19 Integration Revisited

Riemann can integrate piecewise continuous functions. However, there are many functions which cannot be integrated by the Riemann integration, although the values of their integrals are more or less obvious. In this section, the basic idea of the Lebesgue integral is given with a practical summary. The theory is a natural prerequisite for understanding Hilbert space. The most natural integral concept for Fourier expansion is the Lebesgue integral. In the Appendix, rudiments of measure theory is outlined.

Key words: measure zero, almost everywhere, Lebesgue integral, dominated convergence theorem, Beppo-Levi's theorem, Fubini's theorem, Gaussian integral, Wick's theorem.

Remember:
(1) Lebesgue integral is defined by the integral of simple functions (= functions taking only countably many values) (**19.7-8**).
(2) There are several very powerful theorems for Lebesgue integration (**19.11-17**). Basically, they justify what looks formally OK to physicists.
(3) Lebesgue integral is the most natural framework to consider Fourier analysis (**19.18**).
(4) Gaussian integrals should be very familiar (**19.19-20**).

[19.0 Practical Check].

Exercise. Before going into the discussion of the Lebesgue integration theory, let us check our practical ability to compute Riemann integrals. (1) Compute the following indefinite integrals:

$$\int dx \frac{ax+b}{cx+d}. \tag{19.1}$$

Here we assume that $a, b, c(\neq 0), d$ are constants.
(2) Let $n \in \boldsymbol{N}$. For

$$I_n \equiv \int_0^{\pi/2} \sin^n x dx \tag{19.2}$$

demonstrate that

$$I_n = \left(1 - \frac{1}{n}\right) I_{n-2}. \tag{19.3}$$

Then, compute I_n.
(3) Find the range of α where

$$\int_0^\infty \frac{\sin^2 x}{x^\alpha} dx \tag{19.4}$$

exists.
(4) [Fresnel integral]. Show that

$$\int_0^\infty \sin(x^2) dx \tag{19.5}$$

exists (as a Riemann integral). cf **8B.8**(1).
(5) Does

$$\int_0^\infty \sin(\cosh x) dx \tag{19.6}$$

exist (as a Riemann integral)?
(6) Show

$$\int_0^\infty \left(\frac{\sin x}{x}\right)^2 dx = \frac{\pi}{2}. \tag{19.7}$$

Use ($\to$**8B.7**)

$$\int_0^\infty e^{-\alpha x} \frac{\sin \lambda x}{x} dx = \frac{\alpha}{\alpha^2 + \lambda^2}. \tag{19.8}$$

(7) Show that

$$\int_0^\infty \frac{\sin ax \cos bx}{x} = \frac{\pi}{2}, \tag{19.9}$$

if $a > b > 0$. What happens otherwise?
(8) Show that

$$\int_0^{\pi/2} \log \sin \theta d\theta = -\frac{\pi}{2} \log 2. \tag{19.10}$$

(9) Compute

$$\lim_{n\to\infty} \frac{1}{n}[1 + \cos\frac{x}{n} + \cos\frac{2x}{n} + \cdots \cos\frac{(n-1)x}{n} + \cos x] \tag{19.11}$$

(10) Compute

$$\frac{d^n}{dx^n} \int_0^x \frac{(x-y)^{n-1}}{(n-1)!} f(y) dy. \tag{19.12}$$

Discussion.
(1) Let

$$I(a,b) \equiv \int_0^\infty \frac{dx}{\sqrt{(a^2+x^2)(b^2+x^2)}} \tag{19.13}$$

for positive a and b. Show that

$$I(a_n, b_n) = I(a,b) \tag{19.14}$$

for any $n = 1, 2, \cdots$, where $a_{n+1} = (a_n + b_n)/2$ and $b_{n+1} = \sqrt{a_n b_n}$, where $a_1 = a$ and $b_1 = b$. a_n and b_n converge to a common limit μ determined by a and b. Gauss ($\to$**7.15**) used the bove observation to compute $\mu = \pi/2I$. Show this conclusion.
(2) Let f be integrable on $[0, 1]$. Then

$$\int_0^1 \exp(f(t))dt \geq \exp\left(\int_0^1 f(t)dt\right). \tag{19.15}$$

Note that $\int_0^1 f(t)dt$ may be understood as the average of f on $[0, 1]$ ($\to$**2A.1**, Discussion (A)).

19.1 Dirichlet function. The Dirichlet function is defined as[275]

$$D(x) = \begin{cases} 0 \text{ for } x \notin \boldsymbol{Q}, \\ 1 \text{ for } x \in \boldsymbol{Q}. \end{cases} \tag{19.16}$$

$\int_0^1 dx D(x)$ must be zero, but obviously this function is not Riemann integrable.

19.2 The area below $D(x)$ must be zero. We know ($\to$**17.18(4)**, **A1.16**) all the rational numbers can be counted, so we may write the totality of rational numbers in $[0, 1]$ as $Q \equiv \{y_n\}_{n=1}^{\infty} = \boldsymbol{Q} \cap [0, 1]$. Let us cover y_n with an interval E_n of length $\epsilon/2^n$ centered at y_n. Obviously, $\cup E_n \supset Q$ for any positive ϵ, but the total length of $\cup E_n$ is not larger than ϵ, because length$(\cup E_n) \leq \sum($ length $E_n) = \epsilon$. This number is any positive number, so it can be indefinitely small. Hence, the total area occupied by Q must be zero. This must be the area below $D(x)$ on $[0, 1]$. Hence, '$\int_0^1 dx D(x)$' $= 0$ ($\to$**19.7**).

19.3 Measure zero. We have demonstrated that $\boldsymbol{Q}$ is *measure zero.* A set $U \subset \boldsymbol{R}$ is called a *measure zero* set, if it can be covered by countably many open intervals the totality of the length of which is less than ϵ for any $\epsilon (> 0)$. **19.2** tells us that any countable set is measure zero. See Appendix a19 for a general discussion about measure ($\to$**a19.4**).

19.4 Lebesgue's characterization of Riemann integrability. In his thesis, Lebesgue showed the following theorem.
Theorem. A bounded function f is integrable in the sense of Riemann on $[a, b]$ if and only if the set of discontinuous points of f is measure zero. □
Obviously, $D(x)$ is <u>not</u> integrable in the sense of Riemann.

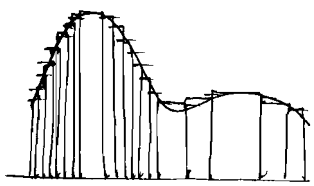
19.5 "Almost everywhere". Lebesgue also introduced the concept of *almost everywhere*: if a property 'A' is true for a function f except on

[275] This is the characteristic function of the set of all the rational numbers.

the measure zero set, we say f has the property 'A' almost everywhere. Thus the theorem above can be restated as: A bounded function f is Riemann integrable if f is almost everywhere continuous.

19.6 Simple function. A function which takes at most countably many ($\rightarrow$**17.18(4)**, **A1.16**) values is called a *simple function.* The Dirichlet function ($\rightarrow$**19.1**) is a simple function, because it assumes only two values, 0 and 1.

19.7 Lebesgue integral of simple functions. Let f be a real-valued simple function defined on an interval I. If the right-hand-side of the following formula converges absolutely, we say f is *Lebesgue integrable* and the limit is denoted by just the same symbol as the Riemann integral:



$$\int_I f(x)dx \equiv \sum_n y_n |I_n|, \tag{19.17}$$

where $|*|$ is the total length of the set $*$, and $I_n \equiv \{x | x \in I, f(x) = y_n\}$. Cantor showed $|\boldsymbol{Q}| = 0$ ($\rightarrow$**19.2**). Hence, the Dirichlet function is Lebesgue integrable and the value of the integral is zero.[276]

Note that the values of a function on measure zero sets are irrelevant to the value of the integral.

19.8 Lebesgue integral of general function: $L_1([a, b])$. The Lebesgue integral of a function f on an interval $[a, b]$ is defined as follows. Make a uniform approximation sequence of Lebesgue integrable simple functions f_i for f:

$$\sup_{x \in [a,b]} |f_i(x) - f(x)| \rightarrow 0 \text{ as } i \rightarrow \infty. \tag{19.18}$$

Then

$$\int_a^b f(x)dx \equiv \lim_{i \rightarrow \infty} \int_a^b f_i(x)dx. \tag{19.19}$$

[Of course, if we cannot find such a sequence, f is not Lebesgue integrable.]

The totality of functions Lebesgue integrable on the interval $[a, b]$ is denoted by $L_1([a, b])$.

[276] In this definition, it is very crucial that all I_n have lengths. Or more generally, if we wish to define an integral of functions on a multidimensional space, then I_n must have a definite volume. Therefore, Lebesgue had to contemplate on the concept 'volume.' This led him to his measure theory ($\rightarrow$**a19**). We say a simple function f is *measurable* if all I_n have well-defined volumes ($\rightarrow$**a19.4**). A function f is said to be *measurable* (more precisely, Borel measurable), if the set $\{x \,|\, a < f(x) < b\}$ has a definite length (measure) for any a and $b(> a)$.

Discussion [Fundamental properties of integrals].
(I) **Double Linearity**. We know that the integral is linear with respect to the integrand. There is one more linearity with respect to the domain as we already noticed in **6.2**:

$$\int_a^c f(t)dt = \int_a^b f(t)dt + \int_b^c f(t)dt \tag{19.20}$$

or

$$\int_{[a,b]+[b,c]} f(t)dt = \int_{[a,b]} f(t)dt + \int_{[b,c]} f(t)dt. \tag{19.21}$$

If we define

$$\int_{\alpha[a,b]} f(t)dt = \alpha \int_{[a,b]} f(t)dt, \tag{19.22}$$

then $\int$ becomes a linear map on geometrical objects (in this case we discussed only 1D objects, but this can be generalized to general dimensional spaces). Notice that the convention is meaningful if we interpret the integral over $-[a,b]$ to be the integral on $[a,b]$ from b to a instead of a to b ($-$ is the reversing of orientation).
(II) **Non-negativity** and **monotonicity**. If the integrand is nonnegative, its integral is nonnegative. Consequently, if $f \geq g$, then $\int_a^b dt f(t) \geq \int_a^b g(t)dt$.
(III) **Boundedness.** If the integrand is bounded, then its integral over a bounded set is bounded.

19.9 Remark. We must demonstrate that the limit in **19.8** does not depend on the choice of the approximation sequences, but it is a technical detail. An important difference between the Riemann and the Lebesgue integrations is that the latter requires absolute convergence. A. N. Kolmogorov and S. V. Fomin, *Introductory Real Analysis* (Revised English edition, Englewood Cliffs, 1970)[277] is an excellent self-study textbook for the measure theory and Lebesgue integration (and standard functional analysis (say, spectral analysis)).

19.10 Relation between Riemann and Lebesgue integrals.
(1) If f is integrable in both the senses, their values are the same.
(2) If f is bounded and Riemann integrable, then it is Lebesgue integrable. But
(3) There are Riemann integrable but not Lebesgue integrable functions, and vice versa.

The practical merit of the Lebesgue integral is that the conditions for exchanging the order of operations (say, limit and integral) can be simpler than those for Riemann integrals (→**19.11**, **19.14**, **19.17**). This simplicity is due to the absolute convergence in the definition

[277]Its original Russian version is an undergraduate textbook for Analysis III (designed by Kolmogorov) of Dept of Engineering Mathematics of Moscow State University.

($\rightarrow$**19.7**).

19.11 Theorem [Lebesgue's dominated convergence theorem]. Let I be an interval. If $\lim_{n\to\infty} f_n(x) = f(x)$ for almost all $x \in I$ (i.e., except on a measure zero set ($\rightarrow$**19.3**), f_n converges to f), and if there is a Lebesgue integrable function ($\rightarrow$**19.8**) $\varphi(x)$ such that $|f_n(x)| < \varphi(x)$ on I, then

$$\lim_{n\to\infty} \int_I f_n(x)dx = \int_I f(x)dx. \tag{19.23}$$

□

19.12 Theorem [Beppo-Levi]. Let f_n be Lebesgue integrable on an interval I, $\int_I f_n(x)dx < K$ for some number K for all n, and $f_1 \le f_2 \le \cdots \le f_n \le \cdots$. Then

$$\lim_{n\to\infty} \int_I f_n(x)dx = \int_I \lim_{n\to\infty} f_n(x)dx. \tag{19.24}$$

□

19.13 Example. Termwise integration of $\sum x^n = (1-x)^{-1}$. For $t \in [0,1)$, we may apply Beppo-Levi's theorem to the partial sums to integrate this termwisely:

$$\int_0^t \sum_{n=0}^{\infty} x^n dx = \sum_{n=0}^{\infty} \int_0^t x^n dx = \sum_{n=1}^{\infty} \frac{t^n}{n} = -\ln(1-t). \tag{19.25}$$

Exercise.
Compute the following integrals in the $n \to \infty$ limit:
(1)

$$\int_0^1 \frac{x}{1+nx}dx. \tag{19.26}$$

(2)

$$\int_0^1 \frac{1}{1+nx^2}dx \tag{19.27}$$

Notice that the exchange of the order of limit and integration does not work for

$$\int_0^1 \frac{n}{1+n^2x^2}dx. \tag{19.28}$$

See **14.19**.

19.14 Theorem [Fubini]. If $\int dx\,(\int dy|f(x,y)|)$ or $\int dy\,(\int dx|f(x,y)|)$ is finite, then we may exchange the order of two integrations in $\int dx \int dy f(x,y)$.
□

Discussion.
(1) Using the integral of $f(x,y) = x^y$ on $[0,1] \times [a,b]$ for $0 < a < b$, demonstrate

$$\int_0^1 \frac{x^b - x^a}{\log x} dx = \log \frac{1+b}{1+a}. \tag{19.29}$$

(2) Demonstrate that

$$\int\int_{x \geq 0, y \geq 0} dx dy f(a^2x^2 + b^2y^2) = \frac{\pi}{4ab} \int_0^\infty x f(x) dx. \tag{19.30}$$

(3) Compute

$$\int_1^2 dx \int_1^x dy \frac{x^2}{y^2}, \tag{19.31}$$

$$\int_0^1 dx \int_0^{\sqrt{1-x^2}} dy (1-y^2)^{3/2}, \tag{19.32}$$

$$\int_0^1 dx \int_{\sqrt{x}}^1 dy \sqrt{1+y^2}. \tag{19.33}$$

19.15 Pathological example. Do not think the order of integrations can be freely changed:

$$\int_0^1 dx \int_0^1 dy \frac{x^2 - y^2}{(x^2+y^2)^2} = \frac{\pi}{4}, \quad \int_0^1 dy \int_0^1 dx \frac{x^2 - y^2}{(x^2+y^2)^2} = -\frac{\pi}{4}. \tag{19.34}$$

Demonstrate that the condition for **19.14** is violated.

Discussion
The reason for the pathology is explained by Legendre with the aid of the following formula:

$$\int_\alpha^1 dx \int_\beta^1 dy \frac{x^2 - y^2}{(x^2+y^2)^2} = \frac{\pi}{2} - \arctan \frac{\beta}{\alpha}. \tag{19.35}$$

Demonstrate the formula and complete the argument.

19.16 Good function principle. In short, if a relation is correct for a simple function ($\rightarrow$**19.6**), then it is correct for integrable functions. This is sometimes called the *good function principle.*

19.17 Exchanging differentiation and integration. Suppose $f(x, \alpha)$ is integrable for any α in its range, and $\partial_\alpha f$ is integrable, then

$$\frac{d}{d\alpha} \int f(x, \alpha) dx = \int \frac{\partial}{\partial \alpha} f(x, \alpha) dx. \tag{19.36}$$

Very crudely peaking, for Lebesgue integration, if the formal result is mathematically meaningful, then the result is (eventually) justifiable.

Discussion.
(1) Let f be continuous. Demonstrate that g defined by

$$g(x) = \int_0^x \frac{(x-y)^{n-1}}{(n-1)!} f(y) dy \tag{19.37}$$

is C^n and $g^{(n)}(x) = f(x)$. [Almost the same as **19.0** (10).]
(2) **Hadamard representation.** Let $f(x, y)$ be C^1 in the ball of radius r centered at (x_0, y_0). Then

$$f(x,y) = f(x_0, y_0) + f_1(x,y)(x - x_0) + f_2(x,y)(y - y_0), \tag{19.38}$$

where

$$f_1(x,y) = \int_0^1 \frac{\partial f}{\partial x}(x_t, y_t) dt, \quad f_2(x,y) = \int_0^1 \frac{\partial f}{\partial y}(x_t, y_t) dt \tag{19.39}$$

with $x_t = tx + (1-t)x_0$ and $y_t = ty + (1-t)y_0$.
Exercise.
(1) Show that

$$F(x) = \int_0^\infty e^{-y^2} \sin 2xy \, dy \tag{19.40}$$

satisfies

$$F'(x) + 2xF(x) = 1. \tag{19.41}$$

(2) A similar question is: Let

$$I(a) = \int_0^\infty e^{-x^2} \cos 2ax dx. \tag{19.42}$$

Show that

$$\frac{dI}{da} = -2aI. \tag{19.43}$$

Use this to demonstrate that

$$I = \frac{\sqrt{\pi}}{2} e^{-a^2}. \tag{19.44}$$

[Hint. The change of variables $z = x + a/x$ works.]
(3) Let

$$I(a) = \int_0^\infty \exp\left\{-b^2\left(x^2 + \frac{a^2}{x^2}\right)\right\} dx. \tag{19.45}$$

Demonstrate that

$$\frac{dI}{da} = -2b^2 I. \tag{19.46}$$

Then, show

$$I(a) = \frac{\sqrt{\pi}}{2|b|} e^{-2ab^2}. \tag{19.47}$$

19.18 Why is the Lebesgue integral most natural for Fourier analysis? As we have already mentioned in **17.10(3)** if f is square

Lebesgue integrable, then its Fourier series is almost everywhere convergent to f. See also Carlson's theorem (→**17.9**). Physicists know that Fourier transform is a powerful tool to disentangle convolution (→**32A.2**). This can be done freely only when we integrate all integrals as Lebesgue integrals. We can make a continuous and absolute integrable function f such that its convolution to itself $\int dx f(t-x)f(x)$ is Lebesgue integrable, but diverges for all rational t (so that it is not Riemann integrable).[278] That is, if we use the Riemann integral, then we cannot freely use Fourier transformation to disentangle the convolution. The Lebesgue integration theory is much more elegant and fundamental in Fourier analysis than the Riemann integration.

19.19 Gaussian integral, 'Wick's theorem'. The following integral (the generator of multidimensional Gaussian distribution) is of vital importance in theoretical physics:

$$I(A, \boldsymbol{b}) \equiv \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dx_1 \cdots dx_n \exp\left(-\frac{1}{2}\sum_{i,j=1}^{n} A_{ij}x_ix_j + \sum_{i=1}^{n} x_ib_i\right), \tag{19.48}$$

where $A = Matr(A_{ij})$ is an $n \times n$ symmetric non-singular matrix, and $\boldsymbol{b}$ is an n-vector. We get

$$I(A, \boldsymbol{b}) = (2\pi)^{n/2}(detA)^{-1/2}\exp\left(\frac{1}{2}\sum_{i,j} A_{ij}b_ib_j\right). \tag{19.49}$$

$I(A, \boldsymbol{b})/I(A, 0)$ is called the generator (generating function) of the Gaussian distribution with mean zero and covariance matrix given by A^{-1}. The standard method to compute this is to shift the origin to the minimum point of the function in the parentheses as

$$y_i = x_i - \sum_j (A^{-1})_{ij}b_j. \tag{19.50}$$

This leads to

$$I(A, \boldsymbol{b}) = \exp\left(\frac{1}{2}\sum_{i,j}(A^{-1})_{ij}b_ib_j\right)\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dy_1 \cdots dy_n \exp\left(-\sum_{i,j=1}^{n} A_{ij}y_iy_j\right) \tag{19.51}$$

The integral can be computed by diagonalizing the matrix.

According to **19.17** we can freely change the order of differentiation with respect to $\boldsymbol{b}$ and integration in (19.48). In this way we arrive at the so-called *Wick's theorem*: For $\boldsymbol{b} = 0$

$$\langle x_a x_b \cdots x_z \rangle = \sum (A^{-1})_{k_1k_2}(A^{-1})_{k_3k_4} \cdots (A^{-1})_{k_{n-1}k_n}, \tag{19.52}$$

[278] See Körner, Example C.6 on p570.

where $\{k_1, \cdots, k_n\} = \{a, \cdots, z\}$ and the sum is over all the possible pairings of $a, b, \cdots, z$. For example,

$$\langle x_1x_2x_3x_4 \rangle = \langle x_1x_2 \rangle \langle x_3x_4 \rangle + \langle x_1x_3 \rangle \langle x_2x_4 \rangle + \langle x_1x_4 \rangle \langle x_2x_3 \rangle. \quad (19.53)$$

Exercise.
(A) Compute the following integrals:
(1)

$$\int\int_{x\geq 0, y\geq 0} dxdy\, e^{-(x^2+2xy\cos\theta+y^2)}. \quad (19.54)$$

(2)

$$\int\int_{\boldsymbol{R}^2} dxdy\, e^{-(x^2+2xy\cos\theta+y^2)}. \quad (19.55)$$

(B) Using the spherical symmetry of the Gaussian integral, find the following integrals in terms of

$$U \equiv \int d^d\boldsymbol{k} e^{-ak^2/2}. \quad (19.56)$$

(1)

$$I = \int d^d\boldsymbol{k} \frac{k_x^2}{k^2} e^{-ak^2/2}. \quad (19.57)$$

(2)

$$J = \int d^d\boldsymbol{k} \frac{k_x^2 k_y^2}{k^4} e^{-ak^2/2}. \quad (19.58)$$

[Hint. (19.53) and $\langle k^4 \rangle = d\langle k_x^4 \rangle + d(d-1)\langle k_x^2 k_y^2 \rangle$. Also differentiation and integration with respect to a (or $-a/2$) is useful.]

19.20 Gaussian integral: complex case. We have the following analogous formula

$$I(A, \boldsymbol{b}) \equiv \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dz_1 d\overline{z}_1 \cdots dz_n d\overline{z}_n \exp\left(-\sum_{i,j=1}^{n} A_{ij}\overline{z}_i z_j + \sum_{i=1}^{n}(\overline{z}_i b_i + z_i \overline{b}_i)\right), \quad (19.59)$$

where A is any nonsingular $n \times n$ matrix, $\boldsymbol{b}$ is a complex n-vector. In terms of real variables x_i and y_i as

$$z_i = (x_i + iy_i)/\sqrt{2}, \quad (19.60)$$

we get $dz_i d\overline{z}_i = dx_i dy_i$.[279] Integration is understood as the integration with respect to these real variables. The result is

$$I(A, \boldsymbol{b}) = (2\pi)^n (detA)^{-1} \exp\left(\sum_{i,j}(A^{-1})_{ij}\overline{b}_i b_j\right). \quad (19.61)$$

[279] although formally, the calculation here seems to justify the equality, it is better to undersdand that $dzd\overline{z}$ is a shorthand notation of $dxdy$.

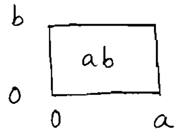
The cleverest proof of this relation is: (i) (if necessary) to slightly perturb A so that all the eigenvalues of $A+\delta A$ are distinct (so that $A+\delta A$ is diagonalizable); (ii) compute the integral analogous to **19.19**; then (iii) use the continuity of the integral as a function of the components of A to obtain the result for the unperturbed case.

APPENDIX a19 Measure

In this appendix the general theory of the Lebesgue measure is outlined. Without measure theory proper understanding of statistical mechanics and dynamical systems is impossible. However, just as all the important topics, the essence of measure theory is not at all hard to understand. The theory could be read as a very nice example of the analysis of a concept that we seem to know intuitively. For a more formal introduction Kolmogorov-Fomin is strongly recommended.

a19.0 Reader's guide to this appendix. (1) + (3) is the minimum of this appendix:
(1) The ordinary Lebesgue measure = volume is explained up to **a19.6**. These entries should be very easy to digest. Remember that Archimedes reached this level of sophistication more than 2000 years ago.
(2) General Lebesgue measure is outlined in **a19.9-11**. This is an abstract repetition of (1), so the essence should be already obvious.
(3) Lebesgue integral is redefined in terms of the Lebesgue measure in **a19.15** with a preparation in **a19.14**. This leads us naturally to the concept of functional and path integrals (**a19.16**).
(4) Probability is a measure with total mass 1 (i.e., normalized) (**a19.19**).
(5) If we read any probability book, we encounter the triplet $(P, X, \mathcal{B})$. The reason why we need such a nonintuitive device is explained in **a19.20-21**.

a19.1 What is volume? For simplicity, we confine our discussion to 2-space, but our discussion can easily be extended to higher dimensional spaces. The question is: what is 'area' ? It is not easy to answer this question for an arbitrary shape.[280] Therefore, we should start with a seemingly obvious example. The area of a rectangle $[0, a] \times [0, b]$ in $\boldsymbol{R}^2$ is ab. Do we actually know this? Why can we say the area of the rectangle is ab without knowing what area is? To be logically conscientious we must accept:

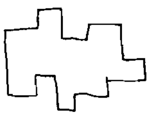


Definition. The area of a rectangle which is congruent[281] to $\langle 0, a\rangle \times \langle 0, b\rangle$ (Here $\langle$ is [or (and $\rangle$ is] or)) is <u>defined</u> to be ab. Notice that

[280]As we will see soon in **a19.21**, if we stick to our usual axiomatic system of mathematics ZF+C (→**17.18(5)** for references), then there are figures without area.

[281]This word is defined by the superposability. That is, if we move (translate, rotate) a figure A and can exactly superpose it on B, we say A and B are congruent. As Hilbert (→**20.4**) realized we must guarantee that the figure does not deform, etc., while being moved, so that we need an axiom, which was never stated in Euclid, although freely used by him (just as the Axiom of Choice in the early 20th century).

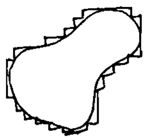
area is defined so that it is not affected by whether the boundary is included or not.



a19.2 Area of fundamental set. A set which is a direct sum (disjoint union) of finite number of rectangles is called a *fundamental set.* The area of a fundamental set is defined by the sum of the areas of constitutive rectangles.

It should be intuitively obvious that the join and the common set of fundamental sets are again fundamental.

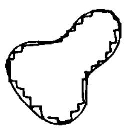
a19.3 Heuristic consideration. For an arbitrary shape, the strategy for defining its area should be to approximate the figure with a sequence of fundamental sets. We should use the idea going back to Archimedes; we must approximate the figure from the inside and from the outside. If both sequences converge to the same area, we should define the area to be the are of the figure.



a19.4 Outer measure. Let A be a set. We consider a cover of A with finite number of rectangles P_k (inclusion or exclusion of their boundaries can be chosen conveniently $\to$**a19.1**), and call it a rectangular cover $P = \{P_k\}$ of A. Let us denote the area of a rectangle P_k by $m(P_k)$. The *outer measure* $m^*(A)$ of A is defined by

$$m^*(A) \equiv \inf \sum_k m(P_k), \tag{19.62}$$

where the infimum is taken over all the finite or countable rectangular covers of A.
$m^*(A) = 0$ is equivalent to A being measure zero ($\to$**19.3** or a *null set*).



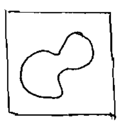
a19.5 Inner measure. For simplicity, let us assume that $A \in E \equiv [0,1] \times [0,1]$. Then, the inner measure $m_*(A)$ of A is defined by

$$m_*(A) = 1 - m^*(E \setminus A). \tag{19.63}$$

Obviously,

$$m^*(A) \geq m_*(A) \tag{19.64}$$

for any figure A.



a19.6 Measurable set, area = Lebesgue measure. Let A be a bounded subset of E.[282] If $m^*(A) = m_*(A)$, then we say A is *measurable* (in the sense of Lebesgue), and $m^*(A)$ written as $\mu(A)$ is called its

[282]It should be obvious how to generalize our argument to a more general bounded set in $\boldsymbol{R}^2$.

area (= *Lebesgue measure*).

a19.7 Additivity. Assume that all the sets here are in a bounded rectangle, say, E above. The join and the common set of finitely many measurable sets are again measurable. This is true even for countably many measurable sets. The second statement follows from the preceding statement thanks to the finiteness of the outer measure of the join or the common set.

a19.8 σ-additivity. Let $\{A_n\}$ be a family of measurable sets satisfying $A_n \cap A_m = \emptyset$ for $n \neq m$. Let $A = \cup_n A_n$. Then,

$$\mu(A) = \sum_n \mu(A_n). \tag{19.65}$$

This is called the *σ-additivity* of the Lebesgue measure. □

[Demo] A is measurable due to **a19.7**. Since $\{A_n\}$ covers A, $\mu(A) \leq \sum \mu(A_n)$. On the other hand $A \supset \cup_{n=1}^N A_n$, so that for any N $\mu(A) \geq \sum_{n=1}^N \mu(A_n)$.

a19.9 Measure, general case. A map from a family of sets to $\boldsymbol{R}$ is called a *set function.* A set function m satisfying the following three conditions is called a *measure.*

(1) m is defined on a semiring[283] $\mathcal{S}$. [Note that the set of all the rectangles is a semiring.]

(2) $m(A) \geq 0$.

(3) m is an additive function: If A is direct-sum-decomposed in terms of the elements of $\mathcal{S}$ as $A = \cup_{k=1}^n A_k$, then $m(A) = \sum_{k=1}^n m(A_k)$.

Therefroe, the area μ defined in **a19.6** is a measure on the set of all the rectangles. In the case of area, the definition of area is extended from rectangles to fundamental sets ($\rightarrow$**a19.2**). This is the next step:

a19A.10 Minimum algebra on $\mathcal{S}$, extension of measure. The totality of sets A which is a finite join of the elements in $\mathcal{S}$ is called the *minimum algebra* generated by $\mathcal{S}$. Notice that the totality of fundamental sets in **a19.2** is the minimum algebra of sets generated by the totality of rectangles. Just as the concept of area could be generalized to the area of a fundamental set, we can uniquely extend m defined on $\mathcal{S}$ to the measure defined on the algebra generated by $\mathcal{S}$.

a19.11 Lebesgue extension. We can repeat the procedure to define μ from m^* and m_* in **a19A.5** for any measure m on $\mathcal{S}$ (in an

[283]If a family of sets $\mathcal{S}$ satisfies the following conditions, it is called a *semiring* of sets:
(i) $\mathcal{S}$ contains $\emptyset$,
(ii) If $A, B \in \mathcal{S}$, then $A \cap B$ and $A \cup B$ are in $\mathcal{S}$,
(iii) if A_1 and A are in $\mathcal{S}$ and $A_1 \subset A$, then $A \setminus A_1$ can be written as a direct sum (the join of disjoint sets) of elements in $\mathcal{S}$.

abstract fashion). We define m^* and m_* with the aid of the covers made of the elements in $\mathcal{S}$. If $m^*(A) = m_*(A)$, we define the *Lebesgue extension* μ of m with $\mu(A) = m^*(A)$, and we say A is *μ-measurable.*

a19.12 Remark. When we simply say the Lebesgue measure, we usually mean the volume (or area) defined as in **a19A.6**. However, there is a different usage of the word. μ constructed in **a19.11** is also called a Lebesgue measure. That is, a measure constructed by the Lebesgue extension is generally called a Lebesgue measure. This concept includes the much narrower usage common to physicists.

a19.13 σ-additivity. (3) in **a19.9** is often replaced by the following *σ-additivity* condition: Let A be a sum of countably many disjoint μ-measurable sets $A = \cup_{n=1}^{\infty} A_n$. If

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A_n), \tag{19.66}$$

we say μ is a σ-additive measure.

The Lebesgue measure defined in **a19.6** is σ-additive. Actually, if m is σ-additive on a semiring of sets, then its Lebesgue extension is also σ-additive.

a19.14 Measurable function. A real function defined on a set D is called a *μ-measurable function* for a Lebesgue measure μ on the set, if any 'level set' $\{x \mid f(x) \in [a, b]\} \cap D$ is μ-measurable. When we simply say a function is measurable, then it means that any level set has a well defined volume in the ordinary sense.

a19.15 Lebesgue integral with measure μ. Let μ be Lebesgue measure on $\boldsymbol{R}^n$. Then the Lebesgue integral of a μ-measurable function on $U \subset \boldsymbol{R}^n$ is defined as

$$\int_U f(x) d\mu(x) = \lim_{\epsilon \to 0} \sum a\, \mu(\{x \mid f(x) \in [a - \epsilon/2, a + \epsilon/2)\} \cap U), \tag{19.67}$$

where the sum is over all the disjoint level sets of 'thickness' ϵ (> 0).[284]

a19.16 Functional integral. As the reader has seen in **a19.15**, if we can define a measure on a set, we can define an integral over the set.

[284]The measures m satisfying $\mu(A) = 0 \Rightarrow m(A) = 0$, where μ is the Lebesgue measure (volume), is said to be ***absolutely continuous*** with respect to μ. If m is absolutely continuous w.r.t. μ, then Lebesgue extension, Lebesgue integral, etc are easy without any technical difficutly just as the volume. However, careful consideration is needed because there are 'singular' measures.

The set need not be an ordinary finite-dimensional set, but can be a function space. In this case the integral is called a *functional integral.* If the set is the totality of paths from time $t = 0$ to T, that is, if the set is the totality of continuous functions: $[0, T] \to \boldsymbol{R}^d$, we call the integral over the set a *path integral.* The Feynman-Kac path integral (→**30.12**) is an example.[285]

a19.17 Uniform measure. The Lebesgue measure defined in **a19.6** is uniform in the sense that the volume of a set does not depend on its absolute location in the space. That is, the measure is translationally invariant (see **a19.20** below for a further comment). However, there is no useful uniform measure in infinite dimensional spaces (→**20.2** Discussion (1)). Thus every measure on a function space or path space must be non-uniform.

a19.18 Borel measure. Usually, we mean by a *Borel measure* a measure which makes measurable all the elements of the smallest algebra (→**a19.10**) of sets containing all the rectangles.

a19.19 Probability. A (Lebesgue) measure P with the total mass 1 is called a *probability measure.* To compute the expectation value with respect to P is to compute the Lebesgue integral w.r.t. the measure P.

When we read mathematical probability books, we always encounter the 'triplet' $(P, X, \mathcal{B})$, where P is a probability measure, X is the totality of elementary events (the event space; $P(X) = 1$) and $\mathcal{B}$ is the algebra of measurable events. This specification is needed, because if we assume that every composite event has a probability, we have paradoxes.[286] This question arose from the characterization of 'uniform measure' in a finite dimensional Euclidean space:

a19.20 Lebesgue's measure problem. Consider d-Euclidean space $\boldsymbol{R}^d$. Is it possible to define a set function (→**a19.9**) m defined on every bounded set $A \in \boldsymbol{R}^d$ such that
(1) The d-unit cube has value 1.

---

[285]However, the definition of the Feynman path integral is too delicate to be discussed in the proper integration theory.

[286]There is at least one problem in which the choice of $\mathcal{B}$ is crucial. This is the *first digit problem.* The first significant digits of a table of natural phenomenon such as the height of mountains do not distribute uniformly: 1 appears much more often than 9. Why is this so? A conclusive mathematical explanation was given recently: T P Hill, The Significant-digit Phenomenon, Am. Math. Month. April 1995, p322. If we apparently need a uniform probability on an infinite space (in this case $[0, \infty)$), the choice of $\mathcal{B}$ seems to be the key (→**a19.17**).

(2) Congruent sets have the same value,
(3) $m(A \cup B) = m(A) + m(B)$ if $A \cap B = \emptyset$, and
(4) σ-additive
?
This is called *Lebesgue's measure problem.*

a19.21 Hausdorff and non-measurable set. Hausdorff demonstrated in 1914 for any d there is no such m satisfying (1)-(4) of **a19.20**. Then, Hausdorff asked in 1914 what if we drop the condition (4). He showed that m does not exist for $d \geq 3$.[287] He showed this by constructing a partition of 2-sphere into sets A, B, C, D such that A, B, C and $B \cup C$ are all congruent and D is countable (→**15.6**). Thus if m existed, then we had to conclude $3 = 2$. Therefore, we must admit non-measurable sets.[288]

[287] Banach demonstrated in 1923 that there is a solution for $d = 1$ and for $d = 2$.
[288] under the current popular axiomatic system ZF + C.