Notes 2

RIEMANN INTEGRATION

2.1 Integrability Criterion

Let f be a function defined on a bounded, closed interval $[a, b]$. We want to consider the Riemann integral of f on $[a, b]$. We will see that this is not always possible; those for which it is possible are called (Riemann) integrable functions on $[a, b]$.

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 $||P|| = \max_j (x_j - x_{j-1})$. 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 $y = f(x)$ (assuming f is non-negative). We call f *Riemann integrable* on $[a, b]$ if there exists $L \in \mathbb{R}$ so 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 $\dot{P}$ on P. 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 $\int_a^b f$. We use $\mathcal{R}[a, b]$ to denote the set of all Riemann integrable functions on $[a, b]$.

It can be shown that any Riemann integrable functions on a closed and bounded interval $[a, b]$ are bounded functions; see textbook for a proof. Henceforth we will work only with bounded functions.

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 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.

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 will show 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

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

provided we choose $\delta < \varepsilon/2n(M + 1)$.

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, let f be a bounded function 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 $m_j = \inf\{f(x) : x \in [x_{j-1}, x_j]\}$. 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. We have

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. Let $[x_{j-1}, x_j]$ be a subinterval of P_1. It can be decomposed into the union of subinterval of P_2 : $[y_{k-1}, y_k] \cup \cdots \cup [y_{l-1}, y_l]$ where $y_{k-1} = x_{j-1}$ and $y_l = x_j$ Then

$$M_j \geq M_k^{'}, ..., M_l^{'},$$

where $M_j = \sup f$ over $[x_{j-1}, x_j]$ and $M_k' = \sup f$ over $[y_{k-1}, y_k]$, etc. From this we conclude that

$$\begin{aligned}\overline{S}(f, P_1) &= \sum M_j \Delta x_j \\ &\geq \sum M_j' \Delta y_j \\ &= \overline{S}(f, P_2).\end{aligned}$$

So (2.1) holds. Similarly, one can prove (2.2). □

The following properties are now clear.

Proposition 2.2. *For any 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 any partition P,*

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

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

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

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

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

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*

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

and

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

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

This proposition implies 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. Alternatively we state

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

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

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. Note that the number of subintervals in P which contain some partition points of Q in its interior must be less than or equal to m. Denote the indices of the collection of these subintervals in P by J. We have

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

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

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

i.e.,

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

Now, we choose

$$\delta < \frac{\varepsilon}{1 + 4Mm},$$

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

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

Similarly, one can prove the second inequality. □

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

Theorem 2.5 (Integrability Criterion I). *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 tagging of the same partition P. 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,

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

Note that the first inequality comes from the definition of the upper/lower Riemann integrals. Since $\varepsilon > 0$ is arbitrary, $\overline{S}(f) = \underline{S}(f)$.

Conversely, suppose $\overline{S}(f) = \underline{S}(f) = L$ for some real number L. Then by Proposition 2.4, we know that for $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\begin{cases} \overline{S}(f, P) - \overline{S}(f) < \varepsilon, \\ \underline{S}(f) - \underline{S}(f, P) < \varepsilon \end{cases}$$

for all partitions P, $\|P\| < \delta$. Then

$$S(f, \dot{P}) - L = S(f, \dot{P}) - \overline{S}(f) \le \overline{S}(f, P) - \overline{S}(f) < \varepsilon,$$

for any tags on P, and similarly,

$$L - S(f, \dot{P}) = \underline{S}(f) - S(f, \dot{P}) \le \underline{S}(f) - \underline{S}(f, P) < \varepsilon.$$

Combining these two inequalities yields

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

for all P, $\|P\| < \delta$, so f is integrable, with $\int_a^b f = L$. □

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

Theorem 2.6. *For any integrable f, $\int_a^b f$ 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 in order to 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

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

By tricky choice of tag points one may evaluate the integrals of all monomials.

Next we formulate our second criterion. Essentially nothing new, but the new formulation is useful in many occasions.

Theorem 2.7. ***(Integrability Criterion II)*** *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*

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

Proof. Let f be Riemann integrable on $[a, b]$, and $\varepsilon > 0$ be given. Then by definition of $\overline{S}(f)$ and $\underline{S}(f)$, there exists partitions Q, R of $[a, b]$ such that

$$\overline{S}(f) + \frac{\varepsilon}{2} > \overline{S}(f, Q), \quad \underline{S}(f) - \frac{\varepsilon}{2} < \underline{S}(f, R).$$

Now by Theorem 2.5, since f is Riemann integrable, we have $\overline{S}(f) = \underline{S}(f)$. Hence the above implies

$$\overline{S}(f, Q) - \underline{S}(f, R) < \varepsilon.$$

Let P be the partition by putting together Q and R. Then by Proposition 2.1,

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

So

$$\overline{S}(f, P) - \underline{S}(f, P) \leq \overline{S}(f, Q) - \underline{S}(f, R) < \varepsilon$$

as desired.

Next, suppose f is bounded on $[a, b]$, and for any $\varepsilon > 0$, there exists a partition P of $[a, b]$ such that

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

Then

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

and since this is true for any ε, we see that

$$\overline{S}(f) = \underline{S}(f).$$

Hence f is Riemann integrable by Theorem 2.5. □

In concluding 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. See Exercises no.16, 17 in 7.1 of Text. 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

Using either one of the integrability criteria above, 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 using the criterion it looks 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 is integrable on every $[c, d] \subset [a, b]$.

Proof. (a). We use the definition to prove (a). The key is the following simple formula for Riemann sum:

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

(verify for yourself!) As f and g are integrable, for any $\varepsilon > 0$, there exists δ such that

$$\left| S(f, \dot{P}) - \int_a^b f \right|, \left| S(g, \dot{P}) - \int_a^b g \right| < \varepsilon,$$

for $\|P\| < \delta$. Using the previous formula for Riemann sums, we then have

$$\left| S(\alpha f + \beta g, \dot{P}) - \alpha \int_a^b f - \beta \int_a^b g \right| \leq (|\alpha| + |\beta|)\varepsilon,$$

so the conclusion follows.
(b). Suppose f, g are Riemann integrable on $[a, b]$. Then they are bounded functions, so there exists $M_1, M_2 > 0$ such that $|f(x)| \leq M_1$ and $|g(x)| \leq M_2$ for all $x \in [a, b]$. It follows that fg is also a bounded function (why?). Now if P is a partition of $[a, b]$, we compute $\overline{S}(fg, P) - \underline{S}(fg, P)$. Let P be the partition given by $a = x_0 < x_1 < \cdots < x_n = b$. Observe that for any $x, y \in [x_{i-1}, x_i]$, we have

$$\begin{aligned} |f(x)g(x) - f(y)g(y)| &\leq |f(x)||g(x) - g(y)| + |g(y)||f(x) - f(y)| \\ &\leq M_1|g(x) - g(y)| + M_2|f(x) - f(y)|. \end{aligned}$$

Hence

$$\sup_{x\in[x_{i-1},x_i]}[f(x)g(x)] - \inf_{x\in[x_{i-1},x_i]}[f(x)g(x)]$$
$$\le M_1\left(\sup_{x\in[x_{i-1},x_i]} g(x) - \inf_{x\in[x_{i-1},x_i]} g(x)\right) + M_2\left(\sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x)\right),$$

which implies that

$$\overline{S}(fg,P) - \underline{S}(fg,P) \le M_1(\overline{S}(g,P) - \underline{S}(g,P)) + M_2(\overline{S}(f,P) - \underline{S}(f,P))$$

Now given $\varepsilon > 0$, there exists partitions P_1, P_2 of $[a,b]$ such that

$$\overline{S}(f,P_1) - \underline{S}(f,P_2) < \frac{\varepsilon}{M_1+M_2}, \quad \overline{S}(g,P_2) - \underline{S}(g,P_2) < \frac{\varepsilon}{M_1+M_2}.$$

Let P be the partition of $[a,b]$ formed by putting together P_1 and P_2. Then we also have

$$\overline{S}(f,P) - \underline{S}(f,P) < \frac{\varepsilon}{M_1+M_2}, \quad \overline{S}(g,P) - \underline{S}(g,P) < \frac{\varepsilon}{M_1+M_2},$$

so the above estimate of $\overline{S}(fg,P) - \underline{S}(fg,P)$ gives

$$\overline{S}(fg,P) - \underline{S}(fg,P) < M_1\frac{\varepsilon}{M_1+M_2} + M_2\frac{\varepsilon}{M_1+M_2} = \varepsilon.$$

This proves that fg is Riemann integrable by Integrability Criterion II.

We leave (c), (d) and (e) to you. □

It is now easy to prove:

Theorem 2.9. *Let g be obtained from an integrable function f on $[a,b]$ by modifying it at finitely many points. Then g is also integrable and the integral of g is equal to the integral of f.*

Proof. Let $\{w_1,\cdots,w_N\}$ be the points at which f and g are different. The function $h = g - f$ is zero except at these points. From Example 3, we know that h is integrable and its integral is 0. By Theorem 8 (a) we conclude that $g = f + h$ is integrable and g and f have the same integral. □

Theorem 2.8 in particular shows that the collection of all Riemann integrable functions form a vector space which is closed under multiplication. Since 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 if $f \in \mathcal{R}[a,b]$ and $g \in C[c,d]$ and $f[a,b] \subset [c,d]$, then $g \circ f \in \mathcal{R}[a,b]$, see exercise.

We will now show that there are indeed plenty of Riemann integrable functions: e.g. every continuous function is integrable, so the inclusion

$$C[a,b] \subset \mathcal{R}[a,b]$$

holds; also every (bounded) monotone function on $[a,b]$ is Riemann integrable on $[a,b]$.

Theorem 2.10. *Every continuous function on a closed and bounded interval $[a,b]$ is Riemann integrable on $[a,b]$.*

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

$$|f(x) - f(y)| < \frac{\varepsilon}{2(b-a)}, \ \forall x, y \in [a,b], |x-y| < \delta.$$

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

$$\begin{aligned} \overline{S}(f,P) &= \sum f(z_j)\Delta x_j, \\ \underline{S}(f,P) &= \sum f(w_j)\Delta x_j, \end{aligned}$$

where $f(z_j) = M_j$ and $f(w_j) = m_j$ by continuity. Therefore,

$$\begin{aligned} 0 \le \overline{S}(f,P) - \underline{S}(f,P) &= \sum (f(z_j) - f(w_j))\Delta x_j, \\ &\le \frac{\varepsilon(b-a)}{2(b-a)} \\ &< \varepsilon. \end{aligned}$$

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

□

In fact, one can allow for finitely many discontinuities of a function defined on $[a,b]$, and still retain Riemann integrability:

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*

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

In particular, every bounded function on a closed and bounded interval $[a, b]$, with (at most) finitely many discontinuities, is Riemann integrable on $[a, b]$.

Here F may not be equal to f_j at some a_j.

Proof. Clearly it suffices to assume $n = 2$, that is, there is some c, $a < c < b$, and two functions f_1 and f_2, integrable on $[a, c]$ and $[c, b]$ respectively, such that $F = f_1$ on (a, c) and $F = f_2$ on (c, b). By Integrability Criterion II, for any $\varepsilon > 0$, we can find partitions P_1 and P_2 of $[a, c]$ and $[c, b]$ respectively such that

$$\overline{S}(f_1, P_1) - \underline{S}(f_1, P_1) < \frac{\varepsilon}{3}, \quad \overline{S}(f_2, P_2) - \underline{S}(f_2, P_2) < \frac{\varepsilon}{3}.$$

We will write P_1 as $a = x_0 < x_1 < \cdots < x_m = c$, and write P_2 as $c = x_m < x_{m+1} < \cdots < x_n = b$. We may also assume that there exists some number $M > 0$ such that $|F|$, $|f_1|$ and $|f_2|$ are all bounded by M, and that $\|P_1\|, \|P_2\| < \frac{\varepsilon}{48M}$. Let P be the partition of $[a, b]$ formed by putting together P_1 and P_2. Then

$$\begin{aligned}
&\overline{S}(F, P) \\
\le&\overline{S}(f_1, P_1) + \overline{S}(f_2, P_2) \\
&\quad + \sup_{x,y\in[x_0,x_1]} |F(x) - f_1(y)|(x_1 - x_0) + \sup_{x,y\in[x_{m-1},x_m]} |F(x) - f_1(y)|(x_{m-1} - x_m) \\
&\quad + \sup_{x,y\in[x_m,x_{m+1}]} |F(x) - f_2(y)|(x_{m+1} - x_m) + \sup_{x,y\in[x_{n-1},x_n]} |F(x) - f_2(y)|(x_{n-1} - x_n) \\
\le&\overline{S}(f_1, P_1) + \overline{S}(f_2, P_2) + 4(2M)\frac{\varepsilon}{48M} \\
=&\overline{S}(f_1, P_1) + \overline{S}(f_2, P_2) + \frac{\varepsilon}{6}.
\end{aligned}$$

Similarly

$$\underline{S}(F, P) \ge \underline{S}(f_1, P_1) + \underline{S}(f_2, P_2) - \frac{\varepsilon}{6}.$$

Hence

$$\begin{aligned}
&\overline{S}(F, P) - \underline{S}(F, P) \\
\le&[\overline{S}(f_1, P_1) - \underline{S}(f_1, P_1)] + [\overline{S}(f_2, P_2) - \underline{S}(f_2, P_2)] + \frac{\varepsilon}{3} \\
<&\frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

By Integrability Criterion II, 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,

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

□

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 enter 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 closed and bounded interval $[a, b]$ is Riemann integrable on $[a, b]$.*

Proof. Assume f is increasing on $[a, b]$. The case where f is decreasing follows by replacing f by $-f$.

Let P be the partition which divides $[a, b]$ into n subintervals of equal lengths. In other words, P is given by $a = x_0 < x_1 < \cdots < x_n = b$ where $x_i - x_{i-1} = \frac{b-a}{n}$ for all i. Then

$$\sup_{x\in[x_{j-1},x_j]} f(x) - \inf_{x\in[x_{j-1},x_j]} f(x) = f(x_j) - f(x_{j-1}).$$

Hence

$$\begin{aligned}\overline{S}(f,P)-\underline{S}(f,P) &= \sum_{j=1}^{n}\left(\sup_{x\in[x_{j-1},x_j]} f(x)-\inf_{x\in[x_{j-1},x_j]} f(x)\right)(x_j-x_{j-1})\\ &= \frac{b-a}{n}\sum_{j=1}^{n}(f(x_j)-f(x_{j-1}))\\ &= \frac{(b-a)(f(b)-f(a))}{n},\end{aligned}$$

and we can make this smaller than any $\varepsilon > 0$ if we choose n so large that $(b-a)(f(b)-f(a))/\varepsilon < n$. By the Integrability Criterion II, f is Riemann integrable on $[a,b]$. □

Monotone functions could have countably many jumps. For instance, let all rational numbers in $(0,1)$ be written in a sequence $\{x_j\}$ and define $\varphi(x) = \sum_{\text{all } j, x_j<x} 2^{-j}$. You can 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

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

where (p,q) denotes the greatest common divisor of p and q. We set $h(0)=1$.

We show that $h\in\mathcal{R}[0,1]$. 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.

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 $\delta\leq 1/4N_0\varepsilon$. Then, for any partition

P with length less than δ, we have

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

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 $\mathcal{D} = \frac{d}{dx}$ is a linear transformation from $D(a, b)$ to $F(a, b)$. On the other hand, for any $f \in \mathcal{R}[a, b]$, the *indefinite integral* F of f, which is defined by

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

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 $\mathcal{R}[a, b]$ to $F(a, b)$. In an ideal setting, one would like to see if there exist $\mathcal{J} : \mathcal{R}[a, b] \to D(a, b)$ and $\mathcal{D} : D(a, b) \to \mathcal{R}[a, b]$ such that $\mathcal{J}\mathcal{D}f = f$, $\forall f \in D(a, b)$, and $\mathcal{D}\mathcal{J}f = f$, $\forall f \in \mathcal{R}[a, b]$. 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 $x \ge 0$, and $= 0$ for $x < 0$, which is in $\mathcal{R}[-1, 1]$, 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 $\mathcal{R}[a, b]$.

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. *Let F be differentiable on $[a,b]$ and $F' \in \mathcal{R}[a,b]$. 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 $f \in \mathcal{R}[a,b]$, 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 the 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

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

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 $\int x^k = x^{k+1}/(k+1)$ 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 consideration of whether $\mathcal{D}\mathcal{J}f = f$. The key is a simple observation: if $f \in \mathcal{R}[a,b]$ with $|f(x)| \le M$ for all $x \in [a,b]$, then

$$\Big|\int_a^b f\Big| \le M(b-a).$$

This inequality holds because any upper and lower sum of f satisfies

$$-M(b-a) \le \underline{S}(f,P) \le \overline{S}(f,P) \le M(b-a);$$

upon taking supremum and infimum over all partitions P, we see that $\int_a^b f = \underline{S}(f) = \overline{S}f \in [-M(b-a), M(b-a)]$ as claimed.

Theorem 2.14. *Let $f \in \mathcal{R}[a,b]$ and $F(x) = \int_a^x f$ be the indefinite integral of f. If f is continuous at some $c \in [a,b]$, then F is differentiable at c, with $F'(c) = f(c)$.*

Proof. 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

$$|f(x) - f(c)| < \varepsilon, \qquad \forall x \in (c-\delta, c+\delta).$$

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

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

(Here we used the observation immediately before the theorem.) 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

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

Then,

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

So, $F'(0)$ does not exist.

As an application of the fundamental theorem, we prove the formula on change of variables.

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 $\varphi' \in \mathcal{R}[\alpha, \beta]$. Then,*

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

Proof. Define

$$F(t) = \int_a^{\varphi(t)} f(x)dx.$$

Regarding F as the composite of two functions: $t \mapsto \varphi(t) = x$ and $x \mapsto \int_a^x f$. By assumption φ is differentiable and by the second fundamental theorem the indefinite integral of f is also differentiable. We use the differential rule for two composite functions to get

$$\frac{dF}{dt}(t) = f(\varphi(t))\varphi'(t).$$

On the other hand, $f(\varphi(t)) \in C[\alpha, \beta]$ and $\varphi' \in \mathcal{R}[\alpha, \beta]$ imply $f(\varphi(t))\varphi'(t)$ is integrable on $[\alpha, \beta]$. By the first fundamental theorem of calculus,

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

i.e.

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt.$$

□

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

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 $[0, 5\pi/2]$. Then

$$\int_0^1 \sqrt{1-x^2}dx = \int_0^{\frac{5\pi}{2}} \sqrt{1-\sin^2 t}\cos t dt = \int_0^{\frac{5\pi}{2}} |\cos t| \cos t dt = \frac{\pi}{4},$$

which yields the same result. Be careful of the cancelation of the integral over $[0, 2\pi]$.

2.4 Improper Integrals

Very often we face the situation where f is unbounded on $[a, b]$ 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 allow these new situations. These generalized integrals are called improper integrals. They are rather common in applications.

We briefly discuss two typical cases: Consider when f 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). We call f has an improper integral on $[a, b]$ if $f \in \mathcal{R}[a', b]$, $\forall a' > a$, and

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

We let

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

Next, let f be in $\mathcal{R}[a, b]$, $\forall b > a$. Then we call $f \in \mathcal{R}[a, \infty)$ if

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

In this case, we define

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

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

Proposition 2.16. *Let f be a function defined on $(a, b]$ which is integrable on $[a', b]$ for all $a' \in (a, b)$. Suppose that for any $\varepsilon > 0$, there exist small $\delta_0 > 0$ such that*

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

for any $\delta, \delta' \in (0, \delta_0)$. Then the improper integral $\int_a^b f$ exists.

Similarly, for the second case we have

Proposition 2.17. *Let f be a function defined on $[a, \infty)$ which is integrable on $[a, b]$ for all $b > a$. Suppose that for any $\varepsilon > 0$, there exists large number $b_0 > a$ such that*

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

for all $b', b \geq b_0$. Then the improper integral $\int_a^\infty f$ exists.

Both proofs are immediate consequence of the fact that any Cauchy sequence converges. Fill in the proofs if you like.

Example 2.7. Let f be a continuous function on $(0, 1]$ satisfying the estimate $|f(x)| \leq Cx^p, p > -1$. We claim that the improper integral $\int_0^1 f$ exists. For, for $\delta < \delta'$ small,

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

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 2.16.

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 $\delta \to 0$,

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

2.5 Integration by Parts and Applications

In this section we discuss the formula on integration by parts and give some of its applications including the important Taylor expansion theorem and the Euler-Maclaurin formula.

Theorem 2.18. ***(Integration by Parts)*** *Let F and G be differentiable on $[a, b]$ and $f = F'$, $g = G'$ be in $\mathcal{R}[a, b]$. 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 Proposition 8, 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 theorem with integral remainder.

Theorem 2.19. ***(Taylor's Theorem with Integral Remainder)*** *Suppose $f', f'', \ldots, f^{(n+1)}$ exist on (a, b) and $f^{(n+1)} \in \mathcal{R}[\alpha, \beta]$ 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 proposition with Theorem 6.4.1 in the textbook. In Theorem 6.4.1 the regularity requirement on f is weaker: $f^{(n+1)} \in \mathcal{R}[\alpha, \beta]$ is not necessary 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.

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. □

Integration by parts is one of the most useful methods in integration. To illustrate its use we shall establish the following formulas obtained by Wallis not long before Newton invented calculus. Due to human's special feeling to π, any formula for this transcendental number catches attention. Wallis' formulas are one of the earliest ones. They are not good in computing π for slow convergence. We shall encounter more formulas in other courses later.

Theorem 2.20. ***(Wallis' Formulas)*** *The following formulas hold true.*

(a) $\lim_{n\to\infty}\left(\frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots\frac{2n}{2n-1}\cdot\frac{2n}{2n+1}\right)=\frac{\pi}{2}.$

(b) $\lim_{n\to\infty}\frac{(n!)^2 2^{2n}}{(2n)!\sqrt{n}}=\sqrt{\pi}.$

Note that (b) relates $(2n)!$ to $n!$, a non-trivial fact to be used later.

Proof. For each $n\in\mathbb{N}$, let

$$I_n=\int_0^{\frac{\pi}{2}}\sin^n x dx,$$

and

$$a_n=\frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdots\frac{2n}{2n-1}\cdot\frac{2n}{2n+1}.$$

We note $I_0=\pi/2$ and $I_1=1$. Now, $n\geq 2$, we use integration by parts to get

$$\begin{aligned}I_n&=-\cos x\sin^{n-1}x\Big|_0^{\frac{\pi}{2}}-\int_0^{\frac{\pi}{2}}(-\cos x)(n-1)\sin^{n-2}x\cos x dx\\&=(n-1)\int_0^{\frac{\pi}{2}}\cos^2 x\sin^{n-2}x dx\\&=(n-1)(I_{n-2}-I_n).\end{aligned}$$

It follows that

$$I_n=\frac{n-1}{n}I_{n-2}.$$

From this recursive formula, we see that, $\forall n\in\mathbb{N}$,

$$I_{2n}=\frac{2n-1}{2n}I_{2n-2}=\cdots=\frac{1\cdot 3\cdots(2n-3)\cdot(2n-1)}{2\cdot 4\cdots(2n-2)\cdot(2n)}I_0,$$

and,

$$I_{2n+1}=\frac{2n}{2n+1}I_{2n-1}=\cdots=\frac{2\cdot 4\cdots(2n-2)\cdot(2n)}{3\cdot 5\cdots(2n-1)\cdot(2n+1)}I_1.$$

Therefore,

$$\frac{I_{2n}}{I_{2n+1}}=\frac{1\cdot 3\cdot 3\cdot 5\cdot 5\cdots(2n-1)\cdot(2n-1)\cdot(2n+1)}{2\cdot 2\cdot 4\cdot 4\cdot 6\cdots(2n-2)(2n)(2n)}\cdot\frac{\pi}{2},$$

or we set

$$a_n=\frac{I_{2n+1}}{I_{2n}}\cdot\frac{\pi}{2}.$$

Noting that (a) will hold if we can show that $I_{2n+1}/I_{2n} \to 1$ as $n \to \infty$, but this is easy to see from

$$\sin^{2n} x \geq \sin^{2n+1} x \geq \sin^{2n+2} x, \quad \forall x \in [0, \frac{\pi}{2}],$$

as it implies $I_{2n} \geq I_{2n+1} \geq I_{2n+2}$ and so

$$1 \geq \frac{I_{2n+1}}{I_{2n}} \geq \frac{I_{2n+2}}{I_{2n}} = \frac{2n+1}{2n+2} \to 1.$$

To prove (b) we simply note that a_n can be written as

$$\frac{\left[2^{2n}(n!)^2\right]^2}{((2n)!)^2(2n+1)},$$

so it follows from (a).

□

Next we present Euler-Maclaurin's formula. This beautiful formula may be regarded as an approximation formula to an integral by finite sums or expressing a finite sum by an integral plus some error terms. To prepare for it we need to introduce a special family of polynomials. We define the *Bernoulli's polynomials* $B_k, k \geq 0$, inductively by setting $B_0(x) \equiv 1$,

$$B_k'(x) = kB_{k-1}(x), \qquad \text{and} \qquad \int_0^1 B_k(x)dx = 0, \quad \forall k \geq 1.$$

The first several Bernoulli's polynomials are given by

$$B_0(x) \equiv 1, \quad B_1(x) = x - \frac{1}{2}, \quad B_2(x) = x^2 - x + \frac{1}{6}, \quad B_3(x) = x^3 - \frac{3}{2}x^2 - \frac{1}{2}x,$$

and

$$B_4(x) = x^4 - 2x^3 + x^2 - \frac{1}{30}.$$

Notice that these polynomials satisfy

$$B_k^{(j)}(x) \equiv k(k-1)\cdots(k-j+1)B_{k-j}, \quad \forall j \leq k. \tag{2.4}$$

In particular, $B_k^{(k)} = k!$. We also have

$$B_k(1) = B_k(0), \quad \forall k \geq 2, \tag{2.5}$$

but $B_1(1) = 1/2 = -B_1(0)$. We call the numbers $b_k = B_k(0)$ the *Bernoulli's* *numbers*. We have

$$b_0 = 1, \quad b_1 = -\frac{1}{2}, \quad b_2 = \frac{1}{6}, \quad b_4 = -\frac{1}{30}, \quad b_6 = \frac{1}{42}, \quad b_8 = -\frac{1}{30},$$

and so on. One can show that $b_{2k+1} = 0$ for $k \geq 1$, but we will not need this fact.

In the following we also set $P_k(x) = B_k(x - [x])$ where $[x]$ is the integral part of x. In other words, for each $k \geq 0$ $P_k = B_k$ on $[0, 1)$ and P_k satisfies $P_k(x + 1) = P_k(x)$ for all $x \in \mathbb{R}$. P_k is the periodic 1 extension of the restriction of B_k on $[0, 1)$. In view of (2.5), each P_k is continuous on $\mathbb{R}$ for $k \geq 2$ but P_1 has a jump at every $x \in \mathbb{Z}$.

Theorem 2.21. ***(Euler-Maclaurin Formula)*** *Let f be k-th continuously differentiable on the interval $[0, n]$. We have*

$$\sum_{j=0}^{n-1} f(j) = \int_0^n f(x)dx + \sum_{j=1}^{k} \frac{b_j}{j!}\big(f^{(j-1)}(n) - f^{(j-1)}(0)\big) - \frac{1}{k!}\int_0^n P_k(-x)f^{(k)}(x)dx. \tag{2.6}$$

In application, usually the second integral term in (2.6) is chosen to be small, so it should be regarded as an error term. This formula was obtained by Euler and Colin Maclaurin independently around 1735. The integration by parts proof presented here is taken from V. Lampret, "The Euler-Maclaurin and Taylor formulas: Twin, elementary derivations," in vol. 74, p.109-122, Mathematics Magazine 2001. You may google for more information on this formula, especially how it is used in numerical integration.

Proof. By using integration by parts, one shows that for any smooth function g on $[0, 1]$,

$$\int_0^1 f(x)g^{(k)}(x)dx = \Big[\sum_{j=0}^{k-1}(-1)^j f^{(j)}(x)g^{(k-1-j)}(x)\Big]_0^1 + (-1)^k \int_0^1 f^{(k)}(x)g(x)dx.$$

By replacing $g(x)$ by $h(1 - x)$ and using $g^{(j)}(x) = (-1)^j h^{(j)}(1 - x)$ this formula becomes

$$\begin{aligned}\int_0^1 f(x)h^{(k)}(1 - x)dx \;&=\; \sum_{j=0}^{k-1}\big(f^{(j)}(0)h^{(k-1-j)}(1) - f^{(j)}(1)h^{(k-1-j)}(0)\big) \\ &\quad + \int_0^1 f^{(k)}(x)h(1 - x)dx.\end{aligned}$$

Now, by taking h to be B_k in this formula and using $B_k^{(k)} \equiv k!$, we have

$$\begin{aligned}
k!\int_0^1 f(x)dx &= \sum_{j=1}^{k} \big(f^{(j-1)}(0)h^{(k-j)}(1) - f^{(j-1)}(1)h^{(k-j)}(0)\big) \\
&\quad + \int_0^1 f^{(k)}(x)h(1-x)dx \\
&= \sum_{j=2}^{k} \big(f^{(j-1)}(0)h^{(k-j)}(1) - f^{(j-1)}(1)h^{(k-j)}(0)\big) + f(0)h^{(k-1)}(1) \\
&\quad - f(1)h^{(k-1)}(0) + \int_0^1 f^{(k)}(x)h(1-x)dx \\
&= \sum_{j=1}^{k} h^{(k-j)}(1)\big(f^{(j-1)}(0) - f^{(j-1)}(1)\big) + k!f(1) \\
&\quad + \int_0^1 f^{(k)}(x)h(1-x)dx,
\end{aligned}$$

after using (2.4) and (2.5). Dividing both sides by $k!$ and moving terms, we get

$$f(1) = \int_0^1 f(x)dx + \sum_{j=1}^{k} \frac{1}{j!}B_j(1)\big(f^{(j-1)}(1) - f^{(j-1)}(0)\big) - \frac{1}{k!}\int_0^1 f^{(k)}(x)P_k(1-x)dx.$$

This formula is valid for any function f defined on $[0,1]$. Now, for a function f defined on $[0,n]$, the functions $f_i(x) = f(x+i)$ is defined on $[0,1]$ for $i = 0,1,\cdots,n-1$. Applying the formula to each f_i and then summing up, we arrive at

$$\begin{aligned}
\sum_{i=1}^{n} f(i) &= \sum_{i=0}^{n-1} f(1+i) \\
&= \sum_{i=0}^{n-1}\int_0^1 f(x+i)dx + \sum_{i=0}^{n-1}\sum_{j=1}^{k} \frac{B_j(1)}{j!}\big(f^{(j-1)}(1+i) - f^{(j-1)}(i)\big) \\
&\quad - \frac{1}{k!}\sum_{i=0}^{n-1}\int_0^1 P_k(1-x)f^{(k)}(x+i)dx \\
&= \int_0^n f(x)dx + \sum_{j=1}^{k} \frac{B_j(1)}{j!}\big(f^{(j-1)}(n) - f^{(j-1)}(0)\big) \\
&\quad - \frac{1}{k!}\int_0^n P_k(-x)f^{(k)}(x)dx.
\end{aligned}$$

Using (2.5) to replace $B_j(1)$ by $B_j(0)$ for $j \geq 2$ and noting that $B_1(0) = -1/2 =$

$B_1(1)$ we obtain (2.6).

□

Euler-Maclaurin formula has many applications. We present two of them.

First, consider the sum of positive powers. The following summation formula is well-known

$$1 + 2 + \cdots + n = \frac{1}{2}n(n+2).$$

One can establish this formula by either using induction or the trick by writing the sum as $n + (n-1) + \cdots + 2 + 1$ and summing up to get twice of the sum. Let us denote

$$S_k(n) = 1^k + 2^k + \cdots + n^k.$$

A general formula for $S_k(n)$ can be obtained by a reduction argument. For instance, consider $S_3(n+1) - S_3(n)$. On one hand, it is equal to $(n+1)^2$ and, on the other hand, it is given by $\sum_{k=0}^{n}(k+1)^3 - \sum_{k=0}^{n} k^3 = 3S_2(n) + 3S_1(n) + n + 1$. Hence $S_2(n)$ can be expressed in terms of $S_1(n)$. Likewise, using $S_{k+1}(n+1) - S_{k+1}(n) = (n+1)^{k+1}$ one can express $S_k(n)$ in terms of $S_{k-1}(n)$, $S_{k-2}(n)$, etc. The first several $S_k(n)$ are given by

$$1 + 2 + 3 + \cdots + n = \frac{1}{2}n(n+1),$$

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1),$$

$$1^3 + 2^3 + 3^3 + \cdots + n^3 = \frac{1}{30}n(n+1)(2n+1)(3n^2+3n-1),$$

etc.

Faulhaber found the formulas for $S_k(n)$ up to $k = 17$ in 1631 and the general formula was found by Jacob Bernoulli. Indeed, by choosing $f(x) = x^k$ and noticing $f^{(k+1)} \equiv 0$, we immediately deduce from Euler-Maclaurin formula (taking $n = n+1$ and $k = k+1$ in (2.6) so the error term vanishes) the Faulhaber's formula

$$S_k(n) = \frac{(n+1)^{k+1}}{k+1} + \sum_{j=1}^{k} \frac{k(k-1)\cdots(k-j+2)}{j!} b_j (n+1)^{k-j+1},$$

or

$$S_k(n) = \frac{1}{k+1} \sum_{j=0}^{k} C_j^{k+1} b_j (n+1)^{k-j+1},$$

where $C_j^{k+1} = (k+1)!/j!(k+1-j)!$.

Next we consider Stirling's formula. This important formula describes an asymptotic form of the factorial function.

Theorem 2.22. ***(Stirling's Formula)*** *We have*

$$n! = C(n)\sqrt{n}\Big(\frac{n}{e}\Big)^n, \qquad \text{where } \lim_{n\to\infty} C(n) = \sqrt{2\pi}.$$

Proof. We choose $f(x) = \log(1+x)$ in (2.6)($k=1$) to obtain

$$\sum_{j=0}^{n-1} \log(1+j) = (1+n)\big(\log(1+n) - 1\big) + 1 - \frac{1}{2}\log(1+n) - \int_0^n \frac{P_1(-x)}{1+x}dx,$$

or

$$\log n! - (n+\frac{1}{2})\log(1+n) + (n+1) = S(n), \tag{2.7}$$

where

$$S(n) = 1 - \int_0^n \frac{P_1(-x)}{1+x}dx.$$

We claim $\lim_{n\to\infty} S(n)$ exists. For, letting

$$a_k = \int_k^{k+1} \frac{P_1(-x)}{1+x}dx,$$

we can express

$$\int_0^n \frac{P_1(-x)}{1+x}dx$$

as $\sum_{k=0}^{n-1} a_k$, and the existence of the limit is equivalent to the fact that $\{\sum_{k=0}^{n} a_k\}$ is a Cauchy sequence. We estimate a_k as follows,

$$\begin{aligned} a_k &= \int_0^1 \frac{P_1(-x-k)}{1+k+x}dx \\ &= \int_0^1 \frac{P_1(-x)}{1+k+x}dx, \qquad (P_1 \text{ is periodic } 1) \\ &= \Big[\frac{-P_2(-x)}{2(x+k+1)}\Big]_0^1 - \int_0^1 \frac{P_2(-x)}{2(x+k+1)^2}dx \\ &= \frac{P_2(0)}{2(k+1)(k+2)} - \int_0^1 \frac{P_2(-x)}{2(x+k+1)^2}dx \qquad (\text{Use } P_2(-1) = P_2(0)). \end{aligned}$$

It follows that

$$|a_k| \le \frac{2M}{(k+1)^2}, \qquad M = \sup_{x\in[0,1]} |P_2(x)|.$$

As $\sum_k 1/k^2 < \infty$, for a given $\varepsilon > 0$, there exists some n_0 such that for all

$n, m \geq n_0$, $\sum_{k=m+1}^{n} 1/k^2 < \varepsilon$. It follows that

$$|\sum_{k=0}^{n} a_k - \sum_{k=0}^{m} a_k| = |\sum_{k=m+1}^{n} a_k| < C \sum_{k=m+1}^{n} \frac{1}{k^2} < C\varepsilon,$$

hence $\{\sum_{k=0}^{n} a_k\}$ is a Cauchy sequence. We conclude that the limit of $S(n)$ exists as $n \to \infty$.

Next, we evaluate this limit. Replacing n by $2n$ in (2.7) we have

$$\log(2n)! - (2n + \frac{1}{2}) \log(2n+1) + 2n + 1 = S(2n).$$

Multiplying (2.7) by 2 and subtracting it from this one we have

$$\log \frac{((n)!)^2 (2n+1)^{2n+1/2} e}{(2n)!(n+1)^{2n+1}} = 2S(n) - S(2n).$$

By letting $n \to \infty$ and applying Wallis' formula to the left hand side we conclude that

$$\begin{aligned} \lim_{n\to\infty} S(n) &= \lim_{n\to\infty} \big(2S(n) - S(2n)\big) \\ &= \log\left(\sqrt{2\pi} e\right) \end{aligned}$$

Now Stirling's formula follows from (2.7). □

In the exercise you are asked to prove Stirling's formula by an elementary method.

We conclude with a very brief discussion on the Gamma function. It is for optional reading.

The Gamma function Γ is a function defined on $(0, \infty)$ given by the improper integral

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt.$$

In an exercise I have asked you to show that this improper integral exists. The Gamma function belongs to a family of functions called special functions. Next to the elementary functions such as polynomials, rational functions, radicals, trigonometric functions, exponential and logarithmic functions, special functions such as the Gamma functions, elliptic functions, Bessel functions, theta functions, etc appear in various contexts. Special functions have been studied since the invention of calculus and many results are known. The Gamma function, one of the earliest special functions, arises from the so-called interpolation problem. As we all know, the factorial function is given by $n! = n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$. It is well-defined only for natural numbers including 0, where we set $0! = 1$. So the question is: How can we define the factorial of a positive number in a reasonable way? In 1731, the young Euler observed that the following

$$n! = \lim_{m\to\infty} \frac{m!(m+1)^n}{(n+1)(n+2)\cdots(n+m)}.$$

(Prove it.) While the left hand side only makes sense for natural numbers n, the right hand side makes sense for any positive number. Motivated by this, Euler proposed the definition

$$x! = \lim_{m\to\infty} \frac{m!(m+1)^x}{(x+1)(x+2)\cdots(x+m)}, \qquad \forall x > 0. \tag{2.8}$$

However, this definition involves the evaluation of a limit and is not convenient for applications. Euler sought different expressions of this general factorial function in subsequent years. Finally, in 1781 he obtained the Gamma function and showed that $\Gamma(x+1) = x!$. In the following we sketch how to establish this fact. First we note some elementary properties of the Gamma function.

Proposition 2.23. *The Gamma function* Γ *satisfies*

(a) $\Gamma(x+1) = x\Gamma(x), \quad \forall x > 0,$

(b) $\Gamma(n+1) = n!,$

(c) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$

In the proof of (c) you need to use the formula

$$\int_{-\infty}^{\infty} e^{-r^2}\,dr = \sqrt{\pi}.$$

Proposition 2.24. *Let* $x!$ *be given in (2.8). We have*

$$\Gamma(x+1) = x!, \quad \forall x > 0.$$

The proof of this proposition may be divided into two steps. Letting

$$P_n(x) = \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1}\,dt,$$

first we show that for each $x > 0$,

$$\lim_{n\to\infty} P_n(x) = \Gamma(x).$$

This can be achieved by using the inequalities

$$0 \le e^{-t} - \Big(1 - \frac{t}{n}\Big)^n \le \frac{t^2}{n} e^{-t}.$$

Next, by a change of variables,

$$P_n(x) = n^x \int_0^1 \big(1 - s\big)^n s^{x-1} ds,$$

and through a series of integration by parts we get $P_n(x) \to (x-1)!$ as $n \to \infty$.

Finally, as an application we compute the volume of the unit ball in $\mathbb{R}^n$. Let

$$B_r = \{x \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 < r^2\}, \qquad \text{and} \qquad S_r = \{x \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = r^2\}$$

be the ball and sphere of radius r centering at the origin respectively. We use $|B_r|$ and $|S_r|$ to denote its volume and surface area. We recall the relation $|S_r| = n|B_r|/r$.

Proposition 2.25. *For $n \ge 1$, we have*

$$|B_1| = \frac{\pi^{n/2}}{\Gamma\Big(1 + \dfrac{n}{2}\Big)}.$$

Proof. Let

$$I = \int_{\mathbb{R}^n} e^{-|x|^2} dx_1 \cdots dx_n.$$

We calculate I in two ways. First, by Fubini's theorem in Advanced Calculus II,

$$I = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} e^{-x_1^2} \cdots e^{-x_n^2} dx_1 \cdots dx_n = \Big(\int_{-\infty}^{\infty} e^{-t^2} dt\Big)^n = \big(\sqrt{\pi}\big)^n.$$

On the other hand, expressing I in the polar coordinates,

$$\begin{aligned}
I &= \int_{S_1} \int_0^\infty e^{-r^2} r^{n-1} dr d\theta \\
&= \frac{|S_1|}{2} \int_0^\infty e^{-t} t^{n/2-1} dt \\
&= \frac{n|B_1|}{2} \times \Gamma\Big(\frac{n}{2}\Big) \\
&= |B_1| \Gamma\Big(1 + \frac{n}{2}\Big),
\end{aligned}$$

and the results follows. □

2.6 Lebesgue's Theorem

We have seen that any function with finite discontinuity is integrable. Also it is not hard to show that a function with countably many discontinuity is still integrable provided these discontinuous points converges to a single points. On the other hand, functions with too many discontinuous points are not integrable. A typical example is $f(x) = 1$ if x is rational, and $= 0$ otherwise. This function is discontinuous everywhere. In this section we prove Lebsegue fundamental theorem characterizing Riemann integrability in terms of the "size" of discontinuity set.

This section is for optional reading.

For any bounded f on $[a, b]$ and $x \in [a, b]$, its *oscillation* at x is defined by

$$\begin{aligned}\omega(f, x) &= \inf_{\delta}\{(\sup f(y) - \inf f(y)) : y \in (x-\delta, x+\delta) \cap [a, b]\} \\ &= \lim_{\delta \to 0^+}\{(\sup f(y) - \inf f(y)) : y \in (x-\delta, x+\delta) \cap [a, b]\}.\end{aligned}$$

It is clear that $\omega(f, x) = 0$ if and only if f is continuous at x. The set of discontinuity of f, D, can be written as

$$D = \bigcup_{k=1}^{\infty} O(k), \tag{2.9}$$

where $O(k) = \{x \in [a, b] : \omega(f, x) \geq 1/k\}$.

A subset E of $\mathbb{R}$ is of *measure zero* if $\forall \varepsilon > 0, \exists$ a sequence of open intervals $\{I_j\}$ such that

$$E \subseteq \bigcup_{j=1}^{\infty} I_j,$$

and

$$\sum_{j=1}^{\infty} |I_j| < \varepsilon.$$

It is not hard to show that

Proposition 2.26. *The following statements hold.*
(a) Any countable set is of measure zero.
(b) Any countable union of measure zero sets is again of measure zero.

Proof. Let $E = \{x_1, x_2, ...\}$ be a countable set. Given $\varepsilon > 0$, the intervals $I_j = (x_j - \frac{\varepsilon}{2^{j+2}}, x_j + \frac{\varepsilon}{2^{j+2}})$ satisfy

$$E \subseteq \bigcup_{j=1}^{\infty} I_j,$$

and

$$\sum_{j=1}^{\infty} |I_j| = \sum_{j=1}^{\infty} \frac{2\varepsilon}{2^{j+2}} = \frac{\varepsilon}{2} < \varepsilon,$$

so E is of measure zero. (a) is proved. (b) can be proved by a similar argument. We leave it as an exercise. □

There are uncountable sets of measure zero. The famous Cantor set is one of them. See Chapter 7 of the textbook. Now, we state the necessary and sufficient condition for Riemann integrability due to Lebesgue.

Theorem 2.27. *A bounded function f on $[a,b]$ is Riemann integrable if and only if its discontinuity set is of measure zero.*

We shall use the compactness of a closed, bounded interval in the proof of this theorem. Recall that compactness is equivalent to the following property: Let K be a compact set in $\mathbb{R}$. Suppose that $\{I_j\}$ is a sequence of open intervals satisfying $K \subseteq \bigcup_{j=1}^{\infty} I_j$. Then we can choose finitely many intervals $I_{j_1}, ..., I_{j_N}$ so that $K \subseteq I_{j_1} \cup \cdots \cup I_{j_N}$.

Proof. Suppose that f is Riemann integrable on $[a,b]$. Recall the formula

$$D = \bigcup_{k=1}^{\infty} O(k).$$

By Proposition 12 (b) it suffices to show that each $O(k)$ is of measure zero. Given $\varepsilon > 0$, by Integrability Criterion II, we can find a partition P such that

$$\overline{S}(f,P) - \underline{S}(f,P) < \varepsilon/2k.$$

Let J be the index set of those subintervals of P which contains some elements of $O(k)$ in their interiors. Then

$$\begin{aligned} \frac{1}{k}\sum_{j\in J} |I_j| &\le \sum_{j\in J} (\sup_{I_j} f - \inf_{I_j} f)\Delta x_j \\ &\le \sum_{j=1}^{n} (\sup_{I_j} f - \inf_{I_j} f)\Delta x_j \\ &= \overline{S}(f,P) - \underline{S}(f,P) \\ &< \varepsilon/2k. \end{aligned}$$

Therefore

$$\sum_{j\in J} |I_j| < \varepsilon/2.$$

Now, the only possibility that an element of $O(k)$ is not contained by one of these I_j is it being a partition point. Since there are finitely many partition points, say N, we can find some open intervals $I'_1, ..., I'_N$ containing these partition points which satisfy

$$\sum |I'_i| < \varepsilon/2.$$

So $\{I_j\}$ and $\{I'_i\}$ together form a covering of $O(k)$ and its total length is strictly less than ε. We conclude that $O(k)$ is of measure zero.

Conversely, given $\varepsilon > 0$, fix a large k such that $\frac{1}{k} < \varepsilon$. Now the set $O(k)$ is of measure zero, we can find a sequence of open intervals $\{I_j\}$ satisfying

$$O(k) \subseteq \bigcup_{j=1}^{\infty} I_j,$$

$$\sum_{j=1}^{\infty} |I_{i_j}| < \varepsilon.$$

One can show that $O(k)$ is closed and bounded, hence it is compact. As a result, we can find $I_{i_1}, ..., I_{i_N}$ from $\{I_j\}$ so that

$$O(k) \subseteq I_{i_1} \cup ... \cup I_{i_N},$$

$$\sum_{j=1}^{N} |I_j| < \varepsilon.$$

Without loss of generality we may assume that these open intervals are mutually disjoint since, whenever two intervals have nonempty intersection, we can put them together to form a larger open interval. Observe that $[a, b] \setminus (I_{i_1} \cup \cdots \cup I_{i_N})$ is a finite disjoint union of closed bounded intervals, call them $V'_i s$, $i \in A$. We will show that for each $i \in A$, one can find a partition on each $V_i = [v_{i-1}, v_i]$ such that the oscillation of f on each subinterval in this partition is less than $1/k$.

Fix $i \in A$. For each $x \in V_i$, we have

$$\omega(f, x) < \frac{1}{k}.$$

By the definition of $\omega(f, x)$, one can find some $\delta_x > 0$ such that

$$\sup\{f(y) : y \in B(x, \delta_x) \cap [a, b]\} - \inf\{f(z) : z \in B(x, \delta_x) \cap [a, b]\} < \frac{1}{k},$$

where $B(y, \beta) = (y - \beta, y + \beta)$. Note that $V_i \subseteq \bigcup_{x \in V_i} B(x, \delta_x)$. Since V_i is closed and bounded, it is compact. Hence, there exist $x_{l_1}, \ldots, x_{l_M} \in V_i$ such that $V_i \subseteq \bigcup_{j=1}^{M} B(x_{i_j}, \delta_{x_{l_j}})$. By replacing the left end point of $B(x_{i_j}, \delta_{x_{l_j}})$ with v_{i-1} if $x_{l_j} - \delta_{x_{l_j}} < v_{i-1}$, and replacing the right end point of $B(x_{i_j}, \delta_{x_{l_j}})$ with v_i if

$x_{l_j} + \delta_{x_{l_j}} > v_i$, one can list out the endpoints of $\{B(x_{l_j}, \delta_{l_j})\}_{j=1}^M$ and use them to form a partition S_i of V_i. It can be easily seen that each subinterval in S_i is covered by some $B(x_{l_j}, \delta_{x_{l_j}})$, which implies that the oscillation of f in each subinterval is less than $1/k$. So, S_i is the partition that we want.

The partitions S_i's and the endpoints of $I_{i_1}, ..., I_{i_N}$ form a partition P of $[a, b]$. We have

$$\begin{aligned}\overline{S}(f, P) - \underline{S}(f, P) &= \sum_{I_{i_j}} (M_j - m_j)\Delta x_j + \sum (M_j - m_j)\Delta x_j \\ &\le 2M \sum_{j=1}^{N} |I_{i_j}| + \frac{1}{k} \sum \Delta x_j \\ &\le 2M\varepsilon + \varepsilon(b - a) \\ &= [2M + (b - a)]\varepsilon,\end{aligned}$$

where $M = \sup_{[a,b]} |f|$ and the second summation is over all subintervals in $V_i, i \in A$. By Integrability Criterion II, f is integrable on $[a, b]$. □