5

Measure Theory

This chapter contains the background from measure theory that is required to understand the Hausdorff dimension. It is true that the Hausdorff dimension can be defined in half a page without reference to measure theory, but when it is done that way there is no indication of the motivation for the definition.

Measure theory will also be indispensable in many of the proofs related to fractal dimension. It will simplify many of the proofs of lower bounds for Hausdorff dimension and of upper bounds for packing dimension. Instead of repeating a combinatorial calculation in each instance, we do the combinatorics once and for all in this chapter, and then repeatedly reap the benefits in Chap. 6.

Since measure theory (like metric topology) is a standard part of graduate mathematics curriculum today, most of the introductory remarks to Chap. 2 are also applicable here.

5.1 Lebesgue Measure

Certain calculations will be done with the symbols ∞ and $-\infty$. They are not real numbers, but they can be useful in connection with calculations involving real numbers. Most of the conventions are sensible when you think about them. Here are some examples:

(1) If $a \in \mathbb{R}$, then $-\infty < a < \infty$.
(2) If $a \in \mathbb{R}$, then $a + \infty = \infty$ and $a - \infty = -\infty$. Also $\infty + \infty = \infty$ and $-\infty - \infty = -\infty$. The combination $\infty - \infty$ is not defined.
(3) If $a \in \mathbb{R}$ is positive, then $a \times \infty = \infty$ and $a \times (-\infty) = -\infty$. If $a \in \mathbb{R}$ is negative, then $a \times \infty = -\infty$ and $a \times (-\infty) = \infty$. The combination $0 \times \infty$ is not defined. [However, we do understand that an infinite series $\sum_{n=1}^{\infty} a_n$, where every term $a_n = 0$, has sum 0.]

The ***length*** of one of the intervals

$$(a, b) \quad (a, b\,] \quad [\,a, b) \quad [\,a, b\,]$$

is $b-a$, where $a, b \in \mathbb{R}$ and $a < b$. The length of the degenerate interval $[\,a, a\,] = \{a\}$ is 0; the length of the empty set $\varnothing$ is 0. The length of an unbounded interval

$$(a, \infty) \quad [\,a, \infty) \quad (-\infty, b) \quad (-\infty, b\,] \quad (-\infty, \infty)$$

is ∞. This follows the conventions on calculation with ∞.

We will be interested in a substantial generalization of the notion of the "length" of a subset of $\mathbb{R}$. The lemma that makes it possible asserts that the length of a countable union of intervals cannot exceed the sum of the lengths of the parts.

Lemma 5.1.1. *Suppose the closed interval $[c, d]$ is covered by a countable family of open intervals:*

$$[c, d] \subseteq \bigcup_{i \in \mathbb{N}} (a_i, b_i).$$

Then

$$d - c < \sum_{i=1}^{\infty} (b_i - a_i).$$

Proof. First, since $[c, d]$ is a compact set, it is in fact covered by a finite number of the intervals:

$$[c, d] \subseteq \bigcup_{i=1}^{n} (a_i, b_i)$$

for some n. I will show that when this happens, the conclusion

$$d - c < \sum_{i=1}^{n} (b_i - a_i)$$

follows. The proof is by induction on n.

If $n = 1$, then $[c, d] \subseteq (a_1, b_1)$, so $a_1 < c$ and $d < b_1$. Thus $d - c < b_1 - a_1$, as required.

Now suppose $n \geq 2$, and the result is true for covers by at most $n-1$ open intervals. Suppose

$$[c, d] \subseteq \bigcup_{i=1}^{n} (a_i, b_i).$$

If some interval (a_i, b_i) is disjoint from $[c, d]$, it may be omitted from the cover; then we have a cover by at most $n - 1$ sets, so we would be finished by the induction hypothesis. So assume $(a_i, b_i) \cap [c, d] \neq \varnothing$ for all i. Among all of the left endpoints a_i, there is one that is no larger than any of the others. By renumbering the intervals, let us assume that it is a_1. Since c is covered, we must have $a_1 < c$. Now if $b_1 > d$, we have $d - c < b_1 - a_1 \leq \sum_{i=1}^{n}(b_i - a_i)$, so we are finished. So suppose $b_1 \leq d$. Since (a_1, b_1) intersects $[c, d]$, we have $b_1 > c$. So $b_1 \in [c, d]$. At least one of the open intervals (a_i, b_i) covers the point b_1. By renumbering, we may assume it is (a_2, b_2). Finally, we have a cover of $[c, d]$ by $n - 1$ sets:

$$[c, d] \subseteq (a_1, b_2) \cup \bigcup_{i=3}^{n}(a_i, b_i).$$

So by the induction hypothesis,

$$\begin{aligned} d - c &< (b_2 - a_1) + \sum_{i=3}^{n}(b_i - a_i) \\ &\leq (b_2 - a_2) + (b_1 - a_1) + \sum_{i=3}^{n}(b_i - a_i) \end{aligned}$$

as required. This completes the proof by induction. □

A useful generalization of the notion of the length of a subset of $\mathbb{R}$ is the ***Lebesgue measure*** of the set. This will be defined in stages. We will use half-open intervals of the form $[a, b)$ in the definition. Intervals of other forms could be used instead, but these have been chosen because of this convenient property:

Lemma 5.1.2. *Let $a < b$ be real numbers, and $\varepsilon > 0$. Then $[a, b)$ can be written as a finite disjoint union*

$$[a, b) = \bigcup_{i=1}^{n}[a_i, b_i),$$

with $b - a = \sum_{i=1}^{n}(b_i - a_i)$ and $b_i - a_i \leq \varepsilon$ for all i.

Proof. Choose $n \in \mathbb{N}$ so large that $(b - a)/n \leq \varepsilon$. Let $b_i = a + i(b - a)/n$ for $0 \leq i \leq n$, and $a_i = b_{i-1}$. □

Now let A be any subset of $\mathbb{R}$. The ***Lebesgue outer measure*** of A is obtained by covering A with countably many half-open intervals of total length as small as possible. In symbols,*

$$\overline{\mathcal{L}}(A) = \inf \sum_{j=1}^{\infty}(b_j - a_j)$$

where the infimum is over all countable families $\{\, [a_j, b_j) : j \in \mathbb{N} \,\}$ of half-open intervals with $A \subseteq \bigcup_{j\in\mathbb{N}}[a_j, b_j)$.

---

* In case you can't tell, the symbol $\mathcal{L}$ is supposed to be a fancy letter L, for "Lebesgue".

Lemma 5.1.3. *Let $A \subseteq \mathbb{R}$ and let $\varepsilon > 0$. Then*

$$\overline{\mathcal{L}}(A) = \inf \sum_{j=1}^{\infty} (b_j - a_j)$$

where the infimum is over all countable families $\{ [a_j, b_j) : j \in \mathbb{N} \}$ of half-open intervals with $A \subseteq \bigcup_{j\in\mathbb{N}} [a_j, b_j)$ and $b_j - a_j \le \varepsilon$ for all j.

Proof. This follows from Lemma 5.1.2. □

We must do some combinatorics on the line to see that the definition is not trivial.*

Theorem 5.1.4. *If A is an interval, then $\overline{\mathcal{L}}(A)$ is the length of A.*

Proof. Suppose $A = [a, b]$, where $a < b$ are real numbers. First, if $\varepsilon > 0$, then the singleton $\{[a, b + \varepsilon)\}$ covers the set A, so $\overline{\mathcal{L}}(A) \le b - a + \varepsilon$. This is true for any $\varepsilon > 0$, so $\overline{\mathcal{L}}(A) \le b - a$.

Now suppose $A \subseteq \bigcup_{j\in\mathbb{N}} [a_j, b_j)$. Let $\varepsilon > 0$, and write $a'_j = a_j - \varepsilon/2^j$. Then $A \subseteq \bigcup_{j\in\mathbb{N}} (a'_j, b_j)$. By Lemma 5.1.1, $\sum_{j=1}^{\infty} (b_j - a'_j) > b - a$. So we have $\sum_{j=1}^{\infty} (b_j - a_j) \ge \sum_{j=1}^{\infty} (b_j - a'_j) - \varepsilon > b - a - \varepsilon$. This is true for any $\varepsilon > 0$, so $\sum_{j=1}^{\infty} (b_j - a_j) \ge b - a$. Therefore $\overline{\mathcal{L}}(A) \ge b - a$. So we have $\overline{\mathcal{L}}\big([a, b]\big) = b - a$.

Next consider $A = (a, b)$. Then $\overline{\mathcal{L}}(A) \le \overline{\mathcal{L}}([a, b]) = b - a$ and on the other hand $\overline{\mathcal{L}}(A) \ge \overline{\mathcal{L}}([a + \varepsilon, b - \varepsilon]) = b - a - 2\varepsilon$ for any $\varepsilon > 0$. Similar arguments cover cases $[a, b)$ and $(a, b]$. If $A = [a, \infty)$, then $A \supseteq [a, a + t]$ for any $t > 0$, and therefore $\overline{\mathcal{L}}(A) \ge t$; this means that $\overline{\mathcal{L}}(A) = \infty$. Similar arguments cover the other cases of infinite length intervals. □

Here are some of the basic properties of Lebesgue outer measure.

Theorem 5.1.5. (1) $\overline{\mathcal{L}}(\varnothing) = 0$;
(2) *if $A \subseteq B$, then $\overline{\mathcal{L}}(A) \le \overline{\mathcal{L}}(B)$;*
(3) $\overline{\mathcal{L}}\left(\bigcup_{n\in\mathbb{N}} A_n\right) \le \sum_{n=1}^{\infty} \overline{\mathcal{L}}(A_n)$.

Proof. For (1), note that $\varnothing \subseteq \bigcup_{i\in\mathbb{N}} [0, \varepsilon/2^i)$, so $\overline{\mathcal{L}}(\varnothing) \le \varepsilon$. For (2), note that any cover of B is also a cover of A.

Now consider (3). If $\overline{\mathcal{L}}(A_n) = \infty$ for some n, then the inequality is clear. So suppose $\overline{\mathcal{L}}(A_n) < \infty$ for all n. Let $\varepsilon > 0$. For each n, choose a countable cover $\mathcal{D}_n$ of A_n by half-open intervals with

$$\sum_{D\in\mathcal{D}_n} \overline{\mathcal{L}}(D) \le \overline{\mathcal{L}}(A_n) + 2^{-n}\varepsilon.$$

Now $\mathcal{D} = \bigcup_{n\in\mathbb{N}} \mathcal{D}_n$ is a countable cover of the union $\bigcup_{n\in\mathbb{N}} A_n$. Therefore

* I can easily write down the same definition for subsets of the rational numbers. But then every set turns out to have outer measure 0, so it is not a very useful definition.

$$\overline{\mathcal{L}}\left(\bigcup_{n\in\mathbb{N}} A_n\right) \le \sum_{D\in\mathcal{D}} \overline{\mathcal{L}}(D) \le \sum_{n=1}^{\infty} \sum_{D\in\mathcal{D}_n} \overline{\mathcal{L}}(D)$$
$$\le \sum_{n=1}^{\infty} \overline{\mathcal{L}}(A_n) + \sum_{n=1}^{\infty} 2^{-n}\varepsilon = \sum_{n=1}^{\infty} \overline{\mathcal{L}}(A_n) + \varepsilon.$$

Since ε was any positive number, we have

$$\overline{\mathcal{L}}\left(\bigcup_{n\in\mathbb{N}} A_n\right) \le \sum_{n=1}^{\infty} \overline{\mathcal{L}}(A_n).$$ □

In general, the inequality in part (3) is not equality, even for two disjoint sets. But we do have equality in some cases. The simplest case is the following:

Theorem 5.1.6. *Let* $A, B \subseteq \mathbb{R}$ *with* $\operatorname{dist}(A, B) > 0$*. Then* $\overline{\mathcal{L}}(A \cup B) = \overline{\mathcal{L}}(A) + \overline{\mathcal{L}}(B)$*.*

Proof. First, the inequality $\overline{\mathcal{L}}(A \cup B) \le \overline{\mathcal{L}}(A) + \overline{\mathcal{L}}(B)$ follows from part (3) of Theorem 5.1.5. Let $\varepsilon = \operatorname{dist}(A, B)/2$, and let $A \cup B \subseteq \bigcup_{j\in\mathbb{N}} [a_j, b_j)$, where $b_j - a_j \le \varepsilon$ for all j. Then each interval $[a_j, b_j)$ intersects at most one of the sets A and B. So the collection $\mathcal{D} = \{\, [a_j, b_j) : j \in \mathbb{N} \,\}$ can be written as the disjoint union of two collections, $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$, where $\mathcal{D}_1$ covers A and $\mathcal{D}_2$ covers B. Now $\overline{\mathcal{L}}(A) \le \sum_{D\in\mathcal{D}_1} \overline{\mathcal{L}}(D)$ and $\overline{\mathcal{L}}(B) \le \sum_{D\in\mathcal{D}_2} \overline{\mathcal{L}}(D)$, so

$$\overline{\mathcal{L}}(A) + \overline{\mathcal{L}}(B) \le \sum_{D\in\mathcal{D}_1} \overline{\mathcal{L}}(D) + \sum_{D\in\mathcal{D}_2} \overline{\mathcal{L}}(D) = \sum_{D\in\mathcal{D}} \overline{\mathcal{L}}(D) \le \sum_{j=1}^{\infty} b_j - a_j.$$

Therefore, by Lemma 5.1.3, we have $\overline{\mathcal{L}}(A) + \overline{\mathcal{L}}(B) \le \overline{\mathcal{L}}(A \cup B)$. □

Corollary 5.1.7. *If* $A, B \subseteq \mathbb{R}$ *are disjoint and compact, then we have* $\overline{\mathcal{L}}(A) + \overline{\mathcal{L}}(B) = \overline{\mathcal{L}}(A \cup B)$*.*

Proof. Apply Theorems 2.3.19 and 5.1.6. □

Theorem 5.1.8. *If* $A \subseteq \mathbb{R}$*, then*

$$\overline{\mathcal{L}}(A) = \inf\left\{\, \overline{\mathcal{L}}(U) : U \supseteq A,\ U \text{ open} \,\right\}.$$

Proof. Certainly $\overline{\mathcal{L}}(A) \le \inf\left\{\, \overline{\mathcal{L}}(U) : U \supseteq A, U \text{ open} \,\right\}$. So I must prove the opposite inequality. If $\overline{\mathcal{L}}(A) = \infty$, it is trivially true. So suppose $\overline{\mathcal{L}}(A) < \infty$. Let $\varepsilon > 0$. Then there exists a cover $\bigcup_{j\in\mathbb{N}} [a_j, b_j)$ of A with $\sum_{j=1}^{\infty} (b_j - a_j) \le \overline{\mathcal{L}}(A) + \varepsilon/2$. Now the set $U = \bigcup_{j\in\mathbb{N}} (a_j - \varepsilon/2^{j+1}, b_j)$ is open, $U \supseteq A$, and $\overline{\mathcal{L}}(U) \le \sum_{j=1}^{\infty} (b_j - a_j) + \varepsilon/2 \le \overline{\mathcal{L}}(A) + \varepsilon$. Therefore $\overline{\mathcal{L}}(A) + \varepsilon \ge \overline{\mathcal{L}}(U)$. This shows that $\overline{\mathcal{L}}(A) \ge \inf\left\{\, \overline{\mathcal{L}}(U) : U \supseteq A \,\right\}$. □

The outer measure $\overline{\mathcal{L}}(A)$ of a set $A \subseteq \mathbb{R}$ is determined by approximating a set from the outside by open sets. There is a corresponding "inner measure", obtained by approximating a set from the inside. This time, however, we will use compact sets.

Let $A \subseteq \mathbb{R}$. The ***Lebesgue inner measure*** of the set A is

$$\underline{\mathcal{L}}(A) = \sup \left\{ \overline{\mathcal{L}}(K) : K \subseteq A, K \text{ compact} \right\}.$$

Again, we need an argument to see that the definition is interesting.

Theorem 5.1.9. *If A is an interval, then $\underline{\mathcal{L}}(A)$ is the length of A.*

Proof. We consider the case of an open interval $A = (a, b)$. Other kinds of intervals follow from this case as before.

If $K \subseteq A$ is compact, then K is covered by the single interval A, so that $\overline{\mathcal{L}}(K) \le b - a$. Therefore $\underline{\mathcal{L}}(A) \le b - a$. On the other hand, if $\varepsilon > 0$, then the set $[a + \varepsilon, b - \varepsilon]$ is compact, so $\underline{\mathcal{L}}(A) \ge \overline{\mathcal{L}}([a + \varepsilon, b - \varepsilon]) = b - a - 2\varepsilon$. This is true for any $\varepsilon > 0$, so $\underline{\mathcal{L}}(A) \ge b - a$. □

Exercise 5.1.10. If $A \subseteq \mathbb{R}$ is any set, then $\underline{\mathcal{L}}(A) \le \overline{\mathcal{L}}(A)$.

It is not possible to prove that $\underline{\mathcal{L}}(A) = \overline{\mathcal{L}}(A)$ in general. A set A is called ***Lebesgue measurable***, roughly speaking, when this equation is true. Precisely: If $\overline{\mathcal{L}}(A) < \infty$, then A is Lebesgue measurable iff $\underline{\mathcal{L}}(A) = \overline{\mathcal{L}}(A)$. If $\overline{\mathcal{L}}(A) = \infty$, then A is Lebesgue measurable iff $A \cap [-n, n]$ is Lebesgue measurable for all $n \in \mathbb{N}$. If A is Lebesgue measurable, we will write $\mathcal{L}(A)$ for the common value of $\overline{\mathcal{L}}(A)$ and $\underline{\mathcal{L}}(A)$, and call it simply the ***Lebesgue measure*** of A. We will often say simply ***measurable*** when we mean Lebesgue measurable.

Theorem 5.1.11. *Let $A_1, A_2, \cdots$ be disjoint Lebesgue measurable sets. Then $\bigcup_n A_n$ is measurable, and $\mathcal{L}(\bigcup_n A_n) = \sum_n \mathcal{L}(A_n)$.*

Proof. It is enough to prove the theorem in the case that $\mathcal{L}(\bigcup A_n) < \infty$, since the general case will then follow by applying this case to sets $A_n \cap [-m, m]$. We know by Theorem 5.1.5 that $\overline{\mathcal{L}}(\bigcup A_n) \le \sum \mathcal{L}(A_n)$. Let $\varepsilon > 0$. For each n, choose a compact set $K_n \subseteq A_n$ with $\overline{\mathcal{L}}(K_n) \ge \underline{\mathcal{L}}(A_n) - \varepsilon/2^n$. Since A_n is measurable, $\overline{\mathcal{L}}(K_n) \ge \overline{\mathcal{L}}(A_n) - \varepsilon/2^n$. Now the sets K_n are disjoint, so by Corollary 5.1.7, the compact set $L_m = K_1 \cup K_2 \cup \cdots \cup K_m$ satisfies $\overline{\mathcal{L}}(L_m) = \overline{\mathcal{L}}(K_1) + \cdots + \overline{\mathcal{L}}(K_m)$. Therefore $\underline{\mathcal{L}}(\bigcup A_n) \ge \sum_{n=1}^{m} \overline{\mathcal{L}}(K_n)$. Now this is true for all m, so $\underline{\mathcal{L}}(\bigcup A_n) \ge \sum_{n=1}^{\infty} \mathcal{L}(K_n) \ge \sum_{n=1}^{\infty} \mathcal{L}(A_n) - \varepsilon$. This is true for any positive ε, so we have $\underline{\mathcal{L}}(\bigcup_n A_n) \ge \sum \mathcal{L}(A_n)$.

So $\overline{\mathcal{L}}(\bigcup A_n) = \underline{\mathcal{L}}(\bigcup A_n)$, and therefore $\bigcup A_n$ is measurable and $\mathcal{L}(\bigcup A_n) = \sum \mathcal{L}(A_n)$. □

Theorem 5.1.12. *Compact subsets, closed subsets, and open subsets of $\mathbb{R}$ are Lebesgue measurable.*

Proof. Let $K \subseteq \mathbb{R}$ be compact. It is bounded, so $K \subseteq [-n, n]$ for some n, and therefore $\overline{\mathcal{L}}(K) < \infty$. The compact set K is a subset of K, so $\underline{\mathcal{L}}(K) \geq \overline{\mathcal{L}}(K)$.

Let $F \subseteq \mathbb{R}$ be a closed set. Then for each $n \in \mathbb{N}$, the intersection $F \cap [-n, n]$ is compact, and therefore measurable. Thus F is measurable.

Let U be an open set. It is enough to do the case $\overline{\mathcal{L}}(U) < \infty$. For each $x \in U$, there is an open interval I with $x \in I \subseteq U$. By the Lindelöf property, U is the union of countably many of these intervals, say $U = \bigcup_{j \in \mathbb{N}} I_j$. Now each set $I_n \setminus \bigcup_{j=1}^{n-1} I_j$ is a finite union of intervals (open, closed, half-open) so that U is a disjoint union of countably many intervals. So U is measurable. □

Theorem 5.1.13. *Let $A \subseteq \mathbb{R}$. Then A is measurable if and only if, for every $\varepsilon > 0$, there exist an open set U and a closed set F with $U \supseteq A \supseteq F$ and $\mathcal{L}(U \setminus F) < \varepsilon$.*

Proof. Suppose first that A is measurable. We consider first of all the case $\mathcal{L}(A) < \infty$. Then there exists an open set $U \supseteq A$ such that $\mathcal{L}(U) < \mathcal{L}(A) + \varepsilon/2$. There exists a compact (therefore closed) set $F \subseteq A$ with $\mathcal{L}(F) > \mathcal{L}(A) - \varepsilon/2$. Now $U \setminus F$ is open, hence measurable, and F is compact, hence measurable, so $\mathcal{L}(U) = \mathcal{L}(U \setminus F) + \mathcal{L}(F)$. Since the terms are all finite, we may subtract, and we get

$$\mathcal{L}(U \setminus F) = \mathcal{L}(U) - \mathcal{L}(F) < \mathcal{L}(A) + \varepsilon/2 - \mathcal{L}(A) + \varepsilon/2 = \varepsilon.$$

Now we take the case $\mathcal{L}(A) = \infty$. All of the sets $A \cap [-n, n]$ are measurable, so there exist open sets $U_n \supseteq A \cap [-n, n]$ and compact sets $F_n \subseteq A \cap [-n, n]$ with $\mathcal{L}(U_n \setminus F_n) < \varepsilon/2^{n+2}$. Define $U'_n = U_n \cap \big((-\infty, -n + 1 + \varepsilon/2^{n+2}) \cup (n - 1 - \varepsilon/2^{n+2}, \infty)\big)$ and $F'_n = F_n \cap \big([-n, -n + 1] \cup [n - 1, n]\big)$, so that U'_n is open, F'_n is compact, $U'_n \supseteq A \cap \big([-n, -n + 1] \cup [n - 1, n]\big) \supseteq F'_n$ and $\mathcal{L}(U'_n \setminus F'_n) < 3\varepsilon/2^{n+2} < \varepsilon/2^n$. Now $U = \bigcup U'_n$ is open, and $F = \bigcup F'_n$ is closed (Exercise 2.2.27). We have $U \supseteq A \supseteq F$, and $U \setminus F \subseteq \bigcup_{n \in \mathbb{N}} (U'_n \setminus F'_n)$, so that $\mathcal{L}(U \setminus F) \leq \sum \mathcal{L}(U'_n \setminus F'_n) < \varepsilon$.

Conversely, suppose sets U and F exist. By Theorem 5.1.12 they are measurable. First assume $\overline{\mathcal{M}}(A) < \infty$. Then $\mathcal{L}(F) < \infty$, and $\mathcal{L}(U) \leq \mathcal{L}(U \setminus F) + \mathcal{L}(F) < \varepsilon + \mathcal{L}(F) < \infty$. Now $\overline{\mathcal{M}}(A) \leq \overline{\mathcal{M}}(U) = \mathcal{L}(U) < \mathcal{L}(F) + \varepsilon = \underline{\mathcal{L}}(F) + \varepsilon \leq \underline{\mathcal{L}}(A) + \varepsilon$. This is true for any $\varepsilon > 0$, so $\overline{\mathcal{M}}(A) = \underline{\mathcal{L}}(A)$, so A is measurable.

For the case $\overline{\mathcal{L}}(A) = \infty$, we have $U \cap (-n - \varepsilon, n + \varepsilon) \supseteq A \cap [-n, n] \supseteq F \cap [-n, n]$, and the previous case may be applied to these sets, using 3ε in place of ε. □

Here are the basic algebraic properties of Lebesgue measurable sets.

Theorem 5.1.14. (1) *Both $\varnothing$ and $\mathbb{R}$ are Lebesgue measurable.*
(2) *If $A \subseteq \mathbb{R}$ is Lebesgue measurable, then so is its complement $\mathbb{R} \setminus A$.*
(3) *If A and B are measurable, then so are $A \cap B$, $A \cup B$, and $A \setminus B$.*
(4) *If A_n is measurable for $n \in \mathbb{N}$, then so are $\bigcup_{n \in \mathbb{N}} A_n$ and $\bigcap_{n \in \mathbb{N}} A_n$.*

Proof. For (1), note that $\overline{\mathcal{L}}(\varnothing) = 0$ and $\mathbb{R} \cap [-n, n]$ is measurable for all n.

For (2), note that if $F \subseteq A \subseteq U$, then $\mathbb{R} \setminus U \subseteq \mathbb{R} \setminus A \subseteq \mathbb{R} \setminus F$ and $(\mathbb{R} \setminus F) \setminus (\mathbb{R} \setminus U) = U \setminus F$.

For the intersection in (3), note that if $F_1 \subseteq A \subseteq U_1$ and $F_2 \subseteq B \subseteq U_2$, then $F_1 \cap F_2 \subseteq A \cap B \subseteq U_1 \cap U_2$ and $(U_1 \cap U_2) \setminus (F_1 \cap F_2) \subseteq (U_1 \setminus F_1) \cup (U_2 \setminus F_2)$. This is enough to show that $A \cap B$ is measurable. Now $A \cup B = \mathbb{R} \setminus \big((\mathbb{R} \setminus A) \cap (\mathbb{R} \setminus B)\big)$, so $A \cup B$ is measurable. And $A \setminus B = A \cap (\mathbb{R} \setminus B)$, so $A \setminus B$ is measurable.

Finally, for (4), note that by (3) we may find disjoint measurable sets B_n with the same union as A_n, so that Theorem 5.1.11 is applicable. The intersection follows by taking complements. □

Note that (4) involves only *countable* unions and intersections.

Proposition 5.1.15. *The Lebesgue measure of the triadic Cantor dust is* 0.

Proof. The Cantor dust C is constructed on p. 2. The set $C_n \supseteq C$ consists of 2^n disjoint intervals of length 3^{-n}. Therefore $\mathcal{L}(C) \le 2^n \cdot 3^{-n}$. This has limit 0, so $\mathcal{L}(C) = 0$. □

Carathéodory Measurability

Carathéodory provided an alternate definition of measurability. Its disadvantage is that the motivation is not as clear. Its advantage is (as we will see in Sect. 5.2) that it can be used in other situations.

A set $A \subseteq \mathbb{R}$ is ***Carathéodory measurable*** iff

$$\overline{\mathcal{L}}(E) = \overline{\mathcal{L}}(E \cap A) + \overline{\mathcal{L}}(E \setminus A)$$

for all sets $E \subseteq \mathbb{R}$.

Proposition 5.1.16. *A set $A \subseteq \mathbb{R}$ is Carathéodory measurable if and only if it is Lebesgue measurable.*

Proof. Suppose A is Lebesgue measurable. Let E be a test set. The inequality $\overline{\mathcal{L}}(E) \le \overline{\mathcal{L}}(E \cap A) + \overline{\mathcal{L}}(E \setminus A)$ is always true. Let $\varepsilon > 0$. There exist an open set U and a closed set F with $F \subseteq A \subseteq U$ and $\mathcal{L}(U \setminus F) < \varepsilon$. Let $V \supseteq E$ be an open set. Then

$$\begin{aligned}
\overline{\mathcal{L}}(E \setminus A) + \overline{\mathcal{L}}(E \cap A) &\le \mathcal{L}(V \setminus F) + \mathcal{L}(V \cap U) \\
&\le \mathcal{L}(V \setminus U) + \mathcal{L}(U \setminus F) + \mathcal{L}(V \cap U) \\
&< \mathcal{L}(V) + \varepsilon.
\end{aligned}$$

Now take the infimum over all such V, to get $\overline{\mathcal{L}}(E \cap A) + \overline{\mathcal{L}}(E \setminus A) < \overline{\mathcal{L}}(E) + \varepsilon$. Therefore $\overline{\mathcal{L}}(E \cap A) + \overline{\mathcal{L}}(E \setminus A) \le \overline{\mathcal{L}}(E)$. This proves that A is Carathéodory measurable.

Conversely, suppose A is Carathéodory measurable. Consider the case in which $\overline{\mathcal{L}}(A) < \infty$. Let $\varepsilon > 0$. Let $U \supseteq A$ satisfy $\mathcal{L}(U) < \overline{\mathcal{L}}(A) + \varepsilon$. Now we have

$$\overline{\mathcal{L}}(U) = \overline{\mathcal{L}}(U \cap A) + \overline{\mathcal{L}}(U \setminus A),$$

so that $\overline{\mathcal{L}}(U \setminus A) < \varepsilon$. So there is an open set $V \supseteq U \setminus A$ with $\mathcal{L}(V) < \varepsilon$. Then $U \setminus V$ is Lebesgue measurable, and $\mathcal{L}(U \setminus V) > \mathcal{L}(U) - \varepsilon$, so there is a closed set $F \subseteq U \setminus V \subseteq A$ with $\mathcal{L}(F) > \mathcal{L}(U) - \varepsilon$. Thus $F \subseteq A \subseteq U$ and $\mathcal{L}(U \setminus F) < \varepsilon$. Therefore A is Lebesgue measurable. □

Theorem 5.1.17. *Let $A \subseteq \mathbb{R}$ be Lebesgue measurable, and let a similarity $f\colon \mathbb{R} \to \mathbb{R}$ with ratio r be given. Then $f[A]$ is Lebesgue measurable and $\mathcal{L}(f[A]) = r\mathcal{L}(A)$.*

Proof. [Strictly speaking, "similarity" disallows $r = 0$, but even if $r = 0$ is allowed, this formula still works: If $r = 0$, then the range of f is a single point, so of course $f[A]$ is measurable and $\mathcal{L}(f[A]) = 0$.] Now suppose $r > 0$.

Consider an interval $I = [a, b)$. The image is an interval, either $[f(a), f(b))$ or $(f(b), f(a)]$, with length $|f(b) - f(a)| = r|b - a|$. Therefore $\overline{\mathcal{L}}(f[I]) = r|b - a|$. Now if $A \subseteq \bigcup_{j\in\mathbb{N}}[a_j, b_j)$, then $f[A] \subseteq \bigcup f\big[[a_j, b_j)\big]$, so $\overline{\mathcal{L}}(f[A]) \le \sum \overline{\mathcal{L}}\big(f\big[[a_j, b_j)\big]\big) = r\sum(b_j - a_j)$. Therefore we have $\overline{\mathcal{L}}(f[A]) \le r\overline{\mathcal{L}}(A)$. If we apply the same thing to the inverse map f^{-1}, which is a similarity with ratio $1/r$, we get $\overline{\mathcal{L}}(f[A]) \ge r\overline{\mathcal{L}}(A)$. Therefore $\overline{\mathcal{L}}(f[A]) = r\overline{\mathcal{L}}(A)$

Now f is a homeomorphism, so the image of an open set is open and the image of a closed set is closed. If $A \subseteq \mathbb{R}$ is measurable, then, for every $\varepsilon > 0$, there exist a closed set F and an open set U with $F \subseteq A \subseteq U$ and $\overline{\mathcal{L}}(U\setminus F) < \varepsilon$. So we have $f[F] \subseteq f[A] \subseteq f[U]$ and $\overline{\mathcal{L}}(f[U] \setminus f[F]) = \overline{\mathcal{L}}\big(f[U \setminus F]\big) < r\varepsilon$. So $f[A]$ is also measurable. □

Next is a preview of how measure theory is related to fractal dimension. In general, we do not yet know that the similarity dimension of a set is unique. However, we can now establish that in one situation we can determine the similarity dimension.

Exercise 5.1.18. Let $(r_1, r_2, \cdots, r_n)$ be a contracting ratio list; let s be its sim-value; let $(f_1, f_2, \cdots, f_n)$ be a corresponding iterated function system in $\mathbb{R}$; and let $A \subseteq \mathbb{R}$ be a nonempty measurable set. Suppose $\mathcal{L}(f_j[A] \cap f_k[A]) = 0$ for $j \ne k$, and

$$A = \bigcup_{j=1}^{n} f_j[A].$$

If $0 < \mathcal{L}(A) < \infty$, then $s = 1$.

Number Systems

Recall the situation from Sect. 1.6. Let b be a real number, and let D be a finite set of real numbers, including 0. We are interested in representing real numbers in the number system they define.

Write W for the set of "whole numbers"; that is, numbers of the form

$$\sum_{j=0}^{M} a_j b^j .$$

Write F for the set of "fractions"; that is numbers of the form

$$\sum_{j=-\infty}^{-1} a_j b^j .$$

We know (Proposition 3.2.21) that there is no number system that has a unique representation for every real number. So we will try to represent all real numbers, and arrange to have as few numbers as possible with more than one representation. One way to specify that the set with multiple representations is small is to require that it have Lebesgue measure 0.

If we analyze the size of the intersections $(F+w_1)\cap(F+w_2)$, $w_1, w_2 \in W$, $w_1 \neq w_2$, then we will know about all numbers with multiple representations:

Exercise 5.1.19. The set of all numbers with multiple representations is a countable union of sets, each of which is similar to one of the sets $(F+w_1)\cap(F+w_2)$, $w_1, w_2 \in W, w_1 \neq w_2$.

The set of all numbers that can be represented is $F+W$, a countable union of sets isometric to F. So if all real numbers can be represented, then $\mathcal{L}(F) > 0$. We know that F is a compact set, so also $\mathcal{L}(F) < \infty$. If the set of all real numbers with multiple representations has Lebesgue measure 0, then the sets $(F+w_1)\cap(F+w_2)$ have Lebesgue measure 0.

Suppose D has k elements. If F has positive Lebesgue measure, but the intersections $(F+w_1)\cap(F+w_2)$ have Lebesgue measure zero, then by Exercise 5.1.18, F has similarity dimension 1. But the similarity dimension is actually $\log k/\log|b|$. Therefore $|b| = k$.

5.2 Method I

We will need to discuss measures other than Lebesgue measure. The basics are contained in this section.

Measures and Outer Measures

A collection $\mathcal{F}$ of subsets of a set X is called an ***algebra*** on X iff:

(1) $\varnothing, X \in \mathcal{F}$;
(2) if $A \in \mathcal{F}$, then $X \setminus A \in \mathcal{F}$;
(3) if $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$.

Note that $A \cap B = X \setminus ((X \setminus A) \cup (X \setminus B))$ and $A \setminus B = A \cap (X \setminus B)$, so an algebra is also closed under these two operations.

A collection $\mathcal{F}$ of subsets of a set X is called a ***σ-algebra*** on X iff:

(1) $\varnothing, X \in \mathcal{F}$;
(2) if $A \in \mathcal{F}$, then $X \setminus A \in \mathcal{F}$;
(3) if $A_1, A_2, \dots \in \mathcal{F}$, then $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{F}$.

Of course (by Theorem 5.1.14), the collection of all Lebesgue measurable subsets of $\mathbb{R}$ is a σ-algebra on $\mathbb{R}$. Combining the clauses of the definition will produce a few more rules: For example, if $A_1, A_2, \dots \in \mathcal{F}$, then $\bigcap_{i\in\mathbb{N}} A_i \in \mathcal{F}$; if $A, B \in \mathcal{F}$ then $A \cap B, A \cup B, A \setminus B \in \mathcal{F}$.

Theorem 5.2.1. *Let X be a set, and let $\mathcal{D}$ be any set of subsets of X. Then there is a set $\mathcal{F}$ of subsets of X such that*

(1) *$\mathcal{F}$ is a σ-algebra on X;*
(2) *$\mathcal{F} \supseteq \mathcal{D}$;*
(3) *if $\mathcal{G}$ is any σ-algebra on X with $\mathcal{G} \supseteq \mathcal{D}$, then $\mathcal{G} \supseteq \mathcal{F}$.*

Proof. First I claim that the intersection of any collection of σ-algebras on X is a σ-algebra. Let Γ be a collection of σ-algebras, and let $\mathcal{B} = \bigcap_{\mathcal{A}\in\Gamma} \mathcal{A}$ be the intersection. Then $\varnothing \in \mathcal{A}$ for all $\mathcal{A} \in \Gamma$, so $\varnothing \in \mathcal{B}$. Similarly $X \in \mathcal{B}$. If $A \in \mathcal{B}$, then $A \in \mathcal{A}$ for all $\mathcal{A} \in \Gamma$, so $X \setminus A \in \mathcal{A}$ for all $\mathcal{A} \in \Gamma$, and therefore $X \setminus A \in \mathcal{B}$. If $A_1, A_2, \dots \in \mathcal{B}$, then each $A_n \in \mathcal{A}$ for all $\mathcal{A} \in \Gamma$, so $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ for all $\mathcal{A} \in \Gamma$ and therefore $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{B}$.

So suppose a set $\mathcal{D}$ of subsets of X is given. Let Γ be the collection of all σ-algebras $\mathcal{G}$ on X with $\mathcal{G} \supseteq \mathcal{D}$. (There is at least one such σ-algebra, namely the family of all subsets of X.) Then the intersection $\mathcal{F} = \bigcap_{\mathcal{G}\in\Gamma} \mathcal{G}$ is a σ-algebra on X. But clearly if $\mathcal{G}$ is any σ-algebra on X with $\mathcal{G} \supseteq \mathcal{D}$, then $\mathcal{G} \in \Gamma$, and therefore $\mathcal{G} \supseteq \mathcal{F}$. □

We say that $\mathcal{F}$ is the ***least σ-algebra*** containing $\mathcal{D}$, or the σ-algebra ***generated*** by $\mathcal{D}$. Let S be a metric space. A subset of S is called a ***Borel set*** iff it belongs to the σ-algebra on S generated by the open sets.

Let X be a set, and let $\mathcal{F}$ be a σ-algebra of subsets of X. A ***measure*** on $\mathcal{F}$ is a set function* $\mathcal{M}\colon \mathcal{F} \to [0, \infty]$ such that:

(1) $\mathcal{M}(\varnothing) = 0$;
(2) If $A_n \in \mathcal{F}$ is a disjoint sequence of sets, then

$$\mathcal{M}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \mathcal{M}(A_n).$$

We call (2) ***countable additivity***.

Let X be a set. An ***outer measure*** on X is a function $\overline{\mathcal{M}}$ defined on all subsets of X, with values in the nonnegative extended real numbers $[0, \infty]$, satisfying:

* A ***set function*** is a function whose domain is a family of sets.

(1) $\overline{\mathcal{M}}(\varnothing) = 0$;
(2) if $A \subseteq B$, then $\overline{\mathcal{M}}(A) \le \overline{\mathcal{M}}(B)$;
(3) $\overline{\mathcal{M}}\left(\bigcup_{n\in\mathbb{N}} A_n\right) \le \sum_{n=1}^{\infty} \overline{\mathcal{M}}(A_n)$.

We call (3) ***countable subadditivity***.

Defining Outer Measures

The Lebesgue outer measure was constructed in Sect. 5.1. The way in which the definition was formulated was not accidental. We will explore a general method for constructing outer measures, known as "method I". We begin with candidate values for the measures of some sets (like the lengths of the half-open intervals), and then attempt to produce an outer measure that is as large as possible, but no larger than the candidate values.

Let X be a set, and let $\mathcal{A}$ be a family of subsets of X that covers X. Let $\mathbf{c}\colon \mathcal{A} \to [0, \infty]$ be any function. The theorem on construction of outer measures is as follows:

Theorem 5.2.2 (Method I Theorem). *There is a unique outer measure $\overline{\mathcal{M}}$ on X such that*

(1) $\overline{\mathcal{M}}(A) \le \mathbf{c}(A)$ *for all* $A \in \mathcal{A}$*;*
(2) *if* $\overline{\mathcal{N}}$ *is any outer measure on* X *with* $\overline{\mathcal{N}}(A) \le \mathbf{c}(A)$ *for all* $A \in \mathcal{A}$*, then* $\overline{\mathcal{N}}(B) \le \overline{\mathcal{M}}(B)$ *for all* $B \subseteq X$*.*

Proof. The uniqueness is easy: if two outer measures satisfy (1) and (2), then each is $\le$ the other, so they are equal.

For any subset B of X, define

$$\overline{\mathcal{M}}(B) = \inf \sum_{A\in\mathcal{D}} \mathbf{c}(A), \tag{I}$$

where the infimum is over all countable covers $\mathcal{D}$ of B by sets of $\mathcal{A}$. (Recall that $\inf \varnothing = \infty$, so if there is no countable cover at all of B by sets of $\mathcal{A}$, then $\overline{\mathcal{M}}(B) = \infty$.)

I claim first that $\overline{\mathcal{M}}$ is an outer measure. First, $\overline{\mathcal{M}}(\varnothing) = 0$, since the empty set is covered by the empty cover, and the empty sum has value 0. If $B \subseteq C$, then any cover of C is also a cover of B, so $\overline{\mathcal{M}}(B) \le \overline{\mathcal{M}}(C)$. Let $B_1, B_2, \cdots$ be given. I must prove

$$\overline{\mathcal{M}}\left(\bigcup_{n\in\mathbb{N}} B_n\right) \le \sum_{n=1}^{\infty} \overline{\mathcal{M}}(B_n).$$

If $\overline{\mathcal{M}}(B_n) = \infty$ for some n, then the inequality is clear. So suppose $\overline{\mathcal{M}}(B_n) < \infty$ for all n. Let $\varepsilon > 0$. For each n, choose a countable cover $\mathcal{D}_n$ of B_n by sets of $\mathcal{A}$ with

$$\sum_{A\in\mathcal{D}_n} \mathbf{c}(A) \le \overline{\mathcal{M}}(B_n) + 2^{-n}\varepsilon.$$

Now $\mathcal{D} = \bigcup_{n\in\mathbb{N}} \mathcal{D}_n$ is a countable cover of the union $\bigcup_{n\in\mathbb{N}} B_n$. Therefore

$$\begin{aligned}\overline{\mathcal{M}}\left(\bigcup_{n\in\mathbb{N}} B_n\right) &\le \sum_{A\in\mathcal{D}} \mathbf{c}(A) \\ &\le \sum_{n=1}^{\infty} \sum_{A\in\mathcal{D}_n} \mathbf{c}(A) \\ &\le \sum_{n=1}^{\infty} \overline{\mathcal{M}}(B_n) + \sum_{n=1}^{\infty} 2^{-n}\varepsilon \\ &= \sum_{n=1}^{\infty} \overline{\mathcal{M}}(B_n) + \varepsilon.\end{aligned}$$

Since ε was any positive number, we have

$$\overline{\mathcal{M}}\left(\bigcup_{n\in\mathbb{N}} B_n\right) \le \sum_{n=1}^{\infty} \overline{\mathcal{M}}(B_n).$$

This completes the proof that $\overline{\mathcal{M}}$ is an outer measure.

Now we may check the two assertions of the theorem. For (1), note that for $A \in \mathcal{A}$, the singleton $\{A\}$ is a cover of A, so

$$\overline{\mathcal{M}}(A) \le \sum_{B\in\{A\}} \mathbf{c}(B) = \mathbf{c}(A).$$

For (2), suppose that $\overline{\mathcal{N}}$ is any outer measure on X with $\overline{\mathcal{N}}(A) \le \mathbf{c}(A)$ for all $A \in \mathcal{A}$. Then for any countable cover $\mathcal{D}$ of a set B by elements of $\mathcal{A}$ we have

$$\sum_{A\in\mathcal{D}} \mathbf{c}(A) \ge \sum_{A\in\mathcal{D}} \overline{\mathcal{N}}(A) \ge \overline{\mathcal{N}}\left(\bigcup_{A\in\mathcal{D}} A\right) \ge \overline{\mathcal{N}}(B).$$

Therefore $\overline{\mathcal{M}}(B) \ge \overline{\mathcal{N}}(B)$. □

When we say that an outer measure is to be constructed by ***method I***, we are referring to this theorem. In practical terms, this means that the outer measure is defined by the formula (I).

Reduced Cover Classes

When a measure is defined by method I, it may be helpful to know that the covers $\mathcal{D}$ in (I) can be chosen from a smaller ("reduced") class of sets.

Proposition 5.2.3. *Let X be a set, and let $\mathbf{c}$ be a set function. For a collection of sets $\mathcal{A}$, let $\overline{\mathcal{M}}_{\mathcal{A}}$ be the method I outer measure defined using the class $\mathcal{A}$ of sets and the restriction of $\mathbf{c}$ to $\mathcal{A}$.*

(a) *If* $\mathcal{B} \subseteq \mathcal{A}$*, then* $\overline{\mathcal{M}}_{\mathcal{A}} \le \overline{\mathcal{M}}_{\mathcal{B}}$.
(b) *Suppose that, for every* $A \in \mathcal{A}$ *and every* $\varepsilon > 0$*, there is* $B \in \mathcal{B}$ *with* $B \supseteq A$ *and* $\mathbf{c}(B) \le \mathbf{c}(A) + \varepsilon$*. Then* $\overline{\mathcal{M}}_{\mathcal{B}} \le \overline{\mathcal{M}}_{\mathcal{A}}$.
(c) *Let* $C > 0$ *be a constant, and suppose that, for every* $A \in \mathcal{A}$*, there is* $B \in \mathcal{B}$ *with* $B \supseteq A$ *and* $\mathbf{c}(B) \le C\,\mathbf{c}(A)$*. Then* $\overline{\mathcal{M}}_{\mathcal{B}} \le C\,\overline{\mathcal{M}}_{\mathcal{A}}$.

Proof. (a) The outer measure $\overline{\mathcal{M}}_{\mathcal{B}}$ is the largest outer measure such that $\overline{\mathcal{M}}_{\mathcal{B}}(E) \le \mathbf{c}(E)$ for all $E \in \mathcal{B}$. But $\overline{\mathcal{M}}_{\mathcal{A}}$ also has this property, so $\overline{\mathcal{M}}_{\mathcal{A}} \le \overline{\mathcal{M}}_{\mathcal{B}}$.

(b) Let $\varepsilon > 0$ be given. Let $\mathcal{D} = \{A_1, A_2, \cdots\} \subseteq \mathcal{A}$ be a countable cover of a set E. For each A_j, choose $B_j \in \mathcal{B}$ with $B_j \supseteq A_j$ and $\mathbf{c}(B_j) \le \mathbf{c}(A_j) + \varepsilon/2^j$. Then $\mathcal{D}' = \{B_1, B_2, \cdots\}$ is also a cover of E, and

$$\varepsilon + \sum_j \mathbf{c}(A_j) \ge \sum_j \mathbf{c}(B_j) \ge \overline{\mathcal{M}}_{\mathcal{B}}(E).$$

Take the infimum over all countable $\mathcal{D} \subseteq \mathcal{A}$ that cover E to get

$$\varepsilon + \overline{\mathcal{M}}_{\mathcal{A}}(E) \ge \overline{\mathcal{M}}_{\mathcal{B}}(E).$$

This is true for all $\varepsilon > 0$, so $\overline{\mathcal{M}}_{\mathcal{A}}(E) \ge \overline{\mathcal{M}}_{\mathcal{B}}(E)$.

(c) Let $\mathcal{D} = \{A_1, A_2, \cdots\} \subseteq \mathcal{A}$ be a countable cover of a set E. For each A_j, choose $B_j \in \mathcal{B}$ with $B_j \supseteq A_j$ and $\mathbf{c}(B_j) \le C\,\mathbf{c}(A_j)$. Then $\mathcal{D}' = \{B_1, B_2, \cdots\}$ is also a cover of E, and

$$C \sum_j \mathbf{c}(A_j) \ge \sum_j \mathbf{c}(B_j) \ge \overline{\mathcal{M}}_{\mathcal{B}}(E).$$

Take the infimum over all countable $\mathcal{D} \subseteq \mathcal{A}$ that cover E to get

$$C\,\overline{\mathcal{M}}_{\mathcal{A}}(E) \ge \overline{\mathcal{M}}_{\mathcal{B}}(E).$$ □

When condition (b) holds, we will say that $\mathcal{B}$ is a ***reduced cover class*** for $\overline{\mathcal{M}}$. When condition (c) holds, we will say that $\mathcal{B}$ is a reduced cover class with factor C for $\overline{\mathcal{M}}$.

Here is an example. Lebesgue measure $\mathcal{L}$ on $\mathbb{R}$ is defined (p. 140) using the class of all intervals $[a, b)$. The ***semi-dyadic intervals*** are sets of the form $[(k-1)/2^n, (k+1)/2^n)$, where $n \in \mathbb{Z}$, $k \in \mathbb{Z}$. The class of all semi-dyadic intervals is a reduced cover class with factor 4 for Lebesgue measure. Indeed, if $a < b$, let n be the integer with $2^{-n-1} < b - a < 2^{-n}$ and k the integer with $k - 1 \le a/2^n < k$, and compute $[k/2^n, (k+1)/2^n) \supseteq [b-a)$ and $(k+1)/2^n - (k-1)/2^n < 4(b-a)$.

The ***dyadic net*** is the class of intervals of the form $[k/2^n, (k+1)/2^n)$. It is not a reduced cover class by itself. Let $\mathcal{R}_n$ consist of the finite disjoint unions of dyadic intervals $[k/2^n, (k+1)/2^n)$ with denominator 2^n. The ***dyadic ring*** is $\mathcal{R} = \bigcup_n \mathcal{R}_n$.

Proposition 5.2.4. *Using the set function* $\mathbf{c} \colon \mathcal{R} \to [0, \infty)$ *defined by* $\mathbf{c}(E) = \mathcal{L}(E)$*, the dyadic ring* $\mathcal{R}$ *is a reduced cover class for Lebesgue measure.*

Proof. Let $a < b$ and let $\varepsilon > 0$. Let $n \in \mathbb{N}$ be so large that $2^{-n} < \varepsilon/2$. Let $j \in \mathbb{Z}$ be such that $j \le 2^n a < j+1$, and $m \ge j$ such that $m \le 2^n b < m+1$. Then

$$[a,b) \subseteq E := \bigcup_{k=j}^{m} \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right),$$

and

$$\mathcal{L}(E) - (b-a) \le \mathcal{L}\left(\left[\frac{j}{2^n}, a\right)\right) + \mathcal{L}\left(\left[b, \frac{m+1}{2^n}\right)\right) \le \frac{2}{2^n} < \varepsilon. \qquad \square$$

Measurable Sets

Let $\overline{\mathcal{M}}$ be an outer measure on a set X. A set $A \subseteq X$ is $\overline{\mathcal{M}}$***-measurable*** (in the sense of Carathéodory) iff $\overline{\mathcal{M}}(E) = \overline{\mathcal{M}}(E \cap A) + \overline{\mathcal{M}}(E \setminus A)$ for all sets $E \subseteq X$.

Theorem 5.2.5. *The collection $\mathcal{F}$ of $\overline{\mathcal{M}}$-measurable sets is a σ-algebra, and $\overline{\mathcal{M}}$ is countably additive on $\mathcal{F}$.*

Proof. First, $\varnothing \in \mathcal{F}$ since for any E, we have $\overline{\mathcal{M}}(E \cap \varnothing) + \overline{\mathcal{M}}(E \setminus \varnothing) = \overline{\mathcal{M}}(\varnothing) + \overline{\mathcal{M}}(E) = \overline{\mathcal{M}}(E)$. It is also easy to see that a set A belongs to $\mathcal{F}$ if and only if its complement $X \setminus A$ does.

Suppose $A_j \in \mathcal{F}$ for $j = 1, 2, \cdots$. Let E be any test set. Then

$$\begin{aligned}
\overline{\mathcal{M}}(E) &= \overline{\mathcal{M}}(E \cap A_1) + \overline{\mathcal{M}}(E \setminus A_1) \\
&= \overline{\mathcal{M}}(E \cap A_1) + \overline{\mathcal{M}}((E \setminus A_1) \cap A_2) + \overline{\mathcal{M}}(E \setminus (A_1 \cup A_2)) \\
&= \cdots \\
&= \sum_{j=1}^{k} \overline{\mathcal{M}}\left(\left(E \setminus \bigcup_{i=1}^{j-1} A_i\right) \cap A_j\right) + \overline{\mathcal{M}}\left(E \setminus \bigcup_{j=1}^{k} A_j\right).
\end{aligned}$$

Hence

$$\overline{\mathcal{M}}(E) \ge \sum_{j=1}^{k} \overline{\mathcal{M}}\left(\left(E \setminus \bigcup_{i=1}^{j-1} A_i\right) \cap A_j\right) + \overline{\mathcal{M}}\left(E \setminus \bigcup_{j \in \mathbb{N}} A_j\right),$$

so (let $k \to \infty$)

$$\overline{\mathcal{M}}(E) \ge \sum_{j=1}^{\infty} \overline{\mathcal{M}}\left(\left(E \setminus \bigcup_{i=1}^{j-1} A_i\right) \cap A_j\right) + \overline{\mathcal{M}}\left(E \setminus \bigcup_{j \in \mathbb{N}} A_j\right).$$

But

$$E \cap \bigcup_{j \in \mathbb{N}} A_j = \bigcup_{j \in \mathbb{N}} \left(\left(E \setminus \bigcup_{i=1}^{j-1} A_i\right) \cap A_j\right),$$

so

$$\begin{aligned}\overline{\mathcal{M}}(E) &\le \overline{\mathcal{M}}\left(E \cap \bigcup_{j\in\mathbb{N}} A_j\right) + \overline{\mathcal{M}}\left(E \setminus \bigcup_{j\in\mathbb{N}} A_j\right)\\ &\le \sum_{j=1}^{\infty} \overline{\mathcal{M}}\left(\left(E \setminus \bigcup_{i=1}^{j-1} A_i\right) \cap A_j\right) + \overline{\mathcal{M}}\left(E \setminus \bigcup_{j\in\mathbb{N}} A_j\right)\\ &\le \overline{\mathcal{M}}(E).\end{aligned}$$

Thus $\bigcup A_j \in \mathcal{F}$. This completes the proof that $\mathcal{F}$ is a σ-algebra.

Now if the sets $A_j \in \mathcal{F}$ are disjoint, we can let $E = \bigcup A_j$ in the previous computation, and we get

$$\overline{\mathcal{M}}\left(\bigcup_{j\in\mathbb{N}} A_j\right) = \sum_{j=1}^{\infty} \overline{\mathcal{M}}(A_j),$$

so $\overline{\mathcal{M}}$ is countably additive on $\mathcal{F}$. □

We will write simply $\mathcal{M}$ for the restriction of $\overline{\mathcal{M}}$ to the σ-algebra $\mathcal{F}$ of measurable sets. It is a measure on $\mathcal{F}$. Thus we see that we have a generalization of Lebesgue measure as constructed in Sect. 5.1.

Corollary 5.2.6. *Every Borel set in $\mathbb{R}$ is Lebesgue measurable.*

Proof. Open sets are measurable by Theorem 5.1.12. The collection of measurable sets is a σ-algebra by Theorem 5.1.14. □

5.3 Two-Dimensional Lebesgue Measure

We next define two-dimensional Lebesgue measure. This is a measure defined for subsets of the plane $\mathbb{R}^2$.

A ***rectangle*** in $\mathbb{R}^2$ is a set R of the form

$$R = [a,b) \times [c,d) = \left\{ (x,y) \in \mathbb{R}^2 : a \le x < b, c \le y < d \right\}$$

for some $a \le b$ and $c \le d$. The ***area*** of this rectangle is $\mathbf{c}(R) = (b-a)(d-c)$, as usual. In particular, if $a = b$ or $c = d$, we see that $\mathbf{c}(\varnothing) = 0$. ***Two-dimensional Lebesgue outer measure*** is the outer measure $\overline{\mathcal{L}}^2$ on $\mathbb{R}^2$ defined by method I from this function $\mathbf{c}$.

Theorem 5.3.1. *Two-dimensional Lebesgue outer measure $\overline{\mathcal{L}}^2$ is a metric outer measure.*

Proof. Suppose A and B are sets with positive separation. Since $\overline{\mathcal{L}}^2$ is an outer measure, we have $\overline{\mathcal{L}}^2(A \cup B) \le \overline{\mathcal{L}}^2(A) + \overline{\mathcal{L}}^2(B)$. So I must prove the opposite inequality.

Let $\mathcal{D}$ be a cover of $A \cup B$ by rectangles. Now a rectangle $R = [a, b) \times [c, d)$ can be written as a union of the four rectangles

$$
\begin{gathered}
[a, (a+b)/2) \times [c, (c+d)/2) \\
[a, (a+b)/2) \times [(c+d)/2, d) \\
[(a+b)/2, b) \times [c, (c+d)/2) \\
[(a+b)/2, b) \times [(c+d)/2, d)\,,
\end{gathered}
$$

and the area of the large rectangle is the sum of the areas of the four small rectangles. So the sum

$$\sum_{R \in \mathcal{D}} \mathbf{c}(R)$$

is unchanged when we replace one of the rectangles by its four parts. Applying this repeatedly, we may assume that the diameters of the rectangles in $\mathcal{D}$ are all smaller than $\operatorname{dist}(A, B)$. Then no rectangle of $\mathcal{D}$ intersects both A and B. So $\mathcal{D}$ is a disjoint union of two families, $\mathcal{A}$ and $\mathcal{B}$, where $\mathcal{A}$ covers A and $\mathcal{B}$ covers B. But then

$$\sum_{R \in \mathcal{D}} \mathbf{c}(R) = \sum_{R \in \mathcal{A}} \mathbf{c}(R) + \sum_{R \in \mathcal{B}} \mathbf{c}(R) \ge \overline{\mathcal{L}}^2(A) + \overline{\mathcal{L}}^2(B).$$

So we conclude that $\overline{\mathcal{L}}^2(A \cup B) \ge \overline{\mathcal{L}}^2(A) + \overline{\mathcal{L}}^2(B)$. □

The sets that are measurable in the sense of Carathéodory for $\overline{\mathcal{L}}^2$ are again called the ***Lebesgue measurable sets***; the restriction of $\overline{\mathcal{L}}^2$ to this σ-algebra is called ***two-dimensional Lebesgue measure***. We will write $\mathcal{L}^2$ for two-dimensional Lebesgue measure.

The fact that two-dimensional Lebesgue measure is not identically zero is left to you:

Exercise 5.3.2. If $a < b$ and $c < d$, then a rectangle of the form $R = [a, b) \times [c, d)$ is Lebesgue measurable and $\mathcal{L}^2(R) = (b - a)(d - c)$.

Now that we know that the Lebesgue measure of a square is what it should be, the usual scheme of approximating an area with a lot of little squares will show that the usual sets of Euclidean plane geometry have two-dimensional Lebesgue measure equal to their areas. In particular, a rectangle with sides not parallel to the coordinate axes has the right area. This should be enough to prove:

Exercise 5.3.3. Let $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ be a similarity with ratio r. If $A \subseteq \mathbb{R}^2$ is Lebesgue measurable, then so is $f[A]$, and $\mathcal{L}^2(f[A]) = r^2\, \mathcal{L}^2(A)$.

This will tell us something about the similarity dimension of a set $A \subseteq \mathbb{R}^2$, as in the one-dimensional case (Exercise 5.1.18).

Exercise 5.3.4. Let $(r_1, r_2, \cdots, r_n)$ be a contracting ratio list; let s be its sim-value; let $(f_1, f_2, \cdots f_n)$ be a corresponding iterated function system in $\mathbb{R}^2$; and let $A \subseteq \mathbb{R}^2$ be a nonempty Borel set. Suppose $\mathcal{L}^2(f_j[A] \cap f_k[A]) = 0$ for $j \neq k$, and $A = \bigcup_{j=1}^n f_j[A]$. If $0 < \mathcal{L}^2(A) < \infty$, then $s = 2$.

This result can be used for complex number systems in the same way as the corresponding result was used for real number systems in 5.1.18

Exercise 5.3.5. Let b be a complex number, and let D be a finite set of complex numbers, including 0. Suppose D has k elements. Suppose every complex number can be represented in the number system defined by base b and digit set D, and the set of complex numbers with multiple representations has two-dimensional Lebesgue measure 0. What does this mean about the relationship between b and k?

Higher Dimensions

Let d be a positive integer. In d-dimensional Euclidean space $\mathbb{R}^d$, we will consider ***hyper-rectangles*** of the form

$$R = [a_1, b_1) \times [a_2, b_2) \times \cdots \times [a_d, b_d),$$

where $a_j < b_j$ for all j. The "hyper-volume" of this hyper-rectangle R is

$$\mathbf{c}(R) = \prod_{j=1}^{d} (b_j - a_j).$$

We define ***d-dimensional Lebesgue outer measure*** to be the method I outer measure defined from this set function $\mathbf{c}$. We define ***d-dimensional Lebesgue measure*** to be the restriction to the measurable subsets. As before, we use the notation $\overline{\mathcal{L}}^d$ and $\mathcal{L}^d$.

Exercise 5.3.6. The outer measure $\overline{\mathcal{L}}^d$ is a metric outer measure. If

$$R = [a_1, b_1) \times [a_2, b_2) \times \cdots \times [a_d, b_d),$$

where $a_j \leq b_j$ for all j, then

$$\mathcal{L}^d(R) = \prod_{j=1}^{d} (b_j - a_j).$$

Dyadic Cubes

A ***semi-dyadic square*** in $\mathbb{R}^2$ is a set of the form

$$\left[\frac{j-1}{2^n}, \frac{j+1}{2^n}\right) \times \left[\frac{k-1}{2^n}, \frac{k+1}{2^n}\right).$$

Exercise 5.3.7. Show that the class of semi-dyadic squares is a reduced cover class with factor 8 for $\mathcal{L}^2$.

A ***dyadic square*** in $\mathbb{R}^2$ is a set of the form

$$\left[\frac{j}{2^n}, \frac{j+1}{2^n}\right) \times \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right).$$

For each n, let $\mathcal{R}_n$ be the set of finite disjoint unions of dyadic squares with denominator 2^n. Then $\mathcal{R} = \bigcup_n \mathcal{R}_n$ is called the ***dyadic ring*** in $\mathbb{R}^2$.

Exercise 5.3.8. Using the set function $\mathbf{c}\colon \mathcal{R} \to [0, \infty)$ defined by $\mathbf{c}(E) = \mathcal{L}^2(E)$, the dyadic ring $\mathcal{R}$ is a reduced cover class for 2-dimensional Lebesgue measure.

After you have completed the preceding two exercises, it should be easy to formulate the corresponding results for d-dimensional Lebesgue measure for any $d \in \mathbb{N}$. You would define semi-dyadic cubes, dyadic cubes, and the dyadic ring in $\mathbb{R}^d$.

5.4 Metric Outer Measure

Consider the following example of a method I outer measure on $\mathbb{R}$; the definition is very close to the definition used for Lebesgue measure. We begin with the collection $\mathcal{A} = \{\, [a,b) : a < b \,\}$ of half-open intervals and the set function $\mathbf{c}\big([a,b)\big) = \sqrt{b-a}$. Let $\overline{\mathcal{M}}$ be the corresponding method I outer measure. I claim that the interval $A = [0,1]$ is not measurable.

Consider the measure of $[0,1)$. Certainly the singleton $\big\{[0,1)\big\}$ covers $[0,1)$, so $\overline{\mathcal{M}}\big([0,1)\big) \le \mathbf{c}\big([0,1)\big) = 1$. If $[0,1) \subseteq \bigcup_{i\in\mathbb{N}}[a_i,b_i)$, then by what we know about Lebesgue measure, we must have $\sum_{i=1}^{\infty}(b_i - a_i) \ge 1$. So we have also

$$\begin{aligned}\left(\sum_{i=1}^{\infty}\sqrt{b_i-a_i}\right)^2 &= \sum_{i=1}^{\infty}\left(\sqrt{b_i-a_i}\right)^2 + 2\sum_{i<j}\sqrt{b_i-a_i}\,\sqrt{b_j-a_j} \\ &\ge \sum_{i=1}^{\infty}(b_i-a_i) \ge 1.\end{aligned}$$

Therefore $\sum_{i=1}^{\infty}\sqrt{b_i-a_i} \ge 1$. This shows that $\overline{\mathcal{M}}\big([0,1)\big) \ge 1$. So we have $\overline{\mathcal{M}}\big([0,1)\big) = 1$.

Similarly $\overline{\mathcal{M}}\big([-1,0)\big) = 1$. The singleton $\big\{[-1,1)\big\}$ covers $[-1,1)$, so as before we have $\overline{\mathcal{M}}\big([-1,1)\big) \le \mathbf{c}\big([-1,1)\big) = \sqrt{2}$. So if $A = [0,1]$ and $E = [-1,1)$, we have

$$\overline{\mathcal{M}}(E \cap A) + \overline{\mathcal{M}}(E \setminus A) = 1 + 1 = 2 > \sqrt{2} \ge \overline{\mathcal{M}}(E).$$

This shows that $A = [0,1]$ is not measurable.

It is often desirable that the sets we work with be measurable sets. When we work with subsets of a metric space (as is common in this book), the sets are often open sets, closed sets, or sets constructed simply from open and closed sets. In particular, the sets are often Borel sets. There is a condition that will insure that all Borel sets are measurable sets.

Two sets A, B in a metric space have ***positive separation*** iff $\operatorname{dist}(A,B) > 0$; that is, there is $r > 0$ with $\varrho(x,y) \ge r$ for all $x \in A$ and $y \in B$. Let $\overline{\mathcal{M}}$ be an outer measure on a metric space S. We say that $\overline{\mathcal{M}}$ is a ***metric outer measure*** iff $\overline{\mathcal{M}}(A \cup B) = \overline{\mathcal{M}}(A) + \overline{\mathcal{M}}(B)$ for any pair A, B of sets with positive separation. (Theorem 5.1.6 shows that $\overline{\mathcal{L}}$ is a metric outer measure.) The measure $\mathcal{M}$ obtained by restricting a metric outer measure $\overline{\mathcal{M}}$ to its measurable sets will be called a ***metric measure***.

The reason that metric outer measures are of interest is that open sets (and therefore all Borel sets) are measurable sets. Before I prove this, I formulate the lemma of Carathéodory.

Lemma 5.4.1. *Let $\overline{\mathcal{M}}$ be a metric outer measure on the metric space S. Let $A_1 \subseteq A_2 \subseteq \cdots$, and $A = \bigcup_{j\in\mathbb{N}} A_j$. Assume $\operatorname{dist}(A_j, A \setminus A_{j+1}) > 0$ for all j. Then $\overline{\mathcal{M}}(A) = \lim_{j\to\infty} \overline{\mathcal{M}}(A_j)$.*

Proof. For all j we have $\overline{\mathcal{M}}(A) \ge \overline{\mathcal{M}}(A_j)$, so $\overline{\mathcal{M}}(A) \ge \lim_{j\to\infty} \overline{\mathcal{M}}(A_j)$. (This inequality is true for any outer measure.) If $\lim_{j\to\infty} \overline{\mathcal{M}}(A_j) = \infty$, then the equation is true. So suppose $\lim_{j\to\infty} \overline{\mathcal{M}}(A_j) < \infty$.

Let $B_1 = A_1$ and $B_j = A_j \setminus A_{j-1}$ for $j \ge 2$. If $i \ge j+2$, then $B_j \subseteq A_j$ and $B_i \subseteq A \setminus A_{i-1} \subseteq A \setminus A_{j+1}$, so B_i and B_j have positive separation. So

$$\overline{\mathcal{M}}\left(\bigcup_{k=1}^{m} B_{2k-1}\right) = \sum_{k=1}^{m} \overline{\mathcal{M}}(B_{2k-1})$$
$$\overline{\mathcal{M}}\left(\bigcup_{k=1}^{m} B_{2k}\right) = \sum_{k=1}^{m} \overline{\mathcal{M}}(B_{2k}).$$

Since $\lim_{j\to\infty} \overline{\mathcal{M}}(A_j) < \infty$, both of these converge (as $m \to \infty$). So

$$\begin{aligned}\overline{\mathcal{M}}(A) = \overline{\mathcal{M}}\left(\bigcup_{j\in\mathbb{N}} A_j\right) &= \overline{\mathcal{M}}\left(A_j \cup \bigcup_{k\ge j+1} B_k\right)\\ &\le \overline{\mathcal{M}}(A_j) + \sum_{k=j+1}^{\infty} \overline{\mathcal{M}}(B_k)\end{aligned}$$

$$\leq \lim_{i\to\infty} \overline{\mathcal{M}}(A_i) + \sum_{k=j+1}^{\infty} \overline{\mathcal{M}}(B_k).$$

Now as $j \to \infty$, the tail of a convergent series goes to 0, so we get

$$\overline{\mathcal{M}}(A) \leq \lim_{i\to\infty} \overline{\mathcal{M}}(A_i). \qquad \square$$

Theorem 5.4.2. *Let $\overline{\mathcal{M}}$ be a metric outer measure on a metric space S. Then every Borel subset of S is $\overline{\mathcal{M}}$-measurable.*

Proof. Since the σ-algebra of Borel sets is the σ-algebra generated by the closed sets, and since the collection $\mathcal{F}$ of measurable sets is a σ-algebra, it is enough to show that every closed set F is measurable. Let A be any test set. I must show that $\overline{\mathcal{M}}(A) \geq \overline{\mathcal{M}}(A \cap F) + \overline{\mathcal{M}}(A \setminus F)$, since the opposite inequality is true for any outer measure.

Let $A_j = \{\, x \in A : \operatorname{dist}(x, F) \geq 1/j \,\}$. Then $\operatorname{dist}(A_j, F \cap A) \geq 1/j$, so

$$\overline{\mathcal{M}}(A \cap F) + \overline{\mathcal{M}}(A_j) = \overline{\mathcal{M}}\big((A \cap F) \cup A_j\big) \leq \overline{\mathcal{M}}(A). \tag{1}$$

Now since F is closed, F contains all points of distance 0 from F, so $A \setminus F = \bigcup_{j\in\mathbb{N}} A_j$. We check the condition of the lemma: If $x \in \big(A \setminus (F \cup A_{j+1})\big)$, then there exists $z \in F$ with $\varrho(x, z) < 1/(j + 1)$. If $y \in A_j$, then

$$\varrho(x, y) \geq \varrho(y, z) - \varrho(x, z) > \frac{1}{j} - \frac{1}{j+1}.$$

Thus

$$\operatorname{dist}\big(A \setminus (F \cup A_{j+1}), A_j\big) \geq \frac{1}{j} - \frac{1}{j+1} > 0.$$

Therefore, applying the lemma, we get $\overline{\mathcal{M}}(A \setminus F) \leq \lim_{j\to\infty} \overline{\mathcal{M}}(A_j)$. Taking the limit in (1), we get $\overline{\mathcal{M}}(A \cap F) + \overline{\mathcal{M}}(A \setminus F) \leq \overline{\mathcal{M}}(A)$, which completes the proof. $\square$

Proposition 5.4.3. *Let $\mathcal{M}$ be a finite metric measure on a compact metric space S. Let $E \subseteq S$ be a Borel set. For any $\varepsilon > 0$, there exist a compact set K and an open set U with $U \supseteq E \supseteq K$ and $\mathcal{M}(U \setminus K) < \varepsilon$.*

Proof. Let $\mathcal{A}$ be the collection of all sets $E \subseteq S$ such that for all $\varepsilon > 0$, there exist a compact set K and an open set U with $U \supseteq E \supseteq K$ and $\mathcal{M}(U \setminus K) < \varepsilon$.

First I claim that all closed sets belong to $\mathcal{A}$. Let $F \subseteq S$ be closed and $\varepsilon > 0$. Now

$$U_n = \left\{\, x \in S : \operatorname{dist}(x, F) < \frac{1}{n} \,\right\}$$

defines open sets U_n with $U_1 \supseteq U_2 \supseteq \cdots$ and $\bigcap_{n\in\mathbb{N}} U_n = F$. So we have $\lim_{n\to\infty} \mathcal{M}(U_n) = \mathcal{M}(F)$. There is n so large that $\mathcal{M}(U_n) - \mathcal{M}(F) < \varepsilon$. Then: $U_n \supseteq F \supseteq F$, U_n is open, F is compact, and $\mathcal{M}(U_n \setminus F) < \varepsilon$.

Clearly $\varnothing \in \mathcal{A}$.

Next, $\mathcal{A}$ is closed under complements. Let $E \in \mathcal{A}$. Consider the complement $E' = S \setminus E$. Let $\varepsilon > 0$. There is an open U and a compact K with $U \supseteq E \supseteq K$ and $\mathcal{M}(U \setminus K) < \varepsilon$. But then $U' = S \setminus U$ is compact, $K' = S \setminus K$ is open, $K' \supseteq E' \supseteq U'$, and $\mathcal{M}(K' \setminus U') = \mathcal{M}(U \setminus K) < \varepsilon$.

Now I claim $\mathcal{A}$ is closed under countable unions. Let $E_n \in \mathcal{A}$ for $n \in \mathbb{N}$. Write $E = \bigcup_{n\in\mathbb{N}} E_n$, and let $\varepsilon > 0$. Then there exist open U_n and compact K_n with $U_n \supseteq E_n \supseteq K_n$ and $\mathcal{M}(U_n \setminus K_n) < \varepsilon/2^{n+1}$. Then $U = \bigcup_{n\in\mathbb{N}} U_n$ is open. Now

$$L_m = \bigcup_{n=1}^{m} K_n$$

is compact, increases with m, and $\bigcup_{m\in\mathbb{N}} L_m = \bigcup_{n\in\mathbb{N}} K_n$. There is m so large that

$$\mathcal{M}\left(\bigcup_{n\in\mathbb{N}} K_n\right) - \mathcal{M}(L_m) < \frac{\varepsilon}{2}.$$

So we have $U \supseteq E \supseteq L_n$ and

$$\begin{aligned} \mathcal{M}(U \setminus L_n) &\le \mathcal{M}\left(\bigcup_{n\in\mathbb{N}} (U_n \setminus K_n)\right) + \mathcal{M}\left(\bigcup_{n\in\mathbb{N}} K_n\right) - \mathcal{M}(L_m) \\ &< \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{n+1}} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Therefore $\mathcal{A}$ includes at least the Borel sets. $\square$

Method II

We have seen that method I may fail to yield a measure where open sets are measurable. There is a related construction, called "method II" that will overcome this difficulty.

Let $\mathcal{A}$ be a family of subsets of a metric space S, and suppose, for every $x \in S$ and $\varepsilon > 0$, there exists $A \in \mathcal{A}$ with $x \in A$ and $\operatorname{diam} A \le \varepsilon$. Suppose $\mathbf{c}\colon \mathcal{A} \to [0, \infty]$ is a given function. An outer measure will be constructed based on this data. For each $\varepsilon > 0$, let

$$\mathcal{A}_\varepsilon = \{\, A \in \mathcal{A} : \operatorname{diam} A \le \varepsilon \,\}.$$

Let $\overline{\mathcal{M}}_\varepsilon$ be the method I outer measure determined by $\mathbf{c}$ using the family $\mathcal{A}_\varepsilon$. Then by Proposition 5.2.3(a), for a given set E, when ε decreases, $\overline{\mathcal{M}}_\varepsilon(E)$ increases. Define

$$\overline{\mathcal{M}}(E) = \lim_{\varepsilon\to 0} \overline{\mathcal{M}}_\varepsilon(E) = \sup_{\varepsilon>0} \overline{\mathcal{M}}_\varepsilon(E).$$

It is easily verified that $\overline{\mathcal{M}}$ is an outer measure. As usual, we will write $\mathcal{M}$ for the restriction to the measurable sets. This construction of an outer measure

$\overline{\mathcal{M}}$ from a set function $\mathbf{c}$ (and a measure $\mathcal{M}$ from $\overline{\mathcal{M}}$) is called ***method II***. It is more complicated than method I, but (unlike method I) it insures that Borel sets are measurable:

Theorem 5.4.4. *The set function $\overline{\mathcal{M}}$ defined by method II is a metric outer measure.*

Proof. Let $A, B \subseteq S$ with $\operatorname{dist}(A, B) > 0$. Since $\overline{\mathcal{M}}$ is an outer measure, we have $\overline{\mathcal{M}}(A \cup B) \leq \overline{\mathcal{M}}(A) + \overline{\mathcal{M}}(B)$. So I must prove the opposite inequality.

Let $\varepsilon > 0$ so small that $\varepsilon < \operatorname{dist}(A, B)$. Let $\mathcal{D}$ be any countable cover of $A \cup B$ by sets of $\mathcal{A}_\varepsilon$. The sets $D \in \mathcal{D}$ have diameter less than $\operatorname{dist}(A, B)$, so such a set D intersects at most one of the sets A, B. Therefore, $\mathcal{D}$ may be divided into two disjoint collections, $\mathcal{D}_1$ and $\mathcal{D}_2$, where $\mathcal{D}_1$ covers A and $\mathcal{D}_2$ covers B. Then

$$\sum_{D \in \mathcal{D}} \mathbf{c}(D) = \sum_{D \in \mathcal{D}_1} \mathbf{c}(D) + \sum_{D \in \mathcal{D}_2} \mathbf{c}(D) \geq \overline{\mathcal{M}}_\varepsilon(A) + \overline{\mathcal{M}}_\varepsilon(B).$$

Now we may take the infimum over all covers, and conclude $\overline{\mathcal{M}}_\varepsilon(A \cup B) \geq \overline{\mathcal{M}}_\varepsilon(A) + \overline{\mathcal{M}}_\varepsilon(B)$. Then we may take the limit as $\varepsilon \to 0$ to conclude $\overline{\mathcal{M}}(A \cup B) \geq \overline{\mathcal{M}}(A) + \overline{\mathcal{M}}(B)$. ☐

Exercise 5.4.5. Let S be a metric space, and let $\mathbf{c}$ be a set function. For a collection of sets $\mathcal{A}$, let $\overline{\mathcal{M}}_\mathcal{A}$ be the method II outer measure defined using the class $\mathcal{A}$ of sets and the restriction of $\mathbf{c}$ to $\mathcal{A}$.

(a) If $\mathcal{B} \subseteq \mathcal{A}$, then $\overline{\mathcal{M}}_\mathcal{A} \leq \overline{\mathcal{M}}_\mathcal{B}$.
(b) Suppose that, for every $\eta > 0$ there is $\delta > 0$ such that for all $A \in \mathcal{A}$ with $\operatorname{diam} A \leq \delta$, and every $\varepsilon > 0$, there is $B \in \mathcal{B}$ with $\operatorname{diam} B \leq \eta$, $B \supseteq A$, and $\mathbf{c}(B) \leq \mathbf{c}(A) + \varepsilon$. Then $\overline{\mathcal{M}}_\mathcal{B} \leq \overline{\mathcal{M}}_\mathcal{A}$.
(c) Let $C > 0$ be a constant, and suppose that, for every $\eta > 0$ there is $\delta > 0$ such that for all $A \in \mathcal{A}$ with $\operatorname{diam} A \leq \delta$, there is $B \in \mathcal{B}$ with $\operatorname{diam} B \leq \eta$, $B \supseteq A$, and $\mathbf{c}(B) \leq C\, \mathbf{c}(A)$. Then $\overline{\mathcal{M}}_\mathcal{B} \leq C\, \overline{\mathcal{M}}_\mathcal{A}$.

When condition (b) holds, we will say that $\mathcal{B}$ is a ***reduced cover class*** for $\overline{\mathcal{M}}$. When condition (c) holds, we will say that $\mathcal{B}$ is a reduced cover class with factor C for $\overline{\mathcal{M}}$.

5.5 Measures for Strings

One of the useful ways we will employ the material of this chapter is by defining measures. Sometimes (for example Lebesgue measure or Hausdorff measure) we will define a measure on subsets of Euclidean space $\mathbb{R}^d$. But also measures will be defined on our string models and path models.

An Example

We will consider an easy example before we attack the more general case.

Begin with the two-letter alphabet $E = \{0, 1\}$. Consider, as usual, the metric space $E^{(\omega)}$ of infinite strings with metric $\varrho_{1/2}$. We will construct a measure on $E^{(\omega)}$. We begin with the family of "basic open sets":

$$\mathcal{A} = \left\{ [\alpha] : \alpha \in E^{(*)} \right\}$$

together with the set function $\mathbf{c}\colon \mathcal{A} \to [0, \infty)$ defined by

$$\mathbf{c}([\alpha]) = \frac{1}{2^{|\alpha|}}.$$

(Recall the notation $|\alpha|$ for the length of the string α.)

Proposition 5.5.1. *The method I outer measure $\overline{\mathcal{M}}_{1/2}$ constructed using this function $\mathbf{c}$ is a metric outer measure and satisfies $\overline{\mathcal{M}}_{1/2}([\alpha]) = \mathbf{c}([\alpha])$ for all $\alpha \in E^{(*)}$.*

Proof. Write $\mathcal{A} = \left\{ [\alpha] : \alpha \in E^{(*)} \right\}$, $\mathcal{A}_\varepsilon = \{ D \in \mathcal{A} : \operatorname{diam} D \le \varepsilon \}$. Let $\overline{\mathcal{N}}_\varepsilon$ be the method I measure defined by the set function $\mathbf{c}$ restricted to $\mathcal{A}_\varepsilon$. If $D \in \mathcal{A}_\varepsilon$, then of course $\mathbf{c}(D) \ge \overline{\mathcal{M}}_{1/2}(D)$, so by the Method I theorem,

$$\overline{\mathcal{N}}_\varepsilon(A) \ge \overline{\mathcal{M}}_{1/2}(A)$$

for all A. Therefore the method II measure $\overline{\mathcal{N}}$ defined by

$$\overline{\mathcal{N}}(A) = \lim_{\varepsilon \to 0} \overline{\mathcal{N}}_\varepsilon(A)$$

satisfies $\overline{\mathcal{N}}(A) \ge \overline{\mathcal{M}}_{1/2}(A)$.

Note that, for any $\alpha \in E^{(*)}$, if k is the length $|\alpha|$, then we have

$$\mathbf{c}([\alpha]) = \frac{1}{2^k} = \frac{1}{2^{k+1}} + \frac{1}{2^{k+1}} = \mathbf{c}([\alpha 0]) + \mathbf{c}([\alpha 1]).$$

Applying this repeatedly, we see that, for any $\varepsilon > 0$, any set $D \in \mathcal{A}$ is a finite disjoint union $D_1 \cup D_2 \cup \cdots \cup D_n$ of sets in $\mathcal{A}_\varepsilon$ with $\mathbf{c}(D) = \sum \mathbf{c}(D_i)$. This means that $\overline{\mathcal{N}}_\varepsilon(D) \le \mathbf{c}(D)$, so by the Method I theorem, $\overline{\mathcal{N}}_\varepsilon(A) \le \overline{\mathcal{M}}_{1/2}(A)$ for all sets A. So $\overline{\mathcal{N}}(A) \le \overline{\mathcal{M}}_{1/2}(A)$

Therefore $\overline{\mathcal{M}}_{1/2} = \overline{\mathcal{N}}$ is a method II outer measure, so it is a metric outer measure. ☐

Exercise 5.5.2. Let $h\colon E^{(\omega)} \to \mathbb{R}$ be the "base 2" addressing function defined on p. 14. If $A \subseteq E^{(\omega)}$ is a Borel set, then $\mathcal{M}_{1/2}(A) = \mathcal{L}(h[A])$.

Measures on String Spaces

Let E be a finite alphabet with at least two letters. Consider the space $E^{(\omega)}$ of infinite strings. This space is a metric space for many different metrics ϱ. But all of the metrics constructed according to the scheme in Proposition 2.6.5 produce the same open sets. A countable base for the open sets is $\{ [\alpha] : \alpha \in E^{(*)} \}$. Also $E^{(\omega)}$ is a compact ultrametric space. An important feature of all these metrics is $\lim_{k\to\infty} \operatorname{diam}[\sigma\restriction k] = 0$ for each $\sigma \in E^{(\omega)}$.

Exercise 5.5.3. It follows that $\lim_{k\to\infty} \left(\sup \{ \operatorname{diam}[\alpha] : \alpha \in E^{(k)} \}\right) = 0$.

Suppose a non-negative number w_α is given for each finite string α. Under what conditions is there a metric outer measure $\overline{\mathcal{M}}$ on $E^{(\omega)}$ with $\overline{\mathcal{M}}([\alpha]) = w_\alpha$ for all α? Since it is a metric outer measure, the open sets $[\alpha]$ are measurable, so $\overline{\mathcal{M}}$ is additive on them. Now the set $[\alpha]$ is the disjoint union of the sets $[\beta]$ as β ranges over the children of α (that is, $\beta = \alpha e$ for $e \in E$).

Theorem 5.5.4. *Suppose the non-negative numbers w_α satisfy*

$$w_\alpha = \sum_{e\in E} w_{\alpha e}$$

for $\alpha \in E^{(*)}$. Then the method I outer measure defined by the set function $\mathbf{c}([\alpha]) = w_\alpha$ is a metric outer measure $\overline{\mathcal{M}}$ on $E^{(\omega)}$ with $\overline{\mathcal{M}}([\alpha]) = w_\alpha$.

Proof. Write $\mathcal{A} = \{ [\alpha] : \alpha \in E^{(*)} \}$, $\mathcal{A}_\varepsilon = \{ D \in \mathcal{A} : \operatorname{diam} D \le \varepsilon \}$. Let $\overline{\mathcal{N}}_\varepsilon$ be the method I measure defined by the set function $\mathbf{c}$ restricted to $\mathcal{A}_\varepsilon$. If $D \in \mathcal{A}_\varepsilon$, then of course $\mathbf{c}(D) \ge \overline{\mathcal{M}}(D)$, so by the Method I theorem,

$$\overline{\mathcal{N}}_\varepsilon(A) \ge \overline{\mathcal{M}}(A)$$

for all A. Therefore the method II measure $\overline{\mathcal{N}}$ defined by

$$\overline{\mathcal{N}}(A) = \lim_{\varepsilon\to 0} \overline{\mathcal{N}}_\varepsilon(A)$$

satisfies $\overline{\mathcal{N}}(A) \ge \overline{\mathcal{M}}(A)$.

Note that, for any $\alpha \in E^{(*)}$, we have

$$\mathbf{c}([\alpha]) = w_\alpha = \sum_{e\in E} w_{\alpha e} = \sum_{e\in E} \mathbf{c}\big([\alpha e]\big).$$

Applying this repeatedly, together with Exercise 5.5.3, we see that, for any $\varepsilon > 0$, any set $D \in \mathcal{A}$ is a finite disjoint union $D_1 \cup D_2 \cup \cdots \cup D_n$ of sets in $\mathcal{A}_\varepsilon$ with $\mathbf{c}(D) = \sum \mathbf{c}(D_i)$. This means that $\overline{\mathcal{N}}_\varepsilon(D) \le \mathbf{c}(D)$, so by the Method I theorem, $\overline{\mathcal{N}}_\varepsilon(A) \le \overline{\mathcal{M}}(A)$ for all sets A. So $\overline{\mathcal{N}}(A) \le \overline{\mathcal{M}}(A)$.

Therefore $\overline{\mathcal{M}} = \overline{\mathcal{N}}$ is a method II outer measure. So we may conclude that it is a metric outer measure. □

How should we formulate the corresponding theorem for the path spaces $E_v^{(\omega)}$ defined by a directed multigraph (V, E, i, t)? We will define measures for each of the spaces $E_v^{(\omega)}$. We only need to define one of them at a time.

Fix a vertex v. Suppose nonnegative numbers w_α are given, one for each $\alpha \in E_v^{(*)}$. They should (of course) satisfy

$$w_\alpha = \sum_{i(e)=t(\alpha)} w_{\alpha e}$$

for $\alpha \in E_v^{(*)}$. Note that this has consequences for the troublesome exceptional cases that came up when we were defining the metric. If α has no children, then (interpreting an empty sum as 0), we see that $w_\alpha = 0$. Similarly, if α has only one child β, then $w_\alpha = w_\beta$.

Exercise 5.5.5. Suppose the non-negative numbers w_α satisfy

$$w_\alpha = \sum_{i(e)=t(\alpha)} w_{\alpha e}$$

for $\alpha \in E_v^{(*)}$. Then the method I outer measure defined by the set function $\mathbf{c}([\alpha]) = w_\alpha$ is a metric outer measure $\overline{\mathcal{M}}$ on $E_v^{(\omega)}$ with $\overline{\mathcal{M}}([\alpha]) = w_\alpha$.

5.6 *Remarks

Henri Lebesgue's measure and integration theory dates from about 1900. It is one of the cornerstones of twentieth century mathematics. I have not discussed integration at all, in order to reduce the amount of material to its bare minimum. (My more advanced text [18] develops integration using the same ideas.) The more abstract measure theory was developed by many others, such as Constantin Carathéodory, during the early 1900's. Theorem 5.5.4 on the existence of measures on the string spaces is due essentially to A. N. Kolmogorov.

The σ-algebra $\mathcal{F}$ generated by a family $\mathcal{D}$ of sets (as in Theorem 5.2.1) is a complicated object to describe constructively. The proof given for 5.2.1 has the advantage of not requiring such a constructive description. Certainly $\mathcal{F}$ contains all countable unions

$$\bigcup_{i \in \mathbb{N}} D_i$$

with $D_i \in \mathcal{D}$; it contains complements of those unions; it contains countable intersections

$$\bigcap_{i \in \mathbb{N}} E_i$$

where each E_i is either a countable union or a complement of a countable union. But that may not be everything in $\mathcal{F}$. (See, for example, the proof of Theorem (10.23) in [34].)

An example of a set in the line not measurable for Lebesgue measure may be found in many texts. For example: [7, pp. 36–37], [11, Theorem 1.4.7], [34, (10.28)], or [58, Chap. 3, Sect. 4].

Exercise 5.3.5: $k = |b|^2$.

Exercise 5.5.2. Both measures are method I measures; use the Method I theorem twice, once to prove an inequality in each direction.

O, wiste a man how manye maladyes
Folwen of excesse and of glotonyes
He wolde been the moore mesurable
—G. Chaucer, *The Pardoner's Tale*