Prof. Dr. H. Brenner **Osnabrück SS 2014**

Analysis II

Vorlesung 46

Totale Differenzierbarkeit und partielle Ableitungen

Im Folgenden wollen wir den Zusammenhang zwischen Richtungsableitungen, partiellen Ableitungen und dem totalen Differential verstehen.

Totale Differenzierbarkeit impliziert richtungsweise Differenzierbarkeit.

PROPOSITION 46.1. *Seien V und W endlichdimensionale $\mathbb{K}$-Vektorräume, $G \subseteq V$ eine offene Teilmenge, und $\varphi\colon G \to W$ eine im Punkt $P \in G$ differenzierbare Abbildung. Dann ist φ in P in jede Richtung v differenzierbar, und es gilt*

$$(D_v\varphi)(P) = (D\varphi)_P(v)\ .$$

Beweis. Da $(D\varphi)_P$ eine lineare Abbildung von V nach W ist, liefert die Anwendung dieser Abbildung auf einen Vektor $v \in V$ einen Vektor in $(D\varphi)_P(v) \in W$. Nach Voraussetzung haben wir

$$\varphi(P+v) = \varphi(P) + (D\varphi)_P(v) + ||\,v\,|| \cdot r(v)$$

(mit den üblichen Bedingungen an r). Insbesondere gilt für (hinreichend kleines) $s \in \mathbb{K}$

$$\varphi(P+sv) = \varphi(P) + s\,(D\varphi)_P(v) + |s| \cdot ||\,v\,|| \cdot r(sv)\,.$$

Also gilt

$$\begin{aligned}\lim_{s\to 0, s\neq 0} \frac{\varphi(P+sv)-\varphi(P)}{s} &= \lim_{s\to 0, s\neq 0} \frac{s\,(D\varphi)_P(v) + |s| \cdot ||\,v\,|| \cdot r(sv)}{s} \\ &= \lim_{s\to 0, s\neq 0} \left((D\varphi)_P(v) + \frac{|s|}{s}\,||\,v\,|| \cdot r(sv) \right) \\ &= (D\varphi)_P(v)\,,\end{aligned}$$

da $\lim_{s\to 0} r(sv) = 0$ und der Ausdruck $\frac{|s|}{s}\,||\,v\,||$ beschränkt ist. □

KOROLLAR 46.2. *Sei $G \subseteq K^n$ offen und $\varphi\colon G \to K^m$ eine in $P \in G$ differenzierbare Abbildung. Dann ist φ in P partiell differenzierbar, und das totale Differential wird bezüglich der Standardbasis durch die Jacobi-Matrix*

$$\left(\frac{\partial \varphi_j}{\partial x_i}(P) \right)_{1\leq i\leq n, 1\leq j\leq m}$$

Beweis. Dies folgt aus Proposition 46.1, Lemma 44.2 und daraus, dass eine lineare Abbildung auf einer Basis festgelegt ist. □

Wir wollen umgekehrt ein handliches Kriterium für die totale Differenzierbarkeit angeben. Vor dem Beweis der nächsten Aussage erinnern wir an die Mittelwertabschätzung für differenzierbare Kurven: Sei $h\colon [a,b] \to \mathbb{K}^n$ differenzierbar. Dann existiert ein $c \in [a,b]$ mit

$$|| h(b) - h(a) || \leq (b-a) \, || h'(c) || \ .$$

Satz 46.3. *Sei $G \subseteq \mathbb{K}^n$ offen und $\varphi\colon G \to \mathbb{K}^m$ eine Abbildung. Seien x_i, $i = 1, \ldots, n$, die Koordinaten von $\mathbb{K}^n$ und $P \in G$ ein Punkt. Es sei angenommen, dass alle partiellen Ableitungen in einer offenen Umgebung von P existieren und in P stetig sind. Dann ist φ in P (total) differenzierbar. Ist die Abbildung φ bezüglich der Standardbasis des $\mathbb{K}^m$ durch die Koordinatenfunktionen $f_1, \ldots, f_m$ gegeben, so wird unter diesen Bedingungen das totale Differential in P durch die Jacobi-Matrix*

$$\left(\frac{\partial f_j}{\partial x_i}(P) \right)_{1 \leq i \leq n, 1 \leq j \leq m}$$

beschrieben.

Beweis. Indem wir G durch eine eventuell kleinere offene Umgebung von P ersetzen, können wir annehmen, dass auf G die Richtungsableitungen

$$Q \longmapsto (D_i \varphi)(Q) := (D_{e_i} \varphi)(Q) = \left(\frac{\partial f_1}{\partial x_i}(Q), \ldots, \frac{\partial f_m}{\partial x_i}(Q) \right) \in \mathbb{K}^m$$

existieren und in P stetig sind. Daher ist nach Proposition 46.1 die lineare Abbildung

$$\mathbb{K}^n \longrightarrow \mathbb{K}^m, \, v = (v_1, \ldots, v_n) \longmapsto \sum_{i=1}^n v_i (D_i \varphi)(P),$$

der einzige Kandidat für das totale Differential. Daher müssen wir zeigen, dass diese lineare Abbildung die definierende Eigenschaft des totalen Differentials besitzt. Setze $P_i := P + v_1 e_1 + \ldots + v_i e_i$ (abhängig von v). Dann gelten mit dem Ansatz

$$r(v) \ := \ \frac{\varphi(P+v) - \varphi(P) - \sum_{i=1}^n v_i (D_i(\varphi))(P)}{|| v ||}$$

(für v hinreichend klein) die Abschätzungen

$$\begin{aligned}
|| r(v) || \ &= \ \frac{|| \varphi(P+v) - \varphi(P) - \sum_{i=1}^n v_i (D_i(\varphi))(P) ||}{|| v ||} \\
&= \ \frac{|| \sum_{i=1}^n (\varphi(P_i) - \varphi(P_{i-1}) - v_i (D_i(\varphi))(P)) ||}{|| v ||} \\
&\leq \ \sum_{i=1}^n \frac{|| \varphi(P_i) - \varphi(P_{i-1}) - v_i (D_i(\varphi))(P) ||}{|| v ||} \\
&= \ \sum_{i=1}^n \frac{|| \varphi(P_{i-1} + v_i e_i) - \varphi(P_{i-1}) - v_i (D_i(\varphi))(P) ||}{|| v ||} .
\end{aligned}$$

Wir betrachten jeden Summanden einzeln. Für fixiertes i ist die Abbildung (die auf dem Einheitsintervall definiert ist)

$$h_i : s \longmapsto \varphi(P_{i-1} + sv_ie_i) - sv_i(D_i(\varphi))(P)$$

differenzierbar (aufgrund der Existenz der partiellen Ableitungen auf G) mit der Ableitung

$$s \longmapsto v_i(D_i(\varphi))(P_{i-1} + sv_ie_i) - v_i(D_i(\varphi))(P)\,.$$

Nach der Mittelwertabschätzung existiert eine reelle Zahl $0 \leq c_i \leq 1$, so dass (dies ist die Norm von $h_i(1) - h_i(0)$)

$$\begin{aligned} & \| \varphi(P_{i-1} + v_ie_i) - \varphi(P_{i-1}) - v_i(D_i(\varphi))(P) \| \\ \leq \; & \| v_i(D_i(\varphi))(P_{i-1} + c_iv_ie_i) - v_i(D_i(\varphi))(P) \| \\ = \; & |v_i| \cdot \| (D_i(\varphi))(P_{i-1} + c_iv_ie_i) - (D_i(\varphi))(P) \| \\ \leq \; & \| v \| \cdot \| (D_i(\varphi))(P_{i-1} + c_iv_ie_i) - (D_i(\varphi))(P) \| \end{aligned}$$

gilt. Aufsummieren liefert also, dass unser Ausdruck $\| r(v) \|$ nach oben beschränkt ist durch

$$\begin{aligned} & \sum_{i=1}^{n} \frac{\| v \| \cdot \| (D_i(\varphi))(P_{i-1} + c_iv_ie_i) - (D_i(\varphi))(P) \|}{\| v \|} \\ \leq \; & \cdot \sum_{i=1}^{n} \| (D_i(\varphi))(P_{i-1} + c_iv_ie_i) - (D_i(\varphi))(P) \|\,. \end{aligned}$$

Da die partiellen Ableitungen $D_i(\varphi)$ stetig in P sind, wird die Summe rechts mit v beliebig klein, da dann $P_{i-1} + c_iv_ie_i$ gegen P konvergiert. Also ist der Grenzwert für $v \to 0$ gleich 0. □

Korollar 46.4. *Polynomfunktionen sind total differenzierbar.*

Beweis. Dies folgt aus Satz 46.2 und daraus, dass die partiellen Ableitungen von Polynomfunktionen wieder Polynomfunktionen und daher nach Satz 34.12 stetig sind. □

Für einen anderen Beweis siehe Aufgabe 45.4 und Aufgabe 45.5.

Extrema

In den nächsten Vorlesungen wollen wir mit der Hilfe von Ableitungen verstehen, wann eine Funktion

$$G \longrightarrow \mathbb{R},$$

$G \subseteq \mathbb{R}^n$, ein Extremum, also ein Maximum oder ein Minimum in einem Punkt $P \in G$ annimmt. Hier stellen wir die relevanten Definitionen zusammen und stellen einige typische Beispiele vor.

Definition 46.5. Sei (M, d) ein metrischer Raum und

$$f\colon M \longrightarrow \mathbb{R}$$

eine Funktion. Man sagt, dass f in einem Punkt $x \in M$ ein *lokales Maximum* besitzt, wenn es ein $\epsilon > 0$ gibt derart, dass für alle $x' \in M$ mit $d(x, x') < \epsilon$ die Abschätzung

$$f(x) \geq f(x')$$

gilt. Man sagt, dass f in $x \in M$ ein *lokales Minimum* besitzt, wenn es ein $\epsilon > 0$ gibt derart, dass für alle $x' \in M$ mit $d(x, x') < \epsilon$ die Abschätzung

$$f(x) \leq f(x')$$

gilt.

Definition 46.6. Sei (M, d) ein metrischer Raum und

$$f\colon M \longrightarrow \mathbb{R}$$

eine Funktion. Man sagt, dass f in einem Punkt $x \in M$ ein *isoliertes lokales Maximum* besitzt, wenn es ein $\epsilon > 0$ gibt derart, dass für alle $x' \in M$ mit $d(x, x') < \epsilon$ und $x' \neq x$ die Abschätzung

$$f(x) > f(x')$$

gilt. Man sagt, dass f in $x \in M$ ein *isoliertes lokales Minimum* besitzt, wenn es ein $\epsilon > 0$ gibt derart, dass für alle $x' \in M$ mit $d(x, x') < \epsilon$ und $x' \neq x$ die Abschätzung

$$f(x) < f(x')$$

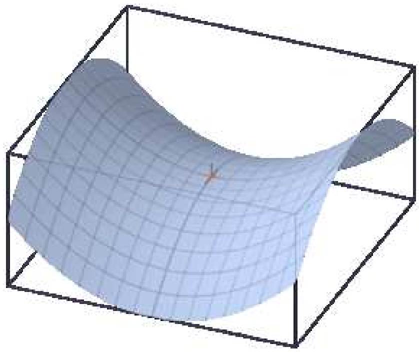
gilt.

Beispiel 46.7. Die Funktion

$$\mathbb{R}^2 \longrightarrow \mathbb{R},\ (x, y) \longmapsto x^2 + y^2,$$

hat in $P = (0, 0)$ den Wert 0 und überall sonst positive Werte, daher liegt in P ein (isoliertes) globales Minimum vor.

Wenn die Funktion $f\colon M \to \mathbb{R}$ ein lokales Minimum im Punkt $P \in M$ besitzt, so gilt dies auch für die Einschränkung von f auf jede Teilmenge $N \subseteq M$, die P enthält. Beispielsweise muss ein (lokales) Minimum einer Funktion der Ebene auch auf jeder Geraden durch diesen Punkt ein (lokales) Minimum sein.



Dies heißt umgekehrt, dass wenn eine Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$ auf einer Geraden L_1 durch P ein isoliertes lokales Maximum und auf einer anderen Geraden L_2 ein isoliertes lokales Minimum besitzt, dass dann kein lokales Extremum vorliegen kann. Solche Punkte nennt man *Sattelpunkt* oder *Passpunkt*, das Standardbeispiel ist das folgende.

Beispiel 46.8. Wir betrachten das Verhalten der Funktion

$$\mathbb{R}^2 \longrightarrow \mathbb{R},\ (x,y) \longmapsto x^2 - y^2.$$

in $P=(0,0)$. Die Einschränkung dieser Funktion auf der durch $y=0$ gegebenen Geraden (also auf der x-Achse) ist die Funktion $x \mapsto x^2$, die in P ein (isoliertes) globales Minimum besitzt. Die Einschränkung dieser Funktion auf der durch $x=0$ gegebenen Geraden (also auf der y-Achse) ist die Funktion $y \mapsto -y^2$, die in P ein (isoliertes) globales Maximum besitzt. Daher kann f in P kein Extremum besitzen. Auf den durch $y=x$ und $y=-x$ gegebenen Geraden ist die Funktion die Nullfunktion.

Es sei

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$$

eine stetige Funktion, die im Nullpunkt $(0,0)$ folgende Eingenschaft erfüllt. Zu jeder Geraden $G \subseteq \mathbb{R}^2$ durch den Nullpunkt besitzt die auf G eingeschränkte Funktion ein lokales isoliertes Maximum. Jeder Wanderer, der durch das durch f gegebene Gebirge schnurstracks in eine bestimmte Richtung durch den Punkt läuft, wird also in diesem Punkt ein Gipfelerlebnis haben. Folgt daraus, dass wirklich ein Gipfel vorliegt? Das folgende Beispiel zeigt, dass das nicht der Fall sein muss.

Beispiel 46.9. Wir betrachten im $\mathbb{R}^2$ die beiden Kreise K_1 und K_2, wobei K_1 den Mittelpunkt $(0,1)$ und Radius 1 und K_2 den Mittelpunkt $(0,2)$ und Radius 2 habe. K_1 liegt innerhalb von K_2, und die beiden Kreise berühren sich in $P=(0,0)$. Durch diese beiden Kreise wird die Ebene (neben den zwei Kreislinien selbst) in drei offene Gebiete aufgeteilt: Das Innere des Kreises K_1 $(=A)$, die große offene Kreisscheibe ohne die kleine abgeschlossene Kreisscheibe $(=B)$ und das Äußere von K_2 $(=C)$. Der innere Kreis K_1 wird als

Nullstelle der Funktion $f_1(x,y) = x^2 + (y-1)^2 - 1$ beschrieben. Im Innern von K_1 ist diese Funktion negativ, auf K_1 hat sie den Wert 0 und außerhalb davon hat sie positive Werte. Entsprechendes gilt für K_2 und die Funktion $f_2(x,y) = x^2 + (y-2)^2 - 4$. Wir setzen

$$\begin{aligned} f(x,y) &:= f_1(x,y) \cdot f_2(x,y) \\ &= \big(x^2 + (y-1)^2 - 1\big) \cdot \big(x^2 + (y-2)^2 - 4\big) \\ &= \big(x^2 + y^2 - 2y\big) \cdot \big(x^2 + y^2 - 4y\big) \\ &= x^4 + y^4 + 2x^2y^2 - 6y^3 - 6x^2y + 8y^2. \end{aligned}$$

Diese Funktion nimmt auf den beiden Kreisen den Wert 0 an, sie ist auf A positiv, auf B negativ und auf C wieder positiv.

Die Funktion f besitzt in P kein lokales Minimum, da sie dort den Wert 0 besitzt und da jede beliebig kleine Ballumgebung $U(P, \epsilon)$ den Bereich B trifft, wo f negative Werte besitzt. Die Einschränkung der Funktion auf jede Gerade durch den Nullpunkt besitzt aber dort ein lokales Minimum. Sei dazu G eine solche Gerade. Wenn G die x-Achse ist, so verläuft diese Gerade (bis auf P selbst) in C, wo f nur positive Werte annimmt, so dass in P ein (sogar globales) Minimum vorliegt. Sei also G eine von der x-Achse verschiedene Gerade durch P. Die eine Hälfte der Geraden verläuft ganz in C, wo die Funktion positiv ist. Die andere Hälfte verläuft, ausgehend von P, zuerst in A, dann in B und schließlich in C. Da die Funktion auf A positiv ist, kann man ein Teilintervall $[-\delta, \delta]$ der Gerade wählen derart, dass dieses Teilstück (abgesehen von P) nur in A und C verläuft. Auf diesem Teilintervall nimmt die Funktion in P den Wert 0 und sonst überall positive Werte annimmt. Daher besitzt die eingeschränkte Funktion ein lokales Minimum. Das dabei zu wählende δ hängt natürlich wesentlich von der Steigung der Geraden ab, es gibt kein gemeinsames δ für alle Geraden.

Abbildungsverzeichnis