Advanced Calculus: MATH 410
Riemann Integrals and Integrability
Professor David Levermore
14 August 2019

9. Riemann Integrals

We now revisit the definite integral that was introduced to you when you first studied calculus. You undoubtedly learned that the definite integral of a positive function f over an interval $[a, b]$,

$$\int_a^b f(x)\,\mathrm{d}x\,,$$

provided it was defined, is a number equal to the area under the graph of f over $[a, b]$. You also may have learned that the definite integral is defined as a limit of Riemann sums. The Riemann sums you most likely saw were constructed by partitioning $[a, b]$ into n uniform subintervals of length $(b - a)/n$ and evaluating f at either the right-hand endpoint, the left-hand endpoint, or the midpoint of each subinterval. At the time your understanding of limits was likely more intuitive than rigorous. This chapter presents the *Riemann Integral*, a development of the definite integral built upon the rigorous notion of limit that we have now developed.

9.1. Partitions and Darboux Sums.

We will consider general partitions of the interval $[a, b]$, not just those with uniform subintervals.

Definition 9.1. *Let $[a, b] \subset \mathbb{R}$. A* partition *of the interval $[a, b]$ is specified by $n \in \mathbb{Z}_+$, and $\{x_i\}_{i=0}^n \subset [a, b]$ such that*

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b\,.$$

The partition P associated with these points is defined to be the ordered collection of subintervals of $[a, b]$ given by

$$P = \big([x_{i-1}, x_i] \,:\, i = 1, \cdots, n\big)$$

This partition is denoted $P = [x_0, x_1, \cdots, x_{n-1}, x_n]$. Each x_i for $i = 0, \cdots, n$ is called a partition point *of P, and for each $i = 1, \cdots, n$ the interval $[x_{i-1}, x_i]$ is called a i^{th}* subinterval *in P. The* partition thickness *or* width*, denoted $|P|$, is defined by*

$$|P| = \max\big\{x_i - x_{i-1} \,:\, i = 1, \cdots, n\big\}\,.$$

The approach to the definite integral taken here is not based on Riemann sums, but rather on Darboux sums. This is because Darboux sums are well-suited for analysis by the tools we have developed to establish the existence of limits. Results about Riemann sums will follow because every Riemann sum is bounded by two Darboux sums.

Let $f : [a, b] \to \mathbb{R}$ be bounded. Set

$$\underline{m} = \inf\big\{f(x) \,:\, x \in [a, b]\big\}\,, \qquad \overline{m} = \sup\big\{f(x) \,:\, x \in [a, b]\big\}\,. \tag{9.1}$$

Because f is bounded, we know that $-\infty < \underline{m} \le \overline{m} < \infty$.

Let $P = [x_0, \cdots, x_n]$ be a partition of $[a, b]$. For each $i = 1, \cdots, n$ set

$$\underline{m}_i = \underline{m}_i(f, P) = \inf\big\{f(x) \,:\, x \in [x_{i-1}, x_i]\big\}\,,$$
$$\overline{m}_i = \overline{m}_i(f, P) = \sup\big\{f(x) \,:\, x \in [x_{i-1}, x_i]\big\}\,.$$

Clearly $\underline{m} \le \underline{m}_i \le \overline{m}_i \le \overline{m}$ for every $i = 1, \cdots, n$. We will use the more cluttered $\underline{m}_i(f, P)$ and $\overline{m}_i(f, P)$ notation only when it is otherwise unclear which function or partition is involved.

Definition 9.2. *The* lower and upper Darboux sums *associated with the function f and the partition P are respectively defined by*

$$L(f,P) = \sum_{i=1}^{n} \underline{m}_i\,(x_i - x_{i-1})\,, \qquad U(f,P) = \sum_{i=1}^{n} \overline{m}_i\,(x_i - x_{i-1})\,. \tag{9.2}$$

Clearly, the Darboux sums satisfy the bounds

$$\underline{m}\,(b-a) \le L(f,P) \le U(f,P) \le \overline{m}\,(b-a)\,. \tag{9.3}$$

These inequalities will all be equalities when f is a constant.

Remark. A Riemann sum associated with the partition P is specified by picking a quadrature point $q_i \in [x_{i-1}, x_i]$ for each $i = 1, \cdots, n$. Let $Q = (q_1, \cdots, q_n)$ be the n-tuple of quadrature points. The associated Riemann sum is then

$$R(f,P,Q) = \sum_{i=1}^{n} f(q_i)\,(x_i - x_{i-1})\,.$$

The Riemann sums usually introduced in elementary calculus courses are given by the so-called *left-hand, right-hand,* and *midpoint* rules, which respectively pick $q_i \in [x_{i-1}, x_i]$ by

$$q_i = x_{i-1}\,, \qquad q_i = x_i\,, \qquad \text{and} \qquad q_i = \tfrac{1}{2}(x_{i-1} + x_i)\,.$$

For any choice of quadrature points Q we have the bounds

$$L(f,P) \le R(f,P,Q) \le U(f,P)\,. \tag{9.4}$$

Moreover, we can show that

$$\begin{aligned} L(f,P) &= \inf\big\{R(f,P,Q) \,:\, Q \text{ are quadrature points for } P\big\}\,, \\ U(f,P) &= \sup\big\{R(f,P,Q) \,:\, Q \text{ are quadrature points for } P\big\}\,. \end{aligned} \tag{9.5}$$

The bounds (9.4) are thereby sharp.

Exercise. Prove (9.5)

9.2. **Refinements.** We now introduce the notion of a refinement of a partition.

Definition 9.3. *Given a partition P of an interval $[a,b]$, a partition P^* of $[a,b]$ is called a* refinement *of P provided every partition point of P is a partition point of P^*.*

If $P = [x_0, x_1, \cdots, x_{n-1}, x_n]$ and P^* is a refinement of P then P^* induces a partition of each $[x_{i-1}, x_i]$, which we denote by P_i^*. For example, if $P^* = [x_0^*, x_1^*, \cdots, x_{n^*-1}^*, x_{n^*}^*]$ and j_i is defined for every $i = 0, \cdots, n$ by the relation $x_{j_i}^* = x_i$ then for every $i = 1, \cdots, n$ we have $P_i^* = [x_{j_{i-1}}^*, \cdots, x_{j_i}^*]$. Observe that

$$L(f,P^*) = \sum_{i=1}^{n} L(f,P_i^*)\,, \qquad U(f,P^*) = \sum_{i=1}^{n} U(f,P_i^*)\,. \tag{9.6}$$

Moreover, upon applying the bounds (9.3) to P_i^* for each $i = 1, \cdots, n$, we obtain the bounds

$$\underline{m}_i\,(x_i - x_{i-1}) \le L(f,P_i^*) \le U(f,P_i^*) \le \overline{m}_i\,(x_i - x_{i-1})\,. \tag{9.7}$$

This observation is key to the proof of the following.

Lemma 9.1. (Refinement) *Let $f : [a, b] \to \mathbb{R}$ be bounded. Let P be a partition of $[a, b]$ and P^* be a refinement of P. Then*

$$L(f, P) \leq L(f, P^*) \leq U(f, P^*) \leq U(f, P) \,. \tag{9.8}$$

Proof. It follows from (9.2), (9.7), and (9.6) that

$$\begin{aligned} L(f, P) &= \sum_{i=1}^{n} \underline{m}_i \, (x_i - x_{i-1}) \leq \sum_{i=1}^{n} L(f, P_i^*) = L(f, P^*) \\ &\leq U(f, P^*) = \sum_{i=1}^{n} U(f, P_i^*) \leq \sum_{i=1}^{n} \overline{m}_i \, (x_i - x_{i-1}) = U(f, P) \,. \end{aligned}$$

□

9.3. **Comparisons.** A key step in our development will be to develop comparisons of $L(f, P^1)$ and $U(f, P^2)$ for any two partitions P^1 and P^2, of $[a, b]$.

Definition 9.4. *Given any two partitions, P^1 and P^2, of $[a, b]$ we define $P^1 \vee P^2$ to be the partition whose set of partition points is the union of the partition points of P^1 and the partition points of P^2. We call $P^1 \vee P^2$ the* supremum *of P^1 and P^2.*

It is easy to argue that $P^1 \vee P^2$ is the smallest partition of $[a, b]$ that is a refinement of both P^1 and P^2. It is therefore sometimes called the *smallest common refinement* of P^1 and P^2.

Lemma 9.2. (Comparison) *Let $f : [a, b] \to \mathbb{R}$ be bounded. Let P^1 and P^2 be partitions of $[a, b]$. Then*

$$L(f, P^1) \leq U(f, P^2) \,. \tag{9.9}$$

Proof. Because $P^1 \vee P^2$ is a refinement of both P^1 and P^2, it follows from the Refinement Lemma that

$$L(f, P^1) \leq L(f, P^1 \vee P^2) \leq U(f, P^1 \vee P^2) \leq U(f, P^2) \,.$$

□

Because the partitions P^1 and P^2 on either side of inequality (9.9) are independent, we may obtain sharper bounds by taking the supremum over P^1 on the left-hand side, or the infimum over P^2 on the right-hand side. Indeed, we prove the following.

Lemma 9.3. (Sharp Comparison) *Let $f : [a, b] \to \mathbb{R}$ be bounded. Let*

$$\begin{aligned} \overline{L}(f) &= \sup \big\{ L(f, P) \,:\, P \text{ is a partition of } [a, b] \big\} \,, \\ \underline{U}(f) &= \inf \big\{ U(f, P) \,:\, P \text{ is a partition of } [a, b] \big\} \,. \end{aligned} \tag{9.10}$$

Let P^1 and P^2 be partitions of $[a, b]$. Then

$$L(f, P^1) \leq \overline{L}(f) \leq \underline{U}(f) \leq U(f, P^2) \,. \tag{9.11}$$

Moreover, if

$$L(f, P) \leq A \leq U(f, P) \quad \text{for every partition } P \text{ of } [a, b] \,,$$

then $A \in [\overline{L}(f), \underline{U}(f)]$.

Remark. Because it is clear from (9.10) that $\overline{L}(f)$ and $\underline{U}(f)$ depend on $[a,b]$, strictly speaking these quantities should be denoted $\overline{L}(f,[a,b])$ and $\underline{U}(f,[a,b])$. This would be necessary if more than one interval was involved in the discussion. However, that is not the case here. We therefore embrace the less cluttered notation.

Proof. If we take the infimum of the right-hand side of (9.9) over P^2, we obtain

$$L(f,P^1) \le \underline{U}(f)\,.$$

If we then take the supremum of the left-hand side above over P^1, we obtain

$$\overline{L}(f) \le \underline{U}(f)\,.$$

The bound (9.11) then follows.

The proof of the last assertion is left as an exercise. □

Exercise. Prove the last assertion of Lemma 9.3.

9.4. **Definition of the Riemann Integral.** We are now ready to define the definite integral of Riemann.

Definition 9.5. *Let $f:[a,b]\to\mathbb{R}$ be bounded. Then f is said to be* Riemann integrable *over $[a,b]$ whenever $\overline{L}(f)=\underline{U}(f)$. In this case we call this common value the* Riemann integral *of f over $[a,b]$ and denote it by $\int_a^b f$:*

$$\int_a^b f = \overline{L}(f) = \underline{U}(f)\,. \tag{9.12}$$

Then f is called the integrand *of the integral, a is called the* lower endpoint *(or* lower limit*) of integration, while b is called the* upper endpoint *(or* upper limit*) of integration.*

Remark. We will call a and b the endpoints of integration rather than the limits of integration. The word "limit" does enough work in this subject. We do not need to adopt terminology that can lead to confusion.

We begin with the following characterizations of integrability.

Theorem 9.1. (Riemann-Darboux) *Let $f:[a,b]\to\mathbb{R}$ be bounded. Then the following are equivalent:*

(1) *f is Riemann integrable over $[a,b]$ (i.e. $\overline{L}(f)=\underline{U}(f)$);*
(2) *for every $\epsilon>0$ there exists a partition P of $[a,b]$ such that*

$$0 \le U(f,P) - L(f,P) < \epsilon\,;$$

(3) *there exists a unique $A\in\mathbb{R}$ such that*

$$L(f,P) \le A \le U(f,P) \quad \text{for every partition } P \text{ of } [a,b]\,. \tag{9.13}$$

Moreover, in case (3) $A=\int_a^b f$.

Remark. Characterizations (2) and (3) of the Riemann-Darboux Theorem are useful for proving the integrability of a function f. The Sharp Comparison Lemma shows that (9.13) holds if and only if $A\in[\overline{L}(f),\underline{U}(f)]$. The key thing to be established when using characterization (3) is therefore the uniqueness of such an A.

Proof. First we show that (1) $\Longrightarrow$ (2). Let $\epsilon > 0$. By the definition (9.10) of $\overline{L}(f)$ and $\underline{U}(f)$, we can find partitions P^L and P^U of $[a,b]$ such that

$$\overline{L}(f) - \frac{\epsilon}{2} < L(f,P^L) \le \overline{L}(f)\,, \qquad \underline{U}(f) \le U(f,P^U) < \underline{U}(f) + \frac{\epsilon}{2}\,.$$

Let $P = P^L \vee P^U$. Because the Comparison Lemma implies that $L(f,P^L) \le L(f,P)$ and $U(f,P) \le U(f,P^U)$, it follows from the above inequalities that

$$\overline{L}(f) - \frac{\epsilon}{2} < L(f,P) \le \overline{L}(f)\,, \qquad \underline{U}(f) \le U(f,P) < \underline{U}(f) + \frac{\epsilon}{2}\,.$$

Hence, if $\overline{L}(f) = \underline{U}(f)$ we conclude that

$$0 \le U(f,P) - L(f,P) < \left(\underline{U}(f) + \frac{\epsilon}{2}\right) - \left(\overline{L}(f) - \frac{\epsilon}{2}\right) = \epsilon\,.$$

This shows that (1) $\Longrightarrow$ (2).

Next we show that (2) $\Longrightarrow$ (3). Suppose that (3) is false. The Sharp Comparison Lemma shows that (9.13) holds for every $A \in [\overline{L}(f), \underline{U}(f)]$, and that this interval is nonempty. So the only way (3) can be false is if uniqueness fails. In that case there exists A_1 and A_2 such that

$$L(f,P) \le A_1 < A_2 \le U(f,P) \quad \text{for every partition } P \text{ of } [a,b]\,.$$

We thereby have that

$$U(f,P) - L(f,P) \ge A_2 - A_1 > 0 \quad \text{for every partition } P \text{ of } [a,b]\,.$$

Hence, (2) must be false. It follows that (2) $\Longrightarrow$ (3).

Finally, we show that (3) $\Longrightarrow$ (1) and that (3) implies $A = \int_a^b f$. The Sharp Comparison Lemma shows that (9.13) holds if and only if $A \in [\overline{L}(f), \underline{U}(f)]$. But (3) states that such an A is unique. Hence, $A = \overline{L}(f) = \underline{U}(f)$, which implies (1) and $A = \int_a^b f$. $\square$

9.5. **Convergence of Riemann and Darboux Sums.** We now make a connection with the notion of a definite integral as the limit of a sequence of Riemann sums.

Recall for any given $f : [a,b] \to \mathbb{R}$ a Riemann sum associated with a partition $P = [x_0, x_1 \cdots, x_n]$ of $[a,b]$ is specified by selecting a quadrature point $q_i \in [x_{i-1}, x_i]$ for each $i = 1, \cdots, n$. Let $Q = (q_1, \cdots, q_n)$ be the n-tuple of quadrature points. The associated Riemann sum is then

$$R(f,P,Q) = \sum_{i=1}^{n} f(q_i)\,(x_i - x_{i-1})\,. \tag{9.14}$$

If $f : [a,b] \to \mathbb{R}$ is bounded (so that the Darboux sums $L(f,P)$ and $U(f,P)$ are defined) then for any choice of quadrature points Q we have the bounds

$$L(f,P) \le R(f,P,Q) \le U(f,P)\,. \tag{9.15}$$

A sequence of Riemann sums for any given $f : [a,b] \to \mathbb{R}$ is therefore specified by a sequence $\{P^n\}_{n=1}^\infty$ of partitions of $[a,b]$ and a sequence $\{Q^n\}_{n=1}^\infty$ of associated quadrature points. The sequence of partitions cannot be arbitrary.

Definition 9.6. *Let $f : [a,b] \to \mathbb{R}$ be bounded. A sequence $\{P^n\}_{n=1}^\infty$ of partitions of $[a,b]$ is said to be* Archimedean *for f provided*

$$\lim_{n\to\infty} \big(U(f,P^n) - L(f,P^n)\big) = 0\,. \tag{9.16}$$

Our main theorem is the following.

Theorem 9.2. (Archimedes-Riemann) *Let $f : [a,b] \to \mathbb{R}$ be bounded. Then f is Riemann integrable over $[a,b]$ if and only if there exists a sequence of partitions of $[a,b]$ that is Archimedean for f. If $\{P^n\}_{n=1}^{\infty}$ is any such sequence then*

$$\lim_{n\to\infty} L(f,P^n) = \int_a^b f\,, \quad \text{and} \quad \lim_{n\to\infty} U(f,P^n) = \int_a^b f\,. \tag{9.17}$$

Moreover, if $\{Q^n\}_{n=1}^{\infty}$ is any sequence of associated quadrature points then

$$\lim_{n\to\infty} R(f,P^n,Q^n) = \int_a^b f\,, \tag{9.18}$$

where the Riemann sums $R(f,P,Q)$ are defined by (9.14).

Remark. The content of this theorem is that once we have found a sequence of partitions P^n such that (9.16) holds, then the integral $\int_a^b f$ exists and may be evaluated as the limit of any associated sequence of Darboux sums (9.17) or Riemann sums (9.18). This theorem thereby splits the task of evaluating a definite integrals into two steps. The first step is by far the easier. It is a rare integrand f for which we can find a sequence of Darboux or Riemann sums that allows one of the limits (9.17) or (9.18) to be evaluated directly.

Proof. If f is Riemann integrable over $[a,b]$ then we can use characterization (2) of the Riemann-Darboux Theorem to construct a sequence of partitions that satisfies (9.16), and is thereby Archimedean for f. Conversely, if you are given a sequence of partitions of $[a,b]$ that is Archimedean for f then the fact that f is integrable over $[a,b]$ follows directly from characterization (2) of the Riemann-Darboux Theorem. The details of these arguments are left as an exercise.

We now establish the limits (9.17) and (9.18). Let $\{P^n\}_{n=1}^{\infty}$ be a sequence of partitions of $[a,b]$ that is Archimedean for f and let $\{Q^n\}_{n=1}^{\infty}$ be a sequence of associated quadrature points. The bounds on Riemann sums given by (9.15) yield the inequalities

$$L(f,P^n) \le R(f,P^n,Q^n) \le U(f,P^n)\,,$$

while, because f is Riemann integrable, we also have the inequalities

$$L(f,P^n) \le \int_a^b f \le U(f,P^n)\,.$$

It follows from these inequalities that

$$\begin{aligned}
L(f,P^n) - U(f,P^n) &\le L(f,P^n) - \int_a^b f \\
&\le R(f,P^n,Q^n) - \int_a^b f \\
&\le U(f,P^n) - \int_a^b f \le U(f,P^n) - L(f,P^n)\,,
\end{aligned}$$

which implies that

$$\left| L(f, P^n) - \int_a^b f \right| \leq U(f, P^n) - L(f, P^n) ,$$
$$\left| R(f, P^n, Q^n) - \int_a^b f \right| \leq U(f, P^n) - L(f, P^n) ,$$
$$\left| U(f, P^n) - \int_a^b f \right| \leq U(f, P^n) - L(f, P^n) .$$

Because $\{P^n\}_{n=1}^{\infty}$ is Archimedean for f, it satisfies (9.16), whereby the right-hand sides above vanish as n tends to ∞. The limits (9.17) and (9.18) follow. $\square$

9.6. **Partitions Lemma.** We now prove a lemma, essentially due to Darboux, that will be used to give new criteria for a function to be Riemann integrable, and to give a simple criterion for a sequence of partitions to be Archimedean for every Riemann integrable function.

Lemma 9.4. (Partitions) *Let $f : [a,b] \to \mathbb{R}$ be bounded. Then for every $\epsilon > 0$ there exists a $\delta > 0$ such that for every partition P of $[a,b]$ we have*

$$|P| < \delta \quad \Longrightarrow \quad \begin{cases} 0 \leq \overline{L}(f) - L(f, P) < \epsilon , \\ 0 \leq U(f, P) - \underline{U}(f) < \epsilon . \end{cases} \tag{9.19}$$

Proof. Let $\epsilon > 0$. There exist partitions P_L^ϵ and P_U^ϵ of $[a,b]$ such that

$$0 \leq \overline{L}(f) - L(f, P_L^\epsilon) < \frac{\epsilon}{2} , \qquad 0 \leq U(f, P_U^\epsilon) - \underline{U}(f) < \frac{\epsilon}{2} .$$

Let $P^\epsilon = P_L^\epsilon \vee P_U^\epsilon$. Let n^ϵ be the number of subintervals in P^ϵ. Pick $\delta > 0$ such that

$$n^\epsilon 2M\delta < \frac{\epsilon}{2} , \quad \text{where} \quad M = \sup\big\{ |f(x)| \,:\, x \in [a,b] \big\} .$$

We will show that (9.19) holds for this δ.

Now let $P = [x_0, x_1, \cdots, x_n]$ be an arbitrary partition of $[a,b]$ such that $|P| < \delta$. Set $P^* = P \vee P^\epsilon$. We consider

$$\begin{aligned} 0 \leq \overline{L}(f) - L(f, P) &= \big(\overline{L}(f) - L(f, P^*)\big) + \big(L(f, P^*) - L(f, P)\big) , \\ 0 \leq U(f, P) - \underline{U}(f) &= \big(U(f, P^*) - \underline{U}(f)\big) + \big(U(f, P) - U(f, P^*)\big) . \end{aligned} \tag{9.20}$$

We will prove the theorem by showing that each of the four terms in parentheses on the right-hand sides above is less than $\epsilon/2$.

Because P^* is a refinement of P^ϵ, which is a refinement of both P_L^ϵ and P_U^ϵ, the Refinement Lemma implies that

$$0 \leq \overline{L}(f) - L(f, P^*) \leq \overline{L}(f) - L(f, P^\epsilon) \leq \overline{L}(f) - L(f, P_L^\epsilon) < \frac{\epsilon}{2} ,$$
$$0 \leq U(f, P^*) - \underline{U}(f) \leq U(f, P^\epsilon) - \underline{U}(f) \leq U(f, P_U^\epsilon) - \underline{U}(f) < \frac{\epsilon}{2} .$$

Thus, the first terms on the right-hand sides of (9.20) are less than $\epsilon/2$.

Because P^* is a refinement of P, for each $i = 1, \cdots, n$ we let P_i^* denote the partition of $[x_{i-1}, x_i]$ induced by P^*. The Refinement Lemma then yields

$$0 \leq L(f, P^*) - L(f, P) = \sum_{i=1}^{n} \left[L(f, P_i^*) - \underline{m}_i(x_i - x_{i-1}) \right],$$

$$0 \leq U(f, P) - U(f, P^*) = \sum_{i=1}^{n} \left[\overline{m}_i(x_i - x_{i-1}) - U(f, P_i^*) \right].$$

Because P^ϵ has at most $n^\epsilon - 1$ partition points that are not partition points of P, there are at most $n^\epsilon - 1$ indices i for which $[x_{i-1}, x_i]$ contains at least one partition point of P_i^* within (x_{i-1}, x_i). For all other indices the terms in the above sums are zero. Each of the nonzero terms in the above sums satisfy the bounds

$$0 \leq L(f, P_i^*) - \underline{m}_i(x_i - x_{i-1}) \leq 2M(x_i - x_{i-1}) < 2M\delta,$$
$$0 \leq \overline{m}_i(x_i - x_{i-1}) - U(f, P_i^*) \leq 2M(x_i - x_{i-1}) < 2M\delta.$$

Because there are at most $n^\epsilon - 1$ such terms, we obtain the bounds

$$0 \leq L(f, P^*) - L(f, P) < n^\epsilon 2M\delta < \frac{\epsilon}{2},$$
$$0 \leq U(f, P) - U(f, P^*) < n^\epsilon 2M\delta < \frac{\epsilon}{2}.$$

This shows the second terms on the right-hand sides of (9.20) are each less than $\epsilon/2$, thereby completing the proof of the lemma. □

An immediate consequence of the Partitions Lemma is the following characterization of Riemann integrable functions due to Darboux.

Theorem 9.3. (Darboux) *Let $f : [a, b] \to \mathbb{R}$ be bounded. Then f is Riemann integrable over $[a, b]$ if and only if for every $\epsilon > 0$ there exists a $\delta > 0$ such that for every partition P of $[a, b]$ we have*

$$|P| < \delta \quad \Longrightarrow \quad 0 \leq U(f, P) - L(f, P) < \epsilon.$$

Proof. The direction ($\Longleftarrow$) follows from the Riemann-Darboux Theorem. The direction ($\Longrightarrow$) follows from the definition of the Riemann integral and the Partitions Lemma. The details of the proof are left as an exercise. □

An immediate consequence of the Darboux Theorem is that there is a simple criterion for a sequence of partitions to be Archimedean for every Riemann integrable function.

Theorem 9.4. (Archimedean Sequence) *Every sequence $\{P^n\}_{n=1}^\infty$ of partitions of $[a, b]$ such that $|P^n| \to 0$ as $n \to \infty$ is Archimedean for every function $f : [a, b] \to \mathbb{R}$ that is Riemann integrable over $[a, b]$.*

Proof. Exercise. □

Remark. A sequence of partitions $\{P^n\}_{n=1}^\infty$ does not need to satisfy the condition $|P^n| \to 0$ as $n \to \infty$ in order to be Archimedean for a given function. For example, every sequence of partitions is Archimedean for every constant function.

The Partitions Lemma also leads to another criterion for a function to be Riemann integrable. The first step towards this criterion is taken by the following lemma.

Lemma 9.5. *Let $f : [a,b] \to \mathbb{R}$ be bounded. Let $\{P^n\}_{n=1}^{\infty}$ be any sequence of partitions of $[a,b]$ such that $|P^n| \to 0$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} L(f,P^n) = \overline{L}(f)\,, \qquad \lim_{n\to\infty} U(f,P^n) = \underline{U}(f)\,. \tag{9.21}$$

Moreover, there exist sequences $\{Q_L^n\}_{n=1}^{\infty}$ and $\{Q_U^n\}_{n=1}^{\infty}$ of associated quadrature points such that

$$\lim_{n\to\infty} R(f,P^n,Q_L^n) = \overline{L}(f)\,, \qquad \lim_{n\to\infty} R(f,P^n,Q_U^n) = \underline{U}(f)\,. \tag{9.22}$$

Proof. The proof of (9.21) follows from the Partitions Lemma. To prove (9.22), it follows from (9.5) that for each partition P^n there exist sets of associated quadrature points Q_L^n and Q_U^n such that

$$0 \le R(f,P^n,Q_L^n) - L(f,P^n) < \frac{1}{2^n}\,, \qquad 0 \le U(f,P^n) - R(f,P^n,Q_U^n) < \frac{1}{2^n}\,.$$

The details of the proof are left as an exercise. □

An immediate consequence of Lemma 9.5 is the following criterion for a function to be Riemann integrable that is stated in terms of limits of Riemann sums associated with a single sequence of partitions.

Theorem 9.5. *Let $f : [a,b] \to \mathbb{R}$ be bounded. Let $\{P^n\}_{n=1}^{\infty}$ be any sequence of partitions of $[a,b]$ such that $|P^n| \to 0$ as $n \to \infty$. Suppose that there exists a unique $A \in \mathbb{R}$ such that for every sequence $\{Q^n\}_{n=1}^{\infty}$ of associated quadrature points we have*

$$\lim_{n\to\infty} R(f,P^n,Q^n) = A\,. \tag{9.23}$$

Then f is Riemann integrable over $[a,b]$ and

$$\int_a^b f = A\,. \tag{9.24}$$

Proof. Exercise. □

9.7. **Power Rule*.** In this section we derive the so-called power rule for definite integrals — specifically, that for any $p \in \mathbb{R}$ and any $[a,b] \subset \mathbb{R}_+$ we have

$$\int_a^b x^p = \begin{cases} \dfrac{b^{p+1} - a^{p+1}}{p+1} & \text{for } p \neq -1\,, \\ \log\!\left(\dfrac{b}{a}\right) & \text{for } p = -1\,. \end{cases} \tag{9.25}$$

Of course, this rule follows easily from the Fundamental Theorem of Calculus. Here we will derive it by taking limits of Riemann sums.

Because for every $p \in \mathbb{R}$ the power function $x \mapsto x^p$ is bounded over every $[a,b] \subset \mathbb{R}_+$, Theorem 9.4 says that a sequence $\{P^n\}_{n=1}^{\infty}$ of partitions of $[a,b]$ will be Archimedean whenever $|P^n| \to 0$ as $n \to \infty$. The problem is thereby reduced to finding such a sequence of partitions and a sequence $\{Q^n\}_{n=1}^{\infty}$ of quadrature sets for which we can show that

$$\lim_{n\to\infty} R(x^p,P^n,Q^n) = \begin{cases} \dfrac{b^{p+1} - a^{p+1}}{p+1} & \text{for } p \neq -1\,, \\ \log\!\left(\dfrac{b}{a}\right) & \text{for } p = -1\,. \end{cases}$$

We will take two approaches to this problem. There are more.

9.7.1. *Uniform Partitions.* Whenever $p \geq 0$ it is clear that the function $x \mapsto x^p$ is Riemann integrable over $[0, b]$. If we use the uniform partitions over $[0, b]$ given by

$$P^n = [x_0, x_1, \cdots, x_n], \qquad x_i = \frac{ib}{n},$$

and the right-hand rule quadrature sets $Q^n = (x_1, \cdots, x_n)$ then

$$R(x^p, P^n, Q^n) = \frac{b}{n} \sum_{i=1}^{n} \left(\frac{ib}{n} \right)^p = \frac{b^{p+1}}{n^{p+1}} S^p(n),$$

where

$$S^p(n) = \sum_{i=1}^{n} i^p.$$

Therefore we must show that for every $p \geq 0$ we have

$$\int_0^b x^p = \lim_{n \to \infty} \frac{b^{p+1}}{n^{p+1}} S^p(n) = \frac{b^{p+1}}{p+1}. \tag{9.26}$$

Once this is done then for every $[a, b] \subset (0, \infty)$ we have

$$\int_a^b x^p = \int_0^b x^p - \int_0^a x^p = \frac{b^{p+1} - a^{p+1}}{p+1},$$

which agrees with (9.25).

In order to prove (9.26) we must establish the limit

$$\lim_{n \to \infty} \frac{1}{n^{p+1}} S^p(n) = \frac{1}{p+1}. \tag{9.27}$$

The details of proving (9.27) are presented in the book for the cases $p = 0, 1, 2$ with $b = 1$. Most calculus books prove this limit for cases when p is a natural number no higher than $p = 3$. They usually proceed by first establishing formulas for $S^p(n)$ like

$$S^0(n) = n, \qquad S^1(n) = \frac{n(n+1)}{2},$$
$$S^2(n) = \frac{n(n+1)(2n+1)}{6}, \qquad S^3(n) = \frac{n^2(n+1)^2}{4}.$$

The first of these formulas is trivial. The others are typically verified by an induction argument on n. However, this approach does not give any insight into how to obtain these formulas, which grow in complexity as p increases. Alternatively, such formulas can be derived by a constructive induction arguement on n by assuming only that $S^p(n)$ is a polynomial in n of degree $p + 1$. Given such an explicit formula for $S^p(n)$, establishing (9.27) is easy.

Proof. Here we will take a different approach that allows us to prove (9.27) for every $p \in \mathbb{N}$. We first show that for every $p \in \mathbb{N}$ and every $n \in \mathbb{N}$ with $n \geq 2$ we have the bounds

$$S^p(n-1) \leq \frac{n^{p+1}}{p+1} \leq S^p(n). \tag{9.28}$$

These imply that

$$\frac{1}{p+1} \leq \frac{1}{n^{p+1}} S^p(n) \leq \frac{1}{p+1} \left(\frac{n+1}{n} \right)^{p+1},$$

from which the limit (9.27) follows easily for every $p \in \mathbb{N}$.

The bounds (9.28) are proved by writing $n^{p+1}/(p+1)$ as the telescoping sum

$$\frac{n^{p+1}}{p+1} = \sum_{i=0}^{n-1} \frac{(i+1)^{p+1} - i^{p+1}}{p+1} \,.$$

By the binomial expansion we have

$$\frac{(i+1)^{p+1} - i^{p+1}}{p+1} = \frac{1}{p+1} \sum_{k=0}^{p} \frac{(p+1)!}{(p+1-k)!k!}\, i^k = \sum_{k=0}^{p} \frac{p!}{(p+1-k)!k!}\, i^k \,.$$

The last sum is bounded below by its $k = p$ term, which is i^p. Again using the binomial expansion, we see that the last sum is bounded above by

$$(i+1)^p = \sum_{k=0}^{p} \frac{p!}{(p-k)!k!}\, i^k \,.$$

Therefore we have the bounds

$$i^p \leq \frac{(i+1)^{p+1} - i^{p+1}}{p+1} \leq (i+1)^p \,.$$

These yield the bounds (9.28) upon summing from $i = 0$ to $n - 1$. □

Remark. The place in the above proof that required p to be a natrual number was the point where we used the binomial formula. This restriction will be removed in the next subsection.

Remark. We can express $S^p(n)$ in terms of all the $S^k(n)$ with $k = 0, \cdots, p-1$. By a telescoping sum, the binomial expansion, and the definition of $S^p(n)$, we obtain the identity

$$\begin{aligned}
(n+1)^{p+1} - 1 &= \sum_{i=1}^{n} \left[(i+1)^{p+1} - i^{p+1}\right] = \sum_{i=1}^{n} \sum_{k=0}^{p} \frac{(p+1)!}{k!(p-k+1)!}\, i^k \\
&= \sum_{k=0}^{p} \frac{(p+1)!}{k!(p-k+1)!}\, S^k(n) = (p+1)\, S^p(n) + \sum_{k=0}^{p-1} \frac{(p+1)!}{k!(p-k+1)!}\, S^k(n) \,.
\end{aligned}$$

Upon solving for $S^p(n)$, we obtain the relation

$$S^p(n) = \frac{1}{p+1} \left[(n+1)^{p+1} - 1 - \sum_{k=0}^{p-1} \frac{(p+1)!}{k!(p-k+1)!}\, S^k(n) \right] . \tag{9.29}$$

Exercise. Relation (9.29) can be used to generate explicit formulas for $S^p(n)$ for any $p \geq 1$. To get an idea of how complicated these explicit formulas become, start with the fact $S^0(n) = n$ and use relation (9.29) to generate explicit formulas for $S^1(n)$, $S^2(n)$, $S^3(n)$, and $S^4(n)$.

9.7.2. *Nonuniform Partitions.* The difficulty with the approach using uniform partitions was that the resulting Riemann sums could not be evaluated easily. Fermat saw that this difficulty can be elegantly overcome by using the nonuniform partitions over $[a, b] \subset (0, \infty)$ given by

$$P^n = [x_0, x_1, \cdots, x_n] \,, \qquad x_i = a \left(\frac{b}{a}\right)^{\frac{i}{n}} .$$

By introducing $r_n = (b/a)^{\frac{1}{n}}$, the partition points can be expressed as $x_i = a\, r_n^i$. If we use the left-hand rule quadrature sets $Q^n = (x_0, \cdots, x_{n-1})$ then

$$R(x^p, P^n, Q^n) = \sum_{i=0}^{n-1} \big(ar_n^i\big)^p \big(ar_n^{i+1} - ar_n^i\big) = a^{p+1}(r_n - 1) \sum_{i=0}^{n-1} r_n^{i(p+1)} .$$

Notice that the last sum is a finite geometric series with ratio $r_n^{(p+1)}$. Therefore it can be evaluated as

$$\sum_{i=0}^{n-1} r_n^{i(p+1)} = \begin{cases} \dfrac{r_n^{n(p+1)} - 1}{r_n^{(p+1)} - 1} & \text{for } p \neq -1\,, \\ n & \text{for } p = -1\,. \end{cases}$$

When $p \neq -1$ the Riemann sums are thereby evaluated as

$$R(x^p, P^n, Q^n) = a^{p+1}(r_n - 1)\, \frac{r_n^{n(p+1)} - 1}{r_n^{(p+1)} - 1} = \big(b^{p+1} - a^{p+1}\big)\, \frac{r_n - 1}{r_n^{(p+1)} - 1} .$$

Here we have used the fact that $r_n^n = b/a$ to see that

$$a^{p+1}\big(r_n^{n(p+1)} - 1\big) = b^{p+1} - a^{p+1} .$$

Given the explicit formula for $R(x^p, P^n, Q^n)$ found above, we only need to show that

$$\lim_{n\to\infty} \frac{r_n - 1}{r_n^{(p+1)} - 1} = \frac{1}{p+1} . \tag{9.30}$$

Then

$$\lim_{n\to\infty} R(x^p, P^n, Q^n) = \big(b^{p+1} - a^{p+1}\big) \lim_{n\to\infty} \frac{r_n - 1}{r_n^{(p+1)} - 1} = \frac{b^{p+1} - a^{p+1}}{p+1} ,$$

which yields (9.25) for the case $p \neq -1$. The case $p = -1$ is left as an exercise. □

Remark. Fermat discovered this beautiful derivation of the power rule (9.25) for the case $p \neq -1$ before Newton and Leibniz developed the fundamental theorems of calculus. He evaluated limit (9.30) for $p \in \mathbb{Q}$ and $p \neq -1$ without using the l'Hospital rule. In other words, there was no "easy way" to do the problem when Fermat discovered the power rule. It took a genius like Fermat to solve a problem that the "easy way" now makes routine. In fact, Fermat's power rule provided a crucial clue that led to the development of the "easy way" by Newton, Leibniz, and others.

Exercise. Prove the limit (9.30) when $p \neq -1$.

Exercise. Prove (9.25) for the case $p = -1$. This requires proving that

$$\lim_{n\to\infty} n(r_n - 1) = \log\left(\frac{b}{a}\right) .$$

Exercise. By taking limits of Riemann sums, show for every positive a and b that

$$\int_0^b a^x = \frac{a^b - 1}{\log(a)} .$$

Hint: Use uniform partitions.

10. Riemann Integrable Functions

In the previous chapter we defined the Riemann integral and gave some characterizations of Riemann integrable functions. However, we did not use those characterizations to identify a large class of Riemann integrable functions. That is what we will do in this chapter. Before beginning that task, we remark that there are many functions that are not Riemann integrable.

Exercise. Let f be the function

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}\,, \\ 0 & \text{if } x \notin \mathbb{Q}\,. \end{cases}$$

Show that the restriction of f to any closed bounded interval $[a,b]$ with $a < b$ is not Riemann integrable.

10.1. **Integrability of Monotonic Functions.**

We first show that the class of Riemann integrable functions includes the class of monotonic functions. Recall that this class is defined as follows.

Definition 10.1. *Let $D \subset \mathbb{R}$. A function $f : D \to \mathbb{R}$ is said to be* nondecreasing *over D provided that*

$$x < y \implies f(x) \le f(y) \quad \textit{for every } x, y \in D\,.$$

A function $f : D \to \mathbb{R}$ is said to be nonincreasing *over D provided that*

$$x < y \implies f(x) \ge f(y) \quad \textit{for every } x, y \in D\,.$$

A function that is either nondecreasing or nonincreasing is said to be monotonic *over D.*

A function that is monotonic over a closed bounded interval $[a,b]$ is clearly bounded by its endpoint values.

Theorem 10.1. (Monotonic Integrability) *Let $f : [a,b] \to \mathbb{R}$ be monotonic. Then f is Riemann integrable over $[a,b]$. Moreover, for every partition P of $[a,b]$ we have*

$$0 \le U(f,P) - L(f,P) \le |P|\,\big|f(b) - f(a)\big|\,. \tag{10.1}$$

Proof. Given (10.1), it is easy to prove that f is Riemann integrable over $[a,b]$. Indeed, let $\epsilon > 0$. Let P be any partition of $[a,b]$ such that $|P|\,|f(b) - f(a)| < \epsilon$. Then by (10.1) we have

$$0 \le U(f,P) - L(f,P) \le |P|\,\big|f(b) - f(a)\big| < \epsilon\,.$$

Hence, f is Riemann integrable by characterization (2) of the Riemann-Darboux Theorem.

All that remains to be done is to prove (10.1). For any partition $P = [x_0, \cdots, x_n]$ we have the following basic estimate. Because f is monotonic, over each subinterval $[x_{i-1}, x_i]$ we have that

$$\overline{m}_i - \underline{m}_i = \big|f(x_i) - f(x_{i-1})\big|\,.$$

We thereby obtain

$$\begin{aligned} 0 \le U(f,P) - L(f,P) &= \sum_{i=1}^{n} (\overline{m}_i - \underline{m}_i)\,(x_i - x_{i-1}) \\ &\le |P| \sum_{i=1}^{n} (\overline{m}_i - \underline{m}_i) = |P| \sum_{i=1}^{n} \big|f(x_i) - f(x_{i-1})\big|\,, \end{aligned}$$

where $|P| = \max\{x_i - x_{i-1} \,:\, i = 1, \cdots, n\}$ is the thickness of P. Because f is monotonic, the terms $f(x_i) - f(x_{i-1})$ are either all nonnegative, or all nonpositive. We may therefore pass the absolute value outside the last sum above, which then telescopes. We thereby obtain the bound

$$\begin{aligned} 0 \leq U(f,P) - L(f,P) &\leq |P| \sum_{i=1}^{n} \big| f(x_i) - f(x_{i-1}) \big| \\ &= |P| \left| \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1}) \big) \right| = |P| \, \big| f(b) - f(a) \big| \,. \end{aligned}$$

This establishes (10.1), and thereby proves the theorem. □

Remark. It is a classical fact that a monotonic function over $[a, b]$ is continuous at all but at most a countable number of points where it has a jump discontinuity. One example of such a function defined over the interval $[0, 1]$ is

$$f(x) = \begin{cases} \dfrac{1}{2^k} & \text{for } \dfrac{1}{2^{k+1}} < x \leq \dfrac{1}{2^k} \,, \\ 0 & \text{for } x = 0 \,. \end{cases}$$

We can show that

$$\int_0^1 f = \frac{2}{3} \,.$$

10.2. **Integrability of Continuous Functions.** The class of Riemann integrable functions also includes the class of continuous functions.

Theorem 10.2. (Continuous Integrability) *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is Riemann integrable over $[a, b]$.*

Remark. The fact that f is continuous over $[a, b]$ implies that it is bounded over $[a, b]$ and that it is uniformly continuous over $[a, b]$. The fact f is bounded is needed to know that the Darboux sums $L(f, P)$ and $U(f, P)$ make sense for any partition P. The fact f is uniformly continuous will play the central role in our proof.

Proof. Let $\epsilon > 0$. Because f is uniformly continuous over $[a, b]$, there exists a $\delta > 0$ such that

$$|x - y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{b - a} \quad \text{for every } x, y \in [a, b] \,.$$

Let $P = [x_0, x_1, \cdots, x_n]$ be any partition of $[a, b]$ such that $|P| < \delta$. Because f is continuous, it takes on extreme values over each subinterval $[x_{i-1}, x_i]$ of P. Hence, for every $i = 1, \cdots, n$ there exist points $\overline{x}_i$ and $\underline{x}_i$ in $[x_{i-1}, x_i]$ such that $\overline{m}_i = f(\overline{x}_i)$ and $\underline{m}_i = f(\underline{x}_i)$. Because $|P| < \delta$ it follows that $|\overline{x}_i - \underline{x}_i| < \delta$, whereby

$$\overline{m}_i - \underline{m}_i = f(\overline{x}_i) - f(\underline{x}_i) < \frac{\epsilon}{b - a} \,.$$

We thereby obtain

$$\begin{aligned} 0 \leq U(f,P) - L(f,P) &= \sum_{i=1}^{n} (\overline{m}_i - \underline{m}_i) \, (x_i - x_{i-1}) \\ &\leq \frac{\epsilon}{b - a} \sum_{i=1}^{n} (x_i - x_{i-1}) = \frac{\epsilon}{b - a} \, (b - a) = \epsilon \,. \end{aligned}$$

This shows that for every partition P of $[a,b]$ we have

$$|P| < \delta \quad \Longrightarrow \quad 0 \leq U(f,P) - L(f,P) < \epsilon\,.$$

But $\epsilon > 0$ was arbitrary. It follows that f is Riemann integrable by the Darboux Theorem (Theorem 9.3). □

Exercise. A function $f : [a,b] \to \mathbb{R}$ is said to be Hölder continuous of order $\alpha \in (0,1]$ if there exists a $C \in \mathbb{R}_+$ such that for every $x, y \in [a,b]$ we have

$$|f(x) - f(y)| \leq C\,|x-y|^{\alpha}\,.$$

Show that for every partition P of $[a,b]$ we have

$$0 \leq U(f,P) - L(f,P) \leq |P|^{\alpha} C\,(b-a)\,.$$

10.3. **Linearity and Order for Riemann Integrals.** Linear combinations of Riemann integrable functions are again Riemann integrable. Riemann integrals respect linearity and order.

10.3.1. *Linearity.* One basic fact about Riemann integrals is that they depend linearly on the integrand. This fact is not completely trivial because we defined the Riemann integral through Darboux sums, which do not depend linearly on the integrand.

Proposition 10.1. (Linearity) *Let $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ be Riemann integrable over $[a,b]$. Let $\alpha \in \mathbb{R}$. Then $f+g$ and αf are Riemann integrable over $[a,b]$ with*

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

Proof. A key step towards establishing the additivity is to prove that if P is any partition of $[a,b]$ then

$$L(f,P) + L(g,P) \leq L\big((f+g),P\big) \leq U\big((f+g),P\big) \leq U(f,P) + U(g,P)\,.$$

A key step towards establishing the scalar multiplicity is to prove that if $\alpha > 0$ and P is any partition of $[a,b]$ then

$$L(\alpha f,P) = \alpha L(f,P)\,, \qquad U(\alpha f,P) = \alpha U(f,P)\,.$$

The proof is left as an exercise. □

Remark. It follows immediately from the above proposition that every linear combination of Riemann integrable functions is also Riemann integrable, and that its integral is the same linear combination of the respective integrals. More precisely, if $f_k : [a,b] \to \mathbb{R}$ is Riemann integrable over $[a,b]$ for every $k = 1, 2, \cdots, n$ then for every $\{\alpha_k\}_{k=1}^n \subset \mathbb{R}$ we know that

$$\sum_{k=1}^n \alpha_k f_k \quad \text{is Riemann integrable over } [a,b]\,,$$

with

$$\int_a^b \left(\sum_{k=1}^n \alpha_k f_k\right) = \sum_{k=1}^n \alpha_k \int_a^b f_k\,.$$

10.3.2. *Nonnegativity and Order.* Another basic fact about definite integrals is that they respect nonnegativity of the integrand.

Proposition 10.2. (Nonnegativity) *Let $f : [a, b] \to \mathbb{R}$ be nonnegative and Riemann integrable over $[a, b]$. Then*

$$0 \leq \int_a^b f \, .$$

Proof. Exercise. ☐

Propositions 10.1 and 10.2 combine to give the following basic comparison property of definite integrals.

Corollary 10.1. (Order) *Let $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ be Riemann integrable over $[a, b]$. If $f(x) \leq g(x)$ for every $x \in [a, b]$ then*

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

Proof. Exercise. ☐

Some basic bounds on definite integrals follow from Corollary 10.1.

Corollary 10.2. (Bounds) *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable (and hence, bounded) over $[a, b]$. Suppose that $\mathrm{Rng}(f) \subset [\underline{m}, \overline{m}]$. Then*

$$\underline{m} \, (b - a) \leq \int_a^b f \leq \overline{m} \, (b - a) \, .$$

Moreover,

$$\left| \int_a^b f \right| \leq M \, (b - a) \, ,$$

where $M = \sup \big\{ |f(x)| \, : \, x \in [a, b] \big\}$.

Proof. Exercise. ☐

10.3.3. *Positivity and Strict Order.* It is natural to ask when a nonnegative integrand will yield a positive definite integral. The following is our first characterization of when this is the case.

Proposition 10.3. (Positivity) *Let $f : [a, b] \to \mathbb{R}$ be nonnegative and Riemann integrable over $[a, b]$. Then*

$$0 < \int_a^b f$$

if and only if there exists $(c, d) \subset [a, b]$ and $\eta > 0$ such that $\eta < f(x)$ for every $x \in [c, d]$.

Proof. Exercise. ☐

Propositions 10.1 and 10.3 combine to give the following basic comparison property of definite integrals.

Corollary 10.3. (Strict Order) *Let $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ be Riemann integrable over $[a, b]$. Let $f(x) \le g(x)$ for every $x \in [a, b]$. Then*

$$\int_a^b f < \int_a^b g$$

if and only if there exists $(c, d) \subset [a, b]$ and $\eta > 0$ such that $f(x) + \eta < g(x)$ for every $x \in [c, d]$.

Proof. Exercise. $\square$

10.4. **Nonlinearity.** Certain nonlinear combinations of Riemann integrable functions are again Riemann integrable.

Theorem 10.3. (Continuous Compositions) *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable over $[a, b]$. Suppose that* $\mathrm{Rng}(f) \subset [\underline{m}, \overline{m}]$. *Let $G : [\underline{m}, \overline{m}] \to \mathbb{R}$ be continuous. Then $G(f) : [a, b] \to \mathbb{R}$ is Riemann integrable over $[a, b]$.*

Proof. We want to show that for every $\epsilon > 0$ there exists a partition P of $[a, b]$ such that

$$0 \le U(G(f), P) - L(G(f), P) < \epsilon \,. \tag{10.2}$$

Then $G(f)$ would be Riemann integrable by characterization (2) of the Riemann-Darboux Theorem.

Let $\epsilon > 0$. Because $G : [\underline{m}, \overline{m}] \to \mathbb{R}$ is continuous, it is both uniformly continuous and bounded over $[\underline{m}, \overline{m}]$. Because G is uniformly continuous over $[\underline{m}, \overline{m}]$, there exists a $\delta > 0$ such that for every $y, z \in [\underline{m}, \overline{m}]$ we have

$$|y - z| < \delta \implies |G(y) - G(z)| < \frac{\epsilon}{2(b - a)} \,.$$

Because G is bounded over $[\underline{m}, \overline{m}]$, there exists $[\underline{m}^*, \overline{m}^*] \subset \mathbb{R}$ such that $\mathrm{Rng}(G) \subset [\underline{m}^*, \overline{m}^*]$. Because f is Riemann integrable over $[a, b]$ there exists a partition P such that

$$0 \le U(f, P) - L(f, P) < \frac{\delta\epsilon}{2(\overline{m}^* - \underline{m}^*)} \,.$$

We claim that partition P satisfies (10.2).

Let $P = [x_0, x_1, \cdots, x_n]$. For every $i = 1, \cdots .n$ define $\underline{m}_i$, $\overline{m}_i$, $\underline{m}_i^*$, and $\overline{m}_i^*$ by

$$\underline{m}_i = \inf\{f(x) \,:\, x \in [x_{i-1}, x_i]\} \,, \qquad \overline{m}_i = \sup\{f(x) \,:\, x \in [x_{i-1}, x_i]\} \,,$$
$$\underline{m}_i^* = \inf\{G(f(x)) \,:\, x \in [x_{i-1}, x_i]\} \,, \qquad \overline{m}_i^* = \sup\{G(f(x)) \,:\, x \in [x_{i-1}, x_i]\} \,.$$

We want to show that

$$0 \le U(G(f), P) - L(G(f), P) = \sum_{i=1}^{n} (\overline{m}_i^* - \underline{m}_i^*)(x_i - x_{i-1}) < \epsilon \,.$$

The key step is to decompose the indices $i = 1, \cdots, n$ into two sets:

$$I_< = \{i \,:\, \overline{m}_i - \underline{m}_i < \delta\} \,, \qquad I_\ge = \{i \,:\, \overline{m}_i - \underline{m}_i \ge \delta\} \,.$$

We analyze the above sum over each of these sets separately.

We first show that the sum over the "good" set $I_<$ is small because each $(\overline{m}_i^* - \underline{m}_i^*)$ is sufficiently small. The values of f over $[x_{i-1}, x_i]$ lie in $[\underline{m}_i, \overline{m}_i]$. Because G is continuous, the

Extreme-Value Theorem implies that G takes on its inf and sup over $[\underline{m}_i, \overline{m}_i]$, say at the points $\underline{y}_i$ and $\overline{y}_i$ respectively. Because $|\overline{y}_i - \underline{y}_i| < \delta$ for every $i \in I_<$, we have

$$\begin{aligned}\overline{m}_i^* - \underline{m}_i^* &\leq \sup\{G(y) \,:\, y \in [\underline{m}_i, \overline{m}_i]\} - \inf\{G(y) \,:\, y \in [\underline{m}_i, \overline{m}_i]\} \\ &= G(\overline{y}_i) - G(\underline{y}_i) < \frac{\epsilon}{2(b-a)}\,,\end{aligned}$$

whereby the sum over $I_<$ satisfies

$$\sum_{i\in I_<} (\overline{m}_i^* - \underline{m}_i^*)(x_i - x_{i-1}) < \frac{\epsilon}{2(b-a)} \sum_{i\in I_<} (x_i - x_{i-1}) \leq \frac{\epsilon}{2}\,. \tag{10.3}$$

We now show that the sum over the "bad" set $I_\geq$ is small because the partition P is refined enough to make the set $I_\geq$ sufficiently small. Indeed, because $\delta \leq \overline{m}_i - \underline{m}_i$ for every $i \in I_\geq$, we have

$$\delta \sum_{i\in I_\geq} (x_i - x_{i-1}) \leq \sum_{i\in I_\geq} (\overline{m}_i - \underline{m}_i)(x_i - x_{i-1}) \leq U(f,P) - L(f,P) < \frac{\delta\epsilon}{2(\overline{m}^* - \underline{m}^*)}\,,$$

whereby the set $I_\geq$ is small in the sense that

$$\sum_{i\in I_\geq} (x_i - x_{i-1}) < \frac{\epsilon}{2(\overline{m}^* - \underline{m}^*)}\,.$$

Hence, the sum over $I_\geq$ satisfies

$$\sum_{i\in I_\geq} (\overline{m}_i^* - \underline{m}_i^*)(x_i - x_{i-1}) < (\overline{m}^* - \underline{m}^*) \sum_{i\in I_\geq} (x_i - x_{i-1}) < \frac{\epsilon}{2}\,. \tag{10.4}$$

Upon combining bounds (10.3) and (10.4), we obtain

$$\begin{aligned}0 &\leq U(G(f),P) - L(G(f),P) \\ &= \sum_{i\in I_<} (\overline{m}_i^* - \underline{m}_i^*)(x_i - x_{i-1}) + \sum_{i\in I_\geq} (\overline{m}_i^* - \underline{m}_i^*)(x_i - x_{i-1}) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\,.\end{aligned}$$

Because ϵ was arbitrary, $G(f)$ is Riemann integrable by characterization (2) of the Riemann-Darboux Theorem. □

An important consequence of the Composition Theorem is that the product of Riemann integrable functions is also Riemann integrable.

Proposition 10.4. (Product) *Let $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ be Riemann integrable over $[a,b]$. Then the product $fg : [a,b] \to \mathbb{R}$ is Riemann integrable over $[a,b]$.*

Remark. Taken together the Linearity and Product Propositions show that the class of Riemann integrable functions is an algebra.

Proof. The proof is based on the algebraic identity

$$fg = \tfrac{1}{4}\big((f+g)^2 - (f-g)^2\big)\,.$$

By the Linearity Proposition the functions $f+g$ and $f-g$ are Riemann integrable over $[a,b]$. By Composition Theorem (applied to $G(z) = z^2$) the functions $(f+g)^2$ and $(f-g)^2$ are Riemann integrable over $[a,b]$. Hence, by applying the Linearity Proposition to the above identity, we see that fg is Riemann integrable over $[a,b]$. □

Remark. We could just as well have built a proof of the Product Lemma based on the identity

$$fg = \tfrac{1}{2}\big((f+g)^2 - f^2 - g^2\big)\,,$$

or the identity

$$fg = \tfrac{1}{2}\big(f^2 + g^2 - (f-g)^2\big)\,.$$

The Composition Theorem also implies that the absolute-value of a Riemann integrable function is also Riemann integrable. When combined with the Order, Bounds, and Product Propositions 10.1, 10.2, and 10.4, this leads to the following useful bound.

Proposition 10.5. (Absolute-Value) *Let $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ be Riemann integrable over $[a,b]$. Suppose that g is nonnegative. Then $fg : [a,b] \to \mathbb{R}$ and $|f|\,g : [a,b] \to \mathbb{R}$ are Riemann integrable over $[a,b]$ and satisfy*

$$\left| \int_a^b fg \right| \leq \int_a^b |f|\,g \leq M \int_a^b g\,,$$

where $M = \sup\{|f(x)| \,:\, x \in [a,b]\}$.

Proof. Exercise. □

The fact the absolute-value of a Riemann integrable function is also Riemann integrable implies that the minimum and the maximum of two Riemann integrable functions are also Riemann integrable.

Proposition 10.6. (Minimum-Maximum) *Let $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ be Riemann integrable over $[a,b]$. Then $\min\{f,g\}$ and $\max\{f,g\}$ are Riemann integrable over $[a,b]$.*

Proof. This follows from the formulas

$$\begin{aligned} \min\{f,g\}(x) &= \frac{f(x)+g(x)}{2} - \frac{|f(x)-g(x)|}{2}\,,\\ \max\{f,g\}(x) &= \frac{f(x)+g(x)}{2} + \frac{|f(x)-g(x)|}{2}\,. \end{aligned}$$

The details are left as an exercise. □

10.5. **Restrictions and Interval Additivity.** A property of the definite integral that you learned when you first studied integration is interval additivity. In its simplest form this property states that, provided all the integrals exist, for every $a, b, c \in \mathbb{R}$ such that $a < b < c$ we have

$$\int_a^c f = \int_a^b f + \int_b^c f\,. \tag{10.5}$$

In elementary calculus courses this formula is often stated without much emphasis on implicit integrability assumptions. As we will see below, Riemann integrals have this property. In that setting this formula assumes that f is Riemann integrable over $[a,c]$, and that the restrictions of f to $[a,b]$ and $[b,c]$ are Riemann integrable over those intervals. As the next lemma shows, these last two assumptions follow from the first.

Lemma 10.1. (Restriction) *Let $f : [a,d] \to \mathbb{R}$ be Riemann integrable over $[a,d]$. Then for every $[b,c] \subset [a,d]$ with $b < c$ the restriction of f to $[b,c]$ is Riemann integrable over $[b,c]$.*

Proof. Let $[b,c] \subset [a,d]$. Let $\epsilon > 0$. Because f is Riemann integrable over $[a,d]$ by characterization (2) of the Riemann-Darboux Theorem there exists a partition P^* of $[a,d]$ such that

$$0 \leq U(f,P^*) - L(f,P^*) < \epsilon\,.$$

By the Refinement Lemma we may assume that b and c are partition points of P^*, otherwise we can simply replace P^* by $P^* \vee [a,b,c,d]$. Let P be the partition of $[b,c]$ induced by P^*. Then

$$0 \leq U(f,P) - L(f,P) \leq U(f,P^*) - L(f,P^*) < \epsilon\,.$$

Hence, f is Riemann integrable over $[b,c]$ by characterization (2) of the Riemann-Darboux Theorem. □

Now return to the interval additivity formula (10.5). More interesting from the viewpoint of building up the class of Riemann integrable functions is the fact that if the restrictions of f to $[a,b]$ and $[b,c]$ are Riemann integrable over those intervals then f is Riemann integrable over $[a,c]$. More generally, we have the following.

Proposition 10.7. (Interval Additivity) *Let $P = [p_0, \cdots, p_k]$ be any partition of $[a,b]$. Then f is Riemann integrable over $[a,b]$ if and only if the restriction of f to $[p_{i-1}, p_i]$ is Riemann integrable for every $i = 1, \cdots, k$. Moerover, in that case we have*

$$\int_a^b f = \sum_{i=1}^{k} \int_{p_{i-1}}^{p_i} f\,. \tag{10.6}$$

Proof. ($\Longrightarrow$) If f is Riemann integrable over $[a,b]$ then the Restriction Lemma implies that the restriction of f to $[p_{i-1}, p_i]$ is Riemann integrable for every $i = 1, \cdots, k$.

($\Longleftarrow$) Because f is Riemann integrable over $[p_{i-1}, p_i]$ for every $i = 1, \cdots, k$ there exists a partition P_i^* of $[p_{i-1}, p_i]$ such that

$$0 \leq U(f,P_i^*) - L(f,P_i^*) < \frac{\epsilon}{k}\,.$$

Let P^* be the refinement of P such that P_i^* is the induced partition of $[p_{i-1}, p_i]$. We then see that

$$0 \leq U(f,P^*) - L(f,P^*) = \sum_{i=1}^{k} \big(U(f,P_i^*) - L(f,P_i^*)\big) < \sum_{i=1}^{k} \frac{\epsilon}{k} = \epsilon\,.$$

Hence, by characterization (2) of the Riemann-Darboux Theorem, f is Riemann integrable over $[a,b]$.

We can use Riemann sums to establish (10.6). This part of the proof is left as an exercise. □

The restriction $a < b < c$ in the interval additivity formula (10.5) can be dropped provided we adopt the following convention.

Definition 10.2. *Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable over $[a,b]$. Define*

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

Exercise. Show that the interval additivity formula (10.5) holds for every a, b, and $c \in \mathbb{R}$, provided that we adopt Definition 10.2 and f is Riemann integrable over every interval involved.

10.6. **Extensions and Piecewise Integrability.** Now we will use interval additivity to build up the class of Riemann integrable functions. This will require a lemma regarding extensions. To motivate the need for this lemma, let us consider the function $f : [-1, 1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x + 2 & \text{for } x \in [-1, 0)\,, \\ 1 & \text{for } x = 0\,, \\ x & \text{for } x \in (0, 1]\,. \end{cases}$$

It is easy to use the Riemann-Darboux Theorem to verify that this function is Riemann integrable with

$$\int_{-1}^{1} f = 2\,,$$

yet this fact does not follow directly from other theorems we have proved. For example, f restricted to either $[-1, 0]$ or $[0, 1]$ is neither monotonic nor continuous because of its behavior at $x = 0$. However, our intuition tells us (correctly) that the value of f at 0 should not effect whether or not it is Riemann integrable. The following lemma shows this to be the case if the points in questions are the endpoints of the interval of integration.

Lemma 10.2. (Extension) *Let $f : (a, b) \to \mathbb{R}$ be bounded. Suppose that for every $[c, d] \subset (a, b)$ the restriction of f to $[c, d]$ is Riemann integrable over $[c, d]$. Let $\hat{f} : [a, b] \to \mathbb{R}$ be any extension of f to $[a, b]$. Then $\hat{f}$ is Riemann integrable over $[a, b]$. Moreover, if $\hat{f}_1$ and $\hat{f}_2$ are two such extensions of f then*

$$\int_a^b \hat{f}_1 = \int_a^b \hat{f}_2\,.$$

Proof. Let $\epsilon > 0$. Let $\mathrm{Rng}(\hat{f}) \subset [\underline{m}, \overline{m}]$. Let $\delta > 0$ such that

$$(\overline{m} - \underline{m})\delta < \frac{\epsilon}{3} \qquad \text{and} \qquad \delta < \frac{b-a}{2}\,.$$

Because the restriction of f to $[a + \delta, b - \delta]$ is Riemann integrable, there exists a partition P of $[a + \delta, b - \delta]$ such that

$$0 \leq U(f, P) - L(f, P) < \frac{\epsilon}{3}\,.$$

Let P^* be the extension of P to $[a, b]$ obtained by adding a and b as partition points. Then

$$\begin{aligned} 0 &\leq U(\hat{f}, P^*) - L(\hat{f}, P^*) \\ &= \big[U(\hat{f}, [a, a + \delta]) - L(\hat{f}, [a, a + \delta])\big] + \big[U(f, P) - L(f, P)\big] \\ &\quad + \big[U(\hat{f}, [b - \delta, b]) - L(\hat{f}, [b - \delta, b])\big] \\ &\leq (\overline{m} - \underline{m})\delta + \frac{\epsilon}{3} + (\overline{m} - \underline{m})\delta < \epsilon\,. \end{aligned}$$

Hence, the extension $\hat{f}$ is Riemann integrable over $[a, b]$ by characterization (2) of the Riemann-Darboux Theorem.

Now let $\hat{f}_1$ and $\hat{f}_2$ be two extensions of f to $[a, b]$. Let $\{P^n\}_{n=1}^{\infty}$ be any sequence of partitions of $[a, b]$ such that $|P^n| \to 0$ as $n \to \infty$. This sequence is Archimedean for both $\hat{f}_1$ and $\hat{f}_2$ by Theorem 9.4. Let $\{Q^n\}_{n=1}^{\infty}$ be any sequence of associated quadrature points such that neither a nor b are quadrature points. Because $\hat{f}_1(x) = \hat{f}_2(x)$ for every $x \in (a, b)$, we have

$R(\hat{f}_1, P^n, Q^n) = R(\hat{f}_2, P^n, Q^n)$ for every $n \in \mathbb{Z}_+$. Therefore the Archimedes-Riemann Theorem yields

$$\int_a^b \hat{f}_1 = \lim_{n\to\infty} R(\hat{f}_1, P^n, Q^n) = \lim_{n\to\infty} R(\hat{f}_2, P^n, Q^n) = \int_a^b \hat{f}_2 \,.$$

$\square$

It is a consequence of the Extension Lemma and interval additivity that two functions that differ at only a finite number of points are the same when is comes to Riemann integrals.

Theorem 10.4. (Piecewise Integrability). *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable over $[a, b]$. Let $g : [a, b] \to \mathbb{R}$ such that $g(x) = f(x)$ at all but a finite number of points in $[a, b]$. Then g is Riemann integrable over $[a, b]$ and*

$$\int_a^b g = \int_a^b f \,.$$

Proof. Exercise. $\square$

Remark. The same cannot be said of two functions that differ at a countable number of points. Indeed, consider the function

$$g(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}\,, \\ 0 & \text{if } x \notin \mathbb{Q}\,. \end{cases}$$

Its restriction to any closed bounded interval $[a, b]$ is not Riemann integrable, yet it differs from $f = 0$ at a countable number of points.

We can now show that all bounded, piecewise monotonic functions over $[a, b]$ are also Riemann intergrable over $[a, b]$. We first recall the definition of piecewise monotonic function.

Definition 10.3. *A function $f : [a, b] \to \mathbb{R}$ is said to be* piecewise monotonic *if there exists a partition $P = [x_0, \cdots, x_n]$ of $[a, b]$ such that f is monotonic over (x_{i-1}, x_i) for every $i = 1, \cdots, n$.*

Theorem 10.5. (Piecewise Monotonic Integrability). *Let $f : [a, b] \to \mathbb{R}$ be bounded and piecewise monotonic. Then f is Riemann integrable over $[a, b]$.*

Proof. This follows from Proposition 10.1, the Extension Lemma, and the Interval Additivity Proposition. The details are left as an exercise. $\square$

We can also show that all bounded, piecewise continuous functions over $[a, b]$ are also intergrable over $[a, b]$. We first recall the definition of piecewise continuous function.

Definition 10.4. *A function $f : [a, b] \to \mathbb{R}$ is said to be* piecewise continuous *if there exists a partition $P = [x_0, \cdots, x_n]$ of $[a, b]$ such that f is continuous over (x_{i-1}, x_i) for every $i = 1, \cdots, n$.*

Theorem 10.6. (Piecewise Continuous Integrability). *Let $f : [a, b] \to \mathbb{R}$ be bounded and piecewise continuous. Then f is Riemann integrable over $[a, b]$.*

Proof. This follows from Proposition 10.2, the Extension Lemma, and the Interval Additivity Proposition. The details are left as an exercise. $\square$

Remark. The class of bounded, piecewise continuous functions includes some fairly wild functions that are not piecewise monotonic. For example, it contains $f : [-1, 1] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 1 + \sin(1/x) & \text{if } x \in (0.1]\,, \\ 4 & \text{if } x = 0\,, \\ -1 + \sin(1/x) & \text{if } x \in [-1, 0)\,. \end{cases}$$

Similarly, the class of bounded, piecewise monotonic functions includes some functions that are not piecewise continuous. Indeed, any function that is monotonic over an interval but has an infinite number of jump discontinuities is not piecewise continuous. Such a function is $f : [0, 1] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 2^{-k} & \text{if } x \in (2^{-(k+1)}, 2^{-k}] \text{ for some } k \in \mathbb{N}\,, \\ 0 & \text{if } x = 0\,. \end{cases}$$

While we have indentified many Riemann intergrable functions, our understanding of the entire class of such functions is still lacking. This will be remedied in the next section.

10.7. **Lebesgue Theorem.** In this section we state a beautiful theorem of Lebesgue that characterizes those functions that are Riemann integrable. In order to do this we need to introduce the following notion of "very small" subsets of $\mathbb{R}$.

Definition 10.5. *A set $A \subset \mathbb{R}$ is said to have* measure zero *if for every $\epsilon > 0$ there exists a countable collection of open intervals $\{(a_i, b_i)\}_{i=1}^{\infty}$ such that*

$$A \subset \bigcup_{i=1}^{\infty} (a_i, b_i)\,, \qquad \textit{and} \qquad \sum_{i=1}^{\infty} (b_i - a_i) < \epsilon\,.$$

In other words, a set has measure zero when it can be covered by a countable collection of open intervals, the sum of whose lengths is arbitrarily small.

Example. Every countable subset of $\mathbb{R}$ has measure zero. In particular, $\mathbb{Q}$ has measure zero. Indeed, consider a countable set $A = \{x_i\}_{i=1}^{\infty} \subset \mathbb{R}$. Let $\epsilon > 0$. Let $r < \frac{1}{3}$ and set $(a_i, b_i) = (x_i - r^i\epsilon, x_i + r^i\epsilon)$ for every $i \in \mathbb{Z}_+$. The collection of open intervals $\{(a_i, b_i)\}_{i=1}^{\infty}$ clearly covers A. Moreover,

$$\sum_{i=1}^{\infty} (b_i - a_i) = \sum_{i=1}^{\infty} 2r^i\epsilon = \frac{2r\epsilon}{1 - r} < \epsilon\,.$$

The fact that measure zero is a reasonable concept of "very small" is confirmed by the following facts.

Proposition 10.8. *If $\{A_n\}_{n=1}^{\infty}$ is a collection of subsets of $\mathbb{R}$ that each have measure zero then*

$$A = \bigcup_{n=1}^{\infty} A_n \quad \textit{has measure zero}\,.$$

If $B \subset \mathbb{R}$ has measure zero and $A \subset B$ then A has measure zero.

Proof. Exercise. □

Of course, we have to show that many sets do not have measure zero. The first step in this direction is provided by the following lemma.

Lemma 10.3. *Let $[c,d] \subset \mathbb{R}$. Let $\{(a_i,b_i)\}_{i=1}^m$ be a finite collection of open intervals such that*

$$[c,d] \subset \bigcup_{i=1}^m (a_i,b_i)\,.$$

Then

$$d-c < \sum_{i=1}^m (b_i - a_i)\,.$$

Proof. Exercise. □

Proposition 10.9. *Sets in $\mathbb{R}$ that contain a nonempty open interval do not have measure zero.*

Remark. There are many such sets!

Proof. Let $A \subset \mathbb{R}$ contain the nonempty open interval (a,b). Let $[c,d] \subset (a,b)$ with $c < d$. Because $[c,d] \subset A$, if we can show that $[c,d]$ does not have measure zero then Proposition 10.8 implies that A does not have measure zero. Let $\{(a_i,b_i)\}_{i=1}^\infty$ be any countable collection of open intervals such that

$$[c,d] \subset \bigcup_{i=1}^\infty (a_i,b_i)\,.$$

We claim that there exists $m \in \mathbb{N}$ such that

$$[c,d] \subset \bigcup_{i=1}^m (a_i,b_i)\,.$$

Then by Lemma 10.3 we see that

$$d-c < \sum_{i=1}^m (b_i - a_i) \le \sum_{i=1}^\infty (b_i - a_i)\,.$$

Therefore $[c,d]$ does not have measure zero. □

Exercise. Prove the claim asserted in the above proof.

The next example shows that there are some very interesting sets that have measure zero.

Example. The Cantor set is an uncountable set that has measure zero. The Cantor set is the subset C of the interval $[0,1]$ obtained by sequentially removing "middle thirds" as follows. Define the sequence of sets $\{C_n\}_{n=1}^\infty$ as follows

$$\begin{aligned}
C_1 &= [0,1] - (\tfrac{1}{3},\tfrac{2}{3}) = [0,\tfrac{1}{3}] \cup [\tfrac{2}{3},1]\,,\\
C_2 &= C_1 - (\tfrac{1}{9},\tfrac{2}{9}) \cup (\tfrac{7}{9},\tfrac{8}{9})\\
&= [0,\tfrac{1}{9}] \cup [\tfrac{2}{9},\tfrac{1}{3}] \cup [\tfrac{2}{3},\tfrac{7}{9}] \cup [\tfrac{8}{9},1]\,,\\
C_3 &= C_2 - (\tfrac{1}{27},\tfrac{2}{27}) \cup (\tfrac{7}{27},\tfrac{8}{27}) \cup (\tfrac{19}{27},\tfrac{20}{27}) \cup (\tfrac{25}{27},\tfrac{26}{27})\\
&= [0,\tfrac{1}{27}] \cup [\tfrac{2}{27},\tfrac{1}{9}] \cup [\tfrac{2}{9},\tfrac{7}{27}] \cup [\tfrac{8}{27},\tfrac{1}{3}] \cup [\tfrac{2}{3},\tfrac{19}{27}] \cup [\tfrac{20}{27},\tfrac{7}{9}] \cup [\tfrac{8}{9},\tfrac{25}{27}] \cup [\tfrac{26}{27},1]\,,\\
&\;\;\vdots
\end{aligned}$$

In general we have

$$C_n = C_{n-1} - \bigcup_{2k<3^n} (\tfrac{2k-1}{3^n},\tfrac{2k}{3^n}) \quad \text{for } n > 3\,.$$

We can show by induction that each C_n is the union of 2^n closed intervals each of which have length $1/3^n$. Therefore each C_n is sequentially compact. Moreover, these sets are nested as

$$C_1 \supset C_2 \supset \cdots \supset C_n \supset C_{n+1} \supset \cdots .$$

The Cantor set C is then defined to be the intersection of these sets:

$$C = \bigcap_{n=1}^{\infty} C_n .$$

Being the intersection of nested sequentially compact sets, this set is nonempty by the Cantor Theorem. It is harder to show that C is uncountable. We will not do so here. However, from the information given above you should be able to show that C has measure zero.

Exercise. Show the Cantor set has measure zero.

Before stating the Lebesgue Theorem we need one more definition.

Definition 10.6. *Let $S \subset \mathbb{R}$. Let $\mathcal{A}(x)$ be any assertion about any point $x \in \mathbb{R}$. Then we say that "$\mathcal{A}(x)$ for almost every $x \in S$" or "$\mathcal{A}$ almost everywhere in S" when*

$$\textit{the set} \quad \big\{x \in S \,:\, \mathcal{A}(x) \text{ is false}\big\} \quad \textit{has measure zero}.$$

Roughly, a property holds almost everywhere if it fails to hold on a set of measure zero.

Example. Let $f : \mathbb{R} \to \mathbb{R}$ be given by

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}\,, \\ 0 & \text{if } x \notin \mathbb{Q}\,. \end{cases}$$

Then $f = 0$ almost everywhere in $\mathbb{R}$.

We are now ready to state the Lebesgue Theorem.

Theorem 10.7. (Lebesgue) *Let $f : [a, b] \to \mathbb{R}$ be bounded. Then f is Riemann integrable over $[a, b]$ if and only if it is continuous almost everywhere in $[a, b]$.*

Proof. The proof is omitted. It is quite involved. One can be found in "Principles of Analysis" by Walter Rudin. $\square$

Remark. Riemann gave a characterization of Riemann integrable functions that was the centerpiece of his theory of integration. This characterization suggested that the nature of the discontinuities of a function mattered. Lebesgue realized that only the size of the set of discontinuities mattered, and gave his elegant characterization in terms of the concept of measure zero. Most proofs of the Lebesgue Theorem borrow heavily from Riemann. Lebesgue introduced the concept of measure zero in developing what is now called the Lebesgue integral, his beautiful and powerful extension of the Riemann integral.

Remark. Much of what we now call the theory of the Riemann integral existed long before Riemann. The idea of defining definite integrals as limits of Riemann sums has its roots in the Method of Archimedes for finding areas and volumes. The fact that continuous functions are integrable was proved by Cauchy. Riemann developed his theory to address the question of what discontinuous functions are integrable. It was presented as part of a paper addressing the question of what functions could be represented by Fourier series and Fourier transforms. Our approach approach follows Darboux, who saw that Darboux sums give an elegant way to control all the Riemann sums.

The Lebesgue Theorem allows us to sharpen our Nonnegativity and Order Propositions for Riemann integrals (Proposition 10.3 and Corollary 10.3).

Proposition 10.10. (Positivity) *Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable. Suppose that $f \geq 0$ and that $f > 0$ almost everywhere over a nonempty $(c,d) \subset [a,b]$. Then*

$$\int_a^b f > 0\,.$$

Proof. The key step is to show that there exists a $p \in (c,d)$ such that $f(p) > 0$ and f is continuous at p. Indeed, consider the sets

$$Y = \{x \in (c,d) \,:\, f(x) = 0\}\,,$$
$$Z = \{x \in (c,d) \,:\, f \text{ is not continous at } x\}\,.$$

The set Y has measure zero by hypothesis. The set Z has measure zero by the Lebesgue Theorem. The set $Y \cup Z$ therefore has measure zero by the first assertion of Proposition 10.8. But then $Y \cup Z$ cannot contain (c,d) because otherwise the last assertion of Proposition 10.8 would imply (c,d) has measure zero, which it does not. Hence, the set $(c,d) - Y \cup Z$ is nonempty. The rest of the proof is left as an exercise. □

Corollary 10.4. (Strict Order) *Let $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ be Riemann integrable. Suppose that $f \leq g$ and that $f(x) < g(x)$ almost everywhere over a nonempty $(c,d) \subset [a,b]$. Then*

$$\int_a^b f < \int_a^b g\,.$$

Proof. Exercise. □

The power of the Lebesgue Theorem becomes even clearer when you realize that it implies the Continuous Composition Theorem (Theorem 10.3), the proof of which was not easy.

New Proof of Proposition 10.3. Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable and $G : [\underline{m}, \overline{m}] \to \mathbb{R}$ be continuous with $\mathrm{Rng}(f) \subset [\underline{m}, \overline{m}]$. Because f is Riemann integrable over $[a,b]$ it is continuous almost everywhere in $[a,b]$. Because G is continuous over $\mathrm{Rng}(f)$, the function $G(f) : [a,b] \to \mathbb{R}$ is continuous wherever f is continuous. Hence, $G(f)$ is continuous almost everywhere in $[a,b]$. Therefore $G(f)$ is Riemann integrable over $[a,b]$. □

Remark. In a similar fashion, the Lebesgue Theorem can be used to prove many of our earlier results on Riemann integrable functions. For example, it can be used to prove Theorem 10.1 (Monotonic Integrability), Theorem 10.2 (Continuous Integrability), Theorem 10.5 (Piecewise Monotonic Integrability), and Theorem 10.6 (Piecewise Continuous Integrability).

Exercise. Use the Lebesgue Theorem to prove Theorem 10.1.

Exercise. Use the Lebesgue Theorem to prove Theorem 10.2.

Exercise. Use the Lebesgue Theorem to prove Theorem 10.5.

Exercise. Use the Lebesgue Theorem to prove Theorem 10.6.

11. Relating Integration with Differentiation

Both integration and differentiation predate Newton and Leibniz. The definite integral has roots that go back at least as far as Eudoxos and Archimedes, some two thousand years earlier. The derivative goes back at least as far as Fermat. The fact they are connected in some instances was understood by Fermat, who worked out special cases, and by Barrow, who extended Fermat's work. Barrow was one of Newton's teachers and his work was known to Leibniz. The big breakthrough of Newton and Leibniz was the understanding that this connection is general. This realization made the job of computing definite intergrals much easier, which enabled major advances in science, engineering, and mathematics. This connection takes form in what we now call the first and second fundamental theorems of calculus.

11.1. First Fundamental Theorem of Calculus.

The business of evaluating integrals by taking limits of Riemann sums is usually either difficult or impossible. However, as you have known since you first studied integration, for many integrands there is a much easier way. We begin with a definition.

Definition 11.1. *Let $f : [a, b] \to \mathbb{R}$. A function $F : [a, b] \to \mathbb{R}$ is said to be a* primitive *or* antiderivative *of f over $[a, b]$ provided*

- *the function F is continuous over $[a, b]$,*
- *there exists a partition $[p_0, \cdots, p_n]$ of $[a, b]$ such that for each $i = 1, \cdots, n$ the function F restricted to (p_{i-1}, p_i) is differentiable and satisfies*

$$F'(x) = f(x) \quad \text{for every } x \in (p_{i-1}, p_i)\,. \tag{11.1}$$

Remark. Definition 11.1 states that F is continuous and piecewise differentiable over $[a, b]$. There are at most a finite number of points in $[a, b]$ at which either F' is not defined or F' is defined but $F' \neq f$. For example, consider the function

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

The function $F(x) = |x|$ is a primitive of f over $[-1, 1]$, yet is not differentiable at $x = 0$. Similarly, consider the function

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

The differentiable function $F(x) = x$ is a primitive of this f over $[-1, 1]$, yet $F'(0) \neq f(0)$. In both examples there is clearly no function F that is differentiable over $[-1, 1]$ such that $F' = f$ because f does not have the intermediate-value property.

Exercise. Let $f : [a, b] \to \mathbb{R}$. Let $F : [a, b] \to \mathbb{R}$ be a primitive of f over $[a, b]$. Let $g : [a, b] \to \mathbb{R}$ such that $g(x) = f(x)$ at all but a finite number of points of $[a, b]$. Show that F is also a primitive of g over $[a, b]$.

If F is a primitive of a function f over $[a, b]$ then so is $F + c$ for any constant c. The next lemma shows that a primitive is unique up to this arbitrary additive constant.

Lemma 11.1. *Let $f : [a, b] \to \mathbb{R}$. Let $F_1 : [a, b] \to \mathbb{R}$ and $F_2 : [a, b] \to \mathbb{R}$ be primitives of f over $[a, b]$. Then there exists a constant c such that $F_2(x) = F_1(x) + c$ for every $x \in [a, b]$.*

Proof. Let $G = F_2 - F_1$. We must show that this function is a constant over $[a, b]$. Let P^1 and P^2 be the partitions associated with F_1 and F_2 respectively. Set $P = P^1 \vee P^2$. Express P in terms of its partition points as $P = [p_0, \cdots, p_n]$. For each $i = 1, \cdots, n$ the restriction of G to $[p_{i-1}, p_i]$ is continuous over $[p_{i-1}, p_i]$ and differentiable over (p_{i-1}, p_i) with

$$G'(x) = F_2'(x) - F_1'(x) = f(x) - f(x) = 0 \quad \text{for every } x \in (p_{i-1}, p_i)\,.$$

It follows from the Lagrange Mean-Value Theorem that restriction of G to each $[p_{i-1}, p_i]$ is constant c_i over that subinterval. But for each $i = 1, \cdots, n-1$ the point p_i is in the subintervals $[p_{i-1}, p_i]$ and $[p_i, p_{i+1}]$, whereby $c_i = G(p_i) = c_{i+1}$. Hence, G must be a constant over $[a, b]$. □

Corollary 11.1. *Let $f : [a, b] \to \mathbb{R}$ have a primitive over $[a, b]$. Let $x_o \in [a, b]$ and $y_o \in \mathbb{R}$. Then f has a unique primitive F such that $F(x_o) = y_o$.*

Proof. Exercise. □

Exercise. Let $f : [0, 3] \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x & \text{for } 0 \le x < 1\,, \\ -x & \text{for } 1 \le x < 2\,, \\ 1 & \text{for } 2 \le x \le 3\,. \end{cases}$$

Find F, the primitive of f over $[0, 3]$ specified by $F(0) = 1$.

We are now ready to for the big theorem.

Theorem 11.1. (First Fundamental Theorem of Calculus) *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable and have a primitive F over $[a, b]$. Then*

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

Remark. This theorem recasts the problem of evaluating definite integrals to that of finding an explicit primitive of f. While such an explicit primitive cannot always be found, it can be found for a wide class of elementary integrands f.

Proof. By characterization (3) of the Riemann-Darboux Theorem, the result will follow if for every partition P of $[a, b]$ we have

$$L(f, P) \le F(b) - F(a) \le U(f, P)\,. \tag{11.2}$$

Let P be an arbitrary partition of $[a, b]$. Let $[p_0, \cdots, p_n]$ be the partition of $[a, b]$ associated with the primitive F. Let $P^* = P \vee [p_0, \cdots, p_n]$. Express P^* in terms of its partition points as $P^* = [x_0, \cdots, x_{n^*}]$. Then for every $i = 1, \cdots, n^*$ we know that $F : [x_{i-1}, x_i] \to \mathbb{R}$ is continuous, and that $F : (x_{i-1}, x_i) \to \mathbb{R}$ is differentiable. Then by the Lagrange Mean-Value Theorem there exists $q_i \in (x_{i-1}, x_i)$ such that

$$F(x_i) - F(x_{i-1}) = F'(q_i)\,(x_i - x_{i-1}) = f(q_i)\,(x_i - x_{i-1})\,.$$

Because $\underline{m}_i \le f(q_i) \le \overline{m}_i$, we see from the above that

$$\underline{m}_i\,(x_i - x_{i-1}) \le F(x_i) - F(x_{i-1}) \le \overline{m}_i\,(x_i - x_{i-1}) \quad \text{for every } i = 1, \cdots, n^*\,.$$

Upon summing these inequalities we obtain

$$L(f, P^*) \le \sum_{i=1}^{n^*} \big(F(x_i) - F(x_{i-1})\big) \le U(f, P^*)\,.$$

Because the above sum telescopes, we see that

$$\sum_{i=1}^{n^*} \big(F(x_i) - F(x_{i-1})\big) = F(b) - F(a)\,.$$

Because P^* is a refinement of P, the Refinement Lemma then yields

$$L(f,P) \leq L(f,P^*) \leq F(b) - F(a) \leq U(f,P^*) \leq U(f,P)\,.$$

Because the partition P was arbitrary, equality (11.2) follows from characterization (3) of the Riemann-Darboux Theorem. $\square$

Remark. Notice that the First Fundamental Theorem of Calculus does not require f to be continuous, or even piecewise continuous. It only requires f to be Riemann integrable and to have a primitive. Also notice how Definition 11.1 of primitives allows the use of the Lagrange Mean-Value Theorem in the above proof.

The following is an immediate corollary of the First Fundamental Theorem of Calculus.

Corollary 11.2. *Let $F : [a,b] \to \mathbb{R}$ be continuous over $[a,b]$ and differentiable over (a,b). Suppose $F' : (a,b) \to \mathbb{R}$ is bounded over (a,b) and Riemann integrable over every $[c,d] \subset (a,b)$. Let f be any extension of F' to $[a,b]$. Then f is Riemann integrable over $[a,b]$, F is a primitive of f over $[a,b]$, and*

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

Example. Let F be defined over $[-1,1]$ by

$$F(x) = \begin{cases} x\cos(\log(1/|x|)) & \text{if } x \neq 0\,, \\ 0 & \text{if } x = 0\,. \end{cases}$$

Then F is continuous over $[-1,1]$ and continuously differentiable over $[-1,0) \cup (0,1]$ with

$$F'(x) = \cos(\log(1/|x|)) + \sin(\log(1/|x|))\,.$$

As this function is bounded, we have

$$\int_{-1}^{1} \big[\cos(\log(1/|x|)) + \sin(\log(1/|x|))\big]\,\mathrm{d}x = F(1) - F(-1) = 2\,.$$

Here the integrand can be assigned any value at $x = 0$. We first apply the above corollary to $[-1,0]$ and to $[0,1]$, and then use interval additivity.

11.2. **Second Fundamental Theorem of Calculus.** It is natural to ask if every Riemann integrable function has a primitive. It is clear from the First Fundamental Theorem that if f is Riemann integrable over $[a,b]$ and has a primitive F that we must have

$$F(x) = F(a) + \int_a^x f\,.$$

So given a function f that is Riemann integrable over $[a,b]$, we can define F by the above formula. We then check if $F'(x) = f(x)$ except at a finite number of points. In general this will not be the case. For example, if $f : [0,1] \to \mathbb{R}$ is the Riemann function given by

$$f(x) = \begin{cases} \frac{1}{q} & \text{if } x \in \mathbb{Q} \text{ with } x = \frac{p}{q} \text{ in lowest terms}\,. \\ 0 & \text{otherwise}\,. \end{cases}$$

This function is continuous at all the irrationals, and so is Riemann integrable by the Lebesgue Theorem. Moreover, we can show that for every $x \in [0,1]$ we have

$$F(x) = \int_0^x f = 0\,.$$

Hence, F is differentiable but $F'(x) \neq f(x)$ at every rational. Therefore F is not a primitive of f. Therefore f has no primitives.

The Second Fundamental Theorem of Calculus shows that the above construction does yield a primitive for a large classes of functions.

Theorem 11.2. (Second Fundamental Theorem of Calculus) *Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable. Define $F : [a,b] \to \mathbb{R}$ by*

$$F(x) = \int_a^x f \qquad \textit{for every } x \in [a,b]\,.$$

Then $F(a) = 0$, F is Lipschitz continuous over $[a,b]$, and if f is continuous at $c \in [a,b]$ then F is differentiable at c with $F'(c) = f(c)$.

In particular, if f is continuous over $[a,b]$ then F is continuously differentiable over $[a,b]$ with $F' = f$. If f is piecewise continuous over $[a,b]$ then F is piecewise continuously differentiable over $[a,b]$ with $F' = f$ at all but a finite number of points in $[a,b]$.

Proof. The fact that $F(a) = 0$ is obvious. Next, we show that F is Lipschitz continuous over $[a,b]$. Let $M = \sup\{|f(x)| : x \in [a,b]\}$. For every $x, y \in [a,b]$ we have

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

This shows that F is Lipschitz continuous over $[a,b]$.

Now let f be continuous at $c \in [a,b]$. Let $\epsilon > 0$. Because f is continuous at c there exists a $\delta > 0$ such that for every $z \in [a,b]$ we have

$$|z - c| < \delta \implies \big|f(z) - f(c)\big| < \epsilon\,.$$

Because $f(c)$ is a constant, for every $x \in [a,b]$ such that $x \neq c$ we have

$$f(c) = \frac{1}{x - c} \int_c^x f(c)\,\mathrm{d}z\,.$$

It follows that

$$\begin{aligned} \frac{F(x) - F(c)}{x - c} - f(c) &= \frac{1}{x - c} \int_c^x f(z)\,\mathrm{d}z - \frac{1}{x - c} \int_c^x f(c)\,\mathrm{d}z \\ &= \frac{1}{x - c} \int_c^x \big(f(z) - f(c)\big)\,\mathrm{d}z\,. \end{aligned}$$

Therefore for every $x \in [a,b]$ we have

$$0 < |x-c| < \delta \implies$$
$$\left|\frac{F(x)-F(c)}{x-c} - f(c)\right| = \left|\frac{1}{x-c}\int_c^x \big(f(z)-f(c)\big)\,\mathrm{d}z\right| \le \frac{1}{|x-c|}\left|\int_c^x \big|f(z)-f(c)\big|\,\mathrm{d}z\right|$$
$$< \frac{\epsilon}{|x-c|}\left|\int_c^x \mathrm{d}z\right| = \frac{\epsilon}{|x-c|}\,|x-c| = \epsilon\,.$$

But this is the ϵ-δ characterization of

$$\lim_{x\to c}\frac{F(x)-F(c)}{x-c} = f(c)\,.$$

Hence, F is differentiable at c with $F'(c) = f(c)$.

The remainder of the proof is left as an exercise. □

Remark. Roughly speaking, the First and Second Fundamental Theorems of Calculus respectively state that

$$F(x) = F(a) + \int_a^x F'(t)\,\mathrm{d}t\,, \qquad f(x) = \frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\,\mathrm{d}t\,.$$

In words, the first states that integration undoes differentiation (up to a constant), while the second states that differentiation undoes integration. In other words, integration and differentiation are (nearly) inverses of each other. This is the realization that Newton and Leibniz had.

Remark. Newton and Leibniz were influenced by Barrow. He had proved the Second Fundamental Theorem for the special case where f was continuous and monotonic. This generalized Fermat's observation that the Second Fundamental Theorem holds for the power functions x^p, which are continuous and monotonic over $x > 0$. Of course, neither Barrow's statement nor his proof of this theorem were given in the notation we use today. Rather, they were given in a highly geometric setting that was commonly used at the time. This made it harder to see that his result could be generalized further. You can get an idea of what he did by assuming that f is nondecreasing and continuous over $[a,b]$ and drawing the picture that goes with the inequality

$$a \le x < y \le b \implies f(x) \le \frac{F(y)-F(x)}{y-x} \le f(y)\,,$$

where $F(x) = \int_a^x f$. By letting $y \to x$ while using the continuity of f, we obtain $F'(x) = f(x)$.

11.3. **Integration by Parts.** An important consequence of the First Fundamental Theorem of Calculus and the Product Rule for derivatives is the following lemma regarding integration by parts.

Proposition 11.1. (Integration by Parts) *Let $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ be Riemann integrable and have primitives F and G respectively over $[a,b]$. Then Fg and Gf are Riemann integrable over $[a,b]$ and*

$$\int_a^b Fg = F(b)G(b) - F(a)G(a) - \int_a^b Gf\,. \tag{11.3}$$

Proof. The functions F and G are Riemann integrable over $[a,b]$ because they are continuous. The functions Fg and Gf are therefore Riemann integrable over $[a,b]$ by the Product Lemma. Suppose we know that $FG : [a,b] \to \mathbb{R}$ is a primitive of $Fg + Gf$ over $[a,b]$. Then equation (11.3) follows from the First Fundamental Theorem of Calculus and the Additivity Lemma.

All that remains to be shown is that FG is a primitive of $Fg+Gf$ over $[a,b]$. It is clear that FG is continuous over $[a,b]$ because it is the product of continuous functions. Now let P and Q be the partitions of $[a,b]$ associated with the primitives F and G respectively. Let $R = P \vee Q$. Express R in terms of its partition points as $R = [r_0, \cdots, r_n]$. Then for every $i = 1, \cdots, n$ the function FG is differentiable over (r_{i-1}, r_i) with (by the Product Rule)

$$\begin{aligned} (FG)'(x) = F(x)G'(x) + G(x)F'(x) \quad &= F(x)g(x) + G(x)f(x) \\ = (Fg + Gf)(x) \quad &\text{for every } x \in (r_{i-1}, r_i)\,. \end{aligned}$$

Therefore FG is a primitive of $Fg + Gf$ over $[a,b]$. □

In the case where f and g are continuous over $[a,b]$ then the Second Fundamental Theorem of Calculus implies that f and g have primitives F and G that are continuously differentiable over $[a,b]$. In that case integration by parts reduces to the following.

Corollary 11.3. *Let $F : [a,b] \to \mathbb{R}$ and $G : [a,b] \to \mathbb{R}$ be continuously differentiable over $[a,b]$. Then*

$$\int_a^b FG' = F(b)G(b) - F(a)G(a) - \int_a^b GF'\,.$$

11.4. **Substitution.** An important consequence of the First Fundamental Theorem of Calculus and the Chain Rule for derivatives is the following proposition regarding changing the variable of integration in a definite integral by monotonic substitution $y = G(x)$.

Proposition 11.2. (Monotonic Substitution) *Let $g : [a,b] \to \mathbb{R}$ be Riemann integrable and have a primitive G that is increasing over $[a,b]$. Let $f : [G(a), G(b)] \to \mathbb{R}$ be Riemann integrable and have a primitive F over $[G(a), G(b)]$. Then if $f(G)\,g$ is Riemann integrable over $[a,b]$ we have the change of variable formula*

$$\int_{G(a)}^{G(b)} f = \int_a^b f(G)\,g\,. \tag{11.4}$$

Remark. If we show the variables of integration explicitly then the change of variable formula (11.4) takes the form

$$\int_{G(a)}^{G(b)} f(y)\,\mathrm{d}y = \int_a^b f(G(x))\,g(x)\,\mathrm{d}x\,.$$

Remark. The assumption that G is increasing over $[a,b]$ will hold if g is positive almost everywhere over $[a,b]$. Because G is a primitive, it is continuous as well as increasing. Its range is therefore the entire interval $[G(a), G(b)]$, the interval overwhich f and F are assumed to be defined. This insures the compositions $f(G)$ and $F(G)$ are defined over $[a,b]$.

Remark. The assumption that G is increasing over $[a,b]$ could have been replaced by the assumption that G is decreasing over $[a,b]$. In that case the interval $[G(b), G(a)]$ replaces the interval $[G(a), G(b)]$ in the hypotheses regarding f and F, but the change of variable formula (11.4) remains unchanged.

Proof. Suppose we know that $F(G) : [a, b] \to \mathbb{R}$ is a primitive of $f(G)g$ over $[a, b]$. Then because $f(G)g$ is Riemann integrable over $[a, b]$, the First Fundamental Theorem of Calculus implies

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

On the other hand, because f is Riemann integrable and F is a primitive of f over $[G(a), G(b)]$, the First Fundamental Theorem of Calculus also implies

$$\int_{G(a)}^{G(b)} f = F(G(b)) - F(G(a)) .$$

The change of variable formula (11.4) immediately follows from these last two equations.

All that remains to be shown is that $F(G)$ is a primitive of $f(G)g$ over $[a, b]$. It is clear that $F(G)$ is continuous over $[a, b]$ because it is the composition of continuous functions. Let $P = [p_0, \cdots, p_l]$ be the partition of $[G(a), G(b)]$ associated with the primitive F. Let $Q = [q_0, \cdots, q_m]$ be the partition of $[a, b]$ associated with the primitive G. Because $G : [a, b] \to [G(a), G(b)]$ is increasing, $G^{-1}(P) = [G^{-1}(p_0), \cdots, G^{-1}(p_l)]$ is a partition of $[a, b]$. Consider the partition $R = Q \vee G^{-1}(P)$ of $[a, b]$. Express R in terms of its partition points as $R = [r_0, \cdots, r_n]$. Then for every $i = 1, \cdots, n$ the function $F(G)$ is differentiable over (r_{i-1}, r_i) with (by the Chain Rule)

$$F(G)'(x) = F'(G(x))\, G'(x) = f(G(x))\, g(x) \quad \text{for every } x \in (r_{i-1}, r_i) .$$

Therefore $F(G)$ is a primitive of $f(G)g$ over $[a, b]$. □

Exercise. The assumption that G is increasing over $[a, b]$ in Proposition 11.2 can be weakened to the assumption that G is nondecreasing over $[a, b]$. Prove this slightly strengthend lemma. The proof can be very similar to the one given above, however you will have to work harder to show that $F(G)$ is a primitive of $f(G)g$ over $[a, b]$. Specifically, because G^{-1} may not exist, you will need to replace the partition $G^{-1}(P)$ in the above proof with a more complicated partition.

It is natural to ask whether we need a hypothesis like G is monotonic over $[a, b]$ in order to establish the change of variable formula (11.4). Indeed, we do not. However, without it we must take care to insure the compositions $f(G)$ and $F(G)$ are defined over $[a, b]$, to insure that $F(G)$ is a primitive of $f(G)g$ over $[a, b]$, and to insure that $f(G)g$ is Riemann integrable over $[a, b]$. Here we do this by assuming that f is continuous over an interval containing $\mathrm{Rng}(G)$.

Proposition 11.3. (Nonmonotonic Substitution) *Let $g : [a, b] \to \mathbb{R}$ be Riemann integrable and have a primitive G over $[a, b]$. Suppose that $\mathrm{Rng}(G) \subset [\underline{m}, \overline{m}]$ and let $f : [\underline{m}, \overline{m}] \to \mathbb{R}$ be continuous over $[\underline{m}, \overline{m}]$. Then the change of variable formula (11.4) holds.*

Proof. By the Second Fundamental Theorem of Calculus f has a continuously differentiable primitive F over $[\underline{m}, \overline{m}]$. It is then easy to show that $F(G)$ is a primitive of $f(G)g$ over $[a, b]$. Because $f(G)$ is continuous (hence, Riemann integrable) while g is Riemann integrable over $[a, b]$, it follows from the Product Lemma that $f(G)g$ is Riemann integrable over $[a, b]$. The rest of the proof proceeds as in the proof of Proposition 11.2. □

11.5. **Integral Mean-Value Theorem.** We will now give a theorem that a first glance may not seem to have a connection either with the Fundamental Theorems of Calculus or with a Mean-Value Theorem for differentiable functions. However, we will see there is a connection.

Theorem 11.3. (Integral Mean-Value) *Let $f : [a, b] \to \mathbb{R}$ be continuous. Let $g : [a, b] \to \mathbb{R}$ be Riemann integrable and positive almost everywhere over $[a, b]$. Then there exists a point $p \in (a, b)$ such that*

$$\int_a^b fg = f(p) \int_a^b g \,. \tag{11.5}$$

Remark. The connection of this theorem to both the First and Second Fundamental Theorem of Calculus and to the Cauchy Mean-Value Theorems for differentiable functions is seen when both f and g are continuous. Then by the Second Fundamental Theorem of Calculus fg and g have continuously differentiable primitives F and G. The Cauchy Mean-Value Theorem applied to F and G then yields a $p \in (a, b)$ such that

$$F(b) - F(a) = \frac{F'(p)}{G'(p)} \big(G(b) - G(a) \big) = f(p) \big(G(b) - G(a) \big) \,.$$

By the First Fundamental Theorem of Calculus we therefore have

$$\int_a^b fg = F(b) - F(a) = f(p) \big(G(b) - G(a) \big) = f(p) \int_a^b g \,.$$

In other words, when both f and g are continuous the Integral Mean-Value Theorem is just the Cauchy Mean-Value Theorem for differentiable functions applied to primitives of fg and g.

Remark. A simpler version of the Integral Mean-Value Theorem only considers the case $g(x) = 1$. In that case, if $f : [a, b] \to \mathbb{R}$ is continuous, there exists a point $p \in (a, b)$ such that

$$f(p) = \frac{1}{b - a} \int_a^b f \,.$$

This is proved by simply applying the Lagrange Mean-Value Theorem to a primitive of f. If we interpret the right-hand side above as the average of f over the interval $[a, b]$ then the theorem asserts that f takes on its average value.

Exercise. Show that Theorem 11.3 does not hold when we replace the hypothesis that f is continuous over $[a, b]$ with the hypothesis that f is Riemann integrable over $[a, b]$.

We now turn to the proof of Theorem 11.3. Because we are not assuming that g is continuous, the proof will take a different form from the ones indicated in the remarks above.

Proof. Because f is continuous over $[a, b]$, the Extreme-Value Theorem implies there exists points $\underline{x}$ and $\overline{x} \in [a, b]$ such that

$$f(\underline{x}) = \inf \big\{ f(x) \,:\, x \in [a, b] \big\} \,, \qquad f(\overline{x}) = \sup \big\{ f(x) \,:\, x \in [a, b] \big\} \,.$$

Then

$$f(\underline{x}) \leq f(x) \leq f(\overline{x}) \quad \text{for every } x \in [a, b] \,,$$

which, because g is nonnegative, implies that

$$f(\underline{x}) \int_a^b g \leq \int_a^b fg \leq f(\overline{x}) \int_a^b g \,.$$

If $f(\underline{x}) = f(\overline{x})$ then f is constant and (11.5) holds for every $p \in (a, b)$.

Now consider the case $f(\underline{x}) < f(\overline{x})$. Because f is continuous there exists $[c, d] \subset [a, b]$ such that $\underline{x} \in [c, d]$ and that $f(x) < \frac{1}{2}(f(\underline{x}) + f(\overline{x}))$. Then

$$f(\overline{x}) - f(x) > \tfrac{1}{2}\big(f(\overline{x}) - f(\underline{x})\big) > 0 \quad \text{for every } x \in (c, d)\,.$$

Because $(f(\overline{x}) - f)g \geq 0$, and because $(f(\overline{x}) - f(x))g(x) > 0$ almost everywhere over the nonempty interval (c, d), the Positivity Proposition implies

$$0 < \int_a^b \big(f(\overline{x}) - f\big)g = f(\overline{x}) \int_a^b g - \int_a^b fg\,.$$

In a similar manner we can argue that

$$0 < \int_a^b fg - f(\underline{x}) \int_a^b g\,.$$

Because g is positive almost everywhere over $[a, b]$, the Positivity Proposition also implies that $\int_a^b g > 0$. Therefore, we see that

$$f(\underline{x}) < \frac{\int_a^b fg}{\int_a^b g} < f(\overline{x})\,.$$

Because f is continuous, the Intermediate-Value Theorem implies there exists a p between $\underline{x}$ and $\overline{x}$ such that (11.5) holds. $\square$

11.6. **Cauchy Riemainder Theorem.** Recall that if f is n-times differentiable over an interval (a, b) and $c \in (a, b)$ then the n^{th} Taylor polynomial approximation of f at c is given by

$$T_c^n f(x) = \sum_{k=0}^{n} f^{(k)}(c)\, \frac{(x - c)^k}{k!}\,. \tag{11.6}$$

Recall too that if f is $(n+1)$-times differentiable over the interval (a, b) then the Lagrange Remainder Theorem states that for every $x \in (a, b)$ there exists a point p between c and x such that

$$f(x) = T_c^n f(x) + f^{(n+1)}(p)\, \frac{(x - c)^{n+1}}{(n+1)!}\,. \tag{11.7}$$

Our proof of the Lagrange Remainder Theorem was based on a direct application of the Lagrange Mean-Value Theorem.

Here we give an alternative representation of the remainder due to Cauchy. Its proof is based on a direct application of the First Fundemental Theorem of Calculus, the proof of which also rests on the Lagrange Mean-Value Theorem. We will see that the resulting representation contains more information than that of Lagrange.

Theorem 11.4. (Cauchy Remainder Theorem) *Let $f : (a, b) \to \mathbb{R}$ be $(n+1)$-times differentiable over (a, b) and let $f^{(n+1)}$ be Riemann integrable over every closed subinterval of (a, b). Let $c \in (a, b)$. Then for every $x \in (a, b)$ we have*

$$f(x) = T_c^n f(x) + \int_c^x f^{(n+1)}(t)\, \frac{(x - t)^n}{n!}\, \mathrm{d}t\,. \tag{11.8}$$

Proof. Let $x \in (a,b)$ be fixed. Then define $F : (a,b) \to \mathbb{R}$ by

$$F(t) = T_t^n f(x) = f(t) + \sum_{k=1}^{n} f^{(k)}(t)\, \frac{(x-t)^k}{k!} \quad \text{for every } t \in (a,b)\,. \tag{11.9}$$

Clearly F is differentiable over (a,b) with (notice the telescoping sum)

$$F'(t) = f'(t) + \sum_{k=1}^{n} \left[f^{(k+1)}(t)\, \frac{(x-t)^k}{k!} - f^{(k)}(t)\, \frac{(x-t)^{k-1}}{(k-1)!} \right] = f^{(n+1)}(t)\, \frac{(x-t)^n}{n!}\,.$$

Because c and x are in (a,b) and because $f^{(n+1)}$ (and hence F') is Riemann integrable over every closed subinterval of (a,b), the First Fundamental Theoren of Calculus yields

$$F(x) - F(c) = \int_c^x F'(t)\, \mathrm{d}t = \int_c^x f^{(n+1)}(t)\, \frac{(x-t)^n}{n!}\, \mathrm{d}t\,. \tag{11.10}$$

However, it is clear from definition (11.9) of $F(t)$ that

$$F(x) = f(x)\,, \quad \text{while} \quad F(c) = T_c^n f(x)\,.$$

Formula (11.8) therefore follows from (11.10). $\square$

Remark. The Lagrange remainder formula can be derived from Cauchy's if we assume that $f^{(n+1)}$ is continuous over (a,b). In that case the Integral Mean-Value Theorem implies that for each $x \in (a,b)$ there exists a point p between c and x such that

$$\int_c^x f^{(n+1)}(t)\, \frac{(x-t)^n}{n!}\, \mathrm{d}t = f^{(n+1)}(p) \int_c^x \frac{(x-t)^n}{n!}\, \mathrm{d}t\,.$$

A direct calculation then shows that

$$\int_c^x \frac{(x-t)^n}{n!}\, \mathrm{d}t = \frac{(x-c)^{n+1}}{(n+1)!}\,,$$

whereby

$$\int_c^x f^{(n+1)}(t)\, \frac{(x-t)^n}{n!}\, \mathrm{d}t = f^{(n+1)}(p)\, \frac{(x-c)^{n+1}}{(n+1)!}\,. \tag{11.11}$$

When this is placed into Cauchy's formula (11.8) we obtain Lagrange's formula (11.7). This argument is not an alternative proof of the Lagrange Remainder Theorem because it assumes that $f^{(n+1)}$ is continuous over (a,b), whereas the Lagrange Remainder Theorem only assumes that $f^{(n+1)}$ exists over (a,b).

Remark. We cannot derive the Cauchy remainder formula from that of Lagrange, even with additional regularity assumptions on f. This is because the Lagrange Remainder formula only tells you that the point p appearing in (11.7) lies between c and x while the Cauchy formula provides you with the explicit formula (11.8) for the remainder. This additional information arises because the Cauchy Remainder Theorem assumes that $f^{(n+1)}$ is Riemann intergrable over (a,b), whereas the Lagrange Remainder Theorem only assumes that $f^{(n+1)}$ exists over (a,b).

Remark. The only way to bound the Taylor remainder using the Lagrange formula (11.7) is to use uniform bounds on $f^{(n+1)}(p)$ over all p that lie between c and x. While this approach is sufficient for some tasks (like showing that the formal Taylor series of e^x, $\cos(x)$, and $\sin(x)$ converge to those functions for every $x \in \mathbb{R}$), it fails for other tasks. However, if you are able to obtain suitable pointwise bounds on the intrgrand in some form of the Cauchy formula (11.8) then you can sometimes obtain bounds on the Taylor remainder that are sufficient for those tasks. This remark is illustrated by the following example.

Example. Let $f(x) = \log(1+x)$ for every $x > -1$. Then

$$f^{(k)}(x) = (-1)^{k-1}\frac{(k-1)!}{(1+x)^k} \quad \text{for every } x > -1 \text{ and } k \in \mathbb{Z}_+\,.$$

The formal Taylor aeries of f about 0 is therefore

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k\,.$$

The Ratio Test shows that this series converges absolutely for $|x| < 1$ and diverges for $|x| > 1$. For $x = -1$ the series is the negative of the harmonic series, and therefore diverges. For $x = 1$ the Alternating Series Test shows the series converges. We therefore conclude that the series converges if and only if $x \in (-1, 1]$. These arguments do not show however that the sum of the series is $f(x)$. This requires showing that for every $x \in (-1, 1]$ the Taylor remainder $f(x) - T_0^n f(x)$ vanishes as $n \to \infty$.

First let us approach this problem using the Lagrange form of the remainder (11.7): there exists a p between 0 and x such that

$$f(x) - T_0^n f(x) = f^{(n+1)}(p)\,\frac{x^{n+1}}{(n+1)!} = \frac{(-1)^n}{n+1}\left(\frac{x}{1+p}\right)^{n+1}.$$

If $x \in (0, 1]$ then $p \in (0, x)$ and we obtain the bound

$$\big|f(x) - T_0^n f(x)\big| < \frac{1}{n+1}\, x^{n+1}\,.$$

This bound clearly vanishes as $n \to \infty$ for every $x \in (0, 1]$. On the other hand, if $x \in (-1, 0)$ then $p \in (-|x|, 0)$ and we obtain the bound

$$\big|f(x) - T_0^n f(x)\big| < \frac{1}{n+1}\left(\frac{|x|}{1-|x|}\right)^{n+1}.$$

This bound will only vanish as $n \to \infty$ for $x \in [-\frac{1}{2}, 0)$. This approach leaves the question open for $x \in (-1, -\frac{1}{2})$.

Now let us approach this problem using the Cauchy form of the remainder (11.8):

$$\begin{aligned} f(x) - T_0^n f(x) &= \int_0^x f^{(n+1)}(t)\,\frac{(x-t)^n}{n!}\,\mathrm{d}t \\ &= (-1)^n \int_0^x \frac{(x-t)^n}{(1+t)^{n+1}}\,\mathrm{d}t = (-1)^n \int_0^x \left(\frac{x-t}{1+t}\right)^n \frac{\mathrm{d}t}{1+t}\,. \end{aligned}$$

Consider the substitution

$$s = \frac{x-t}{1+t} = \frac{1+x}{1+t} - 1\,, \qquad t+1 = \frac{1+x}{1+s}\,, \qquad \frac{\mathrm{d}t}{1+t} = -\frac{\mathrm{d}s}{1+s}\,.$$

Notice that s goes monotonically from x to 0 as t goes monotonically from 0 to x. This substitution yields

$$f(x) - T_0^n f(x) = (-1)^n \int_0^x \frac{s^n}{1+s}\,\mathrm{d}s\,.$$

Because $|x| < 1$ and because

$$\frac{1}{1+s} \leq \frac{1}{1-|s|} \quad \text{for every } s \in (-1,1)\,,$$

we obtain the bound

$$\begin{aligned}\big|f(x) - T_0^n f(x)\big| &= \left| \int_0^x \frac{s^n}{1+s}\,\mathrm{d}s \right| \\ &\leq \int_0^{|x|} \frac{s^n}{1-s}\,\mathrm{d}s \leq \frac{1}{1-|x|} \int_0^{|x|} s^n\,\mathrm{d}s = \frac{1}{1-|x|}\,\frac{|x|^{n+1}}{n+1}\,.\end{aligned}$$

This bound clearly vanishes as $n \to \infty$ for every $x \in (-1,1)$.

Collecting all of our results, we have shown that

$$\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}\,x^k \quad \text{for every } x \in (-1,1]\,,$$

and that the series diverges for all other values of x.

Exercise. When $q \in \mathbb{N}$ the binomial expansion yields

$$(1+x)^q = \sum_{k=0}^{q} \frac{q!}{k!(q-k)!}\,x^k = 1 + \sum_{k=1}^{q} \frac{q(q-1)\cdots(q-k+1)}{k!}\,x^k\,.$$

Now let $q \in \mathbb{R} - \mathbb{N}$. Let $f(x) = (1+x)^q$ for every $x > -1$. Then

$$f^{(k)}(x) = q(q-1)\cdots(q-k+1)(1+x)^{q-k} \quad \text{for every } x > -1 \text{ and } k \in \mathbb{Z}_+\,.$$

The formal Taylor series of f about 0 is therefore

$$1 + \sum_{k=1}^{\infty} \frac{q(q-1)\cdots(q-k+1)}{k!}\,x^k\,.$$

Show that this series converges absolutely to $(1+x)^q$ when $|x| < 1$ and diverges when $|x| > 1$. (This formula is Newton's extension of the binomial expansion to powers q that are real.)

Exercise. Show that for every $q > -1$ we have

$$2^q = 1 + \sum_{k=1}^{\infty} \frac{q(q-1)\cdots(q-k+1)}{k!}\,,$$

while for every $q \leq -1$ the above series diverges. (This is the case $x = 1$ for the series in the previous problem.)

The Cauchy Remainder Theorem 11.4 gives another representation of the Taylor remainder that has fewer regularity requirements on f.

Corollary 11.4. (Alternative Cauchy Remainder) *Let $f : (a,b) \to \mathbb{R}$ be n-times differentiable over (a,b) and let $f^{(n)}$ be Riemann integrable over every closed subinterval of (a,b). Let $c \in (a,b)$. Then for every $x \in (a,b)$ we have*

$$f(x) = T_c^n f(x) + \int_c^x \big(f^{(n)}(t) - f^{(n)}(c)\big)\,\frac{(x-t)^{n-1}}{(n-1)!}\,\mathrm{d}t\,. \tag{11.12}$$

Proof. Exercise. ☐

This corollary allows us to bound the Taylor remainder in cases when f is not $(n+1)$-times differentiable at c. Specifically, we have the following bound on the remainder when $f^{(n)}$ is only Hölder continuous at c.

Proposition 11.4. (Taylor Remainder Bound) *Let $f : (a, b) \to \mathbb{R}$ be n-times differentiable over (a, b) and let $f^{(n)}$ be Riemann integrable over every closed subinterval of (a, b). Let $c \in (a, b)$. Let $\alpha \in (0, 1]$ and $K \in (0, \infty)$ such that $f^{(n)}$ satisfies the Hölder bound*

$$\left|f^{(n)}(t) - f^{(n)}(c)\right| \leq K|t - c|^{\alpha} \quad \text{for every } t \in (a, b)\,. \tag{11.13}$$

Then for every $x \in (a, b)$ we have

$$\begin{aligned}\left|f(x) - T_c^n f(x)\right| &\leq \frac{K}{(n-1)!}\int_0^1 (1-s)^{n-1}s^{\alpha}\,\mathrm{d}s\,|x-c|^{n+\alpha} \\ &= \frac{K}{(\alpha+n)(\alpha+n-1)\cdots(\alpha+1)\alpha}\,|x-c|^{n+\alpha}\,.\end{aligned} \tag{11.14}$$

Proof. Exercise. (Hint: The last integral can be evaluated using integration by parts.) ☐

Remark. This bound is better than the one obtained in Chapter 7.