Chapter 8

Integration

8.1 Partitions

A *partition* of the interval $[a, b]$ is a finite set $P \subset [a, b]$ such that $\{a, b\} \subset P$. The set of all partitions of $[a, b]$ is denoted $\mathcal{P}([a, b])$. Basically, a partition should be thought of as a way to divide an interval into a finite number of subintervals.

If $P \in \mathcal{P}([a, b])$, then the elements of P can be ordered in a list as $a = x_0 < x_1 < \cdots < x_n = b$. The adjacent points of this partition determine n compact intervals of the form $I_k^P = [x_{k-1}, x_k]$, $1 \le k \le n$. If the partition is clear from the context, we write I_k instead of I_k^P. It's clear that these intervals only intersect at their common endpoints.

Since it's inconvenient to always list each part of a partition, we'll use the partition of the previous paragraph as the generic partition. Unless it's necessary within the context to specify some other form for a partition, assume any partition is the generic partition.

If I is any interval, its length is written $|I|$. Using the notation of the previous paragraph, it follows that

$$\sum_{k=1}^{n} |I_k| = \sum_{k=1}^{n} (x_k - x_{k-1}) = x_n - x_0 = b - a.$$

The *norm* of a partition P is

$$\|P\| = \max\{|I_k^P| : 1 \le k \le n\}.$$

In other words, the norm of P is just the length of the longest subinterval determined by P. If $|I_k| = \|P\|$ for every I_k, then P is called a *regular* partition.

Suppose $P, Q \in \mathcal{P}([a, b])$. If $P \subset Q$, then Q is called a *refinement* of P. When this happens, we write $P \ll Q$. In this case, it's easy to see that $P \ll Q$ implies $\|P\| \ge \|Q\|$. It also follows at once from the definitions that $P \cup Q \in \mathcal{P}([a, b])$ with $P \ll P \cup Q$ and $Q \ll P \cup Q$. The partition $P \cup Q$ is called the *common refinement* of P and Q.

8.2 Riemann Sums

Let $f : [a, b] \to \mathbb{R}$ and $P \in \mathcal{P}([a, b])$. Choose $x_k^* \in I_k$ for each k. The set $\{x_k^* : 1 \le k \le n\}$ is called a *selection* from P. The expression

$$\mathcal{R}(f, P, x_k^*) = \sum_{k=1}^{n} f(x_k^*)|I_k|$$

is the *Riemann sum* for f with respect to the partition P and selection x_k^*. Notice that given a particular function f and partition P, there are an uncountably infinite number of different possible Riemann sums, depending on the selection x_k^*. This sometimes makes working with Riemann sums quite complicated.

Example 8.2.1. Suppose $f : [a, b] \to \mathbb{R}$ is the constant function $f(x) = c$. If $P \in \mathcal{P}([a, b])$ and $\{x_k^* : 1 \le k \le n\}$ is any selection from P, then

$$\mathcal{R}(f, P, x_k^*) = \sum_{k=1}^{n} f(x_k^*)|I_k| = c \sum_{k=1}^{n} |I_k| = c(b - a).$$

Example 8.2.2. Suppose $f(x) = x$ on $[a, b]$. Choose any $P \in \mathcal{P}([a, b])$ where $\|P\| < 2(b - a)/n$. (Convince yourself this is always possible.[1]) Make two specific selections $l_k^* = x_{k-1}$ and $r_k^* = x_k$. If x_k^* is any other selection from P, then $l_k^* \le x_k^* \le r_k^*$ and the fact that f is increasing on $[a, b]$ gives

$$\mathcal{R}(f, P, l_k^*) \le \mathcal{R}(f, P, x_k^*) \le \mathcal{R}(f, P, r_k^*).$$

With this in mind, consider the following calculation.

$$\begin{aligned}
\mathcal{R}(f, P, r_k^*) - \mathcal{R}(f, P, l_k^*) &= \sum_{k=1}^{n} (r_k^* - l_k^*)|I_k| \qquad (8.1)\\
&= \sum_{k=1}^{n} (x_k - x_{k-1})|I_k|\\
&= \sum_{k=1}^{n} |I_k|^2\\
&\le \sum_{k=1}^{n} \|P\|^2\\
&= n\|P\|^2\\
&< \frac{4(b - a)^2}{n}
\end{aligned}$$

This shows that if a partition is chosen with a small enough norm, all the Riemann sums for f over that partition will be close to each other.

[1] This is with the generic partition

In the special case when P is a regular partition, $|I_k| = (b-a)/n$, $r_k = a + k(b-a)/n$ and

$$\begin{aligned}
\mathcal{R}(f, P, r_k^*) &= \sum_{k=1}^{n} r_k |I_k| \\
&= \sum_{k=1}^{n} \left(a + \frac{k(b-a)}{n} \right) \frac{b-a}{n} \\
&= \frac{b-a}{n} \left(na + \frac{b-a}{n} \sum_{k=1}^{n} k \right) \\
&= \frac{b-a}{n} \left(na + \frac{b-a}{n} \frac{n(n+1)}{2} \right) \\
&= \frac{(b-a)\ (a(n-1) + b(n+1))}{2n}.
\end{aligned}$$

In the limit as $n \to \infty$, this becomes the familiar formula $(b^2 - a^2)/2$, for the integral of $f(x) = x$ over $[a, b]$.

Definition 8.2.1. The function f is *Riemann integrable* on $[a, b]$, if there exists a number $\mathcal{R}(f)$ such that for all $\varepsilon > 0$ there is a $\delta > 0$ so that whenever $P \in \mathcal{P}([a, b])$ with $\|P\| < \delta$, then

$$|R(f) - R(f, P, x_k^*)| < \varepsilon$$

for any selection x_k^* from P.

Theorem 8.2.1. *If $f : [a, b] \to \mathbb{R}$ and $\mathcal{R}(f)$ exists, then $\mathcal{R}(f)$ is unique.*

Proof. Suppose $R_1(f)$ and $R_2(f)$ both satisfy the definition and $\varepsilon > 0$. For $i = 1, 2$ choose $\delta_i > 0$ so that whenever $\|P\| < \delta_i$, then

$$|R_i(f) - R(f, P, x_k^*)| < \varepsilon/2,$$

as in the definition above. If $P \in \mathcal{P}([a, b])$ so that $\|P\| < \delta_1 \wedge \delta_2$, then

$$|R_1(f) - R_2(f)| \leq | R_1(f) - R(f, P, x_k^*)| + |R_2(f) - R(f, P, x_k^*)| < \varepsilon$$

and it follows $R_1(f) = R_2(f)$. □

Theorem 8.2.2. *If $f : [a, b] \to \mathbb{R}$ and $\mathcal{R}(f)$ exists, then f is bounded.*

Proof. Left as an exercise. □

8.3 Darboux Integration

A difficulty with handling Riemann sums is that there are an uncountably infinite number of Riemann sums associated with each partition. One way to resolve this problem was shown in Example 8.2.2, where it was shown there were largest and smallest Riemann sums associated with each partition. However, that's not always the case, so to use that idea, a little more care must be taken.

Definition 8.3.1. Let $f : [a, b] \to \mathbb{R}$ be bounded and $P \in \mathcal{P}([a, b])$. For each I_k determined by P, let

$$M_k = \sup\{f(x) : x \in I_k\} \quad \text{and} \quad m_k = \inf\{f(x) : x \in I_k\}.$$

The *upper and lower Darboux sums* for f on $[a, b]$ are

$$\overline{\mathcal{D}}(f, P) = \sum_{k=1}^{n} M_k |I_k| \quad \text{and} \quad \underline{\mathcal{D}}(f, P) = \sum_{k=1}^{n} m_k |I_k|.$$

Theorem 8.3.1. *If $f : [a, b] \to \mathbb{R}$ is bounded and $P, Q \in \mathcal{P}([a, b])$ with $P \ll Q$, then*

$$\underline{\mathcal{D}}(f, P) \le \underline{\mathcal{D}}(f, Q) \le \overline{\mathcal{D}}(f, Q) \le \overline{\mathcal{D}}(f, P).$$

Proof. Let P be the generic partition and let $Q = P \cup \{\overline{x}\}$, where $\overline{x} \in (x_{k_0-1}, x_{x_0})$ for some k_0. Clearly, $P \ll Q$. Let

$$\begin{aligned}
M_l &= \sup\{f(x) : x \in [x_{k_0-1}, \overline{x}]\}\\
m_l &= \inf\{f(x) : x \in [x_{k_0-1}, \overline{x}]\}\\
M_r &= \sup\{f(x) : x \in [\overline{x}, x_{k_0}]\}\\
m_r &= \inf\{f(x) : x \in [\overline{x}, x_{k_0}]\}
\end{aligned}$$

Then

$$m_{k_0} \le m_l \le M_l \le M_{k_0} \quad \text{and} \quad m_{k_0} \le m_r \le M_r \le M_{k_0}$$

so that

$$\begin{aligned}
m_{k_0}|I_{k_0}| &= m_{k_0}\left(|[x_{k_0-1}, \overline{x}]| + |[\overline{x}, x_{k_0}]|\right)\\
&\le m_l |[x_{k_0-1}, \overline{x}]| + m_r |[\overline{x}, x_{k_0}]|\\
&\le M_l |[x_{k_0-1}, \overline{x}]| + M_r |[\overline{x}, x_{k_0}]|\\
&\le M_{k_0} |[x_{k_0-1}, \overline{x}]| + M_{k_0} |[\overline{x}, x_{k_0}]|\\
&= M_{k_0}|I_{k_0}|.
\end{aligned}$$

This implies

$$
\begin{aligned}
\underline{\mathcal{D}}(f,P) &= \sum_{k=1}^{n} m_k|I_k| \\
&= \sum_{k=1}^{k_0-1} m_k|I_k| + m_{k_0}|I_{k_0}| + \sum_{k=k_0+1}^{n} m_k|I_k| \\
&\le \sum_{k=1}^{k_0-1} m_k|I_k| + m_l|[x_{k_0-1},\overline{x}]| + m_r|[\overline{x},x_{k_0}]| + \sum_{k=k_0+1}^{n} m_k|I_k| \\
&= \underline{\mathcal{D}}(f,Q) \\
&\le \overline{\mathcal{D}}(f,Q) \\
&= \sum_{k=1}^{k_0-1} M_k|I_k| + M_l|[x_{k_0-1},\overline{x}]| + M_r|[\overline{x},x_{k_0}]| + \sum_{k=k_0+1}^{n} M_k|I_k| \\
&\le \sum_{k=1}^{n} M_k|I_k| \\
&= \overline{\mathcal{D}}(f,P)
\end{aligned}
$$

The argument given above shows that the theorem holds if Q has one more point that P. Using induction, this same technique also shows that the theorem holds when Q has an arbitrarily larger number of points than P. □

The main lesson to be learned from Theorem 8.3.1 is that refining a partition causes the lower Darboux sum to increase and the upper Darboux sum to decrease. Moreover, if $P,Q \in \mathcal{P}([a,b])$ and $f : [a,b] \to [-B,B]$, then,

$$\underline{\mathcal{D}}(f,P) \le \underline{\mathcal{D}}(f,P\cup Q) \le \overline{\mathcal{D}}(f,P\cup Q) \le \overline{\mathcal{D}}(f,Q).$$

Therefore *every Darboux lower sum is less than or equal to every Darboux upper sum.* Consider the following definition with this in mind.

Definition 8.3.2. The *upper and lower Darboux integrals* of a bounded function $f : [a,b] \to \mathbb{R}$ are

$$\overline{\mathcal{D}}(f) = \inf\{\overline{\mathcal{D}}(f,P) : P \in \mathcal{P}([a,b])\} \quad \text{and} \quad \underline{\mathcal{D}}(f) = \sup\{\underline{\mathcal{D}}(f,P) : P \in \mathcal{P}([a,b])\},$$

respectively.

As a consequence of the observations preceding the definition, it follows that $\overline{\mathcal{D}}(f) \ge \underline{\mathcal{D}}(f)$ always. In the case $\overline{\mathcal{D}}(f) = \underline{\mathcal{D}}(f)$, the function is said to be *Darboux integrable* on $[a,b]$, and the common value is written $\mathcal{D}(f)$. The following is obvious.

Corollary 8.3.2. *A bounded function $f : [a,b] \to \mathbb{R}$ is Darboux integrable if and only if for all $\varepsilon > 0$ there is a $P \in \mathcal{P}([a,b])$ such that $\overline{\mathcal{D}}(f,P) - \underline{\mathcal{D}}(f,P) < \varepsilon$.*

Which functions are Darboux integrable? The following corollary gives a first approximation to an answer.

Corollary 8.3.3. *If $f \in C([a,b])$, then $\mathcal{D}(f)$ exists.*

Proof. Let $\varepsilon > 0$. According to Corollary 6.6.3, f is uniformly continuous, so there is a $\delta > 0$ such that whenever $x, y \in [a,b]$ with $|x-y| < \delta$, then $|f(x) - f(y)| < \varepsilon/(b-a)$. Let $P \in \mathcal{P}([a,b])$ with $\|P\| < \delta$. By Corollary 6.5.4, in each subinterval I_i determined by P, there are $x_i^*, y_i^* \in I_i$ such that

$$f(x_i^*) = \inf\{f(x) : x \in I_i\} \quad \text{and} \quad f(y_i^*) = \sup\{f(x) : x \in I_i\}.$$

Since $|x_i^* - y_i^*| \le |I_i| < \delta$, we see $0 \le f(x_i^*) - f(y_i^*) < \varepsilon/(b-a)$, for $1 \le i \le n$. Then

$$\begin{aligned}
\overline{\mathcal{D}}(f) - \underline{\mathcal{D}}(f) &\le \overline{\mathcal{D}}(f,P) - \underline{\mathcal{D}}(f,P) \\
&= \sum_{i=1}^n f(x_i^*)|I_i| - \sum_{i=1}^n f(y_i^*)|I_i| \\
&= \sum_{i=1}^n (f(x_i^*) - f(y_i^*))|I_i| \\
&< \frac{\varepsilon}{b-a} \sum_{i=1}^n |I_i| \\
&= \varepsilon
\end{aligned}$$

and the corollary follows. □

This corollary should not be construed to imply that *only* continuous functions are Darboux integrable. In fact, the set of integrable functions is much more extensive than only the continuous functions. Consider the following example.

Example 8.3.1. Let f be the salt and pepper function of Example 6.3.6. It was shown that $C(f) = \mathbb{Q}^c$. We claim that f is Darboux integrable over any compact interval $[a,b]$.

To see this, let $\varepsilon > 0$ and $N \in \mathbb{N}$ so that $1/N < \varepsilon/2(b-a)$. Let

$$\{q_{k_i} : 1 \le i \le m\} = \{q_k : 1 \le k \le N\} \cap [a,b]$$

and choose $P \in \mathcal{P}([a,b])$ such that $\|P\| < \varepsilon/2m$. Then

$$\begin{aligned}\overline{\mathcal{D}}(f,P) &= \sum_{\ell=1}^{n} \operatorname{lub}\{f(x) : x \in I_\ell\}|I_\ell| \\ &= \sum_{q_{k_i} \notin I_\ell} \operatorname{lub}\{f(x) : x \in I_\ell\}|I_\ell| + \sum_{q_{k_i} \in I_\ell} \operatorname{lub}\{f(x) : x \in I_\ell\}|I_\ell| \\ &\le \frac{1}{N}(b-a) + m\|P\| \\ &< \frac{\varepsilon}{2(b-a)}(b-a) + m\frac{\varepsilon}{2m} \\ &= \varepsilon.\end{aligned}$$

Since $f(x) = 0$ whenever $x \in \mathbb{Q}^c$, it follows that $\underline{\mathcal{D}}(f,P) = 0$. Therefore, $\overline{\mathcal{D}}(f) = \underline{\mathcal{D}}(f) = 0$ and $\mathcal{D}(f) = 0$.

8.4 The Integral

There are now two different definitions for the integral. It would be embarassing, if they gave different answers. The following theorem shows they're really different sides of the same coin.[2]

Theorem 8.4.1. *Let $f : [a,b] \to \mathbb{R}$.*

(a) $\mathcal{R}(f)$ exists iff $\mathcal{D}(f)$ exists.

(b) If $\mathcal{R}(f)$ exists, then $\mathcal{R}(f) = \mathcal{D}(f)$.

Proof. (a) ($\Longrightarrow$) Suppose $\mathcal{R}(f)$ exists and $\varepsilon > 0$. By Theorem 8.2.2, f is bounded. Choose $P \in \mathcal{P}([a,b])$ such that

$$|\mathcal{R}(f) - \mathcal{R}(f,P,x_k^*)| < \varepsilon/4$$

for all selections x_k^* from P. From each I_k, choose $\overline{x}_k$ and $\underline{x}_k$ so that

$$M_k - f(\overline{x}_k) < \frac{\varepsilon}{4(b-a)} \quad \text{and} \quad f(\underline{x}_k) - m_k < \frac{\varepsilon}{4(b-a)}.$$

Then

$$\begin{aligned}\overline{\mathcal{D}}(f,P) - \mathcal{R}(f,P,\overline{x}_k) &= \sum_{k=1}^{n} M_k|I_k| - \sum_{k=1}^{n} f(\overline{x}_k)|I_k| \\ &= \sum_{k=1}^{n} (M_k - \overline{x}_k)|I_k| \\ &< \frac{\varepsilon}{4(b-a)}(b-a) = \frac{\varepsilon}{4}.\end{aligned}$$

[2]Theorem 8.4.1 shows that the two integrals presented here are the same. But, there are many other integrals, and not all of them are equivalent. For example, the well-known Lebesgue integral includes all Riemann integrable functions, but not all Lebesgue integrable functions are Riemann integrable. The Denjoy integral is another extension of the Riemann integral which is not the same as the Lebesgue integral. For more discussion of this, see [3].

In the same way,

$$\mathcal{R}(f,P,\underline{x}_k) - \underline{\mathcal{D}}(f,P) < \varepsilon/4.$$

Therefore,

$$\begin{aligned}
\overline{\mathcal{D}}(f) - \underline{\mathcal{D}}(f) &= \inf\{\overline{\mathcal{D}}(f,Q) : Q \in \mathcal{P}([a,b])\} - \sup\{\underline{\mathcal{D}}(f,Q) : Q \in \mathcal{P}([a,b])\}\\
&\le \overline{\mathcal{D}}(f,P) - \underline{\mathcal{D}}(f,P)\\
&< \left(\mathcal{R}(f,P,\overline{x}_k) + \frac{\varepsilon}{4}\right) - \left(\mathcal{R}(f,P,\underline{x}_k) - \frac{\varepsilon}{4}\right)\\
&\le |\mathcal{R}(f,P,\overline{x}_k) - \mathcal{R}(f,P,\underline{x}_k)| + \frac{\varepsilon}{2}\\
&< |\mathcal{R}(f,P,\overline{x}_k) - \mathcal{R}(f)| + |\mathcal{R}(f) - \mathcal{R}(f,P,\underline{x}_k)| + \frac{\varepsilon}{2}\\
&< \varepsilon
\end{aligned}$$

Since ε is an arbitrary positive number, this shows that $\mathcal{D}(f)$ exists and equals $\mathcal{R}(f)$, which is part (b) of the theorem.

($\Longleftarrow$) Suppose $f : [a,b] \to [-B,B]$, $\mathcal{D}(f)$ exists and $\varepsilon > 0$. Since $\mathcal{D}(f)$ exists, there is a $P_1 = \{p_i : 0 \le i \le m\} \in \mathcal{P}([a,b])$ such that

$$\overline{\mathcal{D}}(f,P_1) - \underline{\mathcal{D}}(f,P_1) < \frac{\varepsilon}{2}.$$

Set $\delta = \varepsilon/8mB$. Choose $P \in \mathcal{P}([a,b])$ with $\|P\| < \delta$ and let $P_2 = P \cup P_1$. Since $P_1 \ll P_2$, according to Theorem 8.3.1,

$$\overline{\mathcal{D}}(f,P_2) - \underline{\mathcal{D}}(f,P_2) < \frac{\varepsilon}{2}.$$

Thinking of P as the generic partition, notice that if $(x_{i-1},x_i) \cap P_1 = \emptyset$, then $\overline{\mathcal{D}}(f,P)$ and $\overline{\mathcal{D}}(f,P_2)$ share a common term, $M_i|I_i|$. There are at most $m-1$ instances where $(x_{i-1},x_i) \cap P_1 \neq \emptyset$ and each such intersection generates a subinterval of P_2 with length less than δ. Therefore,

$$\overline{\mathcal{D}}(f,P) - \overline{\mathcal{D}}(f,P_2) < (m-1)2B\delta < \frac{\varepsilon}{4}.$$

In the same way,

$$\underline{\mathcal{D}}(f,P_2) - \underline{\mathcal{D}}(f,P) < (m-1)2B\delta < \frac{\varepsilon}{4}.$$

Putting these estimates together yields

$$\begin{aligned}
&\overline{\mathcal{D}}(f,P) - \underline{\mathcal{D}}(f,P) =\\
&\quad \left(\overline{\mathcal{D}}(f,P) - \overline{\mathcal{D}}(f,P_2)\right) + \left(\overline{\mathcal{D}}(f,P_2) - \underline{\mathcal{D}}(f,P_2)\right) + \left(\underline{\mathcal{D}}(f,P) - \underline{\mathcal{D}}(f,P)\right)\\
&\qquad < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon
\end{aligned}$$

This shows that, given $\varepsilon > 0$, there is a $\delta > 0$ so that $\|P\| < \delta$ implies

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

Since

$$\underline{\mathcal{D}}(f,P) \le \mathcal{D}(f) \le \overline{\mathcal{D}}(f,P) \text{ and } \underline{\mathcal{D}}(f,P) \le \mathcal{R}(f,P,x_i^*) \le \overline{\mathcal{D}}(f,P)$$

for every selection x_i^* from P, it follows that $\mathcal{R}(f,P,x_i^*) - \mathcal{D}(f) < \varepsilon$ when $\|P\| < \delta$. We conclude f is Riemann integrable and $\mathcal{R}(f) = \mathcal{D}(f)$. □

From Theorem 8.4.1, we are justified in using a single notation for both $\mathcal{R}(f)$ and $\mathcal{D}(f)$. The obvious choice is the familiar $\int_a^b f(x)\,dx$, or, more simply, $\int_a^b f$.

Example 8.4.1. If f is the salt and pepper function of Example 6.3.6, then $\int_a^b f = 0$ for any interval $[a,b]$.

To see this let $\varepsilon > 0$. There is a finite set $\{q_{k_1}, q_{k_2}, \ldots, q_{k_n}\} \subset \mathbb{Q}$ consisting of all the rational numbers in $[a,b]$ where $f(q_{k_j}) > \varepsilon/2(b-a)$.

8.5 The Cauchy Criterion

We now face a conundrum. In order to show that $\int_a^b f$ exists, we must know its value. It's often very hard to determine the value of an integral, even if the integral exists. We've faced this same situation before with sequences. The basic definition of convergence for a sequence, Definition 3.1.2, requires the limit of the sequence be known. The path out of the dilemma in the case of sequences was the Cauchy criterion for convergence, Theorem 3.5.1. The solution is the same here, with a Cauchy criterion for the existence of the integral.

Theorem 8.5.1 (Cauchy Criterion)**.** *Let $f : [a,b] \to \mathbb{R}$. The following statements are equivalent.*

(a) $\int_a^b f$ exists.

(b) Given $\varepsilon > 0$ there exists $P \in \mathcal{P}([a,b])$ such that if $P \ll Q_1$ and $P \ll Q_2$, then

$$|\mathcal{R}(f,Q_1,x_k^*) - \mathcal{R}(f,Q_2,y_k^*)| < \varepsilon \tag{8.2}$$

for any selections from Q_1 and Q_2.

Proof. ($\Longrightarrow$) Assume $\int_a^b f$ exists. According to Definition 8.2.1, there is a $\delta > 0$ such that whenever $P \in \mathcal{P}([a,b])$ with $\|P\| < \delta$, then $|\mathcal{R}(f,P,x_i^*,|,<)\varepsilon/2$ for every selection. If $P \ll Q_1$ and $P \ll Q_2$, then $\|Q_1\| < \delta$, $\|Q2\| < \delta$ and a simple application of the triangle inequality shows

$$|\mathcal{R}(f,Q_1,x_k^*) - \mathcal{R}(f,Q_2,y_k^*)| \le \left|\mathcal{R}(f,Q_1,x_k^*) - \int_a^b f\right| + \left|\int_a^b f - \mathcal{R}(f,Q_2,y_k^*)\right| < \varepsilon.$$

($\Longleftarrow$) Let $\varepsilon > 0$ and choose $P \in \mathcal{P}([a,b])$ satisfying (8.2) with $\varepsilon/2$ in place of ε.

We first claim that f is bounded. To see this, suppose it is not. Then it must be unbounded on an interval I_{k_0} determined by P. Fix a selection $\{x_k \in I_k : 1 \le k \le n\}$ and let $y_k = x_k$ for $k \ne k_0$ with y_{k_0} any element of I_{k_0}. Then

$$\frac{\varepsilon}{2} > |\mathcal{R}(f,P,x_k^*) - \mathcal{R}(f,P,y_k^*)| = |f(x_{k_0}) - f(y_{k_0}\| \, |I_{k_0}| \, .$$

But, the right-hand side can be made bigger than $\varepsilon/2$ with an appropriate choice of y_{k_0} because of the assumption that f is unbounded on I_{k_0}. This contradiction forces the conclusion that f is bounded.

Thinking of P as the generic partition and using m_k and M_k as usual with Darboux sums, for each k, choose $x_k^*, y_k^* \in I_k$ such that

$$M_k - f(x_k^*) < \frac{\varepsilon}{4n|I_k|} \text{ and } f(y_k^*) - m_k < \frac{\varepsilon}{4n|I_k|}.$$

With these selections,

$$\begin{aligned}
\overline{\mathcal{D}}(f,P) - \underline{\mathcal{D}}(f,P) &= \sum_{k=1}^{n} (M_k - m_k)|I_k| \\
&\le \sum_{k=1}^{n} \left(|M_k - f(x_k^*)| + |f(x_k^*) - f(y_k^*)| + |f(y_k^*) - m_k|\right)|I_k| \\
&< \sum_{k=1}^{n} \left(\frac{\varepsilon}{4n|I_k|} + |f(x_k^*) - f(y_k^*)| + \frac{\varepsilon}{4n|I_k|}\right)|I_k| \\
&\le \frac{\varepsilon}{2} + |\mathcal{R}(f,P,x_k^*) - \mathcal{R}(f,P,y_k^*)| < \varepsilon
\end{aligned}$$

Corollary 8.3.2 implies $\mathcal{D}(f)$ exists and Theorem 8.4.1 finishes the proof. □

Corollary 8.5.2. *If $\int_a^b f$ exists and $[c,d] \subset [a,b]$, then $\int_c^d f$ exists.*

Proof. Let $P_0 = \{a,b,c,d\} \in \mathcal{P}([a,b])$ and $\varepsilon > 0$. Choose a partition P_ε such that $P_0 \ll P_\varepsilon$ and whenever $P_\varepsilon \ll P$ and $P_\varepsilon \ll P'$, then

$$|\mathcal{R}(f,P,x_k^*) - \mathcal{R}(f,P',y_k^*)| < \varepsilon.$$

Let $P_\varepsilon^1 \in \mathcal{P}([a,c])$, $P_\varepsilon^2 \in \mathcal{P}([c,d])$ and $P_\varepsilon^3 \in \mathcal{P}([d,b])$ so that $P_\varepsilon = P_\varepsilon^1 \cup P_\varepsilon^2 \cup P_\varepsilon^3$. Suppose $P_\varepsilon^2 \ll Q_1$ and $P_\varepsilon^2 \ll Q_2$. Then $P_\varepsilon^1 \cup Q_i \cup P_\varepsilon^3$ for $i = 1,2$ are refinements of P_ε and

$$\begin{aligned}
&|\mathcal{R}(f,Q_1,x_k^*) - \mathcal{R}(f,Q_2,,)x_k^*| = \\
&\qquad |\mathcal{R}(f,P_\varepsilon^1 \cup Q_1 \cup P_\varepsilon^3, x_k^*) - \mathcal{R}(f,P_\varepsilon^1 \cup Q_2 \cup P_\varepsilon^3, x_k^*)| < \varepsilon
\end{aligned}$$

An application of (8.5.1) shows $\int_a^b f$ exists. □

8.6 Properties of the Integral

Theorem 8.6.1. *If $\int_a^b f$ and $\int_a^b g$ both exist, then*

(a) If $\alpha, \beta \in \mathbb{R}$, then $\int_a^b (\alpha f + \beta g)$ exists and $\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g$.

(b) $\int_a^b fg$ exists.

(c) $\int_a^b |f|$ exists.

Proof. (a) Let $\varepsilon > 0$. If $\alpha = 0$, in light of Example 8.2.1, it is clear αf is integrable. So, assume $\alpha \neq 0$, and choose a partition $P_f \in \mathcal{P}([a,b])$ such that whenever $P_f \ll P$, then

$$\left| \mathcal{R}(f, P, x_k^*) - \int_a^b f \right| < \frac{\varepsilon}{2|\alpha|}.$$

Then

$$\begin{aligned} \left| \mathcal{R}(\alpha f, P, x_k^*) - \alpha \int_a^b f \right| &= \left| \sum_{k=1}^n \alpha f(x_k^*)|I_k| - \alpha \int_a^b f \right| \\ &= |\alpha| \left| \sum_{k=1}^n f(x_k^*)|I_k| - \int_a^b f \right| \\ &= |\alpha| \left| \mathcal{R}(f, P, x_k^*) - \int_a^b f \right| \\ &< |\alpha| \frac{\varepsilon}{2|\alpha|} \\ &= \frac{\varepsilon}{2}. \end{aligned}$$

This shows αf is integrable and $\int_a^b \alpha f = \alpha \int_a^b f$.

Assuming $\beta \neq 0$, in the same way, we can choose a $P_g \in \mathcal{P}([a,b])$ such that when $P_g \ll P$, then

$$\left| \mathcal{R}(g, P, x_k^*) - \int_a^b g \right| < \frac{\varepsilon}{2|\beta|}.$$

Let $P_\varepsilon = P_f \cup P_g$ be the common refinement of P_f and P_g, and suppose $P_\varepsilon \ll P$. Then

$$\begin{aligned} &|\mathcal{R}(\alpha f + \beta g, P, x_k^*) - \left(\alpha \int_a^b f + \beta \int_a^b g \right)| \\ &\qquad \leq |\alpha||\mathcal{R}(f, P, x_k^*) - \int_a^b f| + |\beta||\mathcal{R}(g, P, x_k^*) - \int_a^b g| < \varepsilon \end{aligned}$$

This shows $\alpha f + \beta g$ is integrable and $\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g$.

(b) Claim: If $\int_a^b h$ exists, then so does $\int_a^b h^2$

To see this, suppose first that $0 \le h(x) \le M$ on $[a,b]$. If $M = 0$, the claim is trivially true, so suppose $M > 0$. Let $\varepsilon > 0$ and choose $P \in \mathcal{P}([a,b])$ such that

$$\overline{\mathcal{D}}(h,P) - \underline{\mathcal{D}}(h,P) \le \frac{\varepsilon}{2M}.$$

For each $1 \le k \le n$, let

$$m_k = \inf\{h(x) : x \in I_k\} \le \sup\{h(x) : x \in I_k\} = M_k.$$

Since $h \ge 0$,

$$m_k^2 = \inf\{h(x)^2 : x \in I_k\} \le \sup\{h(x)^2 : x \in I_k\} = M_k^2.$$

Using this, we see

$$\begin{aligned}
\overline{\mathcal{D}}(h^2,P) - \underline{\mathcal{D}}(h^2,P) &= \sum_{k=1}^n (M_k^2 - m_k^2)|I_k| \\
&= \sum_{k=1}^n (M_k + m_k)(M_k - m_k)|I_k| \\
&\le 2M \left(\sum_{k=1}^n (M_k - m_k)|I_k| \right) \\
&= 2M \left(\overline{\mathcal{D}}(h,P) - \underline{\mathcal{D}}(h,P) \right) \\
&< \varepsilon.
\end{aligned}$$

Therefore, h^2 is integrable when $h \ge 0$.

If h is not nonnegative, let $m = \inf\{h(x) : a \le x \le b\}$. Then $h - m \ge 0$, and $h - m$ is integrable by (a). From the claim, $(h-m)^2$ is integrable. Since

$$h^2 = (h-m)^2 + 2mh - m^2,$$

it follows from (a) that h^2 is integrable.

Finally, $fg = \frac{1}{4}((f+g)^2 - (f-g)^2)$ is integrable by the claim and (a).

(c) Claim: If $h \ge 0$ is integrable, then so is $\sqrt{h}$.

To see this, let $\varepsilon > 0$ and choose $P \in \mathcal{P}([a,b])$ such that

$$\overline{\mathcal{D}}(h,P) - \underline{\mathcal{D}}(h,P) < \varepsilon^2.$$

For each $1 \le k \le n$, let

$$m_k = \inf\{\sqrt{h(x)} : x \in I_k\} \le \sup\{\sqrt{h(x)} : x \in I_k\} = M_k.$$

and define

$$A = \{k : M_k - m_k < \varepsilon\} \quad \text{and} \quad B = \{k : M_k - m_k \ge \varepsilon\}.$$

Then

$$\sum_{k\in A}(M_k - m_k)|I_k| < \varepsilon(b-a). \tag{8.3}$$

Using the fact that $m_k \geq 0$, we see that $M_k - m_k \leq M_k + m_k$, and

$$\begin{aligned}\sum_{k\in B}(M_k - m_k)|I_k| &\leq \frac{1}{\varepsilon}\sum_{k\in B}(M_k + m_k)(M_k - m_k)|I_k| \\ &= \frac{1}{\varepsilon}\sum_{k\in B}(M_k^2 - m_k^2)|I_k| \\ &\leq \frac{1}{\varepsilon}\left(\overline{\mathcal{D}}(h,P) - \underline{\mathcal{D}}(h,P)\right) \\ &< \varepsilon\end{aligned} \tag{8.4}$$

Combining (8.3) and (8.4), it follows that

$$\overline{\mathcal{D}}(\sqrt{h},P) - \underline{\mathcal{D}}(\sqrt{h},P) < \varepsilon(b-a) + \varepsilon = \varepsilon((b-a)+1)$$

can be made arbitrarily small. Therefore, $\sqrt{h}$ is integrable.

Since $|f| = \sqrt{f^2}$ an application of (b) and the claim suffice to prove (c). □

Theorem 8.6.2. *If $\int_a^b f$ exists, then*

(a) If $f \geq 0$ on $[a,b]$, then $\int_a^b f \geq 0$.

(b) $|\int_a^b f| \leq \int_a^b |f|$

(c) If $a \leq c \leq b$, then $\int_a^b f = \int_a^c f + \int_c^b f$.

Proof. (a) Since all the Riemann sums are nonnegative, this follows at once.

(b) It is always true that $|f| \pm f \geq 0$ and $|f| - f \geq 0$, so by (a), $\int_a^b(|f|+f) \geq 0$ and $\int_a^b(|f|-f) \geq 0$. Rearranging these shows $-\int_a^b f \leq \int_a^b |f|$ and $\int_a^b f \leq \int_a^b |f|$. Therefore, $|\int_a^b f| \leq \int_a^b |f|$, which is (b).

(c) By Corollary 8.5.2, all the integrals exist. Let $\varepsilon > 0$ and choose $P_l \in \mathcal{P}([a,c])$ and $P_r \in \mathcal{P}([c,b])$ such that whenever $P_l \ll Q_l$ and $P_r \ll Q_r$, then,

$$\left|\mathcal{R}(f,Q_l,x_k^*) - \int_a^c f\right| < \frac{\varepsilon}{2} \quad \text{and} \quad \left|\mathcal{R}(f,Q_r,y_k^*) - \int_c^b f\right| < \frac{\varepsilon}{2}.$$

If $P = P_l \cup P_r$ and $Q = Q_l \cup Q_r$, then $P, Q \in \mathcal{P}([a,b])$ and $P \ll Q$. The triangle inequality gives

$$\left|\mathcal{R}(f,Q,x_k^*) - \int_a^c f - \int_c^b f\right| < \varepsilon.$$

Since every refinement of P has the form $Q_l \cup Q_r$, part (c) follows. □

8.7 The Fundamental Theorem of Calculus

Theorem 8.7.1 (Fundamental Theorem of Calculus 1)**.** *Suppose* $f, F : [a, b] \to \mathbb{R}$ *satisfy*

(a) $\int_a^b f$ *exists*

(b) $F \in C([a, b]) \cap D((a, b))$

(c) $F'(x) = f(x),\ \forall x \in (a, b)$

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

Proof. Let $\varepsilon > 0$ and choose $P_\varepsilon \in \mathcal{P}([a, b])$ such that whenever $P_\varepsilon \ll P$, then

$$\left| \mathcal{R}(f, P, x_k^*) - \int_a^b f \right| < \varepsilon.$$

On each interval $[x_{k-1}, x_k]$ determined by P, the function F satisfies the conditions of the Mean Value Theorem. (See Corollary 7.3.3.) Therefore, for each k, there is an $x_k^* \in (x_{k-1}, x_k)$ such that $F(x_k) - F(x_{k-1}) = F'(x_k^*)(x_k - x_{k-1}) = f(x_k^*)|I_k|$. Therefore,

$$\begin{aligned} \left| \int_a^b f - (F(b) - F(a)) \right| &= \left| \int_a^b f - \sum_{k=1}^n (F(x_k) - F(x_{k-1})) \right| \\ &= \left| \int_a^b f - \sum_{k=1}^n f(x_k^*)|I_k| \right| \\ &= \left| \int_a^b f - \mathcal{R}(f, P, x_k^*) \right| \\ &< \varepsilon \end{aligned}$$

and the theorem follows. □

Corollary 8.7.2 (Integration by Parts)**.** *If* $f, g \in C([a, b]) \cap D((a, b))$ *and both* $f'g$ *and* fg' *are integrable on* $[a, b]$*, then*

$$\int_a^b fg' + \int_a^b f'g = f(b)g(b) - f(a)g(a).$$

Proof. Use Theorems 7.1.2(c) and 8.7.1. □

Example 8.7.1. Suppose f and its first n derivatives are all continuous on $[a, b]$. There is a function $R_n(x, t)$ such that

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(t)}{k!}(x - t)^k + R_n(x, t)$$

for $a \le t \le b$. Differentiate both sides of the equation with respect to t to get

$$\frac{d}{dt}R_n(x,t) = -\frac{(x-t)^{n-1}}{(n-1)!}f^{(n)}(t)$$

Using Theorem 8.7.1 gives

$$\begin{aligned} R_n(x,c) &= R_n(x,c) - R_n(x,x) \\ &= \int_x^c \frac{d}{dt}R_n(x,t)\,dt \\ &= \int_c^x \frac{(x-t)^{n-1}}{(n-1)!}f^{(n)}(t)\,dt \end{aligned}$$

which is the integral form of the remainder from Taylor's formula.

Suppose $\int_a^b f$ exists. By Corollary 8.5.2, f is integrable on every interval $[a,x]$, for $x \in [a,b]$. This allows us to define a function $F : [a,b] \to \mathbb{R}$ as $F(x) = \int_a^x f$, called the *indefinite integral* of f on $[a,b]$.

Theorem 8.7.3 (Fundamental Theorem of Calculus 2)**.** *Let f be integrable on $[a,b]$ and F be the indefinite integral of f. Then $F \in C([a,b])$ and $F'(x) = f(x)$ whenever $x \in C(f) \cap (a,b)$.*

Proof. To show $F \in C([a,b])$, let $x_0 \in [a,b]$ and $\varepsilon > 0$. Since $\int_a^b f$ exists, there is an $M > \operatorname{lub}\{|f(x)| : a \le x \le b\}$. Choose $0 < \delta < \varepsilon/M$ and $x \in (x_0-\delta, x_0+\delta) \cap [a,b]$. Then

$$|F(x) - F(x_0)| = \left|\int_{x_0}^x f\right| \le M|x-x_0| < M\delta < \varepsilon$$

and $x_0 \in C(F)$.

Let $x_0 \in C(f) \cap (a,b)$ and $\varepsilon > 0$. There is a $\delta > 0$ such that $x \in (x_0-\delta, x_0+\delta) \subset (a,b)$ implies $|f(x) - f(x_0)| < \varepsilon$. If $0 < h < \delta$, then

$$\begin{aligned} \left|\frac{F(x_0+h)-F(x_0)}{h} - f(x_0)\right| &= \left|\frac{1}{h}\int_{x_0}^{x_0+h} f - f(x_0)\right| \\ &= \left|\frac{1}{h}\int_{x_0}^{x_0+h} (f(t)-f(x_0))\,dt\right| \\ &\le \frac{1}{h}\int_{x_0}^{x_0+h} |f(t)-f(x_0)|\,dt \\ &< \frac{1}{h}\int_{x_0}^{x_0+h} \varepsilon\,dt \\ &= \varepsilon. \end{aligned}$$

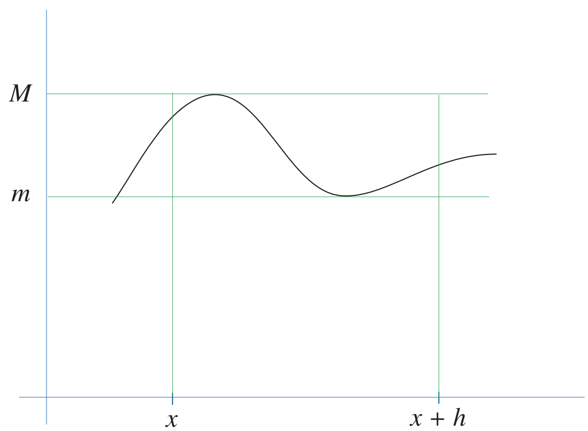
This shows $F'_+(x_0) = f(x_0)$. It can be shown in the same way that $F'_-(x_0) = f(x_0)$. Therefore $F'(x_0) = f(x_0)$. $\square$

The right picture makes Theorem 8.7.3 almost obvious. Suppose $x \in C(f)$ and $\varepsilon > 0$. There is a $\delta > 0$ such that

$$f((x-d, x+d) \cap [a,b]) \subset (f(x) - \varepsilon/2, f(x) + \varepsilon/2).$$

Let

$$m = \text{glb}\,\{fy : |x-y| < \delta\} \le \text{lub}\,\{fy : |x-y| < \delta\} = M.$$



Apparently $M - m < \varepsilon$ and for $0 < h < \delta$,

$$mh \le \int_x^{x+h} f \le Mh \implies m \le \frac{F(x+h) - F(x)}{h} \le M.$$

Since $M - m \to 0$ as $h \to 0$, a "squeezing" argument shows

$$\lim_{h \downarrow 0} \frac{F(x+h) - F(x)}{h} = f(x).$$

A similar argument establishes the limit from the left and $F'(x) = f(x)$.

It's easy to read too much into the Fundamental Theorem of Calculus. We are tempted to start thinking of integration and differentian as opposites of each other. But, this is far from the truth. The operations of integration and antidifferentiation are different operations, that happen to sometimes be tied together by the Fundamental Theorem of Calculus. Consider the following examples.

Example 8.7.2. Let

$$f(x) = \begin{cases} |x|/x, & x \ne 0 \\ 0, & x = 0 \end{cases}$$

It's easy to prove that f is integrable over any compact interval, and that $F(x) = \int_{-1}^{x} f = |x| - 1$ is an indefinite integral of f. But, F is not differentiable at $x = 0$ and f is not a derivative, according to Theorem 7.3.6.

Example 8.7.3. Let

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

It's straightforward to show that f is differentiable and

$$f'(x) = \begin{cases} 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Since f' is unbounded near $x = 0$, it follows from Theorem 8.2.2 that f' is not integrable over any interval containing 0.

Example 8.7.4. Let f be the salt and pepper function of Example 6.3.6. It was shown in Example 8.4.1 that $\int_a^b f = 0$ on any interval $[a, b]$. If $F(x) = \int_0^x f$, then $F(x) = 0$ for all x and $F' = f$ on $C(f) = \mathbb{Q}^c$.

8.8 Integral Mean Value Theorems

Theorem 8.8.1. *Suppose $f, g : [a, b] \to \mathbb{R}$ are such that*

(a) $g(x) \geq 0$ on $[a, b]$,

(b) f is bounded and $m \leq f(x) \leq M$ for all $x \in [a, b]$, and

(c) $\int_a^b f$ and $\int_a^b fg$ both exist.

There is a $c \in [m, M]$ such that

$$\int_a^b fg = c \int_a^b g.$$

Proof. Obviously,

$$m \int_a^b g \leq \int_a^b fg \leq M \int_a^b g. \tag{8.5}$$

If $\int_a^b g = 0$, we're done. Otherwise, let

$$c = \frac{\int_a^b fg}{\int_a^b g}.$$

Then $\int_a^b fg = c \int_a^b g$ and from (8.5), it follows that $m \leq c \leq M$. □

Corollary 8.8.2. *Let f and g be as in Theorem 8.8.1, but additionally assume f is continuous. Then there is a $c \in (a, b)$ such that*

$$\int_a^b fg = f(c) \int_a^b g.$$

Proof. This follows from Theorem 8.8.1 and Corollaries 6.5.4 and 6.5.7. □

Theorem 8.8.3. *Suppose* $f, g : [a, b] \to \mathbb{R}$ *are such that*

(a) $g(x) \geq 0$ *on* $[a, b]$,

(b) f *is bounded and* $m \leq f(x) \leq M$ *for all* $x \in [a, b]$, *and*

(c) $\int_a^b f$ *and* $\int_a^b fg$ *both exist.*

There is a $c \in [a, b]$ *such that*

$$\int_a^b fg = m \int_a^c g + M \int_c^b g.$$

Proof. For $a \leq x \leq b$ let

$$G(x) = m \int_a^x g + M \int_x^b g.$$

By Theorem 8.7.3, $G \in C([a, b])$ and

$$\inf G \leq G(b) = m \int_a^b g \leq \int_a^b fg \leq M \int_a^b g = G(a) \leq \sup G.$$

Now, apply Corollary 6.5.7 to find c where $G(c) = \int_a^b fg$. □