1

A Panorama of Lebesgue Integration

The idea of measuring area and volume by infinitesimal (exhaustion) methods was already known to the ancient Greeks. This may be seen as the first example of 'integration'. The precursor of our modern notion of integration begins with the creation of the infinitesimal calculus by Newton and Leibniz. For Newton, the derivative was the primary operation of calculus and the integral was just the primitive, i.e. the antiderivative. Leibniz followed a more geometric approach, defining the integral as a sum of infinitesimal quantities which represent the area below the graph of a curve, thus establishing the integral as an object in its own right. Of course, both Newton and Leibniz were describing essentially the same object, and the history of integration is, in some sense, the attempt to reconcile both definitions. A short overview of the early history of integration is given in Section 1.11 at the end of this chapter. For us, the modern theory of integration starts in the year 1854 with Riemann's Habilitationsschrift [143].

1.1 Modern Integration. 'Also zuerst: Was hat man unter $\int_a^b f(x)\,dx$ zu verstehen?'[1]

Riemann's answer to (t)his question is the following definition [143, Section 4]:

Definition 1.1 A bounded function $f : [a,b] \to \mathbb{R}$ defined on a compact interval $[a,b] \subseteq \mathbb{R}$ is **integrable** (in the sense of Riemann) if the limit

$$\int_a^b f(x)\,dx = \lim_{|\Pi|\to 0} \sum_{i=1}^{n} (x_i - x_{i-1}) f(\xi_i), \qquad |\Pi| := \max_{1\leqslant i\leqslant n} |x_i - x_{i-1}|, \tag{1.1}$$

taken along all finite partitions $\Pi = \{a = x_0 < x_1 < \cdots < x_n = b\}$ and for any choice of intermediary points $\xi_i \in [x_{i-1}, x_i]$ exists and is finite.

[1] Riemann [143, p. 239] – To begin with: What is the meaning of $\int_a^b f(x)\,dx$?

Riemann immediately gives two necessary and sufficient conditions for the convergence of (1.1), cf. [MIMS, p. 443] for a modern proof. Denote by Π a finite partition of $[a,b]$ and write $D_i := \sup_{[x_{i-1},x_i]} f - \inf_{[x_{i-1},x_i]} f$ for the oscillation of a function f in the ith partition interval $[x_{i-1}, x_i]$.

(R1) The limit in (1.1) exists if, and only if, for all finite partitions Π of $[a,b]$,

$$\lim_{|\Pi|\to 0} \sum_{i=1}^{n} D_i \cdot |x_i - x_{i-1}| = 0.$$

(R2) The limit in (1.1) exists if, and only if,

$$\forall \epsilon > 0,\ \sigma > 0 \quad \exists \delta > 0 \quad \forall \Pi,\ |\Pi| \leqslant \delta \quad \sum_{i\,:\,D_i > \sigma} |x_{i-1} - x_i| < \epsilon.$$

In retrospect, Riemann's condition (R2) marks the beginning of the study of outer (Lebesgue) measure. We will see in Theorem 1.28 below that a bounded function f is Riemann integrable if, and only if, the set of its discontinuity points is a Lebesgue null set.

From (R2) it is clear that the Riemann integral is capable of dealing with functions which are discontinuous on a (countable) dense subset. This fact was already illustrated by Riemann in [143] using the function

$$f(x) = \sum_{n=1}^{\infty} \frac{h(nx)}{n^2}, \qquad h(x) = \begin{cases} x - k, & \text{if } x \in \left(k - \frac{1}{2}, k + \frac{1}{2}\right),\ k \in \mathbb{Z}, \\ 0, & \text{if } x = k \pm \frac{1}{2},\ k \in \mathbb{Z}, \end{cases}$$

which is discontinuous on the set $Q = \{p/(2n)\,;\ \gcd(p, 2n) = 1\}$; see Fig. 3.3. Hankel [73, pp. 199–200] observed that f is an example of a function such that

$$F(x) := \int_0^x f(t)\,dt$$

is continuous, but $F'(x) = f(x)$ fails if $x \in Q$, i.e. F is not a primitive of f [☞ 3.7].

After the publication of Riemann's 1854 thesis in 1867, his definition of the integral was widely accepted, and it is still one of the most important and widely used notions of integration. The presentation was quickly streamlined, notably by the introduction of upper and lower sums and integrals which make Riemann's criterion (R1) more tractable.

Definition 1.2 (Thomae [183], Darboux [37], Volterra [195]) For a bounded function $f : [a,b] \to \mathbb{R}$ we call

$$S_\Pi[f] := \sum_{x_{i-1},x_i \in \Pi} m_i \cdot (x_i - x_{i-1}), \qquad m_i := \inf_{x \in [x_{i-1},x_i]} f(x),$$

$$S^{\Pi}[f] := \sum_{x_{i-1},x_i\in\Pi} M_i \cdot (x_i - x_{i-1}), \qquad M_i := \sup_{x\in[x_{i-1},x_i]} f(x),$$

the **lower** and **upper Darboux sums** and

$$\underline{\int_a^b} f(x)\,dx := \sup_{\Pi\subseteq[a,b]} S_{\Pi}[f] \quad \text{and} \quad \overline{\int_a^b} f(x)\,dx := \inf_{\Pi\subseteq[a,b]} S^{\Pi}[f]$$

(sup and inf range over all finite partitions Π of $[a,b]$) the **lower** and **upper Riemann–Darboux integrals**.

Using the lower and upper integrals we can show the following integrability criterion.

Theorem 1.3 ([MIMS, p. 443]) *A bounded function $f : [a,b] \to \mathbb{R}$ is Riemann integrable if, and only if,*

$$-\infty < \underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx < \infty.$$

The common (finite) value is the Riemann integral $\int_a^b f(x)\,dx$.

The development of the Riemann integral and the concept of a function go hand in hand. Up to Cauchy, functions were (implicitly) thought to be smooth, after Cauchy to be continuous; from 1867, Riemann integrable functions were seen to be the most general and still reasonable functions. But soon there were first examples of non-Riemann integrable functions, and other shortcomings of the Riemann integral were discovered:

1° The rather limited scope of Riemann integrable functions. The (proper) Riemann integral makes sense only on bounded sets and for bounded functions [☞ 3.2, 3.3], it behaves badly under compositions [☞ 3.11] and there are rather natural and simple non-integrable functions [☞ 3.4, 3.5].

2° If the Riemann integral is extended to two dimensions, the familiar formula

$$\iint_{[a,b]\times[c,d]} f(x,y)\,dx\,dy = \int_a^b \left[\int_c^d f(x,y)\,dy\right] dx = \int_c^d \left[\int_a^b f(x,y)\,dx\right] dy$$

may become senseless since some, or all, of the one-dimensional integrals might not exist [☞ 3.20].

3° Riemann's theory does not fix the difference between integral and primitive. There are integrable functions f such that $F(x) = \int_a^x f(t)\,dt$ is not everywhere differentiable, i.e. not a proper primitive. Worse, there are everywhere differentiable functions F whose derivative F' is not integrable [☞ 14.2, 14.3].

4° The Riemann integral behaves rather badly if one wants to interchange limits and integrals. Among other pathologies, one can construct a uniformly bounded sequence of Riemann integrable functions $(f_n)_{n\in\mathbb{N}}$ on $[0,1]$ such that $\lim_{n\to\infty} f_n(x) = f(x)$ but f is not Riemann integrable [☞ 3.13, 3.14, 3.16].

1.2 The Idea Behind Lebesgue Integration

Part of the problem with Riemann's definition is that the approximation procedure used in (1.1) is based on given partitions of the domain $[a,b]$ of the function $f : [a,b] \to \mathbb{R}$, i.e. these partitions need not relate to the behaviour of f.

Lebesgue's idea in [100, 101] is to split the range $f([a,b])$ of a bounded function $f : [a,b] \to \mathbb{R}$ into equal intervals, say $J_1, \dots, J_k$, and to determine those sets $I_1, \dots, I_k \subseteq [a,b]$ such that $I_i = f^{-1}(J_i)$. The corresponding approximations of the integral would be

$$U = \sum_{i=1}^{k} |I_i| \cdot \sup J_i \quad \text{and} \quad L = \sum_{i=1}^{k} |I_i| \cdot \inf J_i, \tag{1.2}$$

where $|A|$ denotes the total length of the set A. If we choose an equidistant partitioning of mesh δ, the value of the upper approximation is $U = L+\delta\cdot|[a,b]| = L+\delta\cdot(b-a)$, i.e. it is enough to restrict one's attention to L. Notice that the resulting partition of the domain depends on f. Before we give proper definitions and discuss the implications of this approach, let us consider a simple example.

Example 1.4 Consider an oscillating periodic function, e.g. $f(x) = \sin^2(n\pi x)$ with $n \in \mathbb{N}$ and $x \in [0,1)$, cf. Fig. 1.1. Using the relation $\sin^2\alpha = \frac{1}{2}(1-\cos 2\alpha)$ it is easy to determine the integral of f,

$$\int_0^1 \sin^2(n\pi x)\,dx = \frac{1}{2}\int_0^1 (1-\cos(2n\pi x))\,dx = \frac{1}{2}\int_0^1 dx = \frac{1}{2},$$

but the upper and lower Darboux sums for an equidistant partition of $[0,1]$, $\Pi = \{0 = x_0 < \dots < x_k = 1\}$, with mesh $|\Pi| = \frac{1}{k} \geqslant \frac{1}{n}$ $\Big($or a general partition such that $\min_i(x_i - x_{i-1}) \geqslant \frac{1}{n}\Big)$ are easily seen to be 1 and 0, respectively.

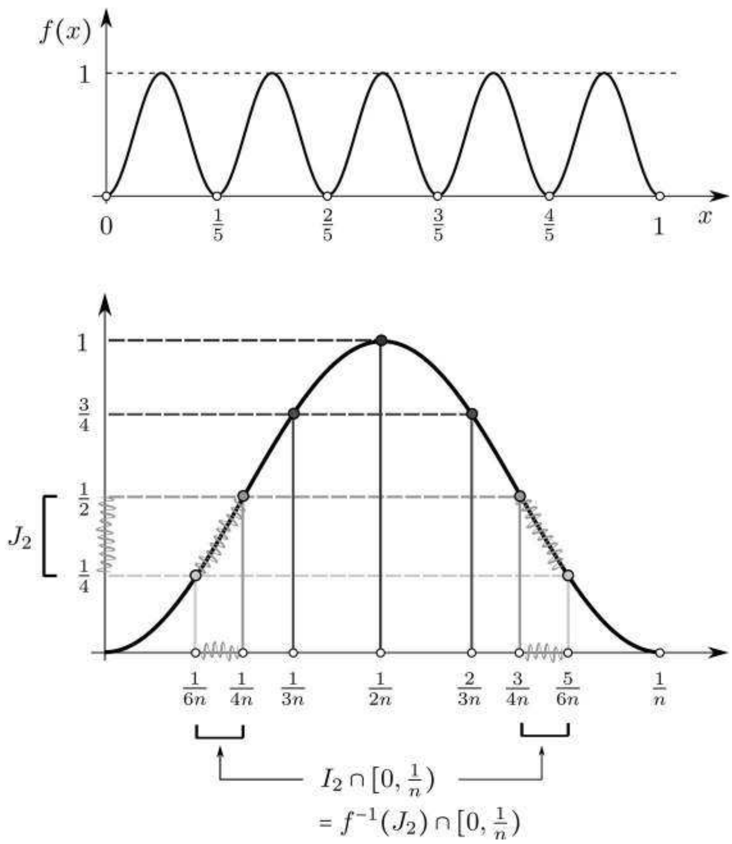

Figure 1.1 The oscillating periodic function $f(x) = \sin^2(n\pi x)$ for $n = 5$ (upper panel) and the choice of the domain partition I_k for an equidistant range partition J_k (lower panel).

However, Lebesgue's approach using $k = 4$ and $J_i = \left[\frac{i-1}{k}, \frac{i}{k}\right)$ gives, over the first period,

$$I_i \cap \left[0, \tfrac{1}{n}\right) = \left\{0 \leqslant x < \tfrac{1}{n}\,;\ \tfrac{i-1}{k} \leqslant f(x) < \tfrac{i}{k}\right\} = [x_{i-1}, x_i) \cup (x_{8-i}, x_{8-(i-1)}]$$

with $x_0 = 0$, $x_1 = 1/6n$, $x_2 = 1/4n$, $x_3 = 1/3n$, $x_4 = 1/2n$, $x_5 = 2/3n$, $x_6 = 3/4n$, $x_7 = 5/6n$ and $x_8 = 1/n$; see Fig. 1.1 (lower panel). Thus, in $[0, 1/n)$ we get

$$\sum_{i=1}^{4} \frac{i-1}{4} \left[(x_i - x_{i-1}) + (x_{8-(i-1)} - x_{8-i})\right] = \left(\frac{1}{24} + \frac{1}{12} + \frac{1}{4}\right)\frac{1}{n} = \frac{3}{8n}.$$

Since there are n periods in $[0, 1]$, we have $L = \frac{3}{8}$ (and $U = \frac{3}{8} + \frac{1}{4} = \frac{5}{8}$). This is already a reasonable approximation of the true value $\frac{1}{2}$ and, if one uses $\frac{1}{2}(U+L)$, it even happens to be the exact value.

This example makes it clear that Lebesgue's approach is better suited to deal with (rapidly) oscillating integrands, in particular, when the oscillations approach a condensation point as is the case for $x \longmapsto \sin^2 \frac{1}{x}$ as $x \to 0$.

1.3 Lebesgue Essentials – Measures and σ-Algebras

Let us recast Lebesgue's approximation of a function $f \geqslant 0$ from below by slicing its range **horizontally** as shown in Fig. 1.2. The level sets

$$A_k^n := \begin{cases} \{k2^{-n} \leqslant f < (k+1)2^{-n}\} & \text{for } k = 0, 1, 2, \dots, n2^n - 1, \\ \{f \geqslant n\} & \text{for } k = n2^n, \end{cases}$$

can be used to define step functions

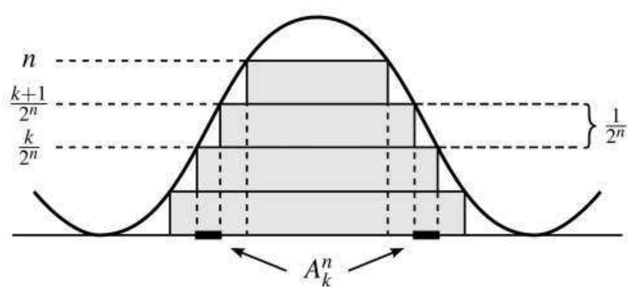

Figure 1.2 The function f sits like a 'Mexican hat' (a sombrero) over the approximating simple functions.

$$\phi_n(x) := \sum_{k=0}^{n2^n} k2^{-n} \mathbb{1}_{A_k^n}(x)$$

which approximate f from below. Coming from below has the advantage that f need not be bounded; instead, we use a moving cut-off level n which kicks in on the set $A_{n2^n}^n$. From Fig. 1.2 we see that

(i) $0 \leqslant \phi_n \leqslant \phi_{n+1} \leqslant f$ and $\phi_n \uparrow f$;
(ii) $|\phi_n(x) - f(x)| \leqslant 2^{-n}$ if $x \in \{f < n\}$; in particular, if f is bounded, the sequence ϕ_n approximates f uniformly.

We are interested in the nature of the level sets A_k^n. Property (i) requires that we are able to subdivide the level sets A_k^n finitely often. If we want to integrate the sum of two functions f, g, the level sets of $f + g$ will be expressed through finite unions and intersections of the level sets of f and g.

Therefore, the level sets form a family of sets which is closed if we repeat the usual set operations (intersection, union, taking complements and differences) finitely or – for limits – countably infinitely often. This requirement leads naturally to the notion of a σ-algebra.

Definition 1.5 Let $X \neq \emptyset$ be any set, and denote by $\mathscr{P}(X)$ its power set. A σ-**algebra** $\mathscr{A}$ on X is a family of subsets of X with the following properties:

$$X \in \mathscr{A}, \tag{Σ_1}$$

$$A \in \mathscr{A} \;\Rightarrow\; A^c \in \mathscr{A}, \tag{Σ_2}$$

$$(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{A} \;\Rightarrow\; \bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}. \tag{Σ_3}$$

Because of (Σ_1) and (Σ_2) we have $\emptyset \in \mathscr{A}$, and using $A_1 \cup A_2 \cup \emptyset \cup \ldots$ in (Σ_3) shows that $\mathscr{A}$ is stable under finite unions. With de Morgan's laws we get that $\mathscr{A}$ is also stable under finite and countably infinite intersections and this is also true for differences as $A \setminus B = A \cap B^c$ is a combination of complementation and intersection.

The second ingredient needed for the construction of the integral is a 'gauge' for the size of the level sets A_k^n. In Example 1.4 we naively took the 'length' of the interval and there was no problem since the level sets were relatively simple. In the general case we need a function defined on all possible level sets which is compatible with (countably often repeated) set operations. This is the rationale for the following definition.

Definition 1.6 Let $X \neq \emptyset$ be any set. A (positive) **measure** is a set function $\mu : \mathscr{A} \to [0, \infty]$ satisfying

$$\mathscr{A} \text{ is a } \sigma\text{-algebra on } X, \tag{M_0}$$

$$\mu(\emptyset) = 0, \tag{M_1}$$

$$(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{A} \text{ pairwise disjoint} \;\Rightarrow\; \mu\left(\dot{\bigcup_{n\in\mathbb{N}}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n). \tag{M_2}$$

The pair $(X, \mathscr{A})$ is called a **measurable space** and $(X, \mathscr{A}, \mu)$ is called a **measure space**. The measure space is called **finite**, if $\mu(X) < \infty$, and σ-**finite**, if there exists a sequence $(F_n)_{n\in\mathbb{N}} \subseteq \mathscr{A}$ such that $X = \bigcup_{n\in\mathbb{N}} F_n$ and $\mu(F_n) < \infty$. A set $A \in \mathscr{A}$ is often called a **measurable set**.

The requirements (M_0)–(M_2) lead to a rich family of set functions with many further properties; see [MIMS, pp. 24, 28]. For example, if $A, B, A_n, B_n \in \mathscr{A}$:

(a) $A \cap B = \emptyset \;\Rightarrow\; \mu(A \mathbin{\dot\cup} B) = \mu(A) + \mu(B)$ (additive)

(b) $A \subseteq B \;\Rightarrow\; \mu(A) \leqslant \mu(B)$ (monotone)

(c) $A \subseteq B,\ \mu(A) < \infty \;\Rightarrow\; \mu(B \setminus A) = \mu(B) - \mu(A)$

(d) $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ (strongly additive)

(e) $\mu(A \cup B) \leqslant \mu(A) + \mu(B)$ (subadditive)

(f) $A_n \uparrow A \;\Rightarrow\; \mu(A) = \sup_n \mu(A_n) = \lim_n \mu(A_n)$ (continuous from below)

(g) $B_n \downarrow B,\ \mu(B_1) < \infty \;\Rightarrow\; \mu(B) = \inf_n \mu(B_n) = \lim_n \mu(B_n)$ (continuous from above)

(h) $\mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) \leqslant \sum_{n\in\mathbb{N}} \mu(A_n)$. ($\sigma$-subadditive)

Example 1.7 Here are some of the most commonly used measures and σ-algebras. Unless otherwise indicated, $(X, \mathscr{A}, \mu)$ is an arbitrary measure space.

σ-**Algebra** $\mathscr{A}$	**Typical measure** on $(X, \mathscr{A})$
(a) The **indiscrete σ-algebra**: $\{\emptyset, X\}$ – this is the smallest possible σ-algebra on X.	▶ The **trivial measure** $\tau(\emptyset) = 0$ and $\tau(X) = \infty$.
(b) The **discrete σ-algebra**: $\mathscr{P}(X)$ – this is the largest possible σ-algebra on X.	▶ As a rule of thumb, rich σ-algebras admit only poor (i.e. simple) measures: $\mathscr{P}(X)$ can support the **trivial measure** from (a), the **counting measure** $\zeta(A) = \#A$, or **Dirac's delta function** (**point mass**) at $x \in X$ $$\delta_x(A) = \begin{cases} 0, & \text{if } x \notin A, \\ 1, & \text{if } x \in A. \end{cases}$$
(c) The **trivial σ-algebra**: $\mathscr{T}_\mu = \{A \in \mathscr{A}; \mu(A) = 0 \text{ or } \mu(A^c) = 0\}$.	▶ This construction works for every measure space $(X, \mathscr{A}, \mu)$.
(d) The **co-countable σ-algebra**: $\{A \subseteq X;\ \#A \leqslant \#\mathbb{N} \text{ or } \#A^c \leqslant \#\mathbb{N}\}$ on an uncountable set X [☞ Example 4A].	▶ The **co-countable** (probability) measure $$\mu(A) = \begin{cases} 0, & \text{if } A \text{ is countable}, \\ 1, & \text{if } A^c \text{ is countable}, \end{cases}$$ [☞ Example 5A].
(e) The **trace σ-algebra**: Let $E \subseteq X$. $\mathscr{A}_E = E \cap \mathscr{A} := \{E \cap A;\ A \in \mathscr{A}\}$.	▶ If $E \in \mathscr{A}$, the restriction $\mu\|_E(A) := \mu(E \cap A)$ is a measure on the trace measurable space $(E, \mathscr{A}_E)$; [☞ 5.9] if $E \notin \mathscr{A}$.
(f) The **pre-image σ-algebra**: Let $\phi : X \to X'$ be any map and $\mathscr{A}'$ a σ-algebra on X'. $$\phi^{-1}(\mathscr{A}') = \{\phi^{-1}(A');\ A' \in \mathscr{A}'\}.$$	▶ If μ is a measure on $(X, \mathscr{A})$, then $$\mu'(A') := \mu(\phi^{-1}(A'))$$ is called the **image measure** or **push-forward measure** of μ under ϕ. Notation: $\mu\circ\phi^{-1}$, $\phi_*\mu$ or $\phi(\mu)$.

(g) The **σ-algebra generated by a set** $F \subseteq X$: This is the smallest σ-algebra on X containing the set F: $\sigma(F) = \{\emptyset, F, F^c, X\}$.

(h) The **σ-algebra generated by a family of sets** $\mathscr{F}$: This is the smallest σ-algebra containing the family $\mathscr{F}$: $\sigma(\mathscr{F}) = \bigcap\{\mathscr{B}\,;\ \mathscr{F} \subseteq \mathscr{B}, \mathscr{B}\ \sigma\text{-algebra}\}$.

(i) Let $\phi_i : X \to X_i$, $i \in I$ be arbitrarily many mappings and assume that $\mathscr{A}_i$ is a σ-algebra in X_i. The **σ-algebra generated by the family of mappings** $(\phi_i)_{i\in I}$, $\sigma(\phi_i, i \in I) = \sigma\left(\bigcup_{i\in I} \phi_i^{-1}(\mathscr{A}_i)\right)$, is the smallest σ-algebra that makes all ϕ_i measurable (see Definition 1.8 further on).

(j) The **completed σ-algebra**: Let $\mathscr{F} \subseteq \mathscr{A}$ be a (not necessarily proper) sub-σ-algebra,

$$\mathscr{N}_\mu = \{N \in \mathscr{A}\,;\ \mu(N) = 0\}$$

the family of all measurable null sets, and

$$\mathscr{N}_\mu^* = \{N^* \subseteq X\,;\ \exists N \in \mathscr{N}_\mu,\ N^* \subseteq N\}$$

the family of all subsets of measurable null sets.
The **completion** of $\mathscr{F}$ is the σ-algebra $\mathscr{F}^* := \sigma(\mathscr{F}, \mathscr{N}^*)$. One can show that

$$\begin{aligned}\mathscr{F}^* &= \{F \vartriangle N^*\,;\ F \in \mathscr{F},\ N^* \in \mathscr{N}^*\}\\ &= \big\{F^*\,;\ \exists A, B \in \mathscr{F},\ A \subseteq F^* \subseteq B,\\ &\qquad\qquad \mu(B \setminus A) = 0\big\}.\end{aligned}$$

▶ The **completion** $\bar\mu$ of the measure μ (defined on $\mathscr{F}$) is the measure $\bar\mu$ on the measurable space $(X, \mathscr{F}^*)$ given by

$$\bar\mu(F^*) := \frac{1}{2}(\mu(A) + \mu(B)), \quad F^* \in \mathscr{F}^*,$$

where the sets $A, B \in \mathscr{F}$ are such that $\mu(B \setminus A) = 0$ and $A \subseteq F^* \subseteq B$. The former ensures that $\bar\mu$ is well-defined, i.e. independent of the choice of the sets A and B.
Since $\mathscr{F} \subseteq \mathscr{F}^*$, $\bar\mu$ is an extension of μ.

(k) Let X be a topological space and $\mathscr{O}$ the family of all open sets. The **Borel** or **topological σ-algebra** is the σ-algebra generated by the open sets $\mathscr{B}(X) = \sigma(\mathscr{O})$.
Since a set is open if its complement is closed, $\mathscr{B}(X)$ is also generated by the closed sets. If X is a metric space which is the union of countably many compact sets $X = \bigcup_{n\in\mathbb{N}} K_n$ (e.g. if X is locally compact and separable), then $\mathscr{B}(X)$ is also generated by the compact sets [☞4.15, 4.16].

(l) The Borel sets in $\mathbb{R}^d$, $\mathscr{B}(\mathbb{R}^d)$, are generated by any of the following families: The open sets, the closed sets, the compact sets, the open balls $B_r(q)$ (radius $r \in \mathbb{Q}^+$, centre $q \in \mathbb{Q}^d$), the rectangles $\bigtimes_{i=1}^d [a_i, b_i)$ (with rational $a_i, b_i \in \mathbb{Q}$).

▶ Most measures used in analysis are defined on the Borel sets (or their completion, the Lebesgue sets, cf. Example (n)). The prime example of a measure on $\mathscr{B}(\mathbb{R}^d)$ is d-dimensional Lebesgue measure λ^d. Since the structure of the Borel sets is quite complicated, one defines λ^d on a sufficiently rich generator

$$\lambda^d\left(\bigtimes_{i=1}^d [a_i, b_i)\right) = \prod_{i=1}^d (b_i - a_i).$$

We will see in Theorem 1.40 that this characterizes λ^d uniquely.

(m) If $A \subseteq \mathbb{R}^d$, then $\mathscr{B}(A)$ is the Borel σ-algebra which is generated by the relatively open subsets of A. It is not hard to see that $\mathscr{B}(A)$ coincides with the trace σ-algebra $A \cap \mathscr{B}(\mathbb{R}^d)$.

▶ Use λ_A^d, the trace of Lebesgue measure λ^d on the trace-σ-algebra; see Example (e).

(n) The **Lebesgue σ-algebra** or **Lebesgue sets** $\mathscr{L}(\mathbb{R}^d)$ are the completion, see Example (j), of the Borel sets with respect to Lebesgue measure.

▶ Use the completion $\bar{\lambda}^d$ of λ^d; see Example (j).

(o) The **product σ-algebra** $\mathscr{A} \otimes \mathscr{B}$ is the σ-algebra $\sigma(\mathscr{A} \times \mathscr{B})$ generated by all generalized 'rectangles', i.e. sets of the form $A \times B \in \mathscr{A} \times \mathscr{B}$ with $A \in \mathscr{A}$ and $B \in \mathscr{B}$.

▶ Let $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$ be σ-finite measure spaces. Similar to the construction of Lebesgue measure, the product measure ρ is defined first on the sets $A \times B \in \mathscr{A} \times \mathscr{B}$ of a generator,

$$\rho(A \times B) := \mu(A)\nu(B),$$

and from the general theory it is known that this characterizes ρ on $\mathscr{A} \otimes \mathscr{B}$, cf. Theorem 1.33.

1.4 Lebesgue Essentials – Integrals and Measurable Functions

Let us return to the original problem of integrating a function. A real-valued function $f : X \to \mathbb{R}$ whose level sets $\{a \leqslant f < b\}$ are in a σ-algebra $\mathscr{A}$ on X is called measurable. The observation

$$\{a \leqslant f < b\} = \{f \geqslant a\} \cap \{f < b\} = f^{-1}([a, \infty)) \cap f^{-1}((-\infty, b))$$

explains the following slightly more general definition.

Definition 1.8 Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be two measurable spaces. A mapping $f : X \to Y$ is called $\mathscr{A}/\mathscr{B}$ **measurable**, if

$$\forall B \in \mathscr{B} \ : \ f^{-1}(B) \in \mathscr{A}. \tag{1.3}$$

If Y is a topological space equipped with its Borel sets, then measurable functions f are also called **Borel maps** or **Borel functions**.

Remark 1.9 (a) If $\mathscr{B}$ is generated by some family $\mathscr{H}$, then (1.3) is equivalent to the requirement that $f^{-1}(H) \in \mathscr{A}$ for all $H \in \mathscr{H}$. In particular, if we consider $\mathbb{R}$ equipped with the Borel σ-algebra $\mathscr{B}(\mathbb{R})$, then measurability of $f : X \to \mathbb{R}$ means that $\{f \leqslant a\} \in \mathscr{A}$ for all $a \in \mathbb{R}$ or $\{f > b\} \in \mathscr{A}$ for all $b \in \mathbb{R}$; see [MIMS, pp. 54, 60].

Since the pre-image of an open set under a continuous function is open, continuous functions are always Borel measurable.