Bergische Universität – Gesamthochschule Wuppertal – Fachbereich Mathematik

Analysis 1

Kapitel 3 Differentialrechnung

Vorlesungsausarbeitung zum WS 2000/01

von Prof. Dr. Klaus Fritzsche

Inhaltsverzeichnis

§ 1 Die Ableitung

Sei $M \subset \mathbb{R}$ eine Teilmenge. Wir setzen voraus, daß jeder Punkt $x \in M$ auch ein Häufungspunkt von M ist. Ist etwa M ein Intervall, so ist diese Bedingung auf jeden Fall erfüllt.

Wir wollen eine Funktion $f : M \to \mathbb{R}$ in einem Punkt $x_0 \in M$ durch eine affin-lineare Funktion $L(x) = mx+b$ approximieren. Dann muß zunächst einmal $L(x_0) = f(x_0)$ sein, also $b = f(x_0) - mx_0$ und

$$L(x) = f(x_0) + m(x - x_0).$$

Der Faktor m beschreibt die Steigung von L. Soll auch noch $L(x_1) = f(x_1)$ sein, für irgend einen Punkt $x_1 \neq x_0$, so ist diese Steigung eindeutig festgelegt:

$$m = \Delta f(x_0, x_1) := \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Man bezeichnet $\Delta f(x_0, x_1)$ auch als den *Differenzenquotienten* von f, bezogen auf x_0 und x_1, und die durch L beschriebene Gerade nennt man die *Sekante* durch x_0 und x_1. Wenn der Graph von f bei x_0 hinreichend „glatt" ist, so kann man erwarten, daß die Richtung der Sekante bei Annäherung von x_1 an x_0 gegen die Richtung der Tangente an den Graphen bei x_0 strebt. Dann liefert die zu der Tangente gehörende affin-lineare Funktion $L_0 = f(x_0) + \Delta(x_0)(x - x_0)$ (mit $\Delta(x_0) = \lim\limits_{x \to x_0} \Delta(x)$) eine optimale Approximation von f.

Der „Fehler" $\varphi(x) := f(x) - L_0(x) = f(x) - f(x_0) - \Delta(x_0)(x - x_0)$ bei dieser Approximation erfüllt die Gleichung

$$\Delta(x) - \Delta(x_0) = \frac{\varphi(x)}{x - x_0}, \qquad \text{also } \lim_{x \to x_0} \frac{\varphi(x)}{x - x_0} = 0.$$

Wir betrachten nun gleich die etwas allgemeinere Situation einer Abbildung

$$f = (f_1, \ldots, f_n) : M \to \mathbb{R}^n.$$

1.1 Satz. *Folgende Aussagen über eine Funktion $f : M \to \mathbb{R}^n$ sind äquivalent:*

1. *Es existiert der Grenzwert $f'(x_0) := \lim\limits_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0}$.*
2. *Es gibt einen Vektor $a \in \mathbb{R}^n$ und eine Funktion $\varphi : M \to \mathbb{R}^n$, mit*

$$f(x) = f(x_0) + (x - x_0) \cdot a + \varphi(x) \qquad \textit{auf } M$$

und

$$\lim_{x \to x_0} \frac{\varphi(x)}{x - x_0} = 0.$$

3. *Es gibt eine in x_0 stetige Abbildung $\Delta : M \to \mathbb{R}^n$ mit*

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

BEWEIS: (1) $\Longrightarrow$ (2): Setze $a := f'(x_0)$ und $\varphi(x) := f(x) - f(x_0) - (x - x_0) \cdot a$. Dann ist

$$f(x) = f(x_0) + (x - x_0) \cdot a + \varphi(x), \text{ und es existiert}$$

$$\lim_{x \to x_0} \frac{\varphi(x)}{x - x_0} = \lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} - a\right) = 0.$$

(2) $\Longrightarrow$ (3): Setze

$$\Delta(x) := \begin{cases} a + \dfrac{\varphi(x)}{x - x_0} & \text{für } x \neq x_0, \\ a & \text{sonst.} \end{cases}.$$

Dann ist $f(x) = f(x_0) + (x - x_0) \cdot \Delta(x)$ und $\lim_{x \to x_0} \Delta(x) = \Delta(x_0)$.

(3) $\Longrightarrow$ (1): Es ist $\Delta(x) = \dfrac{f(x) - f(x_0)}{x - x_0}$ für $x \neq x_0$. Wegen der Stetigkeit von Δ in x_0 existiert $\lim_{x \to x_0} \Delta(x)$. ∎

Definition. $f : M \to \mathbb{R}^n$ heißt *differenzierbar* in x_0, falls die äquivalenten Bedingungen des Satzes erfüllt sind. Der Vektor $f'(x_0)$ heißt die *Ableitung* von f in x_0. Im Falle $n > 1$ spricht man auch vom *Tangentenvektor*.

Im Falle $n = 1$ nennt man die lineare Abbildung $(df)_{x_0} : \mathbb{R} \to \mathbb{R}$ mit $(df)_{x_0}(u) := u \cdot f'(x_0)$ das *Differential* von f in x_0.

Ist f auf ganz M differenzierbar, so ist die *Ableitung* von f auf M die Funktion $f' : x \mapsto f'(x)$ auf M.

Beispiele.

1. Die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) \equiv x$ ist in jedem Punkt $x_0 \in \mathbb{R}$ differenzierbar, denn es ist $x = x_0 + 1 \cdot (x - x_0)$. Also ist $f'(x_0) = 1$, für jedes x_0, und $(dx)_{x_0}(u) = u$. Daraus folgt für eine beliebige in x_0 differenzierbare Funktion f die Beziehung

   $$(df)_{x_0} = f'(x_0) \cdot (dx)_{x_0}.$$

   Man schreibt deshalb auch $f'(x_0) = \dfrac{df}{dx}(x_0)$ und spricht vom *Differentialquotienten*. Natürlich handelt es sich nicht wirklich um einen Quotienten.

2. Sei $c \in \mathbb{R}$. Wegen $c = c + 0 \cdot (x - x_0)$ ist jede konstante Funktion $f(x) \equiv c$ in jedem Punkt $x_0 \in \mathbb{R}$ differenzierbar, und die Ableitung ist $= 0$.

3. Ist f in $x_0 \in M$ differenzierbar, so ist f dort auch stetig (wegen der Darstellung $f(x) = f(x_0) + \Delta(x)(x - x_0)$). Die Umkehrung gilt nicht. So ist z.B. $f(x) = |x|$ in $x_0 = 0$ stetig, aber nicht differenzierbar:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{|x|}{x} = \begin{cases} 1 & \text{für } x > 0 \\ -1 & \text{für } x < 0 \end{cases}$$

besitzt keinen Grenzwert für $x \to x_0$.

4. Durch $f(t) := x_0 + t \cdot v$ wird die Gerade durch x_0 mit Richtung v parametrisiert. Es ist

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

also $f'(t) \equiv v$. Physikalisch gesehen ist v die Geschwindigkeit, mit der die Gerade durchlaufen wird.

5. Eine vektorwertige Funktion $f = (f_1, \ldots, f_n) : M \to \mathbb{R}^n$ ist genau dann in $x_0 \in M$ differenzierbar, wenn alle Komponenten f_i es sind, und dann ist $f'(x_0) = (f_1'(x_0), \ldots, f_n'(x_0))$. Das sieht man unmittelbar, wenn man Kriterium (3) für die Differenzierbarkeit verwendet.

   Insbesondere gilt für komplexwertige Funktionen $f = g + \mathrm{i}h : M \to \mathbb{C}$: f ist genau dann differenzierbar in $x_0 \in M$, wenn g und h es sind, und dann ist $(g + \mathrm{i}h)' = g' + \mathrm{i}h'$.

1.2 Satz. *Sei $f : M \to \mathbb{R}^n$ differenzierbar in x_0 und $u : \mathbb{R}^n \to \mathbb{R}^k$ linear. Dann ist auch $u \circ f : M \to \mathbb{R}^k$ in x_0 differenzierbar, und es ist*

$$(u \circ f)'(x_0) = u(f'(x_0)).$$

Beweis:

$$\frac{u \circ f(x) - u \circ f(x_0)}{x - x_0} = u\left(\frac{f(x) - f(x_0)}{x - x_0}\right),$$

und da u stetig ist, konvergiert der rechte Ausdruck für $x \to x_0$ gegen $u(f'(x_0))$. ■

Beispiele.

1. Sind $f, g : M \to \mathbb{R}$ differenzierbar, so auch $f \pm g$, und es ist $(f \pm g)' = f' \pm g'$.

   Beweis: $(f, g) : M \to \mathbb{R}^2$ ist differenzierbar, und $u : \mathbb{R}^2 \to \mathbb{R}$ mit $u(x_1, x_2) := x_1 \pm x_2$ ist linear. ■

   Analog folgt: Sind $f, g : M \to \mathbb{R}^n$ differenzierbar, so ist auch $f \pm g : M \to \mathbb{R}^n$ differenzierbar, und $(f \pm g)' = f' \pm g'$.

2. Sei $f = g + \mathrm{i}h$ eine komplexwertige differenzierbare Funktion auf M und $c = a + \mathrm{i}b$ eine komplexe Zahl. Dann kann man f auch als Funktion $(g, h) : M \to \mathbb{R}^2$ auffassen. Die Abbildung $M_c : \mathbb{C} \to \mathbb{C}$ mit $M_c(z) := c \cdot z$ ist in reellen Koordinaten gegeben durch $(x, y) \mapsto (ax - by, ay + bx)$, stellt also eine $\mathbb{R}$-lineare Abbildung $M_c : \mathbb{R}^2 \to \mathbb{R}^2$ dar. Offensichtlich folgt jetzt:

   $c \cdot f$ ist differenzierbar, mit $(c \cdot f)' = c \cdot f'$.

1.3 Verallgemeinerte Produktregel. *Sind $f, g : M \to \mathbb{R}^n$ beide in x_0 differenzierbar, so ist auch $f \bullet g : M \to \mathbb{R}$ in x_0 differenzierbar, und es gilt:*

$$(f \bullet g)'(x_0) = f'(x_0) \bullet g(x_0) + f(x_0) \bullet g'(x_0).$$

Beweis: Man benutzt einen kleinen Trick:

$$\begin{aligned}\frac{(f \bullet g)(x) - (f \bullet g)(x_0)}{x - x_0} &= \frac{(f(x) - f(x_0)) \bullet g(x) + f(x_0) \bullet (g(x) - g(x_0))}{x - x_0} \\ &= \frac{f(x) - f(x_0}{x - x_0} \bullet g(x) + f(x_0) \bullet \frac{g(x) - g(x_0)}{x - x_0}.\end{aligned}$$

strebt für $x \to x_0$ gegen $f'(x_0) \bullet g(x_0) + f(x_0) \bullet g'(x_0)$. ∎

Hierin ist die gewöhnliche Produktregel $(f \cdot g)' = f' \cdot g + f \cdot g'$ enthalten. Ein einfacher Induktionsbeweis liefert nun die Regel

$$(x^n)' = n \cdot x^{n-1} \quad (\text{denn} \quad (x \cdot x^{n-1})' = 1 \cdot x^{n-1} + (n-1) \cdot x \cdot x^{n-2}).$$

Aus der Produktregel folgt sogleich auch die bekannte „Quotientenregel“

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

die man erhält, indem man die rechte Seite der Gleichung $f' = (g \cdot \dfrac{f}{g})'$ weiter ausrechnet.

Insbesondere ist $(x^{-n})' = -n \cdot x^{-n-1}$, denn $\quad (\dfrac{1}{x^n})' = \dfrac{-n \cdot x^{n-1}}{x^{2n}}$.

1.4 Kettenregel. *Sei $f : M \to \mathbb{R}^n$ differenzierbar in x_0, $I \subset \mathbb{R}$ ein Intervall und $g : I \to M$ differenzierbar in $t_0 \in I$, $g(t_0) = x_0$. Dann ist auch $f \circ g : I \to \mathbb{R}^n$ differenzierbar in t_0, und es gilt:*

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

Beweis: Sei $f(x) = f(x_0) + (x - x_0) \cdot \Delta_f(x)$, mit einer in x_0 stetigen Funktion $\Delta_f : M \to \mathbb{R}^n$, sowie $g(t) = g(t_0) + (t - t_0) \cdot \Delta_g(t)$, mit einer in t_0 stetigen Funktion $\Delta_g : I \to \mathbb{R}$. Dann folgt:

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

Weil $\Delta_g \cdot (\Delta_f \circ g)$ stetig in t_0 ist, ist $f \circ g$ dort differenzierbar. Der Wert von $\Delta_g \cdot (\Delta_f \circ g)$ an der Stelle t_0 ergibt den Wert der Ableitung. ■

Definition. Sei $f : M \to \mathbb{R}^n$ überall differenzierbar. Ist die abgeleitete Funktion $f' : M \to \mathbb{R}^n$ in $x_0 \in M$ erneut differenzierbar, so sagt man, f ist in x_0 *zweimal differenzierbar*, und

$$f''(x_0) := (f')'(x_0)$$

heißt die *zweite Ableitung* von f in x_0.

Induktiv definiert man nun für $n \in \mathbb{N}$: f heißt in x_0 *n-mal differenzierbar*, falls f auf ganz M (oder zumindest auf einer Umgebung von x_0) $(n-1)$-mal differenzierbar und die $(n-1)$-te Ableitung $f^{(n-1)}$ in x_0 ein weiteres Mal differenzierbar ist. $f^{(n)}(x_0) := (f^{(n-1)})'(x_0)$ heißt dann die *n-te Ableitung* von f in x_0.

Bemerkung. Manchmal benutzt man auch die Leibnizsche Schreibweise:

$$\frac{d^n f}{dx^n} \text{ statt } f^{(n)}.$$

Beispiele.

1. Wir betrachten die reelle Exponentialfunktion, und zwar zunächst im Nullpunkt. Es ist

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

   Da $\Delta(x) := \sum_{n=0}^{\infty} \frac{x^n}{(n+1)!}$ den Konvergenzradius ∞ hat, ist Δ stetig. Außerdem ist $\Delta(0) = 1$. Es folgt, daß exp in $x = 0$ differenzierbar ist, mit $\exp'(0) = 1$.

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

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

   Daraus folgt, daß exp in x_0 differenzierbar ist, und $\exp'(x_0) = \exp(x_0)$. Also ist allgemein

   $$(e^x)' = e^x.$$

   Ersetzen wir x durch $\mathrm{i}t$ und x_0 durch $\mathrm{i}t_0$, so erhalten wir mit dem gleichen Beweis die Formel

   $$\exp(\mathrm{i}t) - \exp(\mathrm{i}t_0) = \mathrm{i}(t - t_0) \cdot \exp(\mathrm{i}t_0) \cdot \Delta(\mathrm{i}(t - t_0)).$$

   Also ist auch $t \mapsto e^{\mathrm{i}t}$ in t_0 differenzierbar, und $(e^{\mathrm{i}t})' = \mathrm{i} \cdot e^{\mathrm{i}t}$.

2. Aus der Eulerschen Formel $e^{\mathrm{i}t} = \cos t + \mathrm{i} \sin t$ erhalten wir:

   $\cos(t)$ und $\sin(t)$ sind auf ganz $\mathbb{R}$ differenzierbar, und es gilt:

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

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

   Sei $r > 0$ und $x_0 = (x_1^{(0)}, x_2^{(0)}) \in \mathbb{R}^2$ ein fester Punkt. Durch $f : [0, 2\pi] \to \mathbb{R}^2$ mit
   $$f(t) := (x_1^{(0)} + r \cdot \cos(t), x_2^{(0)} + r \cdot \sin(t))$$
   wird der Kreis um x_0 mit Radius r parametrisiert. Es ist

   $$f'(t) = (-r \cdot \sin(t), r \cdot \cos(t)).$$

   Sind $v, w \in \mathbb{R}^n$ zwei beliebige Vektoren $\neq 0$, so ist $\|v\| \neq 0$ und $\|w\| \neq 0$. Außerdem gilt die Schwarzsche Ungleichung: $|v \bullet w| \leq \|v\| \cdot \|w\|$. Daraus folgt:

   $$-1 \leq \frac{v \bullet w}{\|v\| \cdot \|w\|} \leq 1.$$

   Das wiederum bedeutet, daß es eine eindeutig bestimmte Zahl $\theta \in [0, \pi]$ gibt, so daß gilt:
   $$v \bullet w = \|v\| \cdot \|w\| \cdot \cos(\theta).$$
   Die Zahl $\angle(v, w) := \theta$ nennt man den *Winkel* zwischen v und w. Ist $\angle(v, w) = \frac{\pi}{2}$, so sagt man, v und w *stehen senkrecht aufeinander* (oder *sind orthogonal zueinander*). Das ist genau dann der Fall, wenn $v \bullet w = 0$ ist.

   Im Falle der Kreislinie f gilt: der „Radius-Vektor“ $f(t) - x_0$ und der Tangentenvektor $f'(t)$ stehen für jedes $t \in [0, \pi]$ aufeinander senkrecht.

3. Aus der Quotientenregel folgt:

   $$\begin{aligned} \tan'(x) &= \frac{\sin'(x)\cos(x) - \sin(x)\cos'(x)}{\cos^2(x)} \\ &= \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} \\ &= \frac{1}{\cos^2(x)} \quad (\text{oder } = 1 + \tan^2(x)\,). \end{aligned}$$

4. Mit Hilfe der Kettenregel erhält man z.B. für $f(x) := e^{x \cdot \sin(x^2)}$:

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

   Allgemein ist $(e^f)' = f' \cdot e^f$, also z.B. $(a^x)' = (e^{x \cdot \ln(a)})' = \ln(a) \cdot a^x$.

5. Sei $f(x) := e^{x^2}$. Dann gilt:

$$\begin{aligned} f'(x) &= 2x \cdot e^{x^2}, \\ f''(x) &= 2 \cdot e^{x^2} + 2x \cdot (2x \cdot e^{x^2}) = (2 + 4x^2) \cdot e^{x^2}, \\ f^{(3)}(x) &= 8x \cdot e^{x^2} + (2 + 4x^2) \cdot (2x \cdot e^{x^2}) = (12x + 8x^3) \cdot e^{x^2}. \end{aligned}$$

Die Versuche lassen folgendes vermuten:

$$f^{(n)}(x) = p_n(x) \cdot e^{x^2},$$

mit einem Polynom $p_n(x)$ vom Grad n, das entweder nur gerade oder nur ungerade Potenzen von x enthält, je nachdem, ob n gerade oder ungerade ist. Für kleine n haben wir das schon verifiziert. Nun führen wir einen Induktionsbeweis nach n. Ist die Formel für ein $n \geq 1$ richtig, so gilt:

$$f^{(n+1)}(x) = (p_n'(x) + 2x \cdot p_n(x)) \cdot e^{x^2}.$$

Ist etwa $n = 2k$, so enthält $p_n(x)$ nur gerade Potenzen von x. Aber dann ist $p_n'(x)$ ein Polynom vom Grad $n-1$ und $2x \cdot p_n(x)$ ein Polynom vom Grad $n+1$, und beide enthalten nur ungerade Potenzen von x. Analog argumentiert man im Falle $n = 2k+1$. Also gilt die Formel auch für $n+1$, und damit für alle $n \in \mathbb{N}$.

1.5 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 $f^{-1} : J \to I$ 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.

Wäre $\Delta(x) = 0$ für ein $x \in M$, so wäre $f(x) = f(x_0)$. Wegen der Injektivität von f müßte dann $x = x_0$ sein, was nach Voraussetzung ausgeschlossen ist. 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 ein, so erhalten wir:

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

Da f abgeschlossene Teilintervalle von I homöomorph abbildet, ist f^{-1} in y_0 stetig. Also ist auch $\dfrac{1}{\Delta(f^{-1}(y))}$ in y_0 stetig und f^{-1} in y_0 differenzierbar, mit

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

■

Beispiele.

1. $\exp : \mathbb{R} \to \mathbb{R}_+$ 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 $(\ln)'(e^x) = \dfrac{1}{e^x}$, also
$$\ln'(y) = \frac{1}{y}.$$

2. Die Funktion x^α ist für festes $\alpha \neq 0$ und $x > 0$ definiert durch $x^\alpha := e^{\alpha \ln(x)}$. Dann 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
$$(\sqrt[n]{x})' = \frac{1}{n} \cdot x^{-\frac{n-1}{n}},$$
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:
$$(x^x)' = (\ln x + 1) \cdot x^x.$$

3. 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:
$$g'(x) = (\ln \circ f)'(x) = \ln'(f(x)) \cdot f'(x) = \frac{f'(x)}{f(x)}.$$
Man nennt diesen Ausdruck auch die „logarithmische Ableitung“ von f.

Definition. $f : I \to \mathbb{R}$ hat in $x_0 \in I$ ein *lokales Maximum* (bzw. *lokales Minimum*), falls gilt:

$$\exists\, \varepsilon > 0, \text{ so daß } f(x) \leq f(x_0) \text{ (bzw. } f(x) \geq f(x_0)\text{) für } x \in I \text{ und } |x - x_0| < \varepsilon \text{ ist.}$$

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! 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 gilt:

$$\exists\, \varepsilon > 0, \text{ so daß } f(x) < f(x_0) \text{ für } |x - x_0| < \varepsilon \text{ und } x \neq x_0 \text{ ist}$$

$$(\text{bzw. } f(x) > f(x_0) \text{ im Falle des Minimums}).$$

Sei $I \subset \mathbb{R}$ ein beliebiges Intervall. Wir nennen einen Punkt $x_0 \in I$ einen *inneren Punkt* von I, falls es ein $\varepsilon > 0$ gibt, so daß $U_\varepsilon(x_0) \subset I$ ist. Das bedeutet, daß I eine Umgebung für x_0 darstellt.

1.6 „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: Wir betrachten den Differenzenquotienten $\Delta f(x_0, x) := \dfrac{f(x) - f(x_0)}{x - x_0}$, für $x \neq x_0$, und behandeln nur den Fall des lokalen Maximums, beim Minimum geht es analog.

Hat f in x_0 ein lokales Maximum, so ist $f(x) \leq f(x_0)$ für x nahe bei x_0. Ist $x < x_0$, so ist $x - x_0 < 0$ und daher $\Delta f(x_0, x) \geq 0$. Ist jedoch $x > x_0$, so ist $\Delta f(x_0, x) \leq 0$. Aber dann muß $f'(x_0) = \lim_{x \to x_0} \Delta f(x_0, x) = 0$ sein. ■

Hinreichende Kriterien behandeln wir später.

Beispiele.

1. Sei $I = [-1, 1]$, $f : I \to \mathbb{R}$ definiert durch $f(x) := x^2$. Dann gilt für alle $x \in I : f(x) \geq 0 = f(0)$. 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.

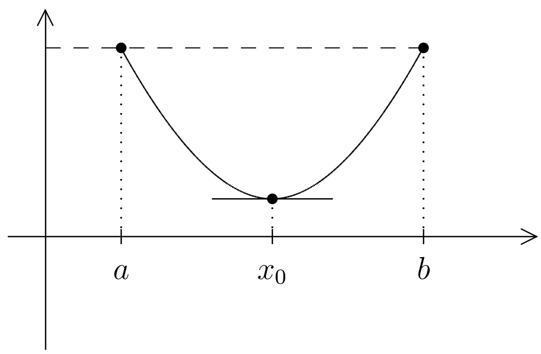
2. $g : I \to \mathbb{R}$ mit $g(x) := |x|$ hat ebenfalls in $x_0 = 0$ ein lokales Minimum. Aber weil $|x|$ dort nicht differenzierbar ist, kann man das Kriterium nicht anwenden.

   In den Punkten $x = -1$ und $x = +1$ hat $|x|$ (als Funktion auf I) jeweils ein lokales Maximum. Aber das kann man nicht aus dem notwendige Kriterium herleiten, denn die Punkte liegen nicht im Innern von I.

3. Sei $f(x) := x^3$. Dann hat $f'(x) = 3x^2$ in $x = 0$ eine Nullstelle. Allerdings liegt kein lokales Extremum vor. Für $x < 0$ ist $f(x) < 0$ und für $x > 0$ ist $f(x) > 0$.

§ 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 $f'(x_0) = 0$.*

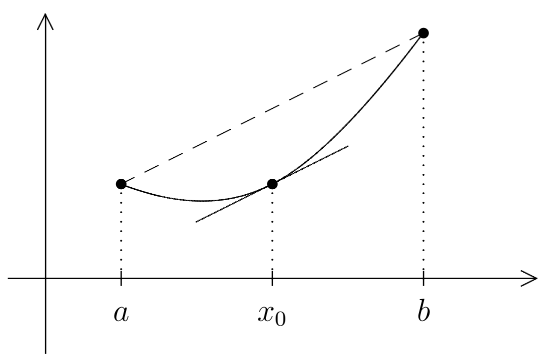


BEWEIS: Sei $c := f(a) = f(b)$. Ist $f(x) \equiv c$ auf ganz I, so ist auch $f'(x) \equiv 0$. Ist f auf I nicht konstant, so muß entweder das Minimum oder das Maximum von f im Innern von I liegen. Und dort muß dann f' verschwinden. ∎

Der folgende wichtige Satz ist eine einfache Folgerung:

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 $f'(x_0) = m$. ∎

So einfach der Beweis, so mächtig die Konsequenzen:

2.3 Satz. *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 vektorwertige Funktionen richtig, also insbesondere für komplexwertige Funktionen. Um die nächste Folgerung ebenfalls für komplexwertige Funktionen zu erhalten, müssen wir zeigen, daß die Produktregel für solche Funktionen gilt:

$f_1 = g_1 + \mathrm{i}h_1$ und $f_2 = g_2 + \mathrm{i}h_2$ seien auf einem Intervall I differenzierbar. Dann ist auch $f_1 \cdot f_2 = (g_1g_2 - h_1h_2) + \mathrm{i}(g_1h_2 + h_1g_2)$ auf I differenzierbar, und es gilt:

$$\begin{aligned}(f_1 \cdot f_2)' &= (g_1'g_2 + g_1g_2' - h_1'h_2 - h_1h_2') + \mathrm{i}(g_1'h_2 + g_1h_2' + h_1'g_2 + h_1g_2') \\ &= [(g_1'g_2 - h_1'h_2) + \mathrm{i}(g_1'h_2 + h_1'g_2)] + [(g_1g_2' - h_1h_2') + \mathrm{i}(g_1h_2' + h_1g_2')] \\ &= f_1' \cdot f_2 + f_1 \cdot f_2'.\end{aligned}$$

Damit erhält man:

2.4 Folgerung. *Sei $f : \mathbb{R} \to \mathbb{C}$ differenzierbar, $c \in \mathbb{C}$ eine Konstante und $f'(x) = c \cdot f(x)$ für alle $x \in \mathbb{R}$.*

Dann ist $f(x) = f(0) \cdot e^{cx}$, für $x \in \mathbb{R}$.

Beweis: Wir setzen $F(x) := f(x) \cdot e^{-cx}$. Dann folgt:

$$F'(x) = (f'(x) - c \cdot f(x)) \cdot e^{-cx} \equiv 0.$$

Also ist $F(x)$ konstant. Wegen $F(0) = f(0)$ ist dann $f(x) = f(0) \cdot e^{cx}$. ∎

Wir haben unsere erste Differentialgleichung gelöst!

2.5 Satz. *Sei $f : I := [a, b] \to \mathbb{R}$ stetig und im Inneren von I differenzierbar. f ist genau dann auf I monoton wachsend (bzw. fallend), wenn $f'(x) \ge 0$ (bzw. $f'(x) \le 0$) für alle $x \in (a, b)$ ist.*

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 den Fall der wachsenden Funktion.

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

2) 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:

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

3) 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. ∎

Beispiele.

1. 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.

2. Die Funktion $\tan : (-\frac{\pi}{2}, +\frac{\pi}{2}) \to \mathbb{R}$ ist differenzierbar, mit $\tan'(x) = 1 + \tan^2(x)$. Da die Ableitung überall positiv ist, ist $\tan(x)$ streng monoton wachsend und damit bijektiv. Der Tangens besitzt eine auf ganz $\mathbb{R}$ definierte Umkehrfunktion, die man mit $\arctan(x)$ (*Arcustangens*) bezeichnet. Offensichtlich ist arctan auf $\mathbb{R}$ differenzierbar, und es gilt:

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

   Dieses Ergebnis sollte man sich für später merken: Durch Differenzieren des Arcustangens erhält man eine rationale Funktion!

3. 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. Die Umkehrfunktion von $\sin : (-\frac{\pi}{2}, \frac{\pi}{2}) \to (-1, 1)$ ist die Funktion

   $$\arcsin : (-1, 1) \to (-\frac{\pi}{2}, \frac{\pi}{2}) \quad (\textit{Arcussinus}).$$

Für die Ableitung gilt:

$$\arcsin'(y) = \frac{1}{\sin'(\arcsin(y))} = \frac{1}{\sqrt{1-y^2}}.$$

Da der Sinus auf allen Intervallen $((k+\frac{1}{2})\pi, (k+\frac{3}{2})\pi)$ streng monoton ist, gibt es auch dazu Umkehrfunktionen $\arcsin_k(y)$. Man spricht von verschiedenen *Zweigen* des Arcussinus. Im Falle $k = -1$ ergibt sich der *Hauptzweig*, den wir oben behandelt haben.

Beim Cosinus kann man analoge Überlegungen anstellen. Wegen $\cos(\frac{\pi}{2} - \arcsin(x)) = \sin(\arcsin(x)) = x$ ist

$$\frac{\pi}{2} - \arcsin(x) = \arccos(x).$$

4. Die Funktionen $\sinh(x) = \dfrac{1}{2}(e^x - e^{-x})$ und $\cosh(x) = \dfrac{1}{2}(e^x + e^{-x})$ 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 $y = \sinh(x) = \dfrac{1}{2}(e^x - e^{-x})$ liefert eine quadratische Gleichung für e^x,

$$2y = \frac{(e^x)^2 - 1}{e^x}, \quad \text{also} \quad (e^x)^2 - 2y \cdot e^x - 1 = 0,$$

und damit

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

Das positive Vorzeichen vor der Wurzel muß gewählt werden, weil $e^x > 0$ ist.

Daraus folgt:

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

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

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

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

Wir können den Mittelwertsatz noch etwas verallgemeinern:

2.6 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

2.7 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 aus dem Satz von Cauchy durch $g(b) - g(a)$ teilen. ■

Eine anschauliche Deutung liefert der folgende Satz, der ebenfalls äquivalent zum Mittelwertsatz ist:

2.8 Satz. *Sei $F : [a, b] \to \mathbb{R}^2$ stetig und auf (a, b) differenzierbar. Ist $F(a) \neq F(b)$, so gibt es ein $c \in (a, b)$, so daß $F'(c)$ parallel zu dem Vektor $F(b) - F(a)$ ist.*

BEWEIS: Wir schreiben $F(t) = (f(t), g(t))$ und können annehmen, daß $g(a) \neq g(b)$ ist. Dann gibt es nach dem Satz von Cauchy ein $c \in (a, b)$ mit

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

Das entspricht der Behauptung. ■

Als Anwendung aus dem 2. Mittelwertsatz ergibt sich:

2.9 Satz (1. Regel von de l'Hospital). *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 $c \in I$ und $f(c) = g(c) = 0$.

Wenn $\lim\limits_{x \to c} \dfrac{f'(x)}{g'(x)}$ existiert, dann existiert auch $\lim\limits_{x \to c} \dfrac{f(x)}{g(x)}$, und die beiden Grenzwerte sind gleich.

Beweis: Wir arbeiten mit einseitigen Grenzwerten. Der Satz gilt dann dementsprechend auch etwas allgemeiner.

Es sei (x_ν) eine Folge von Zahlen mit $c < x_\nu < b$ und $\lim\limits_{\nu \to \infty} x_\nu = c$ (nicht notwendig monoton). Nach dem 2. Mittelwertsatz gibt es Zahlen c_ν mit $c < c_\nu < x_\nu$ und

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

Da auch $\lim\limits_{\nu \to \infty} c_\nu = c$ ist, strebt der letzte Quotient nach Voraussetzung gegen $\dfrac{f'(c)}{g'(c)}$. Aber das bedeutet, daß

$$\lim_{x \to c+} \frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)}$$

ist, und analog schließt man für den linksseitigen Grenzwert. ■

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

Beispiele.

1. Sei $f(x) := \sin x$ und $g(x) := x$.

   Da $f(0) = g(0) = 0$ ist und $\dfrac{f'(x)}{g'(x)} = \dfrac{\cos x}{1}$ für $x \to 0$ gegen 1 strebt, ist auch

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

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

   $$\frac{f'(x)}{g'(x)} = \frac{1}{(x-1)(1 - \sin x)} \to -1, \text{ für } x \to 0.$$

Aber man darf l'Hospital gar nicht anwenden, denn es ist zwar $f(0) = 0$, aber $g(0) = 1$.

Tatsächlich ist $\displaystyle\lim_{x \to 0} \frac{\ln(1-x)}{x + \cos x} = 0$.

2.10 Satz (2. Regel von de l'Hospital). *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 $\displaystyle\lim_{x \to a+} g(x) = +\infty$.

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.

Die gleiche Aussage gilt für die Annäherung an a von links.

BEWEIS: Die Schlußweise ist etwas komplizierter als bei der 1. Regel von de l'Hospital.

Es sei ein $\varepsilon > 0$ vorgegeben. Dann gibt es ein $\delta > 0$, so daß gilt:

$$\left|\frac{f'(x)}{g'(x)} - c\right| < \varepsilon \quad \text{für } a < x < a + \delta.$$

Wegen $\displaystyle\lim_{x \to a+} g(x) = +\infty$ kann man annehmen, daß $g(x) > 0$ für $a < x < a + \delta$ ist

Sei jetzt ein x_0 mit $a < x_0 < a + \delta$ fest gewählt. Zu jedem x mit $a < x < x_0$ gibt es nach dem 2. Mittelwertsatz ein $\xi = \xi(x)$ mit $x < \xi < x_0$ und

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

Daraus folgt:

$$c - \varepsilon < \frac{f(x) - f(x_0)}{g(x) - g(x_0)} < c + \varepsilon \quad \text{für } a < x < x_0.$$

Also existieren Konstanten $k_1, k_2 \in \mathbb{R}$ mit

$$(c - \varepsilon) \cdot g(x) + k_1 < f(x) < k_2 + (c + \varepsilon) \cdot g(x), \quad \text{für } a < x < x_0,$$

und das ergibt die Ungleichungs-Kette

$$(c - \varepsilon) + \frac{k_1}{g(x)} < \frac{f(x)}{g(x)} < \frac{k_2}{g(x)} + (c + \varepsilon).$$

Es gibt ein δ_0 mit $0 < \delta_0 < x_0 - a$ und

$$-\varepsilon < \frac{k_1}{g(x)} < \frac{k_2}{g(x)} < \varepsilon \quad \text{für } a < x < a + \delta_0.$$

So erhalten wir schließlich:

$$c - 2\varepsilon < \frac{f(x)}{g(x)} < c + 2\varepsilon \qquad \text{für } a < x < a + \delta_0.$$

Das zeigt, daß $\lim\limits_{x \to a+} \dfrac{f(x)}{g(x)} = c$ ist. ■

Beispiele.

1. Wir wollen $\lim\limits_{x \to 0+} x \cdot \ln(x)$ berechnen.

   Es ist $\quad x \cdot \ln(x) = \dfrac{\ln(x)}{x^{-1}} \quad$ und $\quad \lim\limits_{x \to 0+} \dfrac{\ln'(x)}{(x^{-1})'} = \lim\limits_{x \to 0+} \dfrac{x^{-1}}{-x^{-2}} = \lim\limits_{x \to 0+} (-x) = 0,$

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

2. Sei $p(x)$ ein Polynom. Ist $\operatorname{grad}(p) = k$, so ist $p^{(k)}$ eine Konstante $c \neq 0$, also

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

   Mehrfache Anwendung von l'Hospital ergibt daher

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

   Die Exponentialfunktion wächst stärker als jedes Polynom.

3. 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.

Eine wichtige Folgerung aus dem Mittelwertsatz ist der

2.11 Schrankensatz. *Sei $I \subset \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ differenzierbar und $m \leq f'(x) \leq M$ auf I. Dann ist*

$$m \leq \frac{f(y) - f(x)}{y - x} \leq M \qquad \textit{für } x, y \in I \textit{ und } x < y.$$

BEWEIS: Es ist $(f - m \cdot \mathrm{id})'(x) = f'(x) - m \geq 0$ auf I, also $f - m \cdot \mathrm{id}$ monoton wachsend.

Sind nun $x, y \in I$, mit $x < y$, so ist $f(x) - mx \leq f(y) - my$, also $m(y - x) \leq f(y) - f(x)$.

Analog folgt die andere Ungleichung. ■

2.12 Folgerung. *Ist $f : I \to \mathbb{R}$ differenzierbar und $|f'(x)| \le C$ auf I, so ist*

$$|f(x) - f(y)| \le C \cdot |x - y| \quad \text{für } x, y \in I.$$

BEWEIS: Es ist $-C \le f'(x) \le C$ für $x \in I$, also auch

$$-C \le \frac{f(y) - f(x)}{y - x} \le C, \text{ für } x < y\,.$$

Daraus folgt: $|f(y) - f(x)| \le C \cdot |y - x|$. Die Formel bleibt auch gültig für $x > y$. ■

Wir werden den Schrankensatz später in der folgenden stark verallgemeinerten Form brauchen:

2.13 Verallgemeinerter Schrankensatz. *Sei $f : I \to \mathbb{R}$ stetig, $Q \subset I$ eine (höchstens) abzählbare Teilmenge und f auf $I \setminus Q$ differenzierbar. Es gebe eine Konstante $C \ge 0$, so daß $|f'(x)| \le C$ in allen Punkten gilt, in denen $f'(x)$ definiert ist.*

Dann ist $|f(x) - f(y)| \le C \cdot |x - y|$ für alle $x, y \in I$.

Man nennt f in einem solchen Fall auch *Lipschitz-stetig*.

BEWEIS: Es sei $x_1 \in I$ ein beliebiger Punkt. Für einen beliebigen Parameter $\varepsilon \ge 0$ definieren wir die Funktion $F_\varepsilon : I \to \mathbb{R}$ durch

$$F_\varepsilon(x) := |f(x) - f(x_1)| - (C + \varepsilon) \cdot (x - x_1).$$

Offensichtlich ist F_ε stetig und $F_\varepsilon(x_1) = 0$. Wir wollen zeigen: Ist $x_2 \in I$, $x_2 > x_1$, so ist $F_\varepsilon(x_2) \le 0$ für jedes ε. Läßt man dann nämlich, bei festgehaltenem x_2, den Parameter ε gegen Null gehen, so erhält man die Aussage des Satzes.

Wir führen einen Beweis durch Widerspruch. Und zwar nehmen wir an, es gebe ein $\varepsilon_0 > 0$ mit $F_{\varepsilon_0}(x_2) > 0$. Weil $F_{\varepsilon_0}(Q)$ höchstens abzählbar ist, gibt es ein y mit

$$0 = F_{\varepsilon_0}(x_1) < y < F_{\varepsilon_0}(x_2) \quad \text{und} \quad y \notin F_{\varepsilon_0}(Q).$$

Nach dem Zwischenwertsatz gibt es ein x mit $x_1 < x < x_2$ und $F_{\varepsilon_0}(x) = y$. Wir setzen

$$x_0 := \sup\{x \in (x_1, x_2) \,:\, F_{\varepsilon_0}(x) = y\}.$$

Dann ist $F_{\varepsilon_0}(x_0) = y$ und $F_{\varepsilon_0}(x) > y$ für $x_0 < x \le x_2$, also

$$\varphi(x) := \frac{F_{\varepsilon_0}(x) - F_{\varepsilon_0}(x_0)}{x - x_0} > 0 \quad \text{für } x_0 < x \le x_2.$$

Setzen wir für F_{ε_0} die Definition dieser Funktion ein, so erhalten wir:

$$\begin{aligned}\varphi(x) &= \frac{|f(x)-f(x_1)| - |f(x_0)-f(x_1)| - (C+\varepsilon_0)\cdot((x-x_1)-(x_0-x_1))}{x-x_0} \\ &= \frac{|f(x)-f(x_1)| - |f(x_0)-f(x_1)|}{x-x_0} - (C+\varepsilon_0).\end{aligned}$$

Wegen $|f(x)-f(x_1)| \le |f(x)-f(x_0)| + |f(x_0)-f(x_1)|$ folgt:

$$\varphi(x) \le \frac{|f(x)-f(x_0)|}{x-x_0} - (C+\varepsilon_0), \qquad \text{für } x_0 < x \le x_2.$$

Da y nicht in $F_{\varepsilon_0}(Q)$ liegt, kann x_0 auch nicht in Q liegen. Also ist f in x_0 differenzierbar und $|f'(x_0)| \le C$. Das bedeutet, daß es ein $\delta > 0$ gibt, so daß gilt:

$$\left|\frac{f(x)-f(x_0)}{x-x_0}\right| < C+\varepsilon_0, \text{ für } x_0 < x < x_0+\delta.$$

Dann ist $\varphi(x) < 0$ für $x_0 < x < x_0+\delta$. Das ist ein Widerspruch! ■

> Eine stetige Funktion $f : I \to \mathbb{R}$ heißt *fast überall differenzierbar*, falls es eine (höchstens) abzählbare Teilmenge $Q \subset I$ gibt, so daß $f|_{I\setminus Q}$ differenzierbar ist. Ist $f'(x) = g(x)$ für $x \in I \setminus Q$, so sagt man, es ist fast überall $f' = g$

2.14 Folgerung. *Ist $f : I \to \mathbb{R}$ stetig und fast überall differenzierbar, und ist außerdem fast überall $f' = 0$, so ist f konstant.*

Beweis: Wir setzen $C = 0$ im verallgemeinerten Schrankensatz und erhalten so, daß $|f(x) - f(y)| = 0$ ist, für $x, y \in I$. Das bedeutet, daß f konstant ist. ■

Kann man die Ableitung einer fast überall differenzierbaren Funktion f in einem Ausnahmepunkt stetig fortsetzen, so ist f dort schon differenzierbar:

2.15 Satz. *Sei $f : I \to \mathbb{R}$ stetig und außerhalb einer höchstens abzählbaren Teilmenge $Q \subset I$ differenzierbar, $x_0 \in I$ ein beliebiger Punkt.*

Wenn $\lim\limits_{\substack{x\to x_0\\ x\in I\setminus Q}} f'(x) =: c$ existiert, dann ist f in x_0 differenzierbar und $f'(x_0) = c$.

Beweis: Sei $\varepsilon > 0$ vorgegeben. Dann gibt es ein $\delta > 0$, so daß für $x \in U_\delta(x_0) \setminus Q$ gilt:

$$c - \varepsilon \le f'(x) \le c + \varepsilon,$$

also $|(f - c\cdot \mathrm{id})'(x)| \le \varepsilon$ fast überall auf $U_\delta(x_0)$. Aus dem verallgemeinerten Schrankensatz folgt nun:

$$|f(x) - f(x_0) - c\cdot(x-x_0)| \le \varepsilon \cdot |x-x_0|, \qquad \text{für } x \in U_\delta(x_0),$$

also

$$c - \varepsilon \le \frac{f(x)-f(x_0)}{x-x_0} \le c + \varepsilon.$$

Das bedeutet, daß $\lim_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0} = c$ ist. ■

2.16 Folgerung. *Sei $f : I \to \mathbb{R}$ stetig, x_0 ein innerer Punkt von I und f in jedem Punkt $x \neq x_0$ differenzierbar. Außerdem sei*

$$\lim_{x \to x_0 -} f'(x) = \lim_{x \to x_0 +} f'(x) =: c.$$

Dann ist f in x_0 differenzierbar und $f'(x_0) = c$.

Folgende Frage bleibt offen:

f erfülle die obigen Voraussetzungen, aber es sei

$$\lim_{x \to x_0 -} f'(x) = c_1 \quad \text{und} \quad \lim_{x \to x_0 +} f'(x) = c_2, \text{ mit } c_1 < c_2 \,.$$

Kann f dann trotzdem differenzierbar sein? Wir haben diese Frage in der Vorlesung nicht behandelt, sie soll aber hier der Vollständigkeit halber beantwortet werden. Dazu brauchen wir folgende verblüffende Tatsache:

2.17 Satz von Darboux. *Sei $f : I \to \mathbb{R}$ differenzierbar, $x, y \in I$, $x < y$ und $f'(x) < c < f'(y)$. Dann gibt es ein x_0 mit $x < x_0 < y$ und $f'(x_0) = c$.*

BEWEIS: Sei $g(t) := f(t) - ct$ auf I. Dann ist $g'(x) < 0$ und $g'(y) > 0$. In der Nähe von x ist g daher streng monoton fallend, und in der Nähe von y streng monoton wachsend. Es muß also Punkte x_1, y_1 mit $x < x_1 < y_1 < y$ geben, so daß $g(x_1) < g(x)$ und $g(y_1) < g(y)$ ist.

Als stetige Funktion muß g auf $[x, y]$ ein Minimum annehmen, und aus den obigen Betrachtungen folgt, daß dies in einem Punkt $x_0 \in (x, y)$ geschehen muß. Dann ist aber $g'(x_0) = 0$, also $f'(x_0) = c$. ■

Die Ableitung einer differenzierbaren Funktion erfüllt also den Zwischenwertsatz. Soll das heißen, daß die Ableitung einer differenzierbaren Funktion immer stetig ist? Mitnichten! Allerdings kann die Ableitung keine Sprungstelle haben:

Sei etwa $f'(x_0-) = \lim_{x \to x_0 -} f'(x) =: c_1 < c_2 := \lim_{x \to x_0 +} f'(x) = f'(x_0+)$. Dann gibt es ein c mit $c_1 < c < c_2$, ein $\varepsilon > 0$ und ein $\delta > 0$, so daß $f'(x) \notin [c - \varepsilon, c + \varepsilon]$ für $|x - x_0| \leq \delta$ und $x \neq x_0$. Wäre f auch in x_0 differenzierbar, so müßte $f'(x)$ zwischen $x_0 - \delta$ und $x_0 + \delta$ jeden Wert aus $[c - \varepsilon, c + \varepsilon]$ annehmen. Das haben wir aber gerade ausgeschlossen.

§ 3 Differentialgleichungen

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 $\mathcal{C}^k(I)$.

Bemerkung. $\mathcal{C}^k(I)$ ist ein $\mathbb{C}$-Vektorraum, und $D : \mathcal{C}^k(I) \to \mathcal{C}^{k-1}(I)$ mit

$$D[f] := f'$$

ist eine lineare Abbildung (auch „linearer Operator“ genannt). $\text{Ker}(D)$ besteht aus den konstanten Funktionen.

Wir definieren $D^q : \mathcal{C}^k(I) \to \mathcal{C}^{k-q}(I)$ durch $D^0 := \text{id}$ und

$$D^q := \underbrace{D \circ D \circ \ldots \circ D}_{q\text{-mal}}, \qquad \text{für } 1 \leq q \leq k.$$

Dann ist $D^q[f] = f^{(q)}$ ebenfalls $\mathbb{C}$-linear. Offensichtlich ist

$$D^p \circ D^q = D^q \circ D^p = D^{p+q}.$$

Ist nun $n \leq k$ und $p(x) = \sum_{\nu=0}^{n} a_\nu x^\nu$ ein Polynom vom Grad n mit reellen oder komplexen Koeffizienten, so setzt man

$$p(D) := \sum_{\nu=0}^{n} a_\nu D^\nu = a_0 \cdot \text{id} + a_1 \cdot D + \cdots + a_n \cdot D^n.$$

Man nennt $p(D)$ einen *Differentialoperator* mit konstanten Koeffizienten. Offensichtlich ist $p(D) : \mathcal{C}^k(I) \to \mathcal{C}^{k-n}(I)$ eine $\mathbb{C}$-lineare Abbildung, und es gilt:

$$p(D)[f] = a_0 \cdot f + a_1 \cdot f' + \cdots + a_n \cdot f^{(n)}.$$

Das Polynom $p(x)$ nennt man das *charakteristische Polynom* von $p(D)$.

3.1 Satz. *Sind p_1, p_2 zwei Polynome, so ist $(p_1 \cdot p_2)(D) = p_1(D) \circ p_2(D) = p_2(D) \circ p_1(D)$.*

BEWEIS: Es ist $(x - c) \cdot (a_i x^i) = a_i x^{i+1} - a_i c x^i$ und

$$(D - c \cdot \text{id}) \circ (a_i \cdot D^i) = a_i \cdot D^{i+1} - (a_i c) \cdot D^i = (a_i \cdot D^i) \circ (D - c \cdot \text{id}),$$

wegen der Linearität von D. Daraus folgt die Behauptung für $p_1(x) = x - c$ und $p_2(x) = a_0 + a_1 x + \cdots + a_n x^n$. Da sich über $\mathbb{C}$ jedes Polynom in Linearfaktoren zerlegen läßt, erhält man per Induktion den Satz für allgemeines p_1. ∎

Der Beweis funktioniert übrigens nur bei Differentialoperatoren mit **konstanten** Koeffizienten.

Beispiele.

1. Sei $p(x)$ ein beliebiges Polynom und $\lambda \in \mathbb{C}$. Es ist $D[e^{\lambda t}] = \lambda \cdot e^{\lambda t}$ und allgemeiner $D^q[e^{\lambda t}] = \lambda^q \cdot e^{\lambda t}$. Daraus ergibt sich:

$$p(D)[e^{\lambda t}] = p(\lambda) \cdot e^{\lambda t}.$$

2. Ist $f \in \mathcal{C}^k(I)$ beliebig, so ist

$$(D - \lambda \cdot \mathrm{id})[f(t)e^{\lambda t}] = f'(t)e^{\lambda t} + \lambda \cdot f(t)e^{\lambda t} - \lambda \cdot f(t)e^{\lambda t} = f'(t)e^{\lambda t},$$

und allgemeiner $\quad (D - \lambda \cdot \mathrm{id})^q[f(t)e^{\lambda t}] = f^{(q)}(t)e^{\lambda t}.$

Definition. Eine *lineare Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten* ist eine Gleichung der Gestalt

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = q(x),$$

mit reellen (oder komplexen) Koeffizienten a_i und einer stetigen Funktion q. Unter dem *charakteristischen Polynom* dieser Differentialgleichung versteht man das Polynom $p(x) := x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$.

Eine *Lösung der Differentialgleichung* über einem (offenen) Intervall I ist eine Funktion $f \in \mathcal{C}^n(I)$ mit

$$p(D)[f] = q.$$

3.2 Satz (von der Eindeutigkeit der Lösung). *$f_1, f_2 \in \mathcal{C}^n(I)$ seien Lösungen einer DGL*

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = q(x)$$

auf dem Intervall I, und es sei $f_1^{(k)}(t_0) = f_2^{(k)}(t_0)$ für $k = 0, 1, 2, \ldots, n-1$ und ein spezielles $t_0 \in I$.

Dann ist $f_1 = f_2$.

BEWEIS: Es genügt, für $f = f_1 - f_2$ zu zeigen:

Ist $p(D)[f] = 0$ und $f^{(k)}(t_0) = 0$ für $k = 0, 1, \ldots, n-1$, so ist $f = 0$.

Nach dem Fundamentalsatz der Algebra gibt es komplexe Zahlen $c_1, \ldots, c_n$, so daß $p(x) = (x - c_1)(x - c_2)\cdots(x - c_n)$ ist. Wir setzen

$$p_0(D) := \mathrm{id} \quad \text{und } p_k(D) := (D - c_{n-k+1}\mathrm{id}) \circ \ldots \circ (D - c_n\mathrm{id}), \text{ für } k = 1, \ldots, n.$$

Dann ist

$$\begin{aligned} p_1(D) &= (D - c_n\mathrm{id}), \\ p_2(D) &= (D - c_{n-1}\mathrm{id}) \circ (D - c_n\mathrm{id}) \\ \text{und} \quad p_n(D) &= p(D). \end{aligned}$$

Weil $f_k := p_k(D)[f]$ eine Linearkombination von $f, f', \ldots, f^{(k)}$ ist, folgt:

$$f_k(t_0) = 0, \text{ für } k = 0, 1, \ldots, n-1.$$

Außerdem ist

$$f_k = f'_{k-1} - c_{n-k+1} f_{k-1}, \text{ für } k = 1, \ldots, n.$$

Es ist $f_n = p(D)[f] = 0$, und wir wissen schon, daß die Gleichung $0 = f'_{n-1} - c_1 f_{n-1}$ die (eindeutig bestimmte) Lösung $f_{n-1}(t) = f_{n-1}(0) \cdot e^{c_1 t}$ hat. Weil $f_{n-1}(t_0) = 0$ ist, muß $f_{n-1} = 0$ sein. Analog folgt, daß $f_{n-2} = f_{n-3} = \ldots = f_1 = 0$ ist. Aus der Gleichung $0 = f'_0 - c_n f_0$ und der Anfangsbedingung $f_0(t_0) = 0$ ergibt sich schließlich, daß $f = f_0 = 0$ ist. ∎

3.3 Satz. *Ist $p(x)$ ein normiertes Polynom n-ten Grades und*

$$L := p(D) : \mathcal{C}^n(I) \to \mathcal{C}^0(I)$$

der zugehörige Differentialoperator über einem (offenen) Intervall I, so ist

$$\dim_{\mathbb{C}} \operatorname{Ker}(L) \le n,$$

d.h., es gibt höchstens n über $\mathbb{C}$ linear unabhängige Lösungen der „homogenen Differentialgleichung“ $L[f] = 0$.

Beweis: Sei $t_0 \in I$ beliebig. Die Abbildung $A : \operatorname{Ker}(L) \to \mathbb{C}^n$ mit

$$A(f) := (f(t_0), f'(t_0), \ldots, f^{(n-1)}(t_0)),$$

die jeder Lösungs-Funktion einen vollständigen Satz von „Anfangsbedingungen“ im Punkt t_0 zuordnet, ist linear, und nach dem Satz über die Eindeutigkeit der Lösungen ist sie auch injektiv. Daraus folgt, daß $A : \operatorname{Ker}(L) \to \operatorname{Im}(A)$ ein Isomorphismus ist. Wegen $\operatorname{Im}(A) \subset \mathbb{C}^n$ folgt $\dim_{\mathbb{C}} \operatorname{Ker}(L) \le n$. ∎

Wir werden sehen, daß sogar die Gleichheit gilt: $\dim_{\mathbb{C}}(\operatorname{Ker}(L)) = n$. Um das zu zeigen, reicht es jetzt aus, n unabhängige Lösungen zu konstruieren.

3.4 Hilfssatz. *Die Zahlen $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ seien paarweise verschieden, und $f_1(x)$, $\ldots$, $f_r(x)$ seien Polynome mit*

$$f_1(x) \cdot e^{\lambda_1 x} + \cdots + f_r(x) \cdot e^{\lambda_r x} \equiv 0.$$

Dann ist $f_1 = \ldots = f_r = 0$.

Beweis: Wir führen Induktion nach r.

Ist $f_1(x) \cdot e^{\lambda_1 x} \equiv 0$, so muß $f_1 = 0$ sein, weil $e^{\lambda_1 x} > 0$ für jedes x ist.

Sei nun $r \ge 2$, und die Behauptung für $r - 1$ bewiesen. Es sei

$$f_1(x) \cdot e^{\lambda_1 x} + \cdots + f_r(x) \cdot e^{\lambda_r x} \equiv 0,$$

also

$$f_1(x) + f_2(x) \cdot e^{(\lambda_2 - \lambda_1)x} + \cdots + f_r(x) \cdot e^{(\lambda_r - \lambda_1)x} \equiv 0.$$

Ist $f_1 = 0$, so sind wir fertig. Ist $f_1 \neq 0$ und $d = \text{grad}(f_1)$, so differenzieren wir die Gleichung $(d+1)$-mal. Dann erhalten wir Polynome $F_2(x), \ldots, F_r(x)$ mit $\text{grad}(F_i) = \text{grad}(f_i)$ für $i = 2, \ldots, r$, so daß gilt:

$$F_2(x) \cdot e^{(\lambda_2 - \lambda_1)x} + \cdots + F_r(x) \cdot e^{(\lambda_r - \lambda_1)x} \equiv 0.$$

Nach Induktionsvoraussetzung muß dann $F_2 = \ldots = F_r = 0$ sein, also auch $f_2 = \ldots = f_r = 0$ (wegen der Gleichheit der Grade). Dann ist selbstverständlich auch $f_1 = 0$. ∎

Jetzt können wir ein „Fundamentalsystem“ von Lösungen der DGL $p(D)[f] = 0$ angeben, d.h. eine Basis des Lösungsraumes.

3.5 Satz über das Fundamentalsystem der DGL $p(D)[f] = 0$**.** *Sei $p(x)$ das charakteristische Polynom der homogenen DGL*

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 = 0.$$

Sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Nullstellen von $p(x)$ in $\mathbb{C}$, mit Vielfachheiten $k_1, \ldots, k_r$, so bilden die Funktionen

$$x^\nu \cdot e^{\lambda_i x}, \text{ mit } i = 1, \ldots, r \text{ und } \nu = 0, \ldots, k_i - 1,$$

ein Fundamentalsystem von Lösungen der Differentialgleichung.

Insbesondere hat der Lösungsraum die Dimension $k_1 + \cdots + k_r = n$ über $\mathbb{C}$.

BEWEIS: 1) Wir zeigen zunächst, daß die angegebenen Funktionen tatsächlich Lösungen sind. Ist λ eine Nullstelle von $p(x)$ mit der Vielfachheit k, so gibt es ein Polynom $q(x)$ vom Grad $n - k$, so daß $p(x) = q(x) \cdot (x - \lambda)^k$ ist. Wegen der Vertauschbarkeit der Differentialoperatoren ist dann auch

$$p(D) = q(D) \circ (D - \lambda)^k,$$

und es folgt:

$$\begin{aligned} p(D)[t^\nu e^{\lambda t}] &= q(D) \circ (D - \lambda)^k [t^\nu \cdot e^{\lambda t}] \\ &= q(D)[(t^\nu)^{(k)} \cdot e^{\lambda t}] = 0. \end{aligned}$$

2) Nun müssen wir noch sehen, daß die Lösungen linear unabhängig sind. Wir müssen mit den beiden Indizes ν und i arbeiten:

Sei $\displaystyle\sum_{i=1}^{r} \sum_{\nu=0}^{k_i - 1} c_{i,\nu} x^\nu e^{\lambda_i x} = 0$. Dann setzen wir $f_i(x) := \displaystyle\sum_{\nu=0}^{k_i - 1} c_{i,\nu} x^\nu$, für $i = 1, \ldots, r$. Das sind alles Polynome, und wir haben

$$f_1(x) \cdot e^{\lambda_1 x} + \cdots + f_r(x) \cdot e^{\lambda_r x} \equiv 0.$$

Also muß $f_1 = \ldots = f_r = 0$ sein. Aber ein Polynom ist nur dann das Nullpolynom, wenn alle seine Koeffizienten verschwinden. Damit ist $c_{i,\nu} = 0$ für alle i und ν, und die Lösungsfunktionen sind linear unabhängig. ∎

Bemerkung. In der Praxis ist man meist an reellen Lösungen interessiert. Sind alle Nullstellen $\lambda_1, \ldots, \lambda_r$ des charakteristischen Polynoms reell, so ist alles in Ordnung. Hat $p(x)$ reelle Koeffizienten und ist $\lambda = \alpha + \mathrm{i}\beta$ eine komplexe Nullstelle, so ist auch $\overline{\lambda}$ eine Nullstelle, und es gilt:

$$\begin{aligned} e^{\lambda x} &= e^{\alpha x}(\cos(\beta x) + \mathrm{i}\sin(\beta x)), \\ e^{\overline{\lambda} x} &= e^{\alpha x}(\cos(\beta x) - \mathrm{i}\sin(\beta x)). \end{aligned}$$

Dann können wir die komplexen Lösungen $x^\nu e^{\lambda x}$ und $x^\nu e^{\overline{\lambda} x}$ durch die linear unabhängigen reellen Lösungen

$$\begin{aligned} \mathrm{Re}(x^\nu e^{\lambda x}) &= x^\nu e^{\alpha x}\cos(\beta x) \\ \text{und } \mathrm{Im}(x^\nu e^{\lambda x}) &= x^\nu e^{\alpha x}\sin(\beta x) \end{aligned}$$

ersetzen.

Beispiel.

Wir betrachten die homogene lineare DGL 2. Ordnung

$$y'' + 2ay' + by = 0,$$

mit reellen Koeffizienten und der Anfangsbedingung $y(0) = 1$ und $y'(0) = 0$. Das charakteristische Polynom ist $p(x) = x^2 + 2ax + b$. Die Nullstellen der Gleichung $p(x) = 0$ sind gegeben durch

$$x = \frac{-2a \pm \sqrt{4a^2 - 4b}}{2} = -a \pm \sqrt{\Delta},$$

mit $\Delta := a^2 - b$.

1. $\Delta > 0$. (starke Dämpfung)

   Dann ist $p(x) = (x - \lambda_1)(x - \lambda_2)$, mit $\{\lambda_1, \lambda_2\} = \{-a \pm \sqrt{\Delta}\}$. Die allgemeine Lösung hat die Gestalt

   $$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}, \quad c_1, c_2 \in \mathbb{R}.$$

   Dann ist $y'(x) = c_1\lambda_1 e^{\lambda_1 x} + c_2\lambda_2 e^{\lambda_2 x}$, und das Einsetzen der Anfangsbedingungen ergibt

   $$c_1 + c_2 = 1 \text{ und } c_1\lambda_1 + c_2\lambda_2 = 0.$$

   Also ist $c_2 = 1 - c_1$ und $c_1\lambda_1 + (1 - c_1)\lambda_2 = 0$ Die gesuchte Lösung hat deshalb die Gestalt

   $$y(x) = \frac{1}{\lambda_2 - \lambda_1}(\lambda_2 e^{\lambda_1 x} - \lambda_1 e^{\lambda_2 x}).$$

2. $\Delta = 0$. (kritische Dämpfung)

   Dann besitzt $p(x)$ die zweifache reelle Nullstelle $-a$. Also hat die allgemeine Lösung die Gestalt

   $$y(x) = (c_1 + c_2 x)e^{-ax}.$$

   Wegen $y'(x) = (c_2 - a(c_1 + c_2 x))e^{-ax}$ ergeben die Anfangsbedingungen:

   $$c_1 = 1 \text{ und } c_2 - ac_1 = 0, \text{ also } c_2 = a.$$

   Die gesuchte Lösung hat die Gestalt

   $$y(x) = (1 + ax)e^{-ax}.$$

3. $\Delta < 0$. (schwache Dämpfung)

   Setzt man $\omega := \sqrt{-\Delta}$ und $\lambda := -a + \mathrm{i}\omega$, so hat $p(x)$ die komplexen Nullstellen λ und $\bar{\lambda}$. In diesem Fall hat die allgemeine Lösung die Gestalt

   $$y(x) = e^{-ax}(c_1 \cos(\omega x) + c_2 \sin(\omega x)).$$

   Dann ist

   $$\begin{aligned} y'(x) &= -ae^{-ax}(c_1 \cos(\omega x) + c_2 \sin(\omega x)) \\ &\quad + e^{-ax}(-\omega c_1 \sin(\omega x) + \omega c_2 \cos(\omega x)) \\ &= e^{-ax}[(c_2\omega - ac_1)\cos(\omega x) - (ac_2 + \omega c_1)\sin(\omega x)]. \end{aligned}$$

   Die Randbedingungen ergeben

   $$c_1 = 1 \text{ und } c_2\omega - ac_1 = 0, \text{ also } c_2 = \frac{a}{\omega}.$$

   Das führt zu der Lösung

   $$y(x) = e^{-ax}(\cos(\omega x) + \frac{a}{\omega}\sin(\omega x)).$$

   Man kann sie umformen zu einer gedämpften Schwingung

   $$y(x) = Ae^{-ax}\sin(\omega x + \varphi).$$

Wie löst man jetzt die „inhomogene Gleichung“ $p(D)[f(x)] = q(x)$?

1. Sind y_1, y_2 zwei Lösungen der inhomogenen Gleichung, so ist $y_1 - y_2$ eine Lösung der homogenen Gleichung.

2. Ist y_0 eine „partikuläre Lösung“ (d.h. eine spezielle Lösung der inhomogenen Gleichung) und y eine Lösung der homogenen Gleichung, so ist $y_0 + y$ wieder eine Lösung der inhomogenen Gleichung.

Das bedeutet, daß die Gesamtheit der Lösungen von $p(D)[y] = q(x)$ einen affinen Raum der Gestalt $y_0 + \text{Ker}(p(D))$ bildet, mit einer beliebigen partikulären Lösung y_0.

Es bleibt also noch das Problem, eine partikuläre Lösung der inhomogenen DGL

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

zu finden.

Hierfür können wir im Augenblick noch kein allgemeines Verfahren angeben. Wir beschränken uns daher auf einen Spezialfall, wo man die partikuläre Lösung durch einen geschickten Ansatz finden kann.

3.6 Satz. *Sei $p(x)$ ein normiertes Polynom (mit reellen Koeffizienten). Die DGL $p(D)[y] = f(x)e^{\lambda x}$ mit einem Polynom $f(x)$ vom Grad s besitzt eine partikuläre Lösung $y_p(x) = g(x)e^{\lambda x}$, mit einem Polynom $g(x)$ vom Grad $s + m$, wobei $m \geq 0$ der höchste Exponent ist, so daß eine Darstellung $p(x) = q(x) \cdot (x - \lambda)^m$ existiert.*

BEWEIS: Wir untersuchen zunächst zwei Spezialfälle.

1) Ist $p(x) = (x - \lambda)^m$, so ist $p(D)[g(x)e^{\lambda x}] = g^{(m)}(x)e^{\lambda x}$. In diesem Falle reicht es, ein Polynom $g(x)$ mit $g^{(m)}(x) = f(x)$ zu finden. Offensichtlich muß g den Grad $s + m$ haben

2) Ist $\mu \neq \lambda$, so ist

$$(D - \mu \cdot \text{id})[g(x)e^{\lambda x}] = (g'(x) + (\lambda - \mu) \cdot g(x)) \cdot e^{\lambda x},$$

wobei das Polynom $g'(x) + (\lambda - \mu) \cdot g(x)$ den gleichen Grad wie $g(x)$ hat. Ist $f(x) = \sum_{i=0}^{s} b_i x^i$, so setzen wir $g(x)$ mit unbestimmten Koeffizienten an:

$$g(x) = \sum_{i=0}^{s} a_i x^i.$$

Zur Abkürzung setzen wir außerdem $c := \lambda - \mu$. Dann ist

$$g'(x) + c \cdot g(x) = \sum_{i=0}^{s-1}((i+1) \cdot a_{i+1} + c \cdot a_i)x^i + c \cdot a_s x^s.$$

Das liefert das folgende lineare Gleichungssystem:

$$\begin{aligned} c \cdot a_s &= b_s, \\ s \cdot a_s + c \cdot a_{s-1} &= b_{s-1}, \\ &\vdots \\ 2 \cdot a_2 + c \cdot a_1 &= b_1, \\ \text{und} \quad a_1 + c \cdot a_0 &= b_0. \end{aligned}$$

Man kann es sehr leicht nach $a_0, a_1, \dots, a_s$ auflösen.

3) Bei zusammengesetztem $p(x)$ wendet man die gewonnenen Ergebnisse mehrfach an. Ist $p(x) = (x - c_1)^{k_1} \cdots (x - c_r)^{k_r}$, so sucht man zunächst ein Polynom $g_1(x)$ mit $(D - c_1\mathrm{id})^{k_1}[g_1(x)e^{\lambda x}] = f(x)e^{\lambda x}$, dann ein Polynom $g_2(x)$ mit $(D - c_2\mathrm{id})^{k_2}[g_2(x)e^{\lambda x}] = g_1(x)e^{\lambda x}$ und so weiter. Schließlich landet man bei einem Polynom $g(x)$ mit $(D - c_r)^{k_r}[g(x)e^{\lambda x}] = g_{r-1}(x)e^{\lambda x}$. Offensichtlich ist g das gesuchte Polynom. ∎

Beispiel.

Das charakteristische Polynom der DGL $y''' + 3y'' + 3y' + y = xe^{-x}$ ist

$$p(x) = x^3 + 3x^2 + 3x + 1 = (x + 1)^3,$$

mit $x = -1$ als dreifacher Nullstelle. Das bedeutet, daß die drei Funktionen

$$f_1(x) = e^{-x}, \quad f_2(x) = xe^{-x} \quad \text{und} \quad f_3(x) = x^2e^{-x}$$

ein Fundamentalsystem bilden.

Die Inhomogenität hat die Form $f(x)e^{\lambda x}$ mit $f(x) = x$ und $\lambda = -1$. Da λ dreifache Nullstelle des charakteristischen Polynoms ist, suchen wir ein Polynom $g(x)$ mit $g^{(3)}(x) = x$. Offensichtlich leistet $g(x) := \frac{1}{24}x^4$ das Verlangte, und wir erhalten für die partikuläre Lösung den Ansatz

$$y_p(x) = \frac{1}{24}x^4e^{-x}.$$

Die Probe zeigt, daß das richtig ist. Die allgemeine Lösung lautet nun

$$y(x) = (c_1 + c_2x + c_3x^2 + \frac{1}{24}x^4)e^{-x}.$$

§ 4 Die Taylor-Entwicklung

4.1 Satz. *Sei $I \subset \mathbb{R}$ ein Intervall und (f_ν) eine Folge von stetigen und fast überall differenzierbaren Funktionen auf I, die in wenigstens einem Punkt von I konvergiert. Weiter sei (g_ν) eine Folge von Funktionen auf I, die auf jedem abgeschlossenen Teilintervall von I gleichmäßig gegen eine auf I definierte Funktion g konvergiert, und es sei fast überall $f'_\nu = g_\nu$.*

Dann ist (f_ν) auch auf jedem abgeschlossenen Teilintervall von I gleichmäßig konvergent, die Grenzfunktion $f = \lim_{\nu\to\infty} f_\nu$ ist fast überall differenzierbar, und es ist fast überall $f' = g$.

Man kann also unter den gegebenen Voraussetzungen Limes und Ableitung miteinander vertauschen.

BEWEIS: Es reicht, ein abgeschlossenes Intervall $J = [a, b] \subset I$ zu betrachten, das einen Punkt c enthält, in dem die Folge (f_ν) konvergiert. Es sei $Q_\nu \subset J$ eine höchstens abzählbare Teilmenge, so daß f_ν auf $J \setminus Q_\nu$ differenzierbar und dort auch $f'_\nu = g_\nu$ ist. Die Menge $Q := \bigcup_\nu Q_\nu$ ist immer noch höchstens abzählbar.

Für $n, m \in \mathbb{N}$ setzen wir $f_{nm} := f_n - f_m$ und $g_{nm} := g_n - g_m$. Nach Voraussetzung ist f_{nm} auf $J \setminus Q$ differenzierbar und dort $f'_{nm} = g_{nm}$. Außerdem gibt es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so daß $\|g_{nm}\| < \varepsilon$ für $n, m \geq n_0$ ist. Die Norm ist die Supremumsnorm auf J, und die Aussage beweist man wie das Cauchy-Kriterium für Zahlenfolgen: Aus der gleichmäßigen Konvergenz der Folge (g_n) auf J erhält man:

$$\|g_{nm}\| = \|(g_n - g) + (g - g_m)\| \leq \|g_n - g\| + \|g_m - g\| < \varepsilon, \text{ für } n, m \geq n_0.$$

Ist $x \in J$, so folgt jetzt aus dem verallgemeinerten Schrankensatz:

$$|f_{nm}(x) - f_{nm}(c)| \leq |x - c| \cdot \varepsilon,$$

also

$$\|f_{nm}\| \leq |f_{nm}(c)| + (b - a) \cdot \varepsilon.$$

Wegen der Konvergenz von $(f_n(c))$ bedeutet das, daß $\|f_{nm}\|$ beliebig klein wird, wenn nur n und m genügend groß gewählt werden. Als stetige Funktionen auf dem abgeschlossenen Intervall J gehören die Funktionen f_n zum Raum $\mathcal{B} = \mathcal{B}(J, \mathbb{R})$ der beschränkten Funktionen auf J, und wir haben gerade gesehen, daß sie eine Cauchyfolge in $\mathcal{B}$ bilden. Da $\mathcal{B}$ vollständig ist, folgt daraus, daß (f_ν) auf J gleichmäßig gegen eine Funktion f konvergiert.

Wir müssen noch zeigen, daß fast überall $f' = g$ ist. Dazu sei $x_0 \in J \setminus Q$ ein beliebiger Punkt und $\varepsilon > 0$ vorgegeben. $n_0 \in \mathbb{N}$ sei wie oben gewählt. Für jedes $x \in J \setminus Q$ liefert der verallgemeinerte Schrankensatz die Abschätzung

$$|f_{nm}(x) - f_{nm}(x_0)| \leq |x - x_0| \cdot \|g_{nm}\|,$$

also

$$\left| \frac{f_n(x) - f_n(x_0)}{x - x_0} - \frac{f_m(x) - f_m(x_0)}{x - x_0} \right| < \varepsilon, \text{ für } n, m \geq n_0,\ x \in I \setminus Q, x \neq x_0.$$

Wir halten ein $m \geq n_0$ fest und lassen n gegen ∞ gehen. Das liefert:

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - \frac{f_m(x) - f_m(x_0)}{x - x_0} \right| \leq \varepsilon, \text{ für } m \geq n_0,\ x \in J \setminus Q,\ x \neq x_0.$$

Es gibt ein $m_0 \geq n_0$, so daß $|g_m(x_0) - g(x_0)| < \varepsilon$ für $m \geq m_0$ ist, und dazu gibt es ein $\delta > 0$, so daß für $x \in J \setminus Q$ und $|x - x_0| < \delta$ gilt:

$$\left| \frac{f_{m_0}(x) - f_{m_0}(x_0)}{x - x_0} - g_{m_0}(x_0) \right| < \varepsilon,$$

weil f_{m_0} in x_0 differenzierbar und $f'_{m_0}(x_0) = g_{m_0}(x_0)$ ist. Jetzt erhalten wir mit Hilfe der Dreiecksungleichung:

$$\begin{aligned}
&\left| \frac{f(x) - f(x_0)}{x - x_0} - g(x_0) \right| \leq \\
&\quad \leq \left| \frac{f(x) - f(x_0)}{x - x_0} - \frac{f_{m_0}(x) - f_{m_0}(x_0)}{x - x_0} \right| + \left| \frac{f_{m_0}(x) - f_{m_0}(x_0)}{x - x_0} - g_{m_0}(x_0) \right| \\
&\qquad + |g_{m_0}(x_0) - g(x_0)| \\
&\quad < 3\varepsilon, \text{ für } |x - x_0| < \delta,\, x \in J \setminus Q.
\end{aligned}$$

Das zeigt, daß f in x_0 differenzierbar und $f'(x_0) = g(x_0)$ ist. ∎

4.2 Folgerung. *Sei (f_ν) eine Folge von differenzierbaren Funktionen auf dem abgeschlossenen Intervall $I = [a, b]$, die in wenigstens einem Punkt $c \in I$ konvergiert. Außerdem sei die Folge der Ableitungen (f'_ν) auf I gleichmäßig konvergent. Dann konvergiert auch (f_ν) auf I gleichmäßig gegen eine differenzierbare Funktion f, und es ist*

$$f' = \lim_{\nu \to \infty} f'_\nu \,.$$

Eine wichtige Anwendung ist der folgende

4.3 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 im Innern des Konvergenzintervalls absolut und auf jedem abgeschlossenen Teilintervall des Konvergenzintervalls gleichmäßig gegen die Ableitung $f'(x)$.

Potenzreihen können also gliedweise differenziert werden!

BEWEIS: Die Funktionenfolge $F_N(x) := \sum_{n=0}^{N} c_n (x - a)^n$ konvergiert auf $(a - R, a + R)$ punktweise (und sogar 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 gleichmäßig gegen eine auf $(a - R, a + R)$ stetige Funktion g

Also ist f sogar stetig differenzierbar, und $f' = g$. ∎

Da die Ableitung einer Potenzreihe wieder eine Potenzreihe ist, kann man das obige Argument wiederholen und erhält:

4.4 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: Wir müssen nur noch die Formel beweisen. Offensichtlich ist

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1) \cdot \ldots \cdot (n-k+1)c_n(x-a)^{n-k}.$$

Setzt man $x = a$ ein, so erhält man:

$$f^{(k)}(a) = k(k-1) \cdot \ldots \cdot (k-k+1) \cdot c_k = k!\, c_k.$$

■

Ist umgekehrt $f : I \to \mathbb{R}$ eine beliebig oft differenzierbare Funktion, so kann man in jedem Punkt $a \in I$ die *Taylorreihe*

$$Tf(x;a) := \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

bilden. Allerdings ist i.a. nicht klar, ob f durch seine Taylorreihe dargestellt wird.

Beispiele.

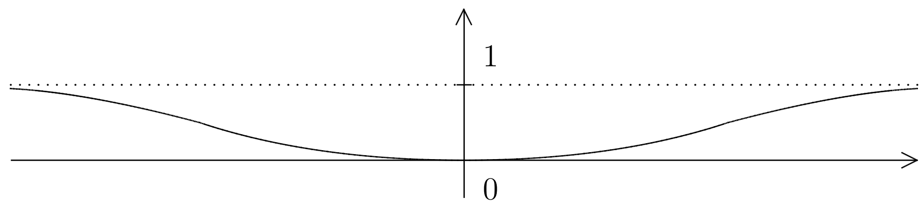
1. $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ sei eine konvergente Potenzreihe, mit Konvergenzradius R. Dann ist $Tf(x;a) = f(x)$ auf $(a-R, a+R)$. Das trifft z.B. auf $\exp(x)$, $\sin(x)$ und $\cos(x)$ zu.

2. 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 $(-1, +1)$.

3. Noch verrückter verhält sich die Funktion

$$f(x) := \begin{cases} \exp\left(-\frac{1}{x^2}\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

Offensichtlich ist f stetig und für $x \neq 0$ beliebig oft differenzierbar. Außerdem ist $x = 0$ die einzige Nullstelle von f.



Behauptung: Es gibt rationale Funktionen $q_k(x)$ (mit 0 als einziger Polstelle), so daß gilt:

$$f^{(k)}(x) = q_k(x) \cdot e^{-1/x^2}, \qquad \text{für } x \neq 0 \text{ und } k \in \mathbb{N}_0.$$

Beweis dafür 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.

Sei nun $I \subset \mathbb{R}$ ein Intervall und $f \in \mathcal{C}^n(I)$. Dann kann man i.a. keine Taylorreihe mehr bilden. Wir wollen jedoch versuchen, f in der Nähe eines fest gewählten Punktes $a \in I$ möglichst gut durch Polynome zu approximieren.

Nach unserer Erfahrung mit den Potenzreihen bieten sich dafür die sogenannten „Taylor-Polynome“ an:

Definition. Sei $f \in \mathcal{C}^n(I)$, $a \in I$ ein fester Punkt. Das Polynom

$$T_n f(x) = T_n f(x; a) := \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

heißt *n-tes Taylorpolynom* von f in a.

Offensichtlich ist $T_1 f(x;a) = f(a) + f'(a)(x-a)$ die lineare Approximation von f in a, und es ist $(T_n f)^{(k)}(a) = f^{(k)}(a)$, für $k = 0, 1, 2, \ldots, n$.

Nun geht es um das Verhalten des *Restgliedes* $R_n(x) := f(x) - T_n f(x)$ in der Nähe von a:

4.5 Satz (von der Taylorentwicklung). *Es sei f auf I n-mal differenzierbar und $R_n(x) := f(x) - T_n f(x)$.*

1. *Es gibt eine Funktion $\eta(x)$ mit $\lim\limits_{x \to a} \eta(x) = 0$, so daß gilt:*

$$R_n(x) = \eta(x) \cdot (x-a)^n.$$

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 „Lagrangeschen Form“ des Restgliedes.

Die Darstellung

$$f(x) = T_n f(x; a) + R_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x-a)^{n+1}$$

nennt man die „Taylorentwicklung“ der Ordnung n von f im Punkte a.

Beweis: Wir beginnen mit dem zweiten Teil. Für $x \neq a$ ist $f(x) = T_n f(x) + \varphi(x) \cdot (x-a)^{n+1}$, wenn man definiert:

$$\varphi(x) := \frac{R_n(x)}{(x-a)^{n+1}}\,.$$

Zähler und Nenner dieses Quotienten sind $(n+1)$-mal differenzierbar, und es ist $R_n(a) = (a-a)^{n+1} = 0$. 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:

$$\varphi(x) = \frac{R_n(x) - R_n(c)}{(x-a)^{n+1}} = \frac{R_n'(c_1)}{(n+1)(c_1-a)^n}\,.$$

Die $(n+1)$-te Ableitung von $(x-a)^{n+1}$ ergibt $(n+1)!$, die $(n+1)$-te Ableitung von $R_n(x)$ 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

$$\begin{aligned}
\varphi(x) &= \frac{R_n'(c_1)}{(n+1)(c_1-a)^n} \\
&= \frac{R_n''(c_2)}{n(n+1)(c_2-a)^{n-1}} \\
&= \cdots \\
&= \frac{R_n^{(n+1)}(c_{n+1})}{(n+1)!} = \frac{f^{(n+1)}(c_{n+1})}{(n+1)!},
\end{aligned}$$

mit geeigneten Punkten $c_i = c_i(x)$ zwischen a und x. Setzt man $c := c_{n+1}$, so erhält man

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Ist f nur n-mal differenzierbar, so betrachten wir die Funktion $g(x) := f(x) - T_n f(x)$. Offensichtlich ist g n-mal differenzierbar, und $g(a) = g'(a) = \ldots = g^{(n)}(a) = 0$. Nach dem ersten Teil des Beweises gibt es zu jedem $x \in I$ ein $c(x)$ zwischen a und x mit

$$g(x) = T_{n-2}g(x) + \frac{g^{(n-1)}(c(x))}{(n-1)!}(x-a)^{n-1} = \eta(x)\cdot(x-a)^n,$$

wobei

$$\eta(x) := \begin{cases} \dfrac{g^{(n-1)}(c(x))}{(n-1)!(x-a)} & \text{für } x \neq a, \\ 0 & \text{für } x = a. \end{cases}$$

Da $g^{(n-1)}$ differenzierbar ist, gibt es eine Funktion Δ mit

$$g^{(n-1)}(c(x)) = g^{(n-1)}(a) + (c(x)-a)\cdot\Delta(c(x)) \quad \text{und} \quad \lim_{y\to a}\Delta(y) = g^{(n)}(a) = 0.$$

Dann ist

$$\eta(x) = \frac{(c(x)-a)}{(n-1)!(x-a)}\cdot\Delta(c(x)),$$

und es 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 *Landauschen Symbols* o abgekürzt:

$$f(x) = \mathrm{o}(g(x)) \qquad (\text{für } x \to a\,).$$

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

$$x \cdot \sin(\frac{1}{x}) = \mathrm{o}(1).$$

Im Falle der Taylorentwicklung kann man nun schreiben: Ist $f \in \mathcal{C}^n(I)$, so ist

$$f(x) = T_n f(x) + \mathrm{o}((x-a)^n), \text{ (für } x \to a).$$

Beispiele.

1. Sei $f(x) = \sin(x)$ und $a = 0$. Es ist

   $$\sin'(x) = \cos(x),\ \sin''(x) = -\sin(x),\ \sin^{(3)}(x) = -\cos(x),\ \sin^{(4)}(x) = \sin(x),$$

   und dann wiederholt sich das wieder. Daraus folgt:

   $$\sin^{(2n)}(0) = 0 \quad \text{und} \quad \sin^{(2n+1)}(0) = (-1)^n.$$

   Das ergibt für die Entwicklung der Ordnung $2n+2$:

   $$\begin{aligned} \sin(x) &= \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} \cdot x^{2k+1} \pm \frac{\cos(c)}{(2n+2)!} x^{2n+3} \\ &= x - \frac{1}{6}x^3 + \frac{1}{120}x^5 \pm \cdots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + \mathrm{o}(x^{2n+2}). \end{aligned}$$

   Weil alle geraden Ableitungen des Sinus bei $x = 0$ verschwinden, gewinnt man eine Ordnung zusätzlich.

2. Analog geht es beim Cosinus. Es ist

   $$\cos^{(2n+1)}(0) = 0 \quad \text{und} \quad \cos^{(2n)}(0) = (-1)^n.$$

   Daraus folgt:

   $$\begin{aligned} \cos(x) &= \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} \cdot x^{2k} + \mathrm{o}(x^{2n+1}) \\ &= 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \frac{1}{720}x^6 \pm \cdots + \frac{(-1)^n}{(2n)!}x^{2n} + \mathrm{o}(x^{2n+1}). \end{aligned}$$

3. Im Falle der Funktion

   $$f(x) := \begin{cases} \exp\left(-\frac{1}{x^2}\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

   ist jedes Taylor-Polynom $T_n f(x; 0) \equiv 0$. Man sagt auch, f ist im Nullpunkt „flach“. Die Taylor-Entwicklung ist dort ziemlich wertlos.

§ 5 Anwendungen

Als Anwendung der Taylor-Entwicklung können wir jetzt das Problem der lokalen Extrema abschließend behandeln:

5.1 Hinreichendes Kriterium für Extremwerte. *Die Funktion f sei in der Nähe von x_0 n-mal stetig differenzierbar. Es sei*

$$\begin{array}{rrcll} & f^{(k)}(x_0) & = & 0 & \textit{für } k = 1, \ldots, n-1 \\ \textit{und} & f^{(n)}(x_0) & \neq & 0. & \end{array}$$

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

$$\begin{array}{rcll} \textit{ein Maximum, falls } f^{(n)}(x_0) & < & 0 & \textit{ist,} \\ \textit{und ein Minimum, falls } f^{(n)}(x_0) & > & 0 & \textit{ist.} \end{array}$$

BEWEIS: Da $f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0$ ist, liefert die Taylor-Entwicklung mit $h := x - x_0$:

$$f(x) = f(x_0 + h) = f(x_0) + \frac{f^{(n)}(c)}{n!} h^n,$$

mit einem geeigneten c zwischen x_0 und x.

Ist $\varepsilon > 0$ klein genug gewählt, so ist $f^{(n)}(x) \neq 0$ für $|x - x_0| < \varepsilon$, und dann hat $f^{(n)}(c)$ das gleiche Vorzeichen wie $f^{(n)}(x_0)$.

Wir betrachten nur den Fall $f^{(n)}(x_0) > 0$, der andere geht analog. Da c von x (und damit von h) abhängt, können wir schreiben:

$$f(x_0 + h) - f(x_0) = \varphi(h) \cdot h^n,$$

mit einer positiven Funktion φ.

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. ∎

Bemerkung. Gelegentlich interessiert man sich für den Fall, daß $f''(x_0) = 0$ und $f'''(x_0) \neq 0$ ist. Dann besitzt f in x_0 kein lokales Extremum, und es muß ein $\varepsilon > 0$ geben, so daß $f''(x) \neq 0$ für $|x - x_0| < \varepsilon$ und $x \neq x_0$ ist, denn sonst gäbe es eine Folge von verschwindenden Differenzenquotienten $\Delta(x_\nu)$ mit $x_\nu \to x_0$, und es wäre $f'''(x_0) = 0$. Da $f''(x_0) = 0$ ist und f'' in x_0 kein Extremum besitzt, muß f'' bei x_0 sein Vorzeichen wechseln, und demnach f' bei x_0 sein Monotonieverhalten. Der Graph einer Funktion mit monoton wachsender Ableitung ist nach links gekrümmt, der einer Funktion mit monoton fallender Ableitung ist nach rechts gekrümmt.

Wenn f von einer Linkskrümmung in eine Rechtskrümmung übergeht, oder umgekehrt, so sagt man, f hat bei x_0 einen *Wendepunkt*. Wenn zusätzlich $f'(x_0) = 0$ ist, sprechen wir von einem *Sattelpunkt*.

Ist etwa $f(x) := x^3$, so ist $f''(x) = 6x$, also $f''(0) = 0$. Für $x < 0$ ist $f''(x) < 0$, und für $x > 0$ ist $f''(x) > 0$. Also wechselt f von einer Rechtskrümmung zu einer Linkskrümmung und hat damit in 0 einen Wendepunkt.

Tatsächlich ist $f'''(0) = 6 \neq 0$.

Als nächstes wollen wir weitere Taylor-Entwicklungen betrachten:

1. Die Funktion $f(x) = \ln(x)$ ist nur für $x > 0$ definiert. Hier nehmen wir $a = 1$ als Entwicklungspunkt. Es ist

$$\ln(1) = 0,\ \ln'(x) = \frac{1}{x},\ \ln''(x) = -\frac{1}{x^2},\ \ln^{(3)}(x) = \frac{2}{x^3},\ \ln^{(4)}(x) = -\frac{6}{x^4},$$

und allgemein

$$\ln^{(k)}(x) = (-1)^{k+1} \cdot \frac{(k-1)!}{x^k}, \text{ für } k \geq 1.$$

Das ergibt, daß $\dfrac{\ln^{(k)}(1)}{k!} = \dfrac{(-1)^{k+1}}{k}$ ist, für $k \geq 1$. Also ist

$$L(x) := \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(x-1)^k}{k}$$

die Taylorreihe von $\ln(x)$ an der Stelle $x = 1$. Im Konvergenzintervall $(0, 2)$ konvergiert sie gegen eine beliebig oft differenzierbare Funktion L mit

$$L'(x) = \sum_{k=1}^{\infty} (-1)^{k+1} (x-1)^{k-1} = \sum_{k=0}^{\infty} (1-x)^k = \frac{1}{1-(1-x)} = \frac{1}{x}.$$

Damit erfüllen $L(x)$ und $\ln(x)$ beide das Anfangswertproblem $y' = 1/x$ und $y(1) = 0$. Das bedeutet, daß $\ln(x) = L(x)$ auf $(0, 2)$ ist.

Die Reihe $L(x)$ hat die reellen Koeffizienten $c_k = (-1)^{k+1}\dfrac{1}{k}$ und den Konvergenzradius 1, und die alternierende harmonische Reihe $\sum_{k=1}^{\infty} c_k$ konvergiert. Außerdem existiert der Grenzwert

$$\lim_{x \to 2-} L(x) = \lim_{x \to 2-} \ln(x) = \ln(2).$$

Mit dem Abelschen Grenzwertsatz folgt nun:

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n} = \ln(2).$$

2. Es ist
$$\arctan'(x) = \frac{1}{1+x^2} = \sum_{n=0}^{\infty}(-1)^n x^{2n}.$$
Setzen wir
$$F(x) := \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n+1}}{2n+1},$$
so erhalten wir eine beliebig oft differenzierbare Funktion auf $(-1,1)$, die dort mit ihrer Taylorreihe übereinstimmt. Es ist $F(0) = 0$ und $F'(x) = \dfrac{1}{1+x^2}$. Also müssen $F(x)$ und $\arctan(x)$ auf $(-1,1)$ übereinstimmen. Die Reihe $\displaystyle\sum_{n=0}^{\infty}(-1)^n \frac{1}{2n+1}$ erfüllt das Konvergenzkriterium von Leibniz. Also folgt mit dem Abelschen Grenzwertsatz:
$$\sum_{n=0}^{\infty}(-1)^n \frac{1}{2n+1} = \lim_{x\to 1}\arctan(x) = \frac{\pi}{4}.$$

Zum Schluß wollen wir noch eine besonders interessante Taylorentwicklung betrachten:

Wir beginnen mit dem Polynom
$$p(x) := (1+x)^n = \sum_{k=0}^{n} \frac{p^{(k)}(0)}{k!} x^k,$$
mit $p^{(0)}(0) = p(0) = 1$ und $p^{(k)}(0) = n(n-1)\cdot\ldots\cdot(n-k+1)$, für $1 \le k \le n$, also
$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k.$$
Das wollen wir jetzt verallgemeinern und die Funktion
$$f(x) := (1+x)^\alpha \quad \text{für } |x| < 1 \text{ und } \alpha \in \mathbb{R}$$
untersuchen. Hier ist
$$\frac{f^{(k)}(0)}{k!} = \frac{\alpha(\alpha-1)\cdot\ldots\cdot(\alpha-k+1)}{k!} =: \binom{\alpha}{k}.$$
Die *verallgemeinerten Binomialkoeffizienten* $\binom{\alpha}{k}$ werden durch die obige Gleichung definiert. Insbesondere setzt man $\binom{\alpha}{0} := 1$. Man beachte, daß diese Zahlen weder ganz noch positiv zu sein brauchen. Es gilt aber z.B.

$$\begin{aligned}\binom{\alpha-1}{k}+\binom{\alpha-1}{k-1} &= \frac{(\alpha-1)\cdots(\alpha-k)}{k!}+\frac{(\alpha-1)\cdots(\alpha-k+1)}{(k-1)!} \\ &= \frac{(\alpha-1)\cdot(\alpha-k+1)((\alpha-k)+k)}{k!} = \binom{\alpha}{k}.\end{aligned}$$

5.2 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

$$\begin{aligned}(1+x)B'(x) &= \alpha\cdot\left(\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^n+\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^{n+1}\right) \\ &= \alpha\cdot\left(1+\sum_{n=1}^{\infty}(\binom{\alpha-1}{n}+\binom{\alpha-1}{n-1})x^n\right) \\ &= \alpha\cdot\left(1+\sum_{n=1}^{\infty}\binom{\alpha}{n}x^n\right) = \alpha\cdot B(x).\end{aligned}$$

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. ∎

Beispiel.

Es ist

$$\sqrt{1+x} = (1+x)^{1/2} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} x^n = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 \pm \cdots .$$

Bemerkung. Sei $M \subset \mathbb{R}$ offen. Eine Funktion $f : M \to \mathbb{R}$ ist bei x_0 *in eine Potenzreihe entwickelbar*, falls ein $\varepsilon > 0$ existiert, so daß gilt:

$$f(x) = \sum_{\nu=0}^{\infty} a_\nu (x - x_0)^\nu, \qquad \text{für } x \in U_\varepsilon(x_0) \cap M.$$

f heißt auf M *analytisch*, falls f bei jedem Punkt $x \in M$ in eine Potenzreihe entwickelbar ist.

Ist f analytisch, so ist f beliebig oft differenzierbar, und die Taylorreihe von f in einem Punkt $x \in M$ konvergiert auf einer ganzen Umgebung von x gegen f. Jede konvergente Potenzreihe ist analytisch (der Beweis ist nicht trivial!), aber eine beliebig oft differenzierbare Funktion braucht nicht analytisch zu sein.