Chapter 8

THE INTEGRAL

8.1 Introduction

Calculus students learn two processes, both of which are described as "integration". The following two examples should be quite familiar:

$$\int x^3\,dx = x^4/4 + C$$

and

$$\int_1^2 x^3\,dx = 2^4/4 - 1/4 = 16/4 - 1/4 = 15/4.$$

The first is called an *indefinite integral* or *antiderivative* and the second a *definite integral*. The use of nearly identical notation, terminology and methods of computation does a lot to confuse the underlying meanings. Many calculus students would be hard pressed to make a distinction.

Indeed even for many eighteenth century mathematicians these two very different procedures were not much distinguished. It was a great discovery that the computation of an area could be achieved by finding an antiderivative. It is attributed to Newton but vague ideas along this line can be found in the thinking of earlier authors. For these mathematicians a definite integral was defined directly in terms of the antiderivative.

This is most unfortunate for the development of a rigorous theory and this was recognized by Cauchy. He saw clearly that it was vital that the meaning of the definite integral be separated from the indefinite integral and given a precise definition independent of it. For this he turned to the geometry of the Greeks who had long ago described a method for computing areas of regions enclosed by curves.

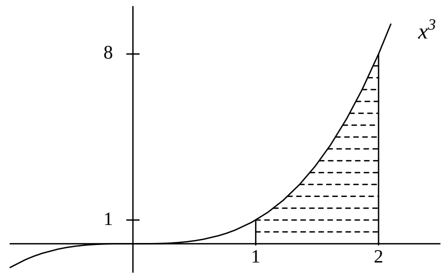

Figure 8.1: Region bounded by $x = 1$, $x = 2$, $y = x^3$, and $y = 0$.

This method, the so-called method of exhaustion, involves computing the areas of simpler figures (squares, triangles, rectangles) that approximate the area of the region.

We return to the example $\int_1^2 x^3\,dx$ above interpreted as an area. The region is that bounded on the left and right by the lines $x = 1$ and $x = 2$, below by the line $y = 0$ and above by the curve $y = x^3$. (See Figure 8.1.)

Using the method of exhaustion we may place this figure inside a collection of rectangles by dividing the interval $[1, 2]$ into n equal sized subintervals each of length $1/n$. This means selecting the points

$$1,\ 1 + 1/n,\ 1 + 2/n,\ \ldots\ 1 + (n - 1)/n$$

and constructing rectangles with vertices at these points. The total area of these rectangles exceeds the true area and is precisely

$$\sum_{k=1}^{n}(1 + (k)/n)^3(1/n).$$

The method of exhaustion requires a lower estimate as well and the true area of the region must be greater than

$$\sum_{k=1}^{n}(1 + (k - 1)/n)^3(1/n).$$

(See Figure 8.2 for an illustration with $n = 4$.)

The method of exhaustion now requires us to show that as n increases *both* approximations, the upper one and the lower one, approach the same number. Cauchy saw that because of the continuity of the function $f(x) = x^3$ these limits would be the same. More than that any choice of points ξ_k from the interval $[1+(k-1)/n, 1+(k)/n]$

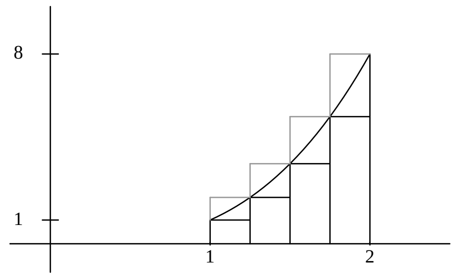

Figure 8.2: Method of exhaustion ($n = 4$).

would have the property that

$$\lim_{n\to\infty} \sum_{k=1}^{n} {\xi_k}^3 (1/n)$$

would exist.

This procedure, borrowed very heavily from the Greeks, will work for any continuous function and thus it offered to Cauchy a way to *define* the integral

$$\int_a^b f(x)\,dx$$

for any function f, continuous on an interval $[a, b]$, without any reference whatsoever to notions of derivatives or antiderivatives. The key ingredients here are first of all dividing the interval $[a, b]$ into a collection of nonoverlapping subintervals called *a partition* of $[a, b]$,

$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n],$$

(it is not important that they have equal size, just that they get small) and then forming the sums

$$\sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) \tag{1}$$

with respect to this partition. The only constraint on the choice of the points ξ_k is that each is taken from the appropriate interval $[x_{k-1}, x_k]$ of the partition; these are often called the *associated points.*

It is an unfortunate trick of fate that the sums (1) which originated with Cauchy are called *Riemann sums* because of Riemann's later (much later) use of them in defining his integral.

In this chapter we start with Cauchy's methods of integration and proceed to Riemann. The important thing for the student to keep track of is how this theory develops in a manner that assigns meaning to the integral of various classes of functions in a way quite distinct from how we would compute an integral in a calculus course. It is of course much easier to compute that $\int_1^2 x^3\,dx = 15/4$ in the familiar way, rather than as a limit of Riemann sums; but the meaning of this statement is properly given in this more difficult way.

8.2 Cauchy's First Method

Cauchy's first goal in defining an integral was to give meaning to the integral for continuous functions. The integral is defined as the limit of Riemann sums. Before such a definition is valid one must show that the limit exists. Thus the first step is the proof of the following theorem.

Theorem 8.1 (Cauchy) *Let f be a continuous function on an interval $[a,b]$. Then there is a number I (called the* definite integral *of f on $[a,b]$) such that for all $\varepsilon > 0$ there is a $\delta > 0$ so that*

$$\left|\sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) - I\right| < \varepsilon$$

whenever $[x_0,x_1]$, $[x_1,x_2]$, ... , $[x_{n-1},x_n]$, is a partition of the interval $[a,b]$ into subintervals of length less than δ and each ξ_k is a point in the interval $[x_{k-1},x_k]$.

Once the theorem is proved, then we can safely define the definite integral of a continuous function as that number I guaranteed by the theorem. Loosely speaking, we say that the integral is defined as a limit of Riemann sums (1).

Definition 8.2 Let f be a continuous function on an interval $[a,b]$. Then we define

$$\int_a^b f(x)\,dx$$

to be that number I whose existence is proved in Theorem 8.1.

Now we must prove Theorem 8.1.

Proof. For any particular partition (let us call it π)

$$[x_0,x_1],\ [x_1,x_2],\ \ldots,\ [x_{n-1},x_n]$$

of the interval $[a, b]$ write

$$M(\pi) = \sum_{k=1}^{n} \max\{f(x) : x \in [x_{k-1}, x_k]\}(x_k - x_{k-1})$$

and

$$m(\pi) = \sum_{k=1}^{n} \min\{f(x) : x \in [x_{k-1}, x_k]\}(x_k - x_{k-1})$$

Here $M(\pi)$ and $m(\pi)$ depend on the partition π. These are called the *upper sums* and *lower sums* for the partition. Note that any Riemann sum over this partition must lie somewhere between the lower sum and the upper sum.

Since f is continuous on $[a, b]$ it is uniformly continuous there (Theorem 5.43). Thus for every $\varepsilon > 0$ there is a $\delta > 0$ that depends on ε so that

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a}$$

if $|x - y| < \delta$. Since we shall need to find a different δ for many choices of ε let us write it as $\delta(\varepsilon)$.

Thus if the partition we are using has the property that every interval is shorter than $\delta(\varepsilon)$ we must have

$$\max\{f(x) : x \in [x_{k-1}, x_k]\} - \min\{f(x) : x \in [x_{k-1}, x_k]\} < \frac{\varepsilon}{b-a}$$

It follows that for such partitions $0 < M(\pi) - m(\pi) < \varepsilon$.

Select a sequence of partitions $\{\pi_n\}$ each one containing all the points of the previous partition and such that every interval in the nth partition π_n is shorter than $\delta(1/n)$. If $M(\pi_n)$ and $m(\pi_n)$ denote the corresponding sums for the nth partition of our sequence of partitions as above then

$$0 < M(\pi_n) - m(\pi_n) < 1/n.$$

One more technical point needs to be raised. As we add points to a partition the upper sums cannot increase nor can the lower sums decrease. Thus $M(\pi_n) \geq M(\pi_{n+1})$ while $m(\pi_n) \leq m(\pi_{n+1})$. (The details just require some inequality work and are left as Exercise 8:2.17.)

Thus the intervals $[m(\pi_n), M(\pi_n)]$ form a descending sequence with lengths shrinking to zero. By Cantor's intersection property (see Section 4.5.2) there is a number I so that $m(\pi_n) \to I$ and $M(\pi_n) \to I$ as $n \to \infty$. We shall show that I has the property of the theorem.

Now let $\varepsilon > 0$ and choose any partition π with the property that every interval is shorter than $\delta(\varepsilon/2)$. By what we have seen above, the interval $[m(\pi), M(\pi)]$ has length smaller than $\varepsilon/2$.

Any Riemann sum over the partition π must evidently belong to the interval $[m(\pi), M(\pi)]$. Let $N > 2/\varepsilon$. Suppose for a moment that the intervals $[m(\pi), M(\pi)]$ and $[m(\pi_N), M(\pi_N)]$ intersect at some point. In that case the Riemann sum over the partition π and the value I (which is inside the interval $[m(\pi_N), M(\pi_N)]$ must be closer together than $\varepsilon/2 + 1/N$ which is smaller than ε. As this is precisely what we want to prove we are done.

It remains to check that the two intervals

$$[m(\pi), M(\pi)] \quad \text{and} \quad [m(\pi_N), M(\pi_N)]$$

intersect at some point. To find a point common to these two intervals combine the two partitions π and π_N to form a partition containing all points in either partition. The Riemann sum over such a partition belongs to the interval $[m(\pi), M(\pi)]$ and also to the interval $[m(\pi_N), M(\pi_N)]$. This completes the proof. (That the number I here is unique is left as Exercise 8:2.2.) ■

A special case of this definition and this theorem allows us to compute an integral as a limit of a sequence. In practice this is seldom the best way to compute it, but it is interesting and useful in some parts of the theory.

Corollary 8.3 *Let f be a continuous function on an interval $[a, b]$ Then*

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + \frac{k}{n}(b-a)\right).$$

8.2.1 Scope of Cauchy's First Method

It is natural to ask whether this method of Cauchy for describing the integral of a continuous function would apply to a larger class of functions. But Cauchy did not ask this question. His goal was to assign a meaning for continuous functions, a class of functions that was quite large enough for most applications. The only limitation he might have seen was that this method would fail for functions having *infinite singularities* (i.e., discontinuity points where the function is unbounded). Thus he was led to the method we discuss in Section 8.4 as Cauchy's second method. Cauchy and other mathematicians of his time were sufficiently confused as to the meaning of the word "function" that they might never have asked such a question.

But we can. And many years later Riemann did too as we shall see in Section 8.6. In the exercises you are asked to prove the following two statements:

> *The first method of Cauchy will fail if applied to an unbounded function f on an interval $[a, b]$.*
>
> *The first method of Cauchy succeeds if applied to any function f that is bounded on an interval $[a, b]$ and has only finitely many discontinuities there.*

The first statement shows that the method used here to define an integral is severely limited. It can never be used for unbounded functions. Since we have restricted it here to continuous functions that is no problem; any function continuous on an interval $[a, b]$ is bounded there.

The second statement shows that the method is not, however, limited only to continuous functions even though that was Cauchy's intention. Later on we will use the method to define Riemann's integral which applies to a large class of (bounded) functions which are permitted to have many, even infinitely many, points of discontinuity.

Exercises

8:2.1 To complete the computations in the introduction to this chapter show that

$$\lim_{n\to\infty}\sum_{k=1}^{n}(1+(k)/n)^3(1/n) = 15/4.$$

This computation alone should be enough to convince you that the definition is intended theoretically and hardly ever used to compute integrals.

8:2.2 Show that the number I in the statement of Theorem 8.1 is unique, i.e., that there cannot be two numbers that would be assigned to the symbol $\int_a^b f(x)\,dx$.

8:2.3 If f is constant and $f(x) = \alpha$ for all x in $[a, b]$ show that

$$\int_a^b f(x)\,dx = \alpha(b-a).$$

8:2.4 If f is continuous and $f(x) \geq 0$ for all x in $[a, b]$ show that

$$\int_a^b f(x)\,dx \geq 0.$$

8:2.5 If f is continuous and $m \le f(x) \le M$ for all x in $[a, b]$ show that

$$m(b-a) \le \int_a^b f(x)\,dx \le M(b-a).$$

8:2.6 Calculate $\int_0^1 x^p\,dx$ (for whatever values of p you can manage) by partitioning $[0, 1]$ into subintervals of equal length.

8:2.7 Calculate $\int_a^b x^p\,dx$ (for whatever values of p you can manage) by partitioning $[a, b]$ into subintervals $[a, aq]$, $[aq, aq^2]$, ... $[aq^{n-1}, b]$ where $aq^n = b$. (Note that the subintervals are not of equal length, but that the lengths form a geometric progression.)

8:2.8 Use the method of the preceding exercise to show that

$$\int_1^2 \frac{dx}{x^2} = \frac{1}{2}$$

and check it by the usual calculus method.

8:2.9 Compute the Riemann sums for the integral $\int_a^b x^{-2}\,dx$ $(a > 0)$ taken over a partition

$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n]$$

of the interval $[a, b]$ and with associated points $\xi_i = \sqrt{x_i x_{i-1}}$. What can you conclude from this?

8:2.10 Compute the Riemann sums for the integral $\int_a^b x^{-1/2}\,dx$ $(a > 0)$ taken over a partition

$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n]$$

of the interval $[a, b]$ and with associated points

$$\xi_i = \left(\frac{\sqrt{x_i} + \sqrt{x_{i-1}}}{2}\right)^2.$$

What can you conclude from this?

8:2.11 Show that

$$\lim_{n\to\infty} n \left\{ \frac{1}{(n+1)^2} + \frac{1}{(n+2)^2} + \frac{1}{(n+3)^2} + \cdots \frac{1}{(2n)^2} \right\} = \frac{1}{2}.$$

8:2.12 Calculate

$$\lim_{n\to\infty} \frac{e^{1/n} + e^{2/n} + \cdots + e^{(n-1)/n} + e^{n/n}}{n}$$

by expressing this limit as a definite integral of some continuous function and then using calculus methods.

8:2.13 Express

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f\left(\frac{k}{n}\right)$$

as a definite integral where f is continuous on $[0, 1]$.

8:2.14 Prove that the conclusion of Theorem 8.1 is false if f is not bounded.

8:2.15 Prove that the conclusion of Theorem 8.1 is true if f is continuous at all but a finite number of points in the interval $[a, b]$ and is bounded.

8:2.16 Prove that the conclusion of Theorem 8.1 is true for the function f defined on the interval $[0, 1]$ as follows: $f(0) = 0$ and $f(x) = 2^{-n}$ for each $2^{-n-1} < x \leq 2^{-n}$ $(n = 0, 1, 2, 3, \dots)$. How many points of discontinuity does f have in the interval $[0, 1]$? What is the value of the number I in this case?

8:2.17 For a bounded function f and any partition π

$$[x_0, x_1],\ [x_1, x_2],\ \dots,\ [x_{n-1}, x_n]$$

of the interval $[a, b]$ write

$$M(f, \pi) = \sum_{k=1}^{n} \max\{f(x) : x \in [x_{k-1}, x_k]\}(x_k - x_{k-1})$$

and

$$m(f, \pi) = \sum_{k=1}^{n} \min\{f(x) : x \in [x_{k-1}, x_k]\}(x_k - x_{k-1})$$

These are called the *upper sums* and *lower sums* for the partition for the function f.

(a) Show that if π_2 contains all of the points of the partition π_1 then

$$m(f, \pi_1) \leq m(f, \pi_2) \leq M(f, \pi_2) \leq M(f, \pi_1).$$

(b) Show that if π_1 and π_2 are arbitrary partitions and f is any bounded function then

$$m(f, \pi_1) \leq M(f, \pi_2).$$

(c) Show that if π is any arbitrary partition and f is any bounded function on $[a, b]$ then

$$c(b - a) \leq m(f, \pi) \leq M(f, \pi) \leq C(b - a)$$

where $C = \sup f$ and $c = \inf f$.

(d) Show that with any choice of associated points the Riemann sum over a partition π is in the interval $[m(f, \pi), M(f, \pi)]$.

(e) Show that, if f is continuous, every value in the interval between $m(f, \pi)$ and $M(f, \pi)$ is equal to some particular Riemann sum over the partition π with an appropriate choice of associated points ξ_k.

(f) Show that if f is not continuous the preceding assertion may be false.

8.3 Properties of the Integral

The integral has thus far been defined just for continuous functions. We ask what properties it must have. Later on we shall have to extend the scope of the integral to much broader classes of functions. It will be important to us then that the collection of elementary properties here will still be valid.

These properties exhibit the structure of the integral. They are the most vital tools to use in handling integrals both for theoretical and practical matters. Since we are restricted to continuous functions in this section, the proofs are very simple. As we enlarge the scope of the integral the proofs may become more difficult, and subtle differences in assertions may arise.

Note. All functions f and g appearing in the statements are assumed to be continuous on the intervals $[a, b]$, $[b, c]$, $[a, c]$ in the statements. Thus the integrals all have meaning. This means we do not have to prove that any of these integrals exist: they do. It is the stated identity that needs to be proved in each case. To prove that we can consider a sequence of partitions π_n chosen so that the points in the partition are closer together than $1/n$. Let us use the notation $S(\pi_n, f)$ to denote a Riemann sum taken over this partition for the function f with associated points chosen (say) at the left hand endpoint of the corresponding intervals. Then

$$\lim_{n\to\infty} S(\pi_n, f) = \int_a^b f(x)\,dx.$$

We shall use this idea in the proofs.

8.4 (Additive Property) *Let f be continuous on $[a, c]$. Then*

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx.$$

Proof. For our sequence of partitions we choose π_n to be a partition of $[a, c]$ chosen so that the points in the partition are closer together than $1/n$ and so that the point b is one of the points. Each partition π_n splits into two parts; π'_n and π''_n where the former is a partition of $[a, b]$ and where the latter is a partition of $[b, c]$. Note that

$$S(\pi_n, f) = S(\pi'_n, f) + S(\pi''_n, f)$$

by elementary arithmetic. If we let $n \to \infty$ in this identity we obtain immediately the identity in the statement we wish to prove. ■

8.5 (Linear Property) *Let f and g be continuous on $[a, b]$.*

$$\int_a^b [\alpha f(x) + \beta g(x)]\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx$$

for all $\alpha,\ \beta \in I\!R$.

Proof. Again consider a sequence of partitions of $[a,b]$, π_n chosen so that the points in the partition are closer together than $1/n$. If $S(\pi_n, f)$ denotes a Riemann sum taken over this partition for the function f then

$$\lim_{n\to\infty} S(\pi_n, f) = \int_a^b f(x)\,dx.$$

In the same way for g we would have

$$\lim_{n\to\infty} S(\pi_n, g) = \int_a^b g(x)\,dx.$$

But it is easy to check that

$$S(\pi_n, \alpha f + \beta g) = \alpha S(\pi_n, f) + \beta S(\pi_n, g)$$

and taking $n \to \infty$ in this identity gives exactly the statement in the property. Note that we do not have to prove that

$$S(\pi_n, \alpha f + \beta g) \to \int_a^b [\alpha f(x) + \beta g(x)]\,dx.$$

This we already know, by Theorem 8.1, because the integrand is continuous. ■

8.6 (Monotone Property) *Let f and g be continuous on $[a,b]$. Then*

$$\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx.$$

if $f(x) \le g(x)$ for all $a \le x \le b$.

Proof. Consider a sequence of partitions π_n chosen so that the points in the partition are closer together than $1/n$. If $S(\pi_n, f)$ denotes a Riemann sum taken over this partition for the function f then

$$\lim_{n\to\infty} S(\pi_n, f) = \int_a^b f(x)\,dx.$$

In the same way for g we would have

$$\lim_{n\to\infty} S(\pi_n, g) = \int_a^b g(x)\,dx.$$

But since $f(x) \le g(x)$ for all x we must have

$$S(\pi_n, f) \le S(\pi_n, g).$$

Taking limits as $n \to \infty$ in this inequality yields the property. ■

8.7 (Absolute Property) *Let f be continuous on $[a,b]$.*

$$-\int_a^b |f(x)|\,dx \le \int_a^b f(x)\,dx \le \int_a^b |f(x)|\,dx$$

or, equivalently,

$$\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx$$

Proof. This follows immediately from the monotone property because $-|f(x)| \le f(x) \le |f(x)|$. ■

Fundamental Theorem of the Calculus The next two properties are known together as the fundamental theorem of the calculus. They establish the close relationship between differentiation and integration and offer, to the calculus student, a useful method for the computation of integrals. This method reduces the computational problem of integration (i.e., computing a limit of Riemann sums) to the problem of finding an antiderivative.

8.8 (Differentiation of the Indefinite Integral) *Let f be continuous on $[a,b]$. The function*

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

has a derivative on $[a,b]$ and $F'(x) = f(x)$ at each point.

Proof. Let $h > 0$ and $x \in [a,b)$. We compute

$$F(x+h) - F(x) - hf(x) = \int_x^{x+h} (f(t) - f(x))\,dt$$

provided only that $x + h \le b$. Thus, using Exercise 8:3.1, we have

$$|F(x+h) - F(x) - hf(x)| \le h \max\{|f(t) - f(x)| : t \in [x, x+h]\}$$

and hence that

$$\left|\frac{F(x+h) - F(x)}{h} - f(x)\right| \le \max\{|f(t) - f(x)| : t \in [x, x+h]\}.$$

As f is continuous at x

$$\max\{|f(t) - f(x)| : t \in [x, x+h]\} \to 0$$

as $h \to 0+$ and this inequality shows that the right hand derivative of F at $x \in [a,b)$ is exactly $f(x)$.

A similar argument would show that the left hand derivative of F at $x \in (a,b]$ is exactly $f(x)$. This proves the property. ■

8.9 (Integral of a Derivative) *If the function F has a continuous derivative on $[a, b]$ then*

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

Proof. Given any $\varepsilon > 0$ there is a $\delta > 0$ so that any Riemann sum for the continuous function F' over a partition of $[a, b]$ into intervals of length less than δ is within ε of $\int_a^b F'(x)\,dx$. If

$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n]$$

is such a partition then observe that, if we choose the associated points $\xi_k \in [x_{k-1}, x_k]$ by the mean value theorem in such a way that

$$F(x_k) - F(x_{k-1}) = F'(\xi_k)(x_k - x_{k-1})$$

then we will have

$$F(b) - F(a) = \sum_{k=1}^n F(x_k) - F(x_{k-1}) = \sum_{k=1}^n F'(\xi_k)(x_k - x_{k-1}).$$

Since the right hand side of the identity is within ε of $\int_a^b F'(x)\,dx$ so too must be the value $F(b) - F(a)$. But this is true for any $\varepsilon > 0$ and hence it follows that these must be equal, i.e., that

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

■

Exercises

8:3.1 If f is continuous on an interval $[a, b]$ and

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

show that

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

8:3.2 (Mean Value Theorem for Integrals) If f is continuous show that there is a point ξ in (a, b) so that

$$\int_a^b f(x)\,dx = f(\xi)(b - a).$$

8:3.3 If f is continuous and $m \leq f(x) \leq M$ for all x in $[a, b]$ show that

$$m\int_a^b g(x)\,dx \leq \int_a^b f(x)g(x)\,dx \leq M\int_a^b g(x)\,dx$$

for any continuous, nonnegative function g.

8:3.4 If f is continuous and nonnegative on an interval $[a,b]$ and

$$\int_a^b f(x)\,dx = 0$$

show that f is identically equal to zero there.

8:3.5 If f and g are continuous on an interval $[a,b]$ show that there is a number $\xi \in (a,b)$ such that

$$\int_a^b f(x)g(x)\,dx = f(\xi)\int_a^b g(x)\,dx.$$

8:3.6 If f is continuous on an interval $[a,b]$ and

$$\int_a^b f(x)g(x)\,dx = 0$$

for every continuous function g on $[a,b]$ show that f is identically equal to zero there.

8:3.7 **(Integration by parts)** Suppose that f, g, f' and g' are continuous on $[a,b]$. Establish the *integration by parts* formula

$$\int_a^b f(x)g'(x)\,dx = [f(b)g(b) - f(a)g(a)] - \int_a^b f'(x)g(x)\,dx.$$

8:3.8 **(Integration by substitution)**

State conditions on f and g so that the *integration by substitution* formula

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(s)\,ds$$

is valid.

8:3.9 State conditions on f, g and h so that the *integration by substitution* formula

$$\int_a^b f(g(h(x)))g'(h(x))h'(x)\,dx = \int_{g(h(a))}^{g(h(b))} f(s)\,ds$$

is valid.

8:3.10 If f and g are continuous on an interval $[a,b]$ show that

$$\left(\int_a^b f(x)g(x)\,dx\right)^2 \le \left(\int_a^b [f(x)]^2\,dx\right)\left(\int_a^b [g(x)]^2\,dx\right)$$

This is called the Cauchy-Schwartz inequality and is the analog for integrals of that inequality in Exercise 3:5.12

8.4 Cauchy's Second Method

Defining an integral only for continuous functions, as we did in the preceding section, is far too limiting. Even in the early nineteenth century the need for considering more general functions was apparent. For Cauchy this meant handling functions that have discontinuities. But Cauchy would not have felt any need to handle badly discontinuous functions, indeed he may not even have considered such objects as functions. In our terminology we could say that Cauchy was interested in extending his integral from continuous functions to functions possessing isolated discontinuities (i.e., the set of discontinuity points contains only isolated points).

We have already noted in Section 8.2.1 that bounded functions with isolated discontinuities present no difficulties. Cauchy's first method can be applied to them. It is the case of unbounded functions that offers real resistance. What should we mean by the integral

$$\int_0^1 \frac{dx}{\sqrt{x}}?$$

While the integral has only one discontinuity (at $x = 0$) the function is unbounded and Cauchy's first method cannot be applied. If the integral did make sense then we would expect that the function

$$F(\delta) = \int_\delta^1 \frac{dx}{\sqrt{x}}$$

would be defined and continuous everywhere on the interval $[0, 1]$ and the value $F(0)$ would equal our integral. But here $F(x)$ is not defined at $x = 0$ although it is defined for all x in $(0, 1]$ since the integrand is continuous on any interval $[x, 1]$ for $x > 0$. If we compute it we see that

$$F(\delta) = \int_\delta^1 \frac{dx}{\sqrt{x}} = 2 - 2\sqrt{\delta}.$$

While we cannot take $F(0)$ itself (it is not defined), we can take the limit,

$$\lim_{\delta \to 0+} F(\delta) = \lim_{\delta \to 0+} \int_\delta^1 \frac{dx}{\sqrt{x}} = 2$$

as a perfectly reasonable value for the integral.

Indeed if we consider this as a problem in determining the area of the unbounded region in Figure 8.3 we should likely come up with the same value 2 as our answer.

This is precisely Cauchy's second method. If you understand this example, you understand the method. Any general write up of the

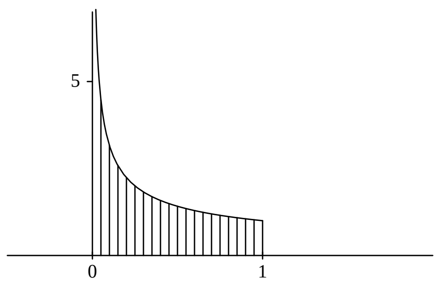

Figure 8.3: Integral $\int_0^1 x^{-1/2}\,dx$ considered as an area.

method might obscure this very simple idea. We need some language, however. The procedure of taking a limit to obtain the final value of the integral may or may not work. If the limit does exist we say that the integral *converges* or is a *convergent integral.* Otherwise the integral is said to be *divergent.* We say that the function f is *integrable by Cauchy's second method* or simply *integrable* if the context is clear.

We give a formal definition valid just in the case that the function has one point of unboundedness and that point occurs at the left hand endpoint of the interval. For more than one point or for a point not at an endpoint the definition is best generalized by splitting the integral into separate integrals each of which can be handled one at a time in this fashion. (See Exercise 8:4.3.)

Definition 8.10 Let f be a continuous function on an interval $(a, b]$ that is unbounded in every interval $(a, a+\delta)$. Then we define

$$\int_a^b f(x)\,dx$$

to be

$$\lim_{\delta \to 0+} \int_{a+\delta}^b f(x)\,dx$$

if this limit exists, and in this case the integral is said to be *convergent.* If both integrals

$$\int_a^b f(x)\,dx \quad \text{and} \quad \int_a^b |f(x)|\,dx$$

converge the integral is said to be *absolutely convergent.*

The role of the extra condition of absolute convergence is much like its role in the study of infinite series. You will recall that absolutely convergent series are more "robust" in the sense that they can be rather freely manipulated, unlike the nonabsolutely convergent series that are rather fragile. The same is true here of absolutely convergent integrals. Note that the integral $\int_0^1 x^{-1/2}\,dx$ considered above is convergent and absolutely convergent merely because the integrand is nonnegative.

Exercises

8:4.1 Formulate a definition of the integral $\int_a^b f(x)\,dx$ for a function continuous on $[a,b)$ and unbounded at the right hand endpoint. Supply an example.

8:4.2 Formulate a definition of the integral $\int_a^b f(x)\,dx$ for a function continuous on $[a,c)$ and on $(c,b]$ and unbounded in every interval containing c. Supply an example.

8:4.3 How would an integral of the form

$$\int_0^3 \frac{f(x)}{\sqrt{|x(x-1)(x-2)(x-3)|}}\,dx$$

be interpreted, where f is continuous?

8:4.4 Let f and g be continuous on $(a,b]$ and such that $|f(x)| \le |g(x)|$ for all $a < x \le b$. If the integral $\int_a^b g(x)\,dx$ is absolutely convergent, show that so also is the integral $\int_a^b f(x)\,dx$.

8:4.5 For what continuous functions f must the integral

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}}\,dx$$

converge?

8:4.6 Let f be a bounded function, continuous on $(a,b]$ and that is discontinuous at the endpoint a. Show that if the second method of Cauchy is applied to f then the result is the same as applying the first method to the entire interval $[a,b]$ (regardless of the value assigned to $f(a)$).

8:4.7 Suppose that f is continuous on $[-1,1]$ except for an isolated discontinuity at $x=0$. If the limit

$$\lim_{\delta\to 0+}\left(\int_{-1}^{-\delta} f(x)\,dx + \int_{\delta}^{1} f(x)\,dx\right)$$

exists does it follow that f is integrable on $[-1,1]$?

8:4.8 As a project determine which of the properties of the integral in Section 8.3 (which apply only to continuous functions on an interval $[a, b]$) can be extended to functions that are integrable by Cauchy's second method on $[a, b]$. Give proofs.

8.5 Cauchy's Second Method (continued)

The same idea that Cauchy used to assign meaning to the integral of unbounded functions he also used to handle functions on unbounded intervals. How should we interpret the integral

$$\int_1^\infty \frac{dx}{x^2}?$$

We might try first to form a partition of the unbounded interval $[1, \infty)$ and seek some kind of limit of Riemann sums. A much simpler idea is to adapt Cauchy's second method to this in the obvious way.

$$\int_1^\infty \frac{dx}{x^2} = \lim_{X\to\infty} \int_1^X \frac{dx}{x^2} = \lim_{X\to\infty} \left(1 - \frac{1}{X}\right) = 1.$$

This is precisely Cauchy's second method applied to unbounded intervals. Again, if you understand this example, you understand the method.

We give a formal definition valid just for an infinite interval of the form $[a, \infty)$. The case $(-\infty, b]$ is similar. The case $(-\infty, +\infty)$ is best split up into the sum of two integrals, from $(-\infty, a]$ and $[a, \infty)$ each of which can be handled in this fashion. (See Exercise 8:5.2.)

Definition 8.11 Let f be a continuous function on an interval $[a, \infty)$. Then we define

$$\int_a^\infty f(x)\,dx$$

to be

$$\lim_{X\to\infty} \int_a^X f(x)\,dx$$

if this limit exists, and in this case the integral is said to be *convergent.* If both integrals

$$\int_a^\infty f(x)\,dx \quad \text{and} \quad \int_a^\infty |f(x)|\,dx$$

converge the integral is said to be *absolutely convergent.*

Again, the role of the extra condition of absolute convergence is much like its role in the study of infinite series. Note that the integral $\int_1^\infty x^{-2}\,dx$ considered above is convergent and also absolutely convergent merely because the integrand is nonnegative.

Exercises

8:5.1 Formulate a definition of the integral $\int_{-\infty}^b f(x)\,dx$ for a function continuous on $(-\infty, b]$. Supply examples of convergent and divergent integrals of this type.

8:5.2 Formulate a definition of the integral $\int_{-\infty}^\infty f(x)\,dx$ for a function continuous on $(-\infty, \infty)$. Supply examples of convergent and divergent integrals of this type.

8:5.3 For what values of p is the integral $\int_1^\infty x^{-p}\,dx$ convergent?

8:5.4 Show that

$$\int_0^\infty x^n e^{-x}\,dx = n!.$$

8:5.5 Let f be a continuous function on $[1, \infty)$ such that $\lim_{x\to\infty} f(x) = \alpha$. Show that if the integral $\int_1^\infty f(x)\,dx$ converges then α must be 0.

8:5.6 Let f be a continuous function on $[1, \infty)$ such that the integral $\int_1^\infty f(x)\,dx$ converges. Can you conclude that $\lim_{x\to\infty} f(x) = 0$?

8:5.7 Let f be a continuous, decreasing function on $[1, \infty)$. Show that the integral $\int_1^\infty f(x)\,dx$ converges if and only if the series $\sum_{n=1}^\infty f(n)$ converges.

8:5.8 Give an example of a function f continuous on $[1, \infty)$ so that the integral $\int_1^\infty f(x)\,dx$ converges but the series $\sum_{n=1}^\infty f(n)$ diverges.

8:5.9 Give an example of a function f continuous on $[1, \infty)$ so that the integral $\int_1^\infty f(x)\,dx$ diverges but the series $\sum_{n=1}^\infty f(n)$ converges.

8:5.10 Show that

$$\int_0^\infty \frac{\sin x}{x}\,dx$$

is convergent but not absolutely convergent. (Note it may seem to require special handling at the left hand endpoint but it does not.)

8:5.11 **(Cauchy Criterion for Convergence)** Let $f : [a, \infty) \to \mathbb{R}$ be a continuous function. Show that the integral $\int_a^\infty f(x)\,dx$ converges if and only if for every $\varepsilon > 0$ there is a number M so that

$$\left|\int_c^d f(x)\,dx\right| < \varepsilon$$

for all $M < c < d$.

8:5.12 (Cauchy Criterion for Absolute Convergence) Let $f : [a, \infty) \to \mathbb{R}$ be a continuous function. Show that the integral $\int_a^\infty f(x)\,dx$ converges absolutely if and only if for every $\varepsilon > 0$ there is a number M so that

$$\int_c^d |f(x)|\,dx < \varepsilon$$

for all $M < c < d$.

8:5.13 As a project determine which of the properties of the integral in Section 8.3 (which apply only to continuous functions on a finite interval) can be extended to integrals on an infinite interval $[a, \infty]$. Give proofs.

8.6 The Riemann Integral

Thus far in our discussion of the integral we have defined the meaning of the symbol

$$\int_a^b f(x)\,dx$$

first for all continuous functions, by Cauchy's first method, and then for functions that may have a finite number of discontinuities at which the function is unbounded, by Cauchy's second method.

Let us return to Cauchy's first method, but this time with rather more ambition. We ask just how far this method can be applied. It can be applied to all continuous functions; that was the content of Theorem 8.1. It can be applied to all bounded functions with finitely many discontinuities (Exercise 8:2.15). It can be applied to *some* bounded functions with infinitely many discontinuities (Exercise 8:2.16).

Rather than search for broader classes of functions to which this method applies we adopt the viewpoint that was taken by Riemann. We simply *define* the class of all functions to which Cauchy's first method can be applied and then seek to characterize that class. This represents a much more modern point of view than Cauchy would have taken with his much more limited idea of what a function is. Note that we need only turn Theorem 8.1 into a definition.

Definition 8.12 Let f be a function on an interval $[a, b]$. Suppose that there is a number I such that for all $\varepsilon > 0$ there is a $\delta > 0$ so that

$$\left| \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) - I \right| < \varepsilon$$

whenever $[x_0, x_1]$, $[x_1, x_2]$, ... , $[x_{n-1}, x_n]$, is a partition of the interval $[a, b]$ into subintervals of length less than δ and each ξ_k is a point in the interval $[x_{k-1}, x_k]$. Then f is said to be *Riemann integrable* on $[a, b]$ and we write

$$\int_a^b f(x)\,dx$$

for that number I.

We can call the set of points

$$\pi = \{x_0, x_1, x_2, \ldots, x_{n-1}, x_n\}$$

the *partition* of the interval $[a, b]$ or equivalently, if it is more convenient, the set of intervals

$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n]$$

can be called the partition. The points ξ_k that are chosen from each interval $[x_{k-1}, x_k]$ are called the *associated points* of the partition. Notice that in the definition the associated points can be freely chosen inside the intervals of the partition.

Loosely a function f is Riemann integrable if the limit of the Riemann sums for f exists over that interval. The program now is to determine what classes of functions are Riemann integrable and to obtain characterizations of Riemann integrability. This we shall investigate in the remainder of this section.

We need also to find out whether the properties of the integral that hold for continuous functions now continue to hold for all Riemann integrable functions. We shall consider that in the next section.

Two observations are immediate from our earlier work and also very important:

All continuous functions are Riemann integrable.

All Riemann integrable functions are bounded.

In light of this last statement we see that the Riemann integral is somewhat limited in that it will not do anything to handle unbounded functions. For that we must still return to Cauchy's second method. But, as we shall see, the Riemann integral will handle many bounded functions that are quite badly discontinuous (but not too badly). As research progressed in the nineteenth century the Riemann integral became the standard tool for discussing integrals of bounded functions. For unbounded functions Cauchy's second method continued to be employed although other methods emerged.

By the early twentieth century the Riemann integral was abandoned by all serious analysts in favor of Lebesgue's integral. The Riemann integral survives in texts such as this mainly because of the technical difficulties of Lebesgue's better, but more difficult, methods.

8.6.1 Some Examples

All Riemann integrable functions are bounded. All continuous functions are Riemann integrable. In order to obtain some insight into the question as to what functions are Riemann integrable we present some examples, first of a bounded function which is not and then, second, of a quite badly discontinuous function which is integrable.

Example 8.13 Here is an example of a function that is bounded but "too discontinuous" to be Riemann integrable. On the interval $[0,1]$ let f be the function equal to 1 for x rational and to 0 for x irrational. Let

$$[x_0,x_1],\ [x_1,x_2],\ \dots,\ [x_{n-1},x_n]$$

be any partition. If we choose associated points $\xi_k \in [x_{k-1},x_k]$ so that ξ_k is rational then the Riemann sum

$$\sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) = \sum_{k=1}^{n}(x_k - x_{k-1}) = 1 \tag{2}$$

while if we choose associated points $\eta_k \in [x_{k-1},x_k]$ so that η_k is irrational

$$\sum_{k=1}^{n} f(\eta_k)(x_k - x_{k-1}) = 0. \tag{3}$$

Because of (2) and (3) the integral $\int_0^1 f(x)\,dx$ cannot exist. ◀

Example 8.14 Recall the Dirichlet function (Section 5.2.6) that provides an example of a function that is discontinuous at every rational number and continuous at every irrational. We show that this function is Riemann integrable. On the interval $[0,1]$ let f be the function equal to $1/q$ for $x = p/q$ rational (assuming that p/q has been expressed in its lowest terms) and to 0 for x irrational.

Let $\varepsilon > 0$. Let q_0 be any positive integer larger than $2/\varepsilon$. We count the number of points x in $[0,1]$ at which $f(x) > 1/q_0$. There are finitely many of these, say M of them. Choose δ_1 sufficiently small so that any two of these points are further apart than $2\delta_1$. Choose $\delta < \delta_1$ so that (for reasons which become clear only after all our computations are done) $M\delta < \varepsilon/2$. This will allow us to use the inequality

$$M\delta + 1/q_0 < \varepsilon. \tag{4}$$

Let

$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n]$$

be any partition chosen so that each of the intervals is shorter than δ.

For any choice of associated points $\xi_k \in [x_{k-1}, x_k]$ we note that either (i) $f(\xi_k) = 0$ if ξ_k is irrational, or (ii) $f(\xi_k) > 1/q_0$ if ξ_k is one of the M points counted above, or (iii) $0 < f(\xi_k) \leq 1/q_0$ is any other rational point. We can estimate the Riemann sum

$$\sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$$

by considering separately these three cases. Case (i) evidently contributes nothing to this sum. Case (ii) can contribute at most $M\delta$ to this sum since each interval in the partition can contain at most one of the points of type (ii) and there are only M such points. Finally case (iii) can contribute in total no more than $1/q_0$. Thus, using the inequality (4), we have

$$0 \leq \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) \leq M\delta + 1/q_0 < \varepsilon.$$

This proves that the integral $\int_0^1 f(x)\,dx = 0$. Considering just how discontinuous this function is (it has a dense set of discontinuities) it is quite startling that it is nonetheless integrable. ◀

✂ 8.6.2 Riemann's Criteria

What bounded functions then are Riemann integrable? The answer is that such functions must be "mostly" continuous. The example of the very discontinuous function in Example 8.13 suggests this. On the other hand Example 8.14 shows that the discontinuities of a Riemann integrable function might even be dense. Riemann first analyzed this by using the oscillation of the function f on an interval. We recall (Definition 6.24) that this is defined as

$$\omega f([c, d]) = \sup_{x\in[c,d]} f(x) - \inf_{x\in[c,d]} f(x).$$

This measures how much the function f changes in the interval $[c, d]$. For a continuous function this is just the difference between the maximum and minimum values of f on $[c, d]$ and will be small if the interval $[c, d]$ is small enough.

Theorem 8.15 (Riemann) *A function f defined on an interval $[a, b]$ is Riemann integrable if and only if for every $\varepsilon > 0$ there is a $\delta > 0$ so that*

$$\sum_{k=1}^{n} \omega f([x_{k-1}, x_k])(x_k - x_{k-1}) < \varepsilon$$

whenever $[x_0, x_1]$, $[x_1, x_2]$, ... , $[x_{n-1}, x_n]$, is a partition of the interval $[a, b]$ into subintervals of length less than δ.

Proof. If f is Riemann integrable on $[a, b]$ with integral I then then for any $\varepsilon > 0$ there must be a $\delta > 0$ so that any two Riemann sums taken over a partition with intervals smaller than δ are both within $\varepsilon/4$ of I. In particular then we have

$$\left| \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) - \sum_{k=1}^{n} f(\eta_k)(x_k - x_{k-1}) \right| < \varepsilon/2$$

whenever $[x_0, x_1]$, $[x_1, x_2]$, ... , $[x_{n-1}, x_n]$, is a partition of the interval $[a, b]$ into subintervals of length less than δ. Here ξ_k and η_k are any choices from $[x_{k-1}, x_k]$. We rewrite this as

$$\left| \sum_{k=1}^{n} [f(\xi_k) - f(\eta_k)](x_k - x_{k-1}) \right| \leq \varepsilon/2 < \varepsilon. \tag{5}$$

Now notice that

$$\sup_{\eta, \xi \in [x_{k-1}, x_k]} (f(\xi) - f(\eta)) = \omega f([x_{k-1}, x_k]).$$

Thus we see that the criterion follows immediately on taking sups over these choices of ξ_k and η_k in the inequality 5.

The other direction of the theorem can be interpreted as a "Cauchy criterion" and proved in a manner similar to all our other Cauchy criteria so far in the text (indeed very like the proof of Theorem 8.1). We omit the details. ■

Theorem 8.15 offers an interesting necessary and sufficient condition for integrability. It is rather awkward to use the sufficiency criterion here since it demands we check that *all* small partitions have a certain property. The following variant is a little easier to apply since we need find only *one* partition for each positive ε.

Theorem 8.16 *A function f on an interval $[a, b]$ is Riemann integrable if and only if for every $\varepsilon > 0$ there is at least one partition $[x_0, x_1]$, $[x_1, x_2]$, ..., $[x_{n-1}, x_n]$, of the interval $[a, b]$ so that*

$$\sum_{k=1}^{n} \omega f([x_{k-1}, x_k])(x_k - x_{k-1}) < \varepsilon.$$

Proof. By Theorem 8.15 we see that if f is Riemann integrable there would have to exist such a partition.

In the opposite direction we must show that the condition here implies integrability. Certainly this condition implies that f is bounded (or else this sum would be infinite) and so we may assume that $|f(x)| \leq M$ for all x. This gives us a useful, if crude, estimate on the size of the oscillation on any interval $[c, d]$:

$$\omega f([c, d]) \leq 2M.$$

Let $\varepsilon > 0$. We shall find a number δ so that the criterion of Theorem 8.15 is satisfied. Let $[x_0, x_1]$, $[x_1, x_2]$, ..., $[x_{n-1}, x_n]$ be the partition whose existence is given. We use that to find our δ. Choose δ sufficiently small so that

$$2Mn\delta < \varepsilon.$$

Now let

$$[y_0, y_1],\ [y_1, y_2],\ \dots,\ [y_{m-1}, y_m],$$

be any partition of the interval $[a, b]$ into subintervals of length less than δ. These intervals are of two types: type (i) are those that are contained entirely inside intervals of our original partition, and type (ii) are those that include as interior points one of the points x_k for $k = 1, 2, \dots, n-1$. In any case there are only $n-1$ of these intervals and each is of length less than δ. Thus, using just a crude estimate on each of these terms, the intervals of type (ii) contribute to the sum

$$\sum_{k=1}^{m} \omega f([y_{k-1}, y_k])(y_k - y_{k-1})$$

no more than $(2M)n\delta$. The sum taken over all the type (i) intervals must be smaller than

$$\sum_{k=1}^{m} \omega f([x_{k-1}, x_k])(x_k - x_{k-1}) < \varepsilon.$$

Thus the total sum

$$\sum_{k=1}^{m} \omega f([y_{k-1}, y_k])(y_k - y_{k-1}) < 2Mn\delta + \varepsilon < 2\varepsilon.$$

It follows by the criterion in Theorem 8.15 that f is Riemann integrable as required. ■

8.6.3 Lebesgue's Criterion

Theorem 8.15 is very beautiful and seemingly characterizes the class of Riemann integrable functions in a meaningful way. But at the time of Riemann there was only an imperfect understanding of sets of real numbers and so it did not occur to Riemann that the property of Riemann integrability for a bounded function f depended exclusively on the nature of the set of points of discontinuity of f. Indeed the condition

$$\sum_{k=1}^{n} \omega f([x_{k-1}, x_k])(x_k - x_{k-1}) < \varepsilon$$

on the oscillation of the function suggests that something more subtle than just this is happening.

In 1901 Henri Lebesgue completed this theorem by using the notion of a set of measure zero. Recall (from Section 6.8) that a set E of real numbers is of measure zero if for every $\varepsilon > 0$ there is a sequence of intervals $\{(c_i, d_i)\}$ covering all points of E and with total length $\sum_{i=1}^{\infty}(d_i - c_i) < \varepsilon$. The exact characterization of Riemann integrable functions is precisely this: they are bounded (as we already well know) and they are continuous at all points except perhaps at the points of a set of measure zero. (In modern language they are said to be continuous *almost everywhere.*)

Theorem 8.17 (Riemann–Lebesgue) *A function f on an interval $[a, b]$ is Riemann integrable if and only if f is bounded and the set of points in $[a, b]$ at which f is not continuous is a set of measure zero.*

Proof. The necessity is not difficult to prove, but is the least important part for us. The sufficiency is more important and harder to prove. Throughout the proof we require a familiarity with the notion of the oscillation $\omega_f(x)$ of a function f at a point x as discussed in Section 6.7. Recall that this value is positive if and only if f is discontinuous at x.

Let us suppose that f is Riemann integrable. Certainly f is bounded. Fix $e > 0$ and consider the set $N(e)$ of points x such that the oscillation of f at x is greater than e, i.e., so that

$$\omega_f(x) > e.$$

Any interval (c, d) that contains a point $x \in N(e)$ will certainly have
$$\omega f([c, d]) \geq e.$$
Let $\varepsilon > 0$ and use Theorem 8.15 to find intervals
$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n],$$
forming a partition of the interval $[a, b]$ and such that
$$\sum_{k=1}^{n} \omega f([x_{k-1}, x_k](x_k - x_{k-1}) < \varepsilon e/2.$$
Select from this collection just those intervals that contain a point from $N(e)$ in their interior. The total length of these intervals cannot exceed $(e\varepsilon)/(2e)$ since for each such interval $[x_{k-1}, x_k]$ we must have $\omega f([x_{k-1}, x_k]) \geq e$.

Thus we have succeeded in covering the set $N(e)$ by a sequence of open intervals (x_{k-1}, x_k) of total length less than $\varepsilon/2$, except for an oversight. One or more of the points $\{x_i\}$ might be in the set $N(e)$ and we have neglected to cover it. Since there are only finitely many such points we can add a few sufficiently short intervals to our collection.

Thus we have proved that for any $\varepsilon > 0$ the set $N(e)$ can be covered by a collection of open intervals of total length less than ε. It follows that $N(e)$ has measure zero. But the set of points of discontinuity of f is the union of the sets $N(1)$, $N(1/2)$, $N(1/4)$, $N(1/8)$, $\ldots$. Since each of these is measure zero it follows from Theorem 6.34 that the set of points of discontinuity of f has measure zero too as required.

This proves the theorem in one direction. In the other suppose that f is bounded, say that $|f(x)| \leq M$ for all x and that the set E of points in $[a, b]$ at which f is not continuous is a set of measure zero. Let $\varepsilon > 0$. By Theorem 8.16 we need to find at least one partition
$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n],$$
of the interval $[a, b]$ so that
$$\sum_{k=1}^{n} \omega f([x_{k-1}, x_k])(x_k - x_{k-1}) < \varepsilon.$$

Let E_1 denote the set of points x in $[a, b]$ at which the oscillation is greater than or equal to $\varepsilon/(2(b-a))$, i.e., for which
$$\omega_f(x) \geq \varepsilon/(2(b-a)).$$
This set is closed (see Theorem 6.27) and, being a subset of E, it must have measure zero. Now closed sets of measure zero can be

covered by a *finite* number of small open intervals of total length smaller than

$$\varepsilon/(4M+1).$$

(See Theorem 6.35.) We can assume that these open intervals do not have endpoints in common. Note that, at points in the intervals that remain, the oscillation of f is smaller than $\varepsilon/(2(b-a))$. Consequently these intervals may be subdivided into smaller intervals on which the oscillation is at least that small (Exercise 8:6.6).

Thus we may construct a partition

$$[x_0, x_1],\ [x_1, x_2],\ \dots,\ [x_{n-1}, x_n],$$

of the interval $[a, b]$ consisting of two kinds of closed intervals: (i) the first kind cover all the points of E_1 and have total length smaller than $\varepsilon/(4M+1)$ and (ii) the remaining kind contain no points of E_1 and the oscillation of f on each of these intervals is smaller than $\varepsilon/(2(b-a))$, i.e.,

$$\omega f([x_{k-1}, x_k]) < \varepsilon/(2(b-a)).$$

The sum

$$\sum_{k=1}^{n} \omega f([x_{k-1}, x_k])(x_k - x_{k-1})$$

splits into two sums depending on the intervals of type (i) or type (ii). The former sum contributes no more than

$$(2M) \times \varepsilon/(4M+1) < \varepsilon/2$$

while the latter sum contributes no more than

$$\varepsilon/(2(b-a)) \times (b-a) < \varepsilon/2.$$

Altogether then

$$\sum_{k=1}^{n} \omega f([x_{k-1}, x_k])(x_k - x_{k-1}) < \varepsilon$$

and the proof is complete. ■

8.6.4 What functions are Riemann integrable? ✂

Theorem 8.17 exactly characterizes those functions that are Riemann integrable as the class of bounded functions that do not have too many points of discontinuity. We should recognize immediately that certain types of functions that we are used to working with are also integrable. We express these as corollaries to our theorem. (Recall that step functions were defined in Section 5.2.6.)

Corollary 8.18 *Every step function on an interval is Riemann integrable there.*

Proof. A step function is bounded and has only finitely many discontinuities. Thus the set of discontinuities has measure zero. Consequently the corollary follows from Theorem 8.17. ■

Corollary 8.19 *Every bounded function with only countably many points of discontinuity in an interval is Riemann integrable there.*

Proof. The corollary follows directly from Theorem 8.17 since countable sets have measure zero. ■

Corollary 8.20 *Every function monotonic on an interval is Riemann integrable there.*

Proof. A monotonic function is bounded and has only countably many discontinuities. Consequently this corollary follows from the preceding corollary. ■

Corollary 8.21 *If a function f is Riemann integrable on an interval $[a, b]$ then so too is the function $|f|$ on that interval.*

Proof. The corollary follows directly from Theorem 8.17 since if f is Riemann integrable on $[a, b]$ it must be bounded and continuous at every point except a set of measure zero. Exercise 8:6.7 shows that $|f|$ has precisely the same properties. ■

Exercises

8:6.1 Show directly from Theorem 8.16 that the characteristic function of the rationals is not Riemann integrable on any interval.

8:6.2 Show that the product of two Riemann integrable functions is itself Riemann integrable.

8:6.3 If f is Riemann integrable on an interval and f is never zero does it follow that $1/f$ is Riemann integrable there? What extra hypothesis could we invoke to make this so?

8:6.4 If f is Riemann integrable on an interval $[a, b]$ show that for every $\varepsilon > 0$ there are a pair of step functions $L(x) \le f(x) \le U(x)$ so that $\int_a^b (U(x) - L(x))\,dx < \varepsilon$.

8:6.5 Let f be a function on an interval $[a, b]$ with the property that for every $\varepsilon > 0$ there are a pair of step functions $L(x) \le f(x) \le U(x)$ so that $\int_a^b (U(x) - L(x))\,dx < \varepsilon$. Show that f is Riemann integrable.

8:6.6 Suppose that the oscillation $\omega_f(x)$ of a function f is smaller than η at each point x of an interval $[c,d]$. Show that there must be a partition $[x_0,x_1]$, $[x_1,x_2]$, ... , $[x_{n-1},x_n]$, of $[c,d]$ so that the oscillation $\omega f([x_{k-1},x_k]) < \eta$ on each member of the partition.

8:6.7 Show that the set of points at which a function F is discontinuous includes all points at which $|F|$ is discontinuous but not conversely. Deduce Corollary 8.21 as a result of this observation from Theorem 8.17.

8:6.8 Deduce Corollary 8.18 directly from Theorem 8.15 rather than from Theorem 8.17.

8:6.9 Deduce Corollary 8.19 directly from Theorem 8.15 rather than from Theorem 8.17.

8:6.10 Deduce Corollary 8.20 directly from Theorem 8.15 rather than from Theorem 8.17.

8:6.11 Show that the converse of Corollary 8.21 does not hold.

8:6.12 This Exercise develops the theory of the *Darboux integral* which is equivalent to Riemann's integral but defined using inf's and sup's of "Darboux sums" rather than limits of Riemann sums. In preparation Exercise 8:2.17 should be consulted. We use the notation $m(f,\pi)$ and $M(f,\pi)$ to denote the upper and lower sums over a partition π for an arbitrary bounded function f. Define the upper and lower integrals as

$$\overline{\int_a^b} f(x)\,dx = \inf M(f,\pi)$$

and

$$\underline{\int_a^b} f(x)\,dx = \inf m(f,\pi)$$

where the inf and sup are taken over all possible partitions π of the interval $[a,b]$. We say f is *Darboux integrable* if the upper and lower integrals are equal.

(a) Show that

$$\underline{\int_a^b} f(x)\,dx \le \overline{\int_a^b} f(x)\,dx.$$

(b) Show that every Riemann integrable function is Darboux integrable.

(c) Show that every Darboux integrable function is Riemann integrable.

(d) Show that if f is Riemann integrable then

$$\overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx.$$

(e) Show that
$$\overline{\int_a^b} (f(x)+g(x))\,dx \le \overline{\int_a^b} f(x)\,dx + \overline{\int_a^b} g(x)\,dx$$
with strict inequality possible.

8.7 Properties of the Riemann Integral

The Riemann integral[1] is an extension of Cauchy's first integral from continuous functions to a larger class of bounded functions—those that are bounded and continuous except at the points of a very small set (a set of measure zero). We have enlarged the class of functions to which the notion of an integral may be applied. Have we lost any of our crucial properties of Section 8.3?

These properties express how we expect integration to behave; it would be distressing to lose any of them. In some cases they remain completely unchanged. In some cases they need to be modified slightly. But our goal was never simply to integrate as many functions as possible; it is to preserve the theory of the integral and to apply that theory sufficiently broadly to handle all necessary applications. If we lose our basic properties we have lost too much. Fortunately the Riemann integral keeps all of the basic properties of the integral of continuous functions. The few differences should be carefully noted. Note especially how some of the properties must be rephrased.

8.22 (Additive Property) *If f is Riemann integrable on both intervals $[a,b]$ and $[b,c]$ then it is Riemann integrable on $[a,c]$ and*
$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx.$$

Proof. The proof of the identity need not change from the way we handled it for continuous functions (check this). It is the first assertion in the statement that must be verified. We prove that f is Riemann integrable on $[a,c]$.

By Theorem 8.17 if f is Riemann integrable on both of these intervals it is bounded on both and the set of points of discontinuity in each interval has measure zero. It follows that f is bounded on $[a,c]$. Also its set of points of discontinuity in $[a,c]$ is the union

[1]The proofs in this section make use of the Lebesgue criterion for integrability. The reader may skip the proofs and just see how the properties are essentially unchanged from Section 8.3 for Cauchy's original integral.

of the set of points of discontinuity in $[a, b]$ and $[b, c]$ together with (possibly) the point b itself. Thus the set of points of discontinuity in $[a, c]$ is also of measure zero. Consequently, by Theorem 8.17, f is Riemann integrable. ■

8.23 (Linear Property) *If f and g are both Riemann integrable on $[a, b]$ then so too is any linear combination $\alpha f + \beta g$ and*

$$\int_a^b [\alpha f(x) + \beta g(x)]\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx.$$

Proof. Again the proof of the identity does not change from the way we handled it for continuous functions (check this). It is the first assertion in the statement that needs to be verified. We must prove that $\alpha f + \beta g$ is Riemann integrable on on $[a, b]$.

The points of discontinuity of the function function $\alpha f + \beta g$ are either points of discontinuity of f or else they are points of discontinuity of g. If both functions f and g are Riemann integrable then then they are both bounded and continuous except at the points of a set of measure zero. It follows that $\alpha f + \beta g$ is bounded and continuous except at the points of a set of measure zero. Hence, by Theorem 8.17, $\alpha f + \beta g$ is Riemann integrable. ■

8.24 (Monotone Property) *If f and g are both Riemann integrable on $[a, b]$ then*

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

if $f(x) \leq g(x)$ for all $a \leq x \leq b$.

Proof. The proof for continuous functions works equally well here. ■

8.25 (Absolute Property) *If f is Riemann integrable on $[a, b]$ then so too is $|f|$ and*

$$-\int_a^b |f(x)|\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b |f(x)|\,dx$$

or, equivalently,

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx$$

Proof. The proof for continuous functions works equally well here because we have already shown, in Corollary 8.21 that if f is Riemann integrable on $[a, b]$ then so too is $|f|$. ■

Fundamental Theorem of Calculus The next two properties, 8.26 and 8.27, are very important. They show how the processes of integration and differentiation are inverses of each other. Together they are known as the fundamental theorem of the calculus for the Riemann integral. The reader should note, however, a weakness in this theory. If we compute F' we cannot immediately conclude from 8.27 that $\int_a^b F'(x)\,dx = F(b) - F(a)$. We need first to check that F' is Riemann integrable. This may not always be easy. Worse yet, it may be false, even for bounded derivatives: see the discussion in Section 9.7. It was this failure of the Riemann integral to integrate *all* derivatives that Lebesgue claimed was his motivation to look for a more general theory of integration.

8.26 (Differentiation of the Indefinite Integral) *If f is Riemann integrable on $[a, b]$ then the function*

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

is continuous on $[a, b]$ and $F'(x) = f(x)$ at each point x at which the function f is continuous.

Proof. Once again the proof for continuous functions works equally well here. Note, however, that we are no longer trying to prove that $F'(x) = f(x)$ at every point x, only at those points x where f is continuous.

It is left to the reader to check that proof and verify that it works here, unchanged. ■

8.27 (Integral of a Derivative) *Suppose that the function F is differentiable on $[a, b]$. Provided it is also true that F' is Riemann integrable on $[a, b]$, then*

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

Proof. Yet again the proof for continuous functions works equally well here. ■

Exercises

8:7.1 Give a set of conditions under which the integration by substitution formula

$$\int_a^b f(\phi(t))\phi'(t)\,dt = \int_{\phi(a)}^{\phi(b)} f(x)\,dx$$

holds.

8:7.2 Give a set of conditions under which the integration by parts formula

$$\int_a^b f(t)g'(t)\,dt = f(b)g(b) - f(a)(g(a) - \int_a^b f'(t)g(t)\,dt$$

holds.

8:7.3 Suppose that f is Riemann integrable on $[a, b]$ and define the function

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

(a) Show that F satisfies a Lipschitz condition on $[a, b]$, i.e., that there exists $M > 0$ such that for every $x, y \in [a, b]$,

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

(b) If x is a point at which f is not continuous is it still possible that $F'(x) = f(x)$?

(c) Is it possible that $F'(x)$ exists but is not equal to $f(x)$?

(d) Is it possible that $F'(x)$ fails to exist?

8:7.4 The function

$$F(x) = \int_0^x \sin(1/t)\,dt$$

has a derivative at every point where the integrand is continuous. Does it also have a derivative at $x = 0$?

8:7.5 Improve Property 8.27 by assuming that F is continuous on $[a, b]$ and allowing that F' exists at all points of $[a, b]$ with finitely many exceptions.

8:7.6 Do much better than the preceding exercise and improve Property 8.27 by assuming that F is continuous on $[a, b]$ and allowing that F' exists at all points of $[a, b]$ with countably many exceptions. ✂

8:7.7 **(More on the Fundamental Theorem of Calculus.)** Let f be bounded on $[a, b]$ and continuous a.e. on $[a, b]$. Suppose that F is defined on $[a, b]$ and that $F' = f$ a.e.. (Recall that "a.e." means everywhere except at the points of some set of measure zero.) ✂

(a) Is it necessarily true that $F(x) - F(a) = \int_a^x f(t)\,dt$ for every $x \in [a, b]$?

(b) Same question as in (a) but assume also that F is continuous.

(c) Same question, but this time assume that F is a Lipschitz function. You may assume the non-elementary fact that a Lipschitz function H with $H' = 0$ a.e. must be constant.

(d) Give an example of a Lipschitz function F such that F is differentiable, F' is bounded, but F' is not integrable.

8:7.8 If f and g are Riemann integrable on an interval $[a, b]$ show that

$$\left(\int_a^b f(x)g(x)\,dx\right)^2 \leq \left(\int_a^b [f(x)]^2\,dx\right)\left(\int_a^b [g(x)]^2\,dx\right).$$

This extends the Cauchy-Schwartz inequality of Exercise 8:3.10.

8:7.9 Show that the integration by parts formula of Exercise 8:3.7 extends to the case where f and g are continuous and f' and g' are Riemann integrable.

8.8 The Improper Riemann Integral

The Riemann integral applies only to bounded functions. What should we mean by the integral

$$\int_0^1 \frac{dx}{\sqrt{x}}?$$

Since the integrand is unbounded on $[0, 1]$ it is not Riemann integrable even though the integrand is continuous at all but one point. There is not much else for us to do but to back track by several decades and return to Cauchy's second method, namely we compute

$$\lim_{\delta\to 0+} F(\delta) = \lim_{\delta\to 0+} \int_\delta^1 \frac{dx}{\sqrt{x}} = 2.$$

What we should probably do now is to create a new hybrid integral by combining the Riemann integral with Cauchy's second method. This is often called the improper Riemann integral. As before we give a definition that considers only one point of unboundedness (at the left endpoint of the interval) with the understanding that the ideas can be applied to any finite number of such points.

Definition 8.28 Let f be a function on an interval $(a, b]$ that is Riemann integrable on $[a+\delta, b]$ and that is unbounded in the interval $(a, a + \delta)$ for every $0 < \delta < b - a$. Then we define

$$\int_a^b f(x)\,dx$$

to be

$$\lim_{\delta\to 0+} \int_{a+\delta}^b f(x)\,dx$$

if this limit exists, and in this case the integral is said to be *convergent.* If both integrals

$$\int_a^b f(x)\,dx \quad \text{and} \quad \int_a^b |f(x)|\,dx$$

converge the integral is said to be *absolutely convergent.*

In the same way we also extend the Riemann integral from bounded intervals to unbounded ones. How should we interpret the integral

$$\int_1^\infty \frac{dx}{x^2}?$$

This cannot exist as a Riemann integral since the definition is very clearly restricted to finite intervals and would not allow any easy interpretation for infinite intervals. As before we use Cauchy's second method to obtain

$$\int_1^\infty \frac{dx}{x^2} = \lim_{X\to\infty}\int_1^X \frac{dx}{x^2} = \lim_{X\to\infty} 1 - \frac{1}{X} = 1.$$

We give a formal definition valid just for an infinite interval of the form $[a,\infty)$. The case $(-\infty,b]$ is similar. The case $(-\infty,+\infty)$ is best split up into the sum of two integrals, from $(-\infty,a]$ and $[a,\infty)$ each of which can be handled in this fashion.

Definition 8.29 Let f be a function on an interval $[a,\infty)$ that is Riemann integrable on every interval $[a,b]$ for $a < b < \infty$. Then we define

$$\int_a^\infty f(x)\,dx$$

to be

$$\lim_{X\to\infty}\int_a^X f(x)\,dx$$

if this limit exists, and in this case the integral is said to be *convergent.* If both integrals

$$\int_a^\infty f(x)\,dx \quad\text{and}\quad \int_a^\infty |f(x)|\,dx$$

converge the integral is said to be *absolutely convergent.*

Both of these definitions extend the Riemann integral to a more general concept. Note that in any applications using an improper Riemann integral of either type, we are obliged to announce whether the integral is convergent or divergent, and frequently whether it is absolutely or nonabsolutely convergent.

It might seem that this theory would be very important to master and represents the final word on the subject of integration. By the end of the nineteenth century it had become increasingly clear that this theory of the Riemann integral itself was completely inadequate to handle the bounded functions that were arising in many applications. The extra step here, using Cauchy's second method, designed to handle unbounded functions also proved far too restrictive. The modern theory of integration was developed in the first decades of the twentieth century. The methods are very much different and even the language needs many changes.

Thus, the material in these last few sections has largely an historical interest. Some mathematicians claim it has only that, others that learning this material is a good preparation for learning the more advanced material.

Exercises

8:8.1 For what values of p, q are the integrals

$$\int_0^1 \frac{\sin x}{x^p}\,dx \text{ and } \int_0^1 \frac{(\sin x)^q}{x}\,dx$$

ordinary Riemann integrals, convergent improper Riemann integrals, or divergent improper Riemann integrals?

8.9 More on the Fundamental Theorem of the Calculus

The Riemann integral does not integrate all bounded derivatives and so the fundamental theorem of the calculus for this integral assumes the awkward form

$$\int_a^b F'(x)\,dx = F(b) - F(a)$$

provided F is differentiable on $[a,b]$ and *the derivative F' is Riemann integrable there.*

The emphasized phrase is unfortunate. It means we have a limited theory and it also means that, in practice, we must always check to be sure that a derivative F' is integrable before proceeding to integrate it. In Section 9.7 we shall show how to construct a function F that is everywhere differentiable on an interval and whose derivative F' is bounded but not itself Riemann integrable on any subinterval.

Let us take another look at the integrability of derivatives so see if we can discover what goes wrong. We take a completely naive approach and start with the definition of the derivative itself. If $F' = f$ everywhere, then, at each point ξ and for every $\varepsilon > 0$, there is a $\delta > 0$ so that

$$|F(x'') - F(x') - f(\xi)(x'' - x')| < \varepsilon(x'' - x') \tag{6}$$

for $x' \leq \xi \leq x''$ and $0 < x'' - x' < \delta$.

We shall attempt to recover $F(b) - F(a)$ as a limit of Riemann sums for f, even though this is a misguided attempt, since we know that the Riemann integral must fail in general to accomplish this. Even so, let us see where the attempt takes us.

Let

$$a = x_0 < x_1 < x_2 \ldots x_n = b$$

be a partition of $[a, b]$, and let $\xi_i \in [x_{i-1}, x_i]$. Then

$$F(b) - F(a) = \sum_{i=1}^{n} (F(x_{i-1}) - F(x_i)) = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) + R$$

where

$$R = \sum_{i=1}^{n} (F(x_i) - F(x_{i-1}) - f(\xi_i)(x_i - x_{i-1})) .$$

Thus $F(b) - F(a)$ has been given as a Riemann sum for f plus some error term R. But it appears now that, if the partition is finer than the number δ so that (6) may be used, we have

$$\begin{aligned} |R| &\leq \sum_{i=1}^{n} \left| F(x_i) - F(x_{i-1}) - f(\xi_i)(x_i - x_{i-1}) \right| \\ &< \sum_{i=1}^{n} \varepsilon(x_i - x_{i-1}) = \varepsilon(b - a). \end{aligned}$$

Evidently, then, *if there are no mistakes here* we have just proved that f is Riemann integrable and that $\int_a^b f(t)\,dt = F(b) - F(a)$.

This is false as we have mentioned above. The error is that the choice of δ depends on the point ξ considered and so is not a constant. This is an error the reader has doubtless made in other contexts: a local condition that holds for *each* point x is misinterpreted as holding uniformly for *all* x.

But, instead of abandoning the argument, one can change the definition of the Riemann integral to allow a variable δ. The definition then changes to look like this.

Definition 8.30 A function f is *generalized Riemann integrable* on $[a, b]$ with value I if for every ε there is a positive function δ on $[a, b]$ so that

$$\left| \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) - I \right| < \varepsilon$$

whenever

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

is a partition of $[a, b]$ with $\xi_i \in [x_{i-1}, x_i]$ and $0 < x_i - x_{i-1} < \delta(\xi_i)$.

The integral $\int_a^b f(x)\,dx$ is taken as this number I that exists. It is easy to check that if f is Riemann integrable it is also generalized Riemann integrable and the integrals have the same value. Thus this new integral is an extension of the old one. To justify the definition requires knowing that such partitions actually exist for any such positive function δ; this is supplied by the Cousin theorem (Lemma 4.26).

This defines a Riemann-type integral that includes the usual Riemann integral and integrates *all* derivatives. The generalized Riemann integral was discovered in the 1950s, independently, by R. Henstock and J. Kurzweil, and these ideas have led to a number of other integration theories that exploit the geometry of the underlying space in the same way that this integral exploits the geometry of derivatives on the real line.

We shall not carry these ideas any further but refer the reader to the recent monographs of Pfeffer[2] or Gordon.[3]

Exercises

8:9.1 Develop the elementary properties of the generalized Riemann integral directly from its definition (e.g., the integral of a sum $f + g$, the integral formula $\int_a^b + \int_b^c = \int_a^c$, etc.).

8:9.2 Show directly from the definition that the characteristic function of the rationals is not Riemann integrable, but is generalized Riemann integrable on any interval, and that $\int_0^1 f(x)\,dx = 0$.

8:9.3 Show that the generalized Riemann integral is closed under the extension procedure of Cauchy from Section 8.4.

[2] W. F. Pfeffer, *The Riemann Approach to Integration: Local Geometric Theory.* Cambridge (1993).

[3] R. A. Gordon, *The Integrals of Lebesgue, Denjoy, Perron and Henstock.* Grad. Studies in Math, Vol. 4, Amer. Math. Soc. (1994).

8.10 Additional Problems for Chapter 8

8:10.1 Let $f : [0,1] \to \mathbb{R}$ be a differentiable function such that $|f'(x)| \leq M$ for all $x \in (0,1)$. Show that

$$\left| \int_0^1 f(x)\,dx - \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right) \right| \leq \frac{M}{n}.$$