 0,}$ so gibt es ein $\delta > 0$ mit
  $$f(x) > 0 \quad \forall \ x \in (a - \delta, a + \delta) \ (\cap I).$$



Beweis:

Das sieht man, wenn man im kontinuierlichen Kriterium $\varepsilon := f(a)$ setzt. Die Bedingung
$$|f(x) - f(a)| < f(a)$$
bedeutet ausgeschrieben ja
$$-f(a) < f(x) - f(a) < f(a),$$
insbesondere
$$f(x) > f(a) - f(a) = 0,$$
und das für alle x mit $|x - a| < \delta.$

Diese harmlose Beobachtung hat immense praktische Bedeutung:

Wenn nämlich ein stetig verlaufender Prozess vorliegt, kann es nicht zu plötzlichem krassem Abändern des Kurses kommen.

Bei unstetigem Verhalten sind sprunghafte Veränderungen vorstellbar bis hin zu chaotischen Verhältnissen (vgl. etwa das Beispiel $\sin\left(\frac{1}{x}\right)$.)

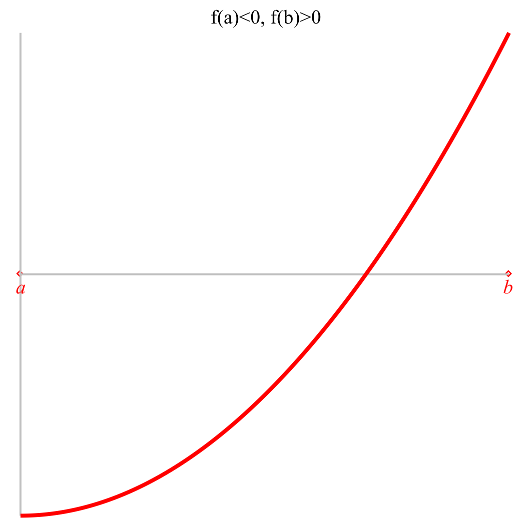
Zu ähnlichen Effekten kann es kommen, wenn die Inputdaten bei einem

Prozess nicht stetig weiterverabeitet werden. Kleine Abweichungen (z.B. bei Messfehlern) können dann zu völlig anderen Verläufen führen.

Stetigkeit führt also zu glattem, überschaubarem, kalkulierbarem, vielleicht auch einschläfernd wirkendem Verhalten.

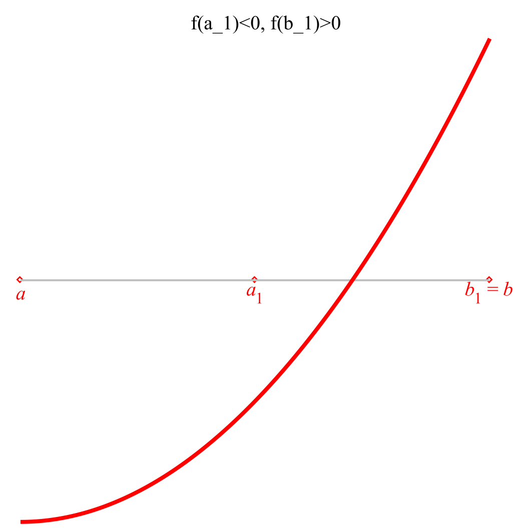
- Ist $f : [a,b] \to \mathbb{R}$ stetig auf $I = [a,b]$ mit $f(a) > 0$ und $f(b) < 0$ (oder umgekehrt), so besitzt f in (a,b) mindestens eine Nullstelle.
  Beweis:

  Um dies zu sehen, kann man konstruktiv vorgehen: Wir führen dies durch für den Fall $f(a) < 0,\ f(b) > 0$:



  Betrachte den Mittelpunkt $\dfrac{a+b}{2}$.

  Ist $f\left(\dfrac{a+b}{2}\right) < 0$, so schiebt man den linken Randpunkt a in diesen Mittelpunkt (und nennt ihn a_1) und lässt den rechten Randpunkt b unverändert (nennt ihn jetzt b_1).



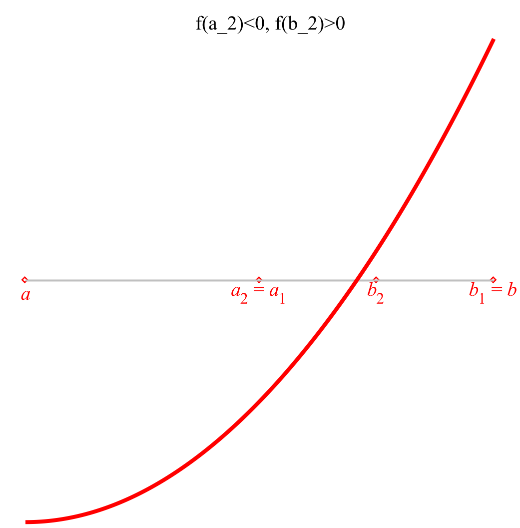
Ist aber $f(\frac{a+b}{2}) \geq 0$, so macht man das genau umgekehrt :

$$a_1 = a, \; b_1 = \frac{a+b}{2}.$$

Nach diesem Schritt haben wir für das neue Intervall $[a_1, b_1]$ dieselbe Situation:

$$f(a_1) \leq 0, \; f(b_1) \geq 0.$$

Jetzt wiederhole man das Verfahren iterativ. Im 2. Schritt haben wir hier tatsächlich die zweite Situation:



Wir bekommen dann Folgen (a_n) von linken und (b_n) von rechten Intervallgrenzen, die dann gegen einen gemeinsamen Wert x_0 konvergieren.

Das ist das Intervallschachtelungsprinzip. Beachte dabei:

$$b_n - a_n = \frac{b-a}{2^n} \to 0.$$

Nach Konstruktion ist aber

$$f(a_n) \leq 0 \leq f(b_n).$$

Für $n \to \infty$ folgt dann (mit der Stetigkeit von f) :

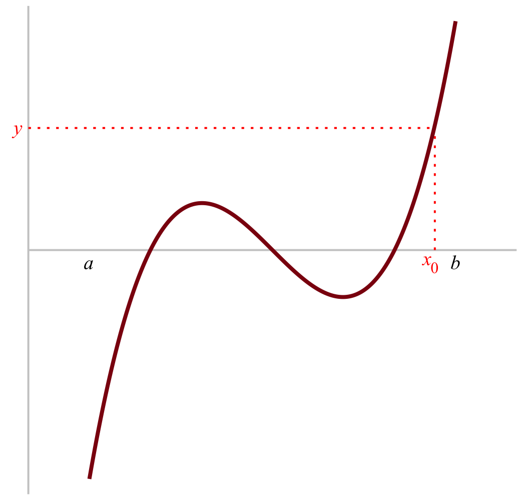
$$f(x_0) = 0.$$

Das hier verwendete Verfahren heißt **Bisektionsverfahren.**Mit ihm kann man Nullstellen stetiger Funktionen konstruktiv bestimmen (vorausgesetzt, man kann die Ausgangssituation $f(a)$ und $f(b)$ haben unterschiedliches Vorzeichen herstellen). Sehr effizient ist es allerdings nicht: Um eine Dezimalstelle zu gewinnen, reichen nicht einmal drei Iterationsschritte, denn

$$\left(\frac{1}{2}\right)^3 = \frac{1}{8} > 0.1$$

Allgemeiner gilt: Ist y ein Wert zwischen $f(a)$ und $f(b)$, und ist f stetig auf $[a,b]$, so gibt es ein $x_0 \in (a,b)$ mit $f(x_0) = y$. Das ist der **Zwischenwertsatz.**

Illustration zum Zwischenwertsatz



Zum Beweis wende man das Nullstellenergebnis an auf die Funktion $f(x) - y$.

Letztlich ist dies der tiefere Grund dafür, dass wir $\sqrt[n]{x}$, $\ln x$ etc definieren konnten, wie wir es in Kapitel I getan haben: Die Graphen von x^n bzw. e^x haben keine Lücken im Wertebereich. Bei unstetigen Funktionen könnten wir dies alles nicht machen.

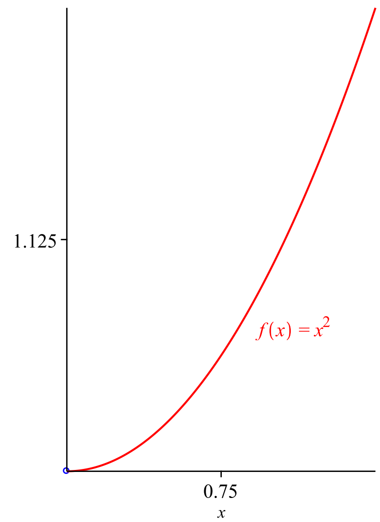
- Stetige Funktionen haben in einem abgeschlossenen und beschränkten Intervall $I = [a,b]$ jeweils Maximal- und Minimalstellen.

  Dieser eher theoretische Satz geht auf **Weierstraß** zurück. Z.B. ist

$$\min_{[-1,3]} x^2 = 0 = f(0) \quad \text{und} \quad \max_{[-1,3]} x^2 = 9 = f(3).$$

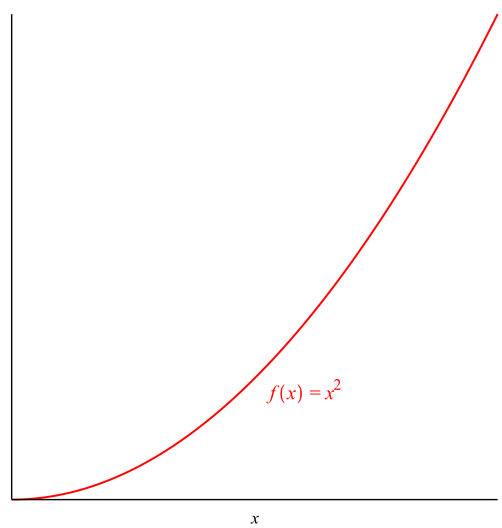
Beachte, dass man auf keine dieser Voraussetzungen verzichten kann, wie folgende Beispiele zeigen:

- I nicht abgeschlossen: Sei $I = (0,1]$, $f(x) = x^2$.



Hier existiert $\min_{x \in I} f(x)$ nicht.

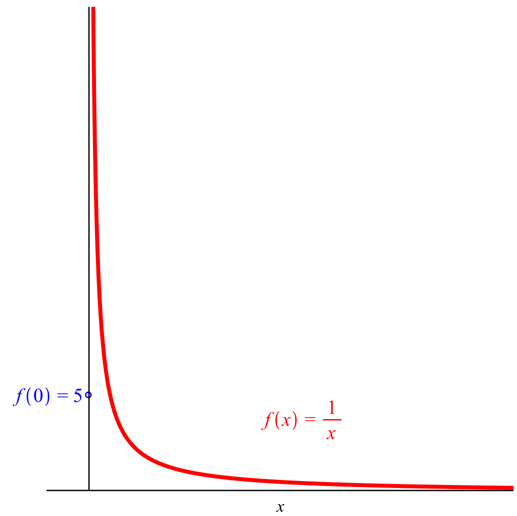
- I nicht beschränkt: Sei $I = [0, \infty)$, $f(x) = x^2$.



Hier existiert $\max_{x \in I} f(x)$ nicht.

- f nicht stetig. Sei $I = [0, 1]$ und

$$f(x) := \begin{cases} \frac{1}{x} & x \neq 0 \\ 5 & x = 0 \end{cases}$$



Hier ist f (nach oben) unbeschränkt.

Bemerkung:

Im Satz von Weierstraß ist von der Existenz von maximalen und minimalen Werten von Teilmengen reeller Zahlen die Rede. Das ist keine Selbstverständlichkeit: Definiert man nämlich für $A \subset \mathbb{R}$

$$M = \max A :\Longleftrightarrow M \in A \quad \text{und} \quad M \geq a \quad \forall\, a \in A$$

das **Maximum** von A und als Gegenstück das **Minimum** von A über

$$m = \min A :\Longleftrightarrow m \in A \quad \text{und} \quad m \leq a \quad \forall\, a \in A,$$

so ist z.B. für $A = [0,1]$ $(= \{x \in \mathbb{R} :\ 0 \leq x \leq 1\})$

$$\max A = 1 \quad \text{und} \quad \min A = 0.$$

Betrachtet man hingegen $A = [0,1)$ $(= \{x \in \mathbb{R} :\ 0 \leq x < 1\})$, so ist $\min A = 0$, $\max A$ existiert jedoch nicht.

Dennoch spielt die Zahl 1 für diese Menge eine Rolle: 1 ist nämlich eine obere Schranke von A, denn $1 \geq a \quad \forall\, a \in A$. Aber unter allen oberen Schranken ist 1 die kleinste. Diese kleinste obere Schranke nennt man dann das **Supremum** von A: $1 = \sup A$.

Natürlich gibt es auch hier ein Gegenstück: Eine untere Schranke von A ist eine Zahl s mit $s \leq a \quad \forall\, a \in A$. Gibt es unter allen unteren Schranken von A eine größte s_0, so nennt man diese das **Infimum** von A: $s_0 = \inf A$.

Supremum und Infimum sind also "Ersatzbegriffe" für Minimum und Maximum, sollten diese nicht existieren. Im Falle, dass $\sup A \in A$ gilt, existiert auch $\max A$, und es ist $\max A = \sup A$. (ebenso für $\inf A$.)

Es ist nun wieder genau die Vollständigkeit von $\mathbb{R}$, die garantiert, dass $\sup A$ und $\inf A$ immer existieren, vorausgesetzt, $A \neq \emptyset$ und A ist nach oben (bzw. nach unten) beschränkt.

Kapitel III
Differenzialrechnung

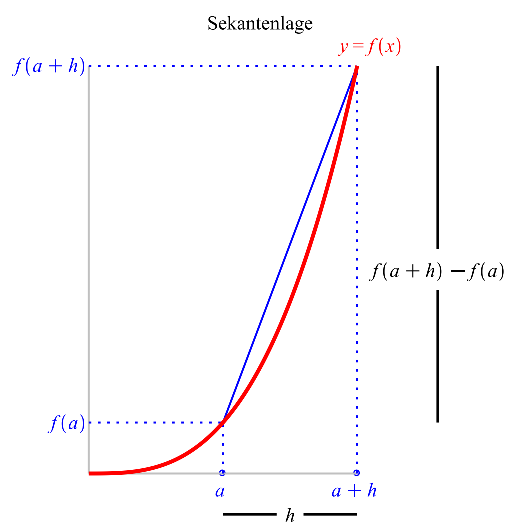
§ 1 Der Begriff der Ableitung

Dieses mächtige Instrument zur Charakterisierung von Funktionsverläufen wurde von Isaac Newton (1643-1727) und Gottfried Wilhelm Leibniz (1646-1716) begründet.

Im Kern will man dabei das lokale Verhalten einer Funktion an einer Stelle a studieren. Ein erster Versuch ist mit dem Begriff der Stetigkeit gelungen:

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

Für Anwenderzwecke reicht das i.a. noch nicht aus. Wir nehmen daher eine Stelle a und eine "benachbarte" Stelle $a + h$ und betrachten die Veränderung $f(a + h) - f(a)$ (absolut)



bzw. die relative Veränderung

$$\frac{f(a + h) - f(a)}{h}.$$

Dies ist geometrisch die Steigung der eingezeichneten Strecke (Sekante). Manchmal, wie z.B. beim Straßenverkehr, wird diese auch in % angegeben. (Was heißt z.B. eine Steigung bzw. ein Gefälle von 5 %?) .

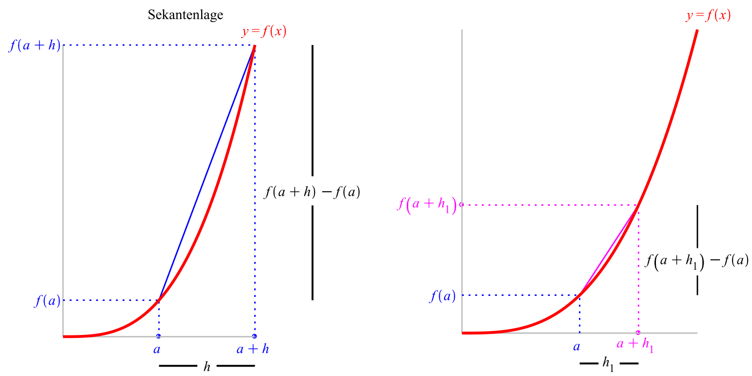
Man nennt

$$\frac{f(a + h) - f(a)}{h}$$

auch den **Differenzenquotienten** von f bei a mit Inkrement h .

Das Unschöne dabei ist, dass man von der Wahl des Inkrements h abhängig ist. Spiegelt etwa $h = 1$ das wider, was man beschreiben möchte, oder doch eher $h = 0.1$ oder?

Wir skizzieren die Situation für h und einen kleineren Wert h_1 :



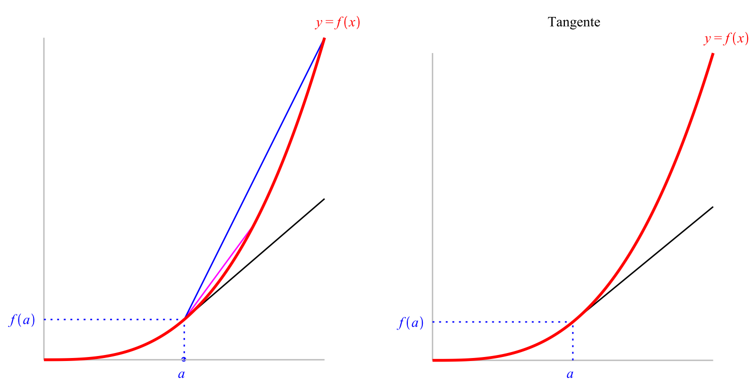
Diese Sorge ist man los, wenn es den Grenzwert

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

gibt. In diesem Falle nennen wir f **differenzierbar** bei a und bezeichnen diesen Grenzwert mit

$$f'(a) \quad \text{oder} \quad \frac{df}{da} \quad \text{oder} \quad \frac{d}{dx} f(x) \Big|_{x=a} .$$

Geometrisch "rutschen" bei diesem Grenzprozess die Sekanten für $h \to 0$ in Tangentiallage:



Deswegen nennt man $f'(a)$ auch die **Tangentensteigung** bei a .

Bemerkungen und Beispiele:

(i) Beispiel:

$f(x) = x^2, \quad a = 3:$

$$\frac{f(3+h) - f(3)}{h} = \frac{(3+h)^2 - 9}{h} = \frac{6h + h^2}{h} = 6 + h \to 6, \ h \to 0$$

$\Rightarrow$ f ist differenzierbar bei $a = 3$ mit $f'(3) = 6$

(ii) Beispiel:

$f(x) = x^2,\ a$ beliebig:

$\Rightarrow\ \dfrac{f(a+h)-f(a)}{h} = \dfrac{(a+h)^2 - a^2}{h} = 2a + h \to 2a.$

Wenn es $f'(a)$ an jeder Stelle a gibt, so nennen wir die Funktion

$$a \to f'(a)$$

die **1.Ableitung(sfunktion)** von f und benutzen i.a. wieder den Original-Variablennamen x statt a. Wir schreiben also:

$$f(x) = x^2 \ \Rightarrow\ f'(x) = 2x \quad \text{oder kurz} \quad \left(x^2\right)' = 2x.$$

(iii) Tangentengleichung: $t(x) = f'(a)(x-a) + f(a)$;

z.B. $f(x) = x^2,\ a = 3 \ \Rightarrow\ t(x) = 6(x-3) + 9 = 6x - 9.$

(iv) Beispiel:

$f(x) = \dfrac{1}{x},\ a \neq 0$

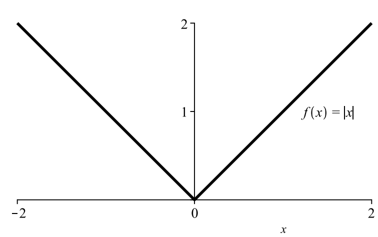
$\Rightarrow\ f'(a) = \lim\limits_{h\to 0} \dfrac{\frac{1}{a+h} - \frac{1}{a}}{h} = \lim\limits_{h\to 0} \dfrac{a-(a+h)}{a(a+h)h} = \lim\limits_{h\to 0} \dfrac{-1}{a(a+h)} = -\dfrac{1}{a^2},$

also

$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2} \quad \forall\ x \neq 0.$$

(iv) Beispiel:

$f(x) = |x|,\ a = 0:$



$$\frac{f(0+h)-f(0)}{h} = \frac{|h|}{h} = \begin{cases} 1 & h > 0 \\ -1 & h < 0 \end{cases}$$

$\Rightarrow\ f'(0)$ existiert nicht. Man kann zwar von linksseitiger Ableitung -1 und rechtsseitiger Ableitung $+1$ reden. Das wollen wir hier aber nicht vertiefen.

Das Beispiel zeigt allerdings auch: Es gibt stetige Funktionen, die nicht differenzierbar sind.

Die Umkehrung ist aber richtig: Alle differenzierbaren Funktionen sind auch stetig. Das sieht man sofort an der Definition: Wenn f differenzierbar ist bei a, so besitzt der Ausdruck

$$\frac{f(a+h)-f(a)}{h}$$

einen endlichen Grenzwert für $h \to 0$. Das kann aber nur dann möglich sein, wenn nicht nur der Nenner h gegen 0 konvergiert, sondern auch der Zähler, d.h. $f(a+h) \to f(a)$, $h \to 0$, und dass ist gerade die Stetigkeit bei a.

§ 2 Die Technik des Differenzierens

A. Ableitungen der elementaren Funktionen:

$$f'(a) = \lim_{h\to 0}\frac{f(a+h)-f(a)}{h} = \lim_{x\to a}\frac{f(x)-f(a)}{x-a}$$
$$(\text{mit}\quad h = x-a)$$
$$\Big(x^n\Big)' = nx^{n-1} \quad (n\in\mathbb{N} \quad \text{oder} \quad \mathbb{Z} \quad \text{oder} \quad \mathbb{R});$$
$$\Big(e^x\Big)' = e^x; \quad \Big(a^x\Big)' = \ln a\cdot a^x \quad (a>0)$$
$$\Big(\sin x\Big)' = \cos x; \quad \Big(\cos x\Big)' = -\sin x;$$
$$\Big(\tan x\Big)' = \frac{1}{\cos^2 x} = 1+\tan^2 x;$$
$$\Big(\ln x\Big)' = \frac{1}{x}.$$
$$\Big(\arcsin x\Big)' = \frac{1}{\sqrt{1-x^2}}; \quad \Big(\arctan x\Big)' = \frac{1}{1+x^2}$$

Wir beweisen diese nach und nach.

(i) $f(x) = x^n$:

$$\frac{(a+h)^n - a^n}{h} \underbrace{=}_{\text{Binomi}} \frac{a^n + na^{n-1}h + \binom{n}{2}a^{n-2}h^2 + \ldots + h^n - a^n}{h} =$$
$$= na^{n-1} + h\binom{n}{2}a^{n-1} + \ldots + h^{n-1} \to na^{n-1},\ h\to 0.$$

(ii) $f(x) = \sin x$:

$$\frac{\sin(a+h)-\sin a}{h} = \frac{\sin a\cos h + \cos a \sin h - \sin a}{h} =$$
$$= \sin a\left(\frac{\cos h - 1}{h}\right) + \cos a\,\frac{\sin h}{h} \to \cos a,\ h\to 0,$$

denn:

- $\lim\limits_{h\to 0}\dfrac{\sin h}{h} = 1$ (s.Kapitel II)
- $\lim\limits_{h\to 0}\dfrac{\cos h - 1}{h} = 0$, denn:

$$\cos 2x = \cos^2 x - \sin^2 x = 1 - 2\sin^2 x \ \Rightarrow\ \cos x = 1 - 2\sin^2\left(\frac{x}{2}\right)$$
$$\Rightarrow\ \frac{\cos h - 1}{h} = \frac{-2\sin^2(\frac{h}{2})}{h} = \underbrace{-\sin(\frac{h}{2})}_{\to 0}\cdot\underbrace{\frac{\sin(\frac{h}{2})}{\frac{h}{2}}}_{\to 1} \to 0.$$

(iii) $\cos x$:

Dies zeigt man ganz analog.

(iv) $f(x) = e^x$:

$$\frac{e^{a+h} - e^a}{h} = \frac{e^a e^h - e^a}{h} = e^a \cdot \underbrace{\frac{e^h - 1}{h}}_{\to 1 \text{ über die Reihe (später)}} \to e^a, \ h \to 0.$$

B. Linearitätsregel

$$\boxed{c \in \mathbb{R} \ \Rightarrow \ \Big(c \cdot f\Big)' = c \cdot f'}$$

Beweis:

$$\frac{cf(a+h) - cf(a)}{h} = c \cdot \frac{f(a+h) - f(a)}{h} \to c \cdot f'(a), \ h \to, 0.$$

So ist z.B.

$$\Big(4x^7\Big)' = 28x^6; \quad \Big(5\cos x\Big)' = -5\sin x.$$

C. Additionsregel:

$$\boxed{\Big(f + g\Big)' = f' + g'}$$

Beweis:

$$\frac{f(a+h) + g(a+h) - f(a) - g(a)}{h} = \frac{f(a+h) - f(a)}{h} + \frac{g(a+h) - g(a)}{h}$$

$$\to f'(a) + g'(a), \ h \to 0.$$

Damit folgt auch (wegen B. mit $c = -1$)

$$\boxed{\Big(f - g\Big)' = f' - g'}$$

So ist z.B.

$$\Big(x^3 - x^2\Big)' = 3x^2 - 2x, \quad \Big(\sin x + e^x\Big)' = \cos x + e^x$$

oder (kombiniert mit B.)

$$\Big(7x^6 - 5x^4 + 2x - 8\Big)' = 42x^5 - 20x^3 + 2, \quad \Big(2\sin x - 8\cos x\Big)' = 2\cos x + 8\sin x.$$

D. Produktregel:

$$\boxed{\Big(f \cdot g\Big)' = f' \cdot g + f \cdot g'}$$

Beweis:

$$\frac{f(a+h)g(a+h) - f(a)g(a)}{h} = f(a+h)\ \frac{g(a+h) - g(a)}{h} + g(a)\ \frac{f(a+h) - f(a)}{h}$$

$$\to f(a)g'(a) + g(a)f'(a),\ h \to 0.$$

So ist z.B.

$$\Big(x^3 \sin x\Big)' = 3x^2 \sin x + x^3 \cos x, \quad \Big(e^x \cos x\Big)' = e^x \cos x - e^x \sin x.$$

E. Quotientenregel:

$$\boxed{\Big(\frac{f}{g}\Big)' = \frac{f'g - fg'}{g^2}}$$

Beweis:

- $$\frac{\frac{1}{g(a+h)} - \frac{1}{g(a)}}{h} = \frac{-\Big(g(a+h) - g(a)\Big)}{g(a+h)g(a)h} \to -\frac{g'(a)}{g^2(a)}\ h \to 0 \quad \text{und}$$
- $$\Big(\frac{f}{g}\Big)' = \Big(f \cdot \frac{1}{g}\Big)' \underbrace{=}_{\text{mit D.}} f' \cdot \frac{1}{g} + f \cdot \Big(-\frac{g'}{g^2}\Big) = \frac{f'g - fg'}{g^2}.$$

So ist z.B.
$$\Big(\frac{x^3 - x}{x^5 + 1}\Big)' = \frac{(3x^2 - 1)(x^5 + 1) - 5x^4(x^3 - x)}{(x^5 + 1)^2};$$

$$\Big(\tan x\Big)' = \Big(\frac{\sin x}{\cos x}\Big)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} \quad \text{(auch als } 1 + \tan^2 x \text{ schreibbar).}$$

$$n \in \mathbb{N} \ \Rightarrow\ \Big(x^{-n}\Big)' = \Big(\frac{1}{x^n}\Big)' = \frac{0 \cdot x^n - nx^{n-1}}{x^{2n}} = -nx^{n-1-2n} = -nx^{-n-1}.$$

F. Kettenregel:

$$\boxed{\Big(f \circ g\Big)' = (f' \circ g) \cdot g'}$$

oder (ausführlich geschrieben)

$$\boxed{\Big(f(g(x))\Big)' = f'(g(x)) \cdot g'(x)}$$

Beweis

$$\frac{f\Big(g(a+h)\Big) - f\Big(g(a)\Big)}{h} = \frac{f\Big(g(a+h)\Big) - f\Big(g(a)\Big)}{g(a+h) - g(a)} \cdot \frac{g(a+h) - g(a)}{h}$$

$$\to f'\Big(g(a)\Big) \cdot g'(a),\ h \to 0.$$

So ist z.B.:

$$\Big(e^{(x^2)}\Big)' = e^{(x^2)} \cdot 2x; \quad \Big(\cos 3x\Big)' = -3 \sin 3x; \quad \Big(e^{-x}\Big)' = -e^{-x}$$

$$\Big((x^{17} + x^{13})^6\Big)' = 6(x^{17} + x^{13})^5 \cdot (17x^{16} + 13x^{12})$$

G. Umkehrfunktionen:

$$\boxed{\Big(f^{-1}\Big)' = \frac{1}{f' \circ f^{-1}},\ \text{sofern}\ \ f' \neq 0\ .}$$

Beweis:

$$f\Big(f^{-1}(x)\Big) = x \underbrace{\Rightarrow}_{\text{Kettenregel}} f'\Big(f^{-1}(x)\Big)\Big(f^{-1}\Big)'(x) = 1 \ \Rightarrow\ \Big(f^{-1}\Big)'(x) = \frac{1}{f'\Big(f^{-1}(x)\Big)}.$$

(Das ist streng genommen nicht ganz korrekt: Die Differenzierbarkeit von f^{-1} wird hier vorausgesetzt.)

Beispiele:

- $f(x) = e^x \ \Rightarrow\ \Big(\ln x\Big)' = \dfrac{1}{e^{\ln x}} = \dfrac{1}{x}.$
- $f(x) = \sin x \ \Rightarrow\ \Big(\arcsin x\Big)' = \dfrac{1}{\cos(\arcsin x)} \underbrace{=}_{\cos^2 + \sin^2 = 1} \dfrac{1}{\sqrt{1 - x^2}}.$
- $(\arctan x)' = \dfrac{1}{\tan'(\arctan x)} = \dfrac{1}{1 + \tan^2(\arctan x)} = \dfrac{1}{1 + x^2}$

H. ... was noch fehlt:

$$\boxed{\Big(x^\alpha\Big)' = \alpha x^{\alpha-1}, \quad \alpha \in \mathbb{R}}$$

Beweis:

$$\Big(x^\alpha\Big)' = \Big(e^{\alpha \ln x}\Big)' = e^{\alpha \ln x} \cdot \frac{\alpha}{x} = x^\alpha \cdot \frac{\alpha}{x} = \alpha x^{\alpha-1}.$$

Das bedeutet konkret etwa:

$$\Big(\frac{1}{x}\Big)' = \Big(x^{-1}\Big)' = -x^{-2} = -\frac{1}{x^2};$$

oder

$$\Big(\sqrt{x}\Big)' = \Big(x^{\frac{1}{2}}\Big)' = \frac{1}{2}x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}}$$

$$\boxed{\Big(a^x\Big)' = a^x \cdot \ln a.}$$

Beweis:

$$\Big(a^x\Big)' = \Big(e^{x \ln a}\Big)' = e^{x \ln a} \cdot \ln a = \ln a \cdot a^x$$

$$\Big(x^x\Big)' = \Big(e^{x \ln x}\Big)' = e^{x \ln x} \cdot \Big(\ln x + x \cdot \frac{1}{x}\Big) = x^x \cdot (\ln x + 1).$$

I. Gemischte Beispiele:

Beachtet man alle diese Ableitungsregeln, kann man - bei ausreichender Konzentration - praktisch jeden Ausdruck ableiten, den man mit den elementaren Funktionen und den einfachen arithmetischen Operationen hinschreiben kann. Wir geben hierfür noch ein Beispiel: Man berechne die Ableitung von

$$f(x) = \frac{e^{(\sin 3x - \sqrt[3]{\ln x})} \cdot \cos \frac{x}{\ln x}}{\sin 2x \cdot \sin(x^2) \cdot \ln(\ln x)}$$

Der Ausdruck ist zunächst einmal ein Quotient. Wenn wir die Quotientenregel anwenden, brauchen wir die Ableitungen von Zähler und Nenner.

- Der Zähler $Z(x)$ ist ein Produkt. Wenn wir die Produktregel anwenden wollen, brauchen wir die Ableitungen der beiden Faktoren.
  - der linke Faktor $e^{(\sin 3x - \sqrt[3]{\ln x})}$ ist zunächst mit der Kettenregel zu bearbeiten:

$$\Big(e^{(\sin 3x - \sqrt[3]{\ln x})}\Big)' = e^{(\sin 3x - \sqrt[3]{\ln x})} \cdot \Big(\sin 3x - (\ln x)^{1/3}\Big)' =$$

$$\underbrace{=}_{\text{nochmal Kettenregel}} \quad e^{(\sin 3x - \sqrt[3]{\ln x})} \cdot \Big(3\cos 3x - \frac{1}{3}(\ln x)^{-2/3} \cdot \frac{1}{x}\Big)$$

  - Der rechte Faktor $\cos \dfrac{x}{\ln x}$ ist ebenfalls zuerst mit der Kettenregel zu bearbeiten:

$$\Big(\cos \frac{x}{\ln x}\Big)' = -\sin \frac{x}{\ln x} \cdot \Big(\frac{x}{\ln x}\Big)' \underbrace{=}_{\text{Quotientenregel}} -\sin \frac{x}{\ln x} \cdot \frac{\ln x - x \cdot \frac{1}{x}}{(\ln x)^2}$$

  Die Ableitung des Zählers ist also

$$Z'(x) = e^{(\sin 3x - \sqrt[3]{\ln x})} \cdot \Big(3\cos 3x - \frac{1}{3}(\ln x)^{-2/3} \cdot \frac{1}{x}\Big) \cdot \cos \frac{x}{\ln x} +$$

$$+ \, e^{(\sin 3x - \sqrt[3]{\ln x})} \cdot \Big(-\sin \frac{x}{\ln x} \cdot \frac{\ln x - 1}{(\ln x)^2}\Big)$$

- Der Nenner $N(x)$ ist ein Produkt. Wir brauchen also die Ableitung der beiden Faktoren
  - Der linke Faktor $\sin 2x \cdot \sin(x^2)$ ist selber ein Produkt mit Ableitung

$$\Big(\sin 2x\Big)' \cdot \sin(x^2) + \sin(2x)\Big(\sin(x^2)\Big)' =$$

$$\underbrace{=}_{\text{mit Kettenregel}} \quad 2\cos 2x \cdot \sin(x^2) + \sin 2x \ \cos(x^2) \cdot (2x)$$

  - Der 2. Faktor $\ln(\ln x)$ besitzt nach der Kettenregel die Ableitung

$$\frac{1}{\ln x} \cdot \frac{1}{x}.$$

  Die Ableitung des Nenners ist also

$$N'(x) = \Big(2\cos 2x \cdot \sin(x^2) + \sin 2x \ \cos(x^2) \cdot (2x)\Big) \cdot \ln(\ln x) + \sin 2x \cdot \sin(x^2) \cdot \frac{1}{x \ln x}$$

Jetzt muss nur noch zusammengebaut werden:

$$f'(x) = \frac{Z'(x)N(x) - Z(x)N'(x)}{N^2(x)},$$

was wir aus Platzgründen nicht mehr ausschreiben.

§ 3 Fehlerrechnung und die Mittelwertsätze der Differenzialrechnung

Wir wollen den Ableitungsbegriff noch einmal unter einem anderen Aspekt betrachten: Wir hatten

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \underbrace{=}_{h:=x-a\to 0 \iff x\to a} \lim_{x\to a} \frac{f(x) - f(a)}{x - a}$$

Besteht also ein funktionaler Zusammenhang

$$y = f(x),$$

und misst man für die unabhängige Größe den Wert a, so ist

$$x - a$$

auffassbar als **(absoluter)** Fehler. Dieser überträgt sich auf den absoluten Fehler bei der abhängigen Göße y auf

$$f(x) - f(a) \approx f'(a) \cdot (x - a).$$

$f'(a)$ gibt also (in etwa) den Faktor an, der beim Fehler entsteht, wenn man von der unabhängigen Größe x zur abhängigen Größe y übergeht. Ein großer Wert von $|f'(a)|$ bedeutet also hohe Fehlersensibilität.

Oftmals stehen aber nicht die absoluten Fehler im Mittelpunkt, sondern die **relativen** Fehler

$$\frac{x-a}{a} \quad \text{bzw.} \quad \frac{f(x) - f(a)}{f(a)}.$$

Diese gibt man üblicherweise prozentual an.

Rechnerisch zeigt sich dies im Differenzierbarkeitsfall über

$$\frac{f(x) - f(a)}{f(a)} \approx \frac{a \cdot f'(a)}{f(a)} \cdot \frac{x-a}{a}$$

Man nennt die Größe

$$E(a) = a \cdot \frac{f'(a)}{f(a)}$$

die

Elastitizität

des funktionalen Zusammenhangs zwischen den Größen x und $y = f(x)$. Sie kann vom Niveau, also von der Stelle a abhängen oder auch nicht.

Beispiele:

(i) $f(x) = x^3$:

$$E(a) = a \cdot \frac{3a^2}{a^3} = 3$$

unabhängig von a. Ein Fehler von 1% bei x bewirkt also in etwa einen Fehler von 3% bei $y = f(x)$ (egal an welcher Stelle)

(ii) $f(x) = e^{2x}$:

$$E(a) = a \cdot \frac{2e^{2a}}{e^{2a}} = 2a,$$

also abhängig von a .

Ist $a \approx 1$, so ist $E(a) \approx 2$. Ein Fehler von 1% bei x bewirkt in diesem Bereich einen Fehler von ca. 2% bei y .

Ist aber $a \approx 10$, so bewirkt ein Fehler von 1% bei x einen Fehler von etwa 20% bei y.

Ein Ziel soll nun sein, von dieser unpräzisen Notation

$$f(x) - f(a) \approx f'(a)(x - a)$$

wegzukommen. Dies benötigt allerdings gewisse Vorbereitungen.

Definition:

Gegeben sei eine Funktion

$$f : I \to \mathbb{R}, \quad (I \subset \mathbb{R} \text{ ein Intervall.})$$

Eine Zahl $c \in I$ heißt **lokale Maximalstelle** (bzw.**lokale Minimalstelle)** von f , wenn es $\delta > 0$ gibt mit

$$f(x) \leq f(c) \quad \forall \; x \in I \quad \text{mit} \quad c - \delta < x < c + \delta)$$

$$(\text{ bzw. } \quad f(x) \geq f(c) \quad \forall \; x \in (c - \delta, c + \delta) \cap I.)$$

Die beiden Begriffe fassen wir zusammen zum Sammelbegriff der

lokalen Extremalstelle.

Satz 3.1 (notwendige Bedingung für lokale Extrema):
Es sei $f : [a,b] \to \mathbb{R}$ differenzierbar in einem

$$\boxed{\text{inneren Punkt}} \;\; c \in (a,b)\ .$$

Ist dann c lokale Extremalstelle von f , so gilt:

$$f'(c) = 0.$$

Beweis:

Sei o.B.d.A. c lokale Maximalstelle von f und sei $\delta > 0$ mit

$$f(x) \le f(c) \quad \forall \;\; x \in [a,b],\ c - \delta < x < c + \delta.$$

$$c < x < c+\delta \;\Rightarrow\; \frac{f(x) - f(c)}{x - c} \le 0 \;\; (\text{Zähler} \le 0, \text{Nenner} > 0) \underbrace{\Rightarrow}_{(x \to c+)} f'(c) \le 0$$

$$c-\delta < x < c \;\Rightarrow\; \frac{f(x) - f(c)}{x - c} \ge 0 \quad (\text{Zähler} \le 0, \text{Nenner} < 0) \underbrace{\Rightarrow}_{(x \to c-)} f'(c) \ge 0$$

Insgesamt folgt: $f'(c) = 0$.

Bemerkungen:

(i) Liegt eine Extremalstelle c **nicht** im Inneren von $[a,b]$, ist also einer der Randpunkte a oder b, so braucht $f'(c) = 0$ nicht zu gelten.

Beispielsweise besitzt

$$f(x) = x^2$$

im Intervall $[0,1]$ eine lokale Maximalstelle bei $c = 1$,

während $f'(1) = 2 \neq 0$ gilt.

Die "Hälfte" des Schlusses im Beweis zu Satz 3.1 ist indes machbar.

In der Tat ist $f'_-(1) \ge 0$.

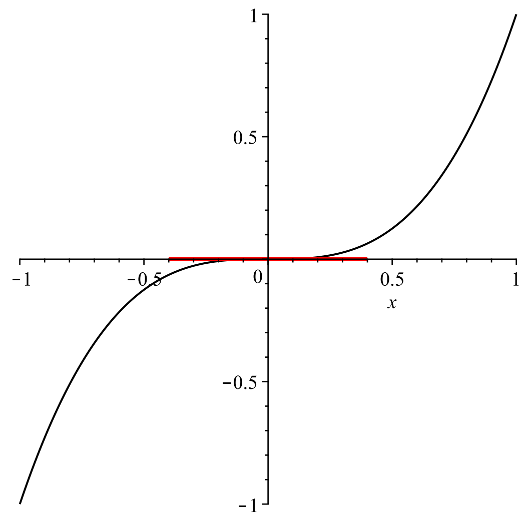
(ii) Hinreichend für lokale Extremalstellen ist $f'(c) = 0$ für $c \in (a, b)$ nicht. Betrachte hierzu das Beispiel

$$f(x) = x^3, \ x \in \mathbb{R} :$$

$f'(0) = 0$, aber 0 ist keine lokale Extremalstelle von f, denn:

$$\delta > 0 \ \Rightarrow \ f(\delta) = \delta^3 > 0; \ \ f(-\delta) = -\delta^3 < 0.$$

0 ist ein sogenannter **Sattelpunkt.**



(iii) Beispiel:

Wir betrachten die Funktion

$$f(x) = xe^{-x} \quad \text{im Intervall } [0, 2].$$

Diese besitzt sowohl ein (absolutes) Maximum als auch ein absolutes Minimum. Das ist von der Theorie her gesichert, denn f ist stetig und $[0, 2]$ ist abgeschlossen und beschränkt. Das war der Satz von Weierstraß.

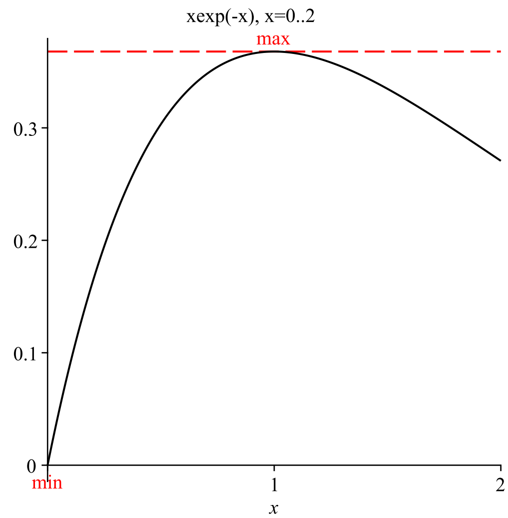
Als Extremalstellen c kommen dabei in Frage:

- der linke Randpunkt $c = 0$ (mit $f(0) = 0$),
- der rechte Randpunkt $c = 2$ (mit $f(2) = 2e^{-2}) \approx 0.27$ sowie
- die Nullstelle(n) von $f'(x) = e^{-x} - xe^{-x} = (1 - x)e^{-x}$, also $c = 1$ mit $f(1) = e^{-1} \approx 0.37$.

Wegen $e^{-1} > 2e^{-2} > 0$ folgt also :

$$\max\{ \ xe^{-x} : \ x \in [0, 2] \ \} = e^{-1} = f(1) \quad \text{und}$$

$$\min\{ \ xe^{-x} : \ x \in [0, 2] \ \} = 0 = f(0).$$



Satz 3.2 (Satz von Rolle):

Es sei

$$f : [a, b] \to \mathbb{R}$$

differenzierbar auf (a, b) und in den Randpunkten a und b (wenigstens) stetig. Weiter gelte

$$\boxed{f(a) = f(b).}$$

Dann gibt es (mindestens) ein $\xi \in (a, b)$ mit

$$f'(\xi) = 0.$$

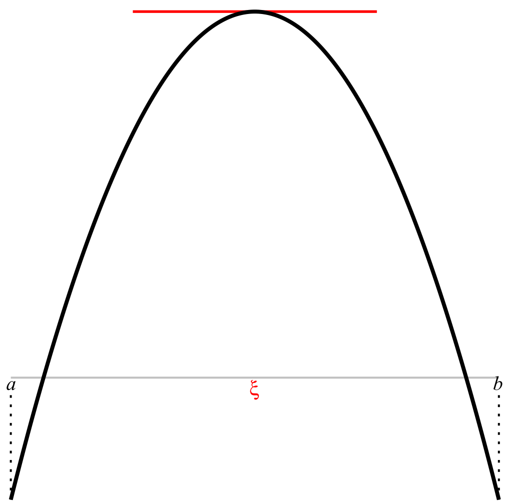
Beweis:

Wegen $f \in C[a, b]$ existieren $\min\limits_{x \in [a,b]} f(x)$ und $\max\limits_{x \in [a,b]} f(x)$.

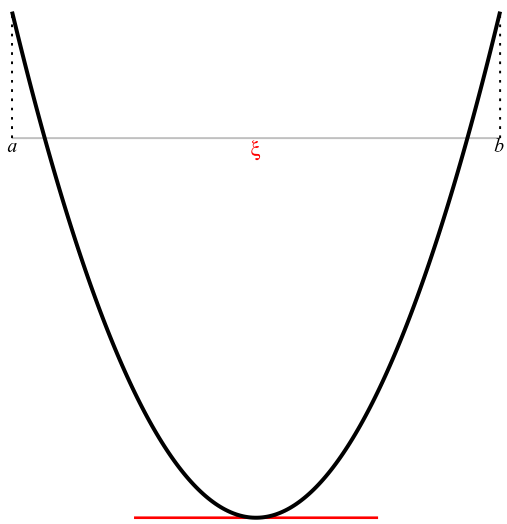
Liegen beide Extremalstellen am Rand, so ist wegen $f(a) = f(b)$ f konstant und damit $f' \equiv 0$.

Liegt aber eine der beiden Extremalstellen, etwa ξ im Inneren, so ist wegen Satz 3.1 $f'(\xi) = 0$.

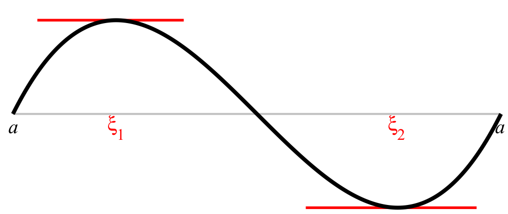
Die nichtkonstanten Situationen im Satz von Rolle können also unterschiedlich aussehen:



oder



oder



(also nicht unbedingt eindeutig).

Satz 3.3 (1. Mittelwertsatz der Differenzialrechnung):
Es sei

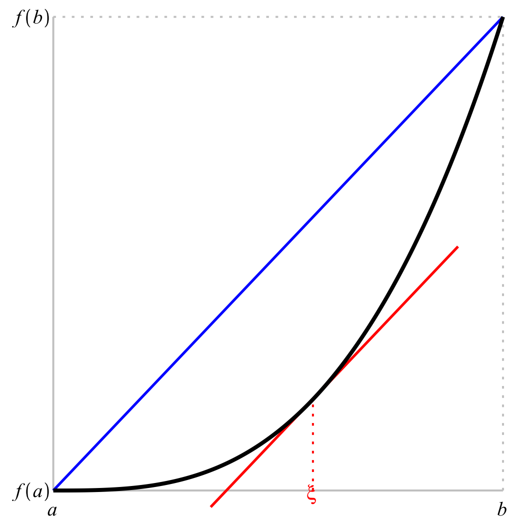
$$f : [a, b] \to \mathbb{R}$$

differenzierbar auf (a, b) und in den Randpunkten a und b immerhin stetig. Dann gibt es eine Stelle $\xi \in (a, b)$ mit

$$\frac{f(b) - f(a)}{b - a} = f'(\xi) \quad (\iff \quad f(b) - f(a) = f'(\xi)(b - a)).$$

(Das ist die oben angedeutete Präzisierung.)

Anschaulich bedeutet die erste Formulierung: Die Kurve mit der Gleichung $y = f(x)$ besitzt im Intervall (a, b) mindestens eine Stelle, an der die Tangente parallel zur Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ ist:



Satz 3.4 (2. Mittelwertsatz der Differenzialrechnung):
Es seien

$$f, g : [a, b] \to \mathbb{R}$$

differenzierbar auf (a, b) und in den Randpunkten a und b stetig, und es sei $g' \neq 0$ auf (a, b) . Dann gibt es $\xi \in (a, b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Beweis der Mittelwertsätze:

Der 1.MWS folgt aus dem 2.MWS mit $g(x) = x$ und $g' \equiv 1$.

Zum 2.MWS: Zunächst ist $g(a) \neq g(b)$. Andernfalls gäbe es nach dem Satz von Rolle ein $\xi \in (a,b)$ mit $g'(\xi) = 0$ im Widerspruch zu $g' \neq 0$ auf (a,b) .
Betrachte nun

$$\varphi : [a,b] \to \mathbb{R},\ \varphi(x) := f(x) - \frac{f(b)-f(a)}{g(b)-g(a)} \cdot g(x)$$

$\Rightarrow \varphi$ differenzierbar auf (a,b) , stetig in den Randpunkten a, b , und es gilt

$$\varphi(b) - \varphi(a) = f(b) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)} \Big(g(b) - g(a)\Big) = 0,$$

also

$$\varphi(a) = \varphi(b).$$

Nach dem Satz von Rolle gibt es dann $\xi \in (a,b)$ mit

$$0 = \varphi'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{g(b)-g(a)}\, g'(\xi),$$

woraus durch Umstellen die Behauptung folgt.

§ 4 Anwendungen der Mittelwertsätze (Kurvendiskussion)

Vorbemerkung:

Wir schließen zunächst an den 1. Mittelwertsatz der Differenzialrechnung an, der im wesentlichen besagt, dass die Differenz $f(b) - f(a)$ einer Funktion f kontrolliert werden kann durch

$$f(b) - f(a) = f'(\xi)\,(b - a),$$

wobei man von ξ letzlich nicht sonderlich viel weiß, außer dass es zwischen a und b liegt.

Dennoch ist dies ein mächtiges Werkzeug, um Differenzialrechnung anzuwenden. Immer dann, wenn man eine Funktion an zwei Stellen vergleichen will, sollte man an den 1. MWS denken.

Wir zeigen dies zunächst am Beispiel einer Ungleichung, etwa

$$\frac{y-x}{y} < \ln y - \ln x < \frac{y-x}{x} \quad \forall\ 0 < x < y.$$

Dies folgt aus $(\ln x)' = \dfrac{1}{x}$, also

$$\ln y - \ln x = \frac{y-x}{\xi} \quad \text{mit einem } \xi \text{ mit } x < \xi < y,$$

und daraus die Behauptung, denn $\dfrac{1}{\xi} < \dfrac{1}{x}$ und $\dfrac{1}{\xi} > \dfrac{1}{y}$.

Wir wollen nun den Mittelwertsatz benutzen, um (wahrscheinlich aus der Schule bekannte) Kriterien zur Kurvendiskussion zu beweisen. Wiederum heißt das, dass wir Funktionswerte an **unterschiedlichen** Stellen vergleichen wollen.

Was könnte das betreffen?

- Bereiche, in denen die Funktionswerte ansteigend bzw. abfallend sind (Monotonie)
- Stellen, an denen Funktionswerte maximal oder minimal sind, (Extremwertprobleme).
- Bereiche, in denen sich die Funktionswerte gar nicht ändern (Das sind dann Identitäten, also quasi "Formeln").

Wir starten mit der ersten Aufgabe: Monotoniefragen.

Satz 4.1 (Monotonie-Kriterien):
Es sei
$$f : [a, b] \to \mathbb{R}$$
differenzierbar auf (a, b) und stetig bei a und b . Dann gilt:

(i) f ist monoton wachsend auf $[a, b] \iff f'(x) \geq 0 \ \forall \, x \in (a, b)$.

(ii) Gilt $f'(x) > 0 \ \forall \ x \in (a, b)$, so ist f streng monoton wachsend auf $[a, b]$.

(iii) f ist monoton fallend auf $[a, b] \iff f'(x) \leq 0 \ \forall \, x \in (a, b)$.

(iv) Gilt $f'(x) < 0 \ \forall \ x \in (a, b)$, so ist f streng monoton fallend auf $[a, b]$.

Beweis:

(i)
- Sei $f \uparrow$ und sei $c \in (a, b) \Rightarrow \dfrac{f(x) - f(c)}{x - c} \geq 0 \ \forall \, x \ (\neq c)$; $x \to c \Rightarrow f'(c) \geq 0$.
- Sei umgekehrt $f' \geq 0$ auf (a, b), und sei $a \leq x < y \leq b \Rightarrow f(y) - f(x) \underbrace{=}_{\substack{\text{(MWS)} \\ \xi \in (x,y)}} \underbrace{f'(\xi)}_{\geq 0} \underbrace{(y - x)}_{\geq 0} \geq 0$, also $f(y) \geq f(x)$.

(ii) $x < y \Rightarrow f(y) - f(x) \underbrace{=}_{\substack{\text{(MWS)} \\ \xi \in (x,y)}} \underbrace{f'(\xi)}_{>0} \underbrace{(y - x)}_{>0} > 0$, also $f(y) > f(x)$.

(iii),(iv) analog

Bemerkung:

(ii) und (iv) lassen sich nicht umkehren. Z.B. ist $f(x) = x^3$ streng monoton wachsend auf $\mathbb{R}$, $f'(x) = 3x^2 \geq 0$ auf $\mathbb{R}$, aber nicht > 0 auf $\mathbb{R}$.

Einschub: Definitionsbereich bei Potenzen

- Für natürliche Zahlen $n \in \mathbb{N}$ definiert man x^n für jede reelle Zahl $x \in \mathbb{R}$ induktiv:
$$x^1 := x, \quad x^{n+1} := x^n \cdot x$$
- Für $x \neq 0$, $n \in \mathbb{N}$ definiert man
$$x^{-n} := \frac{1}{x^n}$$
und schließlich noch $x^0 := 1$ für jedes $x \in \mathbb{R}$. (also auch $0^0 = 1$)
- Für $n \in \mathbb{N}$ definiert man
$$x^{1/n} := \sqrt[n]{x},$$
und zwar für $x \in \mathbb{R}$, wenn n ungerade ist und für $x \geq 0$, wenn n gerade ist.
- Will man allgemeine Potenzen x^α, $\alpha \in \mathbb{R}$ defimieren, so setzt man
$$x^\alpha := e^{\alpha \ln x}$$
Das gilt dann aber **nur** für $x > 0$.

  Voraussetzung dafür ist aber:
  - eine klare Definition von e^x (das fehlt uns eigentlich noch!)
  - das Potenzgesetz $x^x \cdot e^y = e^{x+y}$
  - die Existenz der Umkehrfunktion $\ln x : (0, \infty) \to \mathbb{R}$.
- Die Verträglichkeit mit den oben genannten Spezialfällen müsste man dann beweisen, z.B.: $n \in \mathbb{N} \Rightarrow$
$$x_{\text{neu}}^{n+1} = e^{(n+1)\ln x} = e^{n \ln x + \ln x} = e^{n \ln x} \cdot e^{\ln x} = x^n \cdot x = x_{\text{alt}}^{n+1}$$

Beispiele:

(i) Wir diskutieren noch einmal die Funktion

$$f(x) = \sin x : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, 1].$$

Sie ist streng monoton wachsend, denn

$$(\sin x)' = \cos x > 0 \quad \forall \ \ x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

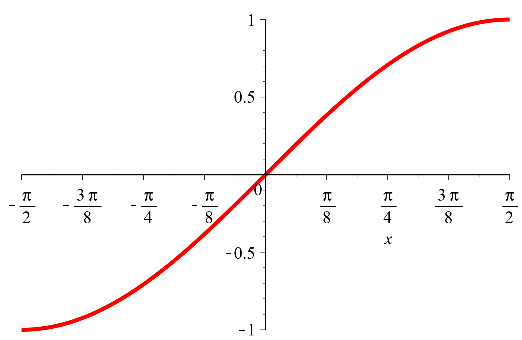
Damit ist $\sin x$ in diesem Bereich injektiv, aber auch surjektiv, letzteres wegen des Zwischenwertsatzes.

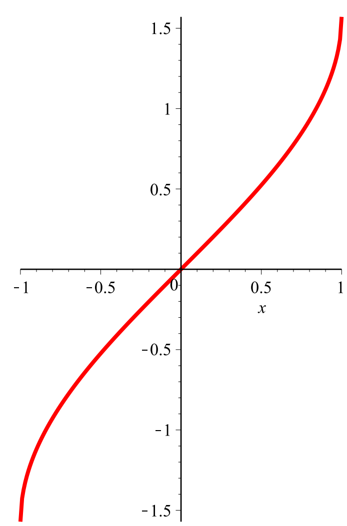
Also existiert die Umkehrfunktion **Arcussinus:**

$$\arcsin : [-1, 1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Deren Ableitung existiert für $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Wegen $\sin(\arcsin(x)) = x$ und der Kettenregel folgt

$$\cos(\arcsin x) \cdot (\arcsin x)' = 1 \;\Rightarrow\; (\arcsin x)' = \frac{1}{\cos(\arcsin x)} \underbrace{=}_{\cos^2 + \sin^2 = 1} \frac{1}{\sqrt{1 - x^2}}$$





(ii) Wir wollen auch die Tangensfunktion unter diesem Aspekt neu diskutieren, also

$$f(x) = \tan x : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to (-\infty, \infty).$$

f ist streng monoton wachsend, denn

$$(\tan x)' = 1 + \tan^2 x > 0 \quad \text{stets}$$

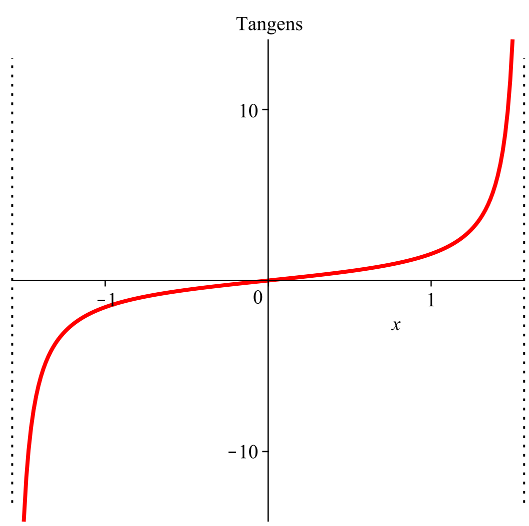
und $\tan x$ ist surjektiv (ZWS).

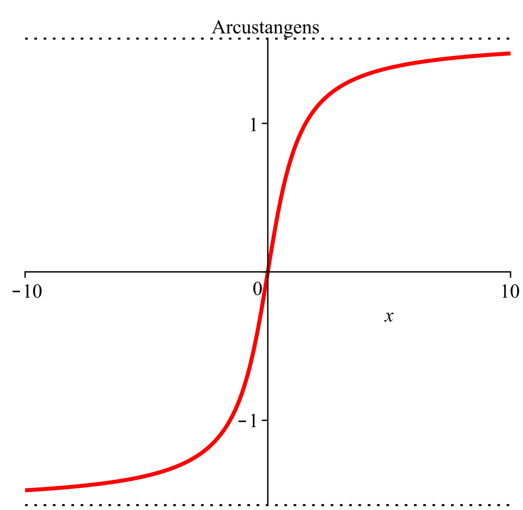
Also gibt es eine Umkehrfunktion **Arcustangens**

$$\arctan : (-\infty, \infty) \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Wegen $\tan(\arctan x) = x$ und der Kettenregel folgt

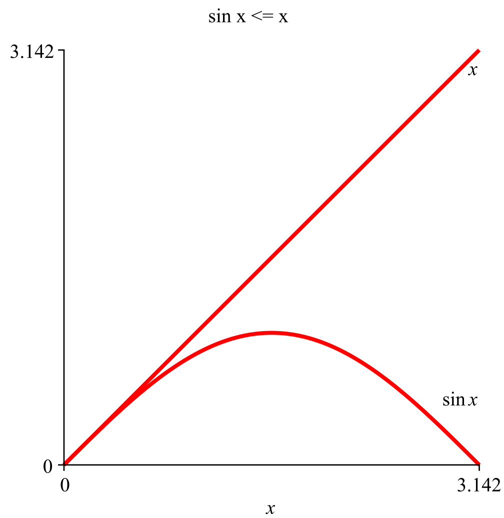
$$(\arctan x)' = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}$$





(iii) Manche Ungleichungen lassen sich über Monotonieargumente herleiten. Als Beispiel zeigen wir

$$\sin x \leq x \quad \forall \ x \geq 0.$$



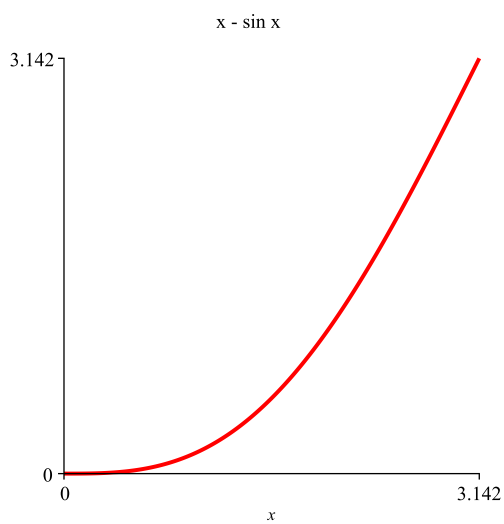
Zum Beweis betrachten wir die Funktion

$$f(x) := x - \sin x$$

f ist differenzierbar, $f(0) = 0$, $f'(x) = 1 - \cos x \geq 0 \ \forall \ x \geq 0$

$\Rightarrow$ f ist monoton wachsend auf $[0, \infty)$. Insbesondere ist also

$$f(x) \geq f(0) = 0 \quad \forall \ x \geq 0.$$



(iv) Ein weiteres Beispiel ist die **verallgemeinerte Bernoullische Ungleichung:**

$$(1+x)^\alpha \geq 1 + \alpha x \quad \forall \, x \geq 0, \ \alpha \geq 1$$

Für den Beweis arbeiten wir mit der Hilfsfunktion

$$f(x) := (1+x)^\alpha - 1 - \alpha x : [0, \infty) \to \mathbb{R}.$$

f ist differenzierbar auf $[0, \infty)$. mit

$$f'(x) = \alpha\Big(\underbrace{(1+x)^{\alpha-1} - 1}_{\geq 0}\Big) \geq 0$$

$\Rightarrow$ f ist monoton wachsend $\Rightarrow$ $f(x) \geq f(0) = 0 \quad \forall \; x \geq 0.$

Satz 4.2 (hinreichendes Kriterium für lokale Minima/Maxima):
Es sei

$$f : [a, b] \to \mathbb{R}$$

differenzierbar auf $[a, b]$, und an einer Stelle c mit $a < c < b$ sei

$$\boxed{f'(c) = 0.}$$

(c erfüllt also die notwendige Bedingung aus Satz 3.1 und ist damit Kandidat für eine Extremalstelle.)

(i) Ist dann für ein $\delta > 0$

$$f'(x) \leq 0 \quad \text{für} \quad x \in (c - \delta, c) \quad \text{und} \quad f'(x) \geq 0 \quad \text{für} \quad x \in (c, c + \delta),$$

so liegt bei c eine lokale Minimalstelle vor.

Mit anderen Worten: Eine Minimalstelle liegt bei c vor, wenn f' bei c sein Vorzeichen wechselt von $-$ (links von c) über 0 (bei c) nach $+$ (rechts von c).

(ii) Wechselt f' sein Vorzeichen von $+$ nach $-$, genauer: $\exists \; \delta > 0$ mit

$$f'(x) \geq 0 \;\; \forall \;\; c - \delta < x < c, \;\; f'(c) = 0, \;\; f'(c) \leq 0 \;\; \forall \;\; c < x < c + \delta),$$

so liegt bei c eine lokale Maximalstelle vor.

(iii) Wechselt f' sein Vorzeichen beim Übergang von $c - \delta$ nach $c + \delta$ <u>nicht</u>, so liegt bei c keine Extremalstelle vor.

Beweis:
zu (i) $f(x) - f(c) = f'(\xi)(x - c)$ mit ξ zwischen c und x (MWS)

- $c - \delta < x < c \Rightarrow x - c < 0, \; f'(\xi) \leq 0 \Rightarrow f(x) - f(c) \geq 0$
- $c < x < c + \delta \Rightarrow x - c > 0, \; f'(\xi) \geq 0 \Rightarrow f(x) - f(c) \geq 0$.

In beiden Fällen ist $f(x) \geq f(c)$, also ist c lokales Minimum.

(ii) beweist man analog.

Im Fall (iii) ist die Kurve im Intervall $(c - \delta, c + \delta)$ monoton.

Bemerkung:

Die Situation (i) in Satz 4.2 liegt beispielsweise dann vor, wenn f' <u>streng</u>

monoton wächst.

Hinreichend hierfür ist (nach Satz 4.1), dass $(f')'$ ($=: f''$) (**2.Ableitung**) > 0 ist in einer Umgebung $(c-\delta, c+\delta)$ von c .

Ist f'' stetig, genügt es dabei, dass $f''(c) > 0$ ist, denn stetige Funktionen wechseln nicht abrupt ihr Vorzeichen. Wir haben daher den folgenden

Satz 4.3 (hinreichendes Kriterium für lokale Extrema, Teil 2):
Es sei

$$f : [a,b] \to \mathbb{R}$$

zweimal differenzierbar, und sei f'' stetig.

Dann gilt für einen inneren Punkt $c \in (a,b)$

(i) $f'(c) = 0,\ f''(c) > 0 \ \Rightarrow\ c$ ist lokale Minimalstelle von f .

(ii) $f'(c) = 0,\ f''(c) < 0 \ \Rightarrow\ c$ ist lokale Maximalstelle von f .

Bemerkung:
Ein Kurvenstück $f : I \to \mathbb{R}$, in dem

$$f''(x) \geq 0 \quad \forall\ x \in I$$

gilt, nennt man **konvex,** manchmal auch **Linkskurve.**

Gilt sogar $f''(x) > 0 \ \ \forall\ x \in I$, so heißt f **streng konvex.**

Analog definiert man die Begriffe **konkav** oder **Rechtskurve** ($f''(x) \leq 0$) bzw. **streng konkav** ($f'' < 0$) .

Beispiel:

Betrachte die Kurve mit der Gleichung

$$f(x) = \frac{1}{1+e^{-x}},\ x \in \mathbb{R}.$$

- $f'(x) = -\dfrac{1}{(1+e^{-x})^2} \cdot (-e^{-x}) = \dfrac{e^{-x}}{(1+e^{-x})^2} > 0$

  Also ist f monoton wachsend (sogar streng).

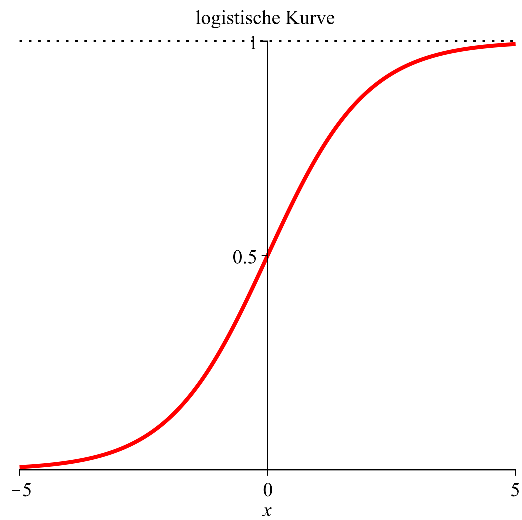
- $$f''(x) = \frac{-e^{-x}(1+e^{-x})^2 - e^{-x} \cdot 2(1+e^{-x}) \cdot (-e^{-x})}{(1+e^{-x})^4} =$$
  $$= \frac{-e^{-x}(1+e^{-x}) + 2e^{-x}e^{-x}}{(1+e^{-x})^3} = \frac{e^{-x}(-1+e^{-x})}{(1+e^{-x})^3}$$

  Also ist f konvex $\iff f'' \geq 0$

  $$\iff -1+e^{-x} \geq 0 \iff e^{-x} \geq 1 \iff e^{x} \leq 1 \iff x \leq 0,$$

  und f ist konkav $\iff f'' \leq 0 \iff x \geq 0.$

- Berücksichtigt man noch $\lim\limits_{x\to-\infty} f(x) = 0$ und $\lim\limits_{x<\to\infty} f(x) = 1$, so hat man einen Überblick über den Verlauf des Graphen:



Diese Kurve wird gerne eingesetzt, um einen Wachstumsprozess zu beschreiben, der anfänglich dynamisch ist (exponentiell), später aber immer mehr abflacht bis zu einem Sättigungsniveau.

In der Ökonomie tritt das etwa auf beim "Lebenszyklus" eines Produktes (anfangs durch Werbung promoviertes schnelles Wachstum bei späterer Marktsättigung).

Bei Populationsentwicklungen trägt das der Tatsache Rechnung, dass exponentielles Wachstum ad infinitum nicht realistisch ist.

In der Literatur wird diese Kurve auch die **logistische Wachstumskurve** genannt.

Satz 4.4:

Es sei

$$f : I \to \mathbb{R},$$

$I \subset \mathbb{R}$ ein Intervall, differenzierbar auf I mit

$$f' \equiv 0 \quad \text{auf} \quad I.$$

Dann ist f konstant auf I .

Beweis:

Sei $c \in I$ fest und $x \in I$ beliebig

$$\Rightarrow \; f(x) - f(c) = f'(\xi)(x - c)$$

(ξ zwischen x und c), also nach Voraussetzung $= 0$.

Bemerkung:

In Anwendungen benutzen wir Satz 4.4 häufig zum Beweis von Identitäten:
$f, g : I \to \mathbb{R}$ differenzierbar mit

$$f' \equiv g' \quad \text{auf } I \text{ und } f(c) = g(c) \text{ für ein } c \in I$$

$$\Rightarrow \; f \equiv g \quad \text{auf } I.$$

(Dies ist Satz 4.4, angewandt auf $f - g$.)

Beispiele dazu:

(i) Behauptung: $|x| < 1 \;\Rightarrow\; 2\arctan x = \arctan \dfrac{2x}{1-x^2}$

Beweis:

- Die Behauptung stimmt offenbar für $x = 0$
- Die Ableitung der linken Seite ist $\dfrac{2}{1+x^2}$
- die Ableitung der rechten Seite ist

$$\frac{1}{1+\frac{4x^2}{(1-x^2)^2}} \cdot \frac{2(1-x^2)+2x\cdot 2x}{(1-x^2)^2} = \frac{2+2x^2}{(1-x^2)^2+4x^2} = \frac{2(1+x^2)}{(1+x^2)^2} = \frac{2}{1+x^2}$$

Daraus folgt die Behauptung mit obiger Bemerkung.

(ii) Erneut folgt

$$\cos^2 x + \sin^2 x \equiv 1,$$

denn:

Bei $x = 0$ stimmt dies, und

$$(\cos^2 x + \sin^2 x)' = -2\cos x \sin x + 2\sin x \cos x \equiv 0.$$

Satz 4.5 (Grenzwertregeln von l'Hospital):

(Guillaume François Antoine Marquis de l'Hospital, 1661-1704)

Es sei $I = [a-\eta, a) \subset \mathbb{R},\ (\eta > 0,\ a \in \mathbb{R})$ oder $I = (K, \infty),\ K > 0$,

und es seien $f, g : I \to \mathbb{R}$ differenzierbar auf I mit

$$\lim_{x\to a-} \frac{f'(x)}{g'(x)} = \alpha \in \mathbb{R} \quad \text{bzw.} \quad \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \alpha \in \mathbb{R}.$$

Gilt dann

(i) $\lim\limits_{x\to a-} f(x) = \lim\limits_{x\to a-} g(x) = 0 \quad (a \in \mathbb{R} \cup \{\infty\})$ oder

(ii) $\lim\limits_{x\to a-} g(x) = \infty$ oder $-\infty \quad (a \in \mathbb{R} \cup \{\infty\}$),

so existiert auch $\lim\limits_{x\to a-} \dfrac{f(x)}{g(x)}$ $\left(\text{bzw. } \lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)}\right)$,

und dieser Grenzwert ist auch $= \alpha$.

Bemerkung:

Entsprechende Aussagen gelten auch für Grenzwerte der Form

$$\lim_{x\to a+} \frac{f(x)}{g(x)} \quad \text{bzw.} \quad \lim_{x\to -\infty} \frac{f(x)}{g(x)}.$$

Beweis:

Wir diskutieren den Fall (i):

- $a \in \mathbb{R}:\ \ \varepsilon > 0$ gegeben. Wähle $\delta > 0$ mit

  $$\left|\frac{f'(x)}{g'(x)} - \alpha\right| < \varepsilon \ \ \forall \ \ a - \delta < x < a.$$

  Setze dann $f(a) = g(a) := 0$ ("stetig ergänzt").

  Dann gilt für $a - \delta < x < a$:

  $$\left|\frac{f(x)}{g(x)} - \alpha\right| = \left|\frac{f(x) - f(a)}{g(x) - g(a)} - \alpha\right| \underbrace{=}_{\substack{\text{2.MWS} \\ \xi \text{ zwischen } a \text{ und } x}} \left|\frac{f'(\xi)}{g'(\xi)} - \alpha\right| < \varepsilon.$$

- $a = \infty:\ \ \varepsilon > 0$ gegeben. Wähle $K > 0$ mit

  $$\left|\frac{f'(x)}{g'(x)} - \alpha\right| < \varepsilon \quad \forall \ \ x \geq K.$$

  Dann gilt für $x \geq K,\ y \geq x$:

  $$\left|\frac{f(y) - f(x)}{g(y) - g(x)} - \alpha\right| \underbrace{=}_{\text{2.MWS}} \left|\frac{f'(\xi)}{g'(\xi)} - \alpha\right| < \varepsilon$$

  Für $y \to \infty$ folgt dann:

  $$\left|\frac{f(x)}{g(x)} - \alpha\right| \leq \varepsilon \quad \forall \ \ x \geq K.$$

Beispiele:

- $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = \lim\limits_{x\to 0} \dfrac{\cos x}{1} = 1$

  Das ist in Kurzform die Anwendung der l'Hospitalschen Regel. Die komplette Argumentation sähe folgendermaßen aus:

  - Wir wollen $\lim\limits_{x\to 0} \dfrac{\sin x}{x}$ berechnen und stellen fest:
  - $\lim\limits_{x\to 0} \sin x = \lim\limits_{x\to 0} x = 0$ (streng genommen wegen der Stetigkeit dieser Funktionen)
  - $\sin x$ und x sind diff'bar mit $(\sin x)' = \cos x$ und $(x)' = 1$

- $\lim\limits_{x\to 0} \dfrac{(\sin x)'}{(x)'} = \lim\limits_{x\to 0} \dfrac{\cos x}{1}$ existiert und ist $= 1$. (eigentlich auch ein Stetigkeitsargument)
- l'Hospital sagt dann: Dann ist auch $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$.

- $\lim\limits_{x\to 0+} \dfrac{\ln(1+x)}{x} = \lim\limits_{x\to 0+} \dfrac{1}{1+x} = 1$
- $\lim\limits_{x\to 0} \dfrac{1-\cos x}{x^2} = \lim\limits_{x\to 0} \dfrac{\sin x}{2x} = \dfrac{1}{2}$
- $\lim\limits_{x\to 0} \dfrac{\sin x - x}{x^3} = \lim\limits_{x\to 0} \dfrac{\cos x - 1}{3x^2} = -\dfrac{1}{6}$
- $\lim\limits_{x\to 0+} x\ln x = \lim\limits_{x\to 0+} \dfrac{\ln x}{\frac{1}{x}} = \lim\limits_{x\to 0+} \dfrac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim\limits_{x\to 0+} -x = 0$

  Damit folgt: $\lim\limits_{x\to 0+} x^x = e^0 = 1$.
- $\lim\limits_{x\to\infty} \dfrac{\ln x}{\sqrt{x}} = \lim\limits_{x\to\infty} \dfrac{\frac{1}{x}}{\frac{1}{2\sqrt{x}}} = \lim\limits_{x\to\infty} \dfrac{2}{\sqrt{x}} = 0$
- $\lim\limits_{x\to\infty} \dfrac{x^2}{e^x} = 0$, denn $\lim\limits_{x\to\infty} \dfrac{x^2}{e^x} = \lim\limits_{x\to\infty} \dfrac{2x}{e^x} = \lim\limits_{x\to\infty} \dfrac{2}{e^x} = 0$.
- $\lim\limits_{x\to\infty} \dfrac{\ln(1+e^{2x})}{x} = \lim\limits_{x\to\infty} \dfrac{e^{2x}\cdot 2}{1+e^{2x}} = \lim\limits_{x\to\infty} \dfrac{2}{e^{-2x}+1} = 2$
- $\lim\limits_{x\to 0} \dfrac{e^{x^2}-1-x^2}{(1-\cos x)^2} = \ldots$ (l'Hospital 4-mal) $\ldots = 2$
- $\lim\limits_{x\to\infty} x\left(\dfrac{\pi}{2} - \arctan x\right) = \lim\limits_{x\to\infty} \dfrac{\frac{\pi}{2} - \arctan x}{\frac{1}{x}} = \lim\limits_{x\to\infty} \dfrac{-\frac{1}{1+x^2}}{-\frac{1}{x^2}} = 1$

§ 5 Das Newton-Verfahren

Eine der schönsten Anwendungen ist das Newton-Verfahren zur näherungsweisen Lösung von Gleichungen. Die Gleichung möge in Nullform vorliegen:

$$f(x) = 0$$

Zum Starten des Verfahrens benötigen wir einen Näherungswert a.

Die Idee ist dann so simpel we genial. Wir ersetzen lokal bei a durch die Tangente an den Graphen von $y = f(x)$ durch den Punkt $\Big(a, f(a)\Big)$. Diese nennen wir $t(x)$. Ihre Gleichung ist gegeben durch

$$t(x) = f(a) + f'(a)(x - a).$$

Wir lösen nun nicht die (schwierige) Gleichung $f(x) = 0$, sondern die viel einfachere, da lineare Gleichung

$$\begin{aligned} & t(x) = 0 \\ \iff\ & f(a) + f'(a)(x - a) = 0 \\ \iff\ & x - a = -\frac{f(a)}{f'(a)} \\ \iff\ & x = a - \frac{f(a)}{f'(a)} \end{aligned}$$

Von diesem x erhoffen wir uns dann eine verbesserte Näherung.

Das **Newton-Verfahren** iteriert diesen Vorgang, und wir erhalten die rekursive Folge

$$\text{Startwert } x_0, \quad x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$$

Die wichtigsten Eigenschaften dieses Verfahrens leitet man in der Numerischen Mathematik her. Im einzelnen gilt:

- Das Verfahren ist **quadratisch konvergent.**

  Damit ist folgendes gemeint: Bezeichnet ξ die wahre Lösung der Gleichung $f(x) = 0$, so misst ja

  $$\varepsilon_n := |x_n - \xi|$$

  den (absoluten) Fehler zwischen $n-$tem Näherungswert und dem wahren Wert.

  Quadratische Konvergenz heißt nun:

  $$\varepsilon_{n+1} = O(\varepsilon_n^2)$$

  oder, was dasselbe bedeutet,

  $$\varepsilon_{n+1} \leq K\ \varepsilon_n^2$$

  mit einer Konstanten $K > 0$.

  Rechnet man dezimal, so bedeutet das, dass sich die Anzahl der richtig berechneten Nachkommastellen pro Schritt im wesentlichen

verdoppelt,

wobei die Formulierung “im wesentlichen” heißen soll, dass es von der Größe der Konstanten K abhängt.

Eine kleine Einschränkung: Quadratische Konvergenz liegt nicht vor, wenn neben $f(\xi) = 0$ auch noch $f'(\xi) = 0$ gilt (“doppelte Nullstelle”).

- Ein gravierender Nachteil ist allerdings die **Startwertproblematik.** Bei schlechtem Startwert x_0 kann es passieren, dass das Verfahren überhaupt nicht konvergiert oder die quadratische Konvergenz erst spät eintritt.

  Das ist insbesondere dann der Fall, wenn $f'(x_0) \approx 0$ ist. (Was würde denn passieren, wenn $f'(x_0) = 0$ wäre?

  Anwender schalten dann gelegentlich die viel langsamere Bisektionsmethode vor, um zu einem brauchbaren Startwert zu gelangen.

Wir illustrieren die Phänomene an Beispielen:

(i) Wir berechnen $\sqrt{2}$ als Nullstelle von $f(x) = x^2 - 2$:

$$x - \frac{f(x)}{f'(x)} = x - \frac{x^2 - 2}{2x} = \frac{x}{2} + \frac{1}{x}$$

Auf TR-Genauigkeit ist $\sqrt{2} = 1,414213562\ldots$

- Die Iterationen mit dem (guten) Startwert $x_0 = 1$ liefern sukzessive folgende Werte:

  $x_1 = \frac{1}{2} + \frac{1}{1} = 1,5$

  $x_2 = \frac{1,5}{2} + \frac{1}{1,5} = 1,4166666\ldots$

  $x_3 = 1,4142156$

  x_4 hat bereits TR-Genauigkeit.

- $x_0 = 100$ ist ein schlechter Startwert mit den ersten Iterationen

$$x_1 = 50,01, \quad x_2 = 25,025, \quad x_3 \approx 12,5, \ldots$$

  (Konvergenz: wohl ja, quadratische Konvergenz: erst spät)

(ii) $f(x) = x^2$ hat bei $\xi = 0$ eine doppelte Nullstelle. Die Newton-Iteration ist auch nicht quadratisch konvergent:

$$x - \frac{f(x)}{f'(x)} = x - \frac{x^2}{2x} = \frac{x}{2}.$$

Startend mit $x_0 = 1$ erhält man $x_n = \left(\frac{1}{2}\right)^n$, was nicht quadratisch konvergiert.

Bemerkung:

Zum Schluss noch ein Hinweis auf eine nette Spielerei: Das wird erst interessant, wenn man das Newton-Verfahren ins Komplexe überträgt. Die Eigenschaften bleiben im wesentlichen erhalten. Nehmen wir nun an, $f(z)$ habe mehrere Nullstellen, sagen wir drei Stück: ξ_1, ξ_2, ξ_3. Nun können wir die Zahlenebene $\mathbb{C}$ in 4 Bereiche färben nach folgendem Prinzip:

Wird ein Punkt zum Startpunkt des Newton-Verfahrens gemacht, so wird er blau gefärbt, wenn das Verfahren gegen ξ_1 konvergiert, grün, wenn es gegen ξ_2 konvergiert, rot, wenn es gegen ξ_3 konvergiert und schwarz, wenn es gar nicht konvergiert.

Auf diese Weise bekommt man ein interessantes Bild: Klare blaue, grüne und rote Umgebungen von ξ_1, ξ_2 bzw. ξ_3, und selbstähnliche Strukturen dazwischen.

§ 6 Der Taylorsche Satz

Die zweite Ableitung

$$f'' := (f')',$$

falls existent, war früher schon aufgetreten. Natürlich kann man das induktiv weiter fortführen, benutzt noch

$$f''' := (f'')',$$

wählt für die weiteren Ableitungen aber eine andere Notation:

$$f^{(n+1)} := (f^{(n)})'$$

Ist eine Funktion f $n-$mal differenzierbar auf einem Intervall I , und ist $f^{(n)}$ auf I auch noch stetig, so schreiben wir

$$f \in C_n(I).$$

Ist f beliebig oft differenzierbar auf I, so schreiben wir

$$f \in C_\infty(I)$$

Ist f auf I lediglich stetig, so schreiben wir

$$f = f^{(0)} \in C(I) \quad \text{oder} \quad f \in C_0(I).$$

Beispiele:

(i) $x^\alpha \in C_\infty(I)$ mit $I = (0, \infty)$;

$(x^\alpha)' = \alpha x^{\alpha-1}$; $(x^\alpha)'' = \alpha(\alpha-1)x^{\alpha-2}$; allgemein:

$$(x^\alpha)^{(n)} = \alpha(\alpha-1) \ \ldots \ (\alpha-n+1) \ x^{\alpha-n} = n! \binom{\alpha}{n} x^{\alpha-n}$$

Das ist zunächst nur für $\alpha \in \mathbb{N}$ erklärt. Es macht aber durchaus für $\alpha \in \mathbb{R}$ Sinn, wenn man definiert:

$$\binom{\alpha}{n} := \frac{\alpha \cdot (\alpha-1) \ldots \cdot (\alpha-n+1)}{n!})$$

(ii) $(\sin x)' = \cos x$; $(\sin x)'' = -\sin x$; $(\sin x)''' = -\cos x$;

$(\sin x)^{(4)} = \sin x$; usw. Allgemein ist:

$$(\sin x)^{(2k)} = (-1)^k \sin x; \ (\sin x)^{(2k+1)} = (-1)^k \cos x.$$

$$\Rightarrow \ \sin x \in C_\infty(\mathbb{R}).$$

Zielsetzung:

Gegeben $f : I \to \mathbb{R}$, $a \in I$. Wir wollen $f(x)$ annähern durch ein Polynom $P(x)$ (vom Grade $\leq n$), welches bei an einer festen Stelle $a \in I$ die Bedingungen

$$P(a) = f(a),\ P'(a) = f'(a),\ P''(a) = f''(a),\ \text{usw.}$$

erfüllt. Die Erwartung ist dann, dass $P(x)$ die Funktion $f(x)$ gut annähert.

Hintergrund ist natürlich, dass es sich mit Polynomen verhältnismäßig einfach arbeiten lässt.

Wie sieht das im einzelnen aus?

- $n = 0:\quad P(x) \equiv f(a)$
- $n = 1:\quad P(x) = f(a) + f'(a)(x - a)$

  $\Big(\Rightarrow\ P(a) = f(a),\ P'(a) = f'(a)\Big):$

  Das ist die Tangente oder lineare Annäherung
- $n = 2:\quad P(x) = f(a) + f'(a)(x - a) + \dfrac{1}{2}\, f^{''}(a)(x - a)^2$

  $\Big(\Rightarrow\ P(a) = f(a),\ P'(a) = f'(a),\ P^{''}(a) = f''(a)\Big)$

  Dies nennt man gelegentlich die **Schmiegparabel** bei a.
- $n = 3:\quad P(x) = f(a) + f'(a)(x-a) + \dfrac{f''(a)}{2}\,(x-a)^2 + \dfrac{f'''(a)}{6}\,(x-a)^3$

Mit einer einfachen Induktion sieht man nun

Satz 5.1:

Es sei f $n-$mal differenzierbar auf einem Intervall $I,\ a \in I$.

Dann gibt es genau ein Polynom $P_n(x)$ mit $\text{grad}\, P_n \leq n$ und

$$P_n^{(k)}(a) = f^{(k)}(a) \quad \forall\ k = 0,\ \dots\ n,$$

und dieses ist gegeben durch

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x - a)^k.$$

Man nennt dieses Polynom das **n-te Taylor-Polynom** zu f beim **Entwicklungspunkt** a .

Beispiele:

(i) Sei $f(x) = e^x \Rightarrow f^{(k)}(x) = e^x \quad \forall\ k$. Also ist das $n-$te Taylorpolynom bei Entwicklungspunkt $a = 0$ gegeben durch

$$P_n(x) = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} = \sum_{k=0}^{n} \frac{x^k}{k!}$$

oder bei beliebigem Entwicklungspunkt a

$$P_n(x) = \sum_{k=0}^{n} \frac{e^a}{k!}\,(x-a)^k$$

(ii) $f(x) = \dfrac{1}{1-x}$; $f'(x) = \dfrac{1}{(1-x)^2}$; $f''(x) = \dfrac{2}{(1-x)^3}$ allgemein:

$$f^{(k)}(x) = \frac{k!}{(1-x)^{k+1}}$$

Induktionsschritt:

$$\frac{d}{dx}\,k!\,(1-x)^{-k-1} = k!(-k-1)(1-x)^{-k-2}\cdot(-1) = \frac{(k+1)!}{(1-x)^{k+2}}$$

Also ist das $n-$te Taylorpolynom bei Entwicklungspunkt $a=0$ gegeben durch

$$P_n(x) = \sum_{k=0}^{n} x^k$$

Das ist die geometrische Summe, die wir früher schon ausgerechnet hatten:

$$P_n(x) = \frac{1-x^{n+1}}{1-x}$$

Wir wissen auch: Ist $|x|<1$, so gilt: $\lim\limits_{n\to\infty} x^{n+1} = 0.$

Also gilt in diesem Beispiel:

$$\lim_{n\to\infty} P_n(x) = f(x) \quad (\text{jedenfalls für } |x|<1.\ )$$

Wir sagen dann, wir haben $f(x) = \dfrac{1}{1-x}$ **in eine Reihe entwickelt:**

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \quad \text{für} \quad |x|<1.$$

Setzt man konkrete Zahlen ein, erhält man interessante Identitäten, z.B.

$$\sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{1-\frac{1}{2}} = 2, \quad \sum_{k=0}^{\infty} \Big(-\frac{1}{2}\Big)^k = \frac{1}{1+\frac{1}{2}} = \frac{2}{3}$$

Fragestellung:

Wie gut stellt das $n-$ te Taylor-Polynom $P_n(x)$ die gegebene Funktion dar? Dazu betrachten wir den 'Fehlerterm' (Rest)

$$R_n(x) := f(x) - P_n(x).$$

Eine Untersuchung von $R_n(x)$, insbesondere die Frage

$$R_n(x) \to 0,\ n \to \infty?$$

oder nicht, wird erleichtert, wenn wir Darstellungsformeln für $R_n(x)$ herleiten. Die wichtigste formulieren wir in folgendem

Satz 5.2 (Taylor): (Brook Taylor, 1685-1731)

Es sei $f : I \to \mathbb{R}$ $(n+1)$ -mal differenzierbar auf $I,\ a \in I,$ und es sei

$$R_n(x) := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

der (Fehler)-Rest zwischen $f(x)$ und dem n-ten Taylorpolynom . Dann gibt es eine Stelle ξ , gelegen zwischen a und x mit

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1} \quad \text{(Formel von Taylor/Lagrange).}$$

(Bemerkung: Dieses Restglied ist leicht zu merken: $R_n(x)$ entspricht dem $(n+1)$-ten Term im Taylorpolynom mit $f^{(n+1)}(\xi)$ statt $f^{(n+1)}(a)$.)

Beweis:

Nach Satz 5.1 gilt:

$R_n^{(k)}(a) = f^{(k)}(a) - P_n^{(k)}(a) = 0 \quad \forall\ k = 0,\ \ldots,\ n.$

Weiter ist $R_n^{(n+1)}(x) = f^{(n+1)}(x) - 0 = f^{(n+1)}(x) \quad \forall\ x \in I$.

Betrachte nun auch $Q_n(x) := \dfrac{(x-a)^{n+1}}{(n+1)!}$

$\Rightarrow\ Q_n^{(k)}(a) = 0 \quad \forall\ 0 \leq k \leq n$ und $Q_n^{(n+1)}(x) \equiv 1$.

$$\Rightarrow \frac{R_n(x)}{Q_n(x)} = \frac{R_n(x) - R_n(a)}{Q_n(x) - Q_n(a)} \underbrace{=}_{\substack{\text{2.MWS} \\ \xi_1 \text{ zwischen } a \text{ und } x}} \frac{R_n'(\xi_1)}{Q_n'(\xi_1)} = \frac{R_n'(\xi_1) - R_n'(a)}{Q_n'(\xi_1) - Q_n'(a)}$$

$$\underbrace{=}_{\substack{\text{2.MWS} \\ \xi_2 \text{ zwischen } a \text{ und } \xi_1}} \frac{R_n''(\xi_2)}{Q_n''(\xi_2)} = \underbrace{\cdots}_{\text{(induktiv)}} = \frac{R_n^{(n+1)}(\xi)}{Q_n^{(n+1)}(\xi)} = f^{(n+1)}(\xi)$$

mit ξ zwischen a und x $\Rightarrow$ $R_n(x) = Q_n(x) \cdot f^{(n+1)}(\xi)$ q.e.d.

Bemerkungen und Beispiele:

(i) Beispiel:

Betrachte noch einmal $f(x) = e^x,\ a = 0:\quad f^{(k)}(x) = e^x \quad \forall\ k$

$$\underbrace{\Rightarrow}_{\text{Taylor}} e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{e^{\xi}}{(n+1)!}\, x^{n+1}$$

(ξ zwischen 0 und x .)

e^{ξ} ist beschränkt (durch e^x oder e^0).

Wegen

$$\frac{x^{n+1}}{(n+1)!} \to 0,\ n \to \infty$$

folgt

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!} = e^x \quad \forall\ x \in \mathbb{R}.$$

Wir schreiben hierfür

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad \forall\ x \in \mathbb{R},$$

und wir können sagen, wir haben e^x in diese Reihe entwickelt.

Als Anwendung wollen wir $e^{0.1}$ auf 10^{-4} genau berechnen:

Dazu sollte

$$R_n(0.1) = \frac{e^{\xi}}{(n+1)!}\, 0.1^{n+1} < 10^{-4}$$

sein. Wegen $0 < \xi < 0.1$ ist sicherlich $e^{\xi} < 3$.

Es reicht also, n so zu wählen, dass

$$3 \cdot \frac{0.1^{n+1}}{(n+1)!} < 10^{-4}$$

gilt. Dies ist für $n = 3$ der Fall. Nun ist

$$1 + 0.1 + \frac{0.1^2}{2} + \frac{0.1^3}{6} \approx 1.1051667$$

$$\Rightarrow\ |e^{0.1} - 1.1051667| < 10^{-4}.$$

TR-Kontrolle zeigt: der wahre Fehler liegt bei $4 \cdot 10^{-6}$.

(ii) Ausgeschrieben lautet der Taylorsche Satz:

- $n = 0:\quad f(x) = f(a) + f'(\xi)(x-a),\ \ \xi$ zwischen a und x .
  Das ist noch einmal der 1.MWS der Differenzialrechnung.
- $n = 1:\quad f(x) = f(a) + f'(a)(x-a) + \dfrac{f''(\xi)}{2}(x-a)^2$.
  Daraus folgt erneut Satz 4.3 (über lokale Extrema). Diesen können wir nunmehr verallgemeinern:

Satz 4.3':

Es sei $f : [a,b] \to \mathbb{R}$ $(n+1)-$mal stetig differenzierbar auf $[a,b]$. Weiter sei $c \in (a,b)$ ein innerer Punkt mit

$$f'(c) = f''(c) = \ \dots\ f^{(n)}(c) = 0, \quad \text{aber} \quad f^{(n+1)}(c) \neq 0.$$

Dann gilt:

(i) Ist n ungerade, so ist im Falle $f^{(n+1)}(c) > 0$ c eine lokale Minimalstelle von f und im Falle $f^{(n+1)}(c) < 0$ c eine lokale Maximalstelle von f .

(ii) Ist n gerade, so liegt bei c kein lokales Extremum vor.

Beweis:

Dies folgt alles aus der Taylorentwicklung von f um den Punkt c :

$$f(x) = f(c) + \underbrace{\sum_{k=1}^{n} \frac{f^{(k)}(c)}{k!}\,(x-c)^k}_{=0} + \frac{f^{(n+1)}(\xi)}{(n+1)!}\,(x-c)^{n+1}$$

sowie: $f^{(n+1)}(c) > 0 \ \Rightarrow\ f^{(n+1)}(\xi) > 0 \ \ \forall\ \ c-\delta < \xi < c+\delta$

(δ genügend klein; beachte die Stetigkeit von $f^{(n+1)}$)

Außerdem ist $(x-c)^{n+1} \geq 0$ im Falle, dass n ungerade ist.

Ist aber n gerade, so ist $(x-c)^{n+1} > 0$ für $x > c$

und < 0 für $x < c$.

(iii) Manchmal erscheint es zweckmäßig, den Taylorschen Satz auch auf ein Polynom selber anzuwenden. Will man z.B. das Polynom

$$f(x) = x^3 - 5x^2 + 8x - 1$$

zentriert um $a = 1$ schreiben, so kann dies via Taylor geschehen:

$f(1) = 3,\ f'(1) = 3x^2 - 10x + 8\Big|_{x=1} = 1,\ f''(1) = 6x - 10\Big|_{x=1} = -4,\ f'''(1) = 6$

Der Fehler ist natürlich $= 0$.

$$\Rightarrow\ x^3 - 5x^2 + 8x - 1 = 3 + (x-1) - 2(x-1)^2 + (x-1)^3,$$

eine Umformung, auf die man (von links nach rechts gesehen) nicht ohne weiteres kommt.

(iv) Es gibt Beispiele dafür, dass nicht zwingend $R_n(x) \to 0$ gilt.

Betrachte etwa

$$f(x) = \begin{cases} \exp\left(-\dfrac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases} \qquad (\in C(\mathbb{R}))$$

$$\Rightarrow \; f'(x) = \begin{cases} \dfrac{2}{x^3}\exp\left(-\dfrac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases} \qquad (\in C(\mathbb{R}))\,,$$

induktiv:

$$f^{(k)}(x) = \begin{cases} P\left(\dfrac{1}{x}\right)\exp\left(-\dfrac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases} \; . \; (P \quad \text{ein Polynom.})$$

Also sind alle Glieder der Taylorreihe $= 0$, und damit stellt sie an keiner einzigen Stelle die Funktion dar (außer trivialerweise am Entwicklungspunkt $a = 0$) .

Kapitel IV Integralrechnung

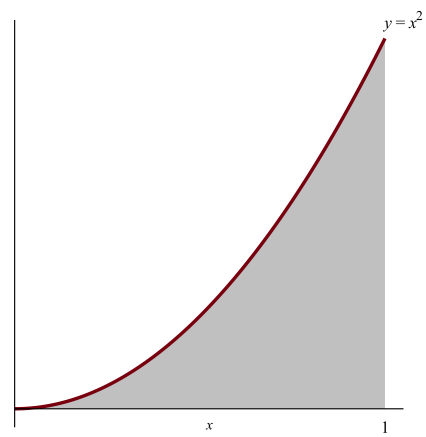
§ 1 Der Begriff des Riemann-Integrals

Eingangsbeispiel:

Das Integral soll, historisch gesehen, die Berechnung von Flächeninhalten ermöglichen. Für einen ersten Zugang dazu stellen wir uns der Frage, der Fläche, die von den Geraden

$$x = 0,\ x = 1,\ y = 0 \quad \text{und der Kurve} \quad y = x^2$$

begrenzt wird, einen sinnvollen Flächeninhalt zuzuordnen.



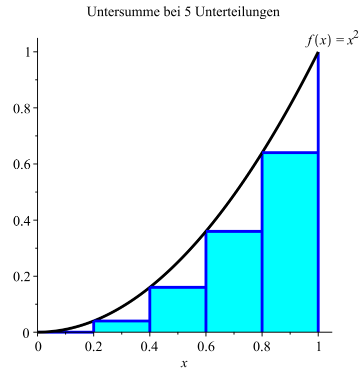
Dazu teilen wir das Intervall $[0, 1]$ in n Teile ein, etwa "äquidistant", d.h. wir wählen (zunächst einmal) gleiche Abstände auf der x -Achse:

$$0,\ \frac{1}{n},\ \frac{2}{n},\ \ldots,\ \frac{n-1}{n},\ \frac{n}{n} = 1,$$

und betrachten als Näherungen für den gesuchten Flächeninhalt die

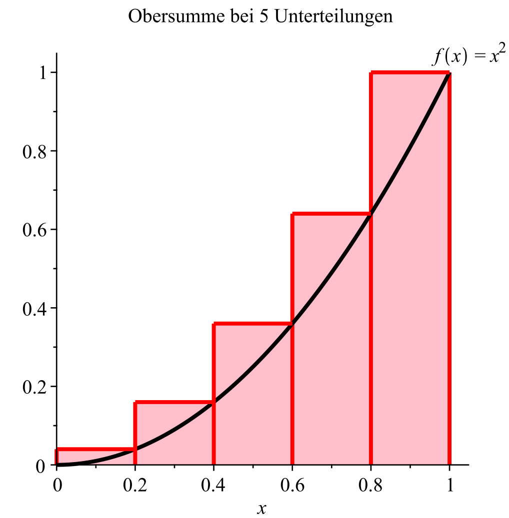
"Untersummen"

$$U_n = 0 + \frac{1}{n} \cdot \frac{1}{n^2} + \frac{1}{n} \cdot \frac{4}{n^2} + \ldots \frac{1}{n} \cdot \frac{(n-1)^2}{n^2}$$



ebenso wie die **"Obersummen"**

$$O_n = \frac{1}{n} \cdot \frac{1}{n^2} + \frac{1}{n} \cdot \frac{4}{n^2} + \ldots + \frac{1}{n} \cdot \frac{n^2}{n^2}.$$



In diesem konkreten Fall verfügen wir über eine Summenformel, nämlich

$$\sum_{k=0}^{n} k^2 = \frac{1}{6}\, n(n+1)(2n+1) \ \ \forall\, n \in \mathbb{N}$$

(s. Übungsblatt 1.) Also ist

$$O_n = \frac{1}{n^3} \sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6n^3} = \frac{2n^3+3n^2+n}{6n^3} \to \frac{1}{3},\ n \to \infty;$$

und

$$U_n = \frac{1}{n^3} \sum_{k=0}^{n-1} k^2 = \frac{(n-1)n(2n-1)}{6n^3} = \frac{2n^3-3n^2+n}{6n^3} \to \frac{1}{3},\ n \to \infty.$$

Daher ist es sinnvoll, den gesuchten Flächeninhalt als $\dfrac{1}{3}$ zu definieren.

Wir schreiben hierfür gleich

$$\int_0^1 x^2\, dx = \frac{1}{3}.$$

Das ist das **Exhaustionsverfahren des Eudoxos** (4.JH v.Chr.).

Zur formalen Definition des bestimmten Integrals verzichten wir aus technischen Gründen auf äquidistante Zerlegungen:

Definition:

Gegeben seien ein Intervall $[a,b] \subset \mathbb{R}$ $(a < b)$ und eine beschränkte Funktion

$$f : [a,b] \to \mathbb{R}.$$

(i) Endlich viele Zahlen $x_k \in [a,b]$, $k = 0, \ldots, n$ mit

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b$$

nennen wir eine **Zerlegung** oder **Partition** von $[a,b]$, welche wir gerne formulieren als Menge

$$Z = \{ x_0, \ldots, x_n \}.$$

Die Größe $\mu(Z) := \max\limits_{k=1,\ldots,n} (x_k - x_{k-1})$ nennen wir die **Feinheit** oder **Maschenweite** dieser Zerlegung.

(ii) Ist $Z = \{ x_0, \ldots, x_n \}$ eine solche Zerlegung von $[a,b]$, so seien

$$m_k := \inf_{x \in [x_{k-1}, x_k]} f(x), \quad \text{und} \quad M_k := \sup_{x \in [x_{k-1}, x_k]} f(x), \; 1 \le k \le n$$

(beachte: $m_k > -\infty$, $M_k < \infty$, denn f ist beschränkt), und wir nennen

$$U = U(Z) = U(Z;f) := \sum_{k=1}^{n} m_k \, (x_k - x_{k-1})$$

die zu Z (und f) gehörige **Untersumme** und

$$O = O(Z) = O(Z;f) := \sum_{k=1}^{n} M_k \, (x_k - x_{k-1})$$

die zu Z (und f) gehörige **Obersumme**.

(iii) Das Infimum aller denkbaren Obersummen, i.e.

$$\inf \{ \; O(Z;f) : \; Z \text{ ist Zerlegung von } [a,b] \; \}$$

nennen wir das **(Darbouxsche) Oberintegral** von f , und wir bezeichnen es mit

$$\overline{\int_a^b} f(x)\, dx.$$

Entsprechend heißt

$$\underline{\int_a^b} f(x)\, dx := \sup\{ \; U(Z;f) : \; Z \text{ ist Zerlegung von } [a,b] \; \}$$

das **(Darbouxsche) Unterintegral** von f .

(iv) Wir nennen f **(Riemann-)integrierbar** über $[a, b]$, wenn Oberintegral und Unterintegral von f übereinstimmen, d.h. wenn

$$\overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx$$

gilt. Dieser (gemeinsame) Wert heißt dann das **(Riemann-)-Integral** von f über $[a, b]$, und wir bezeichnen es kurz mit

$$\int_a^b f(x)\,dx$$

(v) Mit $R[a, b]$ bezeichnen wir die Menge aller über $[a, b]$ Riemann-integrierbaren Funktionen $f : [a, b] \to \mathbb{R}$.

Beispiele:

(i) Wir betrachten die konstante Funktion

$$f(x) \equiv c,\ a \leq x \leq b.$$

Dann gilt für jede Zerlegung $Z = \{\ x_0,\ \ldots,\ x_n\ \}$ von $[a, b]$:

$$m_k = M_k = c,$$

also

$$U(Z; f) = O(Z; f) = \sum_{k=1}^{n} c(x_k - x_{k-1}) \underbrace{=}_{\text{Teleskop}} c(x_n - x_0) = c(b - a)$$

$$\Rightarrow \overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx = c(b - a) = \int_a^b f(x)\,dx.$$

(ii) Wir betrachten $[a, b] = [0, 1]$ und die Funktion

$$f(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & \text{sonst} \end{cases} \qquad \text{(Dirichletfunktion)}$$

Für dieses Beispiel nutzen wir eine Eigenschaft der Zahlen aus $\mathbb{Q}$ (rationale Zahlen) bzw. $\mathbb{R} \setminus \mathbb{Q}$ aus, nämlich:

Zu je zwei Zahlen $\alpha,\ \beta \in \mathbb{R}$ mit $\alpha < \beta$ gibt es immer eine Zahl $x \in \mathbb{Q}$ und $y \in \mathbb{R} \setminus \mathbb{Q}$ mit

$$\alpha < x < \beta \quad \text{und} \quad \alpha < y < \beta.$$

(Man sagt dazu: $\mathbb{Q}$ und $\mathbb{R} \setminus \mathbb{Q}$ liegen **dicht** in $\mathbb{R}$.)

Das ist gar nicht schwer zu sehen: Für x setze man an

$$x = \frac{[q\beta] - 1}{q}$$

mit einer geeigneten Zahl $q \in \mathbb{N}$. Dann gilt nämlich:

$$x \leq \frac{q\beta - 1}{q} < \frac{q\beta}{q} = \beta$$

und

$$x > \frac{q\beta - 2}{q} = \beta - \frac{2}{q} > \alpha, \text{ falls nur } q > \frac{2}{\beta - \alpha}.$$

Für y wende man das eben gezeigte an auf die Zahlen $\frac{\alpha}{\sqrt{2}}$ und $\frac{\beta}{\sqrt{2}}$.
Das so gefundene $x \in \mathbb{Q}$ erfüllt dann

$$\frac{\alpha}{\sqrt{2}} < x < \frac{\beta}{\sqrt{2}} \iff \alpha < x\sqrt{2} < \beta$$

und

$$y := x\sqrt{2} \notin \mathbb{Q}.$$

Zurück zum Beispiel:

Sei $Z : 0 = x_0 < x_1 < \ldots x_n = 1$ eine beliebige Zerlegung von $[0, 1]$

Nach den Ausführungen von eben ist dann

$$M_k = 1 \quad \text{und} \quad m_k = 0 \quad \forall \; k.$$

$$\Rightarrow \; O(Z; f) = \sum_{k=1}^{n} 1 \cdot (x_k - x_{k-1}) = x_n - x_0 = 1 - 0 = 1 \quad \text{und} \quad U(Z; f) = 0.$$

$$\Rightarrow \int\limits_{0}^{-1} f(x) \, dx = 1, \; \int\limits_{-0}^{1} f(x) \, dx = 0,$$

und damit

$$f \notin R[0, 1].$$

Der mathematisch exakte Aufbau der Integralrechnung ist etwas mühsam. Wir zitieren die wichtigsten Resultate:

(i) Riemannsches Integrabilitätskriterium

Eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann (Riemann)-integrierbar, wenn es zu jedem $\varepsilon > 0$ eine Zerlegung Z von $[a,b]$ gibt mit

$$O(Z;f) - U(Z;f) < \varepsilon.$$

Man nennt $O(Z;f)-U(Z;f)$ die **Schwankungssumme** von f bzgl der Zerlegung Z .

(ii) Jede monotone Funktion $f : [a,b] \to \mathbb{R}$ ist integrierbar über $[a,b]$.

Zum Beispiel ist $f(x) = [x]$ integrierbar über jedem Intervall $[a,b]$.

(iii) Jede auf einem Intervall $[a,b]$ erklärte stetige Funktion $f : [a,b] \to \mathbb{R}$ ist auch integrierbar über $[a,b]$.

Wegen (ii) und (iii) ist es ziemlich schwierig, ein Beispiel für eine nicht-integrierbare Funktion zu finden.

(iv) (Integrationsregeln):

Gegeben seien $f, g \in R[a,b], (a < b),\ c \in \mathbb{R}$. Dann gilt:

(α) $f+g \in R[a,b]$ mit $\displaystyle\int_a^b f(x)+g(x)\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx$

sowie

$c \cdot f \in R[a,b]$ mit $\displaystyle\int_a^b c \cdot f(x)\, dx = c \cdot \int_a^b f(x)\, dx$ (Linearität)

(β) $f \cdot g \in R[a,b]$ und $\dfrac{f}{g} \in R[a,b]$, falls $\dfrac{1}{g}$ beschränkt ist.

(leider **ohne** Formel !)

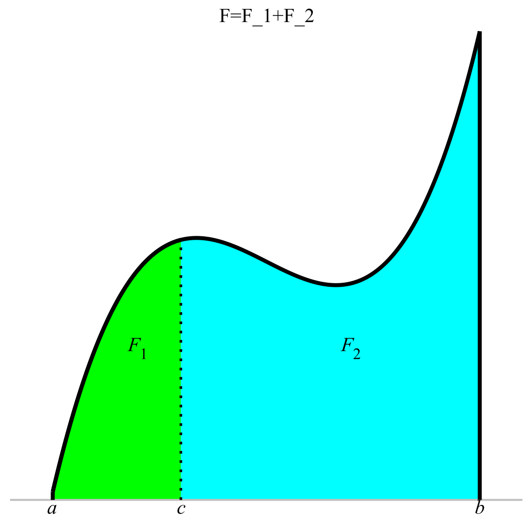
(γ) $|f| \in R[a,b]$ mit $\displaystyle\left|\int_a^b f(x)\, dx\right| \leq \int_a^b |f(x)|\, dx$ (Dreiecksungleichung)

(δ) $f \leq g$ auf $[a,b]$ $\Rightarrow$ $\displaystyle\int_a^b f(x)\, dx \leq \int_a^b g(x)\, dx.$ (Monotonie)

(v) Additivität des Integrals:

Es sei $f \in R[a,b]$, und es sei $a < c < b$. Dann gilt $f \in R[a,c]$ und $f \in R[c,b]$ mit

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$



Bemerkung:

Beim Integral $\int_a^b f(x)\,dx$ gingen wir bisher immer davon aus, dass $a < b$ ist. Sollte in irgendeinem Kontext mal $a > b$ sein, so setzen wir

$$\int_a^b f(x)\,dx := -\int_b^a f(x)\,dx,$$

und, der Vollständigkeit halber

$$\int_a^a f(x)\,dx := 0.$$

Man überzeugt sich leicht davon, dass die oben genannten Rechenregeln auch für diese Situationen gültig sind.

§ 2 Die Hauptsätze der Differenzial- und Integralrechnung

Definition:

Gegeben sei $f \in R[a,b]$. Dann heißt

$$\mathcal{F} : [a,b] \to \mathbb{R},\ \mathcal{F}(x) := \int\limits_a^x f(t)\, dt$$

(welches wegen der Additivität((v) aus § 1 existiert) das

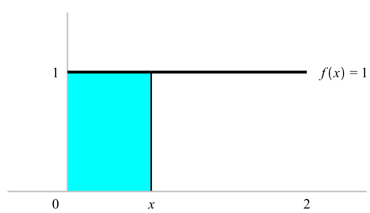
unbestimmte Integral von f .

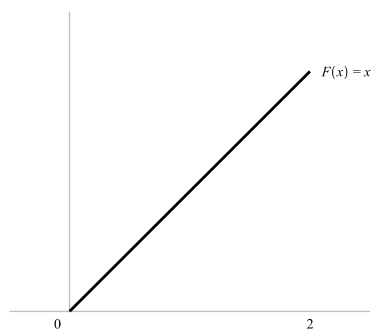
Beispiele:

(i) Wir betrachten

$$f : [0,2] \to \mathbb{R},\ f(x) \equiv 1$$

$$\Rightarrow\ \mathcal{F}(x) = \int\limits_0^x 1\, dt = x \quad \text{(s. Beispiel (i) aus § 1).}$$



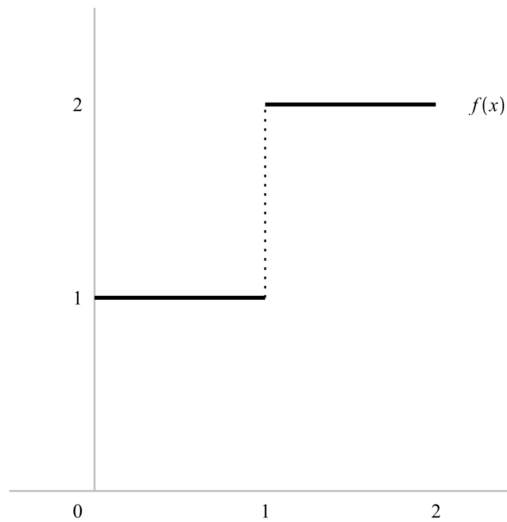


Beobachtung:

$\mathcal{F}'(x)$ existiert für jedes $x \in [0,2]$, und es ist $\mathcal{F}'(x) = f(x)$.

(ii) Wir betrachten nun

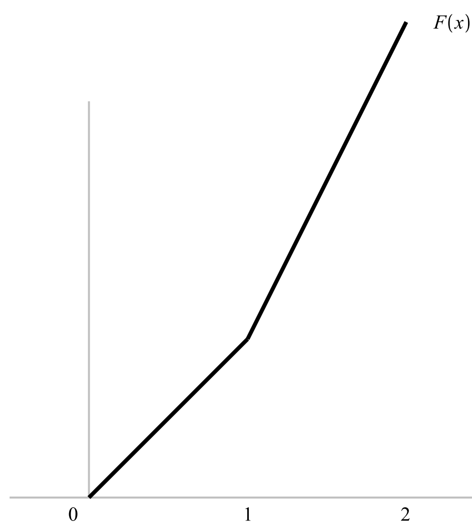
$$f : [0,2] \to \mathbb{R},\ f(x) := \begin{cases} 1 & 0 \le x \le 1 \\ 2 & 1 < x \le 2 \end{cases}$$



f ist $\uparrow \;\Rightarrow\; f \in R[0,2]$, und wir haben

$$\mathcal{F}(x) = \begin{cases} x & 0 \le x \le 1 \\ 1 + \int\limits_1^x 2\,dt & 1 < x \le 2 \end{cases} =$$

$$= \begin{cases} x & 0 \le x \le 1 \\ 2x - 1 & 1 < x \le 2 \end{cases}$$



Beobachtung:

$\mathcal{F}$ ist stetig, $\mathcal{F}'(1)$ existiert nicht wegen $\mathcal{F}'(1-) = 1$, $\mathcal{F}'(1+) = 2$.

Satz 2.1:

Ist f $\boxed{\text{(Riemann-)integrierbar}}$ über $[a,b]$, so ist das unbestimmte Integral $\mathcal{F}(x) = \int\limits_a^x f(t)\,dt$ stetig auf dem ganzen Intervall $[a,b]$.

Beweis:

f ist beschränkt. Wähle $K > 0$ mit $|f(x)| \le K \quad \forall\ x \in [a,b]$.

Für $x, y \in [a,b]$ gilt dann:

$$|\mathcal{F}(y) - \mathcal{F}(x)| = \Big|\int\limits_a^y f(t)\,dt - \int\limits_a^x f(t)\,dt\Big| = \Big|\int\limits_x^y f(t)\,dt\Big| \le K|y-x| \to 0,\ y \to x.$$

Definition:

Es sei $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ ein Intervall. Eine Funktion $F : I \to \mathbb{R}$ heißt **Stammfunktion** zu f , falls F auf I differenzierbar ist mit

$$F'(x) = f(x)\ \forall\, x \in I.$$

Bemerkungen und Beispiele:

(i) Beispiel: $f(x) \equiv 1$ besitzt auf $I = \mathbb{R}$ eine Stammfunktion, etwa $F(x) = x$.

(ii) Betrachte $f(x) = \dfrac{1}{x}$ auf $I = (0,\infty)$.

f besitzt eine Stammfunktion, etwa $F(x) = \ln x$.

Betrachte noch einmal $f(x) = \dfrac{1}{x}$, jetzt aber auf dem Intervall $(-\infty, 0)$.

Hier besitzt f auch eine Stammfunktion, etwa $F(x) = \ln(-x)$.

Manchmal sagt man auch kurz:

Eine Stammfunktion von $f(x) = \dfrac{1}{x}$ ist gegeben durch $F(x) = \ln|x|$ auf $\mathbb{R} \setminus \{0\}$.

(ii) Sind F_1, F_2 zwei Stammfunktionen von f auf I, so gibt es eine Konstante $c \in \mathbb{R}$ mit

$$F_1(x) = F_2(x) + c \quad \forall\ x \in I,$$

d.h. eine Stammfunktion ist eindeutig bis auf eine additive Konstante.

Beweis: $(F_1 - F_2)' = F_1' - F_2' = f - f \equiv 0 \Rightarrow F_1 - F_2 \equiv$ const.

Satz 2.2 (1. Hauptsatz der D/I-Rechnung):
Es sei $f : [a,b] \to \mathbb{R}$ $\boxed{\text{stetig}}$ auf $[a,b]$. Dann ist das unbestimmte Integral $\mathcal{F}(x) = \int\limits_a^x f(t)\, dt$ eine Stammfunktion von f , also:

$\mathcal{F}'$ existiert $\forall\ x \in [a,b]$ und es ist $\mathcal{F}'(x) = f(x) \quad \forall\ x \in [a,b]$.

Beweis:
Sei $c \in [a,b]$ fest und $\varepsilon > 0$ beliebig. Wähle $\delta > 0$ so, dass gilt:

$$|t-c| < \delta \ \Rightarrow\ |f(t) - f(c)| < \varepsilon \quad \text{das (geht, denn} \quad f \quad \text{ist stetig bei} \quad c)$$

Für $|x - c| < \delta$ ist dann

$$\Big|\frac{\mathcal{F}(x) - \mathcal{F}(c)}{x-c} - f(c)\Big| = \Big|\frac{1}{x-c}\int\limits_c^x f(t)\, dt - \frac{1}{x-c}\int\limits_c^x f(c)\, dt\Big| =$$

$$\frac{1}{|x-c|}\Big|\int\limits_c^x \Big(f(t) - f(c)\Big)\, dt\Big| \le \frac{1}{|x-c|}\Big|\int\limits_c^x \underbrace{|f(t)-f(c)|}_{\le\varepsilon}\, dt\Big| \le$$

$$\le \frac{1}{|x-c|}\Big|\int\limits_c^x \varepsilon\, dt\Big| = \varepsilon.$$

Beispiele:

- $\displaystyle \frac{d}{dx}\int\limits_1^x e^t\, dt = e^x$

- $\displaystyle \frac{d}{dx}\int\limits_1^{x^2} \frac{e^t}{t}\, dt = \frac{e^{x^2}}{x^2}\cdot 2x = 2\,\frac{e^{x^2}}{x}, \quad (x \ne 0).$

  (beachte hierbei auch die Kettenregel)

Bemerkung:
Dier Sätze 2.1 und 2.2 fasst man gerne zusammen in der Formulierung

“Integrieren glättet.”

Damit soll zum Ausdruck gebracht werden, dass man beim Übergang von f zu $\mathcal{F}$ die nächst bessere Funktionenklasse erreicht, wenn man von der Hierarchie

$$f \quad \text{differenzierbar} \quad \Rightarrow f \quad \text{stetig} \quad \Rightarrow f \quad \text{integrierbar}$$

ausgeht.

Satz 2.3 (2. Hauptsatz der D/I-Rechnung):
Es sei $f \in R[a,b]$, und es sei $F : [a,b] \to \mathbb{R}$ eine Stammfunktion von f .
Dann gilt:

$$\int_a^b f(x)\,dx = F(b) - F(a) \quad \Big(= F(x)\Big|_a^b \Big).$$

Beweis:
Sei $Z : a = x_0 < x_1 < \ldots < x_n = b$ eine beliebige Zerlegung von $[a,b]$.

$$\Rightarrow F(b) - F(a) \underbrace{=}_{\text{Teleskop}} \sum_{k=1}^{n} F(x_k) - F(x_{k-1})$$

$$\underbrace{=}_{\substack{\text{MWS} \\ \xi_k \in (x_{k-1}, x_k)}} \sum_{k=1}^{n} F'(\xi_k)(x_k - x_{k-1}) =$$

$$\underbrace{=}_{F'=f} \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$$

$$\leq \sum_{k=1}^{n} M_k(x_k - x_{k-1}) = O(Z;f)$$

und

$$\geq \sum_{k=1}^{n} m_k(x_k - x_{k-1}) = U(Z;f).$$

Z war beliebig

$$\Rightarrow F(b) - F(a) \leq \int_a^{-b} f(x)\,dx \underbrace{=}_{f \in R[a,b]} \int_a^b f(x)\,dx$$

und

$$F(b) - F(a) \geq \int_{-a}^{b} f(x)\,dx \underbrace{=}_{f \in R[a,b]} \int_a^b f(x)\,dx.$$

Beispiele:

Ein bestimmtes Integral $\int_a^b f(x)\,dx$ ist also leicht berechenbar, **wenn man eine Stammfunktion** F **zu** f **gefunden hat,** etwa

- $\int_1^3 1\,dx = x\Big|_1^3 = 2;$
- $\int_0^1 x^2\,dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$ (s. Eingangsbeispiel)

- $\displaystyle\int_1^2 \frac{1}{x}\,dx = \ln x\Big|_1^2 = \ln 2.$

- $\displaystyle\int_0^{2\pi} \sin x\,dx = -\cos x\Big|_0^{2\pi} = -1+1 = 0.$

  $\displaystyle\int_0^{\pi} \sin x\,dx = -\cos x\Big|_0^{\pi} = 1-(-1) = 2.$

- $\displaystyle\int_{-T}^{T} \frac{1}{1+x^2}\,dx = \arctan T - \arctan(-T).$

  Interessant ist hier: Der Grenzwert für $T \to \infty$ existiert und ist $= \dfrac{\pi}{2} - \left(-\dfrac{\pi}{2}\right) = \pi$. Wir werden später schreiben

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \pi.$$

§ 3 Integrationstechniken

Das Auffinden von Stammfunktionen ist oft recht schwierig. Nützlich können dabei die folgenden Hinweise sein:

A. Grundintegrale

Diese ergeben sich aus den Differentiationsregeln der elementaren Funktionen,etwa :

$$(e^x)' = e^x \;\Rightarrow\; \int e^x\, dx = e^x$$

$$\Big(\int f(x)\, dx \;\text{ stehe dabei für "ist eine Stammfunktion von } f(x)\text{".}\Big)$$

$$\int x^\alpha\, dx = \frac{x^{\alpha+1}}{\alpha+1}\ (\alpha \neq -1) \qquad \int \frac{1}{x}\, dx = \ln|x| \qquad \int e^x\, dx = e^x$$

$$\int a^x\, dx = \frac{a^x}{\ln a} \qquad \int \sin x\, dx = -\cos x \qquad \int \cos x\, dx = \sin x$$

$$\int \frac{1}{1+x^2}\, dx = \arctan x \qquad \int \frac{1}{\sqrt{1-x^2}}\, dx = \arcsin x$$

B. Partielle Integration

Diese Technik ergibt sich aus der Produktregel: $(f \cdot g)' = f'g + fg'$, also

$$\int f' \cdot g\, dx = f \cdot g - \int f \cdot g'\, dx.$$

Beispiele dazu:

(i) $\displaystyle\int xe^x\, dx = xe^x - \int e^x\, dx = xe^x - e^x.$

(ii)
$$\begin{aligned}\int x^2 \cos x\, dx &= x^2 \sin x - \textstyle\int 2x \sin x\, dx = \\ &= x^2 \sin x + 2x\cos x - \textstyle\int 2\cos x\, dx \\ &= x^2 \sin x + 2x\cos x - 2\sin x\ .\end{aligned}$$

Man sieht: so kann man alle Integranden vom Typ Polynom $*f$ mit $f \in \{\, e^x, \cos x, \sin x, \cosh x, \sinh x\,\}$ behandeln.

(iii) $\displaystyle\int x \ln x\, dx = \frac{x^2}{2}\ln x - \int \frac{x^2}{2}\cdot\frac{1}{x}\, dx = \frac{x^2}{2}\ln x - \frac{x^2}{4}.$

(iv) $\displaystyle\int e^x \cos x\, dx = e^x \cos x + \int e^x \sin x\, dx =$

$$= e^x \cos x + e^x \sin x - \int e^x \cos x\, dx$$

$$\Rightarrow \int e^x \cos x\, dx = \frac{1}{2}\Big(e^x \cos x + e^x \sin x\Big).$$

(Das ist die "Phönix-aus-der-Asche-Methode".)

(v) $\displaystyle\int \cos^2 x\, dx = \cos x \sin x + \int \underbrace{\sin^2 x}_{=1-\cos^2 x}\, dx = \frac{1}{2}\Big(\cos x \sin x + x\Big).$

C. Substitution:

Diese Technik ergibt sich aus der Kettenregel: $(f \circ g)' = (f' \circ g) \cdot g'$, also

$$\int f'\Big(g(x)\Big) \cdot g'(x)\, dx = f\Big(g(x)\Big).$$

formales Vorgehen: Man rechnet dabei mit Differentialen dx, wie wenn das reelle Zahlen wären, also:

$$g(x) = s \;\Rightarrow\; g'(x) = \frac{ds}{dx} \;\Rightarrow\; dx = \frac{ds}{g'(x)}$$

Das ist gerechtfertigt, wie folgende Rechnung zeigt:

$$\int f'\Big(g(x)\Big) g'(x)\, dx = \int f'(s)\, ds = f(s) \underbrace{=}_{\text{Rücksubstitution}} f\Big(g(x)\Big).$$

Beispiele:

(i) $\displaystyle\int 2x \cos(x^2)\, dx = \sin(x^2)$ (mit $x^2 = s \;\Rightarrow\; 2x\, dx = ds$)

(ii) $\displaystyle\int \sin(\pi x)\, dx = -\frac{1}{\pi}\cos(\pi x)$ (mit $\pi x = s \;\Rightarrow\; \pi\, dx = ds$)

(iii) $\displaystyle\int x^5 (x^6+1)^6\, dx = \frac{1}{42}\Big(x^6+1\Big)^7$ (mit $x^6 + 1 = u \;\Rightarrow\; 6x^5\, dx = du$)

(iv) $\displaystyle\int e^{-x}\sqrt{1+3e^{-x}}\, dx = -\frac{2}{9}\Big(1+3e^{-x}\Big)^{3/2}$

(mit $1 + 3e^{-x} = w \;\Rightarrow\; -3e^{-x}\, dx = dw$ und

$\displaystyle\int \sqrt{w}\, dw = \int w^{1/2}\, dw = \frac{w^{3/2}}{3/2}$)

Häufig auftretende Spezialfälle:

- $\displaystyle\int f'(ax+b)\, dx = \frac{1}{a}\, f(ax+b) \quad (a \neq 0)$

  (lineare Verschiebung des Arguments: $ax + b = s;\ a\, dx = ds$)

- $\displaystyle\int \frac{f'(x)}{f(x)}\, dx = \ln|f(x)|$ (mit $f(x) = s$),

  z.B. $\displaystyle\int \tan x\, dx = \int \frac{\sin x}{\cos x}\, dx = -\ln|\cos x|$

- $\displaystyle\int f(x) \cdot f'(x)\, dx = \frac{1}{2}\, f^2(x)$ (auch mit $f(x) = s$),

  z. B. $\displaystyle\int \frac{\ln x}{x}\, dx = \frac{1}{2}\, \ln^2 x.$

Bemerkungen:

(i) Natürlich kann man Techniken auch kombinieren, z.B.

$$\int \arctan x \, dx = x \arctan x - \int \frac{x}{x^2+1} \, dx = x \arctan x - \frac{1}{2} \ln(x^2+1).$$

(ii) Manchmal führt auch eine ganz "freche" Substitution zum Ziel, wie das Beispiel $\int \sqrt{1-x^2} \, dx$ zeigt.

Mit $x = \sin u$ ist etwa $dx = \cos u \, du$ und $\sqrt{1-x^2} = \cos u$

$$\int \sqrt{1-x^2} \, dx = \int \cos^2 u \, du \underbrace{=}_{\text{(s.o.)}} \frac{1}{2}\Big(\cos u \sin u + u\Big) =$$

$$= \frac{1}{2}\Big(\sqrt{1-x^2} \cdot x + \arcsin x\Big).$$

D. Rationale Funktionen

Das sind bekanntlich Funktionen vom Typ

$$f(x) = \frac{p(x)}{q(x)}, \; p, q \quad \text{Polynome,.}$$

Diese lassen sich grundsätzlich alle integrieren:

Zunächst kann man, falls grad $p \geq$ grad q ist, p durch q dividieren (Polyomdivision mit Rest).

Beispiel:

$f(x) = \dfrac{x^4 + 3x^3 + 2x^2 + 1}{x^2 + 2x}$: Der Divisionsalgoritmus liefert:

$$\begin{array}{l} (x^4 + 3x^3 + 2x^2 + 1) : (x^2 + 2x) = x^2 + x \\ \underline{x^4 + 2x^3 } \\ x^3 + 2x^2 + 1 \\ \underline{x^3 + 2x^2 } \\ 1 \end{array}$$

Also ist $f(x) = x^2 + x + \dfrac{1}{x^2 + 2x}$.

Für den Rest zitieren wir aus der Algebra den

Satz von der **Partialbruchzerlegung:**

Gegeben seien (reelle) Polynome $p(x)$, $q(x)$ mit grad $p <$ grad q . Dann lässt sich der rationale Ausdruck $f(x) = \dfrac{p(x)}{q(x)}$ schreiben als Summe von endlich vielen sogenannten **Partialbrüchen.** Ein Partialbruch ist dabei ein Ausdruck von folgender Gestalt:

- $\dfrac{\text{const.}}{x-a}, \dfrac{\text{const.}}{(x-a)^2}, \ldots, \dfrac{\text{const.}}{(x-a)^k},$ falls a $k-$ fache Nullstelle von $q(x)$ ist, d.h. $q(x) = (x-a)^k r(x)$, $r(x)$ Polynom mit $r(a) \neq 0$.

- $\dfrac{c_1 x + d_1}{x^2 + ax + b}, \ldots, \dfrac{c_k x + d_k}{(x^2 + ax + b)^k}$, c_i, d_i Konstanten falls $x^2 + ax + b$ über $\mathbb{R}$ nicht zerlegbar ist und als $k-$ facher Faktor in $q(x)$ auftritt. (Bem.: $x^2 + ax + b$ ist in $\mathbb{R}$ nicht zerlegbar $\iff a^2 - 4b < 0$.)

Beispiele:

(i) $\dfrac{1}{x(x+2)} = \dfrac{c_1}{x} + \dfrac{c_2}{x+2}$. Die Koeffizienten c_1 bzw. c_2 erhält man am einfachsten durch die "Zuhaltemethode":

$$c_1 = \frac{1}{x+2}\Big|_{x=0} = \frac{1}{2} \quad \text{bzw.} \quad c_2 = \frac{1}{x}\Big|_{x=-2} = -\frac{1}{2}.$$

(Zuhalten heißt: In Gedanken mit dem Linearfaktor durchmultiplizieren und die Nullstelle einsetzen).

(ii) $\dfrac{1}{x^2(x+1)} = \dfrac{c_1}{x} + \dfrac{c_2}{x^2} + \dfrac{c_3}{x+1}$. Die Koeffizienten c_2 und c_3 erhalten wir mit der Zuhaltemethode zu $c_2 = 1$, $c_3 = 1$. Den Koeffizienten c_1 erhält man durch Einsetzen einer festen Stelle. Setzen wir z.B. $x = 1$ ein, so folgt:

$$\frac{1}{2} = c_1 + 1 + \frac{1}{2} \Rightarrow c_1 = -1.$$

Alternativ könnte man die Differenz ausrechnen:

$$\frac{1}{x^2(x+1)} - \frac{1}{x^2} - \frac{1}{x+1} = \frac{1 - x - 1 - x^2}{x^2(x+1)} = \frac{-x(x+1)}{x^2(x+1)} = \frac{-1}{x}.$$

(iii) $\dfrac{1}{x(x^2+1)} = \dfrac{c_1}{x} + \dfrac{c_2 x + c_3}{x^2+1}$. Für den Koeffizienten c_1 erhalten wir mit der Zuhaltemethode $c_1 = 1$.

Für c_2, c_3 setzt man am besten zwei Werte ein, etwa

$$x = 1 \left(\Rightarrow \frac{1}{2} = 1 + \frac{c_2 + c_3}{2} \right) \quad \text{und} \quad x = -1 \left(\Rightarrow -\frac{1}{2} = -1 + \frac{-c_2 + c_3}{2} \right)$$

und löst das auftretende (lineare) Gleichungssystem, hier mit

$$c_2 = -1, c_3 = 0.$$

Die einzelnen Partialbrüche sind leicht zu integrieren:

(i) $\displaystyle\int \frac{1}{x-a}\,dx = \ln|x-a|.$

(ii) $\displaystyle\int \frac{1}{(x-a)^k}\,dx = \frac{1}{1-k}\cdot\frac{1}{(x-a)^{k-1}} \quad (k \geq 2).$

(iii) $\displaystyle a^2-4b<0 \;\Rightarrow\; \int \frac{1}{x^2+ax+b}\,dx = \frac{2}{\sqrt{4b-a^2}}\arctan\frac{x+a/2}{\sqrt{b-a^2/4}}\quad .$

$\displaystyle\int \frac{2x+a}{x^2+ax+b}\,dx = \ln(x^2+ax+b)$ und für

$\displaystyle\int \frac{cx+d}{(x^2+ax+b)^k}\,dx \quad (k \geq 2)$ gibt es Rekursionsformeln

(s. einschlägige Formelsammlungen).

In unseren Beispielen von oben ist also

(i) $\displaystyle\int \frac{1}{x(x+2)}\,dx = \int\Big(\frac{1/2}{x}-\frac{1/2}{x+2}\Big)\,dx = \frac{1}{2}\ln|x|-\frac{1}{2}\ln|x+2|$

$$\underbrace{=}_{\text{(wenn man will)}} \ln\sqrt{\left|\frac{x}{x+2}\right|}.$$

(ii) $\displaystyle\int \frac{1}{x^2(x+1)}\,dx = \int\Big(-\frac{1}{x}+\frac{1}{x^2}+\frac{1}{x+1}\Big)\,dx = -\ln|x|-\frac{1}{x}+\ln|x+1|.$

(iii) $\displaystyle\int \frac{dx}{x(x^2+1)} = \int\Big(\frac{1}{x}-\frac{x}{x^2+1}\Big)\,dx = \ln|x|-\frac{1}{2}\ln(x^2+1).$

Bemerkungen:

(i) Die beiden in (iii) angegebenen Formeln reichen völlig aus. Betrachte als Beispiel

$$\int\frac{7x-4}{x^2+x+1}\,dx = \int\frac{\frac{7}{2}(2x-\frac{8}{7})}{x^2+x+1}\,dx = \int\frac{\frac{7}{2}(2x+1-\frac{15}{7})}{x^2+x+1}\,dx =$$

$$= \frac{7}{2}\ln(x^2+x+1) - \frac{15}{2}\cdot\frac{2}{\sqrt{3}}\arctan\frac{x+1/2}{\sqrt{3/4}}$$

(ii) Über $\mathbb{C}$ ist Partialbruchzerlegung noch einfacher. Hier sind nämlich auch die in $\mathbb{R}$ nicht zerlegbaren quadratischen Terme in Linearfaktoren zerlegbar. Zum Beispiel ist

$$\frac{1}{x^2+1} = \frac{1}{(x+i)(x-i)} = \frac{c_1}{x+i}+\frac{c_2}{x-i}$$

mit

$$c_1 = \frac{1}{x-i}\Big|_{x=-i} = \frac{1}{-2i} = \frac{i}{2} \quad\text{und}\quad c_2 = \frac{1}{x+i}\Big|_{x=i} = \frac{1}{2i} = -\frac{i}{2}.$$

(iii) Die Partialbruchzerlegung ist auch an anderen Stellen nützlich. So gelingt z.B. eine Summenformel für $\displaystyle\sum_{k=1}^{n} \frac{1}{k(k+1)}$

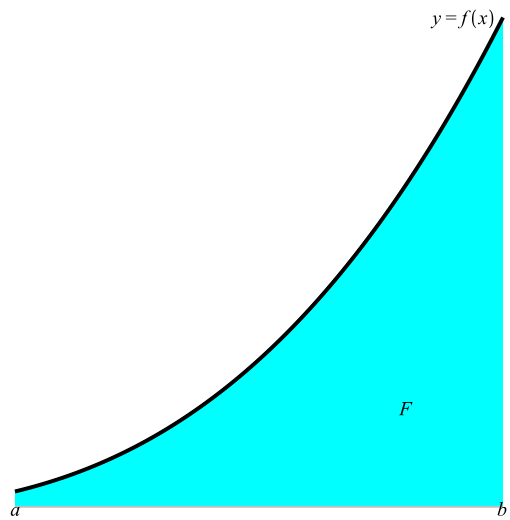
(s. Beispiel zur vollständigen Induktion):

$$\sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) \underbrace{=}_{\text{Teleskop}} 1 - \frac{1}{n+1} = \frac{n}{n+1}.$$

§ 4 Anwendungen der Integralrechnung

Wir wissen bereits (seit der Definition des bestimmten Integrals):

Ist $f(x) \geq 0$ auf $[a, b]$, so gibt $\int_a^b f(x)\,dx$ den Flächeninhalt der von $x = a,\ x = b,\ y = 0$ und $y = f(x)$ begrenzten Fläche an,

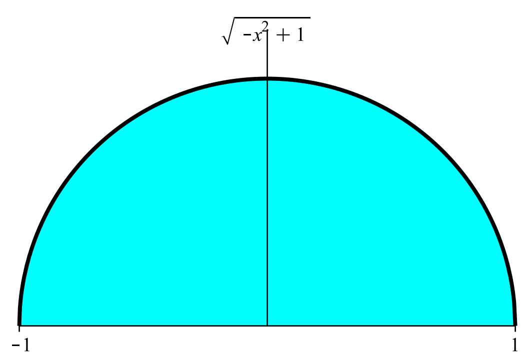


und wir wissen nach dem Hauptsatz

$$\int_a^b f(x)\,dx = F(b) - F(a),$$

wenn F eine Stammfunktion zu f ist, (d.h. $F'(x) = f(x) \quad \forall\ a \leq x \leq b.$)

Als Beispiel können wir die Fläche F des Halbkreises unterhalb $\sqrt{1 - x^2}$ berechnen:



$$F = \int_{-1}^{1} \sqrt{1-x^2}\, dx = \frac{1}{2}\Big(x\sqrt{1-x^2} + \arcsin x\Big)\Big|_{-1}^{1} = \frac{\pi}{2}.$$

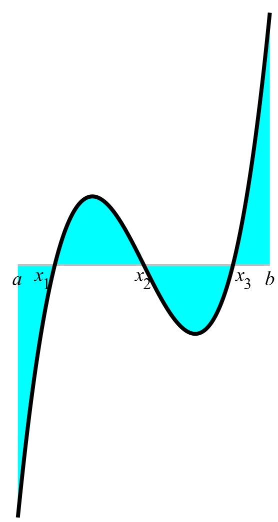
Beim allgemeinen Halbkreis $\int\limits_{-r}^{r} \sqrt{r^2-x^2}\, dx$ ergibt sich $\dfrac{\pi}{2}\, r^2$.

A. Weitere 2d-Flächenberechnungen

(i) Ist $f \ngeq 0$ auf $[a,b]$, so kann man den Flächeninhalt zwischen

$$x = a,\ x = b,\ y = 0 \quad \text{und} \quad y = f(x)$$

trotzdem mit Hilfe von Integralen berechnen.



Man muss nur bedenken, dass diejenigen Teile, die unterhalb der $x-$ Achse liegen, (für die also $f(x) \le 0$ ist), negativ berechnet werden. Man bekommt den Flächeninhalt also in obiger Skizze über

$$\int_{a}^{x_1} -f(x)\, dx + \int_{x_1}^{x_2} f(x)\, dx + \int_{x_2}^{x_3} -f(x)\, dx + \int_{x_3}^{b} f(x)\, dx,$$

was wir natürlich formal abkürzen können über

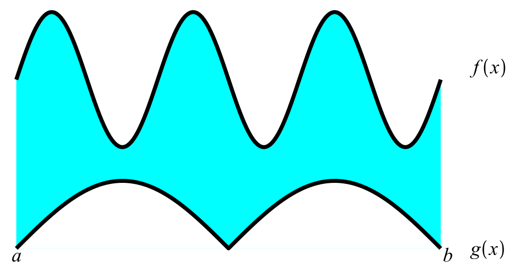
$$\int_{a}^{b} |f(x)|\, dx.$$

Beispielsweise ist der Flächeninhalt zwischen $x-$ Achse und der Sinuskurve im Bereich $[0, 2\pi]$ gegeben durch

$$\int_{0}^{2\pi} |\sin x|\, dx = \int_{0}^{\pi} \sin x\, dx + \int_{\pi}^{2\pi} -\sin x\, dx = -\cos x\Big|_{0}^{\pi} + \cos x\Big|_{\pi}^{2\pi} = 4.$$

(ii) Flächeninhalt zwischen $x = a,\ x = b,\ y = f(x)$ und $y = g(x)$:

(a) $f \geq g$ auf ganz $[a, b]$.



Dann ist der gesuchte Flächeninhalt offenbar gleich

$$\int_a^b f(x)\ dx - \int_a^b g(x)\ dx = \int_a^b \Big(f(x) - g(x)\Big)\ dx.$$

Es spielt dabei keine Rolle, ob dabei $g \geq 0$ ist oder nicht.

(b) Ist $f(x) \ngeq g(x)$ auf ganz $[a, b]$, so ist werden wie in (i) gewisse Teilflächen negativ berechnet. Man bekommt den richtigen Flächeninhalt hier daher in der Form

$$\int \Big|f(x) - g(x)\Big|\ dx.$$

(Dies wird übrigens gerne als Maß für den Abstand zweier Funktionen benutzt.)

Weitere Flächenberechnungen werden wir mit höherdimensionaler Integralrechnung behandeln.

B. Volumen von Rotationskörpern

Gegeben sei $f : [a, b] \to \mathbb{R},\ f(x) \geq 0$ auf $[a, b]$. Die Kurve $y = f(x)$ rotiere um die $x-$Achse. Wir wollen das Volumen des so entstandenen Rotationskörpers berechnen:

Teilt man $[a, b]$ in Teilintervalle

$$a = x_0 < x_1 < \ldots < x_n = b$$

ein, so ergeben sich als Näherung für diesen Körper Ober- bzw. Untersummen über

$$\sum_{i=1} \pi \Big(\sup_{[x_{i-1}, x_i]} f(x) \Big)^2 (x_i - x_{i-1}) \quad \text{bzw.} \quad \sum_{i=1} \pi \Big(\inf_{[x_{i-1}, x_i]} f(x) \Big)^2 (x_i - x_{i-1})$$

(Wir benutzen dabei: Das Volumen eines senkrechten Kreiszylinders mit Grundkreisradius r und Höhe h ist $= \pi r^2 h$.) Nach Definition des Integrals ist also das gesuchte Volumen V gegeben durch

$$V = \pi \int_a^b f^2(x)\, dx.$$

Beispiele:

(i) Kugel: $f(x) = \sqrt{r^2 - x^2},\ -r \le x \le r$

$$\Rightarrow\ V = \int_{-r}^{r} \pi(r^2 - x^2)\, dx = \pi\Big(r^2 x - \frac{x^3}{3}\Big)\Big|_{-r}^{r} = \frac{4}{3}\,\pi r^3.$$

(ii) senkrechter Kreiskegel mit Grundkreisradius r und Höhe h :

$$f(x) = \frac{r}{h}\, x,\ 0 \le x \le h$$

$$\Rightarrow\ V = \pi \int_0^h \frac{r^2}{h^2}\, x^2\, dx = \pi\, \frac{r^2}{h^2}\, \frac{x^3}{3}\Big|_0^h = \frac{1}{3}\,\pi r^2 h.$$

Weitere 3d-Volumina werden wir mit Hilfe mehrdimensionaler Integralrechnung diskutieren.

C. Bogenlänge von Kurven:

Gegeben $f : [a,b] \to \mathbb{R}$. Die Länge der Kurve $y = f(x),\ a \le x \le b,$ ist bei Wahl einer Zerlegung

$$a = x_1 < x_1 < \ldots < x_n = b$$

in etwa die Länge des Streckenzuges durch die Punkte

$$\Big(x_k, f(x_k)\Big),\ k = 0,\ \ldots,\ n,$$

also nach Pythagoras

$$\approx \sum_{k=1}^{n} \sqrt{\Big(x_k - x_{k-1}\Big)^2 + \Big(f(x_k) - f(x_{k-1})\Big)^2} =$$

$$\sum_{k=1}^{n} (x_k - x_{k-1}) \sqrt{1 + \Big(\underbrace{\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}}_{= f'(\xi_k)\ (\text{MWS})}\Big)^2}\ \underbrace{\rightarrow}_{\text{Feinheit } \to 0}$$

$$\int_a^b \sqrt{1 + f'^2(x)}\, dx.$$

Beispiel (Einheitskreis):

$$f(x) = \sqrt{1-x^2},\ 0 \le x \le c \ \ (c \le 1 \quad \text{fest})$$

$$\text{Bogenlänge} \ = \int_0^c \sqrt{1+\frac{x^2}{1-x^2}}\, dx = \int_0^c \frac{1}{\sqrt{1-x^2}}\, dx = \arcsin c.$$

(daher übrigens auch der Name "arcus" = (lat.) Bogen).

Die Bogenlänge ist auch für $c = 1$ sinnvoll, denn $\arcsin 1 = \frac{\pi}{2}$. Allerdings hat der Integrand $\frac{1}{\sqrt{1-x^2}}$ bei $x = 1$ eine Polstelle, und damit ist $\int_0^1 \frac{1}{\sqrt{1-x^2}}\, dx$ im Riemannschen Sinne nicht erklärt. Solche Integrale nennt man **uneigentlich.** Wir diskutieren sie in § 5.

D. Integralrestglied beim Satz von Taylor

Mit dem Integralbegriff können wir eine weitere Darstellung des Fehlerterms im Satz von Taylor herleiten:

Satz (Integralrestgliedformel von Taylor):

Gegeben $f : I \to \mathbb{R}, f \in C^{(n+1)}(I)$ und $a \in I$. Dann gilt:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{1}{n!}\int_a^x f^{(n+1)}(t)\ (x-t)^n\ dt.$$

(Diese Darstellungsformel für das Restglied im Satz von Taylor ist sogar exakt, kommt also ohne Zwischenstelle aus.)

Beweis: (Induktion nach n)

- $n = 0:\ f(x) = f(a) + \int_a^x f'(t)\ dt$ stimmt (Hauptsatz)

- $n \to n{+}1 : f(x) \underbrace{=}_{\text{(IV)}} \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x{-}a)^k + \frac{1}{n!}\int_a^x f^{(n+1)}(t)\ (x{-}t)^n\ dt \underbrace{=}_{\text{(partiell)}}$

  $$\sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x{-}a)^k + \frac{1}{n!} f^{(n+1)}(t) \frac{-(x-t)^{n+1}}{n+1}\Big|_{t=a}^{t=x} + \frac{1}{n!}\int_a^x f^{(n+2)}(t)\frac{(x-t)^{n+1}}{n+1}\ dt =$$

  $$\sum_{k=0}^{n+1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{1}{(n+1)!}\int_a^x f^{(n+2)}(t)\ (x-t)^{n+1}\ dt.$$

§ 5 Uneigentliche Integrale

Manchmal, vor allem in der Stochastik, treten auch Integrale auf, bei denen eine der Grenzen ∞ oder $-\infty$ ist (oder sogar beide), z.B.

$$\int_0^\infty e^{-x}\,dx \quad \text{oder} \quad \int_{-\infty}^\infty e^{-x^2}\,dx.$$

Solche Integrale nennt man **uneigentlich.**

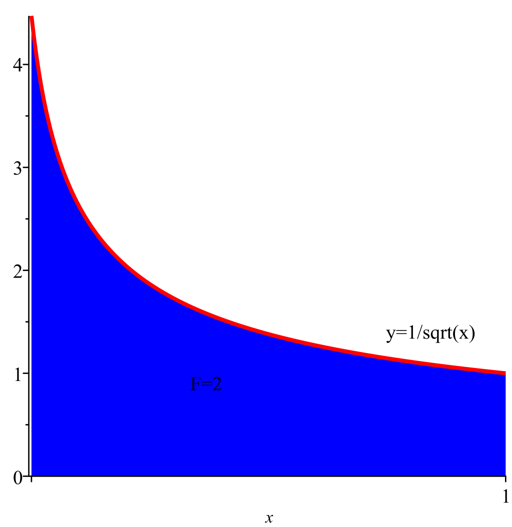
Manche davon haben trotzdem noch einen endlichen Wert. Diese nennen wir **(konvergente Integrale)**.Bei manchen ist der Wert nicht mehr endlich **(divergente Integrale).**

Gleiches kann auch bei endlichen Bereichen $[a, b]$ vorkommen, wenn der Integrand bei a oder/und b Probleme bereitet (z.B. einen Pol hat).

Die Vorgehensweise ist denkbar einfach. Wir integrieren nur bis in die "Nähe" T dieser kritischen Stelle, und untersuchen, ob dann ein Grenzwert existiert.

Beispiele:

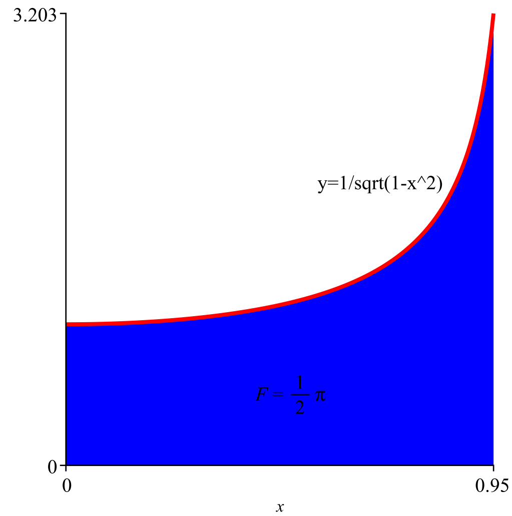
(i) $\displaystyle\int_0^1 \frac{1}{\sqrt{x}}\,dx$:



Rechnung:

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx = \lim_{T\to 0+} \int_T^1 \frac{1}{\sqrt{x}}\,dx = \lim_{T\to 0+} 2\sqrt{x}\Big|_T^1 = \lim_{T\to 0+}(2 - 2\sqrt{T}) = 2.$$

(ii) $\displaystyle\int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx$

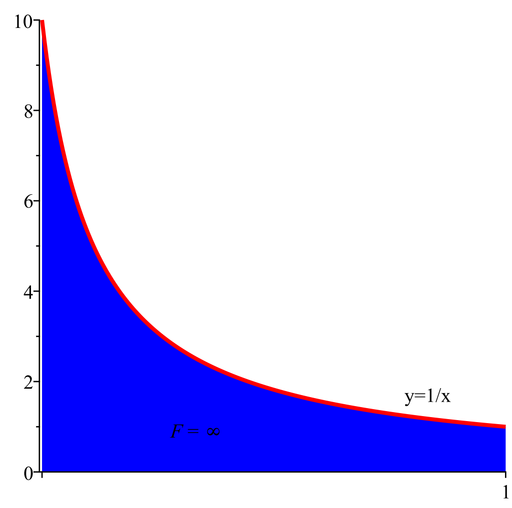


Rechnung:

$$\int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx = \lim_{T\to 1-}\int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx = \lim_{T\to 1-} \arcsin T = \frac{\pi}{2}.$$

(iii) $$\int_0^\infty e^{-x}\,dx = \lim_{T\to\infty}\int_0^T e^{-x}\,dx = \lim_{T\to\infty}(1-e^{-T}) = 1;$$

(iv) $$\int_0^1 \frac{1}{x}\,dx = \lim_{T\to 0+}\int_T^1 \frac{1}{x}\,dx = \lim_{T\to 0+} -\ln T = \infty \quad \text{(also divergent)}$$



(v) $$\int_1^\infty \frac{1}{x^2}\,dx = \lim_{T\to\infty}\int_1^T \frac{1}{x^2}\,dx = \lim_{T\to\infty} 1 - \frac{1}{T} = 1$$

(vi) $\displaystyle\int_1^\infty \frac{1}{x}\,dx = \lim_{T\to\infty}\int_1^T \frac{1}{x}\,dx = \lim_{T\to\infty} \ln T = \infty$

(vii) $\displaystyle\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx = \lim_{T\to\infty}\int_{-T}^T \frac{1}{1+x^2}\,dx = \lim_{T\to\infty} \arctan T - \arctan(-T) = \pi$

(viii) Sei $\alpha \neq 1$:

$$\int_1^\infty \frac{1}{x^\alpha}\,dx = \lim_{T\to\infty}\int_1^T \frac{1}{x^\alpha}\,dx = \lim_{T\to\infty} \frac{T^{-\alpha+1}}{-\alpha+1} + \frac{1}{\alpha-1} \to \begin{cases} \infty & \alpha < 1 \\ \frac{1}{\alpha-1} & \alpha > 1 \end{cases}$$

$$\int_0^1 \frac{1}{x^\alpha}\,dx = \lim_{T\to 0+}\int_T^1 \frac{1}{x^\alpha}\,dx = \lim_{T\to\infty} -\frac{T^{-\alpha+1}}{-\alpha+1} + \frac{1}{1-\alpha} \to \begin{cases} \infty & \alpha > 1 \\ \frac{1}{1-\alpha} & \alpha < 1 \end{cases}$$

(Damit ist $\displaystyle\int_0^\infty \frac{1}{x^\alpha}\,dx$ für jedes α divergent.)

Bemerkung:

Gelingt es nicht, ein uneigentliches Integral explizit auszurechnen (etwa über eine Stammfunktion), so kann man u.U. durch Abschätzen oder Analysieren des Integranden auf Konvergenz oder Divergenz schließen. Beispiele dazu:

(i) $\displaystyle\int_1^\infty \frac{\sin x}{x^2}\,dx$ ist konvergent, denn $\displaystyle\left|\frac{\sin x}{x^2}\right| \leq \frac{1}{x^2}$ und $\displaystyle\int_1^\infty \frac{1}{x^2}\,dx$ ist konvergent (s. Bsp. (v)). Man spricht von einer konvergenten Majorante.

(ii) $\displaystyle\int_1^\infty e^{-x^2}\,dx$ konvergiert, denn $e^{-x^2} \leq e^{-x}$ und $\displaystyle\int_1^\infty e^{-x}\,dx$ konvergiert

(iii) $\displaystyle\int_0^1 \frac{\sin x}{x}\,dx$ konvergiert, denn $\displaystyle\lim_{x\to 0}\frac{\sin x}{x} = 1$ existiert. Der Integrand ist also bei 0 stetig ergänzbar.

(iv) $\displaystyle\int_0^1 \frac{\sin x}{x^2}\,dx$ divergiert, denn $\displaystyle\frac{\sin x}{x^2} = \frac{\sin x}{x}\cdot\frac{1}{x} \approx \frac{1}{x}$, $x \sim 0$ und $\displaystyle\int_0^1 \frac{1}{x}\,dx$ ist divergent.

(v) $\displaystyle\int_0^\infty e^{-t}t^{x-1}\,dt$ konvergiert für jedes $x > 0$, denn:

- $\int\limits_0^1 e^{-t}t^{x-1}\,dt$ konvergiert, denn:

  $$e^{-t}t^{x-1} \leq t^{x-1},\ 0 \leq t \leq 1 \ \text{ und } \int\limits_0^1 t^{x-1}\,dt = \frac{t^x}{x}\Big|_{t=0}^{t=1} = \frac{1}{x} < \infty.$$

- $\int\limits_1^\infty e^{-t}t^{x-1}\,dt$ konvergiert auch, denn:

  $$\lim_{t\to\infty} e^{-t/2}t^{x-1} = 0 \ \Rightarrow\ \exists\, K > 0 :\ |e^{-t/2}t^{x-1}| \leq K$$

  $$\Rightarrow\ |e^{-t}t^{x-1}| \leq K \cdot e^{-t/2} \ \text{ und } \int\limits_1^\infty Ke^{-t/2}\,dt < \infty$$

Das letzte Beispiel führt zu folgender

Definition:

Für $x > 0$ definieren wir $\Gamma(x) := \int\limits_0^\infty t^{x-1}e^{-t}\,dt$ und nennen

$\Gamma : (0,\infty) \to \mathbb{R}$ die **Eulersche Gamma-Funktion.**

Diese besitzt folgende Eigenschaft:

- $\Gamma(x+1) = x\Gamma(x)\ \ \forall\, x > 0$, denn:

  $$\Gamma(x+1) = \int\limits_0^\infty t^x e^{-t}\,dt \underbrace{=}_{\text{(partiell)}} \underbrace{-t^x e^{-t}\Big|_{t=0}^{t=\infty}}_{=0} + \int\limits_0^\infty x t^{x-1}e^{-t}\,dt = x\Gamma(x)$$

- $\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \ldots = n(n-1)\ldots\Gamma(1) =$
  $n!\underbrace{\int\limits_0^\infty e^{-t}\,dt}_{=1} = n!\ \ \forall\ n \in \mathbb{N}.$

  Das bedeutet: $\Gamma(x+1)$ interpoliert die diskrete Folge $n!$.

*

ENDE

* *

Anhang: Allgemeine Grundlagen

§ 1 Grundbegriffe der Mengenlehre

Wir benutzen einen naiven Mengenbegriff, der auf Georg Ferdinand Ludwig Philipp **Cantor** (1845-1918) zurückgeht:

Definition:

Unter einer **Menge** verstehen wir die Zusammenfassung

bestimmter wohlunterschiedener

Objekte unserer Anschauung oder unseres Denkens zu einem Ganzen.

Diese Objekte nennen wir die **Elemente** dieser Menge.

Ausgehend von Mengen $A, B, \ldots$ benutzen wir folgende Bezeichnungen und Begriffe:

(i)

$$x \in A : \quad x \quad \text{ist Element der Menge } A.$$

$$x \notin A : \quad x \quad \text{ist nicht Element der Menge } A.$$

Unsere Definition beinhaltet also, dass für jedes Objekt x **genau** eine der Aussagen

$$x \in A \quad \text{bzw.} \quad x \notin A \quad \text{zutrifft.}$$

(ii) Zur Beschreibung von Mengen gibt es verschiedene Möglichkeiten:

- $$A = \{\, a, b, c, \ \ldots \,\}$$

  Dann besteht A aus den Elementen $a, b, c, \ \ldots$.

  Dies ist die aufzählende, nicht immer präzise Notation.

- Sei E eine Eigenschaft. Dann besteht die Menge

  $$A = \{\, x : \ x \quad \text{besitzt die Eigenschaft E} \,\}$$

  genau aus den Objekten, welche die Eigenschaft E besitzen. E kann man verbal ausdrücken oder mathematisch formulieren.

  Beispiel:

  $$A = \{\, n : \ n \quad \text{ist eine gerade natürliche Zahl} \,\}$$

  $$\text{alternativ:} \quad A = \{\, x : \ \exists\, k \in \mathbb{N} \quad \text{mit} \quad x = 2k \,\}$$

  Wir benutzen dabei das Symbol $\mathbb{N}$ für die Menge der natürlichen Zahlen, die wir als bekannt voraussetzen. Das Symbol $\exists$ steht für "es existiert" oder "es gibt". Eine übliche Beschreibung unserer Menge A ist auch

  $$A = \{\, 2k : \ k \in \mathbb{N} \,\},$$

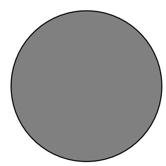
  was selbsterklärend ist.

(iii) Mit $\emptyset$ (gelegentlich auch $\{\}$) bezeichnen wir die Menge ohne Elemente, die sogenannte **leere Menge.** Zum Beispiel ist bekanntlich

$$\{ x : x \quad \text{ist eine reelle Zahl mit} \quad x^2 = -1 \} = \emptyset.$$

Die Menge der reellen Zahlen bezeichnen wir mit $\mathbb{R}$, und sehen diese als bekannt an. Ebenso setzen wir die Mengen $\mathbb{Z}$ der ganzen Zahlen und $\mathbb{Q}$ der rationalen Zahlen mit ihrer Arithmetik als bekannt voraus. Die dazu grundlegenden algebraischen Strukturen diskutieren wir später.

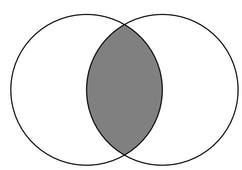
(iv) Zur Veranschaulichung benutzt man gerne sogenannte **Venn-Diagramme** (nach John Venn (1834-1923))



(v)

$$A \cap B := \{ x : x \in A \quad \text{und} \quad x \in B \}$$

heißt der **Durchschnitt** oder die **Schnittmenge** zwischen A und B. Es ist üblich, das Symbol $:=$ zu verwenden, wenn man die Gleichheit als Definition der entsprechenden Größe verstanden haben will.

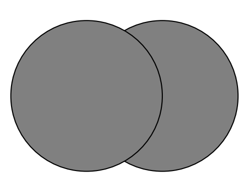


(vi)

$$A \cup B := \{ x : x \in A \quad \text{oder} \quad x \in B \}$$

heißt die **Vereinigung** von A mit B . Das Wort "oder" ist dabei inclusiv gemeint:

$$x \in A \cup B :\Longleftrightarrow \quad x \in A \quad \text{oder} \quad x \in B \quad \text{oder auch} \quad x \in A \cap B$$



(vii) Man kann auch mehr als zwei Mengen miteinander schneiden oder vereinigen. Bei endlich vielen Mengen ist es klar, wie man deren Durchschnitt bzw. deren Vereinigung zu definieren hat. Für beliebig viele

Mengen hat sich folgende Notation durchgesetzt: Fassen wir diese beliebig vielen Mengen zu einer neuen Menge $\mathcal{A}$ zusammen, so definieren wir

$$\bigcup_{A\in\mathcal{A}} A := \{\, x :\ \exists\, A \in \mathcal{A} \quad \text{mit} \quad x \in A \,\} \quad \text{bzw.}$$

$$\bigcap_{A\in\mathcal{A}} A := \{\, x :\ x \in A \quad \forall\, A \in \mathcal{A} \,\}$$

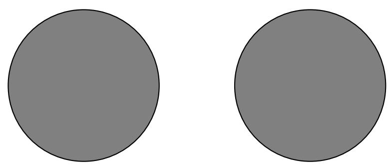
Das aufgetretene Symbol $\forall$ steht für "für alle".

(viii) A und B heißen **disjunkt** oder **elementefremd,** falls gilt:

$$A \cap B = \emptyset.$$

Will man dies bei der Vereinigung ausdrücklich betonen, so schreibt man auch

$$A + B \quad \text{statt} \quad A \cup B$$



(ix) A heißt **Teilmenge** von B, in Zeichen $A \subset B$, falls gilt:

$$x \in A \;\Rightarrow\; x \in B$$

(also in Worten: jedes Element von A ist auch Element von B.)

Das gleiche soll auch die Notation $B \supset A$ bzw. die Sprechweise

"B ist **Obermenge** von A" bedeuten.

(x) A und B heißen **gleich**, in Zeichen $A = B$, falls gilt:

$$A \subset B \quad \text{und} \quad B \subset A, \quad \text{d.h.} \quad x \in A \iff x \in B$$

Ist $A \subset B$, aber $A \neq B$, so schreibt man auch $A \subsetneqq B$ und sagt:

A ist eine **echte** Teilmenge von B.

(xi) Die Menge

$$\mathcal{P}(B) := \{\, A :\ A \subset B \,\}$$

nennt man die **Potenzmenge** von B.

Beispielsweise ist für die zweielementige Menge $B = \{\, a, b \,\}$

$$\mathcal{P}(B) = \Big\{ \{a,b\}, \{a\}, \{b\}, \emptyset \Big\}.$$

$\mathcal{P}(B)$ besteht hier also aus vier Elementen. Die Elementezahl einer Menge A bezeichnet man oft mit $|A|$.

Frage: $|B| = n \in \mathbb{N} \;\Rightarrow\; |\mathcal{P}(B)| =?$ (Antwort: 2^n.)

(xii) Die mengentheoretische Differenz "A ohne B" definiert man über

$$A \setminus B := \{x : \ x \in A \quad \text{und} \quad x \notin B\}$$

Ist $A \subset X$, so nennen wir $X \setminus A$ das **Komplement** von A bezüglich X und schreiben hierfür A_X^C oder kurz A^C, wenn die Obermenge X aus dem Kontext heraus klar ist.

(xiii)

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

also die Menge aller geordneten Paare, heißt das **kartesische Produkt** von A mit B, benannt nach René Descartes (1596-1650). Dabei wollen wir festlegen:

$$(a, b) = (a', b') \quad \text{mit} \quad a, a' \in A, \ b, b' \in B :\Longleftrightarrow \ a = a' \quad \text{und} \quad b = b'.$$

Allgemeiner sei für Mengen $A_1, \ \ldots, \ A_n$ $(n \in \mathbb{N})$

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

die Menge aller geordneten $n-$Tupel mit analoger Gleichheitsdefinition.

Statt $A \times A$ schreiben wir auch A^2, statt $A \times A \times \ldots \times A$ mit n gleichen Faktoren A auch A^n.

Die Gleichheitsdefinition zweier Mengen in (x) birgt ein Beweisprinzip. Wir illustrieren dies an einem Beispiel: Wir behaupten:

$$\{ \ (x, y) \in \mathbb{R}^2 : \ x + y = 1 \ \} = \{ \ (t, 1 - t) : \ t \in \mathbb{R} \ \}$$

Zum Beweis setzen wir

$$A := \{ \ (x, y) \in \mathbb{R}^2 : \ x + y = 1 \ \}, \ B := \{ \ (t, 1 - t) : \ t \in \mathbb{R} \ \}.$$

- Sei $(x, y) \in A \ \Rightarrow \ x, y \in \mathbb{R}$ und $x + y = 1$

  $\Rightarrow$ mit $t := x$ ist $t \in \mathbb{R}$ und $y = 1 - t \ \Rightarrow \ (x, y) = (t, 1 - t) \in B$.

  Da (x, y) beliebig war, folgt $A \subset B$.

- Sei $(t, 1 - t) \in B \ \Rightarrow \ t \in \mathbb{R}$. Mit $x := t, \ y := 1 - t$ folgt:

  $x, y \in \mathbb{R}$ und $x + y = t + (1 - t) = 1 \ \Rightarrow \ (t, 1 - t) = (x, y) \in A$.

  Da $(t, 1 - t)$ beliebig war, folgt $B \subset A$.

- Wegen $A \subset B$ und $B \subset A$ folgt letztendlich $A = B$.

Wir werden dies später nicht mehr in dieser Ausführlichkeit zeigen. Zu diesem Beispiel ist noch zu sagen, dass es sich bei den beiden Mengen A und B um die Beschreibung einer Geraden im 2d-Raum handelt. Bei A ist es die Darstellung in der Form

$$x + y = 1 \ \Longleftrightarrow \ y = 1 - x,$$

bei B um die sogenannte Parameterdarstellung:

$$(t, 1-t) = (0,1) + t(1,-1):$$

"Aufpunkt" (0,1), "Richtungsvektor" $(1,-1)$. Die beiden Rechnungen in unserem Beweis liefen auf Auflösen der Gleichung $x+y=1$ nach einer Variablen (hier war es nach y) bzw. dem Einsetzen der Auflösung ("Probe") hinaus.

Von einer Vielzahl von Rechenregeln seien die folgenden erwähnt:

Seien A, B, C, X Mengen. Dann gilt:

(i) $A \cap B = B \cap A,\ A \cup B = B \cup A$ (Kommutativgesetze)

(ii) $(A \cap B) \cap C = A \cap (B \cap C)$ und
$(A \cup B) \cup C = A \cup (B \cup C)$ (Assoziativgesetze)
Deswegen lässt man hier Klammern i.a. weg.

(iii) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ und
$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ (Distributivgesetze)
Deswegen ist hier Klammersetzung nötig.

(iv) $A \subset X \ \Rightarrow\ \left(A^C\right)^C = A.$

(v) $A \cup B = (A \setminus B) + (A \cap B) + (B \setminus A).$

(vi) $A, B \subset X \ \Rightarrow\ (A \cap B)^C = A^C \cup B^C$ und
$(A \cup B)^C = A^C \cap B^C.$ (de Morgansche Regeln)
(nach Augustus de Morgan (1806-1871)).

Beweis:

Die meisten dieser Regeln sind unmittelbar klar und sind einfache Übungsaufgaben. Wir zeigen den ersten Teil von (vi):

- Sei $x \in (A \cap B)^C \ \Rightarrow\ x \notin A \cap B \ \underbrace{\Rightarrow}_{\text{(Logik)}}\ x \notin A$ oder $x \notin B$
  $\Rightarrow\ x \in A^C$ oder $x \in B^C \ \Rightarrow\ x \in A^C \cup B^C.$
  Da x beliebig war, folgt insgesamt $(A \cap B)^C \subset A^C \cup B^C$.
- Sei $x \in A^C \cup B^C \ \Rightarrow\ x \in A^C$ oder $x \in B^C \ \Rightarrow\ x \notin A$ oder $x \notin B$
  $\Rightarrow x \notin A \cap B \ \Rightarrow\ x \in (A \cap B)^C.$
  Wieder war x beliebig. Also folgt: $A^C \cup B^C \subset (A \cap B)^C$.

Zusammen mit der Definition der Mengengleichheit folgt daraus nun die Behauptung.

Natürlich hätte man den Beweis abkürzen können, wenn man sich in jedem Schritt überlegt, ob der Schluss in beide Richtungen in Ordnung ist.

Für allgemeine Vereinigungen und Durchschnitte gelten die de Morganschen Regeln auch. Sei dazu $\mathcal{A}$ eine Menge von Teilmengen von X. Dann gilt:

$$\Big(\bigcap_{A\in\mathcal{A}} A\Big)^C = \bigcup_{A\in\mathcal{A}} A^C \quad \text{und} \quad \Big(\bigcup_{A\in\mathcal{A}} A\Big)^C = \bigcap_{A\in\mathcal{A}} A^C$$

Wir beweisen den ersten Teil (jetzt kürzer als vorhin):

$x \in \Big(\bigcap_{A\in\mathcal{A}} A\Big)^C \iff x \notin \bigcap_{A\in\mathcal{A}} A \iff \exists\, A \in \mathcal{A}$ mit $x \notin A \iff$
$\exists\, A \in \mathcal{A}$ mit $x \in A^C \iff x \in \bigcup_{A\in\mathcal{A}} A^C$.

§ 2 Funktionen

Definition:

Eine **Funktion** oder **Abbildung** $f : A \to B$ besteht aus

- einer Menge A, dem **Definitionsbereich** von f,
- einer zweiten Menge B, dem **Bildbereich** von f und
- einer Zuordnungsvorschrift, die

  $$\underline{\text{jedem}}\text{ Element } a \in A \ \underline{\text{genau}}\text{ ein Element } b \in B$$

  zuordnet.

Wir schreiben dann $b = f(a)$, nennen b das **Bild** von a (unter f), und a (ein) **Urbild** von b (unter f).
Schreibweise: $f : A \to B,\ a \to f(a)$.

Beispiele:

(i) $A = \{a, b, c\},\ B = \{\alpha, \beta, \gamma\},\ f : a \to \alpha,\ b \to \beta,\ c \to \gamma$.
Dies ist eine Funktion; a besitzt das Bild α, β besitzt das (einzige) Urbild b.

(ii) $A = \{a, b, c\},\ B = \{\alpha, \beta, \gamma\},\ f : a \to \alpha,\ a \to \beta,\ b \to \beta,\ c \to \gamma.$
Das ist keine Funktion in unserem Sinn, denn die Zuordnung von a ist nicht eindeutig.

(iii) $A = \{a, b, c\},\ B = \{\alpha, \beta\},\ f : a \to \alpha,\ b \to \beta$.
Das ist keine Funktion in unserem Sinn, denn $c \in A$ wird nichts zugeordnet.

(iv) $A = \{a, b, c\},\ B = \{\alpha, \beta\},\ f : a \to \alpha,\ b \to \beta,\ c \to \alpha$.
Dies ist eine Funktion in unserem Sinn; das Bild von a unter f ist α. α besitzt zwei Urbilder, nämlich a und c.

(v) $A = \{a, b\},\ B = \{\alpha, \beta, \gamma\},\ f : a \to \alpha,\ b \to \beta.$
Dies ist eine Funktion in unserem Sinne; $\gamma \in B$ besitzt kein Urbild unter f.

Definition:

Gegeben seien eine Funktion $f : A \to B$ sowie Teilmengen $A_1 \subset A$, $B_1 \subset B$. Dann heißt

(i) $G(f) := \{\big(a, f(a)\big) : a \in A\} \subset A \times B$ der **Graph** von f;

(ii) $f(A_1) := \{f(a) : a \in A_1\}\ (= \{b \in B : \exists\ a \in A_1 \text{ mit } f(a) = b\}) \subset B$ das **Bild** von A_1 unter f

(iii) $f^{-1}(B_1) := \{a \in A : f(a) \in B_1\} \subset A$ das **Urbild** von B_1 unter f ;

(iv) $f\Big|_{A_1} : A_1 \to B,\ f\Big|_{A_1}(a) := f(a)$ die **Restriktion** oder **Einschränkung** von f auf A_1.

(v) f **surjektiv** oder Abbildung von A <u>auf</u> B , falls $f(A) = B$ gilt;

(vi) f **injektiv** oder **eineindeutig**, falls gilt:

$$a_1, a_2 \in A \text{ mit } a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2);$$

(vii) f **bijektiv**, falls f sowohl surjektiv als auch injektiv ist

Bemerkungen und Beispiele:

(i) In obigem Beispiel (i) ist $f(\{a,b\}) = \{\alpha, \beta\}$, in Beispiel (iv) ist $f^{-1}(\{\alpha\}) = \{a, c\}$.

(ii)

$$\begin{aligned} f \quad \text{ist surjektiv} \quad &\iff f(A) = B \\ &\iff B \subset f(A) \quad (\text{denn } B \supset f(A) \text{ gilt ja immer}) \\ &\iff \forall\ b \in B\ \exists\ a \in A \quad \text{mit} \quad f(a) = b \\ &\iff \text{jedes } b \in B \text{ besitzt } \underline{\text{mindestens}} \text{ ein Urbild in } A . \end{aligned}$$

Demnach sind die Funktionen in den Beispielen (i),(iv) surjektiv, nicht aber die Funktion in Beispiel (v).

(iii)

$$f \quad \text{ist injektiv} \iff [a_1, a_2 \in A \text{ mit } f(a_1) = f(a_2) \Rightarrow a_1 = a_2]$$

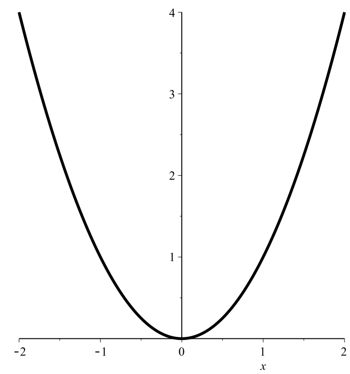
$$\iff \quad \text{jedes Element aus } B \text{ besitzt } \underline{\text{höchstens}} \text{ ein Urbild in } A .$$

Demnach sind die Funktionen aus den Beispielen (i) und (v) injektiv, nicht aber die aus Beispiel (iv).

(iv) f ist demnach bijektiv $\iff$ jedes Element aus B besitzt <u>genau ein</u> Urbild in A . In obigen Beispielen ist lediglich die Funktion aus (i) bijektiv.

(v) Konkret werden wir die Begriffe anwenden auf Funktionen, die per Formelvorschrift gegeben werden. Betrachte dazu als Beispiel die Zuordnungsvorschrift $x \to x^2$:

- $A = B = \mathbb{R} : G(f) = \{(x, x^2) : x \in \mathbb{R}\} \subset \mathbb{R}^2$, oft in einem kartesischen Koordinatensystem zeichnerisch dargestellt.



So ist f

– nicht injektiv, denn z.B. ist

$$f(-1) = f(1) = 1,$$

aber

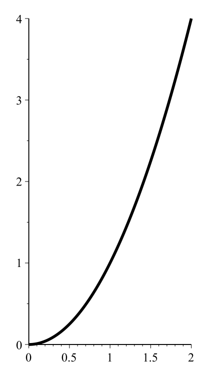
$$-1 \neq 1.$$

– und auch nicht surjektiv, denn z.B. besitzt

$$-1 \in B$$

kein Urbild.

- $A = \mathbb{R},\ B = [0, \infty)\ (:= \{y \in \mathbb{R} : y \geq 0\})$: Jetzt ist f surjektiv, aber immer noch nicht injektiv.
- $A = B = [0, \infty)$.



Jetzt ist f bijektiv. Details dazu lernen Sie in der Analysis kennen.

Definition:

Es sei $f : A \to B$ bijektiv. Dann definieren wir die **Umkehrfunktion**

$$f^{-1} : B \to A,$$

indem wir jedem $b \in B$ dasjenige $a \in A$ zuordnen, für das $f(a) = b$ gilt.

Bemerkungen:

(i) Diese Definition bedarf einer genauen Begründung. Es ist möglich, jedem $b \in B$ ein Element $a \in A$ so zuzuordnen (also mit $f(a) = b$), denn f ist surjektiv, aber auch nur eines, denn f ist injektiv. Man beachte die Probleme, die auftreten würden, wenn f nicht injektiv ist (s. Beispiel (iv): Was soll $f^{-1}(\alpha)$ sein) bzw. nicht surjektiv ist (s. Beispiel (v): $f^{-1}(\gamma) =?$).

Die Umkehrfunktion unserer bijektiven Funktion aus Beispiel (i) ist gegeben durch

$$f^{-1}: \ \alpha \to a, \ \beta \to b, \ \gamma \to c.$$

(ii) Offenbar gilt: Ist $f : A \to B$ bijektiv,so gilt

$$f^{-1}\Big(f(a)\Big) = a \ \ \forall\, a \in A \quad \text{und} \quad f\Big(f^{-1}(b)\Big) = b \ \ \forall\, b \in B.$$

(iii) Praktisch bestimmt man f^{-1}, indem man die Gleichung $f(a) = b$ nach a auflöst:

$$f(a) = b \iff a = f^{-1}(b).$$

Definition:

(i) Gegeben seien Funktionen $g : A \to B$, $f : B \to C$. Dann heißt die Funktion $f \circ g : A \to C$, definiert durch

$$(f \circ g)(a) := f\Big((g(a)\Big) \quad \text{für} \quad a \in A$$

die **zusammengesetzte Funktion** oder **Komposition** f nach g .

(ii) Die Funktion $\mathrm{id}_A : A \to A$, $\mathrm{id}_A(a) := a \ \ \forall\, a \in A$ heißt die **Identität** auf A .

Bemerkungen:

(i) $f \circ g$ existiert also nur, wenn der Definitionsbereich von f mit dem Bildbereich von g übereinstimmt.

(ii) $f : A \to B$ bijektiv $\Rightarrow f^{-1} \circ f = \mathrm{id}_A$ und $f \circ f^{-1} = \mathrm{id}_B$.

(iii) Beispiel: $A = B = C = \mathbb{R}$, $f(x) = x + 1$, $g(x) = 2x \Rightarrow$

$$(f \circ g)(x) = f(2x) = 2x + 1 \quad \text{und}$$

$$(g \circ f)(x) = g(x + 1) = 2(x + 1) = 2x + 2.$$

Man sieht: Auch wenn beide Kompositionen $f \circ g$ und $g \circ f$ bildbar sind, müssen diese nicht gleich sein.

(iv) Es sei $f : A \to B$ bijektiv. Wir zeigen: Dann ist auch $f^{-1} : B \to A$ bijektiv, und es gilt:

$$(f^{-1})^{-1} = f.$$

Dazu:

- Surjektivität: Sei $a \in A$: Setze $b := f(a) \Rightarrow$

$$b \in B \quad \text{und} \quad f^{-1}(b) = f^{-1}\Big((f(a)\Big) = a$$

- Injektivität: Seien $b_1, b_2 \in B$ mit $f^{-1}(b_1) = f^{-1}(b_2) \Rightarrow$

$$f\Big(f^{-1}(b_1)\Big) = f\Big(f^{-1}(b_2)\Big) \Rightarrow b_1 = b_2.$$

- Betrachte die Funktionen $f : A \to B$ und $(f^{-1})^{-1} : A \to B$. Für $a \in A$ gilt dann:

$$a = f^{-1}\Big((f^{-1})^{-1}(a)\Big) \Rightarrow f(a) = f\Big(f^{-1}(f^{-1})^{-1}(a))\Big) = (f^{-1})^{-1}(a).$$

(iv) Es seien $g : A \to B, f : B \to C$ beide bijektiv. Dann ist $f \circ g : A \to C$ bijektiv, und es gilt:

$$(f \circ g)^{-1} = g^{-1} \circ f^{-1}.$$

§ 3 Relationen und Äquivalenzrelationen

Definition:

Es sei A eine nichtleere Menge. Eine Teilmenge

$$\sim\ \subset A \times A$$

nennen wir eine **Relation** auf A .

Bemerkung:

Diese Definition ist reichlich formal und entspricht auch nicht der "üblichen" Schreibweise:

Statt $(x, y) \in\ \sim$ mit $x, y \in A$ schreiben wir kurz: $x \sim y$.

Beispiel:

$A = \{\ 1, 2, 3\ \}$; $\sim\ = \{\ (1,2), (1,3), (2,3)\ \}$, also

$$1 \sim 2,\ 1 \sim 3,\ 2 \sim 3.$$

Diese Relation nennen wir landläufig die "Kleiner-Relation" und schreiben das auch anders:

$$1 < 2,\ 1 < 3,\ 2 < 3$$

Definition:

Eine Relation auf $A\ (\neq \emptyset)$ heißt **Äquivalenzrelation**, falls gilt:

(i) $x \sim x \quad \forall\ x \in A$ **(Reflexivität)**

(ii) $x, y \in A,\ x \sim y\ \Rightarrow\ y \sim x$ **(Symmetrie)**

(iii) $x, y, z \in A,\ x \sim y,\ y \sim z\ \Rightarrow\ x \sim z$ **(Transitivität)**

Beispiele:

(i) A wie in obigem Beispiel. Die Kleiner-Relation ist zwar transitiv, aber weder reflexiv noch symmetrisch.

(ii) $A \neq \emptyset$ beliebig. Durch $x \sim y :\Longleftrightarrow\ x = y$ ist eine Äquivalenzrelation definiert.

(iii) Die Elemente von A sei eine Gruppe von Menschen. Durch

$$x \sim y :\Longleftrightarrow\ x \quad \text{und} \quad y \quad \text{besitzen dasselbe Geschlecht,}$$

wird eine Äquivalenzrelation definiert.

Bei der Überprüfung der einzelnen Eigenschaften einer Äquivalenzrelation treten u.U. aktuelle gesellschaftliche Problemstellungen zutage (gender-Fragen).

(iv) Nun noch ein Beispiel aus der Mathematik: Sei $A = \mathbb{Z}$, und

$x \sim y :\Longleftrightarrow x{-}y$ ist eine gerade Zahl, d.h. $\exists\ k \in \mathbb{Z}$ mit $x{-}y = 2k$.

- Reflexivität: $x \sim x\ \ \forall\, x \in A$, denn

$$x - x = 0 = 2 \cdot 0 \quad \text{und} \quad 0 \in \mathbb{Z}$$

- Symmetrie: Sei $x \sim y$, etwa $x - y = 2k$ mit $k \in \mathbb{Z}$

$$\Rightarrow\ y - x = 2 \cdot (-k) \quad \text{und} \quad -k \in \mathbb{Z}\ \Rightarrow\ y \sim x$$

- Transitivität: Seien $x \sim y$ und $y \sim z$, etwa

$$x - y = 2k,\ y - z = 2m \quad \text{mit} \quad k, m \in \mathbb{Z}$$

$$\Rightarrow\ x - z = (x - y) + (y - z) = 2k + 2m = 2 \cdot (k + m).$$

Wegen $k + m \in \mathbb{Z}$ folgt $x \sim z$.

Damit ist $\sim$ eine Äquivalenzrelation auf $\mathbb{Z}$.

Definition:
Es sei $\sim$ eine Relation auf $A\ (\neq \emptyset)$. Dann nennen wir für $a \in A$ die Menge

$$K(a) := \{\ b :\ b \in A \quad \text{mit} \quad b \sim a\ \}$$

die a zugeordnete **Klasse**.

Satz (Klasseneinteilung):
Es sei $\sim$ eine Äquivalenzrelation auf $A\ (\neq \emptyset)$. Dann gilt:

(i) Für jedes $a \in A$ ist $K(a) \neq \emptyset$.

(ii) $\bigcup\limits_{a \in A} K(a) = A$

(iii) Für $a, b \in A$ gilt entweder $K(a) = K(b)$ oder $K(a) \cap K(b) = \emptyset$.

Dies bedeutet anschaulich: Über eine Äquivalenzrelation wird A zerlegt in nichtleere, paarweise disjunkte Klassen.

Beweis:

(i) das ist klar, denn $a \in K(a)\ \ \forall\, a \in A$, denn $\sim$ ist reflexiv

(ii) $\subset$ ist klar und $\supset$ auch wegen (i)

(iii) Wir zeigen: Ist $K(a) \cap K(b) \neq \emptyset$, so gilt bereits $K(a) = K(b)$.

Sei also $c \in K(a) \cap K(b)$. Dann gilt also $c \sim a$ und $c \sim b$. Wegen der Symmetrie gilt dann aber auch $a \sim c$ und $b \sim c$.

- Sei $x \in K(a) \Rightarrow x \sim a \underbrace{\Rightarrow}_{\text{(trans.)}} x \sim c \underbrace{\Rightarrow}_{\text{(trans.)}} x \sim b \Rightarrow x \in K(b)$.
- Sei $x \in K(b) \Rightarrow x \sim b \underbrace{\Rightarrow}_{\text{(trans.)}} x \sim c \underbrace{\Rightarrow}_{\text{(trans.)}} x \sim a \Rightarrow x \in K(a)$.

Also gilt: $K(a) \subset K(b)$ und $K(b) \subset K(a)$, also $K(a) = K(b)$.

Bemerkungen:

(i) Bei einer Äquivalenzrelation $\sim$ auf A nennen wir die Mengen $K(a)$ **Äquivalenzklassen** und jedes $b \in K(a)$ einen **Repräsentanten** dieser Klasse.

Eine Teilmenge von A , welche aus jeder Äquivalenzklasse genau einen Repräsentanten enthält, heißt ein **vollständiges Repräsentantensystem**.

(ii) (vgl. Bsp.(ii)) Gleichheit liefert triviale, weil einelementige Äquivalenzklassen: $K(a) = \{a\} \ \forall\, a \in A$.

(iii) (vgl. Bsp.(iii)) Die Äquivalenzklassen könnte man hier bezeichnen als Frauen und Männer (bei allen politischen Schwierigkeiten)..

(iv) In Beispiel (iv) ist

$K(0) = \{\ldots, -4, -2, 0, 2, 4, \ldots\}$, also die Menge der geraden Zahlen und

$K(1) = \{\ldots, -3, -1, 1, 3, \ldots\}$ die Menge der ungeraden Zahlen.

$K(0) + K(1) = \mathbb{Z}$.

Vollständige Repräsentantensysteme sind etwa $\{0, 1\}$ oder $\{-2, 17\}$.

Bemerkung: Man kann übrigens mit diesen Klassen wieder rechnen. Definiert man nämlich + und $\cdot$ wie unten, so entsteht eine algebraische Struktur, die man einen Körper nennt (s. Kapitel II, § 3):

+	K(0)	K(1)
K(0)	K(0)	K(1)
K(1)	K(1)	K(0)

$\cdot$	K(0)	K(1)
K(0)	K(0)	K(0)
K(1)	K(0)	K(1)

Das bedeutet z.B., dass die Summe zweier ungerader Zahlen gerade ist.

(v) (vgl.Bsp.(i)) $K(1) = \emptyset,\ K(2) = \{1\},\ K(3) = \{1, 2\}$.

Hier klappt die Konstruktion nicht. Klassen können leer sein oder sich überlappen oder A nicht komplett zerlegen. Alle features treten hier auf. Aber die Kleiner-Relation war ja auch keine Äquivalenzrelation.