5 Theory of Riemann–Stieltjes Integrals

The main goal of this chapter is to construct Riemann–Stieltjes integrals, which is a generalization of Riemann integrals. If you have already seen the construction of Riemann integrals, you will notice that most of the steps and properties are similar, but with some subtilites that you have to be careful with. Otherwise, you will see how Riemann–Stieltjes integrals are specialized to Riemann integrals, and that it does not cost much to consider this more general theory.

5.1 Functions of bounded variation

In this section, we are going to define functions of bounded variation defined on a segment $[a, b] \subseteq \mathbb{R}$ for $a < b$. In Section 5.1.2, we will introduce the notion of *partitions* that will allow us to define the *total variation* of a function. Before closing the section, Theorem 5.1.17 will give us an important and useful characterization of functions with bounded variation.

5.1.1 Reminders on monotonic functions

Definition 5.1.1 : Let $I \subseteq \mathbb{R}$ be an interval, and $f : I \to \mathbb{R}$ be a function. We say that f is

(1) non-increasing (非遞增) or decreasing (遞減) if $f(x) \geqslant f(y)$ for all $x, y \in I$ with $x \leqslant y$;

(2) non-decreasing (非遞減) or increasing (遞增) if $f(x) \leqslant f(y)$ for all $x, y \in I$ with $x \leqslant y$;

(3) monotonic (單調) if one of the above is satisfied.

Definition 5.1.2 : Let $I \subseteq \mathbb{R}$ be an interval, $f : I \to \mathbb{R}$ be a monotonic function, and $x \in I$. We may define the *left limit* (左極限) and the *right limit* (右極限) of f at x as below,

$$f(x-) := \lim_{\substack{y \to x \\ y < x}} f(y), \quad \text{and} \quad f(x+) := \lim_{\substack{y \to x \\ y > x}} f(y).$$

The left limit $f(x-)$ is well defined if $(x - \varepsilon, x) \cap I$ is nonempty for all $\varepsilon > 0$. Similarly, the right limit $f(x+)$ is well defined if $(x, x + \varepsilon) \cap I$ is nonempty for all $\varepsilon > 0$.

Proposition 5.1.3 : *Let $f : [a, b] \to \mathbb{R}$ be a monotonic function. Then, the set of its discontinuities D is countable.*

Proof : We have already shown this in Exercise 1.15. We give a quick sketch of the proof below. Whether f is continuous or not at a and b, the countability of D does not change. Therefore, it is enough to look at the discontinuities of f on (a, b). Let us define

$$D = \{x \in (a, b) : f(x-) \neq f(x+)\}.$$

Without loss of generality, we may assume that f is non-decreasing. For every given $x \in D$, we have $f(x-) < f(x+)$, and the density of $\mathbb{Q}$ in $\mathbb{R}$ implies the existence of $q_x \in \mathbb{Q} \cap (f(x-), f(x+))$. Then, the map $D \to \mathbb{Q}$, $x \mapsto q_x$ is injective, so D is countable. (Corollary 1.4.9) □

5.1.2 Partitions and functions of bounded variation

Below, let us consider $a < b$ and real-valued functions defined on the segment $[a, b]$.

Definition 5.1.4 : Let $[a, b] \subseteq \mathbb{R}$ be a segment.

- A *partition* or a *subdivision* (分割) of the segment $[a, b]$ is a finite sequence $P = (x_k)_{0 \leqslant k \leqslant n}$ satisfying $a = x_0 < x_1 < \cdots < x_n = b$.
- Given a partition $P = (x_k)_{0 \leqslant k \leqslant n}$, its length is denoted by n, the points $x_0, \ldots, x_n$ are called *subdivision points* of P, and $\operatorname{Supp}(P) = \{x_k : 0 \leqslant k \leqslant n\}$ is called the *support* (支集) of P.
- Given a finite subset $A \subseteq [a, b]$ containing a and b, there exists a unique partition P such that $\operatorname{Supp}(P) = A$. It is called the *partition corresponding to* A.
- For $1 \leqslant k \leqslant n$, the segment $[x_{k-1}, x_k]$ is called the k-th *subinterval* of P, and we write $\Delta x_k = x_k - x_{k-1}$.
- The *mesh size* (網格大小) of a partition P is defined by $\|P\| := \max_{1 \leqslant k \leqslant n}(x_k - x_{k-1})$.
- Given two partitions P and P' of $[a, b]$. P' is said to be *finer* than P, denoted $P \subseteq P'$ or $P' \supseteq P$, if $\operatorname{Supp}(P) \subseteq \operatorname{Supp}(P')$. This also implies $\|P'\| \leqslant \|P\|$.
- For two partitions P_1 and P_2 of $[a, b]$, we may define their *joint partition* (聯集分割), or *smallest common refinement* (最小共同分割), denoted $P := P_1 \vee P_2$, which is the partition corresponding to the support $\operatorname{Supp}(P_1) \cup \operatorname{Supp}(P_2)$. Note that P is finer than both P_1 and P_2.
- We write $\mathcal{P}([a, b])$ for the collection of all possible partitions of $[a, b]$.

Remark 5.1.5 : If $P = (x_0, \ldots, x_n)$ is a partition of $[a, b]$, we have $b - a = \sum_{k=1}^{n} \Delta x_k$.

Definition 5.1.6 : Let $f : [a, b] \to \mathbb{R}$ be a function on $[a, b]$. If $P = (x_0, \ldots, x_n)$ is a partition of $[a, b]$, we may write $\Delta f_k = f(x_k) - f(x_{k-1})$ for $1 \leqslant k \leqslant n$ and define

$$V_P(f) := \sum_{k=1}^{n} |\Delta f_k|.$$

We say that f is of *bounded variation* (有界變差) on $[a, b]$ if

$$V_f = V_f([a, b]) := \sup_{P \in \mathcal{P}([a,b])} V_P(f) < \infty.$$

The quantity $V_f([a, b])$ is called the *total variation* (總變差) of f on $[a, b]$. And we write $\mathcal{BV}([a, b], \mathbb{R})$ or $\mathcal{BV}([a, b])$ for the collection of functions on $[a, b]$ of bounded variation.

Example 5.1.7 : Consider the function $f : [0, 1] \to \mathbb{R}$ defined as follows,

$$f(x) = \begin{cases} x \cos(\frac{\pi}{x}) & \text{if } x \in (0, 2\pi], \\ 0 & \text{if } x = 0. \end{cases}$$

For an integer $n \geqslant 1$, let P be the partition corresponding to the finite set

$$\Big\{0, \frac{1}{2n}, \frac{1}{2n-1}, \dots, \frac{1}{2}, 1\Big\}.$$

That is, $x_0 = 0$, and $x_k = \frac{1}{2n+1-k}$ for $1 \leqslant k \leqslant 2n$. Then, we have

$$\begin{aligned} V_P(f) = \sum_{k=1}^{2n} |\Delta f_k| &= \left|\frac{(-1)^{2n}}{2n} - 0\right| + \sum_{k=2}^{2n} \left|\frac{(-1)^{k-1}}{2n+1-k} - \frac{(-1)^k}{2n+2-k}\right| \\ &= \frac{1}{2n} + \sum_{k=2}^{2n} \left(\frac{1}{2n+1-k} + \frac{1}{2n+2-k}\right) \\ &= 1 + \sum_{k=2}^{2n-1} \frac{2}{k} + \frac{1}{2n}. \end{aligned}$$

Due to the harmonic series in the above formula, we know that the sum is not bounded. This allows us to conclude that the above function f is not of bounded variation.

Proposition 5.1.8 : *Let $f : [a, b] \to \mathbb{R}$ be a function on $[a, b]$ with bounded variation. Then, the following properties hold.*

(1) For any partitions $P \subseteq P'$, we have $V_P(f) \leqslant V_{P'}(f)$.

(2) For any $\varepsilon > 0$, there exists a partition $P_\varepsilon \in \mathcal{P}([a, b])$ such that for any finer partition $P \supseteq P_\varepsilon$, the following holds

$$V_P(f) \leqslant V_f \leqslant V_P(f) + \varepsilon.$$

Proof :

(1) Let $P = (x_0, \dots, x_n) \in \mathcal{P}([a, b])$ be a partition of $[a, b]$. By induction, it is sufficient to show the inequality in the case where P' is a partition with one more subdivision point than P in the support. Let us assume that P' is the partition whose support is given by $\mathrm{Supp}(P) \cup \{c\}$ with

$c \in (x_{i-1}, x_i)$ for some $1 \leqslant i \leqslant n$. We have

$$\begin{aligned} V_{P'}(f) &= \sum_{\substack{k=1 \\ k \neq i}}^{n} |f(x_k) - f(x_{k-1})| + |f(c) - f(x_{i-1})| + |f(x_i) - f(c)| \\ &\geqslant \sum_{\substack{k=1 \\ k \neq i}}^{n} |f(x_k) - f(x_{k-1})| + \big|(f(c) - f(x_{i-1})) + (f(x_i) - f(c))\big| \\ &= \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| = V_P(f). \end{aligned}$$

(2) Let $\varepsilon > 0$. By the characterization of supremum, we may find a partition $P_\varepsilon \in \mathcal{P}([a, b])$ such that

$$V_f \leqslant V_{P_\varepsilon}(f) + \varepsilon.$$

Then, for any partition $P \supseteq P_\varepsilon$, we find from (1) that

$$V_f \leqslant V_{P_\varepsilon}(f) + \varepsilon \leqslant V_P(f) + \varepsilon.$$

□

5.1.3 Examples of bounded variation functions

Below we are going to discuss some criteria so that a function f defined on a segment $[a, b]$ is of bounded variation.

Proposition 5.1.9 : *If $f : [a, b] \to \mathbb{R}$ is monotonic, then $f \in \mathcal{BV}([a, b])$ and $V_f = |f(b) - f(a)|$.*

Proof : Without loss of generality, by replacing f with $-f$, we may assume that f is non-decreasing. For any given partition $P = (x_0, \dots, x_n) \in \mathcal{P}([a, b])$, we have

$$V_P(f) = \sum_{k=1}^{n} |\Delta f_k| = \sum_{k=1}^{n} \Delta f_k = \sum_{k=1}^{n} [f(x_k) - f(x_{k-1})] = f(b) - f(a).$$

□

Proposition 5.1.10 : *If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, differentiable on (a, b) with bounded derivative, then $f \in \mathcal{BV}([a, b])$.*

Proof : Let $P = (x_0, \dots, x_n) \in \mathcal{P}([a, b])$ be a partition of $[a, b]$. For $1 \leqslant k \leqslant n$, we may apply the mean-value theorem (Section 4.1.2) on the k-th subinterval of P, and find

$$\Delta f_k = f(x_k) - f(x_{k-1}) = f'(t_k)(x_k - x_{k-1}), \quad \text{where } t_k \in (x_{k-1}, x_k).$$

This implies

$$V_P(f) = \sum_{k=1}^{n} |\Delta f_k| = \sum_{k=1}^{n} |f'(t_k)|\Delta x_k \leqslant \sup_{t\in[a,b]} |f'(t)| \cdot \sum_{k=1}^{n} \Delta x_k = \sup_{t\in[a,b]} |f'(t)| \cdot (b-a).$$

□

Proposition 5.1.11 : *If $f : [a,b] \to \mathbb{R}$ is a function of bounded variation, then it is also bounded. In other words, the inclusion relation $\mathcal{BV}([a,b]) \subseteq \mathcal{B}([a,b])$ holds.*

Proof : Let $M = V_f([a,b])$. For a given $x \in (a,b)$, we may consider a specific partition given by $P = (a,x,b)$. We have

$$|f(x) - f(a)| + |f(b) - f(x)| \leqslant M,$$

which leads to $|f(x)| - |f(a)| \leqslant |f(x) - f(a)| \leqslant M$, that is $|f(x)| \leqslant M + |f(a)|$. □

5.1.4 Properties

Proposition 5.1.12 : *Let $f, g : [a,b] \to \mathbb{R}$ be functions of bounded variation. Then, the functions $f+g$, $f-g$ and fg are also of bounded variation.*

Proof : For any partition $P \in \mathcal{P}([a,b])$, by the triangular inequality, we have

$$V_P(f \pm g) \leqslant V_P(f) + V_P(g).$$

Therefore, the total variation of $f \pm g$ satisfies

$$\begin{aligned} V_{f\pm g} &= \sup_{P\in\mathcal{P}([a,b])} V_P(f \pm g) \leqslant \sup_{P\in\mathcal{P}([a,b])} \left[V_P(f) + V_P(g)\right] \\ &\leqslant \sup_{P\in\mathcal{P}([a,b])} V_P(f) + \sup_{P\in\mathcal{P}([a,b])} V_P(g) = V_f + V_g. \end{aligned}$$

For the product $h := fg$, let us be given a partition $P = (x_0, \ldots, x_n) \in \mathcal{P}([a,b])$. For $1 \leqslant k \leqslant n$, we have

$$\begin{aligned} |\Delta h_k| &= |f(x_k)g(x_k) - f(x_{k-1})g(x_{k-1})| \\ &\leqslant |f(x_k) - f(x_{k-1})||g(x_k)| + |g(x_k) - g(x_{k-1})||f(x_{k-1})| \\ &\leqslant |\Delta f_k| \cdot \sup|g| + |\Delta g_k| \cdot \sup|f|. \end{aligned}$$

By summing the above inequality over k, we find

$$V_P(h) \leqslant V_P(f) \cdot \sup|g| + V_P(g) \cdot \sup|f|.$$

Therefore, the total variation $V_h = V_{fg}$ satisfies

$$V_{fg} \leqslant V_f \cdot \sup|g| + V_g \cdot \sup|f|.$$

□

Proposition 5.1.13 : *Let $f : [a,b] \to \mathbb{R}$ be a function of bounded variation such that $|f| \geqslant m$ for some constant $m > 0$. Then, $g = \frac{1}{f}$ is also a function of bounded variation.*

Proof : For any partition $P = (x_0, \ldots, x_n) \in \mathcal{P}([a,b])$, we have

$$|\Delta g_k| = \left|\frac{1}{f(x_k)} - \frac{1}{f(x_{k-1})}\right| = \left|\frac{\Delta f_k}{f(x_k)f(x_{k-1})}\right| \leqslant \frac{|\Delta f_k|}{m^2}.$$

Then, it follows that

$$V_g \leqslant \frac{V_f}{m^2}.$$

□

Proposition 5.1.14 : *Let $f : [a,b] \to \mathbb{R}$ be a function of bounded variation and $c \in (a,b)$. Then, f is of bounded variation on $[a,c]$ and on $[c,b]$, and we have*

$$V_f([a,b]) = V_f([a,c]) + V_f([c,b]).$$

Proof : First, let us show that $V_f([a,c])$ and $V_f([c,b])$ are well defined, that is f is of bounded variation on both $[a,c]$ and $[c,b]$. Let $P_1 = (x_0, \ldots, x_n) \in \mathcal{P}([a,c])$ and $P_2 = (y_0, \ldots, y_m) \in \mathcal{P}([c,b])$ be partitions. Then, $P = P_1 \vee P_2 = (x_0, \ldots, x_n = y_0, y_1, \ldots, y_m)$ is a partition of $[a,b]$. And it follows that

$$V_{P_1}(f) + V_{P_2}(f) = V_P(f) \leqslant V_f([a,b]). \tag{5.1}$$

The above relation shows that $f \in \mathcal{BV}([a,c])$ and $f \in \mathcal{BV}([c,b])$.

Next, by taking the suprema $P_1 \in \mathcal{P}([a,c])$ and $P_2 \in \mathcal{P}([c,b])$ in Eq. (5.1), we find

$$V_f([a,c]) + V_f([c,b]) \leqslant V_f([a,b]).$$

Next, let us show the reverse inequality. We are given a partition $P = (x_0, \ldots, x_n) \in \mathcal{P}([a,b])$, and want to construct another partition $P' \in \mathcal{P}([a,b])$ from P, so that c is a subdivision point of P'. If $c \in P$, we take $P' = P$; otherwise, there exists a unique m such that $c \in (x_{m-1}, x_m)$, and we define $P' = (y_0, \ldots, y_{n+1})$ as below

$$y_k = \begin{cases} x_k & \text{if } k \leqslant m-1, \\ c & \text{if } k = m, \\ x_{k-1} & \text{if } k \geqslant m+1. \end{cases}$$

Moreover, we may divide P' into two partitions $P_1 = (x_0, \ldots, x_{m-1}, c) \in \mathcal{P}([a,c])$ and $P_2 = (c, x_m, \ldots, x_n) \in \mathcal{P}([c,b])$. We also note that $|f(x_m) - f(x_{m-1})| \leqslant |f(x_m) - f(c)| + |f(c) - f(x_{m-1})|$, so

$$V_P(f) \leqslant V_{P'}(f) = V_{P_1}(f) + V_{P_2}(f) \leqslant V_f([a,c]) + V_f([c,b]).$$

By taking the supremum over $P \in \mathcal{P}([a,b])$, we find

$$V_f([a,b]) \leqslant V_f([a,c]) + V_f([c,b]).$$

□

Definition 5.1.15 : Let $f : [a,b] \to \mathbb{R}$ be a function of bounded variation. We define its *variation function* $V : [a,b] \to \mathbb{R}$ as follows,

$$V(x) = \begin{cases} V_f([a,x]) & \text{if } x \in (a,b], \\ 0 & \text{if } x = a. \end{cases} \tag{5.2}$$

Lemma 5.1.16 : *Let $f : [a,b] \to \mathbb{R}$ be a function of bounded variation and V be its variation function defined in Eq. (5.2). Then, both V and $V - f$ are non-decreasing functions on $[a,b]$.*

Proof : For $x > a$, we clearly have $V(x) \geqslant 0 = V(a)$. For x, y such that $b \geqslant y > x > a$, we have

$$V(y) - V(x) = V_f([a,y]) - V_f([a,x]) = V_f([x,y]) \geqslant 0.$$

So we can conclude that V is a non-decreasing function on $[a,b]$.

Let $D := V - f$. For x, y such that $b \geqslant y > x \geqslant a$, we have

$$D(y) - D(x) = V_f([x,y]) - [f(y) - f(x)].$$

By considering the trivial partition $P = (x,y) \in \mathcal{P}[x,y]$, we see that $|f(y) - f(x)| \leqslant V_f([x,y])$. Therefore, $D(y) - D(x) \geqslant 0$. □

The following decomposition theorem gives us a characterization of functions of bounded variation.

Theorem 5.1.17 : *Let $f : [a,b] \to \mathbb{R}$ be a function. Then, the following two properties are equivalent.*

(a) f is of bounded variation.

(b) There exist two non-decreasing functions g_1 and g_2 so that $f = g_1 - g_2$.

Proof :

- (b) ⇒ (a). It follows from Proposition 5.1.9 that monotonic functions are of bounded variation, then it follows from Proposition 5.1.12 that so is their difference.
- (a) ⇒ (b). We use the variation function V as defined in Eq. (5.2). We know that V and $V - f$ are non-decreasing, so we may conclude by writing $f = V - (V - f)$.

□

Remark 5.1.18 : We note that in Theorem 5.1.17, there is no uniqueness in (b). If g_1 and g_2 are non-decreasing functions such that $f = g_1 - g_2$, then by taking an arbitrary non-decreasing function h, we also have the decomposition $f = (g_1 + h) - (g_2 + h)$.

Proposition 5.1.19 : *Let $f : [a,b] \to \mathbb{R}$ be a function of bounded variation on $[a,b]$. For $x \in [a,b]$, the function f is continuous at x if and only if V is continuous at x.*

Proof : By Definition 5.1.2, we know that for any monotonic function $g : [a,b] \to \mathbb{R}$, its left limit $g(x-)$ at every $x \in (a,b]$ exists; similarly, its right limit $g(x+)$ at every $x \in [a,b)$ also exists. Moreover, thanks to Theorem 5.1.17, we know that both $f(x-)$ and $V(x-)$ are well defined for $x \in (a,b]$; and both $f(x+)$ and $V(x+)$ are well defined for $x \in [a,b)$.

- Let $x \in [a,b)$ and suppose that V is right continuous at x. We want to show that f is right continuous at x. For any $y \in (x,b]$, we have

  $$0 \leqslant |f(y) - f(x)| \leqslant V_f([x,y]) = V(y) - V(x),$$

  where the last equality comes from Proposition 5.1.14. By taking $y \to x+$, we find

  $$0 \leqslant |f(x+) - f(x)| \leqslant V(x+) - V(x),$$

  so the right continuity of V at x implies that of f. The same also holds for the left continuity in the case that $x \in (a,b]$.

- Let $x \in [a,b)$ and suppose that f is right continuous at x. Given $\varepsilon > 0$, take $\delta > 0$ such that

  $$\forall y \in [x, x+\delta), \quad |f(y) - f(x)| \leqslant \varepsilon. \tag{5.3}$$

  Moreover, it follows from Proposition 5.1.8 that we may find a partition $P_\varepsilon \in \mathcal{P}([x,b])$ such that

  $$V_f([x,b]) \leqslant V_P(f) + \varepsilon, \tag{5.4}$$

  for any partition $P \supseteq P_\varepsilon$. We may take a partition $P = (x_k)_{0 \leqslant k \leqslant n} \supseteq P_\varepsilon$ such that $x_1 \in [x, x+\delta)$, so that Eq. (5.3) is satisfied. This means that $|\Delta f_1| \leqslant \varepsilon$. Then, Eq. (5.4) rewrites as below,

  $$V_f([x,b]) \leqslant V_P(f) + \varepsilon \leqslant 2\varepsilon + \sum_{k=2}^{n} |\Delta f_k| \leqslant 2\varepsilon + V_f([x_1,b]),$$

  where the last equality comes from the fact that $(x_1, \ldots, x_n)$ is a partition of $[x_1,b]$. Thus, we find

  $$\begin{aligned} 0 \leqslant V(x_1) - V(x) &= V_f([a,x_1]) - V_f([a,x]) = V_f([x,x_1]) \\ &= V_f([x,b]) - V_f([x_1,b]) \leqslant 2\varepsilon. \end{aligned}$$

  This shows that

  $$\forall y \in [x, x+\delta), \quad 0 \leqslant V(y) - V(x) \leqslant 2\varepsilon.$$

  Since $\varepsilon > 0$ can be arbitrarily small, we conclude that $V(x+) = V(x)$, that is V is continuous from the right at x. The proof is similar for the left continuity, so we omit here.

□

Theorem 5.1.20 : *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then, the following two properties are equivalent.*

(a) f is of bounded variation.

(b) There exist two non-decreasing continuous functions g_1 and g_2 so that $f = g_1 - g_2$.

Proof : This theorem is a direct consequence of the decomposition theorem (Theorem 5.1.17) and the fact that the function f and its variation function V share the same continuities (Proposition 5.1.19). □

5.2 Riemann-Stieltjes integrals

In this section, we will discuss the theory of Riemann–Stieltjes integrals. We first construct such integrals using Riemann–Stieltjes sums, which generalize the notion of Riemann sums, see Definition 5.2.1. In the rest of Section 5.2.1, we talk about useful properties of Riemann–Stieltjes integrals, which share many similarities with Riemann integrals. In Section 5.2.2, we look at the special case of step functions as integrators, which can be related to discrete sums, see Corollary 5.2.21 and Corollary 5.2.23. In Section 5.3, we define the notion of Darboux summations for Riemann–Stieltjes integrals, which allow us to have a better characterization for the existence of such integrals. This is known as Riemann's condition, see Definition 5.3.8 and Theorem 5.3.10.

5.2.1 Definition and properties

Below, let us take $a < b$ and consider the segment $[a, b]$ in $\mathbb{R}$. The functions f, g, α, β that we are going to consider are all bounded real-valued functions defined on $[a, b]$. In other words, $f, g, \alpha, \beta \in \mathcal{B}([a, b], \mathbb{R})$.

Definition 5.2.1 : Let $P = (x_0, \dots, x_n)$ be a partition of $[a, b]$. For $1 \leqslant k \leqslant n$, let $t_k \in [x_{k-1}, x_k]$, and write $t = (t_1, \dots, t_n)$, called *tagged points*. The pair (P, t) is called a *tagged partition*. We define the following *Riemann–Stieltjes sum* of f with respect to α,

$$S_{P,t}(f, \alpha) = \sum_{k=1}^{n} f(t_k)\Delta\alpha_k. \tag{5.5}$$

We say that f is *Riemann–Stieltjes integrable*, or simply *integrable*, with respect to α on $[a, b]$, if the following property is satisfied.

(RS) There exists $L \in \mathbb{R}$ such that for every $\varepsilon > 0$, there exists a partition $P_\varepsilon \in \mathcal{P}([a, b])$, such that for every $P \in \mathcal{P}([a, b])$ with $P \supseteq P_\varepsilon$, and every choice of tagged points t, we have

$$|S_{P,t}(f, \alpha) - L| \leqslant \varepsilon.$$

If f is Riemann–Stieltjes integrable with respect to α on $[a, b]$, we write $f \in R(\alpha; a, b)$ or $f \in R(\alpha)$ if the dependency on the segment $[a, b]$ is clear from the context.

Remark 5.2.2 :

(1) If f is Riemann–Stieltjes integrable with respect to α on $[a,b]$, then the constant $L \in \mathbb{R}$ in the property (RS) is unique.

(2) The unique value of $L \in \mathbb{R}$ satisfying (RS) is denoted by

$$\int_a^b f \,\mathrm{d}\alpha \quad \text{or} \quad \int_a^b f(x)\,\mathrm{d}\alpha(x). \tag{5.6}$$

It is called the Riemann–Stieltjes integral.

(3) The function f is called *integrand* (被積分函數) and α called *integrator* (積分函數) .

(4) If $\alpha(x) = x$, then the integral in Eq. (5.6) is called Riemann integral. The set of Riemann-integrable functions is denoted by $R(x; a, b)$ or $R(x)$.

(5) In Definition 5.1.4, a partition $P = (x_k)_{0 \leqslant k \leqslant n}$ needs to satisfy $x_{k-1} < x_k$ for all $1 \leqslant k \leqslant n$. In fact, this condition can be relaxed to $x_{k-1} \leqslant x_k$ because if there exists some k with $x_{k-1} = x_k$, the corresponding term in the Riemann–Stieltjes sum Eq. (5.5) does not have any contribution. This allows us to have redundant points in a partition. As a consequence, the above notions can also be defined in the case when the domain of the functions is reduced to a singleton $[a,a] = \{a\}$.

(6) If the integrand $f : [a,b] \to V$ takes values in a finite-dimensional vector space V, by considering a basis $(e_1, \dots, e_d)$ of V, we may rewrite f as $f = \sum_{i=1}^d f_i e_i$, where $f_i : [a,b] \to \mathbb{R}$ is a real-valued function. Then, we may define the Riemann–Stieltjes integral of f with respect to α coordinate-wise, that is

$$\int_a^b f\,\mathrm{d}\alpha := \sum_{i=1}^d \Big(\int_a^b f_i\,\mathrm{d}\alpha\Big) e_i,$$

provided that each of the Riemann–Stieltjes integral $\int_a^b f_i\,\mathrm{d}\alpha$ is well defined. This is the reason why we may generalize everything easily to $\mathbb{C} \cong \mathbb{R}^2$ or $\mathbb{R}^d$ for any $d \geqslant 1$.

Example 5.2.3 :

(1) If $\alpha : [a,b] \to \mathbb{R}$ is a constant function, then for any bounded function $f : [a,b] \to \mathbb{R}$, any partition $P \in \mathcal{P}([a,b])$ and any tagged points t, the Riemann–Stieltjes sum always satisfies $S_{P,t}(f,\alpha) = 0$. Therefore, we have $f \in R(\alpha; a, b)$ and $\int_a^b f\,\mathrm{d}\alpha = 0$.

(2) In the case of the Riemann integral, we know from the first-year calculus that for any continuous function $f : [a,b] \to \mathbb{R}$, we have $f \in R(x)$. This is also a consequence of Theorem 5.3.21.

(3) Let $\alpha : [-1,1] \to \mathbb{R}$ be defined by $\alpha(x) = \mathbb{1}_{x \geqslant 0}$ and $f = \alpha$. Consider a partition $P \in \mathcal{P}([a,b])$ such that $x_k = 0 \in P$ for some $k \in \mathbb{N}$, and tagged points t, so the corresponding Riemann–Stieltjes sum writes

$$S_{P,t}(f,\alpha) = f(t_k)\big[\alpha(0) - \alpha(x_{k-1})\big] = f(t_k),$$

where $x_{k-1} \leqslant t_k \leqslant x_k = 0$. This sum depends on the choice of the tagged point t_k,

$$f(t_k) = \begin{cases} 1 & \text{if } t_k = x_k = 0, \\ 0 & \text{otherwise,} \end{cases}$$

so cannot satisfy (RS).

Lemma 5.2.4 : *We follow the notations in Definition 5.2.1. Let us define the following condition.*

(RS') There exists $L \in \mathbb{R}$ such that for every $\varepsilon > 0$, there exists $\delta > 0$ such that for any partition $P \in \mathcal{P}([a,b])$ with $\|P\| < \delta$, and every choice of tagged points t, we have

$$|S_{P,t}(f,\alpha) - L| \leqslant \varepsilon.$$

If f satisfies (RS'), then f also satisfies (RS).

Proof : Suppose that f and α are such that (RS') holds. Let $\varepsilon > 0$, we take $\delta > 0$ such that (RS') is satisfied for any $P \in \mathcal{P}([a,b])$ with $\|P\| < \delta$ and any choice of tagged points t. Then, let $P_\varepsilon \in \mathcal{P}([a,b])$ be any partition such that $\|P_\varepsilon\| < \delta$. It is clear that (RS) is satisfied for this choice of P_ε, because any finer partition $P \supseteq P_\varepsilon$ also satisfies $\|P\| < \delta$. □

Remark 5.2.5 : Note that the converse of Lemma 5.2.4 is false in general. Let us consider the following functions $f, \alpha : [0,1] \to \mathbb{R}$, defined by

$$\alpha(x) = \mathbb{1}_{x \geqslant \frac{1}{2}} \quad \text{and} \quad f(x) = \mathbb{1}_{x > \frac{1}{2}}.$$

- (RS) is satisfied, see Exercise 5.14 or Theorem 5.2.20.
- (RS') is not satisfied. Indeed, for any given $\delta \in (0,1)$, we may consider a partition $P = (x_k)_{0 \leqslant k \leqslant n} \in \mathcal{P}([0,1])$, defined by

  $$x_0 = 0, x_n = 1 \quad \text{and} \quad x_{k-1} = \frac{1}{2}(1-\delta), x_k = \frac{1}{2}(1+\delta) \text{ for some } 1 \leqslant k \leqslant n.$$

  Then, for any tagged points t, we have

  $$S_{P,t}(f,\alpha) = f(t_k)\big[\alpha(x_k) - \alpha(x_{k-1})\big] = f(t_k) = \mathbb{1}_{t_k > \frac{1}{2}},$$

  whose value depends on whether $t_k \in [x_{k-1}, \frac{1}{2}]$ or $t_k \in (\frac{1}{2}, x_k]$.

Remark 5.2.6 : We are able to show that in the case of Riemann-integrability, that is when $\alpha(x) = x$, (RS) implies (RS'). It is a direct consequence of Theorem 5.3.10, see Exercise 5.29.

Proposition 5.2.7 : *Let $\alpha : [a,b] \to \mathbb{R}$ be a bounded function and $f, g \in R(\alpha)$. Then, for any constant $c \in \mathbb{R}$, we have $f + cg \in R(\alpha)$ and*

$$\int_a^b (f + cg)\,\mathrm{d}\alpha = \int_a^b f\,\mathrm{d}\alpha + c\int_a^b g\,\mathrm{d}\alpha.$$

In other words, $R(\alpha)$ is a vector space over $\mathbb{R}$, and the integral operator $f \mapsto \int_a^b f\,\mathrm{d}\alpha$ is a linear form on $R(\alpha)$, that is an element of $\mathcal{L}(R(\alpha), \mathbb{R})$.

Proof : Let $c \in \mathbb{R}$ and $h = f + cg$. For any partition $P \in \mathcal{P}([a,b])$ and tagged points t, we have

$$\begin{aligned} S_{P,t}(h,\alpha) &= \sum_{k=1}^{n} h(t_k)\Delta\alpha_k = \sum_{k=1}^{n} f(t_k)\Delta\alpha_k + c\sum_{k=1}^{n} g(t_k)\Delta\alpha_k \\ &= S_{P,t}(f,\alpha) + c\, S_{P,t}(g,\alpha). \end{aligned}$$

Let $\varepsilon > 0$. Since $f \in R(\alpha)$, we may find a partition $P'_\varepsilon \in \mathcal{P}([a,b])$ such that for every partition $P \supseteq P'_\varepsilon$ and tagged points t, we have

$$\left| S_{P,t}(f,\alpha) - \int_a^b f \,\mathrm{d}\alpha \right| \leqslant \varepsilon.$$

Similarly, we may find a partition $P''_\varepsilon \in \mathcal{P}([a,b])$ such that for every partition $P \supseteq P''_\varepsilon$ and tagged points t, we have

$$\left| S_{P,t}(g,\alpha) - \int_a^b g \,\mathrm{d}\alpha \right| \leqslant \varepsilon.$$

To conclude, we define $P_\varepsilon := P'_\varepsilon \vee P''_\varepsilon$, then for any $P \supseteq P_\varepsilon$, it follows from the inequalities above that

$$\begin{aligned} &\left| S_{P,t}(h,\alpha) - \int_a^b f \,\mathrm{d}\alpha - c\int_a^b g \,\mathrm{d}\alpha \right| \\ &\leqslant \left| S_{P,t}(f,\alpha) - \int_a^b f \,\mathrm{d}\alpha \right| + |c| \left| S_{P,t}(g,\alpha) - \int_a^b g \,\mathrm{d}\alpha \right| \\ &\leqslant (1+|c|)\varepsilon. \end{aligned}$$

□

Proposition 5.2.8 : *Let $\alpha, \beta : [a,b] \to \mathbb{R}$ be two bounded functions and $f \in R(\alpha) \cap R(\beta)$. Then, for any $c \in \mathbb{R}$, we have $f \in R(\alpha + c\beta)$ and*

$$\int_a^b f \,\mathrm{d}(\alpha + c\beta) = \int_a^b f \,\mathrm{d}\alpha + c\int_a^b f \,\mathrm{d}\beta.$$

Proof : You may follow similar steps as in the proof of Proposition 5.2.7. See Exercise 5.17. □

Definition 5.2.9 : For $a < b$, any bounded function $\alpha : [a,b] \to \mathbb{R}$ and $f \in R(\alpha; a, b)$, we may define

$$\int_b^a f \,\mathrm{d}\alpha = \int_b^a f(x) \,\mathrm{d}\alpha(x) := -\int_a^b f(x) \,\mathrm{d}\alpha(x), \tag{5.7}$$

and we write $R(\alpha; b, a) = R(\alpha; a, b)$. We also define $\int_a^a f \,\mathrm{d}\alpha = 0$ for any function f defined at a.

Proposition 5.2.10 : *Let $I \subseteq \mathbb{R}$ be a segment and $a, b, c \in I$. Let $\alpha : I \to \mathbb{R}$ be a bounded function and $f \in R(\alpha; a, b) \cap R(\alpha; b, c)$. Then, $f \in R(\alpha; a, c)$ and we have*

$$\int_a^b f \,\mathrm{d}\alpha + \int_b^c f \,\mathrm{d}\alpha + \int_c^a f \,\mathrm{d}\alpha = 0$$

Proof : Without loss of generality, we may assume that $a < b < c$. Let $\varepsilon > 0$. Since $f \in R(\alpha; a, b)$, we may find a partition $P_\varepsilon^{(1)} \in \mathcal{P}([a, b])$ such that for any finer partition $P^{(1)} \supseteq P_\varepsilon^{(1)}$ with tagged points $t^{(1)}$, we have

$$\left| S_{P^{(1)},t^{(1)}}(f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| < \frac{\varepsilon}{2}. \tag{5.8}$$

Similarly, we may also find a partition $P_\varepsilon^{(2)} \in \mathcal{P}([b, c])$ such that for any finer partition $P^{(2)} \supseteq P_\varepsilon^{(2)}$ with tagged points $t^{(2)}$, we have

$$\left| S_{P^{(2)},t^{(2)}}(f, \alpha) - \int_b^c f \, \mathrm{d}\alpha \right| < \frac{\varepsilon}{2}. \tag{5.9}$$

Let us set $P_\varepsilon := P_\varepsilon^{(1)} \vee P_\varepsilon^{(2)}$.

Let $P \in \mathcal{P}([a, c])$ be a partition of $[a, c]$ that is finer than P_ε, and t be tagged points. Let

$$P^{(1)} = P \cap [a, b] \quad \text{and} \quad P^{(2)} = P \cap [b, c]$$

denote the corresponding restricted partitions on $[a, b]$ and $[b, c]$, so that we also have $P^{(1)} \supseteq P_\varepsilon^{(1)}$ and $P^{(2)} \supseteq P_\varepsilon^{(2)}$. We also denote $t^{(1)}$ and $t^{(1)}$ to be the corresponding restricted tagged points. Then, we have the following relation between the Riemann–Stieltjes sums,

$$S_{P,t}(f, \alpha) = S_{P^{(1)},t^{(1)}}(f, \alpha) + S_{P^{(2)},t^{(2)}}(f, \alpha). \tag{5.10}$$

It follows from Eq. (5.8), Eq. (5.9), Eq. (5.10) and the triangle inequality that

$$\left| S_{P,t}(f, \alpha) - \int_a^b f \, \mathrm{d}\alpha - \int_b^c f \, \mathrm{d}\alpha \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This implies that $f \in R(\alpha; a, c)$ and

$$\int_a^b f \, \mathrm{d}\alpha + \int_b^c f \, \mathrm{d}\alpha = \int_a^c f \, \mathrm{d}\alpha$$

To conclude, we use the notation from Eq. (5.7). □