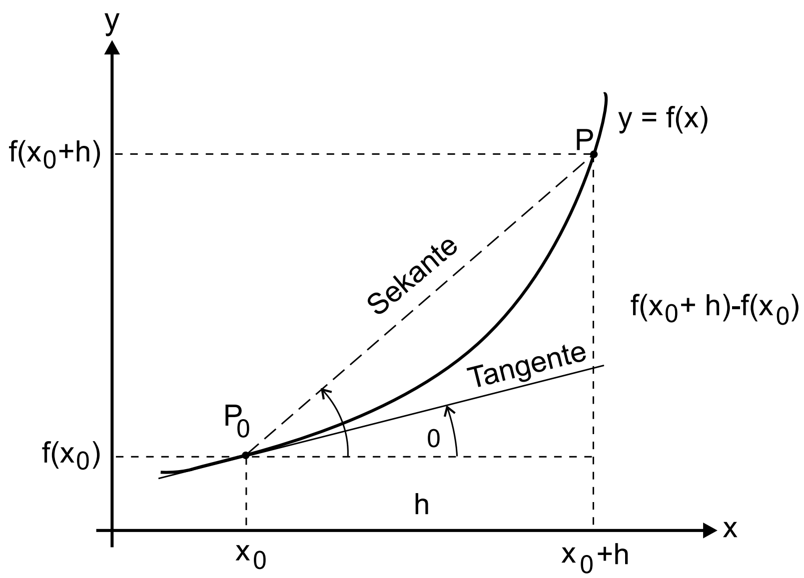
Kapitel 6

Differentialrechnung

- Der Ableitungsbegriff
- Ableitung elementarer Funktionen und höhere Ableitungen
- Ableitungstechniken
- Extrema und Kurvendiskussion
- Numerische Lösung nichtlinearer Gleichungen
- Taylorpolynome
- Funktionen in mehreren Veränderlichen
- Anwendungen

Geometrische Bedeutung und Definition der Ableitung

Gegeben sei jetzt eine Funktion $y = f(x)$.



Die Steigung der Sekante ergibt sich offensichtlich aus dem *Differenzenquotienten*

$$\tan \alpha = \frac{f(x_0 + h) - f(x_0)}{h}.$$

Die *Steigung der Tangente* ergibt sich aus dem *Differentialquotienten*

$$\tan \alpha_0 = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Definition der Ableitung

Definition

Falls der folgende Grenzwert (Differentialquotient) existiert, heißt

$$f'(x_0) := \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

Ableitung von f im Punkt x_0. f heißt dann in x_0 differenzierbar.

Man kann nun jedem Punkt x aus dem Definitionsbereich von $f(x)$ den Wert $f'(x)$ (falls existent!) zuordnen. Durch diese Zuordnungsvorschrift erhält man wieder eine Funktion:

Definition

Die durch die Zuordnung $x \longmapsto f'(x)$ erklärte Funktion heißt (erste) Ableitungsfunktion bzw. kurz (erste) Ableitung.

Äquivalenzen

Völlig äquivalente Schreibweisen für die Ableitung $f'(x_0)$:

$$y'(x_0), \qquad y'|_{x=x_0}, \qquad \left.\frac{dy}{dx}\right|_{x=x_0}, \qquad \left.\frac{df}{dx}\right|_{x=x_0}.$$

Die Größen dy bzw. df und dx nennt man auch *Differentiale*. Dies erklärt den Namen Differentialquotient für $f'(x_0)$.

Manchmal ist es praktischer, den Differentialquotienten in einer anderen Form zu benutzten. Hierzu setzt man $x = x_0 + h$. Dann ist $h = x - x_0$ und $h \to 0$ äquivalent zu $x \to x_0$:

Die Ableitung lässt sich auch durch folgenden Differentialquotienten berechnen:

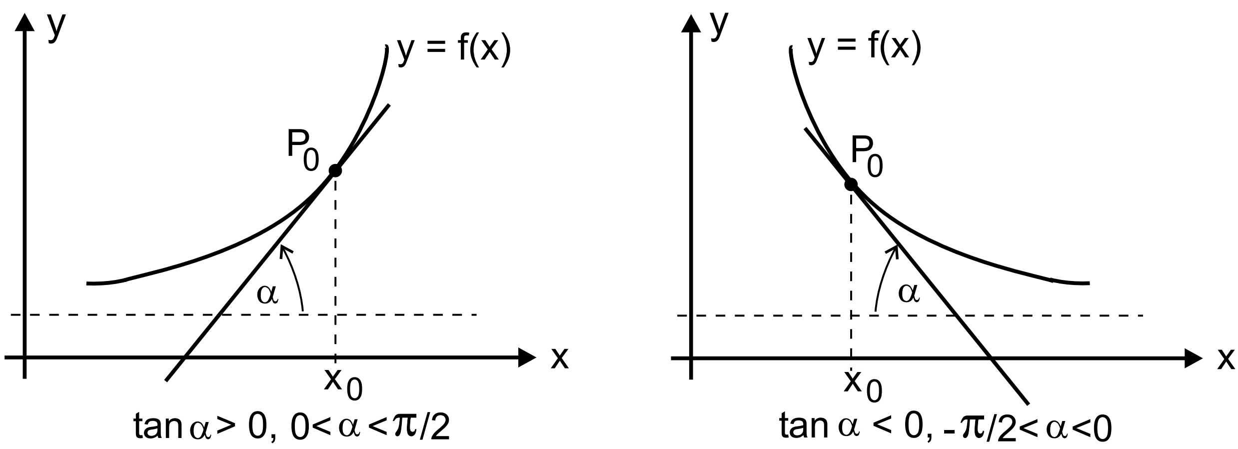
$$\boldsymbol{f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.}$$

Tangente, Monotonieverhalten

Geometrische Deutung:

$f'(x_0)$ ist der Tangens des Steigungswinkels α der Tangente an die Funktion $f(x)$ im Punkt x_0:

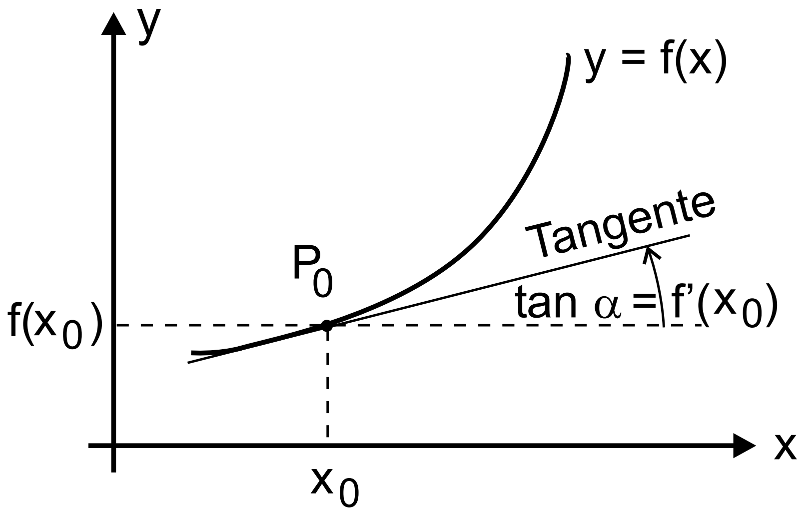
$$f'(x_0) = \tan \alpha.$$



Monotonieverhalten:

Gilt für alle x aus einem Intervall I $f'(x) > 0$ bzw. $f'(x) < 0$, so ist f(x) in I streng monoton wachsend bzw. fallend.

Tangentengleichung



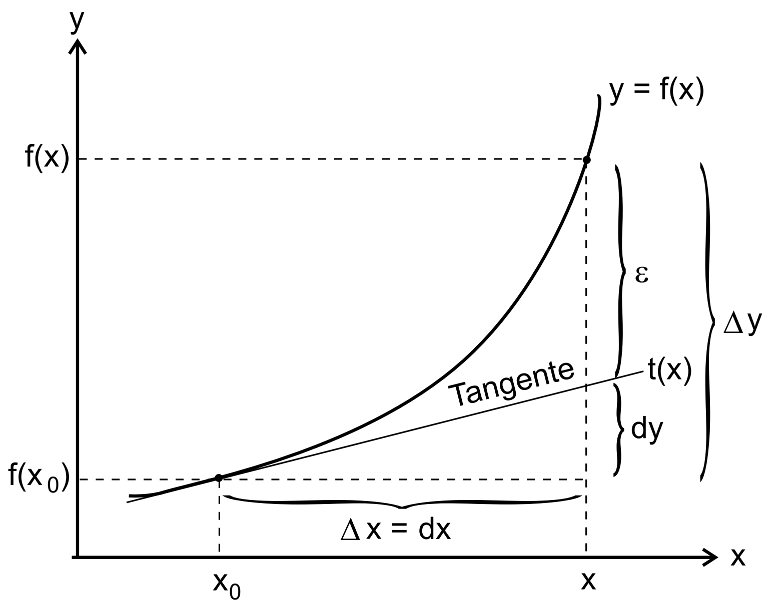
Die Tangente geht durch den *Punkt* $P_0 = (x_0, f(x_0))$ und hat dort definitionsgemäß die *Steigung* $f'(x_0)$. Damit lautet die *Tangentengleichung* offensichtlich:

$$\boldsymbol{t(x) = f(x_0) + f'(x_0)(x - x_0).}$$

Da eine Gerade durch einen Punkt und ihre Steigung in diesem Punkt eindeutig bestimmt ist, verifiziert sich obige Formel wegen $t(x_0) = f(x_0)$ und $t'(x_0) = f'(x_0)$ sofort.

Lineare Approximation

Wir interessieren uns nun für die Differenz ε der Zuwächse zwischen der Funktion und der Tangente, wenn das Argument sich von x_0 auf x ändert.



- Argumentzuwachs: $dx = x - x_0$,
- Zuwachs längs Funktion: $\Delta y = f(x) - f(x_0)$,
- Zuwachs längs Tangente: $dy = t(x) - t(x_0) = f'(x_0)(x - x_0) = f'(x_0)\,dx$,
- Zuwachsdifferenz von f und Tangente t: $\varepsilon = \Delta y - dy$.

Lineare Approximation

Unter Beachtung der Äquivalenz von $x \to x_0$ und $dx \to 0$ lässt sich

$$\lim_{x \to x_0} \frac{f(x) - [\mathbf{f(x_0) + f'(x_0)(x - x_0)}]}{x - x_0} = 0$$

mit den eben eingeführten Abkürzungen auch schreiben als

$$\begin{aligned} 0 &= \lim_{x \to x_0} \frac{(f(x) - f(x_0)) - f'(x_0)(x - x_0)}{x - x_0} \\ &= \lim_{dx \to 0} \frac{\Delta y - dy}{dx} = \lim_{dx \to 0} \frac{\varepsilon}{dx}. \end{aligned}$$

Gemäß Abb. können wir nun aber feststellen, dass gilt

$$\begin{aligned} f(x) &= f(x_0) + dy + \varepsilon \\ &= f(x_0) + f'(x_0)dx + \varepsilon \\ &= f(x_0) + dx\left(f'(x_0) + \frac{\varepsilon}{dx}\right). \end{aligned}$$

Wegen $lim_{dx \to 0} \varepsilon/dx = 0$ kann man die Summe $f'(x_0) + \varepsilon/dx$ für *genügend kleine* $|dx|$ allein durch $f'(x_0)$ ersetzen, ohne eine gegebene Genauigkeitsschranke für $f(x)$ zu verletzen.

Differentiale

Näherungsformel:

$$f(x) \approx f(x_0) + f'(x_0) \cdot dx.$$

Wenn man die Näherungsformel anwendet, so ersetzt man die Funktion f *in der Nähe von* x_0 durch ihre Tangente, also durch eine lineare Funktion (Gerade). Man spricht daher auch von *lokaler linearer Approximation* der Funktion f.

Definition

Man nennt die Größen dx Differential der unabhängigen Variablen x bzw. $dy = f'(x_0) \cdot dx$ Differential der Funktion f an der Stelle x_0.

Beispiel

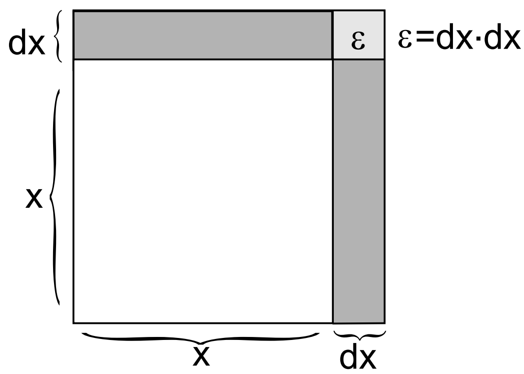
Gegeben sei ein Quadrat mit der Länge $x_0 = 5$ *und dem Flächeninhalt* $f(x_0) = 25$*. Jetzt vergrößern wir die Seiten des Quadrats um* $dx = 0.005$ *und berechnen den neuen Flächeninhalt mit der Näherungsformel: Wegen* $f(x) = x^2,\ f'(x) = 2x$ *lautet diese*

$$f(x_0 + dx) \approx x_0^2 + 2x_0 \cdot dx.$$

Wir setzen $x_0 = 5$ *und* $dx = 0.005$ *und erhalten*

$$f(5.005) \approx 5^2 + 10 \cdot 0.005 = 25.05\,.$$

Man beachte, dass wegen $(x_0+dx)^2 = x_0^2+2x_0 \cdot dx + (dx)^2$*, also* $\varepsilon = (dx)^2 = 0.000025$ *die beiden Ziffern nach dem Komma richtig sind. Der korrekte Flächeninhalt würde sich zu* 25.050025 *ergeben.*



Differentielle Fehleranalyse

Praxis: häufig Größen y_0 zu berechnen, die in einem funktionalen Zusammenhang mit gemessenen Ausgangsgrößen x_0 stehen: $y_0 = f(x_0)$.

Ist dabei x_0 mit einem Messfehler dx behaftet, so stellt sich die Frage, wie sich dieser auf die zu berechnende Größe y_0 auswirkt. Hierzu benutzt man die *differentielle Fehleranalyse*. Sie liefert eine Näherung für den auftretenden Berechnungsfehler $\Delta y = f(x_0+dx)-f(x_0)$ durch Anwendung unserer Näherungsformel. Dieser kann auf zwei Arten berechnet werden:

- *Absoluter Fehler:*

$$|\Delta y| \approx |dy| = |f'(x_0) \cdot dx|,$$

- *Relativer oder prozentualer Fehler:*

$$\left|\frac{\Delta y}{y}\right| \approx \left|\frac{dy}{y}\right| = \left|\frac{f'(x_0) \cdot dx}{f(x_0)}\right| = \left|\frac{f'(x_0)x_0}{f(x_0)}\right| \cdot \left|\frac{dx}{x_0}\right|.$$

Stetigkeit und Differenzierbarkeit

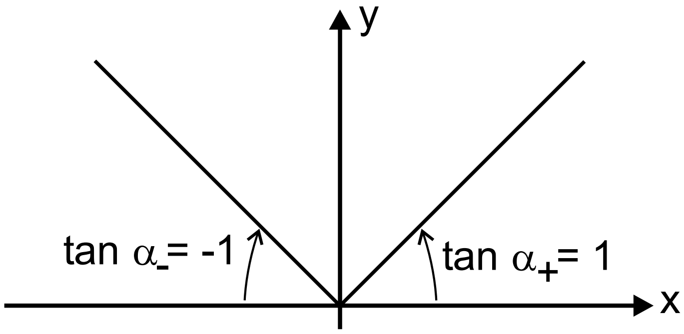
Ist $y = f(x)$ an der Stelle x_0 differenzierbar, so ist $y = f(x)$ an der Stelle x_0 auch stetig.

Betrachten wir nun die Betragsfunktion, so folgt, dass die Umkehrung nicht gilt. Eine stetige Funktion muss *nicht überall differenzierbar* sein: Für $f(x) = |x|$ gilt in $x_0 = 0$

$$\lim_{h \to 0+} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0+} \frac{h - 0}{h} = 1$$

und

$$\lim_{h \to 0-} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0+} \frac{-h - 0}{h} = -1.$$



Stetigkeit und Differenzierbarkeit

Links- und rechtsseitiger Grenzwert stimmen nicht überein, d.h. der Grenzwert existiert nicht. Also ist die in $x_0 = 0$ stetige Funktion $|x|$ dort nicht differenzierbar. Sie hat hier eine Knickstelle, da die Steigung $\alpha_- = -45°$ sofort in die Steigung $\alpha_+ = 45°$ übergeht. Anschaulich gesprochen gilt:

Hat der Graph einer Funktion keine Sprünge, so ist sie stetig; hat ihr Graph zudem keine Knickstellen, so ist sie differenzierbar.

Ableitung elementarer Funktionen

Funktion $y = f(x)$	Ableitung $y' = f'(x)$
$x^{\alpha},\ \alpha \in \mathbb{R}$	$\alpha x^{\alpha-1}$
e^x $\ln x$	e^x $\frac{1}{x}$
$\sin x$ $\cos x$ $\tan x$ $\cot x$	$\cos x$ $-\sin x$ $\frac{1}{\cos^2 x}$ $-\frac{1}{\sin^2 x}$
$\arcsin x$ $\arccos x$ $\arctan x$ $\operatorname{arccot} x$	$\frac{1}{\sqrt{1-x^2}}$ $-\frac{1}{\sqrt{1-x^2}}$ $\frac{1}{1+x^2}$ $-\frac{1}{1+x^2}$
$\lvert x\rvert,\ x \neq 0$	$\operatorname{sgn} x,\ x \neq 0$

Höhere Ableitungen

Wenn die Ableitung $y' = f'(x)$ einer differenzierbaren Funktion $y = f(x)$ ebenfalls wieder differenzierbar ist, so spricht man von der *zweiten Ableitung* $y'' = f''(x)$ von f. Analog werden Ableitungen höherer Ordnung definiert. Folgende Notationen sind üblich:

$$\begin{array}{lccccccccc}
\text{1. Ableitung:} & y' & = & f'(x) & = & \frac{dy}{dx} & = & \frac{df}{dx} \\
\text{2. Ableitung:} & y'' & = & f''(x) & = & \frac{d^2y}{dx^2} & = & \frac{d^2f}{dx^2} \\
\text{3. Ableitung:} & y''' & = & f'''(x) & = & \frac{d^3y}{dx^3} & = & \frac{d^3f}{dx^3} \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots \\
\text{n. Ableitung:} & y^{(n)} & = & f^{(n)}(x) & = & \frac{d^ny}{dx^n} & = & \frac{d^nf}{dx^n}
\end{array}$$

Ableitungsregeln

Unter der Voraussetzung, dass die Funktionen $f(x)$ und $g(x)$ differenzierbar sind, gilt:

$$y = cf(x) \quad \Longrightarrow \quad y' = cf'(x),\ c = const,$$

$$y = f(x) \pm g(x) \quad \Longrightarrow \quad y' = f'(x) \pm g'(x),$$

$$y = f(x) \cdot g(x) \quad \Longrightarrow \quad y' = f'(x) \cdot g(x) + f(x) \cdot g'(x),$$

$$y = \frac{f(x)}{g(x)} \quad \Longrightarrow \quad y' = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$$

Summen- und Differenzenregel

Beispiel

a) *Die Ableitung des Polynoms*

$$p(x) = 5x^8 - 3x^3 + 4x - 2$$

ergibt sich zu

$$\begin{aligned} p'(x) &= (5x^8)' - (3x^3)' + (4x)' - (2)' \\ &= 40x^7 - 9x^2 + 4 \end{aligned}$$

b) *Die Ableitung von*

$$y(x) = 4x^7 + 3e^x + \ln x,\ x > 0,$$

erhält man durch Ableiten einzelner Terme zu

$$y'(x) = 28x^6 + 3e^x + 1/x.$$

Produktregel

Beispiel

a) *Gegeben sei die Funktion*

$$y(x) = x^3 \cdot \ln x,\ x > 0.$$

Die Ableitungen von $f(x) = x^3$ *bzw.* $g(x) = \ln x$ *sind* $3x^2$ *bzw.* $1/x$. *Die beiden neuen Terme ergeben sich also zu* $(x^3)' \cdot \ln x = 3x^2 \cdot \ln x$ *und* $x^3 \cdot (\ln x)' = x^3 \cdot 1/x$. *Somit lautet die Ableitung von* $y(x)$ *nun*

$$y'(x) = 3x^2 \cdot \ln x + x^3 \cdot \frac{1}{x} = x^2(3 \ln x + 1).$$

b) *Durch wiederholte Anwendung der Produktregel erhält man zum Beispiel für die Ableitung eines Produktes von drei Funktionen* $y = f \cdot g \cdot h$:

$$\begin{aligned} y' &= [f \cdot (g \cdot h)]' \\ &= f' \cdot (g \cdot h) + f \cdot (g \cdot h)' \\ &= f' \cdot g \cdot h + f \cdot (g' \cdot h + g \cdot h') \\ &= f' \cdot g \cdot h + f \cdot g' \cdot h + f \cdot g \cdot h' \end{aligned}$$

Quotientenregel

Die Ableitung eines „Bruches“ Z/N ist ebenfalls ein Bruch.
Der Zähler des „Ableitungsbruches“ ergibt sich folgendermaßen: Ableitung von Z mal N minus Z mal Ableitung von N.
Der neue Nenner ist einfach der alte Nenner zum Quadrat.

Beispiel

a) *Wir betrachten die Funktion*

$$y(x) = \mathrm{e}^x/x,\ x \neq 0.$$

Es ist $Z = \mathrm{e}^x$, $N = x$ *und damit* $Z' = \mathrm{e}^x$, $N' = 1$. *Daher gilt* Z' *mal* N *ist* $\mathrm{e}^x \cdot x$ *und* Z *mal* N' *ist* $\mathrm{e}^x \cdot 1$. *Die Differenz von erstem minus zweitem Term ergibt den Zähler des „Ableitungsbruches“. Der Nenner ergibt sich zu* $N^2 = x^2$. *Es gilt also*

$$\begin{aligned} y'(x) &= \frac{\mathrm{e}^x \cdot x - \mathrm{e}^x \cdot 1}{x^2} \\ &= \mathrm{e}^x \left(\frac{x-1}{x^2} \right). \end{aligned}$$

Kettenregel

Funktionen werden häufig auch durch „Hintereinanderschalten“ zweier Funktionen, einer *inneren Funktion* $u = g(x)$ und einer *äußeren Funktion* $y = f(u)$, definiert: $y = f(g(x))$. Für diesen Fall gibt es eine weitere Ableitungsregel:

Sind $y = f(u)$ und $u = g(x)$ differenzierbar, dann ist auch $y(x) = f(g(x))$ differenzierbar und es gilt

$$y'(x) = f'(g(x)) \cdot g'(x) \text{ bzw. } \frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}.$$

Die Bildung des Faktors $g'(x) = du/dx$ bezeichnet man als „Nachdifferenzieren“; manchmal wird die Kettenregel auch als „äußere Ableitung (Ableitung der äußeren Funktion $f(u)$) mal innere Ableitung (Ableitung der inneren Funktion $g(x)$)“ formuliert.

Beispiel

Gegeben sei $y(x) = \sqrt[5]{x^3 + 7x + 5}$. *Wir setzen*

$$f(u) = \sqrt[5]{u} = u^{1/5}$$

und

$$u = g(x) = x^3 + 7x + 5.$$

Dann gilt

$$f'(u) = \frac{1}{5} u^{\frac{1}{5}-1} = \frac{1}{5} u^{-\frac{4}{5}} = \frac{1}{5\sqrt[5]{u^4}}$$

und

$$g'(x) = 3x^2 + 7,$$

woraus nach Kettenregel folgt:

$$y'(x) = \frac{1}{5\sqrt[5]{(x^3 + 7x + 5)^4}} \cdot (3x^2 + 7).$$

Ableitung der Umkehrfunktion

Es sei f^{-1} die Umkehrfunktion der umkehrbaren Funktion f. Dann gilt für deren Ableitung:

$$\boldsymbol{(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))},\ f'(f^{-1}(x)) \neq 0.}$$

Beispiel

Wir betrachten die Umkehrfunktion $f^{-1}(x) = \ln x$ von $f(y) = \mathrm{e}^y$. Es ist $f'(y) = \mathrm{e}^y$. Damit ergibt sich die Ableitung von $\ln x$ nach obiger Formel zu

$$(\ln x)' = \frac{1}{f'(\ln x)} = \frac{1}{\mathrm{e}^{\ln x}} = \frac{1}{x}.$$

Grenzwertregeln von L'Hospital

Verfahren, das die Berechnung von Grenzwerten der Form

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} \quad \text{und} \quad \lim_{x\to \pm\infty} \frac{f(x)}{g(x)},$$

ermöglicht, bei denen Zähler und Nenner gemeinsam entweder den Grenzwert 0 annehmen oder gegen $\pm\infty$ streben (hierfür schreiben wir symbolisch „$0/0$" bzw. „∞/∞").

Wenn der Grenzwert $\lim_{x\to x_0} f(x)/g(x)$ die Form „$0/0$" oder „∞/∞" hat, aber andererseits der Grenzwert $\lim_{x\to x_0} f'(x)/g'(x)$ existiert, dann gilt

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}.$$

Diese Formel kann man anstelle von $x \to x_0$ auch für $x \to \infty$ oder $x \to -\infty$ anwenden.

Beispiel

Wir betrachten

$$\lim_{x \to 0} \frac{\sin x}{x}.$$

Dieser Grenzwert besitzt die Form „$0/0$“. Mit

$$f(x) = \sin x,\ g(x) = x$$

und den Ableitungen

$$(\sin x)' = \cos x,\ (x)' = 1$$

ergibt sich

$$\lim_{x \to x_0} \frac{f'(x)}{g'(x)} = \lim_{x \to 0} \frac{\cos x}{1} = \frac{1}{1} = 1.$$

Es folgt daher

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

Man beachte aber, dass die Regel von L'Hospital nicht immer (auch nicht durch wiederholte Anwendung) zu einem Ergebnis führt, selbst wenn der zu untersuchende Grenzwert existiert. In diesem Fall müssen andere Grenzwertregeln benutzt werden.

Extremwerte

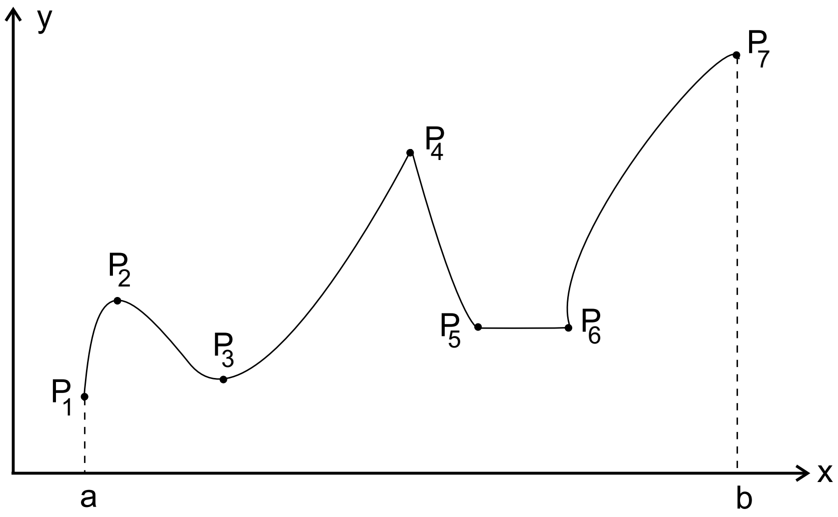
Definition

Eine Funktion $y = f(x)$ mit dem Definitionsbereich D hat an der Stelle $x_0 \in D$ ein globales Maximum bzw. globales Minimum, wenn für alle $x \in D$ gilt

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

Ist eine der beiden obigen Ungleichungen nur für Argumente x aus einer Umgebung von x_0 erfüllt, dann spricht man von einem relativen Maximum bzw. relativen Minimum. Ein relatives Extremum bezeichnet man auch als lokales Extremum (lokales Maximum bzw. lokales Minimum).
Gilt in den Ungleichungen für kein $x \neq x_0$ Gleichheit, so heißen die Extrema streng.

Potentielle Extremwertstellen



Strenge relative Extrema liegen in den Punkten P_1, P_3 (Minima), P_2, P_4 und P_7 (Maxima) vor, während alle Punkte zwischen P_5 und P_6 nicht strenge relative Minima sind. In P_1 und P_7 treten zudem globale Extrema auf. Unmittelbar aus der Abbildung ersichtlich ist:

Mögliche Kanditaten für relative Extrema sind

- **Randpunkte des Definitionsbereichs D,**
- **„Nichtdifferenzierbarkeits"-Stellen,**
- **Stellen mit horizontaler Tangente.**

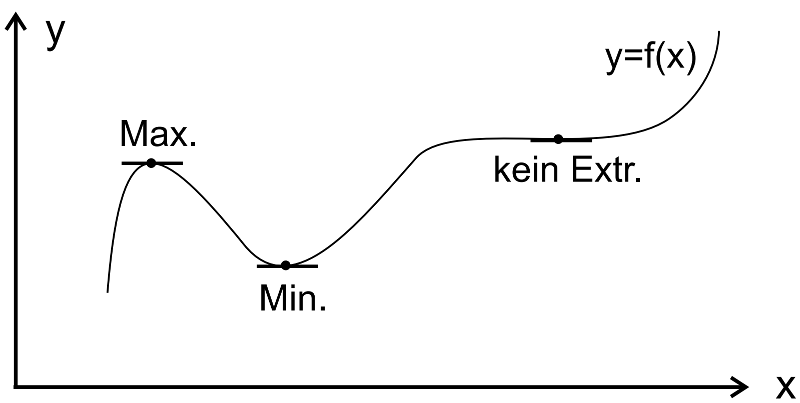
Notwendiges Extremwertkriterium

Bei Extrema in Differenzierbarkeitsstellen müssen die Tangenten *notwendigerweise* immer horizontal verlaufen. Es lässt sich daher festhalten:

Wenn die Funktion $f(x)$ in x_0 differenzierbar und x_0 ein Extremwert von $f(x)$ ist, dann muss gelten:

$$\boldsymbol{f'(x_0) = 0.}$$

Diese Bedingung ist jedoch nur *notwendig*, aber *nicht hinreichend*. So hat beispielsweise die Funktion in der Abb. drei horizontale Tangenten, aber nur zwei Extrema.



Hinreichendes Kriterium

Wenn für die in x_0 differenzierbare Funktion

$$f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0$$

und $f^{(n)}(x_0) \neq 0$ für ein gerades $n \in \mathbb{N}$ gilt, dann hat $f(x)$ in x_0 ein strenges relatives Extremum; genauer ein

- **strenges relatives Maximum, falls $f^{(n)}(x_0) < 0$,**
- **strenges relatives Minimum, falls $f^{(n)}(x_0) > 0$.**

Häufig angewandt wird der - aus dem Schulunterricht bekannte - Fall $n = 2$, d.h.

$$\left.\begin{array}{l} f'(x_0) = 0 \\ f''(x_0) < 0 \end{array}\right\} \Longrightarrow \quad x_0 \text{ ist strg. rel. Maximum,}$$

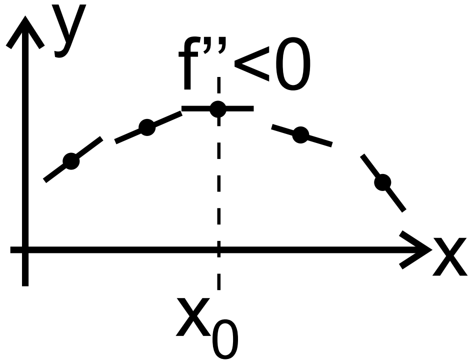
$$\left.\begin{array}{l} f'(x_0) = 0 \\ f''(x_0) > 0 \end{array}\right\} \Longrightarrow \quad x_0 \text{ ist strg. rel. Minimum.}$$

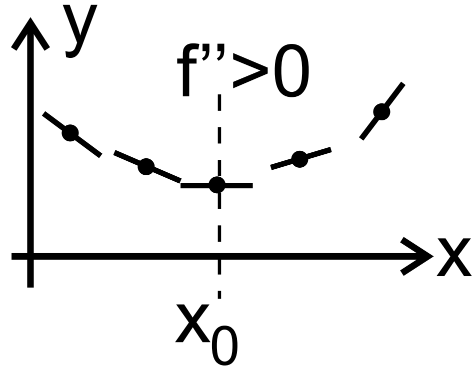
Hinreichendes Kriterium

Dieser Spezialfall ist auch geometrisch einsichtig: Die zweite Ableitung $f''(x)$ kann auch als erste Ableitung von $f'(x)$ interpretiert werden.

Ist $f''(x) < 0$, so fällt die erste Ableitung monoton, d.h. die Steigungen der Funktionstangenten durch verschiedene Punkte (von links nach rechts in einer Umgebung von x_0) werden immer kleiner.

Umgekehrt werden bei $f''(x) > 0$ die Tangentensteigungen immer größer. Aus der Abb. wird die Art des Extremums daher sofort deutlich.





Bestimmung globaler Extrema

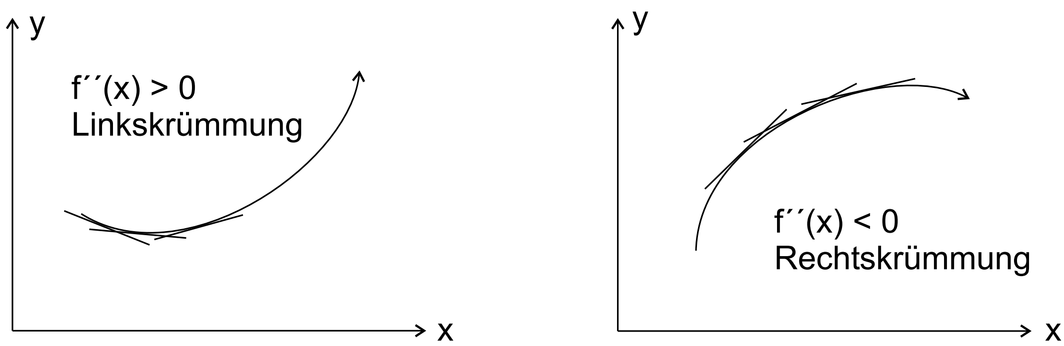
Hat man mit Hilfe obiger Kriterien alle relativen Extrema ermittelt, kann man jetzt auch die *Bestimmung der globalen Extrema* angehen:

Zur Bestimmung der globalen Extrema muss man

- **die zu relativen Extrema gehörigen Funktionswerte vergleichen,**
- **das Verhalten der Funktion für $x \to \pm\infty$ und in eventuell vorhandenen Unendlichkeitsstellen untersuchen.**

Wendepunkte und Krümmungsverhalten

Da die Steigung der Tangente einer Funktion $y = f(x)$ im Punkt x_0 durch $f'(x_0)$ gegeben ist und die Ableitung einer Funktion deren Wachstum charakterisiert, ist die zweite Ableitung $f''(x_0)$ offensichtlich ein Maß für die *Änderung der Tangentensteigung*.



Ist $f''(x) > 0$ in einem Intervall I, so nimmt $f'(x)$ dort streng monoton zu, d.h. die Tangentensteigung wächst ständig. Man spricht in diesem Fall von *Linkskrümmung* und nennt die Funktion *konvex*.

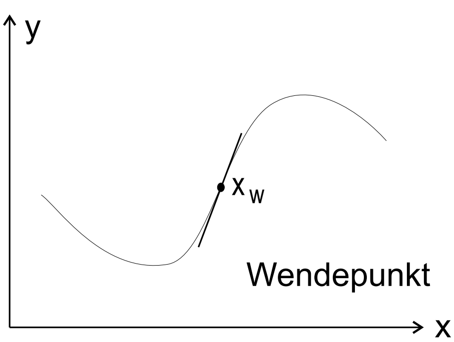
Analog erhält man für $f''(x) < 0$ ständig fallende Tangentensteigungen und damit eine *Rechtskrümmung (konkave Funktion)*.

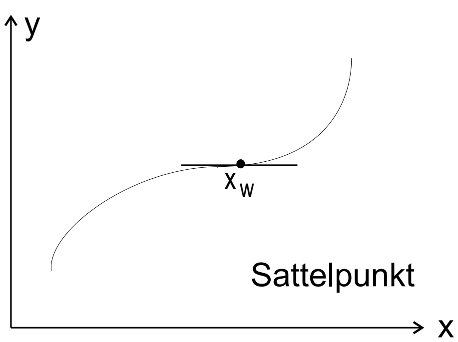
Wendepunkt, Sattelpunkt

Die Krümmung ändert sich nun offensichtlich in Punkten, in denen steigende Tangenten in fallende (oder umgekehrt) übergehen. Dies sind aber genau die Punkte, für die die erste Ableitung Extremwerte annimmt. Daher definiert man:

Definition

Ein Punkt x_w, an dem die erste Ableitung f' einer differenzierbaren Funktion f ein relatives Extremum besitzt ($f''(x_w) = 0$), heißt Wendepunkt von f. Ein Wendepunkt mit horizontaler Tangente (zusätzlich $f'(x_w) = 0$) heißt Sattelpunkt.





Hinreichendes Wendepunkt-Kriterium

Um Wende- und Sattelpunkte zu finden, sind die Extremwertkriterien nicht auf f, sondern auf die Funktion f' anzuwenden. Das liefert folgendes Kriterium:

Hinreichend dafür, dass in x_w ein Wendepunkt vorliegt, ist

$$\boldsymbol{f''(x_w) = 0 \quad \textbf{und} \quad f'''(x_w) \neq 0\,.}$$

Elementare Kurvendiskussion

Mit einer elementaren Kurvendiskussion kann man sich mit recht geringem Arbeitsaufwand einen Überblick über den Graphen einer gegebenen Funktion $y = f(x)$ verschaffen. Empfehlenswert ist dabei die Bestimmung/Ermittlung von

a) Polstellen, Asymptoten und Verhalten am Rand des Definitionsbereichs,

b) Symmetrie (Achsen oder Ursprung) und Periodizität,

c) Nullstellen der Funktion,

d) Stetigkeits- und Differenzierbarkeitsintervalle,

e) Extrema, Wendepunkte mit Krümmungsintervallen,

f) und letztlich die Skizzierung des Funktionsgraphen.

Newton'sches Iterationsverfahren

Newton-Verfahren setzt die Differenzierbarkeit der Funktion f voraus. Es kann daher zur Bestimmung der Nullstelle x^* anstelle von Sekanten die Tangenten an f benutzen.

Zunächst sucht man eine gute Näherung x_0 für x^*, die man beispielsweise aus einer Skizze oder Funktionstabelle gewinnen kann. Im Punkt $(x_0, f(x_0))$ stellt man nun die Tangentengleichung auf:

$$t(x) = f(x_0) + f'(x_0)(x - x_0).$$

Da die Tangente eine gute Näherung der gegebenen Funktion f darstellt, ist zu erwarten, dass die einfach zu berechnende Nullstelle x_1 von t eine bessere Näherung von x^* ist als x_0:

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

Nun setzen wir das Verfahren mit dem neuen Näherungswert x_1 fort. Dies liefert

$$x_2 = x_1 - f(x_1)/f'(x_1).$$

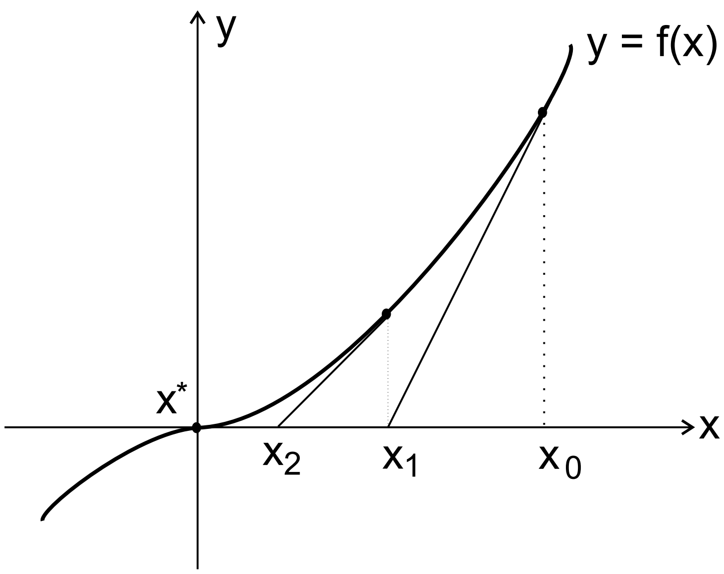
Newton'sches Iterationsverfahren

Wiederholte Anwendung (Iteration) - natürlich nur für $f'(x_k) \neq 0$ - liefert:

Wähle x_0 und berechne die Folge (x_k) mittels

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \ldots .$$

STOP, falls $|x_{k+1} - x_k| < \varepsilon$ (Fehlerschranke).



Beispiel

Gegeben sei nochmals das Polynoms

$$p(x) = 3x^3 - 4x^2 - 2x + 2.$$

Man wählt nun einen Startpunkt x_0 hinreichend nahe bei x_1^*. Für $x_0 = 0.01$ ergeben sich die Iterierten aus folgender Tab.:

k	x_k
0	0.0100000000
1	0.9621441970
2	0.2627451476
3	$0.\underline{6}226741590$
4	$0.\underline{602}0648709$
5	$0.\underline{6022491}765$
6	$0.\underline{6022491900}$

Bei jeder Iterierten sind die gültigen Ziffern unterstrichen. Man erkennt daran deutlich die quadratische Konvergenz. Bereits nach 6 Iterationen sind 10 Nachkommastellen korrekt.

Nimmt man einen „besseren“ Startwert, z. B. $x_0 = 0.1$, so erreicht man 10 gültige Nachkomastellen bereits nach 5 Iterationen.

Taylorpolynome

Es wurde bereits gezeigt, dass man eine Funktion $f(x)$ in der Nähe eines Punktes x_0 *linear approximieren* kann (man beachte, dass $dx = x - x_0$ gilt):

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) =: p_1(x).$$

Beim Gebrauch dieser Formel verwendet man die leicht berechenbare Ersatzfunktion $p_1(x)$, die ein Polynom 1. Grades ist. $p_1(x)$ hat im Punkt x_0 denselben Funktions- und Ableitungswert wie f:

$$p_1(x_0) = f(x_0) \quad \text{und} \quad p_1'(x_0) = f'(x_0).$$

Kennt man auch die zweite Ableitung von f, so sucht man eine „bessere“ Ersatzfunktion $p_2(x)$ mit

$$p_2(x_0) = f(x_0),$$

$$p_2'(x_0) = f'(x_0) \quad \text{und} \quad p_2''(x_0) = f''(x_0).$$

Man verifiziert leicht, dass das Polynom zweiten Grades

$$p_2(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2$$

die geforderten Eigenschaften erfüllt.

Taylorpolynom bzw. Taylorentwicklung

Kennt man nun die Ableitungen von f bis zur n-ten Ordnung, so sucht man ein Polynom $p_n(x)$ n-ten Grades mit der Eigenschaft

$$p_n(x_0) = f(x_0) \quad \text{und} \quad p_n^{(k)}(x_0) = f^{(k)}(x_0)$$

für $k = 1, \ldots n$.

Definition

Das Taylorpolynom n-ten Grades von f in x_0 (auch Taylorentwicklung von f um x_0 bis zum Grad n genannt) lautet (mit $f^{(0)}(x_0) := f(x_0)$):

$$\begin{aligned} p_n(x) &:= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \\ &= f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \\ &\quad + \frac{f''(x_0)}{2!}(x - x_0)^2 \\ &\quad + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n. \end{aligned}$$

Taylor'sche Formel

Die Näherung ist in der Praxis aber nur dann verwertbar, wenn wir Abschätzungen für den Approximationsfehler $R_n(x) = f(x) - p_n(x)$ angeben können. Hierfür gibt es die Taylor'sche Formel.

Ist die Funktion $f(x)$ in einer Umgebung des Punktes x_0 genügend oft differenzierbar, dann gilt in dieser Umgebung

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + R_n(x)$$

mit dem so genannten Lagrange'schen Restglied (Approximationsfehler)

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}.$$

Dabei ist ξ eine Zahl zwischen x_0 und x ($x_0 < \xi < x$ bzw. $x < \xi < x_0$), die i.Allg. nicht bekannt ist.

Beispiel

Gegeben sei die Funktion $f(x) = \mathrm{e}^x$. Wegen

$$f^{(k)}(x) = \mathrm{e}^x \text{ und damit } f^{(k)}(0) = 1,$$

für alle $k = 0, 1, 2, \ldots$, ergibt sich ihr Taylorpolynom n-ten Grades um den Punkt $x_0 = 0$ zu

$$\sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}(x-0)^k = \sum_{k=0}^{n} \frac{x^k}{k!}\,.$$

Auch das Lagrange'sche Restglied wird sehr einfach ($0 < \xi < x$ bzw. $x < \xi < 0$):

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-0)^{n+1} = \frac{\mathrm{e}^\xi}{(n+1)!}x^{n+1}.$$

Approximation der Zahl e (Taylorpolynom bei $x = 1$ auswerten):

$$\mathrm{e} = \mathrm{e}^1 \approx 1^0/0! + 1^1/1! + 1^2/2! + \ldots$$

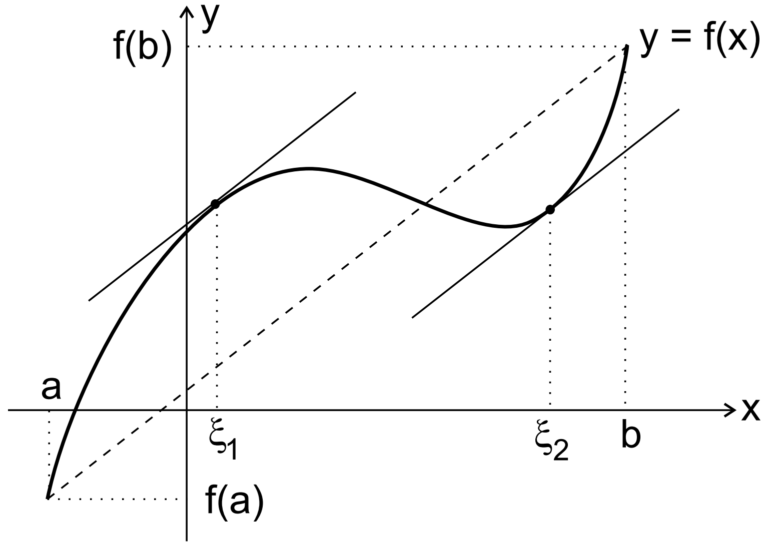
Mittelwertsatz der Differentialrechnung

Spezialfall der Taylorformel: $n = 0$ ($x = b$ und $x_0 = a$)

Sei $f(x)$ stetig auf $[a, b]$ und differenzierbar auf (a, b). Dann gibt es ein $\xi \in (a, b)$ mit

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

Geometrisch besagt dieser Satz, dass es wenigstens eine Tangente an $f(x)$ gibt, die parallel zur Geraden durch die Punkte $f(a)$ und $f(b)$ liegt.



Definition

Eine reellwertige Funktion f mit n unabhängigen Variablen ist eine Zuordnung, die jedem n-Tupel $(x_1, x_2, \ldots, x_n) \in D$ genau eine reelle Zahl z zuordnet:

$$(x_1, x_2, \ldots, x_n) \longmapsto z = f(x_1, x_2, \ldots, x_n)$$

Dabei heißen D Definitionsbereich,

$$x_1, x_2, \ldots, x_n$$

Argumente (bzw. unabhängige Veränderliche oder unabhängige Variable).
z ist die abhängige Veränderliche oder abhängige Variable, $f(x_1, x_2, \ldots, x_n)$ der Funktionswert an der Stelle $(x_1, x_2, \ldots x_n)$.

Die Menge aller Bilder heißt Wertebereich W.

Dimensionen

Die Menge aller n-Tupel $(x_1, x_2, \ldots, x_n)$ mit $x_i \in \mathbb{R}$ für $i = 1, \ldots, n$ wird mit $\mathbb{R}^n$ bezeichnet.
$\mathbb{R}^1$ ist also die Zahlengerade (d.h. die Menge aller reellen Zahlen $\mathbb{R}$), $\mathbb{R}^2$ ist die zweidimensionale Ebene und $\mathbb{R}^3$ der dreidimensionale Raum.

Beispiel

a) *Eine Ebene im $\mathbb{R}^3$ kann durch die Gleichung $ax + by + cz - d = 0$ dargestellt werden. Falls $c \neq 0$, ist diese nach z auflösbar:*

$$z = (-ax - by + d)/c = f(x, y).$$

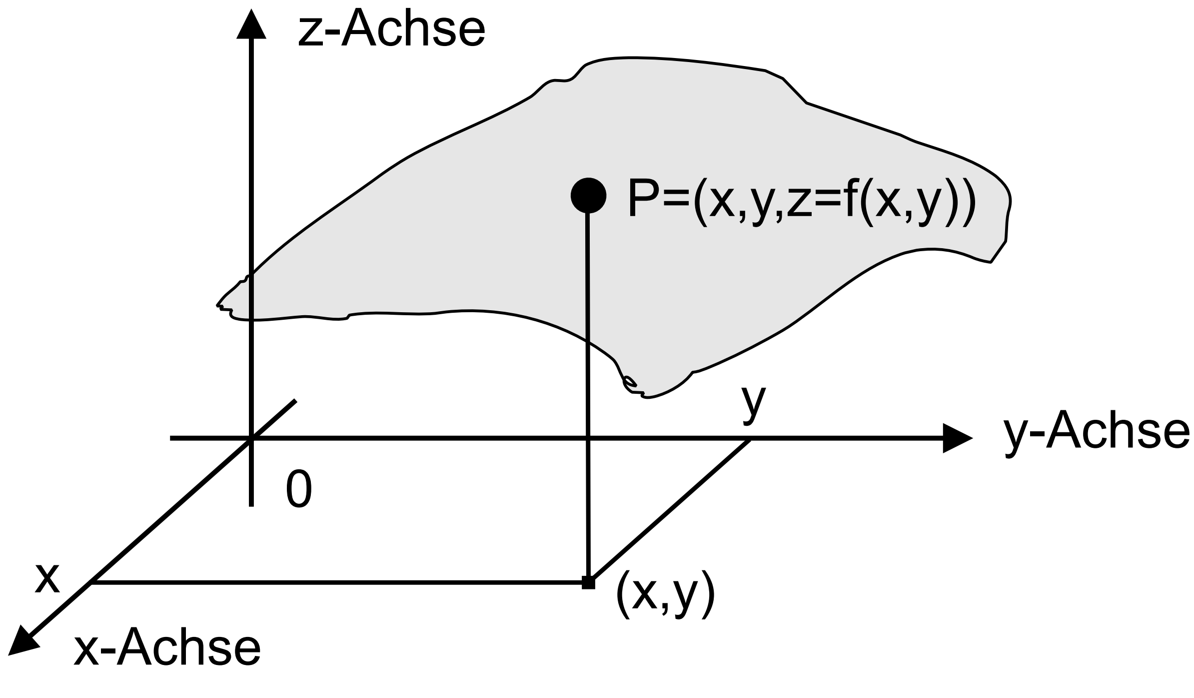
Eine Ebene ist also auch durch eine Funktion mit zwei unabhängigen Variablen x, y darstellbar. Für diese gilt $D = \mathbb{R}^2$ und $W = \mathbb{R}$, falls $a \neq 0$ oder $b \neq 0$.

Funktionsgraph

Funktionen zweier unabhängiger Variablen

$$z = f(x, y)$$

kann man – im Gegensatz zu höherdimensionalen Funktionen – geometrisch anschaulich darstellen, nämlich als Fläche im Raum:
Man interpretiert (x, y, z) mit $z = f(x, y)$ als kartesische Koordinaten eines Punktes im Raum ($\mathbb{R}^3$).
Die so erhaltene Fläche wird auch als *Funktionsgraph* bezeichnet.



Schnittkurven

Hilfsmittel zur Veranschaulichung der in obigem Sinn erzeugten „Funktionsfläche“ sind nun Schnitte mit speziellen Ebenen, die auf folgende Schnittkurven führen:

- Schnittkurve mit Ebene $x = x_0$:

$$z = f(x_0, y),\ x = x_0$$

- Schnittkurve mit Ebene $y = y_0$:

$$z = f(x, y_0),\ y = y_0$$

- Schnittkurve mit Ebene $z = c$:

$$f(x, y) = c,\ z = c$$

Letztere ist Kurvengleichung in impliziter Form. Sie wird auch als *Höhenlinie* oder *Niveaulinie* bezeichnet, da ihre senkrechte Projektion in die (x, y)-Ebene der geometrische Ort aller Punkte (x, y) mit gleichem Funktionswert c ist. Mathematisch ist die Niveaulinie charakterisiert durch die Menge

$$I_c = \{(x_1, x_2) \in D \,|\, f(x_1, x_2) = c\}.$$

Beispiel
Gegeben sei die Funktion $z = x^2 + y^2,\ D = \mathbb{R}^2$*.*
Es ergibt sich:

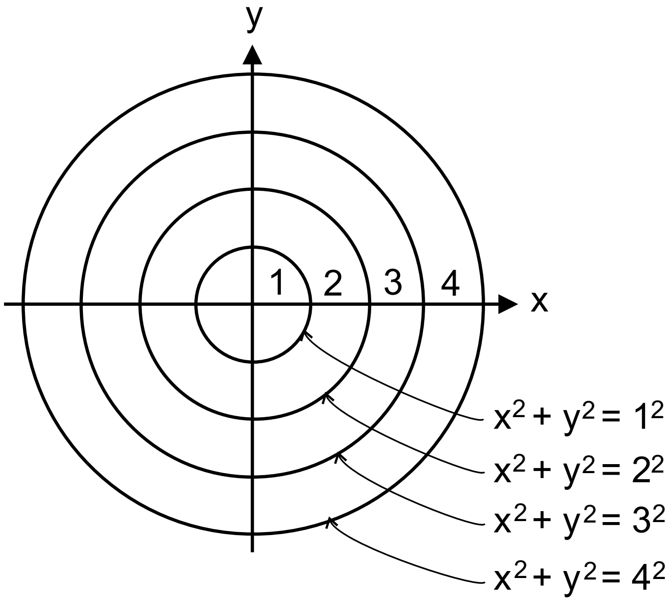
a) *Schnittkurve mit* (y, z)*-Ebene (*$x = 0$*):*
$z = y^2, x = 0$*: Parabel* $z = y^2$ *in der* (y, z)*-Ebene*

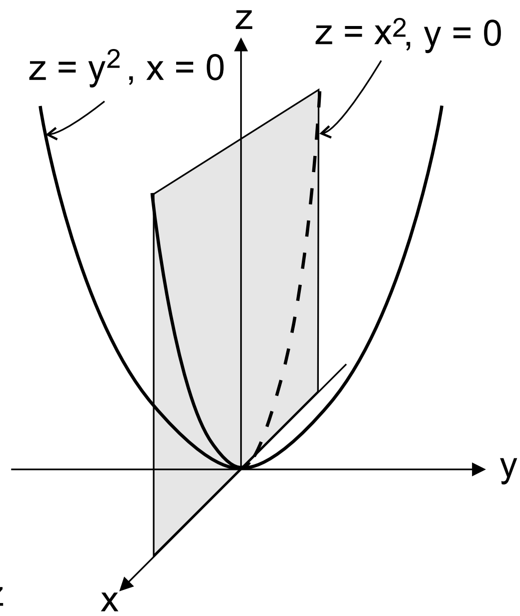
b) *Schnittkurve mit* (x, z)*-Ebene (*$y = 0$*):*
$z = x^2, y = 0$*: Parabel* $z = x^2$ *in der* (x, z)*-Ebene*

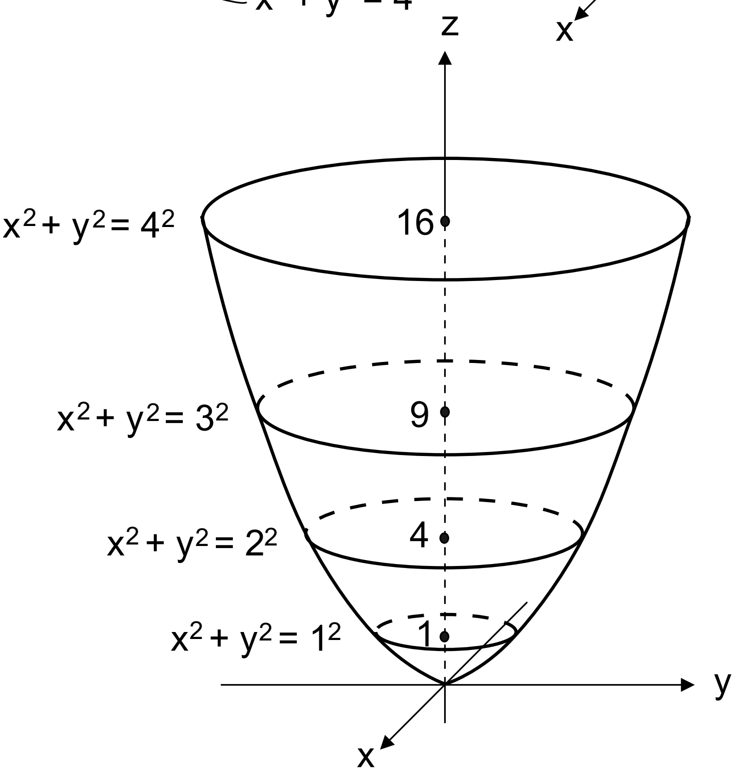
c) *Schnittkurve mit der Ebene* $z = c,\ c \geq 0$*:*
$x^2 + y^2 = c,\ z = c$*: Kreis in der Ebene* $z = c$ *mit Radius* $\sqrt{c}$ *und dem Mittelpunkt* $(0, 0)$ *(Niveaulinie).*
Insbesondere: $c = 0$*: Punkt* $(0, 0, 0)$

Als Fläche ergibt sich daher ein Rotationsparaboloid*.*

Beispiel – Fortsetzung







Grenzwertbegriff

Während bei einer Funktion $y = f(x)$ auf der Zahlengeraden nur zwei Möglichkeiten der Annäherung an einen Punkt $x = x_0$ existieren (nämlich von links $x \to x_0-$ bzw. von rechts $x \to x_0+$), gibt es in der Ebene natürlich *unendlich* viele Möglichkeiten der Annäherung.

Definition

Eine Funktion $z = f(x, y)$ hat an der Stelle (x_0, y_0) den Grenzwert G, falls sich bei beliebiger Annäherung (also auf allen möglichen Wegen) an den Punkt (x_0, y_0) stets der Grenzwert G ergibt. Man schreibt dann $G = \lim_{(x,y)\to(x_0,y_0)} f(x, y)$.

Beispiel

a) *Gegeben sind*

$$z = f(x, y) = x + y$$

und $(x_0, y_0) = (2, 5)$*. Dann gilt:*

$$\lim_{(x,y)\to(2,5)} x+y = \lim_{x\to 2} x + \lim_{y\to 5} y = 2+5 = 7.$$

b) *Zu betrachten ist*

$$f(x, y) = \sin(x^2 + y^2)/(x^2 + y^2)$$

im Punkt $(x_0, y_0) = (0, 0)$*: Setze*

$$v := x^2 + y^2,$$

dann gilt

$$\lim_{(x,y)\to(0,0)} v = 0.$$

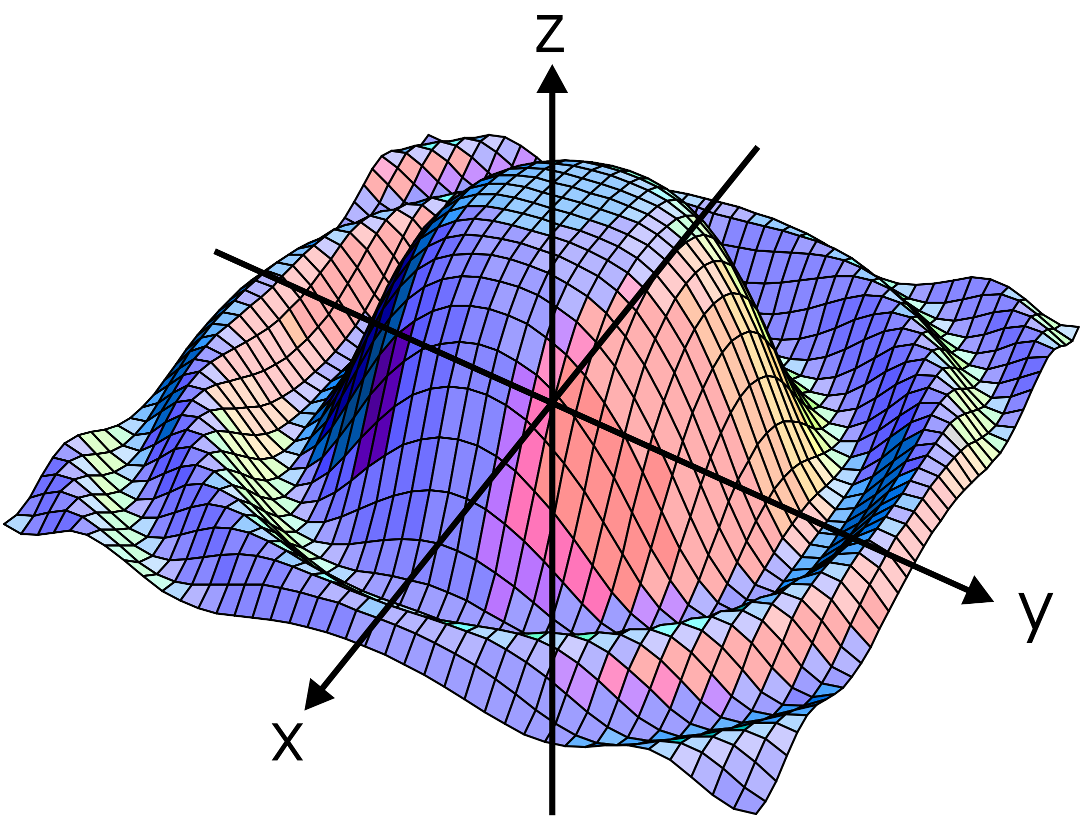
Also folgt für den Grenzwert (L'Hospital!):

$$\lim_{(x,y)\to(0,0)} \sin(x^2 + y^2)/(x^2 + y^2) =$$

$$= \lim_{v\to 0} \sin v/v = 1.$$

Beispiel – Fortsetzung

Mexikanerhut



Stetigkeit

Definition

Eine Funktion $z = f(x, y)$ heißt stetig an der Stelle (x_0, y_0), falls

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0).$$

Wie man am vorangegangenen Beispiel feststellen kann, hat der Graph einer stetigen Funktion („Mexikaner-Hut“) keine „Bruchstellen“.
Die Funktion aus dem Beispiel ist nämlich im Punkt $(0, 0)$ mit $f(0, 0) = 1$ stetig ergänzbar.
Die „Steilklippe“ ist im Punkt $(0, 0)$ nicht stetig ergänzbar.

Partielle Ableitungen 1.Ordnung

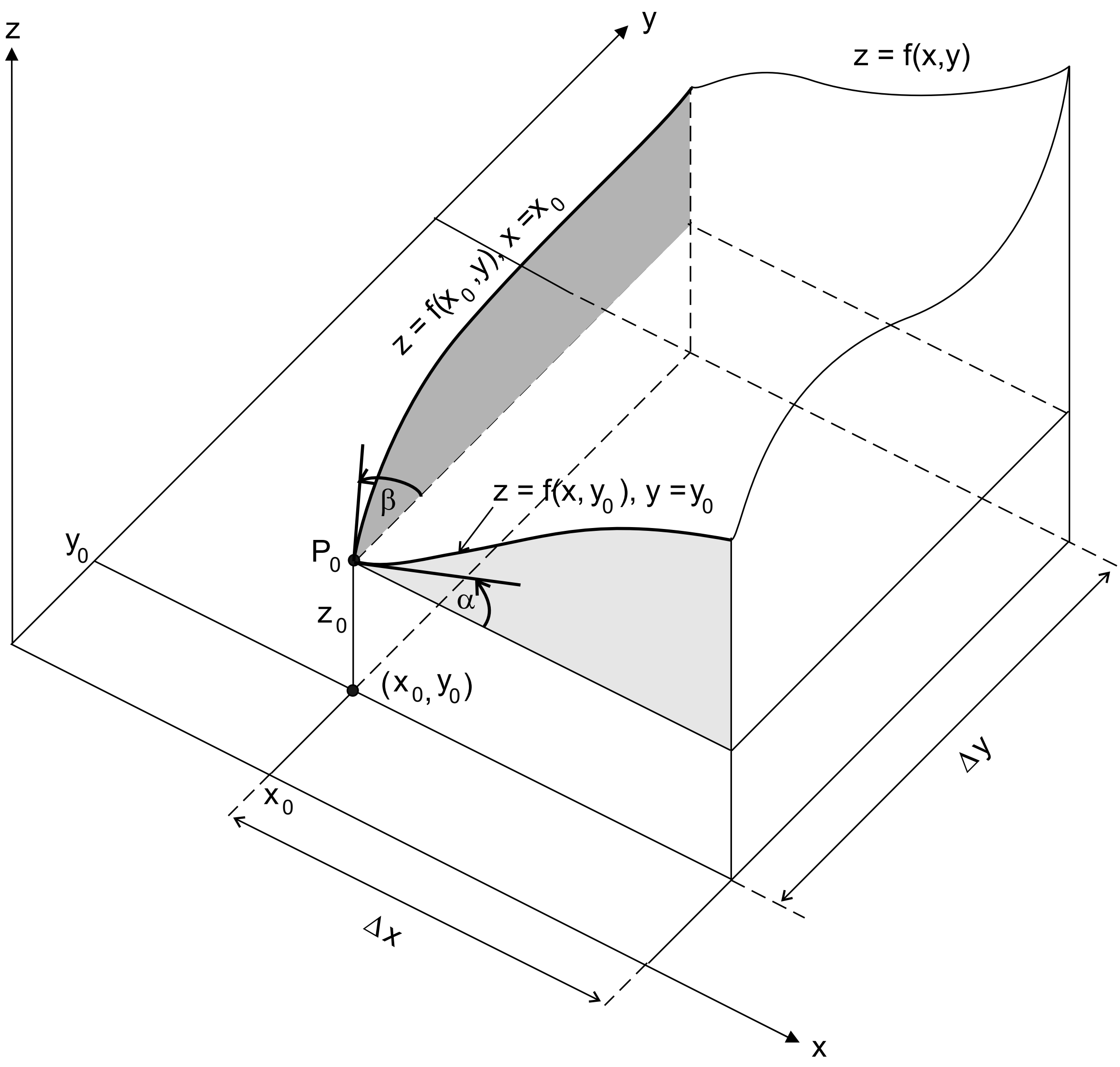
Gegeben sei die Funktion zweier Veränderlicher $z = f(x, y)$. Schneidet man die Fläche $z = f(x, y)$ mit der Ebene $y = y_0$, so erhält man die Schnittkurve $z = f(x, y_0),\ y = y_0$. Die Funktion $z = f(x, y_0) = g(x)$ ist nun eine Funktion in nur einer Variablen (nämlich x). Ihre Ableitung (nach x) für $x = x_0$ gibt die Steigung der Tangente an die Schnittkurve im Punkt $P_0 = (x_0, y_0, z_0 = f(x_0, y_0))$ an, d.h.

$$\tan\alpha = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}.$$

Ebenso ergibt sich für $x = x_0$ die Schnittkurve $z = f(x_0, y),\ x = x_0$. Die Funktion $z = f(x_0, y) = h(y)$ ist ebenfalls eine Funktion in nur einer Variablen (nämlich y). Ihre Ableitung (nach y) für $y = y_0$ gibt die Steigung der Tangente an die Schnittkurve im Punkt $(x_0, y_0, z_0 = f(x_0, y_0))$ an, d.h.

$$\tan\beta = \lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}.$$

Geometrie partieller Ableitungen



Definition

Als partielle Ableitungen (1.Ordnung) der Funktion $z = f(x,y)$ nach x bzw. nach y an der Stelle (x_0, y_0) definiert man die Grenzwerte

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x} =: f_x(x_0, y_0)$$

bzw.

$$\lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y} =: f_y(x_0, y_0).$$

Für $f_x(x_0, y_0)$ bzw. $f_y(x_0, y_0)$ sind auch die Symbole $\partial f/\partial x|_{(x_0,y_0)}$ bzw. $\partial f/\partial y|_{(x_0,y_0)}$ üblich. Die durch die Zuordnungsvorschrift $(x,y) \longmapsto f_x(x,y)$ bzw. $(x,y) \longmapsto f_y(x,y)$ gegebene Funktion heißt partielle Ableitung(sfunktion) nach x bzw. y.

Praktische Ermittlung der partiellen Ableitungen

a) Bildung von f_x: Behandle y als Konstante und differenziere nach x.

b) Bildung von f_y: Behandle x als Konstante und differenziere nach y.

Man bedenke aber, dass obiges Vorgehen nur dann möglich ist, wenn die jeweils zu betrachtenden Funktionen einer unabhängigen Variablen selbst differenzierbar sind.

Beispiel

Die partiellen Ableitungen von

$$f(x, y) = \mathrm{e}^{-x} \cos y$$

ergeben sich zu

$$f_x(x, y) = -\mathrm{e}^{-x} \cos y$$

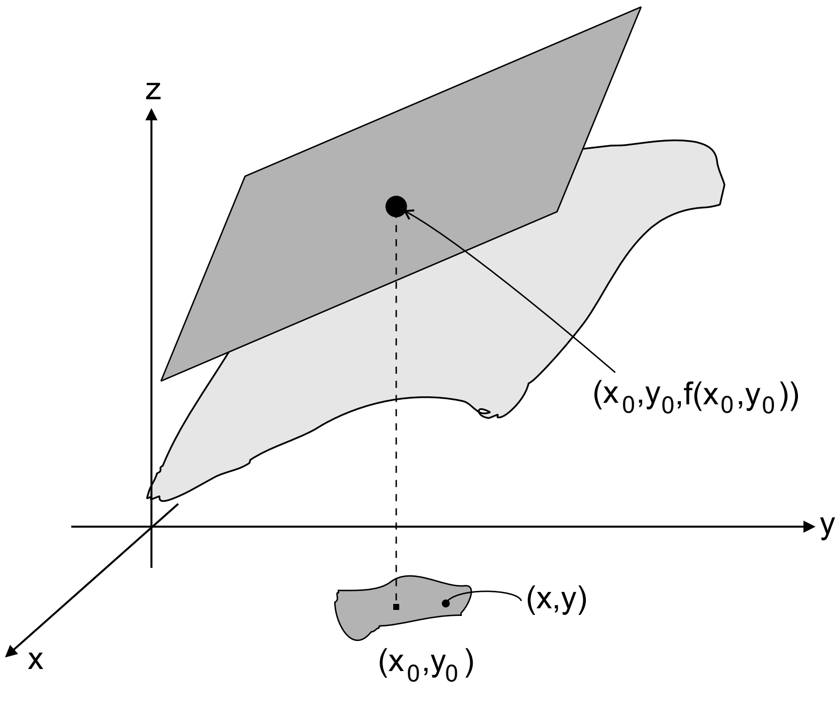
(y und damit $\cos y$ wurde als konstant angesehen) und

$$f_y(x, y) = -\mathrm{e}^{-x} \sin y$$

(x und damit e^{-x} wurde als konstant angesehen).

Tangentialebene, totale Differenzierbarkeit

Partiellen Ableitungen = Steigungen der Tangenten an einen beliebigen Flächenpunkt P_0 bzgl. zweier spezieller Schnittkurven, die parallel zur x- bzw. y-Achse verlaufen. Diese Tangenten spannen im $\mathbb{R}^3$ eine Ebene auf. Wenn der Graph von f im Punkt $(x_0, y_0, f(x_0, y_0))$ genügend „glatt" ist (d.h. keine „Knicke" aufweist), dann ist diese Ebene offensichtlich tangential zur Funktionsfläche und man nennt sie *Tangentialebene*.



Gleichung der Tangentialebene

Aufstellen der Gleichung der Tangentialebene an die Fläche $z = f(x, y)$ im Punkt (x_0, y_0, z_0) mit $z_0 = f(x_0, y_0)$:
Als Ansatz für die Ebene wählen wir die übliche Ebenendarstellung $z = l(x, y) = ax + by + c$. Die folgenden Forderungen dienen der Bestimmung von a, b, c:

a) Die Steigungen der Ebene in x- und y-Richtung müssen mit den Steigungen der Funktion bzgl. dieser Richtungen identisch sein, d.h.

$$f_x(x_0, y_0) = l_x(x_0, y_0) = a$$

und

$$f_y(x_0, y_0) = l_y(x_0, y_0) = b.$$

b) Der Punkt (x_0, y_0, z_0) muss natürlich in der Ebene liegen, d.h. $f(x_0, y_0) = l(x_0, y_0)$. Damit gilt

$$l(x_0, y_0) = ax_0 + by_0 + c$$

und somit

$$c = f(x_0, y_0) - ax_0 - by_0.$$

Gleichung der Tangentialebene

Die Gleichung der Tangentialebene an die Fläche

$$z = f(x, y)$$

im Flächenpunkt $P_0 = (x_0, y_0, z_0)$ **mit** $z_0 = f(x_0, y_0)$ **lautet:**

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

Totale Differenzierbarkeit

Definition

Die Funktion $z = f(x,y)$ heißt total differenzierbar oder linear approximierbar im Punkt (x_0, y_0), falls die partiellen Ableitungen $f_x(x_0, y_0)$, $f_y(x_0, y_0)$ existieren und gilt

$$\lim_{(x,y)\to(x_0,y_0)} \frac{|f(x,y) - t(x,y)|}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} = 0.$$

Dabei ist

$$t(x,y) = f(x_0, y_0) - f_x(x_0, y_0)(x - x_0) - f_y(x_0, y_0)(y - y_0)$$

die Tangentialebene in (x_0, y_0).

Wichtige Eigenschaften

Die folgenden Zusammenhänge gelten allgemein:

a) $z = f(x, y)$ **ist in** (x_0, y_0) **total differenzierbar, falls die partiellen Ableitungen** f_x **und** f_y **existieren und in einer Umgebung des Punktes** (x_0, y_0) **stetig sind.**

b) Ist $z = f(x, y)$ **in** (x_0, y_0) **total differenzierbar, dann ist** f **in** (x_0, y_0) **auch stetig.**

Man beachte aber, dass aus der bloßen Existenz der partiellen Ableitungen i. Allg. nicht die totale Differenzierbarkeit folgt. Ebenso lässt sich aus der totalen Differenzierbarkeit i. Allg. nicht die Stetigkeit der partiellen Ableitungen folgern.

Totales Differential

Definition

Das totale Differential von $z = f(x,y)$ an der Stelle (x_0, y_0) (zu den Zuwächsen dx, dy) ist definiert durch

$$dz := f_x(x_0, y_0)\,dx + f_y(x_0, y_0)\,dy\,.$$

Analog nennt man

$$dz = f_x\,dx + f_y\,dy$$

totales Differential von $z = f(x,y)$ an der (laufenden) Stelle (x, y). In einer hinreichend kleinen Umgebung von (x_0, y_0) ist das totale Differential dz eine gute Näherung für den Funktionszuwachs

$$\Delta z := f(x_0 + dx, y_0 + dy) - f(x_0, y_0).$$

Beispiel

a) *Die Funktion*

$$z = f(x, y) = xy$$

läßt sich geometrisch deuten: Sie misst die Fläche des Rechtecks mit den Seitenlängen x und y. Änderungen von x um dx und y um dy liefern den Funktionszuwachs

$$\Delta z = (x+dx)(y+dy)-xy = y\,dx+x\,dy+dx\,dy.$$

Das totale Differential ergibt sich zu

$$dz = f_x\,dx + f_y\,dy = y\,dx + x\,dy,$$

woraus für den Unterschied

$$\varepsilon = \Delta z - dz = dx\,dy$$

folgt. Man sagt daher auch, der Approximationsfehler ist klein von zweiter Ordnung.

Gradient

Die partiellen Ableitungen einer total differenzierbaren Funktion werden zu einem Vektor zusammengefasst:

Definition

Der zweidimensionale Spaltenvektor

$$\textbf{grad}\ f(x,y) = \nabla f(x,y) := \left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right)^T$$

heißt Gradient von f an der Stelle (x,y). Für $\vec{x} = (x,y)^T$ wird er auch $\nabla f(\vec{x})$ geschrieben.

Beispiel

Der Gradient der Funktion

$$z = f(x, y) = x^2y^2 + y + 1$$

ergibt sich zu

$$\begin{aligned}\nabla f(x, y) &= \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y)\right)^T \\ &= (2xy^2, 2x^2y + 1)^T.\end{aligned}$$

Damit gilt beispielsweise

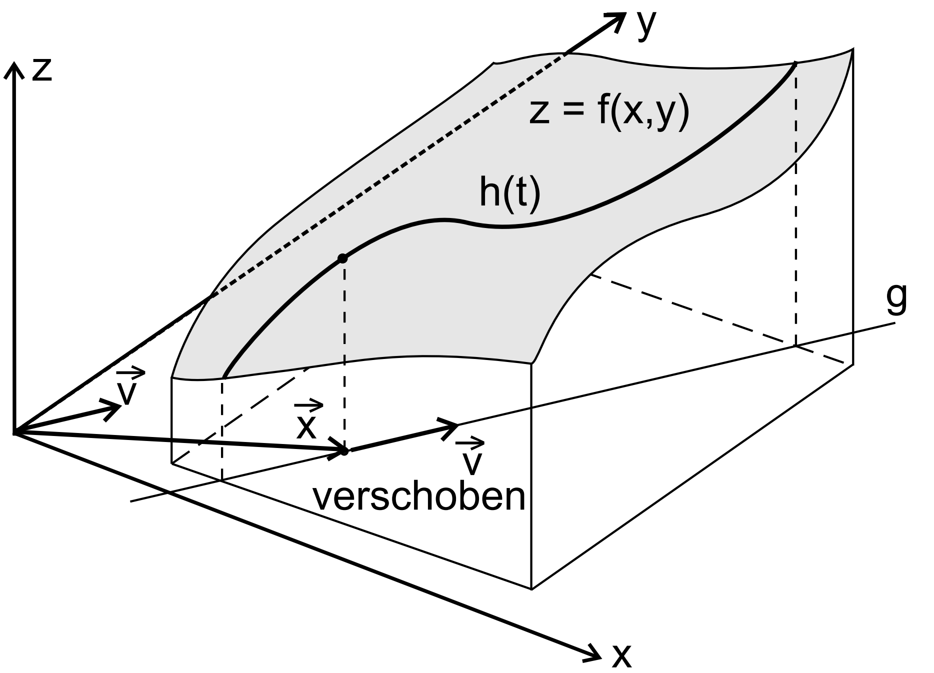
$$\nabla f(0, 0) = (0, 1)^T.$$

Richtungsableitung

Seien jetzt $f : I\!R^2 \to I\!R$ eine total differenzierbare Funktion und $\vec{x}, \vec{v} \in I\!R^2$ zwei fest gewählte Vektoren. Dann betrachte man die Funktion $h : I\!R \to I\!R$, die durch

$$h(t) := f(\vec{x} + t\vec{v})$$

definiert ist. Die Menge aller Punkte der Form $\vec{x} + t\vec{v}$, $t \in I\!R$ ist offensichtlich die Gerade g durch den Punkt $\vec{x}$ parallel zum Vektor $\vec{v}$. Daher stellt die Funktion h die Funktion f eingeschränkt auf die Gerade g dar.



Richtungsableitung

Von Interesse ist nun, wie stark sich die Funktionswerte von f längs der Geraden g im Punkt $\vec{x}$ ändern. Auskunft darüber erhält man aber, wie aus der eindimensionalen Differentialrechnung bekannt, durch die Ableitung $h'(t)$ an der Stelle $t = 0$.

Definition

Für eine total differenzierbare Funktion $f : I\!R^2 \to I\!R$ ist die Richtungsableitung von f im Punkt $\vec{x}$ in Richtung des Einheitsvektors $\vec{v}$ (d.h. $\|\vec{v}\| = 1$) gegeben durch

$$\frac{d}{dt} f(\vec{x} + t\vec{v})|_{t=0}.$$

Richtungsableitung

Da $h(t) := f(\vec{x} + t\vec{v})$ eine Funktion der *einen* Veränderlichen t ist, kann man deren Ableitung natürlich nach Einsetzen des Argumentes durch Ableiten bzgl. t berechnen. Es gibt aber auch eine Formel, die die Berechnung mittels Gradienten gestattet und meist weniger aufwendig ist.

Für die Richtungsableitung im Punkt $\vec{x}$ in Richtung $\vec{v} = (v_1, v_2)^T$, $\|\vec{v}\| = 1$, gilt:

$$\frac{d}{dt} f(\vec{x} + t\vec{v})|_{t=0} = \nabla f^T(\vec{x}) \cdot \vec{v} = f_x(\vec{x}) \cdot v_1 + f_y(\vec{x}) \cdot v_2.$$

Die Richtungsableitung ergibt sich aus dem Skalarprodukt von Gradient und Richtungsvektor.

Beispiel

Zu bestimmen sei die Richtungsableitung der Funktion

$$z = f(x, y) = x^2y^2 + y + 1$$

in $\vec{x} = (0, 0)^T$ *in Richtung* $\vec{v} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$*. Hierfür gibt es zwei Möglichkeiten:*

a) *Man stellt die Funktion* $h(t)$ *auf und differenziert nach* t*:*

$$\begin{aligned} h(t) &= f(\vec{x} + t\vec{v}) = f(t/\sqrt{2}, t/\sqrt{2}) \\ &= t^4/4 + t/\sqrt{2} + 1. \end{aligned}$$

Es ergibt sich $h'(t) = t^3 + 1/\sqrt{2}$ *und daraus die Richtungsableitung* $h'(0) = 1/\sqrt{2}$*.*

b) *Man berechnet die Richtungsableitung mittels Gradient nach obiger Formel:*

$$\nabla f^T(\vec{x}) \cdot \vec{v} = (0, 1) \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}.$$

Durch die Einschränkung auf „Einheitsrichtungen“ erhält man eine *eindeutige* Definition.

Eigenschaften des Gradienten

Der Gradient hat zwei wesentliche Eigenschaften, die nachfolgend aufgeführt sind:

Für den Gradienten gilt:

a) **Falls $\nabla f(\vec{x}) \neq \vec{0}$, dann zeigt $\nabla f(\vec{x})$ in die Richtung, längs der die Funktion f am schnellsten ansteigt.**

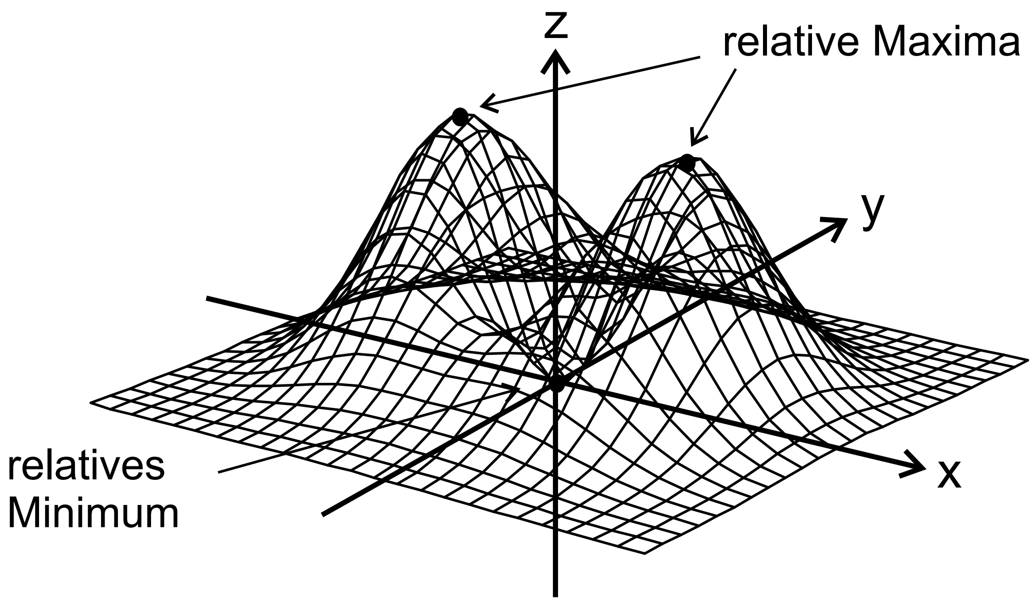
b) **Falls $\nabla f(\vec{x}) \neq \vec{0}$ und $\vec{x}$ auf einer Niveaulinie C von f liegt, dann ist $\nabla f(\vec{x})$ senkrecht zur Tangente an die Niveaulinie C im Punkt $\vec{x}$.**

Definition von Extremwerten

Der Begriff *Extremwert* lässt sich im $\mathbb{R}^2$ leicht aus der Definition für Funktionen einer Variablen verallgemeinern. Gilt

$$f(\vec{x}_0) \leq f(\vec{x}) \quad \text{bzw.} \quad f(\vec{x}_0) \geq f(\vec{x})$$

in einer hinreichend kleinen Umgebung von $\vec{x}_0$, dann heißt $\vec{x}_0$ *lokales Minimum* bzw. *lokales Maximum* von f. Strenge und globale Extremwerte sind analog zum $\mathbb{R}^1$ definiert. Die Abb. zeigt drei Extrema der Funktion $f(x, y) = (x^2 + 3y^2)\mathrm{e}^{1-x^2-y^2}$.



Bestimmung von Extremwerten

Falls $\vec{x}_0$ ein lokales Extremum von $f(\vec{x})$ ist, dann hat die durch

$$h(t) = f(\vec{x}_0 + t\vec{v})$$

($\vec{v}$ beliebig!) definierte Funktion h in $t = 0$ ein lokales Extremum. Notwendigerweise muss daher natürlich $h'(0) = 0$ gelten, d.h.

$$0 = h'(0) = \nabla f^T(\vec{x}_0) \cdot \vec{v}$$

für alle $\vec{v} \in I\!R^2$. Also steht der Vektor $\nabla f(\vec{x}_0)$ auf allen Vektoren $\vec{v} \in I\!R^2$ senkrecht. Dies ist aber nur für den Nullvektor möglich, d.h $\nabla f(\vec{x}_0) = \vec{0}$. Geometrisch bedeutet dies, dass die Tangentialebene in $\vec{x}_0$ *parallel* zur xy-Ebene verläuft.

Falls $\vec{x}_0$ ein lokales Extremum einer total differenzierbaren Funktion ist, dann gilt:

$$\boldsymbol{\nabla f(\vec{x}_0) = \vec{0},}$$

d.h. die partiellen Ableitungen müssen verschwinden:

$$\boldsymbol{f_x(\vec{x}_0) = f_y(\vec{x}_0) = 0.}$$

Beispiel

a) *Rotationsparaboloid $f(x,y) = x^2 + y^2$: Nach notwendigen Kriterium ergeben sich mögliche Kandidaten für Extremwerte als Lösung der Gleichungen*

$$\frac{\partial f}{\partial x} = 2x = 0 \quad \textit{und} \quad \frac{\partial f}{\partial y} = 2y = 0.$$

Einzig möglicher Punkt ist daher $(x,y) = (0,0)$. Da $f(x,y) \geq 0$ überall, ist dieser Punkt tatsächlich ein relatives Minimum.

b) *Gegeben sei $f(x,y) = x^2y^2+y+1$. Die Funktion hat den Gradienten*

$$\nabla f(x,y) = (2xy^2, 2x^2y + 1).$$

Nach dem notwendigen Kriterium muss folgendes Gleichungssystem gelöst werden:

$$\frac{\partial f}{\partial x} = 2xy^2 = 0, \quad \frac{\partial f}{\partial y} = 2x^2y + 1 = 0$$

Aus der ersten Gleichung folgt $x = 0$ oder $y = 0$. In beiden Fällen ist dann aber $2x^2y + 1 = 1 \neq 0$. Das System besitzt deshalb keine Lösung. Damit hat f keine Extrema.

Partielle Ableitungen 2. Ordnung

Für eine gegebene Funktion $z = f(x, y)$ berechnet man die partiellen Ableitungen (1. Ordnung): f_x und f_y.

Wenn diese ebenfalls partielle Ableitungen haben, dann sagt man, dass f *zweimal partiell differenzierbar* ist. Die Ableitungen

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

und

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \qquad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

heißen *partielle Ableitungen 2.Ordnung*, die beiden letzten Ableitungen speziell *gemischte partielle Ableitungen*.

In Analogie zur Notation $f_x,\ f_y$ für die partiellen Ableitungen 1.Ordnung sind hier folgende Abkürzungen üblich:

$$f_{xx},\ f_{yy},\ f_{yx} = (f_y)_x \quad \text{und} \quad f_{xy} = (f_x)_y.$$

Beispiel

a) *Für* $f(x, y) = xy + (x + 2y)^2$ *gilt*

$$\begin{aligned} f_x &= y + 2(x + 2y), \\ f_y &= x + 4(x + 2y). \end{aligned}$$

Die partiellen Ableitungen 2.Ordnung ergeben sich nun zu

$$f_{xx} = 2,\ f_{yy} = 8 \quad \text{und} \quad f_{xy} = f_{yx} = 5.$$

b) *Für* $f(x, y) = x^2y^2 + y + 1$ *gilt*

$$\begin{aligned} f_x &= 2xy^2, \\ f_y &= 2x^2y + 1 \end{aligned}$$

und damit

$$f_{xx} = 2y^2,\ f_{yy} = 2x^2$$

und

$$f_{xy} = f_{yx} = 4xy.$$

Satz von Schwarz

In den beiden Beispielen fällt auf, dass die gemischten Ableitungen jeweils gleich sind. Diese Eigenschaft gilt allgemein:

Falls die gemischten partiellen Ableitungen stetig sind, so ist die Reihenfolge der Ableitungen vertauschbar, d.h. es gilt stets

$$f_{xy} = f_{yx}.$$

Die Hessematrix

Die partiellen Ableitungen 2.Ordnung fasst man nun üblicherweise zu einer Matrix zusammen:

Definition

Sei $f(x, y)$ eine Funktion zweier unabhängiger Variabler mit stetigen partiellen Ableitungen 2.Ordnung. Dann heißt die Matrix

$$H(x_0, y_0) := \begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix}$$

Hessematrix von f im Punkt (x_0, y_0).

Durch ähnliche Untersuchungen wie im $\mathbb{R}^1$ erhält man mit Hilfe dieser Matrix nun hinreichende Bedingungen für Extremwerte. Dabei nutzt man aus, dass sich die Determinante $\det H(x_0, y_0)$ der Hessematrix H nach dem Satz von Schwarz ($f_{xy} = f_{yx}$!) ergibt zu:

$$f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0).$$

Hinreichendes Kriterium

Der Punkt (x_0, y_0) ist ein

a) relatives Minimum von f, falls die folgenden drei Bedingungen erfüllt sind:

- i) $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$,
- ii) $f_{xx}(x_0, y_0) > 0$,
- iii) $f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) > 0$.

b) relatives Maximum von f, falls die folgenden drei Bedingungen erfüllt sind:

- i) $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$,
- ii) $f_{xx}(x_0, y_0) < 0$,
- iii) $f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) > 0$.

c) Sattelpunkt von f, falls gilt:

$f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) < 0$.

Beispiel

a) *Gegeben sei die Funktion*

$$f(x,y) = x^2 + y^2,$$

für die in früherem Beispiel bereits der Punkt $(x_0, y_0) = (0,0)$ *als möglicher Kandidat ermittelt wurde (*$f_x(0,0) = f_y(0,0) = 0$*). Wegen*

$$f_x(x,y) = 2x,\ f_y(x,y) = 2y$$

folgt

$$\begin{aligned} f_{xx}(x,y) &= 2, \\ f_{xy}(x,y) &= f_{yx}(x,y) = 0 \\ f_{yy}(x,y) &= 2. \end{aligned}$$

Somit gilt

$$f_{xx}(0,0) = 2 > 0$$

und

$$f_{xx}(0,0) f_{yy}(0,0) - f_{xy}^2(0,0) = 4 > 0.$$

Daher ist der Ursprung nach dem hinreichenden Kriterium (Teil a)) ein Minimum.

Verallgemeinerung auf den $\mathbb{R}^n$

Sei jetzt $z = f(x_1, x_2, \ldots, x_n)$ eine Funktion von n unabhängigen Variablen x_i, $i = 1, \ldots, n$. Dann werden deren partielle Ableitungen 1. Ordnung durch

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ \ldots \ , \frac{\partial f}{\partial x_n} \quad \text{bzw.} \quad f_{x_1}, f_{x_2}, \ \ldots \ , f_{x_n}$$

bezeichnet. Die Ableitung nach x_i ergibt sich, indem man mit Ausnahme von x_i alle Variablen als konstant betrachtet und die Funktion nach x_i differenziert. Es gibt nun natürlich n^2 partielle Ableitungen 2.Ordnung: $f_{x_i x_j}$, $i, j = 1, \ldots, n$.

Sinngemäß gelten alle Aussagen der vorangegangenen Abschnitte. So erhält man beispielsweise in Analogie zum $\mathbb{R}^2$ das *totale Differential*

$$dz = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \ldots + \frac{\partial f}{\partial x_n} dx_n.$$

Auch im $\mathbb{R}^n$ ist das totale Differential in einer hinreichend kleinen Umgebung eine gute Näherung für den Funktionszuwachs.

Verallgemeinerung auf den $\mathbb{R}^n$

Zur Bestimmung potentieller Extremwertstellen muss man nun ein Gleichungssystem von n Gleichungen mit den n Unbekannten $x_1, \ldots, x_n$ lösen:

$$f_{x_1} = 0,\ f_{x_2} = 0, \quad \ldots \quad , f_{x_n} = 0.$$

Im $\mathbb{R}^3$ setzt man häufig

$$x_1 = x,\ x_2 = y,\ x_3 = z$$

und betrachtet dann Funktionen $f(x, y, z)$. Zur Bestimmung von Extremwertkandidaten ist dann

$$f_x = 0,\ f_y = 0,\ f_z = 0$$

zu lösen.

Ökonomie: Preiselastizität der Nachfrage

Untersucht man auf einem „Markt“ den Absatz eines bestimmten Gutes in Abhängigkeit von dessen Preis, so ergibt sich häufig ein funktionaler Zusammenhang. Dieser lässt sich durch eine sog. **Nachfragefunktion**

$$y = N(p)$$

angeben, wobei y die nachgefragte Menge des Gutes und p der Produktpreis ist. Solche Funktionen werden in der Regel von Ökonomen geschätzt, haben oft die Form $y = cp^\alpha$ mit Konstanten $c, \alpha \in \mathbb{R}$ und sind differenzierbar.

Eine wichtige Fragestellung ist nun die Sensitivität der Nachfrage $y = N(p)$:

Um welchen Prozentsatz ändert sich diese, wenn der Preis ausgehend vom aktuellen Niveau p_0 um x% steigt oder fällt?

Preiselastizität und relativer Fehler

Da die gesuchte Änderung dem aus der differentiellen Fehleranalyse bekanntem relativen Fehler

$$\left|\frac{\Delta y}{y}\right| \approx \left|\frac{dy}{y}\right| = \left|\frac{f'(x_0) \cdot dx}{f(x_0)}\right| = \left|\frac{f'(x_0)x_0}{f(x_0)}\right| \cdot \left|\frac{dx}{x_0}\right|$$

entspricht, gilt mit $x = p$ und $f = N$

$$\left|\frac{\Delta y}{y}\right| \approx \left|\frac{N'(p_0)p_0}{N(p_0)}\right| \cdot \left|\frac{dp}{p_0}\right| .$$

Die Kennzahl

$$\varepsilon(N, p_0) := \frac{N'(p_0)p_0}{N(p_0)}$$

heißt **Preiselastizität der Nachfrage**.

Diese gibt näherungsweise an, um wie viel Prozent sich die Nachfrage bei einer 1%-Preismodifikation ändert. Mit Hilfe der Preiselastitzität lassen sich nun ökonomische Sachverhalte ermitteln.

Fallunterscheidungen

- *Unelastische Preisnachfrage:*
  Dieser Fall gilt für $|\varepsilon| < 1$. Die Nachfrage reagiert kaum auf Preisänderungen. Dieses Verhalten zeigt sich bei Produkten, die den Grundbedarf der Marktteilnehmer decken und nicht substituierbar sind (z.B. Grundnahrungsmittel, Energie, Kraftstoff).

- *Elastische Preisnachfrage:*
  Diesen Fall hat man für $|\varepsilon| > 1$. Die Nachfrage reagiert relativ stark auf (auch kleine) Preisänderungen. Das gilt vor allem für Luxusgüter.

- *Vollkommen elastische bzw. unelastische Nachfrage:*
  Diese beiden Grenzfälle ergeben sich für $|\varepsilon| \to \infty$ bzw. $\varepsilon = 0$. Im ersten Fall bewirken sogar ganz geringe Preisänderungen sehr starke Nachfrageänderungen. Im zweiten Fall reagiert die Nachfrage nicht auf Preisänderungen.

Preisflexibilität der Nachfrage

Welche Preissenkung ist nötig, wenn ausgehend vom derzeitgen Absatz $y_0 = N(p_0)$ die Nachfrage um x % wachsen soll?

Nachfragefunktion ist i.A. invertierbar: Aus $y = N(p)$ kann man die Umkehrfunktion $p = P(y) = N^{-1}(y)$ (*Preisabsatzfunktion*) bestimmen und wie oben deren Elastizität, die sog. **Preisflexibilität der Nachfrage**, berechnen:

$$\varepsilon(P, y_0) := \frac{P'(y_0)y_0}{P(y_0)}.$$

Es geht aber auch einfacher: Benutzt man nämlich die Formel für die Ableitung der Umkehrfunktion, dann gilt:

$$\begin{aligned}
\frac{1}{\varepsilon(N, p_0)} &= \frac{N(p_0)}{N'(p_0)p_0} \\
&= \frac{1}{N'(p_0)} \cdot \frac{1}{p_0} \cdot N(p_0) \\
&= P'(y_0) \cdot \frac{1}{P(y_0)} \cdot y_0 = \varepsilon(P, y_0).
\end{aligned}$$

Die Preisflexibilität ergibt sich also als der reziproke Wert der Preiselastizität.

Nachfragefunktion und Marktgesetz

Realistischere Nachfragefunktionen hängen meistens nicht nur von einer Variablen ab.

Neben dem Preis p spielt natürlich auch das durchschnittlich verfügbare Einkommen e der potentiellen Kunden eine große Rolle: $y = N(p, e)$.

Empirisch ermittelte Funktionen haben dabei häufig die Form

$$N(p, e) = cp^{-\alpha} e^{\beta}$$

mit konstantem $c \in \mathbb{R}$ und reellen Größen $\alpha > 0, \beta > 0$. An der Gestalt der Funktion erkennt man eine Gesetzmäßigkeit des Marktes:

In der Regel ist die Größe der Nachfrage umgekehrt proportional zur Höhe des Preises — je teurer ein Produkt, desto weniger Konsumenten können es sich leisten.

Mit Hilfe der partiellen Ableitungen lassen sich nun analog zum Eindimensionalen *partielle Elastizitäten* bilden.