Chapter 1

Double And Triple Integrals

1.1 Integral Over An Interval

We start by reviewing integration theory of functions of a single variable.

Given an interval $[a, b]$, a partition P on $[a, b]$ is a collection of points $\{x_j\}$ satisfying $a = x_0 < x_1 < \cdots < x_n = b$. The norm of the partition P, denoted by $\{P\}$, is the maximum of $\Delta x_j = x_j - x_{j-1}, j = 1, \cdots, n$. It measures how fine the partition is. Let f be a function defined on an interval $[a, b]$. The Riemann sum of f with respect to the partition P is defined to be

$$R(f, P) = \sum_{j=1}^{n} f(z_j)\Delta x_j \ ,$$

where the tag z_j is an arbitrary point taken from the subinterval $[x_{j-1}, x_j]$. The Riemann also depends on the choice of tag points, but we simplify things by using the same notation.

The function f is called integrable if there exists a real number α such that for every $\varepsilon > 0$, there is some $\delta > 0$ so that

$$|R(f, P) - \alpha| < \varepsilon, \quad \forall P \ , \|P\| < \delta \ .$$

We call α the integral of f over $[a, b]$ and denote it by

$$\int_a^b f \ , \quad \int_a^b f \, dx \ , \quad \text{or} \int_a^b f(x) \, dx \ .$$

When f is non-negative, obviously the Riemann sums are approximate areas and the integral is the area of the set bounded by the x-axis, the graph of f, and the vertical lines $x = a$ and $x = b$.

An immediate question arise: Are there any non-integrable functions? The answer is yes. Let me give you two examples.

First, consider the function $f(x) = 1/x$, $x \in (0,1]$ and $f(0) = 0$. This is a function defined on $[0,1]$, which is unbounded near 0. Suppose on the contrary that f is integrable. For $\varepsilon = 1$, there is some δ such that

$$|R(f,P) - \alpha| < 1 \ , \quad \forall P, \ \|P\| < \delta.$$

Fix one such P. The inequality $|R(f,P) - \alpha| < 1$ is equivalent to $-1 < R(f,P) - \alpha < 1$. In particular, $R(f,P) - \alpha < 1$, that is, $R(f,P) < 1 + \alpha$, so $f(z_1)\Delta x_1 < 1 + \alpha + \sum_{j=2}^{n} f(z_j)\Delta x_j \equiv \beta$ which is a fixed number. It shows that $1/z_1 \Delta x_1 < \beta$. Here β and Δx_1 are fixed number, but the tag point z_1 can be chosen arbitrarily from $(0,1]$. By choosing it as small as you like, you can make $1/z_1\Delta x_1$ as large as you like, and this contradicts the inequality $1/z_1\Delta x_1 < \beta$. Hence f is not integrable. In fact, it can be shown that all unbounded functions are not integrable.

Second, not all bounded functions are integrable. Consider the function g on $[0,1]$ defined by $g(x) = 0$ if x is irrational and $g(x) = 1$ if x is rational. g is a function bounded between 0 and 1. As there are rational and irrational numbers in any interval, for each partition P, when we pick a rational number z_j from $[x_j, x_{j+1}]$ to form a tagged partition, the Riemann sum $R(g,P) = \sum_j g(z_j)\Delta x_j = \sum_j \Delta x_j = 1$. On the other hand, picking tag points w_j to be irrational instead, $g(w_j) = 0$ so $R(g,P) = \sum_j g(w_j)\Delta x_j = 0$. You can see that by choosing different tags, the Riemann sums equal to 1 or 0. It cannot converge to a single number α.

Fortunately, most functions people encountered in applications are integrable. It suffices to know that all continuous functions are integrable. In fact, all functions with jump discontinuity are also integrable.

Coming to the evaluation of an integral, from the definition of integrability we have the following approach, namely, take a sequence of tagged partitions $\{P_n\}$ whose norms tend to 0, then

$$\int_a^b f\,dx = \lim_{n\to\infty} R(f,P_n) \ .$$

Although looking very simple, this method is not practical since it involves a limit process which becomes quite complicated even for very simple functions. You may try it on the functions $f(x) = x^2$ or $\sin x$. Now, we are thankful to Issac Newton for his discovery that the evaluation of an integral can be achieved by the following scheme. First, call a function F a primitive function for a given function f if F is differentiable and its derivative is equal to f, that is, $F' = f$. When f is integrable, Newton's fundamental theorem of calculus asserts that

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

As a result, using the simple fact that a primitive function of x^2 is $x^3/3$,

$$\int_a^b x^2\,dx = \frac{b^3}{3} - \frac{a^3}{3} \ .$$

Likewise, a primitive function of $\sin x$ is given by $-\cos x$, hence

$$\int_a^b \sin x \, dx = \cos a - \cos b \ .$$

Integrals that had been troubled people since the ancient times are evaluated in this way.

1.2 Double Integral in an Rectangle

Now we come to the integration of functions of two variables. This is a direct extension of what we did in the single variable case where now an interval is replaced by a rectangle.

Let $R = [a, b] \times [c, d]$ be a rectangle and f a bounded function defined in R. Likewise, here a finite set of points

$$\{(x_i, y_j) : \ a = x_0 < x_1 < \cdots < x_n = b, \ c = y_0 < y_1 < \cdots < y_m = d, \}$$

is called a *partition on* R. We denote

$$R_{ij} = [x_i, x_{i+1}] \times [y_j, y_{j+1}] \ ,$$

$$\Delta x_i = x_{i+1} - x_i \ , \quad \Delta y_j = y_{j+1} - y_j \ ,$$

and let $\|P\|$ be the maximum of all $\Delta x_i, \Delta y_j$'s. Pick a point p_{ij} from R_{ij} for each (i, j) we form a collection of tags. A partition together with a choice of tags is called a *tagged partition.*

Let f be a function defined in R. Associate to each tagged partition $(P, \{p_{ij})$, we form a *Riemann sum*

$$R(f, P) = \sum_{i,j} f(p_{ij}) |R_{ij}| \ ,$$

where $|R_{ij}| = \Delta x_i \Delta y_j$. A function is called (*Riemann*) *integrable* if there exists a number α such that, for each $\varepsilon > 0$, there is some $\delta > 0$ so that

$$|R(f, P) - \alpha| < \varepsilon \ , \quad \forall P, \ \|P\| < \delta \ .$$

The number α is called the (*Riemann*) *integral* of f over R and is usually denoted by

$$\iint_R f \, dA \ , \quad \text{or} \ \iint_R f \, dA(x, y) \ , \ \text{or} \ \iint_R f(x, y) \, dA(x, y) \ .$$

When f is nonnegative, the Riemann sums are approximate volume and the Riemann integral is the volume of the solid formed between the graph $z = f(x, y)$ and the xy-plane over R.

Just as in the single variable case, unbounded functions are not integrable. In the following discussion, it is implicitly assumed all functions in concern are bounded.

Using the definition of Riemann integral, it is routine to verify that the following basic properties hold:

Theorem 1.1. *Let f and g be integrable in R. For $\alpha, \beta \in \mathbb{R}$.*

1. *$\alpha f + \beta g$ is integrable and*
$$\iint_R (\alpha f + \beta g)\, dA = \alpha \iint_R f\, dA + \beta \iint_R g\, dA \ .$$
2. *fg is integrable.*
3. *f/g is integrable provided that $|g| \geq C$ for some positive constant C.*
4. $$\iint_R f\, dA \geq 0 \ ,$$
whenever f is non-negative.

The first property shows that all integrable functions form a real vector space and the mapping
$$f \mapsto \iint_R f\, dA$$
is a linear mapping from this vector space to the space of real numbers. The second and third properties show how nice the integration interacts with algebraic operations of functions.

The fourth property, which may be termed as positivity preserving (or more precisely non-negativity preserving), is an essential one. Note that $f \geq 0$ and $\iint_R f\, dA = 0$ do not necessarily implies $f \equiv 0$. It suffices to observe that a nonnegative function which vanishes everywhere except at finitely points satisfy these two conditions. On the other hand, it is true that they imply $f \equiv 0$ when f is continuous.

Combining linearity and positivity preserving, we have
$$\iint_R g\, dA \geq \iint f\, dA \ ,$$
provided $g \geq f$ in R.

Theorem 1.2. *1. The constant function c is integrable and*
$$\iint_R c\, dA = c|R| \ , \quad |R| \equiv (b-a)(d-c) \ .$$

2. *There are non-integrable functions in each rectangle.*
3. *Every continuous function is integrable.*

(a) is easily proved. (b) can be shown by considering the function $\varphi(x, y) = 0$ if x is a rational number in $[a, b]$ and $\varphi(x, y) = 1$ when x is irrational. Since there are rational and irrational points in each subrectangle R_{ij}, by choosing suitable tags, $\varphi(p_{ij})$ could be 0 or 1. Consequently, each $f(p_{ij})|R_{ij}|$ is either equal to 0 or $|R_{ij}|$. It follows that the Riemann sum of the same partition could be 0 or $\sum_{i,j} |R_{ij}| = (b - a)(d - c)$. It is impossible to find a number α such that $|R(f, P) - \alpha| < \varepsilon$ for all tags.

We will not prove (c), but simply point out that it is based on a fundamental result, which will be used later.

Theorem 1.3. ***(Uniform Continuity Theorem)*** *Every continuous function in a region R satisfies the following property: Given $\varepsilon > 0$, there is some $\delta > 0$ such that*

$$|f(x, y) - f(x', y')| < \varepsilon \ ,$$

for all $(x, y), (x', y') \in R$, $\sqrt{(x - x')^2 + (y - y')^2} < \delta$.

Here our concern is how to evaluate a double integral. Thankfully we have the following result which reduces it to an iterated integral (two integrals of a single variable). We do not need a new version of the fundamental theorem of calculus.

Theorem 1.4. ***(Fubini's Theorem)*** *Let f be a continuous function in R. Then*

$$\iint_R f\, dA = \int_a^b \int_c^d f(x, y)\, dydx \ .$$

The idea is simple. The double integral can be approximated by Riemann sums. Taking tags of the form (x_i^*, y_j^*), we have

$$\iint_R f\, dA \approx \sum_{i,j} f(x_i^*, y_j^*)\Delta x_i \Delta y_j = \sum_i \left(\sum_j f(x_i^*, y_j^*)\Delta y_j \right) \Delta x_i \ .$$

When $\|P\|$ is very small, both Δy_j and Δx_i are also very small,

$$\sum_i \left(\sum_j f(x_i^*, y_j^*)\Delta y_j \right) \Delta x_i \ \approx \sum_i \int_c^d f(x_i^*, y)\, dy\ \Delta x_i \approx \int_a^b \left(\int_c^d f(x, y) dy \right) dx \ .$$

A similar result holds when the role of x and y are switched. In other words,

$$\iint_R f\, dA = \int_c^d \int_a^b f(x, y)\, dxdy \ .$$

It implies the "commutative relation"

$$\int_a^b \int_c^d f(x,y)\,dydx = \int_c^d \int_a^b f(x,y)\,dxdy \ .$$

Example 1.1 Evaluate

$$\iint_R xy^2\,dA \ ,$$

where R is the rectangle $[0,1]\times[0,2]$. By Fubini's Theorem,

$$\begin{aligned}\iint_R xy^2\,dA &= \int_0^2 \int_0^1 xy^2\,dy\,dx \\ &= \int_0^2 \frac{xy^3}{3}\Big|_{y=0}^{y=1} dx \\ &= \int_0^2 \frac{x}{3}\,dx \\ &= \frac{2}{3} \ .\end{aligned}$$

Alternatively,

$$\begin{aligned}\iint_R xy^2\,dA &= \int_0^1 \int_0^2 xy^2\,dx\,dy \\ &= \int_0^1 \frac{x^2y^2}{2}\Big|_{x=0}^{x=2} dy \\ &= \int_0^1 2y^2\,dy \\ &= \frac{2}{3} \ .\end{aligned}$$

Sometimes, the order of integration matters.

Example 1.2 Evaluate

$$\iint_R x\sin xy\;dA \ ,$$

where $R=[0,1]\times[0,\pi]$.

We have

$$\begin{aligned}\iint_R x\sin xy\,dA &= \int_0^\pi \int_0^1 x\sin xy\,dxdy \\ &= \int_0^\pi \left(\frac{-\cos y}{y} + \frac{\sin y}{y^2}\right)\,dy \ .\end{aligned}$$

At this point we don't know how to proceed further. So we change the order of integration.

$$\begin{aligned}
\iint_R x\sin xy\,dA &= \int_0^1\int_0^\pi x\sin xy\,dy\,dx \\
&= \int_0^1 x\times\frac{-\cos xy}{x}\Big|_{y=0}^{y=\pi}dx \\
&= \int_0^{x^2} 0\,dy+\int_0^1(-\cos\pi x+1)\,dx \\
&= 1\ .
\end{aligned}$$

Example 1.3 Let f be the function in $R=[-1,1]\times[0,1]$ given by $f(x,y)=2,\ x^2<y$ and $f(x,y)=0,\ x^2>y$. Evaluate

$$\iint_R x^2 f(x,y)\,dA\ .$$

By Fubini's Theorem,

$$\iint_R x^2 f(x,y)\,dA=\int_{-1}^1\int_0^1 x^2 f(x,y)\,dydx\ .$$

As

$$\begin{aligned}
\int_0^1 f(x,y)\,dy &= \int_0^{x^2} f(x,y)\,dy+\int_{x^2}^1 f(x,y)\,dy \\
&= \int_{x^2}^1 2\,dy \\
&= 2(1-x^2)\ ,
\end{aligned}$$

we have

$$\iint_R x^2 f(x,y)\,dA=\int_{-1}^1 x^2\times 2(1-x^2)\,dx=2\left(\frac{x^3}{3}-\frac{x^5}{5}\right)\Big|_{-1}^1=\frac{8}{15}\ .$$

1.3 Regions In The Plane

First of all, a C^1-curve is a curve that admits a tangent at every point and the tangent changes continuously as the points vary. In a suitable coordinates, the curve can be locally expressed as the graph $(x,f(x))$ of a C^1-function, that is, a function whose derivative exists and is continuous. A curve is simple if it has no self-intersection point. It is closed if it closes up and has no endpoints. Intuitively speaking, a simple closed curve looks like a deformed circle. We will also consider piecewise C^1-curves, that is, those continuous curves which are C^1 except at finitely many points. A set which is bounded by one or several closed piecewise C^1-curves is called a *region* or a *domain*. This definition is not consistent with the usual definition of a region in mathematics literature. However, we will adopt this definition by following our textbook.

Here are some examples of regions.

- $D_r = \{(x, y) : x^2+y^2 < r^2\}$ is the disk, the region bounded by the unit circle which is a simple closed C^1-curve.
- $\{(x, y) : x^2/a^2 + y^2/b^2 = 1\}$. The ellipse is also a simple closed C^1-curve which bounds a region.
- Let C_1 and C_2 be two circles with C_1 contained in C_2. These two circles bound a region. The punctured disk $D_r \setminus \{(0, 0)\}$ is also a region where the point $\{(0, 0)\}$ may be viewed as a degenerate circle.
- Let Δ be the points lying on or inside a triangle. A triangle is a simple, closed, piecewise C^1-curve composed of three line segments. Tangents do not exist at the three vertices.
- Similarly, every polygon whose boundary is a simple, closed piecewise C^1-curve is a region.
- The cardioid $\{(r, \theta) : r = 1 + \cos\theta)\}$ (in polar coordinates) is a simple closed, piecewise C^1-curve which admits a non-differentiable point (ie, a cusp) at the origin. It also bounds a region.

A region must be bounded from its definition. It consists of interior points and boundary points. In this chapter,

A curve always means a simple, piecewise C^1-curve and a region is the plane set bounded by one or several simple, closed piecewise C^1-curves or points.

In Advanced Calculus I, the objects of study are continuous and differentiable functions. In integration theory the classes of functions are wider. Just like we are able to integrate functions with discontinuity jumps in a single variable, we can integrate functions which admit discontinuous points along some curves.

1.4 Double Integral In A Region

Now we consider double integrals over a region D which is not necessarily a rectangle. A quick way to achieve this goal is to extend f which is only defined in D to the entire space by setting it to be zero outside D. We may call it the extension of f from D. By picking a rectangle R containing D, we may simply define

$$\iint_D f dA = \iint_R \tilde{f}\, dA ,$$

where $\tilde{f}$ is the extended function of f from D. To justify this approach, we need to clarify two points. The first one is the definition must be independent of the choice of the rectangle. The next one seems more serious. Namely, even if the function f is continuous in R, the extended function $\tilde{f}$ may develop a jump discontinuity across the boundary of D.

Theorem 1.5. *Let R_1 and R_2 be two rectangles containing D in their interior. Then*

$$\iint_{R_1} \tilde{f}\, dA = \iint_{R_2} \tilde{f}\, dA ,$$

provided $\tilde{f}$ is integrable in R_1 and R_2.

Proof. Since $D \subset R_1 \cap R_2 \subset R_i, i = 1, 2$, it suffices to show

$$\int_{R_1} \tilde{f}\, dA = \int_{R_3} \tilde{f}\, dA ,$$

where $D \subset R_3 \subset R_1$. But this is obvious as $\tilde{f} \equiv 0$ in $R_1 \setminus R_3$.

□

Theorem 1.6. *Every bounded function in a rectangle R which is continuous except on one or several piecewise C^1-curves is integrable.*

This is a reasonable generalization of the integrability of functions which are piecewise continuous in the single variable case.

Proof. We will give a proof for the special case where f in continuous in the rectangle $R = [a, b] \times [c, d]$ except at a horizonal line $y = \alpha,\ \alpha \in (c, d)$. Given $\varepsilon > 0$, we first fix a small number ρ such that $M\rho(b - a) < \varepsilon/3$ (M is a bound on $|f|$). Next we define a new function f_ρ which is equal to f in $[a, b] \times [c, \alpha - \rho]$ and $[a, b] \times [\alpha + \rho, d]$, and, for linear from $(x, \alpha - \rho)$ to $(x, \alpha + \rho)$. f_ρ is continuous in R and integrable. Now, for the given ε, we can find some δ such that

$$\left| R(f_\rho, P) - \iint_R f_\rho\, dA \right| < \frac{\varepsilon}{2} , \quad \forall P, \quad \|P\| < \delta .$$

Letting $R_1 = [a,b] \times [c,\alpha]$ and $R_2 = [a,b] \times [\alpha, d]$, we have

$$\begin{aligned}
\left| R(f,P) - \int_a^b \int_c^d f(x,y)\, dydx \right| &\le |R(f,P) - R(f_\rho, P)| \\
&\quad + \left| R(f_\rho, P) - \int_a^b \int_c^d f_\rho \, dydx \right| \\
&\quad + \left| \int_a^b \int_c^d f_\rho \, dydx - \int_a^b \int_c^d f \, dydx \right| \\
&\equiv A + B + C \ .
\end{aligned}$$

To estimate (A), observing that $R(f,P) - R(f_\rho, p) = \sum_{i,j} (f(z_{ij}) - f_\rho(z_{ij}))|R_{ij}|$ where the summation is over all those subrectangles R_{ij} that touch the strip $[a,b] \times [\alpha - \rho, \alpha + \rho]$. We have

$$A = |R(f,P) - R(f_\rho, p)| \le \sum_{i,j} |f(z_{ij}) - f_\rho(z_{ij})||R_{ij}| \le 2M \times (b-a) \times 2(\rho + \delta) \ .$$

Next, by Fubini's Theorem, $B \le \varepsilon/2$ by our choice of P. Third, since $f_\rho = f$ outside $[a,b] \times [\alpha - \rho, \alpha + \rho]$, we have

$$C \le 2M(b-a) \times 2\rho \ .$$

Putting these three estimates together, we see that

$$\left| R(f,P) - \int_a^b \int_c^d f \, dydx \right| < \frac{\varepsilon}{2} + 4M(b-a)(\rho + \delta) + 4M(b-a)\rho \ .$$

The left hand side of this estimate is independent of ρ. Letting $\rho \to 0$, we obtain

$$\left| R(f,P) - \int_a^b \int_c^d f \, dydx \right| < \frac{\varepsilon}{2} + 4M(b-a)\delta \ .$$

Now we restrict δ so that $4M(b-a)\delta < \varepsilon/2$, we conclude that

$$\left| R(f,P) - \int_a^b \int_c^d f \, dA \right| < \varepsilon \ ,$$

whenever $\|P\| < \delta$. So f is integrable in R and in fact

$$\iint_R f \, dA = \int_a^b \int_c^d f \, dydx \ .$$

$\square$

Based on these two theorems, we are able to define the integral of a bounded function f over any bounded subset E in $\mathbb{R}^2$ by setting

$$\iint_E f\,dA \equiv \iint_R \tilde{f}\,dA \ , \tag{1.1}$$

where R is any rectangle containing E. The function f is called *integrable* over E provided $\tilde{f}$ is integrable over R. Similar to what is asserted in Theorem 1.5, the integrability of f is independent of the choice of R. When f is nonnegative, the integral of f over D is the volume of the solid bounded between the graph of f and the xy-plane over the region D. When we take $f \equiv 1$, the integral, which becomes

$$\iint_D 1\,dA \ ,$$

reduces to the area of D.

Using (1.1) we have the following extension of Theorem 1.1.

Theorem 1.1'. *Let f and g be integrable in the (bounded) region D. For $\alpha, \beta \in \mathbb{R}$.*

1. *$\alpha f + \beta g$ is integrable in D and*

$$\iint_D (\alpha f + \beta g)\,dA = \alpha \iint_D f\,dA + \beta \iint_D g\,dA \ .$$

2. $$\iint_D f\,dA \geq 0 \ ,$$

   provided f is non-negative.

In Theorem 1.1 (2) and (3) are concerned with the product and quotient of integrable functions. However, functions appearing in applications are mostly continuous ones. For this reason we do not formulate them in Theorem 1.1'.

Next we show the double integration over a curve vanishes. In other words, a curve is too thin to support a volume.

Theorem 1.7. *Let f be a bounded function on a curve C. Then $\tilde{f}$ is integrable and*

$$\iint_C f\,dA = 0 \ .$$

Proof. Let us take $R = [a, b] \times [c, d]$ and consider the special case that $C = \{(x, \varphi(x)\}$ is the graph of a continuous function in $[a, b]$ in R. By Theorem 1.6, $\tilde{f}$ is integrable in R. We have

$$\begin{aligned} \iint_C f\, dA &= \iint_R \tilde{f}\, dA \quad \text{(by definition)} \\ &= \int_a^b \int_c^d \tilde{f}\, dydx \quad \text{(Fubini's)} \\ &= \int_a^b \int_c^{\varphi(x)} \tilde{f}(x,y)\, dydx + \int_a^b \int_{\varphi(x)}^d \tilde{f}(x,y)\, dydx \\ &= 0\ , \end{aligned}$$

since $\tilde{f}(x, y) = 0$ for all y in $[c, \varphi(x))$ and $(\varphi(x), d]$. □

We will associate a set with a function. In this way, sets can be manipulated as functions.

Let E be a nonempty set in $\mathbb{R}^2$ (actually it could be defined in $\mathbb{R}^n$ for any n.) Its *characteristic function* χ_E is defined to be $\chi_E(x, y) = 1\ , \quad (x, y) \in E$, and $\chi_E(x, y) = 0$ otherwise. Also set $\chi_\phi \equiv 0$. We point out the following relations:

- $\chi_{A\cup B} = \chi_A + \chi_B - \chi_{A\cap B}$.
- $\chi_{A\cap B} = \chi_A \cdot \chi_B$.
- $\chi_A \le \chi_B$ if and only if $A \subset B$.

Combining the first two, we have

$$\chi_{A\cup B} = \chi_A + \chi_B - \chi_A \cdot \chi_B\ .$$

We are ready to prove the following frequently used result.

Theorem 1.8. *Divide the region D by a piecewise C^1-curve C to obtain two regions D_1 and D_2. For any integrable function f in D, f is also integrable in $D_i, i = 1, 2$. Moreover,*

$$\iint_D f\, dA = \iint_{D_1} f\, dA + \iint_{D_2} f\, dA\ .$$

Proof. Since the boundary of D_i are composed of piecewise C^1-curves, according to Theorem 5, the characteristic functions χ_{D_i} are integrable, so $f\chi_{D_i}$, as product of two integrable functions, is also integrable. From $C = D_1 \cap D_2$ and $\chi_D = \chi_{D_1} + \chi_{D_2} - \chi_{D_1\cap D_2}$, we have $\chi_D = \chi_{D_1} + \chi_{D_2} - \chi_C$. Let R be a rectangle containing D in its interior. We have

$$\begin{aligned} \iint_D f\, dA &= \iint_R \tilde{f}\, dA \\ &= \iint_R \tilde{f}\chi_{D_1}\, dA + \iint_R \tilde{f}\chi_{D_2} - \iint_R \tilde{f}\chi_C\ . \end{aligned}$$

The function $\tilde{f}\chi_{D_1}$ is equal to f in D_1 and 0 outside D_1. Therefore, it is the extension of f from D_1, that is,

$$\iint_R \tilde{f}\chi_{D_1}\, dA = \iint_{D_1} f\, dA \ .$$

Similarly, we have

$$\iint_R \tilde{f}\chi_{D_2}\, dA = \iint_{D_2} f\, dA \ ,$$

and

$$\iint_R \tilde{f}\chi_{C}\, dA = \iint_{C} f\, dA \ .$$

Thus,

$$\iint_D f\, dA = \iint_{D_1} f\, dA + \iint_{D_2} f\, dA - \iint_C f\, dA \ ,$$

and the desired formula holds as the last term vanishes according to Theorem 1.7. □

Now we come to evaluation of a double integral in a region. We have discussed how to do it in a rectangle. We will work on two types of special regions. Type I is of the form

$$\{(x,y):\ f_1(x) \le y \le f_2(x),\ a \le x \le b\}\ , \quad f_i,\ i=1,2,\ \text{is continuous}\ ,$$

and Type II is

$$\{(x,y):\ g_1(y) \le x \le g_2(y),\ c \le y \le d\}\ , \quad g_i,\ i=1,2,\ \text{is continuous}\ .$$

More complicated regions could be decomposed to a union of Type I and Type II regions, with the help from Theorem 1.8.

Theorem 1.9. ***(Fubini's Theorem)***

(a) Let D be a Type I region. For a continuous function f in D,

$$\iint_D f(x,y)\, dA = \int_a^b \int_{f_1(x)}^{f_2(x)} f(x,y)\, dydx \ .$$

(b) Let D be a Type II region. For a continuous function F in D,

$$\iint_D f(x,y)\, dA = \int_c^d \int_{g_1(y)}^{g_2(y)} f(x,y)\, dxdy \ .$$

Proof. We prove (a) only. Let $R = [a, b] \times [c, d]$ be a rectangle containing the Type I region D. By Theorem 1.4, Fubini's Theorem on a rectangle,

$$\begin{aligned}
\iint_D f(x,y)\,dA &= \iint_R \tilde{f}(x,y)\,dA \\
&= \int_a^b \int_c^d \tilde{f}(x,y)\,dA \\
&= \int_a^b \left(\int_c^{f_1(x)} \tilde{f}(x,y)\,dA + \int_{f_1(x)}^{f_2(x)} \tilde{f}(x,y)\,dA + \int_{f_2(x)}^{d} \tilde{f}(x,y)\,dA \right) \\
&= \int_a^b \int_{f_1(x)}^{f_2(x)} f(x,y)\,dA.
\end{aligned}$$

□

Example 1.3 Evaluate

$$\iint_D (2y+1)\,dA \ ,$$

where D is the region bounded by $y = 2x$ and $y = x^2$.

The curves of $y = 2x$ and $y = x^2$ intersect at $(0, 0)$ and $(0, 2)$. The region of integration is expressed as

$$D = \{(x,y) : \ x^2 \le y \le 2x, \ x \in [0,2] \ \} \ .$$

By Fubini's Theorem,

$$\begin{aligned}
\iint_D (2y+1)\,dA &= \int_0^2 \int_{x^2}^{2x} (2y+1)\,du\,dx \\
&= \int_0^2 (y^2+y)\Big|_{x^2}^{2x} dx \\
&= \frac{28}{5} \ .
\end{aligned}$$

The region D can also be expressed as

$$D = \{(x,y) : \ \frac{y}{2} \le x \le \sqrt{y}, \ y \in [0,4]\}.$$

We have

$$\begin{aligned}
\iint_D (2y+1)\,dA &= \int_0^4 \int_{y/2}^{\sqrt{y}} (2y+1)\,dx\,dy \\
&= \int_0^4 (2y+1) \int_{y/2}^{\sqrt{y}} dx\,dy \\
&= \int_0^4 (2y+1)(\sqrt{y} - \frac{y}{2})\,dy \\
&= \frac{28}{5} \ .
\end{aligned}$$

Example 1.4 Evaluate the iterated integral

$$\int_0^1 \int_y^1 \frac{\sin x}{x}\, dxdy \ .$$

It is hard to integrate $\sin x/x$, so we switch the order of integration. First, recognize this iterated integral is equal to the double integral

$$\iint_D \frac{\sin x}{x}\, dA \ ,$$

where D is the triangle bounded between $y = 0, y = x$ for $x \in [0, 1]$. By Fubini's Theorem,

$$\begin{aligned} \int_0^1 \int_y^1 \frac{\sin x}{x}\, dxdy &= \iint_D \frac{\sin x}{x}\, dA \\ &= \int_0^1 \int_0^x \frac{\sin x}{x}\, dydx \\ &= \int \frac{\sin x}{x} \times x\, dx \\ &= \int_0^1 \sin x\, dx \\ &= 1 - \cos 1 \ . \end{aligned}$$

Example 1.5 Evaluate the double integral

$$\iint_D x\, dA \ ,$$

where D is the region bounded by $y = 0, x + y = 0$, and the unit circle on the half plane $x \geq 0$.

The line $x + y = 0$ intersection the circle $x^2 + y^2 = 1$ at $(\sqrt{2}/2, -\sqrt{2}/2)$, so D can be described as

$$D = \{(x, y) : \ -y \leq x \leq \sqrt{1 - y^2}, \ y \in [-\sqrt{2}/2, 0]\} \ .$$

Hence,

$$\begin{aligned} \iint_D x\, dA &= \int_{-\sqrt{2}/2}^0 \int_{-y}^{\sqrt{1-y^2}} x\, dxdy \\ &= \frac{1}{2} \int_{-\sqrt{2}/2}^0 (1 - y^2 - y)\, dy \\ &= \frac{\sqrt{2}}{6} \ . \end{aligned}$$

If one insists to integrate in y first, we observe that D can be expression as the union of D_1 and D_2:

$$D_1 = \{(x,y) : \ 0 \le y \le -x, \ x \in [0, \sqrt{2}/2]\}$$

and

$$D_2 = \{(x,y) : \ -\sqrt{1-x^2} \le y \le 0, \ x \in [\sqrt{2}/2, 1]\} \ .$$

We have

$$\begin{aligned}
\iint_D x\, dA &= \iint_{D_1} x\, dA + \iint_{D_2} x\, dA \\
&= \int_0^{\sqrt{2}/2} \int_{-x}^{0} x\, dydx + \int_{\sqrt{2}/2}^{1} \int_{-\sqrt{1-x^2}}^{0} x\, dydx \\
&= \int_0^{\sqrt{2}/2} x^2\, dx + \int_{\sqrt{2}/2}^{1} x\sqrt{1-x^2}\, dx \\
&= \frac{\sqrt{2}}{6} \ .
\end{aligned}$$

Example 1.6 Find the area of the region which is bounded between $y = x^2 - 4, y = x^2 - 1$ and $x \ge 0, y \le 0$.

After sketching the figure, we see that the area of this region D is given by

$$\iint_D 1\, dA = \int_0^1 \int_{x^2-4}^{x^2-1} dydx + \int_1^2 \int_{x^2-4}^{0} dydx \ .$$

A straightforward calculation yields

$$\iint_D 1\, dA = \frac{14}{3} \ .$$

As an application of what has been developed, we introduce a definition of the area of a set. Let E be a nonempty set in $\mathbb{R}^2$ (actually in $\mathbb{R}^n$). E is called *rectifiable* if χ_E is integrable. For a rectifiable set E, its area is given by

$$|E| = \iint_E 1\, dA = \iint_R \chi_E\, dA \ , \ E \subset R.$$

We know that every region is rectifiable since its boundary is composed of piecewise C^1-curves. An interesting property is the Euclidean invariant of area. Any Euclidean motion is a composition of translation, rotation and reflection with respect to the x- and y-axes. One can show that the area of a rectifiable set is invariant under any Euclidean motion. This looks like an obvious fact, but can you prove it?

By introducing curves on a region D, D can be decomposed into a union of subregions D_k whose interiors are mutually disjoint. We may call $\{D_k\}$ a *generalized partition* on D. (A partition divides the rectangle into subrectangles $R_{ij}, i = 1, \cdots, n, j = 1, \cdots, m$, but now we cannot use i, j as indices, so we use a single index instead.) Choosing a tag point p_k from each D_k we can form a *generalized Riemann sum* $R(f, P) = \sum_k f(p_k)|D_k|$ for any bounded function f in D. Denote by $\|P\|$ the maximum among all diameters of D_k's. In case the region is a rectangle R, the diameter of the subrectangle R_{ij} is $\sqrt{\Delta x_i^2 + \Delta y_j^2}$, hence the norm $\|P\|$ is small if and only if the maximum of all diameters are small. We see that in measuring smallness, the norm defined here is equivalent to the one defined before.

Theorem 1.10. *Let f be continuous in a region D and let P be a generalized partition in D. For $\varepsilon > 0$, there is some $\delta > 0$ such that*

$$\left| R(f, P) - \iint_D f \, dA \right| < \varepsilon \ , \quad \forall P, \ \|P\| < \delta \ .$$

We will use the following fact: Let M and m be the maximum/minimum of a continuous f in D. Then for any $\alpha \in [m, M]$, there is some $z \in D$ such that $f(z) = \alpha$. It sounds quite natural. For, let $m = f(p)$ and $M = f(q)$ where p, q are two points in D. We connect p to q by a continuous curve C in D. As we go along C from p to q, the values of f changes continuously from m to M. Since f is continuous and α lies between m and M, there must a point z on C such that $f(z) = \alpha$.

Proof. By the Uniform Continuity Theorem (which continues to hold on a region), given $\varepsilon' > 0$, there is some δ such that $|f(x, y) - f(x', y')| < \varepsilon'$ whenever (x, y) and (x', y') are two points in D whose distance is less than δ. We will take $\varepsilon' = \varepsilon/|D|$ where $\varepsilon > 0$ is given.

Now, let m_k and M_k be the minimum and maximum of f over D_k. From $m_k \leq f \leq M_k$ in D_k, we integrate over D_k to get

$$m_k|D_k| \leq \iint_{D_k} f \, dA \leq M_k|D_k| \ .$$

By what we have said above, there is some $f(p_k^*) \in D_k$ such that

$$f(p_k^*) = \frac{1}{|D_k|} \iint_{D_k} f \, dA \ .$$

Therefore, for any Riemann sum $R(f,P)=\sum_k f(q_k^*)|D_k|$ with $\|P\|<\delta$, we have

$$\begin{aligned} & \left|\sum_k f(q_k^*)|D_k| - \iint_D f\,dA\right| \\ = & \left|\sum_k f(q_k^*)|D_k| - \sum_k f(p_k^*)|D_k|\right| \\ = & \left|\sum_k (f(q_k^*) - f(p_k^*))|D_k|\right| \\ < & \sum_k \frac{\varepsilon}{|D|}|D_k| \\ = & \frac{\varepsilon}{|D|}|D| = \varepsilon\ , \end{aligned}$$

and the desired result follows. □

We also have

Theorem 1.11. *Let f and g be two continuous functions in the region D. Let p_k and q_k be tag points for the generalised partition P. Then*

$$\lim_{\|P\|\to 0} \sum_k f(p_k)g(q_k)|D_k| = \iint_D fg\,dA\ .$$

Note that here the functions f and g take different tag points.

Proof. We need to show that for any $\varepsilon>0$, there is some δ such that

$$\left|\sum_k f(p_k)g(q_k)|D_k| - \iint_D fg\,dA\right| < \varepsilon\ , \quad \forall P, \|P\|<\delta.$$

Since fg is continuous so integrable in D, we can find δ_1 such that

$$\left|\sum_k f(p_k)g(p_k)|D_k| - \iint_D fg\,dA\right| < \frac{\varepsilon}{2}\ , \quad \forall P, \|P\|<\delta_1.$$

On the other hand, by the Uniform Continuity Theorem, there is some δ_2 such that

$$|g(p)-g(q)| < \frac{\varepsilon}{2M|D|}\ , \quad |p-q|<\delta_2\ ,$$

where M is a bounded on $|f|$. Therefore, for P satisfying $\|P\| < \delta = \min\{\delta_1, \delta_2\}$,

$$
\begin{aligned}
& \left| \sum_k f(p_k) g(q_k) |D_k| - \iint_D fg \, dA \right| \\
\le & \left| \sum_k f(p_k)(g(q_k) - g(p_k)) |D_k| \right| + \left| \sum_k f(p_k) g(p_k) |D_k| - \iint_D fg \, dA \right| \\
\le & \frac{\varepsilon}{2M|D|} \times M|D| + \frac{\varepsilon}{2} \\
< & \ \varepsilon \ ,
\end{aligned}
$$

done. $\square$

1.5 Double Integral in the Polar Coordinates

Each point in the plane (x, y) (except $(0, 0)$) can be expressed as $x = r\cos\theta$ and $y = r\sin\theta$ for a *unique* pair (r, θ), $r > 0$, $\theta \in [0, 2\pi)$. (r, θ) is called the *polar coordinates* of (x, y). Let Φ be the map $\Phi(r, \theta) = (r\cos\theta, r\sin\theta)$. It maps the strip $[0, \infty) \times [0, 2\pi]$ onto $\mathbb{R}^2$ and is one-to-one from $(0, \infty) \times [0, 2\pi)$ onto $\mathbb{R}^2 \setminus \{(0, 0)\}$. Alternatively, it maps $[0, \infty) \times [-\pi, \pi]$ onto $\mathbb{R}^2$ and is one-to-one from $(0, \infty) \times (-\pi, \pi]$ onto $\mathbb{R}^2 \setminus \{(0, 0)\}$. A curve expressed in polar coordinates could look very different from its form in rectangular coordinates. Here are some examples.

- The horizontal line $y = c \neq 0$ becomes $r = c/\sin\theta$, $\theta \in (0, \pi)$ in polar coordinates.
- The circle $x^2 + y^2 = a^2$ becomes $r = a$.
- The circle $(x - a/2)^2 + y^2 = a^2/4$ becomes $r = a\cos\theta$, $\theta \in [-\pi/2, \pi/2]$.
- The parabola $y = a^2 - x^2$ becomes

  $$r(\theta) = \frac{-\sin\theta + \sqrt{\sin^2\theta + 4a^2\cos^2\theta}}{2\cos^2\theta} \ , \quad \theta \in [0, 2\pi], \ \theta \neq 3\pi/2 \ .$$

  Note that the ray at $\theta = 3\pi/2$ does not hit the parabola.

Sometimes a curve is simpler when expressed in polar coordinates. For instance, the cardioid is

$$r = 1 + a\cos\theta, \ \theta \in [0, 2\pi],$$

where $a \in (0, 1]$. To express it in rectangular coordinates, we proceed as follows. First, multiple the equation by r to get

$$x^2 + y^2 = \sqrt{x^2 + y^2} + ax \ .$$

Then move ax to the left and square to get

$$(x^2 + y^2 - ax)^2 = x^2 + y^2 \ .$$

In the rectangular coordinates, the cardioid is a quartic equation.

A rectangle $R = [r_1, r_2] \times [\theta_1, \theta_2]$ in the (r, θ)-plane is mapped under Φ to the region

$$S = \{(x, y) : \ x = r\cos\theta, y = r\sin\theta, \ r_1 \le r \le r_2, \ \theta_1 \le \theta \le \theta_2, \ \theta_1, \theta_2 \in [0, 2\pi)\} \ .$$

Any partition P on R introduces a generalized partition on S via Φ. Denote its subregions by $S_{ij} = \Phi(R_{ij})$.

Theorem 1.12. *Let f be a bounded function which is continuous in S except along some piecewise C^1-curves. Then*

$$\iint_S f(x, y)\, dA = \iint_R f(r\cos\theta, r\sin\theta) r\, dr d\theta \ .$$

Proof. Let us assume f is continuous in S. The area of S_{ij} is given by

$$\frac{1}{2} r_{i+1}^2 \Delta\theta_j - \frac{1}{2} r_i^2 \Delta\theta_j = \frac{1}{2}(r_{i+1} + r_i)\Delta r_i \Delta\theta_j \ .$$

Let P be a partition on R with tags τ_{ij} tags for R_{ij}. Then $p_{ij} = \Phi(\tau_{ij})$ is a tag for S_{ij}. When $\|P\|$ is small, the generalized partition $S_{ij} = \Phi(R_{ij})$ is also small in norm. We have

$$\iint_S f(x, y)\, dA \approx \sum_{i,j} f(p_{ij})|S_{ij}| = \sum_{i,j} f(\Phi(\tau_{ij})) r_i^* \Delta r_i \Delta\theta_j \ ,$$

where $r_i^* = (r_i + r_{i+1})/2$. This sum can be viewed as $\sum_{i,j} h(\tau_{ij}) g(r_i^*, \theta_j^*)$ where $h(r, \theta) = f(\Phi(r, \theta))$ and $g(r, \theta) = r$. We choose some θ_j^* to make (r_i^*, θ_j^*) a tag point in R_{ij}. Since $f \circ \Phi$ is continuous in R, By Theorem 1.11, as $\|P\| \to 0$,

$$\begin{aligned} \sum_{i,j} f(\Phi(\tau_{ij})) r_i^* \Delta r_i \Delta\theta_j &= \sum_{i,j} h(\tau_{ij}) g(r_i^*, \theta_j^*) \Delta r_i \Delta\theta_j \\ &\to \iint_R h(r, \theta) g(r, \theta)\, dA(r, \theta) \\ &= \iint_R f(\Phi(r, \theta)) r\, dA(r, \theta) \ . \end{aligned}$$

On the other hand,

$$\sum_{i,j} f(\Phi(\tau_{ij})) r_i^* \Delta r_i \Delta\theta_j = \sum_{i,j} f(p_{ij})|S_{ij}| \to \iint_S f(x, y)\, dA \ ,$$

the theorem holds. $\square$

When a region D is expressed as

$$\{(x, y) : \ x = r\cos\theta, y = r\sin\theta, \ \varphi_1(\theta) \le r \le \varphi_2(\theta), \ \theta_1 \le \theta \le \theta_2, \ \theta_1, \theta_2 \in [0, 2\pi)\} .$$

We have

Theorem 1.13. *For a continuous function f in D,*

$$\iint_D f(x, y)\, dA(x, y) = \int_{\theta_1}^{\theta_2} \int_{\varphi_1(\theta)}^{\varphi_2(\theta)} f(r\cos\theta, r\sin\theta) r\, dr d\theta .$$

Proof. Pick $r_1 < r_2$ so that the sector S formed by r_1, r_2, θ_1, and θ_2 contains D. Let $\tilde{D}$ be the preimage of D under Φ. Then $\tilde{D}$ is of the form

$$\{(r, \theta) : \varphi_1(\theta) \le r \le \varphi_2(\theta), \ \theta_1 \le \theta \le \theta_2\}$$

and is contained in the rectangle $[r_1, r_2] \times [\theta_1, \theta_2]$. Let $\tilde{f}$ be the usual extension of f being zero outside D. We have

$$\begin{aligned}
\iint_D f(x, y)\, dA(x, y) &= \iint_S \tilde{f}(x, y)\, dA(x, y) \\
&= \iint_R \tilde{f}(r\cos\theta, r\sin\theta) r\, dA(r, \theta) \\
&= \int_{\theta_1}^{\theta_2} \int_{r_1}^{r_2} \tilde{f}(r\cos\theta, r\sin\theta) r dr \theta \\
&= \int_{\theta_1}^{\theta_2} \left(\int_{r_1}^{\varphi_1(\theta)} + \int_{\varphi_1(\theta)}^{\varphi_2(\theta)} + \int_{\varphi_2(\theta)}^{r_2} \right) \tilde{f}(r\cos\theta, r\sin\theta) r dr d\theta \\
&= \int_{\theta_1}^{\theta_2} \int_{\varphi_1(\theta)}^{\varphi_2(\theta)} f(r\cos\theta, r\sin\theta) r dr d\theta .
\end{aligned}$$

□

Example 1.6 Find the area of the lemniscate $r^2 = 4\cos 2\theta$.

Always sketch the figure before integrating. The lemniscate is a two-leaves like figure symmetric with respect to both axes. For $\theta \in [0, 2\pi], 2\theta \in [0, 4\pi]$. We see that $\cos 2\theta$ is nonnegative on the intervals $[0, \pi/4], [3\pi/4, \pi], [\pi, 5\pi/4], [7\pi/4, 2\pi]$ only. By symmetry it suffices to integrate over the range $\theta \in [0, \pi/4]$. Any ray emitting from the origin with $\theta \in [0, \pi/4]$ hits the lemniscate at one point. Hence the area of the lemniscate is given by

$$\begin{aligned}
\iint_D dA &= 4 \int_0^{\pi/4} \int_0^{(4\cos 2\theta)^{1/2}} r\, dr d\theta \\
&= 4 \int_0^{\pi/4} \frac{1}{2} \times 4\cos 2\theta\, d\theta \\
&= 4 .
\end{aligned}$$

Example 1.7 Evaluate the iterated integral

$$\int_1^2 \int_0^{\sqrt{2x-x^2}} y\,dydx\ .$$

We use polar coordinates to evaluate this integral. First of all, the graph of $y = \sqrt{2x - x^2}$ is the circle of radius 1 at $(1,0)$. The region of integration is given by

$$G = \{(x,y):\ 0 \le y \le \sqrt{2x-x^2}\ ,\ x \in [1,2]\}\ .$$

To express it in polar coordinates, observe that every ray with $\theta \in [0, \pi/4]$ first hits the vertical line $x = 1$ and then the circle. A ray out of this range does not hit G. In polar coordinates, $x = 1$ is given by $r = 1/\cos\theta$ and $y = \sqrt{2x - x^2}$ becomes $r = 2\cos\theta$. Therefore,

$$\begin{aligned}
\int_1^2 \int_0^{\sqrt{2x-x^2}} y\,dydx &= \iint_G y\,dA \\
&= \int_0^{\pi/4} \int_{1/\cos\theta}^{2\cos\theta} r\sin\theta\ r\,drd\theta \\
&= \int_0^{\pi/4} \frac{1}{3}\left(8\cos^3\theta - \frac{1}{\cos^3\theta}\right)\sin\theta\,d\theta \\
&= \frac{1}{2} - \frac{1}{6} \\
&= \frac{1}{3}\ .
\end{aligned}$$

Example 1.8 Find the area pinched between the curves $r = 3/2$ and $r = 1 + \cos\theta$.

The circle $r = 3/2$ and the cardioid $r = 1+\cos\theta$ intersect $1+\cos\theta = 3/2$ at $\theta = \pm\pi/3$. When $\theta \in [-\pi/3, \pi/3]$, the cardioid lies on outside and the circle inside. When $\theta \in [\pi/3, \pi]$ or $[-\pi, -\pi/3]$, the circle lies outside and the cardioid inside. By symmetry, it suffices to calculate things in the first and the second quadrants. We have

$$\begin{aligned}
\frac{1}{2}\ \text{Area} &= \int_0^{\pi/3}\int_{3/2}^{1+\cos\theta} r\,drd\theta + \int_{\pi/3}^{\pi}\int_{1+\cos\theta}^{3/2} r\,drd\theta \\
&= \frac{1}{2}\int_0^{\pi/3}\Big(-\frac{3}{4} + 2\cos\theta + \frac{1}{2}\cos 2\theta\Big)\,d\theta \\
&\quad +\frac{1}{2}\int_{\pi/3}^{\pi}\Big(\frac{3}{4} - 2\cos\theta - \frac{1}{2}\cos 2\theta\Big)\,d\theta \\
&= \frac{\pi}{8} + \frac{5\sqrt{3}}{4}\ .
\end{aligned}$$

Hence the area is given by $(\pi + 10\sqrt{3})/4$.

Example 1.9. Express the integral

$$\int_0^{\sqrt{5}} \int_{x^2}^5 f(x,y)\, dydx$$

in polar coordinates.

Well, the region is the one sitting in the first quadrant bounded by the y-axis, horizontal line $y=5$ and the parabola $y=x^2$. The latter two curves intersect at $(\sqrt{5},5)$ and $(-\sqrt{5},5)$. Any ray from $\theta \in [0,\alpha], \alpha = \tan^{-1}\sqrt{5}/5$, hits the parabola once. On the other hand, any ray from $\theta \in [\alpha, \pi/2]$ hits the horizontal line $y=5$ once. We have

$$\begin{aligned} & \int_0^{\sqrt{5}} \int_{x^2}^5 f(x,y)\, dydx \\ = \; & \int_0^{\alpha} \int_0^{\sin\theta/\cos^2\theta} f(r\cos\theta, r\sin\theta) r\, drd\theta + \int_{\alpha}^{\pi/2} \int_0^{5/\sin\theta} f(r\cos\theta, r\sin\theta) r\, drd\theta \; . \end{aligned}$$

1.6 Improper Integral

In Riemann integrals the functions under consideration are always bounded and the regions of integration are bounded. In practise we need to consider some situations either the functions or the regions are unbounded. In this section we consider two typical situations. First, the function becomes infinity at a point (point singularity). Second, unbounded regions.

Let D be a bounded region and f a function in D which is continuous everywhere except at a point $p_0 = (x_0, y_0)$ and $f(x,y)$ becomes positive or negative as $(x,y) \to (x_0,y_0)$. We say the *improper integral* of f over D exists if

$$\lim_{a \to 0} \iint_{D \setminus D_a} f(x,y)\, dA$$

exists, where D_a is the disk of radius a centered at p_0. When it holds, let

$$\iint_D f\, dA = \lim_{a\to 0} \iint_{D\setminus D_a} f(x,y)\, dA \; . \tag{1.2}$$

We use the same notation to denote the improper integral whenever it exists.

Example 1.9 Determine the range of α such the improper integral

$$\iint_D (x^2+y^2)^{\alpha}\, dA \; ,$$

where D is any region containing the origin.

When $\alpha \geq 0$, the integrand is continuous. We do not have to consider its improper integrability. So we always assume $\alpha < 0$. It is also clear it suffices to take D to be the disk of radius 1 at the origin. Introducing polar coordinates, when $2\alpha \neq -1$,

$$\begin{aligned}\iint_{D\setminus D_a} (x^2+y^2)^\alpha \, dA &= \int_0^{2\pi}\int_a^1 r^{2\alpha} r \, dr d\theta \\ &= \frac{2\pi}{2\alpha+2}\left(1-a^{2\alpha+2}\right) \\ &\rightarrow \frac{\pi}{\alpha+1} \ ,\end{aligned}$$

if and only if $\alpha + 1 > 0$. Hence the improper integral exists for $\alpha \in (-1, 0)$. When $\alpha = -1$, we have instead

$$\iint_{D\setminus D_a} (x^2+y^2)^\alpha \, dA = 2\pi|\log a| \rightarrow \infty \ ,$$

as $a \rightarrow 0$. The improper integral does not exist when $\alpha = -1$.

Next, when the region D is unbounded, we call the improper integral exists if

$$\iint_D f \, dA = \lim_{a\to\infty} \iint_{D\cap D_a} f \, dA \ . \tag{1.3}$$

We consider an interesting case.

Example 1.10 Evaluate

$$\int_{-\infty}^{\infty} e^{-x^2} \, dx \ .$$

This is an improper integral of a single variable. The trick is to make it a double integral. We have

$$\iint_{D_a} e^{-x^2-y^2} \, dA = \int_0^{2\pi}\int_0^a e^{-r^2} r \, dr d\theta = \pi(1-e^{-a^2}) \rightarrow \pi \ ,$$

as $a \rightarrow \infty$. It follows that the improper integral

$$\iint_{\mathbb{R}^2} e^{-x^2-y^2} \, dA$$

exists and is equal to π. Let R_a be the square with side length $2a$ at the origin. Using $D_a \subset R_a \subset D_{\sqrt{2}a}$, we see that

$$\lim_{a\to\infty} \iint_{R_a} e^{-x^2-y^2} \, dA = \lim_{a\to\infty} \iint_{D_a} e^{-x^2-y^2} \, dA = \pi \ .$$

Now,

$$\begin{aligned}
\int_{-a}^{a} e^{-x^2}\,dx \times \int_{-a}^{a} e^{-y^2}\,dy &= \int_{-a}^{a}\int_{-a}^{a} e^{-x^2+y^2}\,dydx \\
&= \iint_{R_a} e^{-x^2-y^2}\,dA \\
&\to \pi\ .
\end{aligned}$$

We conclude that

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}\ . \tag{1.4}$$

1.7 Triple Integrals

The theory of triple integrals is essentially the same as the double integral. It suffices to point out that a region in space is bounded by one or several closed surfaces, each of which are composed of pieces of C^1-surfaces meet along some C^1-curves. We will not give a precise definition here, but the concept is clear in a intuitive way. Let us look at some examples:

- The sphere $\{(x,y,z):\ (x-1)^2+(y-b)^2+(z-c)^2=r^2\}$ is a C^1-surface with center (a,b,c) and radius r. The region bounded by the sphere is a ball.
- The rectangular box is the region bounded by the planes $x=a,b, y=c,d, z=e,f$. Its boundary is composed by six pieces of C^1-surfaces (rectangles in fact) meeting along line segments.
- The circular cone $\{(x,y,z) : z=\sqrt{x^2+y^2}\}$ is an unbounded surface which has a sharp corner at the origin. We could truncate it to get a bounded one $\{(x,y,z) : z=\sqrt{x^2+y^2},\ z=h\}$ to get a region bounded by two surfaces, one being the circular cone and the other the plane $z=h$
- The torus obtained by rotating the circle $(y-a)^2+z^2=b^2, a<b$, around the z-axis. It is a C^1-surface which bounds a region.

Parallel to the double integral, a rectangular box B is given by $[a,b]\times[c,d]\times[e,f]$ and a partition P on B is the collection of points

$$a=x_0<x_1<\cdots<x_n=b,\quad c=y_0<y_1<\cdots<y_m=d,\quad e=z_0<z_1<\cdots<z_l=f\ .$$

The partition P divides B into subrectangular boxes $B_{ijk}=[x_i,x_{i+1}]\times[y_j,y_{j+1}]\times[z_k,z_{k+1}]$. For a bounded function f in B, its Riemann sum is given by

$$R(f,P)=\sum_{i,j,k} f(p_{ijk})|B_{ijk}|,\ ,$$

where $|B_{ijk}| = \Delta x_i \Delta y_j \Delta z_k$. The function f is called integrable if there is a number α such that for every $\varepsilon > 0$, there is some $\delta > 0$ such that

$$|\sum_{i,j,k} f(p_{ijk})|B_{ijk}| - \alpha| < \varepsilon \ , \quad \forall P,\ \|P\| < \delta \ ,$$

where $\|P\|$ is the maximum among all $\Delta x_i, \Delta y_j, \Delta z_k$. The integral α will be denoted by

$$\iiint_B f\, dV, \quad \text{or} \iiint_B f(x,y,z)\, dV, \quad \text{or} \iiint_B f(x,y,z)\, dV(x,y,z) \ .$$

The analog of Theorems 1.1, 1.2, and 1.3 hold for triple integrals, and I trust you to formulate them. We also have

Theorem 1.14. ***(Fubini's Theorem)*** *Let f be a continuous function in a rectangular box B. Then*

$$\begin{aligned} \iiint_R f\, dV &= \iint_R \int_e^f f(x,y,z)\, dz\, dA(x,y) \quad (R = [a,b] \times [c,d]) \\ &= \int_a^b \int_c^d \int_e^f f(x,y,z)\, dz\, dA(x,y) \ . \end{aligned}$$

This theorem still holds for bounded functions that are continuous everywhere except at some surfaces, curves or points in B. No new ideas are involved in the proof. Let us sketch the proof. The triple integral can be approximated by Riemann sums. Taking tags of the form (x_i^*, y_j^*, z_k^*), we have

$$\iiint_R f\, dV \approx \sum_{i,j,k} f(x_i^*, y_j^*, z_k^*) \Delta x_i \Delta y_j \Delta z_k = \sum_{i,j} \left(\sum_k f(x_i^*, y_j^*, z_k^*) \Delta z_k \right) \Delta x_i \Delta y_j \ .$$

When $\|P\|$ is very small, $\Delta x_i, \Delta y_j, \Delta z_k$ are also very small,

$$\begin{aligned} \sum_{i,j} \left(\sum_k f(x_i^*, y_j^*, z_k^*) \Delta z_k \right) \Delta x_i \Delta y_j &\approx \sum_{i,j} \int_e^f f(x_i^*, y^* - j, z)\, dz\ \Delta x_i \Delta y_j \\ &\approx \iint_R \left(\int_e^f f(x,y,z) dz \right) dA(x,y) \ . \end{aligned}$$

A similar result holds when the order of x, y and z are interchanged.

This formula reduces the evaluation of a triple integral to a single integral and a double integrable. A further application reduces the double integral to two single integrals.

For functions defined in a region Ω in space, we take a rectangular box B containing Ω and define

$$\iiint_\Omega f(x,y,z)\,dV = \iiint_B \tilde{f}(x,y,z)\,dV \ ,$$

where $\tilde{f}$ is the trivial extension of f to the entire space (that is, setting $\tilde{f}=0$ outside Ω).

For a region of the form

$$\Omega = \{(x,y,z):\ f_1(x,y) \le z \le f_2(x,y),\ (x,y)\in D\} \ ,$$

where D is a region in the plane, Fubini's theorem becomes

$$\iiint_\Omega f(x,y,z)\,dV = \iint_D \int_{f_1(x,y)}^{f_2(x,y)} f(x,y,z)\,dz\,dA(x,y) \ . \tag{1.5}$$

Corresponding formulas when the role of $z=f_i(x,y)$ is replaced by $y=g_i(x,z)$ or $x=h_i(y,z), i=1,2$, hold.

When f is positive, the triple integral

$$\iiint_\Omega f\,dV$$

gives the mass of Ω with density f. When $f\equiv 1$,

$$|\Omega| \equiv \iiint_\Omega dV$$

is the volume of the region Ω.

Example 1.11 Evaluate

$$\iiint_\Omega xy\,dV$$

in two ways: $dzdA(x,y)$ and $dxdA(y,z)$ where Ω is the region bounded between $x+2y+3z=1$ and the coordinate planes in $x,y,z\ge 0$.

The region Ω is given by

$$\Omega = \{(x,y,z):\ 0\le z\le (1-x-2y)/3,\ (x,y)\in D\} \ ,$$

where D is the triangle with vertices at $(0,0),(1,0),(0,1/2)$ in the xy-plane. By Fubini's Theorem,

$$\begin{aligned}
\iiint_\Omega xy\,dV &= \iint_D \int_0^{(1-x-2y)/3} xy\,dz\,dA(x,y)\\
&= \frac{1}{3}\iint_D xy(1-x-2y)\,dA\\
&= \frac{1}{3}\int_0^1\int_0^{(1-x)/2} xy\,dydx\\
&= \frac{1}{144}\ .
\end{aligned}$$

Next, Ω projects to the triangle Δ with vertices at $(0,0),(1/2,0),(0,1/3)$ in the yz-plane. We have

$$\Omega = \{(x,y,z):\ 0 \le x \le 1-2y-3z,\ (y,z)\in\Delta\}\ .$$

$$\begin{aligned}
\iiint_\Omega xy\,dV &= \iint_\Delta \int_0^{1-2y-3z} xy\,dx\,dA(y,z) \\
&= \iint_\Delta \frac{1}{2}(1-2y-3z)^2\,dA \\
&= \int_0^{1/2}\int_0^{(1-2y)/3} \frac{1}{2}(1-2y-3z)^2\,dz\,dy \\
&= \frac{1}{144}\ .
\end{aligned}$$

Example 1.11' Express the triple integral of a function f over the tetrahedron formed by the vertices $(0,0,0),(0,1,0),(1,1,0)$ and $(0,1,1)$ by an iterated integral in $dzdydx$.

The tetrahedron T has four faces given by triangles lying in the xy-plane, yz-plane, the plane $y=1$ and the plane $x-y+z=0$. See the figure in pg 912, Text. When projecting onto the xy-plane, it is described by $f_1(x,y)\equiv 0 \le z \le f_2(x,y)\equiv y-z$ over the triangle Δ with vertices at $(0,0),(0,1)$ and $(1,1)$. Therefore,

$$\iiint_T f(x,y,z)\,dV = \iint_\Delta \int_0^{y-z} f(x,y,z)\,dzdA(x,y) = \int_0^1\int_0^y\int_0^{y-x} f(x,y,z)\,dzdydx\ .$$

We may also express the triple integral in other orders. For instance, we have

$$\iiint_T f(x,y,z)\,dV = \int_0^1\int_0^y\int_0^{y-z} f(x,y,z)\,dxdzdy\ ,$$

and

$$\iiint_T f(x,y,z)\,dV = \int_0^1\int_0^{1-x}\int_{x+z}^1 f(x,y,z)\,dydzdx\ .$$

When the region D can be expressed in polar coordinates, for instance, it is of the form

$$\{(r\cos\theta, r\sin\theta):\ h_1(\theta)\le r\le h_2(\theta),\ \theta\in[\theta_1,\theta_2]\}\ ,$$

(1.5) becomes

$$\iiint_\Omega f\,dV = \int_{\theta_1}^{\theta_2}\int_{h_1(\theta)}^{h_2(\theta)}\int_{f_1(r\cos\theta,r\sin\theta)}^{f_2(r\cos\theta,r\sin\theta)} f(r\cos\theta,r\sin\theta,z)r\,dz\,drd\theta\ . \tag{1.6}$$

The representation of a point (x,y,z) in the form (r,θ,z) is called the *cylindrical coordinates* of (x,y,z).

Example 1.12 Find the volume of the region R bounded between $z = \sqrt{x^2+y^2}$ and $x^2+y^2+z^2=2$.

These two graphs intersect at $z=1$ and its projection to the xy-plane is the disk $x^2+y^2 \le 1$. Using cylindrical coordinates,

$$\begin{aligned}|R| &= \iiint_R 1\, dV \\ &= \int_0^{2\pi}\int_0^1\int_r^{\sqrt{2-r^2}} r\, dz\, dr d\theta \\ &= 2\pi \int_0^1 (\sqrt{2-r^2}-r) r\, dr \\ &= \frac{2\sqrt{2}}{3} .\end{aligned}$$

Another useful special coordinates is the spherical coordinates.

For each (x,y,z) in $\mathbb{R}^3$, we can find $(\rho,\varphi,\theta) \in [0,\infty)\times[0,\pi]\times[0,2\pi]$ such that $x=\rho\sin\varphi\cos\theta, y=\rho\sin\varphi\sin\theta, z=\rho\cos\varphi$. (ρ,φ,θ) is called the *spherical coordinates* of (x,y,z). These formulas set up a mapping Φ from $[0,\infty)\times[0,\pi]\times[0,2\pi]$ to $\mathbb{R}^3$. It is one-to-one and onto $\mathbb{R}^3$ (with the origin removed) when restricted to $(0,\infty)\times[0,\pi]\times[0,2\pi)$.

Let Ω_1 and Ω be two regions in (ρ,φ,θ)-space and (x,y,z)-space respectively that satisfy $\Phi(\Omega_1)=\Omega$. Given any function f in the (x,y,z)-space, $f\circ\Phi$ becomes a function in the (ρ,φ,θ)-space. The following formula holds:

$$\iiint_\Omega f(x,y,z)\, dV = \iiint_{\Omega_1} f(\rho\sin\varphi\cos\theta,\rho\sin\varphi\sin\theta,\rho\cos\varphi)\rho^2\sin\varphi\, dV(\rho,\varphi,\theta) .$$

In applications, the region Ω is usually of the form:

$$\Omega = \{(x,y,z) : \rho_1(\varphi,\theta) \le \rho \le \rho_2(\varphi,\theta),\ (\varphi,\theta)\in D\ \},$$

for some region D. Then we have

Theorem 1.15. *For a continuous function f in Ω,*

$$\iiint_\Omega f(x,y,z)\, dV = \iint_D \int_{\rho_1(\varphi,\theta)}^{\rho_2(\varphi,\theta)} f(\rho\sin\varphi\cos\theta,\rho\sin\varphi\sin\theta,\rho\cos\varphi)\rho^2\sin\varphi\, d\rho dA(\varphi,\theta) . \tag{1.7}$$

We refer to the text book for a proof of this theorem. We will re-derive it in the next chapter when we discuss the change of variables formula.

Example 1.13 Use spherical coordinates to find the volume of the circular cone whose base radius is R and height h.

In rectangular coordinates, the solid cone is described as

$$\Omega = \{(x,y,z): \ \frac{h}{R}\sqrt{x^2+y^2} \le z \le h, \ x^2+y^2 \le R^2\} \ .$$

In spherical coordinates, it is

$$\tilde{\Omega} = \{(\rho,\varphi,\theta): \ 0 \le \rho \le \rho_2(\varphi,\theta), \ 0 \le \varphi \le \varphi_0, \ 0 \le \theta \le 2\pi \ \}.$$

Here $z = h$ turns into $\rho_2 \cos\varphi = h$, that is,

$$\rho_2(\varphi,\theta) = \frac{h}{\cos\varphi} \ .$$

On the other hand, φ_0, which is determined by the perpendicular triangle with sides R and H, satisfies $h\tan\varphi_0 = R$. Hence

$$\varphi_0 = \tan^{-1} R/h.$$

Only rays from the original can hit $z = h$ when $\varphi \in [0,\varphi_0]$. Henceforth, the volume of the circular cone is given by

$$\begin{aligned} |\Omega| &= \int_0^{2\pi}\int_0^{\varphi_0}\int_0^{h/\cos\varphi} 1 \times \rho^2 \sin\varphi \, d\rho d\varphi d\theta \\ &= 2\pi \int_0^{\varphi_0} \frac{1}{3}\frac{h^3 \sin\varphi}{\cos^3\varphi} \, d\varphi \\ &= \frac{1}{3}\pi R^2 h \ . \end{aligned}$$

Example 1.14 Express the integral

$$\int_0^3 \int_0^{\sqrt{9-y^2}} \int_{\sqrt{x^2+y^2}}^{\sqrt{18-x^2-y^2}} f(x,y,z) \, dzdxdy$$

in cylindrical and spherical coordinates.

This is an ice-cream cone given by

$$\{(x,y,z): \ \sqrt{x^2+y^2} \le z \le \sqrt{18-x^2-y^2}, \ 0 \le r \le 3, \ 0 \le \theta \le 2\pi\},$$

in cylindrical coordinates. Therefore, this integral is equal to

$$\int_0^{2\pi}\int_0^3\int_r^{\sqrt{18-r^2}} f(r\cos\theta, r\sin\theta, z) r \, dzdrd\theta \ .$$

Next, in spherical coordinates, the ice-cream is described by

$$\{(x, y, z) : \ 0 \le \rho \le \rho_2, \ 0 \le \varphi \le \varphi_0, \ 0 \le \theta \le 2\pi \ \}.$$

Here ρ_2 describes the surface of the ice-cream which is given by $\rho_2 = \sqrt{18}$. On the other hand, $x^2+y^2 = 18-x^2-y^2$ implies $x^2+y^2 = 9$. That is, the circular cone and the spherical intersect at a disk of radius of 3 centered at the origin. The angle φ_0 is determined from the perpendicular triangle with sides 3 and $z = 3$, hence $\varphi_0 = \pi/4$. Our integral is equal ton

$$\int_0^{2\pi} \int_0^{\pi/4} \int_0^{\sqrt{18}} f(\rho \sin\varphi \cos\theta, \rho \sin\varphi \sin\theta, \rho\cos\varphi)\rho^2 \sin\varphi \, d\rho d\varphi d\theta \ ,$$

in spherical coordinates.

Example 1.15 Express the triple integral of a function f over the region which is bounded between $z = 3, z = 0$ and $x^2 + y^2 + z^2 = 16$ in spherical coordinates.

The sphere $x^2 + y^2 + z^2 = 16$ and $z = 3$ intersects at a circle which is projected down to the xy-plane as $x^2 + y^2 = 16 - 9 = 7$. Any ray of $\varphi \in [0, \varphi_0], \varphi_0 = \sin^{-1} \sqrt{7}/4$, hits the sphere. On the other hand, any ray of $\varphi \in [\varphi_0, \pi/2]$ hits the plane $z = 3$ or $\rho = 3/\cos\varphi$. Therefore, the triple integral is the sum of two integrals given by

$$\begin{aligned} \iiint_\Omega f \, dV &= \int_0^{2\pi} \int_{\varphi_0}^{\pi/2} \int_0^4 f(\rho \sin\varphi \cos\theta, \rho \sin\varphi \sin\theta, \rho\cos\varphi)\rho^2 \sin\varphi \, d\rho d\varphi d\theta \\ &\quad + \int_0^{2\pi} \int_0^{\varphi_0} \int_0^{3/\cos\varphi} f(\rho \sin\varphi \cos\theta, \rho \sin\varphi \sin\theta, \rho\cos\varphi)\rho^2 \sin\varphi \, d\rho d\varphi d\theta \ . \end{aligned}$$

Example 1.16 The same setting as in the previous example but now the region is the portion bounded between the sphere and the plane $z = 3$.

Now, observe every ray from the origin hits the plane $z = 3$ and then the sphere $\rho = 4$ when $\varphi \in [0, \varphi_0]$ and none otherwise. The triple integral should be

$$\int_0^{2\pi} \int_0^{\varphi_0} \int_{3/\cos\varphi}^{4} f(\rho \sin\varphi \cos\theta, \rho \sin\varphi \sin\theta, \rho\cos\varphi)\rho^2 \sin\varphi \, d\rho d\varphi d\theta \ .$$

1.8 A Variant Of Fubini's Theorem

In the derivation of the formula in Theorem 1.14, if we put the bracket differently, we will have

$$\iiint_B f \, dV \approx \sum_{i,j,k} f(x_i^*, y_j^*, z_k^*)\Delta x_i \Delta y_j \Delta z_k = \sum_k \left(\sum_{i,j} f(x_i^*, y_j^*, z_k^*)\Delta x_i \Delta y_j \right) \Delta z_k \ .$$

When $\|P\|$ is very small, $\Delta x_i, \Delta y_j, \Delta z_k$ are also very small,

$$\sum_k \left(\sum_{i,j} f(x_i^*, y_j^*, z_k^*) \Delta x_i \Delta y_j \right) \Delta z_k \approx \sum_k \iint_R f(x,y,z^*) \Delta z_k \ ,$$

where $R = [a,b] \times [c,d]$. Letting $\|P\| \to 0$, we get

$$\iiint_B f \, dV = \int_e^f \left(\iint_R f(x,y,z) dz \right) dA(x,y) \, dz \ .$$

When f is defined in Ω, let

$$\Omega(z) = \{(x,y) : \ (x,y,z) \in \Omega\}$$

the z-cross section of Ω. Suppose that $\Omega(z)$ is a region for each $z \in [e,f]$ and becomes empty elsewhere. We have the formula

$$\iiint_\Omega f \, dV = \int_e^f \iint_{\Omega(z)} f(x,y,z) \, dA(x,y) \, dz \ . \tag{1.8}$$

Taking $f \equiv 1$, the volume of Ω can be expressed as an integral of the area of its cross sections:

$$|\Omega| = \int_e^f |\Omega(z)| \, dz \ . \tag{1.9}$$

Example 1.14 Find the volume of the cone whose vertex is $(0,0,h)$ and base is a region D in the xy-plane.

By proportion, a line of length l on the xy-plane and the length x of its corresponding line in the xy-plane at z satisfy

$$\frac{h}{h-z} = \frac{l}{x} \ ,$$

that is, $x = l(h-z)/h$. Therefore, the area of the cross section of the cone at z is equal to

$$\frac{(h-z)^2}{h^2} |D| \ .$$

The volume of the cone is

$$\int_0^h \frac{(h-z)^2}{h^2} |D| \, dz = \frac{1}{3} |D| h \ .$$

Example 1.15 Show the volume of the ball $\{(x,y,z,w) : \ x^2+y^2+z^2+w^2 \le r^2\}$ in $\mathbb{R}^4$ is given by $\pi^2 r^4/2$.

It suffices to calculate the volume for the upper half ball. For each $w \in [0, r]$, the cross section $B(w)$ is a three dimensional ball of radius $\sqrt{r^2 - w^2}$. Therefore, the volume of the ball is equal to

$$\begin{aligned} 2\int_0^r \frac{4\pi}{3}(r^2 - w^2)^{3/2}\, dw &= 2\frac{4\pi}{3} r^4 \int_0^{\pi/2} \cos^4\theta \, d\theta \\ &= \frac{1}{2}\pi^2 r^4 \ . \end{aligned}$$

1.9 A Characterization Of Riemann Integral

From the view point of an analyst, the interpretation of integrals as area is not satisfying, let alone the physical point of view such as mass and centroid. Analysts would like to understand Riemann integral (in all dimensions) what the view of point of analysis. Here we present a theorem in this direction.

In the following we let V be the real vector space consisting of all piecewise continuous functions which vanish outside some bounded set in the plane. We will work on this setting for simplicity. You will see the same ideas also work in any dimension.

Theorem 1.16. *Let T be a map from V to $\mathbb{R}$ satisfying the following properties:*

1. *(Linearity) T is linear.*
2. *(Positivity preserving) $T(f) \geq 0$ provided $f \in V$ is nonnegative.*
3. *(Translation invariant) $T(f) = T(f')$ where f' is a translate of f.*
4. *(Normalization) $T(\chi_{R_0}) = 1$ where $R_0 = (0,1) \times (0,1)$.*

Then

$$T(f) = \iint_R f \, dA$$

for all $f \in V$.

f' is a *translate* of f if $f'(p) = f(p + p_0)$ for some $p_0 \in \mathbb{R}^2$.

Proof. Step 1. Divide R_0 into n many subsquares where a typical one is $(0, 1/n) \times (0, 1/n)$ and denote them by R_{ij}. All R_{ij} are translates of the typical one. By translational invariance, all $T(\chi_{R_{ij}})$ are equal. Therefore, from $\cup_{i,j} R_{ij} \subset R_0$ and positivity preserving we get

$$n^2 T(\chi_{(0,1/n)^2}) \leq \sum_{i,j} T(\chi_{R_{ij}}) \leq T(\chi_{R_0}) = 1 \ ,$$

which implies, together with translational invariance,

$$T(\chi_S) \le 1/n^2 ,$$

for any square S of the form $(a, a+1/n) \times (b, b+1/n)$.

Step 2. For any horizontal line segment L, $T(\chi_L) = 0$. WLOG assume L is a natural number. We can fully cover L by $2nL$ many squares S_k of side length $1/n$. From $L \subset \cup_k S_k$ we get $\chi_L \le \sum_k \chi_{S_k}$, so

$$T(\chi_L) \le 2nL \times T(\chi_{S_1}) \le 2L/n \to 0 , \quad \text{as } n \to \infty.$$

Hence $T(\chi_L) = 0$. The same result holds for vertical line segments.

Step 3. $T(\chi_S) = 1/n^2$ where S is a square of side $1/n$, including or excluding its boundary points. This follows from combining Step 1 and Step 2 since the boundary are horizontal or vertical lines.

Step 4. Let $R(a,b)$ be a rectangle of length a and height b. I leave it as an exercise to show $T(\chi_{R(a,b)}) = ab$. Show this for a, b rational numbers first and then for irrational numbers.

Step 5. Let f be a continuous function vanishing outside some rectangle R. Let P be a partition on R into R_{ij}. Let m_{ij} and M_{ij} be the minimum and maximum of f over R_{ij} respectively. From $f \le \sum_{i,j} M_{ij}\chi_{R_{ij}}$ we deduce

$$T(f) \le T(\sum_{i,j} M_{ij}\chi_{R_{ij}}) = \sum_{i,j} M_{ij}|R_{ij}| .$$

As $\|P\| \to 0$, we get

$$T(f) \le \iint_R f\,dA .$$

On the other hand, $\sum_{i,j} m_{ij}\chi_{R'_{ij}} \le f$ where R'_{ij} is the subrectangle without counting in the boundary points. Then $\sum_{i,j} m_{ij}|R_{ij}| \le T(f)$. Letting $\|P\| \to 0$, we get

$$\iint_R f\,dA \le T(f) .$$

We have proved the theorem for continuous functions. The general case can be established via an approximation argument.

□