IV. Grundlagen der Integral- und Differentialrechnung

17 Grundlagen der Integralrechnung

Sowohl die *Integralrechnung* als auch die *Differentialrechnung* gehören zum Kernbestand der Analysis. Beider gehen ursprünglich von geometrischen Fragestellungen aus:
Bei der *Integralrechnung* will man *Flächeninhalte* oder allgemeiner *Volumina* von geometrischen Objekten bestimmen, aber auch *Längen von Kurven*, Dichten, Schwerpunkte, Mittelwerte oder auch Arbeit (Energie) in einem nicht konstanten Kraftfeld berechnen. Auch die Frage, wie man aus einer *Änderungsrate einer mathematischen oder physikalischen Größe* auf diese Größe selber zurückschließen kann, lässt sich mit Hilfe der Integralrechnung beantworten. Ein Fahrtenschreiber z.B. zeichnet die Geschwindigkeit eines Fahrzeugs auf, mit Hilfe der Integralrechnung lässt sich die Fahrtstrecke rekonstruieren.
Das Prinzip der *Trägheitsnavigation* beruht auf dem Messen der Beschleunigung, der daraus rekonstruierten Geschwindigkeit und dem wiederum hieraus rekonstruieren zurückgelegten Weg.
Geometrische Aspekte bei der Differentialrechnung sind etwa das *Tangentenproblem* (Tangente an eine Kurve), physikalische etwa die *momentane Änderung* einer Größe etwa die Momentangeschwindigkeit oder die Momentanbeschleunigung. Mathematisch handelt es sich hierbei jeweils um das gleiche Problem. Die Differentialrechnung ist auch an sich starkes Hilfsmittel zur Untersuchung qualitativer Eigenschaften von Funktionen (Abbildungen).
Für Funktionen auf Intervallen werden wir sehen, dass die Differential- und Integralrechnung, die zunächst nichts untereinander zu tun zu haben scheinen, über den sogenannten Hauptsatz der *Differential- und Integralrechnung* untereinander schon eng verbandelt sind.
Bei Archimedes (†212 v.Chr.) findet man schon als Vorstufe der Integration die Bestimmung von Flächen- und Rauminhalten, den Zusammenhang zwischen Differential- und Integralrechnung haben Leibniz und Newton um ca. 1670 entdeckt.

Eine systematische Fassung des Flächen- und des Volumenbegriffs für allgemeine geometrische Objekte beginnt mit J.Kepler (1615, Volumenbestimmung von Weinfässern), wichtige Beiträge stammen von dem Galilei-Schüler B.Cavalieri (1635, sog. Cavalierisches Prinzip), dann Newton's Lehrer I.Barrow und viele anderen. Strenge systematische Behandlungen solcher Probleme stehen im engen Zusammenhang mit der Entwicklung der Maß- und Integrationstheorie am Anfang des 20. Jahrhunderts.

Es gibt verschiedene Integralbegriffe und die Mengen, denen man in vernünftiger Weise mit Hilfe des Integrals eine Zahl zuordnen kann, die man als Volumen der Menge ansprechen kann, hängen vom verwendeten Integralbegriff ab. Wir beschränken uns hier auf das sog. *Regel-Integral* (auch Cauchy-Integral genannt, Cauchy 1823).
Das sog. *Riemann-Integral* (B.Riemann) streifen wir nur am Rande.
In *mehreren Veränderlichen* betrachten wir später das von H.Lebesgue 1902 eingeführte *Lebesgue -Integral*, das allen anderen Integralbegriffen weit überlegen ist.
Unser Ziel ist zunächst für eine einfache Klasse von Funktionen -die *Treppenfunktionen*- ein Integral zu definieren und dieses dann auf eine größere Klasse von Funktionen -die *Regelfunktionen*- zu erweitern.

17.1 Das Integral für Treppenfunktionen

Im Folgenden legen wir bis auf Widerruf ein kompaktes Intervall

$$M := [a, b] := \{x \in \mathbb{R};\ a \le x \le b\} \qquad (a, b \in \mathbb{R},\ a < b)$$

zu Grunde.

Schränkt man die Größte-Ganze-Funktion

$$\begin{array}{rcl} [\,]:\mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & [x] = \max\{k \in \mathbb{Z};\ k \leq x\} \end{array}$$

auf ein kompaktes Intervall, etwa $[-1, 4]$ ein, erhält man ein typisches Beispiel für eine Treppenfunktion im Sinne der folgenden allgemeinen Definition (vgl. Abb 9).

17.1.1 Definition (Treppenfunktion, Zerlegung)

Eine Funktion

$$t:\ M \to \mathbb{R}$$

heißt *Treppenfunktion*, wenn es ein $r \in \mathbb{N}$ und reelle Zahlen $x_0, x_1, \ldots, x_r$ mit

$$a = x_0 < x_1 < x_2 < \ldots < x_r = b$$

gibt, so dass die Einschränkungen

$$f|]x_{k-1}, x_k[\text{ für } k = 1, \ldots, r$$

konstant, etwa $= c_k \in \mathbb{R}$, sind.
Man sagt: Die endliche Menge

$$\{a = x_0, x_1, \ldots, x_r = b\}$$

bildet eine *Zerlegung* Z des Intervalls $[a, b]$ und schreibt suggestiv

$$Z := \{a = x_0 < x_1 < \ldots < x_r = b\}.$$

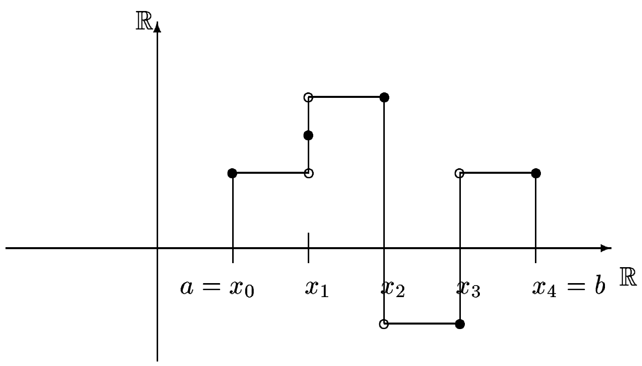
Die Punkte x_k $(0 \leq k \leq r)$ nennt man auch *Teilpunkte* (oder *Stützstellen*) der Zerlegung Z.
Man nennt dann t auch eine Treppenfunktion zur Zerlegung Z.

17.1.2 Bemerkungen

(a) Da t eine Funktion ist, hat t natürlich auch eindeutig bestimmte Werte in den Teilpunkten

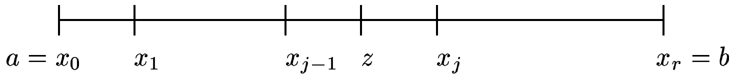
$$x_k\ (0 \leq k \leq r).$$

Diese sind für unsere Integraldefinition allerdings unerheblich.



(b) Bei gegebener Treppenfunktion $t:\ [a,b] \to \mathbb{R}$ ist die Zerlegung Z nicht eindeutig bestimmt, so kann man beispielsweise zu einer gegebenen Zerlegung Z stets Teilpunkte hinzufügen. Ist etwa z ein Punkt mit $x_{j-1} < z < x_j$ für ein geeignetes $j \in \{1, 2, \ldots, r\}$, dann erhält man aus Z durch das Hinzufügen von z die Zerlegung

$$Z' := \{a = x_0 < x_1 < \ldots < x_{j-1} < z < x_j < \ldots < x_r = b\}$$



(c) Wir nennen eine Zerlegung

$$Z^* = \{a = x_0^* < x_1^* < \ldots < x_s^* = b\} \qquad (s \in \mathbb{N})$$

feiner als die Zerlegung

$$Z = \{a = x_0 < x_1 < \ldots < x_{j-1} < z < x_j < \ldots < x_r = b\} \qquad (r \in \mathbb{N})$$

wenn $Z^* \supset Z$ gilt, wenn also jedes x_j $(0 \leq j \leq r)$ unter den x_l^* $(0 \leq l \leq s)$ vorkommt.

(d) Zu je zwei Zerlegungen Z und Z' von $[a,b]$ gibt es stets eine Zerlegung Z^*, die feiner ist als Z und auch feiner als Z'. Offensichtlich ist die Zerlegung $Z^* := Z \vee Z'$, die aus Z durch Hinzunahme der (nicht in Z gelegenen) Punkten von Z' entsteht, eine *gemeinsame Verfeinerung* von Z und Z'.

17.1.3 Definition

Ist $t:\ M \to \mathbb{R}$ eine Treppenfunktion, die bezüglich der Zerlegung

$$Z := \{a = x_0 < x_1 < \ldots < x_{j-1} < z < x_j < \ldots < x_r = b\}$$

definiert ist und gilt etwa $f|\ ]x_{k-1}, x_k[= c_k$ für $k = 1, \ldots, n$, dann heißt die reelle Zahl

$$I_Z(t) := \sum_{k=1}^{r} c_k (x_k - x_{k-1})$$

das *Integral von t bezüglich der Zerlegung Z*.

Wir werden gleich sehen, dass dieses Integral in der Wirklichkeit nicht von der Zerlegung Z abhängt, mit deren Hilfe t erklärt ist.

17.1.4 Satz

Ist $t:\ M \to \mathbb{R}$ eine Treppenfunktion bezüglich der Zerlegung Z und Z', dann gilt

$$I_Z(t) = I_{Z'}(t).$$

Beweis : Wir betrachten zunächst die Zerlegung Z_1, die aus Z durch Hinzufügen eines weiteren Teilpunktes z, etwa Intervall $]x_{j-1}, x_j[$, entsteht. z sei also ein neuer Teilpunkt mit $x_{j-1} < z < x_j$ für ein $j \in \{1, 2, \ldots, n\}$.
Dann ist

$$I_{Z_1}(t) = \sum_{k=1}^{j-1} c_k(x_k - x_{k-1}) + c_j(z - x_{j-1}) + c_j(x_j - z) + \sum_{k=j+1}^{r} c_k(x_k - x_{k-1})$$

Nun ist aber

$$c_j(z - x_{j-1}) + c_j(x_j - z) = c_j z - c_j x_{j-1} + c_j x_j - c_j z = c_j(x_j - x_{j-1}).$$

Daher ist

$$\begin{aligned} I_{Z_1}(t) &= \sum_{k=1}^{j-1} c_k(x_k - x_{k-1}) + c_j(x_j - x_{j-1}) + \sum_{k=j+1}^{r} c_k(x_k - x_{k-1}) \\ &= \sum_{k=1}^{r} c_k(x_k - x_{k-1}) = I_Z(t). \end{aligned}$$

Beim Hinzufügen *eines* weiteren Teilpunktes ändert sich also $I_Z(t)$ nicht. Durch Induktion folgt etwa, dass sich $I_Z(t)$ durch Hinzufügen endlich vieler weiteren Teilpunkte nicht ändert.

Wir betrachten nur zwei Zerlegungen Z und Z' die *gemeinsame Verfeinerung* $Z^* := Z \vee Z'$. Sie entsteht aus Z bzw. Z' durch Hinzunahme endlich vieler weiteren Teilpunkte, daher ist

$$I_Z(t) = I_{Z \vee Z'}(t) = I_{Z'}(t),$$

d.h. also $I_Z(t) = I_{Z'}(t)$.

□

17.1.5 Satz und Definition

Ist $t : [a, b] \to \mathbb{R}$ eine Treppenfunktion und

$$Z = \{a = x_0 < x_1 < \ldots < x_{j-1} < z < x_j < \ldots < x_r = b\}$$

irgendeine Zerlegung von $[a, b]$ mit $t|\ ]x_{k-1}, x_k[= c_k \in \mathbb{R}$ für $k = 1, \ldots, n$, dann heißt die von der Zerlegung unabhängige Zahl

$$I(t) := I_Z(t) := \sum_{k=1}^{r} c_k(x_k - x_{k-1})$$

das *Integral der Treppenfunktion* über das Intervall $M = [a, b]$.

Schreibweise: $I(t) := \int\limits_a^b t = \int\limits_a^b t(x)dx.$

x heißt auch *Integrationsvariable*. Sie ist eine sogenannte freie Variable, kann also mit jedem anderen sinnvollen Buchstaben bezeichnet werden.

Das Integral einer Treppenfunktion ist eine endliche Summe von Produkten.

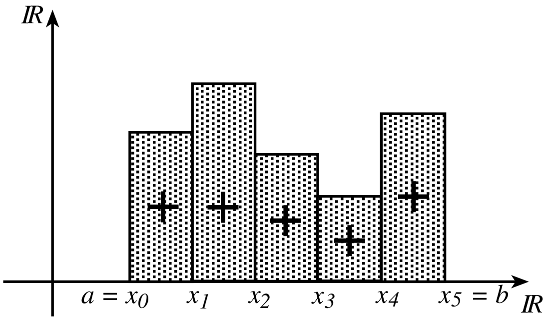
17.1.6 Geometrische Interpretation

Nimmt die Treppenfunktion t nur nicht negative Werte an, dann kann man

$$I(t) = \sum_{k=1}^{r} c_k(x_k - x_{k-1})$$

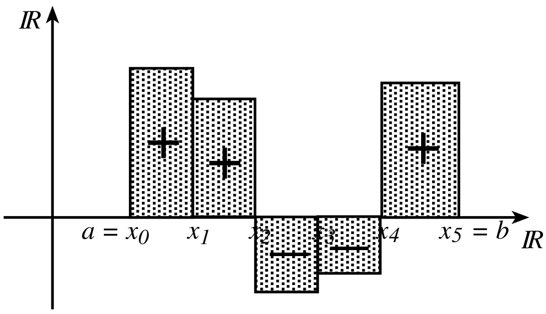
als Flächeninhalt der Punktmenge (Ordinatenmenge)

$$O(t) := \{(x,y) \in \mathbb{R} \times \mathbb{R};\ x \in [a,b];\ 0 \le y \le t(x)\}$$



deuten: $O(t)$ ist eine Vereinigung von Rechtecken und $I(t)$ liefert den elementargeometrischen Flächeninhalt dieser Rechtecke.

Nimmt t auch negative Werte an, dann ist $I(t)$ ein Maß für den *orientierten* Flächeninhalt (d.h. Flächeninhalt mit Vorzeichen) der entsprechenden Rechtecke.



Bezeichnung: Sei $T(M) = \{t : M \to \mathbb{R};\ t\ \text{ Treppenfunktion}\}$ die Menge aller Treppenfunktionen auf $M = [a,b]$, dann gilt

17.1.7 Satz (Eigenschaften von $T(M)$ und $I : M \to \mathbb{R}$)

(a) $T(M)$ ist ein Unterraum von $B(M) := \{f : M \to \mathbb{R};\ f\ \text{ beschränkt}\}$ und damit selbst ein $\mathbb{R}$-Vektorraum und die Abbildung

$$\begin{aligned} I:\ T(M) &\to \mathbb{R} \\ t &\mapsto I(t) \end{aligned}$$

ist linear, d.h. es gilt

(α) $I(t_1 + t_2) = I(t_1) + I(t_2)$ für alle $t_1, t_2 \in T(M)$

(β) $I(ct) = cI(t)$ für alle $c \in \mathbb{R}$ und alle $t \in T(M)$.

Das Integral für Treppenfunktionen ist also ein *lineares Funktional* auf dem Vektorraum $T(M)$.

(b) Für alle $t \in T(M)$ gilt

$$|I(t)| \leq \underbrace{\sup\{|t(x)|;\ x \in M\}}_{||t||_\infty}(b-a) = ||t||_\infty(b-a) \qquad \text{(Standardabschätzung)}$$

(c) Ist $t(x) \geq 0$ für alle $x \in M$ (kurz $t \geq 0$), dann gilt auch $I(t) \geq 0$.
$I : T(M) \to \mathbb{R}$ ist ein nicht negatives lineares Funktional.

(c') Gilt für $t_1, t_2 \in T(M)$, $t_1(x) \leq t_2(x)$ für alle $x \in M$ (kurz: $t_1 \leq t_2$), dann gilt auch

$$I(t_1) \leq I(t_2).$$

Man sagt: I ist ein monotones Funktional.

(d) Mit $t \in T(M)$ gilt auch $|t| \in T(M)$, und es ist

$$|I(t)| \leq I(|t|).$$

Beweis (a): Seien $t_1, t_2 \in T(M)$. Wir können oBdA annehmen, dass t_1 und t_2 bezüglich der gleichen Zerlegung Z von $[a,b]$ definiert sind, dann ist aber klar, dass $t_1 + t_2$ wieder Treppenfunktion ist, die Funktionswerte in den Teilintervallen $]x_{k-1}, x_k[$ sind einfach zu addieren. Sei etwa $t_1|\ ]x_{k-1}, x_k[)c_k$ und $t_2|\ ]x_{k-1}, x_k[= d_k$, dann gilt

$$\begin{aligned} I(t_1+t_2) &= I_Z(t_1+t_2) \\ &= \sum_{k=1}^{r}(c_k+d_k)(x_k-x_{k-1}) \\ &= \sum_{k=1}^{r} c_k(x_k-x_{k-1}) + \sum_{k=1}^{r} d_k(x_k-x_{k-1}) \\ &= I_Z(t_1)+I_Z(t_2) = I(t_1)+I(t_2). \end{aligned}$$

□

Dass mit $t \in T(M)$ auch $ct \in T(M)$ ist evident, ebenso die Regel $I(ct) = cI(t)$

Beweis (b): Es gilt

$$\begin{aligned} |I(t)| &= \left|\sum_{k=1}^{r} c_k(x_k-x_{k-1})\right| \\ &\leq \sum_{k=1}^{r} |c_k|(x_k-x_{k-1}) \qquad (Dreiecksungleichung) \\ &\leq \sum_{k=1}^{r} \underbrace{\sup\{|t(x)|;\ x \in M\}}_{||t||_\infty}(x_k-x_{k-1}) \\ &= ||t||_\infty(b-a), \end{aligned}$$

da $\sum\limits_{k=1}^{r}(x_k - x_{k-1})$ als endliche teleskopische Summe den Wert $b-a$ hat:

$$\begin{aligned} \sum_{k=1}^{r}(x_k-x_{k-1}) &= (x_1-x_0)+(x_2-x_1)+\ldots+(x_r-x_{r-1}) \\ &= -x_0+x_r = x_r-x_0 = b-a \end{aligned}$$

□

Beweis (c): Ist $t(x) \geq 0$, dann gilt speziell $c_k \geq 0$ für alle $k = 1, 2, \ldots, n$ und es folgt

$$I(t) = \sum_{k=1}^{r} c_k \underbrace{(x_k - x_{k-1})}_{>0} \geq 0.$$

□

Beweis (c'): $t_1 \leq t_2 \iff t_2 - t_1 \geq 0$. Aus der Linearität von I und (c) folgt dann

$$0 \leq I(t_2 - t_1) = I(t_2) - I(t_1) \;\text{ oder }\; I(t_1) \leq I(t_2).$$

□

Beweis (d): Aus $t \leq |t|$ folgt nach (c') daher $I(t) \leq I(|t|)$. Weil auch $-t \leq |t|$ gilt, folgt mit (a) und (c')

$$-I(t) = I(-t) \leq I(|t|),$$

also insgesamt

$$|I(t)| \leq I(|t|).$$

□

Wir betrachten ein Beispiel:

17.1.8 Beispiel

Wir betrachten die Treppenfunktion

$$t : [0,1] \to \mathbb{R} \;\text{ mit}$$

$$t(x) := \frac{[nx]}{n},$$

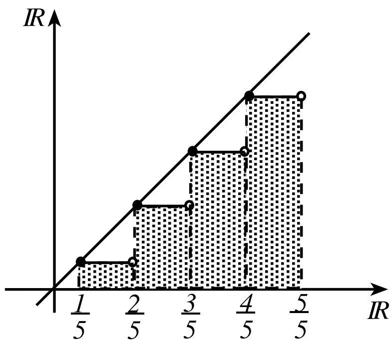
dabei ist $n \in \mathbb{N}$ fest.
Eine zu t passende Zerlegung von $[0,1]$ ist

$$z := 0 = x_0 < \frac{1}{n} < \frac{2}{n} < \ldots < \frac{k}{n} < \ldots < \frac{n-1}{n} < \frac{n}{n} = 1 = x_n.$$

Es ist also $x_k = \frac{k}{n}$ und $x_k - x_{k-1} = \frac{1}{n}$ und $c_k := t|\,]x_k - x_{k-1}[= \frac{k-1}{n}$ für $k = 1, \ldots, n$.
Das Integral $I(t)$ ist leicht zu berechnen:

$$\begin{aligned} I(t) = \sum_{k=1}^{r} c_k (x_k - x_{k-1}) &= \sum_{k=1}^{r} \frac{k-1}{n} \cdot \frac{1}{n} = \frac{1}{n^2} \sum_{k=1}^{r} (k-1) \\ &= \frac{1}{n^2} \frac{n(n-1)}{2} \quad \text{(nach Gauß)} \\ &= \frac{1}{2}\left(1 - \frac{1}{n}\right). \end{aligned}$$

Veranschaulichung für $n = 5$



Lässt man n in alle natürlichen Zahlen durchlaufen und schreibt $t_n(x) = \frac{[nx]}{n}$, dann ist (t_n) eine Folge von Treppenfunktionen, die auf $[0,1]$ gleichmäßig gegen f mit $f(x) = x$ für $x \in [0,1]$ konvergiert.
Beachte dabei auch $\lim\limits_{n\to\infty} \frac{1}{2}\left(1 - \frac{1}{n}\right) = \frac{1}{2}$.

Wir kommen in Abschnitt 17.2 hierauf zurück.

Man kommt zu einer kleinen Erweiterung der Integralfunktion, in der man

$$\int_a^a t = 0$$

definiert und beachtet, dass für eine Treppenfunktion $t : [0,1] \to \mathbb{R}$ für jedes c mit $a < c < b$

$$t|\ [a,c] \quad \text{und} \quad t|\ [c,b]$$

wieder Treppenfunktionen sind, die wir der Einfachheit halber auch wieder mit t bezeichnen.

17.1.9 Satz

Für jedes $t \in T(M)$ und für jedes $c \in [a,b]$ gilt

$$\boxed{\int_a^c t(x)dx + \int_c^b t(x)dx = \int_a^b t(x)dx}$$

Diese Formel bringt eine *Intervalladditivität* des Integrals I zum Ausdruck.

Zum **Beweis** braucht man nur c unter die Teilpunkte der Zerlegung aufzunehmen, mit deren Hilfe man $I(t)$ berechnet.

17.1.10 Bemerkung

Zum Abschluss sei bemerkt, dass mit $t_1, t_2 \in T(M)$ auch $t_1 t_2 \in T(M)$ gilt, das Produkt von zwei Treppenfunktionen ist also wieder eine Treppenfunktion, aber wie einfache Beispiele zeigen (vgl. Übungsblatt 1, Aufgabe 4) gilt i.A. nicht $I(t_1 t_2) = I(t_1)I(t_2)$.

Es gilt jedoch die *Cauchy-Schwarzsche Ungleichung* in der Form

$$\boxed{(I(t_1 t_2))^2 \leq I(t_1^2)I(t_2^2)}$$

Vergleiche hierzu auch die Musterlösung von Aufgabe 4 von Blatt 1.

17.2 Das Integral für Regelfunktionen

Ziel diese Abschnitts ist es, die Integraldefinition, die sich bis jetzt nur auf Treppenfunktionen erstreckt hat, auf eine größere Klasse von Funktionen, die sog. *Regelfunktionen* so auszudehnen, dass die Eigenschaften des Integrals, d.h. der Abbildung

$$\begin{array}{rcl} I: T(M) & \to & \mathbb{R} \qquad (M := [a,b]) \\ t & \mapsto & I(t) = \int\limits_a^b t(x)dx \end{array}$$

erhalten bleiben (I ist ein nicht negatives, beschränktes lineares Funktional). Wir werden sehen, dass stetige bzw. monotone Funktionen auf M Regelfunktionen sind. Da viele der uns geläufigen Funktionen (Polynome höheren Grades, rationale Funktionen, sin, cos, exp, log) keine Treppenfunktionen sind, werden wir versuchen, sie in geeigneter Weise durch Treppenfunktionen zu approximieren.

17.2.1 Definition

Sei $M = [a,b]$ $(a, b \in \mathbb{R},\ a < b)$ ein kompaktes Intervall in $\mathbb{R}$.
Eine Funktion $f : M \to \mathbb{R}$ heißt *Regelfunktion*,wenn es eine Folge (t_n) von Treppenfunktionen $t_n \in T(M)$ gibt, die *gleichmäßig* gegen f konvergiert.

Beachte:
Gleichmäßige Konvergenz ist die Konvergenz bezüglich der Supremumsnorm:
Das die Folge (t_n) gleichmäßig gegen f konvergiert bedeutet also, dass die (reelle) Zahlenfolge $(\|t_n - f\|)$ der Normen eine Nullfolge ist.
Die Supremumsnorm einer beschränkten Funktion $f : M \to \mathbb{R}$ ist dabei die nicht negative reelle Zahl

$$\|f\|_\infty := \|f\|_M = \sup\{|f(x)|;\ x \in M\}$$

Da der Definitionsbereich M zunächst festgehalten wird, lassen wir den Index M auch der Einfachheit halber weg und schreiben nur $\|f\|$.
Die *Supremumsnorm* ist eine *Norm* auf dem Vektorraum $B(M) = \{f : M \to \mathbb{R};\ f \text{ beschränkt}\}$ der beschränkten Funktionen auf M im Sinne von Def. ???.
Für die gleichmäßige Konvergenz der Folge (t_n) gegen f verwenden wir die Kurzschreibweise

$$t_n \rightrightarrows f.$$

Gilt nun $t_n \rightrightarrows f$, d.h. konvergiert (t_n) gleichmäßig gegen f, dann liegt es nahe für die Funktion f ein Integral durch

$$I(f) := \lim_{n\to\infty} I(t_n)$$

zu erklären.
Damit dies sinnvoll ist, ist zweierlei zu zeigen:

(a) Die Folge $(I(t_n))$ der Integrale der Treppenfunktionen t_n ist überhaupt konvergent.

(b) Ist $(\tilde{t}_n)$ eine weitere Folge von Regelfunktionen $\tilde{t}_n \in T(M)$ mit $\tilde{t}_n \rightrightarrows f$, dann gilt

$$\lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} I(\tilde{t}_n).$$

Dass beide Aussagen richtig sind, ist die Aussage von

17.2.2 Lemma

Sei $f: M \to \mathbb{R}$ eine Regelfunktion.

(a) Ist $t_n \in T(M)$ für $n \in \mathbb{N}$ und gilt $t_n \rightrightarrows f$, dann ist die Folge $(I(t_n))$ der Integrale eine konvergente reelle Zahlenfolge.

(b) Gilt auch $\tilde{t}_n \rightrightarrows f$ mit $\tilde{t}_n \in T(M)$, dann ist

$$\lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} I(\tilde{t}_n)$$

Beweis (a): Da wir den potenziellen Grenzwert nicht kennen, verwenden wir das Cauchy-Kriterium, um die Konvergenz der Folge $(I(t_n))$ zu zeigen.
Da (t_n) gleichmäßig gegen f konvergiert, gibt es zu jedem $\varepsilon > 0$ einen Index $N \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt

$$\|t_n - f\| < \frac{\varepsilon}{2(b-a)}$$

Für alle $m \in \mathbb{N}$ mit $m \geq N$ gilt dann auch

$$\|t_m - f\| < \frac{\varepsilon}{2(b-a)}$$

Daher ist

$$\begin{aligned} \|t_m - t_n\| &= \|t_m - f + f - t_n\| \\ &\leq \|t_m - f\| + \|f - t_n\| \\ &= \|t_m - f\| + \|t_n - f\| < \frac{\varepsilon}{2(b-a)} + \frac{\varepsilon}{2(b-a)} = \frac{\varepsilon}{b-a} \end{aligned}$$

für alle $m, n \geq N$.
Nach dieser Vorbereitung betrachten wir die Folge $(I(t_n))$ der Integrale und zeigen, dass sie eine Cauchy-Folge reeller Zahlen, also konvergent ist.
Dazu nutzen wir die Linearität des Integrals für Treppenfunktionen und die Standardabschätzung aus:

$$\begin{aligned} |I(t_m) - I(t_n)| \overset{\text{Linearität}}{=} |I(t_m - t_n)| &\leq (b-a)\|t_m - t_n\| \\ &< (b-a)\frac{\varepsilon}{b-a} = \varepsilon \end{aligned}$$

für alle m, n mit $m, n \geq N$. Die Folge $(I(t_n))$ ist daher als Cauchy-Folge reeller Zahlen konvergent.

□

Beweis (b): Für (b) geben wir zwei Beweise. Da wir nach (a) schon wissen, dass die Folgen $(I(t_n))$ und $(I(\tilde{t}_n))$ konvergent sind, genügt es zu zeigen, dass die Folge der Differenzen eine Nullfolge ist:
Aus

$$\|t_n - \tilde{t}_n\| \leq \|t_n - f\| + \|\tilde{t}_n - f\|$$

ist ersichtlich, dass die Folge $(\|t_n - \tilde{t}_n\|)$ eine Nullfolge ist.
Nun gilt

$$|I(t_n) - I(\tilde{t}_n)| \overset{\text{Linearität}}{=} |I(t_n - \tilde{t}_n)| \overset{\text{Standardabschätzung}}{\leq} (b-a)\|t_n - \tilde{t}_n\|$$

Weil die rechte Seite eine Nullfolge ist, ist auch die linke Seite eine Nullfolge und nach den Rechenregeln für konvergente reelle Zahlenfolgen folgt

$$\lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} I(\tilde{t}_n)$$

□

Ein zweiter Beweis beruht auf dem Prinzip der *Folgenmischung*, das wir schon mehrfach verwendet haben:
Hier bedeutet es:
falls $t_n \rightrightarrows f$ und $\tilde{t}_n \rightrightarrows f$, dann gilt auch, dass die durch Mischung entstehende „Zickzackfolge“ („Reißverschlussfolge“)

$$(t_n^*) = (t_1, \tilde{t}_1, t_2, \tilde{t}_2, t_3, \tilde{t}_3, \ldots)$$

gleichmäßig gegen f konvergiert.
Zum **Beweis** beachte man

$$t_n^* = \begin{cases} t_{\frac{n+1}{2}}, & \text{falls } n = 2m-1,\ m \in \mathbb{N}; \\ \tilde{t}_{\frac{n}{2}}, & \text{falls } n = 2m,\ m \in \mathbb{N}. \end{cases}$$

Nach (a) folgt dann aber, dass die Folge $I(t_n^*))$ der Integrale konvergiert, sei etwa $A := \lim\limits_{n\to\infty} I(t_n^*)$.
Da aber $(I(t_n))$ und $(I(\tilde{t}_n))$ beides Teilfolgen von $(I(t_n^*))$ sind, konvergieren $(I(t_n))$ und $(I(\tilde{t}_n))$ auch beide gegen A, also ist

$$\lim_{n\to\infty} I(t_n) = A = \lim_{n\to\infty} I(\tilde{t}_n).$$

□

Aufgrund des Lemmas ist die folgende Definition sinnvoll.

17.2.3 Definition und Satz (Integral einer Regelfunktion)

Ist $f : M \to \mathbb{R}$ eine Regelfunktion und ist (t_n) irgend eine Folge von Treppenfunktionen t_n auf M, die gleichmäßig gegen f konvergiert: $t_n \rightrightarrows f$, dann existiert der Grenzwert

$$I(f) := \lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} \int_a^b t_n(x)dx.$$

Es ist unabhängig von der Approximationsfolge (t_n) und heißt das *Integral von* f über $[a, b]$.

Bezeichnung: $I(f) = \int_a^b f = \int_a^b f(x)dx$.

x nennt man *Integrationsvariable* und die Funktion f auch *Integrand*.

17.2.4 Bemerkungen

(a) Falls f noch von zusätzlichen Parametern abhängt, z.B. $f(x) = x^2 y$, sollte man die ausführliche Schreibweise verwenden, auf die Bezeichnung der Integrationsvariablen kommt es dabei aber nicht an:

$$\int_a^b f(x)dx = \int_a^b f(u)du = \int_a^b = f(t)dt = \cdots .$$

Wie wir bald sehen werden, gilt z.B.

$$\int_0^1 x^2 y dx = \frac{y}{3}, \quad \text{aber} \quad \int_0^1 x^2 y dy = \frac{x^2}{2}.$$

(b) Wir haben für das Integral einer Regelfunktion keine neue Bezeichnung eingeführt, das macht Sinn, da jede Treppenfunktion $f = t$ eine Regelfunktion ist, als approximierende Folge (t_n) kann man $t_n = t$ für alle $n \in \mathbb{N}$ verwenden:

$$I(f) = I(t) = I(t_n).$$

17.2.5 Beispiel

Ein erstes Beispiel:
Um $\int\limits_0^1 x\ dx$ zu berechnen, betrachten wir die Folge (t_n) von Treppenfunktionen mit $t_n(x) := \frac{[nx]}{n}$, von der wir wissen, dass sie gleichmäßig gegen f mit $f(x) = x$ (sogar auf ganz $\mathbb{R}$) konvergiert.
Für jedes $n \in \mathbb{N}$ ist die Zerlegung $Z^{(n)} := \{0 = x_0 < x_1 < \ldots < x_{k-1} < x_k < \ldots < x_n = 1\}$ mit

$$x_k := \frac{k}{n} \text{ für } k = 0, \ldots, n$$

eine passende Zerlegung zu t_n und $I(t_n) = \frac{1}{2}\left(1 - \frac{1}{n}\right)$ nach ???.
Da $t_n \rightrightarrows id_{[0,1]}$ gilt, ist die Folge $(I(t_n))$ konvergent (das folgt hier auch aus den Rechenregeln für konvergente Folgen), und es gilt

$$\int\limits_0^1 x\ dx = \lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} \frac{1}{2}\left(1 - \frac{1}{n}\right) = \frac{1}{2}.$$

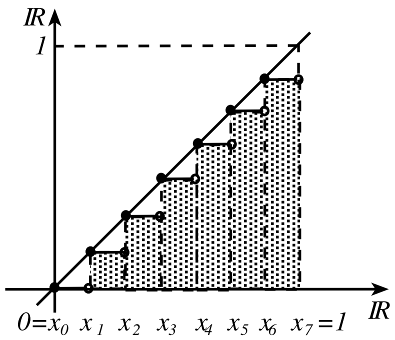


Abbildung 15: Visualisierung von t_7

Wir beschäftigen uns im Folgenden mit *Permanenzeigenschaften des Regelintegrals* und der Frage, wie umfangreich die Klasse der Regelfunktionen ist, einigen speziellen Integralberechnungen und einer Charakterisierung von Regelfunktionen durch innere Eigenschaften.
Da eine Treppenfunktion $t : [a, b] \to \mathbb{R}$ nur endlich viele Werte annimmt, ist jede Treppenfunktion $t : [a, b] \to \mathbb{R}$ beschränkt. Diese Eigenschaft überträgt sich auf Regelfunktionen:

Bezeichnung: $R(M) := \{f : M \to \mathbb{R};\ f \text{ Regelfukntion}\}$

Dann gilt:

17.2.6 Bemerkung

Es gilt $R(M) \subset B(M)$, d.h. jede Regelfunktion ist beschränkt.

Beweis : Gilt $t_n \rightrightarrows f$ mit $t_n \in T(M)$ und beachtet man

$$f = f - t_n + t_n \text{ und}$$

$$(*) \qquad \|f\| \leq \|f - t_n\| + \|t_n\|$$

und die Tatsache, dass Treppenfunktionen beschränkt sind, folgt die Behauptung unmittelbar aus $(*)$.

□

17.2.7 Satz (Eigenschaften des Regelintegrals)

Sind $f, g \in R(M)$ (also Regelfunktionen), dann gilt:

(a) $\alpha) f + g \in R(M)$ und $\beta) cf \in R(M)$ für beliebiges $c \in \mathbb{R}$, das bedeutet. $R(M)$ ist eine $\mathbb{R}$-Vektorraum (Untervektorraum von $B(M)$). Ferner ist

$$I(f+g) = I(f) + I(g) \text{ und } I(cf) = cI(f)$$

($I : R(M) \to \mathbb{R}$ ist also ein lineares Funktional)

Zusatz: Es ist auch $fg \in R(M)$, $R(M)$ ist also sogar eine *Funktionenalgebra*.

(b) $|I(f)| \leq (b-a)\|f\|$ (Standartabschätzung)
(I ist also ein beschränktes lineares Funktional)

(c) Aus $f \geq 0$ folgt $I(f) \geq 0$
(I ist ein nicht negatives lineares Funktional).

(c') Aus $f \leq g$ (d.h. $f(x) \leq g(x)$ für alle $x \in M$) folgt

$$I(f) \leq I(g)$$

(I ist also ein monotones lineares Funktional).

(d) Aus $m_1 \leq f(x) \leq m_2$ für alle $x \in M$ ($m_1, m_2 \in \mathbb{R}$ fest) folgt

$$m_1(b-a) \leq I(f) \leq m_2(b-a)$$

(e) Mit $f \in R(M)$ ist auch $|f| \in R(M)$, und es gilt

$$|I(f)| \leq I(|f|).$$

Beweis (a): Aus $t_n \rightrightarrows f$, $\tilde{t}_n \rightrightarrows g$ $(t_n, \tilde{t}_n \in T(M))$ folgt

$$t_n + \tilde{t}_n \rightrightarrows f + g$$

und damit

$$\begin{aligned} I(f+g) = \lim_{n\to\infty} I(t_n + \tilde{t}_n) &= \lim_{n\to\infty} (I(t_n) + I(\tilde{t}_n)) \\ &= \lim_{n\to\infty} I(t_n) + \lim_{n\to\infty} I(\tilde{t}_n) \\ &= I(f) + I(g). \end{aligned}$$

Wegen $ct_n \rightrightarrows cf$ folgt

$$I(cf) = \lim_{n\to\infty} I(ct_n) = \lim_{n\to\infty} cI(t_n) = c \lim_{n\to\infty} I(t_n) = cI(f).$$

□

Will man den Zusatz beweisen, dann muss man zeigen:
Aus $t_n \rightrightarrows f$ und $\tilde{t}_n \rightrightarrows g$ folgt $t\tilde{t}_n \rightrightarrows fg$. Dazu verwendet man Standardschlüsse:

$$\|fg - t_n\tilde{t}_n\| = \|(f - t_n)g + t_n(g - \tilde{t}_n)\| \leq \|g\|\,\|f - t_n\| + \|t_n\|\,\|g - \tilde{t}_n\|.$$

Mit der Beschränktheit von $\|g\|$ und $\|t_n\|$ folgt die Behauptung.

Beweis (b): Wegen $|\,\|t_n\| - \|f\|\,| \leq \|t_n - f\|$ folgt aus $t_n \rightrightarrows f$ auch $\lim\limits_{n\to\infty} \|t_n\| = \|f\|$, daher

$$\begin{aligned} |I(f)| = |\lim_{n\to\infty} I(t_n)| &= \lim_{n\to\infty} |I(t_n)| \\ &\leq \lim_{n\to\infty} \|t_n\|(b-a) = \|f\|(b-a). \end{aligned}$$

□

Beweis (c): Aus $t_n \rightrightarrows f$ folgt auch $|t_n| \rightrightarrows |f| = f$
(beachte für alle $c, d \in \mathbb{R}$ gilt

$$|\,|c| - |d|\,| \leq |c - d| \text{ und damit (für } alle\ x \in [a,b])$$

$$|\,|t_n(x)| - |f(x)|\,| \leq |t_n(x) - |f(x)|\,| \leq |t_n(x) - f(x)| \leq \|t_n - f\|)$$

Damit gilt auch $|t_n| \rightrightarrows |f|$ und somit $|f| \in R(M)$ und

$$I(f) = I(|f|) = \lim_{n\to\infty} I(|t_n|) \geq 0,$$

da $I(|t_n|) \geq 0$ für alle $n \in \mathbb{N}$ gilt.

□

Beweis (c'): (c') folgt aus (c) wegen der Linearität von I:

$$f \leq g \iff g - f \geq 0,$$

daher

$$I(g) - I(f) = I(g - f) \geq 0,$$

also $I(f) \leq I(g)$.

□

Beweis (d): Wegen $\int\limits_a^b 1 \cdot dx = b - a$, folgt (d) sofort durch mehrfache Anwendung von (c').
Aus $m_1 \cdot 1 \leq f(x)$ folgt $m_1(b-a) \leq I(f)$ und aus $f(x) \leq m_2 \cdot 1$ folgt $I(f) \leq m_1(b-a)$, also

$$m_1(b-a) \leq I(f) \leq m_2(b-a).$$

□

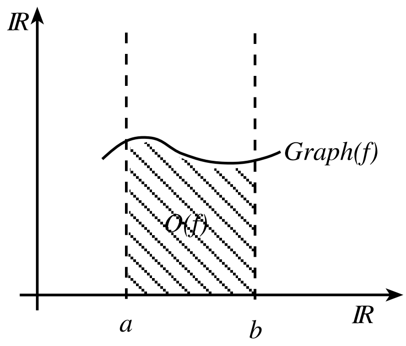
Beweis (e): Aus $t_n \rightrightarrows f$ folgt auch $|t_n| \rightrightarrows |f|$ wie unter (b) schon gezeigt, daher ist

$$\begin{aligned} |I(f)| = |\lim_{n\to\infty} I(t_n)| &= \lim_{n\to\infty} |I(t_n)| \\ &\leq \lim_{n\to\infty} I(|t_n|) = I(|f|). \end{aligned}$$

□

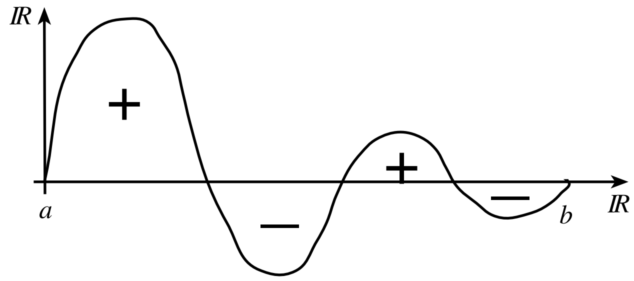
17.2.8 Geometrische Interpretation

Wie bei Treppenfunktionen (vgl. §17.1.6) hat man die folgende geometrische Inteerpretation:
Sei $f : [a, b] \to \mathbb{R}$ eine Regelfunktion, etwa f stetig und $f(X) \geq 0$ für alle $x \in [a, b]$ und $O(f) := \{(x, y) \in \mathbb{R} \times \mathbb{R};\ x \in (a, b);\ o \leq y \leq f(x)\} \subset \mathbb{R} \times \mathbb{R}$ ihre Ordinatenmenge. Dann ist $I(f) \geq 0$ und $I(f) = \int\limits_a^b f(x)dx$ kann man dann als Maß für den Flächeninhalt von $O(f)$ *definieren*.



(ein Flächeninhalt für solche (spezielle) Teilmengen von $\mathbb{R}^2$ ist bisjetzt nihct definiert!).

Nimmt f auch negative Werte an, so liefert $I(f) = \int\limits_a^b f$ ein Maß für den „orientierten Flächeninhalt“ (d.h. Flächeninhalt mitt Vorzeichen) der Menge, deri zwischen dem Graphen von f und der x-Achse liegt:



Indem wir gleichmäßige Limites von Folgen von Treppenfunktionen betrachtet haben, ist es und gelungen, den Raum der Treppenfunktionen zu erweitern zum Raum der Regelfunktionen. Man könnte versuchen, diesen Prozess nochmals zu wiederholen, d.h. gleichmäßige Limites von Folgen von Regelfunktionen zu betrachten und man könnte erwarten, nochmals eine größere Klasse von Funktionen zu erhalten. Der folgende Satz besagt aber, dass der Raum der Regelfunktionen *stabil* gegenüber der Bildung von gleichmäßigen Limites von Regelfunktionen ist (hieraus folgt auch, dass $R(M)$ ein Banach-Raum beütglich der Supremumsnorm ist) und beinhaltet gleichzeitig eine wichtige *Vertauschungseigenschaft des Regelintegrals*.

17.2.9 Satz (Stabilitätssatz, Vertauschungssatz)

(a) Seien $f_n \in R(M)$ für $n \in \mathbb{N}$ und es gelte $f_n \rightrightarrows f : [a, b] \to \mathbb{R}$, dann gilt $f \in R(M)$ und es ist $I(f) = \lim\limits_{n\to\infty} I(f_n)$, oder vielleicht suggestiver

$$\int\limits_a^b f(x)dx = \int\limits_a^b \lim_{n\to\infty} f_n(x)dx = \lim_{n\to\infty} \int\limits_a^b f(x)dx.$$

D.h. der gleichmäßige limes einer Folge von Regelfunktionen ist wieder eine Regelfunktion und man darf zur Berechnung Integration und Grenzwertbildung vertauschen.

Da die Integration ja auch durch ein Grenzprozess definiert ist, handelt es sich dabei um die Vertauschung zweier Grenzprozesse.

(b) Ist (f_k) eine Folge von Regelfunktionen $f_k \in R(M)$, so dass die Funktionsreihe $\sum\limits_{k=0}^{\infty} f_k$ (d.h. die Folge $F_n := \sum\limits_{k=0}^{n} f_k$ der Partialsummen) gleichmäßig auf M gegen $F : [a,b] \to \mathbb{R}$ konvergiert, dann ist F auch eine Regelfunktion und man „darf Summation und Integration vertauschen".

$$\int_a^b F(x)dx = \int_a^b \left(\sum_{k=0}^{\infty} f_k(x) \right) dx = \sum_{k=0}^{\infty} \int_a^b f_k(x)dx.$$

Beweis (a): Wegen $f_n \in R(M)$ für $n \in \mathbb{N}$, kann man für jedes n ein $t_n \in T(M)$ mit $\|t_n - f_n\| < \frac{1}{n}$ wählen. Aus

$$\begin{aligned} \|t_n - f\| &= \|t_n - f_n + f_n - f\| \\ &\leq \|t_n - f_n\| + \|f_n - f\| \\ &\leq \frac{1}{n} + \|f_n - f\| \end{aligned}$$

und $f_n \rightrightarrows f$ folgt daher $\lim\limits_{n\to\infty} \|t_n - f\| = 0$, also $t_n \rightrightarrows f$, also $f \in R(M)$.
Nach uns jetzt schon vertrauter Schlussweise folgt nun

$$|I(f_n) - I(f)| = |I(f_n - f)| \leq (b-a)\|f_n - f\|, \quad \text{also}$$

$$\lim_{n\to\infty} I(f_n) = I(f) \quad \text{wegen} \quad \lim_{n\to\infty} \|f_n - f\| = 0$$

□

Beweis (b): Man braucht (a) nur auf die Partialsummen $F_n := \sum\limits_{k=0}^{n} f_k$ anzuwenden und die Vertauschbarkeit der Indizes mit endlichen Summen beachten, via Induktion folgt aus der Additivität:

$$\int_a^b F_n(x)dx = \int_a^b \left(\sum_{k=0}^{n} f_k(x)dx \right) = \sum_{k=0}^{n} \int_a^b f_k(x)dx.$$

Der Grenzübergang $n \to \infty$ liefert dann nach (a)

$$\begin{aligned} \int_a^b \left(\sum_{k=0}^{\infty} f_k(x) \right) dx &= \int_a^b F(x)dx \\ &= \lim_{n\to\infty} \int_a^b F_n(x)dx \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} \int_a^b f_k(x)dx = \sum_{k=0}^{\infty} \int_a^b f_k(x)dx. \end{aligned}$$

□

Gegenbeispiel:
Bei nur punktweisen Konvergenz ist der Vertauschungssatz i.A. falsch:
Betrachte $t_n : [0,1] \to \mathbb{R}$ mit

$$t_n(x) = \begin{cases} 0, & \text{für } x = 0; \\ n, & \text{für } 0 < x < \frac{1}{n}; \\ 0, & \text{für } \frac{1}{n} < x \leq 1. \end{cases}$$

Die Folge (t_n) konvergiert punktweise gegen die Nullfunktion, es gibt aber $I(t_n) = \int\limits_0^1 t_n(x)dx = 1$ und damit auch $\lim\limits_{n\to\infty} I(t_n) = 1 \neq 0 = \int\limits_0^1 \lim\limits_{n\to\infty} t_n(x)dx$.

Wie umfangreich die Klasse der Regelfunktionen ist zeigt

17.2.10 Theorem

Jede stetige Funktion bzw. jede monotone Funktion $f : M \to \mathbb{R}$ ist eine Regelfunktion.

Der Beweis für *stetige* Funktionen beruht auf der Tatsache, dass eine stetige Funktion $f : K \to \mathbb{K}$, wobei $K \subset \mathbb{K}$ kompakt ist, auf K *gleichmäßig stetig* ist.

17.2.11 Definition

Eine Funktion $f : D \to \mathbb{K}$ $(D \subset K,\ D \neq \emptyset)$ heißt *gleichmäßig stetig auf* D, wenn folgendes gilt:
Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für *alle* $x, y \in D$ mit $|x - y| < \delta$ gilt $|f(x) - f(y)| < \varepsilon$.

Bemerkung:
Wählt man einen festen Punkt $a \in D$, dann sieht man sofort, dass eine gleichmäßige stetige Funktion in jedem Punkt $a \in D$ stetig ist. Der Unterschied zur gewöhnlichen Stetigkeit besteht darin, dass das zu jedem $\varepsilon > 0$ bei der gewöhnlichen Stetigkeitsdefinition existierende $\delta > 0$ neben seiner Abhängigkeit von ε i.A. von der betrachtenden Stelle a abhängt (man vergleiche dazu die Beispiele zur $\varepsilon - \delta-$Definition der Stetigkeit in §12).
Manchmal lässt sich jedoch ein universelles δ wählen, das nur von ε und nicht von der betrachtenden Stelle $a \in D$ abhängt, z.B. bei

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x \end{aligned} \qquad \text{oder} \qquad \begin{aligned} g : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto |x| \end{aligned}$$

Hier ist die Wahl $\delta = \varepsilon$ möglich. Ist jedoch der Definitionsbereich einer stetigen Funktion kompakt (vgl. ???), dann ist f gleichmäßig stetig.

17.2.12 Satz

Ist $K \subset \mathbb{K}$ kompakt($\neq \emptyset$) und $f : K \to \mathbb{K}$ stetig auf K.
Da ein Intervall $[a, b] \to \mathbb{R}$ kompakt ist, ist jede stetige Funktion $f : [a, b] \to \mathbb{R}$ gleichmäßig stetig auf $[a, b]$.

Beweis : Wir führen den Beweis für eine beliebige kompakte Teilmenge $K \subset \mathbb{R}$ oder $K \subset \mathbb{C}$, da beweistechnisch kein Unterschied besteht.
Wir führen einen indirekten Beweis, nehmen also an, dass f stetig auf K aber nicht gleichmäßig

stetig ist. Dann gibt es eine Ausnahme ε, nennen wir es ε_0, so dass kein geeignetes δ mit der genannten Eigenschaft existiert, das heißt für *jedes* $\delta > 0$ gibt es Paare $(x_\delta, y_\delta) \in K \times K$ für die zwar $|x_\delta - y_\delta| < \delta$ aber *nicht* $|f(x_\delta) - f(y_\delta)| < \varepsilon_0$ gilt, das heißt, es gilt $|f(x_\delta) - f(y_\delta)| \geq \varepsilon_0$. Wir setzen $\delta = 1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$ und bezeichnen die entsprechenden Punkte $x_{\delta_n} \in K$ und $y_{\delta_n} \in K$ einfach mit x_n und y_n. Für diese Punkte gilt also

$$|x_n - y_n| < \frac{1}{n}, \text{ aber } |f(x_n) - f(y_n)| \geq \varepsilon_0.$$

Da K kompakt ist, also insbesondere beschränkt ist, ist die Folge (x_n) beschränkt, besitzt also nach Bolzano-Weierstrass eine konvergente Teilfolge (x_{n_k}), die wegen der Kompaktheit (an der Stelle geht die Abgeschlossenheit von K ein) gegen einen Punkt $\xi \in K$ konvergiert:

$$\xi := \lim_{k\to\infty} x_{n_k}.$$

Wegen $|x_n - y_n| < \frac{1}{n}$ gilt dann auch $\lim\limits_{k\to\infty} y_{n_k} = \xi$.
Wegen der Stetigkeit von f in ξ folgt daher

$$\lim_{k\to\infty} f(x_{n_k}) = f(\xi) = \lim_{k\to\infty} f(y_{n_k}),$$

oder auch $\lim\limits_{k\to\infty} |f(x_{n_k}) - f(y_{n_k})| = 0$ im Widerspruch zu $|f(x_n) - f(y_n)| \geq \varepsilon_0$ für alle $n \in \mathbb{N}$.

Dieser Widerspruch zeigt, dass unsere Annahme falsch war, f also doch gleichmäßig stetig auf K ist.

□

Wir kommen nun zum Beweis von Theorem 17.2.10 und zeigen zunächst, dass jede stetige Funktion

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

eine Treppenfunktion ist.
Die Idee ist, das Intervall $[a, b]$ hinreichend fein zu unterteilen und eine geeignete Folge von Treppenfunktionen (t_n) mit $t_n \rightrightarrows f$ zu konstruieren. Dazu wählen wir eine Folge $(Z^{(n)})$ von äquidistanten Zerlegungen

$$Z^{(n)} = \{a = x_0^{(n)} < x_1^{(n)} < \ldots < x_{k-1}^{(n)} < x_k^{(n)} < \ldots < x_n^{(n)}\}$$

mit $x_k^{(n)} = a + k \cdot \frac{b-a}{n},\ k = 0, 1, \ldots, n$ und definieren $t_n : [a, b] \to \mathbb{R}$ durch

$$t_n(x) = \begin{cases} f(x_{k-1}), & \text{für } x \in \left[x_{k-1}^{(n)}, x_k^{(n)}\right[,\ 1 \leq k \leq n \\ f(b), & \text{für } x = x_n^{(n)} = b \end{cases}$$

Für diese speziell ausgewählte Folge (t_n) von Treppenfunktionen zeigen wir, dass $t_n \rightrightarrows f$ gilt.
Dazu benutzen wir die *gleichmäßige Stetigkeit* von f auf $[a, b]$:
Zu jedem $\varepsilon > 0$ gibt es daher ein Universal-$\delta > 0$, so dass für alle $x, y \in [a, b]$ mit $|x - y| < \delta$ gilt $|f(x) - f(y)| < \varepsilon$.
Zunächst wählen wir eine natürliche Zahl N so groß, dass $\frac{b-a}{N} < \delta$ gilt.
Für $n \geq N$ betrachten wir $|f(x) - t_n(x)|$ für $x \in [a, b]$. Jedes $x \in [a, b],\ x \neq b$, liegt in genau einem der Intervalle $\left[x_{k-1}^{(n)},\ x_k^{(n)}\right[$, $k = 1, \ldots, n$. Sei also $x \in \left[x_{k-1}^{(n)},\ x_k^{(n)}\right[$, dann ist nach Definition von t_n

$$f(x) - t_n(x) = f(x) - f\left(x_{k-1}^{(n)}\right)$$

und weiter

$$\left|x - x_{k-1}^{(n)}\right| \leq \left|x_k^{(n)} - x_{k-1}^{(n)}\right| = x_k^{(n)} - x_{k-1}^{(n)} \leq \frac{b-a}{n} < \frac{b-a}{N} < \delta$$

Aus $\left|x - x_{k-1}^{(n)}\right| < \delta$ folgt dann wegen der gleichmäßigen Stetigkeit von f

$$|f(x) - t_n(x)| = \left|f(x) - f(x_{k-1}^{(n)})\right| < \varepsilon$$

Da für $x = b$ nach der Definition

$$|f(b) - t_n(b)| = |f(b) - f(b)| = 0$$

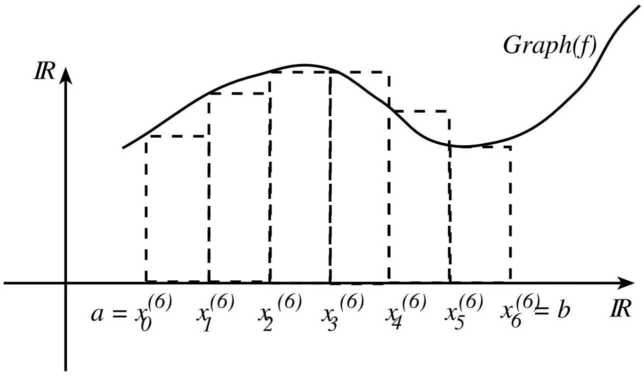
gilt, ist also für alle $n \geq N$ und alle $x \in [a, b]$

$$|f(x) - t_n(x)| < \varepsilon$$

und damit

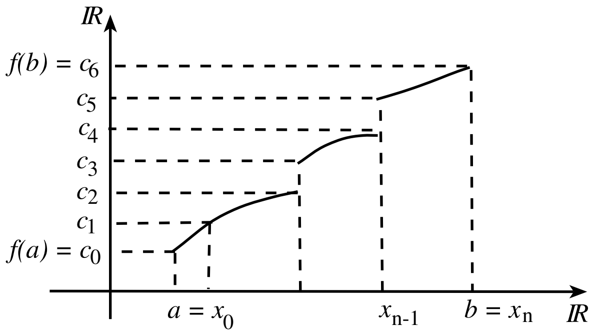
$$\|f - t_n\| \leq \varepsilon$$

für alle $n \geq N$, d.h. $t_n \rightrightarrows f$.
Damit haben wir also nach einem ganz speziellen Verfahren (äquidistante Festlegung des Funktionswertes) eine Folge (t_n) von Treppenfunktionen $t_n : [a, b] \to \mathbb{R}$ konstruiert, die gleichmäßig gegen die gegebene stetige Funktion f konvergiert. Damit ist f also eine Regelfunktion (Die Abbildung veranschaulicht den Fall $n = 6$.):



Wir kehren zum Beweis vom Theorem 17.2.10 zurück und zeigen jetzt, dass jede *monotone* Funktion $f : [a, b] \to \mathbb{R}$ eine Regelfunktion ist.

OBdA sei f monoton wachsend (sonst gehe man von f zu $-f$ über)



Die Bildmenge $f([a, b])$ ist dann im Intervall $[f(a), f(b)]$ enthalten. Wir unterteilen $[f(a), f(b)]$ äquidistant, etwa

$$f(a) = c_0 < c_1 < c_2 < \ldots < c_{\nu-1} < c_\nu < \ldots < c_n = f(b)$$

mit $c_\nu := f(a) + \frac{\nu(f(b)-f(a))}{n}$ für $0 \leq \nu \leq n$. Außerdem setzen wir

$$B_\nu = [c_{\nu-1}, c_\nu] \text{ für } 1 \leq \nu \leq n \text{ und}$$

$$A_\nu = f^{-1}(B_\nu) := \{x \in [a, b];\ c_{\nu-1} \leq f(x) \leq c_\nu\}.$$

Aus der Monotonie von f folgt, dass A_ν ein (möglicherweise leeres) Intervall ist. Wir interessieren uns nur für die „echten" Intervalle, also diejenigen, die mehr als einen Punkt enthalten. Wir nummerieren die Randpunkte dieser Intervalle durch und erhalten in

$$Z := \{a = x_0 < x_1 < \ldots < x_{j-1} < x_j < \ldots < x_n := b\}$$

eine Zerlegung von $[a, b]$.
Wir wählen nun noch irgendeinen Zwischenpunkt $\xi_j \in]x_{j-1}, x_j[$ mit $1 \leq j \leq k$. (z.B. kann man $\xi_j = \frac{x_{j-1}+x_j}{2}$, also der Mittelpunkt des Intervalls wählen) und definieren nun

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

durch

$$t_n(x) = \begin{cases} f(\xi_j), & \text{für } x \in]x_{j-1}, x_j[\ (1 \leq j \leq k); \\ f(x_j), & \text{für } x = x_j, \ (j = 0, \ldots, k). \end{cases}$$

Dann ist t_n eine Treppenfunktion bezüglich der Zerlegung Z.
Wir schätzen jetzt $f(x) - t_n(x)$ ab. Nach Konstruktion ist das Bild irgendeines der Intervalle $]x_{j-1}, x_j[$ enthalten in einem Intervall der Länge $\frac{f(b)-f(a)}{n}$. Hieraus und aus $t_n(x_j) = f(x_j)$ folgt dann für alle $x \in [a, b]$

$$|f(x) - t_n(x)| \leq \frac{f(b) - f(a)}{n},$$

und damit

$$\lim_{n\to\infty} \|f - t_n\| = 0.$$

Damit ist gezeigt, dass die Folge (t_n) gleichmäßig gegen f konvergiert. Damit ist f eine Regelfunktion.

□

Mit der stetigen Funktion und der monotonen Funktion haben wir einen großen Vorrat von Regelfunktionen.
Da die Funktionen, für die wir ein Integral erklärt haben genau die Regelfunktionen $f : [a, b] \to \mathbb{R}$ sind, nennen wir die Regelfunktionen auch die *integrierbare Funktionen*:

$$\boxed{\text{Integrierbare Funktion = Regelfunktion}}$$

Bei Zugrundelegung anderer Integralbegriffe erhält man andere Funktionsklassen.

Riemann-Integral $\Longleftrightarrow$ Riemann-integrierbare Funktion

Lebesque-Integral $\Longleftrightarrow$ Lebesque-integrierbare Funktion

Obwohl wir häufig mit äquidistanten Zerlegungen gearbeitet haben, erhält man eine größere Flexibilität bei der konkreten Integralberechnung, wenn man auch andere Zerlegungen (oder Folgen von Zerlegungen) betrachtet, die nur die Eigenschaft haben müssen, dass sie „fein genug" sind. Diesen Begriff werden wir gleich präzisieren.
Außerdem wollen wir auch eine größere Flexibilität bei der Auswahl der Zwischenpunkte ξ_j im j-ten Teilungintervall. Wir führen dazu die folgende Sprechweise ein.

17.2.13 Definition

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

$$Z^{(n)} := \{a = x_0^{(n)} < x_1^{(n)} < \ldots < x_{k-1}^{(n)} < x_k^{(n)} < \ldots < x_{r_n}^{(n)} = b\}$$

eine (zunächst feste Zerlegung von $[a, b]$ ($n \in \mathbb{N}$ fest). Wählt man für jedes k mit $1 \leq k \leq r_n$ einen Punkt $\xi_k^{(n)} \in [x_{k-1}, x_k],\ 1 \leq k \leq r_n$, dann heißt der Vektor

$$\xi_k^{(n)} = (\xi_1^{(n)}, \ldots, \xi_{r_n}^{(n)})$$

ein „Zwischenvektor“ und die Summe

$$S(f; Z^{(n)}, \xi^{(n)}) = \sum_{k=0}^{r_n} f(\xi_k^{(n)})(x_k^{(n)} - x_k^{(n-1)})$$

eine *Riemannsche Summe* für f zur Zerlegung $Z^{(n)}$ und zum Zwischenvektor $\xi^{(n)}$.
Die maximale Länge der Teilintervalle

$$\eta(Z^{(n)}) = \max\{x_k^{(n)} - x_{k-1}^{(n)};\ 1 \leq k \leq r_n\}$$

heißt die *Feinheit*(Feinheitsmaß, Feinheitsgrad) der Zerlegung $Z^{(n)}$.
Beachte: $S(f, Z^{(n)}, \xi^{(n)})$ ist das Interval einer geeigneten Treppenfunktion (welcher?) und jedes Integral einer Treppenfunktion ist eine spezielle Riemannsche Summe.

17.2.14 Satz (Berechnung von Integralen mit der Riemannschen Summe)

Ist $f : [a, b] \to \mathbb{R}$ eine Regelfunktion und $Z^{(n)} := \{a = x_0^{(n)} < x_1^{(n)} < \ldots < x_k^{(n)} < \ldots < x_{r_n}^{(n)} = b\}$ irgendeine Folge von Zerlegungen mit $\lim\limits_{n\to\infty} \eta(Z^{(n)}) = 0$ („Zerlegungsnullfolge“), dann konvergiert für *jede* Folge von Zwischenvektoren $(\xi^{(n)})$, $\xi^{(n)} = (\xi_1^{(n)}, \ldots, \xi_{r_n}^{(n)})$ mit $\xi_k^{(n)} \in [x_{k-1}^{(n)}, x_k^{(n)}];\ 1 \leq k \leq r_n$ die Folge der Riemannschen Summen $\big(S(f, Z^{(n)}, (\xi^{(n)})\big)_{n\geq 1}$ gegen das Integral von f:

$$\lim_{n\to\infty} S(f, Z^{(n)}, (\xi^{(n)})) = \lim_{n\to\infty} \sum_{k=0}^{r_n} f(\xi_k^{(n)})(x_k^{(n)} - x_{k-1}^{(n)}) = I(f) = \int_a^b f.$$

Beweisskizze:
Wir beweisen die Behauptung nur für *stetige* Funktionen $f : [a, b] \to \mathbb{R}$.
Für eine *beliebige* Regelfunktion vergleiche man die Musterlösung zu Aufgabe ?? von Blatt 2.
Man imitiere den Beweis vom Theorem und wähle eine Folge von $\eta(Z^{(n)})$ der Zerlegungen, so dass $\eta(Z^{(n)}) < \delta$ für $n \geq N$ gilt (dieses δ ist das Universaldelta, was es auf Grund der gleichmäßigen Stetigkeit von f zu einem vorgegebenem $\varepsilon > 0$ gibt).
Definiert man dann t_n $(n \geq N)$ in Analogie zum Beweis von 17.2.10 für stetige Funktionen, wobei man die Zwischenpunkte $\xi_k^{(n)} \in [x_{k-1}^{(n)}, x_k^{(n)}]$ beliebig wählen kann, dann gilt $t_n \rightrightarrows f$ und

$$S(f, Z^{(n)}, (\xi^{(n)})) = I(t_n)$$

und damit

$$\lim_{n\to\infty} S(f, Z^{(n)}, (\xi^{(n)})) = \lim_{n\to\infty} I(t_n) = I(f).$$

17.2.15 Beispiele und Bemerkungen

(a) Häufig verwendet man eine Folge $Z^{(n)}$ von äquidistanten Zerlegungen von $[a, b]$, hier gilt also $r_n = n$ für alle $n \in \mathbb{N}$ und $x_k = a + k\frac{b-a}{n},\ h = 0, \ldots, n$.

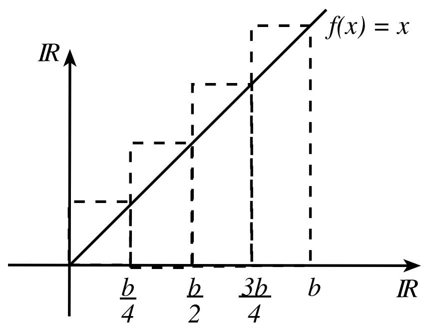
Als erstes Beispiel hatten wir in §17.2.5 das Integral $\int\limits_0^1 x\ dx$ berechnet und den Wert $\frac{1}{2}$ erhalten (was zu erwarten war). Wir wollen etwas allgemeiner das Integral $\int\limits_0^b x\ dx$ berechnen und

verwenden dazu Riemansche Summen.
Wir wählen für jedes $n \in \mathbb{N}$ äquidistante Zerlegungen

$$Z^{(n)} := \left\{0, \frac{b}{n}, \frac{2b}{n}, \dots, b\right\}$$

sowie die Zwischenpunkte (Stützpunkte): $\xi_k^{(n)} = x_k^{(n)} = k\frac{b}{n}$ für $b = 1, \dots, n$. (als Zwischenpunkte haben wir also die rechten Randpunkte des Intervall $[x_{k-1}^{(n)}, x_k^{(n)}]$ gewählt). Es gilt also

$$S_n := S(f;\ Z^{(n)}, \xi^{(n)}) = \sum_{k=1}^{n} \left(\frac{kb}{n}\right) \frac{b}{n} = \frac{b^2}{n^2} \sum_{k=1}^{n} k = \frac{b^2}{n^2} \frac{n(n+1)}{2}$$



Also ist (beachte $\lim\limits_{n\to\infty} \eta(Z^{(n)}) = \lim\limits_{n\to\infty} \frac{b-a}{n} = 0$)

$$\lim_{n\to\infty} S_n = \int_a^b f(x)dx = \int_0^b x\ dx = \frac{b^2}{2}.$$

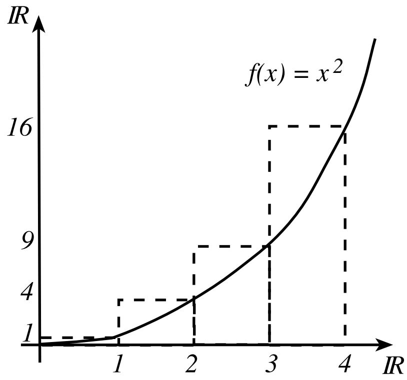
$\square$

(b) Mit der Summenformel (vgl. §2.4(1))

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

findet man völlig analog

$$\int_0^b x^2 dx = \frac{b^3}{3}.$$



(c) Wir wollen allgemeiner das Integral $\int\limits_a^b p_\alpha(x)dx$ mit $p_\alpha(x) = x^\alpha,\ \alpha \in \mathbb{R},\ \alpha \neq -1$ und $0 < a < b$ berechnen.
Wir wählen dazu schon von Fermat benutzte „geometrische Zerlegung “ von $[a, b]$.
Für $n \in \mathbb{N}$ sei

$$Z_q^{(n)} := \{a = x_0^{(n)} < x_1^{(n)} = aq < x_2^{(n)} = aq^2 < \ldots < x_k^{(n)} = aq^k < \ldots < x_n^{(n)} = aq^n = b\}$$

mit $q := \sqrt[n]{\frac{b}{a}}$.

$\xi_k^{(n)} = x_k^{(n)} = aq^{k-1}$ für $k = 1, \ldots, n$.
Es ist $\eta(Z_q^{(n)}) = \max\{qq^{k-1}(q-1);\ k = 1, \ldots, n\} \leq b(q-1)$ und wegen $\lim\limits_{n\to\infty} \sqrt[n]{\frac{b}{a}} = 1$ gilt also $\lim\limits_{n\to\infty} \eta(Z_g^{(n)}) = 0,\ (Z_g^{(n)})$ ist also eine Zerlegungsnullfolge.

Es ist dann

$$\begin{aligned}
S_n := S(p_\alpha, Z_g^{(n)}, \eta^{(n)}) &= \sum_{k=1}^{n} p_\alpha(\eta_k^{(n)})(x_k^{(n)} - x_{k-1}^{(n)}) \\
&= a \cdot \sum_{k=1}^{n} (aq^{k-1})\alpha(q^k - q^{k-1}) \\
&= a^{\alpha+1}(q-1) \sum_{k=1} q^{(k-1)(\alpha+1)} \\
&= a^{\alpha+1} \frac{q^{(\alpha+1)n} - 1}{q^{n+1} - 1} \quad \text{(Summenformel für die endliche geometrische Reihe)} \\
&= (b^{\alpha+1} - a^{\alpha+1}) \frac{q-1}{q^{n+1} - 1}.
\end{aligned}$$

Nun gilt aber für $q := q(n) := \sqrt[n]{\frac{b}{a}} > 1$

$$\lim_{n\to\infty} q(n) = 1$$

und mit der Substitution $q := \exp(t)$ erhält man mit $\beta = \alpha + 1$

$$\begin{aligned}
\lim_{q\to 1} \frac{q^\beta - 1}{q-1} &= \lim_{t\to 0} \frac{\exp(\beta t) - 1}{\exp t(t) - 1} \\
&= \lim_{t\in\infty} \beta \frac{\exp(\beta t) - 1}{\beta t} \cdot \frac{t}{\exp(t) - 1} \\
&= \beta \lim_{t\to 0} \frac{\exp(\beta t) - 1}{\beta t} \lim_{t\to 0} \left(\frac{\exp(t) - 1}{t}\right)^{-1} \\
&= \beta \cdot 1 \cdot 1 = \beta = \alpha + 1.
\end{aligned}$$

(Hierbei wurde $\lim\limits_{y\to 0} \frac{\exp(y)-1}{y} = 1$ verwendet).

Daher ist $\lim\limits_{n\to\infty} S_n = \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha+1}$, also

$$\boxed{\int_a^b x^\alpha dx = \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha+1}}.$$

Mit dem Hauptsatz der Differenzial-und Integralrechnung (vgl. §20) ist ist dieses doch etwas mühselig gewonnene Resultat trivial.

(d) Als weiteres Beispiel wollen wir das Integral $\int_a^b \frac{1}{x}\, dx$ für $0 < a < b$ berechnen.

Wir wählen wieder für jedes $n \in \mathbb{N}$ die geometrische Zerlegung

$$Z_g^{(n)} := \{a = x_0^{(n)} < x_1^{(n)} = aq < \ldots < a_k^{(n)} = aq^k < \ldots < x_n^{(n)} = aq^n = b\}$$

mit $q := \sqrt[n]{\frac{b}{a}}$.

Wir wählen $\eta_k^{(n)} = x_{k-1}^{(n)} = aq^{k-1}$ für $k = 1, \ldots, n$ und erhalten mit $J(x) := \frac{1}{x}$

$$\begin{aligned}
S_n := S(J; Z_g^{(n)}, \xi^{(n)}) &= \sum_{k=1}^n J(\xi_k^{(n)})(x_k^{(n)} - x_{k-1}^{(n)}) \\
&= \sum_{k=1}^n \frac{1}{aq^{k-1}}(aq^k - aq^{k-1}) \\
&= \sum_{k=1}^n (q-1) \\
&= n \cdot (q-1) = n\left(\sqrt[n]{\frac{b}{a}} - 1\right).
\end{aligned}$$

Nach Übungsaufgabe ??? von Blatt ??? gilt aber

$$\lim_{n\to\infty} n\left(\sqrt[n]{\frac{b}{a}} - 1\right) = \log\frac{b}{a}.$$

Dies kann man auch direkt sehen, denn es ist

$$\begin{aligned}
\lim_{n\to\infty} n\left(\sqrt[n]{\frac{b}{a}} - 1\right) &= \lim_{n\to\infty} \frac{\exp\left(\frac{1}{n}\log\frac{b}{a} - 1\right)}{\log\frac{b}{a}} \log\frac{b}{a} \\
&= \log\frac{b}{a} \lim_{t\to 0} \frac{\exp(t) - 1}{t} = \log\frac{b}{a} \quad (\text{mit } t := \tfrac{\log\frac{b}{a}}{n})
\end{aligned}$$

Daher ist

$$\lim_{n\to\infty} S_n = \int_a^b J(x)dx = \int_a^b \frac{1}{x}dx = \log\frac{b}{a} = \log b - \log a$$

(Letzteres nach der Funktionalgleichung des Logarithmus).
Auch hier fällt uns das Ergebnis mit dem Hauptsatz in den Schoss, denn es ist $\log'(x) = \frac{1}{x} = J(x)$, der Logarithmus ist also eine Stammfunktion von J.

(e) Mit Hilfe einer Folge von äquidistanten Zerlegungen von $[a,b]$ lässt sich leicht $\int\limits_a^b \exp(x)dx = \exp(b) - \exp(a)$ zeigen.
Wir wollen das Integral mit Hilfe des Vertauschungssatzes (vgl. §17.2.9), aber einfacher berechnen:
Da die Exponentialreihe das Konvergenzradius ∞ hat, gilt für alle $a, b \in \mathbb{R}$ mit $a < b$

$$\begin{aligned}
\int\limits_a^b \exp(x)dx &= \int\limits_a^b \sum_{k=0}^{\infty} \frac{x^k}{k!} dx \\
&= \sum_{k=0}^{\infty} \int\limits_a^b x^k dx \\
&= \sum_{k=0}^{\infty} \frac{b^{k+1} - a^{k+1}}{(k+1)!} \\
&= 1 + \sum_{k=0}^{\infty} \frac{b^{k+1}}{(k+1)!} - \left(1 + \sum_{k=0}^{\infty} \frac{a^{k+1}}{(k+1)!}\right) \\
&= \exp(b) - \exp(a).
\end{aligned}$$

Genauso zeigt man etwa

$$\int\limits_b^a \cos x \; dx = \sin b - \sin a$$

Wir wollen uns noch mit der Frage beschäftigen, welche Funktionen $f : [a,b] \to \mathbb{R}$ *keine* Regelfunktion sind.
Dazu stellen wir fest:

17.2.16 Bemerkung

Für jede Treppenfunktion $t : M := [a,b] \to \mathbb{R}$ existieren für jeden Punkt x_0 mit $a < x_0 < b$ die einseitigen Grenzwerte

$$f(x_{0_+}) := \lim_{\substack{x \to x_0 \\ x \in M \\ x > x_0}} f(x) \quad \text{(rechtsseitige Grenzwerte)}$$

und

$$f(x_{0_-}) = \lim_{\substack{x \to x_0 \\ x \in M \\ x < x_0}} f(x)$$

und für die Endpunkte a und b existieren

$$f(a_+) = \lim_{\substack{x \to a \\ x > a}} f(x) \text{ und } f(b_-) = \lim_{\substack{x \to b \\ x < b}} f(x).$$

Diese Eigenschaft überträgt sich sofort auch auf einen gleichmäßigen limes vom Treppenfunktionen (Skizzieren Sie einen Beweis!), also Regelfunktionen. Damit haben wir:

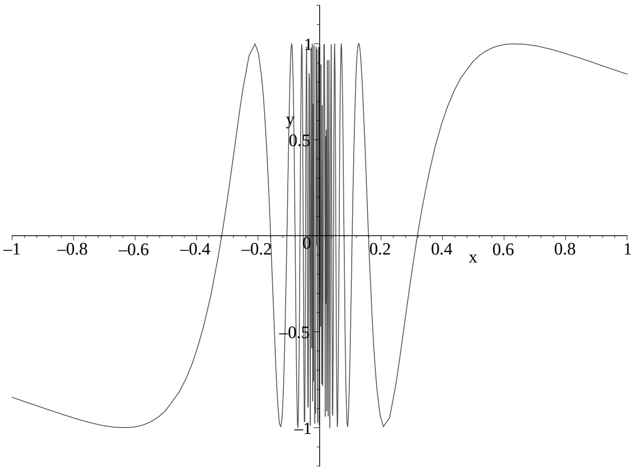


Abbildung 16: Graph der Funktion $\sin\frac{1}{x}$

17.2.17 Satz

Für jede Regelfunktion $f : M \to \mathbb{R}$ existieren für alle $x_0 \in]a,b[$ Grenzwerte $f(x_{0_+})$, $f(x_{0_-})$ und für a bzw. b existieren $f(a_+)$ bzw. $f(b_-)$.

Im Umkehrschluss bedeutet dies:
Wenn für eine Funktion $f : [a,b] \to \mathbb{R}$ einer dieser Grenzwerte *nicht existiert*, dann ist f keine Regelfunktion. So ist also zum Beispiel die Dirichlet- Funktion $\delta : [0,1] \to \mathbb{R}$

$$\delta(x) = \begin{cases} 1, & \text{falls } x \in \mathbb{Q} \\ 0, & \text{sonst} \end{cases}$$

keine Regelfunktion, weil in keinem Punkt $x_0 \in [0,1]$ einer der möglichen Grenzwerte existiert. Mit dem obigen Kriterium erhält man auch, dass die oszilierende Funktionen $f : [0,2\pi] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} \sin\frac{1}{x}, & \text{falls } x \neq 0 \\ 0, & \text{falls } x = 0 \end{cases}$$

keine Regel Funktion ist. (siehe Abb. 16)
Der Grenzwert $f(0_+)$ existiert nicht, denn für die beiden Folgen (x_n) und (y_n) mit $x_n = \frac{2}{(4n+1)\pi}$ bzw. $y_n = \frac{2}{(4n-1)\pi}$ gilt $x_n > 0$ und $y_n > 0$ und $\lim\limits_{n\to\infty} f(x_n) = 1 \neq -1 = \lim\limits_{n\to\infty} f(y_n)$ (beachte: es gilt zum Beispiel schon $f(x_n) = 1$).
Das f keine Regelfunktion sein kann, kann man auch so einsehen:
Ist $Z = \{0 = x_0 < x_1 < \ldots < x_n = 2\pi\}$ eine Zerlegung von $[0,2\pi]$ und t eine Treppenfunktion zu dieser Zerlegung, dann ist t auf dem Interval $]0,x_1[$ konstant. Zu diesem Interval gibt es sowohl Punkte x mit $f(x) = 1$, als auch solche mit $f(x) = -1$. (Man braucht bei den Folge (x_n) und (y_n) n nur hinreichend groß zu wählen.) Daher ist

$$\|f - t\| = \sup\{|f(x) - t(x)|;\ x \in [0,2\pi]\} \geq 1$$

für jede Treppenfunktion $t : [0,2\pi] \to \mathbb{R}$, es kann daher auch keine Folge $(t_n), t_n \in T([0,2\pi])$ mit $t_n \rightrightarrows f$ geben.
f ist auch aus diesem Grund keine Regelfunktion (jedoch Riemann-integrierbar).

Nach Satz ??? gilt auch die Umkehrung, das heißt es gilt:

17.2.18 Charakterisierung von Regelfunktionen durch innere Eigenschaft

Eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ ist genau dann eine Regelfunktion, wenn in jedem Punkt $x_0 \in [a, b]$ die dort möglichen einseitigen Grenzwerte existieren, das heißt es existieren $f(x_{0_+})$ für $a \leq x_0 < b$ und $f(x_{0_-})$ für $a < x_0 \leq b$.

Beweis : Eine Regelfunktion $f : [a, b] \to \mathbb{R}$ ist beschränkt und es existieren die möglichen einseitigen Grenzwerte.
Für die Umkehrung nehmen wir an, dass für $x \in [a, b[$ bzw. $x \in]a, b]$ die Grenzwerte $f(x_+)$ bzw. $f(x_-)$ existieren.
Zu vorgegebenem $\varepsilon > 0$ bilden wir rekursiv

$$x_0 = a, x_1, x_2, \ldots$$

als

$$x_j := sup\{x \in]x, b[;\ |f(x) - f(x_{j-1}\,_+| < \varepsilon\}$$

für $j \geq 1$.

Dann gilt offensichtlich $a = x_0 < x_1 < x_2 < \ldots$. Dieser Prozess muss aber nach endlich vielen Stellen abbrechen, und es gilt dann $x_m = b$ mit einem geeigneten $m \in \mathbb{N}$. Würde der Prozess nämlich nicht abbrechen, dann gebe es einen Punkt $x_* \in]a, b]$, gegen den die Folge (x_j) streng monoton konvergiert. Dann kann aber der einseitige Grenzwert $f(x_{*_-})$ nicht existieren, weil in jedem Interval $]x_* - \delta, x_*[$ mit $\delta > 0$ gelten würde

$$sup\{|f(x) - f(y)|;\ x, y \in]x_* - \delta, x_*[\} \geq \varepsilon.$$

Also gibt es ein $m \in \mathbb{N}$ mit $x_m = b$ und mit Hilfe der Zerlegung

$$Z = \{a = x_0 < x_1 < \ldots < x_m = b\}$$

definieren wir eine Treppenfunktion $t : M = [a, b] \to \mathbb{R}$ durch

$$t(x) := \begin{cases} f(x_{j-1_+}), & \text{für } x_{k-1} \leq x < x_k;\ 1 \leq k \leq m \\ f(b) & \text{für } x = b. \end{cases}$$

Dann folgt $|f(x) - t(x)| < \varepsilon$ für alle $x \in [a, b]$ und damit $\|f - t\| \leq \varepsilon$.
Setzt man nun die Reihe nach $\varepsilon = 1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots$ein, so erhält man eine Folgen (t_n) von Treppenfunktionen $t_n : [a, b] \to \mathbb{R}$ mit $t_n \rightrightarrows f$, also ist f eine Regelfunktion.
Dieser Beweis folgt im Wesentlichen dem von *S.Hildebrandt in S.Hildebrandt: Analysis 1, Springer-Verlag 2002, Seite 330.*

$\square$

Einen weiteren Beweis für die schwierige Beweisrichtung findet man zum Beispiel bei: *K.Königsberger: Analysis 1, Springer-Verlag, 5.Auflage, 2000, S.194.*
Einen weiteren (relativ einfachen) Beweis findet man bei: *Barmer-Floh: Analysis 1*;
Dort wird die „Überdeckungskompaktheit“ des kompakten Intervals benutzt.

17.2.19 Bemerkungen und Beispiele

(a) Aus Satz 17.2.18 folgt nochmals, dass jede *stetige* Funktion $f : [a, b] \to \mathbb{R}$ bzw. jede *monotone* Funktion $f : [a, b] \to \mathbb{R}$ eine Regelfunktion ist. Insbesondere ist auch jede *stückweise stetige* Funktion $f : [a, b] \to \mathbb{R}$ eine Regelfunktion. Dabei heißt f stückweise stetig, wenn es eine Zerlegung

$$Z := \{a = x_0 < x_1 < \ldots < x_{k-1} < x_k < \ldots < x_n = b\}$$

gibt, so dass $f_k := f|\,]x_{k-1}, x_k[$ stetig ist für $1 \leq k \leq n$ und die f_k zu stetigen Funktionen auf $[x_{k-1}, x_k]$ fortgesetzt werden kann.

(b) Nach Satz 17.2.18 ist die Riemann-Funktion (vgl. §11.1.2(b)) $f:[0,1]\to\mathbb{R}$ mit

$$f(x)=\begin{cases} 0, & \text{falls } x \text{ irrational} \\ \frac{1}{q_x}, & \text{falls } x=\frac{p_x}{q_x} \text{ mit } (p_x,q_x)\in\mathbb{N}_0\times\mathbb{N}, q_x \text{ minimal} \end{cases}$$

eine Regelfunktion, weil für *jeden* Punkt $x_0\in[0,1]$ gilt $\lim\limits_{x\to x_0} f(x)=0$. (Vgl. dazu auch die Musterlösung zur Aufgabe ??? von Blatt ???).

Man kann auch direkt eine ε-Approximation durch eine Treppenfunktion für f angegeben: In der Übungsaufgabe wurde gezeigt, dass es zu vorgegebene $\varepsilon>0$ nur endlich viele rationale Zahlen $r_1,\dots,r_N\in[0,1]$ gibt mit $f(r_j)>\varepsilon$. Definiert man t_ε als die Summe der charakteristischen Funktion der endlich vieler Punkte, das heißt $t_\varepsilon := \sum\limits_{k=1}^{N}\chi_{\{x_k\}}$, dann ist $t_\varepsilon\in T([0,1])$, $I(t_\varepsilon)=0$ und $\|f-t_\varepsilon\|\le\varepsilon$.

Das zeigt, dass f ein Regelfunktion ist und dass gleichzeitig $I(f)=\int\limits_0^1 f(x)dx=0$ gilt.

17.2.20 Definition

Eine Unstetigkeitsstelle x_0 einer beschränkten Funktion $f:[a,b]\to\mathbb{R}$ heiße *Sprungstelle*, wenn $f(x_{0_+})$ im Fall $a\le x_0<b$ und $f(x_{0_-})$ im Falle $a<x_0\le b$ existieren.
Eine Umformulierung von Satz ist dann:

17.2.21 Satz

Für eine beschränkte Funktion $f:[a,b]\to\mathbb{R}$ gilt:
f ist genau dann eine Regelfunktion, wenn f (höchstens) abzählbar viele Sprungstellen als Unstetigkeitsstellen hat. ((Höchstens) Abzählbar soll heißen: endlich oder abzählbar unendlich.)
Ist f eine Regelfunktion, dann hat nach Satz 17.2.18 f höchstens Sprungstellen als Unstetigkeitsstellen. (für eine stetige Funktion ist diese Menge leer).
Ferner kann man sich leicht überlegen, dass es für jedes $n\in\mathbb{N}$ höchste endlich viele Sprungstellen $x\in M$ gibt, für welche eine der Zahlen

$$|f(x_{0_+})-f(x_{0_-}),\ |f(x_{0_+})-f(x_0|,\ |f(x_{0_-})-f(x_0)|$$

nicht kleiner als $\frac{1}{n}$ ist. Daher gibt es (höchstens) abzählbar viele Sprungstellen von f. Hat umgekehrt f nur Sprungstellen als Unstetigkeitsstellen, so existieren die möglichen einseitigen Grenzwerte für alle $x_0\in[a,b]$, f ist also nach 17.2.18 eine Regelfunktion.
In 17.2.18 haben wir eine Intervalladditivität für das Integral für Treppenfunktionen bewiesen. Diese überträgt sich sofort auf das Integral für Regelfunktion. Dazu stellen wir fest: Ist $f:[a,b]\to\mathbb{R}$ eine Regelfunktion und $c\in[a,b]$ mit $a<c<b$, dann ist $f_1:=f|\ [a,b]$ und $f_2:=f|\ [a,b]$ ebenfalls eine Regelfunktion.
Sind umgekehrt $f_1:[a,c]\to\mathbb{R}$ und $f_2:[a,b]\to\mathbb{R}$ Regelfunktion, dann ist auch $f:[a,b]\to\mathbb{R}$ mit

$$f(x):=\begin{cases} f_1(x), & \text{für } x\in[a,c]; \\ f_2(x), & \text{für } x\in]c,b]. \end{cases}$$

eine Regelfunktion. Hieraus ergibt sich dann die Intervalladditivität auch für Regelfunktion.

17.2.22 Satz: Intervalladditivität des Regelintegrals

Ist $f : [a, b] \to \mathbb{R}$ eine Regelfunktion und gilt $a < c < b$, dann ist

$$\boxed{\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^d f(x)dx}$$

Beweis : Ist (t_n) eine Folge von Treppenfunktionen $t_n : [a, b] \to \mathbb{R}$ mit $t_n \rightrightarrows f$, dann gilt auch

$$t_n|| [a, b] \rightrightarrows f| [a, c] \text{ und } t_n| [c, b] \rightrightarrows f| [c, b].$$

Daher ist (weil für Treppenfunktionen die Intervalladditivität schon gezeigt ist)

$$\begin{aligned}
\int_a^b f = \lim_{n\to\infty} \int_a^b t_n &= \lim_{n\to\infty} \left(\int_a^c t_n + \int_c^b t_n \right) \\
&= \lim_{n\to\infty} \int_a^c t_n + \lim_{n\to\infty} \int_c^b t_n \\
&= \int_a^c f + \int_c^b f.
\end{aligned}$$

□

Um das Integral etwas flexibler handhaben zu können, wollen wir uns von der Generalvoraussetzung $a < b$ für das Integrationsintervall $[a, b]$ befreien und definieren $\int_a^a f = 0$ für jede Regelfunktion f und falls $b < a$ gilt

$$\int_a^b f = -\int_b^a f.$$

Mit diesen Festlegungen gilt dann

17.2.23 Satz

Sind $a, b, c \in \mathbb{R}$ und ist $f : [\min\{a, b, c\}, \max\{a, b, c\}] \to \mathbb{R}$ eine Regelfunktion, dann gilt bei beliebiger Lage von a, b, c

$$\boxed{\int_a^b f + \int_b^c f + \int_c^a f = 0}$$

Der einfache Beweis durch Fallunterscheidung ($3! = 6$ Fälle) sei den geneigten Lesern überlassen.

Zum Abschluss gehen wir auf zwei Sätze ein, die uns von Nutzen sein werden:

- Mittelwertsatz der Integralrechnung und
- ein „Verschwindungslemma“.

17.2.24 Theorem (Mittelwertsatz der Integralrechnung)

Sei $a, b \in \mathbb{R}$, $a < b$ und $f : [a,b] \to \mathbb{R}$ stetig und $p : [a,b] \to \mathbb{R}$ eine Regelfunktion mit $p \geq 0$ (das heißt $p(x) \geq 0$ für alle $x \in [a,b]$), dann gibt es ein $\xi \in [a,b]$ mit

$(*)$ $\qquad \int\limits_a^b f(x)p(x)dx = f(\xi)\int\limits_a^b p(x)dx$ (verallgemeinerter Mittelwertsatz).

Im Spezialfall $p(x) = 1$ für alle $x \in [a,b]$ folgt:

$(**)$ $\qquad \int\limits_a^b f(x)dx = f(\xi)(b-a)$ mit $\xi \in [a,b]$ (spezieller Mittelwertsatz)

$\mu := \frac{\int_a^b f(x)dx}{b-a}$ nennt man auch *Integral-Mittelwert* (von f im Intervall $[a,b]$). Der spezielle MWS besagt also, dass der Integralmittelwert ein Funktionswert ist.

Beweis : Sei $m_1 = \min\{f(x);\ x \in [a,b]\}$ und $m_2 = \max\{f(x);\ x \in [a,b]\}$.
Nach ??? gilt $f \cdot p \in R([a,b])$ und aus $m_1 p(x) \leq f(x)p(x) \leq m_2 p(x)$ folgt nach ???

$$m_1 \int\limits_a^b p(x)dx \leq \int\limits_a^b f(x)p(x)dx \leq m_2 \int\limits_a^b p(x)dx.$$

Ist nun $\int\limits_a^b p(x)dx = 0$, dann ist auch

$$\int\limits_a^b f(x)p(x)dx = 0$$

und die Gleichung $(*)$ ist für alle $\xi \in [a,b]$ richtig. Gilt jedoch $\int\limits_a^b p(x)dx \neq 0$, so folgt

$$m_1 \leq \frac{\int_a^b f(x)p(x)dx}{\int_a^b p(x)dx} \leq m_2$$

Der Quotient in der Mitte liegt also im Intervall $[m_1, m_2]$, ist also ein Funktionswert von f, daher gibt es nach dem Zwischenwert Satz ein $\xi \in [a,b]$ mit

$$f(\xi) = \frac{\int_a^b f(x)p(x)dx}{\int_a^b p(x)dx},$$

d.h. es gilt

$$\int\limits_a^b f(x)p(x)dx = f(\xi)\int\limits_a^b p(x)dx.$$

$\square$

17.2.25 Beispiele und Bemerkungen

(a) Für Treppenfunktionen ist der MWS nicht richtig, denn für die Treppenfunktionen

$$t : [0,2] \to \mathbb{R} \text{ mit } t(x) = \begin{cases} 1, & \text{falls } 0 \le x \le 1; \\ 0, & \text{falls } 1 < x \le 2. \end{cases}$$

gilt $\frac{1}{2}\int\limits_0^2 t(x)dx = \frac{1}{2} \neq f(\xi)$ für alle $\xi \in [a,b]$.

(b) Man braucht denn Zwischenwertsatz nur im Fall

$$m_1 < \frac{\int\limits_a^b f(x)p(x)dx}{\int\limits_a^b p(x)dx} < m_2,$$

dann kann man ξ sogar im $]a,b[$ wählen. Gilt etwa $m_1 = \frac{\int\limits_a^b f(x)p(x)dx}{\int\limits_a^b p(x)dx}$, dann gibt es ein $x_1 \in [a,b]$ mit $m_1 = f(x_1)$ und $(*)$ ist mit $\xi = x_1$ erfüllt.

Analog schließt man, falls $\frac{\int\limits_a^b f(x)p(x)dx}{\int\limits_a^b p(x)dx} = m_2$ gilt.

(c) Der verallgemeinern Mittelwertsatz macht die Aussage über *gewichtete Mittelwerte* einer stetigen Funktionen. Es gilt auch im Fall $p \le 0$ (das heißt $p(x) \le 0$ für alle $x \in [a,b]$), jedoch *nicht* für Regelfunktionen mit Vorzeichenwechsel in $[a,b]$.

Gegenbeispiel: $p(x) = x - \frac{1}{2}$ mit $\int\limits_0^1 p(x)dx = \frac{1}{2} - \frac{1}{2} = 0$ und $f(x) = x$ für $x \in [0,1]$.

Hier ist

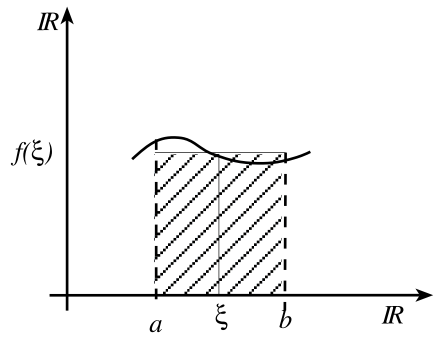
$$\int\limits_0^1 f(x)p(x)dx = \int\limits_0^1 x^2 dx - \frac{1}{2}\int\limits_0^1 x\ dx = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \neq 0,$$

während $f(\xi)\int\limits_0^1 p(x)dx = f(\xi)\cdot 0 = 0$ für alle $\xi \in [0,1]$ gilt.

(d) Der spezieller Mittelwertsatz bedeutet geometrische interpretiert, wegen

$$I(f) = \int\limits_a^b f(x)dx = f(\xi)(b-a),$$

dass das Integral gleicht dem Flächeninhalt eines Rechtecks ist, dessen eine Seitenlänge $(b-a)$ und dessen anderer Seitenlänge ein Funktionswert ist:



Die Mittelwertsätze beruhen auf der Monotonie des Integrals. Die folgende Aussage, die wir häufig benutzen werden, beruht auch wesentlich auf der Monotonie des Integral:

17.2.26 Lemma („Verschwindungslemma“)

Ist $f : [a,b] \to \mathbb{R}$ eine Regelfunktion mit $f \geq 0$ und $\int_a^b f(x)dx = 0$, dann ist $f(x_0) = 0$ an jeder Stetigkeitsstelle $x_0 \in [a,b]$. Insbesondere ist $f(x) = 0$ für fast alle $x \in [a,b]$, d.h. die Menge der $x \in [a,b]$ mit $f(x) \neq 0$ ist höchstens abzählbar.

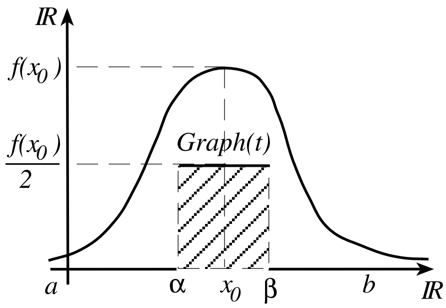
Beweis : Wir nehmen an $x_0 \in [a,b]$ sei eine Stetigkeitsstelle mit $f(x_0) > 0$ und $[\alpha,\beta] \subset [a,b]$ ein Intervall mit $x_0 \in [\alpha,\beta]$ und $f(x) > \frac{1}{2}f(x_0)$ für alle $x \in [\alpha,\beta]$ (ein solches Intervall existiert wegen der Stetigkeit von f in x_0). Dann gilt für die durch

$$t(x) = \begin{cases} \frac{1}{2}f(x_0), & \text{für } x \in [\alpha,\beta]; \\ 0, & \text{für } x \in [a,b] - [\alpha,\beta]; \end{cases}$$

definierte Treppenfunktion

$$\int_a^b f(x)dx \geq \int_a^b t(x)dx = (\beta - \alpha)\frac{f(x_0)}{2} > 0$$

im Widerspruch zur Voraussetzung $\int_a^b f(x)dx = 0$.



Also gilt für jede stetig Stetigkeitsstelle x_0 unter den genannten Voraussetzungen $f(x_0) = 0$. Der Zusatz, dass für fast alle $x \in [a,b]$ auch $f(x) = 0$ gelten muss, folgt aus der Eigenschaft, dass eine Regelfunktion höhchstens abzählbar viele Unstetigkeitsstellen hat (vgl. 17.2.21). Ist f also stetig auf $[a,b]$, $f \geq 0$ und $I(f) = 0$, dann muss $f(x) = 0$ für alle $x \in [a,b]$ gelten, denn hier ist $S_f := \{x \in [a,b];\ x\ \text{ Sprungstelle von }\ f\} = \emptyset$.

$\square$

18 Grundlagen der Differentialrechnung

Der Ausgangspunkt für die Differentialrechnung war bei I.Newton (1643-1727) das Problem der Bestimmung von *Momentangeschwindigkeiten* bzw. *Momentanbeschleunigungen* für ungleichförmigen Bewegungen, bei G.W.Leibniz (1646-1716) stand das Problem der Bestimmung von *Tangenten* an zunächst ebenen Kurven (das heißt Kurven in $\mathbb{R}^2$) im Mittelpunkt des Interesses. Wer untersuchen dieses Problem zunächst für solche Kurven in $\mathbb{R}^2$, die als Graphen von Funktionen auftreten. Wer werde sehen, dass beide Probleme den selben mathematischen Kern haben und darauf hinaus laufen, gegebene Funktionen (hier die differenzierbaren) durch eine einfach Klasse von Funktionen (wenigstens lokal, d.h. in der Umgebung eines festen Punktes) „gut“ zu approximieren. Im ersten Abschnitt beschäftigen wir uns mit dem *Begriff der Ableitung*, dann mit *Ableitungsregeln* (z.B. *Quotientenregel, Kettenregel, Differenzierbarkeit der Umkehrfunktion*). Dann werden wir sehen, wie die Differenzialrechnung geeignete Mittel bereitstellt, um zum Beispiel *lokale Extrema* differenzierbaren Funktionen zu ermitteln. Von großen theoretischen Interesse ist der *Mittelwertsatz der Differenzialrechnung* (abgekürzt MWSD). Wichtig sind die Anwendungen des Mittelwertsatzes: *Monotoniekriterium, Konvexität, Schrankensatz, Charakterisierung konstanter Funktionen* etc.
Wir werden dann auch sehen (siehe den Abschnitt über den Hauptsatz der Differenzial-und Integral Rechnung), dass Differentiation und Integration *Umkehroperationen* voneinander sind.

18.1 Der Begriff der Ableitung, erste Beispiele differenzierbaren Funktionen

18.1.1 Definition: Begriff der Differenzierbarkeit

Sei (zunächst) $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$ eine Funktion.
f heißt *differenzierbar* (ableitbar) im Punkt $a \in D$, falls der Grenzwert

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

existiert, der dann eindeutig bestimmt ist. $f'(a)$ heißt dann die *Ableitung* oder *Differentialquotient* von f in a.

Schreibweise: $f'(a) = \mathrm{D}f(a) = \frac{df}{dx}(a) = \frac{d}{dx}f(a) = \dot{f}(a)$ oder ähnlich.

Die Funktion $\Delta f(a, \cdot) : D - \{a\} \to \mathbb{R}$ mit $\Delta f(a; x) = \frac{f(x)-f(a)}{x-a}$ heißt *Differenzenquotient* von f in a.
Differenzierbarkeit von f in a bedeutet also, dass

$$\lim_{x \to a} \Delta f(a; x) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

existiert.
f heißt *differenzierbar* in D, wenn f in jedem Punkt $x \in D$ differenzierbar ist.
Die Funktion

$$\begin{array}{rccc} f' : & D & \to & \mathbb{R} \\ & x & \mapsto & f'(x) \end{array}$$

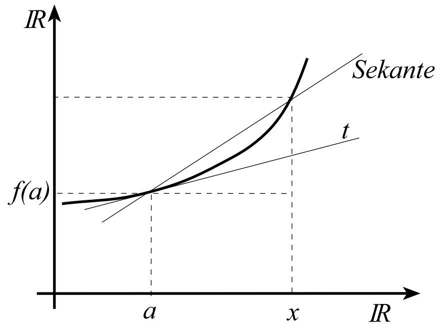
heißt dann Ableitungsfunktion von f.
Schreibweisen: $f' = \mathrm{D}f = \frac{d}{dx}f = \dot{f}$ oder ähnlich.

18.1.2 Geometrische Interpretation

Ist $f : D \to \mathbb{R}$ in $a \in D$ differenzierbar, so kann man den Differenzenquotient

$$\Delta f(a;x) = \frac{f(x)-f(a)}{x-a} \quad (x \neq a)$$

interpretieren als Steigerung der Geraden durch die Punkte $(a, f(a))$ und $(x, f(x))$ („Sekante“)



Ist f differenzierbaren in a, so konvergiert $\Delta f(a;x)$ für $x \to a$ gegen $f'(a)$. Die durch die Gleichung $y = f(a) + f'(a)(x-a)$ für $x \in \mathbb{R}$ definierte Gerade t heißt *Tangente an den Graphen* von f im Punkt $(a, f(a))$.

18.1.3 Physikalische Interpretation

Ist $s(t)$ der Ort eines sich auf einer Geraden bewegten Massenpunktes zur Zeit („tempus“) $t \in \mathbb{R}$, dann ist (für $t \neq t_0$) $\Delta s(t;t_0) = \frac{s(t)-s(t_0)}{t-t_0}$ die *mittlere Geschwindigkeit* (Durchschnittsgeschwindigkeit) des Massenpunktes im Zeit Intervall $[t, t_0]$. Existiert die Ableitung von s in t_0, existiert also

$$\dot{s}(t) := \lim_{t \to t_0} \frac{s(t)-s(t_0)}{t-t_0},$$

so kann man $\dot{s}(t_0)$ als momentane Geschwindigkeit des Massenpunktes zum Zeitpunkt t_0 interpretieren. Auch wenn das Weg-Zeit-Gesetz $t \to s(t)$ nicht durch eine Gerade beschrieben wird, kann man $\dot{s}(t_0)$ als Momentangeschwindigkeit interpretieren. In der Physik wird die schon auf Newton zurückgehende Notation „$\dot{} = \frac{d}{dt}$“ für Ableitung nach der Variablen t (Zeit) verwendet.

18.1.4 Ableitung komplexwertiger Funktionen

Für Funktionen $f : D \to \mathbb{C}$ kann man denn Begriff der Differenzierbarkeit durch den limes des Differenzenquotienten wie in 18.1.1 erklären ($\mathbb{C}$ ist ja ein Körper!). Es gilt: Ist $f = u + iv$, $u = Re\ f$, $v = Im\ f$, dann ist f genau dann in $a \in D$ differenzierbar, wenn die reellwertigen Funktionen u und v in a differenzierbar sind, und es gilt dann

$$f'(a) = u'(a) + iv'(a).$$

18.1.5 Erste Beispiele

(a) Jede konstanter Funktionen $f : D \to \mathbb{C}$ ist differenzierbar, und es gilt $f'(x) = 0$ für alle $z \in D$. Denn ist etwa $f(x) = c$, $c \in \mathbb{C}$ ($c \in \mathbb{R}$ ist natürlich zulässig), dann ist schon ($x \neq c$) $\Delta f(a;x) = \frac{0-0}{x-a} = 0$, also auch

$$f'(a) = \lim_{x \to a} \Delta f(a;x) = 0.$$

(b)

$$\begin{aligned} p_n : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^n \quad (n \in \mathbb{N}) \end{aligned}$$

ist differenzierbar und es gilt

$$p_n'(x) = nx^{n-1}.$$

Für $x \neq a$ ist

$$\Delta p_n(a;x) = \frac{x^n - a^n}{x-a} = x^{n-1} + x^{n-2}a + \ldots + a^{n-1}$$

und damit $\lim\limits_{x\to a} \Delta p_n(a,x) = na^{n-1}$.

(c) Für $\alpha \in \mathbb{C}$, ist

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{C} \\ x &\mapsto \exp(\alpha x) \end{aligned}$$

differenzierbar, und es gilt für alle $x \in \mathbb{R}$

$$f'(x) = \alpha \exp(\alpha x) = \alpha f(x).$$

Es ist nämlich für $x \neq a$

$$\frac{\exp(\alpha x) - \exp(\alpha a)}{x-a} = \exp(\alpha a)\frac{\exp(\alpha(x+a)) - 1}{x-a} = \alpha\exp(\alpha a)\frac{\exp(\alpha(x-a)) - 1}{\alpha(x-a)}$$

und daher

$$\lim_{x\to a} \frac{\exp(\alpha x) - \exp(\alpha a)}{x-a} = \alpha\exp(\alpha a).$$

Insbesondere ist für $\alpha = 1$

$$exp'(x) = exp(x)$$

für alle $x \in \mathbb{R}$ und für $\alpha = i$ gilt mit $f(x) = \exp(ix)$

$$f'(x) = i\exp(ix) = if(x).$$

(d) Die Betragsfunktion $f = abs = |\ | : \mathbb{R} \to \mathbb{R}$ ist in $a = 0$ *nicht* differenzierbar, denn es ist für $x \neq 0$

$$\Delta f(0,x) = \frac{|x| - 0}{x - 0} = \begin{cases} 1, & \text{falls } x > 0; \\ -1, & \text{falls } x < 0; \end{cases}$$

und dieser Differenzenquotient hat keinen Grenzwert für $x \to 0$ (für die Folge $x_n = \frac{1}{n}$ gilt $\lim\limits_{n\to\infty} \Delta f(0;x_n) = 1 \neq -1 = \lim\limits_{n\to\infty} \Delta f(0,y_n)$, $y_n = -\frac{1}{n}$).

Betrachtet man jedoch die Einschränkung g von $f = abs$ auf $\mathbb{R}_+ = \{x \in \mathbb{R}, x \geq 0\}$, dann existiert

$$g'(0) = \lim_{x\to 0} \frac{g(x) - g(0)}{x - 0} = \lim_{x\to 0} \frac{x}{x} = 1.$$

Man definierte $f'_+(0) := g'(0)$ und nennt $f'_+(0)$ die rechtsseitige Ableitung von f im Nullpunkt. Entsprechend ist die linksseitige Ableitung von f im Nullpunkt -1.

18.1.6 Feststellung

Dass der Grenzwert $f'(a) := \lim\limits_{x\to a} \frac{f(x)-f(a)}{x-a}$ in 18.1.1 existiert ist nach der Definition des Grenzwertes äquivalent mit der Fortsetzbarkeit des Differentialquotienten $\Delta f(a;\ ) : D \to \{a\} \to \mathbb{R}$ einer auf ganz D definierte Funktion, die in a stetig ist. Nennen wir die in a stetige Fortsetzung der Einfachheit halber φ (sie hängt von a und f ab), damit ist also $\varphi : D \to \mathbb{R}$ mit

$$\varphi(x) = \begin{cases} \frac{f(x)-f(a)}{x-a}, & \text{für } x \neq a \\ f'(a), & \text{für } x = a \end{cases}$$

stetig in a.
Für $x \neq a$ kann man $f(x) - f(a) = (x-a)\varphi(x)$ schreiben. Diese Gleichung gilt auch für $x = a$. Damit haben wir:
$f : D \to \mathbb{R}$ ist genau dann differenzierbar eben $a \in D$, wenn es eine Funktion $\varphi : D \to \mathbb{R}$ gibt, die in a stetig ist und für die

$(*)$ $\qquad f(x) = f(a) + (x-a)\varphi(x)$

für alle $x \in D$ und $\varphi(a) = f'(a)$ gilt. Setzt man $\varrho(x) := \varphi(x) - f'(x)$ für $x \in D$ und setzt diese in $(*)$ ein, so erhält man

$(**)$ $\qquad f(x) = f(a) + (x-a)f'(a) + (x-a)\varrho(x),$

dabei ist $g : D \to \mathbb{R}$ stetig in a und $g(a) = 0$. Definiert man $r(x) := (x-a)\varrho(x)$ für $x \in D$, dann ist $(**)$ äquivalent zu

$(***)$ $\qquad f(x) = f(a) + (x-a)f'(a) + r(x)$

mit

$$\lim_{x\to a} \frac{r(x)}{x-a} = \lim_{x\to a} g(x) = g(a) = 0.$$

Damit haben wir die meisten Teile des folgenden Äquivalenssatzes für Differenzierbarkeit:

18.1.7 Theorem (Äquivalentssatz für Differenzierbarkeit)

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f : D \to \mathbb{R}$ eine Funktion, $a \in D$. Dann sind folgende Aussagen äquivalent:

(1) f ist in a differenzierbar

(2) Es gibt eine in a stetige Funktionen $\varphi : D \to \mathbb{R}$ mit $f(x) = f(a) + (x-a)\varphi(x)$ für alle $x \in D$

(3) Es gibt eine Zahl $l \in \mathbb{R}$ und eine in a stetige Funktionen $g : D \to \mathbb{R}$ mit $g(a) = 0$ und

$$f(x) = f(a) + l(x-a) + (x-a)g(x)$$

(4) Es gibt eine Zahl $l \in \mathbb{R}$, so dass die durch die Gleichung

$$f(x) = f(a) + l(x-a) + r(x)$$

definierte Funktion („Rest“) $r : D \to \mathbb{R}$ die Bedingung $\lim\limits_{x\to a} \frac{r(x)}{x-a} = 0$ erfüllt.

(5) Es gibt eine lineare Abbildung $L : \mathbb{R} \to \mathbb{R}$ mit $\lim\limits_{x\to a} \frac{f(x)-f(a)-L(x-a)}{x-a} = 0$

Falls die fünf Eigenschaften zu treffen, so sind die Funktionen φ, g, r und die linear Abbildung L und die Konstante $l \in \mathbb{R}$ eindeutig bestimmt, und es gilt $l = \varphi(a) = f'(a) = L(1)$.
Zum **Beweis** zeigen wir noch $(4) \Longrightarrow (5)$ und $(5) \Longrightarrow (1)$, da wir in der Vorüberlegung schon die paarweise Äquivalenz von (1) bis (4) gezeigt haben.

Beachte: Jede lineare Abbildung $L : \mathbb{R} \to \mathbb{R}$ hat die Gestalt $h \mapsto lh$ mit $l := L(1) \in \mathbb{R}$.
Ist nun f in a differenzierbar, so gilt etwa (4) und wenn man $L(h) := lh := f'(a)h$ definiert hat man eine lineare Abbildung $L : \mathbb{R} \to \mathbb{R}$ gefunden, die

$$\lim_{x \to a} \frac{f(x) - f(a) - L(x-a)}{x-a} = 0$$

erfüllt.
Ist umgekehrt $L : \mathbb{R} \to \mathbb{R}$ eine lineare Abbildung, wie die die Gleichung in (5) erfüllt und gilt etwa $L(h) = lh$ mit einem $l \in \mathbb{R}$, so folgt

$$0 = \lim_{x \to a} \frac{f(x) - f(a) - l(x-a)}{x-a} = \lim_{x \to a} \frac{f(x) - f(a)}{x-a} - l,$$

das bedeutet, dass f in a differenzierbar ist und $l = f'(a)$ gelten muss.
Da Äquivalenz der Eigenschaft (4) oder (5) mit der Differenzierbarkeit stellt den Aspekt der „linearen Approximierbarkeit" in den Vordergrund. Die lineare Abbildung $L : \mathbb{R} \to \mathbb{R}$ mit $h \mapsto f'(a)h$ nennt man auch das Differenzial (oder Linearisierung) von f im Punkt a und bezeichnet es mit $d_a f$ oder $df(a)$. Nach Definition gilt also

$$d_a f(h) = df(a)(h) = f'(a)h.$$

Die linear-affine Funktion (Tangente)

$$\begin{aligned} t : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto f(a) + L(x-a) = f(a) + f'(a)(x-a) \end{aligned}$$

heißt auch *lineare Approximation* von f in a.

18.1.8 Feststellung

Ist f differenzierbar in $a \in D$, dann ist f auch stetig in a.

Das folgt unmittelbar aus 18.1.7(2) oder auch einfach aus

$$f(x) - f(a) = \frac{f(x) - f(a)}{x-a}(x-a) \quad (\text{für } x \neq a),$$

da die rechte Seite für $x \to a$ dem Grenzwert Null hat. Dass die Umkehrung hier von im Allgemeinen nicht gilt, seit das einfache Beispiel der Betragsfunktion im Punkt $a = 0$.
Das folgende Beispiel demonstriert aber, dass aus der Differenzierbarkeit von f in a nichts über die Stetigkeit in einer Umgebung von a zu folgen braucht: die Funktion

$$f : \mathbb{R} \to \mathbb{R} \ \text{ mit } \ f(x) = \begin{cases} x, & \text{falls } x \in \mathbb{Q}, \\ x + x^2, & \text{falls } x \notin \mathbb{Q}, \end{cases}$$

ist an der Stelle $a = 0$ differenzierbar und hat dort die Ableitung 1. Für alle Punkte $a \neq 0$ ist f nicht stetig und damit auch erst recht nicht differenzierbar. Für $a = 0$ und $x \neq 0$ ist der Differenzenquotient

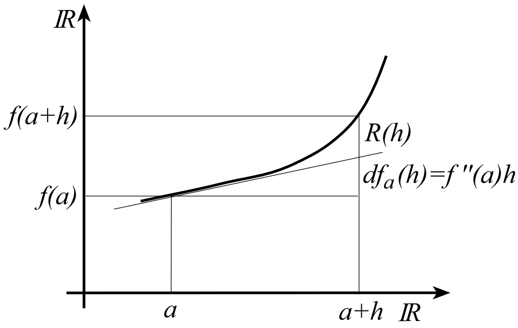
$$\frac{f(x) - f(0)}{x - 0} = \begin{cases} 1, & \text{falls } x \in \mathbb{Q}; \\ 1 + x, & \text{falls } x \notin \mathbb{Q}. \end{cases}$$

Die Funktion φ mit

$$\varphi(x) = \begin{cases} 1, & \text{falls } x \in \mathbb{Q} \\ 1 + x, & \text{falls } x \notin \mathbb{Q} \end{cases} \qquad (x \neq 0)$$

ist aber durch $\varphi(0) := 1$ stetig nach Null fortsetzbar, also ist $f'(0) = 1$.

18.1.9 Geometrische Interpretation des Differenzials



setzt man $x = a + h$, dann wird

$$f(a+h) - f(a) = L(h) + R(h).$$

Die Änderungsrate der Funktion beim Übergang von Argument a Argument $a + h$ wird also beschrieben durch eine lineare Funktion L und einen Rest R, der mit $h \to 0$ schneller gegen Null geht als h selber:

$$\lim_{h\to 0} \frac{R(h)}{h} = 0.$$

Ist $f : D \to \mathbb{R}$ in jedem Punkt $a \in D$ differenzierbar, dann besitzt also f in jedem $a \in D$ ein Differenzial:

$$\begin{aligned} d_a f : \mathbb{R} &\to \mathbb{R} \\ h &\mapsto d_a f(h) := f(a)h. \end{aligned}$$

Beachtet man speziell die Identität eingeschränkt auf D, dann gilt also $id(x) = x$ für alle $x \in D$ und $id'(x) = 1$ für alle $x \in D$. Das Differenzial der Identität

$$\begin{aligned} d\, id_a : \mathbb{R} &\to \mathbb{R} \\ h &\mapsto d\, id_a(h) = 1 \cdot h = h, a \in D, \end{aligned}$$

ist also von a unabhängig und wirkt als Identität.
Wegen $id(x) = x$ schreibt man für denn Differenzial der Identität einfach dx.
Wohlgemerkt dx ist die $\mathbb{R}$-lineare Abbildung

$$\begin{aligned} dx : \mathbb{R} &\to \mathbb{R} \\ h &\mapsto dx(h) = 1 \cdot h = h. \end{aligned}$$

Wie jede linear Abbildung $L : \mathbb{R} \to \mathbb{R}$ ein Vielfaches der identische Abbildung ist ($L = L(1) \cdot id_{\mathbb{R}}$), ist auch das Differenzial einer differenzierbaren Funktionen ein skalares vielfaches von dx:

$$\boxed{df = f'dx}$$

Das bedeutet, dass für alle $a \in D$ und alle $h \in \mathbb{R}$ gilt

$$d_a f(h) = f'(a)dx(h) = f'(a)h.$$

Betrachtet man die konstante Abbildung $h \mapsto f'(a)$, dann kann sie als Quotient der Differenziale (linear Funktionen) df_a und dx auffassen, denn für $h(\neq 0)$ gilt

$$\frac{d_a f}{dx}(h) = \frac{d_a f(h)}{dx(h)} = \frac{f'(a)h}{h} = f'(a).$$

Im eindimensionalen $\mathbb{R}$-Vektorraum $\mathbb{R}$ unterscheidet man in der Regel nicht zwischen linear Abbildungen $L : \mathbb{R} \to \mathbb{R}$ und Zahlen: jede linear Abbildung $L : \mathbb{R} \to \mathbb{R}$ hat die Gestalt $L(h) = L(1)h = lh$

mit einer eindeutig bestimmten Zahl $l(= L(1)) \in \mathbb{R}$. Umgekehrt bestimmt jede reelle Zahlen l eine linear Abfindung $L : \mathbb{R} \to \mathbb{R}$, nämlich die Linksmultiplikation mit l:

$$\begin{array}{rcl} L : \mathbb{R} & \to & \mathbb{R} \\ h & \mapsto & lh. \end{array}$$

Algebraisch bedeutet das:

$$\boxed{Hom(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}}$$

Erst in der Differenzialrechnung mehrerer Veränderlichen, wird die Bedeutung der Differenziale voll zum Tragen kommen.

18.1.10 Definition (höhere Ableitungen)

Ist $f : D \to \mathbb{R}$ differenzierbar ($D \subset \mathbb{R}$ echtes Intervall) und $f' : D \to \mathbb{R}$ die Ableitung von f. Ist f' wiederum differenzierbar, dann heißt f zweimal differenzierbar und man setzt $f'' := (f')'$ und nennt f'' die *zweite* Ableitung von f.
Allgemein definiert man rekursiv höhere Ableitungen von f:
$f : D \to \mathbb{R}$ heißt k-mal differenzierbar ($k \in \mathbb{N}$), falls $f^{(k-1)} : D \to \mathbb{R}$ differenzierbar ist und man setzt dann

$$f^{(k)} := (f^{(k-1)})'.$$

Zur Ergänzung setzt man noch $f^{(0)} := f$. *Jede* Funktion ist als Null-mal differenzierbar.

f heißt im Punkt $a \in D$ k-mal differenzierbar, $k \in \mathbb{N}$, $k \geq 1$, falls es ein $r > 0$ gibt, so dass f in $U := U_r(a) \cap D$ $(k-1)$-mal differenzierbar ist und $f^{(k-1)} : U \to \mathbb{R}$ in a differenzierbar ist. Eine Funktion $f : D \to \mathbb{R}$ ist also genau dann k-mal differenzierbar, wenn sie jedem Punkt $a \in D$ k-mal differenzierbar ist.
Besitzt $f : D \to \mathbb{R}$ eine k-te Ableitung $f^{(k)} : D \to \mathbb{R}$, dann sind alle

$$f^{(0)} = f, f' = f =^{(1)}, \ldots, f^{(k-1)}$$

stetig. Für $f^{(k)}$ braucht dies jedoch nicht zu zutreffen. Ist jedoch auch $f^{(k)}$ stetig, dann heißt f *k-mal stetig differenzierbar.*

Standardbezeichnung: $C^k(D) := \{f : D \to \mathbb{R};\ f k\text{-mal stetig diffferenzierbar}\}$; Alle $C^k(D)$ sind $\mathbb{R}$-Vektorräume, und es gelten die Inklusionen

$$C(D) := C^0(D) \supset C^1(D) \supset \ldots \supset C^k(D) \supset \ldots, \qquad k \in \mathbb{N}_0$$

wobei die Inklusionen echt sind. Existiert $f^{(k)}$ für *jedes* $k \in \mathbb{N}_0$, so heißt f beliebig oft (stetig) differenzierbar und man setzt $C^\infty(D) := \bigcap\limits_{k=0}^{\infty} C^k(D)$. $C^\infty(D)$ ist ebenfalls ein $\mathbb{R}$-Vektorraum. Es gilt etwa $\exp \in C^\infty(\mathbb{R})$.

Vom physikalischen Standpunkt besonders wichtig ist wegen der Newtonschen Bewegungsgesetze die zweite Ableitung.

18.1.11 Bemerkung

Die Definition der k-maligen Differenzierbarkeit in einem Punkt scheint gekünstelt, wir haben sie deshalb gewählt, weil wir auch Funktionen und ihre Ableitungen betrachten wollen, wenn die Definitions Bereiche keinen Intervalle sind. Die in der Praxis vorkommenden Definitionsbereiche sind aber meist endliche oder abzählbar unendliche Vereinigungen von (echte) Intervallen: $D = \bigcup\limits_{j \in J} D_j$, J endliche oder abzählbar unendliche Indexmenge.

So ist zum Beispiel für

$$\begin{array}{rcl} j : D := \mathbb{R} - \{0\} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{x} \end{array}$$

$D = D_1 \cup D_2$ mit $D_1 := \{x \in \mathbb{R};\ x > 0\}$ und $D_2 := \{x \in \mathbb{R};\ x < 0\}$ und für den Definitionsbereich des Tangens gilt

$$D = Def(\tan) = \bigcup_{j \in \mathbb{Z}} \left] \frac{\pi}{2}(2j-1);\ \frac{\pi}{2}(2j+1) \right[= \left\{ x \in \mathbb{R};\ x \neq \frac{\pi}{2}(2k+1),\ k \in \mathbb{R} \right\} .$$

Eine Funktion $f : D = \bigcup\limits_{j \in \mathbb{Z}} D_j \to \mathbb{R}$ wollen wir differenzierbar nennen, wenn $f|\ D_j$ differenzierbar ist ($j \in \mathbb{Z}$). Wie wir im nächsten Abschnitt sehen werden, sind zum Beispiel *rationale Funktionen* in ihrem Definitionsbereich (der die Vereinigung von endlich viele Intervallen ist) differenzierbar.
Allgemein benötigt man zur Definition der Ableitung einer Funktion f im Punkt $a \in D$ ($D \neq \emptyset$ beliebig) lediglich die Eigenschaft, dass „hinreichend nahe bei a“ Punkte $x \in D$ geben muss, die von a verschieden sind. Das ist sicherlich dann gegeben, wenn $a \in D$ ein Häufungspunkt von D ist. Ist D ein (echtes) Intervall, dann ist jeder Punkt $a \in D$ auch Häufungspunkt von D, die obige Bedingung also automatisch erfüllt.
Die Differenzierbarkeit von f in einem Punkt $a \in D$ ist eine *lokale Eigenschaft*, das heißt f ist genau dann differenzierbar in a, wenn es ein $r > 0$ gibt, so dass $f|\ D \cap U_r(a)$ in a differenzierbar ist. Man beweise diese Feststellung.

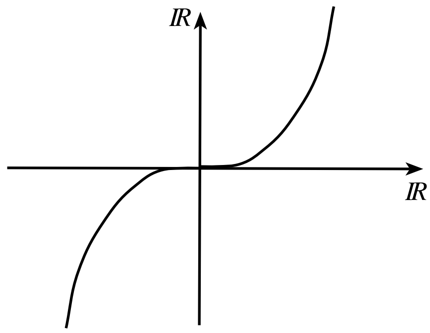
18.1.12 Abschreckende Beispiele

(a) Die Ableitung einer differenzierbaren Funktion $f : D \to \mathbb{R}$ muss nicht notwendig selbst wieder differenzierbar sein. Ein einfaches Beispiel ist

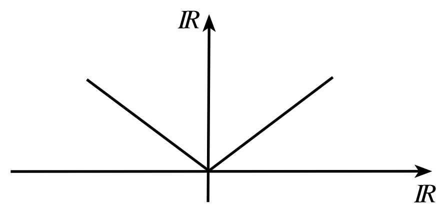
$$\begin{array}{rcl} f : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{2}|x|x, \end{array}$$

es ist also

$$f(x) = \begin{cases} \frac{1}{2}x^2, & \text{falls } x \geq 0 \\ -\frac{1}{2}x^2, & \text{falls } x \leq 0. \end{cases}$$



Graph(f)



Graph(f ')

Offensichtlich ist $f'(x) = |x|$, f' ist aber in Null nicht differenzierbar.
Nach dem gleichen Prinzip kann man Funktionen konstruieren, die in einem Punkt k-mal, aber nicht $(k+1)$-mal differenzierbar sind.

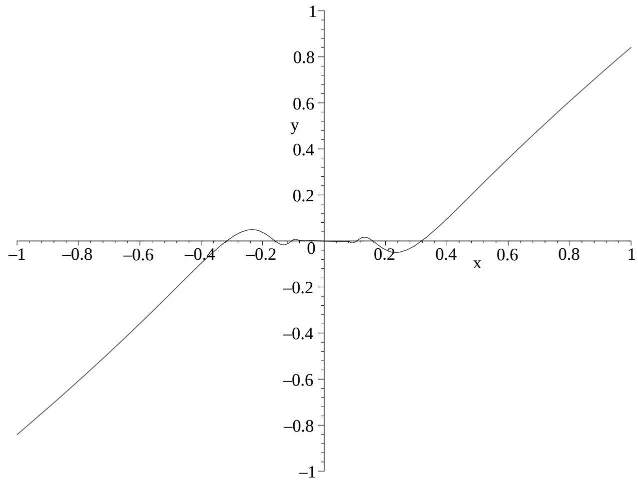


Abbildung 17: Graph von $f(x) = x^2 \sin \frac{1}{x}$

(b) Die Ableitung einer differenzierbaren Funktion $f : D \to \mathbb{R}$ braucht auch nicht stetig zu sein, wie das Beispiel

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \begin{cases} x^2 \sin \frac{1}{x}, & \text{falls } x \neq 0 \\ 0, & \text{falls } x = 0 \end{cases} \end{aligned}$$

zeigt. (siehe Abb. 17)

Hier ist (wie man sich leicht überzeugt)

$$f'(x) = \begin{cases} 2x \sin(\frac{1}{x}), & \text{falls } x \neq 0 \\ 0, & \text{falls } x = 0 \end{cases}$$

aber f' ist in $a = 0$ stetig.

(c) Es gibt Beispiele für stetige Funktionen $f : \mathbb{R} \to \mathbb{R}$, die in *keinem* Punkt differenzierbar sind. ein solches Beispiel ist etwa $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \sum \ldots \sum_{n=0}^{\infty} \frac{\{10^n x\}}{10^n}$, wo $\{y\}$ der kleinste Abstand von y zur nächsten ganzen Zahl ist.

18.2 Die Technik des Differenzierens

Um mit differenzierbaren Funktionen erfolgreich arbeiten zu können, stellen wir in diesem Abschnitt mit den *Ableitungsregeln* das entsprechende Handwerkszeug zur Verfügung. Wir teilen die Regel in drei Gruppen ein:

I. die algebraischen Regeln (18.2.1)

II. die Kettenregel (18.2.3)

III. der Satz über die Differentiation der Umkehrfunktion (18.2.5)

Die Kombination dieser Regel wird es uns ermöglichen, die Ableitungen auch komplizierter Funktionen zu bestimmen. Die Definitionsbereiche der betrachteten Funktionen sind in der Regel echte Intervalle.

18.2.1 Satz (algebraische Regeln)

Sei $D \subset \mathbb{R}$ ein echtes Intervall, $f, g : D \to \mathbb{R}$ Funktionen, die in $a \in D$ differenzierbar sind. So gilt dies auch für jede ihrer Linearkombinationen $\lambda f + \mu g$ $(\lambda, \mu \in \mathbb{R})$, ihr Produkt fg und falls $g(a) \neq 0$ gilt auch für den Quotienten $\frac{f}{g}$ und zwar gilt dann

(a) $(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$ (Linearität der Summenregel)
(b) $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$ (Produktregel)
(c) $\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}$ (Quotientenregel)

Zusatz:
Sind f und g in a k-mal differenzierbar ($n \in \mathbb{N}$), dann sind auch $\lambda f + \mu g$ $(\lambda, \mu \in \mathbb{R})$ und fg in a k-mal differenzierbar, und es gilt

$(d)(\lambda f + \mu g)^{(n)} = \lambda f^{(n)}(a) + \mu g^{(n)}(a)$ (Linearität)
$(e)(fg)^{(n)}(a) = \sum\limits_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} f^{(n-k)}(a) g^{(n)}(a) = \sum\limits_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} f^{(k)}(a) g^{(n-k)}(a)$ (Leibniz'sche Regel)

speziell also für $n = 2$ und $n = 3$

$$(fg)''(a) = f''g + f'g' + f'g' + fg'' = f''g + 2f'g' + fg''$$

$$(fg)'''(a) = f'''g + 3f''g' + 3f'g'' + g'''.$$

Zum Beweis von (a), (b), (c) benutzt man am besten die Äquivalenz von (1) und (2) aus dem Äquivalenssatz:
Wegen der Differenzierbarkeit von f in a gibt es eine im a stetige Funktionen $\varphi : D \to \mathbb{R}$ mit

$$(*) \qquad f(x) = f(a) + (x-a)\varphi(x) \text{ und } \varphi(a) = f'(a).$$

Analog gibt es wegen der Differenzierbarkeit von f in a eine in a stetige Funktionen $\psi : D \to \mathbb{R}$ mit

$$(**) \qquad f(x) = g(a) + (x-a)\psi(x) \text{ und } \psi(a) = g'(a).$$

Dann gilt aber

$$\begin{aligned}
(\lambda f + \mu g)(x) &= \lambda f(x) + \mu g(x) \\
&= \lambda f(a) + (x-a)\lambda\varphi(x) + \mu g(a) + (x-a)\mu\psi(x) \\
&= \lambda f(a) + \mu g(a) + (x-a)\underbrace{(\lambda\varphi(x) + \mu\psi(x))}_{=\rho(3x)} \\
&= (\lambda f + \mu g)(a) + (x-a)\rho(3x),
\end{aligned}$$

da g in a differenzierbar ist, und $\rho(a) = \lambda\varphi(a) + \mu\psi(a) = \lambda f'(a) + \mu g'(a)$ gilt, folgt also $(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$. Damit ist (a) bewiesen.
Aus $(*)$ und $(**)$ folgt für das Produkt

$$(fg)(x) = f(a)g(a) + (\varphi(x)g(a) + f(a)\psi(x) + \varphi(x)\psi(x)(x-a))(x-a)$$

Da als Faktor von $(x-a)$ stehende Funktion ist also Zusammensetzung in a stetiger Funktionen stetig in a und ihr Wert in a ist

$$\varphi(a)g(a) + f(a)\psi(a) = f'(a)g(a) + f(a)g'(a),$$

das beweist (b).
Für die Quotientenregel (c) beweisen wir zunächst den Spezialfall

$$\left(\frac{f}{g}\right)'(a) = -\frac{g'(a)}{(g(a))^2},$$

der allgemeine Fall folgt dann wegen $\frac{f}{g} = f \cdot \frac{1}{g}$ aus der Produktregel.
Wegen der Stetigkeit von g gibt es eine Umgebung $U_r(a)$ mit $g(x) \neq 0$ für alle $x \in D \cap U_r(a)$. Für dieses x gilt dann

$$\frac{1}{g(x)} - \frac{1}{g(a)} = -\frac{g(x) - g(a)}{g(a)g(x)} = \left(-\frac{\psi(x)}{g(a)g(x)}\right)(x-a).$$

Dies als Faktor $(x-a)$ stehende Funktion ist dann (wegen der Stetigkeit von f und ψ in a) auch stetig in a, und ihr Wert in a ist

$$-\frac{g'(a)}{(g(a))^2}.$$

Das beweist den Spezialfall der Quotientenregel. Der allgemeine Fall ergibt sich -wie schon bemerkt- durch Anwendung der Produktregel an

$$\frac{f}{g} = f \cdot \frac{1}{g}.$$

Die Regel (d) ist evident.
Die allgemeine Leibniz'sche Regel (e) für die n-te Ableitung eines Produkts ergibt sich durch Induktion nach n in völliger Analogie zum Beweis des Binomischen Satzes (vgl. §2.4.1).

□

18.2.2 Beispiele und Bemerkungen

(a) Für $n \in \mathbb{N}$ ergibt sich nochmals als Ableitung für die Potenzfunktion

$$\begin{array}{rcl} p_n : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^n \end{array}$$

$p_n'(x) = nx^{n-1} = np_{n-1}(x)$ aus $p_{n+1} = p_1 p_n$ durch vollständige Induktion nach n unter Verwendung der Produktregel.

(b) Für

$$\begin{array}{rcl} f : \mathbb{R}_+^* & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{x^n} \quad (n \in \mathbb{N}) \end{array}$$

liefert die Quotientenregel

$$f'(x) = -\frac{(nx^{n-1})}{(x^n)^2} = -nx^{-n-1}.$$

Diese Regel gilt auch für $x < 0$.

Zusammengefasst: Es gilt für alle $n \in \mathbb{Z}$

$$\mathrm{D}x^n = \frac{d}{dx}(x^n) = nx^{n-1}$$

und alle $x \in \mathbb{R}$ mit $x \geq 0$ bzw. alle $x \in \mathbb{R}$ mit $x < 0$

(c)

$$\begin{array}{rcl} tan : \left]-\frac{\pi}{2}; \frac{\pi}{2}\right[& \to & \mathbb{R} \\ x & \mapsto & \tan x = \dfrac{\sin x}{cos x} \end{array}$$

ist differenzierbar, und es gilt nach der Quotientenregel

$$\tan'(x) = \frac{\sin'(x)\cos(x) - \sin(x)\cos'(x)}{\cos^2(x)} = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)} = 1 + \tan^2(x).$$

In jedem Intervall $(k \in \mathbb{Z})\ \big](2k-1)\frac{\pi}{2}, (2k+1)\frac{\pi}{2}\big[$ hat den Tangens diese Ableitung.

(d) Aus 18.2.2 folgt insbesondere, dass jedes Polynom

$$\begin{array}{rcl} p : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & a_0 + a_1 x + \ldots + a_n x^n,\ a_j \in \mathbb{R}, 0 \le j \le n \end{array}$$

auf ganz $\mathbb{R}$ differenzierbar ist, und dass gilt

$$p'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1}.$$

Beim (einmaligen) Differenzieren erniedrigt sich also der Grad eines Polynoms um 1. Es ist $a_0 = p(0)$ und $a_1 = p'(0)$ und allgemein $a_j := \frac{p^{(j)}(0)}{j!},\ 0 \le j \le n$.
Die Koeffizienten a_j sind also durch p eindeutig bestimmt:

$$\boxed{a_j = \frac{p^{(j)}(0)}{j!}} \qquad 0 \le j \le n.$$

Offensichtlich gilt $p^{(n+1)}(x) = 0$ für alle $x \in \mathbb{R}$ und damit auch $p^{(k)}(x) = 0$ für alle $k \ge n+1$. Offensichtlich gilt damit für jedes Polynom $p \in C^\infty(\mathbb{R})$.

Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ kann damit kein Polynom sein, für sie gilt $\exp \in C^\infty(\mathbb{R})$, aber $\exp^{(k)}(x) \neq 0$ für alle $k \in \mathbb{N}_0$ und alle $x \in \mathbb{R}$.

(e) Gilt $f, g \in C^{(n)}(D)$, dann ist nach (c) auch $fg \in C^{(n)}(D)$, $C^{(n)}(D)$ ist daher sogar eine Funktionenalgebra.

(f) Eine rationale Funktion $f = \frac{P}{Q}$, P, Q Polynome (oBdA Grad Q minimal) ist in ihrem natürlichen Definitionsbereich $D = \{x \in \mathbb{R};\ Q(x) \neq 0\}$ nach unserer Vereinbarung differenzierbar. D ist die Vereinigung von endlich vielen Intervallen, welche durch die endlich vielen Nullstellen von Q bestimmt sind.

(g) Beispiel:

$$f(x) = \frac{1}{1 - x^2} = \frac{1}{Q(x)}$$

Die Nullstellen des Nenners Q sind 1 und -1. Der Definitionsbereich von f ist daher

$$D :=]-\infty, -1[\ \cup\ ]-1, 1[\ \cup\ ]1, \infty[,$$

also eine Vereinigung von Intervallen. In jedem Teilintervall ist f differenzierbar und hat dort die Ableitung

$$f'(x) = \frac{2x}{(1-x)^2},$$

$f'(x)$ ist also wieder eine rationale Funktion. (siehe Abb. 18)

(h) Sind f und g in allen Punkten $x \in D$ differenzierbar, so besagen die Regeln aus 18.2.2 insbesondere, dass das Differentialquotient

$$\begin{array}{rcl} \mathrm{D} : V & \to & Abb(D, \mathbb{R}) \\ f & \mapsto & f', \end{array}$$

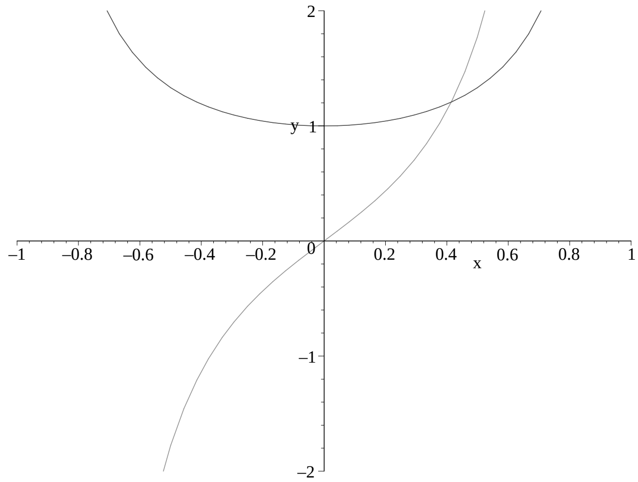


Abbildung 18: Graphen von $f(x) = \frac{1}{1-x^2}$ und $f'(x)$

der also jeder differenzierbaren Funktion auf D ihre Ableitung D$f = f'$ zuordnet, ein linearer Operator (eine lineare Abbildung) ist. V ist dabei der $\mathbb{R}$-Vektorraum der differenzierbaren Funktionen auf D.

Beachte: Es gilt auch $fg \in V$, falls $f, g \in V$ gilt.

Ist etwa $f(x) = x^2 \exp(x)$, dann ist

$$f^{(2002)}(x) = x^2 \exp(x) + 2002 \cdot 2x \exp(x) + \binom{2002}{2} \cdot 2 \exp(x) = (x^2 + 4004x + 2002) \exp(x)$$

Wenn man die Ableitung von

$$\begin{array}{rcl} h : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & (x^2+1)^{1001} \end{array}$$

ausrechnen sollte, könnte man $(x^2+1)^{1001}$ nach der Binomischen Formel berechnen, man erhält ein Polynom der Gestalt

$$x \mapsto x^{2002} + 1001x^{2000} + \ldots + 1$$

und kann dieses differenzieren.
Einfacher geht es mit der Kettenregel, wenn man beachtet, dass sich h aus zwei Funktionen zusammensetzt

$$x \mapsto x^2 + 1 := y \text{ und } y \mapsto y^{1001}$$

18.2.3 Satz (Kettenregel)

Sind $D, D^* \subset \mathbb{R}$ echte Intervalle und $f : D \to \mathbb{R}$ und $g : D^* \to \mathbb{R}$ Funktionen und es gelte $f(D) \subset D^*$.
Dann gilt: Ist f differenzierbar in $a \in D$, g ist differenzierbar in $b = f(a) \in D^*$, dann ist auch die Zusammensetzung

$$g \circ f : D \to \mathbb{R}$$

differenzierbar in a und es gilt

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

Zum Beweis sei wieder

$$f(x) = f(a) + (x-a)\varphi(x), \ \varphi \text{ stetig in } a \text{ und } \varphi(a) = f'(a)$$

bzw.

$$g(y) = g(b) + (y-b)\psi(y),\ \psi \text{ stetig in } b = f(a) \text{ und } \psi(b) = g'(b).$$

Für die Zusammensetzung $g \circ f$ ergibt sich

$$(g \circ f)(x) = g(f(x)) = g(b) + \psi(b + \varphi(x)(x-a))\varphi(x)(x-a).$$

Da $x \mapsto \psi(b+\varphi(x)(x-a))\varphi(x)$ an der Stelle a stetig ist und dort den Wert $\psi(b)\varphi(a) = g'(f(a))f'(a)$ hat, ergibt sich die Behauptung.

18.2.4 Bemerkungen und Beispiele

(a) Der naheliegende Beweisansatz mit Hilfe des Differenzenquotienten ($x \neq a$)

$$\frac{g(f(x)) - g(f(a))}{x-a} = \frac{g(f(x)) - g(f(a))}{f(x) - f(a)} \cdot \frac{f(x) - f(a)}{x-a}$$

erfordert, dass wenigstens in einer Umgebung von a $f(x) \neq f(a)$ gelten muss.

In jeder Umgebung von a kann es jedoch Punkte $x \neq a$, für die $f(x) - f(a) = 0$ gilt. Mit etwas Mühe (man unterscheidet die Fälle: $f'(a) \neq 0$ und $f'(a) = 0$) kann man aber auch den obigen Beweisansatz retten, einfacher geht es mit dem von uns verwendeten Ansatz.

(b) Unser Eingangsbeispiel

$$\begin{aligned} h : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto (x^2+1)^{1001} \end{aligned}$$

wird jetzt erledigt durch die Beobachtung, dass $h = g \circ f$ mit $f(x) = x^2 + 1$ und $g(y) = y^{1001}$ ist. Es ist also (für alle $x \in \mathbb{R}$)

$$h'(x) = (g \circ f)'(x) = g'(f(x)) \cdot f'(x) = 1001y^{1000} \cdot 2x = 2002x(x^2+1)^{1000}.$$

(c) Die allgemeinen Exponentialfunktionen

$$\begin{aligned} \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \exp_a(x) = \exp(x \log a) =: a^x \quad (a \in \mathbb{R}_+^*) \end{aligned}$$

sind differenzierbar, und es gilt

$$\exp_a'(x) = \log a \exp(x \log a) = \log a \exp_a(x)$$

Als letzte Regel für die Technik des Differenzierens wollen wir eine Regel ableiten, die es gestattet, aus Eigenschaften einer differenzierbaren Funktion f, die umkehrbar ist, z.B. weil sie streng monoton ist, auf die Differenzierbarkeit der Umkehrfunktion zu schließen und einen Zusammenhang zwischen der Ableitung des Umkehrfunktion und der Funktion abzuleiten. Wie dieser Zusammenhang aussieht, ist leicht zu erraten.
Wir nehmen an, dass g die Funktion f umkehrt, also $g(f(y)) = y$ für alle $y \in D = Def(f)$ und $f(g(x)) = x$ für alle $x \in Def(g)$ gilt. Nehmen wir weiter an, dass f in $b \in D$ und g in $a := f(b) \in Def(g)$ differenzierbar ist. Dann können wir die Kettenregel verwenden und erhalten aus $(f \circ g)(x) = x$ für alle $y \in D$

$$(f \circ g)'(a) = f'(g(a))g'(a) = 1.$$

Also eine notwendige Bedingung für die Differenzierbarkeit von g in $f(b)$, dass $f'(b) \neq 0$ sein muss.

Wir wollen jedoch die Differenzierbarkeit von f aus Eigenschaften von f schließen, die Bedingung $f'(b) \neq 0$ ist aber ein wichtiger Hinweis.
Wir formulieren den Satz über die Differenzierbarkeit der Umkehrfunktion gleich so, wie wir in der meisten Fällen anwenden werden.
Man beachte in diesem Zusammenhang auch, dass für stetige Funktionen f auf (echten) Intervallen die *strenge Monotonie*(d.h. f ist streng monoton wachsend oder streng monoton fallend) und die Injektivität äquivalente Begriffe sind (vgl. ???). Erinnern Sie sich noch an den Beweis? Die eine Richtung ist offensichtlich. Welche?

18.2.5 Theorem (Differenzierbarkeit der Umkehrfunktion)

Sei $D \subset \mathbb{R}$ ein echtes Intervall: $f : D \to \mathbb{R}$ differenzierbar, streng monoton und $f'(y) \neq 0$ für alle $y \in D$.
Dann besitzt $f : D \to f(D) =: D^*$ (D^* ist wieder Intervall) eine differenzierbare Umkehrfunktion $g : D^* \to D \subset \mathbb{R}$,

$$g(x) = y \iff x = f(y),$$

und es gilt für alle $x \in D^*$

$$g'(x) = \frac{1}{f'(y)}.$$

Zum **Beweis** sei f differenzierbar in $b \in D$ und es sei $a := f(b)$ ($\iff g(a) = b$).

Die Existenz der Umkehrfunktion ist nach der Voraussetzung klar, nach Satz ???. Da f insbesondere stetig ist, ist $D^* = f(D)$ auch wieder ein Intervall.
Die Umkehrfunktion g ist selbst wieder streng monoton (im selben Sinne wie f) und auch stetig (da $g(D^*) = D$ ein Intervall ist). Wegen der Differenzierbarkeit von f in b gibt es eine in b stetige Funktion $\varphi : D \to \mathbb{R}$ mit

$$f(y) - f(b) = (y - b)\varphi(y) \;\text{ und }\; \varphi(b) = f'(b).$$

Nach Voraussetzung ist $\varphi(b) = f'(b) \neq 0$, daher ist wegen der strengen Monotonie von f sogar $\varphi(y) \neq 0$ für alle $y \in D$.
Daher folgt

$$y - b = \frac{1}{\varphi(y)}\,(f(y) - f(b))$$

oder

$$g(x) - g(a) = \frac{1}{\varphi(g(x))}(x - a)$$

Da aber φ und g stetig in a sind, ist auch $\varphi \circ g$ stetig in a und damit auch $\frac{1}{\varphi \circ g}$, daher ist g differenzierbar in a und

$$g'(a) = \frac{1}{\varphi(g(a))} = \frac{1}{f'(g(a))} = \frac{1}{f'(b)}.$$

□

18.2.6 Beispiele und Bemerkungen

(a) Die Voraussetzung, dass $f'(x) \neq 0$ ist wesentlich, so ist etwa die Potenzfunktion

$$\begin{aligned} p_3 : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^3 \end{aligned}$$

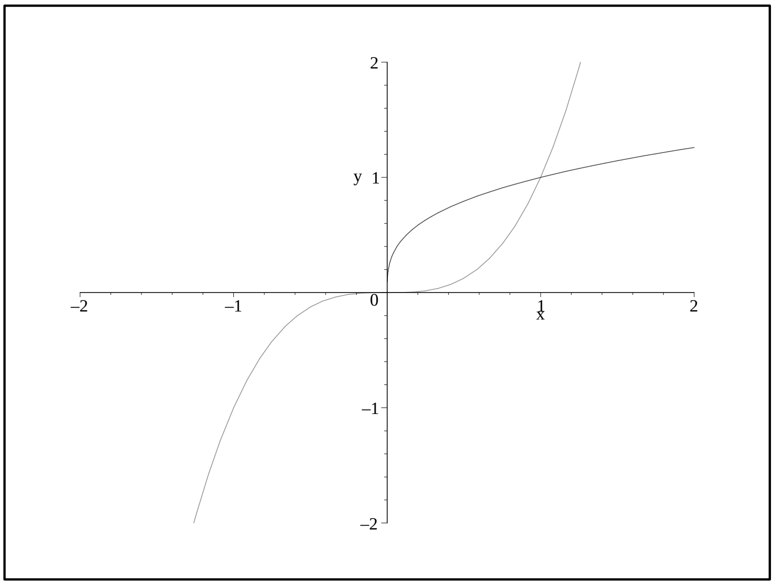


Abbildung 19: Graphen der Funktionen x^3 und $\sqrt[3]{x}$

streng monoton wachsend, differenzierbar auf ganz $\mathbb{R}$, und es gilt $P_3'(x) = 3x^2$ und $p_3(x) = 0 \iff x = 0$.
Die Umkehrfunktion $\sqrt[3]{x}$, ist aber in $x = 0$ nicht differenzierbar, jedoch für alle $x > 0$ oder für alle $x < 0$. (siehe Abb. 19)

(b) Wegen $\exp'(y) = \exp(y) \neq 0$ für alle $y \in \mathbb{R}$ ist $\log : \mathbb{R}_+^* \to \mathbb{R}$ differenzierbar, und es gilt

$$\log'(x) = \frac{1}{\exp'(\log x)} = \frac{1}{\exp(\log x)} = \frac{1}{(\exp \circ \log)(x)} = \frac{1}{x}$$

für alle $x > 0$.

(c) Für alle $\alpha > 0$ sind die Potenzfunktionen

$$\begin{array}{rcl} p_\alpha : \mathbb{R}_+^* & \to & \mathbb{R} \\ x & \mapsto & x^\alpha := \exp(\alpha \log x) \end{array}$$

differenzierbar, und es gilt für alle $x \in \mathbb{R}_+^*$

$$p_\alpha'(x) = \alpha x^{\alpha-1} = \alpha p_{\alpha-1}(x).$$

Der Beweis ist klar nach der Kettenregel und (b)

$$p_\alpha'(x) = \exp(\alpha \log x) \cdot \alpha \cdot \frac{1}{x},$$

also

$$p_\alpha'(x) = \alpha x^{\alpha-1} = \alpha p_{\alpha-1}(x) \quad (\alpha \in \mathbb{R},\ x > 0$$

(d) Ist $f : D \to \mathbb{R}$ eine *nullstellfreie* differenzierbare Funktion, dann ist nach dem Zwischenwertsatz $f(x) > 0$ für alle $x \in D$ oder $f(x) > 0$ für alle $x \in D$. In jedem Fall ist

$$(\log \circ |f|)' = \frac{f'}{f}.$$

Man nennt allgemein $\frac{f'}{f}$ *logarithmische Ableitung* von f' oder $|f|$.

19 Lokale Extrema, Mittelwertsätze, Anwendungen

Eine der wesentlichen Aufgaben der Analysis besteht im Studium des Änderungsverhaltens von Funktionen in Abhängigkeit von Änderungen der betreffenden Argumente. Die Ableitungen im geometrischen Bild, die Steigung der Tangente an den Graphen der Funktion, geben diese Änderung „lokal" an. Es zeigt sich nun, dass man aus der Kenntnis des *lokalen* Veränderungsverhaltens einer differenzierbaren Funktion auf das *globale* Verhalten schließen kann. Viele Eigenschaften einer differenzierbaren Funktion spiegeln sich in Eigenschaften ihrer Ableitungen f' bzw. der höheren Ableitungen. So können etwa *lokale Extrema*, das *Monotonieverhalten*, die *Konvexität* bzw. *Konkavität* einer Funktion mit Hilfe der Ableitungen untersuchen werden. Der zentrale Satz ist der (erste) *Mittelwertsatz der Differentialrechnung* (1.MWSD). Wir folgern den klassischen Beweisschema und folgern ihn aus dem Satz *von Rolle*, diesen wiederum aus einem schon *Fermat* geläufigen notwendigen Kriterium für das Vorliegen eines Extremum für differenzierbare Funktionen in offenen Intervallen. In Wirklichkeit sind alle die Sätze äquivalent (in Wahrheit sogar äquivalent zum Vollständigkeitsaxiom für $\mathbb{R}$) und auch äquivalent zum zweiten MWSD, mit dessen Hilfe wir auch die sog. *Bernoulli-l'Hospital*'schen Regeln ableiten werden. Als Anwendung werden wir auch einige *Extremalprobleme* (aus verschiedenen Anwendungsbereichen) behandeln und weitere *nützliche Ungleichungen* ableiten.

19.1 Lokale Extrema, der MWSD

Wir erinnern an folgende Definition

19.1.1 Definition (lokale und globale Extrema)

Ist $D \neq \emptyset$ und $f : D \to \mathbb{R}$ eine Funktion. Man sagt: f hat in $x_0 \in D$

(a) ein *globales Maximum* (bzw. globales Minimum), wenn für *alle* $x \in D$ gilt

$$f(x) \leq f(x_0) \text{ bzw. } f(x) \geq f(x_0)$$

(b) ein *lokales Maximum* (bzw. lokales Minimum), wenn es ein $r > 0$ gibt, so dass für alle $x \in U_r(x_0) \cap D$ gilt

$$f(x) \leq f(x_0)(\text{ bzw. } f(x) \geq f(x_0))$$

(c) In allen Fällen nennt man x_0 eine lokale bzw. globale Extremalstelle von f und der Funktionswert $f(x_0)$ das globale bzw. lokales Extremum von f in x_0.

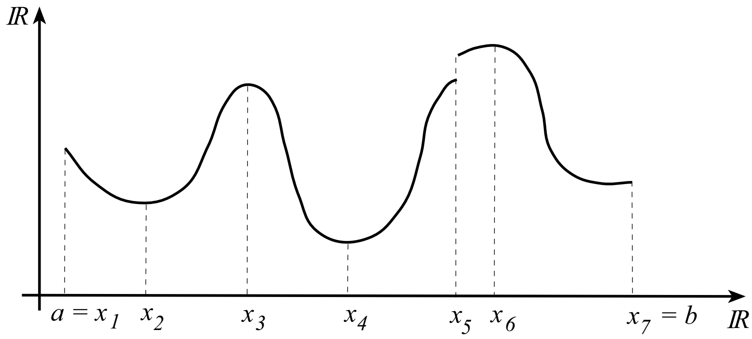
(d) Gilt an den entsprechenden Ungleichungen das Gleichheitszeichen jeweils nur für $x = x_0$, so spricht man von einem *strikten* lokalen oder globalen Maximum bzw. Minimum an der Stelle x_0.

(e) Satt globales Maximum bzw. Minimum sagt man gelegentlich auch *absolutes* Maximum bzw. minimum und statt lokales Maximum bzw. Minimum auch *relatives* Maximum bzw, Minimum.

19.1.2 Bemerkungen

(a) Eine auf einem kompakten Intervall $[a, b]$ stetige Funktionen, hat nach dem Satz von Weierstrass dort ein globales Maximum und ein globales Minimum. Dieser Satz ist jedoch ein reiner Existenzsatz und die kaum einen Hinweis, wie man die jeweilige Extrema Stelle ermittelt.

(b) Eine Funktion kann viele lokale Extrema haben, ein globales Maximum (bzw. Minimum) ist natürlich auch ein lokales Maximum (bzw. Minimum), aber im Allgemeinen nicht umgekehrt. Auch er kann ein globales Maximum (bzw. Minimum) an verschiedenen Stellen des Definition Bereiches angenommen werden. Der geneigte Leser macht sich klar, an welcher Stelle die Funktion f, deren Graf in der folgenden Abbildung skizziert ist, jeweils lokale bzw. globalen Maxima besitzt.



19.1.3 Fermat, ≈ 1638

Ist $D \subset \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion, $a \in D$ und es gebe ein $r > 0$ mit $U_r(a) \subset D$. Hat dann f in a ein lokales Extremum und ist f in a differenzierbar, dann gilt

$$\boxed{f'(a) = 0}$$

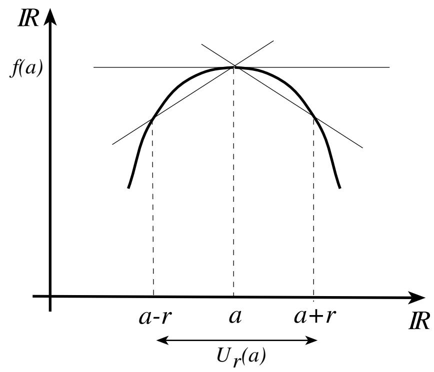
Beweis : Wir nehmen oBdA an, das f an der Stelle a ein lokales Maximum hat (im Falle eines Minimums ersetzt man f durch $-f$). Wir greifen auf die Definition der Differenzierbarkeit als limes des Differenzenquotient zurück.

Für alle $x \in U_r(a)$ mit $x > a$ gilt wegen $f(x) \le f(a)$

$$\frac{f(x) - f(a)}{x - a} \le 0$$

also ist auch

$$f'(a) = \lim_{\substack{x \to a \\ x > a}} \frac{f(x) - f(a)}{x - a} \le 0.$$



Für alle $x \in U_r(a)$ mit $x < a$ gilt auch wegen $f(x) \le f(a)$

$$\frac{f(x) - f(a)}{x - a} \ge 0$$

und damit auch

$$f'(a) = \lim_{\substack{x \to a \\ x < a}} \frac{f(x) - f(a)}{x - a} \geq 0,$$

also insgesamt $f'(a) = 0$.

$\square$

19.1.4 Bemerkungen

(a) Geometrisch besagt der Satz von Fermat:

Wenn f in einem *inneren Punkt* $a \in D$ (das heißt es gibt ein $r > 0$ mit $U_r(a) \subset D$) ein lokales Extremum besitzt und wenn f in a differenzierbar ist, besitzt der Graph von f im Punkt $(a, f(a)) \in \mathbb{R}^2$ eine waagerechte Tangente.

(b) Die Aussage des Fermatschen Lemmas wird für Randpunkte im Allgemein falsch.

Ist etwa $D := [0, 1]$ und

$$\begin{aligned} f : D &\to \mathbb{R} \\ x &\mapsto x \end{aligned}$$

die Einschränkung von $id_{\mathbb{R}}$ auf $[0, 1]$, dann hat f an der Stelle 0 ein (sogar globales) Minimum und an der Stelle 1 ein (sogar globales) Maximum, es ist aber $f'(x) = 1 \neq 0$ für alle $x \in [0, 1]$.

(c) Wie das einfache Beispiel des Polynoms

$$\begin{aligned} p_3 : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^3 \end{aligned}$$

und $a = 0$ zeigt, bedeutet das Verschwinden der Ableitung in einem (inneren) Punkt keineswegs, dass an dieser Stelle ein lokales Extrema vorliegen muss.

Das Verschwinden der Ableitung in einen inneren Punkt (Differenzierbarkeit natürlich vorausgesetzt), ist also lediglich eine *notwendige Bedingung* für das Vorliegen eines lokalen Extremum.

(d) Ist $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$ eine Funktion, die auf Extremalstellen zu untersuchen ist, dann hat man folgende Kandidaten ins Auge zu fassen:

1. Die Randpunkte vor D (also 2 Punkte);
2. Die Punkte von D, in denen f nicht differenzierbar ist.
3. Die inneren Punkte $x_0 \in D$, in denen f differenzierbar ist und für die $f'(x_0) = 0$ gilt. Solche Punkte nennt man auch *kritische Punkte*.

In keinem dieser Punkte braucht aber eine Extremalstelle vorzuliegen.

ABBILDUNG

Wir werden aber bald auch *hinreichende Kriterien* dafür ableiten. Obwohl der Mittelwertsatz der Differenzialrechnung eine sehr anschauliche geometrische Interpretation hat, muss er bewiesen werden. Wir verweisen zunächst einen Spezialfall, aus dem der MWSD dann einfach abzuleiten ist.

19.1.5 Satz (von Rolle, Michael Rolle (1652-1719))

Sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall ($a < b$) und $f : [a,b] \to \mathbb{R}$ eine stetige Funktionen, die im offenen Intervall $]a,b[$ differenzierbar ist und für welche $f(a) = f(b)$ gilt. Dann gibt es einen Punkt $\xi \in]a,b[$ mit $f'(\xi) = 0$.
Insbesondere liegt zwischen zwei Nullstellen von f $(f(a) = f(b) = 0)$ stets eine Nullstellen der Ableitung.

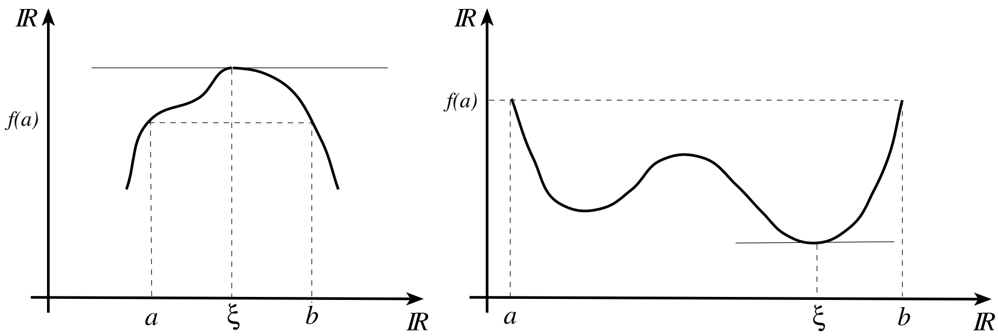
Beweis : Ist f konstant auf $[a,b]$, dann gilt $f'(x) = 0$ für alle $x \in [a,b]$, insbesondere gilt etwa für $\xi := \frac{a+b}{2} \in]a,b[$ auch $f'(\xi) = 0$.

Ist f aber nicht konstant, dann hat f als stetige Funktionen auf der kompakte Intervall $[a,b]$ ein globales Maximum und ein globales Minimum. Diese können wegen $f(a) = f(b)$ nicht bei der in den Endpunkten des Intervall vorliegen. Wenn ein globales Minimum bei a vorliegt und damit auch bei b, dann liegt ein existierendes globales Maximum in einem Punkt $\xi \in]a,b[$. Wegen $]a,b[= U_r(\xi)$ mit $\xi := \frac{a+b}{2}$ und $r := \frac{b-a}{2}$ sind Voraussetzungen des Fermatschen Lemmas erfüllt. Es gilt also $f'(\xi) = 0$.

Im Fall, dass in a (und damit auch in b) ein globales Maximum vorliegt, schließt man völlig analog: Ein sicher vorhandenes globales Minimum muss dann wieder in $[a,b[$ liegen etc.

□

19.1.6 Geometrische Veranschaulichung



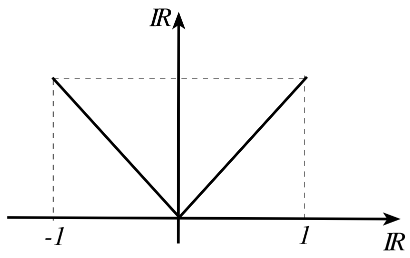
19.1.7 Bemerkungen zu den Voraussetzungen des Satzes von Rolle

Die Voraussetzungen des Satzes von Rolle sind im folgenden Sinne „scharf“:
Verzichtet man auf eine der Voraussetzungen, dann wird der Satz i.A. falsch:

(a) Verzichtet man auf die Voraussetzung der Differenzierbarkeit auch nur in einem Punkt $x_0 \in]a,b[$, dann wird der Satz i.A. falsch.
Ist z.B. $[a,b] = [-1,1]$ und

$$\begin{aligned} f : [-1,1] &\to \mathbb{R} \\ x &\mapsto |x| \end{aligned}$$

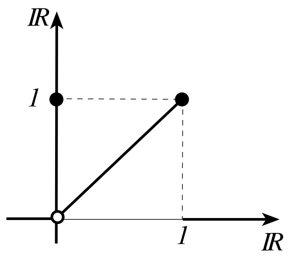
die Betragsfunktion, dann gilt $f(-a) = f(a) = 1$ und f ist stetig in $[-1,1]$, aber f ist nicht differenzierbar in 0, es gibt bei $\xi \in]-1,1[$ mit $f'(\xi) = 0$.



(b) $f : [0,1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 1, & \text{für } x = 0; \\ x, & \text{für } 0 \leq x \leq 1 \end{cases}$$

erfüllt $f(0) = f(1)(= 1)$, ist differenzierbar in $]0,1[$ mit Ableitung $f'(x) = 1$ für alle $x \in]0,1[$, aber es gibt kein $\xi \in]0,1[$ mit $f'(\xi) = 0$. Hier ist f nicht stetig in Null.



Der Mittelwertsatz ist nun eine einfache Konsequenz aus dem Satz von Rolle, er hat zunächst die gleichen Voraussetzungen, man verzichtet jedoch auf die Bedingung

$$f(a) = f(b).$$

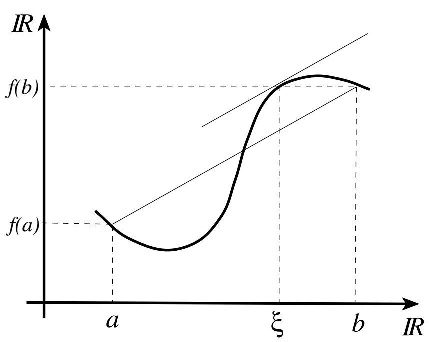
19.1.8 Theorem (MWSD, Lagrange (1797) (J.L.Lagrange, 1736-1813))

Ist $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall ($a < b$) und $f : [a,b] \to \mathbb{R}$ eine stetige Funktion, die im offenen Intervall $]a,b[$ differenzierbar ist. Dann gibt es ein $\xi \in]a,b[$ mit

$$\boxed{\tfrac{f(b)-f(a)}{b-a} = f'(\xi)}$$

19.1.9 Geometrische Veranschaulichung

Auf dem Graphen von f gibt es mindestens einen Punkt $(\xi, f(\xi))$, in welchem die Tangente parallel zur Sekanten durch die Punkte $(a, f(a))$ und $(b, f(b))$ ist.



Zum **Beweis** betrachtet man die Hilfsfunktion $h : [a,b] \to \mathbb{R}$ mit

$$h(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

(Geometrisch bedeutet dies eine Scherung des Graphen von f.)

Dann ist auch h stetig in $[a,b]$ und differenzierbar in $]a,b[$, und es gilt

$$h(a) = f(a) \text{ und } h(b) = f(b) - \frac{f(b)-f(a)}{b-a}(b-a) = f(a),$$

also auch $h(a) = h(b)$.
damit erfüllt h die Voraussetzung des Satzes von Rolle, es gibt also wegen $h'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$ ein $\xi \in]a,b[$ mit

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

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

19.1.10 Bemerkungen

(a) Der „Nachteil" des MWSD liegt in der Tatsache, dass die Zwischenstelle $\xi \in]a,b[$ nicht näher spezifiziert wird, wie einfache Beispiele zeigen, kann es mehrere ξ mit der genannten Eigenschaft geben.

(b) Häufig verwendet man den MWSD in der folgenden Form:
Ist f stetig auf dem Intervall $[a, a+h]$ $(h > 0)$ bzw.$[a+h, a]$ (falls $h < 0$) und differenzierbar in $]a, a+h[$ bzw. $]a+h, a[$, dann gibt es ein ϑ mit $0 < \vartheta < 1$ und

$$f(a+h) = f(a) + f'(a + \vartheta h)h.$$

(c) Eine kleine Anwendung:
Der MWSD eignet sich zur näherungsweisen Berechnung von Funktionswerten.
Um etwa $\sqrt{2}$ näherungsweise zu berechnen setzen wir $2 = a + h$ mit $a = 1,96$ und $h = 0,04$ und erhalten wegen

$$\sqrt{a+h} = \sqrt{a} + \frac{1}{2\sqrt{a+\vartheta h}}$$

mit $0 < \vartheta < 1$

$$\sqrt{2} = \sqrt{1,96 + 0,04} = \sqrt{1,96} + 0,04 \cdot \frac{1}{2\sqrt{3}}$$

mit $1,96 < \xi < 2$ $(< 2,0164)$.
Hieraus folgt mit $1,4 < \sqrt{\xi} < 1,42$ bereits $1,41414 < \sqrt{2} < 1,41429$.
Gar nicht mal so schlecht!

In der anspruchsvolleren Schulbuchliteratur zu Analysis wird die folgende Variante der MWSD diesem meist vorgezogen.

19.1.11 Schrankensatz

Ist $f : [a,b] \to \mathbb{R}$ eine stetige und in $]a,b[$ differenzierbare Funktion.
Für die Ableitung gelte $m \leq f'(\xi) \leq M$ für alle $\xi \in]a,b[$ mit festen Konstanten $m, M \in \mathbb{R}$. Dann gilt für alle $x, y \in [a,b]$ die Abschätzung

$$\boxed{m(y-x) \leq f(y) - f(x) \leq M(y-x)}$$

Der Beweis folgt unmittelbar aus

$$f(y) - f(x) = f'(\xi)(y-x)$$

19.1.12 Folgerung

Eine differenzierbare Funktion f auf einem echten Intervall $D \subset \mathbb{R}$ mit beschränkter Ableitung ist dort Lipschitz-stetig. Genauer gilt:
Ist $f : D \to \mathbb{R}$ differenzierbar und gilt $|f'(\xi)| \leq L$ für alle $x \in D$, dann gilt für alle $x, y \in D$

$$(*) \qquad |f(y) - f(x)| \leq L|y - x|$$

Der Beweis folgt unmittelbar aus dem Schrankensatz oder direkt aus dem MWSD. Zum Beispiel ist also $\sin : \mathbb{R} \to \mathbb{R}$ Lipschitz-stetig (und damit gleichmäßig stetig auf ganz $\mathbb{R}$), dann gilt wegen $\sin'(\xi) = \cos(\xi)$ und $|\cos \xi| \leq 1$ gilt für alle $x, y \in \mathbb{R}$

$$|\sin y - \sin x| \leq 1 \cdot |y - x|$$

mit der Lipschitz-Konstante $L = 1$.
Die Ungleichung $(*)$ gilt auch für komplexwertige (differenzierbare) Funktionen $f : D \to \mathbb{C}$ ($D \subset \mathbb{R}$ Intervall) obwohl für solche Funktionen der Satz von Rolle bzw. der MWSD i.A. nicht gilt. Weitere Folgerungen aus dem MWSD bzw. dem Schrankensatz sind:

19.1.13 Satz (Charakterisierung konstanter Funktionen und von Stammfunktionen)

(a) Eine differenzierbare Funktion $f : D \to \mathbb{R}$ ($D \subset \mathbb{R}$ echtes Intervall) ist genau dann konstant, wenn $f'(x) = 0$ für alle inneren Punkte $x \in D$ gilt.

(b) Ist $F : D \to \mathbb{R}$ *eine* Stammfunktion von $f : D \to \mathbb{R}$, d.h. ist F differenzierbar und gilt $F'(x) = f(x)$ für alle $x \in D$, dann ist $\{F + c;\ c \in \mathbb{R}\}$ die Menge *aller* Stammfunktionen von f auf dem Intervall D.

Die eine Richtung ist evident: Ist $f(x) = const.$ für alle $x \in D$, dann ist $f'(x) = 0$ für alle $x \in D$, insbesondere für alle inneren Punkte.
Die Umkehrung folgt aus dem Schrankensatz mit $m = M = 0$ oder direkt aus dem MWSD (Ist $x_0 \in D$ ein fester Punkt, dann hat man $f(x) = f(x_0)$ zu zeigen; Man kann auf $[x_0, x]$, $[\bar{x}_0, x]$ bzw. $[\bar{x}, x_0]$ den MWSD anwenden).
Zum Beweis von (b) bemerken wir, dass natürlich $G = F + c$, c konstant, ebenfalls eine Stammfunktion ist. Die Umkehrung folgt aus (a), den G und F dieselbe Ableitung haben, ist $(G - F)' = 0$ gilt und damit $G = F + c$ gelten muss.

Die Funktionen

$$\begin{array}{rcl} f_c : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \exp(cx) \quad (c \in \mathbb{R}) \end{array}$$

erfüllen die Gleichung $f_c' = cf_c$. Hiervon gilt eine gewisse Umkehrung.

19.1.14 Folgerung (Charakterisierung von exp durch eine Differentialgleichung)

Ist $c \in \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion mit $f'(x) = cf(x)$ für alle $x \in \mathbb{R}$, dann gilt für alle $x \in \mathbb{R}$

$$f(x) = A\exp(cx) \ \text{ mit } \ A := f(0).$$

Insbesondere ist $\exp : \mathbb{R} \to \mathbb{R}$ die eindeutig bestimmte Lösung des Anfangswertproblems $f' = f$ und $f(0) = 1$.

19.1.15 Monotoniekriterien

Ist $D \subset \mathbb{R}$ ein (echtes) Intervall und $f : D \to \mathbb{R}$ eine differenzierbare Funktion. Dann gilt

(a) f ist *genau dann* monoton wachsend (bzw. monoton fallend) auf D, wenn $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$) für alle inneren Punkte $x \in D$ gilt.

(b) Wenn $f'(x) > 0$ (bzw. $f'(x) < 0$) ist für alle inneren Punkte $x \in D$, dann ist f sogar streng monoton wachsend (bzw. fallend) auf D.

Der **Beweis** ergibt sich in der einen Richtung („$\Longleftarrow$") sofort aus dem MWSD: Sind $x, y \in D$ und gilt $x < y$, dann gibt es ein $\xi \in]x, y[$ (ξ ist dann innerer Punkt von D) mit $f(y) - f(x) = f'(\xi)(y - x) \geq 0$, also ist

$$f(x) \leq f(y).$$

$\Longrightarrow$ Ist $x \in M$ fest, dann ist für alle $y \in D,\ y \neq x$ stets der Differenzenquotient $\frac{f(y)-f(x)}{y-x} \geq 0$, also auch $f'(x) \geq 0$.
Ist sogar $f'(x) > 0$ für alle inneren Punkte $x \in D$, dann gilt auch $f'(\xi) > 0$, also folgt $f(y) - f(x) = f'(\xi)(y - x) > 0$.

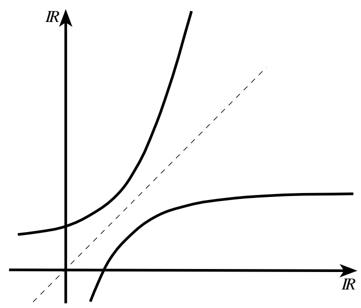
19.1.16 Bemerkungen und Beispiele

(a) Die Potenzfunktion

$$\begin{array}{rcl} p_3 : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 \end{array}$$

ist streng monoton wachsend, obwohl $p_3'(0) = 0$ gilt.
Die Positivität der Ableitung ist also lediglich ein *hinreichendes* Kriterium für die strenge Monotonie.

(b) Wegen $\exp'(x) = \exp(x) > 0$ für alle $x \in \mathbb{R}$ ist $\exp : \mathbb{R} \to \mathbb{R}$ streng monoton wachsend;
Wegen $\log'(x) = \frac{1}{x} > 0$ für $x > 0$ ist auch $\log : \mathbb{R}_+ \to \mathbb{R}$ streng monoton wachsend.



(c) sin ist streng monoton wachsend in $[-\frac{\pi}{2}, \frac{\pi}{2}]$;
cos ist streng monoton fallend in $[0, \pi]$;
tan ist streng monoton wachsend in $]-\frac{\pi}{2}, \frac{\pi}{2}[$;
cot ist streng monoton fallend in $]0, \pi[$.

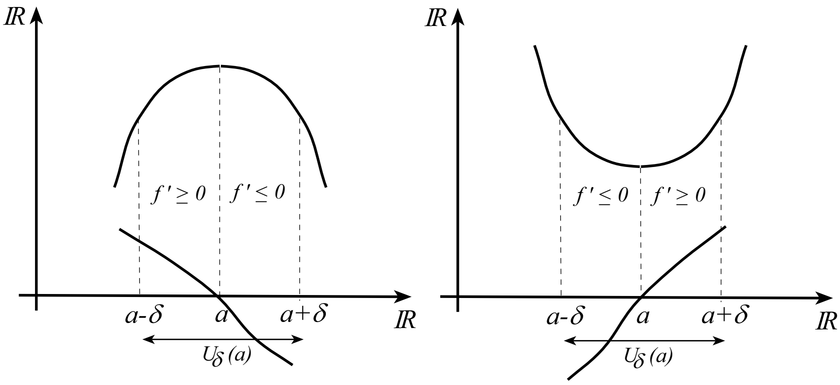
Dort haben diese (auf diese Intervalle eingeschränkte) Funktionen Umkehrfunktionen, die im gleichen Sinne wie die Funktionen monoton sind (vgl. ???)

Eine weitere Anwendung des Monotoniekriteriums sind *hinreichende Bedingungen* für Extremwerte. Wir wissen, dass das Verschwinden der Ableitung in einem inneren Punkt a des Definitionsbereichs für eine in a differenzierbare Funktion ein *notwendiges* Kriterium für ein lokales Extremum ist, wir wissen auch - das obige Beispiel von $p_3 : \mathbb{R} \to \mathbb{R},\ x \mapsto x^3$ und $a = 0$, zeigt, dass i.A. es nicht hinreichend ist.

19.1.17 Satz (hinreichende Kriterien für lokale Extrema)

Ist $D = U_\delta(a),\ \delta > 0$, eine δ-Umgebung von $a \in \mathbb{R}$ und $f : D \to \mathbb{R}$ differenzierbar und $f'(a) = 0$. Dann gilt

(a) Wechselt f' beim Durchgang durch a das Vorzeichen, das heißt gilt für alle $x \in U_\delta(a)$ $(x-a)f'(x) \geq 0$ (bzw. ≤ 0), dann hat f in a ein lokales Minimum (bzw. Maximum)



a ist die einzige Minimal bzw. Maximalstelle von f in $U_\delta(a)$, wenn a die einzige Nullstelle von f' in $U_\delta(a)$ ist.

(b) Ist f in a zweimal differenzierbar und gilt $f''(a) > 0$ (bzw. $f''(a) < 0$), dann hat f in a ein (sogar striktes) lokales Minimum (bzw. Maximum).

Beweis von (a): Der Beweis ergibt such unmittelbar aus dem Monotoniekriterium:
Gilt etwa $(x-a)f'(x) \leq 0$, dann ist f in $]a-\delta, a]$ monoton fallend und in $[a, a+\delta]$ monoton wachsend. f muss also in a ein lokales Minimum haben. Im Fall $(x-a)f'(x) \geq 0$ ist f in $]a-\delta, a]$ monoton fallend und in $[a, a+\delta]$ monoton wachsend, bei a liegt also ein lokales Maximum.

□

Beweis von (b): Wir behandeln exemplarisch den Fall $f''(a) > 0$ (man gehe nicht von f zu $-f$ über).
Wegen

$$f''(a) = \lim_{\substack{x \to a \\ x \neq a}} \frac{f'(x) - f'(a)}{x-a} > 0$$

gibt es ein δ^* $(0 < \delta' \leq \delta)$, so dass $\frac{f'(x)-f'(a)}{x-a} > 0$ gilt für alle $x \in U_\delta(a)$ mit $0 < |x-a| < \delta^*$. Wegen $f'(a) = 0$ folgt hieraus $f'(x) < 0$ für $a - \delta^* < x < a$ und $f'(x) > 0$ für $a < x < a + \delta^*$; f ist also in $]a-\delta^*, a[$ streng monoton fallend und in $]a, a+\delta^*[$ streng monoton wachsend. f besitzt also in a ein striktes lokales Minimum.

□

Bemerkung zum letzten Beweis:
Hätten wir vorausgesetzt, dass sogar $f \in \mathcal{C}^2(D)$ gilt, dann hätten man aus der Stetigkeit von f'

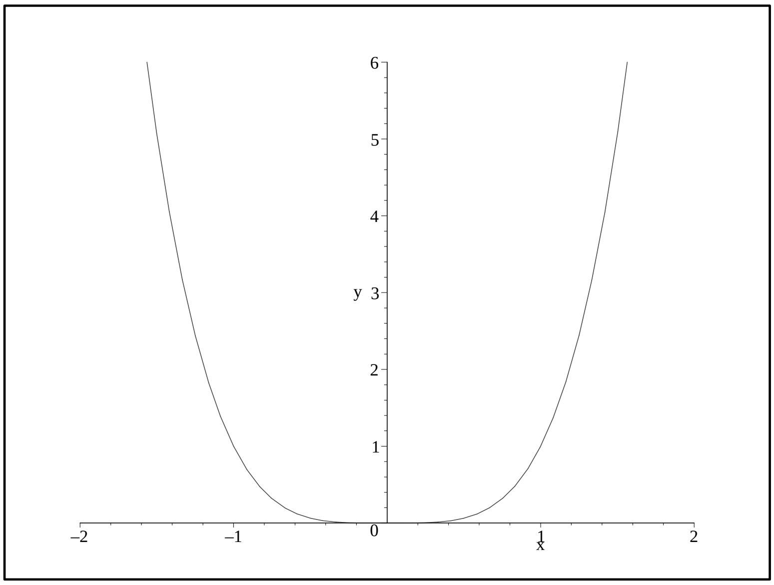


Abbildung 20: Graph der Funktion x^4

auch direkt auf die Existenz eines weiteren δ^* schließen können.

19.1.18 Beispiele und Bemerkungen

(a) Das einfache Beispiel

$$\begin{array}{rcl} p_4 : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^4 \end{array}$$

mit $a = 0$ zeigt, dass das Kriterium (b) nicht notwendig für das Vorliegen eines lokalen Extremum ist.
Es gilt $p_4'(0) = 0$ und p_4 hat in $x = 0$ ein lokales Minimum. (siehe Abb. 20)

Zur Entscheidung kann man in einem solchen Fall die höheren Ableitungen herausziehen: Offensichtlich gilt $p_4'(x) = 4x^3,\ p_4''(x) = 12x^2,\ p_4''' = 24x$ und $p_4^{(4)} = 24$, insbesondere $p_4^{(4)} = 24 > 0$.

(b) Ob ein Extremum vorliegt, kann man häufig mit Hilfe der Taylorschen Formel (vgl. ???) entscheiden. Im konkreten Fall ist offensichtlich, dass an der Stelle $a < 0$ sogar ein globales Minimum von p_4 vorliegt.

Das aus der Schulmathematik bekannte Verfahren zur Extremwertbestimmung findet also durch Satz 19.1.17 eine Rechtfertigung. Ob eine gegebene Funktion an einer Stelle $a \in D$ ein lokales Extremum hat, ist häufig ein schwer zu entscheidendes Problem. Hat man nur notwendige Kriterien zur Verfügung, so ergibt sich häufig aus der Fragestellung selbst oder auf Grund der Herkunft aus einem naturwissenschaftlichen oder geometrischen Problem oder einem Problem aus einem anderen Anwendungsgebiet die Möglichkeit zu entscheiden, welche nach den notwendigen Bedingungen in Frage kommende Stellen die gesuchten Extremalstellen sind. Eine solche Überlegung ist häufig aufschlussreicher, als die formale Anwendung hinreichender Kriterien. Dennoch geben wir ein paar Beispiele für diese, um die Bandbreite der Anwendungen auszudehnen.

(c) Wir bemerken noch eine interessante Eigenschaft der Ableitung einer differenzierbaren Funktion f auf einem Intervall (in der Literatur häufig als **Satz von Darboux** zitiert:
Ist $f[a,b] \to \mathbb{R}$ differenzierbar und ist c eine Zahl mit

$$f'(a) < c < f'(b) \text{ oder } f'(b) < c < f'(a),$$

dann gibt es ein $\xi \in]a,b[$ mit $f'(\xi) = c$.

Zum **Beweis** betrachte man

$$\begin{array}{rcl} h : [a,b] & \to & \mathbb{R} \\ x & \mapsto & f(x) - cx \end{array}$$

Die Ableitung einer differenzierbaren Funktion hat also die „Zwischenwerteigenschaft“. Da es differenzierbare, aber nicht stetig differenzierbare Funktionen gibt (vgl. ???), zeigt dies nochmals, dass die Zwischenwerteigenschaft die stetigen Funktionen nicht charakterisiert. Außerdem zeigt der Satz von Darboux, dass eine Funktion $f : D \to \mathbb{R}$ höchstens dann eine Stammfunktion F haben kann, wenn $f(= F')$ die Zwischenwerteigenschaften hat.

19.1.19 Beispiele für Extremwertaufgaben (wird nachgetragen)

19.1.20 Konvexe Funktionen

Ist eine Funktion $f : D \to \mathbb{R}$ zweimal differenzierbar, dann gibt die zweite Ableitung die Änderung der Tangenrichtungen an, beschreibt also wie sich der Graph von f „krümmt“.
Eine präzise Definition des Begriffs „Krümmung“ geben wir erst später.
Völlig ohne Differezierbarkeitsvoraussetzungen kann man den Begriff der *konvexen Funktionen* einführen, ist f allerdings zusätzlich zweimal differenzierbar, so ergibt sich ein einfaches Konvexitätskriterium über die zweite Ableitung ($f''(x) \geq 0$ für alle $x \in D$).
Zentrale Ungleichungen der Analysis beruhen auf der Konvexität einfacher Funktionen.

19.1.21 Definition (konvexe und konkave Funktion)

Ist $D \subset \mathbb{R}$ ein echtes Intervall:
Eine Funktion $f : D \to \mathbb{R}$ heißt *konvex*, wenn für je zwei Punkte $x_1, x_2 \in D$ und alle $\lambda \in [0,1]$ gilt

$$\boxed{f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2)}$$

f heißt *konkav* $\iff -f$ konvex ist.

Die Fälle $\lambda = 0$ oder $\lambda = 1$ können wir dabei unberücksichtigt lassen, da die obige Ungleichung für diese Werte automatisch erfüllt ist. Auch der Fall $x_1 = x_2$ ist ohne Bedeutung, so dass wir oBdA immer $x_1 < x_2$ und $\lambda \in]0,1[$ annehmen können und sollen.

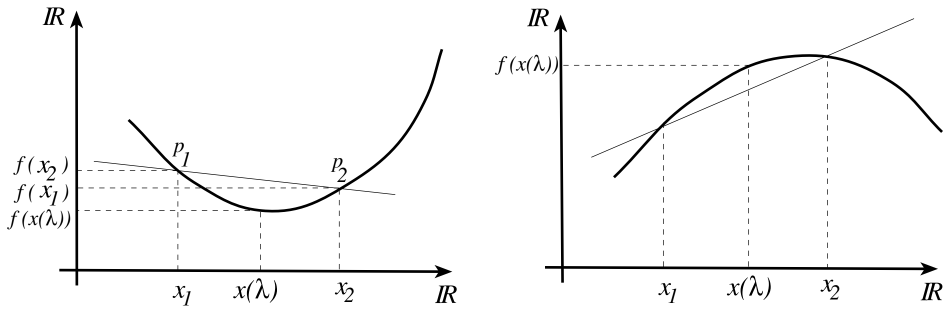
19.1.22 Geometrische Interpretation der Konvexität

Ist $x_1 < x_2$, dann durchläuft der Punkt $x(\lambda) = (1-\lambda)x_1 + \lambda x_2 = x_1 + \lambda(x_2 - x_1)$ gerade das Intervall $]x_1, x_2[$, wenn λ das Intervall $]0,1[$ durchläuft. Durchläuft λ das Intervall $]0,1[$, so durchläuft wegen

$$x(\lambda) = (1-\lambda)f(x_1) + \lambda f(x_2) = f(x_1) + \lambda(f(x_2) - f(x_1)) = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x(\lambda) - x_1)$$

der Punkt $(x(\lambda), f(x))$ gerade das Geradenstück s zwischen den Punkten $P_1 = (x_1, f(x_1))$ und $P_2 = (x_2, f(x_2))$ auf dem Graphen von f. s ist ein Stück der Sekanten durch die Punkte P_1 und P_2.

Konvexität von f bedeutet also, dass der Graph von f zwischen zwei Punkten x_1 und x_2 immer unterhalb der Sehne (Sekante) s verläuft. Konkavität bedeutet, dass s immer *unterhalb* der Graphen von f liegt.



Man könnte etwas unpräzise, aber anschaulich sagen:
f konvex bedeutet, dass der Graph von f mit wachsenden Argumenten eine „Linkskurve“ beschreibt (stellen Sie sich vor, der Graph von f ist in einer Ebene irgendwie markiert und Sie fahren auf der Markierung mit dem Fahrrad entlang).

Übungsaufgabe:
Ist $f : D \to \mathbb{R}$ konvex und besitzt f in $x_0 \in D$ ein *lokales* Minimum, so ist dieses auch *globales* Minimum für f.

Konvexität bzw. Konkavität lässt sich bei *differenzierbaren* Funktionen mit Hilfe der zweiten Ableitung charakterisieren:

19.1.23 Satz

Ist $f : D \to \mathbb{R}$ zweimal differenzierbar, dann ist f genau dann konvex (bzw. konkav), wenn $f''(x) \geq 0$ (bzw. $f''(x) \leq 0$) für alle inneren Punkte $x \in D$ gilt.

Beweis : „$\Longleftarrow$“

Sei $f''(x) \geq 0$ für alle inneren Punkt $x \in D$. Dann ist nach 19.1.15 $f| \ : D \to \mathbb{R}$ monoton wachsend.
Seien $x_1, x_2 \in D$ und oBdA $x_1 < x_2$ und $\lambda \in]0,1[$ und $x = x(\lambda) = (1-\lambda)x_1 + \lambda x_2$. Es ist dann $x_1 < x < x_2$. Wir wenden den MWSD auf die Intervalle $[x_1, x]$ und $[x, x_2]$ an:
Es existieren daher $\xi_1 \in]x_1, x[$ und $\xi_2 \in]x, x_2[$ mit

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) \leq f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}$$

(die mittlere Ungleichung gilt, da $f|$ monoton wächst). Da nun

$$x - x_1 = \lambda(x_2 - x_1) \text{ und } x_2 - x = (1-\lambda)(x_2 - x_1)$$

gilt, folgt

$$\frac{f(x)-f(x_1)}{\lambda} \leq \frac{f(x_2)-f(x)}{1-\lambda}$$

oder auch

$$f(x)=f(x(\lambda)) \leq (1-\lambda)f(x_1)+\lambda f(x_2),$$

d.h.: f ist konvex.

„$\Longrightarrow$“
Sei umgekehrt $f : D \to \mathbb{R}$ konvex.
Wir führen einen indirekten Beweis und nehmen an, dass *nicht* $f''(x) \geq 0$ für alle inneren Punkte $x \in D$ gilt. Dann gibt es einen inneren Punkt $x_0 \in D$, für den $f''(x_0) \geq 0$ gilt. Sei $l := f'(x_0)$ und $\varphi : D \to \mathbb{R}$ definiert durch $x \mapsto f(x) - l(x-x_0)$.
Dann ist φ zweimal differenzierbar und $\varphi(x_0)=0$ und $\varphi''(x_0)=f''(x_0)<0$. φ hat also nach 19.1.17(b) in x_0 ein striktes lokales Maximum. Es gibt also ein $h>0$, so dass $[x_0-h, x_0+h] \subset D$ und $\varphi(x_0-h)<\varphi(x_0)$ und $\varphi(x_0+h)<\varphi(x_0)$ gilt. Hieraus folgt aber

$$f(x_0)=\varphi(x_0) > \frac{1}{2}(\varphi(x_0-h)+\varphi(x_0+h)) = \frac{1}{2}(f(x_0-h)+f(x_0+h))$$

Setzt man nun

$$x_1 := x_0-h,\ x_2 := x_0+h, \lambda := \frac{1}{2},$$

so ist $x_=(1-\lambda)x_1+\lambda x_2$, also

$$f((1-\lambda)x_1+\lambda x_2) > (1-\lambda)f(x_1)+\lambda f(x_2).$$

Diese Ungleichung steht aber im Widerspruch zur vorausgesetzten Konvexität von f.
Also war unsere Annahme falsch. Wenn f konvex ist, gilt daher $f''(x)$ für alle inneren Punkte $x \in D$.

$\square$

Eine einfache, aber wichtige Anwendung der Konvexität von exp ist die folgende Ungleichung.

19.1.24 Satz (Youngsche Ungleichung)

Seien $p,q \in]1,\infty[$ mit $\frac{1}{p}+\frac{1}{q}=1$.
Dann gilt für alle $a,b \in \mathbb{R}_+$ die Unlgeichung

$$\boxed{a \cdot b \leq \frac{a^p}{p}+\frac{b^q}{q}}$$

Der Beweis ergibt sich einfach für $a,b>0$. (für $ab=0$ ist die Ungleichung richtig) Mit $x := \log a$ und $y := \log b$ folgt wegen der Konvexität von exp

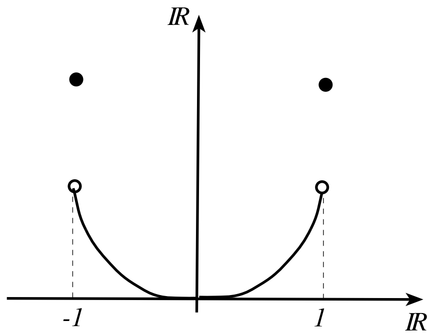
$$\begin{aligned} ab=\exp(x)\exp(y) &= \exp(x+y) \\ &= \exp(\frac{1}{p}px+\frac{1}{q}qy) \\ &\leq \frac{1}{p}\exp(px)+\frac{1}{q}(\exp qy) \\ &= \frac{ap}{p}+\frac{bq}{q} \end{aligned}$$

(man verwende die Konvexitätsbedingung mit $1-\lambda=\frac{1}{p}$ und $\lambda=\frac{1}{q}$).
Die Youngsche Ungleichung ist die *Quelle für viele wichtige Ungleichungen* etwa die *Höldersche Ungleichung* für die p-Normen im $\mathbb{R}^n$ oder $\mathbb{C}^n$ als Spezialfall die CSU für das Standardskalarprodukt im $\mathbb{K}^n$ die Minkowskische Ungleichung in $\mathbb{K}^n$ und auch die entsprechende Ungleichungen für Integrale.

Zum Schluss sei bemerkt, dass die Konvexität eine Funktion eine sehr starke Eigenschaft ist. Ist $f : D \to \mathbb{R}$ konvex, dann ist f in allen inneren Punkten stetig. (man kann sogar zeigen, dass für alle inneren Punkte $x \in D$ f links- und rechtsseitig differenzierbar und, dass $f'_-(x) \leq f'_+(x)$ gilt.
In Randpunkten gilt dies i.A. nicht, wie das einfache Beispiel $f : [-1,1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 2, & \text{falls } x=-1 \text{ oder } x=1; \\ x^2, & \text{falls } x \in]-1,1[\end{cases}$$

zeigt.



19.1.25 Beispiel (Wendepunkte)

Sei $f : D \to \mathbb{R}$ eine Funktion (auf einem echten Intervall D). Dann heißt $a \in D$ Wendepunkt von f, wenn a innerer Punkt von D ist und ein $\delta > 0$ existiert, so dass f auf $[a-\delta, a]$ konvex und auf $[a, a+\delta]$ konkav ist oder umgekehrt. Ist f in a differenzierbar, dann heißt die Tangente $x \mapsto f(a) + f'(a)(x-a)$ an den Graphen von f im Punkt $(a, f(a))$ *Wendepunkt* von f in a. Sie durchsetzt den Graphen im Punkt $(a, f(a))$.

ZEICHNUNG

Ist a innerer Punkt und in einer Umgebung von a differenzierbar und im Punkt a zweimal differenzierbar, dann ist $f''(a)=0$ eine notwendige Bedingung für das Vorliegen eines Wendepunkts von f in a, da f' im Wendepunkt dann ein lokales Extremum hat. Ein hinreichendes Kriterium enthält

19.1.26 Satz

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $a \in D$ ein innerer Punkt. Ferner sei $n \geq 3$ ungerade. Ist dann $f^{(n-1)}$ in a differenzierbar und gilt

$$f^{(3)}(a) = \ldots = f^{(n-1)} = 0,\ f^{(n)}(a) \neq 0,$$

so hat f in a einen Wendepunkt.

Beweis : f'' wechselt beim Durchgang durch a das Vorzeichen.

□

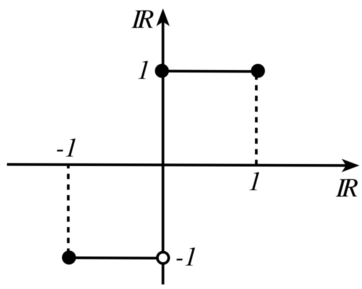
20 Zusammenhang zwischen Integral- und Differenzialrechnung: Der Hauptsatz und seine Anwendungen

In diesem Abschnitt zeigen wir, dass Integralrechnung und Differentialrechnung, die aus völlig verschiedenen Motiven entwickelt wurden (etwa Flächeninhaltsproblem bzw. Tangentenproblem) über den sog. Hauptsatz der Differential- und Integralrechnung eng miteinander verquickt sind. Wir werden auch sehen, wie man das Umkehrproblem der Differentialrechnung, d.h. zu einer vorgegebenen Funktion $f : D \to \mathbb{R}$(oder $\mathbb{C}$) eine *Stammfunktion* zu finden, d.h. eine differenzierbare Funktion $F : D \to \mathbb{R}$(oder $\mathbb{C}$) mit $F' = f$, zumindest für *stetige* F auf Intervallen lösen kann. Der Hauptsatz wird uns auch schlagkräftige *Methoden zur Integralberechnung* zur Verfügung stellen.
Dass Integration und Differentiation Umkehroperationen von einander sind, kommt insbesondere in der *algebraischen Formulierung* des Hauptsatzes zum Ausdruck. Schließlich werden wir und mit einigen Anwendungen des Hauptsatzes beschäftigen.

20.1 Der Hauptsatz der Differential- und Integralrechnung

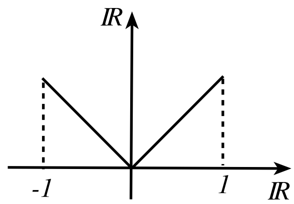
Ist $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$(oder $\mathbb{C}$) eine Funktion, dann ist bis jetzt nicht klar, welche Eigenschaften f erfüllen muss, damit es eine differenzierbare Funktion $F : D \to \mathbb{R}$(oder $\mathbb{C}$) mit $F' = f$ gibt (Problem der Existenz einer Stammfunktion). Das „abschreckende Beispiel“ aus ??? zeigt, dass auch unstetige Funktionen Stammfunktionen haben können. Natürlich kann man in konkreten Fällen mit den Rechenregeln für Ableitungen sofort bestätigen, dass eine Funktion F Stammfunktion einer Funktion f ist.
Wählt man z.B. $D = \mathbb{R}$ und $f := \cos$ und $F := \sin$, dann ist $F' = \sin' = \cos$, also ist der Sinus (und zwar auf ganz $\mathbb{R}$) eine Stammfunktion von $\cos$. Sicherlich gibt es nicht für jede Funktion $f : D \to \mathbb{R}$ eine Stammfunktion, denn nach dem Satz von Darboux (vergl. ???), muss f jedes Teilintervall $D_0 \subset D$ wieder auf ein Intervall abbilden. So kann etwa die Treppenfunktion $t : [-1, 1] \to \mathbb{R}$ mit

$$t(x) = \begin{cases} 1, & \text{für } 0 \leq x \leq 1; \\ -1, & \text{für } -1 \leq x < 0 \end{cases}$$



keine Stammfunktion haben (obwohl t als Treppenfunktion und damit als Regelfunktion) natürlich integrierbar ist (das Integral ist Null).
Ein Kandidat für eine Stammfunktion wäre

$$\begin{aligned} F : [-1, 1] &\to \mathbb{R} \\ x &\mapsto |x|, \end{aligned}$$



also die Einschränkung der Betrags auf $[-1, 1]$, aber F ist in 0 nicht differenzierbar.
Fasst man den Stammfunktionsbegriff allgemeiner (vergl. Königsberger, Analysis 1, 5.Auflage., Seite ???), lässt also abzählbare Ausnahmemengen $A \subset D$ zu, in denen F nicht differenzierbar zu ein braucht und fordert von einer Stammfunktion $F : D \to \mathbb{R}$ die Eigenschaften

(a) F ist stetig;
(b) In $D - A$ ist F differenzierbar und für $x \in D - A$ gilt $F'(x) = f(x)$;

dann wäre F mit $F(x) = |x|$ für $-1 \leq x \leq 1$ eine Stammfunktion von f auf $[-1, 1]$. Nach dem Zwischenwertsatz haben aber stetige reellwertige Funktionen auf Intervallen die Zwischenwerteigenschaft („das stetige Bild eines Intervalls ist eine Intervall“), wenn wir daher f als stetig voraussetzen, widerspricht der Satz von Dauboux auf jeden Fall nicht der möglichen Existenz einer Stammfunktion. Der Hauptsatz besagt u.a., dass die Stetigkeit auch hinreichend für Existenz einer Stammfunktion ist und dass umgekehrt einer Regelfunktion f, die auf einem kompakten Intervall $[a, b]$ eine Stammfunktion besitzt, auch stetig ist. Wir wollen folgende Vereinbarungen treffen:

Ist $D \subset \mathbb{R}$ ein echtes Intervall, und $f : D \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$) eine Funktion, dann heißt f Regelfunktion, wenn die Einschränkung von f auf *jedes* kompakte Teilintervall $[a, b] \subset D$ eine Regelfunktion (im bisherigen Sinne) ist.

20.1.1 Theorem (Hauptsatz der Differential- und Integralrechnung, 1. Version)

(a) Ist $D \subset \mathbb{R}$ ein echtes Intervall und $f : D \to \mathbb{R}$(oder $\mathbb{C}$) eine Regelfunktion, dann besitzt f genau dann eine Stammfunktion, wenn f stetig ist.
Ist f stetig, $a \in D$, so ist *eine* Stammfunktion $F : D \to \mathbb{R}$ gegeben durch die Integralfunktion $x \mapsto \int_a^x f(t)dt$.

Zusatz:
$\{F + c;\ c \in \mathbb{R}\}$ ist dann die Menge *aller* Stammfunktionen von f.

(b) Ist G eine *beliebige* Stammfunktion von $f \in C([a, b])$, dann gilt

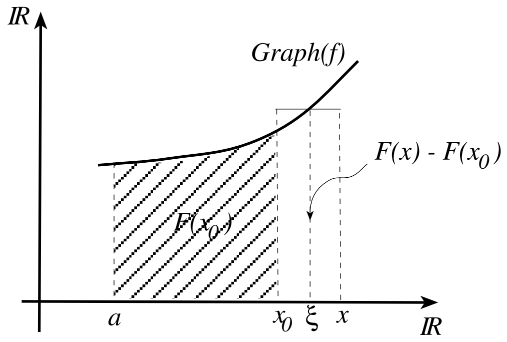
$$\int_a^b f(t)dt = G(b) - G(a)$$

Beweis: (a) Sei zunächst f stetig in $x_0 \in D$ und $F(x) := \int_a^x f(t)dt$.

Wir zeigen, dass F in x_0 differenzierbar ist und $F'(x) = f(x_0)$ gilt. Wenn f in jedem Punkt stetig ist, ist also gezeigt, dass f eine Stammfunktion auf D besitzt, nämlich F.
Wir geben zwei Beweisvarianten:

Bei der ersten wird der 1. Mittelwertsatz der Integralrechnung (vgl. ???) verwendet:
Nach diesem gibt es ein ξ in $[b, x]$ bzw. $[x, b]$ (m.a.W. $\xi = x_0 + \vartheta(x - x_0),\ 0 \leq \vartheta \leq 1$)

$$F(x) - F(x_0) = \int_a^x f(t)dt - \int_a^{x_0} f(t)dt = \int_{x_0}^x f(t)dt = f(\xi)(x - x_0)$$



Für $x \neq x_0$ folgt

$$\frac{F(x)-F(x_0)}{x-x_0} = f(\xi).$$

Für $x \to x_0$ folgt dabei wegen der Stetigkeit von f in x_0

$$F'(x_0) = \lim_{x \to x_0} \frac{F(x)-F(x_0)}{x-x_0} = f(x_0)$$

Wegen der Wichtigkeit des Satzes geben wir einen zweiten Beweis ($\varepsilon - \delta-$Beweis):
Dabei schreiben wir die Differenz $F(x) - F(x_0)$ in der Form

$$\begin{aligned} \int_{x_0}^{x} f(t)dt &= \int_{x_0}^{x} f(x_0)dt + \int_{x_0}^{x} (f(t) - f(x_0))dt \\ & \qquad \uparrow \text{ man beachte: der Integrand } x_0 \text{ ist konstant} \\ &= (x-x_0)f(x_0) + (x-x_0)\varrho(x) \end{aligned}$$

mit

$$\varrho(x) := \begin{cases} \dfrac{\int_{x_0}^{x}(f(t)-f(x_0))dt}{x-x_0}, & \text{falls } x \neq x_0; \\ 0, & \text{falls } x = x_0. \end{cases}$$

Damit haben wir also

$$F(x) - F(x_0) = (x-x_0)f(x_0) + (x-x_0)\varrho(x).$$

Nach dem Äquivalenzsatz über Differenzierbarkeit müssen wir lediglich noch nachweisen, dass ϱ an der Stelle x_0 stetig ist (denn dann ist $F'(x_0) = f(x_0)$).
Sei $\varepsilon > 0$ beliebig vorgegeben. Wegen der Stetigkeit von f in x_0 gibt es dazu ein $\delta > 0$, mit

$$|f(t) - f(x_0)| < \varepsilon|$$

für alle $t \in D$ mit $|t - x_0| < \delta$.
Dann gilt zunächst für $x > x_0$

$$\left| \int_{x_0}^{x} (f(t) - f(x_0))dt \right| \leq \int_{x_0}^{x} |f(t) - f(x_0)|dt \leq \int_{x_0}^{x} \varepsilon \, dt = \varepsilon|x - x_0|$$

Diese Abschätzung funktioniert genauso, wenn $x < x_0$ sein sollte.
Insgesamt folgt also $|\varrho(x)| \leq \varepsilon$ für alle $x \in D$ mit $|x - x_0| \leq \delta$. ϱ ist also stetig in x_0, damit gilt $F'(x_0) = f(x_0)$.

Damit haben wir also gezeigt:

Jede stetige Funktion auf einem Intervall besitzt eine Stammfunktion.

Jetzt sei umgekehrt $f : D \to \mathbb{R}$ eine Regelfunktion, die eine Stammfunktion F besitzt. Wir zeigen die Stetigkeit von f in $x_0 \in D$ in folgender Weise:
Da f eine Regelfunktion ist, existieren in jedem inneren Punkt $x_0 \in D$ die links- und rechtsseitigen Grenzwerte $f(x_{0_-})$ und $f(x_{0_+})$ (in den Randpunkten nur die einseitigen).
Sei (x_n) eine Folge mit $x_n \in D$ und $x_n > x_0$ und $\lim\limits_{n\to\infty} x_n = x_0$. Nach dem MWSD gibt es dann ein $\xi \in]x_0, x_n[$ mit

$$\frac{F(x) - F(x_0)}{x - x_0} = F'(\xi_n) = f(\xi_n),\ x_0 < \xi_n < x_n.$$

Die linke Seite konvergiert nach Voraussetzung gegen $F'(x_0) = f(x_0)$, die rechte Seite aber gegen $f(x_{0_+})$, also folgt

$$f(x_0) = f(x_{0_+}).$$

Analog zeigt man $f(x_0) = f(x_{0_-})$ (falls x_0 nicht linker Randpunkt von D ist.
Insgesamt gilt also: f ist stetig in x_0. Da wir schon wissen (vgl. ???), dass wenn eine Funktion $f : D \to \mathbb{K}$ eine Stammfunktion F besitzt, eine weitere Funktion $G : D \to \mathbb{K}$ *genau dann* auch Stammfunktion von f ist, wenn $G = F + c$ mit irgendeiner Konstanten Funktion gilt, ist die Aussage über die Menge aller Stammfunktionen klar.

(b) Nach (a) ist für stetige $f : [a, b] \to \mathbb{K}$ die Integralfunktion

$$\begin{aligned} F : [a, b] &\to \mathbb{K} \\ x &\mapsto \int_a^x f(t)dt \end{aligned}$$

eine Stammfunktion von f mit $F(a) = \int_a^a f = 0$. Für eine beliebige Stammfunktion G von f gilt nach (a) dann $G = F + c$ mit einer Konstanten c und damit

$$\begin{aligned} G(b) - G(a) &= (F(b) + c) - (F(a) + c) \\ &= F(b) - F(a) \\ &= F(b) = \int_a^b f. \end{aligned}$$

□

20.1.2 Bemerkungen

(a) Der Hauptsatz garantiert die *Existenz* von Stammfunktionen zu allen stetigen Funktionen auf Intervallen. Darin liegt seine große theoretische Bedeutung. Für konkrete Funktionen kann es jedoch *schwierig* oder *sogar unmöglich* sein, Stammfunktionen mit Hilfe bereits bekannter Funktionen auszudrücken.

(b) Diese Phänomen kann man schon an folgenden Beispielen sehen:
Betrachtet man

$$\begin{aligned} j :]0, \infty[&\to \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned}$$

dann ist der Logarithmus

$$L := \log :]0, \infty[\to \mathbb{R}$$

eine Stammfunktion von j:

$$L'(x) = \frac{1}{x} = j(x)$$

für alle $x \in]0, \infty[$. Man beachte hier zum einen: j ist eine rationale Funktion, L aber nicht (warum?). Der Übergang zu Stammfunktionen kann also durchaus aus einer Klasse von Funktionen herausführen. Zum anderen: Wenn wir den Logarithmus nicht schon gekannt hätten, der ja als Umkehrung von exp definiert wurde, hätten wir keine Stammfunktion explizit angeben können.
Wie in Aufgabe ??? von Blatt ??? zu sehen ist, könnte man sogar umgekehrt $L := \log :]0, \infty[\to \mathbb{R}$ durch $x \mapsto L(x) = \int\limits_1^x \frac{1}{t} dt$ definieren, Eigenschaften von L beweisen (Differenzierbarkeit, strenge Monotonie, Funktionalgleichung) und dann exp als Umkehrfunktion von L definieren.
Diesen Weg in der Schulmathematik einzuschlagen, wurde von Felix Klein angeregt. Nach den gängigen Lehrplänen ist dies aber kaum möglich, da die Exponentialfunktion (Wachstum und Zerfallsprozesse) sogar auch in anderen Fächern (Physik, Biologie) benötigt wird, zu einem Zeitpunkt also, wo der Integralbegriff noch nicht zur Verfügung steht.
Nach den Rechenregeln für Ableitungen hat arcsin in $] - 1, 1[$ die Ableitung

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

Auch hier könnte man umgekehrt vorgehen und für $x \in] - 1, 1[$

$$a(x) := arcsin(x) := \int\limits_0^x \frac{1}{\sqrt{1 - t^2}} dt$$

definieren und dann den Sinus als Umkehrfunktion. Die Kreiszahl π wird in diesem Fall als $\frac{\pi}{2} := \lim\limits_{\substack{x \to 1 \\ x < 0}} a(x)$ definiert.

Dieses Vorgehen hat Vorteile: Man hat gleich die richtige Interpretation von x als Bogenlänge. Der entscheidende Nachteil ist, dass man zur Definition das Integral braucht, das man auch noch in die Punkte $x = -1$ und $x = 1$ fortsetzen muss. Außerdem ist die Periodizität von sin nicht offensichtlich: Man muss sin zu einer periodischen Funktion auf ganz $\mathbb{R}$ fortsetzen.

(c) Die große *praktische Bedeutung* des Hauptsatzes liegt darin, dass es eine äußerst elegante und bequeme Methode darstellt, Integrale zu berechnen (deren Existenz ohnehin gesichert ist), wenn man eine explizite Stammfunktion kennt: es ist ja

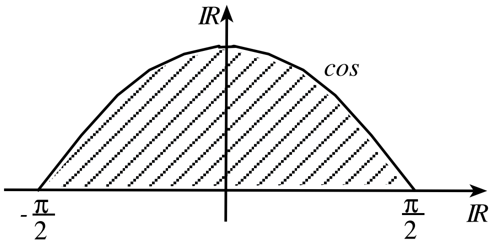
$$(*) \qquad \int\limits_a^b f(x) dx = G(b) - G(a)$$

mit irgendeiner Stammfunktion G von f.
Man führt für die rechte Seite von (x) auch die Notation

$$G(x)|_a^b \quad \text{oder} \quad [G(x)]_a^b$$

ein. So ist z.B.

$$\int\limits_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(x) dx = \sin(x)|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \sin(\pi) - \sin(-\frac{\pi}{2}) = -1 - (-1) = 2.$$



Die in ??? mit etwas Mühe berechneten Integrale fallen uns nun in den Schoss, etwa

$$\int_a^b \exp(x)dx = \exp(b) - \exp(a)$$

oder

$$\int_a^b x^\alpha \, dx = \left.\frac{x^\alpha + 1}{\alpha + 1}\right|_a^b = \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha + 1} \quad (\alpha \neq 1, \; a, b > 0)$$

(d) Für *eine* Stammfunktion F von f wird in der Literatur häufig auch das Symbol

$$\int f(x)dx$$

verwendet und „unbestimmtes Integral" von f genannt (im Gegensatz zum „bestimmten" Integral $\int_a^b f(x)dx$ mit festen Grenzen).

Man beachte aber, dass in der Literatur unter diesem Symbol auch häufig die *Menge aller Stammfunktionen* von f verstanden wird.

Aus

$$\int x^3 dx = \frac{1}{4}x^4 \text{ und } \int x^3 dx = \frac{1}{4}x^4 + 17$$

kann man jedoch nicht auf $\frac{x^4}{4} = \frac{x^4}{4} + 17$ für alle $x \in \mathbb{R}$ schließen. Also ist

$$\int f(x)dx = F(x)$$

zu lesen: F ist eine Stammfunktion von f. Der Nachteil dieser Schreibweise ist auch, dass kein Definitionsbereich angegeben ist. Aus $\int f(x)dx)F(x)$ und $\int f(x) = G(x)$ darf man lediglich $G = F + c$ schließen.

Aus jedem aus der Differentialrechnung gewonnenen Resultat, ergibt sich umgekehrt eine Aussage über Integrale (sofern die Ableitungen stetig sind). Bevor wir eine Liste von *Grundintegralen* zusammenstellen, seien noch *weitere Versionen des Hauptsatzes* aufgeführt:

Ist $f : D \to \mathbb{R}$ stetig differenzierbar, dann liefert der Hauptsatz für $x \in D$ und $x_0 \in D$

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)dt,$$

speziell für stetig differenzierbares $f : [a, b] \to \mathbb{R}$

$$f(b) = f(a) + \int_a^b f'(t)dt.$$

Man gewinnt also die Funktion f (bis auf eine Konstante) aus ihrer Ableitung („Änderungsrate") zurück.

Diese Version des Hauptsatzes wird von vielen Autoren als das wichtigere Resultat betrachtet (wegen weittragenden Anwendungen in den Naturwissenschaften).

20.1.3 Theorem (2.Version des Hauptsatzes)

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f \in C^1(D)$, dann gilt für alle $x, x_0 \in D$

$$\boxed{f(x) = f(x_0) + \int_{x_0}^{x} f'(t)dt}$$

speziell für $[a, b] \subset D$ ist

$$\boxed{f(b) = f(a) + \int_{a}^{b} f'(t)dt}$$

Wir geben noch eine weitere Formulierung des Hauptsatzes in der Sprache der Linearen Algebra, indem wir Ableitung und Integration als lineare Operatoren (lineare Abbildungen) auffassen. Diese Version zeigt überdeutlich, dass Integration und Differentiation Umkehroperationen voneinander sind.
Für ein kompaktes Intervall $[a, b] \subset \mathbb{R}$ ($a < b$) betrachten wir die $\mathbb{R}$-Vektorräume

$$V = \{f : [a, b] \to \mathbb{R},\ f \text{ stetig }\} =: C([a, b])$$

und den Untervektorraum

$$W = \{f : [a, b] \to \mathbb{R},\ f\text{stetig differenzierbar}\} =: C^1([a, b]).$$

Man definiert den Ableitungsoperator $\mathrm{D} : W \to V$ durch $f \mapsto \mathrm{D}f$ mit $(\mathrm{D}f)(x) = f'(x)$.
Nach den Rechenregeln ??? ist D ein linearer Operator (eine lineare Abbildung). Weiter definiert man

$$\begin{aligned} I : V &\to W \\ f &\mapsto If \text{ mit } (If)(x) = \int_a^x f(t)dt. \end{aligned}$$

Nach der ersten Version des Hauptsatzes liegt $F := If$ in W, ferner ist nach Grundeigenschaften des Integrals I auch eine lineare Abbildung (Integrationsoperator) und es gilt wegen $\mathrm{D}F = \mathrm{D}(If) = f$.

$$(*) \qquad \mathrm{D} \circ I = id_V.$$

$$V \underset{\mathrm{D}}{\overset{I}{\rightleftarrows}} W$$

Startet man umgekehrt mit einem $f \in W$, dann erhält man nach der 2. Version des Hauptsatzes

$$((I \circ \mathrm{D})f)(x) = (I(\mathrm{D}f))(x) = \int_a^x (\mathrm{D}f)(t)dt = \int_a^x f'(t) = f(x) - f(a).$$

Man erhält also f bis auf eine Konstante zurück. Will man den Sachverhalt, dass Integration und Differentiation wirklich Umkehroperationen voneinander sind (also Bijektionen), dann muss man den Unterraum $W_0 := \{f \in W;\ f(a) = 0\}$ von W einführen. Für $f \in W_0$ gilt dann auch

$$(**) \qquad I \circ \mathrm{D} = id_{W_0}$$

Da $\mathrm{D} \circ I = id_V$ immer gilt, haben wir die folgende algebraische Formulierung des Hauptsatzes mit Hilfe linearer Operatoren:

20.1.4 Theorem (algebraische Form des Hauptsatzes)

Die Abbildung

$$\begin{array}{rcl} I;V & \to & W_0 \\ f & \mapsto & If \end{array}$$

ist ein Isomorphismus mit der Umkehrabbildung

$$\begin{array}{rcl} \mathsf{D} : W_0 & \to & V \\ f & \mapsto & \mathsf{D}f. \end{array}$$

20.1.5 Bemerkung

Der Vektorraum $V = C([a,b])$ ist isomorph zum echten Unterraum W_0. Aus der Linearen Algebra sollte man wissen, dass ein *endlich dimensionaler* Vektorraum *nie* zu einem echten Unterraum isomorph sein kann.

Wir stellen eine (sehr kleine) Liste von Stammfunktionen stetiger Funktionen zusammen, links steht eine stetige Funktion f, in der Mitte eine Stammfunktion von f, in der dritten Spalte stehen Angaben zum Gültigkeitsbereich:
Diese Liste sollte man beherrschen ohne ein Tafelwerk oder ein CAS zur Hilfe zu nehmen. In klassischen Tafelwerken findet man Tausende von Stammfunktionen.

20.1.6 Eine kleine Liste von Stammfunktionen (Grundintegralen)

$f(x)$	$F(x) = \int f(x)dx$	Gültigkeitsbereich, Bemerkungen
$x^k,\ k \in \mathbb{Z},\ k \neq -1$	$\frac{x^{k+1}}{k+1}$	$x \in \mathbb{R}$ für $k \geq 0,\ x \neq 0$ für $k \leq -2$
$x^\alpha,\ \alpha \in \mathbb{R},\ \alpha \neq -1$	$\frac{x^{\alpha+1}}{\alpha+1}$	$x > 0$
$\frac{1}{x}$	$\log \lvert x\rvert$	$x > 0$ oder $x < 0$
$\sin(x)$	$-\cos(x)$	$x \in \mathbb{R}$
$\cos(x)$	$\sin(x)$	$x \in \mathbb{R}$
$\tan(x)$	$-\log \lvert\cos(x)\rvert$	$x \neq (k+\frac{1}{2})\pi,\ k \in \mathbb{Z}$
$\cot(x)$	$\log \lvert\sin(x)\rvert$	$x \neq k\pi,\ k \in \mathbb{Z}$
$\frac{1}{\cos^2(x)} = 1 + \tan^2(x)$	$\tan(x)$	$x \neq (k+\frac{1}{2})\pi, k \in \mathbb{Z}$
$\frac{1}{\sin^2(x)}$	$-\cot(x)$	$x \neq k\pi, k \in \mathbb{Z}$
$a^x\ (a \neq 1)$	$\frac{a^x}{\log a}$	$x \in \mathbb{R}$
$\exp(ax),\ a \in \mathbb{C},\ a \neq 0$	$\frac{1}{a}\exp(ax)$	$x \in \mathbb{R}$
$\exp(ix)$	$-i\exp(ix)$	$x \in \mathbb{R}$
$\sinh(x)$	$\cosh(x)$	$x \in \mathbb{R}$
$\cosh(x)$	$\sinh(x)$	$x \in \mathbb{R}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$	$\lvert x\rvert < 1$
$\frac{1}{\sqrt{1+x^2}}$	Arsinh(x)	$x \in \mathbb{R}$
$\frac{1}{\sqrt{x^2-1}}$	$\log \lvert x + \sqrt{x^2-1}\rvert$	$\lvert x\rvert > 1$
$\frac{1}{1+x^2}$	$\arctan(x)$	$x \in \mathbb{R}$
$\frac{f'(x)}{f(x)}$	$\log \lvert f(x)\rvert$	Intervalle, in denen $f(x) \neq 0$ ist.
$a_0 + a_1x + \ldots + a_nx^n$ $(n \in \mathbb{N}_0, a_j \in \mathbb{R})$	$a_0x + \frac{a_1x^2}{2} + \ldots + \frac{a_n}{n+1}x^{n+1}$	$x \in \mathbb{R}$
$\sum\limits_{k=0}^{\infty} a_kx^k\ (a_k \in \mathbb{R})$	$\sum\limits_{k=0}^{\infty} \frac{a_k}{k+1}x^{k+1}$	x im Konvergenzintervall der Potenzreihe

20.1.7 Bemerkung

Verschafft man sich Stammfunktionen mit Hilfe eines Computeralgebrasystems, etwa Maple oder Mathematica, so sollte man zur Sicherheit immer den Differentiations-Operator D anwenden, um zu testen, ob tatsächlich die Ausgangsfunktion herauskommt.
Mit Hilfe des Hauptsatzes lassen sich Aussagen der Differentialrechnung in Aussagen der Integralrechnung transformieren. Hierauf beruhen die folgenden *Integrationstechniken*.

20.2 Integrationstechniken, erste Anwendungen des Hauptsatzes

„Differentiation ist Technik, Integration ist Kunst“ (N..N.)

Auch im Zeitalter der Coputeralgebrasysteme, in dem einem etwa für die Funktion f mit

$$f(x) = \frac{x^5 - x^4 + 1}{x^2\sqrt{x^2-1}} \qquad (x \neq 0,\ x \neq \pm 1)$$

in Bruchteilen von Sekunden eine Stammfunktion angeboten wird, etwa F mit

$$F(x) = x^2\sqrt{x^2-1} - \frac{2}{3}(x^2-1)^{3/2} - \frac{1}{2}x\sqrt{x^2-1} + \frac{1}{2}\log(x+\sqrt{x^2-1}) + \frac{\sqrt{x^2-1}}{x},$$

sollte jede(r) einige *Grundtechniken zur Integralrechnung* beherrschen. Zwei dieser Techniken beruhen auf der Umkehrung der *Produktregel* bzw. der *Kettenregel* der Differentiation. Man kann sie für das unbestimmte oder das bestimmte Integral formulieren. Diese Regeln ermöglichen es in Fällen, in denen eine Stammfunktion nicht unmittelbar erkennbar ist, den Integranden in geeigneter Form in der Hoffnung umzuformen, doch eine Stammfunktion zu finden oder ein bestimmtes Integral explizit zu berechnen.

20.2.1 Satz (Partielle Integration oder Produktintegration)

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f : D \to \mathbb{R}$(oder $\mathbb{C}$) stetig und F eine Stammfunktion von f(auf D) und $g : D \to \mathbb{R}$ (oder $\mathbb{C}$) sei stetig differenzierbar.
Dann gilt

$$\int f(x)g(x)dx = F(x)g(x) - \int F(x)g'(x)dx.$$

Ist $D = [a, b]$, dann gilt

$$\int_a^b f(x)g(x)dx = F(x)g(x)|_a^b - \int_a^b F(x)g'(x)dx.$$

20.2.2 Korollar

Sind $u, v : D \to \mathbb{R}$(oder $\mathbb{C}$) stetig differenzierbar, dann gilt

$$\int u'(x)v(x)dx = u(x)v(x) - \int u(x)v'(x)dx$$

Kürzer: $\int vdu = uv - \int udv$ und für $D = [a, b]$ gilt

$$\boxed{\int_a^b u'(x)v(x)dx = u(x)v(x)|_a^b - \int_a^b u(x)v'(x)dx}$$

Der Beweis von 22.2.1 ergibt sich unmittelbar aus der Produktregel:

$$(Fg)' = F'g + Fg' = fg + Fg',$$

also ist Fg eine Stammfunktion von $fg + Fg'$ und daher (die Bildung von Stammfunktionen ist eine lineare Operation)

$$\int f(x)g(x)dx = F(x)g(x) - \int F(x)g'(x)dx.$$

Das Korollar ergibt sich mit $u' = f,\ F = u$ und $g = v$ oder auch direkt aus $(uv)' = u'v + uv'$.

20.2.3 Beispiel

Zu berechnen sei $\int x\exp(x)dx$.

1.Ansatz: $u'(x) = x,\ v(x) = \exp(x)$.
Dann ist $x \to \frac{1}{2}x^2$ eine Stammfunktion von u' und $v'(x)\exp(x)$. Folglich ist

$$\int x\exp(x)dx = \frac{1}{2}x^2\exp(x) - \int \frac{x^2}{2}\exp(x)dx.$$

Das letzte Integral ist aber komplizierter als das Ausgangsintegral, der Ansatz hat keine Vereinfachung gebracht.

2.Ansatz: $u'(x) = \exp(x),\ v(x) = x$, also $u(x) = \exp(x)$ und $v'(x) = 1$.
Man erhält so

$$\int x\exp(x) = x\exp(x) - \int 1\cdot\exp(x)dx = x\exp(x) - \exp(x) = (x-1)\exp(x),$$

also ist $x \mapsto (x-1)\exp(x)$ eine Stammfunktion von $x \mapsto x\exp(x)$ (auf ganz $\mathbb{R}$).

20.2.4 Bemerkungen und Beispiele

(a) Strategie bei der Anwendung der Methode der partiellen Integration sollte es sein, auf der rechten Seite der Formel im nicht ausintegrierten Bestandteil ein Integral zu erhalten, das einfacher ist als das Ausgangsintegral. Das erfordert ein gewisses Geschick und Übung. Ferner ist es zweckmäßig im Ausgangsintegral, den Faktor der im Laufe der Rechnung differenziert wird mit einem nach unten weisenden Pfeil („↓") und den Faktor, der integriert wird, mit einem nach oben weisenden Pfeil („↑") zu bezeichnen. Nach Möglichkeit sollte man den Faktor differenzieren, der bei der Differetiation einfacher wird und den Faktor integrieren, der sich bei der Integration wenigstens nicht allzu verkompliziert.

(b) Um eine Stammfunktion von $x \mapsto \log x\ (x > 0)$ also $\int \log x$, zu berechnen, beachte man $1 \cdot \log x = \log x$. So erhält man

$$\begin{aligned} \int \log x dx = \int \underset{\uparrow}{1}\underset{\downarrow}{\log x} dx &= x\log x - \int x\frac{1}{x}dx \\ &= x\log x - \int 1\ dx \\ &= x\log x - x \\ &= x(\log x - 1). \end{aligned}$$

(c) Bei Integration vom Typ

$$\int x^n\exp(\lambda x)dx\ (\lambda \in \mathbb{C}) \ \text{ oder } \int x^n\sin x dx \ \text{ oder } \int x^n\cos x dx, \qquad n \in \mathbb{N}_0,$$

ergeben sich so Rekursionsformeln. Ist z.B.

$$I_n := \int \underset{\downarrow}{x^n}\underset{\uparrow}{\exp}(x)dx = x^n\exp(x) - n\int x^{n-1}\exp(x)dx = x^n\exp(x) - nI_{n-1}.$$

Diese Rekursionsformeln führen I_n schließlich auf $I_0 = \int \exp(x)dx = \exp(x)$ zurück.

(d) Für $m, n \in \mathbb{N}_0$ sei $I_{n,m} = \int_0^1 x^n(1-x)^m dx$ zu berechnen.

Partielle Integration $I_{n,m} = \int_0^1 \underset{\uparrow}{x^n}(1-\underset{\downarrow}{x})^m dx$ ergibt die Rekursionsformel

$$I_{n,m} = \frac{m}{n+1}I_{n+1,m-1}.$$

Mit $I_{n+m,0} = \frac{1}{m+n+1}$ folgt schließlich

$$I_{n,m} = \frac{n!m!}{(n+m+1)!}:$$

damit kann man auch das Integral $I := \int\limits_{-1}^{1} (1+x)^n(1-x)^m dx$ leicht berechnen.

Setzt man $x := 2t - 1$, so erhält man mit der Substitutionsformel (vgl. 22.2.1), die wir hier vorwegnehmen,

$$I = 2^{n+m+1} \int\limits_0^1 t^n(1-t^m)dt = 2^{n+m+1} I_{n,m}.$$

(e) Um $\int \arctan x dx$ zu berechnen, beachten wir wieder $1 \cdot \arctan x = \arctan x$ und erhalten

$$\begin{aligned}\int \arctan x dx = \int \underset{\uparrow}{1} \cdot \underset{\downarrow}{\arctan x} dx &= x \arctan x - \int \frac{x}{1+x^2}dx\\ &= x\arctan x - \frac{1}{2}\int \frac{2x}{1+x^2}dx\\ &= x\arctan x - \frac{1}{2}\log(1+x^2).\end{aligned}$$

Man beachte, dass im letzten Integral $\int \frac{2x}{1+x^2}dx$ im Zähler die Ableitung des Nenners steht, ein solches Integral ist ein Grundintegral (vgl. 20.1.6).

(f) Für $x \in [-1,1]$ gilt

$$\int \sqrt{1-x^2}dx = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x).$$

Zunächst sei $x \in]-1,1[$, dann erhält man

$$\begin{aligned}I := \int \sqrt{1-x^2}dx = \int \underset{\uparrow}{1} \cdot \underset{\downarrow}{\sqrt{1-x^2}}dx &= x\sqrt{1-x^2} - \int \frac{x(-2x)}{2\sqrt{1-x^2}}dx\\ &= x\sqrt{1-x^2} + \int \frac{x^2}{\sqrt{1-x^2}}dx\\ &= x\sqrt{1-x^2} + \int \frac{1}{\sqrt{x^2-1}}dx - \int \frac{1-x^2}{\sqrt{1-x^2}}dx\\ &= x\sqrt{1-x^2} + \arcsin x - \int \sqrt{1-x^2}dx\\ &= x\sqrt{1-x^2} + \arcsin x - I,\end{aligned}$$

also

$$2I = x\sqrt{1-x^2} + \arcsin x \text{ oder } I = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x).$$

Da die Ableitungsregel $\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$ nur für $x \in]-1,1[$ gilt, gilt die abgeleitete Formel zunächst nur für die x im offenen Intervall $]-1,1[$.

Dass sie auch für die Randpunkte, also für alle $x \in [-1,1]$ gilt, folgt aus dem folgenden allgemeinen bemerkenswerten

Hilfssatz:

Ist $D \subset \mathbb{R}$ ein echtes Intervall und sind f und G stetige Funktionen auf D und ist $a \in D$ ein fester Punkt. Für alle $x \in D$, $x \neq a$, sei G differenzierbar bei x und $G'(x) = f(x)$. Dann ist G auch in a differenzierbar und es gilt $G'(a) = f(a)$.

Zum Beweis beachten wir, dass f nach dem Hauptsatz eine Stammfunktion $F : D \to \mathbb{R}$

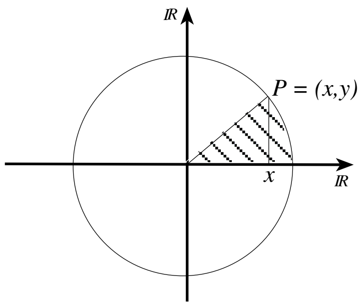
besitzt. Die Menge $D - \{a\}$ besteht entweder aus zwei getrennten Intervallen oder ist (im Fall, dass a ein Randpunkt ist) ein Teilintervall von D. Auf diesem Intervall ist die Differenz $G - F$ jeweils konstant, aber weil beide Funktionen auf D stetig sind, ist $G - F$ insgesamt konstant, etwa $G - F = c$ oder $G = F + c$. Da F in a differenzierbar ist und $F'(a) = f(a)$ gilt, ist auch G in a differenzierbar und es gilt $G'(a) = f(a)$.

Wendet man diesen Hilfssatz auf $f(x) = \sqrt{1-x^2}$ und $G(x) = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)$ an (einmal für $a = 1$ und einmal für $a = -1$), erhält man für alle $x \in [-1, 1]$

$$\boxed{\int \sqrt{1-x^2} = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)}$$

□

(g) **Anwendung: Die Fläche des Einheitskreises**
Als Anwendung von (f) erhalten wir, dass der Flächeninhalt des Einheitskreises π ist. Wir betrachten dazu zunächst den straffierten Sektor in der und nennen seinen Flächeninhalt $\frac{1}{2}F$.



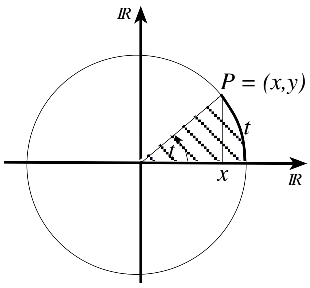
(Es setzt sich aus dem Dreieck OPx und dem durch $\int_x^1 \sqrt{1-t^2}dt$ gegebenen Flächeninhalt zwischen dem Graphen von $x \mapsto \sqrt{1-x^2}$ und der reellen Achse zwischen Punkten x und 1 zusammen).
Wegen der Kreisgleichung $x^2 + y^2 = 1$ gilt

$$\begin{aligned} F = xy + 2\int_x^1 \sqrt{1-t^2}dt &= x\sqrt{1-x^2} + (t\sqrt{1-t^2} + \arcsin t)\Big|_x^1 \\ &= \arcsin 1 - \arccos x \\ &= \frac{\pi}{2} - \arcsin x = \arccos x. \end{aligned}$$

Speziell für $x = -1$ erhält man wegen $\arccos(-1) = \pi$ den Flächeninhalt des Einheitskreises (besser der Einheitskreisscheibe).

$$\boxed{\text{Der Flächeninhalt der Einheitskreisscheibe ist also } \pi.}$$

Man beachte, dass wir $\frac{\pi}{2}$ als kleinste positive Nullstelle den $\cos : \mathbb{R} \to \mathbb{R}$ eingeführt haben. π hat also auch eine klassische Bedeutung. Wie wir wissen (und auch nochmals sehen werden) ist 2π die *Länge* der Einheitskreislinie.
Gehört zum Punkt $P \in S^1 := \{(x, y) \in \mathbb{R}^2;\ x^2 + y^2 = 1\}$ der im Bogenmaß gemessene Winkel t, so ist also $\cos t = x$ oder $t = \arccos x$. Der im Bogenmaß gemessene Winkel ist also gleich dem doppelten Flächeninhalt des straffierten Sektors oder:
Der straffierte Sektor hat den Flächeninhalt $\frac{1}{2}t$.



Die die Einheitskreislinie die Länge 2π hat, ergibt sich nochmals, das π der Flächeninhalt der Einheitskreisscheibe ist.

Eine schöne Anwendung der Produktintegration ist der sog. *Wallis'sche Produktformel* für $\frac{\pi}{2}$.

20.2.5 Die Wallische Produktformel für $\frac{\pi}{2}$ (John Wallis, Arithmetica Infiniforum, 1666)

Ist

$$w_n = \frac{2 \cdot 2}{1 \cdot 3} \cdot \ldots \cdot \frac{2n \cdot 2n}{(2n-1)(2n+1)} = \prod_{k=1}^{n} \frac{(2k)^2}{(2k-1)(2k+1)} = \prod_{k=1}^{n} \frac{4k^2}{4k^2-1}$$

für $n \in \mathbb{N}$, dann ist die Folge (w_n) konvergent und es gilt

$$\boxed{\lim_{n\to\infty} w_n = \lim_{n\to\infty} \prod_{k=1}^{n} \frac{4k^2}{4k^2-1} = \frac{\pi}{2}}$$

Zum **Beweis** betrachten wir die Folge (I_k), $k \in \mathbb{N}_0$ mit

$$I_k := \int\limits_0^{\frac{\pi}{2}} \sin^k x dx$$

Dann ist $I_0 = \frac{\pi}{2}$ und $I_1 = 1$ und für $k \geq 2$ erhält man

$$\begin{aligned} I_k &= \int\limits_0^{\frac{\pi}{2}} \underset{\uparrow}{\sin(x)}\, \underset{\downarrow}{\sin}^{k-1}(x) dx \\ &= -\cos x \sin^{k-1}(x)\Big|_0^{\frac{\pi}{2}} + (k-1) \int\limits_0^{\frac{\pi}{2}} \cos^2(x) \sin^{k-2}(x) dx \\ &= (k-1) \int\limits_0^{\frac{\pi}{2}} (1-\sin^2(x)) \sin^{k-2}(x) dx, \end{aligned}$$

also

$$I_k = (k-1) I_{k-2} - (k-1) I_k \quad \text{oder} \quad I_k = \frac{k-1}{k} I_{k-2} \quad \text{für} \quad k \geq 2.$$

Ist k gerade, also $k = 2n,\ n \in \mathbb{N}_0$, so endet der iterative Abstieg bei 0 und man erhält wegen $I_0 = \frac{\pi}{2}$

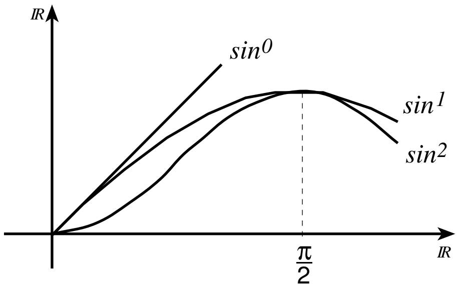
$$I_{2n} = \frac{2n-1}{2n} \cdot \frac{2n-3}{2n-2} \cdot \ldots \cdot \frac{3}{4} \cdot \frac{1}{2} \cdot I_0.$$

Ist jedoch k ungerade; $k = 2n+1,\ n \in \mathbb{N}_0$, so ergibt sich analog wegen $I_1 = 1$

$$I_{2n+1} = \frac{2n}{2n+1} \cdot \frac{2n-2}{2n-1} \cdot \ldots \cdot \frac{4}{5} \cdot \frac{2}{3} \cdot 1.$$

Weil für $x \in [0, \frac{\pi}{2}]$ gilt $\sin^{2n+2} x \leq \sin^{2n+1} x \leq \sin^{2n}$ folgt

$$(*) \qquad I_{2n+2} \leq I_{2n+1} \leq I_{2n}.$$



Nach Definition ist $w_n = \frac{\pi}{2} \frac{I_{2n+1}}{I_{2n}}$. Es bleibt also $\lim\limits_{n\to\infty} \frac{I_{2n+1}}{I_{2n}} = 1$ zu zeigen.
Aus $(*)$ folgt aber

$$(**) \qquad \frac{I_{2n+2}}{I_{2n}} \leq \frac{I_{2n+1}}{I_{2n}} \leq 1$$

und wegen

$$\lim_{n\to\infty} \frac{I_{2n+2}}{I_{2n}} = \lim_{n\to\infty} \frac{2n+1}{2n+2} = \lim_{n\to\infty} \left(1 - \frac{1}{2n+2}\right) = 1$$

folgt nach dem Sandwich-Theorem auch $\lim\limits_{n\to\infty} \frac{I_{2n+2}}{I_{2n}} = 1$, also ist

$$\boxed{\lim_{n\to\infty} w_n = \lim_{n\to\infty} \prod_{k=1}^{n} \frac{4k^2}{4k^2-1} = \frac{\pi}{2}}$$

□

Für den Limes in der Mitte verwendet man auch die Bezeichnung $\prod\limits_{k=1}^{\infty} \frac{4k^2}{4k^2-1}$ und nennt einen solchen Ausdruck auch ein *unendliches Produkt*.
Wegen der Sonderrolle der Null bezüglich der Multiplikation ist die Konvergenztheorie für unendliche Produkte schwieriger als für Reihen. Wir gehen hier nicht darauf ein.

20.2.6 Bemerkung

Die Wallis'sche-Produktformel für $\frac{\pi}{2}$ wird uns nützlich sein im Zusammenhang mit der Theorie der Gamma-Funktion (sie erlaubt eine einfache Bestimmung des Funktionswertes $\Gamma(\frac{1}{2})(=\sqrt{\pi})$ und im Zusammenhang mit der Stirlingschen Formel (Wachstum von $n!$).
Man vergleiche hierzu den Abschnitt 22. Zur numerischen Berechnung von π ist die Wallis'sche Produktformel nicht besonders gut geeignet, da die Folge (w_n) sehr langsam konvergiert:

n	w_n			
1	w_1	$=$	$\frac{4}{3}$	$= 1,333333\ldots$
2	w_2	$=$	$\frac{64}{45}$	$= 1,422222\ldots$
3	w_3	$=$	$\frac{256}{175}$	$= 1,46286\ldots$
4	w_4	$=$	$\frac{16834}{11025}$	$= 1,48608\ldots$
5	w_5	$=$		$= 1,50108\ldots$
10	w_{10}	$=$		$= 1,53385\ldots$
100	w_{100}	$=$		$= 1,56689\ldots$
1000	w_{1000}	$=$		$= 1,57040\ldots$

Der „exakte" Wert von $\frac{\pi}{2}$ ist $\frac{\pi}{2} = 1,5707963267\ldots$.
Immerhin stimmt also w_{1000} in den ersten drei Nachkommastellen mit dem exakten Wert überein.

Die heutigen schnellen Algorithmen zur Berechnung von π liefern auch die Möglichkeit, einzelne (Dezimal- bzw. Hexadezimal-)Stellen „irgendwo“ in π zu berechnen. Eine solche Formel ist etwa

$$\boxed{\pi = \sum_{k=0}^{\infty} \frac{1}{16^k}\left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6}\right)}$$

Diese sog. BBP-Reihe wurde mit dem Computer entdeckt, man kann sie mit Hilfe des Vertauschungssatzes für Integration und Summation relativ einfach beweis. Vergleiche hierzu:

Davis H.Bailey, Jonathan M.Borwein, Peter B.Borwein und Simon Plouffe:
The Quest for Pi, The mathematical Intelligencer,
Vol.19(1997), No.1, 50-56.

oder auch

Hort S. Holdgrün: Analysis, Band 1, Leins-Verlag, 1998,
dort Abschnitt 38.20 und 38.21

Ist s_n die n-te Partialsumme der obigen Reihe, so gilt $|s_n - \pi| < \frac{1}{30(n+1)16^n}$. Berechnet man 25 Summanden der Partialsumme s_{24}, so erhält man

$$s_{24} = 3{,}14\ 15\ 92\ 653\ 589\ 793\ 238\ 462\ 643\ 383\ 279\ 50\ldots$$

Der „exakte“ Wert von π ist

$$\pi = 3{,}14\ 15\ 92\ 653\ 589\ 793\ 238\ 462\ 643\ 383\ 279\ 502\ 884\ldots,$$

die ersten 32 Nachkommastellen von s_{24} sind also korrekt.
Die Darstellung von π mit der BBP-Reihe erlaubt es auch, eine beliebige Ziffer in der Hexa-Dezimal-Darstellung von π zu berechnen, ohne dass man die vorausgehende Ziffern zu kennen braucht. dabei genügt etwas Ausdauer und ein einfacher Taschenrechner! Versuchen Sie es!
Nebenbei: Die 20 billionste Hexadezimalstelle von π ist „A“, die 10 milliardste Hexadezimalstelle ist eine „9“.

Die Substitutionsmethode erhält man durch Umkehrung der Kettenregel der Differentiation. Dazu seien $M, N \subset \mathbb{R}$ echte Intervalle, $\varphi : M \to \mathbb{R}$ stetige differenzierbare Funktion mit $\varphi(M) \subset N$ und $F : N \to \mathbb{R}$ stetig differenzierbar. Dazu ist auch die Zusammensetzung $G = F \circ \varphi : M \to \mathbb{R}$ stetig differenzierbar und für $t \in M$ gilt

$$G'(t) = F'(\varphi(t))\varphi'(t).$$

Übergang zu Stammfunktionen auf beiden Seiten liefert

$$\int F'\varphi(t)\varphi'(t)dt = G(t) = F(\varphi(t))$$

Setzt man $x = \varphi(t)$ und $F'(x) = f(x)$, so erhält man

$$\int f(\varphi(t))\varphi'(t)dt = \int f(x)dx|_{x=\varphi(t)},$$

Wobei die Schreibweise rechts bedeuten soll, dass man in der Stammfunktion $F(x)$ und $f(x)$ für x die Funktion $\varphi(t)$ einzusetzen hat.

20.2.7 Satz (Substitutionsregel)

(a) Sind $M, N \subset \mathbb{R}$ echte Intervalle, $\varphi : M \to \mathbb{R}$ stetig differenzierbar mit $\varphi(M) \subset N$ und $f : N \to \mathbb{R}$ stetig, so gilt mit $x = \varphi(t)$

$$\int f(\varphi(t))\varphi'(t)dt = \int f(x)dx|_{x=\varphi(t)}$$

(b) Ist $M = [a, b]$, so gilt speziell

$$\int\limits_a^b f(\varphi(t))\varphi'(t)dt = \int\limits_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

Dabei ergibt sich (b) so: Ist F eine Stammfunktion von f, dann ist $F \circ \varphi$ eine Stammfunktion von $(f \circ \varphi)\varphi'$, also

$$\int\limits_a^b f(\varphi(t))\varphi'(t)dt = F(\varphi(t))|_a^b = F(\varphi(b)) - F(\varphi(a)) = \int\limits_{\varphi(a)}^{\varphi(b)} f(x)dx$$

20.2.8 Bemerkung

Unter Verwendung der Leibniz'scher Schreibweise (man kann diese durch das Rechnen mit Differentialen rechtfertigen) $dx = d\varphi(t) = \varphi'(t)dt$ lautet die Substitutionsregel einfach

$$\int f(\varphi(t))\varphi'(t)dt = \int f(\varphi(t))d\varphi(t) = \int f(x)dx.$$

Analog für das bestimmte Integral

$$\int\limits_a^b f(\varphi(t))\varphi'(t)dt = \int\limits_a^b f(\varphi(t))d\varphi(t) = \int\limits_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

Man hat einfach x durch $\varphi(t)$ zu ersetzen: Läuft t von a nach b, so läuft $x = \varphi(t)$ von $\varphi(a)$ nach $\varphi(b)$. dabei hat man die Konvention $\int\limits_c^d f = -\int\limits_d^c f$ zu beachten.

20.2.9 Beispiele und Bemerkungen

Die Substitutionsregel lässt sich für unbestimmte und bestimmte Integrale in zwei Richtungen anwenden: Von links nach rechts und von rechts nach links. Wir behandeln jeweils typische Beispiele.

(a) Zu berechnen sei $\int e^{t^2} \cdot tdt$.
Setzt man versuchsweise $x = \varphi(t) = t^2$, so folgt $\varphi'(t) = 2t$, also mit $f(x) = e^x$

$$2\int e^{t^2} \cdot tdt = \int e^x dx \quad (x = t^2),$$

folglich

$$\int e^{t^2} \cdot tdt = \frac{1}{2}\int e^x = \frac{e^x}{2} = \frac{e^{t^2}}{2}.$$

Mit einem scharfer Blick hätte man auch sofort erkennen können, dass $\frac{1}{2}\int e^{t^2} \cdot 2tdt = \frac{1}{2}\int e^x dx = \frac{e^x}{2} = \frac{e^{t^2}}{2}$ ist.

(b) Zu berechnen sei $\int \left(\sin^3 t + e^{\sin t} \cos t dt\right)$.
Man sieht direkt, dass G mit $G(t) = \frac{1}{4}\sin^4 t + e^{\sin t}$ eine Stammfunktion ist, also ist

$$\int (\sin^3 t + e^{\sin t}) \cos t dt = \frac{1}{4}\sin^4 t + e^{\sin t}$$

(c) $I := \int\limits_0^{\frac{\pi}{2}} \sin^3 t \cdot \cos t dt$ ist zu berechnen mit $f(x) = x^3$ und $x = \varphi(t) = \sin t$ erhält man

$$I = \int\limits_0^{\frac{\pi}{2}} \sin^3 t \cos t dt = \int\limits_0^{\frac{\pi}{2}} f(\varphi(t))\varphi'(t)dt = \int\limits_0^1 x^3 dx = \left.\frac{x^4}{4}\right|_0^1 = \frac{1}{4}.$$

(d) Das Integral $I = \int(\cos t + \cos^3 t)dt$ hat zunächst nicht die benötigte Form, um 20.2.7(a) von links nach rechts anwenden zu können. Beachtet man aber

$$I = \int (1 + \cos^2 t) \cos t dt = \int (2 - \sin^2 t) \cos t dt$$

und setzt jetzt $\sin t := \varphi(t) = x$, $dx = \cos t dt$, so erhält man

$$I = \int \left. (2 - x^2)dx \right|_{x=\sin t} = \left.\left[2x - \frac{x^3}{3}\right]\right|_{x=\sin t} = (2\sin t - \frac{1}{3}\sin^3 t)$$

Für das bestimmte Integral $I_0 = \int\limits_{\frac{\pi}{6}}^{\pi} (\cos t + \cos^3 t)dt$ erhält man so z.B.

$$I_0 = \int\limits_{\frac{1}{2}}^{0} (2 - x^2)dx = \left.\left(2x - \frac{x^3}{3}\right)\right|_{\frac{1}{2}}^{0} = -\left(1 - \frac{1}{24}\right) = -\frac{23}{24}.$$

(e) Ist $f = F'$ und $\alpha \neq 0$, so gilt

$$\int f(\alpha t + \beta)dt = \frac{1}{\alpha}\int f(\alpha t + \beta)\alpha dt = \frac{1}{\alpha}\int \left. f(x)dx\right|_{x=\alpha t+\beta} = \frac{1}{\alpha}F(\alpha t + \beta).$$

Hiermit ergibt sich z.B. (für $a \neq 0$)

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

und für bestimmte Integrale als Spezialfälle

$(*)$ $\int\limits_a^b f(t + a)dt = \int\limits_{a+c}^{b+c} f(x)dx$ und

$(**)$ $\int\limits_a^b f(\alpha t)dt = \frac{1}{\alpha}\int\limits_{\alpha a}^{\alpha b} f(x)dx \quad (\alpha \neq 0)$

(f) Ist $\varphi : M \to \mathbb{R}$ stetig differenzierbar und $\varphi'(t) \neq 0$ für alle $t \in [a, b]$, dann ist

$$\int \frac{\varphi'(t)}{\varphi(t)}dt = \log|\varphi(t)| \;\; (f(x) = \frac{1}{x};\; x = \varphi(t))$$

Für $a, b \in M$ ist also

$$\int\limits_a^b \frac{\varphi'(t)}{\varphi(t)}dt = \log|\varphi(t)|\Big|_a^b$$

(siehe auch die Tabelle der Grundintegrale)

(g) Für $a, b \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ gilt

$$\int_a^b \tan t dt = \int_a^b \frac{\sin t}{\cos t} dt = -\log \cos t|_a^b$$

(h) Zu berechnen sei $\int_1^4 (1+x)\sqrt{x}dx$.

Hier wird man versuchen die Substitutionsformel von rechts nach links anzuwenden. Um die Wurzel zu beseitigen, setzen wir $x = \varphi(t) = t^2$, $dx = \varphi'(t)dt = 2tdt$ und rechnen die Grenzen um: $x_1 = 1$ wird für $t_1 = 1$ und $x_2 = 4$ wird für $t_2 = 2$ angenommen. Damit erhalten wir

$$\begin{aligned}\int_1^4 (1+x)\sqrt{x}dx = \int_1^2 (1+t^2)t \cdot 2tdt &= 2\int_1^2 (t^2+t^4)dt \\ &= 2\left(\frac{1}{3}t^3 + \frac{1}{-}5t^5\right)\Big|_1^2 \\ &= \frac{14}{3} + \frac{62}{5} = \frac{256}{15}.\end{aligned}$$

(i) Zu berechnen sei $I = \int_1^2 x\sqrt{5x-1}dx$.

Wir wollen φ so bestimmen, dass $\sqrt{5\varphi(t)-1} = t$ wird. Dazu lösen wir $\sqrt{5x-1} = t$ nach x auf und erhalten

$$x = \frac{1}{5}(t^2+1),\ dx = \frac{2}{5}tdt$$

und erhalten mit $t_1 = \sqrt{5x_1 - 1} = 2$ und $t_2 = \sqrt{5x_2 - 1} = 3$

$$\begin{aligned}\int_1^2 x\sqrt{5x-1} &= \int_2^3 \frac{1}{5}(t^2+1)t \cdot \frac{2}{5}tdt \\ &= \frac{2}{25}\int_2^3 (t^4+t^2)dt \\ &= \frac{2}{25}\left(\frac{t^5}{5} + \frac{t^3}{3}\right)\Big|_2^3.\end{aligned}$$

Durch Einsetzen der Grenzen erhält man den gesuchten Integralwert.
Hier haben wir eine geeignete Subsitutionsfunktion gefunden, indem wir die Gleichung $\sqrt{5x-1} = t$ nach x aufgelöst haben, d.h. wir haben die Umkehrfunktion von $x \mapsto \sqrt{5x-1}$ bestimmt oder $\sqrt{5x-1} = \psi(t)$ gesetzt, wobei ψ die Umkehrfunktion von φ ist.
Das hat funktioniert, weil

$$\begin{aligned}\varphi : [2,3] &\to [1,2] \\ t &\mapsto \frac{1}{5}(t^2+1)\end{aligned}$$

bijektiv ist mit der Umkehrfunktion

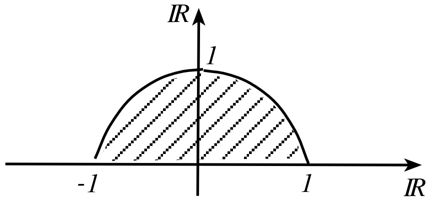
$$\begin{aligned}\psi : [1,2] &\to [2,3] \\ x &\mapsto \sqrt{5x-1}.\end{aligned}$$

Ist $\varphi : M \to N$ bijektiv mit der Umkehrabbildung $\psi : N \to M$, dann kann man die Substitutionsformel auch in der Form $(x_1, x_2 \in N)$

$$\int_{\psi(x_1)}^{\psi(x_2)} f(\varphi(t))\varphi'(t)dt = \int_{x_1}^{x_2} f(x)dx$$

schreiben.

(j) Zu berechnen sei $\int\limits_{-1}^{1} \sqrt{1-x^2}dx$.



In der geometrischen Interpretation des Integrals erhält man also den halben Flächeninhalt der Einheitskreises. Unsere Berechnung sollte also den Wert $\frac{\pi}{2}$ ergeben.
Hier liegt die Substitution $x = \varphi(t) = \cos t$ nahe:
Wir integrieren unbestimmt und erhalten

$$\int \sqrt{1-x^2}dx = -\int \sqrt{1-\cos^2 t}\sin t dt = -\int \sin^2 t dt.$$

$\int \sin^2 t$ kann man mit partieller Integration berechnen oder unter Beachtung von $\sin^2 t = \frac{1}{2}(1-\cos(2t))$. Das folgt aus den Additionstheoremen

$$\cos(2x) = \cos^2 t - \sin^2 t \text{ und } 1 = \cos^2 t + \sin^2 t.$$

Also ergibt sich

$$\begin{aligned} \int \sqrt{1-x^2}dx &= -\int \sin^2 t dt \\ &= -\frac{1}{2}(t - \sin t \cos t) \\ &= -\frac{1}{2}\arccos x + \frac{x}{2}\sqrt{1-x^2}. \end{aligned}$$

Für das bestimmte Integral ergibt sich daher

$$\begin{aligned} \int_{-1}^{1} \sqrt{1-x^2}dx &= \left(-\frac{1}{2}\arccos x + \frac{x}{2}\sqrt{1-x^2}\right)\Bigg|_{-1}^{1} \\ &= -\frac{1}{2}\arccos 1 + \frac{1}{2}\arccos(-1) \\ &= 0 + \frac{1}{2}\pi = \frac{\pi}{2} \end{aligned}$$

in Übereinstimmung mit unseren früheren Resultaten.

Die Beispiele zeigen, dass ein wesentlicher Teil der „Kunst der Integration“ in der Auffindung einer geeigneten Substitution besteht. Durch raffinierte Substitutionen, die wie ein Zaubertrick wirken, lassen sich manchmal sehr komplizierte Integrale so vereinfachen, dass man Stammfunktion angeben oder die Integrale berechnen kann.

Um z.B. das Integral $\int \frac{x^7}{x^4+2}dx$ zu berechnen, kann man durch $u := x^4$ eine neue Variable einführen, und man erhält sofort

$$\int \frac{x^7}{x^4+2}dx = \frac{1}{4}\int \frac{u}{u+2}du = \frac{1}{4}\int \frac{u+2-2}{u+2}du = \frac{1}{4}u - \frac{1}{2}\log(u+2) = \frac{1}{4}x^4 - \frac{1}{2}\log(x^4+2).$$

Das gleiche Resultat, allerdings mit erheblich mehr Aufwand, hätte man mit der *Methode der Partialbruchzerlegung* einer rationalen Funktion erhalten können.
Wie schon festgestellt, garantiert der Hauptsatz zwar, das jede stetige Funktion f auf einem Intervall D eine Stammfunktion besitzt. Doch ist damit noch lange nicht gesagt, dass eine Stammfunktion zu Klasse der *elementaren Funktionen* gehören muss. Unter der *elementaren Funktion* versteht man dabei die Gesamtheit aller Funktionen, die sich aus Polynomen, der Exponentialfunktionen, dem Sinus und allen denjenigen Abbildungen zusammensetzt, die sich hieraus mittels der vier „Grundrechenarten" (Addition, Subtraktion, Multiplikation und Division), sowie den Operationen „Zusammensetzen" und Bildung der Umkehrfunktionen in endlich vielen Schritten gewonnen werden können.
Aus unseren früheren Überlegungen folgt:

Die Ableitung einer elementaren Funktion ist wieder eine elementare Funktion.

Bei Stammfunktionen elementarer Funktionen braucht dies nicht der Fall zu sein.
Systematisch hat sich wohl zuerst J.Liouville (1833) damit beschäftigt. Er hat für eine große Klasse stetiger Funktionen nachgewiesen, dass sie nicht elementar invertierbar sind. In jüngster Zeit hat man sich mit dem Aufkommen der Coputeralgebrasystemenen wieder mit dieser Frage beschäftigt und dabei große Fortschritte erzielt (vgl. z.B. R.H.Risch: The problem of integration in finite terms, Transac.Amer.Math.Soc.139 (1969)167-183).
Durch Integration lassen sich aus gegebenen Funktionen neue Funktionen gewinnen, manche haben wegen ihrer Wichtigkeit einen eigenen Namen:
Der *Integralsinus* $Si(x)$ ist definiert durch

$$Si(x) = \int_0^x \frac{\sin t}{t} dt = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)(2k+1)!}, \quad x \in \mathbb{R}.$$

Die Fehlerfunktion (Errorfunction) ist definiert durch

$$Err(f) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)k!}$$

Durch die angegebenen Reihenentwicklungen rechts lassen sich jedoch die Funktionswerte beliebig genau berechnen.
Der *Integralalgorithmus* definiert durch

$$Li(x) = \int_2^x \frac{1}{\log t} dt$$

Dazu gehören auch die sog. *elliptischen Integrale*, wie

$$\int \frac{1}{\sqrt{(1-t^2)(1-k^2t^2)}} dt \text{ und } \int \sqrt{\frac{1-k^2t^2}{1-t^2}} dt \quad (0 < k^2 \neq 1)$$

(sog. Legendiesche Normalintegrale) oder die Weierstrass'schen Normalintegrale

$$\int \frac{1}{\sqrt{4t^3 - g_2 t - g_3}} dt \text{ oder } \int \frac{t}{\sqrt{4t^3 - g_2 t - g_3}} dt,$$

wobei die Konstanten g_2, g_3 die Bedingung $\Delta(g_2, g_3) := g_2^3 - 27g_3^2 \neq 0$ erfüllen sollen. Das bedeutet, dass das kubische Polynom unter der Wurzel keine mehrfachen Nullstellen hat.
Historisch am Anfang dieser Integralbehandlungen (G.C.Fagnano, 1718) steht das elliptische integral

$$E(x) = \int_0^x \frac{1}{\sqrt{1-t^4}} dt, \quad 0 \leq x < 1.$$

Die Umkehrfunktion von E, nennen wir sie G, besitzt eine Fortsetzung als (meromorphe) Funktion nach $\mathbb{C}$, die Fortsetzung $\tilde{G}$ ist *doppelt-periodisch*, d.h. es gibt zwei über $\mathbb{R}$-linear unabhängige komplexe Zahlen w_1, w_2 (Perioden), mit $\tilde{G}(z+w_1) = \tilde{G}(z+w_2) = \tilde{G}(z)$. Da unter den elementaren Funktionen keine doppelt-periodischen vorkommen, kann $\tilde{G}$ nicht zu den elementaren Funktionen zählen.

Im folgenden Abschnitt werden wir jedoch zeigen, dass *rationale Funktionen elementar* invertierbar sind. Diese Aussage werden wir noch wesentlich präzisieren können.

20.3 Partialbruchzerlegung und die Integration der rationalen Funktionen

Dieser Abschnitt wird nachgetragen.

Als weitere Anwendung des Hauptsatzes beschäftigen wir uns mit der Frage der *Vertauschbarkeit von Differentiation mit Grenzprozessen.*

20.4 Verträglichkeit der Differentiation mit Grenzprozessen

Wir gehen von folgender Fragestellung aus: Ist $D \subset \mathbb{R}$ ein echtes Intervall und (f_n) eine Folge von differenzierbaren Funktionen $f_n : D \to \mathbb{R}$, die gegen eine Funktion $f : D \to \mathbb{R}$ konvergiert (etwa punktweise oder gleichmäßig).
Die Frage, die sich automatisch stellt: Ist dann f auch differenzierbar und gilt etwa $f' = \lim\limits_{n\to\infty} f_n'$?

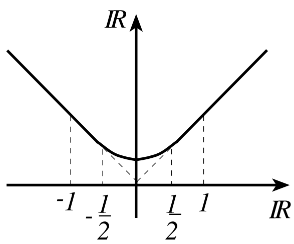
(1) Die schlechten Erfahrungen, die wir bei punktweisen Konvergenz gemacht haben, lassen vermuten, dass f i.A. nicht mehr differenzierbar ist. Das einfache Beispiel $f_n : [0,1] \to \mathbb{R}$ mit $f_(x) = x^n$ zeigt, dass die Grenzfunktion nicht einmal stetig, geschweige denn differenzierbar ist.

(2) Auch die gleichmäßige Konvergenz garantiert nicht die Differenzierbarkeit der Grenzfunktion: Dazu geben wir eine Folge (f_n) von Funktionen $f_n : [-1,1] \to \mathbb{R}$ an, die gleichmäßig gegen

$$\begin{array}{rcl} abs : [-1,1] & \to & \mathbb{R} \\ x & \mapsto & abs(x) = |x| \end{array}$$

konvergiert. Dazu definieren wir

$$f_n(x) = \begin{cases} \frac{n}{2}x^2 + \frac{1}{2n}, & \text{für } |x| \leq \frac{1}{2} \\ |x|, & \text{für } |x| \geq \frac{1}{2} \end{cases}$$



Kritische sind die Punkte $-\frac{1}{n}$ und $\frac{1}{n}$. Dort haben beide Funktionen rechts in der Definition die Steigung -1 bzw. 1. Das Minimum von f_n liegt bei 0 und hat den Wert $\frac{1}{2}n$, daher ist

$$\|f_n - abs\| \leq \frac{1}{2n}.$$

f_n konvergiert also gleichmäßig gegen abs, alle f_n sind differenzierbar, die Grenzfunktion abs ist aber nicht differenzierbar (in Null).
Nach dem Weierstrass'schen Approximationssatz (siehe Kap VI) ist jede stetige Funktion $f : [a,b] \to \mathbb{R}$ gleichmäßige Limes einer Folge von Polynomfunktionen p_n. Jedes $p_n \in C^\infty([a,b])$, eine beliebige stetige Funktion braucht aber nicht differenzierbar zu sein.

(3) Selbst wenn die Grenzfunktion f eine gleichmäßig konvergente Folge differenzierbarer Funktionen f_n differenzierbar ist, braucht nicht $f' = \lim\limits_{n\to\infty} f_n'$ zu gelten.
Sei dazu für $n \in \mathbb{N}$

$$\begin{aligned} f_n : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \frac{\sin nx}{n}, \end{aligned}$$

(f_n) ist wegen $|\sin x| \le 1$ gleichmäßig konvergent mit dem Nullfunktion $f = 0$ als Grenzfunktion.
Es ist aber $f_n'(x) = \cos nx$, $(f_n'(x))$ konvergiert z.B. für $x = 0$, $f_n'(0) = \cos 0 = 1$, aber $\lim\limits_{n\to\infty} f_n'(0) = 1 \neq 0 = f'(0)$.

(4) Analoge Beispiele lassen sich bei Reihen differenzierbarer Funktionen finden:
Die Reihe $\sum\limits_{n=1}^{\infty} \frac{\sin(nx)}{n^2}$ konvergiert gleichmäßig auf $\mathbb{R}$, die Ableitung $\sum\limits_{n=1}^{\infty} \frac{\cos nx}{n}$ konvergiert aber z.B. nicht für $x = 0$.

Ein Vertauschungssatz für die Differentiation und Grenzwertbildung muss also etwa komplizierter aussehen.

20.4.1 Satz (Vertauschbarkeit von Differentiation mit Grenzprozessen)

Ist $D \subset \mathbb{R}$ ein echtes Intervall und gegeben sei eine Folge (f_n) von differenzierbaren Funktionen $f_n : D \to \mathbb{R}$.
Wir machen folgende Voraussetzungen:

(a) Die Folge (f_n) konvergiert in mindestens einem Punkt $x_0 \in D$.
(b) Die Funktionen f_n', $n \in \mathbb{N}$, sind stetig (m.a.W. $f_n \in C^1(D)$)
(c) Die Folge (f_n') konvergiere gleichmäßig. (es reicht: gleichmäßig auf jedem kompakten Teilintervall $[\alpha, \beta] \subset D$.

Dann konvergiert die Folge (f_n) gegen eine differenzierbare Funktion $f : D \to \mathbb{R}$ und es gilt $f = \lim\limits_{n\to\infty} f_n'$ oder suggestiver

$$\boxed{(\lim_{n\to\infty} f_n)' = \lim_{n\to\infty} f_n'}$$

Zusatz: Ist $D = [a,b]$ eine kompaktes Intervall, so konvergiert die Folge (f_n) gleichmäßig gegen f. Die Grenzfunktion der Folge (f_n') sei g. Als gleichmäßiges Limes stetiger Funktionen ist g stetig (vgl. ???). Auf Grund des Hauptsatzes gilt

$$f_n(x) = f_n(x_0) + \int_{x_0}^{x} f_n'(t)dt \quad \text{für } x \in D \text{ und } n \in \mathbb{N}.$$

Nach dem Stabilitätssatz (vgl. ???) gilt

$$\lim_{n\to\infty} \int_{x_0}^{x} f_n'(t)dt = \int_{x_0}^{x} g(t)dt \quad \text{für alle } x \in D$$

und damit existiert

$$f(x) := \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} f_n(x_0) + \int_{x_0}^{x} g(t)dt = f(x_0) + \int_{x_0}^{x} f(t)dt \quad \text{für alle} \quad x \in D$$

Der Hauptsatz liefert nun für alle $x \in D$

$$f'(x) = g(x) = \lim_{n\to\infty} f_n'(x).$$

Beweis des Zusatzes:
Ist $D = [a, b]$ ein kompaktes Intervall, so folgt aus der Ungleichung

$$\begin{aligned} |f_n(x) - f(x)| &\leq |f_n(x_0) - f(x_0)| + \left| \int_{x_0}^{x} f_n'(t)dt - \int_{x_0}^{x} g(t)dt \right| \\ &\leq |f_n(x_0) - f(x_0)| + |x - x_0| \|f_n' - g\| \\ &\leq |f_n(x_0) - f(x_0)| + (b-a)\|f_n' - g\| \end{aligned}$$

für alle $x \in [a, b]$ die gleichmäßige Konvergenz von (f_n) gegen f.
Wendet man den Satz auf die Partialsummen einer Funktionenreihe an, so erhält man

20.4.2 Korollar

Ist $D \subset \mathbb{R}$ ein echtes Intervall, (f_n) eine Folge von stetig differenzierbaren Funktionen mit folgenden Eigenschaften:

Es gibt einen Punkt $x_0 \in D$, so dass die Folge $(F_n) = \left(\sum\limits_{k=0}^{n} f_k\right)$ der Partialsummen in x_0 konvergiert.

Die Folge $(F_n') = \left(\sum\limits_{k=0}^{n} f_k'\right)$ konvergiere gleichmäßig auf D(es genügt gleichmäßig auf jedem kompakten Teilintervall). Dann konvergiert die Folge (F_n) gegen eine differenzierbare Funktion $F : D \to \mathbb{R}$ und es gilt

$$F' = \left(\sum_{k=0}^{\infty} f_k\right)' = \sum_{k=0}^{\infty} f_k'.$$

Ist D kompakt, dann konvergiert (F_n) gleichmäßig gegen F. Die wichtigste Anwendung dieses Satzes betrifft Potenzreihen:

20.4.3 Korollar

Ist $\sum\limits_{k=0}^{\infty} a_k(x-a)^k$ eine Potenzreihe mit Konvergenzradius $r > 0$, $D := U_r(a)$ und ist

$$\begin{aligned} p : D &\to \mathbb{R} \\ x &\mapsto \sum_{k=0}^{\infty} a_k(x-a)^k \end{aligned}$$

die durch die Potenzreihe dargestellte Funktion, dann gilt

$$p'(x) = \sum_{k=1}^{\infty} k a_k (x-a)^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} (x-a)^k.$$

Die Ableitung einer Potenzreihe erhält man also (wie bei Polynomen) durch gliedweise Differentiation der Reihe.

Zusatz: $q : D \to \mathbb{R}$ mit $q(x) := \sum\limits_{k=0}^{\infty} \frac{a_k}{k+1}(x-a)^{k+1}$ ist eine Stammfunktion von p.

Beweis: Wir wissen, dass die formal differenzierbare Reihe $\sum\limits_{k=1}^{\infty} ka_k(x-a)^{k-1}$ den gleichen Konvergenzradius wie die Ausgangsreihe hat (das folgt auch einfach aus der Cauchy-Hadamardschen Formel für den Konvergenzradius). Daher erfüllt p die Voraussetzungen des Korollars, weil jedes $x \in U_r(a)$ in einem abgeschlossenen Teilintervall $[a-\varrho, a+\varrho] \in U_r(a)$, in welchem die formal abgeleitete Reihe gleichmäßig konvergiert ($0 < \varrho < r$ geeignet).
Durch mehrfache Anwendung folgt $p \in C^\infty(D)$ und für $k \in \mathbb{N}_0$ gilt

$$\boxed{a_k = \frac{f^{(k)}(a)}{k!}}$$

Hieraus kommen wir im nächsten Abschnitt zurück.

Da auch $\sum \frac{a_k}{k+1}(x-a)^{k+1}$ den Konvergenzradius r hat, gilt auch der Zusatz.

Durch Anwendung des Korollars erhält man völlig neue Beweise für die Ableitungen von exp, cos, sin, cosh, sinh, etc.

20.4.4 Beispiele und Bemerkungen

(a) Wegen $\exp(x) = \sum\limits_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^k}{k!} + \ldots$ folgt

$$\begin{aligned} \exp'(x) &= 1 + \frac{2}{2!}x + \ldots + \frac{k+1}{(k+1)!}x^{k-1} + \ldots \\ &= 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^k}{k!} + \ldots \\ &= \exp(x) \end{aligned}$$

Analog folgt für $x \in \mathbb{R}$

$$\sin' x = \cos x,\ \cos' x = -\sin x,\ cosh' x = \sinh x \text{ und } sinh' x = \cosh x.$$

(b) Beim Beweis der Ableitungsformel für exp wurde die *Funktionalgleichung* ($x, y \in \mathbb{R}$) $\exp(x + y) = \exp(x)\exp(y)$ wesentlich benutzt.
Weiß man jetzt über die gliedweise Differentiation der Reihe, dass $\exp' = \exp$ gilt, so kann man die Funktionalgleichung (das Additionstheorem) so beweisen:
Bei festem, aber beliebigen $y \in \mathbb{R}$, betrachte man

$$\begin{aligned} g : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \exp(x+y)\exp(-x) \end{aligned}$$

Nach der Produktregel ist

$$g'(x) = \exp(x+y)\exp(y) - \exp(x+y)\exp(-x) = (\exp(x+y) - \exp(x+y))\exp(-x) = 0.$$

Da $\mathbb{R}$ ein Intervall ist, ist also $g(x) = g(0) = \exp(y)$, also

$$\boxed{\exp(x+y) = \exp(x)\exp(y)}$$

Damit ist die Funktionalgleichung (das Additionstheorem) von exp mit Hilfe der Differentialgleichung von exp bewiesen.
Man beachte, dass wir in einen circuluc vitiosus geraten wären, wenn wir $\exp' = \exp$ nicht mit Hilfe der gliedweisen Differentiation der Reihe bewiesen hätten.

20.5 Taylor'sche Formel, Taylor-Reihen, Abelscher Grenzwertsatz

Wir haben die Funktionen exp, sin, cos, sinh, cosh etc durch Potenzreihen eingeführt (sogar für komplexe Argumente). Potenzreihen sind offensichtliche Verallgemeinerungen von Polynomen,, diese wiederum sind sehr gut handhabbare Funktionen mit angenehmen Eigenschaften (stetig, beliebig oft differenzierbar). In diesem Abschnitt beschäftigen wir uns systematisch mit der Entwicklung von Funktionen in Potenzreihen. Wir werden sehen: Ist eine Funktion überhaupt in eine Potenzreihe entwickelbar, dann ist diese Potenzreihe notwendig die Taylor-Reihe, zunächst zur Erinnerung.
Ist $f : D \to \mathbb{R}$ ($D \subset \mathbb{R}$ ein echtes Intervall) $a \in D$ und f in a differenzierbar, dann approximiert die Tangente

$$\begin{aligned} T : D &\to \mathbb{R} \\ x &\mapsto f(a) + f'(a)(x-a) \end{aligned}$$

an den Graphen von f die Funktion f in einer Umgebung von a so gut, dass sogar

$$(*) \qquad \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - T(x)}{x - a} = 0$$

gilt.
Ferner ist $f(a) = T(a)$ und $f'(a) = T'(a)$. $T_1 := T$ ist ein Polynom von Grad ≤ 1. Wir wollen in Verallgemeinerung dieser Situation von folgender *Problemstellung* ausgehen:
Gegeben sei eine echtes Intervall $D \subset \mathbb{R}$, eine hinreichend oft differenzierbare Funktion, sagen wir n-mal differenzierbare Funktion ($n \in \mathbb{N}$). Gesucht ist ein Polynom $T_n : \mathbb{R} \to \mathbb{R}$ vom Grad $\leq n$ mit der folgenden Eigenschaft:

$$\begin{aligned} (**) \qquad & T_n(a) = f(a) = f^{(0)}(a), \\ & T_n'(a) = f'(a), \ldots, T^{(n)}(a) = f^{(n)}(a). \end{aligned}$$

Wir machen für das Polynom T_n versuchsweise den Ansatz

$$T_n(x) = \sum_{k=0}^{n} a_k (x-a)^k \qquad (a_k \in \mathbb{R};\ 0 \leq k \leq n)$$

weil wir f in einer Umgebung von a approximieren wollen.
Wegen $T_n^{(k)}(a) = k! a_k$ für $0 \leq k \leq n$, ergibt sich, dass die a_k durch f eindeutig bestimmt sind: $a_k = \frac{T_n^{(k)}(a)}{k!} = \frac{f^{(k)}(a)}{k!}$, $k = 0, \ldots, n$.

20.5.1 Satz und Definition

Es gibt genau ein Polynom $T_n : \mathbb{R} \to \mathbb{R}$ vom Grad $\leq n$, das $(**)$ erfüllt ist, nämlich

$$\begin{aligned} T : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto T_n(x) := T_{n,a}(f)(x) := f(a) + f'(a)(x-a) + \ldots + \frac{f^{(k)}(a)}{k!}(x-a)^k = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \end{aligned}$$

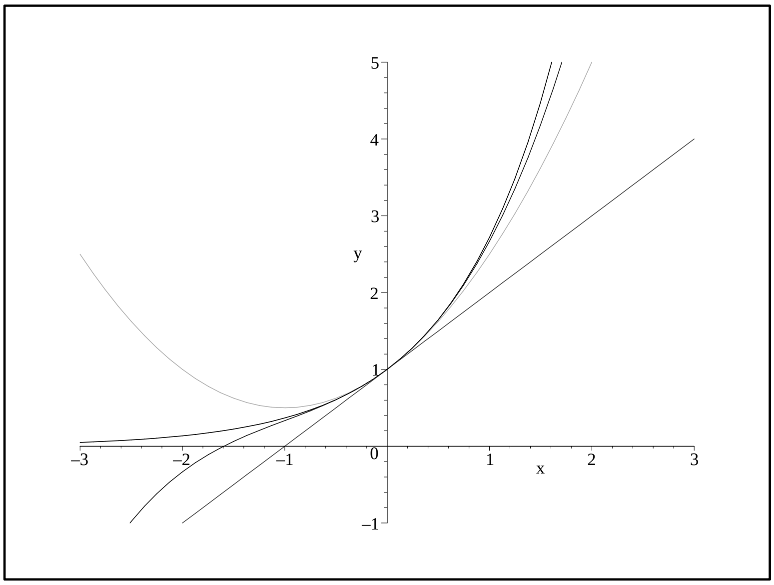


Abbildung 21: Schmiegparabeln der Grade 1, 2, 3 der Exponentialfunktion am Punkt 0

T_n heißt *n-tes Taylor-Polynom* von f zum Punkt a. (B. Taylor 1685-1721, Schüler von I.Newton)
$a_b := \frac{f^{(k)}(a)}{k!}$ heißt der k-te Taylor-Koeffizient von f im Punkt a.
Wir werden sehen, dass für den Fall, dass f n-mal stetig differenzierbar ist, in Verallgemeinerung von $(*)$ gilt

$$\lim_{x \to a} \frac{f(x) - T_{n,a}(a)}{(x-a)^n} = 0$$

Das Taylor-Polynom T_1 ist die Tangente an den Graphen von f, das Taylor-Polynom T_2 eine Parabel, falls $f''(a) \neq 0$. Den Graphen von $T_{n,a}(f)$ nennt man auch Schmiegeparabel n-ten Grades von f in a. (siehe Abb.21, Abb. 22 und Abb. 23)
Um die Güte der Approximation von f durch $T_{n,a}(f)$ zu messen, bezeichnen wir die Differenz (Abweichung) $f - T_{n,a}(f)$ mit R_{n+1}:

$$R_{n+1}(x) = f(x) - T_{n,a}(f)(x).$$

Ob und in wieweit Taylor-Polynome als brauchbare Näherungen für f (wenigstens „nahe bei a") zu betrachten sind, kann erst durch Analyse dieses Fehlers beantwortet werden. Es gilt nun

20.5.2 Satz (Taylor'sche Formel mit Integralrestglied)

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f \in C^{n+1}(D)$ (f also $n+1$-mal stetig differenzierbar), $a \in D$. Dann gilt für alle $x \in D$

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!} + R_{n+1}(x)$$

mit

$$R_{n+1}(x) = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t)dt,$$

also $f(x) = T_{n,a}(f)(x) + R_{n+1}(x)$.

Bemerkung: Der Rest R_{n+1} hängt natürlich auch (wie das Taylor-Polynom) von f und der Stelle a.

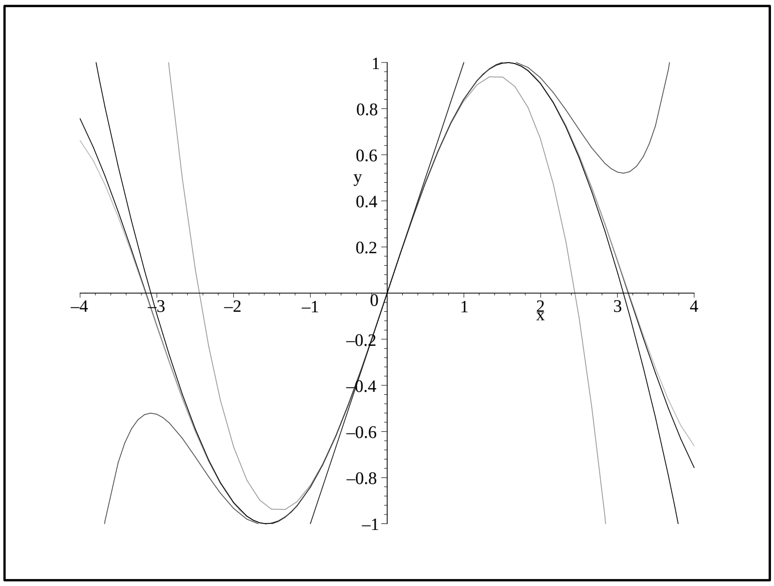


Abbildung 22: Schmiegparabeln der Grade 1, 3, 5, 7, 9 des Sinus am Punkt 0

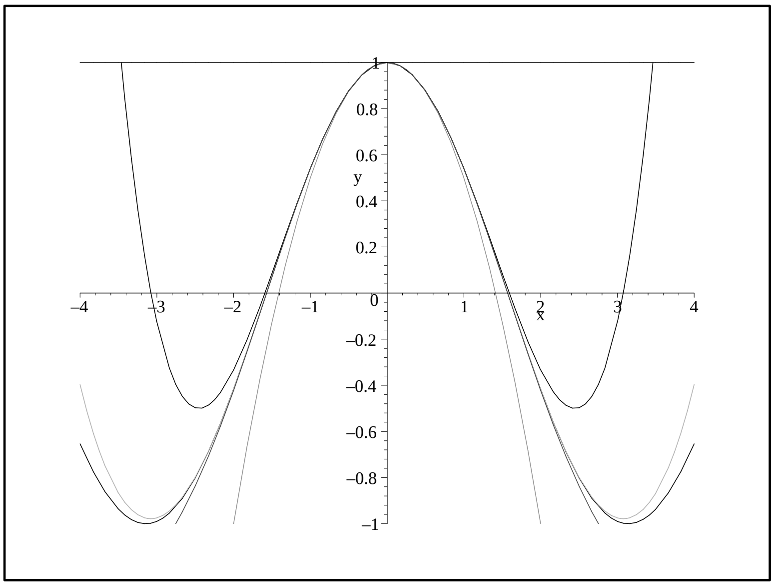


Abbildung 23: Schmiegparabeln der Grade 0, 2, 4, 6, 8 des Cosinus am Punkt 0

Diese Abhängigkeit bringen wir jedoch in der Notation (im Gegensatz zum Taylor-Polynom) nicht extra zum Ausdruck.

Beweis von 20.5.2: Im Fall $n = 0$ lautet die Aussage $f(x) = f(a) + \int_a^x f'(t)dt$, das ist genau die Aussage des Hauptsatzes (Version 2 für stetig differenzierbare Funktionen f). Es bietet sich also ein Induktionsbeweis mit der Induktionsverankerung $n = 0$ an.
Schluss von $n-1$ auf n:
Wir nehmen an, dass die Formel für $n-1$ gilt:

$$(*) \qquad f(x) = T_{n-1,a}(f)(x) + \underbrace{\frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f^{(n)}(t)dt}_{R_n(x)}$$

Mit partieller Integration folgt

$$R_n(x) = -\left.\frac{(x-t)^n}{n!} f^{(n)}(t)\right|_a^x + \underbrace{\frac{1}{n!}\int_a^x (x-a)^n f^{(n-1)}(t)dt}_{R_{n+1}(x)} = \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n+1}(x).$$

Setzt man diesen Wert für $R_n(x)$ in die Darstellung $(*)$ ein, so folgt

$$f(x) = T_{n-1,a}(f)(x) + \frac{f^n(a)}{n!}(x-a)^n + R_{n+1}(x) = T_{n,a}(f)(x) + R_{n+1}(x)$$

und das ist die Aussage für n.

□

Obwohl man an der Integralform des Restgliedes schon einiges ablesen kann, ist die folgende Form des Restgliedes für Anwendungen häufig günstiger.

20.5.3 Satz (Lagrange'sche Form des Restgliedes)

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f \in C^{n+1}(D)$, $a, x \in D$. Dann gibt es ein ξ mit $\xi \in [a, x]$ oder $\xi \in [x, a]$ mit

$$f(x) = T_{n,a}(f)(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

Nach dem Mittelwertsatz der Integralrechnung (vgl. ???) existiert ein ξ, so dass gilt

$$R_{n+1}(x) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)dt = f^{(n+1)}(\xi)\int_a^x \frac{(x-t)^n}{n!}dt = -f^{(n+1)}(\xi)\left[\frac{(x-t)^{n+1}}{(n+1)!}\right]_a^x = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

Man kann die Voraussetzung noch abschwächen.

20.5.4 Satz

Sei $f \in C^n(D)$ und $f^{(n)} : D \to \mathbb{R}$ noch differenzierbar in alle inneren Punkten von D. Dann gibt es zu $a \in D$ und $x \in D$ ein ϑ mit $0 < \vartheta < 1$, so dass gilt

$$f(x) = T_{n,a}(f)(x) + \frac{f^{(n+1)}(a + \vartheta(x-a))}{(n+1)!}(x-a)^{n+1}$$

Beweis : Man betrachte für festes x

$$g(u) = f(x) - f(u) - f'(u)(x-u) - \ldots - \frac{f^{(n)}(u)}{n!}(x-u)^n - \sigma\frac{1}{(n+1)!}(x-n)^{n+1}$$

wobei die Konstante $\sigma \in [a,x]$ bzw $[x,a]$ gewählt sei und so bestimmt wird, dass $g(a) = 0$ gilt. Da außerdem $g(x) = 0$, gibt es nach dem Satz von Rolle ein $\vartheta \in]0,1[$ mit $g'(a_0 + \vartheta(x-a)) = 0$. Andererseits folgt

$$g'(u) = -\frac{1}{n!}f^{(n+1)}(u)(x-u)^n + \frac{\sigma}{n!}(x-u)^n$$

und damit folgt $\sigma = f^{(n+1)}(a_0 + \vartheta(x-a))$.

□

Man beachte, dass dieser Satz im Fall $n = 0$ gerade der MWSD ist.

20.5.5 Folgerung

Ist $f \in C^n(D)$ und $f^{(n+1)} = 0$, dann ist f ein Polynom vom Grad $\leq n$.
Denn das Restglied im letzten Satz ist dann immer Null und f reduziert sich auf das Taylor-Polynom von Grad $\leq n$.

> Alle Polynome vom Grad $\leq n$ sind also Lösungen der Differentialgleichung $f^{(n+1)} = 0$ (sogar auf ganz $\mathbb{R}$)

Die verschiedenen Darstellungen des Restgliedes werden wir für Abschätzungen der Größe des Fehlers als auch (speziell die Lagrange Form) zur Bestimmung des Vorzeichens des Fehlers benutzen.

20.5.6 Beispiele und Bemerkungen

(a) Es gilt $\sin, \cos \in C^\infty(D)$ und für $a = 0$ ist

$$\sin x = x - \frac{x^3}{3!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + R_{2n+3}(x)$$

mit

$$R_{2n+3}(x) = \frac{\sin^{2n+3}(\xi)}{(2n+3)!}x^{2n+3} = (-1)^{n+1}\frac{\cos\xi}{(n+1)!}x^{2n+3}$$

mit ξ zwischen 0 und x, und damit gilt für *alle* $x \in \mathbb{R}$

$$\left|\sin x - \sum_{k=0}^{n}(-1)^k\frac{x^{2k+1}}{(2k+1)!}\right| \leq \frac{|x|^{2n+3}}{(2n+3)!}$$

Als einfaches Beispiel berechnen wir den relativen Fehler $r(x)$, wenn man im Intervall $]0, \frac{\pi}{4}]$ den Sinus durch des Taylor-Polynom $P(x) := T_3(x) := x - \frac{x^3}{3!}$ approximiert. Es ist (beachte $\sin x > 0$)

$$0 < r(x) = \frac{\sin x - P(x)}{\sin x} \leq \frac{x^5}{5! \sin x}.$$

Für $0 < x \leq \frac{\pi}{4}$ ist aber

$$\sin x \geq x - \frac{x^3}{3!} = x(1 - \frac{x^2}{3!}) \geq x(1 - \frac{\pi^2}{4^2 \cdot 3!}),$$

also

$$0 < r(x) \le \frac{x^4}{5!(1-\frac{\pi^2}{4^2\cdot 3!})} \le \frac{\pi^4}{4^4\cdot 5!}\frac{1}{(1-\frac{\pi^2}{4^2\cdot 3!})} < 0,0036 \text{ (Taschenrechner oder Maple)}$$

Also ist der relative Fehler kleiner als $0,36\%$.

Analog gilt für *alle* $x \in \mathbb{R}$:

$$\left|\cos x - \sum_{k=0}^{n}(-1)^k \frac{x^{2k}}{(2k)!}\right| \le \frac{|x|^{2n+2}}{(2n+2)!}$$

(b) Für $\exp \in C^\infty(\mathbb{R})$ und $a = 0$ folgt

$$e^x := \exp(x) = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!}e^{\vartheta x}$$

mit $0 < \vartheta < 1$.
Will man etwa $\exp$ im Intervall $[-1,1]$ durch ein Polynom bis auf 2 Stellen nach dem Komma genau approximieren, so leistet das Taylor-Polynom T_5 das Gewünschte:
So ist nämlich

$$P(x) := T_{5,0}(\exp)(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120}$$

und damit

$$|e^x - P(x)| = \frac{x^6}{720}e^{\vartheta x} < \frac{e}{720} < 0,0038.$$

Aus der obigen Darstellung von $\exp x$ ergibt sich für $x = 1$ auch die Abschätzung

$$0 < e - E_n < \frac{3}{(n+1)!}$$

für alle $n \in \mathbb{N}_0$, dabei ist $E_n = \sum\limits_{k=0}^{n} \frac{1}{k!}$.
Aus dieser Abschätzung ergibt sich (vgl. ???) leicht die *Irrationalität* von e.

Die hinreichende Kriterien für lokale Extrema oder Wendepunkte (vgl. ???) versagen schon in einfachen Beispielen: $P_4(x) = x^4$ oder $P_5(x) = x^5$

Mit Hilfe des Lagrange'schen Restgliedes erhält man folgendes hinreichende Kriterium für lokale Extrema.

20.5.7 Satz (hinreichendes Kriterium für lokale Extrema)

Sei $D \subset \mathbb{R}$ ein echtes Intervall, $f \in C^{n+1}(D)$. In einem Punkt $a \in D$ gelte $f'(a) = f''(a) = \ldots = f^{(n)}(a) = 0$, jedoch $f^{(n+1)}(a) \ne 0$. Dann hat f in a

(a) ein *strenges (striktes) lokales Minimum*, falls n ungerade und $f^{(n+1)}(a) < 0$ gilt;
(b) ein strenges *lokales Maximum*, falls n un gerade und $f^{(n+1)}(a) < 0$,
(c) kein Extremum, falls n gerade ist.

Man hat die Differenz $f(x) - f(a)$ zu betrachten. Zunächst besagt die Voraussetzung

$$f(x) - f(a) = \frac{f^{(n+1)(\xi)}}{(n+1)!}(x-a)^{n+1} = R_{n+1}(x)$$

Wir betrachten den Fall $f^{n+1}(a) > 0$ und wählen ein möglicherweise kleines Intervall $U_\delta(a)$, so dass für alle $x \in U_\delta(a) \cap D$ gilt $f^{(n+1)}(x) > 0$, insbesondere ist $f^{(n+1)}(\xi) > 0$.

Ist nun n ungerade, also $n+1 = 2k$ gerade, dann gilt für $x \in U_\delta(a) \cap D$ immer $(x-a)^{2k \geq 0}$, wobei das Gleichheitszeichen nur für $x = a$ gilt. Es ist also $R_{n+1}(x) > 0$ für alle $x \in U_\delta(a);\ x \neq a$. An der Stelle a liegt also ein strenges lokales Minimum vor. Die Aussage für den Maximum erhält man durch Übergang zu $-f$.
Ist jedoch n gerade, also $n+1$ ungerade, so ist $R_{n+1}(x) > 0$ für $x > a$ und $R_{n+1}(x) < 0$ für $x < a$; In a kann in diesem fall weder ein Maximum noch ein Minimum vorliegen.

□

Wir beweisen nun die angekündigte Verallgemeinerung der Formel $(*)$ aus der Einführung zu diesem Abschnitt:

20.5.8 Satz (qualitative Form der Taylorschen Formel)

Ist $D \subset \mathbb{R}$ ein echtes Intervall und $f \in C^n(D)$, $n \in \mathbb{N}$, $a \in D$. Dann gilt für alle $x \in D$

$$(*) \qquad f(x) = T_{n,a}(f)(x) + (x-a)^n \varrho(x),$$

dabei ist $\varrho : D \to \mathbb{R}$ eine in a stetige Funktion mit $\varrho(a) = 0$.

Man sagt hierfür auch: f und $T_{n,a}f$ stimmen in a in n-ten Ordnung überein.
Zum Beweis gehen wir von (Lagrangesches Restglied)

$$\begin{aligned} f(x) - T_{n-1,a}(f)(x) &= f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k \\ &= \frac{f^{(n)}(\xi)}{n!}(x-a)^n \end{aligned}$$

aus.
Man beachte: wir haben nur $f \in C^n(D)$ vorausgesetzt.
Dann folgt

$$f(x) - T_{n-1,a}(f)(x) = \frac{f^{(n)}(a)}{n!}(x-a) + \underbrace{\frac{f^{(n)}(\xi) - f^{(n)}(a)}{n!}}_{\varrho(x)}(x-a)^n$$

Da ξ zwischen a und x liegt, also die Gestalt $\xi = \xi(x) = a + \vartheta(x-a)$ hat mit $0 \leq \vartheta \leq 1$, folgt wegen der Stetigkeit von $f^{(n)}$

$$\lim_{\substack{x \to a \\ x \neq a}} \varrho(x) = \lim_{x \to a} \frac{f^{(n)}(\xi(n)) - f^{(n)}(a)}{n!} = 0,$$

d.h. ϱ ist stetig in a und $\varrho(a) = 0$.

20.5.9 Beispiele und Bemerkungen

(a) Die Gleichung $(*)$ in 20.5.(8) ist äquivalent mit

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - T_{n,a}(f)(x)}{(x-a)^n} = 0$$

Im Fall $n = 1$ erhält man also einfach die Charakterisierung der Differezierbarkeit durch die lineare Approximierbarkeit. Das Taylor-Polynom $T_{n,a}(f)$ ist durch $(*)$ charakterisiert, denn es gilt:

Ist $p_n : \mathbb{R} \to \mathbb{R}$ ein Polynom vom Grad $\leq n$, mit $\lim\limits_{\substack{x \to a \\ x \neq a}} \frac{f(x)-p_n(a)}{(x-a)^n} = 0$, dann ist $p_n = T_{n,a}(f)$, denn aus der Voraussetzung folgt dann auch

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{T_{n,a}(f)(x) - p_n(x)}{(x-a)^n} = 0$$

das gehr aber nur wenn $T_{n,a}(f) - p_n$ das Nullpolynom ist.

(b) Häufig verwendete *Näherungsformeln* beruhen auf der qualitativen Form des Taylorschen Satzes.

(α) $\sqrt{1+x} \approx 1 + \frac{x}{2} - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4$
(β) $\frac{1}{\sqrt{1+x}} \approx 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{6}{16}x^8$
(γ) $\frac{1}{1-x} \approx 1 + x$ ($|x|$ immer hinreichend klein).

Häufig benügt man sich z.B. im Fall (α) schon mit $\sqrt{1+x} = 1 + \frac{x}{2} + \varrho(x)x \quad (\lim\limits_{x\to 0} \varrho(x) = 0)$ (hier wird

$$\begin{array}{rcl} f:]-1,1[& \to & \mathbb{R} \\ x & \mapsto & \sqrt{1+x} \end{array}$$

betrachtet, $a = 0$)

Wir benutzen die obige Formel, um ganz einfach

$$\boxed{\lim_{n\to\infty} (\sqrt{n+\sqrt{n}} - \sqrt{n}) = \tfrac{1}{2}}$$

zu zeigen.
Für $n > 1$ ist nämlich

$$\begin{aligned} \sqrt{n+\sqrt{n}} = \sqrt{n(1+\frac{1}{\sqrt{n}}} &= \sqrt{n}\sqrt{1+\frac{1}{\sqrt{n}}} \\ &= \sqrt{n}(1+\frac{1}{2\sqrt{n}} + \varrho(\frac{1}{\sqrt{n}})\frac{1}{\sqrt{n}}) \\ &= \sqrt{n} + \frac{1}{2} + \varrho(\frac{1}{\sqrt{n}}). \end{aligned}$$

Hieraus folgt aber wegen $\lim\limits_{n\to\infty} \varrho(\frac{1}{\sqrt{n}} = \varrho(0) = 0$

$$\lim_{n\to\infty} (\sqrt{n+\sqrt{n}} - \sqrt{n}) = \frac{1}{2}.$$

Ist nun $D \subset \mathbb{R}$ ein echtes Intervall und $f \in C^\infty(D)$, $a \in D$, dann ist ja insbesondere $f \in C^n(D)$ für jedes $n \in \mathbb{N}_0$ und man kann für jedes n das n-te Taylor-Polynom $T_{n,a}(f)(x) = \sum\limits_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$ betrachten und fragen: Existiert $\lim\limits_{n\to\infty} T_{n,a}(f)(x) = \lim\limits_{n\to\infty} \sum\limits_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$?
Das führt zu folgender Definition:

20.5.10 Definition (Taylor-Reihe von $f \in C^\infty(D)$

Ist $D \subset \mathbb{R}$ ein echtes Intervall, $f \in C^\infty(D)$, $a \in D$. Dann heißt die Potenzreihe $T_{\infty,a}(f)(x) := \sum\limits_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$ die Taylor-Reihe von f im Punkt a (oder zum Entwicklungspunkt a).
Dabei stellt sich die Frage:
Konvergiert die Taylor-Reihe -wenn sie denn konvergiert- immer gegen f ?

20.5.11 Bemerkungen und Beispiele

(a) Auf Grund des Taylor'schen Formel mit Restglied konvergiert die Taylorreihe $T_{\infty,a}(f)(x)$ für $x \in D$ genau dann gegen $f(x)$, wenn $\lim\limits_{n\to\infty} R_{n+1}(x) = 0$. Oder äquivalent: Die Folge $(T_{n,a}(f))$ der Taylor-Polynome konvergiert genau dann und $T_{\infty,a}f$ stellt dann die Funktion f ander Stelle x dar, wenn die Folge $(R_{n+1}(x)$ der Reste eine Nullfolge ist.

(b) Weiss man, dass f irgendwie um a in eine Potenzreihe entwickelt werden kann, gibt es also eine Potenzreihe $\sum a_k(x-a)^k$ mit positivem konvergenzradius $r > 0$, so dass für alle $x \in U_r(a) \cap D$ gilt

$$f(x) = \sum_{k=0}^{\infty} a_k(x-a)^k,$$

dann folgt aus dem Satz über die gliedweise Differenzierbarkeit einer Potenzreihe für $k \in \mathbb{N}_0$,

$$\boxed{a_k = \frac{f^{(k)}(a)}{k!}}$$

d.h. die Potenzreihe stimmt mit der Taylor-Reihe überein. Es gillt also $T_{\infty,a}(f)(x) = f(x)$ und unsere obige Folge ist positiv zu beantworten.

(c) Eine Potenzreihe kann aber auch nur für den Entwicklungsopunkt a konvergieren (Konvergenzradius Null). Es kann tatsächlich vorkommen, dass die Taylorreihe einer Funktion $f \in C^\infty(D)$ nur für den Entwicklungspunkt a konvergiert. Ein solches Beispiel wird z.B. durch die Funktion $f \in C^\infty(\mathbb{R})$ mit $(x \in \mathbb{R}, a = 0)$

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{2^k} \cos(kx)$$

geliefert. (vgl. z.B. Barner-Flohr: Analysis 1, 9.5)

(d) Für die Funktion $f \subset C^\infty(\mathbb{R})$ mit

$$f(x) = \begin{cases} \exp(-\frac{1}{x^2}, & \text{für } x \neq 0 \\ 0, & \text{für } x = 0 \end{cases}$$

gilt (vgl. Übungsaufgabe ??? vom Blatt ???) $f \in C^\infty(\mathbb{R})$ und $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}_0$.

$T_{\infty,0}f(x)$ konvergiert also für alle $x \in \mathbb{R}$, aber nur für $x = 0$ wird die Funktion dargestellt. Obwohl hier der Konvergenzradius der Taylor-Reihe positiv ist, stellt sie nur für den Entwicklungspunkt die Funktion dar.
Aus der Bemerkung (b) folgt der bemerkenswerte

20.5.12 Eindeutigkeitssatz (Identitätssatz) für Potenzreihen

Konvergieren die Potenzreihen

$$\sum_{k=0}^{\infty} a_k(x-a)^k \; und \; \sum b_k(x-a)^k$$

in $U_R(a)$ $(R > 0)$ und stellen sie dort dieselbe Funktion dar, dann gilt $a_k = b_k$ für alle $k \in \mathbb{N}_0$.

Beweis : Die dargestellte Funktion sei etwa f. Dann gilt nach (b) einerseits $\frac{f^{(k)}}{k!} = a_k$, aber auch $\frac{f^{(k)}(a)}{k!} = b_k$ für alle $k \in \mathbb{N}_0$, d.h. $a_k = b_k$.

□

Die beiden wichtigsten Interpretationen des Satzes sind:

(1) Wenn es überhaupt möglich ist eine Funktion in $U_R(a) =]a-R, a+R[$ als Potenzreihe mit Entwicklungspunkt a darzustellen, dann nur als Taylor-Reihe:

$$f(x) = T_{\infty,a}(f)(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

(2) Wird eine Funktion f auf zwei verschieden Weisen als Potenzreihe mit Entwicklungspunkt a dargestellt, dann sind die Koeffizienten entsprechender $(x-a)^n$-Potenzen gleich.
In dieser Form nennt man Satz auch *Satz von Koeffizientenvergleich*.
Mit ihm gelangt man häufig zu nicht-trivialen Beziehungen.
Der Satz enthält auch als Spezialfall den Identitätssatz für Polynome.

20.5.13 Satz (Hinreichnende Bedingung für die Darstellbarkeit einer C^∞-Funktion durch ihre Taylor-Reihe)

Ist $f \in C^\infty(D)$, $a \in D$ und gibt es Konstanten A und B, so dass für alle $x \in D$ und alle $n \in \mathbb{N}_0$ die Abschätzung $|f^{(n)}(x)| \le AB^n$ gilt, dann gilt

$$f(x) = T_{\infty,a}(f)(x)$$

Unter dieser Voraussetzung ist nämlich

$$|R_{n+1}(x)| \le \frac{AM^{n+1}|x-a|^{n+1}}{(n+1)!} = A\frac{(B|x-a|)^{n+1}}{(n+1)!}$$

und die Folge rechts ist eine Nullfolge.

Zusatz: Gilt sogar $|f^{(n)}(x)| \le M$ für alle $x \in D$ und alle $n \in \mathbb{N}_0$, dann gilt $\lim\limits_{n\to\infty} R_{n+1}(x) = 0$.

20.5.14 Weitere Beispiele für Taylor-Reihen

(a) Die C^∞-Funktionen sin, cos, sinh, cosh, exp wurden durch ihre Taylor-Reihe zum Entwicklungspunkt 0 eingeführt.
Es gilt aber für *alle* $a \in \mathbb{R}$, z.B. $\sin x = T_{\infty,a}(\sin)(x)$ bzw. $\cos x = T_{\infty,a}(\cos)(x)$ bzw.

$$\begin{aligned} \exp(x) = T_{\infty,a}(\exp)(x) &= \sum_{k=0}^{\infty} \frac{\exp^{(k)}(a)}{k!}(x-a)^k \\ &= \sum_{k=0}^{\infty} \frac{\exp(a)}{k!}(x-a)^k. \end{aligned}$$

sin, cos, exp werden also für alle $a \in \mathbb{R}$ durch ihre Taylor-Reihe dargestellt und zwar auf ganz $\mathbb{R}$.
Denn es gilt $|\sin^{(k)}(x)| \le 1$ und $|\cos^{(k)}(x)| \le 1$ für alle $x \in \mathbb{R}$ und alle $k \in \mathbb{N}_0$, daher folgt aus dem Zusatz von 20.5.13 die Behauptung.
Für die Exponentialfunktion kann man analog schließen:
Man nimmt die Lagrange'sche Form des Restgliedes $R_{n+1}(x) = \frac{\exp^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$, wobei $\xi \in [a,x]$ oder $\xi \in [x,a]$ gilt. Ist dann $k = \max\{a,x\}$, so ist

$$|\exp^{(k)}(\xi)| = |\exp(\xi)| \le \exp(b) =: M,$$

also ist,

$$|R_{n+1}(x)| \leq \frac{M|x-a|^{n+1}}{(n+1)!}$$

und damit $\lim\limits_{n\to\infty} R_{n+1}(x) = 0$ für alle $x \in \mathbb{R}$.
Im Fall der Exponentialfunktion hätte man auch mittels der Funktionalgleichung auf $\exp(x) = T_{\infty,a}(\exp)(x)$ für alle $a \in \mathbb{R}$ und alle $x \in \mathbb{R}$ schließen können:
Es ist ja

$$\begin{aligned}
\exp(x) &= \exp(a)\exp(x-a) \\
&= \exp(a)\sum_{k=0}^{\infty}\frac{(x-a)^k}{k!} \qquad \text{(nach Def. von exp)} \\
&= \sum_{k=0}^{\infty}\exp(a)\frac{(x-a)^k}{k!} \\
&= \sum_{k=0}^{\infty}\frac{\exp^{(k)}(a)}{k!}(x-a)^k = T_{\infty,a}(\exp)(x).
\end{aligned}$$

(b) In Spezialfällen kann man ohne Benutzung der Darstellungen in ??? das Restglied erhalten: Sei dazu

$$\begin{aligned}
f:\ \mathbb{R}-\{1\} &\to \mathbb{R} \\
x &\mapsto \frac{1}{1-x}
\end{aligned}$$

Nach der Summenformel für die geometrische Reihe ist für $|x| < 1$

$$f(x) = \sum_{k=0}^{\infty} x^k = T_{\infty,0}(f)(x).$$

Daher $P_n(x) := T_{n,0}(f)(x) = \sum\limits_{k=0}^{n} x^k$.
Statt aber das Restglied aus der Potenzreihendarstellung abzulesen, was dieses nur für $|x| < 1$ liefern würde, beachten wir, dass für $x \neq 1$ gilt

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

Es ist also für alle $x \neq 1$ (d.h. in $]-\infty, 1[$ bzw. $]1,\infty[$)

$$R_{n+1}(x) = \frac{x^{n+1}}{1-x}.$$

(c) Für $x > -1$ sei $f(x) = \log(1+x)$ (beachte $f'(x) = \frac{1}{1+x}$).
Wir wollen die Taylor-Reihe von f zum Entwicklungspunkt $a = 0$ aufstellen und nach Möglichkeit nur wenige Ableitungen ausrechnen.
Dazu beachten wir, dass nach dem Hauptsatz gilt

$$f(x) = \int_0^x \frac{1}{1+t}dt.$$

Für $|x| < 1$ gilt daher

$$f(x) = \log(1+t)|_0^x = \int_0^x \frac{1}{1+t}dt = \int_0^x \left(\sum_{k=0}^{\infty}(-1)^k t^k\right) dt.$$

Nun konvergiert $\sum\limits_{k=0}^{\infty}(-1)^k t^k$ gleichmäßig auf $[-|x|,|x|]$. Mit ??? folgt daher

$$f(x) = \log(1+x) = \sum_{k=0}^{\infty}(-1)^k \int_0^x t^k\,dt = \sum_{k=0}^{\infty}\frac{(-1)^k}{k+1}x^{k+1}$$

Damit haben wir: Für $|x| < 1$ gilt

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

Da die Reihe rechts auch noch (nach dem Leibniz-Kriterium) für $x = 1$ konvergiert, ist zu vermuten, dass die Darstellung auch noch für $x = 1$ gilt, dass also

$$\boxed{\log 2 = \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \mp \ldots}$$

gilt.
Dazu benutzen wir die Taylorsche Formel mit dem Lagrange'schen Restglied ($0 \leq \vartheta \leq 1$)

$$\log(1+x) = \sum_{k=1}^{n}(-1)^{k-1}\frac{x^k}{k} + \underbrace{\frac{(-1)^n}{n+1}\frac{1}{(1+\vartheta x)^{n+1}}x^{n+1}}_{R_{n+1}(x)}$$

Für $0 \leq x \leq 1$ ist dann $1 + \vartheta x \leq 1$, also $0 \leq \frac{x}{1+\vartheta x} \leq 1$ und damit $|R_{n+1}(x)| \leq \frac{1}{n+1}$, also $\lim\limits_{n\to\infty} R_{n+1}(x) = 0$, speziell ist $\lim\limits_{n\to\infty} R_{n+1}(1) = 1$.

Fassen wir zusammen:

20.5.15 Satz

Für alle $x \in]-1,1]$ gilt

$$\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k,$$

Speziell gilt:

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \mp \ldots \approx 0$$

20.5.16 Bemerkung

Da $\lim\limits_{\substack{x \to 1 \\ x > -1}} \log(1+x) = -\infty$ gilt und die rechte Seite für $x \to -1$ $(x > -1)$ ebenfalls diesen uneigentlichen Grenzwert hat, gilt die Darstellung bei großzügiger Interpretation auch noch für $x = -1$.

Auf die Konvergenz der Taylor-Reihe im Punkt $x = 1$ hätte man auch mit Hilfe des folgenden sehr nützlichen Satzes schließen können.

20.5.17 Satz (Abelscher Grenzwertsatz, N.H.Abel, 1826)

Die Potenzreihe $\sum a_k x^k$ besitze den endlichen und positiven Konvergenzradius r. Die Reihe $\sum a_k x^k$ sei noch für $x = r$ konvergent. Dann ist die durch die Potenzreihe dargestellte Funktion $f :]-r, r] \to \mathbb{R}$ in r noch (rechtsseitig) stetig, d.h. es gilt

$$\lim_{\substack{x \to r \\ x < r}} f(x) = \sum_{k=0}^{\infty} a_k r^k =: f(r).$$

Beweis : Wir können oBdA $r = 1$ annehmen, denn hat die Potenzreihe $\sum a_n x^n$ den Konvergenzradius $r > 0$, so hat die Reihe $\sum b_k x^k$ mit $b_k := a_k r^k$ offensichtlich den Konvergenzradius 1 und $\sum b_k x^k$ ist dann und nur dann für $x = 1$ konvergent, wenn $\sum a_k x^k$ für $x = r$ konvergent ist.

Wir setzen $s_{-1} := 0$ und $s_n := a_0 + a_1 + \ldots + a_n$ und erhalten für $|x| < 1$

$$\sum_{k=0}^{n} a_k x^k = \sum_{k=0}^{n} (s_k - s_{k-1}) x^n = (1-x) \sum_{k=0}^{n-1} s_k x^k + s_n x^n$$

und damit ($n \to \infty$) für $|x| < 1$

$$f(x) = (1-x) \sum_{k=0}^{\infty} s_k x^k.$$

Ist

$$s = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^{\infty} a_k = \sum_{k=0}^{\infty} a_k,$$

dann folgt wegen $(1-x) \sum\limits_{k=0}^{\infty} x^k = 1$ $(|x| < 1)$ (Summenformel für die geometrische Reihe)

$$(*) \qquad f(x) - s = (1-x) \sum_{k=0}^{\infty} (s_k - s) x^k.$$

Wegen $s = \lim_{n\to\infty} s_n$ gibt es zu beliebig vorgegebenem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $|s_n - s| < \frac{\varepsilon}{2}$.
Für $0 \leq x < 1$ folgt dann aus $(*)$

$$\begin{aligned} |f(x) - s| &\leq (1-x)\sum_{k=0}^{\infty} |s_n - s| x^k \\ &\leq (1-x)\sum_{k=0}^{N-1} |s_k - s| + (1-x)\frac{\varepsilon}{2}\sum_{k=N}^{\infty} x^k \\ &\leq (1-x)\sum_{k=0}^{N-1} |s_k - s| + \frac{\varepsilon}{2}. \end{aligned}$$

Wählen wir nun ein $\delta \in]0,1[$, so dass für alle $x \in]1-\delta, 1[$ gilt

$$(1-x)\sum_{k=0}^{N-1} |s_k - s| < \frac{\varepsilon}{2},$$

so folgt für diese x

$$|f(x) - s| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

d.h. es ist

$$\lim_{\substack{x\to 1\\ x<1}} f(x) = s = \sum_{k=0}^{\infty} a_k.$$

□

Um die Tragweite des Abelschen Grenzwertsatzes zu demonstrieren, beweisen wir an dieser Stelle einen Satz über das Cauchy-Produkt *nicht notwendig absolut konvergenter* Reihen.

20.5.18 Satz (N.H. Abel, 1826)

Sind die drei (reellen) Reihen $\sum a_k,\ \sum b_l$ und $\sum c_n$ mit $c_n := \sum\limits_{k=0}^{n} a_k b_{n-k}$ konvergent und sind $A,\ B$ bzw. C ihre Summen, dann gilt $C = A \cdot B$ oder explizit

$$\boxed{\sum_{n=0}^{\infty} c_n = \sum_{k=0}^{\infty} a_k \cdot \sum_{l=0}^{\infty} b_l}$$

Man beachte, dass hier keine absolute Konvergenz vorausgesetzt wird.
Für $x \in [0,1[$ sei

$$A(x) = \sum_{k=0}^{\infty} a_k x^k,\ B(x) = \sum_{l=0}^{\infty} b_l x^l \ \text{ und } \ C(x) = \sum_{n=0}^{\infty} c_n x^n$$

und es gilt $A(x)B(x) = C(x)$ für $0 \leq x \leq 1$. Wegen der vorausgesetzten Konvergenz von $\sum a_k,\ \sum b_l$ und $\sum c_n$ besitzen die Funktionen nach dem Abelschen Grenzwertsatz einen Grenzwert für $x \to 1$ $(x < 1)$.

Aus $A(x)B(x) = C(x)$ folgt daher wegen $\lim\limits_{\substack{x\to 1\\ x<1}} A(x) = A,\ \lim\limits_{\substack{x\to 1\\ x<1}} B(x) = B$ und $\lim\limits_{\substack{x\to 1\\ x<1}} C(x) = C$, auch

$$A \cdot B = C.$$

(d) Als weiteres Beispiel behandeln wir die Taylor-Reihe von arctan und gehen ähnlich wie bei Beispiel (c) vor.

Wir nehmen das Ergebnis vorweg:

20.5.19 Satz

Für alle $x \in \mathbb{R}$ mit $|x| \leq 1$ gilt

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \pm \ldots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}$$

Speziell gilt

$$\arctan 1 = \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \pm \ldots$$

Denn für $|x| < 1$ gilt

$$\arctan x = \int_0^x \frac{1}{1+t^2} dt = \int_0^x \sum_{k=0}^{\infty} (-1)^k t^{2k} dt = \sum_{k=0}^{\infty} (-1)^k \int_0^x t^{2k} dt = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}.$$

(da $\sum (-1)^k t^{2k}$ auf $[0, |x|]$ gleichmäßig konvergiert)

Für $x = 1$ folgt die Gültigkeit der Formel aus dem Abelschen Grenzwertsatz, da nach dem Leibniz-Kriterium die Reihe $\sum\limits_{k=0}^{\infty} \frac{(-1)^k}{k+1}$ konvergiert.

Da der Abelsche Grenzwertsatz aber auch für den Fall gilt, dass die Potenzreihe $\sum a_k x^k$ noch für $x = -r$ konvergiert, gilt die Darstellung für alle $x \in \mathbb{R}$ mit $|x| \leq 1$.

Bemerkung: Zur praktischen Berechnung von π ist die Formel

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{1} \pm \ldots$$

nicht sonderlich gut geeignet, da die Reihe zu langsam konvergiert.

Der BBP-Algorithmus (vgl. ???) ist da wesentlich effizienter.
Im Vor-Computer-Zeitalter hat man mit der Darstellung (sog. Machinsche Formel)

$$\frac{\pi}{4} = 4 \arctan \frac{1}{5} - \arctan \frac{1}{239}$$

gearbeitet, die schneller konvergiert.

(e) **Binomische Reihe**
Eine wichtige Reihe, die als Spezialfall sowohl die binomische Formel als auch die geometrische Reihe umfasst, ist die *binomische Reihe*, dahinter versteht man die Reihe

$$b_\alpha(x) := \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \qquad \alpha \in \mathbb{R}.$$

$\binom{\alpha}{k}$ ist dabei der schon in ??? definierte Binomialkoeffizient

$$\binom{\alpha}{0} := 1; \quad \binom{\alpha}{k} = \frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{1 \cdot 2 \cdot 3 \cdot \ldots \cdot k} \quad (k \geq 1).$$

Wir werden sehen, dass durch diese Reihe die Potenzfunktion $x \mapsto (1+x)^\alpha = \exp(\alpha \log(1+x))$ wenigstens in $]-1,1[$ dargestellt wird. Berechnet man nämlich die Taylor-Koeffizienten von $f(x) = (1+x)^\alpha$ zum Entwicklungspunkt $a = 0$, so erhält man

$$f^{(k)}(x) = \alpha(\alpha-1)\dots(\alpha-k+1)(1+x)^{\alpha-k} = k! \binom{\alpha}{k} (1+x)^{\alpha-k}$$

und damit $\frac{f^{(k)}(0)}{k!} = \binom{\alpha}{k}$ und daher

$$T_{\infty,0}(f)(x) = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k = b_\alpha(x).$$

Die Frage ist zunächst, für welche $x \in \mathbb{R}$ die Binomialreihe $b_\alpha(x)$ konvergiert.
Um triviale Fälle auszuschließen, nehmen wir $\alpha \notin \mathbb{N}_0$ und $x \neq 0$ an.
Wir wenden das Quotientenkriterium mit $a_n := \binom{\alpha}{n} x^n$ an:

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \left| \frac{\binom{\alpha}{n+1} x^{n+1}}{\binom{\alpha}{n} x^n} \right| = |x| \underbrace{\lim_{n\to\infty} \left| \frac{\alpha-n}{n+1} \right|}_{=1} = |x|.$$

Ist also $|x| < 1$, so ist die Binomialreihe nach der Limesform des Quotientenkriteriums konvergent.

Um zu zeigen, dass für $|x| < 1$ gilt

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k := b_\alpha(x)$$

oder m.a.W. (mit $f(x) = (1+x)^\alpha$)

$$f(x) = T_{\infty,0}(f)(x) = b_\alpha(x),$$

könnte man zeigen, dass das entsprechende Restglied $R_{n+1}(x)$ für $|x| < 1$ gegen Null konvergiert. Das ist im Prinzip möglich, wir gehen aber anders vor und benutzen einen Differetialgleichungstrick.
Aus

$$b_\alpha(x) = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \quad (|x| < 1)$$

folgt

$$b'_\alpha(x) = \sum_{k=1}^{\infty} k \binom{\alpha}{k} x^{k-1} = \sum_{k=0}^{\infty} (k+1) \binom{\alpha}{k+1} x^{k+1}.$$

Da aber

$$(k+1)\binom{\alpha}{k+1} = (k+1)\frac{\alpha(\alpha-1)\dots(\alpha-k+1)(\alpha-k)}{1\cdot 2\cdot \ldots \cdot k \cdot (k+1)} = \alpha \binom{\alpha-1}{k}$$

gilt, ergibt sich weiter

$$b'_\alpha(x) = \alpha b_{\alpha-1}(x).$$

Nun ist aber

$$\begin{aligned}(1+x)b_{\alpha-1}(x) &= (1+x)\sum_{k=0}^{\infty}\binom{\alpha-1}{k}x^k\\ &= 1+\sum_{k=1}^{\infty}\underbrace{\left\{\binom{\alpha-1}{k}+\binom{\alpha-1}{k-1}\right\}}_{=\binom{\alpha}{k}}x^k\\ &= \sum_{k=0}^{\infty}\binom{\alpha}{k}x^k.\end{aligned}$$

Also gilt für alle $x \in \mathbb{R}$ mit $|x| < 1$ die (Differential-)Gleichung

$$(*) \qquad (1+x)b'_{\alpha}(x) = \alpha b_{\alpha}(x).$$

Da für $|x| < 1$ aber $(1+x)^{\alpha} \neq 0$ gilt, so folgt aus $(*)$, dass $x \mapsto b_{\alpha}(x)(1+x)^{-\alpha}$ in $]-1,1[$ die Ableitung Null hat, dass also $x \mapsto b_{\alpha}(x)(1+x)^{-\alpha}$ dort konstant ist.
Für $x = 0$ erhält man, dass die Konstante gleich 1 sein muss. Daher gilt

$$(1+x)^{\alpha} = b_{\alpha}(x)$$

für alle $x \in]-1,1[$ und beliebige $\alpha \in \mathbb{R}$.
Ohne Beweis sei mitgeteilt, dass die Formel auch noch im Fall $x = -1$ gilt, falls $\alpha > 0$ ist und im Fall $x = 1$, falls $\alpha > -1$. Zu allen anderen Fällen divergiert die Binomialreihe $b_{\alpha}(x)$ (man beachte $\alpha \notin \mathbb{N}_0$ für $\alpha \in \mathbb{N}_0$ „erstarrt“ die Binomialreihe zur Binomischen Formel).

Fassen wir zusammen:

20.5.20 Satz

Für $\alpha \in \mathbb{R} - \mathbb{N}_0$ gilt

$$(1+x)^{\alpha} = \sum_{k=0}^{\infty}\binom{\alpha}{k}x^k, \text{ falls } \begin{cases} |x| < 1, & \\ x = -1, & \text{und } \alpha > 0 \\ x = 1, & \text{und } \alpha > -1. \end{cases}$$

Für alle anderen $x \in \mathbb{R}$ divergiert die Binomialreihe.

20.5.21 Beispiele und Bemerkungen zur Binomialreihe

(a) Für $\alpha = -1$ mit $\binom{-1}{k} = (-1)^k$, daher ergibt sich für $\alpha = -1$ aus der binomischen Reihe für $|x| < 1$

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 \pm \ldots = \sum_{k=0}^{\infty}(-1)^k x^k.$$

Ersetzt man x durch $-x$, so ergibt sich

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots = \sum_{k=0}^{\infty} x^k.$$

Die geometrische Reihe ist also ein Spezialfall der Binomischen Reihe.
Die häufig verwendeten Näherungsformel

$$\frac{1}{1-x} \approx 1+x \text{ bzw. } \frac{1}{1+x} \approx 1-x$$

(für kleine $|x|$) sind nun evident.
Wichtige Spezialfälle sind $\alpha = \frac{1}{2}$ und $\alpha = -\frac{1}{2}$. Berechnet man im Fall $\alpha = \frac{1}{2}$ die ersten Binomialkoeffizienten, so erhält man

$$\binom{\frac{1}{2}}{0} = 1, \quad \binom{\frac{1}{2}}{1} = \frac{1}{2}, \quad \binom{\frac{1}{2}}{2} = \frac{\frac{1}{2}(-\frac{1}{2})}{1\cdot 2} = -\frac{1}{8},$$
$$\binom{\frac{1}{2}}{3} = \frac{(\frac{1}{2})(-\frac{1}{2})(-\frac{3}{2})}{1\cdot 2\cdot 3} = \frac{1}{16}, \quad \binom{\frac{1}{2}}{4} = \binom{\frac{1}{2}}{3}\frac{-\frac{5}{2}}{4} = -\frac{5}{128}$$

und damit für $|x| < 1$

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 - \frac{5}{128}x^4 + \mathcal{O}(x^6),$$

dabei steht $\mathcal{O}(x^6)$ für eine Potenzreihe, die mit dem Glied cx^6 $(c \in \mathbb{R}_+^*)$ beginnt.

Wie schon früher im Fall der linearen Approximation $\sqrt{1+x} \approx 1 + \frac{x}{2}$ ausgeführt, kann man diese Formel benutzen um Wurzeln näherungsweise zu berechnen und den dabei gemachten Fehler abzuschätzen. Um etwa $\sqrt{10}$ näherungsweise zu berechnen, könnte man in erster Näherung $\sqrt{10} = \sqrt{1+9} \approx 1+\frac{9}{2} = 5,5$ rechnen, der Näherungswert ist aber viel zu schlecht. Um die Formel für $|x| < 1$ zu benutzen, schreiben wir

$$\sqrt{10} = \sqrt{9 \cdot \frac{10}{9}} = 3 \cdot \sqrt{1 + \frac{1}{9}} = 3\left(1 + \frac{1}{2\cdot 9} - \frac{1}{8 \cdot 9^2} + \ldots\right) = 3 + \frac{1}{6} - \frac{1}{8\cdot 27} + \ldots \approx 3,162.$$

Der „exakte“ Wert ist

$$\sqrt{10} = 3,1622776602\ldots$$

Für $\alpha = -\frac{1}{2}$ erhält man

$$\binom{-\frac{1}{2}}{0} = 1 \text{ und } \binom{-\frac{1}{2}}{k} = (-1)^k \frac{1\cdot 3\cdot 5\cdot \ldots (2k-1)}{2\cdot 4 \cdot \ldots \cdot 2k}$$

für $k \geq 1$, also z.B.

$$\binom{-\frac{1}{2}}{1} = -\frac{1}{2};, \quad \binom{-\frac{1}{2}}{2} = \frac{3}{8}, \quad \binom{-\frac{1}{2}}{3} = -\frac{5}{16}$$

und daher (für $|x| < 1$)

$$\boxed{\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + \frac{35}{128}x^4 \mp \ldots}$$

Anwendung: Nach A.Einstein beträgt die Energie einer „relativistischen“ Teilchens der Masse $m = \frac{m_0}{\sqrt{1-(\frac{v}{c})^2}}$ (m_0 Ruhemasse, v die Geschwindigkeit des Teilchens, c Lichtgeschwindigkeit)

$$E = mc^2.$$

Die kinetische Energie ist definiert durch

$$E_{kin} = mc^2 - m_0c^2.$$

Ist $v < c$, so kann man zur Berechnung von E_{kin} die binomische Reihe (mit $x = -(\frac{v}{c})^2$) verwenden:

$$\begin{aligned} E_{kin} &= mc^2 - m_0c^2 \\ &= m_0c^2\left(\frac{1}{\sqrt{1-(\frac{v}{c})^2}} - 1\right) \\ &= m_0c^2\left(\frac{1}{2}(\frac{v}{c})^2 + \frac{3}{8}(\frac{v}{c})^4 + \frac{35}{128}(\frac{v}{c})^8 + \ldots\right) \\ &= \frac{1}{2}m_0v^2 + \frac{3}{8}m_0v^2(\frac{v}{c})^2 + \text{ Glieder höherer Ordnung} \end{aligned}$$

Der Term $\frac{1}{2}m_0v^2$ repräsentiert die kinetische Energie im klassischen Fall (v sehr klein gegenüber c), der sog. „3/8-Term" $\frac{3}{8}m_0v^2(\frac{v}{c})^2$ ist das Glied niedrigster Ordnung der Abweichung zwischen dem relativistischen und nicht-relativistischen Fall.

Die Beispiele haben gezeigt, dass zum Aufstellen der Taylor-Entwicklung einer Funktion die Benutzung der Taylorschen Formel mit Restglied und dem Nachweis $\lim\limits_{n\to\infty} R_{n+1}(x) = 0$ häufig schwerfällig ist. Es ist häufig günstiger, bekannte Reihen zu differenzieren oder zu integrieren (vgl. die Beispiele ???).
Ferner kann man versuchen, die gegebene Funktion als Summe und/oder Produkt von Funktionen mit bekannten Reihenentwicklung darzustellen, z.B.

$$\begin{aligned} \cosh x &= \frac{1}{2}(\exp(x) + \exp(-x)) \\ &= \frac{1}{2}\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \ldots\right) \\ &= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6} + \ldots = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \quad (x \in \mathbb{R}) \end{aligned}$$

$$\begin{aligned} \frac{\cos x}{1-x} &= \cos x \frac{1}{1-x} \\ &= \left(\sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!}\right) \cdot \sum_{l=0}^{\infty} x^l \\ &= 1 + x + \left(1 - \frac{1}{2!}\right)x^2 + \left(1 - \frac{1}{2!}\right)x^3 + \left(1 - \frac{1}{2!} + \frac{1}{4!}\right)x^4 + \ldots, \quad |x| < 1 \end{aligned}$$

In der folgenden Tabelle sind einige häufig benutzte Taylor-Entwicklungen zusammengestellt.

Funktion	Taylor-Reihe um $a=0$	Gültigkeitsbereich
$\frac{1}{1-x}$	$\sum\limits_{k=0}^{\infty} x^k$ (geom. Reihe)	$\lvert x\rvert < 1$
$\exp x$	$\sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$	$x \in \mathbb{R}$
$\sin x$	$\sum\limits_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots$	$x \in \mathbb{R}$
$\cos x$	$\sum\limits_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} - \frac{x^4}{4!} + \ldots$	$x \in \mathbb{R}$
$\sinh x$	$\sum\limits_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots$	$x \in \mathbb{R}$
$\cosh x$	$\sum\limits_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots$	$x \in \mathbb{R}$
$(1+x)^\alpha$	$\sum\limits_{k=0}^{\infty} \binom{\alpha}{k} x^k$	$\lvert x\rvert < 1$ $(\alpha \in \mathbb{R} - \mathbb{N}_0)$
$(1+x)^{\frac{1}{2}}$	$1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{10}x^3 - \frac{5}{128}x^4 \pm \ldots$	$\lvert x\rvert < 1$
$(1+x)^{-\frac{1}{2}}$	$1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + \frac{35}{128}x^4 \mp$	$\lvert x\rvert < 1$
$\log(1+x)$	$\sum\limits_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \pm \ldots$	$-1 \le x \le 1$
$\arctan x$	$\sum\limits_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \pm \ldots$	$-1 \le x \le 1$
$\arcsin x$	$\sum\limits_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} \frac{x^{2k+1}}{2k+1} = x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \ldots$	$-1 \le x \le 1$
	beachte: $(-1)^k \binom{-\frac{1}{2}}{k} = \frac{1\cdot 3\cdot 5\cdot \ldots \cdot (2k-1)}{2\cdot 4\cdot \ldots \cdot 2k}$ für $k \ge 1$	

Die Reihenentwicklung von arcsin erhält man aus der Reihenentwicklung der Ableitung $\arcsin' x = \frac{1}{\sqrt{1-x^2}}$ $(\lvert x\rvert < 1)$ und durch Anwendung des Abelschen Grenzwertsatzes durch gliedweise Integration.