Math 6810
(Probability and Fractals)

Spring 2016

Lecture notes

Pieter Allaart
University of North Texas

February 22, 2016

Recommended reading: (Do not purchase these books before consulting with your instructor!)

1. *Real Analysis* by H. L. Royden (4th edition), Prentice Hall.
2. *Probability and Measure* by P. Billingsley (3rd edition), Wiley.
3. *Probability with Martingales* by D. Williams, Cambridge University Press.
4. *Fractal Geometry: Foundations and Applications* by K. Falconer (2nd edition), Wiley.

Chapter 3

Introduction to fractal geometry

There is no universally accepted definition of the idea of a “fractal”. But the sets that most mathematicians call fractals have one thing in common: when viewed at microscopic scales they look just the same, or approximately the same, as the set viewed at macro scale. This phenomenon is made mathematically precise through the notions of a self-similar, self-affine or self-conformal set. Contrast this with a smooth curve, which when viewed under a magnifying glass, is virtually indistinguishable from a straigt line.

Another common property of sets we consider to be fractals is that they usually have Lebesgue measure zero. As a result, Lebesgue measure is not a useful tool to measure the size of a fractal, and we need a new kind of measure, called *Hausdorff measure*, which can distinguish between different sets of Lebesgue measure zero. Associated with Hausdorff measure is a notion of a fractional dimension which we call *Hausdorff dimension*. One reasonable definition of a fractal is any set whose Hausdorff dimension is not an integer, though this would still exclude many interesting sets having an approximately self-similar structure. Other commonly used notions of (fractal) dimension include packing dimension, box-counting dimension, Assouad dimension and quantization dimension. Most of these are beyond the scope of this course, however. We will focus exclusively on the Hausdorff and box-counting dimensions of sets.

Pictures of the fractals presented in these notes, and many others, can be found in Falconer’s book or online.

3.1 Hausdorff measure and Hausdorff dimension

Before we can introduce Hausdorff measure (and prove that it really is a measure!) we need some preliminary work. We follow Royden, sec. 20.4.

Definition 3.1. Two sets are *separated by the function* ϕ if there exist real numbers $a > b$ such that $\phi > a$ on one of the sets, and $\phi < b$ on the other.

Observe that this condition is stronger than the sets being disjoint.

Recall that an *outer measure* on a set X is a set function $\mu^* : 2^X \to [0, \infty]$ satisfying

(i) $\mu^*(\emptyset) = 0$,

(ii) $\mu^*(\bigcup_{i=1}^\infty E_i) \le \sum_{i=1}^\infty \mu^*(E_i)$ for any sequence $(E_i)_i$ of subsets of X.

Definition 3.2. Let Γ be a set of real-valued functions on a set X. An outer measure μ^* on X is called a *Carathéodory outer measure* with respect to Γ if, whenever $A \subset X$ and $B \subset X$ are separated by some function in Γ, we have $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$.

Lemma 3.3. *If μ^* is a Carathéodory outer measure with respect to Γ, then each function in Γ is μ^*-measurable.*

Proof. Fix $\phi \in \Gamma$ and $a \in \mathbb{R}$, and let $E := \{x : \phi(x) > a\}$. We need to show that E is μ^*-measurable; that is,

$$\mu^*(A) \ge \mu^*(A \cap E) + \mu^*(A \backslash E), \qquad \forall A \subset X. \tag{3.1}$$

We may assume $\mu^*(A) < \infty$, for otherwise (3.1) is trivial. Set $B := A \cap E$, $C := A \backslash E$, and define the sets

$$B_n := B \cap \left\{ x : \phi(x) > a + \frac{1}{n} \right\}, \qquad R_n := B_n \backslash B_{n-1}.$$

Note that for each n,

$$B = B_n \cup \bigcup_{k=n+1}^{\infty} R_k. \tag{3.2}$$

Now ϕ separates R_n and B_{n-2} for each n, and hence ϕ separates R_{2k} and $\bigcup_{j=1}^{k-1} R_{2j}$ for each k. Thus, by induction and using the hypothesis of the lemma,

$$\mu^* \left(\bigcup_{j=1}^{k} R_{2j} \right) = \sum_{j=1}^{k} \mu^*(R_{2j}).$$

Since $\bigcup_{j=1}^k R_{2j} \subset B \subset A$ and $\mu^*(A) < \infty$, it follows that $\sum_{j=1}^\infty \mu^*(R_{2j}) < \infty$. By a similar argument, $\sum_{j=1}^\infty \mu^*(R_{2j-1}) < \infty$. Hence,

$$\sum_{k=1}^{\infty} \mu^*(R_k) < \infty.$$

Now let $\varepsilon > 0$ be given, and choose n so that $\sum_{k=n+1}^\infty \mu^*(R_k) < \varepsilon$. Then by (3.2),

$$\mu^*(B) \le \mu^*(B_n) + \sum_{k=n+1}^{\infty} \mu^*(R_k) < \mu^*(B_n) + \varepsilon.$$

On the other hand, since ϕ separates B_n and C,

$$\mu^*(A) \ge \mu^*(B_n \cup C) = \mu^*(B_n) + \mu^*(C),$$

and therefore,

$$\mu^*(A) > \mu^*(B) + \mu^*(C) - \varepsilon.$$

Since ε was arbitrary, this gives (3.1). $\square$

Notation: Given a metric space (X, d), define

$$d(x, E) := \inf_{y \in E} d(x, y), \qquad E \subset X,$$
$$d(A, B) := \inf_{x \in A} d(x, B) = \inf_{x \in A, y \in B} d(x, y), \qquad A, B \subset X.$$

Definition 3.4. Let (X, d) be a metric space. A *Carathéodory outer measure* on X is an outer measure μ^* such that

$$d(A, B) > 0 \quad \Rightarrow \quad \mu^*(A \cup B) = \mu^*(A) + \mu^*(B). \tag{3.3}$$

Lemma 3.5. *Let (X, d) be a metric space and μ^* a Carathéodory outer measure on X. Then every closed set (and hence, every Borel set) in X is μ^*-measurable.*

Proof. Let Γ be the set of functions $\phi(x) = d(x, E)$, where $E \subset X$. Fix $\phi \in \Gamma$ and corresponding set E, and suppose sets A and B are separated by ϕ, so there exist $a, b \in \mathbb{R}$ with $a > b$ such that, WLOG, $\phi > a$ on A and $\phi < b$ on B. It is an easy exercise to show that $d(A, B) \geq a - b > 0$. Thus, by (3.3), μ^* is a Carathéodory outer measure with respect to Γ, and by Lemma 3.3 every $\phi \in \Gamma$ is μ^*-measurable. But if $F \subset X$ is closed, then $F = \{x : d(x, F) \leq 0\}$ and so F is μ^*-measurable. $\square$

Now let (X, d) be a metric space. For $B \subset X$, let $|B| := \sup_{x,y \in B} d(x, y)$ denote the *diameter* of B.

Definition 3.6. For $\delta > 0$, a *δ-cover* of a set $E \subset X$ is a countable collection $\{U_i\}$ of sets with $|U_i| < \delta$ for all i such that $E \subset \bigcup_{i=1}^{\infty} U_i$.

For $s \geq 0$, $\delta > 0$ and $E \subset X$, define

$$\mathcal{H}^s_\delta(E) := \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ is a } \delta\text{-cover of } E \right\}. \tag{3.4}$$

Note that $\mathcal{H}^s_\delta(E)$ increases as $\delta \downarrow 0$, so the limit

$$\mathcal{H}^s(E) := \lim_{\delta \downarrow 0} \mathcal{H}^s_\delta(E) \tag{3.5}$$

is well defined.

Lemma 3.7. *The set function $\mathcal{H}^s$ is a Carathéodory outer measure.*

Proof. It is straightforward to check that $\mathcal{H}^s$ is an outer measure. We must check that it satisfies (3.3). Let $E, F \subset X$ with $d(E, F) > 0$. Choose $\varepsilon > 0$ so that $d(E, F) > \varepsilon$, and let $0 < \delta < \varepsilon$. If $\{U_i\}$ is a collection of sets with $|U_i| < \delta$ for all i, then no U_i can intersect both E and F. Thus,

$$\mathcal{H}^s_\delta(E \cup F) \geq \mathcal{H}^s_\delta(E) + \mathcal{H}^s_\delta(F). \tag{3.6}$$

Letting $\delta \downarrow 0$ completes the proof. $\square$

As a consequence of Lemmas 3.5 and 3.7, every Borel set in X is $\mathcal{H}^s$-measurable. Thus, the restriction of $\mathcal{H}^s$ to the Borel sets in X is a measure, which we call *s-dimensional Hausdorff measure*. Note that if $X = \mathbb{R}^n$ and $s = n$, then $\mathcal{H}^s$ is a constant multiple (depending on n) of n-dimensional Lebesgue measure. But $\mathcal{H}^s$ is defined also for noninteger s, which is why it is useful in fractal geometry.

From here we follow mostly Falconer.

Lemma 3.8. *For each Borel set $E \subset X$, there is a unique number $s_0 \geq 0$ such that*

(i) $\mathcal{H}^s(E) = \infty$ if $s < s_0$; and

(ii) $\mathcal{H}^s(E) = 0$ if $s > s_0$.

Proof. Let $t > s$ and suppose $\mathcal{H}^s(E) < \infty$. Let $\{U_i\}$ be a δ-cover of E. Then

$$\sum_{i=1}^{\infty} |U_i|^t = \sum_{i=1}^{\infty} |U_i|^{t-s}|U_i|^s \leq \delta^{t-s} \sum_{i=1}^{\infty} |U_i|^s.$$

Taking infima, we get that $\mathcal{H}^t_\delta(E) \leq \delta^{t-s}\mathcal{H}^s_\delta(E)$, and letting $\delta \downarrow 0$ this gives

$$\mathcal{H}^t(E) \leq \left(\lim_{\delta\downarrow 0} \delta^{t-s}\right) \mathcal{H}^s(E) = 0.$$

Hence, $\mathcal{H}^t(E) = 0$, and the lemma follows. □

Definition 3.9. The number s_0 in Lemma 3.8 is called the *Hausdorff dimension* of E, and denoted $\dim_H E$. Thus,

$$\dim_H E = \inf\{s : \mathcal{H}^s(E) = 0\} = \sup\{s : \mathcal{H}^s(E) = \infty\}.$$

Note that Lemma 3.8 does not say anything about the value of $\mathcal{H}^s(E)$ for $s = \dim_H E$. In general this value can be any nonnegative real number, or $+\infty$. For many sufficiently regular sets, however, it is true that $0 < \mathcal{H}^s(E) < \infty$ for $s = \dim_H E$, as we will see.

Example 3.10. Let C denote the ternary Cantor set in $[0, 1]$. Then $\dim_H C = \frac{\log 2}{\log 3} \approx 0.63$. To see this, let $s = \frac{\log 2}{\log 3}$, so that $3^s = 2$. We will show that $1/2 \leq \mathcal{H}^s(C) \leq 1$. Note first that for each k, C is covered by 2^k "basic intervals" of length 3^{-k}. Given $\delta > 0$, choose k so that $3^{-k} < \delta$. Then the 2^k basic intervals of level k form a δ-cover of C, and so

$$\mathcal{H}^s_\delta(C) \leq 2^k(3^{-k})^s = \left(\frac{2}{3^s}\right)^k = 1. \tag{3.7}$$

Since this holds for all $\delta > 0$, $\mathcal{H}^s(C) \leq 1$.

The lower bound requires more work. Let $\delta > 0$ and let $\{U_i\}$ be an arbitrary δ-cover of C. Assume first that $\{U_i\}$ consists of *finitely many* closed intervals. For each i, let $k \in \mathbb{N}$ be such that

$$3^{-(k+1)} \leq |U_i| < 3^{-k}.$$

Then U_i intersects at most one basic interval at level k. For $j \geq k$, each level-k basic interval contains 2^{j-k} level-j basic intervals, so U_i intersects at most $2^{j-k} = 2^j 3^{-sk} \leq 2^j 3^s |U_i|^s$

level-j basic intervals. Choose j so large that $3^{-(j+1)} \leq |U_i|$ for each i. (This is possible since the collection $\{U_i\}$ is finite!) Since collectively the U_i must intersect all 2^j basic intervals at level j, we have

$$2^j \leq \sum_i 2^j 3^s |U_i|^s,$$

and as a result,

$$\sum_i |U_i|^s \geq 3^{-s} = \frac{1}{2}. \tag{3.8}$$

Now suppose $\{U_i\}$ is an arbitrary δ-cover of C. For any $\varepsilon > 0$, we can find open intervals $\{U_i'\}$ such that $U_i \subset U_i'$ and $|U_i'| < (1+\varepsilon)|U_i|$. Then $\{U_i'\}$ is an open cover of C, and since C is compact, it contains a finite subcover, say $\{U_1', \dots, U_n'\}$. Replacing each open interval U_i' with its closure $U_i'' := \mathrm{cl}(U_i')$ gives a finite cover $\{U_i''\}$ of C by closed intervals, which satisfies (3.8) as shown above. But then

$$(1+\varepsilon)\sum_i |U_i|^s \geq \sum_i |U_i'|^s \geq \sum_{i=1}^n |U_i'|^s = \sum_{i=1}^n |U_i''|^s \geq \frac{1}{2}.$$

Since ε was arbitrary, it follows that $\{U_i\}$ satisfies (3.8). Thus, $\mathcal{H}^s(C) \geq \frac{1}{2}$. Finally, since $0 < \mathcal{H}^s(C) < \infty$, we conclude that $\dim_H C = s$.

(By refining the above argument it can in fact be shown that $\mathcal{H}^s(C) = 1$.)

Definition 3.11. Let (X, d) and (Y, ρ) be metric spaces. A function $f : X \to Y$ is *bi-Lipschitz* if there are constants $0 < C_1 < C_2$ such that

$$C_1 d(x,y) \leq \rho(f(x), f(y)) \leq C_2 d(x,y) \qquad \forall x, y \in X. \tag{3.9}$$

Proposition 3.12. *Hausdorff dimension has the following properties:*

(i) (Monotonicity) If $E \subset F$ then $\dim_H E \leq \dim_H F$;

(ii) (Countable set) If E is a countable set, then $\dim_H E = 0$;

(iii) (Countable stability) $\dim_H(\bigcup_{n=1}^\infty E_n) = \sup_{n\in\mathbb{N}} \dim_H E_n$.

(iv) (Lipschitz invariance) If $f : X \to Y$ is bi-Lipschitz, then $\dim_H f(E) = \dim_H E$.

Proof. Exercise. □

As example 3.10 suggests, calculating the Hausdorff dimension of a set straight from the definition can be rather tedious, especially where the lower bound is concerned. Fortunately, there are several useful tools to facilitate the computation. One of them is the following.

Lemma 3.13 (Distribution of mass principle)**.** *Suppose there is a Borel measure μ (distribution of mass) on X and a finite constant $C > 0$ such that $\mu(E) > 0$ and, for any Borel set $U \subset X$,*

$$\mu(U) \leq C|U|^s. \tag{3.10}$$

Then $\mathcal{H}^s(E) \geq \mu(E)/C > 0$, and in particular, $\dim_H E \geq s$.

Proof. For any δ-cover $\{U_i\}$ of E,

$$0 < \mu(E) \leq \sum_{i=1}^{\infty} \mu(U_i) \leq C \sum_{i=1}^{\infty} |U_i|^s.$$

Taking infima completes the proof. □

3.2 Box-counting dimension

Because Hausdorff dimension can be difficult to compute, often a different notion of dimension, called *box-counting dimension*, is used instead. Box-counting dimension is generally easier to determine than Hausdorff dimension, but as we will see, it lacks some of the properties of Proposition 3.12.

Notation: For a set $E \subset X$ and $\delta > 0$, let $N_\delta(E)$ denote the smallest number of sets of diameter at most δ which can cover E. For example, $N_\delta([0,1]) = \lceil 1/\delta \rceil$.

Definition 3.14. The *upper* and *lower box-counting dimension* of $E \subset X$ are defined by

$$\overline{\dim}_B E := \limsup_{\delta \downarrow 0} \frac{\log N_\delta(E)}{-\log \delta} \tag{3.11}$$

and

$$\underline{\dim}_B E := \liminf_{\delta \downarrow 0} \frac{\log N_\delta(E)}{-\log \delta}, \tag{3.12}$$

respectively. If $\overline{\dim}_B E = \underline{\dim}_B E$, we call the common value the *box-counting dimension* (also *Minkowski dimension*) of E, and write

$$\dim_B E := \lim_{\delta \downarrow 0} \frac{\log N_\delta(E)}{-\log \delta}. \tag{3.13}$$

Example 3.15. Let $E = \mathbb{Q} \cap [0,1]$ (or any other countable dense subset of $[0,1]$). Then $N_\delta(E) = \lceil 1/\delta \rceil$ for every $\delta > 0$, and so

$$\dim_B E = \lim_{\delta \downarrow 0} \frac{\log \lceil 1/\delta \rceil}{-\log \delta} = 1,$$

whereas $\dim_H E = 0$ by Proposition 3.12(ii). This shows that box-counting dimension and Hausdorff dimension can sometimes dramatically disagree, and also that box-counting dimension fails to satisfy properties (ii) and (iii) of Proposition 3.12.

Example 3.16. Let $E = \{0, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$. We hardly think of E as a fractal, yet $\dim_B E = \frac{1}{2}$. (For the calculation, see Falconer, Example 3.5.)

Despite these shortcomings, box-counting dimension does have its uses. For many fractals of interest, Hausdorff and box-counting dimensions do agree. In general, we have:

Proposition 3.17. *For any $E \subset X$,*

$$\dim_H E \leq \underline{\dim}_B E \leq \overline{\dim}_B E. \tag{3.14}$$

Proof. Only the first inequality needs proof. Let $s = \underline{\dim}_B E$ and $\varepsilon > 0$. Then for every $\delta > 0$ there is $0 < \delta_1 < \delta$ such that

$$\frac{\log N_{\delta_1}(E)}{-\log \delta_1} < s + \varepsilon,$$

and so

$$N_{\delta_1}(E) < (1/\delta_1)^{s+\varepsilon}.$$

Thus,

$$\mathcal{H}^{s+\varepsilon}_\delta(E) \leq N_{\delta_1}(E)\delta_1^{s+\varepsilon} < 1.$$

Letting $\delta \downarrow 0$ we obtain $\mathcal{H}^{s+\varepsilon}(E) \leq 1$, and so $\dim_H E \leq s + \varepsilon$. Since ε was arbitrary, $\dim_H E \leq s$. □

Example 3.18. Let C be the ternary Cantor set. Then $\dim_B C = \dim_H C = \frac{\log 2}{\log 3}$. On the one hand, $\underline{\dim}_B C \geq \dim_H C$ by Proposition 3.17. On the other hand, given $\delta > 0$, let $n \in \mathbb{N}$ such that $3^{-n} \leq \delta < 3^{-n+1}$. Since C is covered by 2^n intervals of diameter $3^{-n} \leq \delta$, $N_\delta(C) \leq 2^n$. Moreover, $-\log \delta > (n-1)\log 3$, and so

$$\overline{\dim}_B C = \limsup_{\delta \downarrow 0} \frac{\log N_\delta(C)}{-\log \delta} \leq \lim_{n\to\infty} \frac{n \log 2}{(n-1)\log 3} = \frac{\log 2}{\log 3}.$$

In general, both inequalities in (3.14) may be strict.

Exercise 3.19. Construct a set $F \subset \mathbb{R}$ for which $\underline{\dim}_B F < \overline{\dim}_B F$. (See Falconer, exercise 3.8 for a hint.)

Constructing a set F with $\dim_H F < \underline{\dim}_B F$ is rather harder. It involves showing that there is a more efficient cover of F by sets of widely different diameters than by sets of (roughly) the same diameter. Nonetheless, such sets F do exist, and are in fact prevalent in fractal geometry.

3.3 Iterated function systems

From now on, we take $X = \mathbb{R}^n$ for some $n \in \mathbb{N}$, with the usual (Euclidian) metric $d(x, y) = |x - y| := \left(\sum_{i=1}^n (x_i - y_i)^2\right)^{1/2}$.

Definition 3.20. A function $S : \mathbb{R}^n \to \mathbb{R}^n$ is a *contraction* if there is a constant $0 < c < 1$ such that

$$|S(x) - S(y)| \leq c|x - y| \qquad \text{for all } x, y \in \mathbb{R}^n. \tag{3.15}$$

Note that a contraction is in particular Lipschitz continuous.

Definition 3.21. A finite set of contractions $\{S_1, \ldots, S_m\}$ on $\mathbb{R}^n$ is called an *iterated function system* (IFS). A nonempty compact set $F \subset \mathbb{R}^n$ is called an *attractor* of the IFS $\{S_1, \ldots, S_m\}$ if

$$F = \bigcup_{i=1}^m S_i(F). \tag{3.16}$$

Informally, (3.16) says that F is made up of finitely many smaller sets that look (roughly) like the set F itself. Many sets that are considered fractals are the attractor of an IFS.

Example 3.22. In $\mathbb{R}$, let $S_1(x) = x/3$ and $S_2(x) = (2+x)/3$. Then the ternary Cantor set C is an attractor of the IFS $\{S_1, S_2\}$.

Our first goal is to show that every IFS has a unique attractor. To do this, we need some additional definitions. Let $\mathcal{K}$ denote the collection of all nonempty compact subsets of $\mathbb{R}^n$. We will turn $\mathcal{K}$ into a metric space as follows. For $A \in \mathcal{K}$ and $\delta > 0$, define the *δ-neighborhood* of A by

$$A_\delta := \{x \in \mathbb{R}^n : |x - a| \le \delta \text{ for some } a \in A\}. \tag{3.17}$$

Definition 3.23. The *Hausdorff metric* on $\mathcal{K}$ is the metric ρ given by

$$\rho(A, B) := \inf\{\delta > 0 : A \subset B_\delta \text{ and } B \subset A_\delta\}, \qquad A, B \in \mathcal{K}. \tag{3.18}$$

Exercise 3.24. Check that ρ is indeed a metric!

Lemma 3.25. *Let S be a contraction on $\mathbb{R}^n$. Then $S(B(0,r)) \subset B(0,r)$ for every sufficiently large r, where $B(0,r) := \{x \in \mathbb{R}^n : |x| \le r\}$.*

Proof. Exercise. (*Hint*: if c satisfies (3.15), take $r \ge (1-c)^{-1}|S(0)|$.) □

For the following theorem, recall that any continuous image of a compact set is compact, and any finite union of compact sets is again compact.

Theorem 3.26. *Every IFS $\{S_1, \dots, S_m\}$ has a unique attractor F. Moreover, if we define the map $S : \mathcal{K} \to \mathcal{K}$ by*

$$S(E) := \bigcup_{i=1}^m S_i(E), \qquad E \in \mathcal{K}, \tag{3.19}$$

then

$$F = \bigcap_{k=1}^\infty S^k(E) \tag{3.20}$$

for any set $E \in \mathcal{K}$ such that $S_i(E) \subset E$ for each i; where S^k denotes the kth iterate of S: $S^k(E) = S(S^{k-1}(E))$ for $k \ge 2$.

Proof. Let $c_1, \dots, c_m$ be the Lipschitz constants of $S_1, \dots, S_m$, respectively, so

$$|S_i(x) - S_i(y)| \le c_i |x - y|, \qquad i = 1, \dots, m. \tag{3.21}$$

Recall that $c_i < 1$. We show first that

$$\rho(S(A), S(B)) \le \left(\max_{1 \le i \le m} c_i\right) \rho(A, B), \qquad A, B \in \mathcal{K}. \tag{3.22}$$

If $S_i(B) \subset (S_i(A))_\delta$ for each i, then $S(B) = \bigcup_{i=1}^m S_i(B) \subset (\bigcup_{i=1}^m S_i(A))_\delta = ((S(A))_\delta$ (check!). This implies that

$$\rho(S(A), S(B)) \leq \max_{1 \leq i \leq m} \rho(S_i(A), S_i(B)).$$

By (3.21), $\rho(S_i(A), S_i(B)) \leq c_i \rho(A, B)$. (Check!) Thus, we have (3.22).

Now let $E \in \mathcal{K}$ such that $S_i(E) \subset E$ for each i. Such a set E exists by Lemma 3.25. Then $S(E) \subset E$, so $\{S^k(E)\}$ is a decreasing sequence of nonempty compact sets, which has a nonempty compact intersection $F = \bigcap_{k=1}^\infty S^k(E)$. Since $S^k(E)$ is decreasing, $S^k(E) \to F$ in the Hausdorff metric. But (3.22) says that $S : \mathcal{K} \to \mathcal{K}$ is (Lipschitz) continuous, so

$$S(F) = S\left(\lim_{k \to \infty} S^k(E)\right) = \lim_{k \to \infty} S\big(S^k(E)\big) = \lim_{k \to \infty} S^{k+1}(E) = F.$$

Thus, F is an attractor of the IFS. The attractor is unique: If A and B are two attractors of the IFS, then $S(A) = A$ and $S(B) = B$, so by (3.22),

$$\rho(A, B) \leq \left(\max_{1 \leq i \leq m} c_i\right) \rho(A, B).$$

Since $\max_i c_i < 1$, this implies $\rho(A, B) = 0$, and hence, $A = B$. □

Observe that (3.20) provides a practical way to approximate the fractal F: Start with a suitable compact set E for which $S_i(E) \subset E$ for each i (say, a large ball or square), and then iterate the map S a large finite number of times to obtain $S^k(E)$ for large k. Since $S^k(E) \to F$ in the Hausdorff metric, $S^k(E)$ will be a good approximation for F when k is large. We also see that for each point $x \in F$ there is a sequence $(i_1, i_2, \dots)$ with $i_k \in \{1, \dots, m\}$ for each k such that

$$x = x_{i_1 i_2 \dots} := \bigcap_{k=1}^\infty (S_{i_1} \circ S_{i_2} \circ \cdots \circ S_{i_k})(E), \tag{3.23}$$

and this point x is independent of the choice of E. Thus, each $x \in F$ is "coded" by a sequence $(i_1, i_2, \dots)$ in $\{1, \dots, m\}^{\mathbb{N}}$. In general, however, this coding need not be unique.

3.4 Self-similar sets

Definition 3.27. A function $S : \mathbb{R}^n \to \mathbb{R}^n$ is called a *similarity (map)* if there is a constant $c > 0$ such that

$$|S(x) - S(y)| = c|x - y| \qquad \text{for all } x, y \in \mathbb{R}^n. \tag{3.24}$$

If $c < 1$ we call c the *contraction ratio* of S. If F is the attractor of an IFS $\{S_1, \dots, S_m\}$ where all of the S_i are (contracting) similarities, we call F a *self-similar set*.

Note that a map $S : \mathbb{R}^n \to \mathbb{R}^n$ is a similarity if and only if there is a constant $c > 0$, a linear isometry $T : \mathbb{R}^n \to \mathbb{R}^n$ and a vector $b \in \mathbb{R}^n$ such that $S(x) = cT(x) + b$.

Example 3.28 (Koch curve)**.** The attractor of the IFS $\{S_1, S_2, S_3, S_4\}$, where

$$S_1(x,y) = \frac{1}{3}\begin{bmatrix} x \\ y \end{bmatrix}, \qquad S_2(x,y) = \frac{1}{3}\begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 1/3 \\ 0 \end{bmatrix},$$
$$S_3(x,y) = \frac{1}{3}\begin{bmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 1/2 \\ \sqrt{3}/6 \end{bmatrix}, \qquad S_4(x,y) = \frac{1}{3}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 2/3 \\ 0 \end{bmatrix},$$

is called the *Koch curve*. (Look at the transformation of the line segment connecting $(0,0)$ and $(1,0)$.) If we fit three copies of the Koch curve on the sides of an equilateral triangle, we obtain the *Koch snowflake*.

Example 3.29 (Sierpinski triangle)**.** The attractor of the IFS $\{S_1, S_2, S_3\}$, where

$$S_1(x,y) = \frac{1}{2}\begin{bmatrix} x \\ y \end{bmatrix}, \qquad S_2(x,y) = \frac{1}{2}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}, \qquad S_3(x,y) = \frac{1}{2}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 1/4 \\ \sqrt{3}/4 \end{bmatrix},$$

is called the *Sierpinski triangle*. (Look at the transformation of the equilateral triangle with vertices $(0,0)$, $(1,0)$ and $(1/2, \sqrt{3}/2)$.)

Definition 3.30. An IFS $\{S_1, \ldots, S_m\}$ satisfies the *open set condition* (OSC) if there is a nonempty bounded open set $V \subset \mathbb{R}^n$ such that

(1) $S_i(V) \subset V$ for $i = 1, \ldots, m$; and

(2) $S_i(V) \cap S_j(V) = \emptyset$ for all $i \neq j$.

The open set condition ensures that the m parts of F in (3.16) do not overlap "too much".

Example 3.31. The IFS from Example 3.22 satisfies the OSC with $U = (0,1)$. A related example is: $S_1(x) = x$ and $S_2(x) = (1+x)/3$. Then $\{S_1, S_2\}$ satisfies the OSC with $U = (0,1)$, but not with any strictly larger open set. What does the attractor look like?

Theorem 3.32. *Let F be the attractor of the IFS $\{S_1, \ldots, S_m\}$, where for each $i = 1, \ldots, m$, S_i is a similarity with contraction ratio c_i. Assume the IFS satisfies the OSC. Then $\dim_H F = \dim_B F = s$, where s is the unique positive root of the equation*

$$\sum_{i=1}^{m} c_i^s = 1. \tag{3.25}$$

Moreover, $0 < \mathcal{H}^s(F) < \infty$.

Preparation for the proof. To prove the lower bound $\dim_H F \geq s$, we will use the mass distribution principle (Lemma 3.13). The standard procedure for doing so is to first define a mass distribution μ on the "code space" $\mathcal{I} := \{1, \ldots, m\}^{\mathbb{N}}$ and then to lift μ to a mass distribution on the fractal F via the projection map

$$\pi((i_1, i_2, \ldots)) := x_{i_1 i_2 \ldots} = \bigcap_{k=1}^{\infty} S_{i_1} \circ \cdots \circ S_{i_k}(E), \tag{3.26}$$

where the compact set E is as in Theorem 3.26. This requires a suitable σ-algebra on $\mathcal{I}$, defined as follows. Let $\mathcal{C}$ denote the collection consisting of $\emptyset$ and all *cylinder sets* in $\mathcal{I}$; that is, sets of the form

$$I_{i_1,\dots,i_k} := \{(j_1, j_2, \dots) \in \mathcal{I} : (j_1, \dots, j_k) = (i_1, \dots, i_k)\}, \quad k \in \mathbb{N},\ i_1, \dots, i_k \in \{1, \dots, m\},$$

where for completeness we put $I_\emptyset := \mathcal{I}$. One checks easily that $\mathcal{C}$ is a semiring. (That is, $\mathcal{C}$ is closed under finite intersections, and if $I, J \in \mathcal{C}$, then $I \backslash J$ is a finite union of sets in $\mathcal{C}$.) Suppose that we now have a set function μ defined on $\mathcal{C}$ such that $\mu(\emptyset) = 0$, $\mu(\mathcal{I}) = 1$, and μ satisfies the *consistency condition*

$$\sum_{j=1}^{m} \mu(I_{i_1,\dots,i_k,j}) = \mu(I_{i_1,\dots,i_k}), \quad \text{for all } k \in \mathbb{N}_0 \text{ and } i_1, \dots, i_k, j \in \{1, \dots, m\}. \tag{3.27}$$

One may check that μ is then a premeasure on $\mathcal{C}$. (Recall that this requires verifying that if $I_1, I_2, \dots$ are in $\mathcal{C}$ and $I := \bigcup_{j=1}^{\infty} I_j \in \mathcal{C}$, then $\mu(I) = \sum_{j=1}^{\infty} \mu(I_j)$. This is a routine exercise.) Thus, by Carathéodory's extension theorem, μ extends uniquely to a probability measure on the σ-algebra $\mathcal{M} := \sigma(\mathcal{C})$, which we again denote by μ.

Next, we wish to define a mass distribution $\tilde{\mu}$ on $\mathbb{R}^n$ by the prescription

$$\tilde{\mu}(B) := \mu(\pi^{-1}(B)) = \mu\{(i_1, i_2, \dots) \in \mathcal{I} : \pi((i_1, i_2, \dots)) \in B\}. \tag{3.28}$$

This, however, makes sense only if the set on the right lies in $\mathcal{M}$ whenever $B \in \mathcal{B}(\mathbb{R}^n)$. We check this as follows. First, verify that the collection $\mathcal{S} := \{B \subset \mathbb{R}^n : \pi^{-1}(B) \in \mathcal{M}\}$ is a σ-algebra, so it is sufficient to show that $\mathcal{S}$ contains all open sets in $\mathbb{R}^n$. Let $O \subset \mathbb{R}^n$ be open, and let $(i_1, i_2, \dots) \in \mathcal{I}$. Since the sets $S_{i_1} \circ \cdots \circ S_{i_k}(E)$ are decreasing in k, it is not difficult to see that

$$\pi(i_1, i_2, \dots) \in O \quad \Leftrightarrow \quad \exists k \ \ni S_{i_1} \circ \cdots \circ S_{i_k}(E) \subset O.$$

As a result,

$$\pi^{-1}(O) = \bigcup_{k=1}^{\infty} \bigcup \{I_{i_1,\dots,i_k} : S_{i_1} \circ \cdots \circ S_{i_k}(E) \subset O\},$$

a countable union of cylinder sets. Thus, $\pi^{-1}(O) \in \mathcal{S}$. Moreover, since $\pi(\mathcal{I}) = F$, we have $\tilde{\mu}(F) = 1$, so $\tilde{\mu}$ is a mass distribution on F.

Lemma 3.33 (Geometry Lemma). *Let $\{V_i\}$ be a collection of disjoint open subsets of $\mathbb{R}^n$ such that each V_i contains a ball of radius $a_1 r$ and is contained in a ball of radius $a_2 r$. Then any open ball with radius r intersects at most $(1 + 2a_2)^n a_1^{-n}$ of the closures $\bar{V}_i$.*

Proof. Recall that the volume of a ball with radius r in $\mathbb{R}^n$ is $c_n r^n$, where c_n depends only on n. Let B be a ball with radius r, and let $\tilde{B}$ be the ball with the same center as B but with radius $(1 + 2a_2)r$. If $\bar{V}_i$ intersects B, then $\bar{V}_i \subset \tilde{B}$. Let q be the number of sets $\bar{V}_i$ that intersect B. Adding volumes we obtain $qc_n(a_1 r)^n \leq \mathrm{Vol}(\tilde{B}) = c_n(1 + 2a_2)^n r^n$, and hence, $q \leq (1 + 2a_2)^n a_1^{-n}$. $\square$

Proof of Theorem 3.32. Let s be given by (3.25). We first show that $\dim_H F \geq s$. Let $\mathcal{I}_k := \{1, \ldots, m\}^k$. For $A \subset \mathbb{R}^n$ and $(i_1, \ldots, i_k) \in \mathcal{I}_k$, put $A_{i_1,\ldots,i_k} := S_{i_1} \circ \cdots \circ S_{i_k}(A)$. Using (3.16) repeatedly, we have

$$F = \bigcup_{\mathcal{I}_k} F_{i_1,\ldots,i_k}. \tag{3.29}$$

Now use (3.25) to check that $\sum_{\mathcal{I}_k} |F_{i_1,\ldots,i_k}|^s = |F|^s$. Given $\delta > 0$, choose k so that $|F_{i_1,\ldots,i_k}| \leq (\max_{1\leq i\leq m} c_i)^k |F| \leq \delta$. Then by (3.29), $\mathcal{H}^s_\delta(F) \leq |F|^s$. Since δ was arbitrary, it follows that $\mathcal{H}^s(F) \leq |F|^s$.

For the lower bound, define a mass distribution on $\mathcal{I} = \{1, \ldots, m\}^{\mathbb{N}}$ by

$$\mu(I_{i_1,\ldots,i_k}) = (c_{i_1} \ldots c_{i_k})^s, \qquad k \in \mathbb{N}_0, \quad (i_1, \ldots, i_k) \in \mathcal{I}_k.$$

(Note that by (3.25), μ satisfies the consistency condition (3.27).) Let $\tilde{\mu}(B) = \mu(\pi^{-1}(B))$ for $B \in \mathcal{B}(\mathbb{R}^n)$, so $\tilde{\mu}$ is a mass distribution on F.

Let V be an open set satisfying the OSC for the IFS $\{S_1, \ldots, S_m\}$. Then $\bar{V}$ satisfies the conditions for E in Theorem 3.26, so the decreasing sequence $\{S^k(\bar{V})\}$ converges to F. In particular, $\bar{V} \supset F$ and $\bar{V}_{i_1,\ldots,i_k} \supset F_{i_1,\ldots,i_k}$ for all $(i_1, \ldots, i_k) \in \mathcal{I}_k$.

Let B be a ball of radius $r < 1$. Write $c_{\min} := \min_{1\leq i\leq m} c_i$. For each sequence $(i_1, i_2, \ldots) \in \mathcal{I}$, there is an integer k such that

$$c_{\min} r \leq c_{i_1} c_{i_2} \cdots c_{i_k} \leq r. \tag{3.30}$$

Choose the smallest such k to obtain a finite sequence $(i_1, \ldots, i_k)$. Let $\mathcal{Q}$ be the set of all (finite) sequences obtained in this way. Note that no sequence in $\mathcal{Q}$ extends any other sequence in $\mathcal{Q}$. Since V satisfies the OSC, the sets $V_1, \ldots, V_m$ are disjoint, and it follows by iteration that the collection of open sets $\{V_{i_1,\ldots,i_k} : (i_1, \ldots, i_k) \in \mathcal{Q}\}$ is disjoint. Moreover,

$$F \subset \bigcup_{\mathcal{Q}} F_{i_1,\ldots,i_k} \subset \bigcup_{\mathcal{Q}} \bar{V}_{i_1,\ldots,i_k}. \tag{3.31}$$

Since V is open and bounded, we can choose a_1 and a_2 so that V contains a ball of radius a_1 and is contained in a ball of radius a_2. Then for all $(i_1, \ldots, i_k) \in \mathcal{Q}$, $V_{i_1,\ldots,i_k}$ contains a ball of radius $c_{i_1} \cdots c_{i_k} a_1$ and hence one of radius $c_{\min} a_1 r$, and is contained in a ball of radius $c_{i_1} \cdots c_{i_k} a_2$ and hence in one of radius $a_2 r$, by (3.30). Let

$$\mathcal{Q}_1 := \{(i_1, \ldots, i_k) \in \mathcal{Q} : \bar{V}_{i_1,\ldots,i_k} \cap B \neq \emptyset\}.$$

By Lemma 3.33,

$$\#\mathcal{Q}_1 \leq C := (1 + 2a_2)^n (a_1 c_{\min})^{-n}.$$

Note that $F \cap B \subset \bigcup_{\mathcal{Q}_1} \bar{V}_{i_1,\ldots,i_k}$, so if $x_{i_1 i_2 \ldots} \in F \cap B$, then there is an integer k such that $(i_1, \ldots, i_k) \in \mathcal{Q}_1$. It follows that

$$\begin{aligned}
\tilde{\mu}(B) = \tilde{\mu}(F \cap B) &= \mu\{(i_1, i_2, \ldots) : x_{i_1 i_2 \ldots} \in F \cap B\} \\
&\leq \mu\left(\bigcup_{\mathcal{Q}_1} I_{i_1,\ldots,i_k}\right) \leq \sum_{\mathcal{Q}_1} \mu(I_{i_1,\ldots,i_k}) \\
&= \sum_{\mathcal{Q}_1} (c_{i_1} \cdots c_{i_k})^s \leq \sum_{\mathcal{Q}_1} r^s \leq C r^s.
\end{aligned}$$

Finally, since any set U is contained in a ball of radius $|U|$, we have $\tilde{\mu}(U) \le C|U|^s$, so by the mass distribution principle, $\mathcal{H}^s(F) > 0$ and $\dim_H F \ge s$.

It remains to show that $\overline{\dim}_B F \le s$; by (3.14), this will complete the proof. Let $\mathcal{Q}$ be the set used in the above argument. Using (3.25) and induction it follows (exercise!) that $\sum_{\mathcal{Q}}(c_{i_1}\cdots c_{i_k})^s = 1$. Thus, by (3.30), $\#\mathcal{Q} \le (c_{\min} r)^{-s}$. If $(i_1,\ldots,i_k) \in \mathcal{Q}$, then $|\bar{V}_{i_1,\ldots,i_k}| = c_{i_1}\cdots c_{i_k}|\bar{V}| \le r|\bar{V}|$, so F can be covered by $(c_{\min} r)^{-s}$ sets of diameter $r|\bar{V}|$ for each $r < 1$. Therefore,

$$\overline{\dim}_B F \le \lim_{r\downarrow 0} \frac{\log(c_{\min} r)^{-s}}{-\log r|\bar{V}|} = s,$$

as required. □

Example 3.34. The Koch curve has Hausdorff and box-counting dimension $\frac{\log 4}{\log 3}$: The IFS from Example 3.28 satisfies the OSC with V being the interior of the triangle with vertices $(0,0)$, $(1,0)$ and $(1/2, \sqrt{3}/6)$. By Theorem 3.32, $\dim_H F = \dim_B F = s$, where $\sum_{i=1}^4 c_i^s = 4(\frac{1}{3})^s = 1$, and this gives $s = \frac{\log 4}{\log 3}$.

Example 3.35. The Sierpinski triangle from Example 3.29 has Hausdorff and box-counting dimension $\frac{\log 3}{\log 2}$. Check this.

Example 3.36. Fix $0 < a < 1/2$, and consider the attractor F of the IFS of 5 transformations which map the unit square $Q := [0,1]^2$ onto the 5 squares (4 of side length a, one of side length $1-2a$) which result when Q is cut by the lines $x = a$, $x = 1-a$, $y = a$ and $y = 1-a$. Taking V to be the interior of Q, we find that $\dim_H F = \dim_B F = s$, where $4a^s + (1-2a)^s = 1$. **Exercise:** Solve this algebraically when (i) $a = 1/3$; (ii) $a = 1/4$.

3.5 Graphs of functions

Many important and interesting fractals arise as graphs of continuous functions. (Consider stock prices, brain waves, etc.) In this section we will develop some tools for computing the dimension of such graphs, and focus on a particular example: the celebrated Weierstrass function.

For a function $f : [a,b] \to \mathbb{R}$, let

$$\mathrm{Graph}(f) := \{(t, f(t) : a \le t \le b\}$$

denote the graph of f.

Exercise 3.37. Show that if f is of bounded variation, then

$$\dim_H \mathrm{Graph}(f) = \dim_B \mathrm{Graph}(f) = 1.$$

(*Hint*: Use Lemma 3.40 below.)

It is a consequence of the Baire category theorem that most continuous functions are of unbounded variation (in fact, nowhere differentiable). The best known example is the *Weierstrass function*

$$f(t) = \sum_{k=1}^{\infty} \lambda^{(s-2)k} \sin(\lambda^k t), \qquad 0 \le t \le 1, \tag{3.32}$$

where $\lambda > 1$ and $1 < s < 2$. It's easy to see that f is the uniform limit of continuous functions (the partial sums of the series), and hence continuous. It is considerably harder to prove that f is nowhere differentiable; the first full proof of this fact was given by G. H. Hardy, and many different proofs now exist. Note that instead of the sine function, the cosine may be used to the same effect.

Definition 3.38. A function $f : [a,b] \to \mathbb{R}$ is *Hölder continuous* of *exponent* α if there is a constant c such that

$$|f(t) - f(u)| \le c|t-u|^\alpha \qquad \text{for all } t, u \in [a,b]. \tag{3.33}$$

Note that the Hölder property with exponent $\alpha = 1$ is precisely the Lipschitz property. The interesting case is when $0 < \alpha < 1$.

Definition 3.39. The *oscillation* (or *maximum range*) of a function f on an interval $I = [t_1, t_2]$ is defined by

$$\operatorname{osc}_f(I) := \sup_{t,u\in I} |f(t) - f(u)|. \tag{3.34}$$

Lemma 3.40. *Let $f : [0,1] \to \mathbb{R}$ be continuous. For $0 < \delta < 1$, let N_δ be the number of δ-*mesh *squares (i.e. squares of the form $[i\delta, (i+1)\delta] \times [j\delta, (j+1)\delta]$ with $i, j \in \mathbb{Z}$) that intersect* Graph(f)*. Put $m := \lceil 1/\delta \rceil$. Then*

$$\frac{1}{\delta}\sum_{i=0}^{m-1} \operatorname{osc}_f[i\delta, (i+1)\delta] \le N_\delta \le 2m + \frac{1}{\delta}\sum_{i=0}^{m-1} \operatorname{osc}_f[i\delta, (i+1)\delta]. \tag{3.35}$$

Proof. Since f is continuous, the number of δ-mesh squares in the column $[i\delta, (i+1)\delta] \times \mathbb{R}$ that intersect Graph(f) is at least $\delta^{-1} \operatorname{osc}_f[i\delta, (i+1)\delta]$, and at most $2 + \delta^{-1} \operatorname{osc}_f[i\delta, (i+1)\delta]$. Summing over all such intervals gives (3.35). □

Proposition 3.41. *Let $f : [0,1] \to \mathbb{R}$ be continuous.*

(i) Suppose f is Hölder continuous with exponent $0 < \alpha \le 1$. Then $\overline{\dim}_B$ Graph$(f) \le 2 - \alpha$ (and hence $\dim_H$ Graph$(f) \le 2 - \alpha$).

(ii) Suppose there are numbers $c > 0$, $\delta_0 > 0$ and $0 < \alpha \le 1$ with the following property: for each $t \in [0,1]$ and $0 < \delta < \delta_0$ there exists u such that $|t - u| \le \delta$ and

$$|f(t) - f(u)| \ge c\delta^\alpha.$$

Then $\underline{\dim}_B$ Graph$(f) \ge 2 - \alpha$.

Proof. (i) Let c be a constant for which (3.33) holds. Then $\operatorname{osc}_f[t_1, t_2] \le c|t_1 - t_2|^\alpha$ for $0 \le t_1, t_2 \le 1$. Given $0 < \delta < 1$, let $m = \lceil 1/\delta \rceil < 1 + \delta^{-1}$. Lemma 3.40 yields

$$N_\delta \le 2m + \delta^{-1} mc\delta^\alpha \le (1 + \delta^{-1})(2 + c\delta^{-1}\delta^\alpha) \le 2(2+c)\delta^{\alpha-2}.$$

It follows that $\overline{\dim}_B$ Graph$(f) \le 2 - \alpha$.

(ii) Let $0 \le t_1 < t_2 \le 1$ such that $|t_1 - t_2| \le 2\delta_0$. Taking $\delta = |t_1 - t_2|/2$ and $t = (t_1 + t_2)/2$ in the hypothesis of (ii), we find that there is $u \in [t_1, t_2]$ such that $|f(t) - f(u)| \ge c\delta^\alpha$. But then

$$\operatorname{osc}_f[t_1, t_2] \ge |f(t) - f(u)| \ge c\left(\frac{|t_1 - t_2|}{2}\right)^\alpha = c_1|t_1 - t_2|^\alpha,$$

where $c_1 = 2^{-\alpha}c$. Hence, for all sufficiently small δ, Lemma 3.40 gives

$$N_\delta \geq \delta^{-1} m c_1 \delta^\alpha \geq c_1 \delta^{\alpha-2},$$

where we used that $m \geq \delta^{-1}$. Since any set of diameter $\leq \delta$ intersects at most four δ-mesh squares, it follows that $\underline{\dim}_B \operatorname{Graph}(f) \geq 2 - \alpha$. $\square$

Theorem 3.42. *For $\lambda > 1$ and $1 < s < 2$, let f be the Weierstrass function as in (3.32). Then, for all sufficiently large λ, $\dim_B \operatorname{Graph}(f) = s$.*

The proof uses the following technical lemma.

Lemma 3.43. *For any $t \in \mathbb{R}$, there is $0 < h < 1$ such that*

$$|\sin(t+h) - \sin t| \geq \frac{2}{25}. \tag{3.36}$$

Proof. We may assume without loss of generality that $-\pi/2 \leq t \leq \pi/2$. If $t \leq 1.16$, we can take $h = \frac{\pi}{2} - 1.16 < 1$, and by the symmetry in the graph of $\sin t$ and the convexity of $\sin t$ on $[-\pi/2, 0]$, we get

$$\begin{aligned}|\sin(t+h) - \sin t| &= \sin(t+h) - \sin t \geq \sin\left(-\frac{\pi}{2} + h\right) - \sin\left(-\frac{\pi}{2}\right) \\ &= \sin(-1.16) + 1 \approx 1 - .9168 > 0.08 = \frac{2}{25}.\end{aligned}$$

On the other hand, if $1.16 < t \leq \pi/2$, then

$$|\sin(2.16) - \sin t| = \sin t - \sin(2.16) > \sin(1.16) - \sin(2.16) > \frac{2}{25}.$$

In both cases, the lemma follows. $\square$

Proof of Theorem 3.42. Given $0 < h < \lambda^{-1}$, let N be the integer such that

$$\lambda^{-(N+1)} \leq h < \lambda^{-N}. \tag{3.37}$$

Then

$$\begin{aligned}|f(t+h) - f(t)| &\leq \sum_{k=1}^{\infty} \lambda^{(s-2)k} \left|\sin(\lambda^k(t+h)) - \sin(\lambda^k t)\right| \\ &\leq \sum_{k=1}^{N} \lambda^{(s-2)k} \lambda^k h + \sum_{k=N+1}^{\infty} 2\lambda^{(s-2)k} \\ &= \sum_{k=1}^{N} \lambda^{(s-1)k} h + \sum_{k=N+1}^{\infty} 2\lambda^{(s-2)k},\end{aligned}$$

where we used the inequality $|\sin u - \sin v| \leq |u - v|$, which follows from the mean value theorem. Summing the geometric series gives

$$|f(t+h) - f(t)| \leq \frac{h\lambda^{(s-1)N}}{1 - \lambda^{1-s}} + \frac{2\lambda^{(s-2)(N+1)}}{1 - \lambda^{s-2}}. \tag{3.38}$$

Now (3.37) implies $\lambda^N < h^{-1} \le \lambda^{N+1}$, and since $s-1>0$ and $s-2<0$, we can estimate (3.38) further by

$$|f(t+h)-f(t)| \le ch^{2-s},$$

where c is independent of h. Hence, f is Hölder continuous with exponent $2-s$, and therefore, $\overline{\dim}_B \operatorname{Graph}(f) \le s$ by Proposition 3.41(i).

Similarly, using that $h < \lambda^{-N}$, we can show

$$\begin{aligned}
\Big| f(t+h) - f(t) - \lambda^{(s-2)N}\big(\sin\lambda^N(t+h) - \sin\lambda^N t\big)\Big| \\
&\le \frac{\lambda^{(s-2)N-s+1}}{1-\lambda^{1-s}} + \frac{2\lambda^{(s-2)(N+1)}}{1-\lambda^{s-2}} \\
&= \left(\frac{1}{\lambda^{s-1}-1} + \frac{2}{\lambda^{2-s}-1}\right)\lambda^{(s-2)N}.
\end{aligned} \tag{3.39}$$

Now assume λ is large enough so that the expression in large parentheses in the last line of (3.39) is less than $1/25$. Given $0<\delta<1$, let N be the integer such that $\lambda^{-N} \le \delta < \lambda^{-(N-1)}$. For each t, Lemma 3.43 allows us to choose $0 < h < \lambda^{-N} \le \delta$ such that

$$\left|\sin\lambda^N(t+h) - \sin\lambda^N t\right| > \frac{2}{25}, \tag{3.40}$$

so by (3.39),

$$|f(t+h)-f(t)| \ge \frac{2}{25}\lambda^{(s-2)N} - \frac{1}{25}\lambda^{(s-2)N} = \frac{1}{25}\lambda^{(s-2)N} \ge \frac{1}{25}\lambda^{s-2}\delta^{2-s}.$$

Thus, by Proposition 3.41(ii), $\underline{\dim}_B \operatorname{Graph}(f) \ge s$. $\square$

Remark 3.44. Falconer (p. 163) claims a lower bound of $\frac{1}{10}$ in (3.40), but this does not appear possible.

It can in fact be proved (with some extra work) that the conclusion of Theorem 3.42 holds for all $\lambda > 1$. It is widely conjectured that even $\dim_H \operatorname{Graph}(f) = s$. This was only recently proved (by much more sophisticated techniques!) to be true for $s \in (s_0, 2)$ where $1 < s_0 < 2$ depends on λ, but remains open for values of s outside this interval.