Math 402 - Real Analysis
The Henstock-Kurzweil Integral

Steven Kao & Jocelyn Gonzales

April 28, 2015

1 Introduction to the Henstock-Kurzweil Integral

Although the Riemann integral is the primary integration technique taught to undergraduates, there are several drawbacks to the Riemann integral.

I. A lot of functions are not Riemann integrable. Recall that a bounded function is only Riemann integrable if its set of discontinuities has measure zero. For example Dirichlet's function:

$$g(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R}\backslash\mathbb{Q} \end{cases}$$

is not Riemann integrable since $U(f, \mathbf{P}) = 1$ and $L(f, \mathbf{P}) = 0$. Thus, $U(f, \mathbf{P}) \neq L(f, \mathbf{P})$.

II. Every derivative is not Riemann integrable. We have seen in class that the function

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases} \tag{1}$$

on $[-1, 1]$ has the derivative

$$f'(x) = \begin{cases} -\frac{2}{x}\cos\left(\frac{1}{x^2}\right) + 2x\left(\sin\left(\frac{1}{x^2}\right)\right) & x \neq 0 \\ 0 & x = 0 \end{cases} \tag{2}$$

however $f'(x)$ is unbounded on $[-1, 1]$, thus it is not Riemann integrable.
Using the Riemann integral, the Fundamental Theorem of Calculus states if $f : [a, b] \to \mathbb{R}$ is Riemann integrable and $F : [a, b] \to \mathbb{R}$ with $F'(x) = f(x)$, $\forall x \in [a, b]$, then

$$\int_a^b f = F(b) - F(a).$$

Thus, the Fundamental Theorem of Calculus requires that f be Riemann integrable. Since not every derivative is Riemann integrable, the Riemann integral places a constraint on the Fundamental Theorem of Calculus. For example, f might have an antiderivative F, but this does not imply that f is Riemann integrable. Thus, the formula from the Fundamental Theorem of Calculus can not be applied.

The Lebesgue integral was introduced in 1902 by Henry Lebesgue. This integral centers around using the range instead of the domain to integrate functions. Although the Lebesgue integral can integrate a larger class of functions then the Riemann integral, it still has limitations. First, there are still a large number of functions which can not be integrated using the Lebesgue integral. Also, the Lebesgue integral does not guarantee that every derivative is integrable. For example, equation (1) is not Lebesgue integral. Thus, it also adds additional constraints on the Fundamental Theorem of Calculus.

The desire to develop an integral that could integrate more functions and generalize the Fundamental Theorem of Calculus motivated mathematicians Ralph Henstock and Jaroslav Kurzweil. In the 1960's, both Henstock and Kurzweil independently developed the Henstock-Kurzweil (HK) integral. The HK integral can integrate a much larger class of functions than the Riemann integral and the Lebesgue integral. We will see later that the HK integral can integrate the Dirichlet function. Also, using the HK integral, every derivative is integrable. We will also see that using the generalized Riemann integral. the Fundamental Theorem of Calculus can be made more general since we can eliminate the assumption that f is Riemann integrable.

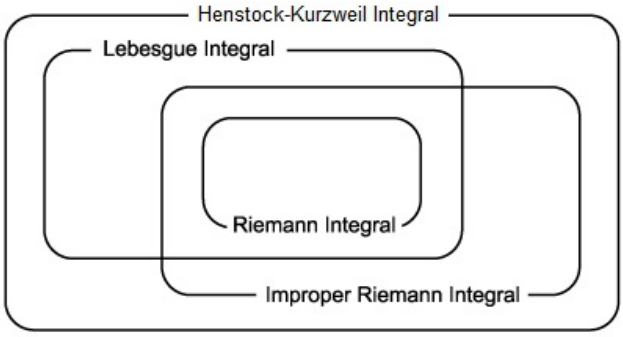

Figure 1: Classification of Integrable Functions

Figure 2: Henstock and Kurzweil

2 Gauges and δ-fine Partitions

We will begin with a few definitions.

Tagged Partition: Let $\mathbf{P}$ be a *partition* of the interval $[a, b]$ with $\mathbf{P} = a = x_o < x_1 < x_2 < ... < x_n = b$. A *tagged partitttion* $(\mathbf{P}, (c_k)_{k=1}^n)$ is a partition which has selected points c_k in each subinterval $[x_{k-1}, x_k]$.

The Riemann sum using the tagged partition can be written

$$R(f, \mathbf{P}) = \sum_{k=1}^{n} f(c_k)[x_k, x_{k-1}].$$

δ-fine: Let $\delta > 0$. A partition $\mathbf{P}$ is *δ-fine* if every subinterval $[x_{k-1}, x_k]$ satisfies $x_k - x_{k-1} < \delta$.

The definition for Riemann integrability can be restated as the following:

> A bounded function $f : [a, b] \to \mathbf{R}$ is Riemann integrable with $\int_a^b f = A$ if and only if for every $\epsilon > 0$, there exists a $\delta > 0$ such that, for any tagged partition $(\mathbf{P}, (c_k))$ that is δ-fine, it follows that $|R(f, \mathbf{P}) - A| < \epsilon$.

Gauge: A function $\delta : [a, b] \to \mathbf{R}$ is called a *gauge* on $[a, b]$ if $\delta(x) > 0$ for all $x \in [a, b]$.

The major difference between the HK integral and the Riemann integral is allowing δ to be a function of x rather than a constant.

$\delta(x)$-fine: Given a particular gauge $\delta(x)$, a tagged partition $(\mathbf{P}, (c_k)_{k=1}^n)$
is *$\delta(x)$-fine* if every subinterval $[x_{k-1}, x_k]$ satisfies $x_k - x_{k-1} < \delta(c_k)$.

If $\delta(x)$ is a constant function, then the definition of a $\delta(x)$-fine partition is equivalent to the definition of a δ-fine partition.

3 Examples of $\delta(x)$-fine tagged partitions

Example 1

Consider the inverval $[0, 1]$. Let $\delta_1(x) = 1/9$. We will find a $\delta_1(x)$-fine tagged partition on $[0, 1]$.

Since the gauge $\delta_1(x)$ is a constant function. Regardless of the choice of tag, $\delta_1(c_k) = 1/9$. Thus, any tagged partition $(\mathbf{P}, (c_k)_{k=1}^n)$ in which $x_k - x_{k-1} < 1/9$ is a $\delta_1(x)$-fine tagged partition.

Consider the following partition, choosing each tag from every interval to be any number in that inverval:

$$m\left(\left[0, \frac{1}{10}\right]\right) < \frac{1}{9} \qquad m\left(\left[\frac{4}{10}, \frac{5}{10}\right]\right) < \frac{1}{9} \qquad m\left(\left[\frac{7}{10}, \frac{8}{10}\right]\right) < \frac{1}{9}$$
$$m\left(\left[\frac{1}{10}, \frac{2}{10}\right]\right) < \frac{1}{9} \qquad m\left(\left[\frac{5}{10}, \frac{6}{10}\right]\right) < \frac{1}{9} \qquad m\left(\left[\frac{8}{10}, \frac{9}{10}\right]\right) < \frac{1}{9}$$
$$m\left(\left[\frac{2}{10}, \frac{3}{10}\right]\right) < \frac{1}{9} \qquad m\left(\left[\frac{6}{10}, \frac{7}{10}\right]\right) < \frac{1}{9} \qquad m\left(\left[\frac{9}{10}, 1\right]\right) < \frac{1}{9}$$
$$m\left(\left[\frac{3}{10}, \frac{4}{10}\right]\right) < \frac{1}{9}$$

This is an example of a $\delta_1(x)$-fine tagged partition.

Example 2

Again considering the interval $[0,1]$, let

$$\delta_2(x) = \begin{cases} 1/4 & x = 0 \\ x/3 & 0 < x \leq 1. \end{cases}$$

Consider the following partition, choosing the first tag from the first interval to be 0 and each tag from every other interval to be the right hand end-point of that interval:

$$m\left(\left[0, \frac{1}{5}\right]\right) = \frac{1}{5} < \delta_2\left(0\right) = \frac{1}{4} \qquad m\left(\left[\frac{2}{5}, \frac{1}{2}\right]\right) = \frac{1}{10} < \delta_2\left(\frac{1}{2}\right) = \frac{1}{6}$$
$$m\left(\left[\frac{1}{5}, \frac{1}{4}\right]\right) = \frac{1}{20} < \delta_2\left(\frac{1}{4}\right) = \frac{1}{12} \qquad m\left(\left[\frac{1}{2}, \frac{3}{5}\right]\right) = \frac{1}{10} < \delta_2\left(\frac{3}{5}\right) = \frac{1}{5}$$
$$m\left(\left[\frac{1}{4}, \frac{1}{3}\right]\right) = \frac{1}{12} < \delta_2\left(\frac{1}{3}\right) = \frac{1}{9} \qquad m\left(\left[\frac{3}{5}, \frac{3}{4}\right]\right) = \frac{3}{20} < \delta_2\left(\frac{3}{4}\right) = \frac{1}{4}$$
$$m\left(\left[\frac{1}{3}, \frac{2}{5}\right]\right) = \frac{1}{15} < \delta_2\left(\frac{2}{5}\right) = \frac{2}{15} \qquad m\left(\left[\frac{3}{4}, 1\right]\right) = \frac{1}{4} < \delta_2\left(1\right) = \frac{1}{3}$$

This is an example of a $\delta_2(x)$-fine tagged partition.

4 Theorem 1

Before we state and prove Theorem 1, we will first prove the Nested Interval Property, which will be used in the proof of Theorem 1.

Nested Interval Property

For each $n \in \mathbb{N}$, assume we are given a closed interval
$I_n = [a_n, b_n] = \{x \in R : a_n \leq x \leq b_n\}$. Assume also that each I_n contains I_{n+1}. Then, the resulting nested sequence of closed intervals

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$$

has a nonempty intersection; that is, $\bigcap_{n=1}^{\infty} I_n \neq \varnothing$.

Proof. Let Let $A = \{a_n : n \in \mathbb{N}\}$ be the left-hand endpoints of each interval. Since every interval is contained inside of the previous interval, the b_n terms are upper bounds for A. By the existence of a least upper bound, since A has an upper bound, it must have a least upper bound. Let $x = \sup A$. Because x is an upper bound for A, then $a_n \leq x$, $\forall n$. Also, since x is a least upper bound of A, $x \leq b_n$, $\forall n$. Then we have $a_n \leq x \leq b_n$, $\forall n$. Thus $x \in I_n$ for every choice of $n \in \mathbb{N}$. Thus $x \in \bigcap_{n=1}^{\infty} I_n$ and the intersection is not empty. □

Theorem 1:

Given a gauge $\delta(x)$ on an interval $[a,b]$, there exists a tagged partition $(\mathbf{P}, (c_k)_{k=1}^n)$ that is $\delta(x)$-fine.

Proof. Let $\delta(x)$ be a gauge on $[a,b]$. An algorithm to find a $\delta(x)$-fine tagged partition is as follows. First consider the trivial partition $\mathbf{P_o} = \{a = x_o < x_1 = b\}$. Then check to see if $\exists\, c_o \in [a,b] \setminus \mathbb{I}$ such that $b - a < \delta(c_o)$. If such c_o exists, then choose $(\mathbf{P_o}, (c_o))$. Thus there exists a tagged partition which is $\delta(x)$-fine. If no c_o exists such that $b - a < \delta(c_o)$, then bisect the interval into two equal halves. Consider the partition
$\mathbf{P_1} = \{a = x_o < x_1 < x_2 = b\}$. Then apply the algorithm to each new half.

We will now prove the algorithm must terminate in a finite number of steps. Assume that the algorithm does not terminate in a finite number of steps. Then we have infinite nested intervals (I_n) where $m(I_n) \to 0$. Since the algorithm has not terminated, this implies that $\delta(x) \leq m(I_n)$, $\forall x \in I_n$. By the Nested Interval Property, we know $\exists\, x_o \in \bigcap_{n=1}^{\infty} I_n$. Then $\delta(x_o) \leq m(I_n)$, $\forall n \in \mathbb{N}$. But since $m(I_n) \to 0$, this implies that $\delta(x_o) = 0$. We draw a contradiction since by definition $\delta(x) > 0$. Thus this algorithm must terminate after a finite number of steps. Thus we can always create a tagged partition of $[a, b]$ that is $\delta(x)$-fine for a given gauge. □

5 Introducing the Henstock-Kurzweil (HK) integral

We now have the machinery to define the Henstock-Kurzweil integral:

Definition: A function $f : [a, b] \to \mathbb{R}$ is *Henstock-Kurzweil (HK) integrable* if $\exists\, A \in \mathbb{R}$ st. $\forall\, \varepsilon > 0$, $\exists$ a gauge $\delta : [a, b] \to \mathbb{R}$ st. for each tagged partition $(\mathbf{P}, (c_k)_{k=1}^n)$ that is $\delta(x)$-fine,

$$|R(f, \mathbf{P}) - A| < \varepsilon$$

In this case, we write $\int_a^b f = A$, or $\mathrm{HK}\int_a^b f = A$ (if we need to distinguish between different versions of integrals) and say that *f has an HK integral value of A*.

Our first order of business is to make sure that if a function has an HK integral value, it can have only one such value:

Theorem 2: If a function is HK integrable, its value is unique.

Proof. Suppose $f : [a, b] \to \mathbb{R}$ is HK integrable and has values A_1 and A_2. Let $\varepsilon > 0 \Rightarrow \varepsilon/2 > 0$.
By definition of HK integrable, $\exists$ a gauge $\delta_1 : [a, b] \to \mathbb{R}$ st. $\forall$ tagged partitions $(\mathbf{P_1}, (a_k)_{k=1}^{n_1})$ which are $\delta_1(x)$-fine,

$$|R(f, \mathbf{P_1}) - A_1| < \varepsilon/2 \tag{3}$$

Similarly, $\exists$ a gauge $\delta_2 : [a, b] \to \mathbb{R}$ st. $\forall$ tagged partitions $(\mathbf{P_2}, (b_k)_{k=1}^{n_2})$ which are $\delta_2(x)$-fine,

$$|R(f, \mathbf{P_2}) - A_2| < \varepsilon/2 \tag{4}$$

Define $\delta : [a, b] \to \mathbb{R}$ to be $\delta(x) := \min(\delta_1(x), \delta_2(x))$.
Clearly δ is a gauge since $\delta(x) = \delta_1(x)$ or $= \delta_2(x)$, both of which > 0, $\forall\, x \in [a, b]$.
Now let $(\mathbf{P}, (c_k)_{k=1}^n)$ be an arbitrary tagged partition which is $\delta(x)$-fine. Notice, $(\mathbf{P}, (c_k)_{k=1}^n)$ is necessarily $\delta_1(x)$-fine and $\delta_2(x)$-fine. This is because for any subinterval $[x_{k-1}, x_k]$ with endpoints taken from $\mathbf{P}$, we have

$$\begin{aligned} x_k - x_{k-1} < \delta(c_k) \leq \delta_1(c_k) \quad \text{and} \\ x_k - x_{k-1} < \delta(c_k) \leq \delta_2(c_k) \end{aligned}$$

Therefore,

$$\begin{aligned} 0 \leq |A_1 - A_2| &= \big|\big(A_1 - R(f, \mathbf{P})\big) + \big(R(f, \mathbf{P}) - A_2\big)\big| \quad \text{“middle-man” trick} \\ &\leq |A_1 - R(f, \mathbf{P})| + |R(f, \mathbf{P}) - A_2| \quad \text{triangle ineq.} \\ &< \varepsilon/2 + \varepsilon/2 \quad \text{from (3), (4) since } R(f, \mathbf{P}) \text{ is } \delta_1(x) \text{ and } \delta_2(x)\text{-fine} \\ &= \varepsilon \end{aligned}$$

And since since ε is arbitrary $\Rightarrow |A_1 - A_2| = 0 \Rightarrow A_1 = A_2$. □

6 Dirichlet's function, revisited

Now that we've shown the value of an HK integral for an HK integrable function is well-defined, we are ready to calculate $\mathrm{HK}\int_0^1 g$, where $g : [0,1] \to \mathbb{R}$ is the Dirichlet function:

$$g(x) := \begin{cases} 1, & \text{if } x \in \mathbb{Q} \\ 0, & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

Claim: $\mathrm{HK}\int_0^1 g = 0$.

Proof. Let $\varepsilon > 0$. In order to prove the claim, we must construct a gauge $\delta : [0,1] \to \mathbb{R}$ st. given an arbitrary tagged partition $(\mathbf{P}, (c_k)_{k=1}^n)$ which is $\delta(x)$-fine,

$$|R(f, \mathbf{P}) - 0| = \sum_{k=1}^{n} g(c_k)(x_k - x_{k-1}) < \varepsilon. \tag{5}$$

Since the rationals are countable, we can list them as $\{r_1, r_2, r_3, \dots\}$.
Now define,

$$\delta(x) := \begin{cases} \varepsilon/2^{i+1}, & \text{if } x = r_i \in \mathbb{Q} \\ 1, & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}.$$

Notice the following three observations:

(i) The function δ is a gauge because $\delta(x) > 0, \ \forall\, x \in [0,1]$ since ε, 2, and 1 are all positive.

(ii) Because $g(c_k) = 0$ when c_k is irrational, we can ignore these irrational tags and pass to the subsequence $(c_{k_j})_{j=1}^m$ consisting of all the rational elements of $(c_k)_{k=1}^n$ when evaluating the finite sum in (5).

(iii) Because a tag can be an endpoint of a subinterval, it's possible for two different tags to have the same value, corresponding to the right endpoint of one subinterval and the left endpoint of the successive subinterval. It is not possible to have equal tags from non-successive subintervals, nor is it possible to have three or more tags with the same value. This is true for both the original sequence of tags $(c_k)_{k=1}^n$ and for the rational subsequence of tags $(c_{k_j})_{j=1}^m$.

For example, suppose we have the sequence $(c_1, c_2, \dots, c_{10})$. Then it's possible for $c_3 = c_4$, but not for $c_3 = c_5$ nor $c_3 = c_4 = c_5$. Similarly, if we have the subsequence $(c_{k_1}, c_{k_2}, \dots, c_{k_6})$, it's possible for $c_{k_3} = c_{k_4}$ but not for $c_{k_3} = c_{k_5}$ nor $c_{k_3} = c_{k_4} = c_{k_5}$.

By observation (ii), we can rewrite the lefthand side of inequality (5) as:

$$\begin{aligned}
|R(f, \mathbf{P}) - 0| &= \sum_{j=1}^{m} g(c_{k_j})(x_{k_j} - x_{k_j - 1}) \\
&= \sum_{j=1}^{m} (x_{k_j} - x_{k_j - 1}) \quad g(c_{k_j}) = 1 \text{ since } c_{k_j} \text{ is rational} \\
&< \sum_{j=1}^{m} \delta(c_{k_j}) \quad \text{since } (\mathbf{P}, (c_k)_{k=1}^n) \text{ is } \delta(x)\text{-fine} \\
&= \sum_{j=1}^{m} \frac{\varepsilon}{2^{i_j+1}} \quad \text{def. of } \delta(x) \text{ where the } i_j\text{'s are finite subseq. of } \mathbb{N}^+ \\
&< 2 \cdot \sum_{i=1}^{\infty} \frac{\varepsilon}{2^{i+1}} \quad \text{by observation (iii)} \\
&= 2\varepsilon \cdot \frac{1/4}{1 - 1/2} = \varepsilon \quad \text{geometric series,}
\end{aligned}$$

which is what we needed to show. □

Remark: Having a gauge allows us to control the length of each subinterval, enabling us to "encapsulate" sets of measure zero. So essentially, the characteristics of the Lebesgue integral is "built into" the HK integral. Also notice, if $\varepsilon < 1$, it is not possible to form a tagged partition which is $\delta(x)$-fine and which consists only of rational tags, because when you add up the lengths of the subintervals in such a partition, the sum will be less than 1.

7 The Fundamental Theorem of Calculus

Now we come to the main feature of the Henstock-Kurzweil integral, that which sets it apart from both the Riemann and Lebesgue integrals and allows for a larger class of functions to be integrable.

Theorem 3: (The Fundamental Theorem of Calculus). Let $F : [a, b] \to \mathbb{R}$ and $f : [a, b] \to \mathbb{R}$ be functions. Suppose F is differentiable on $[a, b]$ st. $F'(x) = f(x),\ \forall\, x \in [a, b]$. Then,

$$\text{HK}\int_a^b f = F(b) - F(a).$$

Remark: Notice that the Fundamental Theorem of Calculus for HK integrals does NOT require any additional assumptions on f, unlike the Riemann integral version which requires f to be Riemann integrable and even the Lebesgue version, which requires f to be absolutely continuous.

Proof. Assume conditions and let $\varepsilon > 0 \Rightarrow \frac{\varepsilon}{2(b-a)} > 0$ since $2(b - a) > 0$. As before, we must construct a gauge $\delta : [0, 1] \to \mathbb{R}$ st. given an arbitrary tagged partition $(\mathbf{P}, (c_k)_{k=1}^n)$ which is $\delta(c)$-fine,

$$|(F(b) - F(a)) - R(f, \mathbf{P})| < \varepsilon \tag{6}$$

(The reason why we switched to using 'c' as our variable instead of 'x' will soon be clear.)

There are three key observations.
First, since $F(c)$ is differentiable $\forall\, c \in [a, b]$ with derivative equal to $f(c)$, by Newton's approximation:

$$\exists\, \delta(c) > 0 \text{ st. } \forall\, |x - c| \le \delta(c),\ |F(x) - F(c) - f(c)(x - c)| \le \frac{\varepsilon|x - c|}{2(b - a)}. \tag{7}$$

In other words, the value of δ *depends* on the choice of c, since we only have regular continuity in our assumption, not uniform continuity. Since $\delta(c) > 0\ \forall\, c \in [a, b]$, it is exactly this $\delta(c)$ that will be our gauge!

Second, for an arbitrary subinterval $[x_{k-1}, x_k]$ and its tag $c_k \in [x_{k-1}, x_k]$, we have the following:

$$\begin{aligned}
|x_k - x_{k-1}| &= x_k - x_{k-1} && \text{since } x_k > x_{k-1} \\
&= (x_k - c_k) + (c_k - x_{k-1}) && \text{adding 0 and associativity} \\
&= |x_k - c_k| + |c_k - x_{k-1}| && \text{since } x_k \ge c_k \text{ and } c_k \ge x_{k-1} \\
\Rightarrow \quad |x_k - c_k| &\le |x_k - x_{k-1}| < \delta(c_k), && \text{by transitivity and because} \quad (8) \\
\Rightarrow |c_k - x_{k-1}| &\le |x_k - x_{k-1}| < \delta(c_k) && (\mathbf{P}, (c_k)_{k=1}^n) \text{ is } \delta(c)\text{-fine} \quad (9)
\end{aligned}$$

From (7) and (8) we get:

$$|F(x_k) - F(c_k) - f(c_k)(x_k - c_k)| \le \frac{\varepsilon|x_k - c_k|}{2(b - a)} = \frac{\varepsilon(x_k - c_k)}{2(b - a)}, \tag{10}$$

since $x_k > c_k$.
And from (7) and (9) we get:

$$\begin{aligned}
&|F(x_{k-1}) - F(c_k) - f(c_k)(x_{k-1} - c_k)| \leq \frac{\varepsilon|x_{k-1} - c_k|}{2(b-a)} = \frac{\varepsilon(c_k - x_{k-1})}{2(b-a)} \\
&\Rightarrow |-1| \cdot |F(x_{k-1}) - F(c_k) - f(c_k)(x_{k-1} - c_k)| \leq \frac{\varepsilon(c_k - x_{k-1})}{2(b-a)} \\
&\Rightarrow |-F(x_{k-1}) + F(c_k) - f(c_k)(-x_{k-1} + c_k)| \leq \frac{\varepsilon(c_k - x_{k-1})}{2(b-a)} \qquad (11)
\end{aligned}$$

Putting all this together,

$$\begin{aligned}
&|F(x_k) - F(x_{k-1}) - f(c_k)(x_k - x_{k-1})| = \\
&\qquad |F(x_k) + (-F(c_k) + F(c_k)) - F(x_{k-1}) - f(c_k)(x_k + (-c_k + c_k) - x_{k-1})|,
\end{aligned}$$

by the middle-man trick.

$$\leq |F(x_k) - F(c_k) - f(c_k)(x_k - c_k)| + |-F(x_{k-1}) + F(c_k) - f(c_k)(-x_{k-1} + c_k)|,$$

by triangle inequality. So,

$$\begin{aligned}
|F(x_k) - F(x_{k-1}) - f(c_k)(x_k - x_{k-1})| &\leq \frac{\varepsilon(x_k - c_k)}{2(b-a)} + \frac{\varepsilon(c_k - x_{k-1})}{2(b-a)} \\
&= \frac{\varepsilon(x_k - x_{k-1})}{2(b-a)} \qquad (12)
\end{aligned}$$

by (10) and (11).
Summing (12) over the tagged partition $(\mathbf{P}, (c_k)_{k=1}^n)$:

$$\begin{aligned}
\sum_{k=1}^{n} |F(x_k) - F(x_{k-1}) - f(c_k)(x_k - x_{k-1})| &\leq \sum_{k=1}^{n} \frac{\varepsilon(x_k - x_{k-1})}{2(b-a)} \\
&= \frac{\varepsilon(x_n - x_0)}{2(b-a)} \\
&= \frac{\varepsilon(b-a)}{2(b-a)} = \frac{\varepsilon}{2} \qquad (13)
\end{aligned}$$

because the sum is telescoping and $x_n = b$ and $x_0 = a$ by construction of our partition $\mathbf{P}$.

Third, notice that:

$$\begin{aligned}
\sum_{k=1}^{n} (F(x_k) - F(x_{k-1})) &= (F(x_1) - F(x_0)) + (F(x_2) - F(x_1)) + \cdots + (F(x_n) - F(x_{n-1})) \\
&= F(x_n) - F(x_0) \quad \text{telescoping finite sum} \\
&= F(b) - F(a) \quad \text{by construction of our partition } \mathbf{P} \qquad (14)
\end{aligned}$$

Finally, we are ready to show (6) holds:

$$\begin{aligned}
|(F(b)-F(a))-R(f,\mathbf{P})| &= \left|\sum_{k=1}^{n}(F(x_k)-F(x_{k-1}))-\sum_{k=1}^{n}f(c_k)(x_k-x_{k-1})\right| && \text{by (14) \& def. of Riem. sum}\\
&= \left|\sum_{k=1}^{n}F(x_k)-F(x_{k-1})-f(c_k)(x_k-x_{k-1})\right| &&\\
&\leq \sum_{k=1}^{n}\left|F(x_k)-F(x_{k-1})-f(c_k)(x_k-x_{k-1})\right| && \text{by finite triangle inequality}\\
&\leq \frac{\varepsilon}{2} < \varepsilon && \text{by (13),}
\end{aligned}$$

which is what we needed to show. □

8 Closing Thoughts

Obviously, we have just barely scratched the surface of the power of the Henstock-Kurzweil integral. But just to give a glimpse of its versatility, we leave the reader with two facts (which we will not prove) of the HK integral:

(i) The HK integral integrates:

  (a) all functions which have anti-derivatives

  (b) all Riemann integrable functions

  (c) all Lebesgue integrable functions

  (d) all functions that can be obtained as "improper integrals"

(ii) The HK integral has theorems which generalize

  (a) the Monotone Convergence Theorem

  (b) the Dominated Convergence Theorems of measure theory

9 References

Abbott, Stephen. Understanding Analysis. New York: Springer, 2001. Print.

Bartle, Robert Gardner. A Modern Theory of Integration. Providence, RI: American Mathematical Society, 2001. Print.

Herschlag, Greg. "A Brief Introduction to Gauge Integration." http://www.math.uchicago.edu/ may/VIGRE/VIGRE2006/PAPERS/Herschlag.pdf.[23 Apr. 2015. Web.]

McKinnis, Erik O. "Gauge Integration." (2002): Web.

"Oberwolfach Photo Collection." Details: Ralph Henstock, Jaroslav Kurzweil. Web. 28 Apr. 2015.

"Prove the Derivative Is Not Lebesgue Integrable." Math Stack Exchange. Web. 28 Apr. 2015. ¡http://math.stackexchange.com/questions/836996/prove-the-derivative-of-x2-sin-1-x2-is-not-lebesgue-integrable-on-0-1¿.

Tao, Terence. Analysis I. New Delhi: Hindustan, 2009. Print.