Differenzierbare Abbildungen in lokalkonvexen Vektorräumen

Seminar zur Funktionalanalysis

Jann D. Zurhelle

11.1.2012

1 Lokalkonvexe Vektorräume

Definition 1.1. a) Ein topologischer Vektorraum (TVR) ist ein Vektorraum X über $\mathbb{R}$ (oder $\mathbb{C}$), der ein topologischer Hausdorffraum ist, so dass die Abbildungen

$$\begin{aligned} X \times X &\to X & (x, y) &\mapsto x + y \\ \mathbb{R} \times X &\to X & (t, x) &\mapsto t \cdot x \end{aligned}$$

stetig sind.

b) Eine Halbnorm auf einem Vektorraum X ist eine Abbildung $p : X \to \mathbb{R}_0^+$, so dass

$$\begin{aligned} p(x + y) &\leq p(x) + p(y) & &\forall x, y \in X \\ p(t \cdot x) &= |t| p(x) & &\forall t \in \mathbb{R}, x \in X \end{aligned}$$

c) Eine Familie von Halbnormen $(p_j)_{j \in J}$ mit $p_j : X \to \mathbb{R}$ trennt die Punkte von X, falls gilt:

$$p_j(x) = 0 \qquad \forall j \in J \Leftrightarrow x = 0$$

Satz 1.2. Sei X ein Vektorraum und $(p_j)_{j \in J}$ eine Familie von Halbnormen, die die Punke von X trennt. Dann definiert

$$\mathcal{B}_0 := \{U(p_{j_1}, \epsilon_1) \cap \ldots \cap U(p_{j_n}, \epsilon_n) : \epsilon_1, \ldots, \epsilon_n > 0, j_1, \ldots, j_n \in J\}$$

eine Basis von Nullumgebungen und

$$\mathcal{B} := \{x + U : x \in X, U \in \mathcal{B}_0\}$$

eine Basis einer Vektorraumtopologie auf X.

Beweis. siehe Werner VIII.1.1-1.2 □

Bemerkung 1.3. Dies ist die Initialtopologie bezüglich der p_j.

Bemerkung 1.4. Für stetige Halbnormen p und lineare Operatoren λ gilt, dass

$$p(x) = \sup_{\lambda \leq p} \{\lambda(x)\}$$

Wobei $\lambda \leq p :\Leftrightarrow \lambda(x) \leq p(x) \forall x \in X$ gilt.

Definition 1.5. Ein lokalkonvexer TVR ist ein TVR, dessen Topologie wie im vorigen Satz durch eine Familie $(p_j)_{j \in J}$ von punktetrennenden Halbnormen beschrieben werden kann.

Bemerkung 1.6. Für einen lokalkonvexen TVR ist die Familie der Halbnormen nicht eindeutig.
Für genauere Informationen siehe Werner, Kor VIII.2.2.

Definition 1.7. Für einen TVR X heißt $X' := \{f : X \to \mathbb{R} : f \text{ linear und stetig}\}$ der Dualraum zu X.

Satz 1.8 (Satz von Hahn-Banach (Trennungsversion))**.** Für einen lokalkonvexen TVR trennt X' die Punkte von X, also

$$f(x) = f(y) \quad \forall f \in X' \quad \Leftrightarrow x = y$$

Beweis. siehe Werner VIII.2.13 (analog zu unserem Beweis von Hahn-Banach in der Vorlesung) $\square$

2 Differentialrechnung in lokalkonvexen Vektorräumen

Definition 2.1. Seien X, Y TVR, und sei $U \subseteq X$ offen und $f : U \to Y$ eine Abbildung.
a) Für $v \in X$ ist dann

$$df(x)(v) := \lim_{h \to 0} \frac{1}{h}[f(x + h \cdot v) - f(x)]$$

die Richtungsableitung von f an der Stelle x in Richtung von v, falls der Limes existiert. f heißt differenzierbar, falls $df(x)(v)$ für alle $(x, v) \in U \times X$ existiert und stetig differenzierbar (C^1-Abbildung), wenn außerdem die Abbildung

$$df : U \times X \to Y \qquad (x, v) \mapsto df(x)(v)$$

stetig ist.

b) f heißt C^n-Abbildung, falls f C^1 ist und df C^{n-1}.
f heißt C^∞ oder glatt, falls f eine C^n-Abbildung ist für alle $n \in \mathbb{N}$.

c)Falls f eine C^n-Abbildung ist, so sind die höheren Abbleitungen von f induktiv definiert:

$$d^n(x)(v_1, \ldots, v_n) := \lim_{h \to 0} \frac{1}{h}[d^{n-1}f(x + h \cdot v_n)(v_1, \ldots, v_{n-1}) - d^{n-1}f(x)(v_1, \ldots, v_{n-1})]$$

Beispiel 2.2. Sei f eine beliebige lineare, stetige Funktion. Dann gilt für die Ableitung am Punkt x in Richtung v:

$$df(x)(v) = \lim_{h \to 0} \frac{1}{h}(f(x + h \cdot v) - f(x)) = \lim_{h \to 0} \frac{1}{h} hf(v) = f(v)$$

Damit existiert die Ableitung in allen Punkten und da f stetig war, ist es somit auch df. Also ist f eine C^1-Abbildung.

Bemerkung 2.3. Im Unterschied zur bekannten Definition der Differenzierbarkeit gilt hier nicht, dass eine differenzierbare Funktion auch stetig sein muss.

Als Beispiel betrachten wir den Raum der Polynome, wobei die Norm durch das Supremum der Koeffizienten bestimmt ist. Auf diesem Raum betrachten wir den Ableitungsoperator L, der einem Polynom seine Ableitung zuordnet.

Dieser ist nicht stetig, da es für jedes $\delta > 0$ und jedes $\epsilon > 0$ ein $n \in \mathbb{N}$ gibt, so dass $\delta \cdot n > \epsilon$ gilt. Dann hat das Polynom $p_x = \delta x^n$ die Norm δ, aber $L(p_x) = \delta \cdot nx^{n-1}$ liegt nicht im ϵ-Ball um Null. Also kann L nicht stetig sein. Allerdings ist L linear, wie nach unseren Ableitungsregeln für Polynome klar ist.

Berechnen wir nun die Richtungsableitungen für ein beliebiges Polynom p_x in Richtung q_x, so erhalten wir:

$$dL(p_x)(q_x) = \lim_{h \to 0} \frac{1}{h}(L(p_x + h \cdot q_x) - L(p_x)) = \lim_{h \to 0} \frac{1}{h} hL(q_x) = L(q_x)$$

Also ist L differenzierbar.

Definition 2.4 (schwaches Integral). Sei X ein lokalkonvexer TVR, $I \subseteq \mathbb{R}$ ein Intervall, $\gamma : I \to X$ stetig und $a, b \in I$.

Falls dann ein $v \in X$ existiert, so dass

$$\lambda(v) = \int_a^b \lambda(\gamma(s))ds \qquad \forall \lambda \in X'$$

gilt, so heißt v dass schwache Integral von γ (von a und b) und wir setzen $\int_a^b \gamma(s)ds := v$

Lemma 2.5. Schwache Integrale sind eindeutig.

Beweis. Falls $v = \int_a^b \gamma = v'$ gilt, so folgt

$$\begin{aligned} &\lambda(v) = \int_a^b \lambda(\gamma(t))dt = \lambda(v') \\ &\Rightarrow \lambda(v - v') = 0 \qquad \forall \lambda \in X' \\ &\Rightarrow v - v' = 0 \end{aligned}$$

nach dem Satz von Hahn-Banach. □

Lemma 2.6. Sei $U \subseteq X$ offen. Ist $f : U \to Y$ C^1-Abbildung und $L : Y \to Z$ stetig und linear, dann ist $L \circ f$ C^1-Abbildung mit $d(L \circ f)(x)(v) = L(df(x)(v))$.

Beweis. Es gilt:

$$\begin{aligned} &d(L \circ f)(x)(v) = \lim_{h \to 0} \frac{1}{h}((L \circ f)(x + hv) - (L \circ f)(x)) \\ &= \lim_{h \to 0} L(\frac{1}{h}(f(x + hv) - f(x))) \\ &= L(\lim_{h \to 0} \frac{1}{h}(f(x + hv) - f(x))) \\ &= L(df(x)(v)) \end{aligned}$$

Es bleibt noch zu zeigen, dass $L \circ f$ eine C^1-Abbildung ist. Da $d(L \circ f)$ aber nach der Rechnung dasselbe ist, wie $L(df(x)(v))$, ist die Ableitung eine Verkettung stetiger Abbildungen und damit selbst stetig. □

Lemma 2.7. Ist X lokalkonvex, $0 \ni I \subseteq \mathbb{R}$ offenes Intervall und $\gamma : I \to X$ eine C^1-Abbildung, dann existiert $\int_0^t \gamma'(s)ds$ für alle $t \in I$ und es gilt

$$\gamma(t) = \gamma(0) + \int_0^t \gamma'(s)ds$$

Beweis. Für $t \in I$ und $\lambda \in X'$ haben wir nach der Definition:

$$\lambda(\gamma(t) - \gamma(0)) = \lambda(\gamma(t)) - \lambda(\gamma(0)) = \int_0^t (\lambda \circ \gamma)'(s)ds = \int_0^t \lambda(\gamma'(s))ds$$

Hier benutzen wir bei der ersten Gleichheit die Linearität von λ, bei der zweiten den Hauptsatz der Differential- und Integralrechung und bei der dritten das Lemma 2.6. Also erfüllt $\gamma(t) - \gamma(0)$ die definierende Eigenschaft des schwachen Integrals $\int_0^t \gamma'(s)ds$. $\square$

Satz 2.8. Sei $U \subseteq X$ offen. Es sei $f : U \to Y$ eine C^1-Abbildung.

a) Für jedes $x \in U$ ist $df(x) : X \to Y$ reell linear und stetig.

b) Falls $x + t \cdot v \subseteq U$ $\forall t \in [0,1]$, so existiert $\int_0^1 df(x + t \cdot v)(v)dt$ und es gilt

$$f(x + v) = f(x) + \int_0^1 df(x + t \cdot v)(v)dt$$

Insbesondere ist f lokal konstant genau dann, wenn $df(x) = 0 \qquad \forall x \in U$.

c) f ist stetig.

d) Ist f C^n-Abbildung für n$\geq$ 2, so sind die Abbildungen

$$X^n \to Y \qquad (v_1, \ldots, v_n) \mapsto d^n f(x)(v_1, \ldots, v_n)$$

für alle $x \in U$ stetig, symmetrisch und n-linear.

Beweis. Wir betrachten für $\lambda \in Y$ und $v_1, v_2 \in X$ die Abbildung

$$(t_1, t_2) \mapsto F(t_1, t_2) := \lambda(f(x + t_1 \cdot v_1 + t_2 \cdot v_2))$$

die für (t_1, t_2) aus einer Nullumgebung von $0 \in \mathbb{R}^2$ definiert ist.

Die partiellen Ableitungen von F sind gegeben durch

$$\frac{\partial F}{\partial t_1}(t_1, t_2) := dF(t_1, t_2)(1, 0) = \lambda(df(x + t_1 \cdot v_1 + t_2 \cdot v_2)(v_1)) \text{ und}$$
$$\frac{\partial F}{\partial t_2}(t_1, t_2) := dF(t_1, t_2)(0, 1) = \lambda(df(x + t_1 \cdot v_1 + t_2 \cdot v_2)(v_2))$$

Denn es gilt:

$$\begin{aligned}
&dF(t_1, t_2)(1, 0) \\
&= \lim_{h \to 0} \frac{1}{h}(F(t_1 + h, t_2) - F(t_1, t_2)) \\
&= \lim_{h \to 0} \frac{1}{h}(\lambda(f(x + (t_1 + h) \cdot v_1 + t_2 \cdot v_2)) - \lambda(f(x + t_1 \cdot v_1 + t_2 \cdot v_2))) \\
&= \lim_{h \to 0} \frac{1}{h}(\lambda(f(x + t_1 \cdot v_1 + t_2 \cdot v_2 + h \cdot v_1) - \lambda(f(x + t_1 \cdot v_1 + t_2 \cdot v_2))) \\
&= d(\lambda \circ f)(x + t_1 \cdot v_1 + t_2 \cdot v_2)(v_1) \\
&= \lambda(df(x + t_1 \cdot v_1 + t_2 \cdot v_2)(v_1))
\end{aligned}$$

Insbesondere sind diese stetig, weshalb F stetig differenzierbar ist. F ist somit eine C^1-Abbildung (siehe Analysis II). Ferner ist $dF(0,0) : \mathbb{R}^2 \to \mathbb{R}$ linear, und somit gilt

$$\begin{aligned}
&\lambda(df(x)(v_1 + v_2)) && = d(\lambda \circ f)(x)(v_1 + v_2) \\
&=: dF(0,0)(1,1) && = dF(0,0)(1,0) + dF(0,0)(0,1) \\
&= \lambda(df(x)(v_1)) + \lambda(df(x)(v_2)) && = \lambda(df(x)(v_1) + df(x)(v_2)) \\
&\Rightarrow \lambda(df(x)(v_1 + v_2)) && = \lambda(df(x)(v_1) + df(v_2)) \forall \lambda \in X', v_1, v_2 \in X \\
&\Rightarrow df(x)(v_1 + v_2) && = df(x)(v_1) + df(x)(v_2) \forall v_1, v_2 \in X
\end{aligned}$$

Wobei wir das Lemma 2.6 und den Satz von Hahn-Banach verwendet haben.
Genauso folgt auch $df(x)(t \cdot v) = t \cdot df(v) \qquad \forall t \in \mathbb{R}, v \in X$.
Da df nach Definition stetig ist und die Inklusion $X \hookrightarrow U \times X$ stetig ist, ist $df(x)$ stetig.

b) Wir betrachten Lemma 2.7 mit $\gamma(t) := f(x + t \cdot v)$ und $t = 1$. Dann folgt $f(x + v) = \gamma(1) = \gamma(0) + \int_0^1 \gamma'(t)dt = f(x + 0 \cdot v) + \int_0^1 df(x + t \cdot v)dt$, was genau die Aussage ist, die wir zeigen wollten.

c) Sei die Topologie auf X durch die punktetrennende Familie $(p_j)_{j \in J}$ definiert. Dann ist zu zeigen, dass $\forall j \in J$ gilt $U \ni x \mapsto p_j(f(x))$ ist stetig.

Aus der Bemerkung über Halbnormen und dem Satz von Hahn-Banach folgt

$$p_j\left(\int_a^b \gamma(t)dt\right) \leq \int_a^b p_j(\gamma(t))dt$$

für jede C^1-Abbildung $\gamma : [a, b] \to X$.

Sei nun $j \in J$ fest. Da df stetig ist, ist $p_j \circ df$ stetig. Also ist für festes x die Abbildung $(t, v) \mapsto p_j(df(x + tv)(v))$ stetig in beiden Argumenten. Da $(t, 0)$ auf 0 abgebildet wird, existieren offene Umgebungen V_t von t und U_t von 0, so dass $p_j(df(x+sw)(w)) < \epsilon$ fuer alle $s \in V_t$ und $w \in U_t$. Die V_t bilden eine offene Überdeckung von [0,1]. Da $[0, 1]$ kompakt ist, gibt es somit eine endliche Teilüberdeckung durch $V_{t_1}, \dots, V_{t_n}$. Zu diesen $t_1, \dots, t_n$ haben wir nur noch endlich viele U_t. Betrachten wir nun ihren Schnitt U_1, so ist dieser offen und da für jedes $t \in [0, 1]$ gilt, dass t in einem der V_{t_k} liegt, gilt:
Es existiert eine offene Nullumgebung $U_1 \subseteq X$ mit $x + U_1 \subseteq U$ und $p_j(df(x + t \cdot v)(v)) \leq \epsilon$ für $t \in [0, 1]$ und $v \in U_1$.

$$\Rightarrow |p_j(f(x + v)) - p_j(f(x))| \leq |p_j(f(x + v) - f(x))| \leq |\int_0^1 p_j(df(x + t \cdot v)(v))dt| \leq \epsilon$$

wobei die letzte Abschätzung aus der für gewöhnliche stetige Funktionen $\mathbb{R} \to \mathbb{R}$ folgt. Also ist $p_j \circ f$ stetig und da j beliebig war, ist somit f stetig.

d)Analog zu der Abbildung in a) definieren wir uns jetzt die Abbildung

$$(t_1, \ldots, t_n) \mapsto f(x + t_1 \cdot v_1 + \ldots + t_n \cdot v_n)$$

für $v_1, \ldots, v_n$ aus einer entsprechenden Nullumgebung in X^n. Dann erhalten wir die gewünschte Aussage aus denselben Rechnungen, wie in a). $\square$