SEMINARARBEIT:
HAUPTSATZ DER DIFFERENTIAL- UND INTEGRALRECHNUNG

MATTHIAS HEINLEIN

1. Einleitung

Oftmals wird das Integral in den Anfängervorlesungen auf zweierlei Weisen eingeführt. Da ist zum einen das formale Integrieren, also das Auffinden einer Stammfunktion und auf der anderen Seite das Berechnen der Fläche zwischen dem Graphen einer Funktion und der x-Achse in einem bestimmten Bereich. Auf den ersten Blick haben diese beiden Dinge nichts miteinander zu tun, erst der Hauptsatz der Differential- und Integralrechnung bringt die beiden Begriffe zusammen.

Man kennt ihn vom Riemann-Integral in zwei Fassungen: (1) Die Integralfunktion F einer stetigen Funktion f ist eine Stammfunktion zu f. (2) Wenn F eine Stammfunktion zu f ist, lässt sich das Integral über f von a bis b als $F(b) - F(a)$ ausdrücken.

Der Hauptsatz fordert im Teil (2) also die Differenzierbarkeit von F in allen Punkten. In der Maßtheorie weiß man aber über eine Funktion oft nur an *fast allen* Punkten bescheid. Dass sich der Hauptsatz nicht einfach auf nur fast-überall differenzierbare Funktionen übertragen lässt, zeigt das Beispiel der Cantorfunktion. Um dennoch einen Hauptsatz für das Lebesgue-Integral formulieren und beweisen zu können (Abschnitt 4), benötigen wir eine zusätzliche Eigenschaft für Funktionen: die absolute Stetigkeit (Abschnitt 3).

2. Grundlagen

Wir benötigen eine Reihe von Grundlagen, die entweder aus der Vorlesung Maßtheorie bekannt sind, in anderen Arbeiten dieses Seminars vorkommen oder leicht aus diesen hergeleitet werden können. Außer der zweiten Aussage sind sie nicht expliziter Bestandteil dieser Arbeit und werden deshalb nicht bewiesen.

(1) Seien im Folgenden λ das Lebesgue-Maß über der Borel-σ-Algebra E über der Grundmenge $\Omega \subset \mathbb{R}$ und η das äußere Lebesgue-Maß (siehe Arbeit von Marcus Heitel). Wenn kein explizites Maß angegeben ist, bedeutet der Ausdruck „fast überall" oder „f.ü." immer „λ-fast überall".

(2) Wenn f integrierbar ist, gibt es für alle $\varepsilon > 0$ ein $\delta > 0$, sodass für alle messbaren Mengen A mit $\lambda(A) < \delta$ gilt: $\int_A |f|\, d\lambda < \varepsilon$. (Elstrodt, Aufgabe 3.7. in Kap. 4)
Hier wollen wir einen kurzen Beweis geben:

Beweis. Falls f beschränkt ist durch $c > 0$, so gilt $\int_A |f| d\lambda \le \int_A c\, d\lambda = c \cdot \lambda(A) \le c \cdot \delta$. D.h. mit $\delta < \frac{\varepsilon}{c}$ folgt die Behauptung.
Sei f nun nicht zwingend beschränkt, dann gilt $\mathbb{1}_{\{|f|>n\}} \cdot |f| \longrightarrow 0 \quad (n \to \infty)$ und mit Lebesgue

$$\int\limits_{\{|f|>n\}} |f| d\lambda \longrightarrow 0 \quad (n \to \infty)$$

Also findet man ein $c \in \mathbb{N}$, sodass

$$\int\limits_{\{|f|>c\}} |f| d\lambda < \frac{\varepsilon}{2}$$

Setzen wir nun $f_c = \mathbb{1}_{\{|f|\leq c\}} \cdot f$. Dann ist f_c beschränkt und man findet nach dem eben bewiesenen Teil ein $\delta > 0$, sodass

$$\int\limits_A |f_c| d\lambda < \frac{\varepsilon}{2} \qquad \forall A \text{ mit } \lambda(A) < \delta$$

Damit gilt insgesamt für alle Mengen A mit $\lambda(A) < \delta$:

$$\int\limits_A |f| d\lambda = \int\limits_{A\cap\{|f|>c\}} |f| d\lambda + \int\limits_{A\cap\{|f|\leq c\}} |f| d\lambda < \int\limits_{\{|f|>c\}} |f| d\lambda + \int\limits_A |f_c| d\lambda < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

□

(3) Seien ein Intervall I mit $a \in I$ gegeben. Falls $\int_a^c f\, d\lambda = 0$ für alle $c \in I$ gilt, so folgt, dass f f.ü. 0 ist in I. (Elstrodt, Aufgabe 5.8. in Kap. 4)

(4) Falls die Totalvariation

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

endlich ist, heißt f *von beschränkter Variation (kurz BV)* Jede BV-Funktion lässt sich als Differenz zweier monotoner Funktionen darstellen (Minimaldarstellung). (Elstrodt, Kap. 1.10)

(5) Jede monoton wachsende Funktion f ist f.ü. differenzierbar und es gilt

$$\int_a^b f'(x) dx \leq f(b) - f(a)$$

Damit ist jede Funktion von beschränkter Variation fast überall differenzierbar. Details siehe Arbeit von Johannes Wiesel.

(6) Aus der Arbeit von Johannes Wiesel wissen wir: Eine Familie $\mathcal{F}$ von nicht-entarteten Intervallen heißt Vitali-Überdeckung einer Menge $A \subset \mathbb{R}$, falls es zu jedem $x \in A$ und $\varepsilon > 0$ ein Intervall $I \in \mathcal{F}$ mit $\lambda(I) < \varepsilon$, sodass $x \in I$.
Der Satz von Vitali besagt, dass es bei jeder Vitali-Überdeckung F einer Menge $A \subset \mathbb{R}$ und für beliebiges $\varepsilon > 0$ endlich viele disjunkte Intervall $I_1, ..., I_n \in F$ gibt, sodass

$$\eta(A \setminus \bigcup_{k=1}^{n} I_k) < \varepsilon\,.$$

.

3. Absolut stetige Funktionen

3.1. **Die Cantorfunktion.** Die Cantorfunktion $c : [0,1] \to [0,1]$ kann anschaulich konstruiert werden, wie in Abb. 3.1 gezeigt. Für eine formale Konstruktion verweisen wir auf Elstrodt, Kapitel II, Beispiel 8.7 oder auch die Wikipedia unter dem Stichwort „Cantor-Verteilung“. Man kann zeigen, dass c stetig und, da es auf einem kompakten Intervall definiert ist, sogar gleichmäßig stetig ist. Außerdem ist c stückweise konstant. Auf diesen Bereichen, in denen c konstant ist, ist es differenzierbar mit Ableitung 0. Die Menge aller Punkte, in denen c nicht differenzierbar ist mit Ableitung 0, bildet die sog. Cantormenge, eine (überabzählbare) Nullmenge. Also ist c f.ü. differenzierbar mit Ableitung 0.
Offensichtlich erfüllt c nicht die Aussage des Hauptsatzes für das Riemann-Integral, Teil 2, denn $\int_0^1 c' d\lambda = 0 \neq c(1) - c(0) = 1$. Das zeigt uns, dass der Hauptsatz in der bisherigen Form tatsächlich nicht auf nur fast überall differenzierbare Funktionen übertragbar ist. Um eine Version des Hauptsatzes auch für solche Funktionen aufstellen und beweisen zu können, führen wir im Folgenden den wichtigen Begriff der absoluten Stetigkeit ein.

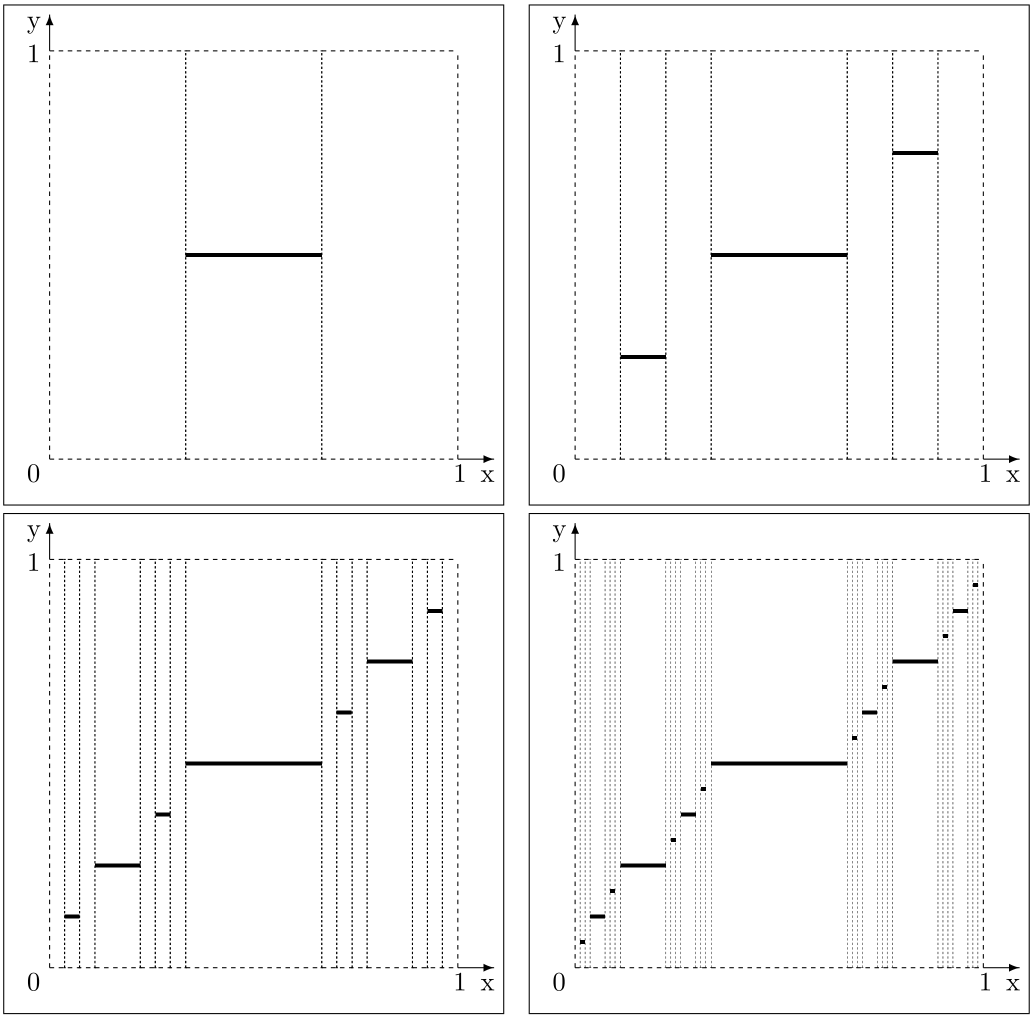

ABBILDUNG 1. Entwicklung der Cantorschen Treppenfunktion

3.2. Absolut stetige Funktionen.

Definition 3.1. $F : [a, b] \to \mathbb{R}$ heißt absolut stetig, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, sodass für alle $n \in \mathbb{N}$ und $a_i, b_i \in [a, b]$ $(i = 1, ..., n)$ mit $a \leq a_1 < b_1 \leq ... < b_n \leq b$ gilt: Wenn

$$\sum_{i=1}^{n}(b_i - a_i) < \delta$$

gilt, dann folgt

$$\sum_{i=1}^{n}|F(b_i) - F(a_i)| < \varepsilon\,.$$

Bemerkung 3.2.

(i) Jede absolut stetige Funktion ist gleichmäßig stetig und damit stetig. Das folgt aus der Definition mit $n = 1$.

(ii) Die Cantorsche Treppenfunktion ist nicht absolut stetig. Der Beweis wird dem geneigten Leser als Übungsaufgabe überlassen.

(iii) Jede absolut stetige Funktion ist von beschränkter Variation und damit f.ü. differenzierbar (Beweis siehe Elstrodt, Folgerung 4.12).

Satz 3.3. *Eine absolut stetige Funktion F, mit $F' = 0$ f.ü., ist konstant.*

Beweis. Wir zeigen, dass $F(a) = F(c)$ für beliebiges $c \in (a, b)$, indem wir $|F(a) - F(c)|$ durch einen Term in ε abschätzen. Mit $\varepsilon \to 0$ folgt die Konstantheit von F.

Seien also $c \in (a, b)$ und $\varepsilon > 0$ beliebig. Wähle das zu ε gehörige $\delta > 0$ aus der Definition der absoluten Stetigkeit (Def. 3.1). Wir fassen alle Punkte, an denen F differenzierbar ist mit Ableitung 0, in einer Menge zusammen:

$$A = \{x \in [a, c) : F'(x) = 0\}$$

Da $F' = 0$ f.ü., ist $N = [a, c) \setminus A$ eine Nullmenge. In jedem Punkt $x \in A$ ist F' ja gleich 0, d.h. betrachtet man Differenzenquotienten $\frac{1}{h} \cdot (F(x + h) - F(x))$ in einer Umgebung von x, so werden diese Differenzenquotienten klein. Es gibt für jedes $x \in A$ ein $H = H(x) > 0$, sodass gilt:

$$\left|\frac{F(x+h) - F(x)}{h}\right| < \varepsilon \quad \text{und damit} \quad |F(x + h) - F(x)| < \varepsilon \cdot h \quad \forall\, 0 < h < H(x)$$

Nun fasst man alle Intervalle $[x, x + h]$ mit $x \in A$ und h klein genug in eine Familie $\mathcal{E}$ zusammen. Da es zu jedem Punkt $x \in A$ ein beliebig kleines Intervall $[x, x + h] \in \mathcal{E}$ gibt, in dem x liegt, handelt es sich bei $\mathcal{E}$ also um eine Vitali-Überdeckung von A. Somit gibt es nach dem Überdeckungssatz von Vitali endlich viele Intervalle $[x_1, y_1], ..., [x_n, y_n] \in \mathcal{E}$, wobei $y_i = x_i + h_i$ für $i = 1, ..., n$, die A bis auf einen kleinen Rest überdecken, genauer

$$\eta(A \setminus \bigcup_{k=1}^{n} [x_k, y_k]) < \delta\,.$$

Definieren wir noch $y_0 := a$ und $x_{n+1} := c$, lässt sich der Rest, der nicht überdeckt wird, wie folgt als „Lücken“ zwischen den Intervallen beschreiben:

$$\delta > \eta(A \setminus \bigcup_{k=1}^{n} [x_k, y_k]) = \eta([a, c) \setminus \bigcup_{k=1}^{n} [x_k, y_k]) = \eta(\bigcup_{k=1}^{n} [x_{k+1}, y_k]) = \sum_{k=1}^{n} (x_{k+1} - y_k)$$

Bei der ersten Gleichung wurde nur die Nullmenge $N = [a, c) \setminus A$ hinzugefügt, die jedoch nichts am äußeren Maß η ändert.
Nun folgt aber aus $\sum_{k=1}^{n} (x_{k+1} - y_k) < \delta$ wegen der absoluten Stetigkeit von F, dass auch

$$\sum_{k=1}^{n} |F(x_{k+1}) - F(y_k)| < \varepsilon. \tag{3.1}$$

Nach Definition der $[x_i, y_i] = [x_i, x_i + h_i]$ ist

$$|F(y_i) - F(x_i)| < \varepsilon \cdot |y_i - x_i| \quad \forall i = 1, ..., n\,. \tag{3.2}$$

Mit diesen beiden Abschätzungen und dem folgenden Trick können wir nun den Beweis abschließen:

$$\begin{aligned}
|F(c) - F(a)| &= |F(x_{n+1}) - F(y_0)| \\
&= \left| \left(\sum_{k=1}^{n+1} F(x_k) - \sum_{k=1}^{n} F(x_k) \right) - \left(\sum_{k=0}^{n} F(y_k) - \sum_{k=1}^{n} F(y_k) \right) \right| \\
&= \left| \sum_{k=0}^{n} [F(x_{k+1}) - F(y_k)] + \sum_{k=1}^{n} [F(y_k) - F(x_k)] \right| \\
&\leq \sum_{k=0}^{n} |F(x_{k+1}) - F(y_k)| + \sum_{k=1}^{n} |F(y_k) - F(x_k)| \\
&\overset{(3.1),(3.2)}{<} \varepsilon + \sum_{k=1}^{n} \varepsilon \cdot |y_k - x_k| < \varepsilon \cdot (c - a + 1)
\end{aligned}$$

Da man ε beliebig klein wählen kann, folgt $|F(c) - F(a)| = 0$ für alle $c \in (a, b)$. □

Bemerkung 3.4. Die Cantorsche Funktion ist ebenfalls fast überall differenzierbar mit Ableitung 0, aber nicht konstant, da sie nicht absolut stetig ist. Man könnte absolute Stetigkeit etwas schwammig auch wie folgt beschreiben:

> *Bei einer absolut stetigen Funktion können wir aus ihrem Verhalten f.ü. schließen, dass sie auch auf der restlichen Nullmenge keine allzu großen Überraschungen bereithält.*

Nun sind wir soweit, den Hauptsatz für das Lebesgue-Integral formulieren und beweisen zu können.

4. Hauptsatz der Differential- und Integralrechnung

Satz 4.1 (HDI)**.**

i) Sei $f : [a,b] \to \mathbb{R}$ Lebesgue-integrierbar und $F : [a,b] \to \mathbb{R}$ definiert über

$$F(x) := \int_a^x f(t)dt, \quad (a \leq x \leq b)$$

Dann ist F absolut stetig und es gilt f.ü. $F' = f$.

ii) Wenn $F : [a,b] \to K$ absolut stetig ist und $F'(x) := 0$ an allen Stellen gesetzt wird, an denen F nicht differenzierbar ist, so ist F' Lebesgue-integrierbar und

$$\int_a^x F'(t)dt = F(x) - F(a)$$

Beweis. Zu (i): Wir beweisen die Aussage in den folgenden Schritten:

- f ist absolut stetig und f.ü. differenzierbar.
- Beweis für beschränkte Funktionen.
- Approximation unbeschränkter Funktionen durch beschränkte.

Sei $F : [a,b] \to \mathbb{R}$ mit $F(x) = \int_a^x f(t)dt$ gegeben. Daraus folgt, dass $F(y) - F(x) = \int_x^y f(t)dt$ für alle $x, y \in [a,b]$ gilt.
Grundlage (2) besagt, dass es für alle $\varepsilon > 0$ ein $\delta > 0$ gibt, sodass für alle Mengen A mit $\lambda(A) < \delta$ dann auch $\int_A |f| d\lambda < \varepsilon$ gilt. Wählt man A spezieller als

$$A = \bigcup_{k=1}^n [a_k, b_k]$$

mit beliebigen Punkten $a_1 < b_1 \leq a_2 < ... < b_n$ (d.h. die Intervalle sind disjunkt), sodass $\lambda(A) = \sum_{k=1}^n (b_k - a_k) < \delta$ ist, dann gilt also

$$\sum_{k=1}^n |F(b_k) - F(a_k)| \overset{\text{Def}}{=} \sum_{k=1}^n \left| \int_{a_k}^{b_k} f(t)dt \right| \leq \sum_{k=1}^n \int_{a_k}^{b_k} |f(t)|dt = \int_A |f| d\lambda < \varepsilon.$$

Das ist genau die Definition der absoluten Stetigkeit, F ist also absolut stetig und damit nach Bemerkung 3.2 (iii) auch f.ü. differenzierbar.

Betrachten wir nun nur *beschränkte* Funktionen, d.h. es gibt ein $M > 0$ mit $|f(x)| \leq M$ für alle $x \in [a,b]$. Wir definieren f_n als gewisse Differenzenquotienten von F:

$$f_n := n \cdot \left(F\left(x + \frac{1}{n}\right) - F(x) \right) \overset{\text{Def}}{=} n \cdot \left(\int_a^{x+\frac{1}{n}} f(t)dt - \int_a^x f(t)dt \right) = n \cdot \int_x^{x+\frac{1}{n}} f(t)dt$$

Die f_n konvergieren gegen die Ableitung von F im Punkt x. Da f beschränkt ist, sind auch die f_n als Integral über einem kleinen Intervall beschränkt:

$$|f_n(x)| = \left| n \cdot \int_x^{x+\frac{1}{n}} f(t)dt \right| \leq n \cdot \int_x^{x+\frac{1}{n}} |f(t)|dt \leq n \cdot \int_x^{x+\frac{1}{n}} M \ dt = M \ \ \forall x \in [a,b]$$

Da weiter der Maßraum $([a,b],\mathcal{B}([a,b]),\lambda)$ endlich ist, kann man bei * den Satz von Lebesgue über die majorisierte Konvergenz anwenden:

$$\begin{aligned}
\int_a^c F'(x)dx &= \int_a^c \lim_{n\to\infty} f_n(x)\ dx \overset{*}{=} \lim_{n\to\infty}\int_a^c f_n(x)dx \\
&= \lim_{n\to\infty}\int_a^c n\cdot\left(F\left(x+\frac{1}{n}\right)-F(x)\right)dx \\
&= \lim_{n\to\infty} n\left[\int_a^c F\left(x+\frac{1}{n}\right)dx-\int_a^c F(x)dx\right] \\
&= \lim_{n\to\infty} n\left[\int_c^{c+1/n} F(x)dx-\int_a^{a+1/n} F(x)dx\right]
\end{aligned}$$

Da F stetig ist, stimmt für F das Lebesgue-Integral mit dem Riemann-Integral überein und F besitzt eine Stammfunktion G, sodass aus dem klassischen Hauptsatz der Differential- und Integralrechnung folgt, dass

$$\begin{aligned}
\int_a^c F'(x)dx &= \lim_{n\to\infty} n\left[\int_c^{c+1/n} F(x)dx-\int_a^{a+1/n} F(x)dx\right] \\
&= \lim_{n\to\infty}\left[\frac{G(c+1/n)-G(c)}{1/n}-\frac{G(a+1/n)-G(a)}{1/n}\right] \\
&= G'(c)-G'(a) = F(c)-F(a) \overset{\text{Def.}F}{=} \int_a^c f(x)dx\ .
\end{aligned}$$

Also ist $\int_a^c (F'-f)d\lambda = 0$ für alle $c\in[a,b]$ und damit nach Grundlage (3) schließlich $F'=f$. Die Ableitung der Integralfunktion ist also der Integrand selbst, was zu beweisen war.

Wenden wir uns nun dem allgemeinen Fall zu, F ist also nicht zwingend beschränkt. Sei oBdA $f\geq 0$, sonst schreibe f als $f=f^+-f^-$ mit $f^+,f^-\geq 0$ und betrachte jeden Teil einzeln. Definiere die Funktionenfolge g_n punktweise durch $g_n(x) := \min(n, f(x))$. Damit hat g_n die folgenden offensichtlichen Eigenschaften:

- g_n ist beschränkt durch n.
- $g_n(x)\leq f(x)$ für alle $x\in[a,b]$, d.h. $f-g_n\geq 0$ auf $[a,b]$.
- $g_n(x)\to f(x)$ für $x\to\infty$ für alle $x\in[a,b]$.

Weiter definieren wir die Integralfunktionen

$$F_n(x) := \int_a^x g_n(t)dt, \quad G_n(x) := \int_a^x (f(t)-g_n(t))dt$$

Da die g_n beschränkt sind, ist nach dem eben bewiesenen Teil $F_n' = g_n$ f.ü. Weil $f-g_n\geq 0$, ist $G_n(x)$ monoton wachsend. Nach der Differenzierbarkeit monotoner Funktionen (Grundlagen (5)) ist G_n also f.ü. differenzierbar mit $G_n'\geq 0$.

Da $F_n(x)+G_n(x) = \int_a^x f(t)dt$, gilt auch $F_n'(x)+G_n'(x) = \frac{d}{dx}\int_a^x f(t)dt = F'(x)$ und somit $F'(x) = F_n'(x)+G_n'(x)\geq g_n(x)+0$ für alle $n\in\mathbb{N}$. Läuft $n\to\infty$, wird aus g_n nun f und aus $F'(x)\geq g_n(x)$ wird $F'\geq f$. Bildet man rechts und links nun das Integral, gilt auch

$$\int_a^b F'(x)dx \geq \int_a^b f(x)dx \overset{\text{Def}}{=} F(b)-F(a)$$

Da F als Integralfunktion der positiven Funktion f monoton wachsend ist, gilt damit nach den Grundlagen (5) hier auch die umgekehrte Relation

$$\int_a^b F'(x)dx \leq F(b)-F(a)$$

also insgesamt die Gleichheit $\int_a^b F'(x)dx = \int_a^b f(x)dx$, d.h. $\int_a^b (F'-f)(x)dx = 0$. Da aber oben gezeigt wurde, dass $F'\geq f$, ist der Integrand nichtnegativ und verschwindet fast überall, d.h. $F'=f$.

Es bleibt noch der zweite Teil zu zeigen. F ist nach Voraussetzung absolut stetig, also nach Bemerkung 3.2 (iii) von beschränkter Variation und damit nach den Grundlagen (4) Linearkombination (Differenz zweier) monotoner Funktionen. Jede dieser monotonen Funktionen ist nach Grundlagen (5) f.ü. differenzierbar und die entstehenden Ableitungen sind wiederum integrierbar. Damit ist auch F' als ihre Linearkombination integrierbar.

Zu zeigen ist, dass die Integralfunktion über F' mit F übereinstimmt. Sei $G(x) := \int_a^x F'(t)dt$. Nach Teil (i) ist G f.ü. differenzierbar mit Ableitung $G' = F'$. Da F nach Voraussetzung und G nach Teil (i) absolut stetig sind, ist es auch $H = G - F$. Für die Ableitung von H gilt $H' = (G - F)' = G' - F' = 0$. Nach Satz 3.3 ist H konstant. Damit gilt auch

$$F(x) - F(a) = G(x) - G(a) = \int_a^x F'(t)dt$$

und der Hauptsatz ist bewiesen. □

Vergleicht man zum Schluss den Hauptsatz für das Riemann- und für das Lebesgue-Integral, fällt das folgende auf: Bei Riemann benötigt man stärkere Voraussetzungen, erhält aber auch stärkere Resultate:

- f muss stetig und nicht nur integrierbar sein,
- Dafür ist die Integralfunktion F dann überall differenzierbar, bei Lebesgue nur fast überall.
- Auch stimmen bei Riemann F' und f überall überein, während wir bei Lebesgue nur eine f.ü.-Identität haben.

5. Quelle

Elstrodt, Jürgen: Maß- und Integrationstheorie, 4., korrigierte Aufl.; Berlin, Heidelberg, New York: Springer, 2005, ISBN: 3-540-21390-2