Structure of Measurable Sets

In these notes we discuss the structure of Lebesgue measurable subsets of $\mathbb{R}$ from several different points of view. Along the way, we will see several alternative characterizations of measurability which might help to make the concept seem more intuitive.

We begin by discussing the measures of open sets. First, recall the following definition.

Definition: Open Set
A subset $U \subseteq \mathbb{R}$ is said to be **open** if there exists a collection $\mathcal{C}$ of open intervals whose union is U.

Note that $\mathbb{R}$ is itself open, being the union of the intervals $(n-1, n+1)$ for $n \in \mathbb{Z}$. The empty set $\emptyset$ is also open, being the union of the empty collection of intervals.

The following proposition highlights the important role that open sets play in analysis.

Proposition 1 Continuity Using Open Sets

Let $f\colon \mathbb{R} \to \mathbb{R}$. Then f is continuous if and only if $f^{-1}(U)$ is open for every open set $U \subseteq \mathbb{R}$.

PROOF Suppose first that $f^{-1}(U)$ is open for every open set $U \subseteq \mathbb{R}$. Let $x \in \mathbb{R}$, and let $\epsilon > 0$. Then $U = \big(f(x) - \epsilon, f(x) + \epsilon\big)$ is open, so $f^{-1}(U)$ is open. Since $x \in f^{-1}(U)$, there must be an open interval that contains x and is contained in $f^{-1}(U)$. Indeed, there must exist a $\delta > 0$ so that $(x - \delta, x + \delta) \subseteq f^{-1}(U)$. Then

$$|y - x| < \delta \quad \Rightarrow \quad \big|f(x) - f(y)\big| < \epsilon$$

for all $y \in \mathbb{R}$, which proves that f is continuous at x.

For the converse, suppose that f is continuous. Let $U \subseteq \mathbb{R}$ be any open set, and let $x \in f^{-1}(U)$. Since U is open, there is an open interval containing $f(x)$ that is contained in U. In particular, there exists an $\epsilon > 0$ so that $\big(f(x) - \epsilon, f(x) + \epsilon\big) \subseteq U$. But since f is continuous at x, there exists a $\delta > 0$ so that

$$|x - y| < \delta \qquad \Rightarrow \qquad \big|f(x) - f(y)\big| < \epsilon$$

for all $y \in \mathbb{R}$. It follows that $f(y) \in \big(f(x) - \epsilon, f(x) + \epsilon\big)$ for all $y \in (x - \delta, x + \delta)$. Then $(x - \delta, x + \delta) \subseteq f^{-1}(U)$, which proves that $f^{-1}(U)$ is open. ■

A priori, there is no reason to think that every open set must be measurable, since by the definition an open set might involve an uncountable union of open intervals. However, the following structure theorem shows that every open set is a countable union of open intervals.

Theorem 2 Structure of Open Sets

Every proper open subset of $\mathbb{R}$ is a countable, disjoint union of open intervals and open rays.

PROOF Let U be a proper open subset of $\mathbb{R}$. Put an equivalence relation $\sim$ on U by $x \sim y$ if U contains every point between x and y. The equivalence classes under this relation are called the **components** of U. We claim that each component of U is either an open interval or an open ray.

Let C be a component of U, and let $a = \inf(U)$ and $b = \sup(U)$, with $a = -\infty$ if U has no lower bound and $b = \infty$ if U has no upper bound. If $x \in (a, b)$, then there must exist points $c_1, c_2 \in C$ so that $a < c_1 < x$ and $x < c_2 < b$. Since $c_1 \sim c_2$ and x lies between c_1 and c_2, it follows that $x \in C$. This proves that $(a, b) \subseteq C$, and we know that $C \subseteq [a, b]$. But if $b \in C$, then since $b \in U$ and U is open there exists an open interval $(d, e) \subseteq U$ that contains b. Then it is easy to see that all of $(a, b) \cup (d, e)$ must lie in C, a contradiction since $e > b$. Thus $b \notin C$, and similar reasoning shows that $a \notin C$, so $C = (a, b)$. Clearly $C \neq \mathbb{R}$ since U is a proper subset of $\mathbb{R}$, so C is either an open interval or an open ray.

Finally, observe that each component of U must contain at least one rational number, and therefore U has only countably many components. Thus U is the countable, disjoint union of its components. ■

Corollary 3

Every open subset of $\mathbb{R}$ is Lebesgue measurable.

Based on the structure of open sets described in Theorem 2, the measure $m(U)$ of an open set U can be interpreted as simply the sum of the lengths of the components of U. Note, however, that an open set may have infinitely many components, and these may form a fairly complicated structure on the real line. Indeed, the following example illustrates that open sets can behave in very counterintuitive ways.

Proposition 4 Small Open Sets Containing $\mathbb{Q}$

For every $\epsilon > 0$, there exists an open set $U \subseteq \mathbb{R}$ such that $m(U) \leq \epsilon$ and U contains the set $\mathbb{Q}$ of rational numbers.

PROOF Let $\epsilon > 0$, let $q_1, q_2, \ldots$ be an enumeration of the rational numbers, and let

$$U = \bigcup_{n \in \mathbb{N}} \left(q_n - \frac{\epsilon}{2^{n+1}}, q_n + \frac{\epsilon}{2^{n+1}} \right).$$

Then U is open, and

$$m(U) \leq \sum_{n \in \mathbb{N}} m\left(\left(q_n - \frac{\epsilon}{2^{n+1}}, q_n + \frac{\epsilon}{2^{n+1}} \right) \right) = \sum_{n \in \mathbb{N}} \frac{\epsilon}{2^n} = \epsilon. \qquad \blacksquare$$

Closed Sets

Recall the following definition.

Definition: Closed Set
A subset $F \subseteq \mathbb{R}$ is **closed** if, for every convergent sequence $\{x_n\}$ of points in F, the point $x \in \mathbb{R}$ that $\{x_n\}$ converges to also lies in F.

That is, a closed set is a set that it closed under the operation of taking limits of sequences. For example, any closed interval $[a, b]$ is closed, since any convergent sequence in $[a, b]$ must converge to a point in $[a, b]$. The entire real line $\mathbb{R}$ is also closed, and technically the empty set $\emptyset$ is closed as well, since the condition is vacuously satisfied.

The following description of closed sets is fundamental.

Proposition 5 Complements of Closed Sets

Let $F \subseteq \mathbb{R}$. Then F is closed if and only if F^c is open.

PROOF Suppose first that F is not closed. Then there exists a sequence $\{x_n\}$ of points in F that converges to a point $x \in F^c$. Then every open interval (a, b) that contains x must contain a point of the sequence, and therefore no open interval (a, b) that contains x is contained in F^c. It follows that F^c is not open.

Conversely, suppose that F^c is not open. Then there must exist a point $x \in F^c$ that does not lie an any open interval contained in F^c. In particular, each of the open intervals $(x - 1/n, x + 1/n)$ must contain a point $x_n \in F$. Then the sequence $\{x_n\}_{n \in \mathbb{N}}$ in F converges to the point $x \in F^c$, so F is not closed. ■

Corollary 6

Every closed subset of $\mathbb{R}$ is Lebesgue measurable.

We now turn to unions and intersections of open and closed sets. Students of point-set topology will recognize parts (1) and (2) of the following proposition as essentially the definition of a topological space.

Proposition 7 Unions and Intersections of Open and Closed Sets

1. *The union of any collection of open sets in $\mathbb{R}$ is open.*
2. *The intersection of finitely many open sets in $\mathbb{R}$ is open.*
3. *The intersection of any collection of closed sets in $\mathbb{R}$ is closed.*
4. *The union of finitely many closed sets in $\mathbb{R}$ is closed.*

PROOF For (1), let $\mathcal{C}$ be a collection of open sets. For each $U \in \mathcal{C}$, let $\mathcal{I}_U$ be a collection of open intervals whose union is U. Then $\mathcal{I} = \bigcup_{U \in \mathcal{C}} \mathcal{I}_U$ is a collection of open intervals whose union is $\bigcup \mathcal{C}$, and therefore $\bigcup \mathcal{C}$ is open.

For (2), it suffices to prove that $U \cap V$ is open if U and V are open. Given such a U and V, let $\mathcal{I}$ and $\mathcal{J}$ be collections of open intervals whose unions are U and V respectively. Then

$$\{I \cap J \mid I \in \mathcal{I},\ J \in \mathcal{J}, \text{ and } I \cap J \neq \emptyset\}$$

is a collection of open intervals whose union is $U \cap V$, and therefore $U \cap V$ is open.

Parts (3) and (4) follow immediately from parts (1) and (2) by taking complements. ■

Though every open set in $\mathbb{R}$ is a disjoint union of countably many open intervals, it is not true that every closed set is a disjoint union of closed intervals. Indeed, there exists a very famous closed set called the **Cantor set** whose structure is much more interesting. To construct the Cantor set, we start with the unit interval:

$$C_0 = [0,1].$$

Next we remove the middle third of this interval, leaving a union of two closed intervals:

$$C_1 = \left[0, \frac{1}{3}\right] \uplus \left[\frac{2}{3}, 1\right].$$

Next we remove the middle third of each of these intervals, leaving a union of four closed intervals:

$$C_2 = \left[0, \frac{1}{9}\right] \uplus \left[\frac{2}{9}, \frac{1}{3}\right] \uplus \left[\frac{2}{3}, \frac{7}{9}\right] \uplus \left[\frac{8}{9}, 1\right].$$

Proceeding in this fashion, we obtain a nested sequence $C_0 \supseteq C_1 \supseteq C_2 \supseteq \cdots$ of closed sets, where C_n is the union of 2^n closed intervals, as illustrated in Figure 1. Then the intersection

$$C = \bigcap_{n \in \mathbb{N}} C_n$$

is the aforementioned Cantor set.

Figure 1: The first six stages in the construction of the Cantor set.

Proposition 8 Properties of the Cantor Set

The Cantor set C has the following properties:

1. *C is closed.*
2. *C is uncountable. Indeed, $|C| = |\mathbb{R}|$.*
3. *$m(C) = 0$.*

PROOF Since C is the intersection of the closed sets C_n, it follows from Proposition 7 that C is closed. Furthermore, since C_n is the disjoint union of 2^n closed intervals of length 3^{-n}, we know that $m(C_n) = 2^n 3^{-n} = (2/3)^n$, and it follows that

$$m(C) \;=\; \inf_{n\in\mathbb{N}} m(C_n) \;=\; \inf_{n\in\mathbb{N}} (2/3)^n \;=\; 0.$$

Finally, to show that C is uncountable, let $\{0,1\}^\infty$ be the (uncountable) set of all infinite binary sequences, and define a function $f\colon \{0,1\}^\infty \to \mathbb{R}$ by

$$f(b_1, b_2, b_3, \ldots) \;=\; \sum_{n\in\mathbb{N}} \frac{2b_n}{3^n}.$$

It is easy to check that f is injective and the image of f lies in the Cantor set (in fact, f is a bijection from $\{0,1\}^\infty$ to C), and therefore C is uncountable. Indeed, we have

$$|\mathbb{R}| \;=\; \bigl|\{0,1\}^\infty\bigr| \;\leq\; |C| \;\leq\; |\mathbb{R}|$$

and hence $|C| = |\mathbb{R}|$. ■

Corollary 9

Let $\mathcal{M}$ be the collection of all Lebesgue measurable subsets of $\mathbb{R}$. Then

$$|\mathcal{M}| \;=\; \bigl|\mathcal{P}(\mathbb{R})\bigr|.$$

PROOF Clearly $|\mathcal{M}| \leq \bigl|\mathcal{P}(\mathbb{R})\bigr|$. But since the Cantor set C has Lebesgue measure zero, every subset of the Cantor set is Lebesgue measurable, i.e. $\mathcal{P}(C) \subseteq \mathcal{M}$. But since $|C| = |\mathbb{R}|$, it follows that $\bigl|\mathcal{P}(C)\bigr| = \bigl|\mathcal{P}(\mathbb{R})\bigr|$, and hence $\bigl|\mathcal{P}(\mathbb{R})\bigr| \leq |\mathcal{M}|$. ■

Incidentally, there is some sense in which the structure of the Cantor set is fairly typical for closed sets. In particular, Theorem 2 tells us that any open set can be described as the union of a nested sequence

$$U_1 \subseteq U_2 \subseteq U_3 \subseteq \cdots$$

where each U_n is a finite disjoint union of open intervals and open rays. Taking complements, we find that any closed set can be described as the intersection of a nested sequence

$$F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$$

where each F_n is a finite disjoint union of closed intervals and closed rays.

Open Sets and Measurability

We are now ready to use open sets and closed sets to give a few alternative descriptions of Lebesgue outer measure and Lebesgue measurability. We begin by describing the Lebesgue outer measure in terms of open sets.

Proposition 10 Open Sets and Outer Measure

If $S \subseteq \mathbb{R}$, then

$$m^*(S) = \inf\{m(U) \mid U \text{ is open and } S \subseteq U\}.$$

PROOF Let x be the value of the infimum. Clearly $m^*(S) \leq m(U)$ for every open set U that contains S, and therefore $m^*(S) \leq x$. For the opposite inequality, let $\epsilon > 0$, and let $\mathcal{C}$ be a cover of S by open intervals so that

$$\sum_{I \in \mathcal{C}} \ell(I) \leq m^*(S) + \epsilon.$$

Then $U = \bigcup \mathcal{C}$ is an open set that contains S, so

$$x \leq m(U) \leq \sum_{I \in \mathcal{C}} m(I) = \sum_{I \in \mathcal{C}} \ell(I) \leq m^*(S) + \epsilon.$$

Since ϵ was arbitrary, it follows that $x \leq m^*(S)$. ■

We shall now use open sets to give a nice characterization of measurability.

Proposition 11 Measurability Using Open Sets

Let $S \subseteq \mathbb{R}$. Then S is Lebesgue measurable if and only if for every $\epsilon > 0$ there exists an open set U containing S so that $m^*(U - S) < \epsilon$.

PROOF Suppose first that S is measurable, and let $\epsilon > 0$. For each $n \in \mathbb{N}$, let $S_n = S \cap [-n, n]$, and let U_n be an open set containing S_n so that

$$m(U_n) \; < \; m(S_n) + \frac{\epsilon}{2^n}.$$

Let $U = \bigcup_{n\in\mathbb{N}} U_n$. Then U is an open set containing S and $U - S \subseteq \bigcup_{n\in\mathbb{N}}(U_n - S_n)$, so

$$m(U - S) \; \leq \; \sum_{n\in\mathbb{N}} m(U_n - S_n) \; \leq \; \sum_{n\in\mathbb{N}} \frac{\epsilon}{2^n} \; = \; \epsilon.$$

For the converse, let $S \subseteq \mathbb{R}$, and suppose that for every $n \in \mathbb{N}$ there exists an open set U_n containing S so that $m^*(U_n - S) < 1/n$. Let $E = \bigcap_{n\in\mathbb{N}} U_n$, and note that E is a measurable set containing S. But $E - S \subseteq U_n - S$ for each n, so

$$m^*(E - S) \; \leq \; m^*(U_n - S) \; \leq \; \frac{1}{n}$$

for each n. We conclude that $m^*(E - S) = 0$, and therefore $E - S$ is Lebesgue measurable. Then $S = E - (E - S)$ is Lebesgue measurable as well. ■

Corollary 12 Measurability Using Closed Sets

Let $S \subseteq \mathbb{R}$. Then S is Lebesgue measurable if and only if for every $\epsilon > 0$ there exists a closed set $F \subseteq S$ containing S so that $m^*(S - F) < \epsilon$.

PROOF Observe that F is a closed set contained in S if and only if $U = F^c$ is an open set containing S^c. Moreover, $S - F = U - S^c$, so $m^*(S - F) < \epsilon$ if and only if $m^*(U - S^c) < \epsilon$, and hence this statement follows directly from applying the previous proposition to S^c. ■

Combining this corollary with the previous proposition yields the following nice result.

Corollary 13

Let $S \subseteq \mathbb{R}$. Then S is Lebesgue measurable if and only if there exists a closed set F and an open set U so that $F \subseteq S \subseteq U$ and $m(U - F) < \epsilon$.

Incidentally, there is another function similar to Lebesgue outer measure that is more closely related to closed sets.

Definition: Inner Measure
If $S \subseteq \mathbb{R}$, the **Lebesgue inner measure** of S is defined by

$$m_*(S) \;=\; \sup\{m(F) \mid F \text{ is closed and } F \subseteq S\}.$$

It follows immediately from Corollary 12 that $m_*(E) = m(E)$ for any measurable set E. It is also apparent that $m_*(S) \leq m^*(S)$ for any set $S \in \mathbb{R}$. The following proposition gives a nice characterization of measurability for sets of finite measure.

Proposition 14 Measurability Using Inner and Outer Measures

Let $S \subseteq \mathbb{R}$, and suppose that $m^*(S) < \infty$. Then S is measurable if and only if

$$m_*(S) \;=\; m^*(S).$$

PROOF If S is measurable, then $m_*(S) = m(S) = m^*(S)$. Conversely, suppose that $m^*(S) < \infty$ and $m_*(S) = m^*(S)$. Let $\epsilon > 0$, and let $F \subseteq S$ be a closed set and $U \subseteq \mathbb{R}$ and open set containing S so that

$$m_*(S) \;\leq\; m(F) + \frac{\epsilon}{2} \qquad \text{and} \qquad m(U) \;\leq\; m^*(S) + \frac{\epsilon}{2}.$$

Then

$$m(U - F) \;=\; m(U) - m(F) \;\leq\; \left(m^*(S) + \frac{\epsilon}{2}\right) - \left(m_*(S) - \frac{\epsilon}{2}\right) \;=\; \epsilon.$$

Since ϵ was arbitrary, it follows from Corollary 13 that S is measurable. ■

F_σ and G_δ Sets

As we have seen, every open or closed subset of $\mathbb{R}$ is Lebesgue measurable. The following definition provides many more examples of measurable sets.

Definition: F_σ and G_δ Sets

1. A subset of $\mathbb{R}$ is $\boldsymbol{F_\sigma}$ if it is a countable union of closed sets.
2. A subset of $\mathbb{R}$ is $\boldsymbol{G_\delta}$ if it is a countable intersection of open sets.

Here F stands for *fermé*, which is French for "closed", and σ stands for *somme*, which is the French word for a union of sets. Similarly, G stands for *Gebiet*, which is the German word for an open set, and δ stands for *Durchschnitt*, which is the German word for an intersection of sets.

The following proposition lists some of the basic properties of F_σ or G_δ sets. The proofs are left to the exercises.

Proposition 15 Properties of F_σ and G_δ Sets

1. *An F_σ or G_δ set is measurable.*
2. *If $S \subseteq \mathbb{R}$, then S has type F_σ if and only if S^c has type G_δ.*
3. *Every countable set is F_σ. In particular, the set $\mathbb{Q}$ of rational numbers is F_σ, and the set $\mathbb{R} - \mathbb{Q}$ of irrational numbers is G_δ.*
4. *Every open or closed set is both F_σ and G_δ.*

The rational numbers provide an example of an F_σ set that is neither open nor closed. Incidentally, it is possible to prove that the rational numbers are *not* a G_δ set using the Baire category theorem (see §48 of Munkres' *Topology*).

Of course, most F_σ sets are not countable. The following example describes an uncountable F_σ set that is neither open nor closed, whose structure is more "typical" for sets of this type.

EXAMPLE 1 Let $C \subseteq [0,1]$ be the Cantor set. Note that any closed interval $[a,b]$ contains a scaled copy of C whose left endpoint is a and whose right endpoint is B. We now define a sequence $F_0 \subseteq F_1 \subseteq F_2 \subseteq \cdots$ of sets as follows:

- We start with $F_0 = C$.

- Let F_1 be the set obtained from C by pasting a scaled copy of C into each interval of $[0,1] - C$.
- For each $n \geq 2$, let F_n be the set obtained from F_{n-1} by pasting a scaled copy of C into each interval of $[0,1] - F_{n-1}$.

Note that each F_n is a closed set, since the complement F_n^c is a union of open intervals. The union $F = \bigcup_{n\in\mathbb{N}} F_n$ is thus an F_σ set. It is not hard to see that neither F nor F^c contains any open intervals, so F is neither open nor closed. Note also that each F_n has measure zero, and therefore F has measure zero.

This set F has a nice description in ternary (base 3). First, observe that the Cantor set C consists of all points $x \in [0,1]$ that have a ternary expansion consisting only of 0's and 2's, with no 1's. Then for each n, the set F_n consists of all points $x \in [0,1]$ that have a ternary expansion with at most n digits that are 1. For example, the number

$$\frac{475}{972} = 0.1110120202020\cdots$$

lies in F_4 but not F_3. Then the F_σ set F consists of all points $x \in [0,1]$ that have a ternary expansion with at most finitely many 1's. ■

We can reinterpret some of our criteria for measurability involving open and closed sets in terms of F_σ and G_δ sets.

Proposition 16 Measurability Using F_σ and G_δ Sets

Let $E \subseteq \mathbb{R}$. Then the following are equivalent:

1. *E is Lebesgue measurable.*

2. *$E = F \cup Z$ for some F_σ set F and some set $Z \subseteq \mathbb{R}$ of measure zero.*

3. *$E = G - Z$ for some G_δ set G and some set $Z \subseteq G$ of measure zero.*

PROOF Clearly (2) and (3) both imply (1). For the converse, suppose that E is Lebesgue measurable. For every $n \in \mathbb{N}$, let F_n be a closed set and U_n be an open set so that $F_n \subseteq E \subseteq U_n$ and $m(U_n - F_n) \leq 1/n$. Then $F = \bigcup_{n\in\mathbb{N}} F_n$ is an F_σ set and $G = \bigcap_{n\in\mathbb{N}} U_n$ is a G_δ set such that $F \subseteq E \subseteq G$. Moreover, since $G - F \subseteq G_n - F_n$ for all n, we know that $m(G - F) = 0$. Then $E = F \cup (E - F) = G - (G - E)$, where $m(E - F) = m(G - E) = 0$. ■

Borel Sets

If X is a set, recall that a **σ-algebra** on X is any nonempty collection of subsets of X that is closed under taking complements and countable unions. For example, the Lebesgue measurable subsets of $\mathbb{R}$ form a σ-algebra on $\mathbb{R}$.

Proposition 17 Intersection of σ-Algebras

Let X be a set, and let $\mathfrak{C}$ be any collection of σ-algebras on X. Then the intersection $\bigcap \mathfrak{C}$ is also a σ-algebra on X.

PROOF Since $\emptyset \in \mathcal{M}$ for every $\mathcal{M} \in \mathfrak{C}$, it follows that $\emptyset \in \bigcap \mathfrak{C}$. Next, if $S \in \bigcap \mathfrak{C}$, then $S \in \mathcal{M}$ for every $\mathcal{M} \in \mathfrak{C}$. Then $S^c \in \mathcal{M}$ for every $\mathcal{M} \in \mathfrak{C}$, and hence $S^c \in \bigcap \mathfrak{C}$. Finally, if $\{S_n\}$ is a sequence in $\bigcap \mathfrak{C}$, then each $\mathcal{M} \in \mathfrak{C}$ must contain the entire sequence $\{S_n\}$. It follows that $\bigcup_{n\in\mathbb{N}} S_n \in \mathcal{M}$ for each $\mathcal{M} \in \mathfrak{C}$, and hence $\bigcup_{n\in\mathbb{N}} S_n \in \bigcap \mathfrak{C}$. ■

You have probably seen propositions similar to this one in other fields of mathematics. For example, a similar fact from group theory is that the intersection of any collection of subgroups of a group G is again a subgroup of G. Similarly, in topology the intersection of any collection of topologies on a set X is again a topology on X.

Definition: Generators for a σ-Algebra
Let X be a set, and let $\mathcal{C}$ be any collection of subsets of X. The **σ-algebra generated by $\mathcal{C}$** is the intersection of all σ-algebras on X that contain $\mathcal{C}$.

It is important for this definition that there is always at least one σ-algebra that contains $\mathcal{C}$, namely the collection $\mathcal{P}(X)$ of all subsets of X.

EXAMPLE 2 Let X be a set, and let $\mathcal{C}$ be the collection of singleton sets in X, i.e.

$$\mathcal{C} = \bigl\{\{x\} \bigm| x \in X\bigr\}.$$

Then it is not hard to check that the σ-algebra generated by $\mathcal{C}$ consists of all sets $S \subseteq X$ for which either S or S^c is countable. ■

Definition: Borel Sets
The **Borel algebra** $\mathcal{B}$ is the σ-algebra on $\mathbb{R}$ generated by the collection of all open sets. A set $B \subseteq \mathbb{R}$ is called a **Borel set** if $B \in \mathcal{B}$.

By definition every open set is a Borel set. Moreover, since the Borel sets are a σ-algebra, the complement of any Borel set is a Borel set, and any countable union of Borel sets is a Borel set.

Proposition 18 Properties of Borel Sets

1. *Every Borel set is measurable.*
2. *Every open set, closed set, F_σ set, or G_δ set is a Borel set.*
3. *The Borel algebra is generated by the collection of all open intervals.*

PROOF For (1), observe that the collection $\mathcal{M}$ of all Lebesgue measurable sets is a σ-algebra that contains the open sets. Since $\mathcal{B}$ is the intersection of all such σ-algebras, it follows that $\mathcal{B} \subseteq \mathcal{M}$.

For (2), every open set lies in $\mathcal{B}$ by definition. Since $\mathcal{B}$ is a σ-algebra, it follows immediately that closed sets, F_σ sets, and G_δ sets lie in $\mathcal{B}$ as well.

For (3), let $\mathcal{A}$ be the σ-algebra generated by the open intervals. Since $\mathcal{B}$ contains the open intervals, we know that $\mathcal{A} \subseteq \mathcal{B}$. But since every open set is a countable union of open intervals, $\mathcal{A}$ contains every open set, and hence $\mathcal{B} \subseteq \mathcal{A}$. ■

As we will see, open sets, closed sets, F_σ sets, and G_δ sets are among the simplest of the Borel sets. In the rest of this section, we describe the overall structure of the Borel algebra. We will not prove any of the theorems below, and indeed any of the proofs would be beyond the scope of this course.

Definition: Finite Borel Hierarchy
The **finite Borel hierarchy** consists of two sequences $\{\Sigma_n\}$ and $\{\Pi_n\}$ of subsets of $\mathcal{B}$ defined as follows:

- Σ_1 is the collection of all open sets in $\mathbb{R}$, and Π_1 is the collection of all closed sets in $\mathbb{R}$.
- For each $n \geq 1$, the collection Σ_{n+1} consists of all countable unions of sets from Π_n, and the collection Π_{n+1} consists of all countable intersections of sets from Σ_n.

For example, Σ_2 is the collection of all F_σ sets, and Π_2 is the collection of all G_δ sets. Similarly, Σ_3 is the collection of all countable unions of G_δ sets, and Π_3 is the collection of all countable intersections of F_σ sets. Thus every set in Σ_3 can be written as

$$\bigcup_{m\in\mathbb{N}} \bigcap_{n\in\mathbb{N}} U_{m,n}$$

for some open sets $U_{m,n}$ and every set in Π_3 can be written as

$$\bigcap_{m\in\mathbb{N}} \bigcup_{n\in\mathbb{N}} F_{m,n}$$

for some closed sets $F_{m,n}$.

As mentioned in the previous section, every open or closed set is both F_σ and G_δ. Thus we have

$$\Sigma_1 \subseteq \Sigma_2, \qquad \Sigma_1 \subseteq \Pi_2, \qquad \Pi_1 \subseteq \Sigma_2, \qquad \text{and} \qquad \Pi_1 \subseteq \Pi_2.$$

Moreover, all four of these inclusions are proper. In particular, $\mathbb{Q}$ is in both $\Sigma_2 - \Sigma_1$ and $\Sigma_2 - \Pi_1$, and $\mathbb{R} - \mathbb{Q}$ is in both $\Pi_2 - \Sigma_1$ and $\Pi_2 - \Pi_1$. The following theorem generalizes all of this.

Theorem 19 Properties of Σ_n and Π_n

For each $n \in \mathbb{N}$, the following statements hold.

1. *For all $S \subseteq \mathbb{R}$, we have $S \in \Sigma_n$ if and only if $S^c \in \Pi_n$.*

2. *We have*

$$\Sigma_n \subseteq \Sigma_{n+1}, \qquad \Sigma_n \subseteq \Pi_{n+1}, \qquad \Pi_n \subseteq \Sigma_{n+1}, \qquad \text{and} \qquad \Pi_n \subseteq \Pi_{n+1}.$$

Moreover, all four of these inclusions are proper.

If $B \subseteq \mathbb{R}$ is a Borel set, the **Borel rank** of B is the minimum n such that B lies in $\Sigma_n \cup \Pi_n$. Thus sets that are open or closed have Borel rank 1, sets that are F_σ or G_δ have Borel rank 2, and so forth.

Amazingly, it is not true that every Borel set has finite rank. For example if $\{S_n\}$ is a sequence of Borel sets such that each S_n is contained in $(n, n+1)$ and has rank n, then the union

$$S = S_1 \cup S_2 \cup S_3 \cup \cdots$$

cannot have any finite rank. Such a set S is said to have **rank $\boldsymbol{\omega}$**, and the collection of all such sets is known as Σ_ω. The complement of any set in Σ_ω is also said to have rank ω, and the collection of all such sets is known as Π_ω.

The Borel hierarchy continues even beyond ω. For example, $\Sigma_{\omega+1}$ consists of all countable unions of sets from Π_ω, and $\Pi_{\omega+1}$ consists of all countable intersections of sets from Σ_ω. Indeed, we have a sequence of sets

$$\Sigma_1 \subseteq \Sigma_2 \subseteq \Sigma_3 \subseteq \cdots \subseteq \Sigma_\omega \subseteq \Sigma_{\omega+1} \subseteq \Sigma_{\omega_2} \subseteq \cdots \subseteq \Sigma_{2\omega} \subseteq \Sigma_{2\omega+1} \subseteq \cdots$$

and similarly for the Π's. The result is that the sets Σ_α and Π_α can be defined for each countable ordinal α (i.e. for each element of a minimal uncountable well-ordered set S_Ω). The resulting families $\{\Sigma_\alpha\}_{\alpha\in S_\Omega}$ and $\{\Pi_\alpha\}_{\alpha\in S_\Omega}$ constitute the full **Borel hierarchy**, and the Borel algebra $\mathcal{B}$ is the union of these:

$$\mathcal{B} = \bigcup_{\alpha\in S_\Omega} \Sigma_\alpha = \bigcup_{\alpha\in S_\Omega} \Pi_\alpha.$$

Incidentally, it is not too hard to prove that each of the sets Σ_α and Π_α has the same cardinality as $\mathbb{R}$. Since $|S_\Omega| \leq |\mathbb{R}|$, it follows that the full Borel algebra $\mathcal{B}$ has cardinality $|\mathbb{R}|$ as well.

Theorem 20 Cardinality of the Borel Algebra

Let $\mathcal{B}$ be the Borel σ-algebra in $\mathbb{R}$. Then

$$|\mathcal{B}| = |\mathbb{R}|.$$

By Corollary 9, the cardinality of the collection of measurable sets is equal to $|\mathcal{P}(\mathbb{R})|$, which is greater than the cardinality of the Borel algebra. This yields the following corollary.

Corollary 21

There exists a Lebesgue measurable set that is not a Borel set.

Since the Borel algebra is a σ-algebra, we could of course restrict Lebesgue measure to the Borel algebra, yielding a measure $m|_{\mathcal{B}} \colon \mathcal{B} \to \mathbb{R}$. However, this measure is not complete, since there exist Borel sets of measure zero (such as the Cantor set) whose subsets are not all Borel. Indeed, by Proposition 16, Lebesgue measure is precisely the completion of the measure $m|_{\mathcal{B}}$.

Exercises

1. Prove that every nonempty open set has positive measure.

2. a) Let $S \subseteq \mathbb{R}$, and let $\mathcal{C}$ be a collection of open sets that covers S. Prove that $\mathcal{C}$ has a countable subcollection that covers S.

   b) A set $S \subseteq \mathbb{R}$ is **locally measurable** if for every point $x \in S$ there exists an open set U containing x so that $S \cap U$ is measurable. Prove that every locally measurable set is measurable.

3. A subset of $\mathbb{R}$ is **totally disconnected** if it does not contain any open intervals (e.g. the Cantor set). Give an example of a closed, totally disconnected subset of $[0, 1]$ that has positive measure.

4. a) If $S \subseteq \mathbb{R}$, prove that $m^*(S) = \inf\{m(E) \mid E \text{ is measurable and } S \subseteq E\}$.

   b) If $S \subseteq \mathbb{R}$, prove that $m_*(S) = \sup\{m(E) \mid E \text{ is measurable and } E \subseteq S\}$.

5. Let $E \subseteq \mathbb{R}$ be a measurable set with $m(E) < \infty$, and let $S \subseteq E$. Prove that

$$m_*(S) \;=\; m(E) - m^*(E - S).$$

6. If $\{S_n\}$ is a sequence of pairwise disjoint subsets of $\mathbb{R}$, prove that

$$m_*\left(\biguplus_{n\in\mathbb{N}} S_n\right) \;\geq\; \sum_{n\in\mathbb{N}} m_*(S_n).$$

7. Prove that a set $S \subseteq \mathbb{R}$ is F_σ if and only if its complement is G_δ.

8. Prove that every countable subset of $\mathbb{R}$ is F_σ.

9. Prove that the intersection of two F_σ sets is F_σ.

10. Prove that every open set is both F_σ and G_δ. Deduce that the same holds true for closed sets.

11. Give an example of a set which is both F_σ and G_δ but is neither open nor closed.

12. Let $\mathcal{C}$ be the collection of all uncountable subsets of $\mathbb{R}$. Prove that the σ-algebra generated by $\mathcal{C}$ is the power set of $\mathbb{R}$.

13. Prove that the σ-algebra generated by the collection $\{(a, \infty) \mid a \in \mathbb{R}\}$ is the Borel sets.

14. Let $\mathcal{M}$ be the collection of Lebesgue measurable sets in $\mathbb{R}$. Prove that $\mathcal{M}$ is the σ-algebra generated by the open intervals together with all sets of measure zero.