Measurable functions

**Notation and Terminology.** Thoughout this section $X$ will denote a Euclidean space.

**Definition 1.** If $A \in \mathcal{M}$, then $f : A \to \mathbb{R}$ is measurable if
\[
\{x : f(x) > a\} \in \mathcal{M} \quad \text{for all } a \in \mathbb{R}.
\]

When we say a function $f$ is measurable without specifying its domain, by default we assume it is defined on all of $X$. If $B$ is a measurable subset of $\text{dom } f$, we say $f$ is measurable on $B$ if $f|B$ is measurable.

Note that if $A \in \mathcal{M}$ and $f : A \to \mathbb{R}$, then $f$ is measurable if and only if the extension of $f$ to $X$ obtained by putting $f = 0$ on $A^c$ is a measurable function. We often find it convenient to tacitly extend $f$ in this way. Thus without loss of generality we can develop the general theory of measurable functions in the context of functions from $X$ to $\mathbb{R}$.

**Observation 2.** If $A, B \in \mathcal{M}$ then $f$ is measurable on $A \cup B$ if and only if it is measurable on both $A$ and $B$. In particular, $f : X \to \mathbb{R}$ is measurable if and only if $f^{-1}(\infty), f^{-1}(-\infty) \in \mathcal{M}$ and $f$ is measurable on $f^{-1}(\mathbb{R})$.

**Lemma 3.** The following are equivalent:

1. $f$ is measurable;
2. $\{x : f(x) \leq a\} \in \mathcal{M}$ for all $a \in \mathbb{R}$;
3. $\{x : f(x) \geq a\} \in \mathcal{M}$ for all $a \in \mathbb{R}$;
4. $\{x : f(x) < a\} \in \mathcal{M}$ for all $a \in \mathbb{R}$.

**Proof.** This follows from the equalities
\[
[-\infty, a] = (a, \infty)^c \\
[a, \infty] = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, \infty\right) \\
[-\infty, a) = [a, \infty)^c.
\]

**Lemma 4.** If $f$ is real-valued then the following are equivalent:

1. $f$ is measurable;
2. $f^{-1}(a, b) \in \mathcal{M}$ for all $a, b \in \mathbb{R}$ with $a < b$;
3. $f^{-1}(A) \in \mathcal{M}$ for all open $A \subset \mathbb{R}$;
4. $f^{-1}(A) \in \mathcal{M}$ for all $A \in \mathcal{B}$.
Proof. (1) \(\implies\) (2). This follows from the equality \((a, b) = (a, \infty) \cap (-\infty, b)\).
(2) \(\implies\) (3). Every open subset of \(\mathbb{R}\) is a countable union of open intervals.
(3) \(\implies\) (iv) It suffices to show that the family
\[A := \{A \subset \mathbb{R} : f^{-1}(A) \in \mathcal{M}\}\]
is a \(\sigma\)-algebra. First, \(\emptyset \in A\) since \(\emptyset\) is open. If \(A \in A\) then
\[f^{-1}(A^c) = f^{-1}(A)^c \in \mathcal{M},\]
so \(A^c \in A\). If \(A_1, A_2, \ldots \in A\) then
\[f^{-1}\left(\bigcup_i A_i\right) = \bigcup_i f^{-1}(A) \in \mathcal{M},\]
so \(\bigcup_i A_i \in A\).
(iv) \(\implies\) (1) This is immediate since \((a, \infty) \in \mathcal{B}\) for all \(a \in \mathbb{R}\). \qed

Corollary 5. Let \(f : X \to \mathbb{R}\).

1. If \(f\) is continuous then \(f\) is measurable.
2. If \(f\) is measurable and \(g : \mathbb{R} \to \mathbb{R}\) is continuous then \(g \circ f\) is measurable.

Proof. (1) If \(A \subset \mathbb{R}\) is open then \(f^{-1}(A)\) is open, hence measurable.
(2) If \(A \subset \mathbb{R}\) is open then \(g^{-1}(A)\) is open, so
\[(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A)) \in \mathcal{M}.\] \qed

Definition 6. The positive and negative parts of \(x \in \mathbb{R}\) are
\[x^+ := \max\{x, 0\}\] and \(x^- := \max\{-x, 0\}\).

Observation 7. With the above notation,

1. \(x^+, x^- \geq 0;\)
2. \(x = x^+ - x^-;\)
3. \(x^+x^- = 0;\)
4. \(|x| = x^+ + x^-;\)
5. \(x^- = (-x)^+;\)
6. \((cx)^+ = \begin{cases} cx^+ & \text{if } c \geq 0 \\ -cx^- & \text{if } c < 0 \end{cases}, \quad \text{and} \quad (cx)^- = \begin{cases} cx^- & \text{if } c \geq 0 \\ -cx^+ & \text{if } c < 0 \end{cases}.\)

Moreover, the pair \((x^+, x^-)\) is uniquely determined by properties (1)–(3).

Definition 8. Let \(f : X \to \mathbb{R}\). The positive and negative parts of \(f\) are the functions \(f^+, f^- : X \to \mathbb{R}\) defined by
\[f^+(x) = f(x)^+ \quad \text{and} \quad f^-(x) = f(x)^-.\]

Corollary 9. If \(f : X \to \mathbb{R}\) is measurable, then so are

1. \(cf\) for all \(c \in \mathbb{R},\)
2. \(f^n\) for all \(n \in \mathbb{N},\) and
3. \(f^+, f^-, \text{ and } |f|\).
Proof. Without loss of generality $f$ is real-valued, so the result follows from Corollary 5. □

**Proposition 10.** If $f$ and $g$ are measurable, then so are $f + g$ and $fg$ (where we assume $\infty - \infty$ does not occur).

*Proof.* For the first, if $a \in \mathbb{R}$ then
\[(f + g)^{-1}(a, \infty) = \bigcup_{r \in \mathbb{Q}} \left( f^{-1}(r, \infty] \cap g^{-1}(a - r, \infty] \right) \in \mathcal{M}. \]

For $fg$, note that on $g^{-1}(\pm \infty)$ we have
\[fg = \begin{cases} g & \text{on } f^{-1}(0, \infty] \\ -g & \text{on } f^{-1}[-\infty, 0) \\ 0 & \text{on } f^{-1}(0) \end{cases}, \]
so it suffices to show $fg$ is measurable on $g^{-1}(\mathbb{R})$. Similarly, it suffices to consider $f^{-1}(\mathbb{R})$, hence $f^{-1}(\mathbb{R}) \cap g^{-1}(\mathbb{R})$. But on this set,
\[fg = \frac{1}{4}((f + g)^2 - (f - g)^2), \]
so the result follows from Corollary 9. □

**Proposition 11.** If $f_1, f_2, \ldots$ are measurable, then so are
\[\sup f_n, \quad \inf f_n, \quad \limsup f_n, \quad \liminf f_n, \quad \text{and} \quad \lim f_n \quad \text{if this exists (in } \mathbb{R}). \]

*Proof.* If $a \in \mathbb{R}$ then
\[\sup f_n(x) > a \iff \text{there exists } n \text{ such that } f_n(x) > a \iff x \in \bigcup_n f_n^{-1}(a, \infty], \]
so $\{x : \sup f_n(x) > a\} \in \mathcal{M}$. Similarly for inf, hence lim sup and lim inf.\(^1\)

For the last part, if $f_n \to f$ pointwise then $f = \limsup f_n$ is measurable. □

**Definition 12.** $P(x)$ almost everywhere, or a.e., means $\{x : P(x) \text{ is false}\}$ is a null set. Here $P(x)$ is a propositional function\(^2\) defined on $X$. $P(x)$ a.e. $x$ is sometimes used to avoid ambiguity, and $P$ a.e. when confusion seems unlikely.

Since every subset of a null set is null, we have $P$ a.e. if and only if there exists a null set $A$ such that $P(x)$ for all $x \not\in A$.

Very often our $P(x)$ will be “$f(x) = 0$”, and then we’ll say $f = 0$ a.e. Note that we have to be careful when negating this condition: it won’t do to say $\neg f \neq 0$ a.e., because this could mean that for a.e. $x$ we have $f(x) \neq 0$. We’ll usually say something like “$f \neq 0$ on a set of positive measure” to mean that it’s false that $f = 0$ a.e.

---

\(^1\) Recall that $\limsup f_n = \inf_k \sup_{n \geq k} f_n$, and similarly for $\liminf$.

\(^2\) That is, for each $x \in X$ the value $P(x)$ is a mathematical proposition — a statement which is either true or false.
**Observation 13.** Since a countable union of null sets is null, if for all \(n \in \mathbb{N}\) we have \(P_n(x)\) a.e. \(x\) (with the complementary null set depending upon \(n\)), we can conclude that a.e. \(x\) we have \(P_n(x)\) for all \(n \in \mathbb{N}\) (so that now the null set is the same for all \(n\)).

**Proposition 14.** If \(f\) is measurable and \(f = g\) a.e., then \(g\) is measurable.

*Proof.* Suppose \(m(A) = 0\) and \(f(x) = g(x)\) for all \(x \notin A\). Then \(g\) is trivially measurable on \(A\), and is measurable on \(A^c\) since \(f\) is. \(\square\)

It’ll be useful to keep in mind that the Lebesgue integral is a generalization of the Riemann integral. Much of the development of the Riemann integral can be based upon step functions. Here’s the analogue for Lebesgue integration:

**Definition 15.** A simple function is a linear combination of characteristic functions\(^3\) of measurable sets.

Thus a simple function has the form \(\sum_{i=1}^{n} c_i \chi_{A_i}\), where \(c_1, \ldots, c_n \in \mathbb{R}\) and \(A_1, \ldots, A_n \in \mathcal{M}\). It’s frequently useful to assume that the \(A_i\)’s are disjoint — this can always be arranged. Simple functions can be characterized as measurable real-valued functions with finite range. Also, simple functions suffice to generate all measurable functions in the following sense:

**Theorem 16.** If \(f\) is measurable, then there exists a sequence \((\phi_n)\) of simple functions such that \(|\phi_n| \leq |f|\) and \(\phi_n \to f\) (in \(\mathbb{R}\)). Moreover, if \(f \geq 0\) we can take \(0 \leq \phi_n \uparrow f\).

*Proof.* First assume \(f \geq 0\). For each \(n \in \mathbb{N}\) put

\[
\phi_n = \sum_{k=0}^{2^n - 1} k2^{-n}\chi_{f^{-1}[k2^{-n},(k+1)2^{-n})} + n\chi_{f^{-1}[n,\infty]}
\]

Then \((\phi_n)\) is an increasing sequence of nonnegative simple functions. To see that \(\phi_n \to f\) pointwise, fix \(x \in X\). If \(f(x) = \infty\) then

\[
\phi_n(x) = n \to f(x),
\]

while if \(f(x) < \infty\) then for all \(n > f(x)\) we have

\[
|f(x) - \phi_n(x)| < 2^{-n} \to 0.
\]

Now remove the restriction \(f \geq 0\). Apply the first part to \(f^+\) and \(f^-\), getting simple functions \(\psi_n\) and \(\xi_n\) such that \(\psi_n \to f^+\) and \(\xi_n \to f^-\). Then each \(\psi_n - \xi_n\) is a simple function, and

\[
\psi_n - \xi_n \to f^+ - f^- = f.
\]

By construction we have

\[
|\psi_n - \xi_n| = \psi_n + \xi_n \leq f^+ + f^- = |f|.
\]

*Definition 17.* A Borel function is a function \(f : X \to \overline{\mathbb{R}}\) such that

\[
\{x : f(x) > a\} \in \mathcal{B} \quad \text{for all } a \in \mathbb{R}.
\]

\(^3\)The characteristic function of a set \(A\) is the function \(\chi_A\) taking the value 1 on \(A\) and 0 on the complement.
Example 18 (Cantor function). This example constructs the Cantor function, which will be an increasing continuous function \( f : [0, 1] \to [0, 1] \) taking the Cantor set \( C \) onto \( [0, 1] \). Let \( F_0 = [0, 1] \), and for \( n \geq 1 \) let \( F_n \) be the closed set remaining after removing the open middle third from every closed interval comprising \( F_{n-1} \). Then the Cantor set \( C \) is \( \bigcap_n F_n \).

Our aim is to define \( f : C \to [0, 1] \) as follows: to each \( x \in C \) we’ll associate a sequence \((x_n)\) of 0’s and 1’s, and we’ll put \( f(x) = \sum_{n=1}^\infty x_n 2^{-n} \). We proceed to define \( x_n \): first find the interval \( I \) in \( F_{n-1} \) which contains \( x \). At the \( n \)th step in the construction of \( C \), \( I \) is split into 2 subintervals, one of which is to the left of the other. Put \( x_n \) equal to 0 if \( x \) is in the left-hand subinterval, and 1 if \( x \) is in the right-hand subinterval. Note that every sequence of 0’s and 1’s occurs in this way, and \( 0.x_1x_2 \ldots \) is (one of) the binary (that is, base 2) expansion(s) of \( f(x) \). Thus \( f \) maps \( C \) onto \( [0, 1] \).

Claim: \( f \) is increasing. Let \( x, y \in C \) with \( x < y \), and let \((x_n)\) and \((y_n)\) be the associated sequences as above. There exists \( k \in \mathbb{N} \) such that:

- \( x_n = y_n \) for \( n < k \),
- \( x_k = 0 \), and
- \( y_k = 1 \)

Then

\[
f(x) = \sum_{n<k} x_n 2^{-n} + \sum_{n>k} x_n 2^{-n}
\leq \sum_{n<k} x_n 2^{-n} + \sum_{n>k} 2^{-n}
= \sum_{n<k} x_n 2^{-n} + 2^{-k}
\leq \sum_{n<k} x_n 2^{-n} + 2^{-k} + \sum_{n>k} y_n 2^{-n}
= f(y).
\]

Thus \( f \) is increasing.

If \((x,y)\) is one of the open intervals removed at the \( k \)th step in the construction of \( C \), then for all \( n > k \) we have \( x_n = 1 \) and \( y_n = 0 \), so

\[
f(x) = \sum_{n<k} x_n 2^{-n} + \sum_{n>k} 2^{-n}
= \sum_{n<k} x_n 2^{-n} + 2^{-k}
= \sum_{n<k} x_n 2^{-n} + 2^{-k} + \sum_{n>k} y_n 2^{-n}
= f(y).
\]

Thus for each open interval \((x,y)\) removed to form \( C \) we have \( f(x) = f(y) \). Consequently, there is exactly one extension of \( f \) to an increasing function on \([0, 1]\), namely let it be constant on the closure \([x,y]\) of each of the open intervals \((x,y)\) removed to form \( C \). We denote this extension also by \( f \). This is the Cantor function.
Since \( f : [0, 1] \to [0, 1] \) is increasing and onto, it is continuous. One of the interesting properties of this function is that it maps the null set \( C \) onto the set \([0, 1]\) of measure 1.

**Example 19** (Modified Cantor function). In this example we modify the Cantor function to get a strictly increasing continuous function mapping the Cantor set onto a set of positive measure, and from this we prove the existence of a measurable set which is non-Borel. Let \( f : [0, 1] \to [0, 1] \) be the Cantor function, and define \( g : [0, 1] \to \mathbb{R} \) by \( g(x) = x + f(x) \). This is the modified Cantor function.

\( g \) is continuous since \( f \) is, and is strictly increasing since \( f \) is increasing. We have \( g(0) = 0 \) and \( g(1) = 2 \) because \( f(0) = 0 \) and \( f(1) = 1 \). Thus \( g \) maps \([0, 1]\) 1-1 onto \([0, 2]\), and has continuous inverse \( g^{-1} : [0, 2] \to [0, 1] \). We have \( m(g(C)) = 1 \) (exercise). Thus \( g(C) \) contains a nonmeasurable set \( A \). Put \( B = g^{-1}(A) \). Then \( B \subset C \), which is a null set, so \( B \) is measurable. If \( B \) were Borel, then the image \( g(B) = A \) under the homeomorphism \( g \) would also be Borel. But \( A \) is not even measurable, so \( B \) is not Borel.

We can use this example to illustrate another pathology: the characteristic function \( \chi_B \) is measurable, and \( g^{-1} \) is continuous, but the composition

\[
\chi_B \circ g^{-1} = \chi_A
\]

is nonmeasurable.