LEBESGUE'S REMARKABLE RESULT

CHARLES A. COPPIN

Abstract. We present a proof based on a 1905 paper by Henri Lebesgue that any continuous function defined on an interval has an antiderivative *without first proving the existence of the definite integral of the function*. We also demonstrate how the definite integral is a byproduct of this proof. Instead of merely presenting an efficient proof using modern techniques, we have chosen to present a more instructive proof actually following the steps of Lebesgue in the spirit of Otto Toeplitz's [8] genetic approach.

H. Lebesgue [4] proved the existence of an antiderivative of a continuous function with domain an interval, done without the benefit of the definite integral; a rather remarkable result established almost 107 years ago. This flies in the face of the traditional approach used in contemporary calculus and analysis courses. Angus Taylor [6] in his review of *Scenes From the History of Real Functions*, by Fyodor A. Medvedev states that the author

> calls attention to another achievement of Lebesgue that is interesting and, I believe, not well-known. Lebesgue proved, without any use of integration, that every continuous function defined on a interval has a primitive defined on that interval.

Needless to say, such a statement piqued our interest. Recently, we were reminded of Lebesgue's perspective on reading the article by David M. Bressoud [2] entitled *Historical Reflections on Teaching the Fundamental Theorem of Integral Calculus*, and communication with others led us to believe that many were not aware of Lebesgue's 1905 proof. We believe it is practically unknown, most likely, because the only proof that exists is in Lebesgue's original French language paper. His proof is relatively easy to understand at some level, even in the French; however, the inner meaning or the "soul" of the proof seemed elusive to some without additional details. Unfortunately, his proof is cursory being essentially an outline of steps of a proof relying on intuition. We attempt to bring clarity to his beautiful result by being more precise. Moreover, Lebesgue's proof will never become mainstream unless it is published in English.

The purpose of this article is two fold: to make this mathematical nugget available in English, and to present a proof in modern mathematical parlance with details added for deeper understanding. We hope this presentation adds to the inherent beauty of Lebesgue's discovery. The reader should find the proof of this result useful and interesting. Moreover, this would be a good starting point for a discussion in an undergraduate analysis course.

Introduction

Consider a set of n points $\{(a_0, d_0), (a_1, d_1), \ldots, (a_n, d_n)\}$ where $a = a_0 < a_1 < \cdots < a_n = b$ points of the interval $[a, b]$. We construct a continuous function ϕ with domain including $[a, b]$ that is linear on each subinterval $[a_i, a_{i+1}], i = 0, 1, \ldots, n-1$ where the linear pieces are joined end-to-end making $\phi(a_i) = d_i, i = 0, 1, \ldots, n$. Thus, for $i = 0, 1, \ldots, n-1$, there are numbers m_i, b_i where $\phi(x) = m_i x + b_i$ for each x in $[a_i, a_{i+1}]$, and the equality $m_i a_{i+1} + b_i = m_{i+1} a_{i+1} + b_{i+1}$ holds for $i = 0, 1, \ldots, n-2$. We now define an antiderivative Φ for ϕ on $[a, b]$. First, define $\Phi_0(x) = (m_0/2)x^2 + b_0 x - (m_0/2)a_0^2 - b_0 a_0$ for each x in $[a_0, a_1]$. Second, define $\Phi_1(x) = (m_1/2)x^2 + b_1 x + \Phi_0(a_1) - (m_1/2)a_1^2 - b_1 a_1$ for each x in $[a_1, a_2]$. Continue defining, in order, $\Phi_i, i = 2, 3, \ldots, n-1$. Now that $\Phi_0(a_0), \Phi_1(a_1), \ldots, \Phi_{n-1}(a_{n-1})$ are well defined, the function Φ is defined as follows:

$$\Phi(x) = \begin{cases} \dfrac{m_0}{2}x^2 + b_0 x - \dfrac{m_0}{2}a_0^2 - b_0 a_0, & \text{if } x \in [a_0, a_1]; \\ \\ \dfrac{m_i}{2}x^2 + b_i x + \Phi_{i-1}(a_i) - \dfrac{m_i}{2}a_i^2 - b_i a_i, & \substack{\text{if } x \in [a_i, a_{i+1}], \\ i=1,2,\ldots,n-1.} \end{cases}$$

We see that Φ is piecewise quadratic, constructed from n second degree polynomials whose left and right slopes at $a_1, a_2, \ldots, a_{n-1}$, respectively, are equal.

Having described and defined Φ, the reader may understand why this construction leads to a function whose derivative is f. We believe this is the significant result of his paper. However, our goal is to present the methods Lebesgue used in his 1905 paper and not merely to prove his result.

To lay a firm foundation for this paper, we state the following. The set of all real numbers inclusively between the real numbers, a and b, where $a < b$, is called the **interval** $[a, b]$. Here, a **partition** of an interval $[a, b]$ is a finite collection of abutting subintervals whose union is $[a, b]$ with no two members having more than one point in common. When P is a partition of $[a, b]$, by $||P||$ we mean $||P|| = \max\{q - p : [p, q] \in P\}$. A **regular partition** is a partition of $[a, b]$ each two members having the same length. Moreover, a **refinement** of a partition of $[a, b]$ is itself a partition of $[a, b]$ where each end point of each member of the original partition is an end point of some member of the refinement. For convenience, at times, for $n = 1, 2, \ldots$, we use the notation P_n to denote a regular partition of $[a, b]$ with 2^{n-1} members or subintervals each of which has length $(b - a)/2^{n-1}$. Thus, P_m is a refinement of P_n for positive integers m, n where $m \geq n$. The functions ϕ_n and Φ_n will be based on the partition P_n for $n = 1, 2, \ldots$. Throughout this paper when the symbols P_n, ϕ_n, and Φ_n are used without any descriptors for the subscripts, the reader should assume that whatever symbol is being used as a subscript is a positive integer unless otherwise stated. Lebesgue did not use the terminology of partition and refinement but the concepts are there.

We are reminded that when f is continuous on $[a, b]$, it is continuous on each subinterval of $[a, b]$, and, as a result, by *the Extreme Value Theorem*, the range of f has both a maximum and a minimum value on any subinterval of $[a, b]$. It is to our advantage to note that the maximum and minimum values of a continuous function on a subinterval "approach" each other as the length of the subinterval "approaches" zero.

Oscillations

We elucidate some of the underlying ideas of the main result of Lebesgue's paper [4] that are almost hidden from the casual reader of the original French. For brevity, we assume throughout this paper that f denotes a continuous function with domain including the interval $[a, b]$.

By the **oscillation** of f on δ (written ω_δ), a subinterval of the domain of f, we mean the real number

$$\omega_\delta = \max_\delta f - \min_\delta f.$$

Of course, by *the Extreme Value Theorem*, the continuity of f is sufficient for the existence of the extreme values and, thus, ω_δ. Moreover, by the **total oscillation** of f on a partition P of $[a, b]$ (written $\Omega(P)$), we mean the maximum value of the finite set of oscillations of f on δ for each δ in P. (Remember that P is finite; thus, guaranteeing the existence of $\Omega(P)$.) By Ω_n, we mean $\Omega(P_n)$.

Whenever each of m and n is a positive integer with $m \geq n$ and $\Omega_n < \epsilon$ for some $\epsilon > 0$, then $\Omega_m < \epsilon$. (Remember that P_m is a refinement of P_n.) This result is an application of the previous definitions of oscillation and total oscillation. We state the following propositions and useful corollaries, without proofs, which are relatively straightforward applications of the finite covering theorem (Heine - Borel Theorem).

Proposition 1. *For each $\epsilon > 0$, there is a partition P of $[a, b]$ such that $\Omega_{P'} < \epsilon$ for any refinement P' of P.*

Corollary 1. *For each $\epsilon > 0$, there is a $\delta > 0$ such that $|f(u) - f(v)| < \epsilon$ for each u, v in $[a, b]$ where $|u - v| < \delta$.*

Corollary 2. *For each $\epsilon > 0$, there is $\delta > 0$ such that $\Omega(P) < \epsilon$ for each partition P of $[a, b]$ where $||P|| < \delta$.*

Proposition 2. *For each $\epsilon > 0$, there is a positive integer n such that $\Omega_m < \epsilon$ for each positive integer $m \geq n$.*

Existence of an Antiderivative of a Continuous Function

Without resorting to integrability, we now prove that any continuous function on an interval "admits" an antiderivative on that interval.

We remind the reader that throughout this paper f denotes a continuous function with domain including $[a, b]$. Moreover, let ϕ_n denote a collection of straight line segments joined end-to-end where ϕ_n and f agree at each end point of each member of P_n. Note that ϕ_n is continuous on $[a, b]$.[1] Moreover, denote by Φ_n an antiderivative of ϕ_n constructed in the manner described in the introduction.

Lemma 1. *For each positive integer n, $|f(x) - \phi_n(x)| \leq \Omega_n$ for each x in $[a, b]$.*

Proof. Suppose x is any point in $[a, b]$, P_n is any partition of $[a, b]$, δ is any member of P_n. Because f is continuous on $[a, b]$, $\min_\delta f$ and $\max_\delta f$ both exist on $[a, b]$, positioning $f(x)$ between $\min_\delta f$ and $\max_\delta f$. Since ϕ_n is a line segment agreeing with f at its end points, $\phi_n(x)$ lies between $\min_\delta f$ and $\max_\delta f$. Thus, we see that $\min_\delta f - \max_\delta f \leq f(x) - \phi_n(x) \leq \max_\delta f - \min_\delta f$. With the definition of ω_δ, we have

$$|f(x) - \phi_n(x)| \leq \max_\delta f - \min_\delta f = \omega_\delta.$$

Therefore, from the definition of Ω_n,

$$|f(x) - \phi_n(x)| \leq \Omega_n.$$

In summary, for each positive integer n, $|f(x) - \phi_n(x)| \leq \Omega_n$ for each x in $[a, b]$. □

The remaining propositions and theorems represent the major points of Lebesgue's paper in a genetic style in the spirit of Otto Toeplitz [8].

Proposition 3. *The sequence $\{\Phi_n\}$ converges uniformly on $[a, b]$.*

[1] Imagine the behavior of ϕ_n as n increases.

Proof. Suppose $\epsilon > 0$. By Proposition 2, there is a positive integer n such that

$$\Omega_m < \epsilon/(4(b-a)) \tag{1}$$

for each positive integer $m \geq n$. Now, let m be an arbitrary positive integer greater than n and $x \in (a, b]$.

Applying *the Mean Value Theorem* to $(\Phi_m - \Phi_n)$ on $[a, x]$, there exists some real number c in (a, x) for which

$$(\Phi_m - \Phi_n)'(c)(x-a) = (\Phi_m - \Phi_n)(x) - (\Phi_m - \Phi_n)(a).$$

Since Φ_m and Φ_n are antiderivatives of ϕ_m and ϕ_n, respectively, and $\Phi_m(a) = \Phi_n(a)$, we have

$$\Phi_m(x) - \Phi_n(x) = (\phi_m(c) - \phi_n(c))(x-a). \tag{2}$$

By Lemma 1,

$$|\phi_m(c) - f(c)| \leq \Omega_m \text{ and } |\phi_n(c) - f(c)| \leq \Omega_n.$$

Combining the foregoing, using the triangle inequality, we obtain

$$|\phi_m(c) - \phi_n(c)| \leq \Omega_m + \Omega_n. \tag{3}$$

Therefore, from (2) and (3), we have

$$|\Phi_m(x) - \Phi_n(x)| \leq (b-a)[\Omega_m + \Omega_n]. \tag{4}$$

Then, by (1) and (4),

$$|\Phi_m(x) - \Phi_n(x)| < \epsilon/2.$$

Or, in a more useful form, using an additional positive integer $m' \geq n$ and the triangle inequality, we obtain

$$|\Phi_m(x) - \Phi_{m'}(x)| < \epsilon.$$

Summarizing, for each $\epsilon > 0$, the case of $x = a$ being trivial, there is a positive integer n such that for each positive m and m' greater than or equal to n we have

$$|\Phi_m(x) - \Phi_{m'}(x)| < \epsilon$$

for all x in $[a, b]$; therefore, proving the proposition. □

We have proved that $\{\Phi_n(x)\}$ is a Cauchy sequence for each x in $[a, b]$. Thus, by *the Axiom of Completeness* (Reed [5]) or some equivalent axiom or theorem, $\{\Phi_n(x)\}$ converges for all x in $[a, b]$. Since all sequences of real numbers converge to a unique value, we now define $F(x)$ as $\Phi_n(x) \to F(x)$ for each x in $[a, b]$.

Proposition 4. *The sequence $\{\Phi_n\}$ is uniformly convergent to F; that is, for each $\epsilon > 0$, there is a positive integer n such that*

$$|\Phi_m(x) - F(x)| < \epsilon$$

for each positive integer $m \geq n$ and for all x in $[a, b]$.[2]

Proof. Suppose $\epsilon > 0$. By the previous proposition, there is a positive integer n_1 such that

$$|\Phi_{m'}(x) - \Phi_m(x)| < \epsilon/2 \tag{5}$$

for each positive integer $m \geq n_1$, for each positive integer $m' \geq n_1$, and for all x in $[a, b]$.

Now, let x be any member of $[a, b]$. From the definition of $F(x)$, there is a positive integer n_2 such that for each positive integer $m \geq n_2$, we have

$$|\Phi_m(x) - F(x)| < \epsilon/2. \tag{6}$$

Let $n = max\{n_1, n_2\}$ and m be any positive integer greater than or equal to n. Since $n \geq n_1, n_2$, then, from (5) and (6),

$$|\Phi_m(x) - \Phi_n(x)| < \epsilon/2 \text{ and } |\Phi_n(x) - F(x)| < \epsilon/2.$$

Applying the triangle inequality to the preceding,

$$|\Phi_m(x) - F(x)| < \epsilon$$

for *all* x in [a,b].

To summarize, for each $\epsilon > 0$, there is a positive integer n such that if m is any positive integer greater than or equal to n, then

$$|\Phi_m(x) - F(x)| < \epsilon$$

for all $x \in [a, b]$. □

Theorem 1. *$F'(x) = f(x)$ for all x in $[a, b]$.*

Proof. Let x be any member of $[a, b]$ and $\epsilon > 0$. By Corollary 1, there is $\alpha > 0$ such that

$$|f(u) - f(v)| < \epsilon/4 \tag{7}$$

for each u, v in $[a, b]$ where $|v - u| < \alpha$. From Lemma 1 and Proposition 2, for each u in $[a, b]$, there is a positive integer n_1 such that

$$|\phi_m(u) - f(u)| < \epsilon/4 \tag{8}$$

for each $m \geq n_1$. Adding (7) and (8), for each u, v in $[a, b]$ where $|v - u| < \alpha$ we have

$$|\phi_m(u) - f(v)| < \epsilon/2 \tag{9}$$

for each $m \geq n_1$.

[2]This proposition ultimately relies on the Heine-Borel Theorem.

Let h be any nonzero number where $|h| < \alpha$ and $[x-h, x]$ and $[x, x+h]$, as appropriate, are subsets of $[a, b]$. By Proposition 4, there is a positive integer n_2 such that

$$|F(t) - \Phi_m(t)| < |h|\epsilon/4$$

for each positive integer $m \geq n_2$ and for each t in $[a, b]$. Then, since $x, x+h$ are in $[a, b]$, we have

$$|F(x+h) - \Phi_m(x+h)| < |h|\epsilon/4 \text{ and } |F(x) - \Phi_m(x)| < |h|\epsilon/4$$

for each $m \geq n_2$. From the above, we derive

$$\left| \frac{F(x+h) - F(x)}{h} - \frac{\Phi_m(x+h) - \Phi_m(x)}{h} \right| < \epsilon/2 \tag{10}$$

for $x, x+h$ in $[a, b]$ and each $m \geq n_2$.

Let $n = max\{n_1, n_2\}$. Applying *the Mean Value Theorem* to Φ_n on $[x, x+h]$, we obtain

$$\Phi_n(x+h) - \Phi_n(x) = (\phi_n(c))h \tag{11}$$

for some real number c between x and $x+h$. Thus, since $n \geq n_1$, substituting (11) into (10),

$$\left| \frac{F(x+h) - F(x)}{h} - \phi_n(c) \right| < \epsilon/2. \tag{12}$$

Since c is between x and $x+h$, and $|c - x| < \alpha$, from (9) and the preceding, we obtain

$$\left| \frac{F(x+h) - F(x)}{h} - f(x) \right| < \epsilon. \tag{13}$$

To summarize, we have shown that for each x in $[a, b]$ and each $\epsilon > 0$, there is $\alpha > 0$, such that

$$\left| \frac{F(x+h) - F(x)}{h} - f(x) \right| < \epsilon$$

for each h where $|h| < \alpha$ and $[x, x+h]$ and $[x-h, x]$, as appropriate, are subsets of $[a, b]$.

By the definition for derivative, we see that $F'(x) = f(x)$ for each x in $[a, b]$. Thus, *each continuous function has an antiderivative.* $\square$

We now have accomplished our main objective, but Lebesgue was not yet finished for, after all, the title of his paper is *Remarques sur la définition de l'intégrale*. We wish to remain true to Lebesgue's approach without resorting to notational devices. That we have done. However, Lebesgue finished his paper with a proof that the integral of f exists on $[a, b]$, but, which of the various definitions that were extant in 1905 did he use? All are equivalent and commonly called the Riemann integral. Cauchy (1823) thought of the integral as the limit of $\sum_{i=1}^{n} f(a_{i-1})(a_i - a_{i-1})$ where f

is a continuous function on an interval with division points $a_0, a_1, \ldots, a_n$. He took the limit as the maximum of the lengths $a_i - a_{i-1}$ approach zero. Riemann generalized Cauchy's definition in two ways. The function f was not necessarily continuous on the interval of integration, and he used a sum of the form $\sum_{i=1}^{n} f(x)(a_i - a_{i-1})$ where x was any arbitrary point between a_{i-1} and a_i. Darboux, on the other hand, defined upper sum $U(P) = \sum \max_{\delta} f(q-p)$ and lower sums $L(P) = \sum \min_{\delta} f(q-p)$ where the sums are taken over all $\delta = [p,q]$ of the partition P, and f is a bounded function with domain $[a,b]$. Then, he defined lower integrals and upper integrals. He defined f to be integrable if the lower integral and upper integral were equal. The Darboux integral is the one most commonly used today. Just check with any undergraduate analysis text. However, this is not the definition that Lebesgue used in his paper. It appears to be very close to Riemann's definition.

Theorem 2. *The function f is integrable on $[a,b]$.*

Proof (Riemann). Suppose $\epsilon > 0$. By Corollary 2, there is $\alpha > 0$ such that $\Omega(P) < \epsilon/(b-a)$ for each partition P of $[a,b]$ where $||P|| < \alpha$. Let Q be any partition of $[a,b]$ where $||Q|| < \alpha$. Note that $F(b)-F(a) = \sum(F(q)-F(p))$ where this sum and all sums to follow are taken over all $\delta = [p,q]$ in Q. Then, as shown above, since $F'(x) = f(x)$ for each x, we apply *the Mean Value Theorem* to each $[p,q]$ in Q; thereby, proving there is a c in $[p,q]$ where $F(q) - F(p) = F'(c)(q-p) = f(c)(q-p)$ for each $[p,q]$ in Q. Thus,

$$F(b) - F(a) = \sum (F(q) - F(p)) = \sum F'(c)(q-p) = \sum f(c)(q-p).$$

Let

$$S(Q) = \sum f(x)(q-p)$$

where the sum is as described above and x is any point in $\delta = [p,q]$ for each δ in Q. From the preceding, by the definition of ω_δ, $|f(x) - f(c)| \le \omega_\delta$ for each x in δ, we obtain

$$\begin{aligned} |F(b) - F(a) - S(Q)| &\le \sum |f(c) - f(x)|(q-p) \\ &\le \sum [\max_{\delta} f - \min_{\delta} f](q-p) = \sum \omega_\delta (q-p). \end{aligned}$$

And, since $\omega_\delta \le \Omega(Q) < \epsilon/(b-a)$ for each δ in Q, we have

$$|F(b) - F(a) - S(Q)| \le \Omega(Q) \sum (q-p) = \Omega(Q)(b-a) < \frac{\epsilon}{b-a}(b-a) = \epsilon.$$

In summary, there is a number $W = f(b) - f(a)$ such that for each $\epsilon > 0$ there is $\alpha > 0$ such that

$$|W - S(Q)| < \epsilon$$

for each partition Q where $||Q|| < \alpha$. Thus, the definition of the Riemann integral is satisfied. $\square$

Conclusion

The following are suggested questions for discussion:

(a) Prove integrability of f using the Darboux integral. It is not what Lebesgue used in his paper; however, it is what most undergraduates know as the integral. The proof is amazingly short. Moreover, proving integrability in this manner will reinforce the ideas of this paper.

(b) In the spirit of Lebesgue's original proof, we used regular partitions P_n. This allowed us to speak of infinite sequences of the functions, $\{\phi_n\}$ and $\{\Phi_n\}$, rather than dealing with a partial order induced by refinements of partitions. Lebesgue likely did not have the use of what we call the ϵ-partition-refinement approach created after 1905. See Apostol [1] for an introduction to this beautiful integral. This would be worthy of debate. Is this approach better for an audience of nascent mathematicians? The proof using this method is slightly shorter than the proof we gave and is less cumbersome.

(c) Why do functions as concrete as ϕ and Φ, linear and quadratic functions, prove useful in showing that any continuous function with domain an interval has an antiderivative? To some readers the foregoing proof may appear "mysterious." Will other functions work as well?

(d) With a somewhat different and modern approach, Brian Thomson [7] proves that every bounded, continuous function on an open interval (a, b) [bounded or unbounded] where there are at most only finitely many discontinuities has an antiderivative that must be Lipschitz on (a, b). However, Thomson does not use the oscillation function defined by Lebesgue in his 1905 paper. We found this particular device to be elegant and quite instructive leading to an easily understandable proof. Comparing Thomson's result with the genetic approach of this paper following Lebesgue's proof would make a good discussion topic.

We have been true to Lebesgue's proof while keeping it at a primitive level albeit with more details, at such a level as to give the reader a "bedrock" understanding, the Zeitgeist of his result. We hope the reader has enjoyed Lebesgue's result as much as we have.

References

[1] T. M. Apostol, *Mathematical Analysis*, Addison-Wesley, Reading, MA, 1957.

[2] D. M. Bressoud, *Historical reflections on teaching the fundamental theorem of integral calculus*, The American Mathematical Monthly, **118** (2011), 99–115.
[3] L. R. Ford, *Interval-additive propositions*, The American Mathematical Monthly, **64** (1957), 106–108.
[4] H. Lebesgue, *Remarques sur la définition de l'intégrale*, Bull. Sci. Math., **29** (1905), 272–275.
[5] M. Reed, *Fundamental Ideas of Analysis*, John Wiley & Sons, Inc., New York, NY, 1998.
[6] A. Taylor, *Review of scenes from the history of real functions*, by Fyodor A. Medvedev, Bulletin of the American Mathematical Society, **28.2** (1993).
[7] B. S. Thomson, *The Calculus Integral*, 2010, `ClassicalRealAnalysis.com.`
[8] O. Toeplitz, *The Calculus: A Genetic Approach*, The University of Chicago Press, Chicago, Illinois, 1963.

MSC2010: 00-01, 26-01

Key words and phrases: antiderivative, Riemann integral

Department of Mathematics, Lamar University, Beaumont, Texas 77710
E-mail address: `charles.coppin@lamar.edu`