Höhere partielle Ableitungen. Ist eine Funktion $f : U \to \mathbb{R}$ partiell differenzierbar, so ist $f_{x_j} := \frac{\partial f}{\partial x_j}$ wieder eine $\mathbb{R}$–wertige Funktion $U \to \mathbb{R}$. Also läßt sich die Frage nach der partiellen Differenzierbarkeit von $\frac{\partial f}{\partial x_j}$ stellen. Falls $\frac{\partial f}{\partial x_j} : U \to \mathbb{R}$ partiell, z.B. nach x_k differenzierbar ist, so schreibt man

$$\frac{\partial^2 f}{\partial x_j \partial x_k}$$

für diese zweite partielle Ableitung. Dieser Vorgang kann beliebig iteriert werden. Man schreibt dan für die k-te partielle Ableitung von f nach $x_{j_1}, \dots, x_{j_k}$ (abgeleitet in dieser Reihenfolge)

$$\frac{\partial^k f}{\partial x_{j_1} \cdots \partial x_{j_k}}$$

$$C^k(U) := \{f : U \to \mathbb{R} : f \text{ } k\text{–fach } \underline{\text{stetig}} \text{ partiell differenzierbar}\}$$

Die Stetigkeit der ersten partiellen Ableitungen hat die folgende Konsequenz:

Satz. Existieren alle partielle Ableitungen $\frac{\partial f_j}{\partial x_k}$ und sind stetig [in x] so ist $f = (f_1, \dots, f_n)$ (total) differenzierbar [in x]. Im Allgemeinen gilt die Umkehrung nicht.

Bei höheren partiellen Ableitungen ist die Reihenfolge der Differenziation wichtig:

Beispiel. $f(x_1, x_2) = \begin{cases} x_1 x_2 \dfrac{x_1^2 - x_2^2}{x_1^2 + x_2^2} & \text{für } (x_1, x_2) \neq (0,0) \\ 0 & \text{für } (x_1, x_2) = (0,0) \end{cases}$ $\qquad \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(0,0) = -1, \quad \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(0,0) = 1$

Satz von Schwarz. Sei $f \in C^k(U)$. Dann hängt

$$\frac{\partial^k f}{\partial x_{j_1} \partial x_{j_2} \cdots \partial x_{j_k}}$$

<u>nicht</u> von der Reihenfolge der partiellen Differenziation ab.

Die beiden folgenden Sätze ergeben sich unmittelbar aus den entsprechenden Aussagen in einer Variable und der Kettenregel:

Mittelwertsatz für reellwertige Funktionen. Sei $f : U \to \mathbb{R}$ (total) differenzierbar, $U \subset \mathbb{R}^m$ offen. Liegt die Verbindungsstrecke zwischen x und $x + h$ in U, so gilt

$$f(x + h) - f(x) = Df_{x+\xi h} \cdot h \qquad \text{für ein } \xi \in [0, 1]$$

Satz von Taylor. Sei $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^m$ offen und $f \in C^{n+1}(U)$. Liegt die Verbindungsstrecke zwischen x und $x + h$ in U, so gilt

$$\begin{aligned} f(x+h) &= \sum_{k=0}^{n} \frac{1}{k!} \Big(h_1 \frac{\partial}{\partial x_1} + \cdots + h_m \frac{\partial}{\partial x_1} \Big)^k f\ (x) + \\ &+ \frac{1}{(n+1)!} \Big(h_1 \frac{\partial}{\partial x_1} + \cdots + h_m \frac{\partial}{\partial x_1} \Big)^{n+1} f\ \ (x + \xi h) \\ &= \sum_{K=0}^{n} \frac{1}{K!} \sum_{\Sigma k_j = K} \binom{K}{k_1, \dots, k_m} h_1^{k_1} \cdots h_m^{k_m} \frac{\partial^K f}{\partial x_1^{k_1} \cdots \partial x_m^{k_m}}\ (x) + \\ &+ \frac{1}{(n+1)!} \sum_{\Sigma k_j = n+1} \binom{n+1}{k_1, \dots, k_m} h_1^{k_1} \cdots h_m^{k_m} \frac{\partial^{n+1} f}{\partial x_1^{k_1} \cdots \partial x_m^{k_m}}\ (x + \xi h) \end{aligned}$$

für ein geeignetes $\xi \in [0, 1]$. Hierbei ist $\dbinom{K}{k_1, \dots, k_m} = \dfrac{K!}{k_1! k_2! \cdots k_m!}$.

Insbesondere gilt für $n = 2$

$$f(x + h) = f(x) + Df_x \cdot h + \frac{1}{2} h^t \cdot \Big(\frac{\partial^2 f}{\partial x_j \partial x_k}(x) \Big) \cdot h + o(\|h\|^2)$$

Implizite Funktionen sind, grob gesagt, Zuordnungen, die nicht explizit durch eine Zuordnugsvorschrift $f : x \mapsto y = f(x)$ gegeben sind, sondern implizit aus einer Gleichung $f(x, y) = c$ bestimmt werden müssen (falls dies überhaupt möglich ist). Einige Vorbemerkungen:

Der Geschwindigkeitsvektor einer Kurve γ: $\dot{\gamma}(t) := \begin{pmatrix} \gamma_1'(t) \\ \vdots \\ \gamma_m'(t) \end{pmatrix}$.

Der Gradient einer reellwertigen Funktion: $\operatorname{grad} f(x) := (Df_x)^T$.

Sei $f : U \to \mathbb{R}$ eine differenzierbare Funktion. Dann steht der Gradient senkrecht auf den Niveauflächen: *Sei $N_c := \{x \in U : f(x) = c\}$ und $\gamma : (-r, r) \to N_c$ eine beliebige differenzierbare Kurve. Die Kettenregel, angewendet auf $f \circ \gamma = c$ ergibt dann die Identität $Df_{\gamma(t)} \cdot \dot{\gamma}(t) = 0$, d.h., $\operatorname{grad} f$ und $\dot{\gamma}$ sind orthogonal bezüglich des kanonischen Skalarproduktes in $\mathbb{R}^m$.*