Chapter 3

Lebesgue and Hausdorff Measures

In the first two chapters we have studied measures, outer measures and Borel measures in a general manner without much attention to particular cases. In this chapter we examine two principal Borel measures—Lebesgue and Hausdorff measures. In Section 1 we define the Lebesgue measure as an outer measure and subsequently show that it coincides with the measure obtained by applying the Riesz representation theorem to the Riemann integral. Various properties of this measure are discussed. In Section 2 we focus on the Lebesgue measure on the real line. First we review non-measurable sets and the Cantor set. Next we use them to illustrate various delicate properties of the Lebesgue measure, especially its incompatibility with the topology on $\mathbb{R}$. In Section 3 we prove the fundamental Brunn-Minkowski inequality. Hausdorff measures and Hausdorff dimension are introduced in Section 4. A main result in this section is a discussion on the equivalence between the n-dimensional Hausdorff measure and the n-dimensional Lebesgue measure on $\mathbb{R}^n$. In the last section we study how Hausdorff measures change under Lipschitz maps, and as an application, we determine the Hausdorff dimension of the Cantor set.

3.1 The Lebesgue Measure

Let $R = \Pi_j [a_j, b_j]$ be a closed rectangle in $\mathbb{R}^n$ and $|R| = \Pi_j (b_j - a_j)$ be its volume. It is a closed cube when $b_j - a_j$ are all equal for $j = 1, 2 \cdots, n$. In this chapter rectangles are referred to those that are parallel to the coordinate planes. We define the n-dimensional *Lebesgue measure* $\mathcal{L}^n$ on $\mathbb{R}^n$ by

$$\mathcal{L}^n(E) = \inf \left\{ \sum_{j=1}^{\infty} |R_j| : E \subset \bigcup_{j=1}^{\infty} R_j, \ R_j \text{ closed cubes} \right\}.$$

In other words, observe $(\mathcal{R}, |\cdot|)$ where $\mathcal{R}$ is the collection of all closed cubes in $\mathbb{R}^n$ and $|R|$ the volume of the cube R forms a gauge on $\mathbb{R}^n$. The Lebesgue measure is the outer measure resulting from this gauge. It is not hard to see that we could also use

$$\inf \left\{ \sum_1^\infty |R_j| : E \subset \bigcup_1^\infty R_j,\ R_j \text{ are open cubes} \right\}$$

to define $\mathcal{L}^n(E)$. For each $\delta > 0$, the Lebesgue measure is also given by

$$\inf \left\{ \sum_1^\infty |R_j| : E \subset \bigcup_1^\infty R_j,\ R_j \text{ are closed cubes of diameter less than } \delta \right\} .$$

I leave these as exercise.

As already covered in previous exercises, we record the following basic facts on the Lebesgue measure.

- The Lebesgue measure $\mathcal{L}^n$ is an outer measure whose measurable sets include the Borel σ-algebra.
- $\mathcal{L}^n(R) = |R|$ for any cube R.
- Every set is outer regular, that is, for any $E \subset \mathbb{R}^n$, $\mathcal{L}^n(E) = \inf \{\mathcal{L}^n(G) : E \subset G,\ G \text{ open }\}$. Every Lebesgue measurable set is inner regular, that is, for any measurable E, $\mathcal{L}^n(E) = \inf \{\mathcal{L}^n(K) :\ K \subset E \text{ compact }\}$.
- The Lebesgue σ-algebra is the completion of the Borel σ-algebra.
- The Lebesgue measure is translational invariant, that is, $\mathcal{L}^n(E + x) = \mathcal{L}^n(E)$, for every $E \subset \mathbb{R}^n$.

In the exercise we know how the translational invariant property characterizes the Lebesgue measure.

The Riesz representation suggests another way of defining the Lebesgue measure. We could use the Riemann integral to play the role of the positive linear functional in the theorem. The Riemann integral on $[a, b]$ is discussed in elementary analysis, and the extension to $\mathbb{R}^n$ for continuous functions of compact support is routine.

Let P be a partition of $\mathbb{R}^n$: $\cdots < x^j_{-1} < x^j_0 < x^j_1 < x^j_2 < \cdots$, $j = 1, \ldots, n$, and $\|P\| = \max \left\{x^j_k - x^j_{k-1} : j = 1, \ldots, n,\ k \in \mathbb{Z}\right\}$. For a bounded function f with compact support in $\mathbb{R}^n$, define its Riemann sum of f w.r.t. the tagged partition $\dot{P}$ to be

$$R(f, \dot{P}) = \sum_J^\infty f(z_J)\Delta x_J,$$

where z_J is a tag point in $R_J = [x^1_{k-1}, x^1_k] \times \cdots \times [x^n_{l-1}, x^n_l]$, $J = (k, \cdots, l)$, and

$$\Delta x_J = (x^1_k - x^1_{k-1}) \times \cdots \times (x^n_l - x^n_{l-1}), \quad J \in \mathbb{Z}^n.$$

Note that the summands in the Riemann sum are zero except for finitely many terms. The Darboux upper and lower sums are given by

$$\overline{R}(f, P) = \sum_J \sup_{R_J} f \, \Delta x_J, \qquad \text{and}$$
$$\underline{R}(f, P) = \sum_J \inf_{R_J} f \, \Delta x_J.$$

As in the one dimensional case, f is called *Riemann integrable* if there exists $L \in \mathbb{R}$ such that for every $\varepsilon > 0$, there is some δ such that

$$\left| R(f, \dot{P}) - L \right| < \varepsilon, \quad \forall P, \ \|P\| < \delta.$$

We will use the notation

$$L = \int f \, dx$$

to denote the Riemann integral of f. The same as in the one dimensional case, it can be shown that f is Riemann integrable if and only if

$$\lim_{n\to\infty} \overline{R}(f, P_n) = \lim_{n\to\infty} \underline{R}(f, P_n)$$

for some P_n, $\|P_n\| \to 0$. We have

- Every f in $C_c(X)$ is Riemann integrable.
- $\displaystyle\int (\alpha f + \beta g) \, dx = \alpha \int f \, dx + \beta \int g \, dx, \qquad \forall f, g \in C_c(\mathbb{R}^n), \ \alpha, \beta \in \mathbb{R}.$
- $\displaystyle\int f \, dx \geq 0$ if $f \geq 0, f \in C_c(\mathbb{R})$.

In view of these,

$$\Lambda_R f = \int f \, dx$$

defines a positive linear functional on $C_c(\mathbb{R}^n)$. By Theorem 2.7, there exists a Borel (outer) measure μ_R such that

$$\int f \, dx = \int f \, d\mu_R, \quad \forall f \in C_c(X).$$

As $\mathbb{R}^n$ can be written as a countable union of compact sets and μ_R is finite on compact sets, $\mathbb{R}^n$ is σ-finite with respect to μ_R. By Proposition 2.8, μ_R enjoys the same regularity properties as the Lebesgue measure. In fact, we have

Proposition 3.1. *The measure μ_R coincides with $\mathcal{L}^n$.*

Proof. According to the properties of the Riesz measure, μ_R is an outer Borel measure where all sets are outer regular. Furthermore, let G be an open set and $f < G, f \in C_c(\mathbb{R}^n)$. For a fixed $x_0 \in \mathbb{R}^n$, the function $g(x) = f(x + x_0$ satisfies $g < G + x_0$. By looking at the Riemann sums and then passing to limit, we see that

$$\int_G f dx = \int_{G+x_0} g dx.$$

Using the characterization

$$\mu_R(G) = \sup\left\{\int_G f dx : \ f < G\right\},$$

and a similar one for $\mu_R(G+x_0)$ we deduce that $\mu_R(G+x_0) = \mu_R(G)$ for all open sets. By outer regularity, it is also true on all sets, so μ_R is translational invariant. By a problem on the characaterization of translational invariant measures in the exercise, μ_R is equal to a constant multiple of the Lebesgue measure on Borel sets and hence on all sets by outer regularity. To complete the proof it remains to show that the constant is equal to 1. For this purpose, let R_1 be the open unit cube $(0,1)^n$ and $R_\varepsilon = (\varepsilon, 1-\varepsilon)^n$ for small $\varepsilon > 0$. For every $f, \overline{R_\varepsilon} < f < R_1$, we have

$$\begin{aligned}
\mu_R(R_1) &\geq \int f d\mu_R \\
&= \int f dx \\
&\geq |\overline{R_\varepsilon}| \quad \text{(by the definition of the Riemann integral)} \\
&= (1-2\varepsilon)^n,
\end{aligned}$$

which implies $\mu_R(R_1) \geq 1$ after letting $\varepsilon \downarrow 0$. On the other hand, by considering cubes $(-\varepsilon, 1+\varepsilon)^n$ enclosing R_1, a similar argument shows that $\mu_R(R_1) \leq 1$. It follows that $\mu_R(R_1) = 1$. As $\mathcal{L}^n(R_1) = |R_1| = 1$, we conclude that the constant is equal to 1. □

Every Euclidean motion, or rigid motion, on $\mathbb{R}^n$ is a finite composition of translations, reflections and rotations. The Lebesgue measure has a nice scaling property which implies that it is invariant under all Euclidean motions.

Proposition 3.2. *Let T be a linear transformation from $\mathbb{R}^n$ to itself. For every measurable set E, TE is measurable and*

$$\mathcal{L}^n(TE) = \Delta(T)\mathcal{L}^n(E),$$

where $\Delta(T)$ is a nonnegative constant depending only T and it is equal to 1 when T is a rotation or reflection.

Proof. Assume that T is a nonsingular linear transformation and so it has a continuous inverse T^{-1}. For $E \subset \mathbb{R}^n$, set

$$\mu(E) = \mathcal{L}^n(TE).$$

It is easy to see that μ is an outer measure. Using

$$|x - y| \leq \|T^{-1}\||Tx - Ty|,$$

we see that $d(TA, TB) > 0$ whenever $A, B \subset \mathbb{R}^n$ with $d(A, B) > 0$, . It follows from Caratheodory's criterion that μ is a Borel measure. Its translational invariance is clear from definition and $0 < \mu(B) < \infty$ on any ball B. According to a problem in the exercise concerning the characterization of translational invariant measures, μ is a constant multiple of the Lebesgue measure.

When T is singular, it maps $\mathbb{R}^n$ into a subspace of less dimension and hence of measure zero. It follows that TE is a null set for all E. The measure $\mu(E) = \mathcal{L}^n(TE)$ is always equal to 0, so we can take $\Delta(T)$ to be 0. When T is nonsingular, ΔT is positive.

To show that $\Delta(T) = 1$ for a rotation or a reflection it suffices to take E to be a ball B centered at the origin, which is unchanged under a rotation or a reflection. Therefore, $\Delta(T) = \mu(B)/\mathcal{L}^n(B) = \mathcal{L}^n(TB)/\mathcal{L}^n(B) = 1$. □

For a general linear transformation T, $\Delta(T)$ is in fact the absolute value of the determinant of the matrix representation of T, see [R].

In the paragraphs above we developed Riemann integral as a tool to define a positive linear functional on $C_c(\mathbb{R}^n)$. In fact, one can mimic the one dimensional situation to define a concept of Riemann integrability for bounded functions with compact support in $\mathbb{R}^n$. We will not give the rather straightforward details, but simply point out that Lebesgue's theorem characterizing Riemann integrability holds in all dimensional, namely, a bounded, Lebesgue measurable function with compact support is Riemann integrable if and only if its set of discontinuity forms a null set in $\mathbb{R}^n$. When this happens, its Riemann integral is equal to its Lebsegue integral. Note that a Riemann integrable function can be approximated by simple functions and hence must be Lebesgue integrable.

3.2 Lebesgue Measure on $\mathbb{R}$

Next we discuss some properties of the Lebesgue measure on the real line. Corresponding results can be established in higher dimensions by using Fubini's theorem. Since Fubini's theorem will not be discussed until Chapter 8, we will focus on the one dimensional case. Many properties have been examined in an undergraduate real analysis. In view of this, we will be sketchy in some of the discussion below.

There are two basic constructions on the real line.

First, non-measurable sets. Introduce a relation $\sim$ on $\mathbb{R}$ by $x \sim y$ if and only if $x - y \in \mathbb{Q}$. "$\sim$" is easily seen to be an equivalence relation. Then $\mathbb{R} = \bigcup_\alpha \mathcal{E}_\alpha$, $\mathcal{E}_\alpha$ its equivalence classes. By the axiom of choice, we pick x_α from $\mathcal{E}_\alpha$ to form $\mathcal{E} = \{x_\alpha\}$. Depending on which points you pick, there are many different $\mathcal{E}$. By construction, we have

$$\mathbb{R} = \bigcup_{q \in \mathbb{Q}} (\mathcal{E} + q)$$

where $\mathcal{E} + q$ are disjoint for different q's.

Proposition 3.3. *Every set in $\mathbb{R}$ with positive measure contains a non-measurable subset.*

Proof. (Following [R]) It is equivalent to proving: Let A be a set whose subsets are all measurable. Then $\mathcal{L}^1(A) = 0$.

Let $A_q = A \bigcap (\mathcal{E} + q)$. As $\mathcal{L}^1(A) = \sum_q^\infty \mathcal{L}^1(A_q)$, it suffices to show $\mathcal{L}^1(A_q) = 0$ for each q.

By our assumption every A_q is measurable, for every compact $K \subset A_q$, we will show that $\mathcal{L}^1(K) = 0$, then $\mathcal{L}^1(A_k) = 0$ by the inner regularity of the Lebesgue measure. By countable additivity $\mathcal{L}^1(A) = \sum_q \mathcal{L}^1(A_q) = 0$. Let $H = \bigcup_j K + r$, $r \in \mathbb{Q} \cap [0, 1]$. One can check that it is a disjoint union. As K is bounded, H is a bounded set. Then

$$\begin{aligned}
\infty > \mathcal{L}^1(H) &= \sum_j \mathcal{L}^1(K + r) \\
&= \mathcal{L}^1(K) \sum_j 1, \quad \text{(translational invariance)}
\end{aligned}$$

which forces $\mathcal{L}^1(K) = 0$. $\square$

The second construction is the Cantor set. It can be described as follows: Let $I_k^n, 1 \le k \le 2^{n-1}, n \ge 1$, be open intervals obtained from the removal of a repeating trisection procedure, for instance, $I_1^1 = (1/3, 2/3), I_1^2 = (1/3^2, 2/3^2), I_2^2 = (7/3^2, 8/3^2)$. Then $\mathcal{C}_n = [0, 1] \setminus \bigcup_{k,m \le n} I_k^m$ is compact and so is the Cantor set $\mathcal{C} = \bigcap_{n=1}^\infty \mathcal{C}_n$. It can be shown that

- C is an uncountable set.

In fact, $x \in \mathcal{C}$ if and only if $x = 0.a_1a_2a_3\cdots$, $a_j \in \{0, 2\}$ in its ternary representation. The Cantor set has the following topological properties

- $\mathcal{C}$ is compact in $[0,1]$.
- $\mathcal{C}$ is a nowhere dense set.
- $\mathcal{C}$ is a perfect set.

Recall that a set in a topological space is nowhere dense if its closure does not contain any nonempty, open set. Since the Cantor set C is closed, it means C does not contain any non-empty, open interval, but this is evident from its construction. A set E is called perfect if for every $x \in E$ and open set V containing x, $V \cap E$ contains infinitely many elements. For the Cantor set this is also clear from its construction.

We also recall the measure-theoretic property:

- $\mathcal{C}$ is a null set.

It is possible to modify the construction of the Cantor set to obtain a Cantor-like set which satisfies all properties of $\mathcal{C}$ listed above except the last one. In other words, these Cantor-like sets are of positive measure, see exercise.

We use these two constructions to illustrate various points.

First of all, the existence of non-measurable sets shows that the inclusion of the Lebesgue σ-algebra $\mathcal{M}_C$ in $\mathcal{P}_{\mathbb{R}}$ is proper. How about the inclusion of $\mathcal{B}$ in $\mathcal{M}_C$? We know that for each $E \in \mathcal{M}_C$, there exists $E_1, E_2 \in \mathcal{B}$ such that $E_1 \subset E \subset E_2$ and $\mathcal{L}^1(E) = \mathcal{L}^1(E_1) = \mathcal{L}^1(E_2)$. But this does not answer our question. In fact, the Borel σ-algebra, by definition the smallest σ-algebra containing all open intervals, is not easy to described explicitly. Using G_δ and F_σ to denote respectively the collections of all countable intersection of open sets and countable union of closed sets, one can show that sets such as half-open and half-closed intervals are G_δ and F_σ sets. We may continue this process to construction, in self-evident notation, $G_{\delta\sigma}, F_{\sigma\delta}, G_{\delta\sigma\delta}, F_{\sigma\delta\sigma}, \cdots$, many many other Borel sets. But to get all Borel sets we need to go beyond countably many steps. Nevertheless, by transfinite induction (see section 10, chapter 3 in [HS]), one can show that the cardinal number of the Borel σ-algebra on $\mathbb{R}$ is equal to $\mathfrak{c}$, the same as the cardinal of $\mathbb{R}$. On the other hand, the Cantor set is of measure zero, so all its subsets are also of measure zero. As the Lebesgue measure is complete, all these subsets are measurable. As the cardinal number of $\mathcal{C}$ is $\mathfrak{c}$, all subsets of $\mathcal{C}$ have the cardinal number $2^{\mathfrak{c}}$. So $|\mathcal{M}_C| \geq |\text{subsets of } \mathcal{C}| = 2^{\mathfrak{c}} > \mathfrak{c} = |\mathcal{B}|$. So there are many more Lebesgue measurable sets than Borel measurable sets.

Next, the unit interval $[0,1]$ carries both a topological structure and a measure-theoretic structure, namely the Lebesgue measure restricted to this interval. The measure of size in a measure space is easy to describe. For instance, a set is small if its measure is zero and is large or full if its measure is equal to one. However, there are various ways to describe largeness in a topological space X. For instance, a set $E \subset X$ is dense if every n'd of any point in X contains some point in E. It is nowhere dense if its closure does not contain open set. It is easy to see that a set of full measure must be dense, but a dense set can be null (for example, the set of all rational numbers in $[0,1]$). Also, one can produce an open dense of in $[0,1]$ whose measure is equal to any prescribed number in $(0,1)$. Moreover, by modifying the construction of the Cantor set slightly to obtain a Cantor-like set which is closed, nowhere dense with measure equal to any number in $(0,1)$.

A more useful topological concept for size is the category. A set E is of the first category if it can be written as a countable union of nowhere dense sets. It is of the second category if it is not of the first category. Baire theorem states that a complete metric space is of the second category. A set of the first category is regarded small in topological setting. However, category and Lebesgue measure are still not compatible. Let us display a set S of full measure in $[0,1]$ which is of first category. Indeed, let C_j be Cantor-like sets of measure $1-1/k$ for each $k \geq 1$. Setting $S = \bigcup_k C_k$, we have $\mathcal{L}^1(S) \geq \mathcal{L}^1(C_j)$ and hence $\mathcal{L}^1(S) = 1$. But, as each C_k is nowhere dense, S is of the first category. On the other, the Cantor set C is closed, so it is a complete metric space under the Euclidean metric. By Baire category theorem, it is of the second category. However, its measure is equal to zero.

The incompatibility between the topological and Lebesgue measure-theoretic properties on the real line can be further reflected from the behavior of mappings between intervals.

First of all, we construct the Cantor function g from the Cantor set. Define, for each $n \geq 1$, $g_n : [0,1] \to [0,1]$ to a continuous, piecewise linear function that satisfies $g_n(0) = 0$, $g_n(x) = (2k-1)/2^m$, $x \in I_k^m$, $m \leq n$, $k = 1, 2, \ldots, 2^{m-1}$, and $g_n(1) = 1$. From the definition of g_n it is not hard to see that $|g_{n+1} - g_n| < 1/2^{n+1}$. By Weierstrass M-test, the series $\sum_{n=1}^\infty h_n$, $h_n = g_{n+1} - g_n$, converges uniformly to some continuous h. It follows that $g = \lim_{n\to\infty} g_n = \lim_{n\to\infty} \sum_{k=1}^n h_k + g_1 = h + g_1$ is also continuous. The function g is called the *Cantor function*. It is increasing, continuous and constant on each $I_k^n, k, n \geq 1$.

The Cantor function maps $\mathcal{C}$, a set of measure zero, to a set of full measure. For, first we have $g([0,1]) = [0,1]$. On the other hand, $\bigcup_{n,k} g(I_k^n) = \{1/2, 1/4, 3/4, \ldots\}$ is a countable set, so $\mathcal{L}^1(g(\mathcal{C})) = 1$. (In fact, one has $g(\mathcal{C}) = [0,1]$ since the endpoints of each I_k^n belong to the Cantor set.) So a continuous map could send a null set to a set of full measure.

We modify g to $h(x) = x + g(x) : [0,1] \to [0,2]$. Then h is strictly increasing and maps $[0,1]$ onto $[0,2]$. Its inverse h^{-1} is a continuous function from $[0,2]$ to $[0,1]$, so h is a homeomorphism between $[0,1]$ and $[0,2]$. Since g is constant on each I_k^n and $\bigcup_{n,k} I_k^n$ has measure 1, it is not hard to see $h(\mathcal{C})$ has measure equal to 1. By Proposition 3.3, there exists some non-measurable $A \subset h(\mathcal{C})$. But then

$h^{-1}(A) \subset \mathcal{C}$ is measurable. We have shown that a non-measurable set can be mapped to a measurable set by a homeomorphism.

3.3 Brunn-Minkowski Inequality

The Euclidean space carries not only a topological structure induced by the Euclidean metric but also a vector space. The translational invariance of the Lebesgue measure reflects the interaction between the measure-theoretic and algebraic properties of the Euclidean space. The Brunn-Minkowski inequality is an inequality of this nature.

Proposition 3.4.

(a) If A and B are open in $\mathbb{R}^n$, $A+B$ is open.

(b) If A and B are compact in $\mathbb{R}^n$, $A+B$ is compact.

(c) If A and B are closed in $\mathbb{R}^n$, $A+B$ is an F_δ-set.

Proof. (a) As

$$A+B = \bigcup_{b\in B} A+b, \quad A+b = \{a+b : a \in A\},$$

$A+B$ is open if we can show that each $A+b$ is open, since the union of open sets is still open. Consider the continuous map $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ given by $\varphi(x) = x+b$ whose inverse given by $\varphi^{-1}(x) = x-b$ is also continuous. so φ is an open map, i.e., it maps open sets to open sets. In particular, $A+b$ is open whenever A is open.

(b) Consider the continuous function $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ given by $\Phi(x,y) = x+y$. When A and B are compact, $A \times B$ is compact in $\mathbb{R}^n \times \mathbb{R}^n$. (This is a general fact in topology, or you can check it directly). So $A+B = \Phi(A,B)$ is compact. (This is again a general fact from topology; the image of any compact set under a continuous map is compact.)

(c) Since every closed set can be expressed as a countable union of compact sets in $\mathbb{R}^n$, (c) follows from (b).

□

Examples show that $A+B$ may not be measurable when A and B are measurable, see exercise in the Chapter 8.

Theorem 3.5 (**Brunn-Minkowski Inequality**). *Let A, B, and $A+B$ be measurable in $\mathbb{R}^n$. Then*

$$\mathcal{L}^n(A+B)^{\frac{1}{n}} \geq \mathcal{L}^n(A)^{\frac{1}{n}} + \mathcal{L}^n(B)^{\frac{1}{n}}.$$

One also has certain characterization on the equality case. For instance, equality sign holds if A and B are convex and B is homothetic to A. But this is not necessary.

Proof. We will divide the proof in five steps. The second step is decisive.
Step 1. Let A and B be two rectangles, A having sides $a_1, \ldots, a_n$ and B sides $b_1, \ldots, b_n$. Then $A + B$ is again a rectangle of sides $a_1 + b_1, \ldots, a_n + b_n$. In this special case Brunn-Minkowski inequality becomes

$$\Big(\prod_j (a_j + b_j)\Big)^{\frac{1}{n}} \geq \Big(\prod_j a_j\Big)^{\frac{1}{n}} + \Big(\prod_j b_j\Big)^{\frac{1}{n}}.$$

Divide both sides to get

$$1 \geq \Big(\prod_j \frac{a_j}{a_j + b_j}\Big)^{\frac{1}{n}} + \Big(\prod_j \frac{b_j}{a_j + b_j}\Big)^{\frac{1}{n}}.$$

By AM-GM inequality,

$$\begin{aligned} \Big(\prod_j \frac{a_j}{a_j + b_j}\Big)^{\frac{1}{n}} &\leq \frac{1}{n} \sum_j \frac{a_j}{a_j + b_j}, \qquad \text{and} \\ \Big(\prod_j \frac{b_j}{a_j + b_j}\Big)^{\frac{1}{n}} &\leq \frac{1}{n} \sum_j \frac{b_j}{a_j + b_j}. \end{aligned}$$

Therefore,

$$\Big(\prod_j \frac{a_j}{a_j + b_j}\Big)^{\frac{1}{n}} + \Big(\prod_j \frac{b_j}{a_j + b_j}\Big)^{\frac{1}{n}} \leq \frac{1}{n} \sum_j \frac{a_j}{a_j + b_j} + \frac{1}{n} \sum_j \frac{b_j}{a_j + b_j} = 1,$$

done.
Step 2. Let A and B both consist of finitely many almost disjoint rectangles. (Almost disjoint means their interiors are mutually disjoint.) We use induction on the total number of rectangles in A and B. When $N = 2$, it is done in Step 1. Now, assume it has been proved for $N \geq 2$ and we prove it for $N + 1$ rectangles.

Pick two rectangles in A (or B) we can find a coordinate hyperplane to separate their interiors. WLOG we may assume the hyperplane is $\{(x_1, \ldots, x_n) : x_n = 0\}$. We let

$$\begin{aligned} A_+ &= \{R \cap H_+ : R \text{ is a rectangle in } A\}, \\ A_- &= \{R \cap H_- : R \text{ is a rectangle in } A\}, \end{aligned}$$

where $H_+ = \{x = (x_1, \ldots, x_n) : x_n \geq 0\}$ is the upper half space and $H_- = \{x : x_n \leq 0\}$ the lower half space. Then A_+ and A_- have at most $m - 1$ many rectangles where m is the number of rectangles in A. Now we translate B up and down so that

$$\frac{\mathcal{L}^n(A_+)}{\mathcal{L}^n(A)} = \frac{\mathcal{L}^n(B_+)}{\mathcal{L}^n(B)} \quad \left(\text{so } \frac{\mathcal{L}^n(A_-)}{\mathcal{L}^n(A)} = \frac{\mathcal{L}^n(B_-)}{\mathcal{L}^n(B)}\right).$$

Here

$$\begin{aligned} B_+ &= \{R \cap H_+ : R \text{ is a rectangle in } B\}, \\ B_- &= \{R \cap H_- : R \text{ is a rectangle in } B\}, \end{aligned}$$

where B is the B after translation. Note that translations of A and B do not change the values of $\mathcal{L}^n(A)$, $\mathcal{L}^n(B)$ and $\mathcal{L}^n(A+B)$.

Now, each of A_+ and B_+ has at most N many rectangles, so have A_- and B_-. By induction hypothesis,

$$\begin{aligned} \mathcal{L}^n(A_+ + B_+) &\geq \left(\mathcal{L}^n(A_+)^{\frac{1}{n}} + \mathcal{L}^n(B_+)^{\frac{1}{n}}\right)^n \\ \mathcal{L}^n(A_- + B_-) &\geq \left(\mathcal{L}^n(A_-)^{\frac{1}{n}} + \mathcal{L}^n(B_-)^{\frac{1}{n}}\right)^n. \end{aligned}$$

As the interiors of $(A_+ + B_+)$ and $(A_- + B_-)$ are disjoint (they belong to the upper and lower half-spaces) and $(A_+ + B_+) \cup (A_- + B_-) \subset A + B$, we have

$$\begin{aligned} \mathcal{L}^n(A+B) &\geq \mathcal{L}^n(A_+ + B_+) + \mathcal{L}^n(A_- + B_-) \\ &\geq \left(\mathcal{L}^n(A_+)^{\frac{1}{n}} + \mathcal{L}^n(B_+)^{\frac{1}{n}}\right)^n + \left(\mathcal{L}^n(A_-)^{\frac{1}{n}} + \mathcal{L}^n(B_-)^{\frac{1}{n}}\right)^n \\ &= \mathcal{L}^n(A_+)\left[1 + \left(\frac{\mathcal{L}^n(B_+)}{\mathcal{L}^n(A_+)}\right)^{\frac{1}{n}}\right]^n + \mathcal{L}^n(A_-)\left[1 + \left(\frac{\mathcal{L}^n(B_-)}{\mathcal{L}^n(A_-)}\right)^{\frac{1}{n}}\right]^n \\ &= \left(\mathcal{L}^n(A_+) + \mathcal{L}^n(A_-)\right)\left[1 + \left(\frac{\mathcal{L}^n(B)}{\mathcal{L}^n(A)}\right)^{\frac{1}{n}}\right]^n \\ &= \left(\mathcal{L}^n(A)^{\frac{1}{n}} + \mathcal{L}^n(B)^{\frac{1}{n}}\right)^n. \end{aligned}$$

By induction, the inequality holds when A and B are finite union of rectangles.

Step 3. A and B are open sets. WLOG $\mathcal{L}^n(A+B) < \infty$. Then $\mathcal{L}^n(A)$, $\mathcal{L}^n(B) < \infty$. It is an exercise to show that every open set can be decomposed into countably many almost disjoint closed cubes. Given $\varepsilon > 0$, there is some A_1 consisting of finitely many almost disjoint cubes such that $A_1 \subset A$ and $\mathcal{L}^n(A \setminus A_1) < \varepsilon$.

Similarly, we have $B_1 \subset B$ and $\mathcal{L}^n(B \setminus B_1) < \varepsilon$. Then $A_1 + B_1 \subset A + B$ and

$$\begin{aligned}\mathcal{L}^n(A+B) \geq \mathcal{L}^n(A_1+B_1) &\geq \left(\mathcal{L}^n(A_1)^{\frac{1}{n}} + \mathcal{L}^n(B_1)^{\frac{1}{n}}\right)^n \\ &\geq \left[(\mathcal{L}^n(A)-\varepsilon)^{\frac{1}{n}} + (\mathcal{L}^n(B)-\varepsilon)^{\frac{1}{n}}\right]^n,\end{aligned}$$

and the inequality follows after letting $\varepsilon \to 0$.

Step 4. A, B compact. Let $A^\varepsilon = \{x \in \mathbb{R}^n : \mathrm{dist}(x, A) < \varepsilon\}$ and $B^\varepsilon = \{x \in \mathbb{R}^n : \mathrm{dist}(x, B) < \varepsilon\}$ be open and $A^\varepsilon \downarrow A$, $B^\varepsilon \downarrow B$ as $\varepsilon \to 0$. So $\mathcal{L}^n(A^\varepsilon) \to \mathcal{L}^n(A)$ and $\mathcal{L}^n(B^\varepsilon) \to \mathcal{L}^n(B)$ as $\varepsilon \to 0$ (see Ex 6). We have $A+B$ is compact and $A + B \subset A^\varepsilon + B^\varepsilon \subset (A+B)^{2\varepsilon}$. So, assuming $\mathcal{L}^n(A+B) < \infty$, for every $\rho > 0$, there exists ε such that

$$\begin{aligned}\rho + \mathcal{L}^n(A+B) &\geq \mathcal{L}^n\left((A+B)^{2\varepsilon}\right) \\ &\geq \mathcal{L}^n\left(A^\varepsilon + B^\varepsilon\right) \\ &\geq \left[\mathcal{L}^n(A^\varepsilon)^{\frac{1}{n}} + \mathcal{L}^n(B^\varepsilon)^{\frac{1}{n}}\right]^n \\ &\geq \left[\mathcal{L}^n(A)^{\frac{1}{n}} + \mathcal{L}^n(B)^{\frac{1}{n}}\right]^n,\end{aligned}$$

and the inequality follows after letting $\rho \to 0$.
Step 5. Now, let A, B and $A+B$ be measurable. WLOG $\mathcal{L}^n(A+B) < \infty$. (So $\mathcal{L}^n(A), \mathcal{L}^n(B) < \infty$.) Given $\varepsilon > 0$, by inner regularity pick $K_1 \subset A$, $K_2 \subset B$, such that $\mathcal{L}^n(A \setminus K_1), \mathcal{L}^n(B \setminus K_2) < \varepsilon$. Then $K_1 + K_2 \subset A + B$, and

$$\begin{aligned}\mathcal{L}^n(A+B) &\geq \mathcal{L}^n(K_1+K_2) \\ &\geq \left[\mathcal{L}^n(K_1)^{\frac{1}{n}} + \mathcal{L}^n(K_2)^{\frac{1}{n}}\right]^n \\ &\geq \left[(\mathcal{L}^n(A)-\varepsilon)^{\frac{1}{n}} + (\mathcal{L}^n(B)-\varepsilon)^{\frac{1}{n}}\right]^n.\end{aligned}$$

Letting $\varepsilon \to 0$, we conclude the proof of Brunn-Minkowski inequality.

□

As an application of Brunn-Minkowski inequalityl, we discuss how to use it to deduce the isoperimetric inequality.

The isoperimetric inequality asserts that for every set E in $\mathbb{R}^n$,

$$A_n(E) \geq c_n \mathcal{L}^n(E)^{\frac{n-1}{n}}, \qquad c_n = \frac{A_n(B)}{\mathcal{L}^n(B)^{(n-1)/n}},$$

where $A_n(E)$ denotes the "surface area" or "perimeter" of E and B is a ball. It seems that this inequality is never proved in full generality. In history, it was first established for convex sets, next for sets with differentiable boundaries and finally for domains bounded by rectifiable boundaries.

In the following we will use $|E|$ to denote the n-dimensional Lebsegue measure of a set E in $\mathbb{R}^n$ instead of $\mathcal{L}^n(E)$.

In fact, for an arbitrary set, there is a very general notion of perimeter replaced by the Minkowski content. It is based on the geometric observation that for a smooth hypersurface Σ in $\mathbb{R}^n$, its δ-tube

$$\Sigma^\delta = \left\{x \in \mathbb{R}^2 : d(x, \Sigma) < \delta\right\}$$

satisfies

$$\lim_{\delta \downarrow 0} \frac{\left|\Sigma^\delta\right|}{2\delta} = \text{the surface area of } \Sigma.$$

Motivated by this, for any set E in $\mathbb{R}^n$, we set

$$A_n(E) = \lim_{\delta \to 0} \frac{|(\partial E)^\delta|}{2\delta},$$

provided the limit exists.

Theorem 3.6. *Let Ω be a bounded open set in $\mathbb{R}^n$ such that $A_n(\Omega)$ exists. Then*

$$A_n(\Omega) \geq n\omega_n^{1/n} |\Omega|^{\frac{n-1}{n}},$$

where ω_n is the volume of the unit ball in $\mathbb{R}^n$. Equality sign in this inequality holds when Ω is a ball.

Proof. Here we set $\Gamma = \partial\Omega$. For a positive δ, let

$$\Omega_+(\delta) = \left\{x \in \mathbb{R}^n : d(x, \overline{\Omega}) < \delta\right\},$$

and

$$\Omega_-(\delta) = \{x \in \mathbb{R}^n : d(x, \Omega') \geq \delta\}$$

and

$$\Gamma^\delta = \{x \in \mathbb{R}^n : d(x, \partial\Omega) < \delta\}.$$

It is readily checked that

$$\Omega_+(\delta) = \Omega_-(\delta) \cup \Gamma^\delta (\text{ disjoint union}),$$

$$\Omega + B_\delta(0) \subset \Omega_+(\delta)$$

and

$$\Omega_-(\delta) + B_\delta(0) \subset \Omega.$$

Applying Brunn-Minkowski inequality to the open set $\Omega + B_\delta(0)$, we have

$$|\Omega_+(\delta)|^{\frac{1}{n}} \geq |\Omega + B_\delta(0)|^{\frac{1}{n}} \geq |\Omega|^{\frac{1}{n}} + |B_\delta(0)|^{\frac{1}{n}}.$$

Using binomial expansion we have

$$|\Omega_+(\delta)| \geq |\Omega| + n\,|\Omega|^{\frac{n-1}{n}}\,|B_\delta(0)|^{\frac{1}{n}}$$

Similarly,

$$|\Omega| \geq |\Omega_-(\delta)| + n\,|\Omega_-(\delta)|^{\frac{n-1}{n}}\,|B_\delta(0)|^{\frac{1}{n}}.$$

By adding up these two inequalities and using $\Omega_+(\delta) \setminus \Omega_-(\delta) = \Gamma^\delta$,

$$\begin{aligned}\left|\Gamma^\delta\right| &= |\Omega_+(\delta)| - |\Omega_-(\delta)| \\ &\geq n\left(|\Omega|^{\frac{n-1}{n}} + |\Omega_-(\delta)|^{\frac{n-1}{n}}\right)|B_\delta(0)|^{\frac{1}{n}}.\end{aligned}$$

Using $|B_\delta(0)| = \omega_n\delta^n$, we deduce

$$\frac{|\Gamma_\delta|}{2\delta} \geq \frac{n}{2}\left(|\Omega|^{\frac{n-1}{n}} + |\Omega_-(\delta)|^{\frac{n-1}{n}}\right)\omega_n^{\frac{1}{n}}.$$

Now the desired inequality follows readily by letting δ go to 0.

Let us verify that it becomes equality when Ω is a ball of radius r. In fact, in this case

$$|\Omega| = \omega_n r^n$$

and

$$A_n(\Omega) = n\omega_n r^{n-1}.$$

The left hand side of this inequality becomes $n\omega_n R^{n-1}$ and its right hand side becomes

$$n\omega_n^{\frac{1}{n}}\left(\omega_n r^n\right)^{\frac{n-1}{n}} = n\omega_n r^{n-1},$$

so equality holds at a ball. □

3.4 Hausdorff Measures

Let $0 \leq s < \infty$ and $A \subset \mathbb{R}^n$. For $0 < \delta \leq \infty$, define

$$\mathcal{H}^s_\delta(A) = \inf\left\{\sum_j d(C_j)^s : A \subset \bigcup_j C_j,\ d(C_j) < \delta, C_j \subset \mathbb{R}^n\right\}$$

where $d(C)$ is the diameter of C, that is, $d(C) = \sup\{|x-y| : x, y \in C\}$ and $|x-y|$ is the Euclidean distance in $\mathbb{R}^n$. The s-dimensional Hausdorff measure (or more precisely, unnormalized Hausdorff measure) of A is defined as

$$\begin{aligned}\mathcal{H}^s(A) &= \lim_{\delta\to 0}\mathcal{H}^s_\delta(A) \\ &= \sup_{\delta>0}\mathcal{H}^s_\delta(A).\end{aligned}$$

We note that $\mathcal{H}^s(A) \in [0, \infty]$.

Theorem 3.7. *(a) $\mathcal{H}^s$ is a Borel measure on $\mathbb{R}^n$.*

(b) For every $A \subset \mathbb{R}^n$, there exists a Borel set $B \supset A$ such that $\mathcal{H}^s(B) = \mathcal{H}^s(A)$.

(c) For every open set G, $\mathcal{H}^s(G) = \sup\{\mathcal{H}^s(K) : K \subset G$ is compact$\}$.

(d) For every Borel set A with $\mathcal{H}^s(A) < \infty$, given any $\varepsilon > 0$, there exists a compact set $K \subset A$ such that $\mathcal{H}^s(A \setminus K) < \varepsilon$.

Proof. Step 1. $\mathcal{H}^s_\delta$ and $\mathcal{H}^s$ are outer measures. As $\phi \subset \{x\}$ for every $x \in \mathbb{R}^n$ and $d(\{x\}) = 0$, $\mathcal{H}^s_\delta(\phi) = 0$ for all $\delta > 0$ and $\mathcal{H}^s(\phi) = 0$. Next, let $A \subset \bigcup_j A_j$, we want to show

$$\mathcal{H}^s_\delta(A) \leq \sum_{j=1}^{\infty} \mathcal{H}^s_\delta(A_j). \tag{3.1}$$

WLOG, $\mathcal{H}^s_\delta(A_j) < \infty, \forall j$. For $\varepsilon > 0$, there exist C^j_k, $A_j \subset \bigcup_k C^j_k$, $d(C^j_k) < \delta$, such that

$$\mathcal{H}^s_\delta(A_j) + \frac{\varepsilon}{2^j} \geq \sum_k d(C^j_k)^s.$$

Thus

$$\begin{aligned}\sum_j \mathcal{H}^s_\delta(A_j) + \varepsilon &\geq \sum_j \sum_k d(C^j_k)^s \\ &= \sum_{j,k} d(C^j_k)^s \\ &\geq \mathcal{H}^s_\delta(A) \qquad (\text{since } A \subset \bigcup_{j,k} C^j_k),\end{aligned}$$

and (3.1) follows by letting $\varepsilon \to 0$. Now, from (3.1)

$$\begin{aligned}\mathcal{H}^s_\delta(A) &\leq \sum_{j=1}^{\infty} \mathcal{H}^s_\delta(A_j) \\ &\leq \sum_{j=1}^{\infty} \mathcal{H}^s(A_j),\end{aligned}$$

and $\mathcal{H}^s(A) \leq \sum_j \mathcal{H}^s(A_j)$ holds after letting $\delta \to 0$.

Step 2. $\mathcal{H}^s$ is a Borel measure. Here we use Caratheodory's criterion. Letting $A, B \subset \mathbb{R}^n$ with $d(A, B) = \delta_0 > 0$, we want to show that

$$\mathcal{H}^s(A \cup B) = \mathcal{H}^s(A) + \mathcal{H}^s(B) \tag{3.2}$$

On one hand, $\mathcal{H}^s(A\cup B) \le \mathcal{H}^s(A)+\mathcal{H}^s(B)$ is evident by countable subadditivity. To show the reverse inequality, let $\{C_j\}$ be a covering of $A\cup B$ with $d(C_j)<\delta<\delta_0/4$. Each C_j can only intersect either A or B but not both. We can divide $\{C_j\}$ into two classes $\{C_j'\}$ and $\{C_j''\}$ such that $A\subset \bigcup_j C_j'$ and $B\subset\bigcup_j C_j''$. Then

$$\begin{aligned}\sum_j d(C_j)^s &= \sum_j d(C_j')^s + \sum_j d(C_j'')^s \\ &\ge \mathcal{H}^s_\delta(A)+\mathcal{H}^s_\delta(B).\end{aligned}$$

Taking infimum over all $\{C_j\}$, we conclude

$$\mathcal{H}^s_\delta(A\cup B)\ge \mathcal{H}^s_\delta(A)+\mathcal{H}^s_\delta(B), \quad \forall\ 0<\delta<\frac{\delta_0}{4}.$$

Letting $\delta\to 0$, we get (3.2).

Step 3. If $\mathcal{H}^s(A)=\infty$, (b) holds by taking $B=\mathbb{R}^n$. When $\mathcal{H}^s(A)<\infty$, for any given k, we fix $\{C_j^k\}$, $A\subset\bigcup_j C_j^k$, $d(C_j^k)<1/k$, such that

$$\mathcal{H}^s_{\frac{1}{k}}(A)+\frac{1}{k}\ge\sum_j d(C_j^k)^s.$$

Let $B_k=\bigcup_{j=1}^\infty C_j^k$ and $B=\bigcap_{k=1}^\infty B_k$. Then $B_k\supset A$ for each k and hence $B\supset A$. Moreover,

$$\begin{aligned}\mathcal{H}^s(A)+\frac{1}{k} &\ge \mathcal{H}^s_{\frac{1}{k}}(A)+\frac{1}{k} \\ &\ge \sum_j d(C_j^k)^s \\ &\ge \mathcal{H}^s_{\frac{1}{k}}(B).\end{aligned}$$

Letting $k\to\infty$, $\mathcal{H}^s(A)\ge\mathcal{H}^s(B)$, and (b) holds.

Step 4. Let $K_j=\{x\in G:\operatorname{dist}(x,G^c)\le 1/j,\ |x|\le j\}$. Then $K_j, j\ge 1$, are compact and $K_j\uparrow G$. By Proposition 1.4, (c) holds.

Step 5. You are referred to Lemma 1 (i) on p. 6 in [EG] for a proof of (d). □

When $s\in(0,n)$ is not an integer, it can be shown that $\mathcal{H}^s(R)=\infty$ on every cube R. It implies that Hausdorff measures $\mathcal{H}^s, s<n$, are not finite on compact sets. In particular, they cannot be constructed by Riesz representation theorem.

Every Euclidean motion is of the form

$$Tx=Ax+b, \quad b\in\mathbb{R}^n,$$

where A is an $n \times n$-orthogonal matrix. The following immediate proposition shows that the Hausdorff measures are geometric measures in the sense that they are invariant under all Euclidean motions.

Proposition 3.8. *Let $A \subset \mathbb{R}^n$.*

(a) $\mathcal{H}^s(TA) = \mathcal{H}^s(A)$, where T is a Euclidean motion.

(b) $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A), \qquad \forall \lambda > 0$.

We now proceed to define the Hausdorff dimension of a set.

Proposition 3.9. *Let $A \subset \mathbb{R}^n$.*

(a) $\mathcal{H}^s(A) = 0$, $\forall s > n$.

(b) If $\mathcal{H}^s(A) < \infty$ for some $s \in [0, \infty)$, then $\mathcal{H}^t(A) = 0$, $\forall t > s$.

(c) If $\mathcal{H}^s(A) \in (0, \infty]$ for some $s \in [0, \infty)$, then $\mathcal{H}^t(A) = \infty$, $\forall t < s$.

Proof. (a) Divide the unit cube Q into subcubes of side $1/k$. There are k^n many such subcubes with diameters equal to $\sqrt{n}/k$. We have,

$$\begin{aligned} \mathcal{H}^s_{\frac{\sqrt{n}}{k}}(Q) &\le \sum_j d(Q_j)^s = k^n \left(\frac{\sqrt{n}}{k}\right)^s \\ &= \sqrt{n}^s \frac{1}{k^{s-n}} \to 0, \qquad \text{as } k \to \infty, \end{aligned}$$

so $\mathcal{H}^s(Q) = 0$ and $\mathcal{H}^s(\mathbb{R}^n) = 0$ for $s > n$.

(b) As $\mathcal{H}^s(A)$ is finite, for each $\delta > 0$, there exists $\{C_j\}$, $d(C_j) < \delta$, such that $A \subset \bigcup_j C_j$ and

$$\sum_j d(C_j)^s \le \mathcal{H}^s(A) + 1.$$

As $d(C_j)^s = d(C_j)^t d(C_j)^{s-t} \ge \delta^{s-t} d(C_j)^t$,

$$\begin{aligned} \mathcal{H}^t_\delta(A) &\le \sum_j d(C_j)^t \\ &\le \delta^{t-s}\left(\mathcal{H}^s(A) + 1\right) \to 0, \qquad \text{as } \delta \to 0. \end{aligned}$$

(c) It follows from (b). $\square$

From this proposition we know that for every $A \subset \mathbb{R}^n$, $\mathcal{H}^s(A) = 0$ for $s > n$. Lowering s passing through n, we will first hit a point $s_0 \in [0, n]$ such that $\mathcal{H}^s(A) = 0$ for all $s > s_0$ and $\mathcal{H}^s(A) = \infty$ for all $s < s_0$. This point is characterized by

$$s_0 = \inf\{s : \mathcal{H}^s(A) = 0\}$$

and is called the *Hausdorff dimension* of A.

Proposition 3.10.

(a) $\mathcal{H}^0$ *is the counting measure in* $\mathbb{R}^n$.

(b) $\mathcal{H}^1 = \mathcal{L}^1$ *on* $\mathbb{R}$.

(c) $\mathcal{H}_N^n = \mathcal{L}^n$ *on* $\mathbb{R}^n$ *for* $n \geq 2$.

Here the *normalized Hausdorff measure* is given by

$$\mathcal{H}_N^s(A) = \sigma(s)\mathcal{H}^s(A) \quad \forall A \subset \mathbb{R}^n,$$

where

$$\sigma(s) = \frac{\pi^{s/2}}{2^s\Gamma(s/2+1)}, \quad \Gamma(s) = \int_0^\infty e^{-x}x^{s-1}dx.$$

When $s = n$, we have $\sigma(n) = \mathcal{L}^n(B)/2^n$ where B is the unit ball in $\mathbb{R}^n$.

Proof. (a) Exercise.

(b) Let $A \subset \mathbb{R}$ and $\delta > 0$. We have

$$\begin{aligned}
\mathcal{L}^1(A) &= \inf\left\{\sum_j |I_j| : A \subset \bigcup_j I_j\right\} \\
&= \inf\left\{\sum_j |I_j| : A \subset \bigcup_j I_j,\ d(I_j) < \delta\right\} \\
&\quad \text{(every interval can be chopped up to subintervals of } d(I_j) < \delta) \\
&= \inf\left\{\sum_j d(I_j) : A \subset \bigcup_j I_j,\ d(I_j) < \delta\right\} \\
&= \inf\left\{\sum_j d(C_j) : A \subset \bigcup_j C_j,\ d(C_j) < \delta\right\} \\
&= \mathcal{H}_\delta^1(A),
\end{aligned}$$

so $\mathcal{L}^1(A) = \mathcal{H}^1(A)$. Here we have used the fact that for each set C in $\mathbb{R}$ there is an interval I satisfying $d(I) = d(C)$.

(c) To show that $c \geq 1/\sigma(n)$ we need the isodiametric inequality from [EG]: For $A \subset \mathbb{R}^n$,

$$\mathcal{L}^n(A) \leq \sigma(n)d(A)^n.$$

It asserts that the ball has the largest volume among all sets with the same diameter. Suppose $A \subset \bigcup_j C_j$, $d(C_j) < \delta$, then

$$\begin{aligned}\mathcal{L}^n(A) &\le \sum_j \mathcal{L}^n(C_j) \\ &\le \sum_j \sigma(n) d(C_j)^n \\ &= \sigma(n) \sum_j d(C_j)^n.\end{aligned}$$

Taking infimum over all $\{C_j\}$, we get

$$\begin{aligned}\mathcal{L}^n(A) &\le \sigma(n)\mathcal{H}^n_\delta(A) \\ &\le \sigma(n)\mathcal{H}^n(A),\end{aligned}$$

hence $c \ge 1/\sigma(n)$.

To show the reverse inequality, we first claim that there exists some constant C such that

$$\mathcal{H}^n(A) \le C\mathcal{L}^n(A), \quad \forall A \subset \mathbb{R}^n \tag{3.3}$$

Assume $\mathcal{L}^n(A) < \infty$. For every $\varepsilon > 0$, we can find congruent cubes $\{R_j\}$, $A \subset \bigcup_j R_j$, such that

$$\mathcal{L}^n(A) + \varepsilon \ge \sum_j |R_j|.$$

Using $|R_j| = \left(\dfrac{d(R_j)}{\sqrt{n}}\right)^n$,

$$\begin{aligned}\mathcal{L}^n(A) + \varepsilon &\ge \left(\frac{1}{\sqrt{n}}\right)^n \sum_j d(R_j)^n \\ &\ge \left(\frac{1}{\sqrt{n}}\right)^n \mathcal{H}^n_l(A),\end{aligned}$$

where l is the side length of R_j. Letting first $l \to 0$ and then $\varepsilon \to 0$,

$$\mathcal{L}^n(A) \ge \left(\frac{1}{\sqrt{n}}\right)^n \mathcal{H}^n(A).$$

At last, for $\varepsilon > 0$, there are R_j, $A \subset \bigcup_j R_j$, $d(R_j) < \delta$, such that

$$\mathcal{L}^n(A) + \varepsilon \ge \sum_j |R_j|.$$

Using the fact that each R_j can be written as a disjoint union of balls B^j_k (p. 28

in [EG]) such that

$$\mathcal{L}^n\left(R_j \setminus \bigcup_k B_k^j\right) = 0,$$

we have

$$\begin{aligned}
\mathcal{L}^n(A) + \varepsilon &\geq \sum_{j,k} \mathcal{L}^n\left(B_k^j\right) \\
&= \sigma(n) \sum d\left(B_k^j\right)^n \\
&\geq \sigma(n)\mathcal{H}_\delta^n\left(\bigcup_{j,k} B_k^j\right).
\end{aligned}$$

As $\mathcal{L}^n\left(R_j \setminus \bigcup_k B_k^j\right) = 0$, by (3.3), $\mathcal{H}^n\left(R_j \setminus \bigcup_k B_k^j\right) = 0$, so $\mathcal{H}^n\left(\bigcup_j R_j \setminus \bigcup_k B_k^j\right) = 0$ and

$$\begin{aligned}
\mathcal{H}_\delta^n\left(\bigcup_{j,k} B_k^j\right) &= \mathcal{H}_\delta^n\left(\bigcup_j R_j\right) - \mathcal{H}_\delta^n\left(\bigcup_j R_j \setminus \bigcup_k B_k^j\right) \\
&= \mathcal{H}_\delta^n\left(\bigcup_j R_j\right) \\
&\geq \mathcal{H}_\delta^n(A).
\end{aligned}$$

That is,

$$\mathcal{L}^n(A) + \varepsilon \geq \sigma(n)\mathcal{H}_\delta^n(A),$$

and $\mathcal{L}^n(A) \geq \mathcal{H}_N^n(A)$ holds. □

3.5 Hausdorff Dimension

We present a preliminary study on the mapping properties of the Hausdorff measures.

A map $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ is *Hölder continuous* with exponent $\alpha \in (0, 1)$ if there exists a constant M such that

$$|f(x) - f(y)| \leq M\,|x - y|^\alpha\,, \qquad \forall x, y \in A.$$

The constant M is called a *Hölder constant.* The map f is called *Lipschitz continuous* when $\alpha = 1$ in this condition and M is called a *Lipschitz constant.* Lipschitz and Höler continuous functions are very much different.

Proposition 3.11. *Let $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ be Hölder or Lipschitz continuous.*

We have

$$\mathcal{H}^{\frac{s}{\alpha}}(f(A)) \le M^{\frac{s}{\alpha}}\mathcal{H}^s(A).$$

Proof. Let $A \subset \bigcup_j C_j$, $d(C_j) < \delta$, satisfy $\mathcal{H}^s(A) + \varepsilon \ge \sum_j d(C_j)^s$. We have

$$d(f(C_j)) \le M d(C_j)^\alpha.$$

Thus

$$\frac{1}{M^{\frac{1}{\alpha}}} d(f(C_j))^{\frac{1}{\alpha}} \le d(C_j),$$

and

$$\mathcal{H}^s(A) + \varepsilon \ge \frac{1}{M^{\frac{s}{\alpha}}} \sum_j d(f(C_j))^{\frac{s}{\alpha}}.$$

As $f(A) \subset \bigcup_j f(C_j)$ and $d(f(C_j)) \le \rho \equiv M\delta^\alpha$,

$$M^{\frac{s}{\alpha}}\left(\mathcal{H}^s(A) + \varepsilon\right) \ge \mathcal{H}^s_\rho(f(A)),$$

and the result follows by letting $\delta \to 0$. $\square$

Let us examine an example. Let $f : [0,1] \to \mathbb{R}$ be Lipschitz continuous and its graph is given by

$$C = \{(x, f(x)) : x \in [0,1]\} \subset \mathbb{R}^2.$$

By Proposition 3.9, we have

$$\mathcal{H}^1(C) \le \sqrt{1+M^2}\, \mathcal{H}^1([0,1]) = \sqrt{1+M^2}.$$

On the other hand, $g : C \to [0,1]$ given by $g(x,y) = x$ is the inverse of f. From $|g(x,y) - g(x_1,y_1)| \le |x - x_1| \le |(x,y) - (x_1,y_1)|$, we see that it is Lipschitz continuous with Lipschitz constant $M = 1$. By Proposition 3.9,

$$1 = \mathcal{H}^1[0,1] \le \mathcal{H}^1(C).$$

We conclude that

$$1 \le \mathcal{H}^1(C) \le \sqrt{1+M^2}.$$

Indeed, one can show that

$$\mathcal{H}^1(C) = \int_0^1 \sqrt{1 + f'^2(x)}\, dx$$

when f is Lipschitz continuous. This formula is a special case of the area formula (see chapter 3 of [EG]). It shows that the Hausdorff measures in special cases such as $\mathcal{H}^1$ in $\mathbb{R}^2$, $\mathcal{H}^1$ and $\mathcal{H}^2$ in $\mathbb{R}^3$, really coincide with the length of curves and surface area of surfaces defined in advanced calculus.

For $A \subseteq \mathbb{R}^n$, the *Hausdorff dimension* of A, denoted as $\dim_H A$, is defined as

$$\dim_H A = \inf\{s \geq 0 : \ \mathcal{H}^s(A) < \infty\}.$$

We end this section by determining the Hausdorff dimension of the Cantor set $\mathcal{C}$.

Proposition 3.12. *The Hausdorff dimension of $\mathcal{C}$ is $\log 2/\log 3$. In fact,*

$$\mathcal{H}^\gamma(\mathcal{C}) \in (0,1], \quad \gamma = \frac{\log 2}{\log 3}.$$

Proof. Observe that $\mathcal{C} = \bigcap_n \mathcal{C}_n$ where $\mathcal{C}_n$ consists of 2^n many intervals of length $1/3^n$. So

$$\begin{aligned}
\mathcal{H}^\gamma_{\frac{1}{3^n}}(\mathcal{C}) &\leq \sum d(\mathcal{C}_n)^\gamma = 2^n \left(\frac{1}{3^n}\right)^\gamma, \quad 3^\gamma = 2 \\
&= \left(\frac{2}{3^\gamma}\right)^n = 1.
\end{aligned}$$

On the other hand, the Cantor function g is the uniform limit of $\{g_n\}$ and we have

$$|g_{n+1}(x) - g_n(x)| \leq \frac{1}{2^n}.$$

Therefore,

$$|g_l(x) - g_n(x)| \leq \sum_{j=0}^{l-n} \frac{1}{2^{n+j}} < \sum_0^\infty \frac{1}{2^{n+j}} = \frac{1}{2^{n-1}}.$$

Letting $l \to \infty$,

$$|g(x) - g_n(x)| \leq \frac{1}{2^{n-1}}.$$

On the other hand, each g_n has slope at most $(3/2)^n$, so

$$|g_n(x) - g_n(y)| \leq \left(\frac{3}{2}\right)^n |x-y|.$$

As a result,

$$\begin{aligned}
|g(x) - g(y)| &\leq |g(x) - g_n(x)| + |g_n(x) - g_n(y)| + |g_n(y) - g(y)| \\
&\leq \left(\frac{3}{2}\right)^n |x-y| + \frac{4}{2^n} \\
&= \frac{1}{2^n}\left(3^n|x-y| + 4\right), \quad \forall n \geq 1.
\end{aligned}$$

For each $x, y \in [0,1]$, we pick n so that

$$\frac{1}{3^n} \leq |x-y| < \frac{1}{3^{n-1}}.$$

Then $1 \leq 3^n |x-y| < 3$ and

$$|g(x) - g(y)| \leq \frac{7}{2^n} = \frac{7}{3^{n\gamma}}.$$

As $3^n |x-y| \geq 1$, $\frac{1}{3^{n\gamma}} \leq |x-y|^\gamma$ and

$$|g(x) - g(y)| \leq 7 |x-y|^\gamma .$$

As $g(\mathcal{C}) = [0,1]$ by Proposition 3.9, $1 = \mathcal{H}^1(g(\mathcal{C})) \leq 7^{\frac{1}{\gamma}} \mathcal{H}^\gamma(\mathcal{C})$. We have shown that the Hausdorff dimension of $\mathcal{C}$ is γ. $\square$

Comments on Chapter 3. Lebesgue measure on the real line was introduced by Lebesgue in 1901/02 together with his integration theory. Nowadays, at least three ways to define the Lebesgue measure are known. First, it is an outer measure using the cubes as a gauge. Second, it is the measure obtained by the positive linear functional defined by the Riemann integral via the representation theorem. Third, it is the n-times product measure of the one dimensional Lebesgue measure. (This is the approach adapted in [EG]. We will discuss it in Chapter 8.) We follow [SS] by employing the first approach, which is intuitive clear, while the second approach is used in [R]. Both approaches can be used to construct the invariant Haar measure on a locally compact topological group, see Nachbin's book, The Haar Integral. Another application of the representation theorem is, on a Riemannian manifold one can use the metric to define an n-form called the volume form on the manifold. Then the analogue of Riemann integral with respect to this volume form is well-defined on the manifold. By the theorem there exists a canonical volume measure defined on the manifold which is invariant under all isometries.

Note that in some books the Lebesgue outer measure and Lebesgue measure are distinguished and used notations like m^* and m. Here Lebesgue measure is always referred to the outer measure and the notation $\mathcal{L}^n$ is applied to measurable and non-measurable sets alike.

The proof of the scaling property of the Lebesgue measure under a linear transformation using the characterization of translational invariant measures is taken from [R]. Different proofs can be found in [SS] either by Fubini's theorem (problem 4 in chapter 2) or in exercise 26, chapter 3. But Rudin's proof is most elegant. In the second part of this chapter we will show that the n-dimensional Lebesgue measure is equal to the n-dimensional Hausdorff measure on $\mathbb{R}^n$. As it is obvious from the definition that the latter is Euclidean invariant, there we will have another proof of the Euclidean invariance of the Lebesgue measure. Any Euclidean-invariant measure is called a *geometric measure*. The Hausdorff measures and Lebesgue measures are geometric measures. The book by Federer, Geometric Measure Theory, is a classic on this topic. You may find many examples of geometric measures in it.

Nonmeasurable sets were first found by Vitali (1905) and every book I know follows his construction. The nonmeasurable sets $\mathcal{E} \cap [0,1]$ are usually called Vitali sets. The existence of nonmeasurable sets depends essentially on the axiom of choice. It destroys the hope to have a geometric measure on all sets in $\mathbb{R}$ and $\mathbb{R}^n$ as well. One may make concession by asking whether there is a translational invariant, finitely additive "measure" on all sets in $\mathbb{R}$ which equal to the Lebesgue measure on Lebesgue measurable sets. Here the requirement of countably additivity has been relaxed to finite additivity. It turns out the answer is affirmative. By using Hahn-Banach theorem, of which proof also depends on the axiom of choice, one can establish the existence of such a measure. See, chapter 1 in Functional Analysis by Stein-Shakarchi for details. In higher dimensions, it is natural to require these finitely additive measures being invariant under translations and rotations. The answer is yes for $n = 2$ by Banach (1920) and no for $n \geq 3$ as a consequence of Banach-Tarski paradox (1923). This striking paradox (theorem) asserts that given any two bounded sets S and T in $\mathbb{R}^n$ with non-empty interior, there is a partition of S into $S_1, S_2, \cdots, S_k$ and a partition of T into $T_1, T_2, \cdots, T_k$ so that S_i and T_i are congruent for $i = 1, 2, \cdots, k$, where the number k depends on S and T. In the special case where S is the unit ball B_1 in $\mathbb{R}^3$ and T is the union of two copies of B_1, it was known that k is equal to 5. Specifically that means we can find a partition of B_1 into five pieces A, B, C, D and E so that A and B can be put up by Euclidean motions to form a copy of B_1 while C, D and F form another B_1 in a similar way. It shows that finitely additive Euclidean invariant "measure" does not exist for $\mathbb{R}^3$. For, If there is such a nontrivial μ, we will have $\mu(B_1) = \mu(A \cup B \cup C \cup D \cup E) = \mu(A \cup B) + \mu(C \cup D \cup E) = 2\mu(B_1)$, which is impossible. The same idea works for all $n > 3$. You may google under Banach-Tarski paradox for more information.

There are many fine, delicate properties of the one dimensional Lebesgue measure we do not cover. The interested reader may consult, for example, Measure Theory by Halmos and Measure and Category by Oxtoby.

The Brunn-Minkowski inequality is taken from [SS]. It is strange that this fundamental result was not included in text books on real analysis until [SS].

In Chapter 7 we will apply it to prove the isoperimetric inequality. We did not discuss the equality case of the inequality. Historically, this inequality was first proved by Brunn for convex sets in the plane and subsequently his teacher, Minkowski established it for convex sets in all dimensions. He also showed that the equality in this inequality holds only in the following cases, first, the two sets are homothetic (they coincide after a translation and dilation), or second, both are degenerate and sit inside two parallel affine sets. The proof presented here is not restricted to convex sets and it was found by Hadwiger and Ohmann in 1956. The recent survey by R. J. Gardner, The Brunn-Minkowski inequality in v.39, 355-405, Bulletin of the American Mathematical Society, 2002, gives a comprehensive account on this inequality as well as many other related inequalities.

For their importance in applications and intricate properties, nowadays Hausdorff measures have become an extensively studied subject. The theory of fractals is devoted to the study of the fine properties of sets of fractional dimension. One may consult, for example, [SS] and the books by Falconer's Fractal Geometry, Mattila's Geometry of Sets and Measures in the Euclidean Spaces, etc, for more. Our exposition merely touches the tip of the iceberg. It is worthwhile to point out that the space filling curve discovered by Peano, a continuous map from $[0, 1]$ onto the unit square, is a fractal map, see [SS].