Generalizations of the Riemann Integral: An Investigation of the Henstock Integral

Jonathan Wells

May 15, 2011

Abstract

The Henstock integral, a generalization of the Riemann integral that makes use of the δ-fine tagged partition, is studied. We first consider Lebesgue's Criterion for Riemann Integrability, which states that a function is Riemann integrable if and only if it is bounded and continuous almost everywhere, before investigating several theoretical shortcomings of the Riemann integral. Despite the inverse relationship between integration and differentiation given by the Fundamental Theorem of Calculus, we find that not every derivative is Riemann integrable. We also find that the strong condition of uniform convergence must be applied to guarantee that the limit of a sequence of Riemann integrable functions remains integrable. However, by slightly altering the way that tagged partitions are formed, we are able to construct a definition for the integral that allows for the integration of a much wider class of functions. We investigate several properties of this generalized Riemann integral. We also demonstrate that every derivative is Henstock integrable, and that the much looser requirements of the Monotone Convergence Theorem guarantee that the limit of a sequence of Henstock integrable functions is integrable. This paper is written without the use of Lebesgue measure theory.

Acknowledgements

I would like to thank Professor Patrick Keef and Professor Russell Gordon for their advice and guidance through this project. I would also like to acknowledge Kathryn Barich and Kailey Bolles for their assistance in the editing process.

Introduction

As the workhorse of modern analysis, the integral is without question one of the most familiar pieces of the calculus sequence. But the integral that most are familiar with, the Riemann integral, is in fact but one of several. And while the Riemann integral has a certain degree of computational finesse, it possesses a handful of theoretical deficiencies as well. In this paper, we will investigate how a small change to the definition of the Riemann integral can have a profound impact on the class of integrable functions. The focus of this paper is on an investigation of one such generalization of the Riemann integral, the Henstock integral. Although alternatively referred to as the Gauge integral, the generalized Riemann integral, and the Henstock-Kurzweil integral, we will follow the most common practice, and refer to it simply as the Henstock integral.

Throughout this paper, we will assume that the reader has completed the standard undergraduate calculus sequence, is familiar with the general concept of the Riemann integral, and is comfortable working with formal mathematical proofs. We will also assume that the reader has had some exposure to the topics common to an undergraduate real analysis course, particularly with respect to the epsilon-delta style proof, and the notions of limits, sequences, series, and continuity. Since the Henstock integral, unlike the Lebesgue integral, does not require that the student first undertake a study of Lebesgue measure theory, before being capable of formulating the integral's definition, this paper will not make use of any advanced results from measure theory, or from the theory of Lebesgue integration. Unless otherwise specified, notation and terminology will follow the conventions of [1].

We begin with a brief look at the historical development of integration theory, which hopefully will offer some context and focus for the problems investigated throughout the rest of this paper. For a more comprehensive discussion of the formal beginnings of calculus and the development of the integral, we refer the reader to chapters 16, 17, and 22 in [2]. In the first section of this paper, we will present important terminology for the study of Riemann and Henstock integration processes, and then investigate the necessary and sufficient conditions for a function to be Riemann integrable. As we will see, a real-valued function is Riemann integrable if and only if it is bounded and continuous almost everywhere. In the second section of this paper, we will introduce the δ-fine tagged partition, a concept that acts as the foundation for the study of the Henstock integral. The third section will present the definition of the Henstock integral and focus on an investigation of several of its properties. In the fourth section, we will explore the relationship between the Henstock integral and the derivative. We will see that one of the greatest deficiencies of the Riemann integral is the inability to integrate all derivatives, and we will demonstrate the Henstock integral does not have the same problem. Finally, in the fifth section, we will present a discussion and proof of the Monotone Convergence Theorem, which under certain conditions, allows for the interchange of limit operations when working with sequences of integrable functions.

The Historical Development of the Integral

The concept of the integral first arose out of attempts to determine the area of curved, geometric figures. Although a few solutions to particular area problems were conjectured prior to the seventeenth century, the first major breakthrough regarding general methods for solving area problems coincided with the development of calculus. Importantly, it was the realization of the fundamental relationship between the tangent problem and the area problem, that is, between differential calculus and integral calculus, that provided the greatest impetus for the full genesis of the modern integral.

Both Isaac Newton and Gottfried Wilhelm Leibniz are credited with the invention of calculus. For Newton, the process of integration was largely seen as an inverse to the operation of differentiation, and the integral was synonymous with the anti-derivative. Leibniz, however, believed that the definite integral could be interpreted as the area under continuous curves, and could be obtained by summing an infinite series of areas corresponding to approximating rectangles of infinitely small width. As will be observed shortly, this notion, once properly formalized, gave rise to the modern conception of the Riemann integral.

While both of these notions of the integral sufficed for solving many previously intractable numerical problems, they possessed several weaknesses. First, the class of integrable curves was limited not only to continuous curves, but also to only those with elementary anti-derivatives. Second, and perhaps more problematic, neither Newton nor Leibniz's integral possessed a rigorous theoretical basis; and while both Newton's notion of fluents and Leibniz's notion of the infinitesimal had intuitive appeal, neither mathematician was able to provide sound justification for the manipulation of these mathematical objects.

In the nineteenth century, calculus underwent a radical reconstitution. Instead of focusing on the solutions to numerical problems, or on methods for calculating integrals, mathematicians began to emphasis the notion of theoretical 'rigor'; they sought to provide mathematics with a firm, and most importantly, logically justified, foundation. It was Augustin-Louis Cauchy, a prolific nineteenth century mathematician, who established calculus on the basis of the modern concept of the limit. Breaking with the tradition of the seventeenth and eighteenth centuries, Cauchy defined the integral as the limit of a sum, rather than in terms of anti-derivatives. Returning to the notion that the area under a function could be approximated by summing together the areas of appropriately selected rectangles, Cauchy noted that for continuous functions, the resulting sum became more accurate as more rectangles of smaller width were used. But rather than using this sum as an estimation for the area under a function, he defined the integral as the limit of the sum of rectangles constructed by subdividing an interval, and using the value of the function at endpoints to determine a rectangle's height. Symbolically, this can be represented as

$$\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)(x_i - x_{i-1}),$$

where x_i represents the right endpoint of each subinterval, for each $1 \leq i \leq n$.

Nevertheless, Cauchy's definition guaranteed the existence of the definite integral only for functions with at most a finite number of discontinuities. For this reason, Georg Bernhard Riemann sought to generalize Cauchy's integral so that a wider class of functions could be integrated. Riemann did so by allowing the height of the approximating rectangles to be determined by any point in the corresponding subinterval, rather than merely by the endpoints. Thus, Riemann's integral took the form of

$$\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}),$$

where t_i represents a sample point taken from the interval $[x_{i-1}, x_i]$. With this definition in hand, Riemann then set about determining precisely which functions did and did not have definite integrals. A comprehensive study of Riemann's integral is usually the subject of the undergraduate real analysis course.

While the Riemann integral at last provided integral calculus with a stable foundation, it was not without its own problems. The integral has several theoretical deficiencies, and as mathematicians in the late nineteenth and early twentieth centuries sought to expand the concepts of analysis beyond the set of real numbers, this necessitated the construction of more abstract definitions of the integral. Perhaps the most famous of these is the Lebesgue integral, which was developed in early twentieth century by Henri Lebesgue. Because of both its complexity and generality, the Lebesgue integral is often studied extensively only at the graduate level and beyond. In the 1950s, mathematicians Jaroslav Kurzweil and Ralph Henstock constructed a new definition of the integral that bore striking similarity to the Riemann definition; in fact, this new integral, the Henstock integral can be seen as a direct generalization of the Riemann integral, and will be the focus of our paper.

1 The Riemann Integral

While the reader is undoubtedly familiar with the general concept of the integral and of several of the methods for performing integration from a study of undergraduate level calculus, we will review in this section the precise definition of the Riemann integral and other requisite terminology and notation, before we dive into an investigation of the necessary and sufficient conditions for a real-valued function to be Riemann integrable. Since the Henstock integral is a direct generalization of the Riemann integral, much of the terminology developed in the study of the latter will be relevant for the study of the former. As we assume that the reader has had some previous exposure to a rigorous treatment of the Riemann integral, we limit our discussion here to only those elements that have direct application to the Henstock integral; for a further investigation of the properties of the Riemann integral, see [1].

We begin by taking a look at several concepts related to subdivisions of closed, bounded intervals.

Definition 1. *A* ***partition*** P *of an interval* $[a, b]$ *is a finite set of points* $\{x_i : 0 \leq i \leq n\}$ *such that* $a = x_0 < x_1 < ... < x_{n-1} < x_n = b$.

Although, strictly speaking, a partition refers to an ordered set of points, since this set naturally subdivides an interval $[a, b]$ into the set of intervals $\{[a, x_1], [x_1, x_2], ..., [x_{n-1}, b]\}$, we will often use the word partition to denote the resulting set of subintervals; nevertheless, when precision is necessary, it should be clear from the context of the discussion which set is meant.

In the context of integration theory, we are often interested in determining the length of the largest of these subintervals.

Definition 2. *The* ***norm*** $\|P\|$ *of the partition* P *is equal to* $\max\{|x_i - x_{i-1}| : 1 \leq i \leq n\}$.

Given a partition of an interval, it is often useful to choose one point as a tag from each of the partition's subintervals; we call the resulting construction a tagged partition.

Definition 3. *A* ***tagged partition*** ${}^tP$ *of an interval* $[a, b]$ *consists of a partition* $P = \{x_i : 0 \leq i \leq n\}$ *of* $[a, b]$ *along with a set* $\{t_i : 1 \leq i \leq n\}$ *of points that satisfy* $x_{i-1} \leq t_i \leq x_i$ *for each* i.

As an example of a tagged partition, consider the interval $[0, 3]$ divided into three subintervals of equal length, where the left endpoint of each subinterval is selected as a tag. The resulting tagged partition is represented as

$$^tP = \{(0, [0, 1]), (1[1, 2]), (2, [2, 3])\}.$$

We recall from our study of calculus that (at least intuitively speaking) the area under a nonnegative curve may be estimated by summing together the areas of carefully selected rectangles. This intuitive approximation gives rise to the following definition of the Riemann Sum.

Definition 4. *Let $f : [a, b] \to \mathbb{R}$ and let $^tP = \{(t_i, [x_{i-1}, x_i]) : 1 \le i \le n\}$ be a tagged partition of $[a, b]$. Then the **Riemann Sum** $S(f,^t P)$ of f on ^tP is defined by*

$$S(f,^t P) = \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}).$$

If we let ^tP be the tagged partition of $[0, 3]$ defined previously, and consider the function $f(x) = x^2$ on this interval, the Riemann Sum of f over ^tP is equal to

$$S(f, ^tP) = 0(1 - 0) + 1(2 - 1) + 4(3 - 2) = 5.$$

Immediately, we notice that this Riemann Sum gives a relatively poor approximation to the area under the given curve, and if we partition the interval into more subintervals, we likely will obtain a better estimate. We say that a function is Riemann integrable if its Riemann sums approach a real limit as the norms of the corresponding tagged partitions approach zero, and formalize this idea with the following definition.

Definition 5. *A function $f : [a, b] \to \mathbb{R}$ is **Riemann integrable** on $[a, b]$ if there exists a number S such that for each $\epsilon > 0$ there exists $\delta > 0$ such that $|S(f,^t P) - S| < \epsilon$ for all tagged partitions ^tP of $[a, b]$ with norms less than δ. We call S the Riemann integral of f on $[a, b]$.*

What follows is an investigation of the necessary and sufficient conditions for a function to be Riemann integrable. Most introductory calculus textbooks restrict the study of integration to the set of continuous functions on closed, bounded intervals. However, there certainly exist functions that are not continuous, but that are nevertheless Riemann integrable (for example, a step function is not continuous, but is certainly integrable). Most standard real analysis texts provide discussion of other sufficient conditions guaranteeing integrability.

As it turns out, a function is Riemann integrable on a closed and bounded interval if and only if it is bounded and continuous almost everywhere on that interval (where by **continuous almost everywhere** we mean that the function is continuous everywhere on that interval except for on a set of points in that interval with measure zero); this result, which was formulated by Henri Lebesgue in the first decade of the twentieth century, is often referred to as Lebesgue's Criterion for Riemann Integrability. While the proof that every Riemann integrable function is bounded and continuous almost everywhere is relatively straightforward, a proof of the converse is somewhat more difficult; thus, we will save the latter until after we have developed some further terminology.

Theorem 1 (Lebesgue's Criterion for Riemann Integrability)**.** *A function is Riemann integrable if and only if it is bounded and continuous almost everywhere.*

Before we begin this study, we will need to make clear what is meant by a set of measure zero.

Definition 6. *For an interval I with endpoints a and b, let the* ***length*** *$\ell(I)$ of the interval be defined as $b-a$. A set S of real numbers has* ***measure zero*** *if for each $\epsilon > 0$, there exists a sequence $\{I_k\}$ of open intervals such that $E \subseteq \bigcup_{k=1}^{\infty} \ell(I_k)$ and $\sum_{k=1}^{\infty} \ell(I_k) < \epsilon$.*

Perhaps the most straightforward approach to proving that every Riemann integrable function is continuous almost everywhere is to use the oscillation of a function at a point.

Definition 7. *If $f : [a,b] \to \mathbb{R}$ is a bounded function $f : [a,b] \to \mathbb{R}$, the* ***oscillation*** *of f on $[a,b]$ is defined to be*

$$\omega(f,[a,b]) = \sup\{|f(x)-f(y)| : x,y \in [a,b]\}.$$

If $c \in (a,b)$, then the oscillation of the function at the point c is defined to be

$$\omega(f,c) = \lim_{r\to 0^+} \omega(f,[c-r,c+r]).$$

With this definition in place, we will first prove three lemmas concerning the oscillation of a continuous function, and then use these to prove our desired result. The first establishes an equivalent condition for the continuity of a function at a point using the idea of oscillation. In brief, this lemma states that the difference between the values of a function at points in a small interval containing a point of continuity is small.

Lemma 1. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function and let $c \in (a,b)$. The function f is continuous at c if and only if the oscillation of f at c is 0.*

Proof. Let $\epsilon > 0$. Suppose f is continuous at $c \in (a,b)$. Then there exists $\delta > 0$ such that $|f(x) - f(c)| < \frac{\epsilon}{2}$ for all x in the interval $[c-\delta, c+\delta]$. Now, by the triangle inequality, and since f is continuous at c, we have that

$$\begin{aligned}
\omega(f,[c-\delta,c+\delta]) &= |\sup\{|f(x)-f(y)| : x,y \in [c-\delta,c+\delta]\}| \\
&< |\sup\{|f(x)-f(c)| + |f(y)-f(c)| : x,y \in [c-\delta,c+\delta]\}| \\
&< |\sup\{|\frac{\epsilon}{2}+\frac{\epsilon}{2}| : x,y \in [c-\delta,c+\delta]\}| \\
&< \epsilon.
\end{aligned}$$

It follows that $\lim_{\delta\to 0^+} \omega(f,[c-\delta,c+\delta]) = 0$, and thus, that the oscillation of f at c is 0.

Now, let $\epsilon > 0$, and suppose that the oscillation of the function f as c is 0. Since $\omega(f,c) = 0$, there exists $r > 0$ such that $\omega(f,[c-r,c+r]) < \epsilon$. Let $x \in [c-r,c+r]$. Then

$$|f(x)-f(c)| < \sup\{|\ f(s)-f(t)\ |: s,t \in [c-r,c+r]\} < \epsilon.$$

Since the choice of x in the interval $[c-r,c+r]$ was arbitrary, it follows that f is continuous at the point c. □

In order to prove the following lemma, we will use an alternative condition for the integrability of a function that is equivalent to the definition of the Riemann integral. A proof of this theorem can be found in [1].

Theorem 2 (Riemann's Criterion)**.** *Let f be a bounded function defined on $[a,b]$. The function f is Riemann integrable on $[a,b]$ if and only if for each $\epsilon > 0$ there exists a partition $P = \{x_i : 0 \leq i \leq n\}$ of $[a,b]$ such that*

$$\sum_{i=1}^{n} \omega(f,[x_{i-1},x_i])(x_i - x_{i-1}) < \epsilon.$$

The purpose of the next lemma is to construct a countable collection of sets of measure zero that cover the set of discontinuities of a function.

Lemma 2. *Suppose that f is a Riemann integrable function on $[a,b]$, let N be any positive integer, and let D_N be the set $\{x \in (a,b) : \omega(f,c) \geq 1/N\}$. The set D_N has measure zero.*

Proof. Since f is Riemann integrable on $[a,b]$, we know that there exists a partition P of $[a,b]$ such that $\displaystyle\sum_{i=1}^{n} \omega(f,[x_{i-1},x_i])(x_i - x_{i-1}) < \frac{\epsilon}{2N}$. Define the sets S_1 and S_2 by

$$S_1 = \{j : D_N \cap [x_{j-1},x_j] \neq \emptyset\} \qquad S_2 = \{k : D_N \cap [x_{k-1},x_k] = \emptyset\}.$$

Therefore, $D_N \subseteq \displaystyle\bigcup_{j \in S_1} [x_{j-1},x_j]$. Now, for each $j \in S_1$, define the open interval I_j by

$$I_j = \left(x_{j-1} - \frac{x_j - x_{j-1}}{2}, x_j + \frac{x_j - x_{j-1}}{2}\right).$$

We see that $\ell(I_j) = 2(x_j - x_{j-1})$ and that $D_N \subseteq \displaystyle\bigcup_{j \in S_1} I_j$. We will now show that $\displaystyle\sum_{j \in S_1} \ell(I_j) < \epsilon$. First, we note that for any $j \in S_1$, $\omega(f,[x_{j-1},x_j]) \geq \frac{1}{N}$, since there exists a point $c \in D_N \cap [x_{j-1},x_j]$ such that the oscillation of the function at that point is greater than or equal to $\frac{1}{N}$. Then

$$\begin{aligned}
\sum_{j \in S_1} \frac{1}{N}(x_j - x_{j-1}) &\leq \sum_{j \in S_1} \omega(f,[x_{j-1},x_j])(x_j - x_{j-1})) \\
&\leq \sum_{j \in S_1} \omega(f,[x_{j-1},x_j])(x_j - x_{j-1})) \\
&+ \sum_{k \in S_2} \omega(f,[x_{k-1},x_k])(x_k - x_{k-1})) \\
&= \sum_{i=1}^{\infty} \omega(f,[x_{i-1},x_i])(x_i - x_{i-1})) \\
&< \frac{\epsilon}{N}.
\end{aligned}$$

and so

$$\sum_{j\in S_1} \ell(I_j) = 2\sum_{j\in S_1}(x_j - x_{j-1}) < 2N \cdot \frac{\epsilon}{2N} = \epsilon.$$

Since the choice of $\epsilon > 0$ was arbitrary, it follows that the set D_N has measure zero. □

The final lemma we will prove before proceeding to a proof of the first part of Theorem 1 is that if we take any countable collection of sets of measure zero, then their union will also be a set of measure zero.

Lemma 3. *A countable union of sets of measure zero has measure zero.*

Proof. Let $\epsilon > 0$. Suppose that $\mathcal{G} = \{S_1, S_2, ...\}$ is a countable collection of sets of measure zero. Now, since each of the sets S_i in $\mathcal{G}$ has measure zero, then for each positive integer i, there exists a sequence of intervals $\{I_{i,k}\}$ such that $S_i \subseteq \bigcup_{k=1}^{\infty} I_{i,k}$ and $\sum_{k=1}^{\infty} \ell(I_{i,k}) < \frac{\epsilon}{2^i}$. We note that the collection of the union of all intervals $I_{i,k}$ is a cover for the union of all S_i in $\mathcal{G}$, that is,

$$\bigcup_{i=1}^{\infty} S_i = \bigcup_{i=1}^{\infty}\bigcup_{k=1}^{\infty} I_{i,k}.$$

Now, we will show that $\sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \ell(I_{i,k}) < \epsilon$. Since for each positive integer i, the series $\sum_{k=1}^{\infty} \ell(I_{i,k}) < \frac{\epsilon}{2^i}$, and since $\sum_{n=1}^{\infty} \frac{1}{2^n}$ is a series that converges absolutely to 1, then

$$\sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \ell(I_{i,k}) < \sum_{i=1}^{\infty} \frac{\epsilon}{2^i} = \epsilon.$$

Since the choice of $\epsilon > 0$ was arbitrary, it follows that the collection $\mathcal{G}$ has measure zero. We can conclude that a countable union of sets of measure zero has measure zero. □

We now have all the tools in place to construct a proof of the first part of Theorem 1.

Theorem 3. *Every Riemann integrable function is continuous almost everywhere.*

Proof. Suppose that f is a Riemann integrable function defined on the interval $[a, b]$, and for each positive integer n, let D_n denote the set $\{x \in (a, b) : \omega(f, c) \geq 1/n\}$. Since the sequence $\{\frac{1}{n}\}$ converges to zero, and since, by Lemma 1, the function f is continuous at a point x in $[a, b]$ if and only if $\omega(f, x) = 0$, then we can conclude that $\bigcup_{n=1}^{\infty} D_n$ is the set of all discontinuities of f. Now, by Lemma 2, we know that for each n, the set D_n has measure zero, and by Lemma 3, we know that every countable union of sets of measure zero has measure zero. Thus, the set $\bigcup_{n=1}^{\infty} D_n$ has measure zero. As the set of discontinuities of f has measure zero, it follows that f is continuous almost everywhere. □

The following result completes the first direction of Theorem 1. Although conceptually straightforward, proving from the definition that every Riemann integrable function is bounded becomes somewhat messy. That is, while it seems intuitive that no matter how small we restrict the width of the approximating rectangles, we can always find one rectangle that has an arbitrarily large height (and thus, arbitrarily large area), ensuring that the resulting Riemann sum always differs from the target limit requires a bit of notational finesse. For a sleek proof using alternate conditions for Riemann integrability, see page 171 in [1].

Theorem 4. *Every Riemann integrable function is bounded.*

Proof. Suppose, by way of contradiction, that f is both Riemann integrable and unbounded, and suppose that $\{a_k\}$ is a sequence of points in $[a,b]$ such that $|f(a_k)| > k$. Choose a positive number δ such that for all tagged partitions ${}^tP$ of $[a,b]$ with norms less than δ, we have that $\left|S(f, {}^tP) - \int_a^b f\right| < \frac{1}{2}$. Let n be a positive integer such that $\frac{1}{n} < \delta$, and let $P = \{x_i : 1 \le i \le n\}$ partition $[a,b]$ into n subintervals of length $\frac{b-a}{n}$. Since the sequence $\{x_k\}$ converges to ∞, there exists an index $1 \le j \le n$ such that the interval $[x_{j-1}, x_j]$ contains an infinite number of terms from $\{a_k\}$. Let K be the smallest integer such that $a_K \in [x_{j-1}, x_j]$, and choose a positive integer K' such that $|f(a_{K'})| \ge |f(a_K)| + \frac{n}{b-a}$. For each interval in P, choose tags t_i such that $f(t_i)$ exists, and define tagged partitions ${}^tP_1$ and ${}^tP_2$ as follows:

$$\begin{aligned} {}^tP_1 =& \{(t_i, [x_{i-1}, x_i]) : i \neq j\} \bigcup \{(x'_K, [x_{j-1}, x_j])\}, \\ {}^tP_2 =& \{(t_i, [x_{i-1}, x_i]) : i \neq j\} \bigcup \{(x_K, [x_{j-1}, x_j])\}. \end{aligned}$$

We observe that both ${}^tP_1$ and ${}^tP_2$ have norms less than δ, and thus, by the triangle inequality, it follows that $|S(f, {}^tP_1) - S(f, {}^tP_2)| < 1$. But since both tagged partitions are equal except on the interval $[x_{j-1}, x_j]$, then

$$\begin{aligned} |S(f, {}^tP_1) - S(f, {}^tP_2)| =& |f(x'_K)(x_j - x_{j-1}) - f(x_K)(x_j - x_{j-1})| \\ \ge& \big||f(x'_K)| - |f(x_K)|\big| \frac{b-a}{n} \\ \ge& \left(|f(a_K)| + \frac{n}{b-a} - |f(x_K)|\right) \frac{b-a}{n} \\ =& 1. \end{aligned}$$

It follows that f must be bounded on $[a,b]$. □

In order to demonstrate that every bounded and continuous almost everywhere function is Riemann integrable, we will need to introduce a further concept: the δ-fine tagged partition. Fortunately, not only does this concept give rise to a sleek proof of Lebesgue's Criterion, but it also stands as the very foundation of the theory of Henstock integration!

2 The δ-fine Tagged Partition

In many ways, the Henstock integral can be viewed as a generalization of the Riemann integral. Where the latter integral considers tagged partitions of an interval with subintervals of a length less than a fixed constant, the former allows the maximum length of the subintervals to vary over the interval. Since the best approximation for the area under a function occurs when we force the lengths of the subintervals to be small when the graph of that function is steep, we are interested in considering tagged partitions where the lengths of the subintervals are dependent on the properties of the function at the tagged points. Informally speaking, each point in the interval is associated with a subinterval of length less than a particular number; given a certain collection of tagged points, the lengths of the intervals of a tagged partition becomes a function of those particular tags. By altering the way we look at tagged partitions in this manner, we are able to integrate a large number of functions that were previously not integrable.

Before we look at a definition of the Henstock integral, it is first necessary to present some preliminary definitions and results. As was implied above, Henstock integration requires that we consider tagged partitions with subintervals of lengths determined by the tags of that partition. We now formalize this idea with the following definition:

Definition 8. *Let $[a, b]$ be a closed, bounded interval and let $\delta(x) : [a, b] \to \mathbb{R}$ be a positive function (that is, $\delta(x) > 0$ for all x in $[a, b]$). A **δ-fine tagged partition** $\{(t_i, [x_{i-1}, x_i]) : 1 \leq i \leq n\}$ of $[a, b]$ is a tagged partition of $[a, b]$ that satisfies $[x_{i-1}, x_i] \subseteq (t_i - \delta(t_i), t_i + \delta(t_i))$ for each $1 \leq i \leq n$.*

When considering tagged partitions with norms less than a positive constant δ, it is easy to see that for each $\delta > 0$, there exists some tagged partition with norm less than δ. We must simply divide the interval into subintervals of length δ, and then choose a tag in each subinterval. In the case above, however, it is not obvious that for each positive function δ, there exists a δ-fine tagged partition. Since, in essence, we are first choosing tags for an interval, and then partitioning the interval based on which tags are selected, it is not at all obvious that for each positive function δ, there exists a δ-fine tagged partition. We will next present a proof of this result using the completeness property of the set of real numbers.

Theorem 5. *For each positive function δ defined on a closed, bounded interval $[a, b]$, there exists a δ-fine tagged partition of $[a, b]$.*

Proof. Let S be the set of all points x in $(a, b]$ such that there exists a δ-fine tagged partition of $[a, x]$. We show that there exists a δ-fine tagged partition of $[a, b]$ by showing that the supremum of the set S exists, that it is an element of S, and that it is equal to b.

Let x be a real number that satisfies $a < x < a + \delta(a)$ and $x < b$. If we consider the interval $[a, x]$ to be a one-interval partition of itself, and if we choose a to be the tag of this interval, then we see that $[a, x]$ is a δ-fine tagged partition of $[a, x]$, since $[a, x] \subseteq (a - \delta(a), a + \delta(a))$. Thus, the point x is in S, and so S is nonempty. Since every nonempty set of real numbers bounded above has a supremum, and as S is bounded above by b, let β denote the supremum of the set S.

Since β is the least upper bound of S, then $\beta \leq b$, and so the function δ is defined at the point β. As β is the supremum of S, there exists a point $y \in S$ such that $\beta - \delta(\beta) < y < \beta$. Let ${}^tP_1$ be a δ-fine tagged partition of $[a, y]$. Then the tagged partition formed by the union of ${}^tP_1$ and the tagged partition

$\{(\beta, [y, \beta])\}$ is δ-fine, since $[y, \beta] \subseteq (\beta - \delta(\beta), \beta + \delta(\beta))$. Thus, β is an element of the set S.

Now, in order to demonstrate that $\beta = b$, we will suppose by way of contradiction that $\beta < b$. Since $\beta < b$, there exists a point $z \in (\beta, b)$ such that $z < \beta + \delta(\beta)$. Let ${}^tP_2$ be a δ-fine tagged partition of $[a, \beta]$. We see that the tagged partition formed by the union of ${}^tP_2$ and the tagged partition $\{(\beta, [\beta, z])\}$ is δ-fine, since $[\beta, z] \subseteq (\beta - \delta(\beta), \beta + \delta(\beta))$. Therefore, $z \in S$. But this contradicts our assumption that β was the supremum of S. It follows that $\beta = b$.

Finally, since b is an element of S, we can conclude that there exists a δ-fine tagged partition of $[a, b]$. $\square$

It must be remarked that this proof bears striking similarity to the standard proof that every closed and bounded interval is compact; both make fundamental use of the completeness of the real numbers. It turns out that the two results are actually equivalent, and consequentially, the fact that every positive function has a corresponding tagged partition can be used to prove many of the standard results in real analysis that rely on either the Heine-Borel theorem or the Completeness Property. As will be seen shortly, the existence of δ-fine tagged partitions stands as the foundation for the theory of Henstock integration. And while the use of the positive δ function affords great flexibility in terms of the ability to integrate real-valued functions, at the same time, it of necessity ties the process of Henstock integration to the set of real numbers, and as a result, we find that the Henstock integral cannot be easily generalized to arbitrary spaces and metrics.

Before we investigate the Henstock integral, we pause to complete the second direction of the proof of Lebesgue's Criterion.

Theorem 6. *If f is bounded and continuous almost everywhere on $[a, b]$, then f is Riemann integrable on $[a, b]$.*

Proof. Let M be a bound for f on $[a, b]$, let D be the set of all points in $[a, b]$ at which f is not continuous, and let $\epsilon > 0$. Since D has measure zero, there exists a sequence of open intervals $\{I_k\}$ such that $D \subseteq \bigcup_{k=1}^{\infty} I_k$ and $\sum_{k=1}^{\infty} \ell(I_k) < \dfrac{\epsilon}{4M}$. Define a function $\delta : [a, b] \to \mathbb{R}^+$ such that $(x - \delta(x), x + \delta(x)) \subseteq I_k$, if $x \in D$ and $x \in I_k$, and such that $|f(x) - f(t)| < \frac{\epsilon}{4(b-a)}$ for all $t \in (x - \delta(x), x + \delta(x))$, otherwise. Let ${}^tP$ be a δ-fine tagged partition of $[a, b]$, where σ_1 denotes the set of all indices of tags such that $t_i \in D$ and where σ_2 denotes the set of all other indices. Then

$$\begin{aligned}
\sum_{i=1}^{n} \omega(f,[x_{i-1},x_i])(x_i - x_{i-1}) &= \sum_{i\in\sigma_1} \omega(f,[x_{i-1},x_i])(x_i - x_{i-1}) \\
&\quad + \sum_{i\in\sigma_2} \omega(f,[x_{i-1},x_i])(x_i - x_{i-1}) \\
&\leq \sum_{i\in\sigma_1} 2M(2\delta(t_i)) + \sum_{i\in\sigma_2} \frac{2\epsilon}{4(b-a)}(x_i - x_{i-1}) \\
&\leq 2M \sum_{k=1}^{\infty} \ell(I_k) + \frac{\epsilon}{2(b-a)}(b-a) \\
&< 2M\frac{\epsilon}{4M} + \frac{\epsilon}{2} \\
&< \epsilon.
\end{aligned}$$

By Riemann's Criterion, it follows that f is Riemann integrable on $[a, b]$. □

Although we have approached it in a roundabout manner, we have now completed the proof of Lebesgue's Criterion, and shown that a function is Riemann integrable if and only if it is bounded and continuous almost everywhere.

3 The Henstock Integral

Although it may not at first be clear, the use of the δ-fine tagged partition in the definition of the Henstock integral has far-reaching consequences.

Definition 9. *A function $f : [a, b] \to \mathbb{R}$ is Henstock integrable on $[a, b]$ if there exists a number L such that for each $\epsilon > 0$ there exists a positive function δ defined on $[a, b]$ such that $|S(f,{}^t P) - L| < \epsilon$ for all δ-fine tagged partitions ${}^t P$ of $[a, b]$. We call L the Henstock integral of f on $[a, b]$, and represent it as $\int_a^b f$.*

The first thing we note about the definition of the Henstock integral is its similarity to the Riemann integral. In fact, the only difference between the two is that, in the definition of the Henstock integral, the condition that the norms of the tagged partitions must be less than a positive constant δ has been replaced by the condition that the tagged partitions must be δ-fine, where δ is a positive function. Moreover, we notice that if we restrict ourselves to the case where δ is a constant, positive function, the definition of the Henstock integral immediately reduces to the definition of the Riemann integral. In fact, if δ has a positive infimum, we can simply define a new, positive and constant function δ' such that δ' is strictly less than δ, in which case the definition once again reduces to the definition of the Riemann integral.

The following theorem, which guarantees that the limit of the Henstock integral is unique, shows that we are justified in using the definite article when talking about the Henstock integral of a function. The structure of the proof is similar to that of the proof that the limit of a convergent sequence is unique. This is not surprising, since the Henstock integral can be seen as a sort of limit of a sequence of Riemann sums.

Theorem 7. *If f is a Henstock integrable function on an interval $[a, b]$, then the Henstock integral of f on $[a, b]$ is unique.*

Proof. Suppose that L_1 and L_2 in $\mathbb{R}$ are both Henstock integrals of f on $[a, b]$. Let $\epsilon > 0$. Then there exist two positive functions δ_1 and δ_2 defined on $[a, b]$ such that $|S(f,^t P) - L_1| < \frac{\epsilon}{2}$ for all δ_1-fine tagged partitions ^tP of $[a, b]$, and such that $|S(f,^t P) - L_2| < \frac{\epsilon}{2}$ for all δ_2-fine tagged partitions ^tP of $[a, b]$. Define a function $\delta : [a, b] \to \mathbb{R}^+$ by $\delta(x) = \min\{\delta_1(x), \delta_2(x)\}$. Then, by the triangle inequality, for all δ-fine tagged partitions of $[a, b]$,

$$|L_1 - L_2| \leq \left|S(f,^t P) - L_1\right| + \left|S(f,^t P) - L_2\right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

This implies that $L_1 = L_2$. It follows that the Henstock integral of f on $[a, b]$ is unique. □

With the definition in hand, one of the first tasks we must do is verify the Henstock integral is a linear operator, and that many of the useful, algebraic properties of the Riemann integral are also valid for the Henstock integral. Since the proofs of each part of the following theorem are straightforward applications of the definition of the Henstock integral, only the first two will be proved in full. Proofs of the third and fourth results will appear in greater generality later in this paper.

Theorem 8. *Suppose that f and g are Henstock integrable functions defined on $[a, b]$ and that k is a constant. Then*

1. *kf is Henstock integrable on $[a, b]$ and $\int_a^b kf = k \int_a^b f$;*
2. *$f + g$ is Henstock integrable on $[a, b]$ and $\int_a^b (f + g) = \int_a^b f + \int_a^b g$;*
3. *$\int_a^b f \leq \int_a^b g$ if $f(x) \leq g(x)$ for all $x \in [a, b]$;*
4. *$|\int_a^b f| \leq M(b - a)$ if $|f(x)| \leq M$ for all $x \in [a, b]$.*

Proof. Let $\epsilon > 0$. Then there exists a real number $\int_a^b f$ and a positive function δ such that $|S(f,^t P) - \int_a^b f| < \frac{\epsilon}{|k|}$ for all δ-fine tagged partition ^tP of $[a, b]$. Thus, for all δ-fine tagged partitions of $[a, b]$,

$$\begin{aligned}\left|\sum_{i=1}^{n} kf(t_i)(x_i - x_{i-1}) - k\int_a^b f\right| &= \left|k\sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) - k\int_a^b f\right| \\ &= |k|\left|\sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) - \int_a^b f\right| \\ &< |k| \cdot \frac{\epsilon}{|k|} = \epsilon.\end{aligned}$$

It follows that kf is Henstock integrable on $[a, b]$ and $\int_a^b kf = k\int_a^b f$.

Now, if g is also Henstock integrable on $[a, b]$, with $\epsilon > 0$, then there exist real numbers $\int_a^b f$ and $\int_a^b g$ and positive functions δ_1 and δ_2 such that $|S(f,^t P_1) - \int_a^b f| < \frac{\epsilon}{2}$ for all δ_1-fine tagged partitions $^tP_1$ of $[a, b]$ and $|S(g,^t P_2) - \int_a^b g| < \frac{\epsilon}{2}$

for all δ_2-fine tagged partitions ${}^tP_2$ of $[a,b]$. Define a function $\delta : [a,b] \to \mathbb{R}$ by $\delta = \min\{\delta_1(x), \delta_2(x)\}$ for each $x \in [a,b]$. Then, for each δ-fine tagged partition ${}^tP$, we have that

$$\begin{aligned}
&\left|\sum_{i=1}^{n}((f+g)(t_i))(x_i - x_{i-1}) - \left(\int_a^b f + \int_a^b g\right)\right| \\
&= \left|\sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) - \int_a^b f + \sum_{i=1}^{n} g(t_i)(x_i - x_{i-1}) - \int_a^b g\right| \\
&\leq \left|\sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) - \int_a^b f\right| + \left|\sum_{i=1}^{n} g(t_i)(x_i - x_{i-1}) - \int_a^b g\right| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Thus, $f+g$ is Henstock integrable on $[a,b]$ and $\int_a^b (f+g) = \int_a^b f + \int_a^b g$. □

Using the first two parts of the preceding theorem and the principle of mathematical induction, it is clear that any linear combination with real coefficients of a finite number of Henstock integrable functions is Henstock integrable. Unlike the Riemann integral, it is not the case that the product of any two Henstock integrals is necessarily Henstock integrable. Constructing a concrete counterexample, however, makes use of the theory of Lebesgue integration, and so is outside the scope of this paper.

When working with δ-fine tagged partitions, it is often easiest to work partitions with tags that lie on the endpoints of the partitioned subinterval, or with partitions with tags that occur only once. The next lemma guarantees that, without a loss of generality, we can always assume that a δ-fine tagged partitions fulfills either of these conditions.

Lemma 4. *Let δ be a positive function defined on $[a,b]$ and let $(t,[a,b])$ be tagged interval. Then $(t,[a,b])$ is a δ-fine tagged partition of $[a,b]$ if and only if $\{(t,[a,t]),(t,[t,b])\}$ is a δ-fine tagged partition of $[a,b]$. Further, $f(t)(b-a) = f(t)(t-a) + f(t)(b-t)$.*

The first statement amounts to saying that, given any tagged partition, if we combine any subintervals that share the same tag (so that each tag occurs only once), or split apart any subinterval at a tagged point (ensuring that tags occur only at endpoints), then this modified tagged partition will also be δ-fine. The second statement, though not particularly abstruse, demonstrates that the Riemann sum on this modified partition is equal to the Riemann sum on the original. Although the proof is presented only for the case of partitions containing one or two subintervals, the result can easily be generalized to any arbitrary tagged partition.

Proof. First, suppose that $(t,[a,b])$ is a δ-fine tagged partition of $[a,b]$. Since $[a,t] \subseteq [a,b]$ and $[t,b] \subseteq [a,b]$, and since $[a,b] \subseteq (t-\delta(t), t+\delta(t))$, it follows that $\{(t,[a,t]),(t,[t,b])\}$ is a δ-fine tagged partition of $[a,b]$.

Now, suppose that $\{(t,[a,t]),(t,[t,b])\}$ is a δ-fine tagged partition of $[a,b]$. Then $[a,t] \subseteq (t-\delta(t), t+\delta(t))$ and $[t,b] \subseteq (t-\delta(t), t+\delta(t))$. Thus, $[a,t] \cup [t,b] = [a,b] \subseteq (t-\delta(t), t+\delta(t))$. It follows that $(t,[a,b])$ is a δ-fine tagged partition of $[a,b]$.

Finally, we note that

$$f(t)(b-a) = f(t)b - f(t)a + f(t)t - f(t)t = f(t)(t-a) + f(t)(b-t).$$

This completes the proof. $\square$

We recall that the characteristic function $\chi_{\mathbb{Q}} : [a,b] \to \{0,1\}$ on the rationals over the close interval $[a,b]$ is defined by

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q} \\ 0, & \text{if } x \notin \mathbb{Q}. \end{cases}$$

The characteristic function is one of the canonical examples of functions that are bounded, but not Riemann integrable. Since every Riemann integrable function is continuous almost everywhere, and since $\chi_{\mathbb{Q}}$ is clearly discontinuous at all points in $[a,b]$, it follows that $\chi_{\mathbb{Q}}$ is not Riemann integrable. Importantly, this fact can be used to demonstrate that not every pointwise limit of sequences of Riemann integrable functions is Riemann integrable. This is one deficiency in the Riemann integral that is rectified by the Henstock integral. The following theorem guarantees not only that the characteristic function on the rationals is Henstock integrable, but that any function that differs from the zero function at only countable number of points is Henstock integrable.

Theorem 9. *Let f be a function defined on $[a,b]$. If $f = 0$ except at a countable number of points in $[a,b]$, then f is Henstock integrable on $[a,b]$ and $\int_a^b f = 0$.*

Proof. Let $\epsilon > 0$, let $S = \{x \in [a,b] : f(x) \neq 0\}$, and let $\{a_n : n \in \mathbb{Z}^+\}$ be a set listing the elements of S. Define the positive function $\delta : [a,b] \to \mathbb{R}^+$ by

$$\delta(x) = \begin{cases} \epsilon/(2^{(n+1)}\,|f(a_n)|), & \text{if } x = a_n \\ 1, & \text{otherwise.} \end{cases}$$

Suppose that ${}^tP$ is a δ-fine tagged partition of $[a,b]$ and suppose that each tag of ${}^tP$ occurs only once. Let σ be the set of indices i of tags such that $t_i \in \{a_n : n \in \mathbb{Z}^+\}$. For each $i \in \sigma$, let n_i be the positive integer that satisfies $t_i = a_{n_i}$. Then

$$\begin{aligned} \left|S(f, {}^tP)\right| &= \left|\sum_{i\in\sigma} f(t_i)(x_i - x_{i-1}) + \sum_{i\notin\sigma} f(t_i)(x_i - x_{i-1})\right| \\ &\leq \sum_{i\in\sigma} |f(t_i)|\,(x_i - x_{i-1}) \\ &< \sum_{i\in\sigma} 2\,|f(a_{n_i})|\,\delta(a_{n_i}) \\ &= \sum_{i\in\sigma} \frac{\epsilon}{2^n} \\ &< \epsilon. \end{aligned}$$

Since the choice of $\epsilon > 0$ was arbitrary, it follows that f is Henstock integrable on $[a,b]$ and that $\int_a^b f = 0$. $\square$

As it turns out, the preceding theorem can be further generalized to show that every function that differs from the zero function except on a set of measure zero is Henstock integrable. However, we note that while this implies that the previous result is actually a consequence of the next theorem, the structure of the two proofs differs considerably, and since the latter makes use of some basic definitions from measure theory, it is appropriate to include proofs of both results.

Theorem 10. *Let f be a function defined on $[a, b]$. If $f = 0$ almost everywhere on $[a, b]$, then f is Henstock integrable on $[a, b]$ and $\int_a^b f = 0$.*

Proof. Let $\epsilon > 0$, and let E be the set $\{x \in [a, b] : f(x) \neq 0\}$. For each positive integer n, let E_n denote the set $\{x \in [a, b] : n - 1 \leq |f(x)| < n\}$. Since $E_n \subseteq E$, and since E is a set of measure zero, it follows that each E_n is also a set of measure zero. Now, for each n, let $\{I_k^n\}_{k=1}^{\infty}$ be a sequence of open intervals such that $E_n \subseteq \displaystyle\bigcup_{k=1}^{\infty} I_k^n$ and $\displaystyle\sum_{k=1}^{\infty} \ell(I_k^n) < \epsilon/n2^{n+2}$. For each $x \in E_n$, choose an open interval in the set $\{I_k^n\}$ such that x is an element of this interval. Call this open interval I_x. Let $\delta : [a, b] \to \mathbb{R}^+$ be the positive function defined by

$$\delta(x) = \begin{cases} 1, & \text{if } x \notin E \\ \ell(I_x)/2, & \text{if } x \in E_n. \end{cases}$$

Suppose that ${}^tP$ is a δ-fine tagged partition of $[a, b]$ and suppose that each tag of ${}^tP$ occurs only once. For each positive n, let σ_n denote the set of indices i of tags such that $t_i \in E_n$, and let $\sigma_0 = \{i : t_i \notin E\}$. We note that since for each n, the set $\{(t_i, [x_{i-1}, x_i]) : i \in \sigma_n\}$ is a collection of nonoverlapping, tagged intervals, and since each interval has length less than $\ell(I_{t_i})$, it follows that $\displaystyle\sum_{i \in \sigma_n} (x_i - x_{i-1}) < \sum_{k=1}^{\infty} 2\ell(I_k^n)$. Now, we have that

$$\begin{aligned}
\left|S(f, {}^tP)\right| &= \left|\sum_{n=0}^{\infty} \sum_{i \in \sigma_n} f(t_i)(x_i - x_{i-1})\right| \\
&\leq \sum_{n=1}^{\infty} \sum_{i \in \sigma_n} |f(t_i)| (x_i - x_{i-1}) \\
&< \sum_{n=1}^{\infty} \sum_{i \in \sigma_n} n(x_i - x_{i-1}) \\
&< \sum_{n=1}^{\infty} n \sum_{k=1}^{\infty} 2\ell(I_k^n) \\
&< \sum_{n=1}^{\infty} 2n \frac{\epsilon}{n2^{n+1}} \\
&= \epsilon.
\end{aligned}$$

Since the choice of $\epsilon > 0$ was arbitrary, we can conclude that f is Henstock integrable on $[a, b]$ and that $\int_a^b f = 0$. □

The following two theorems are consequences of the previous theorem. Both generalize algebraic properties of the Henstock integral that had appeared in Theorem 8, and require only that functions possess certain properties almost everywhere.

Theorem 11. *Let f be a Henstock integrable function defined on $[a, b]$. If g is a function defined on $[a, b]$ and $f = g$ almost everywhere, then g is Henstock integrable on $[a, b]$ and $\int_a^b f = \int_a^b g$*

Proof. Define a function $h : [a, b] \to \mathbb{R}$ by $h(x) = f(x) - g(x)$. Since $f = g$ almost everywhere on $[a, b]$, then $h = 0$ almost everywhere on $[a, b]$. By Theorem (10), we know that h is Henstock integrable and that $\int_a^b h = 0$. Further, by a previous result, we know that the sum of two Henstock integrable functions is Henstock integrable. Thus, the function $g = f + h$ is Henstock integrable on $[a, b]$, and $\int_a^b g = \int_a^b f + \int_a^b h = \int_a^b f$. □

Theorem 12. *Let f and g be Henstock integrable functions defined on $[a, b]$. If $f \leq g$ almost everywhere on $[a, b]$, then $\int_a^b f \leq \int_a^b g$.*

Proof. Define functions $f^* : [a, b] \to \mathbb{R}$ and $g^* : [a, b] \to \mathbb{R}$ by

$$f^*(x) = \begin{cases} f(x), & \text{for all } x \text{ such that } f(x) \leq g(x) \\ 0, & \text{for all } x \text{ such that } f(x) > g(x); \end{cases}$$

$$g^*(x) = \begin{cases} g(x), & \text{for all } x \text{ such that } f(x) \leq g(x) \\ 0, & \text{for all } x \text{ such that } f(x) > g(x). \end{cases}$$

Since $f = f^*$ and $g = g^*$ almost everywhere on $[a, b]$, it follows from Theorem (11) that $\int_a^b f = \int_a^b f^*$ and $\int_a^b g = \int_a^b g^*$. Now, since $g^* \geq f^*$ on $[a, b]$, then $g^* - f^* \geq 0$, and by a previous result, we know that $\int_a^b (g^* - f^*) \geq 0$. Therefore, using the linearity properties of the Henstock integral, we have that

$$\int_a^b g - \int_a^b f = \int_a^b g^* - \int_a^b f^* = \int_a^b (g^* - f^*) \geq 0.$$

We can conclude that $\int_a^b g \geq \int_a^b f$. □

From a geometric perspective, the following result seems rather intuitive; the area under a curve should be equal to the sum of the area under one part of the curve and the area under the other part of the curve. A proof of this result using the definition of the Henstock integral, however, requires slightly more finesse. The crucial step in the following proof comes about by defining the δ-function in such a way as to force a particular point to be a tag for every δ-fine tagged partition. It is not difficult to see that the same method can be used to force any finite number of points in an interval to be tags for every δ-fine tagged partition of that interval.

Theorem 13. *Let $f : [a, b] \to \mathbb{R}$ and let $c \in (a, b)$. If f is Henstock integrable on the intervals $[a, c]$ and $[c, b]$, then f is Henstock integrable on $[a, b]$ and $\int_a^b f = \int_a^c f + \int_c^b f$.*

Proof. Let $\epsilon > 0$. Since f is Henstock integrable on $[a, c]$ and $[c, b]$, there exist functions $\delta_1 : [a, c] \to \mathbb{R}^+$ and $\delta_2 : [c, b] \to \mathbb{R}^+$ such that $|S(f, {}^tP_1) - \int_a^c f| < \frac{\epsilon}{2}$ for all δ_1-fine tagged partitions ${}^tP_1$ of $[a, c]$ and $|S(f, {}^tP_2) - \int_c^b f| < \frac{\epsilon}{2}$ for all δ_2-fine tagged partitions ${}^tP_2$ of $[c, b]$. Define a positive function δ on $[a, b]$ by

$$\delta(x) = \begin{cases} \min\{\delta_1(x), c - x\}, & x \in [a, c) \\ \min\{\delta_1(c), \delta_2(c)\}, & x = c \\ \min\{\delta_2(x), x - c\}, & x \in (c, b]. \end{cases}$$

Let ${}^tP = \{(t_i, [x_{i-1}, x_i]) : 1 \leq i \leq n\}$ be a δ-fine tagged partition of $[a, b]$. We note that for any tag less than c, the right endpoint of the tag's interval is less than c, and for any tag greater than c, the left endpoint of the tag's interval is greater than c. Thus, c must be a tag of ${}^tP$. Now, let N be a positive integer such that $t_i < c$ for all indices $i \leq N$ and $t_i > c$ for all indices $i > N$. Then ${}^tP_1 = \{(t_i, [x_{i-1}), x_i] : 1 \leq i \leq N\} \cup (c, [x_N, c])$ is a δ_1-fine tagged partition of $[a, c]$ and ${}^tP_2 = \{(t_i, [x_{i-1}, x_i] : N + 1 < i \leq n\} \cup (c, [c, x_{N+1}])$ is a δ_2-fine tagged partition of $[c, b]$. Let σ_1 denote the set of indices of tags of ${}^tP_1$ and let σ_2 denote the set of indices of tags of ${}^tP_2$. Thus,

$$\begin{aligned} \left| S(f, {}^tP) - \left(\int_a^c f + \int_c^b f \right) \right| &\leq \left| \sum_{i \in \sigma_1} f(t_i)(x_i - x_{i-1}) - \int_a^c f \right| \\ &\quad + \left| \sum_{i \in \sigma_2} f(t_i)(x_i - x_{i-1}) - \int_a^c f \right| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon. \end{aligned}$$

It follows that f is Henstock integrable on $[a, b]$ and $\int_a^b f = \int_a^c f + \int_c^b f$. □

Given the preceding result, the natural next step is to demonstrate that if a function is Henstock integrable on some interval $[a, b]$, then it must also be Henstock integrable on every subinterval of $[a, b]$. To do this, however, requires us to introduce a Cauchy criterion for Henstock integrability.

We recall that a **Cauchy sequence** of real numbers is a sequence of points $\{x_n\}$ such that for each $\epsilon > 0$, there exists a positive integer N such that $|x_m - x_n| < \epsilon$ for all $m, n \geq N$. Loosely speaking, this means that the terms of a Cauchy sequence are all eventually close to each other. It is relatively straightforward to verify that every convergent sequence is also a Cauchy sequence. Conversely, since the real numbers are complete, every Cauchy sequence of real numbers converges to a real number. This result can be demonstrated using the property that every set of real numbers that is bounded above has a supremum.

As both the Riemann integral and the Henstock integral can be seen as types of limits of sequences of real numbers, Cauchy sequences can be used to provide an alternate condition for establishing integrability. This condition is often referred to as the Cauchy criterion for Riemann or Henstock integrability, and is equivalent to the definition of the respective integral. One advantage of the Cauchy criterion is that it does not require that we first have a candidate for value of the integral before we verify that the integral exists. For this reason, the criterion is particularly useful as a theoretical tool for verifying that certain classes of functions are integrable. However, since the variability between two δ-fine tagged partitions is often too cumbersome to manipulate effectively, the Cauchy criterion has limited application when attempting to prove that particular functions are integrable.

The following proof makes use of the fact that if δ_1 and δ_2 are two positive functions defined on a closed, bounded interval and if δ_2 is less than or equal to δ_1 at all points on that interval, then any δ_2-fine tagged partition is also a δ_1-fine tagged partition.

Theorem 14 (Cauchy Criterion for Henstock Integrals)**.** *Let f be a function defined on the interval $[a,b]$. Then f is Henstock integrable on $[a,b]$ if and only if for each $\epsilon > 0$ there exists a function $\delta : [a,b] \to \mathbb{R}^+$ such that $|S(f, {}^tP_1) - S(f, {}^tP_2)| < \epsilon$ for all δ-fine tagged partitions ${}^tP_1$ and ${}^tP_2$ of $[a,b]$.*

Proof. Suppose first that f is Henstock integrable on $[a,b]$. Let $\epsilon > 0$, and choose $\delta : [a,b] \to \mathbb{R}^+$ such that $|S(f, {}^tP) - \int_a^b| < \frac{\epsilon}{2}$ for all δ-fine tagged partitions of $[a,b]$. Now, let ${}^tP_1$ and ${}^tP_2$ be two δ-fine tagged partitions of $[a,b]$. Then

$$|S(f, {}^tP_1) - S(f, {}^tP_2)| \leq \left| S(f, {}^tP_1) - \int_a^b \right| + \left| S(f, {}^tP_2) - \int_a^b \right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

It follows that $|S(f, {}^tP_1) - S(f, {}^tP_2)| < \epsilon$ for all δ-fine tagged partitions ${}^tP_1$ and ${}^tP_2$ of $[a,b]$.

We will now prove the converse of this statement. Suppose that for each $\epsilon > 0$, there exists a positive function δ defined on $[a,b]$ such that $|S(f, {}^tP_1) - S(f, {}^tP_2)| < \epsilon$ for all δ-fine tagged partitions ${}^tP_1$ and ${}^tP_2$ of $[a,b]$. For each positive integer n, choose a positive function δ_n such that $|S(f, {}^tP_1) - S(f, {}^tP_2)| < \frac{1}{n}$ for all δ-fine tagged partitions ${}^tP_1$ and ${}^tP_2$ of $[a,b]$. Without loss of generality, we may assume that the sequence $\{\delta_n\}$ is decreasing (that is, for each n, $\delta_n(x) \geq \delta_{n+1}(x)$ for all $x \in [a,b]$). Now, for each n, let ${}^tP_n$ be a δ-fine tagged partition of $[a,b]$. If K is a positive integer, and m, n are positive integers greater than or equal to K, then the tagged partitions ${}^tP_m$ and ${}^tP_n$ are δ_K-fine tagged partitions, since $\{\delta_n\}$ is decreasing. It follows that

$$|S(f, {}^tP_n) - S(f, {}^tP_m)| < \frac{1}{K}.$$

We can conclude that $\{S(f, {}^tP_n)\}$ is a Cauchy sequence. Since every Cauchy sequence of real numbers converges, define $L \in \mathbb{R}$ to be the limit of this sequence. Let $\epsilon > 0$. Since $\{S(f, {}^tP_n\}$ converges to L, there exists a positive number N such that $\frac{1}{N} < \frac{\epsilon}{2}$ and $|S(f, {}^tP_n) - L| < \frac{\epsilon}{2}$ for all $n \geq N$. Define a positive function δ on $[a,b]$ by $\delta(x) = \delta_N(x)$, and suppose that ${}^tP$ is a δ-fine tagged partition of $[a,b]$. Then

$$|S(f, {}^tP) - L| \leq |S(f, {}^tP) - S(f, {}^tP_N)| + |S(f, {}^tP_N) - L| < \frac{1}{N} + \frac{\epsilon}{2} < \epsilon.$$

It follows that f is Henstock integrable on $[a,b]$ and that $\int_a^b = L$. □

Theorem 15. *If f is Henstock integrable on $[a,b]$, then f is Henstock integrable on each subinterval of $[a,b]$.*

Suppose that $[c, d]$ is a subinterval of $[a, b]$, and let $\epsilon > 0$. Since f is Henstock integrable on $[a, b]$, there exists a positive function δ defined on $[a, b]$ such that $\left|S(f, {}^tP) - \int_a^b\right| < \frac{\epsilon}{2}$ for all δ-fine tagged partitions of $[a, b]$. Let ${}^tP_a$ be a tagged partition of the interval $[a, c]$ such that $[x_{i-1}, x_i] \subseteq (t_i - \delta(t_i), t_i + \delta(t_i)$ and let ${}^tP_b$ be a tagged partition of the interval $[d, b]$ such that $[x_{i-1}, x_i] \subseteq (t_i - \delta(t_i), t_i + \delta(t_i)$. Assume that, for each partition, each tag occurs only once. Let ${}^tP_1$ and ${}^tP_2$ be two δ-fine tagged partition of $[c, d]$, and suppose that, in each partition, each tag occurs only once. We note that both ${}^tP_1' = {}^tP_a \cup {}^tP_1 \cup {}^tP_b$ and ${}^tP_2' = {}^tP_a \cup {}^tP_2 \cup {}^tP_b$ are δ_1-fine tagged partitions of $[a, b]$. Now,

$$\begin{aligned}
|S(f, {}^tP_1) - S(f, {}^tP_2)| &\leq \left|S(f, {}^tP_1) + S(f, {}^tP_a) + S(f, {}^tP_b) - \int_a^b f\right| \\
&\quad + \left|S(f, {}^tP_a) + S(f, {}^tP_b) + S(f, {}^tP_2) - \int_a^b f\right| \\
&= \left|S(f, {}^tP_1') - \int_a^b\right| + \left|S(f, {}^tP_2') - \int_a^b\right| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

By Theorem 14, it follows that f is Henstock integrable on $[c, d]$.

4 Derivatives

One of the original motivations for the development of the Henstock integral was to address some of the theoretical deficiencies of the Riemann integral. Given the fundamental, inverse relationship between differentiation and integration that underlies the the theory of modern calculus, it stands to reason that we ought to be able to differentiate any integrated function, or integrate any differentiated function, and recover the original function. Unfortunately, the Riemann integral does not guarantee that this latter process will always work. It is relatively painless to come up with examples of functions that are continuous on a closed interval, and differentiable on the open interval, that have unbounded derivatives. The function given by $\sqrt{x}$ is one that springs immediately to mind. In this case, we can get around this problem by introducing the improper Riemann integral. However, as the following example demonstrates, there exist functions that are continuous and differentiable on every closed, bounded interval that nevertheless have derivatives that are not Riemann integrable.

Consider the function $F : \mathbb{R} \to \mathbb{R}$ defined by

$$F(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right), & \text{for } x \neq 0 \\ 0, & \text{for } x = 0 \end{cases}$$

We note that F is continuous and differentiable at every nonzero point in $\mathbb{R}$, since it is the product of a composition of continuous and differentiable functions. Further, since the limit of F at $x = 0$ is 0, and since $F(0) = 0$, it follows that F is continuous at the 0 as well. Finally, we see that as

$$\lim_{x\to 0} \frac{F(x) - F(0)}{x - 0} = \lim_{x\to 0} \frac{F(x)}{x} = \lim_{x\to 0} x \sin\left(\frac{1}{x^2}\right) = 0,$$

it follows that F is differentiable at 0. Using elementary rules of differentiation, we have that

$$F'(x) = \begin{cases} 2x \sin\left(\frac{1}{x^2}\right) - \frac{2}{x}\cos\left(\frac{1}{x^2}\right), & \text{for } x \neq 0 \\ 0, & \text{for } x = 0 \end{cases}$$

Let $b \in \mathbb{R}^+$. To show that F' is not Riemann integrable on $[0, b]$, we will show that F' is unbounded on $[0, b]$. Let n be an arbitrary positive integer, and choose a positive integer N such that $n \leq \sqrt{2\pi N}$ and $0 < (\sqrt{2\pi N})^{-1} < b$ (the Archimedean property of the real numbers guarantees that this is always possible). Then

$$\begin{aligned} \left|F'(\sqrt{2\pi N}^{-1})\right| &= \left|\frac{2}{\sqrt{2\pi N}} \sin\left(\frac{1}{\frac{1}{\sqrt{2\pi N}}^2}\right) - \frac{2}{\frac{1}{\sqrt{2\pi N}}} \cos\left(\frac{1}{\frac{1}{\sqrt{2\pi N}}^2}\right)\right| \\ &= \left|\frac{2}{\sqrt{2\pi N}} \sin(2\pi N) - 2\sqrt{2\pi N} \cos(2\pi N)\right| \\ &= 2\sqrt{2\pi N} \\ &\geq n \end{aligned}$$

Since the choice of $n > 0$ was arbitrary, it follows that F' is unbounded on $[0, b]$, and since every Riemann integrable function is bounded, we can conclude that F' is not Riemann integrable on $[0, b]$.

As a pathological example, we note that there even exist functions that have non-integrable, bounded derivatives. Since the construction of such functions utilizes the Cantor set, it steps too far afield from the current discussion, and so will not appear in this paper. We simply note that in this case, it will be impossible to use the improper Riemann integral to evaluate these type of derivatives.

Fortunately, the slight modification to the definition of the Riemann integral that results in the Henstock integral allows us to overcome this deficiency. Thus, by allowing δ to be defined as a positive function, we find that every derivative is Henstock integrable. In fact, this statement can be further generalized and we see that every function that is differentiable except perhaps at a countable number of points can be recovered via Henstock integration from its derivative.

Since the definition of the Henstock integral requires that every Henstock integrable function is defined at all points between the bounds of integration, we make use of the idea that if two functions differ only on a set of measure zero, then the Henstock integral of the two functions is equal. Thus, in order to ensure that the derivative of a function is defined at all points in an interval, we assign an arbitrary value to the derivative at those points where the function is not differentiable. For the sake of simplicity, we say that the derivative is zero at these points.

Theorem 16. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function on $[a, b]$. If f is differentiable nearly everywhere on $[a, b]$, then f' is Henstock integrable on $[a, b]$ and $\int_a^z f' = f(z) - f(a)$ for each $x \in [a, b]$.*

Proof. Let $\epsilon > 0$, and let $S_1 = \{a_n : n \in \mathbb{Z}^+\}$ be the set of points in $[a, b]$ where f is not differentiable. Let S_2 be the complement of S_1 in $[a, b]$. By convention, let $f' : [a,b] \to \mathbb{R}$ be the function defined by

$$f'(x) = \begin{cases} \lim\limits_{t \to x} \dfrac{f(t) - f(x)}{t - x}, & \text{for } x \in S_2 \\ 0, & \text{for } x \in S_1 \end{cases}$$

Since f is continuous on $[a,b]$, we know that for each $a_n \in S_1$, there exists a positive number δ_{a_n} such that $|f(t) - f(a_n)| < \frac{\epsilon}{2^{n+1}}$ for all $t \in [a,b]$ that satisfy $t \in (a_n - \delta_{a_n}, a_n + \delta_{a_n})$. Now, as f is differentiable at each point in S_2, then for each $x \in S_2$, there exists a positive number δ_x such that $|f(t) - f(x) - f'(x)(t-x)| \leq \frac{\epsilon|t-x|}{2(b-a)}$ for all $t \in [a,b]$ that satisfy $t \in (x - \delta_x, x + \delta_x)$. Define a function $\delta : [a,b] \to \mathbb{R}^+$ by $\delta(x) = \delta_x$ for all $x \in [a,b]$. Fix a point $x \in [a,b]$. Note that if $x = a$, then by standard convention, we say that $\int_a^a f' = 0 = f(a) - f(a)$. Suppose then that $x \in (a,b]$, and let ${}^tP = \{(t_i, [x_{i-1}, x_i]\}$ be a δ-fine tagged partition of $[a,x]$, where each tag appears only once. Let σ_1 denote the set of tags i such that $t_i \in S_1$ and let σ_2 denote the set of tags i such that $t_i \in S_2$. Recalling that $x_0 = a$ and $x_n = x$, we use a telescoping sum to show that

$$\begin{aligned}
\left|S(f', {}^tP) - (f(x) - f(a))\right| &= \left|\sum_{i=1}^{n} f'(t_i)(x_i - x_{i-1}) - \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\right| \\
&\leq \sum_{i=1}^{n} \left| f'(t_i)(x_i - x_{i-1}) - (f(x_i) - f(x_{i-1})) \right| \\
&= \sum_{i \in \sigma_1} \left| f'(t_i)(x_i - x_{i-1}) - (f(x_i) - f(x_{i-1})) \right| \\
&\quad + \sum_{i \in \sigma_2} \left| f'(t_i)(x_i - x_{i-1}) - (f(x_i) - f(x_{i-1})) \right| \\
&\leq \sum_{i \in \sigma_1} \left| (f(x_i) - f(x_{i-1})) \right| \\
&\quad + \sum_{i \in \sigma_2} \left| f'(t_i)(x_i - t_i) - (f(x_i) - f(t_i)) \right| \\
&\quad + \sum_{i \in \sigma_2} \left| f'(t_i)(t_i - x_{i-1}) - (f(t_i) - f(x_{i-1})) \right| \\
&< \sum_{i \in \sigma_1} \frac{\epsilon}{2^{i+1}} + \sum_{i \in \sigma_2} \frac{\epsilon|t_i - x_{i-1}|}{2(b-a)} + \sum_{i \in \sigma_2} \frac{\epsilon|x_i - t_i|}{2(b-a)} \\
&< \frac{\epsilon}{2} + \frac{\epsilon(b-a)}{2(b-a)} = \epsilon
\end{aligned}$$

Since the choice of $x \in [a,b]$ was arbitrary, it follows that f' is Henstock integrable on $[a,b]$ and that $\int_a^t f' = f(x) - f(a)$ for each $x \in [a,b]$ □

Having read this theorem, it might be natural to wonder: if a function is differentiable almost everywhere (that is, if it is differentiable except on a set of measure zero), is it necessarily Henstock integrable? The answer, unfortunately, is no. Although outside the scope of this paper, one can construct a counterexample through the careful manipulation of the Cantor function.

As an example of the power of the Henstock integral, we will show by the definition that the derivative of the function $f(x) = \sqrt{x}$ is Henstock integrable, and that integrating the derivative retrieves the original function as desired. We restrict ourselves to the interval $[0, 1]$ and define $f : [0, 1] \to \mathbb{R}$ by $f(x) = \sqrt{x}$. We note that f is continuous on $[0, 1]$ and differentiable on $(0, 1]$. Using the usual calculus operations, we define $f' : [0, 1] \to \mathbb{R}$ as follows:

$$f'(x) = \begin{cases} \frac{1}{2\sqrt{x}}, & \text{for } x \geq 0 \\ 0, & \text{for } x = 0. \end{cases}$$

We wish to show that $\int_0^1 f' = f(1) - f(0) = 1$. Let $\epsilon > 0$, and define a positive function δ as follows: if $x = 0$, then $\delta(x) = \frac{1}{16}\epsilon$. If $x > 0$, choose $\delta(x)$ such that, for any $[u, v] \subseteq (x - \delta(x), x + \delta(x))$, we have

$$\left| \frac{1}{\sqrt{u} + \sqrt{v}} - \frac{1}{\sqrt{x}} \right| < \frac{\epsilon}{2}.$$

Now, let ${}^tP$ be a δ-fine tagged partition. For each $1 \leq i \leq n$, we use the above inequality and multiply through by $(x_i - x_{i-1})$, so that

$$\left|(\sqrt{x_i} - \sqrt{x_{i-1}}) - f'(t_i)(x_i - x_{i-1})\right| < \frac{\epsilon}{2}(x_i - x_{i-1}).$$

Therefore,

$$\begin{aligned} |S(f, {}^tP) - 1| &= \left| \sum_{i=1}^{n} (x_i - x_{i-1}) - \sum_{i=1}^{n} f'(t_i)(x_i - x_{i-1}) \right| \\ &\leq \sum_{i=1}^{n} |(x_i - x_{i-1}) - f'(t_i)(x_i - x_{i-1})| \\ &\leq \sum_{i=1}^{n} 2 \left|(\sqrt{x_i} - \sqrt{x_{i-1}}) - f'(t_i)(x_i - x_{i-1})\right| \\ &< \sum_{i=1}^{n} 2\frac{\epsilon}{2}(x_i - x_{i-1}) \\ &= \epsilon. \end{aligned}$$

Thus, f' is integrable, and $\int_0^1 f' = 1$.

5 Monotone Convergence Theorem

Although the Henstock integral allows us to successfully integrate all derivatives, this is not the only shortcoming of the Riemann integral that it overcomes. Another arises in the study of sequences of functions. One of the most intuitive ways to define the limit of a sequence of functions is to, for each point in an interval, look at the limit of the sequence of functions evaluated at that point. Consider a sequence of functions $\{f_n\}$ defined on a closed interval $[a, b]$ such that for each x in I, the sequence $\{f_n(x)\}$ converges. Define a function $f : I \to \mathbb{R}$ by $f(x) = \lim_{n\to\infty} \{f_n(x)\}$. We call f the pointwise limit of $\{f_n\}$. As it turns out, supposing that each function in the sequence is Riemann integrable is not enough to guarantee that the limit function is also Riemann integrable. A straightforward counterexample follows.

Suppose $\{a_n\}$ is a listing of the rational numbers in the interval $[0,1]$. For each positive integer n, let f_n be the function defined on $[0,1]$ by $f_n(x) = 1$, if $x \in \{a_k : 1 \le k \le n\}$, and $f_n(x) = 0$ otherwise. Since each function in the sequence is nonzero only on a finite number of points, each is Riemann integrable. However, we note that the sequence of functions converges pointwise to the characteristic function of the rationals on the interval $[0,1]$, which, as we demonstrated earlier, is not Riemann integrable. It follows that not every sequence of Riemann integrable functions converges pointwise to a Riemann integrable function (Even if that limit function is bounded). As is shown in most real analysis texts, however, the stricter conditions required for uniform convergence are enough to guarantee that the limit function of a uniformly convergent sequence of Riemann integrable functions is Riemann integrable.

The following theorem provides one set of conditions that guarantee that the limit of a pointwise convergent sequence of Henstock integrable functions is also Henstock integrable. Before we begin, however, we will need to introduce an important lemma.

Lemma 5 (Henstock's Lemma)**.** *Let $f : [a,b] \to \mathbb{R}$ be a Henstock integrable function, and let $\epsilon > 0$. Suppose that δ is a positive function on $[a,b]$ such that $|S(f, {}^tP) - \int_a^b f| < \epsilon$ for all δ-fine tagged partitions ${}^tP$ of $[a,b]$. If ${}^tP_0 = \{(t_i, [x_{i-1}, x_i]); 1 \le i \le m\}$ is a δ-fine collection of disjoint, tagged intervals, then*

$$\left| S(f, {}^tP_0) - \sum_{i=1}^{m} \int_{x_{i-1}}^{x_i} f \right| \le \epsilon.$$

Further,

$$\sum_{i=1}^{m} \left| f(t_i)(x_i - x_{i-1}) - \int_{x_{i-1}}^{x_i} f \right| \le 2\epsilon.$$

Proof. Let S denote the set $\bigcup_{i=1}^{m}[x_{i-1}, x_i]$, and let $\{I_j : 1 \le j \le n\}$ be a collection of nonoverlapping (except possibly at endpoints) closed intervals such that $\{I_j\} \cup S$ partitions the interval $[a,b]$. Let $\alpha > 0$. For each positive integer j, let ${}^tP_j$ be a δ-fine tagged partition of I_j such that $\left| S(f, {}^tP_j) - \int_{I_j} f \right| < \frac{\alpha}{n}$. Define a tagged partition ${}^tP$ of $[a,b]$ by ${}^tP = \bigcup_{j=0}^{n} {}^tP_j$. We note that ${}^tP$ is a δ-fine tagged partition of $[a,b]$, that $S(f, {}^tP) = \sum_{j=0}^{n} S(f, {}^tP_j)$, and that $\int_a^b f = \int_S f +$

$\sum_{j=1}^{n} \int_{I_j} f$. Therefore, we have that

$$\begin{aligned}
\left| S(f, {}^tP_0) - \int_S f \right| &= \left| \left(S(f, {}^tP) - \sum_{j=1}^{n} S(f, {}^tP_j) \right) - \left(\int_a^b f - \sum_{j=1}^{n} \int_{I_j} f \right) \right| \\
&\leq \left| S(f, {}^tP) - \int_a^b f \right| + \left| \sum_{j=1}^{n} S(f, {}^tP_j) - \sum_{j=1}^{n} \int_{I_j} f \right| \\
&\leq \left| S(f, {}^tP) - \int_a^b f \right| + \sum_{j=1}^{n} \left| S(f, {}^tP_j) - \int_{I_j} f \right| \\
&< \epsilon + n\frac{\alpha}{n} = \epsilon + \alpha.
\end{aligned}$$

Since α can be made arbitrarily small, it follows that $\left| S(f, {}^tP_0) - \int_S f \right| \leq \epsilon$.

Now, let σ_1 be the set of indices i of tagged intervals in ${}^tP_0$ such that $f(t_i)(x_i - x_{i-1}) - \int_{x_{i-1}}^{x_i} f \geq 0$ and let σ_2 be the set of indices i of tagged intervals in ${}^tP_0$ such that $f(t_i)(x_i - x_{i-1}) - \int_{x_{i-1}}^{x_i} f < 0$. By the first part of the lemma, we have that

$$\sum_{i \in \sigma_1} \left| f(t_i)(x_i - x_{i-1}) - \int_{x_{i-1}}^{x_i} f \right| = \sum_{i \in \sigma_1} \left(f(t_i)(x_i - x_{i-1}) - \int_{x_{i-1}}^{x_i} f \right) \leq \epsilon,$$

$$\sum_{i \in \sigma_3} \left| f(t_i)(x_i - x_{i-1}) - \int_{x_{i-1}}^{x_i} f \right| = - \sum_{i \in \sigma_3} \left(f(t_i)(x_i - x_{i-1}) - \int_{x_{i-1}}^{x_i} f \right) \leq \epsilon.$$

It follows that $\sum_{i=1}^{m} \left| f(t_i)(x_i - x_{i-1}) - \int_{x_{i-1}}^{x_i} f \right| \leq 2\epsilon$. □

Theorem 17 (Monotone Convergence Theorem)**.** *Let $\{f_n\}$ be a monotone sequence of Henstock integral functions defined on $[a, b]$ that converges pointwise to a limit function f. If $\lim_{n\to\infty} \int_a^b f_n$ exists, then f is Henstock integrable and $\int_a^b f = \lim_{n\to\infty} \int_a^b f_n$.*

Proof. Suppose first that $\{f_n\}$ is an increasing sequence of functions. A proof of the case when $\{f_n\}$ is decreasing is analogous. Let $\epsilon > 0$ and let $L = \lim_{n\to\infty} \int_a^b f_n$. Since $\{\int_a^b f_n\}$ converges, we know that there exists a positive integer N such that $\left|\int_a^b f_n - L\right| < \frac{\epsilon}{3}$ for all $n \geq N$. Further, since f_n is Henstock integrable for each n, let δ_n be a positive function defined on $[a, b]$ such that $\left| S(f_n, {}^tP) - \int_a^b f_n \right| < \frac{\epsilon}{3 \cdot 2^n}$ for all δ_n-fine tagged partitions of $[a, b]$; without loss of generality, we can suppose that $\delta_n \geq \delta_{n+1}$. Finally, since $\{f_n\}$ converges pointwise to f on $[a, b]$, for each $x \in [a, b]$, choose a positive integer $M_x \geq N$ such that $|f_n(x) - f(x)| <$

$\frac{\epsilon}{3(b-a)}$. Since $\{f_n\}$ is increasing, we also know that $|f_n(x) - f(x)| = f(x) - f_n(x)$. Now, define a function $\delta : [a,b] \to \mathbb{R}^+$ by $\delta(x) = \delta_{M_x}(x)$, and let ${}^tP = \{(t_i, [x_{i-1}, x_i] : 1 \le i \le m\}$ be a δ-fine tagged partition of $[a,b]$.

We next partition the set of tagged subintervals of ${}^tP$ into classes based on the lengths of the subintervals. Recall that if a tagged partition is δ_n-fine, then it is also δ_{n-1}-fine. Let ${}^tP_m$ be the collection of $\delta_{M_{t_m}}$-fine tagged subintervals of ${}^tP$, and for each $m-1 \ge i \ge 1$, let ${}^tP_i$ be the collection of $\delta_{M_{t_i}}$-fine tagged partitions of ${}^tP - \{\bigcup_{j=i+1}^{m} {}^tP_j\}$. We note that the set of ${}^tP_i$ for $1 \le i \le m$ is pairwise disjoint, and that the union elements in this set is equal to ${}^tP$. For each $1 \le k \le m$, let K_i be the collection of intervals used in the collection of tagged intervals ${}^tP_i$. For improved readability, the remainder of the proof has been broken down into a series of smaller steps.

First, by the triangle inequality, we have that

$$
\begin{aligned}
|S(f, {}^tP) - L| \le & \left| S(f, {}^tp) - \sum_{i=1}^{m} S(f_{M_{t_i}}, {}^tP_i) \right| \\
& + \left| \sum_{i=1}^{m} S(f_{M_{t_i}}, {}^tP_i) - \sum_{i=1}^{m} \int_{K_i} f_{M_{t_i}} \right| + \left| \sum_{i=1}^{m} \int_{K_i} f_{M_{t_i}} - L \right|.
\end{aligned}
$$

Looking at the first of the three quantities on the right side of this inequality, we note that $\sum_{i=1}^{m} S(f_{M_{t_i}}, {}^tP_i) = \sum_{i=1}^{m} f_{M_{t_i}}(t_i)(x_i - x_{i-1})$. Thus,

$$
\begin{aligned}
& \left| S(f, {}^tP) - \sum_{i=1}^{m} S(f_{M_{t_i}}(t_i), {}^tP_i) \right| \\
& \qquad = \left| \sum_{i=1}^{m} f(t_i)(x_i - x_{i-1}) - \sum_{i=1}^{m} f_{M_{t_i}}(t_i)(x_i - x_{i-1}) \right| \\
& \qquad = \left| \sum_{i=1}^{m} \left(f(t_i) - f_{M_{t_i}}(t_i) \right) (x_i - x_{i-1}) \right| \\
& \qquad \le \sum_{i=1}^{m} \left| f(t_i) - f_{M_{t_i}}(t_i) \right| (x_i - x_{i-1}) \\
& \qquad < \frac{\epsilon}{3(b-a)}(b-a) = \frac{\epsilon}{3}.
\end{aligned}
$$

Next, we will use Henstock's Lemma to show that the second of the three quantities above is less that $\frac{\epsilon}{3}$. By definition, we know that the collection of intervals ${}^tP_{M_{t_i}}$ is $\delta_{M_{t_i}}$-fine for each $1 \le i \le m$. Since $\left| S(f_n, {}^tP) - \int_a^b f_n \right| < \frac{\epsilon}{3 \cdot 2^n}$, it follows by Henstock's Lemma that $\left| S(f_{M_{t_i}}, {}^tP_i) - \int_{K_i} f_{M_{t_i}} \right| < \frac{\epsilon}{3 \cdot 2^{M_{t_i}}}$ for each

$1 \leq i \leq m$. Thus,

$$\begin{aligned}
\left|\sum_{i=1}^{m} S(f_{M_{t_i}}, {}^tP_i) - \sum_{i=1}^{m}\int_{K_i} f_{M_{t_i}}\right| &\leq \sum_{i=1}^{m}\left|S(f_{M_{t_i}}, {}^tP_i) - \int_{K_i} f_{M_{t_i}}\right| \\
&< \sum_{i=1}^{m} \frac{\epsilon}{3 \cdot 2^{M_{t_i}}} \\
&< \frac{\epsilon}{3}.
\end{aligned}$$

We now turn our attention to the third term in the inequality above. Since $\{f_n\}$ is an increasing sequence of functions, and as each $M_{t_i} \geq N$, then $f_{M_{t_i}} \geq f_N$. Further, since $\{f_n\}$ converges pointwise to f from below, then $\{\int_a^b f_n\}$ converges to L from below, and $\left|\int_{K_i} f_{M_{t_i}} - L\right| \leq \left|\int_{K_i} f_N - L\right|$ for each $1 \leq i \leq m$. Therefore, we have that

$$\left|\sum_{i=1}^{m}\int_{K_i} f_{M_{t_i}} - L\right| \leq \left|\sum_{i=1}^{m}\int_{K_i} f_N - L\right| = \left|\int_a^b f_N - L\right| < \frac{\epsilon}{3}.$$

Combining these three steps, we see that

$$\begin{aligned}
|S(f, {}^tP) - L| \leq &\left|S(f, {}^tp) - \sum_{i=1}^{m} S(f_{M_{t_i}}, {}^tP_i)\right| \\
&+ \left|\sum_{i=1}^{m} S(f_{M_{t_i}}, {}^tP_i) - \sum_{i=1}^{m}\int_{K_i} f_{M_{t_i}}\right| + \left|\sum_{i=1}^{m}\int_{K_i} f_{M_{t_i}} - L\right| \\
< &\frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

It follows that f is Henstock integrable, and that $\displaystyle\int_a^b f = L = \lim_{n\to\infty}\int_a^b f_n$. □

Conclusion

We have now completed our survey of the Henstock integral. We have seen how a small change to the definition of the Riemann integral and the introduction of the δ-fine tagged partition has provided us with a generalized and much more robust integral. Importantly, we have seen how the Henstock integral is able to overcome many of the deficiencies of the Riemann integral; in particular, we have verified that every derivative is necessarily Henstock integrable, and that the Henstock integral provides a much greater degree of flexibility when working with limit interchanges. Although we have conducted this investigation without the aid of measure theory, we have brushed against these limits at several points; we hope that by having eschewed this more advanced theory, we have kept this paper accessible at the undergraduate level. Nevertheless, it may be fruitful in the future to explore the subject more thoroughly using the full power of Lebesgue measure theory, in order to see what further results may be discovered.

References

[1] Gordon, R. 2002, *Real Analysis: A First Course*, 2nd edition, Addison-Wesley Higher Mathematics.

[2] Katz, V. 2009, *A History of Mathematics: An Introduction*, 3rd edition, Addison-Wesley Higher Mathematics.