

## Chapter 10

# Integration in Several Variables

Integration theory for functions of several variables has much in common with integration for functions of a single variable. Many of the proofs are almost identical. However, there are some fundamental differences.

In one variable, we only have to worry about integrating over an interval. However, in several variables the sets we integrate over can be much more complicated. There are issues concerning the boundary of the set and how large it can be. Such issues don't arise in the theory of integration of a function of one variable. In one variable, the change of variable formula for integration (the substitution formula) is quite simple and has a simple proof – it follows directly from the chain rule for differentiation and the Fundamental Theorem of Calculus. The analogous formula in several variables is much more complicated – it involves the determinant of the differential of the change of variables transformation. Its proof is long and complicated.

We begin with a definition of the integral of a function over a multidimensional rectangle.

### 10.1 Integration over a Rectangle

An *aligned rectangle* in  $\mathbb{R}^d$  is a set of the form

$$R = [a_1, b_1] \times \cdots \times [a_d, b_d] = \{(x_1, \dots, x_d) \in \mathbb{R}^d : a_k \leq x_k \leq b_k, k = 1 \cdots d\}.$$

We call such a rectangle *aligned* because each of its edges is parallel to a coordinate axis. Unless otherwise specified, in this chapter the term *rectangle* will mean *aligned rectangle*.

The *d-volume* of a rectangle is the product of the lengths of its edges – that

is, the  $d$ -volume  $V(R)$  of the rectangle  $R$  above is

$$V(R) = \prod_{k=1}^d (b_k - a_k).$$

Thus, the 1-volume of a rectangle (an interval) in  $\mathbb{R}$  is its length; the 2-volume of a rectangle in  $\mathbb{R}^2$  is its area. The 3-volume of a rectangle in  $\mathbb{R}^3$  is its ordinary volume.

Note that it is possible for one of the intervals  $[a_k, b_k]$  defining a rectangle in  $\mathbb{R}^d$  to be degenerate – that is, it could be that  $a_k = b_k$ . In this case, the rectangle has  $d$ -volume 0. This makes sense, because it is actually a rectangle of dimension  $d - 1$  in this case.

As long as the dimension of the ambient space  $\mathbb{R}^d$  is understood, we will drop the  $d$  and just refer to the  $d$ -volume of a rectangle as its volume.

An *aligned partition*  $P$  of an aligned rectangle  $R = [a_1, b_1] \times \cdots \times [a_d, b_d]$  is a partition

$$\{a_k = x_{0,k} \leq x_{1,k} \leq \cdots \leq x_{d,k} = b_k\}$$

of each of the intervals  $[a_k, b_k]$ . Such a thing divides  $R$  up into subrectangles of the form

$$\begin{aligned} & [x_{j_1-1,1}, x_{j_1,1}] \times \cdots \times [x_{j_d-1,d}, x_{j_d,d}] \\ & = \{(x_1, \cdots, x_d) \in \mathbb{R}^d : x_{j_k-1,k} \leq x_k \leq x_{j_k,k}, k = 1 \cdots d\}. \end{aligned}$$

Each of these will be called a subrectangle for the partition  $P$  of the rectangle  $R$ . If  $n$  is the number of subrectangles for  $P$ , then we will number these subrectangles in some fashion so that we have a list  $\{R_1, R_2, \cdots, R_n\}$  of all the subrectangles for  $P$ . We will not attempt to arrange this numbering scheme in a way that has anything to do with the indexing of the points in the corresponding partitions of the individual intervals  $[a_k, b_k]$ . To do so would lead to an awful mess.

Note that  $R$  is the union of the subrectangles determined by a partition of  $R$  and any two of these subrectangles are either disjoint or have a lower dimensional rectangle as intersection. The volume of  $R$  is the sum of the volumes of the subrectangles determined by the partition.

Unless otherwise specified, in this chapter, the term *partition* will mean *aligned partition*.

## Upper and Lower Sums

Let  $f$  be a bounded real valued function defined on a rectangle  $R$  and let  $P$  be a partition of  $R$  determining a list of subrectangles  $R_1, R_2, \cdots, R_n$ ,

**Definition 10.1.1.** If  $f$ ,  $R$ ,  $P$ , and  $\{R_1, R_2, \cdots, R_n\}$  are as above, then we

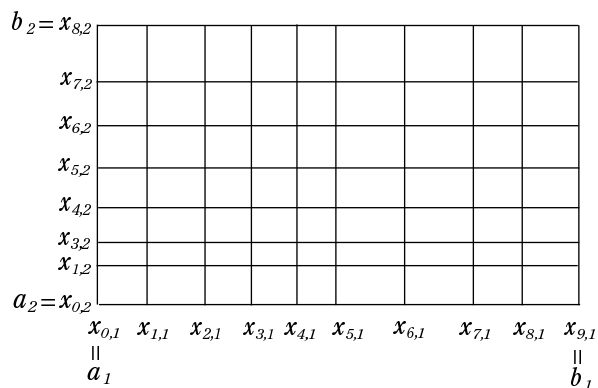


Figure 10.1: Partition of a Rectangle

define the *upper and lower sums* for  $f$  and  $P$  by

$$\begin{aligned}
 U(f, P) &= \sum_{j=1}^n M_j V(R_j), \\
 L(f, P) &= \sum_{j=1}^n m_j V(R_j),
 \end{aligned}
 \tag{10.1.1}$$

where

$$M_j = \sup_{R_j} f \quad \text{and} \quad m_j = \inf_{R_j} f.$$

This is exactly the way we defined the upper and lower sums for  $f$  and the partition  $P$  in Definition 5.1.1, except there we were partitioning intervals into subintervals and here we are partitioning  $d$ -dimensional rectangles into subrectangles.

As in Section 5.1, a *Riemann Sum* for  $f$  and  $P$  on  $R$  is a sum of the form

$$\sum_{j=1}^n f(u_j) V(R_j) \tag{10.1.2}$$

where, for each  $j$ ,  $u_j$  is some point in the rectangle  $R_j$ . For each  $j$ , the term  $f(u_j)V(R_j)$  represents the volume (or minus the volume, if  $f(u_j) < 0$ ) of a  $d+1$ -dimensional rectangle with base  $R_j$  and with height  $|f(u_j)|$ . Now, for each  $j$  we have

$$m_j \leq f(u_j) \leq M_j,$$

which implies

$$L(f, P) \leq \sum_{j=1}^n f(u_j) V(R_j) \leq U(f, P).$$

Thus, as in Section 5.1, every Riemann sum for  $f$  and  $P$  lies between the lower and upper sums for  $f$  and  $P$ .

### Refinement

If  $R$  is a rectangle in  $\mathbb{R}^d$ , and  $P$  and  $Q$  partitions of  $R$ , then  $Q$  is said to be a *refinement* of  $P$  if every subrectangle of  $R$  determined by  $Q$  is a subset of some subrectangle determined by  $P$ .

If  $R = [a_1, b_1] \times \cdots \times [a_d, b_d]$ , then the partition  $P$  consists of a partition of each of the intervals  $[a_k, b_k]$ , as does the partition  $Q$ . It is not difficult to see that  $Q$  is a refinement of  $P$  if and only if, for  $k = 1, \dots, d$ , the partition of  $[a_k, b_k]$  determined by  $Q$  is a refinement of the partition of this same interval determined by  $P$ . For this reason, it is also easy to see that any two partitions  $P, Q$  of  $R$  have a common refinement, since this is true for partitions of intervals.

If  $Q$  is a refinement of  $P$ , then since  $R$  is the union of the subrectangles of itself determined by a given partition, each subrectangle for  $P$  is a union of the subrectangles for  $Q$  which it contains. This is the key fact needed to prove the following theorem in essentially the same way as the analogous theorem in one variable (Theorem 5.1.4). The details are left to the exercises.

**Theorem 10.1.2.** *Let  $f$  be a bounded function on a rectangle  $R$  in  $\mathbb{R}^d$ . If  $Q$  and  $P$  are partitions of  $R$  and  $Q$  is a refinement of  $P$ , then*

$$L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P). \quad (10.1.3)$$

Let  $P_1$  and  $P_2$  be any two partitions of  $R$  and let  $Q$  be a common refinement of  $P_1$  and  $P_2$ , then (10.1.3) holds with  $P$  replaced by  $P_1$  and with  $P$  replaced by  $P_2$ . The resulting inequalities imply the following.

**Theorem 10.1.3.** *If  $P_1$  and  $P_2$  are partitions of  $R$ , then*

$$L(f, P_1) \leq U(f, P_2).$$

Thus, any lower sum for  $f$  is less than or equal to any upper sum for  $f$ .

### Upper and Lower Integrals

**Definition 10.1.4.** Let  $R$  be a rectangle in  $\mathbb{R}^d$  and  $f$  a bounded real valued function on  $R$ . The *upper and lower integrals* of  $f$  on  $R$  are defined by

$$\begin{aligned} \overline{\int}_R f(x) dV(x) &= \inf\{U(f, P) : P \text{ a partition of } \mathbb{R}\} \\ \underline{\int}_R f(x) dV(x) &= \sup\{L(f, P) : P \text{ a partition of } \mathbb{R}\} \end{aligned} \quad (10.1.4)$$

The set of all upper sums for  $f$  is bounded below by any lower sum and the set of lower sums is bounded above by any upper sum. Thus, the inf (greatest lower bound) of the set of upper sums is greater than or equal to any lower sum and, hence, also greater than or equal to the sup (least upper bound) of the set of all lower sums. Thus,

**Theorem 10.1.5.** *If  $f$  is a bounded real valued function on a rectangle  $R$  and if  $P$  and  $Q$  are arbitrary partitions of  $R$  then*

$$L(f, P) \leq \int_{\underline{R}} f(x) dV(x) \leq \int_{\overline{R}} f(x) dV(x) \leq U(f, Q)$$

## The Integral

A bounded function on  $R$  is *integrable* if its upper and lower integrals are the same. That is:

**Definition 10.1.6.** Let  $R$  be a rectangle in  $\mathbb{R}^d$  and  $f$  a bounded real valued function on  $R$ . If  $\int_{\underline{R}} f(x) dV(x) = \int_{\overline{R}} f(x) dV(x)$ , then we will say that  $f$  is *integrable* on  $R$ . In this case, we will call the common value of these two expressions the *Riemann integral* of  $f$  on  $R$  and denote it by

$$\int_R f(x) dV(x).$$

The proofs of the following two theorems are exactly the same as the proofs of Theorems 5.1.7 and 5.1.8 and we will not repeat them here.

**Theorem 10.1.7.** *If  $f$  is a bounded function on a rectangle  $R$ , then  $f$  is Riemann integrable on  $R$  if and only if, for each  $\epsilon > 0$ , there is a partition  $P$  of  $R$  such that*

$$U(f, P) - L(f, P) < \epsilon. \quad (10.1.5)$$

**Theorem 10.1.8.** *With  $f$  and  $R$  as above,  $f$  is Riemann integrable on  $R$  if and only if there is a sequence  $\{P_n\}$  of partitions of  $R$  such that*

$$\lim(U(f, P_n) - L(f, P_n)) = 0. \quad (10.1.6)$$

*In this case,*

$$\int_R f(x) dV(x) = \lim S_n(f)$$

*where, for each  $n$ ,  $S_n(f)$  may be chosen to be  $U(f, P_n)$ ,  $L(f, P_n)$  or any Riemann sum (10.1.2) for  $f$  and the partition  $P_n$ .*

**Remark 10.1.9.** The preceding two theorems both involve the difference between the upper and lower Riemann sums for  $f$  and  $P$ . This can be written as

$$U(f, P) - L(f, P) = \sum_{j=1}^n (M_j - m_j) V(R_j). \quad (10.1.7)$$

The factors  $M_j - m_j$  that appear in this expression are non-negative numbers, as are the numbers  $V_j$ . Hence, any operation that reduces or eliminates some of the terms in this sum will result in a smaller sum.

### Properties of the Integral

The next theorem states one of the most important properties of the integral. The proof of this theorem differs in no essential way from the proof of the analogous theorem for functions of one variable (Theorem 5.2.3). In fact, the only difference is that intervals on the line are replaced by aligned rectangles in  $\mathbb{R}^d$ . We will not repeat the proof here.

**Theorem 10.1.10.** *If  $f$  and  $g$  are integrable functions on an aligned rectangle  $R$  in  $\mathbb{R}^d$  and  $c$  is a constant, then*

$$(a) \text{ } cf \text{ is integrable and } \int_R cf(x)dV(x) = c \int_R f(x)dV(x);$$

$$(b) \text{ } f+g \text{ is integrable and } \int_R (f+g)(x)dV(x) = \int_R f(x)dV(x) + \int_R g(x)dV(x).$$

Taken together, the statements of the above theorem mean that the integrable functions on  $R$  form a vector space under pointwise addition and scalar multiplication of functions, and the integral is a linear transformation from this vector space to the vector space  $\mathbb{R}$ .

The order preserving property is another key property of the integral. The version stated in the next theorem is somewhat more general than the analogous result, proved earlier for functions of a single variable (Theorem 5.2.4), and it has a different proof. Hence, we include the proof.

**Theorem 10.1.11.** *If  $f$  and  $g$  are functions on an aligned rectangle  $R$  in  $\mathbb{R}^d$ , and  $f(x) \leq g(x)$  for all  $x \in [a, b]$ , then*

$$(a) \int_R^{\overline{}} f(x)dV(x) \leq \int_R^{\overline{}} g(x)dV(x) \text{ and } \int_{\underline{}} f(x)dV(x) \leq \int_{\underline{}} g(x)dV(x);$$

$$(b) \int_R f(x)dV(x) \leq \int_R g(x)dV(x) \text{ if } f \text{ and } g \text{ are integrable.}$$

*Proof.* We will prove this result for the upper integrals. The result for the lower integrals has an analogous proof. The result for the integral in the case of integrable functions then follows because upper integral, lower integral, and integral are all the same for an integrable function.

Given a partition  $P$  of  $R$ , determining subrectangles  $\{R_1, \dots, R_n\}$  of  $R$ , we set

$$M_j(f) = \sup_{R_j} f \quad \text{and} \quad M_j(g) = \sup_{R_j} g.$$

Then  $M_j(f) \leq M_j(g)$  for all  $j$  because  $f(x) \leq g(x)$  for all  $x \in R$ . Hence,

$$U(f, P) = \sum_{j=1}^n M_j(f)V(R_j) \leq \sum_{j=1}^n M_j(g)V(R_j) = U(g, P).$$

It follows that

$$\overline{\int}_R f(x)dV(x) = \inf_P U(f, P) \leq \inf_P U(g, P) = \overline{\int}_R g(x)dV(x)$$

This completes the proof.  $\square$

## A Simple Example

So far we have not computed a single integral or shown that a single function is integrable. We do so now. The function we will integrate is very simple, though not continuous, but the computation of its integral is an important step in our development of integration theory.

**Definition 10.1.12.** Let  $E$  be a subset of  $\mathbb{R}^d$ . Then the *characteristic function* of  $E$ , denoted  $\chi_E$  is the real valued function on  $\mathbb{R}^d$  defined by

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E. \end{cases}$$

Our example is as follows:

**Example 10.1.13.** Let  $R$  and  $S$  be aligned rectangles with  $S \subset R$ . Show that  $\chi_S$  is an integrable function on  $R$  and

$$\int_R \chi_S(x)dV(x) = V(S).$$

**Solution:** Let

$$R = [a_1, b_1] \times \cdots \times [a_d, b_d] \text{ and} \\ S = [s_1, t_1] \times \cdots \times [s_d, t_d],$$

where  $a_j \leq s_j \leq t_j \leq b_j$  for each  $j$ . Given  $\epsilon > 0$ , We choose a partition of  $R$  as follows: for each  $j$ , we partition each interval  $[a_j, b_j]$  with the points  $\{a_j \leq u_j \leq s_j \leq t_j \leq v_j \leq b_j\}$ , where the points  $u_j$  and  $v_j$  are chosen so that if  $A$  is the rectangle

$$A = [u_1, v_1] \times \cdots \times [u_d, v_d]$$

Then  $V(A) < V(S) + \epsilon$  (see Figure 10.2 for a two dimensional version of this setup).

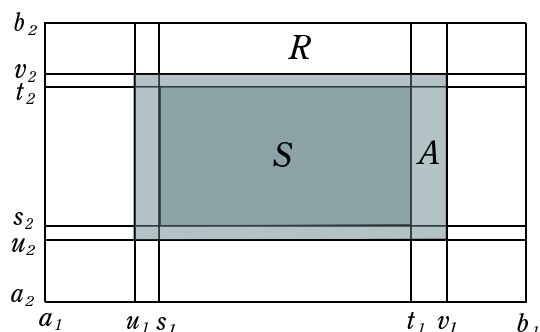
The sup of  $\chi_S$  on a given subrectangle  $R_j$  is 1 if  $R_j \cap S \neq \emptyset$  and is 0 otherwise. The inf of  $\chi_S$  on  $R_j$  is 1 if  $R_j \subset S$  and is 0 otherwise.

There is only one subrectangle for this partition which is contained in  $S$  and that is  $S$  itself. Thus,

$$L(\chi_S, P) = V(S).$$

The union of the subrectangles  $R_j$  that meet  $S$  is  $A$ . Hence,

$$U(\chi_S, P) = V(A).$$

Figure 10.2: Computing the Integral of  $\chi_S$ 

Since  $V(S) < V(A) < V(S) + \epsilon$ , we have  $V(A) - V(S) < \epsilon$ . Hence,

$$U(\chi_S, P) - L(\chi_S, P) < \epsilon.$$

By Theorem 10.1.7,  $\chi_S$  is integrable on  $R$ . Its integral is within  $\epsilon$  of  $L(\chi_S, P) = V(S)$  for every  $\epsilon > 0$  and so  $\int_R \chi_S(x) dV(x) = V(S)$ .

### Exercise Set 10.1

- Let  $R = [0, 1] \times [0, 1]$  be the square with vertices at  $(0, 0)$ ,  $(1, 0)$ ,  $(1, 1)$ , and  $(0, 1)$  and let  $P$  be the partition of  $R$  consisting of the partition  $\{0, 1/4, 1/2, 3/4, 1\}$  in both factors of  $[0, 1] \times [0, 1]$ . Find  $U(f, P)$  and  $L(f, P)$  if  $f(x, y) = xy$ .
- With  $R$  and  $P$  as in the previous problem, find  $U(\chi_\Delta, P)$  and  $L(\chi_\Delta, P)$  if  $\Delta$  is the closed, solid triangle with vertices at  $(0, 0)$ ,  $(1, 0)$ ,  $(1, 1)$ .
- Suppose  $f$  and  $g$  are functions defined on an aligned rectangle  $R$ . Suppose there is a positive constant  $K$  such that  $|f(x) - f(y)| \leq K|g(x) - g(y)|$  for all  $x, y \in R$ . Prove that if  $g$  is integrable on  $R$ , then so is  $f$ .
- Use the result of the preceding exercise to prove that if  $f$  is an integrable function on an aligned rectangle  $R$ , then  $|f|$  is also integrable on  $R$ .
- Prove that if  $f$  is integrable on  $R$ , then  $f^2$  is also integrable on  $R$ .
- Use the result of the preceding exercise to prove that if  $f$  and  $g$  are integrable on  $R$ , then  $fg$  is also integrable on  $R$ .
- Show that each constant function  $k$  is integrable and  $\int_R k dV(x) = kV(R)$ .
- If  $f$  is an integrable function defined on the rectangle  $R$  and  $|f(x)| \leq M$  on  $R$ , where  $M$  is a positive constant, then prove that

$$\left| \int_R f(x) dV(x) \right| \leq MV(R).$$

9. Prove that if  $R$  is an aligned rectangle and  $f$  is a continuous function on  $R$ , then  $f$  is integrable on  $R$ .
10. If  $A$  and  $B$  are subsets of  $\mathbb{R}^d$ , then
- describe  $\chi_{A \cap B}$  in terms of  $\chi_A$  and  $\chi_B$ ;
  - describe  $\chi_{A \cup B}$  in terms of  $\chi_A$  and  $\chi_B$ ;
  - describe the meaning of  $B \subset A$  in terms of  $\chi_A$  and  $\chi_B$ ;
  - if  $B \subset A$ , describe  $\chi_{A \setminus B}$  in terms of  $\chi_A$  and  $\chi_B$ .

## 10.2 Jordan Regions

The concept of characteristic function of a set (Definition 10.1.12) allows us to define the volume of a set in terms of the integral that we just defined. The volume (or inner or outer volume) of a set  $E$ , as defined below, depends very much on the dimension of the ambient space  $\mathbb{R}^d$  and so, technically, it should be called the  $d$ -volume (or inner or outer  $d$ -volume) of the set. However, as with rectangles, we will drop the  $d$  when the dimension of the ambient space is understood.

**Definition 10.2.1.** If  $E$  is a bounded subset of  $\mathbb{R}^d$ , let  $R$  be an aligned rectangle containing  $E$ . Then we define the outer volume  $\overline{V}(E)$ , inner volume  $\underline{V}(E)$ , and volume  $V(E)$  (if it exists) for  $E$  by

- $\overline{V}(E) = \overline{\int}_R \chi_E(x) dV(x)$ ;
- $\underline{V}(E) = \underline{\int}_R \chi_E(x) dV(x)$ ; and
- $V(E) = \int_R \chi_E(x) dV(x)$  if the latter exists – that is if  $\underline{\int}_R \chi_E dV(x) = \overline{\int}_R \chi_E(x) dV(x)$ .

If  $V(E)$  exists, then we call  $E$  a *Jordan region*.

Note that  $E$  is a Jordan region if and only if  $\underline{V}(E) = \overline{V}(E)$  and, in this case,  $V(E)$  is their common value.

Note also that, if  $E$  is an aligned rectangle, then  $E$  is a Jordan region and the above definition of  $V(E)$  agrees with our earlier definition. This is demonstrated in Example 10.1.13.

Implicit in the above definition is the fact that the upper and lower integrals of  $\chi_E$  over  $R$  do not depend on the rectangle  $R$ , as long as  $R$  contains  $E$ . We leave a proof of this to the exercises (Exercise 10.2.1).

**Example 10.2.2.** Show that the closed, solid right triangle  $\Delta$  in  $\mathbb{R}^2$  with vertices at  $(0, 0)$ ,  $(a, 0)$ , and  $(0, b)$  is a Jordan region and has area (2-volume)  $ab/2$ .

**Solution:** We choose  $R$  to be the rectangle  $[0, a] \times [0, b]$ . This contains the triangle  $\Delta$ . For each  $n$ , we choose a partition  $P_n$  of  $R$  consisting of partitions

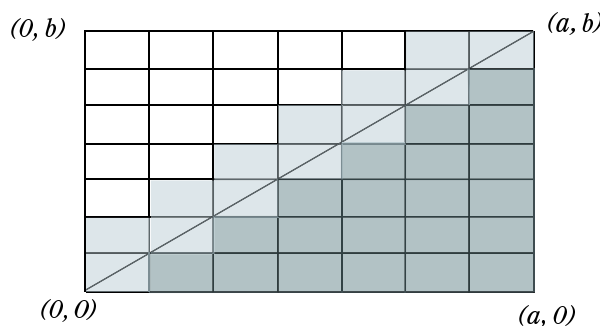


Figure 10.3: Computing the Area of a Triangle

$\{0, a/n, 2a/n, \dots, na/n = a\}$  of  $[0, a]$  and  $\{0, b/n, 2b/n, \dots, nb/n = b\}$  of  $[0, b]$ . This determines  $n^2$  subrectangles of  $R$ , each of volume  $ab/n^2$ .

Now for each of these subrectangles  $R_j$ , the sup,  $M_j$ , and inf,  $m_j$ , of  $\chi_\Delta$  on  $R_j$  is either 1 or 0. In fact,

$$\begin{aligned} M_j &= 1 && \text{if and only if } R_j \cap \Delta \neq \emptyset \\ m_j &= 1 && \text{if and only if } R_j \subset \Delta. \end{aligned}$$

Thus, the only subrectangles  $R_j$  on which  $M_j \neq m_j$  are those which are not contained in  $\Delta$  but have non-empty intersection with it (the light grey subrectangles in Figure 10.3). There are two kinds of these, those of the form  $[(k-1)a/n, ka/n] \times [(k-1)b/n, kb/n]$  which are bisected by the line from  $(0,0)$  to  $(a,b)$  and those of the form  $[(k-1)a/n, ka/n] \times [kb/n, (k+1)b/n]$  which just have a lower right vertex on this line. There are  $n$  of the former and  $n-1$  of the latter. The difference  $U(\chi_\Delta, P_n) - L(\chi_\Delta, P_n)$  is just the sum of the areas of these  $2n-1$  rectangles, which is  $(2n-1)ab/n^2$ . Hence,

$$\lim_{n \rightarrow \infty} (U(\chi_\Delta, P_n) - L(\chi_\Delta, P_n)) = \lim_{n \rightarrow \infty} \frac{(2n-1)ab}{n^2} = 0.$$

By Theorem 10.1.8, the Riemann integral  $\int_R \chi_\Delta(x) dV(x)$  exists and so the 2-volume (area) of the set  $\Delta$  exists – that is,  $\Delta$  is a Jordan region.

Also by Theorem 10.1.8 the integral  $\int_R \chi_\Delta(x) dV(x)$  is the limit of the sequence  $\{L(\chi_\Delta, P_n)\}$ . However,  $L(\chi_\Delta, P_n)$  is the sum of the areas of the subrectangles that are contained in  $\Delta$  (the dark grey subrectangles in Figure 10.3). There are  $n(n-1)/2$  of these (half the number remaining after the ones that are bisected by the line from  $(0,0)$  to  $(a,b)$  are removed). Hence,

$$V(\Delta) = \int_R \chi_\Delta(x) dV(x) = \lim_{n \rightarrow \infty} \frac{n(n-1)ab}{2n^2} = \frac{ab}{2}.$$

### Properties of Volume

Many properties of the integral translate directly into properties of volume. For example, Theorem 10.1.11 implies that

**Theorem 10.2.3.** *If  $E$  and  $F$  are bounded subsets of  $\mathbb{R}^d$  and  $E \subset F$ , then*

$$\underline{V}(E) \leq \underline{V}(F) \quad \text{and} \quad \overline{V}(E) \leq \overline{V}(F).$$

*If  $E$  and  $F$  are Jordan regions, then  $V(E) \leq V(F)$ .*

Theorem 10.1.10 and the fact that  $\chi_{E \cup F} = \chi_E + \chi_F - \chi_{E \cap F}$  (Exercise 10.1.10) imply

**Theorem 10.2.4.** *If  $E$ ,  $F$  and  $E \cap F$  are Jordan regions and  $V(E \cap F) = 0$ , then  $E \cup F$  is a Jordan region and*

$$V(E \cup F) = V(E) + V(F).$$

*In particular, this identity holds if  $E$  and  $F$  are disjoint Jordan regions.*

In particular, if  $R$  is an aligned rectangle in  $\mathbb{R}^d$  and  $R_j \neq R_k$  are two of the subrectangles determined by a partition  $P$ , then  $R_j \cap R_k$  is either empty or is a degenerate aligned rectangle in  $R$  – that is, its dimension is lower than that of  $R$ . Hence,  $V(R_j \cap R_k) = 0$ . Thus, by Theorem 10.1.10,

$$V(R_j \cup R_k) = V(R_j) + V(R_k).$$

An induction argument then shows that if  $F$  is the union of any number of the subrectangles determined by  $P$ , then  $F$  is a Jordan region and  $V(F)$  is the sum of the volumes of these subrectangles. This is used in the proof of the following theorem.

**Theorem 10.2.5.** *If  $E$  is a bounded subset of  $\mathbb{R}^d$ , then  $\overline{V}(E) = \overline{V}(\overline{E})$  and  $\underline{V}(E) = \underline{V}(E^\circ)$ .*

*Proof.* Let  $R$  be an aligned rectangle containing  $E$ , let  $P$  be a partition of  $R$ , and let  $\{R_j\}$  be the list of subrectangles of  $R$  determined by  $P$ . Then  $U(\chi_E, P)$  is the sum of the volumes of the rectangles  $R_j$  in this list that have a non-empty intersection with  $E$  (those for which  $\chi_E$  takes on the value 1 somewhere on  $R_j$ ). If we set

$$F = \bigcup \{R_j : E \cap R_j \neq \emptyset\},$$

then  $U(\chi_E, P) = V(F)$ , by the paragraph preceding this theorem.

Now  $F$  is a finite union of closed sets and so it is also closed. Since  $E \subset F$ , we also have  $\overline{E} \subset F$ . Then

$$\overline{V}(E) \leq \overline{V}(\overline{E}) \leq \overline{V}(F) = V(F) = U(\chi_E, P).$$

Since  $\overline{V}(E) = \inf\{U(\chi_E, P) : P \text{ a partition of } R\}$ , we have

$$\overline{V}(E) \leq \overline{V}(\overline{E}) \leq \overline{V}(E).$$

Thus,  $\overline{V}(E) = \overline{V}(\overline{E})$ .

Similarly, if we set

$$G = \bigcup \{R_j : R_j \subset E\},$$

then, since  $G^\circ \subset E^\circ$ ,

$$\underline{V}(G^\circ) \leq \underline{V}(E^\circ) \leq \underline{V}(E).$$

However,  $V(G^\circ) = V(G) = L(\chi_E, P)$ , since the boundary of  $G$  consists of a finite union of rectangles of dimension lower than  $d$ , and these all have volume 0. Since  $\sup_P L(\chi_E, P) = \underline{V}(E)$ , we conclude that  $\underline{V}(E^\circ) = \underline{V}(E)$ . This completes the proof.  $\square$

**Theorem 10.2.6.** *If  $E$  is a Jordan region, then so are  $\overline{E}$  and  $E^\circ$ . Furthermore,  $V(E) = V(\overline{E}) = V(E^\circ)$ .*

*Proof.* In view of the previous theorem,

$$\underline{V}(E) \leq \underline{V}(\overline{E}) \leq \overline{V}(\overline{E}) \leq \overline{V}(E).$$

If  $E$  is a Jordan region, then  $\underline{V}(E) = \overline{V}(E)$  and, hence, each of the above inequalities is an equality. This implies  $\overline{E}$  is a Jordan region and  $V(\overline{E}) = V(E)$ . The proof of the statement for  $E^\circ$  is similar.  $\square$

## Sets of Volume Zero

We leave the proof of the following theorem to the exercises.

**Theorem 10.2.7.** *If  $E$  is a bounded set with  $\overline{V}(E) = 0$ , then  $E$  is a Jordan region with volume 0. Any subset of a Jordan region of volume 0 is also a Jordan region of volume 0. A finite union of Jordan regions of volume 0 is also a Jordan region of volume 0.*

We will, henceforth, refer to a set  $E$  with  $\overline{V}(E) = 0$  as simply a *set of volume 0*.

**Theorem 10.2.8.** *A set  $E$  is a set of volume 0 if and only if, for each  $\epsilon > 0$ , there is a finite set  $\{R_1, \dots, R_n\}$  of aligned rectangles such that*

$$E \subset \bigcup_{j=1}^n R_j \quad \text{and} \quad \sum_{j=1}^n V(R_j) < \epsilon.$$

*Proof.* If  $\overline{V}(E) = 0$ , then there exist an aligned rectangle  $R$  with  $E$  in its interior and a partition  $P$  of  $R$  such that  $U(\chi_E, P) < \epsilon$ . This just means that those subrectangles determined by  $P$  which meet  $E$  have volumes which add up to a number less than  $\epsilon$ . Since  $E$  is contained in the union of these rectangles, the proof of the "only if" part of the theorem is complete.

On the other hand, if  $E \subset F = \cup_{j=1}^n R_j$  for a set of aligned rectangles with volumes adding up to a number less than  $\epsilon$ , then  $\overline{V}(F) < \epsilon$  since

$$\chi_F \leq \sum_{j=1}^n \chi_{R_j},$$

which, together with the fact that each  $\chi_{R_j}$  is integrable, implies

$$\begin{aligned} \overline{V}(F) &= \int_R \chi_F(x) dV(x) \leq \int_R \sum_{j=1}^n \chi_{R_j}(x) dV(x) \\ &= \sum_{j=1}^n \int_R \chi_{R_j}(x) dV(x) = \sum_{j=1}^n V(R_j) < \epsilon. \end{aligned}$$

This proves the "if" part of the theorem.  $\square$

## A Characterization of Jordan Regions

**Theorem 10.2.9.** *A bounded set  $E$  is a Jordan region if and only if its boundary,  $\partial E$ , is a set of volume 0.*

*Proof.* If  $P$  is a partition of  $R$  determining a list of subrectangles  $\{R_j\}$ , then  $L(\chi_{E^\circ}, P)$  is the sum of the areas of those  $R_j$  which are entirely contained in  $E^\circ$ , while  $U(\chi_{\overline{E}}, P)$  is the sum of the areas of those  $R_j$  which have non-empty intersection with  $\overline{E}$ . It follows that

$$U(\chi_{\overline{E}}, P) - L(\chi_{E^\circ}, P) = U(\chi_{\partial E}, P).$$

Hence, a sequence  $\{P_n\}$  of partitions has the property that  $\lim U(\chi_{\partial E}, P_n) = 0$  if and only if it has the property that

$$\lim(U(\chi_{\overline{E}}, P_n) - L(\chi_{E^\circ}, P_n)) = 0.$$

Since, for an appropriately chosen sequence of partitions, this limit is

$$\overline{V}(\overline{E}) - \underline{V}(E^\circ) = \overline{V}(E) - \underline{V}(E),$$

by Theorem 10.2.5, we conclude that  $\underline{V}(E) = \overline{V}(E)$  if and only if  $\overline{V}(\partial E) = 0$  — that is,  $E$  is a Jordan region if and only if  $\partial E$  is a set of volume 0.  $\square$

**Theorem 10.2.10.** *If  $A$  and  $B$  are Jordan regions, then  $A \cap B$ ,  $A \cup B$ , and  $A \setminus (A \cap B)$  are also Jordan regions. Furthermore,*

$$\begin{aligned} V(A \cup B) &= V(A) + V(B) - V(A \cap B), \text{ and} \\ V(A \setminus (A \cap B)) &= V(A) - V(A \cap B). \end{aligned} \tag{10.2.1}$$

*Proof.* Each of the sets  $A \cap B$ ,  $A \cup B$ , and  $A \setminus (A \cap B)$  has its boundary contained in  $\partial A \cup \partial B$ . Since  $A$  and  $B$  are Jordan regions,  $\partial A$  and  $\partial B$  are sets of volume 0. Then Theorem 10.2.7 implies that  $\partial A \cup \partial B$  has volume 0, as does each of its subsets. It follows from the previous theorem that  $A \cap B$ ,  $A \cup B$ , and  $A \setminus (A \cap B)$  are Jordan regions.

The second statement of the theorem follows from the identities

$$\begin{aligned}\chi_{A \cup B} &= \chi_A + \chi_B - \chi_{A \cap B}, \text{ and} \\ \chi_{A \setminus (A \cap B)} &= \chi_A - \chi_{A \cap B}.\end{aligned}\tag{10.2.2}$$

□

**Example 10.2.11.** Let  $K$  be a compact subset of  $\mathbb{R}^{d-1}$  and let  $f : K \rightarrow \mathbb{R}$  be a continuous function. Show that the graph  $G(f)$  of  $f$  is a set of  $d$ -volume 0, where  $G(f) = \{(x, f(x)) : x \in K\}$ .

**Solution:** Since  $K$  is compact, it is bounded, and so we may choose a rectangle  $R$  in  $\mathbb{R}^{d-1}$  which contains  $K$ . Let  $W$  be the  $(d-1)$ -volume of  $R$ .

Since  $K$  is compact and  $f$  is continuous,  $f$  is actually uniformly continuous. Thus, given  $\epsilon > 0$  we may choose a  $\delta > 0$  such that

$$|f(x) - f(y)| < \epsilon/W \quad \text{whenever} \quad \|x - y\| < \delta.$$

We let  $P$  be a partition of  $R$  such that the diameter of each subrectangle for the partition is less than  $\delta$  (diameter in this case means maximal distance between two points in the subrectangle). Let  $R_1, R_2, \dots, R_n$  be a list of those subrectangles for this partition which meet  $K$ . If

$$m_j = \min\{f(x) : x \in K \cap R_j\} \quad \text{and} \quad M_j = \max\{f(x) : x \in K \cap R_j\},$$

then

$$G(f) \subset \bigcup_j (R_j \times [m_j, M_j]).$$

The sum of the volumes of the rectangles  $R_j \times [m_j, M_j]$  is

$$\sum_j V(R_j)(M_j - m_j) \leq \frac{\epsilon}{W} \sum V(R_j) \leq \frac{\epsilon}{W} W = \epsilon.$$

By Theorem 10.2.8 the graph  $G(f)$  of  $f$  is a set of volume 0.

### Exercise Set 10.2

1. Prove that  $\int_R \chi_E(x) dV(x)$  and  $\int_{\underline{R}} \chi_E(x) dV(x)$  do not depend on the choice of the aligned rectangle  $R$  as long as it contains  $E$ .
2. Prove Theorem 10.2.7 – that is, show that if a subset  $A$  of  $\mathbb{R}^d$  has outer volume zero, then it and each of its subsets is a Jordan region of volume 0.

3. Show that a finite set in  $\mathbb{R}^d$  has volume 0.
4. If  $E$  is the subset of the unit square  $[0, 1] \times [0, 1]$  consisting of points with both coordinates rational numbers, find its inner volume  $\underline{V}(E)$  and outer volume  $\overline{V}(E)$ . Is  $E$  a Jordan region?
5. Show that if  $A$  and  $B$  are sets of volume 0 in  $\mathbb{R}^d$ , then  $A \cup B$  is also a set of volume 0.
6. Let  $U$  be an open subset of  $\mathbb{R}^2$  and  $K \subset U$  a compact set. Suppose  $f : U \rightarrow \mathbb{R}$  is a smooth function and  $E = \{(x, y) \in K : f(x, y) = 0\}$ . If  $df$  is never 0 on  $E$ , then show that  $E$  is a set of area 0 in  $\mathbb{R}^2$ .
7. Show that an ellipse in  $\mathbb{R}^2$  is a set of area 0 in  $\mathbb{R}^2$  and the solid ellipse that it bounds is a Jordan region.
8. Show that a bounded subset of  $\mathbb{R}^2$  whose boundary is a finite union of smooth parameterized curves, is a Jordan region.
9. Consider the following three reflection transformations of  $\mathbb{R}^2$ :

$$T_1(x, y) = (-x, y), \quad T_2(x, y) = (x, -y) \quad \text{and} \quad T_3(x, y) = (y, x).$$

These are reflection through the  $y$ -axis, reflection through the  $x$ -axis, and reflection through the line  $y = x$ , respectively. Prove that if  $E$  is a Jordan region, then, for  $j = 1, 2, 3$ , so is  $T_j(E)$  and  $V(T_j(E)) = V(E)$ . Hint: what do these reflections do to aligned rectangles and their volumes?

10. Using the previous two exercises and theorems from this section, but without using Example 10.2.2, give a proof that the area of a triangle with one side parallel to a coordinate axis is one half its base times its height. Hint: prove this first for right triangles with legs parallel to the axes.
11. Using the result of the preceding exercise, show that a parallelogram in  $\mathbb{R}^2$  with one side parallel to a coordinate axis has area equal to its base times its height.
12. Suppose  $B \subset \mathbb{R}^d$  is a compact Jordan region and  $f$  and  $g$  continuous real valued functions on  $B$  with  $g(x) \leq f(x)$ . Show that the set

$$A = \{(x, t) \in \mathbb{R}^{d+1} : x \in B, \text{ and } g(x) \leq t \leq f(x)\}$$

is also a Jordan region.

### 10.3 The Integral over a Jordan Region

In this section we extend the definition of the integral to cover integration over a Jordan region. We also prove an existence theorem which shows that the class of integrable functions is quite large.

### An Existence Theorem

So far we have only proved the existence of the integral for a few functions of the form  $\chi_E$ . Our next objective is to prove a general existence theorem for the integral over an aligned rectangle. We will then extend this theorem to integrals over Jordan regions.

**Theorem 10.3.1.** *Let  $f$  be a bounded function on an aligned rectangle  $R$ . If the set of points of  $R$  at which  $f$  is not continuous is a set of volume 0, then  $f$  is integrable on  $R$ .*

*Proof.* Let  $E$  be the set of points of  $R$  at which  $f$  is not continuous. Since  $E$  is a set of volume 0, its outer volume  $\overline{V}(E)$  is 0. Hence, given  $\epsilon > 0$ , there is a partition  $P$  of  $R$  such that  $U(\chi_E, P) < \epsilon/(4M)$ , where  $M$  is the sup of  $|f|$  on  $R$ . If  $A$  is the union of the subrectangles for  $P$  which meet  $E$ , then this means that

$$V(A) = U(\chi_E, P) < \frac{\epsilon}{4M}.$$

Let  $B$  be the union of the subrectangles for  $P$  which do not meet  $E$ . Note that  $A \cup B = R$  and  $B$  is a closed, bounded (hence compact) set on which  $f$  is continuous. Hence,  $f$  is uniformly continuous on  $B$  by Theorem 8.2.12. This implies that we may choose a  $\delta > 0$  such that

$$|f(x) - f(y)| < \frac{\epsilon}{2V(R)} \quad \text{whenever} \quad \|x - y\| < \delta.$$

We next choose a refinement  $Q$  for the partition  $P$  in such a way that the diameter of each subrectangle for  $Q$  is at most  $\delta$ . If  $R_1, R_2, \dots, R_n$  is a list of the subrectangles for  $Q$ , then each  $R_j$  is either in  $A$  or in  $B$ . We let  $S$  be the set of integers  $j$  in  $[1, n]$  such that  $R_j \subset A$  and  $T$  the set of integers  $j$  in this interval such that  $R_j \subset B$ . If  $M_j$  and  $m_j$  are the sup and inf of  $f$  on  $R_j$ , then

$$\begin{aligned} U(f, Q) - L(f, Q) &= \sum_{j=1}^n (M_j - m_j)V(R_j) \\ &= \sum_{j \in S} (M_j - m_j)V(R_j) + \sum_{j \in T} (M_j - m_j)V(R_j) \\ &\leq 2MV(A) + \frac{\epsilon}{2V(R)}V(B) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

In view of Theorem 10.1.7, the proof is complete. □

### The Integral over a Jordan Region

**Definition 10.3.2.** Let  $A$  be a Jordan region and  $f$  a bounded function defined on a set containing  $A$ . We define a new function  $f_A$ , with domain all of  $\mathbb{R}^d$ , as follows:

$$f_A(x) = \begin{cases} f(x) & \text{if } x \in A \\ 0 & \text{if } x \in \mathbb{R}^d \setminus A. \end{cases}$$

Thus,  $f_A$  is a function defined on all of  $\mathbb{R}^d$ . It agrees with  $f$  on  $A$  and is 0 on the complement of  $A$ . Note that  $f$  may be originally defined on a larger set than  $A$  or it may be defined just on  $A$ . In the definition of  $f_A$ , it doesn't matter.

**Example 10.3.3.** Let  $A = D_1(0,0)$  in  $\mathbb{R}^2$ . Find  $f_A$  and  $g_A$  if  $f$  is defined on  $\mathbb{R}^2$  by  $f(x,y) = x^2 + y^2$  and  $g$  is defined on  $A$  by  $g(x,y) = \sqrt{1 - x^2 - y^2}$ .

**Solution:** From the above definition, we have

$$f_A(x,y) = \begin{cases} x^2 + y^2 & \text{if } (x,y) \in D_1(0) \\ 0 & \text{if } (x,y) \notin D_1(0). \end{cases}$$

and

$$g_A(x,y) = \begin{cases} \sqrt{1 - x^2 - y^2} & \text{if } (x,y) \in D_1(0) \\ 0 & \text{if } (x,y) \notin D_1(0). \end{cases}$$

Note that here  $f$  is defined originally on all of  $\mathbb{R}^2$  while  $g$  is defined only on  $A$ .

**Definition 10.3.4.** With  $A$ ,  $f$  and  $f_A$  as in the preceding definition, let  $R$  be an aligned rectangle containing  $A$ . If  $f_A$  is integrable on  $R$  we say  $f$  is integrable on  $A$  and we write

$$\int_A f(x)dV(x) = \int_R f_A(x)dV(x).$$

Implicit in the above definition is the assumption that  $\int_R f_A(x)dV(x)$  does not depend on which rectangle  $R$  is chosen, as long as it contains  $A$ . We leave the proof of this to the exercises.

If  $A$  happens to be an aligned rectangle, then one choice for  $R$  in the above definition is  $R = A$ . Then  $f = f_A$  on the rectangle  $R$  and

$$\int_A f(x)dV(x) = \int_R f_A(x)dV(x) = \int_R f(x)dV(x),$$

where, on the right, the integral over  $R$  is the one defined in Section 10.1, while the one on the left is our new definition of the integral over a Jordan region. Fortunately, the two agree.

## Existence of the Integral over a Jordan Region

**Theorem 10.3.5.** *Let  $A$  be a Jordan region and  $f$  a bounded function defined on  $A$ . If the set  $E$  of points of  $A$  at which  $f$  is not continuous is a set of volume 0, then  $f$  is integrable on  $A$ .*

*Proof.* Since both  $E$  and  $\partial A$  are sets of volume 0, their union  $F = E \cup \partial A$  is also. We choose an aligned rectangle  $R$  such that  $\overline{A} \subset R$ . Then  $f_A$  is continuous on  $R \setminus F$ . It follows from Theorem 10.3.1 that  $f_A$  is integrable on  $R$  and, by definition,  $f$  is integrable on  $A$ .  $\square$

### Properties of the Integral

For integrals over rectangles, the following theorem is Exercise 10.1.6. The extension of this result to integrals over Jordan regions is left to the exercises.

**Theorem 10.3.6.** *If  $A$  is a Jordan region and  $f$  and  $g$  are integrable functions on  $A$ , then  $fg$  is also integrable on  $A$ .*

**Example 10.3.7.** Prove that if  $B \subset A$  and  $A$  and  $B$  are Jordan regions, then each function  $f$  which is integrable on  $A$  is also integrable on  $B$ .

**Solution:** This follows immediately from the preceding theorem and the observation that  $f_B = \chi_B f_A$ .

The next three theorems follow from Theorems 10.1.10, 10.1.11, and 10.3.6 and some observations about the passage from  $f$  to  $f_A$ . We leave the details to the exercises.

**Theorem 10.3.8.** *Let  $A$  be a Jordan region,  $f$  and  $g$  integrable functions on  $A$  and  $c$  a scalar constant. Then  $f + g$  and  $cf$  are integrable on  $A$ , and*

$$(a) \int_A 1 dV(x) = V(A);$$

$$(b) \int_A (f(x) + g(x))dV(x) = \int_A f(x)dV(x) + \int_A g(x)dV(x);$$

$$(c) \int_A cf(x)dV(x) = c \int_A f(x)dV(x).$$

Parts (b) and (c) mean that the integral over  $A$  is a linear transformation.

**Theorem 10.3.9.** *Let  $A$  and  $B$  be Jordan regions with  $V(A \cap B) = 0$  and let  $f$  be a bounded function on  $A \cup B$ . Then  $f$  is integrable on  $A$  and on  $B$  if and only if it is integrable on  $A \cup B$ . In this case,*

$$\int_{A \cup B} f(x)dV(x) = \int_A f(x)dV(x) + \int_B f(x)dV(x).$$

**Theorem 10.3.10.** *If  $A$  is a Jordan region and  $f$  and  $g$  are integrable functions on  $A$  with  $f(x) \leq g(x)$  for all  $x \in A$ , then*

$$\int_A f(x)dV(x) \leq \int_A g(x)dV(x).$$

### Integral of a Sequence

**Theorem 10.3.11.** *Let  $A$  be a Jordan region and  $\{f_n\}$  a sequence of integrable functions on  $A$ . If  $\{f_n\}$  converges uniformly on  $A$  to a function  $f$ , then  $f$  is integrable and*

$$\lim_{n \rightarrow \infty} \int_A f_n(x)dV(x) = \int_A f(x)dV(x).$$

*Proof.* We prove this first in the case where  $A$  is an aligned rectangle  $R$ .

Given  $\epsilon > 0$ , there is an  $N$  such that  $|f(x) - f_n(x)| < \epsilon/V(A)$  whenever  $x \in R$  and  $n \geq N$ . this means that, for  $n \geq N$ ,

$$f_n(x) - \frac{\epsilon}{V(R)} < f(x) < f_n(x) + \frac{\epsilon}{V(R)},$$

for all  $x \in R$ . By Theorem 10.1.11 this implies that

$$\begin{aligned} \int_{\underline{R}} (f_n(x) - \epsilon/V(R))dV(x) &\leq \int_{\underline{R}} f(x)dV(x) \\ &\leq \int_{\overline{R}} f(x)dV(x) \leq \int_{\overline{R}} (f_n(x) + \epsilon/V(R))dV(x). \end{aligned}$$

Since  $f_n$  and the constant  $\epsilon/(2V(R))$  are integrable, their upper and lower integrals are the same and are equal to their integrals. Thus,

$$\int_{\underline{R}} f_n(x)dV(x) - \epsilon \leq \int_{\underline{R}} f(x)dV(x) \leq \int_{\overline{R}} f(x)dV(x) \leq \int_{\overline{R}} f_n(x)dV(x) + \epsilon.$$

Since  $\epsilon$  is an arbitrary positive number, we conclude that

$$\int_{\underline{R}} f(x)dV(x) = \int_{\overline{R}} f(x)dV(x)$$

and, hence, that  $f$  is integrable on  $R$ . These inequalities also show that

$$\left| \int_{\overline{R}} f_n(x)dV(x) - \int_{\overline{R}} f(x)dV(x) \right| < \epsilon \quad \text{whenever } n \geq N.$$

Thus,  $\lim \int_{\overline{R}} f_n(x)dV(x) = \int_{\overline{R}} f(x)dV(x)$ .

Now if  $A$  is not an aligned rectangle, we simply choose an aligned rectangle  $R$  which contains  $A$  and replace  $f$  and  $f_n$  by  $f_A$  and  $(f_n)_A$  in the above argument. We note that  $\{(f_n)_A\}$  converges uniformly to  $f_A$  on  $R$  if  $\{f_n\}$  converges uniformly to  $f$  on  $A$ . The conclusion is that  $f_A$  is integrable on  $R$  and

$$\lim \int_{\overline{R}} (f_n)_A(x)dV(x) = \int_{\overline{R}} f_A(x)dV(x).$$

This implies that  $f$  is integrable on  $A$  and

$$\lim \int_A f_n(x)dV(x) = \int_A f(x)dV(x).$$

□

**Example 10.3.12.** Show that if  $f$  is a bounded function on a Jordan region  $A$  and if  $\{x \in A : f(x) < r\}$  is a Jordan region for each  $r \in \mathbb{R}$ , then  $f$  is integrable on  $A$ .

**Solution:** Since  $f$  is bounded, there is an  $M > 0$  such that  $-M < f(x) < M$  for all  $x \in A$ . We set

$$g(x) = \frac{f(x) + M}{2M} \quad \text{so that} \quad f(x) = 2Mg(x) - M.$$

The function  $g$  also satisfies the hypothesis of the theorem, and  $0 < g(x) < 1$  for all  $x \in A$ . We will show that  $g$  is integrable. This clearly implies that  $f$  is integrable.

We will show that  $g$  is integrable by expressing it as a uniform limit of a sequence of integrable functions. This sequence is constructed as follows. For each positive integer  $n$  and each positive integer  $k \leq n$ , we set

$$\begin{aligned} E(n, k) &= \{x \in A : (k-1)/n \leq f(x) < k/n\} \\ &= \{x \in A : f(x) < k/n\} \setminus \{x \in A : f(x) < (k-1)/n\}. \end{aligned}$$

By hypothesis,  $E(n, k)$  is a Jordan region and so  $\chi_{E(n, k)}$  is integrable. Also, for each  $n$ ,  $A = \cup_{k=1}^n E(n, k)$ . We define an integrable function  $g_n$  on  $A$  by

$$g_n(x) = \sum_{k=1}^n \frac{k-1}{n} \chi_{E(n, k)}.$$

That is,

$$g_n(x) = \frac{k-1}{n} \quad \text{if} \quad x \in E(n, k).$$

Since  $g_n$  is a linear combination of integrable functions, it is integrable. Also

$$0 \geq g(x) - g_n(x) < k/n - (k-1)/n = 1/n \quad \text{if} \quad x \in E(n, k).$$

Since every  $x \in A$  is in  $E(n, k)$  for some  $k$ , we conclude that

$$|g(x) - g_n(x)| < 1/n \quad \text{for all} \quad x \in A.$$

This implies that  $\{g_n\}$  converges uniformly to  $g$  on  $A$ . By the previous theorem,  $g$  is integrable on  $A$ . Hence,  $f$  is integrable on  $A$ .

### Exercise Set 10.3

1. Prove that the integral  $\int_R f_A(x) dV(x)$  that appears in Definition 10.3.4 does not depend on the choice of  $R$  as long as  $R$  contains  $A$ .
2. Prove Theorem 10.3.6. You may use the result of Exercise 10.1.6.
3. Prove Theorem 10.3.8.
4. Prove Theorem 10.3.9.
5. Prove Theorem 10.3.10.

6. Prove that if  $A$  and  $B$  are Jordan regions with  $B \subset A$  and  $f$  is a non-negative integrable function on  $A$ , then  $\int_B f(x)dV(x) \leq \int_A f(x)dV(x)$ .
7. Prove that if  $f$  is an integrable function on a Jordan region  $A$ , then  $|f|$  is integrable and

$$\left| \int_A f(x)dV(x) \right| \leq \int_A |f(x)|dV(x).$$

8. Let  $A$  be a Jordan region and  $f$  an integrable function on  $A$ . For each  $x \in A$  define  $f^+(x)$  and  $f^-(x)$  by

$$f^+(x) = \max\{f(x), 0\} \quad \text{and} \quad f^-(x) = \max\{-f(x), 0\} = (-f(x))^+.$$

Prove that  $f^+$  and  $f^-$  are non-negative functions on  $A$  with  $f = f^+ - f^-$  and  $|f| = f^+ + f^-$ . Then prove that  $f^+$  and  $f^-$  are integrable.

9. Prove that if  $f$  is a bounded function on a set  $A$  of volume 0, then  $f$  is integrable on  $A$  and  $\int_A f(x)dV(x) = 0$ .
10. Let  $U$  be an open Jordan region and  $\{K_n\}$  an increasing sequence of compact Jordan subsets of  $U$  such that  $U = \cup_n K_n^\circ$ . Prove that, for each integrable function  $f$  on  $U$ ,

$$\int_U f(x)dV(x) = \lim_n \int_{K_n} f(x)dx.$$

11. Prove that if  $U$  is an open Jordan region, then there always exists a sequence  $\{K_n\}$  like the one in the previous exercise.
12. Let  $A$  be a Jordan region and  $f$  an integrable function on  $A$ . The average value of  $f$  on  $A$  is defined to be the number

$$\text{avg}(f, A) = \frac{1}{V(A)} \int_A f(x)dV(x).$$

If  $A$  is compact and connected and  $f$  is continuous on  $A$ , prove that there is a point  $x_0 \in A$  at which  $f(x_0) = \text{avg}(f, A)$ .

13. Suppose  $A$  is a Jordan region in  $\mathbb{R}^d$  and  $g_k$  is an integrable function on  $A$  for  $k = 1, 2, \dots$ . Prove that if

$$g(x) = \sum_{k=1}^{\infty} g_k(x),$$

where this series converges uniformly on  $A$ , then  $g$  is integrable and

$$\int_A g(x)dV(x) = \sum_{k=1}^{\infty} \int_A g_k(x)dV(x).$$

14. Prove that the function  $g$  on  $\mathbb{R}^2$ , defined by

$$g(x, y) = \sum_{k=1}^{\infty} \frac{1}{k^2} \sin(kx) \sin(ky),$$

is integrable on any Jordan region in  $\mathbb{R}^2$ .

## 10.4 Iterated Integrals

Integrals of functions of a single variable may be calculated exactly in a wide range of situations. The theorem that makes this possible is the Fundamental Theorem of Calculus. We calculate an integral by finding (if we can) an antiderivative for the integrand, then evaluating at the endpoints and subtracting. Fortunately, there is a theorem which often makes it possible to use this same procedure to compute integrals in several variables. This theorem is Fubini's Theorem, and it tells us that, in many situations, we may calculate an integral in several variables by integrating with respect to one variable at a time.

### An Additivity Lemma

We begin our discussion of Fubini's Theorem with a lemma that will play an important role in the proof.

Theorem 10.3.9 says that if  $A$  and  $B$  are Jordan regions with  $V(A \cap B) = 0$ , then the integral of an integrable function over  $A \cup B$  is the sum of the integrals of the function over  $A$  and over  $B$ . If  $f$  is not integrable, only bounded, the analogous result holds for the upper integral of  $f$  and for the lower integral of  $f$ . We will only need the following special case of this result.

**Lemma 10.4.1.** *Suppose  $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d]$  is an aligned rectangle in  $\mathbb{R}^d$  and  $f$  is a bounded function on  $R$ . Suppose that  $R = R_1 \cup R_2$ , where  $R_1$  and  $R_2$  are obtained from  $R$  by partitioning one of the intervals  $[a_j, b_j]$  into two adjacent subintervals  $[a_j, c]$ ,  $[c, b_j]$  and leaving the others alone. Then*

$$\int_{\underline{R}} f(x) dV(x) = \int_{\underline{R}_1} f(x) dV(x) + \int_{\underline{R}_2} f(x) dV(x),$$

and

$$\overline{\int}_R f(x) dV(x) = \overline{\int}_{R_1} f(x) dV(x) + \overline{\int}_{R_2} f(x) dV(x).$$

*Proof.* The proof of this is exactly the same as the proof of the interval additivity theorem for the single variable integral (Theorem 5.2.7). The key to the proof is that a partition  $P_1$  of  $R_1$  and a partition  $P_2$  of  $R_2$ , together form a partition  $P$  of  $R$ , and this partition has the property that

$$L(f, P) = L(f, P_1) + L(f, P_2) \quad \text{and} \quad U(f, P) = U(f, P_1) + U(f, P_2).$$

Furthermore, each partition of  $R$  has a refinement which is of this form.  $\square$

### Fubini's Theorem

Let  $S$  be an aligned rectangle in  $\mathbb{R}^p$  and  $T$  an aligned rectangle in  $\mathbb{R}^q$ . Let  $f$  be a bounded function on the aligned rectangle  $R = S \times T$  in  $\mathbb{R}^{p+q}$ . We will denote the typical point of  $S \times T$  by  $(x, y)$  where  $x \in S$  and  $y \in T$ .

If we hold  $x \in S$  fixed and consider  $f(x, y)$  as a function of  $y \in T$ , then this function may or may not be integrable on  $T$ . In general, it will be integrable for some values of  $x$  and not for others. However, the upper and lower integrals of this function of  $y$  exist for all  $x$  and yield new functions of  $x$  on  $S$  which also have upper and lower integrals. The key step in the proof of Fubini's Theorem is the following theorem which relates these to the upper and lower integrals of  $f$  over  $S \times T$ .

**Theorem 10.4.2.** *With  $S, T$ , and  $f$  as above,*

$$\begin{aligned} \int_{\underline{S \times T}} f(x, y) dV(x, y) &\leq \int_{\underline{S}} \int_{\underline{T}} f(x, y) dV(y) dV(x) \\ &\leq \overline{\int_S} \overline{\int_T} f(x, y) dV(y) dV(x) \leq \overline{\int_{S \times T}} f(x, y) dV(x, y). \end{aligned} \quad (10.4.1)$$

*Proof.* The typical partition of  $S \times T$  has the form  $P \times Q$ , where  $P$  is a partition of  $S$  and  $Q$  is a partition of  $T$ . Recall that a partition of  $S$  consists of a partition of each of the intervals whose cartesian product is  $S$ , while a partition of  $T$  consists of a partition of each of the intervals whose cartesian product is  $T$ . Taken together, these partitions yield partitions of each of the intervals whose product is  $S \times T$ . It is this partition of  $S \times T$  that we denote by  $P \times Q$ .

Let  $\{S_i\}_{i=1}^n$  be a list of the subrectangles of  $S$  determined by the partition  $P$  and  $\{T_j\}_{j=1}^m$  be a list of the subrectangles of  $T$  determined by the partition  $Q$ . Then  $\{S_i \times T_j\}_{i,j=1}^{n,m}$  is a list of the subrectangles for the partition  $P \times Q$ . Let

$$M_{ij} = \sup_{S_i \times T_j} f \quad \text{and} \quad m_{ij} = \inf_{S_i \times T_j} f.$$

Then, for  $x \in S_i$ , Theorem 10.1.11 implies

$$m_{ij}V(T_j) \leq \int_{\underline{T_j}} f(x, y) dV(y) \leq \overline{\int_{T_j}} f(x, y) dV(y) \leq M_{ij}V(T_j).$$

Applying Theorem 10.1.11 again, in the variable  $x$ , implies

$$\begin{aligned} m_{ij}V(S_i)V(T_j) &\leq \int_{\underline{S_i}} \int_{\underline{T_j}} f(x, y) dV(y) dV(x) \\ &\leq \overline{\int_{S_i}} \overline{\int_{T_j}} f(x, y) dV(y) dV(x) \leq M_{ij}V(S_i)V(T_j). \end{aligned}$$

If we sum this inequality over  $i$  and  $j$ , note that  $V(S_i)V(T_j) = V(S_i \times T_j)$ , and

make repeated use of the preceding lemma, the result is

$$\begin{aligned} L(f, P \times Q) &\leq \int_{\underline{S}} \int_{\underline{T}} f(x, y) dV(y) dV(x) \\ &\leq \int_{\overline{S}} \int_{\overline{T}} f(x, y) dV(y) dV(x) \leq U(f, P \times Q) \end{aligned}$$

Since the two expressions in the middle of this inequality give an upper bound for  $\{L(f, P \times Q)\}$  and a lower bound for  $\{U(f, P \times Q)\}$ , and since the least upper bound for  $\{L(f, P \times Q)\}$  is  $\int_{\underline{S \times T}} f(x, y) dV(x, y)$  and the greatest lower bound for  $\{U(f, P \times Q)\}$  is  $\int_{\overline{S \times T}} f(x, y) dV(x, y)$ , we conclude that (10.4.1) holds.  $\square$

In the case where  $f$  is integrable on  $S \times T$ , this yields Fubini's Theorem:

**Theorem 10.4.3.** *Let  $S$  and  $T$  be aligned rectangles in  $\mathbb{R}^p$  and  $\mathbb{R}^q$ , respectively, and let  $f$  be an integrable function on  $S \times T$ , then*

$$\begin{aligned} \int_{S \times T} f(x, y) dV(x, y) \\ = \int_{\underline{S}} \int_{\underline{T}} f(x, y) dV(y) dV(x) = \int_{\overline{S}} \int_{\overline{T}} f(x, y) dV(y) dV(x). \end{aligned} \quad (10.4.2)$$

Furthermore, if  $f(x, y)$  is an integrable function of  $y$  on  $T$  for each fixed  $x \in S$ , then  $\int_T f(x, y) dV(y)$  is an integrable function of  $x$  on  $S$ , and

$$\int_{S \times T} f(x, y) dV(x, y) = \int_S \int_T f(x, y) dV(y) dV(x). \quad (10.4.3)$$

*Proof.* If  $f$  is integrable on  $S \times T$ , then the first and last expressions in the string of inequalities (10.4.1) are equal. Hence, each of the inequalities in (10.4.1) is actually an equality in this case. This proves (10.4.2).

If  $f(x, y)$  is an integrable function of  $y$  on  $T$  for each  $x \in S$ , then

$$\int_{\underline{T}} f(x, y) dV(y) = \int_{\overline{T}} f(x, y) dV(y) = \int_T f(x, y) dV(y)$$

for each  $x \in S$ . Then (10.4.2) implies that

$$\int_{\underline{S}} \int_T f(x, y) dV(y) dV(x) = \int_{\overline{S}} \int_T f(x, y) dV(y) dV(x),$$

which means that  $\int_T f(x, y) dV(y)$  is an integrable function of  $x$ . Then (10.4.2) implies (10.4.3).  $\square$

**Remark 10.4.4.** In (10.4.2) there is nothing special about the order in which the iterated integrals are taken. The theorem is equally valid if we integrate first with respect to  $x$  and then with respect to  $y$ . Of course, for the analogue of (10.4.3) to be valid with the order of integration reversed, we must assume that  $f(x, y)$  is an integrable function of  $x$  for each fixed  $y$ .

This leads to the following consequence of Fubini's Theorem.

**Theorem 10.4.5.** *Let  $S$  and  $T$  be aligned rectangles in  $\mathbb{R}^p$  and  $\mathbb{R}^q$ , respectively, and let  $f(x, y)$  be an integrable function on  $S \times T$  which is also integrable as a function of  $x$  for each fixed  $y$  and integrable as a function of  $y$  for each fixed  $x$ . Then  $\int_S f(x, y)dV(x)$  is an integrable function of  $y$  on  $T$  and  $\int_T f(x, y)dV(y)$  is an integrable function of  $x$  on  $S$ , and*

$$\begin{aligned} \int_{S \times T} f(x, y)dV(x, y) \\ = \int_S \int_T f(x, y)dV(y)dV(x) = \int_T \int_S f(x, y)dV(x)dV(y). \end{aligned} \quad (10.4.4)$$

Note that the integrability conditions in this theorem will all be satisfied if  $f$  is a continuous function on the rectangle  $S \times T$ .

The ability to reverse the order of integration in an iterated integral is a real advantage, as the following example shows.

**Example 10.4.6.** Find  $\int_0^1 \int_0^{\sqrt{\pi}} y^3 \sin(xy^2) dy dx$ .

**Solution:** Computing the inside integral looks difficult. However, if we reverse the order of integration, the inside integral is just  $\int_0^1 y^3 \sin(xy^2) dx = y - y \cos(y^2)$  and the iterated integral becomes

$$\int_0^{\sqrt{\pi}} \int_0^1 y^3 \sin(xy^2) dx dy = \int_0^{\sqrt{\pi}} (y - y \cos(y^2)) dy = \pi/2.$$

## Iterated Integrals over Non-rectangular Regions

A great advantage of integrals in one real variable is that we can often use the Fundamental Theorem of Calculus to calculate them. In order to take advantage of this, we would like to interpret an integral over a Jordan region  $A$  in  $R^d$  as the result of repeated applications of integration in one variable. Fubini's Theorem is the tool which allows us to do this.

The issue is complicated by the fact that we wish to integrate over a Jordan region, rather than over a rectangle. To do this, we replace the function  $f$  to be integrated with  $f_A$ , where  $f$  is an integrable function on  $A$  (then  $f_A$  is an integrable function on any aligned rectangle containing  $A$ ). We then attempt to apply Fubini's Theorem repeatedly to express the integral of  $f_A$  over a rectangle containing  $A$  as the result of a succession of single variable integrations. In order for this to work,  $A$  must have a special form.

We begin with a result which is a direct application of Fubini's Theorem. It will form the basis for the induction argument in the proof of our main theorem. It concerns the case of an integral over a compact Jordan region  $A \subset \mathbb{R}^{k+1}$ , which is constructed as follows: Suppose there is a compact Jordan region  $B \subset R^k$  such that  $A$  has the form

$$A = \{(x, t) : x \in B, \text{ and } \psi(x) \leq t \leq \phi(x)\},$$

where  $\psi$  and  $\phi$  are continuous functions on  $B$ . In this case,  $f_A(x, t) = 0$  if  $x \notin B$  or if  $t \notin [\psi(x), \phi(x)]$ . Then (10.4.3) implies

**Theorem 10.4.7.** *With  $A$ ,  $B$ ,  $\psi$ , and  $\phi$  as above and  $f$  an integrable function on  $A$ ,*

$$\int_A f(x, t) dV(x, t) = \int_B \int_{\psi(x)}^{\phi(x)} f(x, t) dt dV(x).$$

*provided  $f(x, t)$  is an integrable function of  $t$  on  $[\psi(x), \phi(x)]$  for each  $x \in B$ .*

If we write

$$g(x) = \int_{\psi(x)}^{\phi(x)} f(x, t) dt,$$

then the above theorem reduces the problem of computing  $\int_A f(x, t) dV(x, t)$  to the problem of computing the lower dimensional integral  $\int_B g(x) dV(x)$ . This is the basis for the induction argument in the proof of Theorem 10.4.9. Before we state and prove that theorem, we need the following technical result.

**Theorem 10.4.8.** *Let  $A$ ,  $B$ ,  $\psi$ ,  $\phi$ , and  $f$  be as in the previous theorem. If  $f$  is continuous on  $A$ , then the function*

$$g(x) = \int_{\psi(x)}^{\phi(x)} f(x, t) dt$$

*is continuous on  $B$ .*

*Proof.* Since  $A$  is compact and  $f$  continuous on  $A$ ,  $|f|$  has a maximum on  $A$ . Let  $M_1$  be a positive number greater than or equal to this maximum.

Since  $\psi$  and  $\phi$  are continuous on  $B$  and  $\psi(x) \leq \phi(x)$ , the non-negative function  $\phi - \psi$  is also continuous and, hence, has a maximum. Let  $M_2$  be a positive number greater than or equal to this maximum.

Let  $x_0$  be a point of  $B$ . We will prove that  $g$  is continuous at  $x_0$ . We need to consider two cases: (1)  $\phi(x_0) - \psi(x_0) = 0$ , and (2)  $\phi(x_0) - \psi(x_0) > 0$ .

In case (1),  $g(x_0) = 0$ . Furthermore, the continuity of  $\phi - \psi$  implies that, given  $\epsilon > 0$ , there is a  $\delta > 0$  such that

$$\phi(x) - \psi(x) < \frac{\epsilon}{M_1} \quad \text{whenever} \quad \|x - x_0\| < \delta.$$

Then,

$$|g(x) - g(x_0)| = |g(x)| = \left| \int_{\psi(x)}^{\phi(x)} f(x, t) dt \right| \leq M_1(\phi(x) - \psi(x)) < \epsilon.$$

This completes the proof in case (1).

In case (2), we have  $\phi(x_0) - \psi(x_0) > 0$ . Given  $\epsilon > 0$ , we may choose a positive number  $\rho$  such that

$$\rho < \frac{1}{2}(\phi(x_0) - \psi(x_0)) \quad \text{and} \quad \rho < \frac{\epsilon}{12M_1}.$$

We then set  $a = \psi(x_0) + \rho$  and  $b = \phi(x_0) - \rho$ . Since  $\psi$  and  $\phi$  are continuous at  $x_0$ , there is a  $\delta > 0$  such that

$$|\psi(x) - \psi(x_0)| < \rho \quad \text{and} \quad |\phi(x) - \phi(x_0)| < \rho,$$

whenever  $x \in B$  and  $\|x - x_0\| < \delta$ . For each such  $x$ , we have

$$\psi(x) < a < b < \phi(x).$$

Also, each of the intervals  $[\psi(x), a]$  and  $[b, \phi(x)]$  has length less than  $2\rho$ , and so the sum of their lengths is less than  $4\rho$ .

Since  $f$  is continuous on the compact set  $A$ , it is uniformly continuous on  $A$ . Hence, we may choose  $\delta$  small enough that it is also true that

$$|f(x_1, t_1) - f(x_2, t_2)| < \frac{\epsilon}{3M_2},$$

whenever  $(x_1, t_1)$  and  $(x_2, t_2)$  are in  $A$  and  $\|(x_1, t_1) - (x_2, t_2)\| < \delta$ . In particular,

$$|f(x, t) - f(x_0, t)| < \frac{\epsilon}{3M_2} \quad \text{whenever} \quad \|x - x_0\| < \delta,$$

provided that  $(x, t)$  and  $(x_0, t)$  are both in  $A$ . Then,

$$\begin{aligned} |g(x) - g(x_0)| &= \left| \int_{\psi(x)}^{\phi(x)} f(x, t) dt - \int_{\psi(x_0)}^{\phi(x_0)} f(x_0, t) dt \right| \\ &\leq \left| \int_{\psi(x)}^{\phi(x)} f(x, t) dt - \int_a^b f(x, t) dt \right| + \left| \int_a^b (f(x, t) - f(x_0, t)) dt \right| \\ &\quad + \left| \int_a^b f(x_0, t) dt - \int_{\psi(x_0)}^{\phi(x_0)} f(x_0, t) dt \right| \\ &\leq 4\rho M_1 + \frac{\epsilon}{3M_2} M_2 + 4\rho M_1 = \epsilon. \end{aligned}$$

This completes the proof in case (2).  $\square$

We can now state and prove the form of Fubini's Theorem which represents an integral over a Jordan region as the result of repeated single variable integrations.

**Theorem 10.4.9.** *Suppose  $f$  is an integrable function on the closed Jordan region  $A$ . Suppose also that  $A$  is the set of  $x = (x_1, \dots, x_d) \in R^d$  which satisfy the inequalities*

$$\begin{aligned} \psi_1 &\leq x_1 \leq \phi_1, \\ \psi_2(x_1) &\leq x_2 \leq \phi_2(x_1) \\ &\vdots \\ \psi_d(x_1, \dots, x_{d-1}) &\leq x_d \leq \phi_d(x_1, \dots, x_{d-1}), \end{aligned}$$

where  $\psi_1$  and  $\phi_1$  are numbers and  $\psi_j(x_1, \dots, x_{j-1})$  and  $\phi_j(x_1, \dots, x_{j-1})$  are continuous functions on the set of  $(x_1, \dots, x_{j-1})$  which satisfy the inequalities in this list that precede the  $j$ th one. Then

$$\begin{aligned} \int_A f(x) dV(x) \\ = \int_{\psi_1}^{\phi_1} \int_{\psi_2(x_1)}^{\phi_2(x_1)} \cdots \int_{\psi_d(x_1, \dots, x_{d-1})}^{\phi_d(x_1, \dots, x_{d-1})} f(x_1, \dots, x_d) dx_d \cdots dx_1. \end{aligned} \quad (10.4.5)$$

provided that each of the successive iterated integrals exists. This condition is satisfied if  $f$  is continuous on  $A$ .

*Proof.* We prove this by induction on  $d$ . If  $d = 1$ , then there is nothing to prove, since the two sides of (10.4.5) are the same integral over an interval in this case.

Now suppose the theorem is true in dimension  $d - 1$ . To complete the proof we need to prove that it is then true in dimension  $d$ . Let  $A$  be a Jordan region defined by  $d$  inequalities as in the hypothesis of the theorem and let  $f$  be an integrable function on  $A$ . Let  $B$  be the set defined by the first  $d - 1$  of these inequalities. Then  $A$ ,  $B$ , and  $f$  satisfy the conditions of Theorem 10.4.7. Hence, if  $x = (\tilde{x}, x_d)$  where  $\tilde{x} = (x_1, \dots, x_{d-1})$ , and  $f(\tilde{x}, x_d)$  is an integrable function of  $x_d$  on  $[\psi_d(\tilde{x}), \phi_d(\tilde{x})]$  for each  $\tilde{x} \in B$ , then this theorem implies that  $g(\tilde{x}) = \int_{\psi(\tilde{x})}^{\phi(\tilde{x})} f(\tilde{x}, x_d) dx_d$  is integrable on  $B$  and

$$\int_A f(x) dV(x) = \int_B \int_{\phi_d(\tilde{x})}^{\psi_d(\tilde{x})} f(\tilde{x}, x_d) dx_d dV(\tilde{x}). \quad (10.4.6)$$

Now the set  $B$  and the function  $g$  satisfy the conditions of our theorem in dimension  $d - 1$ . Since we are assuming the theorem is true in dimension  $d - 1$ , we have

$$\int_B g(\tilde{x}) dV(\tilde{x}) = \int_{\psi_1}^{\phi_1} \int_{\psi_2(x_1)}^{\phi_2(x_1)} \cdots \int_{\psi_d(x_1, \dots, x_{d-2})}^{\phi_d(x_1, \dots, x_{d-2})} g(x_1, \dots, x_{d-1}) dx_{d-1} \cdots dx_1.$$

If we combine this with (10.4.6), the result is (10.4.5).

It remains to prove that each of the successive iterated integrals exists if  $f$  is continuous on  $A$ . However, this also follows from induction on  $d$ . It is clearly true if  $d = 1$  since a continuous function on an interval is integrable. Assuming it is true in dimension  $d - 1$ , then if  $f$  is continuous on an  $A$  of the form describe in the theorem in dimension  $d$ , we conclude that  $f$  is continuous, hence, integrable in its last variable and the function  $g$ , defined by integrating in this last variable is continuous on the corresponding set  $B$  by Theorem 10.4.8. Since we are assuming the result to be true in dimension  $d - 1$ , we conclude that each of the successive iterated integrals of  $g$  exists. Hence, the same thing is true of  $f$ .  $\square$

**Example 10.4.10.** Find  $\int_A xyz dV(x, y, z)$  if  $A$  is the Jordan region in  $\mathbb{R}^3$  defined by the inequalities  $0 \leq x \leq 1$ ,  $0 \leq y \leq x$ ,  $0 \leq z \leq 1 - x^2$ .

**Solution:** According to the previous theorem,

$$\begin{aligned} \int_A xyz \, dV(x, y, z) &= \int_0^1 \int_0^x \int_0^{1-x^2} xyz \, dz \, dy \, dx \\ &= \int_0^1 \int_0^x \frac{1}{2} xy(1-x^2)^2 \, dy \, dx \\ &= \int_0^1 \frac{1}{4} x^3(1-x^2)^2 \, dx = \frac{1}{4} \int_0^1 (x^3 - 2x^5 + x^7) \, dx \\ &= \frac{1}{4} \left( \frac{1}{4} - \frac{1}{3} + \frac{1}{8} \right) = \frac{1}{96}. \end{aligned}$$

### Exercise Set 10.4

- Find the integral of the function  $g$  of Exercise 10.3.14 over the square  $[-\pi, \pi] \times [-\pi, \pi]$ .
- Evaluate  $\int_0^1 \int_0^1 \frac{y^3 x}{(1+y^2 x^2)^2} \, dy \, dx$ .
- Find the area of the triangle  $\Delta$  with vertices at  $(0, 0)$ ,  $(a, 0)$ ,  $(a, b)$  by calculating  $\int_{\Delta} 1 \, dV(x, y)$  (use Theorem 10.4.9).
- Calculate the area of a disc of radius one by representing it as the integral of 1 over the disc, expressing this integral as an iterated integral, and then evaluating this iterated integral.
- Interpret the iterated integral  $\int_0^1 \int_{x^2}^x (x^2 + y^2) \, dy \, dx$  as an integral of  $x^2 + y^2$  over a certain Jordan region in  $\mathbb{R}^2$ . This, in turn, is equal to a certain iterated integral, first with respect to  $x$  and then with respect to  $y$ . Describe this integral and then evaluate it.
- Write down an integral in  $\mathbb{R}^3$  which represents the volume of a sphere of radius 1. Then express this as a triple iterated integral. You do not need to evaluate this integral.
- Find  $\int_A x \, dV(x, y, z)$  if  $A$  is defined by the inequalities

$$0 \leq x \leq 1, \quad 0 \leq y \leq x^2, \quad 0 \leq z \leq x + y.$$

- Show that if  $f$  and  $g$  are continuous real valued functions on a Jordan region  $B \subset \mathbb{R}^d$  and  $g(x) \leq f(x)$  for all  $x \in B$ , then the Jordan region  $A = \{(x, t) \in \mathbb{R}^{d+1} : x \in B \text{ and } g(x) \leq t \leq f(x)\}$  of Exercise 10.2.12 has volume

$$V(A) = \int_B (f(x) - g(x)) \, dV(x).$$

9. Prove that if  $A$  is any bounded subset of  $\mathbb{R}^p$  and  $B$  is a subset of  $\mathbb{R}^q$  of volume 0, then  $A \times B$  is a subset of  $\mathbb{R}^{p+q}$  of volume 0. Use this to prove that the Cartesian product  $A \times B$  of two Jordan regions is a Jordan region.
10. Use Fubini's Theorem and the previous exercise to prove that if  $A \subset \mathbb{R}^p$  and  $B \subset \mathbb{R}^q$  are Jordan regions, then  $V(A \times B) = V(A)V(B)$ .
11. Suppose  $A$  is a compact Jordan region in  $\mathbb{R}^p$ ,  $B$  is a compact subset of  $\mathbb{R}^q$ , and  $f$  is a continuous function on  $B \times A$ . Prove that  $\int_A f(x, y) dV(y)$  is a continuous function of  $x$  on  $B$ . Hint: this is similar to but not exactly the same as Theorem 10.4.8.
12. Prove that if  $f(t, x)$  is a continuous function on  $I \times A$ , where  $I$  is an open interval in  $\mathbb{R}$  and  $A$  is a compact Jordan region in  $\mathbb{R}^d$ , and if  $\frac{\partial f}{\partial t}(t, x)$  exists and is continuous on  $I \times A$ , then

$$\frac{d}{dt} \int_A f(t, x) dV(x) = \int_A \frac{\partial f}{\partial t}(t, x) dV(x).$$

Hint: fix  $t$  and consider the function

$$g(h, x) = \begin{cases} \frac{f(t+h, x) - f(t, x)}{h} & \text{if } h \neq 0 \\ \frac{\partial f}{\partial t}(t, x) & \text{if } h = 0. \end{cases}$$

Show that this is a continuous function of  $(h, x)$  on  $J \times A$  for some interval  $J$  containing 0 (the Mean Value Theorem is useful in proving this). Then apply the preceding exercise.

## 10.5 The Change of Variables Formula

Recall the substitution formula (Theorem 5.3.6) from Chapter 5:

$$\int_a^b f(g(t))g'(t) dt = \int_{g(a)}^{g(b)} f(u) du.$$

Here, if  $I = [a, b]$  and  $J = g(I)$ , then  $f$  is assumed continuous on  $J$  and  $g$  is assumed differentiable with an integrable derivative on  $I$ .

This can be thought of as a change of variables formula, where  $u = g(t)$  is the transformation from the variable  $t$  to the variable  $u$ , and the integral formula relates the integral of  $f$  as a function of  $u$  to an integral involving the composite function  $f \circ g$  as a function of  $t$ . The formula requires an extra factor  $g'(t)$  in the integrand of the latter integral. This is related to how the transformation  $g$  changes lengths.

In this section we will derive a similar formula for integrals in several variables. In this case, the extra factor that is needed measures how the transformation changes volume.

## Factorization of Matrices

We begin by studying how a linear transformation effects the volume of a Jordan region. The simple way to do this is to factor a given linear transformation as a product of elementary linear transformations whose effect on volume is easy to determine. Such a factorization is given by the process of Gauss elimination (row reduction). The elementary linear transformations in this factorization correspond to the elementary matrices as described below.

The elementary  $d \times d$  matrices are of three types:

1. The interchange matrices  $E_{ij}$ . For  $i \neq j$ , the interchange matrix  $E_{ij}$  is obtained from the identity matrix by interchanging its  $i$ th and  $j$ th rows.
2. The shear matrices  $S_{ij}$ . For  $i \neq j$  the shear matrix  $S_{ij}$  is obtained from the identity matrix by adding its  $j$ th row to its  $i$ th row – that is, by adding a 1 to the  $ij$  position in the identity matrix.
3. The scale matrices  $T_i(a)$ . For  $i = 1, \dots, d$  and  $a \neq 0$ ,  $T_i(a)$  is obtained from the identity matrix by multiplying its  $i$ th row by the scalar  $a$ . – that is, it is the matrix that is  $a$  in the  $i$ th position on the main diagonal, 1 in the other positions on the main diagonal and 0 in all other positions.

Note that if  $A$  is any  $d \times d$  matrix, then  $E_{ij}A$  is the result of interchanging the  $i$ th and  $j$ th rows in  $A$  and leaving the other rows unchanged,  $S_{ij}A$  is the result of adding the  $j$ th row of  $A$  to its  $i$ th row and leaving all but the  $i$ th row unchanged, while  $T_i(a)A$  is the result of multiplying the  $i$ th row of  $A$  by  $a$  and leaving the other rows unchanged.

The process of Gauss elimination is that of successively multiplying a matrix  $A$  on the left by elementary matrices until what is left is a matrix of reduced row echelon form. In the case of a non-singular matrix  $A$  its reduced row echelon form is just the identity matrix. Thus, for each non-singular  $d \times d$  matrix  $A$  there is a matrix  $B$  which is a product of elementary matrices and satisfies  $BA = I$ . Then

$$A = B^{-1}.$$

Note that the inverse of an elementary matrix is an elementary matrix or a product of elementary matrices (Exercise 10.5.1) and so  $B^{-1}$  is also a product of elementary matrices. Thus, we have proved

**Theorem 10.5.1.** *Each non-singular  $d \times d$  matrix  $A$  is a product of matrices of the form  $E_{ij}, S_{ij}, T_i(a)$ .*

The determinants of the elementary matrices are easily calculated.

**Theorem 10.5.2.** *For each  $i$  and each  $j \neq i$  we have  $\det E_{ij} = 1$ ,  $\det S_{ij} = 1$ , and  $\det T_i(a) = a$ .*

Since the determinant is multiplicative ( $\det AB = \det A \det B$  for all pairs  $A, B$  of  $d \times d$  matrices), it follows that the determinant of a given non-singular matrix  $A$  is just the product of the scale factors  $a$  that appear in its factorization as a product of elementary matrices.

## Linear Transformations and Volume

We wish to understand how the volume of a Jordan region is effected by a linear transformation. Some linear transformations clearly have no effect on volume. A transformation that takes each aligned rectangle to an aligned rectangle of the same volume has no effect on the volume of a Jordan region. The elementary interchanges  $E_{ij}$  have this property. The shear matrices  $S_{ij}$  also preserve volumes of Jordan regions, but the proof of this fact is a little more complicated.

**Theorem 10.5.3.** *A shear transformation  $S_{ij}$  takes a Jordan region to a Jordan region of the same volume.*

*Proof.* The shear matrix  $S_{12}$  on  $\mathbb{R}^2$  is the matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

It takes the aligned rectangle  $[a, b] \times [c, d]$ , which has vertices  $(a, c)$ ,  $(b, c)$ ,  $(b, d)$ , and  $(a, d)$  to the parallelogram with vertices  $(a + c, c)$ ,  $(b + c, c)$ ,  $(b + c, d)$ , and  $(a + d, d)$ . This parallelogram has base of length  $(b + c) - (a + c) = b - a$  and height  $d - c$ . Thus, its area is  $(b - a)(d - c)$  (Exercise 10.2.11), which is the same as the volume of the original rectangle.

In general, an aligned rectangle  $R$  in  $\mathbb{R}^d$  for  $d > 2$  has the form  $S \times T$  where  $S$  is an aligned rectangle in  $\mathbb{R}^2$  and  $T$  is an aligned rectangle in  $\mathbb{R}^{d-2}$ . The shear transformation  $S_{12}$  on  $\mathbb{R}^d$  sends this to  $P \times T$  where, by the above discussion,  $P$  is a parallelogram with the same area as  $S$ . It follows from this and Exercise 10.4.10 that  $S_{12}$  sends  $R$  to a Jordan region with the same volume as  $R$ . Since, for any  $i \neq j$ ,  $S_{ij}$  is just  $S_{12}$  composed with some elementary interchanges, it follows that it also takes an aligned rectangle to a Jordan region with the same volume.

Let  $A$  be a Jordan region,  $R$  an aligned rectangle containing  $A$ , and  $P$  a partition of  $R$ . Let  $R_1, R_2, \dots, R_n$  be a list of the subrectangles of  $R$  determined by the partition  $P$ . Set

$$E = \bigcup \{R_k : R_k \subset A\}$$

$$F = \bigcup \{R_k : R_k \cap A \neq \emptyset\}.$$

Then  $U(\chi_A, P) = V(F)$  and  $L(\chi_A, P) = V(E)$ . Since  $A$  is a Jordan region, given  $\epsilon > 0$ , there is a partition  $P$  such that  $V(F) - V(E) < \epsilon$ . Of course, regardless of how the partition is chosen

$$V(E) \leq V(A) \leq V(F). \tag{10.5.1}$$

Note  $S_{ij}F$  is the union of those  $S_{ij}R_k$  such that  $R_k \cap A \neq \emptyset$ , and any two of these sets meet (if at all) in a set of volume 0. Since  $V(S_{ij}R_k) = V(R_k)$ , we conclude that

$$V(S_{ij}F) = V(F).$$

A similar argument shows that

$$V(S_{ij}E) = V(E).$$

Hence,

$$V(E) = V(S_{ij}E) \leq \underline{V}(S_{ij}A) \leq \overline{V}(S_{ij}A) \leq V(S_{ij}F) = V(F). \quad (10.5.2)$$

Since,  $V(F) - V(E) < \epsilon$ , we conclude that

$$\overline{V}(S_{ij}A) - \underline{V}(S_{ij}A) < \epsilon.$$

Since  $\epsilon$  was arbitrary, this difference is actually 0. This proves that  $S_{ij}A$  is a Jordan region. That it has the same volume as  $A$  follows from (10.5.1) and (10.5.2).  $\square$

**Theorem 10.5.4.** *If  $L : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a linear transformation and  $E$  is a Jordan region, then  $L(E)$  is also a Jordan region and  $V(L(E)) = |\det L|V(E)$ , where  $\det L$  denotes the determinant of the matrix corresponding to  $L$ .*

*Proof.* We first note that if this theorem is true for linear transformations  $L_1$  and  $L_2$ , then it is also true for the composition  $L_1 \circ L_2$ , by the following computation:

$$\begin{aligned} V(L_1 \circ L_2(E)) &= |\det L_1|V(L_2(E)) \\ &= |\det L_1||\det L_2|V(E) = |\det L_1L_2|V(E), \end{aligned}$$

since determinant and absolute value are both multiplicative functions.

The elementary interchanges  $E_{ij}$  and shear transformations  $S_{ij}$  do not effect volume and they are matrices of determinant  $\pm 1$ . Thus, the theorem is true for these linear transformations.

The scale matrix  $T_i(a)$  takes each aligned rectangle to an aligned rectangle with edges of the same length as the original except for the  $i$ th edge, which has its length multiplied by  $|a|$ . Hence, each aligned rectangle is sent to an aligned rectangle of volume  $|a|$  times the volume of the original. It follows that  $T_i(a)$  takes a Jordan region to another Jordan region with volume  $|a|$  times the volume of the original. Since  $a = \det T_i(a)$ , the theorem is true for the transformations  $T_i(a)$ .

Since every non-singular  $d \times d$  matrix is a product of interchanges, shear transformations, and scale transformations, the theorem is true for all non-singular linear functions from  $\mathbb{R}^d$  to  $\mathbb{R}^d$ .

If  $L$  is singular, then its determinant is 0. Thus, to finish the proof, we need to show that if  $L$  is a singular linear transformation, then  $L(E) = 0$  for every Jordan region  $E$ . We leave this as an exercise.  $\square$

**Example 10.5.5.** If  $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is the linear transformation with matrix

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

what is the area of the image of the unit disc  $D_1(0, 0)$  under the transformation  $L$ ?

**Solution:** The unit disc has area  $\pi$ . By the previous theorem, its image under  $L$  has area  $|\det L|\pi = 2\pi$ .

**Example 10.5.6.** What is the area of an ellipse, with two vertices at distance 3 from  $(0, 0)$  along the line  $y = x$  and two vertices at distance 2 from  $(0, 0)$  along the line  $y = -x$ ?

**Solution:** This ellipse may be obtained from the unit disc by first applying the transformation with matrix

$$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

and then applying the linear transformation which is rotation through the angle  $\pi/4$ . The first transformation has determinant 6, while the second has determinant 1. Hence the area of the indicated ellipse is  $6\pi$ .

### Smooth Image of a Rectangle

We will prove that, under appropriate conditions, the image of an aligned rectangle under a smooth map is a Jordan region. We first prove the the image of a degenerate rectangle under such a map is a set of volume 0.

**Theorem 10.5.7.** *Let  $\phi$  be a one-to-one smooth transformation from an open set  $U \subset \mathbb{R}^p$  to  $\mathbb{R}^p$  and suppose  $d\phi(x)$  is non-singular at each point of  $U$ . If  $R$  is a degenerate aligned rectangle contained in  $U$ , then  $\phi(R)$  is a set of volume 0 in  $\mathbb{R}^p$ .*

*Proof.* Since  $R$  is degenerate, it is a rectangle of dimension at most  $p - 1$ . We may as well assume that it is contained in  $\mathbb{R}^{p-1} = \{x = (x_1, \dots, x_p) : x_p = 0\}$ . Let  $a$  be a point of  $R$ . We will show first that there is a neighborhood of  $b = \phi(a)$  whose intersection with  $\phi(R)$  has volume 0. If we can do this for each  $a \in R$ , then, since  $\phi(R)$  is compact, we may cover  $\phi(R)$  with finitely many open sets whose intersections with  $\phi(R)$  have volume 0. It follows from this that  $\phi(R)$  itself has volume 0.

Since translations do not effect volume, we may as well assume that  $a$  and  $b = \phi(a)$  are both equal to 0. Also, since applying a non-singular linear transformation does not effect whether or not a set has volume 0, we may replace  $\phi$  by  $(d\phi(0))^{-1}\phi$ . In other words, we may as well assume that  $d\phi(0) = I$  – the identity transformation.

If  $\phi = (\phi_1, \dots, \phi_p)$ , and points of  $\mathbb{R}^p$  are denoted  $(x, y)$  with  $x \in \mathbb{R}^{p-1}$  and  $y \in \mathbb{R}$ , then we define  $g : U \cap \mathbb{R}^{p-1} \rightarrow \mathbb{R}^{p-1}$  by

$$g(x) = (\phi_1(x, 0), \dots, \phi_{p-1}(x, 0)).$$

Then  $dg(0)$  is the upper left  $(p - 1) \times (p - 1)$  subdeterminant of  $d\phi(0)$  and so it too is the identity transformation. The Inverse Function Theorem then implies

that there are neighborhoods  $V$  and  $W$  of 0 in  $\mathbb{R}^{p-1}$  such that  $g$  maps  $V$  onto  $W$  and has a smooth inverse  $g^{-1} : W \rightarrow V$ . Then

$$\phi(g^{-1}(x), 0) = (x, \phi_p \circ g^{-1}(x))$$

for  $x \in W$ . That is, the part of  $\phi(R)$  consisting of points with first coordinate in  $W$  is the graph of the smooth function  $\phi_p \circ g^{-1}$ . It therefore has volume 0 by Example 10.2.11. This completes the proof.  $\square$

**Theorem 10.5.8.** *Let  $\phi : U \rightarrow \mathbb{R}^p$  satisfy the conditions of the previous theorem. If  $R$  is a rectangle in  $U$ , then  $\phi(R)$  is a Jordan region.*

*Proof.* If  $R$  is a rectangle in  $U$ , then its boundary is a union of finitely many rectangles of dimension  $p-1$  – that is, it is the union of finitely many degenerate rectangles. The image of each of these under  $\phi$  has volume 0 by the previous theorem. Hence,  $\phi(\partial R)$  has volume zero. The proof will be complete if we can show that  $\partial\phi(R) = \phi(\partial R)$ .

The image of  $\phi$  is an open set  $V$  by Exercise 9.6.8, and  $\phi : U \rightarrow V$  is one-to-one and onto. Thus,  $\phi$  has an inverse transformation  $\phi^{-1} : V \rightarrow U$  which is a smooth transformation, by the Inverse Function Theorem. It is, in particular, continuous. Since both  $\phi$  and  $\phi^{-1}$  are continuous, a subset  $A \subset U$  is open if and only if its image  $\phi(A) \subset V$  is open. It follows that  $\phi$  takes the interior of  $R$  to the interior of  $\phi(R)$  and, hence, the boundary of  $R$  to the boundary of  $\phi(R)$ .  $\square$

## Integral over the Smooth Image of a Rectangle

Our next objective is to prove the change of variables formula for integration over a rectangle. We will need the following lemma, which says that the relative error in approximating the volume of the image of a rectangle under a smooth map by the volume of its image under the differential of the map can be made arbitrarily small. In the lemma, it is crucial that we don't allow rectangles  $R$  to become too skinny. By this, we mean that we don't want the ratio of the length of the shortest edge of  $R$  to the diameter of  $R$  (greatest distance between two points of  $R$ ) to be too small. We will call this ratio the *aspect ratio* of the rectangle.

**Lemma 10.5.9.** *Let  $\lambda$  and  $K$  be positive constants. Let  $U$  be an open subset of  $\mathbb{R}^p$  and  $\phi : U \rightarrow \mathbb{R}^p$  a smooth one-to-one transformation. Suppose  $d\phi(a)$  is non-singular and  $|\det d\phi(a)| \leq K$  for all  $a \in U$ . Then, given  $\epsilon > 0$ , there is a  $\delta > 0$  such that if  $R$  is a rectangle in  $U$  with diameter less than  $\delta$  and aspect ratio at least  $\lambda$ , then  $|V(\phi(R)) - V(d\phi(a)R)| < \epsilon V(R)$ , where  $a$  is the center of the rectangle  $R$ .*

*Proof.* Let  $R$  be a rectangle in  $U$  with diameter less than a positive number  $\delta$  to be determined below and aspect ratio at least  $\lambda$ . Note that  $\phi(R)$  is a Jordan region, by the previous theorem and, hence, it has volume.

Since translation does not effect volume, we may assume that the center of the rectangle  $R$  is 0 and  $\phi(0) = 0$ . By hypothesis

$$|\det d\phi(0)| \leq K. \quad (10.5.3)$$

If  $0 < \rho < 1$ , then  $(1 + \rho)R$  is the rectangle created from  $R$  by expanding each edge in a symmetric way about its center by the factor  $(1 + \rho)$ . Similarly,  $(1 - \rho)R$  is the rectangle created from  $R$  by shrinking each edge in a symmetric way about its center by the factor  $1 - \rho$ . Also,

$$(1 - \rho)R \subset R \subset (1 + \rho)R,$$

and, since  $d\phi(0)$  is linear,

$$(1 - \rho)d\phi(0)R \subset d\phi(0)R \subset (1 + \rho)d\phi(0)R.$$

Comparing volumes and using (10.5.3) yields,

$$\begin{aligned} V((1 + \rho)d\phi(0)R) - V((1 - \rho)d\phi(0)R) & \\ &= ((1 + \rho)^d - (1 - \rho)^d)V(d\phi(0)R) \\ &= ((1 + \rho)^d - (1 - \rho)^d)|\det d\phi(0)|V(R) \quad (10.5.4) \\ &\leq 2\rho d(1 + \rho)^{d-1}|\det d\phi(0)|V(R) \\ &\leq 2^d \rho d K V(R). \end{aligned}$$

If we choose

$$\rho = \frac{\epsilon}{2^d d K},$$

then it follows from (10.5.4) that

$$V((1 + \rho)d\phi(0)R) - V((1 - \rho)d\phi(0)R) \leq \epsilon V(R).$$

The proof will be complete if we can show that, for small enough  $\delta$ , any rectangle  $R$  containing 0, of diameter less than  $\delta$ , satisfies

$$(1 - \rho)d\phi(0)R \subset \phi(R) \subset (1 + \rho)d\phi(0)R, \quad (10.5.5)$$

since these containments are also satisfied with  $\phi(R)$  replaced by  $d\phi(0)R$ .

If  $x$  is any non-zero vector in  $\mathbb{R}^d$ , then

$$\|x\| = \|(d\phi(0))^{-1}d\phi(0)x\| \leq \|(d\phi(0))^{-1}\| \|d\phi(0)x\|.$$

Thus,  $\|d\phi(0)x\| \geq \|(d\phi(0))^{-1}\|^{-1}\|x\|$ . In other words, if  $L$  is any line segment in  $\mathbb{R}^d$ , then the length of the line segment  $d\phi(0)L$  is at least the factor

$$A = \|(d\phi(0))^{-1}\|^{-1}$$

times the length of  $L$ . It follows that the distance from  $d\phi(0)R$  to the complement of  $(1 + \rho)d\phi(0)R$  is at least  $A\rho r$ , where  $r$  is one half the length of the

shortest edge of  $R$ . By the definition of the differential  $d\phi(0)$ , we may choose  $\delta$  such that  $\|x\| < \delta$  and  $x \in R$  implies

$$\|\phi(x) - d\phi(0)x\| < A\rho\lambda\|x\| < A\rho r.$$

This implies that  $\phi(x) \in (1 + \rho)d\phi(0)R$ . A similar argument shows that, with,  $\delta$  chosen as above,  $x \in R$  implies that  $(1 - \rho)d\phi(0)x \in \phi(R)$ . Hence, (10.5.5) holds if  $R$  has diameter less than  $\delta$ . This completes the proof.  $\square$

**Theorem 10.5.10.** *Let  $U$  be an open subset of  $\mathbb{R}^p$  and  $\phi : U \rightarrow \mathbb{R}^p$  a smooth one-to-one transformation with  $d\phi$  non-singular at each point of  $U$ . Let  $R$  be an aligned rectangle in  $U$  and  $f$  a continuous function on  $\phi(R)$ . Then*

$$\int_{\phi(R)} f(u) dV(u) = \int_R f(\phi(x)) |\det d\phi(x)| dV(x).$$

*Proof.* For each subrectangle  $S$  of  $R$  we set

$$\begin{aligned} \Delta(S) &= \int_{\phi(S)} f(u) dV(u) - \int_S f(\phi(x)) |\det d\phi(x)| dV(x), \\ Q(S) &= \frac{\Delta(S)}{V(S)}. \end{aligned}$$

To prove the theorem, we need to show that  $\Delta(R) = 0$ . This is equivalent to showing that  $Q(R) = 0$ .

Let  $h$  be the diameter of  $R$ . We will choose inductively a downwardly nested sequence  $\{R_i\}_{i=0}^{\infty}$  of subrectangles of  $R$  in such a way that  $R_i$  has diameter  $h/2^i$  and  $|Q(R_i)| \geq |Q(R)|$ . We begin by setting  $R_0 = R$ .

Suppose  $R_0, \dots, R_m$  have been chosen in such a way that the conditions of the previous paragraph are met. If  $R_m = [a_1, b_1] \times \dots \times [a_p, b_p]$ , we partition  $R_m$  by partitioning each interval  $[a_k, b_k]$  into two subintervals of equal length. There are  $2^p$  subrectangles of  $R_m$  for this partition and each of them has diameter  $h/2^{m+1}$  since  $R_m$  has diameter  $h/2^m$ . If  $\{S_1, \dots, S_n\}$  is a list of these subrectangles of  $R_m$ , then  $R_m = \cup_j S_j$  and

$$\Delta(R_m) = \sum_{j=1}^n \Delta(S_j) = \sum_{j=1}^n Q(S_j)V(S_j).$$

For at least one of the rectangles  $S_j$ , we must have  $|Q(S_j)| \geq |Q(R_m)|$ , for if  $|Q(S_j)| < |Q(R_m)|$  for all  $j$ , then

$$\Delta(R_m) = \sum_{j=1}^n Q(S_j)V(S_j) < \sum_{j=1}^n Q(R_m)V(S_j) = Q(R_m)V(R_m) = \Delta(R_m),$$

which is impossible. Thus, for some  $j$ , we have  $|Q(S_j)| \geq |Q(R_m)|$ . We choose  $R_{m+1}$  to be an  $S_j$  which satisfies this inequality. This proves by induction that a sequence  $\{R_i\}$  with the required properties can be chosen.

Since the sequence  $\{R_i\}$  is a downwardly nested sequence of compact sets, it has a non-empty intersection. Let  $a$  be a point in this intersection.

Since  $\phi$  is smooth, we may choose a neighborhood  $V$  of  $a$  in which  $|\det d\phi(x)|$  is bounded above by a positive constant  $K$ . If  $\lambda$  is the aspect ratio of  $R$ , then each of the rectangles  $R_j$  has the same aspect ratio. By the previous lemma, there is a  $\delta > 0$  such that each rectangle  $R$  in  $V$  with aspect ratio at least  $\lambda$  and with diameter less than  $\delta$  satisfies

$$|V(\phi(R)) - V(d\phi(b)R)| < \epsilon V(R),$$

where  $b$  is the center of the rectangle  $R$ . These conditions will be met for all  $R_j$  with  $R_j \subset B_\delta(a)$ . We will denote the center of  $R_j$  by  $a_j$ . If we also choose  $\delta$  small enough that

$$|f(\phi(x)) - f(\phi(y))| < \epsilon, \quad \text{and} \quad |f(\phi(x))|\det d\phi(x) - f(\phi(y))\det d\phi(y)| < \epsilon$$

for all  $x, y \in B_\delta(a)$ , then

$$\begin{aligned} |\Delta(R_j)| &= \left| \int_{\phi(R_j)} f(u) dV(u) - \int_{R_j} f(\phi(x)) |\det d\phi(x)| dV(x) \right| \\ &\leq \left| \int_{\phi(R_j)} f(\phi(a_j)) dV(u) - \int_{R_j} f(\phi(a_j)) |\det d\phi(a_j)| dV(x) \right| \\ &\quad + \int_{\phi(R_j)} |f(u) - f(\phi(a_j))| dV(u) \\ &\quad + \int_{R_j} |f(\phi(x)) |\det d\phi(x)| - f(\phi(a_j)) |\det d\phi(a_j)|| dV(x) \\ &\leq |f(\phi(a_j))| |V(\phi(R_j)) - V(d\phi(a_j)R_j)| + \epsilon V(\phi(R_j)) + \epsilon V(R_j). \end{aligned}$$

Since  $|V(\phi(R_j)) - V(d\phi(a_j)R_j)| < \epsilon V(R_j)$  and  $V(d\phi(a_j)R_j) = |\det d\phi(a_j)|V(R_j)$ , it follows that

$$|\Delta(R_j)| \leq \epsilon V(R_j) (|f(\phi(a_j))| + |\det d\phi(a_j)| + \epsilon + 1).$$

Since  $\epsilon$  was arbitrary and  $\phi(a_j) \rightarrow \phi(a)$  and  $d\phi(a_j) \rightarrow d\phi(a)$  as  $j \rightarrow \infty$ , this implies that  $Q(R_j) = \Delta(R_j)/V(R_j)$  can be made smaller than any positive number by choosing  $j$  large enough. Since  $Q(R) \leq Q(R_j)$  for all  $j$ , this implies that  $Q(R) = 0$ , as required.  $\square$

This has the following corollary, the proof of which is left to the exercises.

**Corollary 10.5.11.** *Let  $U$  be an open subset of  $\mathbb{R}^d$  and  $\phi : U \rightarrow \mathbb{R}^d$  a smooth one-to-one transformation with non-singular differential on  $U$ . If  $R$  is an aligned rectangle in  $U$ , then*

$$V(\phi(R)) = \int_R |\det d\phi(x)| dV(x).$$

Furthermore, if  $M = \sup_R |\det d\phi|$  and  $m = \inf_R |\det d\phi|$ , then

$$mV(R) \leq V(\phi(R)) \leq MV(R).$$

### Integral over the Smooth Image of a Jordan Region

We can now prove the general change of variables formula. The proof uses the following lemma, which follows easily from the previous corollary. The proof is left to the exercises.

**Lemma 10.5.12.** *If  $\phi : U \rightarrow \mathbb{R}^d$  is a smooth one-to-one function with  $d\phi$  non-singular on  $U$  and if  $K \subset U$  is a compact set of volume 0, then  $\phi(K)$  is also a set of volume 0.*

**Theorem 10.5.13.** *Let  $A$  be a compact Jordan region contained in an open set  $U \subset \mathbb{R}^d$ . Let  $\phi : U \rightarrow \mathbb{R}^d$  be a smooth one-to-one function with a differential which is non-singular on  $A$ , and let  $f$  be a function which is bounded on  $\phi(A)$  and continuous except on a subset  $E$  of  $\phi(A)$  of volume 0. Then,  $\phi(A)$  is a Jordan region,  $f$  is integrable on  $\phi(A)$ ,  $f \circ \phi$  is integrable on  $A$  and*

$$\int_{\phi(A)} f(u) dV(u) = \int_A f(\phi(x)) |\det d\phi(x)| dV(x).$$

*Proof.* Let  $V = \phi(U)$ . By the Inverse Function Theorem,  $V$  is an open set and  $\phi^{-1} : V \rightarrow U$  is a smooth function with non-singular differential.

The boundary of  $A$  is a set of volume 0 since  $A$  is a Jordan region. Since  $\phi$  and  $\phi^{-1}$  are both continuous,  $\partial\phi(A) = \phi(\partial A)$ . It follows from the previous lemma that  $\partial\phi(A)$  is also a set of volume 0 and, hence, that  $\phi(A)$  is a Jordan region. Hence, we may extend  $f$  to be 0 on the complement of  $\phi(A)$  in  $V$  and it will still be a function which is continuous except on a set of volume 0. It follows from Theorem 10.3.5 that  $f$  is integrable on  $\phi(A)$ .

Let  $K$  be the closure of  $\partial\phi(A) \cup E$ . Then  $f$ , extended to be 0 on the complement of  $\phi(A)$ , is continuous on the complement of  $K$ . The set  $K$  has volume 0. Hence, by the previous lemma,  $\phi^{-1}(K)$  is a set of volume 0. Since  $f \circ \phi$  is continuous on  $U$  except at points of  $\phi^{-1}(K)$ , it follows that  $f \circ \phi$  is integrable on  $A$ .

Let  $\epsilon$  be any positive number. Let  $R$  be a rectangle containing  $A$  and  $P$  a partition of  $R$ . We choose  $P$  so that  $R_1, R_2, \dots, R_n$  is a list of those rectangles for this partition which are contained in  $U$ . If the partition is fine enough, then it will be true that  $A \subset \cup_j R_j$ . Also, the partition may be chosen fine enough that, if  $S$  is the set of  $j$  for which  $R_j \cap K \neq \emptyset$ , then

$$\sum_{j \in S} V(R_j) < \epsilon.$$

If  $K \cap R_j = \emptyset$ , then either  $A \cap R_j = \emptyset$  or  $R_j$  is a rectangle contained in the interior of  $A$  and  $f$  is continuous on  $\phi(R_j)$ . If the latter is true, then

$$\int_{\phi(R_j)} f(u) dV(u) = \int_{R_j} f(\phi(x)) |\det d\phi(x)| dV(x).$$

Since  $f$  is 0 on the complement of  $\phi(A)$ , we have

$$\begin{aligned}
 & \left| \int_{\phi(A)} f(u) dV(u) - \int_A f(\phi(x)) |\det \phi(x)| dV(x) \right| \\
 &= \left| \sum_j \left( \int_{\phi(R_j)} f(u) dV(u) - \int_{R_j} f(\phi(x)) |\det \phi(x)| dV(x) \right) \right| \\
 &= \left| \sum_{j \in S} \left( \int_{\phi(R_j)} f(u) dV(u) - \int_{R_j} f(\phi(x)) |\det \phi(x)| dV(x) \right) \right| \\
 &\leq \sum_{j \in S} \left( \int_{\phi(R_j)} M dV(u) + \int_{R_j} MK dV(x) \right) \\
 &= \sum_{j \in S} (MV(\phi(R_j)) + MKV(R_j)) \leq 2MK\epsilon.
 \end{aligned}$$

where  $M = \sup_A |f(\phi(x))|$  and  $K = \sup_A |\det d\phi(x)|$ . Since,  $\epsilon$  is arbitrary, this implies the equality of the theorem.  $\square$

With some additional hypotheses, the above theorem can be strengthened so as to apply to integrals over the full open sets  $U$  and  $\phi(U)$  rather than just to integrals over compact subsets. The next theorem is such a result.

**Theorem 10.5.14.** *Let  $U$  be an open Jordan region in  $\mathbb{R}^d$  and let  $\phi : A \rightarrow \mathbb{R}^d$  be a one to one smooth function on  $U$  with image  $\phi(U)$  which is also a Jordan region. Suppose  $d\phi$  is non-singular on  $U$  and  $f$  is bounded on  $\phi(U)$  and continuous except on a subset of volume 0. Then  $f$  is integrable on  $\phi(U)$ . If, in addition,  $f \circ \phi |\det \phi|$  is bounded on  $U$ , then it too is integrable on  $U$  and*

$$\int_{\phi(U)} f(u) dV(u) = \int_U f(\phi(x)) |\det d\phi(x)| dV(x).$$

*Proof.* Since  $d\phi$  is non-singular on  $U$ , Theorem 9.6.5 implies that  $\phi : U \rightarrow \mathbb{R}^d$  is a one-to-one open map onto an open set  $V$ .

Since  $f$  is bounded on  $\phi(U)$  and is continuous except on a set of volume 0, it is integrable on  $\phi(U)$ . The function  $g(x) = f(\phi(x)) |\det d\phi(x)|$  is continuous and bounded and, hence, is an integrable function on  $U$ .

Let  $K_n$  be a sequence of compact Jordan subsets of  $U$  such that  $\cup_n K_n^\circ = U$ . Such a sequence exists by Exercise 10.3.11. Then, by Exercise 10.3.10,

$$\int_U g(x) dV(x) = \lim_n \int_{K_n} g(x) dV(x). \tag{10.5.6}$$

Also, since  $\{\phi(K_n)\}$  is a sequence of compact subsets of  $V = \phi(U)$  with the union of the interiors of the sets in the sequence equal to  $V$ , we conclude

$$\int_{\phi(U)} f(u) dV(u) = \lim_n \int_{\phi(K_n)} f(u) dV(u). \tag{10.5.7}$$

The previous theorem implies that

$$\int_{\phi(K_n)} f(u) dV(u) = \int_{K_n} g(x) dV(x),$$

for each  $n$ . This, together with (10.5.6) and (10.5.7), completes the proof.  $\square$

The change of variables theorem has the following corollary, the proof of which is left to the exercises.

**Corollary 10.5.15.** *Let  $U$  be an open Jordan region in  $\mathbb{R}^d$  and  $\phi : U \rightarrow \mathbb{R}^d$  a function satisfying the conditions of the previous theorem. Then*

$$V(\phi(U)) = \int_U |\det d\phi(x)| dV(x).$$

Note that, in the change of variables formulas in the above theorem and its corollary, the sets  $U$  and  $\phi(U)$  may be replaced by their closures, even though the transformation  $\phi$  may not be defined on the closure of  $U$ . This is due to the fact that the boundaries of  $U$  and  $\phi(U)$  have volume 0.

**Example 10.5.16.** Use the preceding corollary to find the area enclosed by an ellipse with major and minor axes of lengths  $2a$  and  $2b$  without assuming knowledge of the area of a circle.

**Solution:** Such an ellipse has equation  $x^2/a^2 + y^2/b^2 = 1$ . The region it encloses is the image of the square  $A = \{(r, \theta) : 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi\}$ , under the transformation  $\phi(r, \theta) = (ar \cos \theta, br \sin \theta)$ . The differential of this map is

$$d\phi(r, \theta) = \begin{pmatrix} a \cos \theta & -ar \sin \theta \\ b \sin \theta & br \cos \theta \end{pmatrix}$$

The determinant of this matrix is  $abr$ , which is non-zero except at  $r = 0$ . Thus, the function  $\phi$  is one to one and smooth with non-singular differential on the interior of the square  $A$ . The interior of  $A$  is taken by  $\phi$  to the interior of the ellipse with the line joining  $(0, 0)$  to  $(1, 0)$  removed. This set differs from the ellipse itself by a set of volume 0. Thus, the area we seek is, by the previous corollary and Fubini's Theorem,

$$\int_0^{2\pi} \int_0^1 abr \, dr \, d\theta = \pi ab.$$

**Example 10.5.17.** Find  $\int_0^1 \int_0^{\sqrt{1-x^2}} \cos(x^2 + y^2) \, dy \, dx$ .

**Solution:** By Fubini's Theorem, this integral is

$$\int_D \cos(x^2 + y^2) \, dV(x, y),$$

where  $D = B_1(0, 0)$ . If we change to polar coordinates using the transformation

$$\phi(r, \theta) = (r \cos \theta, r \sin \theta),$$

then  $\det d\phi(r, \theta) = r$  and  $D = \phi(R)$ , where  $R$  is the rectangle  $[0, 1] \times [0, 2\pi]$ . On  $R$ ,  $\phi$  is smooth with non-singular differential except when  $r = 0$ , and so Theorem 10.5.14 applies with  $U = R^\circ$ . Hence,

$$\int_{\phi(R)} \cos(x^2 + y^2) dV(x, y) = \int_R \cos(r^2) r dr d\theta.$$

Applying Fubini's Theorem again yields

$$\int_0^1 \int_0^{\sqrt{1-x^2}} \cos(x^2 + y^2) dy dx = \int_0^{2\pi} \int_0^1 \cos(r^2) r dr d\theta = \pi \sin 1.$$

### Exercise Set 10.5

1. Compute the inverse of each elementary matrix  $E_{ij}$ ,  $S_{ij}$ , and  $T_i(a)$ . Show that each inverse is itself an elementary matrix or a product of elementary matrices.
2. Show that if  $E$  is a Jordan region and  $L$  is a linear transformation whose matrix is singular, then  $L(E)$  has volume 0.
3. Let  $u$  and  $v$  be two vectors in the plane and define  $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  by  $\phi(s, t) = su + tv$ . Let  $A$  be the parallelogram which is the image of  $[0, 1] \times [0, 1]$  under  $\phi$ . If  $f$  is a continuous function on  $A$ , express  $\int_A f(x, y) dV(x, y)$  as an integral over  $[0, 1] \times [0, 1]$ .
4. Use the result of the previous exercise to find a formula for the area of the parallelogram determined by two vectors  $u$  and  $v$ .
5. An orthogonal transformation is a linear transformation  $A$  that preserves inner products – that is,  $Au \cdot Av = u \cdot v$  for each pair of vectors  $u, v$ . Note that a rotation is an orthogonal transformation. Prove that a  $d \times d$  orthogonal transformation preserves volume in  $\mathbb{R}^d$ .

6. Compute  $\int_0^a \int_0^{\sqrt{a^2-x^2}} e^{x^2+y^2} dy dx$ .

7. Let  $A = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0, x^2 + y^2 \leq 4, x^2 - y^2 \geq 1\}$ . Compute

$$\int_A \frac{xy}{x^2 + y^2} dV(x, y)$$

by making a change of variables  $u = x^2 + y^2, v = x^2 - y^2$  for  $x \geq 0, y \geq 0$ .

8. Compute the volume of a sphere  $S$  of radius  $r$  by computing the integral

$$\int_S 1 dV(x).$$

Compute this integral by first converting to spherical coordinates.

9. Compute the volume of a right circular cone with height  $h$  and radius  $a$ .  
Hint: such a cone can be described in cylindrical coordinates as the set of points

$$\{(r, \theta, z) : 0 \leq r \leq \frac{a}{h}z, 0 \leq \theta \leq 2\pi\}.$$

Here  $x = r \cos \theta, y = r \sin \theta, z = z$  describes the transformation from cylindrical to rectangular coordinates.

10. Show by example that the conclusion of Theorem 10.5.13 does not hold if the function  $\phi$  is not one-to-one on  $A$ .
11. Prove Corollary 10.5.11
12. Prove Lemma 10.5.12.