3 Differential- und Integralrechnung

3.1 Differenzierbare Funktionen

Gegeben sei eine beliebige Funktion $f : I = [a, b] \to \mathbb{R}$ und ein fester Punkt $x_0 \in I$. Außerdem sei $h \in \mathbb{R}$ so klein, dass auch noch $x_0 + h$ in I liegt.

Wir beschreiben die Gerade durch $P = (x_0, f(x_0))$ und $Q = (x_0 + h, f(x_0 + h))$ durch eine affin-lineare Funktion $\lambda_h(x) = m_h x + b$. Dann ist

$$\lambda_h(x_0) = f(x_0) \quad \text{und} \quad \lambda_h(x_0 + h) = f(x_0 + h). \tag{*}$$

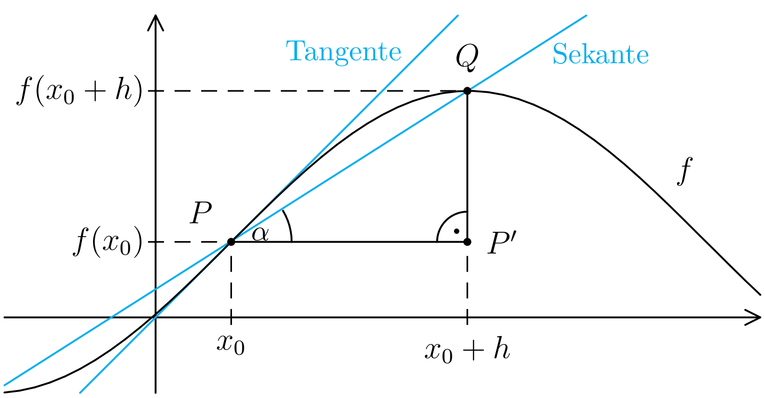
Eine solche Gerade (also den Graphen von λ_h) nennt man eine ***Sekante***, weil sie den Graphen von f in den beiden Punkten P und Q schneidet. Aus (*) erhält man die Steigung m_h und damit die endgültige Form der Sekante,

$$m_h = \frac{f(x_0 + h) - f(x_0)}{h} \quad \text{und} \quad \lambda_h(x) = f(x_0) + m_h \cdot (x - x_0).$$

Bezeichnet α den ***Steigungswinkel*** in dem „Steigungsdreieck“, das aus den Punkten P, $P' = (x_0 + h, f(x_0))$ und Q gebildet wird, so ist $\tan(\alpha) = m_h$. Wenn m_h für $h \to 0$ gegen einen wohlbestimmten Wert m konvergiert, so nennt man die affin-lineare Funktion

$$T(x) := f(x_0) + m \cdot (x - x_0)$$

(oder genauer ihren Graphen) die ***Tangente*** an (den Graphen von) f in x_0.



Definition

Eine Funktion $f : I \to \mathbb{R}$ heißt ***differenzierbar*** in $x_0 \in I$, falls der Grenzwert

$$\boxed{f'(x_0) := \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \qquad (\text{für } h \neq 0 \text{ und } x_0 + h \in I)}$$

existiert. Die Zahl $\dfrac{df}{dx}(x_0) := f'(x_0)$ nennt man die ***Ableitung*** von f in x_0.

1.1. Beispiel

Die Funktion $f(x) = x^2$ ist überall differenzierbar, denn es ist

$$\frac{(x+h)^2 - x^2}{h} = \frac{2xh + h^2}{h} = 2x + h,$$

und dieser Ausdruck strebt für $h \to 0$ gegen $2x$. Also ist $f'(x) = 2x$.

Etwas allgemeiner erhält man im Falle der Funktion $f(x) = x^n$ die Beziehung

$$\begin{aligned}\frac{f(x+h) - f(x)}{h} &= \frac{(x+h)^n - x^n}{h} = \frac{1}{h}\left(\sum_{i=0}^{n}\binom{n}{i}x^i h^{n-i} - x^n\right)\\ &= \sum_{i=0}^{n-1}\binom{n}{i}x^i h^{n-i-1} = n \cdot x^{n-1} + \text{Terme, die } h \text{ enthalten.}\end{aligned}$$

Lässt man h gegen Null gehen, so fallen die Terme weg, die h als Faktor enthalten, und der ganze Ausdruck strebt gegen $n \cdot x^{n-1}$. Also ist auch die Funktion $f(x) = x^n$ differenzierbar und $f'(x) = n \cdot x^{n-1}$.

Definition

Die Funktion heißt ***auf*** I ***differenzierbar***, falls sie in jedem Punkt $t \in I$ differenzierbar ist. Ist die Funktion f auf ganz I differenzierbar, so nennt man die auf I definierte Funktion $f' : t \mapsto f'(t)$ die ***abgeleitete Funktion*** (oder kurz: die ***Ableitung***) von f auf I.

1.2. Äquivalente Formulierungen der Differenzierbarkeit

Sei $I \subset \mathbb{R}$ ein beliebiges Intervall und $t_0 \in I$. Folgende Aussagen über eine Funktion $f : I \to \mathbb{R}$ sind äquivalent:

1. *f ist differenzierbar in t_0.*
2. *Es gibt ein $a \in \mathbb{R}$ und eine Funktion $\delta : I \to \mathbb{R}$ mit*

$$f(t) = f(t_0) + a \cdot (t - t_0) + \delta(t) \cdot (t - t_0) \text{ auf } I \text{ und } \lim_{t \to t_0} \delta(t) = 0.$$

3. *Es gilt das „Grauert-Kriterium“:*

   Es gibt eine in t_0 stetige Abbildung $\Delta : I \to \mathbb{R}$ mit

$$f(t) = f(t_0) + (t - t_0) \cdot \Delta(t).$$

Ist f in t_0 differenzierbar, so ist $f'(t_0) = a = \Delta(t_0)$.

Beweis: (1) $\Longrightarrow$ (3): Ist f in t_0 differenzierbar, so setzen wir $\Delta(t) = \dfrac{f(t) - f(t_0)}{t - t_0}$ für $t \neq t_0$. Im Grenzwert $t \to t_0$ strebt $\Delta(t)$ gegen $f'(t_0)$. Man kann Δ also durch $f'(t_0)$ in t_0 stetig ergänzen. Die gewünschte Gleichung gilt auf ganz I.

(3) $\Longrightarrow$ (2): Gilt das Grauert-Kriterium, so setzen wir $a := \Delta(x_0)$ und $\delta(t) := \Delta(t) - \Delta(t_0)$. Dann ist offensichtlich Eigenschaft (2) erfüllt.

(2) $\Longrightarrow$ (1): Gilt Eigenschaft (2), so strebt

$$\frac{f(t) - f(t_0)}{t - t_0} = a + \delta(t)$$

für $t \to t_0$ gegen a, und f ist in t_0 differenzierbar, mit $f'(t_0) = a$. ∎

Definition

$f = g + \mathrm{i}\,h : I \to \mathbb{C}$ heißt ***differenzierbar in*** $t_0 \in I$, falls g und h in t_0 differenzierbar sind. Dann nennt man $f'(t_0) := g'(t_0) + \mathrm{i}\,h'(t_0)$ die ***Ableitung*** von f in t_0.

Auch im Falle einer komplexwertigen differenzierbaren Funktion ist

$$f'(t_0) := \lim_{t \to t_0} \frac{f(t) - f(t_0)}{t - t_0}\,,$$

und die Kriterien des vorigen Satzes gelten auch in dieser Situation.

1.3. Beispiele

A. Ist $f(t) \equiv c$ eine konstante (reell- oder komplexwertige) Funktion, so ist $f(t) - f(t_0) \equiv 0$ für alle t, t_0, also $f'(t) \equiv 0$.

B. Ist $f(t) := a + bt$ eine affin-lineare Funktion, so ist

$$\frac{f(t) - f(t_0)}{t - t_0} = b \text{ für beliebige Parameter } t,\ t_0,$$

also $f'(t) \equiv b$.

C. Die Ableitung der Potenzfunktion x^n haben wir schon weiter oben berechnet. Betrachten wir jetzt die Exponentialfunktion! Es ist

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

Diese Umformung ist erlaubt, weil $\Delta(x) := \sum_{n=0}^{\infty} x^n/(n+1)!$ genau wie $\exp(x)$ den Konvergenzradius ∞ hat. Außerdem folgt sofort, dass Δ (auf ganz $\mathbb{R}$ und damit insbesondere in $x = 0$) stetig ist. Die Zerlegung

$$\exp(x) = \exp(0) + x \cdot \Delta(x)$$

zeigt, dass exp in $x = 0$ differenzierbar ist, mit $\exp'(0) = \Delta(0) = 1$.

Ist nun $x_0 \in \mathbb{R}$ ein beliebiger Punkt, so ist

$$\begin{aligned}
\exp(x) - \exp(x_0) &= \exp(x_0 + (x - x_0)) - \exp(x_0) \\
&= \exp(x_0) \cdot \exp(x - x_0) - \exp(x_0) \\
&= \exp(x_0) \cdot (\exp(x - x_0) - 1).
\end{aligned}$$

Die Zerlegung $\exp(x) = 1 + x \cdot \Delta(x)$ liefert:

$$\exp(x - x_0) - 1 = (x - x_0) \cdot \Delta(x - x_0).$$

Da die Funktion $\Delta_0(x) := \exp(x_0)\cdot\Delta(x-x_0)$ überall (und damit insbesondere in $x = x_0$) stetig und $\exp(x) = \exp(x_0) + (x - x_0) \cdot \Delta_0(x)$ ist, folgt:

exp ist in x_0 differenzierbar und $\exp'(x_0) = \Delta_0(x_0) = \exp(x_0)$. Also ist allgemein

$$\boxed{(e^x)' = e^x.}$$

D. Jetzt betrachten wir die komplexwertige Funktion $f(t) := \exp(\mathrm{i}\,t)$. Genau wie oben erhalten wir eine Zerlegung

$$f(t) = \sum_{n=0}^{\infty} \frac{(\mathrm{i}\,t)^n}{n!} = 1 + (\mathrm{i}\,t) \cdot \sum_{n=0}^{\infty} \frac{(\mathrm{i}\,t)^n}{(n+1)!} = f(0) + t \cdot \widetilde{\Delta}(t)\,,$$

mit der stetigen komplexwertigen Funktion $\widetilde{\Delta}(t) := \mathrm{i} \cdot \sum_{n=0}^{\infty}(\mathrm{i}\,t)^n/(n+1)!$. Also ist f in $t = 0$ differenzierbar, und es ist $f'(0) = \widetilde{\Delta}(0) = \mathrm{i}$.

Ist $t_0 \in \mathbb{R}$ beliebig, so folgt aus dem Additionstheorem für die Exponentialfunktion die Formel

$$\exp(\mathrm{i}\,t) - \exp(\mathrm{i}\,t_0) = \exp(\mathrm{i}\,t_0) \cdot (\exp(\mathrm{i}\,(t - t_0)) - 1),$$

also $f(t) = f(t_0)+(t-t_0)\cdot f(t_0)\cdot\widetilde{\Delta}(t-t_0)$. Damit ist f auch in t_0 differenzierbar, und

$$f'(t_0) = f(t_0) \cdot \widetilde{\Delta}(0) = \mathrm{i} \cdot f(t_0).$$

Aus der Euler'schen Formel $e^{\mathrm{i}\,t} = \cos t + \mathrm{i}\,\sin t$ erhalten wir nun, dass $\cos(t)$ und $\sin(t)$ auf ganz $\mathbb{R}$ differenzierbar sind, und es gilt:

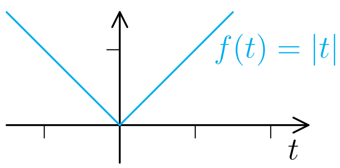
$$\cos'(t) + \mathrm{i}\ \sin'(t) = \mathrm{i} \cdot (\cos(t) + \mathrm{i}\ \sin(t)),$$

also $\boxed{\cos'(t) = -\sin(t)}$ und $\boxed{\sin'(t) = \cos(t).}$

E. Bis jetzt haben wir nur Beispiele von differenzierbaren Funktionen kennengelernt. Die Funktion $f(t) := |t|$ ist allerdings in $t = 0$ nicht differenzierbar, denn

$$\frac{f(t) - f(0)}{t} = \frac{|t|}{t} = \begin{cases} 1 & \text{für } t > 0 \\ -1 & \text{für } t < 0 \end{cases}$$

besitzt für $t \to 0$ unterschiedliche Grenzwerte von links und von rechts.



Tatsächlich ist f in allen Punkten außerhalb 0 differenzierbar, aber bei $x = 0$ hat die Funktion einen „Knick“. Sie ist dort nicht genügend glatt und besitzt deshalb keine eindeutig bestimmte Tangente.

Ist eine Funktion $f : I \to \mathbb{R}$ in t_0 differenzierbar, so ist f in t_0 erst recht stetig. Das folgt sofort aus der Darstellung $f(t) = f(t_0) + (t - t_0) \cdot \Delta(t)$, mit einer in t_0 stetigen Funktion Δ.

Wie man an Hand des obigen Beispiels sehen kann, braucht umgekehrt eine stetige Funktion durchaus nicht differenzierbar zu sein!

1.4. Die Berechnung von $(\mathbf{u} \circ f)'$ für lineares u

Sei $f : I \to \mathbb{C}$ differenzierbar in t_0 und $\mathbf{u} : \mathbb{C} \to \mathbb{R}$ linear. Dann ist auch die Abbildung $\mathbf{u} \circ f : I \to \mathbb{R}$ in t_0 differenzierbar, und es ist

$$(\mathbf{u} \circ f)'(t_0) = \mathbf{u}(f'(t_0)).$$

BEWEIS: Wegen der Linearität von $\mathbf{u}$ ist

$$\frac{\mathbf{u} \circ f(t) - \mathbf{u} \circ f(t_0)}{t - t_0} = \mathbf{u}\left(\frac{f(t) - f(t_0)}{t - t_0}\right),$$

und da $\mathbf{u}$ stetig ist, konvergiert der rechte Ausdruck für $t \to t_0$ gegen $\mathbf{u}(f'(t_0))$. ∎

1.5. Beispiele

A. Sind $f, g : I \to \mathbb{R}$ differenzierbar, so ist auch die Funktion $F := f + \mathrm{i}\,g : I \to \mathbb{C}$ differenzierbar. Die Abbildung $\mathbf{u} : \mathbb{C} \to \mathbb{R}$ mit $\mathbf{u}(x + \mathrm{i}\,y) := x + y$ ist linear. Also ist auch $f + g = \mathbf{u} \circ (f + \mathrm{i}\,g)$ differenzierbar, und es ist $(f + g)' = \mathbf{u}(f' + \mathrm{i}\,g') = f' + g'$.

B. Genauso folgt: Mit $f : I \to \mathbb{C}$ und $c \in \mathbb{C}$ ist auch $c \cdot f$ differenzierbar, und es ist $(c \cdot f)' = c \cdot f'$, für beliebiges $c \in \mathbb{R}$ (man benutze die lineare Abbildung $\mathbf{u} : z \mapsto c \cdot z$).

1.6. Produktregel

Sind die Funktionen $f, g : I \to \mathbb{R}$ beide in t_0 differenzierbar, so ist auch ihr Produkt $f \cdot g : I \to \mathbb{R}$ in t_0 differenzierbar, und es gilt:

$$(f \cdot g)'(t_0) = f'(t_0) \cdot g(t_0) + f(t_0) \cdot g'(t_0).$$

BEWEIS: Nach Voraussetzung sind die einzelnen Funktionen in t_0 differenzierbar, es gibt also in t_0 stetige Funktionen Δ_f und Δ_g, so dass gilt:

$$f(t) = f(t_0) + (t - t_0) \cdot \Delta_f(t) \text{ und } g(t) = g(t_0) + (t - t_0) \cdot \Delta_g(t),$$

sowie $\Delta_f(t_0) = f'(t_0)$ und $\Delta_g(t_0) = g'(t_0)$. Jetzt benutzen wir einen kleinen Trick:

$$\begin{aligned} (f \cdot g)(t) - (f \cdot g)(t_0) &= (f(t) - f(t_0)) \cdot g(t) + f(t_0) \cdot (g(t) - g(t_0)) \\ &= (t - t_0) \cdot [\Delta_f(t) \cdot g(t) + f(t_0) \cdot \Delta_g(t)]\,. \end{aligned}$$

Da die Funktion in der eckigen Klammer stetig in t_0 ist, ist $f \cdot g$ dort differenzierbar, und es ist $(f \cdot g)'(t_0) = f'(t_0) \cdot g(t_0) + f(t_0) \cdot g'(t_0)$. ∎

Bemerkung: Man rechnet leicht nach, dass die Produktregel auch für komplexwertige Funktionen gilt.

Aus der Produktregel folgt scheinbar auch gleich die bekannte „**Quotientenregel**“:

$$\Big(\frac{f}{g}\Big)'(t) = \frac{f'(t)g(t) - f(t)g'(t)}{g(t)^2}\,, \text{ falls } g(t) \neq 0 \text{ ist.}$$

Es ist nämlich $f' = (g \cdot (f/g))' = g' \cdot (f/g) + g \cdot (f/g)'$, also

$$(f/g)' = [f' - g' \cdot (f/g)]/g = (f'g - g'f)/g^2.$$

Leider taugt dieser Beweis nur als Merkregel, wir müssen zuvor noch die Differenzierbarkeit von f/g beweisen. Es reicht, dies im Falle $f = 1$ zu tun.

Ist g in t_0 differenzierbar (und damit dort auch stetig) und $g(t_0) \neq 0$, so gilt $g(t) \neq 0$ auf einer ganzen Umgebung von t_0, und außerdem haben wir eine Darstellung $g(t) = g(t_0) + (t - t_0) \cdot \Delta_g(t)$ mit einer in t_0 stetigen Funktion Δ_g. Dann gilt:

$$\frac{1}{g(t)} - \frac{1}{g(t_0)} = \frac{g(t_0) - g(t)}{g(t)g(t_0)} = -\frac{(t - t_0)\Delta_g(t)}{g(t)g(t_0)}\,.$$

Da $\Delta^*(t) := -\Delta_g(t)/(g(t)g(t_0))$ in t_0 stetig ist, folgt die Differenzierbarkeit von $1/g$ in t_0 mit $(1/g)'(t_0) = -g'(t_0)/g(t_0)^2$. Mit Hilfe der Produktregel ergibt sich daraus die Differenzierbarkeit allgemeiner Quotienten. ∎

1.7. Beispiel

Als Anwendung berechnen wir die Ableitung des Tangens:

$$\tan'(t) = \left(\frac{\sin}{\cos}\right)'(t) = \frac{\cos^2(t) - (-\sin^2(t))}{\cos^2(t)} = 1+\tan^2(t), \text{ für } t \in (-\frac{\pi}{2}, +\frac{\pi}{2}).$$

1.8. Die Kettenregel

Sei $f : I \to \mathbb{R}$ differenzierbar in t_0, J ein weiteres Intervall mit $f(I) \subset J$ und $g : J \to \mathbb{R}$ differenzierbar in $s_0 := f(t_0) \in J$. Dann ist auch $g \circ f : I \to \mathbb{R}$ differenzierbar in t_0, und es gilt:

$$(g \circ f)'(t_0) = g'(f(t_0)) \cdot f'(t_0).$$

BEWEIS: Wir schreiben

$$\begin{aligned} f(t) &= f(t_0) + (t - t_0) \cdot \Delta_f(t) \\ \text{und} \quad g(s) &= g(s_0) + (s - s_0) \cdot \Delta_g(s), \end{aligned}$$

wobei Δ_f in t_0 und Δ_g in s_0 stetig ist, sowie $\Delta_f(t_0) = f'(t_0)$ und $\Delta_g(s_0) = g'(s_0)$. Einfaches Einsetzen ergibt:

$$\begin{aligned} g \circ f(t) &= g(s_0) + (f(t) - s_0) \cdot \Delta_g(f(t)) \\ &= g(f(t_0)) + (f(t) - f(t_0)) \cdot \Delta_g(f(t)) \\ &= g \circ f(t_0) + (t - t_0) \cdot \Delta_f(t) \cdot \Delta_g(f(t)). \end{aligned}$$

Dabei ist $t \mapsto \Delta_f(t) \cdot \Delta_g(f(t))$ in $t = t_0$ stetig. Also ist $g \circ f$ in t_0 differenzierbar und $(g \circ f)'(t_0) = \Delta_g(f(t_0)) \cdot \Delta_f(t_0) = g'(f(t_0)) \cdot f'(t_0)$. ■

1.9. Beispiele

A. Ist f differenzierbar, so ist auch $t \mapsto e^{f(t)}$ differenzierbar, und es gilt:

$$(e^f)'(t) = f'(t) \cdot e^{f(t)}.$$

Will man etwa die Funktion $g(x) := e^{x \cdot \sin(x^2)}$ differenzieren, so kann man $f(x) := x \cdot \sin(x^2)$ setzen. Es ist

$$f'(x) = 1 \cdot \sin(x^2) + x \cdot \cos(x^2) \cdot 2x = \sin(x^2) + 2x^2 \cdot \cos(x^2),$$

unter Verwendung von Produkt- und Kettenregel. Daraus folgt:

$$g'(x) = e^{x \cdot \sin(x^2)} \cdot (\sin(x^2) + 2x^2 \cdot \cos(x^2)).$$

B. Die allgemeine Potenzfunktion $t \mapsto a^t$ kann in der Form

$$a^t = \exp(\ln(a) \cdot t)$$

geschrieben werden. Daher ist $\boxed{(a^t)' = \ln(a) \cdot a^t, \text{ für } a > 0,\ t \in \mathbb{R}.}$

Wir beschäftigen uns jetzt mit höheren Ableitungen. Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine in jedem Punkt $t \in I$ differenzierbare Funktion. Ist die Ableitung $f' : I \to \mathbb{R}$ in $t_0 \in I$ noch ein weiteres Mal differenzierbar, so sagt man, f ist in t_0 ***zweimal differenzierbar***, und man schreibt:

$$f''(t_0) := (f')'(t_0).$$

Entsprechend definiert man die dritte Ableitung $f'''(t_0) := (f'')'(t_0)$, die 4. Ableitung $f^{(4)}(t_0) := (f''')'(t_0)$ und schließlich induktiv:

Definition

Ist f auf I $(n-1)$-mal differenzierbar und die $(n-1)$-te Ableitung $f^{(n-1)}$ in t_0 noch ein weiteres Mal differenzierbar, so sagt man, f ist in t_0 ***n-mal differenzierbar***, und die ***n-te Ableitung*** in t_0 wird definiert durch

$$f^{(n)}(t_0) := (f^{(n-1)})'(t_0).$$

Bemerkung: Manchmal benutzt man auch die Leibniz'sche Schreibweise: An Stelle von $f^{(n)}$ schreibt man auch $\dfrac{d^n f}{dt^n}$. Höhere Ableitungen von komplexwertigen Funktionen definiert man analog.

Definition

Ist $f : I \to \mathbb{C}$ überall k-mal differenzierbar und $f^{(k)}$ auf I noch stetig, so nennt man f auf I ***k-mal stetig differenzierbar***.

Die Menge der k-mal stetig differenzierbaren Funktionen auf I bezeichnet man mit $\mathscr{C}^k(I)$, die Menge der beliebig oft differenzierbaren Funktionen auf I mit $\mathscr{C}^\infty(I)$.

Bemerkung: $\mathscr{C}^k(I)$ ist ein $\mathbb{C}$-Vektorraum.

1.10. Beispiel

Sei $f(t) := e^{t^2}$. Wir wollen sehen, dass f beliebig oft differenzierbar ist, und wir suchen nach einer Beschreibung der n-ten Ableitung. Dazu berechnen wir zunächst die ersten drei Ableitungen:

$$\begin{aligned} f'(t) &= 2t \cdot e^{t^2}, \\ f''(t) &= 2 \cdot e^{t^2} + 2t \cdot (2t \cdot e^{t^2}) \;=\; (2 + 4t^2) \cdot e^{t^2}, \\ f^{(3)}(t) &= 8t \cdot e^{t^2} + (2 + 4t^2) \cdot (2t \cdot e^{t^2}) \;=\; (12t + 8t^3) \cdot e^{t^2}. \end{aligned}$$

Die Versuche lassen folgendes vermuten:

$$f^{(n)}(t) = p(t) \cdot e^{t^2},$$

mit einem Polynom $p(t)$ vom Grad n, das nur gerade bzw. nur ungerade Potenzen von t enthält, je nachdem, ob n gerade oder ungerade ist. Für kleine n haben wir das verifiziert. Also können wir einen Induktionsbeweis führen. Ist die Formel für $n \ge 1$ richtig, so gilt:

$$f^{(n+1)}(t) = (p'(t) + 2t \cdot p(t)) \cdot e^{t^2}.$$

Ist etwa $n = 2k$, so enthält $p(t)$ nur gerade Potenzen von t. Aber dann ist $p'(t)$ ein Polynom vom Grad $n-1$ und $2t \cdot p(t)$ ein Polynom vom Grad $n+1$, und beide enthalten nur ungerade Potenzen von t. Ähnlich funktioniert es im Falle $n = 2k + 1$.

Also gilt die Formel auch für $n + 1$ und damit für alle n.

Wir wollen uns jetzt mit bijektiven Funktionen und ihren Umkehrfunktionen befassen. Im zweiten Kapitel haben wir schon gezeigt, dass eine stetige injektive Funktion f auf einem abgeschlossenen Intervall $I = [a, b]$ streng monoton, $J := f(I)$ wieder ein abgeschlossenes Intervall und $f^{-1} : J \to I$ stetig ist. Ist I ein beliebiges (z.B. offenes) Intervall, so ist jeder Punkt von I Element eines abgeschlossenen Teilintervalls $I' \subset I$. Also überträgt sich die obige Aussage auch auf beliebige Intervalle.

1.11. Ableitung der Umkehrfunktion

Es seien $I, J \subset \mathbb{R}$ Intervalle. Ist $f : I \to J$ bijektiv, stetig, in $x_0 \in I$ differenzierbar und $f'(x_0) \neq 0$, so ist die Umkehrfunktion $f^{-1} : J \to I$ in $y_0 := f(x_0)$ differenzierbar, und es gilt:

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}.$$

Beweis: Wir schreiben $f(x) = f(x_0) + \Delta(x) \cdot (x - x_0)$, mit einer in x_0 stetigen Funktion Δ. Dann ist $\Delta(x_0) = f'(x_0) \neq 0$.

Wäre $\Delta(x) = 0$ für ein $x \in I$, so wäre $x \neq x_0$ und $f(x) = f(x_0)$. Wegen der Injektivität von f ist das ausgeschlossen. Also ist $\Delta(x) \neq 0$ für alle $x \in I$, und wir können die obige Gleichung nach x auflösen:

$$x = x_0 + \frac{f(x) - f(x_0)}{\Delta(x)}.$$

Setzen wir $f^{-1}(y)$ für x und $f^{-1}(y_0)$ für x_0 ein, so erhalten wir:

$$f^{-1}(y) = f^{-1}(y_0) + \frac{1}{\Delta(f^{-1}(y))} \cdot (y - y_0).$$

Da f das Intervall I bijektiv und stetig auf das Intervall J abbildet, ist auch die Umkehrfunktion f^{-1} in y_0 stetig. Also ist $1/\Delta(f^{-1}(y))$ in y_0 stetig und f^{-1} in y_0 differenzierbar, mit

$$(f^{-1})'(y_0) = \frac{1}{\Delta(x_0)} = \frac{1}{f'(x_0)}.$$

■

1.12. Beispiele

A. $\exp : \mathbb{R} \to \mathbb{R}_+$ ist bijektiv und differenzierbar, und $\exp'(x) = \exp(x)$ ist stets > 0. Also ist auch die Umkehrfunktion $\ln : \mathbb{R}_+ \to \mathbb{R}$ überall differenzierbar, mit $(\ln)'(e^x) = 1/e^x$, also

$$\boxed{\ln'(y) = \frac{1}{y}.}$$

B. Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}_+$ eine differenzierbare Funktion. Dann ist auch $g := \ln \circ f : I \to \mathbb{R}$ differenzierbar, und es gilt:

$$g'(x) = (\ln \circ f)'(x) = \ln'(f(x)) \cdot f'(x) = f'(x)/f(x).$$

Man nennt diesen Ausdruck auch die „logarithmische Ableitung“ von f.

C. Die Funktion x^α ist für festes $\alpha \neq 0$ und $x > 0$ definiert durch $x^\alpha := e^{\alpha \ln(x)}$. Daher ist $(x^\alpha)' = \alpha \cdot x^{-1} \cdot e^{\alpha \ln(x)}$, also

$$(x^\alpha)' = \alpha \cdot x^{\alpha-1}, \text{ für } x > 0 \text{ und } \alpha \neq 0.$$

Speziell ist $\sqrt[n]{x} = x^{1/n}$ für $x > 0$ differenzierbar, mit

$$(\sqrt[n]{x})' = \frac{1}{n} \cdot x^{-(n-1)/n},$$

also z.B. $(\sqrt{x})' = \dfrac{1}{2\sqrt{x}}$ und $(\sqrt[3]{x})' = \dfrac{1}{3\sqrt[3]{x^2}}$, jeweils für $x > 0$.

Die Funktion $x^x = e^{x \ln(x)}$ ergibt dagegen beim Differenzieren:

$$(x^x)' = (\ln x + 1) \cdot x^x.$$

D. Die Funktion $\tan = \sin / \cos : (-\pi/2, \pi/2) \to \mathbb{R}$ ist differenzierbar. Da der Cosinus auf $[0, \pi/2)$ streng monoton fällt (und der Sinus dementsprechend streng monoton steigt), ist der Tangens dort streng monoton wachsend. Weil der Cosinus eine gerade und der Sinus eine ungerade Funktion ist, wächst der Tangens auch auf $(-\pi/2, 0]$ streng monoton. Weil $\tan x$ für $x < \pi/2$ und $x \to \pi/2$ gegen $+\infty$ strebt, bildet der Tangens $(-\pi/2, \pi/2)$ bijektiv auf $\mathbb{R}$ ab. Die Ableitung $\tan'(x) = 1 + \tan^2(x)$ hat auf $(-\pi/2, \pi/2)$ keine Nullstelle. Also ist die Umkehrfunktion auf ganz $\mathbb{R}$ differenzierbar. Man bezeichnet diese Umkehrfunktion als ***Arcustangens*** (in Zeichen arctan). Dann gilt:

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

Dieses Ergebnis sollte man sich für später merken: Die Ableitung des Arcustangens ist eine rationale Funktion!

Definition

$f : I \to \mathbb{R}$ hat in $x_0 \in I$ ein ***lokales Maximum*** (bzw. ein ***lokales Minimum***), falls es ein $\varepsilon > 0$ gibt, so dass gilt:

$$f(x) \le f(x_0) \text{ (bzw. } f(x) \ge f(x_0) \text{) für } x \in I \text{ und } |x - x_0| < \varepsilon\,.$$

In beiden Fällen sagt man, f hat in x_0 einen ***(lokalen) Extremwert***.

Man beachte: Ist f in der Nähe von x_0 konstant, so hat f dort nach unserer Definition auch einen Extremwert! Das entspricht nicht ganz der Vorstellung von „Hochpunkten“ und „Tiefpunkten“. Wir führen deshalb noch einen zusätzlichen Begriff ein:

Definition

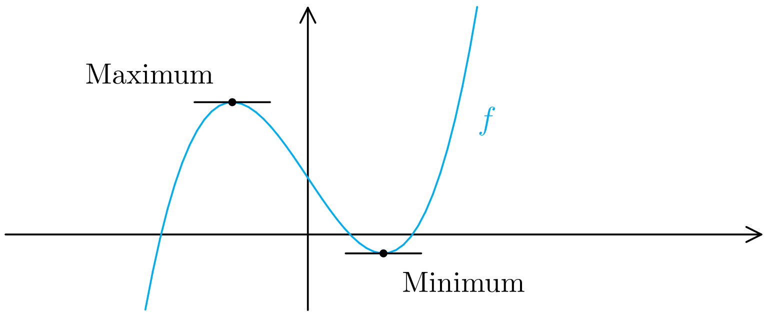
$f : I \to \mathbb{R}$ hat in $x_0 \in I$ ein ***isoliertes Maximum*** (bzw. ein ***isoliertes Minimum***), falls es ein $\varepsilon > 0$ gibt, so dass gilt:

$f(x) < f(x_0)$ (im Falle des Maximums), bzw. $f(x) > f(x_0)$ (im Falle des Minimums), für $x \in I$, $|x - x_0| < \varepsilon$ und $x \ne x_0$.

1.13. „Notwendiges Kriterium“ für Extremwerte

Sei I ein Intervall, x_0 ein innerer Punkt von I und $f : I \to \mathbb{R}$ in x_0 differenzierbar. Wenn f in x_0 ein lokales Extremum besitzt, dann ist $f'(x_0) = 0$.

Beweis: Anschaulich ist klar, dass der Graph von f in x_0 eine waagerechte Tangente besitzen muss:



Wir müssen nun einen Beweis finden, der nicht auf der Anschauung beruht. Dazu verwenden wir die Darstellung $f(x) = f(x_0) + (x - x_0) \cdot \Delta(x)$, mit einer in x_0 stetigen Funktion Δ.

Hat f in x_0 ein lokales Maximum, so ist $f(x) \le f(x_0)$ für x nahe bei x_0, also $(x - x_0) \cdot \Delta(x) \le 0$. Ist $x < x_0$, so ist $x - x_0 < 0$ und daher $\Delta(x) \ge 0$. Ist dagegen $x > x_0$, so ist $\Delta(x) \le 0$. Aber dann muss $\lim\limits_{x \to x_0} \Delta(x) = 0$ sein, also $f'(x_0) = 0$.

Im Falle eines lokalen Minimums ist $\Delta(x) \le 0$ links von x_0 und ≥ 0 rechts von x_0. Genau wie oben folgt auch dann, dass $f'(x_0) = 0$ ist. ∎

Aus der Existenz eines lokalen Extremums folgt also das Verschwinden der ersten Ableitung. Das bedeutet, dass die Eigenschaft „$f'(x_0) = 0$“ dafür notwendig ist, dass in x_0 ein lokales Extremum vorliegt, und deshalb spricht man vom „notwendigen Kriterium“.

1.14. Beispiele

A. Sei $I = [-1, 1]$, $f : I \to \mathbb{R}$ definiert durch $f(x) := x^2$. Dann gilt für alle $x \in I : f(x) \ge 0 = f(0)$. Also hat f in $x_0 := 0$ ein (sogar isoliertes) lokales Minimum. Und tatsächlich besitzt $f'(x) = 2x$ in x_0 eine Nullstelle.

Die Ableitung der Funktion $h(x) := x^3$ verschwindet auch in $x = 0$, aber h hat kein lokales Extremum in diesem Punkt, denn für $x < 0$ ist $f(x) < 0$ und für $x > 0$ ist $f(x) > 0$. Das zeigt, dass das notwendige Kriterium allein noch nicht hinreichend ist.

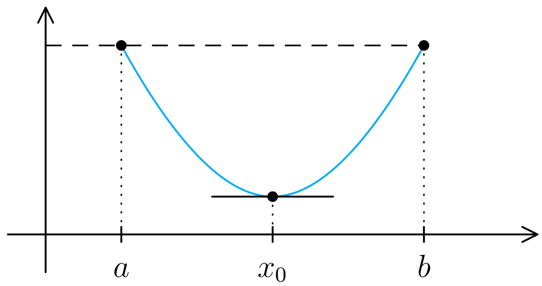
B. $g : I \to \mathbb{R}$ mit $g(x) := |x|$ hat ebenfalls in $x_0 = 0$ ein lokales Minimum. Aber weil $|x|$ dort nicht differenzierbar ist, kann man das Kriterium nicht anwenden.

In den Punkten $x = -1$ und $x = +1$ hat $|x|$ (als Funktion auf I) jeweils ein lokales Maximum. Auch dazu liefert das notwendige Kriterium keinen Beitrag, denn die Punkte liegen nicht im Innern von I.

3.2 Der Mittelwertsatz

2.1. Der Satz von Rolle

Sei $f : I := [a, b] \to \mathbb{R}$ *stetig und im Innern von* I *differenzierbar. Ist* $f(a) = f(b)$*, so gibt es einen Punkt* x_0 *im Innern von* I *mit* $f'(x_0) = 0$.



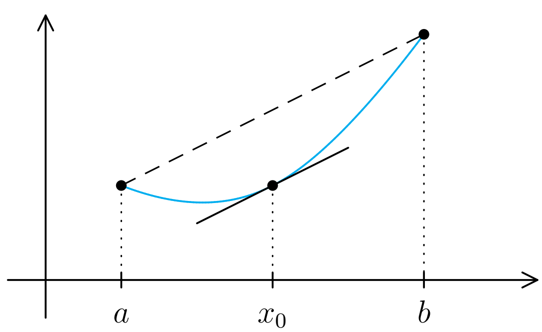
BEWEIS: Sei $c := f(a) = f(b)$. Ist $f(x) \equiv c$ auf ganz I, so ist auch $f'(x) \equiv 0$. Ist f auf I nicht konstant, so muss das Minimum oder das Maximum von f im Innern von I liegen. Und dort muss dann f' verschwinden. ■

Jetzt folgt:

2.2. Der 1. Mittelwertsatz der Differentialrechnung

Sei $f : I := [a, b] \to \mathbb{R}$ *stetig und im Innern von* I *differenzierbar. Dann gibt es einen Punkt* x_0 *im Innern von* I *mit*

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$



BEWEIS: Sei $L : \mathbb{R} \to \mathbb{R}$ die eindeutig bestimmte affin-lineare Funktion mit $L(a) = f(a)$ und $L(b) = f(b)$ (also die Sekante durch $(a, f(a))$ und $(b, f(b))$). Dann ist

$$L(x) = f(a) + m \cdot (x - a), \qquad \text{mit } m := \frac{f(b) - f(a)}{b - a}.$$

Setzen wir $g := f - L$ auf I, so ist $g(a) = g(b) = 0$ und $g'(x) = f'(x) - m$. Nach dem Satz von Rolle, angewandt auf die Funktion g, existiert ein Punkt x_0 im Innern von I mit $g'(x_0) = 0$, also $f'(x_0) = m$. ■

Bemerkung: Ein paar Worte zu den Voraussetzungen: Die Stetigkeit von f auf dem ganzen Intervall wird gebraucht, um – beim Satz von Rolle – die Existenz eines globalen Maximums **und** eines globalen Minimums zu sichern. Ist f nicht konstant, so muss wenigstens eins von beiden im Innern des Intervalls liegen. Um dort das Verschwinden der Ableitung zu erhalten, brauchen wir natürlich die Differenzierbarkeit in allen inneren Punkten des Intervalls.

2.3. Funktionen mit verschwindender Ableitung

Sei $f : I \to \mathbb{R}$ stetig und im Inneren von I differenzierbar. Ist $f'(x) \equiv 0$, so ist f konstant.

BEWEIS: Sei $I = [a, b]$, $a \le x_1 < x_2 \le b$. Nach dem Mittelwertsatz existiert ein x_0 mit $x_1 < x_0 < x_2$ und

$$0 = f'(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Das ist nur möglich, wenn $f(x_1) = f(x_2)$ ist. Und da die Punkte x_1 und x_2 beliebig gewählt werden können, ist f konstant. ∎

Der Satz bleibt auch für komplexwertige Funktionen richtig.

2.4. Hilfssatz

Ist $c \in \mathbb{C}$, so ist $f(t) := e^{ct}$ auf ganz $\mathbb{R}$ differenzierbar und $f'(t) = c \cdot e^{ct}$ für $t \in \mathbb{R}$.

BEWEIS: Sei $c = a + \mathrm{i}\,b$. Wir wissen, dass $(e^{at})' = a \cdot e^{at}$ und auch $(e^{\mathrm{i}\,t})' = \mathrm{i}\; e^{\mathrm{i}\,t}$ ist. Mit der Kettenregel folgt nun, dass $(e^{\mathrm{i}\,bt})' = \mathrm{i}\,be^{\mathrm{i}\,bt}$ ist, und die Produktregel liefert

$$(e^{ct})' = \left(e^{at} \cdot e^{\mathrm{i}\,bt}\right)' = a \cdot e^{at} \cdot e^{\mathrm{i}\,bt} + e^{at} \cdot (\mathrm{i}\;b) \cdot e^{\mathrm{i}\,bt} = c \cdot e^{ct}.$$

∎

2.5. Die Differentialgleichung $y' - c\,y = 0$

Seien $c, \gamma \in \mathbb{C}$ zwei Konstanten und $t_0 \in \mathbb{R}$. Dann gibt es genau eine differenzierbare Funktion $f : \mathbb{R} \to \mathbb{C}$, so dass gilt:

1. *$f'(t) = c \cdot f(t)$ für alle $t \in \mathbb{R}$.*
2. *$f(t_0) = \gamma$.*

Die eindeutig bestimmte Lösung ist gegeben durch $f(t) = \gamma \cdot e^{c(t-t_0)}$.

Beweis: Wäre $c \in \mathbb{R}$ und f eine reellwertige differenzierbare Funktion ohne Nullstellen, so könnte man schließen: $(\ln \circ f)'(t) = f'(t)/f(t) = c$, also $\ln(f(t)) = ct + d$ und $f(t) = \exp(ct + d)$. Diesen Schluss können wir hier nicht durchführen, deshalb verwenden wir einen kleinen Trick.

1) Ist f eine Lösung, so setzen wir $F(t) := f(t) \cdot e^{-ct}$. Dann folgt:

$$F'(t) = (f'(t) - c \cdot f(t)) \cdot e^{-ct} \equiv 0.$$

Also ist F konstant. Wegen $F(t) = F(t_0) = \gamma e^{-ct_0}$ ist dann $f(t) = \gamma \cdot e^{c(t-t_0)}$. Das zeigt die Eindeutigkeit der Lösung.

2) Wählt man umgekehrt f wie oben, so ist $f(t_0) = \gamma$ und $f'(t) = c\gamma e^{c(t-t_0)} = c \cdot f(t)$. ∎

Als nächstes befassen wir uns mit der **Differentialgleichung**

$$y'' + ay' + by = 0 \quad (\text{mit } a, b \in \mathbb{R}).$$

Unter einer ***Lösung*** dieser DGL (über einem Intervall I) verstehen wir eine zweimal differenzierbare Funktion $f : I \to \mathbb{C}$ mit $f''(t) + a \cdot f'(t) + b \cdot f(t) \equiv 0$ auf I.

Zur Bestimmung der Lösungsgesamtheit wird etwas lineare Algebra benötigt.

Erinnerung:

Elemente $X_1, \ldots, X_n$ eines (komplexen) Vektorraumes V heißen *linear unabhängig*, falls aus $c_1X_1 + \cdots + c_nX_n = 0$ (mit $c_i \in \mathbb{C}$) stets $c_1 = \ldots = c_n = 0$ folgt. Ein System $\{X_1, \ldots, X_n\}$ von linear unabhängigen Vektoren heißt eine *Basis* von V, wenn sich jeder Vektor $X \in V$ als Linearkombination der X_i schreiben lässt:

$$X = c_1X_1 + \cdots + c_nX_n.$$

In der Linearen Algebra wird gezeigt: Besitzt V eine Basis aus n Elementen, so besteht auch jede andere Basis von V aus genau n Elementen. Die Zahl n nennt man dann die *Dimension* von V. Besitzt ein Vektorraum keine endliche Basis, so nennt man ihn unendlich-dimensional. Letzteres ist bei Räumen von Funktionen leider der Normalfall.

Es seien V und W zwei (komplexe) Vektorräume. Eine Abbildung $f : V \to W$ heißt *linear* oder ein *Homomorphismus*, falls gilt:

$$\begin{aligned} f(X+Y) &= f(X) + f(Y) \text{ für alle } X, Y \in V \\ \text{und} \quad f(cX) &= c \cdot f(X) \text{ für } c \in \mathbb{C} \text{ und } X \in V. \end{aligned}$$

Im Falle unendlich-dimensionaler Räume spricht man auch gerne von einem *(linearen) Operator*.

Die Menge $\text{Ker}(f) := \{X \in V : f(X) = 0\}$ bezeichnet man als *Kern* der linearen Abbildung. Der Kern ist ein Untervektorraum von V, der zumindest den Nullvektor enthält. Besteht er nur aus dem Nullvektor, so ist f injektiv: Sind nämlich $X_1, X_2 \in V$ mit $f(X_1) = f(X_2)$, so ist $f(X_1 - X_2) = f(X_1) - f(X_2) = 0$, also $X_1 - X_2 \in \text{Ker}(f)$. Daraus folgt, dass $X_1 - X_2 = 0$, also $X_1 = X_2$ ist.

Die Menge $\text{Im}(f) := f(V) = \{Y \in W : \exists\, X \in V \text{ mit } f(X) = Y\}$ nennt man das *Bild* von f. Das Bild ist ein Unterraum von W. Ist sogar $\text{Im}(f) = W$, so ist f surjektiv.

Sind V und W endlich-dimensional, so gilt die folgende Dimensionsformel:

$$\dim V = \dim(\operatorname{Ker} f) + \dim(\operatorname{Im} f).$$

Ein *Isomorphismus* zwischen zwei Vektorräumen V und W ist eine bijektive lineare Abbildung $f : V \to W$. Dann ist auch $f^{-1} : W \to V$ linear, und natürlich ist $\operatorname{Ker}(f) = \{0\}$ und $\operatorname{Im}(f) = W$. Sind die Vektorräume endlich-dimensional, so folgt aus der Dimensionsformel, dass es nur dann einen Isomorphismus zwischen V und W geben kann, wenn $\dim(V) = \dim(W)$ ist.

Definition

Sei $D : \mathscr{C}^1(I) \to \mathscr{C}^0(I)$ definiert durch $D[f] := f'$. Auf $\mathscr{C}^2(I)$ definiert man $D^2 := D \circ D$ durch $D^2[f] = f''$.

Dann sind D und D^2 „lineare Operatoren".

2.6. Satz

Sei $p(x) = x^2 + ax + b = (x - c_1)(x - c_2)$ ein Polynom zweiten Grades. Dann ist

$$(D - c_1 \cdot \mathrm{id}) \circ (D - c_2 \cdot \mathrm{id}) = D^2 + a \cdot D + b \cdot \mathrm{id} \quad \text{(auf } \mathscr{C}^2(I)\text{)}.$$

Beweis: Es ist $(x - c_1)(x - c_2) = x^2 - (c_1 + c_2)x + c_1c_2$, also $c_1 + c_2 = -a$ und $c_1c_2 = b$.

Dann ist

$$\begin{aligned}
(D - c_1 \cdot \mathrm{id}) \circ (D - c_2 \cdot \mathrm{id})[h] &= (D - c_1 \cdot \mathrm{id})[h' - c_2h] \\
&= h'' - c_2h' - c_1h' + c_1c_2h \;=\; h'' + ah' + bh \\
&= (D^2 + a \cdot D + b \cdot \mathrm{id})[h].
\end{aligned}$$

■

2.7. Satz

Ist c eine Nullstelle des Polynoms $p(x) = x^2 + ax + b$, so ist $f(t) := e^{ct}$ eine Lösung der Differentialgleichung $y'' + ay' + by = 0$.

Beweis: Es ist $f'(t) = c \cdot e^{ct}$ und $f''(t) = c^2 \cdot e^{ct}$, also

$$f''(t) + af'(t) + bf(t) = (c^2 + ac + b)e^{ct} \equiv 0.$$

■

2.8. Satz

1) Sei $c \in \mathbb{C}$. Dann sind die Funktionen $f(t) := e^{ct}$ und $g(t) := t \cdot e^{ct}$ linear unabhängig über $\mathbb{C}$.

2) Sind $c_1 \neq c_2$ zwei komplexe Zahlen. Dann sind die Funktionen $f_1(t) := e^{c_1 t}$ und $f_2(t) := e^{c_2 t}$ linear unabhängig über $\mathbb{C}$.

Beweis: 1) Sei $\alpha \cdot e^{ct} + \beta \cdot te^{ct} \equiv 0$. Multipliziert man beide Seiten mit e^{-ct}, so erhält man: $\alpha + \beta \cdot t \equiv 0$, also $\alpha = 0$ ($t = 0$ einsetzen) und dann $\beta = 0$ ($t = 1$ einsetzen).

2) Sei $\alpha \cdot e^{c_1 t} + \beta e^{c_2 t} \equiv 0$, also $\alpha + \beta \cdot e^{(c_2 - c_1)t} \equiv 0$. Differentiation nach t ergibt: $\beta(c_2 - c_1) \cdot e^{(c_2-c_1)t} \equiv 0$, also $\beta = 0$ und dann $\alpha = 0$. ∎

Das Polynom $p(x) = x^2 + ax + b$ nennt man das ***charakteristische Polynom*** der DGL $y'' + ay' + by = 0$. Hat $p(x)$ zwei verschiedene Nullstellen c_1, c_2, so sind $f_1(t) = e^{c_1 t}$ und $f_2(t) = e^{c_2 t}$ zwei linear unabhängige Lösungen der DGL.

Hat $p(x) = (x - c)^2$ eine Nullstelle c der Ordnung 2, dann ist $a = -2c$ und $b = c^2$. Wir wissen schon, dass $f(t) = e^{ct}$ eine Lösung der DGL ist. Sei $g(t) := te^{ct}$. Dann ist

$$g'(t) = (1 + tc)e^{ct} \quad \text{und} \quad g''(t) = (2c + tc^2)e^{ct},$$

also $g''(t) - 2cg'(t) + c^2 g(t) = (2c + tc^2 - 2c - 2tc^2 + tc^2)e^{ct} \equiv 0$. Das bedeutet, dass in diesem Fall f und g zwei linear unabhängige Lösungen der DGL sind.

2.9. Satz

Sind $A, B \in \mathbb{C}$ gegeben, so gibt es höchstens eine Lösung f der DGL $y'' + ay' + by = 0$ mit $f(t_0) = A$ und $f'(t_0) = B$.

Beweis: Es reicht zu zeigen, dass $h(t) \equiv 0$ die einzige Lösung der DGL mit $h(t_0) = h'(t_0) = 0$ ist. Sind nämlich f_1, f_2 zwei Lösungen mit den gegebenen Anfangsbedigungen, so betrachte man $h := f_1 - f_2$.

Sei nun $L := D^2 + a \cdot D + b \cdot \mathrm{id} = (D - c_1 \cdot \mathrm{id}) \circ (D - c_2 \cdot \mathrm{id})$ und h eine Funktion mit $L[h] = 0$ und $h(t_0) = h'(t_0) = 0$. Wir setzen $h_1 := h' - c_2 h$. Dann ist $h_1(t_0) = h'(t_0) - c_2 h(t_0) = 0$ und $h_1' - c_1 h_1 = L[h] = 0$, also $h_1(t) = h_1(t_0)e^{c_1(t-t_0)} \equiv 0$. Daraus folgt, dass $h(t) = h(t_0)e^{c_2(t-t_0)} \equiv 0$ ist. ∎

Sei weiterhin $L := D^2 + a \cdot D + b \cdot \mathrm{id}$. Dann ist $\mathrm{Ker}(L) \subset \mathscr{C}^2(I)$ der Raum der Lösungen der DGL $y'' + ay' + by = 0$ über I.

2.10. Folgerung

Es ist $\dim_{\mathbb{C}} \mathrm{Ker}(L) = 2$.

Beweis: Es ist schon klar, dass $\dim_{\mathbb{C}} \operatorname{Ker}(L) \geq 2$ ist. Nun sei $A : \operatorname{Ker}(L) \to \mathbb{C}^2$ definiert durch

$$A(f) := \big(f(t_0), f'(t_0)\big).$$

Die Abbildung A ist offensichtlich linear, und nach dem Satz von oben ist A injektiv. Also ist $\dim_{\mathbb{C}} \operatorname{Ker}(L) \leq 2$. ■

In jedem Fall kann eine Basis des Lösungsraumes (ein sogenanntes „Fundamentalsystem von Lösungen") angegeben werden.

Nun noch ein paar Worte über die inhomogene Gleichung $L[f] = q(t)$, mit einer stetigen Funktion q. Den allgemeinen Fall können wir mit unseren derzeit verfügbaren Mitteln nicht lösen. Wir beschränken uns auf den Fall

$$q(t) = h(t)e^{\mu t}, \text{ mit einem Polynom } h(t).$$

1. Ist $\lambda = \mu$, so ist $(D - \lambda \cdot \mathrm{id})[g(t)e^{\mu t}] = g'(t)e^{\mu t}$.

2. Ist $\mu \neq \lambda$, so ist $(D - \lambda \cdot \mathrm{id})[g(t)e^{\mu t}] = \big(g'(t) + (\mu - \lambda)g(t)\big)e^{\mu t}$.

Offensichtlich kann man in beiden Fällen $g(t)$ als Polynom wählen, so dass $g(t)e^{\mu t}$ eine Lösung der Gleichung $(D - \lambda \cdot \mathrm{id})[f] = q(t)$ ist. Weil es stets Zahlen c_1, c_2 mit $L = (D - c_1 \cdot \mathrm{id}) \circ (D - c_2 \cdot \mathrm{id})$ gibt, kann auch die Gleichung $L[f] = q(t)$ gelöst werden.

Wir kommen nun zu weiteren Anwendungen des Mittelwertsatzes.

2.11. Ableitung und Monotonie

Sei $f : I := [a, b] \to \mathbb{R}$ stetig und im Inneren von I differenzierbar.

1. *f ist genau dann auf I monoton wachsend (bzw. fallend), wenn $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$) für alle $x \in (a, b)$ ist.*

2. *Ist sogar $f'(x) > 0$ für alle $x \in I$ (bzw. $f'(x) < 0$ für alle $x \in I$), so ist f streng monoton wachsend (bzw. fallend).*

Beweis: Wir beschränken uns auf wachsende Funktionen.

1. a) Ist f monoton wachsend, so sind alle Differenzenquotienten ≥ 0, und daher ist auch überall $f'(x) \geq 0$.

b) Nun sei $f'(x) \geq 0$ für alle $x \in (a, b)$. Ist $x_1 < x_2$, so gibt es nach dem Mittelwertsatz ein x_0 mit $x_1 < x_0 < x_2$, so dass gilt:

$$0 \leq f'(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}, \text{ also } f(x_1) \leq f(x_2).$$

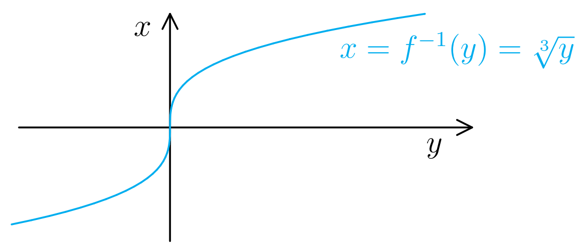
2) Ist f monoton wachsend, aber nicht streng monoton wachsend, so gibt es Punkte $x_1, x_2 \in I$ mit $x_1 < x_2$ und $f(x_1) = f(x_2)$. Dann muss f auf dem Teilintervall

$[x_1, x_2]$ konstant sein und die Ableitung dort verschwinden. Also kann f' nicht überall positiv sein. ■

2.12. Beispiele

A. Sei $f(x) := x^3$. Dann ist $f'(x) = 3x^2$ und $f''(x) = 6x$. Da überall $f'(x) \geq 0$ ist, wächst f auf ganz $\mathbb{R}$ monoton. Außerhalb des Nullpunktes ist $f'(x)$ sogar positiv, also wächst f dort streng monoton. Aber eine monotone Funktion, die auf keinem Intervall positiver Länge konstant ist, muss insgesamt streng monoton sein. Obwohl $f(x) = x^3$ überall streng monoton steigt, ist $f'(0) = 0$.

$f : \mathbb{R} \to \mathbb{R}$ ist bijektiv und stetig, besitzt also eine überall stetige Umkehrfunktion. Offensichtlich ist $f^{-1}(y) = \sqrt[3]{y}$. Wäre diese Funktion in $y = 0$ differenzierbar, so müsste $0 = f'(0) \cdot (f^{-1})'(0) = (f \circ f^{-1})'(0) = 1$ sein. Widerspruch!



So sieht man, dass f^{-1} im Nullpunkt nicht differenzierbar sein kann. Der Grund dafür ist, dass f^{-1} im Nullpunkt eine senkrechte Tangente besitzt.

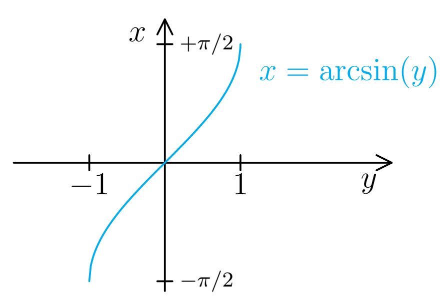
B. Es ist $\sin'(x) = \cos(x) > 0$ für $-\dfrac{\pi}{2} < x < \dfrac{\pi}{2}$. Also ist $\sin(x)$ dort streng monoton wachsend und damit injektiv (was wir natürlich schon wissen). Die Umkehrfunktion von

$$\sin : (-\frac{\pi}{2}, \frac{\pi}{2}) \to (-1, 1)$$

ist die Funktion

$$\arcsin : (-1, 1) \to (-\frac{\pi}{2}, \frac{\pi}{2}),$$

der sogenannte ***Arcussinus***.



Für die Ableitung gilt:
$$\begin{aligned}\arcsin'(y) &= \frac{1}{\sin'(\arcsin(y))} = \frac{1}{\cos(\arcsin(y))} \\ &= \frac{1}{\sqrt{1 - \sin^2(\arcsin(y))}} = \frac{1}{\sqrt{1 - y^2}}.\end{aligned}$$

Da der Sinus auf allen Intervallen $((k - \frac{1}{2})\pi, (k + \frac{1}{2})\pi)$ streng monoton ist, gibt es auch dazu Umkehrfunktionen $\arcsin_k(y)$. Man spricht von verschiedenen

Zweigen des Arcussinus. Im Falle $k = 0$ ergibt sich der ***Hauptzweig***, den wir oben behandelt haben.

Beim Cosinus kann man analoge Überlegungen anstellen. Wegen

$$\cos(\pi/2 - \arcsin(x)) = \sin(\arcsin(x)) = x$$

erhält man den ***Arcuscosinus*** als $\arccos(x) = \pi/2 - \arcsin(x)$.

C. Die Hyperbelfunktionen $\sinh(x) = \frac{1}{2}(e^x - e^{-x})$ und $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$ sind offensichtlich überall differenzierbar, und es gilt:

$$\sinh'(x) = \cosh(x) \quad \text{und} \quad \cosh'(x) = \sinh(x).$$

Da $\sinh'(x) > 0$ für alle x ist, ist sinh streng monoton wachsend und somit umkehrbar. Die Umkehrfunktion wird mit arsinh (***Area-Sinus hyperbolicus***) bezeichnet. Die Beziehung $y = \sinh(x) = (e^x - e^{-x})/2$ liefert eine quadratische Gleichung für e^x, nämlich $2y = \big((e^x)^2 - 1\big)/e^x$, also $(e^x)^2 - 2y \cdot e^x - 1 = 0$. Damit ist

$$\boxed{x = \operatorname{arsinh}(y) = \ln(y + \sqrt{1 + y^2}).}$$

Das positive Vorzeichen vor der Wurzel muss gewählt werden, weil $e^x > 0$ ist. Daraus folgt:

$$\operatorname{arsinh}'(y) = 1/\sqrt{1 + y^2}.$$

Der Cosinus hyperbolicus lässt sich nur für $x \geq 0$ oder für $x \leq 0$ umkehren. Die Umkehrfunktion arcosh (***Area-Cosinus hyperbolicus***) ist jeweils für $y \geq 1$ erklärt. Sie ist gegeben durch

$$\boxed{\operatorname{arcosh}(y) = \ln(y \pm \sqrt{y^2 - 1}).}$$

Das Vorzeichen vor der Wurzel hängt davon ab, welchen Teil des Cosinus hyperbolicus man umkehren möchte.

Wir können den Mittelwertsatz noch etwas verallgemeinern:

2.13. Der Satz von Cauchy

Es seien f und g auf $I := [a, b]$ stetig und im Innern von I differenzierbar. Dann gibt es einen Punkt c im Innern von I mit

$$f'(c) \cdot (g(b) - g(a)) = g'(c) \cdot (f(b) - f(a)).$$

Für $g(x) = x$ erhält man den 1. Mittelwertsatz zurück.

Beweis: Ist $g(a) = g(b)$, so gibt es ein $c \in (a, b)$ mit $g'(c) = 0$, und die Gleichung ist offensichtlich erfüllt.

Ist $g(a) \neq g(b)$, so benutzen wir die Hilfsfunktion

$$h(x) := f(x) - \frac{g(x) - g(a)}{g(b) - g(a)} \cdot (f(b) - f(a)).$$

Weil $h(a) = f(a) = h(b)$ ist, gibt es nach dem Satz von Rolle ein c im Innern des Intervalls, so dass $h'(c) = 0$ ist. Aber offensichtlich ist

$$h'(c) = f'(c) - \frac{g'(c)}{g(b) - g(a)} \cdot (f(b) - f(a)).$$

Daraus folgt die gewünschte Gleichung. ■

Es hat sich ergeben, dass der Satz von Cauchy sogar äquivalent zum 1. Mittelwertsatz ist! Sofort folgt nun

2.14. Der 2. Mittelwertsatz der Differentialrechnung

Es seien f und g auf $I := [a, b]$ stetig und im Innern von I differenzierbar. Außerdem sei $g'(x) \neq 0$ im Innern von I.

Dann gibt es einen Punkt c im Innern von I mit

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

Beweis: Wäre $g(a) = g(b)$, so müsste $g'(c) = 0$ für mindestens ein c sein. Das haben wir aber ausgeschlossen. Also können wir in der Formel im Satz von Cauchy durch $g(b) - g(a)$ teilen. ■

Als Anwendung aus dem 2. Mittelwertsatz ergibt sich:

2.15. Erste Regel von de l'Hospital (der Grenzwert $0/0$)

Die Funktionen f und g seien auf dem offenen Intervall $I := (a, b)$ differenzierbar, und es sei $g'(x) \neq 0$ für alle $x \in I$. Außerdem sei

$$\lim_{x \to a+} f(x) = \lim_{x \to a+} g(x) = 0.$$

Wenn $\displaystyle\lim_{x \to a+} \frac{f'(x)}{g'(x)}$ existiert, dann existiert auch $\displaystyle\lim_{x \to a+} \frac{f(x)}{g(x)}$,

und die beiden Grenzwerte sind gleich.

Eine entsprechende Aussage gilt auch für den linksseitigen Grenzwert bei b.

Beweis: Wir betrachten nur den Fall des rechtsseitigen Grenzwertes. Nach Voraussetzung kann man f und g stetig nach a (durch den Wert Null) fortsetzen. Nun sei (x_ν) eine beliebige Folge von Zahlen mit $a < x_\nu < b$ und $\lim_{\nu\to\infty} x_\nu = a$ (nicht notwendig monoton). Nach dem 2. Mittelwertsatz gibt es Zahlen c_ν mit $a < c_\nu < x_\nu$ und

$$\frac{f(x_\nu)}{g(x_\nu)} = \frac{f(x_\nu) - f(a)}{g(x_\nu) - g(a)} = \frac{f'(c_\nu)}{g'(c_\nu)}.$$

Da auch $\lim_{\nu\to\infty} c_\nu = a$ ist, strebt der letzte Quotient nach Voraussetzung gegen $\lim_{x\to a+} f'(x)/g'(x)$. Aber dann ist $\lim_{x\to a+} f(x)/g(x) = \lim_{x\to a+} f'(x)/g'(x)$. ∎

Bemerkung: An Stelle der Annäherung an eine endliche Intervallgrenze kann man auch den Fall $x \to \pm\infty$ betrachten, es gelten analoge Aussagen.

2.16. Beispiele

A. Sei $f(x) := \sin x$ und $g(x) := x$. Da $f(0) = g(0) = 0$ ist und $f'(x)/g'(x) = \cos x$ für $x \to 0$ gegen 1 strebt, ist auch $\lim_{x\to 0} \sin(x)/x = 1$.

B. Sei $f(x) := \ln(1 - x)$ und $g(x) := x + \cos x$. Dann gilt:

$$\frac{f'(x)}{g'(x)} = \frac{1}{(x-1)(1-\sin x)}$$

konvergiert für $x \to 0$ gegen -1. Aber trotzdem darf man l'Hospital **nicht** anwenden, denn es ist zwar $f(0) = 0$, aber $g(0) = 1$.

Tatsächlich ist $\displaystyle\lim_{x\to 0} \frac{f(x)}{g(x)} = \frac{f(0)}{g(0)} = 0.$

2.17. Zweite Regel von de l'Hospital (der Grenzwert ∞/∞)

Die Funktionen f und g seien auf dem offenen Intervall $I := (a, b)$ differenzierbar, und es sei $g'(x) \neq 0$ für alle $x \in I$. Außerdem sei

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

Wenn $\displaystyle\lim_{x\to a+} \frac{f'(x)}{g'(x)} =: c$ *existiert, dann existiert auch* $\displaystyle\lim_{x\to a+} \frac{f(x)}{g(x)}$,

und die beiden Grenzwerte sind gleich.

Eine entsprechende Aussage gilt für die Annäherung an b von links.

Beweis: Man würde erwarten, dass der Beweis nichts Neues bringt, aber die Schlussweise ist doch etwas komplizierter als bei der 1. Regel von de l'Hospital.

Es sei ein $\varepsilon > 0$ vorgegeben, sowie $\varepsilon_1 := \varepsilon/(2 + |c|)$. Dann gibt es ein $\delta > 0$, so dass gilt:

$$\Big|\frac{f'(x)}{g'(x)} - c\Big| < \varepsilon_1 \qquad \text{für } a < x < a + \delta.$$

Wegen $\lim_{x\to a+} g(x) = +\infty$ kann man außerdem annehmen, dass $g(x) > 0$ für $a < x < a + \delta$ ist. Wir wählen ein festes x_0 zwischen a und $a + \delta$. Zu jedem x mit $a < x < x_0$ gibt es nach dem Satz von Cauchy ein $\xi = \xi(x)$ mit $x < \xi < x_0$ und

$$g'(\xi) \cdot \big(f(x) - f(x_0)\big) = f'(\xi) \cdot \big(g(x) - g(x_0)\big).$$

Weil g' auf I nirgends verschwindet, kann man durch $g'(\xi)$ teilen, und anschließend durch $g(x)$. Dann erhält man:

$$\frac{f(x)}{g(x)} = \frac{f(x_0)}{g(x)} + \frac{f'(\xi)}{g'(\xi)}\Big(1 - \frac{g(x_0)}{g(x)}\Big),$$

also

$$\frac{f(x)}{g(x)} - c = \frac{f(x_0)}{g(x)} + \Big(\frac{f'(\xi)}{g'(\xi)} - c\Big) \cdot \Big(1 - \frac{g(x_0)}{g(x)}\Big) - c \cdot \frac{g(x_0)}{g(x)}\,.$$

Ist x nahe genug bei a, so ist $\Big|\dfrac{f(x_0)}{g(x)}\Big| < \varepsilon_1$ und $\Big|\dfrac{g(x_0)}{g(x)}\Big| < \varepsilon_1$.

Dann folgt:

$$\Big|\frac{f(x)}{g(x)} - c\Big| \le 2\varepsilon_1 + |c| \cdot \varepsilon_1 < \varepsilon \text{ für } a < x < x_0.$$

Das zeigt, dass $\lim\limits_{x\to a+} f(x)/g(x) = c$ ist. ∎

2.18. Beispiele

A. Wir wollen $\lim_{x\to 0+} x \cdot \ln(x)$ berechnen. Dazu müssen wir $x \cdot \ln(x)$ zunächst mal als Quotient schreiben:

$$x \cdot \ln(x) = -\frac{-\ln(x)}{x^{-1}},$$

Es ist

$$\lim_{x\to 0+} (-\ln(x)) = \lim_{x\to 0+} x^{-1} = +\infty \qquad \text{und} \qquad \lim_{x\to 0+} \frac{-\ln'(x)}{(x^{-1})'} = \lim_{x\to 0+} x = 0,$$

also auch $\lim_{x\to 0+} x \cdot \ln(x) = 0$.

B. Sei $p(x) = a_k x^k + \cdots + a_1 x + a_0$ ein Polynom vom Grad k, $a_k > 0$. Dann ist $p'(x) = k a_k x^{k-1} + \cdots + a_1$ ein Polynom vom Grad $k - 1$, $p''(x)$ ein Polynom vom Grad $k - 2$, und schließlich $p^{(k)}(x) = k! a_k =: c$ eine Konstante > 0.

Jedes Polynom, dessen höchster Koeffizient positiv ist, strebt für $x \to +\infty$ gegen $+\infty$, denn $1/x$ strebt für $x \to \infty$ gegen Null, und

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = x^n\bigl(1 + a_{n-1}x^{-1} + \cdots + a_1 x^{-n+1} + a_0 x^{-n}\bigr)$$

verhält sich deshalb für $x \to \infty$ wie x^n.

Auch sämtliche Ableitungen der Exponentialfunktion streben für $x \to +\infty$ gegen $+\infty$. Mehrfache Anwendung von l'Hospital ergibt daher

$$\lim_{x\to\infty} \frac{e^x}{p(x)} = \lim_{x\to\infty} \frac{(e^x)'}{p'(x)} = \ldots = \lim_{x\to\infty} \frac{(e^x)^{(k)}}{p^{(k)}(x)} = \lim_{x\to\infty} \frac{e^x}{c} = +\infty.$$

Die Exponentialfunktion wächst stärker als jedes Polynom!

Dagegen gilt:

$$\lim_{x\to\infty} \frac{\ln(x)}{p(x)} = \lim_{x\to\infty} \frac{1}{x \cdot p'(x)} = 0.$$

Die Logarithmusfunktion wächst also schwächer als jedes Polynom.

2.19. Erstes „Hinreichendes Kriterium“ für Extremwerte

Sei I ein Intervall, x_0 ein innerer Punkt von I, $f : I \to \mathbb{R}$ differenzierbar und $f'(x_0) = 0$. Außerdem gebe es ein $\varepsilon > 0$, so dass $f'(x) \le 0$ für $x_0 - \varepsilon < x < x_0$ und $f'(x) \ge 0$ für $x_0 < x < x_0 + \varepsilon$ ist. Dann besitzt f in x_0 ein lokales Minimum. Ist $f'(x) \ge 0$ links von x_0 und ≤ 0 rechts von x_0, so liegt ein lokales Maximum vor.

Beweis: Wir betrachten zunächst den Fall, dass $f'(x) \le 0$ für $x_0 - \varepsilon < x < x_0$ und ≥ 0 für $x_0 < x < x_0 + \varepsilon$ ist.

Anschaulich: Links von x_0 ist die Richtung der Tangente an den Graphen von f negativ (genauer: nicht positiv), rechts von x_0 wird sie positiv (genauer: nicht negativ). Fassen wir den Graphen als Wanderweg auf, so muss ein Wanderer zunächst bergab und nach x_0 wieder bergauf laufen. Bei x_0 wird ein „Tiefpunkt“ erreicht.

Der formale Beweis ist fast noch einfacher. Aus dem Monotoniekriterium folgt nämlich, dass f auf $[x_0 - \varepsilon, x_0]$ monoton fällt und auf $[x_0, x_0 + \varepsilon]$ monoton wächst. Ist also $|x - x_0| < \varepsilon$ und $x < x_0$, so muss auf jeden Fall $f(x) \ge f(x_0)$ sein. Ist $x > x_0$, so ist $x_0 < x$ und wieder $f(x_0) \le f(x)$. Das bedeutet, dass f in x_0 ein lokales Minimum besitzt. Im zweiten Fall wird analog argumentiert. ■

2.20. Zweites „Hinreichendes Kriterium“ für Extremwerte

Sei I ein Intervall, x_0 ein innerer Punkt von I, $f : I \to \mathbb{R}$ einmal differenzierbar und in x_0 zweimal differenzierbar. Außerdem sei $f'(x_0) = 0$.

Ist $f''(x_0) > 0$, so besitzt f in x_0 ein isoliertes lokales Minimum. Ist $f''(x_0) < 0$, so liegt ein isoliertes lokales Maximum vor.

Beweis: Der Beweis wäre ganz einfach, wenn wir wüssten, dass f in einer ganzen Umgebung von x_0 zweimal differenzierbar und f'' in x_0 stetig ist. Wir schaffen den Beweis aber auch unter unseren eingeschränkten Bedingungen. Dazu benutzen wir wieder einmal die alternative Beschreibung der Differenzierbarkeit, angewandt auf die abgeleitete Funktion f' :

$$f'(x) = f'(x_0) + (x - x_0) \cdot \Delta(x),$$

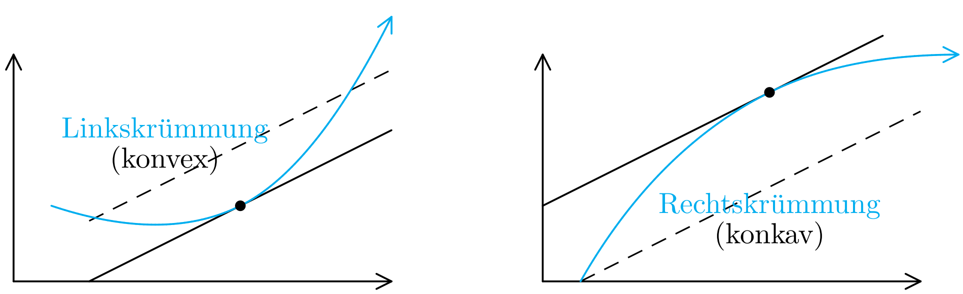
mit einer in x_0 stetigen Funktion Δ und $\Delta(x_0) = f''(x_0)$.

Ist $f''(x_0) > 0$, so gibt es eine Umgebung $U = U_\varepsilon(x_0) \subset I$, so dass $\Delta > 0$ auf U ist. Aber für $x \neq x_0$ ist

$$\Delta(x) = \frac{f'(x) - f'(x_0)}{x - x_0}.$$

Ist also $x < x_0$, so muss $f'(x) < f'(x_0) = 0$ sein, und im Falle $x > x_0$ muss $f'(x) > f'(x_0) = 0$ sein. Damit fällt $f(x)$ streng monoton für $x < x_0$ und wächst streng monoton für $x > x_0$. Das bedeutet, dass f in x_0 ein isoliertes Minimum besitzt. Im Falle des Maximums verläuft der Beweis analog. ■

Was ist die anschauliche Bedeutung der zweiten Ableitung? Wir beschäftigen uns zunächst mit Krümmungen.



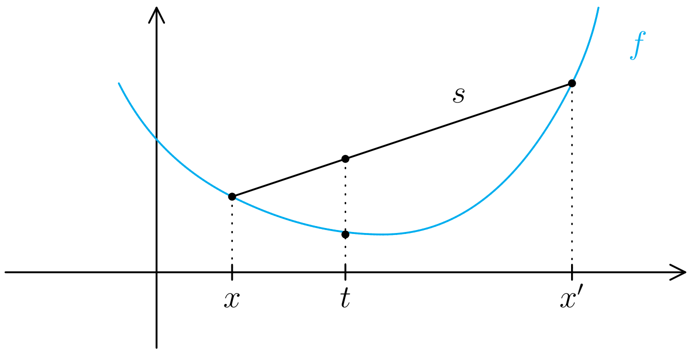
Im Falle einer Linkskrümmung verläuft der Funktionsgraph oberhalb der Tangente, im Falle einer Rechtskrümmung unterhalb. Ersetzt man die Tangente durch eine dazu parallel verlaufende Sekante, so verhält es sich genau umgekehrt.

Definition

Eine Funktion $f : I \to \mathbb{R}$ heißt auf I ***konvex***, falls für alle $x, x' \in I$ mit $x < x'$ und alle $t \in [x, x']$ gilt: $(t, f(t))$ liegt unterhalb der Verbindungsstrecke von $(x, f(x))$ und $(x', f(x'))$.

Die Funktion $f : I \to \mathbb{R}$ heißt auf I ***strikt konvex***, falls für alle $x, x' \in I$ mit $x < x'$ und alle $t \in (x, x')$ gilt: $(t, f(t))$ liegt echt unterhalb der Verbindungsstrecke von $(x, f(x))$ und $(x', f(x'))$.

Ersetzt man „unterhalb“ durch „oberhalb“, so bekommt man die Begriffe ***konkav*** und ***strikt konkav***.



Die Verbindungsstrecke von $(x, f(x))$ und $(x', f(x'))$ ist der Graph der affin-linearen Funktion $s : [x, x'] \to \mathbb{R}$ mit

$$s(t) = s(x, x'; t) := f(x) + (t - x) \cdot \frac{f(x') - f(x)}{x' - x}.$$

Dass $(t, f(t))$ unterhalb (bzw. echt unterhalb) der Strecke liegt, bedeutet, dass $f(t) \le s(t)$ (bzw. $f(t) < s(t)$) ist.

Für $x < t \le x'$ führen wir den Differenzenquotienten

$$D(x, t) := \frac{f(t) - f(x)}{t - x}$$

ein. Dann ist $s(x, x'; t) = f(x) + (t - x) \cdot D(x, x')$. Offensichtlich gilt:

2.21. Hilfssatz

Die folgenden Aussagen über eine Funktion $f : I \to \mathbb{R}$ sind äquivalent:

1. *f ist konvex.*
2. *Für $x, t, y \in I$ mit $x < t < y$ ist $D(x, t) \le D(x, y)$.*

Beweis: Es ist $f(t) \le s(x, y; t) \iff D(x, t) \le D(x, y)$. ■

2.22. Konvexität und zweite Ableitung

Ist $f : I \to \mathbb{R}$ zweimal differenzierbar, so ist f genau dann auf I konvex, wenn dort $f''(x) \ge 0$ ist.

Ist sogar $f''(x) > 0$ auf ganz I, so ist f strikt konvex.

Beweis: Dass $f''(x) \ge 0$ ist, ist gleichbedeutend damit, dass $f'(x)$ monoton wachsend ist.

1) Sei zunächst f' monoton wachsend und $x < t < y$ (in I). Nach dem Mittelwertsatz gibt es Punkte c, d mit $x < c < t < d < y$, so dass $D(x, t) = f'(c)$ und $D(t, y) = f'(d)$ ist. Daraus folgt die Ungleichung $D(x, t) \le D(t, y)$, also

$$\big(f(t) - f(x)\big) \cdot (y - t) \leq \big(f(y) - f(t)\big) \cdot (t - x).$$

Addiert man auf beiden Seiten den Term $\big(f(t) - f(x)\big) \cdot (t - x)$, so folgt:

$$\big(f(t) - f(x)\big) \cdot (y - x) \leq \big(f(y) - f(x)\big) \cdot (t - x),$$

also $D(x,t) \leq D(x,y)$. Also ist f konvex.

Wächst f' sogar streng monoton, so ist $D(x,t) < D(t,y)$ für alle x,t,y mit $x < t < y$, und es folgt, dass f strikt konvex ist.

2) Jetzt setzen wir die Konvexität voraus. Ist $x < y$, so wählen wir einen beliebigen Zwischenpunkt t mit $x < t < y$. Für alle x' mit $x < x' < t$ ist $D(x,x') \leq D(x,t)$. Lässt man x' gegen x gehen, so strebt $D(x,x')$ gegen $f'(x)$. Also ist $f'(x) \leq D(x,t)$.

Sei nun $t < y' < y$. Wegen der Konvexität ist $D(x,t) \leq D(x,y')$. Wir werden zeigen, dass $D(x,y') \leq D(y',y)$ ist. Lässt man dann y' gegen y gehen, so strebt $D(y',y)$ gegen $f'(y)$, und man erhält die Ungleichung $D(x,t) \leq f'(y)$. Damit folgt, dass f' monoton wächst und $f'' \geq 0$ ist.

Aus der Ungleichung $D(x,y') \leq D(x,y)$, die aus der Konvexität folgt, ergibt sich:

$$\big(f(y') - f(x)\big) \cdot (y - x) \leq \big(f(y) - f(x)\big) \cdot (y' - x).$$

Subtrahiert man auf beiden Seiten den Term $\big(f(y') - f(x)\big) \cdot (y' - x)$, so erhält man

$$\big(f(y') - f(x)\big) \cdot (y - y') \leq \big(f(y) - f(y')\big) \cdot (y' - x),$$

also $D(x,y') \leq D(y',y)$. Damit ist alles gezeigt. ■

Wir haben gelernt: Beschreibt der Graph von f eine Linkskurve, so ist f strikt konvex. In einer Rechtskurve ist f dagegen strikt konkav. Die Punkte, an denen der Graph von einer Rechtskurve in eine Linkskurve übergeht (oder umgekehrt) besitzen einen besonderen Namen.

Definition

Sei $f : I \to \mathbb{R}$ zweimal differenzierbar und x_0 ein innerer Punkt von I. Wir sagen, f hat in x_0 einen ***Wendepunkt***, falls es ein $\varepsilon > 0$ gibt, so dass f in $U_\varepsilon(x_0)$ auf der einen Seite von x_0 strikt konvex und auf der anderen strikt konkav ist.

Leider werden Wendepunkte in der Literatur nicht einheitlich definiert. Wir haben hier eine sehr geometrische Beschreibung gewählt. Ist f links von x_0 strikt konvex und rechts von x_0 strikt konkav, so ist $f''(x) \geq 0$ für $x < x_0$ und $f''(x) \leq 0$ für $x > x_0$. Dann wächst f' monoton für $x < x_0$ und fällt monoton für $x > x_0$. Das bedeutet, dass f' in x_0 ein lokales Minimum besitzt. Geht dagegen f bei x_0 von einer konkaven Phase in eine konvexe Phase über, so besitzt f' dort ein lokales Maximum. In beiden Fällen erhalten wir:

2.23. Notwendiges Kriterium für Wendepunkte

Sei $f : I \to \mathbb{R}$ *zweimal differenzierbar und* $x_0 \in I$ *ein innerer Punkt. Besitzt* f *in* x_0 *einen Wendepunkt, so hat* f' *dort ein lokales Extremum, und es ist insbesondere* $f''(x_0) = 0$.

Die Umkehrung stimmt nicht: Wenn f' in x_0 ein Extremum besitzt, dann kann x_0 z.B. ein Häufungspunkt einer ganzen Folge von Wendepunkten sein, ohne selbst ein Wendepunkt zu sein.

2.24. Hinreichendes Kriterium für Wendepunkte

Sei $f : I \to \mathbb{R}$ *zweimal differenzierbar und in einem inneren Punkt* $x_0 \in I$ *ein drittes Mal differenzierbar. Ist* $f''(x_0) = 0$ *und* $f'''(x_0) \neq 0$*, so besitzt* f *in* x_0 *einen Wendepunkt.*

Beweis: Sei etwa $f'''(x_0) > 0$. Wie im Beweis des 2. hinreichenden Kriteriums für Extremstellen folgt, dass $f''(x) < 0$ für $x < x_0$ und $f''(x) > 0$ für $x > x_0$ ist. Daraus ergeben sich die gewünschten Konvexitätsaussagen. ∎

3.3 Integrale und Stammfunktionen

Sei $f : I = [a, b] \to \mathbb{R}$ eine beschränkte (aber ansonsten beliebige) Funktion.

Unter einer ***Zerlegung*** des Intervalls $I = [a, b]$ verstehen wir eine endliche Menge $\mathfrak{Z} = \{x_0, x_1, \dots, x_n\} \subset I$ mit $a = x_0 < x_1 < \dots < x_n = b$. Für $i = 1, \dots, n$ sei

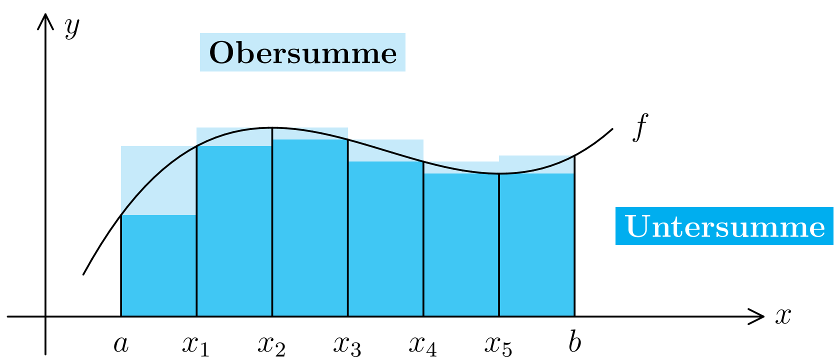
$$\begin{aligned} m_i &= m_i(f, \mathfrak{Z}) &&:= \inf\{f(x) : x_{i-1} \le x \le x_i\} \\ \text{und} \quad M_i &= M_i(f, \mathfrak{Z}) &&:= \sup\{f(x) : x_{i-1} \le x \le x_i\}. \end{aligned}$$

Damit kann man die folgenden Größen definieren:

$$\begin{aligned} \text{Die } \textbf{\textit{Untersumme}}\ U(f, \mathfrak{Z}) &:= \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \\ \text{und die } \textbf{\textit{Obersumme}}\ O(f, \mathfrak{Z}) &:= \sum_{i=1}^{n} M_i(x_i - x_{i-1}). \end{aligned}$$

Eine Zerlegung $\mathfrak{Z}'$ heißt ***feiner*** als die Zerlegung $\mathfrak{Z}$, falls $\mathfrak{Z} \subset \mathfrak{Z}'$ ist. Das bedeutet, dass $\mathfrak{Z}'$ auf jeden Fall die Teilungspunkte von $\mathfrak{Z}$ enthält, eventuell aber noch mehr.

Zu zwei Zerlegungen $\mathfrak{Z}_1, \mathfrak{Z}_2$ gibt es immer eine *gemeinsame Verfeinerung*, die Zerlegung $\mathfrak{Z} = \mathfrak{Z}_1 \cup \mathfrak{Z}_2$.



3.1. Eigenschaften von Ober- und Untersumme

Unter den obigen Voraussetzungen sei noch $m := \inf_I(f)$ und $M := \sup_I(f)$. Dann gilt:

1. $m(b-a) \le U(f, \mathfrak{Z}) \le O(f, \mathfrak{Z}) \le M(b-a)$ *für jede Zerlegung* $\mathfrak{Z}$.
2. *Ist $\mathfrak{Z}'$ eine Verfeinerung von $\mathfrak{Z}$, so ist*
$$U(f, \mathfrak{Z}) \le U(f, \mathfrak{Z}') \le O(f, \mathfrak{Z}') \le O(f, \mathfrak{Z}).$$
3. *Sind $\mathfrak{Z}_1, \mathfrak{Z}_2$ zwei beliebige Zerlegungen von I, so ist $U(f, \mathfrak{Z}_1) \le O(f, \mathfrak{Z}_2)$.*

Beweis: 1) ist klar, denn es ist stets $m \le m_i \le M_i \le M$.

2) Sei $\mathfrak{Z} = \{x_0, x_1, \ldots, x_n\}$. Es reicht, den Fall

$$\mathfrak{Z}' = \{x_0, x_1, \ldots, x_{i-1}, z, x_i, \ldots, x_n\}$$

zu betrachten, mit $x_{i-1} < z < x_i$. Dann sei

$$m_i' := \inf_{[x_{i-1},z]} f \quad \text{und} \quad m_i'' := \inf_{[z,x_i]} f.$$

Offensichtlich ist $m_i \leq \min(m_i', m_i'')$, also $m_i(x_i - x_{i-1}) \leq m_i'(z - x_{i-1}) + m_i''(x_i - z)$. Daraus folgt, dass $U(f, \mathfrak{Z}) \leq U(f, \mathfrak{Z}')$ ist. Eine entsprechende Ungleichung (in der anderen Richtung) erhält man für die Obersumme.

3) Sei $\mathfrak{Z}$ eine gemeinsame Verfeinerung der zwei Zerlegungen $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$. Dann folgt: $U(f, \mathfrak{Z}_1) \leq U(f, \mathfrak{Z}) \leq O(f, \mathfrak{Z}) \leq O(f, \mathfrak{Z}_2)$. ∎

Das sogenannte ***Unterintegral***

$$I_*(f) := \sup\{U(f, \mathfrak{Z}) \,:\, \mathfrak{Z} \text{ Zerlegung von } I\}$$

ist die beste Approximation des Flächeninhaltes von unten, und das ***Oberintegral***

$$I^*(f) := \inf\{O(f, \mathfrak{Z}) \,:\, \mathfrak{Z} \text{ Zerlegung von } I\}$$

ist die beste Approximation des Flächeninhaltes von oben.

Definition

Sei $f : [a, b] \to \mathbb{R}$ eine beschränkte Funktion. Ist $I_*(f) = I^*(f)$, so nennen wir f ***integrierbar*** und den gemeinsamen Wert

$$\boxed{I_{a,b}(f) := I_*(f) = I^*(f).}$$

das ***(bestimmte) Integral*** von f über $[a, b]$.
Sehr oft benutzen wir auch das klassische Integralsymbol

$$\boxed{\int_a^b f(x)\, dx := I_{a,b}(f)\,.}$$

Bemerkung: Man bezeichnet das hier eingeführte Integral auch als das ***Riemann–(Darboux–)Integral***. In Analysis 2 werden wir ein weiteres, allgemeineres Integral (das Integral von Lebesgue) einführen.

Ist $f(x) \geq 0$ auf dem ganzen Intervall, so können wir $I_{a,b}(f)$ als den ***Flächeninhalt*** des Gebietes zwischen der x-Achse und dem Graphen von f auffassen. Nimmt f auch negative Werte an, so ist $I_{a,b}(f)$ die Differenz aus dem Flächenanteil oberhalb der x-Achse und dem Flächenanteil unter der x-Achse. Speziell ist $I_{a,a}(f) = 0$, und wenn $a > b$ ist, setzen wir $I_{a,b}(f) := -I_{b,a}(f)$.

3.2. Integrierbarkeit stetiger Funktionen

Ist $f : [a,b] \to \mathbb{R}$ stetig, so ist $I_*(f) = I^*(f)$, also f integrierbar.

BEWEIS: Wir zeigen: Zu jedem $\varepsilon > 0$ gibt es eine Zerlegung $\mathfrak{Z}_\varepsilon = \{x_0, x_1, \ldots, x_n\}$, so dass für $i = 1, \ldots, n$ gilt:

$$M_i(f, \mathfrak{Z}_\varepsilon) - m_i(f, \mathfrak{Z}_\varepsilon) < \frac{\varepsilon}{b-a}.$$

Dann ist nämlich

$$O(f, \mathfrak{Z}_\varepsilon) - U(f, \mathfrak{Z}_\varepsilon) = \sum_{i=1}^{n} \big(M_i(f, \mathfrak{Z}_\varepsilon) - m_i(f, \mathfrak{Z}_\varepsilon)\big)(x_i - x_{i-1}) < \varepsilon$$

und auch $I^*(f) - I_*(f) \le O(f, \mathfrak{Z}_\varepsilon) - U(f, \mathfrak{Z}_\varepsilon) < \varepsilon$. Weil ε beliebig klein gewählt werden kann, folgt der Satz.

Es sei jetzt ein $\varepsilon > 0$ beliebig vorgegeben. Wie beweisen die Anfangsbehauptung. Als stetige Funktion auf einem abgeschlossenen Intervall ist f sogar gleichmäßig stetig. Daher gibt es ein $\delta > 0$, so dass gilt: Für alle $x, y \in [a, b]$ mit $|x - y| < \delta$ ist $|f(x) - f(y)| < \varepsilon/(b - a)$.

Jetzt sei $\mathfrak{Z}_\varepsilon = \{x_0, x_1, \ldots, x_n\}$ eine Zerlegung, bei der $x_i - x_{i-1} < \delta$ für alle i gilt. Zu jedem i gibt es Punkte $x_{i,u}, x_{i,o} \in [x_{i-1}, x_i]$ mit $f(x_{i,u}) = m_i(f, \mathfrak{Z}_\varepsilon)$ und $f(x_{i,o}) = M_i(f, \mathfrak{Z}_\varepsilon)$. Weil $|x_{i,o} - x_{i,u}| < \delta$ ist, folgt:

$$M_i(f, \mathfrak{Z}_\varepsilon) - m_i(f, \mathfrak{Z}_\varepsilon) = |f(x_{i,o}) - f(x_{i,u})| < \frac{\varepsilon}{b-a}.$$

Damit ist alles bewiesen. ∎

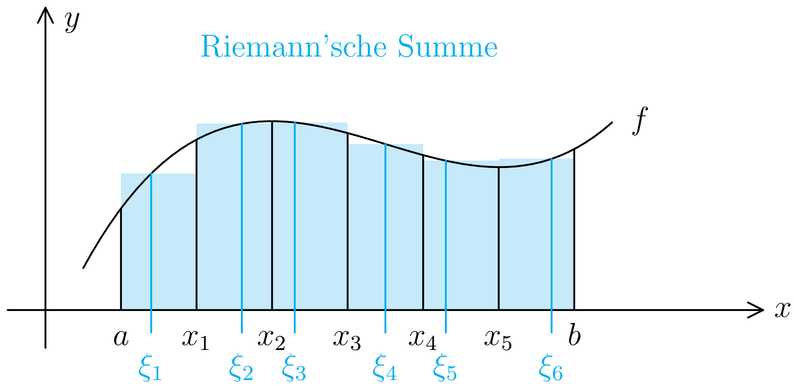
Man kann die Zahl $I_{a,b}(f)$ auch noch auf anderem Wege berechnen.

Definition

Es sei $\mathfrak{Z} = \{x_0, x_1, \ldots, x_n\}$ eine Zerlegung von I, und für jedes i sei ein Punkt $\xi_i \in [x_{i-1}, x_i]$ gewählt. Wir fassen die ξ_i zu einem Vektor $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ zusammen. Dann nennt man

$$\Sigma(f, \mathfrak{Z}, \boldsymbol{\xi}) := \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$$

die ***Riemann'sche Summe*** von f zur Zerlegung $\mathfrak{Z}$ und zur Wahl $\boldsymbol{\xi}$ von Zwischenpunkten.



3.3. Das Integral als Grenzwert Riemann'scher Summen

Sei $f : [a,b] \to \mathbb{R}$ eine integrierbare Funktion. Dann gibt es zu jedem $\varepsilon > 0$ eine Zerlegung $\mathfrak{Z}_0$, so dass für jede feinere Zerlegung $\mathfrak{Z}$ und jede dazu passende Wahl von Zwischenpunkten $\boldsymbol{\xi}$ gilt:

$$\left|\Sigma(f,\mathfrak{Z},\boldsymbol{\xi}) - I_{a,b}(f)\right| < \varepsilon.$$

BEWEIS: Weil $I_*(f) = I^*(f) = I_{a,b}(f)$ ist, gibt es Folgen $(\mathfrak{Z}'_n)$ und $(\mathfrak{Z}''_n)$ von Zerlegungen, so dass die Untersummen $U(f,\mathfrak{Z}'_n)$ monoton wachsend und die Obersummen $O(f,\mathfrak{Z}''_n)$ monoton fallend gegen $I_{a,b}(f)$ konvergieren.

Ist $\varepsilon > 0$ vorgegeben, so gibt es ein n_0, so dass $O(f,\mathfrak{Z}''_n) - U(f,\mathfrak{Z}'_n) < \varepsilon$ für $n \ge n_0$ ist. Sei $\mathfrak{Z}_0$ eine gemeinsame Verfeinerung von $\mathfrak{Z}'_{n_0}$ und $\mathfrak{Z}''_{n_0}$. Dann gilt für jede feinere Zerlegung $\mathfrak{Z}$ und jede dazu passende Wahl von Zwischenpunkten $\boldsymbol{\xi}$:

1. $O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) \le O(f,\mathfrak{Z}''_{n_0}) - U(f,\mathfrak{Z}'_{n_0}) < \varepsilon$.
2. $U(f,\mathfrak{Z}) \le \Sigma(f,\mathfrak{Z},\boldsymbol{\xi}) \le O(f,\mathfrak{Z})$ (weil $m_i(f,\mathfrak{Z}) \le f(\xi_i) \le M_i(f,\mathfrak{Z})$ ist).

Dann ist $-\varepsilon < U(f,\mathfrak{Z}) - O(f,\mathfrak{Z}) \le U(f,\mathfrak{Z}) - I_{a,b}(f) \le \Sigma(f,\mathfrak{Z},\boldsymbol{\xi}) - I_{a,b}$ und $\Sigma(f,\mathfrak{Z},\boldsymbol{\xi}) - I_{a,b} \le O(f,\mathfrak{Z}) - I_{a,b}(f) \le O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) < \varepsilon$. ∎

3.4. Beispiel

Wir betrachten die Funktion $f(x) = x^2$ auf einem Intervall $[a,b]$ mit $0 < a < b$. Als Zerlegung wählen wir $\mathfrak{Z}_n := \{a, a+\delta_n, a+2\delta_n, \ldots, b\}$, mit $\delta_n := (b-a)/n$, als Zwischenpunkte wählen wir $\xi_i^{(n)} := a + i\delta_n$. Dann ist

$$\begin{aligned}
\Sigma(f,\mathfrak{Z}_n,\boldsymbol{\xi}^{(n)}) &= \sum_{i=1}^{n}(a+i\delta_n)^2\cdot\delta_n \\
&= a^2\delta_n\cdot n + 2a\delta_n^2\cdot\sum_{i=1}^{n} i + \delta_n^3\cdot\sum_{i=1}^{n} i^2 \\
&= a^2(b-a) + a(b-a)^2\cdot\frac{n+1}{n} + \frac{1}{6}(b-a)^3\cdot\frac{(n+1)(2n+1)}{n^2}.
\end{aligned}$$

Für $n \to \infty$ strebt dieser Ausdruck gegen

$$\begin{aligned}\int_a^b f(x)\,dx &= a^2b - a^3 + ab^2 - 2a^2b + a^3 + \frac{1}{3}(b^3 - 3b^2a + 3ba^2 - a^3)\\ &= \frac{1}{3}(b^3 - a^3).\end{aligned}$$

Im nächsten Kapitel werden wir lernen, wie ein solches Integral viel leichter berechnet werden kann.

3.5. Elementare Eigenschaften des Integrals

Seien $f, g : [a, b] \to \mathbb{R}$ integrierbar, $\lambda \in \mathbb{R}$:

1. *Ist $c \le f(x) \le C$, so ist $c(b-a) \le I_{a,b}(f) \le C(b-a)$.*
2. *Ist $f \ge 0$, so ist auch $I_{a,b}(f) \ge 0$ (Monotonie des Integrals).*
3. *Ist $a < c < b$, so ist $I_{a,c}(f) + I_{c,b}(f) = I_{a,b}(f)$.*
4. *Es ist $I_{a,b}(\lambda f) = \lambda \cdot I_{a,b}(f)$ und $I_{a,b}(f+g) = I_{a,b}(f) + I_{a,b}(g)$ (Linearität des Integrals).*

Beweis: 1) Sei $\mathfrak{Z}_0 = \{a, b\}$ die „triviale“ Zerlegung von $[a, b]$. Dann ist

$$c(b-a) \le U(f, \mathfrak{Z}_0) \le I_{a,b}(f) \le O(f, \mathfrak{Z}_0) \le C(b-a).$$

2) Ist $f \ge 0$, so ist auch $U(f, \mathfrak{Z}) \ge 0$ für jede Zerlegung $\mathfrak{Z}$, und damit erst recht $I_{a,b}(f) \ge 0$.

3) ist anschaulich völlig klar. Im Detail argumentiert man folgendermaßen:

Zu jeder natürlichen Zahl n kann man Zerlegungen $\mathfrak{Z}_n'$ von $[a, c]$ und $\mathfrak{Z}_n''$ von $[c, b]$ finden, so dass für die zugehörigen Riemannschen Summen Σ_n' und Σ_n'' (mit beliebigen Zwischenpunkten) gilt:

$$|\Sigma_n' - I_{a,c}(f)| < \frac{1}{3n}, \quad |\Sigma_n'' - I_{c,b}(f)| < \frac{1}{3n} \quad \text{und} \quad |(\Sigma_n' + \Sigma_n'') - I_{a,b}(f)| < \frac{1}{3n}.$$

Daraus folgt:

$$\begin{aligned}&|(I_{a,c}(f) + I_{c,b}(f)) - I_{a,b}(f)| = \\ &\quad = |(I_{a,c}(f) - \Sigma_n') + (I_{c,b}(f) - \Sigma_n'') + (\Sigma_n' + \Sigma_n'' - I_{a,b}(f))| \\ &\quad \le |I_{a,c}(f) - \Sigma_n'| + |I_{c,b}(f) - \Sigma_n''| + |(\Sigma_n' + \Sigma_n'') - I_{a,b}(f)| \\ &\quad < 1/3n + 1/3n + 1/3n = 1/n\,.\end{aligned}$$

Weil das für jedes n gilt, ist $I_{a,b}(f) = I_{a,c}(f) + I_{c,b}(f)$.

4) Es ist $\Sigma(\lambda f, \mathfrak{Z}, \boldsymbol{\xi}) = \lambda \cdot \Sigma(f, \mathfrak{Z}, \boldsymbol{\xi})$ und $\Sigma(f+g, \mathfrak{Z}, \boldsymbol{\xi}) = \Sigma(f, \mathfrak{Z}, \boldsymbol{\xi}) + \Sigma(g, \mathfrak{Z}, \boldsymbol{\xi})$. Daraus folgt die Linearität. ∎

Eine Funktion $f : [a, b] \to \mathbb{R}$ heißt ***stückweise stetig***, falls es eine Zerlegung $\mathfrak{Z}_f = \{x_0, \ldots, x_n\}$ von $[a, b]$ gibt, so dass gilt:

1. f ist auf jedem offenen Teilintervall (x_{i-1}, x_i) stetig.
2. Es existieren die einseitigen Grenzwerte $f(a+)$ und $f(b-)$.
3. Für $i = 1, \ldots, n-1$ existieren die einseitigen Grenzwerte $f(x_i-)$ und $f(x_i+)$.

Ist ein $\varepsilon > 0$ vorgegeben, so gibt es eine Zerlegung $\mathfrak{Z}_0$, die feiner als $\mathfrak{Z}_f$ ist, so dass für jede noch feinere Zerlegung $\mathfrak{Z}$ von $[a, b]$ gilt:

Ist $\mathfrak{Z}_i$ die Einschränkung von $\mathfrak{Z}$ auf $[x_{i-1}, x_i]$, so ist $O(f, \mathfrak{Z}_i) - U(f, \mathfrak{Z}_i) < \varepsilon/n$ (denn $f|_{[x_{i-1},x_i]}$ ist ja jeweils integrierbar).

Aber dann ist $O(f, \mathfrak{Z}) - U(f, \mathfrak{Z}) < \varepsilon$. Das bedeutet, dass f integrierbar ist, und natürlich ist

$$\int_a^b f(x)\,dx = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)\,dx.$$

3.6. Der 1. und 2. Mittelwertsatz der Integralrechnung

1. Die Funktion f sei stetig über $[a, b]$. Dann gibt es ein $c \in [a, b]$ mit

$$\int_a^b f(x)\,dx = f(c) \cdot (b - a).$$

2. Sind $f, p : [a, b] \to \mathbb{R}$ stetig, $p \geq 0$, so gibt es ein $c \in [a, b]$ mit

$$\int_a^b f(x)p(x)\,dx = f(c) \cdot \int_a^b p(x)\,dx.$$

BEWEIS: Die erste Aussage folgt sofort aus der zweiten, wenn man $p(x) \equiv 1$ einsetzt. Wir brauchen also nur die zweite Aussage zu beweisen:

Die stetige Funktion f nimmt auf $[a, b]$ ein globales Minimum m und ein globales Maximum M an. Dann ist $m \leq f(x) \leq M$ auf $[a, b]$ und daher

$$m \cdot p(x) \leq f(x)p(x) \leq M \cdot p(x) \text{ auf } [a, b].$$

Wegen der Linearität und der Monotonie des Integrals ist dann auch

$$m \cdot \int_a^b p(x)\,dx \leq \int_a^b f(x)p(x)\,dx \leq M \cdot \int_a^b p(x)\,dx.$$

Durch $F(x) := f(x) \cdot \int_a^b p(t)\,dt$ wird nun eine stetige Funktion $F : [a, b] \to \mathbb{R}$ definiert, die die Werte $m \cdot \int_a^b p(t)\,dt$ und $M \cdot \int_a^b p(t)\,dt$ in $[a, b]$ annimmt. Nach dem Zwischenwertsatz und wegen der obigen Ungleichung muss F dann in einem geeigneten Punkt $c \in [a, b]$ auch den Wert $\int_a^b f(x)p(x)\,dx$ annehmen. Also ist

$$f(c) \cdot \int_a^b p(t)\,dt = F(c) = \int_a^b f(x)p(x)\,dx.$$

■

3.7. Die Standard-Abschätzung

Ist $f : [a, b] \to \mathbb{R}$ stetig, so ist

$$\Big|\int_a^b f(x)\,dx\Big| \le \int_a^b |f(x)|\,dx \le \sup_{[a,b]}|f| \cdot (b - a).$$

BEWEIS: Weil $-|f| \le f \le +|f|$ ist, folgt die erste Ungleichung aus der Monotonie des Integrals. Die zweite Ungleichung ist trivial. ■

Definition

Sei $I = [a, b]$ und $f : I \to \mathbb{R}$ eine beliebige Funktion. Eine Funktion $F : I \to \mathbb{R}$ heißt eine ***Stammfunktion*** von f, falls gilt:

1. F ist stetig.
2. Es gibt eine (leere oder) endliche Teilmenge $M \subset I$, so dass F auf $I \setminus M$ differenzierbar und dort $F' = f$ ist.

3.8. Beispiele

A. $F(x) := \dfrac{1}{n+1}\,x^{n+1}$ ist auf $\mathbb{R}$ eine Stammfunktion von $f(x) = x^n$.

B. $F(x) := -\cos(x)$ ist auf $\mathbb{R}$ eine Stammfunktion von $f(x) = \sin(x)$

C. $F(x) := -\ln(\cos(x))$ ist auf $(-\pi/2, +\pi/2)$ eine Stammfunktion von $f(x) := \tan(x)$. Man überprüft sofort, dass das stimmt. Aber wie kommt man darauf? Das wird das Thema des nächsten Abschnittes sein.

D. Sei $f : [0, 3] \to \mathbb{R}$ definiert durch

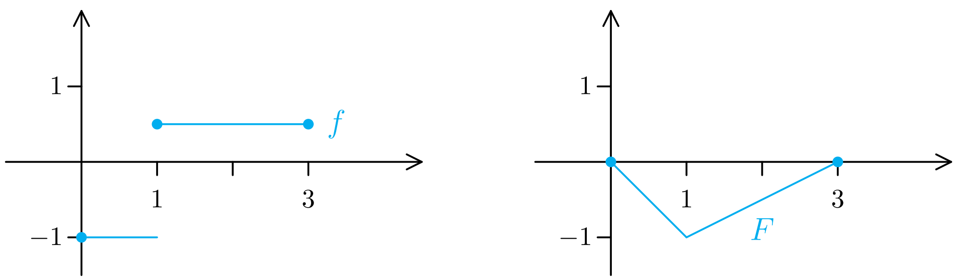
$$f(x) := \begin{cases} -1 & \text{für } 0 \le x < 1, \\ 1/2 & \text{für } 1 \le x \le 3. \end{cases}$$

Auf dem Intervall $[0, 1)$ ist $F_1(x) := -x$ eine Stammfunktion, auf dem Intervall $[1, 3]$ können wir $F_2(x) := x/2$ als Stammfunktion benutzen. Bei $x = 1$

passen die beiden Funktionen nicht stetig aneinander. Als Grenzwert von links erhalten wir -1, als Grenzwert von rechts den Wert $1/2$. Subtrahieren wir die Differenz $3/2$ von F_2, so erhalten wir wieder eine Stammfunktion auf $[1,3]$. Aber jetzt ist die zusammengesetzte Funktion

$$F(x) := \begin{cases} -x & \text{für } 0 \le x < 1, \\ x/2 - 3/2 & \text{für } 1 \le x \le 3, \end{cases}$$

in $x = 1$ stetig, und bis auf den Punkt $x = 1$ ist $F' = f$. Also ist F auf $[0,3]$ eine Stammfunktion von f.



3.9. Die Gesamtheit der Stammfunktionen

Ist F_1 Stammfunktion einer Funktion $f : I \to \mathbb{R}$ und $c \in \mathbb{R}$, so ist auch $F_1 + c$ eine Stammfunktion von f. Ist F_2 eine weitere Stammfunktion von f, so ist $F_1 - F_2$ auf I konstant.

BEWEIS: Die erste Aussage ist trivial. Seien also F_1, F_2 zwei Stammfunktionen von f. Nimmt man aus I die endlich vielen Ausnahmepunkte heraus, in denen F_1 oder F_2 nicht differenzierbar ist, so bleibt eine Vereinigung von (endlich vielen) Intervallen J_ν übrig, so dass $F_1 - F_2$ auf jedem J_ν differenzierbar ist und dort gilt: $(F_1 - F_2)' = F_1' - F_2' = f - f = 0$. Es folgt, dass $F_1 - F_2$ auf jedem J_ν gleich einer Konstanten c_ν ist. Wegen der Stetigkeit von $F_1 - F_2$ müssen alle diese Konstanten übereinstimmen. ■

3.10. Eigenschaften von Stammfunktionen

1. *Besitzen die Funktionen f und g Stammfunktionen F und G, so ist $F + G$ Stammfunktion von $f + g$.*
2. *Ist F Stammfunktion von f und $c \in \mathbb{R}$, so ist $c \cdot F$ Stammfunktion von $c \cdot f$.*
3. *Ist $f \ge 0$ und F Stammfunktion von f, so ist F monoton wachsend.*
4. *Ist F Stammfunktion von f und f in x_0 stetig, so ist F in x_0 differenzierbar.*

BEWEIS: Die ersten beiden Aussagen sind trivial.

Ist $f \geq 0$ auf $I = [a, b]$, so gibt es Punkte $a = t_0 < t_1 < \ldots < t_n = b$, so dass F auf jedem Intervall $I_\nu = (t_{\nu-1}, t_\nu)$ monoton wächst. Da F außerdem stetig ist, muss F auf ganz I monoton wachsen.

Die vierte Aussage zeigt man folgendermaßen: Ist F Stammfunktion von f auf I und $x_0 \in I$, so können wir annehmen, dass es eine Umgebung U von x_0 gibt, so dass F in $I \cap U \setminus \{x_0\}$ differenzierbar und dort $F' = f$ ist. Ist f in x_0 stetig, so existiert der Grenzwert $\lim\limits_{x \to x_0} F'(x) = f(x_0)$.

Sei (x_ν) eine beliebige Folge in I, die gegen x_0 konvergiert. Es reicht zu zeigen, dass die Differenzenquotienten

$$D_F(x_\nu, x_0) := \frac{F(x_\nu) - F(x_0)}{x_\nu - x_0}$$

gegen $f(x_0)$ konvergieren. Dazu sei ein $\varepsilon > 0$ vorgegeben. Es gibt dann ein $\delta > 0$, so dass $|f(x) - f(x_0)| < \varepsilon$ für $|x - x_0| < \delta$ ist. Ist ν so groß, dass $|x_\nu - x_0| < \delta$ ist, sowie c_ν ein Punkt zwischen x_ν und x_0 mit $D_F(x_\nu, x_0) = F'(c_\nu)$ (Mittelwertsatz!), so ist auch $|c_\nu - x_0| < \delta$ und daher

$$|D_F(x_\nu, x_0) - f(x_0)| = |F'(c_\nu) - f(x_0)| = |f(c_\nu) - f(x_0)| < \varepsilon.$$

Es folgt, dass die Differenzenquotienten $D_F(x_\nu, x_0)$ gegen $f(x_0)$ konvergieren. Also ist F in x_0 differenzierbar und $F'(x_0) = f(x_0)$. ∎

Wir wollen jetzt zeigen, dass jede stetige Funktion eine Stammfunktion besitzt.

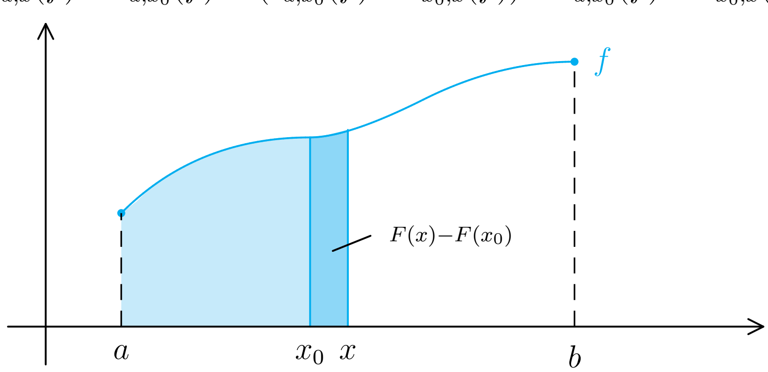
3.11. Hauptsatz der Differential- und Integralrechnung

Sei $f : [a, b] \to \mathbb{R}$ stetig. Für $x \in [a, b]$ sei $F(x) := I_{a,x}(f) = \int_a^x f(t)\,dt$. Dann ist F eine Stammfunktion von f. Insbesondere ist

$$F(a) = 0 \quad \textit{und} \quad F(b) = \int_a^b f(t)\,dt.$$

Beweis: Sei $x_0 \in [a, b]$ und $x > x_0$. Dann ist

$$F(x) - F(x_0) = I_{a,x}(f) - I_{a,x_0}(f) = (I_{a,x_0}(f) + I_{x_0,x}(f)) - I_{a,x_0}(f) = I_{x_0,x}(f).$$



Nun sei $m(f,x) := \inf_{[x_0,x]}(f)$ und $M(f,x) := \sup_{[x_0,x]}(f)$. Es folgt:

$$m(f,x) \cdot (x - x_0) \le I_{x_0,x}(f) \le M(f,x) \cdot (x - x_0),$$

also

$$m(f,x) \le \frac{F(x) - F(x_0)}{x - x_0} \le M(f,x).$$

Für $x \to x_0$ streben $m(f,x)$ und $M(f,x)$ gegen $f(x_0)$. Also ist

$$f(x_0) = \lim_{x \to x_0+} \frac{F(x) - F(x_0)}{x - x_0}.$$

Ist $x < x_0$, so ist $F(x) - F(x_0) = -I_{x,x_0}(f)$ und

$$m(f,x) \cdot (x_0 - x) \le I_{x,x_0}(f) \le M(f,x) \cdot (x_0 - x),$$

also

$$m(f,x) \le \frac{F(x_0) - F(x)}{x_0 - x} \le M(f,x).$$

Daraus folgt:

$$f(x_0) = \lim_{x \to x_0-} \frac{F(x) - F(x_0)}{x - x_0}.$$

Zusammen bedeutet das, dass $F'(x_0) = f(x_0)$ ist. ∎

Wir können den Satz problemlos auf stückweise stetige Funktionen übertragen.

3.12. Stammfunktionen stückweise stetiger Funktionen

Sei $f : [a,b] \to \mathbb{R}$ eine stückweise stetige Funktion. Dann gilt:

1. *f besitzt genau eine Stammfunktion F_0 mit $F_0(a) = 0$.*
2. *Ist F eine beliebige Stammfunktion von f, so ist $\int_a^b f(t)\,dt = F(b) - F(a)$.*

Beweis: 1) Es gibt eine Zerlegung $\mathfrak{Z} = \{x_0, x_1, \ldots, x_n\}$ von $[a,b]$, so dass f auf jedem offenen Intervall $I_k = (x_{k-1}, x_k)$ stetig ist und in x_i jeweils der rechts- und linksseitige Grenzwert existiert. Dann sei $f_k : [x_{k-1}, x_k] \to \mathbb{R}$ definiert durch

$$f_k(x) := \begin{cases} \lim\limits_{x \to x_{k-1}+} f(x) & \text{in } x_{k-1}, \\ f(x) & \text{auf } I_k, \\ \lim\limits_{x \to x_k-} f(x) & \text{in } x_k. \end{cases}$$

Die stetige Funktion f_k besitzt auf $[x_{k-1}, x_k]$ eine Stammfunktion F_k. Bei x_k tritt jeweils zwischen F_k und F_{k+1} ein Sprung der Höhe $C_k := F_{k+1}(x_k) - F_k(x_k)$ auf. Eine auf ganz $[a,b]$ stetige Funktion $\widetilde{F}$ erhalten wir wie folgt:

$$\widetilde{F}(x) := \begin{cases} F_1 & \text{auf } [x_0, x_1], \\ F_2 - C_1 & \text{auf } (x_1, x_2], \\ F_3 - C_1 - C_2 & \text{auf } (x_2, x_3], \\ \dots & \\ F_n - \sum_{\nu=1}^{n-1} C_\nu & \text{auf } (x_{n-1}, x_n]. \end{cases}$$

Offensichtlich ist $\widetilde{F}' = f$ außerhalb der Punkte x_k, also $\widetilde{F}$ eine Stammfunktion von f. Ersetzt man $\widetilde{F}$ durch $F_0 := \widetilde{F} - \widetilde{F}(a)$, so erhält man wieder eine Stammfunktion von f, diesmal aber mit $F_0(a) = 0$.

Sind F_1, F_2 zwei Stammfunktionen von f mit $F_1(a) = F_2(a) = 0$, so ist $F_1 - F_2$ konstant und $F_1(a) - F_2(a) = 0$. Also muss $F_1 = F_2$ sein.

2) Sei F eine beliebige Stammfunktion von f.

Wir setzen zunächst voraus, dass f überall stetig ist. Dann ist $F_0(x) := \int_a^x f(t)\,dt$ die Stammfunktion von f mit $F_0(a) = 0$, und es gibt eine Konstante c, so dass $F(x) = F_0(x) + c$ ist. Daraus folgt:

$$F(b) - F(a) = F_0(b) - F_0(a) = \int_a^b f(t)\,dt.$$

Nun sei f stetig bis auf die Punkte $x_0, \dots, x_n$ (mit $a = x_0 < \dots < x_n = b$). Dann ist f integrierbar und

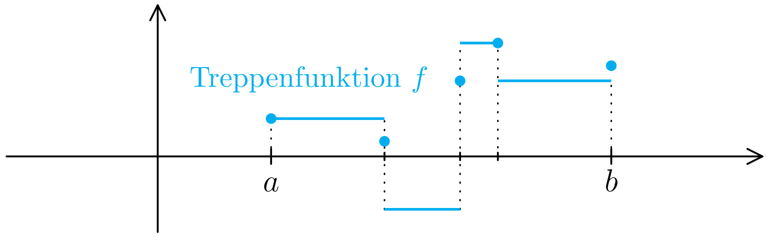
$$\int_a^b f(t)\,dt = \sum_{\nu=1}^{n} \int_{x_{\nu-1}}^{x_\nu} f(t)\,dt = \sum_{\nu=1}^{n} \big(F(x_\nu) - F(x_{\nu-1})\big) = F(b) - F(a).$$

■

3.13. Beispiel

Ein typisches Beispiel für stückweise stetige Funktionen sind die Treppenfunktionen.

Eine Funktion $f : [a, b] \to \mathbb{R}$ heißt ***Treppenfunktion***, falls es eine Zerlegung $\mathfrak{Z} = \{x_0, x_1, \dots, x_n\}$ von $[a, b]$ und Konstante c_ν gibt, so dass $f|_{(x_{\nu-1}, x_\nu)} \equiv c_\nu$ ist, für $\nu = 1, \dots, n$. In den Punkten x_ν kann f ganz beliebige Werte annehmen.



Auf dem Intervall $I_\nu = (x_{\nu-1}, x_\nu)$ ist $F_\nu(x) = c_\nu(x - x_{\nu-1})$ eine Stammfunktion für f mit $F_\nu(x_{\nu-1}) = 0$. Der Sprung bei x_ν von $F_\nu(x_\nu)$ auf $F_{\nu+1}(x_\nu) = 0$ hat jeweils die Höhe $C_\nu = -F_\nu(x_\nu) = -c_\nu(x_\nu - x_{\nu-1})$. Die eindeutig bestimmte Stammfunktion F_0 von f auf I mit $F_0(a) = 0$ hat deshalb in b den Wert $F_0(b) = F_n(x_n) - C_1 - C_2 - \cdots - C_{n-1}$. Also ist

$$\int_a^b f(t)\,dt = F_0(b) = \sum_{\nu=1}^{n} c_\nu(x_\nu - x_{\nu-1}).$$

Das stimmt mit dem anschaulichen Flächeninhalt unter dem Graphen überein.

Es sollen nun weitere Beispiele betrachtet werden. Dabei verwenden wir die folgende Notation: Ist F Stammfunktion von f, so setzt man

$$F(x)\Big|_a^b := F(b) - F(a) = \int_a^b f(x)\,dx.$$

3.14. Beispiele

A. Es soll $\int_{-2}^{+2} |x-1|\,dx$ berechnet werden. Eine Möglichkeit besteht darin, die Nullstellen zu ermitteln und stückweise zu integrieren.

Es gibt nur eine Nullstelle bei $x = 1$. Für $x < 1$ ist $|x-1| = 1-x$, für $x > 1$ ist $|x-1| = x-1$. Also ist

$$\begin{aligned}\int_{-2}^{+2} |x-1|\,dx &= \int_{-2}^{1} (1-x)\,dx + \int_1^2 (x-1)\,dx \\ &= \left(x - \frac{x^2}{2}\right)\Big|_{-2}^{1} + \left(\frac{x^2}{2} - x\right)\Big|_1^2 \\ &= \Big(\frac{1}{2} - (-4)\Big) + \Big(0 - (-\frac{1}{2})\Big) \;=\; 5.\end{aligned}$$

Eine zweite Möglichkeit ist die Verwendung **einer** Stammfunktion über dem ganzen Intervall $[-2,2]$. Die muss dann allerdings stetig sein. Wir benutzen

$$F(x) := \begin{cases} x - x^2/2 & \text{für } -2 \le x \le 1, \\ x^2/2 - x + 1 & \text{für } 1 < x \le 2. \end{cases}$$

Offensichtlich ist $F'(x) = f(x)$ für $x \ne 1$ und beide Zweige der Funktion streben für $x \to 1$ gegen den Wert $1/2$. Dann ist

$$\int_{-2}^{2} |x-1|\,dx = F(2) - F(-2) = 1 - (-4) = 5.$$

Natürlich muss bei beiden Methoden das gleiche Ergebnis herauskommen.

B. Da $\cos'(x) = -\sin(x)$ und $\sin'(x) = \cos(x)$ ist, folgt:

$$\begin{aligned} \int_0^{\pi} \sin(x)\, dx &= -(\cos(\pi) - \cos(0)) \;=\; 2 \\ \text{und} \quad \int_0^{2\pi} \sin(x)\, dx &= -(\cos(2\pi) - \cos(0)) \;=\; 0. \end{aligned}$$

Im zweiten Fall heben sich positive und negative Flächenteile gerade auf.

C. Bekanntlich ist $\arctan'(x) = \dfrac{1}{x^2+1}$, $\arctan(0) = 0$ und $\arctan(1) = \pi/4$. Daraus folgt:

$$\arctan(x) = \int_0^x \frac{1}{1+u^2}\, du \quad \text{und} \quad \frac{\pi}{4} = \int_0^1 \frac{1}{1+u^2}\, du.$$

Das liefert eine interessante Definition für die Zahl π.

Bemerkung: Ein allgemeiner Vertreter der Stammfunktionen von f wird gerne mit dem Leibnizschen Symbol $\int f(x)\, dx$ bezeichnet. Man spricht dann von einem ***unbestimmten Integral***. Ist F eine spezielle Stammfunktion, so schreibt man $\int f(x)\, dx = F(x) + c$, mit einer unbestimmten Konstanten c. Damit soll angedeutet werden, dass die Stammfunktion nur bis auf eine additive Konstante bestimmt ist.

Wir wollen nun noch Integrale komplexwertiger Funktionen betrachten. Dabei werden wir uns auf stetige Funktionen beschränken.

Definition

Sei $f = g + \mathrm{i}\, h : [a, b] \to \mathbb{C}$ stetig. Dann setzen wir

$$\int_a^b f(t)\, dt := \int_a^b g(t)\, dt + \mathrm{i} \int_a^b h(t)\, dt\,.$$

Dann gilt:

1. $\displaystyle\int_a^b \overline{f(t)}\, dt = \overline{\int_a^b f(t)\, dt}.$
2. Ist c eine **komplexe** Zahl, so ist $\displaystyle\int_a^b c \cdot f(t)\, dt = c \cdot \int_a^b f(t)\, dt.$
3. Ist $F : [a, b] \to \mathbb{C}$ eine differenzierbare Funktion mit $F' = f$, so ist

$$\int_a^b f(t)\, dt = F(b) - F(a).$$

Die sehr einfachen Beweise seien dem Leser überlassen.

3.15. Beispiel

Das Rechnen im Komplexen kann manchmal recht bequem sein. Da

$$\left(\frac{1}{c}e^{ct}\right)' = e^{ct}$$

(für komplexes c) ist, folgt z.B.

$$\begin{aligned}
\int_a^b \cos(nt)\,dt + \mathsf{i}\int_a^b \sin(nt)\,dt &= \int_a^b e^{\mathsf{i}\,nt}\,dt = \frac{1}{\mathsf{i}\,n}e^{\mathsf{i}\,nt}\,\Big|_a^b \\
&= -\frac{\mathsf{i}}{n}\big(\cos(nt) + \mathsf{i}\,\sin(nt)\big)\,\Big|_a^b \\
&= \frac{1}{n}\big(\sin(nt) - \mathsf{i}\,\cos(nt)\big)\,\Big|_a^b,
\end{aligned}$$

also

$$\int_a^b \sin(nt)\,dt = \frac{1}{n}\big(\cos(nb) - \cos(na)\big)$$

und

$$\int_a^b \cos(nt)\,dt = \frac{1}{n}\big(\sin(na) - \sin(nb)\big).$$

3.4 Integrationsmethoden

Die ***Regel der partiellen Integration*** (auch ***Produktintegration*** genannt) leitet sich aus der Produktregel der Differentiation her:

$$(f \cdot g)' = f\,'g + f\,g\,'.$$

Das kann man so lesen, dass $f \cdot g$ eine Stammfunktion von $f\,'g + f\,g\,'$ ist. Nun kommt es selten vor, dass der Integrand deutlich sichtbar diese Form besitzt, aber umso häufiger ist er von der Form $f\,'g$. Wenn die Stammfunktion von $f\,g\,'$ leichter als die von $f\,'g$ zu bestimmen ist, hilft die Formel weiter. Wir formulieren das Ergebnis gleich etwas allgemeiner, aber dazu brauchen wir noch den Begriff der „stückweise glatten" Funktion.

Definition

Eine stetige Funktion $f : [a, b] \to \mathbb{R}$ heißt ***stückweise glatt***, falls es eine Zerlegung $a = x_0 < x_1 < \ldots < x_n = b$ und zu jedem ν eine in der Nähe von x_ν definierte Funktion Δ_ν gibt, so dass gilt:

1. $f(x) = f(x_\nu) + (x - x_\nu) \cdot \Delta_\nu(x)$ in der Nähe von x_ν.
2. Es existieren die einseitigen Grenzwerte $\lim\limits_{x \to x_0-} \Delta(x)$ und $\lim\limits_{x \to x_0+} \Delta(x)$

Die Grenzwerte bezeichnet man als ***linksseitige Ableitung*** $f'_-(x_\nu)$ und ***rechtsseitige Ableitung*** $f'_+(x_\nu)$.

Ist f stückweise stetig auf $I = [a, b]$ und F Stammfunktion von f, so ist F stückweise glatt. Zum Beweis beachte man, dass $\Delta_\nu(x)$ für $x \neq x_\nu$ auf jeden Fall mit dem Differenzenquotienten $(F(x) - F(x_\nu))/(x - x_\nu)$ übereinstimmt. Zu jedem $x > x_0$ gibt es dann nach dem Mittelwertsatz ein $c = c(x)$ mit $x_0 < c < x$, so dass $\Delta(x) = F'(c) = f(c)$ ist. Daher existiert

$$\lim_{x \to x_0+} \Delta(x) = \lim_{x \to x_0+} f\big(c(x)\big) = f(x_0+).$$

Analog funktioniert es bei der Annäherung von links.

4.1. Satz von der partiellen Integration

Ist $f : [a, b] \to \mathbb{R}$ eine stückweise glatte Funktion (also Stammfunktion einer stückweise stetigen Funktion f') und g über $[a, b]$ stetig differenzierbar, so ist

$$\int_a^b f'(x)g(x)\,dx = \big(f(x) \cdot g(x)\big)\Big|_a^b - \int_a^b f(x)g'(x)\,dx.$$

Beweis: f ist stetig, und es gibt eine Zerlegung

$$a = x_0 < x_1 < \ldots < x_n = b,$$

so dass f auf jedem offenen Teilintervall $J_k = (x_{k-1}, x_k)$ stetig differenzierbar ist. Dann ist dort $(fg)' = f'\,g + f\,g'$. Außerdem ist fg auf dem ganzen Intervall $[a, b]$ stetig. Also ist fg auf $[a, b]$ Stammfunktion von $f'\,g + f\,g'$. Daraus folgt:

$$\big(f(x) \cdot g(x)\big)\,\Big|_a^b = \int_a^b f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx.$$

∎

Auch in dieser Form ist der Satz sehr nützlich, aber wir werden hier nur Beispiele betrachten, bei denen f und g beide stetig differenzierbar sind.

4.2. Beispiele

A. Es soll das Integral $\displaystyle\int_{-\pi/2}^{\pi/2} x\,\sin x\,dx$ berechnet werden. Hier ist x die Ableitung von $\dfrac{1}{2}\,x^2$ und $\sin x$ die Ableitung von $-\cos x$.

Setzt man $f(x) := x^2/2$ und $g(x) := \sin x$, so ist $f'(x)g(x) = x \sin x$, und $f(x)g'(x) = (x^2/2)\cos x$ deutlich komplizierter als der ursprüngliche Integrand. Setzt man dagegen $f(x) := -\cos x$ und $g(x) := x$, so ist auch $f'(x)g(x) = x \sin x$, aber $f(x)g'(x) = -\cos x$ ist diesmal einfacher. Daher sollte der zweite Weg gewählt werden:

$$\begin{aligned} \int_{-\pi/2}^{\pi/2} x\,\sin x\,dx &= \int_{-\pi/2}^{\pi/2} x\,(-\cos x)'\,dx \\ &= -x\cos x\,\Big|_{-\pi/2}^{\pi/2} - \int_{-\pi/2}^{\pi/2} (-\cos x)\,dx \\ &= (-x\cos x + \sin x)\,\Big|_{-\pi/2}^{\pi/2} = 1 - (-1) = 2. \end{aligned}$$

B. Das folgende Beispiel ist besonders typisch:

$$\begin{aligned} \text{Es ist } \int_a^b e^x \cdot \sin x\,dx &= \int_a^b e^x \cdot (-\cos' x)\,dx \\ &= -(e^x \cdot \cos x)\,\Big|_a^b - \int_a^b (e^x)' \cdot (-\cos x)\,dx \\ &= -(e^x \cdot \cos x)\,\Big|_a^b + \int_a^b e^x \cdot \cos x\,dx. \end{aligned}$$

Es sieht so aus, als wäre alles umsonst gewesen. Aber eine zweite Rechnung liefert

$$\int_a^b e^x \cdot \cos x\,dx = (e^x \cdot \sin x)\,\Big|_a^b - \int_a^b e^x \cdot \sin x\,dx.$$

Jetzt ist das Integral, das wir ausrechnen wollten, wieder aufgetaucht! Trotzdem hilft uns das weiter: Setzt man die rechte Seite in das vorige Ergebnis ein, so erhält man

$$\int_a^b e^x \cdot \sin x\,dx = e^x \cdot \big(\sin x - \cos x\big)\,\Big|_a^b - \int_a^b e^x \cdot \sin x\,dx,$$

also

$$\int_a^b e^x \cdot \sin x\,dx = \frac{e^x}{2} \cdot \big(\sin x - \cos x\big)\,\Big|_a^b .$$

Man hätte natürlich auch auf die Idee kommen können, zu Anfang e^x als die Ableitung von e^x aufzufassen. Das Ergebnis wäre das Gleiche gewesen.

Man kann an diesem Beispiel sehen, wie sich Rechnungen vereinfachen, wenn man komplexwertige Funktionen benutzt: Es ist nämlich

$$\begin{aligned}
&\int_a^b e^x \cos x\,dx + \mathrm{i} \int e^x \sin x\,dx \;= \\
&= \int_a^b e^{(1+\mathrm{i})x}\,dx \;=\; \frac{1}{1+\mathrm{i}} e^{(1+\mathrm{i})x}\,\Big|_a^b = \frac{e^x}{2}(1-\mathrm{i})(\cos x + \mathrm{i}\sin x)\,\Big|_a^b \\
&= \frac{e^x}{2}(\cos x + \sin x)\,\Big|_a^b + \mathrm{i}\,\frac{e^x}{2}(\sin x - \cos x)\,\Big|_a^b .
\end{aligned}$$

Das liefert das obige reelle Integral und ein zweites dazu.

C. Sei $0 < a < b$. Berechnet werden soll das Integral $\displaystyle\int_a^b \ln x\,dx$. Weil der Integrand nicht wie ein Produkt aussieht, überracht es um so mehr, dass hier partielle Integration weiterhelfen soll. Wir benutzen einen Trick: Man kann $\ln x$ als Produkt $(\ln x) \cdot 1$ schreiben und 1 als Ableitung der Funktion x ansehen. Dann gilt:

$$\begin{aligned}
\int_a^b \ln x\,dx &= \int_a^b \ln x \cdot x'\,dx \\
&= (\ln x) \cdot x\,\Big|_a^b - \int_a^b (\ln' x) \cdot x\,dx \quad \text{(Partielle Integration)} \\
&= (x \cdot \ln x)\,\Big|_a^b - x\,\Big|_a^b \qquad \text{(denn es ist } \ln' x \cdot x = 1).
\end{aligned}$$

Also ist $x \cdot \ln x - x$ eine Stammfunktion von $\ln x$.

Eine anspruchsvollere Methode ist die ***Substitutionsregel***, deren Grundlage die Kettenregel ist:

Sei $f : I \to \mathbb{R}$ stetig, $F : I \to \mathbb{R}$ eine Stammfunktion von f. Weiter sei $\varphi : [\alpha, \beta] \to \mathbb{R}$ eine stetig differenzierbare Funktion, mit $\varphi([\alpha, \beta]) \subset I = [a, b]$.

Dann ist auch die Verknüpfung $F \circ \varphi : [\alpha, \beta] \to \mathbb{R}$ stetig differenzierbar, es ist

$$(F \circ \varphi)'(t) = F'(\varphi(t)) \cdot \varphi'(t) = (f \circ \varphi)(t) \cdot \varphi'(t).$$

Also ist $F \circ \varphi$ eine Stammfunktion von $(f \circ \varphi) \cdot \varphi'$.

4.3. Substitutionsregel

Sei $\varphi : [\alpha, \beta] \to \mathbb{R}$ stetig differenzierbar, $I \subset \mathbb{R}$ ein Intervall, $\varphi([\alpha, \beta]) \subset I$ und $f : I \to \mathbb{R}$ stückweise stetig. Dann gilt:

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\,dx = \int_\alpha^\beta f(\varphi(t)) \cdot \varphi'(t)\,dt.$$

Beweis: Wir betrachten eine stetige Funktion f. Der stückweise stetige Fall erfordert nur geringfügige Modifikationen. Ist F eine Stammfunktion von f, so ist $F \circ \varphi$ eine Stammfunktion von $(f \circ \varphi) \cdot \varphi'$. Also ist

$$\begin{aligned}
\int_\alpha^\beta f(\varphi(t)) \cdot \varphi'(t)\,dt &= (F \circ \varphi)(\beta) - (F \circ \varphi)(\alpha) \\
&= F(\varphi(\beta)) - F(\varphi(\alpha)) \\
&= \int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\,dx.
\end{aligned}$$

■

Eine besondere Situation liegt vor, wenn φ **streng monoton wachsend** ist. Dann bildet φ das Intervall $[\alpha, \beta]$ bijektiv auf ein Intervall $[a, b]$ ab, und es ist $\varphi(\alpha) = a$ und $\varphi(\beta) = b$. Man erhält also:

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(\varphi(t)) \cdot \varphi'(t)\,dt = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t)) \cdot \varphi'(t)\,dt.$$

Ist φ **streng monoton fallend**, so ist $\varphi' < 0$, $\varphi(\alpha) = b$, $\varphi(\beta) = a$ und wieder

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx = -\int_\alpha^\beta f(\varphi(t)) \cdot \varphi'(t)\,dt = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t)) \cdot \varphi'(t)\,dt.$$

Will man die Stammfunktionen durch unbestimmte Integrale bezeichnen, so erhält man folgende Formeln:

$$\Big(\int f(x)\,dx\Big) \circ \varphi = \int f(\varphi(t))\varphi'(t)\,dt$$

bzw.

$$\int f(x)\,dx = \Big(\int f(\varphi(t))\varphi'(t)\,dt\Big) \circ \varphi^{-1}, \text{ falls } \varphi \text{ umkehrbar ist.}$$

Auch hier zeigt sich der Nutzen der Regel am besten an Hand von Beispielen. Dabei unterscheiden wir verschiedene Fälle.

Zunächst betrachten wir Beispiele, bei denen der Integrand (deutlich sichtbar) in der Form $f(\varphi(t))\varphi'(t)$ gegeben ist.

4.4. Beispiele

A. Häufig möchte man eine Funktion der Form $x \mapsto f(x+c)$ integrieren. Hier wird in f die Funktion $\varphi(t) := t+c$ eingesetzt. Da $\varphi'(t) \equiv 1$ ist, ist $f(x+c) = f(\varphi(t)) \cdot \varphi'(t)$, also

$$\int_\alpha^\beta f(t+c)\,dt = \int_{\alpha+c}^{\beta+c} f(x)\,dx.$$

Dies ist ein Beispiel für eine streng monoton wachsende Substitution. Die Funktion $\psi(t) := c - t$ ist streng monoton fallend. Ist $\alpha < \beta$, so ist $\psi(\alpha) = c - \alpha > c - \beta = \psi(\beta)$. Mit $\psi'(t) \equiv -1$ erhält man

$$\int_\alpha^\beta f(c-t)\,dt = \int_{c-\beta}^{c-\alpha} f(x)\,dx.$$

B. Es sei $f(x) := 1/x$ und $\varphi(t)$ eine stetig differenzierbare Funktion ohne Nullstellen über $[\alpha, \beta]$. Dann gilt:

$$f(\varphi(t)) \cdot \varphi'(t) = \frac{\varphi'(t)}{\varphi(t)}.$$

Also ist

$$\int_\alpha^\beta \frac{\varphi'(t)}{\varphi(t)}\,dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} \frac{1}{x}\,dx = \big(\ln|x|\big)\,\Big|_{\varphi(\alpha)}^{\varphi(\beta)} = \big(\ln|\varphi(t)|\big)\,\Big|_\alpha^\beta .$$

Das hätte man auch über die logarithmische Ableitung erhalten. Da $\varphi'(t)/\varphi(t) = \big(\ln\circ|\varphi|\big)'(t)$ ist, ist $\ln\circ|\varphi|$ eine Stammfunktion von φ'/φ. Zum Beispiel ist

$$\int_a^b \tan t\,dt = \int_a^b \frac{-\cos' t}{\cos t}\,dt = -\big(\ln|\cos t|\big)\,\Big|_a^b .$$

Schwieriger wird es, wenn der Integrand in der Form $f(x)$ gegeben ist und die Substitution erst gefunden werden muss. Dann ist mehr Kreativität gefordert. Doch auch in der Situation gibt es so etwas wie ein Rezept.

Ein typisches Beispiel ist das Integral $\int x\sqrt{x+1}^3\,dx$. Es hat die Gestalt

$$\int F(x, g(x))\,dx$$

mit $F(x,y) = xy^3$ und dem „störenden" Term $g(x) = \sqrt{x+1}$.

Bei Anwendern ist in diesem Fall die ***„dx-du-Methode"*** besonders beliebt, die explizit mit den Leibniz'schen Differentialen arbeitet. Diese Methode können wir hier nur rein formal benutzen. Um ihr einen inhaltlichen Sinn zu geben, müssten wir in die Theorie der „Differentialformen" einsteigen, und das geht über unsere augenblicklichen Möglichkeiten hinaus. Die Methode funktioniert folgendermaßen:

1. Der störende Term wird mit u bezeichnet: $u := g(x) = \sqrt{x+1}$.

2. Die Gleichung wird nach x aufgelöst: $x = u^2 - 1$.

   Was zu tun ist, wenn dieses Auflösen nicht möglich ist (wenn also g nicht umkehrbar ist), sei erst mal dahingestellt.

3. Nun wird auf beiden Seiten das „Differential" gebildet: $dx = 2u\,du$.

   Dafür gilt die folgende Regel: Ist $h = h(u)$ eine differenzierbare Funktion, so versteht man unter dem Differential dh den formalen Ausdruck $h'(u)\,du$.

4. Jetzt wird im Integral x durch $u^2 - 1$ und dx durch $2u\,du$ ersetzt:

$$\int x\sqrt{x+1}^3\,dx = \int (u^2-1)\cdot u^3 \cdot 2u\,du = 2\int (u^6 - u^4)\,du.$$

5. Es wird – wenn möglich – eine Stammfunktion bestimmt:

$$2\int (u^6 - u^4)\,du = S(u) := \frac{2}{7}u^7 - \frac{2}{5}u^5 + C.$$

6. Zum Schluss muss u wieder durch x ersetzt werden:

$$\begin{aligned}\int x\sqrt{x+1}^3\,dx \;=\; S(u(x)) &= \frac{2}{7}(x+1)^{7/2} - \frac{2}{5}(x+1)^{5/2} + C \\ &= \sqrt{x+1}\Big(\frac{2}{7}(x+1)^3 - \frac{2}{5}(x+1)^2\Big) + C\,.\end{aligned}$$

Jetzt wollen wir sehen, was das mit der Substitutionsregel zu tun hat.

1. Auch bei einer regelgerechten Anwendung der Substitutionsregel sucht man nach einem störenden Term, hier $g(x) := \sqrt{x+1}$.

2. Um den störenden Term $g(x)$ zu beseitigen, führt man eine Substitution ein: $x = \varphi(t)$. Im allgemeinen ist das der Augenblick für Kreativität. $\varphi(t)$ soll so gewählt werden, dass der Integrand einfacher wird, wenn man $g(x)$ durch $g(\varphi(t))$ ersetzt (und mit $\varphi'(t)$ multipliziert). Wenn die Funktion $g(x)$– wie oben bei der dx-du-Methode – umkehrbar ist, kann man einfach $\varphi(t) := g^{-1}(t)$ setzen, hier also $\varphi(t) := t^2 - 1$. Dann ist natürlich $g(\varphi(t)) = t$. Es passiert also exakt das Gleiche wie bei der dx-du-Methode, nur habe ich hier t statt u geschrieben.

3. Nun wird $\varphi'(t)$ berechnet, hier $\varphi'(t) = 2t$.

4. Laut Substitutionsregel ist jetzt

$$\Big(\int F(x, g(x))\,dx\Big) \circ \varphi = \int F\big(\varphi(t), g(\varphi(t))\big) \cdot \varphi'(t)\,dt,$$

hier also

$$\Big(\int x\sqrt{x+1}^3\,dx\Big) \circ \varphi = \int \varphi(t)\sqrt{\varphi(t)+1}^3\varphi'(t)\,dt = \int (t^2-1)t^3 \cdot 2t\,dt$$

und daher (weil φ invertierbar ist)

$$\int x\sqrt{x+1}^3\,dx = \big(\frac{2}{7}t^7 - \frac{2}{5}t^5\big) \circ \varphi^{-1} = \frac{2}{7}(x+1)^{7/2} - \frac{2}{5}(x+1)^{5/2} + C.$$

Tatsächlich sind das die gleichen Vorgänge wie bei der *dx*-*du*-Methode.

Welchen Vorteil bietet nun die *dx*-*du*-Methode? Sie erfordert weniger Intelligenz, weil sie leichter zu merken ist und rein maschinell abläuft. Beim genaueren Hinsehen zeigt sich aber, dass man bei beiden Methoden die gleichen Arbeitsschritte durchführt. Bei der *dx*-*du*-Methode identfiziert man den störenden Term $u = u(x)$ und löst die Gleichung nach x auf (wenn es denn geht). Bei der Anwendung der Substitutionsregel löst man zuerst auf (und schreibt $x = x(u)$ als Substitution $x = \varphi(u)$, das entspricht eher dem modernen Umgang mit Funktionen) und setzt $x(u)$ anschließend ein. Während man bei der *dx*-*du*-Methode das Differential dx durch das Differential $x'(u)\,du$ ersetzt (mit welchem Recht?), wendet man bei der anderen Methode die Substitutionsregel an: Der Ausdruck $f(x)\,dx$ (auf der linken Seite) wird dabei durch $f(\varphi(u))\varphi'(u)\,du$ (auf der rechten Seite) ersetzt.

Problematisch wird es bei der *dx*-*du*-Methode, wenn die Auflösung nicht funktioniert. Und man muss damit leben, dass man sein Tun nicht so recht begründen kann.

Verwendet man bestimmte Integrale, so sieht die korrekte Anwendung der Substitutionsregel in unserem Beispiel folgendermaßen aus:

$$\begin{aligned}
\int_\alpha^\beta x\sqrt{x+1}^3\,dx &= \int_{\varphi^{-1}(\alpha)}^{\varphi^{-1}(\beta)} \varphi(t)\sqrt{\varphi(t)^2+1}^3 \cdot \varphi'(t)\,dt \\
&= \int_{\varphi^{-1}(\alpha)}^{\varphi^{-1}(\beta)} (t^2-1) \cdot t^3 \cdot 2t\,dt \quad (\text{für } \varphi(t) := t^2-1) \\
&= 2\int_{\varphi^{-1}(\alpha)}^{\varphi^{-1}(\beta)} (t^6 - t^4)\,dt.
\end{aligned}$$

4.5. Beispiele

A. Es soll $\displaystyle\int e^{\sqrt{x}}\,dx$ berechnet werden:

Will man unbedingt die dx-du-Methode verwenden, so könnte man dies in einer Nebenrechnung tun:

$$\begin{aligned} u &= \sqrt{x} && \text{(Störender Term)}\\ x &= u^2 && \text{(Auflösen nach } x\text{)}\\ dx &= 2u\,du && \text{(Differential).}\end{aligned}$$

Jetzt ist klar: Als Substitution sollte man $\varphi(u) := u^2$ wählen. Dann ist

$$\int e^{\sqrt{x}}\,dx = \Big(\int 2ue^u\,du\Big)\circ\varphi^{-1} = \big(2(u-1)e^u\big)\circ\varphi^{-1}+C = 2(\sqrt{x}-1)e^{\sqrt{x}}+C\,.$$

Die Stammfunktion von ue^u findet man durch Probieren oder über eine simple Anwendung der partiellen Integration.

Hat man erst einmal verstanden, worum es geht, so kann man natürlich auch direkt erkennen, dass die störende Wurzel durch die Substitution

$$x =: \varphi(t) = t^2$$

beseitigt wird. Dann schreibt man:

$$\begin{aligned}\int e^{\sqrt{x}}\,dx &= \Big(\int e^{\sqrt{\varphi(t)}}\varphi'(t)\,dt\Big)\circ\varphi^{-1} = \Big(\int e^t\cdot 2t\,dt\Big)\circ\varphi^{-1}\\ &= \big(2(t-1)e^t\big)\circ\varphi^{-1}+C = 2(\sqrt{x}-1)\cdot e^{\sqrt{x}}+C.\end{aligned}$$

B. Man kann zeigen, dass man jede rationale Funktion (außerhalb ihrer Polstellen) integrieren kann. Deshalb möchte man Integranden gerne mit Hilfe einer geeigneten Substitution rational machen:

Beim Integral $\displaystyle\int \frac{dx}{\sqrt[3]{x}+\sqrt{x}}$ versagt die dx-du-Methode. Wo ist der störende Term? Es gibt zwei davon! Andererseits sieht man sofort, dass die Substitution $x = \varphi(t) := t^6$ beide Wurzeln simultan beseitigt. Mit $\varphi'(t) = 6t^5$ und $t = \varphi^{-1}(x) = \sqrt[6]{x}$ folgt:

$$\int \frac{dx}{\sqrt[3]{x}+\sqrt{x}} = \Big(\int \frac{6t^5}{t^2+t^3}\,dt\Big)\circ\varphi^{-1} = 6\Big(\int \frac{t^3}{1+t}\,dt\Big)\circ\varphi^{-1}.$$

Polynomdivision ergibt $t^3/(1+t) = t^2-t+1-1/(t+1)$, und diese Funktion kann jeder integrieren.

Manche Integrale sind nicht mit der Standardmethode zu knacken. Betrachten wir zum Beispiel das Integral

$$\int_a^b \sqrt{1-x^2}\,dx, \text{ für } -1 < a < b < +1.$$

Wir suchen nach einer Substitution, durch die der Integrand einfacher wird. Hier versagt die *dx-du*-Methode, denn wenn man etwa $u = \sqrt{1-x^2}$ setzt, dann erhält man $x = \sqrt{1-u^2}$ und nichts ist gewonnen.

Man muss sich etwas anderes einfallen lassen. Die Gleichung $y = \sqrt{1-x^2}$ erinnert an die Gleichung $\cos t = \sqrt{1-\sin^2 t}$, deshalb kann man es ja einmal mit der Substitution $x = \varphi(t) := \sin t$ versuchen. Da die Sinus-Funktion das Intervall $[-\pi/2, +\pi/2]$ bijektiv (streng monoton wachsend) auf das Intervall $[-1, +1]$ abbildet, mit Umkehrfunktion $y = \arcsin x$, erhält man:

$$\begin{aligned} \int_a^b \sqrt{1-x^2}\,dx &= \int_{\arcsin(a)}^{\arcsin(b)} \sqrt{1-\sin^2 t}\cdot \cos t\,dt \\ &= \int_{\arcsin(a)}^{\arcsin(b)} \cos^2 t\,dt. \end{aligned}$$

Tatsächlich hat sich die Situation vereinfacht, das neue Integral kann in der bekannten Weise mit Hilfe partieller Integration berechnet werden. Mit $\alpha := \arcsin(a)$ und $\beta := \arcsin(b)$ erhält man

$$\int_\alpha^\beta \cos^2 t\,dt = \frac{1}{2}\cdot(\cos t\cdot \sin t + t)\,\Big|_\alpha^\beta,$$

also

$$\int_a^b \sqrt{1-x^2}\,dx = \frac{1}{2}\cdot\left(t\cdot\sqrt{1-t^2} + \arcsin t\right)\Big|_a^b .$$

Im Grenzübergang für $a \to -1$ und $b \to +1$ wird damit die Fläche des halben Einheitskreises berechnet:

$$\int_{-1}^{+1} \sqrt{1-x^2}\,dx = \frac{\pi}{2}.$$

Hätte man hier die *dx-du*-Methode verwenden wollen, so hätte man $u = \arcsin x$ setzen müssen. Das liegt nun wirklich nicht auf der Hand, auf die Idee der Substitution $x = \sin t$ kann man eher kommen.

Zum Schluss soll andeutungsweise gezeigt werden, dass jede rationale Funktion „elementar" integrierbar ist.

$$\text{Es sei}\quad f(x) = \frac{P(x)}{Q(x)}, \text{ mit Polynomen } P \text{ und } Q.$$

Indem man notfalls eine Polynomdivision durchführt, kann man erreichen, dass $\text{grad}(P) < \text{grad}(Q)$ ist.

Für den Nenner $Q(x)$ ergibt sich eine Zerlegung

$$Q(x) = (x - c_1)^{k_1} \cdot \ldots \cdot (x - c_r)^{k_r} \cdot q_1(x)^{s_1} \cdot \ldots \cdot q_l(x)^{s_l},$$

mit reellen Zahlen $c_1, \ldots, c_r$ und in $\mathbb{R}$ unzerlegbaren quadratischen Polynomen $q_1(x), \ldots, q_l(x)$.
Dieser Schritt kann natürlich in der Praxis ein unüberwindbares Hindernis darstellen, aber theoretisch existiert die Zerlegung.

Jetzt führt man die (komplexe) Partialbruchzerlegung durch. Ist

$$Q(x) = (x - c_1)^{k_1} \cdots (x - c_r)^{k_r},$$

mit paarweise verschiedenen komplexen Zahlen c_i, so gibt es eine eindeutig bestimmte Darstellung

$$\frac{P(x)}{Q(x)} = \sum_{j=1}^{r} \sum_{k=1}^{k_j} \frac{a_{jk}}{(x - c_j)^k}, \text{ mit } a_{jk} \in \mathbb{C}.$$

Fasst man paarweise auftretende Linearfaktoren $x - c$ und $x - \overline{c}$ zu einem reellen quadratischen Polynomen $q(x)$ zusammen, so erhält man einen Term der Gestalt $h(x)/q(x)$ mit einer (reellen) affin-linearen Funktion $h(x) = Bx + C$.

Es bleibt, Stammfunktionen von Funktionen der Art

$$\frac{A}{(x-c)^k} \quad \text{mit } A, c \in \mathbb{C} \quad \text{und} \quad \frac{Bx + C}{x^2 + ax + b} \quad \text{mit } a, b, B, C \in \mathbb{R}.$$

zu suchen. In einigen Fällen können wir die Stammfunktionen direkt hinschreiben:

a) Ist $c \in \mathbb{R}$, so ist $\displaystyle\int \frac{1}{x - c}\,dx = \ln|x - c| + C$.

b) $\displaystyle\int \frac{1}{t^2 + 1}\,dt = \arctan(t) + C$.

c) $\displaystyle\int \frac{t}{t^2 + 1}\,dt = \frac{1}{2}\ln(t^2 + 1) + C$.

d) $\displaystyle\int \frac{t}{(t^2 + 1)^n}\,dt = \frac{1}{2}\int \frac{\varphi'(t)}{\varphi(t)^n}\,dt = \frac{1}{2}\int \frac{1}{x^n}\,dx$, für $\varphi(t) = t^2 + 1$ (und $n > 1$).

Also ist $\displaystyle\int \frac{t}{(t^2 + 1)^n}\,dt = -\frac{1}{2(n-1)} \cdot \frac{1}{(t^2 + 1)^{n-1}}$.

e) Bei Integralen der Form $\displaystyle\int \frac{h(x)}{x^2+ax+b}\,dx$ mit affin-linearem $h(x)$ müssen wir zunächst eine Substitution vornehmen. Da wir nur den Fall betrachten, dass der Nenner **keine** reelle Nullstelle besitzt, ist $a^2 - 4b < 0$.

Wir setzen
$$x = \varphi(t) := ct - \frac{a}{2}, \qquad \text{mit } c := \frac{1}{2}\sqrt{4b - a^2}.$$
Dann ist $\varphi'(t) = c$ und
$$\varphi(t)^2 + a\varphi(t) + b = c^2t^2 + b - \frac{a^2}{4} = c^2(t^2+1).$$
Also ist
$$\int \frac{h(x)}{x^2+ax+b}\,dx = \int \frac{h(\varphi(t))\cdot\varphi'(t)}{c^2(t^2+1)}\,dt = \frac{1}{c}\int \frac{h(\varphi(t))}{t^2+1}\,dt.$$
Der Zähler $h(\varphi(t))$ ist wieder eine affin-lineare Funktion. Mit den Ergebnissen von (b) und (c) ist dieser Fall dann erledigt.

f) Für $k \geq 2$ ist
$$\int \frac{1}{(x-c)^k}\,dx = \frac{-1}{(k-1)(x-c)^{k-1}} + C, \ .$$

4.6. Beispiel

Sei $f(x) := \dfrac{1}{x^3 - x^2 + x - 1} = \dfrac{1}{(x-1)(x^2+1)} = \dfrac{a}{x-1} + \dfrac{b+cx}{x^2+1}$.

Dann muss $(a+c)x^2 + (b-c)x + a - b = 1$ sein, und das führt zu dem

Gleichungssystem $\quad a + c = 0, \qquad b - c = 0 \quad$ und $\quad a - b = 1$.

Also muss $a = \dfrac{1}{2}$ und $b = c = -\dfrac{1}{2}$ sein, d.h. $f(x) = \dfrac{1}{2}\cdot\left[\dfrac{1}{x-1} - \dfrac{x+1}{x^2+1}\right]$.
Daraus folgt:
$$\begin{aligned}\int f(x)\,dx &= \frac{1}{2}\cdot\left[\int \frac{1}{x-1}\,dx - \int \frac{1}{x^2+1}\,dx - \int \frac{x}{x^2+1}\,dx\right] \\ &= \frac{1}{2}\cdot\left[\ln|x-1| - \arctan(x) - \frac{1}{2}\ln(x^2+1)\right] + C.\end{aligned}$$

3.5 Uneigentliche Integrale

Wir beginnen mit einem Beispiel:

Sei $f(x) := 1/\sqrt{x}$. Dann ist $\lim\limits_{x\to 0+} f(x) = +\infty$, also f über $[0,1]$ nicht integrierbar.

Andererseits ist $F(x) := 2\sqrt{x}$ eine Stammfunktion von $f(x)$, und daher

$$\int_\varepsilon^1 f(x)\,dx = F(1) - F(\varepsilon) = 2(1-\sqrt{\varepsilon}).$$

Nun lassen wir ε gegen Null gehen und erhalten $\lim\limits_{\varepsilon\to 0+} \displaystyle\int_\varepsilon^1 f(x)\,dx = 2$.

Da dieser Grenzwert existiert, wollen wir ihn gerne als Integral von f über $(0,1]$ auffassen. Damit das möglich ist, müssen wir den Integralbegriff erweitern.

Definition

Sei $f : [a,b) \to \mathbb{R}$ eine stückweise stetige Funktion. Der Grenzwert

$$\boxed{\int_a^b f(t)\,dt := \lim_{x\to b-} \int_a^x f(t)\,dt}$$

wird als ***uneigentliches Integral*** bezeichnet. Falls er existiert, nennt man das uneigentliche Integral ***konvergent***, andernfalls ***divergent***.

Analog erklärt man das uneigentliche Integral einer stückweise stetigen Funktion $f : (a,b] \to \mathbb{R}$ durch den rechtsseitigen Limes, wie im obigen Beispiel. Man kann zeigen, dass dieser Begriff auf abgeschlossenen Intervallen nichts Neues bringt.

Ist f eine stückweise stetige Funktion auf einem offenen Intervall (a,b), so bildet man das uneigentliche Integral, indem man einen Punkt $c \in (a,b)$ wählt und die uneigentlichen Integrale von f über $(a,c]$ und über $[c,b)$ bildet und dann addiert. Das Ergebnis hängt nicht von der Wahl des Punktes c ab. Wichtig ist nur, dass man beide Grenzübergänge unabhängig voneinander durchführt!

5.1. Beispiel

Wir betrachten $f(x) := 1/x^\alpha$ auf $(0,b]$ für verschiedene α.

a) Ist $\boxed{\alpha = 1}$, so ist $F(x) := \ln(x)$ eine Stammfunktion für $f(x)$, und daher

$$\int_\varepsilon^b \frac{1}{x}\,dx = \ln(b) - \ln(\varepsilon) \longrightarrow +\infty \text{ für } \varepsilon \to 0.$$

Das uneigentliche Integral divergiert!

b) Ist $\boxed{\alpha \neq 1}$, so ist $F(x) := -\dfrac{1}{(\alpha-1)x^{\alpha-1}}$ Stammfunktion für f.

Wir betrachten zunächst den Fall $\boxed{\alpha < 1}$: dann ist

$$\int_\varepsilon^b \frac{1}{x^\alpha}\,dx = -\frac{1}{\alpha-1}\cdot\left(\frac{1}{b^{\alpha-1}} - \frac{1}{\varepsilon^{\alpha-1}}\right).$$

Da $1-\alpha > 0$ ist, strebt $1/\varepsilon^{\alpha-1} = \varepsilon^{1-\alpha}$ gegen Null für $\varepsilon \to 0$.

Also existiert $\displaystyle\int_0^b \frac{1}{x^\alpha}\,dx = -\frac{1}{(\alpha-1)b^{\alpha-1}}$ für $\alpha < 1$.

Insbesondere ist $\displaystyle\int_0^1 \frac{1}{x^\alpha}\,dx = \frac{1}{1-\alpha}$, z.B. $\displaystyle\int_0^1 \frac{1}{\sqrt{x}}\,dx = 2$.

c) Ist $\boxed{\alpha > 1}$, so ist $\alpha - 1 > 0$, und $1/\varepsilon^{\alpha-1}$ strebt gegen $+\infty$ für $\varepsilon \to 0$. In diesem Fall divergiert das uneigentliche Integral.

Das bedeutet z.B., dass $\displaystyle\int_0^1 \frac{1}{x^2}\,dx$ nicht konvergiert.

Bisher haben wir nur über beschränkte offene Intervalle integriert. Jetzt wollen wir das auf ganz $\mathbb{R}$ oder zumindest auf eine Halbachse ausdehnen.

Definition

Sei $f : [a, +\infty) \to \mathbb{R}$ eine stückweise stetige Funktion. Dann wird auch der Grenzwert

$$\int_a^\infty f(t)\,dt := \lim_{x\to\infty} \int_a^x f(t)\,dt$$

als ***uneigentliches Integral*** bezeichnet. Konvergenz und Divergenz des uneigentlichen Integrals erklärt man wie oben, uneigentliche Integrale über $(-\infty, b]$ definiert man analog.

5.2. Beispiel

Wir betrachten noch einmal $f(x) = 1/x^\alpha$ für verschiedene α.

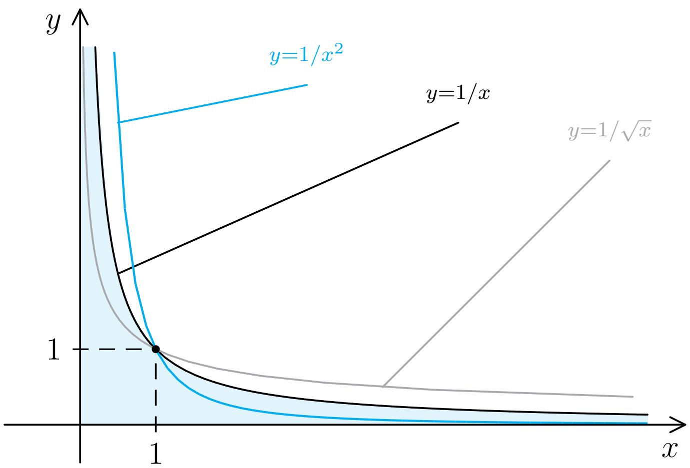
a) $\boxed{\alpha = 1}$: $\displaystyle\int_a^x \frac{1}{t}\,dt = \ln(x) - \ln(a)$ strebt für $x \to +\infty$ gegen $+\infty$.

Das uneigentliche Integral divergiert also auch hier.

b) Ist $\boxed{\alpha < 1}$, so ist $\displaystyle\int_a^x \frac{1}{t^\alpha}\,dt = -\frac{1}{\alpha-1}\cdot\left(\frac{1}{x^{\alpha-1}} - \frac{1}{a^{\alpha-1}}\right)$, wobei $\displaystyle\frac{1}{x^{\alpha-1}} = x^{1-\alpha}$ gegen $+\infty$ strebt, für $x \to \infty$. Auch dieses Integral divergiert.

c) Ist $\boxed{\alpha > 1}$, so konvergiert das uneigentliche Integral offensichtlich. Das bedeutet, dass insbesondere das Integral $\int_1^\infty \frac{1}{x^2}\,dx$ konvergiert, während $\int_1^\infty \frac{1}{\sqrt{x}}\,dx$ divergiert. So ist z.B. $\int_1^\infty \frac{1}{x^2}\,dx = 1$.

Zur besseren Visualisierung folgt noch eine Illustration. Jede Funktion der Form $1/x^\alpha$, deren Graph über $(0,1)$ (bzw. über $(1,\infty)$) ganz innerhalb der hellblau gefärbten Fläche verläuft, ist integrierbar.



Insbesondere ergibt sich, dass

$$\int_0^\infty \frac{1}{x^\alpha}\,dx$$

für **kein** α konvergiert!

Uneigentliche Integrale über ganz $\mathbb{R}$ werden durch zwei voneinander unabhängige Grenzprozesse berechnet:

Definition

Sei $f : \mathbb{R} \to \mathbb{R}$ eine stückweise stetige Funktion. Das uneigentliche Integral $\int_{-\infty}^{+\infty} f(t)\,dt$ konvergiert genau dann, wenn die beiden uneigentlichen Integrale

$$\int_{-\infty}^{0} f(t)\,dt \text{ und } \int_{0}^{+\infty} f(t)\,dt$$

konvergieren. Der Wert des Integrals über ganz $\mathbb{R}$ ist gleich der Summe der Werte der Teilintegrale, d.h. es ist

$$\int_{-\infty}^{+\infty} f(t)\,dt = \lim_{\substack{a\to-\infty \\ b\to+\infty}} \int_a^b f(t)\,dt.$$

Man beachte, dass in der letzten Formel a und b unabhängig voneinander gegen die Grenzen streben. Manchmal findet man auch noch den folgenden Grenzwert:

$$HW \int_{-\infty}^{+\infty} f(t)\,dt := \lim_{r\to\infty} \int_{-r}^{+r} f(t)\,dt.$$

Man spricht vom ***Cauchy'schen Hauptwert***. Es kann passieren, dass dieser Hauptwert existiert, obwohl das uneigentliche Integral von f über $\mathbb{R}$ divergiert.

Wenn keine explizite Stammfunktion gegeben ist, wird es schwierig mit dem Nachweis von Konvergenz oder Divergenz eines uneigentlichen Integrals. Für den Fall gibt es aber gewisse Vergleichskriterien. Wir betrachten nur einen Typ von uneigentlichen Integralen, für die anderen Typen gelten analoge Aussagen.

Definition

[absolute Konvergenz] Das uneigentliche Integral $\displaystyle\int_a^\infty f(x)\,dx$ ***konvergiert absolut***, falls $\displaystyle\int_a^\infty |f(x)|\,dx$ konvergiert. Für andere Typen von uneigentlichen Integralen definiert man die absolute Konvergenz entsprechend.

5.3. Gewöhnliche und absolute Konvergenz

Sei f stückweise stetig über $[a,\infty)$. Konvergiert das uneigentliche Integral $\int_a^\infty f(x)\,dx$ absolut, so auch im gewöhnlichen Sinne, und es ist

$$\Big|\int_a^\infty f(x)\,dx\,\Big| \le \int_a^\infty |f(x)|\,dx\,.$$

Beweis: Für $x \in [a,\infty)$ ist $0 \le |f(x)| - f(x) \le 2|f(x)|$. Ist $A_n \ge a$ eine monoton wachsende Folge, die gegen $+\infty$ konvergiert, so ist

$$0 \le \int_a^{A_n} \big(|f(x)| - f(x)\big)\,dx \le 2\int_a^{A_n} |f(x)|\,dx.$$

Das Integral $Y_n := \int_a^{A_n}|f(x)|\,dx$ konvergiert monoton wachsend gegen $Y := \int_a^\infty|f(x)|\,dx$. Das Integral $X_n := \int_a^{A_n}\big(|f(x)|-f(x)\big)\,dx$ in der Mitte wächst ebenfalls mit n monoton (da der Integrand nicht negativ wird) und ist durch $2Y$ nach oben beschränkt, konvergiert also nach dem Satz von der monotonen Konvergenz (auch wenn wir den Grenzwert i.a. nicht kennen). Die Differenz $\int_a^{A_n} f(x)\,dx = Y_n - X_n$ konvergiert dann ebenfalls für $n \to \infty$. Das zeigt die Konvergenz des uneigentlichen Integrals $\int_a^\infty f(x)\,dx$ und damit auch die Konvergenz von $|\int_a^{A_n} f(x)\,dx|$ gegen $|\int_a^\infty f(x)\,dx|$.

Die Ungleichungen $|\int_a^{A_n} f(x)\,dx| \le \int_a^{A_n}|f(x)|\,dx \le \int_a^\infty|f(x)|\,dx$ bleiben beim Grenzübergang $n \to \infty$ erhalten. ∎

Die Umkehrung des Satzes ist falsch, dafür werden wir weiter unten ein Beispiel angeben.

5.4. Majorantenkriterium für uneigentliche Integrale

Es seien f und g zwei stückweise stetige Funktionen über einem Intervall I, mit $|f| \leq g$. Konvergiert das uneigentliche Integral über g, so konvergiert das uneigentliche Integral über f absolut.

BEWEIS: Wieder sei $A_n \geq a$ eine monoton wachsende Folge, die gegen $+\infty$ konvergiert. Dann ist

$$0 \leq \int_a^{A_n} |f(x)| \, dx \leq \int_a^{A_n} g(x) \, dx.$$

Das Integral auf der rechten Seite konvergiert monoton wachsend (da der Integrand nicht negativ ist) gegen $\int_a^\infty g(x)\,dx$. Das Integral in der Mitte wächst ebenfalls mit n monoton und ist nach oben beschränkt, konvergiert also nach dem Satz von der monotonen Konvergenz. Das zeigt die absolute Konvergenz des uneigentlichen Integrals über f. ∎

Ähnlich wie bei den Reihen nennt man hier die Funktion g eine ***Majorante*** für f.

5.5. Folgerung

Sei $f : [a, \infty) \to \mathbb{R}$ stetig

1. *Gibt es ein $C > 0$, ein $\alpha > 1$ und ein $b \geq a$, so dass $|f(t)|t^\alpha \leq C$ für $t \geq b$ ist, so konvergiert das uneigentliche Integral $\int_a^\infty f(t)\,dt$ absolut.*
2. *Gibt es ein $C > 0$ und ein $b \geq a$, so dass $f(t) \cdot t \geq C$ für $t \geq b$ ist, so divergiert das uneigentliche Integral.*

BEWEIS: Im ersten Fall ist $C \cdot t^{-\alpha}$ eine geeignete Majorante, im zweiten Fall ist $C \cdot t^{-1}$ eine Minorante. ∎

5.6. Beispiele

A. $\displaystyle\int_1^\infty \frac{\cos x}{x^2}\,dx$ konvergiert, weil $\displaystyle\int_1^\infty \frac{1}{x^2}\,dx$ konvergiert.

B. Wir zeigen die Konvergenz des „Fehlerintegrals“ $\int_{-\infty}^{+\infty} e^{-x^2}\,dx$. Es ist

$$\int_1^r e^{-x}\,dx = -\int_1^r (e^{-x})'\,dx = -(e^{-r} - e^{-1}),$$

und dieser Ausdruck konvergiert gegen $1/e$ für $r \to \infty$. Also konvergiert das uneigentliche Integral $\int_1^\infty e^{-x}\,dx$.

Analog ist

$$\int_{-\infty}^{-1} e^x\,dx = \lim_{r\to\infty}\int_{-r}^{-1}(e^x)'\,dx = \lim_{r\to\infty}(e^{-1}-e^{-r}) = e^{-1}.$$

Für $x \geq 1$ ist $x^2 \geq x$, also $e^{-x^2} \leq e^{-x}$.

Für $x \leq -1$ ist $x^2 = |x|^2 \geq |x| = -x$, also $-x^2 \leq x$. Damit ist dort $e^{-x^2} \leq e^x$. Setzt man alles zusammen, so folgt die Konvergenz des Fehlerintegrals.

C. Die (überall stetige) Funktion $f(x) := (\sin x)/x$ ist über $[0,\infty)$ nicht absolut integrierbar, denn es ist

$$\begin{aligned}\int_0^{k\pi}\Big|\frac{\sin x}{x}\Big|\,dx &= \sum_{\nu=1}^{k}\int_{(\nu-1)\pi}^{\nu\pi}\Big|\frac{\sin x}{x}\Big|\,dx\\ &\geq \sum_{\nu=1}^{k}\frac{1}{\nu\pi}\int_{(\nu-1)\pi}^{\nu\pi}|\sin x|\,dx = \frac{2}{\pi}\sum_{\nu=1}^{k}\frac{1}{\nu},\end{aligned}$$

und dieser Ausdruck strebt für $k\to\infty$ gegen Unendlich.

Es ist also nicht so ohne Weiteres möglich, die etwaige Konvergenz von $\int_0^\infty(\sin x/x)\,dx$ mit Hilfe des Majorantenkriteriums zu zeigen. Man muss sich etwas Trickreicheres einfallen lassen.

Wir können die Integration bei 1 beginnen und setzen

$$F(x) := \int_1^x \sin t\,dt.$$

Dann ist F differenzierbar, $F(1)=0$, $F'(x)=\sin x$ und

$$|F(x)| = |\cos(1)-\cos(x)| \leq 2.$$

Wir benutzen nun partielle Integration:

$$\begin{aligned}\int_1^x\frac{\sin t}{t}\,dt &= \int_1^x\frac{1}{t}\cdot F'(t)\,dt = \frac{F(t)}{t}\Big|_1^x - \int_1^x\Big(-\frac{F(t)}{t^2}\Big)\,dt\\ &= \frac{F(x)}{x}+\int_1^x\frac{F(t)}{t^2}\,dt\,.\end{aligned}$$

Da F beschränkt ist und $1/x$ für $x\to\infty$ gegen Null konvergiert, brauchen wir nur den zweiten Term zu betrachten. Es ist $|F(t)/t^2| \leq 2/t^2$, und

$$\int_1^x\frac{2}{t^2}\,dt = -\frac{2}{t}\Big|_1^x = 2-\frac{2}{x}$$

strebt für $x\to\infty$ gegen 2. Also konvergiert $\int_1^\infty F(t)/t^2\,dt$ und damit auch das Integral $\int_1^\infty(\sin t)/t\,dt$.

Zum Schluss wollen wir noch den engen Zusammenhang zwischen Reihen und uneigentlichen Integralen hervorheben:

5.7. Vergleichssatz

Sei $m \in \mathbb{N}$ und $f : [m, \infty) \to \mathbb{R}$ positiv, stetig und monoton fallend. Dann haben die Reihe $\sum_{k=m}^{\infty} f(k)$ und das uneigentliche Integral $\int_m^\infty f(x)\,dx$ das gleiche Konvergenzverhalten.

Beweis: Auf dem Intervall $[k, k+1]$ ist

$$f(k) \geq f(x) \geq f(k+1),$$

also auch

$$f(k) \geq \int_k^{k+1} f(x)\,dx \geq f(k+1)$$

und damit

$$\sum_{k=m}^{N} f(k) \geq \int_m^{N+1} f(x)\,dx \geq \sum_{k=m+1}^{N+1} f(k).$$

Daraus folgt die Behauptung. ∎

5.8. Beispiel

Aus dem Vergleichssatz und unseren Kenntnissen über uneigentliche Integrale folgt sofort:

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} \text{ konvergiert genau dann, wenn } \alpha > 1 \text{ ist.}$$

So ist etwa $\sum_{n=1}^{\infty} \frac{1}{n}$ divergent und $\sum_{n=1}^{\infty} \frac{1}{n\sqrt{n}} = \sum_{n=1}^{\infty} \frac{1}{n^{1.5}}$ konvergent.

Die Theorie der uneigentlichen Integrale liefert neue interessante Funktionen, z.B. die Gammafunktion:

Die ***Gammafunktion*** $\Gamma : (0, \infty) \to \mathbb{R}$ wird definiert durch

$$\boxed{\Gamma(x) := \int_0^\infty e^{-t} t^{x-1}\,dt.}$$

Dabei ist $t^{x-1} = e^{(x-1)\ln t}$.

Das uneigentliche Integral $\Gamma(1) = \int_0^\infty e^{-t}\,dt = -(e^{-t})\Big|_0^\infty = 1$ konvergiert offensichtlich.

Nun sei $x > 0$ und $x \neq 1$.

1) Wir untersuchen zunächst die Konvergenz des uneigentlichen Integrals auf $(0, 1]$. Dort ist $e^t > 1$, also $|e^{-t}t^{x-1}| < t^{x-1}$.

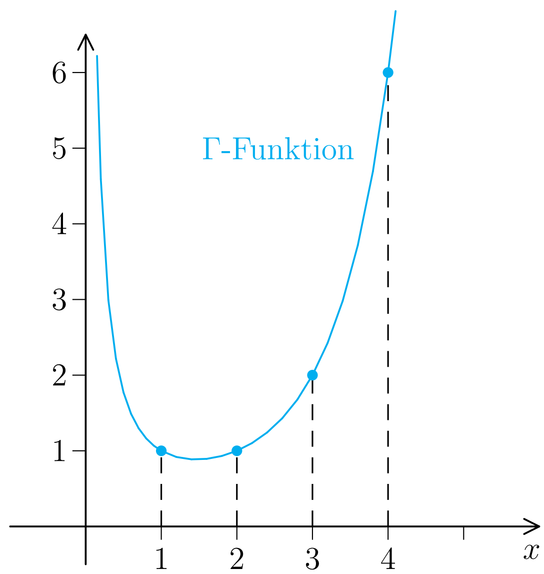
Ist $x > 1$, also $x - 1 > 0$, so ist t^{x-1} über $[0, 1]$ integrierbar, und es gibt keine Probleme. Für $0 < x < 1$ ist aber auch $0 < 1 - x < 1$, und das uneigentliche Integral über $t^{x-1} = 1/t^{1-x}$ konvergiert. Auch in diesem Fall ist $\Gamma(x)$ definiert.

2) Nun zeigen wir die Konvergenz des Integrals auf $[1, \infty)$. Da die Exponentialfunktion stärker als jede Potenz wächst, strebt $t^2 \cdot (t^{x-1}e^{-t}) = e^{-t}t^{x+1}$ für $t \to \infty$ gegen Null. Also gibt es eine Zahl $C > 0$ und ein t_0, so dass $e^{-t}t^{x+1} \le C$ für $t \ge t_0$ ist. Damit ist $e^{-t}t^{x-1} \le C/t^2$, und es folgt die Konvergenz des Integrals.

Es gilt:

a) $\Gamma(1) = 1$.

b) $\Gamma(x + 1) = x \cdot \Gamma(x)$.



Zum Beweis von (b):

Es ist $\Gamma(x+1) = \displaystyle\int_0^\infty e^{-t}t^x\,dt$. Mit partieller Integration erhält man:

$$\int_\varepsilon^r e^{-t}t^x\,dt = -e^{-t}t^x \Big|_\varepsilon^r + x\int_\varepsilon^r e^{-t}t^{x-1}\,dt.$$

Der erste Summand strebt gegen 0, der zweite gegen $x \cdot \Gamma(x)$, für $r \to \infty$ und $\varepsilon \to 0$.

Insbesondere folgt nun:

$$\begin{aligned}
\Gamma(2) &= \Gamma(1+1) = 1 \cdot \Gamma(1) = 1, \\
\Gamma(3) &= \Gamma(2+1) = 2 \cdot \Gamma(2) = 2 \\
\Gamma(4) &= \Gamma(3+1) = 3 \cdot \Gamma(3) = 2 \cdot 3
\end{aligned}$$

und allgemein $\Gamma(n+1) = n \cdot \Gamma(n) = n!$ für $n \in \mathbb{N}$. Die Gammafunktion interpoliert also die Fakultäten!

3.6 Gleichmäßige Konvergenz

Erinnerung: Es sei $M \subset \mathbb{R}$ eine beliebige Teilmenge. Eine Folge von Funktionen $f_n : M \to \mathbb{R}$ ***konvergiert punktweise*** gegen eine Funktion $f : M \to \mathbb{R}$, falls gilt:

$$\text{Für jeden Punkt } x \in M \text{ ist } \lim_{n\to\infty} f_n(x) = f(x)\,.$$

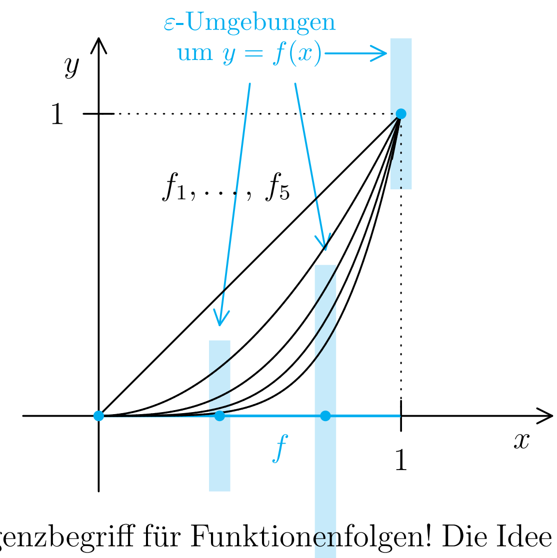
Das Verhalten der Funktionenfolge wird also in jedem einzelnen Punkt $x \in M$ gesondert untersucht. Das globale Verhalten der beteiligten Funktionen spielt dabei keine Rolle.

6.1. Beispiel

Sei $I := [0,1] \subset \mathbb{R}$ und $f_n : I \to \mathbb{R}$ definiert durch $f_n(x) := x^n$. Dann konvergiert diese Funktionenfolge punktweise gegen die Funktion

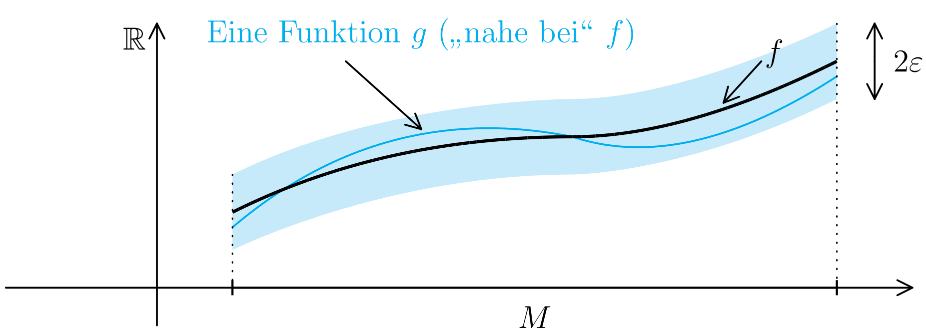
$$f(x) := \begin{cases} 0 & \text{für } 0 \le x < 1\,, \\ 1 & \text{für } x = 1\,. \end{cases}$$

Obwohl alle Funktionen f_n stetig sind, ist die Grenzfunktion f nicht stetig. Das ist ein wenig wünschenswertes Verhalten.



Wir brauchen also einen besseren Konvergenzbegriff für Funktionenfolgen! Die Idee ist, statt zu jedem einzelnen $y = f(x)$ eine eigene ε-Umgebung zu vorzugeben, einen „ε–Schlauch“ um den Graphen von f zu legen. Definieren wir also erst mal einen solchen ε–***Schlauch*** um eine Funktion $f : M \to \mathbb{R}$:

$$\boxed{S_\varepsilon(f) := \{(x,y) \in M \times \mathbb{R} \,:\, f(x) - \varepsilon < y < f(x) + \varepsilon\}.}$$



Eine Funktion $g : M \to \mathbb{R}$ liegt in dem ε-Schlauch, wenn der Graph von g zwischen den Graphen von $f - \varepsilon$ und $f + \varepsilon$ liegt.

Definition

Eine Folge von Funktionen $f_n : M \to \mathbb{R}$ ***konvergiert gleichmäßig*** auf M gegen eine Funktion $f : M \to \mathbb{R}$, wenn es zu jedem $\varepsilon > 0$ ein n_0 gibt, so dass für $n \geq n_0$ und **alle** $x \in M$ gilt:

$$|f_n(x) - f(x)| < \varepsilon.$$

(f_n) konvergiert also genau dann gleichmäßig gegen f, wenn in jedem ε–Schlauch um f fast alle Graphen G_{f_n} liegen.

Dass wir nun den richtigen Konvergenzbegriff gefunden haben, zeigt sich gleich:

6.2. Der gleichmäßige Limes stetiger Funktionen ist stetig

(f_n) konvergiere auf M gleichmäßig gegen f, alle f_n seien stetig. Dann ist auch die Grenzfunktion f auf M stetig.

BEWEIS: Stetigkeit ist eine lokale Eigenschaft, die nur in der Nähe eines (beliebigen) Punktes $\mathbf{x}_0 \in M$ gezeigt werden muss.

Sei $\mathbf{x}_k \in M$ eine Folge von Punkten mit $\lim\limits_{k\to\infty} \mathbf{x}_k = \mathbf{x}_0$. Wir müssen zeigen, dass dann auch $\lim_{k\to\infty} f(\mathbf{x}_k) = f(\mathbf{x}_0)$ ist.

Sei $\varepsilon > 0$. Da (f_n) gleichmäßig auf M gegen f konvergiert, gibt es ein n_0, so dass

$$|f_n(\mathbf{x}) - f(\mathbf{x})| < \frac{\varepsilon}{3} \text{ für } n \geq n_0 \text{ und alle } \mathbf{x} \in M$$

ist. Wir wählen **ein** solches $n \geq n_0$. Da f_n stetig ist, gibt es ein k_0, so dass

$$|f_n(\mathbf{x}_k) - f_n(\mathbf{x}_0)| < \frac{\varepsilon}{3} \text{ für } k \geq k_0$$

ist. Für solche k gilt dann:

$$\begin{aligned} |f(\mathbf{x}_k) - f(\mathbf{x}_0)| &\leq |f(\mathbf{x}_k) - f_n(\mathbf{x}_k)| + |f_n(\mathbf{x}_k) - f_n(\mathbf{x}_0)| + |f_n(\mathbf{x}_0) - f(\mathbf{x}_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Das zeigt, dass $(f(\mathbf{x}_k))$ gegen $f(\mathbf{x}_0)$ konvergiert. ■

6.3. Beispiele

A. Die Folge $f_n(x) := x^n$ kann auf $[0,1]$ nicht gleichmäßig konvergent sein, weil die Grenzfunktion nicht stetig ist. Aber wie steht es mit der Konvergenz auf $I_r := [0, r]$, mit $0 < r < 1$?

Die Folge r^n konvergiert gegen Null, und für $0 \leq x \leq r$ ist $0 \leq x^n \leq r^n$. Zu gegebenem $\varepsilon > 0$ gibt es daher ein n_0, so dass für $n \geq n_0$ gilt:

$$|f_n(x) - 0| = x^n \leq r^n \leq r^{n_0} < \varepsilon.$$

Also konvergiert (f_n) auf dem kleineren Intervall $[0, r]$ gleichmäßig gegen die Nullfunktion.

B. Sei $f_n(x) := \sin(nx)/n$ auf $\mathbb{R}$.

Ist $\varepsilon > 0$ und $n_0 > 1/\varepsilon$, so gilt für $n \geq n_0$:

$$|f_n(x) - 0| = \frac{1}{n}|\sin(nx)| \leq \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon.$$

Also konvergiert (f_n) auf ganz $\mathbb{R}$ gleichmäßig gegen Null.

6.4. Die Konvergenz von $f_n \pm g_n$ und $c \cdot f_n$

Wenn (f_n) auf M gleichmäßig gegen f und (g_n) auf M gleichmäßig gegen g konvergiert, dann konvergiert auch $f_n \pm g_n$ gleichmäßig auf M gegen $f \pm g$ und $(c \cdot f_n)$ gleichmäßig auf M gegen $c \cdot f$.

Auf den Beweis verzichten wir hier, es würden nur Argumente aus den Beweisen der Grenzwertsätze wiederholt.

6.5. Normale Konvergenz $\Longrightarrow$ gleichmäßige Konvergenz

Sei $M \subset \mathbb{R}^n$ beliebig. Die Folge der Partialsummen einer normal konvergenten Reihe von Funktionen auf M ist gleichmäßig konvergent.

Beweis: Zunächst folgt aus der normalen Konvergenz die punktweise Konvergenz der Funktionenreihe gegen eine Grenzfunktion $f : M \to \mathbb{R}$.

Sei $F_N(\mathbf{x}) := \sum_{n=0}^{N} f_n(\mathbf{x})$ die N–te Partialsumme. Dann gibt es wegen des Cauchykriteriums für die normale Konvergenz zu jedem $\varepsilon > 0$ ein n_0, so dass gilt:

$$|F_m(\mathbf{x}) - F_{n_0}(\mathbf{x})| < \frac{\varepsilon}{3} \text{ für alle } \mathbf{x} \in M \text{ und alle } m > n_0.$$

Sei nun $n > n_0$ und $\mathbf{x} \in M$ beliebig. Weil die Zahlenfolge $F_N(\mathbf{x})$ gegen $f(\mathbf{x})$ konvergiert, gibt es ein $m = m(\mathbf{x}, \varepsilon) > n_0$, so dass $|F_m(\mathbf{x}) - f(\mathbf{x})| < \varepsilon/3$ ist. Daraus folgt:

$$\begin{aligned}
|F_n(\mathbf{x}) - f(\mathbf{x})| &\leq |F_n(\mathbf{x}) - F_m(\mathbf{x})| + |F_m(\mathbf{x}) - f(\mathbf{x})| \\
&\leq |F_n(\mathbf{x}) - F_{n_0}(\mathbf{x})| + |F_m(\mathbf{x}) - F_{n_0}(\mathbf{x})| + |F_m(\mathbf{x}) - f(\mathbf{x})| \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

Also konvergiert (F_N) gleichmäßig gegen f. ∎

6.6. Beispiel

Die Umkehrung dieses Satzes gilt nicht. Dazu betrachten wir die Reihe

$$f(x) = \sum_{n=1}^{\infty} (-1)^n \frac{x^n}{n} \qquad \text{auf } [0,1].$$

Es sei $f_n(x) := (-1)^n x^n/n$ und $F_N(x) := \sum_{n=1}^{N} f_n(x)$. Weil $\|f_n\| = 1/n$ ist und die harmonische Reihe divergiert, ist die Funktionenreihe nicht normal konvergent. Sie konvergiert aber nach dem Leibniz-Kriterium punktweise gegen eine Grenzfunktion $f(x)$.

Nun ist $F_{2N-1}(x) < f(x) < F_{2N}(x)$ für alle $x \in [0,1]$, sowie

$$F_{2N}(x) - F_{2N-1}(x) = \frac{x^{2N}}{2N} \leq \frac{1}{2N}.$$

Die rechte Seite strebt – unabhängig von x – gegen Null. Daraus folgt, dass die Folge der Partialsummen F_N auf $[0,1]$ gleichmäßig gegen f konvergiert.

6.7. Vertauschung von Limes und Integral

Eine Folge von stetigen Funktionen $f_n : [a,b] \to \mathbb{R}$ konvergiere gleichmäßig gegen eine Grenzfunktion $f : [a,b] \to \mathbb{R}$. Dann ist

$$\lim_{n\to\infty} \int_a^b f_n(t)\,dt \;=\; \int_a^b f(t)\,dt.$$

BEWEIS: Natürlich ist f als Grenzwert einer gleichmäßig konvergenten Folge von stetigen Funktionen wieder stetig.

Sei $\varepsilon > 0$ vorgegeben. Dann gibt es ein N, so dass für $n \geq N$ und alle $x \in [a,b]$ gilt:

$$|f(x) - f_n(x)| < \frac{\varepsilon}{b-a}.$$

Also ist

$$\begin{aligned} \Big|\int_a^b f(t)\,dt - \int_a^b f_n(t)\,dt\Big| \;&=\; \Big|\int_a^b (f(t) - f_n(t))\,dt\Big| \\ &\leq\; \int_a^b |f(t) - f_n(t)|\,dt \;<\; \frac{\varepsilon}{b-a}\cdot(b-a) \;=\; \varepsilon. \end{aligned}$$

Das ergibt die gewünschte Formel. ∎

Bemerkung: Die Integrierbarkeit von f folgt aus der Stetigkeit. Setzen wir nur voraus, dass die f_n stückweise stetig sind, so können wir mit unseren Mitteln nicht wie oben schließen! In Wirklichkeit gilt der Satz aber viel allgemeiner: Wenn die f_n höchstens abzählbar viele Unstetigkeitsstellen besitzen und dort jeweils die einseitigen Grenzwerte existieren, dann sind die f_n und die Grenzfunktion f im Riemannschen Sinne integrierbar und es gilt der Satz über die Vertauschbarkeit von Limes und Integral. Im zweiten Semester werden wir im Rahmen der Lebesgue-Theorie einen noch allgemeineren Satz beweisen.

6.8. Folgerung (Integration von Reihen)

*Die Funktionen $f_n : [a,b] \to \mathbb{R}$ seien stetig, und die Reihe $\sum_{n=1}^{\infty} f_n$ sei auf $[a,b]$ **normal** konvergent gegen f. Dann ist*

$$\int_a^b f(x)\,dx = \sum_{n=1}^{\infty} \int_a^b f_n(x)\,dx.$$

6.9. Vertauschung von Limes und Ableitung

*Eine Folge von stetig differenzierbaren Funktionen $f_n : [a,b] \to \mathbb{R}$ sei **punktweise** konvergent gegen eine Funktion $f : [a,b] \to \mathbb{R}$. Außerdem sei f_n' **gleichmäßig** konvergent. Dann ist f stetig differenzierbar und*

$$\lim_{n\to\infty} f_n'(x) = f'(x).$$

Beweis: Sei $I := [a,b]$ und $f^* := \lim_{n\to\infty} f_n'$.

Ist $x_0 \in I$ fest gewählt und $x \in I$ ein weiterer Punkt, so ist

$$\begin{aligned} f(x) - f(x_0) &= \lim_{n\to\infty} (f_n(x) - f_n(x_0)) \\ &= \lim_{n\to\infty} \int_{x_0}^{x} f_n'(t)\,dt \;=\; \int_{x_0}^{x} f^*(t)\,dt. \end{aligned}$$

Also ist $f(x) = f(x_0) + \displaystyle\int_{x_0}^{x} f^*(t)\,dt$ eine differenzierbare Funktion mit Ableitung $f'(x) = f^*(x)$. ∎

6.10. Folgerung (Differentiation von Reihen)

*Die Funktionen $f_n : [a,b] \to \mathbb{R}$ seien stetig differenzierbar, die Reihe $\sum_{n=1}^{\infty} f_n$ sei **punktweise** konvergent gegen eine Funktion $f : [a,b] \to \mathbb{R}$ und die Reihe $\sum_{n=1}^{\infty} f_n'$ sei auf $[a,b]$ **normal** konvergent. Dann ist f stetig differenzierbar und*

$$f'(x) = \sum_{n=1}^{\infty} f_n'(x) \text{ für alle } x \in [a,b].$$

Bemerkung: Alle Definitionen und Sätze lassen sich mühelos auf komplexwertige Funktionen übertragen.

6.11. Beispiel

Sei $f_n : [0, 2\pi] \to \mathbb{R}$ definiert durch

$$f_n(x) := \frac{1}{n} \sin(n^2 x).$$

Dann ist

$$\|f_n - 0\| = \sup\{|f_n(x)| \,:\, x \in [0, 2\pi]\,\} \le \frac{1}{n},$$

also (f_n) gleichmäßig konvergent gegen die Nullfunktion.

$f_n'(x) = n \cos(n^2 x)$ oszilliert mit zunehmendem n immer stärker, und mit wachsender Amplitude. So ist z.B.

$$f_1'(\pi) = -1,\ f_2'(\pi) = 2,\ f_3'(\pi) = -3,\ f_4'(\pi) = 4$$

und allgemein $f_n'(\pi) = n \cdot (-1)^n$. Das bedeutet, dass (f_n') nicht konvergiert. Der Satz über die Vertauschbarkeit von Limes und Ableitung ist hier nicht anwendbar.

3.7 Die Taylorentwicklung

Der Satz über die Differenzierbarkeit von Reihen liefert:

7.1. Satz

Sei $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ eine Potenzreihe mit reellen Koeffizienten, Entwicklungspunkt $a \in \mathbb{R}$ und Konvergenzradius $R > 0$. Dann ist die Grenzfunktion $f(x)$ auf dem Konvergenzintervall $(a-R, a+R)$ stetig differenzierbar, und die Reihe

$$\sum_{n=1}^{\infty} n \cdot c_n(x-a)^{n-1}$$

konvergiert auf jedem abgeschlossenen Teilintervall des Konvergenzintervalls normal (und damit gleichmäßig) gegen die Ableitung $f'(x)$.

Potenzreihen können also gliedweise differenziert werden! Wir wussten das bisher nur für die Ableitung im Entwicklungspunkt.

Beweis: Die Funktionenfolge $F_N(x) := \sum_{n=0}^{N} c_n(x-a)^n$ konvergiert auf dem Konvergenzintervall $(a-R, a+R)$ punktweise und auf abgeschlossenen Teilintervallen sogar normal (und damit gleichmäßig) gegen eine stetige Funktion f. Jede der Funktionen F_N ist stetig differenzierbar, und die Folge der Ableitungen F_N' konvergiert nach dem Satz über das Konvergenzverhalten von Potenzreihen auf jedem abgeschlossenen Teilintervall normal (und damit auch wieder gleichmäßig) gegen eine auf $(a-R, a+R)$ stetige Funktion g Also ist f sogar stetig differenzierbar und $f' = g$. ∎

7.2. Folgerung

Eine (reelle) Potenzreihe $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ ist im Konvergenzintervall beliebig oft differenzierbar, und es gilt:

$$c_n = \frac{f^{(n)}(a)}{n!}, \qquad \text{für alle } n \geq 0.$$

Beweis: Da die Ableitung einer Potenzreihe wieder eine Potenzreihe ist, kann man das Argument von oben wiederholen und erhält die beliebige Differenzierbarkeit. Wir müssen nur noch die Formel beweisen. Offensichtlich ist

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)c_n(x-a)^{n-k}.$$

Für $x = a$ erhält man $f^{(k)}(a) = k(k-1)\cdots(k-k+1) \cdot c_k = k!\, c_k$. ∎

Ist umgekehrt $f : I \to \mathbb{R}$ eine in $a \in I$ beliebig oft differenzierbare Funktion, so kann man in a die ***Taylorreihe***

$$\boxed{Tf(x;a) := \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n}$$

bilden. Allerdings ist i.a. nicht klar, ob f durch seine Taylorreihe dargestellt wird.

7.3. Beispiele

A. Sei $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ eine konvergente Potenzreihe mit Konvergenzradius R. Dann ist die Grenzfunktion f auf dem Konvergenzintervall beliebig oft differenzierbar, und dort ist auch $Tf(x;a) = f(x)$. Das lässt sich z.B. auf $\exp x$, $\sin x$ und $\cos x$ anwenden.

B. Wir betrachten die Potenzreihe $\sum_{n=0}^{\infty}(-1)^n x^{2n}$. Dann ist $a = 0$ und $R = 1$, und da es sich um eine geometrische Reihe handelt, ergibt sich als Grenzwert die Funktion

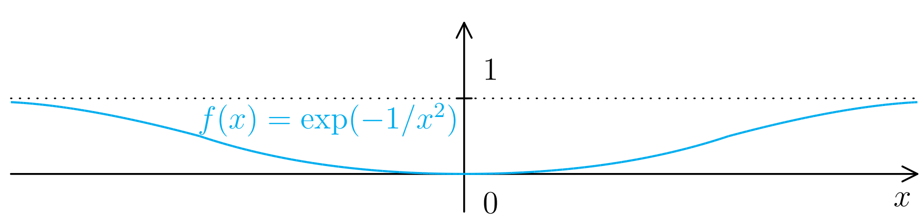
$$\sum_{n=0}^{\infty}(-1)^n x^{2n} = \sum_{n=0}^{\infty}(-x^2)^n = \frac{1}{1+x^2}.$$

Die Grenzfunktion ist auf ganz $\mathbb{R}$ definiert und beliebig oft differenzierbar, aber die Taylorreihe konvergiert nur auf $(-1, +1)$.

C. Noch verrückter verhält sich die Funktion

$$f(x) := \begin{cases} \exp\left(-1/x^2\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

Offensichtlich ist f stetig und für $x \neq 0$ beliebig oft differenzierbar. Außerdem ist $x = 0$ die einzige Nullstelle von f.



Wir wollen zeigen, dass f nur im Nullpunkt mit seiner Taylorreihe übereinstimmt. Dazu zeigen wir: *Es gibt rationale Funktionen $q_k(x)$ (die höchstens in 0 eine Polstelle besitzen), so dass gilt:*

$$f^{(k)}(x) = q_k(x) \cdot e^{-1/x^2}, \qquad \text{für } x \neq 0 \text{ und } k \in \mathbb{N}_0.$$

Den Beweis dazu führen wir durch vollständige Induktion:

Offensichtlich können wir $q_0(x) := 1$ setzen. Ist nun $k \geq 1$ und die Behauptung für $k-1$ schon bewiesen, so folgt:

$$\begin{aligned} f^{(k)}(x) &= {f^{(k-1)}}'(x) \\ &= q'_{k-1}(x) \cdot e^{-1/x^2} + q_{k-1}(x) \cdot 2x^{-3} \cdot e^{-1/x^2} \\ &= (q'_{k-1}(x) + 2q_{k-1}(x)x^{-3}) \cdot e^{-1/x^2}. \end{aligned}$$

Also hat $f^{(k)}$ die gewünschte Gestalt. ∎

Da die Exponentialfunktion stärker als jedes Polynom wächst, folgt:

$$\lim_{x \to 0} f^{(k)}(x) = \lim_{x \to 0} \frac{q_k(x)}{e^{1/x^2}} = \lim_{x \to \infty} \frac{q_k(1/x)}{e^{x^2}} = 0.$$

Also ist f auch im Nullpunkt beliebig oft differenzierbar, und $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}_0$.

Damit folgt, dass $Tf(x;0) \equiv 0$ ist. Da die Funktion außerhalb des Nullpunktes stets $\neq 0$ ist, konvergiert die Taylorreihe nur im Entwicklungspunkt selbst gegen die Funktion.

Ist f nicht beliebig oft differenzierbar, so kann man die Taylorreihe nicht bilden. Aber zumindest die ersten Terme dieser Reihe existieren.

Definition

Sei $I \subset \mathbb{R}$ ein Intervall, f n-mal differenzierbar auf I und $a \in I$ ein fester Punkt. Das Polynom

$$\boxed{T_n f(x) = T_n f(x;a) := \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k}$$

heißt ***n-tes Taylorpolynom*** von f in a.

Offensichtlich ist

$$\begin{aligned} T_1 f(x;a) &= f(a) + f'(a)(x-a) \qquad (= \text{lineare Approximation von } f \text{ in } a) \\ \text{und } T_2 f(x;a) &= f(a) + f'(a)(x-a) + \frac{1}{2} f''(a)(x-a)^2. \end{aligned}$$

Weil $(T_n f)^{(k)}(a) = f^{(k)}(a)$ für $k = 0, 1, 2, \ldots, n$ gilt, ist das Taylorpolynom ein guter Kandidat für eine Approximation von f in der Nähe von a. Wie gut die Approximation in der Nähe von a ist, darüber gibt das Verhalten des ***Restgliedes*** $R_n f(x) := f(x) - T_n f(x)$ Auskunft. Deshalb soll dieses nun genauer studiert werden:

7.4. Satz von der Taylorentwicklung

Es sei f auf I n-mal differenzierbar und $R_n f(x) := f(x) - T_n f(x)$.

1. *Es gibt eine Funktion $\eta(x)$ mit $\lim_{x \to a} \eta(x) = 0$, so dass gilt:*

$$R_n f(x) = \eta(x) \cdot (x-a)^n.$$

2. *Ist f auf I sogar $(n+1)$-mal differenzierbar, so gibt es zu jedem $x \neq a$ ein $c = c(x)$ zwischen a und x, so dass gilt:*

$$R_n f(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x-a)^{n+1}.$$

Man spricht dann auch von der „Lagrange'schen Form“ des Restgliedes.

Bemerkung: Die Darstellung $f(x) = T_n f(x;a) + R_n f(x)$ nennt man die ***Taylorentwicklung*** der Ordnung n von f im Punkte a.

BEWEIS: Wir beginnen mit dem zweiten Teil. Für $x \neq a$ ist

$$f(x) = T_n f(x) + \varphi(x) \cdot (x-a)^{n+1}, \qquad \text{mit } \varphi(x) := \frac{R_n f(x)}{(x-a)^{n+1}}.$$

Zähler und Nenner dieses Quotienten sind $(n+1)$-mal differenzierbar, und es ist $R_n f(a) = f(a) - f(a) = 0$. Der 2. Mittelwertsatz liefert zu jedem $x \neq a$ die Existenz einer Zahl $c_1 = c_1(x)$ zwischen a und x, so dass gilt:

$$\varphi(x) = \frac{R_n f(x) - R_n f(a)}{(x-a)^{n+1}} = \frac{(R_n f)'(c_1)}{(n+1)(c_1 - a)^n}.$$

Die $(n+1)$-te Ableitung von $(x-a)^{n+1}$ ergibt $(n+1)!$, die $(n+1)$-te Ableitung von $R_n f(x)$ ergibt $f^{(n+1)}(x)$, weil $T_n f(x)$ nur den Grad n hat. Eine $(n+1)$-fache Anwendung des 2. Mittelwertsatzes ergibt somit

$$\begin{aligned}
\varphi(x) &= \frac{(R_n f)'(c_1)}{(n+1)(c_1 - a)^n} \\
&= \frac{(R_n f)''(c_2)}{n(n+1)(c_2 - a)^{n-1}} \\
&\;\vdots \\
&= \frac{(R_n f)^{(n+1)}(c_{n+1})}{(n+1)!} = \frac{f^{(n+1)}(c_{n+1})}{(n+1)!},
\end{aligned}$$

mit geeigneten Punkten $c_i = c_i(x)$ zwischen a und x. Setzt man $c := c_{n+1}$, so erhält man

$$R_n f(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Wir zeigen jetzt die erste Aussage des Satzes, indem wir den schon bewiesenen zweiten Teil auf die Funktion $g := R_n f = f - T_n f$ anwenden. Es ist

$$g(a) = g'(a) = \ldots = g^{(n)}(a) = 0, \text{ also } T_k g(x) \equiv 0 \text{ für } k = 1, \ldots, n.$$

Weil f nur n-mal differenzierbar ist, ist auch g nur n-mal differenzierbar. Uns genügt allerdings zunächst schon die $(n-1)$-malige Differenzierbarkeit. Dann gibt es nämlich zu jedem $x \in I$ ein $c(x)$ zwischen a und x mit

$$g(x) = g(x) - T_{n-2}g(x) = \frac{g^{(n-1)}(c(x))}{(n-1)!}(x-a)^{n-1} = \eta(x) \cdot (x-a)^n,$$

wobei

$$\eta(x) := \begin{cases} \dfrac{g^{(n-1)}(c(x))}{(n-1)!(x-a)} & \text{für } x \neq a, \\ 0 & \text{für } x = a. \end{cases}$$

Um das Verhalten von $\eta(x)$ für $x \to a$ zu studieren, benutzen wir die nochmalige Differenzierbarkeit von $g^{(n-1)}$. Es gibt eine Funktion Δ mit

$$g^{(n-1)}(y) = g^{(n-1)}(a) + (y-a)\cdot\Delta(y) = (y-a)\cdot\Delta(y) \quad \text{und} \quad \lim_{y\to a} \Delta(y) = g^{(n)}(a) = 0.$$

Weil $\left| \dfrac{c(x)-a}{x-a} \right| \leq 1$ ist, folgt:

$$|\eta(x)| = \left| \frac{(c(x)-a)}{(n-1)!(x-a)} \cdot \Delta(c(x)) \right| \leq \frac{|\Delta(c(x))|}{(n-1)!},$$

und damit ist klar, dass $\eta(x)$ für $x \to a$ gegen Null konvergiert. Daraus folgt die erste Behauptung. ∎

Bemerkung: Die Beziehung $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)} = 0$ wird gerne mit Hilfe des *Landau'schen Symbols* o abgekürzt:

$$f(x) = \mathrm{o}(g(x)) \qquad (\text{für } x \to a\,).$$

Man kann dafür sagen: „$f(x)$ ist von der Ordnung $g(x)$“. Zum Beispiel ist

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \ldots = 1 + nx + \mathrm{o}(x), \ (\text{für } x \to 0)$$

oder $\quad x \cdot \sin(1/x) = \mathrm{o}(1)$ (für $x \to 0$).

Aus dem Satz von der Taylorentwicklung folgt: Ist f auf I n-mal differenzierbar, so ist

$$f(x) = T_n f(x) + \mathrm{o}((x-a)^n), \ (\text{für } x \to a).$$

Als Anwendung der Taylorentwicklung können wir jetzt das Problem der lokalen Extrema abschließend behandeln:

7.5. Hinreichendes Kriterium für Extremwerte

Die Funktion f sei in der Nähe von x_0 n-mal stetig differenzierbar. Es sei

$$\begin{array}{rrcll} & f^{(k)}(x_0) & = & 0 & \text{für } k = 1, \ldots, n-1 \\ \text{und} & f^{(n)}(x_0) & \neq & 0. & \end{array}$$

*Ist n ungerade, so besitzt f in x_0 **kein** lokales Extremum.*
Ist n gerade, so liegt ein lokales Extremum in x_0 vor, und zwar

$$\begin{array}{rcll} \text{ein Maximum, falls } f^{(n)}(x_0) & < & 0 & \text{ist,} \\ \text{und ein Minimum, falls } f^{(n)}(x_0) & > & 0 & \text{ist.} \end{array}$$

Beweis: Da $f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0$ ist, liefert die Taylorentwicklung mit $h := x - x_0$:

$$f(x_0 + h) = f(x_0) + \left(\frac{f^{(n)}(x_0)}{n!} + \eta(x_0 + h)\right) \cdot h^n,$$

mit $f^{(n)}(x_0) \neq 0$ und $\lim\limits_{x \to x_0} \eta(x) = 0$.

Ist $\varepsilon > 0$ klein genug gewählt, so hat $\dfrac{f^{(n)}(x_0)}{n!} + \eta(x_0 + h)$ für $|h| < \varepsilon$ das gleiche Vorzeichen wie $f^{(n)}(x_0)$.

Wir betrachten nur den Fall $f^{(n)}(x_0) > 0$, der andere geht analog. Wir können schreiben:

$$f(x_0 + h) - f(x_0) = \varphi(h) \cdot h^n, \text{ mit einer nahe 0 positiven Funktion } \varphi.$$

Ist n ungerade, so wechselt h^n bei $h = 0$ sein Vorzeichen, und es kann kein Extremwert vorliegen. Ist n gerade, so bleibt h^n immer ≥ 0 und verschwindet bei $h = 0$. Dann besitzt f in x_0 ein Minimum. ∎

Kommen wir nun zurück zu beliebig oft differenzierbaren Funktionen und ihren Taylorreihen.

7.6. Konvergenzkriterium für die Taylorreihe

Sei $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f : I \to \mathbb{R}$ beliebig oft differenzierbar. Gibt es Konstanten $C, r > 0$, so dass $|f^{(n)}(x)| \leq C \cdot r^n$ für alle $n \in \mathbb{N}$ und alle $x \in I$ gilt, so konvergiert die Taylorreihe Tf von f in x_0 auf ganz I gegen f.

Beweis: Zu jedem $x \in I$ gibt es ein c zwischen x und x_0 mit

$$|R_n f(x)| = \left|\frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x - x_0)^{n+1}\right| \leq C \cdot \frac{(r \cdot |x - x_0|)^{n+1}}{(n+1)!},$$

und die rechte Seite konvergiert gegen Null, wegen der Konvergenz der Exponentialreihe. Das gilt unabhängig von x. ■

Zwar kann man normalerweise über das Konvergenzverhalten einer Potenzreihe und die Stetigkeit der Grenzfunktion auf dem Rand des Konvergenzkreises keine Aussage machen, in einem ganz besonderen Fall geht es aber doch.

7.7. Abel'scher Grenzwertsatz

Es sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit reellen Koeffizienten und dem Konvergenzradius $R = 1$. Dann gilt:

$$\textit{Ist } a := \sum_{n=0}^{\infty} a_n < \infty, \ \textit{ so ist } \lim_{x \to 1-} f(x) = a.$$

Die Grenzfunktion wird also bei $x = 1$ durch $f(x) := a$ stetig fortgesetzt.

Beweis: Zur Abkürzung setzen wir $S_{m,k} := \sum_{n=m+1}^{m+k} a_n$.

Ist $\varepsilon > 0$ vorgegeben, so gibt es ein $m_0 \in \mathbb{N}$, so dass $|S_{m,k}| < \varepsilon$ für $m \geq m_0$ und $k \geq 1$ ist (Cauchykriterium). Für solche m und k ist dann

$$\begin{aligned} \sum_{n=m+1}^{m+k} a_n x^n &= S_{m,1} x^{m+1} + (S_{m,2} - S_{m,1}) x^{m+2} + \cdots + (S_{m,k} - S_{m,k-1}) x^{m+k} \\ &= S_{m,1}(x^{m+1} - x^{m+2}) + S_{m,2}(x^{m+2} - x^{m+3}) + \cdots + S_{m,k} x^{m+k}. \end{aligned}$$

Da $|S_{m,k}| < \varepsilon$ für $m \geq m_0$ und $k \geq 1$ und $x^{n+1} \leq x^n \leq 1$ für $x \in [0,1]$ ist, folgt:

$$\Big| \sum_{n=m+1}^{m+k} a_n x^n \Big| < \varepsilon \cdot x^{m+1} \leq \varepsilon \text{ für } x \in [0,1] \text{ und alle } m \geq m_0 \text{ und } k \geq 1.$$

Der Übergang zum Grenzwert $k \to \infty$ liefert dann für $F_m(x) := \sum_{n=0}^{m} a_n x^n$:

$$|f(x) - F_m(x)| \leq \varepsilon \text{ für } m \geq m_0 \text{ und alle } x \in [0,1].$$

Das bedeutet, dass die Folge der stetigen Polynome F_m auf $[0,1]$ gleichmäßig gegen $f(x)$ konvergiert, also f auf $[0,1]$ stetig ist. ■

7.8. Beispiele

A. Es ist

$$\sin(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} + o(x^{2n+2}).$$

Weil alle geraden Ableitungen des Sinus bei $x = 0$ verschwinden, gewinnt man eine Ordnung zusätzlich.

In diesem Fall ist die Funktion durch die Taylorreihe definiert worden. Die Frage, ob die Taylorreihe gegen die Funktion konvergiert, stellt sich also nicht.

Die Cosinus-Reihe verhält sich analog.

B. Die Funktion $f(x) := \ln(1+x)$ ist im Intervall $I = (-1,1)$ beliebig oft differenzierbar. Die Ableitung $f'(x)$ kann durch die geometrische Reihe

$$f'(x) = \frac{1}{1+x} = \sum_{n=0}^{\infty}(-1)^n x^n$$

dargestellt werden. Als Potenzreihe ist diese Reihe gleich ihrer eigenen Taylorreihe. Allerdings ist $f'(x)$ auch für $x > 1$ definiert, während die Reihe dort divergiert. Weiter gilt

$$\begin{aligned} f(x) \;=\; f(x) - f(0) \;&=\; \int_0^x f'(t)\,dt \;=\; \sum_{n=0}^{\infty}(-1)^n \int_0^x t^n\,dt \\ &=\; \sum_{n=0}^{\infty}(-1)^n \frac{t^{n+1}}{(n+1)}\Big|_0^x = \sum_{n=0}^{\infty}(-1)^n \frac{x^{n+1}}{(n+1)}\,. \end{aligned}$$

Also ist

$$\ln(1+x) = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n}\,.$$

Die alternierende Reihe $\sum_{n=1}^{\infty}(-1)^n/n$ konvergiert nach dem Leibniz–Kriterium. Die Grenzfunktion $f(x) = \ln(1+x)$ der Potenzreihe ist in $x = 1$ noch stetig, mit $f(1) = \ln(2)$. Aus dem Abel'schen Grenzwertsatz folgt jetzt:

$$\boxed{\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n} = \ln(2).}$$

C. Wir betrachten die Funktion $f(x) := \arctan(x)$. Es ist

$$f'(x) = \frac{1}{1+x^2} = \sum_{n=0}^{\infty}(-1)^n x^{2n},$$

nach der Summenformel für die geometrische Reihe. Die Gleichung gilt für $|x| < 1$. Für solche x ist dann

$$f(x) = \int_0^x \frac{dt}{1+t^2} = \sum_{n=0}^{\infty}(-1)^n \int_0^x t^{2n}\,dt = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n+1}}{2n+1}\,.$$

Auch hier gilt die Darstellung nicht über $(-1,1)$ hinaus. Lediglich am rechten Randpunkt des Intervalls geht noch etwas. Nach dem Leibniz–Kriterium ist $\sum_{n=0}^{\infty}(-1)^n/(2n+1)$ konvergent, und $f(x) = \arctan(x)$ ist bei $x = 1$ noch stetig, mit $f(1) = \pi/4$. Der Abel'sche Grenzwertsatz liefert

$$\boxed{\sum_{n=0}^{\infty}(-1)^n \frac{1}{2n+1} = \frac{\pi}{4}.}$$

Zum Schluss wollen wir noch eine besonders interessante Taylorentwicklung betrachten. Ausgangspunkt ist das Polynom

$$p(x) := (1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k. \text{ Dabei ist } \binom{n}{k} = \frac{p^{(k)}(0)}{k!}.$$

Diese Situation wollen wir jetzt verallgemeinern und die Funktion

$$f(x) := (1+x)^\alpha \qquad \text{für } |x| < 1 \text{ und } \alpha \in \mathbb{R}$$

untersuchen. Hier ist

$$\frac{f^{(k)}(0)}{k!} = \frac{\alpha(\alpha-1)\cdot\ldots\cdot(\alpha-k+1)}{k!}.$$

Man definiert deshalb die ***verallgemeinerten Binomialkoeffizienten*** $\binom{\alpha}{k}$ durch

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\cdot\ldots\cdot(\alpha-k+1)}{k!} \qquad \text{und} \qquad \binom{\alpha}{0} := 1.$$

Man beachte, dass diese Zahlen weder ganz noch positiv zu sein brauchen. Es gilt aber z.B.

$$\begin{aligned}\binom{\alpha-1}{k} + \binom{\alpha-1}{k-1} &= \frac{(\alpha-1)\cdots(\alpha-k)}{k!} + \frac{(\alpha-1)\cdots(\alpha-k+1)}{(k-1)!} \\ &= \frac{(\alpha-1)\cdots(\alpha-k+1)((\alpha-k)+k)}{k!} = \binom{\alpha}{k}.\end{aligned}$$

7.9. Satz

Die Binomialreihe $B(x) := \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$ *konvergiert für* $|x| < 1$ *gegen* $(1+x)^\alpha$.

BEWEIS: 1) Den Konvergenzradius der Reihe bestimmen wir mit der Quotientenregel: Es ist $c_n = \binom{\alpha}{n}$ und

$$\begin{aligned} \left|\frac{c_n}{c_{n+1}}\right| &= \left|\frac{\alpha(\alpha-1)\cdot\ldots\cdot(\alpha-n+1)\cdot(n+1)!}{\alpha(\alpha-1)\cdot\ldots\cdot(\alpha-n)\cdot n!}\right| \\ &= \frac{n+1}{|\alpha-n|} = \frac{1+(1/n)}{|1-(\alpha/n)|}, \end{aligned}$$

und dieser Ausdruck konvergiert gegen 1.

2) $B(x)$ ist also eine differenzierbare Funktion auf $(-1,+1)$, und es gilt:

$$\begin{aligned} B'(x) &= \sum_{n=1}^{\infty} n\cdot\binom{\alpha}{n}x^{n-1} = \sum_{n=0}^{\infty}(n+1)\cdot\binom{\alpha}{n+1}x^n \\ &= \alpha\cdot\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^n, \end{aligned}$$

also

$$\begin{aligned} (1+x)B'(x) &= \alpha\cdot\left(\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^n + \sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^{n+1}\right) \\ &= \alpha\cdot\left(1+\sum_{n=1}^{\infty}\left[\binom{\alpha-1}{n}+\binom{\alpha-1}{n-1}\right]x^n\right) \\ &= \alpha\cdot\left(1+\sum_{n=1}^{\infty}\binom{\alpha}{n}x^n\right) = \alpha\cdot B(x). \end{aligned}$$

Setzen wir $h(x) := \dfrac{B(x)}{(1+x)^\alpha}$, so folgt:

$$\begin{aligned} h'(x) &= \frac{(1+x)^\alpha B'(x) - B(x)\alpha(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} \\ &= \frac{(1+x)^{\alpha-1}}{(1+x)^{2\alpha}}\cdot((1+x)B'(x)-\alpha B(x)) \equiv 0. \end{aligned}$$

Demnach muss $h(x)$ konstant sein. Weil $h(0)=1$ ist, folgt, dass $B(x)=(1+x)^\alpha$ für $|x|<1$ ist. ■

7.10. Beispiel

Es ist

$$\sqrt{1+x} = (1+x)^{1/2} = \sum_{n=0}^{\infty}\binom{\frac{1}{2}}{n}x^n = 1+\frac{1}{2}x-\frac{1}{8}x^2+\frac{1}{16}x^3 \pm \cdots .$$