7

Differentialrechnung

7.1 Grundbegriffe

Die Ableitung, auf neue Art betrachtet

Anmerkung: Es ist zu vermuten, dass der Leser schon von der Ableitung gehört hat. Zur Abwechslung und im Hinblick auf die mehrdimensionale Differentialrechnung beginnen wir daher die Sache etwas anders.

Es sei $I \subset \mathbb{R}$ ein Intervall und

$$f: \quad I \to \mathbb{X}, \qquad t \mapsto f(t) \tag{1}$$

eine $\mathbb{X}$-wertige (zum Beispiel reellwertige) Funktion der reellen Variablen t, die man als "Zeit" interpretieren kann. Wir halten einen "Arbeitspunkt" $t_0 \in I$ bis auf weiteres fest und machen den von t_0 aus gemessenen Zeitzuwachs $\Delta t = h$ zur neuen unabhängigen Variablen. Der zugehörige, von $y_0 := f(t_0)$ aus gemessenen **Wertzuwachs** Δf wird damit zu einer Funktion von h:

$$\Delta f(h) := f(t_0 + h) - f(t_0) .$$

In anderen Worten: Der Punkt (t_0, y_0) wird zum Ursprung eines $(\Delta t, \Delta y)$-Koordinatensystems gemacht, und der Wertzuwachs Δf wird an der Stelle $\Delta t = h$ nach oben abgetragen (Fig. 3.1.1). Dieses h hat man sich betragsmässig klein vorzustellen. Ist t_0 ein Endpunkt von I, so muss man sich auf $h \geq 0$ bzw. $h \leq 0$ beschränken.

Es ist ein fundamentales Faktum der Analysis, dass bei "guten" Funktionen der Wertzuwachs Δf im Limes $h \to 0$ *linear* von h abhängt: Es gibt eine Konstante A, die **momentane Zuwachsrate** von f an der Stelle t_0, mit

$$\Delta f(h) \doteq A\,h \qquad (h \to 0) . \tag{2}$$

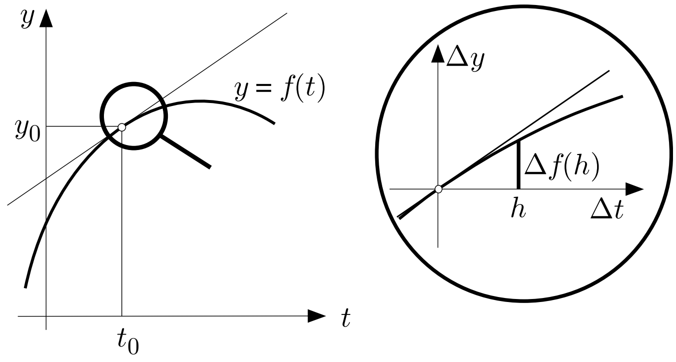

Fig. 7.1.1

Ist f vektorwertig, so sind auch Δf und damit A vektorwertig.

Wir müssen der Formel (2) einen präzisen Sinn erteilen. Die Aussage (2) besitzt nur dann einen tatsächlichen Gehalt, wenn der durch das Zeichen '$\doteq$' implizierte Fehler

$$r(h) := \Delta f(h) - A\,h$$

für $h \to 0$ wesentlich ("um Grössenordnungen") kleiner ist als der rechts in (2) hingeschriebene Term $A\,h$. Es müsste also

$$\frac{r(h)}{|A\,h|} \to 0 \qquad (h \to 0)$$

gelten, und das ist (im Normalfall $A \neq 0$) äquivalent mit $\dfrac{r(h)}{h} \to 0 \ \ (h \to 0)$.

Aufgrund dieser Überlegungen definieren wir definitiv: Die Funktion (1) heisst **an der Stelle** $t_0 \in I$ **differenzierbar**, wenn es eine Konstante $A \in \mathbb{X}$ $\big($und eine Funktion $r(\cdot)\big)$ gibt, so dass folgendes gilt:

$$f(t_0 + h) - f(t_0) \;=\; A\,h + r(h)\,, \qquad \lim_{h \to 0} \frac{r(h)}{h} = 0\;. \tag{3}$$

Die Gleichung links in (3) definiert die Funktion $r(\cdot)$, und diese Funktion muss der entscheidenden Bedingung rechts in (3) genügen.

(7.1) (a) *Ist f an der Stelle t_0 differenzierbar im Sinne von* (3), *so existiert der Grenzwert*

$$\lim_{h \to 0} \frac{f(t_0 + h) - f(t_0)}{h} \;=\; \lim_{t \to t_0} \frac{f(t) - f(t_0)}{t - t_0} =: f'(t_0)\,, \tag{4}$$

und es gilt $A = f'(t_0)$.

(b) *Umgekehrt: Existiert der Grenzwert* (4), *so ist* f *an der Stelle* t_0 *differenzierbar im Sinne von* (3), *und zwar gilt*

$$f(t_0 + h) - f(t_0) = f'(t_0)\, h + r(h)\,, \qquad \lim_{h\to 0} \frac{r(h)}{h} = 0\;. \tag{5}$$

⌜ (a) Aus (3) folgt

$$\frac{f(t_0+h) - f(t_0)}{h} - A = \frac{r(h)}{h} \to 0 \qquad (h \to 0)\,,$$

somit existiert die Ableitung $f'(t_0)$ und hat den Wert A. — (b) Die Funktion

$$r(h) := f(t_0 + h) - f(t_0) - f'(t_0)\, h$$

genügt der Gleichung links in (5), und man hat

$$\frac{r(h)}{h} = \frac{f(t_0+h) - f(t_0)}{h} - f'(t_0) \to 0 \qquad (h \to 0)\;.$$

Somit gilt (3) mit $A := f'(t_0)$. ⌟

Der Grenzwert (4) der **Differenzenquotienten**

$$\frac{f(t_0+h) - f(t_0)}{h} \quad \text{bzw.} \quad \frac{f(t) - f(t_0)}{t - t_0}$$

heisst bekanntlich **Ableitung von** f **an der Stelle** t_0. Die **einseitigen Ableitungen** $f'(t_0+)$ bzw. $f'(t_0-)$ werden sinngemäss erklärt.

Gelegentlich ist die folgende (nennerfreie!) Charakterisierung der Differenzierbarkeit nützlich:

(7.2) (a) *Ist die Funktion* (1) *an der Stelle* $t_0 \in I$ *differenzierbar, so gilt*

$$f(t) - f(t_0) = m(t)\,(t - t_0) \qquad (t \in I) \tag{6}$$

mit einer an der Stelle t_0 *stetigen Funktion* $m(\cdot)\colon I \to \mathbb{X}$, *die dort den Wert* $f'(t_0)$ *annimmt.*

(b) *Umgekehrt: Gilt* (6) *mit einer an der Stelle* t_0 *stetigen Funktion* $m(\cdot)$, *so ist* f *an der Stelle* t_0 *differenzierbar und besitzt dort die Ableitung* $f'(t_0) = m(t_0)$.

⌜ (a) Die Funktion

$$m(t) := \begin{cases} \dfrac{f(t) - f(t_0)}{t - t_0} & (t \neq t_0), \\ f'(t_0) & (t = t_0) \end{cases}$$

besitzt die behaupteten Eigenschaften. — (b) Aus (6) folgt

$$\frac{f(t) - f(t_0)}{t - t_0} = m(t) \to m(t_0) \qquad (t \to t_0)\;.$$ ⌟

Wir nennen die in Satz **(7.2)** auftretende Funktion $m(\cdot)$ die **Trendfunktion von** f **an der Stelle** t_0 und bezeichnen sie, wenn nötig, mit m_{f,t_0}. — Aus **(7.2)**(a) ziehen wir noch den folgenden Schluss:

(7.3) *Ist f an der Stelle t_0 differenzierbar, so ist f dort stetig.*

⌜ Aus (6) folgt $f(t) = f(t_0) + m(t)(t - t_0)$, und hier ist die rechte Seite an der Stelle t_0 stetig. ⌟

Wie das nachfolgende Beispiel ① zeigt, ist die Umkehrung des Satzes **(7.3)** *falsch*.

Die Ableitung als Funktion f'

Die Punkte $t_0 \in I$, für die $f'(t_0)$ existiert, bilden den Definitionsbereich der Funktion

$$f' : \qquad t \mapsto f'(t) ,$$

genannt **Ableitung** von f. Die Ableitung f' gibt für jeden "Zeit"punkt $t \in \operatorname{dom}(f')$ die momentane Zuwachsrate der Ausgangsfunktion f an. Anstelle von f' sind auch Bezeichnungen wie $\dot{f}$, Df, $\dfrac{df}{dt}$ und andere in Gebrauch. Ist $\operatorname{dom}(f') = \operatorname{dom}(f)$, so heisst die Funktion f **differenzierbar**.

① Betrachte die (jedenfalls stetige) Funktion

$$\operatorname{abs} t := |t| \qquad (t \in \mathbb{R}) .$$

Ist zunächst $t_0 \neq 0$, so besitzen alle hinreichend nahe bei t_0 gelegenen t dasselbe Vorzeichen wie t_0 (Fig. 7.1.2). Für diese t gilt daher

$$\operatorname{abs} t = \operatorname{sgn} t \cdot t = \operatorname{sgn} t_0 \cdot t .$$

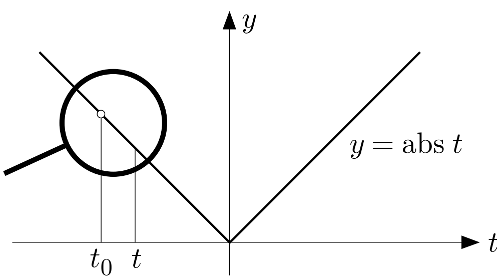

Fig. 7.1.2

Damit erhalten wir

$$\frac{\operatorname{abs} t - \operatorname{abs} t_0}{t - t_0} = \frac{\operatorname{sgn} t_0 \cdot t - \operatorname{sgn} t_0 \cdot t_0}{t - t_0} = \operatorname{sgn} t_0 \qquad (t \neq t_0) ,$$

und es folgt

$$\text{abs}'(t_0) = \operatorname{sgn} t_0 \qquad (t_0 \neq 0) \ .$$

Ist jedoch $t_0 = 0$, so hat man

$$\frac{\text{abs}\, t - \text{abs}\, t_0}{t - t_0} = \frac{|t| - |0|}{t - 0} = \operatorname{sgn} t \qquad (t \neq 0) \ .$$

Der $\lim_{t \to 0} \operatorname{sgn} t$ existiert nicht; folglich ist abs an der Stelle 0 nicht differenzierbar. Hingegen existieren dort die einseitigen Ableitungen

$$\text{abs}'(0+) = \operatorname{sgn}(0+) = 1 \ , \qquad \text{abs}'(0-) = \operatorname{sgn}(0-) = -1 \ .$$

Damit hat sich ergeben:

$$\text{abs}' \ = \ \operatorname{sgn}\restriction \dot{\mathbb{R}} \ .$$

○

② Für die Exponentialfunktion hat man

$$\frac{e^t - e^{t_0}}{t - t_0} = e^{t_0}\, \frac{e^{t - t_0} - 1}{t - t_0} \qquad (t \neq t_0)$$

und somit nach **(5.21)**: $\exp'(t_0) = e^{t_0}$. Da dies für alle $t_0 \in \mathbb{R}$ zutrifft, gilt

$$\exp' \ = \ \exp \qquad \text{bzw.} \qquad \frac{d}{dt}\, e^t \ = \ e^t \ .$$

Wir sehen: Die Exponentialfunktion wird durch Differentiation reproduziert.

○

③ Eine konstante Funktion $c(\cdot)$ besitzt die Ableitung $c'(t) \equiv 0$, die Funktion $\text{id}(t) :\equiv t$ besitzt die Ableitung $\text{id}'(t) \equiv 1$; denn für alle $t_0 \in \mathbb{R}$ gilt

$$\lim_{t \to t_0} \frac{c(t) - c(t_0)}{t - t_0} = 0 \ , \qquad \lim_{t \to t_0} \frac{\text{id}(t) - \text{id}(t_0)}{t - t_0} = 1 \ .$$

○

Geometrische Interpretation der Ableitung

Im Fall einer "Funktion $y = f(x)$" ist der Differenzenquotient

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

die Steigung der Sekante durch die Graphenpunkte $P_0 := \big(x_0, f(x_0)\big)$ und $P := \big(x, f(x)\big)$ (Fig. 7.1.3). Strebt x gegen x_0, so wandert P auf dem Graphen gegen den Punkt P_0, und die Sekante durch P_0 und P geht über in die Graphentangente im Punkt P_0. Die Steigung dieser Tangente ist gleich dem Grenzwert der Sekantensteigungen, also gleich $f'(x_0)$.

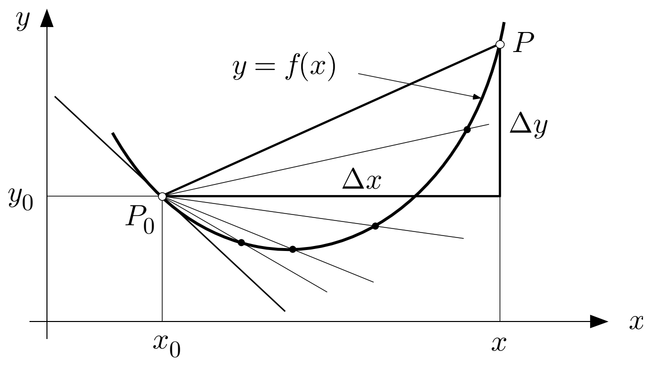

Fig. 7.1.3

Bei vektorwertigen Funktionen lässt sich die Ableitung folgendermassen geometrisch interpretieren (Fig. 7.1.4):

(7.4) *Die Funktion*

$$\mathbf{x}(\cdot): \quad I \to \mathbb{X}, \qquad t \mapsto \mathbf{x}(t) \tag{7}$$

werde als Parameterdarstellung einer Kurve $\gamma \subset \mathbb{X}$ *aufgefasst. Ist* $\mathbf{x}'(t_0) \neq \mathbf{0}$, *so gilt*

$$\frac{\mathbf{x}'(t_0)}{|\mathbf{x}'(t_0)|} = \lim_{h \to 0+} \frac{\Delta\mathbf{x}}{|\Delta\mathbf{x}|}\ ,$$

wobei $\Delta\mathbf{x}/|\Delta\mathbf{x}|$ *den normierten Sekantenvektor bezeichnet. In anderen Worten: Der Ableitungsvektor* $\mathbf{x}'(t_0)$ *zeigt in die Richtung der (als Grenzlage von Sekanten aufgefassten) Tangente an* γ *im Punkt* $\mathbf{x}(t_0)$.

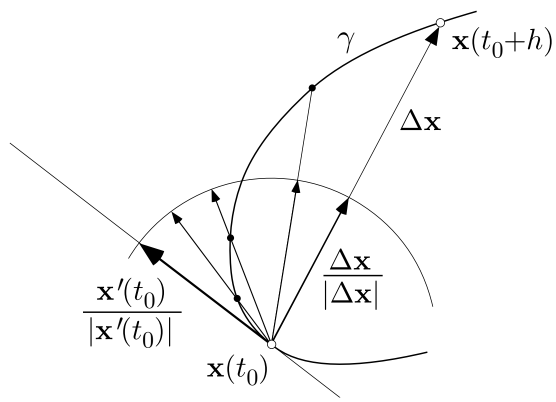

Fig. 7.1.4

⌜ Wir benützen die Proposition **(7.2)** mit $t := t_0 + h$. Wegen $h > 0$ bzw. $h = |h|$ ergibt sich

$$\frac{\Delta\mathbf{x}}{|\Delta\mathbf{x}|} = \frac{\mathbf{x}(t_0+h) - \mathbf{x}(t_0)}{|\mathbf{x}(t_0+h) - \mathbf{x}(t_0)|} = \frac{\mathbf{m}(t_0+h)\,h}{|\mathbf{m}(t_0+h)\,h|} \to \frac{\mathbf{x}'(t_0)}{|\mathbf{x}'(t_0)|} \qquad (h \to 0+)\ .$$ ⌟

Wird die unabhängige Variable t der Parameterdarstellung (7) als “Zeit” interpretiert, so stellt der Ableitungsvektor $\mathbf{x}'(t_0)$ eine Geschwindigkeit dar: Der Differenzenquotient

$$\frac{\Delta\mathbf{x}}{\Delta t} := \frac{\mathbf{x}(t) - \mathbf{x}(t_0)}{t - t_0} \qquad (t > t_0)$$

ist die “Ortsveränderung pro Zeiteinheit”, anders ausgedrückt: die **mittlere Geschwindigkeit** im Zeitintervall $[\,t_0, t\,]$, und der für $t \to t_0+$ resultierende Grenzwert $\mathbf{x}'(t_0+) = \mathbf{x}'(t_0) =: \mathbf{v}(t_0)$ ist die **Momentangeschwindigkeit** oder kurz: **Geschwindigkeit** des laufenden Punktes im Zeitpunkt t_0. Nach **(7.4)** hat $\mathbf{v}(t_0)$ die Richtung der (gerichteten) Bahntangente im Punkt $\mathbf{x}(t_0)$.

5 Betrachte die Parameterdarstellung

$$\mathbf{x}(t) \;:=\; t\mathbf{a} + t^2\mathbf{b} \qquad (-\infty < t < \infty)\ .$$

Es ergibt sich

$$\frac{\mathbf{x}(t) - \mathbf{x}(t_0)}{t - t_0} = \frac{(t - t_0)\mathbf{a} + (t^2 - t_0^2)\mathbf{b}}{t - t_0} = \mathbf{a} + (t + t_0)\mathbf{b}$$

und somit

$$\mathbf{v}(t_0) = \mathbf{x}'(t_0) = \lim_{t \to t_0} \big(\mathbf{a} + (t + t_0)\mathbf{b}\big) = \mathbf{a} + 2t_0\mathbf{b}\ .$$

Sind die Vektoren $\mathbf{a}$ und $\mathbf{b}$ linear unabhängig, so ist $\mathbf{x}'(t_0) \neq \mathbf{0}$ für alle $t_0 \in \mathbb{R}$; die Bahnkurve (eine Parabel) besitzt daher in allen Punkten eine Tangente. Ist aber zum Beispiel $\mathbf{a} = \mathbf{0}$ und $\mathbf{b} \neq \mathbf{0}$, so besitzt die Bahnkurve (eine zweimal durchlaufene Halbgerade) im Ursprung einen Rückkehrpunkt. Der in **(7.4)** ausdrücklich wegbedungene Sachverhalt $\mathbf{x}'(0) = \mathbf{0}$ hat prompt zu einer geometrischen Singularität der Bahnkurve geführt. ○

Exkurs über die *o*-Notation

Zum Schluss dieses Abschnitts wenden wir uns nocheinmal den Formeln (3) zu. Es geht dort um die Grössenordnung einer gewissen Fehlerfunktion $r(h)$ beim Grenzübergang $h \to 0$. Mit Hilfe des **Landauschen *o*-Symbols** lassen sich derartige Sachverhalte in besonders kompakter Weise ausdrücken; man muss sich allerdings ein wenig an diese *o*-Schreibweise gewöhnen.

Also: Da der Quotient $\dfrac{r(h)}{h}$ mit $h \to 0$ gegen 0 geht, sagt man, $r(h)$ sei “klein oh von h”, und schreibt

$$r(h) = o(h) \qquad (h \to 0)\ .$$

Der Term $o(h)$ bezeichnet hier *nicht* einen Funktionswert an der Stelle h, sondern erklärt $r(\cdot)$ zu einer Funktion, von der man folgendes weiss: Dividiert man diese Funktion durch das in der o-Klammer angegebene h, so geht der Quotient mit $h \to 0$ gegen 0.

Allgemein: Es ist die Rede von einem bestimmten Grenzübergang $x \to \xi$. Ein Term $o\bigl(q(x)\bigr)$ in einer Gleichung bezeichnet eine letzten Endes durch diese Gleichung definierte Funktion $R(x)$, *von der man aber weiss*, dass sie für $x \to \xi$ von wesentlich kleinerer Grössenordnung ist als der angegebene Funktionsterm $q(x)$. Damit ist genau folgendes gemeint: Der Quotient $R(x)/q(x)$ geht mit $x \to \xi$ gegen 0. In anderen Worten: Die nennerfreie Formel

$$f(x) = g(x) + o\bigl(q(x)\bigr) \qquad (x \to \xi)$$

ist äquivalent mit dem Sachverhalt

$$\lim_{x \to \xi} \frac{f(x) - g(x)}{q(x)} = 0 \ .$$

Bsp:

$$\begin{aligned}
&\frac{1}{t} = o(1) \qquad (t \to \infty)\ ,\\
&t^{1000} = o(e^t) \qquad (t \to \infty)\ ,\\
&t^n + a_{n-1}t^{n-1} + \ldots + a_0 = t^n\,\bigl(1 + o(1)\bigr) \qquad (t \to \infty)\ ,\\
&\frac{3t^2 - 5t - 7}{t+1} = 3t - 8 + o(1) \qquad (t \to \infty)\ ,\\
&\sqrt{1+t} = 1 + \frac{t}{2} + o(t) \qquad (t \to 0)\ .
\end{aligned}$$

(Die zweitletzte Beziehung ergibt sich durch Ausführung der Polynomdivision; die letzte ist ein Anwendungsfall der nachstehenden Formel (8).)

Wegen $A = f'(t_0)$ sind wir damit in der Lage, den Inhalt von (3) in einer einzigen nennerfreien Formel auszudrücken:

$$\boxed{f(t_0 + h) - f(t_0) = f'(t_0)\,h + o(h) \qquad (h \to 0)} \tag{8}$$

Aufgaben

1. Für welche $\alpha \in \mathbb{R}$ ist die Funktion

$$f_\alpha(t) := \begin{cases} t^\alpha \sin(1/t) & (t > 0) \\ 0 & (t = 0) \end{cases}$$

an der Stelle $t = 0$ bzw. im Intervall $[\,0, 1\,]$

(a) beschränkt, (b) stetig, (c) differenzierbar,

(d) stetig differenzierbar, (e) von beschränkter Ableitung?

2. (a) Die Funktionen f, $g : I \to \mathbb{R}$ seien differenzierbar, und es sei $h(t) := \max\{f(t), g(t)\}$. Zeige: Die Funktion h ist in allen Punkten $c \in I$ mit $f(c) \neq g(c)$ differenzierbar, in den Punkten c mit $f(c) = g(c)$ genau dann, wenn $f'(c) = g'(c)$ ist.

   (b) Formuliere und beweise eine analoge Aussage für die Funktion $k(t) := \bigl|f(t)\bigr|$.

3. Ausgehend von der stetig differenzierbaren Funktion $f : I \to \mathbb{R}$ wird die Funktion

$$\phi(x, y) := \begin{cases} \dfrac{f(y) - f(x)}{y - x} & (y \neq x) \\ f'(x) & (y = x) \end{cases}$$

gebildet. Zeige: Die Funktion ϕ ist auf $I \times I$ stetig. *Hinweis:* Zeige, dass ϕ in jedem einzelnen Punkt (x_0, y_0) stetig ist. Dabei sind verschiedene Fälle zu unterscheiden.

4. Die Ableitung einer differenzierbaren Funktion genügt automatisch dem Zwischenwertsatz: Ist $f : [\,a, b\,] \to \mathbb{R}$ differenzierbar und $f'(a) < \sigma < f'(b)$, so gibt es ein $\xi \in\,]\,a, b\,[$ mit $f'(\xi) = \sigma$. *Hinweis:* Betrachte an Stelle von f die Hilfsfunktion $g(t) := f(t) - \sigma t$. Diese Funktion ist weder bei a noch bei b lokal minimal.

5. Die Funktion $f : [\,a, b\,] \to \mathbb{R}$ sei stetig und auf $]\,a, b\,[$ differenzierbar. Existiert dann der $\lim_{t \to a+} f'(t) =: \alpha$, so ist f auch an der Stelle a differenzierbar, und es gilt $f'(a) = \alpha$.

7.2 Rechenregeln

Die Ableitungsoperation $(d/dt)\colon f \mapsto f'$ genügt verschiedenen Rechenregeln. Deren Beweis stützt sich natürlich auf die Definition der Ableitung an einer festen Stelle t_0. Einmal bewiesen, dienen diese Regeln nicht zuletzt zur "mechanischen" Ableitung von Funktionstermen — siehe dazu die Übungsaufgaben.

(7.5) *Eine vektorwertige Funktion* $\mathbf{f} = (f_1, \ldots, f_m)$ *ist genau dann (an der Stelle* t_0*) differenzierbar, wenn jede einzelne Koordinatenfunktion* f_i *(dort) differenzierbar ist. Es gilt*

$$\mathbf{f}' = \bigl(f_1', \ldots, f_m'\bigr) \ ;$$

insbesondere ist $(\operatorname{Re} f)' = \operatorname{Re}(f')$ *und* $(\operatorname{Im} f)' = \operatorname{Im}(f')$.

⌜ Die Behauptung folgt unmittelbar aus

$$\frac{\mathbf{f}(t) - \mathbf{f}(t_0)}{t - t_0} = \left(\frac{f_1(t) - f_1(t_0)}{t - t_0}, \ \ldots\ , \frac{f_m(t) - f_m(t_0)}{t - t_0}\right)$$

und Regel **(3.14)**(a). ⌟

(7.6) *Wo immer die rechte Seite der folgenden Gleichungen definiert ist, gilt*

(a) $$(f+g)' = f' + g' \ , \qquad (\alpha f)' = \alpha\, f' \quad (\alpha \in \mathbb{X}) \ ,$$

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

(c) $$(f/g)' = \frac{f'g - fg'}{g^2} \ .$$

(d) *Ist* f *reellwertig, so gilt* $(g \circ f)' = (g' \circ f) \cdot f'$ *bzw.*

$$\frac{d}{dt}\, g\bigl(f(t)\bigr) = g'\bigl(f(t)\bigr) \cdot f'(t) \ .$$

(e) *Für beliebige reell- oder komplexwertige Funktionen* f *gilt*

$$\frac{d}{dt}\, e^{f(t)} \ = \ f'(t)\, e^{f(t)} \ .$$

Regel (b) ist die **Produktregel**; sie gilt für alle in den Grundstrukturen erklärten Produkte. Die **Quotientenregel** (c) gilt auch für komplexwertige

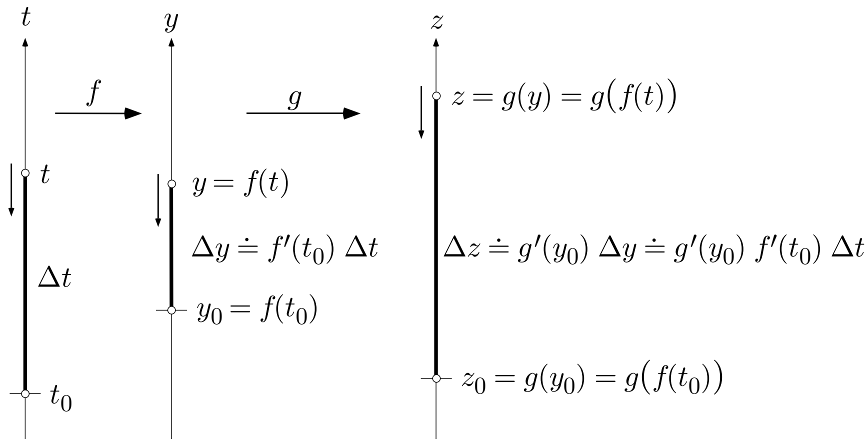

Fig. 7.2.1

Funktionen und für vektorwertige Ausdrücke der Form $\mathbf{f}(t)/g(t)$ mit einem reellwertigen Nenner. Die **Kettenregel** (d) bezieht sich auf die fundamentalste "Verknüpfungsoperation", nämlich $\circ$, und ist damit letzten Endes die wichtigste. Sie dient natürlich als Regel für das Ableiten von zusammengesetzten Funktionstermen, besitzt aber auch einen geometrischen Gehalt, der in der Fig. 7.2.1 dargestellt ist.

⌜ Es genügt, diese Regeln für ein festes $t_0 \in \operatorname{dom}(f)$ zu beweisen. — (a) ist klar. — Aufgrund von **(7.2)**(a) hat man

$$\begin{aligned} f(t)g(t) - f(t_0)g(t_0) &= (f(t) - f(t_0))g(t) + f(t_0)(g(t) - g(t_0)) \\ &= m_f(t)(t - t_0)\, g(t) + f(t_0)\, m_g(t)(t - t_0) \\ &= \bigl(m_f(t)g(t) + f(t_0)m_g(t)\bigr)(t - t_0)\ . \end{aligned}$$

Mit **(7.3)** und **(7.2)**(b) folgt hieraus (b). — Für (c) müssen wir $g(t_0) \neq 0$ voraussetzen. Dann ist $1/g$ in einer ganzen Umgebung von t_0 definiert, und es gilt auf Grund von **(7.2)**(a):

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

Wie bei (b) folgt hieraus $(1/g)' = -g'/g^2$, und wegen $f/g = f \cdot (1/g)$ lässt sich damit (c) auf (b) zurückführen. — Zum Beweis der Kettenregel wenden wir ebenfalls das Prinzip **(7.2)** an. Wir setzen $f(t_0) =: y_0$ und schreiben

$$g(y) - g(y_0) = m_{g,y_0}(y)(y - y_0)\ , \qquad f(t) - f(t_0) = m_{f,t_0}(t)(t - t_0)\ .$$

Damit ergibt sich

$$\begin{aligned} g\bigl(f(t)\bigr) - g\bigl(f(t_0)\bigr) &= g\bigl(f(t)\bigr) - g(y_0) = m_{g,y_0}\bigl(f(t)\bigr)\,(f(t) - y_0) \\ &= m_{g,y_0}\bigl(f(t)\bigr)\, m_{f,t_0}(t)\,(t - t_0)\ . \end{aligned}$$

Hieraus folgt die Behauptung auf Grund der Rechenregeln über stetige Funktionen. — Zum Beweis von (e) benutzen wir die an der Stelle 0 stetige Hilfsfunktion

$$p(z) := \begin{cases} \dfrac{e^z - 1}{z} & (z \neq 0) \\ 1 & (z = 0) \end{cases} .$$

Für alle $z \in \mathbb{C}$ gilt $e^z - 1 = p(z) \cdot z$. Damit ergibt sich

$$\begin{aligned} e^{f(t)} - e^{f(t_0)} &= e^{f(t_0)} \left(e^{f(t) - f(t_0)} - 1 \right) = e^{f(t_0)} \, p\big(f(t) - f(t_0)\big)\big(f(t) - f(t_0)\big) \\ &= e^{f(t_0)} \, p\big(f(t) - f(t_0)\big) m_{f,t_0}(t) \, (t - t_0) \\ &=: \tilde{m}(t) \, (t - t_0) \ . \end{aligned}$$

Hier ist der Faktor $\tilde{m}(\cdot)$ an der Stelle t_0 stetig und hat dort wegen $p(0) = 1$ den erforderlichen Wert $e^{f(t_0)} f'(t_0)$. $\lrcorner$

Eine weitere Regel bezieht sich auf die Umkehrfunktion einer streng monotonen reellwertigen Funktion:

(7.7) *Es sei $f : I \to \mathbb{R}$ streng monoton und differenzierbar auf dem Intervall I. Dann ist die Umkehrfunktion $g := f^{-1}$ in allen Punkten $y \in f(I)$, für die f' im zugehörigen Punkt $t \in I$ nicht verschwindet, differenzierbar, und zwar gilt*

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

$\ulcorner$ Betrachte ein festes $y_0 \in f(I)$ und setze $g(y_0) =: t_0$. Es gilt $f(t) - f(t_0) = m_f(t)(t - t_0)$ und somit

$$y - y_0 = f\big(g(y)\big) - f\big(g(y_0)\big) = m_f\big(g(y)\big) \cdot \big(g(y) - g(y_0)\big) \ .$$

Hieraus folgt

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{1}{m_f(g(y))} \qquad (y \neq y_0) \ .$$

Da g nach Satz **(4.24)** stetig ist, ergibt sich mit $y \to y_0$ die Behauptung. $\lrcorner$

Die Ableitung der elementaren Grundfunktionen

Die Sätze **(7.6)** und **(7.7)** setzen uns in Stand, die Ableitungen der "elementaren Grundfunktionen" zu berechnen:

(7.8)(a) $$\frac{d}{dt}t^k = k\,t^{k-1} \qquad (k \in \mathbb{Z})\,,$$

(b) $$\frac{d}{dt}e^t = e^t\,,$$

(c) $$\frac{d}{dy}\log y = \frac{1}{y}\,,$$

(d) $$\frac{d}{dt}t^\alpha = \alpha\,t^{\alpha-1} \qquad (\alpha \in \mathbb{R})\,,$$

(e) $$\frac{d}{dt}\cosh t = \sinh t\,, \qquad \frac{d}{dt}\sinh t = \cosh t\,,$$

(f) $$\frac{d}{dt}\tanh t = 1 - \tanh^2 t = \frac{1}{\cosh^2 t}\,,$$

(g) $$\frac{d}{dy}\operatorname{arcosh} y = \frac{1}{\sqrt{y^2-1}}\,, \qquad \frac{d}{dy}\operatorname{arsinh} y = \frac{1}{\sqrt{y^2+1}}\,,$$

(h) $$\frac{d}{dy}\operatorname{artanh} y = \frac{1}{1-y^2} \qquad (-1 < y < 1)\,,$$

(i) $$\frac{d}{dt}\cos t = -\sin t\,, \qquad \frac{d}{dt}\sin t = \cos t\,,$$

(j) $$\frac{d}{dt}\tan t = 1 + \tan^2 t = \frac{1}{\cos^2 t}\,,$$

(k) $$\frac{d}{dy}\arcsin y = \frac{1}{\sqrt{1-y^2}}\,,$$

(l) $$\frac{d}{dy}\arctan y = \frac{1}{1+y^2}\,.$$

⌜ (a) Für $k = 0$ und $k = 1$ siehe Beispiel 7.1.③. Die Produktregel erlaubt den Schluss von k auf $k+1$:

$$\frac{d}{dt}(t^{k+1}) = \frac{d}{dt}(t^k \cdot t) = \frac{d}{dt}(t^k)\cdot t + t^k \cdot \frac{d}{dt}t = kt^{k-1}\cdot t + t^k \cdot 1 = (k+1)t^k\,,$$

und mit Hilfe der Quotientenregel lässt sich der Fall $k < 0$ auf den schon erledigten Fall $k > 0$ zurückspielen:

$$\frac{d}{dt}\frac{1}{t^k} = \frac{-kt^{k-1}}{t^{2k}} = \frac{-k}{t^{k+1}} \ .$$

(c) Der Logarithmus ist die Umkehrfunktion von $\exp\restriction\mathbb{R}$. Da $\exp'(t) = e^t$ nirgends verschwindet, ergibt sich mit **(7.7)**:

$$\log'(y) = \frac{1}{\exp'(\log y)} = \frac{1}{\exp(\log y)} = \frac{1}{y} \ .$$

(d) Im Bereich $t > 0$ gilt $t^\alpha = \exp(\alpha \log t)$. Somit folgt mit Hilfe der Kettenregel:

$$\frac{d}{dt}t^\alpha = \exp'(\alpha \log t) \cdot \alpha \ \log'(t) = \exp(\alpha \log t) \cdot \frac{\alpha}{t} = t^\alpha \cdot \frac{\alpha}{t} = \alpha t^{\alpha-1} \ .$$

Für $\alpha \geq 1$ ist die Funktion $p_\alpha : t \mapsto t^\alpha$ auch noch an der Stelle $t = 0$ differenzierbar. Man erhält

$$p'_\alpha(0) = \lim_{t\to 0+} \frac{t^\alpha - 0}{t - 0} = \lim_{t\to 0+} t^{\alpha-1} = \begin{cases} 1 & (\alpha = 1) \\ 0 & (\alpha > 1) \end{cases} \qquad ,$$

und dies ist jedenfalls $= \alpha \cdot 0^{\alpha-1}$.

Die Formeln (e) folgen unmittelbar aus (b) und der Definition von cosh und sinh, und (f) ergibt sich aus

$$(\sinh / \cosh)' = (\sinh' \cosh - \cosh' \sinh)/ \cosh^2 = (\cosh^2 - \sinh^2)/ \cosh^2 \ .$$

(g) Der Areacosinus ist die Umkehrfunktion von $\cosh\restriction[\,0, \infty\,[$. Für alle $t > 0$ gilt $\cosh'(t) = \sinh t > 0$; folglich ergibt sich mit **(7.7)** und der Identität **(6.8)**(a):

$$\frac{d}{dy}\operatorname{arcosh} y = \frac{1}{\sinh(\operatorname{arcosh} y)} = \frac{1}{\sqrt{y^2 - 1}} \qquad (y > 1)\ ;$$

analog schliesst man für den Areasinus. — In ähnlicher Weise werden (k) und (l) bewiesen.

(i) Wird die Identität $\cos t + i \sin t = e^{it}$ abgeleitet, so ergibt sich nach **(7.5)** und **(7.6)**(f):

$$\cos'(t) + i \sin'(t) = ie^{it} = i(\cos t + i \sin t) = -\sin t + i \cos t \ .$$

Mit Hilfe der Quotientenregel folgt hieraus schliesslich (j). ┘

① Die Funktion
$$F(t) := e^{\sqrt{1-t^2}}$$
ist aus den drei Funktionen
$$f : t \mapsto 1 - t^2 =: y\,, \qquad g : y \mapsto \sqrt{y} =: z\,, \qquad h : z \mapsto e^z$$
zusammengesetzt. Ihre Ableitung berechnet sich zu
$$\begin{aligned} F'(t) &= h'\bigl(g(f(t))\bigr) \cdot g'\bigl(f(t)\bigr) \cdot f'(t) \\ &= e^{\sqrt{1-t^2}} \frac{1}{2\sqrt{1-t^2}} (-2t) = \frac{-t}{\sqrt{1-t^2}}\, e^{\sqrt{1-t^2}}\ . \end{aligned}$$
○

Beachte, dass transzendente Funktionen, zum Beispiel arcsin oder log, durchaus algebraische oder sogar rationale Funktionen als Ableitungen haben können. Umgekehrt werden uns die obigen Formeln bei der Integration der rationalen Funktionen und von Wurzelausdrücken weiterhelfen.

Funktionen, die sich mit Hilfe der vier Grundrechenarten und Zusammensetzen aus Konstanten, t^α $(\alpha \in \mathbb{R})$, $\log t$, e^t, $\cos t$, $\sin t$ sowie deren Umkehrfunktionen erhalten lassen, gelten als **elementare Funktionen**.

Bsp:
$$f(t) := \frac{e^{\sqrt{1-\log^2 t}} \cos(\sin t)}{\pi + t^{1/5}}$$
Die Ableitung einer derartigen Funktion ist wieder eine elementare Funktion — dies folgt mit vollständiger Induktion aus den Ableitungsregeln. Es gibt aber elementare Funktionen, deren "Stammfunktionen" (s.u.) nicht elementar sind; hierzu gehören zum Beispiel die Funktion $\sin t/t$ oder die in der Wahrscheinlichkeitstheorie dauernd gebrauchte Funktion $e^{-t^2/2}$. Der Umfang einer Ellipse ist keine elementare Funktion der Halbachsen (sonst hätten Sie die Formel schon gesehen!).

Funktionen der Klasse C^r

Im folgenden ist $I \subset \mathbb{R}$ ein beliebiges Intervall und $f : I \to \mathbb{X}$ eine gegebene Funktion. Die **höheren Ableitungen** von f sind formal definiert durch
$$f^{(0)} := f\,, \qquad f^{(k+1)} := \Bigl(f^{(k)}\Bigr)' \quad (k \geq 0)\ .$$
An Stelle von $f^{(k)}$ schreibt man auch $d^k f/dt^k$, $D^k f$ oder ähnlich. Die Funktion f heisst **r-mal differenzierbar an der Stelle** $t_0 \in I$, wenn f, f', …, $f^{(r-1)}$ in einer ganzen Umgebung von t_0 existieren und $f^{(r)}$ wenigstens im Punkt t_0. Besitzt f auf dem Intervall I *stetige* Ableitungen bis zur Ordnung $r \geq 0$,

so heisst f eine **Funktion der Klasse** C^r; man schreibt dafür $f \in C^r$. Ist f beliebig oft differenzierbar, so nennt man f eine C^∞**-Funktion**. Es bestehen die natürlichen Inklusionen

$$C^0 \supset C^1 \supset C^2 \supset \ \ldots \ \supset C^\infty \ .$$

Wenn nötig, verwendet man auch präzisere Angaben wie $C^r(I)$ oder sogar $C^r(I, \mathbb{C})$, falls auch der intendierte Wertevorrat mitgeteilt werden soll.

Bsp: $\quad \exp \in C^\infty(\mathbb{R})\,, \quad \mathrm{abs} \in C^0(\mathbb{R})\,, \quad |t|^{3/2} \in C^1(\mathbb{R}).$

Bezüglich der höheren Differenzierbarkeit zusammengesetzter Funktionen sowie der Umkehrfunktion gelten die folgenden Sätze, die wir der Einfachheit halber gerade in der "C^r-Sprache" formulieren.

(7.9) *Sind f und g in $C^r(I, \mathbb{X})$, so ist auch das Produkt $f \cdot g$ in $C^r(I, \mathbb{X})$, und zwar gilt*

$$(f \cdot g)^{(r)} = \sum_{k=0}^{r} \binom{r}{k} f^{(r-k)}\, g^{(k)} \ . \tag{1}$$

⌜ Berechnet man mit Hilfe der Produktregel die sukzessiven Ableitungen von $f \cdot g$:

$$\begin{aligned}
(f \cdot g)^{(0)} &= fg \\
(f \cdot g)^{(1)} &= f'g + fg' \\
(f \cdot g)^{(2)} &= f''g + 2f'g' + fg'' \\
&\ \vdots \\
(f \cdot g)^{(r)} &= f^{(r)}g + \ \ldots \ + fg^{(r)} \quad ,
\end{aligned}$$

so erkennt man, dass rechter Hand ein Pascalsches Dreieck aufgebaut wird, und zwar sind alle dort stehenden Funktionen auf dem Intervall I stetig. ⌟

Die Formel (1) wird zuweilen als **Leibnizsche Formel** bezeichnet.

(7.10) *Ist $f \in C^r(I, \mathbb{R})$ und $g \in C^r\big(f(I)\big)$, so ist auch die Zusammensetzung $g \circ f \colon I \to \mathbb{X}$ in C^r.*

⌜ Die Behauptung trifft zu für $r = 0$ und sei richtig für $r - 1$ (≥ 0). Liegen jetzt f und g in C^r, so folgt mit der Kettenregel und **(7.9)**:

$$(g \circ f)' = (g' \circ f) \cdot f' \in (C^{r-1} \circ C^{r-1}) \cdot C^{r-1} \subset C^{r-1} \ ,$$

das heisst: $(g \circ f)' \in C^{r-1}$. Somit ist $g \circ f \in C^r$. ⌟

(7.11) *Es sei $f \in C^r(I, \mathbb{R})$, und es sei $f'(t) \neq 0$ auf I. Dann ist die Umkehrfunktion $f^{-1} =: g$ in $C^r\big(f(I)\big)$.*

⌜ Nach dem Zwischenwertsatz ist f' von einerlei Vorzeichen auf I. Folglich ist f streng monoton auf I und besitzt in der Tat eine Umkehrfunktion.

Die Formel **(7.7)** für die Ableitung der Umkehrfunktion lässt sich lesen als

$$g' = h \circ f' \circ g\,, \tag{2}$$

wobei das Nehmen des Kehrwerts als C^∞-Funktion $h\colon u \mapsto 1/u$ interpretiert wird. Die Funktion g ist jedenfalls in C^0 und sei in C^s für ein $s \in [\,0\,..\,(r-1)\,]$. Aus (2) folgt dann mit **(7.10)**:

$$g' \in C^\infty \circ C^{r-1} \circ C^s \subset C^s\,,$$

woraus man auf $g \in C^{s+1}$ schliesst. ⌟

Wir leiten noch eine zu **(7.7)** analoge Formel für die zweite Ableitung der Umkehrfunktion her. Aus (2) folgt mit der Kettenregel

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

Wegen $h'(u) = -1/u^2$ ist das gleichbedeutend mit

$$g''(y) = -\frac{1}{\big(f'\big(g(y)\big)\big)^2} \cdot f''\big(g(y)\big) \cdot \frac{1}{f'\big(g(y)\big)}\,,$$

so dass wir definitiv die Formel

$$\textbf{(7.12)} \qquad g''(y) = -\frac{f''\big(g(y)\big)}{\big(f'\big(g(y)\big)\big)^3} \qquad (g := f^{-1})$$

erhalten.

Aufgaben

1. Berechne die Ableitungen der folgenden Ausdrücke:

   (a) $\log \dfrac{1+\sqrt{1-t^2}}{t}$, (b) $\sqrt{\dfrac{\alpha+\beta t}{\alpha-\beta t}}$ $(\alpha,\ \beta > 0)$,

   (c) $t^{1/3}(1-t)^{2/3}(1+t)^{1/2}$, (d) t^t,

   (e) $(\log\tan t)^{-1/3}$.

   Bestimme in jedem Fall den Definitionsbereich D sowie den Definitionsbereich D' der Ableitung.

2. Berechne die hundertste Ableitung der Funktion $f(t) := t^2\sin(2t)$.

7.3 Extrema

Siehtu quadratische Gleichung, muttu auflösen;
hörtu "Maximum", muttu differenzieren.
Hasenmutter

Globale und lokale Extrema

In diesem und in den folgenden Abschnitten werden wir die Ableitung f' einer Funktion f heranziehen, um Aufschluss über Verhalten und Eigenschaften von f selbst zu erhalten.

Es sei f eine reellwertige Funktion mit einem vereinbarten Definitionsbereich $A \subset \mathbb{X}$. Neben den globalen, das heisst: auf ganz A bezüglichen Extremalstellen werden **lokale Extremalstellen** von f wie folgt erklärt: Die Funktion $f\colon A \to \mathbb{R}$ ist im Punkt $\xi \in A$ **lokal maximal**, wenn es ein $\delta > 0$ gibt mit

$$f(x) \leq f(\xi) \qquad \forall x \in A \cap U_\delta(\xi) \ .$$

Gemeint ist: Weit weg von ξ darf f schon noch grössere Werte annehmen. Es ist dann auch klar, was **lokal minimal** heisst; und wir sagen, f sei im Punkt ξ **lokal extremal**, wenn dort ein lokales Maximum oder ein lokales Minimum vorliegt. Nimmt $f(x) - f(\xi)$ in beliebig kleinen Umgebungen von ξ beiderlei Vorzeichen an, so ist f an der Stelle ξ *nicht* lokal extremal und somit erst recht nicht global extremal.

① Es sei $A := [\,-2, 4\,]$ und f die in der Figur 7.3.1 dargestellte Funktion. Dann ist f an den Stellen -1 und 4 lokal maximal, bei 4 global maximal, an den Stellen -2 und 1 lokal minimal und bei 1 global minimal. ○

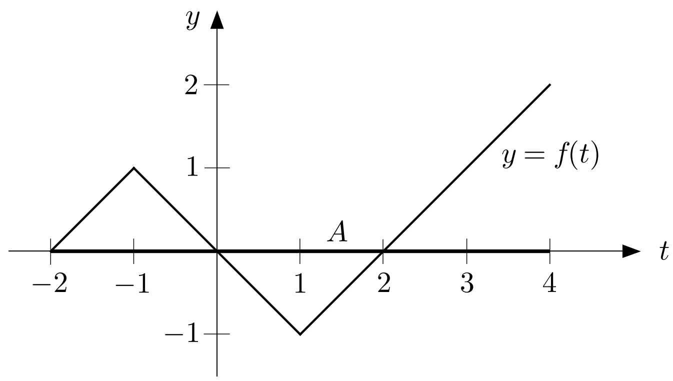

Fig. 7.3.1

Von jetzt ab sei $\operatorname{dom}(f) =: I$ wieder ein Intervall. Wir beginnen mit dem Hilfssatz

(7.13) *Ist t_0 ein innerer Punkt des Intervalls I und $f'(t_0) \neq 0$, so ist f an der Stelle t_0 n i c h t lokal extremal.*

Gemäss dem Prinzip **(7.2)** schreiben wir $f(t)-f(t_0) = m(t)(t-t_0)$, wobei die Funktion $m(\cdot)$ an der Stelle t_0 stetig ist und dort den Wert $f'(t_0) \neq 0$ annimmt. Wegen **(3.9)** gibt es daher eine ganze Umgebung $U_\delta(t_0) \subset I$, in der das Vorzeichen von $m(\cdot)$ mit dem Vorzeichen von $m(t_0) = f'(t_0)$ übereinstimmt. Hiernach gilt

$$\operatorname{sgn}\bigl(f(t) - f(t_0)\bigr) = \operatorname{sgn} f'(t_0) \cdot \operatorname{sgn}(t - t_0) \qquad (t \in U_\delta(t_0))\ ,$$

so dass die Differenz $f(t) - f(t_0)$ in $U_\delta(t_0)$ wie $t - t_0$ beiderlei Vorzeichen annimmt.

Damit werden die Nullstellen von f' interessant. Ein innerer Punkt $t_0 \in \operatorname{dom}(f)$ heisst ein **kritischer** oder **stationärer Punkt** der reellwertigen Funktion f, wenn $f'(t_0) = 0$ ist. Es gilt dann $f(t) - f(t_0) = m(t)(t - t_0)$ mit

$$\lim_{t \to t_0} m(t) = m(t_0) = f'(t_0) = 0\ .$$

Der Wertzuwachs von f ist hiernach für $t \to t_0$ von kleinerer Grössenordnung als das Inkrement $h := t - t_0$ der unabhängigen Variablen; daher der Name "stationär". Die Menge der kritischen Punkte von f bezeichnen wir mit $\operatorname{krit}(f)$.

② Betrachte die Funktion

$$f(t) := \sin t + \frac{1}{2}\sin(2t) \qquad (0 \leq t \leq 2\pi)$$

(Fig. 7.3.2) mit der Ableitung

$$f'(t) = \cos t + \cos(2t) = \cos t + (2\cos^2 t - 1) = (2\cos t - 1)(\cos t + 1)\ .$$

Die kritischen t genügen daher der Gleichung $2\cos t - 1 = 0$ $(\Rightarrow \cos t = 1/2)$ oder der Gleichung $\cos t + 1 = 0$ $(\Rightarrow \cos t = -1)$. Es ergibt sich

$$\operatorname{krit}(f) = \left\{\frac{\pi}{3}, \frac{5\pi}{3}, \pi\right\}\ .$$

○

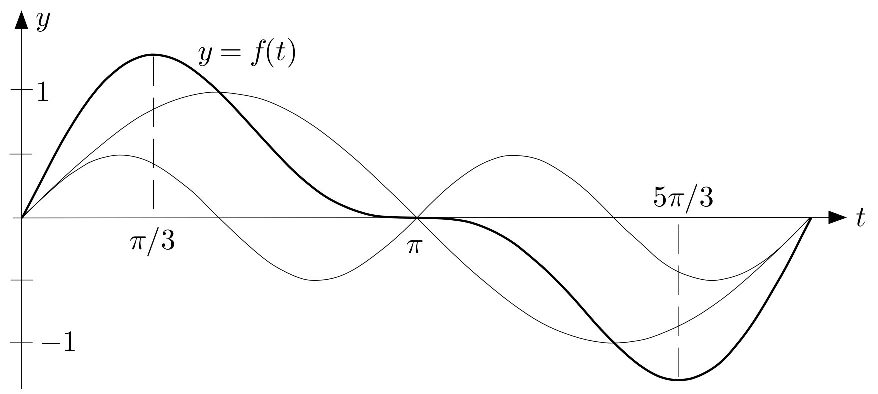

Fig. 7.3.2

Zahlreiche Sätze dieses Kapitels haben die folgende Situation zur Voraussetzung:

($*$) Die Funktion $f\colon [\,a,b\,] \to \mathbb{R}$ ist stetig und auf $]\,a,b\,[$ differenzierbar.

Der Verzicht auf die Differenzierbarkeit in den Endpunkten macht die betreffenden Sätze auch auf Funktionen wie $\sqrt{1-t^2}$ $\bigl([\,a,b\,] := [\,-1,1\,]\bigr)$ anwendbar. Aus ($*$) folgt insbesondere nach Satz **(4.17)**, dass f auf $[\,a,b\,]$ ein globales Maximum und ein globales Minimum annimmt. Die globalen Extrema eines derartigen f lassen sich nun folgendermassen bestimmen:

(7.14) *Genügt f der Voraussetzung* ($*$) *und sind $t_1, \ldots, t_r$ die kritischen Punkte von f auf $]\,a,b\,[$, so gilt*

$$\max_{a\le t\le b} f(t) = \max\{f(a), f(b), f(t_1), \ldots, f(t_r)\}\,, \tag{1}$$

analog für das Minimum.

⌜ Die “Kandidatenliste” rechter Hand in (1) enthält alle Punkte $t \in [\,a,b\,]$, ausgenommen die inneren Punkte mit $f'(t) \ne 0$. Letztere kommen nach **(7.13)** als Extremalstellen ohnehin nicht in Frage. ⌟

Beachte: Falls nur der Maximalwert $\max_{a\le t\le b} f(t)$ und die Menge der Maximalstellen gefragt sind, ist es nach diesem Satz nicht notwendig, zweite Ableitungen auszurechnen. Etwas anderes ist es, wenn zum Beispiel für eine Graphendiskussion oder für Stabilitätsbetrachtungen der Charakter der einzelnen kritischen Punkte untersucht werden soll (s.u.).

Beispiele

② (Forts.) Um die globalen Extrema von f auf $[0, 2\pi]$ zu bestimmen, müssen wir $\mathrm{krit}(f)$ durch die beiden Randpunkte 0 und 2π ergänzen und erhalten damit die definitive “Kandidatenliste”. Hierauf ist der Wertvergleich durchzuführen. Es ergibt sich die folgende Tabelle:

t	0	2π	$\pi/3$	$5\pi/3$	π
$f(t)$	0	0	$3\sqrt{3}/4$	$-3\sqrt{3}/4$	0

Die globalen Extremalstellen und -werte von f können hier sofort abgelesen werden. ○

③ Es sei σ die Verbindungsstrecke der beiden Punkte $A := (-1, 1)$, $B := (0, 2)$, und es sei $P := (p, 0)$ ein Punkt auf der x-Achse (Fig. 7.3.3). Welcher Punkt von σ liegt P am nächsten?

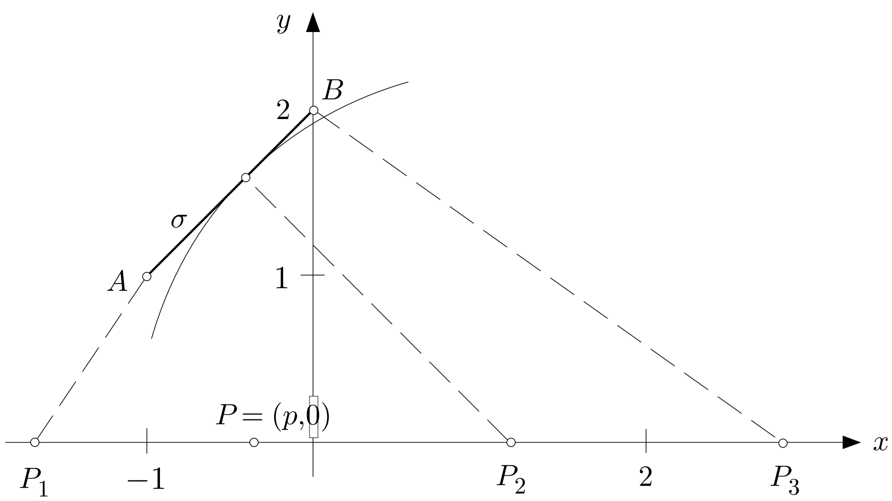

Fig. 7.3.3

Der Figur entnimmt man ohne weiteres die folgende Lösung dieser Aufgabe: Ist $p \leq 0$, so liegt der Endpunkt A am nächsten, und ist $p \geq 2$, so liegt B am nächsten. Für $0 < p < 2$ ist der nächste Punkt ein innerer Punkt von σ, nämlich der Fusspunkt des Lotes von P auf σ.

Was können wir hieraus lernen? Unsere Aufgabe enthält einen Parameter p. Der kürzeste Abstand von P zu σ hängt natürlich “zahlenmässig” von p ab, aber nicht nur das: Auch die *Gestalt* der Extremalsituation hängt von p ab und verändert sich an bestimmten Stellen der p-Achse (bei $p = 0$ und $p = 2$) radikal. Damit sind wir auf ein ziemlich universelles “Katastrophenprinzip” gestossen: Enthält ein mathematisches Modell Parameter p, λ, …, so muss man von vorneherein damit rechnen, dass für gewisse spezielle Werte der Parameter die Gesamtsituation umkippt zu einer vollständig neuen Gestalt.

Die rechnerische Behandlung der obigen Aufgabe überlassen wir dem Leser. ○

④ Wir betrachten die mit λ parametrisierte Funktionenschar

$$f_\lambda(t) := t^3 - 3\lambda^2 t + 4\lambda$$

auf dem Intervall $[-2, 2]$. Jedes f_λ nimmt auf diesem Intervall ein globales Maximum M_λ und ein globales Minimum m_λ an. Diese beiden Grössen sollen nun (als Funktionen von λ) bestimmt werden. Hierzu müssen wir für jedes feste $\lambda \in \mathbb{R}$ eine "Kandidatenliste" herstellen.

Zunächst ist

$$f_\lambda(-2) = -8 + 6\lambda^2 + 4\lambda =: \phi_1(\lambda)\,, \qquad f_\lambda(2) = 8 - 6\lambda^2 + 4\lambda =: \phi_2(\lambda)\ .$$

Ferner gilt $f'_\lambda(t) = 3t^2 - 3\lambda^2$, und dies verschwindet in den beiden Punkten $t = \pm\lambda$. Diese beiden Punkte fallen für $|\lambda| \geq 2$ ausser Betracht, da sie dann nicht im t-Intervall $]-2, 2[$ liegen. Die zugehörigen Funktionswerte

$$f_\lambda(-\lambda) = 2\lambda^3 + 4\lambda =: \phi_3(\lambda)\,, \qquad f_\lambda(\lambda) = -2\lambda^3 + 4\lambda =: \phi_4(\lambda)$$

sind somit nur für $-2 < \lambda < 2$ zur Konkurrenz zugelassen. Die Funktionen $\phi_1(\lambda)$, …, $\phi_4(\lambda)$ sind in der Fig. 7.3.4 simultan dargestellt, ϕ_3 und ϕ_4 nur in dem angegebenen Bereich. Die gesuchten Funktionen M_λ und m_λ lassen sich nun unmittelbar ablesen: Aufgrund von Satz **(7.14)** ist

$$M_\lambda = \max\{\phi_1(\lambda), \ldots, \phi_4(\lambda)\}$$

und analog für m_λ. Damit erhalten wir die folgende Tabelle:

	$\lambda \leq -2$	$-2 < \lambda < -1$	$-1 \leq \lambda \leq 1$	$1 < \lambda < 2$	$\lambda \geq 2$
M_λ	$\phi_1(\lambda)$	$\phi_4(\lambda)$	$\phi_2(\lambda)$	$\phi_3(\lambda)$	$\phi_1(\lambda)$
m_λ	$\phi_2(\lambda)$	$\phi_3(\lambda)$	$\phi_1(\lambda)$	$\phi_4(\lambda)$	$\phi_2(\lambda)$

Die Funktionen f_λ hingen differenzierbar von dem Parameter λ ab. Der Fig. 7.3.4 lässt sich nun entnehmen, dass die beiden Funktionen $\lambda \mapsto M_\lambda$ und $\lambda \mapsto m_\lambda$ zwar stetig sind (das muss so sein auf Grund von **(4.18)**); ihre Graphen haben aber Knickstellen. ○

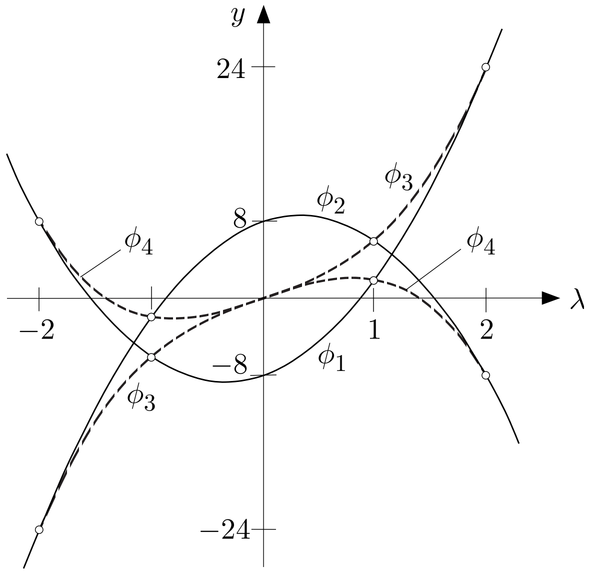

Fig. 7.3.4

Aufgaben

1. Eine Kugel soll in einen aufrechten Kreiskegel von minimalem Volumen gepackt werden. Bestimme den halben Öffnungswinkel des Kegels.

2. Die Funktion $f(t) := \sqrt{t}$ soll auf dem Intervall $[\,0,1\,]$ durch eine lineare Funktion $l(t) := t + c$ möglichst gut approximiert werden, das heisst so, dass
$$\max\bigl\{|f(t) - l(t)| \bigm| 0 \le t \le 1\bigr\}$$
minimal wird. Für welche Wahl von c ist das der Fall? *Hinweis:* Diskutiere die Funktion $g(t) := \sqrt{t} - t$.

3. Auf der Ellipse $x^2 + 4y^2 = 4$ bestimme man diejenigen Punkte, die von dem Punkt $(c, 0)$, $0 < c < 2$, minimalen Abstand haben. Man zeichne die gesuchten Punkte für die Fälle $c := 3/4$ und $c := 9/5$.

4. Eine Zahl $a \ge 1$ soll in $n \ge 1$ gleiche Teile geteilt werden, so dass das Produkt der Teile möglichst gross wird. Bestimme n in Abhängigkeit von a. *Hinweis:* Die Funktion $\phi(t) := (a/t)^t$ ist unimodal.

5. Ein periodischer Vorgang wird beschrieben durch die Funktion
$$f(t) := \alpha \cos t + \cos(2t)\ .$$
Für welchen Wert des reellen Parameters α ist der Maximalausschlag (nach oben oder unten) minimal?

6. Bestimme die grösste Zahl, die als Produkt von positiven ganzen Zahlen der Summe 2003 dargestellt werden kann.

7. Unter allen einem Kreise einbeschriebenen Dreiecken hat das gleichseitige den grössten Flächeninhalt. Beweise dies elementargeometrisch, das heisst: ohne Differentialrechnung, und zwar
   (a) unter der Annahme, dass es ein flächengrösstes Dreieck gibt,
   (b) ohne diese Annahme.

8. Es sei
$$f_\lambda(t) := e^t - \lambda t\ .$$
   Zeichne die Graphen der Funktionen
$$m(\lambda) := \min\bigl\{ f_\lambda(t) \bigm| 0 \le t \le 1 \bigr\}\ , \qquad M(\lambda) := \max\bigl\{ f_\lambda(t) \bigm| 0 \le t \le 1 \bigr\}\ .$$

9. (a) Die Funktion
$$f(t) := \arctan(q \tan t) \qquad \Bigl(-\frac{\pi}{2} < t < \frac{\pi}{2}\Bigr)$$
   spielt eine gewisse Rolle in der Theorie der elliptischen Funktionen; dabei ist q ein fester Parameter, $0 < q < 1$. Diese Funktion ist an den Stellen $\pm\pi/2$ nur scheinbar singulär und besitzt in Wirklichkeit eine überaus natürliche Fortsetzung $f^\sim$ auf die ganze t-Achse. Man finde einen Ausdruck für $f^\sim$. *Hinweis:* Betrachte die Hilfsfunktion $\phi(t) := f(t) - t$.
   (b) Berechne $\max\bigl\{ |f^\sim(t) - t| \bigm| t \in \mathbb{R} \bigr\}$.

7.4 Mittelwertsatz

Reellwertige Funktionen

Wir beginnen mit dem intuitiv einleuchtenden **Satz von Rolle**. Er stützt sich wiederum auf die Voraussetzung

$(*)$ Die Funktion $f\colon [\,a,b\,] \to \mathbb{R}$ ist stetig und auf $]\,a,b\,[$ differenzierbar.

(7.15) *Genügt f der Voraussetzung $(*)$ und ist $f(a) = f(b)$, so gibt es einen Punkt $\tau \in\,]\,a,b\,[$ mit $f'(\tau) = 0$.*

⌜ Ist f nicht konstant, so gibt es zum Beispiel Punkte $t \in\,]\,a,b\,[$ mit $f(t) > f(a)$, und das globale Maximum von f auf $[\,a,b\,]$ wird notwendigerweise in (mindestens) einem inneren Punkt τ angenommen. An einer derartigen Stelle τ ist $f'(\tau) = 0$ nach Lemma **(7.13)**. ⌟

Der angekündigte **Mittelwertsatz** erscheint in verschiedenen Varianten:

(7.16) *Genügt f der Voraussetzung $(*)$, so gibt es einen Punkt $\tau \in\,]\,a,b\,[$ mit*

$$f(b) - f(a) = f'(\tau)(b-a) \qquad \textit{bzw.} \qquad \frac{f(b)-f(a)}{b-a} = f'(\tau)\ . \tag{1}$$

Die Formel (1) gilt auch im Fall $b < a$; natürlich liegt dann der Punkt τ im Intervall $]\,b,a\,[$. Mit **(7.16)** äquivalent ist zum Beispiel

(7.17) *Genügt f auf dem Intervall mit den Endpunkten t_0 und t_0+h $(h > 0$ oder $h < 0)$ der Voraussetzung $(*)$, so gibt es ein $\theta \in\,]\,0,1\,[$ mit*

$$f(t_0 + h) - f(t_0) = f'(t_0 + \theta h)\, h\,;$$

denn der Punkt $t_0 + \theta h$ liegt in jedem Fall zwischen den Punkten t_0 und $t_0 + h$. — Anstelle von **(7.16)** beweisen wir etwas allgemeiner:

(7.18) *Genügen f und g den Voraussetzungen $(*)$ und ist $g'(t) \neq 0$ für alle $t \in\,]\,a,b\,[$, so gibt es einen Punkt $\tau \in\,]\,a,b\,[$ mit*

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\tau)}{g'(\tau)}\ . \tag{2}$$

⌜ Wir setzen zur Abkürzung

$$f(b) - f(a) =: \Delta f\ , \qquad g(b) - g(a) =: \Delta g\ .$$

Dann ist jedenfalls $\Delta g \neq 0$, denn sonst besässe g' nach dem Satz von Rolle eine Nullstelle. Betrachten wir nun die Hilfsfunktion

$$h(t) := \Delta g\, f(t) - \Delta f\, g(t)\ ,$$

so genügt h der Voraussetzung $(*)$, und es gilt

$$h(b) - h(a) = \Delta g\, \Delta f - \Delta f\, \Delta g = 0\ .$$

Nach dem Satz von Rolle gibt es daher einen Punkt $\tau \in\,]\, a, b\, [$ mit $h'(\tau) = 0$, also $\Delta g\, f'(\tau) - \Delta f\, g'(\tau) = 0$. Hieraus folgt aber

$$\frac{\Delta f}{\Delta g} = \frac{f'(\tau)}{g'(\tau)}\ ,$$

wie behauptet. — Wählt man speziell $g(t) :\equiv t$, so ist $g(b) - g(a) = b - a$ und $g'(t) \equiv 1$. Damit geht (2) in (1) über, womit auch **(7.16)** bewiesen ist. ⌟

① Satz **(7.16)** gibt keine Auskunft darüber, wo genau der Punkt $\tau \in\,]\, a, b\, [$ im Einzelfall zu finden ist. Wir stellen uns hier die Aufgabe, diejenigen Funktionen $f : \mathbb{R} \to \mathbb{R}$ zu bestimmen, für die stets $\tau = (a + b)/2$ $\big($bzw. in **(7.17)** stets $\theta = \frac{1}{2}\big)$ ist. Eine derartige Funktion genügt der Identität

$$f'(t) = \frac{f(t+h) - f(t-h)}{2h} \tag{3}$$

in den beiden Variablen t und h, $h > 0$. Die rechte Seite von (3) ist für festes h differenzierbar nach t, also auch die linke Seite. Somit ist f von selbst zweimal differenzierbar. Wir schreiben (3) in der Form

$$2h\, f'(t) = f(t+h) - f(t-h)$$

und differenzieren dies zweimal für festes t nach der Variablen h. Wir erhalten nacheinander $2f'(t) = f'(t+h) + f'(t-h)$ und

$$0 = f''(t+h) - f''(t-h)\ . \tag{4}$$

Da man irgend zwei Punkte auf der Zahlengeraden durch geeignete Wahl von t und $h > 0$ als $t - h$ und $t + h$ auffassen kann, ergibt sich aus (4), dass f'' eine Konstante ist: $f''(t) \equiv 2\alpha$. Hieraus schliesst man mit Hilfe von Satz **(7.20)** auf $f'(t) = 2\alpha t + \beta$, und nochmalige Anwendung von **(7.20)** liefert schliesslich

$$f(t) = \alpha t^2 + \beta t + \gamma\ , \qquad \alpha, \beta, \gamma \in \mathbb{R}\ .$$

Demnach können höchstens die Polynome vom Grad ≤ 2 der Funktionalgleichung (3) genügen. Man rechnet leicht nach, dass derartige Polynome in der Tat diese besondere Eigenschaft besitzen. ○

Komplex- und vektorwertige Funktionen

② Ist $f : [\,a, b\,] \to \mathbb{X}$ eine komplex- oder vektorwertige Funktion, so gibt es im allgemeinen *kein* $\tau \in \,]\,a, b\,[$, für das die Gleichung (1) zutrifft. Betrachte zum Beispiel die Funktion $f(t) := e^{it}$ auf dem Intervall $[\,0, 2\pi\,]$. Die Ableitung $f'(t) = ie^{it}$ ist $\neq 0$ für alle t; somit ist

$$0 = f(2\pi) - f(0) \neq f'(\tau) \cdot 2\pi \qquad \forall \tau \in \,]\,0, 2\pi\,[\ .$$

○

Die folgende handliche Abschätzung nimmt nicht mehr auf einen unfassbaren Punkt τ Bezug und ist auch für komplex- oder vektorwertige Funktionen gültig.

(7.19) *Es sei* I *ein Intervall und* $\mathbf{f} : I \to \mathbb{X}$ *eine differenzierbare Funktion mit*

$$|\mathbf{f}'(t)| \le M \qquad \forall t \in I\ .$$

Dann gilt für beliebige $t_1,\ t_2 \in I$*:*

$$|\mathbf{f}(t_2) - \mathbf{f}(t_1)| \le M\,|t_2 - t_1|\ .$$

⌜ Wir dürfen $t_1 < t_2$ und $\mathbf{c} := \mathbf{f}(t_2) - \mathbf{f}(t_1) \neq 0$ annehmen (für $\mathbf{c} = \mathbf{0}$ ist nichts zu beweisen). Wir führen die reellwertige Hilfsfunktion

$$\phi(t) := \mathbf{c} \bullet \mathbf{f}(t)$$

ein und betrachten die Grösse

$$\phi(t_2) - \phi(t_1) = \mathbf{c} \bullet \big(\mathbf{f}(t_2) - \mathbf{f}(t_1)\big) = |\mathbf{c}|^2\ .$$

Nach **(7.16)** gibt es ein $\tau \in \,]\,t_1, t_2\,[$ mit

$$|\mathbf{c}|^2 = \phi(t_2) - \phi(t_1) = \phi'(\tau)(t_2 - t_1) = \mathbf{c} \bullet \mathbf{f}'(\tau)\,|t_2 - t_1|\ ,$$

und mit Hilfe der Schwarzschen Ungleichung folgt

$$|\mathbf{c}|^2 \le |\mathbf{c}| \cdot |\mathbf{f}'(\tau)|\,|t_2 - t_1| \le |\mathbf{c}|\,M\,|t_2 - t_1|\ .$$

Nach Division mit $|\mathbf{c}|$ ergibt sich hieraus $|\mathbf{c}| \le M|t_2 - t_1|$, wie behauptet. ⌟

(7.20) *Es sei* I *ein Intervall und* $\mathbf{f} : I \to \mathbb{X}$ *eine differenzierbare Funktion mit* $\mathbf{f}'(t) \equiv \mathbf{0}$ *auf* I*. Dann ist* $\mathbf{f}$ *konstant.*

⌜ Ist $\mathbf{f}'(t) \equiv \mathbf{0}$, so treffen die Voraussetzungen von **(7.19)** mit $M := 0$ zu, und es folgt $\mathbf{f}(t_1) = \mathbf{f}(t_2)$ für beliebige $t_1,\ t_2 \in I$. ⌟

Satz **(7.20)** ist ein "Integralsatz": Aus "differentiellen", in kleinsten Teilintervallen überprüfbaren Eigenschaften von $\mathbf{f}$ werden Folgerungen gezogen, die die Funktionswerte in beliebig weit auseinanderliegenden Punkten miteinander verknüpfen. Man kann es auch so sehen: Die Voraussetzung $\mathbf{f}'(t) \equiv \mathbf{0}$ enthält *eine* freie Variable, die Konklusion $\mathbf{f}(t_1) \equiv \mathbf{f}(t_2)$ deren *zwei.*

Die Regel von Bernoulli–de l'Hôpital

Wir benutzen nun den Mittelwertsatz zum Beweis der beliebten **Regel von Bernoulli–de l'Hôpital**. Es handelt sich dabei um eine einfache Methode zur Berechnung von gewissen Grenzwerten, die zunächst auf Ausdrücke der Form $0/0$ oder ∞/∞ führen.

(7.21) *Die reellwertigen Funktionen f und g seien differenzierbar auf dem Intervall $[\,a, b\,[$, $a < b \leq \infty$; dabei seien g und g' durchwegs $\neq 0$. Gilt*

(I) $$\lim_{t \to b-} f(t) = \lim_{t \to b-} g(t) = 0$$

oder

(II) $$\lim_{t \to b-} g(t) = \infty\,,$$

so ist

$$\lim_{t \to b-} \frac{f(t)}{g(t)} = \lim_{t \to b-} \frac{f'(t)}{g'(t)} \qquad (=: \lambda)\,,$$

falls der Grenzwert rechter Hand existiert ($\lambda = \pm\infty$ zugelassen).

⌜ Im Fall (I)$\wedge(b < \infty)$ kann man f und g durch $f(b) := 0$, $g(b) := 0$ stetig in den Punkt b hinein fortsetzen und erhält mit **(7.18)**:

$$\frac{f(t)}{g(t)} = \frac{f(b) - f(t)}{g(b) - g(t)} = \frac{f'(\tau)}{g'(\tau)} \tag{5}$$

für ein gewisses τ zwischen t und b. Mit $t \to b-$ strebt auch τ gegen $b-$, die rechte Seite von (5) gegen λ und folglich auch die linke Seite.

Nochmals von vorn, aber so, dass uneigentliche Grenzlagen mit erfasst werden. Dies zwingt dazu, die in der Grenzwertdefinition vorkommenden Umgebungen durch ihre Endpunkte und nicht in der Form $U_\varepsilon(x)$ zu spezifizieren.

Aus Symmetriegründen genügt es, das folgende Lemma zu beweisen:

(7.22) *Es sei $-\infty \leq \lambda < \infty$ und $q > \lambda$. Dann gibt es ein $t_0 \in [\,a, b\,[$ mit*

$$\frac{f(t)}{g(t)} < q \qquad \forall t \in\,]\,t_0, b\,[\;.$$

⌜ Wähle ein p mit $\lambda < p < q$ (Fig. 7.4.1). Nach Voraussetzung über f'/g' gibt es dann ein $t_0 < b$ mit

$$\frac{f'(t)}{g'(t)} < p \qquad \forall t \in\,]\,t_0, b\,[\;.$$

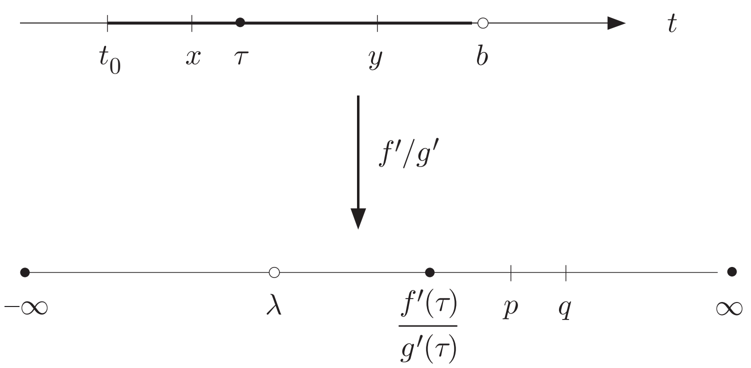

Fig. 7.4.1

Betrachte ein beliebiges Intervall $[\,x, y\,]$ mit $t_0 \le x < y < b$. Aufgrund von Satz **(7.18)** gibt es ein $\tau \in\,]\,x, y\,[$ mit

$$\frac{f(y) - f(x)}{g(y) - g(x)} = \frac{f'(\tau)}{g'(\tau)} < p \; . \tag{6}$$

Wir erledigen zunächst den Fall (I). Führt man in (6) *für festes* x den Grenzübergang $y \to b-$ durch, so ergibt sich im Limes

$$\frac{f(x)}{g(x)} \le p < q \; ,$$

und zwar gilt dies für beliebiges $x > t_0$, wie behauptet.

Im Fall (II) setzen wir $x := t_0$ und erweitern (6) mit $1/g(y)$. Es ergibt sich

$$\frac{\dfrac{f(y)}{g(y)} - \dfrac{f(t_0)}{g(y)}}{1 - \dfrac{g(t_0)}{g(y)}} < p \; .$$

Hieraus folgt

$$\frac{f(y)}{g(y)} < \frac{f(t_0)}{g(y)} + \left(1 - \frac{g(t_0)}{g(y)}\right) p \; .$$

Mit $y \to b-$ strebt hier die rechte Seite nach Voraussetzung über g gegen p $(< q)$; somit gibt es ein $y_0 < b$ mit

$$\frac{f(y)}{g(y)} < q \qquad \forall y \in\,]\,y_0, b\,[\; .$$

$\lrcorner$ $\lrcorner$

③

$$\lim_{t \to 1+} \frac{2t^2 + t - 3}{t^3 - 3t + 2} = \lim_{t \to 1+} \frac{4t + 1}{3t^2 - 3} = \lim_{t \to 1+} \left[\frac{4t + 1}{3(t + 1)} \cdot \frac{1}{t - 1}\right] \; .$$

Hier strebt der erste Faktor rechter Hand mit $t \to 1+$ gegen $5/6$, der zweite gegen ∞, das Produkt also gegen ∞. ○

④ Es sei $\alpha\beta \neq 0$. Bei der Funktion

$$h(t) := \frac{\log\cosh(\alpha t)}{\log\cosh(\beta t)}$$

streben Zähler und Nenner mit $t \to 0$ beide gegen 0 und mit $t \to \infty$ beide gegen ∞. Wegen

$$\frac{d}{dy}\log\cosh y = \log'(\cosh y)\,\cosh'(y) = \frac{\sinh y}{\cosh y} = \tanh y$$

erhalten wir daher einerseits

$$\lim_{t\to 0} h(t) = \lim_{t\to 0}\frac{\alpha\tanh(\alpha t)}{\beta\tanh(\beta t)} = \lim_{t\to 0}\frac{\alpha^2(1-\tanh^2(\alpha t))}{\beta^2(1-\tanh^2(\beta t))} = \frac{\alpha^2}{\beta^2}\ ,$$

wobei wir **(7.21)** gleich zweimal angewandt haben, und andererseits

$$\lim_{t\to\infty} h(t) = \lim_{t\to\infty}\frac{\alpha\tanh(\alpha t)}{\beta\tanh(\beta t)} = \frac{\alpha\operatorname{sgn}\alpha}{\beta\operatorname{sgn}\beta} = \left|\frac{\alpha}{\beta}\right|\ .$$

○

⑤ Die bedenkenlose Anwendung der Regel von Bernoulli–de l'Hôpital auf komplexwertige Funktionen führt zu falschen Resultaten. Betrachte hierzu das einfache Beispiel

$$f(t) := t\ , \qquad g(t) := te^{-i/t}$$

mit $\lim_{t\to 0+} f(t) = \lim_{t\to 0+} g(t) = 0$.
Es ist $f'(t) \equiv 1$ und $g'(t) = (1+i/t)e^{-i/t}$. Wegen $\left|e^{-i/t}\right| = 1$ gilt folglich

$$\lim_{t\to 0+}\frac{f'(t)}{g'(t)} = \lim_{t\to 0+}\frac{t}{t+i}e^{i/t} = 0\ .$$

Der $\lim_{t\to 0+}\dfrac{f(t)}{g(t)} = \lim_{t\to 0+} e^{i/t}$ existiert aber gar nicht!
Um dieses beunruhigende Phänomen zu verstehen, muss man sich daran erinnern, dass **(7.21)** letzten Endes mit Hilfe des Satzes von Rolle bewiesen wurde, der nur für reellwertige Funktionen jenes nützliche $\tau \in\,]\,a,b\,[$ liefert.

○

Ein einfaches Monotoniekriterium

Für differenzierbare Funktionen gibt es ein einfaches Monotoniekriterium:

(7.23) *Es sei I ein beliebiges Intervall und I° das Innere von I. Eine stetige und in I° differenzierbare Funktion $f : I \to \mathbb{R}$ ist genau dann monoton wachsend auf I, wenn folgendes zutrifft:*

$$f'(t) \geq 0 \qquad \forall t \in I^\circ\,, \tag{7}$$

und genau dann streng monoton wachsend, wenn (7) *gilt und f' auf keinem Teilintervall positiver Länge identisch verschwindet.*

⌜ Ist f monoton wachsend, so sind alle Differenzenquotienten

$$\frac{f(t) - f(t_0)}{t - t_0} \qquad (t \neq t_0)$$

nichtnegativ, und dasselbe gilt dann auch für deren Grenzwerte. Gibt es dabei ein Teilintervall $[\,x, y\,] \subset I$ mit $f'(t) \equiv 0$ für $x < t < y$, so folgt mit **(7.20)**: $f(x) = f(y)$, und f ist nicht streng monoton.

Umgekehrt: Die Bedingung (7) sei erfüllt, und es seien x, y zwei beliebige Punkte in I mit $x < y$. Nach dem Mittelwertsatz **(7.17)**, angewandt auf das Intervall $[\,x, y\,]$, gibt es dann ein $\tau \in\,]\,x, y\,[$ mit

$$f(y) - f(x) = f'(\tau)(y - x) \geq 0\,,$$

somit ist f auf I monoton. Gilt dabei sogar $f(x) = f(y)$, so ist die monotone Funktion f auf dem Intervall $[\,x, y\,]$ konstant, und f' ist identisch 0 auf $]\,x, y\,[$. ⌟

⑥ Aus

$$\sin'(t) = \cos t > 0\,, \qquad \tan'(t) = 1 + \tan^2 t > 0 \qquad \left(-\frac{\pi}{2} < t < \frac{\pi}{2}\right)$$

folgt, dass der Sinus im Intervall $\left[\,-\frac{\pi}{2}, \frac{\pi}{2}\,\right]$, der Tangens im Intervall $\left]\,-\frac{\pi}{2}, \frac{\pi}{2}\,\right[$ streng monoton wächst. (Wir haben diese zwei Sachverhalte schon in Abschnitt 6.6 benötigt und dort auf wesentlich umständlichere Weise *ad hoc* nachgewiesen.)

Betrachte die Funktion

$$g(u) := u - \sin u \qquad (-\infty < u < \infty)\ .$$

Die Ableitung $g'(u) = 1 - \cos u$ ist überall ≥ 0 und verschwindet nur in den isolierten Punkten $u_k := 2k\pi,\ k \in \mathbb{Z}$. Somit ist g streng monoton wachsend, und zwar gegen ∞ für $u \to \infty$. Mit Hilfe von g bilden wir nun die auf $[\,0,1\,]$ stetige Funktion

$$f(t) := \begin{cases} \dfrac{\pi}{g(\pi/t)} & (0 < t \leq 1) \\ 0 & (t = 0) \end{cases}$$

(Fig. 7.4.2). Obwohl die Ableitung von f im Intervall $]\,0,1\,]$ unendlich viele Nullstellen besitzt, ist $f : [\,0,1\,] \to \mathbb{R}$ als Zusammensetzung von streng monotonen Funktionen ebenfalls streng monoton. ○

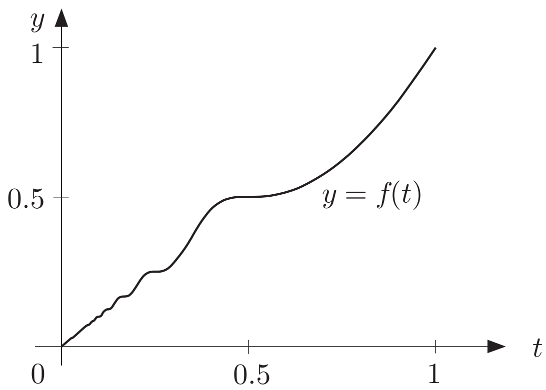

Fig. 7.4.2

Aufgaben

1. Betrachte die Funktion

$$f(t) := \begin{cases} t + 2t^2 \sin(1/t) & (t \neq 0) \\ 0 & (t = 0) \end{cases}$$

und verifiziere, was folgt:

(a) f ist differenzierbar, (b) f' ist auf $]-1,1\,[$ beschränkt,

(c) $f'(0) = 1$,

(d) f ist auf keinem noch so kleinen Intervall $]-h,h\,[$, $h > 0$, monoton wachsend.

2. Berechne die folgenden Grenzwerte:

(a) $\displaystyle\lim_{t\to 0} \frac{\sin\bigl(\sqrt{1+t}-1\bigr)}{t}$, (b) $\displaystyle\lim_{t\to \pi/2} \frac{\sin t + \sin(3t)}{\cos(2t)}$,

(c) $\displaystyle\lim_{t\to 0} \frac{a^t - 1}{b^t - 1}$ $(a,\, b > 0)$, (d) $\displaystyle\lim_{t\to 0} \frac{\sqrt{1-\cos t}}{t}$,

(e) $\displaystyle\lim_{t\to 1} \frac{t^\alpha - t^\beta}{t^{1/\beta} - t^{1/\alpha}}$ $\bigl(\alpha\beta(\alpha-\beta) \neq 0\bigr)$,

(f) $\displaystyle\lim_{t\to 0} (1 + 2\sin t)^{\cot t}$ (*Hinweis:* Logarithmieren).

7.5 Konvexe Funktionen

In den Abschnitten 7.3 und 7.4 haben wir gesehen, dass aus Eigenschaften der Ableitung f' einer Funktion auf das qualitative und quantitative Verhalten von f selbst geschlossen werden kann. In diesem Abschnitt werden wir in ähnlicher Weise einen Zusammenhang zwischen f'' und f herstellen. Hierzu müssen wir etwas weiter ausholen.

Im folgenden ist $I \subset \mathbb{R}$ ein beliebiges Intervall. Eine Funktion $f : I \to \mathbb{R}$ heisst **konvex** auf I, wenn für *jedes* Teilintervall $[\,t_1, t_2\,] \subset I$ gilt:

$$f(t) \le f^*(t) := \frac{t_2 - t}{t_2 - t_1} f(t_1) + \frac{t - t_1}{t_2 - t_1} f(t_2) \qquad (t_1 < t < t_2)\,, \tag{1}$$

und **streng konvex**, wenn gilt: $f(t) < f^*(t) \quad (t_1 < t < t_2)$. Die Funktion f^* **interpoliert** f in den Punkten t_1 und t_2 **linear**: f^* ist linear in t und nimmt an den Stellen t_1, t_2 bzw. die Werte $f(t_1)$, $f(t_2)$ an. Somit ist f konvex, wenn der Graph von f in jedem Teilintervall $[\,t_1, t_2\,] \subset I$ *unterhalb* der betreffenden Sehne verläuft (Fig. 7.5.1). Gilt (1) mit dem umgekehrten Ungleichheitszeichen, so heisst f **konkav** oder auch **nach oben konvex**.

Der Einfachheit halber argumentieren wir im weiteren nur für streng konvexe Funktionen. Um ihre besonderen Eigenschaften zu ergründen, betrachten wir das Verhalten des Differenzenquotienten

$$m_f(s,t) := \frac{f(t) - f(s)}{t - s} \qquad (s, t \in I,\ s < t)$$

als Funktion der beiden Variablen s und t. Zunächst ein einfaches Lemma:

(7.24) *Ist $f : I \to \mathbb{R}$ streng konvex, so ist $m_f(s,t)$ für festes s als Funktion von t und auch für festes t als Funktion von s streng monoton wachsend.*

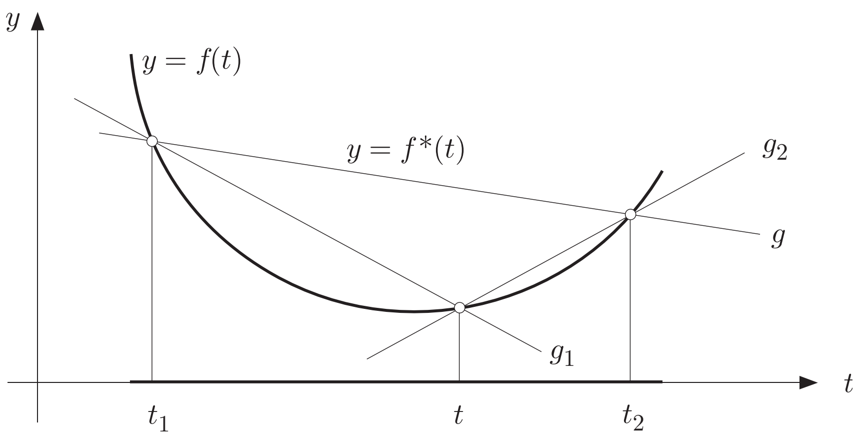

Fig. 7.5.1

⌜ Vergleicht man die Steigungen der zwei Geraden g_1 und g in der Figur 7.5.1, so ergibt sich die Monotonie von $m_f(s,\cdot)$ für festes s bezüglich t; und vergleicht man die Steigungen der Geraden g und g_2, so ergibt sich analog die Monotonie von $m_f(\cdot,t)$ bezüglich s. ⌟

Damit sind wir schon in der Lage, den folgenden überraschenden Satz zu beweisen:

(7.25) *Es sei I ein offenes Intervall und $f\colon I\to\mathbb{R}$ streng konvex.*

(a) *Dann ist f von selbst stetig auf I.*

(b) *In jedem Punkt $t_0\in I$ existieren sogar die beiden einseitigen Ableitungen $f'(t_0-)$ und $f'(t_0+)$, und zwar gilt*

$$f'(t_0-)\le f'(t_0+)\ . \tag{2}$$

(c) *Für jedes feste $m\in[f'(t_0-),f'(t_0+)]$ gilt:*

$$f(t)>f(t_0)+m(t-t_0)\qquad \forall t\ne t_0\ . \tag{3}$$

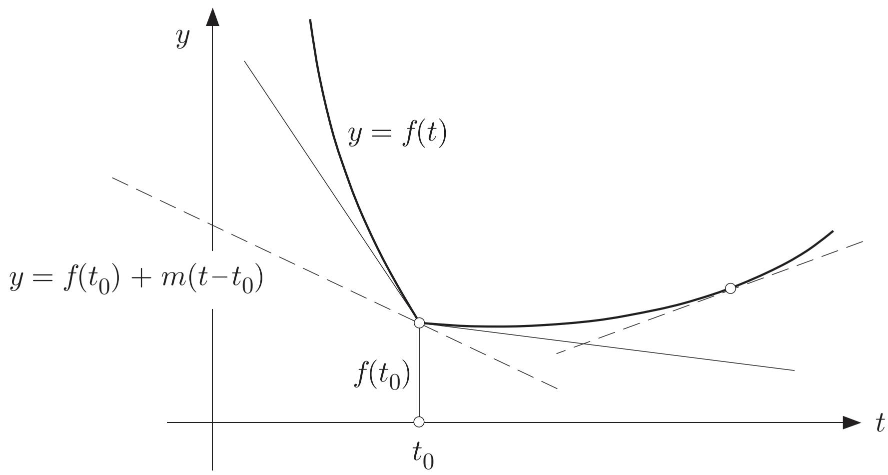

Fig. 7.5.2

Die linearen Funktionen auf der rechten Seite von (3) heissen **Stützfunktionen** von f im Punkt t_0, ihre Graphen **Stützgeraden** des Graphen von f (Fig. 7.5.2). Ist $f'(t_0-)=f'(t_0+)$, so gibt es genau eine Stützgerade, nämlich die Tangente an den Graphen von f im Punkt $\bigl(t_0,f(t_0)\bigr)$.

⌜ Ist $x<t_0<y$, so gilt nach Lemma **(7.24)**:

$$m_f(x,t_0)<m_f(x,y)<m_f(t_0,y)\ .$$

Führt man hier für festes x den Grenzübergang $y \to t_0+$ durch, so folgt aus **(7.24)** und Satz **(4.3)** die Existenz von

$$f'(t_0+) = \lim_{y \to t_0+} m_f(t_0, y) \geq m_f(x, t_0) \ .$$

Mit $x \to t_0-$ erhalten wir hieraus, wiederum wegen **(7.24)** und Satz **(4.3)**, die Existenz von $f'(t_0-)$, ferner die Relation (2). — A fortiori ist f an der Stelle t_0 beidseitig stetig.

Für jedes $t > t_0$ gilt, immer noch wegen **(7.24)**:

$$\frac{f(t) - f(t_0)}{t - t_0} = m_f(t_0, t) > f'(t_0+) \geq m \ ,$$

und hieraus folgt (3) für $t > t_0$; analog schliesst man für $t < t_0$. ┘

Der angekündigte Zusammenhang zwischen f'' und f wird nun durch das folgende Konvexitätskriterium hergestellt:

(7.26) *Es sei I ein beliebiges Intervall und I° das Innere von I. Eine auf I stetige und in I° zweimal differenzierbare Funktion $f\colon I \to \mathbb{R}$ ist genau dann streng konvex auf I, wenn folgendes zutrifft:*

$$f''(t) \geq 0 \qquad \forall t \in I^\circ$$

und wenn dabei f'' auf keinem Teilintervall positiver Länge identisch verschwindet.

┌ Ist f streng konvex und sind t_1, t_2 zwei innere Punkte von I mit $t_1 < t_2$, so gilt wegen **(7.24)**:

$$f'(t_1) = f'(t_1+) < m_f(t_1, t_2) < f'(t_2-) = f'(t_2) \ ;$$

folglich ist $f' : I^\circ \to \mathbb{R}$ streng monoton wachsend, und f'' hat auf Grund des Monotoniekriteriums **(7.23)** die behaupteten Eigenschaften.

Umgekehrt folgt aus den angegebenen Voraussetzungen über f'', dass f' auf I° streng monoton wächst. Es seien jetzt t_1, t, t_2 drei beliebige Punkte von I mit $t_1 < t < t_2$. Dann gibt es nach dem Mittelwertsatz **(7.16)** ein $\tau_1 \in \,]\, t_1, t\, [$ und ein $\tau_2 \in \,]\, t, t_2\, [$ mit

$$\frac{f(t) - f(t_1)}{t - t_1} = f'(\tau_1) < f'(\tau_2) = \frac{f(t_2) - f(t)}{t_2 - t} \ .$$

Hieraus ergibt sich durch Heraufmultiplizieren und Umordnen:

$$(t_2 - t_1) f(t) < (t_2 - t) f(t_1) + (t - t_1) f(t_2) \ ;$$

und dies ist äquivalent mit (1). ┘

Wendepunkte

Wechselt f'' an der Stelle t_0 das Vorzeichen, das heisst: Gilt

$$\operatorname{sgn} f''(t_0 - \delta) \cdot \operatorname{sgn} f''(t_0 + \varepsilon) = -1$$

für alle hinreichend kleinen δ, $\varepsilon > 0$, so geht f an der Stelle t_0 vom konvexen zum konkaven Verhalten über oder umgekehrt: Der Graph von f besitzt dort einen **Wendepunkt**.

① Es sei

$$f(t) := at^3 + bt^2 + ct + d\,, \qquad a \neq 0$$

ein beliebiges Polynom dritten Grades. Zweimalige Ableitung liefert

$$f'(t) = 3at^2 + 2bt + c\,, \qquad f''(t) = 6at + 2b = 6a\bigl(t + \frac{b}{3a}\bigr)\ .$$

Da f'' an der Stelle $t_0 := -b/(3a)$ das Vorzeichen wechselt, besitzt $\mathcal{G}(f)$ an der Stelle $W := (t_0, f(t_0))$ einen Wendepunkt (Fig. 7.5.3). $\mathcal{G}(f)$ ist übrigens bezüglich W symmetrisch. Um dies einzusehen, berechnen wir die Ableitung der Hilfsfunktion $g(h) := f(t_0 + h) + f(t_0 - h)$. Es ergibt sich

$$\begin{aligned} g'(h) &= f'(t_0 + h) - f'(t_0 - h) \\ &= 3a\bigl((t_0 + h)^2 - (t_0 - h)^2\bigr) + 2b\bigl((t_0 + h) - (t_0 - h)\bigr) \\ &= 3a \cdot 4t_0 h + 4b\,h = 0\,, \end{aligned}$$

unabhängig von h; folglich besitzt g den konstanten Wert $2f(t_0)$. Die hieraus folgende Identität

$$\forall h \in \mathbb{R}: \qquad f(t_0 - h) - f(t_0) = -\bigl(f(t_0 + h) - f(t_0)\bigr)$$

drückt gerade die behauptete Symmetrie aus. ○

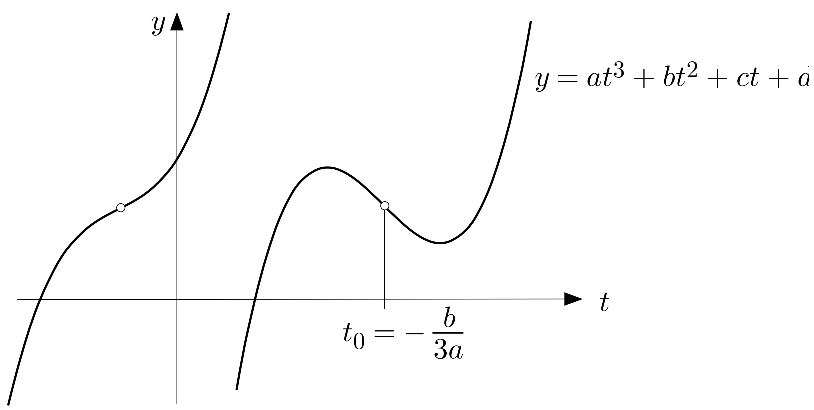

Fig. 7.5.3

Einige allgemeine Ungleichungen

Allgemeine Ungleichungen, das heisst: Ungleichungen mit einer oder mehreren freien Variablen, spielen eine grosse Rolle sowohl in der Analysis wie in der Geometrie. So manche von ihnen lassen sich beweisen, indem man eine passende konvexe Funktion ins Spiel bringt. Wir beginnen mit der folgenden Verallgemeinerung der **Bernoullischen Ungleichung (2.2)**:

(7.27) *Für alle* $t > -1,\ \ t \neq 0$, *gilt*

$$(1+t)^\alpha > 1 + \alpha t \qquad (\alpha < 0 \ \ \textit{oder} \ \ \alpha > 1)$$

bzw.

$$(1+t)^\alpha < 1 + \alpha t \qquad (0 < \alpha < 1)\ .$$

⌜ Die Funktion $f(t) := (1+t)^\alpha$ besitzt die Ableitungen

$$f'(t) = \alpha(1+t)^{\alpha-1}\,, \qquad f''(t) = \alpha(\alpha-1)(1+t)^{\alpha-2}\,;$$

somit gilt für alle betrachteten Werte von α und t:

$$f'(0) = \alpha\,, \qquad \operatorname{sgn} f''(t) = \operatorname{sgn}\alpha \cdot \operatorname{sgn}(\alpha - 1)\ .$$

Hiernach ist f streng konvex, falls $\alpha < 0$ oder $\alpha > 1$, und streng konkav für $0 < \alpha < 1$ (siehe die Fig. 7.5.4). An der Stelle 0 besitzt f die Stützfunktion

$$s(t) := f(0) + f'(0)t = 1 + \alpha t\,;$$

die Behauptung folgt daher mit **(7.25)**(c). ⌟

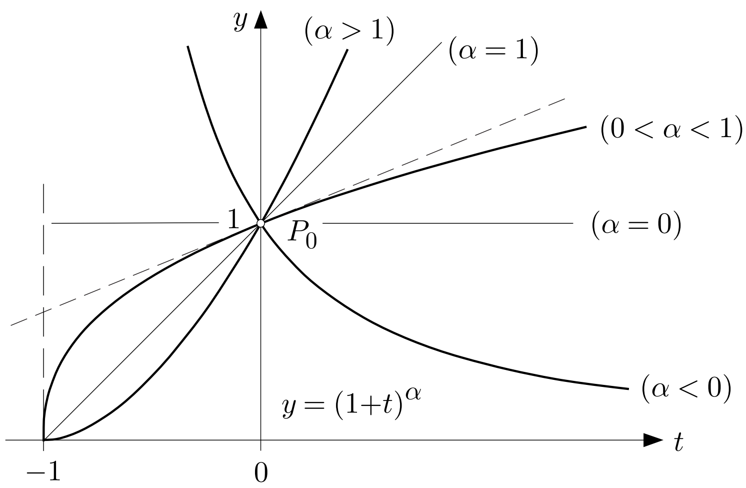

Fig. 7.5.4

Für das Weitere benötigen wir eine allgemeinere, das heisst: mit zusätzlichen Variablen “aufgeladene” Fassung der definierenden Ungleichung (1). Hierzu führen wir zunächst den folgenden Begriff ein: Ein reelles n-Tupel $\mathbf{p} = (p_1, \ldots, p_n)$ mit

$$p_k \geq 0 \quad (1 \leq k \leq n) , \qquad \sum_{k=1}^{n} p_k = 1 \tag{4}$$

heisst ein **Gewichtssatz**. Ist $\mathbf{p}$ ein Gewichtssatz und sind $t_1, \ldots, t_n$ beliebige reelle Zahlen, so lässt sich die Zahl

$$t^* := \sum_{k=1}^{n} p_k t_k$$

als **Schwerpunkt** von n in den Punkten t_k angebrachten Punktmassen p_k betrachten. Wie man leicht verifiziert, gilt

$$\min_{1 \leq k \leq n} t_k \leq \sum_{k=1}^{n} p_k t_k \leq \max_{1 \leq k \leq n} t_k ,$$

so dass auch der Punkt t^* in dem Intervall I liegt, falls das für die gegebenen Punkte $t_1, \ldots, t_n$ zutrifft.

Hier also die sogenannte **Jensensche Ungleichung**:

(7.28) *Ist f streng konvex auf dem Intervall I, so gilt für beliebige Punkte $t_1, \ldots, t_n \in I$ und beliebigen Gewichtssatz* $\mathbf{p}$ *die Beziehung*

$$f\left(\sum_{k=1}^{n} p_k t_k\right) \leq \sum_{k=1}^{n} p_k f(t_k) . \tag{5}$$

Ist dabei $p_k > 0$ für wenigstens zwei verschiedene t_k, so gilt in (5) *die strenge Ungleichung.*

In Worten: Bei einer konvexen Funktion ist der Funktionswert im Schwerpunkt von n gewichteten Punkten stets kleiner als das gewichtete Mittel der Funktionswerte in diesen n Punkten.

⌜ Ist p_k für keine zwei verschiedene t_k positiv, so gilt in (5) trivialerweise das Gleichheitszeichen. Wir wollen daher von nun an (nach eventueller Umnumerierung)

$$p_1 p_2 (t_2 - t_1) > 0 \tag{6}$$

annehmen. — Im Fall $n = 2$ ist (5) eine Paraphrase der definierenden Ungleichung (1). Um dies einzusehen, betrachten wir den Schwerpunkt

$$t := p_1 t_1 + p_2 t_2 .$$

Mit (4) ergibt sich

$$\begin{aligned} t_2 - t &= (1-p_2)t_2 - p_1 t_1 = p_1(t_2 - t_1)\ , \\ t - t_1 &= (p_1 - 1)t_1 + p_2 t_2 = p_2(t_2 - t_1) \end{aligned}$$

(und damit insbesondere $t_1 < t < t_2$), wir haben daher

$$\frac{t_2 - t}{t_2 - t_1} = p_1\ , \qquad \frac{t - t_1}{t_2 - t_1} = p_2\ .$$

Damit geht (1) über in die behauptete Ungleichung

$$f(p_1 t_1 + p_2 t_2) < p_1 f(t_1) + p_2 f(t_2)\ .$$

Wir nehmen jetzt an, die Behauptung sei richtig für n, $n \ge 2$, und betrachten $n+1$ Punkte $t_1, \ldots, t_{n+1} \in I$ sowie einen Gewichtssatz $(p_1, \ldots, p_{n+1})$. Wegen (6) ist $1 - p_{n+1} = \sum_{k=1}^n p_k > 0$; somit wird durch

$$p'_k := \frac{p_k}{1 - p_{n+1}} \qquad (1 \le k \le n)$$

ein Satz $\mathbf{p}'$ von n Gewichten festgelegt. Wir setzen

$$t^* := \sum_{k=1}^n p'_k t_k\ ,$$

dann ist $t^* \in I$, und wir erhalten mit Hilfe der Induktionsvoraussetzung

$$\begin{aligned} f\Bigl(\sum_{k=1}^{n+1} p_k t_k\Bigr) &= f\bigl((1-p_{n+1})t^* + p_{n+1}t_{n+1}\bigr) \\ &\le (1-p_{n+1}) f(t^*) + p_{n+1} f(t_{n+1}) \\ &< (1-p_{n+1}) \sum_{k=1}^n p'_k f(t_k)\ +\ p_{n+1} f(t_{n+1}) \\ &= \sum_{k=1}^{n+1} p_k f(t_k)\ . \end{aligned}$$

┘

② Für jedes reelle $p \ge 1$ ist die sogenannte ***p*-Norm** eines Vektors $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ bzw. $\in \mathbb{C}^n$ definiert durch

$$\|\mathbf{x}\|_p := \Bigl(\sum_{k=1}^n |x_k|^p\Bigr)^{1/p}\ ,$$

sie stellt wegen $\|\mathbf{x}\|_2 = |\mathbf{x}|$ eine Verallgemeinerung der Betragsfunktion dar. Beispiel 3.4.⑥ legt nahe, auch noch eine ∞-Norm zu definieren durch

$$\|\mathbf{x}\|_\infty := \max_{1 \le k \le n} |x_k| \ .$$

Wie man leicht verifiziert, besitzt $\|\cdot\|_p$ für jedes p, $1 \le p \le \infty$, die Eigenschaften

$$\|\mathbf{x}\|_p \ge 0 \,, \qquad \|\mathbf{x}\|_p = 0 \Leftrightarrow \mathbf{x} = \mathbf{0} \,;$$

$$\|\lambda\mathbf{x}\|_p = |\lambda| \, \|\mathbf{x}\|_p \qquad (\lambda \in \mathbb{R} \ \text{ bzw. } \in \mathbb{C}) \ .$$

Weniger selbstverständlich ist die Dreiecksungleichung, die in dieser allgemeinen Form auch als **Minkowskische Ungleichung** bezeichnet wird:

(7.29) *Für jedes $p \ge 1$ (auch für $p := \infty$) gilt*

$$\|\mathbf{x} + \mathbf{y}\|_p \le \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$$

bzw.

$$\left(\sum_{k=1}^n |x_k + y_k|^p\right)^{1/p} \le \left(\sum_{k=1}^n |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^n |y_k|^p\right)^{1/p} . \tag{7}$$

⌜ Die Fälle $p = 1$ und $p = \infty$ sind trivial. — Wegen

$$|x_k + y_k|^p \le \bigl(|x_k| + |y_k|\bigr)^p$$

genügt es, (7) für nichtnegative reelle x_k, y_k zu beweisen. Hierzu setzen wir

$$x_k + y_k =: t_k$$

und definieren einen Gewichtssatz $\mathbf{p}$ durch

$$p_k := \frac{t_k^p}{T} \quad (1 \le k \le n), \qquad T := \sum_{k=1}^n t_k^p \tag{8}$$

(im Fall $T = 0$ ist nichts zu beweisen). Ferner lassen sich Zahlen ξ_k, $\eta_k \in [\,0,1\,]$ bestimmen mit

$$x_k = \xi_k t_k \,, \quad y_k = \eta_k t_k \,, \quad \xi_k + \eta_k = 1 \qquad (1 \le k \le n) \ . \tag{9}$$

Nach diesen Vorbereitungen kommt nun Satz **(7.28)** folgendermassen zum Zug: Die Funktion $\xi \to \xi^p$, $p > 1$, ist auf dem Intervall $[\,0,1\,]$ streng konvex. Somit haben wir

$$\sum_{k=1}^{n} x_k^p = T \sum_{k=1}^{n} p_k\, \xi_k^p \geq T \left(\sum_{k=1}^{n} p_k \xi_k \right)^p .$$

Folglich ist

$$\|\mathbf{x}\|_p \geq T^{1/p} \sum_{k=1}^{n} p_k \xi_k$$

und analog

$$\|\mathbf{y}\|_p \geq T^{1/p} \sum_{k=1}^{n} p_k \eta_k \ .$$

Addiert man diese beiden Ungleichungen, so ergibt sich wegen (8) und (9):

$$\|\mathbf{x}\|_p + \|\mathbf{y}\|_p \geq T^{1/p} = \|\mathbf{x} + \mathbf{y}\|_p \ .$$

③ Aus

$$\log''(t) = -\frac{1}{t^2} < 0 \qquad \forall t > 0$$

folgt: Die Logarithmusfunktion ist streng konkav auf ihrem ganzen Definitionsbereich. Somit gilt für beliebige positive Zahlen $t_1, \ldots, t_n$ und beliebigen Gewichtssatz $\mathbf{p}$ die Relation

$$\log\left(\sum_{k=1}^{n} p_k t_k \right) \geq \sum_{k=1}^{n} p_k \log t_k = \sum_{k=1}^{n} \log\bigl(t_k^{p_k}\bigr) = \log\left(\prod_{k=1}^{n} t_k^{p_k} \right) ,$$

und wir erhalten

(7.30)
$$\sum_{k=1}^{n} p_k t_k \geq \prod_{k=1}^{n} t_k^{p_k} \qquad (t_1, \ldots, t_n > 0) \ .$$

Dies ist die **Ungleichung zwischen dem arithmetischen und dem geometrischen Mittel**. Im Spezialfall $p_k := 1/n \ \ (1 \leq k \leq n)$ lautet sie:

$$\frac{t_1 + \ldots + t_n}{n} \geq \sqrt[n]{t_1 \cdot \ldots \cdot t_n} \, ,$$

und zwar gilt das Gleichheitszeichen genau dann, wenn alle t_k gleich sind.

④ Die **Höldersche Ungleichung** (10) des nachstehenden Satzes ist das zur p-Norm gehörige Analogon zur Schwarzschen Ungleichung 2.7.(2).

(7.31) *Stehen die beiden Zahlen $p > 1$ und $q > 1$ in der Beziehung*

$$\frac{1}{p} + \frac{1}{q} = 1 ,$$

so gilt für beliebige Vektoren $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$ bzw. $\in \mathbb{C}^n$:

$$\left| \sum_{k=1}^{n} x_k \bar{y}_k \right| \leq \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} \cdot \left(\sum_{k=1}^{n} |y_k|^q \right)^{1/q} . \tag{10}$$

⌜ Wegen $\left|\sum_{k=1}^{n} x_k \bar{y}_k\right| \leq \sum_{k=1}^{n} |x_k|\,|y_k|$ genügt es, (10) für nichtnegative reelle x_k, y_k zu beweisen.

Ist $\mathbf{x} = \mathbf{0}$ oder $\mathbf{y} = \mathbf{0}$, so ist (10) trivialerweise richtig. Aus "Homogenitätsgründen" dürfen wir daher ohne Einschränkung der Allgemeinheit $\|\mathbf{x}\|_p = 1$, $\|\mathbf{y}\|_q = 1$ annehmen; die Behauptung erhält damit die Form

$$\sum_{k=1}^{n} x_k y_k \leq 1 .$$

Nun ist $x_k \cdot y_k$ das geometrische Mittel der beiden Zahlen x_k^p und y_k^q mit den (exponentiellen) Gewichten $1/p$ und $1/q$. Mit **(7.30)** ergibt sich daher

$$x_k y_k \leq \frac{1}{p}\, x_k^p + \frac{1}{q}\, y_k^q \qquad (1 \leq k \leq n) .$$

Addiert man diese n Ungleichungen, so folgt in der Tat

$$\sum_{k=1}^{n} x_k y_k \leq \frac{1}{p} \sum_{k=1}^{n} x_k^p + \frac{1}{q} \sum_{k=1}^{n} y_k^q = \frac{1}{p} + \frac{1}{q} = 1 .$$ ⌟

○

⑤ Wir zeigen: Unter allen einem Kreise umbeschriebenen n-Ecken hat das reguläre den kleinsten Umfang.

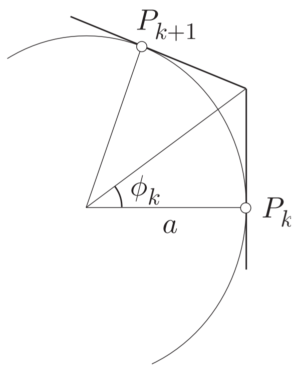

Fig. 7.5.5

Wir verweisen auf die Figur 7.5.5. — Je zwei aufeinanderfolgende Berührungspunkte P_k, P_{k+1} bestimmen einen Zentriwinkel $2\phi_k$. Der zugehörige Beitrag zum Umfang U des ganzen n-Ecks beträgt $2a\tan\phi_k$, wobei a den Kreisradius bezeichnet. Damit ergibt sich

$$U = 2a\sum_{k=1}^{n}\tan\phi_k = 2a\,n\sum_{k=1}^{n}\frac{1}{n}\tan\phi_k \ .$$

Die ϕ_k liegen im Intervall $\left]\,0, \frac{\pi}{2}\,\right[$, und dort ist tan eine streng konvexe Funktion, wie man leicht nachrechnet. Wegen $\sum_{k=1}^{n}\phi_k = \pi$ gilt somit nach **(7.28)**:

$$\sum_{k=1}^{n}\frac{1}{n}\tan\phi_k \geq \tan\left(\sum_{k=1}^{n}\frac{1}{n}\phi_k\right) = \tan\frac{\pi}{n} \ .$$

Setzen wir dies oben ein, so folgt

$$U \geq 2a\,n\tan\frac{\pi}{n} = U_{\mathrm{reg}} \ ,$$

und zwar gilt das Gleichheitszeichen nur, wenn alle ϕ_k gleich sind. In anderen Worten: Alle nicht-regulären Tangenten-n-Ecke haben einen echt grösseren Umfang als das reguläre derartige n-Eck. ○

Aufgaben

1. (a) Unter allen einem Kreise einbeschriebenen n-Ecken hat das reguläre den grössten Flächeninhalt. (Man darf offenbar annehmen, der Kreismittelpunkt liege im Innern des n-Ecks.)

   (b) Unter allen Dreiecken von gegebenem Umfang $2s$ hat das gleichseitige den grössten Flächeninhalt. *Hinweis:* Heronsche Formel.

2. Ein Versuch habe n mögliche Ergebnisse, die bzw. mit Wahrscheinlichkeiten $p_k > 0 \quad (1 \leq k \leq n)$ eintreten, $\sum_{k=1}^{n} p_k = 1$. Ein natürliches Mass für die Ungewissheit über den Ausgang des Versuchs ist die sogenannte **Entropie**

   $$H := -\sum_{k=1}^{n} p_k \log p_k \ .$$

   Zeige: Die Ungewissheit ist am grössten, wenn alle Ergebnisse gleich wahrscheinlich sind.

3. (a) Für beliebige nichtnegative Zahlen a, b, c gilt:

   $$\left(\frac{a+b+c}{3}\right)^3 \leq \frac{a^3+b^3+c^3}{3} \ .$$

   (b) Formuliere und beweise eine möglichst allgemeine Ungleichung dieser Art.

7.6 Taylor-Approximation

Qualität von Nullstellen

Ist $a \in \mathbb{R}$ eine Nullstelle des Polynoms $p(\cdot)$, so gibt es bekanntlich eine wohlbestimmte Zahl $m > 0$, die **Vielfachheit** dieser Nullstelle, und ein Polynom p_0 mit

$$p(t) = (t-a)^m \, p_0(t) \, , \qquad p_0(a) \neq 0 \; .$$

Hieraus folgt

$$p'(t) = (t-a)^{m-1}\bigl(mp_0(t) + (t-a)p_0'(t)\bigr) =: (t-a)^{m-1}\, p_1(t)$$

mit $p_1(a) = mp_0(a) \neq 0$ und weiter mit vollständiger Induktion:

$$p^{(k)}(t) = (t-a)^{m-k}\, p_k(t) \, , \quad p_k(a) \neq 0 \qquad (0 \leq k \leq m) \; .$$

Insbesondere gilt unter diesen Umständen

$$p(a) = p'(a) = \; \ldots \; = p^{(m-1)}(a) = 0 \, , \qquad p^{(m)}(a) \neq 0 \; .$$

Es sei jetzt $g : \mathbb{R} \curvearrowright \mathbb{X}$ eine beliebige Funktion. Ist g hinreichend oft differenzierbar, so lässt sich die Qualität der Nullstellen von g mit der Qualität von Polynom-Nullstellen vergleichen. Das zeigt sich schon in den beiden folgenden Hilfssätzen:

(7.32) *Die reellwertige Funktion g sei n-mal, $n \geq 0$, stetig differenzierbar auf dem Intervall $[\,a, t\,]$ und $(n+1)$-mal differenzierbar auf $]\,a, t\,[$; ferner sei*

$$g(a) = g'(a) = \; \ldots \; = g^{(n)}(a) = 0 \; .$$

Dann gibt es ein $\tau \in \,]\,a, t\,[$ mit

$$g(t) = \frac{g^{(n+1)}(\tau)}{(n+1)!}\,(t-a)^{n+1} \; .$$

$\ulcorner$ Für $n = 0$ ergibt sich die Behauptung unmittelbar aus dem Mittelwertsatz **(7.16)**; wir dürfen daher annehmen, dass sie für $n-1$ zutrifft.
Die Hilfsfunktion

$$\phi(s) := g(t)(s-a)^{n+1} - g(s)(t-a)^{n+1}$$

verschwindet an den Stellen a und t. Nach dem Satz von Rolle gibt es daher ein $\sigma \in \,]\,a, t\,[$ mit $\phi'(\sigma) = 0$, und das heisst

$$g(t)\,(n+1)(\sigma-a)^n = g'(\sigma)(t-a)^{n+1} \; . \tag{1}$$

Die Funktion g' genügt auf dem Intervall $[\,a,\sigma\,]$ den Voraussetzungen mit $n-1$ anstelle von n; nach Induktionsannahme gibt es daher ein $\tau \in]\,a,\sigma\,[$ mit

$$g'(\sigma) = \frac{g'^{(n)}(\tau)}{n!}(\sigma - a)^n ,$$

und zusammen mit (1) folgt

$$g(t) = \frac{g^{(n+1)}(\tau)}{(n+1)!}(t-a)^{n+1} ,$$

was zu beweisen war. $\lrcorner$

(7.33) *Die Funktion $g : \mathbb{R} \curvearrowright \mathbb{X}$ sei n-mal, $n \geq 1$, differenzierbar an der Stelle a, und es gelte*

$$g(a) = g'(a) = \ \ldots \ = g^{(n)}(a) = 0 \ .$$

Dann gibt es eine an der Stelle a stetige Funktion $\tilde{g}$ mit

$$g(t) = (t-a)^n \, \tilde{g}(t) \ , \qquad \tilde{g}(a) = 0 \ . \tag{2}$$

$\ulcorner$ Der Fall $n = 1$ wurde in **(7.2)**(a) behandelt; im weiteren sei daher $n \geq 2$. Wir können (2) als Definition der Funktion $\tilde{g}\colon \mathrm{dom}\,(g) \to \mathbb{X}$ betrachten und haben zu zeigen, dass $\tilde{g}$ an der Stelle a stetig ist. Dabei dürfen wir g reellwertig annehmen, und es genügt, die rechtsseitige Stetigkeit nachzuweisen.

Für alle hinreichend nahe bei a gelegenen $t > a$ genügt g den Voraussetzungen von **(7.32)** mit $n-2$ anstelle von n; darüber hinaus ist

$$g^{(n-1)}(a) = 0 \, , \qquad g^{(n)}(a) = 0 \ . \tag{3}$$

Es gibt somit einen Punkt $\tau \in]\,a,t\,[$ mit

$$g(t) = \frac{1}{(n-1)!}\,(t-a)^{n-1} g^{(n-1)}(\tau) \ . \tag{4}$$

Die Funktion $g^{(n-1)}$ ist an der Stelle a differenzierbar. Nach **(7.2)**(a) und (3) gibt es daher weiter eine in a stetige und dort verschwindende Funktion $m(\cdot)$, so dass für alle t gilt:

$$g^{n-1}(t) = (t-a)m(t) \ .$$

Aufgrund von (4) erhalten wir nunmehr

$$\tilde{g}(t) := \frac{g(t)}{(t-a)^n} = \frac{1}{(n-1)!}\,\frac{\tau - a}{t-a}\,m(\tau) \qquad (t > a) \, ,$$

und hier strebt die rechte Seite mit $t \to a+$ gegen 0, wie behauptet. $\lrcorner$

Konstruktion der Taylor-Polynome

Im weiteren geht es darum, eine gegebene (hinreichend oft differenzierbare) Funktion $f\colon \mathbb{R} \curvearrowright \mathbb{X}$ in der Umgebung einer festen Stelle $a \in \operatorname{dom}(f)$ durch Polynome p zu approximieren, und zwar so, dass der Fehler $f-p$ an der Stelle a "von möglichst hoher Ordnung verschwindet". Satz **(7.33)** zeigt, dass der Fehler mit $t \to a$ *schneller* als $(t-a)^n$ gegen 0 geht, falls die Ableitungen von p bis zur Ordnung n mit denjenigen von f übereinstimmen. Soll p ausserdem einen Grad $\leq n$ haben, so ist p damit eindeutig festgelegt.

⌜ Besitzen sowohl p_1 wie p_2 die genannten Eigenschaften, so genügt die Funktion $g(t) := p_1(t) - p_2(t)$ den Voraussetzungen von **(7.32)**; überdies ist $g^{(n+1)}(t) \equiv 0$. Hieraus folgt $g(t) \equiv 0$. ⌟

Damit stehen wir vor dem folgenden algebraischen Problem: Gesucht ist eine Folge

$$j_a^0 f, \quad j_a^1 f, \quad j_a^2 f, \ \ldots$$

von Polynomen mit den folgenden Eigenschaften:

$$\text{(a)} \qquad \deg(j_a^n f) \leq n \qquad (n \geq 0)\,,$$

$$\text{(b)} \qquad \bigl(j_a^n f\bigr)^{(k)}(a) \ = \ f^{(k)}(a) \qquad (0 \leq k \leq n)\ .$$

Man verifiziert ohne weiteres, dass

$$j_a^{-1} f(t) :\equiv 0\,, \qquad j_a^0 f(t) :\equiv f(a)\,, \qquad j_a^1 f(t) := f(a) + f'(a)(t-a)$$

diese Eigenschaften besitzen.

Wir nehmen nun an, $j_a^{r-1} f$ sei bestimmt, und setzen

$$j_a^r f := j_a^{r-1} f + p\,,$$

wobei das Polynom p noch unbekannt ist. Damit alles richtig hinkommt, muss p den folgenden Bedingungen genügen:

$$\text{(a}'\text{)} \qquad \deg(p) \leq r\,,$$

$$\text{(b}'\text{)} \qquad p(a) = p'(a) = \ \ldots \ = p^{(r-1)}(a) = 0\,, \qquad p^{(r)}(a) = f^{(r)}(a)\ .$$

Es gilt nämlich zu beachten, dass j_a^{r-1} an der Stelle a bis zur Ordnung $r-1$ schon die vorgeschriebenen Ableitungswerte besitzt und dass die r-te Ableitung von j_a^{r-1} identisch verschwindet. Mit

$$p(t) := \frac{f^{(r)}(a)}{r!}\,(t-a)^r$$

sind (a′) und (b′) offensichtlich erfüllt. Damit ergibt sich für die eigentlich gesuchten Polynome $j_a^r f$ die Rekursionsformel

$$j_a^r f(t) = j_a^{r-1} f(t) + \frac{f^{(r)}(a)}{r!}\,(t-a)^r\ ,$$

und durch Aufsummieren erhalten wir schliesslich definitiv

$$\begin{aligned} j_a^n f(t) &= f(a) + \frac{f'(a)}{1!}(t-a) + \frac{f''(a)}{2!}(t-a)^2 + \ \ldots\ + \frac{f^{(n)}(a)}{n!}(t-a)^n \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(t-a)^k\ . \end{aligned}$$

Man nennt $j_a^n f$ den **n-Jet** oder das **n-te Taylorsche Approximationspolynom von f an der Stelle a**. Die formale Reihe

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

heisst **Taylor-Reihe von f an der Stelle a**.

Die Figur 7.6.1 zeigt einige n-Jets der Sinusfunktion an der Stelle $a := 0$. Man erkennt deutlich, wie sich der t-Bereich, in dem j^n eine brauchbare Approximation liefert, mit wachsendem n vergrössert.

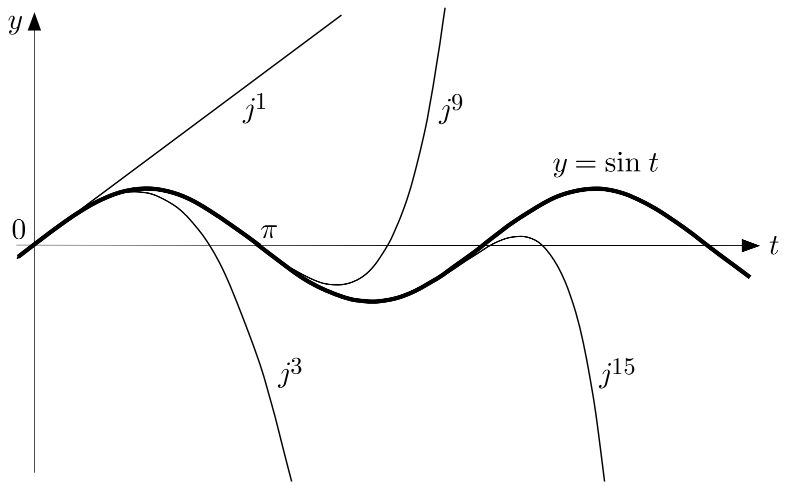

Fig. 7.6.1

① Betrachte die Funktion

$$f(t) := \log(1+t) \qquad (t > -1)$$

in der Umgebung von $a := 0$. Man hat

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

und allgemein

$$f^{(k)}(t) = \frac{(-1)^{k-1}(k-1)!}{(1+t)^k} \qquad (k \ge 1)\ . \tag{5}$$

Damit ergibt sich

$$f(0) = 0\,, \qquad \frac{f^{(k)}(0)}{k!} = \frac{(-1)^{k-1}}{k} \quad (k \ge 1)\,,$$

so dass wir für f an der Stelle 0 die folgende Taylor-Entwicklung erhalten:

$$j_0^\infty f(t) = t - \frac{t^2}{2} + \frac{t^3}{3} - \frac{t^4}{4} + \ldots \quad .$$

○

Untersuchung des Restglieds

Damit stellen sich die folgenden Fragen:

— Wie gut approximiert $j_a^n f$ die Funktion f?

— Gilt $j_a^\infty f = f$?

Man nennt die Differenz

$$R_n(t) := f(t) - j_a^n f(t)$$

das **n-te Restglied** der Taylor-Entwicklung. Im Hinblick auf $j_a^\infty f = f$ (was in vielen Fällen zutrifft) spricht man auch hier vom **Abbrechfehler**. Wir werden im weiteren verschiedene Darstellungen des Restglieds herleiten. Man benötigt sie für Konvergenzuntersuchungen und für die numerische Abschätzung des Abbrechfehlers, nicht aber, um den Funktionswert auf einem unerhörten Umweg doch noch exakt zu berechnen.

(7.34) *Die reellwertige Funktion f sei n-mal stetig differenzierbar auf dem Intervall $[\,a,t\,]$ und $(n+1)$-mal differenzierbar auf $]\,a,t\,[$. Dann gibt es ein $\tau \in\,]\,a,t\,[$ mit*

$$R_n(t) = \frac{f^{(n+1)}(\tau)}{(n+1)!}(t-a)^{n+1}\ ;$$

das heisst, es gilt

$$f(t) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(t-a)^k \ + \frac{f^{(n+1)}(\tau)}{(n+1)!}(t-a)^{n+1}\ .$$

⌜ Erstens genügt die Funktion $R_n := f - j_a^n f$ auf dem Intervall $[\,a,t\,]$ den Voraussetzungen von **(7.32)**, und zweitens ist $R_n^{(n+1)}(\tau) = f^{(n+1)}(\tau)$, da das Polynom $j_a^n f$ einen Grad $\le n$ hat. ⌟

Hieraus ergibt sich das folgende Korollar:

(7.35) *Die Funktion* $\mathbf{f} : [\,a, b\,] \to \mathbb{X}$ *sei* $(n+1)$*-mal differenzierbar, und es gelte*

$$\left|\mathbf{f}^{(n+1)}(t)\right| \le M \qquad \forall t \in \,]\,a, b\,[\ .$$

Dann ist

$$\left|\mathbf{f}(b) - j_a^n \mathbf{f}(b)\right| \le \frac{M}{(n+1)!}(b-a)^{n+1}\ .$$

⌜ (Vergleiche den Beweis von Satz **(7.19)**.) Wir setzen $\mathbf{f}(b) - j_a^n \mathbf{f}(b) =: \mathbf{c}$ und betrachten die reellwertige Hilfsfunktion

$$\phi(t) := \mathbf{c} \bullet \mathbf{f}(t)\ .$$

Es gilt $\phi^{(k)} = \mathbf{c} \bullet \mathbf{f}^{(k)}$ für alle k und somit $j_a^n \phi = \mathbf{c} \bullet j_a^n \mathbf{f}$. Damit haben wir einerseits

$$\phi(b) - j_a^n \phi(b) = \mathbf{c} \bullet \bigl(\mathbf{f}(b) - j_a^n \mathbf{f}(b)\bigr) = |\mathbf{c}|^2\ ;$$

anderseits gibt es nach **(7.34)** ein $\tau \in \,]\,a, b\,[$ mit

$$\phi(b) - j_a^n \phi(b) = \phi^{(n+1)}(\tau)\frac{(b-a)^{n+1}}{(n+1)!} = \mathbf{c} \bullet \mathbf{f}^{(n+1)}(\tau)\,\frac{(b-a)^{n+1}}{(n+1)!}\ .$$

Somit ist

$$\begin{aligned}|\mathbf{c}|^2 &= \mathbf{c} \bullet \mathbf{f}^{(n+1)}(\tau)\,\frac{(b-a)^{n+1}}{(n+1)!} \le |\mathbf{c}| \cdot \left|\mathbf{f}^{(n+1)}(\tau)\right|\frac{(b-a)^{n+1}}{(n+1)!} \\ &\le |\mathbf{c}| \cdot \frac{M}{(n+1)!}\,(b-a)^{n+1}\ .\end{aligned}$$

Nach Division mit $|\mathbf{c}|$ ergibt sich hieraus die Behauptung. ⌟

① (Forts.) Aus (5) folgt

$$\left|f^{(n+1)}(\tau)\right| = \frac{n!}{(1+\tau)^{n+1}} \le \frac{n!}{(1-|t|)^{n+1}} \qquad (0 < |\tau| < |t| < 1)\ ,$$

so dass Satz **(7.34)** die folgende Abschätzung liefert:

$$|R_n(t)| \le \frac{1}{n+1}\left(\frac{|t|}{1-|t|}\right)^{n+1} \qquad (|t| < 1)\ . \tag{6}$$

Ist $|t| \le \frac{1}{2}$, so ist $\dfrac{|t|}{1-|t|} \le 1$; für diese t gilt daher

$$\lim_{n\to\infty}\bigl(f(t) - j_0^n f(t)\bigr) = \lim_{n\to\infty} R_n(t) = 0\ .$$

Damit haben wir folgendes bewiesen: Ist $|t| \leq \frac{1}{2}$, so gilt

$$\log(1+t) = \lim_{n\to\infty} j_0^n f(t) = j_0^\infty f(t) = t - \frac{t^2}{2} + \frac{t^3}{3} - \ \dots \quad .$$

Die rechter Hand aufgebaute Potenzreihe heisst sinngemäss **Logarithmusreihe**. Wie wir noch sehen werden, stellt sie die Funktion $\log(1+t)$ sogar in dem grösseren t-Intervall $]-1, 1]$ dar.

Wir geben noch ein numerisches Beispiel: Es soll $\log(4/5)$ berechnet werden. Hierzu verwenden wir $j_0^3 f$ und erhalten

$$\log\frac{4}{5} = \log\left(1 + (-\frac{1}{5})\right) \doteq -\frac{1}{5} - \frac{1}{2\cdot 25} - \frac{1}{3\cdot 125} = -0.2226667\dots \quad ,$$

wobei (6) einen Fehler

$$\leq \frac{1}{4}\left(\frac{1/5}{4/5}\right)^4 = \frac{1}{1024}$$

garantiert. Der Tabellenwert ist -0.2231435. ○

Im Anschluss an **(7.33)** wurde bereits angedeutet, wie schnell R_n nach 0 geht, wenn t gegen a strebt:

(7.36) *Die Funktion $f : \mathbb{R} \curvearrowright \mathbb{X}$ sei n-mal, $n \geq 1$, differenzierbar an der Stelle a. Dann gilt*

$$R_n(t) = r(t)(t-a)^n$$

bzw.

$$f(t) = j_a^{n-1} f(t) + \left(\frac{f^{(n)}(a)}{n!} + r(t)\right)(t-a)^n$$

mit

$$r(a) = \lim_{t\to a} r(t) = 0 \ . \tag{7}$$

⌜ R_n genügt den Voraussetzungen von Hilfssatz **(7.33)**. ⌟

Die o-Schreibweise erlaubt, den Satz **(7.36)** in der folgenden kompakten Weise zu formulieren:

(7.36′) $$f(t) = j_a^n f(t) + o\big((t-a)^n\big) \qquad (t \to a) \ .$$

Anwendung: Analyse von kritischen Punkten

Satz **(7.36)** setzt uns in Stand, das Verhalten einer hinreichend oft differenzierbaren Funktion in der Umgebung eines kritischen Punktes genauer zu untersuchen:

(7.37) *Es sei t_0 ein kritischer Punkt der Funktion $f : \mathbb{R} \curvearrowright \mathbb{R}$, und zwar gelte für ein $n \geq 2$:*

$$f'(t_0) = f''(t_0) = \ \ldots \ = f^{(n-1)}(t_0) = 0 \, , \qquad f^{(n)}(t_0) =: A \neq 0 \; .$$

Ist n gerade, so besitzt f an der Stelle t_0 ein lokales Minimum, falls $A > 0$, und ein lokales Maximum, falls $A < 0$. Ist n ungerade, so besitzt f an der Stelle t_0 kein lokales Extremum, hingegen einen Wendepunkt.

⌜ Der Punkt t_0 ist definitionsgemäss ein innerer Punkt von $\operatorname{dom}(f)$; die Grösse $t - t_0$ kann daher im folgenden beiderlei Vorzeichen annehmen. Satz **(7.36)** liefert in der vorliegenden Situation die Darstellung

$$f(t) = f(t_0) + \left(\frac{A}{n!} + r(t) \right) (t - t_0)^n \, ,$$

denn alle übrigen Glieder des $(n-1)$-Jets von f an der Stelle t_0 entfallen. Wegen $A \neq 0$ und (7) gilt daher für alle hinreichend nahe bei t_0 gelegenen t:

$$\operatorname{sgn}\bigl(f(t) - f(t_0)\bigr) = \operatorname{sgn} A \cdot \bigl(\operatorname{sgn}(t - t_0)\bigr)^n \; .$$

Hieraus folgen alle Behauptungen bezüglich eines allfälligen Extremums.

Es sei weiter $n \geq 3$ ungerade. Wenden wir Satz **(7.36)** mit $n-2$ anstelle von n auf f'' an, so erhalten wir in ähnlicher Weise

$$f''(t) = \left(\frac{A}{(n-2)!} + r_1(t) \right) (t - t_0)^{n-2}$$

und damit für hinreichend kleine $|t - t_0|$:

$$\operatorname{sgn} f''(t) = \operatorname{sgn} A \cdot \operatorname{sgn}(t - t_0) \; .$$

Folglich wechselt f'' an der Stelle t_0 das Vorzeichen. ⌟

② Es sollen die kritischen Stellen der Funktion

$$f(t) := (t-3)^3 (t+2)^2$$

untersucht werden. — Wir unterlassen tunlichst, das Produkt rechter Hand auszumultiplizieren, und berechnen nacheinander

$$\begin{aligned} f'(t) &= (t-3)^2(t+2)\bigl[3(t+2)+2(t-3)\bigr] \\ &= 5(t-3)^2(t+2)t\ , \\ f''(t) &= 5(t-3)\bigl[2(t+2)t+(t-3)t+(t-3)(t+2)\bigr] \\ &= 5(t-3)(4t^2-6)\ , \\ f'''(t) &= 5(4t^2-6)+40t(t-3)\ . \end{aligned}$$

Es gibt daher die drei kritischen Stellen -2, 0, 3. Wir legen die folgende Tabelle an, in die wir auch noch die zwei weiteren Nullstellen $\pm\sqrt{3/2}$ von f'' aufnehmen:

t_k	$f(t_k)$	$f'(t_k)$	$f''(t_k)$	$f'''(t_k)$	Typ
-2	0	0	-250		lokales Maximum
0	-108	0	90		lokales Minimum
3	0	0	0	150	Wendepunkt
$\pm\sqrt{3/2}$		$\neq 0$	0	$\neq 0$	Wendepunkte

Aufgrund dieser Daten sieht der Graph von f aus, wie in Fig. 7.6.2 dargestellt. ○

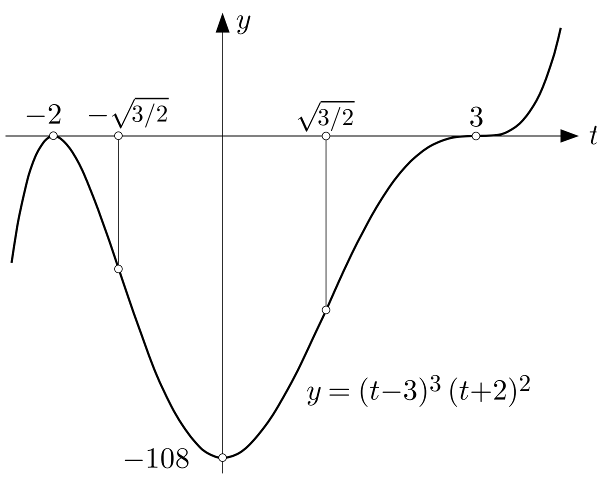

Fig. 7.6.2

Die Taylor-Reihe

Ist die Funktion f in einer Umgebung des Punktes a beliebig oft differenzierbar, so kann man die Taylor-Reihe

$$j_a^\infty f(t) := \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(t-a)^k \tag{8}$$

bilden. Unsere Approximationssätze berechtigen zur Hoffnung, dass (8) die Funktion f in einer geeigneten Umgebung von a **darstellt**, das heisst: dass es ein $\rho > 0$ gibt mit

$$j_a^\infty f(t) \;\equiv\; f(t) \qquad (a-\rho < t < a+\rho)\ . \tag{9}$$

Ist etwa f ein Polynom vom Grad $\leq n$, so ist dies trivialerweise erfüllt. In diesem Fall gilt sogar

$$j_a^\infty f = j_a^n f = f\ ,$$

denn f ist das einzige Polynom vom Grad $\leq n$, dessen Ableitungen an der Stelle a bis zur Ordnung n mit denen von f übereinstimmen. — Für die Exponentialfunktion erhält man wegen $\exp^{(k)}(0) = \exp 0 = 1$ $(k \geq 0)$:

$$j_0^\infty \exp(t) = \sum_{k=0}^{\infty} \frac{1}{k!} t^k = \exp t \qquad (t \in \mathbb{R})\ ;$$

in Worten: Die an der Stelle 0 angesetzte Taylor-Reihe von exp ist nichts anderes als die exp definierende Reihe. Analog verifiziert man, dass auch die bei 0 angesetzten Taylor-Reihen von cos und sin gerade die definierenden Reihen dieser Funktionen sind. Hier liegt ein allgemeiner Sachverhalt vor; wir werden in Kapitel 11 ausführlich darauf eingehen.

Ist eine Funktion nicht von vorneherein als Potenzreihe gegeben, so müsste man zur Etablierung von (9) untersuchen, ob das Restglied

$$R_n(t) = \frac{f^{(n+1)}(a)}{(n+1)!}\,(t-a)^{n+1}$$

auf einem t-Intervall $]\,a-\rho, a+\rho\,[$ mit $n \to \infty$ gegen 0 strebt. In vielen Fällen trifft dies zu (siehe Beispiel ①, Forts.); es gibt aber auch Gegenbeispiele.

③ Wir betrachten die Funktion

$$f(t) := \begin{cases} e^{-1/t} & (t > 0) \\ 0 & (t \leq 0) \end{cases}$$

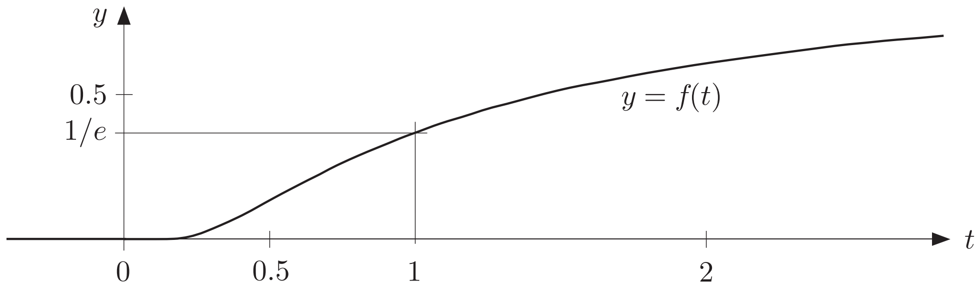

Fig. 7.6.3

(Fig. 7.6.2) und zeigen zunächst: f ist beliebig oft (stetig) differenzierbar auf ganz $\mathbb{R}$, und zwar sind die Ableitungen von f gegeben durch

$$f^{(k)}(t) = \begin{cases} p_{2k}\left(\dfrac{1}{t}\right)e^{-1/t} & (t > 0) \\ 0 & (t \le 0) \end{cases} \quad ; \tag{10}$$

dabei bezeichnet $p_{2k}(u)$ ein gewisses Polynom vom Grad $2k$ der Variablen u.

(10) ist trivialerweise richtig für $t < 0$ und alle k sowie für $t \ge 0$ und $k = 0$. Die Behauptung treffe weiter zu für alle $t \ge 0$ und ein $k \ge 0$. Betrachten wir jetzt erstens ein festes $t > 0$, so gilt in einer ganzen Umgebung von t die obere Zeile von (10); wir erhalten daher:

$$f^{(k+1)}(t) = p'_{2k}\left(\frac{1}{t}\right)\left(-\frac{1}{t^2}\right) \cdot e^{-1/t} + p_{2k}\left(\frac{1}{t}\right) \cdot e^{-1/t}\,\frac{1}{t^2} =: p_{2(k+1)}\left(\frac{1}{t}\right)e^{-1/t}\,,$$

wie behauptet. Zweitens gilt wegen **(6.4)**(c):

$$\begin{aligned} f^{(k+1)}(0+) &= \lim_{t\to 0+} \frac{f^{(k)}(t) - f^{(k)}(0)}{t - 0} = \lim_{t \to 0+} \left(\frac{1}{t}\,p_{2k}\left(\frac{1}{t}\right)e^{-1/t}\right) \\ &= \lim_{u\to\infty} \left(u\,p_{2k}(u)e^{-u}\right) = 0\,, \end{aligned}$$

und es ist trivialerweise $f^{(k+1)}(0-) = 0$. Damit ist auch $f^{(k+1)}(0) = 0$ bewiesen.

Aus (10) folgt natürlich

$$j_0^\infty f(t) \equiv 0\,,$$

wohingegen f selbst für alle $t > 0$ positive Werte annimmt.

Aufgaben

1. Bestimme $j_0^2\tan(t)$ sowie eine im Intervall $-\frac{\pi}{6} \le t \le \frac{\pi}{6}$ gültige Fehlerabschätzung. *Hinweis:* Verwende wiederholt $\tan' = 1 + \tan^2$.

2. Aus dem Additionstheorem für den Tangens folgt relativ leicht die Formel

$$\frac{\pi}{4} = \arctan\frac{1}{2} + \arctan\frac{1}{3}\ .$$

   Man benütze diese Formel und die Taylor-Entwicklung des Arcustangens, um $\pi/4$ mit einem Fehler von höchstens 10^{-4} zu berechnen.

3. Jemand möchte $\sinh 1$ auf hundert Dezimalstellen genau berechnen. Wieviele Glieder der Taylor-Entwicklung $j_0^\infty\sinh$ muss sie berücksichtigen?

4. Betrachte die Funktion $g(t) := \cos t - j_0^{14}\cos(t)$. Besitzt g an der Stelle $t = 0$ ein (lokales) Maximum, ein Minimum oder einen Wendepunkt? Die Antwort ist zu begründen.

5. Die Funktion $f\colon \mathbb{R} \curvearrowright \mathbb{R}$ sei in der Umgebung des Nullpunkts beliebig oft differenzierbar. Wann lässt sich auf Grund der Werte $a_k := f^{(k)}(0)$ $(k \in \mathbb{N})$ entscheiden, ob f im Nullpunkt ein lokales Maximum, ein Minimum oder kein lokales Extremum besitzt?

6. Konstruiere eine C^∞-Funktion $\phi\colon \mathbb{R} \to [\,0,1\,]$, die den folgenden Bedingungen genügt:

$$\phi(t) \equiv 1 \quad (|t| \le 1)\,, \qquad \phi(t) \equiv 0 \quad (|t| \ge 2)\ .$$

7.7 Das Newtonsche Verfahren

Die Idee

In diesem Abschnitt geht es um die approximative Lösung von Gleichungen der Form

$$f(x) = 0 ,$$

f eine gegebene Funktion. Die unabhängige Variable bezeichnen wir hier mit x, um den Eindruck einer "Unbekannten" zu vermitteln und auch, um anzudeuten, dass sich das Verfahren auf mehrdimensionale Situationen, das heisst: auf n Gleichungen in n Unbekannten, übertragen lässt.

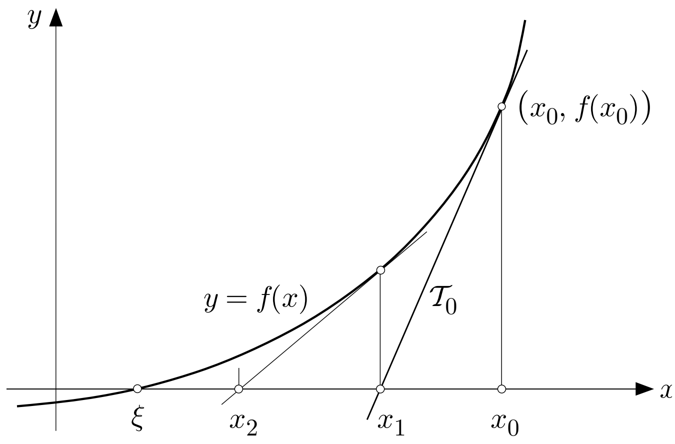

Fig. 7.7.1

Die Idee ist genial einfach (Fig. 7.7.1). Gesucht ist der Schnittpunkt ξ des Graphen $\mathcal{G}\colon\ y = f(x)$ mit der x-Achse. Es sei x_0 ein irgendwie gefundener Näherungswert für ξ. Wir ersetzen nun $\mathcal{G}$ durch die Tangente $\mathcal{T}_0$ im Punkt $\bigl(x_0, f(x_0)\bigr)$ und schneiden statt $\mathcal{G}$ diese Tangente mit der x-Achse. Es ist zu erwarten, dass der Schnittpunkt x_1 näher bei ξ liegt als x_0. Das Verfahren lässt sich beliebig oft wiederholen und liefert eine Folge $x_.$, die unter günstigen Umständen gegen ξ konvergiert. (Es kann aber auch schief gehen, wie Figur 7.7.2 zeigt.)

Die geometrische Konstruktion lässt sich analytisch wie folgt nachvollziehen: Die Graphentangente $\mathcal{T}_0$ besitzt die Gleichung $y = f(x_0) + f'(x_0)\,(x - x_0)$ und schneidet daher die x-Achse an der Stelle

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

Dieselbe Formel gilt auch für den allgemeinen Schritt $x_n \rightsquigarrow x_{n+1}$. Infolgedessen wird das Newtonsche Verfahren durch folgende Rekursionsvorschrift beschreiben:

$$\boxed{\begin{array}{rl} x_0 : & \text{hinreichend nahe bei } \xi \\[1ex] x_{n+1} := x_n - \dfrac{f(x_n)}{f'(x_n)} & (n \geq 0) \end{array}}$$

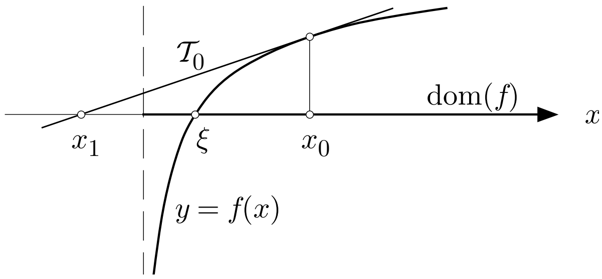

Fig. 7.7.2

Konvergenzverhalten

Wir wollen nun das Konvergenzverhalten der Folge $x_.$ untersuchen , wobei wir geeignete Annahmen treffen, die sicherstellen, dass x_0 unter den gegebenen Umständen “hinreichend nahe bei ξ” liegt. (In der Praxis werden diese Bedingungen natürlich nicht immer verifiziert. Man beginnt einfach mit einem plausiblen Näherungswert x_0 zu rechnen und sieht dann sehr bald, ob es klappt.)

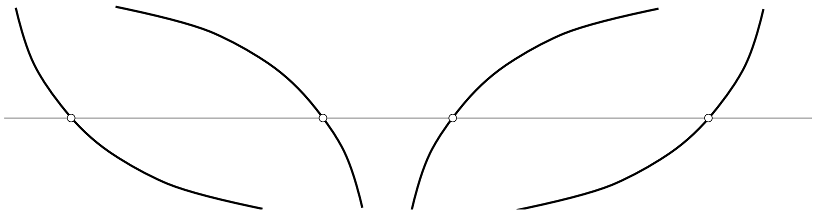

Fig. 7.7.3

Es sei ein Intervall $I := [\,a, b\,]$ so bestimmt, dass gilt:

$$\begin{gathered} f(a) \cdot f(b) < 0 \;, \\ \forall x \in I : \quad f'(x) f''(x) \;\neq\; 0 \;. \end{gathered}$$

Der Graph von $f \restriction I$ hat dann eine der vier in Fig. 7.7.3 dargestellten Formen. Insbesondere ist f streng monoton auf I und besitzt somit genau eine Nullstelle $\xi \in I$. Wir nehmen etwa an, es sei $f' > 0$ und $f'' > 0$ auf I und setzen $x_0 := b$. (Ein Blick auf Fig. 7.7.3 zeigt, dass man $x_0 := b$ setzen soll, wenn

f' und f'' auf I dasselbe Vorzeichen haben; $x_0 := a$ im andern Fall.) Dann treffen die folgenden Bedingungen für $n = 0$ zu:

$$\xi \;<\; x_n \;, \tag{1}$$

$$\forall x \in \,]\,\xi, x_n\,] : \qquad f(x) > 0 \,, \quad f'(x) > 0 \,, \quad f''(x) > 0 \;. \tag{2}$$

Wir berechnen nun x_1 und haben erstens

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} < x_0 \;. \tag{3}$$

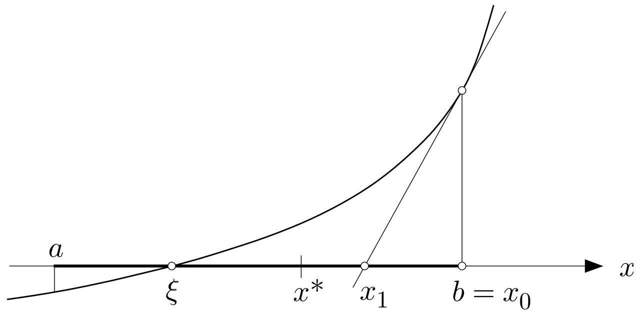

Fig. 7.7.4

Weiter gibt es nach Satz **(7.34)** einen Punkt x^* zwischen ξ und x_0 (siehe die Fig. 7.7.4) mit

$$0 = f(\xi) = f(x_0) + f'(x_0)(\xi - x_0) + \frac{f''(x^*)}{2!}(\xi - x_0)^2 \;. \tag{4}$$

Nach (3) ist $f(x_0) - f'(x_0)\,x_0 = -f'(x_0)\,x_1$; somit folgt aus (4):

$$x_1 - \xi = \frac{f''(x^*)}{2f'(x_0)}\,(x_0 - \xi)^2 \;. \tag{5}$$

Hier ist die rechte Seite > 0 wegen (2). Es gilt daher $\xi < x_1 < x_0$, folglich treffen die Bedingungen (1) und (2) auch für $n = 1$ zu. Mit vollständiger Induktion ergibt sich: Die Folge $x_.$ ist monoton fallend und nach unten beschränkt durch ξ. Sie besitzt daher einen Grenzwert $\xi_1 \in [\,\xi, x_0\,]$. Wir führen nun in der Rekursionsformel

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad (n \geq 0)$$

den Grenzübergang $n \to \infty$ durch und erhalten

$$\xi_1 = \xi_1 - \frac{f(\xi_1)}{f'(\xi_1)} \;.$$

Es folgt $f(\xi_1) = 0$, das heisst: $\xi_1 = \xi$, wie erwartet.

Das Newtonsche Verfahren liefert also in der Tat die gesuchte Nullstelle ξ. Man kann aber noch mehr sagen: Ist n hinreichend gross, so ist x_n schon recht nahe bei ξ und ebenso der in (5) auftretende Punkt $x^* \in\,]\,\xi, x_n\,[$. Wir dürfen daher schreiben

$$x_{n+1} - \xi \doteq C(x_n - \xi)^2 \qquad (n \text{ hinreichend gross})$$

mit

$$C := \frac{f''(\xi)}{2f'(\xi)} .$$

Dies ist folgendermassen zu interpretieren: Das Newtonsche Verfahren **konvergiert quadratisch**, da sich die Abweichung $x_n - \xi$ mit jedem Schritt im wesentlichen quadriert. Anders ausgedrückt: Ist C von der Grössenordnung 1, so wird (nach einer gewissen Anfangsphase) die Anzahl der richtigen Dezimalstellen in der Näherung $x_n \doteq \xi$ mit jedem Schritt verdoppelt.

① Wir haben im Beweis von Satz **(4.3)** bzw. im anschliessenden Beispiel 4.1.② durch heuristische Betrachtungen die folgende Rekursionsformel zur Berechnung von $\sqrt{c}$ gefunden:

$$x_{n+1} := \frac{1}{2}\left(x_n + \frac{c}{x_n}\right) .$$

Wenn man auf die Gleichung

$$f(x) := x^2 - c = 0$$

das Newtonsche Verfahren anwendet, so kommt man auf dieselbe Formel; es ist nämlich

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

Damit wird nachträglich klar, warum diese Methode des "Wurzelziehens" so effizient ist. ○

② Von einem kreisförmigen Kuchen (Fig. 7.7.5) soll durch einen geradlinigen Schnitt ein Drittel abgeteilt werden. In welchem Abstand vom Mittelpunkt ist der Schnitt zu führen?

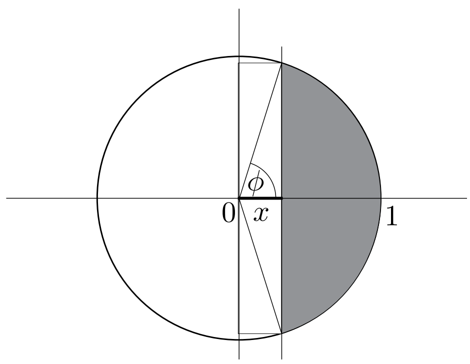

Fig. 7.7.5

Verwenden wir als Hilfsvariable den halben Zentriwinkel ϕ, so ist die Fläche F des abgeteilten Stückes gegeben durch

$$F = \frac{1}{2}\,2\,\phi - \frac{1}{2}\sin(2\phi)\ ,$$

so dass wir die Gleichung

$$f(\phi)\ :=\ \phi - \frac{1}{2}\sin(2\phi) - \frac{\pi}{3} = 0$$

auflösen müssen. Es ist $f(0) < 0$ und $f\left(\frac{\pi}{2}\right) > 0$. Wir leiten ab und erhalten

$$f'(\phi) = 1 - \cos(2\phi)\ , \qquad f''(\phi) = 2\sin(2\phi)\ .$$

Die Rekursionsformel lautet daher

$$\phi_{n+1}\ =\ \phi_n - \frac{\phi_n - \frac{1}{2}\sin(2\phi_n) - \frac{\pi}{3}}{1 - \cos(2\phi_n)}\ .$$

Da f' und f'' im Intervall $]\,0, \frac{\pi}{2}\,[$ gleiches Vorzeichen besitzen, wählen wir mit Vorteil $\phi_0 := \frac{\pi}{2}$ und erhalten

$$\phi_1\ =\ \frac{\pi}{2} - \frac{\frac{\pi}{2} - 0 - \frac{\pi}{3}}{1 - (-1)} = \frac{5\pi}{12}\ .$$

Die weiteren Werte sind:

$\phi_0 = 1.57079$
$\phi_1 = 1.30899$
$\phi_2 = 1.3026736661$
$\phi_3 = 1.3026628373\ldots$
$\phi_4 = 1.3026628373\ldots$

Der gesuchte Abstand x beträgt daher $x = \cos\phi \doteq 0.2649$. ○

Aufgaben

1. Die Gleichung $z^2 - 3z + 27 = 0$ besitzt eine Lösung in der Nähe von $z_0 := 1 + 5i$. Man führe zwei Newtonschritte durch zur Bestimmung eines besseren Näherungswerts z_2.
2. Ausgehend vom Näherungswert $x_0 := 1$ bestimme man mit Hilfe des Newtonschen Verfahrens die Nullstelle der Funktion $f(x) := x^2$. Man berechne explizit den n-ten Näherungswert x_n. Wie gut ist die Konvergenz, und warum ist sie nicht besser?

3. Ein divisionsfreier Algorithmus zur Berechnung von Kehrwerten: Fasse den Kehrwert einer Zahl $c \in \,]\,0,2\,[$ als Lösung der Gleichung

$$\frac{1}{x} - c = 0$$

auf und wende das Newtonsche Verfahren an. Es resultiert eine einfache Rekursionsformel, die mit $x_0 := 1$ beginnend eine schnell gegen $1/c$ konvergente Folge $x_.$ produziert. Für eine Fehlerabschätzung betrachte man die weitere Folge $y_n := 1 - cx_n$.