

# LENGTH FUNCTIONS AND THE DIMENSION OF POINTS IN SELF-SIMILAR FRACTAL TREES

CHRISTOPHER P. PORTER

**ABSTRACT.** In this paper, we study the effective dimension of points in infinite fractal trees generated recursively by a finite tree over some alphabet. Using unequal costs coding, we associate a length function with each such fractal tree and show that the channel capacity of the length function is equal to the similarity dimension of the fractal tree (up to a multiplicative constant determined by the size of the alphabet over which our tree is defined). Using this result, we derive formulas for calculating the effective dimension and strong effective dimension of points in fractal trees, establishing analogues of several results due to Lutz and Mayordomo, who studied the effective dimension of points in self-similar fractals in Euclidean space. Lastly, we explore the connections between the channel capacity of a length function derived from a finite tree and the measure of maximum entropy on a related directed multigraph that encodes the structure of our tree, drawing on work by Abram and Lagarias on path sets, where a path set is a generalization of the notion of a sofic shift.

## 1. OPENING

The tools from algorithmic information theory, particularly the notion of effective dimension of an individual point, have found successful application in the study of fractal geometry. As a particularly striking example, Lutz and Mayordomo in [LM08] provide a general formula for calculating the effective dimension  $\dim(x)$  of a point  $x$  in a self-similar fractal  $F$  in  $\mathbb{R}^n$ :

$$\dim(x) = \text{sdim}(F)\dim^\mu(y),$$

where  $\text{sdim}(F)$  is the similarity dimension of  $F$ ,  $\dim^\mu(\cdot)$  is an effective analogue of the Billingsley dimension with respect to a specific probability measure  $\mu$  defined in terms of the fractal  $F$ , and  $y$  is an infinite sequence over some finite alphabet that serves as a code for  $x$  as an element of  $F$  (we will define all of these notions in Section 4 below). They further established an analogous for effective strong dimension  $\text{Dim}(x)$  of points  $x \in F$ .

In this study, we prove similar results for points in infinite self-similar trees over some finite alphabet. Our main proof relies on the machinery of coding with unequal costs from information theory, which in our context, amounts to considering notions of algorithmic information theory in terms of length functions that do not necessarily measure the length of a string of symbols merely in terms of the number of symbols in the string. More generally, we will identify, for a given  $m$ -ary tree  $T$  for some  $m \in \mathbb{N}$ , the relationship between the following quantities from seemingly disparate areas:

- (1)  $\alpha$ , the channel capacity of a code with unequal letter costs that is defined in terms of  $T$ ,
- (2)  $\text{sdim}(F)$ , the similarity dimension of an infinite self-similar tree  $F$  generated by  $T$ , and
- (3)  $\log(\lambda)$ , the negative logarithm of the Perron eigenvalue of the adjacency matrix of a specific directed graph  $G$  determined by  $T$ ,

are given by the equalities

$$\alpha = \log(m) \cdot \text{sdim}(F) = \log(\lambda)$$

(where the logarithm here and henceforth is taken to the base 2). In the case that  $T$  is a binary tree, we get the equality of the three quantities (1)-(3) listed above. From the relationship between (1) and (2), we will derive our analogue of the above-mentioned Lutz/Mayordomo result that holds for points in self-similar trees. Using (3), we show how the measure  $\mu$  used in the calculation of the effective version of Billingsley dimension mentioned above can be obtained

from a transformation of the unique measure of maximal entropy on a specific subshift related to the tree  $T$  (in this case, the Parry measure on a specific sofic shift).

The outline of the remainder of this article is as follows. In Section 2, we provide background material, including the basics of effective dimension. In Section 3, we lay out a brief survey of previous work at the intersection of algorithmic information theory, fractal geometry, and symbolic dynamics in order to provide a broader context for the reader. Section 4 begins the original contributions of this work, in which we establish the above-described relationship between the channel capacity of a length function generated from a finite tree  $T$  and the similarity dimension of the self-similar fractal tree generated from  $T$ . Using this relationship, we derive the analogue of the Lutz/Mayordomo formulas for the dimension and strong dimension of points in the self-similar fractal generated from  $T$ . Finally, in Section 5, we take a dynamical systems perspective on self-similar fractal trees. In particular, we show how to obtain a directed multi-graph  $G$  from a given finite tree  $T$ , where the one-sided infinite walks through  $G$  are precisely the infinite paths through the infinite self-similar fractal  $F$  generated by  $T$ , thereby establishing the equivalence between channel capacity of the length function with the value  $\log(\lambda)$  described in (3) above. Using properties of  $G$ , we derive the measure of maximum entropy on the closure of  $F$  under the shift operator and prove that this measure is equivalent to one naturally defined in terms of the length function associated with the tree  $T$ .

## 2. BACKGROUND

For  $k \in \mathbb{N}$ ,  $\Sigma_k^*$  consists of all finite strings over the alphabet  $\Sigma_k = \{0, 1, \dots, k-1\}$ . Similarly,  $\Sigma_k^\infty$  consists of all infinite sequences over the same alphabet. We will write elements of  $\Sigma_k^*$  as  $\sigma, \tau$ , and so on, while elements of  $\Sigma_k^\infty$  will be written as  $x, y, z$ , and so on. The concatenation of strings  $\sigma, \tau \in \Sigma_k^*$  is written as  $\sigma \frown \tau$ . The standard length of a string  $\sigma \in \Sigma_k^*$  is written as  $|\sigma|$  (we will consider more general length functions in Section 4). Given  $\sigma, \tau \in \Sigma_k^*$ , we write  $\sigma \preceq \tau$  to indicate that  $\sigma$  is an initial segment of  $\tau$ ; similarly, if  $\sigma$  is an initial segment of some  $x \in \Sigma_k^\infty$ , we write  $\sigma \prec x$ . Given  $x \in \Sigma_k^\infty$  and  $i \in \mathbb{N}$ ,  $x \upharpoonright i$  is the initial segment of  $x$  of length  $i$ . More generally, given  $x \in \Sigma_k^\infty$  and  $m, i \in \mathbb{N}$ ,  $x \upharpoonright [m, m+i)$  is the string  $\tau \in \Sigma_k^*$  of length  $i$  such that  $\tau(j) = x(m+j)$  for  $j = 0, \dots, i-1$ . A set  $S \subseteq \Sigma_k^*$  is *prefix-free* if for  $\sigma, \tau \in S$ ,  $\sigma \preceq \tau$  implies that  $\sigma = \tau$ .

The trees that we consider here are subsets of  $\Sigma_k^*$  that are closed downwards under  $\prec$ . Given a prefix-free set  $S \subseteq \Sigma_k^*$ , we can define a tree  $T$  by closing  $S$  downwards under  $\prec$ . In this case, in an abuse of terminology, we will refer to the members of  $S$  as the *terminal nodes* of  $T$ ; any element of  $T$  that is not a terminal node will be referred to as a *non-terminal node* of  $T$ . Hereafter, we will specify a finite tree simply by specifying its set of terminal nodes. Given an infinite tree  $T \subseteq \Sigma_k^*$ , we will write  $[T]$  as the set of infinite paths through  $T$ , i.e.,  $[T] = \{x \in \Sigma_k^\infty : (\forall n) x \upharpoonright n \in T\}$ .

For  $\sigma \in \Sigma_k^*$ , the *cylinder set defined by  $\sigma$*  is the set  $[\sigma] = \{x \in \Sigma_k^\infty : \sigma \prec x\}$ . Let  $(\sigma_i)_{i \in \mathbb{N}}$  be the enumeration of  $\Sigma_k^*$  in length-lexicographical order. A set  $S \subseteq \Sigma_k^\infty$  is *effectively open* (or a  $\Sigma_1^0$  class) if there is some computable function  $f : \mathbb{N} \rightarrow \mathbb{N}$  such that  $S = \bigcup_{i \in \mathbb{N}} [\sigma_{f(i)}]$ .  $P \subseteq \Sigma_k^\infty$  is *effectively closed* (or a  $\Pi_1^0$  class) if  $\Sigma_k^\infty \setminus P$  is effectively open. Equivalently,  $P$  is a  $\Pi_1^0$  class if  $P = [T]$  for some infinite computable tree  $T \subseteq \Sigma_k^*$ .

A measure  $\mu$  on  $\Sigma_k^\infty$  is determined by its values on cylinder sets. Hereafter, for  $\sigma \in \Sigma_k^*$ , we will write  $\mu([\sigma])$  as  $\mu(\sigma)$ .  $\mu$  is a *computable* measure if  $\mu(\sigma)$  can be computably approximated to an arbitrary precision uniformly in  $\sigma \in \Sigma_k^*$ . A measure  $\mu$  on  $\Sigma_k^\infty$  is a *Bernoulli measure* if  $\mu(\sigma \frown i) = \mu(\sigma)\mu(i)$  for all  $\sigma \in \Sigma_k^*$  and  $i \in \Sigma_k$ .

We assume the reader is familiar with the basics of computability theory and algorithmic randomness; see, for instance, [Nie09], [DH10], or [SUV17]. See also [FP20] for an recent survey on algorithmic randomness.

For  $m, k \in \mathbb{N}$ , a Turing functional  $\Phi : \Sigma_k^\infty \rightarrow \Sigma_m^\infty$  is an effective map defined in terms of a computable function  $\phi : \Sigma_k^* \rightarrow \Sigma_m^*$  satisfying the property that for  $\sigma, \tau \in \Sigma_k^*$ ,  $\sigma \preceq \tau$  implies  $\phi(\sigma) \preceq \phi(\tau)$ . For such a function  $\phi$ , we can define  $\Phi$  on  $x \in \Sigma_k^\infty$  by setting  $\Phi(x) =$

$\lim_{n \rightarrow \infty} \phi(x \upharpoonright n)$  (where this limit is the longest element of  $\Sigma_m^* \cup \Sigma_m^\infty$  that has  $\phi(x \upharpoonright n)$  as an initial segment for all  $n \in \mathbb{N}$ ). In our context, we will only consider *total* Turing functionals, i.e., functionals given in terms of a computable function  $\phi$  for which  $\lim_{n \rightarrow \infty} |\phi(x \upharpoonright n)| = \infty$  for all  $x \in \Sigma_k^\infty$ .

Notions of the effective dimension of a sequence  $x \in \Sigma_k^\infty$  were introduced by Lutz in [Lut00] and further developed in, for instance, [May02], [Lut03], and [AHLM04]. Lutz originally defined effective dimension in terms of certain betting strategies he referred to as gales, but a characterization of dimension in terms of Kolmogorov complexity was later given by Mayordomo (in [May02]). We will use this latter characterization in the present study.

Fix  $m \in \mathbb{N}$ . Let  $M : \Sigma_2^* \rightarrow \Sigma_m^*$  be a prefix-free Turing machine (recall that a Turing machine  $M$  is prefix-free if the domain of  $M$  forms a prefix-free set). For  $\tau \in \Sigma_m^*$ , the *prefix-free Kolmogorov complexity of  $\tau$  with respect to  $M$*  is defined to be  $K_M(\tau) = \min\{|\sigma| : M(\sigma) = \tau\}$ . Let  $(M_i)_{i \in \mathbb{N}}$  be an effective enumeration of all prefix-free machines. We can define a universal prefix-free machine  $U : \Sigma_2^* \rightarrow \Sigma_m^*$  by setting, for each  $e \in \mathbb{N}$  and  $\sigma \in \Sigma_2^*$ ,  $U(1^e 0 \sigma) = M_e(\sigma)$  when the latter is defined (otherwise,  $U(1^e 0 \sigma)$  is undefined). For  $\tau \in \Sigma_m^*$ , we then define the *prefix-free Kolmogorov complexity of  $\tau$*  to be  $K(\tau) := K_U(\tau)$ . As shown by Kolmogorov [Kol65], for every prefix-free machine  $M$ ,  $K(\sigma) \leq K_M(\sigma) + O(1)$  for all  $\sigma \in \Sigma_m^*$ , where the additive constant only depends on  $M$ . Using Kolmogorov complexity, we can define two notions of effective dimension as follows. For  $x \in \Sigma_m^\infty$ ,

$$\dim(x) = \liminf_{n \rightarrow \infty} \frac{K(x \upharpoonright n)}{n \log(m)}$$

and

$$\text{Dim}(x) = \limsup_{n \rightarrow \infty} \frac{K(x \upharpoonright n)}{n \log(m)},$$

where the former notion is known as the *effective dimension* of  $x$  and the latter notion is known as the *effective strong dimension* of  $x$ .

In [LM08], Lutz and Mayordomo introduced two effective notions of Billingsley dimension that generalize the above two notions. For a computable measure  $\mu$  on  $\Sigma_m^\infty$  and  $x \in \Sigma_m^\infty$ , we have

$$\dim^\mu(x) = \liminf_{n \rightarrow \infty} \frac{K(x \upharpoonright n)}{-\log \mu(x \upharpoonright n)}$$

and

$$\text{Dim}^\mu(x) = \limsup_{n \rightarrow \infty} \frac{K(x \upharpoonright n)}{-\log \mu(x \upharpoonright n)},$$

the *effective dimension of  $x$  with respect to  $\mu$*  and the *effective strong dimension of  $x$  with respect to  $\mu$* , respectively.

### 3. THE BIGGER PICTURE

Before we turn to the main contributions of this study, we first offer a brief survey of the key developments that lie at the intersection of algorithmic information theory, fractal geometry, and symbolic dynamics, which should in turn provide the reader with a broader context to understand and appreciate the results contained herein.

The relationship between Kolmogorov complexity and various notions of entropy was observed early in the development of algorithmic information theory. Indeed, as early as 1970, merely five years after Kolmogorov first offered his definition of complexity in [Kol65], Zvonkin and Levin [ZL70] proved that if we consider binary sequences produced as the result of a sequence of independent, identically distributed random variables with probabilities  $(p, 1 - p)$  for some computable  $p \in (0, 1)$ , for  $\mu$ -almost every sequence  $x \in \Sigma_2^\infty$ , where  $\mu$  is the Bernoulli measure on  $\Sigma_2^\infty$  with parameters  $(p, 1 - p)$ , we have

$$\lim_{n \rightarrow \infty} \frac{K(x \upharpoonright n)}{n} = H(p) = -p \log(p) - (1 - p) \log(1 - p)$$

(in fact, Zvonkin and Levin showed the stronger result that the above expression holds for all sequences that are Martin-Löf random with respect to the measure  $\mu$ ; the details of this notion of randomness is not necessary for this study, but see, for instance, [Nie09, Chapter 3], [DH10, Chapter 6], or [SUV17, Chapter 3]). A similar result was obtained soon thereafter by Heim [Hei79] using a variant of Kolmogorov complexity defined in terms of block codes. Subsequent work by Brudno [Bru82] and White [Whi93] characterized the complexity of the codes of trajectories of points in a dynamical system on a compact space using a complexity measure defined explicitly in terms of the asymptotic behavior of  $K(x|n)/n$  (which we do not define here), where  $x$  is the code of a trajectory of a point in the space in question. In this setting, for an ergodic measure  $\mu$ , the complexity for  $\mu$ -almost every point what shown to be equal to the Kolmogorov-Sinai entropy of  $\mu$ .

The first explicit results on algorithmic notions of fractal dimension are due to Lutz [Lut00], who defined an effective version of Hausdorff dimension applicable to both individual sequences and collections of sequences using certain betting strategies known as gales. Soon after Lutz's initial contribution, Mayordomo [May02] characterized Lutz's effective notion of Hausdorff dimension in terms of Kolmogorov complexity (a characterization we have used in our definition of effective dimension given at the end of the previous section). In addition, Athreya, Hitchcock, Lutz, and Mayordomo [AHLM04] defined an effective version of packing dimension and provided alternative characterizations of the notion in terms of both gales and Kolmogorov complexity (with the latter characterization is the definition of strong effective dimension given at the end of the previous section).

With this initial framework of effective notions of fractal dimension in place, a considerable amount of work followed. Highlights relevant to this study include:

- Simpson's work [Sim15] on the effective Hausdorff dimension of subshifts  $X \subseteq A^G$  (where  $A$  is a finite set of symbols and  $G = \mathbb{N}^d$  or  $G = \mathbb{Z}^d$  for some  $d \in \mathbb{N}$ ), which established that the effective dimension of  $X$  is equal to the topological entropy of  $X$ ;
- the extension of effective notions of dimension to Euclidean space, for instance to study dimension of points in a self-similar fractal by Lutz and Mayorodomo [LM08], as discussed in the introduction, as well as the effective dimension of points in a random subfractal of a self-similar fractal [GLMM14];
- the improvement by Hoyrup [Hoy12] of the original Zvonkin-Levin result described above to the context of computable shift-invariant, ergodic measures;
- the study of the dimensions of points on various lines in the plane by N. Lutz and Stull [LS17], [LS20];
- the use of point-to-set principles for effective notions of dimension to characterize the classical dimension of various sets.

This latter point bears further unpacking. As shown by Lutz and N. Lutz [LL18], for every  $E \subseteq \mathbb{R}^n$ , the classical Hausdorff dimension  $\dim_H(E)$  can be calculated from an analogue of the effective dimension of the points in  $E$  via the expression

$$\dim_H(E) = \min_{A \subseteq \mathbb{N}} \sup_{x \in E} \dim^A(x),$$

where  $\dim^A(x)$  is the effective dimension of  $x$  relative to the oracle  $A$ ; a similar expression holds for classical packing dimension. These point-to-set principles have proven to be powerful tools, allowing for the solution of open problems in classical fractal geometry, for instance, problems involving the intersections and products of fractals in Euclidean space due to N. Lutz [Lut21] and a generalization of Marstrand's projection theorem due to N. Lutz and Stull [LS18]. In short, the point-to-set principles have allowed for information-theoretic techniques to be imported into the study of fractal geometry to great effect.

The present study contributes to work in this area in several respects. First, we introduce the machinery of unequal costs coding into the study of effective notions of fractal dimension, a new development in this area. This study can thus be seen as providing a proof of concept of the utility of unequal costs coding in a relatively constrained setting, i.e., in the context

of self-similar fractal trees, obtaining new results on the effective dimension of points in such trees in a way that makes fundamental use of aspects of unequal costs coding (namely, the relationship between the similarity dimension of a self-similar fractal and the channel capacity of an unequal costs code, as laid out in Theorem 4.1), providing a more efficient approach than using previously developed techniques in this area.

Second, in the constrained setting of self-similar fractal trees, we unearth new connections between the structure of such trees, the properties of the associated unequal cost codes, and certain symbolic dynamical features of directed graphs derived from these trees, connections that do not readily generalize to a broader setting in any obvious way. In addition, this connection relies on a new result on the characteristic polynomial of adjacency matrices of the above-mentioned directed graphs, which we establish using ideas from spectral graph theory, an area with which the literature on algorithmic information theory has engaged very little. Moreover, previous work on the overlap between symbolic dynamics and algorithmic information theory does not explicitly discuss the significance of the Parry measure in this context, whereas the Parry measure features prominently in the discussion in Section 5.

Finally, we anticipate that the machinery of unequal costs coding will have a wider range of applicability in the study of effective fractal dimensions. For instance, although the point-to-set principles are not needed to understand the structure of the self-similar fractal trees we consider here (as they are essentially finitary objects, being recursively generated from finite trees), if one were to consider a natural extension of self-similar fractal trees that involve a more general type of length function (for instance, one in which the cost of a symbol may vary depending on its location in a given sequence), this might allow one to obtain similar results with more general self-similar systems using sophisticated tools such as the point-to-set principles.

With this context in mind, we now turn to the main results of this study.

#### 4. CODING WITH UNEQUAL COSTS AND SELF-SIMILAR TREES

**4.1. Unequal costs coding.** Let us consider the finite alphabet  $\Sigma_k$  for some fixed  $k \geq 2$ . As defined in the information theory literature on coding with unequal costs, we define a length function to simply be a function  $\ell : \Sigma_k \rightarrow \mathbb{N}$ ; see, for instance, [Var71, Cha86, Abr94, Abr97, GKY02]. Note that we could consider a broader notion of length functions, for instance, length functions that are real-valued, length functions that are not defined merely in terms of individual symbols but may give a symbol varying cost depending on where it is located in a given string or depending on the symbols that precede it, and so on. However, for our purposes, we only need the simple notion of a length function as defined above.

A key notion in the study of coding with unequal letter costs that will be useful in this study is that of a channel capacity [BBM<sup>+</sup>10], [CK11, Chapter 4]. For a length function  $\ell$  defined on alphabet  $\Sigma_k$ , the *channel capacity* associated to  $\ell$  is the unique  $\alpha \in \mathbb{R}^{\geq 0}$  that satisfies the expression

$$\sum_{i=0}^{k-1} 2^{-\alpha \ell(i)} = 1. \quad (\dagger)$$

Informally, the channel capacity of a length function measures the average amount of information transmitted per unit cost (see the discussion in [CK11, Chapter 4]). It is not difficult to see that the channel capacity associated to  $\ell$  is unique. Indeed, if we set  $r = 2^{-\alpha}$ , we can rewrite Equation  $(\dagger)$  as

$$\sum_{i=0}^{k-1} r^{\ell(i)} = 1.$$

As the equation  $f(r) = \sum_{i=0}^{k-1} r^{\ell(i)} - 1$  is strictly increasing on  $[0,1]$  with  $f(0) < 0$  and  $f(1) > 0$ , uniqueness follows. Observe more generally that the channel capacity  $\alpha$  associated to a length function  $\ell$  on  $\Sigma_k$  allows us to define a Bernoulli measure  $\mu_\ell$  on  $\Sigma_k^\infty$  induced by setting, for each  $i \in \Sigma_k$ ,

$$\mu_\ell(i) = r^{\ell(i)},$$

where  $-\log(r) = \alpha$ . We will refer to  $\mu_\ell$  as the *measure derived from  $\ell$* .

**4.2. Self-similar fractal trees.** We now turn to defining the similarity dimension of a self-similar fractal generated by a finite tree. Let  $T \subseteq \Sigma_m^*$  be a finite tree for some  $m \in \mathbb{N}$ . Suppose that  $T$  has  $k$  terminal nodes  $\tau_0, \dots, \tau_{k-1}$ . We define an infinite, self-similar tree  $T^*$  from  $T$  simply by concatenating a copy of  $T$  at each of its terminal nodes and then repeating this process recursively.

Let us define  $F_T = [T^*]$ . Clearly,  $F_T$  is a self-similar fractal, as above each terminal node of  $T$ , we place a full copy of  $F_T$ ; hereafter, we will refer to  $F_T$  as the *self-similar fractal tree generated by  $T$* .

Since each such copy of  $F_T$  begins at some level of the tree deeper than the root, this amounts to scaling that copy. In our context, this scaling factor is precisely the ratio of the size of the copy of  $F_T$  to the size of  $F_T$ , which is precisely the similarity ratio from fractal geometry (see [Fal04, Section 9.2]). Here, size is given by the standard metric  $d$  on  $\Sigma_m^\infty$  according to which two sequences  $x, y \in \Sigma_m^\infty$  that agree on the first  $j$  symbols but disagree on the  $(j+1)$ -st satisfy  $d(x, y) = m^{-j}$ .

Let us calculate the similarity ratios for these copies of  $F_T$ . Given a terminal node  $\tau$  in  $T$ , by concatenating a copy of  $F_T$  at the end of  $\tau$ , this amounts to scaling  $F$  by a factor of  $m^{-|\tau|}$ . Thus, the similarity ratios corresponding to the terminal nodes  $\tau_0, \dots, \tau_{k-1}$  are equal to  $m^{-|\tau_0|}, \dots, m^{-|\tau_{k-1}|}$ , respectively. Then the similarity dimension  $\text{sdim}(F)$  of  $F$  is the unique real number  $\beta$  satisfying

$$\sum_{i=0}^{k-1} m^{-\beta|\tau_i|} = 1. \quad (\dagger\dagger)$$

**4.3. Relating length functions and self-similarity.** We now spell out how coding with unequal costs is related to the similarity dimension of a self-similar tree.

Given  $T$  as above with terminal nodes  $\tau_0, \dots, \tau_{k-1}$ , we associate to each terminal node  $\tau_i$  the symbol  $i \in \Sigma_k$  and define a length function  $\ell_T$  satisfying

$$\ell_T(i) = |\tau_i|$$

for each  $i \in \{0, \dots, k-1\}$ . Hereafter, let us refer to  $\ell_T$  as the *length function induced by  $T$* . We then have the following.

**Theorem 4.1.** *Let  $T \subseteq \Sigma_m^*$  be a finite tree for some  $m \in \mathbb{N}$ . Let  $\ell_T$  be the length function induced by  $T$  and  $F_T$  the self-similar fractal generated by  $T$ . Then the channel capacity of  $\ell_T$  is equal to  $\log(m) \cdot \text{sdim}(F)$ . In particular, in the case that  $m = 2$ , the channel capacity of  $\ell_T$  is equal to  $\text{sdim}(F_T)$ .*

*Proof.* From Equation  $(\dagger)$ , the channel capacity  $\alpha$  of  $\ell_T$  satisfies

$$\sum_{i=0}^{k-1} 2^{-\alpha \ell_T(i)} = 1.$$

Since  $\ell_T(i) = |\tau_i|$  for  $i \in \Sigma_k$ , changing the base of the exponential from 2 to  $m$ , we can rewrite above equation as

$$\sum_{i=0}^{k-1} m^{-\frac{\alpha}{\log(m)}|\tau_i|} = 1.$$

But from Equation  $(\dagger\dagger)$ ,  $\log(m) \cdot \text{sdim}(F_T)$  is the unique solution to the above equation, and hence we can conclude that  $\log(m) \cdot \text{sdim}(F_T) = \alpha$ .  $\square$

**4.4. Deriving the Lutz/Mayordomo formulas for self-similar trees.** The observation that the similarity dimension of a self-similar  $F \subseteq \Sigma_m^\infty$  generated by a finite tree  $T \subseteq \Sigma_m^*$  is equal to the channel capacity of the unequal costs code determined by  $T$  allows us to derive analogues of the Lutz/Mayordomo formulas for the effective dimension of points in self-similar fractals as stated in the introduction by means of a relatively straightforward proof.

First, we need to define one additional notion. Given a finite tree  $T \subseteq \Sigma_m^*$  with  $k$  terminal nodes  $S = \{\tau_0, \dots, \tau_{k-1}\}$  for some  $m, k \in \mathbb{N}$  and the corresponding self-similar fractal  $F_T$ , we define a one-to-one correspondence  $\Psi$  between elements of  $F_T$  and elements of  $\Sigma_k^\infty$  as follows. First we define a coding map  $\psi : S \rightarrow \Sigma_k$  by setting  $\psi(\tau_i) = i$  for  $i \in \Sigma_k$ . We extend  $\psi$  to a functional  $\Psi : F_T \rightarrow \Sigma_k^\infty$  as follows. Given  $x \in F_T$ , there is a sequence  $(n_i)_{i \in \mathbb{N}} \subseteq \mathbb{N}$  with  $n_0 = 0$  such that for each  $i \in \mathbb{N}$ ,  $x \upharpoonright [n_i, n_{i+1})$  is a terminal node of  $T$ ; let us call this the  $T$ -sequence of  $x$ . Note that since  $S$  is prefix-free, the  $T$ -sequence of each  $x \in F_T$  is unique. We then define  $\Psi_T$  by setting

$$\Psi_T(x) = \psi(x \upharpoonright [n_0, n_1)) \frown \psi(x \upharpoonright [n_1, n_2)) \frown \dots \frown \psi(x \upharpoonright [n_i, n_{i+1})) \frown \dots$$

Setting  $y = \Psi_T(x)$ , we will hereafter refer to  $y$  as the *coding sequence* of  $x$ . Note that we can similarly extend  $\psi$  to a function  $\hat{\psi} : S^* \rightarrow \Sigma_k^*$ , which we will use shortly. We can now state the analogue of the Lutz/Mayordomo formulas in our context.

**Theorem 4.2.** *Let  $T \subseteq \Sigma_m^*$  be a finite tree with  $k$  terminal nodes for some  $m, k \in \mathbb{N}$  with corresponding length function  $\ell = \ell_T$ , self-similar fractal  $F_T$ , and the measure  $\mu_\ell$  derived from  $\ell$ . Then for  $x \in F_T$ ,*

$$\dim(x) = \text{sdim}(F_T) \dim^{\mu_\ell}(y)$$

and

$$\text{Dim}(x) = \text{sdim}(F_T) \text{Dim}^{\mu_\ell}(y),$$

where  $y \in \Sigma_k^\infty$  is the coding sequence of  $x \in F_T$ .

*Proof.* Let  $S = \{\tau_0, \dots, \tau_k\}$  be the set of terminal nodes of  $T$ , and let  $\hat{\psi} : S^* \rightarrow \Sigma_k^*$  be the extension of  $\psi$  as discussed above. Note that  $\ell(i) = |\tau_i|$  for  $i \in \Sigma_k$ . Given  $x \in F_T$ , let  $(n_i)_{i \in \mathbb{N}} \subseteq \mathbb{N}$  be the  $T$ -sequence of  $x$  and  $y \in \Sigma_k^\infty$  the coding sequence for  $x$ . We first establish the following claim:

*Claim:* For  $j \in \mathbb{N}$ ,  $K(x \upharpoonright n_j) = K(y \upharpoonright j) + O(1)$ .

The argument to establish this claim is fairly routine, but we include it here for the sake of completeness. First note that  $\hat{\psi}(x \upharpoonright n_j) = y \upharpoonright j$ . Indeed, since the  $T$ -sequence of  $x$  is unique, we calculate the value  $\hat{\psi}(x \upharpoonright n_j)$  first by decomposing  $x \upharpoonright n_j$  into the concatenation of  $j$  elements of  $S$ . Applying  $\psi$  to each of these in turn and concatenating the outputs yields the string  $y \upharpoonright j$ .

Next we define a machine  $M : \Sigma_2^* \rightarrow \Sigma_m^*$  as follows. For  $\sigma \in \Sigma_2^*$ , if  $U(\sigma) \downarrow$  and  $U(\sigma)$  can be written as the concatenation of members of  $S$  (so that  $\hat{\psi}(U(\sigma))$  is defined), then we set  $M(\sigma) = \hat{\psi}(U(\sigma))$ . In particular,  $U(\sigma) \downarrow = x \upharpoonright n_j$  if and only if  $M(\sigma) \downarrow = y \upharpoonright j$ . By the optimality of  $K$ , we have  $K(y \upharpoonright j) \leq K_M(y \upharpoonright j) + O(1) \leq K(x \upharpoonright n_j) + O(1)$ . Similarly, as  $\hat{\psi}^{-1}$  is invertible, we can use  $\hat{\psi}^{-1}$  we can establish that  $K(x \upharpoonright n_j) \leq K(y \upharpoonright j) + O(1)$ , and so we have established the claim.

There exists a sequence  $\tau_{i_0}, \dots, \tau_{i_{j-1}}$  of elements of  $S$  such that  $x \upharpoonright n_j = \tau_{i_0} \frown \dots \frown \tau_{i_{j-1}}$ . Then if  $-\log(r)$  is the channel capacity associated to  $\ell$ ,

$$\begin{aligned} n_j &= \sum_{e=0}^{j-1} |\tau_{i_e}| = \sum_{e=0}^{j-1} \ell(i_e) = \ell(y \upharpoonright j) \\ &= \left( \frac{1}{-\log(r)} \right) \left( -\log(r^{\ell(y \upharpoonright j)}) \right) \\ &= \left( \frac{1}{-\log(r)} \right) \left( -\log(\mu_\ell(y \upharpoonright j)) \right). \end{aligned} \tag{1}$$

Then we have

$$\begin{aligned} \frac{K(x \upharpoonright n_j)}{n_j \log(m)} &= \frac{K(y \upharpoonright j) + O(1)}{n_j \log(m)} \\ &= \left( \frac{-\log(r)}{\log(m)} \right) \frac{K(y \upharpoonright j) + O(1)}{-\log(\mu_\ell(y \upharpoonright j))} \\ &= \text{sdim}(F_T) \frac{K(y \upharpoonright j) + O(1)}{-\log(\mu_\ell(y \upharpoonright j))}, \end{aligned}$$

where the first equality follows from the above claim, the second equality follows from Equation (1), and the third equality follows from the fact that  $-\log(r)$  is the channel capacity of  $\ell$  and  $-\log(r) = \log(m) \cdot \text{sdim}(F_T)$  by Theorem 4.2. Taking the limit infimum and limit supremum of both sides of this equality yields

$$\liminf_{j \rightarrow \infty} \frac{K(x \upharpoonright n_j)}{n_j \log(m)} = \text{sdim}(F_T) \dim^{\mu_\ell}(y) \tag{2}$$

and

$$\limsup_{j \rightarrow \infty} \frac{K(x \upharpoonright n_j)}{n_j \log(m)} = \text{sdim}(F_T) \text{Dim}^{\mu_\ell}(y). \tag{3}$$

Moreover, we have

$$\dim(x) = \liminf_{n \rightarrow \infty} \frac{K(x \upharpoonright n)}{n \log(m)} \leq \liminf_{j \rightarrow \infty} \frac{K(x \upharpoonright n_j)}{n_j \log(m)}$$

and

$$\limsup_{j \rightarrow \infty} \frac{K(x \upharpoonright n_j)}{n_j \log(m)} \leq \limsup_{n \rightarrow \infty} \frac{K(x \upharpoonright n)}{n \log(m)} = \text{Dim}(x),$$

from which, combined with Equations (2) and (3), we conclude

$$\dim(x) \leq \text{sdim}(F_T) \dim^{\mu_\ell}(y) \tag{4}$$

and

$$\text{sdim}(F_T) \dim^{\mu_\ell}(y) \leq \text{Dim}(x). \tag{5}$$

To conclude the proof, we have to consider the values  $\frac{K(x \upharpoonright n)}{n}$  for  $n \in (n_j, n_{j+1})$  for  $j \in \mathbb{N}$ . We use the subadditivity of  $K$ : there is some  $c \in \mathbb{N}$  such that for all  $\sigma, \tau \in \Sigma_k^*$ ,  $K(\sigma \frown \tau) \leq K(\sigma) + K(\tau) + c$ . Let  $S_0$  consist of all proper suffixes of the terminal nodes of  $T$ , and we let  $s_0 = \max\{K(\sigma) : \sigma \in S_0\}$ . Given  $n \in (n_j, n_{j+1})$ , we can write  $n = n_j + i$  for some  $i < n_{j+1} - n_j$ . Then

$$\begin{aligned} \frac{K(x \upharpoonright n_{j+1})}{n_{j+1}} &\leq \\ \frac{K(x \upharpoonright (n_j + i)) + K(x \upharpoonright [n_j + i, n_{j+1}]) + c}{n_{j+1}} &\leq \\ \frac{K(x \upharpoonright (n_j + i)) + c + s_0}{n_{j+1}}, \end{aligned}$$

where the first inequality follows from the subadditivity of  $K$  and the second from the fact that  $x \upharpoonright [n_j + i, n_{j+1}) \in S_0$ . Rearranging yields

$$\begin{aligned} \frac{K(x \upharpoonright n_{j+1}) - O(1)}{n_{j+1}} &\leq \frac{K(x \upharpoonright (n_j + i))}{n_{j+1}} \\ &\leq \frac{K(x \upharpoonright (n_j + i))}{n_j + i} = \frac{K(x \upharpoonright n)}{n}. \end{aligned}$$

Thus, for all  $n$  such that  $n \neq n_j$  for all  $j$ , there is some  $j$  such that  $n \in (n_j, n_{j+1})$  for which

$$\frac{K(x \upharpoonright n_j) - O(1)}{n_j} \leq \frac{K(x \upharpoonright n)}{n}.$$

Applying the limit infimum of both sides and applying Equation (2), we get

$$\text{sdim}(F_T) \dim^{\mu_\ell}(y) \leq \dim(x),$$

which, combined with Equation (4), gives us

$$\dim(x) = \text{sdim}(F_T) \dim^{\mu_\ell}(y).$$

Similarly, let  $S_1$  consist of all proper initial segments of the terminal nodes of  $T$ . Then we let  $s_1 = \max\{K(\sigma) : \sigma \in S_1\}$ . Given  $n \in (n_j, n_{j+1})$  with  $n = n_j + i$  for some  $i < n_{j+1} - n_j$  as above, we have

$$\begin{aligned} \frac{K(x \upharpoonright n)}{n} &= \frac{K(x \upharpoonright (n_j + i))}{n_j + i} \\ &\leq \frac{K(x \upharpoonright n_j) + K(x \upharpoonright [n_j, n_j + i)) + c}{n_j + i} \\ &\leq \frac{K(x \upharpoonright n_j) + c + s_1}{n_j + i} \\ &\leq \frac{K(x \upharpoonright n_j) + O(1)}{n_j} \end{aligned}$$

where the first inequality follows from the subadditivity of  $K$  and the second from the fact that  $x \upharpoonright [n_j, n_j + i) \in S_1$ . Thus, for all  $n$  such that  $n \neq n_j$  for all  $j$ , there is some  $j$  such that  $n \in (n_j, n_{j+1})$  for which

$$\frac{K(x \upharpoonright n)}{n} \leq \frac{K(x \upharpoonright n_j) + O(1)}{n_j}.$$

Thus

$$\begin{aligned} \text{Dim}(x) &= \limsup_{n \rightarrow \infty} \frac{K(x \upharpoonright n)}{n \log(m)} \\ &\leq \limsup_{n \rightarrow \infty} \frac{K(x \upharpoonright n_j) + O(1)}{n_j \log(m)} \end{aligned}$$

which, combined with Equation (5), gives us

$$\text{Dim}(x) = \text{sdim}(F_T) \text{Dim}^{\mu_\ell}(y).$$

□

## 5. A DYNAMICAL SYSTEMS PERSPECTIVE ON SELF-SIMILAR FRACTAL TREES

**5.1. Self-similar trees and path sets.** We next take a dynamical systems perspective on self-similar fractal trees generated by a finite tree. In this context, such a self-similar tree can be seen as coding the set of one-sided infinite walks through a certain directed graph, where every walk begins at a distinguished node, providing an instance of what has been referred to as a *path set* by Abrams and Lagarias in [AL14].

Fix  $n \in \mathbb{N}$ . A *pointed graph*  $(\mathcal{G}, v)$  over  $\Sigma_n$  consists of an edge-labeled finite directed graph  $\mathcal{G} = (G, \mathcal{E})$ , where  $G = (V, E)$  is a directed graph (in which loops and multiple edges are

permitted) with a distinguished vertex  $v \in V$  and  $\mathcal{E} \subseteq E \times \Sigma_n$  (i.e., the edges of  $G$  are labeled with elements of  $\Sigma_n$ ). We further require that the labels on different edges between any two fixed vertices must be distinct. We then define the *path set*  $P = X_{\mathcal{G}}(v) \subseteq \Sigma_n^\infty$  given by the pointed graph  $\mathcal{G}$  to be the set of infinite sequences obtained from the edge labels of all possible one-sided infinite walks in  $G$  that begin at the distinguished vertex  $v$ .

Path sets are more general than the sets of paths through self-similar fractal trees that we are considering here, but we can characterize the fractals under consideration in terms of a restricted class of path sets. First, note that a path set can be obtained from different pointed graphs; we shall refer to an underlying pointed graph for a path set  $P$  as a *presentation* of  $P$ . A directed graph  $G$  with edge labels  $\mathcal{E}$  is *right-resolving* if for every vertex  $v$  in  $G$ , all directed edges starting from  $v$  have distinct edge labels. In addition, a directed graph  $G$  is *irreducible* (or *strongly connected*) if for every pair of vertices  $u$  and  $v$  in  $G$ , there is a directed walk in  $G$  from  $u$  to  $v$ . It can be shown that every path set has a right-resolving presentation, but not every path set has an irreducible presentation (see [AL14]).

**Theorem 5.1.** *Fix  $m \in \mathbb{N}$ .*

- (i) *Every self-similar fractal tree  $F$  generated by a finite tree  $T \subseteq \Sigma_m^*$  can be obtained as a path set with a right-resolving, irreducible presentation.*
- (ii) *Every path set over  $\Sigma_m$  with a right-resolving, irreducible presentation in which (a) every multiple edge terminates in the distinguished vertex and (b) every cycle contains the distinguished vertex is a self-similar fractal tree generated by some finite tree  $T$ .*

*Proof.* (i) First we show how to associate a pointed directed graph  $\mathcal{G}$  with a given finite tree  $T$ . First let  $n$  be the number of non-terminal nodes in  $T$ , written  $\sigma_1, \dots, \sigma_n$  in length-lexicographic order, so that  $\sigma_1$  is the root of the tree (we begin counting at 1 instead of 0 to align with conventions from linear algebra that we will use shortly). For each such node  $\sigma_i$ , we will have a corresponding vertex  $v_i$  in  $G$ ; the vertex  $v_1$  corresponding to  $\sigma_1$  will be the distinguished vertex. Moreover, for each pair of nodes  $\sigma_i$  and  $\sigma_j$  such that  $\sigma_i$  is a parent of  $\sigma_j$ , we include in  $G$  a directed edge from  $v_i$  to  $v_j$  with the same label on the edge from  $\sigma_i$  to  $\sigma_j$ . Lastly, for each node  $\sigma_i$  that is the parent of a terminal node, we will include in  $G$  one directed edge from  $v_i$  to  $v_1$  for each terminal node connected to  $\sigma_i$ , labeling the directed edge with the same label on the edge connecting  $\sigma_i$  to the corresponding terminal node in  $T$ . It is not difficult to verify that the set of infinite paths  $F$  through  $T^*$  is equal to the set of infinite one-way walks through  $G_T$  that begin at  $v_1$ .

(ii) Suppose  $\mathcal{G} = (G, v)$  is a right-resolving, irreducible directed graph satisfying the conditions (a) and (b) as above. The construction of the tree  $T$  is straightforward. Let  $u$  be a vertex for which there is a directed edge that has the distinguished vertex  $v$  as its terminal vertex. For each such directed edge, we disconnect it from  $v$  and connect it to a new vertex  $w$  that is not connected to any other vertex in the graph, keeping the same edge label as in the original directed graph (note that there may be multiple new vertices connected to  $u$  to which we perform this operation). Repeating this process for all such vertices  $u$  yields a new pointed graph  $\mathcal{G}' = (G', v)$ . We claim that the underlying graph of  $G'$  is a tree with root  $v$ .

First, note that there are no multiple edges in  $G'$ , since by condition (a), the only multiple edges in  $G$  have been removed and replaced with new edges terminating in distinct vertices. Second, we claim that  $G'$  is acyclic. Indeed, if there were a cycle in  $G'$ , then it would be a cycle in  $G$  (since no cycles are created in transforming  $G$  to  $G'$ ). By condition (b), this cycle contains the distinguished vertex  $v$ . But this contradicts the fact that no directed edge terminating in  $v$  in  $G$  is contained in  $G'$  by our construction. Finally, since  $G$  is irreducible,  $G'$  is connected (but not strongly connected), since the transformation from  $G$  to  $G'$  only results in the removing of cycles, keeping all other edge relations intact. If we take  $T$  to be the underlying graph of  $G'$  (removing the orientation of the directed edges), this is the desired tree  $T$ . As in the proof of (i), it is routine to verify that the set of infinite one-way walks through  $\mathcal{G}$  beginning at  $v$  is equal to the set of infinite paths  $F$  through  $T^*$ .  $\square$

See Figure 1 for an example of a finite tree and its corresponding pointed directed graph.

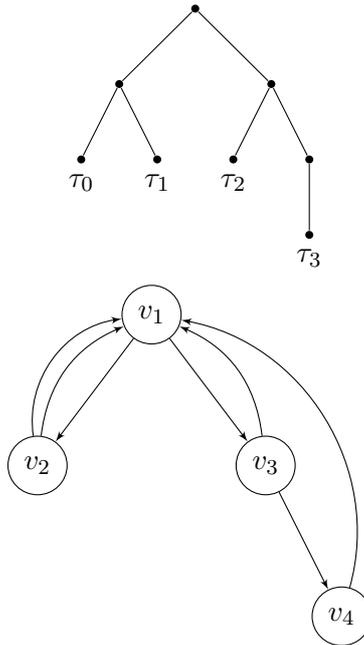


FIGURE 1. A finite tree  $T$  and its corresponding pointed directed graph  $G$

**5.2. Spectral analysis of the pointed graph associated with a finite tree.** The relationship between self-similar fractal trees and path sets bears further fruit for us, particularly when we study the measure  $\mu_\ell$  associated with the length function  $\ell_T$  derived from a finite tree  $T$ . In particular, for a given finite tree  $T$  and associated directed graph  $G$ , we can derive the Perron eigenvalue associated with  $G$  (which we define below) solely in terms of properties of the tree  $T$ . From this Perron eigenvalue, we can derive the so-called Parry measure on the set of infinite walks through  $G$  that begin at any vertex. We will further show the measure  $\mu_\ell$  of the collection of coding sequences for members of  $F_T$  can be derived from this Parry measure.

Fix  $m \in \mathbb{N}$ . Given a finite binary tree  $T \subseteq \Sigma_m^*$  with  $n$  non-terminal nodes, let  $\mathcal{G} = (G, \mathcal{E})$  be the corresponding pointed directed graph on  $n$  vertices  $v_1, \dots, v_n$  as in the proof of Theorem 5.1(i). Let  $A$  be the adjacency matrix of  $G$ , i.e., the  $n \times n$  matrix where for  $1 \leq i, j \leq n$ ,  $a_{i,j}$  is the number of directed edges from vertex  $v_i$  to  $v_j$  in  $G$ . Since the entries of  $A$  are non-negative and the graph  $G$  is irreducible (so that the matrix  $A$  is irreducible), by the Perron-Frobenius theorem (see, e.g., [LM21, Theorem 4.2.3]),  $A$  has a largest positive real eigenvalue  $\rho$  (the so-called *spectral radius* of  $A$ ). As noted in the introduction, we can show that  $\log(\rho)$  is equal to the channel capacity of  $\ell_T$ , the length function induced by  $T$ , which we obtain using the following theorem.

**Theorem 5.2.** *Given a finite tree  $T$  with  $n$  non-terminal nodes, let  $A$  be the  $n \times n$  adjacency matrix of the directed graph associated to  $T$ , and let  $p_A(z) = z^n + c_1 z^{n-1} + \dots + c_{n-1} z + c_n$  be the characteristic polynomial of  $A$ . Then for  $i = 1, \dots, n$ ,  $-c_i$  is equal to the number of terminal nodes in  $T$  of depth  $i$ .*

*Proof.* We proceed by induction on  $k \leq n$ . Let  $G$  be the directed graph derived from  $T$  as in the proof of Theorem 5.1(i). For  $k = 1$ , it is a standard fact in linear algebra that for the characteristic polynomial associated with a matrix  $A$ , we have  $c_1 = -\text{trace}(A)$ . Note that the trace of  $A$  counts the number of self-loops in  $G$ . By the construction of  $G$  from  $T$ , the only vertex of  $G$  that can have a self-loop is the vertex  $v_1$ , where each such self-loop in  $G$  corresponds to a terminal node in  $T$  of length 1. Thus,  $-c_1$  is the number of terminal nodes of  $T$  of depth 1.

For a fixed  $k$  with  $2 \leq k \leq n$ , assume that we have established the result for all  $j < k$ . We will use the Leverrier's algorithm for calculating the coefficients of the characteristic polynomial of an  $n \times n$  matrix  $A$  (see, e.g., [Gan59]), according to which for  $2 \leq k < n$ ,

$$c_k = -\frac{1}{k} \text{trace}(A^k + c_1 A^{k-1} + \dots + c_{k-1} A).$$

We begin this case by noting that the number of terminal nodes of  $T$  of depth  $k$  is equal to the number of  $k$ -cycles in  $G$  (that is, cycles of length  $k$ ) that begin at  $v_1$  and only pass through  $v_1$  once. We further observe that the entries along the diagonal of  $A^k$ , each of which contributes to the trace of  $A^k$ , count the total number of  $k$ -cycles in  $G$ , where this number is the sum of

- (a) the number of  $k$ -cycles in  $G$  that begin at  $v_1$  and only pass through  $v_1$  once,
- (b) the number of  $k$ -cycles in  $G$  that begin at some  $v \neq v_1$  and only pass through  $v_1$  once, and
- (c) the number of  $k$ -cycles in  $G$  that pass through  $v_1$  more than once (regardless of where they start).

As the number of cycles counted in (a) and (b) above is equal to  $k$  times the number of terminal nodes of  $T$  of depth  $k$ , each cycle counted in (a) is counted  $k - 1$  additional times in (b) (corresponding to the  $k - 1$  different vertices other than  $v_1$  at which the each such cycle can start).

Next, for any  $j$  with  $1 \leq j \leq k - 1$ , by our inductive hypothesis, we have

$$\begin{aligned} -\text{trace}(c_j A^{k-j}) &= \\ & (\# \text{ of terminal nodes in } T \text{ of depth } j) \\ & \quad \times (\# \text{ of } (k-j)\text{-cycles in } G) \\ &= -(\# \text{ of } j\text{-cycles in } G \text{ satisfying (a)}) \\ & \quad \times (\# \text{ of } (k-j)\text{-cycles in } G). \end{aligned}$$

This latter expression counts all  $k$ -cycles in  $G$  that have one of the following two forms:

- (i) if the cycle begins at  $v_1$ , the cycle consists of a  $j$ -cycle that passes through  $v_1$  once followed by a  $(k - j)$ -cycle (with no constraints on the number of times it passes through  $v_1$ ), or
- (ii) if the cycle does not begin at  $v_1$ , it begins with a  $(k - j)$ -cycle until it reaches  $v_1$ , completes a  $j$ -cycle that begins and ends at  $v_1$  (without passing through  $v_1$  any other time), and then completes the rest of the original  $(k - j)$ -cycle.

If we take the sum of all such terms  $-\text{trace}(c_j A^{k-j})$  for  $1 \leq j \leq k - 1$ , this will count all  $k$ -cycles that fall under the condition (c) given above.

It thus follows that the term

$$\begin{aligned} \text{trace}(A^k + c_1 A^{k-1} + \dots + c_{k-1} A) &= \\ \text{trace}(A^k) - (-\text{trace}(c_1 A^{k-1}) + \dots & \\ - (-\text{trace}(c_{k-1} A)) & \quad (6) \end{aligned}$$

counts the number of  $k$ -cycles that fall under conditions (a) and (b) above, which as noted above, equals  $k$  times the number of terminal nodes of  $T$  of depth  $k$ . Lastly, dividing the expression given above by (6) by  $k$  and using Leverrier's Method, we conclude that  $-c_k$  is equal to the number of terminal nodes of  $T$  of depth  $k$ . This establishes the inductive step, and so the result holds. □

**Theorem 5.3.** *Let  $T$  be a finite tree that induces a length function  $\ell_T$ , and let  $\rho$  be the Perron eigenvalue of the adjacency matrix of the directed graph associated to  $T$ . Then  $\log(\rho)$  is the channel capacity of  $\ell_T$ .*

*Proof.* Let  $A$  be the adjacency matrix of the directed graph associated to  $T$ , and let  $p_A(z) = z^n + c_1 z^{n-1} + \dots + c_{n-1} z + c_n$  be the characteristic polynomial of  $A$ . If  $\rho \geq 1$  is the Perron eigenvalue of  $p_A$ , we have

$$\rho^n + c_1 \rho^{n-1} + \dots + c_{n-1} \rho + c_n = 0.$$

Dividing through by  $\rho^n$  and rearranging yields

$$-c_1 \left(\frac{1}{\rho}\right) - \dots - c_{n-1} \left(\frac{1}{\rho}\right)^{n-1} - c_n \left(\frac{1}{\rho}\right)^n = 1.$$

Setting  $\alpha = \log(\rho)$ , so that  $2^{-\alpha} = \frac{1}{\rho}$ , by substitution we have

$$-c_1 2^{-\alpha} - \dots - c_{n-1} 2^{-\alpha(n-1)} - c_n 2^{-\alpha n} = 1. \quad (7)$$

By Theorem 5.2, for  $i = 1, \dots, n$ ,  $-c_i$  is equal to the number of terminal nodes in  $T$  of depth  $i$ , so we can rewrite Equation (7) as

$$\sum_{\tau \in S} 2^{-\alpha|\tau|} = 1, \quad (8)$$

where  $S$  is the set of terminal nodes of  $T$ . As the unique positive solution of Equation (8) is the channel capacity of  $\ell$ , the conclusion follows.  $\square$

**5.3. Restricting the Parry measure on a sofic shift.** In their work on path sets, Abram and Lagarias showed that path sets are a generalization of sofic shifts. A *sofic shift* is subset of  $\Sigma_n^\infty$  for some  $n \in \mathbb{N}$  that consists of all sequences of symbols obtained from an one-sided infinite walks though a finite directed graph with edge labels, where, unlike the case with path sets, the walks can start from any vertex of the graph. Abram and Lagarias further proved the equality of several notions of topological entropy relating path sets and sofic shifts. We review the definitions. First, for a path set  $P$ , let  $N_n(P)$  denote the number of distinct blocks of length  $n$  occurring in any member of  $P$ . Similarly,  $N_n^I(P)$  denotes the number of distinct initial blocks of length  $n$  from symbols in  $P$ . Then we have the following definitions.

**Definition 5.4.**

- (1) The *path topological entropy* of a path set  $P$  is defined to be

$$H_p(P) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log N_n(P).$$

- (2) The *topological entropy* of a path set  $P$  is

$$H_{top}(P) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log N_n^I(P).$$

Finally, given a path set  $P \subseteq \Sigma_n^\infty$ , the *one-sided shift closure*  $\overline{P}$  of  $P$  is  $\bigcup_{i \in \mathbb{N}} S^i(P)$ , where  $S$  is the shift operator on  $\Sigma_n^\infty$ . Note that for a path set  $P$ , its shift closure  $\overline{P}$  is a sofic shift, for which the notion of topological entropy  $H_{top}(\overline{P})$  is well-defined. Then we have:

**Theorem 5.5** (Abram, Lagarias [AL14]). *Let  $P \subseteq \Sigma_n^\infty$  be a path shift. Then  $H_p(P) = H_{top}(P) = H_{top}(\overline{P})$ .*

We apply this result in our own context as follows. Given a finite tree  $T$  and its associated pointed directed graph  $\mathcal{G} = (G, v_1)$ , we can assume by Theorem 5.1(i) that  $G$  is right-solving and irreducible. As noted above, the one-sided shift closure of the set of sequences obtained from one-sided infinite walks through  $G$  is a sofic shift, which we will write as  $X_G$ . In this context,  $G$ , equipped with its edge labels, is referred to as a *presentation* of  $X_G$ . Moreover, since the cycles through  $G$  that begin and end at the distinguished vertex  $v_1 \in G$  correspond to the terminal nodes of  $T$ , it is not difficult to establish that  $G$  is *minimal*, in the sense that there is no presentation  $G'$  of  $X_G$  that has strictly fewer vertices than  $G$ . Finally, as is known in

the symbolic dynamics literature, a sofic shift given by an irreducible, right-resolving, minimal presentation has a unique measure of maximal entropy, which is known as the *Parry measure*.

We should note that the entropy of a measure on a graph is slightly more general than the entropy of a measure on a sequence space such as  $\Sigma_n^\infty$ . For a measure  $\mu$  on a graph, we first specify initial probabilities for the initial vertex of each edge, written  $i(e)$  for an edge  $e$ , and then we specify conditional probabilities  $\mu(e | i(e))$  for each edge of the graph. Thus, the entropy of  $\mu$  is

$$h(\mu) = - \sum_{e \in E(G)} \mu(i(e)) \log(\mu(e | i(e)) \mu(e | i(e))).$$

Finally, the measure  $\mu$  of maximal entropy on a sofic shift  $X$  is one that attains the value  $h(\mu) = H_{top}(X)$ .

For the case of  $X_G$  defined above, let us look more closely at the derivation of the Parry measure on  $X_G$ , following the presentation found in see [LM21, Section 13.3]. Assuming that  $T$  contains exactly  $n$  non-terminal nodes, so that  $G$  consists of vertices  $v_1, \dots, v_n$ , let  $A$  be the  $n \times n$  adjacency matrix of  $G$  as in Section 5.2 above. Let  $\rho$  be the Perron eigenvalue of  $A$ , and let  $\vec{r} = \langle r_1, \dots, r_n \rangle$  be the right eigenvector corresponding to  $\rho$ . Then for each directed edge  $e$  from vertex  $v_i$  to vertex  $v_j$ , in a walk through  $G$ , given that we are at the vertex  $v_i$ , the probability of transitioning along  $e$  is equal to  $\frac{r_j}{r_i \rho}$ .

From this we can derive the following (which can be derived from results in [LM21, Section 13.3]):

*Fact:* The probability of traversing an  $n$ -cycle in  $G$  given that it starts at  $v_1$  is equal to  $\rho^{-n}$ .

Indeed, given an  $n$ -cycle that begins at  $v_1$ , which passes through the vertices  $v_1, v_{i_1}, v_{i_2}, \dots, v_{i_{n-1}}, v_1$ , the probability of transitioning along this  $n$ -cycle is

$$\left( \frac{r_{i_1}}{r_1 \rho} \right) \left( \frac{r_{i_2}}{r_{i_1} \rho} \right) \dots \left( \frac{r_0}{r_{i_{n-1}} \rho} \right) = \frac{1}{\rho^n}.$$

Using this, we can establish the following, which relates the material from the present section to the material from Section 4:

**Theorem 5.6.** *Let  $T$  be a finite tree with associated pointed directed graph  $\mathcal{G} = (G, v_1)$  and length function  $\ell$ . Then the restriction of the Parry measure on the sofic shift  $X_G$  with presentation  $G$  to the path set determined by  $\mathcal{G}$  is equal to the pushforward of  $\mu_\ell$  under the function  $\Psi_T : \Sigma_k^\infty \rightarrow F_T$ , which maps coding sequences in  $\Sigma_k^\infty$  to elements in  $F_T$ .*

*Proof.* Let  $\rho$  be the Perron eigenvalue of the adjacency matrix given by  $G$ . By the above fact, for the shift closure of the set of one-way infinite walks through  $G$ , the value that the Parry measure assigns to each cylinder set given by an  $n$ -cycle in  $G$  that begins at  $v_1$ , namely the value  $\rho^{-n}$ . Similarly, if we restrict the Parry measure to the path set given by  $\mathcal{G}$ , which by Theorem 5.1 is precisely the fractal  $F_T$ , we get a measure  $\xi$  on  $F_T$  that assigns to each cylinder set defined by a  $n$ -cycle in  $G$  that begins at  $v_1$  the value  $\rho^{-n}$ . Moreover, as we saw in the proof of Theorem 5.1, each  $n$ -cycle in  $G$  that begins at  $v_1$  has the same edge labels as some terminal node in  $T$  of length  $n$ . Thus,  $\xi$  assigns to each cylinder set defined by a terminal node of  $T$  the value  $\rho^{-n}$ .

Next, by Theorem 5.3,  $\log(\rho)$  is the channel capacity of  $\ell$ . As shown at the end of Subsection 4.1,  $\mu_\ell(i) = r^{\ell(i)}$ , where  $r = 1/\rho$ . Then the pushforward of  $\mu_\ell$  under the function  $\Psi_T$  is a measure  $\nu$  on  $F_T$  that is completely determined by its behavior on cylinders given by the terminal nodes  $\tau_0, \dots, \tau_{k-1}$  in  $T$ , where  $\nu(\tau_i) = \mu_\ell(i)$ , so that  $\nu(\tau_i) = \rho^{-\ell(i)}$ . In particular, for a terminal  $\tau_i$  node of length  $n$ ,  $\ell(i) = |\tau_i| = n$ , so that  $\nu(\tau_i) = \rho^{-n}$ . By what we derived in the previous paragraph,  $\nu$  and  $\xi$  assign the same values to the same cylinders, thereby establishing the theorem.  $\square$

## 6. ACKNOWLEDGEMENTS

I am thankful for the feedback and suggestions made by two anonymous referees that have improved the presentation of this article.

## REFERENCES

- [Abr94] Julia Abrahams. Varn codes and generalized Fibonacci trees. In *National Conference Publication-Institution Of Engineers Australia NCP*, pages 1171–1173, 1994.
- [Abr97] Julia Abrahams. Code and parse trees for lossless source encoding. *Proceedings. Compression and Complexity of SEQUENCES 1997 (Cat. No. 97TB100171)*, pages 145–171, 1997.
- [AHLM04] Krishna B Athreya, John M Hitchcock, Jack H Lutz, and Elvira Mayordomo. Effective strong dimension in algorithmic information and computational complexity. In *STACS 2004*, pages 632–643. Springer, 2004.
- [AL14] William C Abram and Jeffrey C Lagarias. Path sets in one-sided symbolic dynamics. *Advances in Applied Mathematics*, 56:109–134, 2014.
- [BBM<sup>+</sup>10] Marie-Pierre Béal, Jean Berstel, Brian H Marcus, Dominique Perrin, Christophe Reutenauer, and Paul H Siegel. Variable-length codes and finite automata. In *Selected topics in information and coding theory*, pages 505–584. World Scientific, 2010.
- [Bru82] AA Brudno. Entropy and the complexity of the trajectories of a dynamic system. *Trudy Moskovskogo Matematicheskogo Obshchestva*, 44:124–149, 1982.
- [Cha86] Derek K Chang. On fibonacci k-ary trees. *Fibonacci Quarterly*, 24(3):258–262, 1986.
- [CK11] Imre Csiszár and János Körner. *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.
- [DH10] Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic randomness and complexity*. Springer, 2010.
- [Fal04] Kenneth Falconer. *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons, 2004.
- [FP20] Johanna N.Y. Franklin and Christopher P. Porter. Key developments in algorithmic randomness. In Johanna N.Y. Franklin and Christopher P. Porter, editors, *Algorithmic Randomness: Progress and Prospects*, volume 50 of *Lecture Notes in Logic*. Cambridge University Press, 2020.
- [Gan59] F. R. Gantmacher. *The theory of matrices. Vols. 1, 2*. Translated by K. A. Hirsch. Chelsea Publishing Co., New York, 1959.
- [GKY02] Mordecai J Golin, Claire Kenyon, and Neal E Young. Huffman coding with unequal letter costs. In *Proceedings of the thirty-fourth annual ACM symposium on theory of computing*, pages 785–791, 2002.
- [GLMM14] Xiaoyang Gu, Jack H Lutz, Elvira Mayordomo, and Philippe Moser. Dimension spectra of random subfractals of self-similar fractals. *Annals of Pure and Applied Logic*, 165(11):1707–1726, 2014.
- [Hei79] Roland Heim. On the algorithmic foundation of information theory. *IEEE Transactions on Information Theory*, 25(5):557–566, 1979.
- [Hoy12] Mathieu Hoyrup. The dimension of ergodic random sequences. In *STACS’12 (29th Symposium on Theoretical Aspects of Computer Science)*, volume 14, pages 567–576. LIPIcs, 2012.
- [Kol65] Andrei Nikolaevich Kolmogorov. Three approaches to the definition of the concept “quantity of information”. *Problemy peredachi informatsii*, 1(1):3–11, 1965.
- [LL18] Jack H Lutz and Neil Lutz. Algorithmic information, plane keakeya sets, and conditional dimension. *ACM Transactions on Computation Theory (TOCT)*, 10(2):1–22, 2018.
- [LM08] Jack H Lutz and Elvira Mayordomo. Dimensions of points in self-similar fractals. *SIAM Journal on Computing*, 38(3):1080–1112, 2008.
- [LM21] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, 2021.
- [LS17] Neil Lutz and Donald M Stull. Dimension spectra of lines. In *Unveiling Dynamics and Complexity: 13th Conference on Computability in Europe, CiE 2017, Turku, Finland, June 12-16, 2017, Proceedings*, pages 304–314. Springer, 2017.
- [LS18] Neil Lutz and Donald M Stull. Projection theorems using effective dimension. In *43rd International Symposium on Mathematical Foundations of Computer Science*, 2018.
- [LS20] Neil Lutz and Donald M Stull. Bounding the dimension of points on a line. *Information and Computation*, 275:104601, 2020.
- [Lut00] Jack H Lutz. Gales and the constructive dimension of individual sequences. In *International Colloquium on Automata, Languages, and Programming*, pages 902–913. Springer, 2000.
- [Lut03] Jack H. Lutz. The dimension of individual strings and sequences. *Information and Computation*, 187(1):49–79, 2003.
- [Lut21] Neil Lutz. Fractal intersections and products via algorithmic dimension. *ACM Transactions on Computation Theory (TOCT)*, 13(3):1–15, 2021.

- [May02] Elvira Mayordomo. A Kolmogorov complexity characterization of constructive hausdorff dimension. *Information Processing Letters*, 84(1):1–3, 2002.
- [Nie09] André Nies. *Computability and randomness*, volume 51 of *Oxford Logic Guides*. Oxford University Press, 2009.
- [Sim15] Stephen G Simpson. Symbolic dynamics: entropy= dimension= complexity. *Theory of Computing Systems*, 56:527–543, 2015.
- [SUV17] Alexander Shen, Vladimir A. Uspensky, and Nikolay Vereshchagin. *Kolmogorov complexity and algorithmic randomness*, volume 220. American Mathematical Soc., 2017.
- [Var71] Ben Varn. Optimal variable length codes (arbitrary symbol cost and equal code word probability). *Information and Control*, 19(4):289–301, 1971.
- [Whi93] Homer S White. Algorithmic complexity of points in dynamical systems. *Ergodic Theory and Dynamical Systems*, 13(4):807–830, 1993.
- [ZL70] Alexander K. Zvonkin and Leonid A. Levin. The complexity of finite objects and the basing of the concepts of information and randomness on the theory of algorithms. *Uspehi Mat. Nauk*, 25(6(156)):85–127, 1970.