Chapter 8

Riemann Integration

8.1 Introduction

The notion of integral calculus is closely related to the notion of area. The earliest evidence of integral calculus can be found in the works of Greek geometers who employed the Method of Exhaustion to define and find the area of planar regions having circular or parabolic boundaries. When Newton and Leibniz developed calculus, they also study integration. However, they considered integration as the inverse process of differentiation. The formula we know and can prove, \( \int_{a}^{b} f(x) \, dx = F(b) - F(a) \), where \( F \) is any function whose derivative is \( f \) was their definition of the integral. This definition remained until the 1820’s.

The modern approach to integration is due to Cauchy. He was the first to construct a theory of integration based on approximating the area below a curve. His definition of the integral did not involve differential calculus. Cauchy’s definition of the integral was very restrictive in the sense that only continuous functions, or functions having a finite number of discontinuities could be integrated. Cauchy was able to prove the fundamental theorem of calculus. The modern definition of the integral is attributed to the German mathematician Bernhard Riemann (1826-1866). While working on trying to characterize which functions were integrable using Cauchy’s definition, he modified Cauchy’s definition and developed the theory of integration which bears his name. He was also able to find necessary and sufficient conditions for a real valued bounded function to be integrable. The Dutch mathematician Thomas Jan Stieltjes (1856-1894) improved on the notion developed by Riemann. Though his improvements involved only minor modifications of Riemann’s definition, the consequences of these modifications are far reaching. At the beginning of the twentieth century, the French mathematician Henri Lebesgue (1875-1942) introduced the new notion of measure of a set which paved the way to the foundations of the modern theory of integration.

Historically, Riemann integration came before Lebesgue integration. In this class, we will study Riemann integration. We will do so mostly because Lebesgue
integration is more difficult, requires more knowledge and is, for that reason, usually the subject of a graduate course in real analysis. However, the reader should not think that these are two different theory. Lebesgue integration is more general that Riemann integration. There are more functions which are Lebesgue integrable than Riemann integrable. However, for functions which are both, the two integrals have the same value.

Before getting deeply into the subject, it is necessary to develop additional terminology and concepts. We will then start by defining the Riemann integral using a modified approach of Riemann’s original definition. This modified approach is due to the French mathematician Jean Gaston Darboux (1842-1917). We will call this integral the Riemann-Darboux integral. We will then study Riemann’s original definition and show the integral obtained with it is the same as the Riemann-Darboux integral. We will then study important theorems and properties about the Riemann integral.

8.2 Definition of the Riemann-Darboux Integral

In this chapter, we discuss the Riemann integral of a function \( f \) over the interval \([a, b]\). In this chapter we assume that \( f : [a, b] \to \mathbb{R} \) is defined and bounded on \([a, b]\).

8.2.1 Notion of a Partition

Definition 482 (Partition) Let \( a, b \in \mathbb{R} \) with \( a \leq b \).

1. A partition \( P \) of \([a, b]\) is a finite set of points \( P = \{x_0, x_1, x_2, ..., x_n\} \) where \( a = x_0 < x_1 < x_2 < ... < x_n = b \).

2. The width of the \( i \)th interval will be denoted \( d_i \). In other words, \( d_i = x_i - x_{i-1} \) for each \( i = 1, 2, ..., n \).

The basic idea is to divide the interval \([a, b]\) into a finite collection of subintervals. In the above definition, we have \( n + 1 \) points and \( n \) subintervals. The \( i \)th interval is the interval \([x_{i-1}, x_i]\). Not all the intervals need to have the same length.

Definition 483 (norm of a partition) Let \( P = \{x_0, x_1, x_2, ..., x_n\} \) be a partition of \([a, b]\). The mesh or norm of \( P \), denoted \( \|P\| \) is defined by:

\[
\|P\| = \max \{d_1, d_2, ..., d_n\}
\]

Remark 484 The following should be obvious to the reader:

1. If all the subintervals have equal length, then \( \|P\| = \frac{b - a}{n} \)

2. \( \sum_{i=1}^{n} d_i = b - a \)
3. The minimum number of subintervals any partition must have is \( \frac{b - a}{\|P\|} \)
and therefore, the minimum number of points such a partition must have is \( \frac{b - a}{\|P\|} + 1 \).

**Example 485** Find a partition of \([-3, 4]\) having norm \( \frac{1}{10} \).
From the above remark, we see that such a partition must have at least 71 members. When we select our members, the following restrictions apply:

1. \( x_0 = -3, \ x_n = 4 \) and \( n \geq 70 \).

2. For any \( i = 1..n \), we must have \( x_i - x_{i-1} \leq \frac{1}{10} \). We must have equality for at least one value of \( i \).

The are infinitely many ways to find such a partition. The simplest way would be to create a partition where all the subintervals would have equal length, that length would be \( \frac{1}{10} \). Another way is as follows: Let \( x_0 = -3, \ x_1 = -2.9 \) (this ensures that one subintervals has a width of \( \frac{1}{10} \). Then, pick \( \delta \) such that \( 0 < \delta \leq \frac{1}{10} \) and set \( x_i = \begin{cases} x_{i-1} + \delta & \text{if } x_{i-1} + \delta < 4 \\ 4 & \text{otherwise} \end{cases} \).

Then, \( n \) is the last \( i \) such that \( x_i = 4 \).

**Definition 486** Let \( P \) and \( Q \) be two partitions of \([a, b]\). We say that \( Q \) is a refinement of \( P \) if \( P \subseteq Q \).

**Remark 487** Saying that \( Q \) is a refinement of \( P \) means that \( Q \) contains all the points of \( P \) and some extra points. One way to obtain a refinement of a partition is simply to add some points to that partition.

**Remark 488** If \( P \) and \( Q \) are two partitions of \([a, b]\) then \( P \cup Q \) is a refinement of both \( P \) and \( Q \) and \( P \) and \( Q \) are refinements of \( P \cap Q \).

**Proposition 489** If \( Q \) is a refinement of \( P \), then \( \|Q\| \leq \|P\| \).

**Proof.** See homework.

### 8.2.2 Darboux Sums

Let \( P = \{x_0, x_1, \ldots, x_n\} \) be any partition of \([a, b]\).

Since \( f \) is bounded on \([a, b]\), then it has both a supremum and an infimum on \([a, b]\). Thus, we can define:

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

and

\[
m = \inf \{f(x) : x \in [a, b]\}
\]
Also, if \( f \) is bounded on \([a, b]\), then it is bounded on each subinterval \([x_{i-1}, x_i]\). Therefore, we can define

\[
M_i = \sup \{ f(x) : x \in [x_{i-1}, x_i] \}
\]

and

\[
m_i = \inf \{ f(x) : x \in [x_{i-1}, x_i] \}
\]

**Definition 490 (Darboux Sums)** The upper Darboux sum of the function \( f \), corresponding to the partition \( P \), denoted \( \mathcal{S}(f, P) \), is defined by:

\[
\mathcal{S}(f, P) = \sum_{i=1}^{n} M_i d_i.
\]

The lower Darboux sum of the function \( f \), corresponding to the partition \( P \), denoted \( \mathcal{S}(f, P) \), is defined by:

\[
\mathcal{S}(f, P) = \sum_{i=1}^{n} m_i d_i.
\]

**Remark 491** Darboux sums are very important to what is about to follow. Before we continue, let us make several important remarks.

1. Each Darboux sum depends on the function \( f \) and the partition \( P \). This is why we denote them \( \mathcal{S}(f, P) \) and \( \mathcal{S}(f, P) \). Since every partition has a finite number of subintervals, and \( f \) is bounded on each subinterval, every Darboux sum is finite.

2. In the case \( f \) is a positive function on \([a, b]\), the upper Darboux sum corresponds to the sum of the areas of the rectangles above the graph of \( f \). The lower Darboux sum corresponds to the sum of the areas of the rectangles below the graph of \( f \). Thus, these Darboux sums will give an upper and a lower bound for the area below the graph of \( f \).

3. When we are dealing with one bounded function \( f \), we can omit it in the notation for the Darboux sum and simply write \( \mathcal{S}(P) \) and \( \mathcal{S}(P) \).

**Lemma 492** If \( P = \{a, b\} \) then \( \mathcal{S}(P) = M(b - a) \) and \( \mathcal{S}(P) = m(b - a) \). Otherwise, if \( P \) is an arbitrary partition of \([a, b]\) then

\[
m(b - a) \leq \mathcal{S}(P) \leq \mathcal{S}(P) \leq M(b - a)
\]

**Proof.** See exercises.

This lemma establishes the fact that for a fixed bounded function, the collection of all the upper sums as well as the collection of all the lower sums is bounded below and above. This fact will prove to be crucial for the development of our theory.
Lemma 493 Let \( f \) and \( g \) be defined on \([a, b]\), and \( P \) be a partition of \([a, b]\).

Then

1. \( \overline{S}(af, P) = a\overline{S}(f, P) \) for any constant \( a \geq 0 \).
2. \( \overline{S}(af, P) = a\overline{S}(f, P) \) for any constant \( a < 0 \).
3. \( |\overline{S}(f, P)| \leq \overline{S}(|f|, P) \)
4. \( \overline{S}(f, P) + \overline{S}(g, P) \geq \overline{S}(f + g, P) \)

Proof. See exercises. \( \blacksquare \)

8.2.3 Riemann-Darboux Integral

Definition and Examples

For various partitions \( P \) of \([a, b]\) and a fixed bounded function \( f \), we set

\[
\{ \overline{S}(P) \} = \{ \overline{S}(f, P) \} = \{ \overline{S}(f, P) : P \text{ is a partition of } [a, b] \}
\]

and

\[
\{ \underline{S}(P) \} = \{ \underline{S}(f, P) \} = \{ \underline{S}(f, P) : P \text{ is a partition of } [a, b] \}
\]

It is important to understand that these sets are taken over all the possible partitions of \([a, b]\). As noticed above, they are both bounded above by \( M(b - a) \) and below by \( m(b - a) \). Thus, each set has a supremum and an infimum. Therefore, we define:

Definition 494 (Riemann-Darboux integral) Let \( f \) be bounded on \([a, b]\).

1. The upper Darboux integral of \( f \) on \([a, b]\), denoted \( \overline{\int}_a^b f \), is defined by
   \[
   \overline{\int}_a^b f = \inf \{ \overline{S}(P) \}.
   \]
2. The lower Darboux integral of \( f \) on \([a, b]\), denoted \( \underline{\int}_a^b f \), is defined by
   \[
   \underline{\int}_a^b f = \sup \{ \underline{S}(P) \}.
   \]
3. If \( \overline{\int}_a^b f = \underline{\int}_a^b f \), the common value is called the Riemann-Darboux integral (R-D integral for short) of the function \( f \) on the interval \([a, b]\) and is denoted \( \int_a^b f \). When this common value exists, we say that \( f \) is Riemann-Darboux integrable (or R-D integrable) on \([a, b]\).

Remark 495 From the definition, the following should be obvious.

\[
\overline{S}(P) \leq \overline{\int}_a^b f
\]

and

\[
\underline{\int}_a^b f \leq \overline{S}(P)
\]
We now need to understand what kind of functions will be R-D integrable. We also need to learn how to compute the Riemann-Darboux integral. The next results and examples will help us achieve that.

**Theorem 496** If \( f \) is a bounded function on \([a, b]\) and \( Q \) is a refinement of a partition \( P \) on \([a, b]\) then
\[
\underline{\mathcal{S}}(P) \leq \underline{\mathcal{S}}(Q) \leq \mathcal{S}(Q) \leq \mathcal{S}(P)
\]

**Proof.** See exercises (hint: First, prove the result assuming \( Q \) is obtained from \( P \) by adjoining one more point. Then, use that result on a sequence of partitions \( Q_1, Q_2, Q_3, \ldots \) where \( Q_i \) is obtained from \( Q_{i-1} \) by adding one more point.). ■

This theorem was about a partition and its refinement. The next result is about any two partitions.

**Theorem 497** If \( f \) is a bounded function on \([a, b]\) and \( P \) and \( Q \) are any two partitions of \([a, b]\), then
\[
\underline{\mathcal{S}}(Q) \leq \mathcal{S}(P)
\]

**Proof.** See exercises. (hint: use the previous theorem and the fact that \( R = P \cup Q \) is a refinement of both \( P \) and \( Q \)). ■

**Theorem 498** If \( f \) is a bounded function on \([a, b]\) then
\[
\int_a^b f \leq \mathcal{S}(P)
\]

**Proof.** See exercises. Hint: use the previous theorem. ■

This result is simply the analytic form of what is obvious geometrically. Though the results obtained above seem simple, we must not forget that they uses a very powerful principle: the supremum principle.

**Remark 499** Combining the inequality of the above theorem and those in remark 495, we have
\[
\underline{\mathcal{S}}(P) \leq \int_a^b f \leq \mathcal{S}(P)
\]

**Example 500** Discuss the Riemann-Darboux integrability of \( f(x) = k \) on \([a, b]\) where \( k \) is a constant.

We need to prove that \( \int_a^b f = \overline{\mathcal{S}}(P) \). To compute these integrals, we need to find \( \underline{\mathcal{S}}(P) \) and \( \mathcal{S}(P) \) for any partition \( P \). If \( P \) is any partition of \([a, b]\), then \( m = m_i = M = M_i = k \) (since the function is constant). Therefore, \( \mathcal{S}(P) = \sum_{i=1}^{n} k d_i = k \sum_{i=1}^{n} d_i = k (b - a) \). Similarly, \( \underline{\mathcal{S}}(P) = \sum_{i=1}^{n} k d_i = k \sum_{i=1}^{n} d_i = k (b - a) \). Hence, for any partition \( P \), \( \underline{\mathcal{S}}(P) = k (b - a) \) and therefore does not depend on \( P \). It follows that \( \inf \{ \mathcal{S}(P) \} = k (b - a) \) and therefore \( \int_a^b f = k (b - a) \). Similarly, we see that \( \int_a^b f = k (b - a) \).
Example 501 Discuss the R-D integrability of $f$ defined on $[a, b]$ by

$$f(x) = \begin{cases} 
  k > 0 & \text{if } x \neq c \in [a, b] \\
  0 & \text{if } x = c
\end{cases}$$

From what we did above, we see that if $P$ is any partition of $[a, b]$, then the components of $\overline{S}(P)$ will agree with the components of $\underline{S}(P)$ except on the intervals which contain $c$ (there can be at most two). Suppose that $c \in [x_{i-1}, x_i]$. The contribution from this subinterval to $\overline{S}(P)$ will be zero, while the contribution to $\underline{S}(P)$ will be $kd_i$. It follows that $\overline{S}(P) \leq \underline{S}(P) + 2k \|P\|$. In other words,

$$|\overline{S}(P) - \underline{S}(P)| \leq 2k \|P\|$$

Since $\|P\|$ can be made as small as we want, we conclude that for any $\epsilon > 0$, there is a partition $P$ such that

$$|\overline{S}(P) - \underline{S}(P)| \leq \epsilon$$

Since

$$\left| \int_a^b f - \int_a^b f \right| \leq |\overline{S}(P) - \underline{S}(P)|$$

it follows that for any $\epsilon > 0$,

$$\left| \int_a^b f - \int_a^b f \right| \leq \epsilon$$

From a theorem studied in class, it follows that $\int_a^b f = \int_a^b f$ thus $f$ is R-D integrable.

Remark 502 The function in the example above was identical to the function in the example before except at one point. They were both integrable. In fact, later on, we will prove a much more powerful result. If a function $g$ is obtained from a function $f$ by altering a finite number of values of $f$ then if $f$ is R-D integrable, so is $g$ and the two integrals have the same value.

Example 503 Discuss the R-D integrability of $f$ defined on $[a, b]$ by

$$f(x) = \begin{cases} 
  1 & \text{if } x \text{ is irrational} \\
  0 & \text{if } x \text{ is rational}
\end{cases}$$

Let $P$ be any partition of $[a, b]$. Because any subinterval of $[a, b]$ contains both irrational and rational numbers, it follows that $m_i = 0$ and $M_i = 1$. Therefore, $\overline{S}(P) = b - a$ and $\underline{S}(P) = 0$ for any partition $P$ of $[a, b]$. It follows that $\int_a^b f = 0$ and $\int_a^b f = b - a$ therefore $f$ is not R-D integrable on $[a, b]$.

Remark 504 This establishes the fact that not all bounded functions are R-D integrable. It also suggests that continuity will play an important role. There was an important difference between the last example and the one before. In the last example, no matter which partition we selected, we could not make $\overline{S}(P) - \underline{S}(P)$ as small as we wanted. We could do that in the previous one. This fact is captured in the next result.
Integrability Theorems

We now look at several theorems about integrability of functions. The first theorem, attributed to Riemann, gives a necessary and sufficient condition for a function to be integrable.

**Theorem 505** A bounded function $f$ on $[a, b]$ is R-D integrable on $[a, b]$ if and only if for every $\epsilon > 0$, there is a partition $P$ of $[a, b]$ such that

$$\mathcal{S}(P) - \mathcal{S}(P) < \epsilon$$

**Proof.** The proof is assigned as an exercise. We give an outline of the proof here. This is an "if and only if", so we must prove both directions.

- $(\Leftarrow)$ Use the definition of R-D integrability and the fact that $\mathcal{S}(P) \leq \int_a^b f \leq \mathcal{S}(P)$.

- $(\Rightarrow)$ We break the proof in several steps:

  1. Use the fact that if $H$ is a non-empty set of real numbers, then $b$ is the supremum of $H$ if and only if for every positive $\epsilon$ there exists $x \in H$ such that $b - \epsilon < x \leq b$ (this is another form of theorem 151 last semester) to prove that $\int_a^b f - \frac{\epsilon}{2} < \mathcal{S}(Q)$ for some partition $Q$.

  2. Similarly, prove that $\mathcal{S}(R) < \int_a^b f + \frac{\epsilon}{2}$ for some partition $R$.

  3. Apply both inequalities to the refinement $P = Q \cup R$ and conclude using the fact that $f$ is R-D integrable.

The remaining theorems give either necessary conditions or sufficient conditions.

**Theorem 506** If $f$ is continuous on $[a, b]$ then $f$ is R-D integrable on $[a, b]$.

**Proof.** The proof is assigned as an exercise. We give an outline of the proof here. We break the proof in several steps.

1. Using the previous theorem, it is enough to show that $\mathcal{S}(P) - \mathcal{S}(P) < \epsilon$.

2. Use the fact that if $f$ is continuous on $[a, b]$, then it is uniformly continuous hence we can make $|f(s) - f(t)|$ as small as we want. Of course, you will have to determine how small we want it to show that $f$ is R-D integrable.

3. From the previous step and properties of continuous functions on a closed interval, conclude that we can make $M_i - m_i$ as small as we want. Of course, you will have to determine how small we want it to show that $f$ is R-D integrable. Here, I am using the notation of the notes.
4. Using the estimate of the previous step, compute \( S(P) - \underline{S}(P) \) and show \( S(P) - \underline{S}(P) < \epsilon \).

---

**Theorem 507** A monotonic function on \([a, b]\) is R-D integrable.

**Proof.** The proof is assigned as an exercise. We give an outline of the proof here in the case \( f \) is monotone increasing. The case \( f \) monotone decreasing is similar.

- **Case 1:** \( f(a) = f(b) \). What does this imply? You should see that we’ve already done it.
- \( f(b) > f(a) \). We prove the result by proving that \( S(P) - \underline{S}(P) < \epsilon \). We break the proof in several steps.

  1. Prove that \( \sum_{i=1}^{n} (M_i - m_i) = f(b) - f(a) \).
  2. If \( P \) is a partition with norm less that a constant you will have to determine, prove that \( S(P) - \underline{S}(P) < \epsilon \).

---

**Theorem 508** If \( f \) is R-D integrable on \([a, b]\) and \( g \) is obtained from \( f \) by altering the values of \( f \) at a finite number of points, then \( g \) is also R-D integrable. Furthermore, \( \int_{a}^{b} f = \int_{a}^{b} g \).

**Proof.** The proof is assigned as an exercise. We give an outline of the proof here. We break the proof in several cases and give an outline for each case.

- Suppose that \( g \) differs from \( f \) at one point that is let \( g(x) = \begin{cases} f(x) & \text{if } x \neq c \in [a, b] \\ y & \text{if } x = c \end{cases} \). We outline the proof in the case \( y > f(c) \).
  The other case is similar.

  1. First, we prove that \( g \) is R-D integrable by proving that \( |S(g, P) - \underline{S}(g, P)| < \epsilon \). For this, using the notation in the notes, prove that \( |S(f, P) - \underline{S}(g, P)| \leq 2 |y - m_i| d_i \) and \( |\overline{S}(f, P) - S(g, P)| \leq 2 |M_i - y| d_i \). Then, explain why one can find partitions \( Q \) and \( R \) such that \( |S(f, Q) - \overline{S}(g, Q)| < \frac{\epsilon}{3} \) and \( |\overline{S}(f, R) - S(g, R)| < \frac{\epsilon}{3} \). Conclude one can find a partition \( P \) such that \( |S(g, P) - \underline{S}(g, P)| < \epsilon \).
  2. Next, we prove that \( \int_{a}^{b} f = \int_{a}^{b} g \). For this explain why \( S(f, R) \leq \int_{a}^{b} f \leq S(g, R) \leq \int_{a}^{b} g \leq S(f, P) \leq \int_{a}^{b} g \leq \overline{S}(g, P) \) and using the fact that \( f \) and \( g \) are R-D integrable, show \( \int_{a}^{b} f = \int_{a}^{b} g \).
Suppose $g$ differs from $f$ at $k$ points. Build a sequence of functions $\{f_i\}$ where $f_0 = f$, $f_k = g$ such that each two consecutive function differ by only one point and apply the result to the sequence.

Corollary 509 A bounded function on a closed interval $[a, b]$ whose set of discontinuities is finite is R-D integrable on $[a, b]$.

**Proof.** The proof is assigned as an exercise.

8.2.4 Exercises

Do the following problems and write your solutions nicely on a piece of paper.

1. Prove proposition 489
2. Prove lemma 492.
3. Prove lemma 493
4. Prove theorem 496 (hint: First, prove the result assuming $Q$ is obtained from $P$ by adjoining one more point. Then, use that result on a sequence of partitions $Q_1, Q_2, Q_3, \ldots$ where $Q_i$ is obtained from $Q_{i-1}$ by adding one more point.)
5. Prove theorem 497 (hint: use the previous theorem and the fact that $R = P \cup Q$ is a refinement of both $P$ and $Q$).
7. Prove that $\left| \int_a^b f - \int_a^b f \right| \leq |\mathcal{S}(P) - \mathcal{L}(P)|$ for any partition $P$.
8. Prove theorem 505 following the outline in the notes.
9. Prove theorem 506 following the outline in the notes.
10. Prove theorem 507 following the outline in the notes.
11. Prove theorem 508 following the outline in the notes.
12. Prove theorem 509 following the outline in the notes.
13. Discuss the R-D integrability of the functions below in the given interval.

   (a) $f(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1 \\ 2 & \text{if } x = 1 \end{cases}$ on $[0, 1]$. In addition, find $\int_0^1 f$ if it exists.
8.2. DEFINITION OF THE RIEMANN-DARBoux INTEGRAL

(b) \[ f(x) = \begin{cases} 
0 & \text{if } x = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots \\
1 & \text{otherwise} 
\end{cases} \]
on \([0, 1]\). In addition, find \( \int_0^1 f \) if it exists.

(c) \[ f(x) = \begin{cases} 
1 & \text{if } x \in \mathbb{Q} \\
-1 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} 
\end{cases} \]
on \([a, b]\) for any \(a < b\).

(d) \[ f(x) = \begin{cases} 
x & \text{if } x \in \mathbb{Q} \\
0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} 
\end{cases} \]
on \([0, 1]\)

(e) \[ f(x) = \begin{cases} 
0 & \text{if } x \text{ is irrational} \\
\frac{1}{q} & \text{if } x = \frac{p}{q} \in \mathbb{Q} \text{ in lowest terms} 
\end{cases} \]
on \([0, 1]\)