Tatsächlich ist die Vereinigung aller E_g sogar gleich U, denn jeder Punkt $x \in U$ ist innerer Punkt, es gibt also ein $\varepsilon > 0$, so daß jeder Punkt $y \in \mathbb{R}^n$ mit $\|y - x\|_\infty < \varepsilon$ in U liegt. Wir wählen ein $g \in \mathbb{N}$ so, daß $2^g > 1/\varepsilon$ ist und setzen $k_i = [2^g x_i]$, wobei x_i die i-te Koordinate von x bezeichnet. Dann liegt x im Baustein $\left[\frac{k_1}{2^g}, \frac{k_1+1}{2^g}\right] \times \cdots \times \left[\frac{k_1}{2^g}, \frac{k_1+1}{2^g}\right]$, und da dessen Kantenlänge 2^{-g} kleiner ist als ε ist $\|y - x\|_\infty < \varepsilon$ für alle y aus diesem Baustein. Somit liegt dieser ganz in U, und $x \in E_g$.

Um U als Vereinigung volumenfremder Quader darzustellen, beginnen wir mit der Menge aller Bausteinen der ersten Generation, die ganz in U enthalten sind, und nehmen dann im g-ten Schritt, $g \geq 2$, alle Bausteine der Generation g dazu, die erstens ganz in U liegen und zweitens nicht in einem der bereits vorhandenen Quader enthalten sind. Nach dem g-ten Schritt ist die Vereinigung aller bis dahin betrachteter Quader gleich E_g; insgesamt erhalten wir also die Vereinigung aller E_g, das heißt die Menge U.

■

Mit diesem Lemma können wir nun das vorige in eine Form bringen, in der es deutlich einfacher angewendet werden kann:

Lemma: $f\colon D \to \mathbb{R}$ sei eine stetig differenzierbare Abbildung auf der offenen Menge $D \subseteq \mathbb{R}^n$, und $Z \subset D$ sei eine Nullmenge. Dann ist auch $f(Z)$ eine Nullmenge.

Beweis: Nach dem gerade bewiesenen Lemma läßt sich D als abzählbare Vereinigung kompakter Quader Q_k schreiben. Auf jedem dieser Quader erfüllt f eine LIPSCHITZ-Bedingung, denn für $x, y \in Q_k$ können wir die Funktion $g(t) = f\big(x + t(y - x)\big)$ betrachten, und nach dem Mittelwertsatz der Differentialrechnung ist

$$f(y) - f(x) = g(1) - g(0) = \frac{g(1) - g(0)}{1 - 0} = g'(\tau)$$

für ein $\tau \in [0,\ 1]$. Nach der Kettenregel ist

$$g'(\tau) = \nabla f\big(x + \tau(y - x)\big) \cdot (y - x)\,,$$

also insgesamt

$$\begin{aligned}\|f(y) - f(x)\|_\infty &= \big\|\nabla f\big(x + \tau(y - x)\big) \cdot (y - x)\big\|_\infty \\ &\leq \max\big\{\|\nabla f(z)\|_\infty \bigm| z \in Q\big\} \cdot \|y - x\|_\infty\ ,\end{aligned}$$

wobei das Maximum existiert, da wir f als *stetig* differenzierbar vorausgesetzt haben und den Gradienten auf einer kompakten Menge betrachten. Somit genügt die Einschränkung von f auf Q_k einer LIPSCHITZ-Bedingung mit diesem Maximum als Konstante, und damit ist $f(Q_k \cap Z)$ nach dem vorletzten Lemma eine Nullmenge. $f(Z)$ ist die Vereinigung dieser abzählbar vielen Nullmengen $f(Q_k \cap Z)$ und daher ebenfalls eine Nullmenge.

■

Korollar: $D \subset \mathbb{R}^m$ sei eine offene Menge und $f\colon D \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Ist $m < n$, so ist $f(Z)$ für jede Teilmenge $Z \subseteq D$ eine Nullmenge in $\mathbb{R}^n$.

Beweis: Wir betten $\mathbb{R}^m$ ein in $\mathbb{R}^n$, indem wir einfach die letzten $n - m$ Koordinaten auf Null setzen. Das Bild von $\mathbb{R}^m$ liegt dann insbesondere in der Hyperebene $x_n = 0$ von $\mathbb{R}^n$, ist also eine Nullmenge; erst recht ist die Einbettung $\widetilde{A}$ von A eine Nullmenge in $\mathbb{R}^n$. Wir definieren eine neue Abbildung $g\colon D \times \mathbb{R}^{n-m} \to \mathbb{R}^n$, die jedem Punkt $x = (x_1, \ldots, x_m)$ den Punkt $f(x_1, \ldots, x_m)$ zuordnet; die letzten $n - m$ Koordinaten werden also einfach ignoriert. Dann ist $f(A) = g(\widetilde{A})$ nach dem vorigen Lemma eine Nullmenge, wie behauptet.

■

Damit sind also Kurven in $\mathbb{R}^2$ sowie Kurven und Flächen in $\mathbb{R}^3$ Nullmengen, falls sie eine stetig differenzierbare Parametrisierung zulassen, genauso alle Vereinigungen endlich oder abzählbar vieler solcher Mengen. Entsprechendes gilt natürlich auch in höheren Dimensionen.

Für die Zwecke des noch zu definierenden LEBESGUE-Integrals wollen wir Nullmengen weitgehend ignorieren; um dies sprachlich kurz und einfach ausdrücken zu können, führen wir eine neue Sprechweise ein:

Definition: *a)* Zwei Funktionen $f, g\colon D \to \mathbb{R}$ auf $D \subseteq \mathbb{R}^n$ heißen *fast überall gleich*, wenn es eine Nullmenge $Z \subseteq D$ gibt, so daß $f(x) = g(x)$ für alle $x \in D \setminus Z$ ist.
b) Eine Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen $f_n\colon D \to \mathbb{R}$ *konvergiert fast überall punktweise* gegen die Funktion $f\colon D \to \mathbb{R}$, wenn es eine Nullmenge $Z \subset D$ gibt, so daß

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{für alle } x \in D \setminus Z\,.$$

c) Die Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen $f_n\colon D \to \mathbb{R}$ *konvergiert fast überall punktweise,* wenn es eine Funktion $f\colon D \to \mathbb{R}$ gibt, gegen die sie fast überall punktweise konvergiert.

Die Strategie zur Definition des LEBESGUE-Integrals ist nun einfach zu formulieren:

Definition: *a)* Eine Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen $f_n \in \mathrm{K}_1^0(\mathbb{R}^n, \mathbb{R})$ heißt *approximierend* für die Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$, wenn sie erstens eine CAUCHY-Folge ist und zweitens fast überall punktweise gegen f konvergiert.
b) Eine Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ heißt *integrierbar,* wenn es dazu eine approximierende Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen aus $\mathrm{K}_1^0(\mathbb{R}^n, \mathbb{R})$ gibt; in diesem Fall bezeichnen wir

$$\int_{\mathbb{R}^n} f \underset{\text{def}}{=} \lim_{k\to\infty} f_k$$

als das LEBESGUE-Integral über f.

HENRI LÉON LEBESGUE (1875–1941) studierte ab 1894 an der Ecole Normale Supérieure. Nach seinem ersten Abschluß 1897 studierte er noch zwei Jahre auf eigene Faust weiter, hauptsächlich in der Bibliothek, bis er 1899 eine Stelle als Gymnasiallehrer in Nancy erhielt. Dort schrieb er auch 1901 seine Arbeit über die Definition des LEBESGUE-Integrals. 1902 promovierte er mit einer weiteren Arbeit in Paris und bekam daraufhin eine Stelle an der Universität von Rennes. 1906 wechselte er nach Poitiers, 1910 an die Sorbonne in Paris. Ab 1921 war er Professor am Collège de France. Seine Arbeiten befassen sich mit Integrationstheorie, Potentialtheorie, Variationsrechnung, Topologie und anderen Gebieten.

Auch wenn es auf den ersten Blick so aussieht, haben wir damit noch nicht wirklich das LEBESGUE-Integral definiert: *Erstens* wissen wir noch nicht, ob der Grenzwert der Folge von Integralen existiert, und *zweitens* wissen wir noch nicht, daß alle approximierenden Folgen für f zum selben Grenzwert führen.

Das erste Problem ist leicht zu lösen: Wegen der Linearität und der Monotonieeigenschaft des Integrals für Funktionen mit kompaktem Träger

gilt für jede approximierende Folge $(f_k)_{k\in\mathbb{N}}$

$$\left|\int_{\mathbb{R}^n} f_\ell - \int_{\mathbb{R}^n} f_k\right| = \left|\int_{\mathbb{R}^n} (f_\ell - f_k)\right| \le \int_{\mathbb{R}^n} |f_\ell - f_k| = \|f_\ell - f_k\|_1 \ .$$

Da eine approximierende Folge nach Definition eine CAUCHY-Folge bezüglich der L^1-Norm ist, gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so daß die rechte Seite kleiner als ε ist für $k, \ell \ge N$. Damit gilt dasselbe erst recht für die linke Seite, d.h. die Folge der Integrale über die f_k ist eine CAUCHY-Folge reeller Zahlen und konvergiert somit nach dem CAUCHYschen Konvergenzkriterium.

Das zweite Problem erfordert mehr Arbeit. Sind $(f_k)_{k\in\mathbb{N}}$ und $(g_k)_{k\in\mathbb{N}}$ zwei approximierende Folgen für dieselbe Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$, so ist $(f_k - g_k)_{k\in\mathbb{N}}$ eine approximierende Folge für die Nullfunktion, denn nach der Dreiecksungleichung ist die Differenz zweier CAUCHY-Folgen wieder eine CAUCHY-Folge, und sind Z_1 *bzw.* Z_2 die Nullmengen, außerhalb derer die Folge der $f_k(x)$ *bzw.* $g_k(x)$ gegen $f(x)$ konvergiert, so ist auch $Z_1 \cup Z_2$ eine Nullmenge, und für $x \notin Z_1 \cup Z_2$ konvergiert die Folge der Funktionswerte $f_k(x) - g_k(x)$ gegen $f(x) - f(x) = 0$. Wenn wir zeigen können, daß für die Folge der $f_k - g_k$ die Folge der Integrale gegen Null konvergiert, zeigt uns die Linearität des Integrals, daß die Folge der Integrale über die f_k *bzw.* g_k gegen denselben Wert konvergieren.

Wir müssen also zeigen, daß für eine approximierende Folge $(f_k)_{k\in\mathbb{N}}$ zur Nullfunktion die Folge der $\int_{\mathbb{R}^n} f_k$ eine Nullfolge ist. Das Problem dabei ist, daß die Folge der Funktionswerte $f_k(x)$ *nicht* für jeden Punkt $x \in \mathbb{R}^n$ gegen Null konvergieren muß; wir müssen uns überlegen, daß diese Abweichungen für das Integral irrelevant sind.

Dazu verallgemeinern wir zunächst den Begriff der Nullmenge:

Definition: A sei eine Teilmenge von $\mathbb{R}^n$. Wir betrachten die Menge M_A aller reeller Zahlen s mit der Eigenschaft, daß A durch eine abzählbare Menge von Quadern Q_k überdeckt werden kann derart, daß $\sum_{k\ge1} \mu(Q_k) < s$ ist. Falls diese Menge leer ist, definieren wie das *äußere Maß* $\mu^*(A)$ als ∞; andernfalls ist $\mu^*(A)$ das Infimum der Menge M_A.

Offensichtlich ist $\mu^*(A)$ stets größer oder gleich Null, denn M_A kann keine negativen Zahlen enthalten. Ist $\mu^*(A) = 0$, so gibt es zu jedem $\varepsilon > 0$ eine Überdeckung durch Quader Q_k mit $\sum_{k\geq 1}\mu(Q_k) < \varepsilon$, die Menge A ist also eine Nullmenge; umgekehrt ist auch das äußere Maß einer jeden Nullmenge gleich Null. Das äußere Maß einer Teilmenge kann selbstverständlich nicht größer sein als das der größeren Menge, denn jede Überdeckung dieser Menge ist erst recht eine Überdeckung der Teilmenge. Außerdem ist klar, daß für zwei Teilmengen $A, B \subseteq \mathbb{R}^n$ stets $\mu^*(A \cup B) \leq \mu^*(A) + \mu^*(B)$ sein muß, wobei wir eine Summe, in der (mindestens) einer der Summanden ∞ ist als ∞ interpretieren. Nicht ganz so offensichtlich ist die Tatsache, daß dies auch für abzählbare Vereinigungen gilt:

$$\mu^*\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu^*(A_k)\,,$$

wobei die Summe gleich ∞ sein soll, wenn sie divergiert oder wenn mindestens einer der Summanden gleich ∞ ist. In diesen beiden Fällen ist die Ungleichung trivialerweise erfüllt; andernfalls beachten wir, daß wir für jede der Mengen A_k und jedes $\varepsilon > 0$ eine Überdeckung von A_k durch Quader finden können, so daß die Summe von deren Volumina kleiner ist als $\mu^*(A_k) + \varepsilon/2^k$. Nehmen wir alle diese Quader zusammen, so überdecken sie die Vereinigung aller A_k, und die Summe ihrer Volumina ist kleiner als

$$\sum_{k=1}^{\infty}\left(\mu^*(A_k) + \frac{\varepsilon}{2^k}\right) = \sum_{k=1}^{\infty}\mu^*(A_k) + \sum_{k=1}^{\infty}\frac{\varepsilon}{2^k} = \sum_{k=1}^{\infty}\mu^*(A_k) + \varepsilon\,.$$

Somit kann das Infimum über die Menge aller $s \in \mathbb{R}$ derart, daß die Vereinigung der A_k eine Überdeckung durch Quader mit Gesamtvolumen höchstens s hat, nicht kleiner sein als die erste Summe, und genau das war zu beweisen. ■

Als nächstes wollen wir uns überlegen, daß die L^1-Norm einer Funktion mit kompaktem Träger zwar keine obere Schranke für die Funktionswerte liefern kann, wohl aber eine obere Schranke für das äußere Maß jener Teilmenge von $\mathbb{R}^n$, auf der die Funktion Werte oberhalb einer gegebenen Schranke annimmt. Die entsprechende Ungleichung

von TSCHEBYSCHEFF gilt, wie wir bald sehen werden, auch noch unter deutlich schwächeren Voraussetzungen als im folgenden Lemma; sie ist auch ein wichtiges Hilfsmittel der Statistik.

PAFNUTI LWOWITSCH TSCHEBYSCHEFF (1821–1894) ist die in Deutschland übliche Transkription von Пафнуты Лвович Чебышев; im Englischen schreibt man heute meist PAFNUTY LVOVICH CHEBYSHEV. Er studierte Mathematik in Moskau und publizierte bereits während seines Studiums in deutschen und französischen Fachzeitschriften. 1847 bekam er eine Stelle an der Universität von St. Petersburg, wo er bis 1882 lehrte, ab 1860 als Professor. Seine Ergebnisse spielen noch heute eine große Rolle in der Wahrscheinlichkeitstheorie, der Zahlentheorie (insbesondere Primzahlverteilung), der Approximationstheorie und in anderen Gebieten der Mathematik.

Lemma: Die Funktion $f \in \mathrm{K}_1^0(\mathbb{R}^n, \mathbb{R})$ nehme keine negativen Werte an. Dann ist für jede reelle Zahl $c > 0$

$$\mu^*\Big(\big\{x \in \mathbb{R}^n \mid f(x) > c\big\}\Big) \leq \frac{\|f\|_1}{c}\,.$$

Beweis: Als Urbild der offenen Menge $\{x \in \mathbb{R} \mid x > c\}$ unter der stetigen Abbildung f ist $U = \{x \in \mathbb{R}^n \mid f(x) > c\}$ eine offene Teilmenge von $\mathbb{R}^n$; da sie im (kompakten) Träger von f liegt ist sie außerdem beschränkt; es gibt also einen Quader Q mit $U \subset Q$. Wie wir vor kurzem gesehen haben, läßt sich U als abzählbare Vereinigung von Quadern Q_k schreiben; nach den bekannten Eigenschaften des Quaderintegrals ist dann für jedes $m \in \mathbb{N}$

$$\|f\|_1 = \int_{\mathbb{R}^n} f = \int_Q f \geq \sum_{k=1}^{m}\int_{Q_k} f \geq \sum_{k=1}^{m} c\mu(Q_k) = c\sum_{k=1}^{m}\mu(Q_k)\,.$$

Für $m \to \infty$ folgt $\|f\|_1 \geq c\sum_{k=1}^{\infty}\mu(Q_k)$ und damit $\sum_{k=1}^{\infty}\mu(Q_k) \leq \dfrac{\|f\|_1}{c}$.
Die Menge U kann somit überdeckt werden durch Quader deren Gesamtvolumen höchstens gleich der rechten Seite ist; damit muß auch das äußere Maß von U kleiner oder gleich dieser Schranke sein. ■

Als erste Anwendung dieser Ungleichung können wir zeigen, daß eine CAUCHY-Folge bezüglich der L^1-Norm zumindest eine Teilfolge hat, die abgesehen von einer kleinen Teilmenge auch punktweise und sogar gleichmäßig konvergiert, genauer:

Lemma: Jede CAUCHY-Folge $(f_k)_{k\in\mathbb{N}}$ aus $K_1^0(\mathbb{R}^n, \mathbb{R})$ hat eine Teilfolge, die fast überall punktweise konvergiert. Außerdem gibt es für jedes $\varepsilon > 0$ eine Teilmenge $Z \subset \mathbb{R}^n$ mit $\mu^*(Z) < \varepsilon$, so daß die Teilfolge auf $\mathbb{R}^n \setminus Z$ sogar gleichmäßig konvergiert.

Beweis: Wir konstruieren die Folge der Indizes ν_k für die Teilfolge rekursiv: ν_1 sei so gewählt, daß $\left\|f_q - f_p\right\|_1 < \frac{1}{4}$ ist für alle $p, q \geq \nu_1$. Die folgenden ν_k werden so gewählt, daß

$$\left\|f_q - f_p\right\|_1 < 4^{-k} \text{ für alle } p, q \geq \nu_k \quad \text{und} \quad \nu_k > \nu_{k-1}$$

ist. Mit $g_k \underset{\text{def}}{=} f_{\nu_k}$ ist dann eine Teilfolge der gegebenen Folge definiert mit der Eigenschaft, daß

$$\|g_\ell - g_k\|_1 < 4^{-k} \quad \text{für alle } \ell \geq k\,.$$

Als Kandidat für eine Grenzfunktion betrachten wir die Funktion $g\colon \mathbb{R}^n \to \mathbb{R}$ mit

$$g(x) = g_1(x) + \sum_{k=1}^{\infty}\bigl(g_{k+1}(x) - g_k(x)\bigr)\,.$$

Ihre $(n-1)$-te Teilsumme ist

$$g_1(x) + \sum_{k=1}^{n-1}\bigl(g_{k+1}(x) - g_k(x)\bigr) = g_n(x)\,,$$

denn jedes g_k mit $k < n$ steht in dieser endlichen Summe einmal mit positivem und einmal mit negativem Vorzeichen. *Falls* die Reihe konvergiert, ist sie also in der Tat ein Kandidat für eine Grenzfunktion.

Hier kommt nun die Ungleichung von TSCHEBYSCHEFF ins Spiel: Danach ist das äußere Maß der Menge

$$Y_k = \left\{x \in \mathbb{R}^n \;\middle|\; |g_{k+1}(x) - g_k(x)| > 2^{-k}\right\}$$

höchstens gleich

$$\frac{\|g_{k+1} - g_k\|_1}{2^{-k}} < \frac{4^{-k}}{2^{-k}} = 2^{-k}\,.$$

Das äußere Maß der Vereinigung Z_k aller Y_ℓ mit $\ell \geq k$ ist daher kleiner als $2^{-k} + 2^{-k-1} + \cdots = 2^{1-k}$. Für $x \notin Z_k$ ist nach Definition dieser Menge $|g_{\ell+1}(x) - g_\ell(x)| < 2^{-\ell}$, also ist die geometrische Reihe $\sum 2^{-\ell}$ eine konvergente Majorante der definierenden Reihe von $g(x)$. Somit konvergiert diese Reihe auf $\mathbb{R}^n \setminus Z_k$ gleichmäßig. Da wir zu jedem $\varepsilon > 0$ ein $k \in \mathbb{N}$ finden können mit $2^{1-k} < \varepsilon$, ist damit die Behauptung über die gleichmäßige Konvergenz bewiesen.

Bleibt noch die punktweise Konvergenz außerhalb einer Nullmenge. Der Durchschnitt Z aller Z_k ist in jeder der Mengen Z_k enthalten, also ist $\mu^*(Z) < 2^{1-k}$ für alle $k \in \mathbb{N}$ und Z somit eine Nullmenge. Für $x \notin Z$ gibt es Mengen Z_k, in denen x nicht enthalten ist, und wie wir gerade gesehen haben, konvergiert die Reihe in jeder der Mengen $\mathbb{R}^n \setminus Z_k$. Somit konvergiert sie für alle $x \in \mathbb{R}^n \setminus Z$, also fast überall. ■

Damit haben wir alles zusammen, was wir zum Nachweis der Wohldefiniertheit des LEBESGUE-Integrals brauchen; wie wir uns bereits überlegt haben, folgt diese aus dem folgenden

Lemma: $(f_k)_{k\in\mathbb{N}}$ sei eine CAUCHY-Folge aus $K_1^0(\mathbb{R}^n, \mathbb{R})$, und die Folge $\bigl(f_k(x)\bigr)_{k\in\mathbb{N}}$ konvergiere für fast alle $x \in \mathbb{R}^n$ gegen Null. Dann konvergiert auch die Folge der L^1-Normen

$$\|f_k\|_1 = \int_{\mathbb{R}^n} f_k$$

gegen Null.

Beweis: Wir betrachten die Teilfolge $(g_k)_{k\in\mathbb{N}}$ aus dem vorigen Lemma. Wenn wir zeigen können, daß die Teilfolge der Zahlen $\|g_k\|_1$ gegen Null konvergiert, folgt dasselbe auch für die Folge der $\|f_k\|_1$, denn zu jedem $\varepsilon > 1$ gibt es, da $(f_k)_{k\in\mathbb{N}}$ eine CAUCHY-Folge ist, ein $N \in \mathbb{N}$, so daß $\|f_\ell - f_k\|_1 < \frac{1}{2}\varepsilon$ ist für alle $k, \ell > N$. Dies gilt insbesondere auch für solche Indizes $\ell = \nu_p$, für die $f_\ell = g_p$ in der Teilfolge liegt. Für die Teilfolge wiederum gibt es ein $M \in \mathbb{N}$, so daß $\left\|g_p\right\|_1 < \frac{1}{2}\varepsilon$ ist für alle $p \geq M$. Wählen wir N so, daß $\nu_N \geq M$ ist, folgt

$$\|f_k\|_1 = \|f_k + (f_\ell - f_k)\|_1 \leq \|f_k\|_1 + \|f_\ell - f_k\|_1 < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

Somit konvergiert auch die Folge der Normen der f_k gegen Null.

Um zu zeigen, daß die Folge der $\|g_k\|_1$ eine Nullfolge ist, betrachten wir ein $\varepsilon > 0$ und einen Quader Q, der den Träger von g_k enthält. Durch eine langwierige Abschätzung werden wir gleich zeigen, daß dann gilt

$$\|g_k\|_1 \le \frac{1}{3 \cdot 4^{k-1}} + \varepsilon\big(\mu(Q) + \|g_k\|_\infty\big)\,.$$

Wenn wir dies für alle $\varepsilon > 0$ gezeigt haben, können wir ε gegen Null gehen lassen und erhalten die Abschätzung

$$\|g_k\|_1 \le \frac{1}{3 \cdot 4^{k-1}}\,,$$

aus der natürlich sofort folgt, das die Normen eine Nullfolge bilden.

Bleibt also noch die obige Ungleichung zu zeigen. Wir definieren für alle Paare (k, ℓ) von natürlichen Zahlen $\ell \ge k$ Hilfsfunktionen $\gamma_{k\ell}\colon \mathbb{R}^n \to \mathbb{R}$ durch

$$\gamma_{k\ell}(x) \underset{\text{def}}{=} \sum_{i=k}^{\ell} |g_{i+1}(x) - g_i(x)|\ ;$$

da die g_i als Elemente von $\mathrm{K}_1^0(\mathbb{R}^n, \mathbb{R})$ stetig sind, sind auch diese Funktionen stetig.

In einem ersten Schritt wollen wir $|g_k(x)|$ durch die gerade definierten Funktionen abschätzen. Dabei unterscheiden wir zwei Fälle:

1. Fall: x liegt nicht in der Teilmenge Z aus dem vorigen Beweis und $\lim\limits_{k\to\infty} g_k(x) = 0$.

Dann ist

$$|g_k(x)| = \left|\lim_{\ell\to\infty}\big(g_{\ell+1}(x) - g_k(x)\big)\right| = \left|\lim_{\ell\to\infty}\sum_{i=k}^{\ell}\big(g_{i+1}(x) - g_k(x)\big)\right|$$
$$\le \lim_{\ell\to\infty}\sum_{i=k}^{\ell}|g_{i+1}(x) - g_i(x)| \le \lim_{\ell\to\infty}\gamma_{k\ell}(x)\,;$$

der letzte Grenzwert existiert, wie wir im Beweis des vorigen Lemmas gesehen haben, da er durch eine geometrische Reihe majorisiert werden kann.

Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so daß sich $\gamma_{k\ell}(x)$ für $\ell \ge N$ um weniger als ε von diesem Grenzwert unterscheidet; für solche ℓ ist dann auch $|g_k(x)| < \gamma_{k\ell}(x) + \varepsilon$. Da beide Seiten dieser Ungleichung stetige Funktionen von x sind, gilt diese Ungleichung auch noch in einer gewissen Umgebung von x; es gibt also insbesondere einen offenen Quader Q_x, der x enthält, so daß $|g_k(y)| < \gamma_{k\ell}(y) + \varepsilon$ für alle $y \in Q_x$.

2. Fall: $x \in Z$ oder $i \lim\limits_{k\to\infty} g_k(x)$ existiert nicht oder existiert und ist von Null verschieden.

Wie wir wissen, ist Z eine Nullmenge, und nach Voraussetzung ist auch die Menge aller $x \in \mathbb{R}^n$, für die nicht $\lim\limits_{k\to\infty} g_k(x) = 0$ gilt, eine Nullmenge; somit ist auch die Vereinigung der beiden Mengen eine Nullmenge Z'. Daher gibt es für jedes $\varepsilon > 0$ eine abzählbare Menge von Quadern Q_j derart, daß Z' in der Vereinigung der Q_j liegt und die Summe von deren Volumina kleiner ist als ε.

Wir betrachten nun die Menge $\mathfrak{U}$, die aus den sämtlichen Quadern Q_x aus dem ersten Fall sowie den sämtlichen Quadern Q_j aus dem zweiten Fall besteht. Sie ist eine offene Überdeckung des kompakten Quaders Q, hat also eine endliche Teilüberdeckung. Diese besteht aus endlich vielen der Q_x – dies seien die Quader $P_1, \ldots, P_r$ – sowie aus endlich vielen der Q_j; dies seien die Quader $R_1, \ldots, R_s$.

Für jeden der Quader P_i gibt es ein M_i, so daß $|g_k(x)| < \gamma_{k\ell}(x)$ ist für alle $\ell \ge M_i$; bezeichnet M das Maximum dieser M_i und P die Vereinigung der P_i, ist also $|g_k(x)| < \gamma_{k\ell}(x)$ ist für alle $x \in P$ und alle $\ell \ge M$. Indem wir die P_i nötigenfalls noch unterteilen, mit Q schneiden und abschließen, können wir annehmen, daß $P \cap Q$ die volumenfremde Vereinigung von Quadern P_i ist. Dann ist

$$\int_Q \gamma_\ell = \int_Q \sum_{i=k}^{\ell}\int_Q |g_{i+1} - g_i| = \sum_{i=k}^{\ell}\int_Q |g_{i+1} - g_o|$$
$$= \sum_{i=k}^{\ell}\int_{R^n} |g_{i+1} - g_o| = \sum_{i=k}^{\ell}\|g_{k+1} - g_k\| < \sum_{i=k}^{\ell} 4^{-i}$$
$$< \sum_{i=k}^{\infty} 4^{-i} = \frac{1}{3 \cdot 4^{k-1}}\,.$$

Ersetzen wir nun auch noch die R_j durch die entsprechenden abgeschlossenen Quader, erhalten wir für $\ell \geq M$ die Abschätzung

$$\begin{aligned}\|g_k\|_1 = \int_{\mathbb{R}^n} |g_k| = \int_Q |g_k| &\leq \sum_{i=1}^r \int_{P_i} |g_k| + \sum_{j=1}^s \int_{R_j} |g_k| \\ &\leq \sum_{i=1}^r \int_{R_i} (\gamma_{k\ell} + \varepsilon) + \sum_{j=1}^s \|g_k\|_\infty \, \mu(R_j) \\ &\leq \sum_{i=1}^r \int_{R_i} (\gamma_{k\ell} + \varepsilon) + \|g_k\|_\infty \sum_{j=1}^s \mu(R_j) \\ &< \int_Q \gamma_{k\ell} + \varepsilon\big(\mu(Q) + \|g_k\|_\infty\big) \\ &\leq \lim_{\ell\to\infty} \int_Q \gamma_{k\ell} + \varepsilon\big(\mu(Q) + \|g_k\|_\infty\big) \\ &\leq \frac{1}{3 \cdot 4^{k-1}} + \varepsilon\big(\mu(Q) + \|g_k\|_\infty\big) \,,\end{aligned}$$

wie behauptet. Damit ist das Lemma und die Wohldefiniertheit des LEBESGUE-Integrals bewiesen.

■

Wir bezeichnen die Menge aller LEBESGUE-integrierbarer Funktionen auf $\mathbb{R}^n$ mit $\mathrm{Leb}_1(\mathbb{R}^n, \mathbb{R})$, wobei der Index 1 bedeuten soll, daß wir auch für $f \in \mathrm{Leb}_1(\mathbb{R}^n, \mathbb{R})$ die Schreibweise

$$\|f\|_1 \underset{\text{def}}{=} \int_{\mathbb{R}^n} f$$

verwenden wollen. Es ist klar, daß $\mathrm{Leb}_1(\mathbb{R}^n, \mathbb{R})$ ein Vektorraum ist, aber $\|\cdot\|_1$ definiert keine Norm auf $\mathrm{Leb}_1(\mathbb{R}^n, \mathbb{R})$: Schließlich ist $\|f\|_1 = 0$ nicht nur für die Nullfunktion, sondern auch für jede Funktion, die nur *fast überall* verschwindet.

Es gibt zwei Ansätze, um mit diesem Problem fertig zu werden: Zum ersten könnten wir den Faktorraum $\mathrm{L}^1(\mathbb{R}^n, \mathbb{R})$ des Vektorraums $\mathrm{Leb}_1(\mathbb{R}^n, \mathbb{R})$ nach dem Untervektorraum der fast überall verschwindenden Funktionen betrachten; nach dem gerade bewiesenen Lemma ist das ein normierter Vektorraum. Elemente von $\mathrm{L}^1(\mathbb{R}^n, \mathbb{R})$ sind dann allerdings nicht mehr Funktionen, sondern Äquivalenzklassen von Funktionen, wobei zwei Funktionen genau dann zur gleichen Äquivalenzklasse gehören, wenn sie fast überall gleich sind. Dieser Ansatz wird oft in der Funktionalanalysis verwendet.

Alternativ können wir auch den Begriff der Norm etwas abschwächen:

Definition: Eine *Halbnorm* auf einem $\mathbb{R}$- oder $\mathbb{C}$-Vektorraum V ist eine Abbildung $\|\cdot\| : V \to \mathbb{R}$ mit den Eigenschaften

a) $\|\lambda v\| = |\lambda|\,\|v\|$ für alle $\lambda \in \mathbb{R}$ *bzw.* $\mathbb{C}$ und $v \in V$

b) $\|v\| \geq 0$ für alle $v \in V$

c) $\|v + w\| \leq \|v\| + \|w\|$ *(Dreiecksungleichung)*

Wir verlangen also alle Eigenschaften einer Norm außer der Forderung, daß sie nur für den Nullvektor verschwinden darf. Offensichtlich ist $\|\cdot\|_1$ eine Halbnorm auf $\mathrm{Leb}_1(\mathbb{R}^n, \mathbb{R})$, und mit dieser wollen wir im folgenden rechnen.

Da jede Funktion aus $\mathrm{Leb}_1(\mathbb{R}^n, \mathbb{R})$ Grenzwert einer approximierenden Folge von Funktionen mit kompaktem Träger ist, können wir die bereits für das Quaderintegral und damit auch für Integrale über Funktionen mit kompaktem Träger bewiesenen Eigenschaften des Integrals ohne großen Aufwand auf das LEBESGUE-Integral übertragen: Insbesondere ist das definierte Integral *linear,* d.h.

$$\int_{\mathbb{R}^n} (f + g) = \int_{\mathbb{R}^n} f + \int_{\mathbb{R}^n} g \,,$$

wir haben die Monotonieeigenschaft

$$\int_{\mathbb{R}^n} f \leq \int_{\mathbb{R}^n} g \quad \text{falls } f(x) \leq g(x) \text{ für alle } x \in \mathbb{R}^n \,,$$

und nach Definition ist auch

$$\int_{\mathbb{R}^n} f = \int_{\mathbb{R}^n} g \quad \text{falls } f \text{ und } g \text{ fast überall gleich sind} \,,$$

(Damit folgt natürlich auch, daß wir bei der Monotonieregel nur fordern müssen, daß $f(x) \leq g(x)$ fast überall gilt, d.h. die eventuellen Ausnahmepunkte bilden eine Nullmenge.)

Um zu sehen, daß das mit so großem Aufwand definierte Integral auch zu etwas nütze ist, wollen wir als nächstes einige Beispiele von LEBESGUE-integrierbaren Funktionen betrachten.

Selbstverständlich existiert das LEBESGUE-Integral für die DIRICHLETsche Sprungfunktion

$$f\colon \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \\ 0 & \text{sonst} \end{cases} \end{cases},$$

denn wegen der Abzählbarkeit von $\mathbb{Q}$ ist diese Funktion fast überall gleich der Nullfunktion, also ist $\int_{R^n} f = 0$, wie erwartet.

Allgemeiner können wir für eine beliebige Teilmenge $A \subseteq \mathbb{R}^n$ die sogenannte *charakteristische Funktion*

$$\chi_A\colon \begin{cases} \mathbb{R}^n \to \mathbb{R} \\ x \mapsto \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{sonst} \end{cases} \end{cases}$$

definieren und fragen, wann diese integrierbar ist.

Beginnen wir mit einen Intervall $A = (a, b) \subset \mathbb{R}$. Dann ist

$$\chi_A(x) = \begin{cases} 1 & \text{falls } a < x < b \\ 0 & \text{sonst} \end{cases}.$$

Völlig analog zum Fall $A = (0, 2)$, anhand dessen wir uns überlegt hatten, daß $K_1^0(\mathbb{R}, \mathbb{R})$ nicht vollständig ist, zeigt man auch hier, daß $\chi_A(x)$ Grenzwert der Folge der Funktionen

$$f_k(x) = \begin{cases} 0 & \text{falls } x \notin (a, b) \\ kx & \text{falls } a \le x \le a + \frac{1}{k} \\ 1 & \text{falls } a + \frac{1}{k} \le x \le b - \frac{1}{k} \\ k(b - x) & \text{falls } b - \frac{1}{k} \le x \le b \end{cases}$$

aus $K_1^0(\mathbb{R}, \mathbb{R})$ ist, wobei wir allerdings nur Werte von k betrachten dürfen, für die $b - a > 2/k$ ist. Diese Folge konvergiert auch punktweise gegen χ_A, ist also eine approximierende Folge.

Nach Definition des Quaderintegrals, das hier einfach das RIEMANN-Integral aus Kapitel 4 ist, haben wir

$$\int_{\mathbb{R}} f_k = \int_a^{a+1/k} kx\,dx + \int_{a+1/k}^{b-1/k} dx + \int_{b-1/k}^{b} k(b - x)\,dx$$
$$= \frac{1}{2k} + \left(b - a - \frac{2}{k}\right) + \frac{1}{2k} = b - a - \frac{1}{k},$$

was für $k \to \infty$ gegen $b - a$ konvergiert. Somit ist

$$\int_{\mathbb{R}} \chi_A = (b - a),$$

was wohl niemanden überraschen dürfte.

Ersetzen wir A durch das abgeschlossene Intervall $[a, b]$ oder eines der beiden halboffenen Intervalle $(a, b]$ oder $[a, b)$, ändert sich daran nichts, denn die charakteristischen Funktionen sind fast überall gleich: Höchstens an den Stellen $x + a$ und $x = b$ kann es Abweichungen geben.

Ist $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$ ein Quader in $\mathbb{R}^n$, so ist im Punkt $x = (x_1, \ldots, x_n)$

$$\chi_Q(x) = \chi_{[a_1, b_1]}(x_1) \cdot \cdots \cdot \chi_{[a_n, b_n]}(x_n),$$

und auch diese Funktion können wir approximieren durch stetige Funktionen mit kompaktem Träger, indem wir einfach als k-te Funktion das Produkt der $f_k(x_i)$ nehmen, wobei wir natürlich in der i-ten Komponente die Funktion f_k für das Intervall $[a_i, b_i]$ nehmen müssen. Nach der rekursiven Definition des Quaderintegrals folgt dann schnell, daß auch χ_Q LEBESGUE-integrierbar ist und daß wir auch hier das zu erwartete Ergebnis

$$\int_Q \chi_Q = \prod_{i=1}^n (b_i - a_i) = \mu(Q)$$

bekommen. Wegen der Linearität der Integration gilt etwas allgemeiner

$$\int_{\mathbb{R}^n} c\chi_Q = c \cdot \mu(Q).$$

Zumindest einfache unstetige Funktionen lassen sich also integrieren.

Die Ungleichung von TSCHEBYSCHEFF kennen wir bislang nur für stetige Funktionen mit kompaktem Träger; wir wollen sie verallgemeinern auf beliebige LEBESGUE-integrierbare Funktionen:

Lemma: Die Funktion $f \in \mathrm{Leb}_1(\mathbb{R}^n, \mathbb{R})$ nehme keine negativen Werte an. Dann ist für jede reelle Zahl $c > 0$

$$\mu^*\Big(\big\{x \in \mathbb{R}^n \bigm| f(x) > c\big\}\Big) \le \frac{\|f\|_1}{c}.$$