Totale Differenzierbarkeit

Wir möchten Abbildungen $\varphi : V \to W$ zwischen Vektorräumen differenzieren, und allgemeiner Abbildungen

$$\varphi : G \longrightarrow W,$$

wobei $G \subseteq V$ eine gewisse offene Teilmenge ist. Wir wiederholen kurz die Situation in einer Variablen: angenommen wir haben eine Abbildung $\varphi : \mathbb{R} \to \mathbb{R}$, dann ist die Grundidee einer differenzierbaren Abbildung und ihrer Ableitung eine „Tangente an den Graphen“ anzulegen. Dabei kann man sagen, dass die Tangente die beste *lineare Approximation* von φ (genauer: der Graph einer affin-linearen Approximation) in einem gegebenen Punkt $x \in \mathbb{R}$ darstellt. Da die Steigung der Tangente wieder eine reelle Zahl ist, wird beim Differenzieren jedem Punkt x wieder eine Zahl zugeordnet. Wir erhalten also eine neue Funktion, welche wir mit φ' bezeichnen. Im höherdimensionalen Fall ist dies komplizierter, aber die Idee einer bestmöglichen *linearen Approximation* bleibt bestehen.

Im Folgenden nehmen wir an, dass alle Vektorräume endlichdimensional und mit einer euklidischen Norm versehen sind. Wie in der 40. Vorlesung gezeigt wurde, hängt die Topologie, also die Konzepte offene Menge, Stetigkeit, Konvergenz, nicht von der gewählten euklidischen Struktur ab. Bei $\mathbb{K} = \mathbb{C}$ statten wir den zugrunde liegenden reellen Vektorraum mit einer euklidischen Struktur aus.

Definition 44.1. Sei $G \subseteq V$ eine offene Menge in einem endlichdimensionalen $\mathbb{K}$-Vektorraum V und $\varphi : G \to W$ eine Abbildung. Dann heißt φ *differenzierbar* (oder *total differenzierbar*) im Punkt $P \in G$, wenn es eine $\mathbb{K}$-lineare Abbildung $L : V \to W$ mit der Eigenschaft

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

gibt, wobei $r : U(0, \delta) \to W$ eine in 0 stetige Abbildung mit $r(0) = 0$ ist und die Gleichung für alle $v \in V$ mit $P + v \in U(P, \delta) \subseteq G$ gilt.

Diese lineare Abbildung heißt, falls sie existiert, das *(totale) Differential* von φ an der Stelle P und wird mit

$$(D\varphi)_P$$

bezeichnet.

Äquivalent zur totalen Differenzierbarkeit ist die Eigenschaft, dass der Ausdruck

$$r(v) = \frac{\varphi(P + v) - \varphi(P) - L(v)}{||\, v\, ||}$$

für $v \to 0$ gegen 0 konvergiert. Ebenfalls äquivalent ist die Eigenschaft, dass der Limes (von Funktionen)

$$\lim_{v \to 0, v \neq 0} \frac{||\, \varphi(P + v) - \varphi(P) - L(v)\, ||}{||\, v\, ||}$$

existiert und gleich 0 ist.

Das Konzept der totalen Differenzierbarkeit ist eher theoretisch und weniger konkreten Berechnungen zugänglich. Wir werden später dieses Konzept mit dem Konzept der partiellen Ableitungen in Verbindung bringen, welches eher für Berechnungen geeignet ist, jedoch von Koordinaten, d.h. von der Auswahl einer Basis, abhängt (siehe auch Beispiel 44.8 weiter unten).

Lemma 44.2. *Seien V und W endlichdimensionale normierte $\mathbb{K}$-Vektorräume und sei die Abbildung $\varphi : G \to W$ auf einer offenen Teilmenge $G \subseteq V$ definiert. Sei $P \in G$ ein Punkt. Dann existiert höchstens eine lineare Abbildung mit den Eigenschaften aus Definition. Ist φ im Punkt P differenzierbar, so ist das totale Differential $(D\varphi)_P$ eindeutig bestimmt.*

Beweis. Angenommen, es gelte $\varphi(P+v) = \varphi(P) + L_1(v) + \| v \| \cdot r_1(v)$ und $\varphi(P+v) = \varphi(P) + L_2(v) + \| v \| \cdot r_2(v)$ mit zwei linearen Abbildungen L_1 und zwei L_2 und im Punkt 0 stetigen Funktionen $r_1, r_2 : U(0,\delta) \to W$ mit $r_1(0) = r_2(0) = 0$. Wir müssen $L_1 = L_2$ zeigen. Dazu ziehen wir die beiden Gleichungen voneinander ab (da es sich hier um Gleichungen von Funktionswerten im Vektorraum W handelt, ist hier werteweises Abziehen gemeint) und erhalten die Gleichung

$$0 = (L_1 - L_2)(v) + \| v \| \cdot (r_1(v) - r_2(v)) \, .$$

Daher müssen wir zeigen, dass die (konstante) Nullabbildung die Eigenschaft besitzt, dass die lineare Abbildung 0 ihre einzige lineare Approximation ist. Wir nehmen daher an, dass

$$0 = L(v) + \| v \| \cdot r(v)$$

gilt, wobei L linear und r eine in 0 stetige Funktion ist mit $r(0) = 0$. Wenn L nicht die Nullabbildung ist, so gibt es einen Vektor $v \in V$ mit $L(v) = w \neq 0$. Dann gilt für $s \in \mathbb{K}$

$$0 = L(sv) + \| sv \| \, r(sv) = sw + | s | \cdot \| v \| \cdot r(sv) \, .$$

Dies impliziert, dass $r(sv) = -sw/(| s | \cdot \| v \|)$ für $s \neq 0$ gilt. Die Norm von $r(sv)$ ist daher konstant gleich $\| w \| \, / \, \| v \| \neq 0$. Also gilt $\lim_{s \to 0} \| r(sv) \| \neq 0$, ein Widerspruch. □

Proposition 44.3. *Sei $L : V \to W$ eine lineare Abbildung. Dann ist L in jedem Punkt $P \in V$ differenzierbar und stimmt in jedem Punkt mit ihrem totalen Differential überein.*

Beweis. Aufgrund der Linearität gilt

$$L(P+v) = L(P) + L(v) \, .$$

Also können wir $r = 0$ wählen. □

Beispiel 44.4. Ist $\varphi : V \to W$ konstant mit $\varphi(v) = w \in W$ für alle $v \in V$, so ist φ differenzierbar mit totalem Differential 0 (siehe Aufgabe 44.1).

Proposition 44.5. *Es seien V und W endlichdimensionale normierte $\mathbb{K}$-Vektorräume und $G \subseteq V$ eine offene Teilmenge. Seien $\varphi_1, \varphi_2 : G \to W$ im Punkt $P \in G$ differenzierbare Abbildungen mit den Differentialen $(D\varphi_1)_P$ und $(D\varphi_2)_P$. Dann ist auch $\varphi = \varphi_1 + \varphi_2$ in P differenzierbar und es gilt*

$$(D(\varphi_1 + \varphi_2))_P = (D\varphi_1)_P + (D\varphi_2)_P \,.$$

Ebenso gilt $(D(a\varphi_1))_P = a(D\varphi_1)_P$.

Beweis. Sei $\varphi_1(P+v) = \varphi_1(P) + L_1(v) + \| v \| \cdot r_1(v)$ und $\varphi_2(P+v) = \varphi_2(P) + L_2(v) + \| v \| \cdot r_2(v)$. Dann gilt

$$\begin{aligned} (\varphi_1 + \varphi_2)(P+v) &= \varphi_1(P+v) + \varphi_2(P+v) \\ &= \varphi_1(P) + L_1(v) + \| v \| \cdot r_1(v) + \varphi_2(P) + L_2(v) + \| v \| \cdot r_2(v) \\ &= (\varphi_1 + \varphi_2)(P) + (L_1 + L_2)(v) + \| v \| \, (r_1(v) + r_2(v)). \end{aligned}$$

Wir erhalten also die gewünschte Gestalt, da auch $r_1 + r_2$ in 0 stetig ist mit $(r_1 + r_2)(0) = 0$. Der Beweis der zweiten Aussage ist ähnlich. □

Proposition 44.6. *Es seien V und W endlichdimensionale normierte Vektorräume und $G \subseteq V$ eine offene Teilmenge. Sei $\varphi : G \to W$ eine in $P \in G$ differenzierbare Abbildung. Dann ist φ auch stetig im Punkt P.*

Beweis. Nach Definition gilt $\varphi(P+v) = \varphi(P) + L(v) + \| v \| \cdot r(v)$. Die rechte Seite ist stetig (nach Definition und Satz 20.11) in $v = 0$. Damit ist φ stetig in P. □

Die Kettenregel

Die Eleganz des totalen Differentials wird in der folgenden allgemeinen Version der Kettenregel deutlich.

Satz 44.7. *Seien V, W und U $\mathbb{K}$-Vektorräume, $G \subseteq V$ und $D \subseteq W$ offene Mengen, und $\varphi : G \to W$ und $\psi : D \to U$ Abbildungen derart, dass $\varphi(G) \subseteq D$ gilt. Es sei weiter angenommen, dass φ in $P \in G$ und ψ in $\varphi(P) \in D$ differenzierbar ist. Dann ist $\psi \circ \varphi : G \to U$ 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 $L : V \to W$ und $M : W \to U$, und mit in 0 stetigen Funktionen $r : U(0,\delta) \to W$ und $s : U(0,\delta') \to U$ (beachte, dass $U(Q,\delta) \subseteq V$ und $U(Q,\delta') \subseteq W$ gilt), die beide in 0 den Wert 0 annehmen. Damit gilt

$$(\psi \circ \varphi)(P+v)$$

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

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

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

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 $L(v)+\,||\,v\,||\,\cdot r(v)$ hat für $v \to 0$ den Grenzwert 0. Damit ist auch $s(L(v)+\,||\,v\,||\,r(v))$ in 0 stetig und hat dort den Grenzwert 0. $\square$

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

Beispiel 44.8. Wir betrachten die Abbildung $f : \mathbb{K}^3 \to \mathbb{K}$, 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:

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

Wir werden in der nächsten Vorlesung sehen, dass diese Abbildung in jedem Punkt total differenzierbar ist, und dass die Jacobi-Matrix das totale Differential beschreibt.

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

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

interessiert sind. E ist also der Kern der linearen Abbildung

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

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

$$\tilde{f} = f|_E : E \longrightarrow \mathbb{K}.$$

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

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

Daher macht es hier Sinn vom totalen Differential zu sprechen.

Es macht allerdings keinen Sinn von partiellen Ableitungen der Abbildung $f|_E : E \to \mathbb{K}$ 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 bspw. gegeben durch $v_1 = (0, 5, 2)$ und $v_2 = (5, 0, 3)$, und eine weitere durch $w_1 = (1, 1, 1)$ und $w_2 = (2, -3, 0)$. Mit solchen Basen erhalten wir Identifikationen $\mathbb{K}^2 \to E$ und somit eine numerische Beschreibung der Abbildung $\mathbb{K}^2 \cong E \to \mathbb{K}$, womit wir die partiellen Ableitungen bzgl. 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.