Chapter 11

The Riemann integral

In the previous chapter we reviewed *differentiation* - one of the two pillars of single variable calculus. The other pillar is, of course, *integration*, which is the focus of the current chapter. More precisely, we will turn to the *definite integral*, the integral of a function on a fixed interval, as opposed to the *indefinite integral*, otherwise known as the antiderivative. These two are of course linked by the *Fundamental theorem of calculus*, of which more will be said later.

For us, the study of the definite integral will start with an interval I which could be open, closed, or half-open, and a function $f : I \to \mathbf{R}$, and will lead us to a number $\int_I f$; we can write this integral as $\int_I f(x)\, dx$ (of course, we could replace x by any other dummy variable), or if I has endpoints a and b, we shall also write this integral as $\int_a^b f$ or $\int_a^b f(x)\, dx$.

To actually *define* this integral $\int_I f$ is somewhat delicate (especially if one does not want to assume any axioms concerning geometric notions such as area), and not all functions f are integrable. It turns out that there are at least two ways to define this integral: the *Riemann integral*, named after Georg Riemann (1826–1866), which we will do here and which suffices for most applications, and the *Lebesgue integral*, named after Henri Lebesgue (1875–1941), which supercedes the Riemann integral and works for a much larger class of functions. The Lebesgue integral will be constructed in Chapter 19. There is also the *Riemann-Steiltjes integral* $\int_I f(x)\, d\alpha(x)$, a generalization of the Riemann integral due to Thomas Stieltjes (1856–1894), which we will discuss in Section 11.8.

Our strategy in defining the Riemann integral is as follows. We begin by first defining a notion of integration on a very simple class of functions - the *piecewise constant* functions. These functions are quite primitive, but their advantage is that integration is very easy for these functions,

as is verifying all the usual properties. Then, we handle more general functions by approximating them by piecewise constant functions.

11.1 Partitions

Before we can introduce the concept of an integral, we need to describe how one can partition a large interval into smaller intervals. In this chapter, all intervals will be bounded intervals (as opposed to the more general intervals defined in Definition 9.1.1).

Definition 11.1.1. Let X be a subset of $\mathbf{R}$. We say that X is *connected* iff the following property is true: whenever x, y are elements in X such that $x < y$, the bounded interval $[x, y]$ is a subset of X (i.e., every number between x and y is also in X).

Remark 11.1.2. Later on, in Section 13.4 we will define a more general notion of connectedness, which applies to any metric space.

Examples 11.1.3. The set $[1, 2]$ is connected, because if $x < y$ both lie in $[1, 2]$, then $1 \le x < y \le 2$, and so every element between x and y also lies in $[1, 2]$. A similar argument shows that the set $(1, 2)$ is connected. However, the set $[1, 2] \cup [3, 4]$ is not connected (why?). The real line is connected (why?). The empty set, as well as singleton sets such as $\{3\}$, are connected, but for rather trivial reasons (these sets do not contain two elements x, y for which $x < y$).

Lemma 11.1.4. *Let X be a subset of the real line. Then the following two statements are logically equivalent:*

(a) *X is bounded and connected.*

(b) *X is a bounded interval.*

Proof. See Exercise 11.1.1. □

Remark 11.1.5. Recall that intervals are allowed to be singleton points (e.g., the degenerate interval $[2, 2] = \{2\}$), or even the empty set.

Corollary 11.1.6. *If I and J are bounded intervals, then the intersection $I \cap J$ is also a bounded interval.*

Proof. See Exercise 11.1.2. □

Example 11.1.7. The intersection of the bounded intervals $[2,4]$ and $[4,6]$ is $\{4\}$, which is also a bounded interval. The intersection of $(2,4)$ and $(4,6)$ is $\emptyset$.

We now give each bounded interval a length.

Definition 11.1.8 (Length of intervals). If I is a bounded interval, we define the *length* of I, denoted $|I|$ as follows. If I is one of the intervals $[a,b]$, (a,b), $[a,b)$, or $(a,b]$ for some real numbers $a < b$, then we define $|I| := b-a$. Otherwise, if I is a point or the empty set, we define $|I| = 0$.

Example 11.1.9. For instance, the length of $[3,5]$ is 2, as is the length of $(3,5)$; meanwhile, the length of $\{5\}$ or the empty set is 0.

Definition 11.1.10 (Partitions). Let I be a bounded interval. A *partition* of I is a finite set $\mathbf{P}$ of bounded intervals contained in I, such that every x in I lies in exactly one of the bounded intervals J in $\mathbf{P}$.

Remark 11.1.11. Note that a partition is a set of intervals, while each interval is itself a set of real numbers. Thus a partition is a set consisting of other sets.

Examples 11.1.12. The set $\mathbf{P} = \{\{1\},(1,3),[3,5),\{5\},(5,8],\emptyset\}$ of bounded intervals is a partition of $[1,8]$, because all the intervals in $\mathbf{P}$ lie in $[1,8]$, and each element of $[1,8]$ lies in exactly one interval in $\mathbf{P}$. Note that one could have removed the empty set from $\mathbf{P}$ and still obtain a partition. However, the set $\{[1,4],[3,5]\}$ is not a partition of $[1,5]$ because some elements of $[1,5]$ are included in more than one interval in the set. The set $\{(1,3),(3,5)\}$ is not a partition of $(1,5)$ because some elements of $(1,5)$ are not included in any interval in the set. The set $\{(0,3),[3,5)\}$ is not a partition of $(1,5)$ because some intervals in the set are not contained in $(1,5)$.

Now we come to a basic property about length:

Theorem 11.1.13 (Length is finitely additive). *Let I be a bounded interval, n be a natural number, and let $\mathbf{P}$ be a partition of I of cardinality n. Then*

$$|I| = \sum_{J\in\mathbf{P}} |J|.$$

Proof. We prove this by induction on n. More precisely, we let $P(n)$ be the property that whenever I is a bounded interval, and whenever $\mathbf{P}$ is a partition of I with cardinality n, that $|I| = \sum_{J\in\mathbf{P}} |J|$.

The base case $P(0)$ is trivial; the only way that I can be partitioned into an empty partition is if I is itself empty (why?), at which point the claim is easy. The case $P(1)$ is also very easy; the only way that I can be partitioned into a singleton set $\{J\}$ is if $J = I$ (why?), at which point the claim is again very easy.

Now suppose inductively that $P(n)$ is true for some $n \geq 1$, and now we prove $P(n+1)$. Let I be a bounded interval, and let $\mathbf{P}$ be a partition of I of cardinality $n+1$.

If I is the empty set or a point, then all the intervals in $\mathbf{P}$ must also be either the empty set or a point (why?), and so every interval has length zero and the claim is trivial. Thus we will assume that I is an interval of the form (a,b), $(a,b]$, $[a,b)$, or $[a,b]$.

Let us first suppose that $b \in I$, i.e., I is either $(a,b]$ or $[a,b]$. Since $b \in I$, we know that one of the intervals K in $\mathbf{P}$ contains b. Since K is contained in I, it must therefore be of the form $(c,b]$, $[c,b]$, or $\{b\}$ for some real number c, with $a \leq c \leq b$ (in the latter case of $K = \{b\}$, we set $c := b$). In particular, this means that the set $I - K$ is also an interval of the form $[a,c]$, (a,c), $(a,c]$, $[a,c)$ when $c > a$, or a point or empty set when $a = c$. Either way, we easily see that

$$|I| = |K| + |I - K|.$$

On the other hand, since $\mathbf{P}$ forms a partition of I, we see that $\mathbf{P} - \{K\}$ forms a partition of $I - K$ (why?). By the induction hypothesis, we thus have

$$|I - K| = \sum_{J\in\mathbf{P}-\{K\}} |J|.$$

Combining these two identities (and using the laws of addition for finite sets, see Proposition 7.1.11) we obtain

$$|I| = \sum_{J\in\mathbf{P}} |J|$$

as desired.

Now suppose that $b \notin I$, i.e., I is either (a,b) or $[a,b)$. Then one of the intervals K also is of the form (c,b) or $[c,b)$ (see Exercise 11.1.3). In particular, this means that the set $I - K$ is also an interval of the form $[a,c]$, (a,c), $(a,c]$, $[a,c)$ when $c > a$, or a point or empty set when $a = c$. The rest of the argument then proceeds as above. □

There are two more things we need to do with partitions. One is to say when one partition is finer than another, and the other is to talk about the common refinement of two partitions.

Definition 11.1.14 (Finer and coarser partitions). Let I be a bounded interval, and let $\mathbf{P}$ and $\mathbf{P}'$ be two partitions of I. We say that $\mathbf{P}'$ is *finer* than $\mathbf{P}$ (or equivalently, that $\mathbf{P}$ is *coarser* than $\mathbf{P}'$) if for every J in $\mathbf{P}'$, there exists a K in $\mathbf{P}$ such that $J \subseteq K$.

Example 11.1.15. The partition $\{[1,2),\{2\},(2,3),[3,4]\}$ is finer than $\{[1,2],(2,4]\}$ (why?). Both partitions are finer than $\{[1,4]\}$, which is the coarsest possible partition of $[1,4]$. Note that there is no such thing as a "finest" partition of $[1,4]$. (Why? recall all partitions are assumed to be finite.) We do not compare partitions of different intervals, for instance if $\mathbf{P}$ is a partition of $[1,4]$ and $\mathbf{P}'$ is a partition of $[2,5]$ then we would not say that $\mathbf{P}$ is coarser or finer than $\mathbf{P}'$.

Definition 11.1.16 (Common refinement). Let I be a bounded interval, and let $\mathbf{P}$ and $\mathbf{P}'$ be two partitions of I. We define the *common refinement* $\mathbf{P}\#\mathbf{P}'$ of $\mathbf{P}$ and $\mathbf{P}'$ to be the set

$$\mathbf{P}\#\mathbf{P}' := \{K \cap J : K \in \mathbf{P} \text{ and } J \in \mathbf{P}'\}.$$

Example 11.1.17. Let $\mathbf{P} := \{[1,3), [3,4]\}$ and $\mathbf{P}' := \{[1,2], (2,4]\}$ be two partitions of $[1,4]$. Then $\mathbf{P}\#\mathbf{P}'$ is the set $\{[1,2], (2,3), [3,4], \emptyset\}$ (why?).

Lemma 11.1.18. *Let I be a bounded interval, and let $\mathbf{P}$ and $\mathbf{P}'$ be two partitions of I. Then $\mathbf{P}\#\mathbf{P}'$ is also a partition of I, and is both finer than $\mathbf{P}$ and finer than $\mathbf{P}'$.*

Proof. See Exercise 11.1.4. □

— Exercises —

Exercise 11.1.1. Prove Lemma 11.1.4. (Hint: in order to show that (a) implies (b) in the case when X is non-empty, consider the supremum and infimum of X.)

Exercise 11.1.2. Prove Corollary 11.1.6. (Hint: use Lemma 11.1.4, and explain why the intersection of two bounded sets is automatically bounded, and why the intersection of two connected sets is automatically connected.)

Exercise 11.1.3. Let I be a bounded interval of the form $I = (a,b)$ or $I = [a,b)$ for some real numbers $a < b$. Let $I_1, \ldots, I_n$ be a partition of I. Prove that one of the intervals I_j in this partition is of the form $I_j = (c,b)$ or $I_j = [c,b)$ for some $a \leq c \leq b$. (Hint: prove by contradiction. First show that if I_j is *not* of the form (c,b) or $[c,b)$ for any $a \leq c \leq b$, then $\sup I_j$ is *strictly* less than b.)

Exercise 11.1.4. Prove Lemma 11.1.18.

11.2 Piecewise constant functions

We can now describe the class of "simple" functions which we can integrate very easily.

Definition 11.2.1 (Constant functions). Let X be a subset of $\mathbf{R}$, and let $f : X \to \mathbf{R}$ be a function. We say that f is *constant* iff there exists a real number c such that $f(x) = c$ for all $x \in X$. If E is a subset of X, we say that f is *constant on E* if the restriction $f|_E$ of f to E is constant, in other words there exists a real number c such that $f(x) = c$ for all $x \in E$. We refer to c as the *constant value* of f on E.

E

Remark 11.2.2. If E is a non-empty set, then a function f which is constant on E can have only one constant value; it is not possible for a function to always equal 3 on E while simultaneously always equalling 4. However, if E is empty, every real number c is a constant value for f on E (why?).

Definition 11.2.3 (Piecewise constant functions I). Let I be a bounded interval, let $f : I \to \mathbf{R}$ be a function, and let $\mathbf{P}$ be a partition of I. We say that f is *piecewise constant with respect to* $\mathbf{P}$ if for every $J \in \mathbf{P}$, f is constant on J.

Example 11.2.4. The function $f : [1,6] \to \mathbf{R}$ defined by

$$f(x) = \begin{cases} 7 & \text{if } 1 \leq x < 3 \\ 4 & \text{if } x = 3 \\ 5 & \text{if } 3 < x < 6 \\ 2 & \text{if } x = 6 \end{cases}$$

is piecewise constant with respect to the partition $\{[1,3), \{3\}, (3,6), \{6\}\}$ of $[1,6]$. Note that it is also piecewise constant with respect to some other partitions as well; for instance, it is piecewise constant with respect to the partition $\{[1,2), \{2\}, (2,3), \{3\}, (3,5), [5,6), \{6\}, \emptyset\}$.

Definition 11.2.5 (Piecewise constant functions II). Let I be a bounded interval, and let $f : I \to \mathbf{R}$ be a function. We say that f

is *piecewise constant on I* if there exists a partition $\mathbf{P}$ of I such that f is piecewise constant with respect to $\mathbf{P}$.

Example 11.2.6. The function used in the previous example is piecewise constant on $[1,6]$. Also, every constant function on a bounded interval I is automatically piecewise constant (why?).

Lemma 11.2.7. *Let I be a bounded interval, let $\mathbf{P}$ be a partition of I, and let $f : I \to \mathbf{R}$ be a function which is piecewise constant with respect to $\mathbf{P}$. Let $\mathbf{P}'$ be a partition of I which is finer than $\mathbf{P}$. Then f is also piecewise constant with respect to $\mathbf{P}'$.*

Proof. See Exercise 11.2.1. □

The space of piecewise constant functions is closed under algebraic operations:

Lemma 11.2.8. *Let I be a bounded interval, and let $f : I \to \mathbf{R}$ and $g : I \to \mathbf{R}$ be piecewise constant functions on I. Then the functions $f+g$, $f-g$, $\max(f,g)$ and fg are also piecewise constant functions on I. Here of course $\max(f,g) : I \to \mathbf{R}$ is the function $\max(f,g)(x) := \max(f(x), g(x))$. If g does not vanish anywhere on I (i.e., $g(x) \neq 0$ for all $x \in I$) then f/g is also a piecewise constant function on I.*

Proof. See Exercise 11.2.2. □

We are now ready to integrate piecewise constant functions. We begin with a temporary definition of an integral with respect to a partition.

Definition 11.2.9 (Piecewise constant integral I). Let I be a bounded interval, let $\mathbf{P}$ be a partition of I. Let $f : I \to \mathbf{R}$ be a function which is piecewise constant with respect to $\mathbf{P}$. Then we define the *piecewise constant integral* $p.c. \int_{[\mathbf{P}]} f$ of f with respect to the partition $\mathbf{P}$ by the formula

$$p.c. \int_{[\mathbf{P}]} f := \sum_{J \in \mathbf{P}} c_J |J|,$$

where for each J in $\mathbf{P}$, we let c_J be the constant value of f on J.

Remark 11.2.10. This definition seems like it could be ill-defined, because if J is empty then every number c_J can be the constant value of f on J, but fortunately in such cases $|J|$ is zero and so the choice of c_J is irrelevant. The notation $p.c. \int_{[\mathbf{P}]} f$ is rather artificial, but we shall only need it temporarily, en route to a more useful definition. Note that since $\mathbf{P}$ is finite, the sum $\sum_{J \in \mathbf{P}} c_J |J|$ is always well-defined (it is never divergent or infinite).

Remark 11.2.11. The piecewise constant integral corresponds intuitively to one's notion of area, given that the area of a rectangle ought to be the product of the lengths of the sides. (Of course, if f is negative somewhere, then the "area" $c_J|J|$ would also be negative.)

Example 11.2.12. Let $f : [1,4] \to \mathbf{R}$ be the function

$$f(x) = \begin{cases} 2 & \text{if } 1 \le x < 3 \\ 4 & \text{if } x = 3 \\ 6 & \text{if } 3 < x \le 4 \end{cases}$$

and let $\mathbf{P} := \{[1,3), \{3\}, (3,4]\}$. Then

$$\begin{aligned} p.c. \int_{[\mathbf{P}]} f &= c_{[1,3)}|[1,3)| + c_{\{3\}}|\{3\}| + c_{(3,4]}|(3,4]| \\ &= 2 \times 2 + 4 \times 0 + 6 \times 1 \\ &= 10. \end{aligned}$$

Alternatively, if we let $\mathbf{P}' := \{[1,2), [2,3), \{3\}, (3,4], \emptyset\}$ then

$$\begin{aligned} p.c. \int_{[\mathbf{P}']} f &= c_{[1,2)}|[1,2)| + c_{[2,3)}|[2,3)| + c_{\{3\}}|\{3\}| \\ &\quad + c_{(3,4]}|(3,4]| + c_{\emptyset}|\emptyset| \\ &= 2 \times 1 + 2 \times 1 + 4 \times 0 + 6 \times 1 + c_{\emptyset} \times 0 \\ &= 10. \end{aligned}$$

This example suggests that this integral does not really depend on what partition you pick, so long as your function is piecewise constant with respect to that partition. That is indeed true:

Proposition 11.2.13 (Piecewise constant integral is independent of partition). *Let I be a bounded interval, and let $f : I \to \mathbf{R}$ be a function. Suppose that $\mathbf{P}$ and $\mathbf{P}'$ are partitions of I such that f is piecewise constant both with respect to $\mathbf{P}$ and with respect to $\mathbf{P}'$. Then $p.c. \int_{[\mathbf{P}]} f = p.c. \int_{[\mathbf{P}']} f$.*

Proof. See Exercise 11.2.3. □

Because of this proposition, we can now make the following definition:

Definition 11.2.14 (Piecewise constant integral II). Let I be a bounded interval, and let $f : I \to \mathbf{R}$ be a piecewise constant function on I. We define the *piecewise constant integral* $p.c. \int_I f$ by the formula

$$p.c. \int_I f := p.c. \int_{[\mathbf{P}]} f.$$

where $\mathbf{P}$ is any partition of I with respect to which f is piecewise constant. (Note that Proposition 11.2.13 tells us that the precise choice of this partition is irrelevant.)

Example 11.2.15. If f is the function given in Example 11.2.12, then $p.c. \int_{[1,4]} f = 10$.

We now give some basic properties of the piecewise constant integral. These laws will eventually be superceded by the corresponding laws for the Riemann integral (Theorem 11.4.1).

Theorem 11.2.16 (Laws of integration). *Let I be a bounded interval, and let $f : I \to \mathbf{R}$ and $g : I \to \mathbf{R}$ be piecewise constant functions on I.*

(a) *We have* $p.c. \int_I (f+g) = p.c. \int_I f + p.c. \int_I g$.

(b) *For any real number c, we have* $p.c. \int_I (cf) = c(p.c. \int_I f)$.

(c) *We have* $p.c. \int_I (f-g) = p.c. \int_I f - p.c. \int_I g$.

(d) *If $f(x) \geq 0$ for all $x \in I$, then* $p.c. \int_I f \geq 0$.

(e) *If $f(x) \geq g(x)$ for all $x \in I$, then* $p.c. \int_I f \geq p.c. \int_I g$.

(f) *If f is the constant function $f(x) = c$ for all x in I, then* $p.c. \int_I f = c|I|$.

(g) *Let J be a bounded interval containing I (i.e., $I \subseteq J$), and let $F : J \to \mathbf{R}$ be the function*

$$F(x) := \begin{cases} f(x) & \text{if } x \in I \\ 0 & \text{if } x \notin I \end{cases}$$

Then F is piecewise constant on J, and $p.c. \int_J F = p.c. \int_I f$.

(h) *Suppose that $\{J, K\}$ is a partition of I into two intervals J and K. Then the functions $f|_J : J \to \mathbf{R}$ and $f|_K : K \to \mathbf{R}$ are piecewise constant on J and K respectively, and we have*

$$p.c. \int_I f = p.c. \int_J f|_J + p.c. \int_K f|_K.$$

Proof. See Exercise 11.2.4. □

This concludes our integration of piecewise constant functions. We now turn to the question of how to integrate bounded functions.

— Exercises —

Exercise 11.2.1. Prove Lemma 11.2.7.

Exercise 11.2.2. Prove Lemma 11.2.8. (Hint: use Lemmas 11.1.18 and 11.2.7 to make f and g piecewise constant with respect to the *same* partition of I.)

Exercise 11.2.3. Prove Proposition 11.2.13. (Hint: first use Theorem 11.1.13 to show that both integrals are equal to $p.c. \int_{[\mathbf{P} \# \mathbf{P}']} f$.)

Exercise 11.2.4. Prove Theorem 11.2.16. (Hint: you can use earlier parts of the theorem to prove some of the later parts of the theorem. See also the hint to Exercise 11.2.2.)

11.3 Upper and lower Riemann integrals

Now let $f : I \to \mathbf{R}$ be a bounded function defined on a bounded interval I. We want to define the Riemann integral $\int_I f$. To do this we first need to define the notion of upper and lower Riemann integrals $\overline{\int}_I f$ and $\underline{\int}_I f$. These notions are related to the Riemann integral in much the same way that the lim sup and lim inf of a sequence are related to the limit of that sequence.

Definition 11.3.1 (Majorization of functions). Let $f : I \to \mathbf{R}$ and $g : I \to \mathbf{R}$. We say that g *majorizes* f *on* I if we have $g(x) \geq f(x)$ for all $x \in I$, and that g *minorizes* f *on* I if $g(x) \leq f(x)$ for all $x \in I$.

The idea of the Riemann integral is to try to integrate a function by first majorizing or minorizing that function by a piecewise constant function (which we already know how to integrate).

Definition 11.3.2 (Upper and lower Riemann integrals). Let $f : I \to \mathbf{R}$ be a bounded function defined on a bounded interval I. We define the *upper Riemann integral* $\overline{\int}_I f$ by the formula

$$\overline{\int}_I f := \inf\{p.c. \int_I g : g \text{ is a p.c. function on } I \text{ which majorizes } f\}$$

and the *lower Riemann integral* $\underline{\int}_I f$ by the formula

$$\underline{\int}_I f := \sup\{p.c. \int_I g : g \text{ is a p.c. function on } I \text{ which minorizes } f\}.$$

We give a crude but useful bound on the lower and upper integral:

Lemma 11.3.3. *Let $f : I \to \mathbf{R}$ be a function on a bounded interval I which is bounded by some real number M, i.e., $-M \leq f(x) \leq M$ for all $x \in I$. Then we have*

$$-M|I| \leq \underline{\int}_I f \leq \overline{\int}_I f \leq M|I|.$$

In particular, both the lower and upper Riemann integrals are real numbers (i.e., they are not infinite).

Proof. The function $g : I \to \mathbf{R}$ defined by $g(x) = M$ is constant, hence piecewise constant, and majorizes f; thus $\overline{\int}_I f \leq p.c. \int_I g = M|I|$ by definition of the upper Riemann integral. A similar argument gives $-M|I| \leq \underline{\int}_I f$. Finally, we have to show that $\underline{\int}_I f \leq \overline{\int}_I f$. Let g be any piecewise constant function majorizing f, and let h be any piecewise constant function minorizing f. Then g majorizes h, and hence $p.c. \int_I h \leq p.c. \int_I g$. Taking suprema in h, we obtain that $\underline{\int}_I f \leq p.c. \int_I g$. Taking infima in g, we thus obtain $\underline{\int}_I f \leq \overline{\int}_I g$, as desired. □

We now know that the upper Riemann integral is always at least as large as the lower Riemann integral. If the two integrals match, then we can define the Riemann integral:

Definition 11.3.4 (Riemann integral). Let $f : I \to \mathbf{R}$ be a bounded function on a bounded interval I. If $\underline{\int}_I f = \overline{\int}_I f$, then we say that f is *Riemann integrable on I* and define

$$\int_I f := \underline{\int}_I f = \overline{\int}_I f.$$

If the upper and lower Riemann integrals are unequal, we say that f is not Riemann integrable.

Remark 11.3.5. Compare this definition to the relationship between the lim sup, lim inf, and limit of a sequence a_n that was established in Proposition 6.4.12(f); the lim sup is always greater than or equal to the lim inf, but they are only equal when the sequence converges, and in this case they are both equal to the limit of the sequence. The definition given above may differ from the definition you may have encountered in your calculus courses, based on Riemann sums. However, the two definitions turn out to be equivalent; this is the purpose of the next section.

Remark 11.3.6. Note that we do not consider unbounded functions to be Riemann integrable; an integral involving such functions is known as an *improper integral*. It is possible to still evaluate such integrals using more sophisticated integration methods (such as the Lebesgue integral); we shall do this in Chapter 19.

The Riemann integral is consistent with (and supercedes) the piecewise constant integral:

Lemma 11.3.7. *Let $f : I \to \mathbf{R}$ be a piecewise constant function on a bounded interval I. Then f is Riemann integrable, and $\int_I f = p.c. \int_I f$.*

Proof. See Exercise 11.3.3. □

Remark 11.3.8. Because of this lemma, we will not refer to the piecewise constant integral $p.c. \int_I$ again, and just use the Riemann integral $\int_I$ throughout (until this integral is itself superceded by the Lebesgue integral in Chapter 19). We observe one special case of Lemma 11.3.7: if I is a point or the empty set, then $\int_I f = 0$ for all functions $f : I \to \mathbf{R}$. (Note that all such functions are automatically constant.)

We have just shown that every piecewise constant function is Riemann integrable. However, the Riemann integral is more general, and can integrate a wider class of functions; we shall see this shortly. For now, we connect the Riemann integral we have just defined to the concept of a *Riemann sum*, which you may have seen in other treatments of the Riemann integral.

Definition 11.3.9 (Riemann sums). Let $f : I \to \mathbf{R}$ be a bounded function on a bounded interval I, and let $\mathbf{P}$ be a partition of I. We define the *upper Riemann sum* $U(f, \mathbf{P})$ and the *lower Riemann sum* $L(f, \mathbf{P})$ by

$$U(f, \mathbf{P}) := \sum_{J \in \mathbf{P}: J \neq \emptyset} (\sup_{x \in J} f(x))|J|$$

and

$$L(f, \mathbf{P}) := \sum_{J \in \mathbf{P}: J \neq \emptyset} (\inf_{x \in J} f(x))|J|.$$

Remark 11.3.10. The restriction $J \neq \emptyset$ is required because the quantities $\inf_{x \in J} f(x)$ and $\sup_{x \in J} f(x)$ are infinite (or negative infinite) if J is empty.

We now connect these Riemann sums to the upper and lower Riemann integral.

Lemma 11.3.11. *Let $f : I \to \mathbf{R}$ be a bounded function on a bounded interval I, and let g be a function which majorizes f and which is piecewise constant with respect to some partition $\mathbf{P}$ of I. Then*

$$p.c. \int_I g \geq U(f, \mathbf{P}).$$

Similarly, if h is a function which minorizes f and is piecewise constant with respect to $\mathbf{P}$, then

$$p.c. \int_I h \leq L(f, \mathbf{P}).$$

Proof. See Exercise 11.3.4. □

Proposition 11.3.12. *Let $f : I \to \mathbf{R}$ be a bounded function on a bounded interval I. Then*

$$\overline{\int}_I f = \inf\{U(f, \mathbf{P}) : \mathbf{P} \text{ is a partition of } I\}$$

and

$$\underline{\int}_I f = \sup\{L(f, \mathbf{P}) : \mathbf{P} \text{ is a partition of } I\}$$

Proof. See Exercise 11.3.5. □

— Exercises —

Exercise 11.3.1. Let $f : I \to \mathbf{R}$, $g : I \to \mathbf{R}$, and $h : I \to \mathbf{R}$ be functions. Show that if f majorizes g and g majorizes h, then f majorizes h. Show that if f and g majorize each other, then they must be equal.

Exercise 11.3.2. Let $f : I \to \mathbf{R}$, $g : I \to \mathbf{R}$, and $h : I \to \mathbf{R}$ be functions. If f majorizes g, is it true that $f + h$ majorizes $g + h$? Is it true that $f \cdot h$ majorizes $g \cdot h$? If c is a real number, is it true that cf majorizes cg?

Exercise 11.3.3. Prove Lemma 11.3.7.

Exercise 11.3.4. Prove Lemma 11.3.11.

Exercise 11.3.5. Prove Proposition 11.3.12. (Hint: you will need Lemma 11.3.11, even though this Lemma will only do half of the job.)

11.4 Basic properties of the Riemann integral

Just as we did with limits, series, and derivatives, we now give the basic laws for manipulating the Riemann integral. These laws will eventually be superceded by the corresponding laws for the Lebesgue integral (Proposition 19.3.3).

Theorem 11.4.1 (Laws of Riemann integration). *Let I be a bounded interval, and let $f : I \to \mathbf{R}$ and $g : I \to \mathbf{R}$ be Riemann integrable functions on I.*

(a) *The function $f + g$ is Riemann integrable, and we have $\int_I (f+g) = \int_I f + \int_I g$.*

(b) *For any real number c, the function cf is Riemann integrable, and we have $\int_I (cf) = c(\int_I f)$.*

(c) *The function $f - g$ is Riemann integrable, and we have $\int_I (f-g) = \int_I f - \int_I g$.*

(d) *If $f(x) \geq 0$ for all $x \in I$, then $\int_I f \geq 0$.*

(e) *If $f(x) \geq g(x)$ for all $x \in I$, then $\int_I f \geq \int_I g$.*

(f) *If f is the constant function $f(x) = c$ for all x in I, then $\int_I f = c|I|$.*

(g) *Let J be a bounded interval containing I (i.e., $I \subseteq J$), and let $F : J \to \mathbf{R}$ be the function*

$$F(x) := \begin{cases} f(x) & \text{if } x \in I \\ 0 & \text{if } x \notin I \end{cases}$$

Then F is Riemann integrable on J, and $\int_J F = \int_I f$.

(h) *Suppose that $\{J, K\}$ is a partition of I into two intervals J and K. Then the functions $f|_J : J \to \mathbf{R}$ and $f|_K : K \to \mathbf{R}$ are Riemann integrable on J and K respectively, and we have*

$$\int_I f = \int_J f|_J + \int_K f|_K.$$

Proof. See Exercise 11.4.1. □

Remark 11.4.2. We often abbreviate $\int_J f|_J$ as $\int_J f$, even though f is really defined on a larger domain than just J.

Theorem 11.4.1 asserts that the sum or difference of any two Riemann integrable functions is Riemann integrable, as is any scalar multiple cf of a Riemann integrable function f. We now give some further ways to create Riemann integrable functions.

Theorem 11.4.3 (Max and min preserve integrability). *Let I be a bounded interval, and let $f : I \to \mathbf{R}$ and $g : I \to \mathbf{R}$ be a Riemann integrable function. Then the functions $\max(f, g) : I \to \mathbf{R}$ and $\min(f, g) : I \to \mathbf{R}$ defined by $\max(f, g)(x) := \max(f(x), g(x))$ and $\min(f, g)(x) := \min(f(x), g(x))$ are also Riemann integrable.*

Proof. We shall just prove the claim for $\max(f, g)$, the case of $\min(f, g)$ being similar. First note that since f and g are bounded, then $\max(f, g)$ is also bounded.

Let $\varepsilon > 0$. Since $\int_I f = \underline{\int}_I f$, there exists a piecewise constant function $\underline{f} : I \to \mathbf{R}$ which minorizes f on I such that

$$\int_I \underline{f} \geq \int_I f - \varepsilon.$$

Similarly we can find a piecewise constant $\underline{g} : I \to \mathbf{R}$ which minorizes g on I such that

$$\int_I \underline{g} \geq \int_I g - \varepsilon,$$

and we can find piecewise functions $\overline{f}, \overline{g}$ which majorize f, g respectively on I such that

$$\int_I \overline{f} \leq \int_I f + \varepsilon$$

and

$$\int_I \overline{g} \leq \int_I g + \varepsilon.$$

In particular, if $h : I \to \mathbf{R}$ denotes the function

$$h := (\overline{f} - \underline{f}) + (\overline{g} - \underline{g})$$

we have

$$\int_I h \leq 4\varepsilon.$$

On the other hand, $\max(\underline{f}, \underline{g})$ is a piecewise constant function on I (why?) which minorizes $\max(f, g)$ (why?), while $\max(\overline{f}, \overline{g})$ is similarly a piecewise constant function on I which majorizes $\max(f, g)$. Thus

$$\int_I \max(\underline{f}, \underline{g}) \leq \underline{\int}_I \max(f, g) \leq \overline{\int}_I \max(f, g) \leq \int_I \max(\overline{f}, \overline{g}),$$

and so

$$0 \leq \overline{\int}_I \max(f, g) - \underline{\int}_I \max(f, g) \leq \int_I \max(\overline{f}, \overline{g}) - \max(\underline{f}, \underline{g}).$$

But we have

$$\overline{f}(x) = \underline{f}(x) + (\overline{f} - \underline{f})(x) \leq \underline{f}(x) + h(x)$$

and similarly

$$\overline{g}(x) = \underline{g}(x) + (\overline{g} - \underline{g})(x) \leq \underline{g}(x) + h(x)$$

and thus

$$\max(\overline{f}(x), \overline{g}(x)) \leq \max(\underline{f}(x), \underline{g}(x)) + h(x).$$

Inserting this into the previous inequality, we obtain

$$0 \leq \overline{\int}_I \max(f, g) - \underline{\int}_I \max(f, g) \leq \int_I h \leq 4\varepsilon.$$

To summarize, we have shown that

$$0 \le \overline{\int}_I \max(f,g) - \underline{\int}_I \max(f,g) \le 4\varepsilon$$

for every ε. Since $\overline{\int}_I \max(f,g) - \underline{\int}_I \max(f,g)$ does not depend on ε, we thus see that

$$\overline{\int}_I \max(f,g) - \underline{\int}_I \max(f,g) = 0$$

and hence that $\max(f,g)$ is Riemann integrable. □

Corollary 11.4.4 (Absolute values preserve Riemann integrability). *Let I be a bounded interval. If $f : I \to \mathbf{R}$ is a Riemann integrable function, then the positive part $f_+ := \max(f,0)$ and the negative part $f_- := \min(f,0)$ are also Riemann integrable on I. Also, the absolute value $|f| = f_+ - f_-$ is also Riemann integrable on I.*

Theorem 11.4.5 (Products preserve Riemann integrability). *Let I be a bounded interval. If $f : I \to \mathbf{R}$ and $g : I \to \mathbf{R}$ are Riemann integrable, then $fg : I \to \mathbf{R}$ is also Riemann integrable.*

Proof. This one is a little trickier. We split $f = f_+ + f_-$ and $g = g_+ + g_-$ into positive and negative parts; by Corollary 11.4.4, the functions f_+, f_-, g_+, g_- are Riemann integrable. Since

$$fg = f_+g_+ + f_+g_- + f_-g_+ + f_-g_-$$

then it suffices to show that the functions f_+g_+, f_+g_-, f_-g_+, f_-g_- are individually Riemann integrable. We will just show this for f_+g_+; the other three are similar.

Since f_+ and g_+ are bounded and positive, there are $M_1, M_2 > 0$ such that

$$0 \le f_+(x) \le M_1 \text{ and } 0 \le g_+(x) \le M_2$$

for all $x \in I$. Now let $\varepsilon > 0$ be arbitrary. Then, as in the proof of Theorem 11.4.3, we can find a piecewise constant function $\underline{f_+}$ minorizing f_+ on I, and a piecewise constant function $\overline{f_+}$ majorizing f_+ on I, such that

$$\int_I \overline{f_+} \le \int_I f_+ + \varepsilon$$

and

$$\int_I \underline{f_+} \ge \int_I f_+ - \varepsilon.$$

Note that $\underline{f_+}$ may be negative at places, but we can fix this by replacing $\underline{f_+}$ by $\max(\underline{f_+},0)$, since this still minorizes f_+ (why?) and still has integral greater than or equal to $\int_I f_+ - \varepsilon$ (why?). So without loss of generality we may assume that $\underline{f_+}(x) \ge 0$ for all $x \in I$. Similarly we may assume that $\overline{f_+}(x) \le M_1$ for all $x \in I$; thus

$$0 \le \underline{f_+}(x) \le f_+(x) \le \overline{f_+}(x) \le M_1$$

for all $x \in I$.

Similar reasoning allows us to find piecewise constant $\underline{g_+}$ minorizing g_+, and $\overline{g_+}$ majorizing g_+, such that

$$\int_I \overline{g_+} \le \int_I g_+ + \varepsilon$$

and

$$\int_I \underline{g_+} \ge \int_I g_+ - \varepsilon,$$

and

$$0 \le \underline{g_+}(x) \le g_+(x) \le \overline{g_+}(x) \le M_2$$

for all $x \in I$.

Notice that $\underline{f_+}\underline{g_+}$ is piecewise constant and minorizes f_+g_+, while $\overline{f_+}\overline{g_+}$ is piecewise constant and majorizes f_+g_+. Thus

$$0 \le \overline{\int}_I f_+g_+ - \underline{\int}_I f_+g_+ \le \int_I \overline{f_+}\overline{g_+} - \underline{f_+}\underline{g_+}.$$

However, we have

$$\overline{f_+}(x)\overline{g_+}(x) - \underline{f_+}(x)\underline{g_+}(x) = \overline{f_+}(x)(\overline{g_+} - \underline{g_+})(x) + \underline{g_+}(x)(\overline{f_+} - \underline{f_+}(x))$$
$$\le M_1(\overline{g_+} - \underline{g_+})(x) + M_2(\overline{f_+} - \underline{f_+}(x))$$

for all $x \in I$, and thus

$$0 \le \overline{\int}_I f_+g_+ - \underline{\int}_I f_+g_+ \le M_1 \int_I (\overline{g_+} - \underline{g_+}) + M_2 \int_I (\overline{f_+} - \underline{f_+})$$
$$\le M_1(2\varepsilon) + M_2(2\varepsilon).$$

Again, since ε was arbitrary, we can conclude that f_+g_+ is Riemann integrable, as before. Similar argument show that f_+g_-, f_-g_+, f_-g_- are Riemann integrable; combining them we obtain that fg is Riemann integrable. □

— Exercises —

Exercise 11.4.1. Prove Theorem 11.4.1. (Hint: you may find Theorem 11.2.16 to be useful. For part (b): First do the case $c > 0$. Then do the case $c = -1$ and $c = 0$ separately. Using these cases, deduce the case of $c < 0$. You can use earlier parts of the theorem to prove later ones.)

Exercise 11.4.2. Let $a < b$ be real numbers, and let $f : [a,b] \to \mathbf{R}$ be a continuous, non-negative function (so $f(x) \geq 0$ for all $x \in [a,b]$). Suppose that $\int_{[a,b]} f = 0$. Show that $f(x) = 0$ for all $x \in [a,b]$. (Hint: argue by contradiction.)

Exercise 11.4.3. Let I be a bounded interval, let $f : I \to \mathbf{R}$ be a Riemann integrable function, and let $\mathbf{P}$ be a partition of I. Show that

$$\int_I f = \sum_{J \in \mathbf{P}} \int_J f.$$

Exercise 11.4.4. Without repeating all the computations in the above proofs, give a short explanation as to why the remaining cases of Theorem 11.4.3 and Theorem 11.4.5 follow automatically from the cases presented in the text. (Hint: from Theorem 11.4.1 we know that if f is Riemann integrable, then so is $-f$.)

11.5 Riemann integrability of continuous functions

We have already said a lot about Riemann integrable functions so far, but we have not yet actually produced any such functions other than the piecewise constant ones. Now we rectify this by showing that a large class of useful functions are Riemann integrable. We begin with the uniformly continuous functions.

Theorem 11.5.1. *Let I be a bounded interval, and let f be a function which is uniformly continuous on I. Then f is Riemann integrable.*

Proof. From Proposition 9.9.15 we see that f is bounded. Now we have to show that $\underline{\int}_I f = \overline{\int}_I f$.

If I is a point or the empty set then the theorem is trivial, so let us assume that I is one of the four intervals $[a,b]$, (a,b), $(a,b]$, or $[a,b)$ for some real numbers $a < b$.

Let $\varepsilon > 0$ be arbitrary. By uniform continuity, there exists a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in I$ are such that $|x - y| < \delta$. By the Archimedean principle, there exists an integer $N > 0$ such that $(b-a)/N < \delta$.

Note that we can partition I into N intervals $J_1, \ldots, J_N$, each of length $(b-a)/N$. (How? One has to treat each of the cases $[a,b]$, (a,b), $(a,b]$, $[a,b)$ slightly differently.) By Proposition 11.3.12, we thus have

$$\overline{\int}_I f \leq \sum_{k=1}^N (\sup_{x \in J_k} f(x))|J_k|$$

and

$$\underline{\int}_I f \geq \sum_{k=1}^N (\inf_{x \in J_k} f(x))|J_k|$$

so in particular

$$\overline{\int}_I f - \underline{\int}_I f \leq \sum_{k=1}^N (\sup_{x \in J_k} f(x) - \inf_{x \in J_k} f(x))|J_k|.$$

However, we have $|f(x) - f(y)| < \varepsilon$ for all $x, y \in J_k$, since $|J_k| = (b-a)/N < \delta$. In particular we have

$$f(x) < f(y) + \varepsilon \text{ for all } x, y \in J_k.$$

Taking suprema in x, we obtain

$$\sup_{x \in J_k} f(x) \leq f(y) + \varepsilon \text{ for all } y \in J_k,$$

and then taking infima in y we obtain

$$\sup_{x \in J_k} f(x) \leq \inf_{y \in J_k} f(y) + \varepsilon.$$

Inserting this bound into our previous inequality, we obtain

$$\overline{\int}_I f - \underline{\int}_I f \leq \sum_{k=1}^N \varepsilon |J_k|,$$

but by Theorem 11.1.13 we thus have

$$\overline{\int}_I f - \underline{\int}_I f \leq \varepsilon(b-a).$$

But $\varepsilon > 0$ was arbitrary, while $(b-a)$ is fixed. Thus $\overline{\int}_I f - \underline{\int}_I f$ cannot be positive. By Lemma 11.3.3 and the definition of Riemann integrability we thus have that f is Riemann integrable. □

Combining Theorem 11.5.1 with Theorem 9.9.16, we thus obtain

Corollary 11.5.2. *Let $[a,b]$ be a closed interval, and let $f : [a,b] \to \mathbf{R}$ be continuous. Then f is Riemann integrable.*

Note that this Corollary is not true if $[a,b]$ is replaced by any other sort of interval, since it is not even guaranteed then that continuous functions are bounded. For instance, the function $f : (0,1) \to \mathbf{R}$ defined by $f(x) := 1/x$ is continuous but not Riemann integrable. However, if we assume that a function is both continuous *and* bounded, we can recover Riemann integrability:

Proposition 11.5.3. *Let I be a bounded interval, and let $f : I \to \mathbf{R}$ be both continuous and bounded. Then f is Riemann integrable on I.*

Proof. If I is a point or an empty set then the claim is trivial; if I is a closed interval the claim follows from Corollary 11.5.2. So let us assume that I is of the form $(a,b]$, (a,b), or $[a,b)$ for some $a < b$.

We have a bound M for f, so that $-M \leq f(x) \leq M$ for all $x \in I$. Now let $0 < \varepsilon < (b-a)/2$ be a small number. The function f when restricted to the interval $[a+\varepsilon, b-\varepsilon]$ is continuous, and hence Riemann integrable by Corollary 11.5.2. In particular, we can find a piecewise constant function $h : [a+\varepsilon, b-\varepsilon] \to \mathbf{R}$ which majorizes f on $[a+\varepsilon, b-\varepsilon]$ such that

$$\int_{[a+\varepsilon,b-\varepsilon]} h \leq \int_{[a+\varepsilon,b-\varepsilon]} f + \varepsilon.$$

Define $\tilde{h} : I \to \mathbf{R}$ by

$$\tilde{h}(x) := \begin{cases} h(x) & \text{if } x \in [a+\varepsilon, b-\varepsilon] \\ M & \text{if } x \in I \backslash [a+\varepsilon, b-\varepsilon] \end{cases}$$

Clearly $\tilde{h}$ is piecewise constant on I and majorizes f; by Theorem 11.2.16 we have

$$\int_I \tilde{h} = \varepsilon M + \int_{[a+\varepsilon,b-\varepsilon]} h + \varepsilon M \leq \int_{[a+\varepsilon,b-\varepsilon]} f + (2M+1)\varepsilon.$$

In particular we have

$$\overline{\int}_I f \leq \int_{[a+\varepsilon,b-\varepsilon]} f + (2M+1)\varepsilon.$$

A similar argument gives

$$\underline{\int}_I f \geq \int_{[a+\varepsilon,b-\varepsilon]} f - (2M+1)\varepsilon$$

and hence

$$\overline{\int}_I f - \underline{\int}_I f \leq (4M+2)\varepsilon.$$

But ε is arbitrary, and so we can argue as in the proof of Theorem 11.5.1 to conclude Riemann integrability. □

This gives a large class of Riemann integrable functions already; the bounded continuous functions. But we can expand this class a little more, to include the bounded *piecewise* continuous functions.

Definition 11.5.4. Let I be a bounded interval, and let $f : I \to \mathbf{R}$. We say that f is *piecewise continuous on I* iff there exists a partition $\mathbf{P}$ of I such that $f|_J$ is continuous on J for all $J \in \mathbf{P}$.

Example 11.5.5. The function $f : [1,3] \to \mathbf{R}$ defined by

$$F(x) := \begin{cases} x^2 & \text{if } 1 \leq x < 2 \\ 7 & \text{if } x = 2 \\ x^3 & \text{if } 2 < x \leq 3 \end{cases}$$

is not continuous on $[1,3]$, but it is piecewise continuous on $[1,3]$ (since it is continuous when restricted to $[1,2)$ or $\{2\}$ or $(2,3]$, and those three intervals partition $[1,3]$).

Proposition 11.5.6. *Let I be a bounded interval, and let $f : I \to \mathbf{R}$ be both piecewise continuous and bounded. Then f is Riemann integrable.*

Proof. See Exercise 11.5.1. □

— Exercises —

Exercise 11.5.1. Prove Proposition 11.5.6. (Hint: use Theorem 11.4.1(a) and (h).)

11.6 Riemann integrability of monotone functions

In addition to piecewise continuous functions, another wide class of functions is Riemann integrable, namely the monotone functions. We give two instances of this:

Proposition 11.6.1. *Let $[a,b]$ be a closed and bounded interval and let $f : [a,b] \to \mathbf{R}$ be a monotone function. Then f is Riemann integrable on $[a,b]$.*

Remark 11.6.2. From Exercise 9.8.5 we know that there exist monotone functions which are not piecewise continuous, so this proposition is not subsumed by Proposition 11.5.6.

Proof. Without loss of generality we may take f to be monotone increasing (instead of monotone decreasing). From Exercise 9.8.1 we know that f is bounded. Now let $N > 0$ be an integer, and partition $[a,b]$ into N half-open intervals $\{[a+\frac{b-a}{N}j, a+\frac{b-a}{N}(j+1)) : 0 \le j \le N-1\}$ of length $(b-a)/N$, together with the point $\{b\}$. Then by Proposition 11.3.12 we have

$$\overline{\int}_I f \le \sum_{j=0}^{N-1} \left(\sup_{x \in [a+\frac{b-a}{N}j, a+\frac{b-a}{N}(j+1))} f(x) \right) \frac{b-a}{N},$$

(the point $\{b\}$ clearly giving only a zero contribution). Since f is monotone increasing, we thus have

$$\overline{\int}_I f \le \sum_{j=0}^{N-1} f\left(a + \frac{b-a}{N}(j+1)\right) \frac{b-a}{N}.$$

Similarly we have

$$\underline{\int}_I f \ge \sum_{j=0}^{N-1} f\left(a + \frac{b-a}{N}j\right) \frac{b-a}{N}.$$

Thus we have

$$\overline{\int}_I f - \underline{\int}_I f \le \sum_{j=0}^{N-1} \left(f\left(a + \frac{b-a}{N}(j+1)\right) - f\left(a + \frac{b-a}{N}j\right)\right) \frac{b-a}{N}.$$

Using telescoping series (Lemma 7.2.15) we thus have

$$\overline{\int}_I f - \underline{\int}_I f \le \left(f\left(a + \frac{b-a}{N}(N)\right) - f\left(a + \frac{b-a}{N}0\right)\right) \frac{b-a}{N}$$
$$= (f(b) - f(a))\frac{b-a}{N}.$$

But N was arbitrary, so we can conclude as in the proof of Theorem 11.5.1 that f is Riemann integrable. □

Corollary 11.6.3. *Let I be a bounded interval, and let $f : I \to \mathbf{R}$ be both monotone and bounded. Then f is Riemann integrable on I.*

Proof. See Exercise 11.6.1. □

We now give the famous integral test for determining convergence of monotone decreasing series.

Proposition 11.6.4 (Integral test). *Let $f : [0,\infty) \to \mathbf{R}$ be a monotone decreasing function which is non-negative (i.e., $f(x) \ge 0$ for all $x \ge 0$). Then the sum $\sum_{n=0}^\infty f(n)$ is convergent if and only if $\sup_{N>0} \int_{[0,N]} f$ is finite.*

Proof. See Exercise 11.6.3. □

Corollary 11.6.5. *Let p be a real number. Then $\sum_{n=1}^\infty \frac{1}{n^p}$ converges absolutely when $p > 1$ and diverges when $p \le 1$.*

Proof. See Exercise 11.6.5. □

— Exercises —

Exercise 11.6.1. Use Proposition 11.6.1 to prove Corollary 11.6.3. (Hint: adapt the proof of Proposition 11.5.3.)

Exercise 11.6.2. Formulate a reasonable notion of a piecewise monotone function, and then show that all bounded piecewise monotone functions are Riemann integrable.

Exercise 11.6.3. Prove Proposition 11.6.4. (Hint: what is the relationship between the sum $\sum_{n=1}^N f(n)$, the sum $\sum_{n=0}^{N-1} f(n)$, and the integral $\int_{[0,N]} f$?)

Exercise 11.6.4. Give examples to show that both directions of the integral test break down if f is not assumed to be monotone decreasing.

Exercise 11.6.5. Use Proposition 11.6.4 to prove Corollary 11.6.5.

11.7 A non-Riemann integrable function

We have shown that there are large classes of bounded functions which are Riemann integrable. Unfortunately, there do exist bounded functions which are not Riemann integrable:

Proposition 11.7.1. *Let $f : [0,1] \to \mathbf{R}$ be the discontinuous function*

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

considered in Example 9.3.21. Then f is bounded but not Riemann integrable.

Proof. It is clear that f is bounded, so let us show that it is not Riemann integrable.

Let $\mathbf{P}$ be any partition of $[0,1]$. For any $J \in \mathbf{P}$, observe that if J is not a point or the empty set, then

$$\sup_{x \in J} f(x) = 1$$

(by Proposition 5.4.14). In particular we have

$$\left(\sup_{x \in J} f(x)\right)|J| = |J|.$$

(Note this is also true when J is a point, since both sides are zero.) In particular we see that

$$U(f, \mathbf{P}) = \sum_{J \in \mathbf{P}: J \neq \emptyset} |J| = [0,1] = 1$$

by Theorem 11.1.13; note that the empty set does not contribute anything to the total length. In particular we have $\overline{\int}_{[0,1]} f = 1$, by Proposition 11.3.12.

A similar argument gives that

$$\inf_{x \in J} f(x) = 0$$

for all J (other than points or the empty set), and so

$$L(f, \mathbf{P}) = \sum_{J \in \mathbf{P}: J \neq \emptyset} 0 = 0.$$

In particular we have $\underline{\int}_{[0,1]} f = 0$, by Proposition 11.3.12. Thus the upper and lower Riemann integrals do not match, and so this function is not Riemann integrable. □

Remark 11.7.2. As you can see, it is only rather "artificial" bounded functions which are not Riemann integrable. Because of this, the Riemann integral is good enough for a large majority of cases. There are ways to generalize or improve this integral, though. One of these is the *Lebesgue integral*, which we will define in Chapter 19. Another is the *Riemann-Stieltjes integral* $\int_I f d\alpha$, where $\alpha : I \to \mathbf{R}$ is a monotone increasing function, which we define in the next section.

11.8 The Riemann-Stieltjes integral

Let I be a bounded interval, let $\alpha : I \to \mathbf{R}$ be a monotone increasing function, and let $f : I \to \mathbf{R}$ be a function. Then there is a generalization of the Riemann integral, known as the *Riemann-Stieltjes integral.* This integral is defined just like the Riemann integral, but with one twist: instead of taking the length $|J|$ of intervals J, we take the α-length $\alpha[J]$, defined as follows. If J is a point or the empty set, then $\alpha[J] := 0$. If J is an interval of the form $[a,b]$, (a,b), $(a,b]$, or $[a,b)$, then $\alpha[J] := \alpha(b) - \alpha(a)$. Note that in the special case where α is the identity function $\alpha(x) := x$, then $\alpha[J]$ is just the same as $|J|$. However, for more general monotone functions α, the α-length $\alpha[J]$ is a different quantity from $|J|$. Nevertheless, it turns out one can still do much of the above theory, but replacing $|J|$ by $\alpha[J]$ throughout.

Definition 11.8.1 (α-length). Let I be a bounded interval, and let $\alpha : X \to \mathbf{R}$ be a function defined on some domain X which contains I. Then we define the *α-length* $\alpha[I]$ of I as follows. If I is a point or the empty set, we set $\alpha[I] = 0$. If I is an interval of the form $[a,b]$, $[a,b)$, $(a,b]$, or (a,b) for some $b > a$, then we set $\alpha[I] = \alpha(b) - \alpha(a)$.

Example 11.8.2. Let $\alpha : \mathbf{R} \to \mathbf{R}$ be the function $\alpha(x) := x^2$. Then $\alpha[[2,3]] = \alpha(3) - \alpha(2) = 9 - 4 = 5$, while $\alpha[(-3,-2)] = -5$. Meanwhile $\alpha[\{2\}] = 0$ and $\alpha[\emptyset] = 0$.

Example 11.8.3. Let $\alpha : \mathbf{R} \to \mathbf{R}$ be the identity function $\alpha(x) := x$. Then $\alpha[I] = |I|$ for all bounded intervals I (why?) Thus the notion of length is a special case of the notion of α-length.

We sometimes write $\alpha|_a^b$ or $\alpha(x)|_{x=a}^{x=b}$ instead of $\alpha[[a, b]]$.

One of the key theorems for the theory of the Riemann integral was Theorem 11.1.13, which concerned length and partitions, and in particular showed that $|I| = \sum_{J \in \mathbf{P}} |J|$ whenever $\mathbf{P}$ was a partition of I. We now generalize this slightly.

Lemma 11.8.4. *Let I be a bounded interval, let $\alpha : X \to \mathbf{R}$ be a function defined on some domain X which contains I, and let $\mathbf{P}$ be a partition of I. Then we have*

$$\alpha[I] = \sum_{J \in \mathbf{P}} \alpha[J].$$

Proof. See Exercise 11.8.1. □

We can now define a generalization of Definition 11.2.9.

Definition 11.8.5 (P.c. Riemann-Stieltjes integral). Let I be a bounded interval, and let $\mathbf{P}$ be a partition of I. Let $\alpha : X \to \mathbf{R}$ be a function defined on some domain X which contains I, and let $f : I \to \mathbf{R}$ be a function which is piecewise constant with respect to $\mathbf{P}$. Then we define

$$p.c. \int_{[\mathbf{P}]} f \, d\alpha := \sum_{J \in \mathbf{P}} c_J \alpha[J]$$

where c_J is the constant value of f on J.

Example 11.8.6. Let $f : [1, 3] \to \mathbf{R}$ be the function

$$f(x) = \begin{cases} 4 & \text{when } x \in [1, 2) \\ 2 & \text{when } x \in [2, 3], \end{cases}$$

let $\alpha : \mathbf{R} \to \mathbf{R}$ be the function $\alpha(x) := x^2$, and let $\mathbf{P}$ be the partition $\mathbf{P} := \{[1, 2), [2, 3]\}$. Then

$$p.c. \int_{[\mathbf{P}]} f \, d\alpha = c_{[1,2)}\alpha[[1, 2)] + c_{[2,3]}\alpha[[2, 3]]$$
$$= 4(\alpha(2) - \alpha(1)) + 2(\alpha(3) - \alpha(2)) = 4 \times 3 + 2 \times 5 = 22.$$

Example 11.8.7. Let $\alpha : \mathbf{R} \to \mathbf{R}$ be the identity function $\alpha(x) := x$. Then for any bounded interval I, any partition $\mathbf{P}$ of I, and any function f that is piecewise constant with respect to $\mathbf{P}$, we have $p.c. \int_{[\mathbf{P}]} f \, d\alpha = p.c. \int_{[\mathbf{P}]} f$ (why?).

We can obtain an exact analogue of Proposition 11.2.13 by replacing all the integrals $p.c. \int_{[\mathbf{P}]} f$ in the proposition with $p.c. \int_{[\mathbf{P}]} f \, d\alpha$ (Exercise 11.8.2). We can thus define $p.c. \int_I f \, d\alpha$ for any piecewise constant function $f : I \to \mathbf{R}$ and any $\alpha : X \to \mathbf{R}$ defined on a domain containing I, in analogy to before, by the formula

$$p.c. \int_I f \, d\alpha := p.c. \int_{[\mathbf{P}]} f \, d\alpha$$

for any partition $\mathbf{P}$ on I with respect to which f is piecewise constant.

Up until now, our function $\alpha : \mathbf{R} \to \mathbf{R}$ could have been arbitrary. Let us now assume that α is *monotone increasing*, i.e., $\alpha(y) \geq \alpha(x)$ whenever $x, y \in X$ are such that $y \geq x$. This implies that $\alpha(I) \geq 0$ for all intervals in X (why?). From this one can easily verify that all the results from Theorem 11.2.16 continue to hold when the integrals $p.c. \int_I f$ are replaced by $p.c. \int_I f \, d\alpha$, and the lengths $|I|$ are replaced by the α-lengths $\alpha(I)$; see Exercise 11.8.3.

We can then define upper and lower Riemann-Stieltjes integrals $\overline{\int}_I f \, d\alpha$ and $\underline{\int}_I f \, d\alpha$ whenever $f : I \to \mathbf{R}$ is bounded and α is defined on a domain containing I, by the usual formulae

$$\overline{\int}_I f \, d\alpha := \inf\{p.c. \int_I g \, d\alpha : g \text{ is p.c. on } I \text{ and majorizes } f\}$$

and

$$\underline{\int}_I f \, d\alpha := \sup\{p.c. \int_I g \, d\alpha : g \text{ is p.c. on } I \text{ and minorizes } f\}.$$

We then say that f is *Riemann-Stieltjes integrable on I with respect to α* if the upper and lower Riemann-Stieltjes integrals match, in which case we set

$$\int_I f \, d\alpha := \overline{\int}_I f \, d\alpha = \underline{\int}_I f \, d\alpha.$$

As before, when α is the identity function $\alpha(x) := x$ then the Riemann-Stieltjes integral is identical to the Riemann integral; thus the Riemann-Stieltjes integral is a generalization of the Riemann integral.

(We shall see another comparison between the two integrals a little later, in Corollary 11.10.3.) Because of this, we sometimes write $\int_I f$ as $\int_I f\,dx$ or $\int_I f(x)\,dx$.

Most (but not all) of the remaining theory of the Riemann integral then can be carried over without difficulty, replacing Riemann integrals with Riemann-Stieltjes integrals and lengths with α-lengths. There are a couple results which break down; Theorem 11.4.1(g), Proposition 11.5.3, and Proposition 11.5.6 are not necessarily true when α is discontinuous at key places (e.g., if f and α are both discontinuous at the same point, then $\int_I f\,d\alpha$ is unlikely to be defined. However, Theorem 11.5.1 is still true (Exercise 11.8.4).

— Exercises —

Exercise 11.8.1. Prove Lemma 11.8.4. (Hint: modify the proof of Theorem 11.1.13.)

Exercise 11.8.2. State and prove a version of Proposition 11.2.13 for the Riemann-Stieltjes integral.

Exercise 11.8.3. State and prove a version of Theorem 11.2.16 for the Riemann-Stieltjes integral.

Exercise 11.8.4. State and prove a version of Theorem 11.5.1 for the Riemann-Stieltjes integral. (Hint: one has to be careful with the proof; the problem here is that some of the references to the length of $|J_k|$ should remain unchanged, and other references to the length of $|J_k|$ should be changed to the α-length $\alpha(J_k)$ - basically, all of the occurrences of $|J_k|$ which appear inside a summation should be replaced with $\alpha(J_k)$, but the rest should be unchanged.)

Exercise 11.8.5. Let $\mathrm{sgn} : \mathbf{R} \to \mathbf{R}$ be the signum function

$$\mathrm{sgn}(x) := \begin{cases} 1 & \text{when } x > 0 \\ 0 & \text{when } x = 0 \\ -1 & \text{when } x < 0. \end{cases}$$

Let $f : [-1, 1] \to \mathbf{R}$ be a continuous function. Show that f is Riemann-Stieltjes integrable with respect to sgn, and that

$$\int_{[-1,1]} f\,d\mathrm{sgn} = 2f(0).$$

(Hint: for every $\varepsilon > 0$, find piecewise constant functions majorizing and minorizing f whose Riemann-Stieltjes integral is ε-close to $2f(0)$.)

11.9 The two fundamental theorems of calculus

We now have enough machinery to connect integration and differentiation via the familiar fundamental theorem of calculus. Actually, there

are two such theorems, one involving the derivative of the integral, and the other involving the integral of the derivative.

Theorem 11.9.1 (First Fundamental Theorem of Calculus). *Let $a < b$ be real numbers, and let $f : [a, b] \to \mathbf{R}$ be a Riemann integrable function. Let $F : [a, b] \to \mathbf{R}$ be the function*

$$F(x) := \int_{[a,x]} f.$$

Then F is continuous. Furthermore, if $x_0 \in [a, b]$ and f is continuous at x_0, then F is differentiable at x_0, and $F'(x_0) = f(x_0)$.

Proof. Since f is Riemann integrable, it is bounded (by Definition 11.3.4). Thus we have some real number M such that $-M \le f(x) \le M$ for all $x \in [a, b]$.

Now let $x < y$ be two elements of $[a, b]$. Then notice that

$$F(y) - F(x) = \int_{[a,y]} f - \int_{[a,x]} f = \int_{(x,y]} f$$

by Theorem 11.4.1(h). By Theorem 11.4.1(e) we thus have

$$\int_{[x,y]} f \le \int_{[x,y]} M = p.c. \int_{[x,y]} M = M(y - x)$$

and

$$\int_{[x,y]} f \ge \int_{[x,y]} -M = p.c. \int_{[x,y]} -M = -M(y - x)$$

and thus

$$|F(y) - F(x)| \le M(y - x).$$

This is for $y > x$. By interchanging x and y we thus see that

$$|F(y) - F(x)| \le M(x - y)$$

when $x > y$. Also, we have $F(y) - F(x) = 0$ when $x = y$. Thus in all three cases we have

$$|F(y) - F(x)| \le M|x - y|.$$

Now let $x \in [a, b]$, and let $(x_n)_{n=0}^{\infty}$ be any sequence in $[a, b]$ converging to x. Then we have

$$-M|x_n - x| \le F(x_n) - F(x) \le M|x_n - x|$$

for each n. But $-M|x_n - x|$ and $M|x_n - x|$ both converge to 0 as $n \to \infty$, so by the squeeze test $F(x_n) - F(x)$ converges to 0 as $n \to \infty$, and thus $\lim_{n\to\infty} F(x_n) = F(x)$. Since this is true for all sequences $x_n \in [a,b]$ converging to x, we thus see that F is continuous at x. Since x was an arbitrary element of $[a,b]$, we thus see that F is continuous.

Now suppose that $x_0 \in [a,b]$, and f is continuous at x_0. Choose any $\varepsilon > 0$. Then by continuity, we can find a $\delta > 0$ such that $|f(x) - f(x_0)| \leq \varepsilon$ for all x in the interval $I := [x_0 - \delta, x_0 + \delta] \cap [a,b]$, or in other words

$$f(x_0) - \varepsilon \leq f(x) \leq f(x_0) + \varepsilon \text{ for all } x \in I.$$

We now show that

$$|F(y) - F(x_0) - f(x_0)(y - x_0)| \leq \varepsilon|y - x_0|$$

for all $y \in I$, since Proposition 10.1.7 will then imply that F is differentiable at x_0 with derivative $F'(x_0) = f(x_0)$ as desired.

Now fix $y \in I$. There are three cases. If $y = x_0$, then $F(y) - F(x_0) - f(x_0)(y - x_0) = 0$ and so the claim is obvious. If $y > x_0$, then

$$F(y) - F(x_0) = \int_{[x_0,y]} f.$$

Since $x_0, y \in I$, and I is a connected set, then $[x_0, y]$ is a subset of I, and thus we have

$$f(x_0) - \varepsilon \leq f(x) \leq f(x_0) + \varepsilon \text{ for all } x \in [x_0, y],$$

and thus

$$(f(x_0) - \varepsilon)(y - x_0) \leq \int_{[x_0,y]} f \leq (f(x_0) + \varepsilon)(y - x_0)$$

and so in particular

$$|F(y) - F(x_0) - f(x_0)(y - x_0)| \leq \varepsilon|y - x_0|$$

as desired. The case $y < x_0$ is similar and is left to the reader. □

Example 11.9.2. Recall in Exercise 9.8.5 that we constructed a monotone function $f : \mathbf{R} \to \mathbf{R}$ which was discontinuous at every rational and continuous everywhere else. By Proposition 11.6.1, this monotone function is Riemann integrable on $[0,1]$. If we define $F : [0,1] \to \mathbf{R}$ by $F(x) := \int_{[0,x]} f$, then F is a continuous function which is differentiable at every irrational number. On the other hand, F is non-differentiable at every rational number; see Exercise 11.9.1.

Informally, the first fundamental theorem of calculus asserts that

$$\left(\int_{[a,x]} f\right)'(x) = f(x)$$

given a certain number of assumptions on f. Roughly, this means that the derivative of an integral recovers the original function. Now we show the reverse, that the integral of a derivative recovers the original function.

Definition 11.9.3 (Antiderivatives). Let I be a bounded interval, and let $f : I \to \mathbf{R}$ be a function. We say that a function $F : I \to \mathbf{R}$ is an *antiderivative* of f if F is differentiable on I and $F'(x) = f(x)$ for all $x \in I$.

Theorem 11.9.4 (Second Fundamental Theorem of Calculus). *Let $a < b$ be real numbers, and let $f : [a,b] \to \mathbf{R}$ be a Riemann integrable function. If $F : [a,b] \to \mathbf{R}$ is an antiderivative of f, then*

$$\int_{[a,b]} f = F(b) - F(a).$$

Proof. We will use Riemann sums. The idea is to show that

$$U(f, \mathbf{P}) \geq F(b) - F(a) \geq L(f, \mathbf{P})$$

for every partition $\mathbf{P}$ of $[a,b]$. The left inequality asserts that $F(b) - F(a)$ is a lower bound for $\{U(f,\mathbf{P}) : \mathbf{P}$ is a partition of $[a,b]\}$, while the right inequality asserts that $F(b) - F(a)$ is an upper bound for $\{L(f,\mathbf{P}) : \mathbf{P}$ is a partition of $[a,b]\}$. But by Proposition 11.3.12, this means that

$$\overline{\int}_{[a,b]} f \geq F(b) - F(a) \geq \underline{\int}_{[a,b]} f,$$

but since f is assumed to be Riemann integrable, both the upper and lower Riemann integral equal $\int_{[a,b]} f$. The claim follows.

We have to show the bound $U(f,\mathbf{P}) \geq F(b) - F(a) \geq L(f,\mathbf{P})$. We shall just show the first inequality $U(f,\mathbf{P}) \geq F(b) - F(a)$; the other inequality is similar.

Let $\mathbf{P}$ be a partition of $[a,b]$. From Lemma 11.8.4 we have

$$F(b) - F(a) = \sum_{J \in \mathbf{P}} F[J] = \sum_{J \in \mathbf{P}: J \neq \emptyset} F[J],$$

while from definition we have

$$U(f,\mathbf{P}) = \sum_{J\in\mathbf{P}:J\neq\emptyset} \sup_{x\in J} f(x)|J|.$$

Thus it will suffice to show that

$$F[J] \leq \sup_{x\in J} f(x)|J|$$

for all $J \in \mathbf{P}$ (other than the empty set).

When J is a point then the claim is clear, since both sides are zero. Now suppose that $J = [c,d], (c,d], [c,d)$, or (c,d) for some $c < d$. Then the left-hand side is $F[J] = F(d) - F(c)$. By the mean-value theorem, this is equal to $(d-c)F'(e)$ for some $e \in J$. But since $F'(e) = f(e)$, we thus have

$$F[J] = (d-c)f(e) = f(e)|J| \leq \sup_{x\in J} f(x)|J|$$

as desired. □

Of course, as you are all aware, one can use the second fundamental theorem of calculus to compute integrals relatively easily provided that you can find an anti-derivative of the integrand f. Note that the first fundamental theorem of calculus ensures that every *continuous* Riemann integrable function has an anti-derivative. For discontinuous functions, the situation is more complicated, and is a graduate-level real analysis topic which will not be discussed here. Also, not every function with an anti-derivative is Riemann integrable; as an example, consider the function $F : [-1,1] \to \mathbf{R}$ defined by $F(x) := x^2\sin(1/x^3)$ when $x \neq 0$, and $F(0) := 0$. Then F is differentiable everywhere (why?), so F' has an antiderivative, but F' is unbounded (why?), and so is not Riemann integrable.

We now pause to mention the infamous "$+C$" ambiguity in anti-derivatives:

Lemma 11.9.5. *Let I be a bounded interval, and let $f : I \to \mathbf{R}$ be a function. Let $F : I \to \mathbf{R}$ and $G : I \to \mathbf{R}$ be two antiderivatives of f. Then there exists a real number C such that $F(x) = G(x) + C$ for all $x \in I$.*

Proof. See Exercise 11.9.2. □

— Exercises —

Exercise 11.9.1. Let $f : [0,1] \to \mathbf{R}$ be the function in Exercise 9.8.5. Show that for every rational number $q \in \mathbf{Q} \cap [0,1]$, the function $F : [0,1] \to \mathbf{R}$ defined by the formula $F(x) := \int_0^x f(y)\,dy$ is not differentiable at q. (Hint: use the mean-value theorem, Corollary 10.2.9.)

Exercise 11.9.2. Prove Lemma 11.9.5. (Hint: apply the mean-value theorem, Corollary 10.2.9, to the function $F - G$. One can also prove this lemma using the second Fundamental theorem of calculus (how?), but one has to be careful since we do not assume f to be Riemann integrable.)

Exercise 11.9.3. Let $a < b$ be real numbers, and let $f : [a,b] \to \mathbf{R}$ be a monotone increasing function. Let $F : [a,b] \to \mathbf{R}$ be the function $F(x) := \int_{[a,x]} f$. Let x_0 be an element of (a,b). Show that F is differentiable at x_0 if and only if f is continuous at x_0. (Hint: one direction is taken care of by one of the fundamental theorems of calculus. For the other, consider left and right limits of f and argue by contradiction.)

11.10 Consequences of the fundamental theorems

We can now give a number of useful consequences of the fundamental theorems of calculus (beyond the obvious application, that one can now compute any integral for which an anti-derivative is known). The first application is the familiar integration by parts formula.

Proposition 11.10.1 (Integration by parts formula). *Let $I = [a,b]$, and let $F : [a,b] \to \mathbf{R}$ and $G : [a,b] \to \mathbf{R}$ be differentiable functions on $[a,b]$ such that F' and G' are Riemann integrable on I. Then we have*

$$\int_{[a,b]} FG' = F(b)G(b) - F(a)G(a) - \int_{[a,b]} F'G.$$

Proof. See Exercise 11.10.1. □

Next, we show that under certain circumstances, one can write a Riemann-Stieltjes integral as a Riemann integral. We begin with piecewise constant functions.

Theorem 11.10.2. *Let $\alpha : [a,b] \to \mathbf{R}$ be a monotone increasing function, and suppose that α is also differentiable on $[a,b]$, with α' being Riemann integrable. Let $f : [a,b] \to \mathbf{R}$ be a piecewise constant function on $[a,b]$. Then $f\alpha'$ is Riemann integrable on $[a,b]$, and*

$$\int_{[a,b]} f\,d\alpha = \int_{[a,b]} f\alpha'.$$

Proof. Since f is piecewise constant, it is Riemann integrable, and since α' is also Riemann integrable, then $f\alpha'$ is Riemann integrable by Theorem 11.4.5.

Suppose that f is piecewise constant with respect to some partition $\mathbf{P}$ of $[a, b]$; without loss of generality we may assume that $\mathbf{P}$ does not contain the empty set. Then we have

$$\int_{[a,b]} f\, d\alpha = p.c. \int_{[\mathbf{P}]} f\, d\alpha = \sum_{J \in \mathbf{P}} c_J \alpha[J]$$

where c_J is the constant value of f on J. On the other hand, from Theorem 11.2.16(h) (generalized to partitions of arbitrary length - why is this generalization true?) we have

$$\int_{[a,b]} f\alpha' = \sum_{J \in \mathbf{P}} \int_J f\alpha' = \sum_{J \in \mathbf{P}} \int_J c_J \alpha' = \sum_{J \in \mathbf{P}} c_J \int_J \alpha'.$$

But by the second fundamental theorem of calculus (Theorem 11.9.4), $\int_J \alpha' = \alpha[J]$, and the claim follows. □

Corollary 11.10.3. *Let $\alpha : [a, b] \to \mathbf{R}$ be a monotone increasing function, and suppose that α is also differentiable on $[a, b]$, with α' being Riemann integrable. Let $f : [a, b] \to \mathbf{R}$ be a function which is Riemann-Stieltjes integrable with respect to α on $[a, b]$. Then $f\alpha'$ is Riemann integrable on $[a, b]$, and*

$$\int_{[a,b]} f\, d\alpha = \int_{[a,b]} f\alpha'.$$

Proof. Note that since f and α' are bounded, then $f\alpha'$ must also be bounded. Also, since α is monotone increasing and differentable, α' is non-negative.

Let $\varepsilon > 0$. Then, we can find a piecewise constant function $\overline{f}$ majorizing f on $[a, b]$, and a piecewise constant function $\underline{f}$ minorizing f on $[a, b]$, such that

$$\int_{[a,b]} f\, d\alpha - \varepsilon \leq \int_{[a,b]} \underline{f}\, d\alpha \leq \int_{[a,b]} \overline{f}\, d\alpha \leq \int_{[a,b]} f\, d\alpha + \varepsilon.$$

Applying Theorem 11.10.2, we obtain

$$\int_{[a,b]} f\, d\alpha - \varepsilon \leq \int_{[a,b]} \underline{f}\alpha' \leq \int_{[a,b]} \overline{f}\, \alpha' \leq \int_{[a,b]} f\, d\alpha + \varepsilon.$$

Since α' is non-negative and $\underline{f}$ minorizes f, then $\underline{f}\alpha'$ minorizes $f\alpha'$. Thus $\int_{[a,b]} \underline{f}\alpha' \leq \underline{\int}_{[a,b]} f\alpha'$ (why?). Thus

$$\int_{[a,b]} f\, d\alpha - \varepsilon \leq \underline{\int}_{[a,b]} f\alpha'.$$

Similarly we have

$$\overline{\int}_{[a,b]} f\alpha' \leq \int_{[a,b]} f\, d\alpha + \varepsilon.$$

Since these statements are true for any $\varepsilon > 0$, we must have

$$\int_{[a,b]} f\, d\alpha \leq \underline{\int}_{[a,b]} f\alpha' \leq \overline{\int}_{[a,b]} f\alpha' \leq \int_{[a,b]} f\, d\alpha$$

and the claim follows. □

Remark 11.10.4. Informally, Corollary 11.10.3 asserts that $f\, d\alpha$ is essentially equivalent to $f \frac{d\alpha}{dx} dx$, when α is differentiable. However, the advantage of the Riemann-Stieltjes integral is that it still makes sense even when α is not differentiable.

We now build up to the familiar change of variables formula. We first need a preliminary lemma.

Lemma 11.10.5 (Change of variables formula I). *Let $[a, b]$ be a closed interval, and let $\phi : [a, b] \to [\phi(a), \phi(b)]$ be a continuous monotone increasing function. Let $f : [\phi(a), \phi(b)] \to \mathbf{R}$ be a piecewise constant function on $[\phi(a), \phi(b)]$. Then $f \circ \phi : [a, b] \to \mathbf{R}$ is also piecewise constant on $[a, b]$, and*

$$\int_{[a,b]} f \circ \phi\, d\phi = \int_{[\phi(a),\phi(b)]} f.$$

Proof. We give a sketch of the proof, leaving the gaps to be filled in Exercise 11.10.2. Let $\mathbf{P}$ be a partition of $[\phi(a), \phi(b)]$ such that f is piecewise constant with respect to $\mathbf{P}$; we may assume that $\mathbf{P}$ does not contain the empty set. For each $J \in \mathbf{P}$, let c_J be the constant value of f on J, thus

$$\int_{[\phi(a),\phi(b)]} f = \sum_{J \in \mathbf{P}} c_J |J|.$$

For each interval J, let $\phi^{-1}(J)$ be the set $\phi^{-1}(J) := \{x \in [a, b] : \phi(x) \in J\}$. Then $\phi^{-1}(J)$ is connected (why?), and is thus an interval. Furthermore, c_J is the constant value of $f \circ \phi$ on $\phi^{-1}(J)$ (why?). Thus, if we define $\mathbf{Q} := \{\phi^{-1}(J) : J \in \mathbf{P}\}$ (ignoring the fact that $\mathbf{Q}$ has been used to represent the rational numbers), then $\mathbf{Q}$ partitions $[a,b]$ (why?), and $f \circ \phi$ is piecewise constant with respect to $\mathbf{Q}$ (why?). Thus

$$\int_{[a,b]} f \circ \phi \, d\phi = \int_{[\mathbf{Q}]} f \circ \phi \, d\phi = \sum_{J \in \mathbf{P}} c_J \phi[\phi^{-1}(J)].$$

But $\phi[\phi^{-1}(J)] = |J|$ (why?), and the claim follows. □

Proposition 11.10.6 (Change of variables formula II). *Let $[a,b]$ be a closed interval, and let $\phi : [a,b] \to [\phi(a), \phi(b)]$ be a continuous monotone increasing function. Let $f : [\phi(a), \phi(b)] \to \mathbf{R}$ be a Riemann integrable function on $[\phi(a), \phi(b)]$. Then $f \circ \phi : [a,b] \to \mathbf{R}$ is Riemann-Stieltjes integrable with respect to ϕ on $[a,b]$, and*

$$\int_{[a,b]} f \circ \phi \, d\phi = \int_{[\phi(a),\phi(b)]} f.$$

Proof. This will be obtained from Lemma 11.10.5 in a similar manner to how Corollary 11.10.3 was obtained from Theorem 11.10.2. First observe that since f is Riemann integrable, it is bounded, and then $f \circ \phi$ must also be bounded (why?).

Let $\varepsilon > 0$. Then, we can find a piecewise constant function $\overline{f}$ majorizing f on $[\phi(a), \phi(b)]$, and a piecewise constant function $\underline{f}$ minorizing f on $[\phi(a), \phi(b)]$, such that

$$\int_{[\phi(a),\phi(b)]} f - \varepsilon \leq \int_{[\phi(a),\phi(b)]} \underline{f} \leq \int_{[\phi(a),\phi(b)]} \overline{f} \leq \int_{[\phi(a),\phi(b)]} f + \varepsilon.$$

Applying Lemma 11.10.5, we obtain

$$\int_{[\phi(a),\phi(b)]} f - \varepsilon \leq \int_{[a,b]} \underline{f} \circ \phi \, d\phi \leq \int_{[a,b]} \overline{f} \circ \phi \, d\phi \leq \int_{[\phi(a),\phi(b)]} f + \varepsilon.$$

Since $\underline{f} \circ \phi$ is piecewise constant and minorizes $f \circ \phi$, we have

$$\int_{[a,b]} \underline{f} \circ \phi \, d\phi \leq \underline{\int}_{[a,b]} f \circ \phi \, d\phi$$

while similarly we have

$$\int_{[a,b]} \overline{f} \circ \phi \, d\phi \geq \overline{\int}_{[a,b]} f \circ \phi \, d\phi.$$

Thus

$$\int_{[\phi(a),\phi(b)]} f - \varepsilon \leq \underline{\int}_{[a,b]} f \circ \phi \, d\phi \leq \overline{\int}_{[a,b]} f \circ \phi \, d\phi \leq \int_{[\phi(a),\phi(b)]} f + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this implies that

$$\int_{[\phi(a),\phi(b)]} f \leq \underline{\int}_{[a,b]} f \circ \phi \, d\phi \leq \overline{\int}_{[a,b]} f \circ \phi \, d\phi \leq \int_{[\phi(a),\phi(b)]} f$$

and the claim follows. □

Combining this formula with Corollary 11.10.3, one immediately obtains the following familiar formula:

Proposition 11.10.7 (Change of variables formula III). *Let $[a,b]$ be a closed interval, and let $\phi : [a,b] \to [\phi(a), \phi(b)]$ be a differentiable monotone increasing function such that ϕ' is Riemann integrable. Let $f : [\phi(a), \phi(b)] \to \mathbf{R}$ be a Riemann integrable function on $[\phi(a), \phi(b)]$. Then $(f \circ \phi)\phi' : [a,b] \to \mathbf{R}$ is Riemann integrable on $[a,b]$, and*

$$\int_{[a,b]} (f \circ \phi)\phi' = \int_{[\phi(a),\phi(b)]} f.$$

— Exercises —

Exercise 11.10.1. Prove Proposition 11.10.1. (Hint: first use Corollary 11.5.2 and Theorem 11.4.5 to show that FG' and $F'G$ are Riemann integrable. Then use the product rule (Theorem 10.1.13(d)).)

Exercise 11.10.2. Fill in the gaps marked (why?) in the proof of Lemma 11.10.5.

Exercise 11.10.3. Let $a < b$ be real numbers, and let $f : [a,b] \to \mathbf{R}$ be a Riemann integrable function. Let $g : [-b,-a] \to \mathbf{R}$ be defined by $g(x) := f(-x)$. Show that g is also Riemann integrable, and $\int_{[-b,-a]} g = \int_{[a,b]} f$.

Exercise 11.10.4. What is the analogue of Proposition 11.10.7 when ϕ is monotone decreasing instead of monotone increasing? (When ϕ is neither monotone increasing or monotone decreasing, the situation becomes significantly more complicated.)