2. Integration

In this chapter we will develop a theory of integration of functions with respect to a general measure.

2.1 Measurable Functions

Just as we cannot usually measure every subset of $X$, we cannot integrate every function on $X$. We need to restrict our attention to the "measurable functions."

To motivate this, recall that if $X$ & $Y$ are topological spaces, then a function $f: X \to Y$ is continuous if

\[ E \subseteq Y \text{ is open } \Rightarrow f^{-1}(E) \subseteq X \text{ is open} \]

where $f^{-1}(E) = \{x \in X : f(x) \in E\}$ is the preimage of $E$.

The definition of measurable function is similar, but is in terms of measurable sets instead of open sets.
Definition

Let \((X, \mathcal{M})\) and \((Y, \mathcal{N})\) be measurable spaces. Then a function \(f: X \to Y\) is \((\mathcal{M}, \mathcal{N})\)-measurable (or simply measurable for short) if

\[ E \in \mathcal{N} \implies f^{-1}(E) \in \mathcal{M}. \]

That is, the inverse image of every measurable set in \(Y\) must be measurable in \(X\).

Abbreviation

\(\mathcal{M}\) = measurable

Notation

Sometimes we will write \(f: (X, \mathcal{M}) \to (Y, \mathcal{N})\) as a shorthand to denote that \(\mathcal{M}\) is the \(\sigma\)-algebra for \(X\), \(\mathcal{N}\) is the \(\sigma\)-algebra for \(Y\), \& \(f: X \to Y\). This doesn’t require \(f\) be \(\mathcal{M}\)-measurable, it just tells us what \(\sigma\)-algebras are.

Remark

Just as for continuity, measurability of a function does not tell us anything about direct images; even if \(f\) is \(\mathcal{M}\)-measurable,

\[ E \in \mathcal{M} \implies f(E) \in \mathcal{N} \text{ in general}. \]

A \(\mathcal{M}\) function need not send \(\mathcal{M}\) sets to \(\mathcal{M}\) sets.
Exercise
Let \( f : [0,1] \to [0,1] \) be a \textit{Lebesgue} function. Let \( N \) be a nonmeasurable subset of \( \mathbb{Q} \) irrationals in \([0,1]\). Let \( C \) be a Cantor set in \([0,1]\).

a. Show that \( f([0,1] \setminus C) \subseteq \mathbb{Q} \).

b. Show that \( E = f^{-1}(N) \subseteq C \) & hence is Lebesgue measurable.

c. Show that \( f(E) = N \), & hence \( f \) maps a measurable set to a nonmeasurable set.

d. Show that \( f \) is a Lebesgue measurable function.

Remarks. In general, if \( f : X \to Y \) & \( A \subseteq Y \), then we have \( f(f^{-1}(A)) \subseteq A \), but we need not have equality in general.
Exercise
To check measurability, it suffices to consider $f^{-1}(E)$ for $E$ in a generating set $\mathcal{E}$ for $\mathcal{N}$.
That is, suppose $\mathcal{N} = \mathcal{M}(\mathcal{E})$, i.e., $\mathcal{N}$ is the $\sigma$-algebra generated by $\mathcal{E}$, and show that the following two statements are equivalent.

a. $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable.

b. $f^{-1}(E) \in \mathcal{M}$ for every $E \in \mathcal{E}$.

As a corollary, we get an important special case for topological spaces, including metric spaces.

Corollary
Let $X$, $Y$ be topological spaces, & $\mathcal{B}_X$, $\mathcal{B}_Y$ be the corresponding Borel $\sigma$-algebras.

Then every continuous function $f : X \to Y$ is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable.

Proof
By definition, $f^{-1}(U)$ is open in $X$ for every
open \( U \subseteq Y \). Thus \( f^{-1}(U) \in B_X \) for every open \( U \subseteq Y \). Since the open subsets of \( Y \) generate \( B_Y \), it follows from the exercise that \( f \) is measurable.

Real-Valued Functions

Another very important special case is real-valued functions \( f : X \to \mathbb{R} \). In this case, we declare that

"\( f \) is measurable" means \((M, B)\)-measurable, where \( M \) is the \( \sigma \)-algebra on \( X \) and \( B \) is the Borel \( \sigma \)-algebra on \( \mathbb{R} \). That is, \( f \) is measurable if

\[
\forall \text{ open } U \subseteq \mathbb{R}, \quad f^{-1}(U) \text{ is measurable in } X,
\]
(because the open sets generate the Borel \( \sigma \)-algebra).
Note, however, that the Borel $\sigma$-algebra is also generated by other collections. Thus we can give an entire collection of equivalent characterizations of measurability for a real-valued function.

**Lemma**

Let $(X, M)$ be a measurable space. If $f: X \to \mathbb{R}$, then TFAE:

a. $f$ is measurable.

b. $f^{-1}(a, \infty) \in M \quad \forall a \in \mathbb{R}$

c. $f^{-1}[a, \infty) \in M \quad \forall a \in \mathbb{R}$

d. $f^{-1}(-\infty, a) \in M \quad \forall a \in \mathbb{R}$

e. $f^{-1}(-\infty, a] \in M \quad \forall a \in \mathbb{R}$

**Notation**

We will use abbreviations such as the following:

\[
\{f > a\} = \{x \in X : f(x) > a\} = f^{-1}(a, \infty)
\]

*short-hand def.*  \hspace{1cm}  \text{def of preimage}
Using the notation, if \( f : X \rightarrow \mathbb{R} \), then

\[
f \text{ is } m \iff \{ f > a \} \text{ is } m \text{ for } a \in \mathbb{R}.
\]

Show that

Indeed, in many texts, \( m \) is taken as the definition of a measurable function.

Exercise

Show that if \( f : X \rightarrow \mathbb{R} \) is \( m \), then \( \{ f = a \} \) is \( m \) for every \( a \in \mathbb{R} \).

However, the converse is false!

Remark

Suppose that \( f : \mathbb{R} \rightarrow \mathbb{R} \) is given. If we say that "\( f \) is Lebesgue measurable", we mean that we are taking \( \mathbb{R} \) as the domain to be the measurable space \((\mathbb{R}, \mathcal{L})\) where \( \mathcal{L} \) is the Lebesgue \( \sigma \)-algebra.

On the other hand, since \( f \) is real-valued, the target space or codomain is \((\mathbb{R}, \mathcal{B})\) where \( \mathcal{B} \) is the Borel \( \sigma \)-algebra.

Thus,

"\( f : \mathbb{R} \rightarrow \mathbb{R} \) is \( m \)" is shorthand for...
"f: (IR, L) → (IR, B) is measurable."

This is equivalent to saying that

\[ U \text{ is a Borel set} \Rightarrow f^{-1}(U) \text{ is Lebesgue measurable} \]

which is itself equivalent to:

\[ \{ f > a \} \text{ is Lebesgue measurable} \forall a \in IR. \]

Whenever \( f: IR \rightarrow IR \) is given \& we just say that "f is measurable", \( \mathfrak{B} \) is what we usually mean.

Sometimes, we also restrict the \( \mathfrak{B} \)-algebra in the domain to be the Borel \( \mathfrak{B} \)-algebra. To denote this, we usually write "f is Borel measurable" to mean

\[ f: (IR, \mathcal{B}) \rightarrow (IR, B) \text{ is } \mathcal{B} \]

This is equivalent to

\[ \{ f > a \} \text{ is a Borel set } \forall a \in IR. \]
Exercise
Suppose that $f : X \to \mathbb{R}$ and $A \subseteq X$ are given. Show that

\[ X \text{ is a } \sigma \text{-function } \iff A \text{ is a } \sigma \text{-set} \] (i.e., $A \in \mathcal{M}$).

Exercise
Suppose that $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable. Show that

\[ g = f \text{ a.e. } \implies g \text{ is } \sigma \text{-measurable.} \]

Remark
As a consequence, in terms of Lebesgue measurability (and often other properties), changing the values of a function on a set of measure zero has "no effect."

In particular, we can even allow functions to be undefined on a set of measure zero, e.g., we can consider the function $f(x) = 1/x$ to be a function on the domain $\mathbb{R}$ even though it is undefined at $x=0$. As far as measurability goes, we could assign $f(0)$ to be any value that we like, since $\{0\}$ is a set of measure zero.
Extended Real-Valued Functions

Another important special case is extended real-valued functions $f : X \to \overline{\mathbb{R}}$. As it turns out, aside from having to avoid undefined quantities such as $\infty - \infty$, this is very similar to the real-valued case.

**Definition**

The Borel $\sigma$-algebra on $\overline{\mathbb{R}}$ is

$$\mathcal{B}_{\overline{\mathbb{R}}} = \{ E \subseteq \overline{\mathbb{R}} : E \cap \mathbb{R} \in \mathcal{B}_\mathbb{R} \}$$

This $\sigma$-algebra is generated by the intervals $[a, \infty]$, $a \in \mathbb{R}$. Consequently, given $f : X \to \overline{\mathbb{R}}$,

$$f \text{ is } \mathcal{M} \iff f^{-1}(a, \infty] = \{ f > a \} \text{ is } \mathcal{M} \forall a \in \mathbb{R}.$$

**Remark**

We will mostly be concerned with functions $f$ that are finite a.e., i.e., for which $\{ f = \pm \infty \}$ is $\mathcal{M}$ &

$$\mu \{ f = \pm \infty \} = 0.$$

In this case we can usually "ignore" issues about undefined quantities, since they will be restricted to a set of measure zero. Note that...
If \( \mu \) is induced from an outer measure \( \mu^* \), then all sets with \( \mu^*(E) = 0 \) are \( \mathcal{M} \).

Hence, in this case we can rephrase

"\( f \) is finite a.e." as

\[ \mu^* \{ f = \pm \infty \} = 0. \]

**Exercise**

Suppose that \( \mu \) is a complete measure on \( X \), and that \( f : X \to \mathbb{R} \) is \( \mathcal{M} \). Show that

\[ g = f \text{ a.e.} \quad \Rightarrow \quad g \text{ is } \mathcal{M}. \]
Complex-valued functions

From a purely topological point of view, the complex plane \( \mathbb{C} \) can be identified with \( \mathbb{R}^2 = \mathbb{R} \times \mathbb{R} \). By considering product topologies, we can reduce the question of measurability of a complex-valued function to measurability of its real & imaginary parts:

\[
f: X \to \mathbb{C} \text{ is } \mathcal{M} \iff \text{Ref, Imf are } \mathcal{M}.
\]

We will omit the technical details & take this as the definition of measurability of a complex-valued function.

Note that a complex-valued function always takes finite (complex) values - there is no analogue of the extended reals in this setting.

Example:
Complex-valued functions arise in many contexts such as harmonic analysis (and its applied version, signal processing). For example, given \( \Re \mathbb{R} \), the function

\[
f(x) = e^{2\pi i x}, \quad x \in \mathbb{R}
\]

is an important function that maps \( \mathbb{R} \to \mathbb{C} \).
Operations on $\otimes$ functions

Now we consider how we can combine $\otimes$ functions to create other $\otimes$ functions.

Compositions

Exercise:
Let $(X, M)$, $(Y, N)$, & $(Z, \mathcal{O})$ be measurable spaces. Show that

$$f: X \rightarrow Y \otimes \quad \& \quad g: Y \rightarrow Z \otimes \quad \Rightarrow \quad g \circ f: X \rightarrow Z \otimes$$

Cautionary note
Suppose that $f, g: \mathbb{R} \rightarrow \mathbb{R}$ are Lebesgue measurable. Recall that $\otimes$ means that

$$f: (\mathbb{R}, \mathcal{L}) \rightarrow (\mathbb{R}, \mathcal{B}) \quad \& \quad g: (\mathbb{R}, \mathcal{L}) \rightarrow (\mathbb{R}, \mathcal{B})$$

Note that $\mathcal{B}$'s $\sigma$-algebras are different.

Consequently, even if $f$ & $g$ are both Lebesgue $\otimes$ measurable, their composition $g \circ f$ might not be Lebesgue $\otimes$!

In order to make this work, we need
$g$ to be Borel $\mathcal{B}$, i.e., $g$ is measurable as a mapping $g: (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$.

Exercise

a. Show that if $g: \mathbb{R} \to \mathbb{R}$, then

$$g \text{ is Borel } \mathcal{B} \Rightarrow g \text{ is Lebesgue } \mathcal{M}.$$

b. Let $f: \mathbb{R}^d \to \mathbb{R}$ be Lebesgue $\mathcal{M}$ & $g: \mathbb{R} \to \mathbb{R}$ Borel $\mathcal{B}$. Show that $g \circ f$ is Lebesgue $\mathcal{M}$.

c. Show that every continuous function $g: \mathbb{R} \to \mathbb{R}$ is Borel $\mathcal{B}$. Combining this with part b, what do you conclude?

d. Show that is $f: \mathbb{R}^d \to \mathbb{R}$ is Lebesgue $\mathcal{M}$, then so are

$$|f|, f^2, |f|^p \text{ for } p > 0, \quad e^{f(x)}, \quad \sin f(x),$$

$$f^+(x) = \max \{ f(x), 0 \}, \quad \& \quad f^-(x) = -\min \{ f(x), 0 \}.$$

e. Extend Exercise 2d to functions $f: \mathbb{R}^d \to \overline{\mathbb{R}}$ that are finite a.e.
Addition & Scalar Multiplication

Next we will show that the set of \( \mathbb{R} \) functions \( f: X \rightarrow \mathbb{R} \) forms a vector space.

Exercise

Show that if \( f, g: X \rightarrow \mathbb{R} \) and \( c \in \mathbb{R} \), then \( cf \) and \( f+g \) are both \( \mathbb{R} \).

In order to show measurability of \( f+g \), we need the following.

Theorem

If \( f, g: X \rightarrow \mathbb{R} \) are \( \mathbb{R} \), then \( \{f \geq g\} \) is \( \mathbb{R} \).

Proof:

Let \( \{r_k\}_{k \in \mathbb{N}} = \mathbb{Q} \) be an enumeration of the rationals. Then

\[
\{f \geq g\} = \bigcup_k \{f > r_k \geq g\}
\]

\[
= \bigcup_k \left( \{f > r_k\} \cap \{g < r_k\} \right)
\]

\(\ominus \) since \( f, g \in \mathbb{R} \).

Thus \( \{f \geq g\} \) is \( \mathbb{R} \).
Theorem

If \( f, g : X \rightarrow \mathbb{R} \) are \( \mathbb{A} \), then \( f + g \) is \( \mathbb{A} \).

Proof.

Fix \( a \in \mathbb{R} \). Then \( a - g = a + (-1)g \) is \( \mathbb{A} \) by an earlier exercise, so the preceding theorem implies that \( \{ f + a - g \} \) is \( \mathbb{A} \). But

\[ \{ f + a - g \} = \{ f + g > a \} \]

so \( f + g \) is \( \mathbb{A} \).

Exercise

Extend these results to functions \( f, g : X \rightarrow \mathbb{R} \) that are finite a.e. In particular, show that it does not matter what value we assign to \( f(x) + g(x) \) when \( \mathbb{R} \) has been undefined form \( \infty - \infty \) or \( -\infty + \infty \). That is, at least for functions that are finite a.e., we can replace undefined values with whatever values we like.

For functions that need not be finite a.e., see Folland's exercise 2.1 #2 on p. 48. In particular, it says that if \( f, g : X \rightarrow \mathbb{R} \) are \( \mathbb{A} \) \& C is any extended real value, the function
\[ h(x) = \begin{cases} 
  c & \text{if } f(x) + g(x) \text{ is undefined} \\
  f(x) + g(x) & \text{otherwise} 
\end{cases} \]
Products & Quotients

We pull a cute trick to show that products of \( m \) functions are \( m \). The same trick can be applied to \( \max \) & \( \min \) of two functions.

**Theorem**

If \( f, g: X \to \mathbb{R} \) are \( m \), then so are

\[ fg, \max \{f,g\}, \min \{f,g\} \]

**Proof**

For \( \max \) & \( \min \), this follows from writing

\[
\max \{f,g\} = \frac{f+g}{2} + \frac{|f-g|}{2}
\]

\[
\min \{f,g\} = \frac{f+g}{2} - \frac{|f-g|}{2}
\]

and appealing to preceding results on sums & compositions.

For products, we write

\[
fg = \frac{(f+g)^2 - (f-g)^2}{4}
\]

and argue similarly.
Exercise
Extend to functions \( f(x) \rightarrow IR \) that are finite a.e.

Suprema & Infima
We can go beyond max & min of two functions to suprema & infima of countable sequences of functions.

Theorem
Suppose \( f_k: X \rightarrow IR \) are each \( \mathbb{C} \) functions. Then so are

\[
g(x) = \sup_k f_k(x) \quad \text{and} \quad h(x) = \inf_k f_k(x).
\]

Proof/Exercise:
Show that

\[
\{ \sup_k f_k > a \} = \bigcup_k \{ f_k > a \}
\]

\[
\{ \inf_k f_k < a \} = \bigcup_k \{ f_k < a \}.
\]

As a corollary, we obtain measurability of lim sup's, lim inf's, and (if they exist) limits.
Exercise
Suppose that $f_k: X \to \mathbb{R}$ is $\infty$ for each $k \in \mathbb{N}$.

a. Show that

$$g(x) = \limsup_{k \to \infty} f_k(x) \quad \& \quad h(x) = \liminf_{k \to \infty} f_k(x)$$

are both $\infty$.

b. Show that if $k(x) = \lim_{k \to \infty} f_k(x)$ exists $\forall x \in X$, then $k$ is $\infty$.

Remark
If the measure $\mu$ on $X$ is complete, then we don't even need to assume that $\lim_{k \to \infty} f_k(x)$ exists for all $x$. In this case, we can assume that $k(x) = \lim_{k \to \infty} f_k(x)$ exists for a.e. $x$, & let it have any value we like at other points. If each $f_k$ is $\infty$, then $k$ will be $\infty$. 
Approximation by Simple Functions

Definition
The characteristic function of \( A \subseteq X \) is
\[
\chi_A(x) = \begin{cases} 
1, & x \in A, \\
0, & x \notin A. 
\end{cases}
\]

Recall that
\( \chi_A \) is a \( \mathbb{C} \) function \( \iff \) \( A \) is a \( \mathbb{C} \) set.

Definition
A simple function on \( X \) is a finite linear combination of measurable sets of characteristic functions, using complex coefficients.

That is, \( f \) is a simple function if
\[
f = \sum_{k=1}^{N} c_k \chi_{E_k}
\]
for some \( \mathbb{C} \) \( E_k \subseteq X \) and \( c_k \in \mathbb{C} \).

In particular, simple functions can be real-valued. But \( \mathbb{C} \) cannot take extended real values — a simple function must only take finite real or complex.
Thus, simple functions are finite everywhere. They can never take 0 values. Also, simple functions are \( m \).

**Exercise**

Show that

\[ f : X \to C \text{ is simple } \iff \text{ range}(f) \text{ is finite } \text{ and } \{ f = a \} \text{ is } m \ \forall a \in C. \]

\[ \iff \text{ range}(f) \text{ is finite } \text{ and } f \text{ is } m. \]

**Exercise**

Show that a simple function \( f \) can always be written

\[ f = \sum_{k=1}^{N} c_k \chi_{E_k} \]

where \( c_k \in C \) and \( E_1, \ldots, E_N \) are disjoint subsets of \( X \).

In particular, if \( \text{ range}(f) = \{ c_1, \ldots, c_N \} \), then

\[ f = \sum_{k=1}^{N} c_k \chi_{E_k} \text{ with } E_k = f^{-1}\{ c_k \}. \]

This is called the standard representation of \( f \).

Note that one of the \( c_k \) could be zero.
Exercise:

The class of simple functions on $X$ is closed under addition, scalar multiplication, & products. That is, if $f, g$ are simple & $c \in \mathbb{C}$, then so are $f + g, \ c f, \ fg$.

Try to express each of these in the form

$$\sum_{k=1}^{N} c_k X_{E_k} \text{ with disjoint } E_k$$

Note: The standard representation requires $c_k$ to all be distinct, but usually it is sufficient to express a simple function using disjoint sets $E_k$ but allowing some $c_k$'s to be the same. That is all that is being asked for here.

Idea: If $f = \sum_{k=1}^{M} c_k X_{A_k}$ & $g = \sum_{k=1}^{N} c_k X_{B_k}$, consider the sets $A_j \cap B_k$. 
Now we come to a result which is far more important than it appears. Every nonnegative function on $X$ can be written as a limit of simple functions.

**Theorem**

Suppose that $f: (X, M) \rightarrow [0, \infty]$ is in $M$.

Then $\exists$ simple functions $\phi_k$ f pointwise.

That is, $\exists$ simple functions $0 \leq \phi_1 \leq \phi_2 \leq \ldots \leq f$

such that

$$\lim_{k \to \infty} \phi_k(x) = f(x) \quad \forall x \in X$$

Further, if $f$ is bounded on some set $E \subseteq X$, then $\phi_k \to f$ uniformly on $E$. 

Proof by picture

$\phi_0$ is obtained by rounding down to the nearest integer $n \leq 1$.

$\phi_k$ is obtained by rounding down to the nearest half-integer $\frac{n}{2} \leq 2$. 
\( \phi_a \) rounds down to the nearest quarter-integer \( \frac{n}{4} \leq a \).

\( \phi_k \) rounds down to the nearest \( \frac{n}{2^k} \leq k + 1 \).

Define

\[
E_k^n = f^{-1}\left(\frac{n}{2^k}, \frac{n+1}{2^k}\right) \quad k = 0, 1, 2, \ldots
\]

and

\[
F_k = f^{-1}\left(k, \infty\right).
\]

Then \( \phi_k \) as described above is

\[
\phi_k = \sum_{n=0}^{k2^k-1} \frac{n}{2^k} E_k^n + k \frac{n}{F_k}.
\]
By construction, except on $F_k$ we have

$$0 \leq f - \phi_k \leq \frac{1}{2^k}.$$ 

Thus if $x \not\in \bigcap F_k = \{ f = 0 \}.$ \[\lim_{k \to \infty} \phi_k(x) = f(x).\]

And if $f(x) = \infty,$ then $\phi_k(x) = k \; \forall k,$ so again

$$\lim_{k \to \infty} \phi_k(x) = \infty = f(x).$$

Finally, if $f$ is bounded on $E \subseteq X,$ then

$$\exists N \; s.t. \; x \not\in F_n \; \forall k > N. \; \text{Hence} \; \forall k > N,$$

$$\sup_{x \in E} |f(x) - \phi_k(x)| \leq \frac{1}{2^k} \to 0,$$

which says $\phi_k \to f$ uniformly on $E.$ \[\square\]