CHAPTER V
INTEGRATION, AVERAGE BEHAVIOR
$A = \pi r^2$.

In this chapter we will derive the formula $A = \pi r^2$ for the area of a circle of radius r. As a matter of fact, we will first have to settle on exactly what is the definition of the area of a region in the plane, and more subtle than that, we must decide what kinds of regions in the plane "have" areas. Before we consider the problem of area, we will develop the notion of the integral (or average value) of a function defined on an interval $[a, b]$, which notion we will use later to compute areas, once they have been defined.

The main results of this chapter include:

(1) The definition of **integrability** of a function, and the definition of the **integral** of an integrable function,
(2) The **Fundamental Theorem of Calculus** (Theorem 5.9),
(3) The **Integral Form of Taylor's Remainder Theorem** (Theorem 5.12),
(4) The **General Binomial Theorem** (Theorem 5.13),
(5) The definition of the **area** of a **geometric set,**
(6) $A = \pi r^2$ (Theorem 5.15), and
(7) The **Integral Test** (Theorem 5.17).

INTEGRALS OF STEP FUNCTIONS

We begin by defining the integral of certain (but not all) bounded, real-valued functions whose domains are closed bounded intervals. Later, we will extend the definition of integral to certain kinds of unbounded complex-valued functions whose domains are still intervals, but which need not be either closed or bounded. First, we recall from Chapter III the following definitions.

DEFINITION. Let $[a, b]$ be a closed bounded interval of real numbers. By a *partition* of $[a, b]$ we mean a finite set $P = \{x_0 < x_1 < \ldots < x_n\}$ of $n+1$ points, where $x_0 = a$ and $x_n = b$.

The n intervals $\{[x_{i-1}, x_i]\}$ are called the *closed subintervals* of the partition P, and the n intervals $\{(x_{i-1}, x_i)\}$ are called the *open subintervals* or *elements* of P.

We write $\|P\|$ for the maximum of the numbers (lengths of the subintervals) $\{x_i - x_{i-1}\}$, and call $\|P\|$ the *mesh size* of the partition P.

If a partition $P = \{x_i\}$ is contained in another partition $Q = \{y_j\}$, i.e., each x_i equals some y_j, then we say that Q is *finer* than P.

Let f be a function on an interval $[a, b]$, and let $P = \{x_0 < \ldots < x_n\}$ be a partition of $[a, b]$. Physicists often consider sums of the form

$$S_{P,\{y_i\}} = \sum_{i=1}^{n} f(y_i)(x_i - x_{i-1}),$$

where y_i is a point in the subinterval (x_{i-1}, x_i). These sums (called Riemann sums) are approximations of physical quantities, and the limit of these sums, as the mesh of the partition becomes smaller and smaller, should represent a precise value of the physical quantity. What precisely is meant by the " limit" of such sums is already a subtle question, but even having decided on what that definition should be, it is as important and difficult to determine whether or not such a limit exists for many (or even any) functions f. We approach this question from a slightly different point of view, but we will revisit Riemann sums in the end.

Again we recall from Chapter III the following.

DEFINITION. Let $[a, b]$ be a closed bounded interval in $\mathbb{R}$. A real-valued function $h : [a, b] \to \mathbb{R}$ is called a *step function* if there exists a partition $P = \{x_0 < x_1 < \ldots < x_n\}$ of $[a, b]$ such that for each $1 \leq i \leq n$ there exists a number a_i such that $h(x) = a_i$ for all $x \in (x_{i-1}, x_i)$.

REMARK. A step function h is constant on the open subintervals (or elements) of a certain partition. Of course, the partition is not unique. Indeed, if P is such a partition, we may add more points to it, making a larger partition having more subintervals, and the function h will still be constant on these new open subintervals. That is, a given step function can be described using various distinct partitions.

Also, the values of a step function at the partition points themselves is irrelevant. We only require that it be constant on the open subintervals.

Exercise 5.1. Let h be a step function on $[a, b]$, and let $P = \{x_0 < x_1 < \ldots < x_n\}$ be a partition of $[a, b]$ such that $h(x) = a_i$ on the subinterval (x_{i-1}, x_i) determined by P.

(a) Prove that the range of h is a finite set. What is an upper bound on the cardinality of this range?

(b) Prove that h is differentiable at all but a finite number of points in $[a, b]$. What is the value of h' at such a point?

(c) Let f be a function on $[a, b]$. Prove that f is a step function if and only if $f'(x)$ exists and $= 0$ for every $x \in (a, b)$ except possibly for a finite number of points.

(d) What can be said about the values of h at the endpoints $\{x_i\}$ of the subintervals of P?

(e) Let h be a step function on $[a, b]$, and let j be a function on $[a, b]$ for which $h(x) = j(x)$ for all $x \in [a, b]$ except for one point c. Show that j is also a step function.

(f) If k is a function on $[a, b]$ that agrees with a step function h except at a finite number of points $c_1, c_2, \ldots, c_N$, show that k is also a step function.

Exercise 5.2. Let $[a, b]$ be a fixed closed bounded interval in $\mathbb{R}$, and let $H([a, b])$ denote the set of all step functions on $[a, b]$.

(a) Using Part (c) of Exercise 5.1, prove that the set $H([a, b])$ is a vector space of functions; i.e., it is closed under addition and scalar multiplication.

(b) Show that $H([a, b])$ is closed under multiplication; i.e., if $h_1, h_2 \in H([a, b])$, then $h_1 h_2 \in H([a, b])$.

(c) Show that $H([a, b])$ is closed under taking maximum and minimum and that it contains all the real-valued constant functions.

(d) We call a function χ an *indicator function* if it equals 1 on an interval (c, d) and is 0 outside $[c, d]$. To be precise, we will denote this indicator function by $\chi_{(c,d)}$. Prove that every indicator function is a step function, and show also that every step function h is a linear combination of indicator functions:

$$h = \sum_{j=1}^{n} a_j \chi_{(c_j, d_j)}.$$

(e) Define a function k on $[0,1]$ by setting $k(x) = 0$ if x is a rational number and $k(x) = 1$ if x is an irrational number. Prove that the range of k is a finite set, but that k is **not** a step function.

Our first theorem in this chapter is a fundamental consistency result about the "area under the graph" of a step function. Of course, the graph of a step function looks like a collection of horizontal line segments, and the region under this graph is just a collection of rectangles. Actually, in this remark, we are implicitly thinking that the values $\{a_i\}$ of the step function are positive. If some of these values are negative, then we must re-think what we mean by the area under the graph. We first introduce the following bit of notation.

DEFINITION. Let h be a step function on the closed interval $[a,b]$. Suppose $P = \{x_0 < x_1 < \ldots < x_n\}$ is a partition of $[a,b]$ such that $h(x) = a_i$ on the interval (x_{i-1}, x_i). Define the *weighted average* of h *relative to* P to be the number $S_P(h)$ defined by

$$S_P(h) = \sum_{i=1}^{n} a_i(x_i - x_{i-1}).$$

REMARK. Notice the similarity between the formula for a weighted average and the formula for a Riemann sum. Note also that if the interval is a single point, i.e., $a = b$, then the only partition P of the interval consists of the single point $x_0 = a$, and every weighted average $S_P(h) = 0$.

The next theorem is not a surprise, although its proof takes some careful thinking. It is simply the assertion that the weighted averages are independent of the choice of partition.

THEOREM 5.1. *Let h be a step function on the closed interval $[a,b]$. Suppose $P = \{x_0 < x_1 < \ldots < x_n\}$ is a partition of $[a,b]$ such that $h(x) = a_i$ on the interval (x_{i-1}, x_i), and suppose $Q = \{y_0 < y_1 < \ldots < y_m\}$ is another partition of $[a,b]$ such that $h(x) = b_j$ on the interval (y_{j-1}, y_j). Then the weighted average of h relative to P is the same as the weighted average of h relative to Q. That is, $S_P(h) = S_Q(h)$.*

PROOF. Suppose first that the partition Q is obtained from the partition P by adding one additional point. Then $m = n+1$, and there exists an i_0 between 1 and $n-1$ such that

(1) for $0 \leq i \leq i_0$ we have $y_i = x_i$.
(2) $\quad x_{i_0} < y_{i_0+1} < x_{i_0+1}$.
(3) For $i_0 < i \leq n$ we have $x_i = y_{i+1}$.

In other words, y_{i_0+1} is the only point of Q that is not a point of P, and y_{i_0+1} lies strictly between x_{i_0} and x_{i_0+1}.

Because h is constant on the interval $(x_{i_0}, x_{i_0+1}) = (y_{i_0}, y_{i_0+2})$, it follows that

(1) For $1 \leq i \leq i_0$, $a_i = b_i$.
(2) $\quad b_{i_0+1} = b_{i_0+2} = a_{i_0+1}$.
(3) For $i_0 + 1 \leq i \leq n$, $a_i = b_{i+1}$.

So,

$$
\begin{aligned}
S_P(h) &= \sum_{i=1}^{n} a_i(x_i - x_{i-1}) \\
&= \sum_{i=1}^{i_0} a_i(x_i - x_{i-1}) + a_{i_0+1}(x_{i_0+1} - x_{i_0}) \\
&\qquad + \sum_{i=i_0+2}^{n} a_i(x_i - x_{i-1}) \\
&= \sum_{i=1}^{i_0} b_i(y_i - y_{i-1}) + a_{i_0+1}(y_{i_0+2} - y_{i_0}) \\
&\qquad + \sum_{i=i_0+2}^{n} b_{i+1}(y_{i+1} - y_i) \\
&= \sum_{i=1}^{i_0} b_i(y_i - y_{i-1}) + a_{i_0+1}(y_{i_0+2} - y_{i_0+1} + y_{i_0+1} - y_{i_0}) \\
&\qquad + \sum_{i=i_0+3}^{n+1} b_i(y_i - y_{i-1}) \\
&= \sum_{i=1}^{i_0} b_i(y_i - y_{i-1}) + b_{i_0+1}(y_{i_0+1} - y_{i_0}) + b_{i_0+2}(y_{i_0+2} - y_{i_0+1}) \\
&\qquad + \sum_{i=i_0+3}^{m} b_i(y_i - y_{i-1}) \\
&= \sum_{i=1}^{m} b_i(y_i - y_{i-1}) \\
&= S_Q(h),
\end{aligned}
$$

which proves the theorem in this special case where Q is obtained from P by adding just one more point.

It follows easily now by induction that if Q is obtained from P by adding any finite number of additional points, then h is constant on each of the open subintervals determined by Q, and $S_Q(h) = S_P(h)$.

Finally, let $Q = \{y_0 < y_1 < \ldots < y_m\}$ be an arbitrary partition of $[a, b]$, for which h is constant on each of the open subintervals (y_{j-1}, y_j) determined by Q. Define R to be the partition of $[a, b]$ obtained by taking the union of the partition points $\{x_i\}$ and $\{y_j\}$. Then R is a partition of $[a, b]$ that is obtained by adding a finite number of points to the partition P, whence $S_R(h) = S_P(h)$. Likewise, R is obtained from the partition Q by adding a finite number of points, whence $S_R(h) = S_Q(h)$, and this proves that $S_Q(h) = S_P(h)$, as desired.

DEFINITION. Let $[a, b]$ be a fixed closed bounded interval in $\mathbb{R}$. We define the *integral* of a step function h on $[a, b]$, and denote it by $\int h$, as follows: If $P = \{x_0 < x_1 < \ldots < x_n\}$ is a partition of $[a, b]$, for which $h(x) = a_i$ for all $x \in (x_{i-1}, x_i)$, then

$$\int h = S_P(h) = \sum_{i=1}^{n} a_i(x_i - x_{i-1}).$$

REMARK. The integral of a step function h is defined to be the weighted average of h relative to a partition P of $[a,b]$. Notice that the preceding theorem is crucial in order that this definition of $\int h$ be unambiguously defined. The integral of a step function should **not** depend on which partition is used. Theorem 5.1 asserts precisely this fact.

Note also that if the interval is a single point, i.e., $a = b$, then the integral of every step function h is 0.

We use a variety of notations for the integral of h :

$$\int h = \int_a^b h = \int_a^b h(t)\,dt.$$

The following exercise provides a very useful way of describing the integral of a step function. Not only does it show that the integral of a step function looks like a Riemann sum, but it provides a description of the integral that makes certain calculations easier. See, for example, the proof of the next theorem.

Exercise 5.3. Suppose h is a step function on $[a,b]$ and that $R = \{z_0 < z_1 < \ldots < z_n\}$ is a partition of $[a,b]$ for which h is constant on each subinterval (z_{i-1}, z_i) of R.

(a) Prove that

$$\int h = S_R(h) = \sum_{i=1}^{n} h(w_i)(z_i - z_{i-1}),$$

where, for each $1 \le i \le n$, w_i is any point in (z_{i-1}, z_i). (Note then that the integral of a step function takes the form of a Riemann sum.)

(b) Show that $\int h$ is independent of the values of h at the points $\{z_i\}$ of the partition R.

Exercise 5.4. Let h_1 and h_2 be two step functions on $[a,b]$.

(a) Suppose that $h_1(x) = h_2(x)$ for all $x \in [a,b]$ except for one point c. Prove that $\int h_1 = \int h_2$.

HINT: Let P be a partition of $[a,b]$, for which both h_1 and h_2 are constant on its open subintervals, and for which c is one of the points of P. Now use the preceding exercise to calculate the two integrals.

(b) Suppose $h_1(x) = h_2(x)$ for all but a finite number of points $c_1, \ldots, c_N \in [a,b]$. Prove that $\int h_1 = \int h_2$.

We have used the terminology "weighted average" of a step function relative to a partition P. The next exercise shows how the integral of a step function can be related to an actual average value of the function.

Exercise 5.5. Let h be a step function on the closed interval $[a,b]$, and let $P = \{x_0 < x_1 < \ldots < x_n\}$ be a partition of $[a,b]$ for which $h(x) = a_i$ on the interval (x_{i-1}, x_i). Let us think of the interval $[a,b]$ as an interval of time, and suppose that the function h assumes the value a_i for the interval of time between x_{i-1} and

x_i. Show that the average value $A(h)$ taken on by h throughout the entire interval ($[a,b]$) of time is given by

$$A(h) = \frac{\int h}{b-a}.$$

THEOREM 5.2. *Let $H([a,b])$ denote the vector space of all step functions on the closed interval $[a,b]$. Then the assignment $h \to \int h$ of $H([a,b])$ into $\mathbb{R}$ has the following properties:*

(1) (Linearity) $H([a,b])$ is a vector space. Furthermore, $\int(h_1+h_2) = \int h_1 + \int h_2$, and $\int ch = c\int h$ for all $h_1, h_2, h \in H([a,b])$, and for all real numbers c.
(2) If $h = \sum_{i=1}^n a_i\chi_{(c_i,d_i)}$ is a linear combination of indicator functions (See part (d) of Exercise 5.2), then $\int h = \sum_{i=1}^n a_i(d_i - c_i)$.
(3) (Positivity) If $h(x) \geq 0$ for all $x \in [a,b]$, then $\int h \geq 0$.
(4) (Order-preserving) If h_1 and h_2 are step functions for which $h_1(x) \leq h_2(x)$ for all $x \in [a,b]$, then $\int h_1 \leq \int h_2$.

PROOF. That $H([a,b])$ is a vector space was proved in part (a) of Exercise 5.2. Suppose $P = \{x_0 < x_1 < \ldots < x_n\}$ is a partition of $[a,b]$ such that $h_1(x)$ is constant for all $x \in (x_{i-1}, x_i)$, and suppose $Q = \{y_0 < y_1 < \ldots < y_m\}$ is a partition of $[a,b]$ such that $h_2(x)$ is constant for all $x \in (y_{j-1}, y_j)$. Let $R = \{z_0 < z_1 < \ldots < z_r\}$ be the partition of $[a,b]$ obtained by taking the union of the x_i's and the y_j's. Then h_1 and h_2 are both constant on each open subinterval of R, since each such subinterval is contained in some open subinterval of P and also is contained in some open subinterval of Q. Therefore, $h_1 + h_2$ is constant on each open subinterval of R. Now, using Exercise 5.3, we have that

$$\begin{aligned}
\int (h_1+h_2) &= \sum_{k=1}^r ((h_1+h_2)(w_k))(z_k - z_{k-1}) \\
&= \sum_{k=1}^r h_1(w_k)(z_k - z_{k-1}) + \sum_{k=1}^r h_2(w_k)(z_k - z_{k-1}) \\
&= \int h_1 + \int h_2.
\end{aligned}$$

This proves the first assertion of part (1).

Next, let $P = \{x_0 < x_1 < \ldots < x_n\}$ be a partition of $[a,b]$ such that $h(x)$ is constant on each open subinterval of P. Then $ch(x)$ is constant on each open subinterval of P, and using Exercise 5.3 again, we have that

$$\begin{aligned}
\int (ch) &= \sum_{i=1}^n ch(w_i)(x_i - x_{i-1}) \\
&= c\sum_{i=1}^n h(w_i)(x_i - x_{i-1}) \\
&= c\int h,
\end{aligned}$$

which completes the proof of the other half of part (1).

To see part (2), we need only verify that $\int \chi_{(c_i,d_i)} = d_i - c_i$, for then part (2) will follow from part (1). But $\chi_{(c_i,d_i)}$ is just a step function determined by the four point partition $\{a, c_i, d_i, b\}$ and values 0 on (a, c_i) and (d_i, b) and 1 on (c_i, d_i). Therefore, we have that $\int \chi_{(c_i,d_i)} = d_i - c_i$.

If $h(x) \geq 0$ for all x, and $P = \{x_0 < x_1 < \ldots < x_n\}$ is as above, then

$$\int h = \sum_{i=1}^{n} h(w_i)(x_i - x_{i-1}) \geq 0,$$

and this proves part (3).

Finally, suppose $h_1(x) \leq h_2(x)$ for all $x \in [a, b]$. By Exercise 5.2, we know that the function $h_3 = h_2 - h_1$ is a step function on $[a, b]$. Also, $h_3(x) \geq 0$ for all $x \in [a, b]$. So, by part (3), $\int h_3 \geq 0$. Then, by part (1),

$$0 \leq \int h_3 = \int (h_2 - h_1) = \int h_2 - \int h_1,$$

which implies that $\int h_1 \leq \int h_2$, as desired.

Exercise 5.6. (a) Let h be the constant function c on $[a, b]$. Show that $\int h = c(b - a)$.

(b) Let $a < c < d < b$ be real numbers, and let h be the step function on $[a, b]$ that equals r for $c < x < d$ and 0 otherwise. Prove that $\int_a^b h(t)\,dt = r(d - c)$.

(c) Let h be a step function on $[a, b]$. Prove that $|h|$ is a step function, and that $|\int h| \leq \int |h|$.

HINT: Note that $-|h|(x) \leq h(x) \leq |h|(x)$. Now use the preceding theorem.

(d) Suppose h is a step function on $[a, b]$ and that c is a constant for which $|h(x)| \leq c$ for all $x \in [a, b]$. Prove that $|\int h| \leq c(b - a)$.

INTEGRABLE FUNCTIONS

We now wish to extend the definition of the integral to a wider class of functions. This class will consist of those functions that are **uniform limits** of step functions. The requirement that these limits be uniform is crucial. Pointwise limits of step functions doesn't work, as we will see in Exercise 5.7 below. The initial step in carrying out this generalization is the following.

THEOREM 5.3. *Let $[a, b]$ be a closed bounded interval, and let $\{h_n\}$ be a sequence of step functions that converges uniformly to a function f on $[a, b]$. Then the sequence $\{\int h_n\}$ is a convergent sequence of real numbers.*

PROOF. We will show that $\{\int h_n\}$ is a Cauchy sequence in $\mathbb{R}$. Thus, given an $\epsilon > 0$, choose an N such that for any $n \geq N$ and any $x \in [a, b]$, we have

$$|f(x) - h_n(x)| < \frac{\epsilon}{2(b - a)}.$$

Then, for any m and n both $\geq N$ and any $x \in [a, b]$, we have

$$|h_n(x) - h_m(x)| \leq |h_n(x) - f(x)| + |f(x) - h_m(x)| < \frac{\epsilon}{b - a}.$$

Therefore,

$$|\int h_n - \int h_m| = |\int (h_n - h_m)| \leq \int |h_n - h_m| \leq \int \frac{\epsilon}{b-a} = \epsilon,$$

as desired.

The preceding theorem provides us with a perfectly good idea of how to define the integral of a function f that is the uniform limit of a sequence of step functions. However, we first need to establish another kind of consistency result.

THEOREM 5.4. *If $\{h_n\}$ and $\{k_n\}$ are two sequences of step functions on $[a,b]$, each converging uniformly to the same function f, then*

$$\lim \int h_n = \lim \int k_n.$$

PROOF. Given $\epsilon > 0$, choose N so that if $n \geq N$, then $|h_n(x) - f(x)| < \epsilon/(2(b-a))$ for all $x \in [a,b]$, and such that $|f(x) - k_n(x)| < \epsilon/(2(b-a))$ for all $x \in [a,b]$. Then, $|h_n(x) - k_n(x)| < \epsilon/(b-a)$ for all $x \in [a,b]$ if $n \geq N$. So,

$$|\int h_n - \int k_n| \leq \int |h_n - k_n| \leq \int \frac{\epsilon}{b-a} = \epsilon$$

if $n \geq N$. Taking limits gives

$$|\lim \int h_n - \lim \int k_n| \leq \epsilon.$$

Since this is true for arbitrary $\epsilon > 0$, it follows that $\lim \int h_n = \lim \int k_n$, as desired.

DEFINITION. Let $[a,b]$ be a closed bounded interval of real numbers. A function $f : [a,b] \to \mathbb{R}$ is called *integrable* on $[a,b]$ if it is the uniform limit of a sequence $\{h_n\}$ of step functions.

Let $I([a,b])$ denote the set of all functions that are integrable on $[a,b]$. If $f \in I([a,b])$, define the *integral* of f, denoted $\int f$, by

$$\int f = \lim \int h_n,$$

where $\{h_n\}$ is some (any) sequence of step functions that converges uniformly to f on $[a,b]$.

As in the case of step functions, we use the following notations:

$$\int f = \int_a^b f = \int_a^b f(t)\,dt.$$

REMARK. Note that Theorem 5.4 is crucial in order that this definition be unambiguous. Indeed, we will see below that this critical consistency result is one place where uniform limits of step functions works while pointwise limits do not. See parts (c) and (d) of Exercise 5.7. Note also that it follows from this definition that $\int_a^a f = 0$, because $\int_a^a h = 0$ for any step function. In fact, we will derive almost everything about the integral of a general integrable function from the corresponding results about the integral of a step function. No surprise. This is the essence of mathematical analysis, approximation.

Exercise 5.7. Define a function f on the closed interval $[0,1]$ by $f(x) = 1$ if x is a rational number and $f(x) = 0$ if x is an irrational number.

(a) Suppose h is a step function on $[0,1]$. Prove that there must exist an $x \in [0,1]$ such that $|f(x) - h(x)| \geq 1/2$.
HINT: Let (x_{i-1}, x_i) be an interval on which h is a constant c. Now use the fact that there are both rationals and irrationals in this interval.

(b) Prove that f is not the uniform limit of a sequence of step functions. That is, f is **not** an integrable function.

(c) Consider the two sequences $\{h_n\}$ and $\{k_n\}$ of step functions defined on the interval $[0,1]$ by $h_n = \chi_{(0,1/n)}$, and $k_n = n\chi_{(0,1/n)}$. Show that both sequences $\{h_n\}$ and $\{k_n\}$ converge pointwise to the 0 function on $[0,1]$.
HINT: All functions are 0 at $x = 0$. For $x > 0$, choose N so that $1/N < x$. Then, for any $n \geq N$, $h_n(x) = k_n(x) = 0$.

(d) Let h_n and k_n be as in part (c). Show that $\lim \int h_n = 0$, but $\lim \int k_n = 1$. Conclude that the consistency result in Theorem 5.4 does not hold for pointwise limits of step functions.

Exercise 5.8. Define a function f on the closed interval $[0,1]$ by $f(x) = x$.

(a) For each positive integer n, let P_n be the partition of $[0,1]$ given by the points $\{0 < 1/n < 2/n < 3/n < \ldots < (n-1)/n < 1\}$. Define a step function h_n on $[0,1]$ by setting $h_n(x) = i/n$ if $\frac{i-1}{n} < x < \frac{i}{n}$, and $h_n(i/n) = i/n$ for all $0 \leq i \leq n$. Prove that $|f(x) - h_n(x)| < 1/n$ for all $x \in [0,1]$, and then conclude that f is the uniform limit of the h_n's whence $f \in I([0,1])$.

(b) Show that

$$\int h_n = \sum_{i=1}^{n} \frac{i}{n^2} = \frac{n(n+1)}{2n^2}.$$

(c) Show that $\int_0^1 f(t)\,dt = 1/2$.

The next exercise establishes some additional properties of integrable functions on an interval $[a,b]$.

Exercise 5.9. Let $[a,b]$ be a closed and bounded interval, and let f be an element of $I([a,b])$.

(a) Show that, for each $\epsilon > 0$ there exists a step function h on $[a,b]$ such that $|f(x) - h(x)| < \epsilon$ for all $x \in [a,b]$.

(b) For each positive integer n let h_n be a step function satisfying the conclusion of part (a) for $\epsilon = 1/n$. Define $k_n = h_n - 1/n$ and $l_n = h_n + 1/n$. Show that k_n and l_n are step functions, that $k_n(x) < f(x) < l_n(x)$ for all $x \in [a,b]$, and that $|l_n(x) - k_n(x)| = l_n(x) - k_n(x) = 2/n$ for all x. Hence, $\int_a^b (l_n - k_n) = \frac{2}{n}(b-a)$.

(c) Conclude from part (b) that, given any $\epsilon > 0$, there exist step functions k and l such that $k(x) \leq f(x) \leq l(x)$ for which $\int (l(x) - k(x)) < \epsilon$.

(d) Prove that there exists a sequence $\{j_n\}$ of step functions on $[a,b]$, for which $j_n(x) \leq j_{n+1}(x) \leq f(x)$ for all x, that converges uniformly to f. Show also that there exists a sequence $\{j'_n\}$ of step functions on $[a,b]$, for which $j'_n(x) \geq j'_{n+1}(x) \geq f(x)$ for all x, that converges uniformly to f. That is, if $f \in I([a,b])$, then f is the

uniform limit of a nondecreasing sequence of step functions and also is the uniform limit of a nonincreasing sequence of step functions.
HINT: To construct the j_n's and j_n''s, use the step functions k_n and l_n of part (b), and recall that the maximum and minimum of step functions is again a step function.

(e) Show that if $f(x) \geq 0$ for all $x \in [a, b]$, and g is defined by $g(x) = \sqrt{f(x)}$, then $g \in I([a, b])$.
HINT: Write $f = \lim h_n$ where $h_n(x) \geq 0$ for all x and n. Then use part (g) of Exercise 3.28.

(f) (Riemann sums again.) Show that, given an $\epsilon > 0$, there exists a partition P such that if $Q = \{x_0 < x_1 < \ldots < x_n\}$ is any partition finer than P, and $\{w_i\}$ are any points for which $w_i \in (x_{i-1}, x_i)$, then

$$|\int_a^b f(t)\,dt - \sum_{i=1}^n f(w_i)(x_i - x_{i-1})| < \epsilon.$$

HINT: Let P be a partition for which both the step functions k and l of part (c) are constant on the open subintervals of P. Verify that for any finer partition Q, $l(w_i) \geq f(w_i) \geq k(w_i)$, and hence

$$\sum_i l(w_i)(x_i - x_{i-1}) \geq \sum_i f(w_i)(x_i - x_{i-1}) \geq \sum_i k(w_i)(x_i - x_{i-1}).$$

DEFINITION. A bounded real-valued function f on a closed bounded interval $[a, b]$ is called *Riemann-integrable* if, given any $\epsilon > 0$, there exist step functions k and l, on $[a, b]$ for which $k(x) \leq f(x) \leq l(x)$ for all x, such that $\int (l - k) < \epsilon$. We denote the set of all functions on $[a, b]$ that are Riemann-integrable by $I_R([a, b])$.

REMARK. The notion of Riemann-integrability was introduced by Riemann in the mid nineteenth century and was the first formal definition of integrability. Since then several other definitions have been given for an integral, culminating in the theory of Lebesgue integration. The definition of integrability that we are using in this book is slightly different and less general from that of Riemann, and both of these are very different and less general from the definition given by Lebesgue in the early twentieth century. Part (c) of Exercise 5.9 above shows that the functions we are calling integrable are necessarily Riemann-integrable. We will see in Exercise 5.10 that there are Riemann-integrable functions that are **not** integrable in our sense. In both cases, Riemann's and ours, an integrable function f must be trapped between two step functions k and l. In our definition, we must have $l(x) - k(x) < \epsilon$ for all $x \in [a, b]$, while in Riemann's definition, we only need that $\int l - k < \epsilon$. The distinction is that a small step function must have a small integral, but it isn't necessary for a step function to be (uniformly) small in order for it to have a small integral. It only has to be small on most of the interval $[a, b]$.

On the other hand, all the definitions of integrability on $[a, b]$ include among the integrable functions the continuous ones. And, all the different definitions of integral give the same value to a continuous function. The differences then in these definitions shows up at the point of saying exactly which functions are integrable. Perhaps the most enlightening thing to say in this connection is that it is impossible to make a "good" definition of integrability in such a way that every function is integrable. Subtle points in set theory arise in such attempts, and many fascinating and deep mathematical ideas have come from them. However, we will stick with our definition, since it is simpler than Riemann's and is completely sufficient for our purposes.

THEOREM 5.5. *Let $[a,b]$ be a fixed closed and bounded interval, and let $I([a,b])$ denote the set of integrable functions on $[a,b]$. Then:*

(1) *Every element of $I([a,b])$ is a bounded function. That is, integrable functions are necessarily bounded functions.*
(2) *$I([a,b])$ is a vector space of functions.*
(3) *$I([a,b])$ is closed under multiplication; i.e., if f and $g \in I([a,b])$, then $fg \in I([a,b])$.*
(4) *Every step function is in $I([a,b])$.*
(5) *If f is a continuous real-valued function on $[a,b]$, then f is in $I([a,b])$. That is, every continuous real-valued function on $[a,b]$ is integrable on $[a,b]$.*

PROOF. Let $f \in I([a,b])$, and write $f = \lim h_n$, where $\{h_n\}$ is a sequence of step functions that converges uniformly to f. Given the positive number $\epsilon = 1$, choose N so that $|f(x) - h_N(x)| < 1$ for all $x \in [a,b]$. Then $|f(x)| \leq |h_N(x)| + 1$ for all $x \in [a,b]$. Because h_N is a step function, its range is a finite set, so that there exists a number M for which $|h_N(x)| \leq M$ for all $x \in [a,b]$. Hence, $|f(x)| \leq M + 1$ for all $x \in [a,b]$, and this proves part (1).

Next, let f and g be integrable, and write $f = \lim h_n$ and $g = \lim k_n$, where $\{h_n\}$ and $\{k_n\}$ are sequences of step functions that converge uniformly to f and g respectively. If s and t are real numbers, then the sequence $\{sh_n + tk_n\}$ converges uniformly to the function $sf + tg$. See parts (c) and (d) of Exercise 3.28. Therefore, $sf + tg \in I([a,b])$, and $I([a,b])$ is a vector space, proving part (2).

Note that part (3) does not follow immediately from Exercise 3.28; the product of uniformly convergent sequences may not be uniformly convergent. To see it for this case, let $f = \lim h_n$ and $g = \lim k_n$ be elements of $I([a,b])$. By part (1), both f and g are bounded, and we write M_f and M_g for numbers that satisfy $|f(x)| \leq M_f$ and $|g(x)| \leq M_g$ for all $x \in [a,b]$. Because the sequence $\{k_n\}$ converges uniformly to g, there exists an N such that if $n \geq N$ we have $|g(x) - k_n(x)| < 1$ for all $x \in [a,b]$. This implies that, if $n \geq N$, then $|k_n(x)| \leq M_g + 1$ for all $x \in [a,b]$.

Now we show that fg is the uniform limit of the sequence $h_n k_n$. For, if $n \geq N$, then

$$\begin{aligned}|f(x)g(x) - h_n(x)k_n(x)| &= |f(x)g(x) - f(x)k_n(x) + f(x)k_n(x) - h_n(x)k_n(x)| \\ &\leq |f(x)||g(x) - k_n(x)| + |k_n(x)||f(x) - h_n(x)| \\ &\leq M_f|g(x) - k_n(x)| + (M_g + 1)|f(x) - h_n(x)|,\end{aligned}$$

which implies that $fg = \lim(h_n k_n)$.

If h is itself a step function, then it is obviously the uniform limit of the constant sequence $\{h\}$, which implies that h is integrable.

Finally, if f is continuous on $[a,b]$, it follows from Theorem 3.20 that f is the uniform limit of a sequence of step functions, whence $f \in I([a,b])$.

Exercise 5.10. Let f be the function defined on $[0,1]$ by $f(x) = \sin(1/x)$ if $x \neq 0$ and $f(0) = 0$.

(a) Show that f is continuous at every nonzero x and discontinuous at 0.

HINT: Observe that, on any interval $(0, \delta)$, the function $\sin(1/x)$ attains both the values 1 and -1.

(b) Show that f is not integrable on $[0, 1]$.

HINT: Suppose $f = \lim h_n$. Choose N so that $|f(x) - h_N(x)| < 1/2$ for all $x \in [0, 1]$. Let P be a partition for which h_N is constant on its open subintervals, and examine the situation for x's in the interval (x_0, x_1).

(c) Show that f **is** Riemann-integrable on $[0, 1]$. Conclude that $I([a, b])$ is a proper subset of $I_R([a, b])$.

Exercise 5.11. (a) Let f be an integrable function on $[a, b]$. Suppose g is a function for which $g(x) = f(x)$ for all $x \in [a, b]$ except for one point c. Prove that g is integrable and that $\int g = \int f$.

HINT: If $f = \lim h_n$, define $k_n(x) = h_n(x)$ for all $x \neq c$ and $k_n(c) = g(c)$. Then use Exercise 5.4.

(b) Again, let f be an integrable function on $[a, b]$. Suppose g is a function for which $g(x) = f(x)$ for all but a finite number of points $c_1, \ldots, c_N \in [a, b]$. Prove that $g \in I([a, b])$, and that $\int g = \int f$.

(c) Suppose f is a function on the closed interval $[a, b]$, that is uniformly continuous on the open interval (a, b). Prove that f is integrable on $[a, b]$.

HINT: Just reproduce the proof to Theorem 3.20.

REMARK. In view of part (b) of the preceding exercise, we see that whether a function f is integrable or not is totally independent of the values of the function at a fixed finite set of points. Indeed, the function needn't even be defined at a fixed finite set of points, and still it can be integrable. This observation is helpful in many instances, e.g., in parts (d) and (e) of Exercise 5.21.

THEOREM 5.6. *The assignment $f \to \int f$ on $I([a, b])$ satisfies the following properties.*

1. (Linearity) $I([a, b])$ is a vector space, and $\int(\alpha f + \beta g) = \alpha \int f + \beta \int g$ for all $f, g \in I([a, b])$and $\alpha, \beta \in \mathbb{R}$.
2. (Positivity) If $f(x) \geq 0$ for all $x \in [a, b]$, then $\int f \geq 0$.
3. (Order-preserving) If $f, g \in I([a, b])$ and $f(x) \leq g(x)$ for all $x \in [a, b]$, then $\int f \leq \int g$.
4. If $f \in I([a, b])$, then so is $|f|$, and $|\int f| \leq \int |f|$.
5. If f is the uniform limit of functions f_n, each of which is in $I([a, b])$, then $f \in I([a, b])$ and $\int f = \lim \int f_n$.
6. Let $\{u_n\}$ be a sequence of functions in $I([a, b])$. Suppose that for each n there is a number m_n, for which $|u_n(x)| \leq m_n$ for all $x \in [a, b]$, and such that the infinite series $\sum m_n$ converges. Then the infinite series $\sum u_n$ converges uniformly to an integrable function, and $\int \sum u_n = \sum \int u_n$.

PROOF. That $I([a, b])$ is a vector space was proved in part (2) of Theorem 5.5. Let f and g be in $I([a, b])$, and write $f = \lim h_n$ and $g = \lim k_n$, where the h_n's and the k_n's are step functions. Then $\alpha f + \beta g = \lim(\alpha h_n + \beta k_n)$, so that, by Theorem

5.2 and the definition of the integral, we have

$$\begin{aligned}\int(\alpha f+\beta g) &= \lim\int(\alpha h_n+\beta k_n)\\ &= \lim(\alpha\int h_n+\beta\int k_n)\\ &= \alpha\lim\int h_n+\beta\lim\int k_n\\ &= \alpha\int f+\beta\int g,\end{aligned}$$

which proves part (1).

Next, if $f\in I([a,b])$ satisfies $f(x)\geq 0$ for all $x\in[a,b]$, let $\{l_n\}$ be a nonincreasing sequence of step functions that converges uniformly to f. See part (d) of Exercise 5.9. Then $l_n(x)\geq f(x)\geq 0$ for all x and all n. So, again by Theorem 5.2, we have that

$$\int f=\lim\int l_n\geq 0.$$

This proves part (2).

Part (3) now follows by combining parts (1) and (2) just as in the proof of Theorem 5.2.

To see part (4), let $f\in I([a,b])$ be given. Write $f=\lim h_n$. Then $|f|=\lim|h_n|$. For

$$||f(x)|-|h_n(x)||\leq|f(x)-h_n(x)|.$$

Therefore, $|f|$ is integrable. Also,

$$\int|f|=\lim\int|h_n|\geq\lim|\int h_n|=|\lim\int h_n|=|\int f|.$$

To see part (5), let $\{f_n\}$ be a sequence of elements of $I([a,b])$, and suppose that $f=\lim f_n$. For each n, let h_n be a step function on $[a,b]$ such that $|f_n(x)-h_n(x)|<1/n$ for all $x\in[a,b]$. Note also that it follows from parts (3) and (4) that

$$|\int f_n-\int h_n|<\frac{b-a}{n}.$$

Now $\{h_n\}$ converges uniformly to f. For,

$$\begin{aligned}|f(x)-h_n(x)| &\leq|f(x)-f_n(x)|+|f_n(x)-h_n(x)|\\ &<|f(x)-f_n(x)|+\frac{1}{n},\end{aligned}$$

showing that $f=\lim h_n$. Therefore, $f\in I([a,b])$. Moreover, $\int f=\lim\int h_n$. Finally, $\int f=\lim\int f_n$, for

$$\begin{aligned}|\int f-\int f_n| &\leq|\int f-\int h_n|+|\int h_n-\int f_n|\\ &\leq|\int f-\int h_n|+\frac{b-a}{n}.\end{aligned}$$

This completes the proof of part (5).

Part (6) follows directly from part (5) and the Weierstrass M Test (Theorem 3.18). For, part (1) of that theorem implies that the infinite series $\sum u_n$ converges uniformly, and then $\int \sum u_n = \sum \int u_n$ follows from part (5) of this theorem.

As a final extension of our notion of integral, we define the integral of certain complex-valued functions.

DEFINITION. Let $[a, b]$ be a fixed bounded and closed interval. A complex-valued function $f = u + iv$ is called *integrable* if its real and imaginary parts u and v are integrable. In this case, we define

$$\int_a^b f = \int_a^b (u + iv) = \int_a^b u + i \int_a^b v.$$

THEOREM 5.7.

(1) *The set of all integrable complex-valued functions on $[a, b]$ is a vector space over the field of complex numbers, and*

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g$$

for all integrable complex-valued functions f and g and all complex numbers α and β.

(2) *If f is an integrable complex-valued function on $[a, b]$, then so is $|f|$, and $|\int_a^b f| \leq \int_a^b |f|$.*

PROOF. We leave the verification of part (1) to the exercise that follows.

To see part (2), suppose that f is integrable, and write $f = u + iv$. Then $|f| = \sqrt{u^2 + v^2}$, so that $|f|$ is integrable by Theorem 5.5 and part (e) of Exercise 5.9. Now write $z = \int_a^b f$, and write z in polar coordinates as $z = re^{i\theta}$, where $r = |z| = |\int_a^b f|$. (See Exercise 4.23.) Define a function g by $g(x) = e^{-i\theta} f(x)$ and notice that $|g| = |f|$. Then $\int_a^b g = e^{-i\theta} \int_a^b f = r$, which is a real number. Writing $g = \widehat{u} + i\widehat{v}$, we then have that $r = \int \widehat{u} + i \int \widehat{v}$, implying that $\int \widehat{v} = 0$. So,

$$
\begin{aligned}
|\int_a^b f| &= r \\
&= \int_a^b g \\
&= \int_a^b \widehat{u} + i \int_a^b \widehat{v} \\
&= \int_a^b \widehat{u} \\
&= |\int_a^b \widehat{u}| \\
&\leq \int_a^b |\widehat{u}| \\
&\leq \int_a^b |g| \\
&= \int_a^b |f|,
\end{aligned}
$$

as desired.

Exercise 5.12. Prove part (1) of the preceding theorem.
HINT: Break $\alpha, \beta, \int f$, and $\int g$ into real and imaginary parts.

THE FUNDAMENTAL THEOREM OF CALCULUS

We begin this section with a result that is certainly not a surprise, but we will need it at various places in later proofs, so it's good to state it precisely now.

THEOREM 5.8. *Suppose $f \in I([a,b])$, and suppose $a < c < b$. Then $f \in I([a,c])$, $f \in I([c,b])$, and*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

PROOF. Suppose first that h is a step function on $[a,b]$, and let $P = \{x_0 < x_1 < \ldots < x_n\}$ be a partition of $[a,b]$ such that $h(x) = a_i$ on the subinterval (x_{i-1}, x_i) of P. Of course, we may assume without loss of generality that c is one of the points of P, say $c = x_k$. Clearly h is a step function on both intervals $[a,c]$ and $[c,b]$.

Now, let $Q_1 = \{a = x_0 < x_1 < \ldots < c = x_k\}$ be the partition of $[a,c]$ obtained by intersecting P with $[a,c]$, and let $Q_2 = \{c = x_k < x_{k+1} < \ldots < x_n = b\}$ be the partition of $[c, b]$ obtained by intersecting P with $[c, b]$. We have that

$$\begin{aligned}\int_a^b h &= S_P(h) \\ &= \sum_{i=1}^{n} a_i(x_i - x_{i-1}) \\ &= \sum_{i=1}^{k} a_i(x_i - x_{i-1}) + \sum_{i=k+1}^{n} a_i(x_i - x_{i-1}) \\ &= S_{Q_1}(h) + S_{Q_2}(h) \\ &= \int_a^c h + \int_c^b h,\end{aligned}$$

which proves the theorem for step functions.

Now, write $f = \lim h_n$, where each h_n is a step function on $[a, b]$. Then clearly $f = \lim h_n$ on $[a, c]$, which shows that $f \in I([a, c])$, and

$$\int_a^c f = \lim \int_a^c h_n.$$

Similarly, $f = \lim h_n$ on $[c, b]$, showing that $f \in I([c, b])$, and

$$\int_c^b f = \lim \int_c^b h_n.$$

Finally,

$$\begin{aligned}\int_a^b f &= \lim \int_a^b h_n \\ &= \lim(\int_a^c h_n + \int_c^b h_n) \\ &= \lim \int_a^c h_n + \lim \int_c^b h_n \\ &= \int_a^c f + \int_c^b f,\end{aligned}$$

as desired.

I's time for the trumpets again! What we call the Fundamental Theorem of Calculus was discovered by Newton and Leibniz more or less simultaneously in the seventeenth century, and it is without doubt the cornerstone of all we call mathematical analysis today. Perhaps the main theoretical consequence of this theorem is that it provides a procedure for inventing "new" functions. Polynomials are rather natural functions, power series are a simple generalization of polynomials, and then what? It all came down to thinking of a function of a variable x as being the area beneath a curve between a fixed point a and the varying point x. By now, we have polished and massaged these ideas into a careful, detailed development of the subject, which has substantially obscured the original ingenious insights of Newton and Leibniz. On the other hand, our development and proofs are complete, while theirs were based heavily on their intuition. So, here it is.

THEOREM 5.9. (Fundamental Theorem of Calculus) Suppose f is an arbitrary element of $I([a,b])$. Define a function F on $[a,b]$ by $F(x) = \int_a^x f$. Then:

(1) F is continuous on $[a,b]$, and $F(a) = 0$.
(2) If f is continuous at a point $c \in (a,b)$, then F is differentiable at c and $F'(c) = f(c)$.
(3) Suppose that f is continuous on $[a,b]$. If G is any continuous function on $[a,b]$ that is differentiable on (a,b) and satisfies $G'(x) = f(x)$ for all $x \in (a,b)$, then

$$\int_a^b f(t)\,dt = G(b) - G(a).$$

REMARK. Part (2) of this theorem is the heart of it, the great discovery of Newton and Leibniz, although most beginning calculus students often think of part (3) as the main statement. Of course it is that third part that enables us to actually compute integrals.

PROOF. Because $f \in I([a,b])$, we know that $f \in I([a,x])$ for every $x \in [a,b]$, so that $F(x)$ at least is defined.

Also, we know that f is bounded; i.e., there exists an M such that $|f(t)| \le M$ for all $t \in [a,b]$. Then, if $x, y \in [a,b]$ with $x \ge y$, we have that

$$\begin{aligned}
|F(x) - F(y)| &= |\int_a^x f - \int_a^y f| \\
&= |\int_a^y f + \int_y^x f - \int_a^y f| \\
&= |\int_y^x f| \\
&\le \int_y^x |f| \\
&\le \int_y^x M \\
&= M(x-y),
\end{aligned}$$

so that $|F(x) - F(y)| \le M|x-y| < \epsilon$ if $|x-y| < \delta = \epsilon/M$. This shows that F is (uniformly) continuous on $[a,b]$. Obviously, $F(a) = \int_a^a f = 0$, and part (1) is proved.

Next, suppose that f is continuous at $c \in (a,b)$, and write $L = f(c)$. Let $\epsilon > 0$ be given. To show that F is differentiable at c and that $F'(c) = f(c)$, we must find a $\delta > 0$ such that if $0 < |h| < \delta$ then

$$|\frac{F(c+h) - F(c)}{h} - L| < \epsilon.$$

Since f is continuous at c, choose $\delta > 0$ so that $|f(t) - f(c)| < \epsilon$ if $|t - c| < \delta$. Now, assuming that $h > 0$ for the moment, we have that

$$\begin{aligned} F(c+h) - F(c) &= \int_a^{c+h} f - \int_a^c f \\ &= \int_a^c f + \int_c^{c+h} f - \int_a^c f \\ &= \int_c^{c+h} f, \end{aligned}$$

and

$$L = \frac{\int_c^{c+h} L}{h}.$$

So, if $0 < h < \delta$, then

$$\begin{aligned} \left|\frac{F(c+h) - F(c)}{h} - L\right| &= \left|\frac{\int_c^{c+h} f(t)\,dt}{h} - \frac{\int_c^{c+h} L}{h}\right| \\ &= \left|\frac{\int_c^{c+h} (f(t) - L)\,dt}{h}\right| \\ &\leq \frac{\int_c^{c+h} |f(t) - L|\,dt}{h} \\ &= \frac{\int_c^{c+h} |f(t) - f(c)|\,dt}{h} \\ &\leq \frac{\int_c^{c+h} \epsilon}{h} \\ &= \epsilon, \end{aligned}$$

where the last inequality follows because for $t \in [c, c+h]$, we have that $|t-c| \leq h < \delta$. A similar argument holds if $h < 0$. (See the following exercise.) This proves part (2).

Suppose finally that G is continuous on $[a, b]$, differentiable on (a, b), and that $G'(x) = f(x)$ for all $x \in (a, b)$. Then, $F - G$ is continuous on $[a, b]$, differentiable on (a, b), and by part (2) $(F - G)'(x) = F'(x) - G'(x) = f(x) - f(x) = 0$ for all $x \in (a, b)$. It then follows from Exercise 4.12 that $F - G$ is a constant function C, whence,

$$G(b) - G(a) = F(b) + C - F(a) - C = F(b) = \int_a^b f(t)\,dt,$$

and the theorem is proved.

Exercise 5.13. (a) Complete the proof of part (2) of the preceding theorem; i.e., take care of the case when $h < 0$.
HINT: In this case, $a < c + h < c$. Then, write $\int_a^c f = \int_a^{c+h} f + \int_{c+h}^c f$.

(b) Suppose f is a continuous function on the closed interval $[a, b]$, and that f' exists and is continuous on the open interval (a, b). Assume further that f' is integrable on the closed interval $[a, b]$. Prove that $f(x) - f(a) = \int_a^x f'$ for all

$x \in [a,b]$. Be careful to understand how this is different from the Fundamental Theorem.

(c) Use the Fundamental Theorem to prove that for $x \geq 1$ we have

$$\ln(x) = F(x) \equiv \int_1^x \frac{1}{t}\,dt,$$

and for $0 < x < 1$ we have

$$\ln(x) = F(x) \equiv -\int_x^1 \frac{1}{t}\,dt.$$

HINT: Show that these two functions have the same derivative and agree at $x = 1$.

CONSEQUENCES OF THE FUNDAMENTAL THEOREM

The first two theorems of this section constitute the basic "techniques of integration" taught in a calculus course. However, the careful formulations of these standard methods of evaluating integrals have some subtle points, i.e., some hypotheses. Calculus students are rarely told about these details.

THEOREM 5.10. (Integration by Parts Formula) Let f and g be integrable functions on $[a,b]$, and as usual let F and G denote the functions defined by

$$F(x) = \int_a^x f, \text{ and } G(x) = \int_a^x g.$$

Then

$$\int_a^b fG = [F(b)G(b) - F(a)G(a)] - \int_a^b Fg.$$

Or, recalling that $f = F'$ and $g = G'$,

$$\int_a^b F'G = [F(b)G(b) - F(a)G(a)] - \int_a^b FG'.$$

Exercise 5.14. (a) Prove the preceding theorem.
HINT: Replace the upper limit b by a variable x, and differentiate both sides. By the way, how do we know that the functions Fg and fG are integrable?

(b) Suppose f and g are integrable functions on $[a,b]$ and that both f' and g' are continuous on (a,b) and integrable on $[a,b]$. (Of course f' and g' are not even defined at the endpoints a and b, but they can still be integrable on $[a,b]$. See the remark following Exercise 5.11.) Prove that

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

THEOREM 5.11. (Integration by Substitution) Let f be a continuous function on $[a,b]$, and suppose g is a continuous, one-to-one function from $[c,d]$ onto $[a,b]$ such that g is continuously differentiable on (c,d), and such that $a=g(c)$ and $b=g(d)$. Assume finally that g' is integrable on $[c,d]$. Then

$$\int_a^b f(t)\,dt = \int_c^d f(g(s))g'(s)\,ds.$$

PROOF. It follows from our assumptions that the function $f(g(s))g'(s)$ is continuous on (a,b) and integrable on $[c,d]$. It also follows from our assumptions that g maps the open interval (c,d) onto the open interval (a,b). As usual, let F denote the function on $[a,b]$ defined by $F(x)=\int_a^x f(t)\,dt$. Then, by part (2) of the Fundamental Theorem, F is differentiable on (a,b), and $F'=f$. Then, by the chain rule, $F\circ g$ is continuous and differentiable on (c,d) and

$$(F\circ g)'(s) = F'(g(s))g'(s) = f(g(s))g'(s).$$

So, by part (3) of the Fundamental Theorem, we have that

$$\begin{aligned}\int_c^d f(g(s))g'(s)\,ds &= \int_c^d (F\circ g)'(s)\,ds \\ &= (F\circ g)(d) - (F\circ g)(c) \\ &= F(g(d)) - F(g(c)) \\ &= F(b) - F(a) \\ &= \int_a^b f(t)\,dt,\end{aligned}$$

which finishes the proof.

Exercise 5.15. (a) Prove the "Mean Value Theorem" for integrals: If f is continuous on $[a,b]$, then there exists a $c\in(a,b)$ such that

$$\int_a^b f(t)\,dt = f(c)(b-a).$$

(b) (Uniform limits of differentiable functions. Compare with Exercise 4.26.) Suppose $\{f_n\}$ is a sequence of continuous functions on a closed interval $[a,b]$ that converges pointwise to a function f. Suppose that each derivative f_n' is continuous on the open interval (a,b), is integrable on the closed interval $[a,b]$, and that the sequence $\{f_n'\}$ converges uniformly to a function g on (a,b). Prove that f is differentiable on (a,b), and $f'=g$.
HINT: Let x be in (a,b), and let c be in the interval (a,x). Justify the following equalities, and use them together with the Fundamental Theorem to make the proof.

$$f(x)-f(c) = \lim(f_n(x)-f_n(c)) = \lim\int_c^x f_n' = \int_c^x g.$$

We revisit now the Remainder Theorem of Taylor, which we first presented in Theorem 4.19. The point is that there is another form of this theorem, the integral form, and this version is more powerful in some instances than the original one, e.g., in the general Binomial Theorem below.

THEOREM 5.12. (Integral Form of Taylor's Remainder Theorem) Let c be a real number, and let f have $n+1$ derivatives on $(c-r, c+r)$, and suppose that $f^{(n+1)} \in I([c-r, c+r])$. Then for each $c < x < c+r$,

$$f(x) - T^n_{(f,c)}(x) = \int_c^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,dt,$$

where T^n_f denotes the nth Taylor polynomial for f.

Similarly, for $c - r < x < c$,

$$f(x) - T^n_{(f,c)}(x) = \int_x^c f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,dt.$$

Exercise 5.16. Prove the preceding theorem.
HINT: Argue by induction on n, and integrate by parts.

REMARK. We return now to the general Binomial Theorem, first studied in Theorem 4.21. The proof given there used the derivative form of Taylor's remainder Theorem, but we were only able to prove the Binomial Theorem for $|t| < 1/2$. The theorem below uses the integral form of Taylor's Remainder Theorem in its proof, and it gives the full binomial theorem, i.e., for all t for which $|t| < 1$.

THEOREM 5.13. (General Binomial Theorem) Let $\alpha = a+bi$ be a fixed complex number. Then

$$(1+t)^\alpha = \sum_{k=0}^\infty \binom{\alpha}{k} t^k$$

for all $t \in (-1, 1)$.

PROOF. For clarity, we repeat some of the proof of Theorem 4.21. Given a general $\alpha = a+bi$, consider the function $g : (-1, 1) \to \mathbb{C}$ defined by $g(t) = (1+t)^\alpha$. Observe that the nth derivative of g is given by

$$g^{(n)}(t) = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{(1+t)^{n-\alpha}}.$$

Then $g \in C^\infty((-1, 1))$.

For each nonnegative integer k define

$$a_k = g^{(k)}(0)/k! = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!} = \binom{\alpha}{k},$$

and set $h(t) = \sum_{k=0}^\infty a_k t^k$. The radius of convergence for the power series function h is 1, as was shown in Exercise 4.31. We wish to show that $g(t) = h(t)$ for all $-1 < t < 1$. That is, we wish to show that g is a Taylor series function around 0. It will suffice to show that the sequence $\{S_n\}$ of partial sums of the power series function h converges to the function g. We note also that the nth partial sum is just the nth Taylor polynomial T^n_g for g.

Now, fix a t strictly between 0 and 1. The argument for t's between -1 and 0 is completely analogous.. Choose an $\epsilon > 0$ for which $\beta = (1+\epsilon)t < 1$. We let C_ϵ be a numbers such that $|\binom{\alpha}{n}| \le C_\epsilon(1+\epsilon)^n$ for all nonnegative integers n. See Exercise

4.31. We will also need the following estimate, which can be easily deduced as a calculus exercise (See part (d) of Exercise 4.11.). For all s between 0 and t, we have $(t-s)/(1+s) \leq t$. Note also that, for any $s \in (0,t)$, we have $|(1+s)^\alpha| = (1+s)^a$, and this is trapped between 1 and $(1+t)^a$. Hence, there exists a number M_t such that $|(1+s)^{\alpha-1}| \leq M_t$ for all $s \in (-0,t)$. We will need this estimate in the calculation that follows.

Then, by the integral form of Taylor's Remainder Theorem, we have:

$$
\begin{aligned}
|g(t) - \sum_{k=0}^{n} a_k t^k| &= |g(t) - T_g^n(t)| \\
&= |\int_0^t g^{(n+1)}(s) \frac{(t-s)^n}{n!}\, ds| \\
&= |\int_0^t \binom{(n+1) \times \alpha}{n+1} (1+s)^{\alpha-n-1}(t-s)^n\, ds| \\
&\leq \int_0^t |\binom{\alpha}{n+1}||(1+s)^{\alpha-1}|(n+1)|(\frac{t-s}{1+s}|^n\, ds \\
&\leq \int_0^t |\binom{\alpha}{n+1}| M_t(n+1)t^n\, ds \\
&\leq C_\epsilon M_t(n+1) \int_0^t (1+\epsilon)^{n+1} t^n\, ds \\
&= C_\epsilon M_t(n+1)(1+\epsilon)^{n+1} t^{n+1} \\
&= C_\epsilon M_t(n+1)\beta^{n+1},
\end{aligned}
$$

which tends to 0 as n goes to ∞, because $\beta < 1$. This completes the proof for $0 < t < 1$.

AREA OF REGIONS IN THE PLANE

It would be desirable to be able to assign to each subset S of the Cartesian plane $\mathbb{R}^2$ a nonnegative real number $A(S)$ called its area. We would insist based on our intuition that (i) if S is a rectangle with sides of length L and W then the number $A(S)$ should be LW, so that this abstract notion of area would generalize our intuitively fundamental one. We would also insist that (ii) if S were the union of two disjoint parts, $S = S_1 \cup S_2$, then $A(S)$ should be $A(S_1) + A(S_2)$. (We were taught in high school plane geometry that the whole is the sum of its parts.) In fact, even if S were the union of an infinite number of disjoint parts, $S = \cup_{n=1}^\infty S_n$ with $S_i \cap S_j = \emptyset$ if $i \neq j$, we would insist that (iii) $A(S) = \sum_{n=1}^\infty A(S_n)$.

The search for such a definition of area for every subset of $\mathbb{R}^2$ motivated much of modern mathematics. Whether or not such an assignment exists is intimately related to subtle questions in basic set theory, e.g., the *Axiom of Choice* and the *Continuum Hypothesis*. Most mathematical analysts assume that the Axiom of Choice holds, and as a result of that assumption, it has been shown that there can be no assignment $S \to A(S)$ satisfying the above three requirements. Conversely, if one does not assume that the Axiom of Choice holds, then it has also been shown that it is perfectly consistent to assume as a basic axiom that such an assignment $S \to A(S)$ does exist. We will not pursue these subtle points here, leaving them to a course in Set Theory or Measure Theory. However, Here's a statement of the Axiom of Choice, and we invite the reader to think about how reasonable it seems.

AXIOM OF CHOICE. *Let $\mathcal{S}$ be a collection of sets. Then there exists a set A that contains exactly one element out of each of the sets S in $\mathcal{S}$.*

The difficulty mathematicians encountered in trying to define area turned out to be involved with defining $A(S)$ for **every** subset $S \in \mathbb{R}^2$. To avoid this difficulty, we will restrict our attention here to certain " reasonable" subsets S. Of course, we certainly want these sets to include the rectangles and all other common geometric sets.

DEFINITION. By a (open) *rectangle* we will mean a set $R = (a, b) \times (c, d)$ in $\mathbb{R}^2$. That is, $R = \{(x, y) : a < x < b \text{ and } c < y < d\}$. The analogous definition of a *closed rectangle* $[a, b] \times [c, d]$ should be clear: $[a, b] \times [c, d] = \{(x, y) : a \leq x \leq b,\ c \leq y \leq d\}$.

By the *area* of a (open or closed) rectangle $R = (a, b) \times (c, d)$ or $[a, b] \times [c, d]$ we mean the number $A(R) = (b - a)(d - c)$.

. The fundamental notion behind our definition of the area of a set S is this. If an open rectangle $R = (a, b) \times (c, d)$ is a subset of S, then the area $A(S)$ surely should be greater than or equal to $A(R) = (b - a)(d - c)$. And, if S contains the disjoint union of several open rectangles, then the area of S should be greater than or equal to the sum of their areas.

We now specify precisely for which sets we will define the area. Let $[a, b]$ be a fixed closed bounded interval in $\mathbb{R}$ and let l and u be two continuous real-valued functions on $[a, b]$ for which $l(x) < u(x)$ for all $x \in (a, b)$.

DEFINITION. Given $[a, b], l,$ and u as in the above, let S be the set of all pairs $(x, y) \in \mathbb{R}^2$, for which $a < x < b$ and $l(x) < y < u(x)$. Then S is called an *open geometric set*. If we replace the $<$ signs with $\leq$ signs, i.e., if S is the set of all (x, y) such that $a \leq x \leq b$, and $l(x) \leq y \leq u(x)$, then S is called a *closed geometric set*. In either case, we say that S is bounded on the left and right by the vertical line segments $\{(a, y) : l(a) \leq y \leq u(a)\}$ and $\{(b, y) : l(b) \leq y \leq u(b)\}$, and it is bounded below by the graph of the function l and bounded above by the graph of the function u. We call the union of these four bounding curves the *boundary* of S, and denote it by C_S.

If the bounding functions u and l of a geometric set S are smooth or piecewise smooth functions, we will call S a *smooth* or *piecewise smooth* geometric set.

If S is a closed geometric set, we will indicate the corresponding open geometric set by the symbol S^0.

The symbol S^0 we have introduced for the open geometric set corresponding to a closed one is the same symbol that we have used previously for the interior of a set. Study the exercise that follows to see that the two uses of this notation agree.

Exercise 5.17. (a) Show that rectangles, triangles, and circles are geometric sets. What in fact is the definition of a circle?

(b) Find some examples of sets that are **not** geometric sets. Think about a horseshoe on its side, or a heart on its side.

(c) Let f be a continuous, nonnegative function on $[a, b]$. Show that the "region" under the graph of f is a geometric set.

(d) Show that the intersection of two geometric sets is a geometric set. Describe the left, right, upper, and lower boundaries of the intersection. Prove that the interior $(S_1 \cap S_2)^0$ of the intersection of two geometric sets S_1 and S_2 coincides with the intersection $S_1^0 \cap S_2^0$ of their two interiors.

(e) Give an example to show that the union of two geometric sets need not be a geometric set.

(f) Show that every closed geometric set is compact.

(g) Let S be a closed geometric set. Show that the corresponding open geometric set S^0 coincides with the interior of S, i.e., the set of all points in the interior of S. HINT: Suppose $a < x < b$ and $l(x) < y < u(x)$. Begin by showing that, because both l and u are continuous, there must exist an $\epsilon > 0$ and a $\delta > 0$ such that $a < x - \delta < x + \delta < b$ and $l(x) < y - \epsilon < y + \epsilon < u(x)$.

Now, given a geometric set S (either open or closed), that is determined by an interval $[a, b]$ and two bounding functions u and l, let $P = \{x_0 < x_1 < \ldots < x_n\}$ be a partition of $[a, b]$. For each $1 \leq i \leq n$, define numbers c_i and d_i as follows:

$$c_i = \sup_{x_{i-1} < x < x_i} l(x), \text{ and } d_i = \inf_{x_{i-1} < x < x_i} u(x).$$

Because the functions l and u are continuous, they are necessarily bounded, so that the supremum and infimum above are real numbers. For each $1 \leq i \leq n$ define R_i to be the open rectangle $(x_{i-1}, x_i) \times (c_i, d_i)$. Of course, d_i may be $< c_i$, in which case the rectangle R_i is the empty set. In any event, we see that the partition P determines a finite set of (possibly empty) rectangles $\{R_i\}$, and we denote the union of these rectangles by the symbol $\mathcal{C}_P. = \cup_{i=1}^n (x_{i-1}, x_i) \times (c_i, d_i)$.

The area of the rectangle R_i is $(x_i - x_{i-1})(d_i - c_i)$ if $c_i < d_i$ and 0 otherwise. We may write in general that $A(R_i) = (x_i - x_{i-1}) \max((d_i - c_i), 0)$. Define the number A_P by

$$A_P = \sum_{i=1}^n (x_i - x_{i-1})(d_i - c_i).$$

Note that A_P is not exactly the sum of the areas of the rectangles determined by P because it may happen that $d_i < c_i$ for some i's, so that those terms in the sum would be negative. In any case, it is clear that A_P is less than or equal to the sum of the areas of the rectangles, and this notation simplifies matters later.

For any partition P, we have $S \supseteq \mathcal{C}_P$, so that, if $A(S)$ is to denote the area of S, we want to have

$$\begin{aligned}
A(S) &\geq \sum_{i=1}^n A(R_i) \\
&= \sum_{i=1}^n (x_i - x_{i-1}) \max((d_i - c_i), 0) \\
&\geq \sum_{i=1}^n (x_i - x_{i-1})(d_i - c_i) \\
&= A_P.
\end{aligned}$$

DEFINITION. Let S be a geometric set (either open or closed), bounded on the left by $x = a$, on the right by $x = b$, below by the graph of l, and above by the graph of u. Define the *area* $A(S)$ of S by

$$A(S) = \sup_P A_P = \sup_{P=\{x_0 < x_1 < \ldots < x_n\}} \sum_{i=1}^n (x_i - x_{i-1})(d_i - c_i),$$

where the supremum is taken over all partitions P of $[a,b]$, and where the numbers c_i and d_i are as defined above.

Exercise 5.18. (a) Using the notation of the preceding paragraphs, show that each rectangle R_i is a subset of the set S and that $R_i \cap R_j = \emptyset$ if $i \neq j$. It may help to draw a picture of the set S and the rectangles $\{R_i\}$. Can you draw one so that $d_i < c_i$?

(b) Suppose S_1 is a geometric set and that S_2 is another geometric set that is contained in S_1. Prove that $A(S_2) \leq A(S_1)$.
HINT: For each partition P, compare the two A_P's.

Exercise 5.19. Let T be the triangle in the plane with vertices at the three points $(0,0), (0,H)$, and $(B,0)$. Show that the area $A(T)$, as defined above, agrees with the formula $A = (1/2)BH$, where B is the base and H is the height.

The next theorem gives the connection between area (geometry) and integration (analysis). In fact, this theorem is what most calculus students think integration is all about.

THEOREM 5.14. *Let S be a geometric set, i.e., a subset of $\mathbb{R}^2$ that is determined in the above manner by a closed bounded interval $[a,b]$ and two bounding functions l and u. Then*

$$A(S) = \int_a^b (u(x) - l(x))\,dx.$$

PROOF. Let $P = \{x_0 < x_1 < \ldots < x_n\}$ be a partition of $[a,b]$, and let c_i and d_i be defined as above. Let h be a step function that equals d_i on the open interval (x_{i-1}, x_i), and let k be a step function that equals c_i on the open interval (x_{i-1}, x_i). Then on each open interval (x_{i-1}, x_i) we have $h(x) \leq u(x)$ and $k(x) \geq l(x)$. Complete the definitions of h and k by defining them at the partition points so that $h(x_i) = k(x_i)$ for all i. Then we have that $h(x) - k(x) \leq u(x) - l(x)$ for all $x \in [a,b]$. Hence,

$$A_P = \sum_{i=1}^{n} (x_i - x_{i-1})(d_i - c_i) = \int_a^b (h-k) \leq \int_a^b (u - l).$$

Since this is true for every partition P of $[a,b]$, it follows by taking the supremum over all partitions P that

$$A(S) = \sup_P A_P \leq \int_a^b (u(x) - l(x))\,dx,$$

which proves half of the theorem; i.e., that $A(S) \leq \int_a^b u - l$.

To see the other inequality, let h be any step function on $[a,b]$ for which $h(x) \leq u(x)$ for all x, and let k be any step function for which $k(x) \geq l(x)$ for all x. Let $P = \{x_0 < x_1 < \ldots < x_n\}$ be a partition of $[a,b]$ for which both h and k are constant on the open subintervals (x_{i-1}, x_i) of P. Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be the numbers such that $h(x) = a_i$ on (x_{i-1}, x_i) and $k(x) = b_i$ on (x_{i-1}, x_i). It follows, since $h(x) \leq u(x)$ for all x, that $a_i \leq d_i$. Also, it follows that $b_i \geq c_i$. Therefore,

$$\int_a^b (h-k) = \sum_{i=1}^{n} (a_i - b_i)(x_i - x_{i-1}) \leq \sum_{i=1}^{n} (x_i - x_{i-1})(d_i - c_i) = A_P \leq A(S).$$

Finally, let $\{h_m\}$ be a nondecreasing sequence of step functions that converges uniformly to u, and let $\{k_m\}$ be a nonincreasing sequence of step functions that converges uniformly to l. See part (d) of Exercise 5.9. Then

$$\int_a^b (u-l) = \lim_m \int_a^b (h_m - k_m) \leq A(S),$$

which proves the other half of the theorem.

OK! Trumpet fanfares, please!

THEOREM 5.15. ($A = \pi r^2$.) If S is a circle in the plane having radius r, then the area $A(S)$ of S is πr^2.

PROOF. Suppose the center of the circle S is the point (h,k). This circle is a geometric set. In fact, we may describe the circle with center (h,k) and radius r as the subset S of $\mathbb{R}^2$ determined by the closed bounded interval $[h-r, h+r]$ and the functions

$$u(x) = k + \sqrt{r^2 - (x-h)^2}$$

and

$$l(x) = k - \sqrt{r^2 - (x-h)^2}.$$

By the preceding theorem, we then have that

$$A(S) = \int_{h-r}^{h+r} 2\sqrt{r^2 - (x-h)^2}\, dx = \pi r^2.$$

We leave the verification of the last equality to the following exercise.

Exercise 5.20. Evaluate the integral in the above proof:

$$\int_{h-r}^{h+r} 2\sqrt{r^2 - (x-h)^2}\, dx.$$

Be careful to explain each step by referring to theorems and exercises in this book. It may seem like an elementary calculus exercise, but we are justifying each step here.

REMARK. There is another formula for the area of a geometric set that is sometimes very useful. This formula gives the area in terms of a "double integral." There is really nothing new to this formula; it simply makes use of the fact that the number (length) $u(x) - l(x)$ can be represented as the integral from $l(x)$ to $u(x)$ of the constant 1. Here's the formula:

$$A(S) = \int_a^b (\int_{l(x)}^{u(x)} 1\, dy)\, dx.$$

The next theorem is a result that justifies our definition of area by verifying that the whole is equal to the sum of its parts, something that any good definition of area should satisfy.

THEOREM 5.16. *Let S be a closed geometric set, and suppose $S = \cup_{i=1}^n S_i$, where the sets $\{S_i\}$ are closed geometric sets for which $S_i^0 \cap S_j^0 = \emptyset$ if $i \neq j$. Then*

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

PROOF. Suppose S is determined by the interval $[a, b]$ and the two bounding functions l and u, and suppose S_i is determined by the interval $[a_i, b_i]$ and the two bounding functions l_i and u_i. Because $S_i \subseteq S$, it must be that the interval $[a_i, b_i]$ is contained in the interval $[a, b]$. Initially, the bounding functions l_i and u_i are defined and continuous on $[a_i, b_i]$, and we extend their domain to all of $[a, b]$ by defining $l_i(x) = u_i(x) = 0$ for all $x \in [a, b]$ that are not in $[a_i, b_i]$. The extended functions l_i and u_i may not be continuous on all of $[a, b]$, but they are still integrable on $[a, b]$. (Why?) Notice that we now have the formula

$$A(S_i) = \int_{a_i}^{b_i} (u_i(x) - l_i(x))\,dx = \int_a^b (u_i(x) - l_i(x))\,dx.$$

Next, fix an x in the open interval (a, b). We must have that the vertical intervals $(l_i(x), u_i(x))$ and $(l_j(x), u_j(x))$ are disjoint if $i \neq j$. Otherwise, there would exist a point y in both intervals, and this would mean that the point (x, y) would belong to both S_i^0 and S_j^0, which is impossible by hypothesis. Therefore, for each $x \in (a, b)$, the intervals $\{(l_i()x), u_i(x))\}$ are pairwise disjoint open intervals, and they are all contained in the interval $(l(x), u(x))$, because the S_i's are subsets of S. Hence, the sum of the lengths of the open intervals $\{(l_i(x), u_i(x))\}$ is less than or equal to the length of $(l(x), u(x))$. Also, for any point y in the closed interval $[l(x), u(x)]$, the point (x, y) must belong to one of the S_i's, implying that y is in the closed interval $[l_i(x), u_i(x)]$ for some i. But this means that the sum of the lengths of the closed intervals $[l_i(x), u_i(x)]$ is greater than or equal to the length of the interval $[l(x), u(x)]$. Since open intervals and closed intervals have the same length, we then see that $(u(x) - l(x) = \sum_{i=1}^n (u_i(x) - l_i(x))$.

We now have the following calculation:

$$\begin{aligned}
\sum_{i=1}^n A(S_i) &= \sum_{i=1}^n \int_{a_i}^{b_i} (u_i(x) - l_i(x))\,dx \\
&= \sum_{i=1}^n \int_a^b (u_i(x) - l_i(x))\,dx \\
&= \int_a^b \sum_{i=1}^n (u_i(x) - l_i(x))\,dx \\
&= \int_a^b (u(x) - l(x))\,dx \\
&= A(S),
\end{aligned}$$

which completes the proof.

EXTENDING THE DEFINITION OF INTEGRABILITY

We now wish to extend the definition of the integral to a wider class of functions, namely to some that are unbounded and Others whose domains are not closed and bounded intervals. This extended definition is somewhat ad hoc, and these integrals are sometimes called "improper integrals."

DEFINITION. Let f be a real or complex-valued function on the open interval (a,b) where a is possibly $-\infty$ and b is possibly $+\infty$. We say that f is *improperly-integrable* on (a,b) if it is integrable on each closed and bounded subinterval $[a',b'] \subset (a,b)$, and for each point $c \in (a,b)$ we have that the two limits $\lim b' \to b-0 \int_c^{b'} f$ and $\lim_{a'\to a+0} \int_{a'}^c f$ exist.

More generally, We say that a real or complex-valued function f, not necessarily defined on all of the open interval (a,b), is *improperly-integrable* on (a,b) if there exists a partition $\{x_i\}$ of $[a,b]$ such that f is defined and improperly-integrable on each open interval (x_{i-1},x_i).

We denote the set of all functions f that are improperly-integrable on an open interval (a,b) by $I_i((a,b))$.

Analogous definitions are made for a function's being integrable on half-open intervals $[a,b)$ and $(a,b]$.

Note that, in order for f to be improperly-integrable on an open interval, we only require f to be defined at almost all the points of the interval, i.e., at every point except the endpoints of some partition.

Exercise 5.21. (a) Let f be defined and improperly-integrable on the open interval (a,b). Show that $\lim_{a'\to a+0} \int_{a'}^c f + \lim_{b'\to b-0} \int_c^{b'} f$ is the same for all $c \in (a,b)$.

(b) Define a function f on $(0,1)$ by $f(x) = (1-x)^{-1/2}$. Show that f is improperly-integrable on $(0,1)$ and that f is **not** bounded. (Compare this with part (1) of Theorem 5.5.)

(c) Define a function g on $(0,1)$ by $g(x) = (1-x)^{-1}$. Show that g is **not** improperly-integrable on $(0,1)$, and, using part (b), conclude that the product of improperly-integrable functions on $(0,1)$ need not itself be improperly-integrable. (Compare this with part (3) of Theorem 5.5.)

(d) Define h to be the function on $(0,\infty)$ given by $h(x) = 1$ for all x. Show that h is not improperly-integrable on $(0,\infty)$. (Compare this with parts (4) and (5) of Theorem 5.5.)

Part (a) of the preceding exercise is just the consistency condition we need in order to make a definition of the integral of an improperly-integrable function over an open interval.

DEFINITION. Let f be defined and improperly-integrable on an open interval (a,b). We define the *integral* of f over the interval (a,b), and denote it by $\int_a^b f$, by

$$\int_a^b f = \lim_{a'\to a+0} \int_{a'}^c f + \lim_{b'\to b-0} \int_c^{b'} f.$$

In general, if f is improperly-integrable over an open interval, i.e., f is defined and improperly-integrable over each subinterval of (a,b) determined by a partition $\{x_i\}$, then we define the *integral* of f over the interval (a,b) by

$$\int_a^b f = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f.$$

THEOREM 5.17. *Let (a, b) be a fixed open interval (with a possibly equal to $-\infty$ and b possibly equal to $+\infty$), and let $I_i((a, b))$ denote the set of improperly-integrable functions on (a, b). Then:*

(1) *$I_i((a, b))$ is a vector space of functions.*

(2) (Linearity) $\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g$ for all $f, g \in I_i((a, b))$and $\alpha, \beta \in \mathbb{C}$.

(3) (Positivity) If $f(x) \geq 0$ for all $x \in (a, b)$, then $\int_a^b f \geq 0$.

(4) (Order-preserving) If $f, g \in I_i((a, b))$ and $f(x) \leq g(x)$ for all $x \in (a, b)$, then $\int_a^b f \leq \int_a^b g$.

Exercise 5.22. (a) Use Theorems 5.5, 5.6, 5.7, and properties of limits to prove the preceding theorem.

(b) Let f be defined and improperly-integrable on (a, b). Show that, given an $\epsilon > 0$, there exists a $\delta > 0$ such that for any $a < a' < a + \delta$ and any $b - \delta < b' < b$ we have $|\int_a^{a'} f| + |\int_{b'}^b f| < \epsilon$.

(c) Let f be improperly-integrable on an open interval (a, b). Show that, given an $\epsilon > 0$, there exists a $\delta > 0$ such that if (c, d) is any open subinterval of (a, b) for which $d - c < \delta$, then $|\int_c^d f| < \epsilon$.
HINT: Let $\{x_i\}$ be a partition of $[a, b]$ such that f is defined and improperly-integrable on each subinterval (x_{i-1}, x_i). For each i, choose a δ_i using part (b). Now f is bounded by M on all the intervals $[x_{i-1} + \delta_i, x_i - \delta_i]$, so $\delta = \epsilon/M$ should work there.

(d) Suppose f is a continuous function on a closed bounded interval $[a, b]$ and is continuously differentiable on the open interval (a, b). Prove that f' is improperly-integrable on (a, b), and evaluate $\int_a^b f'$.
HINT: Fix a point $c \in (a, b)$, and use the Fundamental Theorem of Calculus to show that the two limits exist.

(e) (Integration by substitution again.) Let $g : [c, d] \to [a, b]$ be continuous on $[c, d]$ and satisfy $g(c) = a$ and $g(d) = b$. Suppose there exists a partition $\{x_0 < x_1 < \ldots < x_n\}$ of the interval $[c, d]$ such that g is continuously differentiable on each subinterval (x_{i-1}, x_i). Prove that g' is improperly-integrable on the open interval (c, d). Show also that if f is continuous on $[a, b]$, we have that

$$\int_a^b f(t)\, dt = \int_c^d f(g(s))g'(s)\, ds.$$

HINT: Integrate over the subintervals (x_{i-1}, x_i), and use part (d).

REMARK. Note that there are parts of Theorems 5.5 and 5.6 that are not asserted in Theorem 5.17. The point is that these other properties do not hold for improperly-integrable functions on open intervals. See the following exercise.

Exercise 5.23. (a) Define f to be the function on $[1, \infty)$ given by $f(x) = (-1)^{n-1}/n$ if $n - 1 \leq x < n$. Show that f is improperly-integrable on $(1, \infty)$, but that $|f|$ is not improperly-integrable on $(1, \infty)$. (Compare this with part (4) of Theorem 5.6.)
HINT: Verify that $\int_1^N f$ is a partial sum of a convergent infinite series, and then verify that $\int_1^N |f|$ is a partial sum of a divergent infinite series.

(b) Define the function f on $(1,\infty)$ by $f(x) = 1/x$. For each positive integer n, define the function f_n on $(1,\infty)$ by $f_n(x) = 1/x$ if $1 < x < n$ and $f_n(x) = 0$ otherwise. Show that each f_n is improperly-integrable on $(1,\infty)$, that f is the uniform limit of the sequence $\{f_n\}$, but that f is not improperly-integrable on $(1,\infty)$. (Compare this with part (5) of Theorem 5.6.)

(c) Suppose f is a nonnegative real-valued function on the half-open interval (a,∞) that is integrable on every closed bounded subinterval $[a,b']$. For each positive integer $n \geq a$, define $y_n = \int_a^n f(x)\,dx$. Prove that f is improperly-integrable on $[a,\infty)$ if and only if the sequence $\{y_n\}$ is convergent. In that case, show that $\int_a^\infty f = \lim y_n$.

We are now able to prove an important result relating integrals over infinite intervals and convergence of infinite series.

THEOREM 5.18. *Let f be a positive function on $[1,\infty)$, assume that f is integrable on every closed bounded interval $[1,b]$, and suppose that f is nonincreasing; i.e., if $x < y$ then $f(x) \geq f(y)$. For each positive integer i, set $a_i = f(i)$, and let S_N denote the Nth partial sum of the infinite series $\sum a_i : S_N = \sum_{i=1}^N a_i$. Then:*

(1) *For each N, we have*
$$S_N - a_1 \leq \int_1^N f(x)\,dx \leq S_{N-1}.$$

(2) *For each N, we have that*
$$S_{N-1} - \int_1^N f(x)\,dx \leq a_1 - a_N \leq a_1;$$
i.e., the sequence $\{S_{N-1} - \int_1^N f\}$ is bounded above.

(3) *The sequence $\{S_{N-1} - \int_1^N f\}$ is nondecreasing.*

(4) (Integral Test) The infinite series $\sum a_i$ converges if and only if the function f is improperly-integrable on $(1,\infty)$.

PROOF. For each positive integer N, define a step function k_N on the interval $[1,N]$ as follows. Let $P = \{x_0 < x_1 < \ldots < x_{N-1}\}$ be the partition of $[1,N]$ given by the points $\{1 < 2 < 3 < \ldots < N\}$, i.e., $x_i = i+1$. Define $k_N(x)$ to be the constant $c_i = f(i+1)$ on the interval $[x_{i-1},x_i) = [i,i+1)$. Complete the definition of k_N by setting $k_N(N) = f(N)$. Then, because f is nonincreasing, we have that $k_N(x) \leq f(x)$ for all $x \in [1,N]$. Also,
$$\begin{aligned}\int_1^N k_N &= \sum_{i=1}^{N-1} c_i(x_i - x_{i-1}) \\ &= \sum_{i=1}^{N-1} f(i+1) \\ &= \sum_{i=2}^{N} f(i) \\ &= \sum_{i=2}^{N} a_i \\ &= S_N - a_1,\end{aligned}$$

which then implies that

$$S_N - a_1 = \int_1^N k_N(x)\,dx \le \int_1^N f(x)\,dx.$$

This proves half of part (1).

For each positive integer $N > 1$ define another step function l_N, using the same partition P as above, by setting $l_N(x) = f(i)$ if $i \le x < i+1$ for $1 \le i < N$, and complete the definition of l_N by setting $l_N(N) = f(N)$. Again, because f is nonincreasing, we have that $f(x) \le l_N(x)$ for all $x \in [1, N]$. Also

$$\begin{aligned}\int_1^N l_N &= \sum_{i=1}^{N-1} f(i)\\ &= \sum_{i=1}^{N-1} a_i\\ &= S_{N-1},\end{aligned}$$

which then implies that

$$\int_1^N f(x)\,dx \le \int_1^N l_N(x)\,dx = S_{N-1},$$

and this proves the other half of part (1).

It follows from part (1) that

$$S_{N-1} - \int_1^N f(x)\,dx \le S_{N-1} - S_N + a_1 = a_1 - a_N,$$

and this proves part (2).

We see that the sequence $\{S_{N-1} - \int_1^N f\}$ is nondecreasing by observing that

$$\begin{aligned}S_N - \int_1^{N+1} f - S_{N-1} + \int_1^N f &= a_N - \int_N^{N+1} f\\ &= f(N) - \int_N^{N+1} f\\ &\ge 0,\end{aligned}$$

because f is nonincreasing.

Finally, to prove part (4), note that both of the sequences $\{S_N\}$ and $\{\int_1^N f\}$ are nondecreasing. If f is improperly-integrable on $[1,\infty)$, then $\lim_N \int_1^N f$ exists, and $S_N \le a_1 + \int_1^\infty f(x)\,dx$ for all N, which implies that $\sum a_i$ converges by Theorem 2.14. Conversely, if $\sum a_i$ converges, then $\lim S_N$ exists. Since $\int_1^N f(x)\,dx \le S_{N-1}$, it then follows, again from Theorem 2.14, that $\lim_N \int_1^N f(x)\,dx$ exists. So, by the preceding exercise, f is improperly-integrable on $[1,\infty)$.

We may now resolve a question first raised in Exercise 2.32. That is, for $1 < s < 2$, is the infinite series $\sum 1/n^s$ convergent or divergent? We saw in that exercise that this series is convergent if s is a rational number.

Exercise 5.24. (a) Let s be a real number. Use the Integral Test to prove that the infinite series $\sum 1/n^s$ is convergent if and only if $s > 1$.

(b) Let s be a complex number $s = a + bi$. Prove that the infinite series $\sum 1/n^s$ is absolutely convergent if and only if $a > 1$.

Exercise 5.25. Let f be the function on $[1, \infty)$ defined by $f(x) = 1/x$.

(a) Use Theorem 5.18 to prove that the sequence $\{\sum_{i=1}^N \frac{1}{i} - \ln N\}$ converges to a positive number $\gamma \le 1$. (This number γ is called Euler's constant.)
HINT: Show that this sequence is bounded above and nondecreasing.

(b) Prove that

$$\sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} = \ln 2.$$

HINT: Write S_{2N} for the $2N$th partial sum of the series. Use the fact that

$$S_{2N} = \sum_{i=1}^{2N} \frac{1}{i} - 2\sum_{i=1}^{N} \frac{1}{2i}.$$

Now add and subtract $\ln(2N)$ and use part (a).

INTEGRATION IN THE PLANE

Let S be a closed geometric set in the plane. If f is a real-valued function on S, we would like to define what it means for f to be "integrable" and then what the "integral" of f is. To do this, we will simply mimic our development for integration of functions on a closed interval $[a, b]$.

So, what should be a "step function" in this context? That is, what should is a "partition" of S be in this context? Presumably a step function is going to be a function that is constant on the "elements" of a partition. Our idea is to replace the subintervals determined by a partition of the interval $[a, b]$ by geometric subsets of the geometric set S.

DEFINITION. The *overlap* of two geometric sets S_1 and S_2 is defined to be the interior $(S_1 \cap S_2)^0$ of their intersection. S_1 and S_2 are called *nonoverlapping* if this overlap $(S_1 \cap S_2)^0$ is the empty set.

DEFINITION. A *partition* of a closed geometric set S in $\mathbb{R}^2$ is a finite collection $\{S_1, S_2, \dots, S_n\}$ of nonoverlapping closed geometric sets for which $\cup_{i=1}^n S_i = S$; i.e., the union of the S_i's is all of the geometric set S.

The open subsets $\{S_i^0\}$ are called the *elements* of the partition.

A *step function* on the closed geometric set S is a real-valued function h on S for which there exists a partition $P = \{S_i\}$ of S such that $h(z) = a_i$ for all $z \in S_i^0$; i.e., h is constant on each element of the partition P.

REMARK. One example of a partition of a geometric set, though not at all the most general kind, is the following. Suppose the geometric set S is determined by the interval $[a, b]$ and the two bounding functions u and l. Let $\{x_0 < x_1 < \ldots < x_n\}$ be a partition of the interval $[a, b]$. We make a partition $\{S_i\}$ of S by constructing vertical lines at the points x_i from $l(x_i)$ to $u(x_i)$. Then S_i is the geometric set determined by the interval $[x_{i-1}, x_i]$ and the two bounding functions u_i and l_i that are the restrictions of u and l to the interval $[x_{i-1}, x_i]$.

A step function is constant on the open geometric sets that form the elements of some partition. We say nothing about the values of h on the "boundaries" of these geometric sets. For a step function h on an interval $[a, b]$, we do not worry about the finitely many values of h at the endpoints of the subintervals. However, in the plane, we are ignoring the values on the boundaries, which are infinite sets. As a consequence, a step function on a geometric set may very well have an infinite range, and may not even be a bounded function, unlike the case for a step function on an interval. The idea is that the boundaries of geometric sets are "negligible" sets as far as area is concerned, so that the values of a function on these boundaries shouldn't affect the integral (average value) of the function.

Before continuing our development of the integral of functions in the plane, we digress to present an analog of Theorem 3.20 to functions that are continuous on a closed geometric set.

THEOREM 5.19. *Let f be a continuous real-valued function whose domain is a closed geometric set S. Then there exists a sequence $\{h_n\}$ of step functions on S that converges uniformly to f.*

PROOF. As in the proof of Theorem 3.20, we use the fact that a continuous function on a compact set is uniformly continuous.

For each positive integer n, let δ_n be a positive number satisfying $|f(z)-f(w)| < 1/n$ if $|z-w| < \delta_n$. Such a δ_n exists by the uniform continuity of f on S. Because S is compact, it is bounded, and we let $R = [a, b] \times [c, d]$ be a closed rectangle that contains S. We construct a partition $\{S_i^n\}$ of S as follows. In a checkerboard fashion, we write R as the union $\cup R_i^n$ of small, closed rectangles satisfying

(1) If z and w are in R_i^n, then $|z - w| < \delta_n$. (The rectangles are that small.)

(2) $R_i^{n\,0} \cap R_j^{n\,0} = \emptyset$. (The interiors of these small rectangles are disjoint.)

Now define $S_i^n = S \cap R_i^n$. Then $S_i^{n\,0} \cap S_j^{n\,0} = \emptyset$, and $S = \cup S_i^n$. Hence, $\{S_i^n\}$ is a partition of S.

For each i, choose a point z_i^n in S_i^n, and set $a_i^n = f(z_i^n)$. We define a step function h_n as follows: If z belongs to one (and of course only one) of the open geometric sets $S_i^{n\,0}$, set $h_n(z) = a_i^n$. And, if z does not belong to any of the open geometric sets $S_i^{n\,0}$, set $h_n(z) = f(z)$. It follows immediately that h_n is a step function.

Now, we claim that $|f(z) - h_n(z)| < 1/n$ for all $z \in S$. For any z in one of the $S_i^{n\,0}$'s, we have

$$|f(z) - h_n(z)| = |f_z) - a_i^n| = |f(z) - f(z_I^n)| < 1/n$$

because $|z - z_i^n| < \delta_n$. And, for any z not in any of the $S_i^{n\,0}$'s, $f(z) - h_n(z) = 0$. So, we have defined a sequence $\{h_n\}$ of step functions on S, and the sequence $\{h_n\}$ converges uniformly to f by Exercise 3.29.

What follows now should be expected. We will define the integral of a step function h over a geometric set S by

$$\int_S h = \sum_{i=1}^n a_i \times A(S_i).$$

We will define a function f on S to be integrable if it is the uniform limit of a sequence $\{h_n\}$ of step functions, and we will then define the integral of f by

$$\int_S f = \lim \int_S h_n.$$

Everything should work out nicely. Of course, we have to check the same two consistency questions we had for the definition of the integral on $[a,b]$, i.e., the analogs of Theorems 5.1 and 5.4.

THEOREM 5.20. *Let S be a closed geometric set, and let h be a step function on S. Suppose $P = \{S_1, \ldots, S_n\}$ and $Q = \{T_1, \ldots, T_m\}$ are two partitions of S for which $h(z)$ is the constant a_i on S_i^0 and $h(z)$ is the constant b_j on T_j^0. Then*

$$\sum_{i=1}^{n} a_i A(S_i) = \sum_{j=1}^{m} b_j A(T_j).$$

PROOF. We know by part (d) of Exercise 5.17 that the intersection of two geometric sets is itself a geometric set. Also, for each fixed index j, we know that the sets $\{T_j \cap S_i^0\}$ are pairwise disjoint. Then, by Theorem 5.16, we have that $A(T_j) = \sum_{i=1}^{n} A(T_j \cap S_i)$. Similarly, for each fixed i, we have that $A(S_i = \sum_{j=1}^{m} A(T_j \cap S_i)$. Finally, for each pair i and j, for which the set $T_j^0 \cap S_i^0$ is not empty, choose a point $z_{i,j} \in T_j^0 \cap S_i^0$, and note that $a_i = h(z_{i,j}) = b_j$, because $z_{i,j}$ belongs to both S_i^0 and T_j^0.

With these observations, we then have that

$$\begin{aligned}
\sum_{i=1}^{n} a_i A(S_i) &= \sum_{i=1}^{n} a_i \sum_{j=1}^{m} A(T_j \cap S_i) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i A(T_j \cap S_i) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} h(z_{i,j}) A(T_j \cap S_i) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} b_j A(T_j \cap S_i) \\
&= \sum_{j=1}^{m} \sum_{i=1}^{n} b_j A(T_j \cap S_i) \\
&= \sum_{j=1}^{m} b_j \sum_{i=1}^{n} A(T_j \cap S_i) \\
&= \sum_{j=1}^{m} b_j A(T_j),
\end{aligned}$$

which completes the proof.

OK, the first consistency condition is satisfied. Moving right along:

DEFINITION. Let h be a step function on a closed geometric set S. Define the *integral* of h over the geometric set S by the formula

$$\int_S h = \int_S H(z)\,dz = \sum_{i=1}^{n} a_i A(S_i),$$

where $S_1, \dots, S_n$ is a partition of S for which h is the constant a_i on the interior S_i^0 of the set S_i.

Just as in the case of integration on an interval, before checking the second consistency result, we need to establish the following properties of the integral of step functions.

THEOREM 5.21. *Let $H(S)$ denote the vector space of all step functions on the closed geometric set S. Then the assignment $h \to \int h$ of $H(S)$ into $\mathbb{R}$ has the following properties:*

(1) (Linearity) $H(S)$ is a vector space, and $\int_S (h_1 + h_2) = \int_S h_1 + \int_S h_2$, and $\int_S ch = c \int_S h$ for all $h_1, h_2, h \in H(S)$, and for all real numbers c.
(2) If $h = \sum_{i=1}^n c_i \chi_{S_i}$ is a linear combination of indicator functions of geometric sets that are subsets of S, then $\int h = \sum_{i=1}^n c_i A(S_i)$.
(3) (Positivity) If $h(z) \geq 0$ for all $z \in S$, then $\int_S h \geq 0$.
(4) (Order-preserving) If h_1 and h_2 are step functions on S for which $h_1(z) \leq h_2(z)$ for all $z \in S$, then $\int_S h_1 \leq \int_S h_2$.

PROOF. Suppose h_1 is constant on the elements of a partition $P = \{S_i\}$ and h_2 is constant on the elements of a partition $Q = \{T_j\}$. Let V be the partition of the geometric set S whose elements are the sets $\{U_k\} = \{S_i^0 \cap T_j^0\}$. Then both h_1 and h_2 are constant on the elements U_k of V, so that $h_1 + h_2$ is also constant on these elements. Therefore, $h_1 + h_2$ is a step function, and

$$\int (h_1 + h_2) = \sum_k (a_k + b_k) A(U_k) = \sum_k a_k A(U_k) + \sum_k b_k A(U_k) = \int h_1 + \int h_2,$$

and this proves the first assertion of part (1).

The proof of the other half of part (1), as well as parts (2), (3), and (4), are totally analogous to the proofs of the corresponding parts of Theorem 5.2, and we omit the arguments here.

Now for the other necessary consistency condition:

THEOREM 5.22. *let S be a closed geometric set in the plane.*

(1) *If $\{h_n\}$ is a sequence of step functions that converges uniformly to a function f on S, then the sequence $\{\int_S h_n\}$ is a convergent sequence of real numbers.*
(2) *If $\{h_n\}$ and $\{k_n\}$ are two sequences of step functions on S that converge uniformly to the same function f, then*

$$\lim \int_S h_n = \lim \int_S k_n.$$

Exercise 5.26. Prove Theorem 5.22. Mimic the proofs of Theorems 5.3 and 5.4.

DEFINITION. If f is a real-valued function on a closed geometric set S in the plane, then f is *integrable on S* if it is the uniform limit of a sequence $\{h_n\}$ of step functions on S.

We define the *integral* of an integrable function f on S by

$$\int_S f \equiv \int_S f(z)\, dz = \lim \int_S h_n,$$

where $\{h_n\}$ is a sequence of step functions on S that converges uniformly to f.

THEOREM 5.23. *Let S be a closed geometric set in the plane, and let $I(S)$ denote the set of integrable functions on S. Then:*

(1) *$I(S)$ is a vector space of functions.*
(2) *If f and $g \in I(S)$, and one of them is bounded, then $fg \in I(S)$.*
(3) *Every step function is in $I(S)$.*
(4) *If f is a continuous real-valued function on S, then f is in $I(S)$. That is, every continuous real-valued function on S is integrable on S.*

Exercise 5.27. (a) Prove Theorem 5.23. Note that this theorem is the analog of Theorem 5.5, but that some things are missing.

(b) Show that integrable functions on S are not necessarily bounded; not even step functions have to be bounded.

(c) Show that, if $f \in I(S)$, and g is a function on S for which $f(x,y) = g(x,y)$ for all (x,y) in the interior S^0 of S, then $g \in I(S)$. That is, integrable functions on S can do whatever they like on the boundary.

THEOREM 5.24. *Let S be a closed geometric set. The assignment $f \to \int f$ on $I(S)$ satisfies the following properties.*

(1) (Linearity) $I(S)$ is a vector space, and $\int_S(\alpha f + \beta g) = \alpha \int_S f + \beta \int_S g$ for all $f, g \in I(S)$and $\alpha, \beta \in \mathbb{R}$.
(2) (Positivity) If $f(z) \geq 0$ for all $z \in S$, then $\int_S f \geq 0$.
(3) (Order-preserving) If $f, g \in I(S)$ and $f(z) \leq g(z)$ for all $z \in S$, then $\int_S f \leq \int_S g$.
(4) If $f \in I(S)$, then so is $|f|$, and $|\int_S f| \leq \int_S |f|$.
(5) If f is the uniform limit of functions f_n, each of which is in $I(S)$, then $f \in I(S)$ and $\int_S f = \lim \int_S f_n$.
(6) Let $\{u_n\}$ be a sequence of functions in $I(S)$, and suppose that for each n there is a number m_n, for which $|u_n(z)| \leq m_n$ for all $z \in S$, and such that the infinite series $\sum m_n$ converges. Then the infinite series $\sum u_n$ converges uniformly to an integrable function, and $\int_S \sum u_n = \sum \int_S u_n$.
(7) If $f \in I(S)$, and $\{S_1, \ldots, S_n\}$ is a partition of S, then $f \in I(S_i)$ for all i, and
$$\int_S = \sum_{i=1}^{n} \int_{S_i} f.$$

Exercise 5.28. Prove Theorem 5.24. It is mostly the analog to Theorem 5.6. To see the last part, let h_i be the step function that is identically 1 on S_i; check that $h_i f \in I(S_i)$; then examine $\sum_i \int_S f h_i$.

Of course, we could now extend the notion of integrability over a geometric set S to include complex-valued functions just as we did for integrability over an interval $[a,b]$. However, real-valued functions on geometric sets will suffice for the purposes of this book.

We include here, to be used later in Chapter VII, a somewhat technical theorem about constructing partitions of a geometric set.

THEOREM 5.25. *Let $S_1, \ldots, S_n$ be closed, nonoverlapping, geometric sets, all contained in a geometric set S. Then there exists a partition $\widehat{S}_1, \ldots, \widehat{S}_M$ of S such* *that for $1 \le i \le n$ we have $S_i = \widehat{S}_i$. In other words, the s_i's are the first n elements of a partition of S.*

PROOF. Suppose S is determined by the interval $[a, b]$ and the two bounding functions u and l. We prove this theorem by induction on n.

If $n = 1$, let S_1 be determined by the interval $[a_1, b_1]$ and the two bounding functions u_1 and l_1. Set $\widehat{S}_1 = S_1$, and define four more geometric sets $\widehat{S}_2, \ldots, \widehat{S}_5$ as follows:

(1) $\widehat{S}_2$ is determined by the interval $[a, a_1]$ and the two bounding functions u and l restricted to that interval.
(2) S_3 is determined by the interval $[a_1, b_1]$ and the two bounding functions u and u_1 restricted to that interval.
(3) S_4 is determined by the interval $[a_1, b_1]$ and the two bounding functions l and l_1 restricted to that interval.
(4) $\widehat{S}_5$ is determined by the interval $[b_1, b]$ and the two bounding functions u and l restricted to that interval.

Observe that the five sets $\widehat{S}_1, \widehat{S}_2, \ldots, \widehat{S}_5$ constitute a partition of the geometric set S, proving the theorem in the case $n = 1$.

Suppose next that the theorem is true for any collection of n sets satisfying the hypotheses. Then, given $S_1, \ldots, S_{n+1}$ as in the hypothesis of the theorem, apply the inductive hypothesis to the n sets $S_1, \ldots, S_n$ to obtain a partition $T_1, \ldots, T_m$ of S for which $T_i = S_i$ for all $1 \le i \le n$. For each $n + 1 \le i \le m$, consider the geometric set $S'_i = S_{n+1} \cap T_i$ of the geometric set T_i. We may apply the case $n = 1$ of this theorem to this geometric set to conclude that S'_i is the first element $S'_{i,1}$ of a partition $\{S'_{i,1}, S'_{i,2}, \ldots, S'_{i,m_i}\}$ of the geometric set T_i.

Define a partition $\{\widehat{S}_k\}$ of S as follows: For $1 \le k \le n$, set $\widehat{S}_k = T_k$. Set $\widehat{S}_{n+1} = \cup_{i=n+1}^{m} S'_{i,1} = S_{n+1}$. And define the rest of the partition $\{\widehat{S}_k\}$ to be made up of the remaining sets $S'_{i,j}$ for $n + 1 \le i \le m$ and $2 \le j \le m_i$. It follows directly that this partition $\{\widehat{S}_k\}$ satisfies the requirements of the theorem.

Exercise 5.29. Let $S_1, \ldots, S_n$ be as in the preceding theorem. Suppose S_k is determined by the interval $[a_k, b_k]$ and the two bounding functions u_k and l_k. We will say that S_k is "below" S_j, equivalently S_j is "above" S_k, if there exists a point x such that $u_k(x) < l_j(x)$. Note that this implies that $x \in [a_k, b_k] \cap [a_j, b_j]$.

(a) Suppose S_k is below S_j, and suppose $(z, y_k) \in S_k$ and $(z, y_j) \in S_j$. Show that $y_j > y_k$. That is, if S_k is below S_j, then no part of S_k can be above S_j.

(b) Suppose S_2 is below S_1 and S_3 is below S_2. Show that no part of S_3 can be above S_1.
HINT: By way of contradiction, let $x_1 \in [a_1, b_1]$ be such that $u_2(x_1) < l_1(x_1)$; let $x_2 \in [a_2, b_2]$ be such that $u_3(x_2) < l_2(x_2)$; and suppose $x_3 \in [a_3, b_3]$ is such that $u_1(x_3) < l_3(x_3)$. Derive contradictions for all possible arrangements of the three points x_1, x_2, and x_3.

(c) Prove that there exists an index k_0 such that S_{k_0} is minimal in the sense that there is no other S_j that is below S_{k_0}.
HINT: Argue by induction on n. Thus, let $\{T_l\}$ be the collection of all S_k's that are below S_1, and note that there are at most $n-1$ elements of $\{T_l\}$. By induction, there is one of the T_l's, i.e., an S_{k_0} that is minimal for that collection. Now, using part (b), show that this S_{k_0} must be minimal for the original collection.

There is one more concept about integrating over geometric sets that we will need in later chapters. We have only considered sets that are bounded on the left and right by straight vertical lines and along the top and bottom by graphs of continuous functions $y = u(x)$ and $y = l(x)$. We equally well could have discussed sets that are bounded above and below by straight horizontal lines and bounded on the left and right by graphs of continuous functions $x = l(y)$ and $x = r(y)$. These additional sets do not provide anything particularly important, so we do not discuss them. However, there are times when it is helpful to work with geometric sets with the roles of horizontal and vertical reversed. We accomplish this with the following definition.

DEFINITION. Let S be a subset of $\mathbb{R}^2$. By the *symmetric image* of S we mean the set $\widehat{S}$ of all points $(x, y) \in \mathbb{R}^2$ for which the point $(y, x) \in S$.

The symmetric image of a set is just the reflection of the set across the $y = x$ line in the plane. Note that the symmetric image of the rectangle $[a, b] \times [c, d]$ is again a rectangle, $[c, d] \times [a, b]$, and therefore the area of a rectangle is equal to the area of its symmetric image. This has the implication that if the symmetric image of a geometric set is also a geometric set, then they both have the same area. The symmetric image of a geometric set doesn't have to be a geometric set itself. For instance, consider the examples suggested in part (b) of Exercise 5.17. But clearly rectangles, triangles, and circles have this property, for their symmetric images are again rectangles, triangles, and circles. For a geometric set, whose symmetric image is again a geometric set, there are some additional computational properties of the area of S as well as the integral of functions over S, and we present them in the following exercises.

Exercise 5.30. Suppose S is a closed geometric set, which is determined by a closed interval $[a, b]$ and two bounding functions $u(x)$ and $l(x)$. Suppose the symmetric image $\widehat{S}$ of S is also a closed geometric set, determined by an interval $[\widehat{a}, \widehat{b}]$ and two bounding functions $\widehat{u}(x)$ and $\widehat{l}(x)$.

(a) Make up an example to show that the numbers $\widehat{a}$ and $\widehat{b}$ need not have anything to do with the numbers a and b, and that the functions $\widehat{u}$ and $\widehat{l}$ need not have anything to do with the functions u and l.

(b) Prove that S and $\widehat{S}$ have the same area.
HINT: use the fact that the area of a geometric set is approximately equal to the sum of the areas of certain rectangles, and then use the fact that the area of the symmetric image of a rectangle is the same as the area of the rectangle.

(c) Show that for every point $(x, y) \in S$, we must have $\widehat{a} \le y \le \widehat{b}$, and for every such y, we must have $\widehat{l}(y) \le x \le \widehat{u}(y)$.
HINT: If $(x, y) \in S$, then $(y, x) \in \widehat{S}$.

(d) Prove that the area $A(S)$ of S is given by the formula

$$A(S) = \int_a^b \int_{l(x)}^{u(x)} 1\, dydx = \int_{\widehat{a}}^{\widehat{b}} \int_{\widehat{l}(y)}^{\widehat{u}(y)} 1\, dxdy.$$

(See the remark preceding Theorem 5.16.)

(e) Let S be the right triangle having vertices (a, c), (b, c), and (b, d), where $d > c$. Describe the symmetric image of S; i.e., find the corresponding $\widehat{a}, \widehat{b}, \widehat{u}$, and $\widehat{l}$. Use part (d) to obtain the following formulas for the area of S :

$$A(S) = \int_a^b \int_c^{d+\frac{t-b}{a-b}(c-d)} 1\, dsdt = \int_c^d \int_{b+\frac{t-d}{c-d}(a-b)} 1\, dsdt.$$

Exercise 5.31. (a) Prove that if S_1 and S_2 are geometric sets whose symmetric images are again geometric sets, then the symmetric image of the geometric set $S_1 \cap S_2$ is also a geometric set.

(b) Suppose T is a closed geometric set that is contained in a closed geometric set S. Assume that both the symmetric images $\widehat{T}$ and $\widehat{S}$ are also geometric sets. If $\widehat{S}$ is determined by an interval $[\widehat{a}, \widehat{b}]$ and two bounding functions $\widehat{u}$ and $\widehat{l}$, prove that

$$A(T) = \int_{\widehat{a}}^{\widehat{b}} \int_{\widehat{l}(s)}^{\widehat{u}(s)} \chi_T(t,s)\, dtds,$$

where χ_T is the indicator function of the set T; i.e., $\chi_T(t,s) = 1$ if $(t,s) \in T$, and $\chi_T(t,s) = 0$ if $(t,s) \notin T$.
HINT: See the proof of Theorem 5.16, give names to all the intervals and bounding functions, and in the end use part (d) of the preceding exercise.

(c) Suppose $\{S_i\}$ is a partition of a geometric set S, and suppose the symmetric images of S and all the S_i's are also geometric sets. Suppose h is a step function that is the constant a_i on the element S_i^0 of the partition $\{S_i\}$. Prove that $\int_S h = \sum_{i=1}^n a_i \int_S \chi_{S_i^0}$, and therefore that

$$\int_S h = \int_a^b \int_{l(t)}^{u(t)} h(t,s)\, dsdt = \int_{\widehat{a}}^{\widehat{b}} \int_{\widehat{l}(s)}^{\widehat{u}(s)} h(t,s)\, dtds.$$

HINT: Use part (b).

(d) Let S be a geometric set whose symmetric image $\widehat{S}$ is also a geometric set, and suppose f is a continuous function on S. Show that

$$\int_S f = \int_a^b \int_{l(t)}^{u(t)} f(t,s)\, dsdt = \int_{\widehat{a}}^{\widehat{b}} \int_{\widehat{l}(s)}^{\widehat{u}(s)} f(t,s)\, dtds.$$

HINT: Make use of the fact that the step functions constructed in Theorem 5.19 satisfy the assumptions of part (c). Then take limits.

(e) Let S be the triangle in part (e) of the preceding exercise. If f is a continuous function on S, show that the integral of f over S is given by the formulas

$$\int_S f = \int_a^b \int_c^{d+\frac{t-b}{a-b}(c-d)} f(t,s)\, dsdt = \int_c^d \int_{b+\frac{s-d}{c-d}(a-b)} f(t,s)\, dtds.$$