

# Mathematik III

(für Informatiker)

Oliver Ernst

Professur Numerische Mathematik

Wintersemester 2016/17



Mathematik!  
TU Chemnitz

## 9 Differentialrechnung in mehreren Variablen

9.1 Vektorfolgen und ihre Grenzwerte

9.2 Grenzwerte von Funktionen und Stetigkeit

9.3 Darstellungsfragen, Anwendungen und Systematisierungsversuch zu Funktionen mehrerer Variablen

9.4 Differenzierbarkeit bei mehreren Variablen

9.5 Differentiation vektorwertiger Funktionen

9.6 Extrema von Funktionen mehrerer Variablen

## 10 Integralrechnung für Funktionen von mehreren Variablen

10.1 Das Riemann-Integral im  $\mathbb{R}^n$

10.2 Kurven und Kurvenintegrale

10.3 Oberflächen und Oberflächenintegrale

## 11 Integraltransformationen

11.1 Allgemeines

11.2 Fourier-Transformation

11.3 Laplace-Transformationen

## 12 Algebraische Strukturen

12.1 Gruppen

12.2 Ringe und Körper

12.3 Elementare Zahlentheorie

12.4 Äquivalenzrelationen und Äquivalenzklassen

12.5 Zahlentheorie und Kryptographie

- 9 Differentialrechnung in mehreren Variablen
- 10 Integralrechnung für Funktionen von mehreren Variablen
- 11 Integraltransformationen
- 12 Algebraische Strukturen

# Multivariate Differential- und Integralrechnung

In diesem Kapitel weiten wir die Infinitesimalrechnung auf Funktionen aus, bei denen sowohl die Argumente als auch die Funktionswerte Vektoren sind.

Es geht also um Funktionen vom Typ

$$\mathbf{f} : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}^m, \quad \mathbf{x} \mapsto \mathbf{f}(\mathbf{x}).$$

Begriffe wie Abbildungsvorschrift, Definitions- und Wertebereich haben dabei die gewohnte Bedeutung.

# Multivariate Differentialrechnung

## Notation

Statt  $f\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)$  schreibt man häufig komfortabler  $f(x_1, \dots, x_n)$  oder  $f(\mathbf{x})$ .

Für  $n = 2$  schreibt man oft  $[x, y]^T$  statt  $[x_1, x_2]^T$  bzw. für  $n = 3$  oft  $[x, y, z]^T$  statt  $[x_1, x_2, x_3]^T$ . Dies hilft vor allem, bei verschiedenen Indizierungen die Übersicht zu behalten.

### Beispiele:

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = 2x + 3y - 1$
- $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad g(x, y, z) = \begin{bmatrix} x^2 + y^2 + z^2 \\ x + e^y \end{bmatrix}$

## 9 Differentialrechnung in mehreren Variablen

### 9.1 Vektorfolgen und ihre Grenzwerte

### 9.2 Grenzwerte von Funktionen und Stetigkeit

### 9.3 Darstellungsfragen, Anwendungen und Systematisierungsversuch zu Funktionen mehrerer Variablen

### 9.4 Differenzierbarkeit bei mehreren Variablen

### 9.5 Differentiation vektorwertiger Funktionen

### 9.6 Extrema von Funktionen mehrerer Variablen

# Multivariate Differentialrechnung

## Vektorfolgen und ihre Grenzwerte

Um den Ableitungsbegriff auf Funktionen mehrerer Variablen zu übertragen, müssen wir zunächst Grenzwerte von Funktionen und damit auch Grenzwerte von Folgen zur Verfügung stellen.

Wir klären zunächst, was wir unter einer Vektorfolge und deren Konvergenz verstehen. Dabei orientieren wir uns konsequent am eindimensionalen Fall.

Als erstes müssen wir beschreiben, was wir unter „nahe beieinander liegen“ für Vektoren verstehen. Dabei hilft uns Euklidische Norm.

Insbesondere brauchen wir eine Verallgemeinerung der „ $\epsilon$ -Umgebung“ – im Eindimensionalen ein Intervall der Form

$$(x - \epsilon, x + \epsilon).$$



### Definition 9.1

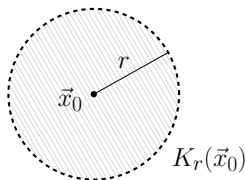
Sei  $\mathbf{x}_0 \in \mathbb{R}^n$  und  $r > 0$ . Dann heißt

$$K_r(\mathbf{x}_0) := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}_0\| < r\}$$

$n$ -dimensionale (offene) **Kugel** um  $\mathbf{x}_0$  mit Radius  $r$ .

Die Kugel  $K_1(\mathbf{0})$  heißt (offene)  $n$ -dimensionale **Einheitskugel**.

$K_r(\mathbf{x}_0)$  ist also die Menge aller Punkte des  $\mathbb{R}^n$ , deren Abstand zu  $\mathbf{x}_0$  kleiner als  $r$  ist.



Wir betrachten jetzt Folgen, deren Glieder Vektoren aus dem  $\mathbb{R}^n$  sind.

### Notation:

- für das  $m$ -te Folgenglied:

$$\mathbf{x}^{(m)} = \begin{bmatrix} x_1^{(m)} \\ \vdots \\ x_n^{(m)} \end{bmatrix}$$

(der Folgenindex steht in Klammern oben, um Verwechslungen mit dem Komponentenindex und mit Exponenten bei Potenzen vorzubeugen),

- für die Folge als Ganzes:  $(\mathbf{x}^{(m)})_{m \in \mathbb{N}}$ ,  $(\mathbf{x}^{(m)})_m$  oder einfach  $(\mathbf{x}^{(m)})$ .

Die reellen Zahlenfolgen  $(x_j^{(m)})_m$  (für  $j \in \{1, \dots, n\}$  jeweils fest) heißen **Komponentenfolgen** von  $(\mathbf{x}^{(m)})_m$ .

### Definition 9.2

Ein Vektor  $\mathbf{x}$  heißt **Grenzwert** der Vektorfolge  $(\mathbf{x}^{(m)})$ , wenn zu jedem  $\epsilon > 0$  ein Index  $m_0 \in \mathbb{N}$  existiert, so dass

$$\|\mathbf{x}^{(m)} - \mathbf{x}\| < \epsilon \quad \text{für alle } m \geq m_0.$$

Besitzt die Vektorfolge  $(\mathbf{x}^{(m)})$  einen Grenzwert, so heißt sie **konvergent**, anderenfalls **divergent**.

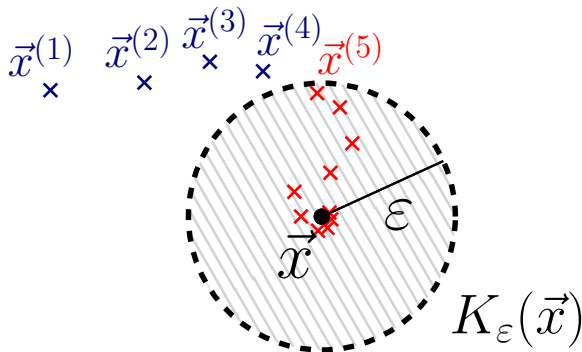
#### Schreibweisen:

- $\mathbf{x} = \lim_{m \rightarrow \infty} \mathbf{x}^{(m)}$
- $\mathbf{x}^{(m)} \rightarrow \mathbf{x}$  für  $m \rightarrow \infty$ , oder kürzer:  $\mathbf{x}^{(m)} \rightarrow \mathbf{x}$ .

Vergleichen Sie mit der Grenzwertdefinition für reelle Zahlenfolgen (Definition 3.7).

# Multivariate Differentialrechnung

## Graphische Interpretation



Für große  $m$  liegen die Folgenglieder beliebig nahe am Grenzwert, d.h. in einer beliebig kleinen Kugel  $K_\epsilon(\vec{x})$  („ $\epsilon$ -Umgebung“ von  $\vec{x}$ ).

Im hier visualisierten Fall liegen für  $m \geq m_0 = 5$  alle Folgenglieder in der gewählten  $\epsilon$ -Umgebung.

# Multivariate Differentialrechnung

## Komponentenweise Konvergenz

Auch wenn sich eine Vektorfolge auf verschiedenste Weise dem Grenzwert nähern kann, ist die Konvergenzanalyse einfacher als erwartet:

### Satz 9.3

Sei  $(\mathbf{x}^{(m)}) \subset \mathbb{R}^n$  eine Vektorfolge und  $\mathbf{x} \in \mathbb{R}^n$ . Dann gilt

$$\lim_{m \rightarrow \infty} \mathbf{x}^{(m)} = \mathbf{x}$$

genau dann, wenn

$$\lim_{m \rightarrow \infty} x_j^{(m)} = x_j \quad \text{für alle } 1 \leq j \leq n.$$

Die Konvergenz einer Vektorfolge ist also äquivalent zur Konvergenz sämtlicher Komponentenfolgen  $(x_j^{(m)})_m$ .

Bei diesen handelt es sich um reelle Zahlenfolgen, für die Sie die bekannten Gesetzmäßigkeiten verwenden können.

Konvergieren die Vektorfolgen

$$\mathbf{x}^{(m)} = \begin{bmatrix} \frac{1}{m^2} \\ 0 \end{bmatrix}, \quad \mathbf{y}^{(m)} = \begin{bmatrix} \frac{1}{m} \\ m \end{bmatrix}, \quad \mathbf{z}^{(m)} = \begin{bmatrix} 1 \\ \sqrt[m]{42} \\ \frac{2m+3}{m+2} \end{bmatrix} ?$$

Bestimmen Sie im Falle der Konvergenz den Grenzwert.

Wegen Satz 9.3 übertragen sich alle Rechenregeln aus Satz 3.18:

### Satz 9.4

Seien  $(\mathbf{x}^{(m)}), (\mathbf{y}^{(m)}) \subset \mathbb{R}^n$  konvergente Vektorfolgen mit  $\mathbf{x}^{(m)} \rightarrow \mathbf{x}$  und  $\mathbf{y}^{(m)} \rightarrow \mathbf{y}$ . Dann gilt:

- (1)  $\lambda \mathbf{x}^{(m)} \rightarrow \lambda \mathbf{x}$  für jede Konstante  $\lambda \in \mathbb{R}$ .
- (2)  $\mathbf{x}^{(m)} + \mathbf{y}^{(m)} \rightarrow \mathbf{x} + \mathbf{y}$ ,
- (3)  $\mathbf{x}^{(m)} - \mathbf{y}^{(m)} \rightarrow \mathbf{x} - \mathbf{y}$ ,
- (4)  $\mathbf{x}^{(m)} \cdot \mathbf{y}^{(m)} \rightarrow \mathbf{x} \cdot \mathbf{y}$ .
- (5) Sind alle Komponenten von  $\mathbf{y}$  von Null verschieden, so gibt es ein  $m_0 \in \mathbb{N}$ , so dass  $y_j^{(m)} \neq 0$  ( $m \geq m_0, j = 1, \dots, n$ ). Die Folge  $\left(\frac{\mathbf{x}^{(m)}}{\mathbf{y}^{(m)}}\right)_{m \geq m_0}$  konvergiert mit  $\frac{\mathbf{x}^{(m)}}{\mathbf{y}^{(m)}} \rightarrow \frac{\mathbf{x}}{\mathbf{y}}$ .

Die Multiplikation und Division in (4) und (5) ist dabei komponentenweise zu verstehen.

Schließlich wollen wir noch eine Verallgemeinerung von Satz 3.10 formulieren:

### Satz 9.5

Seien  $(\mathbf{x}^{(m)}) \subset \mathbb{R}^n$  eine Vektorfolge. Dann gilt:

$$\mathbf{x}^{(m)} \rightarrow \mathbf{x} \quad \Longleftrightarrow \quad \|\mathbf{x}^{(m)} - \mathbf{x}\| \rightarrow 0. \quad (9.1)$$

Die Vektorfolge  $(\mathbf{x}^{(m)})$  konvergiert also genau dann gegen  $\mathbf{x}$ , wenn der Abstand ihrer Glieder zu  $\mathbf{x}$  für große  $m$  beliebig klein wird.

Beachten Sie, dass die Folge in (9.1) rechts wieder eine reelle Zahlenfolge ist, die mit herkömmlichen Mitteln untersucht werden kann.



## 9 Differentialrechnung in mehreren Variablen

9.1 Vektorfolgen und ihre Grenzwerte

9.2 Grenzwerte von Funktionen und Stetigkeit

9.3 Darstellungsfragen, Anwendungen und Systematisierungsversuch zu Funktionen mehrerer Variablen

9.4 Differenzierbarkeit bei mehreren Variablen

9.5 Differentiation vektorwertiger Funktionen

9.6 Extrema von Funktionen mehrerer Variablen

# Multivariate Differentialrechnung

## Grenzwerte von Funktionen und Stetigkeit

In diesem Kapitel werden wir das Konzept der Stetigkeit auf Funktionen mehrerer Variablen erweitern.

Intuitiv wollen wir mit diesem Begriff wieder folgende Eigenschaften einer Funktion erfassen:

- (Hinreichend) kleine Änderungen an den Argumenten führen zu (beliebig) kleinen Änderungen der Funktionswerte.
- Durch hinreichend feines „Justieren“ der Eingabewerte lassen sich die Ausgabewerte einer Funktion hinreichend fein „einstellen“.

Der Zugang zur Stetigkeit erfolgt wie im Eindimensionalen über Grenzwerte von Funktionen. Erforderlich ist also die Verallgemeinerung des Grenzwertbegriffs auf Funktionen mehrerer Variablen.

### Definition 9.6

Sei  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}^m$  und  $\mathbf{x}_0 \in \mathbb{R}^n$ . Man sagt,  $f$  **konvergiert** für  $\mathbf{x}$  gegen  $\mathbf{x}_0$  gegen  $\mathbf{a}$ , wenn für alle Folgen  $(\mathbf{x}^{(m)}) \subset D_f$  mit

$$\mathbf{x}^{(m)} \rightarrow \mathbf{x}_0 \quad \text{und} \quad \mathbf{x}^{(m)} \neq \mathbf{x}_0 \quad (\text{für alle } m \in \mathbb{N}) \quad (9.2)$$

die Beziehung  $f(\mathbf{x}^{(m)}) \rightarrow \mathbf{a}$  gilt. Man nennt  $\mathbf{a}$  den **Grenzwert** von  $f$  für  $\mathbf{x}$  gegen  $\mathbf{x}_0$ .

**Schreibweise:**  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{a}$  oder  $f(\mathbf{x}) \rightarrow \mathbf{a}$  für  $\mathbf{x} \rightarrow \mathbf{x}_0$ .

Strenggenommen muss man natürlich noch fordern, dass man sich dem Punkt  $\mathbf{x}_0$  auf die beschriebene Weise aus  $D_f$  heraus nähern kann.

### Definition 9.7

Die Funktion  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}^m$  heißt **stetig** an der Stelle  $\mathbf{x}_0 \in D_f$ , wenn der Grenzwert  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$  existiert, und

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$$

gilt.  $f$  heißt stetig auf der Menge  $M \subseteq D_f$ , wenn  $f$  an jeder Stelle  $\mathbf{x}_0 \in M$  stetig ist.

Machen Sie sich klar, dass es sich bei den Definitionen 9.6 und 9.7 um Verallgemeinerungen der eindimensionalen Situation handelt.

# Multivariate Differentialrechnung

## Beispiel

Wir betrachten die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = 1 - 2x^2 - y^2$$

und untersuchen ihr Verhalten im Punkt  $[x_0, y_0]^T = [0, 2]^T$ .

Für jede Vektorfolge mit  $[x_m, y_m]^T \rightarrow [0, 2]^T$  und  $[x_m, y_m]^T \neq [0, 2]^T$  gilt  $x_m \rightarrow 0$  und  $y_m \rightarrow 2$  nach Satz 9.3. Damit

$$f(x_m, y_m) = 1 - 2x_m^2 - y_m^2 \rightarrow 1 - 2 \cdot 0^2 - 2^2 = -3$$

und somit

$$\lim_{[x, y]^T \rightarrow [0, 2]^T} f(x, y) = -3 = f(0, 2).$$

Die Funktion besitzt also für  $[x, y]^T \rightarrow [0, 2]^T$  einen Grenzwert und ist dort stetig.

Natürlich greift die Argumentation analog für jeden beliebigen Punkt  $[x_0, y_0]^T$  – die Funktion  $f$  ist auf  $\mathbb{R}^2$  überall stetig.

# Multivariate Differentialrechnung

## Beispiel

Ändert man nun den Funktionswert im Punkt  $[0, 2]^T$  beispielsweise zu  $-42$ , dann bleibt der Grenzwert zwar erhalten, aber die entstehende Funktion

$$g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g(x, y) = \begin{cases} 1 - 2x^2 - y^2, & \text{für } [x, y]^T \neq [0, 2]^T; \\ -42, & \text{für } [x, y]^T = [0, 2]^T \end{cases}$$

ist im Punkt  $[0, 2]^T$  nicht mehr stetig. Es gilt

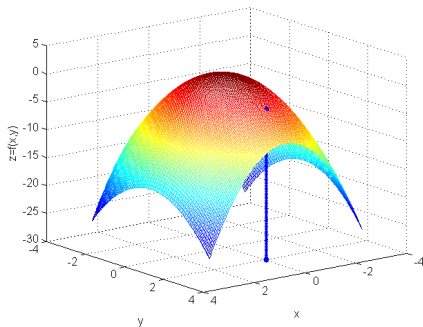
$$\lim_{[x, y]^T \rightarrow [0, 2]^T} g(x, y) = -3 \neq -42 = g(0, 2).$$

In allen anderen Punkten bleibt die Stetigkeit dagegen erhalten.

# Multivariate Differentialrechnung

## Beispiel

Wir versuchen, die Situation graphisch umzusetzen, und zeichnen den Graphen von  $z = f(x, y)$  als Fläche über der  $x$ - $y$ -Ebene.



Wenn man sich nahe des Punktes  $[0, 2]^T$  oben auf der Fläche befindet, dann wird man sich in etwa auf Höhe  $-3$  befinden.

Wenn man dagegen an besagter Stelle ein (unendlich dünnes) Loch bis zur Tiefe  $-42$  bohrt, kann man zwar beliebig dicht auf altem Höhenniveau herantreten, fällt allerdings bei exaktem Erreichen des Punktes  $[0, 2]^T$  in das Loch hinein.

Untersuchen Sie die folgenden Funktionen auf Stetigkeit.

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \begin{cases} 1, & \text{für } y < 0; \\ 0, & \text{für } y \geq 0, \end{cases}$$

$$g : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad g(x, y) = \begin{bmatrix} xe^y \\ x^4 \end{bmatrix},$$

$$h : \mathbb{R} \rightarrow \mathbb{R}^2, \quad h(t) = \begin{bmatrix} t \\ t^4 \end{bmatrix}.$$

Greifen Sie im Falle der Stetigkeit auf eindimensionale Argumente zurück.

### Satz 9.8

Eine Funktion  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}^m$  ist genau dann stetig in  $x \in D_f$ , wenn alle ihre Komponentenfunktionen  $f_i : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  ( $i = 1, \dots, m$ ) in  $x$  stetig sind.



# Multivariate Differentialrechnung

## Exkurs: Die Parabelfalte – ein interessantes Lehrbeispiel

Am Beispiel der **Parabelfalte** kann man eindrucksvoll studieren, wie wichtig sauberes Arbeiten bei mehrdimensionalen Grenzwertuntersuchungen ist. Sie ist gegeben durch die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \begin{cases} \frac{2xy^2}{x^2 + y^4}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

- Für  $y \neq 0$  ist  $f(0, y) = 0$ , d.h. mit Ausnahme möglicherweise von  $(0, 0)$  schließen sich beide Definitionsabschnitte stetig an.
- Für  $x > 0$  gilt  $f(x, y) = f(x, -y)$ , d.h. die Funktion ist symmetrisch zur  $y$ -Achse mit  $f(x, 0) \equiv 0$ .
- Für  $y \neq 0$  führt Erweitern mit  $1/y^4$  auf

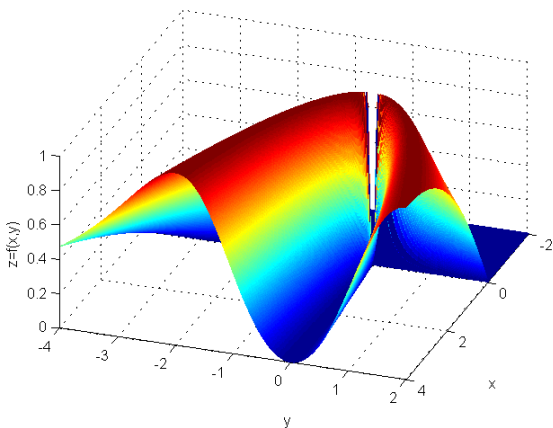
$$f(x, y) = \frac{2t}{t^2 + 1}, \quad \text{mit } t = \frac{x}{y^2} \geq 0.$$

Wegen  $2t \leq 1 + t^2$  (da  $(1 + t)^2 \geq 0$ ) folgt hieraus  $0 \leq f(x, y) \leq 1$ .

# Multivariate Differentialrechnung

## Exkurs: Die Parabelfalte – ein interessantes Lehrbeispiel

Man visualisiert die Parabelfalte wieder als „Gebirgsoberfläche“. Frage: was passiert beim „Talausgang“  $(0,0)$ ? Ist die Lücke in der Computer-Darstellung wirklich vorhanden?



# Multivariate Differentialrechnung

## Exkurs: Die Parabelfalte – ein interessantes Lehrbeispiel

- Wir interessieren uns für das Verhalten dieser Funktion für  $(x, y) \rightarrow (0, 0)$ .
- Wir betrachten Vektorfolgen  $(x_m, y_m) \rightarrow (0, 0)$ , zunächst solche, die sich entlang einer Geraden  $y = kx$  oder  $x = 0$  durch  $(0, 0)$  diesem Punkt nähern.
- Für  $x = 0$  erhält man die eindimensionale **Profilfunktion**

$$f(0, y) \equiv 0 \quad \text{d.h.} \quad f(0, y) \rightarrow 0 \quad \text{für } y \rightarrow 0.$$

- Betrachtet man die Funktionswerte entlang der Geraden  $y = kx$  ( $k \in \mathbb{R}$ ), so erhält man die Profilfunktion

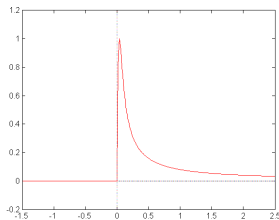
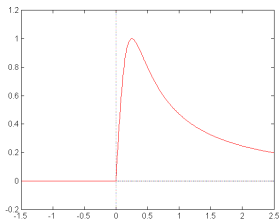
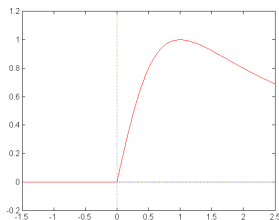
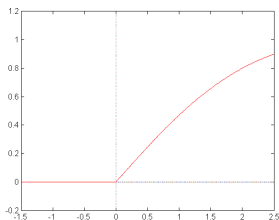
$$f(x, kx) = \begin{cases} \frac{2k^2 x^3}{x^2 + k^4 x^4} = \frac{2k^2 x}{1 + k^4 x^2}, & \text{für } x > 0, \\ 0, & \text{für } x \leq 0, \end{cases}$$

und somit ebenfalls  $f(x, kx) \rightarrow 0$  für  $x \rightarrow 0$  gilt.

- Mit anderen Worten, die Profile des Funktionsgraphen der Parabelfalte entlang von Geraden durch  $(0, 0)$  sind allesamt stetig.

# Multivariate Differentialrechnung

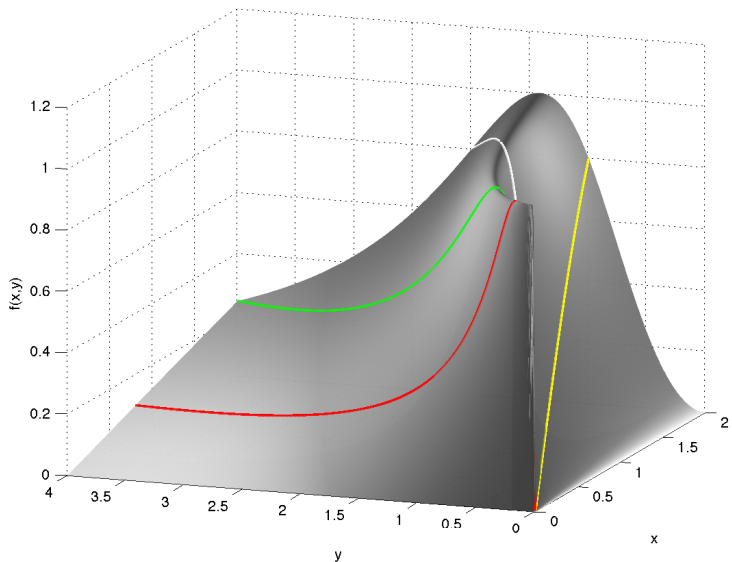
## Exkurs: Die Parabelfalte – ein interessantes Lehrbeispiel



Profilbilder der Parabelfalte entlang der Geraden  $y = kx$  für  $k = 0.5$ ,  $k = 1$ ,  $k = 2$  und  $k = 5$ . Man übersteigt jeweils den Gebirgskamm und läuft dann über einen mehr oder weniger steilen Hang zum Punkt  $[0, 0]^T$ .

# Multivariate Differentialrechnung

Exkurs: Die Parabelfalte – ein interessantes Lehrbeispiel



# Multivariate Differentialrechnung

## Exkurs: Die Parabelfalte – ein interessantes Lehrbeispiel

Da man zu allen Geraden durch  $(0, 0)$  stetige Profilkfunktionen erhält, könnte man versucht sein, daraus auf die Stetigkeit der Parabelfalte in diesem Punkt schließen.

Dies ist aber falsch! Betrachten Sie die Folge  $(x_m, y_m) = \left(\frac{1}{m^2}, \frac{1}{m}\right)$ . Hierfür gilt

$$f(x_m, y_m) \equiv 1 \quad \text{für alle } m \in \mathbb{N}, \quad \text{und somit } f(x_m, y_m) \rightarrow 1.$$

Die Parabelfalte besitzt also für  $(x, y) \rightarrow (0, 0)$  keinen Grenzwert und ist damit erst recht **nicht stetig** an der Stelle  $(0, 0)$ .

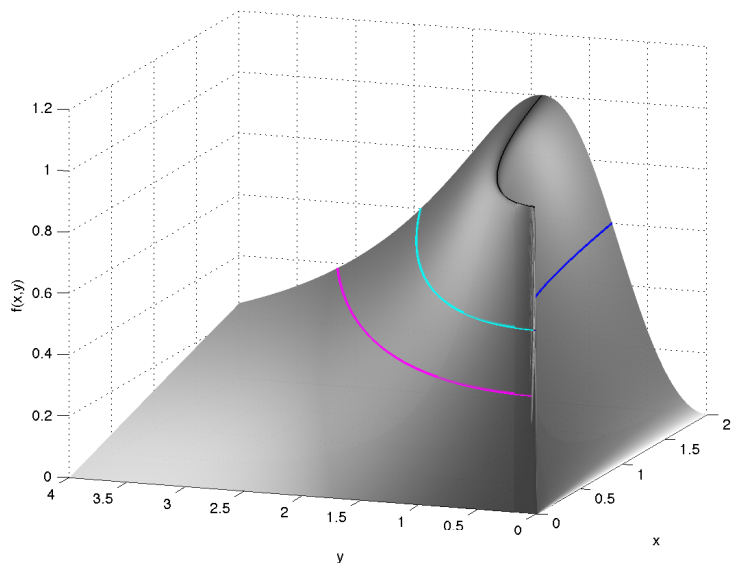
Das folgende Bild zeigt Profilkurven mit  $x = ky^2$ , d.h.  $x/y^2 = k$  sodass

$$f(x, y) = \frac{2k}{k^2 + 1}, \tag{9.3}$$

mit  $k = 1$ , (schwarz),  $k = 1/9$  (cyan),  $k = 1/25$  (magenta) und  $k = 9$  (blau).

# Multivariate Differentialrechnung

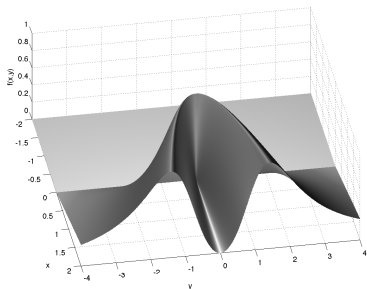
Exkurs: Die Parabelfalte – ein interessantes Lehrbeispiel



# Multivariate Differentialrechnung

## Exkurs: Die Parabelfalte – ein interessantes Lehrbeispiel

- Da der Ausdruck (9.3) für  $k \geq 0$  das gesamte Intervall  $[0, 1]$  durchläuft, sieht man, dass es zu allen Niveaus  $M \in [0, 1]$  eine Parabelkurve gibt, längs derer der Punkt  $(0, 0)$  mit konstantem Niveau  $M$  angesteuert werden kann. Somit existiert tatsächlich keine Lücke am Ende des „Tals“.



- Die betrachtete Folge  $(\frac{1}{m^2}, \frac{1}{m})$  liegt auf der Parabelkurve für  $k = 1$ , was dem (höchsten) Niveau  $M = 1$  entspricht. Diese Folge strebt dem Punkt  $(0, 0)$  somit längs dem „Gebirgskamm“ entgegen.

### Merke also:

Bei Grenzwerten ist es immer notwendig, in Definition 9.6 wirklich sämtliche Möglichkeiten der Annäherung der Folge  $(x^{(m)})$  an  $x_0$  zu betrachten.



Schließlich soll noch ein Ergebnis über Extrema reellwertiger Funktionen diskutiert werden, welches Satz 4.23 verallgemeinert. (Für vektorwertige Funktionen macht die Frage keinen Sinn.)

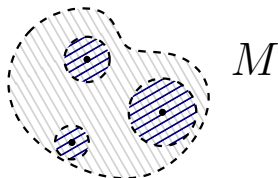
Zunächst einige Begriffe:

### Definition 9.9

Eine Punktmenge  $M \subset \mathbb{R}^n$  heißt **offen**, wenn es zu jedem Punkt  $x \in M$  eine Kugel  $K_r(x)$  gibt, die vollständig in  $M$  enthalten ist.

Eine Menge  $M \subset \mathbb{R}^n$  heißt **abgeschlossen**, wenn ihr Komplement offen ist.

Eine Menge  $M \subset \mathbb{R}^n$  heißt **beschränkt**, wenn  $\|x\| \leq C$  für eine Konstante  $C > 0$  und alle  $x \in M$ .



zum Begriff der offenen Menge

### Satz 9.10

Eine Funktion  $f : \mathbb{R}^n \supset M \rightarrow \mathbb{R}$ , die auf der abgeschlossenen und beschränkten Menge  $M$  stetig ist, nimmt auf  $M$  ihr Minimum und ihr Maximum an.

Es existieren also  $\mathbf{x}_{\min}$  und  $\mathbf{x}_{\max} \in M$  mit

$$f(\mathbf{x}_{\min}) \leq f(\mathbf{x}) \quad \text{bzw.} \quad f(\mathbf{x}_{\max}) \geq f(\mathbf{x})$$

für alle  $\mathbf{x} \in M$ .

## 9 Differentialrechnung in mehreren Variablen

9.1 Vektorfolgen und ihre Grenzwerte

9.2 Grenzwerte von Funktionen und Stetigkeit

9.3 Darstellungsfragen, Anwendungen und Systematisierungsversuch zu Funktionen mehrerer Variablen

9.4 Differenzierbarkeit bei mehreren Variablen

9.5 Differentiation vektorwertiger Funktionen

9.6 Extrema von Funktionen mehrerer Variablen

# Multivariate Differentialrechnung

## Darstellung

Bisher sind wir meist vom allgemeinen Fall  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  ausgegangen, ohne  $m$  und  $n$  konkret zu wählen. Zum Erschließen der Theorie ist dies sinnvoll, da man alle Fälle in einem Zug behandeln kann.

In konkreten Anwendungen wird man sich jedoch immer zuerst klarmachen, welche Objekte als Argumente und welche als Funktionswerte auftreten. Generell ist eine gute Anschauung von Vorteil.

Wie man Funktionen mehrerer Variablen veranschaulicht, hängt aber stark von den gewählten Raumdimensionen  $m$  und  $n$  ab. Wir wollen dies systematischer ergründen.

Unter einem **skalaren Feld** versteht man eine reellwertige Funktion von mehreren Variablen, also vom Typ

$$f : \mathbb{R}^n \rightarrow \mathbb{R}.$$

Eine solche Funktion ordnet jedem Punkt/Vektor im Raum (oder in der Ebene) eine reelle Zahl als Funktionswert zu.

Typische Anwendungsbeispiele in den Naturwissenschaften sind:

- Temperatur- oder Druckverteilungen im Raum oder an Oberflächen,
- räumlich variable Leitfähigkeiten oder Dichten im Boden/in Probekörpern/im menschlichen Körper,
- ortsabhängige Potentiale (z. B. Gravitationspotential der Erde),
- Höhenreliefs (z. B. Geländeoberflächen).

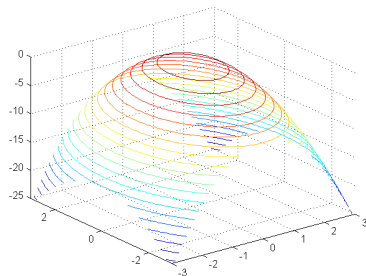
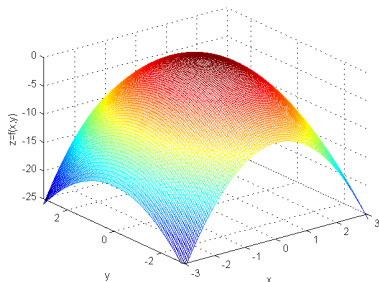
# Multivariate Differentialrechnung

## Skalare Felder

Eine vollständige Darstellungsmöglichkeit ergibt sich meist nur für den Fall  $n = 2$ . Hier kann man die Funktion in der Form

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

schreiben und als Fläche („Relief“) über der  $x$ - $y$ -Ebene visualisieren.



Relief zur Funktion  $f(x, y) = 1 - 2x^2 - y^2$  (vgl. S. 27). Rechts eine Variante, bei der nur Punkte zu bestimmten Höhenniveaus gezeichnet sind.

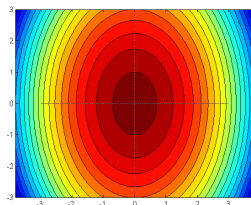
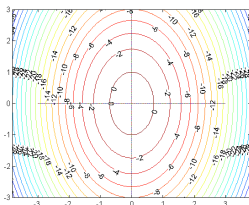
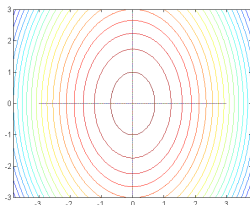
# Multivariate Differentialrechnung

## Skalare Felder

Eine weitere Möglichkeit ist die Darstellung von Höhenlinien in der Ebene (Karte, Konturplot). Diese Höhenlinien erhält man durch geometrische Interpretation der Gleichung

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

für verschiedene „Höhenniveaus“  $c \in \mathbb{R}$ .



Geben Sie eine analytische Darstellung für die hier dargestellten Höhenlinien der Funktion  $f(x, y) = 1 - 2x^2 - y^2$  an. Um welche geometrischen Figuren handelt es sich?

Konturplots findet man häufig zur Geländebeschreibung auf Landkarten:

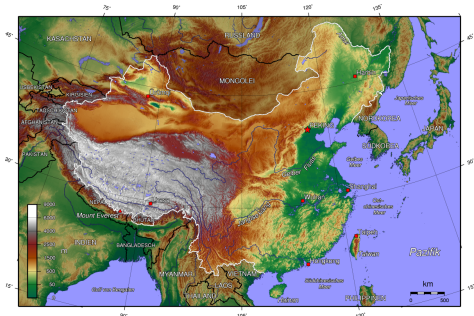
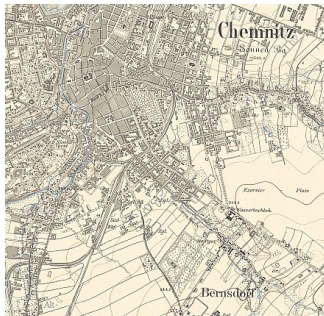


Bild links: Höhenlinien auf einem Messtischblatt von 1905 (Deutsche Fotothek, Kartenforum)

Bild rechts: farbige Geländedarstellung auf einer topografischen Karte von China (Captain Blood, Wikimedia Commons)



# Multivariate Differentialrechnung

## Skalare Felder

... oder auf Wetter- und Klimakarten:

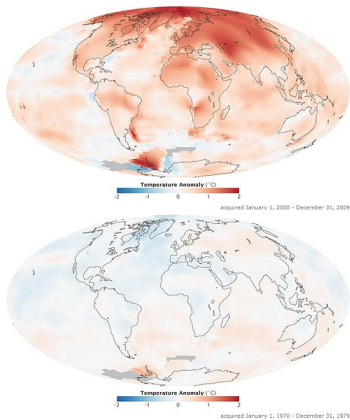
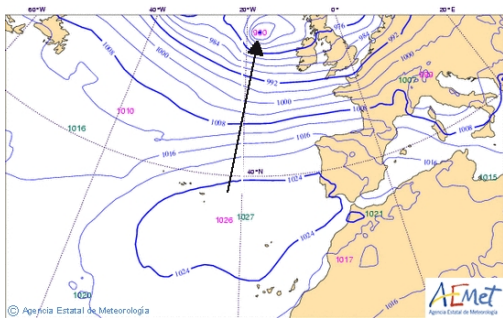


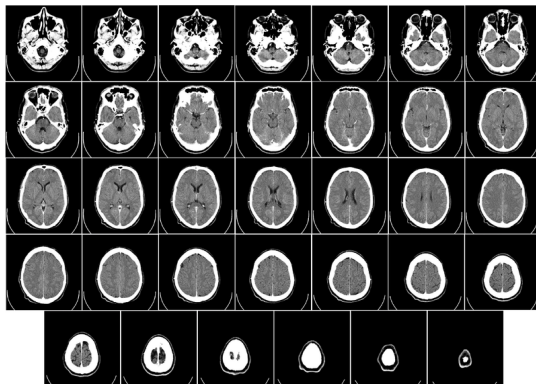
Bild links: Wetterkarte mit Isobaren (AEMET)

Bild rechts: Temperaturanomalien 1970-79 und 2000-09 im Vergleich zum Mittel von 1951-80 (NASA)

# Multivariate Differentialrechnung

## Skalare Felder

Für skalare Felder mit  $n = 3$  (oder gar  $n > 3$ ) ist die Darstellung schwieriger. Zum Beispiel kann man Schnittbilder zu erstellen. Dies geschieht häufig bei tomografischen Verfahren (CT, MRT, PET).



CT eines menschlichen Schädels. Dargestellt sind die ortsabhängigen Schwächungskoeffizienten für Röntgenstrahlung in verschiedenen transversalen Ebenen. (Bild: Uppsala University Hospital, Mikael Häggström)

Unter einem **Vektorfeld** versteht man eine vektorwertige Funktion vom Typ

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

Eine solche Funktion ordnet jedem Punkt/Vektor im Raum (oder in der Ebene) einen Vektor mit der gleichen Anzahl Komponenten zu.

Typische Beispiele in den Naturwissenschaften sind:

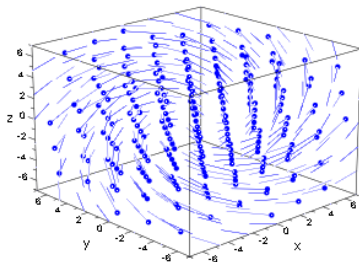
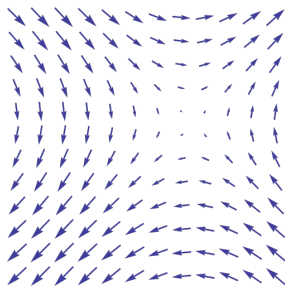
- (räumlich variable) elektrische oder magnetische Felder,
- Kraftfelder (z. B. die durch die Erde auf einen Satelliten wirkende Gravitationskraft),
- räumlich variable Geschwindigkeit eines Fluids in der Strömungsmechanik,
- Gradientenfelder (dazu später).

# Multivariate Differentialrechnung

## Vektorfelder

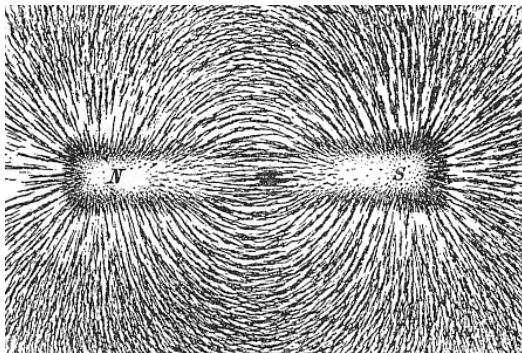
Bei der Darstellung von Vektorfeldern zeichnet man zumeist eine repräsentative Menge von Funktionswerten  $f(x)$  als Pfeile, die man an den zugehörigen Punkt  $x$  anheftet (quiver plot).

Die Pfeillänge gibt wie üblich den Betrag des Funktionswerts an.



Quiver plots der Vektorfelder  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $f(x, y) = \begin{bmatrix} \sin y \\ \sin x \end{bmatrix}$  und  $g : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $g(x, y, z) = \begin{bmatrix} -y \\ z \\ x \end{bmatrix}$ . (Bilder: Jim.belk/Mth77777)

„Quiver plots“ entstehen manchmal auf ganz natürliche Weise:



Die Eisenspäne richten sich an den Magnetfeldlinien aus und erzeugen einen ähnlichen Eindruck wie in unseren Plots.

Bild: Newton Henry Black (1913), Wikimedia Commons

Hier stehen zunächst Funktionen vom Typ

$$f : \mathbb{R} \rightarrow \mathbb{R}^n$$

zur Diskussion, an die man üblicherweise noch näher zu spezifizierende Glattheitsvoraussetzungen stellt.

Solche Funktionen beschreiben

- für  $n = 2$  Kurven in der Ebene,
- für  $n = 3$  Kurven im Raum.

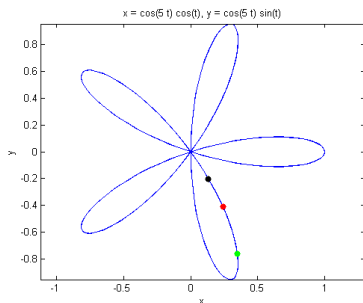
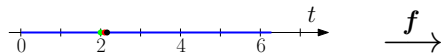
Zumeist interpretiert man das Argument  $t$  als Zeit, so dass  $f(t)$  den Ort zur Zeit  $t$  beim Durchlaufen der Kurve darstellt.

# Multivariate Differentialrechnung

## Parametrisierungen

Bei Kurven stellt man zumeist nur die Menge der Funktionswerte dar (rechtes Bild). Die Information über die zugehörigen Argumente (Zeiten) geht dabei verloren.

Bei Bedarf kann man aber zumindest einzelne Zeitpunkte wie im Bild rechts markieren (hier  $t = 2, 2.1, 2.15$ ).



Gezeigt ist ein Bild der Blütenblattkurve  $f : [0, 2\pi] \rightarrow \mathbb{R}^2, f(t) = \begin{bmatrix} \cos(5t) \cos t \\ \cos(5t) \sin t \end{bmatrix}$ .  
Beachten Sie, dass sich diese Kurve nicht als Funktion  $y = g(x)$  darstellen lässt.

Als letzten Punkt besprechen wir die Möglichkeit der Parametrisierung von Flächen im Raum. Dazu verwendet man Funktionen vom Typ

$$\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3,$$

wieder mit entsprechenden Glattheitseigenschaften. Wie bei Kurven wird wieder nur die Menge der Funktionswerte (also das Bild) von  $\mathbf{f}$  dargestellt. Allerdings benötigt man statt einem jetzt zwei Parameter.

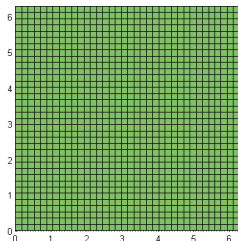
Ein Beispiel ist der Torus

$$\mathbf{f} : \mathbb{R}^2 \supset [0, 2\pi] \times [0, 2\pi] \rightarrow \mathbb{R}^3, \quad \mathbf{f}(u, v) = \begin{bmatrix} (10 + 5 \cos v) \cos u \\ (10 + 5 \cos v) \sin u \\ 5 \sin v \end{bmatrix},$$

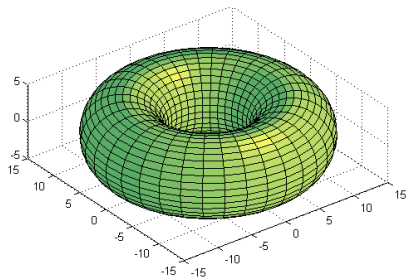
dessen Bild Sie auf der nächsten Seite finden.



### Grafische Darstellung



$f \rightarrow$



Gezeigt ist ein Bild des auf S. 52 beschriebenen Torus. Zu jedem Punkt des Rechtecks links finden Sie einen Punkt auf der Torusoberfläche.

Anschaulich können Sie sich das Netz (links) als Gummihaut vorstellen, welche Sie über den Torus (rechts) spannen.

Der Torus besitzt keine Funktionsdarstellung der Form  $z = h(x, y)$ .

## 9 Differentialrechnung in mehreren Variablen

9.1 Vektorfolgen und ihre Grenzwerte

9.2 Grenzwerte von Funktionen und Stetigkeit

9.3 Darstellungsfragen, Anwendungen und Systematisierungsversuch zu Funktionen mehrerer Variablen

9.4 Differenzierbarkeit bei mehreren Variablen

9.5 Differentiation vektorwertiger Funktionen

9.6 Extrema von Funktionen mehrerer Variablen

# Multivariate Differentialrechnung

## Wiederholung

Beim Verallgemeinern des Differenzierbarkeitsbegriffs versuchen wir, analog zum eindimensionalen Fall vorzugehen. Daher wiederholen wir zunächst die wichtigsten Begriffe.

Eine reelle Funktion  $f : \mathbb{R} \supset D_f \rightarrow \mathbb{R}$  heißt in  $z \in D_f$  differenzierbar, wenn der Grenzwert

$$f'(z) := \lim_{h \rightarrow 0} \frac{f(z+h) - f(z)}{h} \quad (9.4)$$

existiert. Diesen hatten wir erste Ableitung von  $f$  an der Stelle  $z$  genannt.

# Multivariate Differentialrechnung

## Wiederholung

Ersetzt man  $z$  und  $h$  gedanklich durch Vektoren  $z$  und  $h$ , sieht man, dass Gleichung (9.4) für eine Verallgemeinerung nicht taugt – denn durch Vektoren kann man nicht teilen.

Verallgemeinerungspotential hat dagegen die äquivalente Charakterisierung aus Satz 5.3:

$f$  ist genau dann in  $z$  differenzierbar, wenn es eine Zahl  $a$  (das ist genau die Ableitung  $f'(z)$ ) und eine Funktion  $\phi : D_f \rightarrow \mathbb{R}$  gibt, so dass

$$f(x) = f(z) + a(x - z) + \phi(x) \quad \text{mit} \quad \frac{\phi(x)}{|x - z|} \rightarrow 0 \text{ für } x \rightarrow z. \quad (9.5)$$

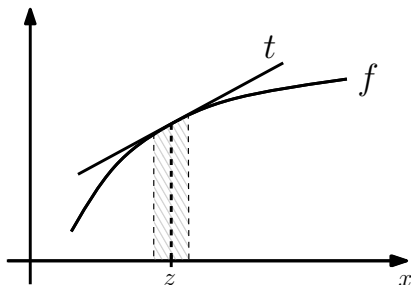
# Multivariate Differentialrechnung

## Wiederholung

Das Kriterium (9.5) sagt, grob gesprochen, dass  $f(x)$  im Falle der Differenzierbarkeit nahe  $z$  gut durch eine Tangente  $t(x)$  approximiert wird:

$$f(x) \approx \underbrace{f(z) + f'(z)(x - z)}_{=: t(x)} \quad \text{für } x \approx z.$$

Graphisch:



# Multivariate Differentialrechnung

## Partielle Ableitungen

In diesem Abschnitt betrachten wir stets reellwertige Funktionen mehrerer Variablen, genauer Funktionen

$$f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$$

mit offenem Definitionsbereich  $D_f$ .

In einem ersten Ansatz wollen wir lediglich eine Komponente  $x_j$  im Argument von  $f$  variieren, während wir die anderen Komponenten festhalten (d. h. als Parameter behandeln).

Dieser Ansatz führt uns direkt zum Begriff der partiellen Ableitung. Wir benötigen dafür lediglich den eindimensionalen Ableitungsbegriff.

### Definition 9.11

Eine Funktion  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  heißt in  $z \in D_f$  partiell nach  $x_j$  differenzierbar, wenn die  $j$ -te partielle Funktion\*

$$f_j : \mathbf{x} \mapsto f(z_1, \dots, z_{j-1}, \mathbf{x}, z_{j+1}, \dots, z_n)$$

in  $z_j$  differenzierbar ist. Die Zahl

$$\frac{\partial f}{\partial x_j}(z) := f'_j(z_j) \tag{9.6}$$

heißt die  $j$ -te **partielle Ableitung** von  $f$  an der Stelle  $z$ .

Eine Funktion  $f$  heißt an der Stelle  $z$  **partiell differenzierbar**, wenn sie in  $z$  nach allen Variablen  $x_1, \dots, x_n$  partiell differenzierbar ist.

\*Beachten Sie, dass es sich hierbei um Funktionen einer Variablen handelt!

Weitere Schreibweisen für die partiellen Ableitungen aus (9.6):

$$\frac{\partial f(\mathbf{z})}{\partial x_j} \quad \text{oder} \quad f_{x_j}(\mathbf{z}),$$

seltener auch  $\partial_j f(\mathbf{z})$  oder  $\partial_{x_j} f(\mathbf{z})$ .

Natürlich kann man auch die zugrundeliegenden Differentialquotienten notieren:

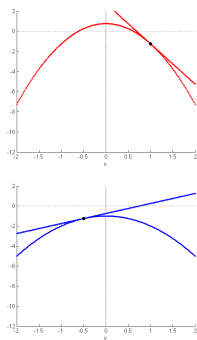
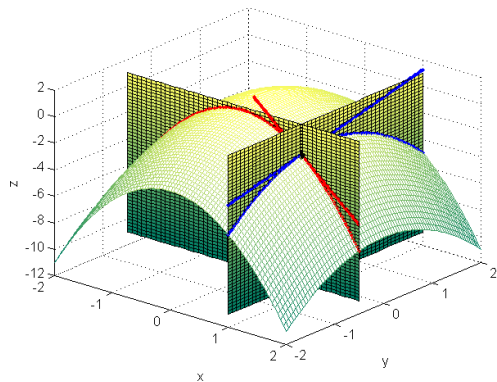
$$\begin{aligned} \frac{\partial f}{\partial x_j}(\mathbf{z}) &= \lim_{h \rightarrow 0} \frac{f(z_1, \dots, z_{j-1}, \mathbf{z}_j + h, z_{j+1}, \dots, z_n) - f(z_1, \dots, z_{j-1}, \mathbf{z}_j, z_{j+1}, \dots, z_n)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(\mathbf{z} + h\mathbf{e}_j) - f(\mathbf{z})}{h}. \end{aligned}$$

Aufgrund letzterer Darstellung spricht man auch von der Ableitung in Richtung des  $j$ -ten Einheitsvektors.



# Multivariate Differentialrechnung

## Visualisierung



Graphen der Funktion  $f(x, y) = 1 - 2x^2 - y^2$  sowie der partiellen Funktionen  $f_1(x)$  (rot) und  $f_2(y)$  (blau) zum Punkt  $(1, -\frac{1}{2})$ .

Die partiellen Ableitungen  $\frac{\partial f}{\partial x}(1, -\frac{1}{2})$  und  $\frac{\partial f}{\partial y}(1, -\frac{1}{2})$  entsprechen den Anstiegen der eingezeichneten Tangenten.

# Multivariate Differentialrechnung

## Berechnung der partiellen Ableitungen

Zur Berechnung der partiellen Ableitung  $\frac{\partial f}{\partial x_j}$  leitet man  $f$  nach  $x_j$  ab und behandelt alle anderen Variablen als Konstanten.

Dabei gelten die gewohnten Ableitungsregeln.

**Beispiel:** Die partiellen Ableitungen zu  $f(x, y) = 1 - 2x^2 - y^2$  lauten

$$\frac{\partial f}{\partial x}(x, y) = -4x \quad \text{und} \quad \frac{\partial f}{\partial y}(x, y) = -2y.$$

Die Anstiege der auf S. 61 visualisierten Tangenten an die partiellen Funktionen sind also

$$\frac{\partial f}{\partial x}\left(1, -\frac{1}{2}\right) = -4 \text{ (rot)} \quad \text{und} \quad \frac{\partial f}{\partial y}\left(1, -\frac{1}{2}\right) = 1 \text{ (blau)}.$$

# Multivariate Differentialrechnung

## Berechnung der partiellen Ableitungen

Man berechne sämtliche partiellen Ableitungen von

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = x^2 + 2y,$
- $g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g(x, y) = x \sin y + y \cos x,$
- $h : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad h(x, y, z) = y(1 + z^2)^2.$

Wie lauten die partiellen Ableitungen der Funktionen

- $f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(\mathbf{x}) = \|\mathbf{x}\|^2,$
- $g : \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}, \quad g(\mathbf{x}) = \|\mathbf{x}\|?$

# Multivariate Differentialrechnung

## Partielle Ableitungen höherer Ordnung

Die partiellen Ableitungen einer Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  sind wiederum Funktionen vom Typ  $\frac{\partial f}{\partial x_j} : \mathbb{R}^n \rightarrow \mathbb{R}$ .

Man kann sie also ggf. erneut partiell ableiten – möglicherweise nach jeder der Variablen  $x_k$  ( $k = 1, \dots, n$ ).

Die dabei entstehenden partiellen Ableitungen höherer Ordnung bezeichnet man mit

$$\frac{\partial^2 f}{\partial x_j^2}, \frac{\partial^2 f}{\partial x_k \partial x_j}, \frac{\partial^3 f}{\partial x_j^3}, \frac{\partial^3 f}{\partial x_k^2 \partial x_j}, \frac{\partial^3 f}{\partial x_k \partial x_j^2}, \frac{\partial^3 f}{\partial x_l \partial x_k \partial x_j}, \dots$$

bzw.  $f_{x_j x_j}$ ,  $f_{x_j x_k}$  etc.

Die Differentiationsreihenfolge liest man im „Nenner“ von rechts nach links; d. h. die zuletzt ausgeführte Differentiation steht ganz rechts im Nenner. Sie spielt jedoch in den meisten praktischen Fällen keine Rolle.

Berechnen Sie zu

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = 1 - 2x^2 - y^2,$
- $g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g(x, y) = x \sin y + ye^{2x},$

sämtliche partiellen Ableitungen bis zur zweiten Ordnung.

Was können Sie beobachten?

Wir gehen dem beobachteten Phänomen näher auf den Grund. Dafür benötigen wir den Begriff der stetigen Differenzierbarkeit.

### Definition 9.12

Eine Funktion  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  heißt in  $z \in D_f$  ( $k$ -mal) **stetig partiell differenzierbar**, wenn alle partiellen Ableitungen (bis zur  $k$ -ten Ordnung) an der Stelle  $z$  existieren und dort stetig sind.

$f$  heißt auf einer offenen Menge  $\Omega \subset D_f$  stetig differenzierbar, wenn  $f$  an jeder Stelle  $z \in \Omega$  stetig partiell differenzierbar ist (Schreibweise  $f \in C^k(\Omega)$ ).

Ist die Funktion  $f(x, y) = 1 - 2x^2 - y^2$  (zweimal) stetig partiell differenzierbar?

### Satz 9.13 (von Schwarz\*)

Sei  $D \subset \mathbb{R}^n$  offen und  $f : D \rightarrow \mathbb{R}$  zweimal stetig partiell differenzierbar, d.h.  $f \in C^2(D)$ . Dann gilt

$$\frac{\partial^2 f}{\partial x_j \partial x_k}(z) = \frac{\partial^2 f}{\partial x_k \partial x_j}(z) \quad \text{für alle } j, k \in \{1, \dots, n\} \text{ und alle } z \in D.$$



\*Hermann Amandus Schwarz, 1843-1921,  
deutscher Mathematiker

# Multivariate Differentialrechnung

## Vertauschbarkeit der Differentiationsreihenfolge

Bei zweimal stetig partiell differenzierbaren Funktionen (bei uns der Normalfall) spielt also die Differentiationsreihenfolge beim zweimaligen partiellen Differenzieren keine Rolle.

Das Ergebnis überträgt sich mühelos auf  $k$ -mal stetig partiell differenzierbare Funktionen und partielle Ableitungen bis zur Ordnung  $k$ .

Berechnen Sie sämtliche partiellen Ableitungen bis zur dritten Ordnung von

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = x \sin(xy).$$



# Multivariate Differentialrechnung

Differenzierbarkeit, Gradient, Richtungsableitung und totales Differential

Wie im vorherigen Abschnitt behandeln wir wieder reellwertige Funktionen

$$f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$$

mit offenem Definitionsbereich  $D_f$ .

Ziel ist die Verallgemeinerung des eindimensionalen Ableitungsbegriffs auf diese Funktionen. Dabei werden wir uns an die Approximierbarkeit von  $f$  durch lineare Funktionen halten.

Zunächst benötigen wir aber noch einen weiteren Begriff.

### Definition 9.14

Sei  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  in  $z$  partiell differenzierbar, dann heißt

$$\nabla f(z) := \begin{bmatrix} \frac{\partial f}{\partial x_1}(z) \\ \vdots \\ \frac{\partial f}{\partial x_n}(z) \end{bmatrix}$$

der **Gradient** von  $f$  an der Stelle  $z$ .

Alternative Schreibweise:  $\text{grad } f(z)$ .

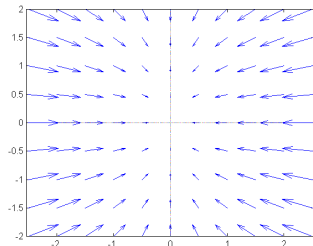
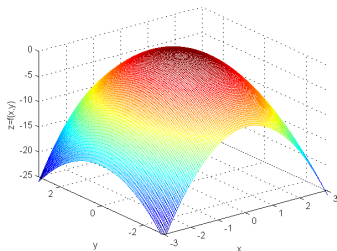
#### Anmerkung:

Das Symbol „ $\nabla$ “ wird „Nabla“ ausgesprochen. Dabei handelt es sich nicht um einen Buchstaben, sondern um eine antike Harfe mit ähnlicher Gestalt.

# Multivariate Differentialrechnung

## Visualisierung

Der Gradient ist ein Vektorfeld – für eine Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  also eine Funktion  $\nabla f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ :



Graph der Funktion  $f(x, y) = 1 - 2x^2 - y^2$  und zugehöriger Gradient.

Wir werden später sehen, dass der Gradient die Richtung des steilsten Anstiegs von  $f$  markiert.

Berechnen Sie den Gradienten der Funktion  $f(x, y) = 1 - 2x^2 - y^2$ .

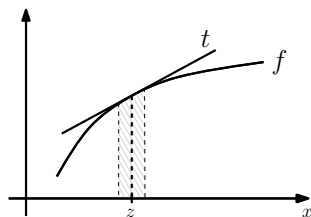
# Multivariate Differentialrechnung

## Differenzierbarkeit

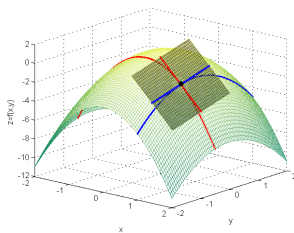
Kommen wir nun zur (totalen) Ableitung.

Im Eindimensionalen bedeutete Differenzierbarkeit von  $f$  in  $z$ , dass  $f$  nahe  $z$  gut durch die Tangente approximiert wird.

Für Funktionen  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  tritt an deren Stelle eine Tangentialebene:



1D



2D

Diese Idee fassen wir jetzt mathematisch exakt und schließen auch höhere Dimensionen mit ein. Ausgangspunkt ist die Charakterisierung aus Satz 5.3.

### Definition 9.15

Eine Funktion  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  heißt in  $\mathbf{z} \in D_f$  (total) **differenzierbar**, wenn es einen Vektor  $\mathbf{a} \in \mathbb{R}^n$  sowie eine Funktion  $\phi : D_f \rightarrow \mathbb{R}$  gibt, so dass

$$f(\mathbf{x}) = f(\mathbf{z}) + \mathbf{a}^T(\mathbf{x} - \mathbf{z}) + \phi(\mathbf{x}), \quad (9.7)$$

mit  $\frac{\phi(\mathbf{x})}{\|\mathbf{x} - \mathbf{z}\|} \rightarrow 0$  für  $\mathbf{x} \rightarrow \mathbf{z}$ .

$f'(\mathbf{z}) := \mathbf{a}^T$  heißt dann die (totale) **Ableitung** von  $f$  an der Stelle  $\mathbf{z}$ .

**Interpretation** Gleichung (9.14) besagt gerade, dass

$$f(\mathbf{x}) \approx f(\mathbf{z}) + f'(\mathbf{z})(\mathbf{x} - \mathbf{z}) \quad \text{für } \mathbf{x} \approx \mathbf{z}.$$

Die Funktion  $\phi$  ist dabei der Approximationsfehler, der für  $\mathbf{z} \rightarrow \mathbf{x}$  schneller als linear gegen Null strebt (also für  $\mathbf{x} \approx \mathbf{z}$  sehr klein ist).

Zur konkreten Berechnung der Ableitung verwendet man allerdings nicht Definition 9.15, sondern folgenden Satz:

### Satz 9.16

Ist  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  in  $z \in D_f$  differenzierbar, so ist  $f$  in  $z$  auch partiell differenzierbar, und es gilt

$$f'(z) = \left[ \frac{\partial f}{\partial x_1}(z), \dots, \frac{\partial f}{\partial x_n}(z) \right] = \nabla f(z)^T. \quad (9.8)$$

Im Fall der Existenz ist die Ableitung also die (Transponierte des) Gradienten.

**Anschauliche Begründung von (9.8):** Die Tangenten der partiellen Funktionen (vgl. S. 61) müssen in der Tangentialebene liegen.

Woher wissen wir aber, dass  $f$  total differenzierbar ist? Partielle Differenzierbarkeit allein reicht dafür nicht!

Es reicht aber, nur ein wenig mehr zu fordern:

### Satz 9.17

Ist  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  in einer **offenen Umgebung** des Punktes  $z \in D_f$  partiell differenzierbar und in  $z$  selbst **stetig** partiell differenzierbar, so ist  $f$  in  $z$  auch total differenzierbar.

Insbesondere ist  $f$  auf ganz  $D_f$  differenzierbar, wenn  $f$  auf ganz  $D_f$  stetig partiell differenzierbar ist. Für die meisten praktisch relevanten Beispiele ist dieses Kriterium erfüllt.

# Multivariate Differentialrechnung

## Differenzierbarkeit, Beispiel

Die Funktion

$$f(x, y) = 1 - 2x^2 - y^2$$

ist auf ganz  $\mathbb{R}^2$  stetig partiell differenzierbar mit

$$f_x(x, y) = -4x \quad \text{und} \quad f_y(x, y) = -2y.$$

Damit ist  $f$  auf ganz  $\mathbb{R}^2$  auch (total) differenzierbar mit

$$f'(x, y) = \nabla f(x, y)^T = [-4x, -2y].$$

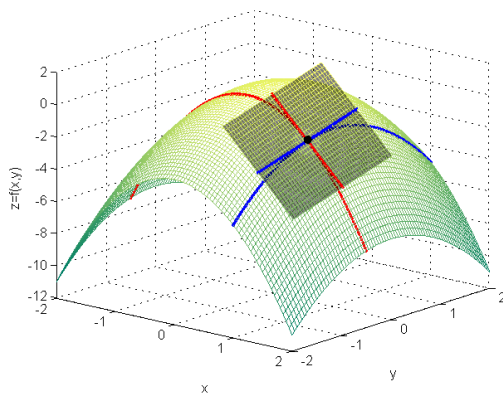
Die Gleichung der Tangentialebene im Punkt  $(1, -\frac{1}{2})$  lautet

$$\begin{aligned} t(x, y) &= f\left(1, -\frac{1}{2}\right) + \nabla f\left(1, -\frac{1}{2}\right)^T \begin{bmatrix} x - 1 \\ y + \frac{1}{2} \end{bmatrix} \\ &= -1.25 + [-4, 1] \begin{bmatrix} x - 1 \\ y + \frac{1}{2} \end{bmatrix} \\ &= -4x + y + 5.75. \end{aligned}$$



# Multivariate Differentialrechnung

## Differenzierbarkeit, Bild zum Beispiel



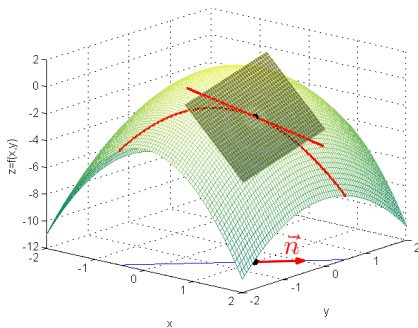
Bemerken Sie auch, dass die Tangenten der partiellen Funktionen von S. 59 in der Tangentialebene liegen.

# Multivariate Differentialrechnung

## Richtungsableitungen

Wenn man sich in unserem Beispiel  $f(x, y) = 1 - 2x^2 - y^2$  vom Punkt  $(1, -\frac{1}{2})$  aus in Richtung eines beliebigen Vektors  $\mathbf{n}$  ( $\|\mathbf{n}\| = 1$ ) bewegt, entsteht auf dem Graphen von  $f$  wieder eine reelle Funktion.

Dieser kann man wieder eine Tangente zuordnen, die natürlich wieder in der Tangentialebene enthalten ist. Ihr Anstieg ist die Richtungsableitung.



Mathematisch exakt:

### Definition 9.18

Eine Funktion  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  heißt in  $z \in D_f$  in Richtung eines Vektors  $\mathbf{n} \in \mathbb{R}^n$  ( $\|\mathbf{n}\| = 1$ ) differenzierbar, wenn der Grenzwert

$$\frac{\partial f}{\partial \mathbf{n}}(z) := \lim_{h \rightarrow 0} \frac{f(z + h\mathbf{n}) - f(z)}{h}$$

existiert.

Wir nennen  $\frac{\partial f}{\partial \mathbf{n}}(z)$  Richtungsableitung von  $f$  an der Stelle  $z$  in Richtung  $\mathbf{n}$ .

# Multivariate Differentialrechnung

## Richtungsableitungen

Da der Gradient das Anstiegsverhalten der Tangential(hyper)ebene vollständig beschreibt, kann man aus  $\nabla f$  und  $\mathbf{n}$  die Richtungsableitung berechnen:

### Satz 9.19

Ist  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  in  $\mathbf{z} \in D_f$  differenzierbar, so existiert zu jeder Richtung  $\mathbf{n}$  (mit  $\|\mathbf{n}\| = 1$ ) die Richtungsableitung, und es gilt

$$\frac{\partial f}{\partial \mathbf{n}}(\mathbf{z}) = \nabla f(\mathbf{z})^T \mathbf{n}. \quad (9.9)$$

Man berechne die Ableitung der Funktion  $f(x, y) = 1 - 2x^2 - y^2$  an der Stelle  $(1, -\frac{1}{2})$  in Richtung  $\mathbf{n} = \frac{1}{\sqrt{2}}[1, 1]^T$ . (Das Bild zu diesem Beispiel finden Sie auf S. 78.)

Aus den Eigenschaften des Skalarprodukts folgt nun für festes  $\mathbf{z} \in D_f$ :

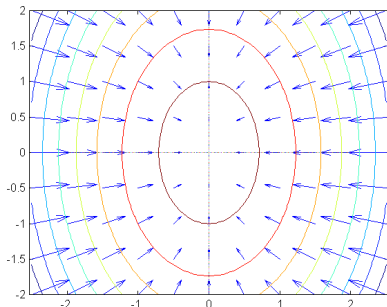
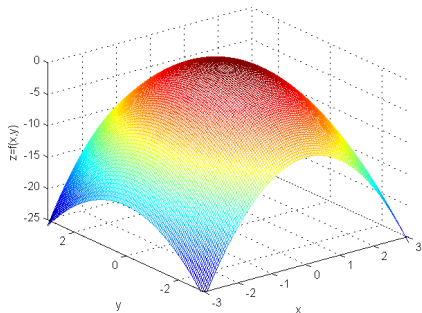
- $\frac{\partial f}{\partial \mathbf{n}}(\mathbf{z})$  wird maximal, wenn  $\mathbf{n}$  in Richtung des Gradienten zeigt, d. h. wenn  $\mathbf{n} = \frac{\nabla f(\mathbf{z})}{\|\nabla f(\mathbf{z})\|}$ .
- Es gilt  $\frac{\partial f}{\partial \mathbf{n}}(\mathbf{z}) = 0$  genau dann, wenn  $\mathbf{n}$  senkrecht auf dem Gradienten  $\nabla f(\mathbf{z})$  steht.

Einfacher ausgedrückt:

- Der Gradient zeigt in Richtung des steilsten Anstiegs von  $f$ .
- Der Gradient steht senkrecht auf den Höhenlinien von  $f$ , genauer auf der Tangente an die Höhenlinie durch  $\mathbf{z}$ .

# Multivariate Differentialrechnung

## Illustration



Der Gradient zeigt in die Richtung des steilsten Anstiegs und steht senkrecht auf den Höhenlinien. Je länger der Gradientenvektor, desto steiler verläuft der Graph von  $f$ .

Gezeichnet ist wieder das Beispiel  $f(x, y) = 1 - 2x^2 - y^2$ .

# Multivariate Differentialrechnung

## Illustration

Berechnen Sie für  $f(x, y) = 1 - 2x^2 - y^2$  den maximalen Wert, den eine Richtungsableitung  $\frac{\partial f}{\partial \vec{n}}(1, -\frac{1}{2})$  annehmen kann.

Formulieren und beweisen Sie eine allgemeingültige Aussage.

In welche Richtung  $\vec{n}$  ist der Anstieg einer differenzierbaren Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  an der Stelle  $\vec{z}$  minimal?

Welcher Zusammenhang besteht zwischen der partiellen Ableitung  $\frac{\partial f}{\partial x_j}(\vec{z})$  und Richtungsableitungen?

Wie lässt sich die Ableitung einer Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  in Richtung  $[0, -1]^T$  am einfachsten über partielle Ableitungen ausdrücken?

Wir stellen hier noch einmal kurz den Zusammenhang zwischen den einzelnen Differenzierbarkeitsbegriffen für reellwertige Funktionen mehrerer Variablen dar:

$f$  stetig partiell differenzierbar in einer offenen Umgebung von  $z$



$f$  (total) differenzierbar in  $z$



Existenz sämtlicher Richtungsableitungen von  $f$  in  $z$



$f$  partiell differenzierbar in einer offenen Umgebung von  $z$

### Faustregel:

Besitzt  $f$  auf ganz  $D_f$  stetige partielle Ableitungen nach allen Variablen, dann ist alles im grünen Bereich.



# Multivariate Differentialrechnung

## Totales Differential und elementare Fehlerrechnung

Wie im Eindimensionalen kann man zu einem gegebenen Vektor  $d\mathbf{x}$  und einer differenzierbaren Funktion  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  das **totale Differential** an der Stelle  $\mathbf{x} \in D_f$  definieren:

$$df(\mathbf{x}) := f'(\mathbf{x}) d\mathbf{x} \quad (9.10)$$

Es beschreibt die Änderung des Funktionswerts der Tangential(hyper)ebene, wenn man von  $\mathbf{x}$  zu  $\mathbf{x} + d\mathbf{x}$  übergeht.

Für kleine  $d\mathbf{x}$  unterscheiden sich Tangential(hyper)ebene und Funktion im Punkt  $\mathbf{x} + d\mathbf{x}$  kaum. Für die Änderung der tatsächlichen Funktionswerte gilt also

$$\Delta f \approx df(\mathbf{x}) \quad \text{für } d\mathbf{x} \text{ klein.}$$

# Multivariate Differentialrechnung

## Anwendung in der elementaren Fehlerrechnung

Bei der Auswertung von Experimenten möchte man oft aus den Messfehlern  $\Delta x_1, \dots, \Delta x_n$  zu Messgrößen  $x_1, \dots, x_n$  auf den resultierenden Fehler  $\Delta f$  einer Größe  $f(x_1, \dots, x_n)$  schließen.

Im totalen Differential

$$df(x_1, \dots, x_n) = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \dots + \frac{\partial f}{\partial x_n} dx_n$$

aus (9.10) identifiziert man dazu  $dx_j$  mit dem Messfehler  $\Delta x_j$  sowie  $df$  mit dem Fehler  $\Delta f$ , und erhält die Abschätzung

$$|\Delta f| \leq \left| \frac{\partial f}{\partial x_1} \right| |\Delta x_1| + \left| \frac{\partial f}{\partial x_2} \right| |\Delta x_2| + \dots + \left| \frac{\partial f}{\partial x_n} \right| |\Delta x_n|. \quad (9.11)$$

# Multivariate Differentialrechnung

## Anwendung in der elementaren Fehlerrechnung

Misst man die Zeit  $t$ , in der ein anfangs ruhender Körper im freien Fall (im Vakuum) den Weg  $s$  zurückzulegen, kann man daraus über

$$g = \frac{2s}{t^2} \quad (9.12)$$

die Fallbeschleunigung  $g$  bestimmen.

In einer Einzelmessung ergeben sich die Werte  $s = (0.5000 \pm 0.0002)\text{m}$  und  $t = (0.319 \pm 0.001)\text{s}$ . Bestimmen Sie daraus  $g$ , und schätzen Sie den zugehörigen Fehler mit (9.11) ab.

Wird man zur Genauigkeitsverbesserung eher an der Weg- oder eher an der Zeitmessung arbeiten müssen?

# Multivariate Differentialrechnung

## Anwendung in der elementaren Fehlerrechnung

Besonders einfach wird die Analyse für Funktionen vom Typ

$$f(x_1, \dots, x_n) = x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \dots \cdot x_n^{\alpha_n} \quad (\alpha_1, \dots, \alpha_n \in \mathbb{R}).$$

Hier gilt für den relativen Fehler

$$\left| \frac{\Delta f}{f} \right| \leq |\alpha_1| \left| \frac{\Delta x_1}{x_1} \right| + |\alpha_2| \left| \frac{\Delta x_2}{x_2} \right| + \dots + |\alpha_n| \left| \frac{\Delta x_n}{x_n} \right|. \quad (9.13)$$

Beachten Sie, dass mit (9.13) auch Quotienten und Wurzeln behandelt werden können.

Bestätigen Sie Formel (9.13) für den Fall  $n = 2$ .

Betrachten Sie das Beispiel von S. 87 mit Hilfe von (9.13) erneut. Machen Sie sich klar, wieviel Aufwand man damit gegenüber der Rechnung von S. 87 spart.

# Multivariate Differentialrechnung

## Zur Einordnung der Methode

Die hier vorgestellte Methode zur Fehlerapproximation ist sehr konservativ und geht vom ungünstigsten Fall der Fehlerverteilung aus.

Sie bietet sich vor allem bei Einzelmessungen und zur groben Abschätzung in einfachen Experimenten an.

Eine fortgeschrittene Fehlerrechnung berücksichtigt dagegen auch den zufälligen Charakter der Messfehler.

Der Umgang mit Messfehlern bei der Fehlerfortpflanzung wird im „Guide to the Expression of Uncertainty in Measurement“ (GUM) beschrieben, der in Deutschland Normcharakter hat.

## 9 Differentialrechnung in mehreren Variablen

9.1 Vektorfolgen und ihre Grenzwerte

9.2 Grenzwerte von Funktionen und Stetigkeit

9.3 Darstellungsfragen, Anwendungen und Systematisierungsversuch zu Funktionen mehrerer Variablen

9.4 Differenzierbarkeit bei mehreren Variablen

9.5 Differentiation vektorwertiger Funktionen

9.6 Extrema von Funktionen mehrerer Variablen

# Multivariate Differentialrechnung

## Differentiation vektorwertiger Funktionen

Wir kommen nun zum allgemeinsten Fall, d. h. zu Funktionen vom Typ

$$\mathbf{f} : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m.$$

Auch im diesem Abschnitt setzen wir stillschweigend einen offenen Definitionsbereich  $D$  voraus.

Solche Funktionen ordnen also jedem Vektor  $\mathbf{x} \in D$  mit  $n$  Komponenten einen Vektor  $\mathbf{f}(\mathbf{x})$  mit  $m$  Komponenten zu:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x} \quad \mapsto \quad \mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{bmatrix}.$$

Wir nennen die Funktionen  $f_j : D \rightarrow \mathbb{R}$  **Komponentenfunktionen** von  $\mathbf{f}$ .

Sämtliche Differenzierbarkeitsbegriffe greifen letztlich auf Grenzwerte zurück.

Eine vektorwertige Funktion  $f$  besitzt für  $x \rightarrow z$  genau dann einen Grenzwert, wenn sämtliche Komponentenfunktionen  $f_j$  ( $j = 1, \dots, m$ ) für  $x \rightarrow z$  einen Grenzwert besitzen (vgl. 9.1 und 9.2).

Daher kann man die Differenzierbarkeitsbegriffe für vektorwertige Funktionen komponentenweise aufbauen und sich die Betrachtungen für reellwertige Funktionen zum Vorbild nehmen.

Wir beginnen also wieder mit partieller Differenzierbarkeit und gehen dann zur (totalen) Differenzierbarkeit weiter.



### Definition 9.20

Eine Funktion  $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m$  heißt an der Stelle  $z$  **partiell differenzierbar**, wenn alle Komponentenfunktionen  $f_j$  ( $j = 1, \dots, m$ ) in  $z$  partiell differenzierbar sind.

Unter der **partiellen Ableitung** von  $f$  nach der Variablen  $x_j$  an der Stelle  $z$  verstehen wir den Vektor

$$\frac{\partial f}{\partial x_j}(z) = \begin{bmatrix} \frac{\partial f_1}{\partial x_j}(z) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(z) \end{bmatrix}.$$

$f$  heißt in  $z$  **stetig partiell differenzierbar**, wenn alle partiellen Ableitungen von  $f$  in einer Umgebung von  $z$  existieren und in  $z$  stetig sind.

Man berechne alle partiellen Ableitungen der Funktion

$$\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad \mathbf{f}(x, y) = \begin{bmatrix} 3x^2 + y \\ x \sin y \\ ye^{x^2+y} \end{bmatrix}.$$

Ist  $\mathbf{f}$  stetig partiell differenzierbar?

Wie im reellwertigen Fall gibt partielle Differenzierbarkeit noch keine Garantie dafür, dass man  $\mathbf{f}$  nahe  $\mathbf{z}$  durch eine affin lineare Funktion

$$\mathbf{t}(\mathbf{x}) = \mathbf{f}(\mathbf{z}) + \mathbf{f}'(\mathbf{z})(\mathbf{x} - \mathbf{z})$$

approximieren kann. Dafür benötigt man wieder totale Differenzierbarkeit.

### Definition 9.21

Eine Funktion  $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m$  heißt in  $z \in D$  (total) **differenzierbar**, wenn es eine Matrix  $A \in \mathbb{R}^{m \times n}$  sowie eine Funktion  $\phi : D \rightarrow \mathbb{R}^m$  gibt, so dass

$$f(x) = f(z) + A(x - z) + \phi(x), \quad \text{mit } \frac{\phi(x)}{\|x - z\|} \rightarrow \mathbf{0} \text{ für } x \rightarrow z. \quad (9.14)$$

$f'(z) := A$  heißt dann die (totale) **Ableitung** von  $f$  an der Stelle  $z$ .

Entscheidendes Kriterium ist also auch hier wieder die Approximierbarkeit durch affin lineare Funktionen.

Wie im reellwertigen Fall folgt aus totaler Differenzierbarkeit die partielle Differenzierbarkeit.

# Multivariate Differentialrechnung

## Totale Differenzierbarkeit

Analog zum reellwertigen Fall benutzt man zur Entscheidung über Differenzierbarkeit und zur Berechnung der Ableitung folgende Sätze:

### Satz 9.22

Ist  $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m$  in  $D$  **stetig** partiell differenzierbar, so ist  $f$  in  $D$  auch total differenzierbar.

### Satz 9.23

Ist  $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m$  in  $z \in D$  (total) differenzierbar, so hat die Ableitung die Form

$$f'(z) = \begin{bmatrix} \nabla f_1(z)^T \\ \nabla f_2(z)^T \\ \vdots \\ \nabla f_m(z)^T \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1(z)}{\partial x_1} & \frac{\partial f_1(z)}{\partial x_2} & \cdots & \frac{\partial f_1(z)}{\partial x_n} \\ \frac{\partial f_2(z)}{\partial x_1} & \frac{\partial f_2(z)}{\partial x_2} & \cdots & \frac{\partial f_2(z)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(z)}{\partial x_1} & \frac{\partial f_m(z)}{\partial x_2} & \cdots & \frac{\partial f_m(z)}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

(**Funktional-** oder **Jacobi-Matrix**).

# Multivariate Differentialrechnung

## Totale Differenzierbarkeit

Machen Sie sich klar, dass der in Folie 69 ff. behandelte Ableitungsbegriff für reellwertige Funktionen einen Spezialfall von Definition 9.21 bzw. Satz 9.23 darstellt.

Berechnen Sie die Ableitungen (Jacobi-Matrizen) zu  $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  und  $\mathbf{g} : \mathbb{R} \rightarrow \mathbb{R}^2$  mit

$$\mathbf{f}(x, y) = \begin{bmatrix} 3x^2 + y \\ x \sin y \\ ye^{x^2+1} \end{bmatrix} \quad \text{und} \quad \mathbf{g}(t) = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}.$$

Zeigen Sie mit Hilfe von Definition 9.21, dass für

$$\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \mathbf{f}(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

mit  $A \in \mathbb{R}^{m \times n}$  und  $\mathbf{b} \in \mathbb{R}^m$  die Beziehung  $\mathbf{f}'(\mathbf{x}) = A$  gilt.

# Multivariate Differentialrechnung

## Eigenschaften differenzierbarer Funktionen, Ableitungsregeln

Wie im eindimensionalen Fall gilt für differenzierbare Funktionen mehrerer Variablen:

### Satz 9.24

Ist die Funktion  $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m$  in  $z \in D$  differenzierbar, so ist sie in  $z$  auch stetig.

Dabei wird der reellwertige Fall  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  selbstverständlich mit erfasst.

**Beweisidee** Für  $x \rightarrow z$  gilt  $f(x) = f(z) + \underbrace{f'(z)(x-z)}_{\rightarrow 0} + \underbrace{\phi(x)}_{\rightarrow 0} \rightarrow f(z)$ .

### Satz 9.25

Sind  $\mathbf{f}, \mathbf{g} : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m$  in  $\mathbf{z} \in D$  differenzierbar, so sind auch  $\mathbf{f} + \mathbf{g}$  und  $\lambda \mathbf{f}$  ( $\lambda \in \mathbb{R}$ ) in  $\mathbf{z}$  differenzierbar, und es gilt

$$\begin{aligned}(\mathbf{f} + \mathbf{g})'(\mathbf{z}) &= \mathbf{f}'(\mathbf{z}) + \mathbf{g}'(\mathbf{z}), \\ (\lambda \mathbf{f})'(\mathbf{z}) &= \lambda \mathbf{f}'(\mathbf{z}).\end{aligned}$$

Man kann also die Jacobi-Matrizen elementweise addieren bzw. mit einem Skalar  $\lambda$  multiplizieren.

### Satz 9.26

Seien  $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^p$  in  $z \in D$  sowie  $g : \mathbb{R}^p \supset f(D) \rightarrow \mathbb{R}^m$  in  $f(z)$  differenzierbar. Dann ist  $h = g \circ f$  in  $z$  differenzierbar, und es gilt

$$h'(z) = (g \circ f)'(z) = g'(f(z)) f'(z). \quad (9.15)$$

Die Struktur von (9.15) ist die gleiche wie im Fall reeller Funktionen, allerdings sind die Ableitungen hier Matrizen.

Man sollte daher immer die Raumdimensionen und Matrizenformate im Auge behalten. Die Ableitungen müssen zwischen den gleichen Räumen vermitteln wie die Funktionen selbst:

$$\begin{array}{ccccc} g(f(z)) & \xleftarrow{g} & f(z) & \xleftarrow{f} & z \\ \mathbb{R}^m & g'(f(z)) \in \mathbb{R}^{m \times p} & \mathbb{R}^p & f'(z) \in \mathbb{R}^{p \times n} & \mathbb{R}^n \end{array}$$



# Multivariate Differentialrechnung

## Kettenregel

Man berechne für

$$\mathbf{f}(x, y) = \begin{bmatrix} 2x^2 + y \\ 3xy \\ x \sin y \end{bmatrix} \quad \text{und} \quad g(x, y, z) = 7x^2 + z$$

die Ableitung der Verkettung  $h = g \circ \mathbf{f}$ .

Berechnen Sie die Ableitung von

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

längs des Kreises um  $\mathbf{0}$  mit Radius 2, d.h. für

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} 2 \cos t \\ 2 \sin t \end{bmatrix} \quad (0 \leq t < 2\pi).$$

# Multivariate Differentialrechnung

## Anmerkungen

Häufig ist es einfacher, die Verkettung  $h = g \circ f$  durch Einsetzen explizit zu bestimmen und dann  $h'(z)$  direkt auszurechnen.

Machen Sie sich das am zweiten Beispiel von S. 101 klar.

Für den Fall  $f : \mathbb{R} \rightarrow \mathbb{R}^n$ ,  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  (passt auf eben genanntes Beispiel) findet man in Tafelwerken auch häufig die Formel

$$(g \circ f)'(t) = \sum_{j=1}^n \frac{\partial g}{\partial x_j}(f(t)) f'_j(t). \quad (9.16)$$

Bestätigen Sie (9.16) mit Hilfe von Satz 9.26.

Zwei reelle Variablen  $x$  und  $y$  seien durch eine Gleichung der Form

$$F(x, y) = 0 \tag{9.17}$$

verknüpft. Kann man diese Gleichung, wenigstens in der Nähe eines Punktes  $(x_0, y_0)$ , der (9.17) selbst erfüllt, eindeutig nach  $y$  „auflösen“ ?

Gesucht ist also eine reelle Funktion  $f$  mit

$$F(x, y) = 0 \quad \Leftrightarrow \quad y = f(x) \quad \text{für } (x, y) \text{ nahe bei } (x_0, y_0).$$

Eine solche Funktion  $f$  nennt man **implizite Funktion**.

Manchmal ist man dabei noch nicht einmal an  $f(x)$  selbst interessiert, sondern eher an der Ableitung  $f'(x)$ .

# Multivariate Differentialrechnung

## Beispiel

Die Gleichung

$$x^2 + y^2 - 1 = 0 \quad (9.18)$$

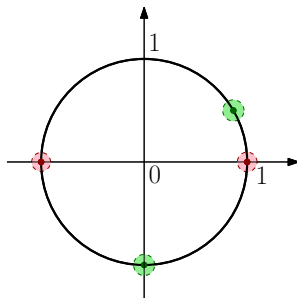
beschreibt einen Kreis um  $(0, 0)$  mit Radius 1.

Nahe der Punkte  $\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$  und  $(0, -1)$  (grün) kann man (9.18) eindeutig nach  $y$  auflösen:

$$y = f_1(x) = \sqrt{1 - x^2} \quad \text{bzw.}$$

$$y = f_2(x) = -\sqrt{1 - x^2}.$$

Nahe der Punkte  $(-1, 0)$  und  $(1, 0)$  (rot) gelingt eine eindeutige Auflösung dagegen nicht, da man nicht weiß, für welchen Funktionswert man sich entscheiden soll.



### Satz 9.27

Die Funktion  $F : \mathbb{R}^2 \supset D \rightarrow \mathbb{R}$  sei auf der offenen Menge  $D$  stetig partiell differenzierbar. Der Punkt  $(x_0, y_0) \in D$  erfülle die Gleichungen

$$F(x_0, y_0) = 0 \quad \text{und} \quad \frac{\partial F}{\partial y}(x_0, y_0) \neq 0.$$

Dann gibt es offene Umgebungen  $U, V \subset \mathbb{R}$  von  $x_0$  und  $y_0$  und eine Funktion  $f : U \rightarrow \mathbb{R}$  mit folgenden Eigenschaften:

- $F(x, y) = 0 \Leftrightarrow y = f(x)$  für  $x \in U, y \in V$ ,
- $f$  ist in  $x_0$  stetig partiell differenzierbar mit

$$f'(x_0) = -\frac{\frac{\partial F}{\partial x}(x_0, y_0)}{\frac{\partial F}{\partial y}(x_0, y_0)}. \quad (9.19)$$

# Multivariate Differentialrechnung

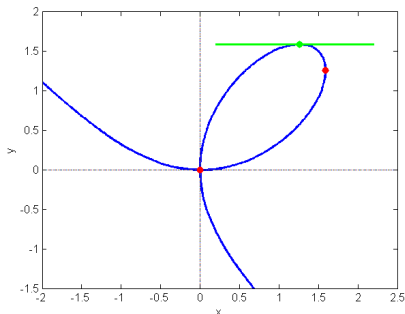
## Der Satz über implizite Funktionen

Wir betrachten das **kartesische Blatt**\*, gegeben durch die Gleichung

$$F(x, y) = x^3 + y^3 - 3xy = 0. \quad (9.20)$$

Bestimmen Sie alle Punkte des kartesischen Blatts, für die keine eindeutige lokale Auflösung nach  $y$  möglich ist.

Bestimmen Sie desweiteren alle Punkte mit horizontaler Tangente sowie den Anstieg der Tangente im Punkt  $(1, 2 \cos \frac{2\pi}{9})$ .



\*benannt nach dem französischen Mathematiker und Philosophen René Descartes.

# Multivariate Differentialrechnung

## Lösungsskizze

Die kritischen Punkte, für die keine Auflösung nach  $y$  möglich ist, ergeben sich aus der Bedingung

$$\frac{\partial F}{\partial y}(x, y) = 3y^2 - 3x \stackrel{!}{=} 0 \quad \Leftrightarrow \quad x = y^2,$$

die man in Gleichung (9.20) einsetzt:

$$\begin{aligned} 0 &\stackrel{!}{=} x^3 + y^3 - 3xy \\ &= y^6 + y^3 - 3y^3 \\ &= y^3(y^3 - 2) \quad \Rightarrow \quad y = 0 \quad \text{oder} \quad y = \sqrt[3]{2}. \end{aligned}$$

Unter Berücksichtigung von  $x = y^2$  erhält man also als kritische Punkte  $(0, 0)$  und  $(\sqrt[3]{4}, \sqrt[3]{2})$ .

Dies sind gerade die „Kreuzung“ im Ursprung sowie der Punkt mit vertikaler Tangente (beide rot markiert).

In der Nähe aller anderen Punkte der Kurve lässt sich (9.20) in die Form  $y = f(x)$  bringen. Die konkrete Angabe von  $f$  ist aber zumindest mühsam, da (9.20) eine Gleichung dritten Grades ist.

Dies müssen wir zur Berechnung der Ableitung aber auch gar nicht leisten, denn Satz 9.27 liefert uns die Ableitung unmittelbar:

$$f'(x) = -\frac{F_x(x, y)}{F_y(x, y)} = -\frac{3x^2 - 3y}{3y^2 - 3x} = \frac{y - x^2}{y^2 - x} \quad (x \neq y^2).$$

Im Punkt  $(1, 2 \cos \frac{2\pi}{9})$  findet man als Tangentenanstieg

$$f'(1) = \frac{2 \cos(\frac{2\pi}{9}) - 1}{4 \cos^2(\frac{2\pi}{9}) - 1} = 0.39493084 \dots$$



Für die Punkte mit horizontaler Tangente muss neben (9.20) auch  $f'(x) = 0$  und somit  $y = x^2$  gelten.

Nach Einsetzen in (9.20) erhält man damit  $(x, y) = (0, 0)$  (rot) und  $(x, y) = (\sqrt[3]{2}, \sqrt[3]{4})$  (grün) als Kandidaten.

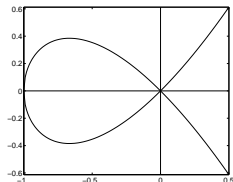
Der erste Punkt muss jedoch ausgesondert werden – es handelt sich um einen der kritischen Punkte, für die keine Auflösung nach  $y$  möglich ist.

**Warnung:** Selbst die Interpretation, es gebe in  $(0, 0)$  eine horizontale und eine vertikale Tangente ist unzulässig. Beim kartesischen Blatt ist dies zwar richtig, es wäre aber noch zu zeigen.

### Illustrierendes Beispiel:

$$F(x, y) = y^2 - x^3 - x^2 = 0 \quad (x \geq -1)$$

$$F_x(0, 0) = F_y(0, 0) = 0$$



# Multivariate Differentialrechnung

## Implizite Funktionen

Analysieren Sie das Kreisbeispiel (S. 104) im Kontext von Satz 9.27.

Gehen Sie dabei auf lokale Auflösbarkeit ein, und berechnen Sie den Anstieg der Tangenten an den Kreis im Punkt  $\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$  sowohl durch Ableiten von  $f_1(x) = \sqrt{1-x^2}$  als auch mit Hilfe von Satz 9.27.

Bestätigen Sie Formel (9.19) durch beidseitiges Ableiten der Gleichung  $F(x, y) = 0$  (mit  $y = f(x)$ ) unter Anwendung der Kettenregel.

## 9 Differentialrechnung in mehreren Variablen

9.1 Vektorfolgen und ihre Grenzwerte

9.2 Grenzwerte von Funktionen und Stetigkeit

9.3 Darstellungsfragen, Anwendungen und Systematisierungsversuch zu Funktionen mehrerer Variablen

9.4 Differenzierbarkeit bei mehreren Variablen

9.5 Differentiation vektorwertiger Funktionen

9.6 Extrema von Funktionen mehrerer Variablen

# Multivariate Differentialrechnung

## Extrema ohne Nebenbedingungen

Wir werden in diesem Abschnitt nach Punkten  $z \in \mathbb{R}^n$  suchen, in denen eine reellwertige differenzierbare Funktion mehrerer Variablen ein lokales Extremum annimmt.

Wir erinnern uns an das Vorgehen im Eindimensionalen, d. h. für  $f : \mathbb{R} \supset (a, b) \rightarrow \mathbb{R}$ :

- Finde alle Punkte  $z \in \mathbb{R}$ , für die  $f'(z) = 0$  gilt (notwendige Bedingung).
- Falls in einem solchen Punkt zusätzlich  $f''(z) > 0$  bzw.  $f''(z) < 0$  gilt, so handelt es sich um ein lokales Minimum bzw. Maximum (hinreichende Bedingung).

Wir versuchen, diese Strategie zu verallgemeinern.

### Definition 9.28

Sei  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$ . Ein Punkt  $z \in D_f$  heißt

- **globales Minimum** von  $f$ , wenn  $f(z) \leq f(x)$  für alle  $x \in D_f$ ,
- **globales Maximum** von  $f$ , wenn  $f(z) \geq f(x)$  für alle  $x \in D_f$ ,
- **lokales Minimum** von  $f$ , wenn  $f(z) \leq f(x)$  für alle  $x \in D_f \cap K_\epsilon(z)$  mit einem  $\epsilon > 0$ ,
- **lokales Maximum** von  $f$ , wenn  $f(z) \geq f(x)$  für alle  $x \in D_f \cap K_\epsilon(z)$  mit einem  $\epsilon > 0$ .

Mit  $K_\epsilon(z)$  ist dabei wieder eine Kugel um  $z$  mit Radius  $\epsilon$  gemeint.

Bei lokalen Extrema vergleicht man  $f(z)$  also nur mit Funktionswerten zu Argumenten, die nahe an  $z$  liegen.

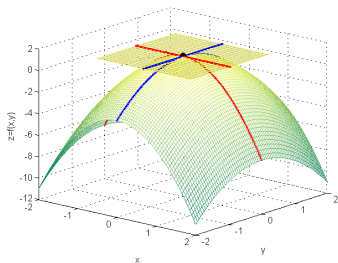
# Multivariate Differentialrechnung

## Notwendige Bedingung für lokale Extrema

Besitzt  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  in  $\mathbf{z}$  ein lokales Maximum (für Minima analog), dann besitzt auch die partielle Funktion

$$f_j : \mathbf{x} \mapsto f(z_1, \dots, z_{j-1}, \mathbf{x}, z_{j+1}, \dots, z_n)$$

in  $z_j$  ein Maximum. Diese ist eine Funktion einer Variablen, daher gilt  $f'_j(z_j) = \frac{\partial f}{\partial x_j}(\mathbf{z}) = 0$ . Das Argument greift dabei für jede Wahl von  $j$ .



Zusammengefasst gilt also  $\nabla f(\mathbf{z}) = \mathbf{0}$ , wie folgender Satz formuliert.

### Satz 9.29 (und Definition)

Sei  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  partiell differenzierbar, und  $D_f$  offen. Besitzt  $f$  in  $z \in D_f$  ein lokales Minimum oder Maximum, so gilt

$$\nabla f(z) = \mathbf{0}. \quad (9.21)$$

Ein Punkt  $z$ , der (9.21) erfüllt, heißt **stationärer Punkt** von  $f$ .

Genausogut könnte man statt (9.21) natürlich auch  $f'(z) = \mathbf{0}^T$  schreiben.

Gleichung (9.21) stellt eine notwendige Bedingung für lokale Extrema dar. Mit ihr kann man Kandidaten für Extrema finden; der Nachweis für das Vorliegen eines Extremums ist jedoch zusätzlich zu erbringen.

# Multivariate Differentialrechnung

## Geometrische Interpretation

In einem lokalen Extremum verläuft die Tangential(hyper)ebene horizontal. Dies kann man an typischen Mittelgebirgsgipfeln wie dem Brocken (1142 m) gut beobachten:



Mittelgebirgsgipfel sind in unmittelbarer Gipfelumgebung flach.



# Multivariate Differentialrechnung

## Extrema ohne Nebenbedingungen

Bestimmen Sie sämtliche stationären Punkte von

$$f(x, y) = 1 - 2x^2 - y^2 \quad \text{und} \quad g(x, y) = x^3 + y^3 - 9xy + 25.$$

Der Gipfel des Siniolchu in Sikkim (6887 m) lässt scheinbar auch „Tangentialebenen“ zu, die nicht horizontal verlaufen (Bild links).



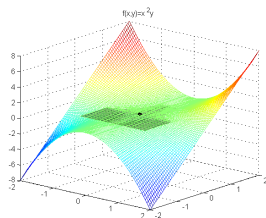
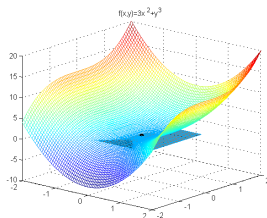
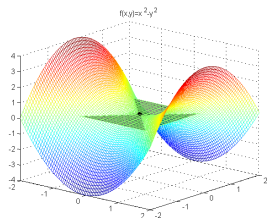
Muss für Indien eine separate Version von Satz 9.29 erstellt werden?  
(Bild rechts: Charles Fanning, 1920)

# Multivariate Differentialrechnung

## Hesse-Matrix

Wie im Eindimensionalen reicht die notwendige Bedingung (9.21) für einen Nachweis von Extrema nicht aus.

Eine horizontale Tangentialebene ( $\nabla f(z) = \mathbf{0}$ ) könnte zum Beispiel folgendermaßen entstehen:



Wir wollen daher die bekannte reelle Bedingung  $f''(z) \neq 0$  verallgemeinern. Dazu formulieren wir zunächst ein Analogon zur zweiten Ableitung.

### Definition 9.30

Sei  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  in  $z \in D_f$  zweimal partiell differenzierbar. Dann heißt

$$H_f(z) := \begin{bmatrix} \frac{\partial^2 f(z)}{\partial x_1^2} & \frac{\partial^2 f(z)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(z)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(z)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(z)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(z)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(z)}{\partial x_n \partial x_1} & \frac{\partial^2 f(z)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(z)}{\partial x_n^2} \end{bmatrix}$$

**Hesse-Matrix** von  $f$  an der Stelle  $z$ .

Aus dem Satz von Schwarz folgt sofort:

### Satz 9.31

Ist  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  in  $z \in D_f$  zweimal **stetig** partiell differenzierbar, dann ist die Hesse-Matrix  $H_f(z)$  symmetrisch.

Als nächstes benötigen wir eine mehrdimensionale Entsprechung zu den Relationen „ $<$ “ und „ $>$ “ in den eindimensionalen Beziehungen  $f''(z) < 0$  bzw.  $f''(z) > 0$ .

### Definition 9.32

Eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  heißt

- **positiv definit**, wenn ihre Eigenwerte sämtlich positiv sind,
- **negativ definit**, wenn ihre Eigenwerte sämtlich negativ sind,
- **indefinit**, wenn sie sowohl positive als auch negative Eigenwerte besitzt.

Im symmetrischen  $2 \times 2$ -Fall ist die Untersuchung auf Definitheit besonders leicht:

### Satz 9.33

Eine symmetrische Matrix

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{1,2} & a_{2,2} \end{bmatrix}$$

ist genau dann

- positiv definit, wenn  $\det A > 0$  und  $a_{1,1} > 0$ ,
- negativ definit, wenn  $\det A > 0$  und  $a_{1,1} < 0$ ,
- indefinit, wenn  $\det A < 0$ .

### Satz 9.34

Seien  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  zweimal stetig partiell differenzierbar und  $z \in D_f$  ein stationärer Punkt von  $f$  (d. h.  $\nabla f(z) = \mathbf{0}$ ).

Dann gilt:

- Ist  $H_f(z)$  positiv definit, dann besitzt  $f$  in  $z$  ein lokales Minimum.
- Ist  $H_f(z)$  negativ definit, dann besitzt  $f$  in  $z$  ein lokales Maximum.
- Ist  $H_f(z)$  indefinit, dann ist  $z$  keine Extremalstelle von  $f$ .

Stationäre Punkte, die keine Extrempunkte sind (betrifft dritten Punkt), werden auch **Sattelpunkte** genannt.

# Multivariate Differentialrechnung

## Rezept zur Bestimmung lokaler Extrema

Für zweimal stetig partiell differenzierbare Funktionen  $f : \mathbb{R}^n \supset D_f \rightarrow \mathbb{R}$  mit offenem Definitionsbereich  $D_f$  gehe wie folgt vor:

- Bestimme alle stationären Punkte von  $f$ , d. h. die Lösungen des (i. Allg. nichtlinearen) Gleichungssystems

$$\nabla f(\mathbf{x}) = \mathbf{0}.$$

- Berechne für sämtliche dieser stationären Punkte die Hesse-Matrix, und analysiere anhand der Eigenwerte oder mit Satz 9.33 deren Definitheitseigenschaften.
- Entscheide mit Satz 9.34, ob Extrema vorliegen und welcher Art diese gegebenenfalls sind.

# Multivariate Differentialrechnung

## Bestimmung lokaler Extrema

Welche Definitheitseigenschaften besitzen folgende symmetrische Matrizen?

$$A = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 42 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -3 \\ -3 & 7 \end{bmatrix}.$$

Bestimmen Sie alle lokalen Extrema von

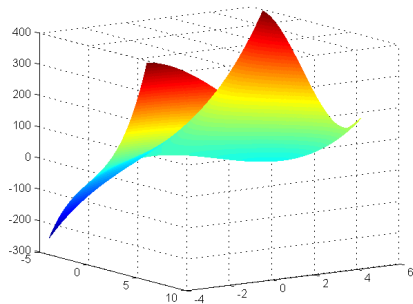
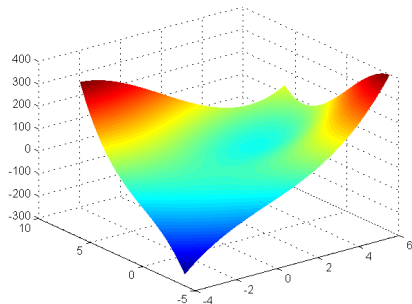
$$f(x, y) = 1 - 2x^2 - y^2 \quad \text{und} \quad g(x, y) = x^3 + y^3 - 9xy + 25.$$

Nutzen Sie dabei ihre Vorarbeiten aus der Aufgabe von S. 117.



# Multivariate Differentialrechnung

## Bestimmung lokaler Extrema



Plot der Funktion  $g(x, y) = x^3 + y^3 - 9xy + 25$  (zweites Beispiel S.124) aus verschiedenen Winkeln.

# Multivariate Differentialrechnung

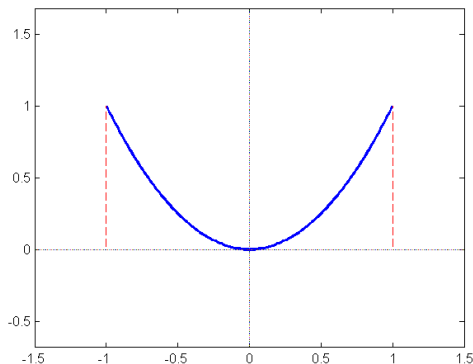
## Extrema am Rand des Definitionsbereichs

Ist der Definitionsbereich  $D_f$  nicht offen, so muss man ggf. dessen Randpunkte gesondert untersuchen, denn sie werden von Satz 9.29 nicht erfasst.

Ein einfaches Beispiel ist die reelle Funktion

$$f : [-1, 1] \rightarrow \mathbb{R}, \quad f(x) = x^2.$$

Diese besitzt an den Rändern des Definitionsbereichs (in  $x_{1/2} = \pm 1$ ) lokale Maxima, in denen die (einseitige) Ableitung nicht 0 ist!



**Problemstellung:** Wir suchen Extrema einer gegebenen **Zielfunktion**

$$f(\mathbf{x}) = f(x_1, \dots, x_n),$$

wobei die Variablen  $x_1, \dots, x_n$  zusätzlich  $m$  Nebenbedingungen erfüllen sollen ( $m < n$ ):

$$g_1(\mathbf{x}) = g_1(x_1, \dots, x_n) = 0,$$

$$\vdots$$

$$g_m(\mathbf{x}) = g_m(x_1, \dots, x_n) = 0.$$

In den meisten Fällen betrachten wir hier den Fall  $n = 2$  und  $m = 1$ , d. h. die Aufgabe

$$\begin{aligned} & f(x, y) \stackrel{!}{\rightarrow} \min \text{ oder } \max \\ \text{unter } & g(x, y) = 0. \end{aligned} \tag{9.22}$$

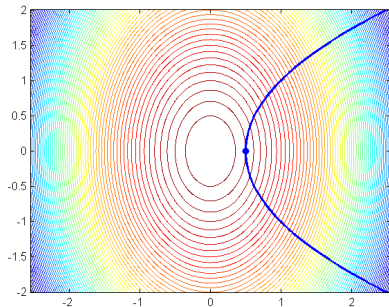
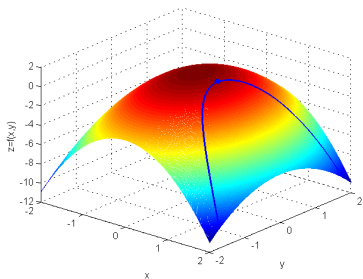
# Multivariate Differentialrechnung

## Beispiel mit grafischer Interpretation

### Die Aufgabe

$$\begin{aligned} f(x, y) &= 1 - 2x^2 - y^2 \xrightarrow{!} \max \\ \text{unter } g(x, y) &= y^2 - 2x + 1 = 0 \end{aligned} \quad (9.23)$$

bedeutet geometrisch: Suche den höchsten Punkt, der beim Durchlaufen des blau markierten Pfades auf der Fläche im linken Bild erreicht wird (rechts die Darstellung in der Höhenkarte).



# Multivariate Differentialrechnung

## Eliminationsmethode

Bei **Eliminationsmethode** löst man die Nebenbedingung(en) nach einer (bzw.  $m$ ) Variablen auf und setzt sie in die Zielfunktion ein.

Dadurch werden in der Zielfunktion eine (bzw.  $m$ ) Variablen eliminiert.

Die verbleibende Funktion von  $n - m$  Variablen kann man mit den Methoden von Abschnitt 112 behandeln.

Voraussetzung ist natürlich, dass das Auflösen nach einer (bzw.  $m$ ) Variablen ohne viel Aufwand möglich ist.

Ist dies der Fall, so ist die Eliminationsmethode der einfachste und effizienteste Weg, das Problem von S. 127 zu lösen.

# Multivariate Differentialrechnung

## Beispiel

Die Nebenbedingung in (9.23) kann man in  $y^2 = 2x - 1$  umschreiben und in die Zielfunktion  $f(x, y) = 1 - 2x^2 - y^2$  einsetzen.

Die entstehende reelle Funktion

$$\tilde{f} : \left[ \frac{1}{2}, \infty \right) \rightarrow \mathbb{R}, \quad \tilde{f}(x) = -2x^2 - 2x + 2$$

besitzt an der Stelle  $x_0 = \frac{1}{2}$  ein lokales Maximum (Achtung: Randpunkt des Definitionsbereichs). Der zugehörige Wert  $y_0 = 0$  ergibt sich aus der Nebenbedingung.

Führen Sie die Details dieser Argumentation aus.

Bestimmen Sie die Abmessungen eines zylindrischen Körpers mit 1 l Volumen und minimaler Oberfläche.

Leider lässt sich das System der Nebenbedingungen

$$g_j(\mathbf{x}) = 0 \quad (j = 1, \dots, m)$$

häufig nicht nach  $m$  Variablen auflösen. Insbesondere ist das der Regelfall, wenn die Nebenbedingungen nichtlineare Gleichungen sind.

Für solche Fälle bietet sich die **Lagrange-Methode** als Alternative an.

Wir machen uns diesen Ansatz am Fall

$$\begin{aligned} f(x, y) &\stackrel{!}{\rightarrow} \min \text{ oder } \max \\ \text{unter } g(x, y) &= 0. \end{aligned}$$

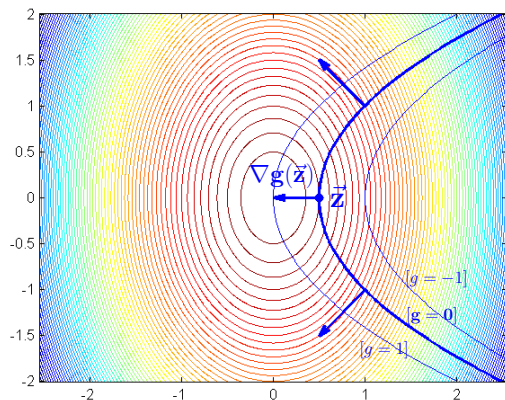
plausibel, wobei wir  $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}$  als stetig partiell differenzierbar voraussetzen.

# Multivariate Differentialrechnung

## Lagrange-Methode

Wir erinnern uns zunächst, dass der Gradient  $\nabla g(x, y)$  senkrecht auf den Höhenlinien von  $g(x, y)$  steht.

Da die Nebenbedingung  $g(x, y) = 0$  aber gerade die Höhenlinie von  $g$  zum Niveau 0 darstellt, steht  $\nabla g(x, y)$  senkrecht auf der durch die Nebenbedingung gegebenen Kurve.

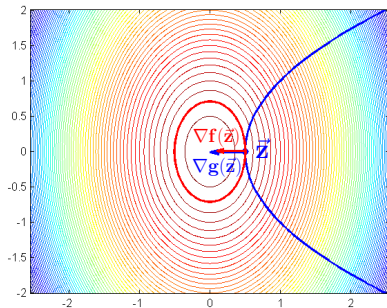




Zumindest im gezeichneten Beispiel erkennen wir:

- Im gesuchten Maximum  $\mathbf{z}$  berühren sich die Höhenlinie  $f(\mathbf{x}) = f(\mathbf{z})$  und die Nebenbedingung  $g(\mathbf{x}) = 0$ ; sie besitzen die gleiche Tangente.
- $\nabla f(\mathbf{z})$  und  $\nabla g(\mathbf{z})$  stehen beide senkrecht auf dieser Tangente und sind daher parallel. Es gilt

$$\nabla f(\mathbf{z}) = -\lambda \nabla g(\mathbf{z}) \quad \text{für ein } \lambda \in \mathbb{R}.$$



Die gesuchten Extrema sind also Lösungen des (i.A. nichtlinearen) Gleichungssystems

$$\begin{aligned}\nabla f(x, y) + \lambda \nabla g(x, y) &= \mathbf{0} \\ g(x, y) &= 0\end{aligned}\tag{9.24}$$

(drei Gleichungen für die drei Variablen  $x, y$  und  $\lambda$ ).

Mit der „Lagrange-Funktion“

$$L(x, y; \lambda) := f(x, y) + \lambda g(x, y)$$

liest sich das System (9.24) kürzer als

$$\nabla L(x, y; \lambda) = \mathbf{0}.\tag{9.25}$$

Machen Sie sich die Äquivalenz von (9.24) und (9.25) klar.

Welche Größe erhalten Sie für die dritte Komponente von  $\nabla L(x, y; \lambda)$ , d. h. für die partielle Ableitung  $L_\lambda(x, y; \lambda)$ ?

Mathematisch präzise ergibt sich die Lagrange-Methode für eine Nebenbedingung aus folgendem Satz:

### Satz 9.35

Die Funktionen  $f, g : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$  seien stetig partiell differenzierbar, und  $\nabla g(\mathbf{x}) \neq \mathbf{0}$  auf  $D$ . Dann gilt:

Ist  $\mathbf{z}$  eine lokale Extremstelle von  $f$  unter der Nebenbedingung  $g(\mathbf{x}) = 0$ , dann gibt es einen **Lagrange-Multiplikator**  $\lambda_0 \in \mathbb{R}$ , so dass

$$\nabla f(\mathbf{z}) + \lambda_0 \nabla g(\mathbf{z}) = \mathbf{0}.$$

Insbesondere ist  $\begin{bmatrix} \mathbf{z} \\ \lambda_0 \end{bmatrix}$  stationärer Punkt der **Lagrange-Funktion**

$$L(\mathbf{x}; \lambda) := f(\mathbf{x}) + \lambda g(\mathbf{x}).$$

Wir kennen nun also eine notwendige Bedingung für lokale Extrema unter Nebenbedingungen.

Wir betrachten wieder das Beispiel von S. 128, d.h.

$$\begin{aligned} f(x, y) &= 1 - 2x^2 - y^2 \xrightarrow{!} \max \\ \text{unter } g(x, y) &= y^2 - 2x + 1 = 0. \end{aligned}$$

Die Lagrange-Funktion des Problems ist

$$L(x, y; \lambda) = 1 - 2x^2 - y^2 + \lambda(y^2 - 2x + 1).$$

Deren stationären Punkte ergeben sich aus den drei Gleichungen

$$\begin{aligned} L_x(x, y, \lambda) &= -4x - 2\lambda = 0 \\ L_y(x, y, \lambda) &= -2y + 2\lambda y = 0 \\ L_\lambda(x, y, \lambda) &= y^2 - 2x + 1 = 0 \end{aligned} \tag{9.26}$$

Man bestimme die Lösungen des Gleichungssystems (9.26), und untersuche deren Extremaleigenschaften anhand einer Höhenlinienkarte.

Für mehrere Nebenbedingungen verwendet man folgende Verallgemeinerung von Satz 9.35:

### Satz 9.36

Die Funktionen  $f, g_i : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$  ( $i = 1, \dots, m$ ;  $m < n$ ) seien stetig partiell differenzierbar, wobei die Gradienten  $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$  für alle  $\mathbf{x} \in D$  linear unabhängig seien. Dann gilt:

Ist  $\mathbf{z}$  eine lokale Extremstelle von  $f$  unter den Nebenbedingungen  $g_i(\mathbf{x}) = 0$ , dann gibt es einen **Lagrange-Multiplikator**  $\boldsymbol{\lambda}^{(0)} \in \mathbb{R}^m$  mit

$$\nabla f(\mathbf{z}) + \sum_{i=1}^m \lambda_i^{(0)} \nabla g_i(\mathbf{x}) = \mathbf{0}.$$

Insbesondere ist  $[\boldsymbol{\lambda}^{(0)}]$  stationärer Punkt der **Lagrange-Funktion**

$$L(\mathbf{x}; \boldsymbol{\lambda}) := f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}).$$

### Rezept zur Lagrange-Methode:

- Ordne jeder Nebenbedingung  $g_i(x_1, \dots, x_n) = 0$  ( $i = 1, \dots, m$ ) einen Lagrange-Multiplikator  $\lambda_i$  zu.
- Stelle die Lagrange-Funktion auf:

$$L(\mathbf{x}; \boldsymbol{\lambda}) = L(x_1, \dots, x_n; \lambda_1, \dots, \lambda_m) := f(\mathbf{x}) + \sum_{j=1}^m \lambda_j g_j(\mathbf{x}).$$

- Berechne die stationären Punkte der Lagrange-Funktion  $L(\mathbf{x}; \boldsymbol{\lambda})$ , d. h. löse die  $(n + m)$  Gleichungen

$$\begin{aligned} L_{x_i}(\mathbf{x}; \boldsymbol{\lambda}) &= 0 & (i = 1, \dots, n), \\ L_{\lambda_j}(\mathbf{x}; \boldsymbol{\lambda}) &= 0 & (j = 1, \dots, m). \end{aligned}$$

- Untersuche die Extremaleigenschaften dieser stationären Punkte (z. B. mittels Höhenlinienkarte oder Kompaktheitsargument).

Sie sollten nun (bzw. nach Abschluss der Übungen/Selbststudium):

- Funktionen  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  anhand der Dimensionen  $m$  und  $n$  typisieren können und zu den verschiedenen Funktionstypen eine gute Vorstellung besitzen,
- die verschiedenen Ableitungsbegriffe (partielle, Richtungs- und totale Ableitungen, Gradient, Jacobi-Matrix) theoretisch gut verstanden haben und sicher berechnen können,
- Gradient und Höhenlinien geometrisch interpretieren können,
- Ableitungen mittels mehrdimensionaler Kettenregel und des Satzes über implizite Funktionen berechnen können,
- lokale Extrema und Sattelpunkte sicher bestimmen können,
- lokale Extrema unter Nebenbedingungen sowohl mit der Eliminationsmethode als auch mit der Lagrange-Methode bei überschaubarer Problemgröße sicher berechnen können.