Seminararbeit

Satz von Rademacher und weiterführende Resultate

ausgeführt am

Institut für
Analysis und Scientific Computing
TU Wien

unter der Anleitung von

Ao.Univ.Prof. Dipl.-Ing. Dr.techn.

Michael Kaltenbäck

durch

Katerina Uncovska

Matrikelnummer: 1427514

Wien, am 27. Juni 2018

Inhaltsverzeichnis

1 Einleitung

Zentral in dieser Arbeit ist der Satz von Rademacher, der eine wichtige Aussage über Differenzierbarkeit in endlichdimensionalen Banachräumen liefert. Sie orientiert sich vor allem an den Ergebnissen in [DiB02] und [EG15]. Nach dem Einführen des relevanten Werkzeugs wird zunächst der eindimensionale, dann der allgemeine Fall betrachtet. Schließlich wird ein kleiner Einblick in Räume mit Radon Nikodým Eigenschaft, vor allem Ergebnisse aus [BL00], in denen ähnliche, wenn auch schwächere Aussagen über Differenzierbarkeit getroffen werden können, geboten.

2 Satz von Rademacher

2.1 Relevante Definitionen

Bevor wir zur Aussage und zum Beweis des Satzes von Rademacher kommen, müssen zunächst einige Begriffe eingeführt werden. Im Folgenden werden alle Definitionen und Sätze für auf $\mathbb{R}$, beziehungsweise auf $\mathbb{R}^N$, mit $N \in \mathbb{N}$ definierte Funktionen formuliert. Analog sind diese auch für offene Teilmengen von $\mathbb{R}$ (bzw. $\mathbb{R}^n$) gültig. Mit$\|\cdot\|$ bezeichnen wir die Zweinorm $\|\cdot\|_2 =$ auf $\mathbb{R}^n$. Da alle Normen auf $\mathbb{R}^n$ äquivalent sind, gelten alle Resultate analog für andere Normen.

Definition 2.1.1 Eine Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ heißt *Lipschitz-stetig*, wenn es eine Konstante L gibt, sodass

$$|f(x) - f(y)| \leq L \|x - y\| \quad \text{für alle } x, y \in \mathbb{R}^n.$$

Definition 2.1.2 Eine Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ heißt dementsprechend *lokal Lipschitz-stetig*, wenn es für jede kompakte Teilmenge K von $\mathbb{R}^n$ eine Konstante L_K gibt, sodass die Lipschitz-Bedingung auf K erfüllt ist.

Jede Lipschitz-stetige Funktion ist auch gleichmäßig stetig. Die Umkehrung gilt nicht, wie man im folgenden Beispiel sieht.

Beispiel 2.1.3 Man betrachte die Wurzelfunktion, definiert auf $[0, \infty)$. Wäre diese Lipschitz stetig, so gäbe es nun ein $L > 0$, sodass $\left|\sqrt{x} - \sqrt{y}\right| \leq L \cdot |x - y|$ auf ganz $[0, \infty)$. Erweitert man die linke Seite mit $\frac{|\sqrt{x}+\sqrt{y}|}{|\sqrt{x}+\sqrt{y}|}$, so ergibt sich $\frac{1}{|\sqrt{x}+\sqrt{y}|} \leq L$. Für $x = \frac{1}{4L^2}$, $y = 0$ erhalten wir einen Widerspruch. Wählt man $\delta = \varepsilon^2$ zu $\varepsilon > 0$, dann folgt, mit o.B.d.A. $y \leq x$, dass $y \leq x + \varepsilon^2 \leq (\sqrt{x} + \varepsilon)^2$ und somit $\left|\sqrt{x} - \sqrt{y}\right| \leq \varepsilon$. Damit ist gleichmäßige Stetigkeit gezeigt.

Definition 2.1.4 Sei $f\colon \mathbb{R}^n \to \mathbb{R}$. Für beliebiges $v \in \mathbb{R}^n$ bezeichnet man mit

$$D_v f(x) := \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}, \quad x \in \mathbb{R}^n,$$

die *Richtungsbleitung von f an der Stelle x nach v*, falls dieser Grenzwert existiert.

Die Richtungsableitungen nach den Einheitsvektoren $e_1, .., e_n \in \mathbb{R}^n$ werden auch mit $f_{x_1}(x), .., f_{x_N}(x)$ bezeichnet, der *Gradient* von f ist $\nabla f(x) = (f_{x_1}(x), .., f_{x_n}(x))^T$.

Weiters ist f *partiell differenzierbar bei $x \in \mathbb{R}^n$*, wenn die Richtungsableitung $D_v(x)$ für $v = e_1, \ldots, e_n$ existiert. Sind diese darüber hinaus stetig, heißt f *stetig partiell differenzierbar*. In dem Fall schreiben schreiben wir $f \in C^1(\mathbb{R}^n)$.

Nicht mit diesen Ausdrücken zu verwechseln ist der Begriff der Differenzierbarkeit, der in der Literatur auch als *Fréchet-Differenzierbarkeit* oder *totale Differenzierbarkeit* bezeichnet wird.

Definition 2.1.5 Eine Funktion $f \colon \mathbb{R}^n \to \mathbb{R}$ ist *differenzierbar bei $x \in \mathbb{R}^n$*, wenn es ein $Df(x) \in \mathbb{R}^{1\times n}$ gibt, sodass

$$f(y) = f(x) + Df(x)(y - x) + \varepsilon(x - y), \text{ für } y \in U_\rho(x),$$

erfüllt ist. Hierbei ist $\rho > 0$ und $\varepsilon \colon U_\rho(0)\backslash 0 \to \mathbb{R}$ derart, dass $\lim_{y\to x} \frac{\varepsilon(x-y)}{\|x-y\|} = 0$.

Elementar zeigt man: Gibt es ein solches $Df(x)$, so stimmt dieses mit $\nabla f(x)^\top$ überein. Also folgt Partielle Differenzierbarkeit aus der Differenzierbarkeit. Umgekehrt lässt sich jedoch nicht von der Existenz von $\nabla f(x)$ auf die Differenzierbarkeit bei x schließen. Aus der stetigen partiellen Differenzierbarkeit folgt aber sehr wohl die Differenzierbarkeit.

2.2 Funktionen auf $\mathbb{R}$

Für den Beweis des Satzes von Rademacher ist noch eine Aussage zur Differenzierbarkeit für Funktionen auf $\mathbb{R}$ nötig.

Definition 2.2.1 Eine Funkion $f \colon \mathbb{R} \to \mathbb{R}$ heißt *absolut stetig*, falls es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, sodass für jede endliche Menge paarweise disjunkter Intervalle $\{(x_k, y_k)\}_{k=1}^n$ mit Gesamtlänge $\sum_{k=1}^n (y_k - x_k) < \delta$

$$\sum_{k=1}^n |f(y_k) - f(x_k)| < \varepsilon$$

gilt.

Ein wichtiges Resultat ist, dass Lipschitz-stetige Funktionen auf $\mathbb{R}$ absolut stetig sind.

Lemma 2.2.2 *Eine Lipschitz-stetige Funktion $f \colon \mathbb{R} \to \mathbb{R}$ ist absolut stetig.*

Beweis. Für $\varepsilon > 0$ wähle $\delta := \frac{\varepsilon}{L}$. Dann folgt für alle $\{(x_k, y_k)\}_{k=1}^n$ mit Gesamtlänge $\sum_{k=1}^n (y_k - x_k) < \delta$

$$\sum_{k=1}^n |f(y_k) - f(x_k)| \leq \sum_{k=1}^n L \cdot (y_k - x_k) = L \cdot \sum_{k=1}^n (y_k - x_k) < L \cdot \delta = L \frac{\varepsilon}{L} = \varepsilon.$$

□

Definition 2.2.3 Eine Funktion $f\colon [a, b] \to \mathbb{R}$ heißt *von beschränkter Variation ($f \in BV([a, b])$)*, falls $V_a^b(f) < +\infty$, wobei

$$V_a^b(f) := \sup\{\sum_{i=1}^n |f(x_i) - f(x_{i-1})| : a = x_0 < x_1 < .. < x_n = b, n \in \mathbb{N}\}.$$

Ein bekanntes Ergebnis aus der Maßtheorie ist, dass absolut stetige Funktionen auf jedem Intervall $[a, b]$ von beschränkter Variation sind [Kus14, Kapitel 12.2/3]. Der Satz von Lebesgue - siehe [Kus14, Kapitel 12.2/3] - liefert dann die Differenzierbarkeit fast überall auf genanntem Intervall; also ist f außerhalb einer Lebesgue-Nullmenge differenzierbar. Da Differenzierbarkeit eine lokale Eigenschaft ist, lässt sich dieses Resultat dann für absolut stetige Funktionen auf ganz $\mathbb{R}$ übertragen. Wir erhalten also:

Lemma 2.2.4 *Eine absolut stetige Funktion $f\colon \mathbb{R} \to \mathbb{R}$ ist f.ü. differenzierbar.*

Dieses Ergebnis lässt sich mit *Lemma 2.2.2* für lokal Lipschitz stetige Funktionen kombinieren. Wir erhalten eine Aussage, die wir für den Beweis vom Satz von Rademacher benötigen werden.

Korollar 2.2.5 *Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine lokal Lipschitz-stetige Funktion. Dann ist f f.ü. differenzierbar.*

2.3 Satz von Rademacher

Wir kommen zur Hauptaussage dieser Arbeit.

Satz 2.3.1 (SATZ VON RADEMACHER). *Sei $f\colon \mathbb{R}^n \to \mathbb{R}$ eine lokal Lipschitz-stetige Funktion. Dann ist f f.ü. differenzierbar.*

Beweis. Der Beweis wird in drei Schritten erfolgen. Für den ersten und zweiten Teil betrachte Richtungsableitungen $D_v f(x)$ für $\|v\| = 1$.

1. Behauptung: $D_v f(x)$ existiert für fast alle $x \in \mathbb{R}^n$.

Wie definieren

$$\overline{D_v}f(x) = \limsup_{t\to 0} \frac{f(x+tv)-f(x)}{t},$$
$$\underline{D_v}f(x) = \liminf_{t\to 0} \frac{f(x+tv)-f(x)}{t}.$$

Bezeichne mit E_v die Menge, auf der $D_v f$ nicht definiert ist, genauer

$$E_v = \{x \in \mathbb{R}^n | \underline{D_v}f(x) < \overline{D_v}f(x)\}.$$

Da f stetig ist, kann der Grenzwert $\overline{D_v}f(x)$ bzw. $\underline{D_v}f(x)$ für $t \in (-1,1)\cap\mathbb{Q}$ gebildet werden. Es folgt, dass $\underline{D_v}f(x)$ und $\overline{D_v}f(x)$ messbar sind und dementsprechend auch E_v messbar ist [Kus14, Kapitel 7.2].
Mit f ist auch $\phi\colon \mathbb{R}\to\mathbb{R}, t\mapsto f(x+tv)$ für jedes $x\in\mathbb{R}^n$ lokal Lipschitz-stetig, da

$$\begin{aligned}|\phi(t)-\phi(s)| = |f(x+tv)-f(x+sv)| &\leq L\,\|x+tv-x-sv\| \\ &= L\cdot|t-s|\,\|v\| = L\,|t-s|\,.\end{aligned}$$

Aus *Korollar 2.2.5* folgt damit für ϕ die Differenzierbarkeit fast überall auf $\mathbb{R}$. Wegen $\phi'(t) = D_v f(x+tv)$ hat der Schnitt von E_v mit jeder zu v parallelen Linie L eindimensionales Lebesgue-Maß null. Mit dem Satz von Fubini [Kus14] folgt

$$\lambda^n(E_v) = \int_{v^\perp} \lambda(E_v\cap L)\; d\lambda^{n-1} = 0.$$

2. Behauptung: $D_v f(x) = \nabla f(x)^\top v$ für fast alle x.

Auf der Menge $A = \mathbb{R}^n \backslash (E_v \cup \bigcup_{j=1}^n E_{e_j})$ existieren alle Richtungsableitungen von f nach v sowie der Gradient. Betrachte $t\in(-1,1)\cap\mathbb{Q}$ und bezeichne mit $e_j,\ j=1,\dots,n$, die Einheitsvektoren von $\mathbb{R}^n$. Da f lokal Lipschitz-stetig ist, gilt für alle beliebig oft differenzierbaren Funktionen mit kompaktem Träger $\zeta\in C_0^\infty(\mathbb{R}^n)$,

$$\left|\frac{f(x+tv)-f(x)}{t}\zeta(x)\right| \leq L_\zeta|\zeta(x)|,$$

wobei L_ζ die Lipschitz Konstante von f auf einer hinreichend großen, den Träger von ζ enthaltenden Kugel ist. Somit sind die Bedingungen für den Satz von dominierter Konvergenz [Kus14] erfüllt. Integriert man auf der Menge A, so folgt

$$\begin{aligned}\int_A D_v f(x)\zeta(x)\;dx &= \lim_{t\to 0}\int_A \frac{f(x+tv)-f(x)}{t}\zeta(x)\;dx \\ &= \lim_{t\to 0}(\int_A \frac{f(x+tv)}{t}\zeta(x)\;dx - \int_A \frac{f(x)}{t}\zeta(x)\;dx)\end{aligned}$$

Substituiert man im linken Integral mit $x-tv$, so erhält man

$$\begin{aligned}&= \lim_{t\to 0}(\int_A \frac{f(x)}{t}\zeta(x-tv)\;dx - \int_A \frac{f(x)}{t}\zeta(x)\;dx) \\ &= -\lim_{t\to 0}\int_A f(x)\frac{\zeta(x)-\zeta(x-tv)}{t}\;dx.\end{aligned}$$

Mit dem selben Argument wie zuvor, kann man jetzt mit dominierter Konvergenz den Limes und das Integral vertauschen. Da ζ aus $C_0^\infty(\mathbb{R}^N)$ gewählt wurde, ergibt das

$$-\sum_{j=1}^{n} v_j \int_A f(x)\zeta_{x_j}(x)\,dx = -\sum_{j=1}^{n} v_j \lim_{t\to 0}\int_A f(x)\frac{\zeta(x)-\zeta(x-te_j)}{t}\,dx.$$

Nun wird, ähnlich wie zuvor, mit $v + te_j$ substituiert und dominierte Konvergenz angewendet. Also stimmt obiger Ausdruck überein mit

$$\begin{aligned}&\sum_{j=1}^{n} v_j \lim_{t\to 0}\int_A \frac{f(x+te_j)-f(x)}{t}\zeta(x)\,dx\\ &=\int_A v\cdot\nabla f(x)\zeta(x)\,dx\end{aligned}$$

Das impliziert

$$\int_A (D_v f(x) - \nabla f(x)^\top v)\zeta(x)\,dx = 0$$

für alle $\zeta \in C_0^\infty(\mathbb{R}^n)$. Nach [Kus14] folgt daraus $D_v(x)f = \nabla f(x)^\top v$ fast überall auf A. Da nach dem ersten Teil $E_v \cup \bigcup_{j=1}^n E_{e_j}$ Lebesgue-Maß Null hat, erhalten wir $D_v(x)f = v\cdot\nabla f(x)$ fast überall auf $\mathbb{R}^n$.

Sei nun $B = \{u_k | k\in\mathbb{N}\}$ eine abzählbare dichte Teilmenge auf dem Rand der Einheitskugel ∂B_1, sodass alle Einheitsvektoren e_j darin enthalten sind. Betrachte

$$F_{u_k} := \{x\in\mathbb{R}^n | D_{u_k}f(x) \text{ existiert und } D_{u_k}f(x) = \nabla f(x)^\top u_k\}.$$

Nach dem bereits Gezeigten hat $\mathbb{R}^n \setminus \bigcap\limits_{k\in\mathbb{N}} F_{u_k}$ Lebesgue-Maß Null.

Um das Theorem zu beweisen, genügt es nun die Differenzierbarkeit von f auf $F = \bigcap\limits_{k\in\mathbb{N}} F_{u_k}$ zu zeigen.

3. Behauptung: f ist differenzierbar auf F.
Für $x\in F$, $u\in\partial B_1$ und $t\in[-1,1]\setminus\{0\}$ definieren wir

$$Q(x,u,t) := \frac{f(x+tu)-f(x)}{t} - \nabla f(x)^\top u.$$

Sei weiters L die Lipschitzkonstante von f auf $K_1(x)$. Für beliebeige $v\in\partial B_1$ gilt

$$\begin{aligned}|Q(x,u,t)-Q(x,v,t)| &\le \left|\frac{f(x+tu)-f(x+tv)}{t}\right| + \left|\nabla f(x)^\top(u-v)\right|\\ &\le L\,\|u-v\| + \|\nabla f(x)\|\,\|u-v\|\end{aligned}$$

$$\leq L(\sqrt{n}+1)\left\|u-v\right\|. \tag{1}$$

Die letzte Ungleichung gilt, da $|f_{x_j}| = \left|\lim_{t\to 0} \frac{f(x+te_j)-f(x)}{t}\right| \leq L \cdot \frac{\|x+t-x\|}{t} = L$ und somit

$$\|\nabla f\| \leq \sqrt{n} \cdot L.$$

Für $\varepsilon > 0$ wähle $N(\varepsilon) \in \mathbb{N}$ groß genug, sodass

$$\sup_{\|u\|=1} \min_{k=1,\dots,N(\varepsilon)} \|u-u_k\| \leq \frac{\varepsilon}{2L(\sqrt{n}+1)}. \tag{2}$$

Wegen der Dichteeigenschaft der Menge B und der Kompaktheit von ∂B gibt es ein solches $N(\varepsilon)$. Weiters gilt nach Definition der Menge F

$$\lim_{t\to 0} Q(x,u_k,t) = 0 \quad \text{für } k \in \{1,\dots,N(\varepsilon)\}.$$

Es folgt die Existenz eines $\delta > 0$ mit o.B.d.A. $\delta \leq 1$, sodass

$$|Q(x,u_k,t)| < \frac{\epsilon}{2} \quad \text{für alle } 0 < |t| < \delta \;\text{ und } k = 1,\dots,N(\varepsilon). \tag{3}$$

Zusammen folgt aus den Gleichungen (1)-(3) für alle $u \in \partial B_1$

$$|Q(x,u,t)| \leq |Q(x,t,u_k)| + |Q(x,u,t) - Q(x,u_k,t)| < \varepsilon \quad \text{für } 0 < |t| < \delta,$$

wobei k so gewählt wird, dass $\|u-u_k\| = \min_{j=1,\dots,N(\varepsilon)} \|u-u_j\|$.
Wähle $y \neq x$ aus $\mathbb{R}^n$ und setze $u := \frac{y-x}{\|y-x\|}$ und $t = \|y-x\|$. Offensichtlich ist dann $y = x + tu$. Es folgt

$$\begin{aligned} f(y) - f(x) - \nabla f(x)^\top (y-x) &= f(x+tu) - f(x) - t\nabla f(x)^\top u \\ &= Q(x, \frac{y-x}{\|y-x\|}, \|y-x\|) \cdot \|y-x\|, \end{aligned}$$

mit

$$\left|Q(x, \frac{y-x}{\|y-x\|}, \|y-x\|)\right| < \varepsilon \qquad \text{für } 0 < \|y-x\| < \delta.$$

Somit ist f bei x differenzierbar, wobei $Df(x) = \nabla f(x)^\top$. □

3 Weiterführende Resultate

Der Satz von Rademacher ist eine sehr schöne und starke Aussage auf endlichdimensionalen Vektorräumen. Aber wie sieht das mit unendlichdimensionalen Räumen aus? Es lassen sich tatsächlich ähnliche Resultate finden. Eines davon betrachtet Räume mit der Radon-Nikodým Eigenschaft.
Da es vom Inhalt dieser Arbeit zu stark abweicht, werden Ausdrücke wie Radon-Nikodým Ableitung, absolute Stetigkeit von Maßen und damit verbundene Eigenschaften vorausgesetzt. Die zuvor verwendeten Definitionen für absolute Stetigkeit und beschränkte Variation sind hier analog, mit der Norm statt dem Betrag auf die Funktionswerte angewandt, zu verwenden.

3.1 Differenzierbarkeit in Räumen mit RNP

Definition 3.1.1 Sei E ein Banachraum, $C \subseteq E$ abgeschlossen, konvex und beschränkt. Man sagt, *C habe RNP (Radon Nikodým Property)*, wenn die folgene Eigenschaft erfüllt ist:

Ist $(\Omega, \mathcal{A})$ ein Maßraum, τ ein E-wertiges Maß und μ ein skalares Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$, sodass $\frac{\tau(A)}{\mu(A)} \in C$ für alle $A \in \mathcal{A}$ mit $\mu(A) \neq 0$, dann existiert ein $f \in L^1(\mu, E)$, sodass $\tau(A) = \int_A f(\omega)\, d\mu(\omega)$ für alle $A \in \mathcal{A}$. Dieses Integral existiert im Sinne von Bochner.

Man sagt E habe RNP, wenn die Einheitskugel RNP hat.

Offenbar hat eine abgeschlossene, konvexe, beschränkte Teilmenge eines Banachraumes E RNP, falls E selbst RNP hat. Diese Definition mag zunächst ein wenig abstrakt wirken. Zum besseren Verständnis ist die folgende charakteristische Eigenschaft hilfreich.

Lemma 3.1.2 *Der Banachraum E hat genau dann RNP, wenn jede absolut stetige Funktion $f : \mathbb{R} \to E$ differenzierbar f.ü. ist.*

Den Beweis kann man bei *Theorem 5.21,* [BL00, Kapitel 5.3] nachlesen.

Auf unendlichdimensionalen Räumen müssen wir zudem die Begriffe der Differenzierbarkeit einführen. Differenzierbarkeit (Fréchet Differenzierbarkeit) lässt sich genau wie auf $\mathbb{R}^N$ auch auf beliebigen Banachräumen definieren. Um ein Analogon zur partiellen Differenzierbarkeit zu erhalten, muss aber noch weiter ausgeholt werden.

Definition 3.1.3 Seien E und F Banachräume. Eine Funktion $f \colon E \to F$ ist *Gâteaux differenzierbar bei $x \in E$*, wenn für jedes $v \in E$ die Richtungsableitung $D_v f(x)$ existiert und $v \mapsto D_v f(x)$ eine lineare, stetige Funktion ist.

Kommen wir nun zu dem zu Beginn angesprochenen Resultat über Differenzierbarkeit in Räumen mit RNP.

Satz 3.1.4 *Sei F ein Banachraum mit RNP und $f \colon \mathbb{R}^n \to F$ lokal Lipschitz-stetig. Dann ist f fast überall Gâteaux differenzierbar auf $\mathbb{R}^n$.*

Für den Beweis des Satzes wird noch ein technisches Lemma benötigt.

Lemma 3.1.5 *Seien E, F Banachräume über $\mathbb{R}$, $U \subseteq E$ offen und $f \colon U \to F$ eine Lipschitz stetige Funktion. Sei weiters $G \subseteq E$ eine dichte, abzählbare Untergruppe von $(E, +)$.*
Angenommen, es gibt ein $x_0 \in E$, sodass für alle $u \in G$ die Richtungsableitungen $D_u f(x_0)$ existieren und die Funktion $u \mapsto D_u f(x_0)$ additiv ist. Dann ist f Gâteaux differenzierbar bei x_0.

Beweis. Betrachte Funktionen $h_t(u) := \frac{f(x_0+tu)-f(x_0)}{t}$, $t \neq 0$. Dabei gilt $\|h_t(u)\| \leq L\|u\|$. Außerdem gilt

$$\|h_t(u) - h_t(v)\| = \left\| \frac{f(x_0 + tu) - f(x_0 + tv)}{t} \right\| \leq L\|u - v\|.$$

Also gibt es zu einem gegebenen $\epsilon > 0$ und $v \in E$ ein $\delta > 0$, sodass für alle $u \in G$ mit $\|u - v\| < \delta$ die Ungleichung $\|h_t(v) - h_t(u)\| \leq \frac{\epsilon}{3}$ für alle $t \in \mathbb{R}, t \neq 0$, gilt. Für solche u existiert $\lim_{t\to 0} h_t(u)$, wodurch

$$\|h_t(v) - h_s(v)\| \leq \|h_t(v) - h_t(u)\| + \|h_t(u) - h_s(u)\| + \|h_s(u) - h_s(v)\| < \epsilon$$

für t und s klein genug, sodass $\|h_t(u) - h_s(u)\| < \frac{\epsilon}{3}$. Also existiert $\lim_{t\to 0} h_t(v)$ für alle $v \in E$. Die Additivität des Grenzwertes auf G hat auch die Additivität auf E zu Folge und per Definition gilt $\lim_{t\to 0} h_t(\alpha \cdot u) = \alpha \lim_{t\to 0} h_t(u)$. Also ist $\lim_{t\to 0} h_t(\cdot)$ ein linearer Operator, der mit $\|h_t(u)\| \leq L\|u\|$ durch die Lipschitzkonstante beschränkt ist. □

Ein weiterer Begriff, der im Beweis wichtig ist, ist der von *Lebesgue-Punkten* einer Funktion.

Definition 3.1.6 Sei $f\colon \mathbb{R}^n \to \mathbb{R}$ eine lokal Lebesgue integrierbare Funktion. Dann heißt $x \in \mathbb{R}^n$ ein *Lebesgue-Punkt von* f, wenn

$$\lim_{r\to 0} \frac{1}{r} \int_{|y|<r} |f(x - y) - f(x)| \; dy = 0.$$

Nach *Theorem 1.37* [EG15, Kapitel 1.7.2] sind für lokal integrierbare Funktionen fast alle $x \in \mathbb{R}$ Lebesgue-Punkte.

Beweis. [von *Satz 3.1.4*] Sei $G \subseteq \mathbb{R}^n$ eine abzählbare, dichte additive Untergruppe. Eine solche ist zum Beispiel $\mathbb{Q}^n$.
Sei $\phi\colon \mathbb{R} \to F, t \mapsto f(x + tv)$. Nach *Lemma 3.1.2* ist ϕ differenzierbar fast überall mit $\phi'(t) = D_v f(x+tv)$. Bezeichne weiters mit E_v die Menge, auf der die Richtungsableitung nach einem $v \in G$ nicht existiert. Analog zum Beweis des Satzes von Rademacher lässt sich der Satz von Fubini anwenden, um zu zeigen, dass E_v eine Lebesgue-Nullmenge ist. Da G abzählbar ist, folgt die Existenz der Richtungsableitungen nach allen $v \in G$ von f fast überall. Nach *Lemma 3.1.5* genügt es zu zeigen, dass diese Richtungsableitungen nach $v \in G$ fast überall additiv sind. Hierzu werden Faltungen verwendet.

Betrachte eine stetig differenzierbare Funktion $\psi\colon \mathbb{R}^n \to [0, +\infty)$ mit kompaktem Träger und $\int_{\mathbb{R}^n} \psi(x)dx = 1$. Dann ist auch

$$g := f * \psi = \int_{\mathbb{R}^n} \psi(x - y)f(y) \; dy$$

stetig differenzierbar. Also ist für jedes $x \in \mathbb{R}^n$

$$D_u g(x) = \int_{\mathbb{R}^n} D_u\psi(x - y)f(y)\, dy = (f * D_u\psi)(x)$$

linear in $u \in \mathbb{R}^n$.
Andererseits folgt für fast alle x und für alle $u \in G$ mit dominierter Konvergenz

$$D_u g(x) = \lim_{t\to 0} \frac{g(x+tu)-g(x)}{t} = \int_{\mathbb{R}^n} \psi(y) \lim_{t\to 0} \frac{f(x-y+tu)-f(x-y)}{t}\, dy.$$

Also gilt $D_u g(x) = (\psi * h_u)(x)$ mit der beschränkten, messbaren Funktion $h_u(x) = \lim_{t\to 0} \frac{f(x+tu)-f(x)}{t}$. Es folgt

$$\begin{aligned} 0 = f * (D_{u+v}\psi - D_u\psi - D_v\psi) &= D_{u+v}g(x) - D_u g(x) - D_v g(x) \\ &= \psi * (h_{u+v} - h_u - h_v) \end{aligned}$$

für alle $u, v \in G$. Diese Argumentation funktioniert ebenso für $\psi_k := k^n \psi(kx)$, mit einer beliebigen natürlichen Zahl k, und somit auch für den Grenzwert $\lim_{k\to\infty} \psi_k * h_u$. Um diesen zu berechnen, betrachte $\lim_{k\to\infty} \psi_k * Th_u(x)$ mit einem linearen und beschränkten Funktional $T\colon F \to \mathbb{R}$. Es ergibt sich

$$\begin{aligned} \lim_{k\to\infty} \psi_k * Th_u(x) - Th_u(x) &= \lim_{k\to\infty} \int_{\mathbb{R}^n} k^n \psi(ky) \cdot Th_u(x-y)\, dy - Th_u(x) \\ &= \lim_{k\to\infty} \int_{\mathbb{R}^n} \psi(y) \cdot (Th_u(x - \frac{y}{k}) - Th_u(x))\, dy. \end{aligned} \tag{4}$$

Als Grenzwert messbarer Funktionen ist h_u und somit auch die Zusammensetzung Th_u messbar. Weiters ist Th_u durch $|Th_u(x)| \leq \|T\| \cdot L\, \|x\|$ beschränkt, also lokal integrierbar und besteht dementsprechend f.ü. aus Lebesgue-Punkten.
Das Integral in (4) ist kleiner als

$$\begin{aligned} &\int_{|y|<1} \psi(y) \left|(Th_u(x-\frac{y}{k}) - Th_u(x))\right|\, dy + \int_{|y|\geq 1} \psi(y) \left|(Th_u(x-\frac{y}{k}) - Th_u(x))\right|\, dy \\ &\leq \sup \psi(y) \int_{|y|<\frac{1}{k}} |(Th_u(x-y) - Th_u(x))|\, dy + \|T\|\, L \frac{1}{k} \int_{|y|\geq 1} \psi(y)\, |y|\, dy. \end{aligned} \tag{5}$$

Für den linken Teil von (5) findet sich nun auf allen Lebesgue-Punkten von Th_u für jedes $\varepsilon > 0$ ein $K \in \mathbb{N}$ groß genug, sodass

$$\int_{|y|<\frac{1}{k}} |(Th_u(x-y) - Th_u(x))|\, dy \leq \frac{1}{k}\varepsilon$$

für alle $k \geq K$.
Für den rechten Teil von (5) gibt es, da ψ einen kompakten Träger hat, ein $R \in \mathbb{R}$, sodass $supp\, \psi \subset B_R(0)$ gilt. Es folgt

$$\int_{|y|\geq 1} \psi(y)\, |y|\, dy = \int_{1\leq|y|\leq R} \psi(y)\, |y|\, dy \leq R^2 \sup \psi(y).$$

Fasst man nun alle Konstanten zu einer Konstante C zusammen, lässt sich (5) durch

$$\leq \frac{1}{k} C(\varepsilon + 1)$$

für alle $k \geq K$ abschätzen. Für $k \to +\infty$ konvergiert der Ausdruck gegen Null und es folgt fast überall

$$\lim_{k\to\infty} \psi_k * Th_u(x) = Th_u(x).$$

Weiters ist $\lim_{k\to\infty} \psi_k * Th_u(x) = T \lim_{k\to\infty} \psi_k * h_u(x)$ für alle $x \in \mathbb{R}^n$. Da F' punktetrennend auf F operiert (vgl. [WKB, Kapitel 5.2]), folgt daraus $\lim_{k\to\infty} \psi_k * h_u(x) = h_u(x)$ fast überall. Da G abzählbar ist, folgt für alle $u, v \in G : h_{u+v} = h_u + h_v$ fast überall auf $\mathbb{R}^n$. $\square$

Literatur

[BL00] Yoav Benyamini and Joram Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000.

[DiB02] Emmanuele DiBenedetto. *Real analysis*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Boston, Inc., Boston, MA, 2002.

[EG15] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.

[Kal14] Michael Kaltenbäck. *Fundament Analysis*, volume 26 of *Berliner Studienreihe zur Mathematik [Berlin Study Series on Mathematics]*. Heldermann Verlag, Lemgo, 2014.

[Kus14] Norbert Kusolitsch. *Maß- und Wahrscheinlichkeitstheorie*. Springer-Lehrbuch. [Springer Textbook]. Springer Spektrum, Berlin Heidelberg, expanded edition, 2014. Eine Einführung. [An introduction].

[WKB] Harald Woracek, Michael Kaltenbäck, and Martin Blümlinger. *Funktionalanalysis*. Vorlesungsskriptum Version SS 2016 (11. Auflage).