

Riemann-Stieltjes Integral

S Kumaresan
kumaresa@gmail.com

27 January 2021

Contents

1 Darboux Approach to Riemann Integration	2
2 Darboux Approach to Riemann-Stieltjes Integral	4
3 Riemann's Criterion for Integrability	6
4 Examples of Integrable Functions	8
5 Riemann's Approach to Integration	12
6 Class of Integrable Functions	13
7 Two Important Examples of Riemann-Stieltjes Integrals	19
8 Fundamental Theorems of Calculus	22

Abstract

We treat the case of Riemann integrals first and motivate the results geometrically. When we prove the results analytically, we observe that the proof can be adopted to prove the corresponding results in Riemann-Stieltjes integrals where the integrator is an increasing function. The aim of this article is to give an outline of the theory of Riemann-Stieltjes integral closely following the development of Riemann integral in our book [1]. The approach may develop geometric ideas of Riemann integrals and hone the analytical skills.

You may also like to watch the videos in [5] on which this article is based.

1 Darboux Approach to Riemann Integration

Definition 1. A *partition* P of an interval $[a, b]$ is a finite set $\{x_0, x_1, \dots, x_n\}$ such that $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$. The points x_i are called the *nodes* of P .

Example 2. (i) $P = \{a = x_0, x_1 = b\}$ is the trivial partition of $[a, b]$.

(ii) For any $n \in \mathbb{N}$, let $x_i = a + \frac{i}{n}(b - a)$ for $0 \leq i \leq n$. Then $\{x_0, \dots, x_n\}$ is a partition, say, P_n . Note that P_n divides $[a, b]$ in subintervals of equal length. (One often says that P_n divides $[a, b]$ into equal parts.) \square

Given two partitions P and Q of $[a, b]$, we say that Q is a *refinement* of P if $P \subset Q$. In the example (ii) above, the partition P_{2k+1} is a refinement of P_{2k} . More generally, P_{mn} is a refinement of P_n where $m, n \in \mathbb{N}$.

If $P = \{a = s_0, \dots, s_m = b\}$ and $Q = \{a = t_0, \dots, t_n = b\}$ are two partitions of $[a, b]$, let us consider $P \cup Q$. It is a finite set of points in $[a, b]$. So we can arrange them in increasing order $\{x_0, \dots, x_N\}$. (Here $x_0 = s_0 = t_0 = a$ and $x_N = s_m = t_n = b$.) This is a partition of $[a, b]$ and by an abuse of notation it is denoted by $P \cup Q$. Note that $P \cup Q$ is a refinement of both P and Q . For example, if $P = \{0, 1/2, 3/4, 1\}$ and $Q = \{0, 1/3, 1/2, 2/3, 1\}$, then $P \cup Q = \{0, 1/3, 1/2, 2/3, 3/4, 1\}$.

Definition 3. Given $f: [a, b] \rightarrow \mathbb{R}$ and a partition $P = \{x_0, \dots, x_n\}$ of $[a, b]$, we let

$$L(f, P) := \sum_{i=0}^{n-1} m_i(f)(x_{i+1} - x_i), \quad \text{where } m_i(f) := \text{GLB} \{f(x) : x \in [x_{i-1}, x_i]\}$$

$$U(f, P) := \sum_{i=0}^{n-1} M_i(f)(x_{i+1} - x_i), \quad \text{where } M_i(f) := \text{LUB} \{f(x) : x \in [x_{i-1}, x_i]\}$$

Observe that, if $f \geq 0$, $L(f, P)$ (respectively, $U(f, P)$) is the sum of areas of rectangles inscribed inside (respectively, circumscribing) the region bounded by the curves $x = a$, $x = b$, $y = 0$ and $y = f(x)$. For simplicity, we refer to this region as the region under the graph in the case of nonnegative functions. The numbers $L(f, P)$ and $U(f, P)$ are respectively called the *lower and the upper Darboux sums* of f with respect to the partition P . They approximate the area under the graph from below and from above.

See Figure 6.3 on page 177 in [1].

We let $m = m(f) := \text{GLB} \{f(x) : x \in [a, b]\}$ and $M = M(f) := \text{LUB} \{f(x) : x \in [a, b]\}$.

We now look at some examples of upper and lower sums. They also teach us the importance of choosing partitions smartly depending on the nature of the functions.

Example 4 (Constant functions). Let $f: [a, b] \rightarrow \mathbb{R}$ be a constant, say, c . Let P be any partition. Then we see that $m_i = M_i = c$. Hence $L(f, P) = \sum_{i=1}^n c(x_i - x_{i-1}) = c(b - a)$. Similarly $U(f, P) = c(b - a)$. This validates our intuition that the "area" under the graph should be $c(b - a)$.

Example 5 (Step Functions). Let $f: [-1, 1] \rightarrow \mathbb{R}$ be defined by $f(x) = 1$ for $-1 \leq x < 0$, $f(0) = 10$ and $f(x) = 2$ for $0 < x \leq 1$. Draw picture of the graph of f . Can you convince yourself that the area under graph should be 3? Our intuition says that the point $x = 0$ needs special attention. Let us consider the partition $P_n := \{-1, x_1 = -\frac{1}{n}, x_2 = \frac{1}{n}, 1\}$. We see that $m_1 = M_1 = 1$, $m_2 = 1$ and $M_2 = 10$ and $m_3 = M_3 = 2$. Hence we find that

$$L(f, P_n) = 1 \left(1 - \frac{1}{n}\right) + 1 \times \frac{2}{n} + 2 \left(1 - \frac{1}{n}\right) = 3 \left(1 - \frac{1}{n}\right) + \frac{2}{n}$$

$$U(f, P_n) = 1 \left(1 - \frac{1}{n}\right) + 10 \times \frac{2}{n} + 2 \left(1 - \frac{1}{n}\right) = 3 \left(1 - \frac{1}{n}\right) + 10 \times \frac{2}{n}.$$

Do you observe that $L(f, P_n) \rightarrow 3$ and $U(f, P_n) \rightarrow 3$? □

Example 6 (Monotone Functions). Let $f: [a, b] \rightarrow \mathbb{R}$ be increasing. Let P be any partition. Then we have $m_i = f(x_{i-1})$ and $M_i = f(x_i)$ for each i . Hence we see that

$$L(f, P) = \sum_i f(x_{i-1})(x_i - x_{i-1}) \quad \text{and} \quad U(f, P) = \sum_i f(x_i)(x_i - x_{i-1})$$

(If f is decreasing what are the lower and upper sums?) If we take finer partitions do the upper and lower sums come close to each other? That is, we ask: can we make $U(f, P) - L(f, P)$ arbitrarily small? Note that

$$U(f, P) - L(f, P) = \sum_{i=1}^n [f(x_i) - f(x_{i-1})](x_i - x_{i-1}).$$

Note that if all the lengths $x_i - x_{i-1}$ of the subintervals are equal, say, $(b - a)/n$, then we have

$$U(f, P) - L(f, P) = \frac{b - a}{n} \sum_{i=1}^n [f(x_i) - f(x_{i-1})] = \frac{b - a}{n} (f(b) - f(a)). \quad (1)$$

In the last equality we used the fact that the sum

$$\begin{aligned} \sum_i [f(x_i) - f(x_{i-1})] &= [f(x_1) - f(x_0)] + [f(x_2) - f(x_1)] + \cdots + [f(x_n) - f(x_{n-1})] \\ &= -f(x_0) + f(x_n), \end{aligned}$$

a telescopic sum. Is our question above answered? □

These examples lead us to believe that as we take finer and finer partitions, we seem to approximate the desired area well.

Example 7 (Dirichlet's Function). Let $f: [0, 1] \rightarrow \mathbb{R}$ be the function $f(x) = 1$ if $x \in \mathbb{Q} \cap [0, 1]$ and $f(x) = 0$ if $x \in [0, 1] \setminus \mathbb{Q}$. Let P be any partition. Then $m_i = 0$ and $M_i = 1$ thanks to the density of rational numbers as well as the irrational numbers. Hence we find that

$$L(f, P) = \sum_i 0 \times (x_i - x_{i-1}) = 0 \quad \text{and} \quad U(f, P) = \sum_i 1 \times (x_i - x_{i-1}) = 1.$$

This is true for any partition P . What do we infer? If at all there is meaning to the area under the graph of the function, our approach is not going to give approximations to the 'area'. So there are two possible policy decisions we have to take. We love our approach and so we declare that the area of the region under the graph of f has no meaning. Or, we should devise a better system. We adopt the first and leave the second to a later course! \square

Ex. 8. Prove that $m(b - a) \leq L(f, P) \leq U(f, P) \leq M(b - a)$ for any partition P of $[a, b]$.

Ex. 9. Let $f: [a, b] \rightarrow \mathbb{R}$ be bounded. Let $c \in (a, b)$. Let P_1 (respectively, P_2) be a partition of $[a, c]$ (respectively, of $[c, b]$). Show that $P := P_1 \cup P_2$ is a partition of $[a, b]$. Show also that $L(f, P) = L(f, P_1) + L(f, P_2)$ and $U(f, P) = U(f, P_1) + U(f, P_2)$.

A very pedantic formulation is as follows. Let f_1 (respectively, f_2) be the restriction of f to $[a, c]$ (respectively, to $[c, b]$). Prove that $L(f, P) = L(f_1, P_1) + L(f_2, P_2)$ and so on.

Definition 10. Given a partition $P = \{x_0, \dots, x_n\}$, we insert a new node, say, t such that $x_i < t < x_{i+1}$ for some i and get a new partition Q . Then drawing pictures of a non-negative function, it is clear that $L(f, Q) \geq L(f, P)$. Similarly, $U(f, Q) \leq U(f, P)$. Draw pictures. See Figures 6.4–6.5 on Page 178 in [1]. (We shall prove this later.) Thus, Q produces a better approximation to the area bound by the graph. This suggests that to get the "real" area we should look at

$$L(f) := \text{LUB} \{L(f, P) : P \text{ is a partition of } [a, b]\}$$

$$U(f) := \text{GLB} \{U(f, P) : P \text{ is a partition of } [a, b]\}.$$

These numbers exist (why?) and are called respectively the *lower and upper integral* of f on $[a, b]$. The upper integral of f on $[a, b]$ may be understood as the best possible approximation to the area of the region under the graph as approximated from above. How do we understand the lower integral?

Ex. 11. Show that $L(f) \leq U(f)$ for any bounded function $f: [a, b] \rightarrow \mathbb{R}$.

Definition 12. We say that f is *Darboux integrable* (or *simply integrable*) on $[a, b]$ if the upper and lower integrals coincide. (This intuitively says that we require that the area should be approximable both from below and from above.) If f is integrable, the common value of the upper and lower integrals is denoted by the symbol $\int_a^b f(x) dx$. This is just a notation; we may as well use $I_a^b(f)$ or $\int_a^b f$ etc.

2 Darboux Approach to Riemann-Stieltjes Integral

Let $\alpha: [a, b] \rightarrow \mathbb{R}$ be an increasing (non-decreasing) function. Let $f: [a, b] \rightarrow \mathbb{R}$ be a bounded function. Let $P := \{x_0, x_1, \dots, x_n\}$ be a partition of $[a, b]$. Let $m_i \equiv m_i(f)$ and $M_i = M_i(f)$ be as usual. Let $\Delta\alpha_i := \alpha(x_i) - \alpha(x_{i-1})$. Note that $\Delta\alpha_i \geq 0$. Also $\sum_{i=1}^n \Delta\alpha_i = \alpha(b) - \alpha(a)$. (Verify this.)

We define upper and lower Riemann-Stieltjes sums of f relative to the partition P in an obvious way:

$$U(f, P, \alpha) := \sum_{i=1}^n M_i \Delta\alpha_i \quad \text{and} \quad L(f, P, \alpha) := \sum_{i=1}^n m_i \Delta\alpha_i.$$

Since $m_i \leq M_i$ and $\Delta\alpha_i \geq 0$, we see that

$$L(f, P, \alpha) := \sum_{i=1}^n m_i \Delta\alpha_i \leq \sum_{i=1}^n M_i \Delta\alpha_i = U(f, P, \alpha).$$

Let m and M be such that $m \leq f(x) \leq M$ for $x \in [a, b]$. Then $m \leq m_i$ and $M_i \leq M$ for all i . Hence it follows that

$$m(\alpha(b) - \alpha(a)) \leq L(f, P, \alpha) \leq U(f, P, \alpha) \leq M(\alpha(b) - \alpha(a)).$$

In particular, the set $\{U(f, P, \alpha)\}$ is bounded below and hence its GLB exists. In a similar way, $\{L(f, P, \alpha)\}$ is bounded above and hence its LUB exists. As in the case of Riemann integral, let $U(f, \alpha) := \text{GLB} \{U(f, P, \alpha)\}$ be the upper α -integral of f on $[a, b]$. How is $L(f, \alpha)$ defined? What do you call it?

The following theorem is a collection of results which are proved by mimicking the proofs of the corresponding results in the theory of Riemann integration.

Theorem 13. *The following are true:*

- (i) *If Q is a refinement of the partition P , then $L(f, P, \alpha) \leq L(f, Q, \alpha)$ and $U(f, P, \alpha) \geq U(f, Q, \alpha)$.*
- (ii) *For any partitions P and Q of $[a, b]$, we have $L(f, P, \alpha) \leq U(f, Q, \alpha)$.*
- (iii) *We have $L(f, \alpha) \leq U(f, \alpha)$.*

Proof. Let $P := \{x_0, \dots, x_{i-1}, x_i, \dots, x_n\}$ be a partition. Let $t \in (x_{i-1}, x_i)$ and

$$Q := P \cup \{t\} = \{x_0, x_1, \dots, x_{i-1}, t, x_i, \dots, x_n\}.$$

Note that $M_i = \text{LUB} \{f(x) : x \in [x_{i-1}, x_i]\} \geq \text{LUB} \{f(x) : x \in [x_{i-1}, t]\}$ and $M_i = \text{LUB} \{f(x) : x \in [x_{i-1}, x_i]\} \geq \text{LUB} \{f(x) : x \in [t, x_i]\}$. Hence we find that

$$\begin{aligned} M_i(\Delta\alpha_i) &= M_i(\alpha(t) - \alpha(x_{i-1})) + M_i(\alpha(x_i) - \alpha(t)) \\ &\geq \text{LUB} \{f(x) : x \in [x_{i-1}, t]\}(\alpha(t) - \alpha(x_{i-1})) \\ &\quad + \text{LUB} \{f(x) : x \in [t, x_i]\}(\alpha(x_i) - \alpha(t)). \end{aligned}$$

Can you state and prove the analogous results for $m_i \Delta\alpha_i$? We then have

$$\begin{aligned} U(f, P, \alpha) &= \sum_{j \neq i} M_j \Delta\alpha_j + M_i \Delta\alpha_i \\ &\geq \sum_{j \neq i} M_j \Delta\alpha_j + \text{LUB} \{f(x) : x \in [x_{i-1}, t]\}(\alpha(t) - \alpha(x_{i-1})) \\ &\quad + \text{LUB} \{f(x) : x \in [t, x_i]\}(\alpha(x_i) - \alpha(t)) \\ &= U(f, Q, \alpha). \end{aligned}$$

The reader should prove that $L(f, P, \alpha) \leq L(f, Q, \alpha)$. Let $Q = P \cup \{t_1, \dots, t_r\}$. Let $Q_i := P \cup \{t_1, \dots, t_i\}$, $1 \leq i \leq r$. Let $Q_0 = P$. Each Q_i is a refinement of Q_{i-1} . Hence $U(f, Q_r, \alpha) \leq$

$U(f, Q_{r-1}, \alpha) \leq \dots \leq U(f, Q_1, \alpha) \leq U(f, P, \alpha)$. The analogous result for lower sum follows in a similar fashion.

(ii) We know that $L(f, P, \alpha) \leq U(f, P, \alpha)$ for any partition P . Let $P' = P \cup Q$. Observe that

$$L(f, P, \alpha) \leq L(f, P', \alpha) \leq U(f, P', \alpha) \leq U(f, Q, \alpha).$$

(Can you justify the first and the third inequalities?)

(iii) is an easy exercise in LUB and GLB. Consider the two subsets A and B of real numbers. Assume that for any pair of points $(x, y) \in A \times B$, $x \leq y$. Then each y is an upper bound of A . Hence $\text{LUB } A \leq y$. Since this holds true for each $y \in B$, we see that $\text{LUB } A$ is a lower bound for B . Hence $\text{LUB } A \leq \text{GLB } B$. Let $A := \{L(f, P, \alpha) : P \text{ a partition of } [a, b]\}$ and $B := \{U(f, Q, \alpha) : Q \text{ a partition of } [a, b]\}$. Then A and B satisfy our assumption in view of (ii). The result (iii) follows. \square

Definition 14. Keep the notation above. We say that f is α -integrable if $L(f, \alpha) = U(f, \alpha)$. We also say that the Riemann-Stieltjes integral of f with respect to the *integrator* α exists. The common value of the upper and lower integrals is denoted by $\int_a^b f d\alpha$.

Let us repeat: if $\alpha(x) = x$ then we simply say that f is integrable on $[a, b]$.

Note that when α is the identity function, this means that we can approximate the area under the graph of a positive function both from inside and outside and get the same result. Thus, this definition insists on a kind of "symmetry".

3 Riemann's Criterion for Integrability

Theorem 15 (Riemann's Criterion). f is α -integrable iff for any $\varepsilon > 0$ there exists a partition P such that $U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon$.¹

Proof. Let the condition be satisfied. We are required to prove that f is α -integrable. Let I_1 and I_2 be the lower and upper integrals. We need to show $I_1 = I_2$. It is enough to show that for any $\varepsilon > 0$, $|I_1 - I_2| < \varepsilon$.

Since $I_1 \leq I_2$, it is enough to show that $I_2 < I_1 + \varepsilon$ for any $\varepsilon > 0$. Given $\varepsilon > 0$, let P be as in the hypothesis. Observe that

$$I_1 \geq L(f, P, \alpha) \text{ and } I_2 \leq U(f, P, \alpha).$$

Hence

$$I_2 - I_1 \leq U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon.$$

Thus we have proved that f α -is integrable.

¹See the video in [5]: Riemann Integration - 3, especially 30:57 – 46:25.

To prove the converse, let $\varepsilon > 0$ be given. Let I_1 and I_2 be the lower and upper integrals. Then we have $I_1 = I_2$. Since I_1 is the LUB of $L(f, P, \alpha)$'s, there exists a partition P_1 of $[a, b]$ such that $I_1 - \varepsilon/2 < L(f, P_1, \alpha)$. That is,

$$I_1 - L(f, P_1, \alpha) < \varepsilon/2. \quad (2)$$

Similarly, there exists a partition P_2 of $[a, b]$ such that

$$U(f, P_2, \alpha) < I_2 + \varepsilon/2 \text{ so that } U(f, P_2, \alpha) - I_2 < \varepsilon/2. \quad (3)$$

Let $P = P_1 \cup P_2$. Since P is a refinement of P_1 and P_2 , we have

$$L(f, P_1, \alpha) \leq L(f, P, \alpha) \leq U(f, P, \alpha) \leq U(f, P_2, \alpha). \quad (4)$$

Now,

$$\begin{aligned} U(f, P, \alpha) - L(f, P, \alpha) &= U(f, P, \alpha) - I_2 + I_2 - L(f, P, \alpha), && \text{adding and subtracting } I_2 \\ &= U(f, P, \alpha) - I_2 + I_1 - L(f, P, \alpha), && \text{since } I_1 = I_2 \\ &\leq U(f, P_2, \alpha) - I_2 + I_1 - L(f, P_1, \alpha), && \text{by (4)} \\ &< \varepsilon/2 + \varepsilon/2, && \text{by (2) and (3)} \\ &= \varepsilon. \end{aligned}$$

Thus, the condition is necessary. □

Remark 16. Let $f: [a, b] \rightarrow \mathbb{R}$ be continuous. Let $\alpha: [a, b] \rightarrow \mathbb{R}$ be increasing. If we wish to use the Riemann's criterion to prove the integrability of f , for any $\varepsilon > 0$ we need to find a partition P such that $U(f, P, \alpha) - L(f, P, \alpha) = \sum_i (M_i - m_i) \Delta \alpha_i < \varepsilon$. Look at the sum. If we wish to estimate it, we may wish to estimate each of the summands $(M_i - m_i) \Delta \alpha_i$. If the function is "well-behaved" (for example, continuous) in $[x_{i-1}, x_i]$, then it is easier to estimate $M_i - m_i$. So a natural way of estimating the sum is to use "Divide and Conquer" trick.²

Let B (B for bad!) be the set of indices i for which we have no good control over $M_i - m_i$. We break the sum into two parts $\sum_i (M_i - m_i) \Delta \alpha_i = \sum_{i \notin B} (M_i - m_i) \Delta \alpha_i + \sum_{i \in B} (M_i - m_i) \Delta \alpha_i$.

For the latter we have a crude estimate $\sum_{i \in B} (M_i - m_i) \Delta \alpha_i \leq 2C \sum_{i \in B} \Delta \alpha_i$, where $|f(x)| \leq C$ for $x \in [a, b]$. Thus to control the second, we need to find a partition P so that $\sum_{i \in B} \Delta \alpha_i$ is "small". Keep these vague ideas in mind while going through the proofs which use Riemann's criterion. You will not miss the wood for trees! Look at Examples 18–19, Theorem 22 (i), Theorem 24, Example 27, and Theorem 32.

Observation 17. As explained in the Remark above, we need to look at $M_i(f) - m_i(f)$. Alternate descriptions of $M_i(f) - m_i(f)$ are useful in the sequel. Let us start with a general observation.

Let $A - A := \{x - y : x, y \in A\}$. We claim that $\text{LUB } A - \text{GLB } A = \text{LUB } (A - A)$. Let $C = \text{LUB } (A - A)$. Let $M := \text{LUB } A$ and $m = \text{GLB } A$. Then $x - y \leq M - m$. Hence $C \leq M - m$. Fix $y \in A$.

²See the video in [5]: Riemann Integration - 5, especially 8:3–21:23.

Since $x - y \leq C$, we see that $x \leq C + y$ for all $x \in A$. Thus $C + y$ is an upper bound for A and so $M \leq C + y$. It follows from this that $M - C \leq y$. This is true for any $y \in A$. Therefore, $M - C$ is a lower bound for A . We conclude that $M - C \leq m$ or $M - m \leq C$. Thus, $M - m = \text{LUB}(A - A)$.

Next question is: Is there a better description of $\text{LUB}(A - A)$? Note that both $x - y, y - x \in A - A$ for $x, y \in A$. Since $|x - y| = \max\{x - y, y - x\}$, we suspect that $C := \text{LUB}(A - A) = \text{LUB}\{|x - y| : x, y \in A\} =: D$. Since $x - y \leq |x - y| \leq D$, it follows that D is an upper bound of $A - A$ and so $C \leq D$. Conversely, since $|x - y|$ is either $x - y$ or $y - x$ and since both lie in $A - A$, we see that $x - y \leq C$ and $y - x \leq C$ and hence $|x - y| \leq C$. That is, C is an upper bound for the set $\{|x - y| : x, y \in A\}$. Hence $D \leq C$. Thus we have established

$$\text{LUB } A - \text{GLB } A = \text{LUB}\{x - y : x, y \in A\} = \text{LUB}\{|x - y| : x, y \in A\}. \quad (5)$$

We shall apply these observations to the set $A := \{f(x) : x \in [x_{i-1}, x_i]\}$. We then arrive at the following:

$$\begin{aligned} M_i(f) - m_i(f) &= \text{LUB}\{f(s) - f(t) : s, t \in [x_{i-1}, x_i]\} \\ &= \text{LUB}\{|f(s) - f(t)| : s, t \in [x_{i-1}, x_i]\}. \end{aligned} \quad (6)$$

4 Examples of Integrable Functions

Example 18 (Step Functions). Let $f : [a, b] \rightarrow \mathbb{R}$ be a bounded function. Assume that $|f(x)| \leq M$ for $x \in [a, b]$. Let $c \in (a, b)$. Assume that $f(x) = A$ on $[a, c)$ and $f(x) = B$ on $(c, b]$. Let $f(c) = C$. We claim that f is integrable. Let P be a partition such that $c \in [x_{j-1}, x_j]$. Then we have

$$U(f, P) - L(f, P) = (\max\{A, B, C\} - \min\{A, B, C\})(x_j - x_{j-1}) \leq 2M(x_j - x_{j-1}),$$

If ε is given, we need only choose any partition P such that the length of the subinterval containing c in its interior is arbitrarily small.

For example, choose $N \in \mathbb{N}$ such that $\frac{4M}{N} < \varepsilon$. Consider the partition $P = \{a, c - \frac{1}{N}, c + \frac{1}{N}, b\}$. We assume that N is so large that $a < c - \frac{1}{N}$ and $c + \frac{1}{N} < b$. Then $U(f, P) - L(f, P) \leq 2M \times \frac{2}{N} < \varepsilon$.

This argument is easily adapted to step-functions. A function $f : [a, b] \rightarrow \mathbb{R}$ is a *step function* if there exists a finite set of points $\{t_i : 1 \leq i \leq r\} \subset [a, b]$ such that on each of the subintervals in $[a, b] \setminus \{t_i : 1 \leq i \leq r\}$, the function is a constant.

Let us be explicit. Assume without loss of generality $a \leq t_1 < t_2 < \dots < t_r \leq b$. We have subintervals of the form $[a, t_1)$, (t_{i-1}, t_i) and $(t_r, b]$. Then the assumption is that f is a constant c_i on the i -th interval. \square

Example 19. We now look at very important but easy example. Fix c such that $a < c \leq b$. Define the Heaviside function at c as follows:

$$H_c(x) := \begin{cases} 0 & \text{if } x < c \\ 1 & \text{if } x \geq c. \end{cases}$$

Note that H_c is an increasing function. We let $\alpha := H_c$. Let $f: [a, b] \rightarrow \mathbb{R}$ be a bounded function which is continuous at c . We claim that f is α -integrable and we have $\int_a^b f d\alpha = f(c)$. Let P be any partition. Choose the unique k -th subinterval of the partition so that $x_{k-1} < c \leq x_k$. Then we have the following:

$$\text{For } i < k, \quad \Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = 0 - 0 = 0,$$

$$\text{For } i = k, \quad \Delta\alpha_k = \alpha(x_k) - \alpha(x_{k-1}) = 1 - 0 = 1,$$

$$\text{For } i > k, \quad \Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = 1 - 1 = 0.$$

It follows that $U(f, P, \alpha) = M_k$ and $L(f, P, \alpha) = m_k$. We wish to apply the Riemann's criterion. We have $U(f, P, \alpha) - L(f, P, \alpha) = M_k - m_k$. Let $\varepsilon > 0$ be given. By the continuity of f at c , there exists $\delta > 0$ such that $f(c) - \varepsilon < f(t) < f(c) + \varepsilon$ for $t \in (c - \delta, c + \delta)$. We choose a partition P such that each of the subintervals is of length at most δ . If k is as above, then we have $f(c) - \varepsilon \leq m_k \leq M_k \leq f(c) + \varepsilon$. As a consequence, we have $U(f, P, \alpha) - L(f, P, \alpha) \leq 2\varepsilon$. This proves the α -integrability of f . What is $\int_a^b f d\alpha$? Observe that for such partitions we have

$$f(c) - \varepsilon \leq L(f, P, \alpha) \leq L(f, \alpha) = U(f, \alpha) \leq U(f, P, \alpha) \leq f(c) + \varepsilon.$$

Thus, $\left| \int_a^b f d\alpha - f(c) \right| \leq 2\varepsilon$. Since $\varepsilon > 0$ is arbitrary, the result follows. \square

Ex. 20. Prove that the result of the last example holds true if we relax the assumption of f by requiring that f is continuous from the left. *Hint:* We may assume that c is the right endpoint of the subinterval $(x_{k-1}, x_k]$.

Ex. 21. This is an exploratory question. (i) What happens if one assumes right continuity.
2. If we replace \geq by \leq in the definition of the Heaviside function.

Theorem 22. (i) *If f is continuous on $[a, b]$, then f is α -integrable.*
(ii) *If f is monotone and if α is continuous, then f is α -integrable.*

Proof. (i) Please review Remark 16. Given $\varepsilon > 0$, we need to find a partition P of $[a, b]$ such that $\sum_{i=1}^n (M_i - m_i)\Delta\alpha_i < \varepsilon$. Since f is continuous on any subinterval $[x_{i-1}, x_i]$, there exists $s_i, t_i \in [x_{i-1}, x_i]$ such that $f(t_i) = M_i$ and $f(s_i) = m_i$. Hence as explained in the Remark quoted above, if s_i, t_i are close to each other we may ensure that $f(t_i) - f(s_i) = M_i - m_i < \varepsilon$. A little thought will convince us to invoke the uniform continuity of f on the closed and bounded interval $[a, b]$. Let δ be such that if $|s - t| < \delta$, then $|f(s) - f(t)| < \varepsilon$. Choose $N \in \mathbb{N}$ such that $(b - a)/N < \delta$. Let P be the partition whose i -th node is $x_i := a + \frac{i}{N}(b - a)$. Then if s_i and t_i as above, it is clear that $M_i - m_i < \varepsilon$. Hence we have

$$U(f, P, \alpha) - L(f, P, \alpha) = \sum_{i=1}^N (M_i - m_i)[\alpha(x_i) - \alpha(x_{i-1})] < \varepsilon[\alpha(b) - \alpha(a)].$$

Hence we conclude that f is α -integrable.

(ii) Revisit Example 6. Note that (1) establishes the integrability of f (where $\alpha(x) = x$). What made the proof work? We focused on partitions in which all subintervals have the same

length. This facilitated the estimate of $U(f, P) - L(f, P)$. If we wish to extend the result to an arbitrary increasing integrator α , we need to find partition of $[\alpha(a), \alpha(b)]$ into subintervals of equal length and this, in turn, gives rise to partition of $[a, b]$. Let us attend to the details.

Let $y_k := \alpha(a) + \frac{k}{n}(\alpha(b) - \alpha(a))$, $0 \leq k \leq n$. Since α is increasing and continuous, we note that $\alpha([a, b]) = [\alpha(a), \alpha(b)]$. Now $y_k \in [\alpha(a), \alpha(b)]$. Hence by the intermediate value theorem, there exists $x_k \in [a, b]$ such that $\alpha(x_k) = y_k$. Note also that $x_k \leq x_{k+1}$ since α is increasing. Let $P = \{x_0, \dots, x_k, \dots, x_n\}$ be the partition of $[a, b]$. We can adapt the argument in Example 6 to arrive at $U(f, P, \alpha) - L(f, P, \alpha) = \frac{\alpha(b) - \alpha(a)}{n} [f(b) - f(a)]$. \square

Remark 23. Go through the proof of (ii) of Theorem 22 again. You may be tempted to say that it suffices to assume that $\alpha: [a, b] \rightarrow [\alpha(a), \alpha(b)]$ is onto. If you remember your real analysis well, especially the study of monotone functions, you may recall that a monotone function on an interval is continuous iff its range is an interval. (This follows from the intermediate value theorem. See pages 88-89 of [1], especially Proposition 3.5.5 and its corollary. It is always worth reviewing as it makes you gain mastery over the subject.) So, nothing gained!

The next theorem improves (i) of Theorem 22

Theorem 24. Let $f: [a, b] \rightarrow \mathbb{R}$ be a bounded function. Assume that f is continuous on $[a, b]$ except at a finite number of points. Let $\alpha: [a, b] \rightarrow \mathbb{R}$ be increasing. Assume that α is continuous at the points of discontinuity of f . Then f is α -integrable.

Proof. We shall walk through a proof assuming that f is discontinuous at only one point $c \in (a, b)$. The general case can be proved along similar lines.

As per the strategy outlined in Remark 16, c is a "bad" point. Let P be a partition such that $c \in [x_{j-1}, x_j]$. Then $U(f, P, \alpha) - L(f, P, \alpha) = \sum_{i \neq j} (M_i - m_i) \Delta \alpha_i + (M_j - m_j)(\alpha(x_j) - \alpha(x_{j-1}))$. Thanks to the continuity of f on $[a, x_{j-1}]$ and $[x_j, b]$ we are confident of ensuring that the first summand on the right is as small as we please. If $C > 0$ is such that $|f(x)| \leq C$ for $x \in [a, b]$, then the term $(M_j - m_j)[\alpha(x_j) - \alpha(x_{j-1})] \leq 2C[\alpha(x_j) - \alpha(x_{j-1})]$. Thus we should aim to "control" $\alpha(x_j) - \alpha(x_{j-1})$. This is where we invoke the continuity of α at c . Choose $s < c < t$ so small that $\alpha(t) - \alpha(s)$ is as small as we please. Let us look at $[a, b] = [a, s] \cup [s, t] \cup [t, b]$. The textbook proof is ready.

Let $\varepsilon > 0$ be given. Let $C > 0$ be such that $|f(x)| \leq C$ for $x \in [a, b]$. Since α is continuous at c , there exists $\delta > 0$ such that $|x - c| < \delta \implies |\alpha(x) - \alpha(c)| < \frac{\varepsilon}{12C}$. Let $s := c - \frac{\delta}{2}$ and $t := c + \frac{\delta}{2}$. Then $|\alpha(t) - \alpha(s)| = \alpha(t) - \alpha(s) < \frac{\varepsilon}{6C}$. Consider the intervals $[a, s]$ and $[t, b]$. Since f is continuous on each of them there exist partitions $P_1 := \{a = s_0 < \dots < s_m = s\}$ and $P_2 := \{t = t_0 < \dots < t_n = b\}$ such that $U(f, P_i, \alpha) - L(f, P_i, \alpha) < \frac{\varepsilon}{3}$ for $i = 1, 2$. (If you are very pedantic, then the " f " in $U(f, P_1, \alpha)$ is the restriction of f to $[a, s]$ and so on!) Consider $P = P_1 \cup P_2$. Let $M' = \text{LUB} \{f(x) : x \in [s, t]\}$ and $m' = \text{GLB} \{f(x) : x \in [s, t]\}$. We then have

$$\begin{aligned} & U(f, P, \alpha) - L(f, P, \alpha) \\ &= [U(f, P_1, \alpha) - L(f, P_1, \alpha)] + (M' - m')[\alpha(t) - \alpha(s)] + [U(f, P_2, \alpha) - L(f, P_2, \alpha)] \\ &< \frac{\varepsilon}{3} + 2C \frac{\varepsilon}{6C} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

This proves that f is α -integrable.

Can you now work out the proof of the general case? If c_1, \dots, c_r are the points of discontinuity of f , choose $\delta > 0$ in such that way that $\{(c_i - \delta, c_i + \delta) : 1 \leq i \leq r\}$ are pairwise disjoint. Let $s_i := c_i - \frac{\delta}{2}$ etc. What do we want about $\alpha(t_i) - \alpha(s_i)$? What kind of estimate you want on $U(f, P_i, \alpha) - L(f, P_i, \alpha)$ where P_i is a partition of $[t_{i-1}, s_i]$? Draw pictures for the final partition P . On how many subintervals the ‘control’ of $U(f, P_i, \alpha) - L(f, P_i, \alpha)$ is easy? What kind of estimate you would like to have on each of them? \square

Remark 25. Revisit Example 18. Let $f: [a, b] \rightarrow \mathbb{R}$ be a step function. Then f is continuous except at finitely many points. If we let $\alpha(x) = x$, then f is integrable on $[a, b]$ by the last theorem. This was proved already in Example 18.

Remark 26. There is a result due to Lebesgue which characterizes the integrability of f in terms of the ‘size’ (measure or length) of the set of discontinuities. (See a forthcoming book) A particular case: if the set of discontinuities of f is countable, then the function is integrable. For example, Thomae’s function is integrable. See Example 27.

Give Ref!

Example 27 (Integrability of Thomae’s Function). Let $f: [0, 1] \rightarrow \mathbb{R}$ be defined by

$$f(x) = \begin{cases} 0, & \text{if } x \text{ is irrational} \\ 1/q, & \text{if } x \in \mathbb{Q}, x = p/q \text{ with } p, q \in \mathbb{N} \text{ and } \gcd(p, q) = 1. \end{cases}$$

Given $\varepsilon > 0$, we must find a partition P such that $U(f, P) - L(f, P) < \varepsilon$. By the density of irrationals, in any subinterval $[x_i, x_{i+1}]$ of a partition P of $[0, 1]$, irrationals exist and hence m_i ’s are zero and hence $L(f, P) = 0$. Hence we need only show $U(f, P) < \varepsilon$. Let $n \in \mathbb{N}$ be such that $1/n < \varepsilon$. The set $A_n := \{r \in \mathbb{Q} \cap [0, 1] : f(r) > 1/n\}$ is finite. If $r \in A_n$ and $r \in [x_{i-1}, x_i]$, $M_i(f) \geq \frac{1}{n}$. We call such interval ‘bad’. We employ the divide and conquer method to show that the sum of the lengths of such bad intervals is ‘small’.

Let $\varepsilon > 0$ be given. Choose $k \in \mathbb{N}$ such that $\frac{1}{k} < \varepsilon/2$. There exists a finite number, say, N of rational numbers p/q with $q \leq k$. Denote them by $\{r_j : 1 \leq j \leq N\}$. Let $\delta < \varepsilon/(4N)$. Choose a partition $P = \{x_0, \dots, x_n\}$ of $[0, 1]$ such that

$$\max\{|x_{i+1} - x_i| : 0 \leq i \leq n-1\} < \delta.$$

Let $A := \{i : r_j \in [x_i, x_{i+1}], \text{ for some } j\}$, and $B := \{0, \dots, n\} \setminus A$. Note that the number of elements in A will be at most $2N$. (Why $2N$? Some r_j could be the left and the right endpoint of adjacent

subintervals!) For $i \in A$, we have $M_i \leq 1$. For $j \in B$, we have $M_j < 1/k$. Hence

$$\begin{aligned}
 U(f, P) &= \sum_{i=0}^{n-1} M_i(x_{i+1} - x_i) \\
 &= \sum_{i \in A} M_i(x_{i+1} - x_i) + \sum_{j \in B} M_j(x_{j+1} - x_j) \\
 &\leq (2N)\delta + \sum_{j \in B} \frac{1}{k}(x_{j+1} - x_j) \\
 &\leq (2N)\delta + \frac{1}{k} \\
 &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
 \end{aligned}$$

In the above, we used the fact that $\sum_{j \in B} (x_{j+1} - x_j)$ is the sum of the lengths of the disjoint subintervals that lie in B , and hence it is at most 1.

Thus, for any $\varepsilon > 0$, we have found a partition P_ε such that $U(f, P_\varepsilon) < \varepsilon$. It follows that $\text{GLB} \{U(f, P) : P \text{ is a partition of } [0, 1]\} = 0$. Hence $\int_0^1 f = 0$. Of course, we could use a simpler argument. Since f is integrable, and since each lower sum is zero, it follows that the lower integral is 0. Hence $\int_0^1 f = 0$. (Always explore different ways of looking at the same thing!)

5 Riemann's Approach to Integration

We now introduce the Riemann sum approach to α -integrability.

Definition 28. Let $f, \alpha: [a, b] \rightarrow \mathbb{R}$ be bounded functions. (At present we do not insist on the monotonicity of α .) Let $P = \{x_0, \dots, x_n\}$ be a partition of $[a, b]$. Let $t_i \in [x_{i-1}, x_i]$ for $1 \leq i \leq n$ be arbitrary. We let $\mathbf{t} := (t_1, \dots, t_n)$ and call it a set of tags. The Riemann-Stieltjes sum is defined as follows:

$$S(f, P, \mathbf{t}, \alpha) = \sum_{i=1}^n f(t_i) \Delta \alpha_i.$$

We say that f is α -integrable on $[a, b]$ if there exists $A \in \mathbb{R}$ such that for any given $\varepsilon > 0$, there exists a partition P such that for *any* refinement Q of P and for *any* set of tags in Q , we have $|S(f, P, \mathbf{t}, \alpha) - A| = |\sum_{i=1}^n f(t_i) \Delta \alpha_i - A| < \varepsilon$. It is easy to see that such an A is unique. We call A the Riemann-Stieltjes integral of f with respect to the integrator α .

We refer to this as the Riemann sum approach to integration. The earlier one will be referred to as Darboux sum (upper/lower sum) approach to integration.

We now show that the present definition of α -integral is equivalent to the earlier definition if the integrator α is increasing.

Theorem 29. Let $f: [a, b] \rightarrow \mathbb{R}$ be bounded. Let $\alpha: [a, b] \rightarrow \mathbb{R}$ be increasing. Then f is α -integrable according to the first definition iff it is α -integrable according to Definition 28. In such a case, both the integrals are the same.

Proof. Let f be integrable according to Definition 28 with the integral A . We need to show that f is α -integrable according to the first definition. We plan to use the Riemann's criterion to achieve this. Let $\varepsilon > 0$ be given. Then by definition there exists a partition P such that for any partition Q , which refines P and for any set of tags \mathbf{t} , we have $|\sum_{i=1}^n f(t_i)\Delta\alpha_i - A| < \varepsilon$. We now choose a sequence $\mathbf{t}_k := (t_{k1}, \dots, t_{kn})$ of tags as follows: $f(t_{ki}) > M_i - \frac{1}{k}$ for $k \in \mathbb{N}$. Note that we have $M_i - \frac{1}{k} < f(t_{ki}) \leq M_i$. It follows that $\lim_k \sum_i f(t_{ki})\Delta\alpha_i = U(f, P, \alpha)$. Since $|\sum_{i=1}^n f(t_{ki})\Delta\alpha_i - A| < \varepsilon$, it leads us to conclude that $|U(f, P, \alpha) - A| \leq \varepsilon$. In a similar way, we arrive at $|L(f, P, \alpha) - A| \leq \varepsilon$. Hence $|U(f, P, \alpha) - L(f, P, \alpha)| \leq 2\varepsilon$. This establishes the α -integrability of f according to the first definition. Also it shows that $A = U(f, \alpha) = L(f, \alpha)$.

We now prove the converse. Let $A := U(f, \alpha) = L(f, \alpha)$. We shall show that f is α -integrable as per Definition 28. Let $\varepsilon > 0$ be given. Then there exist partitions P_1 and P_2 such that

$$A - \varepsilon < L(f, P_1, \alpha) \quad \text{and} \quad U(f, P_2, \alpha) < A + \varepsilon. \quad (7)$$

Note that (7) remains true for any partition P which refines both P_1 and P_2 . Let \mathbf{t} be a set of tags in such a partition P . We then have

$$-\varepsilon < L(f, P, \alpha) - A \leq \sum_{i=1}^n f(t_i)\Delta\alpha_i - A \leq U(f, P, \alpha) - A < \varepsilon.$$

We conclude that f is α -integrable as per Definition 28 with the integral A . \square

Remark 30. Definition 28 has certain advantages. (i) We do not have to impose any condition on the integrator except that it is bounded. (ii) The proofs of the analogues of (i)–(iii) of Theorem 31 are easier to prove using the second definition, that too, in a more general setting of any integrator α ! But a look at (iv) and (v) of the same theorem will show that we cannot hope to prove analogous results in the general setting.

6 Class of Integrable Functions

You may like to watch the videos 7 and 8 in [5] either as you read this section or you may read this section after watching them.

Theorem 31. Let $\alpha: [a, b] \rightarrow \mathbb{R}$ be a bounded function. Let $\mathcal{R}(\alpha)$ denote the set of α -integrable functions on $[a, b]$. Then the following are true.

- i) $\mathcal{R}(\alpha)$ is a vector space over \mathbb{R} and the map $f \mapsto \int_a^b f d\alpha$ is linear.
- ii) Let $\beta: [a, b] \rightarrow \mathbb{R}$ be bounded. Let $\sigma := \lambda\alpha + \mu\beta$ where $\lambda, \mu \in \mathbb{R}$. If f is α as well as β -integrable, then f is σ integrable and we have

$$\int_a^b f d\sigma = \lambda \int_a^b f d\alpha + \mu \int_a^b f d\beta. \quad (8)$$

- iii) Let $a < c < b$. Assume that any two of the integrals in (9) exist. Then the third integral exists and we have

$$\int_a^b f d\alpha = \int_a^c f d\alpha + \int_c^b f d\alpha. \quad (9)$$

iv) Assume that α is increasing. Let $f \in \mathcal{R}(\alpha)$ and $f \geq 0$ on $[a, b]$. Then $\int_a^b f d\alpha \geq 0$. More generally, let $f, g \in \mathcal{R}(\alpha)$. Assume that $f(x) \leq g(x)$ for $x \in [a, b]$. Then we have

$$\int_a^b f d\alpha \leq \int_a^b g d\alpha. \quad (10)$$

v) Assume that α is increasing. Let $f \in \mathcal{R}(\alpha)$. Assume that $|f(x)| \leq M$ for $x \in [a, b]$. Then $|f| \in \mathcal{R}(\alpha)$ and we have

$$\left| \int_a^b f d\alpha \right| \leq \int_a^b |f| d\alpha \leq M(\alpha(b) - \alpha(a)). \quad (11)$$

vi) Assume that α is increasing. Let $f \in \mathcal{R}(\alpha)$. Then $f^2 \in \mathcal{R}(\alpha)$.

vii) Assume that α is increasing. Let $f, g \in \mathcal{R}(\alpha)$. Then $fg \in \mathcal{R}(\alpha)$.

Proof. Just to make sure that you are alert while reading the statements, (i) – (iii) are true for any integrator α . We do not require α to be increasing for the validity of (i)–(iii).

The proofs will be instructive in the sense that we shall employ one of the equivalent definitions (for the case on hand) which makes our lives easy!

(i) Let $\lambda, \mu \in \mathbb{R}$. Observe that

$$S(\lambda f + \mu g, P, \mathbf{t}, \alpha) = \sum_{i=1}^n (\lambda f(t_i) + \mu g(t_i)) \Delta \alpha_i = \lambda S(f, P, \mathbf{t}, \alpha) + \mu S(g, P, \mathbf{t}, \alpha).$$

Hence we obtain

$$\begin{aligned} & \left| S(\lambda f + \mu g, P, \mathbf{t}, \alpha) - \left(\lambda \int_a^b f d\alpha + \mu \int_a^b g d\alpha \right) \right| \\ & \leq |\lambda| \left| S(f, P, \mathbf{t}, \alpha) - \int_a^b f d\alpha \right| + |\mu| \left| S(g, P, \mathbf{t}, \alpha) - \int_a^b g d\alpha \right|. \end{aligned} \quad (12)$$

Do you see how to write a textbook proof now?

Given $\varepsilon > 0$, choose a partition P_1 and P_2 such that

$$\left| S(f, P_1, \mathbf{t}, \alpha) - \int_a^b f d\alpha \right| < \frac{\varepsilon}{2(1+|\lambda|)} \quad \& \quad \left| S(g, P_2, \mathbf{t}, \alpha) - \int_a^b g d\alpha \right| < \frac{\varepsilon}{2(1+|\mu|)}$$

Let $P := P_1 \cup P_2$. Then, for any partition Q finer than P , it follows from (12) that

$$\left| S(\lambda f + \mu g, P, \mathbf{t}, \alpha) - \left(\lambda \int_a^b f d\alpha + \mu \int_a^b g d\alpha \right) \right| < \varepsilon.$$

This completes the proof of (i).

Note that (ii) has no analogue in the theory of Riemann integrals.

Proof of (ii) is very similar to that of (i). Start with the question: What is the relation between $S(f, P, \lambda\alpha + \mu\beta)$ in terms of $\lambda, \mu, S(f, P, \alpha)$ and $S(f, P, \beta)$?

Remark: This may be omitted on first reading.

Just for fun: If you wish to prove (ii) using Darboux sum approach, you need to assume λ, μ are positive. Let us go through a proof when α is increasing and $\lambda = \mu = 1$.

To prove ii), we start with an observation: $\Delta\sigma_i = \Delta\alpha_i + \Delta\beta_i$. An immediate consequence is: $L(f, P, \sigma) = L(f, P, \alpha) + L(f, P, \beta)$ for any partition P . Similar result holds for the upper sums.

We first show that f is σ -integrable, using Riemann's criterion. Let $\varepsilon > 0$ be given. Then we can find partitions P_1 and P_2 so that

$$U(f, P_1, \alpha) - L(f, P_1, \alpha) < \varepsilon/2 \quad \text{and} \quad U(f, P_2, \beta) - L(f, P_2, \beta) < \varepsilon/2.$$

For any common refinement P of P_1 and P_2 , we have

$$U(f, P, \sigma) - L(f, P, \sigma) = U(f, P, \alpha) - L(f, P, \alpha) + U(f, P, \beta) - L(f, P, \beta) < \varepsilon.$$

This implies that f is σ -integrable.

We now prove (8). For *any* partition P , we have

$$\begin{aligned} L(f, P, \sigma) &= L(f, P, \alpha) + L(f, P, \beta) \\ &\leq \int_a^b f d\alpha + \int_a^b f d\beta \\ &\leq U(f, P, \alpha) + U(f, P, \beta) \\ &= U(f, P, \sigma). \end{aligned}$$

Can you see how (8) follows from this?

Let $A, B \subset \mathbb{R}$ be nonempty. Assume that for each $x \in A$ and $y \in B$, we have $x \leq y$. Let $\alpha := \text{LUB } A$ and $\beta := \text{GLB } B$. (Why do they exist?) Let us assume that $\alpha = \beta$. Let $c \in \mathbb{R}$ be such that $x \leq c \leq y$ for $x \in A$ and $y \in B$. Then we claim that $\alpha = c = \beta$.

For, c is an upper bound for A and a lower bound for B . Hence $\alpha \leq c$ and $c \leq \beta$. Since $\alpha = \beta$, the claim follows.

Thus we have proved ii).

We shall now prove one of the three cases of (iii). Assume that $\int_a^c f d\alpha$ and $\int_c^b f d\alpha$ exist. Then we are required to show that f is α -integrable on $[a, b]$ and that (9) holds.

Let P be a partition of $[a, b]$ such that c is a node, that is, $c \in P$. Let $P_1 = P \cap [a, c]$ and $P_2 = P \cap [c, b]$. Then P_1 and P_2 provide partitions of $[a, c]$ and $[c, b]$ respectively.

What is the relation between $S(f, P_1, \mathbf{t}_1, \alpha)$, $S(f, P_2, \mathbf{t}_2, \alpha)$ and $S(f, P, \mathbf{t}, \alpha)$? Of course, its easily seen that

$$S(f, P, \mathbf{t}, \alpha) = S(f, P_1, \mathbf{t}_1, \alpha) + S(f, P_2, \mathbf{t}_2, \alpha).$$

Let $\varepsilon > 0$ be given. Choose partitions P_1 and P_2 of $[a, c]$ and $[c, b]$ respectively with the property that for any set of tags we have

$$\left| S(f, P_1, \mathbf{t}_1, \alpha) - \int_a^c f d\alpha \right| < \varepsilon/2 \text{ and } \left| S(f, P_2, \mathbf{t}_2, \alpha) - \int_c^b f d\alpha \right| < \varepsilon/2.$$

Let $P = P_1 \cup P_2$. Let Q be any partition finer than P . Note that $Q_i = Q \cap P_i$ will be finer than P_i , $i = 1, 2$. Hence for any set of tags in Q we have

$$\left| S(f, Q, \mathbf{t}, \alpha) - \int_a^c f d\alpha - \int_c^b f d\alpha \right| < \varepsilon.$$

(Can you justify the inequality quoting the observation made above?) This proves (iii).

Let $f \in \mathcal{R}(\alpha)$ and $f \geq 0$. To prove (iv), we need only ask the question: Which of the definitions is likely to yield a simple proof? Yes, the Darboux sum approach. Again, the next question is since $\int_a^b f d\alpha$ is the LUB of lower sums and the GLB of upper sums, which sum should we use? An obvious choice is lower sums. Observe that $L(f, P, \alpha) \geq 0$. Since $\int_a^b f d\alpha \geq L(f, P, \alpha)$, the result follows.

An aside. You could work with upper sums as well as Riemann sums. In each of these cases, assume that the integral is negative, say $A < 0$. Choose $\varepsilon = -\frac{A}{2}$ to arrive at a contradiction. Recall how you prove: If (x_n) is a sequence of nonnegative terms converging to x , then $x \geq 0$.

The general case can be deduced easily from this. Consider $g - f$. By (i), $g - f \in \mathcal{R}(\alpha)$ and also $g - f \geq 0$. Hence $\int_a^b (g - f) d\alpha \geq 0$. Again by the linearity of the integral, $\int_a^b (g - f) d\alpha = \int_a^b g d\alpha - \int_a^b f d\alpha \geq 0$. Thus we have proved (iv).

(v) I am sure all of us will think of Darboux sum approach to prove this. The key idea of the proof is to find the relation between $M_i(f) - m_i(f)$ and $M_i(|f|) - m_i(|f|)$. If you think for a while, we can formulate a more general question. Let $A \subset \mathbb{R}$ be a nonempty set which is both bounded above and bounded below. Let $B := \{|x| : x \in A\}$. Recall the relation between LUB A - GLB A and LUB B - GLB B made in Observation 17.

We are now ready to prove (v). We let $A := \{f(x) : x \in [x_{i-1}, x_i]\}$ and apply the observations made in the last two paragraphs above. Recall that for any $s, t \in \mathbb{R}$, we have $||s| - |t|| \leq |s - t|$. Observe that for $s, t \in [x_{i-1}, x_i]$ we have

$$||f(s)| - |f(t)|| \leq |f(s) - f(t)|.$$

The LUB of the left side is $M_i(|f|) - m_i(|f|)$. Similarly the LUB of the right side is $M_i(f) - m_i(f)$. It follows that

$$M_i(|f|) - m_i(|f|) \leq M_i(f) - m_i(f).$$

Since $\Delta\alpha_i \geq 0$, we see that $[M_i(|f|) - m_i(|f|)] \Delta\alpha_i \leq [M_i(f) - m_i(f)] \Delta\alpha_i$. Hence

$$U(|f|, P, \alpha) - L(|f|, P, \alpha) \leq U(f, P, \alpha) - L(f, P, \alpha).$$

Do you see how to prove $|f| \in \mathcal{R}(\alpha)$? If yes, which results are needed to establish (11)?

We now prove (vi). The key observation is $x^2 = |x|^2$ for $x \in \mathbb{R}$. We have

$$\begin{aligned} M_i(f^2) - m_i(f^2) &= (M_i(|f|))^2 - (m_i(|f|))^2 \\ &= [M_i(|f|) + m_i(|f|)] \cdot [M_i(|f|) - m_i(|f|)] \\ &\leq 2C [M_i(|f|) - m_i(|f|)]. \end{aligned}$$

Here $C := \text{LUB } \{|f(x)| : x \in [a, b]\}$. Can you complete the proof now?

Proof (vii) depends on the algebraic identity: $ab = \frac{1}{2}[(a+b)^2 - a^2 - b^2]$. We have $2f(x)g(x) = [f(x) + g(x)]^2 - f(x)^2 - g(x)^2$. Since $f, g \in \mathcal{R}(\alpha)$, $f + g \in \mathcal{R}(\alpha)$ and hence $(f + g)^2, f^2, g^2 \in \mathcal{R}(\alpha)$. Hence the right side lies in $\mathcal{R}(\alpha)$ and so the left side $2fg \in \mathcal{R}(\alpha)$.

You may compare the proofs of i), iii)–v) with those in Section 6.2 of [1]. □

The next theorem is a very useful result.

Theorem 32. *Let $\alpha: [a, b] \rightarrow \mathbb{R}$ be increasing. Let $f \in \mathcal{R}(\alpha)$. Assume that $m \leq f(x) \leq M$ for $x \in [a, b]$. Let $g: [m, M] \rightarrow \mathbb{R}$ be continuous. Then $g \circ f \in \mathcal{R}(\alpha)$.*

Proof. Let $C > 0$ be such that $|g(y)| \leq C$ for $y \in [m, M]$. Let $\varepsilon > 0$ be given. We need to find a partition P such that

$$U(g \circ f, P, \alpha) - L(g \circ f, P, \alpha) = \sum_i (M_i(g \circ f) - m_i(g \circ f)) \Delta \alpha_i < \varepsilon.$$

We shall use the “Divide and Conquer” trick to estimate the sum. By the uniform continuity of g on $[m, M]$, for the given θ , there corresponds a $\delta > 0$. Hence if we let $G := \{i : M_i(f) - m_i(f) < \delta\}$, then we have a control over $M_i(g \circ f) - m_i(g \circ f)$. Let $B := \{i : M_i(f) - m_i(f) \geq \delta\}$. The sum over B is

$$\sum_{i \in B} (M_i(g \circ f) - m_i(g \circ f)) \Delta \alpha_i \leq 2C \sum_{i \in B} \Delta \alpha_i.$$

As per our strategy outlined in Remark 16, we need to choose the partition P so that $\sum_{i \in B} \Delta \alpha_i$ is “small”. Note that $\sum_{i \in G} (M_i(g \circ f) - m_i(g \circ f)) \Delta \alpha_i \leq \sum_{i \in G} \theta \Delta \alpha_i \leq \theta(\alpha(b) - \alpha(a))$. Let P be a partition of $[a, b]$ such that $U(f, P, \alpha) - L(f, P, \alpha) < \eta$. Presumably η will be something to do with ε and will be decided later. We then arrive at the following estimate

$$\begin{aligned} U(g \circ f, P, \alpha) - L(g \circ f, P, \alpha) &= \sum_{i \in G} (M_i(g \circ f) - m_i(g \circ f)) \Delta \alpha_i + \sum_{i \in B} (M_i(g \circ f) - m_i(g \circ f)) \Delta \alpha_i \\ &< \sum_{i \in G} \theta \Delta \alpha_i + 2C \sum_{i \in B} \Delta \alpha_i \\ &< \theta(\alpha(b) - \alpha(a)) + 2C \sum_{i \in B} \Delta \alpha_i. \end{aligned}$$

We want the first term on the last right expression to be less than $\varepsilon/2$. So, you choose $0 < \theta < \frac{\varepsilon}{2[1 + \alpha(b) - \alpha(a)]}$.

We need to estimate $\sum_{i \in B} \Delta \alpha_i$ so that it is less than $\varepsilon/2$. Observe that

$$\eta > U(f, P, \alpha) - L(f, P, \alpha) \geq \sum_{i \in B} (M_i - m_i) \Delta \alpha_i \geq \delta \sum_{i \in B} \Delta \alpha_i.$$

We find that $\sum_{i \in B} \Delta \alpha_i < \eta/\delta$. So this suggests that we choose a partition P so that $U(f, P, \alpha) - L(f, P, \alpha) < \frac{\varepsilon \delta}{4C}$.

Do you think you can write a textbook proof now?

Let $\varepsilon > 0$ be given. Let $C > 0$ be such that $|g(t)| \leq C$ for $t \in [m, M]$. Since g is uniformly continuous on $[m, M]$, there exists $\delta > 0$ such that

$$|s - t| < \delta \implies |g(s) - g(t)| < \frac{\varepsilon}{2[1 + \alpha(b) - \alpha(a)]}, \quad s, t \in [m, M].$$

Since f is integrable, there exists a partition P of $[a, b]$ such that

$$U(f, P, \alpha) - L(f, P, \alpha) < \frac{\varepsilon \delta}{4C}.$$

Let $G := \{i : M_i(f) - m_i(f) < \delta\}$ and $B := \{i : M_i(f) - m_i(f) \geq \delta\}$. Then we have

$$\delta \sum_{i \in B} \Delta \alpha_i \leq \sum_{i \in B} (M_i - m_i) \Delta \alpha_i < \frac{\varepsilon \delta}{4C} \implies \sum_{i \in B} \Delta \alpha_i < \frac{\varepsilon}{4C}. \quad (13)$$

For $x, y \in [x_{i-1}, x_i]$, $i \in G$, we see that $|f(x) - f(y)| \leq M_i - m_i < \delta$ so that $g(f(x)) - g(f(y)) < \frac{\varepsilon}{2[1 + \alpha(b) - \alpha(a)]}$. Observation 17 implies that

$$M_i(g \circ f) - m_i(g \circ f) \leq \frac{\varepsilon}{2[1 + \alpha(b) - \alpha(a)]} \quad \text{for } i \in G. \quad (14)$$

Let us estimate $U(g \circ f, P, \alpha) - L(g \circ f, P, \alpha)$.

$$\begin{aligned} & U(g \circ f, P, \alpha) - L(g \circ f, P, \alpha) \\ &= \sum_{i \in G} [M_i(g \circ f) - m_i(g \circ f)] \Delta \alpha_i + \sum_{i \in B} [M_i(g \circ f) - m_i(g \circ f)] \Delta \alpha_i \\ &< \frac{\varepsilon}{2[1 + \alpha(b) - \alpha(a)]} \sum_{i \in G} \Delta \alpha_i + 2C \sum_{i \in B} \Delta \alpha_i && \text{by (14)} \\ &< \frac{\varepsilon}{2[1 + \alpha(b) - \alpha(a)]} \sum_{i=1}^n \Delta \alpha_i + 2C \frac{\varepsilon}{4C} && \text{by (13)} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}. \end{aligned}$$

We hope that you enjoyed the proof. This is an example of how one thinks through a proof and then writes a textbook proof. \square

Remark 33. Do you realize that (v) and (vi) of Theorem 31 are easy consequences of the last Theorem 32?

7 Two Important Examples of Riemann-Stieltjes Integrals

You may like to watch the video 11 in [5] either as you read this section or you may read this section after watching it.

We now give two important examples of Riemann-Stieltjes integrals which are very useful. Also, the first example shows that an infinite series is a special case of α -integrals.

Example 34. Let $\{t_k : 1 \leq k \leq N\}$ be a finite set of points in (a, b) . (Did you notice the brackets? Do you see why they are different?) Let $\{c_k : 1 \leq k \leq N\}$ be a set of nonnegative numbers. Let $\alpha(x) := \sum_{k=1}^N c_k H_{t_k}(x)$. It is clear that α is increasing. Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous. We then have

$$\int_a^b f d\alpha = \sum_k c_k f(t_k). \quad (15)$$

(What results did we use to assert (15)?)

Ex. 35. Let $\alpha : [0, N] \rightarrow \mathbb{R}$ be the greatest integer function $\alpha(x) = [x]$. Write it as “a sum of Heaviside functions”. Let $f : [0, N] \rightarrow \mathbb{R}$ be continuous. Prove that $\int_0^N f d\alpha = \sum_{k=0}^N f(k)$.

Let now $\{t_k : k \in \mathbb{N}\}$ be a countable subset of (a, b) . Let (c_k) be a sequence of nonnegative numbers such that the associated series $\sum_k c_k$ is convergent. We define

$$\alpha(x) := \sum_k c_k H_{t_k}(x), \quad x \in [a, b]. \quad (16)$$

Then $\alpha(x)$ makes sense (why?) and is an increasing function. We now claim that for any continuous function f on $[a, b]$ we have $\int_a^b f d\alpha = \sum_k c_k f(t_k)$. This is an easy exercise with the following hint: Given $\varepsilon > 0$, choose $N \in \mathbb{N}$ such that $\sum_{k \geq N+1} c_k < \varepsilon/2$. Write $\alpha = \sum_{k=1}^N c_k H_{t_k} + \sum_{k \geq N+1} c_k H_{t_k} = \beta_n + \sigma_n$, say.

We need to show that $\sum_{k=1}^n c_k f(t_k) \rightarrow \int_a^b f d\alpha$. Let $\varepsilon > 0$ be given. Then we need to find N such that for $n \geq N$, $\left| \int_a^b f d\alpha - \sum_{k=1}^n c_k f(t_k) \right| < \varepsilon$. Keep the notation of the hint. For $n \geq N$, let us write $\alpha = \beta_n + \sigma_n$, in an obvious (?) notation. Then we have $\int_a^b f d\alpha = \int_a^b f d\beta_n + \int_a^b f d\sigma_n$. We know that $\int_a^b f d\beta_n = \sum_{k=1}^n c_k f(t_k)$. Hence we have, for $n \geq N$,

$$\begin{aligned} \left| \int_a^b f d\alpha - \sum_{k=1}^n c_k f(t_k) \right| &= \left| \int_a^b f d\alpha - \int_a^b f d\beta_n \right| && \text{(Why?)} \\ &= \left| \int_a^b f d\sigma_n \right| && \text{(Why?)} \\ &\leq \int_a^b |f| d\sigma_n && \text{(Why?)} \\ &\leq M(\sigma_n(b) - \sigma_n(a)) && \text{(What is } M? \text{ Why?)} \\ &= M \sum_{k \geq n+1} c_k < M\varepsilon. \end{aligned}$$

Hence we have proved the following theorem.

Theorem 36. Let $\{t_k : k \in \mathbb{N}\}$ be a countable subset of (a, b) . Let (c_k) be a sequence of nonnegative numbers such that the associated series $\sum_k c_k$ is convergent. Define

$$\alpha(x) := \sum_k c_k H_{t_k}(x), \quad x \in [a, b]. \quad (17)$$

Then for any continuous function $f: [a, b] \rightarrow \mathbb{R}$, we have

$$\int_a^b f d\alpha = \sum_k c_k f(t_k). \quad (18)$$

The second example is given by the following theorem.

Theorem 37. Let $f, \alpha: [a, b] \rightarrow \mathbb{R}$ be bounded. Assume that α is continuously differentiable and that f is α -integrable. Then $f\alpha'$ is integrable on $[a, b]$ and we have $\int_a^b f d\alpha = \int_a^b f \alpha' dx$.

Proof. It behooves us to consider the Riemann sums. Let P be any partition with tags \mathbf{t} . Then we look at

$$\sum_{i=1}^n f(t_i) \Delta\alpha_i - \sum_{i=1}^n f(t_i) \alpha'(t_i) \Delta x_i. \quad (19)$$

Again it is clear that we wish to apply the mean value theorem to $\Delta\alpha_i$: $\Delta\alpha_i = \alpha'(s_i) \Delta x_i$. Using this in (19) we arrive at

$$\sum_{i=1}^n f(t_i) \Delta\alpha_i - \sum_{i=1}^n f(t_i) \alpha'(t_i) \Delta x_i = \sum_{i=1}^n f(t_i) (\alpha'(s_i) - \alpha'(t_i)) \Delta x_i. \quad (20)$$

Our aim is to show that these two Riemann sums are close to each other. So the next obvious step is to use the uniform continuity of α' . Give $\varepsilon > 0$, by the uniform continuity of α' there exists a $\delta > 0$ such that $|s - t| < \delta \implies |\alpha'(s) - \alpha'(t)| < \varepsilon$. Then for any a partition Q , which is a refinement of P , with the property that the maximum length of the subintervals is less than δ , it follows from (20)

$$\left| \sum_{i=1}^n f(t_i) \Delta\alpha_i - \sum_{i=1}^n f(t_i) \alpha'(t_i) \Delta x_i \right| \leq M\varepsilon(b-a). \quad (21)$$

(M has the usual meaning!) We choose any partition Q' which is refinement of Q so that $\left| \int_a^b f d\alpha - \sum_{i=1}^n f(t_i) \Delta\alpha_i \right| < \varepsilon$ for any set of tags. We then see that for any such Q' and any set of tags

$$\left| \int_a^b f d\alpha - \sum_{i=1}^n f(t_i) \alpha'(t_i) \Delta x_i \right| < (M(b-a) + 1)\varepsilon.$$

This completes the proof. □

Remark 38. The result above can be proved without assuming that α' is continuous but under the hypothesis that α' is integrable. See Theorem 39.

Theorem 39. Let $f \in \mathcal{R}[a, b]$. Let $\alpha: [a, b] \rightarrow \mathbb{R}$ be differentiable. Assume further that $\alpha' \in \mathcal{R}[a, b]$. Then $f \in \mathcal{R}(\alpha)$ and we have

$$\int_a^b f d\alpha = \int_a^b f \alpha' dx. \quad (22)$$

Proof. Note that $f, \alpha' \in \mathcal{R}[a, b]$ implies $f\alpha' \in \mathcal{R}[a, b]$. Hence the right side of (22) makes sense. Let $A := \int_a^b f \alpha'$. We wish to show that $\int_a^b f d\alpha = A$. Let $|f(x)| \leq C$ for $x \in [a, b]$.

Let $\varepsilon > 0$ be given. Since $f\alpha'$ is integrable and α' is integrable, there exists a (common) partition P such that for any partition $Q \geq P$, and for any set of tags in Q , we have

$$|S(f\alpha', Q, \mathbf{t}) - A| = \left| \sum_i f(t_i)\alpha'(t_i)(x_i - x_{i-1}) - A \right| < \frac{\varepsilon}{2}, \quad \text{and} \quad U(\alpha', Q) - L(\alpha', Q) < \frac{\varepsilon}{2C}. \quad (23)$$

We claim that $|S(f, Q, \mathbf{t}, \alpha) - A| < \varepsilon$. Applying the mean value theorem to $\alpha(x_i) - \alpha(x_{i-1})$, we find $s_i \in [x_{i-1}, x_i]$ such that

$$S(f, Q, \mathbf{t}, \alpha) = \sum_i f(t_i)\Delta\alpha_i = \sum_i f(t_i)\alpha'(s_i)(x_i - x_{i-1}).$$

We have

$$\begin{aligned} |S(f\alpha', Q, \mathbf{t}) - S(f, Q, \mathbf{t}, \alpha)| &= \left| \sum_i f(t_i)(\alpha'(t_i) - \alpha'(s_i))(x_i - x_{i-1}) \right| \\ &\leq \sum_i |f(t_i)| |\alpha'(t_i) - \alpha'(s_i)| (x_i - x_{i-1}) \\ &\leq \sum_i |f(t_i)| (M_i(\alpha') - m_i(\alpha'))(x_i - x_{i-1}) \quad (\text{Why?}) \\ &\leq C(U(\alpha', Q) - L(\alpha', Q)) < \varepsilon/2. \quad (\text{Why?}) \end{aligned} \quad (24)$$

It follows from (23) - (24) that

$$\begin{aligned} |S(f, Q, \mathbf{t}, \alpha) - A| &\leq |S(f, Q, \mathbf{t}, \alpha) - S(f\alpha', Q, \mathbf{t}) + S(f\alpha', Q, \mathbf{t}) - A| \\ &\leq |S(f, Q, \mathbf{t}, \alpha) - S(f\alpha', Q, \mathbf{t})| + |S(f\alpha', Q, \mathbf{t}) - A| \\ &< \varepsilon. \end{aligned}$$

This shows that $\int_a^b f d\alpha = A$. □

Remark 40. Did you observe something interesting about the proof above? We did not hesitate to use both the definitions of the integral! To show that f is α -integrable, we used the Riemann sum approach and to estimate a sum, we used the Riemann's criterion for the Darboux sum approach. I thought of the proof after I sent you the file! I have given a textbook proof above. Can you 'discover' the proof starting from a wish to estimate $|S(f, Q, \mathbf{t}, \alpha) - A|$?

8 Fundamental Theorems of Calculus

You may like to watch the videos 9 and 10 in [5] either as you read this section or you may read this section after watching them.

We now look at one of the most important results in the theory of integration, namely, the fundamental theorems of calculus. These theorems establish the validity of the computation of integrals via Newtonian calculus, as learned in high school. In some sense, they justify the high-school way of defining integration as finding an anti-derivative.

Theorem 41 (First Fundamental Theorem of Calculus). *Let $f: [a, b] \rightarrow \mathbb{R}$ be differentiable. Assume that f' is integrable on $[a, b]$. Then*

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

Proof. We wish to use the Riemann sum approach. A Riemann sum for the integral $\int_a^b f'(t) dt$ looks like $\sum_{i=1}^n f'(t_i)(x_i - x_{i-1})$. The summands remind us of the mean value theorem of differential calculus. Let $\varepsilon > 0$ be given. Then there exists a partition P such that for any partition $Q \geq P$ and set of tags \mathbf{t} , we have

$$\left| S(f', Q, \mathbf{t}) - \int_a^b f'(t) dt \right| < \varepsilon. \quad (25)$$

Let $Q \geq P$ be partition. Then observe that

$$f(b) - f(a) = \sum_{i=1}^n [f(x_i) - f(x_{i-1})] = \sum_{i=1}^n f'(t_i)(x_i - x_{i-1}), \quad (26)$$

where we used the mean value theorem to $f(x_i) - f(x_{i-1})$, $1 \leq i \leq n$. But the term on the extreme right is $S(f', Q, \mathbf{t})$. Hence it follows from (25)-(26) that

$$\left| f(b) - f(a) - \int_a^b f'(t) dt \right| < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the result follows. \square

Remark 42. Note that this justifies what you learned in school about the integral being anti-derivative. That is, to find $\int_a^b f(x) dx$, we find a function g such that $g' = f$ and then in this case we have $\int_a^b f(x) dx = g(b) - g(a)$.

Let $f: [a, b] \rightarrow \mathbb{R}$ be integrable. Then for any $x \in [a, b]$, we know that f is integrable on $[a, x]$. Hence we have a function $F: x \mapsto \int_a^x f(t) dt$, $x \in [a, b]$. The new function F is called the *indefinite integral* of f . This is the area under the curve $y = f(x)$ between the x -axis, $x = a$, and $x = b$. Draw a picture

Theorem 43 (Second Fundamental Theorem of Calculus). *Let $f: [a, b] \rightarrow \mathbb{R}$ be integrable. The indefinite integral F of f is continuous (in fact, Lipschitz) on $[a, b]$ and is differentiable at x if f is continuous at $x \in [a, b]$. In fact, $F'(x) = f(x)$, if f is continuous at x .*

Why is this result plausible? Draw a picture. It seems that $\int_x^{x+h} f(t) dt$ is approximately the area of the rectangle whose base is h and height is $f(x)$. Hence $\frac{1}{h} \int_x^{x+h} f(t) dt \approx f(x)$. Observe that

$$\frac{1}{h} \left(\int_a^{x+h} f(t) dt - \int_a^x f(t) dt \right) = \frac{1}{h} \int_x^{x+h} f(t) dt \approx f(x).$$

Proof. Since f is bounded, there exists M such that $|f(x)| \leq M$ for $x \in [a, b]$. Then we have

$$\begin{aligned} |F(x) - F(y)| &= \left| \int_a^x f(t) dt - \int_a^y f(t) dt \right| \\ &= \left| \int_x^y f(t) dt \right| \\ &\leq M \left| \int_x^y 1 dt \right| \\ &= M|x - y|. \end{aligned}$$

Thus F is Lipschitz and hence continuous on $[a, b]$.

Let f be continuous at $c \in [a, b]$. We shall show that F is differentiable at c and $F'(c) = f(c)$. Observe that, for $x > c$,

$$\frac{F(x) - F(c)}{x - c} = \frac{1}{x - c} \int_c^x f(t) dt \quad \text{and} \quad f(c) = \frac{1}{x - c} \int_c^x f(c) dt.$$

Hence, we obtain

$$\begin{aligned} \left| \frac{F(x) - F(c)}{x - c} - f(c) \right| &= \left| \frac{1}{x - c} \int_c^x f(t) dt - \frac{1}{x - c} \int_c^x f(c) dt \right| \\ &= \left| \frac{1}{x - c} \int_c^x [f(t) - f(c)] dt \right| \\ &\leq \frac{1}{x - c} \int_c^x |f(t) - f(c)| dt. \end{aligned} \tag{27}$$

Given $\varepsilon > 0$, by the continuity of f at c , we can find a $\delta > 0$ such that $|f(t) - f(c)| < \varepsilon$ for $|t - c| < \delta$. Hence for $x \in [a, b]$ such that $|x - c| < \delta$, we see that the RHS of (27) is estimated above by ε . Similar argument applies when $x < c$.

This shows that F is differentiable at c and $F'(c) = f(c)$. □

Remark 44. Look at the inequality (27). One of the terms in LHS is an integral while the other is a number $f(c)$. We re-wrote this as a sum of two integrals by observing that $f(c)$ is the average $\frac{1}{x-c} \int_c^x f(c) dt$ and then applied the linearity, continuity, and the standard estimate for the integral. Learn this well as this trick is often used.

Remark 45. We can deduce a weaker version of the first fundamental Theorem 41 from the second fundamental theorem of calculus.

Let $f: [a, b] \rightarrow \mathbb{R}$ be differentiable with f' continuous on $[a, b]$. Then

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

Proof. Since f' is continuous, it is integrable and its indefinite integral, say, $G(x) = \int_a^x f'(t) dt$, exists. By the last item, G is differentiable with derivative $G' = f'$. Hence the derivative of $f - G$ is zero on $[a, b]$ and hence the function $f - G$ is a constant on $[a, b]$. In particular, $f(a) - G(a) = f(b) - G(b)$, that is, $f(a) = f(b) - \int_a^b f'(x) dx$. \square

Theorem 46 (Integration by Parts). *Let $u, v: [a, b] \rightarrow \mathbb{R}$ be differentiable. Assume that u', v' are integrable on $[a, b]$. Then*

$$\int_a^b u(x)v'(x) dx = u(x)v(x) \Big|_a^b - \int_a^b u'(x)v(x) dx. \quad (28)$$

Strategy: Let $g := uv$. Then g is integrable, and $g' = u'v + uv'$ is integrable. (Why? If we assume that u and v are continuously differentiable, then the integrability of g' etc. are clear.) Apply the first fundamental theorem of calculus to $\int_a^b g'(x) dx$ to arrive at the result.

Proof. We assume that u and v are continuously differentiable functions. Then $g = uv$ is continuous and hence integrable. Also $g' = u'v + uv'$. Furthermore, g' is continuous and hence integrable. Applying the first fundamental theorem of calculus, we obtain

$$\begin{aligned} g(b) - g(a) &= u(b)v(b) - u(a)v(a) \\ &= \int_a^b u(x)v'(x) dx + \int_a^b u'(x)v(x) dx \end{aligned}$$

The term $u(b)v(b) - u(a)v(a)$, we write as $u(x)v(x) \Big|_a^b$. Hence we have

$$\int_a^b u(x)v'(x) dx = u(x)v(x) \Big|_a^b - \int_a^b u'(x)v(x) dx.$$

\square

One of the most basic tools for computing integration in high school is integration by substitution or the change of variables. The following result justifies this process.

Theorem 47 (Change of Variables). *Let I, J be closed and bounded intervals. Let $u: J \rightarrow \mathbb{R}$ be continuously differentiable. Let $u(J) \subset I$ and $f: I \rightarrow \mathbb{R}$ be continuous. Then $f \circ u$ is continuous on J and we have*

$$\int_a^b f(u(x))u'(x) dx = \int_{u(a)}^{u(b)} f(y) dy, \quad a, b \in J. \quad (29)$$

Proof. Fix $c \in I$. Let $F(y) := \int_c^y f(t) dt$. Then by the second fundamental theorem of calculus (Theorem 43), F is differentiable and $F'(y) = f(y)$. Let $g(x) := (F \circ u)(x)$. Then g is differentiable, and by the chain rule we have

$$g'(x) = F'(u(x))u'(x) = f(u(x))u'(x).$$

We apply the first fundamental theorem of calculus to g' :

$$\begin{aligned} \int_a^b f(u(x))u'(x) dx &= \int_a^b g'(x) dx \\ &= g(b) - g(a) \\ &= F(u(b)) - F(u(a)) \\ &= \int_c^{u(b)} f(t) dt - \int_c^{u(a)} f(t) dt \\ &= \int_{u(a)}^{u(b)} f(t) dt. \end{aligned}$$

This completes the proof. □

We now prove a general version of integration by parts formula for Riemann-Stieltjes integrals.

Theorem 48 (Integration by Parts Formula). *Let $f, g: [a, b] \rightarrow \mathbb{R}$ be bounded functions. Assume that f is g -integrable on $[a, b]$, that is, assume that $\int_a^b f dg$ exists. Then g is f -integrable and we have*

$$\int_a^b g df = f(b)g(b) - f(a)g(a) - \int_a^b f dg, \quad (30)$$

Strategy: Let $A := \int_a^b f dg$. Let $\varepsilon > 0$ be given. Choose $P = \{x_0, \dots, x_n\}$ such that for any partition Q finer than P and for any set of tags \mathbf{t} in Q , we have $|A - S(f, P, \mathbf{t}, g)| < \varepsilon$.

We wish to show that $|-A + f(b)g(b) - f(a)g(a) - S(g, P, \mathbf{t}, f)| < \varepsilon$. Note that we need to move from a Riemann sum of the form

$$S(f, P, \mathbf{t}, g) = \sum_i f(t_i)(g(x_i) - g(x_{i-1})) \quad \text{to} \quad S(g, P, \mathbf{t}, f) = \sum_i g(t_i)(f(x_i) - f(x_{i-1})).$$

This suggests that we consider a partition Q in which the nodes are the nodes of P along with the tags in P . Also, the tags of Q should include the tags \mathbf{t} and the nodes of P . Let $Q = P \cup \mathbf{t}$. Let $\mathbf{s} = \{s_1 = a, s_2 = x_1, s_3 = x_1, \dots, x_{n-1}, x_{n-1}, s_{2n} = b\}$ be the set of tags in Q . Then $|A - S(f, Q, \mathbf{s}, g)| < \varepsilon$. Keeping in mind what we want, rearrange the terms in $S(f, Q, \mathbf{s}, g)$ to arrive at (31).

Proof. Let $A := \int_a^b f dg$. Let $\varepsilon > 0$ be given. Choose $P = \{x_0, \dots, x_n\}$ such that for any partition Q finer than P and for any set of tags \mathbf{t} in Q , we have $|A - S(f, P, \mathbf{t}, g)| < \varepsilon$.

We wish to show that $|-A + f(b)g(b) - f(a)g(a) - S(g, P, \mathbf{t}, f)| < \varepsilon$.

Let $Q = P \cup \mathbf{t}$. Let $\mathbf{s} = \{s_1 = a, s_2 = x_1, s_3 = x_1, \dots, x_{n-1}, x_{n-1}, s_{2n} = b\}$ be the set of tags in Q . Then $|A - S(f, Q, \mathbf{s}, g)| < \varepsilon$.

We claim that

$$S(f, Q, \mathbf{s}, g) = f(b)g(b) - f(a)g(a) - S(g, P, \mathbf{t}, f). \quad (31)$$

The result follows from this.

In order to make the ideas clear, we look at a specific case where $P = \{a = x_0, x_1, x_2, b\}$. Let $Q = \{a, t_1, x_1, t_2, x_2, t_3, b\}$. Note that there are six subintervals. We assume that the set $\mathbf{s} = \{a, x_1, x_1, x_2, x_2, b\}$ of tags are chosen. (Can you say to which subinterval the third x_1 belong? Draw pictures. The point x_i is the tag of the consecutive subintervals $[t_i, x_i]$ and $[x_i, t_{i+1}]$.)

$$\begin{aligned} S(f, Q, \mathbf{s}, g) &= f(a)(g(t_1) - g(a)) + f(x_1)(g(x_1) - g(t_1)) + f(x_1)(g(t_2) - g(x_1)) \\ &\quad + f(x_2)(g(x_2) - g(t_2)) + f(x_2)(g(t_3) - g(x_2)) + f(b)(g(b) - g(t_3)) \\ &= f(b)g(b) - f(a)g(a) - g(t_1)[f(x_1) - f(a)] - g(t_2)[f(x_2) - f(x_1)] - g(t_3)[f(b) - f(x_2)] \\ &= f(b)g(b) - f(a)g(a) - S(g, P, \mathbf{t}, f). \end{aligned}$$

Thus (31) is proved. From it, we obtain

$$|S(f, Q, \mathbf{s}, g) - A| = |f(b)g(b) - f(a)g(a) - S(g, P, \mathbf{t}, f) - A|.$$

Since $Q \geq P$, the left side is less than ε . So the right side is also less than ε . We deduce that

$$-\int_a^b g df = \int_a^b f dg - f(b)g(b) + f(a)g(a).$$

Thus (30) follows. □

Acknowledgment. I thank Vikram Aithal, Tamoghna Kar, Manmohan Sahu, Shivam Bajpai, V Balakumar and Ruby Pinto for a very meticulous proof-reading of a (deliberately minimally proofread) draft. If this article of mine has less number of irritating typos, you should thank them. I also thank Tamoghna for suggesting specific references to our book on real analysis and the YouTube videos.

References

- (1) Ajit Kumar and S Kumaresan, *A Basic Course in Real Analysis*, CRC Press.
- (2) Tom Apostol, *Analysis*, Indian Reprint.
- (3) Walter Rudin, *Principles of Mathematical Analysis*, International Student Edition.
- (4) S Kumaresan and G Santhanam, *Analysis on \mathbb{R}^n* , Forthcoming.
- (5) This article is based on the series of my videos in the YouTube playlist on Riemann Integration:

<https://www.youtube.com/playlist?list=PLDzvuf9Uf4FNveNqPpjBtZqtfWEIj3DKr>

You may either watch them first and then read this article or the other way around.