v: 2018-01-30

REAL ANALYSIS LECTURE NOTES

RASUL SHAFIKOV

3. Integration: from Riemann to Lebesgue

3.1. **Riemann Integral.**

Definition 3.1. *For $a < b$, a partition of an interval $[a, b] \subset \mathbb{R}$ is a finite collection of points $P = \{x_0, \ldots, x_m\}$, $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$. A step function $s(x)$ for a partition P is a function which is constant on each interval (x_i, x_{i+1}), and arbitrary at all other points.*

For a domain $B = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ a partition is a set of the form $P = P_1 \times \cdots \times P_n$, where P_i is a partition of $[a_i, b_i]$. For a multi-index $I = (i_1, \ldots, i_n)$ denote by $\Box_I$ the set of the form $(x_{i_1}, x_{i_1+1}) \times \cdots \times (x_{i_n}, x_{i_n+1})$ and call it a *brick* of the partition P. A function $s(x)$ is a step function for a partition P if it is constant on every brick of P. The volume of a brick $\Box_I$ is the usual Euclidean volume, i.e.,

$$\operatorname{vol}(\Box_I) = (x_{i_1+1} - x_{i_1}) \cdot \ldots \cdot (x_{i_n+1} - x_{i_n}).$$

Definition 3.2. *A partition Q is a refinement of a partition P if $P_i \subset Q_i$ for all $i = 1, \ldots, n$.*

Lemma 3.3. *Any two partitions of a domain B have a common refinement.*

Proof. Given partitions $P = P_1 \times \cdots \times P_n$ and $P' = P'_1 \times \cdots \times P'_n$, the partition $(P_1 \cup P'_1) \times \cdots \times (P_n \cup P'_n)$ is a common refinement. $\Box$

Given a step function $s(x)$ for a partition P of $B \subset \mathbb{R}^n$, we define

$$\mathcal{I}(s, P) = \sum_{I \in P} s_I \operatorname{vol}(\Box_I),$$

where s_I is the value of $s(x)$ on the brick $\Box_I$, and the summation is taken over all bricks in the partition.

Lemma 3.4. *If $s(x)$ is a step function for partitions P and P' then $\mathcal{I}(s, P) = \mathcal{I}(s, P')$.*

Proof. Obvious. $\Box$

It follows from the above lemma that $\mathcal{I}(s, P)$ does not depend on the choice of the partition P for which s is a step function. Therefore, we simply denote this number by $\mathcal{I}(s)$.

Lemma 3.5. *If $s(x)$ is a step function for a partition P and $t(x)$ is a step function for P', then $s(x) \le t(x)$ implies $\mathcal{I}(s) \le \mathcal{I}(t)$.*

Proof. Pass to a common refinement and use the preceding lemma. $\Box$

Definition 3.6. *Let $B = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$. A function $f : B \to \mathbb{R}$ is called (Riemann) integrable on B if for any $\varepsilon > 0$ there exist step functions $s(x)$ and $t(x)$ such that $s(x) \le f(x) \le t(x)$ for all x and $\mathcal{I}(t) - \mathcal{I}(s) < \varepsilon$. For a function f integrable on a domain B define*

$$\int_B f(x)dx = \sup_{s \le f} \mathcal{I}(s) = \inf_{f \le t} \mathcal{I}(t),$$

where the the supremum (resp. infimum) is taken over all step function s (resp. t) with $s \leq f$ (resp. $f \leq t$).

Proposition 3.7. *Continuous functions on $\mathbb{R}^n$ are Riemann integrable on any domain $B = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$.*

Proof. Let $\varepsilon > 0$ be given. Recall that a continuous function on a compact set is uniformly continuous, i.e., for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \varepsilon$. Thus, there exists $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \frac{\varepsilon}{\mathrm{vol}\,(B)}$. Select a partition P sufficiently fine so that the diameter of each brick of P is less than δ. Choose step functions $s(x)$ and $t(x)$ to be respectively the minimum and the maximum of f on each brick. Then $s \leq f \leq t$, and

$$\int_B t(x) - \int_B s(x) \leq \frac{\varepsilon}{\mathrm{vol}\,(B)} \sum_{I \in P} \mathrm{vol}\,(\square_I) = \varepsilon.$$

$\square$

We remark that what we defined above is, in fact, called the *Darboux integral*. However, it can be shown that Darboux's definition of integral is equivalent to that of Riemann.

3.2. **What is wrong with the Riemann integral?** There are several reasons why the Riemann integral defined in the previous section does not seem to be adequate. It all boils down to the fact that certain reasonable functions are not Riemann integrable. The following three examples will illustrate that. We begin with a definition.

Definition 3.8. *Given a set $S \subset \mathbb{R}^n$, the characteristic function χ_S of S is defined to be*

$$\chi_S(x) = \begin{cases} 1, & \text{if } x \in S, \\ 0, & \text{if } x \notin S. \end{cases}$$

Example 3.1. The so-called Dirichlet function $\chi_{\mathbb{Q}}$, is clearly not Riemann integrable on $[0, 1]$, since $\int_{[0,1]} s(x) = 0$ and $\int_{[0,1]} t(x) = 1$ for any step functions s and t with $s \leq \chi_{\mathbb{Q}} \leq t$. This is because both rational numbers $\mathbb{Q}$ and irrational numbers $\mathbb{R} \setminus \mathbb{Q}$ are dense in $\mathbb{R}$. $\diamond$

Proposition 3.9. *Every open set $U \subset \mathbb{R}$ can be written in a unique way as an at most countable union of disjoint open intervals.*

We leave the proof of the proposition as an exercise for the reader. With the help of this proposition we can make the following definition. Given an open set $U \subset \mathbb{R}$ we define the *Lebesgue measure* of U to be

$$m(U) = \sum_I |U_I|,$$

where $|(U_i)|$ is the length of the interval U_I, and the summation is taken over the disjoint union of open intervals whose union is U. It is immediate that the Lebesgue measure of every open interval is equal to its length.

While the previous example can be dismissed by declaring $\chi_{\mathbb{Q}}$ to be "too irregular" to be integrable, the next example shows that there exist open sets whose characteristic functions are not integrable.

Example 3.2. Suppose $U \subset [0, 1]$ is an open set with the following properties: U is dense in $[0, 1]$, and $m(U) < 1$. We claim that χ_U is not Riemann integrable. For the proof of the claim consider any two step functions $s(x) \leq \chi_U(x) \leq t(x)$ for a partition P of $[0, 1]$. Since U is dense, any brick

$[x_i, x_{i+1}]$ will have a nonempty intersection with U, and so $\int_{[0,1]} t(x) = 1$. On the other hand, (using the multidimensional notation, although we are in $\mathbb{R}$), let

$$\int_{[0,1]} s(x) = \sum_I s_I \operatorname{vol}(\Box_I).$$

Separate the partition P into $R \cup S$, where $R = \{J \in P : s_J > 0\}$ and $S = \{J \in P : s_J \leq 0\}$. Then $J \in R$ implies $0 < s_J < 1$, and $\Box_J \subset U$. It follows then that

$$\int_{[0,1]} s(x) = \sum_{J \in S} s_J \operatorname{vol}(\Box_J) + \sum_{J \in R} s_J \operatorname{vol}(\Box_J) \leq \sum_{J \in R} \operatorname{vol}(\Box_J) \leq m(U) < 1.$$

This shows that χ_U is not Riemann integrable.

It remains to show that there indeed exist dense open subsets of $[0,1]$ with the Lebesgue measure less than 1. To construct such a set, enumerate $\mathbb{Q} \cap (0,1)$ as $\{r_1, r_2, \ldots, \}$. Suppose that $0 < b < 1$. For every $l \in \mathbb{N}$ select an open interval J_l such that $r_l \in J_l$, $J_l \subset (0,1)$ and the length of J_l equals $b/2^l$. Then the union U of all U_l is an open subset of $[0,1]$ which is clearly dense in $[0,1]$. Let U_i be the disjoint union of open intervals, with $\cup_j U_j = U$ (these exist by Proposition 3.9). Then,

$$m(\cup_l U_l) = \sum_j \operatorname{vol}(U_j) \leq \sum_l \operatorname{vol}(J_l) = b.$$

Thus, U has the required properties. $\diamond$

Example 3.3 (Cantor-type sets)**.** Let $I = [p, q]$ be an interval in $\mathbb{R}$, and let the length of I be equal to $b > 0$. For $b > a > 0$, write $I = [p, r] \cup (r, s) \cup [s, q]$, such that $|r - p| = |q - s| = (b - a)/2$, and $|s - r| = a$. We call $[p, r]$ and $[s, q]$ the remnants of I and (r, s) the middle part of I.

Select $a_0, a_1, \ldots,$ positive real numbers such that $\sum_{n=0}^{\infty} 2^n a_n = a$. For each $n \geq 1$, let

$$b_n = 2^{-n}\left(1 - \sum_{k=0}^{n-1} 2^k a_k\right),$$

so that $b_n > a_n$ and $b_{n+1} = \frac{b_n - a_n}{2}$ for all n. Let $S_0 = \{I_0\}$ be the middle part of $[0,1]$, $T_1 = \{J_1, J_2\}$ be the corresponding a_0-remnants. Then these have length $b_1 > a_1$. Let S_1 be the middle a_1-parts of T_1, and let T_2 be the set of a_1-remnants of T_1. Their length is $b_2 > a_2$. Note that S_1 has 2 elements, while T_2 has 4. We continue inductively: construct S_{n+1} by taking the middle a_n-parts of T_n, while T_{n+1} will consist of the remnants of T_n.

Let $U = S_0 \cup S_1 \cup \ldots$. By construction, this union is disjoint, and $m(U) = \sum 2^n a_n = a$. Let J be an arbitrary subinterval of $[0,1]$ of length b_n. Then J intersects $S_0, \ldots S_n$, and hence, U. Indeed, otherwise, J is contained in the disjoint union in T_{n+1}. But the intervals in T_{n+1} have length $b_{n+1} < b_n$, a contradiction. Since $b_n \to 0$ as $n \to \infty$, we conclude that any interval of positive length will have a nonempty intersection with U. Thus U is dense in $[0,1]$. It follows from the previous example that χ_U is not integrable on $[0,1]$.

For a concrete Cantor-type set, consider $a_n = \frac{1}{4^{n+1}}$. Then $a = \sum_{n=0}^{\infty} 2^n \frac{1}{4^{n+1}} = 1/2$, and thus the set U obtained for this choice of a_n has a nonintegrable characteristic function. $\diamond$

3.3. **Lebesgue Integral.** In this subsection we briefly outline the construction of the Lebesgue integral. We begin with Lebesgue measurable sets.

Let B be a "brick" domain in $\mathbb{R}^n$ defined by $B = I_1 \times \ldots. \times I_n$ where I_j are intervals in $\mathbb{R}$ of the form (a_j, b_j), $(a_j, b_j]$, $[a_j, b_j)$, or $[a_j, b_j]$, $a_j \leq b_j$. Define a map $m : \mathcal{P} \to [0, +\infty)$ on the set $\mathcal{P}$ of all bricks by setting $m(B) = \prod_j (b_j - a_j)$. Thus m is just the usual Euclidean volume (resp. length, area) of a brick. We also add the empty set to $\mathcal{P}$ and define $m(\varnothing) = 0$. If a set E is a finite disjoint union of bricks, i.e.,

$$E = \cup_{j=0}^{k} B_j, \quad B_j \in \mathcal{P}, \; B_i \cap B_j = \varnothing, \quad \forall i \neq j, \tag{1}$$

then clearly

$$m(E) = \sum_{j=0}^{k} m(B_k). \tag{2}$$

It is possible to extend m as a positive function to a wider class of sets still keeping the additivity property (2). We say that a subset E of $\mathbb{R}^n$ is *elementary* if it admits representation (1). Then we view (2) as the definition of $m(E)$. Note that this definition is independent of the choice of B_k in (1). It is easy to see that if E_1 and E_2 are two elementary sets, then $E_1 \cup E_2$, $E_1 \cap E_2$, $E_1 \backslash E_2$ are elementary sets. We denote the class of elementary sets by $\mathcal{E}$. The crucial property of the function $m : \mathcal{E} \to \mathbb{R}^+ \cup \{0\}$ is the following: if (E_j) is a finite or countable collection of elementary sets and $E \in \mathcal{E}$ satisfies $E \subset \cup_j E_j$, then $m(E) \le \sum_j m(E_j)$.

Let now A be a subset of $\mathbb{R}^n$. We define its outer measure m^* by

$$m^*(A) = \inf \left\{ \sum_j m(E_j) : A \subset \cup_j E_j, E_j \in \mathcal{E} \right\},$$

where the infimum is taken over all finite or countable coverings of A by elementary sets. Recall that a symmetric difference of two sets A and B is defined by $A \Delta B = (A \cup B) \backslash (A \cap B)$.

Definition 3.10. *A set $A \subset \mathbb{R}^n$ is called Lebesgue measurable if for every $\varepsilon > 0$ there exists $E \in \mathcal{E}$ such that $m^*(A \Delta E) < \varepsilon$. If A is a measurable set, the Lebesgue measure of A is defined as $m(A) := m^*(A)$.*

Denote by $\mathcal{M}$ the class of all measurable sets in $\mathbb{R}^n$. Clearly, every brick domain is measurable. One can show that $\mathcal{M}$ is closed with respect to finite or countable application of unions, intersections and differences. Further, one can show that any open or closed subset of $\mathbb{R}^n$ is measurable, and that a set X is measurable if and only if for any $\varepsilon > 0$ there exists an open set G (resp. closed F) such that $X \subset G$ (resp. $F \subset X$) such that $m^*(G \setminus X) < \varepsilon$ (resp. $(m^*(X \setminus F) < \varepsilon)$.

Perhaps the most important property of the Lebesgue measure is its σ-*additivity*: if (A_j) is a disjoint sequence of measurable sets and $A = \cup_j A_j$, then $m(A) = \sum_j m(A_j)$. It is also monotone: if $A \subset B$ then $m(A) \le m(B)$.

Lemma 3.11. *Any countable set S in $\mathbb{R}^n$ has measure zero.*

Proof. Enclose every point a_n of $S = \{a_0, a_1, \dots\}$ in a brick domain of volume $\varepsilon/2^n$. □

Note that the converse to the lemma is false: there exist sets of measure zero which are not countable. A primary example of such domain is the *Cantor set*. Following the construction in Example 3.3 we produce an open set U by taking $a_n = 1/3$ for all $n \in \mathbb{N}$. Then the set $[0,1] \setminus U$ is called the Cantor set. It is a compact set of measure zero, and can be shown to have cardinality of $\mathbb{R}$. We leave details to the reader.

We now move from sets to functions. Let X be a measurable subset of $\mathbb{R}^n$. A function $f : X \to \mathbb{R}$ is called measurable if all subsets $f^{-1}((-\infty, a))$, $f^{-1}((-\infty, a])$, $f^{-1}([a, \infty))$, $f^{-1}((a, \infty))$ are measurable for every $a \in f(X)$. In particular, suppose that f admits at most a finite set of values $y_0, y_1, ..., y_k$. Then f is measurable if and only if every set $f^{-1}(y_j)$ is measurable. Measurable functions that admit only finitely many values will be called *simple*. The *Lebesgue integral* over X of a simple function ψ is defined by

$$\int_X \psi(x)dx := \sum_j y_j\, m(\psi^{-1}(y_j)). \tag{3}$$

Definition 3.12. *Let $f : X \to \mathbb{R}$ be a bounded measurable function defined on $X \in \mathcal{M}$ with $m(X) < \infty$. Then define*

$$\int_X f(x)dx = \sup_{\psi \le f} \int_X \psi(x)dx, \tag{4}$$

where the supremum is taken over all simple functions ψ on X satisfying $\psi \le f$.

It can be shown that for a measurable function $f : X \to \mathbb{R}$, the Lebesgue integral can be also defined as $\int_X f(x)dx = \inf_{\phi \ge f} \int_X \phi(x)dx$ for simple functions $\phi \ge f$. Both definitions agree.

Proposition 3.13. *If $f : X \to \mathbb{R}$ is Riemann integrable for a brick domain $X \subset \mathbb{R}^n$, then the integral in* (4) *well-defined and finite.*

Proof. Note that every step function on X in particular is a simple function. Hence, for step functions $s(x)$ and $t(x)$ satisfying $s \le f \le t$, we have

$$\mathcal{I}(s) = \int_X s(x)dx \le \sup_{\phi \le f} \int_X \phi(x)dx \le \inf_{f \le \psi} \int_X \psi(x)dx \le \int_X t(x)dx = \mathcal{I}(t).$$

Since f is Riemann integrable, $\mathcal{I}(t) - \mathcal{I}(s)$ can be made arbitrarily small, and we conclude that the function f is Lebesgue integrable. $\square$

If now $f \ge 0$ on $X \in \mathcal{M}$, we define

$$\int_X f(x)dx = \sup_{h \le f} \int_X h(x)dx, \tag{5}$$

where the supremum is taken over all bounded measurable functions h such that $m\{x : h(x) \neq 0\} < \infty$. This last assumption ensures that $\int_X h(x)dx$ on the right-hand side of (5) is well-defined even if $m(X) = \infty$. Indeed, we simply have

$$\int_X h(x)dx = \int_{\{x : h(x) \neq 0\}} h(x)dx.$$

For a general measurable $f : X \to \mathbb{R}$ we set $f^+ = \max\{f, 0\}$, and $f^- = \max\{-f, 0\}$. Then $f = f^+ - f^-$, and $|f| = f^+ + f^-$.

Definition 3.14. *For $X \in \mathcal{M}$ and a measurable $f : X \to \mathbb{R}$ we define*

$$\int_X f(x)dx = \int_X f^+(x)dx - \int_X f^-(x)dx.$$

If both integrals on the right are finite we say that f is (Lebesgue) integrable on X. The class of integrable functions is denoted by $L^1(X)$.

A property of functions defined on a domain in $\mathbb{R}^n$ is said to hold *almost everywhere* if it does not hold on a set of measure zero. The common notation for that is *a.e.*. For example two functions $f = g$ a.e. means that the set of points where f is not equal to g has measure zero. It follows then that $\int f = \int g$. Another example is convergence a.e.: we say $\lim f_n = f$ a.e., if the set of points x for which $\lim f_n(x) \neq f(x)$ has measure zero.

Using the definition of the integral and properties of measurable sets one can prove basic properties of integration, such as $\int af + bg = a\int f + b\int g$ for $a, b \in \mathbb{R}$; $f \le g \Rightarrow \int f \le \int g$; $\int_{A\cup B} f = \int_A f + \int_B f$ for disjoint A, B; etc. A more delicate property is taking the limit under the integral sign. The following two theorems provide sufficient conditions under which the operations of taking a limit and integration commute.

Theorem 3.15 (Fatou's lemma)**.** *If $\{f_n\}$ is a sequence of nonnegative measurable functions and $f_n(x) \to f(x)$ a.e. on $X \in \mathcal{M}$, then*

$$\int_X f(x)dx \le \liminf \int_X f_n. \tag{6}$$

Proof. Without loss of generality we may assume $f_n(x) \to f(x)$ for all x. By the definition of the Lebesgue integral, it is enough to show that (6) holds if we replace f with any non-negative simple function $\phi \le f$. Suppose that $\phi = \sum_{k=1}^m a_k \chi_{A_k}$, where A_k are disjoint measurable sets, and $a_k > 0$. Let $0 < t < 1$. Since $\phi(x) \le f(x)$, we see that $a_k \le \liminf f_n(x)$ for each k and $x \in A_k$. It follows that for a fixed k the sequence of sets

$$B_{kn} = \{x \in A_k : f_p(x) \ge ta_k \ \text{ for all } \ p \ge n\}$$

increases to A_k. Consequently, $m(B_{kn}) \to m(A_k)$ as $n \to \infty$. The simple function $\sum_{k=1}^m ta_k\chi_{B_{kn}}$ is everywhere less than f_n, and so

$$\int_X f_n\, dx \ge \sum_{k=1}^m ta_k m(B_{kn}).$$

Taking $\liminf$ in this inequality yields

$$\liminf_{n\to\infty} \int_X f_n dx \ge \sum_{k=1}^m ta_k m(A_k) = t\int_X \phi dx.$$

Finally, by letting $t \to 1$ we get (6). $\square$

Theorem 3.16 (Lebesgue Convergence theorem)**.** *Let $X \subset \mathbb{R}^n$ be a measurable subset and $\{f_n\}$ be a sequence of measurable functions. Suppose that $f_n(x) \longrightarrow f(x)$ for almost every $x \in X$. Furthermore, assume that there exists a function $g \in L^1(X)$ such that*

$$|f_n(x)| \le g(x), n = 1, 2, ...$$

Then $f \in L^1(X)$, and

$$\lim_{n\longrightarrow\infty} \int_X f_n dx = \int_X f\, dx.$$

Proof. The function $g - f_n$ is nonnegative, and so by Fatou's lemma

$$\int_X (g - f)\, dx \le \liminf \int_X (g - f_n)\, dx.$$

Since $|f| \le g$, f is integrable, and we have

$$\int_X g\, dx - \int_X f\, dx \le \int_X g\, dx - \limsup \int_X f_n\, dx,$$

from which we conclude that

$$\int_X f\, dx \ge \limsup \int_X f_n\, dx.$$

Similarly, considering $g + f_n$ we get

$$\int_X f\,dx \leq \liminf \int_X f_n\,dx,$$

and the theorem follows. $\square$

Note that Fatou's lemma has a weaker hypothesis than the Lebesgue Convergence theorem, and as a result its conclusion is also weaker. The advantage of Fatou's lemma is that it is applicable even if f is not known to be integrable and so it is often a good way of showing that f is integrable.