

Part A Integration Notes

Tim Hosgood*
`contact@timhosgood.co.uk`

September 10, 2014

Abstract

These notes summarise the contents of a second-year integration course at the University of Oxford. Because of this, there are many references to certain proofs being ‘non-examinable’, though the contents are still intended to be as self contained as possible.

Please email any corrections to `contact@timhosgood.co.uk`.

CONTENTS

1	Introduction	2
1.1	Important results from Riemann integration	2
1.2	Overview/Summary	2
2	General measure spaces	3
2.1	Extended real number system	3
2.2	Measure spaces	4
3	The Lebesgue measure	6
3.1	The outer measure	6
3.2	Null sets and almost-sure properties	9
3.3	Lebesgue measurable sets and measurable functions	10
4	Measurable functions	12
4.1	Further convergence of functions	15
5	Lebesgue integration	17
5.1	Non-negative simple (measurable) functions	17
5.2	Non-negative measurable functions	19
5.3	Integrable functions	22
5.4	Integrals depending on parameters	27
6	Fubini and Tonelli	29
6.1	A speedy outline of Lebesgue theory on \mathbb{R}^2	29
6.2	Fubini’s Theorem and Tonelli’s Theorem	29
6.3	Change of variables	31

*Based on Z. Qian’s 2014 University of Oxford lecture notes.

7	\mathcal{L}^p and L^p spaces	31
7.1	Convergence in measure	35
7.2	Convergence in L^p space	35

INTRODUCTION

Important results from Riemann integration

In Riemann integration, we defined some set \mathcal{L} of step functions. That is, $\varphi \in \mathcal{L}$ means that $\varphi = \sum_{i=1}^n c_i \mathbb{1}_{A_i}$, where the A_i are finite intervals. Then it was possible to define some form of integration on \mathcal{L} : for $\varphi \in \mathcal{L}$, let

$$I(\varphi) = \sum_{i=1}^n c_i \int \mathbb{1}_{A_i} = \sum_{i=1}^n c_i |A_i|$$

where $|A_i|$ is the ‘length’ of A_i . We then proceeded to define upper and lower integrals of functions $f : [a, b] \rightarrow \mathbb{R}$ by

$$\underline{\int_a^b} f = \sup\{I(\varphi) \mid \varphi \in \mathcal{L}, \varphi \leq f \mathbb{1}_{[a,b]}\} \quad \text{and} \quad \overline{\int_a^b} f = \inf\{I(\psi) \mid \psi \in \mathcal{L}, \psi \geq f \mathbb{1}_{[a,b]}\}$$

This then gave us the condition: $f : [a, b] \rightarrow \mathbb{R}$ is Riemann integrable if and only if $\underline{\int_a^b} f = \overline{\int_a^b} f =: \int_a^b f$.

We obtained two important theorems in our development of Riemann integration which we shall be able to make use of in applying the theory of Lebesgue integration:

1. f is Riemann integrable on $[a, b]$, where $a, b \in \mathbb{R}$, if and only if f continuous on $[a, b]$ (and thus bounded and uniformly continuous);
2. (*The Fundamental Theorem of Calculus*) If F has continuous derivative on $[a, b]$ then $\int_a^b F'(x) dx = F(b) - F(a)$.

The main problem with this system is that \mathcal{L} is quite a small set, and we are led to the idea that increasing the size of the set of ‘simple’ function that we know how to integrate might be able to reduce the difference between $\underline{\int}$ and $\overline{\int}$, and thus be able to integrate even more functions.

Overview/Summary

One thing to bear in mind when we define Lebesgue integration is that, from the point of view of ‘integration is the area under the graph’, whereas Riemann integration involves partitioning the x -axis, Lebesgue integration in a sense instead partitions the y -axis and looks at the preimage of these partitions (which is evident in the definition of a measurable function).

First of all we define a measure space, then a semi-measure m^* (the outer measure), then the set \mathcal{M}_{Leb} of subsets of \mathbb{R} such that m^* becomes a measure (and thus the idea of a (Lebesgue) measurable set). Next we define what it means for a function to be measurable in a general measure space, and then in particular look at Lebesgue measurable functions (which we from now on refer to simply as measurable functions).

In what follows, we then ignore measure spaces in general (and so we simply say ‘measurable’ to mean Lebesgue measurable) and focus solely on our measure space $(\mathbb{R}, \mathcal{M}_{\text{Leb}}, m)$. The big step that all the previous work has been leading up to (and our motivation for the preceeding content): we define our new class of simple functions,

which turn out to be measurable functions. We next define integrals for non-negative simple functions, and further for non-negative measurable functions (via approximation by simple functions, just as in Riemann integration), and finally for any measurable function. Here we end up with Fatou's lemma, the MCT, DCT, etc. etc.

Penultimately, we look at products of measure spaces (but really only just \mathbb{R}^2), and then at what it means to be integrable in these spaces. Both Tonelli's and Fubini's theorem arise here, in an attempt to categorise the functions that are now integrable.

Lastly, we see how the set of all (Lebesgue) p -integrable (that is, integrable when raised to the p^{th} power) functions form a vector space, and that, when we quotient this space by the equivalence relation given by two functions be equal almost everywhere, a norm arises from integrating f^p . It turns out that this space is complete, so that we have a complete normed vector space, i.e. a Banach space. Along the way we also find out some things about functions converging under this norm, and how that relates to them converging in other ways.

GENERAL MEASURE SPACES

Extended real number system

We introduce the idea of infinity into our real number system: $[-\infty, \infty] = \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$. Everything is extended in the expected way, apart from the fact that $\frac{\infty}{\infty}$, $\infty - \infty$, and $\frac{a}{0}$ are undefined, and we set $0 \cdot \infty = 0 \cdot (-\infty) = 0$.

This turns out to be useful for various reasons: any subset of \mathbb{R} now has an infimum and a supremum (which are both obtained); and any monotonic sequence has a limit; any series of non-negative terms has a sum.

We now consider a different way of looking at infinite sums.

PROPOSITION (2.1) 1. Let $(a_n)_{n \geq 1}$ be a sequence of non-negative terms. Then

$$\sum_{n=1}^{\infty} a_n = \sup \left\{ \sum_{n \in J} a_n : J \subseteq \mathbb{N}, |J| < \infty \right\}.$$

2. Let $(b_{mn})_{m,n \geq 1}$ be a double sequence of non-negative terms, and $\{(m_k, n_k) : k \geq 1\}$ be any enumeration of $\mathbb{N} \times \mathbb{N}$. Then

$$\begin{aligned} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} b_{mn} &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} b_{mn} = \sum_{k=1}^{\infty} b_{m_k n_k} \\ &= \sup \left\{ \sum_{(m,n) \in J} b_{mn} : J \subseteq \mathbb{N} \times \mathbb{N}, |J| < \infty \right\}. \end{aligned}$$

That is, these sums can be arbitrarily rearranged.

DEFINITION

(2.2)

For a sequence $(a_n) \subset [-\infty, \infty]$ we define

1. $\limsup_{n \rightarrow \infty} a_n = \lim_{m \rightarrow \infty} \left(\sup_{n \geq m} a_n \right);$
2. $\liminf_{n \rightarrow \infty} a_n = \lim_{m \rightarrow \infty} \left(\inf_{n \geq m} a_n \right).$

We can be sure that these exist since $(\sup_{n \geq m} a_n)_{m \geq 1}$ is a decreasing sequence.

The use of these definitions is that $\limsup_{n \rightarrow \infty} a_n$ is the largest number ℓ such that there is a subsequence of (a_n) converging to ℓ .

PROPOSITION

(2.3)

Some useful properties of \liminf and \limsup are as follows:

1. $\liminf_{n \rightarrow \infty} a_n = -\limsup_{n \rightarrow \infty} (-a_n);$
2. $\liminf_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} a_n;$
3. $\lim_{n \rightarrow \infty} a_n$ exists if and only if $\liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n$, and then all three are equal;
4. if $a_n \leq b_n$ for all n then $\limsup_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} b_n;$
5. $\limsup_{n \rightarrow \infty} (a_n + b_n) \leq \limsup_{n \rightarrow \infty} a_n + \limsup_{n \rightarrow \infty} b_n$ (if all sums exist).

Measure spaces

DEFINITION Algebras, σ -algebras, and measurable spaces:

(2.4)

Let Ω be a set (called a *space*), and let \mathcal{F} be a collection of some subsets of Ω .

1. \mathcal{F} is called an *algebra* on Ω if
 - (a) $\emptyset, \Omega \in \mathcal{F};$
 - (b) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
 - (c) $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}.$
2. \mathcal{F} is further called a σ -*algebra* on Ω if it is an algebra, and also closed under countable unions (that is, $A_1, A_2, \dots \in \mathcal{F} \implies \cup_{n=1}^{\infty} A_n \in \mathcal{F}$).
3. A pair (Ω, \mathcal{F}) , where \mathcal{F} is a σ -algebra on Ω , is called a *measurable space*. If $E \in \mathcal{F}$ then E is said to be a *measurable subset* of Ω (with respect to \mathcal{F}).

DEFINITION Measures and measure spaces: (2.5)

Let (Ω, \mathcal{F}) be a measurable space. A function $\mu: \mathcal{F} \rightarrow [0, \infty]$ is called a *measure* on (Ω, \mathcal{F}) if

1. $\mu(\emptyset) = 0$;
2. if $A_i \in \mathcal{F}$ for $i = 1, 2, \dots$ and the A_i are disjoint then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

A triple $(\Omega, \mathcal{F}, \mu)$, where \mathcal{F} is a σ -algebra and μ is a measure on (Ω, \mathcal{F}) , is called a *measure space*.

PROPOSITION (2.6)

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Then

1. if $A \subset B$ then $\mu(A) \leq \mu(B)$;
2. if $A_n \in \mathcal{F}$ and $A_n \uparrow$ (that is, $A_n \subset A_{n+1}$ for all n) then $\mu(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mu(A_n)$;
3. if $A_n \in \mathcal{F}$ and $A_n \downarrow$ (that is, $A_{n+1} \subset A_n$ for all n) and $\mu(A_1) < \infty$ then $\mu(\bigcap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mu(A_n)$

Proof. 1. If $A \subset B$ then $B = A \cup (B \cap A^c)$. Then, since A and $B \cap A^c$ are disjoint, we have that $\mu(B) = \mu(A) + \mu(B \cap A^c) \geq \mu(A)$.

2. Let $E_1 = A_1$ and $E_n = A_n \setminus A_{n-1}$ for $n \geq 2$. Then the E_n are disjoint, and $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} E_n$. So

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

But also, since $A_n \uparrow$, we have $\bigcup_{k=1}^n E_k = A_n$, and thus

$$\sum_{n=1}^{\infty} \mu(E_n) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \mu(E_k) = \lim_{n \rightarrow \infty} \mu\left(\bigcup_{k=1}^n E_k\right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

3. Let $B_n = A_1 \setminus A_n$. Then $B_n \uparrow$. Since $A_1 = B_n \cup A_n$ and B_n and A_n are disjoint, we have that $\mu(A_1) = \mu(B_n) + \mu(A_n)$, and hence $\mu(B_n) = \mu(A_1) - \mu(A_n)$ (as $\mu(A_n) \leq \mu(A_1) \leq \infty$). By de Morgan's law,

$$\bigcup_{n=1}^{\infty} B_n = A_1 \setminus \bigcap_{n=1}^{\infty} A_n$$

and so, by applying the previous part of the proposition to B_n , we obtain

$$\begin{aligned}\mu(A_1) - \mu\left(\bigcap_{n=1}^{\infty} A_n\right) &= \mu\left(A_1 \setminus \bigcap_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) \\ &= \lim_{n \rightarrow \infty} \mu(B_n) \\ &= \lim_{n \rightarrow \infty} (\mu(A_1) - \mu(A_n)) \\ &= \mu(A_1) - \lim_{n \rightarrow \infty} \mu(A_n)\end{aligned}$$

□

THE LEBESGUE MEASURE

The outer measure

From here on in, \mathcal{J} denotes the set of all intervals of \mathbb{R} . That is, we will simply say $I \in \mathcal{J}$ to mean that I is an interval of \mathbb{R} . Recall that for $I \in \mathcal{J}$, $|I|$ denotes the length of I , which might be ∞ if I is unbounded. Also, when we say ‘subset’, we mean a subset of \mathbb{R} unless otherwise stated.

DEFINITION Outer measure:

(3.1)

Let $A \subset \mathbb{R}$. Then the *outer measure* of A , denoted by $m^*(A)$, is defined by

$$m^*(A) = \inf \left\{ \sum_{n=1}^{\infty} |J_n| : J_n \in \mathcal{J}, A \subseteq \bigcup_{n=1}^{\infty} J_n \right\}.$$

Note that it follows immediately from this definition that

1. $m^*(\emptyset) = 0$;
2. $m^*(A) \geq 0$ for all $A \subset \mathbb{R}$;
3. $A \subset B \implies m^*(A) \leq m^*(B)$.

A very important point is that m^* is *not* a measure: even though it is defined on all subsets of \mathbb{R} it is not always countably additive. But it turns out that this is the fault of \mathbb{R} , not m^* , since really $\mathcal{P}(\mathbb{R})$ (the power set of \mathbb{R}) is just ‘too big’ for any measure to make sense of it all. Thus our next goal is to find a σ -algebra (which we call shall \mathcal{M}_{Leb}) on which m^* is countably additive. We build up to this in a sequence of lemmas and propositions, but it is a good idea to keep our goal in mind throughout.

LEMMA Countable subadditivity of the outer measure:

(3.2)

Let $\{A_n : n = 1, 2, \dots\}$ be a sequence of subsets. Then

$$m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m^*(A_n).$$

Proof. Clearly if $\sum_{n=1}^{\infty} m^*(A_n) = \infty$ then the lemma is vacuously true, so suppose that $\sum_{n=1}^{\infty} m^*(A_n) < \infty$. Then $m^*(A_n) < \infty$ for each n . By definition of $m^*(A_n)$, for every $\varepsilon > 0$ there is exists a countable cover $\{J_i^{(n)} \in \mathcal{J} : i = 1, 2, \dots\}$ of A_n such that $\sum_{i=1}^{\infty} |J_i^{(n)}| \leq m^*(A_n) + \varepsilon/2^n$. Then $\{J_i^{(n)} : n, i = 1, 2, \dots\}$ forms a countable cover

of $\bigcup_{n=1}^{\infty} A_n$, and so

$$\begin{aligned} m^*\left(\bigcup_{n=1}^{\infty} A_n\right) &\leq \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} |J_i^{(n)}| \leq \sum_{n=1}^{\infty} \left(m^*(A_n) + \frac{\varepsilon}{2^n}\right) \\ &= \sum_{n=1}^{\infty} m^*(A_n) + \varepsilon \end{aligned}$$

Thus, since $\varepsilon > 0$ is arbitrary, we are done. \square

An important thing to note is that our definition of the outer measure is only one of a few equivalent definitions. We could, instead, have said that the infimum need run over all possible countable coverings consisting only of open intervals, or half-closed, half-open intervals, or closed intervals, etc. First of all we introduce some new notation: \mathcal{J}_o is the collection of all open intervals in \mathbb{R} and similarly \mathcal{J}_c is the collection of all closed intervals in \mathbb{R} . That is, $J \in \mathcal{J}_o$ if and only if $J = (a, b)$, for some $a, b \in [-\infty, \infty]$.

So just for completeness, we now prove that the definition where the infimum is over countable open covers is equivalent to the more general one that we have given.

PROPOSITION

(3.3)

$$m^*(A) := \inf \left\{ \sum_{n=1}^{\infty} |J_n| : J_n \in \mathcal{J}, A \subseteq \bigcup_{n=1}^{\infty} J_n \right\} = \inf \left\{ \sum_{n=1}^{\infty} |J_n| : J_n \in \mathcal{J}_o, A \subseteq \bigcup_{n=1}^{\infty} J_n \right\}.$$

Proof. Suppose that $m^*(A) < \infty$, then for every $\varepsilon > 0$ there exists a countable cover $\{J_i : i = 1, 2, \dots\}$ of A such that $m^*(A) \geq \sum_{n=1}^{\infty} |J_i| - \varepsilon/2$. Let $a_i \leq b_i$ be the ends of the interval J_i , and define the slightly larger open intervals $O_i = (a_i - \varepsilon/2^{i+2}, b_i + \varepsilon/2^{i+2})$. Then $\{O_i : i = 1, 2, \dots\}$ is also a countable cover of A , and clearly $|O_i| = |J_i| + \varepsilon/2^{i+1}$. Hence

$$\begin{aligned} m^*(A) &\geq \sum_{i=1}^{\infty} |J_i| - \frac{\varepsilon}{2} = \sum_{i=1}^{\infty} |O_i| - \sum_{i=1}^{\infty} \frac{\varepsilon}{2^{i+1}} - \frac{\varepsilon}{2} \\ &= \sum_{i=1}^{\infty} |O_i| - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} \\ &= \sum_{i=1}^{\infty} |O_i| - \varepsilon. \end{aligned}$$

Since this is true for arbitrary $\varepsilon > 0$, and $m^*(A)$ is defined as the infimum, we see that our proposition holds. \square

LEMMA

(3.4)

Let $J \in \mathcal{J}$ have endpoints $a < b$. Then $m^*(J) = b - a$.

Proof. By the definition of m^* , we have $m^*(J) \leq b - a$. On the other hand, for every $\varepsilon > 0$ there exists a countable cover $\{I_k \in \mathcal{J}_o : k = 1, 2, \dots\}$ of J such that $m^*(J) \geq \sum_{k=1}^{\infty} |I_k| - \varepsilon$.

First assume that $J \in \mathcal{J}_c$, so that $J = [a, b]$. Since $[a, b]$ is compact, there is a finite subcover $\{I_{k_1}, \dots, I_{k_n}\}$ of $[a, b]$. We may assume that $I_{k_i} = (a_i, b_i)$ and arrange them in the order $a_1 \leq a_2 \leq \dots \leq a_n$. Then we must have $a_1 \leq a$, $b_n \geq b$, and $b_k \geq a_{k+1}$.

Thus

$$\sum_{k=1}^{\infty} |I_k| \leq \sum_{i=1}^n |J_{k_i}| = \sum_{i=1}^n (b_i - a_i) \geq b - a.$$

Therefore $m^*(J) \geq (b - a) - \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we have that $m^*(J) = |J|$ for all $J \in \mathcal{J}_c$.

Now let $J \in \mathcal{J}_o$. Then for every $\varepsilon \in (0, \frac{b-a}{2})$ we have $m^*(J) \geq m^*([a + \varepsilon, b - \varepsilon]) = b - a - 2\varepsilon$. Since $\varepsilon > 0$ is arbitrary, we have that $m^*(J) = |J|$ for all $J \in \mathcal{J}_o$ as well.

Thus for any $J \in \mathcal{J}$ with endpoints $a < b$ we have

$$b - a \geq m^*(J) \geq m^*((a, b)) = b - a$$

and we are done. \square

LEMMA

(3.5)

Let $A = \cup_{i=1}^n J_i$, where $J_i \in \mathcal{J}$ are disjoint. Then $m^*(A) = \sum_{i=1}^n |J_i|$.

Proof. We may assume that each J_i is finite, otherwise the lemma is vacuously true. Now, for every $\varepsilon > 0$ there exists a cover $\{I_k \in \mathcal{J} : k = 1, 2, \dots\}$ of A such that

$$m^*(A) \leq \sum_{k=1}^{\infty} |I_k| \leq m^*(A) + \varepsilon.$$

Note that $(I_k \cap J_i) \subset I_k$ is an interval (potentially empty), and $\{I_k \cap J_i : i = 1, \dots, n\}$ is a set of disjoint subintervals of I_k for each fixed k . Thus $\sum_{i=1}^n |I_k \cap J_i| \leq |I_k|$.

On the other hand,

$$\begin{aligned} |J_i| = m^*(J_i) &= m^*\left(\bigcup_{k=1}^{\infty} (I_k \cap J_i)\right) \leq \sum_{k=1}^{\infty} m^*(I_k \cap J_i) \\ &= \sum_{k=1}^{\infty} |I_k \cap J_i|. \end{aligned}$$

Thus

$$\begin{aligned} \sum_{i=1}^n |J_i| &\leq \sum_{i=1}^n \sum_{k=1}^{\infty} |I_k \cap J_i| = \sum_{k=1}^{\infty} \sum_{i=1}^n |I_k \cap J_i| \\ &\leq \sum_{k=1}^{\infty} |I_k| \\ &\leq m^*(A) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we may conclude that $\sum_{i=1}^n |J_i| \leq m^*(A)$. The partial converse to this comes from the definition of m^* , and thus we have that $m^*(A) = \sum_{i=1}^n |J_i|$. \square

Note that even though this isn't countable additivity, it is still a promising sign that intervals seem to be very well behaved with m^* , and so it seems very likely that we will want $\mathcal{J} \subseteq \mathcal{M}_{\text{Leb}}$ when we define \mathcal{M}_{Leb} . But the way in which we shall proceed is slightly less constructive: first we define \mathcal{M}_{Leb} in a more general manner, then show that m^* is a measure on the measure space $(\mathbb{R}, \mathcal{M}_{\text{Leb}})$, and then, finally, we will get around to examining the kind of subsets that make up \mathcal{M}_{Leb} , and it will turn out that, indeed, $\mathcal{J} \subset \mathcal{M}_{\text{Leb}}$.

Before this, we examine the idea of a set with measure zero, since these sets turn out to be both important and useful later on.

Null sets and almost-sure properties

DEFINITION Null sets: (3.6)
A subset $A \subset \mathbb{R}$ is called a *null set* if it has zero outer measure. That is, $m^*(A) = 0$.

LEMMA (3.7)
Let $\{A_n : n = 1, 2, \dots\}$ be a sequence of null subsets. Then $\bigcup_{n=1}^{\infty} A_n$ is also null.

Proof. By the countable subadditivity of m^* :

$$m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m^*(A_n) = 0.$$

□

REMARK (3.8)
All countable subsets $A \subset \mathbb{R}$ are null, but not every null set is countable.

DEFINITION The Cantor ternary set: (3.9)
Consider the closed interval $J_1^{(0)} = [0, 1]$. First divide $J_1^{(0)}$ into three equal subintervals: the middle open interval $I_1^{(1)} = (\frac{1}{3}, \frac{2}{3})$ is removed from $J_1^{(0)}$, and the remaining two closed subintervals are labelled $J_1^{(1)} = [0, \frac{1}{3}]$ and $J_2^{(1)} = [\frac{2}{3}, 1]$.

Now we repeat this idea on each of the closed intervals $J_1^{(1)}$ and $J_2^{(1)}$ by taking out their middle thirds (labelled $I_1^{(2)}$ and $I_2^{(2)}$, respectively), and naming the four remaining closed intervals $J_1^{(2)}$, $J_2^{(2)}$, $J_3^{(2)}$, and $J_4^{(2)}$.

Repeating this process, for each n we have 2^{n-1} disjoint open intervals $I_i^{(n)}$ (for $i = 1, \dots, 2^{n-1}$) with equal length $\frac{1}{3^n}$, and 2^n disjoint closed intervals $J_i^{(n)}$ (for $i = 1, \dots, 2^n$) with equal length $\frac{1}{3^n}$. So, for each n , we have

$$\left(\bigcup_{k=1}^n \bigcup_{i=1}^{2^{k-1}} I_i^{(k)}\right) \cup \left(\bigcup_{k=1}^{2^n} J_k^{(n)}\right) = [0, 1]$$

Then the Cantor ternary set C is defined to be the subset of $[0, 1]$ obtained by removing all middle thirds $I_i^{(n)}$. That is,

$$C = [0, 1] \setminus \left(\bigcup_{n=1}^{\infty} \bigcup_{i=1}^{2^{n-1}} I_i^{(n)}\right).$$

PROPOSITION (3.10)
The Cantor ternary set C is null but uncountable.

Proof. By definition, for every n , we have that $J_i^{(n)}$ (for $i = 1, \dots, 2^n$) is a finite cover

of C , so that

$$m^*(C) \leq \sum_{i=1}^{2^n} |J_i^{(n)}| = 2^n \frac{1}{3^n} \rightarrow 0$$

Thus C is null. Showing that C is uncountable is left ‘as an exercise to the reader’. \square

DEFINITION Almost-sure properties: (3.11)

Suppose that $P(x)$ is a property depending on $x \in \mathbb{R}$. We say that the property P hold *almost surely*, or *almost everywhere*, if the set $\{x \in \mathbb{R} : \neg P(x)\}$ is a null set.

Lebesgue measurable sets and measurable functions

Really, there are many equivalent ways of defining \mathcal{M}_{Leb} , the set of all Lebesgue measurable sets. The following way (with the Carathéodory condition) has a non-examinable proof of its properties (which shall thus be omitted) and seems unintuitive at first glance (and in fact at second and third glances as well), but turns out to be rather simple to work with.

DEFINITION Lebesgue measurable sets and the Carathéodory condition: (3.12)

A subset $E \subset \mathbb{R}$ is *Lebesgue measurable* if E satisfies the *Carathéodory condition*. That is

$$E \in \mathcal{M}_{\text{Leb}} \iff m^*(F) = m^*(F \cap E) + m^*(F \cap E^c) \quad \forall F \subset \mathbb{R}$$

The collection of all Lebesgue measurable subsets is denoted by \mathcal{M}_{Leb} .

REMARK (3.13)

Since $(F \cap E)$ and $(F \cap E^c)$ are disjoint, and their union is F , by the subadditivity of m^* we have $m^*(F) \leq m^*(F \cap E) + m^*(F \cap E^c)$ for all $F \subset \mathbb{R}$. Thus the Carathéodory condition is equivalent to the inequality

$$m^*(F) \geq m^*(F \cap E) + m^*(F \setminus E) \quad \forall F \subset \mathbb{R}.$$

(Note that $F \cap E^c = F \setminus E$.)

THEOREM (3.14)

\mathcal{M}_{Leb} is a σ -algebra on \mathbb{R} and m^* is a measure on \mathcal{M}_{Leb} .

Proof. The proof of this theorem is not examinable, and has thus been omitted (due to of lack of time...). \square

DEFINITION Lebesgue measure space: (3.15)

The restriction of m^* to \mathcal{M}_{Leb} is simply denoted m , and is called the *Lebesgue measure*. The measure space $(\mathbb{R}, \mathcal{M}_{\text{Leb}}, m)$ is called the *Lebesgue measure space*.

Now we come back to the question of wondering what the σ -algebra \mathcal{M}_{Leb} actually looks like.

PROPOSITION (3.16)

Let A be a null set. Then $A \in \mathcal{M}_{\text{Leb}}$.

Proof. For every subset $F \subset \mathbb{R}$ we have $m^*(F \cap A) \leq m^*(A) = 0$. Then

$$m^*(F) \geq m^*(F \cap A^c) = m^*(F \cap A) + m^*(F \cap A^c).$$

□

PROPOSITION

(3.17)

Let $J \in \mathcal{J}$ be finite. Then $J \in \mathcal{M}_{Leb}$ and $m(J) = J$.

Proof. We already know that $m^*(J) = J$, so if we show that $J \in \mathcal{M}_{Leb}$ then we prove both parts of the proposition. That is, we need to show that $m^*(F) \geq m^*(F \cap J) + m^*(F \cap J^c)$ for all $F \subset \mathbb{R}$. This inequality is trivial if $m^*(F) = \infty$, so we can assume that $m^*(F) < \infty$. Then for every $\varepsilon > 0$ there exist intervals J_k such that $F \subset \bigcup_{k=1}^{\infty} J_k$ and $m^*(F) \leq \sum_{k=1}^{\infty} |J_k| \leq m^*(F) + \varepsilon$.

Now $J_k \cap J$ is an interval, and $J_k \cap J^c$ is a union of at most two intervals, so that

$$|J_k| = |J_k \cap J| + m^*(J_k \cap J^c) = m^*(J_k \cap J) + m^*(J_k \cap J^c)$$

for every k . Therefore

$$\begin{aligned} m^*(F) &\geq \sum_{k=1}^{\infty} |J_k| - \varepsilon = \sum_{k=1}^{\infty} (m^*(J_k \cap J) + m^*(J_k \cap J^c)) - \varepsilon \\ &= \sum_{k=1}^{\infty} m^*(J_k \cap J) + \sum_{k=1}^{\infty} m^*(J_k \cap J^c) - \varepsilon \\ &\geq m^*(F \cap J) + m^*(F \cap J^c) - \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary we must have $m^*(F) \geq m^*(F \cap J) + m^*(F \cap J^c)$ for all $F \subset \mathbb{R}$. □

DEFINITION Borel algebra:

(3.18)

The Borel σ -algebra \mathcal{M}_{Bor} is defined to be the smallest σ -algebra containing all finite intervals.

PROPOSITION

(3.19)

$\mathcal{M}_{Bor} \subsetneq \mathcal{M}_{Leb}$. That is, any Borel subset is Lebesgue measurable.

Proof. Since \mathcal{M}_{Leb} is a σ -algebra and contains all finite intervals, by definition $\mathcal{M}_{Bor} \subset \mathcal{M}_{Leb}$. □

REMARK

(3.20)

In particular, $\mathcal{J} \subsetneq \mathcal{M}_{Bor} \subsetneq \mathcal{M}_{Leb}$.

COROLLARY

(3.21)

Open and closed sets are measurable.

Proof. If U is an open set in \mathbb{R} then U can be written as a union of at most countably many open intervals, so that $U \in \mathcal{M}_{Leb}$. Since \mathcal{M}_{Leb} is a σ -algebra (and thus contains complements) any closed set is also measurable. □

PROPOSITION

(3.22)

Let $E \in \mathcal{M}_{Leb}$. Then there exist $A, B \in \mathcal{M}_{Bor}$ such that $A \subset E \subset B$, and both $B \setminus E$ and $E \setminus A$ are null sets.

Proof. This proof has been left, oh so generously, as an exercise for the reader. \square

MEASURABLE FUNCTIONS

Now we identify a class of functions, called *measurable functions*, for which we aim to define integrals. The concept of measurable functions can be developed independent of a choice of measure, so we work now with a general measurable space (Ω, \mathcal{F}) . In particular, we continue to use words such as ‘measurable’ in their most general sense. After this chapter, however, we will then move onto using such words to mean specifically ‘Lebesgue measurable’, but we will make this transition explicit.

DEFINITION \mathcal{F} -measurable functions:

(4.1)

Let $f: \Omega \rightarrow \mathbb{R}$ be a function on Ω . Then f is \mathcal{F} -measurable if $f^{-1}(I) \in \mathcal{F}$ for every $I \in \mathcal{I}$.

REMARK

(4.2)

We often write $\{f \in I\}$ to mean $f^{-1}(I) = \{x \in \Omega : f(x) \in I\}$. Similarly, we also have notations such as $\{f > a\}$ to mean $f^{-1}((a, \infty)) = \{x \in \Omega : f(x) > a\}$, and $\{f \neq g\}$ to mean $\{x \in \Omega : f(x) \neq g(x)\}$, etc.

PROPOSITION

(4.3)

Let \mathcal{B} be a collection of any type of intervals (all intervals, only open intervals, half-open half-closed intervals, half-infinite intervals, etc.) or the collection of all open subsets of \mathbb{R} . Let $\sigma\{\mathcal{B}\}$ denote the smallest σ -algebra containing \mathcal{B} . Then $\sigma\{\mathcal{B}\} = \mathcal{M}_{Bor}$ for all of the above \mathcal{B} .

Proof. Once again, the proof has been left as an exercise for the reader. \square

PROPOSITION

(4.4)

Let $f: \Omega \rightarrow \mathbb{R}$ be a function. Then $\mathcal{F}_f = \{B \subset \mathbb{R} : f^{-1}(B) \in \mathcal{F}\}$ is a σ -algebra.

Proof. Clearly both \emptyset and Ω belong to \mathcal{F}_f . Now assume that $B \in \mathcal{F}_f$. Then

$$\begin{aligned} f^{-1}(B^c) &= \{\omega : f(\omega) \in B^c\} = \{\omega : f(\omega) \notin B\} \\ &= \Omega \setminus f^{-1}(B) \\ &= (f^{-1}(B))^c. \end{aligned}$$

Thus, since \mathcal{F} is closed under complements, $B^c \in \mathcal{F}_f$ as well. Finally, suppose that $B_i \in \mathcal{F}_f$ for $i = 1, 2, \dots$. Then

$$f^{-1}\left(\bigcup_{i=1}^{\infty} B_i\right) = \bigcup_{i=1}^{\infty} f^{-1}(B_i) \in \mathcal{F}_f.$$

Hence \mathcal{F}_f is a σ -algebra. \square

THEOREM (4.5)

Let $f: \Omega \rightarrow \mathbb{R}$. Then f is \mathcal{F} -measurable if and only if $f^{-1}(G) \in \mathcal{F}$ for every $G \in \mathcal{M}_{\text{Bor}}$.

Proof. Suppose that f is \mathcal{F} -measurable, then any interval $I \in \mathcal{F}_f$. Since \mathcal{M}_{Bor} is the smallest σ -algebra containing all intervals, and since \mathcal{F}_f is a σ -algebra, it follows that $\mathcal{M}_{\text{Bor}} \subseteq \mathcal{F}_f$. Thus $f^{-1}(G) \in \mathcal{F}$ for any $G \in \mathcal{M}_{\text{Bor}}$. The other direction follows from the definition of a measurable function. \square

PROPOSITION (4.6)

$f: \Omega \rightarrow \mathbb{R}$ is \mathcal{F} -measurable if one of the following holds for all $a \in \mathbb{R}$:

1. $\{f > a\} \in \mathcal{F}$;
2. $\{f < a\} \in \mathcal{F}$;
3. $\{f \geq a\} \in \mathcal{F}$;
4. $\{f \leq a\} \in \mathcal{F}$.

Proof. This follows directly from the previous two propositions, but we shall explicitly prove, say, the first for clarity. So assume that $\{f > a\} \in \mathcal{F}$. By definition, this says that $f^{-1}((a, \infty)) \in \mathcal{F}$. Now \mathcal{F}_f is a σ -algebra, and $(a, \infty) \in \mathcal{F}_f$ for all $a \in \mathbb{R}$. Since \mathcal{M}_{Bor} is the smallest σ -algebra containing all intervals of the form (a, ∞) , we must have that $\mathcal{M}_{\text{Bor}} \subseteq \mathcal{F}_f$. But this simply says that $f^{-1}(G) \in \mathcal{F}$ for every $G \in \mathcal{M}_{\text{Bor}}$, which, by the previous theorem, is exactly what we need to say that f is \mathcal{F} -measurable. \square

PROPOSITION (4.7)

Let $A \subset \Omega$. Then $\mathbb{1}_A$ is \mathcal{F} -measurable if and only if $A \in \mathcal{F}$.

Proof. Since $\mathbb{1}_A: \Omega \rightarrow \{0, 1\}$, there are only four sets to consider the preimages of: $\{0\}$, $\{1\}$, $\{0, 1\}$, and \emptyset . The preimages of these sets are, respectively, $\{A^c\}$, $\{A\}$, Ω , and \emptyset . Thus $\mathbb{1}_A$ is measurable if and only if $A \in \mathcal{F}$. \square

DEFINITION Lebesgue measurable functions: (4.8)

A function $f: \mathbb{R} \rightarrow \mathbb{R}$ which is \mathcal{M}_{Leb} -measurable is called a *Lebesgue measurable function*, or simply a measurable function, if no confusion may arise. Similarly, a function $f: \mathbb{R} \rightarrow \mathbb{R}$ which is \mathcal{M}_{Bor} -measurable is called a *Borel measurable function*. Note that, since $\mathcal{M}_{\text{Bor}} \subset \mathcal{M}_{\text{Leb}}$, a Borel measurable function is Lebesgue measurable.

REMARK (4.9)

A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is continuous if the preimage of every open set is also open. But since \mathcal{M}_{Bor} contains all open sets, this says (after some working) that $f^{-1}(G) \in \mathcal{M}_{\text{Bor}}$ for any $G \in \mathcal{M}_{\text{Bor}}$. Hence a continuous function is Borel measurable, and thus Lebesgue measurable.

PROPOSITION (4.10)

Let $f: \Omega \rightarrow \mathbb{R}$ be \mathcal{F} -measurable, and $h: \mathbb{R} \rightarrow \mathbb{R}$ be Borel measurable. Then $h \circ f$ is \mathcal{F} -measurable.

Proof. We have that $h^{-1}(G) \in \mathcal{M}_{\text{Bor}}$ for any $G \in \mathcal{M}_{\text{Bor}}$. Thus $(h \circ f)^{-1}(G) = f^{-1}(h^{-1}(G)) \in \mathcal{F}$, and so $h \circ f$ is \mathcal{F} -measurable. \square

PROPOSITION

(4.11)

Let $f, g: \Omega \rightarrow \mathbb{R}$ be \mathcal{F} -measurable and $a \in \mathbb{R}$. Then the following are measurable:

1. af ;
2. $f \pm g$;
3. fg and f/g (if $g \neq 0$);
4. $|f|$, $f \wedge g$, and $f \vee g$;
5. f^+ and f^- .

Proof. 1. There is nothing to prove if $a = 0$, so assume that $a > 0$. Then for any $b \in \mathbb{R}$ we have $\{af > b\} = \{f > \frac{b}{a}\} \in \mathcal{F}$. Similarly for $a < 0$, we have $\{af > b\} = \{f < \frac{b}{a}\} \in \mathcal{F}$.

2. Since ag is \mathcal{F} -measurable, in particular $-g$ is \mathcal{F} -measurable, so it suffices to prove the case $f + g$. For any $b \in \mathbb{R}$ we have

$$\{f + g > b\} = \bigcup_{q \in \mathbb{Q}} [\{f > q\} \cap \{g > b - q\}].$$

Since f and g are both \mathcal{F} -measurable we have $\{f > q\}, \{g > b - q\} \in \mathcal{F}$. Therefore $\{f > q\} \cap \{g > b - q\} \in \mathcal{F}$ for all $q \in \mathbb{Q}$. Since \mathbb{Q} is countable, it follows that the union of this intersection over all $q \in \mathbb{Q}$ is also a \mathcal{F} -measurable set, and so we are done.

3. Since $h(x) = x^2$ is continuous and therefore Borel measurable, by a previous proposition f^2 is \mathcal{F} -measurable. Then $fg = \frac{1}{4}[(f + g)^2 - (f - g)^2]$ is \mathcal{F} -measurable. For f/g consider the continuous (for non-zero x) function $h(x) = 1/x$ and its composition with g .
4. Since $h(x) = |x|$ is continuous, $|f|$ is \mathcal{F} -measurable. Then $f \wedge g = \frac{1}{2}[(f + g) - |f - g|]$ is \mathcal{F} -measurable, and similarly for $f \vee g$.
5. Note that $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$.

 \square

DEFINITION Measurable functions with infinity:

(4.12)

A function $f: \Omega \rightarrow [-\infty, \infty] = [-\infty, \infty]$ is \mathcal{F} -measurable if $\{f = -\infty\}, \{f = \infty\} \in \mathcal{F}$ and $\{f \in I\} \in \mathcal{F}$ for every $I \in \mathcal{J}$.

Now we prove a very important point. One of the main problems with the Riemann integral was its inability to exchange the order of limit functions and integration except under certain specific conditions. It turns out that Lebesgue integration combats this problem much more slickly, and one of the reasons it does so is the fact that measurable functions are also very well behaved under limit operations.

THEOREM

(4.13)

The class of \mathcal{F} -measurable functions on a measurable space (Ω, \mathcal{F}) is closed under limit operations.

Proof. Let $\{f_n\}$ be a sequence of \mathcal{F} -measurable functions on Ω . Recall our definitions of \limsup and \liminf , and apply them here pointwise to our sequence of functions $\{f_n\}$. That is, $(\limsup_{n \rightarrow \infty} f_n)(x) = \sup_{m \geq n} \{f_m(x)\}$. Now let $g_n(x) = \sup_{m \geq n} \{f_m(x)\}$, so that (g_n) is a pointwise increasing sequence, and $g_n \uparrow g$, where $g = \limsup_{n \rightarrow \infty} f_n$. Similarly define (h_n) to be the pointwise decreasing sequence where $h_n(x) = \inf_{m \geq n} \{f_m(x)\}$, so that $h_n \downarrow h$, where $h = \liminf_{n \rightarrow \infty} f_n$.

For any $a \in \mathbb{R}$ we have

$$\{g_n \leq a\} = \bigcap_{m=n}^{\infty} \{f_m \leq a\}$$

which implies that g_n is \mathcal{F} -measurable (since σ -algebras are closed under complements and countable unions, and thus countable intersections). Then $h_n = -\sup_{m \geq n} \{-f_m(x)\}$, so that h_n is also \mathcal{F} -measurable.

Finally, since $\limsup_{n \rightarrow \infty} f_n = \inf_{n \geq 1} \{g_n\}$ we have that $\limsup_{n \rightarrow \infty} f_n$ is \mathcal{F} -measurable, and thus so too is $\liminf_{n \rightarrow \infty} f_n = -\limsup_{n \rightarrow \infty} (-f_n)$. So if $f = \lim_{n \rightarrow \infty} f_n$ exists then it is \mathcal{F} -measurable. \square

The following definition seems rather pointless now, but eventuallys make our notation quite a bit simpler when we get around to talking about integrals.

DEFINITION Lebesgue subspaces: (4.14)
Let $E \subset \mathbb{R}$ be Lebesgue measurable. Then

$$\begin{aligned} \mathcal{M}_{\text{Leb}}(E) &= \{E \cap A : A \in \mathcal{M}_{\text{Leb}}\} \\ &= \{A \subseteq E : A \in \mathcal{M}_{\text{Leb}}\} \end{aligned}$$

is a σ -algebra on E . Also, $\mathcal{M}_{\text{Leb}}(E) \subseteq \mathcal{M}_{\text{Leb}}$, so that the restriction of the Lebesgue measure m , denoted by $m|_E$, or simply m if the meaning is clear, is obviously a measure on $(E, \mathcal{M}_{\text{Leb}}(E))$, called the *Lebesgue measure on E* . The triple $(E, \mathcal{M}_{\text{Leb}}(E), m)$ is therefore a measure space, called a *Lebesgue subspace*. A function f defined on E which is $\mathcal{M}_{\text{Leb}}(E)$ -measurable is said to be *Lebesgue measurable on E* .

Further convergence of functions

Let (Ω, \mathcal{F}) be a measurable space. Suppose that $f_n: \Omega \rightarrow \mathbb{R}$ are \mathcal{F} -measurable. Recall that $\{f_n\}$ converges (to a finite number) at $x \in \Omega$ if and only if $\{f_n(x)\}$ is a Cauchy sequence. Hence both

$$\{f_n \rightarrow \ell \in \mathbb{R}\} = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{m,n=N}^{\infty} \left\{ |f_n - f_m| < \frac{1}{k} \right\} \quad \text{and} \quad \{f_n \nrightarrow \ell \in \mathbb{R}\} = \bigcup_{k=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{m,n=N}^{\infty} \left\{ |f_n - f_m| \geq \frac{1}{k} \right\}$$

are measurable sets.

In particular, if $f: \Omega \rightarrow \mathbb{R}$ is measurable, then both

$$\{f_n \rightarrow f\} = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \left\{ |f_n - f| < \frac{1}{k} \right\} \quad \text{and} \quad \{f_n \nrightarrow f\} = \bigcup_{k=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} \left\{ |f_n - f| \geq \frac{1}{k} \right\}$$

are measurable sets.

Now let $E \subset \mathbb{R}$, and $f, f_n: E \rightarrow \mathbb{R}$ for $n = 1, 2, \dots$ be the Lebesgue measurable functions on E . Then $f_n \rightarrow f$ almost everywhere on E (with respect to the Lebesgue measure) if and only if $m(\{f_n \nrightarrow f\} \cap E) = 0$ (by definition). But by the above, this is

to say that, for all $\varepsilon > 0$ we have

$$m\left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} \{|f_n - f| \geq \varepsilon\}\right).$$

Now assume that $m(E) < \infty$. Then, if we set $A_N = \bigcup_{n=N}^{\infty} \{|f_n - f| \geq \varepsilon\}$, we have that $A_N \in \mathcal{F}$, and $A_N \downarrow$ (that is, $A_{N+1} \subset A_N$ for each $N \in \mathbb{N}$), so by a previous lemma, we can rewrite the above condition: $f_n \rightarrow f$ almost everywhere on E if and only if, for all $\varepsilon > 0$, we have

$$\lim_{N \rightarrow \infty} m\left(\bigcup_{n=N}^{\infty} \{|f_n - f| \geq \varepsilon\}\right) = 0.$$

EXAMPLE

(4.15)

Let $E = (0, \infty)$ and $f_n(x) = \mathbb{1}_{(0, n]}$. Then $f_n \rightarrow f = \mathbb{1}_{(0, \infty)}$ everywhere. But now, for every $\varepsilon \in (0, 1)$ we have $\{|f_n - f| \geq \varepsilon\} = (n, \infty)$. Thus $m(\{|f_n - f| \geq \varepsilon\}) = \infty$. So we see that the previous equivalent statement of almost everywhere convergence does indeed require $m(E) < \infty$ in general (though it may be the case that $m(E) = \infty$ and it is true nonetheless).

On the other hand, we have

$$\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} \{|f_n - f| \geq \varepsilon\} = \bigcap_{N=1}^{\infty} (N, \infty) = \emptyset.$$

This is an example of how $f_n \rightarrow f$ everywhere (and thus, clearly, almost everywhere as well), but f_n does *not* converge to f in *measure*. This is something that will be expanded upon later on in the notes, when we come to \mathcal{L}^p spaces.

THEOREM Egorov's Theorem:

(4.16)

Let $E \subset \mathbb{R}$ be measurable and $m(E) < \infty$. Let $f_n, f: E \rightarrow \mathbb{R}$ be measurable and such that $f_n \rightarrow f$ almost everywhere on E . Then, for every $\delta > 0$ there exists a measurable $E_\delta \subset E$ such that $m(E \setminus E_\delta) < \delta$ and $f_n \rightarrow f$ uniformly on E_δ .

Proof. We use all the previous stuff in this section to prove this. Since $f_n \rightarrow f$ almost everywhere on E and $m(E) < \infty$, we have that, for every $\delta > 0$ and $k = 1, 2, \dots$, there exists some $n_k \in \mathbb{N}$ such that, for $n \geq n_k$, we have

$$m\left(\bigcup_{i=n}^{\infty} \left\{|f_i - f| \geq \frac{1}{k}\right\}\right) \leq \frac{\delta}{2^k}.$$

Now let

$$E_k = \bigcap_{i=n_k}^{\infty} \left\{|f_i - f| < \frac{1}{k}\right\} = \left\{|f_i - f| < \frac{1}{k} : i \geq n_k\right\} \quad \text{and} \quad E_\delta = \bigcap_{k=1}^{\infty} E_k.$$

Then $m(E_k^c) = \frac{\delta}{2^k}$, and

$$m(E \setminus E_\delta) = m\left(\bigcup_{k=1}^{\infty} E_k^c\right) \leq \sum_{k=1}^{\infty} m(E_k^c) \leq \sum_{k=1}^{\infty} \frac{\delta}{2^k} = \delta.$$

We now claim that $f_n \rightarrow f$ uniformly on E_δ . In fact, we see that $x \in E_\delta$ if and only if $x \in E_k$ for all $k = 1, 2, \dots$, and this is if and only if, for all $i \geq n_k$ we have $|f_i(x) - f(x)| < \frac{1}{k}$. Therefore, for all $i \geq n_k$ we have

$$\sup_{x \in E_\delta} |f_i(x) - f(x)| \leq \frac{1}{k}$$

which says that $f_n \rightarrow f$ uniformly on E_δ . \square

LEBESGUE INTEGRATION

Now, at long last, we reach the point in the notes that is the namesake of the course. First of all we define the class of simple (and, not surprisingly, measurable) functions, and then we define integrals for non-negative measurable functions using these simple functions. Finally we define Lebesgue integrals for measurable functions.

Although we only deal with the Lebesgue space, the approach we use applies to a general measure space, so could be transferred to, say, Lebesgue-Stieltjes integrals, without too much hassle.

Non-negative simple (measurable) functions

DEFINITION Simple (measurable) functions: (5.1)

Let (Ω, \mathcal{F}) be a measurable space. Then a function $\varphi: \Omega \rightarrow \mathbb{R}$ is *simple* if $\varphi = \sum_{i=1}^k c_i \mathbb{1}_{E_i}$ for some $k \in \mathbb{N}$, $c_i \in \mathbb{R}$, and $E_i \in \mathcal{F}$. If the E_i are all disjoint then we say that φ is in *standard form*.

Clearly, but importantly, a simple function is \mathcal{F} -measurable.

LEMMA (5.2)

Let $\varphi = \sum_{i=1}^k c_i \mathbb{1}_{E_i}$ be a simple function, where the E_i are all disjoint. Then φ is a non-negative function if and only if $c_i \geq 0$ for all $i = 1, \dots, k$.

The collection of all non-negative, simple functions is denoted by $\mathcal{S}^+(\Omega, \mathcal{F})$, or simply \mathcal{S}^+ if no confusion will arise.

Proof. This is obvious by definition. \square

THEOREM (5.3)

Suppose that $f: \Omega \rightarrow [0, \infty]$ is a \mathcal{F} -measurable function. Then there exists an increasing sequence of simple function (f_n) such that $f_n \uparrow f$.

Proof. For each n we define

$$f_n = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} \mathbb{1}_{E_k^{(n)}} + 2^n \mathbb{1}_{A_n}$$

where

$$E_k^{(n)} = \left\{ x : \frac{k}{2^n} \leq f(x) \leq \frac{k+1}{2^n} \right\} \quad \text{and} \quad A_n = \{x : f(x) \geq 2^n\}.$$

Now each $E_k^{(n)}$ and A_n are measurable, and $0 \leq f_n \uparrow f$. \square

COROLLARY (5.4)

A function $f: \Omega \rightarrow [-\infty, \infty]$ is measurable if and only if there is a sequence of simple measurable functions (f_n) such that $f_n \rightarrow f$.

Proof. Note that $f = f^+ - f^-$ is measurable if and only if both f^+ and f^- are, but we see that this is indeed the case by the previous theorem. \square

REMARK (5.5)

From now on, we work entirely with the Lebesgue space $(\mathbb{R}, \mathcal{M}_{\text{Leb}}, m)$, so when we use words such as ‘measurable’ we mean with respect to the Lebesgue space. Some results may hold more generally, but here we are only really interested in the Lebesgue case. If it is ever the case that we go back to talking about more general measure space theory, it will be explicitly stated beforehand.

DEFINITION Lebesgue integrals of indicator functions: (5.6)

Let $E \in \mathcal{M}_{\text{Leb}}$ be a measurable subset of \mathbb{R} . Then we define the integral of the indicator function of E to be

$$\int_{\mathbb{R}} \mathbb{1}_E dm = m(E).$$

We write dm to make clear that we are working with our Lebesgue measure m , but this can be omitted. Note that this definition is consistent with our idea of what integration should represent.

DEFINITION Lebesgue integrals of simple functions: (5.7)

Let $\varphi = \sum_{i=1}^k c_i \mathbb{1}_{E_i}$ be a simple measurable function, where $c_i \geq 0$ (that is, $\varphi \in \mathcal{S}^+$). Then the (Lebesgue) integral of φ is defined by extending our notion for indicator functions linearly:

$$\int_{\mathbb{R}} \varphi dm = \int_{\mathbb{R}} \sum_{i=1}^k c_i \mathbb{1}_{E_i} dm = \sum_{i=1}^k \int_{\mathbb{R}} \mathbb{1}_{E_i} dm = \sum_{i=1}^k c_i m(E_i)$$

where we use the convention $0 \cdot \infty = 0$.

An important point to note is that the definition of $\int_{\mathbb{R}} \varphi$ doesn’t depend on the representation of φ , as long as $\varphi \in \mathcal{S}^+$. That is, if $\varphi = \sum_{i=1}^k c_i \mathbb{1}_{E_i} = \sum_{i=1}^{\ell} d_i \mathbb{1}_{F_i}$ then $\sum_{i=1}^k c_i m(E_i) = \sum_{i=1}^{\ell} d_i m(F_i)$. This follows from the additivity of m .

If $\varphi \in \mathcal{S}^+$ then $\int_{\mathbb{R}} \varphi < \infty$ if and only if $m(E_i) < \infty$ for $i = 1, \dots, k$. If this is the case, then we say that the non-negative simple function φ is (Lebesgue) integrable.

PROPOSITION (5.8)

Let $\varphi, \psi \in \mathcal{S}^+$ and $\lambda \geq 0 \in \mathbb{R}$. Then

1. $\int_{\mathbb{R}} (\varphi + \psi) = \int_{\mathbb{R}} \varphi + \int_{\mathbb{R}} \psi$;
2. $\int_{\mathbb{R}} \lambda \varphi = \lambda \int_{\mathbb{R}} \varphi$;
3. if $\varphi \leq \psi$ then $\int_{\mathbb{R}} \varphi \leq \int_{\mathbb{R}} \psi$.

Proof. We may choose measure sets E_i for $i = 1, \dots, k$ that are all disjoint such that

$$\varphi = \sum_{i=1}^k c_i \mathbb{1}_{E_i} \quad \text{and} \quad \psi = \sum_{i=1}^k d_i \mathbb{1}_{E_i}.$$

1. Now $\varphi + \psi = \sum_{i=1}^k (c_i + d_i) \mathbb{1}_{E_i}$, so that $\int_{\mathbb{R}} (\varphi + \psi) = \sum_{i=1}^k (c_i + d_i) m(E_i) = \int_{\mathbb{R}} \varphi + \int_{\mathbb{R}} \psi$.
2. $\int_{\mathbb{R}} \lambda \varphi = \int_{\mathbb{R}} \lambda \sum_{i=1}^k c_i \mathbb{1}_{E_i} = \int_{\mathbb{R}} \sum_{i=1}^k (\lambda c_i) \mathbb{1}_{E_i} = \sum_{i=1}^k (\lambda c_i) m(E_i) = \lambda \left(\sum_{i=1}^k c_i m(E_i) \right) = \lambda \int_{\mathbb{R}} \varphi$.
3. If $\varphi \leq \psi$ then $c_i \leq d_i$ for all i , and hence $\int_{\mathbb{R}} \varphi = \sum_{i=1}^k c_i m(E_i) \leq \sum_{i=1}^k d_i m(E_i) = \int_{\mathbb{R}} \psi$.

□

Non-negative measurable functions

We can now proceed to the next step in our development of the Lebesgue integral: defining integration for all non-negative measurable functions, not just simple ones.

DEFINITION Lebesgue integrals of non-negative measurable functions: (5.9)

Let $f: \mathbb{R} \rightarrow [0, \infty]$ be measurable. Then the (Lebesgue) integral of f is defined by

$$\int_{\mathbb{R}} f = \sup \left\{ \int_{\mathbb{R}} \varphi : \varphi \in \mathcal{S}^+, \varphi \leq f \right\}.$$

A non-negative measurable function f is then said to be (*Lebesgue*) *integrable* if $\int_{\mathbb{R}} f < \infty$.

PROPOSITION (5.10)

Let $f: \mathbb{R} \rightarrow [0, \infty]$ be integrable. Then $m(\{f = \infty\}) = 0$. That is, f is finite almost everywhere.

Proof. Let $E = \{f = \infty\}$, which is measurable, and define $\varphi_n = n \mathbb{1}_E$. Then $\varphi_n \in \mathcal{S}^+$ are $\varphi_n \leq f$, so that $\int_{\mathbb{R}} f \geq \int_{\mathbb{R}} \varphi_n = n \cdot m(E)$. Thus $m(E) \leq \frac{1}{n} \int_{\mathbb{R}} f$ for all $n \in \mathbb{N}$. Letting $n \rightarrow \infty$ we get that $m(E) \leq 0$. Thus, since m is non negative, we get that $m(E) = 0$. □

The above definition of integrals for non-negative measurable functions applies equally well to measurable functions on a general measure space. In particular, it applies also to a Lebesgue subspace $(E, \mathcal{M}_{\text{Leb}}(E), m)$, where $E \in \mathcal{M}_{\text{Leb}}$, where we can define the Lebesgue integral $\int_E f$ for some function f on E . That is,

$$\int_E f = \sup \left\{ \int_E \varphi : \varphi \in \mathcal{S}^+(E, \mathcal{M}_{\text{Leb}}(E)), \varphi \leq f \right\}.$$

Similarly to before, we say that a non-negative measurable function $f: E \rightarrow [0, \infty]$ is integrable over E if $\int_E f < \infty$.

PROPOSITION (5.11)

Let $E \in \mathcal{M}_{\text{Leb}}$, and $f: E \rightarrow [0, \infty]$ be measurable. Then $\int_E f = \int_{\mathbb{R}} f \mathbb{1}_E$.

Proof. This is immediate from the definition of $\int_E f$. \square

This means that we can omit the \mathbb{R} or E from the integral sign, if no confusion may arise, and simply write $\int f$ instead.

Before we go on to finish our definition of integration, and answer the question of ‘which functions are integrable, and how do we define their integrals’, we first linger with these non-negative measurable functions a bit more to see if we can gain any useful knowledge of how they behave.

PROPOSITION (5.12)

Let f, g be non-negative measurable functions, with $f \leq g$, and $\lambda \geq 0$. Then

1. $\int \lambda f = \lambda \int f$;
2. $\int f \leq \int g$.

Proof. Both of these are immediate from the definitions, and any working follows that which we did before for non-negative simple functions. \square

The following theorem is one of the most important results in the theory of Lebesgue integration, and the proof is not examinable (and thus omitted).

THEOREM Monotone Convergence Theorem: (5.13)

Let (f_n) be an increasing sequence (that is, $f_n \leq f_{n+1}$ for all $n \in \mathbb{N}$) of non-negative measurable functions. Then

$$\int \lim_{n \rightarrow \infty} f_n = \lim_{n \rightarrow \infty} \int f_n.$$

Proof. As mentioned, the proof is not examinable. \square

Now we can obtain quite a few useful consequences from the MCT.

COROLLARY (5.14)

Let f be a non-negative measurable function. Then there exists a sequence (φ_n) with $\varphi_n \in \mathcal{S}^+$ such that $\varphi_n \uparrow f$ and $\int f = \lim_{n \rightarrow \infty} \int \varphi_n$.

Proof. This follows directly from the MCT and a previous theorem. \square

COROLLARY (5.15)

Let f, g be non-negative functions. Then $\int (f + g) = \int f + \int g$.

Proof. Choose $\varphi_n, \psi_n \in \mathcal{S}^+$ such that $\varphi_n \uparrow f$ and $\psi_n \uparrow g$. Then $\varphi_n + \psi_n \uparrow f + g$. Thus

$$\begin{aligned} \int (f + g) &= \lim_{n \rightarrow \infty} \int (\varphi_n + \psi_n) = \lim_{n \rightarrow \infty} \left(\int \varphi_n + \int \psi_n \right) \\ &= \lim_{n \rightarrow \infty} \int \varphi_n + \lim_{n \rightarrow \infty} \int \psi_n \\ &= \int f + \int g. \end{aligned}$$

\square

THEOREM MCT for series of non-negative measurable functions: (5.16)

Let (f_n) be a sequence of non-negative measurable functions. Then

$$\int \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int f_n.$$

Therefore f is integrable if and only if $\sum_{n=1}^{\infty} \int f_n < \infty$.

Proof. Apply the MCT to the sequence of partial sums (S_n) , where $S_n = \sum_{i=1}^n f_i$, to get that

$$\int \sum_{n=1}^{\infty} f_n = \lim_{n \rightarrow \infty} \int \sum_{i=1}^n f_i = \lim_{n \rightarrow \infty} \sum_{i=1}^n \int f_i = \sum_{i=1}^{\infty} \int f_i.$$

□

THEOREM (5.17)

Let $E_n \in \mathcal{M}_{Leb}$ be disjoint measurable sets for $n = 1, 2, \dots$ and $f: E \rightarrow [0, \infty]$ be measurable. Then

$$\int_{\bigcup_{n=1}^{\infty} E_n} f = \sum_{n=1}^{\infty} \int_{E_n} f.$$

Proof. Let $f_n = \sum_{i=1}^n f \mathbb{1}_{E_i}$ and $E = \bigcup_{n=1}^{\infty} E_n$. Then each f_n is measurable, and $f_n \uparrow f$ on E . Therefore, by the MCT,

$$\begin{aligned} \int_E f &= \lim_{n \rightarrow \infty} \int_E f_n = \lim_{n \rightarrow \infty} \int_E \sum_{i=1}^n f \mathbb{1}_{E_i} \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \int_E f \mathbb{1}_{E_i} = \lim_{n \rightarrow \infty} \sum_{i=1}^n \int_{E_i} f = \sum_{i=1}^{\infty} \int_{E_i} f. \end{aligned}$$

□

COROLLARY (5.18)

Let $h: \mathbb{R} \rightarrow [0, \infty]$ be Lebesgue measurable. Define $\mu(E) = \int_E h$ for any $E \in \mathcal{M}_{Leb}$. Then μ is a measure on $(\mathbb{R}, \mathcal{M}_{Leb})$. We denote $d\mu$ by $h dm$, where h is called the density of the measure μ with respect to the Lebesgue measure.

Now let $f: \mathbb{R} \rightarrow [0, \infty]$ be Lebesgue measurable. Then

$$\int f d\mu = \int fh dm.$$

Proof. First we prove this statement for simple functions. Let $\varphi \in \mathcal{S}^+$. Then $\varphi = \sum_{i=1}^k c_i \mathbb{1}_{E_i}$ for some $c_i \geq 0$ and disjoint E_i . Now, by definition,

$$\begin{aligned} \int \varphi d\mu &= \sum_{i=1}^k c_i \mu(E_i) = \sum_{i=1}^k c_i \int_{E_i} h dm \\ &= \sum_{i=1}^k c_i \int \mathbb{1}_{E_i} h dm = \int \sum_{i=1}^k c_i \mathbb{1}_{E_i} h dm = \int fh dm. \end{aligned}$$

Since f is Lebesgue measurable, there exists a sequence (φ_n) of non-negative simple functions such that $\varphi_n \uparrow f$. In particular then, (φ_n) is an increasing sequence of non-

negative functions.

Now each φ_n is Lebesgue measurable, and since μ is also a measure on \mathcal{M}_{Leb} , each φ_n is also μ measurable. Thus, by the MCT on the measure space $(\mathbb{R}, \mathcal{M}_{\text{Leb}}, \mu)$, we have that

$$\int f \, d\mu = \int \lim_{n \rightarrow \infty} \varphi_n \, d\mu = \lim_{n \rightarrow \infty} \int \varphi_n \, d\mu.$$

Since $\varphi_n \in \mathcal{S}^+$, and we have proven the statement for simple functions, we know that

$$\lim_{n \rightarrow \infty} \int \varphi_n \, d\mu = \lim_{n \rightarrow \infty} \int \varphi_n h \, d\mu.$$

Then by the MCT, since the product $\varphi_n h$ is measurable, increasing, and non negative, we have that

$$\lim_{n \rightarrow \infty} \int \varphi_n h \, d\mu = \int \lim_{n \rightarrow \infty} \varphi_n h \, d\mu = \int f h \, d\mu.$$

□

THEOREM Fatou's Lemma: (5.19)

Let $E \in \mathcal{M}_{\text{Leb}}$ and (f_n) be a sequence of measurable functions, where $f_n: E \rightarrow [0, \infty]$. Then

$$\int_E \liminf_{n \rightarrow \infty} f_n \leq \liminf_{n \rightarrow \infty} \int_E f_n.$$

Proof. This proof is also not examinable, but is included for completeness' sake.

Let $g_n(x) = \inf_{i \geq n} f_i(x)$, so that $g_n \leq f_i$ for all $i \geq n$, and $g_n \uparrow \liminf_{n \rightarrow \infty} f_n$. Now applying the MCT to (g_n) we have that $\int \liminf_{n \rightarrow \infty} f_n = \lim_{n \rightarrow \infty} \int g_n$. Also, since $\int g_n \leq \int f_i$ for all $i \geq n$, in particular $\int g_n \leq \inf_{i \geq n} \int f_i$. Thus

$$\int \liminf_{n \rightarrow \infty} f_n = \lim_{n \rightarrow \infty} \int g_n \leq \lim_{n \rightarrow \infty} \inf_{i \geq n} \int f_i = \liminf_{n \rightarrow \infty} \int f_n.$$

□

Integrable functions

We can now give a definition of integrals for all measurable functions, not just non negative ones.

DEFINITION Lebesgue integrals of measurable functions: (5.20)

Let $f: E \rightarrow [-\infty, \infty]$ be a measurable function. Then its positive part f^+ and negative part f^- are both measurable. Thus $\int f^+$ and $\int f^-$ are both well defined. If both $\int f^+ < \infty$ and $\int f^- < \infty$ then we say that f is (Lebesgue) integrable on E , and define its integral:

$$\int_E f = \int_E f^+ - \int_E f^-.$$

If f is Lebesgue integrable on E then we say that $f \in \mathcal{L}^1(E)$. From the definition we see that, if $f: E \rightarrow [-\infty, \infty]$ is measurable, then $f \in \mathcal{L}^1(E)$ if and only if $\int_E |f| < \infty$. Then $m(\{|f| = \infty\}) = 0$, by a previous proposition.

There are some important properties of this definition that we now take some time to examine.

1. Let $N \subset \mathbb{R}$ be a null set. Then any function f on N is measurable, and $\int_N f = 0$, for any such f .

2. Let $E \in \mathcal{M}_{\text{Leb}}$ and $f, g: E \rightarrow [-\infty, \infty]$ be two functions. Now let f be measurable, and g be such that $f = g$ almost everywhere on E . That is, $N = \{f \neq g\}$ is a null set. Then g is measurable too, and both N and $E \setminus N$ are measurable. Thus

$$\int_E g^+ = \int_{E \setminus N} g^+ + \int_N g^+ = \int_{E \setminus N} g^+ = \int_{E \setminus N} f^+ = \int_E f^+,$$

and similarly $\int_E g^- = \int_E f^-$. Therefore $f \in \mathcal{L}^1(E)$ if and only if $g \in \mathcal{L}^1(E)$, and if so then $\int_E f = \int_E g$.

This means that the definition of Lebesgue integrals on a measurable set E applies to measurable functions defined only almost everywhere on E .

3. Let $E, F \in \mathcal{M}_{\text{Leb}}$ with $F \subset E$ and $E \setminus F$ a null set. Let $f: E \rightarrow [-\infty, \infty]$ be measurable. Then

$$\int_E f = \int_F f + \int_{E \setminus F} f = \int_F f.$$

In particular, if $a < b$ then

$$\int_{(a,b)} f = \int_{(a,b]} f = \int_{[a,b)} f = \int_{[a,b]} f.$$

So we can use $\int_a^b f(x) dx$ to denote the Lebesgue integral of f on an interval with ends $a < b$.

Note, however, that this last point is actually a lot deeper than it sounds, since we're saying that the Lebesgue integral in this case is actually equal to the Riemann integral. A proper proof of this will come later in the notes.

THEOREM

(5.21)

Let $E \in \mathcal{M}_{\text{Leb}}$. Then $\mathcal{L}^1(E)$ is a linear space, and the integral operation $\int_E: f \mapsto \int_E f$ is linear.

Proof. This follows from the definition and a previous corollary about addition of non-negative measurable functions.

Take $f, g \in \mathcal{L}^1(E)$. Then $\int_E f^+, \int_E f^- < \infty$. Also, $(f + g)^+ \leq f^+ + g^+$, and similarly $(f + g)^- \leq f^- + g^-$. Thus $\int_E (f + g)^+, \int_E (f + g)^- < \infty$, so that $f + g \in \mathcal{L}^1(E)$.

Moreover, $(f + g)^+ - (f + g)^- = f^+ - f^- + g^+ - g^-$. Rearranging gives

$$(f + g)^+ - f^+ - g^+ = (f + g)^- - f^- - g^-,$$

and integrating this yields

$$\int_E (f + g)^+ - \int_E f^+ - \int_E g^+ = \int_E (f + g)^- - \int_E f^- - \int_E g^-.$$

Thus

$$\int_E (f + g) = \int_E f + \int_E g.$$

□

Next we state and prove a very simple comparison proposition for integrability, which turns out to be very useful in showing that various functions are integrable. However, much of this will be superseded by a later theorem stating the conditions under which Riemann and Lebesgue integrals are equivalent, but this is still an immeasurably useful proposition at times.

PROPOSITION (5.22) 1. Let $E \in \mathcal{M}_{Leb}$, and $f, g: E \rightarrow [-\infty, \infty]$ be measurable with $g \in \mathcal{L}^1(E)$ and $|f| \leq g$ almost everywhere on E . Then $f \in \mathcal{L}^1(E)$ and $\int_E |f| \leq \int_E g$.

2. Let $E \in \mathcal{M}_{Leb}$ be such that $m(E) < \infty$, and $f: E \rightarrow \mathbb{R}$ be measurable and bounded. Then $f \in \mathcal{L}^1(E)$.

Proof. 1. This is immediate by definition, since both f^+ and f^- are dominated by g under the assumptions.

2. Let $\varphi \in \mathcal{S}^+$ with $\varphi \leq |f|$. Then in particular $\varphi \leq \sup |f| < \infty$. Thus

$$\int_E \varphi \leq \int_E \sup |f| \mathbb{1}_E = \sup |f| m(E).$$

Hence

$$\int_E |f| = \sup_{\substack{\varphi \leq |f| \\ \varphi \in \mathcal{S}^+}} \int_E \varphi \leq \sup |f| m(E) < \infty.$$

Thus $f \in \mathcal{L}^1(E)$. □

PROPOSITION (5.23)

Let $E \in \mathcal{M}_{Leb}$ and $f, g \in \mathcal{L}^1(E)$ with $\int_E |f - g| = 0$. Then $f = g$ almost everywhere on E .

Proof. Let $E_n = \{|f - g| \geq \frac{1}{n}\}$. Since $\frac{1}{n} \mathbb{1}_{E_n} \in \mathcal{S}^+(E_n)$ and $\frac{1}{n} \mathbb{1}_{E_n} \leq |f - g|$ on E_n , by definition we have that

$$\int_{E_n} \frac{1}{n} \leq \int_{E_n} |f - g| \leq \int_E |f - g| = 0.$$

Thus E_n is null for $n = 1, 2, \dots$, and so the countable union $\cup_{n \in \mathbb{N}} E_n = \{|f - g| > 0\}$ is also null. □

We now state another derivative of the MCT.

THEOREM MCT for sequences of integrable functions: (5.24)

Let $E \in \mathcal{M}_{Leb}$ and (f_n) be a sequence of functions with $f_n \in \mathcal{L}^1(E)$ such that $f_n \uparrow f$ almost everywhere on E . (That is, there exists a null set $N \subset E$ such that $f_n \leq f_{n+1}$ on $E \setminus N$, and we define $f = \lim_{n \rightarrow \infty} f_n$ on $E \setminus N$.) Also let the sequence $(\int_E f_n)$ be bounded above. Then $f \in \mathcal{L}^1(E)$ and $\int_E f = \lim_{n \rightarrow \infty} \int_E f_n$.

Proof. Apply the MCT to $f_n - f_1$ (which is definitely non negative) to obtain

$$\int_E (f - f_1) = \lim_{n \rightarrow \infty} \int_E (f_n - f_1) = \lim_{n \rightarrow \infty} \int_E f_n - \int_E f_1,$$

which is finite, and thus $f - f_1 \in \mathcal{L}^1(E)$. Hence $f \in \mathcal{L}^1(E)$ and $\int_E f = \lim_{n \rightarrow \infty} \int_E f_n$. □

At long last we describe explicitly the link between Riemann integrals and Lebesgue integrals. This turns out to be extremely useful, since after this we can use everything we know about Riemann integration, as well as all the theory that we have built up about Lebesgue integration.

REMARK

(5.25)

Generally, in both these notes and texts elsewhere, $\int_{[a,b]} f$ is the Lebesgue integral of f (if it exists), whereas $\int_a^b f(x) dx$ is the Riemann integral of f (if it exists). The fact that we previously said that we could write $\int_a^b f$ for a Lebesgue integral over any interval with end points $a < b$ really relies on this theorem.

But, however, and this is an important but, this is all purely notational convention. If there is ever any risk of confusion it is most likely that it will be explicitly stated that such and such integral is to be understood ‘in the sense of Lebesgue’, say.

THEOREM

(5.26)

Let $f: [a, b] \rightarrow \mathbb{R}$ be Riemann integrable. Then f is measurable. Further, the Lebesgue integral $\int_{[a,b]} f$ coincides with the Riemann integral of $\int_a^b f(x) dx$.

Proof. (Since f is Riemann integrable on $[a, b]$ we can quote any results about Riemann integrable functions from the Prelims integration module. The main two results are stated in the introduction to these notes, and we use them here without proof. Further, we use notation from the Prelims module as well, but most of this is stated again in the introduction.)

Since f is Riemann integrable on $[a, b]$ it is also bounded and uniformly continuous on this interval, and in particular thus measurable. Also, since it is Riemann integrable, for every $n \in \mathbb{N}$ there exist step functions (which are a specific example of simple measurable functions) φ_n and ψ_n such that

1. $\varphi_n \leq f \mathbb{1}_{(a,b)} \leq \psi_n$;
2. $\int \varphi_n \leq \int_a^b f(x) dx \leq \int \psi_n$;
3. $\int \psi_n - \int \varphi_n < \frac{1}{n}$.

(Note that we use the notation $\int \varphi$ instead of $I(\varphi)$, since for simple measurable functions the two are interchangeable, but in both cases we use the definition for simple measurable functions, not for step functions.)

Clearly we can choose $\varphi_n \uparrow$ and $\psi_n \downarrow$. Let $g = \lim_{n \rightarrow \infty} \varphi_n$ and $h = \lim_{n \rightarrow \infty} \psi_n$. Then both g and h are measurable, and

$$\int (\varphi_n - \varphi_1) = \int \varphi_n - \int \varphi_1 \leq \int_a^b f(x) dx - \int \varphi_1.$$

Thus, thanks to the MCT, $g - \varphi_1 \in \mathcal{L}^1([a, b])$, and therefore $g \in \mathcal{L}^1([a, b])$. Similarly we can show that $h \in \mathcal{L}^1([a, b])$.

Now since $I(\psi_n) - I(\varphi_n) < \frac{1}{n}$, by the MCT again we have that

$$\int \lim_{n \rightarrow \infty} \psi_n - \int \lim_{n \rightarrow \infty} \varphi_n \leq \lim_{n \rightarrow \infty} \frac{1}{n} = 0,$$

so that $\int_{[a,b]} g = \int_{[a,b]} h$. Also since $\int \varphi_n \leq \int_a^b f(x) dx \leq \int \psi_n$ for all $n \in \mathbb{N}$ we have that

$$\int_{[a,b]} g = \int_{[a,b]} h = \int_a^b f(x) dx.$$

Finally, $h - g \geq 0$ and $\int_{[a,b]}(h - g) = 0$, we have that $h = g$ almost surely on $[a, b]$. However, $g \leq f \leq h$, so that $f = g = h$ almost surely, and thus their Lebesgue integrals over $[a, b]$ will be the equivalent. Hence f is Lebesgue integrable on $[a, b]$, and

$$\int_{[a,b]} f = \int_{[a,b]} h = \int_{[a,b]} g = \int_a^b f(x) dx.$$

□

THEOREM Dominated Convergence Theorem:

(5.27)

Let $E \in \mathcal{M}_{Leb}$ and (f_n) be a sequence of measurable functions on E . Let $f = \lim_{n \rightarrow \infty} f_n$ almost surely on E . If there exists a $g \in \mathcal{L}^1(E)$ such that $|f_n| \leq g$ almost everywhere on E , then

1. f and f_n are integrable, for all $n \in \mathbb{N}$;
2. $\int_E f = \lim_{n \rightarrow \infty} \int_E f_n$.

Proof. Since f is measurable and $|f| \leq g$ almost surely, by comparison all f_n and f are integrable on E . We then apply Fatou's Lemma to $g - f_n$ to obtain

$$\int_E (g - f) = \int_E \lim_{n \rightarrow \infty} (g - f_n) \leq \limsup \int_E (g - f_n) = \int_E g - \limsup \int_E f_n.$$

Thus $\limsup \int_E f_n \leq \int_E f$. Similarly, by applying Fatou's Lemma to $g + f_n$ we get that $\int_E f \leq \liminf \int_E f_n$. That is,

$$\limsup \int_E f_n \leq \int_E f \leq \liminf \int_E f_n,$$

and since $\liminf \int_E f_n \leq \limsup \int_E f_n$, we must have that

$$\limsup \int_E f_n = \int_E f = \liminf \int_E f_n.$$

So $\lim_{n \rightarrow \infty} \int_E f_n$ exists, and equals $\int_E f$.

□

COROLLARY Bounded Convergence Theorem:

(5.28)

Let $E \in \mathcal{M}_{Leb}$ be such that $m(E) < \infty$, and (f_n) be a sequence of measurable functions with $f_n \rightarrow f$ almost everywhere on E such that $|f_n| \leq K$ for all $n \in \mathbb{N}$ almost everywhere on E , where K is a constant. Then f_n and f are integrable on E , and

$$\int_E f = \lim_{n \rightarrow \infty} \int_E f_n.$$

Proof. Apply the DCT with $g = K$.

□

REMARK

(5.29)

So in summary, we have proved in this chapter the following convergence theorems (all of which are used extremely often, and therefore necessary to remember):

1. Fatou's Lemma;
2. (a) MCT (for sequences of non-negative measurable functions);
 (b) MCT for series of non-negative measurable functions;
 (c) MCT for sequences of integrable functions;
3. DCT.

Integrals depending on parameters

Suppose we have some function f_λ , defined for some values of λ , say $\lambda \in G$, on some measurable set E . That is, we think of a function of two variables, $f(\lambda, x)$, as instead a collection of functions of a single variable, $\{f_\lambda(x) : \lambda \in G\}$. Then we can define a function $F: G \rightarrow \mathbb{R}$ by setting $F(\lambda) = \int_E f_\lambda$ (which seems arbitrary, but actually turns out to have at least some uses). This leads us to seek sufficient conditions to ensure that such an F is continuous on G .

THEOREM

(5.30)

Let $E \in \mathcal{M}_{Leb}$ and $J \in \mathcal{J}$. Let $f_\lambda: E \rightarrow [-\infty, \infty]$ be measurable for each $\lambda \in J$, and be such that $f_\lambda \rightarrow f_{\lambda_0}$ as $\lambda \rightarrow \lambda_0$ almost everywhere on E for any $\lambda_0 \in J$. If there exists a $g \in \mathcal{L}^1(E)$ such that $|f_\lambda| \leq g$ almost everywhere on E for every $\lambda \in J$ (that is, one g 'works' for all λ), then $f_\lambda \in \mathcal{L}^1(E)$ and $F(\lambda) = \int_E f_\lambda$ is continuous on J .

Proof. Since f_λ is bounded by $g \in \mathcal{L}^1(E)$, we have that $f_\lambda \in \mathcal{L}^1(E)$ for every $\lambda \in J$, and so $F(t) = \int_E f_\lambda$ is a well-defined real function on J .

To see the continuity at $\lambda_0 \in J$, consider any sequence $(\lambda_n) \subset J$ such that $\lambda_n \rightarrow \lambda_0$. Then apply the DCT to f_{λ_n} to see that

$$F(\lambda_n) = \int_E f_{\lambda_n} \xrightarrow{\lambda_n \rightarrow \lambda_0} \int_E f_{\lambda_0} = F(\lambda_0).$$

□

We can now use the above theorem to prove the following result (which can be used to solve integrals by 'differentiating under the integral').

THEOREM

(5.31)

Let $E \in \mathcal{M}_{Leb}$ and $J \in \mathcal{J}$, and $f_\lambda: E \rightarrow \mathbb{R}$ be measurable for each $\lambda \in J$, and also $f_\lambda \in \mathcal{L}^1(E)$ for each $\lambda \in J$. Define $F(\lambda) = \int_E f_\lambda$ for $\lambda \in J$. Further, let the following conditions be satisfied:

1. for every $x \in E$ the partial derivative

$$\frac{\partial}{\partial \lambda} f_\lambda(x) = \lim_{\substack{h \rightarrow 0 \\ \lambda+h \in J}} \frac{f_{\lambda+h}(x) - f_\lambda(x)}{h}$$

exists for every $\lambda \in J$;

2. there exists some $g \in \mathcal{L}^1(E)$ such that

$$\left| \frac{\partial}{\partial \lambda} f_\lambda \right| \leq g$$

almost everywhere on E for all $\lambda \in J$ (that is, one g ‘works’ for all λ).

Then F is differentiable on J , and $F'(\lambda) = \int_E \frac{\partial}{\partial \lambda} f_\lambda$. Or, more explicitly,

$$\frac{d}{d\lambda} \int_E f_\lambda = \int_E \frac{\partial}{\partial \lambda} f_\lambda.$$

Proof. Let $\lambda \in J$. We may assume that there exists $\varepsilon > 0$ such that $[\lambda, \lambda + \varepsilon) \subset J$ (and if not, we can similarly consider the case that $(\lambda - \varepsilon, \lambda] \subset J$). Consider

$$g_h(x) = \frac{f_{\lambda+h}(x) - f_\lambda(x)}{h}$$

for $x \in E$ and $h \in (0, \varepsilon)$.

By the Mean Value Theorem, there exists $\lambda_0 \in [\lambda, \lambda + \varepsilon)$ such that $\frac{\partial}{\partial \lambda} f_{\lambda_0}(x) = g_h(x)$. Thus $|g_h| \leq g$ almost surely on E for all $h \in (0, \varepsilon)$.

Applying the DCT to g_{h_n} , where $(h_n) \subset (0, \varepsilon)$ such that $h_n \rightarrow 0$, we obtain

$$\frac{F(\lambda + h_n) - F(\lambda)}{h_n} = \int_E g_{h_n} \rightarrow \int_E \frac{\partial}{\partial \lambda} f_\lambda.$$

Thus F' exists, and $F'(\lambda) = \int_E \frac{\partial}{\partial \lambda} f_\lambda$. □

EXAMPLE The Gamma function:

(5.32)

Consider the Gamma function, defined as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

Our aim is to show that Γ is continuous on $(0, \infty)$.

Let $\alpha_0 > 0$, and $f_\alpha(x) = x^{\alpha-1} e^{-x}$. Choose $J = (a, b)$ such that $0 < a < \alpha_0 < b$. Define

$$g(x) = x^{a-1} \mathbb{1}_{(0,1]} + x^{b-1} e^{-x} \mathbb{1}_{(1,\infty)}.$$

Then $g \in \mathcal{L}^1((0, \infty))$ and $|f_\alpha| \leq g$ for all $x \in (0, \infty)$ and $\alpha \in (a, b)$. By the previous theorem then, $\Gamma(\alpha)$ is continuous at α_0 . Since α_0 was arbitrary, Γ is continuous on $(0, \infty)$.

FUBINI AND TONELLI

We may construct the Lebesgue measure on \mathbb{R}^n in a similar way to the case where $n = 1$, and then proceed with our theory of integration on \mathbb{R}^n , just as before. We now consider the specific case \mathbb{R}^2 , and briefly bring ourselves up to speed in this new space, to the same point that we were with \mathbb{R} .

A speedy outline of Lebesgue theory on \mathbb{R}^2

We consider \mathbb{R}^2 as $\mathbb{R} \times \mathbb{R}$. Let $A, B \in \mathcal{M}_{\text{Leb}}$, then $A \times B \subset \mathbb{R}^2$ is called a *measurable rectangle type set*, and its measure is naturally defined to be $\mu(A \times B) = m(A)m(B)$, where we again use the convention that $0 \cdot \infty = \infty \cdot 0 = 0$. The collection of all subsets $E \subset \mathbb{R}^2$ of the form $E = \bigcup_{i=1}^k A_i \times B_i$, where $A_i, B_i \in \mathcal{M}_{\text{Leb}}$, is denoted by \mathcal{R} . For $E \in \mathcal{R}$ with the above representation, but where the $A_i \times B_i$ are all disjoint, we then define $m(E) = \sum_{i=1}^k m(A_i)m(B_i)$. This definition is independent of the choice of representation, as one would hope.

Back to our goal of constructing the Lebesgue measure m on \mathbb{R}^2 , we proceed as we did for \mathbb{R} : define the outer measure on \mathbb{R}^2 by

$$m^*(G) = \inf \left\{ \sum_{i=1}^{\infty} m(A_i)m(B_i) : A_i, B_i \in \mathcal{M}_{\text{Leb}}, G \subseteq \bigcup_{i=1}^{\infty} A_i \times B_i \right\}$$

for any subset $G \subseteq \mathbb{R}^2$.

We have the following properties of m^* on \mathbb{R}^2 which are completely similar to those of the outer measure on the real line:

1. $m^*(G) \geq 0$;
2. $m^*(E) = m(E)$ for $E \in \mathcal{R}$;
3. $m^*(G) \leq m^*(Q)$ for $G \subset Q$;
4. (countable subadditivity:) $m^*(\bigcup_{i=1}^{\infty} G_i) \leq \sum_{i=1}^{\infty} m^*(G_i)$.

We say that a subset $E \subset \mathbb{R}^2$ is Lebesgue measurable if $m^*(F) = m^*(F \cap E) + m^*(F \cap E^c)$ for any $F \subset \mathbb{R}^2$. Then the collection of all Lebesgue measurable subsets of \mathbb{R}^2 is a σ -algebra, denoted again by \mathcal{M}_{Leb} , and m^* is a measure on the measurable space $(\mathbb{R}^2, \mathcal{M}_{\text{Leb}})$, which is denoted by m . As expected, $\mathcal{R} \subset \mathcal{M}_{\text{Leb}}$.

If $E \in \mathcal{M}_{\text{Leb}}$, then $\mathcal{M}_{\text{Leb}}(E) = \{E \cap F : F \in \mathcal{M}_{\text{Leb}}\}$ is a σ -algebra on E and $\mathcal{M}_{\text{Leb}}(E) \subset \mathcal{M}_{\text{Leb}}$. Thus the restriction of the Lebesgue measure m on $\mathcal{M}_{\text{Leb}}(E)$ is a measure on $(E, \mathcal{M}_{\text{Leb}}(E))$.

Lebesgue's theory of integration is applicable to the measure spaces $(\mathbb{R}^2, \mathcal{M}_{\text{Leb}}, m)$ and $(E, \mathcal{M}_{\text{Leb}}(E), m)$ for $E \in \mathcal{M}_{\text{Leb}}$.

The integral of a non-negative measurable function or integrable function f on $E \in \mathcal{M}_{\text{Leb}}$ is denoted by $\int_E f(x, y) dx dy$ or simply $\int_E f$, and is called the *double (Lebesgue) integral*.

Let $E \subset \mathbb{R}^2$ be bounded and closed, and $f: E \rightarrow \mathbb{R}$ continuous. Then the double Riemann integral $\iint_E f(x, y) dx dy$ exists. In this case, $f \in \mathcal{L}^1(E)$ and the Lebesgue integral $\int_E f$ coincides with the Riemann double integral.

Fubini's Theorem and Tonelli's Theorem

LEMMA

(6.1)

Let φ be a non-negative simple function on \mathbb{R}^2 :

$$\varphi = \sum_{i=1}^k c_i \mathbb{1}_{A_i \times B_i},$$

where $A_i, B_i \in \mathcal{M}_{Leb}$ and $c_i \geq 0$. Then

1. $\varphi_y \in \mathcal{S}^+$ for all $y \in \mathbb{R}$, where $\varphi_y(x) = \varphi(x, y)$, and $F \in \mathcal{S}^+$ as well, where $F(y) = \int_{\mathbb{R}} \varphi_y$;
2. $\int_{\mathbb{R}} F = \int_{\mathbb{R}^2} \varphi$. That is,

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} \varphi(x, y) dx \right) dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \varphi(x, y) dy \right) dx = \int_{\mathbb{R}^2} \varphi(x, y) dx dy.$$

Proof. According to the definition of the Lebesgue measure on \mathbb{R}^2 we have

$$\int_{\mathbb{R}^2} \varphi(x, y) dx dy = \sum_{i=1}^k c_i m(A_i \times B_i) = \sum_{i=1}^k c_i m(A_i) m(B_i).$$

On the other hand, for each y , we have $\varphi_y = \varphi(\cdot, y) = \sum_{i=1}^k c_i \mathbb{1}_{B_i}(y) \mathbb{1}_{A_i}$, which is a non-negative measurable simple function on \mathbb{R} . Then

$$\begin{aligned} F(y) &= \int_{\mathbb{R}} \varphi_y = \int_{\mathbb{R}} \varphi(x, y) dx \\ &= \sum_{i=1}^k c_i \mathbb{1}_{B_i}(y) m(A_i) \\ &= \sum_{i=1}^k c_i m(A_i) \mathbb{1}_{B_i}(y), \end{aligned}$$

which is a non-negative simple function on \mathbb{R} , and we can calculate its integral:

$$\begin{aligned} \int_{\mathbb{R}} F &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \varphi(x, y) dx \right) dy = \sum_{i=1}^k c_i m(A_i) m(B_i) \\ &= \int_{\mathbb{R}^2} \varphi(x, y) dx dy. \end{aligned}$$

□

PROPOSITION

(6.2)

Let $f: \mathbb{R}^2 \rightarrow [0, \infty)$ be measurable. Then

1. $x \mapsto f(x, y)$ is measurable for almost all $y \in \mathbb{R}$;
2. $y \mapsto \int_{\mathbb{R}} f(x, y) dx$ (defined for almost all $y \in \mathbb{R}$) is non-negative and measurable;
- 3.

$$\int_{\mathbb{R}^2} f(x, y) dx dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x, y) dx \right) dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x, y) dy \right) dx.$$

Proof. The proof is a careful application of the DCT, together with the definition of the outer measure, but the details are omitted (and therefore probably, or at least hopefully, not examinable). \square

THEOREM Fubini's Theorem: (6.3)

Let $A, B \in \mathcal{M}_{Leb}$, so $A \times B \in \mathcal{M}_{Leb}(\mathbb{R}^2)$, and $f \in L^2(A \times B)$. Then

1. $f_y \in L^2(B)$ for almost all $y \in B$, where $f_y(x) = f(x, y)$ for $x \in A$, so $F(y) = \int_A f_y$ is well defined for almost all $y \in B$;
2. F , as defined above, is integrable on B (so, in particular, F is Lebesgue measurable), and $\int_B F = \int_{A \times B} f$. Therefore

$$\int_B \left(\int_A f(x, y) dx \right) dy = \int_A \left(\int_B f(x, y) dy \right) dx = \int_{A \times B} f(x, y) dx dy.$$

Proof. Apply the previous proposition to f^+ and f^- . \square

THEOREM Tonelli's Theorem: (6.4)

Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ be measurable. If either of the repeated integrals

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x, y)| dx \right) dy \quad \text{or} \quad \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x, y)| dy \right) dx$$

exists and is finite, then $f \in \mathcal{L}^1(\mathbb{R}^2)$, so that Fubini's Theorem is applicable to both f and $|f|$.

Proof. Applying the previous proposition to the assumptions gives us that $\int_{\mathbb{R}^2} |f| < \infty$, so that $f \in \mathcal{L}^1(\mathbb{R}^2)$. \square

Change of variables

THEOREM (6.5)

Let G be an open subset of \mathbb{R}^2 , and $T: G \rightarrow T(G) = E$ be injective and differentiable. Let $f: E \rightarrow \mathbb{R}$ be measurable. Then f is integrable if and only if $f \circ T(|\det J_T|)$ is integrable on G , where $\det J_T$ is the Jacobian of T . In that case,

$$\int_E f = \int_G f \circ T(|\det J_T|).$$

Recall that, if we write $T(u, v) = (x(u, v), y(u, v))$, then

$$J_T = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

For example, if we use polar coordinates, $T(r, \theta) = (x, y)$, where $x = r \cos \theta$, $y = r \sin \theta$, then $|\det J_T| = r$. Let $E = T(G)$ and f be measurable. Then f is integrable on E if and only if $f(r \cos \theta, r \sin \theta)r$ is integrable on G .

\mathcal{L}^p AND L^p SPACES

DEFINITION \mathcal{L}^p spaces: (7.1)

Let $E \in \mathcal{M}_{\text{Leb}}$ be a subset of \mathbb{R} or \mathbb{R}^2 , and $f: E \rightarrow [-\infty, \infty]$ be measurable. Then $|f|^p$ is measurable for any positive constant p . If f is measurable and $|f|^p$ is integrable on E , that is, $\int_E |f|^p < \infty$, then we say that f is p -th integrable. For fixed $p > 0$, the collection of all such functions on E is denoted by $\mathcal{L}^p(E)$.

An important property of $\mathcal{L}^p(E)$ is that it is a vector space. The fact that $\lambda f \in \mathcal{L}^p$ for some constant λ follows from definitions.

To show that $f + g \in \mathcal{L}^p$ for $f, g \in \mathcal{L}^p$, we consider the definition of a convex function. We say that a function $h: (a, b) \rightarrow \mathbb{R}$ is convex on (a, b) if, for all $s, t \in (a, b)$ and $\lambda \in [0, 1]$,

$$h(\lambda s + (1 - \lambda)t) \leq \lambda h(s) + (1 - \lambda)h(t)$$

(The left-hand side of this inequality corresponds to the segment of the graph of h varying between the points s and t , and the right-hand side corresponds to the straight line joining $h(s)$ and $h(t)$.) From last year's analysis courses, we know that, if the second derivative of h exists and is non negative on (a, b) , then h is convex on (a, b) .

Consider the function $h(t) = t^p$ for $t \geq 0$, where $p \geq 1$. Then $h''(t) = p(p-1)t^{p-2}$ is non negative on $(0, \infty)$. Thus

$$(\lambda s + (1 - \lambda)t)^p \leq \lambda s^p + (1 - \lambda)t^p$$

for all $s, t \geq 0$ and $\lambda \in [0, 1]$.

Applying this now to $(\frac{1}{2}|f| + \frac{1}{2}|g|)^p$, i.e. where $\lambda = \frac{1}{2}$, gives us

$$\left| \frac{1}{2}f + \frac{1}{2}g \right|^p \leq \left(\frac{1}{2}|f| + \frac{1}{2}|g| \right)^p \leq \frac{1}{2}|f|^p + \frac{1}{2}|g|^p.$$

Thus

$$|f + g|^p \leq \frac{1}{2}|2f|^p + \frac{1}{2}|2g|^p = 2^{p-1}(|f|^p + |g|^p),$$

and as the right-hand side is integrable, so too is the left, and thus $f + g \in \mathcal{L}^p$.

The main issue that we have with \mathcal{L}^p though is trying to define a norm on it. One of the properties of a norm $N: \mathcal{L}^p \rightarrow [0, \infty)$ is that $N(v) = 0$ if and only if $v = 0$, but if we try to define a norm on \mathcal{L}^p by something like $\|v\| = \int |f|^p$, we quickly run into the issue that there are many (in fact, infinitely many) functions who have a 'norm' of zero without being zero themselves (in fact, the best we can develop is a *seminorm*, which is simply a norm without this zero property). So we are led to the following definition.

DEFINITION L^p spaces: (7.2)

If we identify the functions in $\mathcal{L}^p(E)$ that are equal almost everywhere on E (that is, consider the quotient space, where the equivalence classes are given by functions that are equal almost everywhere) then we denote the resulting vector space by L^p .

For $p = 1$ we could try to define a norm on $L^1(E)$ by considering $\int_E |f|$. The triangle inequality for this norm follows from the definition of integration, and the other properties can be verified.

We can try to extend this to a definition of a metric for L^p .

DEFINITION The L^p ‘norm’: (7.3)

Let $p > 0$, $E \in \mathcal{M}_{\text{Leb}}$, f be measurable, and $|f|^p$ be integrable on E . Define the L^p ‘norm’, denoted by $\|\cdot\|_p$, for $f \in L^p$, as

$$\|f\| = \left(\int_E |f|^p \right)^{1/p},$$

which is a norm on $L^p(E)$ only if $p \geq 1$. Then we can define the L^p distance, denoted by d_p , as

$$d_p(f, g) = \|f - g\|_p,$$

which is a metric on $L^p(E)$ only if $p \geq 1$.

The proof that d_p is a metric is equivalent to a proof of Minkowski’s inequality, which we shall look at next, since all other properties required of d_p we have already, or can see from definitions.

REMARK (7.4)

From now on we assume that $p \geq 1$ and $E \in \mathcal{M}_{\text{Leb}}$, unless stated otherwise.

As we said previously, to see that $L^p(E)$ is a metric space under d_p for $p \geq 1$ we need to prove Minkowski’s inequality, and one of the most common ways of doing so is by using another inequality, called *Hölder’s inequality*.

LEMMA Hölder’s inequality: (7.5)

Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$, and $f \in L^p(E)$, $g \in L^q(E)$. Then $fg \in L^1(E)$, and

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

That is,

$$\int_E |fg| \leq \left(\int_E |f|^p \right)^{1/p} \left(\int_E |g|^q \right)^{1/q}$$

Proof. The proof of this inequality can be quite fiddly, and is thus not examinable. However, there is yet another inequality, called *Young’s inequality*, which states that, for p and q satisfying the assumptions of Hölder’s inequality, $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ for all $a, b \geq 0$. Then it is often the case that we use Young’s to prove Hölder’s, and Hölder’s to prove Minkowski’s. \square

LEMMA Minkowski’s inequality: (7.6)

Let $p \geq 1$ and $f, g \in L^p(E)$. Then $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

Proof. Technically this proof is not examinable, but it is included here because, by using Hölder’s inequality, it is not too messy.

If $\|f + g\|_p = 0$ then there is nothing to really prove, so assume that $\|f + g\|_p \neq 0$. Now

$$\begin{aligned} \|f + g\|_p^p &= \int_E |f + g|^p \leq \int_E (|f| + |g|)|f + g|^{p-1} \\ &= \int_E |f||f + g|^{p-1} + \int_E |g||f + g|^{p-1}. \end{aligned}$$

We know that $f, g, f + g \in L^p$, and so we can apply Hölder’s inequality here (albeit

rather confusingly, at first glance):

$$\begin{aligned} & \int_E |f| |f + g|^{p-1} + \int_E |g| |f + g|^{p-1} \\ & \leq \left(\int_E |f|^p \right)^{1/p} \left(\int_E (|f + g|^{(p-1)})^{p/(p-1)} \right)^{(p-1)/p} \\ & \quad + \left(\int_E |g|^p \right)^{1/p} \left(\int_E (|f + g|^{(p-1)})^{p/(p-1)} \right)^{(p-1)/p} \\ & = (\|f\|_p + \|g\|_p) \frac{\|f + g\|_p^p}{\|f + g\|_p}. \end{aligned}$$

So we have shown that

$$\|f + g\|_p^p \leq (\|f\|_p + \|g\|_p) \frac{\|f + g\|_p^p}{\|f + g\|_p},$$

which gives us Minkowski's inequality when we multiply both sides by $\frac{\|f + g\|_p}{\|f + g\|_p^p}$. \square

PROPOSITION

(7.7)

Let $p \geq 1$. Then $\|\cdot\|_p$ is a norm on $L^p(E)$. That is, $\|\cdot\|_p$ possesses the following three properties:

1. $\|f\|_p = 0$ if and only if $f = 0$;
2. $\|\lambda f\|_p = |\lambda| \|f\|_p$ for $f \in L^p(E)$ and $\lambda \in \mathbb{R}$;
3. $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

Remember that, since we are talking about L^p , and not \mathcal{L}^p , when we mention a function f , what we really mean is the equivalence class \bar{f} , where

$$\bar{f} = \{g \in \mathcal{L}^p : g = f \text{ almost everywhere}\}.$$

Proof. This follows from all the above, and anything that doesn't should follow reasonably immediately from definitions. \square

COROLLARY

(7.8)

Let $E \in \mathcal{M}_{Leb}$ be such that $m(E) < \infty$, and $1 \leq p \leq q$. Then $L^q(E) \subset L^p(E)$.

Proof. Let $f \in L^q(E)$. Define $g = |f|^p$ and $r = q/p$. Then $g \in L^r(E)$ and $1 \in L^s(E)$, where $s = r/(r-1)$, so that $1/r + 1/s = 1$. By Hölder's inequality on g and 1 , we then have that $g \in L^1(E)$, so that $|f|^p \in L^1(E)$. Thus $\int_E |f|^p < \infty$, and so $f \in L^p(E)$.

Hölder's inequality also tells us that $\|g\|_1 \leq \|g\|_r \|1\|_s$, which, after some algebra, gives us $\|f\|_p \leq m(E)^{1/p-1/q} \|f\|_q$, which we use to prove the next corollary. \square

COROLLARY

(7.9)

Let $E \in \mathcal{M}_{Leb}$ be such that $m(E) = 1$. Then $p \mapsto \|f\|_p$ is increasing. That is, for $1 \leq p \leq q$, we have $\|f\|_p \leq \|f\|_q$.

Proof. This follows from the previous corollary, since $m(E) = 1$. \square

Convergence in measure

Recall that we say $f_n \rightarrow f$ almost everywhere on E if $\{f_n \not\rightarrow f\}$ is a null set. We have seen before that, if $E \in \mathcal{M}_{\text{Leb}}$ is such that $m(E) < \infty$, and $f_n, f: E \rightarrow \mathbb{R}$ are measurable, then $f_n \rightarrow f$ almost everywhere on E if and only if, for all $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} m\left(\bigcup_{i=n}^{\infty} \{|f_i - f| > \varepsilon\}\right) = 0.$$

So we introduce the following definition, for reasons that will become clear.

DEFINITION Convergence in measure: (7.10)

Let $E \in \mathcal{M}_{\text{Leb}}$ and $f_n, f: E \rightarrow \mathbb{R}$ be measurable. Then $f_n \rightarrow f$ in measure if, for every $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} m(\{|f_n - f| > \varepsilon\}) = 0.$$

One reason for this definition is the following corollary.

COROLLARY (7.11)

Let $E \in \mathcal{M}_{\text{Leb}}$ be such that $m(E) < \infty$. Then $f_n \rightarrow f$ almost everywhere on E implies that $f_n \rightarrow f$ in measure.

This has a partial converse, which we will state (but not prove).

PROPOSITION (7.12)

Let $f_n \rightarrow f$ in measure. Then there exists a subsequence (f_{n_k}) such that $f_{n_k} \rightarrow f$ almost everywhere.

Proof. Just as in a race, however well paced it may be, one always ends sprinting. In this dash to the finish, it seems that we are going too quick to be expected to prove everything we pass, and so most proofs are non-examinable ones as we approach the finish line. This one, in particular, is not examinable. \square

Convergence in L^p space

DEFINITION Convergence in L^p : (7.13)

Let $p \geq 1$. We say that $f_n \rightarrow f$ in the L^p norm if $\|f_n - f\|_p \rightarrow 0$ as $n \rightarrow \infty$.

PROPOSITION (7.14)

Let $f_n \rightarrow f$ in the L^p norm. Then $f_n \rightarrow f$ in measure (and there thus exists a subsequence (f_{n_k}) such that $f_{n_k} \rightarrow f$ almost everywhere).

Proof. For any $\varepsilon > 0$ we have

$$\begin{aligned} m(\{|f_n - f| > \varepsilon\}) &= \int_E \mathbb{1}_{\{|f_n - f| > \varepsilon\}} \leq \int_E \frac{|f_n - f|^p}{\varepsilon^p} \mathbb{1}_{\{|f_n - f| > \varepsilon\}} \\ &\leq \frac{1}{\varepsilon^p} \int_E |f_n - f|^p \\ &= \frac{1}{\varepsilon^p} \|f_n - f\|_p^p, \end{aligned}$$

which tends to zero as $n \rightarrow \infty$, and so $f_n \rightarrow f$ in measure. \square

THEOREM

(7.15)

Let $p \geq 1$ and $E \in \mathcal{M}_{Leb}$. Then $L^p(E)$, with the metric $d_p(f, g) = \|f - g\|_p$, is a complete metric space. That is, every Cauchy sequence (f_n) is convergent to some $f \in L^p(E)$ in the L^p norm, and, in particular, $f_n \rightarrow f$ in measure as well, and there exists a subsequence (f_{n_k}) such that $f_{n_k} \rightarrow f$ almost everywhere.

Therefore $L^p(E)$ equipped with the norm $\|\cdot\|_p$ is a Banach space.

Proof. Being the last theorem in the module, we are given it for free, as it were. The proof is not examinable. \square