Chapter 3

Multiple Integrals

3.1 Integrable Functions

Let us start our discussion on the integrability of functions of two variables.

Definition 3.1. Let $A \subseteq \mathbb{R}^2$ be a bounded set. Define

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

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 $[a_2, b_2]$,

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

such that

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

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

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

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

Definition 3.2. Let $A \subseteq \mathbb{R}^2$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. For any partition $\mathcal{P} = \big\{\Delta_{ij} \,\big|\, \Delta_{ij} = (x_i, x_{i+1}) \times (y_j, y_{j+1}), i = 0, \cdots, n-1, j = 0, \cdots, m-1\big\}$, 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

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

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

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

and

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

are called the ***upper integral*** and ***lower integral*** of f over A, 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 ***double integral*** of f over A.

Similar to the case of double integrals, we can consider the integrability of a bounded function f defined on a bounded set $A \subseteq \mathbb{R}^{\mathrm{n}}$ as follows

Definition 3.3. Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded set. Define the numbers $a_1, a_2, \cdots, a_n$ and $b_1, b_2, \cdots, b_n$ by

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

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

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

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

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

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

Definition 3.4. Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. For any partition

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

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

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

where $\nu_{\mathrm{n}}(\Delta)$ is the n-dimensional ***volume*** of the rectangle Δ given by

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

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

$$\overline{f}^A(x) = \begin{cases} f(x) & x \in A\,, \\ 0 & x \notin A\,. \end{cases} \tag{3.1}$$

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 ***n-tuple integral*** of f over A.

Definition 3.5. 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.

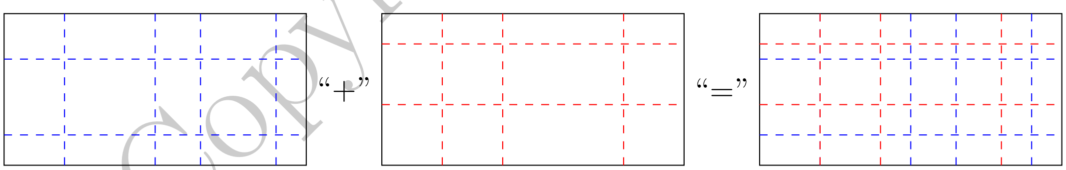

Figure 3.1: The common refinement of two partitions

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 $\mathcal{P}_i^{(j)}$ for some $i = 1, \cdots, k$.

Proposition 3.6. *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 3.7. *Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ 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*

$$L(f, \mathcal{P}_1) \leqslant U(f, \mathcal{P}_2)\,.$$

Proof. Let $\mathcal{P}$ be the common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Then Proposition 3.6 implies that

$$L(f, \mathcal{P}_1) \leqslant L(f, \mathcal{P}) \leqslant U(f, \mathcal{P}) \leqslant U(f, \mathcal{P}_2)\,. \qquad \square$$

Corollary 3.8. *Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded subset, and $f : A \to \mathbb{R}$ be a bounded function. Then*

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

Proof. Noting that for each given partition $\mathcal{P}$ of A, $L(f, \mathcal{P})$ is a lower bound for all possible upper sum; thus

$$L(f, \mathcal{P}) \leqslant \overline{\int_A} f(x)dx \qquad \text{for all partitions } \mathcal{P} \text{ of } A$$

which further implies that $\underline{\int_A} f(x)dx \leqslant \overline{\int_A} f(x)dx\,.$ $\square$

Proposition 3.9 (Riemann's condition)**.** *Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ 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*

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

Proof. "$\Rightarrow$" Let $\varepsilon > 0$ be given. By the definition of infimum and supremum, there exist partition $\mathcal{P}_1$ and $\mathcal{P}_2$ of A such that

$$\underline{\int_A} f(x)\,dx - \frac{\varepsilon}{2} < L(f, \mathcal{P}_2) \quad \text{and} \quad \overline{\int_A} f(x)\,dx + \frac{\varepsilon}{2} > U(f, \mathcal{P}_1)\,.$$

Let $\mathcal{P}$ be a common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Since f is Riemann integrable over A, $\underline{\int_A} f(x)dx = \overline{\int_A} f(x)dx$; thus Proposition 3.6 implies that

$$\begin{aligned} U(f, \mathcal{P}) - L(f, \mathcal{P}) &\leqslant U(f, \mathcal{P}_1) - L(f, \mathcal{P}_2) \\ &< \overline{\int_A} f(x)\,dx + \frac{\varepsilon}{2} - \Big(\underline{\int_A} f(x)\,dx - \frac{\varepsilon}{2}\Big) = \varepsilon\,. \end{aligned}$$

"$\Leftarrow$" Let $\varepsilon > 0$ be given. By assumption there exists a partition $\mathcal{P}$ of A such that $U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon$. Then

$$0 \leqslant \overline{\int_A} f(x)\,dx - \underline{\int_A} f(x)\,dx \leqslant U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon\,.$$

Since $\varepsilon > 0$ is given arbitrary, we must have $\underline{\int_A} f(x)dx = \overline{\int_A} f(x)dx$; thus f is Riemann integrable over A. □

Definition 3.10. 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 $\mathcal{P}$.

Theorem 3.11 (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 (3.1). 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*

$$\Big|\sum_{k=1}^{N} \overline{f}^A(\xi_k)\nu(\Delta_k) - \mathrm{I}\Big| < \varepsilon \tag{3.2}$$

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

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

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

"$\Rightarrow$" Let $\mathrm{I} = (\mathrm{D})\displaystyle\int_A f(x)dx$, and $\varepsilon > 0$ be given. Since f is Darboux integrable on A, there exists a partition $\mathcal{P}_1$ of A such that $U(f,\mathcal{P}_1) - L(f,\mathcal{P}_1) < \dfrac{\varepsilon}{2}$. 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 $\delta > 0$.

Assume that $\mathcal{P} = \{\Delta_1, \Delta_2, \cdots, \Delta_N\}$ is a given partition of A with $\|\mathcal{P}\| < \delta$, and $\Xi = \{\xi_1, \cdots, \xi_N\}$ is a set satisfying that $\xi_k \in \Delta_k$ for all $1 \leqslant k \leqslant N$. 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,

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

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

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

and

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

The equalities above further imply that

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

Moreover, for each fixed i, j,

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

thus

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

Therefore,

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

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

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

Therefore,

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

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. □

Definition 3.12. A bounded set $A\subseteq\mathbb{R}^n$ is said to ***have volume*** if the constant function $f(x)=1$ for all $x\in A$ is Riemann integrable on A. The number $\int_A 1\,dx$ is called the ***volume*** of A and is denoted by $\nu(A)$. If $\nu(A)=0$, then A is said to have volume zero or be a set of volume zero.

Remark 3.13. 1. For a set $A\subseteq\mathbb{R}^n$, the characteristic function or indicator function of A, denoted by $\mathbf{1}_A$ or χ_A, is given by

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

Then a bounded set A has volume if and only if $\mathbf{1}_A$ is Riemann integrable on A.

2. Having defined the indicator function, then for a bounded function $f : A \to \mathbb{R}$ with bounded domain A, any given partition $\mathcal{P}$ of A we have $\overline{f}^{A} = f\mathbf{1}_A$; thus

$$U(f,\mathcal{P}) = \sum_{\Delta\in\mathcal{P}} \sup_{x\in\Delta}(f\mathbf{1}_A)(x)\nu(\Delta) \qquad \text{and} \qquad L(f,\mathcal{P}) = \sum_{\Delta\in\mathcal{P}} \inf_{x\in\Delta}(f\mathbf{1}_A)(x)\nu(\Delta)\,.$$

3.2 Properties of the Integrals

Proposition 3.14. *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 $\overline{\int}_A (cf)(x)\,dx = c\overline{\int}_A f(x)\,dx$. If $c < 0$, then $\underline{\int}_A (cf)(x)\,dx = c\overline{\int}_A f(x)\,dx$ and $\overline{\int}_A (cf)(x)\,dx = c\underline{\int}_A f(x)\,dx$.*

(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) is 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 3.6 implies that

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

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

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

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

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

which further implies that

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

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

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

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 $c = -1$. Let $\varepsilon > 0$ be given. Then there exist partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of A such that

$$\underline{\int}_A -f(x)\,dx - \varepsilon < L(-f,\mathcal{P}_1) \qquad \text{and} \qquad U(f,\mathcal{P}_2) < \overline{\int}_A f(x)\,dx + \varepsilon\,.$$

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

$$\underline{\int}_A -f(x)\,dx - \varepsilon < L(-f,\mathcal{P}_1) \leqslant L(-f,\mathcal{P}) \leqslant \underline{\int}_A -f(x)\,dx$$

and

$$\overline{\int}_A f(x)\,dx \leqslant U(f,\mathcal{P}) \leqslant U(f,\mathcal{P}_2) < \overline{\int}_A f(x)\,dx + \varepsilon\,.$$

By the fact that

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

we find that

$$\underline{\int}_A -f(x)\,dx - \varepsilon < L(-f,\mathcal{P}) = -U(f,\mathcal{P}) \leqslant -\overline{\int}_A f(x)\,dx$$

and

$$\underline{\int}_A -f(x)\,dx \geqslant L(-f,\mathcal{P}) = -U(f,\mathcal{P}) > -\overline{\int}_A f(x)\,dx - \varepsilon\,.$$

Therefore,

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

Since $\varepsilon > 0$ is given arbitrarily, we conclude (c).

(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, $-1_A \leqslant \dfrac{f}{M} \leqslant 1_A$ 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 3.8 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 3.15. 1. Let $A \subseteq \mathbb{R}^n$ be bounded, and $f : A \to \mathbb{R}$ be bounded. Then (a) of Proposition 3.14 shows that if $B \subseteq A$, then f is Riemann integrable on B if and only if $f\mathbf{1}_B$ is Riemann integrable on A.

2. Let $A \subseteq \mathbb{R}^n$ be bounded and $f, g : A \to \mathbb{R}$ be bounded. Then (b) of Proposition 3.14 also implies that

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

Corollary 3.16. *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 3.14 and Remark 3.15 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 3.14 (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 3.14, it is easy to see the following theorem (except (c)). The proof is left as an exercise.

Theorem 3.17. *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* $\displaystyle\Big|\int_A f(x)\,dx\Big| \leqslant \int_A |f(x)|dx$.

(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* $\displaystyle\Big|\int_A f(x)\,dx\Big| \leqslant M\nu(A)$.

Definition 3.18. 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 two theorems are direct consequences of (a) of Proposition 3.14 and Corollary 3.16.

Theorem 3.19. *Let $A, B \subseteq \mathbb{R}^n$ be bounded, $B \subseteq A$, and $f : A \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable over B if and only if $f|_B$ is Riemann integrable over A. In either cases,*

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

Theorem 3.20. *Let A, B be bounded subsets of $\mathbb{R}^n$ be such that $A \cap B$ has volume zero, and $f : A \cup B \to \mathbb{R}$ be bounded such that $f\big|_A$ and $f\big|_B$ are all Riemann integrable over $A \cup B$. Then f is Riemann integrable over $A \cup B$, and*

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

3.3 Integrability for Almost Continuous Functions

Lemma 3.21. *Let $A \subseteq \mathbb{R}^n$ be a bounded set of volume zero. If $B \subseteq A$, then B has volume zero.*

Proof. By (a), (d) and (e) of Proposition 3.14,

$$0 = \int_A \mathbf{1}_B(x)\,dx = \underline{\int}_A \mathbf{1}_B(x)\,dx = \underline{\int}_B \mathbf{1}_B(x)\,dx$$

and

$$0 = \int_A \mathbf{1}_B(x)\,dx = \overline{\int}_A \mathbf{1}_B(x)\,dx = \overline{\int}_B \mathbf{1}_B(x)\,dx\,.$$

Therefore, $\displaystyle\int_B \mathbf{1}_B(x)\,dx = 0$ which implies that B has volume zero. □

Lemma 3.22. *Let $A_1, \cdots, A_k \subseteq \mathbb{R}^n$ be bounded sets of volume zero. Then $\displaystyle\bigcup_{j=1}^{k} A_j$ has volume zero.*

Proof. It suffices to prove the case for $k = 2$. Suppose that A_1 and A_2 are bounded sets of volume zero, and $A = A_1 \cup A_2$. By Lemma 3.21, $A_1 \cap A_2$ has volume zero; thus (e) of Proposition 3.14 and Corollary 3.16 imply that

$$\underline{\int}_A \mathbf{1}_A(x)\,dx = \underline{\int}_{A_1 \cup A_2} \mathbf{1}_A(x)\,dx \geqslant \underline{\int}_{A_1} \mathbf{1}_A(x)\,dx + \underline{\int}_{A_2} \mathbf{1}_A(x)\,dx = 0$$

and

$$\overline{\int}_A \mathbf{1}_A(x)\,dx = \overline{\int}_{A_1 \cup A_2} \mathbf{1}_A(x)\,dx \leqslant \overline{\int}_{A_1} \mathbf{1}_A(x)\,dx + \overline{\int}_{A_2} \mathbf{1}_A(x)\,dx = 0\,.$$

Therefore, $\displaystyle\int_A \mathbf{1}_A(x)\,dx = 0$ which implies that A has volume zero. □

Theorem 3.23. *Let $A \subseteq \mathbb{R}^n$ be a bounded set such that ∂A has volume zero, and $f : A \to \mathbb{R}$ be a bounded function. If f is continuous except perhaps on a set of volume zero, then f is Riemann integrable over A.*

Proof. Let R be a closed cube such that $A \subseteq R$ and $\partial A \cap \partial R = \varnothing$. We show that $\overline{f}^A = f\mathbf{1}_A$ is Riemann integrable over R and by (a) of Proposition 3.14, we then obtain that

$$\underline{\int}_A f(x)\,dx = \underline{\int}_R (f\mathbf{1}_A)(x)\,dx = \int_R (f\mathbf{1}_A)(x)\,dx = \overline{\int}_R (f\mathbf{1}_A)(x)\,dx = \overline{\int}_A f(x)\,dx$$

which implies that f is Riemann integrable over A.

Let $\varepsilon > 0$ be given. Suppose that the collection of discontinuities of f is D, and $B = \partial A \cup D$. Since ∂A and D has volume zero, Lemma 3.22 implies that B has volume zero; thus (a) of Proposition 3.14 then implies (with $B \subseteq R$ in mind) that

$$\underline{\int}_R \mathbf{1}_B(x)\,dx = \underline{\int}_B \mathbf{1}_B(x)\,dx = 0 \qquad \text{and} \qquad \overline{\int}_R \mathbf{1}_B(x)\,dx = \overline{\int}_B \mathbf{1}_B(x)\,dx = 0\,.$$

Therefore, $\displaystyle\int_R \mathbf{1}_B(x)\,dx = 0$, so there exists a partition $\mathcal{P}_1$ of R such that

$$\sum_{\Delta\in\mathcal{P}_1, \Delta\cap B\neq\varnothing} \nu(\Delta) = U(\mathbf{1}_B, \mathcal{P}_1) < \frac{\varepsilon}{2\big[\sup \overline{f}^A(R) - \inf \overline{f}^A(R) + 1\big]}\,.$$

Let $\mathcal{U} \equiv \operatorname{int}\Big(\displaystyle\bigcup_{\Delta\in\mathcal{P}_1, \Delta\cap B\neq\varnothing} \Delta\Big)$. Then $B \subseteq \mathcal{U}$. Since the discontinuity of $\overline{f}^A$ is a subset of B, $\overline{f}^A : R \cap \mathcal{U}^\complement \to \mathbb{R}$ is continuous. Since $R \cap \mathcal{U}^\complement$ is closed and bounded, $\overline{f}^A$ is uniformly continuous; thus there exists $\delta > 0$ such that

$$\big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| < \frac{\varepsilon}{2\nu(R)} \qquad \text{if } x_1, x_2 \in R \cap \mathcal{U}^\complement \text{ and } \|x_1 - x_2\| < \delta\,.$$

Let $\mathcal{P}$ be a refinement of $\mathcal{P}_1$ such that $\|\mathcal{P}\| < \delta$, and define two classes C_1, C_2 of rectangles in $\mathcal{P}$ by $C_1 = \big\{\Delta' \in \mathcal{P} \,\big|\, \Delta' \nsubseteq \Delta \text{ for all } \Delta \in \mathcal{P}_1 \text{ satisfying } \Delta \cap B \neq \varnothing\big\}$ and $C_2 = \big\{\Delta' \in \mathcal{P} \,\big|\, \Delta' \notin C_1\big\}$. Then if $\Delta' \in C_1$, then $\Delta' \subseteq R \backslash \mathcal{U}^\complement$; thus

$$\begin{aligned}
U(\overline{f}^A, \mathcal{P}) - L(\overline{f}^A, \mathcal{P}) &= \sum_{\Delta'\in\mathcal{P}} \Big[\sup_{x\in\Delta'} (\overline{f}^A \mathbf{1}_R)(x) - \inf_{x\in\Delta'} (\overline{f}^A \mathbf{1}_R)(x)\Big]\nu(\Delta') \\
&= \Big(\sum_{\Delta'\in C_1} + \sum_{\Delta'\in C_2}\Big)\Big[\sup_{x\in\Delta'} \overline{f}^A(x) - \inf_{x\in\Delta'} \overline{f}^A(x)\Big]\nu(\Delta') \\
&\leqslant \frac{\varepsilon}{2\nu(R)} \sum_{\Delta'\in C_1} \nu(\Delta') + \Big[\sup \overline{f}^A(R) - \inf \overline{f}^A(R)\Big] \sum_{\Delta'\in C_2} \nu(\Delta') \\
&= \frac{\varepsilon}{2\nu(R)} \nu(R) + \Big[\sup \overline{f}^A(R) - \inf \overline{f}^A(R)\Big] \sum_{\Delta\in\mathcal{P}_1, \Delta\cap B\neq\varnothing} \nu(\Delta) \\
&< \frac{\varepsilon}{2} + \frac{\big[\sup \overline{f}^A(R) - \inf \overline{f}^A(R)\big]\varepsilon}{2\big[\sup \overline{f}^A(R) - \inf \overline{f}^A(R) + 1\big]} < \varepsilon\,,
\end{aligned}$$

and we conclude that f is Riemann integrable over A by Riemann's condition. $\square$

3.4 The Fubini theorem

If $f : [a,b] \to \mathbb{R}$ is continuous, the fundamental theorem of Calculus 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 $A \subseteq \mathbb{R}^{\mathrm{n}}$, where $\mathrm{n} \geqslant 2$, can be computed if the integral exists.

Definition 3.24. Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ and $B \subseteq \mathbb{R}^{\mathrm{m}}$ be bounded sets, $S = A \times B$ be a product set in $\mathbb{R}^{\mathrm{n+m}}$, and $f : S \to \mathbb{R}$ be bounded. For each fixed $x \in A$, the lower integral of the function $f(x,\cdot) : B \to \mathbb{R}$ is denoted by $\underline{\int}_B f(x,y)\,dy$, and the upper integral of $f(x,\cdot) : B \to \mathbb{R}$ is denoted by $\overline{\int}_B f(x,y)\,dy$. If for each $x \in A$ the upper integral and the lower integral of $f(x,\cdot) : B \to \mathbb{R}$ are the same, we simply write $\int_B f(x,y)\,dy$ for the integrals of $f(x,\cdot)$ over B. The integrals $\overline{\int}_A f(x,y)\,dx$, $\underline{\int}_A f(x,y)\,dx$ and $\int_A f(x,y)\,dx$ are defined in a similar way.

Theorem 3.25 (Fubini's Theorem)**.** *Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ and $B \subseteq \mathbb{R}^{\mathrm{m}}$ be bounded sets, and $f : A \times B \to \mathbb{R}$ be bounded. For $x \in \mathbb{R}^{\mathrm{n}}$ and $y \in \mathbb{R}^{\mathrm{m}}$, write $z = (x,y)$. Then*

$$\underline{\int}_{A\times B} f(z)\,dz \leqslant \underline{\int}_A \Big(\underline{\int}_B f(x,y)dy\Big)dx \leqslant \overline{\int}_A \Big(\overline{\int}_B f(x,y)dy\Big)dx \leqslant \overline{\int}_{A\times B} f(z)\,dz \tag{3.3}$$

and

$$\underline{\int}_{A\times B} f(z)\,dz \leqslant \underline{\int}_B \Big(\underline{\int}_A f(x,y)dx\Big)dy \leqslant \overline{\int}_B \Big(\overline{\int}_A f(x,y)dx\Big)dy \leqslant \overline{\int}_{A\times B} f(z)\,dz\,. \tag{3.4}$$

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 (3.3). 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 $\mathcal{P} = \big\{\Delta = R \times S \,\big|\, R \in \mathcal{P}_x, S \in \mathcal{P}_y\big\}$. By Proposition

3.14 and Corollary 3.16, we find that

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

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 (3.3) is concluded. ▫

Corollary 3.26. *Let $S \subseteq \mathbb{R}^{\mathrm{n}}$ be a closed and bounded set such that ∂S has volume zero, $\varphi_1, \varphi_2 : S \to \mathbb{R}$ be continuous maps such that $\varphi_1(x) \leqslant \varphi_2(x)$ for all $x \in S$, $A = \big\{(x,y) \in \mathbb{R}^{\mathrm{n}} \times \mathbb{R} \,\big|\, x \in S, \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\big\}$, and $f : A \to \mathbb{R}$ be continuous. Then f is Riemann integrable over A, and*

$$\int_A f(x,y)\,d(x,y) = \int_S \Big(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy\Big)dx\,. \tag{3.5}$$

Proof. To establish that f is Riemann integrable over A, by Theorem 3.23 it suffices to show that ∂A has volume zero. Let $m = \min\limits_{x\in S} \varphi_1(x)$ and $M = \max\limits_{x\in S} \varphi_2(x)$. Since

$$\partial A \subseteq \big\{(x,\varphi_1(x)) \,\big|\, x \in S\big\} \cup \big\{(x,\varphi_2(x)) \,\big|\, x \in S\big\} \cup \big(\partial S \times [m,M]\big)\,,$$

to see ∂A has volume zero it suffices to show that $\partial S \times [m,M]$, $\big\{(x,\varphi_1(x)) \,\big|\, x \in S\big\}$ and $\big\{(x,\varphi_2(x)) \,\big|\, x \in S\big\}$ have volume zero because of Lemma 3.21 and 3.22. Note that Theorem 3.23 implies that φ_1 is Riemann integrable over S; thus for a given $\varepsilon > 0$ there exists partition $\mathcal{P}$ of S such that

$$U(\varphi_1,\mathcal{P}) - L(\varphi_1,\mathcal{P}) < \varepsilon\,.$$

Let $B = \bigcup_{\Delta\in\mathcal{P},\Delta\cap S\neq\varnothing} \Delta \times \big[\inf_{x\in\Delta} \overline{\varphi_1}^S(x), \sup_{x\in\Delta} \overline{\varphi_1}^S(x)\big]$. Then $C \equiv \big\{(x,\varphi_1(x)) \,\big|\, x \in S\big\} \subseteq B$ and

$$\begin{aligned}0 \leqslant \overline{\int_C} \mathbf{1}_C(z)\,dz \leqslant \overline{\int_B} \mathbf{1}_B(z)\,dz &\leqslant \sum_{\Delta\in\mathcal{P},\Delta\cap S\neq\varnothing} \big(\sup_{x\in\Delta}\overline{\varphi_1}^S(x) - \inf_{x\in\Delta} \overline{\varphi_1}^S(x)\big) \times \nu_{\rm n}(\Delta) \\ &\leqslant U(\varphi_1,\mathcal{P}) - L(\varphi_1,\mathcal{P}) < \varepsilon\,.\end{aligned}$$

Therefore, $C = \big\{(x,\varphi_1(x)) \,\big|\, x \in S\big\}$ has volume zero and similarly, $\big\{(x,\varphi_2(x)) \,\big|\, x \in S\big\}$ has volume zero.

Now we show that $\partial S \times [m,M]$ has volume zero. Since ∂S has volume zero in $\mathbb{R}^{\rm n}$, for a given $\varepsilon > 0$ there exists a partition $\mathcal{P}$ of ∂S such that

$$U(\mathbf{1}_S,\mathcal{P}) < \frac{\varepsilon}{M-m+1}\,.$$

Then $\partial S \times [m,M] \subseteq \bigcup_{\Delta\in\mathcal{P},\Delta\cap\partial S\neq\varnothing} \Delta \times [m,M]$, and as above

$$\overline{\int_{\partial S\times[m,M]}} \mathbf{1}_{\partial S\times[m,M]}(z)\,dz \leqslant \sum_{\Delta\in\mathcal{P},\Delta\cap\partial S\neq\varnothing} \nu_{\rm n}(\Delta) \times (M-m) \leqslant (M-m)U(\mathbf{1}_S,\mathcal{P}) < \varepsilon\,.$$

Therefore, $\partial S \times [m,M]$ has volume zero; thus we establish that f is Riemann integrable over A.

Next we prove (3.5). Note that $A \subseteq S \times [m,M]$; thus Theorem 3.20 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 $y = \varphi_2(x)$, Theorem 3.23 implies that $\overline{f}^A(x,\cdot)$ is Riemann integrable over $[m,M]$ for each $x \in S$; thus $\underline{\int_m^M} \overline{f}^A(x,y)\,dy = \overline{\int_m^M} \overline{f}^A(x,y)\,dy$ which further implies that

$$\int_A f(x,y)\,d(x,y) = \int_S \Big(\int_m^M \overline{f}^A(x,y)\,dy\Big)dx\,. \tag{3.6}$$

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 3.14 (a) implies that

$$\int_m^M \overline{f}^A(x,y)\,dy = \int_{A_x} f(x,y)\,dy = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy \qquad \forall\, x \in S\,. \tag{3.7}$$

Combining (3.6) and (3.7), we conclude (3.5). □

Example 3.27. Let $A = \{(x,y) \in \mathbb{R}^2 \,|\, 0 \leqslant x \leqslant 1, x \leqslant y \leqslant 1\}$, and $f : A \to \mathbb{R}$ be given by $f(x,y) = xy$. Then Corollary 3.26 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 $A = \{(x,y) \in \mathbb{R}^2 \,|\, 0 \leqslant y \leqslant 1, 0 \leqslant x \leqslant y\}$, 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 3.28. 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 3.26 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

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

Example 3.29. 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 $\overline{f} : \mathbb{R}^3 \to \mathbb{R}$ be the extension of f by zero outside A. Then Theorem 3.23 implies that f is Riemann integrable. Write $\widehat{x}_1 = (x_2,x_3)$, $\widehat{x}_2 = (x_1,x_3)$ and $\widehat{x}_3 = (x_1,x_2)$. Theorem 3.20 implies that

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

and Theorem 3.25 implies that

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

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

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

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 3.30. In this example we compute the volume $\omega_{\rm n}$ of the n-dimensional unit ball. By the Fubini theorem,

$$\omega_{\rm n} = \int_{-1}^1 \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{\rm n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{\rm n-1}^2}} dx_{\rm n} \cdots dx_1 \,.$$

Note that the integral $\displaystyle\int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{\rm n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{\rm n-1}^2}} dx_{\rm n} \cdots dx_2$ is in fact $\omega_{\rm n-1}(1 - x_1^2)^{\frac{\rm n-1}{2}}$, the volume of $(n-1)$-dimensional ball of radius $\sqrt{1 - x_1^2}$; thus

$$\omega_{\rm n} = \int_{-1}^1 \omega_{\rm n-1}(1 - x^2)^{\frac{\rm n-1}{2}} dx = 2\,\omega_{\rm n-1} \int_0^{\frac{\pi}{2}} \cos^{\rm n}\theta \, d\theta \,. \tag{3.8}$$

Integrating by parts,

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

which implies that

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

As a consequence,

$$\int_0^{\frac{\pi}{2}} \cos^{\rm n}\theta\,d\theta = \left\{\begin{array}{ll} \dfrac{({\rm n}-1)({\rm n}-3)\cdots 2}{{\rm n}({\rm n}-2)\cdots 3}\displaystyle\int_0^{\frac{\pi}{2}} \cos\theta\,d\theta & \text{if n is odd}\,, \\ \dfrac{({\rm n}-1)({\rm n}-3)\cdots 1}{{\rm n}({\rm n}-2)\cdots 2}\displaystyle\int_0^{\frac{\pi}{2}} d\theta & \text{if n is even}\,, \end{array}\right.$$

and the recursive formula (3.8) implies that $\omega_{\rm n} = \dfrac{2\,\omega_{{\rm n}-2}}{n}\pi$. Further computations shows that

$$\omega_{\rm n} = \left\{\begin{array}{ll} \dfrac{(2\pi)^{\frac{{\rm n}-1}{2}}}{n({\rm n}-2)\cdots 3}\omega_1 & \text{if n is odd}\,, \\ \dfrac{(2\pi)^{\frac{{\rm n}-2}{2}}}{n({\rm n}-2)\cdots 4}\omega_2 & \text{if n is even}\,. \end{array}\right.$$

Let Γ be the Gamma function defined by $\Gamma(t) = \displaystyle\int_0^\infty x^{t-1}e^{-x}\,dx$ 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 $\omega_{\rm n}$ as

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

3.5 The 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 is easier to be evaluated. In this section, we establish the change of variables formula for the integral of functions of several variables.

Theorem 3.31 (Change of Variables Formula)**.** *Let $\mathcal{U} \subseteq \mathbb{R}^{\rm n}$ be an open set with volume, and $\psi : \mathcal{U} \to \mathbb{R}^{\rm n}$ be an one-to-one $\mathscr{C}^1$-mapping with $\mathscr{C}^1$-inverse; that is, $\psi^{-1} : \psi(\mathcal{U}) \to \mathcal{U}$ is also continuously differentiable. Assume that the Jacobian of ψ, ${\rm J} = \det([D\psi])$, does not vanish in $\mathcal{U}$. If $f : \psi(\mathcal{U}) \to \mathbb{R}$ is Riemann integrable, then $(f\circ\psi){\rm J}$ is Riemann integrable* *over* $\mathcal{U}$*, and*

$$\int_{\psi(\mathcal{U})} f(y)\,dy = \int_{\mathcal{U}} (f\circ\psi)(x)\big|\mathrm{J}(x)\big|\,dx = \int_{\mathcal{U}} (f\circ\psi)(x)\Big|\frac{\partial(\psi_1,\cdots,\psi_{\mathrm{n}})}{\partial(x_1,\cdots,x_{\mathrm{n}})}\Big|\,dx\,.$$

The proof of Theorem 3.31 is very lengthy and requires a bit more knowledge about the integration, so we only present the proof of a much simpler case.

Theorem 3.32. *Let* $\mathrm{D}\subseteq\mathbb{R}^{\mathrm{n}}$ *be an open rectangle, and* $\psi:\mathbb{R}^{\mathrm{n}}\to\mathbb{R}^{\mathrm{n}}$ *be an one-to-one* $\mathscr{C}^2$ *mapping such that* $\psi=\mathrm{Id}$ *outside* $B(0,r)$ *for some* $r>0$*; that is,* $\psi(x)=x$ *if* $|x|\geqslant r$*. Assume that the Jacobian of* ψ*,* $\mathrm{J}=\det(\nabla\psi)$*, does not vanish in* $\mathbb{R}^{\mathrm{n}}$*. If* $f:\mathrm{D}\to\mathbb{R}$ *is of class* $\mathscr{C}^1$ *and is compactly supported in* D*; that is,* $\mathrm{cl}\big(\{x\in\mathrm{D}\,|\,f(x)\neq 0\}\big)\subseteq\mathrm{D}$*, then*

$$\int_{\mathrm{D}} f(y)\,dy = \int_{\psi^{-1}(\mathrm{D})} (f\circ\psi)(x)\mathrm{J}(x)\,dx\,.$$

Proof. W.L.O.G. we can assume that $\mathrm{D}=[-R,R]^n$ is a cube and $B(0,r)\subset\subset\mathrm{D}$ (or equivalently, $0<r<R$). Then $\psi^{-1}(\mathrm{D})=\mathrm{D}$ since $\psi=\mathrm{Id}$ outside $B(0,R)$. Define

$$g(y_1,\cdots,y_n)=\int_{-R}^{y_1} f(z,y_2,\cdots,y_n)\,dz\,,$$

and $\mathrm{M}=\begin{bmatrix}[D(g\circ\psi)]\\ [D\psi_2]\\ [D\psi_3]\\ \vdots\\ [D\psi_n]\end{bmatrix}$. By the property of determinants and the chain rule, we find that

$$\det(\mathrm{M})=\det\left(\begin{bmatrix}\sum\limits_{j=1}^n\big(\frac{\partial g}{\partial y_j}\circ\psi\big)\frac{\partial\psi_j}{\partial x_1} & \sum\limits_{j=1}^n\big(\frac{\partial g}{\partial y_j}\circ\psi\big)\frac{\partial\psi_j}{\partial x_2} & \cdots & \sum\limits_{j=1}^n\big(\frac{\partial g}{\partial y_j}\circ\psi\big)\frac{\partial\psi_j}{\partial x_n}\\ \frac{\partial\psi_2}{\partial x_1} & \frac{\partial\psi_2}{\partial x_2} & \cdots & \frac{\partial\psi_2}{\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial\psi_n}{\partial x_1} & \frac{\partial\psi_n}{\partial x_2} & \cdots & \frac{\partial\psi_n}{\partial x_n}\end{bmatrix}\right)$$

$$= \det\left(\begin{bmatrix} \big(\frac{\partial g}{\partial y_1}\circ\psi\big)\frac{\partial\psi_1}{\partial x_1} & \big(\frac{\partial g}{\partial y_1}\circ\psi\big)\frac{\partial\psi_1}{\partial x_2} & \cdots & \big(\frac{\partial g}{\partial y_1}\circ\psi\big)\frac{\partial\psi_1}{\partial x_n} \\ \frac{\partial\psi_2}{\partial x_1} & \frac{\partial\psi_2}{\partial x_2} & \cdots & \frac{\partial\psi_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\psi_n}{\partial x_1} & \frac{\partial\psi_n}{\partial x_2} & \cdots & \frac{\partial\psi_n}{\partial x_n} \end{bmatrix}\right)$$

$$= \big(\frac{\partial g}{\partial y_1}\circ\psi\big)\det\left(\begin{bmatrix} \frac{\partial\psi_1}{\partial x_1} & \frac{\partial\psi_1}{\partial x_2} & \cdots & \frac{\partial\psi_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\psi_n}{\partial x_1} & \frac{\partial\psi_n}{\partial x_2} & \cdots & \frac{\partial\psi_n}{\partial x_n} \end{bmatrix}\right) = (f\circ\psi)\mathrm{J}\,.$$

On the other hand, letting $\mathrm{A} = (D\psi)^{-1}$, then

$$\mathrm{Adj}(\mathrm{M})_{j1} = (-1)^{1+j}\det\big(\mathrm{M}(\widehat{1},\widehat{j})\big) = \mathrm{Adj}([D\psi])_{j1} = \mathrm{J}\mathrm{A}^j_1\,.$$

Computing the determinant by expanding along the first row, we obtain that

$$\det(\mathrm{M}) = \sum_{j=1}^n \mathrm{M}_{1j}\mathrm{Adj}(\mathrm{M})_{j1} = \sum_{j=1}^n \frac{\partial(g\circ\psi)}{\partial x_j}\mathrm{J}\mathrm{A}^j_1\,;$$

thus we conclude the identity

$$(f\circ\psi)\mathrm{J} = \sum_{j=1}^n \frac{\partial(g\circ\psi)}{\partial x_j}\mathrm{J}\mathrm{A}^j_1\,.$$

Therefore, with $\widehat{dx_j}$ denoting $dx_1\cdots dx_{j-1}dx_{j+1}\cdots dx_n$, the Fubini theorem and the Piola identity imply that

$$\begin{aligned}\int_{\mathrm{D}}\big[(f\circ\psi)\mathrm{J}\big](x)\,dx &= \sum_{j=1}^n\int_{-R}^R\int_{-R}^R\cdots\int_{-R}^R\frac{\partial(g\circ\psi)}{\partial x_j}\mathrm{J}\mathrm{A}^j_1\,dx_j\widehat{dx_j} \\ &= \sum_{j=1}^n\int_{-R}^R\int_{-R}^R\cdots\int_{-R}^R\big[(g\circ\psi)\mathrm{J}\mathrm{A}^j_1\big]\Big|_{x_j=-R}^{x_j=R}\widehat{dx_j}\,.\end{aligned}$$

Since $\psi = \mathrm{Id}$ outside $B(0,r)$, we find that $\mathrm{J} = 1$ and $\mathrm{A}^j_1 = \delta_{1j}$ on $\partial\mathrm{D}$; thus by the definition of g,

$$\int_{\mathrm{D}}\big[(f\circ\psi)\mathrm{J}\big](x)\,dx = \int_{-R}^R\int_{-R}^R\cdots\int_{-R}^R g(R,x_2,\cdots,x_n)\,\widehat{dx_1} = \int_{\mathrm{D}} f(x)\,dx\,. \qquad \square$$

Example 3.33. Suppose that $f : [0,1] \to \mathbb{R}$ is Riemann integrable and $\displaystyle\int_0^1 (1-x)f(x)\,dx = 5$. We would like to evaluate the iterated integral $\displaystyle\int_0^1 \int_0^x f(x-y)\,dydx$.

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

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

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

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

Example 3.34. 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 $(x,y) = g(u,v) = \big(u, \frac{u-v}{2}\big)$; 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 $A = g(E)$.

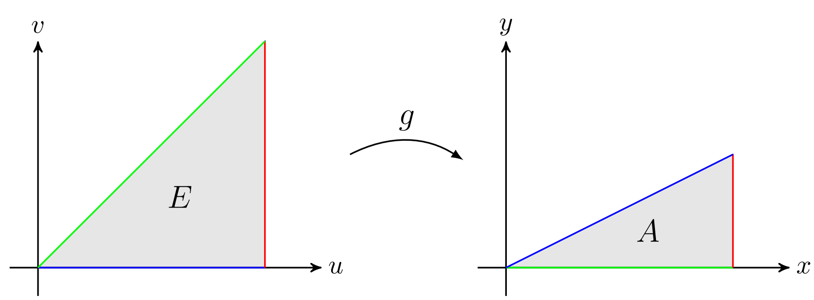

Figure 3.2: The image of E under g

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 3.35. Let A be the region in the first quadrant of the plane bounded by the curves $xy - x + y = 0$ and $x - y = 1$, and $f : A \to \mathbb{R}$ be given by

$$f(x,y) = x^2y^2(x+y)e^{-(x-y)^2}\,.$$

We would like to evaluate the integral $\displaystyle\int_A f(x,y)\,d(x,y)$.

Let $(u,v) = (xy - x + y, x - y)$. Unlike the previous two examples we do not want to solve for (x,y) in terms of (u,v) but still assume that $(x,y) = g(u,v)$. By the inverse function theorem,

$$\mathrm{J}_g(u,v)\Big|_{(u,v)=g^{-1}(x,y)} = \Big(\frac{\partial(u,v)}{\partial(x,y)}\Big)^{-1} = \begin{vmatrix} y-1 & x+1\\ 1 & -1\end{vmatrix}^{-1} = \frac{1}{-y+1-x-1} = -\frac{1}{x+y}\,.$$

Moreover, the curve $xy - x + y = 0$ corresponds to $u = 0$, while the lines $x - y = 1$ and $y = 0$ correspond to $v = 1$ and $u + v = 0$, respectively; thus if E is the region enclosed by $u = 0$, $v = 1$ and $u + v = 0$, then $A = g(E)$.

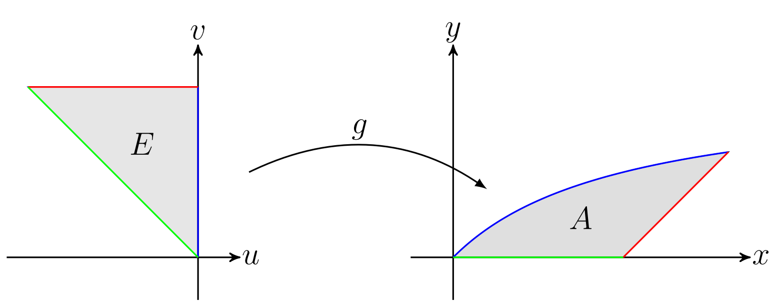

Figure 3.3: The image of E under g

Therefore,

$$\begin{aligned}\int_A f(x,y)d(x,y) &= \int_{g(E)} f(x,y)\,d(x,y) = \int_E (f\circ g)(u,v)|\mathrm{J}_g(u,v)|\,d(u,v)\\ &= \int_0^1\int_{-v}^0 (u+v)^2 e^{-v^2}\,dudv = \frac{1}{3}\int_0^1 v^3 e^{-v^2}\,dv\\ &= \frac{1}{6}\int_0^1 we^{-w}\,dw = -\frac{1}{6}(w+1)e^{-w}\Big|_{w=0}^{w=1} = -\frac{1}{6}\big(\frac{2}{e}-1\big)\,.\end{aligned}$$

Example 3.36 (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 3.37 (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 3.38 (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

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