A SHORT STORY OF MEASURE THEORY

1. Introduction

The Riemann integral is founded on the following idea: divide up the domain of a function $f : [a,b] \longrightarrow \mathbb{R}$ into subintervals, estimate f from above and below on each interval, and approximate the integral of f by the *upper* and *lower sums*—the summation of the widths of the intervals times the upper and lower estimates on each. The limit over partitions of $[a,b]$ of these two approximations, should they exist and agree, is declared to be the integral of f.

Sadly, this definition of the integral lacks some desirable properties. In particular, the space of absolutely integrable functions is not complete—a sequence of functions which is Cauchy in the norm $\|f\|_1 = \int_a^b |f(x)|\, dx$ need not converge to a Riemann-integrable function.

As a remedy to such deficiencies, the Lebesgue integral is founded on a different idea: namely, divide up the *range* of f into subintervals, and approximate the integral by the summation of the lower endpoints times the *volume*, or *measure*, of the interval's preimage under f. To make this idea precise, we require

(1) a notion of measure for appropriate sets,
(2) a class of "measurable" functions which can be so approximated, and
(3) a definition of the integral of a measurable function on a measurable set.

As with most ideas in math, it is possible to develop this in a fairly general setting. In this note, we outline this development in the general setting, with particular mention of the Lebesgue measure on $\mathbb{R}$. As this is a "story", not a course in measure theory, you are meant to provide your own proofs (or look them up). Most are straightforward, if tedious. Folland's *Real Analysis* is the treatment we mostly follow here.

2. Measures

It is an unfortunate fact that we often cannot assign a coherent measure to *all* subsets of a given space. We can, however, require some nice conditions of those sets to be 'measured'.

A collection $\mathcal{A} \subseteq \mathcal{P}(X)$ of subsets of X is an **algebra** if it contains $\emptyset$ and is closed under pairwise (hence finite) union and complements:

$$A_1,\ A_2 \in \mathcal{A} \implies A_1 \cup A_2,\ A_1^c \in \mathcal{A}.$$

$\mathcal{A}$ is a **σ-algebra** if in addition it is closed under *countable* unions:

$$\{A_n : n \in \mathbb{N}\} \subset \mathcal{A} \implies \bigcup_n A_n \in \mathcal{A}.$$

It follows that $\mathcal{A}$ is likewise closed under countable intersections.

Often we start with a collection of sets of interest, and take the smallest σ-algebra generated by these. If X is a topological space, the **Borel** σ-algebra, $\mathcal{B}_X$, is the one generated by all open (equivalently closed) sets.

Proposition 2.1. *The Borel σ-algebra on $\mathbb{R}$ is equivalently generated by any of the following collections of subsets:*

$$\{(a,b) : a,b \in \mathbb{R}\} \qquad \{[a,b) : a,b \in \mathbb{R}\} \qquad \{(a,b] : a,b \in \mathbb{R}\}$$
$$\{[a,b] : a,b \in \mathbb{R}\} \qquad \{(a,\infty) : a \in \mathbb{R}\} \qquad \{[a,\infty) : a \in \mathbb{R}\}$$
$$\{(-\infty,a) : a \in \mathbb{R}\} \qquad \{(-\infty,a] : a \in \mathbb{R}\}$$

In measure theory it is often useful to work with the **extended real numbers** $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, a 2-point compactification of $\mathbb{R}$ with the obvious topology (i.e., $(a, \infty]$ and $[-\infty, b)$ are open for all $a, b \in \mathbb{R}$) and total order. Then $\mathcal{B}_{\overline{\mathbb{R}}}$ is generated by the collection $\{[a, \infty]\}$, for instance.

Let $\mathcal{A}$ be a σ-algebra on a set X. A **measure** on $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \longrightarrow [0, \infty]$ satisfying

(M1) $\mu(\emptyset) = 0$, and

(M2) (**Countable additivity**) if $\{A_n : n \in \mathbb{N}\}$ are mutually disjoint then $\mu\Big(\bigcup_{n=1}^\infty A_n\Big) = \sum_{n=1}^\infty \mu(A_n)$.

The following properties follow easily from the definition.

Proposition 2.2. *Let μ be a measure on $(X, \mathcal{A})$. Then*

(M4) **(Monotonicity)** $A \subset B \implies \mu(A) \leq \mu(B)$,
(M5) **(Countable sub-additivity)** $\mu\big(\bigcup_{n=1}^\infty A_n\big) \leq \sum_{n=1}^\infty A_n$,
(M6) **(Continuity from below)** $A_1 \subset A_2 \subset \cdots \implies \mu\big(\bigcup_n A_n\big) = \lim_n \mu(A_n)$,
(M7) **(Continuity from above)** $A_1 \supset A_2 \supset \cdots \implies \mu\big(\bigcap_n A_n\big) = \lim_n \mu(A_n)$.

We defer the existence and construction of useful measures until §6.

3. Measurable functions

Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be spaces with σ-algebras (aka "measurable spaces"). A function $f : X \longrightarrow Y$ is **measurable** if

$$B \in \mathcal{B} \implies f^{-1}(B) \in \mathcal{A}.$$

In particular, a (possibly extended) real-valued function $f : X \longrightarrow \overline{\mathbb{R}} = (\overline{\mathbb{R}}, \mathcal{B}_{\overline{\mathbb{R}}})$ is measurable if and only if $f^{-1}([a, \infty]) \in \mathcal{A}$ for all $a \in \mathbb{R}$. The set of measurable $\overline{\mathbb{R}}$-valued functions has particularly nice limit properties:

Proposition 3.1. *Let $\{f_n\}$ be a sequence of $\overline{\mathbb{R}}$-valued measurable functions on $(X, \mathcal{A})$. Then*

$$g_1(x) = \sup_n f_n(x), \quad g_2(x) = \inf_n f_n(x),$$
$$g_3(x) = \limsup_n f_n(x), \quad \textit{and } g_4(x) = \liminf_n f_n(x)$$

are all measurable. In particular if the sequence converges pointwise then $\lim_n f_n$ *is measurable.*

A **step function** is a measurable function given by a finite linear combination

$$\phi = \sum a_k \chi_{A_k}, \quad A_k \in \mathcal{A}, \quad a_k \in \mathbb{C},$$

where χ_A denotes the **indicator function**

$$\chi_A(x) = \begin{cases} 1 & x \in A, \\ 0 & x \notin A. \end{cases}$$

(Note that the a_k are not allowed to be infinite, and that by requiring the A_k to be disjoint, we can arrange for a unique representation of ϕ.) For a step function, the definition of the integral is almost obvious; however we run into issues whenever some of the A_k have infinite measure.

Initially then, we restrict attention to the *positive* measurable functions:

$$L^+ = L^+(X) = \{f : X \longrightarrow [0, \infty] \text{ measurable}\}.$$

Proposition 3.2. *$f \in L^+$ if and only if there is an increasing sequence of positive step functions $\{\phi_n\}$ such that $\phi_n \longrightarrow f$ pointwise.*

4. The integral

For a positive step function $\phi = \sum a_k \chi_{A_k}$, $a_k \in [0, \infty)$, the integral is defined by

$$\int \phi \, d\mu = \sum a_k \mu(A_k), \tag{1}$$

with the convention that $0 \cdot \infty = 0$. Note that $\int \phi \, d\mu$ may have the value ∞.

Proposition 4.1. *The integral (on step functions) has the following properties:*
(a) $\int (\phi + \psi) \, d\mu = \int \phi \, d\mu + \int \psi \, d\mu$.
(b) $\int c\phi \, d\mu = c \int \phi \, d\mu$, $c \in [0, \infty)$.
(c) *If* $\phi \le \psi$, *then* $\int \phi \, d\mu \le \int \psi \, d\mu$.
(d) $A \longmapsto \int_A \phi \, d\mu = \sum a_k \mu(A \cap A_k)$ *is a measure on* $\mathcal{A}$.

For a positive measurable function $f \in L^+$, the integral is defined by estimating from below by step functions:

$$\int f \, d\mu := \sup \left\{ \int \phi \, d\mu : 0 \le \phi \le f, \ \phi \text{ step} \right\}$$

This extends (1) when f is a step function, since the supremum is then achieved by $\phi = f$.

Theorem 4.2 (Monotone Convergence Theorem)**.** *Let* $\{f_n\}$ *be a sequence in* L^+ *such that* $f_n \le f_{n+1}$ *for all* n *and* $f_n \to f \in L^+$. *Then*

$$\int f \, d\mu = \lim_n \int f_n \, d\mu.$$

Instead of taking the supremum over all step functions bounded by $f \in L^+$, we can thus represent each f by a pointwise increasing limit of step functions by Proposition 3.2 and exchange limits and integral signs by Theorem 4.2.

Corollary 4.3. *Proposition 4.1 extends to the integral on* L^+*; in fact the latter is countably additive:* $\int \sum_{n=1}^{\infty} f_n \, d\mu = \sum_{n=1}^{\infty} \int f_n \, d\mu$.

Note that, without the monotone increasing hypothesis, Theorem 4.2 may fail. For instance, $f_n = \chi_{[n,n+1]}$ and $g_n = n\chi_{[0,1/n]}$ are two sequences of step functions on $\mathbb{R}$ converging pointwise to 0, but for which $\int f_n \, dx = \int g_n \, dx = 1$ for all n. A general inequality holds however:

Corollary 4.4 (Fatou's Lemma)**.** *Let* $\{f_n\}$ *be any sequence in* L^+. *Then*

$$\int \liminf_n f_n \, d\mu \le \liminf_n \int f_n \, d\mu$$

We are tempted to suppose that $0 \le f$, $\int f \, d\mu = 0$ implies $f = 0$, but this is generally false, as can be seen already for step functions. Indeed, if $\phi = a\chi_A$ where the A has measure zero ($\mu(A) = 0$), then $\int \phi \, d\mu = 0$ even if $a \ne 0$. We say that a property that holds off of a set of measure zero holds **almost everywhere**, or **a.e.**, for short[1]

Proposition 4.5. *If* $f \in L^+$ *and* $\int f \, d\mu = 0$, *then* $f = 0$ *almost everywhere.*

Evidently we are free to alter measurable functions on a set of measure zero without altering their integrals. It follows that Theorem 4.2 holds under the relaxed condition that $f_n \nearrow f$ pointwise a.e. (hereafter we just say "$f_n \nearrow f$ a.e."), rather than pointwise everywhere.

[1]Given a measure space $(X, \mathcal{A}, \mu)$, it is technically useful to suppose that $\mathcal{A}$ contains all unions of sets of μ measure 0, which can always be arranged by enlarging $\mathcal{A}$. Then μ is said to be **complete**.

5. Integrating real and complex functions

If f is a $\mathbb{R}$-valued measurable function, then $f = f_+ - f_-$ where $f_+ = \max\{f, 0\}$ and $f_- = -\min\{f, 0\}$ are measurable (c.f. Prop. 3.1) and positive. Note that $|f| = f_+ + f_-$ is also measurable and positive. We say f is **integrable** if

$$\int |f|\, d\mu < \infty,$$

which implies that both $\int f_+\, d\mu$ and $\int f_-\, d\mu$ are finite, and we define

$$\int f\, d\mu = \int f_+\, d\mu - \int f_-\, d\mu.$$

We denote the set of real valued integrable functions by $L(X; \mathbb{R})$.

Proposition 5.1. *$L(X, \mathbb{R})$ is a vector space, $\int \cdot\, d\mu : L(X; \mathbb{R}) \longrightarrow \mathbb{R}$ is a linear functional, and $\left|\int f\, d\mu\right| \le \int |f|\, d\mu$.*

Likewise, we say a complex valued function g is integrable if $\int |g|\, d\mu < \infty$, which holds if and only if $\operatorname{Re} g$ and $\operatorname{Im} g$ are integrable real functions, and

$$\int g\, d\mu = \int \operatorname{Re} g\, d\mu + i \int \operatorname{Im} g\, d\mu.$$

Denote the set of complex valued integrable functions by $L(X; \mathbb{C})$. Proposition 5.1 extends to $L(X; \mathbb{C})$.

The workhorse limit theorem in Lebesgue integration theory is the following.

Theorem 5.2 (Lebesgue Dominated Convergence Theorem)**.** *Let $\{f_n\}$ be a sequence in $L(X; \mathbb{C})$ such that $f_n \longrightarrow f$ pointwise a.e., and suppose there exists a real valued $g \ge 0$ with $\int g\, d\mu < \infty$ and $|f_n| \le g$ for all n. Then f is integrable and*

$$\int f\, d\mu = \lim_n \int f_n\, d\mu.$$

6. Construction of measures

How do we come up with useful measures in practice? One way is to start with a putative measure defined on some collection of sets, not necessarily a σ-algebra, and try to extend it.

For example, in $\mathbb{R}^n$ we agree that the standard volume of a product of intervals $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$ is $\lambda(A) = \prod_{i=1}^n (b_i - a_i)$. This is the starting point for Lebesgue measure, and we wish to extend λ to a measure on some σ-algebra which at least contains the σ-algebra generated by such A (the latter being the Borel algebra $\mathcal{B}_{\mathbb{R}^n}$).

Let $\lambda : \mathcal{A} \longrightarrow [0, \infty]$ satisfy the conditions of a measure for some collection $\mathcal{A}$ of subsets of X, not necessarily a σ-algebra. Then for any subset $E \subset X$, we define the **outer measure** by

$$\lambda^*(E) = \inf\left\{ \sum_{n=1}^{\infty} \lambda(A_n) : E \subset \bigcup_{n=1}^{\infty} A_n,\ A_n \in \mathcal{A} \right\}$$

Then $\lambda^* : \mathcal{P}(X) \longrightarrow [0, \infty]$ is not necessarily a measure, but satisfies the weaker properties

(a)

¡++¿

7. L^p spaces