

Mathematik II für MB, WI/MB und andere
Prof. Dr. Wilhelm Stannat

Inhalt:

1. Folgen und Reihen von Funktionen
2. Kurven im \mathbb{R}^n
3. Funktionen in mehreren Variablen
4. Differentialrechnung von Funktionen mehrerer Veränderlicher
5. Implizite Funktionen und Extrema mit Nebenbedingungen

Das vorliegende Skript ist eine **korrigierte** Zusammenfassung des ersten Teils der Vorlesung Mathematik II für MB, WI/MB und andere, die im SS 2008 an der TU Darmstadt gehalten wurde. Die Lektüre des Skriptes ist kein gleichwertiger Ersatz für den Besuch der Vorlesung.

Korrekturen bitte per Email an stannat@mathematik.tu-darmstadt.de

1 Folgen und Reihen von Funktionen

1.1 Allgemeines

Wir wollen im folgenden einige allgemeine Eigenschaften von Funktionenreihen zusammenfassen. Als Beispiele dienen uns dabei die in Mathematik I eingeführten Potenzreihen und Taylorreihen.

Im ganzen Abschnitt sei $I \subset \mathbb{R}$ ein Intervall.

Eine **Funktionenfolge (auf I)** $(f_n)_{n \geq n_0}$ ist eine Abbildung $n \mapsto f_n$, die jeder ganzen Zahl n ($n \geq n_0$) eine Funktion $f_n : I \rightarrow \mathbb{R}$ zuordnet.

Ist $(f_n)_{n \geq n_0}$ eine Funktionenfolge, so heißt die Folge $(s_n)_{n \geq n_0}$ der Partialsummen

$$s_n(x) = \sum_{k=n_0}^n f_k(x), \quad x \in I$$

eine **Funktionenreihe (auf I)** und wir schreiben hierfür auch $\sum_{k=n_0}^{\infty} f_k$.

Punktweise Konvergenz

Für alle $x \in I$ erhält man aus einer Funktionenfolge $(f_n)_{n \geq n_0}$ eine Folge reeller Zahlen $(f_n(x))_{n \geq n_0}$. Die Funktionenfolge $(f_n)_{n \geq n_0}$ heißt **punktweise konvergent mit Grenzfunktion f** , wenn für alle $x \in I$ die Folge der Funktionswerte $(f_n(x))_{n \geq n_0}$ gegen $f(x)$ konvergiert

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad \forall x \in I$$

Analog für Funktionenreihen: Die Funktionenreihe $\sum_{k=n_0}^{\infty} f_k$ heißt **punktweise konvergent mit Summenfunktion s** , falls die Funktionenfolge der Partialsummen punktweise gegen s konvergiert

$$\lim_{n \rightarrow \infty} \sum_{k=n_0}^n f_k(x) = \lim_{n \rightarrow \infty} s_n(x) = s(x) \quad \forall x \in I$$

Beispiele 1.1

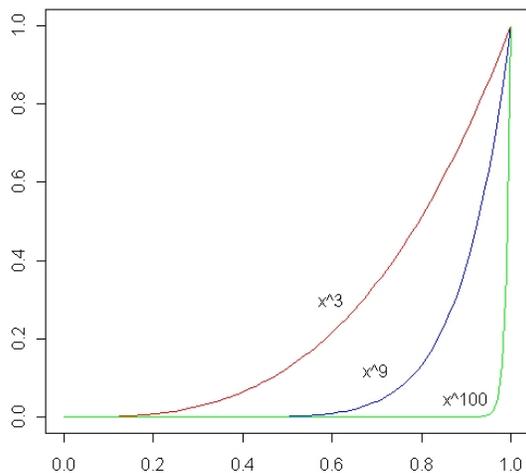
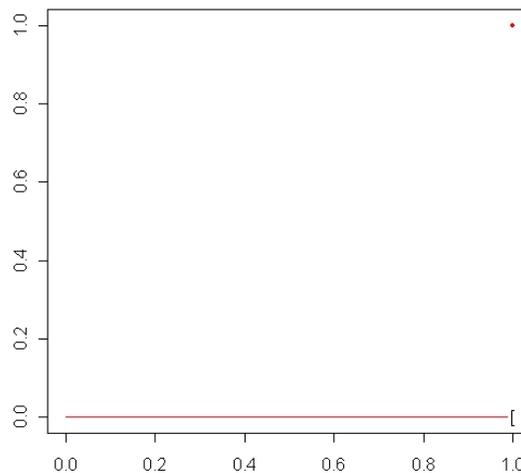
(i) $f_n = x^n$ auf $I = [0, 1]$ ist konvergent gegen $f(x) = \begin{cases} 0 & \text{für } x \in [0, 1[\\ 1 & \text{für } x = 1. \end{cases}$

(ii) $f_n(x) = nxe^{-nx^2}$ auf $I = [0, 1]$ ist konvergent gegen $f(x) = 0$, denn für $x \neq 0$ gilt

$$f_n(x) = \frac{1}{x} \underbrace{(nx^2)e^{-nx^2}}_{\rightarrow 0 \text{ für } n \rightarrow \infty}.$$

(iii) $f_n(x) = \frac{1}{\sqrt{n}} \sin(\pi nx)$ auf $I = [0, 1]$ konvergiert gegen $f(x) = 0$.

Beispiel (i) zeigt: Grenzfunktionen stetiger Funktionen müssen nicht wieder stetig sein

Funktionsfolge x^n 

Grenzfunktion

Beispiel (ii) zeigt: Integration darf im allgemeinen mit dem Funktionsgrenzwert nicht vertauscht werden

$$\int_0^1 f_n(x) dx = \int_0^1 nxe^{-nx^2} dx = -\frac{1}{2}e^{-nx^2} \Big|_0^1 = \frac{1}{2} - \frac{1}{2}e^{-n} \xrightarrow{n \rightarrow \infty} \frac{1}{2}$$

aber $\int_0^1 f(x) dx = 0$.

Beispiel (iii) zeigt: Differentiation darf im allgemeinen mit dem Funktionsgrenzwert nicht vertauscht werden

$$f'_n(x) = \pi\sqrt{n} \cos(\pi nx) \quad \text{also } f'_n(0) = \pi\sqrt{n} \quad \text{aber } f'(x) = 0.$$

Um Differentiation und Integration mit Funktionsgrenzwerten vertauschen zu können, benötigen wir den stärkeren Konvergenzbegriff der **gleichmäßigen Konvergenz**.

Definition Eine Funktionenfolge $(f_n)_{n \geq n_0}$ (auf I) heißt **gleichmäßig konvergent** gegen die Grenzfunktion f , falls es eine Nullfolge $(a_n)_{n \geq n_0}$ gibt mit

$$|f_n(x) - f(x)| \leq a_n \quad \text{für alle } x \in I, n \geq n_0$$

Entsprechend heißt die Funktionenreihe $\sum_{k=n_0}^{\infty} f_k$ gleichmäßig konvergent, falls die Funktionenfolge der Partialsummen gleichmäßig konvergiert.

Damit gilt nun

$$(f_n)_{n \geq n_0} \text{ gleichmäßig konvergent} \Rightarrow (f_n)_{n \geq n_0} \text{ punktweise konvergent}$$

Die Umkehrung gilt aber nicht, wie folgendes Beispiel zeigt

Beispiel (siehe Beispiel 1.1 (i))

$f_n(x) = x^n$ ist punktweise konvergent auf $[0, 1]$ gegen die Grenzfunktion

$$f(x) = \begin{cases} 0 & \text{für } x \in [0, 1[\\ 1 & \text{für } x = 1 \end{cases}$$

aber nicht gleichmäßig, denn

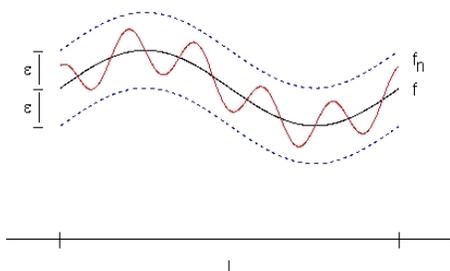
$$|f_n(x) - f(x)| = \begin{cases} x^n & \text{für } x \in [0, 1[\\ 0 & \text{für } x = 1 \end{cases}$$

und speziell für $x_n = 1 - \frac{1}{n}$ gilt

$$|f_n(x_n) - f(x_n)| = \left(1 - \frac{1}{n}\right)^n \xrightarrow{n \rightarrow \infty} e^{-1}$$

Daher kann es keine Nullfolge (a_n) mit $|f_n(x) - f(x)| \leq a_n$ für alle $x \in [0, 1]$, wie in der Definition für gleichmäßige Konvergenz gefordert, geben.

Graphische Veranschaulichung gleichmäßiger Konvergenz



$(f_n)_{n \geq n_0}$ ist genau dann gleichmäßig konvergent gegen f , wenn zu $\varepsilon > 0$ ein N_ε existiert mit der Eigenschaft, dass alle Funktionen f_n für $n \geq N_\varepsilon$ ganz im ε -Schlauch um f liegen.

Eigenschaften gleichmäßig konvergenter Funktionenfolgen

Es sei $(f_n)_{n \geq n_0}$ eine punktweise konvergente Funktionenfolge (auf $I = [a, b]$) mit Grenzfunktion f . Dann gilt

- (i) (f_n) gleichmäßig konvergent, f_n stetig für alle $n \Rightarrow f$ stetig
- (ii) (f_n) gleichmäßig konvergent, f_n integrierbar für alle $n \Rightarrow f$ integrierbar

und
$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx$$

- (iii) (f_n) stetig differenzierbar, (f'_n) gleichmäßig konvergent $\Rightarrow f$ stetig differenzierbar, $f'(x) = \lim_{n \rightarrow \infty} f'_n(x)$ und (f_n) gleichmäßig konvergent.

Entsprechende Aussagen gelten für Funktionenreihen, wenn man obige Aussagen auf die Funktionenfolgen der Partialsummen anwendet. Insbesondere gilt also für eine punktweise konvergente Funktionenreihe $\sum_{k=n_0}^{\infty} f_k$

- $\sum_{k=n_0}^{\infty} f_k$ gleichmäßig konvergent, f_n integrierbar für alle n
 $\Rightarrow \int_a^b \sum_{k=n_0}^{\infty} f_k(x) dx = \sum_{k=n_0}^{\infty} \int_a^b f_k(x) dx$
- (f_n) stetig differenzierbar, $\sum_{k=n_0}^{\infty} f'_k$ gleichmäßig konvergent
 $\Rightarrow \left(\sum_{k=n_0}^{\infty} f_k(x)\right)' = \sum_{k=n_0}^{\infty} f'_k(x)$

Beispiele

- (i) $\sum_{k=0}^{\infty} x^k$ ist gleichmäßig konvergent auf $[0, b]$ für alle $b < 1$ mit Grenzfunktion $\frac{1}{1-x}$, denn

$$s_n(x) = \sum_{k=0}^n x^k = \frac{1-x^{n+1}}{1-x} \quad \text{also}$$

$$\left|s_n(x) - \frac{1}{1-x}\right| = \frac{x^{n+1}}{1-x} \leq \underbrace{\frac{b^{n+1}}{1-b}}_{=: a_n} \xrightarrow{n \rightarrow \infty} 0$$

Insbesondere folgt

$$\sum_{k=0}^{\infty} \int_0^b x^k dx = \sum_{k=0}^{\infty} \frac{1}{k+1} b^{k+1} = \int_0^b \frac{1}{1-x} dx = -\ln(1-b)$$

oder äquivalent

$$\sum_{k=0}^{\infty} (-1)^k \frac{(-b)^{k+1}}{k+1} = \ln(1-b) \quad \text{für } 0 < b < 1$$

Dies entspricht gerade der in Kapitel 10, Mathematik I, aufgeführten Potenzreihendarstellung von $\ln(1+x)$

$$\ln(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1}, \quad |x| < 1$$

Die Funktionenreihe der Ableitungen $\sum_{k=1}^{\infty} kx^{k-1}$ ist ebenfalls gleichmäßig konvergent auf $[0, b]$ für $b < 1$ und damit folgt

$$\sum_{k=1}^{\infty} kx^{k-1} = \frac{d}{dx} \left(\frac{1}{1-x} \right) = \frac{1}{(1-x)^2} \quad \text{für alle } x \in [0, b].$$

Dabei ergibt sich die gleichmäßige Konvergenz der Funktionenreihe $\sum_{k=1}^{\infty} kx^{k-1}$ auf $[0, b]$, $b < 1$, aus den allgemeinen Sätzen zum Konvergenzbereich von Potenzreihen (siehe Kapitel 10, Mathematik I und Ergänzung hierzu am Ende dieses Abschnittes). In diesem Spezialfall kann man die gleichmäßige Konvergenz jedoch auch wie folgt direkt einsehen

$$\sum_{k=1}^n kx^{k-1} = \frac{d}{dx} \left(\sum_{k=0}^n x^k \right) = \frac{d}{dx} \left(\frac{1-x^{n+1}}{1-x} \right) = \frac{1}{(1-x)^2} - \underbrace{(n+1) \frac{x^n}{1-x}}_{\rightarrow 0 \text{ glm. auf } [0, b]}$$

Für Funktionenreihen gilt folgendes **einfaches Kriterium für gleichmäßige Konvergenz**

Gibt es eine Folge $(a_n)_{n \geq n_0}$ mit

$$|f_k(x)| \leq a_k \quad \text{für alle } x \in I, k \geq n_0$$

und ist die Reihe $\sum_{k=n_0}^{\infty} a_k$ konvergent, so ist die Funktionenreihe $\sum_{k=n_0}^{\infty} f_k$ gleichmäßig konvergent (auf I).

Beispiel Die Funktionenreihe $\sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2}$ konvergiert gleichmäßig auf \mathbb{R} , denn

$$\left| \frac{\cos(kx)}{k^2} \right| \leq \frac{1}{k^2}$$

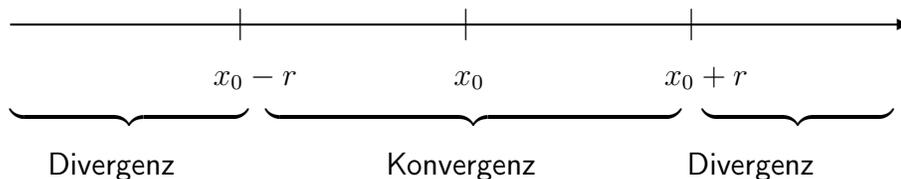
und die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^2}$ ist konvergent.

Potenzreihen

Als Beispiele für Funktionenreihen hatten wir bereits in Kapitel 10, Mathematik I, Potenzreihen

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k \quad (1.1)$$

kennengelernt. Der Konvergenzbereich von (1.1) hat die Form



für ein $r \geq 0$ (auch $r = +\infty$). Für den Konvergenzradius r gilt dabei die Formel

$$r = \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}} \quad (1.2)$$

Hinweis Ist $(\sqrt[n]{|a_n|})$ keine konvergente Folge, so muss man in der Formel (1.2) " $\lim_{n \rightarrow \infty}$ " durch " $\limsup_{n \rightarrow \infty}$ " (" n limes superior" ersetzen).

Einschub: limes superior/limes inferior

Es sei $(b_n)_{n \geq n_0}$ eine Folge reeller Zahlen. Unter dem $\limsup_{n \rightarrow \infty} b_n$ der Folge $(b_n)_{n \geq n_0}$ versteht man den **größten** Wert b (auch $b = +\infty$) für den es eine **Teilfolge**

$$b_{n_1}, b_{n_2}, b_{n_3}, \dots (n_0 \leq n_1 < n_2 < n_3 < \dots)$$

der ursprünglichen Folge $(b_n)_{n \geq n_0}$ gibt, die gegen b konvergiert

$$\lim_{k \rightarrow \infty} b_{n_k} = b.$$

Entsprechend definiert man $\liminf_{n \rightarrow \infty} b_n$ als den **kleinsten** Wert b (auch $b = -\infty$) gegen den eine **Teilfolge** $(b_{n_k})_{k \geq 1}$ der ursprünglichen Folge $(b_n)_{n \geq n_0}$ konvergiert.

Beispiel Für $b_n = (-1)^n$, $n \geq 0$, gilt

$$\limsup_{n \rightarrow \infty} b_n = +1, \quad \liminf_{n \rightarrow \infty} b_n = -1.$$

Man bezeichnet $\limsup_{n \rightarrow \infty} b_n$ auch als **größten Häufungspunkt** der Folge $(b_n)_{n \geq n_0}$, denn in jeder ε -Umgebung von $\limsup_{n \rightarrow \infty} b_n$ liegen unendlich viele Folgenglieder der Folge, aber oberhalb dieser Umgebung liegen jeweils nur endlich viele Folgenglieder. Der $\limsup_{n \rightarrow \infty} b_n$ ist also der größte Wert, um den sich die Folgenglieder der Folge $(b_n)_{n \geq n_0}$ häufen. Entsprechend bezeichnet man $\liminf_{n \rightarrow \infty} b_n$ auch als **kleinsten Häufungspunkt** der Folge $(b_n)_{n \geq n_0}$.

Zurück zu Potenzreihen: Alternativ gilt für den Konvergenzradius die einfachere Formel

$$r = \lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right|$$

falls der Grenzwert existiert.

(i) $\sum_{k=0}^{\infty} \frac{1}{k!} x^k$ hat Konvergenzradius $r = \infty$, denn

$$\left| \frac{a_k}{a_{k+1}} \right| = \left| \frac{(k+1)!}{k!} \right| = k+1 \rightarrow \infty.$$

(ii) $\sum_{k=0}^{\infty} k 2^k (x-2)^k$ hat Konvergenzradius $r = \frac{1}{2}$, denn

$$\left| \frac{a_k}{a_{k+1}} \right| = \left| \frac{k 2^k}{(k+1) 2^{k+1}} \right| = \frac{k}{k+1} \cdot \frac{1}{2} \rightarrow \frac{1}{2}.$$

(iii) Die Potenzreihe $\sum_{k=1}^{\infty} a_k x^k$ mit

$$a_k = \begin{cases} 1 & \text{für } k \text{ gerade} \\ \frac{1}{k} & \text{für } k \text{ ungerade} \end{cases}$$

hat Konvergenzradius 1, aber

$$\lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right| \text{ existiert nicht.}$$

Innerhalb ihres Konvergenzbereiches, genauer, auf allen abgeschlossenen Teilintervallen der Form $[x_0 - r + \varepsilon, x_0 + r - \varepsilon] \subset]x_0 - r, x_0 + r[$ für alle $\varepsilon > 0$, ist die Potenzreihe (1.1) **gleichmäßig konvergent**. Die durch die Potenzreihe (1.1) dargestellte Funktion

$$f :]x_0 - r, x_0 + r[\rightarrow \mathbb{R}, x \mapsto \sum_{k=0}^{\infty} a_k (x - x_0)^k$$

ist also stetig differenzierbar mit Ableitung

$$f'(x) = \sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1}$$

und integrierbar mit Stammfunktion

$$F(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (x - x_0)^{k+1}.$$

Anwendungsbeispiel Die Funktion $f(x) = e^{-x^2}$ ist in der Fehlerrechnung von großer Bedeutung, besitzt aber keine elementare Stammfunktion. Daher sind die Integrale

$$\Phi(x) = \int_0^x e^{-t^2} dt$$

nicht elementar zu berechnen. f besitzt jedoch die Potenzreihendarstellung

$$e^{-x^2} \stackrel{y=-x^2}{=} e^y = \sum_{k=0}^{\infty} \frac{y^k}{k!} = \sum_{k=0}^{\infty} \underbrace{\frac{(-x^2)^k}{k!}}_{=(-1)^k \frac{x^{2k}}{k!}} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{k!}.$$

Für die Stammfunktion $\Phi(x)$ ergibt sich die Reihendarstellung

$$\Phi(x) = \int_0^x \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{k!} dt = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)k!}.$$

Also

$$\Phi(x) \sim \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)k!}$$

und für den Fehler gilt die Abschätzung

$$\left| \Phi(x) - \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)k!} \right| \leq \frac{|x|^{2n+1}}{2n+1} e^{x^2}.$$

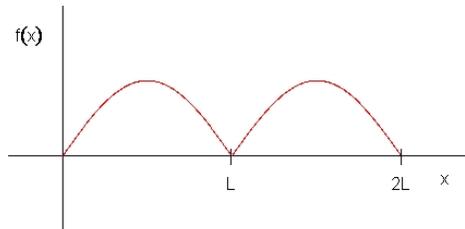
1.2 Fourierreihen

Eine besonders wichtige Klasse von Funktionenreihen bilden die Fourierreihen. Sie eignen sich insbesondere zur Behandlung periodischer Probleme.

Periodische Funktionen

Es sei $L > 0$. Eine Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ heißt **periodisch mit Periode L** , falls

$$f(x + L) = f(x) \quad \forall x \in \mathbb{R}$$



Bemerkung

(i) Die trigonometrischen Funktionen sind periodisch:

- $\sin(x), \cos(x)$ haben Periode 2π
- $\sin(nx), \cos(nx)$ haben ebenfalls Periode 2π für alle $n \geq 1$

(ii) Ist f periodisch mit Periode L , so ist

$$\tilde{f}(x) = f\left(\frac{L}{2\pi}x\right) \quad x \in \mathbb{R}$$

periodisch mit Periode 2π , denn

$$\tilde{f}(x + 2\pi) = f\left(\frac{L}{2\pi}(x + 2\pi)\right) = f\left(\frac{L}{2\pi}x + L\right) = f\left(\frac{L}{2\pi}x\right) = \tilde{f}(x).$$

Bei der Diskussion periodischer Funktionen kann man sich also auf 2π -periodische Funktionen beschränken. Zur Approximation 2π -periodischer Funktionen betrachtet man statt Polynomen oder Potenzreihen geeigneterweise **trigonometrische Polynome bzw. Reihen**

$$t(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \quad (1.3)$$

für Koeffizienten $a_n, b_n \in \mathbb{R}$.

Eigenschaften

(i) Eine trigonometrische Reihe der Form (1.3) ist gleichmäßig konvergent auf \mathbb{R} , falls die Reihe

$$\sum_{n=1}^{\infty} (|a_n| + |b_n|)$$

konvergiert. Hinreichend für die Konvergenz dieser Reihe ist

$$\lim_{n \rightarrow \infty} n^\alpha a_n = 0, \quad \lim_{n \rightarrow \infty} n^\beta b_n = 0 \quad \text{für } \alpha, \beta > 1.$$

- (ii) Eine trigonometrische Reihe der Form (1.3) ist k -mal stetig differenzierbar, falls die Reihe

$$\sum_{n=1}^{\infty} n^k (|a_n| + |b_n|)$$

konvergiert. Hinreichend für die Konvergenz dieser Reihe ist

$$\lim_{n \rightarrow \infty} n^\alpha a_n = 0, \quad \lim_{n \rightarrow \infty} n^\beta b_n = 0 \quad \text{für } \alpha, \beta > k + 1.$$

In diesem Falle lässt sich die Ableitung von t durch gliedweise Differentiation bestimmen:

$$t'(x) = \sum_{n=1}^{\infty} (nb_n \cos(nx) - na_n \sin(nx))$$

und dies ist wieder eine trigonometrische Reihe.

Es sei f eine 2π -periodische Funktion. Gibt es dann eine trigonometrische Reihe

$$t(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

mit $t(x) = f(x)$ für alle x , so heißt t **Fourierreihe zu f** und man sagt, dass f **in eine Fourierreihe entwickelbar** ist.

Welche Beziehung besteht zwischen f und den Koeffizienten a_n, b_n ? Dazu halten wir zunächst folgende wichtige Eigenschaft trigonometrischer Funktionen fest.

Orthogonalitätsrelationen

Für ganze Zahlen $n, m \geq 0$ gilt

$$\int_0^{2\pi} \cos(mx) \sin(nx) dx = 0$$

$$\int_0^{2\pi} \cos(mx) \cos(nx) dx = \begin{cases} 0 & \text{für } m \neq n \\ \pi & \text{für } m = n \geq 1 \\ 2\pi & \text{für } m = n = 0 \end{cases}$$

$$\int_0^{2\pi} \sin(mx) \sin(nx) dx = \begin{cases} 0 & \text{für } m \neq n \\ \pi & \text{für } m = n \geq 1 \\ 0 & \text{für } m = n = 0 \end{cases}$$

Insbesondere gilt also

$$\int_0^{2\pi} \cos(mx) dx = 0, \quad \int_0^{2\pi} \sin(mx) dx = 0 \quad \text{für alle } m \geq 1.$$

Ist f in eine Fourier-Reihe entwickelbar, also

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

und ist die trigonometrische Reihe gliedweise integrierbar (etwa falls gleichmäßig konvergent), so folgt

$$\begin{aligned} \int_0^{2\pi} f(x) \cos(nx) dx &= \int_0^{2\pi} \left(\frac{a_0}{2} + \sum_{m=1}^{\infty} (a_m \cos(mx) + b_m \sin(mx)) \right) \cos(nx) dx \\ &= \frac{a_0}{2} \int_0^{2\pi} \cos(nx) dx + \sum_{m=1}^{\infty} a_m \int_0^{2\pi} \cos(mx) \cos(nx) dx + b_m \int_0^{2\pi} \sin(mx) \cos(nx) dx \\ &= a_n \pi. \end{aligned}$$

Analog zeigt man

$$\int_0^{2\pi} f(x) \sin(nx) dx = b_n \pi \quad \text{also}$$

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx \quad \text{für } n = 0, 1, 2, \dots \\ b_n &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx \quad \text{für } n = 1, 2, \dots \end{aligned}$$

Die Koeffizienten der Fourierreihe zu f sind also eindeutig bestimmt.

Definition Es sei f integrierbar auf $[0, 2\pi]$. Die Zahlen

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx, \quad b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx$$

heißen **Fourierkoeffizienten von f** . Die mit den Fourierkoeffizienten gebildete Reihe

$$F_f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

heißt **Fourierreihe** von f .

Bemerkungen

- (i) Zur Berechnung der Koeffizienten kann statt über $[0, 2\pi]$ über ein beliebiges anderes Intervall der Länge 2π integriert werden, z.B. $[-\pi, \pi]$.
- (ii) Für den Koeffizienten a_0 gilt speziell

$$\frac{a_0}{2} = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx$$

Daher heißt $\frac{a_0}{2}$ **Mittelwert der Funktion f** .

- (iii) Ist f **gerade**, also $f(-x) = f(x)$ für $x \in \mathbb{R}$, so folgt

$$\begin{aligned} a_n &= \frac{2}{\pi} \int_0^{\pi} f(x) \cos(nx) dx \quad \text{für } n \geq 0 \\ b_n &= 0 \quad \text{für } n \geq 1 \end{aligned}$$

Denn es gilt

$$a_n \pi = \int_0^{2\pi} f(x) \cos(nx) dx = \int_{-\pi}^{\pi} f(x) \cos(nx) dx = 2 \underbrace{\int_0^{\pi} f(x) \cos(nx) dx}_{f \text{ und } \cos(nx) \text{ gerade}}$$

$$\begin{aligned} b_n \pi &= \int_0^{2\pi} f(x) \sin(nx) dx = \int_{-\pi}^{\pi} f(x) \sin(nx) dx \\ &= \underbrace{\int_{-\pi}^0 f(x) \sin(nx) dx}_{=-\int_0^{\pi} f(x) \sin(nx) dx} + \int_0^{\pi} f(x) \sin(nx) dx = 0 \end{aligned}$$

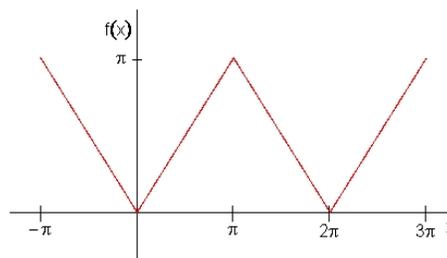
Analog gilt: Ist f **ungerade**, also $f(-x) = -f(x)$ für $x \in \mathbb{R}$, so folgt

$$\begin{aligned} a_n &= 0 \quad \text{für } n \geq 0 \\ b_n &= \frac{2}{\pi} \int_0^{\pi} f(x) \sin(nx) dx \quad \text{für } n \geq 1 \end{aligned}$$

Beispiele 1.1

(i) f sei 2π -periodisch mit

$$f(x) = \begin{cases} x & \text{für } 0 \leq x \leq \pi \\ 2\pi - x & \text{für } \pi < x \leq 2\pi \end{cases}$$



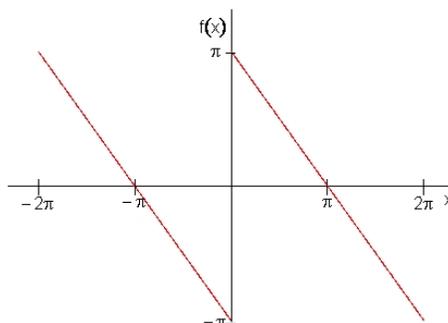
Da f gerade, ist $b_n = 0$ für alle n und

$$a_n = \frac{2}{\pi} \int_0^{\pi} f(x) \cos(nx) dx = \frac{2}{\pi} \int_0^{\pi} x \cos(nx) dx = \begin{cases} \pi & \text{für } n = 0 \\ 0 & \text{für } n \text{ gerade} \\ -\frac{4}{\pi n^2} & \text{für } n \text{ ungerade} \end{cases}$$

Folglich ist

$$F_f(x) = \frac{\pi}{2} - \frac{4}{\pi} \left(\cos(x) + \frac{1}{3^2} \cos(3x) + \frac{1}{5^2} \cos(5x) + \dots \right)$$

(ii) f sei 2π -periodisch mit $f(0) = 0$ und $f(x) = \pi - x$ für $0 < x < 2\pi$



'Sägezahnfunktion'

Da f ungerade, ist $a_n = 0$ für alle n und

$$b_n = \frac{2}{\pi} \int_0^\pi f(x) \sin(nx) dx = \frac{2}{\pi} \int_0^\pi (\pi - x) \sin(nx) dx = \frac{2}{n}$$

Folglich ist

$$F_f(x) = 2 \left(\sin(x) + \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} + \dots \right) = 2 \sum_{n=1}^{\infty} \frac{\sin(nx)}{n}$$

Wie bei Taylorreihen gilt auch bei Fourierreihen:

- 1) Die Fourierreihe muss nicht konvergieren.
- 2) Ist F_f konvergent in x , so muss der Wert der Reihe, $F_f(x)$, nicht mit $f(x)$ übereinstimmen.

Hinreichendes Kriterium für die Konvergenz der Fourierreihe F_f gegen f

Satz Ist $f : [0, 2\pi] \rightarrow \mathbb{R}$ stückweise stetig differenzierbar (d.h. es gibt eine Unterteilung $0 = x_0 < x_1 < \dots < x_n = 2\pi$ von $[0, 2\pi]$, so dass f' auf $]x_{i-1}, x_i[$ existiert und stetig fortsetzbar auf $[x_{i-1}, x_i]$ für alle i), so konvergiert die Fourierreihe F_f für alle x und

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) = \frac{f(x_+) + f(x_-)}{2}$$

Hierbei ist

$$f(x_+) = \lim_{y \rightarrow x^+} f(y) \text{ der rechtsseitige Grenzwert, und}$$

$$f(x_-) = \lim_{y \rightarrow x^-} f(y) \text{ der linksseitige Grenzwert von } f \text{ in } x.$$

In allen Punkten x also, in denen f stetig ist, folgt unter den Annahmen des Satzes, dass

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

Auf jedem abgeschlossenen Teilintervall, auf dem eine stückweise stetig differenzierbare Funktion f stetig ist, konvergiert die Fourierreihe F_f dann sogar gleichmäßig gegen f .

Beispiele 1.2

- (i) f aus Beispiel 1.1 (i) ist stetig und stückweise stetig differenzierbar. Daher ist die zugehörige Fourierreihe

$$F_f(x) = \frac{\pi}{2} - \frac{4}{\pi} \left(\cos(x) + \frac{1}{3^2} \cos(3x) + \frac{1}{5^2} \cos(5x) + \dots \right)$$

gleichmäßig konvergent gegen f .

Insbesondere gilt zum Beispiel

$$0 = f(0) = F_f(0) = \frac{\pi}{2} - \frac{4}{\pi} \left(1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots \right)$$

also

$$\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$$

- (ii) f aus Beispiel 1.2 (ii) ist stückweise stetig differenzierbar, aber nicht stetig in 0. Es ist also

$$F_f(x) = 2 \sum_{k=1}^{\infty} \frac{\sin(kx)}{k} = \frac{f(x_+) + f(x_-)}{2} \quad \text{für alle } x.$$

Gibbssches Phänomen

Ist f stückweise stetig differenzierbar, aber nicht stetig in x , so konvergiert $F_f(x)$ gegen den Mittelwert $\frac{1}{2}(f(x_+) + f(x_-))$. Die Konvergenz ist nicht gleichmäßig, vielmehr beobachtet man in x ein Überschwingen der Fourierpolynome mit asymptotisch etwa 9% der Sprunghöhe (genauer: 8,949%).

Beispiel 1.3 (Sägezahnfunktion aus Beispiel 1.1 (ii))

$$F_f(x) = 2 \sum_{k=1}^{\infty} \frac{\sin(kx)}{k}$$

Für die Partialsummen

$$s_n(x) = 2 \sum_{k=1}^n \frac{\sin(kx)}{k} = 2 \left(\sin(x) + \frac{\sin(2x)}{2} + \dots + \frac{\sin(nx)}{n} \right)$$

gilt

$$s_n(x) = \int_0^x \frac{\sin\left(\left(n + \frac{1}{2}\right)t\right)}{\sin\left(\frac{t}{2}\right)} dt - \underbrace{x}_{=f(x)-\pi}$$

Für die Ableitung des Fehlers $r_n(x) = s_n(x) - f(x)$ folgt also

$$r'_n(x) = \frac{\sin\left(\left(n + \frac{1}{2}\right)x\right)}{\sin\left(\frac{1}{2}x\right)}$$

mit kleinster positiver Nullstelle in $x_n = \frac{\pi}{n+\frac{1}{2}}$. Hier besitzt der Fehler r_n in der Tat ein globales Maximum mit

$$\begin{aligned} r_n(x_n) &= \int_0^{x_n} \frac{\sin\left(\left(n+\frac{1}{2}\right)t\right)}{\sin\left(\frac{1}{2}t\right)} dt - \pi = \int_0^\pi \frac{\sin(u) du}{\left(n+\frac{1}{2}\right) \sin\left(\frac{u}{2\left(n+\frac{1}{2}\right)}\right)} - \pi \\ &> 2 \underbrace{\int_0^\pi \frac{\sin(u)}{u} du}_{\sim 1.8519} - \pi \sim 0.1789 \cdot \pi \sim 0.09 \cdot \underbrace{2\pi}_{\text{Sprunghöhe}} \end{aligned}$$

Nebenrechnung zur Darstellung der Partialsumme:

$$\begin{aligned} s'_n(x) &= 2 \sum_{k=1}^n \cos(kx) = 2 \operatorname{Re} \left(\sum_{k=1}^n e^{ikx} \right) = 2 \operatorname{Re} \left(e^{ix} \frac{1 - e^{inx}}{1 - e^{ix}} \right) \\ &= 2 \operatorname{Re} \left(\frac{e^{i\frac{x}{2}} - e^{i\left(n+\frac{1}{2}\right)x}}{e^{-i\frac{x}{2}} - e^{i\frac{x}{2}}} \right) = -\frac{1}{\sin\left(\frac{x}{2}\right)} \cdot \left(\sin\left(\frac{x}{2}\right) - \sin\left(\left(n+\frac{1}{2}\right)x\right) \right) \\ &= \frac{\sin\left(\left(n+\frac{1}{2}\right)x\right)}{\sin\left(\frac{x}{2}\right)} - 1 \end{aligned}$$

und damit folgt

$$s_n(x) = \int_0^x \frac{\sin\left(\left(n+\frac{1}{2}\right)t\right)}{\sin\left(\frac{t}{2}\right)} dt - x$$

Allgemeine Perioden

Ist f periodisch mit Periode $L > 0$, so ersetzt man $\cos(nx)$, $\sin(nx)$ durch $\cos\left(\frac{2\pi}{L}nx\right)$, $\sin\left(\frac{2\pi}{L}nx\right)$. Die zugehörige Fourierreihe lautet also

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{2\pi}{L}nx\right) + b_n \sin\left(\frac{2\pi}{L}nx\right) \right)$$

mit

$$a_n = \frac{2}{L} \int_0^L f(x) \cos\left(\frac{2\pi}{L}nx\right) dx, \quad n = 0, 1, 2, \dots$$

$$b_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{2\pi}{L}nx\right) dx, \quad n = 1, 2, \dots$$

Alle Aussagen zum 2π -periodischen Fall übertragen sich sinngemäß auf den Fall allgemeiner Periode.

Komplexe Schreibweise

Nach der Eulerschen Formel gilt

$$e^{it} = \cos t + i \sin t, \quad t \in \mathbb{R}.$$

Daher können wir eine Fourierreihe

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

auch in der Form

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}, \quad x \in \mathbb{R}$$

schreiben, mit den **komplexen Fourierkoeffizienten**

$$\begin{aligned} c_n &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx \\ &:= \frac{1}{2\pi} \int_0^{2\pi} f(x) \cos(-nx) dx + i \frac{1}{2\pi} \int_0^{2\pi} f(x) \sin(-nx) dx \\ &= \begin{cases} \frac{a_0}{2} & \text{für } n = 0 \\ \frac{1}{2}(a_n - i b_n) & \text{für } n > 0 \\ \frac{1}{2}(a_n + i b_n) & \text{für } n < 0 \end{cases} \end{aligned}$$

2 Kurven im \mathbb{R}^n

Zur Erinnerung:

$$\mathbb{R}^n = \left\{ X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} : x_1, \dots, x_n \in \mathbb{R} \right\}$$

ist der n -dimensionale **Punktraum** mit den bekannten Rechenoperationen Addition und Skalarmultiplikation. Für einen Punkt $X = [x_1, \dots, x_n]^T$ definieren wir seine **Norm** durch

$$\|X\| := \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}$$

mit den bekannten Eigenschaften

- (i) **Positiv-Definitheit** $\|X\| \geq 0$ und $\|X\| = 0$ genau dann, wenn $X = 0$.
- (ii) **Homogenität** $\|\alpha X\| = |\alpha| \|X\|$ für $\alpha \in \mathbb{R}$
- (iii) **Dreiecksungleichung** $\|X + Y\| \leq \|X\| + \|Y\|$

Für Punkte $X = [x_1, \dots, x_n]^T, Y = [y_1, \dots, y_n]^T$ ist das Skalarprodukt definiert durch

$$\langle X, Y \rangle := x_1 y_1 + x_2 y_2 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i.$$

Es gilt die **Cauchy-Schwarzsche Ungleichung**

$$|\langle X, Y \rangle| \leq \|X\| \cdot \|Y\| \quad \text{für } X, Y \in \mathbb{R}^n.$$

Abstände und Konvergenz von Folgen

Die Norm definiert einen Abstand zwischen Punkten X und Y im \mathbb{R}^n , indem wir als Abstand gerade $\|X - Y\|$ nehmen.

Für $\varepsilon > 0$ und $X \in \mathbb{R}^n$ heißt die Menge

$$U_\varepsilon(X) := \{Y \in \mathbb{R}^n : \|X - Y\| < \varepsilon\},$$

also die Menge aller Punkte $Y \in \mathbb{R}^n$, deren Abstand zu X kleiner als ε ist, die ε -**Umgebung von** X . Geometrisch beschreibt $U_\varepsilon(x)$ im

- \mathbb{R}^1 das offene Intervall $]X - \varepsilon, X + \varepsilon[$
- \mathbb{R}^2 eine Kreisscheibe mit Mittelpunkt X und Radius ε (ohne Rand)
- \mathbb{R}^3 einen Ball mit Mittelpunkt X und Radius ε (ohne die Balloberfläche).

Mithilfe des Abstandsbegriffes lässt sich nun auch Konvergenz von Punktfolgen im \mathbb{R}^n definieren:

Definition Es sei $(X_k)_{k \geq k_0}$ eine Folge von Punkten im \mathbb{R}^n . Die Folge heißt **konvergent** gegen $X \in \mathbb{R}^n$, falls gilt: Für alle $\varepsilon > 0$ existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$\|X_k - X\| < \varepsilon \quad \text{für alle } k \geq N_\varepsilon$$

X heißt **Grenzwert** der Folge $(X_k)_{k \geq k_0}$ und wir schreiben

$$\lim_{k \rightarrow \infty} X_k = X.$$

Bemerkung

- (i) Die Folge (X_k) konvergiert gegen X genau dann, wenn für alle $\varepsilon > 0$ nur endlich viele Folgenglieder außerhalb der ε -Umgebung um X liegen, d.h. alle, bis auf endlich viele, X_k liegen in $U_\varepsilon(X)$.
- (ii) Ist $X_k = [x_{1,k}, \dots, x_{n,k}]^T$ und $X = [x_1, \dots, x_n]^T$, so gilt:

$$\lim_{k \rightarrow \infty} X_k = X \quad \text{genau dann, wenn} \quad \lim_{k \rightarrow \infty} x_{i,k} = x_i \quad \text{für alle } i = 1, \dots, n$$

D.h. die Punktfolge $(X_k)_{k \geq k_0}$ konvergiert im \mathbb{R}^n gegen X genau dann, wenn die Komponentenfolgen $(x_{i,k})_{k \geq k_0}$ gegen die Komponenten x_i von x konvergieren.

Definition Es sei $[a, b] \subset \mathbb{R}$ ein Intervall. Eine Abbildung $X : [a, b] \rightarrow \mathbb{R}^n$ heißt **Kurve** (bzw. **Weg**) im \mathbb{R}^n . Die Punktmenge

$$\{X(t) : t \in [a, b]\}$$

heißt **Bahn** (bzw. **Spur**) der Kurve X .

Eine Kurve $X(t) = [x_1(t), \dots, x_n(t)]^T$, $t \in [a, b]$, im \mathbb{R}^n setzt sich zusammen aus insgesamt n **Komponentenfunktionen**.

Beispiele Kurven dienen der Parameterdarstellung von Punktmengen im \mathbb{R}^n , etwa

- (i) $n = 2$ **Kreislinie** Die Kurve

$$X(t) = [\cos(t), \sin(t)]^T, t \in [0, 2\pi]$$

durchläuft den Einheitskreis im \mathbb{R}^2 einmal im entgegengesetzten Uhrzeigersinn. Entsprechend

$$X(t) = r[\cos(t), \sin(t)]^T, t \in [0, 2\pi], r > 0$$

für einen Kreis mit Radius r .

- (ii) $n = 3$ **Schraublinie**

$$X(t) = [r \cos(t), r \sin(t), ct]^T, t \in [0, 2\pi n], r > 0, c \neq 0$$

durchläuft eine Schraublinie im \mathbb{R}^3 .

- r heißt Radius,

- n heißt Windungszahl,
- $2\pi|c|$ heißt Ganghöhe.

(iii) Ist $f : [a, b] \rightarrow \mathbb{R}$ eine Funktion, so beschreibt

$$X(t) = [t, f(t)], t \in [a, b]$$

den Funktionsgraph von f als Kurve im \mathbb{R}^2 .

(iv) Sind $P, R \in \mathbb{R}^n, R \neq 0$, so beschreibt die Kurve

$$X(t) = P + t \cdot R, \quad t \in \mathbb{R}$$

die Gerade mit Aufpunkt P und Richtungsvektor R .

Definition Eine Kurve $X = [x_1, \dots, x_n]^T : [a, b] \rightarrow \mathbb{R}^n$ heißt

- stetig**, falls alle Komponentenfunktionen $x_i : [a, b] \rightarrow \mathbb{R}$ stetig sind,
- differenzierbar**, falls alle Komponentenfunktionen $x_i : [a, b] \rightarrow \mathbb{R}$ differenzierbar sind. In diesem Falle schreiben wir

$$\dot{X}(t) := \frac{d}{dt} X(t) := \lim_{h \rightarrow 0} \frac{1}{h} (X(t+h) - X(t)) = [\dot{x}_1(t), \dots, \dot{x}_n(t)]^T$$

$\dot{X}(t)$ heißt **Tangentialvektor** der Kurve X zum Parameterwert t . Falls $\dot{X}(t) \neq 0$, so heißt $T(t) = \frac{\dot{X}(t)}{\|\dot{X}(t)\|}$ **Tangenteneinheitsvektor**.

(iii) **regulär**, falls X stetig differenzierbar und $\dot{X}(t) \neq 0$ für alle $t \in [a, b]$.

Interpretation

Geometrisch $\dot{X}(t) = \lim_{h \rightarrow 0} \frac{1}{h} \underbrace{(X(t+h) - X(t))}_{\text{Sekante}}$

Geometrisch beschreibt $\dot{X}(t)$ also die **Tangente** an die Bahn von X im Punkt $X(t)$.

Physikalisch Stellen wir uns X als Bahnkurve eines Massenpunktes vor, so beschreibt $\dot{X}(t)$ die momentane Ortsveränderung, also den Geschwindigkeitsvektor des Massenpunktes im Zeitpunkt t .

Kurvenlänge

Es sei $X : [a, b] \rightarrow \mathbb{R}^n$ eine Kurve im \mathbb{R}^n . Zu gegebener Zerlegung Z von $[a, b]$ der Form

$$a = t_0 < t_1 < \dots < t_n = b$$

beschreibt

$$L(X, Z) = \sum_{i=1}^n \|X(t_i) - X(t_{i-1})\|$$

die Länge des Polygonzuges durch die Punkte

$$X(a) = X(t_0), X(t_1), \dots, X(t_n) = X(b)$$

Ist X stetig differenzierbar, so können wir die Länge des i -ten Teilstücks approximieren durch

$$\|X(t_i) - X(t_{i-1})\| \sim \|\dot{X}(t_{i-1})\|(t_i - t_{i-1})$$

und damit

$$L(X, Z) \sim \sum_{i=1}^n \|\dot{X}(t_{i-1})\|(t_i - t_{i-1}) \quad (2.1)$$

Der Ausdruck auf der rechten Seite (2.1) ist eine Riemann-Summe zur Funktion $f(t) = \|\dot{X}(t)\|$ und Zerlegung Z . Konvergiert die Feinheit der Zerlegung Z

$$\delta(Z) = \max\{t_i - t_{i-1} : 1 \leq i \leq n\}$$

gegen 0, so passt sich der Polygonzug immer genauer dem exakten Verlauf der Kurve an und die zugehörigen Riemann-Summen konvergieren

$$L(X, Z) = \sum_{i=1}^n \|X(t_i) - X(t_{i-1})\| \sim \sum_{i=1}^n \|\dot{X}(t_{i-1})\|(t_i - t_{i-1}) \xrightarrow{\delta(Z) \rightarrow 0} \int_a^b \|\dot{X}(t)\| dt$$

Der Fehler, der bei der Approximation von $L(X, Z)$ durch die Riemann-Summe in (2.1) gemacht wurde, geht dabei asymptotisch gegen 0, d.h. es gilt

$$\lim_{\delta(Z) \rightarrow 0} L(X, Z) = \int_a^b \|\dot{X}(t)\| dt.$$

Definition Ist $X : [a, b] \rightarrow \mathbb{R}^n$ stetig differenzierbare Kurve im \mathbb{R}^n , so ist die Länge $L(X)$ von X definiert als das Riemann-Integral

$$L(X) = \int_a^b \|\dot{X}(t)\| dt = \int_a^b \sqrt{\dot{X}_1^2(t) + \dots + \dot{X}_n^2(t)} dt$$

Beispiele

- (i) $X(t) = r [\cos(t), \sin(t)]^T$, $t \in [0, 2\pi]$, ist stetig differenzierbar, $\dot{X}(t) = r [-\sin(t), \cos(t)]^T$, also gilt

$$\|\dot{X}(t)\| = \underbrace{\sqrt{(r(-\sin(t)))^2 + (r \cos(t))^2}}_{=r^2((\sin(t))^2 + (\cos(t))^2)=r^2} = r$$

und somit

$$L(X) = \int_0^{2\pi} \|\dot{X}(t)\| dt = \int_0^{2\pi} r dt = r \cdot 2\pi = \text{Umfang des Kreises mit Radius } r$$

- (ii) Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar, $X(t) = [t, f(t)]^T$ die Parametrisierung des Funktionsgraphen, so ist

$$L(X) = \int_a^b \|\dot{X}(t)\| dt = \int_a^b \sqrt{1 + (f'(t))^2} dt$$

seine Länge.

Zum Beispiel ergibt sich für die Normalparabel $f(t) = t^2$, $t \in [-1, 1]$, als Länge für den dazugehörigen Funktionsgraphen

$$L(X) = \int_{-1}^1 \sqrt{1 + 4t^2} dt = \sqrt{5} + \frac{1}{4} \ln \left(\frac{\sqrt{5} + 2}{\sqrt{5} - 2} \right)$$

Haben dabei verwandt, dass $\frac{1}{4} (2t\sqrt{1 + 4t^2} + \ln(2t + \sqrt{1 + 4t^2}))$ Stammfunktion zu $\sqrt{1 + 4t^2}$ ist.

Bemerkung Die Länge einer Kurve kann auch für stetige, stückweise stetig differenzierbare Kurven $X : [a, b] \rightarrow \mathbb{R}^n$ definiert werden. (X heißt stückweise stetig differenzierbar, falls eine Unterteilung $a = t_0 < t_1 < \dots < t_n = b$ existiert, so dass X auf $]t_{i-1}, t_i[$ zu einer stetig differenzierbaren Kurve auf $[t_{i-1}, t_i]$ fortsetzbar ist.)

Beispiel

$$X(t) = \begin{cases} [0, t]^T & \text{für } t \in [0, 1] \\ [t - 1, 2 - t]^T & \text{für } t \in]1, 2] \end{cases}, \quad \text{also} \quad \|\dot{X}(t)\| = \begin{cases} 1 & \text{für } t \in [0, 1] \\ \sqrt{2} & \text{für } t \in]1, 2] \end{cases}$$

und somit $L(X) = \int_0^1 1 dt + \int_1^2 \sqrt{2} dt = 1 + \sqrt{2}$.

Krümmung von Kurven

Es sei X zweimal stetig differenzierbare reguläre Kurve. Die Änderung der Geschwindigkeitsrichtung einer Kurve, bezogen auf die Kurvenlänge

$$\frac{\dot{T}(t)}{\|\dot{X}(t)\|} = \lim_{h \rightarrow 0} \frac{T(t+h) - T(t)}{\|X(t+h) - X(t)\|}$$

beschreibt die **Krümmungsrichtung** von X in $X(t)$.

Seine Länge $\kappa(t) = \frac{\|\dot{T}(t)\|}{\|\dot{X}(t)\|}$ heißt **Krümmung**.

Beispiel Kreislinie mit Radius r

$$X(t) = r [\cos(t), \sin(t)]^T, t \in [0, 2\pi]$$

Für den Tangenteneinheitsvektor gilt

$$T(t) = \frac{\dot{X}(t)}{\|\dot{X}(t)\|} = [-\sin(t), \cos(t)]^T \quad \text{unabhängig vom Radius } r!$$

Der Krümmungsvektor, also die Änderung der Tangentialrichtung,

$$\dot{T}(t) = -[\cos(t), \sin(t)]^T = -\frac{1}{r} X(t)$$

zeigt zum Kreismittelpunkt. Für die Krümmung errechnet man $\kappa(t) = \frac{1}{r}$. Sie ist also umgekehrt proportional zum Radius des Kreises.

Kurven im \mathbb{R}^3

Es sei $X : [a, b] \rightarrow \mathbb{R}^3$ eine zweimal stetig differenzierbare reguläre Kurve im \mathbb{R}^3 mit $\dot{T}(t) \neq 0$ für alle $t \in [a, b]$. Dann stehen die drei Einheitsvektoren

$$\begin{aligned} T(t) &= \frac{\dot{X}(t)}{\|\dot{X}(t)\|} \quad (\text{Tangenteneinheitsvektor}) \\ N(t) &= \frac{\dot{T}(t)}{\|\dot{T}(t)\|} \quad (\text{Hauptnormaleneinheitsvektor}) \\ B(t) &= T(t) \times N(t) \quad (\text{Binormaleneinheitsvektor}) \end{aligned}$$

orthogonal aufeinander und bilden in dieser Reihenfolge ein Rechtssystem (das sogenannte **begleitende Dreibein**).

Die von $T(t)$ und $N(t)$ aufgespannte Ebene

$$E(t) : X(t) + \lambda T(t) + \mu N(t) \quad \lambda, \mu \in \mathbb{R}$$

heißt **Schmiegeebene** von X an der Stelle $X(t)$.

Es sei X dreimal differenzierbare reguläre Kurve mit $\dot{T}(t) \neq 0$ für alle $t \in [a, b]$. Die Änderung der Binormalenrichtung bezogen auf die Kurvenlänge

$$\frac{\dot{B}(t)}{\|\dot{X}(t)\|} = \lim_{h \rightarrow 0} \frac{B(t+h) - B(t)}{\|X(t+h) - X(t)\|}$$

beschreibt das Herauswinden der Kurve aus der Schmiegeebene und wird als **Torsionsvektor** bezeichnet. $\dot{B}(t)$ ist orthogonal zu $B(t)$ und auch orthogonal zu $T(t)$, denn

$$\dot{B}(t) = \frac{d}{dt} (T(t) \times N(t)) = \dot{T}(t) \times N(t) + T(t) \times \dot{N}(t) = T(t) \times \dot{N}(t).$$

Damit gibt es $\tau(t)$ mit

$$\frac{\dot{B}(t)}{\|\dot{X}(t)\|} = -\tau(t)N(t).$$

$\tau(t)$ heißt **Torsion**. Für den Absolutbetrag gilt $|\tau(t)| = \frac{\|\dot{B}(t)\|}{\|\dot{X}(t)\|}$.

Beispiel (Schraublinie)

$$X(t) = [r \cos(t), r \sin(t), ct]^T, t \in [0, 2\pi n], r > 0, c > 0$$

$$\dot{X}(t) = \begin{bmatrix} -r \sin(t) \\ r \cos(t) \\ c \end{bmatrix}, \quad \ddot{X}(t) = \begin{bmatrix} -r \cos(t) \\ -r \sin(t) \\ 0 \end{bmatrix}, \quad \ddot{\ddot{X}}(t) = \begin{bmatrix} r \sin(t) \\ -r \cos(t) \\ 0 \end{bmatrix}$$

Also gilt

$$T(t) = \frac{1}{R} [-r \sin(t), r \cos(t), c]^T, \quad R = \sqrt{r^2 + c^2}$$

$$N(t) = [-\cos(t), -\sin(t), 0]^T, \quad B(t) = T(t) \times N(t) = \frac{1}{R} [c \sin(t), -c \cos(t), r]^T$$

Hieraus erhält man schließlich Krümmung $\kappa(t) \equiv \frac{r}{r^2+c^2}$ und Torsion $\tau(t) \equiv \frac{c}{r^2+c^2}$.

3 Funktionen in mehreren Variablen

Es sei $D \subset \mathbb{R}^n$. Eine Funktion

$$f : D \rightarrow \mathbb{R}, \\ X = [x_1, \dots, x_n] \mapsto f(X) = f(x_1, \dots, x_n)$$

heißt **Funktion in n -Variablen**.

Beispiele

$$\begin{aligned} f(x, y) &= 3xy + 2x \quad \text{auf } \mathbb{R}^2 \\ f(x, y, z) &= e^z + \sin(x + y) \quad \text{auf } \mathbb{R}^3 \\ f(X) &= \|X\|^2 = x_1^2 + \dots + x_n^2 \quad \text{auf } \mathbb{R}^n \end{aligned}$$

Darstellungsweisen

Wie im Falle $n = 1$ definieren wir den **Funktionsgraphen von f** als die Menge

$$\Gamma_f = \{(x, f(x)) : x \in D\} \subset D \times \mathbb{R} \quad (\subset \mathbb{R}^{n+1})$$

Beispiele Für $n = 2$ beschreibt der Funktionsgraph von f eine Fläche im \mathbb{R}^3 , etwa ein Paraboloid im Falle von $f(x, y) = x^2 + y^2$.

Für $n \geq 3$ können wir den Funktionsgraphen als Fläche im \mathbb{R}^3 nicht mehr darstellen. Als Alternativen bieten sich an

Schnittkurvendiagramme

Hält man jeweils $n - 1$ Variablen fest, etwa x_1, \dots, x_{n-1} , so erhält man aus f eine Funktion in einer Variablen:

$$t \mapsto f(x_1, \dots, x_{n-1}, t)$$

Der zugehörige Funktionsgraph ist die Menge

$$\{(t, f(x_1, \dots, x_{n-1}, t)) : (x_1, \dots, x_{n-1}, t) \in D\} \subset \mathbb{R}^2$$

und er wird als Schnittkurve (von Γ_f mit der (x_1, \dots, x_{n-1}) -Ebene) bezeichnet.

Beispiel ($n = 2$) $f(x, y) = x^2 + y^2$ Hält man x (bzw. y) fest, so erhält man in der jeweils anderen Variablen eine Parabel.

Höhenliniendiagramme

Als weitere Alternative zur Darstellung einer Funktion in mehreren Variablen betrachtet man Höhenliniendiagramme.

Definition Zu gegebenem $c \in \mathbb{R}$ heißt die Punktmenge

$$N_c = \{X \in D : f(X) = c\}$$

die **Niveaumenge zum Niveau c** (und für $n = 2$ speziell Höhenlinie zur Höhe c).

Man bestimmt die Menge N_c durch **Auflösen der Gleichung**

$$f(X) = c \quad \text{nach } X.$$

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

Die Gleichung $x^2 + y^2 = c$ hat für

- $c < 0$ keine Lösung
- $c = 0$ die Lösung $x = y = 0$
- $c > 0$ einen Kreis mit Mittelpunkt 0 und Radius \sqrt{c} als Lösung.

Stetigkeit von Funktionen in mehreren Variablen

Aufbauend auf dem Konvergenzbegriff von Folgen im \mathbb{R}^n können wir nun auch den Stetigkeitsbegriff auf Funktionen mehrerer Veränderlicher übertragen.

Definition Es sei $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$ und $X \in \mathbb{R}^n$ so dass X Häufungspunkt von D (d.h. in jeder ε -Umgebung $U_\varepsilon(X)$ gibt es Punkte aus D):

- (i) (Grenzwerte von Funktionen) f besitzt in X den Grenzwert c , falls

$$\lim_{k \rightarrow \infty} f(X_k) = c$$

für **jede Folge** $(X_k) \subset D$ mit $\lim_{k \rightarrow \infty} X_k = X$. Wir schreiben in diesem Falle $\lim_{Y \rightarrow X} f(Y) = c$.

- (ii) (Stetigkeit) Ist $X \in D$, so heißt f **stetig in** X , falls

$$\lim_{Y \rightarrow X} f(Y) = f(X)$$

- (iii) f heißt **stetig auf** D , falls f in allen Punkten $X \in D$ stetig ist.

Bemerkungen

- (i) Die **Rechenregeln für Grenzwerte** für Funktionen einer Veränderlichen übertragen sich auf Funktionen mehrerer Veränderlicher, also

$$\begin{aligned} \lim_{Y \rightarrow X} f(Y) + g(Y) &= \lim_{Y \rightarrow X} f(Y) + \lim_{Y \rightarrow X} g(Y) \\ \lim_{Y \rightarrow X} f(Y) \cdot g(Y) &= \lim_{Y \rightarrow X} f(Y) \cdot \lim_{Y \rightarrow X} g(Y) \\ \lim_{Y \rightarrow X} cf(Y) &= c \lim_{Y \rightarrow X} f(Y) \\ \lim_{Y \rightarrow X} \frac{f(Y)}{g(Y)} &= \frac{\lim_{Y \rightarrow X} f(Y)}{\lim_{Y \rightarrow X} g(Y)} \quad \text{falls } \lim_{Y \rightarrow X} g(Y) \neq 0 \end{aligned}$$

- (ii) Aus den Rechenregeln für Grenzwerte folgt wie im Falle von Funktionen einer Veränderlicher: Sind $f, g : D \rightarrow \mathbb{R}$ stetig, so sind auch folgende Funktionen stetig:

$$f + g, f \cdot g, c \cdot f \text{ und } \frac{f}{g} \text{ auf } D_0 = \{ X \in D \mid g(X) \neq 0 \}$$

Ist $h : E \rightarrow \mathbb{R}$ stetig, $f(D) \subset E$, so ist auch die Verkettung $h \circ f$ stetig.

Beispiele 3.1

- (i) Die Projektionen $p_i : \mathbb{R}^n \rightarrow \mathbb{R}, x = [x_1, \dots, x_n] \mapsto x_i$ sind stetig. Damit sind dann auch die folgenden Funktionen stetig:

- lineare Funktionen $f(x) = a_1x_1 + \dots + a_nx_n$
- Polynome in n -Variablen $p(x) = p(x_1, \dots, x_n) = \sum_{1 \leq k_i \leq m} a_{k_1 \dots k_n} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$
- rationale Funktionen (also Quotienten von Polynomen) $\frac{p(x)}{q(x)}$ auf $D_0 = \{ x \in \mathbb{R}^n \mid q(x) \neq 0 \}$

- (ii) Die Funktion

$$f(x, y) = \begin{cases} \frac{x^2y^2}{x^2+y^2} & \text{für } [x, y]^T \neq [0, 0]^T \\ 0 & \text{für } [x, y]^T = [0, 0]^T \end{cases}$$

ist stetig in 0, denn für **jede** Folge $X_k = [x_k, y_k]^T$, die gegen 0 konvergiert, gilt:

$$|f(x_k, y_k) - 0| = \frac{x_k^2 y_k^2}{x_k^2 + y_k^2} \leq y_k^2 \rightarrow 0$$

daher ist $\lim_{X \rightarrow 0} f(X) = 0 = f(0)$.

- (iii) Die Funktion

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

ist **nicht stetig in 0**, denn für die Folge $X_k = [\frac{1}{k}, \frac{1}{k}]^T$ gilt $\lim_{k \rightarrow \infty} X_k = 0$, aber

$$f(X_k) = f\left(\frac{1}{k}, \frac{1}{k}\right) = \frac{\frac{1}{k} \cdot \frac{1}{k}}{\left(\frac{1}{k}\right)^2 + \left(\frac{1}{k}\right)^2} = \frac{1}{2}$$

und damit

$$\lim_{k \rightarrow \infty} f(X_k) = \frac{1}{2} \neq f(0).$$

Eigenschaften stetiger Funktionen

Für stetige Funktionen $f : [a, b] \rightarrow \mathbb{R}$ einer reeller Veränderlicher hatten wir bereits gesehen:

- f ist beschränkt
- f besitzt Maximum und Minimum

Entsprechendes gilt für die Funktionen in mehreren Veränderlichen. Dazu muss jedoch zuerst eine Verallgemeinerung abgeschlossener Intervalle im \mathbb{R}^n gefunden werden.

Definition Es sei $D \subset \mathbb{R}^n$ eine Teilmenge

- (i) $X \in D$ heißt **innerer Punkt von D** , wenn es eine ε -Umgebung von X gibt, die ganz in D enthalten ist.
- (ii) D heißt **offen**, wenn jeder Punkt von D ein innerer Punkt ist.
- (iii) $X \in \mathbb{R}^n$ heißt **Randpunkt von D** , wenn jede ε -Umgebung von X sowohl Punkte aus D , als auch Punkte aus dem Komplement $\mathbb{R}^n \setminus D = \{Y \in \mathbb{R}^n : Y \notin D\}$ enthält.
Die Menge aller Randpunkte von D heißt **Rand** und wird mit ∂D bezeichnet.
- (iv) D heißt **abgeschlossen**, wenn D alle Randpunkte enthält.

Beispiele

- (i) Die ε -Umgebung eines Punktes X

$$U_\varepsilon(X) = \{Y \in \mathbb{R}^n : \|X - Y\| < \varepsilon\}$$

ist selber eine offene Menge mit Rand

$$\partial U_\varepsilon(X) = \{Y \in \mathbb{R}^n : \|X - Y\| = \varepsilon\}$$

Folglich ist die Menge $\{Y \in \mathbb{R}^n : \|X - Y\| \leq \varepsilon\}$ abgeschlossen.

- (ii) Rechtecke $Q \subset \mathbb{R}^2$ der Form

$$Q = \{[x, y]^T : 1 < x < 2, -2 < y < 5\}$$

sind offen, Rechtecke Q der Form

$$Q = \{[x, y]^T : 1 \leq x \leq 2, -2 \leq y \leq 5\}$$

abgeschlossen, Rechtecke der Form

$$Q = \{[x, y]^T : 1 < x \leq 2, -2 \leq y < 5\}$$

sind weder offen noch abgeschlossen.

Entsprechendes gilt für Quader $Q \subset \mathbb{R}^3$.

Definition Es sei $D \subset \mathbb{R}^n$ eine Teilmenge

- (i) D heißt **beschränkt**, falls eine Konstante M existiert mit

$$\|X\| \leq M \quad \text{für alle } X \in D$$

(ii) D heißt **kompakt**, falls D abgeschlossen und beschränkt ist.

Die kompakten Mengen sind gerade die Analoga abgeschlossener, beschränkter Intervalle $[a, b]$, denn es gilt: Ist $F : D \rightarrow \mathbb{R}$ stetig und $D \subset \mathbb{R}^n$ kompakt. Dann gilt

(i) f ist beschränkt.

(ii) **Existenz des Maximums / Minimums:**

Es existieren $X_{\max}, X_{\min} \in D$ mit

$$f(X_{\min}) \leq f(X) \leq f(X_{\max}) \quad \forall X \in D$$

4 Differentialrechnung von Funktionen mehrerer Veränderlicher

4.1 Partielle Ableitungen

Es sei $D \subset \mathbb{R}^n$, $f : D \rightarrow \mathbb{R}$ eine Funktion und $X = [x_1, \dots, x_n]^T \in D$ ein innerer Punkt. Für kleine h ist

$$h \mapsto f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n)$$

wohldefiniert als Funktion von h .

Existiert dann der Grenzwert der Differenzenquotienten in der i -ten Koordinatenrichtung

$$\lim_{h \rightarrow 0} \frac{1}{h} (f(x_1, \dots, x_i + h, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)) =: \frac{\partial f}{\partial x_i}(X)$$

so heißt f an der Stelle X **partiell differenzierbar nach x_i** , $1 \leq i \leq n$. Der Grenzwert $\frac{\partial f}{\partial x_i}(X)$ heißt **partielle Ableitung** von f nach x_i an der Stelle X .

f heißt an der Stelle X **partiell differenzierbar**, falls alle partiellen Ableitungen $\frac{\partial f}{\partial x_i}(X)$, $1 \leq i \leq n$, existieren. Der Vektor

$$\text{grad } f(X) := \left[\frac{\partial f}{\partial x_1}(X), \dots, \frac{\partial f}{\partial x_n}(X) \right]^T$$

der partiellen Ableitungen von f in X heißt **Gradient von f in X** .

Bemerkung Bezeichnen wir mit $e_i = [0, \dots, 0, 1, 0, \dots, 0]^T$ den i -ten Einheitsvektor im \mathbb{R}^n , so können wir schreiben

$$\frac{\partial f}{\partial x_i}(X) = \lim_{h \rightarrow 0} \frac{1}{h} (f(X + he_i) - f(X))$$

Beispiele 4.1

(i) $f(x, y) = 2x^2 + y^2$ auf \mathbb{R}^2 .

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \lim_{h \rightarrow 0} \frac{1}{h} (f(x+h, y) - f(x, y)) = \lim_{h \rightarrow 0} \frac{1}{h} (2(x+h)^2 + y^2 - (2x^2 + y^2)) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} (4xh + 2h^2) = 4x \end{aligned}$$

Entsprechend gilt $\frac{\partial f}{\partial y}(x, y) = 2y$

(ii) $f(x, y) = \sin(xy^2)$ auf \mathbb{R}^2

$$\frac{\partial f}{\partial x}(x, y) = y^2 \cos(xy^2) \qquad \frac{\partial f}{\partial y}(x, y) = 2yx \cos(xy^2)$$

Bei partieller Ableitung nach x behandelt man also y wie eine Konstante und bei partieller Ableitung nach y behandelt man x wie eine Konstante.

Alternative Bezeichnungen für partielle Ableitungen: $f_{x_i}, f_x, f_y, f_z, f_t, \dots$

Partielle Ableitungen höherer Ordnung

Ist f partiell differenzierbar nach x_i , so definiert

$$X \mapsto \frac{\partial f}{\partial x_i}(X)$$

wieder eine Funktion in n Veränderlichen, die man wieder gegebenenfalls nach x_j partiell differenzieren kann:

$$\frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (X) =: \frac{\partial^2 f}{\partial x_j \partial x_i} (X) \quad \text{für } 1 \leq i, j \leq n$$

Man erhält auf diese Weise die **partiellen Ableitungen zweiter Ordnung** von f .

Beispiele 4.2 (i) $f(x, y) = 2x^2 + y^2$, also

$$\frac{\partial f}{\partial x}(x, y) = 4x \qquad \frac{\partial f}{\partial y}(x, y) = 2y$$

und damit

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = 0 \qquad \frac{\partial^2 f}{\partial y \partial x}(x, y) = 0 \qquad \frac{\partial^2 f}{\partial x^2}(x, y) = 4 \qquad \frac{\partial^2 f}{\partial y^2}(x, y) = 2$$

(ii) $f(x, y) = \sin(xy^2)$, also

$$\frac{\partial^2 f}{\partial x^2}(x, y) = -y^4 \sin(xy^2)$$

Analog zu den partiellen Ableitungen zweiter Ordnung bildet man partielle Ableitungen höherer Ordnung:

$$\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(x), \frac{\partial^3 f}{\partial x_i^2 \partial x_j}(x), \dots$$

Bemerkungen

- (i) Bei Funktionen f in einer reellen Veränderlichen folgt aus der Differenzierbarkeit von f die Stetigkeit von f . Das ist für Funktionen in mehr als einer Veränderlichen im allgemeinen nicht richtig.

Beispiel

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

ist nicht stetig in 0 (siehe Beispiel 3.1(iii)). Jedoch ist

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

- (ii) Im allgemeinen ist die Reihenfolge der partiellen Ableitungen nicht vertauschbar, d.h. im allgemeinen ist z.B.

$$\frac{\partial^2 f}{\partial x \partial y} \neq \frac{\partial^2 f}{\partial y \partial x}$$

Beispiel

$$f(x, y) = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{für } [x, y]^T \neq [0, 0]^T \\ 0 & \text{sonst} \end{cases}$$

Für $[x, y]^T \neq 0$ gilt:

$$\begin{aligned} f_x(x, y) &= y \frac{x^2 - y^2}{x^2 + y^2} + xy \left(\frac{2x}{x^2 + y^2} - \frac{2x(x^2 - y^2)}{(x^2 + y^2)^2} \right) \\ &= y \left(\frac{x^2 - y^2}{x^2 + y^2} + \frac{4x^2 y^2}{(x^2 + y^2)^2} \right). \end{aligned}$$

Analog gilt

$$f_y(x, y) = x \left(\frac{x^2 - y^2}{x^2 + y^2} - \frac{4x^2 y^2}{(x^2 + y^2)^2} \right)$$

und $f_x(0, 0) = f_y(0, 0) = 0$. Folglich

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\partial f}{\partial x}(0, h) - \frac{\partial f}{\partial x}(0, 0) \right) = \lim_{h \rightarrow 0} \frac{1}{h} (h \cdot (-1)) = -1$$

aber

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\partial f}{\partial y}(h, 0) - \frac{\partial f}{\partial y}(0, 0) \right) = \lim_{h \rightarrow 0} \frac{1}{h} (h \cdot 1) = 1$$

Es gilt jedoch allgemein: Sind die partiellen Ableitungen k -ter Ordnung **stetig**, so ist die Reihenfolge der partiellen Ableitungen bis zur k -ten Ordnung vertauschbar.

- (iii) Für die partiellen Ableitungen von Summen und Produkten gelten dieselben **Rechenregeln** wie im Falle $n = 1$: Sind $f, g : D \mapsto \mathbb{R}$ an der Stelle X_0 partiell nach x_i differenzierbar, so sind auch die folgenden Funktionen an der Stelle X_0 partiell differenzierbar nach x_i :

- (a) **Linearität** $\alpha f + \beta g$ für $\alpha, \beta \in \mathbb{R}$ mit

$$\frac{\partial(\alpha f + \beta g)}{\partial x_i}(X_0) = \alpha \frac{\partial f}{\partial x_i}(X_0) + \beta \frac{\partial g}{\partial x_i}(X_0)$$

- (b) **Produktregel** $f \cdot g$

$$\frac{\partial(f \cdot g)}{\partial x_i}(X_0) = \frac{\partial f}{\partial x_i}(X_0) \cdot g(X_0) + f(X_0) \cdot \frac{\partial g}{\partial x_i}(X_0)$$

- (c) **Quotientenregel** Ist $g(X) \neq 0$ für $X \in D$, so ist auch $\frac{f}{g}$ an der Stelle X_0 partiell differenzierbar nach x_i mit

$$\frac{\partial}{\partial x_i} \left(\frac{f}{g} \right) (X_0) = \frac{\frac{\partial f}{\partial x_i}(X_0) \cdot g(X_0) - f(X_0) \cdot \frac{\partial g}{\partial x_i}(X_0)}{g^2(X_0)}$$

Weiterhin gilt:

Satz (Kettenregel): Es sei $D \subset \mathbb{R}^n$ offen, $f : D \rightarrow \mathbb{R}, Y \mapsto f(Y) = f(y_1, \dots, y_n)$ stetig partiell differenzierbar. Weiterhin sei $U \subset \mathbb{R}^m$ ($m \geq 1$) offen,

$$g_i : U \rightarrow \mathbb{R} \quad \text{stetig partiell differenzierbar für } 1 \leq i \leq n$$

und es gelte

$$[g_1(X), \dots, g_n(X)]^T \in D \quad \text{für alle } X \in U.$$

Dann ist die Verkettung

$$\begin{aligned} F : U &\rightarrow \mathbb{R} \\ X &\mapsto f(g_1(X), \dots, g_n(X)) \end{aligned}$$

stetig partiell differenzierbar und für die partiellen Ableitungen gilt:

$$\frac{\partial F}{\partial x_i}(X) = \sum_{j=1}^n \underbrace{\frac{\partial f}{\partial y_j}(g_1(X), \dots, g_n(X))}_{\text{äußere Abl.}} \cdot \underbrace{\frac{\partial g_j}{\partial x_i}(X)}_{\text{innere Abl.}}$$

Beispiele

(i) Polarkoordinaten im \mathbb{R}^2

Für einen Punkt $P = \begin{bmatrix} x \\ y \end{bmatrix}$ bezeichne

$$r(x, y) = \left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\| = \sqrt{x^2 + y^2} \quad (\text{Länge des zugehörigen Ortsvektors})$$

$$\varphi(x, y) = \arctan\left(\frac{y}{x}\right) \quad (\text{Winkel des Ortsvektors zu } P \text{ mit der x-Achse})$$

die Polarkoordinaten von P .

Für einen Punkt $P = \begin{bmatrix} x \\ y \end{bmatrix}$ mit Polarkoordinaten (r, φ) gilt also

$$x = r \cdot \cos \varphi = g_1(r, \varphi)$$

$$y = r \cdot \sin \varphi = g_2(r, \varphi)$$

für $(r, \varphi) \in]0, \infty[\times [0, 2\pi[$.

Ist eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben, so beschreibt die Verkettung

$$F(r, \varphi) = f(r \cos \varphi, r \sin \varphi)$$

also dieselbe Funktion in Polarkoordinaten.

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

$$F(r, \varphi) = e^{-((r \cos \varphi)^2 + (r \sin \varphi)^2)} = e^{-r^2} \quad (!)$$

Mit Hilfe der Kettenregel lassen sich die partiellen Ableitungen von F aus den partiellen Ableitungen von f berechnen. Aus $F(r, \varphi) = f(g_1(r, \varphi), g_2(r, \varphi))$ folgt

$$\begin{aligned}\frac{\partial F}{\partial r}(r, \varphi) &= \frac{\partial f}{\partial x}(r \cdot \cos \varphi, r \sin \varphi) \cos \varphi + \frac{\partial f}{\partial y}(r \cos \varphi, r \sin \varphi) \cdot \sin \varphi \\ \frac{\partial F}{\partial \varphi}(r, \varphi) &= \frac{\partial f}{\partial x}(r \cos \varphi, r \sin \varphi)(-r \sin \varphi) + \frac{\partial f}{\partial y}(r \cos \varphi, r \sin \varphi)(r \cos \varphi)\end{aligned}\quad (4.1)$$

In obigem Beispiel etwa:

$$\begin{aligned}\frac{\partial F}{\partial r} &= -2r \cos \varphi \cdot e^{-r^2} \cdot \cos \varphi - 2r \sin \varphi \cdot e^{-r^2} \cdot \sin \varphi \\ &= -2re^{-r^2} \quad (\text{hätte man auch direkt sehen können!})\end{aligned}$$

(ii) **Ableitung entlang von Kurven** Ist $\gamma : [a, b] \rightarrow D$ stetig differenzierbare Kurve und $f : D \rightarrow \mathbb{R}$ stetig partiell differenzierbar, so ist die Funktion

$$f \circ \gamma : [a, b] \mapsto f(\gamma(t)) = f(\gamma_1(t), \dots, \gamma_n(t))$$

stetig differenzierbar nach t und es gilt

$$(f \circ \gamma)'(t) = \frac{d}{dt}(f \circ \gamma)(t) = \sum_{i=1}^n \underbrace{\frac{\partial f}{\partial y_i}(\gamma(t))}_{\text{äußere}} \cdot \underbrace{\gamma'_i(t)}_{\text{innere}} \quad \text{Ableitung}$$

Beispiel $f(x, y) = e^{-(x^2+y^2)}$ und $\gamma(t) = \begin{bmatrix} r \cos t \\ r \sin t \end{bmatrix}$

$$\begin{aligned}\Rightarrow \frac{d}{dt}(f \circ \gamma)(t) &= -2\gamma_1(t)e^{-(\gamma_1(t)^2+\gamma_2(t)^2)} \cdot \gamma'_1(t) - 2\gamma_2(t)e^{-(\gamma_1(t)^2+\gamma_2(t)^2)} \cdot \gamma'_2(t) \\ &= (2r^2 \cos(t) \sin(t) - 2r^2 \sin(t) \cos(t)) e^{-r^2} = 0\end{aligned}$$

Setzt man speziell für $X_0 \in D$

$$\gamma(t) = X_0 + t \cdot \vec{v}$$

für einen Vektor $\vec{v} \in \mathbb{R}^n$, so ist $\gamma(t) \in D$ für t in einer Umgebung der 0, also etwa für $t \in]-\varepsilon, \varepsilon[$.

Für die Ableitung von f entlang γ erhalten wir in $t = 0$ wegen $\gamma'(t) = \vec{v}$

$$\begin{aligned}(f \circ \gamma)'(0) &= \lim_{h \rightarrow 0} \frac{1}{h} (f(\gamma(h)) - f(\gamma(0))) = \lim_{h \rightarrow 0} \frac{1}{h} (f(X_0 + h\vec{v}) - f(X_0)) \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X_0) \cdot v_i = \text{grad } f(X_0)^T \cdot \vec{v}\end{aligned}$$

Definition Ist $f : D \rightarrow \mathbb{R}$ in einer Umgebung von $X_0 \in D$ stetig partiell differenzierbar und $\vec{v} \in \mathbb{R}^n$ ein Richtungsvektor, so heißt

$$\partial_{\vec{v}} f(X_0) = \lim_{h \rightarrow 0} \frac{1}{h} (f(X_0 + h\vec{v}) - f(X_0)) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0) \cdot v_i = \text{grad } f(X_0)^T \cdot \vec{v}$$

die **Ableitung von f längs \vec{v}** (im Punkt X_0).

Ist $\|\vec{v}\| = 1$, d.h. hat der Richtungsvektor \vec{v} Länge 1, so heißt $\partial_{\vec{v}} f(X_0)$ die **Richtungsableitung von f in Richtung \vec{v}** (im Punkt X_0).

Speziell für $\vec{v} = \vec{e}_i$ erhalten wir:

$$\partial_{\vec{v}} f(X_0) = \frac{\partial f}{\partial x_i}(X_0)$$

- Die Richtungsableitung in Richtung der i -ten Einheitsvektoren stimmt also mit der partiellen Ableitung nach x_i überein.
- **Geometrische Interpretation** des Gradienten: Aus der Cauchy-Schwarzschen Ungleichung folgt für $\|\vec{v}\| = 1$

$$\lim_{h \rightarrow 0} \frac{1}{h} (f(X_0 + h\vec{v}) - f(X_0)) = \partial_{\vec{v}} f(X_0) = \text{grad } f(X_0)^T \vec{v} \leq \|\text{grad } f(X_0)\|$$

und falls $\text{grad } f(X_0) \neq 0$, so gilt die Gleichheit dann, wenn

$$\vec{v} = \frac{\text{grad } f(X_0)}{\|\text{grad } f(X_0)\|}$$

Der **Gradient von f in X_0** zeigt also stets in die **Richtung des steilsten Anstiegs** der Funktion!

Beispiele

(i) $f(x, y) = x^2 + y^2$, also $\text{grad } f(x, y) = 2 \begin{bmatrix} x \\ y \end{bmatrix}$

Für die Richtungsableitungen erhält man $\partial_{\vec{v}} f(x, y) = 2(xv_1 + yv_2)$.

Insbesondere wird $\partial_{\vec{v}} f(x, y)$ maximal für $\vec{v} = \frac{1}{\sqrt{x^2 + y^2}} \begin{bmatrix} x \\ y \end{bmatrix}$ und 0 für $\vec{v} \perp \begin{bmatrix} x \\ y \end{bmatrix}$

(ii) $f(x, y) = x^2 - y^2$, also $\text{grad } f(x, y) = 2 \begin{bmatrix} x \\ -y \end{bmatrix}$

Für die Richtungsableitungen erhält man $\partial_{\vec{v}} f(x, y) = 2(xv_1 - yv_2)$

Insbesondere ist $\partial_{\vec{v}} f(x, y) = 0$ für $\vec{v} = \begin{bmatrix} y \\ x \end{bmatrix}$

4.2 Die totale Ableitung

Ist $f : [a, b] \rightarrow \mathbb{R}$ differenzierbar in $x_0 \in]a, b[$, so beschreibt die Tangentengleichung

$$g : x \mapsto f(x_0) + f'(x_0)(x - x_0)$$

eine **lineare Näherung** von f in x_0 .

Die Güte der Approximation kann man dabei wie folgt beschreiben:

$$f(x) - g(x) = f(x) - (f(x_0) + f'(x_0)(x - x_0)) = \underbrace{\left(\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right)}_{\rightarrow 0 \text{ für } x \rightarrow x_0} (x - x_0) = o(|x - x_0|)$$

Notation (Landau Symbol)

Es sei $D \subset \mathbb{R}^n$ und $X_0 \in D$. Für zwei Funktionen $f, g : D \rightarrow \mathbb{R}$ schreiben wir

$$f(X) = g(X) + o(\|X - X_0\|^k)$$

falls

$$\lim_{X \rightarrow X_0} \frac{f(X) - g(X)}{\|X - X_0\|^k} = 0.$$

Das Landausche Symbol $o(\|X - X_0\|^k)$ besagt also, dass der bei der Approximation von f durch g in der Nähe um X_0 gemachte Fehler von einer kleineren Ordnung als $\|X - X_0\|^k$ ist.

Wir wollen den Gedanken der linearen Approximation von Funktionen einer reellen Veränderlichen auf Funktionen in mehreren Veränderlichen übertragen. Statt Näherungen von f durch eine Geradengleichung

$$g(x) = f(x_0) + f'(x_0)(x - x_0)$$

betrachten wir die Näherung von f durch Abbildungen der Form $f(X_0) + \vec{a}^T \cdot (X - X_0)$ für einen Vektor $\vec{a} \in \mathbb{R}^n$.

Definition Es sei $D \subset \mathbb{R}^n$, $X_0 \in D$ innerer Punkt und $f : D \rightarrow \mathbb{R}$ eine Funktion. f heißt in X_0 **total differenzierbar** (oder **linear approximierbar**), falls ein Vektor $\vec{a} \in \mathbb{R}^n$ existiert, so dass

$$f(X) = f(X_0) + \vec{a}^T \cdot (X - X_0) + o(\|X - X_0\|)$$

für X in einer Umgebung von X_0 gilt.

Ist f in $X_0 \in D$ total differenzierbar, so gilt

- f ist in X_0 stetig
- f ist in X_0 partiell differenzierbar und es gilt

$$\vec{a} = \text{grad } f(X_0) = \left[\frac{\partial f}{\partial x_1}(X_0), \dots, \frac{\partial f}{\partial x_n}(X_0) \right]^T$$

In den Komponenten des Vektors \vec{a} stehen also die partiellen Ableitungen von f in X_0 . Insbesondere ist der Vektor \vec{a} eindeutig bestimmt.

Aus der partiellen Differenzierbarkeit folgt im allgemeinen nicht die totale Differenzierbarkeit. Jedoch gilt, falls $D \subset \mathbb{R}^n$ offen:

$f : D \rightarrow \mathbb{R}$ stetig partiell differenzierbar

$\implies f$ total differenzierbar (in allen Punkten $X_0 \in D$)

Beispiele

(i) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = 2 - x^2 - y^2$, also $\text{grad} f(x, y) = -2[x, y]^T$

Da f stetig partiell differenzierbar, ist f insbesondere total differenzierbar und die lineare Näherung von f im Punkte $[x_0, y_0]^T \in \mathbb{R}^2$ ist

$$f(x, y) = 2 - x_0^2 - y_0^2 - 2x_0(x - x_0) - 2y_0(y - y_0) + o\left(\left\| \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \right\|\right)$$

etwa in $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}$ gilt

$$f(x, y) = 0 - 2\sqrt{2}(x - \sqrt{2}) + o\left(\sqrt{(x - \sqrt{2})^2 + y^2}\right)$$

4.3 Die Taylorformel für Funktionen in n Variablen

Die Taylorformel für die Approximation einer differenzierbaren Funktion

$$f :]a, b[\rightarrow \mathbb{R}$$

durch Polynome der Form

$$f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(m)}(x_0)}{m!}(x - x_0)^m$$

lässt sich auf differenzierbare Funktionen in mehreren Variablen verallgemeinern:

Satz Es sei $D \subset \mathbb{R}^n$ offen, $f : D \rightarrow \mathbb{R}$ $(m + 1)$ -mal stetig partiell differenzierbar, $X_0 \in D$, $\vec{v} \in \mathbb{R}^n$, so dass die Verbindungsstrecke

$$\{X_0 + t\vec{v} : t \in [0, 1]\}$$

zwischen X_0 und $X_0 + \vec{v}$ ganz in D verläuft. Dann gilt die **Taylorformel**

$$f(X_0 + \vec{v}) = f(X_0) + \partial_{\vec{v}} f(X_0) + \frac{1}{2!} \partial_{\vec{v}}^2 f(X_0) + \dots + \frac{1}{m!} \partial_{\vec{v}}^m f(X_0) + R_{m+1}(X_0, \vec{v})$$

mit dem Restglied

$$R_{m+1}(X_0, \vec{v}) = \frac{1}{(m+1)!} \partial_{\vec{v}}^{m+1} f(X_0 + \xi \vec{v})$$

für ein $\xi \in [0, 1]$

Mit der Substitution $\vec{v} = X - X_0$ erhält man aus

$$f(X) = f(X_0) + \partial_{\vec{v}} f(X_0) + \dots + \frac{1}{m!} \partial_{\vec{v}}^m f(X_0)$$

ein Polynom $p(X)$ vom Grade m in den Variablen x_1, \dots, x_n . Wie im Falle $n = 1$ heißt p **m -tes Taylorpolynom von f im Entwicklungspunkt X_0** . Mit diesem Polynom gilt dann

$$f(X) = p(X) + o(\|X - X_0\|^m)$$

Die wichtigen Spezialfälle $m = 1, 2$

$m = 1$: **Lineare Approximation** (f zweimal stetig partiell differenzierbar)

$$\begin{aligned} f(X_0 + \vec{v}) &= f(X_0) + \partial_{\vec{v}} f(X_0) + R_2(X_0, \vec{v}) \\ &= f(X_0) + \text{grad } f(X_0)^T \cdot \vec{v} + R_2(X_0, \vec{v}) \end{aligned}$$

mit $R_2(X_0, \vec{v}) = \frac{1}{2!} \partial_{\vec{v}}^2 f(X_0 + \xi \vec{v})$ für ein $\xi \in [0, 1]$.

Mit der Substitution $\vec{v} = X - X_0$ erhält man hieraus

$$f(X) = \underbrace{f(X_0) + \text{grad } f(X_0)^T \cdot (X - X_0)}_{p(X)} + o(\|X - X_0\|)$$

$m = 2$: **Quadratische Approximation** (f dreimal stetig partiell differenzierbar)

$$\begin{aligned} f(X_0 + \vec{v}) &= f(X_0) + \partial_{\vec{v}} f(X_0) + \frac{1}{2!} \partial_{\vec{v}}^2 f(X_0) + R_3(X_0, \vec{v}) \\ &= f(X_0) + \text{grad } f(X_0)^T \cdot \vec{v} + \frac{1}{2} \vec{v}^T \cdot H_f(X_0) \vec{v} + R_3(X_0, \vec{v}) \end{aligned} \quad (4.2)$$

mit $R_3(X_0, \vec{v}) = \frac{1}{3!} \partial_{\vec{v}}^3 f(X_0 + \xi \vec{v})$ für ein $\xi \in [0, 1]$.

Hierbei ist

$$H_f(X_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(X_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(X_0) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(X_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(X_0) & \frac{\partial^2 f}{\partial x_2^2}(X_0) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(X_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(X_0) & \frac{\partial^2 f}{\partial x_n \partial x_2}(X_0) & \dots & \frac{\partial^2 f}{\partial x_n^2}(X_0) \end{bmatrix} = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(X_0) \right]$$

die $n \times n$ -Matrix der partiellen Ableitungen zweiter Ordnung von f in X_0 .

Die Matrix $H_f(X_0)$ heißt **Hesse-Matrix von f in X_0** . Da die partiellen Ableitungen vertauscht werden können

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(X_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(X_0) \quad \text{für } 1 \leq i, j \leq n$$

ist $H_f(X_0)$ eine **symmetrische Matrix**, d.h. es gilt $H_f(X_0)^T = H_f(X_0)$.

Mit der Substitution $\vec{v} = X - X_0$ erhält man aus (4.2)

$$f(X) = \underbrace{f(X_0) + \text{grad } f(X_0)^T \cdot (X - X_0) + \frac{1}{2} (X - X_0)^T \cdot H_f(X_0) (X - X_0)}_{=p(X)} + o(\|X - X_0\|^2)$$

Beispiele

(i) $f(x, y) = e^{xy}$ ist 3-mal stetig partiell differenzierbar

$$\text{grad } f(x_0, y_0) = e^{x_0 y_0} \begin{bmatrix} y_0 \\ x_0 \end{bmatrix}$$

$$H_f(x_0, y_0) = e^{x_0 y_0} \begin{bmatrix} y_0^2 & 1 + x_0 y_0 \\ 1 + x_0 y_0 & x_0^2 \end{bmatrix}$$

Also ist

$$\begin{aligned} & e^{x_0 y_0} + e^{x_0 y_0} \begin{bmatrix} y_0 \\ x_0 \end{bmatrix}^T \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \frac{1}{2} e^{x_0 y_0} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}^T \begin{bmatrix} y_0^2 & 1 + x_0 y_0 \\ 1 + x_0 y_0 & x_0^2 \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ &= e^{x_0 y_0} + e^{x_0 y_0} (y_0(x - x_0) + x_0(y - y_0)) \\ & \quad + \frac{1}{2} e^{x_0 y_0} (y_0^2(x - x_0)^2 + 2(1 + x_0 y_0)(x - x_0)(y - y_0) + x_0^2(y - y_0)^2) \end{aligned}$$

die quadratische Näherung für f im Punkte $\begin{bmatrix} y_0 \\ x_0 \end{bmatrix}$

(ii) $f(x, y) = x^2 + 4xy + 8y^2 + 3x + 5y + 1$

$$\text{grad } f(x_0, y_0) = \begin{bmatrix} 2x + 4y + 3 \\ 4x + 16y + 5 \end{bmatrix}$$

$$H_f(x_0, y_0) = \begin{bmatrix} 2 & 4 \\ 4 & 16 \end{bmatrix}$$

also ist die quadratische Näherung in $[x_0, y_0]^T$ gegeben durch

$$\begin{aligned} & (x_0^2 + 4x_0 y_0 + 8y_0^2 + 3x_0 + 5y_0 + 1) \\ & \quad + ((2x_0 + 4y_0 + 3)(x - x_0) + (4x_0 + 16y_0 + 5)(y - y_0)) \\ & \quad + \frac{1}{2} (2(x - x_0)^2 + 8(x - x_0)(y - y_0) + 16(y - y_0)^2) \\ &= \dots = 1 + 3x + 5y + x^2 + 4xy + 8y^2 = f(x, y) \quad (!) \end{aligned}$$

d.h., wie im Falle $n = 1$ stimmt das Taylorpolynom zweiter Ordnung mit f überein, wenn f ein Polynom vom Grad ≤ 2 ist. Analoges gilt für Polynome höheren Grades.

Anwendung: Bestimmung von Extremalstellen

Wie im Falle einer Funktion einer reellen Variablen definieren wir:

Eine Funktion $f : D \rightarrow \mathbb{R}$ hat in $x_0 \in D$ ein **lokales Maximum (Minimum)**, falls ein $\varepsilon > 0$ existiert mit

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

für alle $x \in U_\varepsilon(x_0)$.

Ist $f(x_0) \geq f(x)$ (bzw. $f(x_0) \leq f(x)$) für alle $x \in D$, so heißt x_0 **absolutes Maximum (Minimum)**.

Notwendiges Kriterium

Wie im Falle einer Funktion einer reellen Variablen gilt: Ist $f : D \rightarrow \mathbb{R}$ partiell differenzierbar und X_0 lokales Extremum (also X_0 eine lokales Maximum oder lokales Minimum), so ist $\text{grad } f(X_0) = 0$. Man nennt die Nullstellen des Gradienten von f , also die Stellen X_0 mit $\text{grad } f(X_0) = 0$ **kritische Stellen**.

Möchte man also (lokale) Maxima (bzw. (lokale) Minima) einer Funktion bestimmen, so sucht man zunächst die Menge der kritischen Stellen. Diese kritischen Stellen sind die Kandidaten für lokale Extrema. Man versucht dann mit den folgenden hinreichenden Kriterien für lokale Maxima und Minima zu entscheiden, ob es sich bei einer gegebenen kritischen Stelle um ein lokales Maximum oder um ein lokales Minimum handelt. Zur Bestimmung aller lokalen Extrema muss man schließlich gegebenenfalls auch noch das Verhalten der Funktion am Rand des Definitionsbereiches untersuchen.

Für hinreichende Kriterien benötigen wir wie im Falle von Funktionen einer reellen Veränderlichen die höheren Ableitungen von f . Bevor wir ein hinreichendes Kriterium formulieren können, benötigen wir noch Informationen über positiv definite Matrizen.

Einschub: Positiv definite Matrizen

Definition Es sei $A = [a_{ij}]$ eine **symmetrische** $n \times n$ -Matrix (also $A^T = A$, d.h. $a_{ij} = a_{ji}$ für $1 \leq i, j \leq n$). A heißt **positiv definit**, falls gilt

$$\vec{x}^T A \vec{x} > 0 \quad \text{für alle } \vec{x} \in \mathbb{R}^n, \vec{x} \neq 0.$$

A heißt **negativ definit**, falls $-A$ positiv definit ist, und **indefinit**, falls $\vec{x}, \vec{y} \in \mathbb{R}^n$ existieren mit

$$\vec{x}^T A \vec{x} > 0 \quad \text{und} \quad \vec{y}^T A \vec{y} < 0$$

Beispiele

(i) $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ist positiv definit, denn

$$\begin{bmatrix} x \\ y \end{bmatrix}^T A \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + y^2 = \left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|^2 > 0 \quad \text{für } \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2, \begin{bmatrix} x \\ y \end{bmatrix} \neq 0$$

(ii) $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ ist indefinit, denn

$$\vec{e}_1^T \underbrace{A \vec{e}_1}_{= -\vec{e}_1} = -\|\vec{e}_1\|^2 = -1$$

und

$$\vec{e}_2^T \underbrace{A \vec{e}_2}_{= -\vec{e}_2} = -\|\vec{e}_2\|^2 = +1$$

Bemerkung 4.3

(i) Wir haben in Mathematik I für MB gesehen, dass für eine $n \times n$ -Matrix A alle Eigenwerte von A reell sind. Weiter gilt dann:

$$\begin{aligned} A \text{ positiv definit} &\Leftrightarrow \text{alle Eigenwerte von } A \text{ positiv} \\ A \text{ negativ definit} &\Leftrightarrow \text{alle Eigenwerte von } A \text{ negativ} \\ A \text{ indefinit} &\Leftrightarrow A \text{ besitzt mindestens einen positiven} \\ &\quad \text{und mindestens einen negativen Eigenwert} \end{aligned}$$

(ii) Das folgende Kriterium charakterisiert positive Definitheit einer Matrix mit Hilfe von Determinanten. Es gilt:

A ist genau dann positiv definit, wenn für $k = 1, 2, \dots, n$ gilt

$$\det \begin{bmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{bmatrix} > 0$$

(iii) Für 2×2 -Matrizen ist dieses Kriterium besonders einfach anzuwenden: die Matrix

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix} \quad \text{ist}$$

ist

$$\begin{aligned} \text{positiv definit} &\Leftrightarrow a > 0 \quad \text{und} \quad \det A = ad - b^2 > 0 \\ \text{negativ definit} &\Leftrightarrow a < 0 \quad \text{und} \quad \det A = ad - b^2 > 0 \\ \text{indefinit} &\Leftrightarrow \det A = ad - b^2 < 0 \end{aligned}$$

Kehren wir zurück zu hinreichenden Kriterien für lokale Extrema einer Funktion in mehreren reellen Veränderlichen. Mit Hilfe der Aussagen über positiv definite Matrizen erhält man nun:

Satz Es sei $f : D \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar und $X_0 \in D$ mit $\text{grad } f(X_0) = 0$. Es sei $H_f(X_0) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(X_0) \right]$ die Hesse-Matrix von f in X_0 . Dann gilt:

- $H_f(X_0)$ positiv definit $\Rightarrow X_0$ ist lokales Minimum
- $H_f(X_0)$ negativ definit $\Rightarrow X_0$ ist lokales Maximum
- $H_f(X_0)$ indefinit $\Rightarrow X_0$ ist kein lokales Extremum (sondern Sattelpunkt)

Spezialfall $n = 2$

Im Spezialfall $n = 2$ erhält man aufgrund von Bemerkung 4.3 (iii) insbesondere für $f(x, y)$:

$$H_f(x_0, y_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\ \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) & \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{bmatrix}$$

und ist

$$\Delta := \det H_f(x_0, y_0) = \frac{\partial^2 f}{\partial^2 x}(x_0, y_0) \cdot \frac{\partial^2 f}{\partial^2 y}(x_0, y_0) - \left(\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \right)^2$$

so gilt:

- $\Delta > 0$ und $\frac{\partial^2 f}{\partial^2 x}(x_0, y_0) > 0 \Rightarrow \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ ist lokales Minimum
- $\Delta > 0$ und $\frac{\partial^2 f}{\partial^2 x}(x_0, y_0) < 0 \Rightarrow \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ ist lokales Maximum
- $\Delta < 0 \Rightarrow \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ ist kein lokales Extremum (sondern Sattelpunkt).

Beispiele Es sei $f(x, y) = -x^4 - y^4 + 2x^2 + 2y^2$. Dann gilt

$$\text{grad } f(x, y) = 4 \begin{bmatrix} -x^3 + x \\ -y^3 + y \end{bmatrix}$$

und dieser besitzt neun Nullstellen:

$$A_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, A_3 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, A_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, A_5 = \begin{bmatrix} 0 \\ -1 \end{bmatrix},$$

$$A_6 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, A_7 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, A_8 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, A_9 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

Für die Hesse-Matrix

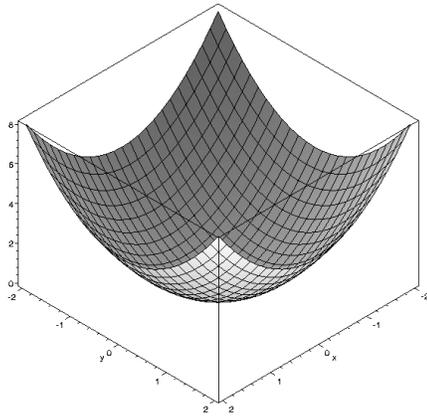
$$H_f(x, y) = 4 \begin{bmatrix} -3x^2 + 1 & 0 \\ 0 & -3y^2 + 1 \end{bmatrix}$$

gilt in diesen kritischen Punkten:

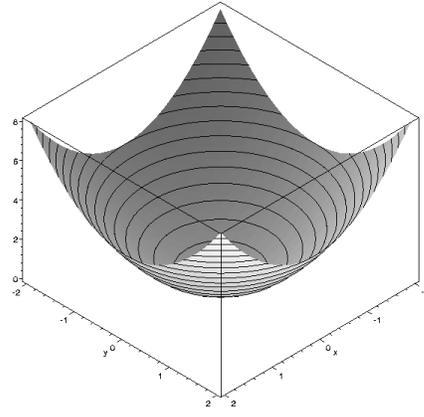
$$H_f(A_1) = 4 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ positiv definit} \Rightarrow A_1 \text{ lokales Minimum}$$

$$H_f(A_2) = 4 \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix} \text{ indefinit} \Rightarrow A_2 \text{ Sattelpunkt (ebenso } A_3, A_4, A_5)$$

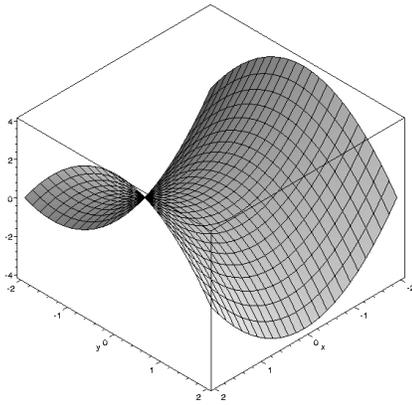
$$H_f(A_6) = 4 \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} \text{ negativ definit} \Rightarrow A_6 \text{ lokales Maximum (ebenso } A_7, A_8, A_9)$$



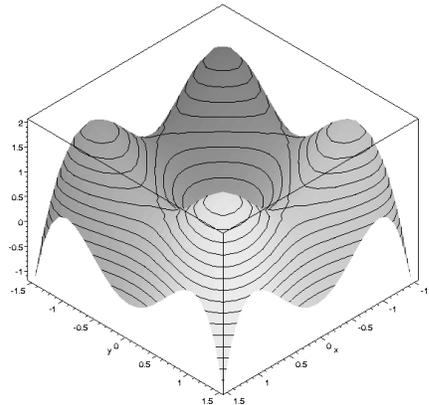
(i) $f(x, y) = x^2 + y^2$ mit Schnittkurven



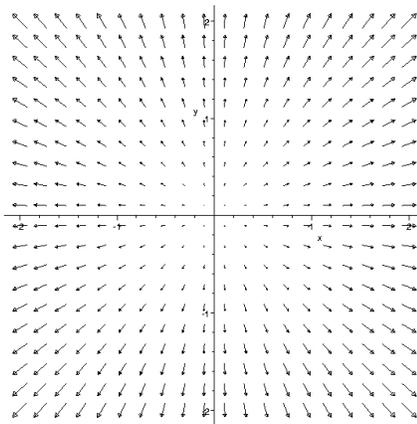
(ii) $f(x, y) = x^2 + y^2$ mit Höhenlinien



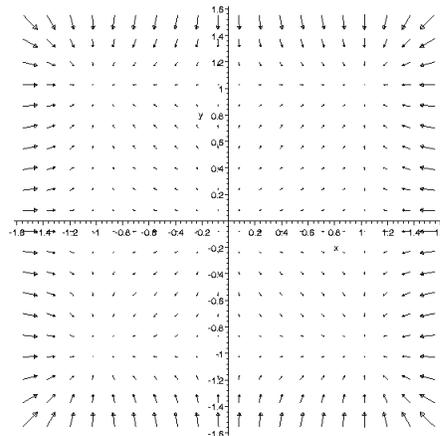
(iii) $f(x, y) = x^2 - y^2$



(iv) $f(x, y) = -x^4 - y^4 + 2x^2 + 2y^2$



(v) Gradientenfeld zu (i)



(vi) Gradientenfeld zu (iv)

5 Implizite Funktionen und Extrema mit Nebendbedingungen

5.1 Vektorfelder

Bisher haben wir nur reellwertige Abbildungen $f : D \rightarrow \mathbb{R}$ betrachtet. Nun wollen wir auch Abbildungen

$$F : D \rightarrow \mathbb{R}^m \quad \text{für } D \subset \mathbb{R}^n, n, m \geq 1$$

betrachten. Eine solche Abbildung nennen wir **Vektorfeld**.

Schreiben wir $F = [F_1, \dots, F_m]^T$, so erhalten wir zu F die **Komponentenfunktionen** des Vektorfeldes

$$F_i : D \rightarrow \mathbb{R}, \quad 1 \leq i \leq m.$$

- **n=1** Ein Vektorfeld $F : D \subset \mathbb{R} \rightarrow \mathbb{R}^m$ ist nichts weiteres als eine Kurve im \mathbb{R}^m .
- **m=1** Ein Vektorfeld $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ist nichts weiteres als eine reellwertige Funktion in n Veränderlichen.

Beispiel Strömung einer zähen Flüssigkeit durch ein Rohr mit Radius r

Es bezeichne $D \subset \mathbb{R}^3$ ein Rohr mit Radius r dessen Mittelpunkt durch die x -Achse gegeben ist. Dann beschreibt das Vektorfeld

$$v : D \rightarrow \mathbb{R}^3 \\ [x, y, z]^T \mapsto c[0, r^2 - y^2 - z^2, 0]^T \quad y^2 + z^2 \leq r^2, c > 0.$$

das Geschwindigkeitsfeld einer zähen Flüssigkeit durch D .

Definition Ein Vektorfeld $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt

- **stetig (in X_0)**, falls alle Komponentenfunktionen F_i stetig (in X_0) sind,
- **partiell differenzierbar (in X_0)**, falls alle Komponentenfunktionen F_i partiell differenzierbar (in X_0) sind,
- **total differenzierbar in X_0 oder linear approximierbar**, falls eine $m \times n$ -Matrix A existiert, so dass

$$F(X) = F(X_0) + A(X - X_0) + o(\|X - X_0\|)$$

für X in einer Umgebung von X_0 .

Ist F total differenzierbar in X_0 , so ist F stetig in X_0 und partiell differenzierbar in X_0 und es gilt

$$A = \left[\frac{\partial F_i}{\partial x_j}(X_0) \right] \quad (= J_F(X_0) \text{ siehe unten}).$$

Fassen wir alle partiellen Ableitungen eines Vektorfeldes F im Punkt X_0 in einer Matrix zusammen, so erhält man die **Funktionalmatrix von F in X_0** :

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_1}(X_0) & \frac{\partial F_1}{\partial x_2}(X_0) & \dots & \frac{\partial F_1}{\partial x_n}(X_0) \\ \frac{\partial F_2}{\partial x_1}(X_0) & \frac{\partial F_2}{\partial x_2}(X_0) & \dots & \frac{\partial F_2}{\partial x_n}(X_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1}(X_0) & \frac{\partial F_m}{\partial x_2}(X_0) & \dots & \frac{\partial F_m}{\partial x_n}(X_0) \end{pmatrix} = \left[\frac{\partial F_i}{\partial x_j}(X_0) \right]$$

$J_F(X_0)$ ist also eine $m \times n$ -Matrix. In den Zeilen stehen die partiellen Ableitungen der Komponenten des Vektorfeldes. Ist $m = n$, so ist $J_F(X_0)$ also eine quadratische Matrix. Ihre Determinante $\det(J_F(X_0))$ heißt **Funktionaldeterminante von F in X_0** .

Beispiel Polarkoordinaten im \mathbb{R}^2

$$G :]0, \infty[\times]0, 2\pi[\rightarrow \mathbb{R}^2 \\ [r, \varphi]^T \mapsto [r \cos \varphi, r \sin \varphi]^T$$

also $G_1(r, \varphi) = r \cos \varphi$, $G_2(r, \varphi) = r \sin \varphi$ und

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

und $\det(J_G(r, \varphi)) = r((\cos \varphi)^2 + (\sin \varphi)^2) = r$.

Mit Hilfe der Funktionalmatrix J_G lässt sich die Kettenregel (4.1) nun wie folgt umschreiben: Ist $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ stetig partiell differenzierbar, so ist

$$F = f \circ G :]0, \infty[\times]0, 2\pi[\rightarrow \mathbb{R}$$

stetig partiell differenzierbar und es gilt

$$J_F(r, \varphi) = J_f(G(r, \varphi)) \cdot J_G(r, \varphi).$$

\cdot bezeichnet hierbei das Matrizenprodukt der beiden Funktionalmatrizen. In dieser Form lässt sich die Kettenregel auf allgemeine Vektorfelder verallgemeinern.

Satz (Kettenregel für Vektorfelder) Es sei $D \subset \mathbb{R}^n$ offen, $U \subset \mathbb{R}^m$ offen,

$$G : D \rightarrow U \quad \text{und} \quad F : U \rightarrow \mathbb{R}^k$$

stetig partiell differenzierbare Vektorfelder. Dann ist die Verkettung

$$H = F \circ G : D \rightarrow \mathbb{R}^k \\ X \mapsto F(G(X))$$

stetig (partiell) differenzierbar und es gilt für $X_0 \in D$

$$J_H(X_0) = J_{F \circ G}(X_0) = \underbrace{J_F(G(X_0))}_{\text{äußere}} \cdot \underbrace{J_G(X_0)}_{\text{innere}} \\ \text{Ableitung}$$

Die Funktionalmatrix der Verkettung berechnet sich also als Matrizenprodukt der beiden Funktionalmatrizen $J_F(G(X_0))$ mit $J_G(X_0)$. Da $J_F(G(X_0))$ eine $k \times m$ -Matrix und $J_G(X_0)$ eine $m \times n$ -Matrix, ist das Matrizenprodukt eine $k \times n$ -Matrix, wie es $J_H(X_0)$ auch sein muss.

5.2 Implizite Funktionen

Um eine Funktion $g : D \subset \mathbb{R} \rightarrow \mathbb{R}$ mit den Methoden der Differentialrechnung zu diskutieren, bedarf es eines expliziten Ausdrucks für die Funktion $g(x) = y$. Mitunter ist der Zusammenhang zwischen x und y nur durch eine Gleichung

$$f(x, y) = 0 \quad (5.1)$$

gegeben, und man hat zu gegebenem x nach $y = g(x)$ aufzulösen. Gesucht ist also eine Funktion g mit

$$f(x, g(x)) = 0. \quad (5.2)$$

Man sagt dann, dass g durch die Gleichung (5.1) **implizit** definiert ist.

Beispiele

(i) $f(x, y) = 3x + 2y - 4$

Die Gleichung $f(x, y) = 0$ führt auf die Gleichung $3x + 2y = 4$, die wir nach y auflösen können:

$$y = -\frac{3}{2}x + 2, \quad \text{also } g(x) = -\frac{3}{2}x + 2.$$

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

Die Gleichung $f(x, y) = 0$ führt auf die Gleichung $y^2 = 1 - x^2$ mit den Lösungen $y = \pm\sqrt{1 - x^2}$, $|x| \leq 1$. Dies führt auf die beiden implizit definierten Funktionen

$$g_1(x) = \sqrt{1 - x^2}, \quad g_2(x) = -\sqrt{1 - x^2}$$

Wenn wir nun einen Punkt $[x_0, y_0]^T$ mit $f(x_0, y_0) = 0$ festhalten, so können wir, falls $y_0 \neq 0$, die Gleichung $f(x, y) = 0$ in einer Umgebung von $[x_0, y_0]^T$ eindeutig auflösen, denn in einer Umgebung wird das Vorzeichen von y nicht wechseln.

Satz Es sei $f : D =]a, b[\times]c, d[\rightarrow \mathbb{R}$ stetig partiell differenzierbar, $[x_0, y_0]^T \in D$ ein Punkt mit $f(x_0, y_0) = 0$ und $f_y(x_0, y_0) \neq 0$. Dann gibt es offene Intervalle $I \subset]a, b[$ mit $x_0 \in I$ und $J \subset]c, d[$ mit $y_0 \in J$, so dass $f_y(x, y) \neq 0$ für alle $[x, y]^T \in I \times J$, und es gibt eine stetig differenzierbare Funktion

$$g : I \rightarrow \mathbb{R} \quad \text{mit } g(I) \subset]c, d[, g(x_0) = y_0,$$

so dass $f(x, g(x)) = 0$ für alle $x \in I$. Für die Ableitung von g gilt

$$g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))}, x \in I. \quad (5.3)$$

Beispiel Es sei $f(x, y) = y^2 + \sin x - 4$

Dann ist $[0, 2]^T$ Lösung von $f(x, y) = 0$ und $f_y(0, 2) = 4$, also ist die Gleichung

$$y^2 + \sin x - 4 = 0$$

in 0 lokal eindeutig nach y auflösbar. Für die durch die eindeutig bestimmte Lösung implizit definierte Funktion g gilt: $g(0) = 2$ und

$$g^2(x) + \sin x - 4 = 0$$

also $g(x) = \sqrt{4 - \sin x}$. Für die Ableitung gilt

$$g'(x) = -\frac{\cos x}{2g(x)} = -\frac{\cos x}{2\sqrt{4 - \sin x}}$$

was man durch direktes Ableiten von g bestätigen kann.

Bemerkung Die Formel (5.3) für die Ableitung der implizit definierten Funktion g folgt aus der Kettenregel, denn aus

$$f(x, g(x)) = 0$$

folgt durch Ableiten nach x :

$$f_x(x, g(x)) + f_y(x, g(x)) g'(x) = 0$$

also

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

Der Satz über implizite Funktionen gilt auch allgemeiner:

Satz (Hauptsatz über implizite Funktionen) Es seien $U \subset \mathbb{R}^n$ offen, $V \subset \mathbb{R}^m$ offen und

$$F : U \times V \rightarrow \mathbb{R}^m, \quad [X, Y]^T \mapsto F(X, Y)$$

stetig differenzierbar. Weiterhin sei $[X_0, Y_0]^T \in U \times V$ eine Lösung von $F(X, Y) = 0$ und die $m \times m$ -Matrix

$$\frac{\partial F}{\partial Y}(X_0, Y_0) := \left[\frac{\partial F_i}{\partial y_j}(X_0) \right]$$

sei invertierbar (also $\det\left(\frac{\partial F}{\partial Y}(X_0, Y_0)\right) \neq 0$). Dann gibt es offene Mengen $U_0 \subset U$ mit $X_0 \in U_0$ und $V_0 \subset V$ mit $Y_0 \in V_0$, so dass $\frac{\partial F}{\partial Y}(X, Y)$ invertierbar für alle $[X, Y]^T \in U_0 \times V_0$, und es gibt ein stetig differenzierbares Vektorfeld

$$G : U_0 \rightarrow V_0 \quad \text{mit} \quad G(X_0) = Y_0$$

und

$$F(X, G(X)) = 0 \quad \text{für alle} \quad X \in U_0$$

sowie

$$J_G(X) = -\frac{\partial F}{\partial Y}(X, G(X))^{-1} \cdot \frac{\partial F}{\partial X}(X, G(X)).$$

5.3 Extrema unter Nebenbedingungen

Gegeben sei eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Mitunter ist es notwendig, ein Extremum von f nicht auf ganz \mathbb{R}^n zu bestimmen, sondern nur auf einer Teilmenge, die mit Hilfe einer weiteren Funktion $g : \mathbb{R}^n \rightarrow \mathbb{R}$ durch eine Gleichung $g(X) = 0$ implizit definiert ist.

Definition Ein Punkt X_0 mit $g(X_0) = 0$ heißt **relatives Maximum** (bzw. **Minimum**) unter der **Nebenbedingung** $g(X) = 0$, falls

$$f(X_0) \geq f(X) \quad (\text{ bzw. } f(X_0) \leq f(X))$$

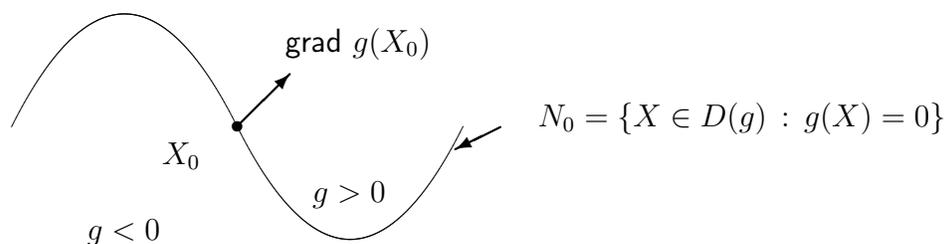
für alle $X \in U_\varepsilon(X_0) \cap \{X \in \mathbb{R}^n : g(X) = 0\}$.

Der folgende Satz liefert eine notwendige Bedingung die zum Auffinden relativer Extrema unter Nebenbedingungen nützlich ist.

Satz Sind f, g auf einer Umgebung $U_\varepsilon(X_0)$ von X_0 stetig differenzierbar, hat f in X_0 ein relatives Extremum unter der Nebenbedingung $g(X) = 0$ und ist $\text{grad } g(X_0) \neq 0$, so gibt es ein $\lambda \in \mathbb{R}$ mit

$$\begin{aligned} \text{grad } f(X_0) + \lambda \text{ grad } g(X_0) &= 0, \\ \text{d.h.} \quad f_{x_i}(X_0) + \lambda g_{x_i}(X_0) &= 0 \quad \text{für } i = 1, \dots, n. \end{aligned} \quad (5.4)$$

Geometrische Interpretation von (5.4)



Der Gradient $\text{grad } g(X_0)$ zeigt stets in eine Richtung, die aus $N_0 := \{X \in D(g) : g(X) = 0\}$ **hinausführt**, und zwar genauer in die Menge $g > 0$ **hinein**. Kurven, die ganz in N_0 verlaufen, haben Tangentialvektoren, die senkrecht zu $\text{grad } g(X_0)$ sind, denn: Ist

$$\gamma :] - \varepsilon, \varepsilon[\rightarrow N_0$$

stetig differenzierbare Kurve mit $\gamma(0) = X_0$, so folgt aus $g(\gamma(t)) = 0$ für alle $t \in] - \varepsilon, \varepsilon[$ mit Hilfe der Kettenregel

$$0 = \frac{d}{dt} g(\gamma(t)) = \text{grad } g(X_0)^T \cdot \gamma'(t)$$

und somit insbesondere für $t = 0$

$$\text{grad } g(X_0)^T \cdot \gamma'(0) = 0.$$

Beispiel 5.1 Gegeben sei die Funktion

$$g(x, y, z) = x^2 + y^2 + z^2 - 1.$$

Dann ist $N_0 = \{[x, y, z]^T \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ also die Oberfläche der Einheitskugel S^1 im \mathbb{R}^3 . Im folgenden bezeichne X_0 den Nordpol, also $X_0 = [0, 0, 1]^T$. Dann steht $\text{grad } g(X_0) = 2[0, 0, 1]^T$ senkrecht zur Tangentialebene

$$E : 0x + 0y + 1z = 1$$

an S^1 im Punkt X_0 . Gegeben sei etwa die Kurve

$$\gamma(t) = [\sin t, 0, \cos t]^T, \quad t \in]-\pi, \pi[.$$

Dann gilt $\gamma(0) = X_0$ und $\gamma'(0) = [\cos 0, 0, \sin 0]^T = [1, 0, 0]^T$.

Es sei nun f eine Funktion, die auf einer Umgebung von X_0 definiert ist. Der Gradient $\text{grad } f(X_0)$ zeigt immer in die Richtung des steilsten Anstieges von f . Ist also X_0 relatives Extremum, so wird der Gradient keinen Richtungsanteil haben, der tangential zu N_0 verläuft. Damit aber muss $\text{grad } f(X_0)$ in dieselbe Richtung wie $\text{grad } g(X_0)$ oder $-\text{grad } g(X_0)$ zeigen. Folglich müssen $\text{grad } g(X_0)$ und $\text{grad } f(X_0)$ **linear abhängig** sein, d.h. es gibt ein $\lambda \in \mathbb{R}$ mit

$$\text{grad } f(X_0) + \lambda \text{grad } g(X_0) = 0.$$

Beispiel 5.2 In der Situation des Beispiels 5.1 sei

$$f(x, y, z) = \frac{1}{2}x^4 + \frac{1}{2}y^4 + \frac{1}{2}z^4$$

gegeben. Für $z > 0$ ist f monoton wachsend in z . Daher steht zu vermuten, dass f in X_0 sein Maximum unter der Nebenbedingung $g(X) = 0$ annimmt. In der Tat gilt

$$\text{grad } f(x, y, z) = 2[x^3, y^3, z^3]^T,$$

also

$$\text{grad } f(X_0) = 2[0, 0, 1]^T = \text{grad } g(X_0).$$

Damit ist die **notwendige Bedingung** für das Vorliegen eines relativen Extremums in X_0 erfüllt. Um zu prüfen, ob X_0 ein Maximum ist, müssen nun alle relativen Extrema untersucht werden. In diesem Falle finden wir alle relativen Extrema unter den Lösungen x, y, z und λ des nichtlinearen Gleichungssystems

$$\text{grad } f(x, y, z) + \lambda \text{grad } g(x, y, z) = 0 \text{ mit } g(x, y, z) = 0,$$

also

$$2 \begin{bmatrix} x^3 \\ y^3 \\ z^3 \end{bmatrix} + 2\lambda \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \quad \text{mit} \quad x^2 + y^2 + z^2 = 1.$$

Dies führt auf

$$\begin{aligned} x^3 + \lambda x &= 0 & \implies & x(x^2 + \lambda) = 0 \\ y^3 + \lambda y &= 0 & & y(y^2 + \lambda) = 0 \\ z^3 + \lambda z &= 0 & & z(z^2 + \lambda) = 0. \end{aligned}$$

1. Fall $x \neq 0$, also $\lambda = -x^2$ und damit

$$y = 0 \text{ oder } y = \pm x$$

$$z = 0 \text{ oder } z = \pm x$$

(I) $y = z = 0$ führt wegen der Nebenbedingung $g(x, y, z) = 0$ auf $x = \pm 1$, also

$$f(0, 0, \pm 1) = \frac{1}{2} = f(X_0).$$

(II) $y = \pm x, z = 0$ führt entsprechend auf $x = \pm \frac{1}{\sqrt{2}}, y = \pm \frac{1}{\sqrt{2}}$, also

$$f\left(\pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}, 0\right) = \frac{1}{4}.$$

(III) $y = \pm x, z = \pm x$ führt auf $x = \pm \frac{1}{\sqrt{3}}, y = \pm \frac{1}{\sqrt{3}}, z = \pm \frac{1}{\sqrt{3}}$, also

$$f\left(\pm \frac{1}{\sqrt{3}}, \pm \frac{1}{\sqrt{3}}, \pm \frac{1}{\sqrt{3}}\right) = \frac{1}{6}.$$

Die übrigen Fälle $y \neq 0$ und $z \neq 0$ sind analog zu behandeln.

Insgesamt gibt es

- 6 Punkte vom Typ (I), die allesamt relative Maxima sind. Insbesondere ist also der Nordpol X_0 ein relatives Maximum.
- 12 Punkte vom Typ (II), die allesamt Sattelpunkte sind.
- 8 Punkte vom Typ (III), die allesamt relative Minima sind.

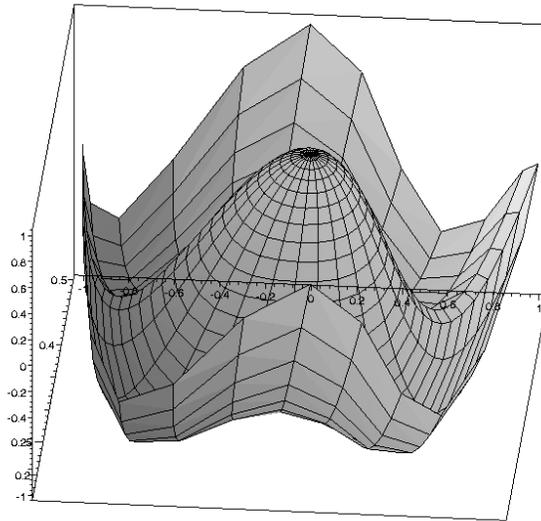
Für einen Punkt $[x, y, z]^T$ mit $z > 0$ können wir schreiben

$$z = z(x, y) = \sqrt{1 - x^2 - y^2}$$

und damit erhält man

$$f(x, y, z) = f(x, y, z(x, y)) = \frac{1}{2}x^4 + \frac{1}{2}y^4 + \frac{1}{2}(1 - x^2 - y^2)^2$$

als Funktion der zwei Variablen x und y . Die folgende Grafik enthält den zugehörigen Funktionsgraphen. Man erkennt deutlich die relativen Extrema vom Typ (I) und (III):



Beispiel Gesucht ist der minimale Abstand des Punktes $P = [1, 1, 1]^T \in \mathbb{R}^3$ zur Kugeloberfläche

$$K := \{X \in \mathbb{R}^3 : \|X\| = 1\}.$$

Gesucht ist also das Minimum des Abstandsquadrats

$$f(X) := \|X - P\|^2 \quad (\text{grad } f(X) = 2(X - P))$$

unter der Nebenbedingung

$$g(X) := \|X\|^2 - 1 = 0 \quad (\text{grad } g(X) = 2X).$$

Mögliche Kandidaten für Minima sind die Punkte $X \in K$ mit

$$\text{grad } f(X) + \lambda \text{grad } g(X) = 0$$

für ein $\lambda \in \mathbb{R}$, also diejenigen $X \in K$ mit

$$2(X - P) + 2\lambda X = 0 \quad \implies \quad (1 + \lambda)X = P.$$

Da $X \in K$, muss $|1 + \lambda| = \|P\| = \sqrt{3}$ sein, also $\lambda = -1 \pm \sqrt{3}$, und damit $X = \pm \frac{1}{\sqrt{3}}P$. Es ist klar, dass $-\frac{1}{\sqrt{3}}P$ (relatives) Maximum und $\frac{1}{\sqrt{3}}P$ (relatives) Minimum der Abstandes ist. Für den minimalen Abstand erhält man also

$$\left\| \frac{1}{\sqrt{3}}P - P \right\| = \sqrt{3} - 1.$$