Notes 2. RIEMANN INTEGRATION

2.1 Integrability Criterion

In MATH1010 we learned that every continuous function on $[a, b]$ is integrable, that is, the area bounded between its graph over $[a, b]$ and the x-axis makes sense. Moreover, functions which are continuous except at finitely many points are also integrable. This shows that the class of integrable functions contains more functions than those continuous ones. It is our aim to characterize these integrable functions. As we are going to see, unlike the definitions of continuity and differentiability which are local, the definition of integrability is global one. It is more subtle to determine whether a function is integrable or not. In this section some integrability criteria will be developed for this purpose.

The setting of the Riemann integral is a bounded function f defined on a bounded, closed interval $[a, b]$. A **partition** of $[a, b]$, P, is a finite collection of points,

$$a = x_0 < x_1 < \cdots < x_n = b,$$

which divides $[a, b]$ into n many subintervals $I_j = [x_{j-1}, x_j], j = 1, \ldots, n$. The length of a partition is given by $||P|| = \max_j (x_j - x_{j-1})$. Given a partition P, a **tag** on this partition is a collection of points $\{z_1, \cdots, z_n\}$ satisfying $z_j \in I_j, j = 1, \cdots, n$. A **tagged partition** is the pair $(P, z_1, ..., z_n)$ where $z_i \in I_j$. We shall use $\dot{P}$ to denote a tagged partition.

Given any tagged partition $\dot{P}$, we define the **Riemann sum** of f with respect to $\dot{P}$ by

$$S(f, \dot{P}) = \sum_{j=1}^{n} f(z_j)\Delta x_j, \qquad \text{where } \Delta x_j = x_j - x_{j-1}.$$

Geometrically, $S(f, \dot{P})$ is an approximate area of the region bounded by $x = a, x = b, y = 0$ and $y = f(x)$ when f is non-negative. We call f **Riemann integrable** on $[a, b]$ if there exists $L \in \mathbb{R}$ such that for every $\varepsilon > 0$, there exists some $\delta > 0$ s.t.

$$|S(f, \dot{P}) - L| < \varepsilon, \qquad \forall P, \quad ||P|| < \delta,$$

for any tag. It is easy to show that such L is uniquely determined whenever it exists. It is called the **Riemann integral** of f over $[a, b]$ and is denoted by

$$\int_a^b f \ , \qquad \text{or} \ \int_a^b f(x)dx \ .$$

When f is non-negative, this number is the area of the region bounded between the graph of the function and the x-axis and the two vertical lines $x = a$ and $x = b$.

Example 2.1. The constant function $f_1(x) = c$ is integrable on $[a, b]$ and $\int_a^b f_1 = c(b-a)$. For, let P be any partition of $[a, b]$, we have $S(f_1, \dot{P}) = \sum_j f_1(z_j)\Delta x_j = \sum_j c\Delta x_j = c(b-a)$, hence

$$\left| S(f, \dot{P}) - c(b-a) \right| = 0 \ ,$$

the conclusion follows.

Example 2.2. Define $f_2(x) = 1$ (x is rational) and $= 0$ (otherwise). In any interval, there are rational and irrational points, hence we can find tags z and w so that $f_2(z) = 1$ and $f_2(w) = 0$. It follows that $S(f_2, \dot{P}) = b - a$ for the former but $S(f_2, \tilde{P}) = 0$ for the latter. Clearly, the number L does not exist, so f_2 is not integrable. Note that f_2 is discontinuous everywhere.

Example 2.3. Let $f_3(x)$ be equal to 0 except at $w_1, \cdots, w_n \in [a, b]$ where $f_3(w_j) \neq 0$. We claim that f_3 is integrable with integral equal to 0. To see this, let P be a partition whose length is δ. Every subinterval of this partition contains or does not contain some w_j's. Hence there are at most $2n$-many subintervals which contain some w_j. Denote the collection of all these subintervals by $\mathcal{B}$. Then

$$\begin{aligned} 0 \leq S(f_3, \dot{P}) - 0 &= \sum_{\mathcal{B}} f_3(z_j)\Delta x_j, \quad M = \{\sup |f_3(x)| : \ x \in [a, b]\} > 0, \\ &\leq M \times 2n \times \delta < \varepsilon, \end{aligned}$$

provided we choose $\delta < \varepsilon/2nM$. Note that f_3 has finitely many discontinuity points.

From these examples we gather the impression that a function is integrable if its points of discontinuity are not so abundant. We will pursue this in the following sections. To proceed, we introduce more concept. First, we use $\mathcal{R}[a, b]$ to denote the set of all Riemann integrable functions on $[a, b]$. For any partition P, we define its **Darboux upper** and **lower sums** respectively by

$$\overline{S}(f, P) = \sum_{j=1}^{n} M_j \Delta x_j,$$

and

$$\underline{S}(f, P) = \sum_{j=1}^{n} m_j \Delta x_j,$$

where $M_j = \sup\{f(x) : x \in [x_{j-1}, x_j]\}$ and $m_j = \inf\{f(x) : x \in [x_{j-1}, x_j]\}$. Since the infimum or supremum of the function may not be attained on a subinterval, a Darboux sum may not be a Riemann sum. However, all Riemann sums are bounded between these two sums. With a given P, although there are infinitely many choices of Riemann sums, there are only two Darboux sums. Essentially the study of Riemann sums is reduced to the study of Darboux sums.

A partition P_2 is called a **refinement** of P_1 if every partition point of P_1 is a partition point of P_2. In the following we show that the Darboux upper sum decreases and the lower sum increases under refinement.

Proposition 2.1. *Let P_2 be a refinement of P_1. Then*

$$\overline{S}(f, P_1) \geq \overline{S}(f, P_2), \tag{2.1}$$

and

$$\underline{S}(f, P_1) \leq \underline{S}(f, P_2). \tag{2.2}$$

Proof. We will only prove (2.1) as (2.2) can be handled in a similar way. Let us consider the simplest case that Q is obtained from P by adding one partition point. Without loss of generality assume $P : x_0 < x_1 < x_2 < \cdots < x_n$ and $Q : x_0 < z < x_1 < x_2 \cdots < x_n$. Letting $M_1 = \sup_{I_1} f$ and $M = \sup_{[x_0,z]} f$, $M' = \sup_{[z,x_1]} f$. We have $M_1 \geq M, M'$. Therefore, observing the cancelation of the terms after x_1,

$$\begin{aligned} \overline{S}(f, Q) - \overline{S}(f, P) &= M(z - x_0) + M'(x_1 - z) - M_1(x_1 - x_0) \\ &= (M - M_1)(z - x_0) + (M' - M_1)(x_1 - z) \\ &\leq 0 , \end{aligned}$$

which shows (2.1) holds when Q has one more partition point of P. In general, (2.1) follows by regarding P_2 as obtained from P_1 in finitely many steps where in every step the partition has one more partition point than its precedent one. □

Now we deduce that a lower Darboux sum cannot be greater than an Darboux upper sum for different partitions.

Proposition 2.2. *For two partitions P and Q,*

$$\underline{S}(f, P) \leq \overline{S}(f, Q). \tag{2.3}$$

Proof. By putting the partition points of P and Q together we obtain a partition R which refines both P and Q. By Proposition 2.1,

$$\underline{S}(f, P) \leq \underline{S}(f, R) \leq \overline{S}(f, R) \leq \overline{S}(f, Q).$$

□

The following proposition is an immediate consequence from the definition of the Darboux sums.

Proposition 2.3. *For every partition P,*

$$\underline{S}(f,P) \le S(f,\dot{P}) \le \overline{S}(f,P).$$

for any tag on P. Moreover, given $\varepsilon > 0$, there exists a tag $\dot{P}$ such that

$$\underline{S}(f,P) + \varepsilon \ge S(f,\dot{P}),$$

and another tag $\ddot{P}$ such that

$$\overline{S}(f,P) - \varepsilon \le S(f,\ddot{P}).$$

Proof. According to the definition of infimum, for every $\varepsilon > 0$, there is some $z_j \in I_j$ such that $m_j + \varepsilon/(b-a) > f(z_j)$. All these z_j's form a tagged partition $\dot{P}$. We have

$$\begin{aligned} \underline{S}(f,P) &= \sum_{j=1}^{n} m_j \Delta x_j \\ &> \sum_{j=1}^{n} \left(f(z_j) - \frac{\varepsilon}{b-a} \right) \Delta x_j \\ &= \sum_{j=1} f(z_j)\Delta x_j - \varepsilon \\ &= S(f,\dot{P}) - \varepsilon \ . \end{aligned}$$

The upper sum can be handled in a similar way.

□

We define the **Riemann upper** and **lower integrals** respectively to be

$$\overline{S}(f) = \inf_P \overline{S}(f,P),$$

and

$$\underline{S}(f) = \sup_P \underline{S}(f,P).$$

Theorem 2.4. *For every $\varepsilon > 0$, there exists some δ such that*

$$0 \le \overline{S}(f,P) - \overline{S}(f) < \varepsilon,$$

and

$$0 \le \underline{S}(f) - \underline{S}(f,P) < \varepsilon,$$

for any partition P, $\|P\| < \delta$.

This theorem asserts that by simply taking any sequence of partitions whose lengths tend to zero, the limit of the corresponding Darboux upper and lower sums always exist and give you the Riemann upper and lower integrals respectively.

Proof. Given $\varepsilon > 0$, there exists a partition Q such that

$$\overline{S}(f) + \varepsilon/2 > \overline{S}(f, Q).$$

Let m be the number of partition points of Q (excluding the endpoints). Consider any partition P and let R be the partition by putting together P and Q. There are at most $2m$ many subintervals of P containing some partition points of Q. (A partition point of Q may be an endpoint of two subintervals of P.) Denote the indices of the collection of these subintervals in P by J. We have

$$0 \leq \overline{S}(f, P) - \overline{S}(f, R) \leq \sum_{j \in J} 2M \Delta x_j \leq 2M \times 2m||P||,$$

where $M = \sup_{[a,b]} |f|$, because the contributions of $\overline{S}(f, P)$ and $\overline{S}(f, Q)$ from the subintervals not in J cancel out. Hence, by Proposition 2.1

$$\overline{S}(f) + \varepsilon/2 > \overline{S}(f, Q) \geq \overline{S}(f, R) \geq \overline{S}(f, P) - 4Mm||P||,$$

i.e.,

$$0 \leq \overline{S}(f, P) - \overline{S}(f) < \varepsilon/2 + 4Mm||P||.$$

Now, we choose

$$\delta < \frac{\varepsilon}{1 + 8Mm},$$

Then for P, $\|P\| < \delta$,

$$0 \leq \overline{S}(f, P) - \overline{S}(f) < \varepsilon.$$

Similarly, one can prove the second inequality. □

Alternatively we can formulate Theorem 2.4 as follows.

Theorem 2.4.' *Let* $\{P_n\}$ *be a sequence of partitions satisfying* $\lim_{n\to\infty} \|P_n\| = 0$. *Then*

$$\lim_{n\to\infty} \overline{S}(f, P_n) = \overline{S}(f),$$

and

$$\lim_{n\to\infty} \underline{S}(f, P_n) = \underline{S}(f).$$

In other words, taking any sequence of partitions whose length tends to 0, its upper and lower Darboux sums form two sequences converging to the upper and lower Riemann sums respectively.

Now we relate the upper/lower Riemann integrals to Riemann integrability.

Theorem 2.5 (The First Integrability Criterion). *Let f be bounded on $[a,b]$. Then f is Riemann integrable on $[a,b]$ if and only if $\overline{S}(f) = \underline{S}(f)$. When this holds, $\int_a^b f = \overline{S}(f) = \underline{S}(f)$.*

Proof. According to the definition of integrability, when f is integrable, there exists some $L \in \mathbb{R}$ so that for any given $\varepsilon > 0$, there is a $\delta > 0$ such that for all partitions P with $||P|| < \delta$,

$$|S(f, \dot{P}) - L| < \varepsilon/2,$$

holds for any tags. Let $\ddot{P}$ be another tagged partition. By the triangle inequality we have

$$|S(f, \dot{P}) - S(f, \ddot{P})| \le |S(f, \dot{P}) - L| + |S(f, \ddot{P}) - L| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Since the tags are arbitrary, it implies

$$\overline{S}(f, P) - \underline{S}(f, P) \le \varepsilon.$$

As a result,

$$0 \le \overline{S}(f) - \underline{S}(f) \le \overline{S}(f, P) - \underline{S}(f, P) \le \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, $\overline{S}(f) = \underline{S}(f)$.

Conversely, by Theorem 2.4, for $\varepsilon > 0$, there exists a δ such that $\overline{S}(f, P) < \overline{S}(f) + \varepsilon$ and $\underline{S}(f) - \varepsilon < \underline{S}(f, P)$ whenever $\|P\| < \delta$. It follows that $S(f, \dot{P}) \le \overline{S}(f, P) < \overline{S}(f) + \varepsilon$ and $S(f, \dot{P}) \ge \underline{S}(f, P) > \underline{S}(f) - \varepsilon$. When $L = \overline{S}(f) = \underline{S}(f)$, this means $L - \varepsilon < S(f, \dot{P}) < L + \varepsilon$, so f is integrable with $L = \underline{S}(f) = \overline{S}(f)$. □

Combining Theorem 2.4' and the Integrability Criterion I, we have the following useful way of evaluating integral.

Theorem 2.6. *For an integrable function f, its integral over $[a,b]$ is equal to the limit of $\overline{S}(f, P_n)$, $\underline{S}(f, P_n)$ or $S(f, \dot{P}_n)$ for any sequence of (tagged) partitions P_n, $||P_n|| \to 0$.*

Keep in mind that you need to know that f is integrable first before you can apply this theorem.

Example 2.4. We show that the linear function $f(x) = x$ is integrable on $[a,b]$ with integral given by $(b^2 - a^2)/2$. To see this we note that f is increasing, so for any partition P, we have

$$\overline{S}(f, P) = \sum_1^n x_j \Delta x_j, \quad \underline{S}(f, P) = \sum_1^n x_{j-1} \Delta x_j.$$

Therefore,

$$\overline{S}(f) - \underline{S}(f) \le \overline{S}(f,P) - \underline{S}(f,P) \le \sum_1^n \Delta x_j \Delta x_j .$$

It follows that

$$\overline{S}(f) - \underline{S}(f) \le (b-a)\|P\|.$$

By taking $P = P_n, \|P_n\| \to 0$ we conclude the upper and lower integrals coincide, so f is integrable by Theorem 2.5.

To evaluate the integral, we make a good of tag points by letting $z_j = (x_j + x_{j-1})/2$, then

$$S(f, \dot{P}_n) = \frac{1}{2}\sum_1^n z_j \Delta x_j = \frac{1}{2}\sum_1^n (x_j^2 - x_{j-1}^2) = \frac{1}{2}(b^2 - a^2),$$

which is independent of n. Letting $n \to \infty$, we conclude from Proposition 2.6 that the integral of the linear function over $[a,b]$ is equal to $(b^2 - a^2)/2$.

By tricky choices of tag points one can evaluate the integrals of all monomials.

Next we formulate our second criterion. Essentially nothing new, this alternative formulation is convenient in application. We define the **oscillation** of a (bounded) function f over an interval I to be

$$\text{osc}_I f = \sup\{|f(x) - f(y)| : \quad x, y \in I\}$$

It is clear that

$$\text{osc}_I f = \sup_I f - \inf_I f.$$

Using this concept we can reformulate our first criterion into our second integrability criterion. Its proof is immediate from the first one.

Theorem 2.7 (**The Second Integrability Criterion**). *Let f be a bounded function on $[a,b]$. Then f is Riemann integrable on $[a,b]$ if and only if for every $\varepsilon > 0$, there exists a partition P such that*

$$\sum_{j=1}^n osc_{I_j} f \Delta x_j < \varepsilon.$$

Here we do not need $\|P\|$ to be small. This is sometimes convenient in applications.

Proof. We have just shown that f is integrable if and only if $\overline{S}(f) = \underline{S}(f)$. By Theorem 2.4, for $\varepsilon > 0$, there exists some partition P such that $\overline{S}(f,P) - \overline{S}(f) < \varepsilon$

and $\underline{S}(f) - \underline{S}(f, P) < \varepsilon$. Then

$$\sum_{1}^{n} \mathrm{osc}_{I_j} f \Delta x_j = \overline{S}(f, P) - \underline{S}(f, P) < \varepsilon \ .$$

On the other hand, suppose for each $\varepsilon > 0$, there is a partition P such that $|\sum_j \mathrm{osc}_{I_j} f \Delta x_j| < \varepsilon$. We have

$$0 \leq \overline{S}(f) - \underline{S}(f) \leq \overline{S}(f, P) - \underline{S}(f, P) = \sum_j \mathrm{osc}_{I_j} f \Delta x_j < \varepsilon \ .$$

Since ε is arbitrary, we conclude that $\overline{S}(f) = \underline{S}(f)$, and hence f is integrable. □

Remark 2.1. It is clear that

$$\mathrm{osc}_I f \leq \mathrm{osc}_J f \ , \quad I \subset J \ .$$

Therefore, whenever Q is a refinement of P,

$$\sum_Q \mathrm{osc}_j f \Delta x_j \leq \sum_P \mathrm{osc}_{J_j} f \Delta y_j \ ,$$

where $I_j = [x_j, x_{j+1}]$ and $J_j = [y_j, y_{j+1}]$ are subintervals of P and Q respectively.

Using either one of these criteria we now show that Riemann integrability is preserved under vector space operations, multiplication and division. One may also deduce it right from the definition, but the application of the integrability criterion makes things clean.

Theorem 2.8. *Let f and g be integrable on $[a, b]$ and $\alpha, \beta \in \mathbb{R}$. We have*
(a) $\alpha f + \beta g$ is integrable on $[a, b]$ and

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g,$$

(b) fg is integrable on $[a, b]$,
(c) f/g is integrable on $[a, b]$ provided $|g| \geq \rho$ for some positive number ρ, and
(d) $|f|$ is integrable on $[a, b]$ and

$$\left| \int_a^b f \right| \leq \int_a^b |f|.$$

(e) $f \leq g$ implies

$$\int_a^b f \leq \int_a^b g \ .$$

(f) f is integrable on every $[c,d] \subset [a,b]$.

Proof. (a). We use the definition to prove (a). As f and g are integrable, for any $\varepsilon > 0$, there exists δ such that

$$\left| S(f,\dot{P}) - \int_a^b f \right| < \frac{\varepsilon}{2(1+|\alpha|)}, \quad \left| S(g,\dot{P}) - \int_a^b g \right| < \frac{\varepsilon}{2(1+|\beta|)},$$

for $\|P\| < \delta$. Using the formula,

$$S(\alpha f + \beta g, \dot{P}) = \alpha S(f,\dot{P}) + \beta S(f,\dot{P}),$$

we have

$$\begin{aligned}
\left| S(\alpha f + \beta g, \dot{P}) - \alpha \int_a^b f - \beta \int_a^b g \right| &= \left| \alpha S(f,\dot{P}) + \beta S(g,\dot{P}) - \alpha \int_a^b f - \beta \int_a^b g \right| \\
&\le |\alpha| \left| S(f,\dot{P}) - \int_a^b f \right| + |\beta| \left| S(g,\dot{P}) - \int_a^b g \right| \\
&< \frac{|\alpha|\varepsilon}{2(1+|\alpha|)} + \frac{|\beta|\varepsilon}{2(1+|\beta|)} \\
&< \varepsilon \ ,
\end{aligned}$$

and the conclusion follows.

(b) Observe that

$$|f(x)g(x) - f(y)g(y)| \le |f(x)||g(x) - g(y)| + |g(y)||f(x) - f(y)|$$

implies

$$\text{osc}_I fg \le M_2 \text{osc}_I f + M_1 \text{osc}_I g \ ,$$

where $M_1 = \sup |f|, M_2 = \sup |g|$ and I is an interval. By Integrability Criterion II, given $\varepsilon > 0$, there exist partitions P_1 and P_1 such that

$$\sum_{P_1} \text{osc} f \Delta x_j < \frac{\varepsilon}{2M_2 + 1} \ , \quad \sum_{P_2} \text{osc} g \Delta y_j < \frac{\varepsilon}{2M_1 + 1} \ .$$

Putting P_1 and P_2 together to form a refinement P, by Remark 2.1,

$$\sum_{P} \text{osc} f \Delta z_j < \frac{\varepsilon}{2M_2 + 1} \ , \quad \sum_{P} \text{osc} g \Delta x_j < \frac{\varepsilon}{2M_1 + 1} \ .$$

Therefore,

$$\begin{aligned}\sum_P \text{osc} fg \Delta x_j &< M_2 \sum_P \text{osc} f \Delta z_j + M_1 \sum_P \text{osc} g \Delta z_j \\ &\le \frac{M_2}{2M_2+1} + \frac{M_1}{2M_1+1} \\ &< \varepsilon \ .\end{aligned}$$

By Integrability Criterion II that fg is integrable.

(c) It suffices to show that $1/g$ is integrable when g is integrable and $|g| \ge \rho > 0$. Together with (b) it implies (c). Using

$$\left|\frac{1}{g(x)} - \frac{1}{g(y)}\right| = \frac{|g(y)-g(x)|}{|g(x)g(y)|} \le \frac{1}{\rho^2}|g(x)-g(y)|,$$

we see that $\text{osc}_I g^{-1} \le \rho^{-2}\text{osc}_I g$ over any interval I, and the desired conclusion follows from the second criterion.

(d) The integrability of $|f|$ comes from the observation $\text{osc}_I|f| \le \text{osc}_I f$. (Why?) For any tagged partition $\dot{P}$,

$$-S(|f|, \dot{P}) \le S(f, \dot{P}) \le S(|f|, \dot{P}) \ .$$

Letting $P = P_n$ with $\|P_n\| \to 0$, we obtain

$$-\int_a^b |f| \le \int_a^b f \le \int_a^b |f| \ .$$

(e) In view of (a), it suffices to show $f \ge 0$ implies its integral is non-negative. But this follows from

$$\int_a^b f = \lim_{n\to\infty} S(f, \dot{P}_n)$$

for any $\|P_n\| \to 0$ and $S(f, \dot{P}) \ge 0$ for all tagged partitions $\dot{P}$ whenever f is non-negative.

(f) f is integrable on $[a, b]$ if and only if there is a partition R with $\sum_j \text{osc}_{I_k} f \Delta x_k < \varepsilon$. Now we refine R to R_1 be putting in two points c and d (if they are not yet in R). Denoting the partition of $[c, d]$ formed by restricting R_1 to the subintervals inside $[c, d]$ by Q. By Remark 2.1,

$$\begin{aligned}\sum_Q \operatorname{osc}_{I_k} f \Delta y_k &\le \sum_{R_1} \operatorname{osc}_{I_k} f \Delta y_k \\ &\le \sum_R \operatorname{osc}_{I_k} f \Delta x_k \\ &< \varepsilon ,\end{aligned}$$

so f is integrable on $[c, d]$ too.

□

A further remark to Theorem 2.8(e). According to this result, the integral of a non-negative function is always non-negative. It is natural to ask: Is it true the function must identically be zero if its integral vanishes? The answer is no. For instance, a bounded function which vanishes everywhere except at finitely many points is integrable and yet it is not a zero function. The characterization of non-negative functions with zero integral has to wait until the notion of a measure zero set is introduced. Then one can show that the integral of a non-negative integrable function vanishes if and only if its discontinuity set is a set of measure zero. We will not go into this direction.

Denoting the collection of all Riemann integrable functions by $R[a, b]$, Theorem 2.8(a)(b) in particular shows that the collection of all Riemann integrable functions form a vector space which is closed under multiplication. Furthermore, the map $\mathcal{J} : R[a, b] \to \mathbb{R}$ given by

$$\mathcal{J}(f) = \int_a^b f$$

is a linear map. In the next section we will show that every continuous function is integrable. Since bounded functions with finitely many discontinuous points are integrable, we have the proper inclusion

$$C[a, b] \subsetneqq R[a, b] \ .$$

As continuity and differentiability (the chain rule) are preserved under composition of functions, it is natural to ask if integrability enjoys the same property. Unfortunately, this is not true. There are examples showing that the composition of two integrable functions may not be integrable. On the other hand, it can be shown that the composition of an integrable function with a continuous function is integrable. More precisely, we have

Theorem 2.9. *Let $f \in R[a, b]$ and Φ continuous on $\mathbb{R}$. Then $\Phi \circ f \in R[a, b]$.*

We will leave the proof of this theorem as an exercise. Note that using this theorem, $|f|, f^n, p(f)$ and e^f where $n \geq 1$ and p is a polynomial are all integrable.

To end this section, we would like to point out that although the two criteria provide efficient means to verify integrability, they do not tell how to compute the integral. To achieve this job, we need to use Theorem 2.6. By choosing a suitable sequence of partitions with length tending to zero and suitable tags on them, the integral can be obtained by evaluating the limit of the Riemann sums. Thus we have the freedom in choosing the partitions as well as the tags. In fact, the concept of using approximate sum of rectangles to calculate areas or volumes were known in many ancient cultures. In particular, in the works of Archimedes the areas and volumes of many common geometric objects were found by using ingenious methods. In terms of modern calculus, he used good choices of partitions and tags. This method, of course, cannot be pushed too far. We have to wait more than one thousand years until Newton related integration to differentiation. Then the evaluation of integrals becomes much easier. We shall discuss this shortly in the fundamental theorem of calculus.

2.2 Integrable Functions

In this section we apply the integrability criteria obtained in the last section to show that continuous functions and monotone functions are Riemann integrable.

Theorem 2.10. *Every continuous function on $[a, b]$ is integrable.*

Proof. That f is continuous on $[a, b]$ implies that it is bounded and uniformly continuous on $[a, b]$. For $\varepsilon > 0$, there is some $\delta > 0$ such that

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a}, \ \forall x, y \in [a, b], \ |x - y| < \delta.$$

Consider any P, $||P|| < \delta$, we have

$$\overline{S}(f, P) = \sum_j f(z_j)\Delta x_j,$$
$$\underline{S}(f, P) = \sum_j f(w_j)\Delta x_j \ .$$

Since f is continuous, we can find points $z_j, w_j \in I_j$ such that $f(z_j) = M_j$ and

$f(w_j) = m_j$ respectively. Therefore,

$$\begin{aligned}\sum_{j=1}^{n} \text{osc}_{I_j} f \Delta x_j &= \overline{S}(f, P) - \underline{S}(f, P) \\ &= \sum_j (f(z_j) - f(w_j))\Delta x_j, \\ &< \frac{\varepsilon}{b-a} \times (b-a) = \varepsilon\ .\end{aligned}$$

By Integrability Criterion II, f is integrable on $[a, b]$.

□

Now consider the following situation. Let $f \in R[a, b]$ and $g \in R[b, c]$. We may put them together to form a new function by setting $F(x) = f(x)$ on $[a, b)$ and $F(x) = g(x)$ on $(b, c]$. As $f(b)$ may not be equal to $g(a)$, there is no preferred way to define $F(b)$. Let us assign any value to it by setting $F(b) = \alpha$, say. Now we have a bounded function F on $[a, c]$. The natural questions are: Is F integrable over $[a, c]$? Is it true that

$$\int_a^c F = \int_a^b f + \int_b^c g\ ?$$

It turns out both questions have positive answers. In the following we formulate a general theorem.

Theorem 2.11. *Let $f_j, j = 0, \cdots, n-1$ be integrable on $[a_j, a_{j+1}]$ where $a < a_0 < a_1 < \cdots < a_{n-1} < a_n = b$. Suppose that F is a function which is equal to f_j on (a_j, a_{j+1}) for all j. Then F is integrable on $[a, b]$ and*

$$\int_a^b F = \sum_{j=1}^{n} \int_{a_j}^{a_{j+1}} f_j\ .$$

Here the values of F may not be equal to those of f_j's at the endpoints.

Proof. Clearly it suffices to assume $n = 2$, that is, there is some c, $a < c < b$ and $f_1, f_2 = f, g$ respectively. By Integrability Criterion II, for any $\varepsilon > 0$, we can find partitions P and Q on $[a, c]$ and $[c, b]$ respectively such that

$$\sum_P \text{osc}\ f \Delta x_j < \frac{\varepsilon}{3}, \quad \sum_Q \text{osc}\ g \Delta y_j < \frac{\varepsilon}{3}\ ,$$

where $P :\ a = x_0 < \cdots < x_n = c$ and $Q :\ c = y_0 < \cdots < y_m = b$. By refining P and Q if necessary, we may assume that the length of P and Q are less than $\varepsilon/(24M + 1)$ where $M = \sup_{[a,b]} |F|$. Let R be the partition on $[a, c]$ formed by putting P and Q together. As $F = f$ on (a, c), osc $F =$ osc f on each subinterval of P except the first and the last one. The same is true for osc $F =$ osc g on each subinterval of Q except the first one and the last one. We have

$$\begin{aligned}
\sum_R \text{osc} F \Delta x_j &= \sum_P \text{osc} F \Delta x_j + \sum_Q \text{osc} F \Delta y_j \\
&= \sum_{j=2}^{n-1} \text{osc} F \Delta x_j + \text{osc} F \Delta x_1 + \text{osc} F \Delta x_n \\
&\quad + \sum_{j=2}^{m-1} \text{osc} F \Delta y_j + \text{osc} F \Delta y_1 + \text{osc} F \Delta y_m \\
&= \sum_{j=2}^{n-1} \text{osc} f \Delta x_j + \text{osc} F \Delta x_1 + \text{osc} F \Delta x_n \\
&\quad + \sum_{j=2}^{m-1} \text{osc} g \Delta y_j + \text{osc} F \Delta y_1 + \text{osc} F \Delta y_m \\
&\leq \sum_{j=1}^{n} \text{osc} f \Delta x_j + \text{osc} F \Delta x_1 + \text{osc} F \Delta x_n \\
&\quad + \sum_{j=1}^{m} \text{osc} g \Delta y_j + \text{osc} F \Delta y_1 + \text{osc} F \Delta y_m \\
&\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + 4 \times 2M \times \frac{\varepsilon}{24M+1} < \varepsilon \ .
\end{aligned}$$

By the second criterion, F is integrable on $[a, b]$.

To find the integral, we let P_n and Q_n be partitions of $[a, c]$ and $[c, b]$ respectively with lengths tending to zero. Then the lengths of the partitions $R_n = P_n \cup Q_n$ tend to zero too. Taking the tags lying on the interior of each subinterval of R_n, then $S(F, \dot{R}_n) = S(f, \dot{P}_n) + S(g, \dot{Q}_n)$ and, according to Theorem 2.6,

$$\begin{aligned}
\int_a^b F &= \lim_{n \to \infty} S(F, \dot{R}_n) \\
&= \lim_{n \to \infty} S(f, \dot{P}_n) + \lim_{n \to \infty} S(g, \dot{Q}_n) \\
&= \int_a^c f + \int_c^b g.
\end{aligned}$$

□

We point out that when applying to the same function f on $[a, b]$ and $[b, c]$, this theorem yields

$$\int_a^c f = \int_a^b f + \int_b^c f.$$

In practise it is frequently encountered the integral limits a, b, and c are unordered. To facilitate this situation we adapt the following convention: For $a < b$,

$$\int_b^a f = -\int_a^b f,$$

and

$$\int_a^a f = 0.$$

Under this convention we have

$$\int_a^c f = \int_a^b f + \int_b^c f,$$

for any a, b, and c regardless of their ordering. Verify it for yourself.

Next we consider another class of integrable functions.

Theorem 2.12. *Every monotone function on $[a, b]$ is integrable.*

Proof. Take f to be increasing. Let P be the partition which divides $[a, b]$ equally. Observing that $M_j - m_j = f(x_j) - f(x_{j-1})$. For any $\varepsilon > 0$,

$$\begin{aligned}
\sum_{j=1}^n \operatorname{osc}_{I_j} f \Delta x_j &= \sum_{j=1}^n (M_j - m_j)\Delta x_j \\
&= \frac{b-a}{n} \sum_{j=1}^n (f(x_j) - f(x_{j-1})) \\
&= \frac{(b-a)(f(b) - f(a))}{n} \\
&< \varepsilon,
\end{aligned}$$

if we choose n so large that $(b-a)(f(b) - f(a))/\varepsilon < n$. By the Integrability Criterion II, f is integrable on $[a, b]$.

□

Monotone functions could be very bad in the sense that they have countably many jumps. For instance, let all rational numbers in $(0, 1)$ be written in a sequence $\{x_j\}$ and define $\varphi(x) = \sum_{\text{all } j, x_j < x} 2^{-j}$. It is a good exercise to verify that φ is strictly increasing and continuous precisely at irrational numbers in $(0, 1)$.

We have shown that continuous functions and monotone functions are integrable. Some more complicated functions may still be integrable. In the following we show that Thomae's function is integrable. In last semester we saw that this function is discontinuous at rational points and continuous at irrational points in the unit interval.

Example 2.5. Recall that Thomae's function $h : [0,1] \to \mathbb{R}$ is given by

$$h(x) = \begin{cases} 0 & \text{if } x \text{ is irrational },\\ \dfrac{1}{q} & \text{if } x = \dfrac{p}{q}, \text{ for some } p,q \in \mathbb{N} \text{ with } (p,q) = 1, \end{cases}$$

where (p,q) denotes the greatest common divisor of p and q. For instance $h(3/12) = 1/4$ but not $1/12$ since $(1,4) = 1$ We set $h(0) = 1$.

We show that $h \in R[0,1]$. The key idea is the following observation: Given $q_0 \in \mathbb{N}$, the number of points in $E_{q_0} = \{x \in [0,1] : h(x) \geq 1/q_0\}$ is a finite set depending on q_0. For, as $h(x) \geq 1/q_0 > 0$, x must be a rational number. Assuming that it is of the form p/q, where $(p,q) = 1$, $0 < p \leq q$. So, $h(x) = 1/q \geq 1/q_0$, it means $1 \leq q \leq q_0$. From the two inequalities $1 \leq q \leq q_0$ and $1 \leq p \leq q$, we see that the number of elements in E_{q_0} must not be more than q_0^2. Although this is a rough estimate, it is sufficient for our purpose.

Now, given $\varepsilon > 0$, we fix $q_0 \in \mathbb{N}$ such that $1/q_0 < \varepsilon/2$. There are at most $N_0 \equiv q_0^2$ many points x_j in $[0,1]$ such that $h(x_j) \geq 1/q_0$, $j = 1, \cdots, N_0$. For any partition P, there are at most $2N_0$ many subintervals touching some x_j, and the rest are disjoint from them. Call the former "bad" and the latter "good" subintervals. Now, let δ be chosen such that $\delta \leq \varepsilon/(4N_0+1)$. Then, for any partition P with length less than δ, we have

$$\begin{aligned} 0 \leq S(h,\dot{P}) - 0 &\leq \sum_j h(z_j)\Delta x_j \\ &\leq \sum_{bad} h(z_j)\Delta x_j + \sum_{good} h(z_j)\Delta x_j \\ &\leq 1 \times \delta \times 2N_0 + \frac{1}{q_0} \times (1-0) \\ &\leq \frac{\varepsilon}{4N_0+1} \times 2N_0 + \frac{1}{q_0} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

From the definition of Riemann integral, h is integrable and its integral is 0 over $[0,1]$.

Thus it is an interesting problem to find necessary and sufficient conditions for Riemann integrability. The solution was found by Lebesgue in the beginning of the twentieth century. It asserts that a bounded function is integrable if and only if its points of discontinuity form a set of measure zero. A countable set is of measure zero. However, there are uncountable sets of measure zero. Further discussion on Lebesgue's theorem can be found in the last section of this chapter.

2.3 The Fundamental Theorem of Calculus

Newton discovered that integration and differentiation are inverse to each other. The word "inverse" here cannot be taken too strict. We have seen that differentiation $\mathcal{D} = d/dx$ is a linear transformation from $D(a,b)$ to $F(a,b)$. On the other hand, for any $f \in R[a,b]$, the **indefinite integral** F of f, which is defined by

$$F(x) = \int_a^x f = \int_a^x f(t)dt$$

is a well-defined function on (a,b). Furthermore, one can define $\mathcal{J}$ by $\mathcal{J}f = F$, which forms a linear transformation from $R[a,b]$ to $F(a,b)$. In an ideal setting, one would like to see if there exist $\mathcal{J} : R[a,b] \to D(a,b)$ and $\mathcal{D} : D(a,b) \to R[a,b]$ such that $\mathcal{J}\mathcal{D}f = f$, $\forall f \in D(a,b)$, and $\mathcal{D}\mathcal{J}f = f$, $\forall f \in R[a,b]$. Unfortunately, this is not true for (at least) two reasons. First, we have already seen that $\mathcal{D}$ is not injective, the derivative of any constant function is equal to zero. As a result, $\mathcal{J}\mathcal{D}f = f$ can never hold for non-zero constant functions. Next, $\mathcal{J}(R[a,b])$ is not contained in $D(a,b)$. According to Darboux theorem, the function $f(x) = 1$ for $x \geq 0$, and $= 0$ for $x < 0$, which is in $R[-1,1]$, cannot be the derivative of any differentiable function. Also, $\mathcal{J}f$ is not differentiable at 0 and so $\mathcal{J}f \notin D(-1,1)$. Hence $\mathcal{D}\mathcal{J}$ may not make sense on $R[a,b]$.

In view of these considerations, additional conditions are needed for the validity of the fundamental theorems. We must be careful in formulating the fundamental theorems. Here is the first form, the one corresponding to the case $\mathcal{J}\mathcal{D}f = f$.

Theorem 2.13 (**First Fundamental Theorem of Calculus**). *Let F be differentiable on $[a,b]$ and $F' \in R[a,b]$. Then,*

$$\int_a^x F'(t)dt = F(x) - F(a), \quad \forall x \in [a,b].$$

Proof. Denote $F' = f$. It suffices to prove the theorem for $x = b$. As $f \in R[a,b]$, for each $\varepsilon > 0$, $\exists \delta > 0$ such that

$$\Big| \int_a^b f - \sum_{j=1}^n f(z_j)\Delta x_j \Big| < \varepsilon, \quad \text{whenever } \|P\| < \delta \text{ and for any tag on } P.$$

We partition $[a,b]$ into $x_0 = a < x_1 < \cdots < x_n = b$ to form a partition P such that $\|P\| < \delta$ and then write

$$F(b) - F(a) = \sum_{j=1}^{n} F(x_j) - F(x_{j-1}).$$

Applying Mean-Value Theorem to F on each $[x_{j-1}, x_j]$, we find $z_j \in (x_{j-1}, x_j)$ such that

$$F(x_j) - F(x_{j-1}) = f(z_j)(x_j - x_{j-1}).$$

Taking z_j to be the tags for this P, we have

$$\left| \int_a^b f - (F(b) - F(a)) \right| = \left| \int_a^b f - \sum_{j=1}^{n} f(z_j)\Delta x_j \right| < \varepsilon.$$

So, the theorem follows as $\varepsilon > 0$ is arbitrary. □

A function F is called a **primitive function** of f if F is differentiable and $F' = f$. This theorem tells us that

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

It reduces the evaluation of **definite integral** to the evaluation of **indefinite integral** (that is, finding a primitive function). This provides the most efficient way to evaluate integrals. For instance, the evaluation of $\int_0^1 x^k$ becomes more and more difficult using the old method of smart choice of tagged points as k increases. However, using the simple fact that $x^{k+1}/(k+1)$ is a primitive function for x^k, by the first fundamental theorem we immediately deduce

$$\int_0^1 x^k = \frac{x^{k+1}}{k+1}\Big|_0^1 = \frac{1}{k+1}.$$

Next, we turn to the case $\mathcal{DJ}f = f$.

Theorem 2.14 (**Second Fundamental Theorem of Calculus**). *Let $f \in R[a,b]$ and*

$$F(x) = \int_a^x f \ .$$

Then,

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

(b) F is differentiable at every continuous point c of f. Moreover, $F'(c) = f(c)$.

Proof. Let $M = \sup_{x\in[a,b]} |f(x)|$. For $h \geq 0$, we have

$$|F(x+h) - F(x)| = \left| \int_x^{x+h} f \right| \leq Mh \ .$$

When $h < 0$, we have

$$|F(x+h) - F(x)| = \left| \int_{x+h}^{x} f \right| \leq M|h| \ .$$

Therefore, for $\varepsilon > 0$, we choose $\delta < \varepsilon/(M+1)$. Then

$$|F(x+h) - F(x)| < \varepsilon \ , \quad |h| < \delta, \ x+h \in [a,b] \ ,$$

and (a) holds.

Next, we prove (b). We only consider the case when $c \in (a,b)$, while the case $c = a$ or b can be treated similarly. Fix $c \in (a,b)$. For $|h| > 0$ small,

$$\frac{F(c+h) - F(c)}{h} = \frac{1}{h}\Big(\int_a^{c+h} f - \int_a^c f \Big) = \frac{1}{h} \int_c^{c+h} f.$$

As f is continuous at c, for each $\varepsilon > 0$, $\exists \delta > 0$ such that

$$|f(x) - f(c)| < \frac{\varepsilon}{2}, \quad \forall x \in (c-\delta, c+\delta).$$

For $0 < |h| < \delta$,

$$\Big| \frac{F(c+h) - F(c)}{h} - f(c) \Big| = \Big| \frac{1}{h} \int_c^{c+h} (f(t) - f(c))dt \Big| \leq \frac{\varepsilon}{2} < \varepsilon.$$

We conclude that $F'(c)$ exists and is equal to $f(c)$. □

F may not be differentiable at c if c is not a continuous point of f. For instance, consider

$$f(x) = \begin{cases} 1 & \text{if } x \in [0,1], \\ -1 & \text{if } x \in [-1,0). \end{cases}$$

Then, $F(x) = |x|$ which is not differentiable at the origin.

As an application of the fundamental theorem, we prove the formula on change of variables, or the substitution rule as sometimes called.

Theorem 2.15 (**Change of Variables**). *Let f be a continuous function on some interval I. Suppose $\varphi : [\alpha, \beta] \to I$ is a differentiable function with $\varphi' \in R[\alpha, \beta]$.*

Then,

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt, \quad \textit{where } a = \varphi(\alpha),\, b = \varphi(\beta).$$

The function $f(\varphi(t))$ is continuous and hence integrable. As a result, as the product of two integrable functions, the integrand in the integral in the right hand side is also integrable. Also note that it is not required $\varphi[\alpha, \beta] = [a, b]$, nor do we need φ to be increasing.

Proof. We will assume $a < b$ in the proof. The other case $a \geq b$ can be handled similarly. The function

$$F(x) = \int_a^x f\ , \quad a = \varphi(\alpha)$$

is differentiable according to the Second Fundamental Theorem and $F' = f$. Therefore, the composite function $F \circ \varphi$ is differentiable and, by the Chain Rule,

$$\frac{d}{dt} F \circ \varphi(t) = f(\varphi(t))\varphi'(t)\ .$$

By the First Fundamental Theorem,

$$F(\varphi(\beta)) - F(\varphi(\alpha)) = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt\ .$$

On the other hand, by integrating $F' = f$ from a to b we have

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

so the formula follows after noting $a = \varphi(\alpha)$ and $b = \varphi(\beta)$.

□

Example 2.6. Evaluate $\displaystyle\int_0^1 \sqrt{1 - x^2}dx$.

Let $x = \varphi(t) = \sin t$, $\forall t \in [0, \pi/2]$. Then, $\varphi'(t) = \cos t$, so

$$\int_0^1 \sqrt{1 - x^2}dx = \int_0^{\frac{\pi}{2}} \sqrt{1 - \sin^2 t} \cos t dt = \int_0^{\frac{\pi}{2}} \cos^2 t dt = \frac{\pi}{4}.$$

We can also use the same function φ but now on a different interval $[0, 5\pi/2]$. It is no longer monotone. The result is the same:

$$\begin{aligned}
\int_0^1 \sqrt{1-x^2}dx &= \int_0^{\frac{5\pi}{2}} \sqrt{1-\sin^2 t}\cos t dt \\
&= \int_0^{5\pi/2} |\cos t|\cos t dt \\
&= \int_0^{\pi/2} \cos^2 t dt + \int_{\pi/2}^{3\pi/2} (-\cos^2 t) dt + \int_{3\pi/2}^{5\pi/2} \cos^2 t dt \\
&= \frac{\pi}{4},
\end{aligned}$$

which yields the same result. Be careful of the cancelation of the integrals over $[\pi/2, 3\pi/2]$ and over $[3\pi/2, 5\pi/2]$.

2.4 Integration by Parts and Applications

In this section we discuss the formula on integration by parts and use it to prove another version of the Taylor expansion theorem.

Theorem 2.16 (**Integration by Parts**). *Let F and G be differentiable on $[a, b]$ and $f = F'$, $g = G'$ be in $R[a, b]$. Then*

$$\int_a^b fG = FG\Big|_a^b - \int_a^b Fg,$$

where $FG\Big|_a^b = F(b)G(b) - F(a)G(a)$.

Proof. By assumption, $(FG)' = fG + Fg$ by the product rule of differentiation. Since F and G are differentiable and hence continuous on $[a, b]$, it follows that F and G are integrable on $[a, b]$. According to Theorem 2.8(b), fG and Fg are integrable, and thus $fG + Fg$ is integrable by the same proposition. Hence this proposition follows from the First Fundamental Theorem of Calculus. □

An interesting application of integration by parts is the following Taylor's Expansion Theorem with integral remainder.

Theorem 2.17 (**Taylor's Expansion Theorem with Integral Remainder**). *Suppose $f, \ldots, f^{(n+1)}$ exist on (a, b) and $f^{(n+1)} \in R[\alpha, \beta]$ for any $a < \alpha < \beta < b$. Then, $\forall x_0, x \in (a, b)$,*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{1}{n!}\int_{x_0}^x f^{(n+1)}(t)(x - t)^n dt.$$

You should compare this theorem with Taylor's Expansion Theorem with Lagrange remainder (Theorem 6.4.1 in the textbook). In Theorem 6.4.1 the regularity requirement on f is weaker: $f^{(n+1)} \in R[\alpha, \beta]$ is not needed and the remainder (error) is given by

$$\frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$$

for some c between x and x_0. Now, with a little more assumption, the remainder term is more precise. The vague c is replaced by an integral.

Proof. Let $F(t) = f^{(n)}(t)$, $G(t) = (x-t)^n/n!$ (and so $g(t) = -(x-t)^{n-1}/(n-1)!$). By integration by parts, one has

$$\begin{aligned} \frac{1}{n!}\int_{x_0}^{x} f^{(n+1)}(t)(x-t)^n dt &= \int_{x_0}^{x} F'(t)G(t)dt \\ &= \frac{1}{n!} f^{(n)}(t)(x-t)^n\Big|_{x_0}^{x} + \frac{1}{(n-1)!}\int_{x_0}^{x} f^{(n)}(t)(x-t)^{n-1}dt \\ &= -\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \frac{1}{(n-1)!}\int_{x_0}^{x} f^{(n)}(t)(x-t)^{n-1}dt. \end{aligned}$$

Keep integrating by parts we get the complete formula. Or you may use mathematical induction. □

2.5 Improper Integrals

Very often we face the situation where f is unbounded on some bounded interval, for instance, $(a, b]$, or the domain of integration is unbounded, for example $[a, \infty)$ or $(-\infty, \infty)$. As the setting of Riemann integration is a bounded function over a closed and bounded interval, we need to extend the concept of integration to accommodate these new situations. These generalized integrals are called improper integrals. They are rather common in applications.

We briefly discuss two typical cases.

The first type: f on $(a, b]$ which is bounded on any subinterval $[a', b]$ of $(a, b]$, where $a' \in (a, b)$ (so f is allowed to become unbounded as x tends to a). For instance, the following integrals belong to this type:

$$\int_0^2 \frac{\sin x}{x^p}\,dx,\ p > 0\ , \quad \int_2^5 \frac{x^2 \log(x-2)}{\sqrt{x^2-4}}\,dx\ .$$

Let f be a function defined on $(a, b]$ which is bounded on $[a', b]$ for each $a', a < a' < b$. We call f has an **improper integral** on $(a, b]$ or **improperly**

integrable if $f \in R[a', b]$, $\forall a' > a$, and

$$\lim_{a' \to a^+} \int_{a'}^{b} f \quad \text{exists.}$$

We define

$$\int_a^b f = \lim_{a' \to a^+} \int_{a'}^b f.$$

In particular, when f is improperly integrable, it implies

$$\int_a^b f = \lim_{n \to \infty} \int_{a_n}^b f \ ,$$

for any sequence $a_n \to a$.

A simple integrability criterion for the first type is the following "Cauchy criterion".

Proposition 2.18. *Let f be a function defined on $(a, b]$ which is integrable on $[a', b]$ for all $a' \in (a, b)$. Then f is improperly integrable over $(a, b]$ if and only if for every $\varepsilon > 0$, there exists a (small) $\delta > 0$ such that*

$$\left| \int_{a'}^{a''} f \right| < \varepsilon,$$

for any $a', a'' \in (a, a + \delta)$.

Proof. When f is improperly integrable over $(a, b]$, for $\varepsilon > 0$, there exists some δ such that

$$\left| \int_{a'}^b f - L \right| < \frac{\varepsilon}{2} \ , \qquad L = \int_a^b f, \qquad 0 < a' - a < \delta \ .$$

Therefore, for $a', a'' \in (a, a + \delta)$, we have

$$\begin{aligned} \left| \int_{a'}^{a''} f \right| &= \left| \int_{a''}^b f - \int_{a'}^b f \right| \\ &\leq \left| \int_{a''}^b f - L \right| + \left| L - \int_{a'}^b f \right| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \ . \end{aligned}$$

Conversely, for $\varepsilon = 1$, there is some δ_1 such that

$$\left| \int_{a'}^b f - \int_{a''}^b f \right| < 1 \ , \quad a', a'' \in (a, a + \delta_1) \ .$$

It implies

$$\left|\int_{a'}^{b} f\right| \le 1 + \left|\int_{a+\delta_1/2}^{b} f\right| ,$$

so the set

$$\left\{\int_{a'}^{b} f \ : \ a' \in (a, a+\delta_1)\right\}$$

is a bounded set. By Bolzano-Weierstrass Theorem, there is $\{a_n\}, a_n \to a^+$, such that

$$L \equiv \lim_{n\to\infty} \int_{a_n}^{b} f$$

exists. We claim that

$$\lim_{a'\to a} \int_{a'}^{b} f = L .$$

For, let $\varepsilon > 0$, by assumption we have some δ such that

$$\left|\int_{a'}^{b} f - \int_{a''}^{b} f\right| < \frac{\varepsilon}{2} , \quad a', a'' \in (a, a+\delta) .$$

We can also find some n_1 such that

$$\left|\int_{a_n}^{b} f - L\right| < \frac{\varepsilon}{2} , \quad \forall n \ge n_1 .$$

Then for $a_n \in (a, a+\delta)$ so we have

$$\begin{aligned} \left|\int_{a'}^{b} f - L\right| &\le \left|\int_{a'}^{b} f - \int_{a_n}^{b} f\right| + \left|\int_{a_n}^{b} f - L\right| \\ &\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon , \end{aligned}$$

for all $a' \in (a, a+\delta)$, hence f is improperly integrable over $(a, b]$. □

Example 2.7. Let f be a continuous function on $(0, 1]$ satisfying the estimate $|f(x)| \le Cx^p,\ p > -1$. We claim that its improper integral over $(0, 1]$ exists. For, for $\delta < \delta'$ small,

$$\left|\int_{\delta}^{\delta'} f\right| \le C\left|\int_{\delta}^{\delta'} x^p\right| \le \frac{C\delta'^{p+1}}{p+1}.$$

It is clear that for any $\varepsilon > 0$ we can find δ and δ' such that the right hand side of this estimate is less than ε. Hence the improper integral exists by Proposition 2.18.

Example 2.8.. Evaluate $\displaystyle\int_0^1 \frac{1}{\sqrt{x}+\sqrt[3]{x}}dx$. By the previous example, this improper integral exists. Let $x = \varphi(t) = t^6$, $\forall t \in [\delta, 1]$. Then, as $\delta \to 0$,

$$\begin{aligned}\int_{\delta^6}^1 \frac{1}{\sqrt{x}+\sqrt[3]{x}}\,dx &= \int_\delta^1 \frac{6t^5}{t^3+t^2}dt \\ &= 6\int_\delta^1 \frac{t^3}{t+1}dt \\ &\to 2t^3 - 3t^2 + 6t - 6\log|1+t|\Big|_0^1 .\end{aligned}$$

Therefore,

$$\int_0^1 \frac{1}{\sqrt{x}+\sqrt[3]{x}}\,dx = 5 - 6\log 2 .$$

The second type: f on $[a,\infty)$ belongs to $R[a,b]$, $\forall b > a$. We call $f \in R[a,\infty)$ if

$$\lim_{b\to\infty}\int_a^b f \quad \text{exists.}$$

In this case, we define

$$\int_a^\infty f = \lim_{b\to\infty}\int_a^b f.$$

When this improper integral exists,

$$\int_a^\infty f = \lim_{n\to\infty}\int_a^{b_n} f ,$$

for any sequence $b_n \to \infty$.

The Cauchy Criterion for the integrability of the second type improper integral is:

Proposition 2.19. *Let f be a function defined on $[a,\infty)$ which is integrable on $[a,b]$ for all $b > a$. The improper integral f over $[a,\infty)$ exists if and only if, for any $\varepsilon > 0$, there exists a (large) number $b_0 > a$ such that*

$$\left|\int_b^{b'} f\right| < \varepsilon,$$

for all $b', b \geq b_0$.

The proof is parallel to that of the first type and is left as an exercise.