8

Ableitung

Die Ableitung einer Funktion $f: t \mapsto f(t)$ in einem Punkt a beschreibt ihr punktuelles *Veränderungsverhalten* in der Nähe eines Punktes a. Dazu betrachtet man die Änderung der abhängigen Größe, $f(t) - f(a)$, *im Verhältnis* zur Änderung der unabhängigen Größe, $t - a$, wenn t sich a nähert.

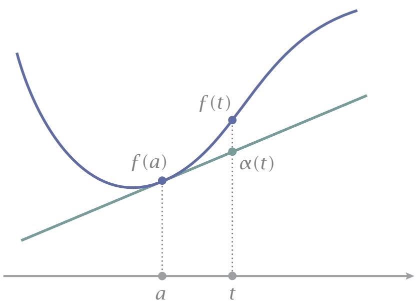
Im einfachsten Fall ist dieses Verhältnis *konstant*, nämlich dann, wenn es sich bei f um eine *lineare* oder *affine* Funktion der Gestalt $\alpha: t \mapsto mt + b$ handelt. Andernfalls kann man versuchen, eine solche affine Funktion zu finden, die die gegebene Funktion in der Umgebung von a am besten *approximiert*. Existiert eine solche bestapproximierende Gerade, so ist sie eindeutig – es handelt sich um die *Tangente* an den Graphen von f im Punkt a. Ihre Steigung m ist dann die Ableitung von f im Punkt a.

Diese Interpretation der Ableitung als Steigung einer bestapproximierenden Geraden werden wir später auf höhere Dimensionen verallgemeinern. Diese Steigung erscheint dort als lineare Abbildung zwischen Vektorräumen.

Doch zunächst beginnen wir mit der klassischen Definition der Ableitung als Grenzwert von Differenzenquotienten.

Abb 1

Tangente an einen Graphen



8.1

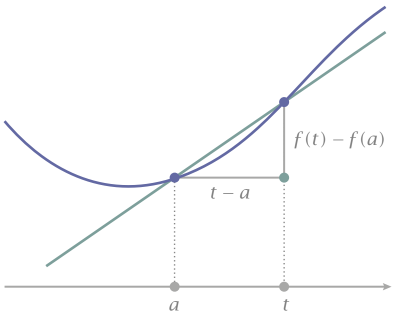


Abb 2

Sekante an einen Graphen

8.1 Definitionen und Rechenregeln

Im Folgenden sei I immer ein *nichtentartetes Intervall*, also ein Intervall mit mehr als einem Punkt. Ferner sei $f\colon I \to \mathbb{R}$ eine reellwertige Funktion auf I. Die folgende Definition der Differenzierbarkeit in einem Punkt ist wahrscheinlich aus der Schule vertraut.

Definition *Eine Funktion $f\colon I \to \mathbb{R}$ heißt differenzierbar im Punkt $a \in I$, wenn der Grenzwert*

$$\lim_{t\to a} \frac{f(t)-f(a)}{t-a} = m$$

existiert. In diesem Fall heißt m die erste Ableitung von f im Punkt a und wird mit $f'(a)$ bezeichnet. ⋊

Da t nur in der Nähe von a zu betrachten ist, schreibt man oft $t = a + h$ und nimmt h als hinreichend klein an. Damit wird

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

Dieser Definition liegt folgende geometrische Anschauung zugrunde. Der Differenzenquotient

$$\frac{f(t)-f(a)}{t-a}$$

beschreibt die Steigung der Sekante durch die Punkte des Graphen von f über a und t. Konvergieren diese Steigungen für $t \to a$ gegen einen Grenzwert m, so kann man diesen als *infinitesimale Steigung* von f im Punkt a auffassen. Die Grenzlage der Sekanten bildet dann eine *Tangente* an den Graphen von f mit der Steigung m.

Unbefriedigend an dieser Definition ist allerdings, dass sie sich in dieser Form nicht auf höhere Dimensionen verallgemeinern lässt, da eine Division durch

Vektoren nicht sinnvoll erklärt werden kann. Um dieses Problem zu umgehen, charakterisieren wir Differenzierbarkeit noch auf andere, äquivalente Weisen.

1 **Differenzierbarkeitssatz** *Für eine Funktion $f\colon I \to \mathbb{R}$ und einen Punkt $a \in I$ sind folgende Aussagen äquivalent:*

(i) *f ist differenzierbar in a mit $f'(a) = m$.*

(ii) *Es gibt eine reelle Zahl m, so dass*

$$\lim_{t\to a} \frac{|f(t) - f(a) - m(t-a)|}{|t-a|} = 0. \tag{1}$$

(iii) *Es gibt eine reelle Zahl m und eine im Punkt a stetige Funktion $\varepsilon\colon I \to \mathbb{R}$ mit $\varepsilon(a) = 0$, so dass*

$$f(t) = f(a) + m(t-a) + \varepsilon(t)(t-a). \tag{2}$$

(iv) *Es gibt eine im Punkt a stetige Funktion $\varphi\colon I \to \mathbb{R}$ mit $\varphi(a) = m$, so dass*

$$f(t) = f(a) + \varphi(t)(t-a), \qquad t \in I. \quad ⋊ \tag{3}$$

⟨⟨⟨⟨⟨ (i) ⇔ (ii) Die Definition von m als Ableitung ist äquivalent mit

$$0 = \lim_{t\to a}\left(\frac{f(t) - f(a)}{t-a} - m\right) = \lim_{t\to a} \frac{f(t) - f(a) - m(t-a)}{t-a},$$

und dies ist äquivalent mit (1).

(ii) ⇔ (iii) Gleichung (2) ist für $t \neq a$ äquivalent mit

$$\varepsilon(t) = \frac{f(t) - f(a) - m(t-a)}{t-a}.$$

Somit ist (1) äquivalent dazu, dass die Funktion ε durch 0 stetig nach a fortgesetzt werden kann.

(iii) ⇔ (iv) Dies ergibt sich mit $\varphi(t) = m + \varepsilon(t)$. ⟩⟩⟩⟩⟩

Wir haben nicht verlangt, dass eine im Punkt a differenzierbare Funktion dort auch stetig ist. Da in Gleichung (3) aber φ im Punkt a stetig ist, ist es auch die gesamte rechte Seite dieser Gleichung. Ist also f im Punkt a differenzierbar, so ist f dort auch stetig.

Gleichungen (1) und (2) erlauben folgende geometrische Interpretation. Die affine Funktion

$$\lambda\colon\ t \mapsto \lambda(t) = f(a) + m(t-a)$$

beschreibt eine Gerade durch den Punkt des Graphen von f über a mit der Steigung m, und für diese gilt

$$\lim_{t\to a} \frac{|f(t) - \lambda(t)|}{|t-a|} = 0.$$

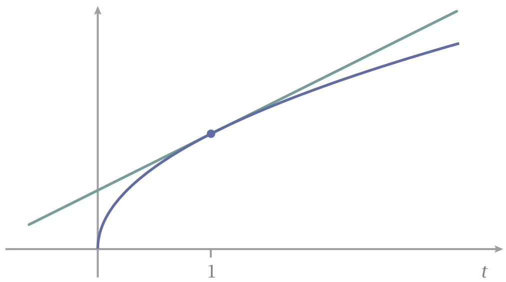


Abb 3

Wurzelfunktion mit Tangente

Der *Approximationsfehler* $|f(t) - \lambda(t)|$ zwischen Funktion und Gerade verschwindet also schneller als $|t-a|$ für $t \to a$. Diese Eigenschaft bestimmt diese Gerade in der Tat *eindeutig* A-33. Mit anderen Worten, die Tangente ist die *bestapproximierende Gerade* an den Graphen von f in diesem Punkt.

Definition *Ist $f: I \to \mathbb{R}$ im Punkt $a \in I$ differenzierbar, so heißt die Gerade*

$$\lambda: \ t \mapsto \lambda(t) = f(a) + f'(a)(t-a)$$

die Linearisierung und ihr Graph die Tangente von f im Punkt a. ⋊

2 ▷ *Beispiele* A. Jede affine Funktion $\alpha: t \mapsto mt + b$ ist in jedem Punkt differenzierbar, denn für alle t und h gilt

$$\frac{\alpha(t+h) - \alpha(t)}{h} = m.$$

Also ist $\alpha'(t) = m$ für alle t. Ihre Linearisierung im Punkt a ist

$$\lambda(t) = \alpha(a) + m(t-a) = ma + b + m(t-a) = \alpha(t).$$

Jede Gerade ist daher in jedem Punkt ihre eigene Tangente.

B. Die Wurzelfunktion $t \mapsto \sqrt{t}$ ist in jedem Punkt $t > 0$ differenzierbar, denn

$$\frac{\sqrt{t+h} - \sqrt{t}}{h} = \frac{1}{h}\frac{h}{\sqrt{t+h} + \sqrt{t}} = \frac{1}{\sqrt{t+h} + \sqrt{t}}$$

und damit

$$(\sqrt{t})' = \lim_{h \to 0} \frac{\sqrt{t+h} - \sqrt{t}}{h} = \frac{1}{2\sqrt{t}}, \qquad t > 0.$$

Im Punkt $a = 1$ beispielsweise ist die Linearisierung $t \mapsto 1 + (t-1)/2$ Abb 3.

Bei $t = 0$ dagegen divergiert der Differenzenquotient, und die Wurzelfunktion ist in 0 *nicht* differenzierbar.

C. Die Betragsfunktion $|\cdot|$ ist im Punkt 0 *nicht* differenzierbar, da es dort keine eindeutige bestapproximierende Tangente gibt. ◁

8.4

3 **Rechenregeln für die Ableitung** *Sind $f, g \colon I \to \mathbb{R}$ im Punkt $a \in I$ differenzierbar, so sind es auch $f + g$, fg, und f/g, falls $g(a) \neq 0$, und es gilt die Summen-, Produkt- und Quotientenregeln*

$$(f + g)'(a) = (f' + g')(a) \qquad \text{(Summenregel)}$$
$$(fg)'(a) = (f'g + fg')(a) \qquad \text{(Produktregel)}$$
$$(f/g)'(a) = ((f'g - fg')/g^2)(a) \qquad \text{(Quotientenregel)}.$$

⟨⟨⟨⟨⟨ Im Folgenden verwenden wir Kriterium (iv) von Satz 1. Wir schreiben

$$f(t) = f(a) + \varphi(t)(t - a), \qquad g(t) = g(a) + \psi(t)(t - a),$$

mit in a stetigen Funktionen φ und ψ, wobei $\varphi(a) = f'(a)$ und $\psi(a) = g'(a)$. Für das Produkt beispielsweise folgt daraus

$$\begin{aligned}(fg)(t) &= f(t)g(t) \\ &= f(a)g(a) + [\varphi(t)g(a) + f(a)\psi(t)]\,(t - a) \\ &\quad + \varphi(t)\psi(t)(t - a)^2 \\ &= (fg)(a) + [\varphi(t)g(a) + f(a)\psi(t) + \varphi(t)\psi(t)(t - a)]\,(t - a).\end{aligned}$$

Der Ausdruck in eckigen Klammern ist stetig im Punkt a und hat dort den Wert

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

Also ist fg in a differenzierbar mit der behaupteten Ableitung. Die übrigen Aussagen werden analog bewiesen. ⟩⟩⟩⟩⟩

4 **Satz** *Für alle $n \in \mathbb{Z}$ gilt*

$$(t^n)' = nt^{n-1},$$

wobei $t \neq 0$ für $n \leqslant 0$. ⋊

⟨⟨⟨⟨⟨ Für $n = 0$ ist dies die Ableitung der konstanten Funktion, $(1)' = 0$. Für $n = 1$ ist es die Ableitung der Identität, $(t)' = 1$. Mit Induktion über n und der Produktregel folgt für $n \geqslant 1$

$$(t^{n+1})' = (t^n t)' = (t^n)'t + t^n(t)' = nt^{n-1}t + t^n = (n + 1)t^n.$$

Für $n < 0$ erhält man hieraus mit der Quotientenregel

$$(t^n)' = \left(\frac{1}{t^{-n}}\right)' = \frac{-(t^{-n})'}{(t^{-n})^2} = \frac{nt^{-n-1}}{t^{-2n}} = \frac{n}{t^{-n+1}} = nt^{n-1}.$$

▷ *Beispiel* Ein reelles Polynom ist überall differenzierbar, und es gilt

$$\left(\sum_{k=0}^{n} a_k t^k\right)' = \sum_{k=0}^{n} k a_k t^{k-1} = \sum_{k=1}^{n} k a_k t^{k-1}. \quad ◁$$

Kettenregel und Umkehrregel

Wie bei der Stetigkeit untersuchen wir nun die Frage, unter welchen Bedingungen die Differenzierbarkeit bei Verkettung und Umkehrung von Funktionen erhalten bleibt. — Zuerst die Verkettung zweier Funktionen.

5 **Kettenregel** *Es seien $f: I \to J$ im Punkt $a \in I$ und $g: J \to \mathbb{R}$ im Punkt $f(a) \in J$ differenzierbar. Dann ist auch $g \circ f$ im Punkt a differenzierbar, und es gilt*

$$(g \circ f)'(a) = g'(f(a))f'(a). \qquad \rtimes$$

⟪⟪ Nach Voraussetzung ist

$$f(t) = f(a) + \varphi(t)(t-a), \qquad g(s) = g(b) + \psi(s)(s-b),$$

mit in a respektive $b = f(a)$ stetigen Funktionen φ und ψ, wobei $\varphi(a) = f'(a)$ und $\psi(b) = g'(b)$. Somit wird

$$\begin{aligned}(g \circ f)(t) &= g(f(t)) \\ &= g(b) + \psi(f(t))(f(t) - b) \\ &= (g \circ f)(a) + [\psi(f(t))\varphi(t)]\,(t-a).\end{aligned}$$

Der Ausdruck in eckigen Klammern ist stetig im Punkt a mit Wert

$$\psi(f(a))\varphi(a)) = g'(f(a))f'(a).$$

Also ist $g \circ f$ in a differenzierbar mit der behaupteten Ableitung. ⟫⟫

Man beachte, dass nur die Differenzierbarkeit von f und g in den Punkten a respektive $b = f(a)$ gefordert wird. Differenzierbarkeit in anderen Punkten wird nicht benötigt. Dasselbe gilt für die gleich folgende Umkehrregel $_6$.

▹ *Beispiel* Die Funktion

$$f: \ t \mapsto \sqrt{1+t^2}$$

kann als Komposition der Wurzel $g: x \mapsto \sqrt{x}$ mit dem Polynom $h: t \mapsto 1 + t^2$ aufgefasst werden. Sie ist auf ganz $\mathbb{R}$ definiert und differenzierbar mit der Ableitung

$$f'(t) = g'(x)\Big|_{x=h(t)} h'(t) = \frac{1}{2\sqrt{x}}\Big|_{x=h(t)} 2t = \frac{t}{\sqrt{1+t^2}}. \qquad ◃$$

6 **Umkehrregel** *Sei $f : I \to J$ stetig und bijektiv. Ist f im Punkt $a \in I$ differenzierbar und $f'(a) \neq 0$, so ist die Umkehrfunktion $f^{-1} : J \to I$ im Punkt $b = f(a)$ differenzierbar, und es gilt*

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))}. \tag{4}$$

⟨⟨⟨⟨⟨ Nach Voraussetzung ist wieder

$$f(t) = f(a) + \varphi(t)(t-a)$$

mit einer im Punkt a stetigen Funktion φ, wobei $\varphi(a) = f'(a) \neq 0$. Es existiert dann ein $\delta > 0$, so dass $\varphi(t) \neq 0$ für alle $t \in U_\delta(a) \cap I$. Für diese t können wir die erste Gleichung umformen zu

$$t = a + \frac{f(t) - f(a)}{\varphi(t)}.$$

Setzen wir $s = f(t)$ und $b = f(a)$, so ist $t = g(s)$ und $a = g(b)$ mit $g = f^{-1}$. Die letzte Gleichung geht dann über in

$$g(s) = g(b) + \frac{s-b}{\varphi(g(s))}.$$

Da der Nenner im Punkt b stetig ist und nicht verschwindet, ist g im Punkt b differenzierbar, und die Ableitung ist wie behauptet.

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

Man kann Gleichung (4) leicht *rekonstruieren*. Aus $f^{-1}(f(t)) = t$ folgt mit der Kettenregel 5

$$(f^{-1})'(f(t))f'(t) = 1,$$

also

$$(f^{-1})'(f(t)) = \frac{1}{f'(t)}.$$

Dies ist allerdings *kein Ersatz* für den vorangehenden Beweis. Wir müssen ja zuerst *wissen*, dass f^{-1} im Punkt b differenzierbar ist, bevor wir die Kettenregel anwenden dürfen.

Beispiel Die Parabel $f : t \mapsto t^2$ ist für $t \geqslant 0$ strikt wachsend, bijektiv und differenzierbar mit Ableitung

$$f'(t) = 2t.$$

Diese ist nicht Null für $t > 0$. Die Umkehrfunktion $g = f^{-1}$ ist deshalb in jedem Punkt $x = t^2 > 0$ differenzierbar mit Ableitung

$$g'(x) = \frac{1}{f'(g(x))} = \frac{1}{2g(x)}.$$

Für die Wurzelfunktion erhalten wir also wieder das Ergebnis von Beispiel 2,

$$(\sqrt{x})' = \frac{1}{2\sqrt{x}}.$$

Andererseits ist $f'(0) = 0$, und in $0 = f(0)$ ist die Wurzelfunktion auch tatsächlich nicht differenzierbar. Die Annahme, dass f' im betrachteten Punkt nicht verschwinden darf, ist also notwendig.

Differenzierbare Funktionen

Wir definieren nun noch die Differenzierbarkeit auf einem Intervall. Dies geschieht genauso wie im Fall der Stetigkeit.

Definition *Eine Funktion $f : I \to \mathbb{R}$ heißt differenzierbar auf I, oder einfach differenzierbar, wenn sie in jedem Punkt von I differenzierbar ist. In diesem Fall heißt die Funktion*

$$f' : I \to \mathbb{R}, \quad t \mapsto f'(t)$$

die Ableitung von f. Ist f' stetig, so heißt f stetig differenzierbar auf I.

Der Raum aller auf I stetig differenzierbaren reellen Funktionen wird mit $C^1(I)$ bezeichnet. Offensichtlich gilt

$$C^1(I) \subsetneq C^0(I) \coloneqq C(I).,$$

denn stetige Funktionen müssen nicht differenzierbar sein. Es gibt sogar stetige Funktionen auf $\mathbb{R}$, die in *keinem einzigen Punkt* differenzierbar sind ??. Daneben gibt es natürlich auch Funktionen, die differenzierbar, aber nicht stetig differenzierbar sind A-9.

8.8

7 **Satz** *Sind f und g in $C^1(I)$, so sind auch $f+g$ und fg in $C^1(I)$, und es gilt*

$$(f+g)' = f' + g', \qquad (fg)' = f'g + fg'.$$

Verschwindet g nirgends auf I, so ist auch f/g in $C^1(I)$, und es gilt

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}. \quad \rtimes$$

Aufgrund dieses Satzes ist $C^1(I)$ nicht nur ein Vektorraum, sondern eine kommutative *Algebra* – das heißt, es ist neben der Addition eine Multiplikation erklärt, für die die Assoziativ-, Kommutativ- und Distributivgesetze gelten. — Nun noch die globale Kettenregel.

8 **Kettenregel** *Sind f und g beide C^1 und ist $g \circ f$ auf dem Definitionsbereich von f erklärt, so ist auch $g \circ f$ eine C^1-Funktion, und es gilt*

$$(g \circ f)' = (g' \circ f)\, f'. \quad \rtimes$$

Die globale Umkehrregel formulieren wir im nächsten Abschnitt $_{15}$, da wir hierfür den Monotoniesatz $_{??}$ benötigen.

8.2 Lokale Extrema und Mittelwertsatz

Es liegt nahe, dass die Ableitung einer Funktion Aufschluss gibt über ihr lokales Verhalten. Man sollte zum Beispiel an ihr erkennen, ob sie monoton ist oder eine Extremstelle ausbildet. Dies wollen wir jetzt untersuchen.

Definition *Eine Funktion $f\colon I \to \mathbb{R}$ besitzt im Punkt $c \in I$ ein lokales Minimum, wenn es eine Umgebung $U_\delta(c)$ gibt, so dass*

$$f(c) \leqslant f(t), \qquad t \in U_\delta(c) \cap I.$$

Sie besitzt in c ein striktes lokales Minimum, wenn außerdem für $t \neq c$ die strikte Ungleichung gilt, also

$$f(c) < f(t), \qquad t \in \dot{U}_\delta(c) \cap I.$$

Entsprechend ist ein (striktes) lokales Maximum definiert. $\rtimes$

Lokale Minima und Maxima werden gemeinsam als *Extrema* bezeichnet. Punkte, an denen ein lokales Extremum vorliegt, werden *Extremstellen* genannt, unterschieden in *Minimal-* und *Maximalstellen*.

Der folgende Satz ist von fundamentaler Bedeutung für das Auffinden von Extremalstellen. Hierbei heißt $c \in I$ *innerer Punkt* des Intervalls I, wenn er kein Randpunkt ist.

9 **Satz von Fermat** *Besitzt $f: I \to \mathbb{R}$ in einem inneren Punkt c von I ein lokales Extremum und ist f in diesem Punkt differenzierbar, so gilt $f'(c) = 0$.* ⋊

⟨⟨⟨⟨⟨ Sei zum Beispiel c eine Minimalstelle von f im Innern von I. Dann existiert ein $\delta > 0$, so dass $U_\delta(c) \subset I$ und

$$f(c+h) \geqslant f(c), \qquad |h| < \delta.$$

Also gilt

$$\frac{f(c+h) - f(c)}{h} \geqslant 0, \qquad 0 < h < \delta,$$

und

$$\frac{f(c+h) - f(c)}{h} \leqslant 0, \qquad 0 > h > -\delta.$$

Da wegen der Differenzierbarkeit von f im Punkt c der Grenzwert für $h \to 0$ existiert, folgt aus diesen Ungleichungen sowohl $f'(c) \geqslant 0$ als auch $f'(c) \leqslant 0$. Also ist $f'(c) = 0$. ⟩⟩⟩⟩⟩

Beide Voraussetzungen des Satzes – Lage der Extremstelle c im Innern und Differenzierbarkeit im Punkt c – sind *notwendig*. In einer Extremstelle am Rand muss die Ableitung nicht verschwinden, und an einer Extremstelle im Innern muss eine Funktion nicht differenzierbar sein.

Das Kriterium von Fermat ist allerdings nur *notwendig*, aber *nicht hinreichend*. Ein kritischer Punkt ist nicht notwendigerweise eine Extremstelle – siehe Abbildung 6.

Punkte, in denen die Ableitung einer Funktion verschwindet, haben eine besondere Bedeutung und deshalb auch einen besonderen Namen.

Abb 4

Verschiedene Minimalstellen

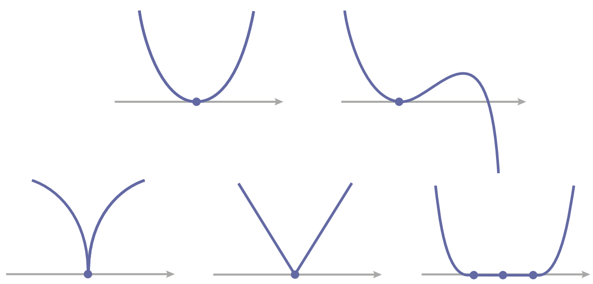
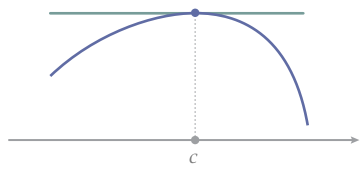


Abb 5

Satz von Fermat für eine Maximalstelle



Definition *Ist $f : I \to \mathbb{R}$ im Punkt c differenzierbar und $f'(c) = 0$, so heißt c ein stationärer oder kritischer Punkt von f.* $\rtimes$

Der Satz von Fermat besagt also, dass eine Extremstelle im Innern notwendig ein kritischer Punkt ist, wenn die Funktion dort differenzierbar ist. Geometrisch bedeutet dies, dass der Graph von f an einer solchen Extremstelle eine *horizontale Tangente* aufweist.

▪ Mittelwertsätze

Als erste Anwendung des Satzes von Fermat erhalten wir sogenannte Mittelwertsätze. Zuerst betrachten wir einen Spezialfall.

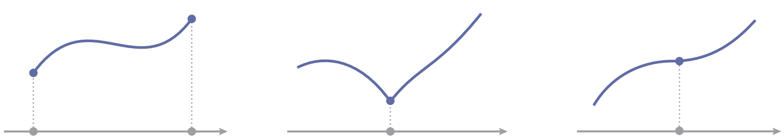
10 **Satz von Rolle** *Sei $f : [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar. Gilt $f(a) = f(b)$, so existiert ein Punkt $c \in (a,b)$ mit*

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

Somit besitzt f einen kritischen Punkt im Innern von $[a,b]$. $\rtimes$

⟨⟨⟨⟨⟨ Ist f konstant, so verschwindet f' überall, und *jeder* Punkt in $[a,b]$ ist ein kritischer Punkt. Sei also f nicht konstant. Die stetige Funktion f nimmt auf $[a,b]$ ihr Minimum und Maximum an, und beide können nicht gleichzeitig am Rand liegen, da f sonst konstant wäre. Also besitzt f wenigstens eine Extremalstelle c im Innern von $[a,b]$. Da f dort auch differenzierbar ist, ist nach dem Satz von Fermat $_9$ $f'(c) = 0$. ⟩⟩⟩⟩⟩

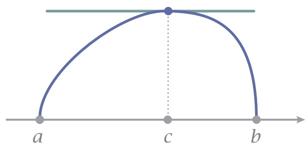
Abb 6 Zum Satz von Fermat



Extremstellen am Rand respektive nicht differenzierbar *Kritischer Punkt nicht extremal*

Abb 7

Satz von Rolle



Der Satz von Rolle gilt natürlich erst recht, wenn f auf ganz $[a,b]$ differenzierbar ist. Es ist aber sinnvoll, hier und im Folgenden nur die Differenzierbarkeit auf (a,b) zu verlangen, um auch Funktionen zu erfassen, die in den Endpunkten des Intervalls stetig, aber *nicht* differenzierbar sind, wie zum Beispiel in Abbildung 7. Und überhaupt vermeidet ein Mathematiker gerne Bedingungen, die er nicht benötigt.

Übrigens war Rolle ein französischer Mathematiker, sein Name wird daher ohne ›e‹ gesprochen. — Nun der allgemeine Fall.

11 **Mittelwertsatz (Mws)** *Sei $f\colon [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar. Dann existiert ein Punkt $c \in (a,b)$, so dass*

$$f(b) - f(a) = f'(c)(b-a). \quad \rtimes$$

⟨⟨⟨⟨⟨ Betrachte die Hilfsfunktion φ mit

$$\varphi(t) = f(t) - m(t-a), \qquad m = \frac{f(b)-f(a)}{b-a}.$$

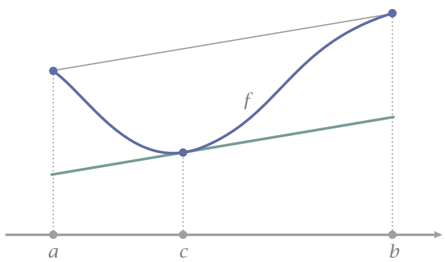
Mit anderen Worten, wir subtrahieren von der Funktion f die Sekante durch die Eckpunkte ihres Graphen. Für diese Funktion gilt $\varphi(a) = \varphi(b)$. Auch ist φ auf $[a,b]$ stetig und auf (a,b) differenzierbar. Nach dem Satz von Rolle $_{10}$ existiert also ein $c \in (a,b)$ mit

$$\varphi'(c) = f'(c) - m = 0.$$

Dies ist äquivalent zur Behauptung. ⟩⟩⟩⟩⟩

Abb 8

Zum Mittelwertsatz



Der Quotient m im Beweis des Mittelwertsatzes stellt die *mittlere Steigung* der Funktion f im Intervall $[a,b]$ dar. Der Mittelwertsatz sagt also aus, dass an wenigstens einer Stelle im Innern des Intervalls die Tangentensteigung gleich der mittleren Steigung ist. Der Satz von Rolle ist hiervon ein Spezialfall, denn für $f(b) = f(a)$ ist die mittlere Steigung 0. Andererseits haben wir den Mittelwertsatz gerade mit dem Satz von Rolle bewiesen. Beide Sätze sind somit *äquivalent*.

Oft wird der Mittelwertsatz in folgender Form angewandt. Sei $f\colon\ [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar Dann existiert zu $t, t+h \in [a,b]$ immer ein $\theta \in (0,1)$, so dass

$$f(t+h) - f(t) = f'(t+\theta h)h.$$

In dieser Formulierung ist das Vorzeichen von h und damit die Anordnung von t und $t+h$ unerheblich.

12 **Korollar** *Sei $f\colon\ [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar. Dann ist*

(i) *$f' = 0$ auf (a,b) gdw f konstant auf $[a,b]$,*

(ii) *$f' \geqslant 0$ auf (a,b) gdw f monoton steigend auf $[a,b]$,*

(iii) *$f' \leqslant 0$ auf (a,b) gdw f monoton fallend auf $[a,b]$,*

(iv) *$\|f'\|_{(a,b)} \leqslant L$ gdw f lipschitz auf $[a,b]$ mit Lipschitzkonstante L.* ⋊

⟨⟨⟨⟨ Wir beweisen zuerst die Richtung $\Rightarrow$.

(i) Für $t \in (a,b]$ wenden wir den Mittelwertsatz auf das Intervall $[a,t]$ an und erhalten ein $c \in (a,t)$ mit

$$f(t) - f(a) = f'(c)(t-a) = 0.$$

Also ist $f(t) = f(a)$ für alle $t \in (a,b]$, und damit f konstant.

(ii) Entsprechend wenden wir den Mittelwertsatz auf zwei Punkte $u < v$ in $[a,b]$ an und erhalten ein $c \in (u,v)$ mit

$$f(v) - f(u) = f'(c)(v-u) \geqslant 0.$$

Somit ist f monoton steigend auf $[a,b]$. Ditto (iii).

(iv) Ebenso folgt hieraus

$$|f(v) - f(u)| \leqslant |f'(c)|\,|v-u| \leqslant \sup_{t\in(a,b)} |f'(t)|\,|v-u| \leqslant L\,|v-u|\,.$$

Die Richtung $\Leftarrow$ folgt aus der Betrachtung der Differenzenquotienten. ⟩⟩⟩⟩

Aus $f' > 0$ auf (a,b) folgt außerdem, dass f auf $[a,b]$ streng monoton steigt. Hiervon gilt allerdings *nicht* die Umkehrung. So ist die kubische Parabel $t \mapsto t^3$ streng monoton steigend, aber ihre Ableitung verschwindet bei $t = 0$.

Wir benötigen noch folgende Verallgemeinerung des Mws. Er beinhaltet den ›normalen‹ Mittelwertsatz mit $g = id$.

13 **Allgemeiner Mittelwertsatz** *Seien $f, g\colon [a,b] \to \mathbb{R}$ stetig, auf (a,b) differenzierbar, und $g' \neq 0$ auf (a,b). Dann gibt es ein $c \in (a,b)$ mit*

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

⟨⟨⟨⟨⟨ Es ist $g(a) \neq g(b)$, da andernfalls g in (a,b) aufgrund des Satzes von Rolle einen krtitischen Punkt hätte. Also ist

$$\varphi(t) = f(t) - m(g(t) - g(a)), \qquad m = \frac{f(b) - f(a)}{g(b) - g(a)},$$

auf $[a,b]$ wohldefiniert, stetig und auf (a,b) differenzierbar. Außerdem ist $\varphi(a) = \varphi(b)$. Mit Rolle $_{10}$ existiert also ein Punkt $c \in (a,b)$ mit

$$0 = \varphi'(c) = f'(c) - mg'(c).$$

Dies ist äquivalent zur Behauptung. ⟩⟩⟩⟩⟩

Aus dem Korollar $_{12}$ ergibt sich ein einfaches Kriterium für das Vorliegen einer Extremalstelle.

14 **Satz** *Die Funktion $f\colon I \to \mathbb{R}$ sei differenzierbar und besitze im Innern von I einen kritischen Punkt c. Ist f' in einer Umgebung von c streng monoton steigend respektive fallend, so besitzt f in c ein striktes lokales Minimum respektive Maximum.* $\rtimes$

⟨⟨⟨⟨⟨ Angenommen, f' ist in einer Umgebung $U_\delta(c) \subset I$ streng monoton steigend. Wegen $f'(c) = 0$ ist dann in dieser Umgebung $f'(t) < 0$ für $t < c$ und $f'(t) > 0$ für $t > c$. Also ist f auf $(c - \delta, c)$ streng monoton fallend und auf $(c, c + \delta)$ streng monoton steigend. Das aber bedeutet, dass

$$f(c) < f(c + h), \qquad 0 < |h| < \delta.$$

Somit liegt bei c ein striktes Minimum. ⟩⟩⟩⟩⟩

▷ *Beispiel* In welchem Verhältnis müssen Höhe h und Radius r einer Konservendose gewählt werden, um bei vorgegebenen Volumen V den Blechbedarf zu minimieren? Für Volumen und Oberfläche gilt

$$V = \pi r^2 h, \qquad A = 2\pi r^2 + 2\pi r h.$$

Auflösen der Volumengleichung nach h auf und Einsetzen in die Oberflächengleichung ergibt

$$A(r) = 2\pi r^2 + \frac{2V}{r}.$$

Da $A(r)$ unbeschränkt ist für $r \searrow 0$ und $r \nearrow \infty$, muss aus Stetigkeitsgründen mindestens ein Minimum auf $(0, \infty)$ vorliegen. Nach dem Satz von Fermat ist dort notwendigerweise

$$0 = A'(r) = 4\pi r - \frac{2V}{r^2} \quad \Leftrightarrow \quad r^3 = \frac{V}{2\pi}.$$

Da es nur einen einzigen kritischen Punkt gibt, muss dieser die Minimalstelle sein, auch ohne weitere Betrachtung von A' oder A''. ◂

Schließlich erhalten wir noch einen handlichen Satz über Umkehrfunktionen.

15 **Satz über C¹-Umkehrfunktionen** *Ist f in $C^1(I)$ und $f' \neq 0$ auf ganz I, so ist f umkehrbar, die Umkehrfunktion f^{-1} ist auf dem Intervall $J = f(I)$ ebenfalls C^1, und es gilt dort*

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}. \qquad \rtimes$$

⟨⟨⟨⟨ Da f' auf I stetig ist und nirgends verschwindet, hat f' festes Vorzeichen. Somit $_{12}$ ist f streng monoton und somit umkehrbar. Aufgrund des Umkehrsatzes $_6$ ist, da f' nirgends verschwindet, die Umkehrfunktion $g = f^{-1}$ auf $f(I)$ differenzierbar mit Ableitung

$$g' = \frac{1}{f' \circ g}.$$

Diese ist offensichtlich ebenfalls stetig. Somit ist f^{-1} ebenfalls C^1. ⟩⟩⟩⟩

8.3
Höhere Ableitungen

Ist $f\colon I \to \mathbb{R}$ differenzierbar, so ist die Ableitung $f'\colon I \to \mathbb{R}$ wieder eine reelle Funktion auf I. Somit kann man fragen, ob f' ebenfalls differenzierbar ist. Ist dies der Fall, so erhält man die *zweite Ableitung* von f,

$$f'' \coloneqq (f')' : I \to \mathbb{R}.$$

So kann man fortfahren, so lange die nächste Ableitung ebenfalls differenzierbar ist, und induktiv die höheren Ableitungen von f erklären. Bezeichnen wir die r-te Ableitung mit $f^{(r)}$, so gilt mit $f^{(0)} \coloneqq f$ also

$$f^{(r)} \coloneqq (f^{(r-1)})', \qquad r \geqslant 1.$$

Insbesondere ist $f^{(1)} = f'$ und $f^{(2)} = f''$. Andere übliche Bezeichnungen hierfür sind $D^r f$ oder auch $\partial^r f$.

Man sagt, f ist *r-mal differenzierbar*, wenn $f, f', .., f^{(r)}$ existieren. Ist außerdem $f^{(r)}$ stetig, so heißt f *r-mal stetig differenzierbar*. Die Klasse dieser

Funktionen wird mit $C^r(I)$ bezeichnet, und man sagt, eine Funktion $f \in C^r(I)$ ist eine *C^r-Funktion* oder kurz C^r.

Schließlich definiert man noch den Raum der unendlich oft differenzierbaren Funktionen auf I,

$$C^\infty(I) \coloneqq \bigcap_{r \geqslant 0} C^r(I).$$

Man erhält somit eine *Skala* von Funktionenräumen

$$C^\infty(I) \subsetneq \,..\, \subsetneq C^{r+1}(I) \subsetneq C^r(I) \subsetneq \,..\, \subsetneq C^1(I) \subsetneq C^0(I).$$

Jede Inklusion ist eine echte Inklusion, denn natürlich gibt es für jedes $r \geqslant 0$ C^r-Funktionen, die nicht C^{r+1} sind.

16 ⊳ A. Jede Potenz $t \mapsto t^n$ mit $n \geqslant 0$ ist unendlich oft differenzierbar, wobei

$$(t^n)^{(r)} = n(n-1)\cdot\cdot(n-r+1)\,t^{n-r} = \frac{n!}{(n-r)!}t^{n-r}, \qquad 0 \leqslant r \leqslant n,$$

insbesondere

$$(t^n)^{(n)} \equiv n!\ .$$

Somit ist auch $(t^n)^{(r)} \equiv 0$ für alle $r > n$.

B. Jedes Polynom ist unendlich oft differenzierbar. Ist das Polynom p vom Grad n, so ist $p^{(n+1)} \equiv 0$.

C. Die im nächsten Kapitel definierten Funktionen exp, sin und cos sind auf $\mathbb{R}$ unendlich oft differenzierbar, und es gilt

$$\exp' = \exp, \qquad \sin' = \cos, \qquad \cos' = -\sin. \quad ⊲$$

Jede Linearkombination von C^r-Funktionen ist wieder eine C^r-Funktion, wie man induktiv mit $(\alpha f + \beta g)' = \alpha f' + \beta g'$ leicht zeigt. Somit ist $C^r(I)$ ein reeller Vektorraum. Es gilt aber noch mehr.

17 **Leibnizsche Formel** *Das Produkt zweier Funktionen $f, g \in C^r(I)$ gehört ebenfalls zu $C^r(I)$, und es gilt*

$$(fg)^{(r)} = \sum_{s=0}^{r} B_s^r f^{(r-s)} g^{(s)}$$

mit den Binomialkoeffizienten $B_s^r = \dfrac{r!}{s!\,(r-s)!} = \dbinom{r}{s}$. ⋊

Den Beweis erfolgt per Induktion wie bei der binomischen Formel A-35. Auch bei der Komposition und der Umkehrung gibt es keine Überraschungen.

18 **Satz über C^r-Komposition** *Ist $f \in C^r(I)$ und $g \in C^r(J)$ mit $J \supset f(I)$, so ist auch $g \circ f \in C^r(I)$.* ⋊

⟨⟨⟨⟨⟨ Für $r = 1$ ist dies die Kettenregel $_8$, und es gilt

$$(g \circ f)' = (g' \circ f)\, f'.$$

Alles Weitere folgt mit Induktion: Sind f und g von der Klasse C^{r+1}, so sind f' und g' von der Klasse C^r. Dann ist auch $g' \circ f$ C^r aufgrund der Induktionsannahme, und damit auch das Produkt mit f'. Somit ist $(g \circ f)'$ C^r und damit $g \circ f$ selbst C^{r+1}. ⟩⟩⟩⟩⟩

19 **Satz über C^r-Umkehrfunktion** *Sei $f \in C^r(I)$. Verschwindet f' nirgends, so ist f umkehrbar und $f^{-1} \in C^r(J)$ mit $J = f(I)$.* ⋊

⟨⟨⟨⟨⟨ Aufgrund des Satzes über C^1-Umkehrfunktionen $_{15}$ ist f auf jeden Fall umkehrbar, und für $\phi = f^{-1}$ gilt $\phi \in C^1(J)$ mit

$$\phi' = \frac{1}{f' \circ \phi}.$$

Alles Weitere folgt mit Induktion: Ist f C^{r+1}, so ist f' C^r ebenso wie ϕ aufgrund der Induktionsannahme. Dann steht auf der rechten Seite ebenfalls eine C^r-Funktion aufgrund des Satzes über die C^-r-Komposition. Somit ist ϕ' C^r und damit ϕ selbst C^{r+1}. ⟩⟩⟩⟩⟩

Bemerkung Die zweite Ableitung von $\phi = f^{-1}$ ist dann

$$\phi'' = -\frac{1}{(f' \circ \phi)^2}\,(f'' \circ \phi)\,\phi' = -\frac{f'' \circ \phi}{(f' \circ \phi)^3} = -\left(\frac{f''}{f'^3}\right) \circ \phi. \qquad ⊸$$

8.4 Die Regel von l'Hospital

Oft sind sogenannte ›unbestimmte Ausdrücke‹ der Form $0/0$ zu bestimmen, also Grenzwerte wie

$$\lim_{t \to 0} \frac{\sin t}{t}, \qquad \lim_{t \to 0} \frac{1 - \cos^2 t}{t^2},$$

wo Zähler und Nenner gleichzeitig gegen Null konvergieren. Für diesen Fall gibt es eine nützliche Regel, die auf dem verallgemeinerten Mittelwertsatz basiert.

20 **Einfache Regel von l'Hospital** *Seien $f, g \in C^1(I)$. Ist in einem Punkt $a \in I$*

$$f(a) = g(a) = 0$$

und $g' \neq 0$ in einer punktierten Umgebung von a, so gilt

$$\lim_{t \to a} \frac{f(t)}{g(t)} = \lim_{t \to a} \frac{f'(t)}{g'(t)}.$$

Insbesondere existiert der linke Grenzwert, wenn der rechte Grenzwert existiert. ⋊

⟨⟨⟨⟨⟨ Aufgrund des allgemeinen Mittelwertsatzes $_{13}$ ist

$$\frac{f(t)}{g(t)} = \frac{f(t) - f(a)}{g(t) - g(a)} = \frac{f'(s)}{g'(s)}$$

mit einem s zwischen a und t. Mit $t \to a$ haben wir auch $s \to a$. Existert also der Grenzwert auf der rechten Seite für $s \to a$, so folgt

$$\lim_{t \to a} \frac{f(t)}{g(t)} = \lim_{s \to a} \frac{f'(s)}{g'(s)}.$$ ⟩⟩⟩⟩⟩

▷ *Beispiele* A. $\displaystyle \lim_{x \to 1} \frac{x^2 - 1}{x^2 + x - 2} = \lim_{x \to 1} \frac{2x}{2x + 1} = \frac{2}{3}$,

B. $\displaystyle \lim_{t \to 0} \frac{\sin t}{t} = \lim_{t \to 0} \frac{\cos t}{1} = 1$. C. $\displaystyle \lim_{t \to a} \frac{f(t) - f(a)}{t - a} = \lim_{t \to a} \frac{f'(t)}{1} = f'(a)$.

D. Man darf nicht vergessen zu überprüfen, ob tatsächlich alle Voraussetzungen der Regel erfüllt sind. So ist zum Beispiel

$$\lim_{t \to 0} \frac{t^2}{\cos t} = 0 \quad \neq \quad \lim_{t \to 0} \frac{2t}{-\sin t} = \lim_{t \to 0} \frac{2}{-\cos t} = -2.$$ ◁

Ist auch $f'(a) = g'(a) = 0$, so kann man den Satz nochmals anwenden. Erfüllen die Ableitungen von f' und g' die entsprechenden Voraussetzungen, so geht man zu den zweiten Ableitungen über.

▷ A. $\displaystyle \lim_{x \to 1} \frac{x^3 - x^2 - x + 1}{x^3 - 3x + 2} = \lim_{x \to 1} \frac{3x^2 - 2x - 1}{3x^2 - 3} = \lim_{x \to 1} \frac{6x - 2}{6x} = \frac{2}{3}$.

B. $\displaystyle \lim_{t \to 0} \frac{1 - \cos^2 t}{t^2} = \lim_{t \to 0} \frac{\sin t \cos t}{t} = \lim_{t \to 0} \frac{\cos^2 t - \sin^2 t}{1} = 1$.

C. $\displaystyle \lim_{h \to 0} \frac{f(a+h) - 2f(a) + f(a-h)}{h^2} = \lim_{h \to 0} \frac{f'(a+h) - f'(a-h)}{2h} = f''(a)$. ◁

▪ Der allgemeine Fall

Oft handelt es sich bei a um einen Randpunkt eines Definitionsintervalls, in dem f' und g' nicht definiert sind. Auch ist der Fall eines uneigentlichen Randpunktes bei ∞ oder $-\infty$ von Interesse, sowie unbestimmte Ausdrücke der Form ∞/∞. Auf alle diese Fälle lässt sich die Regel von l'Hospital verallgemeinern. Der Beweis ist etwas technisch und kann beim ersten Lesen übergangen werden.

21 **Regel von l'Hospital** *Es seien $f, g \in C^1(I)$. Ist a ein Randpunkt von I mit*

$$\lim_{t\to a} f(t) = \lim_{t\to a} g(t) = 0 \qquad \textit{oder} \qquad \lim_{t\to a} g(t) = \pm\infty,$$

und ist $g' \neq 0$ in einer punktierten Umgebung von a, so gilt

$$\lim_{t\to a} \frac{f(t)}{g(t)} = \lim_{t\to a} \frac{f'(t)}{g'(t)}, \tag{5}$$

falls der zweite Grenzwert auf der erweiterten Zahlengerade existiert. ⋊

⟪⟪⟪ Sei

$$\lambda = \lim_{t\to a} \frac{f'(t)}{g'(t)}. \tag{6}$$

Wir zeigen, dass im Fall $\lambda < \infty$ zu jedem $\lambda^+ > \lambda$ ein Intervall $I_\delta = \dot{U}_\delta(a) \cap I$ existiert, so dass

$$\frac{f(t)}{g(t)} \leqslant \lambda^+, \qquad t \in I_\delta. \tag{7}$$

Analog zeigt man im Fall $\lambda > -\infty$, dass zu jedem $\lambda^- < \lambda$ ein ähnliches Intervall J_δ existiert, so dass

$$\frac{f(t)}{g(t)} \geqslant \lambda^-, \qquad t \in J_\delta.$$

Beides zusammen ergibt die Behauptung (5).

Sei also $\lambda < \lambda^+ < \infty$. Da der Grenzwert in Gleichung (6) existiert, gibt es ein $\varepsilon > 0$ und ein Intervall $I_\delta = \dot{U}_\delta(a) \cap I$ so, dass

$$\frac{f'(t)}{g'(t)} \leqslant \lambda^+ - \varepsilon, \qquad t \in I_\delta.$$

Mit dem verallgemeinerten Mittelwertsatz folgt hieraus

$$\frac{f(t) - f(s)}{g(t) - g(s)} = \frac{f'(u)}{g'(u)} \leqslant \lambda^+ - \varepsilon, \qquad s, t \in I_\delta, \tag{8}$$

mit einem von s und t abhängenden $u \in I_\delta$.

Gilt jetzt

$$\lim_{t\to a} f(t) = \lim_{t\to a} g(t) = 0,$$

so folgt mit $s \to a$ hieraus (7), und wir sind fertig.

Gilt dagegen beispielsweise $\lim_{t \to a} g(t) = \infty$, so schreiben wir zuerst

$$\frac{f(t)}{g(t)} = \frac{f(t) - f(s)}{g(t) - g(s)} \frac{g(t) - g(s)}{g(t)} + \frac{f(s)}{g(t)}.$$

Für irgendein festes $s \in I_\delta$ und t hinreichend nahe bei a ist $g(t) - g(s) > 0$, so dass mit (8)

$$\frac{f(t)}{g(t)} \leqslant (\lambda^+ - \varepsilon) \frac{g(t) - g(s)}{g(t)} + \frac{f(s)}{g(t)}, \qquad t \in I_\theta,$$

auf einem hinreichend kleinen Intervall $I_\theta \subset I_\delta$. Für $t \to a$ konvergiert die rechte Seite gegen $\lambda^+ - \varepsilon$. Wählen wir also I_θ nochmal kleiner, so gilt auch

$$\frac{f(t)}{g(t)} \leqslant \lambda^+, \qquad t \in I_\theta,$$

und damit wieder (7). Der Fall $\lim_{t \to a} g(t) = -\infty$ geht genauso. ⟫⟫⟫

Bemerkung Der Beweis gilt für den Fall $t \to a$ und $t \to \pm\infty$ als auch für den Fall, dass es sich in (5) um einen uneigentlichen Grenzwert handelt. ⊸

▷ A. $\displaystyle\lim_{t\to\infty} t^2 e^{-t} = \lim_{t\to\infty} \frac{t^2}{e^t} = \lim_{t\to\infty} \frac{2t}{e^t} = \lim_{t\to\infty} \frac{2}{e^t} = 0.$

B. Ganz allgemein gilt $\displaystyle\lim_{t\to\infty} t^m e^{-t} = 0$ für $m \geqslant 0$.

C. $\displaystyle\lim_{t\to\infty} \frac{\log t}{t} = \lim_{t\to\infty} \frac{1/t}{1} = 0.$

Andere unbestimmte Ausdrücke wie $0 \cdot \infty$ lassen sich oft in eine Form bringen, auf die man die Regel von l'Hospital anwenden kann:

D. $\displaystyle\lim_{t\to 0} (t \log t) = \lim_{t\to 0} \frac{\log t}{1/t} = -\lim_{t\to 0} \frac{1/t}{1/t^2} = -\lim_{t\to 0} t = 0.$

E. $\displaystyle\lim_{t\to\infty} t \log\left(1 + \frac{1}{t}\right) = \lim_{s\to 0} \frac{\log(1+s)}{s} = \lim_{s\to 0} \frac{1/(1+s)}{1} = 1.$

F. $\displaystyle\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = \lim_{n\to\infty} \exp\left(n \log\left(1 + \frac{1}{n}\right)\right) = e^1 = e.$ ◁

8.5 Taylorpolynome

Sei f eine reelle Funktion auf einem Intervall I und $a \in I$. Die *Stetigkeit* von f im Punkt a können wir auffassen als lokale Approximierbarkeit von f durch eine *konstante Funktion*, denn es gilt

$$f(a+h) = f(a) + \varepsilon(h)$$

mit einer in 0 stetigen und dort verschwindenden Funktion ε. Die *Differenzierbarkeit* von f in a ist gleichbedeutend mit der lokalen Approximierbarkeit durch eine *lineare Funktion*, denn

$$f(a+h) = f(a) + f'(a)h + \varepsilon(h)h$$

wiederum mit einer in 0 stetigen und dort verschwindenden Funktion ε_1.

Man kann daher fragen, ob sich eine Funktion lokal auch durch *Polynome höherer Ordnung* so approximieren lässt, dass

$$f(a+h) = f(a) + \sum_{k=1}^{n} a_k h^k + \varepsilon(h)h^n$$

mit einer im Punkt 0 stetigen und dort verschwindenden Funktion ε.

Vernächlässigen wir den ε-Term, so sind die Koeffizienten a_k eindeutig bestimmt durch die Ableitungen von f an der Stelle a, denn

$$f^{(k)}(a) = \sum_{l=1}^{n} a_l \partial^k h^l \Big|_{h=0} = k!a_k, \qquad 1 \leqslant k \leqslant n.$$

Dies nehmen wir zum Anlass für folgende Definition.

Definition *Für $f \in C^n(I)$ und $a \in I$ heißt*

$$T_a^n f(t) := \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(t-a)^k$$

das n-te Taylorpolynom von f im Entwicklungspunkt a. ⋊

Insbesondere ist $T_a^0 f(t) = f(a)$ und

$$T_a^1 f(t) = f(a) + f'(a)(t-a),$$
$$T_a^2 f(t) = f(a) + f'(a)h + \frac{1}{2}f''(a)(t-a)^2,$$

und $T_a^1 f$ ist genau die Linearisierung von f im Punkt a.

Schreibt man $t = a + h$, so lautet das n-te Taylorpolynom

$$T_a^n f(a+h) = f(a) + f'(a)h + \frac{f''(a)}{2}h^2 + .. + \frac{f^{(n)}(a)}{n!}h^n.$$

8.21

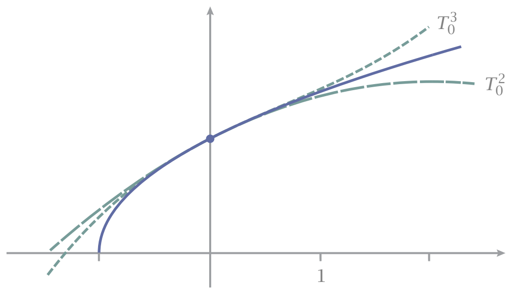


Abb 9

Taylorpolynome zu $\sqrt{1+t}$ im Punkt 0

Beispiele A. Für $f : t \mapsto \sqrt{1+t}$ und $a = 0$ ist

$$T_0^2 f(t) = 1 + \frac{t}{2} - \frac{t^2}{4}.$$

Übringens schreibt man $1+t$, um den Entwicklungspunkt in den Nullpunkt zu legen.

B. Für das Polynom $p : t \mapsto (1+t)^n$ gilt

$$\frac{p^{(k)}(0)}{k!} = \frac{n \cdots (n-k+1)}{k!} = \frac{n!}{k!(n-k)!} = \binom{n}{k}.$$

Also ist

$$T_0^n p(t) = \sum_{k=0}^{n} \binom{n}{k} t^k = (1+t)^n = p(t)$$

aufgrund der binomischen Formel $_{3.34}$.

C. Entwickeln wir p an einer anderen Stelle $a \neq 0$, so erhalten wir ein Polynom n-ter Ordnung mit anderen Koeffizienten.

Restgliedformel

Das Taylorpolynom allein sagt wenig aus, solange man den Approximationsfehler, das sogenannte *Restglied*

$$R_a^n f(t) = f(t) - T_a^n f(t)$$

nicht kontrollieren kann. Eine erste Abschätzung gibt der folgende Satz.

22 **Satz von Taylor mit Restglied von Lagrange** *Sei $f \in C^{n+1}(I)$ und $a \in I$. Dann existiert zu jedem $a + h \in I$ ein $\theta \in [0,1]$, so dass*

$$R_a^n f(a+h) = \frac{f^{(n+1)}(a+\theta h)}{(n+1)!} h^{n+1}.$$

⟨⟨⟨⟨⟨ Betrachte

$$R(h) = f(a+h) - T_a^n f(a+h), \qquad S(h) = h^{n+1},$$

als Funktionen von h. Alle Ableitungen von R und S im Punkt 0 bis zur Ordnung n verschwinden:

$$R^{(k)}(0) = S^{(k)}(0) = 0, \qquad 0 \leqslant k \leqslant n,$$

während

$$S^{(n+1)} = (n+1)! \neq 0.$$

Durch n-fache Anwendung des allgemeinen Mittelwertsatzes 13 erhalten wir eine Folge von Punkten $\tau_1, \ldots, \tau_{n+1}$ zwischen 0 und h mit

$$\begin{aligned} \frac{R(h)}{S(h)} &= \frac{R(h) - R(0)}{S(h) - S(0)} = \frac{R'(\tau_1)}{S'(\tau_1)} \\ &= \frac{R'(\tau_1) - R'(0)}{S'(\tau_1) - S'(0)} = \frac{R''(\tau_2)}{S''(\tau_2)} \\ &\ldots \\ &= \frac{R^{(n+1)}(\tau_{n+1})}{S^{(n+1)}(\tau_{n+1})} = \frac{f^{n+1}(a+\theta h)}{(n+1)!}, \end{aligned}$$

denn die $n+1$-te Ableitung des n-ten Taylorpolynoms verschwindet. Dies ist äquivalent zur Behauptung. ⟩⟩⟩⟩⟩

Eine Funktion $f \in C^{n+1}(I)$ kann man also lokal durch sein Taylorpolynom vom Grad n so approximieren, dass

$$f(t) = T_a^n f(t) + O((t-a)^{n+1}).$$

Hierbei steht der ›O-Ausdruck‹ für eine Funktion ϕ mit der Eigenschaft, dass

$$|\phi(h)| \leqslant M\,|h|^{n+1}$$

für alle kleinen $|h|$ mit einer Konstanten $M \geqslant 0$. Man sagt, ϕ *verschwindet bei Null mit der Ordnung* $n+1$. Diese Information ist für viele Zwecke bereits völlig ausreichend.

▹ *Beispiel* Die Bewegungsgleichung eines Oszillators mit Masse m ist gegeben durch

$$m\ddot{x} = -F(x),$$

wobei x die Auslenkung aus der Ruhelage bei $x = 0$, $\ddot{x}$ ihre zweite Ableitung nach der Zeit, und F die dort wirkende Rückstellkraft bezeichnen. Es gilt also

$F(x) \geqslant 0 = F(0)$, und Entwickeln bei 0 ergibt

$$F(x) = F(0) + F'(0)x + O(x^2) = \omega^2 x + O(x^2)$$

mit einer gewissen Konstanten $\omega^2 > 0$. Für kleine Auslenkungen kann man in erster Näherung den quadratischen Term vernachlässigen und erhält die Bewegungsgleichung für den harmonischen Oszillator,

$$m\ddot{x} = -\omega^2 x.$$ ◁

▷ *Beispiel* Die kinetische Energie eines relativistischen Teilchens mit Ruhemasse m_0 und Geschwindigkeit v ist

$$E_{\text{rel}} = m_0 c^2 \left(\frac{1}{\sqrt{1-(v/c)^2}} - 1 \right),$$

wobei c die Lichtgeschwindigkeit bezeichnet. Für die Funktion ϕ mit

$$\phi(t) = \frac{1}{\sqrt{1-t}} = (1-t)^{-1/2}, \qquad |t| < 1,$$

finden wir

$$\phi(t) = 1 + \frac{1}{2}t + \frac{3}{8}t^2 + O(t^3).$$

Somit erhalten wir mit $u = v^2/c^2$ näherungsweise

$$E_{\text{rel}} = m_0 c^2 \left(\frac{1}{2}u + \frac{3}{8}u^2 \right) = \frac{1}{2}m_0 v^2 + \frac{3}{8}m_0 \frac{v^4}{c^2}.$$ ◁

8.6 Taylorreihe

Ist f beliebig oft differenzierbar, so können wir Taylorpolynome jeder Ordnung bilden. Dies führt zum Begriff der *Taylorreihe.*

Definition *Für $f \in C^\infty(I)$ und $a \in I$ heißt*

$$T_a f(t) \coloneqq T_a^\infty f(t) \coloneqq \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(t-a)^k$$

die Taylorreihe von f am Entwicklungspunkt a. ⋊

Die Taylorreihe $T_a f$ konvergiert immer im Entwicklungspunkt a selbst und hat dort den Wert $f(a)$. Das ist trivial. Eine ganz andere Frage ist, ob sie auch in anderen Punkten konvergiert, und ob ihr Wert dort mit der Funktion selbst übereinstimmt.

Definition *Konvergiert die Taylorreihe von f in einer Umgebung von a gegen f, so heißt f um a entwickelbar in seine Taylorreihe.* ⋊

In diesem Fall gilt also

$$0 = \lim_{n\to\infty}(f(t) - T_a^n f(t)) = \lim_{n\to\infty} R_a^n f(t).$$

Somit erhalten wir sofort folgendes

23 **Entwicklungskriterium** *Es gilt $f(t) = T_a f(t)$ genau dann, wenn*

$$\lim_{n\to\infty} R_a^n f(t) = 0. \quad ⋊$$

Dies muss man nun für die jeweilige Funktion nachprüfen, so wie man auch die Konvergenz einer Reihe in jedem Fall einzeln prüfen muss. Ein Allzweckkriterium gibt es nicht. Am einfachsten ist noch folgende Situation, welche wir auch im nächsten Kapitel antreffen werden. — Zur Erinnerung: Es ist $\|\phi\|_I = \sup_{t\in I}|\phi(t)|$ die Supremumsnorm von ϕ über I.

24 **Satz** *Sei $f \in C^\infty(I)$. Gilt*

$$\sup_{n\geqslant 0} \sqrt[n]{\|f^{(n)}\|_I} < \infty,$$

so wird f durch seine Taylorreihe in jedem Punkt von I dargestellt. ⋊

⟨⟨⟨⟨⟨ Sei $a \in I$ und $r > 0$. Aufgrund der Annahme existiert ein $M \geqslant 0$, so dass $\|f^{(n)}\|_I < M^n$ für alle $n \geqslant 1$. Auf dem Intervall $I_r = U_r(a) \cap I$ gilt für das Restglied von Lagrange 22 dann

$$\|R_a^{n-1} f\|_{I_r} \leqslant \|f^{(n)}\|_I \frac{r^n}{n!} \leqslant \frac{M^n r^n}{n!}, \qquad n \geqslant 1.$$

Für $n \to \infty$ konvergiert dies gegen Null 5.12. ⟩⟩⟩⟩⟩

▪ Binomische Reihe

Als Beispiel betrachten wir die binomische Formel für beliebige reelle Exponenten. Wir greifen dabei der Definition von t^α für $t > 0$ und $\alpha \in \mathbb{R}$ 9.2 und der verallgemeinerten Ableitungsregel $(t^\alpha)' = \alpha t^{\alpha-1}$ vor.

25 **Binomische Reihe für reelle Exponenten** *Für jedes $\alpha \in \mathbb{R}$ gilt*

$$(1+t)^\alpha = 1 + \sum_{k\geqslant 1} B_k^\alpha t^k, \qquad |t| < 1,$$

mit den allgemeinen Binomialkoeffizienten

$$B_k^\alpha \coloneqq \frac{\alpha\cdot(\alpha-1)\cdots(\alpha-k+1)}{1\cdot 2\cdots k}, \qquad k \geqslant 1. \quad ⋊$$

Man beachte, dass die Identität nur für $|t| < 1$ gilt, während die Funktion selbst für alle $t > -1$ erklärt ist.

⟨⟨⟨⟨⟨ Die Funktion $f : t \mapsto (1+t)^\alpha$ ist auf $(-1, \infty)$ unendlich oft differenzierbar mit

$$f^{(k)}(t) = \alpha(\alpha-1)\cdot\cdot(\alpha-k+1)\,(1+t)^{\alpha-n}.$$

Also ist

$$\frac{f^{(k)}(0)}{k!} = \frac{\alpha(\alpha-1)\cdot\cdot(\alpha-k+1)}{1\cdot 2\cdot\cdot k} = B_k^\alpha.$$

Für das Restglied erhalten wir die Abschätzung

$$|R_0^{n-1}f(t)| = |B_n^\alpha|\,\frac{|t|^n}{|1+s_n|^{n-\alpha}}$$

mit s_n zwischen t und 0. Dies konvergiert gegen Null für $-1/2 < t < 1$. Für $-1 < t \leqslant -1/2$ benötigen wir allerdings eine bessere Restgliedabschätzung durch ein Integral $_{10.21}$. ⟩⟩⟩⟩⟩

▷ A. Für $\alpha = 0$ verschwinden alle Binomialkoeffizienten, und die Gleichung reduziert sich auf $(1+t)^0 = 1$, was ja stimmt.

B. Für natürliche Exponenten $\alpha = n \geqslant 1$ gilt $B_k^n = 0$ für $k > n$, und wir erhalten die klassische binomische Formel $_{3.34}$.

C. Für $\alpha = -1$ ist $B_k^\alpha = (-1)^k$ und somit

$$\frac{1}{1+t} = 1 + \sum_{k\geqslant 1}(-1)^k t^k = 1 - t + t^2 - t^3 \pm \cdot\cdot, \qquad |t| < 1.$$

Also gilt auch

$$\frac{1}{1-t} = 1 + \sum_{k\geqslant 1} t^k = 1 + t + t^2 + \cdot\cdot\ , \qquad |t| < 1,$$

wie es sich für die geometrische Reihe auch gehört.

D. Für $\alpha = 1/2$ erhalten wir

$$\sqrt{1+t} = 1 + \sum_{k\geqslant 1} b_k t^k = 1 + \frac{1}{2}t - \frac{1}{8}t^2 + \frac{1}{16}t^3 - \frac{5}{128}t^4 \pm \cdot\cdot$$

mit

$$b_k = \frac{(-1)^k}{2^k}\,\frac{(-1)\cdot 1\cdot 3\cdot\cdot\cdot(2k-3)}{1\cdot 2\cdot\cdot k}.$$ ◁

Funktionen in $C^\infty(I)$, die sich in jedem Punkt in ihre Taylorreihe entwickeln lassen, heißen *reell analytisch*. Ihre Klasse wird mit $C^\omega(I)$ bezeichnet. Diese Funktionen können also lokal immer durch eine Potenzreihe dargestellt werden.

8.26

Dies gilt aber nicht für jede Funktion in $C^\infty(I)$! Selbst wenn die Taylorreihe einer Funktion in einer Umgebung des Entwicklungspunktes konvergiert, so bedeutet dies keineswegs, dass sie auch die Funktion darstellt. Das klassische Beispiel hierfür geben wir in Abschnitt **??**.

Noch einmal Potenzreihen

Offen ist noch die Frage, ob die Summenfunktion einer Potenzreihe innerhalb ihres Konvergenzbereichs eine analytische Funktion darstellt. Dies ist tatsächlich der Fall. Die Grundlage dafür bildet der folgende Satz.

26 **Satz** *Jede Potenzreihe*

$$\phi(t) = \sum_{n\geqslant 0} a_n t^n$$

definiert im Innern ihres Konvergenzintervalls eine differenzierbare Funktion, deren Ableitung man durch gliedweises Differenzieren erhält:

$$\phi'(t) = \sum_{n\geqslant 1} n a_n t^{n-1}.$$

Diese Reihe hat denselben Konvergenzradius wie ϕ. ⋊

⟪⟪⟪ Wie im Fall der Potenzreihe selbst 6.18 zeigt man, dass die ϕ'-Reihe für $|t| < |t_0|$ konvergiert, wenn die ϕ-Reihe im Punkt t_0 konvergiert, und umgekehrt. Daher haben beide Reihen denselben Konvergenzradius A-34.

Für den Differenzenquotienten erhalten wir mit dem Mittelwertsatz 11

$$\frac{\phi(t+h)-\phi(t)}{h} = \sum_{n=0}^{\infty} a_n \frac{(t+h)^n - t^n}{h} = \sum_{n=1}^{\infty} n a_n s_n^{n-1}$$

mit Punkten $(s_n)_{n\geqslant 1}$ zwischen t und $t+h$. Mit

$$\psi(t) = \sum_{n\geqslant 1} n a_n t^{n-1}$$

erhalten wir also

$$\begin{aligned} \left| \frac{\phi(t+h)-\phi(t)}{h} - \psi(t) \right| &= \left| \sum_{n=1}^{\infty} n a_n (s_n^{n-1} - t^{n-1}) \right| \\ &\leqslant \sum_{n=1}^{\infty} n|a_n|\, |s_n^{n-1} - t^{n-1}| \\ &= \left\{ \sum_{1\leqslant n\leqslant N} + \sum_{n>N} \right\} n|a_n|\, |s_n^{n-1} - t^{n-1}|\,. \end{aligned}$$

Fur t und $t+h$ in einem abgeschlossenen Konvergenzintervall wird die zweite Summe kleiner als ein vorgegebenes $\varepsilon > 0$, indem man N hinreichend groß wählt. Anschließend wird die erste Summe kleiner als ε, indem man $|h| < \delta$ hinreichend klein wählt. Daher gilt

$$\lim_{h \to 0} \frac{\phi(t+h) - \phi(t)}{h} = \psi(t).$$

Also ist ϕ im Punkt t differenzierbar, und es ist $\phi'(t) = \psi(t)$. ⟫⟫⟫

Da der Konvergenzradius einer Reihe unter Differenziation derselbe bleibt, können wir diesen Vorgang beliebig oft wiederholen. Das Ergebnis ist der

27 **Potenzreihensatz** *Eine Potenzreihe*

$$\phi(t) = \sum_{n \geqslant 0} a_n t^n$$

mit positivem Konvergenzradius definiert im Innern ihres Konvergenzintervalles I eine analytische Funktion ϕ, deren Taylorreihe im Punkt 0 die Potenzreihe selbst ist:

$$T_0 \phi = \phi. \qquad \rtimes$$

⟪⟪⟪ Aus dem vorangehenden Satz folgt durch Induktion, dass eine Potenzreihe im Innern ihres Konvergenzintervalls beliebig oft differenzierbar ist, mit

$$\phi^{(k)}(t) = \sum_{n \geqslant k} \frac{n!}{(n-k)!} a_n t^{n-k}, \qquad k \geqslant 0.$$

Also ist $\phi^{(k)}(0) = k!\, a_k$ für alle $k \geqslant 0$ und damit

$$T_0 \phi(t) = \sum_{k \geqslant 0} a_k t^k = \phi(t). \qquad ⟫⟫⟫$$

Im nächsten Kapitel werden wir mithilfe von Potenzreihen einige der wichtigsten Funktionen der Analysis definieren.