Exercise 1. Let $E$ be a given set. Prove that the following statements are equivalent.

(i) $E$ is measurable.
(ii) Given $\varepsilon > 0$, there exists an open set $O \supseteq E$ so that $m^*(O \setminus E) < \varepsilon$.
(iii) Given $\varepsilon > 0$, there exists a closed set $F \subseteq E$ so that $m^*(E \setminus F) < \varepsilon$.
(iv) There exists $G$ in $G_\delta$ (i.e. $G$ is a countable intersection of open sets) so that $E \subseteq G$ and $m^*(G \setminus E) = 0$.
(v) There exists $F$ in $F_\sigma$ (i.e. $F$ is a countable union of closed sets) so that $F \subseteq E$ and $m^*(E \setminus F) = 0$.

If $m^*(E) < \infty$, then the above statements are also equivalent to

(vi) Given $\varepsilon > 0$, there exists a finite union $U$ of open intervals so that $m^*(U \Delta E) < \varepsilon$, where $U \Delta E = (U \setminus E) \cup (E \setminus U)$ is the symmetric difference of $U$ and $E$.

Answer:

We first show that the first three conditions are equivalent by showing inclusions in the order (i)\(\iff\)(ii)\(\implies\)(iii)\(\implies\)(i). We then show that (ii)\(\iff\)(iv) and (iii)\(\iff\)(v), and consider the case (vi) in the end separately.

(i)\(\implies\)(ii). Assume that (i) holds and fix $\varepsilon > 0$. We handle the cases $m^*(E) < \infty$ and $m^*(E) = \infty$ separately. Assume that $m^*(E) < \infty$. By the definition of $m^*(E)$ there exists a countable collection of open intervals $\{I_n\}_n$ so that $E \subseteq \bigcup_n I_n$ and

$$\sum_n \ell(I_n) < m^*(E) + \varepsilon.$$

We then denote $O = \bigcup_n I_n$, which is an open set and $E \subseteq O$. Since $E$ is measurable then

$$m^*(O) = m^*(O \setminus E) + m^*(O \cap E) = m^*(O \setminus E) + m^*(E),$$

where
and since $m^*(E) < \infty$ then moreover

$$m^*(O \setminus E) = m^*(O) - m^*(E) = m^*(\bigcup^n I_n) - m^*(E)$$

$$\leq \sum_n m^*(I_n) - m^*(E) = \sum_n \ell(I_n) - m^*(E) < \varepsilon.$$ 

So the statement (ii) holds if $m^*(E) < \infty$. For the case $m^*(E) = \infty$, consider for each $k \in \mathbb{N}$ the sets $E_k = E \cap B(0, k)$, where $B(0, k)$ is the $k$-radius open ball centered at 0. Each $E_k$ is measurable as an intersection of two measurable sets and $E = \bigcup_{k=1}^{\infty} E_k$. We also have

$$m^*(E_k) \leq m^*(B(0, k)) = 2k < \infty$$

for all $k \in \mathbb{N}$, so we can apply the earlier reasoning to each set $E_k$. In other words, for each $k \in \mathbb{N}$ there exists an open set $O_k \supseteq E_k$ so that

$$m^*(O_k \setminus E_k) < \frac{\varepsilon}{2^k}.$$ 

We then denote $O = \bigcup_{k=1}^{\infty} O_k$, which is an open set and $O \supseteq E$. Moreover, we have

$$m^*(O \setminus E) = m^*(\left( \bigcup_{k=1}^{\infty} O_k \right) \setminus E) = m^*(\bigcup_{k=1}^{\infty} (O_k \setminus E_k))$$

$$\leq \sum_{k=1}^{\infty} m^*(O_k \setminus E_k) \leq \sum_{k=1}^{\infty} m^*(O_k \setminus E_k)$$

$$\leq \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon.$$ 

So we have shown that the statement (i) implies the statement (ii).

(ii)⇒(i). Assume that (ii) holds and let $\varepsilon > 0$ be fixed and $A \subseteq \mathbb{R}$ arbitrary. By (ii) we find an open set $O \supseteq E$ so that $m^*(O \setminus E) < \varepsilon$. Since open sets are measurable then

$$m^*(A) = m^*(A \cap O) + m^*(A \setminus O).$$

Since

$$A \setminus E = (A \setminus O) \cup ((O \cap A) \setminus E),$$

we have

$$m^*(A) = m^*(A \cap O) + m^*(A \setminus O) + m^*(A \setminus E).$$

Since $m^*(A \setminus E) < \varepsilon$, we have

$$m^*(A) < m^*(A \cap O) + m^*(A \setminus O) + \varepsilon.$$ 

Therefore, (i) holds.
then
\[ m^*(A \cap E) + m^*(A \setminus E) = m^*((A \cap E) \cup ((A \cap E) \setminus E)) \leq m^*(A \cap O) + m^*(A \setminus O) + m^*((A \cap O) \setminus E) \leq m^*(A) + m^*(O \setminus E) < m^*(A) + \varepsilon. \]

This holds for all \( \varepsilon > 0 \) and thus
\[ m^*(A \cap E) + m^*(A \setminus E) \leq m^*(A). \]

By subadditivity of \( m^* \) the reverse inequality follows, whence
\[ m^*(A \cap E) + m^*(A \setminus E) = m^*(A). \]

Since the choice of \( A \) was arbitrary then \( E \) is measurable.

(ii) \( \Rightarrow \) (iii). Assume that (ii) holds. By the previous part it follows that \( E \) is measurable, whence \( E^c \) is also measurable. We apply part (ii) to \( E^c \) to find an open set \( O \supseteq E^c \) so that \( m^*(O \setminus E^c) < \varepsilon \). Choose \( F \) to be the closed set \( F = O^c \), whence \( F \subseteq E \) and
\[ m^*(E \setminus F) = m^*(F^c \setminus E^c) = m^*(O \setminus E^c) < \varepsilon, \]
so we have established (iii).

(iii) \( \Rightarrow \) (i). Assume that (iii) holds and let \( \varepsilon > 0 \) be fixed and \( A \subseteq \mathbb{R} \) arbitrary. By (iii) we find a closed set \( F \subseteq E \) so that \( m^*(E \setminus F) < \varepsilon \). Since closed sets are measurable then
\[ m^*(A) = m^*(A \cap F) + m^*(A \setminus F). \]

Since
\[ A \cap E = (A \cap F) \cup (A \cap (E \setminus F)) \]
then
\[ m^*(A \cap E) + m^*(A \setminus E) = m^*((A \cap F) \cup (A \cap (E \setminus F))) + m^*(A \setminus E) \leq m^*(A \cap F) + m^*(A \cap (E \setminus F)) + m^*(A \setminus E) \leq m^*(A) + m^*(E \setminus F) < m^*(A) + \varepsilon. \]

This holds for all \( \varepsilon > 0 \) and thus
\[ m^*(A \cap E) + m^*(A \setminus E) \leq m^*(A). \]
By subadditivity of $m^*$ the reverse inequality follows, whence

$$m^*(A \cap E) + m^*(A \setminus E) = m^*(A).$$

Since the choice of $A$ was arbitrary then $E$ is measurable.

Thus far we have shown that (i)$\Leftrightarrow$(ii)$\Leftrightarrow$(iii). We then show that (ii)$\Rightarrow$(iv)$\Rightarrow$(i).

(ii)$\Rightarrow$(iv). Assume that (ii) holds. For each $n \in \mathbb{N}$ we find by (ii) an open set $O_n \supseteq E$ so that $m^*(O_n \setminus E) < \frac{1}{n}$. By choosing $G = \bigcap_{n=1}^{\infty} O_n$ we have $E \subseteq G \subseteq O_n$ for all $n \in \mathbb{N}$ and $G$ is in $G_\delta$. Now

$$m^*(G \setminus E) \leq m^*(O_n \setminus E) < \frac{1}{n}$$

for all $n \in \mathbb{N}$, which implies that $m^*(G \setminus E) = 0$. So (iv) is true.

(iv)$\Rightarrow$(i). Assume that (iv) holds and let $A \subseteq \mathbb{R}$ be arbitrary. By (iv) there exists $G$ in $G_\delta$ so that $E \subseteq G$ and $m^*(G \setminus E) = 0$. Since $G$ is in $G_\delta$ then $G$ is measurable, so we have

$$m^*(A) = m^*(A \cap G) + m^*(A \setminus G).$$

Since

$$A \setminus E = (A \setminus G) \cup ((G \cap A) \setminus E),$$

then

$$m^*(A \cap E) + m^*(A \setminus E) = m^*(A \cap E) + m^*((A \setminus G) \cup ((A \cap G) \setminus E))$$

$$\leq m^*(A \cap G) + m^*(A \setminus G) + m^*((A \cap G) \setminus E)$$

$$\leq m^*(A) + m^*(G \setminus E) = m^*(A).$$

By subadditivity of $m^*$ the reverse inequality follows, whence

$$m^*(A \cap E) + m^*(A \setminus E) = m^*(A).$$

Since the choice of $A$ was arbitrary then $E$ is measurable and (i) holds.

We then show that (iii)$\Rightarrow$(v)$\Rightarrow$(i).
Assume that (iii) holds. For each $n \in \mathbb{N}$ we find by (ii) a closed set $F_n \subseteq E$ so that $m^*(E \setminus F_n) < \frac{1}{n}$. By choosing $F = \bigcup_{n=1}^{\infty} F_n$ we have $F_n \subseteq F \subseteq E$ for all $n \in \mathbb{N}$ and $F$ is in $F_\sigma$. Now

$$m^*(E \setminus F) \leq m^*(E \setminus F_n) < \frac{1}{n}$$

for all $n \in \mathbb{N}$, which implies that $m^*(E \setminus F) = 0$. So (v) is true.

(v) $\Rightarrow$ (i). Assume that (v) holds and let $A \subseteq \mathbb{R}$ be arbitrary. By (v) we find a set $F$ in $F_\sigma$ so that $F \subseteq E$ and $m^*(E \setminus F) = 0$. Since $F_\sigma$ sets are measurable then

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

Since

$$A \cap E = (A \cap F) \cup (A \cap (E \setminus F))$$

then

$$m^*(A \cap E) + m^*(A \setminus E) = m^*((A \cap F) \cup (A \cap (E \setminus F))) + m^*(A \setminus E) \leq m^*(A \cap F) + m^*(A \cap (E \setminus F)) + m^*(A \setminus E) \leq m^*(A) + m^*(E \setminus F) = m^*(A).$$

By subadditivity of $m^*$ the reverse inequality follows, whence

$$m^*(A \cap E) + m^*(A \setminus E) = m^*(A).$$

Since the choice of $A$ was arbitrary then $E$ is measurable and (i) holds.

We have now shown that all the conditions (i) to (v) are equivalent for a given set $E$. We will then turn to the final part of this exercise by assuming that $m^*(E) < \infty$ and showing that (vi) is equivalent to (i).

(i) $\Rightarrow$ (vi). Assume that (i) holds and let $\varepsilon > 0$ be given. Since (i) $\Rightarrow$ (ii) then by (ii) we find an open set $O \supseteq E$ so that

$$m^*(O \setminus E) < \frac{\varepsilon}{2}.$$ 

Since every open set in $\mathbb{R}$ can be expressed as a countable union of disjoint open intervals, we find a countable collection of disjoint
open intervals \( \{I_k\}_k \) so that \( O = \bigcup_k I_k \). Since the collection is disjoint and \( E \) is measurable then

\[
\sum_{k=1}^{\infty} m^*(I_k) = m^*\left( \bigcup_k I_k \right) = m^*(O) = m^*(O \setminus E) + m^*(E)
\]

\[
< \frac{\varepsilon}{2} + m^*(E) < \infty
\]

and thus

\[
\lim_{n \to \infty} \sum_{k=n}^{\infty} m^*(I_k) = 0.
\]

Hence there exists \( n_0 \in \mathbb{N} \) so that

\[
\sum_{k=n_0+1}^{\infty} m^*(I_k) < \frac{\varepsilon}{2}.
\]

We then denote \( U = \bigcup_{k \leq n_0} I_k \), which is a finite union of disjoint open intervals. We now have \( (E \setminus U) \subseteq (O \setminus U) \) so

\[
m^*(E \setminus U) \leq m^*(O \setminus U) = m^*\left( \bigcup_k I_k \setminus \left( \bigcup_{k \leq n_0} I_k \right) \right)
\]

\[
= m^*\left( \bigcup_{k=n_0+1}^{\infty} I_k \right) = \sum_{k=n_0+1}^{\infty} m^*(I_k) < \frac{\varepsilon}{2}.
\]

And also

\[
m^*(U \setminus E) \leq m^*(O \setminus E) < \frac{\varepsilon}{2}.
\]

So putting together

\[
m^*(U \Delta E) = m^*((U \setminus E) \cup (E \setminus U))
\]

\[
\leq m^*(U \setminus E) + m^*(E \setminus U) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\]

We have shown that (vi) holds.

(vi)⇒(i). Assume that (vi) holds and let \( \varepsilon > 0 \) be given and \( A \subseteq \mathbb{R} \) arbitrary. By (vi) there exists a finite union \( U \) of open intervals so that \( m^*(U \Delta E) < \varepsilon \). Now observe that

\[
A \cap E = ((A \cap E) \setminus U) \cup (A \cap E \setminus U)
\]
and
\[ A \setminus E = (A \setminus (E \cup U)) \cup ((A \cap U) \setminus E). \]

Since \( U \) is measurable then
\[ m^*(A) = m^*(A \cap U) + m^*(A \setminus U), \]
so
\[
m^*(A \cap E) + m^*(A \setminus E) \leq m^*((A \cap E) \setminus U) + m^*(A \cap E \cap U)
+ m^*(A \setminus (E \cup U)) + m^*((A \cap U) \setminus E)
\leq m^*(E \setminus U) + m^*(A \cap U) + m^*(A \setminus U) + m^*(U \setminus E)
\leq 2m^*(E \Delta U) + m^*(A)
< 2\varepsilon + m^*(A). \]

Since the choice of \( \varepsilon > 0 \) was arbitrary it follows that
\[ m^*(A \cap E) + m^*(A \setminus E) \leq m^*(A). \]

By subadditivity of \( m^* \) the reverse inequality follows, whence
\[ m^*(A \cap E) + m^*(A \setminus E) = m^*(A). \]

Since the choice of \( A \) was arbitrary then \( E \) is measurable and (i) follows.

**Exercise 2.** Show that if \( E \) is a measurable set then each translate
\[ E + y = \{a + y : a \in E\} \]
is measurable.

**Answer:**
Let \( A \subseteq \mathbb{R} \) be arbitrary. Now observe that \( x \in (A - y) \setminus E \) iff \( x = a - y \) for some \( a \in A \) and \( x \notin E \), iff \( x + y \in A \) and \( x + y \notin E + y \), iff \( x + y \in A \setminus (E + y) \), iff \( x \in (A \setminus (E + y)) - y \). Hence
\[ (A - y) \setminus E = (A \setminus (E + y)) - y. \]

Similarly we see that
\[ (A - y) \cap E = (A \cap (E + y)) - y. \]
Now since $E$ is measurable and $m^*$ is translation invariant then

\[
m^*(A) = m^*(A - y) = m^*((A - y) \setminus E) + m^*((A - y) \cap E)
\]

\[
= m^*((A \setminus (E + y)) - y) + m^*((A \cap (E + y)) - y)
\]

\[
= m^*(A \setminus (E + y)) + m^*(A \cap (E + y)).
\]

Since this applies for all $A \subseteq \mathbb{R}$ then $E + y$ is measurable.

**Exercise 3.**

(i) Show that if $E$ is measurable and $E \subseteq P$ where $P$ is the set used in constructing non-measurable sets then $m^*(E) = 0$.

(ii) Show that if $A$ is any set with $m^*(A) > 0$ then there exists a non-measurable $E \subseteq A$.

**Answer:**

(i) Let $E$ be a measurable set with $E \subseteq P$. Assume towards contradiction that $m(E) > 0$. Recall that $P$ is a collection of representatives of equivalence classes (of the equivalence relation $x \sim y$ iff $x - y \in \mathbb{Q}$) produced by using the axiom of choice. This means that $E$ is a subcollection of such representatives. Consider $E_i = E + r_i$ for all $i \in \mathbb{N}$ where $\{r_i\}_{i=1}^{\infty}$ is an enumeration of $\mathbb{Q}$. By Exercise 2 the sets $E_i$ are measurable and by translation invariance $m(E) = m(E_i)$ for all $i \in \mathbb{N}$. Now the proof in the discussion at Royden’s textbook after Lemma 16 shows that $E_i \cap E_j = \emptyset$ for all $i \neq j$, so it follows that

\[
\infty = \sum_{i=1}^{\infty} m(E) = \sum_{i=1}^{\infty} m(E_i) = m\left(\bigcup_{i=1}^{\infty} E_i\right) \leq m([0,1]) = 1,
\]

which is a contradiction. Hence $m(E) = 0$.

(ii) Let $A$ be a set with $m^*(A) > 0$. Assume first that $A \subseteq (0,1)$ and let $E_i = A \cap P_i$ for all $i \in \mathbb{N}$. If $E_i$ is not measurable for some $i \in \mathbb{N}$ then $E_i \subseteq A$ is the desired set and we are done. If $E_i$ is measurable for all $i \in \mathbb{N}$, then $E_i \subseteq P_i$ implies $E_i + (1 - r_i) \subseteq P$, so $E_i + (1 - r_i)$ is a measurable subset of $P$ for all $i \in \mathbb{N}$. Part (i) implies

\[
m(E_i) = m(E_i + (1 - r_i)) = 0
\]
for all \( i \in \mathbb{N} \). But

\[
A = A \cap [0,1) = A \cap \left( \bigcup_{i=1}^{\infty} P_i \right) = \bigcup_{i=1}^{\infty} (A \cap P_i) = \bigcup_{i=1}^{\infty} E_i,
\]

so

\[
0 < m^*(A) = m^* \left( \bigcup_{i=1}^{\infty} E_i \right) \leq \sum_{i=1}^{\infty} m(E_i) = 0,
\]

which is a contradiction. Hence there exists \( i \in \mathbb{N} \) so that \( E_i \subseteq A \) is not measurable.

Assume then that \( A \) is any set with \( m^*(A) > 0 \). The non-measurable set construction can be repeated in any interval \( I_k = [k,k+1) \) for \( k \in \mathbb{Z} \) and we can thus consider the disjoint sets \( A_k = A \cap I_k \subseteq I_k \). Whenever \( m^*(A_k) > 0 \) (there must exist at least one such \( k \in \mathbb{N} \) because \( A = \bigcup_{k=1}^{\infty} A_k \)) we extract a non-measurable set \( N_k \subseteq A_k \) by repeating the argument in the previous part. Then \( N_k \subseteq A_k \subseteq A \) is the desired non-measurable subset.

Exercise 4. Prove the following. Let \( f \) be a measurable function defined on an interval \([a,b]\), and assume that \( f \) takes the values \( \pm \infty \) only on a set of measure zero. Then given \( \varepsilon > 0 \), we can find a step function \( g \) and a continuous function \( h \) so that \( |f - g| < \varepsilon \) and \( |f - h| < \varepsilon \) except on a set of measure \( < \varepsilon \). If in addition \( m \leq f \leq M \) then we may choose \( g \) and \( h \) so that \( m \leq g \leq M \) and \( m \leq h \leq M \).

Answer:

Let \( f \) be a measurable function on \([a,b]\) that takes the values \( \pm \infty \) only on a set of measure zero, and let \( \varepsilon > 0 \) be fixed.

We do the exercise by first establishing the following lemmas:

Lemma (a). There exists \( M > 0 \) so that \( |f| \leq M \) except on a set of measure less than \( \varepsilon \).

Proof:

Denote \( A_k = f^{-1}(B(0,k)^c) \) for all \( k \in \mathbb{N} \), where \( B(0,k) \) is the open \( k \)-radius ball centered at 0 and the complement is taken in \( \mathbb{R} \cup \{\pm \infty\} \). The sets \( A_k \) are measurable since \( f \) is measurable.
Now $A_{k+1} \subseteq A_k$ for all $k \in \mathbb{N}$ and $m(A_k) \leq m([a,b]) < \infty$ for all $k \in \mathbb{N}$, so

$$\lim_{k \to \infty} m(A_k) = m\left(\bigcap_{k=1}^{\infty} A_k\right) = m\left(\bigcap_{k=1}^{\infty} f^{-1}(B(0,k)^c)\right)$$

$$= m\left(f^{-1}\left(\bigcap_{k=1}^{\infty} B(0,k)^c\right)\right) = m(f^{-1}\{\pm \infty\}) = 0.$$ 

Thus there exists $M \in \mathbb{N}$ so that

$$m(f^{-1}(B(0,M)^c)) = m(A_M) < \varepsilon.$$ 

In other words, $|f(x)| \leq M$ for all $x \in A_M$ and $m(A_M) < \varepsilon$. Hence $|f| \leq M$ except on a set of measure $\varepsilon$.

Lemma (b). Given $M > 0$ there exists a simple function $\varphi$ such that

$$|f(x) - \varphi(x)| < \varepsilon$$

except where $|f(x)| > M$. If $m \leq f \leq M$, then we may take $\varphi$ so that $m \leq \varphi \leq M$.

Proof:

Let $M > 0$ be given. We partition $\overline{B(0,M)}$ to equal length disjoint intervals $\{I_k\}_{k=1}^{n}$ so that $\ell(I_k) < \varepsilon$ for all $k$. For each $k$ denote $A_k = f^{-1}(I_k)$, which are measurable since $f$ is. In addition, $\{A_k\}_{k=1}^{n}$ is a disjoint partition of the set $B := \{x : |f(x)| \leq M\}$. For each $k$ choose $a_k \in I_k$ and define $\varphi : [a,b] \to \mathbb{R}$ so that

$$\varphi(x) = \sum_{k=1}^{n} a_k \chi_{A_k}(x)$$

for all $x \in [a,b]$, making $\varphi$ a simple function. Let $x$ be such that $|f(x)| \leq M$. Then $x \in A_k$ and $f(x) \in I_k$ for some unique $k$, and thus

$$|f(x) - \varphi(x)| = |f(x) - \sum_{k=1}^{n} a_k \chi_{A_k}(x)| = |f(x) - a_k| \leq \ell(I_k) < \varepsilon.$$ 

So $\varphi$ is the desired simple function. If $m \leq f \leq M$ then we can repeat the above argument by taking a partition of $[m, M]$ instead
Lemma (c). Let \( \varphi : [a, b] \to \mathbb{R} \) be a simple function. There then exists a step function \( g \) on \([a, b]\) such that \( g(x) = \varphi(x) \) except on a set of measure less than \( \varepsilon \). If \( m \leq \varphi \leq M \) then we can take \( g \) so that \( m \leq g \leq M \).

Proof:
We first prove the statement for a characteristic function \( \varphi = \chi_A \), with a measurable \( A \subseteq [a, b] \). Since \( A \) is measurable then by Exercise 1 we find an open set \( O \supseteq A \) so that \( m(O \setminus A) < \frac{\varepsilon}{2} \).
Since every open set in \( \mathbb{R} \) can be obtained as a countable union of disjoint open intervals, then there exists a disjoint countable collection \( \{I_n\}_n \) of open intervals so that \( O = \bigcup_n I_n \). Now setting \( L_k = \bigcup_{n \geq k} I_n \) we have \( L_{k+1} \subseteq L_k \) for all \( k \) so

\[
\lim_{k \to \infty} m(L_k) = m\left(\bigcap_{k=1}^{\infty} \bigcup_{n \geq k} I_k \right) = m(\emptyset) = 0
\]

since the sets \( I_k \) are disjoint. There thus exists \( n_0 \in \mathbb{N} \) so that

\[
m\left(\bigcup_{n \geq n_0+1} I_n \right) = m(L_{n_0}) < \frac{\varepsilon}{2}.
\]

We then set

\[
g(x) = \sum_{n=1}^{n_0} \chi_{I_n}(x)
\]

for all \( x \in [a, b] \). Note that if \( x \in A \cap \left(\bigcup_{n=1}^{n_0} I_n \right) \) then \( x \in A \) and \( x \in I_k \) for some \( k \leq n_0 \). Thus \( \varphi(x) = 1 = g(x) \). If \( x \notin O \) then \( \varphi(x) = 0 = g(x) \). So it suffices to show that the set

\[
O \setminus \left(A \cap \left(\bigcup_{n=1}^{n_0} I_n \right) \right) = (O \setminus A) \cup \left( O \setminus \left( \bigcup_{n=1}^{n_0} I_n \right) \right)
\]

\[
= (O \setminus A) \cup \left( \bigcup_{n \geq n_0+1} I_n \right)
\]
has measure less than \( \varepsilon \). But this is true because

\[
m((O \setminus A) \cup \left( \bigcup_{n \geq n_0 + 1} I_n \right)) \leq m(O \setminus A) + m\left( \bigcup_{n \geq n_0 + 1} I_n \right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\]

Hence we have shown that there exists a step function \( g \) so that \( g(x) = \varphi(x) \) except on a set of measure less than \( \varepsilon \).

If we then assume that \( \varphi \) is any simple function

\[
\varphi(x) = \sum_{n=1}^{k} a_n \chi_{A_n}(x)
\]

where the sets \( A_n \) are disjoint and measurable, we can apply the earlier part to produce step functions \( g_n \) for each \( n \leq k \) for which \( a_n \chi_{A_n}(x) = g_n(x) \) except on a set \( E_n \) with \( m(E_n) < \frac{\varepsilon}{k} \). Define a step function \( g : [a, b] \to \mathbb{R} \) so that

\[
g(x) = \sum_{n=1}^{k} g_n(x)
\]

for all \( x \in [a, b] \). Now for all \( x \in \left( \bigcup_{n=1}^{k} E_n \right)^c \) we have \( x \in E_n^c \) for all \( n \leq k \), i.e. \( x \notin E_n \) for all \( n \leq k \). Thus \( a_n \chi_{A_n}(x) = g_n(x) \) for all \( n \leq k \) so

\[
\varphi(x) = \sum_{n=1}^{k} a_n \chi_{A_n}(x) = \sum_{n=1}^{k} g_n(x) = g(x).
\]

Also we have

\[
m\left( \bigcup_{n=1}^{k} E_n \right) \leq \sum_{n=1}^{k} m(E_n) < \sum_{n=1}^{k} \frac{\varepsilon}{k} = \varepsilon.
\]

So we have shown that there exists a step function \( g \) so that \( \varphi(x) = g(x) \) except on a set with measure less than \( \varepsilon \).

Now lastly if we assume that \( m \leq \varphi \leq M \) then by repeating the above argument with all the intervals \( I_n \) intersected with \([m, M]\) we produce \( g \) so that \( m \leq g \leq M \).
Lemma (d). Let $g : [a, b] \to \mathbb{R}$ be a step function. Then there is a continuous function $h$ such that $g(x) = h(x)$ except on a set of measure less than $\varepsilon$. If $m \leq g \leq M$ then we may take $h$ so that $m \leq h \leq M$.

Proof:
Since $g$ is a step function there exists a disjoint partition of $[a, b]$ to intervals $\{I_k\}_{k=1}^n$ and constants $(a_k)_{k=1}^n \subseteq \mathbb{R}$ so that

$$g(x) = \sum_{k=1}^n a_k \chi_{I_k}(x)$$

for all $x \in [a, b]$. Denote the endpoints of each interval $I_k$ by $c_k$ and $d_k$ respectively as the left and right endpoint. Choose

$$\rho = \min_{k \leq n} \left( \frac{d_k - c_k}{2} \right) > 0$$

and $\delta = \min\{\frac{\varepsilon}{n}, \rho\}$. Now for each $k$ we consider the intervals

$$J_k = (c_k + \frac{\delta}{2n}, d_k - \frac{\delta}{2n}).$$

Note that the intervals $J_k$ are disjoint and $J_k \subseteq I_k$ for all $k$. We then define $h$ so that $h(x) = a_k$ when $x \in J_k$ for every $k \leq n$, and between all the intervals $J_i$ and $J_{i+1}$ we linearly interpolate from the right endpoint of $J_i$ to the left endpoint of $J_{i+1}$. We thus get a continuous function $h$ on $[a, b]$. Note that the interval between each $J_i$ and $J_{i+1}$ has measure $\frac{\varepsilon}{n} \leq \frac{\varepsilon}{2n} < \frac{\varepsilon}{n}$. Now if $x \in \bigcup_{k=1}^n J_k$ then $x \in J_k \subseteq I_k$ for some $k \leq n$ and thus $h(x) = a_k = g(x)$. Hence $h(x) = g(x)$ for all $x \in \bigcup_{k=1}^n J_k$. Note that the set $\left( \bigcup_{k=1}^n J_k \right)^c$ is a union of $n$ intervals each length less than $\frac{\varepsilon}{n}$. Thus by subadditivity

$$m\left( \left( \bigcup_{k=1}^n J_k \right)^c \right) < \sum_{k=1}^n \frac{\varepsilon}{n} = \varepsilon.$$ 

We have shown that there exists a continuous function $h$ so that $h(x) = g(x)$ except on a set of measure less than $\varepsilon$.

Assume then finally that $m \leq g \leq M$. Since $h$ was produced by linearly interpolating between subintervals in the definition of $g$, then

$$m \leq \min g \leq \min h \leq h \leq \max h \leq \max g \leq M.$$
Thus \( m \leq h \leq M \).

We can now solve the exercise. By Lemma (a) there exists \( M > 0 \) so that \( |f| \leq M \) except on a set \( E_1 \) with \( m(E_1) < \frac{\varepsilon}{3} \). By Lemma (b) we get a simple function \( \varphi \) so that \( |f - \varphi| < \varepsilon \) except on the set \( E_1 \). By Lemma (c) we find a step function \( g \) so that \( \varphi = g \) except on a set \( E_2 \) with \( m(E_2) < \frac{\varepsilon}{3} \). Now finally by Lemma (d) we find a continuous function \( h \) so that \( g = h \) except on a set \( E_3 \) with \( m(E_3) < \frac{\varepsilon}{3} \). Now putting all together we have

\[
|f(x) - h(x)| = |f(x) - g(x)| = |f(x) - \varphi(x)| < \varepsilon
\]

except for \( x \in E_1 \cup E_2 \cup E_3 \), and

\[
m(E_1 \cup E_2 \cup E_3) \leq m(E_1) + m(E_2) + m(E_3) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\]

Hence there exists a continuous function \( h \) and a step function \( g \) so that \( |f - h| = |f - g| < \varepsilon \) except on a set of measure less than \( \varepsilon \). In addition, if \( m \leq f \leq M \) then each of the Lemmas imply that \( m \leq g \leq M \) and \( m \leq h \leq M \).

**Exercise 5.** A function \( f \) is said to be Borel measurable if for each \( \alpha \) the set \( \{ x : f(x) > \alpha \} \) is a Borel set. Show that each of the statements of Lecture 3 remain valid if we replace "Lebesgue measurable" by "Borel measurable". In other words, show that the following statements are equivalent

(i) For each real number \( \alpha \) the set \( \{ x : f(x) > \alpha \} \) is a Borel set.
(ii) For each real number \( \alpha \) the set \( \{ x : f(x) \geq \alpha \} \) is a Borel set.
(iii) For each real number \( \alpha \) the set \( \{ x : f(x) < \alpha \} \) is a Borel set.
(iv) For each real number \( \alpha \) the set \( \{ x : f(x) \leq \alpha \} \) is a Borel set.

Also show that

(a) Every Borel measurable function is Lebesgue measurable.
(b) If \( f \) is a Borel measurable and \( B \) is a Borel set then \( f^{-1}(B) \) is a Borel set.
(c) If \( f \) and \( g \) are Borel measurable then so is \( f \circ g \).
(d) If $f$ is Borel measurable and $g$ is Lebesgue measurable then $f \circ g$ is Lebesgue measurable: However, if $f$ is Lebesgue measurable and $g$ is Borel measurable, then $f \circ g$ may not be Lebesgue measurable.

(e) If $f$ is a continuous function defined on a Borel set then $f$ is Borel measurable.

**Answer:**

The statements (i) to (iv) are equivalent with identical arguments as with Lebesgue measurability, as the set of Borel sets is a $\sigma$-algebra. We then verify the statements (a) to (e). For simplicity we will denote the usual topology on $\mathbb{R}$ by $\tau$ and the collection of Borel sets, i.e. the $\sigma$-algebra generated by open sets, by $\sigma(\tau)$.

(a) Take a Borel measurable function $f$. Fix $\alpha$ and consider the set $A = \{x : f(x) > \alpha\}$. Since $f$ is Borel measurable then $A$ is a Borel set. Since Borel sets are all Lebesgue measurable then $A$ is Lebesgue measurable. Since $\alpha$ was arbitrary, then $f$ is Lebesgue measurable.

(b) Let $f$ be Borel measurable and $B$ a Borel set. We want to show that $f^{-1}(B)$ is a Borel set. Denote $\mathcal{D} = \{A \in \sigma(\tau) : f^{-1}(A) \in \sigma(\tau)\}$.

Now $\mathcal{D} \subseteq \sigma(\tau)$ by definition. It suffices to show that $\mathcal{D}$ is a $\sigma$-algebra that contains all open sets, as this would imply the converse inclusion $\mathcal{D} \supseteq \sigma(\tau)$ by the definition of $\sigma(\tau)$ because $\sigma(\tau)$ is the smallest $\sigma$-algebra that contains all open sets. This will imply that the preimage of every Borel set under $f$ is a Borel set.

So we will first verify that $\mathcal{D}$ is a $\sigma$-algebra:

1. Since $\emptyset \in \sigma(\tau)$ and $f^{-1}(\emptyset) = \emptyset$ then $\emptyset \in \mathcal{D}$.
2. Let $E \in \mathcal{D}$. Thus $E \in \sigma(\tau)$ and $f^{-1}(E) \in \sigma(\tau)$. Since $\sigma(\tau)$ is a $\sigma$-algebra then $E^c \in \sigma(\tau)$ and $f^{-1}(E^c) = (f^{-1}(E))^c \in \sigma(\tau)$. Hence $E^c \in \mathcal{D}$.
3. Assume that $(E_n)_{n=1}^{\infty} \subseteq \mathcal{D}$. Now $(E_n)_{n=1}^{\infty} \subseteq \sigma(\tau)$ and $(f^{-1}(E_n))_{n=1}^{\infty} \subseteq \sigma(\tau)$. Since $\sigma(\tau)$ is a $\sigma$-algebra then $\bigcup_{n=1}^{\infty} E_n \in \sigma(\tau)$ and

$$f^{-1}\left(\bigcup_{n=1}^{\infty} E_n\right) = \bigcup_{n=1}^{\infty} f^{-1}(E_n) \in \sigma(\tau).$$

Hence $\bigcup_{n=1}^{\infty} E_n \in \mathcal{D}$. 

15
By (1.), (2.) and (3.) the collection $\mathcal{D}$ is a $\sigma$-algebra. To show that $\mathcal{D}$ contains all open sets, let $I = (a, b)$ be an open interval. Now $I$ is a Borel set and since $f$ is a Borel function then
\[
f^{-1}(a, b) = f^{-1}((\infty, b) \cap (a, \infty)) = f^{-1}(\infty, b) \cap f^{-1}(a, \infty) = \{x : f(x) < b\} \cap \{x : f(x) > a\},
\]
which is a Borel set as an intersection of two Borel sets. Here we assumed the earlier part of this exercise where the conditions (i) to (iv) were verified to be equivalent. Hence $f^{-1}(a, b)$ is a Borel set, so we have shown that $\mathcal{D}$ contains all open intervals. Any open set $O$ can be written as a countable union of open intervals (open intervals with rational endpoints constitute a countable basis for the usual topology $\tau$) so it follows that if $O$ is open then $O = \bigcup_n I_n$ where each $I_n$ is an open interval. Now $O \in \sigma(\tau)$ as open sets are Borel sets and
\[
f^{-1}(O) = f^{-1}\left(\bigcup_n I_n\right) = \bigcup_n f^{-1}(I_n) \in \sigma(\tau)
\]
since $\sigma(\tau)$ is a $\sigma$-algebra. Hence $O \in \mathcal{D}$. So we have shown that $\mathcal{D}$ is a $\sigma$-algebra that contains all open sets, and thus $\mathcal{D} = \sigma(\tau)$. Thus $f^{-1}(B)$ is a Borel set.

(c) Let $f$ and $g$ be Borel measurable and fix $\alpha \in \mathbb{R}$. Since $f$ is Borel measurable then by condition (i) we have that $f^{-1}(\alpha, \infty)$ is a Borel set. By part (b) it follows that $g^{-1}(f^{-1}(\alpha, \infty))$ is a Borel set. In other words, the set $\{x : (f \circ g)(x) > \alpha\}$ is a Borel set. Hence by (i) the function $f \circ g$ is Borel measurable.

(d) Let $f$ be a Borel measurable and $g$ a Lebesgue measurable. We show that $f \circ g$ is Lebesgue measurable. Note that since the collection of Borel sets is included (with strict inclusion) in the Lebesgue measurable sets then by similar reasoning as in part (b) we conclude that $g^{-1}(B)$ is Lebesgue measurable for every Borel set $B$. Now note that since $f$ is Borel measurable then by (i) we have that $f^{-1}(\alpha, \infty)$ is a Borel set. It thus follows that $g^{-1}(f^{-1}(\alpha, \infty))$ is Lebesgue measurable. In other words, the set $\{x : (f \circ g)(x) > \alpha\}$ is Lebesgue measurable. Thus $f \circ g$ is Lebesgue measurable.

Counter example to this statement was done in the lecture by using the Cantor ternary function $f : [0, 1] \to [0, 2]$ and using
Exercise 3 (b) to take a non-measurable subset of the cantor set under $f$ (which must have positive outer measure).

(d) Note that as in part (b) we noted that the conditions (i) and (iii) imply that for a Borel measurable function the preimage of every open set is a Borel set. And if on the converse we assume that the preimage of every open set is a Borel set then the condition (i) holds. Hence a function is Borel measurable if and only if the preimage of every open set is a Borel set. Thus if $f : B \to \mathbb{R}$ is a continuous function defined on a Borel set $B$, then for every open set $O \subseteq \mathbb{R}$ we have that $f^{-1}(O)$ is an open set in the subspace topology of $B$. In other words, $f^{-1}(O)$ is the intersection of an open set in $\mathbb{R}$ and the Borel set $B$ and is thus Borel. Hence $f$ is a Borel measurable function.