A Definition of the Riemann–Stieltjes Integral

Let \( a < b \) and let \( f, \alpha : [a, b] \to \mathbb{R} \). In these notes I will state one of several closely related, but not 100% equivalent, standard definitions of the Riemann–Stieltjes integral \( \int_a^b f(x) \, d\alpha(x) \). Let’s start by reviewing the first year Calculus definition of the Riemann integral \( \int_a^b f(x) \, dx \). The definition goes in several steps.

**Step 1:** Slice up the interval \([a, b]\) into a finite number of subintervals.

A partition of \([a, b]\) is a finite set \( \mathcal{P} = \{x_0, x_1, \ldots, x_n\} \) of points with
\[
a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b
\]

The mesh
\[
\|\mathcal{P}\| = \max_{1 \leq i \leq n} |x_i - x_{i-1}|
\]
of this partition is the width of the widest subinterval.

**Step 2:** For each \(1 \leq i \leq n\), choose a point to represent the subinterval \([x_{i-1}, x_i]\).

A choice for the partition \(\mathcal{P}\) is a finite set \(\mathcal{T} = \{t_1, \ldots, t_n\}\) of points with
\[
x_{i-1} \leq t_i \leq x_i \quad \text{for each } 1 \leq i \leq n
\]

**Step 3:** Calculate an approximate value for the integral, by approximating the \(i\)th strip by a rectangle of width \(x_i - x_{i-1}\) and height \(f(t_i)\).

The Riemann sum with partition \(\mathcal{P}\) and choice \(\mathcal{T}\) is
\[
S(\mathcal{P}, \mathcal{T}, f) = \sum_{i=1}^{n} f(t_i) \,(x_i - x_{i-1})
\]

**Step 4:** Make the partition finer and finer.

A partition \(\mathcal{P}'\) is said to be finer than the partition \(\mathcal{P}\) if \(\mathcal{P} \subset \mathcal{P}'\).

**Step 5:** Hope the Riemann sums converge.

A function \(f : [a, b] \to \mathbb{R}\) is said to be Riemann integrable on \([a, b]\), denoted “\(f \in \mathcal{R} \) on \([a, b]\)”, if there exists an \(I \in \mathbb{R}\) such that
\[
\forall \varepsilon > 0 \ \exists \text{a partition } \mathcal{P}_\varepsilon \text{ of } [a, b] \quad \text{s.t.} \quad |S(\mathcal{P}, \mathcal{T}, f) - I| < \varepsilon
\]
for all partitions \(\mathcal{P}\) of \([a, b]\) finer than \(\mathcal{P}_\varepsilon\) and all choices \(\mathcal{T}\) for \(\mathcal{P}\).

If so, \(I\) is denoted \(\int_a^b f(x) \, dx\) and we write
\[
\lim_{\mathcal{P} \to \mathcal{P}_\varepsilon} S(\mathcal{P}, \mathcal{T}, f) = \int_a^b f(x) \, dx
\]

To generalize this to the Riemann–Stieltjes integral \(\int_a^b f(x) \, d\alpha(x)\) we simply replace the strip width \(x_i - x_{i-1}\) with \(\alpha(x_i) - \alpha(x_{i-1})\). That is, Steps 1, 2 and 4 are left as is, and Steps 3 and 5 are replaced with...
Step 3: Calculate an approximate value for the integral.
The Riemann-Stieltjes sum with partition $\mathcal{P}$ and choice $T$ is

$$S(\mathcal{P}, T, f, \alpha) = \sum_{i=1}^{n} f(t_i) \left[ \alpha(x_i) - \alpha(x_{i-1}) \right]$$

Step 5: Hope the Riemann-Stieltjes sums converge.
A function $f : [a, b] \to \mathbb{R}$ is said to be Riemann integrable with respect to $\alpha$ on $[a, b]$, denoted “$f \in \mathcal{R}(\alpha)$ on $[a, b]$”, if there exists an $I \in \mathbb{R}$ such that

$$\forall \varepsilon > 0 \ \exists \text{a partition } \mathcal{P}_\varepsilon \text{ of } [a, b] \text{ s.t. } |S(\mathcal{P}, T, f, \alpha) - I| < \varepsilon$$

for all partitions $\mathcal{P}$ of $[a, b]$ finer than $\mathcal{P}_\varepsilon$ and all choices $T$ for $\mathcal{P}$.

If so, $I$ is denoted $\int_{a}^{b} f(x) \, d\alpha(x)$ or $\int_{a}^{b} f \, d\alpha$ and we write

$$\lim_{\mathcal{P}} S(\mathcal{P}, T, f, \alpha) = \int_{a}^{b} f(x) \, d\alpha(x)$$

Remark.

1. If $\alpha(x) = x$, then the Riemann-Stieltjes integral $\int_{a}^{b} f(x) \, d\alpha(x)$ reduces to the Riemann integral $\int_{a}^{b} f(x) \, dx$.

2. We shall eventually prove that

   (a) If $\alpha$ is monotonic and $f$ is continuous, then $f \in \mathcal{R}(\alpha)$.

   (b) If $\alpha$ is continuous and $f$ is monotonic, then $f \in \mathcal{R}(\alpha)$.

   (c) If $\alpha$ is strictly monotonic and $f$ is unbounded, then $f \notin \mathcal{R}(\alpha)$.

3. Observe that

   $$S(\mathcal{P}, T, f, \alpha) = \sum_{i=1}^{n} f(t_i) \frac{\alpha(x_i) - \alpha(x_{i-1})}{x_i - x_{i-1}} \left[ x_i - x_{i-1} \right]$$

   $$\approx \sum_{i=1}^{n} f(t_i) \alpha'(t_i) \left[ x_i - x_{i-1} \right]$$

We shall eventually prove that if $\alpha$ has a continuous derivative, then

$$\int_{a}^{b} f(x) \, d\alpha(x) = \int_{a}^{b} f(x)\alpha'(x) \, dx$$

Example.

1. If $\alpha(b) \neq \alpha(a)$ and

   $$f(x) = \begin{cases} 
   1 & \text{if } x \text{ is rational} \\
   0 & \text{if } x \text{ is irrational}
   \end{cases}$$

then $f \notin \mathcal{R}(\alpha)$ on $[a, b]$ because, for any choice $T$ consisting only of rational points,

$$S(\mathcal{P}, T, f, \alpha) = \sum_{i=1}^{n} f(t_i) \left[ \alpha(x_i) - \alpha(x_{i-1}) \right] = \sum_{i=1}^{n} 1 \times \left[ \alpha(x_i) - \alpha(x_{i-1}) \right] = \alpha(b) - \alpha(a)$$
while, for any choice $T$ consisting only of irrational points,

$$S(\mathbf{P}, T, f, \alpha) = \sum_{i=1}^{n} f(t_i) \left[ \alpha(x_i) - \alpha(x_{i-1}) \right] = \sum_{i=1}^{n} 0 \times \left[ \alpha(x_i) - \alpha(x_{i-1}) \right] = 0$$

So if $\varepsilon < \frac{\alpha(b) - \alpha(a)}{2}$, there is no number $I$ such that both of these partial sums are within $\varepsilon$ of $I$, no matter what $\mathbf{P}$ is.

2. If $\alpha(x) = x$ and

$$f(x) = \begin{cases} \frac{1}{n} & \text{if } x \text{ is rational with } x = \frac{m}{n} \text{ in lowest terms} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

then $f \in \mathcal{R}$ on $[a, b]$ and $\int_{a}^{b} f(x) \, dx = 0$.

**Proof:** Let $\varepsilon > 0$. There are only finitely many, say $N$, points in $[a, b]$ with $f(x) > \frac{\varepsilon}{2(b-a)}$. Let $\mathbf{P}$ be any partition with mesh $\|\mathbf{P}\| < \frac{\varepsilon}{4N}$.

- For any subinterval that does contain a point $x$ where $f(x) > \frac{\varepsilon}{2(b-a)}$

  $$|f(t_i)| |x_i - x_{i-1}| \leq 1 \cdot \|\mathbf{P}\| < \frac{\varepsilon}{4N}$$

  There are at most $2N$ such subintervals, because each of the $N$ points with $f(x) > \frac{\varepsilon}{2(b-a)}$ can appear in at most 2 subintervals.

- For any subinterval that does not contain a point $x$ where $f(x) > \frac{\varepsilon}{2(b-a)}$

  $$|f(t_i)| |x_i - x_{i-1}| \leq \frac{\varepsilon}{2(b-a)} |x_i - x_{i-1}|$$

  so that

  $$|S(\mathbf{P}, T, f)| < 2N \cdot \frac{\varepsilon}{4N} + \sum_{i} \frac{\varepsilon}{2(b-a)} |x_i - x_{i-1}| \leq \varepsilon$$

3. The constant function $1 \in \mathcal{R}(\alpha)$ on $[a, b]$ and

$$\int_{a}^{b} d\alpha = \alpha(b) - \alpha(a)$$

This is true for any function $\alpha : [a, b] \to \mathbb{R}$, because

$$S(\mathbf{P}, T, 1, \alpha) = \sum_{i=1}^{n} 1 [\alpha(x_i) - \alpha(x_{i-1})] = \alpha(b) - \alpha(a)$$

for any partition $\mathbf{P}$ and any choice $T$. 