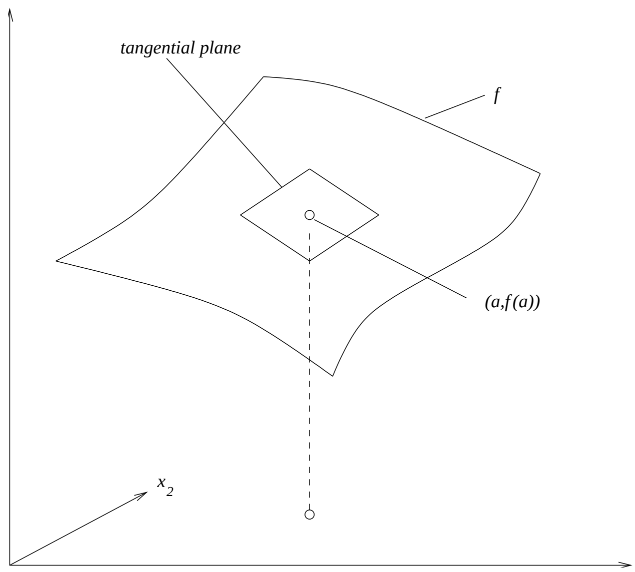
5 Differentialrechnung im $\mathbb{R}^n$

5 a.) Definition der Ableitung

Es soll der Begriff der Ableitung von reellen Funktionen auf Funktionen von n Veränderlichen verallgemeinert werden. Die Grundidee dabei ist folgende: Seien $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ und $a \in U$. Unter der Ableitung von f an der Stelle a versteht man diejenige lineare Abbildung, die f in einer Umgebung des Punktes a „am besten approximiert".
Ich betrachte als Beispiel Abbildungen der Form $f : \mathbb{R}^2 \to \mathbb{R}$. Der Graph dieser Abbildung ist eine Fläche im $\mathbb{R}^3$. Der Graph einer linearen Abbildung $T : \mathbb{R}^2 \to \mathbb{R}$ ist eine durch den Nullpunkt gehende Ebene. Man bestimme nun diejenige lineare Abbildung T, deren Graph parallel ist zu der Tangentialebene an den Graphen von f im Punkt $(a, f(a))$. Dann heißt T „Ableitung von f im Punkt a", und die Funktion

$$x \mapsto f(a) + T(x - a)$$

approximiert f in einer Umgebung von a.



Diese auf der Anschauung beruhende Idee wird in der folgenden Definition mathematisch genau gefasst:

Definition: Sei $U \subseteq \mathbb{R}^n$ offen. $f : U \to \mathbb{R}^m$ heißt differenzierbar an der Stelle $a \in U$, wenn es eine lineare Abbildung $T : \mathbb{R}^n \to \mathbb{R}^m$ und eine an der Stelle a stetige Abbildung $r : U \to \mathbb{R}^m$ mit $r(a) = 0$ gibt derart, daß für alle $x \in U$ gilt

$$f(x) = f(a) + T(x - a) + r(x)\|x - a\| .$$

Um zu prüfen, ob f an der Stelle $a \in D$ differenzierbar ist muß man zuerst eine geeignete lineare Abbildung T finden und dann prüfen, ob für

$$r(x) = \frac{f(x) - f(a) - T(x-a)}{\|x-a\|}$$

$\lim_{x \to a} r(x) = 0$ gilt. Ich werde später angeben, wie man T findet. Es kann höchstens ein solches T geben. Denn es gilt:

Lemma: T ist eindeutig bestimmt.

Beweis: Seien T_1, T_2 lineare Abbildungen mit

$$\begin{aligned} f(x) &= f(a) + T_1(x-a) + r_1(x)\|x-a\| \\ f(x) &= f(a) + T_2(x-a) + r_2(x)\|x-a\|\,. \end{aligned} \quad , \quad \lim_{x\to a} r_1(x) = \lim_{x \to a} r_2(x) = 0$$

Dann folgt

$$(T_1 - T_2)(x-a) = \Big(r_2(x) - r_1(x)\Big)\|x-a\|\,.$$

Sei $h \in \mathbb{R}^n$. Für alle hinreichend kleinen $t > 0$ gilt dann $x = a + th\, eU$, und es folgt

$$(T_1 - T_2)(th) = t(T_1 - T_2)(h) = \Big(r_2(a+th) - r_1(a+th)\Big)\|th\|\,,$$

also

$$(T_1 - T_2)(h) = \lim_{t\to 0}(T_1 - T_2)(h) = \lim_{t \to 0}\Big(r_2(a+th) - r_1(a+th)\Big)\|h\| = 0\,,$$

also $T_1 = T_2$, da h beliebig gewählt war. ■

Definition: Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ sei in $a \in U$ differenzierbar. Dann heißt die eindeutig bestimmte lineare Abbildung $T : \mathbb{R}^n \to \mathbb{R}^m$ mit

$$f(x) = f(a) + T(x-a) + r(x)\|x-a\|\,, \quad \lim_{x\to a} r(x) = 0\,,$$

die Ableitung von f an der Stelle a. Bezeichnung: $T = f'(a)$.

Bei linearen Abbildungen läßt man häufig die Klammern um das Argument weg und schreibt $T(h) = Th = f'(a)h$.

Ist f reellwertig, dann ist $f'(a)$ eine lineare Abbildung $f'(a) : \mathbb{R}^m \to \mathbb{R}$. Solche lineare Abbildungen nennt man auch Linearformen. In diesem Fall bezeichnet man $f'(a)$ auch als Gradienten von f und schreibt $\text{grad} f(a) := f'(a)$. Aus der linearen Algebra weiß man, daß jede Linearform auf $\mathbb{R}^n$ mit Hilfe des Skalarproduktes in eindeutiger Weise durch einen Vektor im $\mathbb{R}^n$ dargestellt werden kann: Es gibt ein eindeutig bestimmtes $y \in \mathbb{R}^n$, so daß für alle $h \in \mathbb{R}^n$ gilt

$$\mathrm{grad} f(a)h = y \cdot h\,.$$

Oft wird auch der Vektor y mit $\mathrm{grad} f(a)$ bezeichnet. Man muß sich aber im klaren darüber sein, daß dies unpräzise ist, weil man für zwei verschiedene Dinge dieselbe Bezeichnung benutzt. Der Vektor $\mathrm{grad} f(a)$ zeigt in die Richtung des größten Zuwachses der Abbildung $f'(a)$, und damit auch in die Richtung des größten Zuwachses der Abbildung f an der Stelle a, weil $f'(a)$ die Abbildung f in einer Umgebung von a approximiert.
Die Tangentialhyperebene der Abbildung $f : U \to \mathbb{R}$ im Punkt $a \in U$ wird definiert durch

$$\Big\{(x,z) \in \mathbb{R}^{n+1}\,\Big|\, z = f(a) + \Big[f'(a)\Big](x-a)\,,\ \ x \in \mathbb{R}^n\Big\}\,.$$

Der Vektor $(-\mathrm{grad} f(a), 1) \in \mathbb{R}^{n+1}$ steht senkrecht auf dieser Hyperebene. Denn für zwei Vektoren (x_1, z_1) und (x_2, z_2) aus dieser Hyperebene gilt

$$\begin{aligned}(x_2, z_2) - (x_1, z_1) &= \Big(x_2 - x_1, \Big[f'(a)\Big](x_2 - a) - \Big[f'(a)\Big](x_1 - a)\Big)\\ &= \Big(x_2 - x_1, \Big[f'(a)\Big](x_2 - x_1)\Big)\,,\end{aligned}$$

also

$$\begin{aligned}&\Big(-\mathrm{grad} f'(a), 1\Big) \cdot \Big[(x_2, z_2) - (x_1, z_1)\Big]\\ &= \Big(-\mathrm{grad} f'(a)\Big) \cdot (x_2 - x_1) + \Big[f'(a)\Big](x_2 - x_1)\\ &= -\Big(\mathrm{grad} f'(a)\Big) \cdot (x_2 - x_1) + \Big(\mathrm{grad} f'(a)\Big) \cdot (x_2 - x_1) = 0\,.\end{aligned}$$

Ist insbesondere $U \subseteq \mathbb{R}$ und $f : U \to \mathbb{R}$ eine reelle Funktion, dann ist die lineare Abbildung $T = f'(a) : \mathbb{R} \to \mathbb{R}$ gegeben durch

$$Th = \frac{df}{dx}(a)h, \ \ h \in \mathbb{R}\,,$$

mit der klassischen Ableitung $\frac{df}{dx}(a) \in \mathbb{R}$ von f an der Stelle a.

Es gilt:

Lemma: Die Abbildung $f : U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$, ist differenzierbar in $a \in U$, genau dann wenn jede der Komponentenfunktionen $f_1, \ldots, f_m : U \to \mathbb{R}$ differenzierbar ist in $a \in U$. Es gilt dann

$$f_j'(a) = \Big(f'(a)\Big)_j\,, \quad j = 1, \ldots, m\,.$$

Beweis: Wenn $f'(a)$ existiert, ist $(f'(a))_j : \mathbb{R}^n \to \mathbb{R}$ linear, und es gilt

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{f_j(h+a) - f_j(a) - (f'(a))_j h}{\|h\|} = 0\,,$$

also ist $(f'(a))_j = f_j'(a)$. Ist umgekehrt $f_j'(a)$ die Ableitung von f_j für $j = 1, \ldots, m$, dann wird durch

$$Th = \begin{pmatrix} f_1'(a)h \\ \vdots \\ f_m'(a)h \end{pmatrix} : \mathbb{R}^n \to \mathbb{R}^m$$

eine lineare Abbildung definiert für die gilt

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(a+h) - f(a) - Th}{\|h\|} = 0\,,$$

also ist $T = f'(a)$. ■

Um $f'(a)v$ für $v \in \mathbb{R}^n$ zu bestimmen, setze man $x = a + tv$ mir $t \in \mathbb{R}\,,\ t \neq 0$. Es gilt dann

$$f(a+tv) = f(a) + f'(a)(tv) + r(tv+a)|t|\,\|v\|\,,$$

also

$$f'(a)v = \frac{f(a+tv) - f(a)}{t} - r(tv+a)\,\frac{|t|}{t}\,\|v\|\,,$$

somit

$$f'(a)v = \lim_{\substack{t\to 0\\ t\neq 0}} f'(a)v = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(a+tv) - f(a)}{t}\,.$$

Der rechtsstehende Grenzwert heißt Richtungsableitung von f an der Stelle a in Richtung von $v \in \mathbb{R}^n$. Für die Richtungsableitung benutze ich die Bezeichnung

$$D_v f(a) := \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(a+tv) - f(a)}{t}\,.$$

Zur Bestimmung der linearen Abbildung $f'(a)$ genügt es, die Richtungsableitungen $D_{v_i} f(a)$ für eine Basis $v_1, \ldots, v_n$ von $\mathbb{R}^n$ zu berechnen, weil man jedes $v \in \mathbb{R}^n$ als Linearkombination $v = \sum_{i=1}^n \alpha_i v_i$ schreiben kann mit $\alpha_i \in \mathbb{R}^n$, also

$$f'(a)v = f'(a)(\sum_{i=1}^n \alpha_i v_i) = \sum_{i=1}^n \alpha_i f'(a) v_i.$$

Es ist naheliegend, als Basis die Standardbasis $e_1, \ldots, e_n$ zu wählen. Die dabei benötigte Richtungsableitung $D_{e_i} f(a)$ nennt man i–te partielle Ableitung. Partielle Ableitungen bezeichnet man durch

$$\frac{\partial f}{\partial x_i}\,,\ \ D_i f\,,\ \ f_{x_i}\,,\ \ f'_{x_i}\,,$$

manchmal auch durch $f_{|i}$ oder f_i . Hierbei können aber Verwechslungen auftreten. Es gilt:

$$\frac{\partial f}{\partial x_i}(a) = \lim_{\substack{x_i \to a_i \\ x_i \neq a_i}} \frac{f(a_1, \ldots, x_i, \ldots, a_n) - f(a_1, \ldots, a_i, \ldots, a_n)}{x_i - a_i},$$

d.h.

$$\frac{\partial f_j}{\partial x_i}(a) = \lim_{\substack{x_i \to a_i \\ x_i \neq a_i}} \frac{f_j(a_1, \ldots, x_i, \ldots, a_n) - f_j(a_1, \ldots, a_i, \ldots, a_n)}{x_i - a_i};$$

$$i = 1, \ldots, n; \quad j = 1, \ldots, m$$

Zur Bestimmung der partiellen Ableitungen von f an der Stelle a genügt somit die Differentialrechnung einer reellen Variablen.

Sei $U \subseteq \mathbb{R}^n$ eine offene Menge und sei $f : U \to \mathbb{R}^m$ differenzierbar an der Stelle $a \in U$. Zur Bestimmung von $f'(a)$ geht man nun folgendermaßen vor: Weil f im Punkt a differenzierbar ist, existieren alle partiellen Ableitungen $D_i f(a) = \frac{\partial f}{\partial x_i}(a)$ an der Stelle a . Für beliebiges $h \in \mathbb{R}^n$ gilt $h = \sum_{i=1}^n h_i e_i$, $h_i \in \mathbb{R}$, also

$$f'(a)h = f'(a)(\sum_{i=1}^n h_i e_i) = \sum_{i=1}^n \Big(f'(a)e_i\Big)h_i = \sum_{i=1}^n D_i f(a) h_i$$

oder, in der üblichen Matrizenschreibweise,

$$f'(a)h = \begin{pmatrix} [f'(a)h]_1 \\ \vdots \\ [f'(a)h]_m \end{pmatrix} = \begin{pmatrix} D_1 f_1(a) & \ldots & D_n f_1(a) \\ \vdots & & \\ D_1 f_m(a) & \ldots & D_n f_m(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

Also ist

$$\begin{pmatrix} D_1 f_1(a) & \ldots & D_n f_1(a) \\ \vdots & & \\ D_1 f_m(a) & \ldots & D_n f_m(a) \end{pmatrix}$$

die zu den Standardbasen $e_1, \ldots, e_n$ in $\mathbb{R}^n$ und $e_1, \ldots, e_m$ in $\mathbb{R}^m$ gehörende Darstellung von $f'(a)$ also $m \times n$ Matrix. Diese Matrix heißt Jacobi–Matrix von f an der Stelle a .

Um zu prüfen ob f an der Stelle a differenzierbar ist, prüft man zuerst, ob alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}(a)$ existieren. Dies ist eine notwendige Bedingung für die Differenzierbarkeit. Wenn alle partiellen Ableitungen existieren braucht aber f nicht differenzierbar zu sein. Daher muß man mit der Matrix

$$T = \Big(\frac{\partial f_j}{\partial x_i}(a)\Big)_{\substack{i=1,\ldots n \\ j=1,\ldots,m}}$$

prüfen, ob

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(a+h) - f(a) - Th}{\|h\|} = 0\,,$$

gilt. Falls dies richtig ist, ist $f'(a) := T$ die Ableitung von f an der Stelle a.

5 b.) Beispiele

1.) Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$, $\quad x = (x_1, x_2) \mapsto f(x) = (f_1(x_1,x_2),\, f_2(x_1,x_2))$, definiert durch

$$\begin{aligned} f_1(x_1,x_2) &= x_1^2 - x_2^2 \\ f_2(x_1,x_2) &= 2x_1x_2\,. \end{aligned}$$

Falls f an der Stelle $a = (a_1, a_2) \in \mathbb{R}^2$ differenzierbar ist, muß gelten

$$f'(a) = \begin{pmatrix} 2a_1 & -2a_2 \\ 2a_2 & 2a_1 \end{pmatrix}.$$

Sei $r(x) = \Big(r_1(x), r_2(x)\Big) = \dfrac{f(x) - f(a) - f'(a)(x-a)}{\|x-a\|}$.

Dann gilt

$$\begin{aligned} r_1(x) &= \frac{x_1^2 - x_2^2 - a_1^2 + a_2^2 - 2a_1(x_1 - a_1) + 2a_2(x_2 - a_2)}{\|x-a\|} = \frac{(x_1-a_1)^2 - (x_2-a_2)^2}{\|x-a\|} \\ r_2(x) &= \frac{2x_1x_2 - 2a_1a_2 - 2a_2(x_1-a_1) - 2a_1(x_2-a_2)}{\|x-a\|} = \frac{2(x_1-a_1)(x_2-a_2)}{\|x-a\|}\,. \end{aligned}$$

Mit der Maximumsnorm ergibt sich

$$\begin{aligned} |r_1(x)| &\leq 2\|x-a\| \\ |r_2(x)| &\leq 2\|x-a\|\,, \end{aligned}$$

also

$$\lim_{x \to a} \|r(x)\| \leq \lim_{x \to a} 2\|x-a\| = 0\,,$$

also ist f an der Stelle a, und weil a beliebig war, in ganz $\mathbb{R}^2$ differenzierbar.

2.) Sei $A : \mathbb{R}^n \to \mathbb{R}^m$ linear, und sei $f : \mathbb{R}^n \to \mathbb{R}^m$ deffiniert durch

$$f(x) = Ax + c\,, \quad c \in \mathbb{R}^m\,.$$

Dann ist f in ganz $\mathbb{R}^n$ differenzierbar, und es gilt $f'(a) = A$. Denn

$$\frac{f(a+h) - f(a) - Ah}{\|h\|} = \frac{A(a+h) + c - Aa - c - Ah}{\|h\|} = 0\,.$$

3.) $f : \mathbb{R}^2 \to \mathbb{R}$ sei definiert durch

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

f ist an der Stelle $a = 0$ nicht differenzierbar, aber die partiellen Ableitungen existieren im Nullpunkt und es gilt

$$\frac{\partial f}{\partial x_1}(0) = 0\,, \quad \frac{\partial f}{\partial x_2}(0) = 0\,.$$

Wäre also f in 0 differenzierbar, müßte

$$\mathrm{grad} f(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

sein. Es gilt aber für

$$r(h) = \frac{f(h) - f(0)}{|h|} = \frac{h_1 h_2}{\sqrt{h_1^2 + h_2^2}\sqrt{h_1^2 + h_2^2}} = \frac{h_1 h_2}{h_1^2 + h_2^2}$$

auf der Diagonalen $h = (h_1, h_1)$

$$\lim_{h \to 0} |r(h)| = \lim_{h \to 0} \frac{h_1^2}{2h_1^2} = \frac{1}{2} \neq 0\,.$$

5 c.) Einfache Eigenschaften und Rechenregeln für differenzierbare Abbildungen

Zur Vorbereitung benötige ich ein Resultat über lineare Abbildungen:

Lemma: Sei $A : \mathbb{R}^n \to \mathbb{R}^m$ linear. Dann ist A stetig, und es existiert eine nicht negative Konstante, die mit $\|A\|$ bezeichnet wird, so daß $\|Ax\| \leq \|A\|\,\|x\|$ gilt für alle $x \in \mathbb{R}^n$.

Beweis: Es existiert eine $m \times n$ Matrix $(a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$, so daß für $y = Ax$ gilt

$$\begin{aligned} y_1 &= a_{11}x_1 + \ldots + a_{1n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + \ldots + a_{mn}x_n\,. \end{aligned}$$

Jede dieser Abbildungsgleichungen definiert eine stetige Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}$, also ist A stetig.

Sei $E = \{x \in \mathbb{R}^n \,\big|\, \|x\| \leq 1\}$. Dies ist eine kompakte Menge. Also existiert

$$\|A\| := \sup_{x \in E} \|Ax\|\,,$$

da A stetig ist. Für alle $x \in \mathbb{R}^n$ gilt nun

$$\|Ax\| = \Big\|A\Big(\|x\|\frac{x}{\|x\|}\Big)\Big\| = \Big\|\,\|x\|A\Big(\frac{x}{\|x\|}\Big)\Big\| = \|x\|\,\Big\|A\Big(\frac{x}{\|x\|}\Big)\Big\| \le \|A\|\,\|x\|\,,$$

wegen $\frac{x}{\|x\|} \in E$. ■

Definition: Die Zahl $\|A\| := \sup_{\|x\|\le 1} \|Ax\|$ heißt Norm der linearen Abbildung A.

Lemma: Es gilt für alle $x \in \mathbb{R}^n$, $c \in \mathbb{R}$ und für alle linearen Abbildungen $A, B : \mathbb{R}^n \to \mathbb{R}^m$

(i) $\|A\| \ge 0$,
$A = 0 \iff \|A\| = 0$,

(ii) $\|cA\| = |c|\,\|A\|$,

(iii) $\|A + B\| \le \|A\| + \|B\|$,

(iv) $\|Ax\| \le \|A\|\,\|x\|$.

Beweis: (ii) ist klar, (iv) wurde schon gezeigt. Zum Beweis von (iii) beachte, daß

$$\begin{aligned} \|(A+B)x\| &= \|Ax + Bx\| \le \|Ax\| + \|Bx\| \\ &\le \|A\|\,\|x\| + \|B\|\,\|x\| = (\|A\| + \|B\|)\|x\| \end{aligned}$$

gilt. Hieraus folgt

$$\|A + B\| = \sup_{\|x\|\le 1} \|(A+B)x\| \le \sup_{\|x\|\le 1} (\|A\| + \|B\|)\|x\| = \|A\| + \|B\|\,.$$

Zum Beweis von (i) sei $\|A\| = 0$. Für alle $x \in \mathbb{R}^n$ folgt dann aus (iv), daß $0 \le \|Ax\| \le \|A\|\,\|x\| = 0$ gilt. Somit ist $A = 0$. Die anderen Aussagen von (i) sind klar. ■

Die Menge $L(\mathbb{R}^n, \mathbb{R}^m)$ der linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ bildet einen Vektorraum, und dieses Lemma zeigt, daß $\|A\|$ wirklich die Eigenschaften einer Norm besitzt. Also wird $L(\mathbb{R}^n, \mathbb{R}^m)$ mit dieser Norm zu einem normierten Raum.

Wir studieren nun wieder differenzierbare Abbildungen.

Seien $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$, $a \in U$. Wenn für f alle Richtungsableitungen im Punkt a existieren, braucht f doch nicht stetig zu sein. Ein Beispiel ist

$$f : \mathbb{R}^2 \to \mathbb{R}\,,$$

$$f(x_1, x_2) = \begin{cases} 0\,, & (x_1, x_2) = 0 \\ \dfrac{x_1 x_2^2}{x_1^2 + x_2^6}\,, & (x_1, x_2) \ne 0\,. \end{cases}$$

Die Richtungsableitungen existieren alle im Nullpunkt, weil für $v = (v_1, v_2) \in \mathbb{R}^2\,,\ \ v \neq 0\,,$ gilt

$$D_v f(0) = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(tv)-f(0)}{t} = \begin{cases} \lim\limits_{t\to 0} \dfrac{v_1 v_2^2}{v_1^2 + t^4 v_2^6} = \dfrac{v_2^2}{v_1}\,, & v_1 \neq 0 \\ 0\,, & v_1 = 0\,. \end{cases}$$

Aber für

$$h = (h_1, \sqrt{h_1})\,, \quad h_1 > 0\,,$$

gilt

$$f(h) = \frac{h_1^2}{h_1^2 + h_1^3} = \frac{1}{1+h_1} \to 1 \neq f(0)\,,$$

für $h_1 \to 0$.

Es gilt aber

Satz: Sei $U \subseteq \mathbb{R}^n$, und $f : U \to \mathbb{R}^m$ sei an der Stelle $a \in U$ differenzierbar. Dann existiert $c > 0$, so daß für alle x aus einer Umgebung von a gilt

$$\|f(x) - f(a)\| \le c\|x-a\|\,.$$

Insbesondere ist f in a stetig.

Beweis: Es gilt

$$f(x) = f(a) + f'(a)(x-a) + r(x)\|x-a\|\,,$$

also

$$\|f(x)-f(a)\| \le \|f'(a)\|\,\|x-a\| + \|r(x)\|\,\|x-a\|\,.$$

Wegen $\lim_{x\to a} r(x) = 0$ folgt

$$\|f(x)-f(a)\| \le c\|x-a\|\,,$$

also

$$\lim_{x\to a} \|f(x)-f(a)\| = 0\,.$$

∎

Satz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$, $g : U \to \mathbb{R}^m$ seien beide an der Stelle $a \in U$ differenzierbar. Dann sind auch $f+g$ und cf $(c \in \mathbb{R})$ an der Stelle a differenzierbar und es gilt

$$\begin{aligned} (f+g)'(a) &= f'(a) + g'(a) \\ (cf)'(a) &= cf'(a)\,. \end{aligned}$$

Beweis: Es gilt

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + r_1(a+h)\|h\|\,, \quad \lim_{h\to 0} r_1(a+h) = 0 \\ g(a+h) &= g(a) + g'(a)h + r_2(a+h)\|h\|\,, \quad \lim_{h\to 0} r_2(a+h) = 0\,. \end{aligned}$$

Also folgt

$$(f+g)(a+h) = (f+g)(a) + \Big(f'(a) + g'(a)\Big)h + (r_1 + r_2)(a+h)\|h\|\,.$$

Hieraus resultiert $(f+g)'(a) = f'(a) + g'(a)$. Die andere Aussage ergibt sich ebenso. ∎

Satz (Produktregel): Die Funktionen $f : U \to \mathbb{R}$ und $g : U \to \mathbb{R}$ seien beide an der Stelle $a \in U$ differenzierbar. Dann ist auch $f \cdot g$ an der Stelle a differenzierbar mit

$$(f \cdot g)'(a)h = f(a)g'(a)h + g(a)f'(a)h\,.$$

Beweis:

$$\begin{aligned} (f\cdot g)(a+h) = \Big(&f(a) + f'(a)h + r_1(a+h)\|h\|\Big) \\ &\cdot \Big(g(a) + g'(a)h + r_2(a+h)\|h\|\Big) \\ = \;& (f\cdot g)(a) + f(a)g'(a)h + g(a)f'(a)h + r(a+h)\|h\|\,, \end{aligned}$$

mit

$$\begin{aligned} r(a+h)\|h\| &= f'(a)h\, g'(a)h + \Big(g(a) + g'(a)h\Big)r_1(a+h)\|h\| \\ &\quad + \Big(f(a) + f'(a)h\Big)r_2(a+h)\|h\| + r_1(a+h)r_2(a+h)\|h\|^2\,. \end{aligned}$$

Es gilt

$$\begin{aligned} \lim_{h\to 0}\|r(a+h)\| \;\leq\; \lim_{h\to 0}\frac{1}{\|h\|}\Big[&\|f'(a)\|\,\|g'(a)\|\,\|h\|^2 \\ &+ \Big(|g(a)| + \|g'(a)\|\,\|h\|\Big)|r_1(a+h)|\,\|h\| \\ &+ \Big(|f(a)| + \|f'(a)\|\,\|h\|\Big)|r_2(a+h)|\,\|h\| \\ &+ |r_1(a+h)|\,|r_2(a+h)|\,\|h\|^2\Big] = 0\,. \end{aligned}$$

∎

Bemerkung: Natürlich kann man die Produktregel auch in der Form

$$\operatorname{grad}(fg)(a) = f(a)\,\operatorname{grad}g(a) + g(a)\,\operatorname{grad}f(a)$$

schreiben.

Satz (Kettenregel): Sei $V \subseteq \mathbb{R}^n$ offen, $g : V \to \mathbb{R}^m$ sei an der Stelle $b \in V$ differenzierbar. Sei $U \subseteq \mathbb{R}^p$ offen, $f : U \to V$ sei an der Stelle $a \in U$ differenzierbar, und es sei $b = f(a)$. Dann ist $g \circ f$ an der Stelle $a \in U$ differenzierbar, und es gilt

$$(g \circ f)'(a) = g'\Big(f(a)\Big) \circ f'(a)\,.$$

Beweis: Zur Abkürzung seien

$$T_2 = g'(b)\,, \quad T_1 = f'(a)\,,$$

und für $h \in \mathbb{R}^p$, $\|h\|$ genügend klein, sei

$$R(h) = (g \circ f)(a+h) - (g \circ f)(a) - T_2 T_1 h\,.$$

Es muß gezeigt werden, daß

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{\|R(h)\|}{\|h\|} = 0$$

ist. Es gilt

$$\begin{aligned} f(x) - f(a) - T_1(x-a) &= r_1(x-a)\|x-a\|\,, \quad \lim_{x \to 0} r_1(x) = 0 \\ g(y) - g(b) - T_2(y-b) &= r_2(y-b)\|y-b\|\,, \quad \lim_{y \to 0} r_2(y) = 0\,. \end{aligned}$$

Somit folgt

$$\begin{aligned} R(h) &= g\Big(f(a+h)\Big) - g\Big(f(a)\Big) - T_2\Big(f(a+h) - f(a)\Big) \\ &\quad + T_2\Big(f(a+h) - f(a) - T_1 h\Big) \\ &= r_2\Big(f(a+h) - f(a)\Big)\|f(a+h) - f(a)\| + T_2\Big(r_1(h)\|h\|\Big)\,, \end{aligned}$$

d.h.

$$\begin{aligned} \lim_{h \to 0} \frac{\|Rh\|}{\|h\|} &\leq \lim_{h \to 0} \Big[\frac{1}{\|h\|}\,\|r_2\Big(f(a+h) - f(a)\Big)\|\,\|f(a+h) - f(a)\|\Big] \\ &\quad + \lim_{h \to 0} \|T_2\Big(r_1(h)\Big)\|\Big]\,. \end{aligned}$$

Wegen der Stetigkeit von T_2 folgt $\lim_{h \to 0} T_2(r_1(h)) = 0$. Wegen $\|f(a+h) - f(a)\| \leq c\|h\|$ ergibt sich

$$\lim_{h \to 0} \frac{\|Rh\|}{\|h\|} \leq c \lim_{h \to 0} \|r_2\Big(f(a+h) - f(a)\Big)\| = 0\,.$$

Damit ist der Satz bewiesen. ■

Für die Jacobi–Matrizen von $f : U \subseteq \mathbb{R}^p \to \mathbb{R}^n$, $\;g : V \subseteq \mathbb{R}^n \to \mathbb{R}^m$ und $h = g \circ f : U \subseteq \mathbb{R}^p \to \mathbb{R}^m$ ergibt sich also

$$\begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_p} \\ \vdots & & \\ \frac{\partial h_m}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_p} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_1}{\partial y_n} \\ \vdots & & \\ \frac{\partial g_m}{\partial y_1} & \cdots & \frac{\partial g_m}{\partial y_n} \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_p} \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_p} \end{pmatrix},$$

wobei die partiellen Ableitungen von h und f an der Stelle a , von g an der Stelle $b = f(a)$ zu bilden sind.

Es ergibt sich also

$$\frac{\partial h_j}{\partial x_i}(a) = \sum_{k=1}^{n} \frac{\partial g_j}{\partial y_k}(b) \frac{\partial f_k}{\partial x_i}(a) , \;\; i = 1, \ldots, p \;\; j = 1, \ldots, m .$$

Folgerung: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ sei in $a \in U$ differenzierbar, und es gelte $f(a) \neq 0$. Dann gilt

$$\operatorname{grad} \frac{1}{f}(a) = (\frac{1}{f})'(a) = -\frac{1}{f(a)^2} \operatorname{grad} f(a) .$$

Beweis: Betrachte die Abbildung $g : \mathbb{R}\backslash\{0\} \to \mathbb{R}$, $\;g(x) := \frac{1}{x}$. Dann gilt

$$\frac{1}{f} = g \circ f ,$$

also

$$\operatorname{grad} \frac{1}{f}(a) = \Big(\frac{1}{f}\Big)'(a) = g'\Big(f(a)\Big) f'(a) = -\frac{1}{f(a)^2} \operatorname{grad} f(a) .$$

■

Man kann die Ableitung der Umkehrabbildung einer bijektiven Abbildung $f : U \to V$, $U, V \subseteq \mathbb{R}^n$ offen, mit der Kettenregel berechnen.

Denn sei $g : V \to U$, die Umkehrabbildung zu f , sei f an der Stelle $a \in U$ und g an der Stelle $b = f(a) \in V$ differenzierbar. Dann gilt

$$g \circ f = \mathrm{id}_U ,$$

also

$$(g \circ f)'(a) = g'\Big(f(a)\Big) \circ f'(a) = \mathrm{id}_{\mathbb{R}^n} ,$$

folglich

$$g'\Big(f(a)\Big) = \Big(f'(a)\Big)^{-1} ,$$

oder

$$g'(b) = \Big[f'\Big(g(b)\Big)\Big]^{-1} .$$

Wenn man voraussetzt, daß die Umkehrabbildung der linearen Abbildung $f'(a)$ existiert, genügt es sogar vorauszusetzen, daß g stetig sei. Nach einem Satz der linearen Algebra existiert die Umkehrabbildung von $f'(a)$, wenn die Determinante $\det f'(a)$ der $f'(a)$ repräsentierenden $n \times n$–Matrix von Null verschieden ist. Man nennt $\det f'(a)$ Jacobi–Determinante.

Satz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ sei umkehrbar, an der Stelle a differenzierbar, und die Jacobi–Determinante $\det f'(a)$ sei von Null verschieden. Sei $f(U)$ offen und die Umkehrabbildung $g : f(U) \to \mathbb{R}^n$ sei an der Stelle $b = f(a)$ stetig. Dann ist g an der Stelle b differenzierbar, und es gilt

$$g'(b) = \Big(f'(a)\Big)^{-1} .$$

Beweis: Zunächst zeige ich: Es gibt eine Umgebung V von b und eine Konstante $c > 0$ mit

$$\frac{\|g(y) - g(b)\|}{\|y - b\|} \le c ,$$

für alle $y \in V \cap f(U)$.

Es gilt

$$f(x) - f(a) = f'(a)(x - a) \quad + \quad r(x)\|x - a\| , \ \lim_{x \to a} r(x) = 0 .$$

Also folgt mit der umgekehrten Dreiecksungleichung

$$\frac{\|g(y) - g(b)\|}{\|y - b\|} = \frac{\|g(y) - g(b)\|}{\|f(g(y)) - f(g(b))\|}$$

$$= \frac{\|g(y) - g(b)\|}{\|f'(a)(g(y) - g(b)) \ + \ r(g(y))\|g(y) - g(b)\| \, \|}$$

$$\le \frac{\|(f'(a))^{-1} f'(a) \, (g(y) - g(b))\|}{\|f'(a)(g(y) - g(b))\| - \|r(g(y))\| \, \|(f'(a))^{-1} f'(a)(g(y) - g(b))\|}$$

$$\le \frac{\|(f'(a))^{-1}\| \, \|f'(a)(g(y) - g(b))\|}{\|f'(a)(g(y) - g(b))\|(1 - \|r(g(y))\| \, \|(f'(a))^{-1}\|)}$$

$$= \frac{\|(f'(a))^{-1}\|}{1 - \|r(g(y))\| \, \|f'(a)^{-1}\|} .$$

Wegen $\lim_{y \to b} r(g(y)) = 0$ folgt die Behauptung. Hierbei wird die Stetigkeit von g benutzt. Nun ergibt sich der Satz folgendermaßen:

Es muß gezeigt werden, daß

$$\lim_{\substack{y\to b\\ y\neq b}} \frac{g(y)-g(b)-f'(a)^{-1}(y-b)}{\|y-b\|} = 0$$

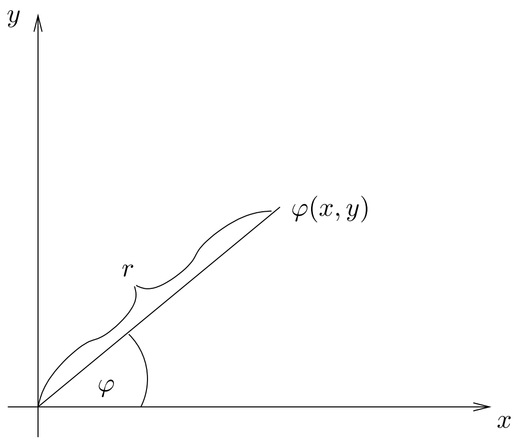
ist. Es gilt

$$\begin{aligned}
&\lim_{\substack{y\to b\\ y\neq b}} \frac{g(y)-g(b)-f'(a)^{-1}(y-b)}{\|y-b\|} \\
&= \lim_{\substack{y\to b\\ y\neq b}} \frac{g(y)-g(b)-f'(a)^{-1}(f(g(y))-f(g(b)))}{\|y-b\|} \\
&= \lim_{\substack{y\to b\\ y\neq b}} \frac{g(y)-g(b)-f'(a)^{-1}\Big(f'(a)(g(y)-g(b))+r(g(y))\|g(y)-g(b)\|\Big)}{\|y-b\|} \\
&= \lim_{\substack{y\to b\\ y\neq b}} f'(a)^{-1}\Big(r\Big(g(y)\Big)\Big)\,\frac{\|g(y)-g(b)\|}{\|y-b\|} = 0\,.
\end{aligned}$$

■

Beispiel (Polarkoordinatenabbildung): Seien $\varepsilon > 0$ und $c_2 > c_1 > 0$, und für $c_1 \leq r \leq c_2$, $0 \leq \varphi \leq 2\pi - \varepsilon$ sei

$$\begin{aligned}
x &= f_1(r,\varphi) = r\cos\varphi \\
y &= f_2(r,\varphi) = r\sin\varphi\,.
\end{aligned}$$



Diese Abbildung ist injektiv, differenzierbar, die Jacobi–Determinante ist von Null verschieden, und die Umkehrabbildung ist stetig, weil f auf einer kompakten Menge definiert ist. Ohne die Umkehrabbildung bestimmen zu müssen, kann die Ableitung der Umkehrabbildung bestimmt werden. Im Punkt $(x,y) = f(r,\varphi)$ gilt

$$\begin{aligned}
[f^{-1}]'(x,y) = f'(r,\varphi)^{-1} &= \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}^{-1} \\
= \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\frac{1}{r}\sin\varphi & \frac{1}{r}\cos\varphi \end{pmatrix} &= \begin{pmatrix} \dfrac{x}{\sqrt{x^2+y^2}} & \dfrac{y}{\sqrt{x^2+y^2}} \\ \dfrac{-y}{x^2+y^2} & \dfrac{x}{x^2+y^2} \end{pmatrix}.
\end{aligned}$$

5 d.) Mittelwertsatz

Der Mittelwertsatz für reelle Funktionen kann auf *reellwertige* Funktionen verallgemeinert werden.

Satz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ sei differenzierbar, und die Verbindungsstrecke der beiden Punkte $a, b \in U$ sei ganz in U enthalten. Dann gibt es einen Punkt c auf dieser Verbindungsstrecke mit

$$f(b) - f(a) = f'(c)(b-a)\,.$$

Beweis: Definiere die Abbildung $\gamma : [0,1] \to U$ durch $t \mapsto \gamma(t) := a + t(b-a)$. Hierdurch wird $[0,1]$ auf die Verbindungstrecke von a und b abgebildet. γ ist differenzierbar mit

$$\gamma'(t) = b - a\,.$$

Auf die differenzierbare Funktion $F : [0,1] \to \mathbb{R}$,

$$F = f \circ \gamma\,,$$

wende man den Mittelwertsatz für reelle Funktionen an. Es folgt mit geeignetem $\vartheta \in (0,1)$

$$f(a) - f(b) = F(1) - F(0) = F'(\vartheta) = f'\Big(\gamma(\vartheta)\Big)\gamma'(\vartheta) = f'(c)(b-a)\,,$$

mit $c = \gamma(\vartheta)$. ∎

Natürlich kann man den Mittelwertsatz auch folgendermaßen formulieren: Zu $x, x+h \in U$ gibt es $\vartheta, 0 < \vartheta < 1$, mit

$$f(x+h) - f(x) = f'(x+\vartheta h)h\,.$$

Folgerung (Schrankensatz): Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ sei differenzierbar, und die Ableitung von f sei auf der Verbindungsstrecke von a und b beschränkt, d.h. es existiere eine Konstante $S > 0$ mit

$$\|f'(c)\| \leq S$$

für alle c aus der Verbindungsstrecke. Dann gilt

$$\|f(x+h) - f(x)\| \leq S\|h\| \, .$$

Beweis: Wegen der Äquivalenz aller Normen auf $\mathbb{R}^m$ genügt es, diese Folgerung für die Maximumsnorm auf $\mathbb{R}^m$ zu beweisen. Wendet man Mittelwertsatz auf die j–te Komponentenfunktion f_j von f an, dann folgt wegen $f_j' = (f')_j$, daß

$$\begin{aligned} |f_j(x+h) - f_j(x)| &= |f_j'(x+\vartheta_j h)h| = |(f')_j(x+\vartheta_j h)h| \\ &\leq \ \|f'(x+\vartheta_j h)h\|_\infty \leq \|f'(x+\vartheta_j h)\| \, \|h\| \leq S\|h\| \, , \end{aligned}$$

also

$$\|f(x+h) - f(x)\|_\infty = \max_{j=1,\ldots,m} |f_j(x+h) - f_j(x)| \leq S\|h\| \, .$$

■

Satz: Sei $U \subseteq \mathbb{R}^n$ offen und wegzusammenhängend. $f : U \to \mathbb{R}^m$ sei differenzierbar. Dann gilt: f ist konstant, genau dann wenn $f'(x) = 0$ ist für alle $x \in U$.

Zum Beweis benützen wir folgendes

Lemma: Sei $U \subseteq \mathbb{R}^n$ offen und wegzusammenhängend, und seien $a, b \in U$. Dann können a, b durch einen ganz in U verlaufenden Streckenzug mit den ‚Eckpunkten'

$$a_0 = a \, , \ \ a_1, \, \ldots, a_{k-1}, \ \ a_k = b$$

verbunden werden.

Dieses Lemma beweise ich nicht. Man findet einen Beweis im Buch von Barner–Flohr, Analysis II, S. 56.

Beweis des Satzes: Falls f konstant, ist $f' = 0$. Zum Beweis der Umkehrung sei $f'(x) = 0$ für alle $x \in U$. Es genügt, die Behauptung für Funktionen $f : U \to \mathbb{R}$ zu beweisen, weil man im allgemeinen Fall die Komponentenfunktionen $f_1, \ldots, f_m$ von f betrachten kann. Sei also f reellwertig.

Seien $a, b \in U$. Man verbinde diese Punkte durch einen Streckenzug in U mit den angegebenen Eckpunkten, und wende den Mittelwertsatz auf jede der Strecken mit den Endpunkten $a_j, \ \ a_{j+1}$ an, $j = 0, 1, \ldots, k-1$. Es folgt

$$f(a_{j+1}) = f(a_j) + f'(c)(a_{j+1} - a_j) = f(a_j) \, ,$$

also

$$f(b) = f(a) \, .$$

$\blacksquare$

Wenn f differenzierbar ist, existieren alle partiellen Ableitungen. Wenn die partiellen Ableitungen existieren, braucht f aber nicht differenzierbar zu sein. Es gilt jedoch:

Satz: Sei $U \in \mathbb{R}^n$ offen. Wenn die Funktion $f : U \to \mathbb{R}^m$ sämtliche partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, besitzt, und diese an der Stelle $a \in U$ stetig sind, dann ist f an der Stelle a differenzierbar.

Beweis: Es genügt zu zeigen, daß jede der Komponentenfunktionen $f_1, \ldots, f_m$ differenzierbar ist. Also kann man annehmen, daß $f : U \to \mathbb{R}$ gilt. Es ist zu zeigen, daß

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(a+h) - f(a) - Th}{\|h\|_\infty} = 0$$

ist mit

$$T := \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right).$$

Für $h \in \mathbb{R}^n$ setze

$$\begin{aligned} a_0 &= a\,, \\ a_1 &= a_0 + h_1 e_1 \\ a_2 &= a_1 + h_2 e_2 \\ &\vdots \\ a + h = a_n &= a_{n-1} + h_n e_n\,, \end{aligned}$$

wobei $e_1, \ldots, e_n \in \mathbb{R}^n$ die kanonische Basis sei. Es gilt dann

$$\begin{aligned} &f(a+h) - f(a) \\ &= \Big(f(a+h) - f(a_{n-1}) \Big) + \Big(f(a_{n-1}) - f(a_{n-2}) \Big) + \ldots + \Big(f(a_1) - f(a) \Big). \qquad (*) \end{aligned}$$

Läuft x auf der Verbindungsstrecke zwischen a_{j-1} und a_j, dann variiert nur die Komponente x_j von x. Da die Abbildung $x_j \mapsto f(x_1, \ldots, x_j, \ldots, x_n)$ nach Voraussetzung differenzierbar ist, kann der Mittelwertsatz auf jeden Summanden in der Formel $(*)$ angewendet werden. c_j sei der Zwischenpunkt auf der Verbindungsstrecke von a_{j-1} und a_j. Dann gilt

$$f(a+h) - f(a) = \sum_{j=1}^{n} \Big(f(a_j) - f(a_{j-1}) \Big) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(c_j) h_j\,.$$

Also folgt

$$
\begin{aligned}
f(a+h) - f(a) - Th &= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(c_j)h_j - Th \\
= \;\; \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(c_j)h_j - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)h_j &= \sum_{j=1}^{n} \Big(\frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(a) \Big) h_j \, ,
\end{aligned}
$$

somit

$$
|f(a+h) - f(a) - Th| \leq \|h\|_\infty \sum_{j=1}^{n} \Big| \frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(a) \Big| .
$$

Wegen $\|c_j - a\|_\infty \leq \|h\|_\infty$ folgt die Behauptung aus der Stetigkeit aller partiellen Ableitungen $\frac{\partial f}{\partial x_j}$, in a. ■

Dieser Satz liefert eine einfache hinreichende Bedingung für die Differenzierbarkeit einer Abbildung

$$
f : U \to \mathbb{R}^m \, , \;\; U \subseteq \mathbb{R}^n \, .
$$

Beispiel: $f : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ sei definiert durch

$$
f(x) = (x_1^2 + \ldots + x_n^2)^s \, , \quad s \in \mathbb{R} \, .
$$

Diese Abbildung ist überall differenzierbar. Denn die partiellen Ableitungen

$$
\frac{\partial f}{\partial x_j} = s \cdot (x_1^2 + \ldots + x_n^2)^{s-1} 2x_j
$$

sind stetig.

5 e.) Stetig differenzierbare Abbildungen

Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ sei in allen Punkten $x \in U$ differenzierbar. Dann wird durch

$$
x \mapsto f'(x) : U \to L(\mathbb{R}^n, \mathbb{R}^m)
$$

eine Abbildung von U in die Menge der linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ definiert. Wendet man die lineare Abbildung $f'(x)$ auf einen beliebigen Vektor $h \in \mathbb{R}^n$ an, dann erhält man einen Vektor in $\mathbb{R}^m$:

$$
f'(x,h) := f'(x)h \in \mathbb{R}^m \, .
$$

Also kann man f' auch als Abbildung von $U \times \mathbb{R}^n$ nach $\mathbb{R}^m$ auffassen:

$$
(x,h) \mapsto f'(x,h) : U \times \mathbb{R}^n \to \mathbb{R}^m \, .
$$

f' ist bezüglich des zweiten Arguments linear. Welche Auffassung man verwendet, ist eine Frage der Zweckmäßigkeit.

Weil $L(\mathbb{R}^n, \mathbb{R}^m)$ mit der am Anfang dieses Abschnittes eingeführten Norm ein normierter Raum ist, ist f' bei beiden Auffassungen eine Abbildung zwischen normierten Räumen. Für Abbildungen zwischen normierten Räumen ist der Begriff der Stetigkeit definiert, und man kann daher untersuchen, ob f' bei einer der beiden verschiedenen Auffassungen eine stetige Abbildung ist. Das folgende Lemma zeigt, daß es bei der Untersuchung der Stetigkeit nicht darauf ankommt, welche Auffassung man zu Grunde legt:

Lemma: Sei $U \subseteq \mathbb{R}^n$ eine offene Menge. Genau dann ist $f' : U \times \mathbb{R}^n \to \mathbb{R}^m$ stetig, wenn $f' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ stetig ist.

Beweis: Auf $\mathbb{R}^n$ und auf $U \times \mathbb{R}^n \subseteq \mathbb{R}^n \times \mathbb{R}^n$ verwende ich die Maximimsnorm. Sei $f' : U \times \mathbb{R}^n \to \mathbb{R}^m$ stetig und sei $a \in U$. Wähle $c > 0$ mit $K = \{x \in \mathbb{R}^n \,\big|\, \|x - a\|_\infty \leq c\} \subseteq U$. Weil f' auf der kompakten Menge

$$K \times \{h \in \mathbb{R}^n \,\big|\, \|h\|_\infty \leq 1\}$$

gleichmäßig stetig ist, existiert zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit

$$\|f'(x,h) - f'(a,h)\| \leq \varepsilon$$

für alle $x, h \in \mathbb{R}^n$ mit $\|x - a\|_\infty < \delta$ und $\|h\|_\infty \leq 1$, weil dann $\|(x,h) - (a,h)\|_\infty = \|(x - a, h)\|_\infty = \|x - a\|_\infty < \delta$ gilt. Also folgt für diese x und für die Norm $\|f'(x) - f'(a)\|$ der linearen Abbildung $f'(x) - f'(a) \in L(\mathbb{R}^n, \mathbb{R}^m)$:

$$\begin{aligned} \|f'(x) - f'(a)\| &= \sup_{\|h\| \leq 1} \|\big[f'(x) - f'(a)\big]h\| \\ &= \sup_{\|h\| \leq 1} \|f'(x)h - f'(a)h\| = \sup_{\|h\| \leq 1} \|f'(x,h) - f'(a,h)\| \leq \varepsilon\,. \end{aligned}$$

Dies bedeutet, daß $f' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ in a stetig ist. Weil a beliebig gewählt war, ist diese Abbildung stetig.

Sei umgekehrt $f' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ stetig und sei $(a, h) \in U \times \mathbb{R}^n$. Zu jedem $\varepsilon > 0$ gibt es dann eine Zahl $\delta > 0$, die kleiner oder gleich $\min(\varepsilon, 1)$ gewählt werden kann, so daß $\|f'(x) - f'(a)\| \leq \varepsilon$ gilt für alle $x \in U$ mit $\|x - a\|_\infty < \delta$. Für $(x, h_1) \in U \times \mathbb{R}^n$ mit

$\|(x, h_1) - (a, h)\|_\infty < \delta$ folgt dann

$$\begin{aligned}
&\|f'(x, h_1) - f'(a, h)\| = \|f'(x)h_1 - f'(a)h\| \\
&= \Big\| \Big[f'(x) - f'(a)\Big] h_1 - f'(a)\,(h_1 - h) \Big\| \\
&\le \|f'(x) - f'(a)\|\,\|h_1\| + \|f'(a)\|\,\|h_1 - h\| \\
&\le \varepsilon(\|h\|_\infty + \|h_1 - h\|_\infty) + \|f'(a)\|\delta \le \varepsilon\Big(\|h\|_\infty + 1 + \|f'(a)\|\Big),
\end{aligned}$$

wegen $\|x - a\|_\infty$, $\|h_1 - h\|_\infty < \delta \le \min(\varepsilon, 1)$. Weil $\|h\|_\infty + 1 + \|f'(a)\|$ unabhängig von (x, h_1) ist, folgt hieraus die Stetigkeit der Abbildung $f' : U \times \mathbb{R}^n \to \mathbb{R}^m$ in (a, h) . Da dieser Punkt beliebig gewählt war, ist diese Abbildung stetig. ∎

Definition: (i) Sei $U \subseteq \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}^m$ differenzierbar. Ist $f' : U \times \mathbb{R}^n \to \mathbb{R}^m$ beziehungsweise $f' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ stetig, dann heißt f stetig differenzierbar.

(ii) Seien $U, V \subseteq \mathbb{R}^n$ offen und sei $f : U \to V$ stetig differenzierbar und umkehrbar. Ist die Umkehrabbildung $f^{-1} : V \to U$ ebenfalls stetig differenzierbar, dann heißt f Diffeomorphismus.

Der folgende Satz gibt ein handhabbares Kriterium, mit dem man nachprüfen kann, ob eine Abbildung stetig differenzierbar ist:

Satz: Sei $U \subseteq \mathbb{R}^n$ offen. Die Abbildung $f : U \to \mathbb{R}^m$ ist stetig differenzierbar, genau dann wenn alle partiellen Ableitungen $\frac{\partial}{\partial x_i} f_j$ in U existieren und stetig sind.

Beweis: Die Abbildung $f' : U \times \mathbb{R}^n \to \mathbb{R}^m$ ist stetig, genau dann wenn jede der Komponentenfunktionen

$$(x, h) \mapsto f_j'(x, h) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f_j(x) h_i \qquad (*)$$

stetig ist. Wählt man für h den Einheitsbasisvektor e_i , dann folgt aus der stetigen Differenzierbarkeit von f , daß die partielle Ableitung

$$x \mapsto \frac{\partial}{\partial x_i} f_j(x) = f_j'(x, e_i) : U \to \mathbb{R}$$

stetig ist. Wenn umgekehrt alle partiellen Ableitungen von f in U existieren und stetig sind, dann ist f in U differenzierbar und die j–te Komponente der Ableitung ist gegeben durch die rechte Seite von $(*)$. Man sieht sofort, daß diese rechte Seite eine stetige Funktion von (x, h) ist. ∎

5 f.) Höhere Ableitungen, Taylorsche Formel

Die Ableitung von $f : U \to \mathbb{R}^m$ ist eine Abbildung $f : U \to L(\mathbb{R}^n, \mathbb{R}^m)$. Die Ableitung von f' wird man als zweite Ableitung f'' von f bezeichnen. Also ist die zweite Ableitung $f''(x)$ von f an der Stelle x eine lineare Abbildung von $\mathbb{R}^n$ in den Raum der linearen Abbildungen $L(\mathbb{R}^n, \mathbb{R}^m)$:

$$f'' : U \to L\Big(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m)\Big) .$$

Es ist möglich, die zweite Ableitung von f so zu definieren, weil $L(\mathbb{R}^n, \mathbb{R}^m)$ ein normierter Raum ist (sogar ein Banachraum). Denn man kann die Definition der Ableitung einer Funktion von $\mathbb{R}^n$ nach $\mathbb{R}^m$ ohne Änderung auf Funktionen zwischen allgemeinen normierten Räumen übertragen. Jedoch will ich die zweite Ableitung weniger abstrakt aber in äquivalenter Weise folgendermaßen definieren:

Definition: Sei $U \subseteq \mathbb{R}^n$ eine offene Menge und sei $f : U \to \mathbb{R}^m$ differenzierbar in U . Die Funktion f heißt zweimal differenzierbar in einem Punkt $x \in U$, wenn zu jeden festen $h \in \mathbb{R}^n$ die durch

$$g_h(x) = f'(x, h) = f'(x)h$$

definierte Funktion $g_h : U \to \mathbb{R}^m$ in x differenzierbar ist. Als zweite Ableitung von f im Punkt x bezeichnet man die durch

$$f''(x, h, k) = g_h'(x)k$$

definierte Funktion $(h, k) \mapsto f''(x, h, k) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$. Ist f in jedem Punkt von U zweimal differenzierbar, dann gilt $f'' : U \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$.

Für jedes $x \in U$ ist

$$(h, k) \mapsto f''(x, h, k) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$$

eine bilineare Abbildung, d.h. eine Abbildung, die in beiden Variablen linear ist. Denn da $g_{h_1+h_2}(x) = f'(x)(h_1 + h_2) = f'(x)h_1 + f'(x)h_2 = g_{h_1}(x) + g_{h_2}(x)$ gilt, folgt

$$\begin{aligned} f''(x, h_1 + h_2, k_1 + k_2) &= g'_{h_1+h_2}(x)(k_1 + k_2) \\ &= \Big[g_{h_1}(x) + g_{h_2}(x)\Big]'(k_1 + k_2) = g'_{h_1}(x)(k_1 + k_2) + g'_{h_2}(x)(k_1 + k_2) \\ &= f''(x, h_1, k_1) + f''(x, h_1, k_2) + f''(x, h_2, k_1) + f''(x, h_2, k_2) . \end{aligned}$$

Ebenso folgt

$$\begin{aligned} f''(x, ch, k) &= cf''(h, k) , \\ f''(x, h, ck) &= cf''(h, k) . \end{aligned}$$

Seien $h = (h_1, \dots, h_n)$ und $k = (k_1, \dots, k_n)$. Dann gilt

$$f''(x,h,k) = g_h'(x)k = \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\, g_h(x)k_j\,.$$

Wegen

$$g_h(x) = f'(x)h = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\, f(x)h_i$$

folgt also

$$\begin{aligned} f''(x,h,k) = g_h'(x)k &= \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\Big(\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\, f(x)h_i\Big)k_j \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i}\, f(x)h_ik_j\,. \end{aligned}$$

Hierbei sieht man, daß die zweiten partiellen Ableitungen $\frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f(x)$ alle existieren, indem man für h und k die Standardbasisvektoren e_i und e_j wählt. Es gilt

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x) = \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f(x) = \begin{pmatrix} \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f_1(x) \\ \vdots \\ \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f_m(x) \end{pmatrix} \in \mathbb{R}^m\,.$$

Man setzt auch

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

Für reellwertiges $f : U \to \mathbb{R}$ erhält man in Matrizenschreibweise

$$\begin{aligned} f''(x,h,k) &= (k_1, \dots, k_n) \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \\ &= k \cdot Hh\,, \end{aligned}$$

wobei man

$$H = \Big(\frac{\partial^2 f}{\partial x_i \partial x_j}\Big)_{i,j=1,\dots,n}$$

als die Hessesche Matrix bezeichnet. Für beliebiges $f : U \to \mathbb{R}^m$ erhält man

$$\Big[f''(x,h,k)\Big]_j = k \cdot H_j h\,,$$

wobei H_j die Hessesche Matrix der j–ten Komponentenfunktion f_j ist. Insbesondere folgt hieraus

$$(f'')_j(x,h,k) = (f_j)''(x,h,k)\,,$$

d.h. die j–te Komponente von f'' ist die zweite Ableitung der Komponentenfunktion f_j. Falls $f : U \to \mathbb{R}^m$ differenzierbar ist, und f in $a \in U$ zweimal differenzierbar ist, dann ist H beziehungsweise H_j eine symmetrische Matrix, d. h. es gilt

$$\frac{\partial^2 f_j}{\partial x_i \partial x_k}(a) = \frac{\partial^2 f_j}{\partial x_k \partial x_i}(a)\,.$$

Dies ergibt sich aus dem folgenden Satz. Man beachte aber, daß alle zweiten partiellen Ableitungen in a existieren können, ohne daß f diese Voraussetzungen erfüllt. Dann braucht H nicht symmetrisch zu sein.

Satz von H.A. Schwartz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ sei differenzierbar und in einem Punkt $x \in U$ zweimal differenzierbar. Dann gilt für alle $h, k \in \mathbb{R}^n$

$$f''(x,h,k) = f''(x,k,h)\,.$$

(Die bilineare Abbildung $(h,k) \mapsto f''(x,h,k) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ ist symmetrisch.)

Beweis: Die Bilinearform $(h,k) \mapsto f''(x,h,k)$ ist symmetrisch, genau dann wenn jede ihrer Komponenten $(h,k) \mapsto (f'')_j(x,h,k) = (f_j)''(x,h,k)$ symmetrisch ist. Es genügt also, die Symmetrie für die Komponentenfunktionen f_j zu beweisen, wobei ich den Index j weglasse und voraussetze, daß $f : U \to \mathbb{R}$ gilt. Zum Beweis des Satzes zeige ich, daß für alle $h, k \in \mathbb{R}^n$

$$\lim_{\substack{s\to 0\\ s>0}} \frac{f(x+sh+sk) - f(x+sh) - f(x+sk) + f(x)}{s^2} = f''(x,h,k) \tag{$*$}$$

gilt. Hieraus folgt die Behauptung, weil sich die linke Seite bei Vertauschen von h und k nicht ändert.

$f''(x,h,k)$ ist die Ableitung der Funktion $x \mapsto f'(x,h)$. Also gilt

$$f'(x+k,h) - f'(x,h) = f''(x,h,k) + R_x(h,k)\|k\|$$

mit

$$\lim_{k\to 0} R_x(h,k) = 0\,.$$

$R_x(h,k)$ ist linear bezüglich h, weil $f'(x+k,h)$, $f'(x,h)$ und $f''(x,h,k)$ linear in h sind, und es existiert eine von h und k abhängige Zahl ϑ mit $0 < \vartheta < 1$, so daß

$$\begin{aligned} &f(x+h+k) - f(x+h) - f(x+k) + f(x) \\ &\quad = f''(x,h,k) + R_x(h,\vartheta h + k)\|\vartheta h + k\| - R_x(h,\vartheta h)\|\vartheta h\| \end{aligned} \tag{$**$}$$

gilt. Zum Beweis dieser Gleichung betrachte man die Hilfsfunktion

$$F : [0,1] \to \mathbb{R}\,, \quad F(t) := f(x+th+k) - f(x+th)\,.$$

Wegen

$$F'(t) = f'(x+th+k)h - f'(x+th)h = f'(x+th+k,h) - f'(x+th,h)$$

und wegen

$$F(1) - F(0) = f(x+h+k) - f(x+h) - f(x+k) + f(x)$$

folgt nach dem Mittelwertsatz

$$F(1) - F(0) = F'(\vartheta)\,,$$

mit geeignetem $\vartheta, 0 < \vartheta < 1$, also

$$\begin{aligned} & f(x+h+k) - f(x+h) - f(x+k) + f(x) \\ & = \quad f'(x+\vartheta h+k,h) - f'(x+\vartheta h,h) \\ & = \quad \Big(f'(x+\vartheta h+k,h) - f'(x,h)\Big) - \Big(f'(x+\vartheta h,h) - f'(x,h)\Big)\,. \end{aligned}$$

Mit

$$\begin{aligned} f'(x+\vartheta h+k,h) - f'(x,h) &= f''(x,h,\vartheta h+k) + R_x(h,\vartheta h+k)\|\vartheta h+k\| \\ f'(x+\vartheta h,h) - f'(x,h) &= f''(x,h,\vartheta h) + R_x(h,\vartheta h)\|\vartheta h\| \end{aligned}$$

und mit

$$f''(x,h,\vartheta h+k) - f''(x,h,\vartheta h) = f''(x,h,k)$$

folgt $(**)$.

Sei $s > 0$. Ersetzt man in $(**)$ den Vektor k durch sk und den Vektor h durch sh , dann kann man auf der rechten Seite wegen der Bilinearität oder Linearität der Terme den Faktor s^2 herausziehen und erhält

$$\begin{aligned} & f(x+sh+sk) - f(x+sh) - f(x+sk) + f(x) \\ & = \quad s^2\Big[f''(x,h,k) + R_x\Big(h,s(\vartheta h+k)\Big)\|\vartheta h+k\| - R_x(h,s\vartheta h)\|\vartheta h\|\Big]\,. \end{aligned}$$

Wegen

$$\lim_{s\to 0} R_x\Big(h,s(\vartheta h+k)\Big) = 0\,, \quad \lim_{s\to 0} R_x(h,s\vartheta h) = 0$$

folgt $(*)$.

Beispiel:

$$f : \mathbb{R}^2 \to \mathbb{R}\,,\quad f(x_1,x_2) = x_1^2 x_2 + x_1 + x_2^3\,.$$

Die partiellen Ableitungen jeder Ordnung existieren und sind stetig, also ist f zweimal differenzierbar. Es gilt

$$\begin{aligned}
\mathrm{grad} f &= \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1x_2+1 \\ \\ x_1^2+3x_2^2 \end{pmatrix} \\
f''(x) := H &= \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_2 \partial x_1} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2x_2 & 2x_1 \\ \\ 2x_1 & 6x_2 \end{pmatrix}.
\end{aligned}$$

Höhere Ableitungen: Höhere Ableitungen definiert man *induktiv*. Die p–te Ableitung von $f : U \to \mathbb{R}^m$ ist eine Abbildung

$$f^{(p)} : U \times \underbrace{\mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{p \text{ Faktoren}} \to \mathbb{R}^m\,,$$

die man folgendermaßen aus $f^{(p-1)}$ erhält: Sind $x \in U$ und $h^{(1)}, \ldots, h^{(p)} \in \mathbb{R}^n$, dann ist $f^{(p)}$ definiert durch

$$f^{(p)}(x, h^{(1)}, \ldots, h^{(p)}) := \Big[y \mapsto f^{(p-1)}(y, h^{(1)}, \ldots, h^{(p-1)}) \Big]' \Big|_{y=x} (h^{(p)})\,.$$

$f^{(p)}$ ist linear in den letzten p Argumenten und ist total symmetrisch: Für $1 \le i \le j \le p$ gilt

$$f^{(p)}(x, \ldots, h^{(i)}, \ldots, h^{(j)}, \ldots) = f^{(p)}(x, \ldots, h^{(j)}, \ldots, h^{(i)}, \ldots)\,.$$

Ist $f^{(p)}$ stetig, dann heißt f p–mal stetig differenzierbar. Wenn $f^{(p)}$ für alle $p \in \mathbb{N}$ existiert, heißt f unendlich oft differenzierbar. Wie für f'' sieht man, daß

$$f^{(p)}(x, h^{(1)}, \ldots, h^{(p)}) = \sum_{i_1=1}^{n} \ldots \sum_{i_p=1}^{n} \frac{\partial^p f}{\partial x_{i_1} \ldots \partial x_{i_p}}(x) h_{i_1}^{(1)} \ldots h_{i_p}^{(p)}$$

gilt.

Satz (Taylorformel): Sei $U \subseteq \mathbb{R}^n$ eine offene Menge $f : U \to \mathbb{R}$ sei $(p+1)$–mal differenzierbar und die Verbindungsstrecke der beiden Punkte x und $x+h$ gehöre zu U. Dann existiert ϑ, $0 < \vartheta < 1$, mit

$$f(x+h) = f(x) + f'(x,h) + \frac{1}{2!} f''(x,h,h) + \ \ldots + \frac{1}{p!} f^{(p)}(x, \underbrace{h, \ldots, h}_{p \text{ mal}}) + R_p(x,h)\,,$$

wobei

$$R_p(x,h) = \frac{1}{(p+1)!}\, f^{(p+1)}(x+\vartheta h,\ \underbrace{h,\,\ldots,h}_{p+1 \text{ mal}})$$

sei.

Beweis: Sei $\gamma : [0,1] \to U\,,\quad \gamma(t) = x + th$. Auf $F = f \circ \gamma : [0,1] \to \mathbb{R}$ wende man den Taylorschen Satz für reelle Funktionen an:

$$F(1) = \sum_{j=0}^{p} \frac{F^{(j)}(0)}{j!} + \frac{1}{(p+1)!}\, F^{(p+1)}(\vartheta)\,.$$

Wegen

$$\begin{aligned}
F'(t) &= f'\Big(\gamma(t)\Big)\gamma'(t) = f'\Big(\gamma(t), \gamma'(t)\Big) = f'\Big(\gamma(t), h\Big)\\
F''(t) &= f''\Big(\gamma(t), h, \gamma'(t)\Big) = f''\Big(\gamma(t), h, h\Big)\\
&\vdots\\
F^{(p+1)}(t) &= f^{(p+1)}\Big(\gamma(t),\ \underbrace{h,\,\ldots,h}_{(p+1)\text{ mal}}\ \Big)
\end{aligned}$$

folgt hieraus die Behauptung. ∎

Man kann die Taylorformel auch in folgender Form schreiben:

$$\begin{aligned}
f(x+h) \;=\; & \sum_{j=0}^{p} \frac{1}{j!}\Big[\sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f(x)}{\partial x_{i_1} \ldots \partial x_{i_j}}\, h_{i_1} \ldots h_{i_j}\Big]\\
& + \frac{1}{(p+1)!} \sum_{i_1=1}^{n} \cdots \sum_{i_{p+1}=1}^{n} \frac{\partial^{p+1} f(x+\vartheta h)}{\partial x_{i_1} \ldots \partial x_{i_{p+1}}}\, h_{i_1} \ldots h_{i_{p+1}}\,.
\end{aligned}$$

Zur Abkürzung führt man die folgenden Bezeichnungen ein: Für $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ und für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ sei

$$\begin{aligned}
|\alpha| &:= \alpha_1 + \ldots + \alpha_n\\
\alpha! &:= \alpha_1! \ldots \alpha_n!\\
x^\alpha &:= x_1^{\alpha_1} \ldots x_n^{\alpha_n}\,,\\
D^\alpha f(a) &:= \frac{\partial^{|\alpha|} f}{\partial^{\alpha_1} x_1 \ldots \partial^{\alpha_n} x_n}\,(a)\,.
\end{aligned}$$

Man bezeichnet $\alpha \in \mathbb{N}_0^n$ als Multiindex und $|\alpha|$ als Länge von α. Bei vorgegebenem Multiindex α mit $|\alpha| = j$ gibt es in

$$\sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f(x)}{\partial x_{i_1} \ldots \partial x_{i_j}}\, h_{i_1} \ldots h_{i_j}$$

$\frac{j!}{\alpha_1! \dots \alpha_n!}$ Glieder, die aus dem Glied $D^\alpha f(x) h^\alpha$ durch Vertauschen der Reihenfolge, in der die Ableitungen gebildet werden, entstehen. Also folgt

$$\begin{aligned} f(x+h) &= \sum_{j=0}^{p} \sum_{|\alpha|=j} \frac{1}{\alpha!} D^\alpha f(x) h^\alpha + \sum_{|\alpha|=p+1} \frac{1}{\alpha!} D^\alpha f(x+\vartheta h) h^\alpha \\ &= \sum_{|\alpha| \le p} \frac{1}{\alpha!} D^\alpha f(x) h^\alpha + \sum_{|\alpha|=p+1} \frac{1}{\alpha!} D^\alpha f(x+\vartheta h) h^\alpha \,. \end{aligned}$$

6 Lokale Extrema, Sätze von der inversen und der impliziten Funktion.

In diesem Kapitel werden die Ergebnisse von Kapitel 5 angewandt.

6 a.) Lokale Extrema

Definition: Sei $U \subseteq \mathbb{R}^n$ offen, sei $f : U \to \mathbb{R}$ differenzierbar und sei $a \in U$. Gilt $\operatorname{grad} f(a) = 0$, dann heißt a kritischer Punkt von f.

Satz: Sei $U \subseteq \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}$ differenzierbar. Ist a lokale Extremalstelle von f, dann ist a kritischer Punkt von f.

Beweis: O.B.d.A. habe f in a ein Maximum. Dann gibt es eine Umgebung V von a mit $f(x) \le f(a)$ für alle $x \in V$. Sei $h \in \mathbb{R}^n$. Wähle $\delta > 0$ so klein, daß $a + th \in V$ ist für alle $t \in \mathbb{R}$ mit $|t| \le \delta$. Sei $F : [-\delta, \delta] \to \mathbb{R}$ definiert durch

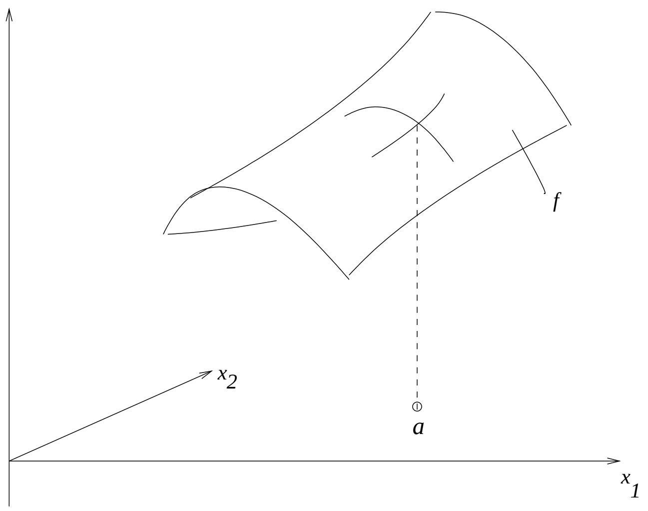
$$F(t) := f(a + th)\,.$$

Dann hat F ein Maximum in $t = 0$, also folgt

$$0 = F'(0) = f'(a)h\,.$$

Weil dies für alle $h \in \mathbb{R}^n$ gilt, resultiert $f'(a) = 0$. ■

Dies ist eine notwendige Bedingung an f für eine lokale Extremalstelle, aber keine hinreichende. Zum Beispiel ist der Sattelpunkt in der folgenden Skizze zwar ein kritischer Punkt, aber keine lokale Extremalstelle:



Wie für reelle Funktionen kann man mit Hilfe der zweiten Ableitung hinreichende Bedingungen erhalten. Hierzu benötigt man Resultate über quadratische Formen, die ich hier ohne Beweis angebe:

Vorbemerkung über quadratische Formen: 1.) Sei $Q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine bilineare Abbildung. Dann heißt die Abbildung $h \mapsto Q(h,h) : \mathbb{R}^n \to \mathbb{R}$ quadratische Form. Man teilt quadratische Formen folgendermaßen ein: Sei

$$\begin{array}{lllllll} Q(h,h) > 0 & \text{für alle} & h \neq 0 & : & \text{Dann heißt} & Q & \text{positiv definit} \\ Q(h,h) \geq 0 & \text{für alle} & h \neq 0 & : & \text{Dann heißt} & Q & \text{positiv semidefinit} \\ Q(h,h) < 0 & \text{für alle} & h \neq 0 & : & \text{Dann heißt} & Q & \text{negativ definit} \\ Q(h,h) \leq 0 & \text{für alle} & h \neq 0 & : & \text{Dann heißt} & Q & \text{negativ semidefinit.} \end{array}$$

Q heißt indefinit, wenn Q sowohl positive wie negative Werte annimmt.

2.) Eine quadratische Form kann man immer in der Form

$$Q(h,h) = \sum_{i,j=1}^{n} c_{ij} h_i h_j = h \cdot Ch$$

darstellen mit einer symmetrischen Koeffizientenmatrix

$$C = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \\ c_{n1} & \dots & c_{nn} \end{pmatrix}, \quad c_{ij} = c_{ji} \,.$$

3.) Ein Kriterium dafür, daß Q positiv definit ist, ist

$$c_{11} > 0\,, \ \det \begin{pmatrix} c_{11} & c_{12} \\ c_{21}, & c_{22} \end{pmatrix} > 0\,, \quad \det \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} > 0\,, \ \dots\,, \det(c_{ij})_{i,j=1,\dots,n} > 0\,.$$

4.) Für eine in $a \in U$ zweimal differenzierbare Abbildung $f : U \to \mathbb{R}$ ist $h \to f''(a,h,h)$ eine quadratische Form. Wegen

$$f''(a,h,h) = \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} f(a) h_i h_j$$

ist die Koeffizientenmatrix dieser quadratischen Form die Hessesche Matrix

$$H = \Big(\frac{\partial^2}{\partial x_i \partial x_j} f(a) \Big)_{i,j=1,\dots,n} .$$

Mit diesen Definitionen und Resultaten für quadratische Formen kann ein hinreichendes Kriterium für Extremalstellen formuliert werden:

Satz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ sei zweimal stetig differenzierbar und sei $a \in U$ kritischer Punkt von f. Ist dann die quadratische Form $f''(a, h, h)$

(i) positiv definit, so ist a Minimalstelle von f

(ii) negativ definit, so ist a Maximalstelle von f

(iii) indefinit, so ist a keine Extremalstelle von f.

Beweis: Aus der Taylorformel ergibt sich

$$f(x) = f(a) + f'(a, x-a) + \frac{1}{2} f''\Big(a + \vartheta(x-a), x-a, x-a\Big)\,,$$

mit geeignetem $0 < \vartheta < 1$, also, wegen $f'(a) = 0$,

$$\begin{aligned} f(x) &= f(a) + \frac{1}{2} f''\Big(a + \vartheta(x-a), x-a, x-a\Big) \\ &= f(a) + \frac{1}{2} f''(a, x-a, x-a) + R(x, x-a, x-a)\,, \end{aligned} \qquad (*)$$

mit

$$R(x, h, k) = \frac{1}{2} f''\Big(a + \vartheta(x-a), h, k\Big) - \frac{1}{2} f''(a, h, k\Big)\,.$$

Es ist $\mathbb{R}(a, h, k) = 0$, und es gilt sogar, daß zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$|R(x, h, h)| \le \varepsilon \|h\|^2\,, \qquad (**)$$

für alle $x \in U$ mit $\|x - a\| < \delta$ und alle $h \in \mathbb{R}^n$. Zum Beweis wähle man $r > 0$ so klein, daß die abgeschlossene Kugel $K_r(a) = \{x \in \mathbb{R}^n \,\big|\, \|x - a\| \le r\}$ ganz zu U gehört. Nach Voraussetzung ist $(x, h) \mapsto f''(x, h, h) : U \times \mathbb{R}^n \to \mathbb{R}$ stetig, also ist diese Funktion auf der abgeschlossenen und beschränkten, folglich kompakten Teilmenge

$$K_r(a) \times \{h \in \mathbb{R}^n \,\big|\, \|h\| = 1\}$$

sogar gleichmäßig stetig. Dies bedeutet, daß zu $\varepsilon > 0$ eine Zahl $\delta > 0$ existiert mit

$$\frac{1}{2} \Big| f''(x, h, h) - f''(a, h, h) \Big| < \varepsilon$$

für alle $h \in \mathbb{R}^n$ mit $\|h\| = 1$ und alle $x \in U$ mit $\|x - a\| < \delta$. Weil für $0 < \vartheta < 1$ auch $\|a + \vartheta(x - a) - a\| = \vartheta \|x - a\| < \delta$ gilt, folgt

$$\begin{aligned} |R(x, h, h)| &= \|h\|^2 \Big| R\Big(x, \frac{h}{\|h\|}, \frac{h}{\|h\|}\Big) \Big| \\ &= \|h\|^2 \frac{1}{2} \Big| f''\Big(a + \vartheta(x-a), \frac{h}{\|h\|}, \frac{h}{\|h\|}\Big) - f''\Big(a, \frac{h}{\|h\|}, \frac{h}{\|h\|}\Big) \Big| < \varepsilon \|h\|^2\,. \end{aligned}$$

Dies beweist $(**)$.

Sei nun $h \mapsto f''(a,h,h)$ eine positiv definite quadratische Form. Dann gilt $f''(a,h,h) > 0$ für alle $h \in \mathbb{R}^n$ mit $h \neq 0$, und da die stetige Abbildung $h \mapsto f''(a,h,h) : \mathbb{R}^n \to \mathbb{R}$ auf der abgeschlossenen und beschränkten, also kompakten Menge $\{h \in \mathbb{R}^n \,\big|\, \|h\| = 1\}$ ihr Minimum an einer Stelle h_0 annimmt, folgt für alle $h \in \mathbb{R}^n$

$$f''(a,h,h) = \|h\|^2 f''\Big(a, \frac{h}{\|h\|}, \frac{h}{\|h\|}\Big) \geq \|h\|^2 \min_{\|\eta\|=1} f''(a,\eta,\eta) = c\|h\|^2\,,$$

mit

$$c = \min_{\|\eta\|=1} f''(a,\eta,\eta) = f''(a,h_0,h_0) > 0\,.$$

Wählt man nun $\varepsilon = c/4$, dann folgt hieraus und aus $(*), (**)$, daß $\delta > 0$ existiert mit

$$\begin{aligned} f(x) - f(a) &= \frac{1}{2} f''(a, x-a, x-a) + R(x, x-a, x-a) \\ &\geq \frac{c}{2}\,\|x-a\|^2 - \frac{c}{4}\,\|x-a\|^2 = \frac{c}{4}\,\|x-a\|^2 \geq 0\,, \end{aligned}$$

für alle x mit $\|x-a\| < \delta$, also ist a ein lokales Minimum.

Entsprechend beweist man, daß bei negativ definitem $f''(a,h,h)$ ein lokales Maximum vorliegt. Wenn $f''(a,h,h)$ indefinit ist, gibt es $h_0 \in \mathbb{R}^n$, $k_0 \in \mathbb{R}^n$ mit $\|h_0\| = \|k_0\| = 1$ und mit

$$f''(a,h_0,h_0) > 0\,, \quad f''(a,k_0,k_0) < 0\,.$$

Hieraus folgt, daß auf der Geraden mit Richtungsvektor h_0 bzw. k_0 für genügend kleines $\|x-a\|, x \neq a$ die Differenz $f(x) - f(a)$ positiv bzw. negativ ist. Dies beweist man wie oben. Also ist a kein lokales Extremum. ■

Beispiel: Sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x,y) = 6xy - 3y^2 - 2x^3$. Für jeden kritischen Punkt (x,y) gilt

$$f'(x,y) = \begin{pmatrix} \dfrac{\partial f}{\partial x}(x,y) \\[2mm] \dfrac{\partial f}{\partial y}(x,y) \end{pmatrix} = \begin{pmatrix} 6y - 6x^2 \\ 6x - 6y \end{pmatrix} = 0\,.$$

Hieraus können die kritischen Punkte bestimmt werden. Man erhält für die kritischen Punkte $(x,y) = (0,0)$ und $(x,y) = (1,1)$.

Um festzustellen, ob diese Punkte Extremalpunkte sind, muß die Hessesche Matrix

$$H(x,y) = \begin{pmatrix} -12x & 6 \\ 6 & -6 \end{pmatrix}$$

von f an den kritischen Punkten untersucht werden. Die durch die Matrix

$$H(0,0) = \begin{pmatrix} 0 & 6 \\ 6 & -6 \end{pmatrix}$$

definierte quadratische Form $f''(0,0,h,h)$ ist indefinit. Denn es gilt für $h = (1,1)$

$$f''(0,0,h,h) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 6 \\ 6 & -6 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ 0 \end{pmatrix} = 6$$

und für $h = (0,1)$

$$f''(0,0,h,h) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 6 \\ 6 & -6 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ -6 \end{pmatrix} = -6\,,$$

also ist $(0,0)$ keine Extremalstelle. Dagegen ist die durch die Matrix

$$H(1,1) = \begin{pmatrix} -12 & 6 \\ 6 & -6 \end{pmatrix}$$

definierte quadratische Form $f(1,1,h,h)$ negativ definit. Denn nach dem oben angegebenen Kriterium ist die Matrix $-H(1,1)$ positiv definit wegen $12 > 0$ und

$$\det \begin{pmatrix} 12 & -6 \\ -6 & 6 \end{pmatrix} = 72 - 36 > 0\,.$$

Somit ist $H(1,1)$ negativ definit und $(1,1)$ ein lokales Maximum.

6 b.) Lokale Umkehrbarkeit von Abbildungen

Früher wurde gezeigt, daß wenn f invertierbar und in einem Punkt a differenzierbar ist, wenn außerdem $\det f'(a) \neq 0$ gilt und die Inverse g in $b = f(a)$ stetig ist, dann ist g in b differenzierbar. Man kann sich fragen, ob aus $\det f'(a) \neq 0$ bereits folgt, daß f in einer Umgebung von a invertierbar ist. Das folgende Beispiel zeigt, daß dies im Allgemeinen nicht richtig ist.

Gegenbeispiel: Sei $f : (-1,1) \to \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} x + 3x^2 \sin \dfrac{1}{x}\,, & x \neq 0 \\ 0\,, & x = 0\,. \end{cases}$$

f ist für alle $|x| < 1$ differenzierbar, $f'(x)$ ist beschränkt, $f'(0) = 1$, aber f ist in keiner Umgebung des Nullpunktes invertierbar.

Jedoch ist in diesem Beispiel f' nicht stetig im Nullpunkt, weil der Grenzwert von

$$f'(x) = 1 + 6x \sin \frac{1}{x} - 3 \cos \frac{1}{x}$$

an der Stelle 0 nicht existiert. Setzt man auch noch die Stetigkeit der Ableitung voraus, dann kann man folgern, daß eine lokale Inverse existiert:

Satz: Sei $U \subseteq \mathbb{R}^n$ offen und $a \in U$, $f : U \to \mathbb{R}^n$ sei stetig differenzierbar, und es sei $\det f'(a) \neq 0$. Sei $b = f(a)$. Dann existieren offene Mengen $V, W \subseteq \mathbb{R}^n$ mit $a \in V$, $b \in W$, so daß $f : V \to W$ bijektiv ist, und so daß die Inverse $g : W \to V$ stetig differenzierbar ist. (Natürlich gilt dann $g'(y) = [f'(g(y))]^{-1}$.)

Beweis: Setze $A := f'(a)$ und $\lambda := \frac{1}{4\|A^{-1}\|}$. Die Inverse A^{-1} existiert, weil nach Voraussetzung $\det A \neq 0$ ist. Da nach Vorasusetzung $f' : U \mapsto L(\mathbb{R}^n, \mathbb{R}^n)$ stetig ist, existiert eine offene Kugel V um a mit

$$\|f'(x) - A\| < 2\lambda$$

für alle $x \in V$.

1.) Zunächst soll gezeigt werden: Für beliebige $x, x+h \in V$ gilt

$$\|f(x+h) - f(x) - Ah\| \leq \frac{1}{2} \|Ah\|, \tag{$*$}$$

also, wegen der umgekehrten Dreiecksungleichung,

$$\begin{aligned} \|f(x+h) - f(x)\| &\geq \|Ah\| - \|f(x+h) - f(x) - Ah\| \geq \frac{1}{2}\|Ah\| \\ &= 2\lambda\|A^{-1}\| \, \|Ah\| \geq 2\lambda\|A^{-1}Ah\| = 2\lambda\|h\|, \end{aligned} \tag{$**$}$$

woraus dann folgt, daß f in V injektiv ist.

Hierzu definiere man $F : [0,1] \to \mathbb{R}^n$ durch

$$F(t) = f(x+th) - tAh.$$

Da V eine Kugel ist, gehört mit x und $x+h$ auch die Verbindungsstrecke $\{x + th \,\big|\, 0 \leq t \leq 1\}$ zu V, und es gilt

$$\begin{aligned} \|F'(t)\| &= \|f'(x+th)h - Ah\| \leq \|f'(x+th) - A\| \, \|h\| \leq 2\lambda\|h\| \\ &= 2\lambda\|A^{-1}Ah\| \leq 2\lambda\|A^{-1}\| \, \|Ah\| = \frac{1}{2} \|Ah\|. \end{aligned}$$

Aus dem Schrankensatz folgt nun $\|F(1) - F(0)\| \leq \frac{1}{2} \|Ah\|$, also $(*)$.

Somit existiert die Inverse $g : W \to V$ mit $W = f(V)$.

2.) Es ist zu zeigen, daß W offen ist. Sei $x_0 \in V$ und sei K_r eine offene Kugel mit Mittelpunkt x_0 und Radius $r > 0$, so daß $\overline{K}_r \subseteq V$ ist. Es soll gezeigt werden, daß $f(K_r)$ eine offene Kugel um $f(x_0)$ mit Radius λr enthält.

Hierzu wähle $y \in \mathbb{R}^n$ mit $\|y - f(x_0)\| < \lambda r$. Es wird das Urbild von y unter f in K_r konstruiert. Wähle $x^* \in \overline{K}_r$ mit

$$\|y - f(x^*)\| = \min_{x \in \overline{K}_r} \|y - f(x)\| .$$

Es wird sich zeigen, daß x^* das Urbild ist. Zunächst muß gezeigt werden, daß ein solches x^* existiert. Hierzu beachte man, daß die durch $\phi(x) := \|y - f(x)\|$ definierte Funktion $\phi : \overline{K}_r \to \mathbb{R}$ stetig ist. Denn es gilt für $z \in \overline{K}_r$

$$\begin{aligned} \lim_{x \to z} |\phi(x) - \phi(z)| &= \lim_{x \to z} \Big| \|y - f(x)\| - \|y - f(z)\| \Big| \\ &\leq \lim_{x \to z} \Big\| \Big(y - f(x)\Big) - \Big(y - f(z)\Big) \Big\| \leq \lim_{x \to z} \|f(x) - f(z)\| = 0 , \end{aligned}$$

wobei wieder die umgekehrte Dreiecksungleichung verwendet wurde. Also nimmt ϕ auf der kompakten Menge $\overline{K}_r$ das Minimum in mindestens einem Punkt x^* an. Es soll nun gezeigt werden, daß $\|y - f(x^*)\| = \phi(x^*) = 0$ ist. Hierzu zeigt man, daß ein $\tilde{x} \in K_r$ existieren würde mit $\|y - f(\tilde{x})\| < \|y - f(x^*)\|$, falls $\|y - f(x^*)\| \neq 0$ wäre. Dies ist ein Widerspruch zur Definition von x^*.

Sei $h = A^{-1}(y - f(x^*))$. Für hinreichend kleines $t \in (0, 1)$ ist

$$\tilde{x} = x^* + th \in V .$$

Es gilt nun nach $(*)$ wegen $x^*, x^* + th \in V$:

$$\begin{aligned} \|f(x^* + th) - y\| &= \|f(x^* + th) - f(x^*) - Ath + f(x^*) - y + Ath\| \\ &\leq \|f(x^* + th) - f(x^*) - Ath\| + \|f(x^*) - y + Ath\| \\ &\leq \frac{1}{2} \|Ath\| + \|f(x^*) - y + Ath\| \\ &= \frac{1}{2} \|A\Big[tA^{-1}\Big(y - f(x^*)\Big)\Big]\| + \|f(x^*) - y + A\Big[tA^{-1}\Big(y - f(x^*)\Big)\Big]\| \\ &= \frac{1}{2} t\|y - f(x^*)\| + \|(1-t)\Big(f(x^*) - y\Big)\| = \Big(1 - \frac{t}{2}\Big)\|y - f(x^*)\| < \|y - f(x^*)\| , \end{aligned}$$

falls $y \neq f(x^*)$. Weil diese Ungleichung für alle $t \in (0, 1)$ mit $x^* + th \in V$ gilt, bleibt nur noch zu zeigen, daß $\tilde{x} = x^* + th \in \overline{K}_r$ ist für alle hinreichend kleines t. Dann ist der Widerspruch konstruiert. Hierzu genügt es zu zeigen, daß x^* nicht auf dem Rand der Kugel K_r liegt. Für einen Randpunkt x von K_r gilt $\|x - x_0\| = r$. Aus $(**)$ folgt somit

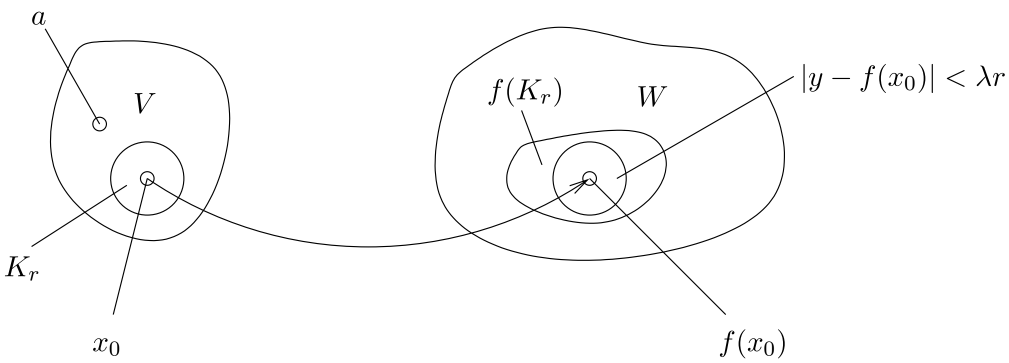
$$2\lambda r \leq \|f(x) - f(x_0)\| \leq \|y - f(x)\| + \|y - f(x_0)\| < \phi(x) + \lambda r ,$$

also

$$\phi(x_0) = \|y - f(x_0)\| < \lambda r < \phi(x)\,,$$

so daß ϕ in keinem Randpunkt das Minimum annehmen kann. Also ist x^* innerer Punkt von K_r .

Damit ist bewiesen, daß $y = f(x^*)$ gilt, somit gehören alle $y \in \mathbb{R}^n$ mit $\|y - f(x_0)\| < \lambda r$ zu $f(V)$, also ist $W = f(V)$ offen, weil $x_0 \in V$ beliebig gewählt war.



3.) Es bleibt zu zeigen, daß die Inverse $g : W \to V$ stetig differenzierbar ist. Da f stetig differenzierbar ist, sind die partiellen Ableitungen $\frac{\partial f_i}{\partial x_j} : V \to \mathbb{R}$ stetig. Da $\det f'(x) = \det(\frac{\partial f_i}{\partial x_j})_{i,j=1,\ldots,n}$ eine Summe aus Produkten der $\frac{\partial f_i}{\partial x_j}(x)$ ist, ist $x \mapsto \det f'(x) : V \to \mathbb{R}$ stetig, also gibt es eine Umgebung von a in der $\det f'(x) \neq 0$ ist, in der also $f'(x)$ invertierbar ist. Man verkleinere die Umgebung V soweit, daß $f'(x)$ invertierbar ist für alle $x \in V$. Nach einem Satz aus Abschnitt 5 c.) folgt dann, daß die Inverse $g : W \to V$ in jedem Punkt von W differenzierbar ist, wenn sie stetig ist. Die Stetigkeit von g folgt unmittelbar aus $(**)$. Denn es gilt für $y, y + k \in W$:

$$\|k\| = \|y + k - y\| = \|f\Big(g(y+k)\Big) - f\Big(g(y)\Big)\| \geq 2\lambda \|g(y+k) - g(y)\|\,,$$

also

$$\lim_{k\to 0} \|g(y+k) - g(y)\| \leq \frac{1}{2\lambda} \lim_{k\to 0} \|k\| = 0\,.$$

Somit ist g stetig, also differenzierbar mit $g'(y) = [f'(g(y))]^{-1}$.

Aus dieser Formel folgt auch, daß g' stetig ist. Denn in Abschnitt 5 e.) wurde gezeigt, daß g' stetig ist, wenn die Elemente der Matrix $g'(y)$, also die partiellen Ableitungen von g , stetige Funktionen von y sind. Weil $g'(y)$ die Inverse der Matrix $f'(g(y))$ ist, werden die Elemente von $g'(y)$ aus den Elementen von $f'(g(y))$ durch Bildung von Determinanten und Quotienten berechnet (Cramersche Regel!), also sind die Elemente von $g'(y)$ stetige Funktionen der Elemente der Matrix $f'(g(y))$, die selber wieder stetige Funktionen von y

sind, weil f' und g stetig sind. Also ist g stetig differenzierbar. ■

Beispiel: $f : \mathbb{R}^3 \to \mathbb{R}^3$ sei erklärt durch

$$\begin{aligned} f_1(x_1, x_2, x_3) &= x_1 + x_2 + x_3 \\ f_2(x_1, x_2, x_3) &= x_2x_3 + x_3x_1 + x_1x_2 \\ f_3(x_1, x_2, x_3) &= x_1x_2x_3\,. \end{aligned}$$

Weil die partiellen Ableitungen alle existieren und stetig sind, ist f stetig differenzierbar mit

$$f'(x) = \begin{pmatrix} 1 & 1 & 1 \\ x_3 + x_2 & x_3 + x_1 & x_2 + x_1 \\ x_2x_3 & x_1x_3 & x_1x_2 \end{pmatrix},$$

also

$$\begin{aligned} \det f'(x) &= \begin{vmatrix} 1 & 0 & 0 \\ x_3 + x_2 & x_1 - x_2 & x_1 - x_3 \\ x_2x_3 & (x_1 - x_2)x_3 & (x_1 - x_3)x_2 \end{vmatrix} \\ &= (x_1 - x_2)(x_1 - x_3)x_2 - (x_1 - x_2)(x_1 - x_3)x_3 \\ &= (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)\,. \end{aligned}$$

Sei also $b = f(a)$ mit $(a_1 - a_2)(a_1 - a_3)(a_2 - a_3) \neq 0$. Dann gilt es Umgebungen V von a und W von b, so daß das Gleichungssystem

$$\begin{aligned} y_1 &= x_1 + x_2 + x_3 \\ y_2 &= x_2x_3 + x_3x_1 + x_1x_2 \\ y_3 &= x_1x_2x_3 \end{aligned}$$

zu jedem $y \in W$ eine eindeutige Lösung $x \in V$ hat.

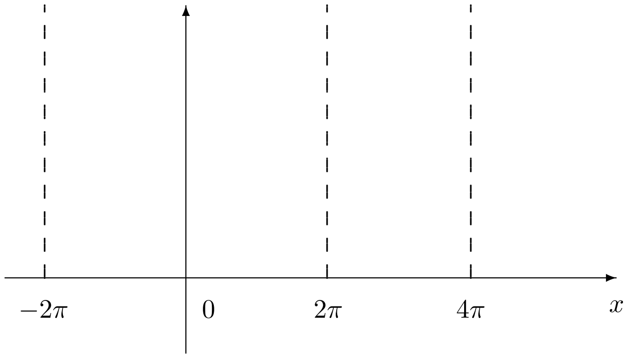
Man beachte aber, daß aus der lokalen Invertierbarkeit nicht die globale folgt. Man sieht dies an folgendem Beispiel: Sei $f : \{(x, y) \in \mathbb{R}^2 \mid y > 0\} \to \mathbb{R}^2$ definiert durch

$$\begin{aligned} f_1(x, y) &= y \cos x \\ f_2(x, y) &= y \sin x\,. \end{aligned}$$

f ist stetig differenzierbar mit

$$\det f'(x, y) = \begin{vmatrix} -y \sin x & \cos x \\ y \cos x & \sin x \end{vmatrix} = -y \sin^2 x - y \cos^2 x = -y \neq 0$$

für alle (x, y) aus dem Definitionsbereich. Also ist f in jedem Punkt lokal invertierbar, jedoch nicht global. Denn sei $b = f(a)$ mit $a = (a_1, a_2)$. Dann gilt auch $b = f(a_1 + 2\pi m, a_2)$, $\;m \in \mathbb{Z}$, weil f bezüglich der x–Koordinate 2π–periodisch ist.



6 c.) Implizite Funktionen

Es sei eine Abbildung $f : \mathbb{R}^{n+m} \to \mathbb{R}^n$ gegeben mit den Komponenten f_j, also $f = (f_1, \ldots, f_n)$, und es sei $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ gegeben. Es liegt nahe zu fragen, ob $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ so bestimmt werden kann, daß

$$\begin{aligned} f_1(x_1, \ldots, x_n, y_1, \ldots, y_m) &= 0 \\ &\vdots \\ f_n(x_1, \ldots, x_n, y_1 \ldots, y_m) &= 0 \end{aligned}$$

gilt. Dies sind n Gleichungen zur Bestimmung von n Unbekannten $x_1, \ldots, x_n$. Zunächst betrachte man den Fall, daß $f = A : R^{n+m} \to \mathbb{R}^n$ eine lineare Abbildung ist,

$$A(x,y) = \begin{pmatrix} A_1(x,y) \\ \vdots \\ A_n(x,y) \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n + b_{11}y_1 + \cdots + b_{1m}y_m \\ \vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n + b_{n1}y_1 + \ldots + b_{nm}y_m \end{pmatrix}.$$

A habe folgende Eigenschaft:

$$A(h,0) = 0 \implies h = 0\,.$$

Dies ist genau dann erfüllt, wenn die Matrix

$$\begin{pmatrix} a_{11} \ldots a_{1n} \\ \vdots \\ a_{n1} \ldots a_{nn} \end{pmatrix} = \begin{pmatrix} \frac{\partial A_1}{\partial x_1} \cdots \frac{\partial A_1}{\partial x_n} \\ \vdots \\ \frac{\partial A_n}{\partial x_1} \cdots \frac{\partial A_n}{\partial x_n} \end{pmatrix}$$

invertierbar ist, also genau dann wenn

$$\det \left(\frac{\partial A_j}{\partial x_i} \right)_{\substack{i=1,\,\dots,n \\ j=1,\,\dots,n}} \neq 0$$

ist. Unter dieser Bedingung ist

$$h \mapsto Ch := A(h,0) : \mathbb{R}^n \to \mathbb{R}^n$$

eine invertierbare lineare Abbildung, folglich hat das Gleichungssystem

$$A(h,k) = A(h,0) + A(0,k) = Ch + A(0,k) = 0$$

für jedes $k \in \mathbb{R}^m$ die eindeutig bestimmte Lösung

$$h = \varphi(k) := -C^{-1}A(0,k)\,.$$

Für $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ gilt

$$A(\varphi(k), k) = 0\,,$$

für alle $k \in \mathbb{R}^m$. Man sagt, φ sei durch diese Gleichung implizit gegeben. Der Satz über implizit gegebene Funktionen betrifft dieselbe Situation für stetig differenzierbare Abbildungen f, die nicht notwendig linear sein müssen:

Satz (über implizite Funktionen): Sei $D \subseteq \mathbb{R}^{n+m}$ offen, $f : D \to \mathbb{R}^n$ sei stetig differenzierbar. Sei $(a,b) \in \mathbb{R}^n \times \mathbb{R}^m$ mit $f(a,b) = 0$ und mit

$$\det \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a,b) & \dots & \frac{\partial f_1}{\partial x_n}(a,b) \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1}(a,b) & \dots & \frac{\partial f_n}{\partial x_n}(a,b) \end{pmatrix} \neq 0\,. \tag{$*$}$$

Dann gibt es eine Umgebung $U \subseteq \mathbb{R}^m$ von b und eine eindeutig bestimmte stetig differenzierbare Abbildung $\varphi : U \to \mathbb{R}^n$ mit $\varphi(b) = a$ und mit

$$f\Big(\varphi(y), y\Big) = 0$$

für alle $y \in U$.

Bemerkung: Sei $A = f'(a,b) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. Die Bedingung $(*)$ ist äquivalent zur Bedingung

$$A(h,0) = 0 \Longrightarrow h = 0\,.$$

Beweis des Satzes: Betrachte die Abbildung $F : D \to \mathbb{R}^{n+m}$,

$$(x,y) \mapsto F(x,y) := \Big(f(x,y), y\Big) \in \mathbb{R}^{n+m} .$$

Es gilt $F(a,b) = (0,b)$. Es soll gezeigt werden, daß F die Voraussetzungen des Satzes über lokale Umkehrbarkeit erfüllt. Aus diesem Satz folgt dann, daß es Umgebungen $V \subseteq \mathbb{R}^{n+m}$ von (a,b) und $W \subseteq \mathbb{R}^{n+m}$ von $(0,b)$ gibt, so daß $F : V \to W$ bijektiv ist und eine stetig differenzierbare Inverse $F^{-1} : W \to V$ besitzt. Die Inverse ist von der Form

$$F^{-1}(z,w) = \Big(\phi(z,w), w\Big) ,$$

mit einer stetig differenzierbaren Funktion $\phi : W \to \mathbb{R}^n$. Aus W und ϕ erhält man die gesuchte Umgebung $U \subseteq \mathbb{R}^m$ von b und die gesuchte Funktion $\varphi : U \to \mathbb{R}^n$ durch die Definitionen

$$U = \{w \in \mathbb{R}^m \mid (0,w) \in W\}$$

und

$$\varphi(w) := \phi(0,w) , \quad w \in U .$$

Denn wegen $(0,b) \in W$ ist U eine Umgebung von b in $\mathbb{R}^m$, und für alle $w \in U$ gilt

$$(0,w) = F\Big(F^{-1}(0,w)\Big) = F\Big(\phi(0,w), w\Big) = F\Big(\varphi(w), w\Big) = \Big(f\Big(\varphi(w), w\Big), w\Big) ,$$

also

$$f\Big(\varphi(w), w\Big) = 0 .$$

Also genügt es, die Voraussetzungen des Satzes über lokale Umkehrbarkeit nachzuprüfen. Weil f nach Voraussetzung stetig differenzierbar ist, folgt aus der Definition von F sofort, daß alle partiellen Ableitungen von F existieren und stetig sind. Also ist F stetig differenzierbar. Der Satz über lokale Umkehrbarkeit kann somit angewandt werden, wenn $F'[a,b]$ invertierbar ist. Mit $A = f'(a,b)$ gilt für $(h,k) \in \mathbb{R}^{n+m}$

$$(h,k) \mapsto F'[a,b](h,k) = \Big(A(h,k), k\Big) \in \mathbb{R}^{n+m} . \qquad (*)$$

Denn da f differenzierbar ist, folgt

$$f(a+h, b+k) = f(a,b) + A(h,k) + r(h,k)\|(h,k)\| ,$$

mit $\lim_{(h,k)\to 0} r(h,k) = 0$. Also gilt

$$\begin{aligned} F(a+h, b+k) &= \Big(f(a+h,b+k) ,\ \ b+k\Big) \\ &= \Big(f(a,b), b\Big) + \Big(A(h,k), k\Big) + \Big(r(h,k)\|(h,k)\|, 0\Big) \\ &= F(a,b) + \Big(A(h,k), k\Big) + \Big(r(h,k), 0\Big)\|(h,k)\| , \end{aligned}$$

mit $\lim_{(h,k)\to 0}\Big(r(h,k),0\Big) = 0$. Dies beweis $(*)$. Hieraus folgt, daß $F'[a,b]$ invertierbar ist. Denn aus

$$F'[a,b](h,k) = \Big(A(h,k),k\Big) = (0,0)$$

resultiert $k=0$, also $A(h,0)=0$, somit $h=0$. Also besteht der Nullraum der linearen Abbildung $F'[a,b]$ nur aus der Menge $\{0\}$, also ist die Abbildung invertierbar. ∎

Man kann auch die Ableitung der Funktion φ berechnen. Nach der Kettenregel gilt für die Ableitung $\frac{d}{dy} f(\varphi(y),y)$ der Funktion $y \mapsto f(\varphi(y),y)$:

$$\begin{aligned} 0 = \frac{d}{dy} f\Big(\varphi(y),y\Big) &= \Big(\frac{\partial}{\partial x} f\Big(\varphi(y),y\Big), \frac{\partial}{\partial y} f\Big(\varphi(y),y\Big)\Big) \begin{pmatrix} \varphi'(y) \\ I_{m\times m} \end{pmatrix} \\ &= \frac{\partial}{\partial x} f\Big(\varphi(y),y\Big) \circ \varphi'(y) + \frac{\partial}{\partial y} f\Big(\varphi(y),y\Big) \end{aligned}$$

mit der Einheitsmatrix $I_{m\times m}$ auf $\mathbb{R}^m$. Hieraus folgt

$$\varphi'(y) = -\Big[\frac{\partial}{\partial x} f\Big(\varphi(y),y\Big)\Big]^{-1} \circ \frac{\partial}{\partial y} f\Big(\varphi(y),y\Big) ,$$

mit

$$\begin{aligned} \frac{\partial}{\partial x} f(x,y) &= \Big(\frac{\partial f_j}{\partial x_i}(x,y)\Big)_{j=1,\dots,n;\, i=1,\dots,n} \\ \frac{\partial}{\partial y} f(x,y) &= \Big(\frac{\partial f_j}{\partial y_i}(x,y)\Big)_{j=1,\dots,n;\, i=1,\dots,m} \end{aligned}$$

Beispiele 1.) Sei eine Gleichung

$$f(x_1, \dots, x_n) = 0$$

gegeben mit stetig differenzierbarem $f:\mathbb{R}^n \to \mathbb{R}$. Zu gegebenen $x_1, \dots, x_{n-1}$ ist x_n gesucht, so daß diese Gleichung erfüllt ist. Angenommen, es existiere $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ mit

$$f(a_1, \dots, a_n) = 0$$

und mit

$$\frac{\partial f}{\partial x_n}(a_1, \dots, a_n) \neq 0 .$$

Dann existiert eine Umgebung $U \subseteq \mathbb{R}^{n-1}$ von $(a_1, \dots, a_{n-1})$, so daß zu jedem $(x_1, \dots, x_{n-1}) \in U$ ein eindeutiges $x_n = \varphi(x_1, \dots, x_{n-1})$ aus einer Umgebung von a_n existiert mit

$$f(x_1 \dots, x_{n-1}, x_n) = 0 .$$

Für die Ableitung von φ gilt

$$\begin{aligned}
\mathrm{grad}\varphi(x_1,\,\ldots,x_{n-1}) &= \frac{-1}{\frac{\partial}{\partial x_n}f(x_1,\,\ldots,x_n)}\,\mathrm{grad}_{n-1}f(x_1,\,\ldots,x_n) \\
&= \frac{-1}{\frac{\partial}{\partial x_n}f}\begin{pmatrix} \dfrac{\partial}{\partial x_1}f \\ \vdots \\ \dfrac{\partial}{\partial x_{n-1}}f \end{pmatrix}.
\end{aligned}$$

2.) $F:\mathbb{R}^3\to\mathbb{R}^2$ sei definiert durch

$$\begin{aligned}
f_1(x,y,z) &= 3x^2+xy-z-3 \\
f_2(x,y,z) &= 2xz+y^3+xy\,.
\end{aligned}$$

Es gilt $f(1,0,0)=0$. Zu gegebenen $z\in\mathbb{R}$ aus einer Umgebung des Nullpunktes ist $(x,y)\in\mathbb{R}^2$ in einer Umgebung von $(1,0)$ gesucht so daß $f(x,y,z)=0$ gilt. Es ist

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 6x+y & x \\ 2z+y & 3y^2+x \end{pmatrix},$$

also

$$\begin{vmatrix} \dfrac{\partial f_1}{\partial x}(1,0,0) & \dfrac{\partial f_1}{\partial y}(1,0,0) \\ \dfrac{\partial f_2}{\partial x}(1,0,0) & \dfrac{\partial f_2}{\partial y}(1,0,0) \end{vmatrix} = \begin{vmatrix} 6 & 1 \\ 0 & 1 \end{vmatrix} = 6\neq 0\,,$$

also kann eine genügend kleine Zahl $\delta>0$ und eine Funktion $\varphi:(-\delta,\delta)\to\mathbb{R}^2$ gefunden werden mit $f(\varphi_1(z),\varphi_2(z),z)=0$ für alle z mit $|z|<\delta$. Es gilt

$$\begin{aligned}
\varphi'(z) &= -\begin{pmatrix} 6x+y & x \\ 2z+y & 3y^2+x \end{pmatrix}^{-1}\begin{pmatrix} -1 \\ 2x \end{pmatrix} \\
&= \frac{-1}{(6x+y)(3y^2+x)-x(2z+y)}\begin{pmatrix} 3y^2+x & -x \\ -(2z+y) & 6x+y \end{pmatrix}\begin{pmatrix} -1 \\ 2x \end{pmatrix} \\
&= \frac{-1}{(6x+y)(3y^2+x)-x(2z+y)}\begin{pmatrix} -3y^2-x-2x^2 \\ +(2z+y)+12x^2+2xy \end{pmatrix},
\end{aligned}$$

mit $x = \varphi_1(z)$ und mit $y = \varphi_2(z)$. Insbesondere gilt

$$\varphi'(0) = -\frac{1}{6} \begin{pmatrix} -3 \\ \\ 12 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -2 \end{pmatrix} ,$$

wegen $\varphi(0) = (1, 0)$, also $\varphi_1(0) = 1, \ \varphi_2(0) = 0$.