

# Multi-scale properties of continued fraction sets

ALEX RUTAR

ABSTRACT. We survey the dimension theory of sets of real numbers with regular continued fraction expansion restricted to a non-empty and possibly infinite subset  $\mathcal{D} \subset \mathbb{N}$ .

## CONTENTS

<b>1</b>	<b>Continued fractions with restricted digits</b>	<b>2</b>
1.1	Regular continued fractions . . . . .	2
1.2	Infinitely-generated self-conformal sets . . . . .	2
1.3	Hausdorff and box dimensions . . . . .	4
1.4	Overview . . . . .	5
1.5	A note on implicit constants . . . . .	6
<b>2</b>	<b>Infinitely-generated self-similar sets</b>	<b>6</b>
2.1	Self-similar iterated function systems . . . . .	6
2.2	Hausdorff dimension of the attractor . . . . .	7
2.3	Upper box dimension of the attractor . . . . .	9
<b>3</b>	<b>Lower box dimension and asymptotics for covering numbers</b>	<b>11</b>
3.1	Branching functions . . . . .	12
3.2	A candidate asymptotic formula . . . . .	14
3.3	Technical results concerning level sets of cylinders . . . . .	15
3.4	Proof of the asymptotic formula . . . . .	16
3.5	The asymptotic formula as an operator . . . . .	17
3.6	Consequences for the lower box dimension . . . . .	19
<b>4</b>	<b>Various loose ends</b>	<b>21</b>
4.1	Infinitely-generated self-conformal sets . . . . .	22
4.2	Alternatives to the fixed point set . . . . .	25
4.3	Sharpness of the bounds on the lower box dimensions . . . . .	26
4.4	Intermediate dimensions . . . . .	29
4.5	Some open problems . . . . .	31
	<b>Acknowledgements</b>	<b>32</b>
	<b>References</b>	<b>32</b>

## 1. CONTINUED FRACTIONS WITH RESTRICTED DIGITS

**1.1. Regular continued fractions.** Let  $x \in \mathbb{R}$ . A (regular) continued fraction expansion of  $x$  is representation of the form

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{\ddots}}}}$$

where  $a_n \in \mathbb{Z}$  and  $a_n \geq 1$  for  $n \neq 0$ . We will write  $x = [a_0; a_1, a_2, a_3, \dots]$ . It is easy to check that the continued fraction expansion is infinite if and only if  $x$  is irrational, in which case the continued fraction expansion is also unique. If  $x$  is rational then there are two finite continued fraction expansions.

Henceforth, we only consider numbers  $x \in \mathbb{I} := (0, 1) \setminus \mathbb{Q}$  and write  $x = [a_1, a_2, a_3, \dots] = [0; a_1, a_2, a_3, \dots]$ . We are primarily interested in the following subsets of  $\mathbb{I}$ . For an arbitrary non-empty (and possibly infinite) subset  $\mathcal{D} \subset \mathbb{N}$ , we write

$$\Lambda_{\mathcal{D}} = \{x = [a_1, a_2, a_3, \dots] : a_n \in \mathcal{D} \text{ for all } n \in \mathbb{N}\}.$$

This is the set of numbers for which the continued fraction expansion only contains digits in  $\mathcal{D}$ . A classical example is the case  $\mathcal{D} = \{1, \dots, \tau\}$  for some  $\tau \in \mathbb{N}$ , in which case  $\Lambda_{\mathcal{D}}$  is the set of  $\tau$ -badly approximable numbers.

The continued fraction expansion can be obtained in a natural way using the *Gauss map*. The Gauss map is the transformation  $\Gamma: \mathbb{I} \rightarrow \mathbb{I}$  defined by the rule

$$(1.1) \quad \Gamma(x) := \left\{ \frac{1}{x} \right\} = \frac{1}{x} - \left[ \frac{1}{x} \right].$$

If  $x = [a_1, a_2, a_3, a_4, \dots]$ , then  $\Gamma(x) = [a_2, a_3, a_4, \dots]$ , so  $\Gamma$  acts like a left shift on the continued fraction expansion of  $x$ . In particular, the sets  $\Lambda_{\mathcal{D}}$  are *forward invariant* for the Gauss map:  $\Gamma(\Lambda_{\mathcal{D}}) = \Lambda_{\mathcal{D}}$ . These are certainly not the only invariant sets! However, they are perhaps the simplest family of invariant sets.

**1.2. Infinitely-generated self-conformal sets.** We now see the representation of the sets  $\Lambda_{\mathcal{D}}$  which will be of primary use in this survey: as the attractor of an infinitely-generated self-conformal IFS.

The Gauss map is far from invertible, in that  $\Gamma^{-1}(y)$  is a countably infinite set for all  $y$ . On the other hand, for  $b \in \mathbb{N}$ , if we restrict  $\Gamma$  to the interval  $J_b := ((b+1)^{-1}, b^{-1}) \cap \mathbb{I}$ , then  $\Gamma$  is invertible and  $\Gamma(J_b) = \mathbb{I}$ . We call the restriction  $\Gamma|_{J_b}$  a *branch* of  $\Gamma$ .

We can invert each branch to obtain a map  $\phi_b = \Gamma_{J_b}^{-1}: \mathbb{I} \rightarrow J_b \subset \mathbb{I}$ . Concretely,  $\phi_b(x) = (b+x)^{-1}$ , which we can think of as a map from  $[0, 1]$  to  $[0, 1]$ . Then the set  $\Lambda_{\mathcal{D}}$  can equivalently be obtained as the maximal invariant set for the family of

maps  $\{\phi_b\}_{b \in \mathcal{D}}$ : that is, the union over all sets  $E \subset [0, 1]$  for which

$$E = \bigcup_{b \in \mathcal{D}} \phi_b(E).$$

This allows us to place the study of the sets  $\Lambda_{\mathcal{D}}$  in a somewhat more general context.

**Definition 1.1.** Suppose  $(X, d)$  is a compact metric space and  $\mathcal{I}$  is a countable index set. A countable family of maps  $\{f_i\}_{i \in \mathcal{I}}$  from  $X$  to itself is (*uniformly contracting*) if there is a constant  $\xi < 1$  such that each  $f_n$  is a Lipschitz contraction with constant  $\xi$ :

$$d(f_n(x), f_n(y)) \leq \xi d(x, y) \quad \text{for all } x, y \in X.$$

Similarly to the case for the sets  $\Lambda_{\mathcal{D}}$ , given a contracting family of maps  $\{f_i\}_{i \in \mathcal{I}}$ , there is a unique maximal invariant set

$$\Lambda = \bigcup_{i \in \mathcal{I}} f_i(\Lambda).$$

We call the family  $\{f_i\}_{i \in \mathcal{I}}$  an *iterated function system* (or *IFS* for short) and the invariant set  $\Lambda$  the *attractor*.

An IFS  $\{f_i\}_{i \in \mathcal{I}}$  comes equipped with a surjective *coding map*  $\pi: \mathcal{I}^{\mathbb{N}} \rightarrow \Lambda$  defined by the rule

$$\{\pi((i_n)_{n=1}^{\infty})\} = \lim_{n \rightarrow \infty} f_{i_1} \circ \cdots \circ f_{i_n}(X).$$

The contracting property and compactness of  $X$  ensures that the function  $\pi$  is well-defined, and it is also easy to see that  $\pi$  is continuous when  $\mathcal{I}^{\mathbb{N}}$  is equipped with the product topology. In particular, we alternatively note that  $\Lambda = \pi(\mathcal{I}^{\mathbb{N}})$ . For example, in the special case of the IFS  $\{\phi_b\}_{b \in \mathcal{D}}$ ,  $\pi$  is the map which sends a sequence  $(a_n)_{n=1}^{\infty}$  to the corresponding continued fraction  $[a_1, a_2, a_3, \dots]$ , and  $\pi \circ \sigma = \Gamma \circ \pi$  where  $\sigma: \mathcal{D}^{\mathbb{N}} \rightarrow \mathcal{D}^{\mathbb{N}}$  is the left shift map.

Studying general attractors of infinite iterated function systems is rather intractable. Later on, we will make more assumptions on the maps  $\{f_i\}_{i \in \mathcal{I}}$ . These assumptions will be satisfied by the IFS  $\{\phi_b\}_{b \in \mathcal{D}}$ .

**Remark 1.2.** The maps  $\phi_b$  are not literally contractions since  $\phi_b'(1) = 1$ . However, the maps  $\phi_b$  for  $b \geq 2$  are uniformly contracting, and we can replace the map  $\phi_1$  with the set of maps  $\{\phi_1 \circ \phi_b : b \in \mathcal{D}\}$  (which is a uniformly contracting family), and the resulting invariant set is unchanged by this operation. We will ignore this problem and just speak of the maps  $\{\phi_b\}_{b \in \mathcal{D}}$  as uniformly contracting even when  $1 \in \mathcal{D}$ .

**Remark 1.3.** Given a finite set of contracting Lipschitz maps  $\{f_1, \dots, f_m\}$ , the transformation  $K \mapsto f_1(K) \cup \cdots \cup f_m(K)$  maps compact sets to compact sets, and is in fact a Lipschitz contraction on the space of compact subsets of  $X$  with the Hausdorff metric, in which case the maximal invariant set  $\Lambda = f_1(\Lambda) \cup \cdots \cup f_m(\Lambda)$  is the unique non-empty compact set with this property.

However, if there are infinitely many maps, then  $\Lambda$  need not be compact. In the general countably infinite case, the set  $\Lambda$  is necessarily analytic, being a continuous image of the Polish space  $\mathcal{I}^{\mathbb{N}}$  under the coding map  $\pi$ . However, as far as I am aware, it is not known if  $\Lambda$  is necessarily Borel. In the special case that the images  $f_n(X)$  do not overlap too much (say, each  $x \in X$  is only contained in finitely many images  $f_n(X)$ ), then  $\Lambda$  is a  $F_{\sigma\delta}$  subset of  $X$ . This is the case for the sets  $\Lambda_{\mathcal{D}}$ .

It is convenient, when working with repeated compositions, to have some shorthand. We denote by  $\mathcal{I}^* = \bigcup_{n=0}^{\infty} \mathcal{I}^n$  all of the finite words of length  $n$ ; we denote by  $\emptyset$  the unique word of length 0. Elements of  $\mathcal{I}^*$  will be denoted in type-writer font, like  $\mathbf{i}$  or  $\mathbf{j}$ . The space  $\mathcal{I}^*$  is equipped with the operation of concatenation. For  $\mathbf{i} = (i_1, \dots, i_n) \in \mathcal{I}^*$ , we write

$$f_{\mathbf{i}} = f_{i_1} \circ \dots \circ f_{i_n}.$$

We say that  $\mathbf{i}$  is a prefix of  $\mathbf{j}$ , and write  $\mathbf{i} \prec \mathbf{j}$ , if there is a  $\mathbf{k} \in \mathcal{I}^*$  such that  $\mathbf{ik} = \mathbf{j}$ . If  $\mathbf{i} = (i_1, \dots, i_n) \neq \emptyset$ , we also write  $\mathbf{i}^- = (i_1, \dots, i_{n-1})$ ; in other words,  $\mathbf{i}^-$  is the unique prefix of  $\mathbf{i}$  of length  $n - 1$ .

**1.3. Hausdorff and box dimensions.** We now briefly introduce the notions of fractal dimension which will be the most important for us in this note. The most fundamental notion is the *Hausdorff dimension* which may be defined as follows. For  $0 < \delta \leq \infty$  and  $s \geq 0$ , we define the subadditive set function

$$\mathcal{H}_{\delta}^s(E) = \inf \left\{ \sum_n (\text{diam } A_n)^s : E \subset \bigcup_{n=1}^{\infty} A_n, \text{diam } A_n \leq \delta \right\}.$$

The *Hausdorff measure*  $\mathcal{H}^s(E) := \lim_{\delta \rightarrow 0} \mathcal{H}_{\delta}^s(E)$  is a metric outer measure and defines a Borel measure on  $\mathbb{R}^d$ . The quantity  $\mathcal{H}_{\infty}^s(E)$  is called *Hausdorff content*. Hausdorff measure (or equivalently Hausdorff content) can be used to define *Hausdorff dimension*:

$$\dim_{\text{H}} E = \sup\{s : \mathcal{H}_{\infty}^s(E) > 0\} = \inf\{s : \mathcal{H}^s(E) < \infty\}$$

Hausdorff dimension is quite measure-theoretically robust. However, there is also a much conceptually simpler notion of dimension called the *box* or *Minkowski dimension*. For a bounded subset  $E \subset \mathbb{R}^d$  and  $r > 0$ , we let  $N_r(E)$  denote the least number of balls of radius  $r$  required to cover  $E$ . Then the *lower* and *upper box dimensions* are given respectively by

$$\underline{\dim}_{\text{B}} E = \liminf_{r \rightarrow 0} \frac{\log N_r(E)}{\log(1/r)} \quad \text{and} \quad \overline{\dim}_{\text{B}} E = \limsup_{r \rightarrow 0} \frac{\log N_r(E)}{\log(1/r)}.$$

When the limits coincide, we call the shared value the *box dimension* and denote it by  $\dim_{\text{B}} E$ .

**Remark 1.4.** The terminology *Minkowski dimension* originates in an alternative formulation of the box dimension due to Minkowski. Given a set  $E$ , let  $E^{(r)}$  denote

the open  $r$ -neighbourhood of  $E$  and let  $m$  denote Lebesgue measure in  $\mathbb{R}^d$ . Then

$$\text{vol}(E^{(r)}) \approx r^d N_r(E)$$

so

$$\overline{\dim}_B E = \limsup_{r \rightarrow 0} \frac{\log(r^{-d} \text{vol}(E^{(r)}))}{\log(1/r)} = d - \liminf_{r \rightarrow 0} \frac{\log \text{vol}(E^{(r)})}{\log r}.$$

Of course, an analogous formula holds for  $\underline{\dim}_B E$ . The definition in terms of covers is typically more convenient since it is independent on the embedding in ambient space.

It is easy to check that

$$\dim_H E \leq \underline{\dim}_B E \leq \overline{\dim}_B E$$

and, for general subsets of  $\mathbb{R}^d$ , the dimensions can be distinct. Essentially, the main difference between the Hausdorff and lower box dimensions is that the Hausdorff dimension permits covers using balls of any radius, whereas the lower box dimension only permits covers of a fixed radius  $r$ .

**1.4. Overview.** We will study the Hausdorff dimension and the lower and upper box dimensions of sets of continued fractions with restricted digits. More generally, we will work with a general IFS  $\{f_i\}_{i \in \mathcal{I}}$  with attractor  $\Lambda$ , for a finite or countably infinite index set  $\mathcal{I}$ , acting on a compact subset  $X \subset \mathbb{R}^d$ .

Of particular interest is the family of maps  $\{\phi_b\}_{b \in \mathcal{D}}$  of inverse branches of the Gauss map, and the corresponding restricted digit sets  $\Lambda_{\mathcal{D}}$ . The non-linearity of the maps  $\phi_b$  results in a decent amount of technical difficulty, which obscures the overall picture of the proof. Therefore, we will first handle the special case that the maps  $f_n$  are in *similarities*, and only later on §4.1 explain what needs to be done to handle general conformal maps.

- In §2 we provide an introduction to infinitely-generated self-similar sets, and prove results concerning the Hausdorff and upper box dimensions. These results are originally due to Mauldin–Williams [MW86] and Mauldin–Urbański [MU96; MU99].
- In §3 we turn our attention to the lower box dimension, still in the self-similar case. These results are originally from [BR24+]. It turns out that there is no elegant formula for the lower box dimension, so we will study the more general question of asymptotics of the covering numbers  $r \mapsto N_r(\Lambda)$ . Then, we will use this to say something about the lower box dimension. For example, we can obtain a simple classification of when the box dimension of  $\Lambda$  exists.
- In §4 we tie up some loose ends, albeit with less detail than the earlier results. Most notably, we discuss the sharpness of the bounds on the lower box dimension, and also provide a summary of the key differences in the self-conformal case and the main ideas of how to handle them. We also spend some time discussing other notions of dimension.

**1.5. A note on implicit constants.** Given functions  $f, g: A \rightarrow [0, \infty)$  defined on some domain  $A$ , we write  $f \lesssim g$  to indicate that there is an implicit constant  $C > 0$  such that  $f(a) \leq Cg(a)$  for all  $a \in A$ . If  $f \lesssim g$  and  $g \lesssim f$ , we also write  $f \approx g$ .

We will fix an IFS  $\{f_i\}_{i \in \mathcal{I}}$  defined on a compact subset  $X \subset \mathbb{R}^d$ . The implicit constants are permitted to depend on the ambient dimension  $d$ , the uniform upper bound  $\xi$  on the contraction ratio, and  $\text{diam } X$ . Any other dependence will be indicated explicitly with a subscript, like  $\lesssim_\varepsilon$ .

## 2. INFINITELY-GENERATED SELF-SIMILAR SETS

**2.1. Self-similar iterated function systems.** Let us begin by introducing the key concepts concerning self-similar iterated function systems.

**Definition 2.1.** Let  $(X, d)$  be a metric space. A map  $f: X \rightarrow X$  is a *similarity* if there is a number  $\lambda > 0$  so that for all  $x, y \in X$ ,

$$d(f(x), f(y)) = \lambda d(x, y).$$

We call the number  $\lambda$  the *similarity ratio*.

In  $\mathbb{R}^d$ , similarity maps can always be written in the form  $f(x) = \lambda O x + t$  where  $O$  is an orthogonal matrix and  $t \in \mathbb{R}^d$  is a translation.

A *self-similar IFS* is an IFS  $\{f_i\}_{i \in \mathcal{I}}$  where the maps  $f_i$  are contracting similarities. Given  $\mathbf{i} \in \mathcal{I}^*$ , we let  $\rho(\mathbf{i}) \in (0, 1]$  denote the similarity ratio of  $f_{\mathbf{i}}$ . Observe that  $\rho$  is multiplicative:  $\rho(\mathbf{i}j) = \rho(\mathbf{i})\rho(j)$ . The quantity  $\rho$  is important since  $f_{\mathbf{i}}(X)$  is essentially a copy of  $X$  scaled down by a factor of  $\rho(\mathbf{i})$ .

Assuming that the maps  $f_i$  are similarity maps is very useful, but without more assumptions it is still rather difficult to understand the dimension theory of the attractor  $\Lambda$ . Essentially, the difficulty is that the sets  $f_{\mathbf{i}}(X)$  may overlap substantially, in which case the geometry of  $X$  depends not only on the similarity ratios of the maps  $f_{\mathbf{i}}$  but also on how the images  $f_{\mathbf{i}}(X)$  are arranged in space. Even in the case that  $\mathcal{I}$  is finite, this is a very difficult problem and dimension of the attractor is not known in general. Some of the most general results concerning the Hausdorff dimension in the overlapping case can be found in [Hoc14; KK25+; Rap22; RV24; Var19; Wu25+].

In the special case of continued fractions, however, the functions  $\phi_b$  were obtained as the inverse branches of the Gauss map, so the sets  $\phi_b(\mathbb{I})$  are disjoint for distinct values of  $n$ . The following assumption is a close sibling to such a disjointness assumption.

We say that a subset  $\mathcal{F} \subset \mathcal{I}^*$  is *mutually incomparable* if  $\mathbf{i}$  is not a prefix of  $\mathbf{j}$  for all  $\mathbf{i}, \mathbf{j} \in \mathcal{F}$  with  $\mathbf{i} \neq \mathbf{j}$ .

**Definition 2.2.** We say that the IFS  $\{f_i\}_{i \in \mathcal{I}}$  satisfies the *bounded neighbourhood condition* if there exists  $M \in \mathbb{N}$  so that for all mutually incomparable  $\mathcal{F} \subset \mathcal{I}^*$ , for all  $x \in X$ , and for all  $r \in (0, 1)$ ,

$$\#\{\mathbf{i} \in \mathcal{F} : \rho(\mathbf{i}) > r \text{ and } f_{\mathbf{i}}(X) \cap B(x, r) \neq \emptyset\} \leq M.$$

**Example 2.3.** A simple example of an IFS of similarities can be obtained from the Lüroth expansion on  $[0, 1]$ . For each  $n \in \mathbb{N}$ , let  $g_n: [0, 1] \rightarrow [0, 1]$  denote the unique orientation-preserving affine map which maps  $[0, 1]$  to the interval  $[1/(n+1), 1/n]$ . More precisely,

$$g_n(x) = \frac{x}{n(n+1)} + \frac{1}{n+1}.$$

The IFS  $\{g_n\}_{n=1}$  is the set of inverse branches of the transformation

$$T(x) = \left\lfloor \frac{1}{x} \right\rfloor \left( \left\lfloor \frac{1}{x} \right\rfloor + 1 \right) x - \left\lfloor \frac{1}{x} \right\rfloor$$

and produces expansions of the form

$$x = \frac{1}{a_1} + \frac{1}{a_1(a_1-1)a_2} + \cdots + \frac{1}{a_n \cdot \prod_{j=1}^{n-1} a_j(a_j-1)} + \cdots.$$

This is a linearized version of the Gauss map, with the branches flipped.

This IFS satisfies the bounded neighbourhood condition with constant 3 since, for a mutually incomparable family  $\mathcal{F}$ , the sets  $f_i((0, 1))$  for  $i \in \mathcal{F}$  are disjoint open intervals of width  $\rho(i)$ .

**2.2. Hausdorff dimension of the attractor.** When  $\mathcal{I}$  is a finite index set and  $\{f_i\}_{i \in \mathcal{I}}$  is an IFS of similarity maps satisfying the bounded neighbourhood condition, it is well-known that the attractor  $\Lambda$  satisfies

$$\dim_{\text{H}} \Lambda = s \quad \text{where} \quad \sum_{i \in \mathcal{I}} \rho(i)^s = 1.$$

The upper bound is immediate, since the family of sets  $\{f_i(X)\}_{i \in \mathcal{I}^n}$  is a cover for  $\Lambda$  with

$$\begin{aligned} \sum_{i \in \mathcal{I}^n} (\text{diam } f_i(X))^s &= (\text{diam } X)^s \sum_{i \in \mathcal{I}^n} \rho(i)^s \\ &= (\text{diam } X)^s \left( \sum_{i \in \mathcal{I}} \rho(i)^s \right)^n \\ &= (\text{diam } X)^s \\ &< \infty. \end{aligned}$$

The lower bound requires slightly more work, but argument is essentially as follows. Using the equation  $\sum_{i \in \mathcal{I}} \rho(i)^s = 1$ , define a product measure on the space  $\mathcal{p}^{\mathbb{N}}$  on  $\mathcal{I}^{\mathbb{N}}$  where  $\mathcal{p}$  is the measure on  $\mathcal{I}$  which assigns mass  $\mathcal{p}(i) = \rho(i)^s$ . Let  $\mu$  denote the pushforward of  $\mathcal{p}^{\mathbb{N}}$  under the coding map  $\pi$ . Then, using the bounded neighbourhood condition, check that  $\mu$  is well-distributed, in that  $\mu(B(x, r)) \leq Cr^s$  for all  $0 < r \leq \text{diam } X$  and  $x \in \Lambda$ . Finally, we can use  $\mu$  to obtain a lower bound on the  $s$ -cost of any cover of  $\Lambda$ : if  $\Lambda \subset \bigcup_{n=1}^{\infty} A_n$ , then

$$1 = \mu(\Lambda) \leq \sum_{n=1}^{\infty} \mu(B(x_n, \text{diam } A_n)) \leq C \sum_{n=1}^{\infty} (\text{diam } A_n)^s$$

where  $x_n \in A_n$ . Since  $A_n$  was an arbitrary cover, it follows that  $\mathcal{H}^s(\Lambda) > 0$  and therefore  $\dim_{\text{H}} \Lambda \geq s$ .

When  $\mathcal{I}$  is infinite, we immediately run into an issue: the equation  $\sum_{i \in \mathcal{I}} \rho(i)^s = 1$  may not have a solution! Regardless, we can still say something useful. Write

$$P(t) = \log \sum_{i \in \mathcal{I}} \rho(i)^t.$$

We call  $P$  the *pressure function* corresponding to the IFS. Then, we set

$$(2.1) \quad h := \inf\{t \geq 0 : P(t) < 0\}.$$

The exact same argument as in the finite case shows that  $\dim_{\text{H}} \Lambda \leq h$ . In fact, for all  $\varepsilon > 0$ ,

$$(2.2) \quad \sum_{\mathbf{i} \in \mathcal{I}^*} \rho(\mathbf{i})^{h+\varepsilon} = \sum_{n=0}^{\infty} \left( \sum_{i \in \mathcal{I}} \rho(i)^{h+\varepsilon} \right)^n \lesssim_{\varepsilon} 1.$$

On the other hand, it need not hold that  $P(h) = 0$ . It can happen that  $P(t) = +\infty$  for small values of  $t$ , and the moment  $P(t)$  is finite it already takes negative values bounded away from 0. To this extent, we write

$$\Omega = \{t \geq 0 : P(t) < \infty\} \quad \text{and} \quad \eta = \inf \Omega.$$

We call  $\eta$  the *finiteness parameter*. Clearly  $0 \leq \eta \leq h$ , and  $\eta = 0$  if and only if  $\mathcal{I}$  is finite. Since  $P$  is strictly decreasing, either  $\Omega = (\eta, \infty)$  or  $\Omega = [\eta, \infty)$ .

Moreover, one can check that  $P$  is a continuous and convex function on  $\Omega$ . In particular, if for instance  $\eta < h$ , then necessarily  $P(h) = 0$ .

If  $P(h) = 0$ , one can (with rather more work) develop a similar theory as in the finite case  $\mathcal{I}$  and define a well-distributed measure  $\mu$  on the attractor  $\Lambda$ , and with that obtain a lower bound for the Hausdorff dimension of  $\Lambda$ . This is done in detail in [MU96]. But if one is only interested in the Hausdorff dimension, it turns out that one only requires the results in the finite case to complete the proof.

Let  $\text{Fin}(\mathcal{I})$  denote the set of all finite subsets of  $\mathcal{I}$ . Moreover, for  $\emptyset \neq \mathcal{F} \subset \mathcal{I}$ , we let  $\Lambda_{\mathcal{F}}$  denote the attractor of the IFS  $\{f_i\}_{i \in \mathcal{F}}$ . The following result is [MU96, Theorem 3.15].

**Theorem 2.4.** *Let  $\{f_i\}_{i \in \mathcal{I}}$  be an IFS of similarities satisfying the bounded neighbourhood condition, with pressure function  $P(t)$  and attractor  $\Lambda$ . Then*

$$\dim_{\text{H}} \Lambda = \inf\{t \geq 0 : P(t) < 0\} = \sup\{\dim_{\text{H}} \Lambda_{\mathcal{F}} : \mathcal{F} \in \text{Fin}(\mathcal{I})\}.$$

*Proof.* We already saw that  $\dim_{\text{H}} \Lambda \leq h$ . Since  $\Lambda_{\mathcal{F}} \subset \Lambda$  for all  $\mathcal{F} \in \text{Fin}(\mathcal{I})$ , it remains to show that

$$h \leq \sup\{\dim_{\text{H}} \Lambda_{\mathcal{F}} : \mathcal{F} \in \text{Fin}(\mathcal{I})\}.$$

If  $h = 0$  there is nothing to prove; otherwise, let  $0 < s < h$  be arbitrary so that  $P(s) > 0$  (note, perhaps, that  $P(s) = +\infty$ ). Equivalently,  $\sum_{i \in \mathcal{I}} \rho(i)^s > 1$ . By the

usual properties of summation, there is a finite subset  $\mathcal{F} \subset \mathcal{I}$  such that

$$\sum_{i \in \mathcal{F}} \rho(i)^s > 1.$$

But recalling the formula for  $\dim_{\text{H}} \Lambda_{\mathcal{F}}$ , this means that  $\dim_{\text{H}} \Lambda_{\mathcal{F}} \geq s$ . Since  $s < h$  was arbitrary, the proof is complete.  $\square$

To summarize, the sets  $f_i(X)$  provide natural families of covers, and the covers obtained in this way are optimal from the perspective of Hausdorff dimension.

**2.3. Upper box dimension of the attractor.** Now that we have a good picture of the Hausdorff dimension, let us turn our attention to the upper box dimension. We continue to write  $h = \inf\{t \geq 0 : P(t) < 0\}$ .

Unlike the Hausdorff dimension, the upper box dimension can be strictly greater than 0 even for countable sets. For example, it is a good exercise to verify that

$$\dim_{\text{B}}\{n^{-1} : n \in \mathbb{N}\} = \frac{1}{2}.$$

The IFS of similarities  $\{f_i\}_{i \in \mathcal{I}}$  contains a countable subset which may be large from the perspective of the upper box dimension. For each map  $f_n$ , let  $x_n$  denote the unique fixed point, and let  $F = \{x_i : i \in \mathcal{I}\}$ . Certainly,  $F \subset \Lambda$ . Moreover, it could happen that  $\overline{\dim}_{\text{B}} F > 0$ , and this is an obstruction to the upper box dimension:

$$(2.3) \quad \overline{\dim}_{\text{B}} \Lambda \geq \max\{h, \overline{\dim}_{\text{B}} F\}.$$

This seems like a rather trivial obstruction, but let us note two facts. Firstly, one must also contend with images  $f_i(F)$  for  $i \in \mathcal{I}^*$ , and this family of sets is rather more substantial. Secondly, since  $\overline{\dim}_{\text{B}} \Lambda \cap V = \overline{\dim}_{\text{B}} \Lambda$  for any open set  $V$  intersecting  $\Lambda$ , by a Baire category argument, it in fact holds that  $\overline{\dim}_{\text{B}} \Lambda = \dim_{\text{P}} \Lambda$ , where  $\dim_{\text{P}} \Lambda$  is the packing dimension which (for our purposes) can be obtained as countably stabilized upper box dimension:

$$\dim_{\text{P}} E = \inf\{\max_i \overline{\dim}_{\text{B}} E_i : E \subset \bigcup_{i=1}^{\infty} E_i\}.$$

Since the packing dimension is countably stable,  $\dim_{\text{P}} F = 0$ , but regardless  $\dim_{\text{P}} \Lambda \geq \overline{\dim}_{\text{B}} F$ .

**Remark 2.5.** The precise choice of  $F$  as the set of fixed points of the maps  $f_i$  is convenient for the proofs, but it is not essential. For example, one could take any point  $x_i \in f_i(X)$  for  $i \in \mathcal{I}$  and let  $F = \{x_i : i \in \mathcal{I}\}$ . See §4.2 for more detail.

The following result was first proven in [MU99], with a somewhat different inductive proof. We present an alternative proof since (at least in my opinion) the proof is more transparent, and moreover the technique generalizes more easily when we handle the lower box dimension.

**Theorem 2.6.** *Let  $\{f_i\}_{i \in \mathcal{I}}$  be an IFS of similarities with attractor  $\Lambda$  and fixed point set  $F$ . Then*

$$\overline{\dim}_B \Lambda \leq \max\{h, \overline{\dim}_B F\}.$$

*In particular, if  $\dim_H \Lambda \geq h$  (for instance, if the IFS satisfies the bounded neighbourhood condition), then equality holds.*

*Proof.* Let  $\varepsilon > 0$  be arbitrary and let  $r > 0$  be small. We will provide an upper bound for  $N_r(\Lambda)$ .

First, iterate the self-similarity relation:

$$N_r(\Lambda) = N_r \left( \bigcup_{i \in \mathcal{I}} f_i(\Lambda) \right).$$

At this point, we cannot swap the covering number  $N_r$  and the union, since the union is infinite. However, we can split the indices in  $\mathcal{I}$  into two families. If  $\rho(i) < r$ , since  $f_i(\Lambda) \cap F \neq \emptyset$ ,  $f_i(\Lambda) \subset F^{(r)}$ . Moreover, since we are in  $\mathbb{R}^d$ , there is a constant  $C > 0$  so that  $N_\delta(E^{(\delta)}) \leq CN_\delta(E)$  for any set  $E \subset \mathbb{R}^d$  and  $\delta > 0$ . For the terms  $\rho(i) > r$ , at least at a heuristic level, swapping the union and covering number does not seem too lossy. Therefore, we may decompose

$$\begin{aligned} N_r(\Lambda) &\leq N_r \left( \bigcup_{\substack{i \in \mathcal{I} \\ \rho(i) < Nr}} f_i(\Lambda) \right) + \sum_{\substack{i \in \mathcal{I} \\ \rho(i) \geq r}} N_r(f_i(\Lambda)) \\ &\leq N_r(F^{(r \cdot \text{diam } X)}) + \sum_{\substack{i \in \mathcal{I} \\ \rho(i) \geq r}} N_r(f_i(\Lambda)) \\ &\leq CN_r(F) + \sum_{\substack{i \in \mathcal{I} \\ \rho(i) \geq r}} N_{r\rho(i)-1}(\Lambda). \end{aligned}$$

In the last line, we used self-similarity: the set  $f_i(\Lambda)$  is a self-similar copy of  $\Lambda$  so  $N_r(f_i(\Lambda)) = N_{r\rho(i)-1}(\Lambda)$ .

Now, we can apply the exact same argument to each set  $N_{r\rho(i)-1}(\Lambda)$ :

$$\begin{aligned} N_r(\Lambda) &\leq CN_r(F) + \sum_{\substack{i \in \mathcal{I} \\ \rho(i) \geq r}} \left( CN_{r\rho(i)-1} + \sum_{\substack{j \in \mathcal{I} \\ \rho(j) \geq r\rho(i)^{-1}}} N_{r\rho(ij)-1} \right) \\ &= CN_r(F) + \sum_{\substack{i \in \mathcal{I} \\ \rho(i) \geq r}} CN_{r\rho(i)-1}(F) + \sum_{\substack{\mathbf{i} \in \mathcal{I}^2 \\ \rho(\mathbf{i}) \geq r}} N_{r\rho(\mathbf{i})-1}(\Lambda). \end{aligned}$$

Now, recall that the IFS is uniformly contracting, so  $\rho(\mathbf{i}) \leq \xi^k$  for  $\mathbf{i} \in \mathcal{I}^k$ . In other words, if we iterate the above procedure  $k$  times where  $k$  is sufficiently large so that  $\xi^k < r$ , the final sum becomes empty and we obtain the general upper bound

$$(2.4) \quad N_r(\Lambda) \leq C \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) \geq r}} N_{r\rho(\mathbf{i})-1}(F).$$

So, we have replaced the question of covering  $\Lambda$  with a problem of bounding the covering numbers of many copies of  $F$  at many different scales. In order to bound the above sum, we find a new representation of the sum in a more convenient form. Define

$$(2.5) \quad \theta_{\mathbf{i}}(r) = 1 - \frac{\log \rho(\mathbf{i})}{\log r}.$$

Equivalently,  $\theta_{\mathbf{i}}(r)$  is chosen so that  $r^{\theta_{\mathbf{i}}(r)} = r\rho(\mathbf{i})^{-1}$ . Note that  $\theta_{\mathbf{i}}(r) \in [0, 1]$ , and therefore

$$(2.6) \quad \begin{aligned} N_{r\rho(\mathbf{i})^{-1}}(F) &\lesssim_{\varepsilon} \left(\frac{\rho(\mathbf{i})}{r}\right)^{\overline{\dim}_{\mathbb{B}} F + \varepsilon} \\ &= \rho(\mathbf{i})^{h+\varepsilon} \left(\frac{1}{r}\right)^{\theta_{\mathbf{i}}(r)(\overline{\dim}_{\mathbb{B}} F + \varepsilon) + (1-\theta_{\mathbf{i}}(r))(h+\varepsilon)} \\ &\leq \rho(\mathbf{i})^{h+\varepsilon} \left(\frac{1}{r}\right)^{\max\{h, \overline{\dim}_{\mathbb{B}} F\} + \varepsilon}. \end{aligned}$$

Substituting this back into (2.4) and then using (2.2),

$$\begin{aligned} N_r(\Lambda) &\lesssim_{\varepsilon} \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) \geq r}} \rho(\mathbf{i})^{h+\varepsilon} \left(\frac{1}{r}\right)^{\max\{h, \overline{\dim}_{\mathbb{B}} F\} + \varepsilon} \\ &\lesssim_{\varepsilon} \left(\frac{1}{r}\right)^{\max\{h, \overline{\dim}_{\mathbb{B}} F\} + \varepsilon} \sum_{\mathbf{i} \in \mathcal{I}^*} \rho(\mathbf{i})^{h+\varepsilon} \\ &\lesssim_{\varepsilon} \left(\frac{1}{r}\right)^{\max\{h, \overline{\dim}_{\mathbb{B}} F\} + \varepsilon}. \end{aligned}$$

Therefore  $\overline{\dim}_{\mathbb{B}} \Lambda \leq \max\{h, \overline{\dim}_{\mathbb{B}} F\} + \varepsilon$ .  $\square$

### 3. LOWER BOX DIMENSION AND ASYMPTOTICS FOR COVERING NUMBERS

We now turn our attention to the lower box dimension. Unlike the upper box dimension, it turns out that the case of the lower box dimension is quite a bit more subtle.

Let us begin with a simple bound, which only uses our results on the Hausdorff and upper box dimensions:

$$(3.1) \quad \max\{h, \underline{\dim}_{\mathbb{B}} F\} \leq \underline{\dim}_{\mathbb{B}} \Lambda \leq \overline{\dim}_{\mathbb{B}} \Lambda = \max\{h, \overline{\dim}_{\mathbb{B}} F\}.$$

In particular, this immediately provides two trivial mechanisms to guarantee that the box dimension exists:

- (i) We have  $h \geq \overline{\dim}_{\mathbb{B}} F$ , in which case  $\underline{\dim}_{\mathbb{B}} \Lambda = h$ . In other words, the Hausdorff dimension is sufficiently large that we can simply ignore the set of fixed points.

(ii) We have  $\underline{\dim}_B F = \overline{\dim}_B F$ , in which case  $\dim_B \Lambda = \max\{h, \dim_B F\}$ . This case occurs when the fixed point set itself is sufficiently regular.

Perhaps surprisingly, it turns out that these are the only two mechanisms under which the box dimension can exist. In fact, we will prove the following result:  $\underline{\dim}_B \Lambda = \overline{\dim}_B \Lambda$  if and only if  $\max\{h, \underline{\dim}_B F\} = \max\{h, \overline{\dim}_B F\}$ . This was first proven in [BR24+].

Unfortunately, there is a problem: it turns out that there is no simple formula for  $\underline{\dim}_B \Lambda$ . More precisely, there is no formula for  $\underline{\dim}_B \Lambda$  which depends only on  $h$ ,  $\underline{\dim}_B F$ , and  $\overline{\dim}_B F$ .

Let us briefly try to understand the difficulty by inspecting the proof of [Theorem 2.6](#), and trying to substitute the lower box dimension in place of the upper box dimension. Heuristically speaking, (2.4) is not so lossy: below, we will give a different shorter proof of a slight generalization of (2.4) below in [Lemma 3.5](#), which is clearly optimal. The main loss occurs in the first inequality in (2.6) if  $\underline{\dim}_B F < \overline{\dim}_B F$ . Moreover, the optimal exponent here depends on the value of  $\theta_1(r)$ , which could take (essentially) any value in the interval  $[0, 1]$ . So, to understand the covering number at scale  $r$ , we need to know something about the covering numbers of  $F$  at all scales between  $r$  and 1.

This motivates us to instead try to answer the following more general question: can we obtain an asymptotic formula for the covering numbers  $r \mapsto N_r(\Lambda)$  given the Hausdorff dimension  $h$  and the covering numbers  $r \mapsto N_r(F)$ ? We are asking for quite a bit more information about  $\Lambda$ , but in exchange we receive quite a bit more information about  $F$ . It will turn out that this question has a concrete answer, and we can use the concrete answer to this question to provide sharp bounds on the lower box dimension (in terms of  $h$ ,  $\underline{\dim}_B F$ , and  $\overline{\dim}_B F$ ) which in turn provides the classification of existence of the box dimension of  $\Lambda$ .

**3.1. Branching functions.** Let us first briefly introduce a formalism around the covering numbers  $r \mapsto N_r(F)$ . We begin with a normalization which is very convenient.

**Definition 3.1.** Given a set  $E \subset \mathbb{R}^d$ , the *branching function* of  $E$  is the function  $\beta_E: [0, \infty) \rightarrow [0, \infty)$  given by

$$\beta_E(u) = \log N_{2^{-u}}(E).$$

Here, the logarithm is in base 2. The reason for choosing this normalization of the covering number is that the function  $\beta_E$  is essentially an increasing  $d$ -Lipschitz function with  $\beta_E(0) = 0$ . The  $d$ -Lipschitz property occurs for the following reason. Given  $u, v \geq 0$ , we need to upper bound  $N_{2^{-(u+v)}}(E)$  in terms of  $N_{2^{-u}}(E)$ . First fix an optimal cover  $\{B(x_1, 2^{-u}), \dots, B(x_n, 2^{-u})\}$  for  $E$ , where  $n = N_{2^{-u}}(E)$ . Then, working in  $\mathbb{R}^d$ , we can cover each ball  $B(x_j, 2^{-u})$  with  $C_d 2^{dv}$  balls of radius  $2^{-(u+v)}$ , where  $C_d$  is some constant depending only on  $d$ . Therefore

$$(3.2) \quad \beta_E(u+v) = \log N_{2^{-(u+v)}}(E) \leq \log (C_d 2^{dv} N_{2^{-u}}(E)) \leq \beta_E(u) + dv + \log C_d.$$

In other words,  $\beta_E$  is  $d$ -Lipschitz up to a fixed additive error term.

Here is a formalization which makes it slightly easier to work with the error terms. Let  $\mathcal{B}(d)$  denote the increasing  $d$ -Lipschitz functions  $g: [0, \infty) \rightarrow [0, \infty)$  with  $g(0) = 0$ .

**Proposition 3.2.** *For all  $d \in \mathbb{N}$ , there is a constant  $M_d \in \mathbb{R}$  such that the following holds:*

*Let  $E \subset \mathbb{R}^d$ . Then there exists  $g \in \mathcal{B}(d)$  so that for all  $u \geq 0$ ,*

$$|\beta_E(u) - g(u)| \leq M_d + d \log(1 + \text{diam } E).$$

*Proof.* Let  $C_d > 0$  be chosen so that

$$(3.3) \quad N_{2^{-u}}(E) \leq C_d (\text{diam } E)^{d2^{ud}}$$

and (3.2) holds. Let  $M_d = \log C_d$ . Now, set

$$g = \sup\{h \in \mathcal{B}(d) : h \leq \beta_E\}.$$

The supremum is non-empty since  $\beta_E \geq 0$ . Moreover, it is easy to check that  $\mathcal{B}(d)$  is closed under suprema, so  $g \in \mathcal{B}(d)$ .

Now, let  $u \geq 0$  be fixed. If  $\beta_E(u) \leq M_d + d \log(1 + \text{diam } E)$ , there is nothing to prove. Otherwise, consider the piecewise linear function  $\psi_u$  such that:

- (a)  $\psi_u(u) = \beta_E(u) - M_d - d \log(1 + \text{diam } E) \geq 0$ ,
- (b)  $\psi_u$  has constant slope  $d$  on  $[0, u]$ , and
- (c)  $\psi_u$  is constant on  $[u, \infty)$ .

By (3.2) and since  $\beta_E$  is increasing, it follows that  $\psi_u \leq \beta_E$  on  $[0, \infty)$ . Moreover, recall from (3.3)

$$\beta_E(u) \leq M_d + d \log(\text{diam } E) + ud.$$

In particular,  $\psi_u(0) \leq 0$ . Thus  $\max\{0, \psi_u\} \in \mathcal{B}(d)$  so

$$g(u) \geq \psi_u(u) = \beta_E(u) - M_d - d \log(1 + \text{diam } E).$$

Since  $u \geq 0$  was arbitrary, the proof is complete.  $\square$

We can also choose branching functions which play nicely with the lower and upper box dimensions.

**Proposition 3.3.** *Let  $E \subset \mathbb{R}^d$  and let  $s = \underline{\dim}_B E$  and  $t = \overline{\dim}_B E$ . Then there exists  $g \in \mathcal{B}(d)$  such that  $g = \beta_E + o(u)$  and  $su \leq g(u) \leq tu$  for all  $u \geq 0$ .*

*Proof.* First apply Proposition 3.2 to get  $f \in \mathcal{B}(d)$  such that  $f = \beta_E + O(1)$ , and then set

$$g(u) = \min\{\max\{f(u), su\}, tu\}.$$

In other words,  $g(u)$  is obtained by clamping  $f(u)$  to the desired range  $[su, tu]$ . Clearly  $g \in \mathcal{B}(d)$ ; we must show that  $g = f + o(u)$ .

Indeed, for any  $\varepsilon > 0$ , for all  $u$  sufficiently large by definition of the lower and upper box dimension,  $(s - \varepsilon)u \leq f(u) \leq (t + \varepsilon)u$ . Therefore either  $f(u) \in [su, tu]$  in which case  $f = g$ , or  $|f(u) - g(u)| \leq \varepsilon u$ . Since this holds for all  $\varepsilon > 0$ , it follows that  $g = f + o(u)$ .  $\square$

**3.2. A candidate asymptotic formula.** Recall our goal in this section: given an IFS of similarities  $\{f_i\}_{i \in \mathcal{I}}$  satisfying the bounded neighbourhood condition, we wish to give essentially sharp bounds on  $\beta_\Lambda(u)$  in terms of  $h$  and  $\beta_F(u)$ . More precisely, we will provide a formula for the equivalence class  $\beta_\Lambda(u) + o(u)$  in terms of  $h$  and  $\beta_F(u) + o(u)$ . The  $o(u)$  error term is acceptable since, for example,

$$\underline{\dim}_B \Lambda = \liminf_{u \rightarrow \infty} \frac{\beta_\Lambda(u)}{u}.$$

**Remark 3.4.** It is an interesting question to try to determine sharper asymptotics, but as we will see this would likely require (at the very least) additional assumptions on the similarity ratios  $\rho(i)$ . For an explicit formulation, see [Question 4.16](#).

Our strategy is essentially as follows. For the upper bound, we will follow the proof of [Theorem 2.6](#) but whenever possible input information about  $\beta_F$ . Then, we prove a matching lower bound by considering finite subsystems.

In the proof of [Theorem 2.6](#), we first obtained the formula (2.4) using an inductive argument, and then used (2.4) along with some algebraic manipulation to obtain the upper bound on the upper box dimension. It is not too difficult to furnish a direct proof of (2.4) with less effort. Given  $r \in (0, 1)$ , define

$$F^*(r) = \bigcup_{\substack{i \in \mathcal{I}^* \\ \rho(i) \geq r}} f_i(F).$$

Of course,  $F^*(r) \subset \Lambda$  for all  $r \in (0, 1)$ , and  $F^*(r) = F$  for  $r$  sufficiently close to 1. Essentially as proven in (2.4), the sets  $F^*(r)$  provide good approximations to  $\Lambda$  at scale  $r$ . We recall that  $E^{(\delta)}$  denotes open  $\delta$ -neighbourhood of  $E$ .

**Lemma 3.5.** *Let  $\{f_i\}_{i \in \mathcal{I}}$  be an IFS of similarities with attractor  $\Lambda$  and fixed point set  $F$ . Then for all  $r \in (0, 1)$ ,*

$$\Lambda \subset (F^*(r))^{(r \cdot \text{diam } X)}.$$

*Proof.* Let  $x \in \Lambda$  be arbitrary. Iterating the self-similarity relation, there is an  $i \in \mathcal{I}^*$  with  $\rho(i) < r$ ,  $\rho(i^-) \geq r$ , and  $x \in f_i(\Lambda)$ . Write  $i = i^-j$  and let  $x_j \in F$  be the unique fixed point of  $f_j$ . Then

$$f_{i^-}(x_j) = f_{i^-}(f_j(x_j)) = f_i(x_j) \in f_i(\Lambda).$$

Since  $\rho(i^-) \geq r$ ,  $f_{i^-}(x_j) \in F^*(r)$  by definition, so  $f_i(\Lambda) \cap F^*(r) \neq \emptyset$ . Since  $\rho(i) < r$ ,  $\text{diam } f_i(\Lambda) < r \text{ diam } X$  so that

$$x \in f_i(\Lambda) \subset (F^*(r))^{(r \cdot \text{diam } X)}.$$

Since  $x \in \Lambda$  was arbitrary, the proof is complete. □

Now, let us try to guess a formula for  $\beta_\Lambda$ , given  $\beta_F$  and  $h$ . Using [Lemma 3.5](#) and

the fact that the  $f_i$  are similarities,

$$N_r(\Lambda) \lesssim \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) \geq r}} N_{r\rho(\mathbf{i})^{-1}}(F).$$

The bounded neighbourhood condition implies this inequality is quite close to being an equality. Now recall the definition of  $\theta_i = \theta_i(r)$  from (2.5). Write  $r = 2^{-u}$  for  $u \geq 0$ , so  $r\rho(\mathbf{i})^{-1} = 2^{-\theta_i u}$ . Taking logarithms and recalling the computation from the proof of Theorem 2.6,

$$\begin{aligned} \beta_\Lambda(u) &\leq \log \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) \geq r}} N_{r\rho(\mathbf{i})^{-1}}(F) = \log \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) \geq r}} \rho(\mathbf{i})^h 2^{(1-\theta_i) \cdot h \cdot u + \beta_F(\theta_i u)} \\ &\leq \sup_{0 \leq z \leq u} (\beta_F(z) + h(u-z)) + \log \sum_{\mathbf{i} \in \mathcal{I}^*} \rho(\mathbf{i})^h \end{aligned}$$

In the second line, we used the fact that  $\theta_i u$  takes values on  $[0, u]$ . Swapping the summation and the supremum is not so lossy since, by pigeonholing and losing a  $\log u$  error term, we may assume that  $r\rho(\mathbf{i})^{-1}$  is essentially constant, which does not cause issues since  $\beta_F$  is essentially continuous. This gives us the following candidate formula for  $\beta_\Lambda$ :

$$\beta_\Lambda(u) = \sup_{0 \leq z \leq u} (\beta_F(z) + h(u-z)) + o(u).$$

In order to make this sketch rigorous, we first prove a few essential technical facts.

**3.3. Technical results concerning level sets of cylinders.** The first technical concerns the summation defining the pressure function  $P$ .

**Lemma 3.6.** *For all  $\varepsilon > 0$ , there is a constant  $C = C(\varepsilon) \geq 1$  so that*

$$\left(\frac{1}{r}\right)^{-\varepsilon} \leq \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ Cr \geq \rho(\mathbf{i}) \geq r}} \rho(\mathbf{i})^h \leq \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) \geq r}} \rho(\mathbf{i})^h \lesssim_\varepsilon \left(\frac{1}{r}\right)^\varepsilon.$$

*Proof.* We begin with the upper bound. Recalling (2.2),

$$\sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) \geq r}} \rho(\mathbf{i})^h \leq \left(\frac{1}{r}\right)^\varepsilon \sum_{\mathbf{i} \in \mathcal{I}^*} \rho(\mathbf{i})^{h+\varepsilon} \lesssim_\varepsilon \left(\frac{1}{r}\right)^\varepsilon.$$

For the lower bound, we approximate the summation using finite subsets of  $\mathcal{I}$ . Since  $\sum_{i \in \mathcal{I}} \rho(i)^{h-\varepsilon} > 1$ , there exists  $\mathcal{F} \in \text{Fin}(\mathcal{I})$  and  $t > h - \varepsilon$  so that  $\sum_{i \in \mathcal{F}} \rho(i)^t = 1$ . Let  $C = \max\{\rho(i)^{-1} : i \in \mathcal{F}\}$ , so every infinite word  $x \in \mathcal{F}^\mathbb{N}$  contains at least one prefix with  $r \leq \rho(\mathbf{i}) \leq Cr$ , and therefore

$$r^{-\varepsilon} \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ Cr \geq \rho(\mathbf{i}) \geq r}} \rho(\mathbf{i})^h \geq \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ Cr \geq \rho(\mathbf{i}) \geq r}} \rho(\mathbf{i})^{h-\varepsilon} \geq \sum_{\substack{\mathbf{i} \in \mathcal{F}^* \\ Cr \geq \rho(\mathbf{i}) \geq r}} \rho(\mathbf{i})^t \geq 1.$$

The proof is complete after rearranging.  $\square$

**Lemma 3.6** shows at the cost of a  $(1/r)^\epsilon$  factor, we may restrict attention to maps with similarity ratio approximately  $r$ . Using the bounded neighbourhood condition, we can transfer the symbolic result concerning the pressure to a geometric fact about the distribution of the cylinders  $f_i(X)$ .

**Lemma 3.7.** *Let  $\{f_i\}_{i \in \mathcal{I}}$  be an IFS of similarities satisfying the bounded neighbourhood condition with constant  $M$ . Then for all  $C \geq 1$  and  $0 < r \leq R < 1$ ,*

$$\frac{1}{MC^{d+1}} \sum_{\substack{i \in \mathcal{I}^* \\ CR \geq \rho(i) \geq R}} N_{r/R}(F) \lesssim N_r \left( \bigcup_{\substack{i \in \mathcal{I}^* \\ CR \geq \rho(i) \geq R}} f_i(F) \right) \lesssim C^d \sum_{\substack{i \in \mathcal{I}^* \\ CR \geq \rho(i) \geq R}} N_{r/R}(F)$$

*Proof.* For the upper bound, note that  $r\rho(i)^{-1} \geq r/CR$ , so the statement holds with  $N_{r/CR}$  in place of  $N_{r/R}$ , but these covering numbers are the same up to a factor  $C^d$  times a constant in  $\mathbb{R}^d$ .

For the lower bound, we find a large subset of  $\mathcal{M} := \{i \in \mathcal{I}^* : R \leq \rho(i) \leq CR\}$  such that the sets  $f_i(F)$  are  $R$ -separated. First, since each infinite  $x \in \mathcal{I}^{\mathbb{N}}$  contains at most  $\lesssim \log(C) + 1 \lesssim C$  prefixes  $i$  in  $\mathcal{M}$ , we may find a mutually incomparable sub-family  $\mathcal{M}_1 \subset \mathcal{M}$  such that  $\#\mathcal{M}_1 \gtrsim \frac{1}{C}\#\mathcal{M}$ . Next, since  $\text{diam } f_i(F) \lesssim CR$ , we can cover  $f_i(F)$  with  $\lesssim C^d$  balls of radius  $R$ . For each such ball  $B(x, R)$ , by the bounded neighbourhood condition, there are at most  $M$  words  $i \in \mathcal{M}_1$  such that  $f_i(F) \cap B(x, R) \neq \emptyset$ . In particular, there are  $\lesssim MC^d$  words  $j \neq i$  such that  $d(f_i(F), f_j(F)) \leq R$ . Thus the greedy algorithm yields a subset  $\mathcal{M}_2 \subset \mathcal{M}_1$  with  $\#\mathcal{M}_2 \gtrsim M^{-1}C^{-d}\#\mathcal{M}_1$  such that  $d(f_i(F), f_j(F)) > R$  for all  $i \neq j$  in  $\mathcal{M}_2$ . Since  $r \leq R$ , the lower bound follows.  $\square$

**3.4. Proof of the asymptotic formula.** We now have our main result on the covering numbers  $r \mapsto N_r(\Lambda)$ . This result was first established in [BR24+], in the more general case that the maps are conformal.

**Theorem 3.8.** *Let  $\{f_i\}_{i \in \mathcal{I}}$  be an IFS of similarities satisfying the bounded neighbourhood condition with attractor  $\Lambda$  and fixed point set  $F$ . Then its branching function satisfies*

$$\beta_\Lambda(u) = \sup_{0 \leq z \leq u} (\beta_F(z) + h(u - z)) + o(u).$$

*Proof.* Let  $\varepsilon > 0$  be arbitrary and let  $C = C_\varepsilon \geq 1$  be the corresponding constant guaranteed by Lemma 3.6. Let  $u \geq 0$  and set  $r = 2^{-u}$  and let  $r \leq R \leq 1$  be arbitrary. Let  $\theta$  be chosen so that  $r/R = 2^{-\theta u}$  and write  $N_{r/R}(F) = R^h 2^{(1-\theta)hu + \beta_F(\theta u)}$ . Since  $R \leq \rho(i) \leq CR$ , by Lemma 3.7,

$$\begin{aligned} N_r \left( \bigcup_{\substack{i \in \mathcal{I}^* \\ CR \geq \rho(i) \geq R}} f_i(F) \right) &\approx_\varepsilon \sum_{\substack{i \in \mathcal{I}^* \\ CR \geq \rho(i) \geq R}} N_{r/R}(F) \\ &= 2^{(1-\theta)hu + \beta_F(\theta u)} \sum_{\substack{i \in \mathcal{I}^* \\ CR \geq \rho(i) \geq R}} R^h \end{aligned}$$

$$\approx_\varepsilon 2^{(1-\theta)hu + \beta_F(\theta u)} \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ CR \geq \rho(\mathbf{i}) \geq R}} \rho(\mathbf{i})^h.$$

By [Lemma 3.5](#) and the pigeonhole principle,

$$\frac{\log C_\varepsilon}{1 + \log(1/r)} \cdot N_r(\Lambda) \leq \sup_{r \leq R \leq 1} N_r \left( \bigcup_{\substack{\mathbf{i} \in \mathcal{I}^* \\ CR \geq \rho(\mathbf{i}) \geq R}} f_{\mathbf{i}}(F) \right) \leq N_r(\Lambda).$$

As  $R$  ranges from  $r$  to 1,  $\theta$  ranges from 1 to 0. Thus applying [Lemma 3.6](#), there is a constant  $M_\varepsilon \geq 0$  so that, taking logarithms,

$$\left| \beta_\Lambda(u) - \sup_{0 \leq z \leq u} (\beta_F(z) + h(u - z)) \right| \leq \varepsilon u + M_\varepsilon.$$

Since  $\varepsilon > 0$  was arbitrary, we are done.  $\square$

**3.5. The asymptotic formula as an operator.** Let  $\mathcal{A}$  denote the family of increasing functions from  $[0, \infty) \rightarrow [0, \infty)$ , and recall the space  $\mathcal{B}(\alpha)$  of functions which are  $\alpha$ -Lipschitz and satisfy  $f(0) = 0$ . Certainly  $\mathcal{B}(\alpha) \subset \mathcal{A}$ ; we recall from [§3.1](#) that  $\mathcal{B}(d)$  is essentially the space of branching functions of subsets of  $\mathbb{R}^d$ .

With [Theorem 3.8](#) in mind, for a number  $0 \leq h$  and  $f \in \mathcal{A}$ , let

$$\Psi_h(f)(u) = \sup_{0 \leq z \leq u} (f(z) + h(u - z)).$$

Since  $f$  is increasing,  $\Psi_0$  is the identity map. On the other hand,  $\Psi_h$  for  $h > 0$  is already a non-trivial operator; for instance  $\Psi_h(f)(u) \geq hu$  for any  $f \in \mathcal{A}$ .

We begin by establish some basic properties of  $\Psi_h$  as an operator.

**Proposition 3.9.** *Let  $h \geq 0$  and  $f, g \in \mathcal{A}$ . The following hold:*

- (i)  $\Psi_h f \in \mathcal{A}$ .
- (ii) If  $h \leq \alpha$  and  $f \in \mathcal{B}(\alpha)$ , then  $\Psi_h f \in \mathcal{B}(\alpha)$ .
- (iii) If  $\eta \in \mathcal{A}$  and  $f \leq g + \eta$ , then  $\Psi_h f \leq \Psi_h g + \eta$ .
- (iv) If  $f \leq g$ , then  $\Psi_h f \leq \Psi_h g$ .
- (v) If  $f = g + o(u)$ , then  $\Psi_h f = \Psi_h g + o(u)$ .
- (vi) If  $f = g + O(1)$ , then  $\Psi_h f = \Psi_h g + O(1)$ .

*Proof.* For  $f \in \mathcal{A}$  and  $z \geq 0$ , define

$$f_z(u) = \begin{cases} f(u) & : 0 \leq u \leq z \\ f(z) + h(u - z) & : z \leq u \end{cases}$$

Observe that  $\Psi_h f = \sup_{z \geq 0} f_z$ . If  $f \in \mathcal{A}$ , then  $f_z \in \mathcal{A}$ , and a supremum of increasing functions is increasing, so  $\Psi_h f \in \mathcal{A}$  giving (i). Similarly, if  $f \in \mathcal{B}(\alpha)$  where  $\alpha \geq h$ , then  $f_z \in \mathcal{B}(\alpha)$ ; since  $\mathcal{B}(\alpha)$  is closed under suprema,  $\Psi_h f \in \mathcal{B}(\alpha)$  giving (ii).

Now, for (iii), observe that

$$\Psi_h f(u) = \sup_{0 \leq z \leq u} (f(z) + h(u - z))$$

$$\begin{aligned} &\leq \sup_{0 \leq z \leq u} (g(z) + \eta(z) + h(u - z)) \\ &\leq \Psi_h g(u) + \eta(u) \end{aligned}$$

since  $\eta$  is increasing. Then (iv) follows by taking  $\eta(u) = 0$ , (v) follows by taking  $\eta$  with  $\lim_{u \rightarrow \infty} u^{-1}\eta(u) = 0$ , and (vi) follows by taking  $\eta(u) = C$ .  $\square$

Recall from [Proposition 3.2](#) that if  $E \subset \mathbb{R}^d$ , then there is an  $f \in \mathcal{B}(d)$  such that  $\beta_E = f + O(1)$ . By [Proposition 3.9](#), the  $O(1)$  error term is preserved by the operator  $\Psi_h$ , and  $\Psi_h$  descends to an operator on  $\mathcal{B}(d)$ . Therefore instead of thinking of  $\Psi_h$  as acting on the branching function  $\beta_F$ , we think of  $\Psi_h$  as acting on a representative  $f \in \mathcal{B}(d)$  for  $\beta_F$ .

A convenient consequence of working with a representative  $f \in \mathcal{B}(d)$  is that, by continuity of  $f$ , for each  $u \geq 0$  there is a maximal  $0 \leq u \leq \zeta_h(u)$  which realizes the supremum:

$$\Psi_h(f)(u) = f(\zeta_h(u)) + h(u - \zeta_h(u)).$$

Now we give a simple description of the operator  $\Psi_h$  restricted to  $\mathcal{B}(\alpha)$ . For  $0 \leq h \leq \alpha$ , set

$$\mathcal{B}_h(\alpha) = \{f \in \mathcal{B}(\alpha) : f(u+z) \geq f(u) + hz \text{ for all } 0 \leq z \leq u\}.$$

The space  $\mathcal{B}_h(\alpha)$  is the set of branching functions which are *lower  $h$ -Lipschitz*: if  $\partial f(u)$  denotes the set of derivatives of  $f$  at  $u$ , then  $\mathcal{B}_h(\alpha)$  is the set of functions  $f: [0, \infty) \rightarrow [0, \infty)$  with  $f(0) = 0$  and  $\partial f(u) \subset [h, \alpha]$  for all  $u \geq 0$ .

**Proposition 3.10.** *Let  $0 \leq h \leq \alpha$  and  $f \in \mathcal{B}(\alpha)$ . Then*

$$\Psi_h(f) = \inf\{g \in \mathcal{B}_h(\alpha) : g \geq f\}.$$

*In particular,  $\Psi_h: \mathcal{B}(\alpha) \rightarrow \mathcal{B}_h(\alpha)$  is surjective and idempotent.*

*Proof.* Let  $f \in \mathcal{B}(\alpha)$ . Clearly  $\Psi_h f \geq f$ . We next show that  $\Psi_h f \in \mathcal{B}_h(\alpha)$ . Recall that  $\Psi_h f \in \mathcal{B}(\alpha)$  from [Proposition 3.9 \(ii\)](#), so it suffices to check the  $h$ -Lipschitz property. Let  $0 \leq v \leq u$ : then

$$\begin{aligned} \Psi_h f(u) &\geq h(u - \zeta_h(v)) + f(\zeta_h(v)) \\ &= h(u - \zeta_h(v)) + \Psi_h f(v) - h(v - \zeta_h(v)) \\ &= \Psi_h f(v) + h(u - v) \end{aligned}$$

as required. To complete the proof, suppose  $g \in \mathcal{B}_h(\alpha)$  and  $g \geq f$ : we must show that  $g \geq \Psi_h f$ . Indeed, using the  $h$ -Lipschitz property of  $g$ ,

$$\Psi_h f(u) = f(\zeta_h(u)) + h(u - \zeta_h(u)) \leq g(\zeta_h(u)) + h(u - \zeta_h(u)) \leq g(u).$$

as required.  $\square$

The formula [Proposition 3.10](#) has a natural geometric interpretation. Momentarily forget about the formula for  $\beta_\Lambda$ , and assume we are given the values  $h$  and the function  $\beta_F$ . What can we say about  $\beta_\Lambda$ ?

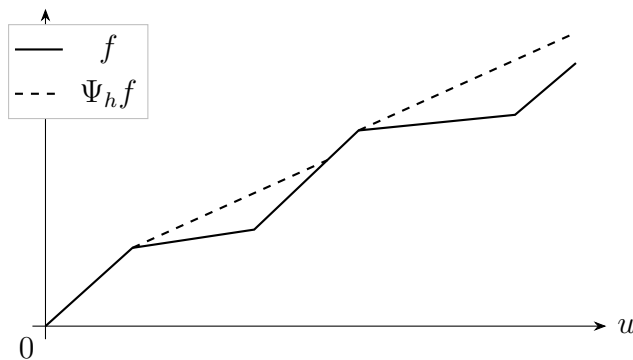


FIGURE 1. A branching function  $f$  and the corresponding function  $\Psi_h f$ .

1. Since  $F \subset \Lambda$ , we must have  $\beta_F \leq \beta_\Lambda$ .
2. Suppose we know the value of  $\beta_\Lambda(u)$ . What can we say about  $\beta_\Lambda(u+v)$ ? Let  $\{B(x_i, 2^{-u})\}$  be a maximal family of disjoint balls with  $x_i \in \Lambda$ . Since  $\Lambda$  is self-conformal, in a heuristic sense, we can find a copy of  $\Lambda$  scaled by a factor of  $2^{-u}$  inside each ball  $B(x_i, 2^{-u})$ . Since  $\Lambda$  is  $h$ -dimensional, this means that  $N_{2^{-(u+v)}}(B(x_i, 2^{-u}) \cap \Lambda) \geq 2^{h(u+v)}$ : in words,  $\Lambda$  is “at least  $h$ -dimensional in all balls”. But this property of being “at least  $h$ -dimensional between all pairs of scales”, at the level of branching functions, is the same as saying that  $\beta_\Lambda$  is lower  $h$ -Lipschitz.

**Proposition 3.10** says that these are the only two obstructions:  $\beta_\Lambda$  is *as small as possible*, given that it is bounded below by  $\beta_F$  and lower  $h$ -Lipschitz.

Note, in the infinite case, that property 2 is not literally true: the issue is that in a ball  $B(x, r)$  with  $x \in \Lambda$ , there need not exist a copy of  $\Lambda$  rescaled by a factor of exactly  $r$ : the rescaled copy could be much smaller. However, this is still true on average in a weak sub-exponential sense, and this is what is formalized in the proof of **Theorem 3.8**.

**3.6. Consequences for the lower box dimension.** Now, we use **Theorem 3.8** to obtain some applications concerning the lower box dimension. The first application is a non-trivial bound on  $\underline{\dim}_B \Lambda$  in terms of  $h$ ,  $\underline{\dim}_B F$ ,  $\overline{\dim}_B F$ , and the ambient dimension  $d$ . Given  $0 \leq s \leq t \leq \alpha$  and  $0 \leq h \leq \alpha$ , let

$$\mathcal{D}(h, s, t, \alpha) = \begin{cases} \{h\} & : t \leq h, \\ \left[ \max\{h, s\}, h + \frac{(t-h)(\alpha-h)s}{\alpha t - h s} \right] & : t > h. \end{cases}$$

Note that  $\mathcal{D}(h, s, t, \alpha)$  is not a singleton if and only if  $0 < h < t$  and  $0 < s < t$ .

**Corollary 3.11.** *Let  $\{f_i\}_{i \in \mathcal{I}}$  be an IFS of similarities satisfying the bounded neighbourhood condition with attractor  $\Lambda$  and fixed point set  $F$ . Then*

$$\underline{\dim}_B \Lambda \in \mathcal{D}(\dim_H \Lambda, \underline{\dim}_B F, \overline{\dim}_B F, d).$$

*Proof.* Write  $s = \underline{\dim}_B F$  and  $t = \overline{\dim}_B F$ . If  $t \leq h$  or  $s = t$ , we already saw that  $\underline{\dim}_B \Lambda = \max\{h, s\}$ , and in these cases  $\mathcal{D}(h, s, t, d) = \{\max\{h, s\}\}$ . If

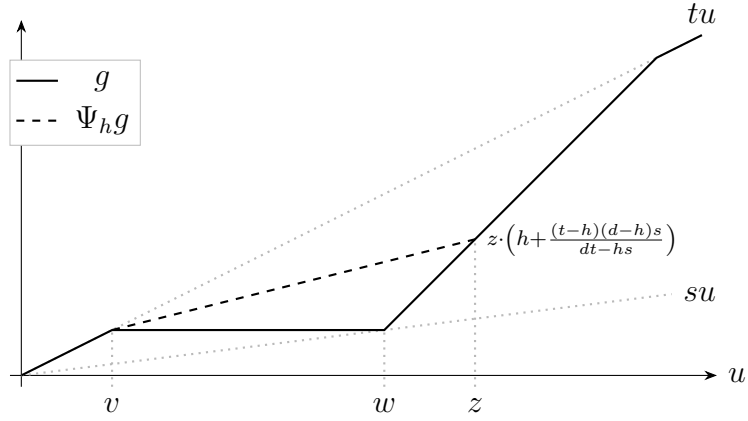


FIGURE 2. The upper bound  $g$  in the proof of [Corollary 3.11](#), in the special case that  $f(w) = sw$ . The dashed line has slope  $h$ . The function  $f$  (not depicted) is some branching function bounded above by  $g$  and below by  $su$ .

$\overline{\dim}_{\mathbb{B}} F \leq h$  or  $\underline{\dim}_{\mathbb{B}} F = \overline{\dim}_{\mathbb{B}} F$ , we already saw that  $\underline{\dim}_{\mathbb{B}} \Lambda = h$ . Therefore we may assume that  $s < t$  and  $h < t$ . By [Proposition 3.3](#), get  $f \in \mathcal{B}(d)$  such that  $f = \beta_F + o(u)$ ,  $f(u) \in [su, tu]$  for all  $u \geq 0$ . Then  $\beta_{\Lambda} = \Psi_h f + o(u)$  by [Theorem 3.8](#) and [Proposition 3.9 \(v\)](#).

Let  $s < s' < t$  be arbitrary. By definition of  $s = \underline{\dim}_{\mathbb{B}} F$ , there exists arbitrarily large  $w \geq 0$  such that  $f(w) \leq s'w$ . Fix such a value  $w$ , let  $v \leq w$  be such that  $tv = s'w$  and consider the function

$$g(u) = \begin{cases} tu & : u \leq v, \\ tv & : v \leq u \leq w, \\ tv + d(u - w) & : w \leq u. \end{cases}$$

Clearly  $g \in \mathcal{B}(d)$ . Moreover, since  $f(u) \leq tu$  and  $f(w) \leq s'w$ ,  $f \leq g$ . In particular, by [Proposition 3.9 \(iv\)](#),  $\Psi_h f \leq \Psi_h g$ .

Since  $h < t \leq d$ ,  $g$  is bounded above by the affine line  $\ell(u) = tv + h(u - v)$  on the interval  $[0, z]$  where  $z$  is the largest intersection point of  $g$  and  $\ell$  (and the only intersection point other than  $(v, \ell(v))$ ). For this value  $z$ ,

$$\Psi_h f(z) \leq \Psi_h g(z) = \ell(z).$$

Writing  $v$  and  $z$  in terms of  $w$  (and the ambient data  $s', t, h$ , and  $d$ ), a direct computation shows that

$$\frac{\ell(z)}{z} = h + \frac{(t-h)(d-h)s'}{dt-hs'}.$$

Since this holds for arbitrarily large  $z$ , and since  $s' > s$  was arbitrary, we conclude that  $\underline{\dim}_{\mathbb{B}} \Lambda \in \mathcal{D}(h, s, t, d)$  as claimed.  $\square$

**Remark 3.12.** In fact, the bounds in [Corollary 3.11](#) are sharp in the following sense: given any values for  $0 < h < d$ ,  $0 \leq s \leq t \leq d$ , and  $\beta \in \mathcal{D}(h, s, t, d)$  there exists an

IFS of similarities on  $\mathbb{R}^d$  satisfying the bounded neighbourhood condition with attractor  $\Lambda$  and fixed point set  $F$  such that  $\dim_{\text{H}} \Lambda = h$ ,  $\underline{\dim}_{\text{B}} F = s$ ,  $\overline{\dim}_{\text{B}} F = t$ , and  $\underline{\dim}_{\text{B}} \Lambda = \beta$ . See §4.3 for more details.

**Remark 3.13.** The sets  $\mathcal{D}(h, s, t, \alpha)$  are monotonically increasing in  $\alpha$ . Moreover,

$$\begin{aligned} \mathcal{D}(h, s, t, \infty) &:= \lim_{\alpha \rightarrow \infty} \mathcal{D}(h, s, t, \alpha) \\ &= \begin{cases} \{h\} & : t \leq h, \\ [\max\{h, s\}, s + (1 - \frac{s}{t})h] & : t > h. \end{cases} \end{aligned}$$

This gives non-trivial bounds which are valid in all dimensions simultaneously.

Using Corollary 3.11, we now obtain our characterization of the existence of the box dimension of  $\Lambda$ . This fact is intuitively clear from Figure 2: as long as  $s < t$  and  $h < t$ , the dashed line intersects the function  $g$  strictly below the line with slope  $t$ .

**Corollary 3.14.** *Let  $\{f_i\}_{i \in \mathcal{I}}$  be an IFS of similarities satisfying the bounded neighbourhood condition with attractor  $\Lambda$  and fixed point set  $F$ . Then  $\underline{\dim}_{\text{B}} \Lambda = \overline{\dim}_{\text{B}} \Lambda$  if and only if  $\overline{\dim}_{\text{B}} F \leq \max\{h, \underline{\dim}_{\text{B}} F\}$ .*

*Proof.* Recalling the earlier discussion in (3.1), if  $\overline{\dim}_{\text{B}} F \leq \max\{h, \underline{\dim}_{\text{B}} F\}$ , then it follows immediately that  $\underline{\dim}_{\text{B}} \Lambda = \overline{\dim}_{\text{B}} \Lambda$ .

Conversely, suppose that  $\overline{\dim}_{\text{B}} F > \max\{h, \underline{\dim}_{\text{B}} F\}$ . In particular,

$$\overline{\dim}_{\text{B}} \Lambda > h \quad \text{and} \quad 1 - \frac{\overline{\dim}_{\text{B}} F}{\underline{\dim}_{\text{B}} F} > 0.$$

Thus by Corollary 3.11 and Remark 3.13,

$$\begin{aligned} \underline{\dim}_{\text{B}} \Lambda &\leq \underline{\dim}_{\text{B}} F + \left(1 - \frac{\overline{\dim}_{\text{B}} F}{\underline{\dim}_{\text{B}} F}\right) h \\ &< \underline{\dim}_{\text{B}} F + \left(1 - \frac{\overline{\dim}_{\text{B}} F}{\underline{\dim}_{\text{B}} F}\right) \overline{\dim}_{\text{B}} F \\ &= \overline{\dim}_{\text{B}} F \\ &\leq \overline{\dim}_{\text{B}} \Lambda \end{aligned}$$

as claimed. □

## 4. VARIOUS LOOSE ENDS

We now tie up a variety of loose ends, sometimes with full proofs and sometimes with less detail. Most notably, we discuss the general self-conformal case, discuss a few other notions of dimension, and conclude with some open problems.

**4.1. Infinitely-generated self-conformal sets.** Let us first discuss what must be changed to handle more general self-conformal iterated function systems. For us, the following definition will be sufficient.

**Definition 4.1.** We say that the IFS  $\{f_i\}_{i \in \mathcal{I}}$  is *weakly conformal* if there are constants  $c > 0$ ,  $K \geq 1$ ,  $\xi \in (0, 1)$ , and a function  $\rho: \mathcal{I}^* \rightarrow (0, 1]$  such that:

- (i)  $\rho(\emptyset) = 1$ ,
- (ii)  $\rho(i) \leq \xi$  for all  $i \in \mathcal{I}$ ,
- (iii)  $K^{-1}\rho(i)\rho(j) \leq \rho(ij) \leq \rho(i)\rho(j)$ , and
- (iv) for all  $r > 0$  sufficiently small,  $x \in X$ , and  $r > 0$ ,

$$(4.1) \quad B(f_i(x), c\rho(i)r) \subseteq f_i(B(x, r)) \subseteq B(f_i(x), \rho(i)r).$$

**Remark 4.2.** Condition (ii) is just strict contractivity of the IFS.

Since the ball  $B(x, r)$  may not be contained in  $X$ , **Definition 4.1** implicitly assumes that there is an open neighbourhood  $V$  of  $X$  such that each  $f_i$  may be extended to map on  $V$ .

**Example 4.3.** Let us verify that these assumptions are satisfied for the inverse branches of the Gauss map. For each function  $\phi_b$ , recall that  $\phi_b(x) = 1/(b+x)$ . Let us prove, by induction on  $n$ , that for each  $b \in \mathbb{N}^n$ ,

$$(4.2) \quad \phi'_b(x) = \frac{(-1)^n}{(px+q)^2}$$

for integers  $1 \leq p \leq q$ . (In fact, the constants  $p$  and  $q$  are precisely the denominators of the convergents at steps  $n-1$  and  $n$  respectively.) Clearly this is true for  $n=1$ . Now let  $n \geq 1$  be arbitrary and suppose  $b \in \mathbb{N}^{n+1}$ . Write  $b = b^-j$  so

$$\phi'_b(x) = \frac{(-1)^n}{(p(j+x)^{-1}+q)^2} \cdot \frac{-1}{(j+x)^2} = \frac{(-1)^{n+1}}{(qx+(qj+p))^2}$$

which is of the claimed form.

Now, it is easy to check using (4.2) that  $|\phi'_b(x)| \leq 4|\phi'_b(y)|$  for all  $x, y \in [0, 1]$ ; and therefore in some small open neighbourhood  $[0, 1] \subset V$  we have that  $|\phi'_b(x)| \leq 5|\phi'_b(y)|$  for all  $x, y \in V$ . Thus we can define the function  $\rho(b) = \sup_{x \in V} |\phi'_b(x)|$  and the IFS is weakly conformal with constants  $c^{-1} = K = 5$ .

Similarly to the definition of the pressure in the self-similar case, the function  $\rho$  allows us to define the pressure in the self-conformal case:

$$P(t) = \lim_{m \rightarrow \infty} \frac{1}{m} \log S_m(t) \quad \text{where} \quad S_m(t) := \sum_{i \in \mathcal{I}^m} \rho(i)^t.$$

By **Definition 4.1 (iii)**, it always holds that  $K^{-t}S_k(t)S_m(t) \leq S_{k+m}(t) \leq S_k(t)S_m(t)$ . Therefore, the limit defining  $P(t)$  always exists and moreover that  $S_m(t) < \infty$  if and only if  $S_1(t) < \infty$ . Moreover, since the IFS is strictly contracting, we can write

$$\frac{1}{m} \log \sum_{i \in \mathcal{I}^m} \rho(i)^t = t \log \xi + \frac{1}{m} \log \sum_{i \in \mathcal{I}^m} (\xi^{-m} \rho(i))^t.$$

Since  $\xi^{-m}\rho(\mathbf{i}) \leq 1$ , it follows that  $P(t) - t \log \xi$  is decreasing, so in particular  $P(t)$  is strictly decreasing and  $\lim_{t \rightarrow \infty} P(t) = -\infty$  as long as  $P(t) < \infty$  for some  $t$ . With this in mind, if  $\eta = \inf\{t \geq 0 : S_1(t) < \infty\}$  then  $P(t) < \infty$  for all  $t > \eta$ . In particular,

$$(4.3) \quad \sum_{\mathbf{i} \in \mathcal{I}^m} \rho(\mathbf{i})^{\eta+\varepsilon} < \infty.$$

If  $\rho$  is multiplicative, then

$$\sum_{\mathbf{i} \in \mathcal{I}^m} \rho(\mathbf{i})^t = \left( \sum_{i \in \mathcal{I}} \rho(i)^t \right)^m$$

so the definition of the pressure coincides with our earlier definition in the special case that the  $f_i$  are similarities. In order to work with inequalities involving the pressure, we must therefore work with the higher iterates  $\{f_{\mathbf{i}}\}_{\mathbf{i} \in \mathcal{I}^n}$ .

For the upper bound, we no longer (in general) control sums of the form  $\sum_{\mathbf{i} \in \mathcal{I}^*} \rho(\mathbf{i})^h$ . Instead, we can obtain bounds of the form

$$\sum_{\substack{\mathbf{i} \in (\mathcal{I}^m)^* \\ \rho(\mathbf{i}) > r}} \rho(\mathbf{i})^h \lesssim_{\varepsilon} \left( \frac{1}{r} \right)^{\varepsilon}$$

where  $m$  must be chosen sufficiently large in terms of  $\varepsilon$ . Since the ‘‘step size’’ is now of order  $m$ , we instead obtain bounds in terms of the set  $F_m$  of fixed points of  $\{f_{\mathbf{i}} : \mathbf{i} \in \mathcal{I}^m\}$ . Note that  $\beta_{F_m}$  may be much larger than  $\beta_{F_1}$ : regardless, we still expect that  $\beta_{F_m}$  is bounded above by  $\Psi_h \beta_{F_1}$  (for instance, this must be true if the main result holds). Moreover, since  $\Psi_h$  is idempotent, this means that  $\Psi_h \beta_{F_1} = \Psi_h \beta_{F_m}$ . So that the argument is not circular, we must show up-front that  $\Psi_h \beta_{F_m} = \Psi_h \beta_{F_1}$ . This can be done directly with an induction argument.

It turns out that instead of working with sets of fixed points, it is more convenient to work with the level  $m$  orbit set

$$F_m := \bigcup_{\mathbf{i} \in \mathcal{I}^m} f_{\mathbf{i}}(x_0)$$

where  $x_0$  is chosen so that  $f_i(x_0) = x_0$  for some  $i \in \mathcal{I}$ . It is not too difficult to show that if  $F$  is the set of fixed points and  $F_1$  is the orbit set above, then  $\beta_F = \beta_{F_1} + O(1)$ ; see [Proposition 4.8](#). With this choice of  $x_0$ , observe that

$$(4.4) \quad F_{m+1} = \bigcup_{\mathbf{i} \in \mathcal{I}^m} f_{\mathbf{i}}(F_1) \subset \Lambda.$$

As before, for each  $\mathbf{i} \in \mathcal{I}^*$  with  $\rho(\mathbf{i}) > r = 2^{-u}$ , let  $\theta_{\mathbf{i}}(r) \in (0, 1]$  be chosen so that  $r\rho(\mathbf{i})^{-1} = r^{\theta_{\mathbf{i}}(r)}$  and therefore

$$(4.5) \quad N_{r, \rho(\mathbf{i})^{-1}}(E) = \rho(\mathbf{i})^h \cdot 2^{(1-\theta_{\mathbf{i}}(r))hu + \theta_{\mathbf{i}}(r)\beta_E(u)}$$

for a general set  $E$ . Also, recall that

$$(4.6) \quad N_r \left( \bigcup_{\substack{\mathbf{i} \in \mathcal{I}^m \\ \rho(\mathbf{i}) \leq r}} f_{\mathbf{i}}(\Lambda) \right) \lesssim N_r(F_m).$$

since each  $f_{\mathbf{i}}(\Lambda)$  where  $\rho(\mathbf{i}) \leq r$  is contained in the  $r \cdot (\text{diam } X)$ -neighbourhood of  $F_m$ . We now establish invariance under higher iterates. The idea here is similar to the idea in the proof of [MU99, Lemma 2.10].

**Lemma 4.4.** *Let  $\{f_i\}_{i \in \mathcal{I}}$  be a weakly conformal IFS. Then for each  $m \in \mathbb{N}$ ,*

$$\Psi_h \beta_{F_m}(u) = \Psi_h \beta_{F_1}(u) + o_m(u).$$

*Proof.* Since  $F_1 \subset F_2 \subset \dots$ , it follows that  $\Psi_1 \beta_{F_1} \leq \Psi_h \beta_{F_m}$  for all  $m \in \mathbb{N}$ . Thus it suffices to prove for all  $\varepsilon > 0$  and all  $m \in \mathbb{N}$  that

$$(4.7) \quad \limsup_{u \rightarrow \infty} (\Psi_h \beta_{F_m}(u) - \Psi_h \beta_{F_1}(u)) \leq 2\varepsilon u.$$

Fix  $\varepsilon > 0$ . We first prove by induction on  $m$  that there are constants  $A_m > 0$  so that for all  $r = 2^{-u} > 0$  sufficiently small,

$$(4.8) \quad N_r(F_m) \leq A_m 2^{\Psi_h \beta_{F_1}(u) + u\varepsilon}.$$

First, recall from (4.3) that

$$(4.9) \quad B_m := \sum_{\mathbf{i} \in \mathcal{I}^m} \rho(\mathbf{i})^{\eta + \varepsilon} < \infty.$$

Now we proceed with the induction. The case  $m = 1$  is trivial since  $\beta_{F_1} \leq \Psi_h \beta_{F_1}$ . Now suppose we have established (4.8) for some  $m \in \mathbb{N}$ . In the computation below the implicit constant depends only on the ambient dimension  $d$ . For each  $r = 2^{-u} > 0$  sufficiently small to apply (4.1), recalling (4.4) and (4.6),

$$\begin{aligned} N_r(F_{m+1}) &= N_r \left( \bigcup_{\mathbf{i} \in \mathcal{I}^m} f_{\mathbf{i}}(F_1) \right) \\ &\lesssim N_r(F_m) + \sum_{\substack{\mathbf{i} \in \mathcal{I}^m \\ \rho(\mathbf{i}) > r}} N_{r \cdot \rho(\mathbf{i})^{-1}}(F_1) \\ &= N_r(F_m) + \sum_{\substack{\mathbf{i} \in \mathcal{I}^m \\ \rho(\mathbf{i}) > r}} \rho(\mathbf{i})^h \cdot 2^{(1-\theta_{\mathbf{i}}(r))hu + \theta_{\mathbf{i}}(r)\beta_{F_1}(u)} \\ &\leq N_r(F_m) + 2^{\Psi_h \beta_{F_1}(u) + u\varepsilon} \sum_{\mathbf{i} \in \mathcal{I}^m} \rho(\mathbf{i})^{h+\varepsilon} \\ &\leq (A_m + B_m) \left( \frac{1}{r} \right)^{\Psi_h \beta_{F_1}(r) + \varepsilon}. \end{aligned}$$

In the last line, we have used the inductive hypothesis and (4.9). Thus (4.8) follows. We have shown that for all  $m \in \mathbb{N}$  that there is a constant  $a_m \in \mathbb{R}$  so that

$$\beta_{F_m}(u) \leq \Psi_h \beta_{F_1}(u) + 2\varepsilon u + a_m.$$

Since  $\varepsilon > 0$  was arbitrary, it follows that  $\beta_{F_m}(u) \leq \Psi_h \beta_{F_1}(u) + o_m(u)$ . Applying the operator  $\Psi_h$  on both sides of this inequality and using Proposition 3.9 (iii) and idempotence of  $\Psi_h$  from Proposition 3.10, the proof is complete.  $\square$

**Remark 4.5.** We actually proved something somewhat stronger: if  $\{f_i\}_{i \in \mathcal{I}}$  is an IFS for which there is a function  $\rho: \mathcal{I}^* \rightarrow (0, 1]$  such that  $\rho(\emptyset) = 1$ ,  $\rho(\mathbf{i}j) \leq \rho(\mathbf{i})\rho(j)$ , and

$$f_{\mathbf{i}}(B(x, r)) \subseteq B(f_{\mathbf{i}}(x), \rho(\mathbf{i})r),$$

then

$$\Psi_\eta \beta_{F_m}(u) = \Psi_\eta \beta_{F_1}(u) + o_m(u)$$

where  $\eta = \inf\{t \geq 0 : \sum_{i \in \mathcal{I}} \rho(i)^t < \infty\}$  is the usual finiteness parameter.

**Example 4.6.** Using Corollary 3.14, or more precisely, the version of Corollary 3.14 for self-conformal IFSs which we have outlined in this section, we can give an explicit example of a set  $I \subset \mathbb{N}$  for which the lower and upper box dimensions do not coincide.

The first thing to keep in mind is that we require  $h < 1/2$ : since  $\dim_{\mathbb{B}} F = \dim_{\mathbb{B}}\{n^{-1} : n \in \mathbb{N}\} = 1/2$ , if  $h \geq 1/2$ , then  $\dim_{\mathbb{H}} \Lambda_{\mathcal{D}} = \dim_{\mathbb{B}} \Lambda_{\mathcal{D}}$ . Since the finiteness parameter is  $1/2$ , in order to lower the Hausdorff dimension we must first lower the finiteness parameter.

So, we first consider the digit set  $\mathcal{D} := \{n^p : n \in \mathbb{N}, n \geq N\}$  for some  $p > 1$  and large  $N \in \mathbb{N}$ . This set has finiteness parameter  $1/2p$  and therefore, taking  $N$  to be large, the Hausdorff dimension of  $\Lambda_{\mathcal{D}}$  can be made arbitrarily close to  $1/2p$ .

On the other hand, one can compute that

$$\dim_{\mathbb{B}} F = \frac{p}{p+1} > \frac{1}{2p} \quad \text{where} \quad F := \{n^{-p} : n \in \mathbb{N}, n \geq N\},$$

independently of the choice of  $N$ . Moreover, the set of fixed points is essentially  $F$ .

So, to complete the construction, by Corollary 3.14, we just need to modify  $\mathcal{D}$  so that the lower box dimension is strictly less than  $p/(p+1)$ . This is quite easy; simply delete long stretches of consecutive entries but sufficiently sparsely so that the upper box dimension is unchanged.

**4.2. Alternatives to the fixed point set.** In this section, we discuss a minor technical fact concerning the choice of the set  $F$  of fixed points.

**Definition 4.7.** Fix an IFS  $\{f_i\}_{i \in \mathcal{I}}$  acting on a compact set  $X$ . We say that a set  $E \subset \bigcup_{i \in \mathcal{I}} f_i(X)$  is an *approximate orbit* if there is an integer  $\ell$  such that

$$1 \leq \#(E \cap f_i(X)) \leq \ell$$

for all  $i \in \mathcal{I}$ .

Assuming the bounded neighbourhood condition, it is easy to check that fixed point set  $F$  and the orbit sets  $\{f_i(x_0) : i \in \mathcal{I}\}$  (for any  $x_0 \in X$ ) are approximate orbits.

From the perspective of branching functions, all approximate orbits are equivalent (with implicit constants depending on the compact set  $X$ , the integer  $\ell$ , and the constant implicit in the definition of the bounded neighbourhood condition). This is essentially [MU99, Proposition 2.9].

**Proposition 4.8.** *Let  $A$  and  $B$  be approximate orbits of an IFS  $\{f_i\}_{i \in \mathcal{I}}$  satisfying the bounded neighbourhood condition. Then*

$$\beta_A = \beta_B + O(1).$$

*Proof.* Let  $r \in (0, 1)$  and let  $\mathcal{F} = \{i \in \mathcal{I} : \rho(i) > r\}$ . We partition  $A = A_{\leq r} \cup A_{> r}$  where  $A_{> r} = A \cap \bigcup_{i \in \mathcal{F}} f_i(X)$ , and similarly for  $B$ .

If  $i \notin \mathcal{F}$ , then  $\text{diam } f_i(X) \leq \rho(i) \text{diam } X \leq r \text{diam } X$ . Since  $B \cap f_i(X) \neq \emptyset$ , it follows that  $A_{\leq r}$  is contained in the  $r \cdot (\text{diam } X)$ -neighbourhood of  $B$ . Of course, the same also holds with  $A$  and  $B$  swapped, and therefore

$$N_r(A_{\leq r}) \lesssim N_r(B) \quad \text{and} \quad N_r(B_{\leq r}) \lesssim N_r(A).$$

To complete the proof, it suffices to show that

$$N_r(A_{> r}) \approx \#\mathcal{F} \approx N_r(B_{> r}).$$

Let  $M$  be the constant implicit in the bounded neighbourhood condition, so that

$$\#\{i \in \mathcal{F} : f_i(X) \cap B(x, r) \neq \emptyset\} \leq M \quad \text{for all } x \in X.$$

Fix a cover  $\{B(x_i, r)\}_{i=1}^n$  where  $n = N_r(A_{> r})$ . On one hand,  $B(x_i, r)$  intersects at most  $M$  images  $f_i(X)$ , so  $\#\mathcal{F} \leq nM$ . On the other hand, each  $f_i(X)$  contains at most  $\ell$  points in  $A$ , so  $n \leq \ell \#\mathcal{F}$ . Of course, the same holds with  $B$  in place of  $A$ , as required.  $\square$

**4.3. Sharpness of the bounds on the lower box dimensions.** We now explain why the bounds in Corollary 3.11 are sharp.

We focus our attention on the case  $d = 1$  for notational simplicity; see [BR24+] for the general case. We will prove the following result:

**Proposition 4.9.** *For any  $0 < h \leq 1$ ,  $0 \leq s \leq t \leq 1$ , and  $\beta \in \mathcal{D}(h, s, t, 1)$  there exists an IFS of similarities on  $[0, 1]$  satisfying the bounded neighbourhood condition with attractor  $\Lambda$  and fixed point set  $F$  such that  $\dim_{\text{H}} \Lambda = h$ ,  $\underline{\dim}_{\text{B}} F = s$ ,  $\overline{\dim}_{\text{B}} F = t$ , and  $\underline{\dim}_{\text{B}} \Lambda = \beta$ .*

We will prove something a bit more general, and obtain Proposition 4.9 by applying the result to a particularly chosen branching function  $f$ .

**Theorem 4.10.** *Let  $f \in \mathcal{B}(1)$  and  $h \in (0, 1]$ . Then there exists an IFS of similarities  $\{f_i\}_{i \in \mathcal{I}}$  with fixed point set  $F$  and attractor  $\Lambda$  such that  $\dim_{\mathbb{H}} \Lambda = h$  and  $f(u) = \beta_F(u) + o(u)$ . In particular,*

$$\beta_{\Lambda}(u) = \Psi_h f(u) + o(u).$$

We prove this theorem in two steps:

1. Given  $f \in \mathcal{B}(1)$ , we construct a countable set  $F \subset [0, 1]$  with at most one accumulation point and with branching function  $\beta_F(u) = f(u) + o(u)$ .
2. Given a countable set  $F \subset [0, 1]$  with at most 1 accumulation point, we construct an IFS  $\{f_i\}_{i \in \mathcal{I}}$  such that  $F$  is approximately the fixed point set of the IFS.

We begin by constructing a compact set corresponding to an arbitrary branching function.

**Lemma 4.11.** *For all  $f \in \mathcal{B}(1)$ , there exists a countable set  $F \subset [0, 1]$  with at most one accumulation point such that*

$$\beta_F(u) = f(u) + o(u).$$

*Proof.* We begin by constructing a compact set  $K_f$  which is quite uniform in space, and then construct  $F$  as a subset of  $K_f$ .

First, by taking a supremum over all integer-valued 1-Lipschitz functions  $h$  with  $h(0) = 0$  bounded above by  $f$ , there exists a function  $h: \mathbb{Z} \cap [0, \infty) \rightarrow \mathbb{Z} \cap [0, \infty)$  which is also an increasing and 1-Lipschitz function with  $h(0) = 0$ , such that

$$f(k) - 1 < h(k) \leq f(k) \quad \text{for all } k \geq 0.$$

Then, define a sequence  $(a_k)_{k=1}^{\infty} \in \{0, 1\}^{\mathbb{N}}$  by the rule  $a_k = h(k) - h(k-1)$ .

We inductively construct nested families of dyadic interval as follows. Let  $\mathcal{Q}_0 = \{[0, 1]\}$ . Now, suppose we have constructed  $\mathcal{Q}_k$  as a union of dyadic intervals at level  $k$ , for some  $k \geq 0$ . Then, replace each interval  $Q \in \mathcal{Q}_k$  with  $2^{a_{k+1}}$  sub-intervals at level  $k+1$ . Finally, set

$$K_f = \bigcap_{k=0}^{\infty} \bigcup_{Q \in \mathcal{Q}_k} Q.$$

Note that  $\mathcal{Q}_k$  is a union of exactly  $2^{h(k)}$  dyadic intervals, so  $\beta_{K_f}(u) = f(u) + O(1)$ .

We now construct a countable subset of  $F \subset K_f$ . For each  $k \in \mathbb{Z}$ ,  $k \geq 0$ , fix a dyadic interval  $Q_k \in \mathcal{Q}_k$  such that  $Q_0 \supset Q_1 \supset Q_2 \supset \dots$ . Let  $n_k$  be a rapidly increasing sequence of integers; for instance,  $n_k = 2^k$  suffices. For each interval  $Q_k$ , let  $F_k \subset K_f \subset Q_k$  be a subset constructed by taking one point from each dyadic interval in  $\mathcal{Q}_{n_k}$  contained in  $Q_k$ . Finally, let  $F = \bigcup_{k=0}^{\infty} F_k$ .

Note that the only accumulation point of  $F$  is  $\bigcap_{k=0}^{\infty} Q_k$  which is a singleton. Moreover, since  $F \subset K_f$  by construction,  $\beta_F \leq \beta_{K_f} = f + O(1)$ , so it suffices to prove the other inequality.

Let  $u \geq 0$  be arbitrary and let  $k$  be the unique integer such that  $2^{k-1} < u \leq 2^k$ . By definition of  $F_k \subset F$ ,

$$N_{2^{-u}}(F) \geq N_{2^{-u}}(F_k) \geq 2^{-k} N_{2^{-u}}(K_f).$$

But  $k = O(\log u)$  by choice of  $n_k$ , so taking logarithms, the proof is complete.  $\square$

Now given the set  $F$ , we construct the IFS of similarities. This is particularly easy in  $\mathbb{R}$ .

**Lemma 4.12.** *Let  $K \subset [0, 1]$  be any compact set of Lebesgue measure 0. Then there exists an IFS of similarities  $\{f_i\}_{i \in \mathcal{I}}$  satisfying the bounded neighbourhood condition with attractor  $\Lambda$  such that  $\dim_{\mathbb{H}} \Lambda = h$  and with an approximate orbit  $F$  such that  $\overline{F} \cup \{0, 1\} = K \cup \{0, 1\}$ .*

*Proof.* Enumerate the maximal disjoint open intervals of  $K^c$  as  $\{J_i\}_{i \in \mathcal{I}}$  for a countable index set  $\mathcal{I}$ . For each  $i \in \mathcal{I}$ , write  $J_i = (a_i, b_i)$ , and let  $f_i: (0, 1) \rightarrow J_i$  be defined by the rule

$$f_i(x) = (b_i - a_i)^{1/h} x + a_i.$$

Since the intervals  $J_i$  are pairwise disjoint, it is easy to see that the IFS satisfies the bounded neighbourhood condition. Moreover, by definition,  $K \cup \{0, 1\} = \overline{\{f_i(0) : i \in \mathcal{I}\}} \cup \{0, 1\}$ .

The IFS  $\{f_i\}_{i \in \mathcal{I}}$  has pressure function

$$P(t) = \sum_{i \in \mathcal{I}} (b_i - a_i)^{t/h}.$$

Since  $F$  has at most 1 accumulation point,  $\sum_{i=1}^{\infty} (b_i - a_i) = 1$ , so  $P(h) = 1$ , as required.  $\square$

*Proof (of Theorem 4.10).* Let  $f \in \mathcal{B}(1)$ , and by Lemma 4.11, get a countable set  $E \subset [0, 1]$  with at most 1 accumulation point such that  $\beta_E(u) = f(u) + o(u)$ . Then apply Lemma 4.12 with the set  $E$  and value  $h$  to get an IFS  $\{f_i\}_{i \in \mathcal{I}}$  satisfying the bounded neighbourhood condition with attractor  $\Lambda$  and fixed point set  $F$ , such that  $\beta_F(u) = \beta_E(u) + o(u) = f(u) + o(u)$  and  $\dim_{\mathbb{H}} \Lambda = h$ . The proof is therefore complete by Theorem 3.8.  $\square$

Now we can complete the proof of the sharpness result.

*Proof (of Proposition 4.9).* If  $\mathcal{D}(h, s, t, 1)$  is a singleton, let  $f \in \mathcal{B}(1)$  be such that  $\liminf_{u \rightarrow \infty} u^{-1} f(u) = s$  and  $\liminf_{u \rightarrow \infty} u^{-1} f(u) = t$ . Apply Theorem 4.10 to get an IFS with fixed point set  $F$  and attractor  $\Lambda$  such that  $\beta_F = f + o(u)$  and  $\dim_{\mathbb{H}} \Lambda = h$ . Then by Corollary 3.11, since  $\mathcal{D}(h, s, t, 1)$  is a singleton, it must be that  $\underline{\dim}_{\mathbb{B}} \Lambda$  is the expected value.

Otherwise,  $\mathcal{D}(h, s, t, 1)$  is not a singleton, so  $0 < h < t$  and  $0 < s < t$ . For simplicity, assume  $\max\{h, s\} < \beta$  and  $t < 1$ ; in the general case, apply the construction

inductively with sequences  $\beta \leq \beta_n \leq t_n \leq t$  such that  $\beta_n > \max\{h, s\}$  decreases to  $\beta$  and  $t_n < 1$  increases to 1.

The building block is a function  $\psi: [0, 1] \rightarrow [0, t]$  which is increasing, 1-Lipschitz, and satisfies  $\psi(u) = tu$  for all  $u$  in a neighbourhood of 0 and  $u = 1$ . The function  $\psi$  is defined in a similar way as depicted in [Figure 2](#): fix values  $0 < v < w < 1$ ; on  $[0, v]$ , let  $\psi$  have slope  $t$ , on  $[v, w]$ , let  $\psi$  have slope 0, and on  $[w, 1]$  let  $\psi$  have slope  $\alpha$ . Adjusting  $v, w$ , and  $\alpha$  and using the fact that  $\beta \in \mathcal{D}(h, s, t, 1)$ , it can be arranged so that  $\psi(z)/z = \beta$  where  $z$  is the unique value for which the segment joining  $(v, tv)$  and  $(z, \psi(z))$  has slope  $h$ .

Finally, to construct the function  $f$ , set  $f(u) = tu$  for  $u \leq 1$ , and for  $v^{n-1} \leq u \leq v^n$  define  $f(u) = v^n \psi(u/v^n)$ . Now  $f$  on  $[v^{n-1}, v^n]$  is precisely a self-similar copy of  $\psi$  on  $[v, 1]$ , and the choice of  $\psi$  ensures, by [Theorem 3.8](#), that the corresponding IFS provided by [Theorem 4.10](#) with branching function  $f$  and value  $h$  has the desired dimensions.  $\square$

**4.4. Intermediate dimensions.** The intermediate dimensions are a continuously parametrized family of dimensions which lie between the Hausdorff and box dimensions, first introduced in [\[FFK20\]](#). The motivation is the following: since the box dimensions are defined using covers consisting of a fixed radius, whereas the Hausdorff dimension is defined using arbitrary covers, one can obtain a notion of dimension by allowing some flexibility of the size of the covers. More precisely, the covers are permitted to lie in a block of scales of the form  $[r, r^\theta]$  where  $0 < \theta \leq 1$ .

**Definition 4.13.** Given  $0 < \theta \leq 1$  and  $r > 0$ , the  $(s, r, \theta)$ -covering cost is the quantity

$$\mathcal{C}_r^{s, \theta}(E) = \inf \left\{ \sum_i (\text{diam } A_i)^s : E \subset \bigcup_i A_i \text{ and } \text{diam } A_i \in [r, r^\theta] \text{ for all } i \right\}.$$

Now, the *upper intermediate dimensions* of  $E$  are given by

$$\overline{\dim}_\theta E = \inf \{ s \geq 0 : \mathcal{C}_r^{s, \theta}(E) \lesssim_s 1 \text{ for all } 0 < r < 1 \}.$$

For example,  $\mathcal{C}_r^{s, 1}(E) = r^s N_r(E)$ , so  $\overline{\dim}_1 E = \overline{\dim}_B E$ .

Analogously to the results for the upper box dimensions, in [\[BF23\]](#), Banaji & Fraser proved that for an infinitely generated self-conformal set satisfying the usual conditions, for all  $\theta \in (0, 1]$ ,

$$\overline{\dim}_\theta \Lambda = \max\{h, \overline{\dim}_\theta F\}.$$

Similarly to the proof due to Mauldin & Urbański in the case of the upper box dimension [\[MU99\]](#), they use an induction argument.

We give a simplified proof of this result in the self-similar case using the approach in [Theorem 2.6](#). In the proof of [Theorem 2.6](#), we needed upper bounds for the covering numbers  $N_r(f_i(F))$  where  $\rho(i) > r$ . By self-similarity, this number is precisely  $N_{r^{\theta_i(r)}}(F)$ , where we recall that  $r^{\theta_i(r)} = r\rho(i)^{-1}$ .

For the upper intermediate dimensions, the bound is slightly different:

$$\mathcal{C}_r^{s, \theta}(f_i(F))$$

$$\begin{aligned}
&= \inf \left\{ \sum_i (\text{diam } A_i)^s : f_i(F) \subset \bigcup_i A_i \text{ and } \text{diam } A_i \in [r, r^\theta] \right\} \\
&= \inf \left\{ \sum_i (\rho(\mathbf{i}) \cdot \text{diam } B_i)^s : F \subset \bigcup_i B_i \text{ and } \text{diam } B_i \in [\rho(\mathbf{i})^{-1}r, \rho(\mathbf{i})^{-1}r^\theta] \right\}.
\end{aligned}$$

In particular, we now have two cases to consider: either  $r \leq \rho(\mathbf{i}) \leq r^\theta$ , or  $r^\theta < \rho(\mathbf{i}) < 1$ . Equivalently, these are the cases  $0 \leq \theta_i(r) \leq 1 - \theta$  and  $1 - \theta < \theta_i(r) < 1$ . The former case is easy, since  $\rho(\mathbf{i})$  is a permitted diameter so we can cover each  $f_i(F)$  with a bounded number of balls. In the latter case, we can use the upper intermediate dimensions with a smaller value of  $\theta$ : set

$$\kappa_i(r, \theta) = \frac{\theta_i(r) + \theta - 1}{\theta_i(r)} \in (0, \theta]$$

so that

$$(4.10) \quad \mathcal{C}_r^{s, \theta}(f_i(F)) = \rho(\mathbf{i})^s \cdot \mathcal{C}_{r, \theta_i(r)}^{s, \kappa_i(r, \theta)}(F) \leq \rho(\mathbf{i})^s \cdot \mathcal{C}_{r, \theta_i(r)}^{s, \theta}(F).$$

Here is the formal version of the above argument.

**Theorem 4.14.** *Let  $\{f_i\}_{i \in \mathcal{I}}$  be an IFS of similarities with attractor  $\Lambda$  and fixed point set  $F$ . Then*

$$\overline{\dim}_\theta \Lambda \leq \max\{h, \overline{\dim}_\theta F\}.$$

*In particular, if  $\dim_{\text{H}} \Lambda \geq h$  (for instance, if the IFS satisfies the bounded neighbourhood condition), then equality holds.*

*Proof.* Let  $r \in (0, 1)$  and  $s > \max\{h, \overline{\dim}_\theta F\}$ . By [Lemma 3.5](#),

$$\Lambda \subset (F^*(r))^{(r \cdot \text{diam } X)}$$

so that

$$\mathcal{C}_r^{s, \theta}(\Lambda) \lesssim \mathcal{C}_r^{s, \theta} \left( \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) > r}} f_i(F) \right) \leq \mathcal{C}_r^{s, \theta} \left( \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ r^\theta \geq \rho(\mathbf{i}) > r}} f_i(F) \right) + \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) > r^\theta}} \mathcal{C}_r^{s, \theta}(f_i(F)).$$

To bound the first term, since  $\text{diam } f_i(F) \lesssim r^\theta$  and  $s > h$ ,

$$\mathcal{C}_r^{s, \theta} \left( \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ r^\theta \geq \rho(\mathbf{i}) > r}} f_i(F) \right) \lesssim \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ r^\theta \geq \rho(\mathbf{i}) > r}} \rho(\mathbf{i})^s \leq \sum_{\mathbf{i} \in \mathcal{I}^*} \rho(\mathbf{i})^s \lesssim_s 1.$$

To bound the second term, by [\(4.10\)](#) and since in addition  $s > \overline{\dim}_\theta F$ ,

$$\sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) > r^\theta}} \mathcal{C}_r^{s, \theta}(f_i(F)) \leq \sum_{\substack{\mathbf{i} \in \mathcal{I}^* \\ \rho(\mathbf{i}) > r^\theta}} \rho(\mathbf{i})^s \mathcal{C}_{r, \theta_i(r)}^{s, \theta}(F) \lesssim_s \sum_{\mathbf{i} \in \mathcal{I}^*} \rho(\mathbf{i})^s \lesssim_s 1.$$

Therefore  $C_r^{s,\theta}(\Lambda) \lesssim_s 1$ , and since  $s > \max\{h, \overline{\dim}_\theta F\}$  was arbitrary this completes the proof.  $\square$

**Remark 4.15.** Banaji & Fraser also obtained some partial results for the Assouad dimension (and other related notions of dimension) in [BF24]. See that paper for definitions and more discussion of the results.

**4.5. Some open problems.** In [Theorem 3.8](#) we proved an asymptotic formula for the covering numbers  $N_r(\Lambda)$  in terms of the Hausdorff dimension  $h$  and  $N_r(F)$ . However, the asymptotic formula was only up to error terms of the form  $C_\varepsilon(1/r)^\varepsilon$ , for arbitrarily small  $\varepsilon$ . There are multiple places where this error term enters the proof, but one notable location where it seems rather unavoidable is when bounding sums of the form  $\sum_{i \in \mathcal{I}^*} \rho(i)^h$ . If such sums do not introduce error terms, is it possible to get better asymptotics? In the best case scenario, it might be possible to get an error term which is an absolute constant independent of  $r$ .

**Question 4.16.** Let  $\{f_i\}_{i \in \mathcal{I}}$  be an IFS of similarities satisfying the bounded neighbourhood condition with attractor  $\Lambda$  and fixed point set  $F$ . Suppose  $\{f_i\}_{i \in \mathcal{I}}$  is regular, in that the value  $h$  satisfies  $\sum_{i \in \mathcal{I}} \rho(i)^h = 1$ . Does its branching function satisfy

$$\beta_\Lambda(u) = \sup_{0 \leq z \leq u} (\beta_F(z) + h(u - z)) + O(1)?$$

What about the more general self-conformal case?

See [\[MU96\]](#) for more discussion concerning the assumption that the IFS is regular.

In [Theorem 4.14](#) we saw that the results concerning the upper box dimension extended quite easily to the case of the upper intermediate dimensions. One can also define the *lower intermediate dimensions*:

$$\underline{\dim}_\theta E = \inf \{s \geq 0 : C_r^{s,\theta}(E) \lesssim_s 1 \text{ for arbitrarily small } r > 0\}.$$

As before,  $\underline{\dim}_1 E = \underline{\dim}_B E$ .

In order to understand the lower box dimension, we required information about the covering numbers  $N_r(F)$  at all scales  $r$ . We can also consider such information for the intermediate dimensions. For values  $0 \leq v \leq u$ , define

$$\mathcal{W}_E(s, u, v) = \inf \left\{ \sum_i (\text{diam } A_i)^s : E \subset \bigcup_i A_i \text{ and } \text{diam } A_i \in [2^{-u}, 2^{-(u-v)}] \right\}.$$

This is essentially [Definition 4.13](#) but with a somewhat different parametrization. One can show that  $s \mapsto \log \mathcal{W}_E(s, u, v)$  is strictly decreasing and continuous, and approximately satisfies  $\log \mathcal{W}_E(0, u, v) \geq 0$  and  $\log \mathcal{W}_E(d, u, v) \leq 0$  in  $\mathbb{R}^d$ . Therefore, one may define (somewhat informally) a function  $\omega_E(u, v)$  so that

$$\log \mathcal{W}_E((u - v)^{-1} \omega_E(u, v), u, v) = 0.$$

For example,

$$0 = \log \mathcal{W}_E(u^{-1} \omega_E(u, 0), u, 0) = \beta_E(u) - \omega_E(u, 0)$$

so  $\omega_E(u, 0) = \beta_E(u)$  recovers the usual branching function. Moreover, one can check for  $0 < \theta \leq 1$  that  $\omega_E$  encodes the intermediate dimensions:

$$\overline{\dim}_\theta E = \limsup_{u \rightarrow \infty} \frac{\omega_E(u, (1 - \theta)u)}{\theta u}$$

$$\underline{\dim}_\theta E = \liminf_{u \rightarrow \infty} \frac{\omega_E(u, (1 - \theta)u)}{\theta u}.$$

Here is an explicit question.

**Question 4.17.** *Let  $\{f_i\}_{i \in \mathcal{I}}$  be an IFS of similarities satisfying the bounded neighbourhood condition with attractor  $\Lambda$  and fixed point set  $F$ . What can be said about  $\underline{\dim}_\theta \Lambda$ ? Can one obtain a similar classification as proven in [Corollary 3.14](#)? More generally, can one provide a formula for  $\omega_\Lambda$  in terms of  $h$  and  $\omega_F$ ?*

## ACKNOWLEDGEMENTS

These notes were prepared for a minicourse taught during the semester program “Continued Fractions, Fractals, Ergodic theory and Dynamics” during the Simons semester at IMPAN in Warsaw. My stay was supported by a Simons Foundation grant (award no. SFI-MPS-T-Institutes-00010825) and from Polish State Treasury funds (agreement no. MNiSW/2025/DAP/491).

## REFERENCES

- [BF23] A. Banaji and J. M. Fraser. *Intermediate dimensions of infinitely generated attractors*. *Trans. Am. Math. Soc.* **376** (2023), 2449–2479. [zbl:1547.28014](#).
- [BF24] A. Banaji and J. M. Fraser. *Assouad type dimensions of infinitely generated self-conformal sets*. *Nonlinearity* **37** (2024). Article No. 045004, 32pp. [zbl:1558.28002](#).
- [BR24+] A. Banaji and A. Rutar. *Lower box dimension of infinitely generated self-conformal sets*. Preprint. [arxiv:2406.12821](#).
- [FFK20] K. J. Falconer, J. M. Fraser, and T. Kempton. *Intermediate dimensions*. *Math. Z.* **296** (2020), 813–830. [zbl:1448.28009](#).
- [Hoc14] M. Hochman. *On self-similar sets with overlaps and inverse theorems for entropy*. *Ann. Math.* **180** (2014), 773–822. [zbl:1337.28015](#).
- [KK25+] S. Kittle and C. Kogler. *Dimension of contracting on average self-similar measures*. Preprint. [arxiv:2501.17795v1](#).
- [MU96] R. D. Mauldin and M. Urbański. *Dimensions and measures in infinite iterated function systems*. *Proc. Lond. Math. Soc.* **73** (1996), 105–154. [zbl:0852.28005](#).
- [MU99] R. D. Mauldin and M. Urbański. *Conformal iterated function systems with applications to the geometry of continued fractions*. *Trans. Am. Math. Soc.* **351** (1999), 4995–5025. [zbl:0940.28009](#).

- [MW86] R. D. Mauldin and S. C. Williams. *Random recursive constructions: asymptotic geometric and topological properties*. Trans. Am. Math. Soc. **295** (1986), 325–346. [zbl:0625.54047](#).
- [Rap22] A. Rapaport. *Proof of the exact overlaps conjecture for systems with algebraic contractions*. Ann. Sci. Éc. Norm. Supér. **55** (5 2022), 1357–1377. [zbl:1527.28009](#).
- [RV24] A. Rapaport and P. P. Varjú. *Self-similar measures associated to a homogeneous system of three maps*. Duke Math. J. **173** (2024), 513–602. [zbmath:07835917](#).
- [Var19] P. P. Varjú. *On the dimension of Bernoulli convolutions for all transcendental parameters*. Ann. Math. **189** (2019), 1001–1011. [zbl:1426.28024](#).
- [Wu25+] M. Wu. *Projection theorems with countably many exceptions and applications to the exact overlaps conjecture*. Preprint. [arxiv:2503.21923v2](#).

**Alex Rutar**

*Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box 35 (MaD), FI-40014 University of Jyväskylä, Finland*

Email: [alex@rutar.org](mailto:alex@rutar.org)