Properties of the Riemann–Stieltjes Integral

Theorem (Linearity Properties)
Let $a < c < d < b$ and $A, B \in \mathbb{R}$ and $f, g, \alpha, \beta : [a,b] \to \mathbb{R}$.

(a) If $f, g \in \mathcal{R}(\alpha)$ on $[a,b]$, then $Af + Bg \in \mathcal{R}(\alpha)$ on $[a,b]$ and

$$\int_a^b \big[Af + Bg\big]\, d\alpha = A\int_a^b f\, d\alpha + B\int_a^b g\, d\alpha$$

(b) If $f \in \mathcal{R}(\alpha) \cap \mathcal{R}(\beta)$ on $[a,b]$, then $f \in \mathcal{R}(A\alpha + B\beta)$ on $[a,b]$ and

$$\int_a^b f\, d(A\alpha + B\beta) = A\int_a^b f\, d\alpha + B\int_a^b f\, d\beta$$

(c) If $f \in \mathcal{R}(\alpha)$ on $[a,c]$ and on $[c,b]$, then $f \in \mathcal{R}(\alpha)$ on $[a,b]$ and

$$\int_a^b f\, d\alpha = \int_a^c f\, d\alpha + \int_c^b f\, d\alpha$$

(d) If $f \in \mathcal{R}(\alpha)$ on $[a,b]$ then $f \in \mathcal{R}(\alpha)$ on $[c,d] \subset [a,b]$.

Proof: (a) Let $\varepsilon > 0$. Then, for any partition $\mathbb{P}$ of $[a,b]$ and choice $\mathbb{T}$ for $\mathbb{P}$,

$$\begin{aligned}
&\left| S(\mathbb{P}, \mathbb{T}, Af + Bg, \alpha) - A\int_a^b f\, d\alpha - B\int_a^b g\, d\alpha \right| \\
&\qquad = \left| A\, S(\mathbb{P}, \mathbb{T}, f, \alpha) + B\, S(\mathbb{P}, \mathbb{T}, g, \alpha) - A\int_a^b f\, d\alpha - B\int_a^b g\, d\alpha \right| \\
&\qquad \le |A| \left| S(\mathbb{P}, \mathbb{T}, f, \alpha) - \int_a^b f\, d\alpha \right| + |B| \left| S(\mathbb{P}, \mathbb{T}, g, \alpha) - \int_a^b g\, d\alpha \right|
\end{aligned}$$

Assume that A and B are nonzero. (The cases that A and/or B are zero are similar, but easier.) Since $f \in \mathcal{R}(\alpha)$ on $[a,b]$ there is a partition $\mathbb{P}_{f,\varepsilon}$ such that

$$\left| S(\mathbb{P}, \mathbb{T}, f, \alpha) - \int_a^b f\, d\alpha \right| < \frac{\varepsilon}{2|A|} \qquad \text{whenever } \mathbb{P} \supset \mathbb{P}_{f,\varepsilon}$$

and since $g \in \mathcal{R}(\alpha)$ on $[a,b]$ there is a partition $\mathbb{P}_{g,\varepsilon}$ such that

$$\left| S(\mathbb{P}, \mathbb{T}, g, \alpha) - \int_a^b g\, d\alpha \right| < \frac{\varepsilon}{2|B|} \qquad \text{whenever } \mathbb{P} \supset \mathbb{P}_{g,\varepsilon}$$

It now suffices to set $\mathbb{P}_\varepsilon = \mathbb{P}_{f,\varepsilon} \cup \mathbb{P}_{g,\varepsilon}$ and observe that

$$\left| S(\mathbb{P}, \mathbb{T}, Af + Bg, \alpha) - A\int_a^b f\, d\alpha - B\int_a^b g\, d\alpha \right| < \varepsilon \qquad \text{whenever } \mathbb{P} \supset \mathbb{P}_\varepsilon$$

(b) See Problem Set 1, #3.

(c) See Problem Set 1, #2.

(d) See Problem Set 2, #3.

■

Theorem (Integration by Parts)
Let $a < b$ and $f, \alpha : [a,b] \to \mathbb{R}$. If $f \in \mathcal{R}(\alpha)$ on $[a,b]$, then $\alpha \in \mathcal{R}(f)$ on $[a,b]$ and

$$\int_a^b f(x)\,d\alpha(x) + \int_a^b \alpha(x)\,df(x) = f(b)\alpha(b) - f(a)\alpha(a)$$

Remark.

(a) The integration by parts formula may also be written

$$\int_a^b f\,d\alpha + \int_a^b \alpha\,df = \int_a^b d(f\alpha)$$

(b) We shall later see that if α has a continuous first derivative, then $\int_a^b f(x)\,d\alpha(x) = \int_a^b f(x)\alpha'(x)\,dx$. So if both f and α have continuous first derivatives, we may write the integration by parts formula as

$$\int_a^b f(x)\alpha'(x)\,dx = f(b)\alpha(b) - f(a)\alpha(a) - \int_a^b \alpha(x)f'(x)\,dx$$

which is the integration by parts formula of first year calculus (though you probably used $f(x) = u(x)$ and $\alpha(x) = v(x)$.

Proof of integration by parts: Our goal is to show that $\int_a^b \alpha\,df$ exists and takes the value

$$\int_a^b \alpha\,df = f(b)\alpha(b) - f(a)\alpha(a) - \int_a^b f\,d\alpha$$

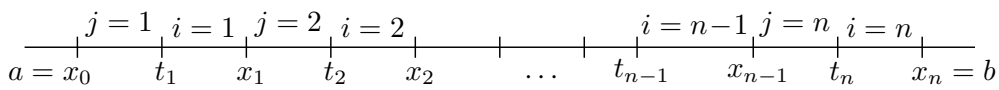
So let's look at the difference between a Riemann sum for $\int_a^b \alpha\,df$ and the right hand side. For any partition $\mathbb{P} = \big\{a = x_0\,,\,x_1\,,\,x_2\,,\,x_3\,,\,\cdots\,,\,x_n = b\big\}$ of $[a,b]$ and choice $\mathbb{T} = \big\{t_1\,,\,t_2\,,\,t_3\,,\,\cdots\,,\,t_n\big\}$ for $\mathbb{P}$,

$$\begin{aligned}
&S\big(\mathbb{P},\mathbb{T},\alpha,f\big) - \Big\{f(b)\alpha(b) - f(a)\alpha(a) - \int_a^b f\,d\alpha\Big\} \\
&\quad = \sum_{i=1}^n \alpha(t_i)\big[f(x_i) - f(x_{i-1}\big] - \sum_{i=1}^n \underbrace{\alpha(x_i)f(x_i)}_{\substack{\alpha(b)f(b)\\ \text{for } i=n}} + \sum_{i=1}^n \underbrace{\alpha(x_{i-1})f(x_{i-1})}_{\substack{\alpha(a)f(a)\\ \text{for } i=1}} + \int_a^b f\,d\alpha \\
&\qquad\qquad \text{(The } 1 \le i \le n-1 \text{ terms of the second sum cancel the} \\
&\qquad\qquad\qquad 2 \le i \le n \text{ terms of the third sum.)} \\
&\quad = -\sum_{i=1}^n f(x_i)\big[\alpha(x_i) - \alpha(t_i)\big] - \sum_{j=1}^n f(x_{j-1})\big[\alpha(t_j) - \alpha(x_{j-1})\big] + \int_a^b f\,d\alpha
\end{aligned}$$

Note that

- $f(x_i)\big[\alpha(x_i)-\alpha(t_i)\big]$ looks like a term in a Riemann sum approximation to $\int_a^b f\,d\alpha$ with subinterval $[t_i,x_i]$ and choice point $x_i\in[t_i,x_i]$ and
- $f(x_{j-1})\big[\alpha(t_j)-\alpha(x_{j-1})\big]$ looks like a term in a Riemann sum approximation to $\int_a^b f\,d\alpha$ with subinterval $[x_{j-1},t_j]$ and choice point $x_{j-1}\in[x_{j-1},t_j]$.

Here is a figure which shows all of these subintervals.



So all of the subintervals fit together perfectly to form the partition[1]

$$\mathbb{P}\cup\mathbb{T}=\big\{x_0\,,\,t_1\,,\,x_1\,,\,t_2\,,\,x_2\,,\,t_3\,,\,\cdots\,,\,x_{n-1}\,,\,t_n\,,\,x_n\big\}$$

and

$$S\big(\mathbb{P},\mathbb{T},\alpha,f\big)-\Big\{f(b)\alpha(b)-f(a)\alpha(a)-\int_a^b f\,d\alpha\Big\}$$
$$=-S\big(\mathbb{P}\cup\mathbb{T},\mathbb{T}',f,\alpha\big)+\int_a^b f\,d\alpha$$

where the choice

$$\mathbb{T}'=\big\{\overbrace{x_0}^{j=1},\overbrace{x_1}^{i=1},\overbrace{x_1}^{j=2},\overbrace{x_2}^{i=2},\overbrace{x_2}^{j=3},\overbrace{x_3}^{i=3},\cdots,\overbrace{x_{n-1}}^{i=n-1},\overbrace{x_{n-1}}^{j=n},\overbrace{x_n}^{i=n}\big\}$$

Now let $\varepsilon>0$. Since $f\in\mathcal{R}(\alpha)$ for $[a,b]$ there is a partition $\mathbb{P}_\varepsilon$ such that

$$\Big|S\big(\tilde{\mathbb{P}},\tilde{\mathbb{T}},f,\alpha\big)-\int_a^b f\,d\alpha\Big|<\varepsilon$$

for all partitions $\tilde{\mathbb{P}}$ finer than $\mathbb{P}_\varepsilon$. If the partitition $\mathbb{P}$ above is finer than $\mathbb{P}_\varepsilon$ then the partition $\mathbb{P}\cup\mathbb{T}$ is also finer than $\mathbb{P}_\varepsilon$ and we have

$$\Big|S\big(\mathbb{P},\mathbb{T},\alpha,f\big)-\Big\{f(b)\alpha(b)-f(a)\alpha(a)-\int_a^b f\,d\alpha\Big\}\Big|=\Big|S\big(\mathbb{P}\cup\mathbb{T},\mathbb{T}',f,\alpha\big)-\int_a^b f\,d\alpha\Big|<\varepsilon$$

■

Theorem (The Change of Variables $x=g(y)$)

Let

- *$a<b$ and $c<d$,*
- *$g:[c,d]\to[a,b]$ be continuous, strictly monotonically increasing, and obey $g(c)=a$ and $g(d)=b$ and*
- *$f,\alpha:[a,b]\to\mathbb{R}$.*

Set

$$h(y)=f\big(g(y)\big)\qquad\beta(y)=\alpha\big(g(y)\big)$$

If $f\in\mathcal{R}(\alpha)$ on $[a,b]$, then $h\in\mathcal{R}(\beta)$ on $[c,d]$ and

$$\int_c^d h(y)\,d\beta(y)=\int_a^b f(x)\,d\alpha(x)$$

[1] There is a subtlety hidden in these definitions. We are not allowed to use subintervals of width zero. So if, for example, $t_i=x_i$, we merge the two subintervals $[x_{i-1},t_i]$, $[t_i,x_i]$ into the single subinterval $[x_{i-1},x_i]$ and use $f(x_{i-1})[\alpha(t_i)-\alpha(x_{i-1})]+f(x_i)[\alpha(x_i)-\alpha(t_i)]=f(x_{i-1})[\alpha(x_i)-\alpha(x_{i-1})]$.

Proof: Our goal is to prove that $\int_c^d h(y)\,d\beta(y)$ exists and equals $\int_a^b f(x)\,d\alpha(x)$, so let's consider the difference between a Riemann sum for $\int_c^d h(y)\,d\beta(y)$ and $\int_a^b f(x)\,d\alpha(x)$. For any partition $\mathbb{P} = \{c = y_0, \cdots, y_n = d\}$ of $[c,d]$ and any choice $\mathbb{T} = \{s_1, \cdots, s_n\}$ for $\mathbb{P}$

$$\begin{aligned}
\left| S(\mathbb{P}, \mathbb{T}, h, \beta) - \int_a^b f\,d\alpha \right| &= \left| \sum_{i=1}^n h(s_i)\big[\beta(y_i) - \beta(y_{i-1})\big] - \int_a^b f\,d\alpha \right| \\
&= \left| \sum_{i=1}^n f\big(g(s_i)\big)\big[\alpha\big(g(y_i)\big) - \alpha\big(g(y_{i-1})\big)\big] - \int_a^b f\,d\alpha \right| \\
&= \left| S(g(\mathbb{P}), g(\mathbb{T}), f, \alpha) - \int_a^b f\,d\alpha \right|
\end{aligned}$$

where

$$\begin{aligned}
g(\mathbb{P}) &= \big\{\ g(y)\ \big|\ y \in \mathbb{P}\ \big\} \\
&= \big\{ \overbrace{g(y_0)}^{x_0} = g(c) = a\ ,\ \overbrace{g(y_1)}^{x_1}\ , \cdots ,\ \overbrace{g(y_n)}^{x_n} = g(d) = b \big\}
\end{aligned}$$

is a partition of $[a,b]$ because g is assumed to be strictly monotonic, so that

$$y_{i-1} < y_i \implies x_{i-1} < x_i$$

and is assumed to obey $x_0 = g(y_0) = a$ and $x_n = g(y_n) = b$ and

$$g(\mathbb{T}) = \big\{ \overbrace{g(s_1)}^{t_1}\ , \cdots ,\ \overbrace{g(s_n)}^{t_n} \big\}$$

is a choice for $g(\mathbb{P})$ because g is assumed to be strictly monotonic so that

$$y_{i-1} \le s_i \le y_i \implies x_{i-1} = g(y_{i-1}) \le g(s_i) = t_i \le g(y_i) = x_i$$

Now let $\varepsilon > 0$.

- We have assumed that $f \in \mathcal{R}(\alpha)$ on $[a,b]$. So there is a partition $\mathbb{P}_{f,\varepsilon}$ of $[a,b]$ such that $\mathbb{P}_f \supset \mathbb{P}_{f,\varepsilon} \implies \left| S(\mathbb{P}_f, \mathbb{T}_f, f, \alpha) - \int_a^b f\,d\alpha \right| < \varepsilon$ for all choices $\mathbb{T}_f$ for $\mathbb{P}_f$.
- The assumptions that we have made on g guarantee that the inverse function $g^{-1} : [a,b] \to [c,d]$ exists and that $g^{-1}\big(\mathbb{P}_{f,\varepsilon}\big)$ is a partition of $[c,d]$. We choose $\mathbb{P}_\varepsilon = g^{-1}\big(\mathbb{P}_{f,\varepsilon}\big)$.

Then

$$\begin{aligned}
&\mathbb{P} \supset \mathbb{P}_\varepsilon \implies g(\mathbb{P}) \supset g(\mathbb{P}_\varepsilon) = \mathbb{P}_{f,\varepsilon} \\
&\implies \left| S(\mathbb{P}, \mathbb{T}, h, \beta) - \int_a^b f\,d\alpha \right| = \left| S(\ \overbrace{g(\mathbb{P})}^{\substack{\text{finer} \\ \text{than } \mathbb{P}_{f,\varepsilon}}}\ ,\ \overbrace{g(\mathbb{T})}^{\substack{\text{allowed} \\ \mathbb{T}_f}}, f, \alpha) - \int_a^b f\,d\alpha \right| < \varepsilon
\end{aligned}$$

as desired.

■

Theorem (Second Fundamental Theorem of Calculus) *Let $a < b$ and $f : [a,b] \to \mathbb{R}$. Assume that*

- *f is differentiable on $[a,b]$ and*
- *$f' \in \mathcal{R}$ on $[a,b]$*

Then

$$\int_a^b f'(x)\,dx = f(b) - f(a)$$

Proof: Let $\mathbb{P} = \{x_0, x_1, x_2, \cdots, x_n\}$ be any partition of $[a,b]$. Then, by the mean value theorem, there exists, for each $1 \le i \le n$ a $t_i \in [x_{i-1}, x_i]$ with

$$f(x_i) - f(x_{i-1}) = f'(t_i)\big(x_i - x_{i-1}\big)$$

So, setting $\mathbb{T} = \{t_1, t_2, \cdots, t_n\}$, we have

$$S(\mathbb{P}, \mathbb{T}, f') = \sum_{i=1}^n f'(t_i)\big(x_i - x_{i-1}\big) = \sum_{i=1}^n \big[f(x_i) - f(x_{i-1})\big] = f(b) - f(a)$$

So now we just have to apply the definition of "$f' \in \mathcal{R}$ on $[a,b]$". ■

Theorem (Basic Bounds)
Let $a < b$ and $f, g, \alpha : [a,b] \to \mathbb{R}$. Assume that $f, g \in \mathcal{R}(\alpha)$ on $[a,b]$ and α is monotonically increasing.

(a) If $f(x) \le g(x)$ for all $x \in [a,b]$, then

$$\int_a^b f\,d\alpha \le \int_a^b g\,d\alpha$$

(a) If $|f(x)| \le g(x)$ for all $x \in [a,b]$, then

$$\left|\int_a^b f\,d\alpha\right| \le \int_a^b g\,d\alpha$$

Proof: Let $\varepsilon > 0$. Since $f \in \mathcal{R}(\alpha)$ on $[a,b]$ there is a partition $\mathbb{P}_{f,\varepsilon}$ such that

$$\left|S(\mathbb{P}, \mathbb{T}, f, \alpha) - \int_a^b f\,d\alpha\right| < \varepsilon$$

whenever $\mathbb{P} \supset \mathbb{P}_{f,\varepsilon}$ and $\mathbb{T}$ is a choice for $\mathbb{P}$. Since $g \in \mathcal{R}(\alpha)$ on $[a,b]$ there is a partition $\mathbb{P}_{g,\varepsilon}$ such that

$$\left|S(\mathbb{P}, \mathbb{T}, g, \alpha) - \int_a^b g\,d\alpha\right| < \varepsilon$$

whenever $\mathbb{P} \supset \mathbb{P}_{g,\varepsilon}$ and $\mathbb{T}$ is a choice for $\mathbb{P}$. Set $\mathbb{P}_\varepsilon = \{a = x_0, x_1, x_2, \cdots, x_n = b\} = \mathbb{P}_{f,\varepsilon} \cup \mathbb{P}_{g,\varepsilon}$ and let $\mathbb{T} = \{t_1, t_2, \cdots, t_n\}$ be a choice for $\mathbb{P}_\varepsilon$.

(a) We have

$$\begin{aligned}
\int_a^b f\,d\alpha &\le S(\mathbb{P}_\varepsilon, \mathbb{T}, f, \alpha) + \varepsilon \\
&= \sum_{i=1}^n f(t_i)\big[\alpha(x_i) - \alpha(x_{i-1})\big] + \varepsilon \\
&\le \sum_{i=1}^n g(t_i)\big[\alpha(x_i) - \alpha(x_{i-1})\big] + \varepsilon \\
&\qquad\qquad \text{(since } f(t_i) \le g(t_i) \text{ and } \alpha(x_i) - \alpha(x_{i-1}) \ge 0) \\
&= S(\mathbb{P}_\varepsilon, \mathbb{T}, g, \alpha) + \varepsilon \\
&\le \int_a^b g\,d\alpha + 2\varepsilon
\end{aligned}$$

As $\int_a^b f\,d\alpha \le \int_a^b g\,d\alpha + 2\varepsilon$ is true for all $\varepsilon > 0$, we also have $\int_a^b f\,d\alpha \le \int_a^b g\,d\alpha$.

(b) We have

$$\begin{aligned}
\left|\int_a^b f\,d\alpha\right| &\le \big|S(\mathbb{P}_\varepsilon, \mathbb{T}, f, \alpha)\big| + \varepsilon \\
&\le \sum_{i=1}^n \big|f(t_i)\big|\big[\alpha(x_i) - \alpha(x_{i-1})\big] + \varepsilon \\
&\le \sum_{i=1}^n g(t_i)\big[\alpha(x_i) - \alpha(x_{i-1})\big] + \varepsilon \\
&\qquad\qquad \text{(since } |f(t_i)| \le g(t_i) \text{ and } \alpha(x_i) - \alpha(x_{i-1}) \ge 0) \\
&= S(\mathbb{P}_\varepsilon, \mathbb{T}, g, \alpha) + \varepsilon \\
&\le \int_a^b g\,d\alpha + 2\varepsilon
\end{aligned}$$

As $\left|\int_a^b f\,d\alpha\right| \le \int_a^b g\,d\alpha + 2\varepsilon$ is true for all $\varepsilon > 0$, we also have $\left|\int_a^b f\,d\alpha\right| \le \int_a^b g\,d\alpha$.

■

Theorem (First Fundamental Theorem of Calculus) *Let $a < b$ and $f : [a, b] \to \mathbb{R}$. Assume that*

- *$f \in \mathcal{R}$ on $[a, b]$*

Set, for $a \le x \le b$,

$$F(x) = \int_a^x f(x)\,dx$$

Then

(a) F is continuous on $[a, b]$ and

(b) if f is continuous at $x_0 \in [a, b]$, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

Proof: (a) Since $f \in \mathcal{R}$ it is bounded. Suppose that $|f(t)| \le M$ for all $a \le t \le b$. Then

$$\big|F(y) - F(x)\big| = \left| \int_x^y f(t)\,dt \right| \le M|y-x|$$

so F is uniformly continous.
(b) Let f be continuous at $x_0 \in [a,b]$. Then, for all $a \le x \le b$

$$\begin{aligned}
\left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| &= \left| \frac{\int_{x_0}^x f(t)\,dt - f(x_0)[x - x_0]}{x - x_0} \right| \\
&= \left| \frac{1}{x - x_0} \int_{x_0}^x \big[f(t) - f(x_0)\big]\,dt \right| \\
&\le \sup_{\substack{t \text{ between} \\ x_0 \text{ and } x}} \big|f(t) - f(x_0)\big|
\end{aligned}$$

Since f is continuous at x_0, the right hand side converges to zero as $x \to x_0$. ■