3.2 Stetigkeit

Die Stetigkeit von Funktionen $f : \mathbb{R} \to \mathbb{R}$ kennen wir bereits aus Teil I. Dort hatten wir zuerst Folgen betrachtet, um den Begriff der Stetigkeit definieren zu können. So gehen wir auch bei mehrdimensionalen Funktionen vor:

Seien $(x_k)_k = x_1, x_2, \ldots$ und $(y_k)_k = y_1, y_2, \ldots$ zwei Folgen in den reellen Zahlen.

$\binom{x_k}{y_k} = \binom{x_1}{y_1}, \binom{x_2}{y_2}, \ldots$ bildet dann eine Folge im $\mathbb{R}^2$ und dies lässt sich natürlich auch in höheren Dimensionen fortführen: Im $\mathbb{R}^n$ betrachten wir Folgen

$$(\vec{x}_k)_k = \vec{x}_1, \vec{x}_2, \vec{x}_3, \ldots = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1n} \end{pmatrix}, \begin{pmatrix} x_{21} \\ \vdots \\ x_{2n} \end{pmatrix}, \begin{pmatrix} x_{31} \\ \vdots \\ x_{3n} \end{pmatrix}, \ldots .$$

Für $i = 1, \ldots, n$ stellt also jedes $(x_{ki})_k = x_{1i}, x_{2i}, \ldots$ eine Folge in $\mathbb{R}$ dar.[4]

3.2.1 Beispiel

$$(\vec{x})_k = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k^2} \end{pmatrix}_k = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{1}{4} \end{pmatrix}, \begin{pmatrix} \frac{1}{3} \\ \frac{1}{9} \end{pmatrix}, \ldots$$

ist eine Folge im $\mathbb{R}^2$ mit $(x_{k1})_k = \left(\frac{1}{k}\right)_k$ und $(x_{k2})_k = \left(\frac{1}{k^2}\right)_k$.

Wenn wir eine Folge im $\mathbb{R}^n$ als n Folgen in $\mathbb{R}$ betrachten, fällt es uns leicht, die Konvergenz mehrdimensionaler Folgen zu definieren:

3.2.2 Definition Eine Folge $(\vec{x}_k)_k = \left(\begin{pmatrix} x_{k1} \\ \vdots \\ x_{kn} \end{pmatrix}\right)_k$ des $\mathbb{R}^n$ ist *konvergent*, wenn alle Komponentenfolgen $(x_{ki})_k$ für alle $i = 1, \ldots, n$ konvergieren.

3.2.3 Beispiele **(1)** $\lim\limits_{k\to\infty} \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k^2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \vec{0}$

(2) $\begin{pmatrix} 1 - \frac{1}{k} \\ k^2 \end{pmatrix}_k$ konvergiert nicht: $(1 - \frac{1}{k})_k$ konvergiert zwar, aber $(k^2)_k$ ist divergent.

Nun können wir uns überlegen, wie man Stetigkeit auch bei Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ definieren kann.

Erinnern wir uns zuerst an die Situation in den reellen Zahlen: Eine Funktion f *konvergiert* in $a \in \mathbb{R}$ gegen einen Grenzwert $g \in \mathbb{R}$ (d.h. $\lim\limits_{x\to a} f(x) = g$), wenn für jede Folge (a_k) mit $\lim\limits_{k\to\infty} a_k = a$ gilt: $\lim\limits_{k\to\infty} f(a_k) = g$.

f ist *stetig* in a, wenn $\lim\limits_{x\to a} f(x) = f(a)$.

Ganz analog definieren wir nun:

3.2.4 Definition Sei $D \subset \mathbb{R}^n$. Eine Funktion $f : D \to \mathbb{R}^m$ *konvergiert* in $\vec{a} \in D$ gegen $\vec{g} \in \mathbb{R}^m$, wenn für jede Folge $(\vec{a}_k)_k$ mit $\lim\limits_{k\to\infty} \vec{a}_k = \vec{a}$ gilt: $\lim\limits_{k\to\infty} f(\vec{a}_k) = \vec{g}$.

f ist *stetig* in $\vec{a}$, wenn f in $\vec{a}$ gegen $f(\vec{a})$ konvergiert, also, wenn

$$\lim_{k\to\infty} f(\vec{a}_k) = f(\vec{a}) \qquad \text{für jede Folge } (\vec{a}_k)_k \text{ mit } \lim_{k\to\infty} \vec{a}_k = \vec{a} \;.$$

[4] Auch, wenn die Schreibweise x_{ki} an Komponenten von Matrizen erinnert, handelt es sich hier um keine Matrix: $k = 1, 2, 3, \ldots$ läuft in Folgen „bis ins Unendliche“.

Wie wir es schon von Funktionen $f : \mathbb{R} \to \mathbb{R}$ kennen, sind aus stetigen Funktionen zusammengesetzte Funktionen wiederum stetig:

3.2.5 Satz Sind $f, g : \mathbb{R}^n \to \mathbb{R}^m$ und $h : \mathbb{R}^m \to \mathbb{R}^l$ stetig, so sind auch $h(f(\vec{x}))$ und $f(\vec{x}) \pm g(\vec{x})$ stetig. Sind $f, g : \mathbb{R}^n \to \mathbb{R}$ stetig, dann ist auch $f(\vec{x}) \cdot g(\vec{x})$ und, falls $g(\vec{x}) \neq 0$, $\dfrac{f(\vec{x})}{g(\vec{x})}$ stetig.

3.2.6 Beispiele

$$f(x, y, z) = \frac{x}{z^2} - x^2 y z^3 + 2 \qquad \text{und} \qquad f(x, y) = e^{x-y}$$

sind (für $z \neq 0$) stetig, da sich diese Funktionen aus stetigen Funktionen zusammensetzen.

Für $f : \mathbb{R}^n \to \mathbb{R}^m$ ist $f(\vec{x}) \in \mathbb{R}^m$ und daher können wir den Funktionswert auch als Vektor schreiben, dessen Komponenten Funktionen $f_i : \mathbb{R}^n \to \mathbb{R}$ sind:

$$f(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{pmatrix} .$$

3.2.7 Satz Sei $D \subset \mathbb{R}^n$. $f : D \to \mathbb{R}^m$ ist genau dann stetig, wenn alle *Komponentenfunktionen* $f_i(\vec{x}) : D \to \mathbb{R}$ für $i = 1, \ldots, n$ stetig sind.

3.2.8 Beispiele **(1)** $f(t) = (\sin t, \cos t)$ ist stetig, da $f_1(t) = \sin t$ und $f_2(t) = \cos t$ stetig sind.

(2) $f(x, y) = \left(2x^2 - y, \sqrt{y}, \dfrac{1}{x-y}\right)$ ist stetig für $x \neq y$ und $y \geq 0$, da $f_1(x, y) = 2x^2 - y$, $f_2(x, y) = \sqrt{y}$ und für $x \neq y$ auch $f_3(x, y) = \frac{1}{x-y}$ stetig sind. Für $x = y$ oder $y < 0$ ist die Funktion nicht definiert, also weder stetig noch unstetig.

(3)

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{für } \binom{x}{y} \neq \vec{0} \\ 0 & \text{für } \binom{x}{y} = \vec{0} \end{cases}$$

ist stetig für $\binom{x}{y} \neq \vec{0}$, da sich die Funktion dort aus stetigen Funktionen zusammensetzt.

Im Punkt $\binom{x}{y} = \vec{0}$ mag man vermuten (vgl. Abbildung 18), dass die Funktion dort unstetig ist. Um dies zu zeigen, müssen wir wegen Definition 3.2.4 nachweisen, dass es eine Folge $\begin{pmatrix} x_k \\ y_k \end{pmatrix}$ mit $\lim\limits_{k\to\infty} \begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ gibt, für die gilt:

$$\lim_{k\to\infty} f(x_k, y_k) \neq f(0, 0) .$$

Betrachten wir dazu $\begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k} \end{pmatrix} \stackrel{k\to\infty}{\longrightarrow} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Es gilt

$$f(x_k, y_k) = f\left(\frac{1}{k}, \frac{1}{k}\right) = \frac{\frac{1}{k} \cdot \frac{1}{k}}{\frac{1}{k^2} + \frac{1}{k^2}} = \frac{1}{2} \stackrel{k\to\infty}{\longrightarrow} \frac{1}{2} \neq 0 = f(0, 0) ,$$

f ist also unstetig in $\vec{0}$.

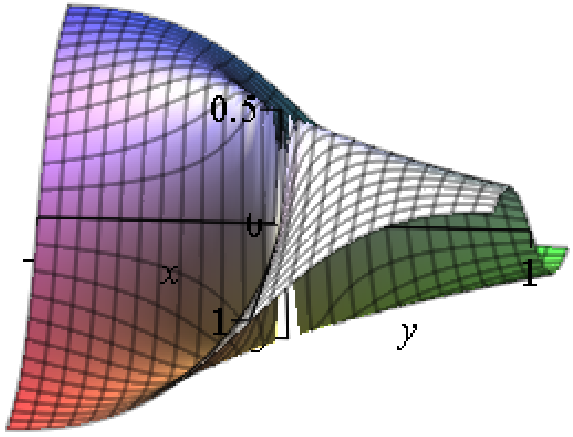

Abbildung 18: $f(x,y) = \frac{xy}{x^2+y^2}$ für $\binom{x}{y} \neq \vec{0}$

3.3 Differenzierbarkeit

Zur Erinnerung: Die Ableitung einer reellen Funktion $f : \mathbb{R} \to \mathbb{R}$ an einer Stelle $x_0 \in \mathbb{R}$ ist

$$\begin{aligned} \frac{d}{dx} f(x_0) = f'(x_0) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\ &= \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \qquad \text{mit } h = x - x_0. \end{aligned} \tag{7}$$

Analog können wir bei Funktionen mehrerer Variablen verfahren: Betrachten wir beispielsweise eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, dann können wir diese Funktion sowohl nach der ersten als auch nach der zweiten Komponente von $\binom{x}{y}$ ableiten. Wir erhalten

$$\begin{aligned} \frac{\partial}{\partial x} f(x_0, y_0) &= \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \ , \\ \frac{\partial}{\partial y} f(x_0, y_0) &= \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} \ . \end{aligned}$$

Außer der etwas anderen Schreibweise „$\frac{\partial}{\partial x}$" statt „$\frac{d}{dx}$", entspricht dies gerade der eindimensionalen Differenzierbarkeit in (7).

Allgemein:

3.3.1 Definition Eine Funktion $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, heißt in $\vec{x} \in D$ *partiell differenzierbar nach* x_k, wenn der Grenzwert

$$\frac{\partial}{\partial x_k} f(\vec{x}) = f_{x_k}(\vec{x}) = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{k-1}, x_k + h, x_{k+1}, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}$$

existiert.

Partielle Ableitungen höherer Ordnung erhält man durch wiederholte partielle Differenziation. Ist f nach allen $x_1, \ldots, x_n$ partiell differenzierbar, dann heißt f *partiell differenzierbar* in $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

3.3.2 Beispiele **(1)** $f(x, y) = 2x^2 + xy - 2y$

Es ist $\dfrac{\partial}{\partial x} f(x, y) = f_x(x, y) = 4x + y$ und $\dfrac{\partial}{\partial y} f(x, y) = f_y(x, y) = x - 2$.

Die höheren partiellen Ableitungen sind

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}f(x,y)\right) &= \frac{\partial^2}{\partial x\partial x}f(x,y) = f_{xx}(x,y) = 4\\
\frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}f(x,y)\right) &= \frac{\partial^2}{\partial y\partial x}f(x,y) = f_{xy}(x,y) = 1\\
\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}f(x,y)\right) &= \frac{\partial^2}{\partial x\partial y}f(x,y) = f_{yx}(x,y) = 1\\
\frac{\partial}{\partial y}\left(\frac{\partial}{\partial y}f(x,y)\right) &= \frac{\partial^2}{\partial y\partial y}f(x,y) = f_{yy}(x,y) = 0
\end{aligned}$$

(2) $f(x,y) = e^{x-y}$

Es folgt

$$\begin{aligned}
f_x(x,y) &= e^{x-y} & f_y(x,y) &= -e^{x-y}\\
f_{xx}(x,y) &= e^{x-y} & f_{yx}(x,y) &= -e^{x-y}\\
f_{xy}(x,y) &= -e^{x-y} & f_{yy}(x,y) &= e^{x-y} \qquad \text{usw.}
\end{aligned}$$

(3) $f(x,y,z) = x^3 + xy^2 + xyz$

Die partiellen Ableitungen (erster Ordnung) sind

$$f_x(x,y,z) = 3x^2 + y^2 + yz \qquad f_y(x,y,z) = 2xy + xz \qquad f_z(x,y,z) = xy$$

und die partiellen Ableitungen zweiter Ordnung lauten

$$\begin{aligned}
f_{xx}(x,y,z) &= 6x & f_{yx}(x,y,z) &= 2y+z & f_{zx}(x,y,z) &= y\\
f_{xy}(x,y,z) &= 2y+z & f_{yy}(x,y,z) &= 2x & f_{zy}(x,y,z) &= x\\
f_{xz}(x,y,z) &= y & f_{yz}(x,y,z) &= x & f_{zz}(x,y,z) &= 0\ .
\end{aligned}$$

Unter den partiellen Ableitungen dritter Ordnung haben wir beispielsweise $f_{xxx}(x,y,z) = 6$ oder $f_{zyx}(x,y,z) = 1$ usw.

Obwohl es „eigentlich" nicht gleichgültig ist, in welcher Reihenfolge man nach den verschiedenen Komponenten partiell ableitet, fällt auf, dass das Ergebnis unabhängig von dieser Reihenfolge zu sein scheint: es ist z.B. im letzten Beispiel $f_{xyz}(x,y,z) = f_{zyx}(x,y,z) = 1$. Tatsächlich gilt

3.3.3 Satz Existieren alle partiellen Ableitungen k-ter Ordnung und sind diese alle stetig, dann kann die Reihenfolge der partiellen Differenziationen vertauscht werden.

Bei einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ erhalten wir also (wenn die Funktion partiell differenzierbar ist) n partielle Ableitungen. Wie sieht es bei Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ mit $m \geq 2$ aus?

In dieser Situation können wir die Funktion wieder in m Komponentenfunktionen zerlegen und diese wie gehabt partiell ableiten:

3.3.4 Definition Eine Funktion $f : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$, $f(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{pmatrix}$, mit $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ist *partiell differenzierbar nach* x_k, wenn alle Komponentenfunktionen $f_1(\vec{x}), \ldots, f_m(\vec{x})$ nach x_k partiell differenzierbar sind.

f heißt *partiell differenzierbar*, wenn f nach allen $x_1, \ldots, x_n$ partiell differenzierbar ist.

3.3.5 Beispiel Betrachten wir $f(x,y) = (\sin x, \cos y, x-y)$.

Hier ist $f_1(x,y) = \sin x$, $f_2(x,y) = \cos y$ und $f_3(x,y) = x - y$ und wir erhalten

$$\begin{aligned}
\frac{\partial}{\partial x} f_1(x,y) &= \cos x & \frac{\partial}{\partial y} f_1(x,y) &= 0 \\
\frac{\partial}{\partial x} f_2(x,y) &= 0 & \frac{\partial}{\partial y} f_2(x,y) &= -\sin y \\
\frac{\partial}{\partial x} f_3(x,y) &= 1 & \frac{\partial}{\partial y} f_3(x,y) &= -1\ .
\end{aligned}$$

3.3.6 Definition Sei $f : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$, eine partiell differenzierbare Funktion mit stetigen partiellen Ableitungen $\frac{\partial}{\partial x_k} f_i(\vec{x})$.

Die (m,n)-Matrix

$$f'(\vec{x}) = Df(\vec{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} f_1(\vec{x}) & \dots & \frac{\partial}{\partial x_n} f_1(\vec{x}) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1} f_m(\vec{x}) & \dots & \frac{\partial}{\partial x_n} f_m(\vec{x}) \end{pmatrix}$$

ist die *Ableitung* von f. Weitere übliche Bezeichnungen sind *Funktionalmatrix* oder *Jacobi-Matrix.*

3.3.7 Beispiel $f(x,y) = (\sin x, \cos y, x-y)$ (wie in Beispiel 3.3.5).

Dann ist die Ableitung von f die Matrix $f'(x,y) = \begin{pmatrix} \cos x & 0 \\ 0 & -\sin y \\ 1 & -1 \end{pmatrix}$.

3.3.8 Bemerkung Im Fall eindimensionaler Funktionen $f : \mathbb{R} \to \mathbb{R}$ gibt die Ableitung in einem Punkt a bekanntlich gerade die Steigung der Tangenten in diesem Punkt an. Die Tangente ist eine Gerade $t(x) = A + B \cdot (x-a)$ (Geradengleichung), wobei $A = f(a)$ den Funktionswert an der Stelle a und $B = f'(a)$ die Steigung bezeichnet:

$$t(x) = f(a) + f'(a) \cdot (x-a)\ .$$

Im mehrdimensionalen Fall, $f : \mathbb{R}^n \to \mathbb{R}^m$, ist dies vergleichbar mit der *Tangentialebene* in einem Punkt $\vec{a} \in \mathbb{R}^n$:

$$T(\vec{x}) = f(\vec{a}) + f'(\vec{a}) \cdot (\vec{x} - \vec{a})\ . \tag{8}$$

3.3.9 Beispiel Betrachten wir $f(x,y) = x^2 - y^3$ im Punkt $\vec{a} = \binom{1}{-1}$ (vgl. Abbildung 19).

f' ist hier die $(1,2)$-Matrix $f'(x,y) = \begin{pmatrix} 2x & -3y^2 \end{pmatrix}$, also haben wir in $\vec{a}$ die Tangentialebene

$$\begin{aligned}
T(x,y) &= f(1,-1) + f'(1,-1) \cdot \left(\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right) \\
&= 1^2 - (-1)^3 + \begin{pmatrix} 2 & -3 \end{pmatrix} \cdot \begin{pmatrix} x-1 \\ y+1 \end{pmatrix} \\
&= 2 + 2(x-1) - 3(y+1) \\
&= 2x - 3y - 3\ .
\end{aligned}$$

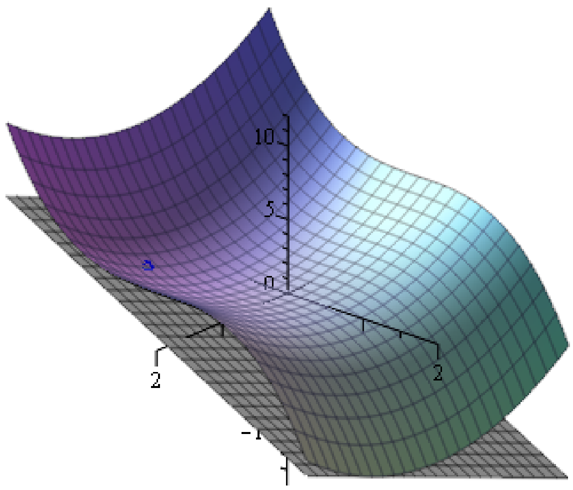

Abbildung 19: $f(x,y) = x^2 - y^3$ mit Tangentialebene in $\binom{1}{-1}$

3.4 Extremwerte

Bei Funktionen $f : \mathbb{R} \to \mathbb{R}$ stellt man sich die Ableitung in einem Punkt gerne als die Steigung des Funktionsgraphen in diesem Punkt vor. Bei Funktionen mit mehrdimensionalen Bildern macht der Begriff der Steigung keinen Sinn mehr, da dort keine „größer“- oder „kleiner“-Relation existiert.

Auch bei Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ (deren Graph eine Art „Berg- und Tallandschaft“ darstellt) können wir nicht von „der“ Steigung in einem Punkt sprechen, da die Steigung in einem Punkt von der Richtung abhängt, in der sie gemessen wird. Statt dessen interessiert man sich eher für die maximale Steigung in einem Punkt.

3.4.1 Definition Ist $f : D \to \mathbb{R}$ mit $D \subset \mathbb{R}^n$, in $\vec{a} \in D$ partiell differenzierbar, dann ist der Vektor der partiellen Ableitungen an der Stelle $\vec{a} \in D$ der *Gradient* von f in $\vec{a}$:

$$\operatorname{grad} f(\vec{a}) = \left(\frac{\partial}{\partial x_1} f(\vec{a}), \frac{\partial}{\partial x_2} f(\vec{a}), \ldots, \frac{\partial}{\partial x_n} f(\vec{a}) \right)$$

3.4.2 Bemerkung Der Gradient entspricht der Ableitung von f bei Skalar-wertigen Funktionen, die in diesem Fall eine $(1,n)$-Matrix ist. Es ist also $\operatorname{grad} f(\vec{a}) = f'(\vec{a})$.

3.4.3 Satz Der Gradient von f zeigt in die Richtung des stärksten Anstiegs von f.

3.4.4 Beispiele **(1)** $f(x,y) = x^2 - y^3$ in $\vec{a} = \binom{1}{-1}$ (vgl. Abbildung 19).

Wir haben $\operatorname{grad} f(x,y) = \big(2x, -3y^2\big)$, also $\operatorname{grad} f(1,-1) = (2,-3)$. Im Punkt $\binom{1}{-1}$ ist der stärkste Anstieg also in Richtung $\binom{2}{-3}$ zu finden.

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

Der Gradient dieser Funktion ist $\operatorname{grad} f(x,y) = (2x, 2y)$, also z.B.

$$\operatorname{grad} f(0,0) = \vec{0}\ .$$

In $\binom{0}{0}$ haben hier also keine Richtung des stärksten Anstiegs!

Grund hierfür ist, dass die Funktion an der Stelle $\vec{0}$ ein Minimum hat und die Tangentialebene in diesem Punkt daher waagerecht ist (vgl. Abbildung 20). Dies entspricht bei eindimensionalen Funktionen einer waagerechten Tangente mit Steigung 0.

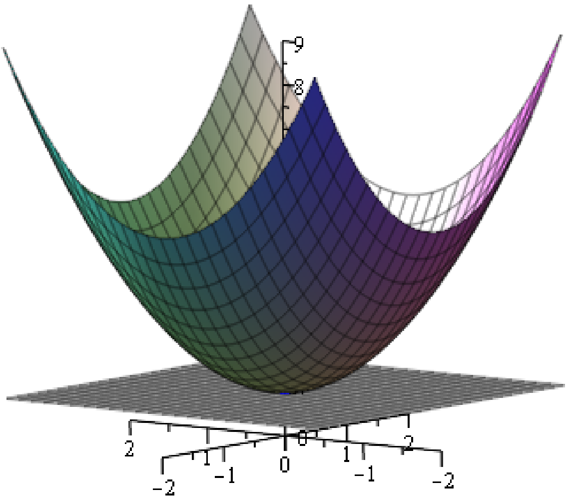

Abbildung 20: $f(x, y) = x^2 + y^2 + 1$ mit Tangentialebene in $\vec{0}$

Bei Funktionen $f : \mathbb{R} \to \mathbb{R}$ wissen wir, dass in einem lokalen Extremum stets $f'(x) = 0$ gilt (das *notwendige Kriterium* für Extrema). Im Mehrdimensionalen ist die Situation ganz ähnlich:

3.4.5 Satz (Notwendiges Kriterium für Extremwerte) Eine Funktion $f : D \to \mathbb{R}$ mit $D \subset \mathbb{R}^n$ sei in $\vec{a} \in D$ partiell differenzierbar. Hat f in $\vec{a}$ ein lokales Extremum, dann gilt

$$\operatorname{grad} f(\vec{a}) = \vec{0} \, .$$

3.4.6 Bemerkung Wie im Eindimensionalen können wir dies nutzen, um *mögliche* Stellen für Extrema zu finden. Es ist aber auch hier unklar, ob an diesen Stellen wirklich Extremwerte vorliegen. Betrachten wir beispielsweise die Sattelfläche $f(x, y) = x^2 - y^2$ (Abbildung 12). Der Gradient ist $\operatorname{grad} f(x, y) = (2x, -2y)$.

In $\vec{0}$ ist also $\operatorname{grad} f(\vec{0}) = \vec{0}$, das notwendige Kriterium ist somit erfüllt. Tatsächlich gibt es in $\vec{0}$ aber keinen Extremwert!

Im Eindimensionalen konnten wir die Entscheidung, ob ein Extremum vorliegt, meistens treffen, indem wir das Vorzeichen der zweiten Ableitung prüften: ist $f'(a) = 0$ und $f''(a) > 0$, haben wir ein Minimum in a, ist $f'(a) = 0$ und $f''(a) < 0$, haben wir ein Maximum.

Im Mehrdimensionalen ist dies leider nicht mehr so einfach, da wir nicht *eine* zweite Ableitung haben, sondern mehrere partielle Ableitungen zweiter Ordnung:

3.4.7 Definition Sei $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, eine zweimal partiell differenzierbare Funktion. Die (n, n)-Matrix der partiellen Ableitungen zweiter Ordnung in $\vec{a} \in D$ heißt *Hesse-Matrix* von f in $\vec{a}$

$$H_f(\vec{a}) = \begin{pmatrix} f_{x_1,x_1}(\vec{a}) & f_{x_1,x_2}(\vec{a}) & \cdots & f_{x_1,x_n}(\vec{a}) \\ f_{x_2,x_1}(\vec{a}) & f_{x_2,x_2}(\vec{a}) & \cdots & f_{x_2,x_n}(\vec{a}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n,x_1}(\vec{a}) & f_{x_n,x_2}(\vec{a}) & \cdots & f_{x_n,x_n}(\vec{a}) \end{pmatrix} .$$

3.4.8 Bemerkung Gibt es alle partiellen Ableitungen zweiter Ordnung und sind diese stetig, dann kann nach Satz 3.3.3 die Reihenfolge der Differenziation vertauscht werden: $f_{x_i,x_k}(\vec{a}) = f_{x_k,x_i}(\vec{a})$. $H_f(\vec{a})$ ist in diesem Fall „symmetrisch“.