Exercise Class 9

Lebesgue Integration Theory[1]

Lebesgue Integration as an Extension Riemann Integration

We have now defined measures and what it means to be measurable (be it for sets or functions).

Let μ be the Lebesgue measure on $\mathbb{R}^n$ and $\Omega \subset \mathbb{R}^n$ a Lebesgue measurable set. We are now able to define the Lebesgue integral.

$$\int_\Omega f d\mu.$$

Just as not every set is μ-measurable, not every function is μ-integrable; the function will need to be μ-measurable. To define the integral with respect to μ, we turn to another more basic notion of integration, namely the Riemann integral

$$\int_a^b f(x)dx$$

of a Riemann-integrable function $f : [a, b] \to \mathbb{R}$. Recall that this integral (if it exists) is equal to the supremum of the lower Riemann sums

$$\int_a^b f(x)dx = \underline{\int_a^b} f(x)dx := \sup_{\substack{g \le f \\ g \text{ step function}}} \int_a^b g(x)dx,$$

where the integral of step functions (or piecewise constant functions) was defined as

$$\int_a^b g(x)dx = \sum_{i=1}^N (x_i - x_{i-1})g(x_i),$$

where the partition $\{x_i \mid i = 0, \ldots N\}$ breaks up the step function into finite linear combinations of characteristic functions χ_I of intervals I. (It is also equal to the infimum of upper Riemann sums, but for the moment we solely rely on lower integrals for brevity.)

It turns out that virtually the same definition allows us to define a *lower Lebesgue integral*

$$\underline{\int_\Omega} f d\mu$$

of any measurable function $f : \mathbb{R}^n \to [-\infty, \infty]$. One simply needs to replace the intervals I with the more general class of μ-measurable sets and thus replace piecewise constant functions with the more general class of simple functions. That is,

$$\underline{\int_\Omega} f d\mu = \sup_{\substack{g \le f \ \mu-\text{a.e.} \\ g \text{ simple function}}} \int_\Omega g d\mu.$$

As we shall see, it obeys all the basic properties one expects of an integral, such as monotonicity and linearity; we will also see that it behaves quite will with respect to

[1] Today's notes are taken from [Tao11]

limits. We will first establish *Fatou's lemma* and the *monotone convergence theorem* before proving what is essentially the cornerstone of Lebesgue integration theory — the *dominated convergence theorem.* This convergence theorem makes the Lebesgue integral (and its abstract generalizations to other measure spaces than $\mathbb{R}^n$) particularly suitable for analysis, as well as related fields that rely heavily on limits of functions, such as PDE theory, probability theory, and ergodic theory.

Remark. This is not the only route to setting up the Lebesgue integral. For instance, in [SS09] the authors proceed in four stages, by progressively integrating

1. Simple functions
2. Bounded functions supported on a set of finite measure
3. Non-negative functions
4. Integrable functions (the general case).

Another approach is to take the metric completion of the Riemann integral with respect to the L^1 metric.

The Lebesgue integral and Lebesgue measure can be viewed as *completions* of the Riemann integral and Jordan measure[2] respectively. That is, Lebesgue integration theory extends the Riemann theory: every Jordan measurable set is Lebesgue measurable, and every Riemann integrable function is Lebesgue measurable, with the measures and integrals from the two theories being compatible. Conversely, the Lebesgue theory can be approximated by the Riemann theory; as we saw in Exercise Class 8 every Lebesgue measurable set can be approximated (in the appropriate sense) by simpler sets such as open sets or elementary sets. In a similar fashion, Lebesgue measurable functions can be approximated by nicer functions, such as continuous functions. Finally, the Lebesgue theory is complete in various ways (as you will see in the chapters on L^p-spaces). The convergence theorems already hint at this completeness. Egorov's theorem is also related to that completeness.

Abstract Integrals

Note that there was no reason for us to restrict ourselves to the Lebesgue measure. The measure μ can be any Radon measure on $\mathbb{R}^n$ and the same construction yields an integral with respect to the measure μ. That integral still satisfies basic properties that we are used to, such as mononotonicity and linearity. However, by changing the measure, one does not necessarily obtain a generalization of the Riemann integral, but sometimes something completely different. We consider three such examples.

Example. We consider what happens when one considers the Dirac measure δ_z defined by

$$\delta_z(A) = \begin{cases} 1, & \text{if } z \in A, \\ 0, & \text{else} \end{cases}$$

for any measurable set A. The corresponding integral is then defined as

$$\int_A f d\delta_z = \begin{cases} f(z), & z \in A, \\ 0 & z \notin A. \end{cases}$$

[2]Strictly seen, the Jordan measure is not a measure.

Example. Another simple example is to consider μ to be the couting measure on $\mathbb{N}$. As it turns out, integration with respect to the counting measure is just summation. One obtains that

$$\int_{\mathbb{N}} f d\mu = \sum_{k=1}^{\infty} f(k).$$

Example. Lastly, recall the Lebesgue–Stieltjes measure, which is defined as the Carathéodory–Hahn extension of

$$\lambda_g([a,b)) = \begin{cases} g(b) - g(a) & \text{if } a < b, \\ 0 & \text{otherwise,} \end{cases}$$

where $g : \mathbb{R} \to \mathbb{R}$ is a non-decreasing[3] and right-continuous function, and $a, b \in \mathbb{R}$. The Lebesgue–Stieltjes integral

$$\int_a^b f d\Lambda_g$$

is then defined as the integral of f with respect to the measure Λ_g. Thus we are able to integrate functions *with respect to (or against)* functions of bounded variation. The Lebesgue–Stieltjes integral is often denoted by

$$\int_a^b f dg$$

and it finds common application in probability theory and stochastic processes, as well as in certain branches of analysis such as potential theory.

Remark. If one wants to generalize this integral to less regular functions one enters the realm of stochastic integration and rough path theory. There, one can construct integrals with respect to stochastic processes such as Brownian motion.

Almost Everywhere

In the lecture, we were able to prove monotonicity in the following sense.

Proposition. *Let $f_1, f_2 : \Omega \to [-\infty, \infty]$ be μ-integrable with $f_1 \geq f_2$ μ-a.e. Then*

$$\int_\Omega f_1 d\mu \geq \int_\Omega f_2 d\mu.$$

As a corollary, one obtains

Corollary. *Let $f_1, f_2 : \Omega \to [-\infty, \infty]$ be μ-integrable with $f_1 = f_2$ μ-a.e. Then*

$$\int_\Omega f_1 d\mu = \int_\Omega f_2 d\mu.$$

We now comment further on the fact that functions that agree almost everywhere have the same integral. We can view this as an assertion that integration is a *noise-tolerant* operation: one can have “noise” or “errors” in a function $f(x)$ on a null set, and this will not affect the final value of the integral. Indeed, once one has this noise tolerance, one can even integrate functions f that are not defined everywhere on $\mathbb{R}^n$,

[3] the monotonicity can be generalized to bounded variation

but merely almost every on $\mathbb{R}^n$, simply by extending f to all of $\mathbb{R}^n$ in some arbitrary fashion (e.g. by setting f equal to zero on the nullset). This is extremely convenient for analysis, as there are many natural functions that are only defined almost everywhere instead of everywhere (often due to division by zero problems when a denominator vanishes). While such functions cannot be evaluated at certain singular points, they can still be integrated (provided they obey some integrability conditions), and so one can still perform a large portion of analysis on such functions.

In fact, in the subfield of analysis known as *functional analysis*, it is convenient to abstract the notion of an almost everywhere defined function somewhat, by replacing any such function f with the equivalence class of almost everywhere defined functions that are equal to f almost everyhwere. Such classes are then no longer functions in the standard set-theoretic sense, but the properties of various function spaces improve when one does this (various pseudo-norm become norms, various topologies become Hausdorff, and so forth).

The "Lebesgue philosophy" is that one is willing to lose control on sets of measure zero. This perspective is what distinguishes Lebesgue-type analysis from other types (most notable descriptive set theory). This loss of control on null sets is the price one has to pay for gaining access to the Lebesgue integral.

References

[SS09] Elias M. Stein and Rami Shakarchi. *Real Analysis*. Princeton University Press, 2009. DOI: doi:10.1515/9781400835560.

[Tao11] Terence Tao. *An introduction to measure theory*. Vol. 126. 2011. URL: https://terrytao.files.wordpress.com/2011/01/measure-book1.pdf.