Prof. Dr. H. Brenner **Osnabrück SS 2012**

Mathematik für Anwender II

Vorlesung 47

Die Kettenregel

Die Eleganz des totalen Differentials wird in der folgenden allgemeinen Version der Kettenregel deutlich. Sie besagt, dass bei einer Verknüpfung von differenzierbaren Abbildungen das totale Differential (also die lineare Approximation) gleich der Verknüpfung der einzelnen totalen Differentiale ist. Der Beweis verwendet an einer Stelle, dass eine lineare Abbildung

$$L\colon V \longrightarrow W$$

zwischen euklidischen Räumen auf der abgeschlossenen Einheitskugel $B(0,1)$ beschränkt ist, d.h. dass es ein $b \in \mathbb{R}$ gibt mit

$$|| L(v) || \leq b$$

für alle v mit $|| v || \leq 1$. Diese Aussage gilt sogar für jede stetige Abbildung, werden wir hier aber nur für eine lineare Abbildung beweisen: Dazu wählen wir eine Orthonormalbasis $v_1, \ldots, v_n$ von V. Sei $v = \sum_{i=1}^n a_i v_i$ aus $B(0,1)$. Wegen

$$|| v ||^2 = \sum_{i=1}^n a_i^2 \leq 1$$

ist

$$|a_i| \leq 1.$$

Somit ist

$$\begin{aligned} || L(v) || &= || L\left(\sum_{i=1}^n a_i v_i\right) || \\ &= || \sum_{i=1}^n a_i L(v_i) || \\ &\leq \sum_{i=1}^n |a_i| \, || L(v_i) || \\ &\leq \sum_{i=1}^n || L(v_i) ||, \end{aligned}$$

das heißt, dass die Beschränktheit mit

$$b := \sum_{i=1}^n || L(v_i) ||$$

gilt.

Satz 47.1. *Seien V, W und U endlichdimensionale $\mathbb{R}$-Vektorräume, $G \subseteq V$ und $D \subseteq W$ offene Mengen, und $\varphi\colon G \to W$ und $\psi\colon D \to U$ Abbildungen derart, dass $\varphi(G) \subseteq D$ gilt. Es sei weiter angenommen, dass φ in $P \in G$ und ψ in $\varphi(P) \in D$ total differenzierbar ist. Dann ist $\psi \circ \varphi\colon G \to U$ in P differenzierbar mit dem totalen Differential*

$$(D(\psi \circ \varphi))_P = (D\psi)_{\varphi(P)} \circ (D\varphi)_P .$$

Beweis. Wir haben nach Voraussetzung (wobei wir $Q := \varphi(P)$ setzen)

$$\varphi(P+v) = \varphi(P) + L(v) + ||v||\, r(v)$$

und

$$\psi(Q+w) = \psi(Q) + M(w) + ||w||\, s(w)$$

mit linearen Abbildungen $L\colon V \to W$ und $M\colon W \to U$, und mit in 0 stetigen Funktionen $r\colon U(0,\delta) \to W$ und $s\colon U(0,\delta') \to U$ (beachte, dass $U(P,\delta) \subseteq V$ und $U(Q,\delta') \subseteq W$ gilt), die beide in 0 den Wert 0 annehmen. Damit gilt

$$\begin{aligned}
& (\psi \circ \varphi)(P+v) \\
= \; & \psi(\varphi(P+v)) \\
= \; & \psi\left(\varphi(P) + L(v) + ||v||\, r(v)\right) \\
= \; & \psi(\varphi(P)) + M(L(v) + ||v||\, r(v)) \\
& + ||L(v) + ||v||\, r(v)||\, s(L(v) + ||v||\, r(v)) \\
= \; & \psi(\varphi(P)) + M(L(v)) + M(||v||\, r(v)) \\
& + ||L(v) + ||v||\, r(v)||\, s(L(v) + ||v||\, r(v)) \\
= \; & \psi(\varphi(P)) + (M \circ L)(v) + ||v||\, M(r(v)) \\
& + ||\, ||v||\, L\left(\frac{v}{||v||}\right) + ||v||\, r(v)||\, s(L(v) + ||v||\, r(v)) \\
= \; & \psi(\varphi(P)) + (M \circ L)(v) \\
& + ||v|| \left(M(r(v)) + ||L\left(\frac{v}{||v||}\right) + r(v)||\, s(L(v) + r(v)) \right).
\end{aligned}$$

Dabei haben wir in der dritten Gleichung die lineare Approximation für $w = L(v) + ||\, v\, ||\, r(v)$ eingesetzt. Die beiden letzten Gleichungen gelten nur für $v \neq 0$. Der Ausdruck

$$t(v) := M(r(v)) + ||L\left(\frac{v}{||v||}\right) + r(v)||\, s(L(v) + r(v))$$

ist unser Kandidat für die Abweichungsfunktion. Der erste Summand $M(r(v))$ ist in $v = 0$ stetig und hat dort auch den Wert 0. Es genügt also den zweiten Summanden zu betrachten. Der $||-||$-Ausdruck ist in einer Umgebung der Null beschränkt, da L auf der kompakten Einheitssphäre beschränkt ist und da r in 0 stetig ist. Daher hängt die Stetigkeit nur von dem rechten Faktor ab. Aber $L(v) + r(v)$ hat für $v \to 0$ den Grenzwert 0. Damit ist auch $s(L(v) + r(v))$ in 0 stetig und hat dort den Grenzwert 0. □

Korollar 47.2. *Es seien $G \subseteq \mathbb{R}^n$ und $D \subseteq \mathbb{R}^m$ offene Mengen, und $f\colon G \to \mathbb{R}^m$ und $g\colon D \to \mathbb{R}^k$ seien Abbildungen derart, dass $f(G) \subseteq D$*

gilt. Es sei weiter angenommen, dass f in $P \in G$ und g in $f(P) \in D$ total differenzierbar ist. Dann ist $h = g \circ f\colon\ G \to U$ in P differenzierbar und zwischen den Jacobi-Matrizen gilt die Beziehung

$$\operatorname{Jak}(h)_P \ = \ \operatorname{Jak}(g \circ f)_P \ = \ \operatorname{Jak}(g)_{f(P)} \circ \operatorname{Jak}(f)_P,$$

also ausgeschrieben

$$\begin{pmatrix} \frac{\partial h_1}{\partial x_1}(P) & \dots & \frac{\partial h_1}{\partial x_n}(P) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_k}{\partial x_1}(P) & \dots & \frac{\partial h_k}{\partial x_n}(P) \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(f(P)) & \dots & \frac{\partial g_1}{\partial y_m}(f(P)) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_k}{\partial y_1}(f(P)) & \dots & \frac{\partial g_k}{\partial y_m}(f(P)) \end{pmatrix} \circ \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(P) & \dots & \frac{\partial f_1}{\partial x_n}(P) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(P) & \dots & \frac{\partial f_m}{\partial x_n}(P) \end{pmatrix}.$$

Beweis. Dies folgt direkt aus Satz 47.1 unter Berücksichtigung von Bemerkung 46.9. □

Bei der vorstehenden Aussage kann man mit Satz 46.10 häufig direkt auf die totale Differenzierbarkeit schließen.

Beispiel 47.3. Wir wollen die Kettenregel an Hand der beiden Abbildungen

$$f\colon\ \mathbb{R}^3 \longrightarrow \mathbb{R}^2,\ (u,\ v,\ w) \longmapsto \left(uv^3w^2,\ u^2 - v^2w\right)$$

und

$$g\colon\ \mathbb{R}^2 \longrightarrow \mathbb{R}^3,\ (x,\ y) \longmapsto \left(xy - y^2,\ \cos x\,,\ x - y\right)$$

illustrieren. Diese Abbildungen sind stetig partiell differenzierbar und daher auch total differenzierbar. Die Jacobi-Matrizen zu diesen Abbildungen (in einem beliebigen Punkt $P = (u, v, w) \in \mathbb{R}^3$ bzw. $Q = (x, y) \in \mathbb{R}^2$) sind

$$\operatorname{Jac}(f)_P \ = \ \begin{pmatrix} \frac{\partial f_1}{\partial u}(P) & \frac{\partial f_1}{\partial v}(P) & \frac{\partial f_1}{\partial w}(P) \\ \frac{\partial f_2}{\partial u}(P) & \frac{\partial f_2}{\partial v}(P) & \frac{\partial f_2}{\partial w}(P) \end{pmatrix} \ = \ \begin{pmatrix} v^3w^2 & 3uv^2w^2 & 2uv^3w \\ 2u & -2vw & -v^2 \end{pmatrix}$$

und

$$\operatorname{Jac}(g)_Q \ = \ \begin{pmatrix} \frac{\partial g_1}{\partial x}(Q) & \frac{\partial g_1}{\partial y}(Q) \\ \frac{\partial g_2}{\partial x}(Q) & \frac{\partial g_2}{\partial y}(Q) \\ \frac{\partial g_3}{\partial x}(Q) & \frac{\partial g_3}{\partial y}(Q) \end{pmatrix} \ = \ \begin{pmatrix} y & x - 2y \\ -\sin x & 0 \\ 1 & -1 \end{pmatrix}.$$

Die zusammengesetzte Abbildung $g \circ f$ ist

$$\begin{aligned} g(f(u,v,w)) \ &= \ \left(uv^3w^2\left(u^2 - v^2w\right) - \left(u^2 - v^2w\right)^2,\ \cos\left(uv^3w^2\right),\ uv^3w^2 - u^2 + v^2w\right) \\ &= \ \left(u^3v^3w^2 - uv^5w^3 - u^4 - v^4w^2 + 2u^2v^2w,\ \cos\left(uv^3w^2\right),\ uv^3w^2 - u^2 + v^2w\right), \end{aligned}$$

die zugehörige Jacobi-Matrix in $P = (u, v, w)$ ist

$$\operatorname{Jac}(g \circ f)_P = \begin{pmatrix} 3u^2v^3w^2 - v^5w^3 - 4u^3 + 4uv^2w & 3u^3v^2w^2 - 5uv^4w^3 - 4v^3w^2 + 4u^2vw & 2u^3v^3w - 3uv^5w^2 - 2v^4w + 2u^2v^2 \\ -v^3w^2\sin\left(uv^3w^2\right) & -3uv^2w^2\sin\left(uv^3w^2\right) & -2uv^3w\sin\left(uv^3w^2\right) \\ v^3w^2 - 2u & 3uv^2w^2 + 2vw & 2uv^3w + v^2 \end{pmatrix}.$$

Die zusammengesetzte lineare Abbildung ist

$$\begin{aligned}
&\mathrm{Jak}(g)_{f(P)} \circ \mathrm{Jak}(f)_P \\
=&\mathrm{Jak}(g)_{(uv^3w^2,u^2-v^2w)} \circ \mathrm{Jak}(f)_P \\
=&\begin{pmatrix} u^2-v^2w & uv^3w^2-2u^2+2v^2w \\ -\sin\left(uv^3w^2\right) & 0 \\ 1 & -1 \end{pmatrix} \circ \begin{pmatrix} v^3w^2 3uv^2w^2 2uv^3w \\ 2u \quad -2vw \quad -v^2 \end{pmatrix} \\
=&\begin{pmatrix} 3u^2v^3w^2 -^5 w^3-4u^3+4uv^2w 3u^3v^2w^2-5uv^4w^3-4v^3w^2+4u^2vw 2u^3v^3w-3uv^5w^2-2v^4w+2u^2v^2 \\ -v^3w^2 \sin\left(uv^3w^2\right) & -3uv^2w^2 \sin\left(uv^3w^2\right) & -2uv^3w \sin\left(uv^3w^2\right) \\ v^3w^2-2u & 3uv^2w^2+2vw & 2uv^3w+v^2 \end{pmatrix}.
\end{aligned}$$

Bemerkung 47.4. Es sei I ein reelles Intervall, V und W seien euklidische Vektorräume und es sei

$$\gamma\colon\, I \longrightarrow V$$

eine differenzierbare Kurve. Es sei

$$L\colon\, V \longrightarrow W$$

eine lineare Abbildung. In Lemma 34.10 wurde gezeigt, dass die zusammengesetzte Abbildung

$$L \circ \gamma\colon\, I \longrightarrow W,\, t \longmapsto L(\gamma(t)),$$

(ebenfalls differenzierbar ist) und dass die Beziehung

$$(L \circ \gamma)'(t) \,=\, L(\gamma'(t))$$

besteht. Hier erhält man also den Richtungsvektor der zusammengesetzten Kurve, indem man den Richtungsvektor der Kurve in die lineare Abbildung einsetzt. Dies ist ein Spezialfall der Kettenregel angewendet auf γ und L. Es ist $(DL)_P = L$ nach Proposition 46.4 und es ist $(D\gamma)_t$ die lineare Abbildung von $\mathbb{R}$ nach V, die 1 auf den Richtungsvektor $\gamma'(t)$ schickt. Gemäß der Kettenregel ist das totale Differential der zusammengesetzten Kurve $L \circ \gamma$ gleich

$$(DL)_{\gamma(t)} \circ (D\gamma)_t \,=\, L \circ (D\gamma)_t \,=\, L \circ (1 \mapsto \gamma'(t))\,.$$

Dies ist die lineare Abbildung von $\mathbb{R}$ nach W, die 1 auf $L(\gamma'(t))$ schickt.

Bemerkung 47.5. Es seien V und W euklidische Vektorräume, $G \subseteq V$ eine offene Teilmenge und

$$f\colon\, G \longrightarrow W$$

eine in $P \in G$ total differenzierbare Abbildung. Es sei $v \in V$ ein Vektor und

$$\gamma\colon\, I \longrightarrow G,\, t \longmapsto P + tv,$$

die zugehörige affin-lineare Abbildung durch diesen Punkt (dabei sei das reelle Intervall $I = [-a, a]$ so gewählt, dass $\gamma(I) \subseteq G$) liegt. Die zusammengesetzte Abbildung

$$I \longrightarrow W,\, t \longmapsto f(\gamma(t)),$$

wird zur Definition der Richtungsableitung von f in P in Richtung v verwendet, es ist

$$(D_vf)(P) \,=\, (f \circ \gamma)'(0).$$

Das zur Kurve $f \circ \gamma$ gehörige totale Differential in 0 von $\mathbb{R}$ nach W, also $(D(f \circ \gamma))_0$, ist durch $1 \mapsto (f \circ \gamma)'(0)$ festgelegt. Andererseits ist nach der Kettenregel

$$(D(f \circ \gamma))_0 = (Df)_{\gamma(0)} \circ (D\gamma)_0$$

und somit ist

$$\begin{aligned}(D_v f)(P) &= (f \circ \gamma)'(0)\\ &= (D(f \circ \gamma))_0(1)\\ &= \left((Df)_{\gamma(0)} \circ (D\gamma)_0\right)(1)\\ &= (Df)_{\gamma(0)}((D\gamma)_0(1))\\ &= (Df)_{\gamma(0)}(\gamma'(0))\\ &= (Df)_{\gamma(0)}(v).\end{aligned}$$

Dies ergibt einen neuen Beweis für Proposition 46.8.

Das folgende Beispiel illustriert, dass das totale Differential unabhängig von der Wahl einer Basis ist, die partiellen Ableitungen aber nicht.

Beispiel 47.6. Wir betrachten die Abbildung $f\colon \mathbb{R}^3 \to \mathbb{R}$, die durch

$$(x, y, z) \longmapsto 2xy^2 + x^2z^3 + z^2$$

gegeben sei. Es ist leicht die partiellen Ableitungen in jedem Punkt zu berechnen, nämlich:

$$(\partial f/\partial x,\, \partial f/\partial y,\, \partial f/\partial z)_{(x,y,z)} = \left(2y^2 + 2xz^3,\, 4xy,\, 3x^2z^2 + 2z\right).$$

Da diese alle stetig sind, haben wir nach Satz 46.10 das totale Differential in jedem Punkt gefunden.

Nehmen wir nun an, dass wir nur an der Restriktion dieser Funktion auf die Ebene

$$E \subset \mathbb{R}^3,\, E = \{(x, y, z) \mid 3x + 2y - 5z = 0\}$$

interessiert sind. E ist also der Kern der linearen Abbildung

$$L\colon \mathbb{R}^3 \longrightarrow \mathbb{R},\, (x, y, z) \longmapsto 3x + 2y - 5z.$$

Als Kern ist E selbst ein (zweidimensionaler) Vektorraum. Die Einschränkung von f auf die Ebene ergibt also die Abbildung

$$\tilde{f} = f|_E\colon E \longrightarrow \mathbb{R}.$$

Diese Abbildung kann man als die Komposition $E \subset \mathbb{R}^3 \to \mathbb{R}$ auffassen und diese ist nach der Kettenregel differenzierbar. Wenn wir die Inklusion von E in $\mathbb{R}^3$ mit N bezeichnen, so ist das totale Differential der Komposition in einem Punkt $P \in E$ gemäß der Kettenregel gerade die Abbildung

$$(D\tilde{f})_P = (Df)_P \circ N\colon E \longrightarrow \mathbb{R}.$$

Daher macht es hier Sinn vom totalen Differential zu sprechen.

Es macht allerdings keinen Sinn von partiellen Ableitungen der Abbildung $f|_E\colon E \to \mathbb{R}$ zu sprechen, da es keine natürliche Basis auf E gibt und daher auch keine natürlichen Koordinaten. Es ist leicht eine Basis von E zu finden und damit Koordinaten, es gibt aber keine „beste Wahl“, und die partiellen Ableitungen sehen in jeder Basis verschieden aus.

Eine Basis von E ist beispielsweise durch $v_1 = (0, 5, 2)$ und $v_2 = (5, 0, 3)$ gegeben, und eine weitere durch $w_1 = (1, 1, 1)$ und $w_2 = (2, -3, 0)$. Mit solchen Basen erhalten wir Identifikationen $\mathbb{R}^2 \to E$ und somit eine numerische Beschreibung der Abbildung $\mathbb{R}^2 \cong E \to \mathbb{R}$, womit wir die partiellen Ableitungen bezüglich der gewählten Basen berechnen können.

In der ersten Basis ist die Identifikation gegeben durch die Abbildung

$$(s, t) \longmapsto sv_1 + tv_2 = s(0, 5, 2) + t(5, 0, 3) = (5t, 5s, 2s + 3t)$$

und dieser Ausdruck wird durch f abgebildet auf

$$\begin{aligned} & 2(5t)(5s)^2 + (5t)^2(2s + 3t)^3 + (2s + 3t)^2 \\ = \; & 250ts^2 + 25t^2(8s^3 + 36s^2t + 54st^2 + 27t^3) + 4s^2 + 9t^2 + 12st \\ = \; & 250ts^2 + 200s^3t^2 + 900s^2t^3 + 1350st^4 + 675t^5 + 4s^2 + 9t^2 + 12st. \end{aligned}$$

Die partiellen Ableitungen dieser Komposition (nennen wir sie g) bezüglich dieser Basis sind gegeben durch

$$\partial g/\partial s = 500ts + 600s^2t^2 + 1800st^3 + 1350t^4 + 8s + 12t$$

und

$$\partial g/\partial t = 250s^2 + 400s^3t + 2700s^2t^2 + 5400st^3 + 3375t^4 + 18t + 12s\,.$$

In der zweiten Basis $w_1 = (1, 1, 1)$ und $w_2 = (2, -3, 0)$ ist die Identifikation gegeben durch

$$(r, u) \longmapsto rw_1 + uw_2 = r(1, 1, 1) + u(2, -3, 0) = (r + 2u, r - 3u, r)$$

und dieser Ausdruck wird unter f abgebildet auf

$$\begin{aligned} & 2(r + 2u)(r - 3u)^2 + (r + 2u)^2r^3 + r^2 \\ = \; & 2r^3 + 4r^2u - 12r^2u - 24ru^2 + 18ru^2 + 36u^3 + r^5 + 4r^4u + 4r^3u^2 + r^2 \\ = \; & 2r^3 - 8r^2u - 6ru^2 + 36u^3 + r^5 + 4r^4u + 4r^3u^2 + r^2. \end{aligned}$$

Die partiellen Ableitungen der Komposition (nennen wir sie h) bezüglich dieser Basis sind

$$\partial h/\partial r = 6r^2 - 16ru - 6u^2 + 5r^4 + 16r^3u + 12r^2u^2 + 2r$$

und

$$\partial h/\partial u = -8r^2 - 12ru + 108u^2 + 4r^4 + 8r^3u\,.$$

Fazit: Koordinaten sind gut für Berechnungen aber schlecht für die Mathematik.