# VII. Differentiation of Lebesgue Integrals on the Line, 395-410

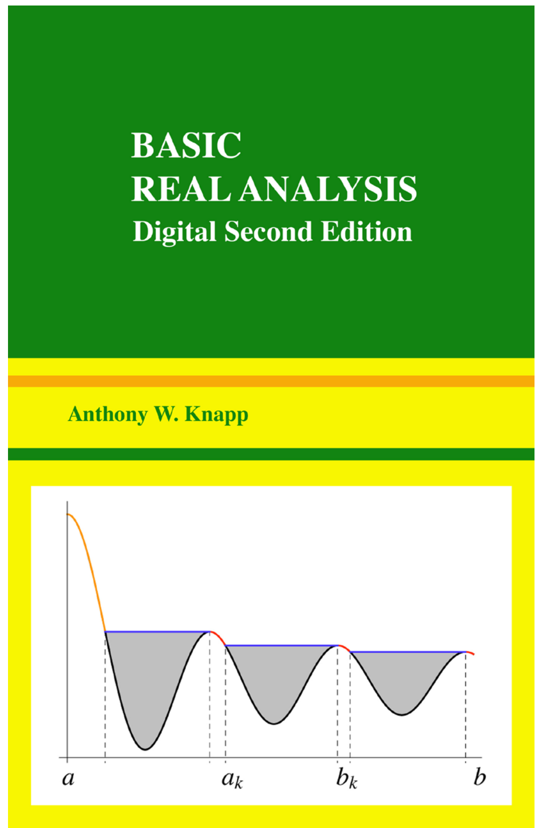
DOI: 10.3792/euclid/9781429799997-7

from

***Basic Real Analysis***
***Digital Second Edition***

Anthony W. Knapp

Full Book DOI: 10.3792/euclid/9781429799997
ISBN: 978-1-4297-9999-7



Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, N.Y. 11733–1729, U.S.A.
Email to: aknapp@math.stonybrook.edu
Homepage: www.math.stonybrook.edu/~aknapp

Title: Basic Real Analysis, with an appendix "Elementary Complex Analysis"
Cover: An instance of the Rising Sun Lemma in Section VII.1.

Mathematics Subject Classification (2010): 28–01, 26–01, 42–01, 54–01, 34–01, 30–01, 32–01.

First Edition, ISBN-13 978-0-8176-3250-2

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

Published by the author

# CHAPTER VII

# Differentiation of Lebesgue Integrals on the Line

**Abstract.** This chapter concerns the Fundamental Theorem of Calculus for the Lebesgue integral, viewed from Lebesgue's perspective but slightly updated.

Section 1 contains Lebesgue's main tool, a theorem saying that monotone functions on the line are differentiable almost everywhere. A relatively easy consequence is Fubini's theorem that an absolutely convergent series of monotone increasing functions may be differentiated term by term. The result that the indefinite integral $\int_a^x f(t)\,dt$ of a locally integrable function $f$ is differentiable almost everywhere with derivative $f$ follows readily.

Section 2 addresses the converse question of what functions $F$ have the property for a particular $f$ that the integral $\int_a^b f(t)\,dt$ can be evaluated as $F(b)-F(a)$ for all $a$ and $b$. The development involves a decomposition theorem for monotone increasing functions and a corresponding decomposition theorem for Stieltjes measures. The answer to the converse question when $f \geq 0$ and $F' = f$ almost everywhere is that $F$ is "absolutely continuous" in a sense defined in the section.

## 1. Differentiation of Monotone Functions

The generalization of the Fundamental Theorem of Calculus to the Lebesgue integral was the crowning achievement of Lebesgue's book. We have already stated and proved a particular result in that direction as Corollary 6.40, using a more recent method that is of continual applicability in analysis. The statement of the part of the Fundamental Theorem in that corollary is that $\int_a^x f(t)\,dt$ is differentiable almost everywhere with derivative $f(x)$ if $f$ is a Borel function on the line that is integrable on every bounded interval.

In this chapter we shall develop that and allied results using something closer to Lebesgue's original method. These allied results are chiefly of historical interest, no longer being of great importance as analytic tools. However, their beauty is undeniable and by itself justifies their inclusion in this book. In addition, these allied results motivate some results in Chapter IX, particularly the Radon–Nikodym Theorem, that might seem strange indeed if the historical background were omitted.

The starting point is the almost-everywhere differentiability of monotone functions on the line, given in Theorem 7.2 below. Since monotone functions include the distribution functions of Stieltjes measures, this differentiability shows at once that functions of the form $\int_a^x f(t)\,dt$ with $f \geq 0$ are differentiable almost everywhere, and then we are well on our way toward a more traditional proof of the Fundamental Theorem for the Lebesgue integral. The advantage of starting with all monotone functions is that one can address at the same time the differentiability of *all* distribution functions of Stieltjes measures, not just those of measures $f(t)\,dt$. From this fact one can attack the question of how close the derivative $f(t)$ is to determining the function of which it is the derivative almost everywhere. This is the second aspect of the traditional Fundamental Theorem of Calculus as in Theorem 1.32: for the case of continuous $f$, any two functions with derivative $f$ *everywhere* on an interval differ by a constant.

There is a certain formal similarity between the theory of differentiation of monotone functions and the theory of the Hardy–Littlewood Maximal Theorem as in Chapter VI. Wiener's Covering Lemma captured the geometric core of the theorem in Chapter VI, and another covering lemma captures the geometric core here. This is the Rising Sun Lemma, which will be given as Lemma 7.1.

By way of preliminaries, any open subset $U$ of $\mathbb{R}^1$ is uniquely the union of countably many disjoint open intervals, the open interval containing a point $x$ in $U$ being the union of all connected subsets of $U$ containing $x$. These sets give the required decomposition of $U$ by Propositions 2.48 and 2.51. An open subset of an interval $(a, b)$ is necessarily open in $\mathbb{R}^1$, and hence it too is uniquely the countable union of disjoint open intervals.

**Lemma 7.1** (Rising Sun Lemma).[1] Let $g : [a, b] \to \mathbb{R}$ be continuous, and define

$$E = \{x \in (a, b) \mid \text{there exists } \xi \in (a, b) \text{ with } \xi > x \text{ and } g(\xi) > g(x)\}.$$

The set $E$ is open in $(a, b)$. If $E$ is written as the disjoint union of open intervals with endpoints $a_k$ and $b_k$, then $g(a_k) \leq g(b_k)$ for each $k$.

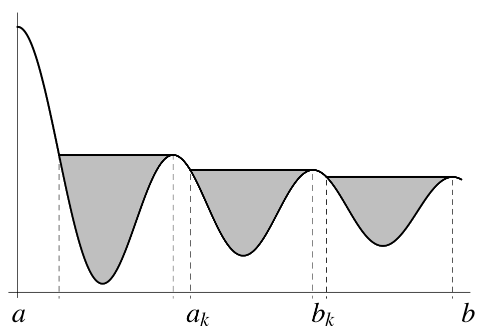



FIGURE 7.1. Rising Sun Lemma. Graph showing three open intervals produced by the lemma.

[1]Some authors call this result Riesz's Lemma.

REMARK. The Rising Sun Lemma is so named because of the situation in Figure 7.1. The sun rises in the east, viewed as the direction of the positive $x$ axis. It casts shadows within the graph of $g$, and the content of the lemma is the nature of those shadows. Although the conclusion of the lemma is that $g(a_k) \le g(b_k)$ for all $k$, the reader can observe in the figure that $g(a_k) = g(b_k)$ for the open intervals that are shown. This observation is valid in general except possibly when $a_k = a$, but the observation is not needed in the proof of Theorem 7.2 below.

PROOF. If $x_0 \in E$ and $\xi \in (a, b)$ have $\xi > x_0$ and $g(\xi) > g(x_0)$, then every $x$ in $(a, \xi)$ with $|g(x) - g(x_0)| < \frac{1}{2}(g(\xi) - g(x_0))$ lies in $E$. Hence $E$ is open.

Let $E$ be the disjoint union of intervals $(a_k, b_k)$. Fix attention on one such interval $(a_k, b_k)$. We make critical use of the fact that the point $b_k$ is not in $E$. If $x_0$ satisfies $a_k < x_0 < b_k$, we prove that $g(x_0) \le g(b_k)$. Once we do so, we can let $x_0$ decrease to $a_k$ and use continuity to obtain the assertion $g(a_k) \le g(b_k)$ of the lemma.

Arguing by contradiction, suppose that $g(x_0) > g(b_k)$. Since $x_0$ is in $E$, there exists $x_1 > x_0$ with $g(x_1) > g(x_0)$. If $x_1 > b_k$, then the inequality $g(x_1) > g(x_0) > g(b_k)$ forces $b_k$ to be in $E$. Since $b_k$ is not in $E$, we conclude that $x_1 \le b_k$. The set of all $x$ with $x_1 \le x \le b_k$ and $g(x) \ge g(x_1)$ is closed, bounded, and nonempty, and we let $x_2$ be its largest element, so that $x_2 \le b_k$.

Since $g(x_2) \ge g(x_1) > g(x_0) > g(b_k)$, we must have $x_2 < b_k$; in fact, $x_2 = b_k$ would yield the contradiction $g(b_k) > g(b_k)$. From $a_k < x_0 < x_2 < b_k$ and $(a_k, b_k) \subseteq E$, we see that $x_2$ is in $E$. Hence there is some $\xi > x_2$ with $g(\xi) > g(x_2)$. Then the conditions $g(\xi) > g(b_k)$ and $b_k \notin E$ together force $\xi$ to be $\le b_k$. So $x_2 < \xi \le b_k$ with $g(\xi) \ge g(x_1)$, in contradiction to the maximality of $x_2$. This contradiction allows us to conclude that $g(x_0) \le g(b_k)$, and the proof is complete. □

**Theorem 7.2** (Lebesgue). If $F$ is a monotone increasing function on an interval, then $F$ is differentiable almost everywhere in this sense: the set where $F$ fails to be differentiable is a Lebesgue measurable set of Lebesgue measure 0. In addition, if the definition of $F'$ is extended so that $F'(x) = 0$ at every point where $F$ is not differentiable, then $F'$ is Lebesgue measurable.

REMARK. Recall that any monotone increasing function $F$ can have only countably many discontinuities, and these are all given by jumps. In other words, $F$ has, at each point $x$, left and right limits $F(x-)$ and $F(x+)$, and the only possible discontinuities occur when one or both of the equalities $F(x-) = F(x)$ and $F(x) = F(x+)$ fail.

PROOF. The second statement is a consequence of the first. In fact, if $E$ is the Lebesgue measurable set of measure 0 where $F$ is nondifferentiable and if $B$ is a Borel set of measure 0 containing $E$, then the sequence of Borel functions

$G_n(x) = \frac{1}{1/n}(F(x+1/n) - F(x))$ converges everywhere on $B^c$ to a function $G$. If $G$ is extended to the domain of $F$ by defining it to be 0 on $B$, then $G$ is a Borel function that equals $F'$ except on a subset of $B$, and hence $F'$ is Lebesgue measurable.

Let us come to the conclusion about differentiability. Possibly by taking the union of countably many sets, we may assume that the domain of $F$ is a bounded interval $[a, b]$. For $a < x < b$, define

$$U_r(x) = \limsup_{h \downarrow 0} \tfrac{1}{h}(F(x+h) - F(x))$$

and

$$L_r(x) = \liminf_{h \downarrow 0} \tfrac{1}{h}(F(x+h) - F(x)),$$

$$U_l(x) = \limsup_{h \uparrow 0} \tfrac{1}{h}(F(x+h) - F(x))$$

and

$$L_l(x) = \liminf_{h \uparrow 0} \tfrac{1}{h}(F(x+h) - F(x)).$$

We shall prove that

$$U_r(x) < +\infty$$

and

$$U_r(x) \le L_l(x)$$

almost everywhere. If the latter inequality is applied to $-F(-x)$, we obtain also

$$U_l(x) \le L_r(x)$$

almost everywhere. Putting these inequalities together, we have $U_l(x) \le L_r(x) \le U_r(x) \le L_l(x) \le U_l(x)$ almost everywhere, and equality must hold throughout, almost everywhere. The points where equality holds throughout and also $U_r(x) < +\infty$ are the points where $F$ is differentiable, and hence the two inequalities $U_r(x) < +\infty$ and $U_r(x) \le L_l(x)$ prove the theorem.

For most of the proof, we shall assume that $F$ is continuous. At the end we return and show how to modify the proof to handle discontinuous $F$. First we consider the inequality $U_r(x) < +\infty$. The subset $E$ of $(a, b)$ where this inequality fails is, for each positive integer $n$, contained in the set where $U_r(x) > n$. If $U_r(x) > n$, then $\dfrac{F(\xi) - F(x)}{\xi - x} > n$ for some $\xi > x$. That is, $g(\xi) > g(x)$ for the continuous function $g(x) = F(x) - nx$. In the notation of Lemma 7.1, $E$ is covered by a system of disjoint open intervals $(a_k, b_k)$ such that $g(a_k) \le g(b_k)$ for each such interval. Thus $n(b_k - a_k) \le F(b_k) - F(a_k)$ for each. Summing on $k$ gives $n \sum_k (b_k - a_k) \le \sum_k (F(b_k) - F(a_k)) \le F(b) - F(a)$. Thus the exceptional set $E$ can be covered by a system of open intervals of total measure $\le \frac{1}{n}(F(b) - F(a))$. Since $n$ is arbitrary, Proposition 5.39 shows that $E$ is Lebesgue measurable of Lebesgue measure 0.

Next we prove that $U_r(x) \le L_l(x)$ almost everywhere on $(a, b)$. If $0 \le p < q$ are rational numbers, we prove that the set $E_{pq}$ where

$$L_l(x) < p < q < U_r(x)$$

has Lebesgue measure 0. The countable union of such sets is the exceptional set in question, and thus we will have proved that the exceptional set has measure 0. If $L_l(x) < p$, then there exists $\xi \in (a, b)$ with $\xi < x$ and $\dfrac{F(\xi) - F(x)}{\xi - x} < p$, hence with $p\xi - F(\xi) < px - F(x)$. Define $g(z) = pz + F(-z)$ for $z$ in $[-b, -a]$. If $y = -x$ and $\eta = -\xi$, then $p\eta + F(-\eta) > py + F(-y)$ and hence $g(\eta) > g(y)$ with $\eta > y$. Applying Lemma 7.1 to $g$ on the interval $[-b, -a]$, we obtain a disjoint system of open intervals $(-b_i, -a_i)$ covering the set of $y$'s where $L_l(-y) < p$ and having $g(-b_i) \le g(-a_i)$ in each case. Thus $-pb_i + F(b_i) \le -pa_i + F(a_i)$. In other words, the set of $x$'s where $L_l(x) < p$ is covered by a disjoint system of open intervals $(a_i, b_i)$ such that

$$F(b_i) - F(a_i) \le p(b_i - a_i) \qquad (*)$$

for each such interval. Applying the lemma to $g_p(x) = F(x) - qx$ on the interval $[a_i, b_i]$, we obtain a disjoint system of open intervals $(a_{ij}, b_{ij})$ indexed by $j$ and having $g_p(a_{ij}) \le g_p(b_{ij})$. Thus $(*)$ and

$$q(b_{ij} - a_{ij}) \le F(b_{ij}) - F(a_{ij}) \qquad (**)$$

hold in each case. Summing $(**)$ over $j$, we obtain

$$q \sum_j (b_{ij} - a_{ij}) \le \sum_j (F(b_{ij}) - F(a_{ij})) \le F(b_i) - F(a_i) \le p(b_i - a_i). \qquad (\dagger)$$

Summing this inequality over $i$ and dividing by $q$ gives

$$m(E_{pq}) \le \sum_{i,j} (b_{ij} - a_{ij}) \le (p/q)(b - a).$$

If we repeat this argument with $[a_{ij}, b_{ij}]$ in place of $[a, b]$, we obtain intervals $(a_{ijuv}, b_{ijuv})$ and an inequality

$$m(E_{pq}) \le \sum_{i,j,u,v} (b_{ijuv} - a_{ijuv}) \le (p/q) \sum_{i,j} (b_{ij} - a_{ij}) \le (p/q)^2 (b - a).$$

Iteration gives $m(E_{pq}) \le (p/q)^n (b - a)$ for every $n$, and therefore $m(E_{pq}) = 0$. This completes the proof in the case that $F$ is continuous.

If $F$ is possibly discontinuous, we modify Lemma 7.1 and the present proof as follows. Each function $g$ that arises has right and left limits $g(x+)$ and $g(x-)$ at each point $x$, and we let $G(x)$ be the largest of $g(x-)$, $g(x)$, and $g(x+)$. A modified Lemma 7.1 says that the set of $x$ in $(a, b)$ for which there is some $\xi \in (a, b)$ with $\xi > x$ and $g(\xi) > G(x)$ is an open set whose component intervals $(a_k, b_k)$ have $g(a_k+) \le G(b_k)$ for each $k$. Going over the proof of Lemma 7.1 carefully and changing $g$ to $G$ as necessary, we obtain a proof of the modified Lemma 7.1.

The modifications necessary to the present proof are as follows. In the proof that $U_r(x) < +\infty$ almost everywhere, the set $E$ is to be taken to be the set where $F$ is continuous and this inequality fails. The inequality that results from applying the modified Lemma 7.1 is $n(b_k - a_k) \le F(b_k+) - F(a_k+)$, and this inequality can be summed on $k$ without any further change. Similarly in the proof that $U_r(x) \le L_l(x)$ almost everywhere, the set $E_{pq}$ is to be taken to be the set where $F$ is continuous and $L_l(x) < p < q < U_r(x)$. Inequality $(*)$ becomes $F(b_i-) - F(a_i+) \le p(b_i - a_i)$. When we consider the interval $[a_i, b_i]$, the value of $F(b_i+)$ is not relevant, and the value of $F(b_i)$ can be adjusted to equal $F(b_i-)$ for purposes of understanding $F$ between $a_i$ and $b_i$. With that understanding, inequality $(**)$ becomes $q(b_{ij} - a_{ij}) \le F(b_{ij}+) - F(a_{ij}+)$, and step (†) is replaced by

$$q \sum_j (b_{ij} - a_{ij}) \le \sum_j (F(b_{ij}+) - F(a_{ij}+)) \le F(b_i-) - F(a_i+) \le p(b_i - a_i).$$

The two inequalities at the ends have come about from $(*)$ and $(**)$, and the critical observation is that the convention $F(b_i) = F(b_i-)$ makes the middle inequality hold. The rest of the argument proceeds as in the case that $F$ is continuous, and then the theorem is completely proved. □

**Theorem 7.3** (Fubini's theorem on the differentiation of series of monotone functions). If $F = \sum F_n$ is an absolutely convergent sequence of monotone increasing functions on $[a, b]$, then $F'(x) = \sum_{n=1}^{\infty} F_n'(x)$ almost everywhere.

PROOF. Without loss of generality, we may assume that $F_n(a) = 0$ for all $n$. Then $F_n(x) \ge 0$ for all $n$ and $x$. Possibly by lumping terms, we may assume also that $F(b) - \sum_{k=1}^{n} F_k(b) \le 2^{-n}$. Since $F(x) - \sum_{k=1}^{n} F_k(x)$ is a monotone increasing function that is 0 for $x = a$, we have

$$0 \le F(x) - \sum_{k=1}^{n} F_k(x) \le 2^{-n} \qquad (*)$$

for $a \le x \le b$. The decomposition $F(x) = \sum_{k=1}^{n} F_k(x) + \left(\sum_{k=n+1}^{\infty} F_k(x)\right)$ exhibits $F$ as the sum of $n+1$ monotone increasing functions, and thus we have

$\sum_{k=1}^{n} F_k'(x) \leq F'(x)$ at all points where all the derivatives exist. In view of Theorem 7.2, this inequality holds almost everywhere. Consequently

$$0 \leq \sum_{k=1}^{\infty} F_k'(x) \leq F'(x) \tag{$**$}$$

almost everywhere. Now consider the series

$$G(x) = \sum_{n=1}^{\infty} \Big( F(x) - \sum_{k=1}^{n} F_k(x) \Big).$$

Then

$$0 \leq G(x) - \sum_{n=1}^{N} \Big( F(x) - \sum_{k=1}^{n} F_k(x) \Big) \leq \sum_{n=N+1}^{\infty} 2^{-n} = 2^{-N}.$$

Thus $G$ satisfies the same kind of inequality that $F$ did in $(*)$, and we can conclude that $G$ satisfies the analog of $(**)$, namely

$$0 \leq \sum_{n=1}^{\infty} \Big( F'(x) - \sum_{k=1}^{n} F_k'(x) \Big) \leq G'(x).$$

The right side is finite almost everywhere by Theorem 7.2, and thus the individual terms $F'(x) - \sum_{k=1}^{n} F_k'(x)$ of the series tend to 0 almost everywhere. This completes the proof. □

From Theorems 7.2 and 7.3, we can derive the first part of Lebesgue's form of the Fundamental Theorem of Calculus. This same result was stated as Corollary 6.40 and was proved in Chapter VI by using the Hardy–Littlewood Maximal Theorem.

**Corollary 7.4** (first part of Lebesgue's form of the Fundamental Theorem of Calculus). If $f$ is integrable on every bounded subset of $\mathbb{R}^1$, then $\int_a^x f(y)\,dy$ is differentiable almost everywhere and

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x) \qquad \text{almost everywhere.}$$

PROOF. It is enough to prove the theorem for functions vanishing off an interval $[a,b]$. Let $\mathcal{A}$ be the set of all Borel sets $E \subseteq [a,b]$ such that $\frac{d}{dx}\int_a^x I_E(t)\,dt = I_E(x)$ almost everywhere. Then $\mathcal{A}$ contains the elementary sets within $[a,b]$, and $\mathcal{A}$ is closed under complements within $[a,b]$. If $\{E_n\}$ is an increasing sequence of sets in $\mathcal{A}$ with $E_0 = \varnothing$ and with union $E$, let us write $I_E = \sum_{n=1}^{\infty} (I_{E_n} - I_{E_{n-1}})$. This is a series of nonnegative functions. Putting $F(x) = \int_a^x I_E(t)\,dt$ and $F_n(x) = \int_a^x (I_{E_n}(t) - I_{E_{n-1}}(t))\,dt$ and applying Corollary 5.27, we obtain $F(x) = \sum_{n=1}^{\infty} F_n(x)$. Then Theorem 7.3 gives $F'(x) = \sum_{n=1}^{\infty} F_n'(x) = \lim_N \sum_{n=1}^{N} F_n'(x) = \lim_N \sum_{n=1}^{N} (I_{E_n}(x) - I_{E_{n-1}}(x)) = \lim_N I_{E_N}(x) = I_E(x)$ almost everywhere. Thus $E$ is in $\mathcal{A}$, and $\mathcal{A}$ is closed under increasing countable unions. Since $\mathcal{A}$ is closed under complements as well, $\mathcal{A}$ is closed under decreasing countable intersections. Then the Monotone Class Lemma (Lemma 5.43) shows that $\mathcal{A}$ contains all Borel sets.

Now consider the set $\mathcal{F}$ of all integrable Borel functions $f$ for which the almost-everywhere equality $\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$ holds. We have just seen that $\mathcal{F}$ contains all indicator functions of Borel subsets of $[a, b]$. By linearity, $\mathcal{F}$ contains all nonnegative simple functions vanishing off $[a, b]$. Let $f \geq 0$ be an integrable function on $[a, b]$, and let $\{s_n\}$ be an increasing sequence of nonnegative simple functions with pointwise limit $f$. The functions $s_n$ are in $\mathcal{F}$. Put $s_0 = 0$, and let $F(x) = \int_a^x f(t)\,dt$ and $F_n(x) = \int_a^x (s_n(t) - s_{n-1}(t))\,dt$. Since $s_n \geq s_{n-1}$, each $F_n$ is monotone increasing. Corollary 5.27 shows that $F(x) = \sum_{n=1}^{\infty} F_n(x)$, and Theorem 7.3 then shows that $F'(x) = \sum_{n=1}^{\infty} F_n'(x) = \lim_N \sum_{n=1}^{N} F_n'(x) = \lim_N \sum_{n=1}^{N} (s_n(x) - s_{n-1}(x)) = \lim_N s_n(x) = f(x)$ almost everywhere. Thus $\mathcal{F}$ contains all nonnegative integrable Borel functions, and by linearity it contains all integrable Borel functions. □

## 2. Absolute Continuity, Singular Measures, and Lebesgue Decomposition

In this section we address questions about the Lebesgue integral raised by the second part of the Fundamental Theorem of Calculus in Theorem 1.32. For continuous integrands $f$, the result is a kind of uniqueness statement, asserting that any function with derivative $f$ differs from $\int_a^x f(t)\,dt$ by a constant function. From a practical point of view, this is the really important part of the theorem for calculus, since it provides a technique for evaluating definite integrals: find any function whose derivative is the given function, evaluate it at the endpoints, and subtract the results. With the Lebesgue integral and equality of derivatives only almost everywhere, the uniqueness result is not as sharp. The practical aspect of a uniqueness theorem is largely lost, and the resulting theory ends up having to be appreciated only as an end in itself. We begin at the following point.

**Proposition 7.5.** Every monotone increasing function on $\mathbb{R}^1$ is uniquely the sum of an indefinite integral $G(x) = \int_0^x f(t)\,dt$, where $f \geq 0$ is integrable on every bounded interval, and a monotone increasing function $H$ such that $H'(x) = 0$ almost everywhere.

PROOF. Let $F$ be a given monotone increasing function on $\mathbb{R}^1$. If $F = G + H$ with $G$ as in the statement of the proposition and with $H'(x) = 0$ almost everywhere, Corollary 7.4 shows that we must have $f = F'$. This proves uniqueness.

For existence we take $f = F'$. Regard $h$ as a positive number tending to 0 through some sequence, so that $h^{-1}(F(t+h) - F(t))$ tends to $f(t)$ for almost every $t$. If $a < b$, then

$$\int_a^b \frac{F(t+h) - F(t)}{h}\,dt = \frac{1}{h}\int_{a+h}^{b+h} F(t)\,dt - \frac{1}{h}\int_a^b F(t)\,dt$$
$$= \frac{1}{h}\int_b^{b+h} F(t)\,dt - \frac{1}{h}\int_a^{a+h} F(t)\,dt.$$

The right side tends to $F(b) - F(a)$ if $a$ and $b$ are points of continuity of $F$. By Fatou's Lemma (Theorem 5.29), $\int_a^b f(t)\,dt \le F(b) - F(a)$ if $a$ and $b$ are points of continuity of $F$. The points of continuity of $F$ are dense, and thus for general $a$ and $b$, we can find sequences of points of continuity decreasing to $a$ and increasing to $b$. Passing to the limit, we obtain

$$\int_a^b f(t)\,dt \le F(b-) - F(a+) \le F(b) - F(a) \qquad (*)$$

for all $a$ and $b$. Hence $f$ is integrable. With $G(x)$ as in the statement of the proposition, $(*)$ gives $G(b)-G(a) \le F(b)-F(a)$. Equivalently, $F(a)-G(a) \le F(b) - G(b)$. Thus the function $H(x) = F(x) - G(x)$ is monotone increasing with $F = G + H$. Since $F$ and $G$ have derivative $f$ almost everywhere, $H$ has derivative 0 almost everywhere. □

Thus we want to identify all monotone increasing functions with derivative zero almost everywhere. The first step is to see that the question of discontinuities of a monotone function can be completely eliminated from the problem.

**Proposition 7.6.** Let $c$ be a real number. If $\{x_n\}$ is a sequence in $[a, b]$ and if $\{c_n\}$ and $\{d_n\}$ are sequences of positive real numbers with $\sum c_n$ finite and $\sum d_n$ finite, then the function

$$F(x) = c + \sum_{\substack{n \text{ with} \\ x_n \le x}} c_n + \sum_{\substack{n \text{ with} \\ x_n < x}} d_n$$

is a monotone increasing function on $[a, b]$ with $F'(x) = 0$ almost everywhere.

PROOF. The function $F$ is certainly monotone increasing. It is the convergent sum of the constant function $c$ and monotone increasing functions of the form

$$F_n(x) = \begin{cases} 0 & \text{for } x < x_n, \\ c_n & \text{for } x = x_n, \\ c_n + d_n & \text{for } x > x_n, \end{cases}$$

and the function $F_n$ has derivative 0 except at the point $x_n$. Thus the proposition follows immediately from Theorem 7.3. □

A monotone increasing function on the line whose restriction to every closed bounded interval is of the form in Proposition 7.6 is called a **saltus function**; the name comes from the Latin word for "jump." Since $\mathbb{R}^1$ is the countable union of closed bounded intervals, it follows from Proposition 7.6 that every saltus function has derivative 0 almost everywhere.

**Proposition 7.7.** Any monotone increasing function $F$ on $\mathbb{R}^1$ is uniquely the sum $F = G + S$ of a continuous monotone increasing function $G$ with $G(0) = 0$ and a saltus function $S$.

PROOF. For existence, it is enough to obtain the decomposition without insisting on the normalization $G(0) = 0$, since the sum of a saltus function and a constant is a saltus function. Let $x_0$ be a point of continuity of $F$, and enumerate the points of discontinuity of $F$ as $x_n$, $n \ge 1$. For each $n \ge 1$, define $c_n = F(x_n) - F(x_n-)$ and $d_n = F(x_n+) - F(x_n)$. Let $S$ be the saltus function

$$S(x) = \begin{cases} \sum_{x_0 \le x_n \le x} c_n + \sum_{x_0 \le x_n < x} d_n & \text{if } x \ge x_0, \\ -\sum_{x < x_n \le x_0} c_n - \sum_{x \le x_n \le x_0} d_n & \text{if } x \le x_0, \end{cases}$$

and put $G = F - S$. Then $G$ is continuous everywhere. To see that $G$ is monotone increasing, let $a < b$ be points of continuity of $F$ and $S$. We start from the equality $S(x_n+) - S(x_n-) = F(x_n+) - F(x_n-)$ and sum for $x_n$ with $a < x_n < b$ to obtain

$$\begin{aligned} S(b) - S(a) &= \sum_{a < x_n < b} (S(x_n+) - S(x_n-)) \\ &= \sum_{a < x_n < b} (F(x_n+) - F(x_n-)) \\ &\le F(b) - F(a). \end{aligned}$$

Hence $F(a) - S(a) \le F(b) - S(b)$, and we conclude that $G(a) \le G(b)$ at all points of continuity $a < b$ of $F$ and $S$. These points are dense, and $G$ is continuous everywhere. Hence $G(a) \le G(b)$ whenever $a < b$, and $G$ is monotone increasing. This proves existence. Uniqueness follows from the fact that $S(b-) - S(a+) = \sum_{a < x_n < b} (F(x_n+) - F(x_n-))$ whenever $a < b$, and the proof is complete. □

Consequently we need to understand the continuous monotone increasing functions $F$ on the line with $F'(x) = 0$ almost everywhere. The Cantor function for the standard Cantor set, constructed as in Section VI.8, is an example. For such a function, $F - F(0)$ satisfies the defining properties of the distribution function of a Stieltjes measure $\mu$ on $\mathbb{R}^1$. The continuity of $F$ is equivalent to the fact that $\mu$ contains no point masses. The following property isolates the meaning of having derivative zero almost everywhere.

**Proposition 7.8.** Suppose that $\mu$ is a Stieltjes measure with no point masses. If the distribution function $F$ of $\mu$ has $F'(x) = 0$ at every point of a Borel set $E$, then $\mu(E) = 0$.

REMARK. The proof will use the Rising Sun Lemma (Lemma 7.1). Problem 3 at the end of the chapter asks for an alternative proof by means of Wiener's Covering Lemma (Lemma 6.41). A proof using Wiener's Covering Lemma does not make use of the continuity of $F$, and therefore it is not necessary to assume in the proposition that $\mu$ has no point masses.

PROOF. We may confine our attention to an interval $[a, b]$, taking $E$ to be a subset of $[a, b]$. Since $\mu$ has no point masses, we may discard $a$ and $b$ from $E$. Fix a positive integer $n$. For every point $x$ in $E$, we have $F'(x) < \frac{1}{n}$. Therefore to each such $x$, we can associate some $\xi > x$ with $\xi$ in $(a, b)$ such that

$$\frac{F(\xi) - F(x)}{\xi - x} < \frac{1}{n}.$$

This inequality says that $\frac{1}{n}\xi - F(\xi) > \frac{1}{n}x - F(x)$, hence that the continuous function $g$ with $g(x) = \frac{1}{n}x - F(x)$ has $g(\xi) > g(x)$. The Rising Sun Lemma (Lemma 7.1) applies and shows that $E$ is covered by countably many disjoint open intervals $(a_k, b_k)$ with $g(a_k) \le g(b_k)$. Thus $\frac{1}{n}a_k - F(a_k) \le \frac{1}{n}b_k - F(b_k)$ for each $k$. Adding, we obtain

$$\mu(E) \le \sum_k \mu((a_k, b_k)) = \sum_k F(b_k) - F(a_k)) \le \tfrac{1}{n} \sum_k (b_k - a_k) \le \tfrac{1}{n}(b - a).$$

Since $n$ is arbitrary, $\mu(E) = 0$. $\square$

Again consider a continuous monotone function $F$ with derivative zero almost everywhere. The function $F - F(0)$ is the distribution function of some Stieltjes measure $\mu$ with no point masses, and Proposition 7.8 shows that there is a Borel set $E$ such that $\mu(E) = 0$ and $m(E^c) = 0$, where $m$ is Lebesgue measure. In other words, $\mu$ is concentrated completely on the set $E^c$ of Lebesgue measure 0. A Stieltjes measure $\mu$ for which there is a Borel set $F$ with $\mu(F^c) = 0$ and

$m(F) = 0$ is called a **singular** Stieltjes measure or a "Stieltjes measure singular with respect to Lebesgue measure." If also it contains no point masses, it is said to be **continuous singular**. The Stieltjes measure associated to the Cantor function for the standard Cantor set is an example. We can summarize matters either in terms of decompositions of monotone functions or in terms of decompositions of Stieltjes measures. The result in the case of monotone functions is a first answer to the question of uniqueness in the Fundamental Theorem of Calculus for the Lebesgue integral; the result in the case of Stieltjes measures gives the **Lebesgue decomposition** of Stieltjes measures.

**Theorem 7.9.** Every monotone increasing function $F$ on $\mathbb{R}^1$ decomposes uniquely as the sum $F = G + H + S$, where $G$ is the indefinite integral $G(x) = \int_0^x f(t)\,dt$ of a function $f \geq 0$ integrable on every bounded interval, $H$ is the distribution function of a continuous singular measure, and $S$ is a saltus function. The function $f$ is the derivative of $F$.

PROOF. Proposition 7.7 allows us to write $F = P + S$ uniquely, where $S$ is a saltus function and $P$ is continuous and monotone increasing with $P(0) = 0$. Proposition 7.5 says that $P = G + H$ uniquely, where $G$ is an indefinite integral $G(x) = \int_0^x f(t)\,dt$ and $H$ is monotone increasing with $H'(x) = 0$ almost everywhere. The function $f$ can be taken as the derivative of $F$. The function $H$ has $H(0) = 0$ and is continuous because $P$ and $G$ have these properties, and therefore $H$ is the distribution function of a Stieltjes measure $\mu$ containing no point masses. Since $H'(x) = 0$ almost everywhere, Proposition 7.8 shows that $\mu$ is singular. □

**Corollary 7.10** (Lebesgue decomposition). Every Stieltjes measure $\mu$ decomposes uniquely as the sum $\mu = f\,dx + \mu_{cs} + \mu_d$, where $f \geq 0$ is a function integrable on every bounded interval, $\mu_{cs}$ is a continuous singular measure, and $\mu_d$ is a countable sum of point masses such that the sum of the weights on any bounded interval is finite.

PROOF. This follows by applying Theorem 7.9 to the distribution function of $\mu$. □

The final question that we address in this section is how to recognize the particular monotone function $G(x) = \int_0^x f(t)\,dt$ from among all the monotone functions $F = G + H + S$ described in Theorem 7.9.

**Proposition 7.11.** With $m$ denoting Lebesgue measure, the following conditions on a Stieltjes measure $\mu_a$ are equivalent:

(a) $\mu_a$ is of the form $\mu_a = f\,dx$ for some function $f \geq 0$ that is integrable on every bounded interval,

(b) for each bounded interval $[a, b]$ and number $\epsilon > 0$, there exists a number $\delta > 0$ such that $\mu_a(E) < \epsilon$ whenever $E$ is a Borel subset of $[a, b]$ with $m(E) < \delta$,

(c) $\mu_a(E) = 0$ whenever $E$ is a Borel subset of $\mathbb{R}^1$ with $m(E) = 0$.

REMARK. A Stieltjes measure $\mu_a$ satisfying the equivalent conditions in this proposition is said to be **absolutely continuous** or "absolutely continuous with respect to Lebesgue measure." From any of these defining conditions, we see right away that an absolutely continuous measure contains no point masses.

PROOF. Corollary 5.24 shows immediately that (a) implies (b). To see that (b) implies (c), let $E$ be a Borel set $E$ in $\mathbb{R}^1$ with $m(E) = 0$. Applying (b) to $E \cap [a, b]$ gives $\mu_a(E \cap [a, b]) < \epsilon$ for every positive $\epsilon$, and hence $\mu_a(E \cap [a, b]) = 0$. Since $[a, b]$ is arbitrary and $\mu_a$ is completely additive, $\mu_a(E) = 0$.

To see that (c) implies (a), we appeal to Corollary 7.10 to decompose $\mu_a$ according to the Lebesgue decomposition as

$$\mu_a = f\,dx + \mu_{cs} + \mu_d, \qquad (*)$$

where $\mu_{cs}$ is continuous singular and $\mu_d$ is discrete. The measures $\mu_{cs}$ and $\mu_d$ have the property that there is a Borel set $E$ with $m(E) = 0$ such that $\mu_{cs}(E^c) = \mu_d(E^c) = 0$. Condition (c) shows that $\mu_a(E) = 0$. Evaluating $(*)$ at $E$, we obtain $0 = \mu_a(E) = 0 + \mu_{cs}(E) + \mu_d(E)$. Therefore $\mu_{cs}(E) = \mu_d(E) = 0$. Since $\mu_{cs}(E^c) = \mu_d(E^c) = 0$ also, we must have $\mu_{cs} = \mu_d = 0$, and then $(*)$ shows that $\mu_a = f\,dx$. □

In Chapter IX the implication (c) implies (a) will be generalized to a result in abstract measure theory known as the Radon–Nikodym Theorem. Meanwhile, it is conditions (b) and (c) that we can translate into a condition on the corresponding distribution function, and then we shall have our second and final answer to the question of uniqueness in the Fundamental Theorem of Calculus for the Lebesgue integral. A monotone increasing function $F$ on the line is said to be **absolutely continuous** if for each bounded interval $[a, b]$ and number $\epsilon > 0$, there exists a $\delta > 0$ such that on any countable disjoint union $\bigcup_k (a_k, b_k)$ of intervals within $[a, b]$ having total length $< \delta$, the variation $\sum_k (F(b_k) - F(a_k))$ of $F$ on that union of intervals is $< \epsilon$.

**Proposition 7.12.** A Stieltjes measure is absolutely continuous if and only if its distribution function is absolutely continuous.

PROOF. Let $\mu$ be a Stieltjes measure with distribution function $F$. Suppose that $\mu$ is absolutely continuous. Fix an interval $[a, b]$, let $\epsilon > 0$ be given, and choose $\delta > 0$ by (b) in Proposition 7.11 such that $m(E) < \delta$ implies $\mu(E) < \epsilon$.

If the set $A = \bigcup_k (a_k, b_k)$ is a countable disjoint union of intervals within $[a, b]$ having total length $< \delta$, then $m(A) < \delta$, and hence $\mu(A) < \epsilon$. Therefore $\sum_k (F(b_k) - F(a_k)) = \sum_k \mu((b_k - a_k)) = \mu(A) < \epsilon$, and we conclude that $F$ is absolutely continuous.

Conversely suppose that $F$ is absolutely continuous, and suppose that $E$ is a Borel set with $m(E) = 0$. Fix an interval $[a, b]$, and let $\epsilon > 0$ be given. By absolute continuity of $F$, there exists a $\delta > 0$ such that on any countable disjoint union $\bigcup_k (a_k, b_k)$ of intervals within $[a, b]$ having total length $< \delta$, the variation $\sum_k (F(b_k) - F(a_k))$ of $F$ on that union of intervals is $< \epsilon$. With $\delta$ defined in this way, we can find a countable disjoint union of intervals $\bigcup_k (a_k, b_k)$ covering $E \cap [a, b]$ and having total length $< \delta$. Then $\mu(E \cap [a, b]) \le \mu\big(\bigcup_k (a_k, b_k)\big) = \sum_k \mu((a_k, b_k)) = \sum_k ((F(b_k) - F(a_k)) < \epsilon$. Since $\epsilon$ is arbitrary, $\mu(E \cap [a, b]) = 0$. Since $[a, b]$ is arbitrary, $\mu(E) = 0$. Therefore $\mu$ satisfies (c) in Proposition 7.11 and is absolutely continuous. □

**Corollary 7.13** (second part of Lebesgue's form of the Fundamental Theorem of Calculus). Let $F$ be a monotone increasing function on $\mathbb{R}^1$, and let $f$ be its almost-everywhere derivative. Then $\int_a^b f(t)\,dt = F(b) - F(a)$ whenever $a < b$ if and only if $F$ is absolutely continuous.

PROOF. By Theorem 7.9 we can write $F(x) = \int_0^x f(t)\,dt + H(x) + S(x)$, where $H$ is the distribution function of a continuous singular measure and $S$ is a saltus function. For $a < b$, we then have

$$F(b) - F(a) = \int_a^b f(t)\,dt + (H(b) - H(a)) + (S(b) - S(a)).$$

If $F(b) - F(a) = \int_a^b f(t)\,dt$ whenever $a < b$, then the monotonicity of $H$ and $S$ forces $H$ and $S$ to be constant functions, say with $H(0) + S(0) = c$. Substituting, we see that $F(x) = \int_0^x f(t)\,dt + c$ for all $x$. The function $\int_0^x f(t)\,dt$ is absolutely continuous by Proposition 7.12, and the additive constant $c$ does not hurt matters. Thus $F$ is absolutely continuous.

Conversely if $F$ is absolutely continuous, then it is continuous, and its monotonicity forces $F - F(0)$ to be a distribution function of some Stieltjes measure $\mu$. Proposition 7.12 shows that the measure $\mu$ is absolutely continuous, and Proposition 7.11 shows that $\mu$ is of the form $\mu = g\,dx$. Therefore $F(x) - F(0) = \int_0^x g(t)\,dt$. By Corollary 7.4, $g = F' = f$ almost everywhere. Hence $F(b) - F(a) = \int_a^b g(t)\,dt = \int_a^b f(t)\,dt$ whenever $a < b$. □

## 3. Problems

1. In the Rising Sun Lemma (Lemma 7.1), show that $g(a_k) = g(b_k)$ if $a_k \ne a$. Give an example of a continuous $g$ for which one of the intervals $(a_k, b_k)$ has $a_k = a$ and $g(a_k) < g(b_k)$.

2. Let $m$ be Lebesgue measure. Does there exist a Lebesgue measurable set $E$ such that $m(E \cap I) = \frac{1}{2}m(I)$ for every bounded interval $I$? Why or why not?

3. Prove Proposition 7.8 using Wiener's Covering Lemma (Lemma 6.41) instead of the Rising Sun Lemma (Lemma 7.1).

4. Find all continuous monotone increasing functions on $\mathbb{R}^1$ with derivative 0 at all but countably many points.

5. Cantor sets within [0, 1] were introduced in Section II.9. Each is associated to a sequence $\{r_n\}_{n\geq 1}$ of numbers with $0 < r_n < 1$, the standard Cantor set being obtained when $r_n = 1/3$ for every $n$. Section VI.8 showed how to associate a distribution function to the standard Cantor set, and in similar fashion one can associate a distribution function to any Cantor set. Let $C$ be a Cantor set, let $F$ be the associated distribution function, and let $\mu$ be the associated Stieltjes measure. The Lebesgue measure of $C$ is the number $P = \prod_{n=1}^{\infty}(1 - r_n)$. Prove that
   (a) $\mu$ is singular if $P = 0$,
   (b) $\mu$ is absolutely continuous if $P > 0$, being of the form $P^{-1}I_C(x)\,dx$.

Problems 6–7 concern the **Lebesgue set** of an integrable function $f$ on an interval $[a, b]$. This is the set where $\frac{d}{dx}\int_a^x |f(t)-f(x)|\,dt$ exists and equals 0. Many almost-everywhere convergence results involving $f$ are valid at *every* point of the Lebesgue set. Such results may be regarded as relatively straightforward consequences of Corollary 7.4. Conversely an almost-everywhere convergence theorem that fails to hold at some point of the Lebesgue set might well be expected to involve some new idea.

6. For $f$ integrable on $[a, b]$, prove that almost every point of $(a, b)$ is in the Lebesgue set of $f$ by showing that the Lebesgue set of $f$ is the same as the set where $\frac{d}{dx}\int_a^x |f(t) - r|\,dt \neq |f(x) - r|$ for some rational $r$.

7. The Fejér kernel, which was defined in Section I.10 and studied further in Section VI.7, is the periodic function defined for $t$ in $[-\pi, \pi]$ by $K_N(t) = \frac{1}{N+1}\frac{1-\cos(N+1)t}{1-\cos t}$. Let $f$ be integrable on $[-\pi, \pi]$, regard $f$ as periodic, and let $x$ be in the Lebesgue set of $f$. Prove that $\lim_N (K_N * f)(x) = f(x)$ by following these steps:
   (a) Check that the estimates $K_N(t) \leq N + 1$ and $K_N(t) \leq c/(Nt^2)$ are valid for all $N$ and for $|t| \leq \pi$.
   (b) Check that the problem is to show that $\int_{|t|\leq\pi} K_N(t)|f(x-t) - f(x)|\,dt$ tends to 0 as $N$ tends to infinity.
   (c) Break the integral in (b) into pieces where $|t| \leq 1/N$, where $2^{k-1}/N \leq |t| \leq 2^k/N$ for $1 \leq k \leq \log_2(N^{3/4})$, and where $1/N^{1/4} \leq |t| \leq \pi$. Using the better of the bounds in (a) in each piece, prove the statement that (b) says needs to be shown.

Problems 8–12 concern singular Stieltjes measures, which for notational convenience we assume are continuous singular. In all these problems it is assumed that $\mu$ is a continuous singular measure and $m$ is Lebesgue measure. Among other things these problems prove that the indefinite integral of $\mu$ has derivative 0 almost everywhere with respect to Lebesgue measure, i.e., $\frac{d}{dx}\int_0^x d\mu(t) = 0$ a.e. $[dx]$, with the tools of Chapter VI and without Theorem 7.2.

8. If $\epsilon > 0$ is given, prove by considering $m + \mu$ that there exists an open set $U$ in $\mathbb{R}^1$ such that $\mu(U) < \epsilon$ and $m(U^c) = 0$.
9. If $U$ is an open subset of $\mathbb{R}^1$ and $\nu$ is a Stieltjes measure with $\nu(U) = 0$, prove that $\lim_{h\downarrow 0}(2h)^{-1}\nu((x-h, x+h)) = 0$ for all $x$ in $U$.
10. Let $\nu$ be any finite Stieltjes measure, and define

$$\nu^*(x) = \sup_{h>0}(2h)^{-1}\nu((x-h, x+h)).$$

Prove for each $\xi > 0$ that $m\{x \mid \nu^*(x) > \xi\} \le 5\nu(\mathbb{R}^1)/\xi$ by imitating the proof of Theorem 6.38.

11. For the singular measure $\mu$, assume that $\mu(\mathbb{R}^1)$ is finite. Let $\epsilon > 0$ be given, and choose an open set $U$ as in Problem 8. Define Stieltjes measures $\mu_1$ and $\mu_2$ by $\mu_1(A) = \mu(A \cap U)$ and $\mu_2(A) = \mu(A - U)$. Use Problem 9 to prove that $\lim_{h\downarrow 0}(2h)^{-1}\mu_2((x-h, x+h)) = 0$ a.e. $[dx]$, and use Problem 10 to prove for all $\xi > 0$ that

$$m\Big\{x \;\Big|\; \limsup_{h\downarrow 0}(2h)^{-1}\mu_1((x-h, x+h)) > \xi\Big\} \le 5\epsilon/\xi.$$

12. Deduce from Problem 11 that $\lim_{h\downarrow 0}(2h)^{-1}\mu((x-h, x+h)) = 0$ a.e. $[dx]$. By reviewing the proof of Corollary 6.40, show how the argument in Problems 8–11 can be adjusted to yield the better conclusion that $\frac{d}{dx}\int_0^x d\mu(t) = 0$ a.e. $[dx]$.