§ 8.5. Höhere Ableitungen.

$$f : \Omega \to \mathbb{R}, \quad \Omega \subset \mathbb{R}^?$$

Defn 8.44 — $f \in C^1(\Omega)$ heisst von Klasse

Struwe 7.5.1, 7.5.2

C^2 falls $\dfrac{\partial f}{\partial x^i} \in C^1(\Omega)$ $\quad 1 \leq i \leq n$

Für beliebiges m, die Funktion $f \in C^1(\Omega)$ heisst von der Klasse C^m, $f \in C^m(\Omega)$ falls $\dfrac{\partial f}{\partial x^i} \in C^{m-1}(\Omega)$ $\quad 1 \leq i \leq n$.

Für eine $f \in C^2(\Omega)$, die Funktionen

$$\frac{\partial^2 f}{\partial x^i \partial x^j} := \frac{\partial}{\partial x^i}\left(\frac{\partial f}{\partial x^j}\right)$$

heissen die zweiten partiellen Ableitungen von f.

Analog definiert man die m-ten partielle Ableitungen von f oder partielle Ableitungen vom Grad m für jedes $m > 0$ (für $f \in C^m(\Omega)$).

Satz 8.45 (Schwarz) Sei $f \in C^2(\Omega)$
(Struwe 7.5.1)
Dann gilt: $\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial}{\partial x^i}\left(\frac{\partial}{\partial x^j} f\right)$

$$= \frac{\partial}{\partial x^j}\left(\frac{\partial}{\partial x^i} f\right) = \frac{\partial^2 f}{\partial x^j x^i}$$

Im Allgemein

Satz 8.46 Für jede C^k-Funktion sind alle Partielle Ableitungen vom Grad $\leq k$ von der Reihenfolge der Ableitungen unabhängig

Von Satz 8.35 erhalten wir folgende notwendige Bedingung für konservativität

Kor 8.47 Sei $v : \Omega \to \mathbb{R}^n$, $v = (v^i)_{1 \leq i \leq n}$
Struwe 7.5.1
ein C^1-Vektorfeld.
Falls v konservativ ist, folgt

$$\frac{\partial v^i}{\partial x^j} = \frac{\partial v^j}{\partial x^i} \quad 1 \leq i, j \leq n.$$

Beweis: Nach Voraussetzung gibt es $f \in C^1(\Omega)$ mit

$v^i(x) = \frac{\partial f}{\partial x^i}$. Da nun $v^i \in C^1$, $1 \le i \le n$

folgt $f \in C^2(\Omega)$, woraus

$$\frac{\partial v^i}{\partial x^j} = \frac{\partial}{\partial x^j}\left(\frac{\partial f}{\partial x^i}\right) = \frac{\partial}{\partial x^i}\left(\frac{\partial f}{\partial x^j}\right) = \frac{\partial}{\partial x^i} v^j$$

folgt ∎

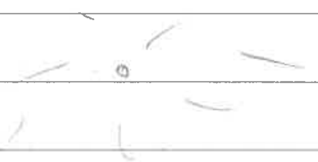
Bsp 8.481) $v(x,y) = \begin{pmatrix} 4xy^2 \\ 2y \end{pmatrix}$

Es gilt $\frac{\partial v^1}{\partial y} = 8xy$ $\qquad \frac{\partial v^2}{\partial x} = 2$

Also ist v nicht konservativ.

2) Sei $\Omega = \{(x,y) \in \mathbb{R}^2 ; (x,y) \neq (0,0)\}$

$= \mathbb{R}^2 \setminus \{0,0\}$



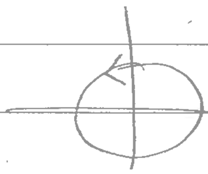
und $v(x,y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \right)$

Dann $v : \Omega \to \mathbb{R}^2$ mindestens C^1

Ausserdem $\frac{\partial v^1}{\partial y} = \frac{y^2-x^2}{(x^2+y^2)^2} = \frac{\partial v^2}{\partial x}$

Jetzt berechnen wir $\int_C v \, ds$

wobei C:



$c(t) = (\cos t, \sin t)$
$t \in (0, 2\pi]$

$$\int_C v \, ds = \int_0^{2\pi} \langle v(c(t)), c'(t) \rangle dt$$

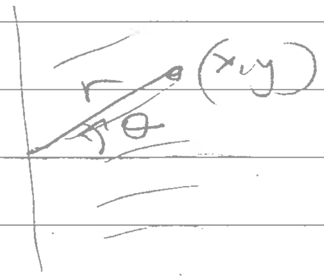
$$= \int_0^{2\pi} (-\sin t, \cos t) \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} dt$$

$$= \int_0^{2\pi} (\sin^2 t + \cos^2 t) \, dt = 2\pi \neq 0$$

$\Rightarrow$ v auf Ω nicht konservativ!

Jetzt betrachten wir

$$\Omega' = \{(x,y) \mid x > 0\}$$



und führen Polarkoordinaten ein

$$x = r\cos\theta, \quad y = r\sin\theta \qquad -\frac{\pi}{2} < \theta < \frac{\pi}{2}$$

Dann ist $\operatorname{tg}\theta = \frac{y}{x}$ und

$$\theta = \arctan\frac{y}{x}$$

Wir betrachten $\theta : \Omega' \to \mathbb{R}$ als ein Funktion der Variablen x, y und berechnen

$$\frac{\partial\theta}{\partial x} = \frac{1}{1+\left(\frac{y}{x}\right)^2}\left(\frac{-y}{x^2}\right) = -y/x^2+y^2$$

$$\frac{\partial\theta}{\partial y} = \frac{x}{x^2+y^2}$$

Also gilt $\nabla\theta(x,y) = v(x,y)$

$\forall (x,y) \in \Omega'$

$\Rightarrow v$ ist konservativ auf Ω'.

Das heisst Konservativität ist eine Eigenschaft zugleich des Vektorfeldes v <u>und</u> der Region Ω.

Defn 8.49 Eine offene Menge $\Omega \subset \mathbb{R}^2$ heisst einfach zusammenhängend falls

① Ω ist stückweise C^1-wegzusammenhängend

② Jeder stückweise C^1-Weg in Ω kann stetig innerhalb Ω auf einen Punkt zusammengezogen werden.

Die Region $\Omega = \mathbb{R}^2 \setminus \{0\}$ ist nicht einfach zu..

$\Omega' = \{(x, y) \mid x > 0\}$ ist es aber.

Satz 8.50 (Poincare) (Struwe 8-4-3) sei $\Omega \in \mathbb{R}^2$ beschränkt zusammenhängend sowie einfachzusammenhängend, sei $v \in C^1(\Omega; \mathbb{R}^2)$ Vektorfeld. Dann sind äquivalent

1) v ist konservativ

2) $\dfrac{\partial v_1}{\partial y} = \dfrac{\partial v_2}{\partial x}$

Taylorentwicklung und der lokale Verhalten von C^m-Funktionen

Wir werden jetzt ein Verallgemeinerung der 1-Variablen Taylorentwicklung herleiten.

Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine C^m-Funktion sowie $x_0, x_1 \in \mathbb{R}^n$

(Allgemein könnte man $\mathbb{R}^n$ durch eine offene konvexe Menge ersetzen).

Sei $\varphi : \mathbb{R} \to \mathbb{R}^n$
$t \mapsto (1-t)x_0 + x_1$

Dann ist $g := f \circ \varphi : \mathbb{R} \to \mathbb{R}$ eine C^m-Funktion
$(g(0) = f(x_0), \quad g(1) = f(x_1))$
und (Nach Taylor entwicklung von Funktionen 1-variable.)

es gibt $\xi \in (0,1)$ so dass

(*) $g(1) = g(0) + g'(0) + \dots + \frac{g^{(m-1)}(0)}{(m-1)!} + \frac{g^{(m)}(\xi)}{m!}$.

Jetzt berechnen wir $g^{(l)}(t)$ im Funktion von f und seinem Ableitungen.

Für $g'(t)$ benützen wir die Kettenregel:

$$g'(t) = df(\varphi(t)).\ \varphi'(t), \quad \text{mit}$$

$$\varphi'(t) = x_1 - x_0 = (x_1^1 - x_0^1, x_1^2 - x_0^2, \dots, x_1^n - x_0^n)$$

Erhalten wir:

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(\varphi(t))\,(x_1^i - x_0^i)$$

$$= \nabla f(\varphi(t)).(x_1 - x_0)$$

$$g'(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x_0)\,(x_1^i - x_0^i) = \nabla f(x_0).(x_1 - x_0)$$

Jetzt berechnen wir $g^{(2)}(t)$:

$$g^{(2)}(t) = \frac{d}{dt}\left(g'(t)\right) = \sum_{i=1}^{n} \frac{d}{dt}\left(\frac{\partial f}{\partial x^i}(\varphi(t))\right)(x_1^i - x_0^i)$$

Analog gilt $\frac{d}{dt}\left(\frac{\partial f}{\partial x^i}(\varphi(t))\right) = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x^j \partial x^i}(\varphi(t))(x_1^j - x_0^j)$

Eingesetzt gilt:

$$g^{(2)}(t) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial^2 f}{\partial x^j \partial x^i}(\varphi(t))\right)(x_1^i - x_0^i)(x_1^j - x_0^j)$$

$$g^{(2)}(0) = \sum_{i,j=1}^{n}\left(\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)\right)(x_1^i - x_0^i)(x_1^j - x_0^j)$$

Daraus schliesst man induktiv das

$$g^{(k)}(t) = \sum_{i_1, i_2, \dots i_k = 1}^{n} \frac{\partial^k f}{\partial x^{i_1} \dots \partial x^{i_k}}(\varphi(t)) \prod_{l=1}^{k}(x_1^{i_l} - x_0^{i_l})$$

Eingesetzt in (*) (s. 190) ergibt

Stimme (Satz 7.5.2)

Satz 8.51 (Taylor entwicklung)

$$f(x_1) = f(x_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x_0)(x_1^i - x_0^i) + \dots$$

$$+ \frac{1}{(m-1)!} \sum_{i_1 \dots i_{m-1}=1}^{n} \frac{\partial^{m-1} f}{\partial x^{i_1} \dots \partial x^{i_{m-1}}}(x_0) \prod_{\ell=1}^{(m-1)} (x_1^{i_\ell} - x_0^{i_\ell})$$

$$+ \frac{1}{m!} \sum_{i_1, i_2 \dots i_m=1}^{n} \frac{\partial^m f}{\partial x^{i_1} \dots \partial x^{i_m}}(x_\xi) \prod_{\ell=1}^{m} (x_1^{i_\ell} - x_0^{i_\ell})$$

mit eine Zahl $\xi \in (0,1)$, $x_\xi = (1-\xi)x_0 + \xi x_1$

Bmk 8.52

Insbesondere für $m=2$ erhalten wir for f

die quadratische Näherung

$$f(x_1) = f(x_0) + \nabla f(x_0)(x_1 - x_0) +$$

$$\frac{1}{2} \sum_{i,j=1}^{2} \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)(x_1^i - x_0^i)(x_1^j - x_0^j)$$

$$+ r_2(f, x_1, x_0) \quad \text{mit}$$

Fehler $\frac{r_2(f, x_1, x_0)}{|x_1 - x_0|} \to 0 \quad (x_1 \to x_0)$

Defn 8.53 Die Matrix der zweiten partiellen Ableitungen heisst die Hesse-Matrix von f, und mit Hess(f) oder $\nabla^2 f$ bezeichnet

$$\mathrm{Hess}(f) = \nabla^2 f := \left(\frac{\partial^2 f}{\partial x^i \partial x^j} \right)_{i,j=1,\dots,n}$$

$$= \begin{pmatrix} \frac{\partial^2 f}{\partial x^1 \partial x^1} & \frac{\partial^2 f}{\partial x^1 \partial x^2} & \cdots & \frac{\partial^2 f}{\partial x^1 \partial x^n} \\ \frac{\partial^2 f}{\partial x^2 \partial x^1} & \frac{\partial^2 f}{\partial x^2 \partial x^2} & \cdots & \frac{\partial^2 f}{\partial x^2 \partial x^n} \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x^n \partial x^1} & & \cdots & \frac{\partial^2 f}{\partial x^n \partial x^n} \end{pmatrix}$$

Seien ∇f, $x_1 - x_0$ Zeilenvektoren und sei $(x - x_0)^t$ der zu $x - x_0$ transponierte Spaltenvektor. Dann wird die Taylorentwicklung von Grad 2 äquivalent zu

$$f(x) = f(x_0) + \nabla f(x_0)(x - x_0)^t$$

$$+ \frac{1}{2}(x - x_0) \cdot \nabla^2 f(x_0) \cdot (x - x_0)^t$$

$$+ r_3(f, x, x_0).$$

__Bmk__. Die Hesse-Matrix von f, nach Satz von Schwarz ist eine symmetrische Matrix.

<u>Bsp.</u> $f(x,y) = e^{x+y} \cos x$ im Punkt $(0,0)$

Die Taylorentwicklung vom Grad 2 ist

$$\frac{\partial f}{\partial x} = e^{x+y} \cos x - e^{x+y} \sin x \;, \quad \frac{\partial f}{\partial x}(0,0) = 1$$

$$\frac{\partial f}{\partial y} = e^{x+y} \cos x \;, \quad \frac{\partial f}{\partial x}(0,0) = 1.$$

$$(\nabla f)(0,0) = (1,1) \qquad f(0,0) = 1$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = e^{x+y} \cos x - e^{x+y} \sin x \;, \quad \frac{\partial^2 f}{\partial x \partial y}((0,0)) = 1$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = e^{x+y} \cos x - e^{x+y} \sin x$$

$$- e^{x+y} \sin x - e^{x+y} \cos x$$

$$= -2 e^{x+y} \sin x$$

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

$$\frac{\partial^2 f}{\partial y^2} = e^{x+y} \cos x \qquad \frac{\partial^2 f}{\partial y^2}(0,0) = 1$$

196

$$\nabla^2 f(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

$$((x,y)-(0,0))\, \nabla^2 f(0,0)\, ((x,y)-(0,0))^T = (x,y)\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= (x,y)\begin{pmatrix} y \\ x+y \end{pmatrix}$$

$$= 2xy + y^2$$

$$f(x,y) = e^{x+y}\cos x = 1 + \begin{pmatrix} 1 \\ 1 \end{pmatrix}(x,y) + \frac{1}{2}(x,y)\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= 1 + (x+y) + \frac{1}{2}(2xy+y^2)$$

$$+ r_3(f,(x,y)).$$

Taylorpolynom vom grad 2 : $1+(x+y)+\frac{1}{2}(2xy+y^2)$

Die Hesse-Matrix bestimmt ob die Funktion f in der Nähe von x konvex oder konkav ist (oder nicht)

Sie "spielt" die gleiche Rolle, wie die zweite Ableitung von Funktionen in einer Variable.

Als nächstes benötigen wir eine mehrdimensionale Entsprechung zu den Positivität (Negativität) in den eindimensionalen Beziehungen

$f''(z) > 0$ bzw $f''(z) < 0$.

<u>Defn</u> <u>8.54</u>. Eine symmetrische Matrix $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ heisst

1) <u>positiv definit</u> wenn $x^t A x = \sum_{i,j=1}^{n} a_{ij} x^i x^j > 0$ $\forall x \in \mathbb{R}^n$.

(oder "wenn ihre Eigenwerte sämtlich positiv sind)

2) <u>negativ definit</u> wenn ${}^t x A x < 0 \quad \forall x \in \mathbb{R}^n$

(= wenn ihre Eigenwerte sämtlich negativ sind)

3) sonst <u>indefinit</u> (wenn sie sowohl positive als auch negative Eigenwerte besitzt).

Im symmetrischen 2x2 Fall ist die Untersuchung auf Definitheit besonders leicht

<u>Satz 8.55</u> Eine symmetrische Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$$ ist genau dann

1) positiv definit, wenn $\det A > 0$ und $a_{11} > 0$

2) negativ definit, wenn $\det A > 0$ und $a_{11} < 0$

3) indefinit wenn $\det A < 0$.

Extrema von Funktionen mehrerer Variablen

Jetzt werden wir nach Punkten $x \in \mathbb{R}^n$ suchen, in denen eine Funktion $F: \mathbb{R}^n \to \mathbb{R}$ ein lokales Extremum annimmt.

Wir erinnern uns an das Vorgehen im $f: \mathbb{R} \to \mathbb{R}$

1) Finde alle Punkte $x \in \mathbb{R}$, für die $f'(x) = 0$ gilt (notwendige Bedingung)

2) Falls in einem solchen Punkt zusätzlich $f''(x) > 0$ (bzw $f''(z) < 0$) gilt so handelt es sich um ein lokales Minimum (bzw Maximum) (hinreichende Bedingung).

Jetzt verallgemeinern wir diese Strategie

Zunächst

Defn 8.55 Ein Punkt $x_0 \in \mathbb{R}^n$ mit $df(x_0) = 0$ heisst kritischer Punkt von f (oder stationärer Punkt von f)

Satz 8.56. Sei $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$
(7.5.3 Skript) $f \in C^2(\Omega)$; $x_0 \in \Omega$.

1) Falls $x_0 \in \Omega$ lokale Extremum (min oder max) von f ist, so gilt $df(x_0) = 0$.

2) Falls $df(x_0) = 0$, und falls $\mathrm{Hess}(f)(x_0)$ positiv definit ist, so ist x_0 eine lokale Minimalstelle

3) Falls $df(x_0) = 0$ und $\mathrm{Hess}_f(x_0) < 0$ negativ definit ist, so ist x_0 eine lokale Maximalstelle

4) Falls $df(x_0) = 0$, und $\mathrm{Hess}_f(x_0)$ indefinite ist, so ist x_0 ein Sattelpunkt

(d.h. jede Umgebung U von x_0 enthält Punkte $p, q \in U$ mit
$f(p) > f(x_0) > f(q)$.)

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

$\nabla f = (2(x-1),\ 2(y+2),\ 2(z+1))$

$\nabla f(x_0) = (0,0,0) \Rightarrow x_0 = (1,-2,-1)$

$$H_f(x_0) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$H_f(x_0)$ ist pos. definit.

$\Rightarrow x_0 = (1,-2,-1)$ ist ein lokales Minimum.

② $f(x,y) = \cos(x+2y) + \cos(2x+3y)$

$\nabla f = (-\sin(x+2y) - 2\sin(2x+3y),$

$-2\sin(x+2y) - 3\sin(2x+3y)) = (0;0)$

$\Rightarrow -\sin(x+2y) - 2\sin(2x+3y) = 0$
$-2\sin(x+2y) - 3\sin(2x+3y) = 0$

$\Rightarrow \sin(2x+3y) = 0 \quad , \quad \sin(x+2y) = 0$

$\Rightarrow \left.\begin{array}{l} 2x+3y = k\pi \\ x+2y = l\pi \end{array}\right\} \Rightarrow y = k\pi$ und $x = l\pi$.

kritische Punkte: $(\pi l, \pi k) \quad k,l \in \mathbb{Z}$.

(202)

$$\frac{\partial f}{\partial y \partial x} = \frac{\partial f}{\partial x \partial y} = -2\cos(x+2y) - 6\cos(2x+3y)$$

$$\frac{\partial^2 f}{\partial x^2} = -\cos(x+2y) - 4\cos(2x+3y)$$

$$\frac{\partial^2 f}{\partial y^2} = -4\cos(x+2y) - 9\cos(2x+3y)$$

$(0,0)$: $\frac{\partial^2 f}{\partial x^2} = -5 < 0$

$$|\nabla^2 f(0,0)| = \begin{vmatrix} -5 & -8 \\ -8 & -13 \end{vmatrix} = (13)(5) - (64) = 1 > 0$$

$\Rightarrow$ $\nabla^2 f(0,0)$ ist negativ definit und $(0,0)$ ist eine lokale maximalle stelle

Auch alle punkte $(2\pi k, 2\pi l)$ sind lok. max.

Analog, Bis auf Addition von vielfachen von 2π:

f hat lokale minimale stelle in (π, π)

und sattelpunkte in $(0, \pi)$ und $(\pi, 0)$