MATH3320 summary

Stanley Li

April 16, 2023

Contents

0 Notation, conventions and perspectives

The *characteristic function* χ_S or $\mathbb{1}_S$ of a set $S \subseteq X$ is the function

$$\chi_S = \mathbb{1}_S : X \to \mathbb{R}$$
$$x \mapsto \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{otherwise} \end{cases}$$

0.1 An analogy from real numbers to sets

As below (and as previously seen) we have two supplementary notions of a limit of a sequence of real numbers $(a_n)_{n=1}^{\infty}$, namely an "upper limit"

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup_{k\geq n} a_k = \inf_{n\in\mathbb{N}} \sup_{k\geq n} a_k$$

and a "lower limit"

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \sup_{k\geq n} a_k = \sup_{n\in\mathbb{N}} \inf_{k\geq n} a_k$$

The lattermost descriptions are independent of the notion of a limit and hence the notion of a metric, and only require the notion of a partial order[1] to state.

Another natural partial order is on the subsets of any set X (where in particular we should think of $X = \mathbb{R}^d$ as we will be almost exclusively doing analysis over Euclidean space), given by subset containment. This means that even though we have no notion of a "limit of sets" (due to the lack of suitable metric on a collection of sets), we can define a natural notion of an "upper limit" and "lower limit" to a sequence of sets, by first describing a suitable notion of supremum and infimum. Naturally, we can take

$$\inf_{i\in I} U_i = \bigcap_{i\in I} U_i$$

and

$$\sup_{i\in I} U_i = \bigcup_{i\in I} U_i$$

Then, following the above descriptions of $\limsup$ and $\liminf$, for a sequence $(U_n)_{n=1}^{\infty}$ we naturally arrive at

$$\limsup_{n\to\infty} U_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} U_n = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} U_n$$

and

$$\liminf_{n\to\infty} U_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} U_n = \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} U_n$$

and it is probably unsurprising from these notions that $\liminf_{n\to\infty} U_n \subseteq \limsup_{n\to\infty} U_n$. These are the notions of $\limsup$ and $\liminf$ we take in the remainder of these notes. For an increasing sequence of sets $(U_n)_{n=1}^{\infty}$ we can treat its limsup as its actual limit, and the same for a decreasing sequence of sets and its liminf[2].

[1] A *partial order* $\preceq$ on a set X is a relation satisfying reflexivity ($x \preceq x$), antisymmetry ($x \preceq y$ and $y \preceq x$ if and only if $x = y$) and transitivity ($x \preceq y$ and $y \preceq z$ necessarily yields $x \preceq z$). This can be thought of as a sensible notion of "ordering".

[2] And in theory if these two limiting notions are equal, it could be possible to define the "limit" of a sequence of sets.

0.2 An overarching idea for Lebesgue integration

The Riemann integration approach is to attempt to directly measure the area beneath the graph, using overestimates and underestimates and attempting to reach a common ground between the two.

In Lebesgue integration we also look at approximations to functions, but we take a different approach. Rather than directly taking a *measure* and attempting to compute upper and lower areas, we instead start from the level of sets to be measured. Measuring the size of sets in this case does not attempt to assign every set a "volume", but rather only the ones which are well-approximated by open sets (equivalently closed sets, and notably finite unions of cubes).

Once we have established which sets can be assigned a reasonable volume, we move to considering which functions can be easily assigned a reasonable notion of an integral, and we find that such functions are "very close" to being continuous. We then use these functions as a building block to define the integral for any arbitrary function, starting from the simplest functions (where we get finite sums), then the previous functions with a reasonable notion of an integral, and then arbitrary functions.

1 Analysis review

Definition 1 (liminf, limsup)**.** *Let* $(a_n)_{n=1}^{\infty} \in \mathbb{R}^{\infty}$ *be a sequence.*

- *If* (a_n) *is bounded above, its* limsup *is*

$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty} \sup_{k\geq n} a_k = \inf_{n\geq 1} \left(\sup_{k\geq n} a_k \right)$$

- *If* (a_n) *is bounded below, its* liminf *is*

$$\liminf_{n\to\infty} a_n := \lim_{n\to\infty} \inf_{k\geq n} a_k = \sup_{n\geq 1} \left(\inf_{k\geq n} a_k \right)$$

Notice that $\lim_{n\to\infty} a_n$ exists iff

$$\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$$

Definition 2 (Partitions)**.** *A* partition *of* $[a,b] \subseteq \mathbb{R}$ *is a set* $\{x_0, \ldots, x_n\} \subseteq [a,b]$ *with* $a = x_0 < \ldots < x_n = b$*.*

Definition 3 (Upper and lower areas; integrals)**.** *Let* $f : [a,b] \to \mathbb{R}$ *be bounded, and* $P = \{x_0, \ldots, x_n\}$ *be a partition of* $[a,b]$*.*

1. - *The* lower area of f with respect to P *is*

$$L(P,f) := \sum_{i=0}^{n-1} \inf_{x\in[x_i,x_{i+1}]} f(x)(x_{i+1} - x_i)$$

   - *The* upper area of f with respect to P *is*

$$U(P,f) := \sum_{i=0}^{n-1} \sup_{x\in[x_i,x_{i+1}]} f(x)(x_{i+1} - x_i)$$

2. • *The* lower Riemann integral of f *is*

$$\underline{\int_{[a,b]}^{R}} f := \sup_P L(P, f)$$

• *The* upper Riemann integral of f *is*

$$\overline{\int_{[a,b]}^{R}} f := \inf_P U(P, f)$$

3. *If* $\underline{\int_{[a,b]}^{R}} f = \overline{\int_{[a,b]}^{R}} f$*, we say that* f *is* Riemann integrable*, and define*

$$\int_{[a,b]}^{R} f := \underline{\int_{[a,b]}^{R}} f = \overline{\int_{[a,b]}^{R}} f$$

Definition 4 (Rectangles, volume)**.** *A rectangle* $R \subseteq \mathbb{R}^d$ *is of the form*

$$R := \prod_{i=1}^{d} [a_i, b_i] = [a_1, b_1] \times \ldots \times [a_d, r_d]$$

with volume

$$|R| := \prod_{i=1}^{d} (b_i - a_i)$$

A *cube* is rectangle with all sidelengths equal.

2 Measure theory

2.1 Simple functions

Rather than partitioning the domain as with Riemann integration, we instead opt to partition the range. The usual approximations of a bounded $f : [a, b] \to \mathbb{R}$ are by the functions

$$\varphi_{L,P} := \sum_{i=0}^{n-1} \inf_{x \in [x_i, x_{i+1}]} \mathbb{1}_{[x_i, x_{i+1}]}$$
$$\varphi_{U,P} := \sum_{i=0}^{n-1} \sup_{x \in [x_i, x_{i+1}]} \mathbb{1}_{[x_i, x_{i+1}]}$$

We instead opt for the following setup: Let $f([a, b]) \subseteq [c, d]$ and $Q = \{y_0, \ldots, y_n\}$ be a partition of $[c, d]$ (that is, $c = y_0 < \ldots < y_n = d$). Then we define

$$\varphi_Q := \sum_{i=0}^{n} y_i \mathbb{1}_{f^{-1}[y_i, y_{i+1}]}$$

To define the Riemann integral from this perspective, we want a notion of the "measure" of a set of the form $f^{-1}[y_i, y_{i+1}]$.

2.2 Elementary sets and Jordan measure

We take $A\dot{\sqcup}B$ to mean that A and B are "almost disjoint", in the sense that $\text{int}(A)$ and $\text{int}(B)$ are disjoint.

Definition 5 (Elementary sets; Jordan measure)**.** *Let $C \subseteq \mathbb{R}^d$. Then C is an* elementary set *if there are rectangles $R_1, \ldots, R_n$ with*

$$C = R_1 \dot{\sqcup} \ldots \dot{\sqcup} R_n$$

1. *If C is elementary as above, its* Jordan measure *(or volume) is*

$$|C| := m^J(C) := \sum_{i=1}^{n} |R_i|$$

2. *For $S \subseteq \mathbb{R}^d$,*
   - *the* inner Jordan measure $m_*^J(S)$ *is*

$$m_*^J(S) := \sup_{C \subseteq S \text{ elementary}} |C|$$

   - *the* outer Jordan measure $m_J^*(S)$ *is*

$$m_J^*(S) := \inf_{C \supseteq S \text{ elementary}} |C|$$

3. *If $m_*^J(S) = m_J^*(S)$, we say S is* Jordan measurable*, and define its* Jordan measure *$m^J(S)$ as $m^J(S) := m_*^J(S) = m_J^*(S)$.*

By rewriting the definitions, we have

Theorem 1 (Classification of Jordan measurable sets)**.** *Let $B \subseteq \mathbb{R}^d$ be bounded. Then B is Jordan measurable if and only if $\mathbb{1}_B$ is Riemann integrable.*

We want a notion of integrability for which we have some notion of commutativity with integrals and limits. For $\mathbb{Q} \cap [0,1] = \{q_1, q_2, \ldots\}$, this means

$$\lim_{n\to\infty} \mathbb{1}_{\{q_1,\ldots,q_n\}} = \mathbb{1}_{\mathbb{Q}\cap[0,1]}$$

and also in the context of integrals.

Theorem 2 (Coverings of open sets)**.** *Every open set $O \subseteq \mathbb{R}^d$ is the almost disjoint union of countably many cubes in $\mathbb{R}^d$.*

The key idea is to restrict to consider O and its relative position to the lattices $\mathbb{Z}^d$, $\left(\frac{1}{2}\mathbb{Z}\right)^d, \ldots$. In 1-dimension, we can look at $O \cap \mathbb{Q} = \{q_1, q_2, \ldots\}$ and take the largest intervals O_i contained in O containing q_i.

2.3 Exterior measure

Definition 6 (Exterior measure)**.** *Let $E \subseteq \mathbb{R}^d$. Its* exterior measure *is*

$$m_*(E) = \inf \left\{ \sum_{i=1}^{k} |Q_i| \;\middle|\; k \in \mathbb{N} \cup \{\infty\}, E \subseteq \bigcup_{i=1}^{k} Q_i \right\}$$

Here m_* is a map $m_* : \mathcal{P}(\mathbb{R}^d) \to [0, \infty]$. The exterior measure has the following properties:

1. If E is countable, then $m_*(E) = 0$;
2. if $A \subseteq B$, $m_*(A) \leq m_*(B)$;
3. $m_*(\mathbb{R}^d) = \infty$; and
4. if Q is a (closed or open) cube, then $m_*(Q) = |Q|$.

We further have

1. **[Volume of rectangles]** If R is a rectangle, then $m_*(R) = |R|$
2. **[Countable additivity]** If $E \subseteq \bigcup_{j=1}^{\infty} E_j$, then
$$m_*(E) \leq \sum_{j=1}^{\infty} m_*(E_j)$$
3. **[Relation to open sets]** Let $\mathcal{O}_E = \{O \supseteq E \mid O \text{ is open}\}$. Then
$$m_*(E) = \inf_{O \in \mathcal{O}_E} m_*(O)$$
4. **[Exact additivity]** If $E = E_1 \sqcup E_2$ with $d(E_1, E_2) > 0$, then $m_*(E) = m_*(E_1) + m_*(E_2)$.
5. **[Almost disjoint unions of cubes]** If $E = \dot{\bigcup}_{j=1}^{\infty} Q_j$, then $m_*(E) = \sum_{j=1}^{\infty} |Q_j|$.

The first statement can intuitively be rephrased as saying that the surface area of a rectangle is $(d-1)$-dimensional, while its volume is d-dimensional, and so the ratio between these two quantities tends towards zero.

Since every open set is an almost disjoint union of cubes, we have

Corollary 1. *If $O_1, O_2, \ldots$ are disjoint open sets, then*
$$m_*\left(\bigsqcup_{j=1}^{\infty} O_j\right) = \sum_{j=1}^{\infty} m_*(O_j)$$

The key idea in the proof of these statements is to use the ε-criterion to show that a value is indeed the infimum of a set (namely to get the harder inequality). Namely, for each $\varepsilon > 0$, note that there is some element $s \in S$ with $s \leq \inf(S) + \varepsilon$, and try to relate your claimed value i_s with s (usually by showing $i_s \leq s$). It often also helps to write $\varepsilon = \varepsilon/2 + \varepsilon/4 + \ldots + \varepsilon/2^k + \ldots$.

2.4 Measurable sets

The ideal properties for a measure defined on all sets to satisfy are incompatible. We describe Lebesgue measurable sets, which are intuitively those which can be well-approximated (in terms of volume) by open sets.

Definition 7 (Lebesgue measurable)**.** *We say $E \subseteq \mathbb{R}^d$ is* Lebesgue measurable *if for $\varepsilon > 0$, we can find $O \supseteq E$ with*
$$m_*(O - E) < \varepsilon$$
In this case, its Lebesgue measure *is $m(E) := m_*(E)$.*

The set of measurable sets satisfies the following properties.

Theorem 3. *The following sets are measurable.*

1. *Any open $O \subseteq \mathbb{R}^d$.*
2. *Any set $E \subseteq \mathbb{R}^d$ of exterior measure zero.*
3. *A countable union of measurable sets.*
4. *Any closed $C \subseteq \mathbb{R}^d$.*
5. *The complement of any measurable set.*
6. *A countable intersection of measurable sets.*

By *(1.)*, *(3.)* and *(5.)*, Lebesgue measurable sets are thus a σ-algebra, as defined below.

Definition 8. *A* σ-algebra *on $\mathbb{R}^d$ is a collection of subsets Σ closed under complement and countable union (and hence countable intersection) with $\emptyset \in \Sigma$.*

We note also that for any measurable set E, taking $\varepsilon = 1/n$ and O_n to be the associated open set, we find that for $O = \bigcap_{n=1}^{\infty} O_n$, $m_*(O - E) = 0$, and so the countable intersection O of open sets differs from E by a set of measure zero, i.e. $O = E \sqcup (O - E)$. We refer to such a set (i.e. a countable intersection of open sets) as a *G_δ-set.*

Similarly, doing the same for E^c, we find open sets $O_n \supseteq E^c$ with $m_*(O_n - E^c) < \frac{1}{n}$, i.e. $m_*(E - O_n^c) < \frac{1}{n}$, and then for $C = \bigcap_{n=1}^{\infty} O_n^c$ (which is a countable intersection of closed sets), we see that $m_*(E - C) = 0$, i.e. E differs from a countable union of closed sets by a set of measure zero ($E = C \sqcup (E - C)$). A countable union of closed sets is referred to as a *F_σ-set.*

Hence all Lebesgue measurable sets can be "approximated" by G_δ and F_σ sets. We denote the set of Lebesgue measureable sets by $\mathcal{M}(\mathbb{R}^d)$

Other examples of σ-algebras are the trivial σ-algebra $\{\emptyset, \mathbb{R}^d\}$, the discrete σ-algebra $2^{\mathbb{R}^d} := \mathcal{P}(\mathbb{R}^d)$, and the Borel σ-algebra[3] $\mathcal{B}(\mathbb{R}^d)$, generated by the open sets in $\mathbb{R}^d$.

We have strict inclusions

$$\mathcal{B}(\mathbb{R}^d) \subsetneq \mathcal{M}(\mathbb{R}^d) \subsetneq \mathcal{P}(\mathbb{R}^d)$$

Indeed, for the second inclusion, we can use the axiom of choice to define a map $[0,1]/\sim \to [0,1]$ where $x \sim y \iff x - y \in \mathbb{Q}$, and this defines a non-measurable set.

For the first inclusion, we can take two Cantor sets – one with positive measure and one with measure 0, and describe a homeomorphism between them. Taking a non-measurable set contained in the Cantor set of positive measure (justified by a later result), its image is necessarily a measurable, non-Borel set.

The Lebesgue measure satisfies some translation invariance properties. Indeed, for any measurable $E \subseteq \mathbb{R}^d$, we have

1. [**Additive translation invariance**] For any $h \in \mathbb{R}^d$, $E + h$ is measurable with
$$m(E) = m(E + h)$$
2. [**Multiplicative translation invariance**] For any $\delta = (\delta_1, \ldots, \delta_d) \in (\mathbb{R}^+)^d$, $m(\delta E)$ (defined by pointwise multiplication) is measurable with
$$m(\delta E) = \left(\prod_{i=1}^{d} \delta_i \right) m(E)$$

[3]The Borel σ-algebra $\mathcal{B}(X)$ is a construction which works on any topological space X.

Unlike the exterior measure, due to our restriction to measurable sets, we have the following additivity with the Lebesgue measure.

Remark 1. *The Lebesgue measure is a measure μ on $\mathbb{R}^d$ with the properties that*

1. *(Open sets are "large") $\mu(O) > 0$ for any open O;*
2. *(Compact sets are "small") $\mu(K) < \infty$ for any compact K; and*
3. *(Additive translation invariance) $\mu(E + h) = \mu(E)$ for any measurable E.*

Such a measure is called a Haar measure*, and as $\mathbb{R}^d$ is a "nice" space (i.e. a locally compact additive topological group), it is actually the unique measure (up to a scalar) on $\mathbb{R}^d$ with these properties.*

Theorem 4 (Countable additivity)**.** *Let $E_j \subseteq \mathbb{R}^d$ be pairwise disjoint measurable sets (for $j \in \mathbb{N}$). Then*

$$m\left(\bigsqcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} m(E_j)$$

By the construction of a non-measurable set, we also see that the following holds.

Proposition 1. *Any set of positive measure contains a non-measurable subset.*

We have the following notion of "convergence" of subsets of $\mathbb{R}^d$.

Definition 9 (Convergence of sets)**.** 1. *Let $\{E_i\}_{i=1}^{\infty}$ be an increasing sequence of sets. We say that E_j* converges to *E and write $E_i \nearrow E$ if $E = \bigcup_{i=1}^{\infty} E_i$.*

2. *Let $\{E_i\}_{i=1}^{\infty}$ be a decreasing sequence of sets. We say that E_j* converges to *E and write $E_i \searrow E$ if $E = \bigcap_{i=1}^{\infty} E_i$.*

We can view these as suprema and infima of the associated sequences of sets. The Lebesgue measure has the following compatibility with this notion of set limit.

Lemma 1. 1. *Let $\{E_i\}_{i=1}^{\infty}$ be an increasing sequence of sets converging to E. Then $m(E) = \lim_{n\to\infty} m(E_n)$.*

2. *Let $\{E_i\}_{i=1}^{\infty}$ be a decreasing sequence of sets converging to E, and suppose that $E_k < \infty$ for some $k \in \mathbb{N}$. Then $m(E) = \lim_{n\to\infty}(E_n)$.*

We thus have the following characterisations of measurable sets.

Theorem 5. *Let $E \subseteq \mathbb{R}^d$ be measurable. Then for every $\varepsilon > 0$,*

1. *(Definition) there is an open set $O \supseteq E$ with $m(O - E) < \varepsilon$.*
2. *(Equivalent definition) There is a closed subset $C \subseteq E$ with $m(E - C) < \varepsilon$.*
3. *If $m(E) < \infty$, then there is an compact set $K \subseteq E$ with $m(E - K) < \varepsilon$.*
4. *If $m(E) < \infty$, then there is a finite almost-disjoint union of cubes[4] F with $m(E \triangle F) < \varepsilon$.*

2.5 Measurable functions

Having described measurable sets, we move to describing measurable functions, en route to defining Lebesgue integration.

[4] where $\triangle$ denotes symmetric difference.

Definition 10. *Let $E \subseteq \mathbb{R}^d$ be measurable and $f : E \to \mathbb{R}$ be a function. We say that f is* measurable *if the preimages $f^{-1}[-\infty, a)$ (denoted $\{f < a\}$) are measurable for all $a \in \mathbb{R}$.*

By the properties of measurable sets, it is equivalent to say that any of $\{f \leq a\} := f^{-1}[-\infty, a]$, $\{f > a\} := f^{-1}(a, \infty]$, $\{f \geq a\} := f^{-1}[a, \infty]$, $f^{-1}(a, b)$, $f^{-1}[a, b]$ are measurable for all a, b.

For examples of measurable functions, we see that any continuous f is measurable as $f^{-1}[-\infty, a)$ is open in each case, and that for a measurable set E, its associated characteristic function χ_E is also measurable.

Measurable functions have the following compatibility with arbitrary functions.

Theorem 6. *Let $(f_n)_{n=1}^{\infty}$ be a sequence of measurable functions. Then*

$$\liminf_{n\to\infty} f_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} f_k \qquad \text{and} \qquad \limsup_{n\to\infty} f_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} f_k$$

are measurable.

We then have the following corollary.

Corollary 2. *If $(f_n)_{n=1}^{\infty}$ is a sequence of measurable functions with $f_n \to f$, then f is measurable.*

We also have the following "closure" properties of measurable functions under (pointwise) addition and multiplication.

Theorem 7. *Let f and g be measurable functions. Then*

1. *f^k is measurable for any $k \in \mathbb{N}$.*
2. *$f + g$ and fg are measurable.*

We evidently have (2) $\implies$ (1), but the statement of (1) is used to prove (2), by using the fact that $fg = \frac{1}{4}((f + g)^2 - (f - g)^2)$.

Remark 2. *Algebraically,* ***Theorem 7*** *shows that the set of measurable functions on $\mathbb{R}^d$ is a subring of the ring of functions $\mathbb{R}^d \to \mathbb{R}$ with pointwise operations, which in turn contains the subring of continuous functions (i.e. those where preimages of open sets are open sets).*

We have the following "equivalence relation" between almost-equal functions.

Definition 11. *Let $E \subseteq \mathbb{R}^d$ be measurable and $f, g : E \to \mathbb{R}$. We say that f and g are* equal almost everywhere *if*

$$m(\{x \in E \mid f(x) \neq g(x)\}) = 0$$

that is, they differ by a measure zero set.

From the perspective of measure theory, the statement that f and g are equal almost everywhere should be viewed as being "good enough" for all intensive purposes. The following proposition is one such way to make this notion explicit.

Proposition 2. *Let f and g be functions which are equal almost everywhere. Then if f is measurable, so is g.*

We will define integrals later in several stages, beginning with relatively basic functions and then extending these definitions to fit larger (and possibly more cathartic) classes of functions. For this, we have the following notion of approximating a measurable function f by an increasing sequence of step functions.

Proposition 3. *Let f be a non-negative measurable function. Then there exists a sequence of simple functions $\varphi_k \nearrow f$ pointwise.*

The previous result applies only to non-negative functions, and we have a result by writing $f = f^+ - f^-$ for non-negative functions $f^\pm$.

Theorem 8. *For each measurable f, there is a sequence of simple functions φ_k with*

$$|\varphi_k(x)| \leq |\varphi_{k+1}(x)|$$

and $\varphi_k \to f$ pointwise.

By instead replacing simple functions we have a slightly weaker approximation (i.e. for which this convergence happens *almost everywhere*).

Theorem 9. *For each measurable f, there is a sequence of step functions φ_k with $\lim_{k\to\infty} \varphi_k = f$ almost everywhere.*

2.6 Littlewood's three principles

We have the following "principles" for approximating functions by rectangles and their associated step functions.

1. Every measurable $E \subseteq \mathbb{R}^d$ is *nearly* a finite union of cubes;
2. Every measurable function f of finite-value is *nearly* continuous; and
3. Every convergent sequence of measurable functions is *nearly* uniformly convergent.

The first principle corresponds to the last statement in **Theorem 5**, and the second and third principles are statements known as Lusin's and Egorov's theorems respectively.

Theorem 10 (Egorov)**.** *Let $(f_k : E \to \mathbb{R})_{k=1}^\infty$ each be measurable, with $f_n \to f$ and $m(E) < \infty$. Then for every $\varepsilon > 0$, there is a closed set $A_\varepsilon \subseteq E$ with $m(E \setminus A_\varepsilon) < \varepsilon$ and $f_k \to f$ uniformly on A_ε.*

Egorov's theorem is extremely powerful, as the weak notion of pointwise convergence is near useless in comparison to uniform convergence, which preserves many important properties (such as integrals and (uniform) continuity). Having established Egorov's theorem, we are then able to prove

Theorem 11 (Lusin)**.** *Let $f : E \to \mathbb{R}$ be measurable and $m(E) < \infty$. Then for every $\varepsilon > 0$ there is a closed set A_ε with $m(E \setminus A_\varepsilon) < \varepsilon$ and $f|_{A_\varepsilon} : A_\varepsilon \to \mathbb{R}$ a continuous function.*

3 Integration theory

Equipped with the notions of measurable sets and functions move to properly defining the notions of integration in a more generalised manner (but only for measurable functions, for which this notion makes sense). This is done in stages, beginning with "basic" functions where integrals are just finite sums and building up to higher generality. In this section all functions are $\mathbb{R}^d \to \mathbb{R}$ unless otherwise stated. From the perspective of integration, every measurable function $f : E \to \mathbb{R}$ can be made into a function $\mathbb{R}^d \to \mathbb{R}$ by extending it to $\mathbb{R}^d$ by zero. This is denoted $f\chi_E$, and we treat $\int_E f = \int_{\mathbb{R}^d} f\chi_E$ (when these are defined).

3.1 Lebesgue integration

3.1.1 Stage 1 – Simple functions

Let φ be a simple function, i.e. of the form $\varphi = \sum_{k=1}^n a_k \chi_{E_k}$ where E_k are measurable sets. Its Lebesgue integral is then

$$\int_{\mathbb{R}^d} \varphi := \int_{\mathbb{R}^d} \varphi(x)dx := \sum_{k=1}^n a_k m(E_k)$$

This is well-defined (independent of the choice of measurable sets and coefficients a_k), and every such simple function has a unique form $\varphi = \sum_{k=1}^n a_k \chi_{E_k}$ where a_k are distinct and E_k are pairwise disjoint.

3.1.2 Stage 2 – Bounded measurable functions

When we say "bounded, measurable" in this sense we mean a function whose domain and range are both bounded in some sense (not just its range). We have the following lemma which says that we can approximate the integrals of bounded measurable functions by those of simple functions.

Lemma 2. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a bounded measurable function with $m(\mathrm{supp}(f)) < \infty$, and let $\varphi_k(x) \to f(x)$ everywhere in $\mathbb{R}^d$. Then*

1. *The limit*
$$\lim_{k\to\infty} \int_E \varphi_k$$
exists; and
2. *If $\psi_k(x) \to f(x)$ everywhere in E, then*
$$\lim_{k\to\infty} \left(\int_E (\varphi_k - \psi_k) \right) = 0$$
and in particular $\lim_{k\to\infty} \int_E \varphi_k = \lim_{k\to\infty} \int_E \psi_k$.

We then define

$$\int_{\mathbb{R}^d} f := \lim_{n\to\infty} \int_{\mathbb{R}^d} \varphi_k$$

for any sequence of simple functions $\varphi_k \to f$ pointwise, and the above lemma shows that this is well-defined. We also have the first integral convergence theorem, applicable to bounded functions with "bounded" domains.

Theorem 12 (Bounded convergence theorem)**.** *Let (f_n) be a sequence of bounded, measurable functions supported on a measurable set E of finite measure. Suppose also that $f_n \to f$ pointwise almost everywhere. Then f is measurable, bounded, and $m(\mathrm{supp}(f) \bigtriangleup E) = 0$, and*

$$\int |f_n - f| \to 0$$

and in particular $\int f_n \to \int f$.

3.1.3 Stage 3 – Positive measurable functions

We define integrals in this case by taking the "largest" of the underestimates by bounded measurable functions. Namely, we set

$$\int_{\mathbb{R}^d} f := \sup\left\{\int_{\mathbb{R}^d} g \;\middle|\; 0 \le g \le f, g \text{ bounded and measurable}\right\}$$

and we say that f (when positive) is integrable when $\int_{\mathbb{R}^d} f$ We have the following properties of integrals of positive, measurable functions.

Lemma 3. *Let $f, g \ge 0$ be measurable, E be measurable and $\lambda, \mu \in \mathbb{R}$.*

1. *(Linearity)*
$$\int_E (\lambda f + \mu g) = \lambda \int_E f + \mu \int_E g$$
2. *(Additivity) If E_1, E_2 are disjoint and measurable, then*
$$\int_{E_1 \sqcup E_2} f = \int_{E_1} f + \int_{E_2} f$$
3. *(Monotonicity) If $f \le g$ on E, then*
$$\int_E f \le \int_E g$$
Further, if g is integrable, then so if f.
4. *(Integrability) If f is integrable, then $m(f^{-1}(\infty)) = 0$.*
5. *(Zero) If $\int_{\mathbb{R}^d} f = 0$, then $f = 0$ almost everywhere.*
6. *(Triangle inequality)*
$$\left|\int_E f\right| \le \int_E |f|$$

3.1.4 Stage 4 – General measurable functions

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a measurable function. Writing $f = f^+ - f^-$ where $f^\pm = \max(\pm f, 0)$, we set its integral to be

$$\int_{\mathbb{R}^d} f := \int_{\mathbb{R}^d} f^+ - \int_{\mathbb{R}^d} f^-$$

To check that this is well-defined, we have that if $f = f_1 - f_2 = g_1 - g_2$ for positive f_i, g_i, then $\int f_1 + \int g_2 = \int g_1 + \int f_2$, so the value of the above definition is independent of the choice of decomposition into positive functions. Such a function is integrable if $\int_{\mathbb{R}^d} |f| < \infty$.

Properties (1) – (4) of Lemma 3 also hold for general measurable functions. In terms of integrals, all convergence will be pointwise almost everywhere.

3.2 Convergence of integrals

The first result we have relates the integrals of f_n and f for $f_n \to f$.

Proposition 4 (Fatou Lemma). *Let (f_n) be a sequence of non-negative measurable functions with $f_n(x) \to f(x)$ almost everywhere. Then*

$$\int f \le \liminf_{n \to \infty} \int f_n$$

This can be thought of as an artifact of the fact that integrals are defined in terms of underestimates.

Corollary 3. *Let (f_n) be a sequence of non-negative measurable functions.*

1. *If $f_n(x) \to f(x)$ and $f_n(x) \leq f(x)$ for almost every x, then*
$$\int f = \lim_{n\to\infty} \int f_n$$
2. *If $a_k \geq 0$ are integrable functions, then*
$$\int \sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} \int a_k$$
which may be infinite. Further, if $\sum_{k=1}^{\infty} \int a_k$ is finite, then $\sum_{k=1}^{\infty} a_k$ converges for almost every x.
3. *(Monotone convergence theorem) If (f_n) are non-negative and $f_n(x) \nearrow f(x)$ pointwise almost everywhere, then*
$$\int f_n \to \int f$$
4. *(Borel-Cantelli Lemma) Let $\{E_i\}_{i=1}^{\infty}$ be a sequence of measurable sets with $\sum_{i=1}^{\infty} m(E_i) < \infty$. Then*
$$m\left(\limsup_{n\to\infty} E_n\right) = m\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} E_n\right) = 0$$
5. *For any $\delta > 0$, the function $f(x) = 1/\left|x\right|^{d+\delta}$ on $\mathbb{R}^d$ is integrable on $|x| \geq \varepsilon$ for any $\varepsilon > 0$.*

We have the following (rather strong) convergence theorem on integrals. It implies the bounded convergence theorem, but not the monotone convergence theorem, as it only deals with finite integrals.

Theorem 13 (Dominated convergence theorem)**.** *Suppose that (f_n) are integrable and f is such that $f_n(x) \to f(x)$ almost everywhere. If there is an integrable function g such that $|f_n(x)| \leq g(x)$ for almost every x, then*
$$\int |f_n - f| \to 0$$
and in particular
$$\lim_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n$$

The proof of this theorem requires the following results.

Lemma 4. *Let g be an integrable function, and $\varepsilon > 0$.*

1. *There is $N > 0$ such that*
$$\int_{B_N(0)^c} |g| < \varepsilon$$
2. *(Absolute continuity) There is $\delta > 0$ so that*
$$\int_E |g| < \varepsilon$$
whenever $m(E) < \delta$.

We have the following relationship between Riemann and Lebesgue integration.

Proposition 5. *Let f be a Riemann integrable function. Then f is Lebesgue integrable, and its Riemann and Lebesgue integrals are equal.*

In this sense Lebesgue integration is more general than and compatible with Riemann integration.

3.3 $L^1(\mathbb{R}^d)$ and its properties

Definition 12. *The* space $L^1(\mathbb{R}^d)$ of integrable functions on $\mathbb{R}^d$ *is*

$$L^1(\mathbb{R}^d) = \left\{ f : \mathbb{R}^d \to \mathbb{R} \;\middle|\; \int_{\mathbb{R}^d} |f| < \infty \right\} \Big/ \sim$$

where $f \sim g$ iff $f = g$ almost everywhere.

This set is naturally a normed vector space over $\mathbb{R}$ with the L^1 *norm*

$$\|f\|_{L^1} := \int_{\mathbb{R}^d} |f|$$

In fact, it is complete (i.e. a Banach space):

Theorem 14 (Riese-Fischer)**.** *$L^1(\mathbb{R}^d)$ is complete with the L^1-norm.*

In general, L^1 and pointwise convergence have no relation, but we have the following relation.

Corollary 4. *Every L^1-convergent sequence contains a pointwise convergent subsequence.*

It is usually convenient to view a metric space through the lens of a "much smaller" dense subset (such as $\mathbb{Q} \subseteq \mathbb{R}$). In the case of L^1, we have the following dense subsets.

Theorem 15 (Dense subsets of $L^1(\mathbb{R}^d)$)**.** *The following subsets of $L^1(\mathbb{R}^d)$ are dense:*

1. *the step functions;*
2. *the simple functions; and*
3. *the (uniformly) continuous functions with compact support.*

In L^1 we have the following invariance properties:

Proposition 6. *Let $f, g \in L^1(\mathbb{R}^d)$, $h \in \mathbb{R}^d$ and $\delta = (\delta_1, \dots, \delta_n) \in (\mathbb{R}^*)^d$. Then*

1. *(Additive invariance) The function $f_h(x) = f(x+h)$ is integrable, with*
$$\int_{\mathbb{R}^d} f(x+h)dx = \int_{\mathbb{R}^d} f_h = \int_{\mathbb{R}^d} f = \int_{\mathbb{R}^d} f(x)dx$$
2. *(Multiplicative invariance) The function $x \mapsto f(\delta x)$ is integrable, with*
$$\int_{\mathbb{R}^d} f(\delta x)dx = \left(\prod_{i=1}^{d} |\delta_i| \right)^{-1} \int_{\mathbb{R}^d} f(x)dx$$
3. *(Convolution is commutative) The map $y \mapsto f(x-y)g(y)$ is integrable, with*
$$\int_{\mathbb{R}^d} f(y)g(x-y)dy = \int_{\mathbb{R}^d} f(x-y)g(y)dy$$
4. *(L^1 continuity) We have $\|f_h - f\|_{L^1} \to 0$ as $\|h\| \to 0$.*

3.4 Fubini and Tonelli's theorems

Fubini and Tonelli's theorems concern the integrability and measurability of slice functions, and the interchanging of integrals.

Theorem 16 (Fubini)**.** *Let $f \in L^1(\mathbb{R}^{d_1+d_2})$. Then for almost every $y \in \mathbb{R}^{d_2}$:*

1. *the function f^y defined by $f^y(x) = f(x,y)$ is in $L^1(\mathbb{R}^{d_1})$;*
2. *the function*
$$\left(y \mapsto \int_{\mathbb{R}^{d_1}} f(x,y)dx\right) \in L^1(\mathbb{R}^{d_2})$$
and
3. $$\int_{\mathbb{R}^{d_1+d_2}} f(x,y)dxdy = \int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x,y)dx\right) dy$$

and the same holds under the interchange $x \leftrightarrow y$.

Tonelli's theorem is the analogous statement with integrable replaced with non-negative and measurable, and proven by taking a sequence of integrable functions tending towards each measurable function.

Theorem 17 (Tonelli)**.** *Let f be non-negative and measurable on $\mathbb{R}^{d_1+d_2}$. Then for almost every $y \in \mathbb{R}^{d_2}$,*

1. *the slice function $f^y(x) = f(x,y)$ is measurable;*
2. *the map $\mathbb{R}^{d_2} \to [0,\infty]$ given by*
$$\left(y \mapsto \int_{\mathbb{R}^{d_1}} f(x,y)dx\right)$$
is measurable; and
3. $$\int_{\mathbb{R}^{d_1+d_2}} f(x,y)dxdy = \int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x,y)dx\right) dy$$

and the same holds under the interchange $x \leftrightarrow y$.

Fubini's (and more generally Tonelli's) theorem allow us to prove statements about measurability, by taking $f = \chi_E$ for measurable E. We have

Proposition 7. *If $E_1 \times E_2 \subseteq \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ is measurable, and $m_*(E_2) > 0$, then E_1 is measurable.*

Alternatively, if we have E_1 and E_2 measurable, their product is measurable and the measures are compatible:

Proposition 8. *Suppose that $E_1 \subseteq \mathbb{R}^{d_1}, E_2 \subseteq \mathbb{R}^{d_2}$ are measurable. Then $E_1 \times E_2 \subseteq \mathbb{R}^{d_1+d_2}$ is measurable with*
$$m(E_1 \times E_2) = m(E_1)m(E_2)$$
with the understanding that $\infty \cdot 0 = 0$

To prove this we approximate $E_1 \times E_2$ by a uniform product of G_δ sets, and show that their difference is a measure zero set using

Lemma 5. *If $E_1 \subseteq \mathbb{R}^{d_1}, E_2 \subseteq \mathbb{R}^{d_2}$ are measurable, then*

$$m_*(E_1 \times E_2) \leq m(E_1)m(E_2)$$

with the understanding that $\infty \cdot 0 = 0$.

Returning to integrals, we have the following proposition on rewriting an integral:

Proposition 9. *Let $g \in L^1(\mathbb{R}^d)$ and $E_\alpha = \left\{x \in \mathbb{R}^d \mid |g(x)| > \alpha\right\}$. Then*

$$\int_{\mathbb{R}^d} |g(x)|\, dx = \int_0^\infty m(E_\alpha) d\alpha$$

We then have

Corollary 5. *The following hold.*

1. *If $f : \mathbb{R}^{d_1} \to \mathbb{R}$ is measurable, then so is $\left(\tilde{f}(x, y) = f(x)\right) : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$;*
2. *if $f \geq 0$ is measurable on $\mathbb{R}^d$, then so is the set*

$$\mathcal{A} = \left\{(x, y) \in \mathbb{R}^d \times \mathbb{R} \mid 0 \leq y \leq f(x)\right\}$$

and

$$m(A) = \int_{\mathbb{R}^d} f(x)dx$$

3. *If f is measurable on $\mathbb{R}^d$, then so is $\tilde{f}(x, y) = f(x - y) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$*
4. *(Convolution) If f and g are measurable on $\mathbb{R}^d$, then so is $(x, y) \mapsto f(x - y)g(y)$*

The only non-trivial part is (3), which follows by showing that the preimage of a measure zero set is also measure zero. As a final proposition, we have the actual definition of the Dirichlet convolution of two functions:

Proposition 10. *Let $f, g \in L^1(\mathbb{R}^d)$. Then*

$$(f * g)(x) := \int_{\mathbb{R}^d} f(x - y)g(y)dy$$

is well-defined and finite-valued almost everywhere.