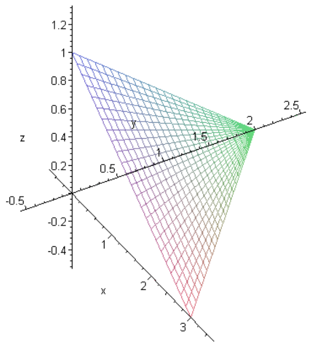
Funktionen mit zwei bzw. mehreren Variablen

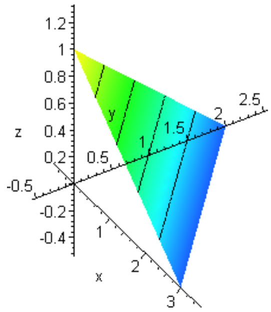


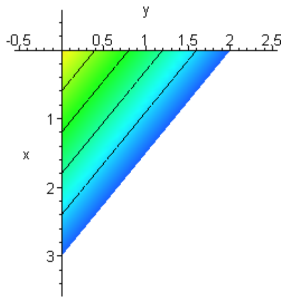
Beispiel 1: $f(x,y) = 1 - \frac{x}{3} - \frac{y}{2}$ mit $x, y \in \mathbb{R}$. Wie lässt sich ein Schaubild dieser Funktion zeichnen?

Setzt man $f(x,y) = z$, so folgt $z = 1 - \frac{x}{3} - \frac{y}{2}$. Nach Multiplikation mit 6 erhält man $2x + 3y + 6z = 6$. Aus der Schule ist bekannt, dass durch diese Gleichung eine Ebene beschrieben wird, welche die drei Achsen in den Punkten $(3,0,0)$, $(0,2,0)$ und $(0,0,1)$ schneidet.

Definition: Die Schnittlinien eines Schaubilds $z = f(x,y)$ mit Ebenen $z = c$, $c \in \mathbb{R}$, die parallel zur x-y-Ebene verlaufen, heißen **„Linien gleicher Höhe“** oder **„Niveaulinien“**.

In der Abbildung unten sind die Linien für die Höhen $c = 0$, $c = 0{,}2$, $c = 0{,}4$, $c = 0{,}6$ und $c = 0{,}8$ eingetragen. Dabei handelt es sich um Geraden parallel zur x-y-Ebene.

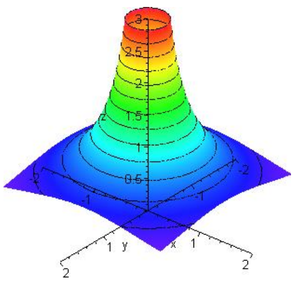


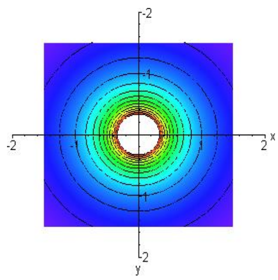


Die Projektion einer Linie gleicher Höhe in die x-y-Ebene wird als **„Höhenlinie“** bezeichnet. Die Menge dieser Höhenlinien bildet das **„Höhenliniendiagramm“.**

Beispiel 2: $f(x,y) = \frac{1}{\sqrt{x^2 + y^2}}$ für $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$.

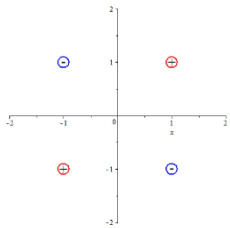
Die Linien gleicher Höhe sind gegeben durch $f(x,y) = c$ mit $c > 0$, d.h. $\sqrt{x^2 + y^2} = \frac{1}{c}$. Dabei handelt es sich um Kreise vom Radius $r = \frac{1}{c}$. Das Höhenliniendiagramm besteht also aus konzentrischen Kreisen. Das Schaubild ist gezeichnet für $-1{,}5 \le x, y \le 1{,}5$.



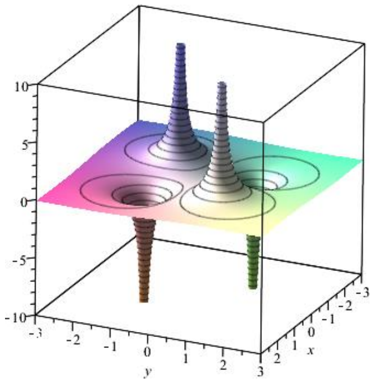


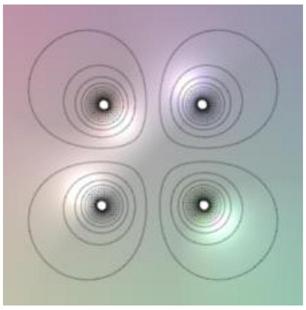
Anwendung: Das elektrische Potenzial einer im Ursprung befindlichen Punktladung Q ist $\varphi(r) = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r}$ mit der elektrischen Feldkonstanten $\varepsilon_0 = 8{,}854 \cdot 10^{-12}\,\frac{F}{m} = 8{,}854 \cdot 10^{-12}\,\frac{As}{Vm}$. Für das Potential φ im Punkt (x,y) folgt wegen $r = \sqrt{x^2 + y^2}$ die Beziehung $\varphi(x,y) = \frac{Q}{4\pi\varepsilon_0}\frac{1}{\sqrt{x^2 + y^2}}$.

Beispiel: Potenzialverlauf eines Quadrupols mit den Ladungen $+Q$ in den beiden Punkten $(1|1)$ und $(-1|-1)$ und $-Q$ in den beiden Punkten $(-1|1)$ und $(1|-1)$. Dann gilt für den Potenzialverlauf



$$\varphi(x,y) = \frac{Q}{4\pi\varepsilon_0} \cdot \left(\frac{1}{\sqrt{(x-1)^2 + (y-1)^2}} + \frac{1}{\sqrt{(x+1)^2 + (y+1)^2}} - \frac{1}{\sqrt{(x+1)^2 + (y-1)^2}} - \frac{1}{\sqrt{(x-1)^2 + (y+1)^2}} \right).$$





Man erkennt die geschlossenen Äquipotenziallinien $\varphi = c$ mit einer Konstanen c. Im Schaubild ist Q so gewählt, dass der Faktor $\frac{Q}{4\pi\varepsilon_0}$ den Zahlenwert 1 besitzt.

Die partielle Ableitung 1. Ordnung

Wir betrachten die Funktion f mit $z = f(x, y)$ und den beiden Variablen x, y.

Definition: Unter den **partiellen Ableitungen 1. Ordnung** einer Funktion $z = f(x, y)$ an der Stelle (x, y) versteht man:

$$\boxed{\frac{\partial f}{\partial x}(x, y) = f_x(x, y) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x, y)}{\Delta x}}$$ partielle Ableitung 1. Ordnung nach x,

$$\boxed{\frac{\partial f}{\partial y}(x, y) = f_y(x, y) = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x, y)}{\Delta y}}$$ partielle Ableitung 1. Ordnung nach y.

Beispiel: $f(x, y) = \frac{1}{\sqrt{x^2 + y^2}}$

$$\frac{\partial f}{\partial x}(x, y) = f_x(x, y) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x, y)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\frac{1}{\sqrt{(x+\Delta x)^2 + y^2}} - \frac{1}{\sqrt{x^2 + y^2}}}{\Delta x} = \lim_{\Delta x \to 0} \frac{\sqrt{x^2 + y^2} - \sqrt{(x+\Delta x)^2 + y^2}}{\Delta x \cdot \sqrt{(x+\Delta x)^2 + y^2} \cdot \sqrt{x^2 + y^2}}$$

Durch Erweitern mit $\sqrt{x^2 + y^2} + \sqrt{(x+\Delta x)^2 + y^2}$ folgt mit Hilfe der 3. Binomischen Formel

$$\lim_{\Delta x \to 0} \frac{x^2 + y^2 - \left((x+\Delta x)^2 + y^2\right)}{\Delta x \cdot \sqrt{(x+\Delta x)^2 + y^2} \cdot \sqrt{x^2 + y^2} \cdot \left(\sqrt{x^2 + y^2} + \sqrt{(x+\Delta x)^2 + y^2}\right)} = \lim_{\Delta x \to 0} \frac{x^2 + y^2 - \left(x^2 + 2x \cdot \Delta x + (\Delta x)^2 + y^2\right)}{\Delta x \cdot \sqrt{(x+\Delta x)^2 + y^2} \cdot \sqrt{x^2 + y^2} \cdot \left(\sqrt{x^2 + y^2} + \sqrt{(x+\Delta x)^2 + y^2}\right)} =$$

$$\lim_{\Delta x \to 0} \frac{\Delta x \cdot (-2x - \Delta x)}{\Delta x \cdot \sqrt{(x+\Delta x)^2 + y^2} \cdot \sqrt{x^2 + y^2} \cdot \left(\sqrt{x^2 + y^2} + \sqrt{(x+\Delta x)^2 + y^2}\right)} =$$

$$\lim_{\Delta x \to 0} \frac{-2x - \Delta x}{\sqrt{(x+\Delta x)^2 + y^2} \cdot \sqrt{x^2 + y^2} \cdot \left(\sqrt{x^2 + y^2} + \sqrt{(x+\Delta x)^2 + y^2}\right)} = \frac{-2x}{\sqrt{x^2 + y^2} \cdot \sqrt{x^2 + y^2} \cdot \left(\sqrt{x^2 + y^2} + \sqrt{x^2 + y^2}\right)} = \frac{-x}{\left(x^2 + y^2\right)^{3/2}}$$

Natürlich geht es mit der Ketteregel einfacher, wenn man beachtet, dass y wie eine Konstante zu behandeln ist.

$$f(x, y) = \frac{1}{\sqrt{x^2 + y^2}} = (x^2 + y^2)^{-1/2}.$$

Dann ist $\frac{\partial f}{\partial x}(x, y) = f_x(x, y) = -\frac{1}{2}(x^2 + y^2)^{-3/2} \cdot 2x = -\frac{x}{(x^2 + y^2)^{3/2}}$.

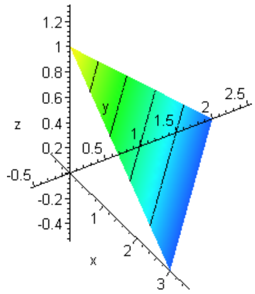
Analog folgt $\frac{\partial f}{\partial y}(x, y) = f_y(x, y) = -\frac{y}{(x^2 + y^2)^{3/2}}$.

Welche Information liefert die partielle Ableitung $\frac{\partial f}{\partial x}(x, y) = f_x(x, y)$?

Bei der partiellen Ableitung f_x bleibt y konstant. Der y-Wert ist konstant für alle Ebenen der Gleichung $y = y_0$, die senkrecht zur y-Achse verlaufen, also parallel zur x-z-Ebene sind. Die Ebene $y = y_0$ schneidet das Schaubild in einer Kurve der Gleichung $z = f(x, y_0)$. Einige dieser Kurven sind in den drei Abbildungen Seite 3 dargestellt.

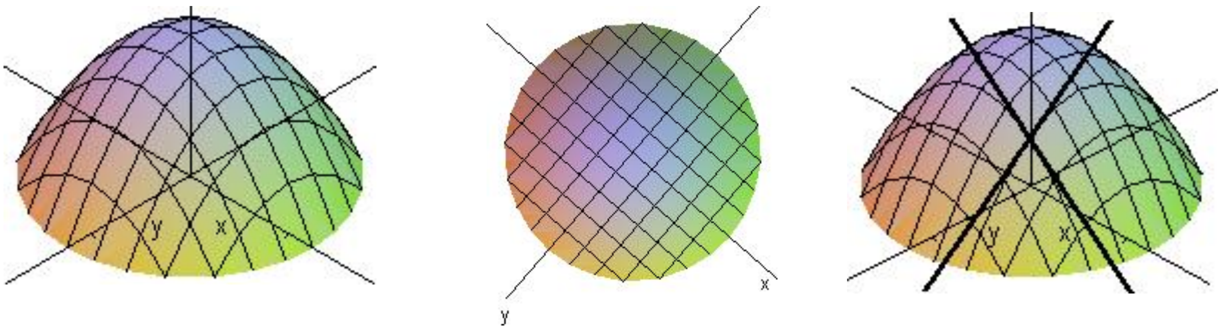
Die partielle Ableitung $\frac{\partial f}{\partial x}(x_0, y_0) = f_x(x_0, y_0)$ gibt die Tangentensteigung dieser Schnittkurve in der Ebene $y = y_0$ im Kurvenpunkt $P_0\left(x_0/y_0/f(x_0, y_0)\right) = P_0\left(x_0/y_0/z_0\right)$ an.

Beispiel 1: Die Achsenschnittpunkte der Ebene $z = 1 - \frac{x}{3} - \frac{y}{2}$ sind $X(3/0/0)$, $Y(0/2/0)$ und $Z(0/0/1)$.



Die Ableitung $\frac{\partial f}{\partial x}(x, y) = f_x(x, y) = -\frac{1}{3}$ ist gleich der Steigung $\frac{\Delta z}{\Delta x} = \frac{1-0}{0-3} = -\frac{1}{3}$ der Geraden XZ und auch jeder ihrer Parallelen, da y wegfällt.
Die Ableitung $\frac{\partial f}{\partial y}(x, y) = f_y(x, y) = -\frac{1}{2}$ ist gleich der Steigung $\frac{\Delta z}{\Delta y} = \frac{1-0}{0-2} = -\frac{1}{2}$ der Geraden YZ und auch jeder ihrer Parallelen, da x wegfällt.

Beispiel 2:
In den Abbildungen unten ist das Schaubild von $z = f(x, y) = 1 - x^2 - y^2$ für $z \ge 0$ dargestellt.
Die Schnittlinien gehören zu den Schnittebenen $x = 0$ (die y-z-Ebene), $x = \pm 0{,}2$, $x = \pm 0{,}4$, $x = \pm 0{,}6$ und $x = \pm 0{,}8$; außerdem $y = 0$ (die x-z-Ebene), $y = \pm 0{,}2$, $y = \pm 0{,}4$, $y = \pm 0{,}6$ und $y = \pm 0{,}8$.
In der rechten Abbildung sind die beiden Tangenten parallel zur x-z- und zur y-z-Ebene im Kurvenpunkt $P_0(0{,}4/0{,}4/0{,}68)$ eingezeichnet. Wegen $\frac{\partial f}{\partial x}(x, y) = f_x(x, y) = -2x$ und $\frac{\partial f}{\partial y}(x, y) = f_y(x, y) = -2y$ haben sie jeweils die Steigungen $-2 \cdot 0{,}4 = -0{,}8$.



Analog wird bei $\frac{\partial f}{\partial y}(x_0, y_0) = f_y(x_0, y_0)$ das Schaubild $z = f(x, y)$ mit Ebenen der Gleichung $x = x_0$, die senkrecht zur x-Achse verlaufen, in Kurven der Gleichung $z = f(x_0, y)$ geschnitten. Die Tangentensteigung dieser Schnittkurve im Kurvenpunkt $P_0\left(x_0/y_0/f(x_0, y_0)\right)$ ist dann durch $\frac{\partial f}{\partial y}(x_0, y_0) = f_y(x_0, y_0)$ gegeben.

Weitere Beispiele zur partiellen Ableitung:

1. $f(x, y) = \ln(1 + 2x - 3y + xy)$. Dann ist $f_x(x, y) = \frac{2 + y}{1 + 2x - 3y + xy}$ und $f_y(x, y) = \frac{-3 + x}{1 + 2x - 3y + xy}$.

2. $f(x, y) = (x \cdot y)^y$ für $x, y > 0$. Dann ist $f_x(x, y) = y \cdot (x \cdot y)^{y-1} \cdot y = y^2 \cdot (x \cdot y)^{y-1} = x^{y-1} \cdot y^{y+1}$.
   $f_y(x, y)$ erhält man durch logarithmische Differenziation. Mit $\ln\left(f(x, y)\right) = y \cdot \ln(x \cdot y)$ folgt
   $\frac{1}{f(x, y)} \cdot f_y(x, y) = 1 \cdot \ln(x \cdot y) + y \cdot \frac{1}{x \cdot y} \cdot x = \ln(x \cdot y) + 1$, also $f_y(x, y) = (x \cdot y)^y \cdot \left(\ln(x \cdot y) + 1\right)$.

3. $f(x_1, x_2, x_3, x_4) = \frac{x_1 \cdot x_2^3 - 4x_3}{x_2} - 3x_4$. Dann ist $f_{x_1}(x_1, x_2, x_3, x_4) = x_2^2$, $f_{x_2}(x_1, x_2, x_3, x_4) = \frac{2x_1 x_2^3 + 4x_3}{x_2^2}$,
   $f_{x_3}(x_1, x_2, x_3, x_4) = -\frac{4}{x_2}$ und $f_{x_4}(x_1, x_2, x_3, x_4) = -3$.

Die Tangentialebene

In der x-y-Ebene ist eine Gerade gegeben durch die Gleichung $y = a \cdot x + b$. Und die Tangente in einem Kurvenpunkt $\left(x_0, f(x_0)\right)$ des Schaubilds $y = f(x)$ hat die Gleichung $t: y = f'(x_0) \cdot (x - x_0) + f(x_0)$, sofern diese Tangente nicht parallel zur y-Achse verläuft. In diesem Fall existiert $f'(x_0)$ nicht.

Zur Erinnerung die Taylor-Reihe: $f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(x_0)}{i!} \cdot (x - x_0)^i = f(x_0) + f'(x_0) \cdot (x - x_0) + \frac{f''(x_0)}{2} \cdot (x - x_0)^2 + \cdots$

Die ersten beiden Summanden gehören zur Tangente.

Im x-y-z-Raum ist eine Ebene gegeben durch die Gleichung $z = a \cdot x + b \cdot y + c$, sofern sie nicht parallel zur z-Achse verläuft. Sie schneidet die z-Achse im Punkt $(0/0/c)$. Gesucht ist nun eine Gleichung der Tangentialebene T im Punkt $P_0\left(x_0, y_0, f(x_0, y_0)\right)$ des dreidimensionalen Schaubilds $z = f(x, y)$.

Die Tangentialebene T: $z = a \cdot x + b \cdot y + c$ im Punkt P_0 muss die beiden Tangenten parallel zur x-z- und zur y-z-Ebene enthalten. Deshalb muss $\frac{\partial z}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)$ und $\frac{\partial z}{\partial y}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0)$ gelten, d.h.

$a = \frac{\partial f}{\partial x}(x_0, y_0) = f_x(x_0, y_0)$ und $b = \frac{\partial f}{\partial y}(x_0, y_0) = f_y(x_0, y_0)$. Außerdem muss die Tangentialebene den Punkt $P_0\left(x_0, y_0, f(x_0, y_0)\right)$ des Schaubilds enthalten. Durch Einsetzen folgt $f(x_0, y_0) = a \cdot x_0 + b \cdot y_0 + c$, so dass $c = -f_x(x_0, y_0) \cdot x_0 - f_y(x_0, y_0) \cdot y_0 + f(x_0, y_0)$. Es ergibt sich die Gleichung der Tangentialebene zu

$\boxed{T: z = f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0) + f(x_0, y_0)}$ bzw.

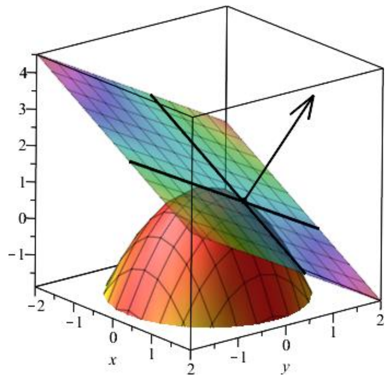
$\boxed{T: z = \begin{pmatrix} f_x(x_0, y_0) \\ f_y(x_0, y_0) \end{pmatrix} \bullet \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + f(x_0, y_0)}$ Dabei bedeutet $\bullet$ das Skalarprodukt.

Die Taylorreihe für eine Funktion $f(x, y)$ mit zwei Variablen um die Stelle (x_0, y_0) lautet

$f(x, y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{\partial^i}{\partial x^i} \frac{\partial^j}{\partial y^j} f(x_0, y_0) \cdot \frac{(x - x_0)^i \cdot (y - y_0)^j}{i! \cdot j!}$. Die ersten Summanden lauten

$$f(x, y) = f(x_0, y_0) \; + \; f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0) +$$
$$+ \frac{1}{2!} f_{xx}(x_0, y_0) \cdot (x - x_0)^2 + f_{xy}(x_0, y_0) \cdot (x - x_0) \cdot (y - y_0) + \frac{1}{2!} f_{yy}(x_0, y_0) \cdot (y - y_0)^2 + \cdots$$

Die erste Zeile gehört zur Tangentialebene.



Beispiel: Es sei $z = f(x, y) = 1 - x^2 - y^2$. In der Abbildung rechts sind die beiden Tangenten im Punkt $P_0(0{,}4/0{,}4/0{,}68)$ eingezeichnet, die parallel zur x-z- und zur y-z-Ebene verlaufen.
Die Gleichung der Tangentialebene T lautet
T: $z = -2x_0 \cdot (x - x_0) - 2y_0 \cdot (y - y_0) + 1 - x_0^2 - y_0^2$.
Speziell für den Kurvenpunkt $P_0(0{,}4/0{,}4/0{,}68)$ folgt dann
T: $z = -0{,}8 \cdot (x - 0{,}4) - 0{,}8 \cdot (y - 0{,}4) + 0{,}68$, d.h.
T: $z = -0{,}8x - 0{,}8y + 1{,}32$.

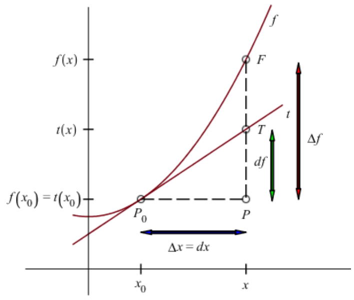
Zusatz: Die Gleichung $T: z = f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0) + f(x_0, y_0)$ der Tangentialebene des Schaubilds von f im Punkt $P_0(x_0/y_0/z_0)$ mit $z_0 = f(x_0, y_0)$ lässt sich mit Hilfe des Skalarprodukts umformen in $T: f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0) - \left(z - f(x_0, y_0)\right) = 0$, also $T: \begin{pmatrix} f_x(x_0, y_0) \\ f_y(x_0, y_0) \\ -1 \end{pmatrix} \bullet \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} = 0$.

Somit steht der Vektor $\vec{n}_T = \begin{pmatrix} f_x(x_0, y_0) \\ f_y(x_0, y_0) \\ -1 \end{pmatrix}$ senkrecht auf dem Vektor $\overrightarrow{P_0P} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{pmatrix}$, der durch die beiden in der Tangentialebene T liegenden Punkten P_0 und P gegeben ist. Siehe Schaubild.

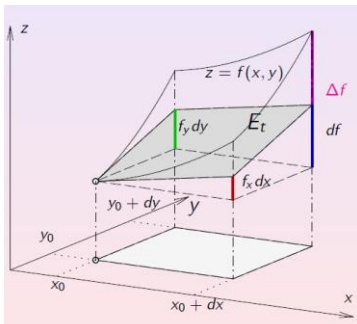
Satz: Es sei $z = f(x,y)$ eine Funktion mit zwei Variablen. Dann ist $\boxed{\vec{n}_T = \begin{pmatrix} f_x(x_0, y_0) \\ f_y(x_0, y_0) \\ -1 \end{pmatrix}}$ **ein Normalenvektor der Tangentialebene T** des Schaubilds von f im Punkt $P_0(x_0 / y_0 / z_0)$, wobei $z_0 = f(x_0, y_0)$.

Eine Parameterform von T ist $T: \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + r \begin{pmatrix} 1 \\ 0 \\ f_x(x_0, y_0) \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ f_y(x_0, y_0) \end{pmatrix}$, denn die beiden Richtungsvektoren $\vec{u} = \begin{pmatrix} 1 \\ 0 \\ f_x(x_0, y_0) \end{pmatrix}$ und $\vec{v} = \begin{pmatrix} 0 \\ 1 \\ f_y(x_0, y_0) \end{pmatrix}$ sind wegen $\vec{u} \bullet \vec{n}_T = 0$ und $\vec{v} \bullet \vec{n}_T = 0$ orthogonal zu $\vec{n}_T$. Dabei verläuft $\vec{u} = \begin{pmatrix} 1 \\ 0 \\ f_x(x_0, y_0) \end{pmatrix}$ parallel zur x-z-Ebene und $\vec{v} = \begin{pmatrix} 0 \\ 1 \\ f_y(x_0, y_0) \end{pmatrix}$ parallel zur y-z-Ebene.

Das totale Differenzial



Im Zweidimensionalen sei die Kurve $y = f(x)$ gegeben. Wenn sich x um $\Delta x = x - x_0$ ändert, ändert sich $y = f(x)$ um $\boxed{\Delta y = f(x) - f(x_0)}$. In unmittelbarer Nähe eines Kurvenpunktes $P_0(x_0, y_0)$ kann man das Schaubild durch seine Tangente t: $y_t = f'(x_0) \cdot (x - x_0) + f(x_0)$ annähern. Diese Gleichung lässt sich umschreiben in $y_t - f(x_0) = f'(x_0) \cdot (x - x_0)$ und neu interpretieren: Wenn sich x um $\Delta x = x - x_0$ ändert, dann ändert sich $f(x)$ näherungsweise um $df = \Delta y_t = f'(x_0) \cdot \Delta x$ bzw. differenziell geschrieben: $\boxed{df = f'(x_0) \cdot dx}$.



Bei zwei Variablen wird das Schaubild von $z = f(x,y)$ im Punkt $P_0(x_0, y_0, z_0)$ näherungsweise durch seine Tangentialebene T (in der Skizze mit E_t bezeichnet) ersetzt:
$T: z_T = f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0) + f(x_0, y_0)$.
Mit $dz = z_T - f(x_0, y_0)$, $dx = x - x_0$ und $dy = y - y_0$ folgt
$\boxed{df = f_x(x_0, y_0) \cdot dx + f_y(x_0, y_0) \cdot dy}$.
Dagegen ist $\boxed{\Delta f = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)}$ die wahre Änderung des Funktionswertes von f aber nur annähernd gleich df.

Beispiel: Aus obigen Schaubildern kennen wir $z = f(x,y) = 1 - x^2 - y^2$ mit $f(0,4 / 0,4) = 0,68$.
Bewegt man sich nur wenig von $(x_0 / y_0) = (0,4 / 0,4)$ nach $(x / y) = (0,41 / 0,38)$, so ändert sich z um
$\Delta z = \Delta f = f(0,41 / 0,38) - f(0,4 / 0,4) = \left(1 - 0,41^2 - 0,38^2\right) - \left(1 - 0,4^2 - 0,4^2\right) = 0,6875 - 0,68 = 0,0075$
Andererseits gilt für das totale Differenzial dz
$df = f_x(x_0, y_0) \cdot dx + f_y(x_0, y_0) \cdot dy = -2x_0 \cdot dx - 2y_0 \cdot dy = -2 \cdot 0,4 \cdot 0,01 - 2 \cdot 0,4 \cdot (-0,02) = 0,008$, was näherungsweise mit Δz übereinstimmt.

Allgemein ist für die Funktion $y = f(x_1, x_2, \ldots, x_n)$ das **totale (vollständige) Differenzial** gleich

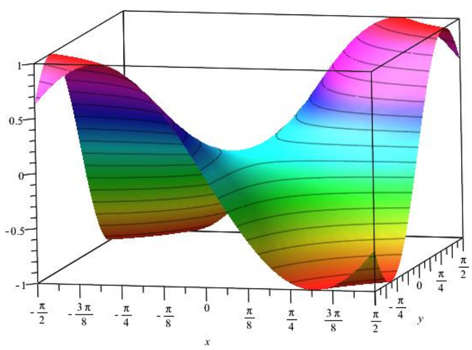
$$\boxed{dy = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \ldots + \frac{\partial f}{\partial x_n} dx_n = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i = \sum_{i=1}^{n} f_{x_i} dx_i}\,.$$

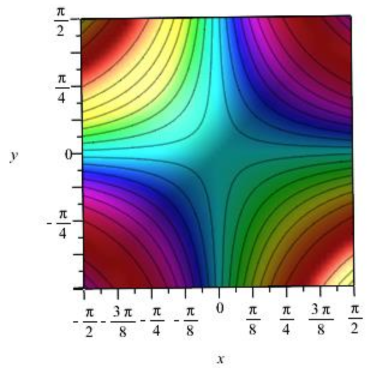
dy gibt die Änderung dy von y an, wenn sich die Variablen x_i um dx_i ändern.

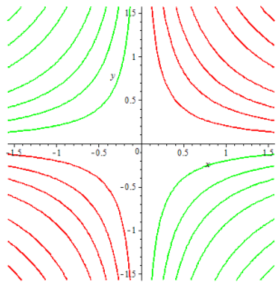
Definition: Der Vektor $\nabla f = \text{grad}(f) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$ heißt **Gradient von f** oder auch **Nabla von f** .

Beispiel: $z = f(x, y) = \sin(x \cdot y)$ für $-\pi/2 \le x \le \pi/2$, $-\pi/2 \le y \le \pi/2$.

Es gilt $\text{grad}(f) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} y \cdot \cos(x \cdot y) \\ x \cdot \cos(x \cdot y) \end{pmatrix}$.



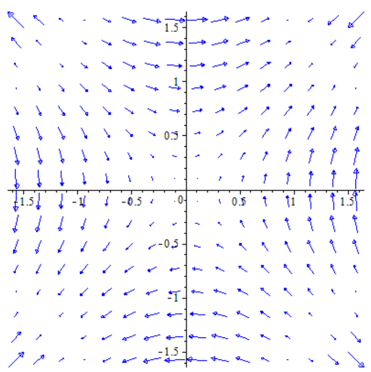


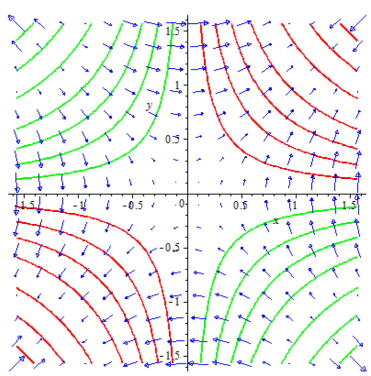


Das linke dreidimensionale Schaubild zeigt die **Niveaulinien** (Linien gleicher Höhe), in denen $f(x, y) = \sin(x \cdot y)$ konstant ist.

Das mittlere Schaubild ist eine Projektion des linken Schaubilds in die x-y-Ebene, das sog. Höhenliniendiagramm.

Das rechte Schaubild ist ebenfalls das Höhenliniendiagramm. Bei ihm sind im ersten und dritten Quadranten von innen nach außen die Höhenlinien (in Rot) $z = 0{,}2$, $z = 0{,}4$, $z = 0{,}6$, $z = 0{,}8$, $z = 0{,}98$, $z = 0{,}98$ und $z = 0{,}8$ gezeichnet. Im zweiten und vierten Quadranten sind von innen nach außen die Höhenlinien (in Grün) $z = -0{,}2$, $z = -0{,}4$, $z = -0{,}6$, $z = -0{,}8$, $z = -0{,}98$, $z = -0{,}98$ und $z = -0{,}8$ gezeichnet.





Das linke Bild zeigt das Gradientenfeld von $f(x,y)=\sin(x\cdot y)$ mit den Gradientenvektoren

$\operatorname{grad}(f)=\begin{pmatrix}\frac{\partial f}{\partial x}\\ \frac{\partial f}{\partial y}\end{pmatrix}$. Die Länge der Pfeile ist proportional zu ihrem Betrag $|\operatorname{grad}(f)|=\sqrt{\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2}$.

Das rechte Bild vereinigt die Gradienten mit den Höhenlinien.

Satz: Der Gradient $\operatorname{grad}(f)=\begin{pmatrix}f_x(x,y)\\ f_y(x,y)\end{pmatrix}$ steht an jeder Stelle (x,y) **senkrecht auf der** durch (x,y) verlaufenden **Höhenlinie** von f und zeigt **in Richtung des größten Zuwachses von f**.

Beweis: Für das totale Differenzial von f gilt $df=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy$. Es lässt sich als Skalarprodukt interpretieren:

$df=\begin{pmatrix}\frac{\partial f}{\partial x}\\ \frac{\partial f}{\partial y}\end{pmatrix}\bullet\begin{pmatrix}dx\\ dy\end{pmatrix}=|\operatorname{grad}(f)|\cdot\left|\begin{pmatrix}dx\\ dy\end{pmatrix}\right|\cdot\cos\varphi$, wobei φ der Winkel zwischen $\operatorname{grad}(f)$ und $d\vec{r}=\begin{pmatrix}dx\\ dy\end{pmatrix}$ ist.

Nun verbinde $d\vec{r}=\begin{pmatrix}dx\\ dy\end{pmatrix}$ zwei verschiedene Punkte auf einer Niveaulinie, so dass $df=|\operatorname{grad}(f)|\cdot\left|\begin{pmatrix}dx\\ dy\end{pmatrix}\right|\cdot\cos\varphi=0$ sein muss, da auf ihr $f(x,y)=c$ konstant ist. Dies ist aber allgemein nur möglich, wenn $\cos\varphi=0$, also $\varphi=90°$ beträgt. Somit steht der **Gradient grad(f) senkrecht auf den Niveaulinien**.

Nun sei $\begin{pmatrix}dx\\ dy\end{pmatrix}=\vec{e}=\begin{pmatrix}e_x\\ e_y\end{pmatrix}$ ein Einheitsvektor. Dann ist $df=\begin{pmatrix}\frac{\partial f}{\partial x}\\ \frac{\partial f}{\partial y}\end{pmatrix}\bullet\vec{e}=|\operatorname{grad}(f)|\cdot 1\cdot\cos\varphi$ genau dann positiv und zugleich maximal, wenn $\cos\varphi=1$, also $\varphi=0°$ beträgt, und somit $\vec{e}$ in Richtung von $\operatorname{grad}(f)$ zeigt.

Der Gradient grad(f) zeigt also im zweidimensionalen Höhenliniendiagramm in Richtung des maximalen Zuwachses von f .

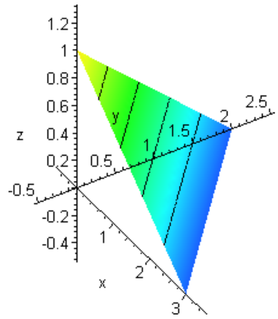
Folgerungen:

① $-\operatorname{grad}(f)=-\begin{pmatrix}\frac{\partial f}{\partial x}\\ \frac{\partial f}{\partial y}\end{pmatrix}$ zeigt in Richtung der maximalen Abnahme von f .

② $\begin{pmatrix}\frac{\partial f}{\partial y}\\ -\frac{\partial f}{\partial x}\end{pmatrix}$ ist ein Richtungsvektor in Richtung der Höhenlinien. Denn mit Hilfe des Skalarprodukts folgt

$\operatorname{grad}(f)\bullet\begin{pmatrix}\frac{\partial f}{\partial y}\\ -\frac{\partial f}{\partial x}\end{pmatrix}=\begin{pmatrix}\frac{\partial f}{\partial x}\\ \frac{\partial f}{\partial y}\end{pmatrix}\bullet\begin{pmatrix}\frac{\partial f}{\partial y}\\ -\frac{\partial f}{\partial x}\end{pmatrix}=\frac{\partial f}{\partial y}\cdot\frac{\partial f}{\partial x}-\frac{\partial f}{\partial x}\cdot\frac{\partial f}{\partial y}=0$, so dass diese beiden Vektoren orthogonal sind.

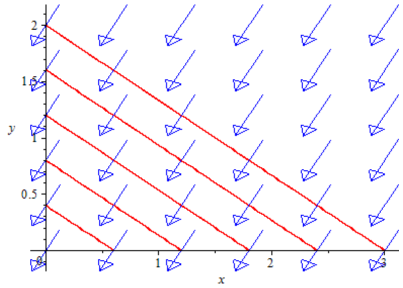
Generell steht $\begin{pmatrix}-b\\ a\end{pmatrix}$ senkrecht auf $\begin{pmatrix}a\\ b\end{pmatrix}$.



Beispiel 1: Für die Ebene $z = f(x,y) = 1 - \frac{x}{3} - \frac{y}{2}$ gilt $\frac{\partial f}{\partial x}(x_0,y_0) = f_x(x_0,y_0) = -\frac{1}{3}$

und $\frac{\partial f}{\partial y}(x_0,y_0) = f_y(x_0,y_0) = -\frac{1}{2}$, so dass $\mathrm{grad}(f) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ -\frac{1}{2} \end{pmatrix} = \frac{1}{6}\begin{pmatrix} -2 \\ -3 \end{pmatrix}$.

Der Vektor $\overrightarrow{XY} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$ von $X(3/0)$ nach $Y(0/2)$ gibt die Richtung der Niveaulinien



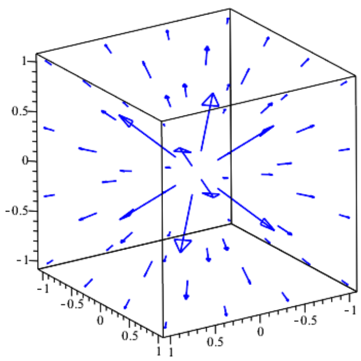
an. Wegen $\mathrm{grad}(f) \bullet \overrightarrow{XY} = \begin{pmatrix} -\frac{1}{3} \\ -\frac{1}{2} \end{pmatrix} \bullet \begin{pmatrix} -3 \\ 2 \end{pmatrix} = 1 - 1 = 0$ steht der Gradient senkrecht auf $\overrightarrow{XY}$. Außerdem nimmt $f(x,y)$ am stärksten zu (Schaubild am steilsten), wenn sich (x,y) in Richtung $\begin{pmatrix} -\frac{1}{3} \\ -\frac{1}{2} \end{pmatrix} = \frac{1}{6}\begin{pmatrix} -2 \\ -3 \end{pmatrix}$ ändert. Gezeichnet sind die Höhenlinien der Richtung $\begin{pmatrix} -3 \\ 2 \end{pmatrix}$ und die dazu orthogonalen Gradientenvektoren der Richtung $\begin{pmatrix} -2 \\ -3 \end{pmatrix}$, die in Richtung des stärksten Zuwachses zeigen.

Beispiel 2 aus der Physik:

Es sei $\varphi(P)$ das Potential im Punkt P eines elektrischen Feldes. Dann gilt für die elektrische Feldstärke $\boxed{\vec{E} = -\mathrm{grad}(\varphi)}$.

Im Ursprung befinde sich eine Punktladung Q, Q mit Vorzeichen. Dann gilt für das Potential in ihrer Umgebung im Abstand r die Beziehung $\varphi(r) = \frac{Q}{4\pi\varepsilon_0 r} = \frac{Q}{4\pi\varepsilon_0\sqrt{x^2+y^2+z^2}}$. Das Potential ist so normiert, dass es im Unendlichen gleich Null ist. Für die elektrische Feldstärke gilt dann

$$\vec{E} = -\mathrm{grad}(\varphi) = -\begin{pmatrix} \frac{\partial \varphi}{\partial x} \\ \frac{\partial \varphi}{\partial y} \\ \frac{\partial \varphi}{\partial z} \end{pmatrix} = +\frac{Q}{4\pi\varepsilon_0 r^3}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{Q}{4\pi\varepsilon_0 r^3}\vec{r} = \frac{Q}{4\pi\varepsilon_0 r^2}\vec{e}_r,$$ wobei $\vec{r} = \overrightarrow{OP} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = r \cdot \vec{e}_r$ der Vektor vom Ursprung O zum Punkt P ist und $\vec{e}_r$ der Einheitsvektor in dieser Richtung darstellt. Wenn die Ladung Q positiv ist, dann zeigt $\vec{E}$ von Q weg. Für den Betrag der elektrischen Feldstärke gilt

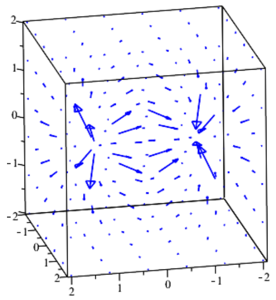


$$E(r) = \left|\vec{E}(r)\right| = \frac{Q}{4\pi\varepsilon_0 r^3}\left|\vec{r}\right| = \frac{Q}{4\pi\varepsilon_0 r^2}.$$

In der Abbildung ist das dreidimensionale Gradientenfeld $-\mathrm{grad}(f)$ von

$f(x,y,z) = \frac{1}{\sqrt{x^2+y^2+z^2}}$ dargestellt.

Es ist $-\mathrm{grad}(f) = \frac{1}{(x^2+y^2+z^2)^{3/2}}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{r^3}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{r^2}\vec{e}_r$.

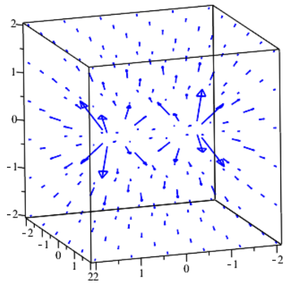


Beispiel 3 aus der Physik:

Im Punkt $(1/0/0)$ befinde sich die positive Ladung Q und im Punkt $(-1/0/0)$ die negative Ladung $-Q$. Dann gilt für das Potential im Punkt (x,y,z) die Beziehung

$$\varphi(x,y,z)=\frac{Q}{4\pi\varepsilon_0}\left(\frac{1}{(x-1)^2+y^2+z^2}-\frac{1}{(x+1)^2+y^2+z^2}\right).$$

Für die elektrische Feldstärke gilt dann $\vec{E}=-\operatorname{grad}(\varphi)$. Das Gradientenfeld zeigt, dass $\vec{E}$ von der positiven zur negativen Ladung zeigt.



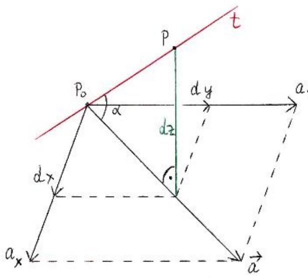
Beispiel 4 aus der Physik:

Im Punkt $(1/0/0)$ und im Punkt $(-1/0/0)$ befindet sich jeweils die positive Ladung Q. Dann gilt für das Potential im Punkt (x,y,z) die Beziehung

$$\varphi(x,y,z)=\frac{Q}{4\pi\varepsilon_0}\left(\frac{1}{(x-1)^2+y^2+z^2}+\frac{1}{(x+1)^2+y^2+z^2}\right).$$

Für die elektrische Feldstärke gilt dann $\vec{E}=-\operatorname{grad}(\varphi)$. Das Gradientenfeld zeigt, dass $\vec{E}$ von der positiven Ladung weg zeigt.

Die Richtungsableitung



Es sei $T:\ z=f_x(x_0,y_0)\cdot(x-x_0)+f_y(x_0,y_0)\cdot(y-y_0)+f(x_0,y_0)$ die Gleichung der Tangentialebene an das Schaubild $z=f(x,y)$ an der Stelle $(x_0\,|\,y_0)$ bzw. im Punkt $P_0(x_0\,|\,y_0\,|\,z_0)$, wobei $z_0=f(x_0,y_0)$. Alle Tangenten an das Schaubild $z=f(x,y)$ im Punkt P_0 liegen in T.

Wir betrachten nun diejenige Tangente $t=P_0P$, die durch die zweidimensionale Richtung $\vec{a}=\begin{pmatrix}a_x\\a_y\end{pmatrix}\neq\vec{0}$ gegeben ist; siehe Skizze.

Ihre Steigung wird symbolisch mit $\frac{\partial f}{\partial\vec{a}}(x_0,y_0)$ bezeichnet. Für sie gilt also $\boxed{\frac{\partial f}{\partial\vec{a}}(x_0,y_0)=\tan\alpha}$.

Durch Umformung der Gleichung von T folgt $z-f(x_0,y_0)=f_x(x_0,y_0)\cdot(x-x_0)+f_y(x_0,y_0)\cdot(y-y_0)$.
Daraus ergibt sich für die kleine Änderung dz von z zu $dz=f_x(x_0,y_0)\cdot dx+f_y(x_0,y_0)\cdot dy$, so dass

$$\frac{\partial f}{\partial\vec{a}}(x_0,y_0)=\tan\alpha=\frac{dz}{\sqrt{(dx)^2+(dy)^2}}=\frac{f_x(x_0,y_0)\cdot dx+f_y(x_0,y_0)\cdot dy}{\sqrt{(dx)^2+(dy)^2}}=f_x(x_0,y_0)\cdot\frac{dx}{\sqrt{(dx)^2+(dy)^2}}+f_y(x_0,y_0)\cdot\frac{dy}{\sqrt{(dx)^2+(dy)^2}}=$$

$$=f_x(x_0,y_0)\cdot\frac{a_x}{|\vec{a}|}+f_y(x_0,y_0)\cdot\frac{a_y}{|\vec{a}|}=\frac{f_x(x_0,y_0)\cdot a_x+f_y(x_0,y_0)\cdot a_y}{|\vec{a}|}=\frac{1}{|\vec{a}|}\operatorname{grad}(f)(x_0,y_0)\bullet\vec{a}\,.$$

$$\boxed{\frac{\partial f}{\partial\vec{a}}(x_0,y_0)=\tan\alpha=\frac{1}{\sqrt{a_x^2+a_y^2}}\left(f_x(x_0,y_0)\cdot a_x+f_y(x_0,y_0)\cdot a_y\right)=\frac{1}{|\vec{a}|}\operatorname{grad}(f)(x_0,y_0)\bullet\vec{a}}$$

Speziell: Mit $\vec{a}=\begin{pmatrix}1\\0\end{pmatrix}$ folgt $\frac{\partial f}{\partial\vec{a}}=\frac{1}{|\vec{a}|}\operatorname{grad}(f)\bullet\vec{a}=\frac{1}{1}\begin{pmatrix}f_x(x,y)\\f_y(x,y)\end{pmatrix}\bullet\begin{pmatrix}1\\0\end{pmatrix}=f_x(x,y)$, so dass $f_x(x,y)$ wie erwartet die partielle Ableitung von $f(x,y)$ in x-Richtung ist. Analog ist für $\vec{a}=\begin{pmatrix}0\\1\end{pmatrix}$ die Richtungsableitung $\frac{\partial f}{\partial\vec{a}}=f_y(x,y)$.

Zusatz 1: Wir wissen, dass der Gradientenvektor $\operatorname{grad}(f)=\begin{pmatrix} f_x(x,y) \\ f_y(x,y) \end{pmatrix}$ senkrecht auf den Höhenlinien steht und in Richtung des maximalen Zuwachses von f zeigt. Jetzt können wir eine Formel für den maximalen Steigungswinkel α_{max} in Richtung von $\vec{a}=\operatorname{grad}(f)$ finden:

$$\frac{\partial f}{\partial \vec{a}}=\tan\alpha_{max}=\frac{1}{|\vec{a}|}\operatorname{grad}(f)(x_0,y_0)\bullet\vec{a}=\frac{1}{|\operatorname{grad}(f)(x_0,y_0)|}\operatorname{grad}(f)(x_0,y_0)\bullet\operatorname{grad}(f)(x_0,y_0)=\frac{f_x^{\,2}(x_0,y_0)+f_y^{\,2}(x_0,y_0)}{\sqrt{f_x^{\,2}(x_0,y_0)+f_y^{\,2}(x_0,y_0)}},$$

so dass $\boxed{\tan\alpha_{max}=\sqrt{f_x^{\,2}(x_0,y_0)+f_y^{\,2}(x_0,y_0)}=\left|\operatorname{grad}(f)(x_0,y_0)\right|}$.

Zusatz 2: Eine Tangente t im Punkt $P_0(x_0\,|\,y_0\,|\,z_0)$ sei durch ihre zweidimensionale Richtung $\vec{a}=\begin{pmatrix} a_x \\ a_y \end{pmatrix}\neq\vec{0}$ gegeben. Wie lautet ihr dreidimensionaler Richtungsvektor $\vec{u}_t=\begin{pmatrix} a_x \\ a_y \\ u_z \end{pmatrix}$?

Die Tangentialebene $T:\ z=f_x(x_0,y_0)\cdot(x-x_0)+f_y(x_0,y_0)\cdot(y-y_0)+f(x_0,y_0)$ besitzt den Normalenvektor $\boxed{\vec{n}_T=\begin{pmatrix} f_x(x_0,y_0) \\ f_y(x_0,y_0) \\ -1 \end{pmatrix}}$. Da die Tangente t in der Tangentialebene T liegt, ist $\vec{n}_T$ orthogonal zu $\vec{u}_t$, d.h. ihr Skalarprodukt muss Null sein. Aus $\boxed{0=\vec{n}_T\bullet\vec{u}_t=\begin{pmatrix} f_x(x_0,y_0) \\ f_y(x_0,y_0) \\ -1 \end{pmatrix}\bullet\begin{pmatrix} a_x \\ a_y \\ u_z \end{pmatrix}}$ folgt $u_z=f_x(x_0,y_0)\cdot a_x+f_y(x_0,y_0)\cdot a_y$.

Somit haben wir einen Richtungsvektor der Tangente t im Punkt $P_0(x_0\,|\,y_0\,|\,f(x_0,y_0))$ in Richtung $\vec{a}=\begin{pmatrix} a_x \\ a_y \end{pmatrix}$:

$$\vec{u}_t=\begin{pmatrix} a_x \\ a_y \\ f_x(x_0,y_0)\cdot a_x+f_y(x_0,y_0)\cdot a_y \end{pmatrix}=\begin{pmatrix} a_x \\ a_y \\ \operatorname{grad}(f)(x_0,y_0)\bullet\vec{a} \end{pmatrix}$$

Eine Gleichung der Tangente ist dann: $t:\ \vec{x}=\begin{pmatrix} x_0 \\ y_0 \\ f(x_0,y_0) \end{pmatrix}+r\cdot\begin{pmatrix} a_x \\ a_y \\ \operatorname{grad}(f)(x_0,y_0)\bullet\vec{a} \end{pmatrix}$ mit $r\in\mathbb{R}$.

Beispiel 1: Es sei $z=f(x,y)=x^2+y^2$. Es ist $\operatorname{grad}(f)=\begin{pmatrix} 2x \\ 2y \end{pmatrix}$.

Gesucht ist die Richtungsableitung an der Stelle $(1|2)$ bzw. im Punkt $P_0(1|2|5)$ in Richtung $\vec{a}=\begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

$\tan\alpha=\frac{\partial f}{\partial\vec{a}}=\frac{1}{\sqrt{a_x^2+a_y^2}}\left(\frac{\partial f}{\partial x}\cdot a_x+\frac{\partial f}{\partial y}\cdot a_y\right)=\frac{1}{5}(2\cdot3+4\cdot4)=\frac{22}{5}$. Es folgt $\alpha=77{,}20°$.

Die Tangentialebene im Punkt P_0 lautet $T:\ z=2\cdot(x-1)+4\cdot(y-2)+5$, bzw. $T:\ 2x+4y-z-5=0$ mit dem Normalenvektor $\vec{n}=\begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}$. Die Tangente t in Richtung $\vec{a}=\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ hat den Richtungsvektor $\vec{u}=\begin{pmatrix} 3 \\ 4 \\ u_z \end{pmatrix}$ mit $\vec{n}\bullet\vec{u}=0$. Es folgt $2\cdot3+4\cdot4-u_z=0$, also $\vec{u}=\begin{pmatrix} 3 \\ 4 \\ 22 \end{pmatrix}$ und die Gleichung $t:\ \vec{x}=\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}+r\cdot\begin{pmatrix} 3 \\ 4 \\ 22 \end{pmatrix}$, $r\in\mathbb{R}$.

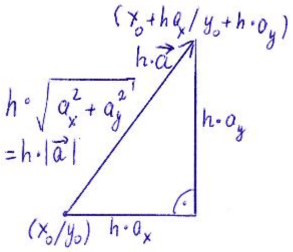
Der steilste Anstieg erfolgt in Richtung $\operatorname{grad}(f)(1/2)=\begin{pmatrix}2\\4\end{pmatrix}$. Für diesen maximalen Steigungswinkel α_{max} gilt

$$\tan\alpha_{max}=\frac{1}{\sqrt{2^2+4^2}}\operatorname{grad}(f)(x_0,y_0)\bullet\begin{pmatrix}2\\4\end{pmatrix}=\frac{1}{\sqrt{20}}\begin{pmatrix}2\\4\end{pmatrix}\bullet\begin{pmatrix}2\\4\end{pmatrix}=\frac{1}{\sqrt{20}}20=\sqrt{20}$$, so dass $\alpha_{max}=77{,}40°$.

Zusatz: Die Richtungsableitung $\frac{\partial f}{\partial \vec{a}}$ kann auch mit Hilfe einer Grenzwertbetrachtung bestimmt werden.

Für $y=f(x)$ gilt $f'(x_0)=\lim\limits_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}$, indem man sich von der Stelle x_0 zur Stelle x_0+h bewegt.

Analog bewegt man sich bei der Richtungsableitung von der Stelle (x_0/y_0) in Richtung $\vec{a}$ um den Vektor $h\cdot\vec{a}=\begin{pmatrix}h\,a_x\\h\,a_y\end{pmatrix}$ zur Stelle $(x_0+h\cdot a_x/y_0+h\cdot a_y)$. Dann



folgt $$\boxed{\frac{\partial f}{\partial \vec{a}}(x_0,y_0)=\lim_{h\to 0}\frac{f(x_0+h\cdot a_x\,,\,y_0+h\cdot a_y)-f(x_0,y_0)}{h\cdot|\vec{a}|}}$$ mit $\vec{a}=\begin{pmatrix}a_x\\a_y\end{pmatrix}$.

Möglich ist auch $$\boxed{\frac{\partial f}{\partial \vec{a}}(x_0,y_0)=\lim_{h\to 0}\frac{f(x_0+h\cdot\frac{a_x}{|\vec{a}|},\,y_0+h\cdot\frac{a_y}{|\vec{a}|})-f(x_0,y_0)}{h}}$$

Zurück zum Beispiel 1 mit $z=f(x,y)=x^2+y^2$ und $\vec{a}=\begin{pmatrix}3\\4\end{pmatrix}$.

$$\frac{\partial f}{\partial \vec{a}}(x_0,y_0)=\lim_{h\to 0}\frac{f(x_0+h\cdot 3\,,\,y_0+h\cdot 4)-f(x_0,y_0)}{h\cdot\sqrt{3^2+4^2}}=\lim_{h\to 0}\frac{(x_0+h\cdot 3)^2+(y_0+h\cdot 4)^2-(x_0^2+y_0^2)}{5h}=$$

$$=\lim_{h\to 0}\frac{6x_0h+8y_0h+25h^2}{5h}=\lim_{h\to 0}\frac{6x_0+8y_0+25h}{5}=\frac{6}{5}x_0+\frac{8}{5}y_0\,.$$

Für $P_0(1|2|5)$ von oben folgt $\frac{\partial f}{\partial \vec{a}}(x_0,y_0)=\frac{6}{5}x_0+\frac{8}{5}y_0=\frac{6}{5}\cdot 1+\frac{8}{5}\cdot 2=\frac{22}{5}$ wie oben.

Beispiel 2: Es sei $z=f(x,y)=\frac{x}{y^2+1}$ und $P_0(2/1/1)$. Bestimmen Sie für den Punkt P_0 eine Gleichung der Tangentialebene T, einen Normalenvektor $\vec{n}$ von T, die Ableitung in Richtung $\vec{a}=\begin{pmatrix}4\\-3\end{pmatrix}$, den dazugehörigen Neigungswinkel α und den zu $\vec{a}$ gehörigen dreidimensionalen Vektor $\vec{u}=\begin{pmatrix}4\\-3\\u_z\end{pmatrix}$.

Mit $f_x(x,y)=\frac{1}{y^2+1}$ und $f_y(x,y)=-\frac{2xy}{\left(y^2+1\right)^2}$ folgt $f_x(2/1)=\frac{1}{2}$ und $f_y(2/1)=-1$.

Dann gilt $T: z=\frac{1}{2}(x-2)-1\cdot(y-1)+1$, d.h. $T:\ z=\frac{1}{2}x-y+1$ mit $\vec{n}=\begin{pmatrix}f_x(x_0,y_0)\\f_y(x_0,y_0)\\-1\end{pmatrix}=\begin{pmatrix}1/2\\-1\\-1\end{pmatrix}$.

$\frac{\partial f}{\partial \vec{a}}(2/1)=\frac{1}{|\vec{a}|}\operatorname{grad}(f)(x_0,y_0)\bullet\vec{a}=\frac{1}{5}\begin{pmatrix}1/2\\-1\end{pmatrix}\bullet\begin{pmatrix}4\\-3\end{pmatrix}=1$. Aus $\tan(\alpha)=1$ ergibt sich $\alpha=45°$.

Aus $0=\vec{n}\bullet\vec{u}=\begin{pmatrix}1/2\\-1\\-1\end{pmatrix}\bullet\begin{pmatrix}4\\-3\\u_z\end{pmatrix}$ folgt $u_z=5$, so dass $\vec{u}=\begin{pmatrix}4\\-3\\5\end{pmatrix}$.

Zusatz: Die Richtungsableitung erhält man auch aus $\vec{u}$ durch $\frac{\partial f}{\partial \vec{a}}(2/1)=\frac{u_z}{\sqrt{u_x^2+u_y^2}}=\frac{5}{\sqrt{4^2+(-3)^2}}=\frac{5}{5}=1$.

Beispiel 3: Es sei $u = f(x,y,z) = x^2 \cdot e^{2y} + x \cdot z^3$. Bestimmen Sie die Ableitung in Richtung $\vec{a} = \begin{pmatrix} 3 \\ -4 \\ 12 \end{pmatrix}$ an der Stelle $(1/0/-1)$ bzw. im Punkt $P(1/0/-1/0)$.

Mit $f_x(x,y,z) = 2x \cdot e^{2y} + z^3$, $f_y(x,y,z) = 2x^2 \cdot e^{2y}$ und $f_z(x,y,z) = 3x \cdot z^2$ folgt

$$\frac{\partial f}{\partial \vec{a}}(1/0/-1) = \frac{1}{|\vec{a}|}\operatorname{grad}(f)(1/0/-1) \bullet \vec{a} = \frac{1}{13}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \bullet \begin{pmatrix} 3 \\ -4 \\ 12 \end{pmatrix} = \frac{31}{13}.$$

Die Elastizität einer Funktion

Es sei $f(x)$ eine beliebige Funktion. Wenn sich das Argument von x zu $x + \Delta x$ ändert, so ändert sich f von $f(x)$ zu $f(x + \Delta x)$.

Δx und $\Delta f = f(x + \Delta x) - f(x)$ sind die **absoluten** Änderungen, $\frac{\Delta x}{x}$ und $\frac{\Delta f}{f}$ sind die dazugehörigen **relativen** Änderungen.

Der Quotient $\dfrac{\frac{\Delta f}{f}}{\frac{\Delta x}{x}} = \dfrac{\Delta f}{\Delta x} \cdot \dfrac{x}{f}$ heißt die **durchschnittliche Elastizität** von f im Intervall $[x / x + \Delta x]$.

Beispiel: Die durchschnittliche Elastizität einer Funktion f sei 2. Dann gilt $\dfrac{\frac{\Delta f}{f}}{\frac{\Delta x}{x}} = \dfrac{\Delta f}{\Delta x} \cdot \dfrac{x}{f} = 2 = \dfrac{2}{1} = \dfrac{2\%}{1\%}$.

Die durchschnittliche Elastizität gibt an, um wie viel Prozent sich $f(x)$ ändert, wenn x um 1% zunimmt.

Oder auch: $\dfrac{\frac{\Delta f}{f}}{\frac{\Delta x}{x}} = \dfrac{\Delta f}{\Delta x} \cdot \dfrac{x}{f} = 2 = \dfrac{2}{1} = \dfrac{2\%}{1\%} = \dfrac{1\%}{0{,}5\%} = \dfrac{8\%}{4\%} = \ldots$. Also gilt:

Die durchschnittliche Elastizität $\varepsilon_{f,x}(x)$ gibt das Verhältnis der relativen (prozentualen) Änderung von $f(x)$ zur relativen (prozentualen) Änderung von x an. Oder: Wenn x um 1% zunimmt dann ändert sich $f(x)$ um etwa $\varepsilon_{f,x}(x)\%$.

Meist verwendet man den Grenzwert für $\Delta x \to 0$. Dann wird $\lim\limits_{\Delta x \to 0} \dfrac{\Delta f}{\Delta x} \cdot \dfrac{x}{f} = f'(x) \cdot \dfrac{x}{f(x)}$.

Definition: Die Funktion $y = f(x)$ sei differenzierbar, und es sei $x \neq 0$ und $f(x) \neq 0$. Dann heißt

$\boxed{\varepsilon_f(x) = f'(x) \cdot \dfrac{x}{f(x)}}$ die **Elastizität** bzw. **Punktelastizität** von $f(x)$ bezüglich x.

Beispiel 1: Es sei $f(x) = a \cdot x^n$ mit $a, n, x \neq 0$. Dann wird $\varepsilon_f(x) = a \cdot n \cdot x^{n-1} \cdot \dfrac{x}{a \cdot x^n} = n$, also konstant.

Wenn nun x um 1% zunimmt, so dass $\Delta x = 0{,}01x$ wird, dann gilt für die durchschnittliche Elastizität:

$$\frac{\frac{\Delta f}{f}}{\frac{\Delta x}{x}} = \frac{\Delta f}{\Delta x} \cdot \frac{x}{f} = \frac{f(x + \Delta x) - f(x)}{\Delta x} \cdot \frac{x}{f(x)} = \frac{a \cdot (1{,}01x)^n - a \cdot x^n}{0{,}01x} \cdot \frac{x}{a \cdot x^n} = \frac{1{,}01^n - 1}{0{,}01} = 100 \cdot \left(1{,}01^n - 1\right).$$

n	2	1	0,5	0	$-0{,}5$	-1	-2
$\frac{\Delta f}{\Delta x} \cdot \frac{x}{f}$	2,01	1	0,499	0	$-0{,}496$	$-0{,}99$	$-1{,}97$

Wert der Elastizität	Eigenschaft der Funktion
$\varepsilon = 0$	vollkommen unelastisch
$\lvert\varepsilon\rvert < 1$	unelastisch
$\lvert\varepsilon\rvert = 1$	Grenzbereich
$\lvert\varepsilon\rvert > 1$	elastisch

Bei der vollkommen unelastischen Funktion ist $f(x)$ konstant. Je mehr sich aber $\Delta f/f$ zusammen mit $\Delta x/x$ ändert, desto elastischer ist die Funktion f.

Beispiel 2: Es sei $f(x) = \ln x$. Dann folgt $\varepsilon(x) = f'(x)\cdot\frac{x}{f(x)} = \frac{1}{x}\cdot\frac{x}{\ln x} = \frac{1}{\ln x}$ für $x > 0$ und $x \neq 1$.

Z.B. Für $x = 2$ folgt $\varepsilon_f(2) = \frac{1}{\ln 2} \approx \frac{1}{0{,}6931} \approx 1{,}44$. Das heißt, dass das Verhältnis der relativen Änderung von f zur relativen Änderung von x in der Nähe von 2 etwa 1,44 beträgt. Oder: Wenn x in der Nähe von 2 um 1% zunimmt, dann nimmt $f(x)$ um etwa $1{,}44\%$ zu.

Probe mit $\frac{\Delta x}{x} = 1\%$: $$\frac{\frac{\Delta f}{f}}{\frac{\Delta x}{x}} = \frac{\frac{f(2{,}02)-f(2)}{f(2)}}{\frac{2{,}02-2}{2}} = \frac{\frac{\ln(2{,}02)-\ln(2)}{f(2)}}{\frac{0{,}02}{2}} = \frac{0{,}0144}{0{,}01} = \frac{1{,}44\%}{1\%} = 1{,}44$$

Probe mit $\frac{\Delta x}{x} = 0{,}5\%$: $$\frac{\frac{\Delta f}{f}}{\frac{\Delta x}{x}} = \frac{\frac{f(2{,}01)-f(2)}{f(2)}}{\frac{2{,}01-2}{2}} = \frac{\frac{\ln(2{,}01)-\ln(2)}{f(2)}}{\frac{0{,}01}{2}} \approx \frac{0{,}0071922}{0{,}005} = \frac{0{,}71922\%}{0{,}5\%} \approx 1{,}44.$$

Und nun zu Funktionen $f(x,y)$ mit zwei Variablen.

Das totale Differenzial von f beträgt $df = \frac{\partial f}{\partial x}\cdot dx + \frac{\partial f}{\partial y}\cdot dy = f_x(x,y)\cdot dx + f_y(x,y)\cdot dy$.

Daraus ergibt sich die relative Änderung von f zu

$$\frac{df}{f(x,y)} = \frac{f_x(x,y)\cdot dx + f_y(x,y)\cdot dy}{f(x,y)} = f_x(x,y)\cdot\frac{x}{f(x,y)}\cdot\frac{dx}{x} + f_y(x,y)\cdot\frac{y}{f(x,y)}\cdot\frac{dy}{y} = \varepsilon_{f,x}(x,y)\cdot\frac{dx}{x} + \varepsilon_{f,y}(x,y)\cdot\frac{dy}{y}.$$

Als gilt $\boxed{\frac{df}{f(x,y)} = \varepsilon_{f,x}(x,y)\cdot\frac{dx}{x} + \varepsilon_{f,y}(x,y)\cdot\frac{dy}{y}}$ mit den **partiellen Elastizitäten**

$$\boxed{\varepsilon_{f,x}(x,y) = f_x(x,y)\cdot\frac{x}{f(x,y)}} \quad \text{und} \quad \boxed{\varepsilon_{f,y}(x,y) = f_y(x,y)\cdot\frac{y}{f(x,y)}}.$$

Beispiel 1: Es sei $z = f(x,y) = a\cdot x^r\cdot y^s$. Dann folgt $\varepsilon_{f,x} = f_x(x,y)\cdot\frac{x}{f(x,y)} = r\cdot a\cdot x^{r-1}\cdot y^s\cdot\frac{x}{a\cdot x^r\cdot y^s} = r$ und

$\varepsilon_{f,y} = f_y(x,y)\cdot\frac{y}{f(x,y)} = s\cdot a\cdot x^r\cdot y^{s-1}\cdot\frac{y}{a\cdot x^r\cdot y^s} = s$, so dass $\frac{df}{f(x,y)} = r\cdot\frac{dx}{x} + s\cdot\frac{dy}{y}$.

Interpretation: Ändert sich z.B. x um p% und y um q%, so ändert sich f um ungefähr $(r\cdot p + s\cdot q)\%$.

Beispiel 2: $f(x,y) = 2x^2 + 4y^3$ und $P(2/1/12)$. x soll um 1% zunehmen und y um 2% abnehmen. Dann gilt

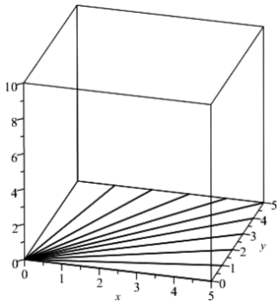
$\frac{\Delta f}{f} = \frac{f(2{,}02/0{,}98) - f(2/1)}{f(2/1)} = \frac{11{,}925568 - 12}{12} = -0{,}0062 = -0{,}62\%$. Und nun in der Näherung:

$\frac{df}{f(x,y)} = \varepsilon_{f,x}(x,y)\cdot\frac{dx}{x} + \varepsilon_{f,y}(x,y)\cdot\frac{dy}{y} = 4x\cdot\frac{x}{2x^2+4y^3}\cdot\frac{dx}{x} + 12y^2\cdot\frac{y}{2x^2+4y^3}\cdot\frac{dy}{y} = \frac{4x^2}{f(x,y)}\cdot\frac{dx}{x} + \frac{12y^3}{2x^2+4y^3}\cdot\frac{dy}{y} =$
$\frac{4\cdot 2^2}{12}\cdot\frac{0,02}{2} + \frac{12\cdot 1^3}{12}\cdot\frac{-0,02}{1} = \frac{-0,08}{12} = -0,00\overline{6} \approx -0,67\%$.

Homogene Funktionen

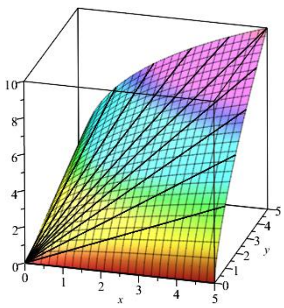
Definition: Eine Funktion $f:\mathbb{R}^2 \to \mathbb{R}$ heißt **homogen vom Grad λ** mit $\lambda \in \mathbb{R}^+$, wenn für alle $x,\ y,\ k \in \mathbb{R}$ die Beziehung $\boxed{f(k\cdot x,\ k\cdot y) = k^\lambda \cdot f(x,y)}$ gilt. Im Fall $\lambda = 1$ spricht man von linearer Homogenität.
In der linearen Algebra bedeutet die Homogenität einer linearen Abbildung $f(k\cdot\vec{a}) = k\cdot f(\vec{a})$.

Beispiel: Die Cobb-Douglas-Produktionsfunktion $f(x,y) = c\cdot x^\alpha \cdot y^\beta$ mit $0<\alpha,\ \beta<1$ und $x,\ y \ge 0$. Dabei stellt x den Kapitaleinsatz und y den Arbeitseinsatz dar, während $f(x,y)$ die Produktionsmenge angibt.
(Nach dem Mathematiker Charles Wiggins Cobb und dem Ökonomen Paul Howard Douglas, 1927)
Die Frage ist nun, wie sich die Produktionsmenge $f(x,y)$ verändert, wenn sowohl der Kapitaleinsatz x als auch der Arbeitseinsatz y ver-k-facht wird.
$f(k\cdot x,\ k\cdot y) = c\cdot(k\cdot x)^\alpha\cdot(k\cdot y)^\beta = k^{\alpha+\beta}\cdot c\cdot x^\alpha\cdot y^\beta = k^{\alpha+\beta}\cdot f(x,y)$, d.h. diese Produktionsfunktion f ist homogen vom Grad $\lambda = \alpha+\beta$. Speziell für $\alpha+\beta = 1$ ist f sogar linear homogen.



Zusatz: Es sei f eine homogene Funktion vom Grad λ. Was lässt sich über die Form der Bilder $z = f(x,y)$ der Ursprungsgeraden $\begin{pmatrix}x\\y\end{pmatrix} = r\cdot\begin{pmatrix}u\\v\end{pmatrix}$ mit dem Richtungsvektor $\begin{pmatrix}u\\v\end{pmatrix} \ne \begin{pmatrix}0\\0\end{pmatrix}$ und $r\in\mathbb{R}$ aussagen?
Es gilt $z = f(r\cdot u, r\cdot v) = r^\lambda \cdot f(u,v) = C\cdot r^\lambda$ mit der Konstanten $C = f(u,v)$.

Satz: Eine homogene Funktion f vom Grad λ verhält sich über einer Ursprungsgeraden wie die Potenzfunktion $z = C\cdot r^\lambda$.



Beispiel: $f(x,y) = 2\cdot x^{1/3}\cdot y^{2/3}$ mit $x,y\in\mathbb{R}^+$ ist linear homogen, da $f(k\cdot x,\ k\cdot y) = 2\cdot(k\cdot x)^{1/3}\cdot(k\cdot y)^{2/3} = k^{1/3+2/3}\cdot 2\cdot x^{1/3}\cdot y^{2/3} = k^1\cdot f(x,y)$ für alle $k\in\mathbb{R}^+$.
Die Bilder von zweidimensionalen Ursprungsgeraden $\begin{pmatrix}x\\y\end{pmatrix} = r\cdot\begin{pmatrix}u\\v\end{pmatrix}$ sind Potenzfunktionen vom Grad 1, also wieder Geraden, aber im $\mathbb{R}^3$
$\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}r\cdot u\\r\cdot v\\f(r\cdot u, r\cdot v)\end{pmatrix} = \begin{pmatrix}r\cdot u\\r\cdot v\\r\cdot f(u,v)\end{pmatrix} = r\cdot\begin{pmatrix}u\\v\\f(u,v)\end{pmatrix}$ durch den Ursprung. Einige dieser Geraden sind in der Figur eingezeichnet. Sie liegen alle auf dem Schaubild von f.

Partielle Ableitungen höherer Ordnung

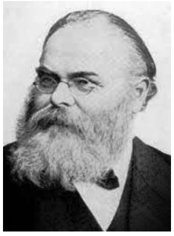
Beispiel: $f(x,y) = x\cdot\ln y^2$ für $x\in\mathbb{R}$ und $y\in\mathbb{R}\setminus\{0\}$. Dann ist

$\frac{\partial f}{\partial x}(x,y) = f_x(x,y) = \ln y^2$, $\frac{\partial f}{\partial y}(x,y) = f_y(x,y) = \frac{2x}{y}$, $\frac{\partial^2 f}{\partial x^2}(x,y) = f_{xx}(x,y) = 0$, $\frac{\partial^2 f}{\partial y^2}(x,y) = f_{yy}(x,y) = -\frac{2x}{y^2}$

$\frac{\partial^2 f}{\partial x\partial y}(x,y) = f_{xy}(x,y) = \frac{2}{y}$, zuerst nach x, dann nach y differenziert

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = f_{yx}(x,y) = \frac{2}{y}$$, zuerst nach y, dann nach x differenziert

$$\frac{\partial^3 f}{\partial x \partial y \partial y}(x,y) = f_{xyy}(x,y) = -\frac{2}{y^2}.$$

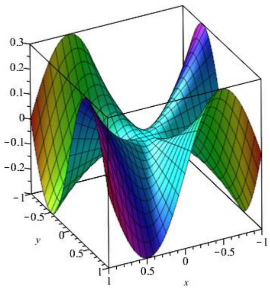


Man rechnet leicht nach, dass hier $f_{xyy}(x,y) = f_{yxy}(x,y) = f_{yyx}(x,y)$, d.h. dass die Reihenfolge der 3 Ableitungen keine Rolle spielt.

Satz von Schwarz (Hermann Amandus Schwarz, Deutscher Mathematiker, 1843-1921)
Es sei $k \geq 2$. Wenn alle k-fachen Ableitungsfunktionen stetig sind, dann darf man die Reihenfolge der k Ableitungen vertauschen, ohne dass sich das Ergebnis ändert.

Def.: Eine Funktion $z = f(x,y)$ ist **stetig** an der Stelle (x_0 / y_0), falls für beliebige Wege nach (x_0 / y_0) der Grenzwert $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$ immer gleich $f(x_0,y_0)$ ist, d.h. $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0,y_0)$.

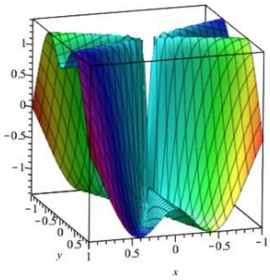
Bemerkung: Für den Beweis der Unstetigkeit muss man nur zwei Wege nach (x_0 / y_0) finden, die verschiedene Grenzwerte ergeben.



Ein **Beispiel**, bei dem die zweiten Ableitungen nicht stetig sind, ist

$$f(x,y) = \begin{cases} x \cdot y \cdot \dfrac{x^2 - y^2}{x^2 + y^2} & \text{für } (x,y) \neq (0,0) \\ 0 & \text{für } (x,y) = (0,0) \end{cases}$$. Siehe Schaubild rechts.

$f(x,y)$ ist stetig in ganz $\mathbb{R}^2$, siehe Hausaufgabe.



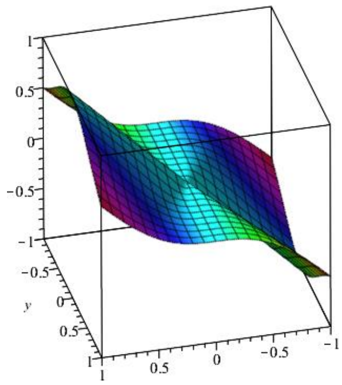
Für $(x,y) \neq (0,0)$ folgt $f_{xy}(x,y) = f_{yx}(x,y) = \dfrac{(x^2 - y^2)\cdot(x^4 + 10x^2 \cdot y^2 + y^4)}{(x^2 + y^2)^3}$.

Siehe Schaubild rechts.

Aber was ist $f_{xy}(0,0)$ bzw. $f_{yx}(0,0)$? Man kann sich der Stelle $(0,0)$ auf verschiedenen Wegen nähern, z.B. auf einer Geraden der Gleichung $y = m \cdot x$ mit $x \neq 0$.
Durch Einsetzen folgt

$$f_{xy}(x, m\cdot x) = f_{yx}(x, m\cdot x) = \frac{(1-m^2)\cdot(1+10m^2+m^4)}{(1+m^2)^3} .$$

Nähert sich (x,y) der Stelle $(0,0)$ auf der Geraden $y = x$, d.h. $m = 1$, so streben beide Funktionen gegen 0.
Nähert sich (x,y) der Stelle $(0,0)$ auf der Geraden $y = 0$, d.h. $m = 0$, so streben beide Funktionen gegen 1.
Somit sind die beiden Funktionen f_{xy} und f_{yx} an der Stelle $(0,0)$ nicht stetig.



Ein **Beispiel** für einen Stetigkeitsbeweis an der Stelle $(0,0)$:

$$f(x,y) = \begin{cases} \dfrac{x^3}{x^2 + y^2} & \text{für } (x,y) \neq (0,0) \\ 0 & \text{für } (x,y) = (0,0) \end{cases} .$$

Für $(x,y) \neq (0,0)$ gilt

$$|f(x,y) - f(0,0)| = \left| \frac{x^3}{x^2+y^2} - 0 \right| = \left| \frac{x^3}{x^2+y^2} \right| = |x| \cdot \underbrace{\frac{x^2}{x^2+y^2}}_{\leq 1} \leq |x| .$$ Da auf jedem Weg nach $(0,0)$ der Betrag $|x|$ gegen 0 strebt, muss auch $f(x,y)$ gegen 0 streben.

Extremwerte bei Funktionen mit zwei Variablen

Aus der Schule ist bekannt:
Notwendige Bedingung für einen Extremwert von $y = f(x)$ an der Stelle $x = x_0$**:** $f'(x_0) = 0$**.**

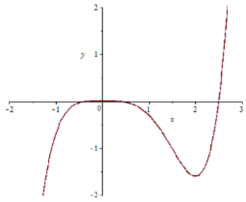
Hinreichende Bedingung 1:

$$\begin{cases} f'(x_0) = 0 \text{ und } f''(x_0) < 0 & \Rightarrow \text{ Das Schaubild } y = f(x) \text{ besitzt an der Stelle } x = x_0 \text{ einen Hochpunkt } H. \\ f'(x_0) = 0 \text{ und } f''(x_0) = 0 & \Rightarrow \text{ keine Entscheidung} \\ f'(x_0) = 0 \text{ und } f''(x_0) > 0 & \Rightarrow \text{ Das Schaubild } y = f(x) \text{ besitzt an der Stelle } x = x_0 \text{ einen Tiefpunkt } T. \end{cases}$$

Hinreichende Bedingung 2:

$$\begin{cases} f'(x_0) = 0 \text{ und } f'(x) \text{ wechselt sein Vorzeichen an der Stelle } x_0 \text{ von + nach} - & \Leftrightarrow \text{ Hochpunkt } H(x_0 \,|\, f(x_0)) \\ f'(x_0) = 0 \text{ und } f'(x) \text{ wechselt sein Vorzeichen an der Stelle } x_0 \text{ nicht} & \Leftrightarrow \text{ Sattelpunkt} \\ f'(x_0) = 0 \text{ und } f'(x) \text{ wechselt sein Vorzeichen an der Stelle } x_0 \text{ von − nach +} & \Leftrightarrow \text{ Tiefpunkt } T(x_0 \,|\, f(x_0)) \end{cases}$$

Beispiel: $f(x) = 0{,}2x^5 - 0{,}5x^4$.
Dann ist $f'(x) = x^4 - 2x^3$ und $f''(x) = 4x^3 - 6x^2$.
$f'(x) = 0$ hat die Lösungen $x_1 = 0$ und $x_2 = 2$.



Zur ersten hinreichenden Bedingung:
$f''(0) = 0$, also keine Entscheidung, ob Hoch- oder Tiefpunkt.
$f''(2) = 8 > 0$, also ist $T(2/-1{,}6)$ ein (relativer) Tiefpunkt.

Im Fall $x_1 = 0$ hilft die zweite hinreichende Bedingung: $f'(-1) = 3 > 0$ und $f'(+1) = -1 < 0$, so dass $H(0|0)$ (relativer) Hochpunkt ist.

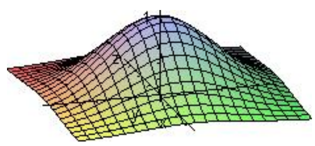
Im Fall $x_1 = 2$ hilft die zweite hinreichende Bedingung ebenfalls: $f'(1) = -1 < 0$ und $f'(3) = 27 > 0$, so dass $T(2/-1{,}6)$ ein (relativer) Tiefpunkt ist.

Und nun zu Funktionen mit zwei Variablen:

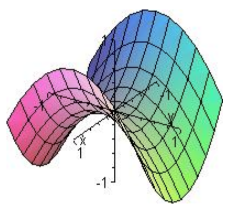
In einem Hochpunkt (x_0, y_0, z_0) oder Tiefpunkt (x_0, y_0, z_0) des Schaubildes von $z = f(x, y)$ muss die Tangentialebene parallel zur x-y-Ebene verlaufen, d.h. die beiden partiellen Ableitungen müssen 0 sein. Dies ist eine **Notwendige Bedingung** für einen Extremwert von $z = f(x, y)$ an der Stelle (x_0, y_0).

$$\boxed{P(x_0, y_0, z_0) \text{ ist ein Hochpunkt oder Tiefpunkt } \Rightarrow f_x(x_0, y_0) = 0 \text{ und } f_y(x_0, y_0) = 0}\ .$$

Beispiel 1: $z = f(x, y) = e^{-x^2 - y^2}$ hat einen Hochpunkt an der Stelle $(0,0)$.
Aus $f_x(x, y) = -2x \cdot e^{-x^2-y^2} = 0$ und $f_y(x, y) = -2y \cdot e^{-x^2-y^2} = 0$ folgt die Stelle $(0,0)$. Der Punkt $(0,0,1)$ ist ein Hochpunkt, wie man aus dem Schaubild erkennt.



Beispiel 2: $z = f(x, y) = x^2 - y^2$. Aus $f_x(x, y) = 2x = 0$ und $f_y(x, y) = -2y = 0$ folgt die Stelle $(0,0)$. Der Punkt $(0,0,0)$ ist aber weder ein Hoch- noch ein Tiefpunkt., denn:
Schnitt mit der x-z-Ebene, d.h. $y = 0$, ergibt die nach oben geöffnete Parabel $z = x^2$,
Schnitt mit der y-z-Ebene, d.h. $x = 0$, ergibt die nach unten geöffnete Parabel $z = -y^2$.



Definition: Eine Funktion $y = f(x_1, x_2, \ldots, x_n) = f(\vec{x})$ hat an der Stelle $\vec{x}_0 = (x_{01}, x_{02}, \ldots, x_{0n})$ ein **relatives (lokales) Maximum**, wenn für alle $\vec{x}$ in einer Umgebung von $\vec{x}_0$ die Beziehung $f(\vec{x}_0) > f(\vec{x})$ gilt, wobei $x \neq x_0$ ist.
Für das **relative (lokales) Minimum** gilt entsprechend die Beziehung $f(\vec{x}_0) < f(\vec{x})$.

<u>Notwendige Bedingung</u> für ein relatives Extremum von $y = f(\vec{x})$:

Wenn die Funktion $y = f(\vec{x})$ an der Stelle $\vec{x}_0$ nach allen Variablen partiell differenzierbar ist und in $\vec{x}_0$ ein relatives Extremum hat, dann gilt: $\boxed{f_{x_1}(\vec{x}_0) = 0,\ f_{x_2}(\vec{x}_0) = 0,\ \dots,\ f_{x_n}(\vec{x}_0) = 0}$, d.h. $\boxed{\operatorname{grad} f(\vec{x}_0) = \vec{0}}$.

Definition: Unter der **Hesse-Matrix** (Ludwig Otto Hesse, 1811 – 1874, deutscher Mathematiker) einer mindestens zweimal partiell differenzierbaren Funktion $y = f(\vec{x})$ versteht man die folgende Matrix

$$H(f) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix} = \begin{pmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_n} \\ f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_n} \\ \vdots & \vdots & & \vdots \\ f_{x_nx_1} & f_{x_nx_2} & \cdots & f_{x_nx_n} \end{pmatrix}.$$

Bemerkung: Nach dem Satz von Schwarz ist die Hesse-Matrix symmetrisch, wenn alle zweiten Ableitungen stetig sind.

Definition: Ein Vektor $\vec{x} \in \mathbb{R}^n$, $\vec{x} \neq \vec{0}$, $n \in \mathbb{N}$, heißt **Eigenvektor** einer quadratischen Matrix A vom Typ (n,n), wenn es ein $\lambda \in \mathbb{R}$ gibt mit $\boxed{A \cdot \vec{x} = \lambda \cdot \vec{x}}$. Die reelle Zahl λ heißt dann **Eigenwert** zum Eigenvektor $\vec{x}$.

Definition: Eine **symmetrische Matrix** A heißt
positiv definit, wenn A nur positive Eigenwerte besitzt,
positiv semidefinit, wenn alle Eigenwerte von A größer oder gleich 0 sind,
negativ definit, wenn $-A$ positiv definit ist, d.h. wenn A nur negative Eigenwerte besitzt,
negativ semidefinit, wenn alle Eigenwerte von A kleiner oder gleich 0 sind,
indefinit, wenn A positive und negative Eigenwerte besitzt.

Beispiel 1: Es sei $A = \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix}$. $A \cdot \vec{x} = \lambda \cdot \vec{x}$ mit $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ ergibt das LGS $\begin{array}{rcrcl} 6x_1 & + & 2x_2 & = & \lambda \cdot x_1 \\ 2x_1 & + & 3x_2 & = & \lambda \cdot x_2 \end{array}$, bzw. $\begin{array}{rcrcl} (6-\lambda) \cdot x_1 & + & 2x_2 & = & 0 \\ 2x_1 & + & (3-\lambda) \cdot x_2 & = & 0 \end{array}$. Dieses homogene LGS besitzt nur dann eine Lösung $\vec{x} \neq \vec{0}$, falls die Determinante $\begin{vmatrix} 6-\lambda & 2 \\ 2 & 3-\lambda \end{vmatrix} = (6-\lambda) \cdot (3-\lambda) - 4 = 0$ ist, d.h. falls $\lambda^2 - 9\lambda + 14 = 0$ gilt. Die Lösungen sind $\lambda_1 = 7$ und $\lambda_2 = 2$. Und damit ist die Matrix A positiv definit.

Beispiel 2: Es sei $B = -A = \begin{pmatrix} -6 & -2 \\ -2 & -3 \end{pmatrix}$. $B \cdot \vec{x} = \lambda \cdot \vec{x}$ mit $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ ergibt das LGS $\begin{array}{rcrcl} -6x_1 & - & 2x_2 & = & \lambda \cdot x_1 \\ -2x_1 & - & 3x_2 & = & \lambda \cdot x_2 \end{array}$, bzw. $\begin{array}{rcrcl} (-6-\lambda) \cdot x_1 & - & 2x_2 & = & 0 \\ -2x_1 & + & (-3-\lambda) \cdot x_2 & = & 0 \end{array}$. Dieses homogene LGS besitzt nur dann eine Lösung $\vec{x} \neq \vec{0}$, falls die Determinante $\begin{vmatrix} -6-\lambda & -2 \\ -2 & -3-\lambda \end{vmatrix} = (-6-\lambda) \cdot (-3-\lambda) - 4 = 0$ ist, d.h. falls $\lambda^2 + 9\lambda + 14 = 0$ gilt. Die Lösungen sind $\lambda_1 = -7$ und $\lambda_2 = -2$. Und damit ist die Matrix $B = -A$ negativ definit.

Bemerkung: Wenn die Matrix A positiv (negativ) definit ist, dann ist die Matrix $-A$ negativ (positiv) definit. Denn aus $A \cdot \vec{x} = \lambda \cdot \vec{x}$ folgt $(-A) \cdot \vec{x} = -\lambda \cdot \vec{x}$, so dass der Eigenwert sein Vorzeichen ändert.

Definition: Die **Hauptuntermatrizen** oder **Hauptminoren** einer Matrix $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$ sind

$$A_1 = (a_{11}),\quad A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},\quad A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},\ \ldots,\quad A_n = A .$$

Satz: Eine **symmetrische** Matrix A ist
positiv definit, falls für alle Hauptuntermatrizen $\det(A_i) > 0$, $i = 1, \ldots, n$, gilt,
negativ definit, falls für alle Hauptuntermatrizen $\det(-A_i) > 0$ oder gleichbedeutend
$(-1)^i \cdot \det(A_i) > 0$, $i = 1, \ldots, n$, gilt, d.h. $\begin{cases} \det(A_i) > 0 & \text{für gerades } i \\ \det(A_i) < 0 & \text{für ungerades } i \end{cases}$.

Beweis später für n = 2.

<u>Hinreichende Bedingung</u> **für ein relatives Extremum von $y = f(\vec{x})$:**

1. $\operatorname{grad}(f(\vec{x}_0)) = \vec{0}$ und $H(f(\vec{x}_0))$ ist positiv definit $\Rightarrow$ f hat an der Stelle $\vec{x}_0$ ein **relatives Minimum**,
2. $\operatorname{grad}(f(\vec{x}_0)) = \vec{0}$ und $H(f(\vec{x}_0))$ ist negativ definit $\Rightarrow$ f hat an der Stelle $\vec{x}_0$ ein **relatives Maximum**,
3. $\operatorname{grad}(f(\vec{x}_0)) = \vec{0}$ und $H(f(\vec{x}_0))$ ist indefinit $\Rightarrow$ f hat an der Stelle $\vec{x}_0$ einen **Sattelpunkt**.

Beweis später für n = 2.

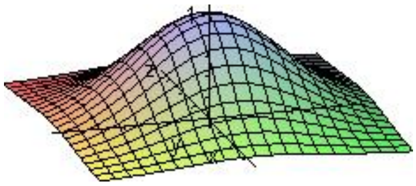
Spezialisierung auf n = 2: Es sei jetzt $z = f(x, y)$.
Die Hesse-Matrix ist $H(f(\vec{x}_0)) = \begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix}$. Sie ist

positiv definit, falls $\det(H_1) = f_{xx}(x_0, y_0) > 0$ und $\det(H_2) = f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2 > 0$,
negativ definit, falls $\det(H_1) = f_{xx}(x_0, y_0) < 0$ und $\det(H_2) = f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2 > 0$.

Also gilt:

1. Es gelte $f_x(x_0, y_0) = 0$, $f_y(x_0, y_0) = 0$ und $f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2 > 0$.
   a. wenn zusätzlich $f_{xx}(x_0, y_0) < 0$, dann hat $f(x, y)$ an der Stelle (x_0, y_0) ein **relatives Maximum**.
   b. wenn zusätzlich $f_{xx}(x_0, y_0) > 0$, dann hat $f(x, y)$ an der Stelle (x_0, y_0) ein **relatives Minimum**.
   $f_{xx}(x_0, y_0)$ und $f_{yy}(x_0, y_0)$ haben das gleiche Vorzeichen.
2. a. $f(x, y)$ habe an der Stelle (x_0, y_0) ein <u>relatives Maximum</u>. Dann gilt
   $f_x(x_0, y_0) = 0$, $f_y(x_0, y_0) = 0$, $f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2 \ge 0$ und
   $f_{xx}(x_0, y_0) \le 0$, $f_{yy}(x_0, y_0) \le 0$
   b. $f(x, y)$ habe an der Stelle (x_0, y_0) ein <u>relatives Minimum</u>. Dann gilt
   $f_x(x_0, y_0) = 0$, $f_y(x_0, y_0) = 0$, $f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2 \ge 0$ und
   $f_{xx}(x_0, y_0) \ge 0$, $f_{yy}(x_0, y_0) \ge 0$
3. Wenn $f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2 < 0$, dann hat $f(x, y)$ an der Stelle (x_0, y_0) <u>kein</u> relatives Extremum, sondern einen **Sattelpunkt**.
4. Wenn $f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2 = 0$, dann **keine Entscheidung**.

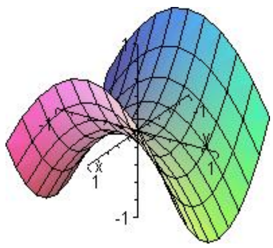
Beispiel 1: $z = f(x,y) = e^{-x^2-y^2}$ mit $(x,y) \in \mathbb{R}^2$. Aus $f_x(x,y) = -2x \cdot e^{-x^2-y^2} = 0$ und $f_y(x,y) = -2y \cdot e^{-x^2-y^2} = 0$ folgt $x_0 = y_0 = 0$. Außerdem ist $f_{xx}(x,y) = (4x^2-2) \cdot e^{-x^2-y^2}$,
$f_{xy}(x,y) = 4xy \cdot e^{-x^2-y^2}$ und $f_{yy}(x,y) = (4y^2-2) \cdot e^{-x^2-y^2}$.



Es folgt $f_{xx}(0,0) \cdot f_{yy}(0,0) - f_{xy}(0,0)^2 = (-2) \cdot (-2) - 0^2 = 4 > 0$. Wegen $f_{xx}(0,0) = -2 < 0$ folgt, dass $f(x,y)$ an der Stelle $(0,0)$ ein Maximum besitzt.

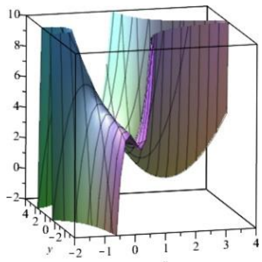
Zusatz: Die Hesse-Matrix ist $H(f(0,0)) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$. Ihre Eigenwerte folgen aus der charakteristischen Gleichung $\begin{vmatrix} -2-\lambda & 0 \\ 0 & -2-\lambda \end{vmatrix} = (-2-\lambda)^2 = 0$. Sie hat die Lösung $\lambda_1 = \lambda_2 = -2 < 0$, d.h. die Hesse-Matrix ist negativ definit, so dass $f(x,y)$ an der Stelle $(0,0)$ ein Maximum besitzt.

Beispiel 2: $z = f(x,y) = x^2 - y^2$ mit $(x,y) \in \mathbb{R}^2$. Aus $f_x(x,y) = 2x = 0$ und $f_y(x,y) = -2y = 0$ folgt die Stelle $(0,0)$. Außerdem ist $f_{xx}(x,y) = 2$, $f_{xy}(x,y) = 0$ und $f_{yy}(x,y) = -2$. Es folgt $f_{xx}(0,0) \cdot f_{yy}(0,0) - f_{xy}(0,0)^2 = 2 \cdot (-2) - 0^2 = -4 < 0$, also existiert kein relativer Extremwert, sondern ein Sattelpunkt.



Zusatz: Die Hesse-Matrix ist $H(f(0,0)) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$. Ihre Eigenwerte folgen aus der charakteristischen Gleichung $\det\begin{pmatrix} 2-\lambda & 0 \\ 0 & -2-\lambda \end{pmatrix} = (2-\lambda) \cdot (-2-\lambda) = 0$ zu $\lambda_1 = 2 > 0$ und $\lambda_2 = -2 < 0$, also kein Extremwer, sondern ein Sattelpunkt.

Beispiel 3: $z = f(x,y) = (2-x)^2 + x^3 \cdot y^2$ mit $(x,y) \in \mathbb{R}^2$. Es muss
$f_x(x,y) = -2(2-x) + 3x^2 \cdot y^2 = 0$ und $f_y(x,y) = 2x^3 \cdot y = 0$ sein.
Aus der zweiten Gleichung folgt $x = 0$ oder $y = 0$. $x = 0$ führt bei der ersten Gleichung zu einem Widerspruch. Mit $y = 0$ folgt $x = 2$, also die Stelle $(2|0)$.
Außerdem ist $f_{xx}(x,y) = 2 + 6x \cdot y^2$, $f_{xy}(x,y) = 6x^2 \cdot y$ und $f_{yy}(x,y) = 2x^3$.
Es folgt $f_{xx}(2|0) \cdot f_{yy}(2|0) - f_{xy}(2|0)^2 = 2 \cdot (2 \cdot 2^3) - 0^2 = 32 > 0$, also ist $T(2|0|0)$ ein relativer Tiefpunkt. T ist kein absoluter Tiefpunkt, da z.B. $P(-1|4|-7)$ tiefer liegt.



Zusatz: Die Hesse-Matrix ist $H(f(2,0)) = \begin{pmatrix} 2 & 0 \\ 0 & 16 \end{pmatrix}$. Ihre Eigenwerte folgen aus der charakteristischen Gleichung $\begin{vmatrix} 2-\lambda & 0 \\ 0 & 16-\lambda \end{vmatrix} = (2-\lambda)(16-\lambda) = 0$. Sie hat die Lösung $\lambda_1 = 2 > 0$ und $\lambda_2 = 16 > 0$, d.h. die Hesse-Matrix ist positiv definit, so dass $f(x,y)$ an der Stelle $(2|0)$ ein Minimum besitzt.

Beispiel 4: $u = f(x,y,z) = x + y + z + \sqrt{1-x^2-y^2-z^2}$ für $x^2+y^2+z^2 \le 1$.

Es ist $f_x(x,y,z) = 1 - \dfrac{x}{\sqrt{1-x^2-y^2-z^2}}$, $f_y(x,y,z) = 1 - \dfrac{y}{\sqrt{1-x^2-y^2-z^2}}$, $f_z(x,y,z) = 1 - \dfrac{z}{\sqrt{1-x^2-y^2-z^2}}$.

Aus $\nabla f(x,y,z) = \vec{0}$ folgt sofort $x = y = z > 0$ und durch Einsetzen das Ergebnis $x = y = z = \dfrac{1}{2}$. Außerdem ist

$$f_{xx}(x,y,z) = \frac{y^2+z^2-1}{(1-x^2-y^2-z^2)^{(3/2)}}, \quad f_{yy}(x,y,z) = \frac{x^2+z^2-1}{(1-x^2-y^2-z^2)^{(3/2)}}, \quad f_{zz}(x,y,z) = \frac{x^2+y^2-1}{(1-x^2-y^2-z^2)^{(3/2)}},$$

$$f_{xy}(x,y,z) = \frac{-xy}{(1-x^2-y^2-z^2)^{(3/2)}}, \quad f_{xz}(x,y,z) = \frac{-xz}{(1-x^2-y^2-z^2)^{(3/2)}}, \quad f_{yz}(x,y,z) = \frac{-yz}{(1-x^2-y^2-z^2)^{(3/2)}}.$$

Die Hesse-Matrix an der Stelle $\left(\frac{1}{2}\middle/\frac{1}{2}\middle/\frac{1}{2}\right)$ ist $H = \begin{pmatrix} -4 & -2 & -2 \\ -2 & -4 & -2 \\ -2 & -2 & -4 \end{pmatrix}$.

Die Determinanten der Hauptuntermatrizen sind $\det(H_1) = \det(-4) = -4$, $\det(H_2) = \det\begin{pmatrix} -4 & -2 \\ -2 & -4 \end{pmatrix} = 12$ und

$\det(H_3) = \det(H) = \det\begin{pmatrix} -4 & -2 & -2 \\ -2 & -4 & -2 \\ -2 & -2 & -4 \end{pmatrix} = -32$. Man erkennt, dass alle $(-1)^i \cdot \det(H_i) > 0$ sind, also ist H negativ definit, so dass f an der Stelle $\left(\frac{1}{2}\middle/\frac{1}{2}\middle/\frac{1}{2}\right)$ ein lokales Maximum hat, nämlich $f\left(\frac{1}{2}\middle/\frac{1}{2}\middle/\frac{1}{2}\right) = 2$.

Zusatz: Die charakteristische Gleichung der Hesse-Matrix ist

$\det\begin{pmatrix} -4-\lambda & -2 & -2 \\ -2 & -4-\lambda & -2 \\ -2 & -2 & -4-\lambda \end{pmatrix} = -\lambda^3 - 12\lambda^2 - 36\lambda - 32 = 0$ mit den Eigenwerten $\lambda_{1/2} = -2$ und $\lambda_3 = -8$, die alle negativ sind. Und damit ist die Hesse-Matrix negativ definit.

<u>Beweis obiger Sätze für den Fall n = 2</u>

Die Taylorreihe für eine Funktion $f(x_1, x_2, ..., x_n)$ mit n Variablen um den Punkt $(a_1, a_2, ..., a_n)$ lautet

$$f(x_1, x_2, \dots, x_n) = \sum_{i_1=0}^{\infty} \dots \sum_{i_n=0}^{\infty} \frac{(x_1 - a_1)^{i_1} \cdots (x_n - a_n)^{i_n}}{i_1! \cdots i_n!} \frac{\partial^{i_1}}{\partial x_1^{i_1}} \cdots \frac{\partial^{i_n}}{\partial x_n^{i_n}} f(a_1, ..., a_n) .$$

Speziell für $n = 2$ Variablen (x / y) um den Punkt (x_0 / y_0) bedeutet dies bis zur 2. Ordnung:

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + \\ &+ f_x(x_0,y_0)\cdot(x-x_0) + f_y(x_0,y_0)\cdot(y-y_0) + \\ &+ \frac{1}{2!} f_{xx}(x_0,y_0)\cdot(x-x_0)^2 + f_{xy}(x_0,y_0)\cdot(x-x_0)\cdot(y-y_0) + \frac{1}{2!} f_{yy}(x_0,y_0)\cdot(y-y_0)^2 = \\ &= f(x_0,y_0) + \begin{pmatrix} f_x(x_0,y_0) \\ f_y(x_0,y_0) \end{pmatrix} \bullet \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \bullet \left(\begin{pmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) \\ f_{xy}(x_0,y_0) & f_{yy}(x_0,y_0) \end{pmatrix} \cdot \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \right) . \end{aligned}$$

Der dickere Punkt bedeutet Skalarprodukt.

Falls f an der Stelle (x_0 / y_0) ein Minimum oder Maximum besitzt, dann muss die Tangentialebene horizontal verlaufen, so dass dann $f_x(x_0, y_0) = 0$ und $f_y(x_0, y_0) = 0$ sind.

Dann wird $f(x,y) = f(x_0,y_0) + \frac{1}{2}\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \bullet \left(H(x_0,y_0) \cdot \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \right)$ mit der Hesse-Matrix H.

Die Hesse-Matrix habe nun die Eigenwerte λ_1 und λ_2 mit den Eigenvektoren $\vec{u}_1$ und $\vec{u}_2$.
Diese Eigenvektoren sind orthogonal, da die Hesse-Matrix symmetrisch ist.

Wenn nun $\vec{u} = \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$ in Richtung eines Eigenvektors mit dem Eigenwert λ zeigt, dann gilt wegen $\vec{u}^2 = \vec{u} \bullet \vec{u} = |\vec{u}| \cdot |\vec{u}| \cdot \cos 0 = |\vec{u}|^2$ die Beziehung

$$f(x,y) = f(x_0,y_0) + \frac{1}{2}\vec{u} \bullet \left(H(x_0,y_0)\cdot\vec{u}\right) = f(x_0,y_0) + \frac{1}{2}\vec{u} \bullet \left(\lambda\cdot\vec{u}\right) = f(x_0,y_0) + \frac{1}{2}\lambda\vec{u}^2 = f(x_0,y_0) + \frac{1}{2}\lambda|\vec{u}|^2 .$$

Falls nun **beide Eigenwerte positiv** sind, dann nimmt f in den beiden orthogonalen Richtungen ihrer Eigenvektoren zu, so dass ein **Minimum** vorliegt.
Falls nun **beide Eigenwerte negativ** sind, dann nimmt f in den beiden orthogonalen Richtungen ihrer Eigenvektoren ab, so dass ein **Maximum** vorliegt.

Falls die **beiden Eigenwerte unterschiedliche Vorzeichen** besitzen, dann liegt ein **Sattelpunkt** vor.
Falls **ein Eigenwert Null** ist, dann lässt sich keine Aussage machen, denn in zweiter Näherung folgt $f(x,y)=f(x_0,y_0)+\frac{1}{2}\lambda|\vec{u}|^2=f(x_0,y_0)$. Dann geben erst die dritten Ableitungen Auskunft über Minimum, Maximum, Sattelpunkt.

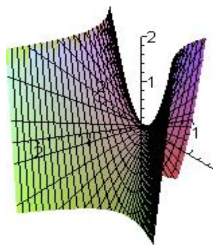
Und nun zur Hesse Matrix:
Wir wissen, dass die Determinante einer Matrix gerade das Produkt ihrer Eigenwerte ist. Deshalb gilt im Fall eines Maximums oder Minimums $\boxed{\det(H)=f_{xx}(x_0,y_0)\cdot f_{yy}(x_0,y_0)-f_{xy}(x_0,y_0)^2=\lambda_1\cdot\lambda_2>0}$.
Das Vorzeichen von $f_{xx}(x_0,y_0)$ entscheidet nun, ob ein Minimum oder ein Maximum vorliegt:

Fall 1: $f_{xx}(x_0,y_0)>0$
Dann muss zwangsläufig auch $f_{yy}(x_0,y_0)>0$ sein, damit $\det(H)>0$ sein kann.
Da die **Spur einer Matrix** gleich die Summe der Eigenwerte ist, gilt in diesem Fall $\mathrm{Spur}(H)=f_{xx}(x_0,y_0)+f_{yy}(x_0,y_0)=\lambda_1+\lambda_2>0$. Da im Fall eines Extremwertes beide Eigenwerte gleiches Vorzeichen besitzen, müssen beide Eigenwerte positiv sein, so dass ein Minimum vorliegt.

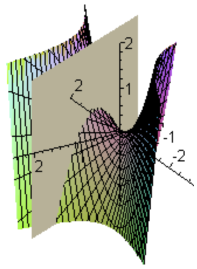
Fall 2: $f_{xx}(x_0,y_0)<0$
Dann muss zwangsläufig auch $f_{yy}(x_0,y_0)<0$ sein, damit $\det(H)>0$ sein kann.
Da die Spur einer Matrix gleich die Summe der Eigenwerte ist, gilt in diesem Fall $\mathrm{Spur}(H)=f_{xx}(x_0,y_0)+f_{yy}(x_0,y_0)=\lambda_1+\lambda_2<0$. Da im Fall eines Extremwertes beide Eigenwerte gleiches Vorzeichen besitzen, müssen beide Eigenwerte negativ sein, so dass ein Maximum vorliegt.

Fall 3: $f_{xx}(x_0,y_0)=0$
Dieser Fall kann nicht auftreten, da $\det(H)=f_{xx}(x_0,y_0)\cdot f_{yy}(x_0,y_0)-f_{xy}(x_0,y_0)^2=-f_{xy}(x_0,y_0)^2$ dann nicht mehr positiv sein könnte.

Extremwerte mit Nebenbedingungen

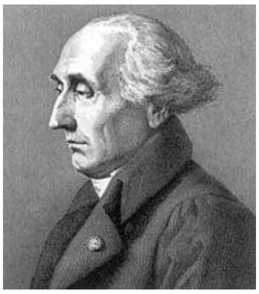


Beispiel 1: $z=f(x,y)=x^2+3xy$.
Es ist $f_x(x,y)=2x+3y$, $f_y(x,y)=3x$, $f_{xx}(x,y)=2$, $f_{xy}(x,y)=3$, $f_{yy}(x,y)=0$
Aus $\nabla f(x,y)=\vec{0}$ folgt sofort $x=y=0$.
Es folgt $f_{xx}(0,0)\cdot f_{yy}(0,0)-f_{xy}(0,0)^2=0^2-3^2=-9<0$, also existiert kein relativer Extremwert.
Die Hesse-Matrix im Punkt $(0/0)$ ist $\begin{pmatrix}2&3\\3&0\end{pmatrix}$. Aus der charakteristischen Gleichung $\lambda^2-2\lambda-9=0$ folgen die beiden Eigenwerte $\lambda_{1/2}=1\pm\sqrt{10}$ mit unterschiedlichen Vorzeichen. Und damit liegt kein Extremum vor.



Nun untersuchen wir die Extremwerte von $z=f(x,y)=x^2+3xy$ unter der **Nebenbedingung** $2x+y=1$, also unter einer Beziehung zwischen den beiden unabhängigen Variablen x und y.
Da sich diese Nebenbedingung nach einer Variablen auflösen lässt, bzw. schon aufgelöst ist, kann sie in f eingesetzt werden: $z=f(x,1-2x)=x^2+3x\cdot(1-2x)=3x-5x^2$.
Mit $z'=3-10x$ und $z''=-10<0$ folgt der Hochpunkt $H\left(0{,}3/0{,}4/0{,}45\right)$.
Im Schaubild erkennt man, dass bei dieser Aufgabe die Schnittlinie des Schaubildes von f mit der zur z-Achse parallelen Ebene $2x+y=1$, welche die Nebenbedingung darstellt, auf Extremwerte untersucht wird.

Problem: Wie lässt sich eine Aufgabe lösen, wenn man die Nebenbedingung nicht nach einer der Variablen auflösen kann?



Joseph-Louis de Lagrange, geboren 1736 als Lodovico Lagrangia in Turin, gestorben 1813 in Paris.

Er fand ein elegantes Verfahren.

Herleitung des Verfahrens von Lagrange:
Gegeben ist die Funktion $z = f(x, y)$, für welche die Extremwerte unter der Bedingung $u = \varphi(x, y) = 0$ gesucht werden.
Für die totalen Differenziale gilt $dz = f_x dx + f_y dy$ und $du = \varphi_x dx + \varphi_y dy$.
Da $u = 0$ konstant für alle x, y gilt, muss auch $du = \varphi_x(x, y)dx + \varphi_y(x, y)dy = 0$ für alle x, y sein.
Am Ort (x_0, y_0) des gesuchten Extremums ist auch $dz = f_x(x_0, y_0)dx + f_y(x_0, y_0)dy = 0$.

Man kann die beiden Gleichungen $\begin{cases} f_x dx + f_y dy = 0 \\ \varphi_x dx + \varphi_y dy = 0 \end{cases}$ als lineares homogenes Gleichungssystem für dx und dy interpretieren. Dieses System besitzt nur dann eine nichttriviale Lösung $(dx, dy) \neq (0, 0)$, falls die Determinante im Punkt (x_0, y_0) gleich $\begin{vmatrix} f_x & f_y \\ \varphi_x & \varphi_y \end{vmatrix} = 0$ ist. Die Determinante gibt den Flächeninhalt des von den beiden Zeilenvektoren oder den beiden Spaltenvektoren aufgespannten Parallelogramms an. Da in unserem Fall dieser Flächeninhalt gleich Null ist, müssen die Zeilenvektoren linear abhängig (Vielfache voneinander) sein:

$\begin{pmatrix} f_x \\ f_y \end{pmatrix} = -\lambda \cdot \begin{pmatrix} \varphi_x \\ \varphi_y \end{pmatrix}$, d.h. $\begin{pmatrix} f_x \\ f_y \end{pmatrix} + \lambda \cdot \begin{pmatrix} \varphi_x \\ \varphi_y \end{pmatrix} = \vec{0}$ im Punkt (x_0, y_0). Außerdem gilt ja $\varphi(x, y) = 0$.

Diese drei Gleichungen erhält man auch mit Hilfe der Hilfsfunktion $\boxed{L(x, y, \lambda) = f(x, y) + \lambda \cdot \varphi(x, y)}$. Sie wird nach x, nach y und nach λ partiell differenziert.

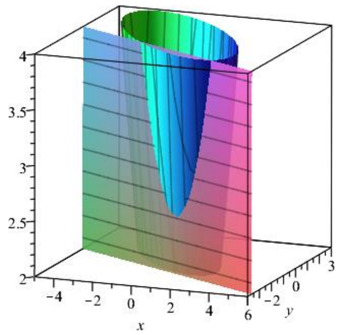
Multiplikatorregel von Lagrange: Die Extremwerte von $z = f(x, y)$ unter der Nebenbedingung $\varphi(x, y) = 0$ findet man auf folgende Weise:
Die Lagrangesche Hilfsfunktion $L(x, y, \lambda) = f(x, y) + \lambda \cdot \varphi(x, y)$ wird nach x, y und λ partiell differenziert und die Ergebnisse werden jeweils Null gesetzt.

$$\begin{cases} L_x(x, y, \lambda) = f_x(x, y) + \lambda \cdot \varphi_x(x, y) = 0 \\ L_y(x, y, \lambda) = f_y(x, y) + \lambda \cdot \varphi_y(x, y) = 0 \\ L_\lambda(x, y, \lambda) = \varphi(x, y) = 0 \end{cases}.$$

Aus diesen drei Gleichungen lassen sich die Koordinaten (x_0, y_0) der Extremwerte und der Lagrangesche Multiplikator λ bestimmen.

Bemerkung: Die Lagrangeschen Gleichungen sind **notwendig, aber nicht hinreichend** für die Existenz eines Extremwertes. Dieses Verfahren ist deshalb dann besonders nützlich, wenn die Existenz eines Extremwertes schon bekannt ist und nur noch die Berechnung fehlt.

Zurück zu unserem Beispiel: $z = f(x, y) = x^2 + 3xy$, Nebenbedingung $\varphi(x, y) = 2x + y - 1 = 0$. Dann wird $L(x, y, \lambda) = x^2 + 3xy + \lambda \cdot (2x + y - 1)$. Die drei Lagrangeschen Gleichungen $L_x(x, y, \lambda) = 2x + 3y + 2\lambda = 0$, $L_y(x, y, \lambda) = 3x + \lambda = 0$, $L_\lambda(x, y, \lambda) = 2x + y - 1 = 0$ haben die Lösungen $x_0 = 0{,}3$, $y_0 = 0{,}4$ und $\lambda = -0{,}9$.
Der Hochpunkt ist wie oben $H(0{,}3/0{,}4/0{,}45)$.

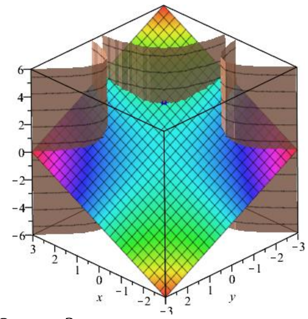


Beispiel 2: $z = f(x,y) = \frac{1}{2}x^2 + y^2 + x\cdot y$ hat unter der Nebenbedingung $\varphi(x,y) = x + 3y + 5 = 0$ einen Tiefpunkt. Bestimmen Sie seine Koordinaten.

Die Lagrange-Funktion lautet $L(x,y,\lambda) = \frac{1}{2}x^2 + y^2 + x\cdot y + \lambda\cdot(x+3y+5)$.

Die drei Lagrangeschen Gleichungen sind $L_x(x,y,\lambda) = x + y + \lambda = 0$, $L_y(x,y,\lambda) = 2y + x + 3\lambda = 0$, $L_\lambda(x,y,\lambda) = x + 3y + 5 = 0$.

Sie haben die Lösungen $x = 1$, $y = -2$ und $\lambda = 1$. Mit $f(1,-2) = 2,5$ folgt der Tiefpunkt $T(1/-2/2,5)$.



Beispiel 3: $z = f(x,y) = x - y$ hat unter der Nebenbedingung $\varphi(x,y) = x^2\cdot y^3 - x^3\cdot y^2 + 2 = 0$ einen Tiefpunkt. Bestimmen Sie seine Koordinaten.

Die Lagrange-Funktion lautet $L(x,y,\lambda) = x - y + \lambda\cdot(x^2\cdot y^3 - x^3\cdot y^2 + 2)$. Die drei Lagrangeschen Gleichungen sind $L_x(x,y,\lambda) = 1 + \lambda\cdot(2x\cdot y^3 - 3x^2\cdot y^2) = 0$, $L_y(x,y,\lambda) = -1 + \lambda\cdot(3x^2\cdot y^2 - 2x^3\cdot y) = 0$, $L_\lambda(x,y,\lambda) = x^2\cdot y^3 - x^3\cdot y^2 + 2 = 0$.

Mit Hilfe der beiden ersten Gleichungen $\lambda\cdot(2x\cdot y^3 - 3x^2\cdot y^2) = -1$ und $\lambda\cdot(3x^2\cdot y^2 - 2x^3\cdot y) = 1$ folgt $-\lambda\cdot(2x\cdot y^3 - 3x^2\cdot y^2) = \lambda\cdot(3x^2\cdot y^2 - 2x^3\cdot y)$, also $\lambda\cdot(-2x\cdot y^3 + 2x^3\cdot y) = 0$, d.h. $2\lambda\cdot x\cdot y\cdot(-y^2 + x^2) = 0$.

Wegen $L_x(x,y,\lambda) = 1 + \lambda\cdot(2x\cdot y^3 - 3x^2\cdot y^2) = 0$ ist sicher $\lambda \neq 0$, $x \neq 0$ und $y \neq 0$. Also muss $x^2 = y^2$ sein.

Fall1: $y = x$: $L_\lambda(x,y,\lambda) = x^2\cdot y^3 - x^3\cdot y^2 + 2 = x^5 - x^5 + 2 = 0$ ist nicht möglich.

Fall2: $y = -x$: $L_\lambda(x,y,\lambda) = x^2\cdot y^3 - x^3\cdot y^2 + 2 = -x^5 - x^5 + 2 = 0$, d.h. $x^5 = 1$, also $x = 1$.

Im Schaubild erkennt man, dass es sich um einen Tiefpunkt handelt. Es ist $T(1/-1/2)$.

Beispiel 4: $u = f(x,y,z) = x^2 + y^2 + z^2$ sei extremal unter der Bedingung $\varphi(x,y,z) = x + 4y + 8z - 81 = 0$.

Die Lagrangesche Hilfsfunktion lautet $L(x,y,z,\lambda) = x^2 + y^2 + z^2 + \lambda\cdot(x + 4y + 8z - 81)$.

Durch partielles Ableiten von $L(x,y,z,\lambda)$ nach x, y, z und λ folgt:

$$\begin{cases} L_x(x,y,z,\lambda) &= 2x + \lambda = 0 \\ L_y(x,y,z,\lambda) &= 2y + 4\lambda = 0 \\ L_z(x,y,z,\lambda) &= 2z + 8\lambda = 0 \\ L_\lambda(x,y,z,\lambda) &= x + 4y + 8z - 81 = 0 \end{cases}$$

Diese lineare Gleichungssytem hat die Lösung $x = 1$, $y = 4$, $z = 8$ und $\lambda = -2$.

Geometrische Deutung: Durch $\varphi(x,y,z) = x + 4y + 8z - 81 = 0$ ist eine Ebene im dreidimensionalen Raum $\mathbb{R}^3$ beschrieben. $u = f(x,y,z) = x^2 + y^2 + z^2$ gibt das Quadrat des Abstandes eines Ebenenpunktes vom Nullpunkt (Ursprung) an. Damit hat der Ebenenpunkt $P(1/4/8)$ den kleinsten Abstand, nämlich $d = \sqrt{1^2 + 4^2 + 8^2} = 9$, vom Ursprung.

Beispiel 5: $u = f(x,y,z) = x^2 + 2y^2 - z^2$ sei maximal unter den beiden Bedingungen $\varphi_1(x,y,z) = x - y + z + 2 = 0$ und $\varphi_2(x,y,z) = 2x + y + z + 5 = 0$

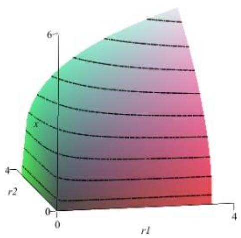
Die Lagrangesche Hilfsfunktion lautet $L(x,y,z,\lambda_1,\lambda_2) = x^2 + 2y^2 - z^2 + \lambda_1\cdot(x - y + z + 2) + \lambda_2\cdot(2x + y + z + 5)$.

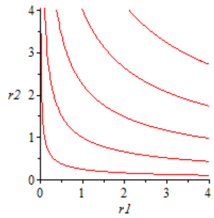
$$\begin{cases} L_x(x,y,z,\lambda_1,\lambda_2) &= 2x + \lambda_1 + 2\lambda_2 = 0 \\ L_y(x,y,z,\lambda_1,\lambda_2) &= 4y - \lambda_1 + \lambda_2 = 0 \\ L_z(x,y,z,\lambda_1,\lambda_2) &= -2z + \lambda_1 + \lambda_2 = 0 \\ L_{\lambda_1}(x,y,z,\lambda_1,\lambda_2) &= x - y + z + 2 = 0 \\ L_{\lambda_2}(x,y,z,\lambda_1,\lambda_2) &= 2x + y + z + 5 = 0 \end{cases}$$

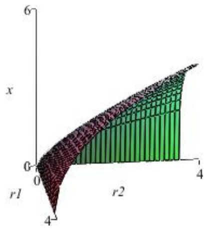
Dieses LGS hat die Lösung $x = -5$, $y = 1$, $z = 4$, $\lambda_1 = 6$, $\lambda_2 = 2$.

Zweiter Weg: Mit $x=-2y-3$ und $z=3y+1$ folgt
$f(x,y,z)=x^2+2y^2-z^2=(-2y-3)^2+y^2-(3y+1)^2=-3y^2+6y+8$.
Mit Hilfe von $f'(y)=-6y+6=0$ folgt $y=1$. Wegen $f''(y)=-6$, also $f''(1)=-6<0$, liegt ein Maximum vor.
Mit $y=1$ folgt aus den beiden Nebenbedungen $x=-2y-3=1$ und $z=3y+1=4$.

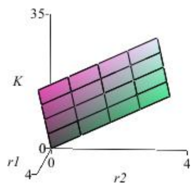
Beispiel 6:
Gegeben ist die Produktionsfunktion $x(r_1,r_2)=2\cdot r_1^{0,3}\cdot r_2^{0,5}$. Dabei bedeuten r_1 und r_2 die Einsatzmengen der beiden Produktionsfaktoren und x die Produktionsmenge.







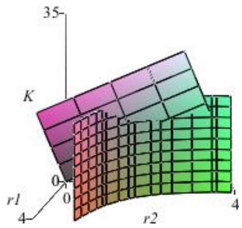
Für konstantes $x=c$ ergeben sich die sog. **Isoquanten** $c=2\cdot r_1^{0,3}\cdot r_2^{0,5}$, bzw. $r_2=\dfrac{c^2}{4r_1^{0,6}}$.



Weiterhin seien $q_1=6$ und $q_2=4$ die Faktorpreise von r_1 und r_2. Die Gesamtkosten K betragen dann $K=6r_1+4r_2$. Das Schaubild von K stellt eine Ebene dar.

Es sollen nun $x=3$ Einheiten hergestellt werden, d.h. $3=2\cdot r_1^{0,3}\cdot r_2^{0,5}$. Wie sind r_1 und r_2 zu wählen, damit die Kosten K minimal werden?

Die Lagrangesche Hilfsfunktion lautet $L(r_1,r_2,\lambda)=6r_1+4r_2+\lambda\cdot(3-2\cdot r_1^{0,3}\cdot r_2^{0,5})$.
Die drei notwendigen Bedingungen für ein Minimum von K lauten dann:
$\dfrac{\partial L}{\partial r_1}=6-0,6\cdot\lambda\cdot\dfrac{r_2^{0,5}}{r_1^{0,7}}=0$, $\dfrac{\partial L}{\partial r_2}=4-\lambda\cdot\dfrac{r_1^{0,3}}{r_1^{0,5}}=0$ und $\dfrac{\partial L}{\partial\lambda}=3-2r_1^{0,3}\cdot r_2^{0,5}=0$.



Aus der ersten Gleichung folgt $\lambda=10\dfrac{r_1^{0,7}}{r_2^{0,5}}$. Setzt man dies in die zweite Gleichung ein, so folgt $r_1=0,4r_2$. Aus der dritten Gleichung folgt daraus $3-2\cdot(0,4r_2)^{0,3}\cdot r_2^{0,5}=0$, d.h. $r_2=\dfrac{3}{4}\cdot 3^{1/4}\cdot 10^{3/8,}\approx 2,34$ und $r_1=\dfrac{3^{5/4}}{10^{5/8}}\approx 0,937$. Die minimalen Kosten sind dann $K(r_1,r_2)=6r_1+4r_2\approx 14,98$.
Im Schaubild ist das Minimum der Schnittkurve bestimmt worden.

Zusatz: Diese Aufgabe lässt sich auch ohne Lagrange lösen:
Aus $3=2\cdot r_1^{0,3}\cdot r_2^{0,5}$ folgt $r_2=\dfrac{2,25}{r_1^{3/5}}$. Eingesetzt in $K=6r_1+4r_2$ ergibt $K(r_1)=6r_1+9\cdot r_1^{-3/5}$ und durch Ableiten $K'(r_1)=6-\dfrac{27}{5}\cdot r_1^{-8/5}$ und $K''(r_1)=\dfrac{216}{25}\cdot r_1^{-13/5}>0$. Aus $K'(r_1)=0$ folgt $r_1=\dfrac{3^{5/4}}{10^{5/8}}\approx 0,937$ wie oben.

Allgemein zur Bedeutung des Lagrange-Multiplikators λ

Die Nebenbedingung $\varphi(x,y)=0$ werde umgeschrieben zu $g(x,y)+c=0$ mit einem $c\in\mathbb{R}$. Dann lässt sich die Langrange-Funktion $L(x,y,\lambda)=f(x,y)+\lambda\cdot\varphi(x,y)$ schreiben in der Form $L(x,y,\lambda)=f(x,y)+\lambda\cdot\big(g(x,y)+c\big)$.
Nach Lagrange bestimmt man die Lösungen der drei Gleichungen $L_x(x,y,\lambda)=f_x(x,y)+\lambda\cdot g_x(x,y)=0$, $L_y(x,y,\lambda)=f_y(x,y)+\lambda\cdot g_y(x,y)=0$ und $L_\lambda(x,y,\lambda)=g(x,y)+c=0$. Dabei können $x=x(c)$ und $y=y(c)$ noch von c abhängen.

Dann gilt $\frac{df(x(c),y(c))}{dc} = \frac{\partial f(x,y)}{\partial x} \cdot \frac{dx(c)}{dc} + \frac{\partial f(x,y)}{\partial y} \cdot \frac{dy(c)}{dc} = f_x(x,y) \cdot x'(c) + f_y(x,y) \cdot y'(c) =$

$= -\lambda \cdot g_x(x,y) \cdot x'(c) - \lambda \cdot g_y(x,y) \cdot y'(c) = -\lambda \cdot \left(g_x(x,y) \cdot x'(c) + g_y(x,y) \cdot y'(c)\right) = -\lambda \cdot \frac{dg(x(c),y(c))}{dc}$.

Da aber die Lösungen $x(c)$ und $y(c)$ die Bedingung $g(x(c),y(c)) + c = 0$ erfüllen, folgt durch Ableiten dieser Gleichung nach c die Beziehung $\frac{dg(x(c),y(c))}{dc} + 1 = 0$, d.h. $\frac{dg(x(c),y(c))}{dc} = -1$.

Somit folgt $\boxed{\frac{df(x(c),y(c))}{dc} = \lambda}$.

In Worten: λ gibt an, wie sich die Funktion $f(x,y)$ <u>im Extrempunkt</u> mit dem Wert von c ändert.

Beispiel: Wir greifen auf obiges Beispiel 1 mit $z = f(x,y) = x^2 + 3xy$ nun aber unter der Nebenbedingung $2x + y + c = 0$ statt $2x + y - 1 = 0$ zurück. Dann lautet die Lagrange-Funktion $L(x,y,\lambda) = x^2 + 3xy + \lambda \cdot (2x + y + c)$. Daraus ergeben sich die drei Gleichungen $L_x(x,y,\lambda) = 2x + 3y + 2\lambda = 0$, $L_y(x,y,\lambda) = 3x + \lambda = 0$, $L_\lambda(x,y,\lambda) = 2x + y + c = 0$ mit den Lösungen $x = 0{,}3c$, $y = 0{,}4c$ und $\lambda = 0{,}9c$.

Der Hochpunkt lautet $H\left(0{,}3c / 0{,}4c / 0{,}45c^2\right)$. Und in der Tat ist $\left(0{,}45c^2\right)' = 0{,}9c$, was mit λ übereinstimmt.

Zwei Ansätze, um zu entscheiden, ob bei Lagrange ein Maximum oder Minimum vorliegt

1. Ansatz mit Hilfe der Hesse-Matrix der Lagrange-Funktion

Es sei $L(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_k)$ eine Lagrange-Funktion mit n Variablen und k Nebenbedingungen. Ihre symmetrische Hesse-Matrix lautet

$$H_L = \begin{pmatrix} L_{x_1x_1} & L_{x_1x_2} & \cdots & L_{x_1x_n} \\ L_{x_2x_1} & L_{x_2x_2} & \cdots & L_{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ L_{x_nx_1} & L_{x_nx_2} & \cdots & L_{x_nx_n} \end{pmatrix}$$

In diese Hesse-Matrix werden die Lösungspunkte samt ihren λ-Werten eingesetzt. Dann gilt die folgende hinreichende Bedingung für die Art des Extremums:

Diese Matrix ist positiv definit	$\Rightarrow$	Es liegt ein Minimum vor
Diese Matrix ist negativ definit	$\Rightarrow$	Es liegt ein Maximum vor

In den anderen Fällen erhält man keine Auskunft.
Ohne Beweis.

<u>Zu obigem Beispiel 1:</u> $L(x,y,\lambda) = x^2 + 3xy + \lambda \cdot (2x + y - 1)$ mit der Lösung $x_0 = 0{,}3$, $y_0 = 0{,}4$ und $\lambda = -0{,}9$.

Die Hesse-Matrix ist $H_L = \begin{pmatrix} L_{xx} & L_{xy} \\ L_{yx} & L_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 3 & 0 \end{pmatrix}$. Ihre Werte sind konstant, so dass die Lösung nicht eingesetzt werden muss. Aus der charakteristischen Gleichung $(2-\lambda)\cdot(-\lambda) - 9 = 0$, also $\lambda^2 - 2\lambda - 9 = 0$ folgen die Eigenwerte $\lambda_{1,2} = 1 \pm \sqrt{10}$ mit unterschiedlichen Vorzeichen. Unsere Matrix ist indefinit, also erhalten wir leider keine Auskunft über die Art der Lösung.

2. Ansatz mit Hilfe der geänderten (berandeten) Hesse-Matrix

Zunächst zum allgemeinen Fall mit n Variablen und k Nebenbedingungen. Die Langrange-Funktion lautet dann $L(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_k) = f(x_1, x_2, ..., x_n) + \lambda_1 \cdot \varphi_1(x_1, x_2, ..., x_n) + ... + \lambda_k \cdot \varphi_k(x_1, x_2, ..., x_n)$.

In der folgenden Matrix ist die Reihenfolge der Variablen vertauscht: Nicht $x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_k$, sondern $\lambda_1, \lambda_2, ..., \lambda_k, x_1, x_2, ..., x_n$, sowohl zeilen- als auch spaltenweise.

$$H_L = \begin{pmatrix} L_{\lambda_1\lambda_1} & L_{\lambda_1\lambda_2} & ... & L_{\lambda_1\lambda_k} & L_{\lambda_1 x_1} & ... & L_{\lambda_1 x_n} \\ L_{\lambda_2\lambda_1} & L_{\lambda_2\lambda_2} & ... & L_{\lambda_2\lambda_k} & L_{\lambda_2 x_1} & ... & L_{\lambda_2 x_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ L_{\lambda_k\lambda_1} & L_{\lambda_k\lambda_2} & ... & L_{\lambda_k\lambda_k} & L_{\lambda_k x_1} & ... & L_{\lambda_k \lambda_n} \\ L_{x_1\lambda_1} & L_{x_1\lambda_2} & ... & L_{x_1\lambda_k} & L_{x_1 x_1} & ... & L_{x_1 x_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ L_{x_n\lambda_1} & L_{x_n\lambda_2} & ... & L_{x_n\lambda_k} & L_{x_n x_1} & ... & L_{x_n x_n} \end{pmatrix} = \begin{pmatrix} 0 & 0 & ... & 0 & \varphi_{1_{x_1}} & ... & \varphi_{1_{x_n}} \\ 0 & 0 & ... & 0 & \varphi_{2_{x_1}} & ... & \varphi_{2_{x_n}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & 0 & \varphi_{k_{x_1}} & ... & \varphi_{k_{x_n}} \\ \varphi_{1_{x_1}} & \varphi_{2_{x_1}} & ... & L_{x_1\lambda_k} & L_{x_1 x_1} & ... & L_{x_1 x_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \varphi_{1_{x_n}} & \varphi_{2_{x_n}} & ... & L_{x_n\lambda_k} & L_{x_n x_1} & ... & L_{x_n x_n} \end{pmatrix}$$

In diese Hesse-Matrix werden die Lösungspunkte samt ihren λ-Werten eingesetzt. Nun berechnet man alle Hauptunterdeterminanten (Hauptminoren) dieser Matrix **mit mehr als $2 \cdot k$ Zeilen** bzw. Spalten. Es gibt $n + k - 2 \cdot k = n - k$ solcher Determinanten. Dann gilt die folgende hinreichende Bedingung für die Art des Extremums:

1. Alle diese Determinanten haben das gleiche Vorzeichen $(-1)^k$ $\Rightarrow$ Es liegt ein **Minimum** vor
2. Diese Determinanten haben wechselnde Vorzeichen, wobei die Gesamtdeterminante der Ordnung $n + k$ das Vorzeichen $(-1)^n$ hat $\Rightarrow$ Es liegt ein **Maximum** vor.
3. Bei einer anderen Vorzeichenfolge liegt kein Extremum vor. Falls dabei Determinanten mit dem Wert 0 vorkommen, darf man sie ignorieren, falls mit den restlichen einer der beiden obigen beiden Fälle zutrifft.

Zu obigem Beispiel 1: $L(x, y, \lambda) = x^2 + 3xy + \lambda \cdot (2x + y - 1)$ mit der Lösung $x_0 = 0,3$, $y_0 = 0,4$ und $\lambda = -0,9$.

Dabei gibt es $n = 2$ Variablen und $k = 1$ Nebenbedingung.

Die berandete Hesse-Matrix ist $H_L = \begin{pmatrix} 0 & \varphi_x & \varphi_y \\ \varphi_x & L_{xx} & L_{xy} \\ \varphi_y & L_{yx} & L_{yy} \end{pmatrix} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 2 & 3 \\ 1 & 3 & 0 \end{pmatrix}$. Hier sind nur Minoren mit mehr als $2 \cdot 1 = 2$ Zeilen zu berechnen, also nur die Gesamtdeterminante. Sie hat den Wert $10 > 0$, also das positive Vorzeichen $(-1)^n = (-1)^2 = 1$. Somit liegt ein Maximum vor.

Zu obigem Beispiel 5: $L(x, y, z, \lambda_1, \lambda_2) = x^2 + 2y^2 - z^2 + \lambda_1 \cdot (x - y + z + 2) + \lambda_2 \cdot (2x + y + z + 5)$ mit $n = 3$ Unbekannten und $k = 2$ Nebenbedingungen.

Die Hesse-Matrix der Langrange-Funktion ist $H_L = \begin{pmatrix} L_{xx} & L_{xy} & L_{xz} \\ L_{yx} & L_{yy} & L_{yz} \\ L_{zx} & L_{zy} & L_{zz} \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{pmatrix}$. Sie hat die drei Eigenwerte 2, 4 und -2, ist folglich indefinit und liefert uns leider keine Information.

Wir versuchen es mit der geränderten Hesse-Matrix

$$H_L = \begin{pmatrix} 0 & 0 & \varphi_{1_x} & \varphi_{1_y} & \varphi_{1_z} \\ 0 & 0 & \varphi_{2_x} & \varphi_{2_y} & \varphi_{2_z} \\ \varphi_{1_x} & \varphi_{2_x} & L_{xx} & L_{xy} & L_{xz} \\ \varphi_{1_y} & \varphi_{2_y} & L_{yx} & L_{yy} & L_{yz} \\ \varphi_{1_z} & \varphi_{2_z} & L_{zx} & L_{zy} & L_{zz} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 0 & 0 \\ -1 & 1 & 0 & 4 & 0 \\ 1 & 1 & 0 & 0 & -2 \end{pmatrix}.$$

Wir müssen alle Minoren vom Grad größer als $2 \cdot k = 4$ bestimmen.
Also nur die Gesamtdeterminante.
Sie hat den Wert $-6 < 0$.

Wegen $(-1)^n = (-1)^3 = -1 < 0$ liegt ein Maximum vor. Wäre diese Determinante positiv, so läge ein Minimum vor. Und wäre ihr Wert gleich 0, so könnte man keine Aussage machen.

Noch ein Beispiel mit $n = 3$ Variablen und $k = 1$ Nebenbedingung.

Die geänderte Hesse-Matrix ist $H_L = \begin{pmatrix} 0 & \varphi_x & \varphi_y & \varphi_z \\ \varphi_x & L_{xx} & L_{xy} & L_{xz} \\ \varphi_y & L_{yx} & L_{yy} & L_{yz} \\ \varphi_z & L_{zx} & L_{zy} & L_{zz} \end{pmatrix}$. Wir müssen alle Minoren vom Grad größer als $2 \cdot k = 2$ bestimmen, also

die Determinante $D_3 = \det \begin{pmatrix} 0 & \varphi_x & \varphi_y \\ \varphi_x & L_{xx} & L_{xy} \\ \varphi_y & L_{yx} & Lyy \end{pmatrix}$ und die Gesamtdeterminante $D_4 = \det(H_L)$.

Dann können folgende Fälle auftreten:

1. Beide Determinanten D_3 und D_4 haben das gleiche Vorzeichen $(-1)^k = (-1)^1 = -1 < 0$. Dann liegt ein Minimum vor.
2. D_4 hat das Vorzeichen $(-1)^n = (-1)^3 = -1 < 0$ und $D_3 > 0$, dann liegt ein Maximum vor.
3. Falls $D_4 > 0$ liegt sicher kein Extrempunkt vor, da für ein Extremum $D_4 < 0$ sein muss.
4. Falls $D_4 < 0$ und $D_3 = 0$, dann fehlt das Vorzeichen von D_3 zur Entscheidung, ob ein Maximum oder Minimum vorliegt.
5. Falls $D_4 = 0$, dann ist keine Aussage möglich.

Die Taylorreihe für mehrere Variablen

Es sei $p(x, y) = a_{00} + (a_{10}x + a_{01}y) + (a_{20}x^2 + a_{11}xy + a_{02}y^2) + (a_{30}x^3 + a_{21}x^2y + a_{12}xy^2 + a_{03}y^3) + \cdots$ ein Polynom in zwei Variablen x und y. Für die partiellen Ableitungen folgt dann

$p_x(x, y) = a_{10} + (2a_{20}x + a_{11}y) + (3a_{30}x^2 + 2a_{21}xy + a_{12}y^2) + \cdots$

$p_y(x, y) = a_{01} + (a_{11}x + 2a_{02}y) + (a_{21}x^2 + 2a_{12}xy + 3a_{03}y^2) + \cdots$

$p_{xx}(x, y) = 2a_{20} + (6a_{30}x + 2a_{21}y) + \cdots$

$p_{xy}(x, y) = p_{yx}(x, y) = a_{11} + (2a_{21}x + 2a_{12}y) + \cdots$

$p_{yy}(x, y) = 2a_{02} + (2a_{12}x + 6a_{03}y) + \cdots$

$p_{xxx}(x, y) = 6a_{30} + \cdots, \quad p_{xxy}(x, y) = 2a_{21} + \cdots, \quad p_{xyy}(x, y) = 2a_{12} + \cdots, \quad p_{yyy}(x, y) = 6a_{03} + \cdots$

Die Funktion $z = f(x, y)$ soll um die Stelle $(0,0)$ entwickelt werden, d.h. $f(x, y) = p(x, y)$. Es gilt

$f(0,0) = p(0,0) = a_{00}$

$f_x(0,0) = p_x(0,0) = a_{10}, \quad f_y(0,0) = p_y(0,0) = a_{01}$

$f_{xx}(0,0) = p_{xx}(0,0) = 2a_{20}, \quad f_{xy}(0,0) = p_{xy}(0,0) = a_{11}, \quad f_{yy}(0,0) = p_{yy}(0,0) = 2a_{02}$

$f_{xxx}(0,0) = p_{xxx}(0,0) = 6a_{30}, \quad f_{xxy}(0,0) = p_{xxy}(0,0) = 2a_{21},$

$f_{xyy}(0,0) = p_{xyy}(0,0) = 2a_{12}, \quad f_{yyy}(0,0) = p_{yyy}(0,0) = 6a_{03}.$

Somit folgt

$$f(x, y) = f(0,0) + f_x(0,0) \cdot x + f_y(0,0) \cdot y + \frac{1}{2} f_{xx}(0,0) \cdot x^2 + f_{xy}(0,0) \cdot xy + \frac{1}{2} f_{yy}(0,0) \cdot y^2 +$$
$$+ \frac{1}{3!} f_{xxx}(0,0) \cdot x^3 + \frac{1}{2! \cdot 1!} f_{xxy}(0,0) \cdot x^2y + \frac{1}{1! \cdot 2!} f_{xyy}(0,0) \cdot xy^2 + \frac{1}{3!} f_{yyy}(0,0) \cdot y^3 + \cdots$$

oder für (x_0, y_0) statt $(0,0)$:

$$f(x,y) = f(x_0,y_0) + f_x(x_0,y_0)\cdot(x-x_0) + f_y(x_0,y_0)\cdot(y-y_0) +$$
$$+\frac{1}{2!}f_{xx}(x_0,y_0)\cdot(x-x_0)^2 + \frac{1}{1!\cdot 1!}f_{xy}(x_0,y_0)\cdot(x-x_0)\cdot(y-y_0) + \frac{1}{2!}f_{yy}(x_0,y_0)\cdot(y-y_0)^2 +$$
$$+\frac{1}{3!}f_{xxx}(x_0,y_0)\cdot(x-x_0)^3 + \frac{1}{2!\cdot 1!}f_{xxy}(x_0,y_0)\cdot(x-x_0)^2\cdot(y-y_0) + \frac{1}{1!\cdot 2!}f_{xyy}(x_0,y_0)\cdot(x-x_0)\cdot(y-y_0)^2 +$$
$$+\frac{1}{3!}f_{yyy}(x_0,y_0)\cdot(y-y_0)^3 + \cdots$$

Zusammengefasst: $\boxed{f(x,y) = \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{1}{(n-k)!\cdot k!}\cdot\frac{\partial^n}{\partial x^{n-k}\partial y^k}f(x_0,y_0)\cdot(x-x_0)^{n-k}\cdot(y-y_0)^k}$

oder $\boxed{f(x,y) = \sum_{i=0}^{\infty}\sum_{j=0}^{n}\frac{1}{i!\cdot j!}\cdot\frac{\partial^{i+j}}{\partial x^i\partial y^j}f(x_0,y_0)\cdot(x-x_0)^i\cdot(y-y_0)^j}$

Die Taylorreihe einer Funktion $f(x_1,x_2,...,x_n)$ mit n Variablen um die Stelle $(a_1,a_2,...,a_n)$ lautet

$$\boxed{f(x_1,x_2,\dots,x_n) = \sum_{i_1=0}^{\infty}\dots\sum_{i_n=0}^{\infty}\frac{1}{i_1!\cdots i_n!}\frac{\partial^{i_1}}{\partial x_1^{i_1}}\cdots\frac{\partial^{i_n}}{\partial x_n^{i_n}}f(a_1,\dots a_n)\cdot(x_1-a_1)^{i_1}\cdots(x_n-a_n)^{i_n}}\ .$$

Beispiel 1: $f(x,y) = e^{x^2+y^2}$ an der Stelle $(0,0)$.

$f_x(x,y) = 2x\,e^{x^2+y^2}$, $f_y(x,y) = 2y\,e^{x^2+y^2}$,

$f_{xx}(x,y) = (2+4x^2)e^{x^2+y^2}$, $f_{xy}(x,y) = 4xy\,e^{x^2+y^2}$, $f_{yy}(x,y) = (2+4y^2)e^{x^2+y^2}$,

$f_{xxx}(x,y) = (12x+8x^3)e^{x^2+y^2}$, $f_{xxy}(x,y) = (4y+8x^2y)e^{x^2+y^2}$, $f_{xyy}(x,y) = (4x+8xy^2)e^{x^2+y^2}$,

$f_{yyy}(x,y) = (12y+8y^3)e^{x^2+y^2}$

$\frac{\partial^4}{\partial_x^4}f(x,y) = (12+48x^2+16x^4)e^{x^2+y^2}$, $\frac{\partial^4}{\partial_x^3\,\partial_y}f(x,y) = (24xy+16x^3y)e^{x^2+y^2}$,

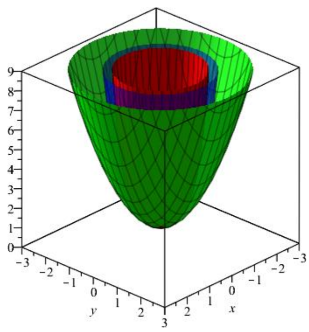
$\frac{\partial^4}{\partial_x^2\,\partial_y^2}f(x,y) = (4+8x^2+8y^2+16x^2y^2)e^{x^2+y^2}$,

$\frac{\partial^4}{\partial_x\,\partial_y^3}f(x,y) = (24xy+16xy^3)e^{x^2+y^2}$, $\frac{\partial^4}{\partial_y^4}f(x,y) = (12+48y^2+16y^4)e^{x^2+y^2}$.

Nach Taylor folgt $f(x,y) = 1+x^2+y^2+\frac{1}{2}x^4+x^2y^2+\frac{1}{2}y^4+\cdots$

Die äußere Figur gehört zu $y = 1+x^2+y^2$, die mittlere zu

$y = 1+x^2+y^2+\frac{1}{2}x^4+x^2y^2+\frac{1}{2}y^4$ und die innerste zu $f(x,y) = e^{x^2+y^2}$.



Beispiel 2: $f(x,y) = \frac{x-y}{x+y}$ um die Stelle $(2/1)$

$f_x(x,y) = \frac{2y}{(x+y)^2}$, $f_y(x,y) = -\frac{2x}{(x+y)^2}$,

$f_{xx}(x,y) = -\frac{4y}{(x+y)^3}$, $f_{xy}(x,y) = \frac{2x-2y}{(x+y)^3}$, $f_{yy}(x,y) = \frac{4x}{(x+y)^3}$,

$f_{xxx}(x,y) = \frac{12y}{(x+y)^4}$, $f_{xxy}(x,y) = \frac{-4x+8y}{(x+y)^4}$, $f_{xyy}(x,y) = \frac{-8x+4y}{(x+y)^4}$, $f_{yyy}(x,y) = -\frac{12x}{(x+y)^4}$.

Nach Taylor folgt $f(x,y) = \frac{1}{3}+\frac{2}{9}(x-2)-\frac{4}{9}(y-1)-\frac{2}{27}(x-2)^2+\frac{2}{27}(x-2)(y-1)+\frac{4}{27}(y-1)^2+$

$$+\frac{2}{81}(x-2)^3-\frac{2}{27}(x-2)(y-1)^2-\frac{4}{81}(y-1)^3+\cdots\ .$$