Chapter 8

Integration of Functions of Several Variables

In this chapter, we focus on the integration of bounded functions on bounded subsets of $\mathbb{R}^n$.

8.1 Integrable Functions

We start with a simpler case $n = 2$.

Definition 8.1. Let $A \subseteq \mathbb{R}^2$ be a bounded set. Define

$$\begin{aligned} a_1 &= \inf\{x \in \mathbb{R} \,|\, (x,y) \in A \text{ for some } y \in \mathbb{R}\}, \\ b_1 &= \sup\{x \in \mathbb{R} \,|\, (x,y) \in A \text{ for some } y \in \mathbb{R}\}, \\ a_2 &= \inf\{y \in \mathbb{R} \,|\, (x,y) \in A \text{ for some } x \in \mathbb{R}\}, \\ b_2 &= \sup\{y \in \mathbb{R} \,|\, (x,y) \in A \text{ for some } x \in \mathbb{R}\}. \end{aligned}$$

A collection of rectangles $\mathcal{P}$ is called a ***partition*** of A if there exists a partition $\mathcal{P}_x$ of $[a_1, b_1]$ and a partition P_y of $[a_2, b_2]$,

$$\mathcal{P}_x = \{a_1 = x_0 < x_1 < \cdots < x_n = b_1\} \quad \text{and} \quad \mathcal{P}_y = \{a_2 = y_0 < y_1 < \cdots < y_m = b_2\},$$

such that

$$\mathcal{P} = \{\Delta_{ij} \,|\, \Delta_{ij} = [x_i, x_{i+1}] \times [y_j, y_{j+1}] \text{ for } i = 0, 1, \cdots, n-1 \text{ and } j = 0, 1, \cdots, m-1\}.$$

The ***mesh size*** of the partition $\mathcal{P}$ and also called the norm of $\mathcal{P}$, denoted by $\|\mathcal{P}\|$, is defined by

$$\|\mathcal{P}\| = \max\left\{\sqrt{(x_{i+1} - x_i)^2 + (y_{j+1} - y_j)^2} \,\middle|\, i = 0, 1, \cdots, n-1, j = 0, 1, \cdots, m-1\right\}.$$

The number $\sqrt{(x_{i+1}-x_i)^2+(y_{j+1}-y_j)^2}$ is often denoted by $\operatorname{diam}(\Delta_{ij})$, and is called the ***diameter*** of Δ_{ij}.

Similar to the integrability of f on a bounded subset of $\mathbb{R}$, we have the following

Definition 8.2. Let $A \subseteq \mathbb{R}^2$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. For any partition $\mathcal{P} = \{\Delta_{ij} \,|\, \Delta_{ij} = [x_i, x_{i+1}] \times [y_j, y_{j+1}], i = 0, \cdots, n-1, j = 0, \cdots, m-1\}$, the ***upper sum*** and the ***lower sum*** of f with respect to the partition $\mathcal{P}$, denoted by $U(f, \mathcal{P})$ and $L(f, \mathcal{P})$ respectively, are numbers defined by

$$U(f,\mathcal{P}) = \sum_{\substack{0\leqslant i\leqslant n-1\\ 0\leqslant j\leqslant m-1}} \sup_{(x,y)\in\Delta_{ij}} \overline{f}^A(x,y)\mathbb{A}(\Delta_{ij}),$$
$$L(f,\mathcal{P}) = \sum_{\substack{0\leqslant i\leqslant n-1\\ 0\leqslant j\leqslant m-1}} \inf_{(x,y)\in\Delta_{ij}} \overline{f}^A(x,y)\mathbb{A}(\Delta_{ij}),$$

where $\mathbb{A}(\Delta_{ij}) = (x_{i+1}-x_i)(y_{j+1}-y_j)$ is the area of the rectangle Δ_{ij}, and $\overline{f}^A$ is an extension of f, called the extension of f by zero outside A, given by

$$\overline{f}^A(x) = \begin{cases} f(x) & x \in A\,, \\ 0 & x \notin A\,. \end{cases}$$

The two numbers

$$\overline{\int_A} f(x,y)d\mathbb{A} \equiv \inf\{U(f,\mathcal{P}) \,|\, \mathcal{P} \text{ is a partition of } A\}$$

and

$$\underline{\int_A} f(x,y)d\mathbb{A} \equiv \sup\{L(f,\mathcal{P}) \,|\, \mathcal{P} \text{ is a partition of } A\}$$

are called the ***upper integral*** and ***lower integral*** of f over A, respectively. The function f is said to be ***Riemann*** (***Darboux***) ***integrable*** (over A) if $\overline{\int_A} f(x,y)d\mathbb{A} = \underline{\int_A} f(x,y)d\mathbb{A}$, and in this case, we express the upper and lower integral as $\int_A f(x,y)d\mathbb{A}$, called the ***integral*** of f over A.

In general, we can consider the integrability of a bounded function f defined on a bounded set $A \subseteq \mathbb{R}^n$ as follows

Definition 8.3. Let $A \subseteq \mathbb{R}^n$ be a bounded set. Define the numbers $a_1, a_2, \cdots, a_n$ and $b_1, b_2, \cdots, b_n$ by

$$\begin{aligned} a_k &= \inf\left\{x_k \in \mathbb{R} \,\middle|\, x = (x_1, \cdots, x_n) \in A \text{ for some } x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n \in \mathbb{R}\right\}, \\ b_k &= \sup\left\{x_k \in \mathbb{R} \,\middle|\, x = (x_1, \cdots, x_n) \in A \text{ for some } x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n \in \mathbb{R}\right\}. \end{aligned}$$

A collection of rectangles $\mathcal{P}$ is called a ***partition*** of A if there exists partitions $\mathcal{P}^{(k)}$ of $[a_k, b_k]$, $k = 1, \cdots, n$, $\mathcal{P}^{(k)} = \left\{a_k = x_0^{(k)} < x_1^{(k)} < \cdots < x_{N_k}^{(k)} = b_k\right\}$, such that

$$\begin{aligned} \mathcal{P} = \Big\{\Delta_{i_1 i_2 \cdots i_n} \,\Big|\, \Delta_{i_1 i_2 \cdots i_n} = [x_{i_1}^{(1)}, x_{i_1+1}^{(1)}] \times [x_{i_2}^{(2)}, x_{i_2+1}^{(2)}] \times \cdots \times [x_{i_n}^{(n)}, x_{i_{n+1}}^{(n+1)}], \\ i_k = 0, 1, \cdots, N_k - 1, k = 1, \cdots, n\Big\}. \end{aligned}$$

The ***mesh size*** of the partition $\mathcal{P}$, denoted by $\|\mathcal{P}\|$ and also called the norm of $\mathcal{P}$, is defined by

$$\|\mathcal{P}\| = \max\left\{\sqrt{\sum_{k=1}^{n} (x_{i_k+1}^{(k)} - x_{i_k}^{(k)})^2} \,\middle|\, i_k = 0, 1, \cdots, N_k - 1, k = 1, \cdots, n\right\}.$$

The number $\sqrt{\sum\limits_{k=1}^{n} (x_{i_k+1}^{(k)} - x_{i_k}^{(k)})^2}$ is often denoted by $\operatorname{diam}(\Delta_{i_1 i_2 \cdots i_n})$, and is called the ***diameter*** of the rectangle $\Delta_{i_1 i_2 \cdots i_n}$.

Definition 8.4. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. For any partition

$$\begin{aligned} \mathcal{P} = \Big\{\Delta_{i_1 i_2 \cdots i_n} \,\Big|\, \Delta_{i_1 i_2 \cdots i_n} = [x_{i_1}^{(1)}, x_{i_1+1}^{(1)}] \times [x_{i_2}^{(2)}, x_{i_2+1}^{(2)}] \times \cdots \times [x_{i_n}^{(n)}, x_{i_{n+1}}^{(n+1)}], \\ i_k = 0, 1, \cdots, N_k - 1, k = 1, \cdots, n\Big\}, \end{aligned}$$

the ***upper sum*** and the ***lower sum*** of f with respect to the partition $\mathcal{P}$, denoted by $U(f, \mathcal{P})$ and $L(f, \mathcal{P})$ respectively, are numbers defined by

$$\begin{aligned} U(f, \mathcal{P}) &= \sum_{\Delta \in \mathcal{P}} \sup_{(x,y) \in \Delta} \overline{f}^A(x, y) \nu(\Delta), \\ L(f, \mathcal{P}) &= \sum_{\Delta \in \mathcal{P}} \inf_{(x,y) \in \Delta} \overline{f}^A(x, y) \nu(\Delta), \end{aligned}$$

where $\nu(\Delta)$ is the ***volume*** of the rectangle Δ given by

$$\nu(\Delta) = (x_{i_1+1}^{(1)} - x_{i_1}^{(1)})(x_{i_2+1}^{(2)} - x_{i_2}^{(2)}) \cdots (x_{i_n+1}^{(n)} - x_{i_n}^{(n)})$$

if $\Delta = [x^{(1)}_{i_1} - x^{(1)}_{i_1+1}] \times [x^{(2)}_{i_2} - x^{(2)}_{i_2+1}] \times \cdots \times [x^{(n)}_{i_n} - x^{(n)}_{i_n+1}]$, and $\overline{f}^A$ is the extension of f by zero outside A given by

$$\overline{f}^A(x) = \begin{cases} f(x) & x \in A\,, \\ 0 & x \notin A\,. \end{cases} \tag{8.1.1}$$

The two numbers

$$\overline{\int_A} f(x)\,dx \equiv \inf\big\{U(f,\mathcal{P})\,\big|\,\mathcal{P} \text{ is a partition of } A\big\}\,,$$

and

$$\underline{\int_A} f(x)\,dx \equiv \sup\big\{L(f,\mathcal{P})\,\big|\,\mathcal{P} \text{ is a partition of } A\big\}$$

are called the ***upper integral*** and ***lower integral*** of f over A, respective. The function f is said to be ***Riemann*** (***Darboux***) ***integrable*** (over A) if $\overline{\int_A} f(x)\,dx = \underline{\int_A} f(x)\,dx$, and in this case, we express the upper and lower integral as $\int_A f(x)\,dx$, called the ***integral*** of f over A.

Definition 8.5. A partition $\mathcal{P}'$ of a bounded set $A \subseteq \mathbb{R}^n$ is said to be a ***refinement*** of another partition $\mathcal{P}$ of A if for any $\Delta' \in \mathcal{P}'$, there is $\Delta \in \mathcal{P}$ such that $\Delta' \subseteq \Delta$. A partition $\mathcal{P}$ of a bounded set $A \subseteq \mathbb{R}^n$ is said to be the ***common refinement*** of another partitions $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ of A if

1. $\mathcal{P}$ is a refinement of $\mathcal{P}_j$ for all $1 \leqslant j \leqslant k$.

2. If $\mathcal{P}'$ is a refinement of $\mathcal{P}_j$ for all $1 \leqslant j \leqslant k$, then $\mathcal{P}'$ is also a refinement of $\mathcal{P}$.

In other words, $\mathcal{P}$ is a common refinement of $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ if it is the coarsest refinement.

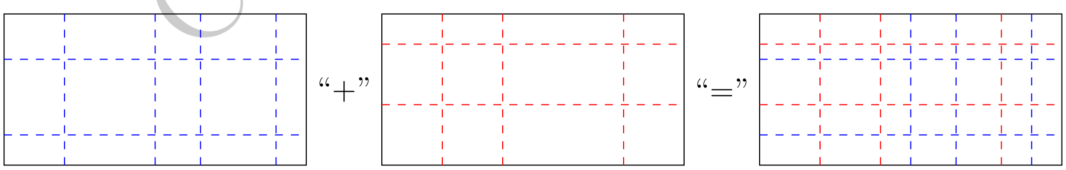

Figure 8.1: The common refinement of two partitions

Qualitatively speaking, $\mathcal{P}$ is a common refinement of $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ if for each $j = 1, \cdots n$, the j-th component c_j of the vertex $(c_1, \cdots, c_n)$ of each rectangle $\Delta \in \mathcal{P}$ belongs to $\mathcal{P}^{(j)}_i$ for some $i = 1, \cdots, k$.

Similar to Proposition 4.78 and Corollary 4.79, we have

Proposition 8.6. *Let $A \subseteq \mathbb{R}^n$ be a bounded subset, and $f : A \to \mathbb{R}$ be a bounded function. If $\mathcal{P}$ and $\mathcal{P}'$ are partitions of A and $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then*

$$L(f,\mathcal{P}) \leqslant L(f,\mathcal{P}') \leqslant U(f,\mathcal{P}') \leqslant U(f,\mathcal{P}) .$$

Corollary 8.7. *Let $A \subseteq \mathbb{R}^n$ be a bounded subset, and $f : A \to \mathbb{R}$ be a bounded function. If $\mathcal{P}_1$ and $\mathcal{P}_2$ are partitions of A, then*

$$L(f,\mathcal{P}_1) \leqslant \underline{\int_A} f(x)dx \leqslant \overline{\int_A} f(x)dx \leqslant U(f,\mathcal{P}_2) .$$

8.2 Conditions for Integrability

In the following two sections, we discuss some equivalent conditions for Riemann integrability of bounded functions (over bounded sets). We recall that in Section 4.7 we have talked about two equivalent conditions for Riemann integrability: the Riemann condition (Proposition 4.80) and the Darboux theorem (Theorem 4.94). This section contributes to the n-dimensional version of Riemann's condition and Darboux theorem.

The proof of the following proposition is identical to the proof of Proposition 4.80.

Proposition 8.8 (Riemann's condition). *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable over A if and only if*

$$\forall\, \varepsilon > 0, \exists \text{ a partition } \mathcal{P} \text{ of } A \ni U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon .$$

Definition 8.9. Let $\mathcal{P} = \{\Delta_1, \Delta_2, \cdots, \Delta_N\}$ be a partition of a bounded set $A \subseteq \mathbb{R}^n$. A collection of N points $\{\xi_1, \cdots, \xi_N\}$ is called a ***sample set*** for the partition $\mathcal{P}$ if $\xi_k \in \Delta_k$ for all $k = 1, \cdots, N$. Points in a sample set are called sample points for the partition $\mathcal{P}$.

Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. A ***Riemann sum*** of f for the the partition $\mathcal{P} = \{\Delta_1, \Delta_2, \cdots, \Delta_N\}$ of A is a sum which takes the form

$$\sum_{k=1}^{N} \overline{f}^A(\xi_i)\nu_n(\Delta_k) ,$$

where the set $\Xi = \{\xi_1, \xi_2, \cdots, \xi_N\}$ is a sample set for the partition $\mathcal{P}$.

Similar to Theorem 4.94, the following theorem establishes the equivalence between the Riemann condition and the Darboux integrals. The idea of the proof of the following theorem are essentially identical to the proof of Theorem 4.94; however, the detail proof requires a slight modification due to the fact that the dimension is bigger than one.

Theorem 8.10 (Darboux). *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function with extension $\overline{f}^A$ given by (8.1.1). Then f is Riemann integrable over A if and only if there exists $\mathrm{I} \in \mathbb{R}$ such that for every given $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of A satisfying $\|\mathcal{P}\| < \delta$, then any Riemann sums for the partition $\mathcal{P}$ belongs to the interval $(\mathrm{I} - \varepsilon, \mathrm{I} + \varepsilon)$. In other words, f is Riemann integrable over A if and only if there exists $\mathrm{I} \in \mathbb{R}$ such that for every given $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\Big|\sum_{k=1}^{N} \overline{f}^A(\xi_k)\nu(\Delta_k) - \mathrm{I}\Big| < \varepsilon \tag{8.2.1}$$

whenever $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ is a partition of A satisfying $\|\mathcal{P}\| < \delta$ and $\{\xi_1, \xi_2, \cdots, \xi_N\}$ is a sample set for $\mathcal{P}$.

Proof. The boundedness of A guarantees that $A \subseteq \big[-\frac{r}{2}, \frac{r}{2}\big]^n$ for some $r > 0$. Let $R = \big[-\frac{r}{2}, \frac{r}{2}\big]^n$.

"$\Leftarrow$" Suppose the right-hand side statement is true. Let $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that if $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ is a partition of A satisfying $\|\mathcal{P}\| < \delta$, then for all sets of sample points $\{\xi_1, \cdots, \xi_N\}$ for $\mathcal{P}$, we must have

$$\Big|\sum_{k=1}^{N} \overline{f}^A(\xi_k)\nu(\Delta_k) - \mathrm{I}\Big| < \frac{\varepsilon}{4}.$$

Let $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ be a partition of A with $\|\mathcal{P}\| < \delta$. Choose two sample sets $\{\xi_1, \cdots, \xi_N\}$ and $\{\eta_1, \cdots, \eta_N\}$ for $\mathcal{P}$ such that

(a) $\sup\limits_{x \in \Delta_k} \overline{f}^A(x) - \dfrac{\varepsilon}{4\nu(R)} < \overline{f}^A(\xi_k) \leqslant \sup\limits_{x \in \Delta_k} \overline{f}^A(x)$;

(b) $\inf\limits_{x \in \Delta_k} \overline{f}^A(x) + \dfrac{\varepsilon}{4\nu(R)} > \overline{f}^A(\eta_k) \geqslant \inf\limits_{x \in \Delta_k} \overline{f}^A(x)$.

Then

$$\begin{aligned} U(f, \mathcal{P}) &= \sum_{k=1}^{N} \sup_{x \in \Delta_k} \overline{f}^A(x)\nu(\Delta_k) < \sum_{k=1}^{N} \Big[\overline{f}^A(\xi_k) + \frac{\varepsilon}{4\nu(R)}\Big]\nu(\Delta_k) \\ &= \sum_{k=1}^{N} \overline{f}^A(\xi_k)\nu(\Delta_k) + \frac{\varepsilon}{4\nu(R)} \sum_{k=1}^{N} \nu(\Delta_k) < \mathrm{I} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \mathrm{I} + \frac{\varepsilon}{2} \end{aligned}$$

and

$$L(f,\mathcal{P})=\sum_{k=1}^{N}\inf_{x\in\Delta_k}\overline{f}^A(x)\nu(\Delta_k)>\sum_{k=1}^{N}\Big[\overline{f}^A(\eta_k)-\frac{\varepsilon}{4\nu(R)}\Big]\nu(\Delta_k)$$
$$=\sum_{k=1}^{N}\overline{f}^A(\eta_k)\nu(\Delta_k)-\frac{\varepsilon}{4\nu(R)}\sum_{k=1}^{N}\nu(\Delta_k)>\mathrm{I}-\frac{\varepsilon}{4}-\frac{\varepsilon}{4}=\mathrm{I}-\frac{\varepsilon}{2}\,.$$

As a consequence, $\mathrm{I}-\frac{\varepsilon}{2}<L(f,\mathcal{P})\leqslant U(f,\mathcal{P})<\mathrm{I}+\frac{\varepsilon}{2}$; thus $U(f,\mathcal{P})-L(f,\mathcal{P})<\varepsilon$.

"⇒" Let $\mathrm{I}=\int_A f(x)dx$, and $\varepsilon>0$ be given. Since f is Riemann integrable over A, there exists a partition $\mathcal{P}_1$ of A such that $U(f,\mathcal{P}_1)-L(f,\mathcal{P}_1)<\frac{\varepsilon}{2}$. Suppose that $\mathcal{P}_1^{(i)}=\big\{y_0^{(i)},y_1^{(i)},\cdots,y_{m_i}^{(i)}\big\}$ for $1\leqslant i\leqslant n$. With M denoting the number $m_1+m_2+\cdots+m_n$, we define

$$\delta=\frac{\varepsilon}{4r^{n-1}(M+n)\big(\sup\overline{f}^A(R)-\inf\overline{f}^A(R)+1\big)}\,.$$

Then $\delta>0$. Our goal is to show that if $\mathcal{P}$ is a partition of A with $\|\mathcal{P}\|<\delta$ and $\{\xi_1,\cdots,\xi_N\}$ is a set of sample points for $\mathcal{P}$, then (8.2.1) holds.

Assume that $\mathcal{P}=\{\Delta_1,\Delta_2,\cdots,\Delta_N\}$ is a given partition of A with $\|\mathcal{P}\|<\delta$. Let $\mathcal{P}'$ be the common refinement of $\mathcal{P}$ and $\mathcal{P}_1$. Write $\mathcal{P}'=\{\Delta'_1,\Delta'_2,\cdots,\Delta'_{N'}\}$ and $\Delta_k=\Delta_k^{(1)}\times\Delta_k^{(2)}\times\cdots\times\Delta_k^{(n)}$ as well as $\Delta'_k=\Delta_k'^{(1)}\times\Delta_k'^{(2)}\times\cdots\times\Delta_k'^{(n)}$. By the definition of the upper sum,

$$U(f,\mathcal{P})=\sum_{k=1}^{N}\sup_{x\in\Delta_k}\overline{f}^A(x)\nu(\Delta_k)$$
$$=\sum_{\substack{1\leqslant k\leqslant N\text{ with}\\ y_j^{(i)}\notin\Delta_k^{(i)}\text{ for all }i,j}}\sup_{x\in\Delta_k}\overline{f}^A(x)\nu(\Delta_k)+\sum_{\substack{1\leqslant k\leqslant N\text{ with}\\ y_j^{(i)}\in\Delta_k^{(i)}\text{ for some }i,j}}\sup_{x\in\Delta_k}\overline{f}^A(x)\nu(\Delta_k)$$

and similarly,

$$U(f,\mathcal{P}')=\sum_{\substack{1\leqslant k\leqslant N'\text{ with}\\ y_j^{(i)}\notin\Delta_k'^{(i)}\text{ for all }i,j}}\sup_{x\in\Delta'_k}\overline{f}^A(x)\nu(\Delta'_k)+\sum_{\substack{1\leqslant k\leqslant N'\text{ with}\\ y_j^{(i)}\in\Delta_k'^{(i)}\text{ for some }i,j}}\sup_{x\in\Delta'_k}\overline{f}^A(x)\nu(\Delta'_k)\,.$$

By the fact that $\Delta_k\in\mathcal{P}'$ if $y_j^{(i)}\notin\Delta_k'^{(i)}$ for all i,j, we must have

$$\sum_{\substack{1\leqslant k\leqslant N\text{ with}\\ y_j^{(i)}\in\Delta_k^{(i)}\text{ for some }i,j}}\nu(\Delta_k)=\sum_{\substack{1\leqslant k\leqslant N'\text{ with}\\ y_j^{(i)}\in\Delta_k'^{(i)}\text{ for some }i,j}}\nu(\Delta'_k)\,.$$

The equality above further implies that

$$U(f,\mathcal{P})-U(f,\mathcal{P}') = \sum_{\substack{1\leqslant k\leqslant N \text{ with} \\ y_j^{(i)}\in\Delta_k^{(i)} \text{for some } i,j}} \sup_{x\in\Delta_k} \overline{f}^A(x)\nu(\Delta_k) - \sum_{\substack{1\leqslant k\leqslant N' \text{ with} \\ y_j^{(i)}\in\Delta_k'^{(i)} \text{for some } i,j}} \sup_{x\in\Delta_k'} \overline{f}^A(x)\nu(\Delta_k')$$
$$\leqslant \big(\sup \overline{f}^A(R) - \inf \overline{f}^A(R)\big) \sum_{\substack{1\leqslant k\leqslant N \text{ with} \\ y_j^{(i)}\in\Delta_k^{(i)} \text{for some } i,j}} \nu(\Delta_k)\,.$$

Moreover, for each fixed i,j,

$$\bigcup_{1\leqslant k\leqslant N \text{ with } y_j^{(i)}\in\Delta_k^{(i)}} \Delta_k \subseteq \Big[-\frac{r}{2},\frac{r}{2}\Big]^{i-1}\times\big[y_j^{(i)}-\delta, y_j^{(i)}+\delta\big]\times\Big[-\frac{r}{2},\frac{r}{2}\Big]^{n-i};$$

thus

$$\sum_{1\leqslant k\leqslant N \text{ with } y_j^{(i)}\in\Delta_k^{(i)}} \nu(\Delta_k) \leqslant 2\delta r^{n-1} \qquad \forall\, 1\leqslant i\leqslant n, 1\leqslant j\leqslant m_i\,.$$

Therefore,

$$U(f,\mathcal{P})-U(f,\mathcal{P}') \leqslant \big(\sup \overline{f}^A(R) - \inf \overline{f}^A(R)\big)\sum_{i=1}^{n}\sum_{j=0}^{m_i} \sum_{1\leqslant k\leqslant N \text{ with } y_j^{(i)}\in\Delta_k^{(i)}} \nu(\Delta_k)$$
$$\leqslant \big(\sup \overline{f}^A(R) - \inf \overline{f}^A(R)\big)\sum_{i=1}^{n}\sum_{j=0}^{m_i} 2\delta r^{n-1}$$
$$\leqslant 2\delta r^{n-1}(m_1+m_2+\cdots+m_n+n)\big(\sup \overline{f}^A(R) - \inf \overline{f}^R(A)\big) < \frac{\varepsilon}{2}\,,$$

and the fact that $U(f,\mathcal{P}_1)-L(f,\mathcal{P}_1)<\dfrac{\varepsilon}{2}$ shows that

$$U(f,\mathcal{P})-\mathrm{I} \leqslant U(f,\mathcal{P})-\mathrm{I}+U(f,\mathcal{P}_1)-U(f,\mathcal{P}_1)$$
$$\leqslant U(f,\mathcal{P})-L(f,\mathcal{P}_1)+U(f,\mathcal{P}_1)-U(f,\mathcal{P}') < \varepsilon\,.$$

Therefore, for any sample set $\{\xi_1,\cdots,\xi_N\}$ for $\mathcal{P}$,

$$\sum_{k=1}^{N}\overline{f}^A(\xi_k)\nu(\Delta_k)\leqslant U(f,\mathcal{P})<\mathrm{I}+\varepsilon\,.$$

Similar argument can be used to show that

$$\sum_{k=1}^{N}\overline{f}^A(\xi_k)\nu(\Delta_k)\geqslant L(f,\mathcal{P})>\mathrm{I}-\varepsilon$$

which concludes the Theorem. □

In Section 5.1, we show that if a sequence of Riemann integrable functions $\{f_k\}_{k=1}^{\infty}$ converges to a function f uniformly on $[a,b]$, then f is also Riemann integrable over $[a,b]$ and the integral of the limit function is the same as the limit of the integrals (of the sequences). This theorem can also be established if the domain A under consideration is a bounded subset of $\mathbb{R}^n$. In fact, the same proof used to establish Theorem 5.16 can be applied to conclude the following

Theorem 8.11. *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f_k : A \to \mathbb{R}$ be a sequence of Riemann integrable functions over A such that $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on A. Then f is Riemann integrable over A, and*

$$\lim_{k\to\infty} \int_A f_k(x)\,dx = \int_A f(x)\,dx\,. \tag{8.2.2}$$

From now on, we will simply use $\bar{f}$ to denote the zero extension of f when the domain outside which the zero extension is made is clear.

8.3 The Lebesgue Theorem

In this section, we talk about another equivalent condition of Riemann integrability, named the Lebesgue theorem. The Lebesque theorem provides a more practical way to check the Riemann integrability in the development of theory. To understand the Lebesgue theorem, we need to talk about a new concept, sets of measure zero.

8.3.1 Volume and Sets of Measure Zero

Definition 8.12. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $\mathbf{1}_A$ (or χ_A) be the characteristic function of A defined by

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

A is said to ***have volume*** if $\mathbf{1}_A$ is Riemann integrable (over A), and the ***volume*** of A, denoted by $\nu(A)$, is the number $\int_A \mathbf{1}_A(x)\,dx$. A is said to have ***volume zero*** or ***content zero*** if $\nu(A) = 0$.

Remark 8.13. Not all bounded set has volume.

Proposition 8.14. *Let $A \subseteq \mathbb{R}^n$ be bounded. Then the following three statements are equivalent.*

(a) *A has volume zero;*

(b) *for every $\varepsilon > 0$, there exists finite open rectangles $S_1, \cdots, S_N$ whose sides are parallel to the coordinate axes such that*

$$A \subseteq \bigcup_{k=1}^{N} S_k \qquad \textit{and} \qquad \sum_{k=1}^{N} \nu(S_k) < \varepsilon \tag{8.3.1}$$

(c) *for every $\varepsilon > 0$, there exists finite rectangles $S_1, \cdots, S_N$ such that* (8.3.1) *holds.*

Proof. It suffices to show (a)$\Rightarrow$(b) and (c)$\Rightarrow$(a) since it is clear that (b)$\Rightarrow$(c).

"(a)$\Rightarrow$(b)" Since A has volume zero, $\overline{\int}_A \mathbf{1}_A(x)\,dx = 0$; thus for any given $\varepsilon > 0$, there exists a partition $\mathcal{P}$ of A such that

$$U(\mathbf{1}_A, \mathcal{P}) < \overline{\int}_A \mathbf{1}_A(x)\,dx + \frac{\varepsilon}{2} = \frac{\varepsilon}{2}\,.$$

Since $\sup\limits_{x\in\Delta} \mathbf{1}_A(x) = \begin{cases} 1 & \text{if } \Delta \cap A \neq \varnothing\,, \\ 0 & \text{otherwise}\,, \end{cases}$ we must have $\sum\limits_{\substack{\Delta\in\mathcal{P}\\ \Delta\cap A\neq\varnothing}} \nu(\Delta) < \dfrac{\varepsilon}{2}$. Now if $\Delta \in \mathcal{P}$ and $\Delta \cap A \neq \varnothing$, we can find an open rectangle $\square$ such that $\Delta \subseteq \square$ and $\nu(\square) < 2\nu(\Delta)$. Let $S_1, \cdots, S_N$ be those open rectangles $\square$. Then $A \subseteq \bigcup\limits_{k=1}^{N} S_k$ and $\sum\limits_{k=1}^{N} \nu(S_k) < \varepsilon$.

"(c)$\Rightarrow$(a)" W.L.O.G. we can assume that the ratio of the maximum length and minimum length of sides of S_k is less than 2 for all $k = 1, \cdots, N$ (otherwise we can divide S_k into smaller rectangles so that each smaller rectangle satisfies this requirement). Then each S_k can be covered by a closed rectangle $\square_k$ whose sides are parallel to the coordinate axes with the property that $\nu(\square_k) \leqslant 2^{n-1}\sqrt{n}^{\,n}\nu(S_k)$. Let $\mathcal{P}$ be a partition of A such that for each $\Delta \in \mathcal{P}$ with $\Delta \cap A \neq \varnothing$, $\Delta \subseteq \square_k$ for some $k = 1, \cdots, N$. Then

$$U(\mathbf{1}_A, \mathcal{P}) = \sum_{\substack{\Delta\in\mathcal{P}\\ \Delta\cap A\neq\varnothing}} \nu(\Delta) \leqslant \sum_{k=1}^{N} \nu(\square_k) \leqslant 2^{n-1}\sqrt{n}^{\,n} \sum_{k=1}^{N} \nu(S_k) < 2^{n-1}\sqrt{n}^{\,n}\varepsilon\,;$$

thus the upper integral $\overline{\int}_A \mathbf{1}_A(x)\,dx = 0$. Since the lower integral cannot be negative, we must have $\overline{\int}_A \mathbf{1}_A(x)\,dx = \underline{\int}_A \mathbf{1}_A(x)\,dx = 0$ which shows that A has volume zero. □

Example 8.15. Each point in $\mathbb{R}^n$ has volume zero.

Example 8.16. The Cantor set (defined in Exercise Problem 2.11) has volume zero.

Definition 8.17. A set $A \subseteq \mathbb{R}^n$ (not necessarily bounded) is said to ***have measure zero*** (測度為零) or be ***a set of measure zero*** (零測度集) if for every $\varepsilon > 0$, there exist countable many rectangles $S_1, S_2, \cdots$ such that $\{S_k\}_{k=1}^{\infty}$ is a cover of A $\big($that is, $A \subseteq \bigcup\limits_{k=1}^{\infty} S_k\big)$ and $\sum\limits_{k=1}^{\infty} \nu(S_k) < \varepsilon$.

Example 8.18. The real line $\mathbb{R} \times \{0\}$ on $\mathbb{R}^2$ has measure zero: for any given $\varepsilon > 0$, let $S_k = [-k, k] \times \big[\frac{-\varepsilon}{2^{k+3}k}, \frac{\varepsilon}{2^{k+3}k}\big]$. Then

$$\mathbb{R} \times \{0\} \subseteq \bigcup_{k=1}^{\infty} S_k \quad \text{and} \quad \sum_{k=1}^{\infty} \nu(S_k) = \sum_{k=1}^{\infty} 2k \cdot \frac{2\varepsilon}{2^{k+3}k} = \sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} = \frac{\varepsilon}{2} < \varepsilon\,.$$

Similarly, any hyperplane in $\mathbb{R}^n$ also has measure zero.

Proposition 8.19. *Let $A \subseteq \mathbb{R}^n$ be a set of measure zero. If $B \subseteq A$, then B also has measure zero.*

Modifying the proof of Proposition 8.14, we can also conclude the following

Proposition 8.20. *A set $A \subseteq \mathbb{R}^n$ has measure zero if and only if for every $\varepsilon > 0$, there exist countable many open rectangles $S_1, S_2, \cdots$ whose sides are parallel to the coordinate axes such that $A \subseteq \bigcup\limits_{k=1}^{\infty} S_k$ and $\sum\limits_{k=1}^{\infty} \nu(S_k) < \varepsilon$.*

Remark 8.21. If a set A has volume zero, then it has measure zero.

Proposition 8.22. *Let $K \subseteq \mathbb{R}^n$ be a compact set of measure zero. Then K has volume zero.*

Proof. Let $\varepsilon > 0$ be given. Then there are countable open rectangles $S_1, S_2, \cdots$ such that

$$K \subseteq \bigcup_{k=1}^{\infty} S_k \qquad \text{and} \qquad \sum_{k=1}^{\infty} \nu(S_k) < \varepsilon\,.$$

Since $\{S_k\}_{k=1}^{\infty}$ is an open cover of K, by the compactness of K there exists $N > 0$ such that $K \subseteq \bigcup\limits_{k=1}^{N} S_k$, while $\sum\limits_{k=1}^{N} \nu(S_k) \leqslant \sum\limits_{k=1}^{\infty} \nu(S_k) < \varepsilon$. As a consequence, K has volume zero. □

Since the boundary of a rectangle has measure zero, we also have the following

Corollary 8.23. *Let $S \subseteq \mathbb{R}^n$ be a bounded rectangle with positive volume. Then S is not a set of measure zero.*

Theorem 8.24. *If $A_1, A_2, \cdots$ are sets of measure zero in $\mathbb{R}^n$, then $\bigcup_{k=1}^{\infty} A_k$ has measure zero.*

Proof. Let $\varepsilon > 0$ be given. Since $A_k's$ are sets of measure zero, there exist countable rectangles $\{S_j^{(k)}\}_{j=1}^{\infty}$, such that

$$A_k \subseteq \bigcup_{j=1}^{\infty} S_j^{(k)} \quad \text{and} \quad \sum_{j=1}^{\infty} \nu(S_j^{(k)}) < \frac{\varepsilon}{2^{k+1}} \qquad \forall\, k \in \mathbb{N}.$$

Consider the collection consisting of all $S_j^{(k)}$'s. Since there are countable many rectangles in this collection, we can label them as $S_1, S_2, \cdots$, and we have

$$\bigcup_{k=1}^{\infty} A_k \subseteq \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} S_j^{(k)} = \bigcup_{\ell=1}^{\infty} S_\ell$$

and

$$\sum_{k=1}^{\infty} \nu(S_\ell) = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \nu(S_j^{(k)}) \leqslant \sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} = \frac{\varepsilon}{2} < \varepsilon\,.$$

Therefore, $\bigcup_{k=1}^{\infty} A_k$ has measure zero. □

Corollary 8.25. *The set of rational numbers in $\mathbb{R}$ has measure zero.*

Theorem 8.26. *Let $A \subseteq \mathbb{R}^n$ be bounded and $B \subseteq \mathbb{R}^m$ be a set of measure zero. Then $A \times B$ has measure zero in $\mathbb{R}^{n+m}$.*

Proof. Let $\varepsilon > 0$ be given. Since A is bounded, there exist a bounded rectangle R such that $A \subseteq R$. Since B has measure zero, there exist countable rectangles $\{S_k\}_{k=1}^{\infty} \subseteq \mathbb{R}^m$ such that

$$B \subseteq \bigcup_{k=1}^{\infty} S_k \quad \text{and} \quad \sum_{k=1}^{\infty} \nu_m(S_k) < \frac{\varepsilon}{\nu(R)}\,.$$

Then $A \times B \subseteq \bigcup_{k=1}^{\infty} (R \times S_k)$, and

$$\sum_{k=1}^{\infty} \nu_{n+m}(R \times S_k) = \sum_{k=1}^{\infty} \nu_n(R)\nu_m(S_k) = \nu_n(R) \sum_{k=1}^{\infty} \nu_m(S_k) < \varepsilon\,.$$

Since $R \times S_k$ is a rectangle for all $k \in \mathbb{N}$, we conclude that $A \times B$ has measure zero. □

8.3.2 The Lebesgue Theorem

在之前我們提到了函數 Riemann 可積的兩個等價條件：Riemann's condition 和 Darboux 定理。在這一節中，我們將引進函數是 Riemman 可積的另一個等價條件。這個等價條件說的是一個函數 f 在 A 上是 Riemann 可積的若且唯若 f 的延拓 $\overline{f}^A$（在函數可積分的定義中有定義）的不連續點所構成的集合其測度為零。為了證明這個敘述，我們先對一個函數的連續點做一個量化的刻劃。這個刻劃的方式，可以很容易用來檢驗一個函數在一個點是否連續。

Definition 8.27. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. For any $x \in \mathbb{R}^n$, the ***oscillation*** of f at x is the quantity

$$\operatorname{osc}(f,x) \equiv \inf_{\delta>0} \sup_{x_1,x_2\in D(x,\delta)} \big|f(x_1)-f(x_2)\big|.$$

我們注意到在上述定義中被取 infimum 的這個量 $h(\delta;x) \equiv \sup\limits_{x_1,x_2\in D(x,\delta)} \big|f(x_1)-f(x_2)\big|$ 是個 δ 的遞減函數（x 固定），而 $\operatorname{osc}(f,x)$ 則是 $h(\delta;x)$ 當 $\delta\to 0$ 時的極限。另外，我們也注意到說 $h(\delta;x)$ 也可以寫成 $\sup\limits_{y\in D(x,\delta)} f(y) - \inf\limits_{y\in D(x,\delta)} f(y)$.

以下的 Lemma 是關於如何檢驗一個函數在一個點是連續的。

Lemma 8.28. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, and $x_0 \in \mathbb{R}^n$. Then f is continuous at x_0 if and only if* $\operatorname{osc}(f,x_0)=0$.

Proof. "⇒" Let $\varepsilon>0$ be given. Since f is continuous at x_0,

$$\exists\,\delta>0 \ni \big|f(x)-f(x_0)\big| < \frac{\varepsilon}{3} \qquad \text{whenever } x\in D(x_0,\delta).$$

In particular, for any $x_1,x_2 \in D(x_0,\delta)$,

$$\big|f(x_1)-f(x_2)\big| \leqslant \big|f(x_1)-f(x_0)\big| + \big|f(x_0)-f(x_2)\big| < \frac{2\varepsilon}{3}\,;$$

thus $\sup\limits_{x_1,x_2\in D(x_0,\delta)} \big|f(x_1)-f(x_2)\big| \leqslant \frac{2\varepsilon}{3}$ which further implies that

$$0 \leqslant \operatorname{osc}(f,x_0) \leqslant \frac{2\varepsilon}{3} < \varepsilon.$$

Since ε is given arbitrarily, $\operatorname{osc}(f,x_0)=0$.

"⇐" Let $\varepsilon>0$ be given. By the definition of infimum, there exists $\delta>0$ such that

$$\sup_{x_1,x_2\in D(x_0,\delta)} \big|f(x_1)-f(x_2)\big| < \varepsilon\,.$$

In particular, $\big|f(x)-f(x_0)\big| \leqslant \sup\limits_{x_1,x_2\in D(x_0,\delta)} \big|f(x_1)-f(x_2)\big| < \varepsilon$ for all $x\in D(x_0,\delta)$. □

Lemma 8.29. *Let $f:\mathbb{R}^n\to\mathbb{R}$ be a function. Then for all $\varepsilon>0$, the set $D_\varepsilon=\big\{x\in\mathbb{R}^n\,\big|\,\operatorname{osc}(f,x)\geqslant\varepsilon\big\}$ is closed.*

Proof. Suppose that $\{y_k\}_{k=1}^\infty\subseteq D_\varepsilon$ and $y_k\to y$. Then for any $\delta>0$, there exists $N>0$ such that $y_k\in D(y,\delta)$ for all $k\geqslant N$. Since $D(y,\delta)$ is open, for each $k\geqslant N$ there exists $\delta_k>0$ such that $D(y_k,\delta_k)\subseteq D(y,\delta)$; thus we find that

$$\sup_{x_1,x_2\in D(y_k,\delta_k)}\big|f(x_1)-f(x_2)\big|\leqslant\sup_{x_1,x_2\in D(y,\delta)}\big|f(x_1)-f(x_2)\big|\qquad\forall\,k\geqslant N\,.$$

The inequality above implies that $\operatorname{osc}(f,y)\geqslant\varepsilon$; thus $y\in D_\varepsilon$ and D_ε is closed. □

Theorem 8.30 (Lebesgue)**.** *Let $A\subseteq\mathbb{R}^n$ be bounded, $f:A\to\mathbb{R}$ be a bounded function, and $\overline{f}^A$ be the extension of f by zero outside A; that is,*

$$\overline{f}^A(x)=\left\{\begin{array}{cl} f(x) & \textit{if } x\in A\,,\\ 0 & \textit{otherwise}\,.\end{array}\right.$$

Then f is Riemann integrable over A if and only if the collection of discontinuity of $\overline{f}^A$ is a set of measure zero.

Proof. Let $D=\big\{x\in\mathbb{R}^n\,\big|\,\operatorname{osc}(\overline{f}^A,x)>0\big\}$ and $D_\varepsilon=\big\{x\in\mathbb{R}^n\,\big|\,\operatorname{osc}(\overline{f}^A,x)\geqslant\varepsilon\big\}$. We remark here that $D=\bigcup\limits_{k=1}^\infty D_{\frac{1}{k}}$.

"⇒" We show that $D_{\frac{1}{k}}$ has measure zero for all $k\in\mathbb{N}$ (if so, then Theorem 8.24 implies that D has measure zero).

Let $k\in\mathbb{N}$ be fixed, and $\varepsilon>0$ be given. By Riemann's condition there exists a partition $\mathcal{P}$ of A such that

$$\sum_{\Delta\in\mathcal{P}}\Big[\sup_{x\in\Delta}\overline{f}^A(x)-\inf_{x\in\Delta}\overline{f}^A(x)\Big]\nu(\Delta)<\frac{\varepsilon}{k}\,.$$

Define

$$D^{(1)}_{\frac{1}{k}}\equiv\big\{x\in D_{\frac{1}{k}}\,\big|\,x\in\partial\Delta\text{ for some }\Delta\in\mathcal{P}\big\}\,,$$
$$D^{(2)}_{\frac{1}{k}}\equiv\big\{x\in D_{\frac{1}{k}}\,\big|\,x\in\operatorname{int}(\Delta)\text{ for some }\Delta\in\mathcal{P}\big\}\,.$$

Then $D_{\frac{1}{k}}=D^{(1)}_{\frac{1}{k}}\cup D^{(2)}_{\frac{1}{k}}$. We note that $D^{(1)}_{\frac{1}{k}}$ has measure zero since it is contained in $\bigcup\limits_{\Delta\in\mathcal{P}}\partial\Delta$ while each $\partial\Delta$ has measure zero. Now we show that $D^{(2)}_{\frac{1}{k}}$ also has measure zero. Let $C = \big\{\Delta \in \mathcal{P} \,\big|\, \text{int}(\Delta) \cap D_{\frac{1}{k}} \neq \varnothing\big\}$. Then $D^{(2)}_{\frac{1}{k}} \subseteq \bigcup\limits_{\Delta \in C} \Delta$. Moreover, we also note that if $\Delta \in C$, $\sup\limits_{x\in\Delta} \overline{f}^A(x) - \inf\limits_{x\in\Delta} \overline{f}^A(x) \geqslant \dfrac{1}{k}$. In fact, if $\Delta \in C$, there exists $y \in \text{int}(\Delta) \cap D_{\frac{1}{k}}$; thus choosing $\delta > 0$ such that $D(y,\delta) \subseteq \text{int}(\Delta)$,

$$\begin{aligned}\sup_{x\in\Delta} \overline{f}^A(x) - \inf_{x\in\Delta} \overline{f}^A(x) &= \sup_{x_1,x_2\in\Delta} \big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| \geqslant \sup_{x_1,x_2\in D(y,\delta)} \big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| \\ &\geqslant \inf_{\delta>0} \sup_{x_1,x_2\in D(y,\delta)} \big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| = \text{osc}(\overline{f}^A, y) \geqslant \frac{1}{k}.\end{aligned}$$

As a consequence,

$$\frac{1}{k} \sum_{\Delta\in C} \nu(\Delta) \leqslant \sum_{\Delta\in\mathcal{P}} \Big[\sup_{x\in\Delta} \overline{f}^A(x) - \inf_{x\in\Delta} \overline{f}^A(x)\Big] \nu(\Delta) = U(f,\mathcal{P}) - L(f,\mathcal{P}) < \frac{\varepsilon}{k}$$

which implies that $\sum\limits_{\Delta\in C} \nu(\Delta) < \varepsilon$. In other words, we establish that $D^{(2)}_{\frac{1}{k}}$ has measure zero. Therefore, $D_{\frac{1}{k}}$ has measure zero for all $k \in \mathbb{N}$; thus D has measure zero.

"$\Leftarrow$" Let R be a bounded closed rectangle with sides parallel to the coordinate axes and $\bar{A} \subseteq \text{int}(R)$, and $\varepsilon > 0$ be given. Define $\varepsilon' = \dfrac{\varepsilon}{2\|f\|_\infty + \nu(R) + 1}$, where $\|f\|_\infty = \sup\limits_{x\in A} |f(x)|$.

1. Since $D_{\varepsilon'}$ is a subset of D, Proposition 8.19 implies that $D_{\varepsilon'}$ has measure zero; thus Proposition 8.20 provides open rectangles $S_1, S_2, \cdots$ whose sides are parallel to the coordinate axes such that $D_{\varepsilon'} \subseteq \bigcup\limits_{k=1}^{\infty} S_k$, and $\sum\limits_{k=1}^{\infty} \nu(S_k) < \varepsilon'$. In addition, we can assume that $S_k \subseteq R$ for all $k \in \mathbb{N}$ since $D_{\varepsilon'} \subseteq R$.
2. Since $D_{\varepsilon'} \subseteq R$ is bounded, Lemma 8.29 implies that $D_{\varepsilon'}$ is compact; thus $D_{\varepsilon'} \subseteq \bigcup\limits_{k=1}^{N} S_k$ for some $N \in \mathbb{N}$.

Let $\square_k = \overline{S_k}$, and $\mathcal{P}_1$ be a partition of R satisfying

(a) For each $\Delta \in \mathcal{P}_1$ with $\Delta \cap D_{\varepsilon'} \neq \varnothing$, $\Delta \subseteq \square_k$ for some $k = 1, \cdots, N$.

(b) For each $k = 1, \cdots, N$, $\square_k$ is the union of rectangles in $\mathcal{P}_1$.

(c) Some collection of $\Delta \in \mathcal{P}_1$ forms a partition $\mathcal{P}_2$ of A.

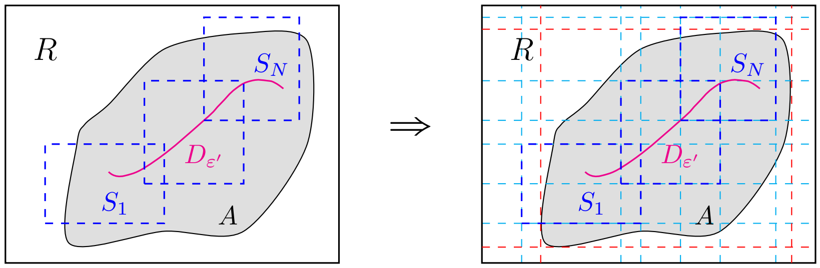

Figure 8.2: Constructing partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ from finite rectangles $S_1, \cdots, S_N$

Rectangles in $\mathcal{P}_1$ fall into two families: $C_1 = \big\{\Delta \in \mathcal{P}_1 \,\big|\, \Delta \subseteq \square_k \text{ for some } k = 1, \cdots, N\big\}$, and $C_2 = \big\{\Delta \in \mathcal{P}_1 \,\big|\, \Delta \nsubseteq \square_k \text{ for all } k = 1, \cdots, N\big\}$. By the definition of the oscillation function, for $x \notin D_{\varepsilon'}$ we let $\delta_x > 0$ be such that

$$\sup_{x \in D(y,\delta_y)} \overline{f}^A(y) - \inf_{x \in D(y,\delta_y)} \overline{f}^A(y) = \sup_{x_1,x_2 \in D(x,\delta_x)} \big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| < \varepsilon' .$$

Since $K = \bigcup\limits_{\Delta \in C_2} \Delta$ is compact, there exists $r > 0$ (the Lebesgue number associated with K and open cover $\bigcup\limits_{x \in K} D(x, \delta_x)\big)$ such that for each $a \in K$, $D(a, r) \subseteq D(y, \delta_y)$ for some $y \in K$. Let $\mathcal{P}'$ be a refinement of $\mathcal{P}_1$ such that $\|\mathcal{P}'\| < r$. Then if $\Delta' \in \mathcal{P}'$ satisfies that $\Delta' \subseteq \Delta$ for some $\Delta \in C_2$, we must have $\Delta' \subseteq D(y, \delta_y)$ for some $y \in K$; thus

$$\begin{aligned}\sup_{x \in \Delta'} \overline{f}^A(x) - \inf_{x \in \Delta'} \overline{f}^A(x) &\leqslant \sup_{x \in D(y,\delta_y)} \overline{f}^A(y) - \inf_{x \in D(y,\delta_y)} \overline{f}^A(y) \\ &= \sup_{x_1,x_2 \in D(y,\delta_y)} \big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| < \varepsilon'\end{aligned}$$

if $\Delta' \subseteq \Delta$ for some $\Delta \in C_2$. Let $\mathcal{P} = \big\{\Delta' \in \mathcal{P}' \,\big|\, \Delta' \subseteq \Delta \text{ for some } \Delta \in \mathcal{P}_2\big\}$. Then $\mathcal{P}$ is a partition of A and

$$\begin{aligned}U(f,\mathcal{P}) - L(f,\mathcal{P}) &= \Big(\sum_{\substack{\Delta' \in \mathcal{P}' \\ \Delta' \subseteq \Delta \in C_1}} + \sum_{\substack{\Delta' \in \mathcal{P}' \\ \Delta' \subseteq \Delta \in C_2}}\Big)\big(\sup_{x \in \Delta'} \overline{f}^A(x) - \inf_{x \in \Delta'} \overline{f}^A(x)\big)\nu(\Delta') \\ &\leqslant 2\|f\|_\infty \sum_{\substack{\Delta' \in \mathcal{P}' \\ \Delta' \subseteq \Delta \in C_1}} \nu(\Delta') + \varepsilon' \sum_{\substack{\Delta' \in \mathcal{P}' \\ \Delta' \subseteq \Delta \in C_2}} \nu(\Delta') \\ &\leqslant 2\|f\|_\infty \sum_{\Delta \in \mathcal{P} \cap C_1} \nu(\Delta) + \varepsilon'\nu(R) \\ &\leqslant 2\|f\|_\infty \sum_{k=1}^{N} \nu(S_k) + \varepsilon'\nu(R) < \big(2\|f\|_\infty + \nu(R)\big)\varepsilon' \leqslant \varepsilon ;\end{aligned}$$

thus f is Riemann integrable over A by Riemann's condition. □

Example 8.31. Let $A = \mathbb{Q} \cap [0,1]$, and $f : A \to \mathbb{R}$ be the constant function $f \equiv 1$. Then

$$\bar{f}(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0,1], \\ 0 & \text{otherwise}. \end{cases}$$

The collection of points of discontinuity of $\bar{f}$ is $[0,1]$ which, by Corollary 8.23, cannot be a set of measure zero; thus f is not Riemann integrable.

Another way to see that f is not Riemann integrable is $U(f,\mathcal{P}) = 1$ and $L(f,\mathcal{P}) = 0$ for all partitions $\mathcal{P}$ of A.

Corollary 8.32. *A bounded set $A \subseteq \mathbb{R}^n$ has volume if and only if the boundary of A has measure zero.*

Proof. It suffices to show that the collection of discontinuities of the function $\mathbf{1}_A$ (which is the same as $\overline{\mathbf{1}_A}^A$) is indeed ∂A.

1. If $x_0 \notin \partial A$, then there exists $\delta > 0$ such that either $D(x_0,\delta) \subseteq A$ or $D(x_0,\delta) \subseteq A^{\complement}$; thus $\mathbf{1}_A$ is continuous at $x_0 \notin \partial A$ since $\mathbf{1}_A(x)$ is constant for all $x \in D(x_0,\delta)$.
2. On the other hand, if $x_0 \in \partial A$, then there exists $x_k \in A$, $y_k \in A^{\complement}$ such that $x_k \to x_0$ and $y_k \to x_0$ as $k \to \infty$. This implies that $\mathbf{1}_A$ cannot be continuous at x_0 since $\mathbf{1}_A(x_k) = 1$ while $\mathbf{1}_A(y_k) = 0$ for all $k \in \mathbb{N}$.

As a consequence, the collection of discontinuity of $\mathbf{1}_A$ is exactly ∂A, and the corollary follows from Lebesgue's theorem. □

Corollary 8.33. *Let $A \subseteq \mathbb{R}^n$ be a bounded set with volume. A bounded function $f : A \to \mathbb{R}$ with a finite or countable number of points of discontinuity is Riemann integrable over A.*

Proof. We note that $\{x \in \mathbb{R}^n \,|\, \text{osc}(\bar{f},x) > 0\} \subseteq \partial A \cup \{x \in A \,|\, f \text{ is discontinuous at } x\}$. □

Remark 8.34. In addition to the set inclusion listed in the proof of Corollary 8.33, we also have

$$\{x \in A \,|\, f \text{ is discontinuous at } x\} \subseteq \{x \in \mathbb{R}^n \,|\, \text{osc}(\bar{f},x) > 0\}.$$

Therefore, if $A \subseteq \mathbb{R}^n$ is a bounded set with volume, then a bounded function $f : A \to \mathbb{R}$ is Riemann integrable if and only if the collection of points of discontinuity of f has measure zero.

Corollary 8.35. *A bounded function is integrable over a compact set of measure zero.*

Proof. If $f : K \to \mathbb{R}$ is bounded, and K is a compact set of measure zero, then the collection of discontinuities of $\bar{f}$ is a subset of K. ▫

Corollary 8.36. *Suppose that $A, B \subseteq \mathbb{R}^n$ are bounded sets with volume, and $f : A \to \mathbb{R}$ is Riemann integrable over A. Then f is Riemann integrable over $A \cap B$.*

Proof. By the inclusion

$$\{x \in \operatorname{int}(A \cap B) \,|\, \operatorname{osc}(\overline{f}^{A\cap B}, x) > 0\} \subseteq \{x \in \mathbb{R}^n \,|\, \operatorname{osc}(\overline{f}^{A}, x) > 0\},$$

we find that

$$\begin{aligned}\{x \in \mathbb{R}^n \,|\, \operatorname{osc}(\overline{f}^{A\cap B}, x) > 0\} &\subseteq \partial(A \cap B) \cup \{x \in \operatorname{int}(A \cap B) \,|\, \operatorname{osc}(\overline{f}^{A\cap B}, x) > 0\} \\ &\subseteq \partial A \cup \partial B \cup \{x \in \mathbb{R}^n \,|\, \operatorname{osc}(\overline{f}^{A}, x) > 0\}.\end{aligned}$$

Since ∂A and ∂B both have measure zero, the integrability of f over $A \cap B$ then follows from the integrability of f over A and the Lebesgue Theorem. ▫

Remark 8.37. Suppose that $A \subseteq \mathbb{R}^n$ is a bounded set of measure zero. Even if $f : A \to \mathbb{R}$ is continuous, f might not be Riemann integrable. For example, the function f given in Example 8.31 is not Riemann integrable even though f is continuous on A.

Remark 8.38. When $f : A \to \mathbb{R}$ is Riemann integrable over A, it is not necessary that A has volume. For example, the zero function is Riemann integrable over $A = \mathbb{Q} \cap [0, 1]$ even though A does not has volume.

Corollary 8.39 (Lebesgue's Differentiation Theorem, a simple version)**.** *Let $A \subseteq \mathbb{R}^n$ be a bounded set with volume, and $f : A \to \mathbb{R}$ be bounded and Riemann integrable over A. Then*

$$\lim_{r\to 0} \frac{1}{\nu(D(x_0, r) \cap A)} \int_{D(x_0,r)\cap A} f(x)\, dx = f(x_0) \tag{8.3.2}$$

for almost every $x_0 \in A$; that is, the equality above does not hold only for x_0 from a set of measure zero.

Proof. Let $\varepsilon > 0$ be given, and suppose that $\bar{f}$, the zero extension of f outside A, is continuous at x_0. Then there exists $\delta > 0$ such that

$$\big|\bar{f}(x) - \bar{f}(x_0)\big| < \frac{\varepsilon}{2} \qquad \forall\, x \in D(x_0, \delta) \cap A\,.$$

Since ∂A has measure zero, by the fact that $\partial(D(x_0,r)\cap A)\subseteq \partial D(x_0,r)\cup\partial A$ we find that $\partial(D(x_0,r)\cap A)$ also has measure zero for all $r>0$. In other words, $D(x_0,r)\cap A$ has volume. Then if $0<r<\delta$,

$$\begin{aligned}
&\Big|\frac{1}{\nu(D(x_0,r)\cap A)}\int_{D(x_0,r)\cap A} f(x)\,dx - f(x_0)\Big| \\
&\quad = \Big|\frac{1}{\nu(D(x_0,r)\cap A)}\int_{D(x_0,r)\cap A}\big(\bar{f}(x)-\bar{f}(x_0)\big)\,dx\Big| \\
&\quad \leqslant \frac{1}{\nu(D(x_0,r)\cap A)}\int_{D(x_0,r)\cap A}\big|\bar{f}(x)-\bar{f}(x_0)\big|\,dx \\
&\quad \leqslant \frac{\varepsilon}{2}\frac{1}{\nu(D(x_0,r)\cap A)}\int_{D(x_0,r)\cap A} 1\,dx = \frac{\varepsilon}{2} < \varepsilon\,.
\end{aligned}$$

This implies that (8.3.2) holds for all x_0 at which $\bar{f}$ is continuous. The theorem then follows from the Lebesgue theorem. ▫

8.4 Properties of the Integrals

Proposition 8.40. *Let $A\subseteq\mathbb{R}^n$ be bounded, and $f,g:A\to\mathbb{R}$ be bounded. Then*

(a) *If $B\subseteq A$, then $\underline{\int_A}(f\mathbf{1}_B)(x)\,dx = \underline{\int_B} f(x)\,dx$ and $\overline{\int_A}(f\mathbf{1}_B)(x)\,dx = \overline{\int_B} f(x)\,dx$.*

(b) $\underline{\int_A} f(x)\,dx + \underline{\int_A} g(x)\,dx \leqslant \underline{\int_A}(f+g)(x)\,dx \leqslant \overline{\int_A}(f+g)(x)\,dx \leqslant \overline{\int_A} f(x)\,dx + \overline{\int_A} g(x)\,dx.$

(c) *If $c\geqslant 0$, then $\underline{\int_A}(cf)(x)\,dx = c\underline{\int_A} f(x)\,dx$ and $\overline{\int_A}(cf)(x)\,dx = c\overline{\int_A} f(x)\,dx$.*

(d) *If $f\leqslant g$ on A, then $\underline{\int_A} f(x)\,dx \leqslant \underline{\int_A} g(x)\,dx$ and $\overline{\int_A} f(x)\,dx \leqslant \overline{\int_A} g(x)\,dx$.*

(e) *If A has volume zero, then f is Riemann integrable over A, and $\int_A f(x)\,dx = 0$.*

Proof. We only prove (a), (b), (c) and (e) since (d) are trivial.

(a) Let $\varepsilon>0$ be given. By the definition of the lower integral, there exist partition $\mathcal{P}_A$ of A and $\mathcal{P}_B$ of B such that

$$\underline{\int_A}(f\mathbf{1}_B)(x)\,dx - \varepsilon < L(f\mathbf{1}_B,\mathcal{P}_A) = \sum_{\Delta\in\mathcal{P}_A}\inf_{x\in\Delta}\overline{f\mathbf{1}_B}^{A}(x)\nu(\Delta)$$

and

$$\underline{\int}_B f(x)\,dx - \frac{\varepsilon}{2} < L(f, \mathcal{P}_B) = \sum_{\Delta \in \mathcal{P}_B} \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta)\,.$$

Let $\mathcal{P}'_A$ be a refinement of $\mathcal{P}_A$ such that some collection of rectangles in $\mathcal{P}'_A$ forms a partition of B. Denote this partition of B by $\mathcal{P}'_B$. Since $\inf\limits_{x\in\Delta} \overline{f}^B(x) \leqslant 0$ if $\Delta \in \mathcal{P}'_A \backslash \mathcal{P}'_B$, Proposition 8.6 implies that

$$\begin{aligned}
\underline{\int}_A (f\mathbf{1}_B)(x)\,dx - \varepsilon < L(f\mathbf{1}_B, \mathcal{P}_A) &\leqslant L(f\mathbf{1}_B, \mathcal{P}'_A) = \sum_{\Delta\in\mathcal{P}'_A} \inf_{x\in\Delta} \overline{f\mathbf{1}_B}^A(x)\nu(\Delta) \\
&= \Big(\sum_{\Delta\in\mathcal{P}'_A\backslash\mathcal{P}'_B} + \sum_{\Delta\in\mathcal{P}'_B} \Big) \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) \\
&\leqslant \sum_{\Delta\in\mathcal{P}'_B} \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) = L(f, \mathcal{P}'_B) \leqslant \underline{\int}_B f(x)\,dx\,.
\end{aligned}$$

On the other hand, let $\widetilde{\mathcal{P}}_A$ be a partition of A such that $\mathcal{P}_B \subseteq \widetilde{\mathcal{P}}_A$ and

$$\sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B\neq\varnothing} \nu(\Delta) \leqslant \frac{\varepsilon}{2(M+1)}\,,$$

where $M > 0$ is an upper bound of $|f|$. Then

$$\sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B\neq\varnothing} \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) \geqslant -M \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B\neq\varnothing} \overline{f}^B(x)\nu(\Delta) \geqslant -\frac{\varepsilon}{2}$$

which further implies that

$$\begin{aligned}
\underline{\int}_A (f\mathbf{1}_B)(x)\,dx \geqslant L(f\mathbf{1}_B, \widetilde{\mathcal{P}}_A) &= \sum_{\Delta\in\widetilde{\mathcal{P}}_A} \inf_{x\in\Delta} (\overline{f\mathbf{1}_B}^A(x)\nu(\Delta) \\
&= \Big(\sum_{\Delta\in\mathcal{P}_B} + \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B\neq\varnothing} + \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B=\varnothing} \Big) \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) \\
&= L(f, \mathcal{P}_B) + \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B\neq\varnothing} \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) > \underline{\int}_B f(x)\,dx - \varepsilon\,.
\end{aligned}$$

Therefore, we establish that

$$\underline{\int}_B f(x)\,dx - \varepsilon < \underline{\int}_A (f\mathbf{1}_B)(x)\,dx < \underline{\int}_B f(x)\,dx + \varepsilon\,.$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that $\underline{\int}_A (f\mathbf{1}_B)(x)\,dx = \underline{\int}_B f(x)\,dx$. Similar argument can be applied to conclude that $\overline{\int}_A (f\mathbf{1}_B)(x)\,dx = \overline{\int}_B f(x)\,dx$.

(b) Let $\varepsilon > 0$ be given. By the definition of the lower integral, there exist partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of A such that

$$\underline{\int}_A f(x)\,dx - \frac{\varepsilon}{2} < L(f,\mathcal{P}_1) \quad\text{and}\quad \underline{\int}_A g(x)\,dx - \frac{\varepsilon}{2} < L(g,\mathcal{P}_2)\,.$$

Let $\mathcal{P}$ be a common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Then

$$\begin{aligned}\underline{\int}_A f(x)\,dx + \underline{\int}_A g(x)\,dx - \varepsilon < L(f,\mathcal{P}_1) + L(f,\mathcal{P}_2) &\leqslant L(f,\mathcal{P}) + L(g,\mathcal{P})\\ &= \sum_{\Delta\in\mathcal{P}} \inf_{x\in\Delta} \bar{f}(x)\nu(\Delta) + \sum_{\Delta\in\mathcal{P}} \inf_{x\in\Delta} \bar{g}(x)\nu(\Delta)\\ &\leqslant \sum_{\Delta\in\mathcal{P}} \inf_{x\in\Delta} (\bar{f}+\bar{g})(x)\nu(\Delta) = L(f+g,\mathcal{P}) \leqslant \underline{\int}_A (f+g)(x)\,dx\,.\end{aligned}$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that

$$\underline{\int}_A f(x)\,dx + \underline{\int}_A g(x)\,dx \leqslant \underline{\int}_A (f+g)(x)\,dx\,.$$

Similarly, we have $\overline{\int}_A (f+g)(x)\,dx \leqslant \overline{\int}_A f(x)\,dx + \overline{\int}_A g(x)\,dx$; thus (b) is established.

(c) It suffices to show the case $c = -1$. For each $k \in \mathbb{N}$, there exist partitions $\mathcal{P}_k$ and $\mathcal{Q}_k$ of A such that

$$\underline{\int}_A -f(x)\,dx - \frac{1}{k} < L(-f,\mathcal{P}_k) \leqslant \underline{\int}_A -f(x)\,dx$$

and

$$\overline{\int}_A f(x)\,dx \leqslant U(f,\mathcal{Q}_k) < \overline{\int}_A f(x)\,dx + \frac{1}{k}\,.$$

Let $\mathcal{R}_k$ be the common refinement of $\mathcal{P}_k$ and $\mathcal{Q}_k$. Then

$$\underline{\int}_A -f(x)\,dx - \frac{1}{k} < L(-f,\mathcal{P}_k) \leqslant L(-f,\mathcal{R}_k) \leqslant \underline{\int}_A -f(x)\,dx$$

and

$$\overline{\int}_A f(x)\,dx \leqslant U(f,\mathcal{R}_k) \leqslant U(f,\mathcal{Q}_k) < \overline{\int}_A f(x)\,dx + \frac{1}{k}\,.$$

which implies that

$$\lim_{k\to\infty} U(f,\mathcal{R}_k) = \overline{\int}_A f(x)\,dx \quad\text{and}\quad \lim_{k\to\infty} L(-f,\mathcal{R}_k) = \underline{\int}_A -f(x)\,dx$$

Since

$$L(-f,\mathcal{R}_k)=\sum_{\Delta\in\mathcal{R}_k}\inf_{x\in\Delta}\overline{(-f)}^{A}(x)\nu(\Delta)=-\sum_{\Delta\in\mathcal{R}_k}\sup_{x\in\Delta}\overline{f}^{A}(x)\nu(\Delta)=U(f,\mathcal{R}_k)\,,$$

we conclude that $\underline{\int_A}-f(x)\,dx=-\overline{\int_A}f(x)\,dx$.

(e) Since f is bounded on A, there exist $M>0$ such that $-M\leqslant f(x)\leqslant M$ for all $x\in A$. Therefore, $-\mathbf{1}_A\leqslant \frac{f}{M}\leqslant \mathbf{1}_A$ on A; thus (c) and (d) imply that

$$0=\int_A\mathbf{1}_A(x)\,dx=\overline{\int_A}\mathbf{1}_A(x)\,dx\geqslant\overline{\int_A}\frac{f(x)}{M}\,dx=\frac{1}{M}\overline{\int_A}f(x)\,dx$$

which implies that $\overline{\int_A}f(x)\,dx\leqslant 0$. Similarly, $\overline{\int_A}-f(x)\,dx\leqslant 0$ which further implies that $\underline{\int_A}f(x)\,dx\geqslant 0$. Therefore, by Corollary 8.7 we conclude that

$$0\leqslant\underline{\int_A}f(x)\,dx\leqslant\overline{\int_A}f(x)\,dx\leqslant 0$$

which implies that f is Riemann integrable over A and $\int_A f(x)\,dx=0$. □

Remark 8.41. Let $A\subseteq\mathbb{R}^n$ be bounded and $f,g:A\to\mathbb{R}$ be bounded. Then (b) of Proposition 8.40 also implies that

$$\underline{\int_A}(f-g)(x)\,dx\leqslant\underline{\int_A}f(x)\,dx-\underline{\int_A}g(x)\,dx\quad\text{and}\quad\overline{\int_A}f(x)\,dx-\overline{\int_A}g(x)\,dx\leqslant\overline{\int_A}(f-g)(x)\,dx\,.$$

Corollary 8.42. *Let $A,B\subseteq\mathbb{R}^n$ be bounded such that $A\cap B$ has volume zero, and $f:A\cup B\to\mathbb{R}$ be bounded. Then*

$$\underline{\int_A}f(x)\,dx+\underline{\int_B}f(x)\,dx\leqslant\underline{\int_{A\cup B}}f(x)\,dx\leqslant\overline{\int_{A\cup B}}f(x)\,dx\leqslant\overline{\int_A}f(x)\,dx+\overline{\int_B}f(x)\,dx\,.$$

Proof. Note that $f\mathbf{1}_A+f\mathbf{1}_B=f\mathbf{1}_{A\cup B}+f\mathbf{1}_{A\cap B}$ on $A\cup B$. Therefore, (a), (b) of Proposition 8.40 and Remark 8.41 implies that

$$\begin{aligned}\underline{\int_A}f(x)\,dx+\underline{\int_B}f(x)\,dx&=\underline{\int_{A\cup B}}(f\mathbf{1}_A)(x)\,dx+\underline{\int_{A\cup B}}(f\mathbf{1}_B)(x)\,dx\leqslant\underline{\int_{A\cup B}}(f\mathbf{1}_A+f\mathbf{1}_B)(x)\,dx\\&=\underline{\int_{A\cup B}}\big(f\mathbf{1}_{A\cup B}-(-f\mathbf{1}_{A\cap B})\big)(x)\,dx\\&\leqslant\underline{\int_{A\cup B}}f\mathbf{1}_{A\cup B}(x)\,dx-\underline{\int_{A\cup B}}(-f\mathbf{1}_{A\cap B})(x)\,dx\\&=\underline{\int_{A\cup B}}f(x)\,dx-\underline{\int_{A\cap B}}(-f)(x)\,dx\end{aligned}$$

which, with the help of Proposition 8.40 (e), further implies that

$$\underline{\int_A} f(x)\,dx + \underline{\int_B} f(x)\,dx \leqslant \underline{\int_{A\cup B}} f(x)\,dx\,.$$

The case of the upper integral can be proved in a similar fashion. □

Having established Proposition 8.40, it is easy to see the following theorem (except (c)). The proof is left as an exercise.

Theorem 8.43. *Let $A \subseteq \mathbb{R}^n$ be bounded, $c \in \mathbb{R}$, and $f, g : A \to \mathbb{R}$ be Riemann integrable. Then*

(a) *$f \pm g$ is Riemann integrable, and* $\displaystyle\int_A (f \pm g)(x)\,dx = \int_A f(x)\,dx \pm \int_A g(x)\,dx$.

(b) *cf is Riemann integrable, and* $\displaystyle\int_A (cf)(x)\,dx = c\int_A f(x)\,dx$.

(c) *$|f|$ is Riemann integrable, and* $\displaystyle\left|\int_A f(x)\,dx\right| \leqslant \int_A |f(x)|dx$.

(d) *If $f \leqslant g$, then* $\displaystyle\int_A f(x)\,dx \leqslant \int_A g(x)\,dx$.

(e) *If A has volume and $|f| \leqslant M$, then* $\displaystyle\left|\int_A f(x)\,dx\right| \leqslant M\nu(A)$.

Theorem 8.44. *Let $A \subseteq \mathbb{R}^n$ be bounded, and $f : A \to \mathbb{R}$ be a bounded integrable function.*

1. *If A has measure zero, then* $\displaystyle\int_A f(x)\,dx = 0$.

2. *If $f(x) \geqslant 0$ for all $x \in A$, and $\displaystyle\int_A f(x)\,dx = 0$, then the set $\{x \in A \,|\, f(x) \neq 0\}$ has measure zero.*

Proof. 1. We show that $L(f, \mathcal{P}) \leqslant 0 \leqslant U(f, \mathcal{P})$ for all partitions $\mathcal{P}$ of A. Let $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ be a partition of A. By Corollary 8.23, $\Delta_k \cap A^\complement \neq \varnothing$ for $k = 1, \cdots, N$; thus we must have $\inf\limits_{x\in\Delta_k} \bar{f}(x) \leqslant 0$ and $\sup\limits_{x\in\Delta_k} \bar{f}(x) \geqslant 0$. As a consequence, if $\mathcal{P}$ is a partition of A,

$$L(f, \mathcal{P}) = \sum_{k=1}^{N} \inf_{x\in\Delta_k} \bar{f}(x)\nu(\Delta_k) \leqslant 0 \quad \text{and} \quad U(f, \mathcal{P}) = \sum_{k=1}^{N} \sup_{x\in\Delta_k} \bar{f}(x)\nu(\Delta_k) \geqslant 0\,;$$

thus $\displaystyle\underline{\int_A} f(x)\,dx \leqslant 0 \leqslant \overline{\int_A} f(x)dx$. Since f is integrable over A, $\displaystyle\int_A f(x)\,dx = 0$.

2. Let $A_k = \left\{x \in A \,\middle|\, f(x) \geqslant \dfrac{1}{k}\right\}$. We claim that A_k has measure zero for all $k \in \mathbb{N}$.

   Let $\varepsilon > 0$ be given. Since $\displaystyle\int_A f(x)dx = 0$, there exists a partition $\mathcal{P}$ of A such that $U(f,\mathcal{P}) < \dfrac{\varepsilon}{k}$. Let $C = \left\{\Delta \in \mathcal{P} \,\middle|\, \Delta \cap A_k \neq \varnothing\right\}$. Then $A_k \subseteq \bigcup\limits_{\Delta \in C} \Delta$, and

$$\frac{1}{k}\sum_{\Delta \in C} \nu(C) \leqslant \sum_{\Delta \in C} \sup_{x \in \Delta} \bar{f}(x)\nu(\Delta) \leqslant \sum_{\Delta \in \mathcal{P}} \sup_{x \in \Delta} \bar{f}(x)\nu(\Delta) = U(f,\mathcal{P}) < \frac{\varepsilon}{k}$$

   which implies that $\sum\limits_{\Delta \in C} \nu(\Delta) < \varepsilon$. Therefore, A_k has measure zero; thus Theorem 8.24 implies that $A = \bigcup\limits_{k=1}^{\infty} A_k$ also has measure zero. □

Remark 8.45. Combining Corollary 8.35 and Theorem 8.44, we conclude that the integral of a bounded function over a compact set of measure zero is zero.

Remark 8.46. Let $A = \mathbb{Q} \cap [0,1]$ and $f : A \to \mathbb{R}$ be the constant function $f \equiv 1$. We have shown in Example 8.31 that f is not Riemann integrable. We note that A has no volume since $\partial A = [0,1]$ which is not a set of measure zero. However, A has measure zero since it consists of countable number of points.

1. Since f is continuous on A, the condition that A has volume in Corollary 8.33 cannot be removed.

2. Since A has measure zero, the condition that f is Riemann integrable in Theorem 8.44 cannot be removed.

Theorem 8.47 (Mean Value Theorem for Integrals)**.** *Let A be a subset of $\mathbb{R}^n$ such that A has volume and is compact and connected. Suppose that $f : A \to \mathbb{R}$ is continuous, then there exists $x_0 \in A$ such that*

$$\int_A f(x)\,dx = f(x_0)\nu(A)\,.$$

The quantity $\dfrac{1}{\nu(A)}\displaystyle\int_A f(x)\,dx$ is called the ***average*** *of f over A.*

Proof. Because of Theorem 8.44, it suffices to show the case that $\nu(A) \neq 0$. Let $m = \min\limits_{x \in A} f(x)$ and $M = \max\limits_{x \in A} f(x)$. Then

$$m\mathbf{1}_A(x) \leqslant f(x) \leqslant M\mathbf{1}_A(x)\,;$$

thus (b) and (d) of Theorem 8.43 imply that

$$m\nu(A) = \int_A m\mathbf{1}_A(x)\,dx \leqslant \int_A f(x)\,dx \leqslant \int_A M\mathbf{1}_A(x)\,dx = M\nu(A)\,.$$

By the connectedness of A and continuity of f, Theorem 4.21 and Theorem 3.38 implies that $f(A) = [m, M]$; thus by the fact that the quantity $\dfrac{1}{\nu(A)}\displaystyle\int_A f(x)\,dx \in [m, M]$, there must be $x_0 \in A$ such that

$$f(x_0) = \frac{1}{\nu(A)}\int_A f(x)\,dx\,. \qquad \square$$

Definition 8.48. Let $A \subseteq \mathbb{R}^n$ be a set and $f : A \to \mathbb{R}$ be a function. For $B \subseteq A$, the ***restriction of*** f ***to*** B is the function $f\big|_B : A \to \mathbb{R}$ given by $f|_B = f\mathbf{1}_B$. In other words,

$$f\big|_B(x) = \begin{cases} f(x) & \text{if } x \in B\,, \\ 0 & \text{if } x \in A\backslash B\,. \end{cases}$$

The following lemma is a direct consequence of Proposition 8.40 (a).

Lemma 8.49. *Let $A \subseteq \mathbb{R}^n$ be bounded, and $f : A \to \mathbb{R}$ be a bounded function. Suppose that $B \subseteq A$, and $f\big|_B$ is Riemann integrable over A. Then f is Riemann integrable over B, and*

$$\int_A f\big|_B(x)\,dx = \int_B f(x)\,dx\,.$$

Theorem 8.50. *Let A, B be bounded subsets of $\mathbb{R}^n$ be such that $A \cap B$ has measure zero, and $f : A \cup B \to \mathbb{R}$ be such that $f\big|_{A\cap B}$, $f\big|_A$ and $f\big|_B$ are all Riemann integrable over $A \cup B$. Then f is integrable over $A \cup B$, and*

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

Proof. Since $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{A\cap B}$, we have

$$f = f\mathbf{1}_{A\cup B} = f\big|_A + f\big|_B - f\big|_{A\cap B}\,;$$

thus Theorem 8.43, Theorem 8.44 and Lemma 8.49 imply that

$$\int_{A\cup B} f(x)\,dx = \int_{A\cup B} f\big|_A(x)\,dx + \int_{A\cup B} f\big|_B(x)\,dx = \int_A f(x)\,dx + \int_B f(x)\,dx\,. \qquad \square$$

8.5 The Fubini Theorem

If $f:[a,b]\to\mathbb{R}$ is continuous, the fundamental theorem of Calculus (Theorem 4.90) can be applied to computed the integral of f over $[a,b]$. In the following two sections, we focus on how the integral of f over $A\subseteq\mathbb{R}^n$, where $n\geqslant 2$, can be computed if the integral exists. We start with the special case $n=2$.

Definition 8.51. Let $S=[a,b]\times[c,d]$ be a rectangle in $\mathbb{R}^2$, and $f:S\to\mathbb{R}$ be bounded. For each fixed $x\in[a,b]$, the lower integral of the function $f(x,\cdot):[c,d]\to\mathbb{R}$ is denoted by $\underline{\int_c^d} f(x,y)\,dy$, and the upper integral of $f(x,\cdot):[c,d]\to\mathbb{R}$ is denoted by $\overline{\int_c^d} f(x,y)\,dy$. If for each $x\in[a,b]$ the upper integral and the lower integral of $f(x,\cdot):[c,d]\to\mathbb{R}$ are the same, we simply write $\int_c^d f(x,y)\,dy$ for the integrals of $f(x,\cdot)$ over $[c,d]$. The integrals $\overline{\int_a^b} f(x,y)\,dx$, $\underline{\int_a^b} f(x,y)\,dx$ and $\int_a^b f(x,y)\,dx$ are defined in a similar way.

Lemma 8.52. *Let $A=[a,b]\times[c,d]$ be a rectangle in $\mathbb{R}^2$, and $f:A\to\mathbb{R}$ be bounded. Then*

$$\underline{\int_A} f(x,y)\,d\mathbb{A}\leqslant\underline{\int_a^b}\Big(\underline{\int_c^d} f(x,y)\,dy\Big)dx\leqslant\overline{\int_a^b}\Big(\overline{\int_c^d} f(x,y)\,dy\Big)dx\leqslant\overline{\int_A} f(x,y)\,d\mathbb{A}\tag{8.5.1}$$

and

$$\underline{\int_A} f(x,y)\,d\mathbb{A}\leqslant\underline{\int_c^d}\Big(\underline{\int_a^b} f(x,y)\,dx\Big)dy\leqslant\overline{\int_c^d}\Big(\overline{\int_a^b} f(x,y)\,dx\Big)dy\leqslant\overline{\int_A} f(x,y)\,d\mathbb{A}\,.\tag{8.5.2}$$

Proof. It suffices to prove (8.5.1). Let $\varepsilon>0$ be given. Choose a partition

$$\mathcal{P}=\big\{\Delta_{ij}\,\big|\,\Delta_{ij}=[x_i,x_{i+1}]\times[y_j,y_{j+1}]\text{ for }i=0,1,\cdots,n-1\text{ and }j=0,1,\cdots,m-1\big\}$$

of A such that $L(f,\mathcal{P})>\underline{\int_A} f(x,y)\,d\mathbb{A}-\varepsilon$. Using (4.7.3) and Remark 4.82, we find that

$$\begin{aligned}\underline{\int_a^b}\Big(\underline{\int_c^d} f(x,y)\,dy\Big)dx&=\sum_{i=0}^{n-1}\underline{\int_{x_i}^{x_{i+1}}}\Big(\sum_{j=0}^{m-1}\underline{\int_{y_j}^{y_{j+1}}} f(x,y)\,dy\Big)dx\\&\geqslant\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\underline{\int_{x_i}^{x_{i+1}}}\Big(\underline{\int_{y_j}^{y_{j+1}}} f(x,y)\,dy\Big)dx\\&\geqslant\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\inf_{(x,y)\in\Delta_{ij}} f(x,y)\nu(\Delta_{ij})=L(f,\mathcal{P})>\underline{\int_A} f(x,y)\,d\mathbb{A}-\varepsilon\,.\end{aligned}$$

Since $\varepsilon > 0$ is given arbitrarily, we must have

$$\underline{\int_a^b} \Big(\underline{\int_c^d} f(x,y)\,dy \Big) dx \geqslant \underline{\int_A} f(x,y)\,d\mathbb{A}\,.$$

Similarly, $\overline{\int_a^b} \Big(\overline{\int_c^d} f(x,y)\,dy \Big) dx \leqslant \overline{\int_A} f(x,y)\,d\mathbb{A}$, so (8.5.1) is concluded. □

Theorem 8.53 (Fubini's Theorem, the case $n = 2$)**.** *Let $A = [a,b] \times [c,d]$ be a rectangle in $\mathbb{R}^2$, and $f : A \to \mathbb{R}$ be Riemann integrable. Then*

1. *the functions $\underline{\int_c^d} f(\cdot,y)\,dy$ and $\overline{\int_c^d} f(\cdot,y)\,dy$ are Riemann integrable over $[a,b]$;*
2. *the functions $\underline{\int_a^b} f(x,\cdot)dx$ and $\overline{\int_a^b} f(x,\cdot)dx$ are Riemann integrable over $[c,d]$, and*
3. *The integral of f over A is the same as the iterated integrals; that is,*

$$\begin{aligned}\int_A f(x,y)\,d\mathbb{A} &= \int_a^b \Big(\underline{\int_c^d} f(x,y)\,dy \Big) dx = \int_a^b \Big(\overline{\int_c^d} f(x,y)\,dy \Big) dx \\ &= \int_c^d \Big(\underline{\int_a^b} f(x,y)\,dx \Big) dy = \int_c^d \Big(\overline{\int_a^b} f(x,y)\,dx \Big) dy\,.\end{aligned}$$

Proof. It suffices to prove that $\underline{\int_c^d} f(x,y)\,dy$ is Riemann integrable over $[a,b]$ and

$$\int_a^b \Big(\underline{\int_c^d} f(x,y)\,dy \Big) dx = \int_A f(x,y)\,d\mathbb{A}\,. \tag{8.5.3}$$

Since $\underline{\int_a^b} \Big(\underline{\int_c^d} f(x,y)\,dy \Big) dx \leqslant \overline{\int_a^b} \Big(\underline{\int_c^d} f(x,y)\,dy \Big) dx$, Lemma 8.52 implies that

$$\begin{aligned}\underline{\int_A} f(x,y)\,d\mathbb{A} &\leqslant \underline{\int_a^b} \Big(\underline{\int_c^d} f(x,y)\,dy \Big) dx \leqslant \overline{\int_a^b} \Big(\underline{\int_c^d} f(x,y)\,dy \Big) dx \\ &\leqslant \overline{\int_a^b} \Big(\overline{\int_c^d} f(x,y)\,dy \Big) dx \leqslant \overline{\int_A} f(x,y)\,d\mathbb{A}\,.\end{aligned}$$

The integrability of $\underline{\int_c^d} f(x,y)\,dy$ and the validity of (8.5.3) are then concluded by the integrability of f over A. □

Remark 8.54. To simplify the notation, sometimes we use $\displaystyle\int_a^b\int_c^d f(x,y)\,dydx$ to denote the iterated integral the iterated integral $\displaystyle\int_a^b\Big(\int_c^d f(x,y)\,dy\Big)dx$. Similar notation applies to the upper and the lower integrals. For example, we also have $\displaystyle\overline{\int_a^b}\underline{\int_c^d} f(x,y)\,dydx = \overline{\int_a^b}\Big(\underline{\int_c^d} f(x,y)\,dy\Big)dx$.

Remark 8.55. For each $x \in [a,b]$, define $\displaystyle\varphi(x) = \underline{\int_c^d} f(x,y)\,dy$ and $\displaystyle\psi(x) = \overline{\int_c^d} f(x,y)\,dy$. Then $\varphi(x) \leqslant \psi(x)$ for all $x \in [a,b]$, and the Fubini Theorem implies that

$$\int_a^b \big[\psi(x) - \varphi(x)\big]\,dx = 0\,.$$

By Theorem 8.44, the set $\big\{x \in [a,b] \,\big|\, \psi(x) - \varphi(x) \neq 0\big\}$ has measure zero. In other words, except on a set of measure zero, $f(x,\cdot)$ is Riemann integrable over $[c,d]$ if f is Riemann integrable over $[a,b]\times[c,d]$. This property can be rephrased as that "$f(x,\cdot)$ is Riemann integrable over $[c,d]$ for **almost every** $x \in [a,b]$ if f is Riemann integrable over the rectangle $[a,b]\times[c,d]$". Similarly, $f(\cdot,y)$ is Riemann integrable for almost every $y \in [c,d]$ if f is Riemann integrable over $[a,b]\times[c,d]$.

Remark 8.56. The integrability of f does not guarantee that $f(x,\cdot)$ or $f(\cdot,y)$ is Riemann integrable. In fact, there exists a function $f : [0,1]\times[0,1] \to \mathbb{R}$ such that f is Riemann integrable, $f(\cdot,y)$ is Riemann integrable for each $y \in [0,1]$, but $f(x,\cdot)$ is not Riemann integrable for infinitely many $x \in [0,1]$. For example, let

$$f(x,y) = \begin{cases} 0 & \text{if } x = 0 \text{ or if } x \text{ or } y \text{ is irrational}\,, \\ \dfrac{1}{p} & \text{if } x,y \in \mathbb{Q} \text{ and } x = \dfrac{q}{p} \text{ with } (p,q) = 1\,. \end{cases}$$

Then

1. For each $y \in [0,1]$, $f(\cdot,y)$ is continuous at all irrational numbers. Therefore, $f(\cdot,y)$ is Riemann integrable, and $\displaystyle\int_0^1 f(x,y)\,dx = \underline{\int_0^1} f(x,y)\,dx = 0$.

2. For $x = 0$ or $x \notin \mathbb{Q}$, $f(x,\cdot)$ is Riemann integrable, and $\displaystyle\int_0^1 f(x,y)\,dy = 0$.

3. If $x = \dfrac{q}{p}$ with $(p,q) = 1$, $f(x,\cdot)$ is nowhere continuous in $[0,1]$. In fact, for each $y_0 \in [0,1]$,
$$\lim_{\substack{y\to y_0\\ y\in\mathbb{Q}}} f(x,y) = \frac{1}{p} \quad\text{while}\quad \lim_{\substack{y\to y_0\\ y\notin\mathbb{Q}}} f(x,y) = 0\,;$$
thus the limit of $f(x,y)$ as $y \to y_0$ does not exist. Therefore, the Lebesgue theorem implies that $f(x,\cdot)$ is not Riemann integrable if $x \in \mathbb{Q}\cap(0,1]$. On the other hand, for $x = \dfrac{q}{p}$ with $(p,q) = 1$ we have
$$\underline{\int_0^1} f(x,y)\,dy = 0 \quad\text{and}\quad \overline{\int_0^1} f(x,y)\,dy = \frac{1}{p}\,.$$

4. Define $\varphi(x) = \underline{\int_0^1} f(x,y)\,dy$ and $\psi(x) = \overline{\int_0^1} f(x,y)\,dy$. Then 2 and 3 imply that φ and ψ are Riemann integrable over $[0,1]$, and $\int_0^1 \varphi(x)dx = \int_0^1 \psi(x)dx = 0$.

5. For each $a \notin \mathbb{Q}\cap[0,1]$ and $b \in [0,1]$, f is continuous at (a,b). In fact, for any given $\varepsilon > 0$, there exists a prime number p such that $\dfrac{1}{p} < \varepsilon$. Let
$$\delta = \min\Big\{\Big|a - \frac{\ell}{k}\Big| \,\Big|\, 0 \leqslant \ell \leqslant k \leqslant p, k\in\mathbb{N}, \ell\in\mathbb{N}\cup\{0\}\Big\}\,.$$
Then $\delta > 0$, and if $(x,y) \in D\big((a,b),\delta\big) \cap ([0,1]\times[0,1])$, we have
$$\big|f(x,y) - f(a,b)\big| = \big|f(x,y)\big| < \frac{1}{p} < \varepsilon\,,$$
where we use the fact that if $(x,y) \in D\big((a,b),\delta\big)$ and $x \in \mathbb{Q}$, then $x = \dfrac{\ell}{k}$ (in reduced form) for some $k > p$.

As a consequence, $\big\{(a,b) \in \mathbb{R}^2 \,\big|\, \bar{f}$ is discontinuous at $(a,b)\big\} \subseteq \mathbb{Q}\times[0,1]$. Since $\mathbb{Q}\times[0,1]$ is a countable union of measure zero sets, it has measure zero; thus f is Riemann integrable by the Lebesgue theorem. The Fubini theorem then implies that
$$\int_{[0,1]\times[0,1]} f(x,y)\,d\mathbb{A} = \int_0^1\int_0^1 f(x,y)\,dxdy = 0\,.$$

Remark 8.57. The integrability of $f(x,\cdot)$ and $f(\cdot,y)$ does not guarantee the integrability of f. In fact, there exists a bounded function $f : [0,1]\times[0,1] \to \mathbb{R}$ such that $f(x,\cdot)$ and $f(\cdot,y)$

are both Riemann integrable over $[0,1]$, but f is not Riemann integrable over $[0,1]\times[0,1]$. For example, let

$$f(x,y)=\begin{cases} 1 & \text{if } (x,y)=\big(\frac{k}{2^n},\frac{\ell}{2^n}\big),\ 0<k,\ell<2^n \text{ odd numbers, } n\in\mathbb{N}\,,\\ 0 & \text{otherwise}\,. \end{cases}$$

Then for each $x\in[0,1]$, $f(x,\cdot)$ only has finite number of discontinuities; thus $f(x,\cdot)$ is Riemann integrable, and

$$\int_0^1 f(x,y)\,dy=0\,.$$

Similarly, $f(\cdot,y)$ is Riemann integrable, and $\displaystyle\int_0^1 f(x,y)\,dx=0$. As a consequence,

$$\int_0^1\int_0^1 f(x,y)\,dydx=\int_0^1\int_0^1 f(x,y)\,dxdy=0\,.$$

However, note that f is nowhere continuous on $[0,1]\times[0,1]$; thus the Lebesgue theorem implies that f is not Riemann integrable. One can also see this by the fact that $U(f,\mathcal{P})=1$ and $L(f,\mathcal{P})=0$ for all partition of $[0,1]\times[0,1]$.

Corollary 8.58. 1. *Let $\varphi_1,\varphi_2:[a,b]\to\mathbb{R}$ be continuous maps such that $\varphi_1(x)\leqslant\varphi_2(x)$ for all $x\in[a,b]$, $A=\big\{(x,y)\,\big|\,a\leqslant x\leqslant b,\varphi_1(x)\leqslant y\leqslant\varphi_2(x)\big\}$, and $f:A\to\mathbb{R}$ be continuous. Then f is Riemann integrable over A, and*

$$\int_A f(x,y)\,d\mathbb{A}=\int_a^b\Big(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy\Big)dx\,.$$

2. *Let $\psi_1,\psi_2:[c,d]\to\mathbb{R}$ be continuous maps such that $\psi_1(y)\leqslant\psi_2(y)$ for all $y\in[c,d]$, $A=\big\{(x,y)\,\big|\,c\leqslant y\leqslant d,\psi_1(y)\leqslant x\leqslant\psi_2(y)\big\}$, and $f:A\to\mathbb{R}$ be continuous. Then f is Riemann integrable over A, and*

$$\int_A f(x,y)\,d\mathbb{A}=\int_c^d\Big(\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\,dx\Big)dy\,.$$

Proof. It suffices to prove 1. First we show that f is Riemann integrable over A. By Lebesgue's theorem, it suffices to show that the set $\big\{(x,y)\in\mathbb{R}^2\,\big|\,\text{osc}\big(\bar{f},(x,y)\big)>0\big\}$ has measure zero, where $\bar{f}$ is the extension of f by zero outside A. Nevertheless, we note that

$$\begin{aligned}\big\{(x,y)\in\mathbb{R}^2\,\big|\,\text{osc}\big(\bar{f},(x,y)\big)>0\big\}\subseteq\{a\}\times\big[\varphi_1(a),\varphi_2(a)\big]\cup\{b\}\times\big[\varphi_1(b),\varphi_2(b)\big]\cup&\\ \cup\big\{\big(x,\varphi_1(x)\big)\,\big|\,x\in[a,b]\big\}\cup\big\{\big(x,\varphi_2(x)\big)\,\big|\,x\in[a,b]\big\}\,.&\end{aligned}$$

It is clear that $\{a\}\times\big[\varphi_1(a),\varphi_2(a)\big]$ and $\{b\}\times\big[\varphi_1(b),\varphi_2(b)\big]$ have measure zero since they have volume zero. Now we claim that the sets $\big\{\big(x,\varphi_1(x)\big)\,\big|\,x\in[a,b]\big\}$ and $\big\{\big(x,\varphi_2(x)\big)\,\big|\,x\in[a,b]\big\}$ also have measure zero.

Let $\varepsilon>0$ be given. Since φ_1 is continuous on a compact set $[a,b]$, φ_1 is uniformly continuous on $[a,b]$; thus there exists $\delta>0$ such that

$$\big|\varphi_1(x_1)-\varphi_1(x_2)\big|<\frac{\varepsilon}{b-a}\quad\text{whenever}\quad |x_1-x_2|<\delta\,.$$

Let $\mathcal{P}=\{a=x_0<x_1<\cdots<x_{n-1}<x_n=b\}$ be a partition of $[a,b]$ such that $|x_{i+1}-x_i|<\delta$ for all $i=0,\cdots,n-1$, and let $\Delta_i=\big[x_i,x_{i+1}\big]\times\big[\min\limits_{x\in[x_i,x_{i+1}]}\varphi_1(x),\max\limits_{x\in[x_i,x_{i+1}]}\varphi_1(x)\big]$. Then

$$\big\{\big(x,\varphi_1(x)\big)\,\big|\,x\in[a,b]\big\}\subseteq\bigcup_{i=0}^{n-1}\Delta_i$$

and

$$\sum_{i=0}^{n-1}\nu(\Delta_i)<\sum_{i=0}^{n-1}\frac{\varepsilon}{b-a}(x_{i+1}-x_i)=\frac{\varepsilon}{b-a}\sum_{i=0}^{n-1}(x_{i+1}-x_i)=\varepsilon\,.$$

Therefore, $\big\{\big(x,\varphi_1(x)\big)\,\big|\,x\in[a,b]\big\}$ has volume zero; thus $\big\{\big(x,\varphi_1(x)\big)\,\big|\,x\in[a,b]\big\}$ has measure zero. Similarly, $\big\{\big(x,\varphi_2(x)\big)\,\big|\,x\in[a,b]\big\}$ also has measure zero. By Theorem 8.24, $\big\{(x,y)\in\mathbb{R}^2\,\big|\,\operatorname{osc}\big(\bar{f},(x,y)\big)>0\big\}$ has measure zero; thus f is Riemann integrable over A.

Let $m=\min\limits_{x\in[a,b]}\varphi_1(x)$, $M=\max\limits_{x\in[a,b]}\varphi_2(x)$, and $S=[a,b]\times[m,M]$. Then $A\subseteq S$. By Lemma 8.49 and the Fubini Theorem,

$$\int_A f(x,y)\,d\mathbb{A}=\int_S\bar{f}(x,y)\,d\mathbb{A}=\int_a^b\Big(\int_m^M\bar{f}(x,y)\,dy\Big)dx=\int_a^b\Big(\int_{\varphi_1(x)}^{\varphi_2(x)}f(x,y)\,dy\Big)dx$$

which concludes 1. □

Example 8.59. Let $A=\big\{(x,y)\in\mathbb{R}^2\,\big|\,0\leqslant x\leqslant 1, x\leqslant y\leqslant 1\big\}$, and $f:A\to\mathbb{R}$ be given by $f(x,y)=xy$. Then Corollary 8.58 implies that

$$\int_A f(x,y)\,d\mathbb{A}=\int_0^1\Big(\int_x^1 xy\,dy\Big)dx=\int_0^1\frac{xy^2}{2}\Big|_{y=x}^{y=1}dx=\int_0^1\Big(\frac{x}{2}-\frac{x^3}{2}\Big)dx=\frac{1}{4}-\frac{1}{8}=\frac{1}{8}\,.$$

On the other hand, since $A=\big\{(x,y)\in\mathbb{R}^2\,\big|\,0\leqslant y\leqslant 1, 0\leqslant x\leqslant y\big\}$, we can also evaluate the integral of f over A by

$$\int_A xy\,d\mathbb{A}=\int_0^1\Big(\int_0^y xy\,dx\Big)dy=\int_0^1\frac{x^2y}{2}\Big|_{x=0}^{x=y}dy=\int_0^1\frac{y^3}{2}dy=\frac{1}{8}\,.$$

Example 8.60. Let $A = \{(x,y) \in \mathbb{R}^2 \,|\, 0 \leqslant x \leqslant 1, \sqrt{x} \leqslant y \leqslant 1\}$, and $f : A \to \mathbb{R}$ be given by $f(x,y) = e^{y^3}$. Then Corollary 8.58 implies that

$$\int_A f(x,y)\, d\mathbb{A} = \int_0^1 \Big(\int_{\sqrt{x}}^1 e^{y^3} dy\Big) dx\,.$$

Since we do not know how to compute the inner integral, we look for another way of finding the integral. Observing that $A = \{(x,y) \in \mathbb{R}^2 \,|\, 0 \leqslant y \leqslant 1, 0 \leqslant x \leqslant y^2\}$, we have

$$\int_A f(x,y)\, d\mathbb{A} = \int_0^1 \Big(\int_0^{y^2} e^{y^3}\, dx\Big) dy = \int_0^1 y^2 e^{y^3} dy = \frac{1}{3} e^{y^3}\Big|_{y=0}^{y=1} = \frac{e-1}{3}\,.$$

Now we prove the general Fubini Theorem.

Theorem 8.61 (Fubini's Theorem)**.** *Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be rectangles, and $f : A \times B \to \mathbb{R}$ be bounded. For $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, write $z = (x,y)$. Then*

$$\underline{\int}_{A\times B} f(z)\, dz \leqslant \underline{\int}_A \Big(\underline{\int}_B f(x,y) dy\Big) dx \leqslant \overline{\int}_A \Big(\overline{\int}_B f(x,y) dy\Big) dx \leqslant \overline{\int}_{A\times B} f(z)\, dz \qquad (8.5.4)$$

and

$$\underline{\int}_{A\times B} f(z)\, dz \leqslant \underline{\int}_B \Big(\underline{\int}_A f(x,y) dx\Big) dy \leqslant \overline{\int}_B \Big(\overline{\int}_A f(x,y) dx\Big) dy \leqslant \overline{\int}_{A\times B} f(z)\, dz\,. \qquad (8.5.5)$$

In particular, if $f : A \times B \to \mathbb{R}$ is Riemann integrable, then

$$\begin{aligned}\int_{A\times B} f(z)\, dz &= \int_A \Big(\underline{\int}_B f(x,y) dy\Big) dx = \int_A \Big(\overline{\int}_B f(x,y) dy\Big) dx \\ &= \int_B \Big(\underline{\int}_A f(x,y) dx\Big) dy = \int_B \Big(\overline{\int}_A f(x,y) dx\Big) dy\,.\end{aligned}$$

Proof. It suffices to prove (8.5.4). Let $\varepsilon > 0$ be given. Choose a partition $\mathcal{P}$ of $A \times B$ such that $L(f,\mathcal{P}) > \underline{\int}_{A\times B} f(z)\, dz - \varepsilon$. Since $\mathcal{P}$ is a partition of $A \times B$, there exist partition $\mathcal{P}_x$ of A and partition $\mathcal{P}_y$ of B such that $\mathcal{P} = \{\Delta = R \times S \,|\, R \in \mathcal{P}_x, S \in \mathcal{P}_y\}$. By Proposition

8.40 and Corollary 8.42, we find that

$$\begin{aligned}
\underline{\int}_A \Big(\underline{\int}_B f(x,y)\,dy\Big)dx &= \underline{\int}_{\bigcup_{R\in\mathcal{P}_x} R} \mathbf{1}_A(x)\Big(\underline{\int}_{\bigcup_{S\in\mathcal{P}_y} S} f(x,y)\mathbf{1}_B(y)\,dy\Big)dx \\
&\geqslant \sum_{R\in\mathcal{P}_x} \underline{\int}_R \Big(\sum_{S\in\mathcal{P}_y} \underline{\int}_S \overline{f}^{A\times B}(x,y)\,dy\Big)dx \\
&\geqslant \sum_{R\in\mathcal{P}_x} \sum_{S\in\mathcal{P}_y} \underline{\int}_R \Big(\underline{\int}_S \overline{f}^{A\times B}(x,y)\,dy\Big)dx \\
&\geqslant \sum_{R\in\mathcal{P}_x, S\in\mathcal{P}_y} \inf_{(x,y)\in R\times S} \overline{f}^{A\times B}(x,y)\nu_m(S)\nu_n(R) \\
&= \sum_{\Delta\in\mathcal{P}} \inf_{(x,y)\in\Delta} \overline{f}^{A\times B}(x,y)\nu_{n+m}(\Delta) = L(f,\mathcal{P}) > \underline{\int}_{A\times B} f(z)dz - \varepsilon\,.
\end{aligned}$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that

$$\underline{\int}_{A\times B} f(z)\,dz \leqslant \underline{\int}_B \Big(\underline{\int}_A f(x,y)dx\Big)dy\,.$$

Similarly, $\overline{\int}_A \Big(\overline{\int}_B f(x,y)dy\Big)dx \leqslant \overline{\int}_{A\times B} f(z)\,dz$; thus (8.5.4) is concluded. □

Corollary 8.62. *Let $S \subseteq \mathbb{R}^n$ be a bounded set with volume, $\varphi_1, \varphi_2 : S \to \mathbb{R}$ be continuous maps such that $\varphi_1(x) \leqslant \varphi_2(x)$ for all $x \in S$, $A = \big\{(x,y) \in \mathbb{R}^n \times \mathbb{R} \,\big|\, x \in S, \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\big\}$, and $f : A \to \mathbb{R}$ be continuous. Then f is Riemann integrable over A, and*

$$\int_A f(x,y)\,d(x,y) = \int_S \Big(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy\Big)dx\,. \tag{8.5.6}$$

Proof. Since ∂A has measure zero, and f is continuous on A, Corollary 8.33 implies that f is Riemann integrable over A. Let $m = \min\limits_{x\in S} \varphi_1(x)$ and $M = \max\limits_{x\in S} \varphi_2(x)$. Then $A \subseteq S \times [m, M]$; thus Theorem 8.50 and the Fubini Theorem imply that

$$\begin{aligned}
\int_A f(x,y)\,d(x,y) = \int_{S\times[m,M]} \overline{f}^{A}(x,y)\,d(x,y) &= \underline{\int}_S \Big(\underline{\int}_m^M \overline{f}^{A}(x,y)\,dy\Big)dx \\
&= \overline{\int}_S \Big(\overline{\int}_m^M \overline{f}^{A}(x,y)\,dy\Big)dx\,.
\end{aligned}$$

Noting that $[m, M]$ has a boundary of volume zero in $\mathbb{R}$, and for each $x \in S$, $\overline{f}^{A}(x,\cdot)$ is continuous except perhaps at $y = \varphi_1(x)$ and $y = \varphi_2(x)$, Corollary 8.33 implies that $\overline{f}^{A}(x,\cdot)$ is

Riemann integrable over $[m, M]$ for each $x \in S$. Therefore, $\underline{\int_m^M} \overline{f}^A(x,y)\,dy = \overline{\int_m^M} \overline{f}^A(x,y)\,dy$ which further implies that

$$\int_A f(x,y)\,d(x,y) = \int_S \Big(\int_m^M \overline{f}^A(x,y)\,dy \Big) dx\,. \tag{8.5.7}$$

For each fixed $x \in S$, let $A_x = \{ y \in \mathbb{R} \,|\, \varphi_1(x) \leqslant y \leqslant \varphi_2(x) \}$. Then $\overline{f}^A(x,y) = f(x,y)\mathbf{1}_{A_x}(y)$ for all $(x,y) \in S \times [m, M]$ or equivalently, $\overline{f}^A(x,\cdot) = f(x,\cdot)|_{A_x}$ for all $x \in S$; thus Proposition 8.40 (a) implies that

$$\int_m^M \overline{f}^A(x,y)\,dy = \int_{A_x} f(x,y)\,dy = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy \qquad \forall\, x \in S\,. \tag{8.5.8}$$

Combining (8.5.7) and (8.5.8), we conclude (8.5.6). ▫

Example 8.63. Let $A \subseteq \mathbb{R}^3$ be the set $\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0, \text{and } x_1 + x_2 + x_3 \leqslant 1 \}$, and $f : A \to \mathbb{R}$ be given by $f(x_1, x_2, x_3) = (x_1 + x_2 + x_3)^2$. Let $S = [0,1] \times [0,1] \times [0,1]$, and $\bar{f} : \mathbb{R}^3 \to \mathbb{R}$ be the extension of f by zero outside A. Then Corollary 8.33 implies that f is Riemann integrable (since ∂A has measure zero). Write $\widehat{x}_1 = (x_2, x_3)$, $\widehat{x}_2 = (x_1, x_3)$ and $\widehat{x}_3 = (x_1, x_2)$. Lemma 8.49 implies that

$$\int_A f(x)dx = \int_S \bar{f}(x)dx\,,$$

and Theorem 8.61 implies that

$$\int_S \bar{f}(x)dx = \int_{[0,1]} \Big(\int_{[0,1]\times[0,1]} \bar{f}(\widehat{x}_3, x_3)d\widehat{x}_3 \Big) dx_3\,.$$

Let $A_{x_3} = \{ (x_1, x_2) \in \mathbb{R}^2 \,|\, x_1 \geqslant 0, x_2 \geqslant 0, x_1 + x_2 \leqslant 1 - x_3 \}$. Then for each $x_3 \in [0,1]$,

$$\int_{[0,1]\times[0,1]} \bar{f}(\widehat{x}_3, x_3)d\widehat{x}_3 = \int_{A_{x_3}} f(\widehat{x}_3, x_3)d\widehat{x}_3 = \int_0^{1-x_3} \Big(\int_0^{1-x_3-x_2} f(x_1,x_2,x_3)dx_1 \Big) dx_2\,.$$

Computing the iterated integral, we find that

$$\begin{aligned} \int_A f(x)dx &= \int_0^1 \Big[\int_0^{1-x_3} \Big(\int_0^{1-x_3-x_2} (x_1+x_2+x_3)^2 dx_1 \Big) dx_2 \Big] dx_3 \\ &= \int_0^1 \Big[\int_0^{1-x_3} \frac{(x_1+x_2+x_3)^3}{3} \Big|_{x_1=0}^{x_1=1-x_3-x_2} dx_2 \Big] dx_3 \\ &= \int_0^1 \Big[\int_0^{1-x_3} \Big(\frac{1}{3} - \frac{(x_2+x_3)^3}{3} \Big) dx_2 \Big] dx_3 \\ &= \int_0^1 \Big(\frac{1}{4} - \frac{x_3}{3} + \frac{x_3^4}{12} \Big) dx_3 = \frac{1}{4} - \frac{1}{6} + \frac{1}{60} = \frac{15-10+1}{60} = \frac{1}{10}\,. \end{aligned}$$

Example 8.64. In this example we compute the volume of the n-dimensional unit ball ω_n. By the Fubini theorem,

$$\omega_n = \int_{-1}^{1} \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{n-1}^2}} dx_n \cdots dx_1 \,.$$

Note that the integral $\int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{n-1}^2}} dx_n \cdots dx_2$ is in fact $\omega_{n-1}(1-x_1^2)^{\frac{n-1}{2}}$; thus

$$\omega_n = \int_{-1}^{1} \omega_{n-1}(1-x^2)^{\frac{n-1}{2}} dx = 2\,\omega_{n-1} \int_0^{\frac{\pi}{2}} \cos^n \theta d\theta \,. \tag{8.5.9}$$

Integrating by parts,

$$\begin{aligned}\int_0^{\frac{\pi}{2}} \cos^n \theta \, d\theta &= \int_0^{\frac{\pi}{2}} \cos^{n-1} \theta \, d(\sin\theta) = \cos^{n-1}\theta \sin\theta \Big|_{\theta=0}^{\theta=\frac{\pi}{2}} + (n-1) \int_0^{\frac{\pi}{2}} \cos^{n-2}\theta \sin^2\theta \, d\theta \\ &= (n-1) \int_0^{\frac{\pi}{2}} \cos^{n-2}\theta (1-\cos^2\theta) \, d\theta\end{aligned}$$

which implies that

$$\int_0^{\frac{\pi}{2}} \cos^n \theta \, d\theta = \frac{n-1}{n} \int_0^{\frac{\pi}{2}} \cos^{n-2}\theta \, d\theta \,.$$

As a consequence,

$$\int_0^{\frac{\pi}{2}} \cos^n \theta \, d\theta = \begin{cases} \dfrac{(n-1)(n-3)\cdots 2}{n(n-2)\cdots 3} \displaystyle\int_0^{\frac{\pi}{2}} \cos\theta \, d\theta & \text{if } n \text{ is odd}\,, \\ \dfrac{(n-1)(n-3)\cdots 1}{n(n-2)\cdots 2} \displaystyle\int_0^{\frac{\pi}{2}} d\theta & \text{if } n \text{ is even}\,; \end{cases}$$

thus the recursive formula (8.5.9) implies that $\omega_n = \dfrac{2\,\omega_{n-2}}{n}\pi$. Further computations shows that

$$\omega_n = \begin{cases} \dfrac{(2\pi)^{\frac{n-1}{2}}}{n(n-2)\cdots 3}\omega_1 & \text{if } n \text{ is odd}\,, \\ \dfrac{(2\pi)^{\frac{n-2}{2}}}{n(n-2)\cdots 4}\omega_2 & \text{if } n \text{ is even}\,. \end{cases}$$

Let Γ be the Gamma function defined by $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\,dx$ for $t > 0$. Then $\Gamma(x+1) = x\Gamma(x)$ for all $x > 0$, $\Gamma(1) = 1$ and $\Gamma\big(\frac{1}{2}\big) = \sqrt{\pi}$. By the fact that $\omega_1 = 2$ and $\omega_2 = \pi$, we can express ω_n as

$$\omega_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\big(\frac{n+2}{2}\big)} \,.$$

8.6 Change of Variables Formula

Fubini theorem can be used to find the integral of a (Riemann integrable) function over a rectangular domain if the iterated integrals can be evaluated. However, like the integral of a function of one variable, in many cases we need to make use of several change of variables in order to transform the integral to another integral that can be easily evaluated. In this section, we establish the change of variables formula for the integral of functions of several variables.

Theorem 8.65 (Change of Variables Formula)**.** *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open bounded set, and $g : \mathcal{U} \to \mathbb{R}^n$ be an one-to-one $\mathscr{C}^1$ mapping with $\mathscr{C}^1$ inverse; that is, $g^{-1} : g(\mathcal{U}) \to \mathcal{U}$ is also continuously differentiable. Assume that the Jacobian of g, $\mathrm{J}_g = \det([Dg])$, does not vanish in $\mathcal{U}$, and $E \subset\subset \mathcal{U}$ has volume. Then $g(E)$ has volume. Moreover, if $f : g(E) \to \mathbb{R}$ is bounded and integrable, then $(f \circ g)\mathrm{J}_g$ is integrable over E, and*

$$\int_{g(E)} f(y)\,dy = \int_E (f \circ g)(x)\big|\mathrm{J}_g(x)\big|\,dx = \int_E (f \circ g)(x)\Big|\frac{\partial(g_1, \cdots, g_n)}{\partial(x_1, \cdots, x_n)}\Big|\,dx\,.$$

Remark 8.66. The condition that g has to be defined on a larger open set $\mathcal{U}$ can be removed. In other words, $E \subseteq \mathcal{U}$ has volume is enough for the change of variable formula to hold; however, we will not prove this more generalized version here.

The proof of the change of variables formula is separated into several steps, and we list each step as a lemma.

First, we show that the map g in Theorem 8.65 has the property that $g^{-1}(Z)$ has measure zero (or volume zero) if Z itself has measure zero (or volume zero). This establishes that if A and B are not overlapping; that is, $\nu(A \cap B) = 0$, then $\nu(g^{-1}(A \cap B)) = 0$.

Lemma 8.67. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open set, and $\phi : \mathcal{U} \to \mathbb{R}^n$ be Lipschitz continuous; that is, there exists $L > 0$ such that $\|\phi(x) - \phi(y)\|_{\mathbb{R}^n} \leqslant L\|x - y\|_{\mathbb{R}^n}$ for all $x, y \in \mathcal{U}$. Suppose that $Z \subseteq \mathcal{U}$ is a set of measure zero (or a set of volume zero) and $\bar{Z} \subseteq \mathcal{U}$. Then $\phi(A)$ has measure zero (or volume zero).*

Proof. We prove the case that Z has measure zero, and the proof for the case that Z has volume zero is obtained by changing the countable sum/union to finite sum/union.

First we note that if $S \subseteq \mathcal{U}$ is a rectangle on which the ratio of the maximum length and minimum length of sides is less than 2, then $\phi(S) \subseteq R$ for some n-cube R with side of length

$L\sqrt{n}\delta$, where δ is the maximum length of sides of S. Therefore, $\nu(\phi(S)) \leqslant (2\sqrt{n}L)^n\nu(S)$. Let $\varepsilon > 0$ be given. Since Z has measure zero, there exists countable rectangles $S_1, S_2, \cdots$ such that $Z \subseteq \bigcup_{k=1}^{\infty} S_k$ and $\sum_{k=1}^{\infty} \nu(S_k) < \frac{\varepsilon}{(2\sqrt{n}L)^n}$. Moreover, as in the proof of Proposition 8.14 we can also assume that the ratio of the maximum length and minimum length of sides of S_k is less than 2 for all $k \in \mathbb{N}$; thus $\phi(Z) \subseteq \bigcup_{k=1}^{\infty} R_k$ and $\sum_{k=1}^{\infty} \nu(R_k) < \varepsilon$ for some rectangles R_k's. □

Next, we prove that it suffices to show the change of variables formula for the case that f is a constant and E is the pre-image of closed rectangle under g in order to establish the full result.

Lemma 8.68. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open bounded set, and $g : \mathcal{U} \to \mathbb{R}^n$ be an one-to-one $\mathscr{C}^1$ mapping that has a $\mathscr{C}^1$ inverse. Assume that the Jacobian of g, $\mathrm{J}_g = \det([Dg])$, does not vanish in $\mathcal{U}$, and*

$$\nu(R) = \int_{g^{-1}(R)} |\mathrm{J}_g(x)| dx \quad \textit{for all closed rectangle } R \subseteq g(\mathcal{U})\,. \tag{8.6.1}$$

Then if $E \subset\subset \mathcal{U}$ has volume and $f : g(E) \to \mathbb{R}$ is bounded and integrable, then $(f \circ g)|\mathrm{J}_g|$ is Riemann integrable over E, and

$$\int_{g(E)} f(y)\, dy = \int_E (f \circ g)(x)|\mathrm{J}_g(x)| dx\,.$$

Proof. Consider the extensions of f and $(f \circ g)|\mathrm{J}_g|$ given by

$$\overline{f}^{g(E)}(x) = \begin{cases} f(x) & \text{if } x \in g(E)\,, \\ 0 & \text{if } x \in g(E)^{\complement}, \end{cases} \quad \text{and} \quad \overline{(f \circ g)|\mathrm{J}_g|}^{E}(x) = \begin{cases} (f \circ g)(x)|\mathrm{J}_g|(x) & \text{if } x \in E\,, \\ 0 & \text{if } x \in E^{\complement}. \end{cases}$$

By the integrability of f over $g(E)$, the set $\big\{y \in \mathbb{R}^n \,\big|\, \overline{f}^{g(E)} \text{ is discontinuous at } y\big\}$ has measure zero. Since

$$\begin{aligned} \big\{x \in \mathbb{R}^n \,\big|\, \overline{(f \circ g)|\mathrm{J}_g|}^{E} \text{ is discontinuous at } x\big\} & \\ \subseteq \partial E \cup \big\{x \in \operatorname{int}(E) \,\big|\, f \text{ is discontinuous at } g(x)\big\} & \\ = \partial E \cup \big\{y \in g(\operatorname{int}(E)) \,\big|\, f \text{ is discontinuous at } y\big\} & \\ \subseteq \partial E \cup \big\{y \in \mathbb{R}^n \,\big|\, \overline{f}^{g(E)} \text{ is discontinuous at } y\big\}\,, & \end{aligned}$$

we conclude that $\left\{x \in \mathbb{R}^n \,\middle|\, \overline{(f \circ g)|\mathrm{J}_g|}^{E} \text{ is discontinuous at } x\right\}$ has measure zero. Therefore, $(f \circ g)|\mathrm{J}_g|$ is Riemann integrable over E. On the other hand, by the fact that

$$(\overline{f}^{g(E)} \circ g)|\mathrm{J}_g| = (\overline{f \circ g})^{E}|\mathrm{J}_g| = \overline{(f \circ g)|\mathrm{J}_g|}^{E} \quad \text{on} \quad \mathcal{U}\,,$$

the Lebesgue theorem also implies that $(\overline{f}^{g(E)} \circ g)|\mathrm{J}_g|$ is Riemann integrable over F if $E \subseteq F \subseteq \mathcal{U}$ since

$$\overline{\overline{(f \circ g)|\mathrm{J}_g|}^{E}}^{F} = \overline{(f \circ g)|\mathrm{J}_g|}^{E} \qquad \forall\, F \supseteq E\,.$$

Moreover, it follows from Lemma 8.49 that

$$\int_F (\overline{f}^{g(E)} \circ g)(x)|\mathrm{J}_g(x)|dx = \int_E (f \circ g)(x)|\mathrm{J}_g(x)|dx \qquad \forall\, E \subseteq F \subseteq \mathcal{U}\,. \tag{8.6.2}$$

Since the Jacobian of g does not vanish in $\mathcal{U}$, Remark 7.2 implies that g is an open mapping; thus $g(\mathcal{U})$ is open. By the fact that $g(\bar{E})$ is compact, there exists an open set $\mathcal{V}$ in $\mathbb{R}^n$ such that $g(\bar{E}) \subseteq \mathcal{V} \subset\subset g(\mathcal{U})$. It then follows from $g^{-1} \in \mathscr{C}^1(g(\mathcal{U}))$ and $\bar{\mathcal{V}} \subseteq \mathcal{U}$ that there exists $L > 0$ such that $\|g^{-1}(y_1) - g^{-1}(y_2)\|_{\mathbb{R}^n} \leqslant L\|y_1 - y_2\|_{\mathbb{R}^n}$ for all $y_1, y_2 \in \mathcal{V}$. In other words, g^{-1} is Lipschitz on $\mathcal{V}$, and Lemma 8.67 implies that $g^{-1}(Z)$ has volume zero if $Z \subseteq \mathcal{V}$ has volume zero.

Note that there exists $\delta > 0$ such that $d(x, y) > \delta$ for all $x \in g(E)$ and $y \in \mathcal{V}^{\complement}$. Let $\mathcal{P}$ be a partition of $g(E)$ such that $\|\mathcal{P}\| < \delta$; that is, $\operatorname{diam}(\Delta) < \delta$ for all $\Delta \in \mathcal{P}$. Then $\Delta \subseteq \mathcal{V}$ if $\Delta \in \mathcal{P}$ and $\Delta \cap g(E) \neq \varnothing$. Since $\inf\limits_{y \in \Delta} \overline{f}^{g(E)}(y) = \inf\limits_{x \in g^{-1}(\Delta)} (\overline{f}^{g(E)} \circ g)(x)$ if $\Delta \subseteq \mathcal{U}$, using (8.6.1) we find that

$$\begin{aligned} L(f, \mathcal{P}) &= \sum_{\substack{\Delta \in \mathcal{P} \\ \Delta \cap g(E) \neq \varnothing}} \inf_{y \in \Delta} \overline{f}^{g(E)}(y)\nu(\Delta) = \sum_{\substack{\Delta \in \mathcal{P} \\ \Delta \cap g(E) \neq \varnothing}} \inf_{x \in g^{-1}(\Delta)} (\overline{f}^{g(E)} \circ g)(x)\nu(\Delta) \\ &= \sum_{\substack{\Delta \in \mathcal{P} \\ \Delta \cap g(E) \neq \varnothing}} \inf_{x \in g^{-1}(\Delta)} (\overline{f}^{g(E)} \circ g)(x) \int_{g^{-1}(\Delta)} |\mathrm{J}_g(x)|dx \\ &\leqslant \sum_{\substack{\Delta \in \mathcal{P} \\ \Delta \cap g(E) \neq \varnothing}} \underline{\int}_{g^{-1}(\Delta)} (\overline{f}^{g(E)} \circ g)(x)|\mathrm{J}_g(x)|dx\,. \end{aligned}$$

Since each "face" of the rectangle $\Delta \in \mathcal{P}$ has volume zero, Lemma 8.67 implies that $g^{-1}(\Delta) \cap$

$g^{-1}(\Delta')$ has volume zero if $\Delta \cap \Delta'$ has volume zero. Therefore, Corollary 8.42 shows that

$$
\begin{aligned}
L(f,\mathcal{P}) &\leqslant \underline{\int}_{\bigcup_{\Delta\in\mathcal{P},\Delta\cap g(E)\neq\varnothing} g^{-1}(\Delta)} (\overline{f}^{g(E)}\circ g)(x)|\mathrm{J}_g(x)|dx \\
&= \underline{\int}_{g^{-1}(\bigcup_{\Delta\in\mathcal{P},\Delta\cap g(E)\neq\varnothing} \Delta)} (\overline{f}^{g(E)}\circ g)(x)|\mathrm{J}_g(x)|dx \\
&= \int_{g^{-1}(\bigcup_{\Delta\in\mathcal{P},\Delta\cap g(E)\neq\varnothing} \Delta)} (\overline{f}^{g(E)}\circ g)(x)|\mathrm{J}_g(x)|dx \\
&= \int_E (f\circ g)(x)|\mathrm{J}_g(x)|\,dx\,,
\end{aligned}
$$

where we have used the integrability of $(\overline{f}^{g(E)}\circ g)|\mathrm{J}_g|$ over the set $g^{-1}(\bigcup_{\Delta\in\mathcal{P},\Delta\cap g(E)\neq\varnothing}\Delta)$ (since this set is a super set of E) and (8.6.2) to conclude the last two equalities.

Similarly, by the fact that $\sup\limits_{y\in\Delta}\overline{f}^{g(E)}(y) = \sup\limits_{x\in g^{-1}(\Delta)}(\overline{f}^{g(E)}\circ g)(x)$ if $\Delta\subseteq\mathcal{U}$, we obtain that

$$
\begin{aligned}
U(f,\mathcal{P}) &= \sum_{\substack{\Delta\in\mathcal{P}\\ \Delta\cap g(E)\neq\varnothing}} \sup_{x\in g^{-1}(\Delta)}(\overline{f}^{g(E)}\circ g)(x)\int_{g^{-1}(\Delta)}|\mathrm{J}_g(x)|dx \\
&\geqslant \sum_{\substack{\Delta\in\mathcal{P}\\ \Delta\cap g(E)\neq\varnothing}} \overline{\int}_{g^{-1}(\Delta)}(\overline{f}^{g(E)}\circ g)(x)|\mathrm{J}_g(x)|dx = \int_{g^{-1}(\bigcup_{\Delta\in\mathcal{P},\Delta\cap g(E)\neq\varnothing}\Delta)}(\overline{f}^{g(E)}\circ g)(x)|\mathrm{J}_g(x)|dx \\
&= \int_E (f\circ g)(x)|\mathrm{J}_g(x)|\,dx\,.
\end{aligned}
$$

The integrability of f over $g(E)$ then implies that $\displaystyle\int_{g(E)} f(y)\,dy = \int_E (f\circ g)(x)|\mathrm{J}_g(x)|dx$. □

Since the differentiability of g implies that locally g is very closed to an affine map; that is, $g(x)\approx \mathrm{L}x+c$ for some $\mathrm{L}\in\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)$ and $c\in\mathbb{R}^n$ (in fact, $g(x)\approx g(x_0)+(Dg)(x_0)(x-x_0)$ in a neighborhood of x_0), our next step is to establish (8.6.1) first for the case that g is an affine map. Since the volume of a set remains unchanged under translation, W.L.O.G. we can assume that g is linear.

Lemma 8.69. *Let $g\in\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)$, and $A\subseteq\mathbb{R}^n$ be a set that has volume. Then $g(A)$ has volume, and*

$$\nu\big(g(A)\big) = \int_{g(A)} 1dy = \int_A |\mathrm{J}_g(x)|dx\,. \tag{8.6.3}$$

Remark 8.70. If $g\in\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)$, then $g(x)=\mathrm{L}x$ for some $n\times n$ matrix. In this case $\mathrm{J}_g(x)=\det(\mathrm{L})$ for all $x\in A$; thus (8.6.3) is the same as that

$$\nu\big(\mathrm{L}(A)\big) = \int_{\mathrm{L}(A)} 1dy = \int_A |\det(\mathrm{L})|dx = |\det(\mathrm{L})|\nu(A)\,. \tag{8.6.4}$$

Therefore, in the following we prove (8.6.4) instead of (8.6.3).

Proof of Lemma 8.69. Since any $n \times n$ matrices can be expressed as the product of elementary matrices, it suffices to prove the validity of the lemma for the case that L is an elementary matrix.

Suppose first that $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$ is a rectangle.

1. If L is an elementary matrix of the type

$$
\mathrm{L} = \left[\begin{array}{ccccccccccc}
1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & \ddots & 0 & & & & & & & & \vdots \\
\vdots & \ddots & 1 & \ddots & & & & & & & \vdots \\
\vdots & & 0 & 0 & 0 & & & 1 & & & \vdots \\
\vdots & & & \ddots & 1 & \ddots & & & & & \vdots \\
\vdots & & & & 0 & \ddots & 0 & & & & \vdots \\
\vdots & & & & & \ddots & 1 & \ddots & & & \vdots \\
\vdots & & & 1 & & & 0 & 0 & 0 & & \vdots \\
0 & & & & & & & \ddots & 1 & \ddots & \vdots \\
0 & & & & & & & & 0 & \ddots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \\ \leftarrow \text{the } i_0\text{-th row} \\ \\ \\ \\ \leftarrow \text{the } j_0\text{-th row} \\ \\ \\ \\
\end{array}
$$

$\uparrow$ the i_0-th column $\qquad$ $\uparrow$ the j_0-th column

then

$$
\begin{aligned}
\mathrm{L}(A) = [a_1, b_1] \times \cdots \times [a_{i_0-1}, b_{i_0-1}] \times [a_{j_0}, b_{j_0}] \times [a_{i_0+1}, b_{i_0+1}] \times \cdots \times \\
\times [a_{j_0-1}, b_{j_0-1}] \times [a_{i_0}, b_{i_0}] \times [a_{j_0+1}, b_{j_0} + 1] \times \cdots \times [a_n, b_n]\,;
\end{aligned}
$$

thus $\nu\big(\mathrm{L}(A)\big) = \nu(A) = |\det(\mathrm{L})|\nu(A)$.

2. If L is an elementary matrix of the type

$$
\mathrm{L} = \left[\begin{array}{ccccccccc}
1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 1 & 0 & & & & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & & & & & \vdots \\
\vdots & & 0 & 1 & 0 & & & & \vdots \\
\vdots & & & 0 & c & 0 & & & \vdots \\
\vdots & & & & 0 & 1 & 0 & & \vdots \\
\vdots & & & & & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & & & & 0 & 1 & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \\ \\ \leftarrow \text{the } k_0\text{-th row} \\ \\ \\ \\ \\
\end{array}
$$

then

$$\mathrm{L}(A) = [a_1, b_1] \times \cdots \times [a_{k_0-1}, b_{k_0-1}] \times [ca_{k_0}, cb_{k_0}] \times [a_{k_0+1}, b_{k_0+1}] \times \cdots \times [a_n, b_n]$$

if $c \geqslant 0$ or

$$\mathrm{L}(A) = [a_1, b_1] \times \cdots \times [a_{k_0-1}, b_{k_0-1}] \times [cb_{k_0}, ca_{k_0}] \times [a_{k_0+1}, b_{k_0+1}] \times \cdots \times [a_n, b_n]$$

if $c < 0$. In either case, $\nu\big(\mathrm{L}(A)\big) = |c|\nu(A) = |\det(\mathrm{L})|\nu(A)$.

3. If L is an elementary matrix of the type

$$\mathrm{L} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & & & & & & 0 \\ \vdots & \ddots & \ddots & \ddots & & & c & & 0 \\ \vdots & & \ddots & \ddots & \ddots & & & & 0 \\ \vdots & & & 0 & 1 & 0 & & & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & & 0 & 1 & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow \text{the } i_0\text{-th row} \\ \\ \\ \\ \\ \\ \\ \end{matrix}$$

↑ the j_0-th column

then $\mathrm{L}(A)$ is a parallelepiped

$$\begin{aligned} L(A) &= \big\{(x_1, \cdots, x_{i_0-1}, x_{i_0} + cx_{j_0}, x_{i_0+1}, \cdots, x_n) \in \mathbb{R}^n \,\big|\, x_i \in [a_i, b_i] \ \forall\, 1 \leqslant i \leqslant n\big\} \\ &= \big\{(x_1, \cdots, x_{i_0-1}, y_{i_0}, x_{i_0+1}, \cdots, x_n) \in \mathbb{R}^n \,\big|\, a_{i_0} + cx_{j_0} \leqslant y_{i_0} \leqslant b_{i_0} + cx_{j_0}, \\ &\qquad\qquad x_i \in [a_i, b_i] \ \forall\, i \neq i_0\big\}\,; \end{aligned}$$

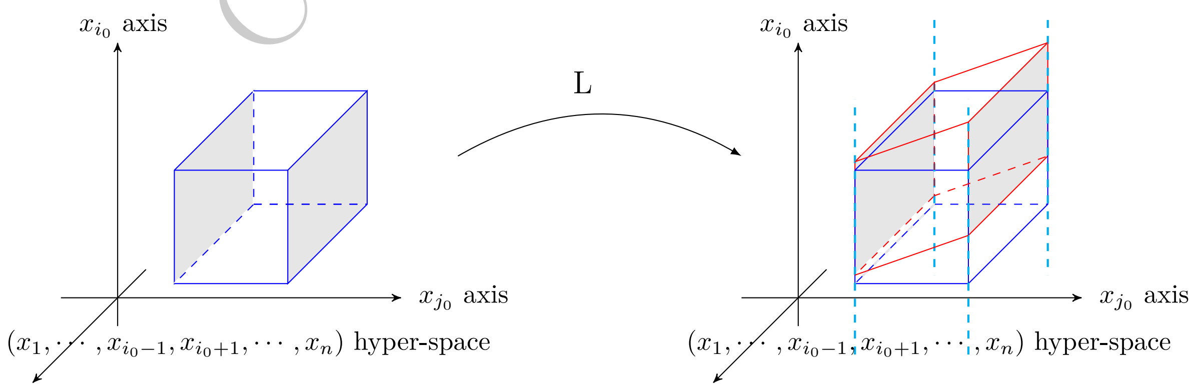

Figure 8.3: The image of a rectangle under a linear map induced by the elementary matrix of the third type

thus the Fubini theorem (or Corollary 8.62) implies that

$$\nu(L(A)) = \int_{[a_1,b_1]\times\cdots\times[a_{i_0-1},b_{i_0-1}]\times[a_{i_0+1},b_{i_0+1}]\times\cdots\times[a_n,b_n]} \Big(\int_{a_{i_0}+cx_{j_0}}^{b_{i_0}+cx_{j_0}} 1 dy_{i_0}\Big) d\widehat{x}_{i_0} = \nu(A)\,.$$

On the other hand, $|\det(\mathrm{L})| = 1$, so $\nu\big(\mathrm{L}(A)\big) = |\det(\mathrm{L})|\nu(A)$ is validated.

Therefore, (8.6.4) holds if A is a rectangle and L is an elementary matrix. An immediate consequence of this observation is that if Z is a set of measure zero, so is $\mathrm{L}(Z)$.

Now suppose that A is an arbitrary set with volume, and L is an elementary matrix.

1. If $\det(\mathrm{L}) = 0$, L must be an elementary matrix of the second type (with $c = 0$), and in this case,

$$\mathrm{L}(A) \subseteq [-r,r]\times\cdots\times[-r,r]\times \underbrace{[-\varepsilon,\varepsilon]}_{\text{the } k_0\text{-th slot}} \times\cdots\times[-r,r]$$

   for some $r > 0$ sufficiently large and arbitrary $\varepsilon > 0$. Therefore, $\mathrm{L}(A)$ has volume zero; thus $\mathrm{L}(A)$ has volume and $\nu\big(\mathrm{L}(A)\big) = |\det(\mathrm{L})|\nu(A)$.

2. Suppose that $\det(\mathrm{L}) \neq 0$. Let $\varepsilon > 0$ be given. Since A has volume, by Riemann's condition there exists a partition of A such that

$$U(1_A,\mathcal{P}) - L(1_A,\mathcal{P}) < \frac{\varepsilon}{|\det(\mathrm{L})|}\,.$$

   Note that the inequality above also implies that

$$U(1_A,\mathcal{P}) - \nu(A) < \frac{\varepsilon}{|\det(\mathrm{L})|} \quad\text{and}\quad \nu(A) - L(1_A,\mathcal{P}) < \frac{\varepsilon}{|\det(\mathrm{L})|}\,.$$

   Let $C_1 = \big\{\Delta\in\mathcal{P}\,\big|\,\Delta\cap A\neq\varnothing\big\}$ and $C_2 = \big\{\Delta\in\mathcal{P}\,\big|\,\Delta\subseteq A\big\}$, and define $\mathrm{R}_1 = \bigcup\limits_{\Delta\in C_1}\Delta$ and $\mathrm{R}_2 = \bigcup\limits_{\Delta\in C_2}\Delta$. Then $\mathrm{R}_2\subseteq A\subseteq\mathrm{R}_1$. Moreover,

$$\begin{aligned}\nu\big(\mathrm{L}(\mathrm{R}_1)\big) &= \sum_{\Delta\in C_1}\nu(\mathrm{L}(\Delta)) = \sum_{\Delta\in C_1}|\det(\mathrm{L})|\nu(\Delta) = |\det(\mathrm{L})|U(1_A,\mathcal{P})\\ &< |\det(\mathrm{L})|\nu(A)+\varepsilon\end{aligned}$$

   and

$$\begin{aligned}\nu\big(\mathrm{L}(\mathrm{R}_2)\big) &= \sum_{\Delta\in C_2}\nu(\mathrm{L}(\Delta)) = \sum_{\Delta\in C_2}|\det(\mathrm{L})|\nu(\Delta) = |\det(\mathrm{L})|L(1_A,\mathcal{P})\\ &> |\det(\mathrm{L})|\nu(A)-\varepsilon\,.\end{aligned}$$

As a consequence, by the fact that $\mathrm{L}(\mathrm{R}_2) \subseteq \mathrm{L}(A) \subseteq \mathrm{L}(\mathrm{R}_1)$ we conclude that

$$\Big| \overline{\int}_{\mathrm{L}(A)} 1 dx - \underline{\int}_{\mathrm{L}(A)} 1 dx \Big| \leqslant \nu\big(\mathrm{L}(\mathrm{R}_1)\big) - \nu\big(\mathrm{L}(\mathrm{R}_2)\big) = |\det(\mathrm{L})| \big(U(1_A, \mathcal{P}) - L(1_A, \mathcal{P})\big) < \varepsilon \,.$$

Since $\varepsilon > 0$ is arbitrary, we find that $\overline{\int}_{\mathrm{L}(A)} 1 dx = \underline{\int}_{\mathrm{L}(A)} 1 dx$ which implies that $1_{\mathrm{L}(A)}$ is Riemann integrable, or equivalently, $\mathrm{L}(A)$ has volume.

On the other hand, observing that

$$|\det(\mathrm{L})|\nu(A) - \varepsilon < \nu\big(\mathrm{L}(\mathrm{R}_2)\big) \leqslant \nu\big(\mathrm{L}(A)\big) \leqslant \nu\big(\mathrm{L}(\mathrm{R}_1)\big) < |\det(\mathrm{L})|\nu(A) + \varepsilon \,,$$

we conclude that $\nu\big(\mathrm{L}(A)\big) = |\det(\mathrm{L})|\nu(A)$ again because $\varepsilon > 0$ is arbitrary. □

Lemma 8.71. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open bounded set, and $g : \mathcal{U} \to \mathbb{R}^n$ be an one-to-one $\mathscr{C}^1$ mapping that has a $\mathscr{C}^1$ inverse. Assume that the Jacobian of g, $\mathrm{J}_g = \det([Dg])$, does not vanish in $\mathcal{U}$. Then*

$$\nu(R) = \int_{g^{-1}(R)} |\mathrm{J}_g(x)| dx \quad \textit{for all closed rectangle } R \subseteq g(\mathcal{U}) \,. \tag{8.6.1}$$

Proof. First, we note that by the the compactness of R, there exist $m > 0$ and $\Lambda > 0$ such that

$$|\mathrm{J}_g(x)| \geqslant m \quad \text{and} \quad \big\|(Dg)(x)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} \leqslant \Lambda \qquad \forall\, x \in g^{-1}(R) \,.$$

Let $0 < \varepsilon < 1$ be given. By the compactness of $g^{-1}(R)$, (Theorem 4.52 implies that) $\mathrm{J}_g : g^{-1}(R) \to \mathbb{R}$ is uniformly continuous; thus there exists $\delta_1 > 0$ such that

$$\big|\mathrm{J}_g(x_1) - \mathrm{J}_g(x_2)\big| < m\varepsilon \quad \text{if} \quad \|x_1 - x_2\|_{\mathbb{R}^n} < \delta_1 \text{ and } x_1, x_2 \in g^{-1}(R) \,.$$

Since g^{-1} is of class $\mathscr{C}^1$, the continuity of g^{-1} and Corollary 6.36 guarantee that there exists $\delta > 0$ such that if $\|y_1 - y_2\|_{\mathbb{R}^n} < \delta$ and $y_1, y_2 \in R$, we have

$$\big\|g^{-1}(y_1) - g^{-1}(y_2)\big\|_{\mathbb{R}^n} < \delta_1$$

and

$$\big\|g^{-1}(y_2) - g^{-1}(y_1) - (Dg^{-1})(y_1)(y_2 - y_1)\big\|_{\mathbb{R}^n} \leqslant \frac{\varepsilon}{2\sqrt{n}\Lambda} \|y_1 - y_2\|_{\mathbb{R}^n} \,.$$

Let $\mathcal{P}$ be a partition of R with mesh size $\|\mathcal{P}\| < \delta$ and the ratio of the maximum length and minimum length of sides of each Δ is less than 2. For $\Delta \in \mathcal{P}$, let c_Δ denote the center of Δ for $\Delta \in \mathcal{P}$, and define $A_\Delta = (Dg)(g^{-1}(c_\Delta))$ as well as $h_\Delta(x) = A_\Delta\big(x - g^{-1}(c_\Delta)\big) + c_\Delta$. Then

$$\big|\mathrm{J}_g(g^{-1}(y)) - \mathrm{J}_g(g^{-1}(c_\Delta))\big| < m\varepsilon \quad \forall\, y \in \Delta\,.$$

Moreover, the inverse function theorem (Theorem 7.1) implies that $A_\Delta^{-1} = (Dg^{-1})(c_\Delta)$; thus for $y \in \Delta$,

$$\begin{aligned}\big\|(h \circ g^{-1})(y) - y\big\|_{\mathbb{R}^n} &= \big\|A_\Delta\big(g^{-1}(y) - g^{-1}(c_\Delta) - (Dg^{-1})(c_\Delta)(y - c_\Delta)\big)\big\|_{\mathbb{R}^n} \\ &\leqslant \|A_\Delta\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\big\|g^{-1}(y) - g^{-1}(c_\Delta) - (Dg^{-1})(c_\Delta)(y - c_\Delta)\big\|_{\mathbb{R}^n} \\ &\leqslant \frac{\varepsilon\|(Dg)(g^{-1}(c_\Delta))\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}}{2\sqrt{n}\Lambda}\|y - c_\Delta\|_{\mathbb{R}^n} \leqslant \frac{\varepsilon}{4\sqrt{n}}\mathrm{diam}(\Delta)\,.\end{aligned}$$

The inequality above implies that for all $\Delta \in \mathcal{P}$,

$$(1-\varepsilon)^n \nu(\Delta) \leqslant \nu\big((h_\Delta \circ g^{-1})(\Delta)\big) \leqslant (1+\varepsilon)^n \nu(\Delta)\,.$$

Since $\mathrm{J}_{h_\Delta} = \det(A_\Delta) = \mathrm{J}_g(g^{-1}(c_\Delta))$, Lemma 8.69 or (8.6.4) provides that

$$\begin{aligned}\int_{g^{-1}(\Delta)} \big|\mathrm{J}_g(x)\big|\,dx &\leqslant \int_{g^{-1}(\Delta)} \big(|\mathrm{J}_g(g^{-1}(c_\Delta))| + m\varepsilon\big)dx = \big(|\mathrm{J}_g(g^{-1}(c_\Delta))| + m\varepsilon\big)\nu\big(g^{-1}(\Delta)\big) \\ &= \big(|\mathrm{J}_g(g^{-1}(c_\Delta))| + m\varepsilon\big)\frac{\nu\big((h_\Delta \circ g)^{-1}(\Delta)\big)}{|\mathrm{J}_g(g^{-1}(c_\Delta))|} \leqslant (1+\varepsilon)^{n+1}\nu(\Delta)\,.\end{aligned}$$

A similar argument provides a lower bounded of the left-hand side, and we conclude that

$$(1-\varepsilon)^{n+1}\nu(\Delta) \leqslant \int_{g^{-1}(\Delta)} |\mathrm{J}_g(x)|dx \leqslant (1+\varepsilon)^{n+1}\nu(\Delta) \qquad \forall\, \Delta \in \mathcal{P}\,.$$

Summing over all $\Delta \in \mathcal{P}$, we find that

$$(1-\varepsilon)^{n+1}\nu(R) \leqslant \sum_{\Delta\in\mathcal{P}} \int_{g^{-1}(\Delta)} |\mathrm{J}_g(x)|dx \leqslant (1+\varepsilon)^{n+1}\nu(R)\,.$$

Identity (8.6.1) is then concluded since $\sum\limits_{\Delta\in\mathcal{P}} \int_{g^{-1}(\Delta)} |\mathrm{J}_g(x)|dx = \int_{g^{-1}(R)} |\mathrm{J}_g(x)|dx$ and $\varepsilon \in (0,1)$ is arbitrary. □

Example 8.72. Let A be the triangular region with vertices $(0,0)$, $(4,0)$, $(4,2)$, and $f : A \to \mathbb{R}$ be given by

$$f(x,y) = y\sqrt{x-2y}\,.$$

Let $(u,v)=(x,x-2y)$. Then $(x,y)=g(u,v)=\big(u,\dfrac{u-v}{2}\big)$; thus

$$\mathrm{J}_g(u,v)=\begin{vmatrix}1 & 0\\ \frac{1}{2} & -\frac{1}{2}\end{vmatrix}=-\frac{1}{2}\,.$$

Define E as the triangle with vertices $(0,0)$, $(4,0)$, $(4,4)$. Then $A=g(E)$.

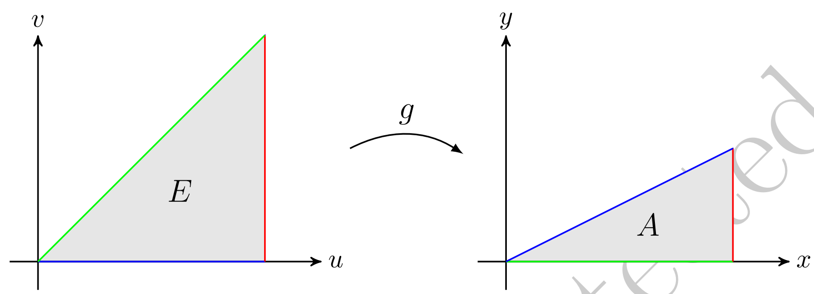

Figure 8.4: The image of E under g

Therefore,

$$\begin{aligned}\int_A f(x,y)d(x,y) &= \int_{g(E)} f(x,y)d(x,y)=\frac{1}{2}\int_E f\big(g(u,v)\big)d(u,v)\\ &=\frac{1}{4}\int_0^4\!\!\int_0^u (u-v)\sqrt{v}dvdu=\frac{1}{4}\int_0^4\Big[\frac{2}{3}uv^{\frac{3}{2}}-\frac{2}{5}v^{\frac{5}{2}}\Big]\Big|_{v=0}^{v=u}du\\ &=\frac{1}{4}\int_0^4\Big(\frac{2}{3}-\frac{2}{5}\Big)u^{\frac{5}{2}}du=\frac{1}{15}\times\frac{2}{7}u^{\frac{7}{2}}\Big|_{u=0}^{u=4}=\frac{256}{105}\,.\end{aligned}$$

Example 8.73. Suppose that $f:[0,1]\to\mathbb{R}$ is Riemann integrable and $\int_0^1(1-x)f(x)\,dx=5$ (note that the function $g(x)=(1-x)f(x)$ is Riemann integrable over $[0,1]$ because of the Lebesgue theorem). We would like to evaluate the iterated integral $\int_0^1\int_0^x f(x-y)\,dydx$.

It is nature to consider the change of variables $(u,v)=(x-y,x)$ or $(u,v)=(x-y,y)$. Suppose the later case. Then $(x,y)=g(u,v)=(u+v,v)$; thus

$$\mathrm{J}_g(u,v)=\begin{vmatrix}1 & 1\\ 1 & 0\end{vmatrix}=-1\,.$$

Moreover, the region of integration is the triangle A with vertices $(0,0)$, $(1,0)$, $(1,1)$, and three sides $y=0$, $x=1$, $x=y$ correspond to $u=0$, $u+v=1$ and $v=0$. Therefore, if

E denotes the triangle enclosed by $u = 0$, $v = 0$ and $u + v = 1$ on the (u, v)-plane, then $g(E) = A$, and

$$\begin{aligned}\int_0^1 \int_0^x f(x-y)\,dydx &= \int_A f(x-y)d(x,y) = \int_{g(E)} f(x-y)d(x,y) \\ &= \int_E f\big(g_1(u,v) - g_2(u,v)\big)|\mathrm{J}_g(u,v)|d(u,v) = \int_0^1 \int_0^{1-u} f(u)\,dvdu \\ &= \int_0^1 (1-u)f(u)\,du = 5\,.\end{aligned}$$

Example 8.74 (Polar coordinates)**.** In $\mathbb{R}^2$, when the domain over which the integral is taken is a disk D, a particular type of change of variables is sometimes very useful for the purpose of evaluating the integral. Let $(x, y) = (x_0 + r\cos\theta, y_0 + r\sin\theta) \equiv \psi(r, \theta)$, where (x_0, y_0) is the center of D under consideration. If the radius of D is R, then D, up to removing a line segment with length R, is the image of $(0, R) \times (0, 2\pi)$ under ψ. Note that the Jacobian of ψ is

$$\mathrm{J}_\psi(r,\theta) = \begin{vmatrix} \dfrac{\partial\psi_1}{\partial r} & \dfrac{\partial\psi_1}{\partial\theta} \\ \dfrac{\partial\psi_2}{\partial r} & \dfrac{\partial\psi_2}{\partial\theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\,.$$

Therefore, if $f : \mathrm{D} \to \mathbb{R}$ is Riemann integrable, then

$$\begin{aligned}\int_{\mathrm{D}} f(x,y)\,d(x,y) &= \int_{\psi((0,R)\times(0,2\pi))} f(x,y)\,d(x,y) = \int_{(0,R)\times(0,2\pi)} (f\circ\psi)(r,\theta)\big|\mathrm{J}_\psi(r,\theta)\big|\,d(r,\theta) \\ &= \int_{(0,R)\times(0,2\pi)} f(x_0 + r\cos\theta, y_0 + r\sin\theta)\,r\,d(r,\theta)\,.\end{aligned}$$

Example 8.75 (Cylindrical coordinates)**.** In $\mathbb{R}^3$, when the domain over which the integral is taken is a cylinder C; that is, $\mathrm{C} = \mathrm{D} \times [a, b]$ for some disk D and $-\infty < a < b < \mathbb{R}$, then the change of variables

$$\psi(r,\theta,z) = (x_0 + r\cos\theta, y_0 + r\sin\theta, z) \qquad 0 < r < R\,, 0 < \theta < 2\pi\,, a \leqslant z \leqslant b\,,$$

where (x_0, y_0) is the center of D and R is the radisu of D, is sometimes very useful for evaluating the integral. Since the Jacobian of ψ is

$$\mathrm{J}_\psi(r,\theta,z) = \begin{vmatrix} \dfrac{\partial\psi_1}{\partial r} & \dfrac{\partial\psi_1}{\partial\theta} & \dfrac{\partial\psi_1}{\partial z} \\ \dfrac{\partial\psi_2}{\partial r} & \dfrac{\partial\psi_2}{\partial\theta} & \dfrac{\partial\psi_2}{\partial z} \\ \dfrac{\partial\psi_3}{\partial r} & \dfrac{\partial\psi_3}{\partial\theta} & \dfrac{\partial\psi_3}{\partial z} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r\,,$$

we must have

$$\begin{aligned}\int_{\mathrm{C}} f(x,y,z)\,d(x,y,z) &= \int_{\psi((0,R)\times(0,2\pi)\times[a,b])} f(x,y,z)\,d(x,y,z)\\ &= \int_{(0,R)\times(0,2\pi)\times[a,b]} (f\circ\psi)(r,\theta,z)\big|\mathrm{J}_\psi(r,\theta,z)\big|\,d(r,\theta,z)\\ &= \int_{(0,R)\times(0,2\pi)\times[a,b]} f(x_0+r\cos\theta, y_0+r\sin\theta, z)\,r\,d(r,\theta,z)\,.\end{aligned}$$

Example 8.76 (Spherical coordinates)**.** In $\mathbb{R}^3$, when the domain over which the integral is taken is a ball B, the change of variables

$$\psi(\rho,\theta,\phi) = (x_0+\rho\cos\theta\sin\phi, y_0+\rho\sin\theta\sin\phi, z_0+\rho\cos\phi) \quad 0<\rho<R, 0<\theta<2\pi, 0<\phi<\pi,$$

where (x_0,y_0,z_0) is the center of B and R is the radius of B, is often used to evaluate the integral a function over B. Since the Jacobian of ψ is

$$\begin{aligned}\mathrm{J}_\psi(\rho,\theta,\phi) &= \begin{vmatrix} \dfrac{\partial\psi_1}{\partial\rho} & \dfrac{\partial\psi_1}{\partial\theta} & \dfrac{\partial\psi_1}{\partial\phi}\\ \dfrac{\partial\psi_2}{\partial\rho} & \dfrac{\partial\psi_2}{\partial\theta} & \dfrac{\partial\psi_2}{\partial\phi}\\ \dfrac{\partial\psi_3}{\partial\rho} & \dfrac{\partial\psi_3}{\partial\theta} & \dfrac{\partial\psi_3}{\partial\phi}\end{vmatrix} = \begin{vmatrix}\cos\theta\sin\phi & -\rho\sin\theta\sin\phi & \rho\cos\theta\cos\phi\\ \sin\theta\sin\phi & \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi\\ \cos\phi & 0 & -\rho\sin\phi\end{vmatrix}\\ &= -\rho^2\cos^2\theta\sin^3\phi - \rho^2\sin^2\theta\sin\phi\cos^2\phi - \rho^2\cos^2\theta\sin\phi\cos^2\phi - \rho^2\sin^2\theta\sin^3\phi\\ &= -\rho^2\sin^3\phi - \rho^2\sin\phi\cos^2\phi = -\rho^2\sin\phi\,,\end{aligned}$$

if the radius of B is R, we must have

$$\begin{aligned}\int_{\mathrm{B}} f(x,y,z)\,d(x,y,z) &= \int_{\psi((0,R)\times(0,2\pi)\times(0,\pi))} f(x,y,z)\,d(x,y,z)\\ &= \int_{(0,R)\times(0,2\pi)\times(0,\pi)} (f\circ\psi)(\rho,\theta,\phi)\big|\mathrm{J}_\psi(\rho,\theta,\phi)\big|\,d(\rho,\theta,\phi)\\ &= \int_{(0,R)\times(0,2\pi)\times(0,\pi)} f(x_0+\rho\cos\theta\sin\phi, y_0+\rho\sin\theta\sin\phi, z_0+\rho\cos\phi)\,\rho^2\sin\phi\,d(r,\theta,z)\,.\end{aligned}$$

8.7 Exercises

§8.2 Conditions for Integrability

Problem 8.1. Let $f : [0,1] \times [0,1] \to \mathbb{R}$ be a bounded function such that $f(x,y) \leqslant f(x,z)$ if $y < z$ and $f(x,y) \leqslant f(t,z)$ if $x < t$. In other words, $f(x,\cdot)$ and $f(\cdot,y)$ are both non-decreasing functions for fixed $x, y \in [0,1]$. Show that f is Riemann integrable over $[0,1] \times [0,1]$.

Problem 8.2. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f_k : A \to \mathbb{R}$ be a sequence of Riemann integrable functions which converges uniformly to f on A. Show that f is Riemann integrable over A, and

$$\lim_{k\to\infty} \int_A f_k(x)\,dx = \int_A \lim_{k\to\infty} f_k(x)\,dx = \int_A f(x)\,dx\,.$$

§8.3 Lebesgue's Theorem

Problem 8.3. Complete the following.

1. Show that if A is a set of volume zero, then A has measure zero. Is it true that if A has measure zero, then A also has volume zero?

2. Let $a, b \in \mathbb{R}$ and $a < b$. Show that the interval $[a,b]$ does not have measure zero (in $\mathbb{R}$).

3. Let $A \subseteq [a,b]$ be a set of measure zero (in $\mathbb{R}$). Show that $[a,b]\backslash A$ does not have measure zero (in $\mathbb{R}$).

4. Show that the Cantor set (defined in Exercise Problem 2.11) has volume zero.

Problem 8.4. Let $A = \bigcup\limits_{k=1}^{\infty} D\big(\frac{1}{k}, \frac{1}{2^k}\big) = \bigcup\limits_{k=1}^{\infty} \big(\frac{1}{k} - \frac{1}{2^k}, \frac{1}{k} + \frac{1}{2^k}\big)$ be a subset of $\mathbb{R}$. Does A have volume?

Problem 8.5. Let $f : [a,b] \to \mathbb{R}$ be bounded and Riemann integrable. Show that the graph of f has volume zero by considering the difference of the upper and lower sums of f.

Problem 8.6. Let $A \subseteq \mathbb{R}^n$ be an open bounded set with volume, and $f : A \to \mathbb{R}$ be continuous. Show that if $\displaystyle\int_B f(x)\,dx = 0$ for all subsets $B \subseteq A$ with volume, then $f = 0$.

Problem 8.7. Prove the following statements.

1. The function $f(x) = \sin\dfrac{1}{x}$ is Riemann integrable over $(0,1)$.

2. Let $f:[0,1]\to\mathbb{R}$ be given by

$$f(x)=\begin{cases}\dfrac{1}{p} & \text{if } x=\dfrac{q}{p}\in\mathbb{Q},\ (p,q)=1\,,\\ 0 & \text{if } x \text{ is irrational.}\end{cases}$$

Then f is Riemann integrable over $[0,1]$. Find $\displaystyle\int_0^1 f(x)dx$ as well.

3. Let $A\subseteq\mathbb{R}^n$ be a bounded set, and $f:A\to\mathbb{R}$ is Riemann integrable. Then f^k（f 的 k 次方）is integrable for all $k\in\mathbb{N}$.

Problem 8.8. Suppose that $f:[a,b]\to\mathbb{R}$ is Riemann integrable, and the set $\big\{x\in[a,b]\,\big|\,f(x)\neq 0\big\}$ has measure zero. Show that $\displaystyle\int_a^b f(x)\,dx=0$.

§8.5 Fubini's Theorem

Problem 8.9. Evaluate the iterated integral $\displaystyle\int_0^1\int_0^x (2y-y^2)^{\frac{2}{3}}\,dydx$.

Problem 8.10. Let $A=[a,b]\times[c,d]$ be a rectangle in $\mathbb{R}^2$, and $f:A\to\mathbb{R}$ be Riemann integrable. Show that the sets

$$\Big\{x\in[a,b]\,\Big|\,\underline{\int_c^d} f(x,y)dy\neq\overline{\int_c^d} f(x,y)dy\Big\}\quad\text{and}\quad\Big\{y\in[c,d]\,\Big|\,\underline{\int_a^b} f(x,y)dx\neq\overline{\int_a^b} f(x,y)dx\Big\}$$

have measure zero (in $\mathbb{R}^1$).

Problem 8.11. Define a set $S\subseteq[0,1]\times[0,1]$ by

$$S=\Big\{\big(\frac{p}{m},\frac{k}{m}\big)\in[0,1]\times[0,1]\,\Big|\,m,p,k\in\mathbb{N}\,,\gcd(m,p)=1\text{ and }1\leqslant k\leqslant m-1\Big\}\,.$$

Show that

$$\int_0^1\Big(\int_0^1 \mathbf{1}_S(x,y)\,dy\Big)dx=\int_0^1\Big(\int_0^1 \mathbf{1}_S(x,y)\,dx\Big)dy=0$$

but $\mathbf{1}_S$ is not Riemann integrable over $[0,1]\times[0,1]$.

Problem 8.12. Let $f:[0,1]\times[0,1]\to\mathbb{R}$ be given by

$$f(x,y)=\begin{cases}2^{2n} & \text{if } (x,y)\in[2^{-n},2^{-n+1})\times[2^{-n},2^{-n+1}),\ n\in\mathbb{N}\,,\\ -2^{2n+1} & \text{if } (x,y)\in[2^{-n},2^{-n+1})\times[2^{-n-1},2^{-n}),\ n\in\mathbb{N}\,,\\ 0 & \text{otherwise}\,.\end{cases}$$

1. Show that $\int_0^1 f(x,y)\,dx = 0$ for all $y \in [0,1)$.
2. Show that $\int_0^1 f(x,y)\,dy = 0$ for all $x \in \big[0, \frac{1}{2}\big)$.
3. Justify if the iterated (improper) integrals $\int_0^1 \int_0^1 f(x,y)dxdy$ and $\int_0^1 \int_0^1 f(x,y)\,dydx$ are identical.

Problem 8.13.

1. Draw the region corresponding to the integral $\int_0^1 \Big(\int_1^{e^x} (x+y)\,dy\Big)dx$ and evaluate.
2. Change the order of integration of the integral in 1 and check if the answer is unaltered.

§8.6 Change of Variables Formula

Problem 8.14. Prove Theorem 4.95 using Theorem 8.65.

Problem 8.15. Find the volume of the set $\big\{(x,y,z) \in \mathbb{R}^3 \,\big|\, 0 \leqslant x^2 + y^2 + xy \leqslant z^2 \leqslant 4\big\}$.

Problem 8.16. Suppose that $\mathcal{U} \subseteq \mathbb{R}^n$ is an nonempty open set, and $f : \mathcal{U} \to \mathbb{R}$ is of class $\mathscr{C}^1$ such that $\mathbf{J}_f(x) \neq 0$ for all $x \in \mathcal{U}$. Show that

$$\lim_{r\to 0^+} \frac{\nu\big(f(D(x_0,r))\big)}{\nu\big(D(x_0,r)\big)} = \mathbf{J}_f(x_0) \qquad \forall\, x_0 \in \mathcal{U}\,.$$

Problem 8.17. 1. Let A be the parallelogram with vertices $(0,0)$, $\big(\frac{2}{3}, -\frac{1}{3}\big)$, $(1,0)$ and $\big(\frac{1}{3}, \frac{1}{3}\big)$. Evaluate the integral

$$\int_A \sqrt{x-y}\sqrt{x+2y}\,d\mathbb{A}\,.$$

2. Let A be the parallelogram bounded by lines $x = 3y$, $x = 1 + 3y$, $y = -2x$ and $y = 1 - 2x$. Evaluate the integral

$$\int_A \sqrt[3]{2x^2 - 5xy - 3y^2}\,d\mathbb{A}\,.$$

3. Let A be the trapezoid with vertices $(1,1)$, $(2,2)$, $(2,0)$ and $(4,0)$. Evaluate the integral

$$\int_A e^{(y-x)/(y+x)}\,d\mathbb{A}\,.$$

Problem 8.18 (True or False). Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. Let $A \subseteq \mathbb{R}^n$ be bounded, and $f : A \to \mathbb{R}$ be Riemann integrable. If $\mathcal{P}$ be a partition of A, and $m \leqslant f(x) \leqslant M$ for all $x \in A$. Then $m\nu(A) \leqslant L(f, \mathcal{P}) \leqslant U(f, \mathcal{P}) \leqslant M\nu(A)$.

2. Let $A \subseteq \mathbb{R}^n$ be a set of measure zero. If $\bar{A} \backslash A$ is countable, then A has volume zero.

3. Let $A \subseteq \mathbb{R}^n$ be a closed rectangle and $f, g : B \to \mathbb{R}$ be Riemann integrable functions. If there exists a set $Z \subseteq A$ such that Z has measure zero and $g(x) = f(x)$ for all $x \in A \backslash Z$, then $\displaystyle\int_A f(x)\,dx = \int_A g(x)\,dx$.

4. Let $A \subseteq \mathbb{R}^n$ be a closed rectangle. Suppose that f and g are two bounded real-valued functions defined on A such that f is continuous and $g = f$ except on a set of measure zero, then f and g are both Riemann integrable over A.

5. Let $A, B \subseteq \mathbb{R}$ be bounded, and $f : A \to \mathbb{R}$ and $g : f(A) \to \mathbb{R}$ be Riemann integrable. Then $g \circ f$ is Riemann integrable over A.

6.