Kapitel 8
Differentialrechnung für Funktionen mehrerer Veränderlicher (Einführung)

8.1 Differenzierbarkeit 8

In Kapitel 7 wurden zwei äquivalente Definitionen für die Differenzierbarkeit von Funktionen mit einer Veränderlichen angegeben. Wir werden jetzt die zweite der Definitionen, die die lineare Approximation benutzt, für Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ bzw. $f : \mathbb{R}^n \to \mathbb{R}^m$ mit mehreren Veränderlichen verallgemeinern. Hierzu machen wir zunächst einen kleinen Exkurs in die lineare Algebra. 8/1/0

Eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ ($\mathbb{R}^n, \mathbb{R}^m$ Vektorräume mit kanonischer Basis) kann als Matrix aufgefaßt werden:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Die Elemente $\bar{x} \in \mathbb{R}^n$, $\bar{y} \in \mathbb{R}^m$ sind dann sog. Spaltenvektoren:

$$\bar{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ \bar{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

und es ist

$$A\bar{x} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n a_{1i}x_i \\ \vdots \\ \sum_{i=1}^n a_{mi}x_i \end{pmatrix} := \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \quad \in \mathbb{R}^m.$$

Ist $m = 1$, dann besteht die Matrix A nur aus einer Zeile, die dann als Vektor in $\mathbb{R}^n$ aufgefaßt werden kann. Hierfür wählen wir die Bezeichnung $A := (a_1, \dots, a_n)$.
In diesem Fall schreibt man (der Einfachheit wegen aus drucktechnischen Gründen) die Vektoren $\bar{x} \in \mathbb{R}^n$ als Zeilen: $\bar{x} = (x_1, \dots, x_n)$. Dann ist

$$A\bar{x} = (a_1, \dots, a_n) \cdot (x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i$$

das Skalarprodukt der beiden Vektoren.

Ist zusätzlich $n = 1$, also $A : \mathbb{R} \to \mathbb{R}$, dann besteht die Matrix (bzw. der Vektor) A nur aus einer Komponente: $A = (a_1)$. Wendet man A auf einen „Vektor“ $\bar{x} = (x_1)$ aus $\mathbb{R}$ an, so erhält man $A\bar{x} = a_1 x_1$.
Für A bzw. $\bar{x}$ schreiben wir dann einfach a_1 bzw. x_1.

Diese Bezeichnungsweise ausnutzend liefert die zweite Definition der Differenzierbarkeit aus Kapitel 7, 7/1/10 folgende Formulierung:

f ist in $c \in \mathbb{R}$ differenzierbar $\iff$
f ist in einer Umgebung $U(c)$ definiert, und es existiert eine lineare Abbildung $b : \mathbb{R} \to \mathbb{R}$ und eine Funktion $o(x) : \mathbb{R} \to \mathbb{R}$ mit $\dfrac{o(x)}{|x-c|} \xrightarrow[x \to c]{} 0$, so daß für jedes $x \in U(c)$ gilt: $f(x) = f(c) + b \cdot (x-c) + o(x)$.
Diese Formulierung läßt sich sofort auf Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ und Elemente $\bar{c} \in \mathbb{R}^n$ erweitern.

Definition. (*Differenzierbarkeit, Ableitung*) **8/1/1**
Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ und $\bar{c} \in \mathbb{R}^n$.
f ist in $\bar{c}$ *differenzierbar* (oder *total differenzierbar*)
$\underset{\text{Df}}{=}$ f ist in einer Umgebung $U(\bar{c})$ definiert, und es existiert eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ und eine Funktion $o(\bar{x}) : \mathbb{R}^n \to \mathbb{R}^m$ mit der Eigenschaft $\lim\limits_{\bar{x} \to \bar{c}} \dfrac{o(\bar{x})}{|\bar{x}-\bar{c}|} = \bar{0}$, so daß für jedes $\bar{x} \in U(\bar{c})$ gilt: $f(\bar{x}) = f(\bar{c}) + A \cdot (\bar{x}-\bar{c}) + o(\bar{x})$.
Die Matrix A heißt dann *1. Ableitung* von f an der Stelle $\bar{c}$.

Bez.: $A := f'(\bar{c})$.

Als wichtigsten Spezialfall erhält man die Differenzierbarkeit bzw. die Ableitung einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ an der Stelle $\bar{c} \in \mathbb{R}^n$. A besteht in diesem Falle nur aus einer Zeile. Die Ableitung $f'(\bar{c})$ ist dann ein Vektor (im allgemeinen ist $f'(\bar{c})$ eine Matrix). **8/1/2**
Dieser Vektor heißt auch *Gradient* von f an der Stelle $\bar{c}$ (oder kurz in $\bar{c}$).

Bez.: $f'(\bar{c}) = \operatorname{grad} f(\bar{c})$.

Das Wesen der Differenzierbarkeit besteht auch hier in der linearen Approximierbarkeit einer Funktion f in einer Umgebung $U(\bar{c})$; d.h., f läßt sich in $U(\bar{c})$ darstellen als $f(\bar{c}) + A(\bar{x}-\bar{c})$ plus einem Rest $o(\bar{x})$, der in $U(\bar{c})$ „klein“ ist, wobei hier 'klein' bedeuten soll, daß $\lim\limits_{\bar{x} \to \bar{c}} \dfrac{o(\bar{x})}{|\bar{x}-\bar{c}|} = \bar{0}$ $(\in \mathbb{R}^m)$. **8/1/3**
Durch $\bar{y} = A(\bar{x}-\bar{c})$ wird eine Punktmenge in $\mathbb{R}^n \times \mathbb{R}^m$ beschrieben.
Die durch $\bar{y} := t(\bar{x}) = f(\bar{c}) + A(\bar{x}-\bar{c})$ definierte Punktmenge heißt *Tangentialebene* von f an der Stelle $\bar{c}$, und $t(\bar{x}) = f(\bar{c}) + A(\bar{x}-\bar{c})$ heißt *Gleichung der Tangentialebene.*
Der Punkt $(\bar{c}, f(\bar{c}))$ erfüllt diese Gleichung, d.h., er liegt in der Tangentialebene.
Der Anteil $A(\bar{x}-\bar{c})$ von der Funktion $t(\bar{x})$ heißt *Differential* (oder *1. Differential*) von f in $\bar{c}$.
Wir werden auf diese Begriffe noch einmal zurückkommen, insbesondere im Zusammenhang mit der Berechnung und Darstellung der Ableitung und damit auch der Tangentialebene und des Differentials.

Um die Analogie bei der Differenzierbarkeit von Funktionen einer und mehrerer Veränderlicher besser zu veranschaulichen, betrachten wir die folgenden beiden Spezialfälle.

Für $n = m = 1$ (also $f : \mathbb{R} \to \mathbb{R}$ und $c \in \mathbb{R}$) definiert $t(x) = f(c) + \underbrace{A}_{:=b}(x - c)$ eine Gerade in $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, die Tangente von f an der Stelle c.

Für $n = 2$, $m = 1$ und $\bar{c} \in \mathbb{R}^2$ wird durch $t(\bar{x}) = f(\bar{c}) + A(\bar{x} - \bar{c})$ eine Ebene in $\mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$ bestimmt, die offenbar den Punkt $(\bar{c}, f(\bar{c}))$ enthält. In diesem Fall ist der Begriff „*Tangentialebene*" wörtlich zu nehmen.

Wir haben bereits die Ableitung einer Funktion mit mehreren Veränderlichen definiert, wir haben aber noch kein praktikabeles Verfahren, um die Ableitung einer konkreten Funktion zu berechnen. Dazu betrachten wir zunächst Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ und definieren den Begriff der partiellen Ableitbarkeit.

Definition. (*partielle Ableitung*) **8/1/4**

Sei $f : \mathbb{R}^n \to \mathbb{R}$, $\bar{c} \in \mathbb{R}^n$, $\bar{x} = (x_1, \ldots, x_n)$ und f in einer Umgebung $U(\bar{c})$ definiert.
f ist in $\bar{c}$ *partiell nach* x_i *differenzierbar* $(i = 1, \ldots, n)$

$\underset{\text{Df}}{\Longleftrightarrow}$ Die Funktion $\varphi(x_i) := f(c_1, \ldots, c_{i-1}, x_i, c_{i+1}, \ldots, c_n)$ ist (als Funktion der einen Veränderlichen x_i) an der Stelle c_i differenzierbar.

Nach der früheren Differenzierbarkeitsdefinition bedeutet dies, daß die folgenden Limites existieren:

$$\lim_{x_i \to c_i} \frac{\varphi(x_i) - \varphi(c_i)}{x_i - c_i} = \lim_{h \to 0} \frac{\varphi(c_i + h) - \varphi(c_i)}{h}, \quad \text{für } h := x_i - c_i$$

$$= \lim_{x_i \to c_i} \frac{f(c_i, \ldots, c_{i-1}, x_i, c_{i+1}, \ldots, c_n) - f(\bar{c})}{x_i - c_i}.$$

Der Limes selbst (falls er existiert) heißt *partielle Ableitung* von f nach x_i an der Stelle $\bar{c}$ (oder kurz: in $\bar{c}$).

Bez.: $\dfrac{\partial f}{\partial x_i}(\bar{c}) = f_{x_i}(\bar{c})$.

Für $h := x_i - c_i$ und $\bar{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^n$, $i = 1, \ldots, n$, wobei die Eins in **8/1/5**
$\bar{e}_i$ an der i-ten Stelle steht, existieren die folgenden Limites:

$$\lim_{x_i \to c_i} \frac{\varphi(x_i) - \varphi(c_i)}{x_i - c_i} = \lim_{h \to 0} \frac{f(\bar{c} + h \cdot \bar{e}_i) - f(\bar{c})}{h}.$$

Beispiel. **8/1/6**

Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 y + 2y$, $\bar{c} = (a, b)$.

Dann ist

$$\frac{\partial f}{\partial x}(\bar{c}) = \lim_{x \to a} \frac{f(x, b) - f(a, b)}{x - a} = \Big(f(x, b)\Big)'(a) = (2xb)(a) = 2ab.$$

Allgemein ist

$$\frac{\partial f}{\partial x}(x, y) := \frac{\partial}{\partial x}(f(x, y)) = \frac{\partial}{\partial x}(x^2 y + 2y) = 2xy.$$

Analog erhält man

$$\frac{\partial f}{\partial y}(x,y) = \frac{\partial}{\partial y}(x^2y+2y) = x^2+2 \implies \frac{\partial f}{\partial y}(\bar{c}) = a^2+2.$$

Die Abbildung 8.1 zeigt die geometrische Veranschaulichung der partiellen Ableitungen für eine Funktion mit zwei Veränderlichen.

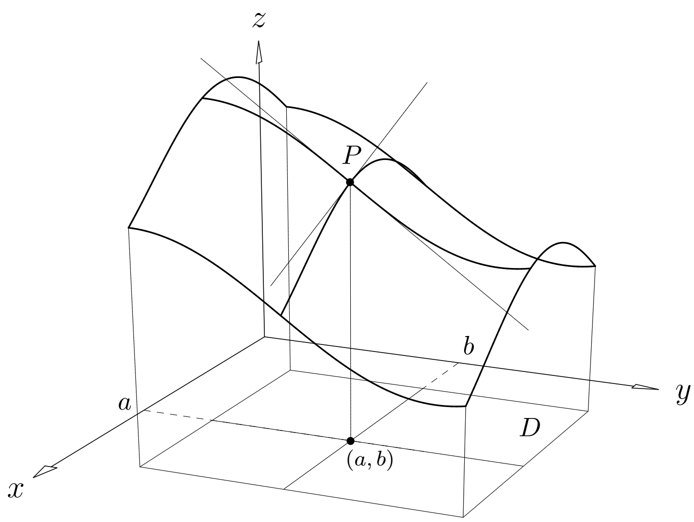

Abb. 8.1 Der Definitionsbereich der hier dargestellten Funktion $f(x,y)$ ist ein Rechteck D. Schränkt man den Definitionsbereich auf $\{(x,y)\in D : y=b\}$ bzw. auf $\{(x,y)\in D : x=a\}$ ein, dann entstehen Kurven auf der Fläche $\{(x,y,f(x,y)) : (x,y)\in D\}$. Die partiellen Ableitungen f_x und f_y in (a,b) geben die Anstiege der Tangenten an diesen Kurven im Punkt (a,b) in Richtung der x-Achse bzw. der y-Achse an.

Die partielle Ableitung nach x_i gibt also den Anstieg der Tangente in Richtung der Achse x_i an. Wir werden diese Art der Ableitung noch einmal verallgemeinern zur sog. *Richtungsableitung*. Dazu geben wir uns (durch einen geeigneten Vektor) eine beliebige Richtung vor und betrachten den Anstieg der Tangente in diese Richtung, falls die zugrundegelegte Funktion dies zuläßt. Daraus ergibt sich folgende Definition.

Definition. (*Richtungsableitung*) **8/1/7**

Es sei $f : \mathbb{R}^n \to \mathbb{R}$, $\bar{c} \in \mathbb{R}^n$ und f in einer Umgebung $U(\bar{c})$ definiert.
Weiterhin sei $\bar{r} \in \mathbb{R}^n$ und $|\bar{r}| = 1$.
f ist an der Stelle $\bar{c}$ *in Richtung* $\bar{r}$ *differenzierbar*
$\overline{\overline{\text{Df}}}$ Die Funktion $\varphi(h) := f(\bar{c} + h\cdot\bar{r})$ ist (als Funktion der einen Veränderlichen h) an der Stelle 0 differenzierbar;

d.h., es existiert $\lim\limits_{h\to 0}\dfrac{\varphi(h)-\varphi(0)}{h} = \lim\limits_{h\to 0}\dfrac{f(\bar{c}+h\cdot\bar{r})-f(\bar{c})}{h}$.

Der Limes heißt dann *Richtungsableitung* von f an der Stelle $\bar{c}$ in Richtung $\bar{r}$.

Bez.: $\dfrac{\partial f}{\partial \bar{r}}(\bar{c}) = f_{\bar{r}}(\bar{c})$.

Bemerkung. **8/1/8**

(1) Der Vektor $\bar{r}$, der die Richtung festlegt, wird stets mit der Länge 1 gewählt, damit

die Richtungsableitung nur von der Richtung und nicht von der Länge des Vektors $\bar{r}$ abhängt.

(2) Wie auch bei Funktionen einer Veränderlichen können Ableitung, partielle Ableitungen und Richtungsableitungen einer Funktion $f(\bar{x})$ in einer Menge $M \subseteq D(f)$ gebildet werden, wenn die entsprechenden Ableitungen in jedem Punkt der Menge existieren. Diese Ableitungen beschreiben neue in M definierte Funktionen, die wir der Reihe nach mit f', $\dfrac{\partial f}{\partial x_i} := f_{x_i}$, $\dfrac{\partial f}{\partial \bar{r}} := f_{\bar{r}}$ bezeichnen.

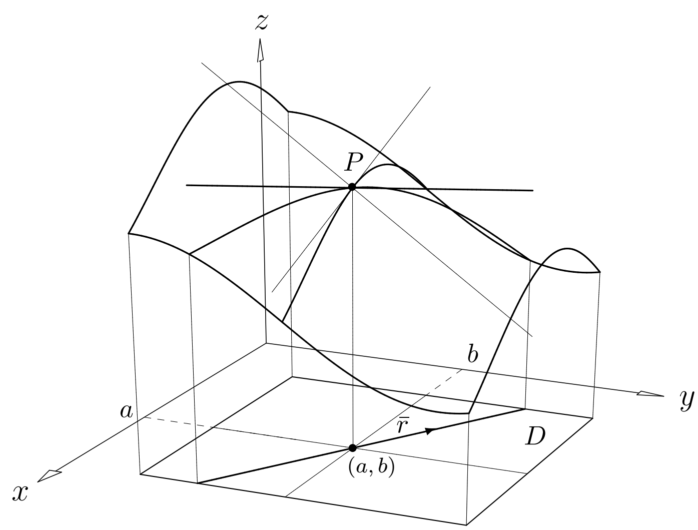

Abb. 8.2 Hier ist die analoge Situation wie in der vorhergehenden Abbildung dargestellt. Es ist zusätzlich eine Richtung durch einen Vektor $\bar{r}$ ausgezeichnet. Schränkt man den Definitionsbereich D auf das im Bild stärker hervorgehobene Geradenstück ein, dann entsteht erneut eine Kurve auf der durch die Funktion gegebenen Fläche. Die Richtungsableitung der Funktion in Richtung $\bar{r}$ gibt den Anstieg der Kurve in dieser Richtung an.

Satz 8.1 *Sei* $f : \mathbb{R}^n \to \mathbb{R}^m$ *und* $\bar{c} \in \mathbb{R}^n$. **8/1/9**
Ist f *in* $\bar{c}$ *differenzierbar, dann ist* f *in* $\bar{c}$ *stetig.*

Beweis. Sei f in $\bar{c}$ differenzierbar. Dann existiert eine Umgebung $U(\bar{c})$, so daß für jedes $\bar{x} \in U(\bar{c})$ gilt: $f(\bar{x}) - f(\bar{c}) = A(\bar{x} - \bar{c}) + o(\bar{x})$. **8/1/10**

g.z.z. : $\lim\limits_{\bar{x}\to\bar{c}} \Big(f(\bar{x} - f(\bar{c})) \Big) = \bar{0}$.

Offenbar gilt $o(\bar{x}) \xrightarrow[\bar{x}\to\bar{c}]{} \bar{0}$.

Nach Satz 6.8 ist eine Vektorfunktion stetig, wenn alle ihre Komponenten stetig sind. Dies trifft insbesondere auf $A(\bar{x} - \bar{c})$ zu. Die Komponenten sind aber als lineare Funktionen (nach Satz 6.11) stetig. Folglich gilt auch $A(\bar{x} - \bar{c}) \xrightarrow[\bar{x}\to\bar{c}]{} \bar{0}$.

Hieraus folgt die Behauptung. ❏

Bemerkung. Aus der partiellen Differenzierbarkeit nach allen Variablen folgt im allgemeinen noch nicht die Stetigkeit, wie das folgende Beispiel zeigt. **8/1/11**

Beispiel. **8/1/12**

Es sei $\quad f(x,y) = \begin{cases} \frac{2xy}{x^2+y^2}, & \text{falls} \quad (x,y) \neq \bar{0}, \\ 0, & \text{falls} \quad (x,y) = \bar{0}. \end{cases}$

Dann ist f an der Stelle $(0,0)$ nach x und y differenzierbar, denn es ist

$$\frac{f(x,0)-f(0,0)}{x-0} = \frac{0-0}{x} = 0 \implies \frac{\partial f}{\partial x}(\bar{0}) = 0$$

und

$$\frac{f(0,y)-f(0,0)}{y-0} = \frac{0-0}{y} = 0 \implies \frac{\partial f}{\partial y}(\bar{0}) = 0.$$

Aber f ist in $\bar{0}$ nicht stetig, denn anderenfalls wäre $\lim\limits_{(x,y)\to(0,0)} f(x,y) = f(0,0) = 0.$
Speziell für $x = y,\ x \neq 0$ und $(x,y) \to (0,0)$ gilt dann

$$f(x,y) = f(x,x) = \frac{2x^2}{x^2+x^2} = 1 \not\longrightarrow 0. \qquad \lightning!$$

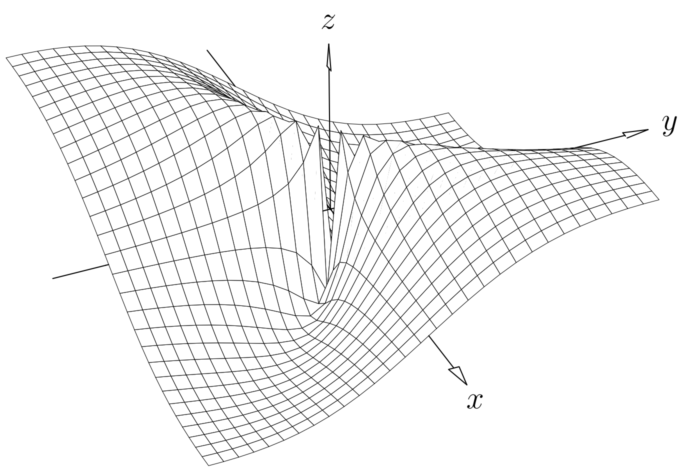

Abb. 8.3 Die Abbildung zeigt die Funktion f aus dem letzten Beispiel. In jeder Umgebung des Nullpunktes nimmt f die Werte ± 1 an, folglich kann die Funktion dort nicht stetig sein. In $(0,0)$ selbst ist der Anstieg von f in Richtung der x-Achse und der y-Achse jeweils null.

Die nächsten drei Sätze stellen wichtige Beziehungen zwischen partieller und totaler Differenzierbarkeit her. **8/1/13**

Satz 8.2 *Sei $f : \mathbb{R}^n \to \mathbb{R}$ und $\bar{c} \in \mathbb{R}^n$.* **8/1/14**
Ist f in einer Umgebung $U(\bar{c})$ definiert und nach allen Variablen partiell differenzierbar und sind alle partiellen Ableitungen in $\bar{c}$ stetig, dann ist f in $\bar{c}$ (total) differenzierbar, und für jedes $\bar{x} \in U(\bar{c})$ gilt $\quad f(\bar{x}) = f(\bar{c}) + \sum\limits_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{c}) \cdot (x_i - c_i) + o(\bar{x}).$

(D.h., die lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}$, die aufgrund der Differenzierbarkeit existiert, ist gegeben durch $A = (a_1, \dots, a_n) = \left(\frac{\partial f}{\partial x_1}(\bar{c}), \dots, \frac{\partial f}{\partial x_n}(\bar{c})\right)$ und $A(\bar{x} - \bar{c}) = \sum\limits_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{c}) \cdot (x_i - c_i)$.)

Beweis. Wir betrachten hier nur den Spezialfall $n = 3$. Dazu sei $\bar{x} = (x, y, z)$ und $\bar{c} = (a, b, c)$ (den allgemeinen Fall beweist man analog). **8/1/15**
Mit Hilfe des 1. Mittelwertsatz der Differentialrechnung (leicht modifiziert) erhält man

$$f(\bar{x}) - f(\bar{c}) = f(x, y, z) - f(a, b, c) =$$

$$f(x, y, z) - f(a, y, z) + f(a, y, z) - f(a, b, z) + f(a, b, z) - f(a, b, c) =$$

$$\frac{\partial f}{\partial x}(\underbrace{a + \vartheta_1(x-a), y, z}_{:= \bar{u}}) \cdot (x - a) + \frac{\partial f}{\partial y}(\underbrace{a, b + \vartheta_2(y-a), z}_{:= \bar{v}}) \cdot (y - b) +$$

$$\frac{\partial f}{\partial z}(\underbrace{a, b, c + \vartheta_3(x-a)}_{:= \bar{w}}) \cdot (z - c) = \qquad \text{(nach dem 1. Mittelwertsatz)}$$

$$\frac{\partial f}{\partial x}(\bar{u}) \cdot (x - a) + \frac{\partial f}{\partial y}(\bar{v}) \cdot (y - b) + \frac{\partial f}{\partial z}(\bar{w}) \cdot (z - c) =$$

$$\frac{\partial f}{\partial x}(\bar{c}) \cdot (x - a) + \frac{\partial f}{\partial y}(\bar{c}) \cdot (y - b) + \frac{\partial f}{\partial z}(\bar{c}) \cdot (z - c) +$$

$$\underbrace{\Big[\frac{\partial f}{\partial x}(\bar{u}) - \frac{\partial f}{\partial x}(\bar{c})\Big] \cdot (x - a) + \cdots + \Big[\frac{\partial f}{\partial x}(\bar{w}) - \frac{\partial f}{\partial x}(\bar{c})\Big] \cdot (z - c)}_{:= r(\bar{x})}.$$

Folglich ist

$$f(\bar{x}) = f(\bar{c}) + \frac{\partial f}{\partial x}(\bar{c}) \cdot (x - a) + \cdots + \frac{\partial f}{\partial z}(\bar{c}) \cdot (z - c) + r(\bar{x}).$$

Behauptung: $\dfrac{r(\bar{x})}{|\bar{x} - \bar{c}|} \xrightarrow[\bar{x} \to \bar{c}]{} 0.$

Nach Voraussetzung sind die partiellen Ableitungen in $\bar{c}$ stetig, folglich gibt es für jedes $\varepsilon > 0$ ein $\delta > 0$, so daß für jedes $\bar{x}$ mit $|\bar{x} - \bar{c}| < \delta$ gilt:

$$\Big|\frac{\partial f}{\partial x}(\bar{x}) - \frac{\partial f}{\partial x}(\bar{c})\Big| < \frac{\varepsilon}{3}, \qquad \text{(im allgemeinen Fall ist der Betrag } < \tfrac{\varepsilon}{n}\text{)}.$$

Die obige Abschätzung gilt auch entsprechend für die partiellen Ableitungen nach y und nach z.
Offenbar gilt:

$$|\bar{x} - \bar{c}| < \delta \implies |\bar{u} - \bar{c}|, \ldots, |\bar{w} - \bar{c}| < \delta,$$

denn $0 < \vartheta_i < 1$ für alle i.
Folglich ist erst recht

$$\Big|\frac{\partial f}{\partial x}(\bar{u}) - \frac{\partial f}{\partial x}(\bar{c})\Big| < \frac{\varepsilon}{3}; \quad \text{analog für } \bar{v} \text{ und } \bar{w}.$$

Damit ergibt sich folgende Abschätzung:

$$|r(\bar{x})| \leq \underbrace{\left|\frac{\partial f}{\partial x}(\bar{u}) - \frac{\partial f}{\partial x}(\bar{c})\right|}_{<\frac{\varepsilon}{3}} \cdot \underbrace{|x-a|}_{\leq|\bar{x}-\bar{c}|} + \cdots + \underbrace{\left|\frac{\partial f}{\partial z}(\bar{w}) - \frac{\partial f}{\partial z}(\bar{c})\right|}_{<\frac{\varepsilon}{3}} \cdot \underbrace{|z-c|}_{\leq|\bar{x}-\bar{c}|}$$

$$\leq 3 \cdot \frac{\varepsilon}{3} \cdot |\bar{x}-\bar{c}| = \varepsilon \cdot |\bar{x}-\bar{c}|.$$

Also

$$\frac{|r(\bar{x})|}{|\bar{x}-\bar{c}|} \leq \varepsilon \quad \text{für} \quad 0 < |\bar{x}-\bar{c}| < \delta.$$

Hieraus folgt die Behauptung. ❏

Bemerkung. Ist f in einer Umgebung $U(\bar{c})$ definiert und nach allen Variablen partiell differenzierbar und sind alle partiellen Ableitungen in $\bar{c}$ stetig, dann läßt sich $f(\bar{x})$ in $U(\bar{c})$ durch die Funktion **8/1/16**

$$t(\bar{x}) := f(\bar{c}) + \sum_{i=1}^{n} \underbrace{\frac{\partial f}{\partial x_i}(\bar{c})}_{:=a_i \in \mathbb{R}} \cdot (x_i - c_i)$$

linear approximieren. Es gilt also $f(\bar{x}) = t(\bar{x}) + o(\bar{x})$.

Satz 8.3 *Sei* $f : \mathbb{R}^n \to \mathbb{R}$ *und* $\bar{c} \in \mathbb{R}^n$. **8/1/17**
Ist f *in* $\bar{c}$ *differenzierbar* (d.h., es gibt reelle Zahlen $a_1, \ldots, a_n$, so daß
$f(\bar{x}) = f(\bar{c}) + \sum_{i=1}^{n} a_i(x_i - c_i) + o(\bar{x})$ für alle $\bar{x}$ in einer Umgebung $U(\bar{c})$),
dann existieren alle partiellen Ableitungen von f *in* $\bar{c}$, *und es ist* $a_i = \frac{\partial f}{\partial x_i}(\bar{c})$.
(Die Ableitung $f'(\bar{c}) := (a_1, \ldots, a_n)$ ist also durch die partiellen Ableitungen eindeutig bestimmt.)

Beweis. Nach Voraussetzung gilt **8/1/18**

$$f(\bar{x}) - f(\bar{c}) = \sum_{i=1}^{n} a_i(x_i - c_i) + o(\bar{x})$$

auch speziell für

$$\bar{x} = (c_1, \ldots, c_{i-1}, x_i, c_{i+1}, \ldots, c_n) = \bar{c} + h\bar{e}_i \quad \text{mit} \quad h = x_i - c_i.$$

Folglich ist

$$f(\bar{x}) - f(\bar{c}) = f(\bar{c} + h\bar{e}_i) - f(\bar{c}) = a_i \cdot h + o(\bar{x}) \quad \Longleftrightarrow$$

$$\frac{f(\bar{c} + h\bar{e}_i) - f(\bar{c})}{h} = a_i + \underbrace{\frac{o(\bar{x})}{h}}_{\xrightarrow[h \to 0]{} 0}.$$

Damit existiert

$$\frac{\partial f}{\partial x_i}(\bar{c}) = \lim_{h\to 0} \frac{f(\bar{c}+h\bar{e}_i) - f(\bar{c})}{h} = a_i. \qquad ❑$$

Satz 8.4 *Sei* $f : \mathbb{R}^n \to \mathbb{R}^m,\ f = (f_1, \ldots, f_m)$ *und* $\bar{c} \in \mathbb{R}^n$. 8/1/19
Ist f *in* $\bar{c}$ *differenzierbar* (d.h., es gibt eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$, also eine $(m\times n)$-Matrix $A = (a_{ji})_{\substack{i=1,\ldots,n\\ j=1,\ldots,m}}$, so daß $f(\bar{x}) = f(\bar{c}) + A(\bar{x}-\bar{c}) + o(\bar{x})$ für alle $\bar{x}$ in einer Umgebung $U(\bar{c})$), *dann existieren alle partiellen Ableitungen* $\frac{\partial f_j}{\partial x_i}(\bar{c}),\ i = 1,\ldots,n,\ j = 1,\ldots,m,$ *und es ist* $a_{ji} = \frac{\partial f_j}{\partial x_i}(\bar{c})$.
(Die Abbildung A ist also eindeutig bestimmt durch die partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}(\bar{c})$.)

Beweis. Den Beweis führt man analog wie zum Satz 8.2. ❑ 8/1/20

Bemerkung. Die Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$, also die 1. Ableitung der Vektorfunktion f, heißt auch *Funktionalmatrix* oder *Jacobimatrix* von f in $\bar{c}$. 8/1/21

$$\textbf{Bez.:}\quad f'(\bar{c}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\bar{c}) & \ldots & \frac{\partial f_1}{\partial x_n}(\bar{c}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\bar{c}) & \ldots & \frac{\partial f_m}{\partial x_n}(\bar{c}) \end{pmatrix} := \frac{\partial(f_1,\ldots,f_m)}{\partial(x_1,\ldots,x_n)}(\bar{c}).$$

Ist $m = 1$, dann besteht die Matrix $A = f'(\bar{c}) = \left(\frac{\partial f}{\partial x_1}(\bar{c}), \ldots, \frac{\partial f}{\partial x_n}(\bar{c})\right)$ nur aus einer Zeile. In diesem Falle hat die 1. Ableitung oder der Gradient von f die Gestalt 8/1/22

$$f'(\bar{c}) = \operatorname{grad} f(\bar{c}) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)(\bar{c}).$$

Ist f in einer ganzen Umgebung $U(\bar{c})$ differenzierbar, dann ist durch $\bar{a} \mapsto f'(\bar{a})$ für jedes $\bar{a} \in U(\bar{c})$ eine Funktion f' definiert. Diese Funktion bezeichnen wir mit

$$f' = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \ldots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \ldots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} = \frac{\partial(f_1,\ldots,f_m)}{\partial(x_1,\ldots,x_n)};$$

für $m = 1$ ist $f' = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right) = \operatorname{grad} f$.

Für das Differential von f an der Stelle $\bar{c}$ schreiben wir auch $df(\bar{x},\bar{c})$ oder kurz $df(\bar{x})$. Mit dieser Bezeichnung läßt sich die Tangentialebene folgendermaßen darstellen:

$$t(\bar{x}) = f(\bar{c}) + df(\bar{x},\bar{c}).$$

Wir werden jetzt noch gewisse Techniken im Umgang mit Differentialen entwickeln, die für manche Anwendungen sehr hilfreich sind. Dazu betrachten wir zunächst den

eindimensionalen Fall $f: \mathbb{R} \to \mathbb{R}$ und $f'(c) = \dfrac{df}{dx}(c) = \dfrac{dy}{dx}(c)$ für $y = f(x)$.
$\dfrac{dy}{dx}(c)$ gibt den Anstieg der Tangente von f in c an. Faßt man die Zahl $f'(c)$ als Bruch $\dfrac{dy}{dx}$ auf, dann erhält man $dy = f'(c)dx$ (dy hängt hierbei natürlich von c ab). Läßt man in der „Verhältniszahl“ $f'(c) = \dfrac{dy}{dx}$ den „Nenner“ dx konstant, dann verändert sich mit c nur noch dy, d.h., dy ist dann eine Funktion von c. (vgl. Abb. 8.4)

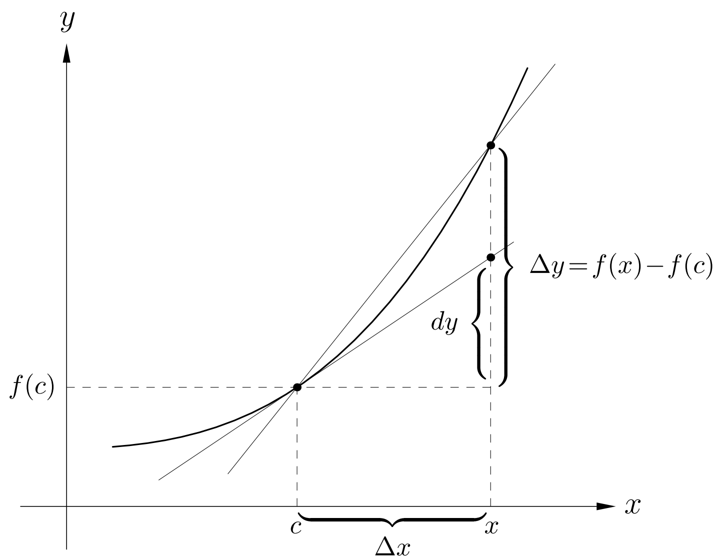

Abb. 8.4 Setzt man $\Delta x := dx$ und läßt dx konstant, dann ist dy nur noch von c abhängig.

Sei jetzt $f: \mathbb{R}^n \to \mathbb{R}$.
Manchmal schreibt man für $x_i - c_i$ auch dx_i und damit

$$\bar{x} - \bar{c} = (x_1 - c_1, \ldots, x_n - c_n) = (dx_1, \ldots, dx_n) := d\bar{x}.$$

Folglich erhält man

$$df(\bar{x}, \bar{c}) = f'(\bar{c}) \cdot (\bar{x} - \bar{c}) = f'(\bar{c}) \cdot d\bar{x} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{c}) \cdot dx_i.$$

Betrachtet man die dx_i als Konstante, dann hängt $df(\bar{x}, \bar{c})$ nur von $\bar{c}$ ab. Damit ist das Differential für die Funktion $f: \mathbb{R}^n \to \mathbb{R}$ wieder eine Abbildung aus $\mathbb{R}^n$ in $\mathbb{R}$, $df(\bar{c}): \mathbb{R}^n \to \mathbb{R}$. Die Ableitung $f'(\bar{c})$ (mit veränderlichem $\bar{c}$) ist hingegen eine Vektorfunktion $f'(\bar{c}) = \left(\dfrac{\partial f}{\partial x_1}, \ldots, \dfrac{\partial f}{\partial x_n}\right)(\bar{c}): \mathbb{R}^n \to \mathbb{R}^n$. Die Ableitung von f' (falls sie existiert) ist dann schon eine Matrix (Funktionalmatrix) usw. Die „Dimension“ der Ableitung wird also größer, die des Differentials nicht. Dies ist ein Vorteil, wenn man mit höheren Ableitungen bzw. Differentialen umgehen will.

Wir betrachten jetzt ein einfaches Beispiel für die Berechnung der Tangentialebene.

Beispiel. **8/1/23**

Es sei $f(x,y) = x^2 + y^2$.
Gesucht ist die Gleichung der Tangentialebene von f an der Stelle $\bar{c} = (1,1)$.
(vgl. Abb. 8.5; ein weiteres Beispiel für die Darstellung einer Funktion und der Tangentialebene an einer Stelle ist in den Abb. 8.6 a und 8.6 b gegeben.)

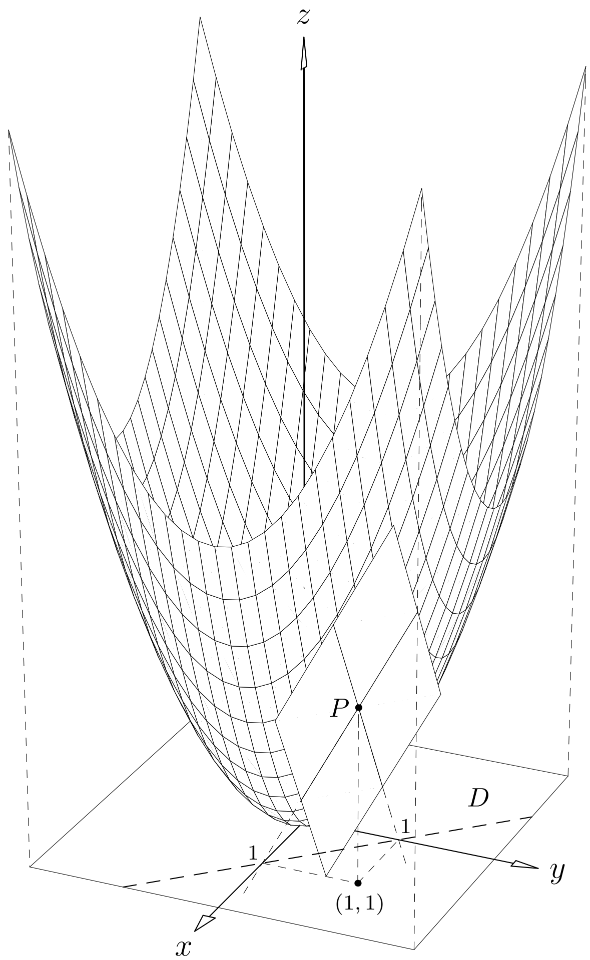

Abb. 8.5 Die Abbildung zeigt die Funktion $f(x,y) = x^2 + y^2$ mit dem Definitionsbereich $D := [-2,2] \times [-2,2]$ und der Tangentialebene an der Stelle $(1,1)$ bzw. im Punkt $P := (1,1,2)$. Das „Kreuz“ in der Tangentialebene entsteht durch die Tangenten an den entsprechenden Kurven im Punkt P in Richtung der Achsen. Die dicker gestrichelte Linie symbolisiert den Schnitt der Tangentialebene mit der durch D gegebenen Ebene.

Es ist

$$f(1,1) = 2,$$

$$\frac{\partial f}{\partial x}(x,y) = 2x \quad \Longrightarrow \quad \frac{\partial f}{\partial x}(1,1) = 2,$$

$$\frac{\partial f}{\partial y}(x,y) = 2y \quad \Longrightarrow \quad \frac{\partial f}{\partial y}(1,1) = 2.$$

Folglich ist

$$
\begin{aligned}
z = t(x,y) &= f(1,1) + \frac{\partial f}{\partial x}(1,1)\cdot(x-1) + \frac{\partial f}{\partial x}(1,1)\cdot(y-1) \\
&= 2 + 2(x-1) + 2(y-1).
\end{aligned}
$$

Die Ebene geht durch die drei Punkte $(1,1,2),(1,0,0),(0,1,0)$, wodurch die Ebene schon eindeutig bestimmt ist.

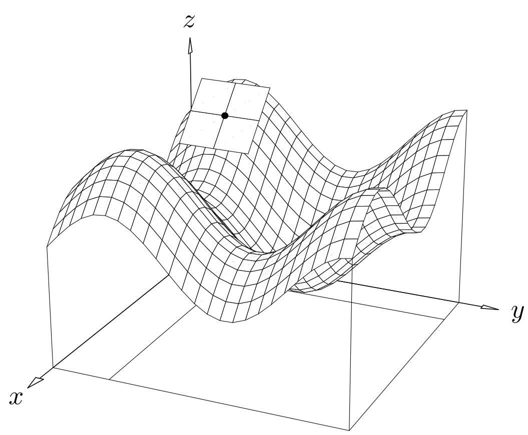

Abb. 8.6 a

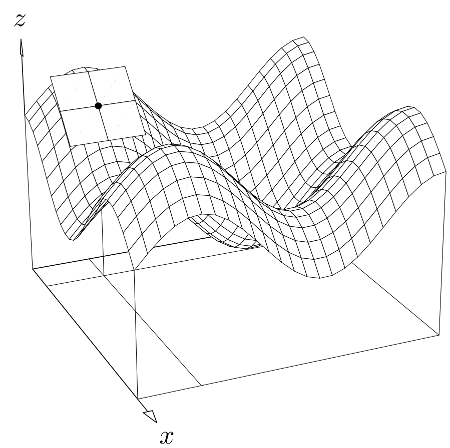

Abb. 8.6 b

In den Abbildungen ist die Funktion $f(x,y) = \cos x + \sin y + 3$ mit dem Definitionsbereich $D := [0, \frac{5}{2}\pi] \times [0, \frac{5}{2}\pi]$ aus zwei verschiedenen Perspektiven dargestellt. Die Tangentialebene wird an der Stelle $(\frac{3}{8}\pi, \frac{1}{2}\pi)$ betrachtet.

In dem folgenden Satz wird nachgewiesen, daß man aus der totalen Differenzierbarkeit **8/1/24**
die Richtungsableitbarkeit erhält und daß sich die Richtungsableitung mit Hilfe der partiellen Ableitungen berechnen läßt.

Satz 8.5 *Sei* $f : \mathbb{R}^n \to \mathbb{R}$ *und* $\bar{c} \in \mathbb{R}^n$. **8/1/25**
Ist f *in* $\bar{c}$ *differenzierbar, dann ist* f *an der Stelle* $\bar{c}$ *in jede Richtung* $\bar{r}$ (*mit* $|\bar{r}| = 1$) *differenzierbar, und es ist* $f_{\bar{r}}(\bar{c}) = f'(\bar{c})\cdot\bar{r} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{c})\cdot r_i$, *wobei* $\bar{r} = (r_1, \ldots, r_n)$.

Beweis. Nach Voraussetzung gilt für alle $\bar{x}$ aus einer Umgebung $U(\bar{c})$: **8/1/26**

$$
f(\bar{x}) = f(\bar{c}) + f'(\bar{c})\cdot(\bar{x}-\bar{c}) + o(\bar{x})
$$

Speziell für $\bar{x} := \bar{c} + h\bar{r} \in U(\bar{c})$ erhält man dann

$$
f(\bar{c}+h\bar{r}) - f(\bar{c}) = f'(\bar{c})\cdot(\underbrace{\bar{c}+h\bar{r}-\bar{c}}_{h\bar{r}}) + o(\bar{x})
$$

und schließlich

$$\frac{f(\bar{c}+h\bar{r})-f(\bar{c})}{h} = f'(\bar{c})\cdot\bar{r} + \underbrace{\frac{o(\bar{x})}{h}}_{\longrightarrow 0}.$$

Folglich ist

$$\lim_{h\to 0}\frac{f(\bar{c}+h\bar{r})-f(\bar{c})}{h} = f'(\bar{c})\cdot\bar{r}.$$

Andererseits ist dieser Limes gleich $f_{\bar{r}}(\bar{c})$. Hieraus folgt die Behauptung. ❑

Wir befassen uns jetzt mit der Differenzierbarkeit zusammengesetzter Funktionen. Da sich die Ableitung einer Vektorfunktion auf die Ableitung ihrer reellwertigen Komponenten zurückführen läßt, werden wir uns im folgenden vorwiegend mit reellwertigen Funktionen befassen. **8/1/27**

Satz 8.6 (*Differentiation rationaler Funktionen*) **8/1/28**
Seien $f,\, g : \mathbb{R}^n \to \mathbb{R}$ *und* $f,\, g$ *in* $\bar{c}$ *differenzierbar. Dann gilt:*

(1) $f \pm g$ *und* $f \cdot g$ *sind in* $\bar{c}$ *differenzierbar, und es ist*
$(f\pm g)'(\bar{c}) = f'(\bar{c}) \pm g'(\bar{c}) \qquad \big(\Longrightarrow \;\; d(f\pm g) = df \pm dg\big),$
$(f\cdot g)'(\bar{c}) = f'(\bar{c})\cdot g(\bar{c}) + f(\bar{c})\cdot g'(\bar{c}) \qquad \big(\Longrightarrow \;\; d(f\cdot g) = g\cdot df + f\cdot dg\big).$

(2) *Ist* $g(\bar{x}) \neq 0$ *für jedes* $\bar{x}$ *in einer Umgebung* $U(\bar{c})$, *dann ist* $\frac{f}{g}$ *in* $\bar{c}$ *differenzierbar, und es ist*
$\left(\frac{f}{g}\right)'(\bar{c}) = \frac{f'(\bar{c})\cdot g(\bar{c}) - f(\bar{c})\cdot g'(\bar{c})}{g^2(\bar{c})} \qquad \left(\Longrightarrow \;\; d\left(\frac{f}{g}\right) = \frac{g\cdot df - f\cdot dg}{g^2}\right).$

Beweis. Den Beweis führt man ähnlich wie für Funktionen einer reellen Veränderlichen; man hat hier lediglich alle Beweisschritte für die partiellen Ableitungen vorzunehmen. ❑ **8/1/29**

Satz 8.7 (*Spezialfall für die Kettenregel*) **8/1/30**
Es sei $g : \mathbb{R}^n \to \mathbb{R}$ *und* $f : \mathbb{R} \to \mathbb{R}$ (*also* $f\circ g : \mathbb{R}^n \to \mathbb{R}$).
Ist g *in* $\bar{c}$ *und* f *in* $g(\bar{c})$ *differenzierbar, dann ist* $f\circ g$ *in* $\bar{c}$ *differenzierbar, und es ist* $(f\circ g)'(\bar{c}) = f'(g(\bar{c}))\cdot g'(\bar{c})$.

Beweis. Sei $b = g(\bar{c})$. Nach Voraussetzung ist g in $\bar{c}$ und f in $b = g(\bar{c})$ differenzierbar. Damit gilt: **8/1/31**

$$g(\bar{x}) - g(\bar{c}) = g'(\bar{c})\cdot(\bar{x}-\bar{c}) + o_1(\bar{x})$$

für alle $\bar{x}$ in einer Umgebung $U(\bar{c})$ und $\frac{o_1(\bar{x})}{|\bar{x}-\bar{c}|} \xrightarrow[\bar{x}\to\bar{c}]{} 0$, und es ist

$$f(y) - f(b) = f'(b) \cdot (y-b) + o_2(y)$$

für alle y in einer Umgebung $U(b)$ und $\dfrac{o_2(y)}{|y-b|} \xrightarrow[y \to b]{} 0.$

Da g in $\bar{c}$ differenzierbar, also dort auch stetig ist, gilt für $\bar{x} \to \bar{c}$ auch $g(\bar{x}) \to g(\bar{c})$. Nach Voraussetzung strebt $o_2(y)$ für $y \to b$ von höherer als 1. Ordnung gegen null, d.h., es gibt eine Funktion $r(y)$, so daß $o_2(y) = r(y) \cdot (y-b)$ und $r(y) \xrightarrow[y \to b]{} 0.$

Wir wählen jetzt $y := g(\bar{x})$. Damit erhält man insgesamt

$$(f \circ g)(\bar{x}) - (f \circ g)(\bar{c}) = f(\underbrace{g(\bar{x})}_{=y}) - f(\underbrace{g(\bar{c})}_{=b}) =$$

$$f'(g(\bar{c})) \cdot (g(\bar{x}) - g(\bar{c})) + o_2(g(\bar{x})) =$$

$$f'(g(\bar{c})) \cdot \Big(g'(\bar{c}) \cdot (\bar{x} - \bar{c}) + o_1(x)\Big) + o_2(g(\bar{x})) =$$

$$f'(g(\bar{c})) \cdot g'(\bar{c}) \cdot (\bar{x} - \bar{c}) + \underbrace{f'(g(\bar{x})) \cdot o_1(x) + o_2(g(\bar{x}))}_{:= o(\bar{x})}.$$

Es bleibt zu zeigen, daß

$$\frac{o(\bar{x})}{|\bar{x} - \bar{c}|} \xrightarrow[\bar{x} \to \bar{c}]{} 0.$$

$f'(g(\bar{c}))$ ist konstant, folglich gilt $\quad f'(g(\bar{c})) \cdot \dfrac{o_1(\bar{x})}{|\bar{x} - \bar{c}|} \xrightarrow[\bar{x} \to \bar{c}]{} 0.$

Weiterhin ist

$$\begin{aligned}
\frac{|o_2(g(\bar{x}))|}{|\bar{x} - \bar{c}|} &= \frac{|r(g(\bar{x})) \cdot (g(\bar{x}) - g(\bar{c}))|}{|\bar{x} - \bar{c}|} \\
&= |r(g(\bar{x}))| \cdot \frac{|g(\bar{x}) - g(\bar{c})|}{|\bar{x} - \bar{c}|} = |r(g(\bar{x}))| \cdot \frac{|g'(\bar{c}) \cdot (\bar{x} - \bar{c}) + o_1(\bar{x})|}{|\bar{x} - \bar{c}|} \\
&\le |r(g(\bar{x}))| \cdot \left(\frac{|g'(\bar{c})| \cdot |\bar{x} - \bar{c}|}{|\bar{x} - \bar{c}|} + \frac{|o_1(\bar{x})|}{|\bar{x} - \bar{c}|} \right) \\
&= |\underbrace{r(g(\bar{x}))}_{\xrightarrow[\bar{x} \to \bar{c}]{} 0}| \cdot \Big(\underbrace{|g'(\bar{c})|}_{\text{konstant}} + \underbrace{\frac{|o_1(\bar{x})|}{|\bar{x} - \bar{c}|}}_{\xrightarrow[\bar{x} \to \bar{c}]{} 0} \Big) \xrightarrow[\bar{x} \to \bar{c}]{} 0.
\end{aligned}$$

Damit ist $f \circ g$ in $U(\bar{c})$ durch

$$t(\bar{x}) := (f \circ g)(\bar{c}) + \Big(\underbrace{f'(g(\bar{c}))}_{\in \mathbf{R}} \cdot \underbrace{g'(\bar{c})}_{\in \mathbf{R}^n} \Big) \cdot (\bar{x} - \bar{c})$$

linear approximiert, folglich ist

$$(f \circ g)'(\bar{c}) = f'(g(\bar{c})) \cdot g'(\bar{c}). \qquad ❑$$

Satz 8.8 (*Kettenregel*) **8/1/32**
Es sei $g : \mathbb{R}^n \to \mathbb{R}^m$ *und* $f : \mathbb{R}^m \to \mathbb{R}^k$ (*also* $f \circ g : \mathbb{R}^n \to \mathbb{R}^k$).
Ist g *in* $\bar{c}$ *und* f *in* $g(\bar{c})$ *differenzierbar, dann ist* $f \circ g$ *in* $\bar{c}$ *differenzierbar, und es ist* $(f \circ g)'(\bar{c}) = f'(g(\bar{c})) \cdot g'(\bar{c})$.
(Das Produkt der „inneren“ und der „äußeren“ Ableitung ist ein Produkt von Matritzen.)

Beweis. Die grundlegende Beweisidee ist die gleiche wie im vorhergehenden Satz. Da der technische Aufwand jedoch wesentlich größer ist, kann hier nur auf die Literatur verwiesen werden. (vgl. z.B. Literaturangabe [4], Teil II, Seite 217) ❑ **8/1/33**

Bemerkung. Für $g : \mathbb{R}^n \to \mathbb{R}^m$ und $f : \mathbb{R}^m \to \mathbb{R}^k$ stellt sich die Ableitung von $(f \circ g)$ an der Stelle $\bar{c}$ wie folgt dar, wobei $\bar{b} = g(\bar{c})$ ist: **8/1/34**

$$\begin{aligned}
(f \circ g)'(\bar{c}) &= f'(g(\bar{c})) \cdot g'(\bar{c}) = f'(\bar{b}) \cdot g'(\bar{c}) = \frac{\partial(f_1, \dots, f_k)}{\partial(y_1, \dots, y_m)}(\bar{b}) \cdot \frac{\partial(g_1, \dots, g_m)}{\partial(x_1, \dots, x_n)}(\bar{c}) \\
&= \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(\bar{b}) & \dots & \frac{\partial f_1}{\partial y_m}(\bar{b}) \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial y_1}(\bar{b}) & \dots & \frac{\partial f_k}{\partial y_m}(\bar{b}) \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\bar{c}) & \dots & \frac{\partial g_1}{\partial x_n}(\bar{c}) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(\bar{c}) & \dots & \frac{\partial g_m}{\partial x_n}(\bar{c}) \end{pmatrix} \\
&= \Big(a_{ji}\Big)_{\substack{i=1,\dots,n \\ j=1,\dots,k}} \qquad \text{und} \qquad a_{ji} = \sum_{\nu=1}^{m} \frac{\partial f_j}{\partial y_\nu}(\bar{b}) \cdot \frac{\partial g_\nu}{\partial x_i}(\bar{c}).
\end{aligned}$$

Beispiele.

(1) Spezialfall einer Verkettung **8/1/35/1**

Sei $g : \mathbb{R} \to \mathbb{R}^2$ und $f : \mathbb{R}^2 \to \mathbb{R}$, also $f \circ g : \mathbb{R} \to \mathbb{R}$, wobei $g := (g_1, g_2)$.
Speziell sei $g_1(t) := t$, $g_2(t) := t^2$, also $g(t) = (g_1(t), g_2(t)) = (t, t^2)$ $(\in \mathbb{R}^2)$ und $f(x, y) := \sin(x \cdot y)$.
Für $x = g_1(t)$ und $y = g_2(t)$ erhält man

$$(f \circ g)(t) = f(g(t)) = f(g_1(t), g_2(t)) = \sin(t \cdot t^2) = \sin t^3.$$

Offenbar ist $f \circ g$ eine reellwertige Funktion einer reellen Veränderlichen, folglich läßt sich die Ableitung nach den Regeln für Funktionen einer Veränderlichen bilden:

$$(f \circ g)'(t) = (\cos t^3) \cdot 3t^2.$$

Wir werden jetzt die Ableitung nach den Regeln für Funktionen mehrerer Veränderlicher berechnen; es wird sich zeigen, daß das gleiche Ergebnis entsteht.
Es ist

$$f'(x,y) = \Big(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\Big),$$

$$g'(t) = \begin{pmatrix} \frac{\partial g_1}{\partial t}(t) \\ \frac{\partial g_2}{\partial t}(t) \end{pmatrix}.$$

Folglich ist

$$(f \circ g)'(t) = f'(g(t)) \cdot g'(t) = \Big(\frac{\partial f}{\partial x}(g(t)), \frac{\partial f}{\partial y}(g(t))\Big) \cdot \begin{pmatrix} \frac{\partial g_1}{\partial t}(t) \\ \frac{\partial g_2}{\partial t}(t) \end{pmatrix} =$$

$$\frac{\partial f}{\partial x}(g(t)) \cdot \frac{\partial g_1}{\partial t}(t) \;+\; \frac{\partial f}{\partial y}(g(t)) \cdot \frac{\partial g_2}{\partial t}(t) := (\star).$$

Es ist

$$\frac{\partial f}{\partial x}(x,y) = \cos(xy) \cdot y \quad\Longrightarrow\quad \frac{\partial f}{\partial x}(g(t)) = \cos\Big(g_1(t) \cdot g_2(t)\Big) \cdot g_2(t) = (\cos\, t^3) \cdot t^2,$$

$$\frac{\partial f}{\partial y}(x,y) = \cos(xy) \cdot x \quad\Longrightarrow\quad \frac{\partial f}{\partial y}(g(t)) = \cos\Big(g_1(t) \cdot g_2(t)\Big) \cdot g_1(t) = (\cos\, t^3) \cdot t,$$

$$\frac{\partial g_1}{\partial t}(t) = 1 \quad \text{und} \quad \frac{\partial g_2}{\partial t}(t) = 2t.$$

Dann ist

$$(f \circ g)'(t) = (\star) = (\cos\, t^3) \cdot t^2 \cdot 1 \;+\; (cos\, t^3) \cdot t \cdot 2t = (\cos\, t^3) \cdot 3t^2.$$

(2) Transformation in Polarkoordinaten **8/1/35/2**

Bei der Lösung mathematischer Probleme, insbesondere in der Physik, der Technik und den Naturwissenschaften überhaupt, ist es oft vorteilhaft, die zu behandelnden Probleme mit Hilfe besonders geeigneter Koordinatensysteme zu beschreiben oder vom kartesischen Koordinatensystem zu einem anderen überzugehen. Da dieser Übergang häufig mit der Verkettung von Funktionen und die Lösung der anstehenden Probleme oft mit der Differenzierbarkeit der Verkettung verbunden ist, wollen wir hier einige wichtige nicht-kartesische Koordinatensysteme behandeln. Zunächst betrachten wir die sog. *Polarkoordinaten*, die schon bei der Behandlung der trigonometrischen Funktionen in Kapitel 5 (vgl. Abb. 5.21) eine gewisse Rolle spielten.

Es sei $P = (a,b) \neq (0,0)$ ein Punkt in der euklidischen Ebene, die mit einem kartesischen Koordinatensystem versehen ist, dessen Achsen mit x bzw. y bezeichnet werden. Offenbar läßt sich der Punkt (a,b) auch eindeutig durch das Paar (r,φ) beschreiben,

wobei r der Abstand von (a, b) zum Nullpunkt ist und φ den Winkel zwischen der x-Achse und der Verbindungsstrecke von $(0, 0)$ nach (a, b) angibt (φ in Bogenmaß gemessen). Damit ist der gleiche Punkt P in der Ebene $\mathbb{R}^2$ durch unterschiedliche Koordinatensysteme eindeutig beschrieben worden (vgl. Abb. 8.7).
a, b sind die kartesischen Koordinaten von P, und r, φ heißen *Polarkoordinaten.*
(Der Nullpunkt ist mit Hilfe der Polarkoordinaten nicht eindeutig darstellbar, da der Winkel φ hierfür beliebig sein könnte.)

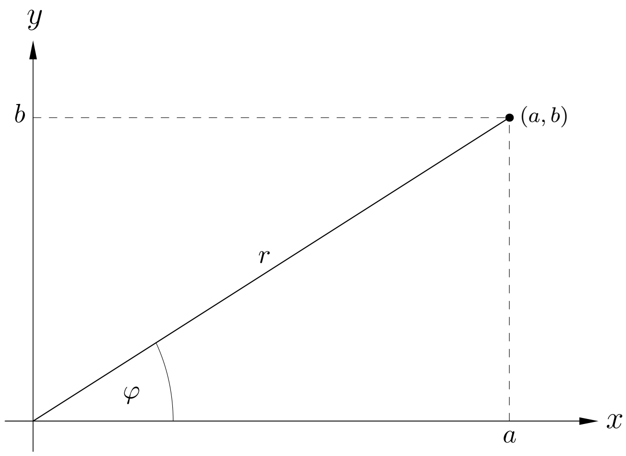

Abb. 8.7 Der Punkt $(a, b) \in \mathbb{R}^2$ ist mit Hilfe von Polarkoordinaten dargestellt. r bezeichnet den Abstand zwischen $(0, 0)$ und (a, b).

Nach Voraussetzung gilt also

$$a = r \cdot \cos\varphi := g_1(r, \varphi)$$

und

$$b = r \cdot \sin\varphi := g_2(r, \varphi).$$

Betrachtet man den Punkt P als variabel, $(a, b) := (x, y)$, dann erhält man eine Abbildung

$$g(r, \varphi) := (g_1(r, \varphi), g_2(r, \varphi)) = (x, y),$$

also

$$g : \mathbb{R}^2 \to \mathbb{R}^2 \quad \text{mit} \quad 0 \leq \varphi < 2\pi \quad \text{und} \quad 0 \leq r < \infty.$$

Wir betrachten jetzt ein Beispiel einer Funktion in kartesischen Koordinaten und in Polarkoordinaten.
Sei $f : \mathbb{R}^2 \to \mathbb{R}$ und $f(x, y) = x^2 + y^2$, wobei der Definitionsbereich von f ein Kreis mit dem Mittelpunkt $(0, 0)$ und dem Radius R sein soll, also $D(f) := \{(x, y) : x^2 + y^2 \leq R^2\}$.
Wir stellen jetzt f in Polarkoordinaten dar. Für

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

und

$$x = r \cdot \cos\varphi := g_1(r,\varphi), \quad y = r \cdot \sin\varphi := g_2(r,\varphi)$$

ist

$$f(x,y) = f(g_1(r,\varphi), g_2(r,\varphi)) = \Big(g_1(r,\varphi)\Big)^2 + \Big(g_2(r,\varphi)\Big)^2 =$$

$$r^2\cos^2\varphi + r^2\sin^2\varphi = r^2(\underbrace{\cos^2\varphi + \sin^2\varphi}_{=1}) = r^2.$$

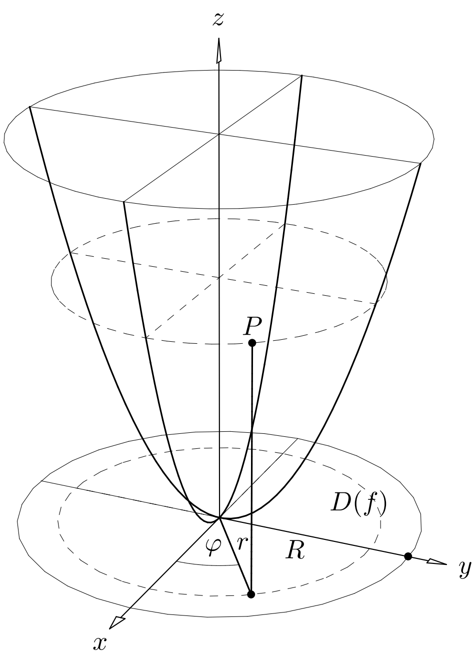

Abb. 8.8 a

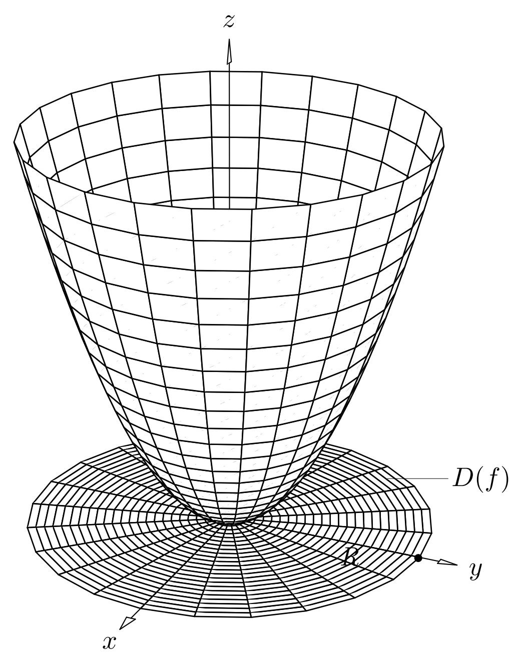

Abb. 8.8 b

In der Abb. 8.8 a ist der Punkt $P = (x, y, f(x,y))$ mit Hilfe von Polarkoordinaten dargestellt. In diesem Koordinatensystem ist P durch (r, φ, r^2) gegeben.
Abb. 8.8 b verdeutlicht den Zusammenhang zwischen Definitionsbereich und Graph der Funktion $z = f(x,y)$ in Polarkoordinaten. In beiden Fällen wurde als Definitionsbereich ein Kreis mit dem Radius R gewählt.
Man vergleiche auch Abb. 8.5, in der dieselbe Funktion dargestellt wird, wobei jedoch der Definitionsbereich ein Rechteck ist.

Insgesamt haben wir

$$g : \mathbb{R}^2 \to \mathbb{R}^2, \quad f : \mathbb{R}^2 \to \mathbb{R} \quad \Longrightarrow \quad f \circ g : \mathbb{R}^2 \to \mathbb{R}.$$

Also

$$(f \circ g)(r,\varphi) = f(g(r,\varphi)) = f(g_1(r,\varphi), g_2(r,\varphi)) = r^2.$$

Die Ableitung der verketteten Funktion ergibt sich wie folgt:

$$(f \circ g)'(r, \varphi) = f'(g(r, \varphi)) \cdot g'(r, \varphi),$$

$$f'(x, y) = \Big(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\Big)(x, y) = (2x, 2y),$$

$$g'(r, \varphi) = \begin{pmatrix} \frac{\partial g_1}{\partial r} & \frac{\partial g_1}{\partial \varphi} \\ \frac{\partial g_2}{\partial r} & \frac{\partial g_2}{\partial \varphi} \end{pmatrix} (r, \varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}.$$

Folglich gilt

$$(f \circ g)'(r, \varphi) = f'(g(r, \varphi)) \cdot g'(r, \varphi) =$$

$$(2r\cos\varphi, 2r\sin\varphi) \cdot \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} =$$

$$(2r\cos\varphi \cdot \cos\varphi + 2r\sin\varphi \cdot \sin\varphi, -2r\sin\varphi \cdot r\sin\varphi + 2r\sin\varphi \cdot r\cos\varphi) =$$

$$(2r, 0).$$

Wir berechnen jetzt die Determinante der Funktionalmatrix von $g(r, \varphi)$.

$$\det\Big(\frac{\partial(g_1, g_2)}{\partial(r, \varphi)}(r, \varphi)\Big) = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r.$$

Bemerkung. (ohne Beweis) **8/1/35/3**

Sei $g : \mathbb{R}^n \to \mathbb{R}^n$, $\bar{c} \in \mathbb{R}^n$, $g(\bar{x}) = g(x_1, \ldots, x_n)$, und $g := (g_1, \ldots, g_n)$.
Ist die Determinante der Funktionalmatrix von g in einer Umgebung $U(\bar{c})$ von null verschieden, also

$$\det\Big(\frac{\partial(g_1, \ldots, g_n)}{\partial(x_1, \ldots, x_n)}(\bar{x})\Big) \neq 0 \quad \text{für alle } \bar{x} \in U(\bar{c}),$$

dann besitzt g in $U(\bar{c})$ eine Umkehrfunktion.
Speziell für unsere Transformationsfunktion $g : \mathbb{R}^2 \to \mathbb{R}^2$ gilt dann

$$\det\Big(\frac{\partial(g_1, g_2)}{\partial(r, \varphi)}(r, \varphi)\Big) \neq 0 \quad \Longleftrightarrow \quad r \neq 0.$$

Die Koordinatentransformation ist demnach außer im Punkt $(0, 0)$ injektiv.

(3) Transformation in Zylinderkoordinaten **8/1/35/4**

Analog wie im Beispiel (2) werden jetzt räumliche Koordinaten transformiert.
Dazu sei $P = (a, b, c) \neq (0, 0, 0)$ ein Punkt im Raum $\mathbb{R}^3$, der mit einem kartesischen

Koordinatensystem versehen ist, dessen Achsen mit x, y bzw. z bezeichnet werden. Dann ist der Punkt (a, b, c) eindeutig durch das Tripel (r, φ, c) darstellbar, wobei r und φ die Polarkoordinaten des Punktes (a, b) in der Ebene $\mathbb{R}^2$ sind und c bei der Transformation unverändert bleibt.

Damit ist der gleiche Punkt P im Raum $\mathbb{R}^3$ durch unterschiedliche Koordinatensysteme eindeutig beschrieben worden (vgl. Abb. 8.9).
Die neuen Koordinaten heißen *Zylinderkoordinaten*.
(Der Nullpunkt ist analog wie im vorhergehenden Beispiel mit Hilfe der Zylinderkoordinaten nicht eindeutig darstellbar.)

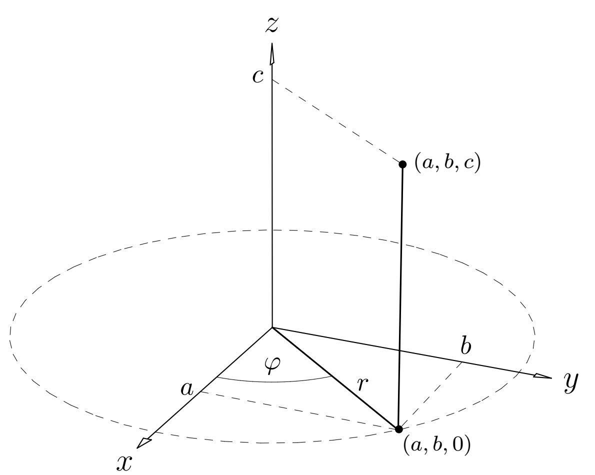

Abb. 8.9 Der Punkt $(a, b, c) \in \mathbb{R}^3$ ist mit Hilfe von Zylinderkoordinaten dargestellt. r bezeichnet den Abstand zwischen $\bar{0} := (0, 0, 0)$ und $(a, b, 0)$.
Der Punkt (a, b, c) besitzt die Zylinderkoordinaten (r, φ, c), wobei (r, φ) die Polarkoordinaten von (a, b) sind und c unverändert bleibt.

Es gilt also

$$a = r \cdot \cos\varphi := g_1(r, \varphi, z),$$

$$b = r \cdot \sin\varphi := g_2(r, \varphi, z),$$

$$c = c.$$

Betrachtet man den Punkt P als variabel, $(a, b, c) := (x, y, z)$, dann erhält man eine Abbildung

$$g(r, \varphi, z) := \Big(g_1(r, \varphi, z), g_2(r, \varphi, z), g_3(r, \varphi, z)\Big) = (x, y, z),$$

also

$$g : \mathbb{R}^3 \to \mathbb{R}^3 \quad \text{mit} \quad 0 \le \varphi < 2\pi, \quad 0 \le r < \infty \quad \text{und} \quad z \in \mathbb{R}.$$

Die Ableitung der Funktion g ergibt sich wie folgt:

$$g'(r, \varphi, z) = \frac{\partial(g_1, g_2, g_3)}{\partial(r, \varphi, z)}(r, \varphi, z) = \begin{pmatrix} \cos\varphi & -r\sin\varphi & 0 \\ \sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Damit ist

$$\det\left(\frac{\partial(g_1,g_2,g_3)}{\partial(r,\varphi,z)}(r,\varphi,z)\right)=r.$$

Für $r \neq 0$ ist die Transformation injektiv.

(4) Transformation in Kugelkoordinaten (oder sphärische Koordinaten) **8/1/35/5**

Wir transformieren hierbei wieder räumliche Koordinaten.
Dazu sei $P=(a,b,c)\neq(0,0,0):=\bar{0}$ ein Punkt in dem Raum $\mathbb{R}^3$, der mit dem kartesischen Koordinatensystem aus Beispiel (3) versehen ist.
Der Punkt (a,b,c) wird erneut durch ein Koordinatentripel (r,φ,ϑ) beschrieben, deren Bedeutung aus der Abbildung 8.10 hervorgeht.

r gibt den Abstand zwischen $\bar{0}$ und P an.
$P' := (a,b,0)$ ist die Projektion von P auf die (x,y)-Ebene, und φ bezeichnet den Winkel, der durch die x-Achse und die Verbindungsstrecke zwischen den Punkten $\bar{0}$ und P' aufgespannt wird.

Offenbar ist dann

$$c=r\sin\vartheta \quad \text{und} \quad r'=r\cos\vartheta, \quad \text{wobei} \quad -\frac{\pi}{2}\leq\vartheta\leq\frac{\pi}{2}.$$

Folglich ist

$$a=r'\cdot\cos\varphi=r\cos\varphi\cos\vartheta:=g_1(r,\varphi,\vartheta),$$

$$b=r'\cdot\sin\varphi=r\sin\varphi\cos\vartheta:=g_2(r,\varphi,\vartheta),$$

$$c=r\sin\vartheta:=g_3(r,\varphi,\vartheta).$$

Die neuen Koordinaten r, φ, ϑ heißen *Kugelkoordinaten* oder auch *sphärische Koordinaten.*

(Die Punkte auf der z-Achse sind analog wie im vorhergehenden Beispiel mit Hilfe der Kugelkoordinaten nicht eindeutig darstellbar.)

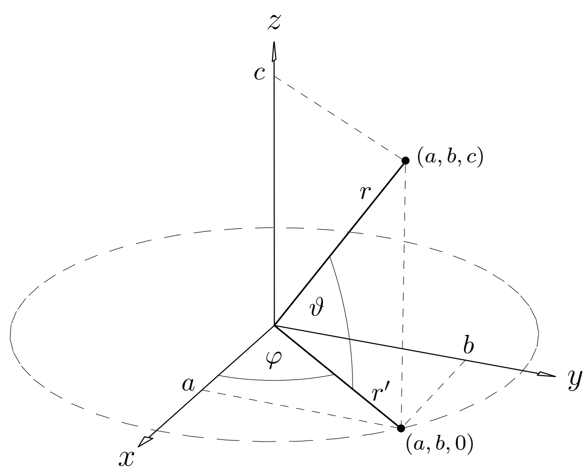

Abb. 8.10 Der Punkt $(a,b,c) \in \mathbb{R}^3$ ist mit Hilfe von Kugelkoordinaten dargestellt. r' und r bezeichnen die Abstände von $\bar{0} := (0,0,0)$ und $(a,b,0)$ bzw. von $\bar{0}$ und (a,b,c).
ϑ ist der Winkel zwischen der (x,y)-Ebene und der durch die Punkte $\bar{0}$ und (a,b,c) verlaufenden Geraden.
Der Punkt (a,b,c) besitzt die Kugelkoordinaten (r,φ,ϑ).

Betrachtet man den Punkt P als variabel, $(a,b,c) := (x,y,z)$, dann erhält man eine Abbildung

$$g(r,\varphi,\vartheta) := \Big(g_1(r,\varphi,\vartheta), g_2(r,\varphi,\vartheta), g_3(r,\varphi,\vartheta)\Big) = (x,y,z),$$

also

$$g : \mathbb{R}^3 \to \mathbb{R}^3 \quad \text{mit} \quad 0 \le \varphi < 2\pi, \;\; 0 \le r < \infty \quad \text{und} \quad \frac{-\pi}{2} \le \vartheta \le \frac{\pi}{2}.$$

Die Ableitung der Funktion g ergibt sich wie folgt:

$$g'(r,\varphi,\vartheta) = \frac{\partial(g_1,g_2,g_3)}{\partial(r,\varphi,\vartheta)}(r,\varphi,\vartheta) = \begin{pmatrix} \frac{\partial g_1}{\partial r} & \frac{\partial g_1}{\partial \varphi} & \frac{\partial g_1}{\partial \vartheta} \\ \frac{\partial g_2}{\partial r} & \frac{\partial g_2}{\partial \varphi} & \frac{\partial g_2}{\partial \vartheta} \\ \frac{\partial g_3}{\partial r} & \frac{\partial g_3}{\partial \varphi} & \frac{\partial g_3}{\partial \vartheta} \end{pmatrix}(r,\varphi,\vartheta) =$$

$$\begin{pmatrix} \cos\varphi \cdot \cos\vartheta & -r\sin\varphi \cdot \cos\vartheta & -r\cos\varphi \cdot \sin\vartheta \\ \sin\varphi \cdot \cos\vartheta & r\cos\varphi \cdot \cos\vartheta & -r\sin\varphi \cdot \sin\vartheta \\ \sin\vartheta & 0 & r\cos\vartheta \end{pmatrix}.$$

Damit ist

$$\det\left(\frac{\partial(g_1,g_2,g_3)}{\partial(r,\varphi,\vartheta)}(r,\varphi,\vartheta)\right) = r^2\cos\vartheta,$$

also

$$\det\left(\frac{\partial(g_1,g_2,g_3)}{\partial(r,\varphi,\vartheta)}(r,\varphi,\vartheta)\right) \ne 0 \quad \Longleftrightarrow \quad r \ne 0 \quad \text{und} \quad \vartheta \ne \pm\frac{\pi}{2}.$$

Für alle Punkte, die nicht auf der z-Achse liegen, ist die Transformation injektiv.

8.2 Partielle Ableitungen und Differentiale höherer Ordnung

Es sei $f(\bar{x}) : \mathbb{R}^n \to \mathbb{R}$ in einer Umgebung eines Punktes nach allen Variablen partiell **8/2/0**
differenzierbar. Dann entstehen offenbar beim partiellen Differenzieren neue Funktionen $\frac{\partial f}{\partial x_i}(\bar{x}) := \frac{\partial f(\bar{x})}{\partial x_i}$, $i = 1, \ldots, n$, die ebenfalls von $\bar{x}$ abhängen. Diese partiellen Ableitungen können wieder nach gewissen Variablen partiell differenzierbar sein, etwa nach der Variablen x_j.

Bildet man $\frac{\partial}{\partial x_j}\Big(\frac{\partial f(\bar{x})}{\partial x_i}\Big)$, dann erhält man die *2. partielle Ableitung* von f nach x_i und x_j in $\bar{x}$.

Bez.: $\frac{\partial}{\partial x_j}\Big(\frac{\partial f(\bar{x})}{\partial x_i}\Big) := \frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j} = f_{x_i x_j}(\bar{x})$

Für $i = j$ schreibt man $\frac{\partial}{\partial x_j}\Big(\frac{\partial f(\bar{x})}{\partial x_i}\Big) := \frac{\partial^2 f(\bar{x})}{\partial x_i^2} = f_{x_i x_i}(\bar{x})$. **8/2/1**

Ist $i \neq j$, dann nennt man die 2. partiellen Ableitungen auch *gemischte partielle Ableitungen.*

Nach dem gleichen Muster definiert man induktiv die *n-ten partiellen Ableitungen.* Hierfür benutzt man die Bezeichnung

$$\frac{\partial^n f(\bar{x})}{\partial x_{i_1} \ldots \partial x_{i_n}} = f_{x_{i_1} \ldots x_{i_n}}(\bar{x}), \quad \text{wobei} \quad i_1, \ldots, i_n \in \{1, \ldots, n\}.$$

Satz 8.9 (*Satz von Schwarz*) **8/2/2**

Es sei $f(x,y) : \mathbb{R}^2 \to \mathbb{R}$ und $\bar{c} \in \mathbb{R}^2$.
Ist f in einer Umgebung $U(\bar{c})$ definiert und existieren in $U(\bar{c})$ die partiellen Ableitungen f_x, f_y, f_{xy} und ist f_{xy} in $\bar{c}$ stetig, dann existiert auch f_{yx} in $\bar{c}$, und es ist $f_{xy}(\bar{c}) = f_{yx}(\bar{c})$.
(Unter den angegebenen Bedingungen sind die gemischten Ableitungen in $\bar{c}$ gleich.)

Beweis. Es sei $\bar{c} = (a,b)$ und $\bar{x} = (x,y)$. **8/2/3**

Wir zeigen, daß $f_{yx}(a,b) = \lim\limits_{x \to a} \frac{f_y(x,b) - f_y(a,b)}{x-a}$ existiert und gleich $f_{xy}(a,b)$ ist.

Nach Voraussetzung ist f in $U(\bar{c})$ partiell nach x differenzierbar. Folglich läßt sich auf f (bei festgehaltenem y) der 1. Mittelwertsatz der Differentialrechnung bezüglich x anwenden. Damit erhält man für alle $(x,y) \in U(\bar{c})$ und $x \neq a$, $y \neq b$:

$$\frac{1}{x-a} \cdot \Big(f_y(x,b) - f_y(a,b)\Big)$$

$$= \frac{1}{x-a} \cdot \Bigg(\lim_{y \to b} \frac{1}{y-b} \cdot \Big(\underbrace{f(x,y) - f(x,b)}_{:= g(x,y)}\Big) - \lim_{y \to b} \frac{1}{y-b} \cdot \Big(\underbrace{f(a,y) - f(a,b)}_{= g(a,y)}\Big)\Bigg)$$

$$= \frac{1}{x-a} \cdot \lim_{y\to b} \frac{1}{y-b} \cdot \Big(g(x,y) - g(a,y) \Big)$$

$$= \lim_{y\to b} \frac{1}{y-b} \cdot \frac{g(x,y)-g(a,y)}{x-a}$$

$$= \lim_{y\to b} \frac{1}{y-b} \cdot g_x(\underbrace{a+\vartheta(x-a)}_{:=u}, y) \qquad \text{(1. Mittelwertsatz, } y \text{ fest, } 0<\vartheta<1)$$

$$= \lim_{y\to b} \frac{1}{y-b} \cdot \Big(f_x(u,y) - f_x(u,b) \Big)$$

$$= f_{xy}(u,b). \qquad (f_{xy} \text{ existiert in } U(\bar{c}))$$

Da nach Voraussetzung f_{xy} in $\bar{c}$ stetig ist, existieren die folgenden Limites und es gilt:

$$f_{yx}(a,b) = \lim_{x\to a} \frac{1}{x-a} \cdot \Big(f_y(x,b) - f_y(a,b) \Big) = \lim_{x\to a} f_{xy}(u,b) = f_{xy}(a,b). \qquad ❏$$

Bemerkung. **8/2/4**

Ist $f(x_1,\ldots,x_n) : \mathbb{R}^n \to \mathbb{R}$ und $\bar{c} \in \mathbb{R}^n$ und existieren in einer Umgebung $U(\bar{c})$ die partiellen Ableitungen f_{x_i}, f_{x_j} und $f_{x_i x_j}$ und ist $f_{x_i x_j}$ in $\bar{c}$ stetig, dann existiert auch $f_{x_j x_i}$ in $\bar{c}$, und es ist $f_{x_i x_j}(\bar{c}) = f_{x_j x_i}(\bar{c})$.

Den Beweis hierzu führt man leicht auf den vorhergehenden Satz zurück.

Wir befassen uns jetzt mit *Differentialen höherer Ordnung.*
Dazu sei $f(\bar{x}) : \mathbb{R}^n \to \mathbb{R}$. Das 1. Differential wurde als Funktion $df : \mathbb{R}^n \to \mathbb{R}$ definiert, die sich darstellen läßt in der Form

$$df := \frac{\partial f}{\partial x_1} \cdot dx_1 + \ldots + \frac{\partial f}{\partial x_n} \cdot dx_n = \frac{\partial f}{\partial x_1}\, dx_1 + \ldots + \frac{\partial f}{\partial x_n}\, dx_n,$$

wobei die dx_i als konstant anzusehen sind.

Wir berechnen (definieren) jetzt das *2. Differential* von f wie folgt.

$$d^2 f := d(df) = d\Big(\frac{\partial f}{\partial x_1} dx_1 + \ldots + \frac{\partial f}{\partial x_n} dx_n\Big)$$

$$= d\Big(\frac{\partial f}{\partial x_1}\Big) dx_1 + \ldots + d\Big(\frac{\partial f}{\partial x_n}\Big) dx_n \qquad (dx_i \text{ konstant !})$$

$$= \Big(\frac{\partial}{\partial x_1}\Big(\frac{\partial f}{\partial x_1}\Big) dx_1 + \ldots + \frac{\partial}{\partial x_n}\Big(\frac{\partial f}{\partial x_1}\Big) dx_n\Big) dx_1 \; + \; \ldots$$

$$+ \Big(\frac{\partial}{\partial x_1}\Big(\frac{\partial f}{\partial x_n}\Big) dx_1 + \ldots + \frac{\partial}{\partial x_n}\Big(\frac{\partial f}{\partial x_n}\Big) dx_n\Big) dx_n$$

$$= \frac{\partial}{\partial x_1}\Big(\frac{\partial f}{\partial x_1}\Big) dx_1 dx_1 + \ldots + \frac{\partial}{\partial x_n}\Big(\frac{\partial f}{\partial x_1}\Big) dx_n dx_1 \ + \ \ldots$$

$$+ \frac{\partial}{\partial x_1}\Big(\frac{\partial f}{\partial x_n}\Big) dx_1 dx_1 + \ldots + \frac{\partial}{\partial x_n}\Big(\frac{\partial f}{\partial x_n}\Big) dx_n dx_n$$

$$= \frac{\partial^2 f}{\partial x_1 \partial x_1} dx_1 dx_1 + \ldots + \frac{\partial^2 f}{\partial x_1 \partial x_n} dx_n dx_1 \ + \ \ldots$$

$$+ \frac{\partial^2 f}{\partial x_n \partial x_1} dx_1 dx_n + \ldots + \frac{\partial^2 f}{\partial x_n \partial x_n} dx_n dx_n$$

$$= \frac{\partial^2 f}{\partial x_1^2} dx_1^2 + \ldots + \frac{\partial^2 f}{\partial x_i \partial x_j} dx_j dx_i + \ldots + \frac{\partial^2 f}{\partial x_j \partial x_i} dx_i dx_j + \ldots + \frac{\partial^2 f}{\partial x_n^2} dx_n^2 \quad (i \neq j).$$

Sind die gemischten Ableitungen gleich, dann gilt

$$d^2 f = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} dx_i^2 + 2 \sum_{\substack{i,j=1 \\ i<j}}^{n} \frac{\partial^2 f}{\partial x_i \partial x_i} dx_j dx_j.$$

Analog definiert man induktiv

$$d^{n+1} f := d(d^n f).$$

(Hierbei ist der Satz von Schwarz sehr nützlich.)

Beispiel. **8/2/5**

Sei $f(x,y) = x^2 + y^2$. Dann ist

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y \ \text{ und } \ \frac{\partial^2 f}{\partial x^2} = 2, \ \frac{\partial^2 f}{\partial x \partial y} = 0, \ \frac{\partial^2 f}{\partial y \partial x} = 0, \frac{\partial^2 f}{\partial y^2} = 2.$$

Folglich ist

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy = 2x\, dx + 2y\, dy.$$

Weiterhin ist

$$d^2 f = d(df) = \frac{\partial^2 f}{\partial x^2}\, dx^2 + 2 \frac{\partial^2 f}{\partial x \partial y}\, dx dy + \frac{\partial^2 f}{\partial y^2}\, dy^2 =$$

$$2dx^2 + 2dy^2,$$

und schließlich

$$d^3 f = 0, \ \text{(denn alle dritten partiellen Ableitungen sind null)}.$$

8.3 Der Satz von Taylor; lokale Extrema für Funktionen mit mehreren Veränderlichen

Wir werden jetzt einige Ergebnisse aus der Differentialrechnung für Funktionen mit einer Veränderlichen auf Funktionen mit mehreren Veränderlichen erweitern. Wir beginnen zunächst mit dem Mittelwertsatz, der sich auch hier als Spezialfall des Taylorschen Satzes erweist. **8/3/0**

Satz 8.10 (*Mittelwertsatz der Differentialrechnung mit mehreren Veränderlichen*) **8/3/1**
Sei $M \subseteq \mathbb{R}^n$ *offen und* $f : \mathbb{R}^n \to \mathbb{R}$ *eine in* M *differenzierbare Funktion. Weiterhin seien* $\bar{a}, \bar{b} \in M$ *und die Verbindungsstrecke* $s(\bar{a}, \bar{b})$ *zwischen* $\bar{a}$ *und* $\bar{b}$ *gehöre zu* M. *Dann gibt es ein* $\bar{c} \in s(\bar{a}, \bar{b})$ *mit* $\bar{c} \neq \bar{a}, \bar{b}$, *so daß* $f(\bar{b}) - f(\bar{a}) = f'(\bar{c}) \cdot (\bar{b} - \bar{a})$.

Beweis. Wir betrachten zunächst eine Parameterdarstellung **8/3/2**

$$g(t) := \bar{b} + t(\bar{b} - \bar{a}), \quad t \in [0,1]$$

von $s(\bar{a}, \bar{b})$ und definieren damit die Funktion $F : [0,1] \to \mathbb{R}$ mit der Eigenschaft

$$F(t) := f(g(t)).$$

Offenbar ist F in $[0,1]$ stetig und in $(0,1)$ differenzierbar. Dann läßt sich auf F der 1. Mittelwertsatz der Differentialrechnung anwenden. Folglich existiert ein $\xi \in (0,1)$, so daß

$$F'(\xi) = \frac{F(1) - F(0)}{1 - 0} = F(1) - F(0).$$

Hieraus folgt

$$\begin{aligned} F(1) - F(0) &= f(\underbrace{g(1)}_{=\bar{b}}) - f(\underbrace{g(0)}_{=\bar{a}}) = f(\bar{b}) - f(\bar{a}) \\ &= F'(\xi) = f'(\underbrace{g(\xi)}_{:=\bar{c}}) \cdot \underbrace{g'(\xi)}_{=\bar{b}-\bar{a}} \\ &= f'(\bar{c}) \cdot (\bar{b} - \bar{a}). \qquad ❏ \end{aligned}$$

Bemerkung. Für Vektorfunktionen ist der Mittelwertsatz im allgemeinen falsch. **8/3/3**
Man betrachte das Beispiel $f : [0,1] \to \mathbb{R}^2$ mit $f(x) = (x^2, x^3)$.
Offenbar gibt es kein $c \in (0,1)$, so daß $f(1) - f(0) = f'(c) \cdot (1 - 0) = (2c, 3c^2)$.

Definition. (*polygonzusammenhängend; Gebiet*) **8/3/4**
Sei $M \subseteq \mathbb{R}^n$.

(1) M ist *polygonzusammenhängend*
$\underset{\text{Df}}{\equiv}$ Zu je zwei Punkten $\bar{a}, \bar{b} \in M$ gibt es endlich viele Elemente $\bar{a}_0 = \bar{a}, \bar{a}_1, \ldots, \bar{a}_{m+1} = \bar{b}$, so daß die Verbindungsstrecken $s(\bar{a}_i, \bar{a}_{i+1})$ zwischen $\bar{a}_i$ und $\bar{a}_{i+1}$, $i = 1, \ldots, m$, stets zu M gehören.
(Da durch Polygonzüge stetige Funktionen gegeben sind, sind polygonzusammenhängende Mengen auch bogenzusammenhängend.)

(2) M ist ein *Gebiet*
$\underset{\text{Df}}{\equiv}$ M ist polygonzusammenhängend und offen.

Satz 8.11 *Sei $M \subseteq \mathbb{R}^n$ ein Gebiet und $f : M \to \mathbb{R}$ in M differenzierbar.* **8/3/5**
Dann gilt: *f ist konstant in M $\iff$ $f'(\bar{x}) = 0$ für alle $\bar{x} \in M$.*

Beweis. ($\longrightarrow$) trivial, da die partiellen Ableitungen von konstanten Funktionen null sind. **8/3/6**

($\longleftarrow$) Es sei $f'(\bar{x}) = 0$ für jedes $\bar{x} \in M$ und es seien $\bar{a}, \bar{b} \in M$. Dann gibt es nach Voraussetzung Elemente $\bar{a}_0 = \bar{a}, \bar{a}_1, \ldots, \bar{a}_{m+1} = \bar{b}$ in M, so daß $s(\bar{a}_i, \bar{a}_{i+1}) \in M$ für $i = 0, \ldots, m$.

Nach dem Mittelwertsatz existiert stets ein $\bar{c}_i \in s(\bar{a}_i, \bar{a}_{i+1})$ mit $\bar{c}_i \neq \bar{a}_i, \bar{a}_{i+1}$, so daß

$$f(\bar{a}_{i+1}) - f(\bar{a}_i) = \underbrace{f'(\bar{c}_i)}_{=0} \cdot (\bar{a}_{i+1} - \bar{a}_i) = 0,$$

also ist $f(\bar{a}_{i+1}) = f(\bar{a}_i)$ für alle i und damit $f(\bar{a}) = f(\bar{b})$ für beliebige $\bar{a}, \bar{b} \in M$.
❑

Beispiel. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$, wobei **8/3/7**

$$f(x,y) = \arctan \frac{x}{y} + \arctan \frac{y}{x} \quad \text{und} \quad D(f) = \{(x,y) : x \neq 0,\ y \neq 0\}.$$

Folglich ist f in den Gebieten

$$M_1 := \{(x,y) : x > 0,\ y > 0\},$$
$$M_2 := \{(x,y) : x > 0,\ y < 0\},$$
$$M_3 := \{(x,y) : x < 0,\ y > 0\},$$
$$M_4 := \{(x,y) : x < 0,\ y < 0\}$$

definiert und offenbar differenzierbar. In jedem der Gebiete gilt:

$$\frac{\partial f(x,y)}{\partial x} = \frac{1}{1 + \left(\frac{x}{y}\right)^2} \cdot \left(\frac{1}{y}\right) + \frac{1}{1 + \left(\frac{y}{x}\right)^2} \cdot \left(-\frac{y}{x^2}\right)$$

$$= \frac{y^2}{(y^2 + x^2) \cdot y} - \frac{x^2 y}{(x^2 + y^2) \cdot x^2} = 0.$$

$$\frac{\partial f(x,y)}{\partial y} = \frac{1}{1+\left(\frac{x}{y}\right)^2}\cdot\Big(-\frac{x}{y^2}\Big) + \frac{1}{1+\left(\frac{y}{x}\right)^2}\cdot\Big(\frac{1}{x}\Big) = 0.$$

Folglich ist $f'(\bar{x}) = 0$, und damit ist f in jedem der M_i konstant. Die Werte von f kann man leicht durch geeignete spezielle Argumente ermitteln.

In M_1 ist $f(x,y) = f(1,1) = 2\arctan 1 = \frac{\pi}{2}$,
in M_2 ist $f(x,y) = f(1,-1) = 2\arctan(-1) = -\frac{\pi}{2}$,
in M_3 ist $f(x,y) = f(-1,1) = 2\arctan(-1) = -\frac{\pi}{2}$,
in M_4 ist $f(x,y) = f(-1,-1) = 2\arctan 1 = \frac{\pi}{2}$.

Es soll jetzt der Taylorsche Satz für Funktionen mit mehreren Veränderlichen verallgemeinert werden. Dies erfordert einen gewissen technischen Aufwand, den wir durch eine geeignete Schreibweise etwas reduzieren wollen. **8/3/8**

Definition. Es sei $M \subseteq \mathbb{R}^n$ und $f : \mathbb{R}^n \to \mathbb{R}$ in M definiert. **8/3/9**

(1) f heißt in M *stetig differenzierbar*
$\underset{\text{Df}}{=}$ f ist in M differenzierbar und f' ist in M stetig.
(Dies ist nach dem Satz 8.2 genau dann der Fall, wenn alle partiellen Ableitungen von f in M vorhanden und stetig sind.)

(2) f ist in M $(k+1)$-*mal stetig differenzierbar*
$\underset{\text{Df}}{=}$ $f^{(k)}$ ist in M stetig differenzierbar.
Bez.: $f \in C^{k+1}(M)$.

Es sei jetzt M eine offene Teilmenge von $\mathbb{R}^2$, $f : \mathbb{R}^2 \to \mathbb{R}$ und $f \in C^{k+1}(M)$. **8/3/10** Weiterhin seien $\bar{a} = (a_1, a_2)$, $\bar{c} = \bar{a} + t\cdot\bar{h}$, wobei $t \in [0,1]$, $\bar{h} = (h_1, h_2)$ und die Verbindungsstrecke $s(\bar{a}, \bar{b})$ ganz zu M gehöre. Dann ist

$$\varphi(t) := f(\bar{a} + t\bar{h}) = f(a_1 + th_1, a_2 + fh_2) \ \text{ für } \ 0 \le t \le 1$$

eine reellwertige Funktion einer reellen Veränderlichen, deren Ableitung sich gemäß der Kettenregel wie folgt berechnet

$$\varphi'(t) = \frac{\partial}{\partial x} f(\bar{c})\cdot h_1 + \frac{\partial}{\partial y} f(\bar{c})\cdot h_2.$$

Für $\dfrac{\partial}{\partial x}$ bzw. $\dfrac{\partial}{\partial y}$ schreiben wir kurz D_1 bzw. D_2. Damit ergibt sich

$$\varphi'(t) = h_1\cdot D_1 f + h_2\cdot D_2 f = (h_1 D_1 + h_2 D_2) f,$$

wobei das Argument von f der Einfachheit halber weggelassen wurde.

Für $n = 2$ ist dann

$$\varphi''(t) = h_1(h_1 D_1 D_1 f + h_2 D_2 D_1 f) + h_2(h_1 D_1 D_2 f + h_2 D_2 D_2 f).$$

Nach dem Satz von Schwarz ist $D_1 D_2 f = D_2 D_1 f$ und somit erhält man für $D_i D_i := D_i^2,\ i = 1, 2,$

$$\varphi''(t) = h_1^2 D_1^2 f + 2h_1 h_2 D_1 D_2 f + h_2^2 D_2^2 f.$$

In Analogie zur binomischen Formel schreiben wir für $h_1^2 D_1^2 f + 2h_1 h_2 D_1 D_2 f + h_2^2 D_2^2 f$ im folgenden auch $(h_1 D_1 + h_2 D_2)^{(2)} f$.

Analog erhält man für $\varphi^{(k)}(t),\ 0 \le t \le 1,$ die Darstellung

$$\varphi^{(k)}(t) = (h_1 D_1 + h_2 D_2)^{(k)} f.$$

(Beweis induktiv über k)

Ist $M \subseteq \mathbb{R}^n$ offen, $f \in C^{m+1}(M)$ und sind $\bar{a} = (a_1, \dots, a_n)$, $\bar{a} + t \cdot \bar{h}$ mit $0 \le t \le 1$ und $\bar{h} = (h_1, \dots, h_n)$ Elemente aus M, deren Verbindungsstrecke ganz zu M gehört, und ist $\varphi(t) = f(\bar{a} + t\bar{h}) = f(a_1 + th_1, \dots, a_n + th_n)$, dann ist

$$\varphi'(t) = \sum_{i=1}^{n} h_i \cdot \frac{\partial}{\partial x_i} f(\bar{a} + t\bar{h}).$$

Schreibt man D_i für $\dfrac{\partial}{\partial x_i}$, so erhält man

$$\varphi'(t) = \sum_{i=1}^{n} h_i D_i f \quad \text{und} \quad \varphi''(t) = \sum_{i,j=1}^{n} h_i h_j D_i D_j f$$

$$= (h_1 D_1 + \dots + h_n D_n)^{(2)} f,$$

wenn man den Satz von Schwarz und eine der binomischen Formel (für n Summanden) analoge Schreibweise benutzt.

Induktiv zeigt man schließlich

$$\varphi^{(k)}(t) = (h_1 D_1 + \dots + h_n D_n)^{(k)} f.$$

Jetzt sind wir in der Lage, den Taylorschen Satz in übersichtlicher Weise zu formulieren.

Satz 8.12 (*Satz von Taylor für Funktionen mit n Veränderlichen*) **8/3/11**

Sei $f : \mathbb{R}^n \to \mathbb{R}$, $\bar{c} \in \mathbb{R}^n$, U eine offene Umgebung von $\bar{c}$ und $f \in C^{m+1}(U)$. Sei $\bar{x} \in U$, so daß die Verbindungsstrecke von $\bar{c}$ und $\bar{x}$ ganz zu U gehört. Für $\bar{x} - \bar{c} = (x_1 - c_1, \dots, x_n - c_n) := (h_1, \dots, h_n) = \bar{h}$ gilt dann: Es gibt ein $\vartheta \in \mathbb{R}$ mit $0 < \vartheta < 1$, so daß $f(\bar{x}) = \sum_{i=0}^{m} \frac{1}{i!}(h_1 D_1 + \dots + h_n D_n)^{(i)} f(\bar{c}) + R_m(\bar{x})$, wobei

$$R_m(\bar{x}) = \frac{1}{(m+1)!} \cdot (h_1 D_1 + \dots + h_n D_n)^{(m+1)} f(\bar{c} + \vartheta \bar{h}).$$

Beweis. Sei $\varphi(t) := f(\bar{c} + t\bar{h}),\ 0 \leq t \leq 1$. Dann ist $\varphi(t)$ als reellwertige Funktion einer reellen Veränderlichen offenbar $(m+1)$-mal differenzierbar. Nach dem Taylorschen Satz für Funktionen einer Veränderlichen gibt es ein ϑ mit $0 < \vartheta < 1$, so daß **8/3/12**

$$\begin{aligned}\varphi(1) &= \varphi(0) + \sum_{i=1}^{m} \frac{\varphi^{(i)}(0)}{i!} \cdot (1-0)^i + \frac{\varphi^{(m+1)}(0 + \vartheta(1-0))}{(m+1)!} \cdot (1-0)^{m+1} \\ &= \varphi(0) + \sum_{i=1}^{m} \frac{\varphi^{(i)}(0)}{i!} + \frac{\varphi^{(m+1)}(\vartheta)}{(m+1)!}.\end{aligned}$$

Es ist

$$\varphi(1) = f(\bar{x}), \quad \varphi(0) = f(\bar{a}), \quad \varphi^{(i)}(0) = (h_1 D_1 + \cdots + h_n D_n)^{(i)} f(\bar{c})$$

für $i = 1, \ldots, m$ und

$$\varphi^{(m+1)}(\vartheta) = (h_1 D_1 + \cdots + h_n D_n)^{(m+1)} f(\bar{c} + \vartheta \bar{h}).$$

Hieraus folgt die Behauptung. ❑

Korollar. **8/3/13**

(1) *Für $m = 0$ liefert der Satz von Taylor* (wie für Funktionen mit einer Veränderlichen) *den Mittelwertsatz als Spezialfall.*

(2) *Für $m = 1$, $n = 2$ und $\bar{a} = (a, b)$, $\bar{x} = (x, y)$, $\bar{h} = \bar{x} - \bar{a} = (x - a, y - b)$ und $\bar{u} := \bar{a} + \vartheta \bar{h}$ erhält man*

$$\begin{aligned}f(x,y) = f(a,b) &+ f_x(\bar{a}) \cdot (x-a) + f_y(\bar{a}) \cdot (y-b) + \\ &\frac{1}{2!}\Big(f_{xx}(\bar{u}) \cdot (x-a)^2 + 2 f_{xy}(\bar{u}) \cdot (x-a)(y-b) + f_{yy}(\bar{u}) \cdot (y-b)^2\Big).\end{aligned}$$

(3) *Gilt zusätzlich zu den Voraussetzungen des Satzes 8.12, daß $f \in C^\infty(M)$* (d.h., f ist in M beliebig oft differenzierbar)*, dann läßt sich f in eine Potenzreihe* (mit mehreren Veränderlichen) *entwickeln.*

Wenn $R_m(\bar{x}) \xrightarrow[m \to \infty]{} 0$, *so ist* $f(\bar{x}) = \sum_{i=0}^{\infty} (h_1 D_1 + \cdots + h_n D_n)^{(i)} f(\bar{a})$.

Beweis. (1) und (2) sind trivial; (3) zeigt man wie im eindimensionalen Fall. ❑ **8/3/14**

Beispiel. Sei $f(x,y) = \mathrm{e}^{x+y}$ und $\bar{a} = (a,b) = (0,0)$. **8/3/15**
Dann ist $D_1 f(x,y) = \mathrm{e}^{x+y}$ und $D_2 f(x,y) = \mathrm{e}^{x+y}$. Folglich ist f beliebig oft stetig partiell differenzierbar und es ist $D_i^k D_j^m f(x,y) = \mathrm{e}^{x+y}$ und somit insbesondere $D_i^k D_j^m f(0,0) = 1$ für $i, j \in \{1, 2\}$.

Wegen $h_1 = x - 0 = x$, $h_2 = y - 0 = y$ und $(h_1 D_1 + h_2 D_2)^{(m)} f(0,0) = (x+y)^m$ gilt

$$\mathrm{e}^{x+y} = \sum_{m=0}^{\infty} \frac{1}{m!} \cdot (x+y)^m = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{i=0}^{m} \binom{m}{i} x^i y^{m-i} = \sum_{m=0}^{\infty} \sum_{i+j=m}^{m} \frac{1}{i!j!} \cdot x^i y^j .$$

(Man hätte diese Reihe natürlich auch anders gewinnen können.)

Wir befassen uns jetzt mit lokalen Extrema bei Funktionen mit zwei und mehr Veränderlichen. **8/3/16**

In den Sätzen 7.15 und 7.16 sind (für differenzierbare Funktionen mit einer Veränderlichen) gewisse Bedingungen für das Vorliegen eines lokalen Extremums an einer Stelle c angegeben worden; und zwar eine notwendige Bedingung: $f'(c) = 0$ und eine hinreichende Bedingung: $f''(c) \neq 0$.

Ähnliche, wenn auch kompliziertere, Bedingungen gibt es auch für Funktionen mit zwei (und mehr) Veränderlichen, mit denen wir uns jetzt befassen.

Definition. (*lokales Extremum*) **8/3/17**

Sei $f : \mathbb{R}^n \to \mathbb{R}$ und $\bar{c}$ ein innerer Punkt von $D(f)$.
f besitzt an der Stelle $\bar{c}$ ein *relatives* oder *lokales Extremum* (:= *lokales Minimum* bzw. *lokales Maximum*)
$\underset{\mathrm{Df}}{=}$ Es gibt eine Umgebung $U(\bar{c})$, so daß für jedes $\bar{x} \in U(\bar{c})$ mit $\bar{x} \neq \bar{c}$ gilt:
$f(\bar{x}) > f(\bar{c})$ für ein lokales Minimum und
$f(\bar{x}) < f(\bar{c})$ für ein lokales Maximum.

Satz 8.13 (*Notwendige Bedingung für die Existenz eines lokalen Extremums*) **8/3/18**

Sei $f(\bar{x}) : \mathbb{R}^n \to \mathbb{R}$ in einer Umgebung von $\bar{c}$ definiert und in $\bar{c}$ nach allen Variablen partiell differenzierbar.
Besitzt f in $\bar{c}$ ein lokales Extremum, dann sind alle (ersten) *partiellen Ableitungen von f in $\bar{c}$ null.*
(Wenn $\dfrac{\partial f}{\partial x_i}(\bar{c}) = 0$ für $i = 1, \ldots, n$, also $\operatorname{grad} f(\bar{c}) = \bar{0}$, dann heißt $\bar{c}$ auch *kritischer* oder *stationärer Punkt* von f.)

Beweis. Habe o.B.d.A. f in $\bar{c}$ ein lokales Minimum (für ein lokales Maximum verläuft der Beweis analog). **8/3/19**
Dann gibt es eine Umgebung $U(\bar{c})$, so daß für jedes $\bar{x} \in U(\bar{c})$ mit $\bar{x} \neq \bar{c}$ gilt: $f(\bar{x}) > f(\bar{c})$. Dies gilt insbesondere für $\bar{x} := \bar{c} + h\bar{e}_i$, wenn h hinreichend klein ist.
Nach Voraussetzung ist die Funktion $\varphi(h) := f(\bar{c} + h\bar{e}_i)$ (als Funktion von h) in $h = 0$ differenzierbar, und es ist $\varphi'(0) = \dfrac{\partial f}{\partial x_j}(\bar{c})$.
Offenbar besitzt φ (als Funktion einer Veränderlichen) in 0 ein lokales Minimum. Folglich gilt nach Satz 7.15:

$$\varphi'(\bar{0}) = 0 = \frac{\partial f}{\partial x_i}(\bar{c}) \quad \text{für } i = 1, \ldots, n. \qquad \square$$

Beispiel. 8/3/20

Sei $f(x,y) = x^2 + y^2$. Wir berechnen die kritischen Stellen von f. Es ist

$$\frac{\partial f}{\partial x}(x,y) = 2x = 0 \implies x = 0$$

und

$$\frac{\partial f}{\partial y}(x,y) = 2y = 0 \implies y = 0.$$

Die einzige kritische Stelle ist $(0,0)$. Höchstens dort kann f ein lokales Extremum besitzen.

Im eindimensionalen Fall haben wir eine hinreichende Bedingung für das Vorliegen eines lokalen Extremums mit Hilfe des Taylorschen Satzes bewiesen. Analoge Überlegungen führen auch bei Funktionen mit mehreren Veränderlichen zum Ziel. Wir beschränken uns hier auf Funktionen mit zwei Veränderlichen, da der technische Aufwand für den n-dimensionalen Fall nicht unerheblich ist. 8/3/21

Ist $\bar{c} \in \mathbb{R}^2$, U eine Umgebung von $\bar{c} = (a,b)$, f zweimal stetig differenzierbar in U und $\bar{x} = (x,y)$ hinreichend dicht bei $\bar{c}$, dann gilt nach dem Satz von Taylor (für $m = 1$, $n = 2$, $\bar{h} = \bar{x} - \bar{a} = (x-a, y-b) := (h,k)$ und $\bar{u} := \bar{a} + \vartheta\bar{h}$)

$$f(\bar{x}) - f(\bar{c}) = f_x(\bar{c})\cdot h + f_y(\bar{c})\cdot k + \frac{1}{2}\Big(f_{xx}(\bar{u})\cdot h^2 + 2f_{xy}(\bar{u})\cdot hk + f_{yy}(\bar{u})\cdot k^2\Big). \quad (\star)$$

Sind die notwendigen Bedingungen für das Vorliegen eines lokalen Extremums erfüllt, d.h., $f_x(\bar{c}) = f_y(\bar{c}) = 0$, dann erhält man

$$f(\bar{x}) - f(\bar{c}) = \frac{1}{2}\Big(h^2\cdot f_{xx}(\bar{u}) + 2hk\cdot f_{xy}(\bar{u}) + k^2\cdot f_{yy}(\bar{u})\Big) = R_1(\bar{x}).$$

Ob f an der Stelle $\bar{c}$ ein lokales Extremum besitzt, hängt allein von dem Restglied ab. Aufgrund der Stetigkeit der zweiten partiellen Ableitungen in U wechseln diese in einer hinreichend kleinen Umgebung von $\bar{c}$ ihr Vorzeichen nicht, falls sie an der Stelle $\bar{c}$ von null verschieden sind. Dies nutzen wir aus, um ein handhabbares Kriterium zur Verfügung zu haben.

Satz 8.14 (*Hinreichende Bedingung für die Existenz eines lokalen Extremums*) 8/3/22

Es sei $f : \mathbb{R}^2 \to \mathbb{R}$, $\bar{c} \in \mathbb{R}^2$ und f sei in einer Umgebung von $\bar{c}$ zweimal stetig differenzierbar.

Weiterhin sei $\bar{c}$ ein kritischer Punkt von f und $D = \begin{vmatrix} f_{xx}(\bar{c}) & f_{xy}(\bar{c}) \\ f_{xy}(\bar{c}) & f_{yy}(\bar{c}) \end{vmatrix}$.

Dann gilt:

(1) *Ist $D > 0$, dann besitzt f in $\bar{c}$ ein lokales Extremum, und zwar ein lokales Minimum, falls $f_{xx}(\bar{c}) > 0$ und ein lokales Maximum, falls $f_{xx}(\bar{c}) < 0$.*

(2) *Ist* $D < 0$, *dann besitzt* f *in* $\bar{c}$ *einen sog. Sattelpunkt* (das ist ein kritischer Punkt, in dem die betrachtete Funktion weder ein lokales Minimum noch ein lokales Maximum besitzt; vgl. Abb. 8.11).

(3) *Ist* $D = 0$, *dann läßt sich* (allein mit Hilfe der zweiten partiellen Ableitungen) *noch keine Aussage treffen.*

Beweis. Wir benutzen die gleichen Bezeichnungen wie in der obigen Formel ($\star$). Da nach Voraussetzung $f_x(\bar{c})$ und $f_y(\bar{c})$ null sind, erhält man aus ($\star$) **8/3/23**

$$f(\bar{x}) - f(\bar{c}) = \frac{1}{2}\Big(h^2\cdot f_{xx}(\bar{u}) + 2hk\cdot f_{xy}(\bar{c}) + k^2\cdot f_{yy}(\bar{u})\Big) = R_1(\bar{x}).$$

(1). Nach Voraussetzung ist

$$D := D(\bar{c}) = f_{xx}(\bar{c})\cdot f_{yy}(\bar{c}) - f_{xy}^2(\bar{c}) > 0.$$

Betrachtet man $D(\bar{u})$ als Funktion von $\bar{u} = \bar{c} + \vartheta(\bar{x} - \bar{c})$, dann ist wegen $f \in C^2(U)$ die Funktion $D(\bar{u})$ in U stetig und somit $D(\bar{u}) > 0$, falls $\bar{x}$ hinreichend nahe bei $\bar{c}$ liegt.

Analog gilt für $f_{xx}(\bar{c}) \lessgtr 0$ auch $f_{xx}(\bar{u}) \lessgtr 0$.

Im folgenden schreiben wir für $f_{xx}(\bar{u})$, $f_{xy}(\bar{u})$, $f_{yy}(\bar{u})$ kurz f_{xx}, f_{xy}, f_{yy}.

Da f_{xx} nicht null ist, erhält man

$$\begin{aligned} f(\bar{x}) - f(\bar{c}) &= \frac{1}{2f_{xx}}\cdot\Big[h^2 f_{xx}^2 + hk f_{xx} f_{xy} + k^2 f_{xx} f_{yy}\Big] \\ &= \frac{1}{2f_{xx}}\Big[\underbrace{(hf_{xx} + kf_{yy})^2}_{\geq 0} + k^2\underbrace{(f_{xx}f_{yy} - f_{xy}^2)}_{>0}\Big]. \end{aligned}$$

Da der Ausdruck in den eckigen Klammern für $k \neq 0$ positiv ist, hängt das Vorzeichen von $f(\bar{x}) - f(\bar{c})$ nur von $f_{xx}(\bar{c})$ ab. (Für $k = 0$ gilt nach ($\star$) schon $f(\bar{x}) - f(\bar{c}) = \frac{h^2}{2}\cdot f_{xx}$.)

Also für $f_{xx}(\bar{c}) > 0$ besitzt f an der Stelle $\bar{c}$ ein lokales Minimum und für $f_{xx}(\bar{c}) < 0$ ein lokales Maximum.

(2). Sei $D < 0$. Setzt man $h = k$ bzw. $h = -k$, dann erhält man

$$f(\bar{x}) - f(\bar{c}) = \frac{h^2}{2}\cdot(f_{xx} + 2f_{xy} + f_{yy}) \quad \text{bzw.}$$

$$f(\bar{x}) - f(\bar{c}) = \frac{h^2}{2}\cdot(f_{xx} - 2f_{xy} + f_{yy}).$$

Es sei zunächst $f_{xx}(\bar{c}) = f_{yy}(\bar{c}) = 0$.

Wegen $D(\bar{c}) < 0$ ist dann $f_{xy}(\bar{c}) \neq 0$ und somit

$$f(\bar{x}) - f(\bar{c}) \xrightarrow[\bar{x}\to\bar{c}]{} 2f_{xy}(\bar{c}), \text{ falls } h = k \quad \text{und}$$

$$f(\bar{x}) - f(\bar{c}) \xrightarrow[\bar{x}\to\bar{c}]{} -2f_{xy}(\bar{c}), \text{ falls } h = k.$$

Dies bedeutet, daß in jeder Umgebung von $\bar{c}$ sowohl positive als auch negative Werte von $f(\bar{x}) - f(\bar{c})$ auftreten. Folglich besitzt f in $\bar{c}$ kein lokales Extremum.

Es bleibt noch der Fall zu betrachten, daß $f_{xx}(\bar{c}) \neq 0$ oder $f_{yy}(\bar{c}) \neq 0$.

Sei o.B.d.A. $f_{xx}(\bar{c}) \neq 0$. Dann erhält man für $k = 0$

$$f(\bar{x}) - f(\bar{c}) = \frac{h^2}{2} \cdot f_{xx}.$$

Für hinreichend nahe bei $\bar{c}$ gelegene $\bar{x}$ haben $f(\bar{x}) - f(\bar{c})$ und $f_{xx}(\bar{c})$ das gleiche Vorzeichen.

Sei jetzt $h = -s \cdot f_{xy}(\bar{c})$ und $k = s \cdot f_{xx}(\bar{c})$ für „kleine" $s \neq 0$. Dann gilt

$$f(\bar{x}) - f(\bar{c}) = \frac{s^2}{2}\Big(f_{xy}^2(\bar{c}) \cdot f_{xx} - 2 f_{xy}(\bar{c}) \cdot f_{xx}(\bar{c}) \cdot f_{xy} + f_{xx}^2(\bar{c}) \cdot f_{yy}\Big).$$

Wegen $f \in C^2(U)$ gilt

$$\begin{aligned} f(\bar{x}) - f(\bar{c}) &\xrightarrow[\bar{x}\to\bar{c}]{} \frac{s^2}{2}\Big(f_{xy}^2(\bar{c}) \cdot f_{xx}(\bar{c}) - 2 f_{xy}^2(\bar{c}) \cdot f_{xx}(\bar{c}) + f_{xx}^2(\bar{c}) \cdot f_{yy}(\bar{c})\Big) \\ &= \frac{s^2}{2} \cdot f_{xx}(\bar{c}) \cdot \Big(\underbrace{f_{xx}(\bar{c}) \cdot f_{yy}(\bar{c}) - f_{xy}^2(\bar{c})}_{<0}\Big). \end{aligned}$$

Folglich haben $f(\bar{x}) - f(\bar{c})$ und $f_{xx}(\bar{c})$ unterschiedliches Vorzeichen, und damit besitzt f in $\bar{c}$ kein lokales Extremum.

Den Fall $f_{yy}(\bar{c}) \neq 0$ beweist man durch ähnliche Überlegungen. ❑

Die folgende Abbildung zeigt eine Funktion mit Sattelpunkt. **8/3/24**

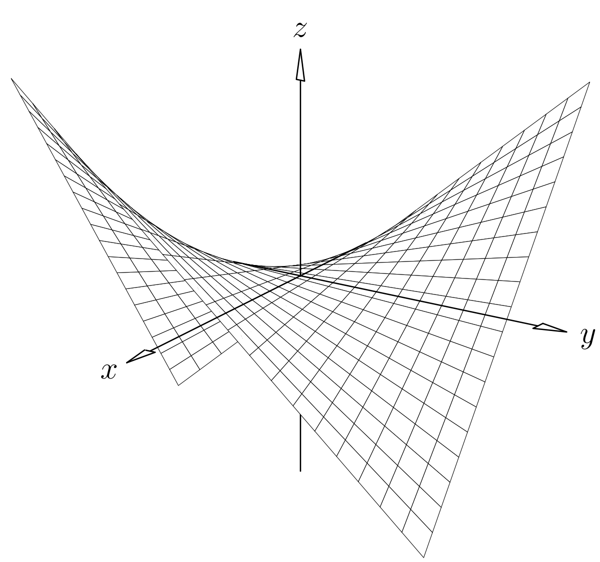

Abb. 8.11

In der Abbildung ist die Funktion $f(x, y) = -xy$ dargestellt. f besitzt in $(0, 0)$ einen Sattelpunkt. Der Graph der Funktion erzeugt eine sog. *Sattelfläche*.

In den Quadranten, in denen x und y jeweils nur positive bzw. nur negative Werte annehmen, ist die Funktion $f(x,y)$ stets negativ, in den Quadranten, wo jeweils ein Wert positiv und ein Wert negativ ist, ist die Funktion positiv. Entlang der x-Achse und der y-Achse ist die Funktion stets null.
Die dargestellte Fläche läßt sich offenbar allein durch Geraden erzeugen.

Wir fahren jetzt fort mit der Untersuchung unseres Beispiels. Hierzu benutzen wir das oben erhaltene Kriterium. **8/3/25**

Offenbar sind die zweiten partiellen Ableitungen an der kritischen Stelle $(0,0)$ stetig. Denn es ist

$$f_{xx}(x,y) = 2, \quad f_{xy}(x,y) = f_{yx}(x,y) = 0, \quad f_{yy}(x,y) = 2,$$

und damit gilt

$$D = \begin{vmatrix} f_{xx}(\bar{0}) & f_{xy}(\bar{0}) \\ f_{xy}(\bar{0}) & f_{yy}(\bar{0}) \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 \neq 0.$$

Wegen $f_{xx}(\bar{0}) = 2 > 0$ besitzt f in $\bar{0}$ ein lokales Extremum (vgl. auch Abb. 8.5 und Abb. 8.8 b).

Wir betrachten jetzt das Beispiel $f(x,y) = x^2 - y^2$. **8/3/26**

Die folgende Abbildung zeigt diese Funktion; sie stellt ebenfalls eine Sattelfläche dar, die an der Stelle $(0,0)$ einen Sattelpunkt besitzt.

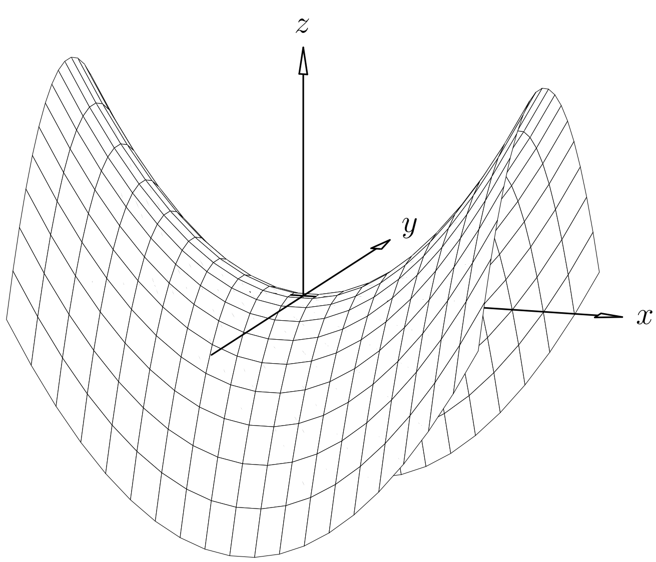

Abb. 8.12
Die Funktion $f(x,y) = x^2 - y^2$ besitzt an der Stelle $(0,0)$ einen Sattelpunkt. Der Graph von f stellt eine Sattelfläche dar.

Es ist

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

$$f_{xx}(x,y)=2,\quad f_{xy}(x,y)=f_{yx}(x,y)=0,\quad f_{yy}(x,y)=-2.$$

Dann ist $\bar{0}$ wieder ein kritischer Punkt von f, aber

$$D=\begin{vmatrix}2 & 0\\ 0 & -2\end{vmatrix}=-4<0,$$

folglich besitzt f in $\bar{0}$ einen Sattelpunkt.

Bemerkung. Nach unserer Definition sind die Extremstellen einer Funktion f immer **8/3/27** innere Punkte des betrachteten Definitionsbereiches (man kann dies auch anders definieren !). Ist man nicht nur an lokalen sondern auch an absoluten (oder globalen) Extrema (im Gegensatz zu lokalen Extremstellen) interessiert, dann muß noch der Teil des Randes des Definitionsbereiches untersucht werden, der selbst zum Definitionsbereich gehört. Schränkt man die Funktion f auf den betreffenden Teil des Randes ein, dann erhält man (in Abhängigkeit von der Kompliziertheit des Randes) oft eine „handhabbare" Funktion g mit einer Veränderlichen. Die lokalen und globalen Extrema für g (falls solche existieren) müssen dann mit den lokalen Extrema von f verglichen werden.

8.4 Implizite Funktionen

In diesem Abschnitt befassen wir uns mit der (nicht einfachen Problematik der) Auflösbarkeit **8/4/0** von Gleichungssystemen. Die lineare Algebra stellt bekanntlich gut nutzbare Werkzeuge für die Auflösung von linearen Gleichungssystemen bereit. Diese Hilfsmittel versagen jedoch im nichtlinearen Fall. Zur Verdeutlichung der Auflösbarkeit solcher Systeme starten wir mit einem linearen Gleichungssystem, das gegeben ist durch

$$A\bar{x}=\bar{b}\ ,\ \text{ wobei }\ A=\begin{pmatrix}a_{11} & \dots & a_{1n}\\ \vdots & & \vdots\\ a_{m1} & \dots & a_{mn}\end{pmatrix},\quad \bar{x}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\quad \bar{b}=\begin{pmatrix}b_1\\ \vdots\\ b_m\end{pmatrix}.$$

Das Gleichungssystem ist bekanntlich genau dann lösbar (d.h. nach $x_1,\dots,x_n$ auflösbar), wenn der Rang der Koeffizientenmatrix mit dem der erweiterten Koeffizientenmatrix übereinstimmt. Für den Fall, daß $m=n$ und der Rang von A ebenfalls n ist (d.h. die Determinante $\det A$ von A nicht null ist), läßt sich das Gleichungssystem stets eindeutig lösen. Ist das Gleichungssystem lösbar, $n>m$ und (nach eventueller Umsortierung der Koeffizienten)

$$\det\begin{pmatrix}a_{11} & \dots & a_{1m}\\ \vdots & & \vdots\\ a_{m1} & \dots & a_{mm}\end{pmatrix}\neq 0\ ,$$

dann gibt es lineare Funktionen $f_1(x_{m+1},\dots,x_n),\ \dots\ ,f_m(x_{m+1},\dots,x_n)$, so daß

$$x_i=f_i(x_{m+1},\dots,x_n)\ \text{ für }\ i=\ \dots,m.$$

Für $\bar{x} = (x_{m+1}, \ldots, x_n)$, $\bar{y} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ und $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$ erhält man somit

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

Wir befassen uns jetzt mit beliebigen Gleichungssystemen. Dazu betrachten wir zunächst eine Gleichung mit zwei Unbekannten.

Definition. (*Implizit definierte Funktion*) **8/4/1**

Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $\bar{a} = (a, b)$ und f in einer Umgebung $U(\bar{x})$ definiert und $f(a, b) = 0$. Weiterhin seien $\varepsilon, \delta > 0$, jedoch so klein, daß $U_\delta(a) \times U_\varepsilon(b) \subseteq U(\bar{a})$.

Durch die Gleichung $f(x, y) = 0$ ist in der Umgebung $U_\delta(a)$ eine *Funktion* $g : \mathbb{R} \to \mathbb{R}$ *implizit definiert*

$\overset{=}{\text{Df}}$ Für jedes $x \in U_\delta(a)$ gibt es genau ein $y \in U_\varepsilon(b)$, so daß $f(x, y) = 0$ und $y = g(x)$ (insbesondere ist $b = g(a)$).

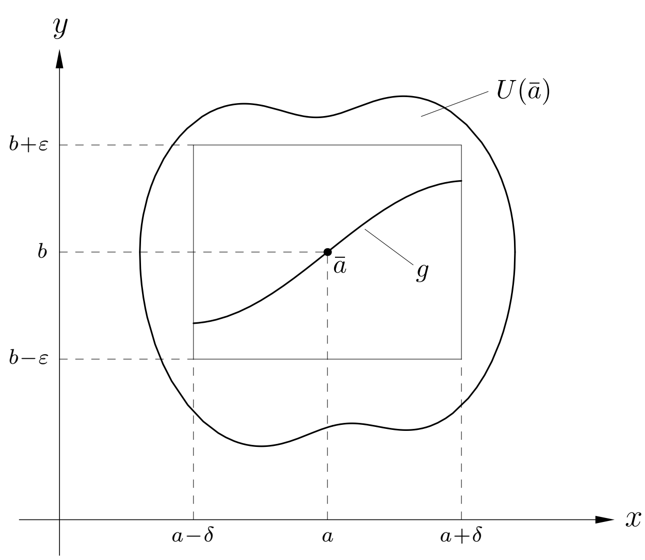

Abb. 8.13 zeigt die durch die Gleichung $f(x, y) = 0$ in dem Intervall $(a-\delta, a+\delta)$ implizit definierte Funktion g. Die entsprechende Kurve ist der Durchschnitt der dreidimensionalen Punktmenge $\{(x, y, z) : z = f(x, y)\}$ mit der Ebene $z = 0$. Hierfür denke man sich die z-Achse senkrecht auf der (x, y)-Ebene stehend.

Beispiel. Sei $f(x, y) = x^2 + y^2 - 1$. **8/4/2**

$f(x, y) = 0$ ($\iff x^2 + y^2 = 1$) definiert in $\mathbb{R}^2$ eine Kreislinie mit dem Radius 1 und dem Mittelpunkt (0,0).

Sei $\bar{a} = (a, b) \in \mathbb{R}^2$ und $f(\bar{a}) = 0$, dann ist $\bar{a}$ ein Punkt der Kreislinie. Für $a \neq \pm 1$ und $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß für jedes $x \in U_\delta(a)$ genau ein $y \in U_\varepsilon(b)$ existiert mit $f(x, y) = 0$. In diesem Fall ist $y = g(x) = \sqrt{x}$.

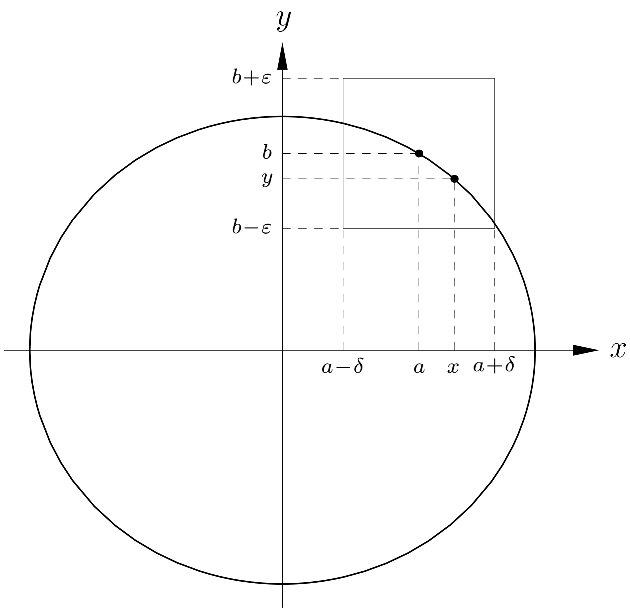

Abb. 8.14 Hier wird die analoge Situation wie in der Abb. 8.13 dargestellt, jedoch für den Spezialfall $f(x,y) = x^2 + y^2 - 1 = 0$.

Satz 8.15 (*Hauptsatz über implizite Funktionen mit zwei Veränderlichen*) **8/4/3**

Voraussetzung:

(1) *Sei* $M \subseteq \mathbb{R}$ *eine offene Menge und* $f : \mathbb{R}^2 \to \mathbb{R}$ *stetig in* M.

(2) *Sei* $\bar{c} = (a,b) \in M$ *und* $f(\bar{c}) = 0$.

(3) f *ist in einer Umgebung* $U(\bar{c})$ *stetig partiell nach* y *differenzierbar und* $f_y(\bar{c}) \neq 0$.

Behauptung:

Es gibt eine stetige Funktion $g : \mathbb{R} \to \mathbb{R}$, *für die gilt*:
Für jedes $\varepsilon > 0$ *gibt es ein* $\delta > 0$, *so daß für jedes* $x \in U_\delta(a)$ *genau ein* $y \in U_\varepsilon(b)$ *existiert mit* $f(x,y) = 0$ *und* $y = g(x)$ (insbesondere ist $b = g(a)$).

Beweis. Da $f_y(\bar{c}) \neq 0$ und f_y in $U(\bar{c})$ stetig ist, gibt es eine Umgebung $U'(\bar{c})$, **8/4/4** so daß f_y dort stets positiv oder stets negativ ist. Sei o.B.d.A. $f_y(\bar{c}) > 0$ ($f_y(\bar{c}) < 0$ analog).

Wir wählen jetzt $d > 0$, jedoch so klein, daß $U_d(a) \times U_d(b) \subseteq U'(\bar{c})$.

Da $f_y(a,y) > 0$ für alle $y \in U_d(b)$, ist f_y in $U_d(b)$ streng monoton wachsend. Sei $0 < \varepsilon < d$ und $y_1 := b - \varepsilon$, $y_2 := b + \varepsilon$. Wegen $f(a,b) = 0$ ist $f(a,y_1) < 0 < f(a,y_2)$. Da $f(x,y_1)$, $f(x,y_2)$ als Funktionen von x stetig sind, gibt es ein δ mit $0 < \delta < \varepsilon$, so daß für jedes $x \in U_\delta(a)$ gilt: $f(x,y_1) < 0 < f(x,y_2)$.

Für $x_0 \in U_\delta(a)$ ist also $f(x_0,y)$ in $[y_1,y_2]$ stetig und

$$f(x_0,y_1) < 0 < f(x_0,y_2).$$

Nach dem Zwischenwertsatz (für Funktionen einer Veränderlichen) gibt es ein $y_0 \in [y_1,y_2]$, so daß $f(x_0,y_0) = 0$.

Da f_y in $U'(\bar{c})$ stets positiv ist, erhält man insbesondere $f_y(x_0, y) > 0$. Folglich ist $f(x_0, y)$ streng monoton wachsend und somit y_0 das einzige Element in $U'_\varepsilon(b)$ mit $f(x_0, y_0)$. Durch $g(x_0) := y_0$ für $x_0 \in U_\delta(a)$ ist eine Funktion g definiert.

Es bleibt noch zu zeigen, daß g in $U_\delta(a)$ stetig ist.

Sei $x_0 \in U_\delta(a)$, $\varepsilon > 0$ (wie oben) und $\delta' > 0$, jedoch so klein, daß $U_{\delta'}(x_0) \subseteq U_\delta(a)$ und $|x - x_0| < \delta'$. Nach den vorhergehenden Betrachtungen existiert für $x \in U_{\delta'}(a)$ genau ein $y \in [b - \varepsilon, b + \varepsilon]$, so daß $f(x, y) = 0$, also $g(x) = y$.

Damit erhält man

$$|g(x) - g(x_0)| = |y - y_0| \leq \underbrace{|y - b|}_{<\varepsilon} + \underbrace{|b - y_0|}_{<\varepsilon} < 2\varepsilon.$$

Hieraus folgt die Stetigkeit von g in x_0. ❑

Korollar. *Gilt zusätzlich zu den Voraussetzungen von Satz 8.15 noch* **8/4/5**

(4) f *ist in* $U(\bar{c})$ *stetig partiell nach* x *differenzierbar,*

dann ist die durch $f(x, y) = 0$ *in* $U_\delta(a)$ *implizit definierte Funktion* g *differenzierbar, und es gilt:* $g'(x) = -\dfrac{f_x(x, y)}{f_y(x, y)}$.

Beweis. Wir wählen die Bezeichnungen wie im vorhergehenden Beweis. **8/4/6**

Sei $|x - x_0| < \delta'$, $x \neq a$ und $y = g(x)$, dann ist $(x, y) \in U_{\delta'}(x_0) \times U_\varepsilon(b) := D$, und die Verbindungsstrecke zwischen (a, b) und (x, y) gehört ganz zu D. Nach dem Mittelwertsatz für Funktionen mit mehreren Veränderlichen gilt (er kann hier angewendet werden, da $f(x, y)$ nach Voraussetzung differenzierbar ist):

$$\underbrace{f(x, y)}_{=0} - \underbrace{f(a, b)}_{=0} = 0 = (x - a) \cdot f_x(\bar{u}) + (y - b) \cdot f_y(\bar{u}),$$

wobei $\bar{u} = \bar{a} + \vartheta(\bar{x} - \bar{c})$, $y = g(x)$ und $b = g(a)$. Hieraus folgt

$$\frac{y - b}{x - a} = \frac{g(x) - g(a)}{x - a} = -\frac{f_x(\bar{u})}{f_y(\bar{u})} \ .$$

Da g in $U_\delta(a)$ stetig ist, gilt für $x \to a$ auch $y = g(x) \to g(a) = b$ und somit $\bar{u} \to \bar{a}$. Folglich existiert

$$\underbrace{\lim_{x \to a} \frac{g(x) - g(a)}{x - a}}_{= g'(a)} = -\frac{f_x(\bar{c})}{f_y(\bar{c})} \ ,$$

also

$$g'(x) = -\frac{f_x(x, y)}{f_y(x, y)} \ . \quad ❑$$

Beispiel. Es sei $f(x,y) = x^3 + y^3 - 2xy - 1$. **8/4/7**

Für f gelten die folgenden Voraussetzungen:

1. $f : \mathbb{R}^2 \to \mathbb{R}$ ist in einer Umgebung $U(\bar{c})$ mit $\bar{c} = (1,0)$ stetig.
2. $f(1,0) = 0$.
3. f ist in $U(\bar{c})$ stetig partiell nach y differenzierbar und $f_y(1,0) = -2 \neq 0$.

Nach dem Satz über implizite Funktionen gibt es dann eine stetige Funktion $g : \mathbb{R} \to \mathbb{R}$, für die gilt:

Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß für alle $x \in U_\delta(1)$ genau ein $y \in U_\varepsilon(0)$ existiert mit $f(x,y) = 0$ und $g(x) = y$.

Offenbar ist auch f stetig partiell nach x differenzierbar. Damit ist (nach dem letzten Korollar) g in $U_\delta(1)$ differenzierbar und

$$y' = g'(x) = -\frac{f_x(x,y)}{f_y(x,y)} = \frac{3x^2 - 2y}{3y^2 - 2x} \ .$$

Hierbei entsteht eine Gleichung, die eine unbekannte Funktion y und deren Ableitung y' enthält. (Gleichungen dieser Art heißen *Differentialgleichungen*.)

Abschließend soll noch der Hauptsatz über implizite Funktionen mit mehreren Veränderlichen ohne Beweis angegeben werden. (Siehe hierzu Literaturangabe [4], Teil II, Seite 235 – 242.) **8/4/8**

Dazu benötigen wir noch die folgende Definition.

Definition. (*implizit definierte Funktionen mit mehreren Veränderlichen*) **8/4/9**

Sei $f : \mathbb{R}^{n+m} \to \mathbb{R}^m$, $\bar{a} = (a_1, \dots, a_n)$, $\bar{b} = (b_1, \dots, b_m)$, $\bar{c} = (\bar{a}, \bar{b})$ und f in einer Umgebung $U(\bar{c}) \subseteq \mathbb{R}^{n+m}$ definiert und $f(\bar{c}) = \bar{0}$.
Weiterhin seien $\varepsilon, \delta > 0$, jedoch so klein, daß $U_\delta(\bar{a}) \times U_\varepsilon(\bar{b}) \subseteq U(\bar{c})$.

Durch die Gleichung $f(\bar{x}, \bar{y}) = \bar{0}$ ist in der Umgebung $U_\delta(\bar{a})$ eine *Funktion* $g : \mathbb{R}^n \to \mathbb{R}^m$ *implizit definiert*

$\underset{\text{Df}}{\equiv}$ Für jedes $\bar{x} \in U_\delta(\bar{a})$ gibt es genau ein $\bar{y} \in U_\varepsilon(\bar{b})$, so daß $f(\bar{x}, \bar{y}) = \bar{0}$ und $\bar{y} = g(\bar{x})$ (insbesondere ist $\bar{b} = g(\bar{a})$).

Satz 8.16 (*Hauptsatz über implizite Funktionen*) **8/4/10**

Voraussetzung:

(1) *Sei* $M \subseteq \mathbb{R}^{n+m}$ *eine offene Menge und* $f = (f_1, \dots, f_m) : \mathbb{R}^{n+m} \to \mathbb{R}^m$ *stetig in* M.

(2) *Sei* $\bar{c} = (\bar{a}, \bar{b})$ *und* $f(\bar{c}) = \bar{0}$.

(3) $f_1, \dots, f_m$ *seien in einer Umgebung* $U(\bar{c})$ *nach allen Variablen* $y_1, \dots, y_m$ *stetig partiell differenzierbar und die Determinante der Funktionalmatrix* $\dfrac{\partial(f_1, \dots, f_m)}{\partial(y_1, \dots, y_m)}(\bar{c})$ *sei* (an der Stelle $\bar{c}$) *nicht null.*

Behauptung:
Es gibt eine stetige Funktion $g : \mathbb{R}^n \to \mathbb{R}$, *für die gilt*:
Für jedes $\varepsilon > 0$ *gibt es ein* $\delta > 0$, *so daß für jedes* $\bar{x} \in U_\delta(\bar{a})$ *genau ein* $\bar{y} \in U_\varepsilon(\bar{b})$ *existiert mit* $f(\bar{x}, \bar{y}) = 0$ *und* $\bar{y} = g(\bar{x})$ (insgesondere ist $g(\bar{a}) = b$).

Beweis. Siehe Literaturangabe [4], Teil II, Seite 235. ❏ **8/4/11**

Schwerpunkte für die Wiederholung von Kapitel 8

- Das Wesen der Differenzierbarkeit (lineare Approximation) **8/5/1**
- Definitionen: Ableitung, partielle Ableitung, Richtungsableitung, Tangentialebene (geometrische Veranschaulichung), Differential, **8/5/2**
- Aus der Differenzierbarkeit folgt die Stetigkeit; aus der partiellen Differenzierbarkeit (nach allen Variablen) folgt noch nicht die Stetigkeit, also auch nicht die Differenzierbarkeit, **8/5/3**
- Beziehungen zwischen Differenzierbarkeit, partieller Differenzierbarkeit und Richtungsableitbarkeit, **8/5/4**
- Differenzierbarkeit zusammengesetzter Funktionen, **8/5/5**
- Satz von Schwarz, **8/5/6**
- Mittelwertsatz der Differentialrechnung mit mehreren Veränderlichen, **8/5/7**
- Definitionen: polygonzusammenhängend, Gebiet, stetig differenzierbar, **8/5/8**
- f ist konstant ... $\iff$ $f'(\bar{x} = 0) \ldots$ (Satz 8.11), **8/5/9**
- Satz von Taylor für Funktionen mit n Veränderlichen + Korollar, **8/5/10**
- Definition: lokale Extrema, **8/5/11**
- Notwendige und (für Funktionen mit zwei Veränderlichen auch) hinreichende Bedingungen für die Existenz lokaler Extrema. **8/5/12**