Valuations and Unambiguity of Languages, with Applications to Fractal Geometry

Henning Fernau*
Lehrstuhl Informatik für Ingenieure und Naturwissenschaftler
Universität Karlsruhe (TH)
D-76128 Karlsruhe

Ludwig Staiger†
Lehrstuhl für Informatik II
RWTH Aachen
Ahornstraße 55
D-52056 Aachen

Abstract

Valuations — morphisms from $(\Sigma^*, \cdot, e)$ to $((0, \infty), \cdot, 1)$ — are a simple generalization of Bernoulli morphisms (distributions, measures) as introduced in [11, 5]. This paper shows that valuations are not only useful within the theory of codes, but also when dealing with ambiguity, especially in regular expressions and contextfree grammars, or for defining outer measures on the space of ω-words which are of some importance for the theory of fractals. These connections yield new formulae to determine the Hausdorff dimension of fractal sets (especially in Euclidean spaces) defined via regular expressions and contextfree grammars. Furthermore, we generalize the classical notion of the entropy of a formal language.

This paper is an enhanced version of the one presented at ICALP'94 [17].

*email: `fernau@ira.uka.de`
†email: `staiger@zeus.informatik.rwth-aachen.de`

Contents

1 Introduction

Unambiguity is an old theme in formal language theory. In many applications, like compilers, text processors, etc. one is interested in describing words in a unique manner. In this paper, we treat three kinds of unambiguities:

- unambiguous language operations
- unambiguous regular expressions
- unambiguous contextfree grammars

We will characterize these unambiguities with the help of so-called valuations. It is also very interesting that our results permit the calculation of the valuation of a language given by a specific finite description from the finite description itself, circumventing the struggle with the (in general) infinite languages.

On the other hand, fractal geometry is now a budding branch of mathematics with a variety of possible applications [23, 12]. One of the main problems encountered in that field is the determination of the Hausdorff dimension of fractals. We show that, for certain types of fractals described by formal languages, this calculation may be simplified considerably by using the results on unambiguous regular expressions and unambiguous contextfree grammars derived in the second section. Surprisingly, we can give conditions in terms of formal language theory that alleviate arguments in topology and measure theory (the mathematical backbones of fractal geometry) considerably. In such a way, results from formal language theory (and algebra in a broader sense) contribute to the theory of fractals.

The bridges between formal language theory and fractal geometry we are going to build are:

- Language specifying iterated function systems, such describing fractals.
- Valuation dimension; this entity corresponds to the similarity dimension encountered with iterated function systems.
- Entropy with respect to a valuation; this entity corresponds to the Besicovitch-Taylor index encountered with iterated function systems [41].
- Special metrizations (induced by valuations) of spaces of ω-words; Hausdorff measure and Hausdorff dimension within these spaces are directly related to these entities within Euclidean spaces.

There have been other approaches using formal languages in order to describe fractals, or, more general, pictures. Closely related to our work are [44, 29]. Probably the most prominent example is L systems [30]. Interestingly, there are close connections to our language-based approach (see our comments below and the section on controlled iterated function systems in [30]). Additionally, we also find hypergraph-based ideas like collage grammars [19] and cellular automata [45, 40, 42]. One could also mention map 0L systems [30] and chain code picture languages [39].

Further material on the topic is contained in other works of the authors [14, 13, 16, 15, 26, 34, 35, 36, 37].

Conventions: $\subseteq$ and $\subset$ denote inclusion and strict inclusion, respectively. $\mathbb{N}$ is the set of nonnegative integers. The cardinality of a set S is denoted by card S.

$\Sigma_n = \{1, \ldots, n\} \subset \mathbb{N}$ denotes our standard alphabet. Σ (without subscript) denotes some at most countable alphabet. A language L is a subset of the word monoid Σ^* generated by the alphabet Σ, where e is the neutral element of the monoid, called empty word. Mostly, the monoid operation called catenation is just denoted by juxtaposition, sometimes made explicit using $\cdot$ between the words. The monoid generated by the language $L \subseteq \Sigma^*$ is denoted by L^*, and the semigroup generated by L is denoted by L^+.

We also consider ω-languages F over the alphabet Σ, i. e. sets of one-sided infinite words, $F \subseteq \Sigma^\omega$ for short. In general, if $L \subseteq \Sigma^*$, $L^\omega = \{v_0 \cdot v_1 \cdot v_2 \cdots : \forall i \in \mathbb{N}(v_i \in L \setminus \{e\})\}$.

Further notions and denotations are introduced throughout the text body.

2 Valuations and Unambiguity

We call a monoid morphism β mapping from $(\Sigma_n^*, \cdot, e)$ to $((0, \infty), \cdot, 1)$ a *valuation.* Any valuation can be extended to languages $L \subseteq \Sigma_n^*$ defining $\beta(L) = \sum_{w \in L} \beta(w)$.

As an example, consider the valuation β_n defined by $\beta_n(a) = 1/n$ for every $a \in \Sigma_n$.

Proofs of the results mentioned in this section can be found in [14, 13, 16].

2.1 Simple Properties of Valuations

Almost by definition, we find that $(\Sigma_n^*, 2^{\Sigma_n^*}, \beta)$ is a measure space. Other properties exploit that β is a morphism. We list some of them below. Let K, L be languages over Σ_n and $\{L_i\}_{i \in \Sigma}$ be an at most countable family of languages over Σ_n.

Determinacy	$\emptyset \neq K \iff 0 < \beta(K)$.
Monotonicity	$K \subseteq L \Rightarrow \beta(K) \leq \beta(L) \leq \beta(\Sigma_n^*)$.
(σ-)Additivity	Let $\{L_i\}_{i \in \Sigma}$ be a family of pairwise disjoint languages L_i over Σ_n. Then $\beta(\bigcup_{i \in \Sigma} L_i) = \sum_{i \in \Sigma} \beta(L_i)$.
(σ-)Subadditivity	In general, $\beta(\bigcup_{i \in \Sigma} L_i) \leq \sum_{i \in \Sigma} \beta(L_i)$.
Continuity	Let $(L_i)_{i \in \mathbb{N}}$ be an increasing chain. Then $\beta(\bigcup_{i \in \mathbb{N}} L_i) = \lim_{i \to \infty} \beta(L_i)$.
Subtractivity	$(\beta(K) < \infty \wedge K \subseteq L) \Rightarrow \beta(L \setminus K) = \beta(L) - \beta(K)$.
Multiplication law	$\beta(KL) \leq \beta(K)\beta(L)$.
Power law	$(\forall m \in \mathbb{N})(\beta(L^m) \leq (\beta(L))^m)$.
Star law	$\beta(L^*) = \beta(\bigcup_{m=0}^{\infty} L^m) \leq \sum_{m=0}^{\infty} \beta(L^m)$; if $\beta(L) < 1$, then additionally $\beta(L^*) \leq \frac{1}{1-\beta(L)} < \infty$.

Observe that determinacy, subadditivity and multiplication law are quite similar to the properties required for real-valued valuations of fields [43].

2.2 Unambiguous Operations

Recall that the product KL is *unambiguous* iff any word in KL is unambiguously decomposable, i.e. for any $w_1, w_1' \in K$ and $w_2, w_2' \in L$ with $w = w_1w_2 = w_1'w_2'$, we have $w_1 = w_1'$ and $w_2 = w_2'$. The union $K \cup L$ is termed *unambiguous* iff K and L are disjoint. The star operation C^*, $C \subseteq \Sigma_n^+$ is called *unambiguous* iff C is a code, i.e. for all $(v_1, \ldots, v_k) \in C^k$ and for all $(u_1, \ldots, u_m) \in C^m$, the equality $u_1 \cdots u_m = v_1 \cdots v_k$ implies $m = k$ and, for all $i \in \Sigma_k$, $u_i = v_i$.

The next lemma is easily proved.

Lemma 1 *Let $K, L, C \subseteq \Sigma_n^*$ and β be a valuation with $\beta(K), \beta(L) < \infty$. Then:*

- $\beta(KL) = \beta(K)\beta(L)$ *iff the product KL is unambiguous.*
- $\beta(K \cup L) = \beta(K) + \beta(L)$ *iff the union $K \cup L$ is unambiguous.*
- *If $\beta(C) < 1$, then $\beta(C^*) = \frac{1}{1-\beta(C)}$ iff C is a code.*
- *If C is a code, then $\beta(C^*) = \sum_{i=0}^{\infty}(\beta(C))^i$.*

Observe that, presupposing the unambiguity of the involved operators (products, unions, stars), we may calculate the valuation of a given language by calculating the valuations of the parts defining the language, and by interpreting concatenation as multiplication, union as addition, and star as an infinite sum. This observation will be exploited in the following.

2.3 Unambiguous Regular Expressions

In this part, we give conditions under which it is possible to calculate the valuation of a language given by a regular expression from the expression itself. This in turn enables us to characterize a certain class of regular expressions known as (strong) unambiguous expressions.

We give a formal definition of regular and unambiguous regular expressions, since, following Brüggemann-Klein [7], we leave out the empty set in the usual definition of regular expressions.

Let $\mathcal{R}_n \subseteq (\Sigma_n \cup \{(,),\cup,*\})^*$ denote the class of *regular expressions* over Σ_n. $\mathcal{R}_n$ is the smallest language over $\Sigma_n \cup \{(,),\cup,*\}$ satisfying:

- $\Sigma_n^* \subset \mathcal{R}_n$.
- If $R_1, R_2 \in \mathcal{R}_n$, then (R_1R_2), $(R_1 \cup R_2)$, and R_1* lie in $\mathcal{R}_n$.

As usual, the language $[R] \subseteq \Sigma_n^*$ described by some regular expression R is defined recursively:

- For $w \in \Sigma_n^*$, let $[w] = \{w\}$.
- If $R_1, R_2 \in \mathcal{R}_n$, then $[(R_1R_2)] = [R_1][R_2]$, $[(R_1 \cup R_2)] = [R_1] \cup [R_2]$, and $[R_1*] = [R_1]^*$.

Let $\mathcal{UR}_n \subseteq (\Sigma_n \cup \{(,),\cup,*\})^*$ denote the class of *unambiguous regular expressions* over Σ_n. $\mathcal{UR}_n$ is the smallest language over $\Sigma_n \cup \{(,),\cup,*\}$ satisfying:

- $\Sigma_n^* \subset \mathcal{UR}_n$.
- If R_1, $R_2 \in \mathcal{UR}_n$, then $(R_1R_2) \in \mathcal{UR}_n$, provided the corresponding language operation $[R_1][R_2]$ is unambiguous.
- If R_1, $R_2 \in \mathcal{UR}_n$, then $(R_1 \cup R_2) \in \mathcal{UR}_n$, provided the corresponding language operation $[R_1] \cup [R_2]$ is unambiguous.
- If $R_1 \in \mathcal{UR}_n$, then $R_1* \in \mathcal{UR}_n$, provided the corresponding language operation $[R_1]^*$ is unambiguous.

Obviously, $\mathcal{UR}_n \subseteq \mathcal{R}_n$, and this inclusion is strict, as the example $e* \in \mathcal{R}_n \backslash \mathcal{UR}_n$ shows.
We develop a characterization of unambiguous regular expressions in terms of valuations in the following. To this end, we inductively define the valuation of a regular expression. Observe that, in this definition, we formally interpret language operations as numerical operations.
Let $\beta : \Sigma_n^* \to (0,\infty)$ be some valuation. The *valuation $\beta_\mathcal{R}$ of a regular expression $R \in \mathcal{R}_n$* is a mapping $\beta_\mathcal{R} : \mathcal{R}_n \to [0,\infty]$ defined inductively as follows.

- If $w \in \Sigma_n^* \subset \mathcal{R}_n$, then $\beta_\mathcal{R}(w) = \beta(w)$.
- If R_1, $R_2 \in \mathcal{R}_n$, then $\beta_\mathcal{R}((R_1R_2)) = \beta_\mathcal{R}(R_1)\beta_\mathcal{R}(R_2)$, $\beta_\mathcal{R}((R_1 \cup R_2)) = \beta_\mathcal{R}(R_1) + \beta_\mathcal{R}(R_2)$ and $\beta_\mathcal{R}(R_1*) = \sum_{i\in\mathbb{N}}(\beta_\mathcal{R}(R_1))^i$.

Note that we consider $[0,\infty]$ as a semiring extending the semiring $([0,\infty),+,\cdot,1,0)$ defining especially $0\cdot\infty = \infty\cdot 0 = 0$.
By the multiplication, subadditivity and star laws, we have immediately $\beta([R]) \leq \beta_\mathcal{R}(R)$ for any regular expression $R \in \mathcal{R}_n$ and any valuation $\beta : \Sigma_n^* \to (0,\infty)$. By induction, Lemma 1 leads to the following corollary.

Corollary 2 *Let $\beta : \Sigma_n^* \to (0,\infty)$ be some valuation and $R \in \mathcal{UR}_n$. Then, $\beta_\mathcal{R}(R) = \beta([R])$.*

Considering the example $R = (1*2*)* \in \mathcal{R}_2$, $\beta(1) = \beta(2) = 1/3$, we find

$$\beta_\mathcal{R}(R) = \sum_{i=0}^{\infty}(\sum_{j=0}^{\infty}(\frac{1}{3})^j + \sum_{j=0}^{\infty}(\frac{1}{3})^j)^i = \infty;$$

on the other hand $[R] = \Sigma_2^*$, yielding $\beta([R]) = \sum_{i=0}^{\infty}(\frac{2}{3})^i < \infty$. Hence, $R \notin \mathcal{UR}_2$.

Theorem 3 *Let $\beta : \Sigma_n^* \to (0,\infty)$ be some valuation and $R \in \mathcal{R}_n$ such that $\beta_\mathcal{R}(R) < \infty$. Then, R is unambiguous iff $\beta_\mathcal{R}(R) = \beta([R])$.*

On the other hand, if we know two regular expressions $R_1, R_2 \in \mathcal{R}_n$ describing the same language, then R_1 is ambiguous provided we have a valuation β such that $\beta_\mathcal{R}(R_1) > \beta_\mathcal{R}(R_2)$.

2.4 Unambiguous Contextfree Grammars

In this part, we examine contextfree grammars. The question is again the next: When is it possible to compute the valuation of a language directly numerically from its finite description, here a contextfree grammar? From our preceding thoughts, the basic idea of transferring a contextfree grammar into a numeric system should be clear: Think of a contextfree grammar as a system of equations involving catenation and union, interpret the language variables (nonterminals) as numerical variables, interpret the terminals by the given valuation, and finally interpret the catenation as multiplication and the union as addition.

In the following, we assume that the reader is familiar with some basics of formal power series, see e.g. [21]: If R is a semiring (with zero 0 and one 1) and Σ_n^* is a monoid, the set of formal power series, i.e. mappings $\Sigma_n^* \to R$, is denoted by $R\ll\Sigma_n^*\gg$. $s \in R\ll\Sigma_n^*\gg$ is written formally as a series $\sum_{w\in\Sigma_n^*} <s, w> w$, where $<s, w> = s(w)$. The support $\mathbf{supp}(s)$ of a series s equals $\{w : <s, w> \neq 0\}$. The set of polynomials, i.e. formal power series having finite support, is written $R<\Sigma_n^*>$. The characteristic series $\underline{L}$ of a language $L \subseteq \Sigma_n^*$ is defined by $\underline{L} = \sum_{w\in L} w$. We consider the semirings $\mathbb{N}$, $[0, \infty)$, and $[0, \infty]$ together with the usual addition and multiplication as semiring operations. Note that any valuation itself can be viewed as a formal power series in $[0, \infty]$. Another connection of valuations and formal power series is the next.

Lemma 4 *Any valuation $\beta : \Sigma_n^* \to (0, \infty)$ induces a semiring morphism*

$$\psi_\beta : [0,\infty]\ll\Sigma_n^*\gg \to [0,\infty] \ \textit{by} \ \sum_{w\in\Sigma_n^*} <s, w> w \mapsto \sum_{w\in\Sigma_n^*} <s, w> \beta(w).$$

To the set of productions P of a contextfree grammar $G = (X, \Sigma_n, P, x_1)$, $X = \{x_1, \ldots, x_m\}$, there corresponds a system of equations of the form

$$x_i = p_i, \quad 1 \leq i \leq m, \quad p_i \in \mathbb{N}<(\Sigma_n \cup X)^*> \tag{1}$$

with $(\forall w \in (\Sigma_n \cup X)^*)(<p_i, w> \in \{0, 1\})$ such that $x_i \to w \in P$ iff $<p_i, w> = 1$. It is well-known that G is unambiguous iff the solution of the $\mathbb{N}$-algebraic system (1) is given by $(\underline{L_1}, \ldots, \underline{L_m})$, where $\underline{L_i}$ denotes the characteristic series of the language L_i, which in turn is generated by the original grammar G, taking as start symbol x_i.

ψ_β can be extended to a semiring morphism $\psi_\beta : ([0,\infty]\ll\Sigma_n^*\gg)\ll X^*\gg \to [0,\infty]\ll X^*\gg$. This provides us with a mathematically sound description of the numerical system corresponding to Eq. (1), namely $x_i = \psi_\beta(p_i)$.

Elaborating the above observations, we obtain the next lemma.[1]

Lemma 5 *Let $\beta : \Sigma_n^* \to (0, \infty)$ be some valuation. Let $L_i \subseteq \Sigma_n^+$ be generated by an unambiguous contextfree grammar $G_i = (\{x_1, \ldots, x_m\}, \Sigma_n, P, x_i)$ in Chomsky normal form. Let $x_i = p_i$ be the corresponding $\mathbb{N}$-algebraic system. Then, $(\beta(L_1), \ldots, \beta(L_m))$ is a solution of the system of equations $x_i = \psi_\beta(p_i)$.*

[1] The Chomsky normal form restriction in the following assertions alleviates the proofs contained in [14] but is not necessary.

Is it possible to calculate $\beta(L_i)$ directly from the numerical system $x_i = \psi_\beta(p_i)$? An answer is given in the next theorem which is proved in [14].

Theorem 6 *Let $\beta : \Sigma_n^* \to (0,\infty)$ be some valuation. Let L be an e-free unambiguous contextfree language. Let $x_i = p_i$, $i = 1,\ldots,m$, be an $\mathbb{N}$-algebraic system of equations in Chomsky normal form with $<p_i, w> \in \{0,1\}$ such that the series $\underline{L}$ is the first component of the uniquely determined solution $s = (s_1,\ldots,s_m)$ of the equation system $x_i = p_i$. We assume furthermore that, without loss of generality, any variable $x_i \in X\backslash\{x_1\}$ is reachable from x_1 in some derivation.*

- *If $\beta(L) < \infty$ and if the corresponding system with valuation $x_i = \psi_\beta(p_i)$ has exactly one solution $b = (b_1,\ldots,b_m) \in [0,\infty)^m$, then $b_1 = \psi_\beta(s_1) = \beta(L)$.*
- *If the corresponding system with valuation $x_i = \psi_\beta(p_i)$ has no solution in $[0,\infty)^m$, then $\beta(L) = \infty$.*

On the other hand, it is possible to obtain a criterion for unambiguity of grammars using valuations.

Theorem 7 *Let $G = (X, \Sigma_n, P, x_1)$ be a contextfree grammar in Chomsky normal form such that from x_1 any nonterminal $x_i \in X\backslash\{x_1\}$ is reachable, inducing the $\mathbb{N}$-algebraic system $x_i = p_i$ with the solution $s = (s_1,\ldots,s_m)$. Let $\beta : \Sigma_n^* \to (0,\infty)$ be a valuation such that $\beta(\mathbf{supp}(s_1)) < \infty$. Assume that the corresponding system with valuation $x_i = \psi_\beta(p_i)$ has exactly one solution $b = (b_1,\ldots,b_m)$ with $b_i = \beta(\mathbf{supp}(s_i))$ in $(0,\infty)^m$. Then, G is unambiguous.*

In Theorem 6, the case when the system with valuation $x_i = \psi_\beta(p_i)$ has more than one solution is left open. We tackle this problem in the following.

In order to do this, we need some definitions and facts about complete partially orderered sets (cpo's), cf. [18]. A partial order $\sqsubseteq$ on a set D is a reflexive, anti-symmetric and transitive binary relation. A subset $M \subseteq D$ is called directed if, for every finite subset $u \subseteq M$, there is an upperbound $x \in M$ for u. $(D, \sqsubseteq)$ is complete (a cpo), if (1) every directed subset $M \subseteq D$ has a least upperbound $\bigsqcup M$ and (2) there is a least element $-_D$ in D. If $(D, \sqsubseteq)$ is a partially ordered set, $(D^m, \sqsubseteq)$ is so, too, defining $a \sqsubseteq b \iff \forall i \in \Sigma_m(a(i) \sqsubseteq b(i))$. If $(D, \sqsubseteq)$ is a cpo, the m^{th} power $(D^m, \sqsubseteq)$ is also complete.

In our case, we consider the following three cpo's and the derived m^{th} powers.

- $D_1 = ([0,\infty], \leq)$
- $D_2 = (2^{\Sigma_n^*}, \subseteq)$ with least element $\emptyset$.
- $D_3 = ((\mathbb{N} \cup \{\infty\})\ll\Sigma_n^*\gg, \leq)$ with $s \leq t$ iff, for all words $w \in \Sigma_n^*$, $<s, w> \leq <t, w>$.

Given cpo's D and E, a function $f : D \to E$ is monotone if $f(x) \sqsubseteq f(y)$ whenever $x \sqsubseteq y$. If f is monotone and $f(\bigsqcup M) = \bigsqcup f(M)$ for every directed M, then f is said to be cpo-continuous. For example, any valuation β, viewed as a map from $(2^{\Sigma_n^*}, \subseteq)$ to $([0,\infty], \leq)$, is cpo-continuous. The basic lemma we need is the next.

Lemma 8 *If D is a cpo and $f : D \to D$ is continuous, then there is a point $\mathit{fix}(f) \in D$ such that $\mathit{fix}(f) = f(\mathit{fix}(f))$ and $\mathit{fix}(f) \leq x$ for any $x \in D$ such that $x = f(x)$. In other words, $\mathit{fix}(f)$ is the least fixed point of f. Moreover, $\mathit{fix}(f) = \bigsqcup\{f^k(-_D) : k \in \mathbb{N}\}$.*

Observe that any contextfree grammar $G = (X, \Sigma_n, P, x_1)$, written as a system of equations $x_i = \tilde{p}_i$, may be viewed as a cpo-continuous mapping $\tilde{p} = (\tilde{p}_1, \ldots, \tilde{p}_m)$ on D_2^m, where $\tilde{p}_j$ is applied to $(L_1, \ldots, L_m)$ by substituting any occurrence of the variable x_i in $\tilde{p}_j$ by the corresponding language L_i. Fortunately, the usual definition of a language (or a tuple of languages) derived from a contextfree grammar coincides with the least fixed point semantic just described.

Similarly, the system of equations of formal power series obtained from a contextfree grammar has a least fixed point semantics which again coincides with the classical interpretation. Just note that our usual assumption that the corresponding grammar is e-free and in Chomsky normal form only serves to guarantee that the least solution does not have ∞ anywhere as a coefficient. Moreover, the mapping $D_2 \to D_3, L \mapsto \underline{L}$ and $\mathbf{supp} : D_3 \to D_2$ are cpo-continuous. Furthermore, if we consider the sequences $L_{(0)} = (L_{0,1}, \ldots, L_{0,m}) = (\emptyset, \ldots, \emptyset)$, $L_{(1)} = (L_{1,1}, \ldots, L_{1,m})$ with $L_{1,j} = \tilde{p}_j(L_{(0)})$, $L_{(2)}$, etc. on the language side, and $s_{(0)} = (s_{0,1}, \ldots, s_{0,m}) = (0, \ldots, 0)$, $s_{(1)} = (s_{1,1}, \ldots, s_{1,m})$ with $s_{1,j} = p_j(s_{(0)})$, $s_{(2)}$, etc. on the side of formal power series, i. e. considering the set of equations $x_i = p_i$, we observe $L_{(k)} = \mathbf{supp}(s_{(k)})$ for every k by induction, and hence by cpo-continuity, the language tuple generated by a contextfree grammar is just the support of the least solution of the corresponding equation of formal power series. The other 'direction' $\underline{L_{(k)}} = s_{(k)}$ is valid if and only if G is unambiguous.

Finally, the system of equations with valuation $x_i = \psi_\beta(p_i)$ obtained from a system of equations $x_i = p_i$ of formal power series has a least fixed point semantics. By induction, it is easy to see that its approximating sequence coincides with $(\psi_\beta(s_{(k)}))_k$, where $s_{(k)}$ is defined as above, yielding the appoximating sequence of the system of equations of formal power series. Since $\psi_\beta : D_3 \to D_1$ is cpo-continuous, $(\psi_\beta(s_1), \ldots, \psi_\beta(s_m))$ is the least solution of the D_1^m-system $x_i = \psi_\beta(p_i)$, where $(s_1, \ldots, s_m)$ is the least formal power series solution of the system $x_i = p_i$.

We summarize our observation in the following theorem.

Theorem 9 *Let $\beta : \Sigma_n^* \to (0, \infty)$ be some valuation. Let L be an e-free unambiguous contextfree language. Let $x_i = p_i$, $i = 1, \ldots, m$, be an $\mathbb{N}$-algebraic system of equations in Chomsky normal form with $<p_i, w> \in \{0, 1\}$ such that the series $\underline{L}$ is the first component of the uniquely determined solution $s = (s_1, \ldots, s_m)$ of the equation system $x_i = p_i$.*

If $(b_1, \ldots, b_m)$ is the least solution of the D_1^m-system $x_i = \psi_\beta(p_i)$, then $b_i = \psi_\beta(s_i)$, since $(s_1, \ldots, s_m)$ is the least solution of the D_3^m-system $x_i = p_i$. Especially, $b_1 = \beta(L)$.

3 β-Entropy of Languages

It was shown in [35], [36], [26] that the entropy of languages introduced by Chomsky and Miller (cf. [20]) is a useful tool for the calculation of the Hausdorff dimension of certain subsets of the Cantor space Σ_n^ω or of the Euclidean space $\mathbb{R}^d$. In this section we derive a generalization of this concept which, as we shall see in the subsequent sections, leads to similar calculation formulae for the Hausdorff dimension of subsets of $(\Sigma_n^\omega, \rho_\beta)$ and can be used to calculate the Hausdorff dimension of certain subsets of $\mathbb{R}^d$ thereby generalizing results of [1], [2], [25].

We call this generalization entropy of languages with respect to a valuation β (short: β-entropy of a language). The aim of this section is to show that the properties of the entropy of languages mentioned in [33] and [36, Section 2] hold as well for this generalization.

Let $\beta\colon \Sigma_n^* \to (0,\infty)$ be a valuation. We call $\beta^s(L) := \sum_{w\in L} \beta(w)^s$ the *s-dimensional valuation* of the language $L \subseteq \Sigma_n^*$. In particular, we allow valuations β having $\beta(w) \geq 1$ for some words.

For a fixed language $L \subseteq \Sigma_n^*$ we consider the s-dimensional valuation β^s as a function $\beta^s(L) : [0,\infty) \to [0,\infty]$. We summarize some properties of the function $\beta^s(L)$.

Property 10 *Let $L \subseteq \Sigma_n^*$ and $\beta\colon \Sigma_n^* \to (0,\infty)$ be a valuation, and let $\beta^s(L) < \infty$ for some $s \in [0,\infty)$. Then there is an $\alpha \in [0,\infty)$ such that $\beta^s(L) = \infty$ for $s < \alpha$, $\beta^s(L) < \infty$ for $s > \alpha$, and the function $\beta^s(L)$ is continuous on (α,∞) and satisfies $\lim_{s\downarrow\alpha} \beta^s(L) = \beta^\alpha(L)$.*[2]
If, moreover, $\beta(w) < 1$ for all $w \in L$, then the function $\beta^s(L)$ is strictly decreasing and $\lim_{s\to\infty} \beta^s(L) = 0$.

Before proceeding to the proof we would like to add two remarks.

Remark 1. The point α defined above is called a "change-over-point" of the function $\beta^s(L)$.

Remark 2. It is possible to construct valuations β and languages L such that $\beta(w) < 1$ for $w \in L$ and nevertheless $\beta^s(L) = \infty$ for all $s \in [0,\infty)$. In the sequel, however, we are not interested in such pathological cases. If $\beta(a) < 1$ for any $a \in \Sigma_n$, then there is a finite change-over-point α of the function $\beta^s(L)$ for any $L \subseteq \Sigma_n^*$.

Proof. If $\beta^s(L) < \infty$ for some $s \in [0,\infty)$, then the set $\{w : \beta(w) \geq 1 \wedge w \in L\}$ is finite.

Since our assertion is obvious for finite languages, we may split L into a disjoint union $L = L' \cup L''$ where L'' is finite and contains $\{w : \beta(w) \geq 1 \wedge w \in L\}$. Thus in virtue of $\beta^s(L) = \beta^s(L') + \beta^s(L'')$ it remains to verify the assertion for functions $\beta^s(L')$ where $\emptyset \neq L' \subseteq \{w : \beta(w) < 1 \wedge w \in L\}$.

First observe that, once $\beta^s(L')$ is finite, it is strictly decreasing and positive. Thus $\alpha := \inf\{s : \beta^s(L') < \infty\}$.

Now let $\theta \geq \alpha$ such that $\beta^\theta(L') < \infty$. As the infinite series $\beta^\theta(L') = \sum_{w\in L'} \beta(w)^\theta$ is the approximation of its finite sums $\beta^\theta(U_i)$ where $U_i := \{w : w \in L' \wedge |w| \leq i\}$ for every $\epsilon > 0$ there is an $i \in \mathbb{N}$ such that $\epsilon > \beta^\theta(L') - \beta^\theta(U_i) = \beta^\theta(L' \setminus U_i) \geq \beta^s(L') - \beta^s(U_i)$ for $s \geq \theta$. Consequently the sequence of continuous functions $\left(\beta^s(U_i)\right)_{i\in\mathbb{N}}$ uniformly converges to the function $\beta^s(L')$ on the interval $[\theta,\infty)$, whence $\beta^s(L')$ is continuous on $[\theta,\infty)$. Since $\beta^\theta(L') < \infty$ for all $\theta > \alpha$ this shows that $\beta^\theta(L')$ is continuous on (α,∞). If, moreover, $\beta^\alpha(L') < \infty$, then $\beta^s(L')$ is continuous on $[\alpha,\infty)$.

Utilizing the same argument and the property that $\lim_{s\to\infty} \beta^s(U_i) = 0$ we can show that $\beta^s(L')$ tends to zero as s approaches infinity.

Finally, it remains to show that $\lim_{s\downarrow\alpha} \beta^s(L) = \beta^\alpha(L) = \infty$ if $\beta^\alpha(L') = \infty$.

---

[2]permitting the value ∞ for $\beta^\alpha(L)$

As it was mentioned above, for finite U the function $\beta^s(U)$ is continuous. Hence, $\beta^s(U_i) = \lim_{\varepsilon\to 0}\beta^{s+\varepsilon}(U_i)$. Taking suprema on both sides yields $\beta^\alpha(L') \leq \lim_{\varepsilon\to 0}\beta^{\alpha+\varepsilon}(L')$, and the assertion follows.

Q.E.D.

The β-*entropy* of the language $L \subseteq \Sigma_n^*$, H_L^β, is defined as the "change-over-point" of the function $\beta^s(L)$.

$$H_L^\beta := \inf\{s : s \geq 0 \wedge \beta^s(L) < \infty\}\ .^3 \tag{2}$$

In particular, $H_L^\beta < \infty$ iff $\exists s(s \in (0,\infty) \wedge \beta^s(L) < \infty)$.

As the usual entropy of languages, our β-entropy satisfies also the following simple identities.

$$H_{W\cup V}^\beta = H_{W\cdot V}^\beta = \max\{H_W^\beta, H_V^\beta\} \quad \text{if } W\cdot V \neq \emptyset, \text{ and} \tag{3}$$

$$H_L^\beta = 0 \quad \text{if } L \text{ is finite.} \tag{4}$$

3.1 The β-Entropy of Regular Languages

Now, we consider regular languages. Here we can characterize our languages having finite β-entropy. To this end we introduce the *state* of a subset $M \subseteq \Sigma_n^* \cup \Sigma_n^\omega$ derived from a word $w \in \Sigma_n^*$.

$$M/w := \{p : p \in \Sigma_n^* \cup \Sigma_n^\omega \wedge w\cdot p \in M\} \tag{5}$$

We call a set $M \subseteq \Sigma_n^* \cup \Sigma_n^\omega$ *finite-state* provided it has only a finite number of distinct states. It is well-known that a language $L \subseteq \Sigma_n^*$ is finite-state iff it is regular, whereas every regular ω-language [4] is finite-state but the converse does not hold (see e.g. [31]).

Furthermore, we say that a word $w \in \Sigma_n^*$ is a *prefix* of a string $p \in \Sigma_n^* \cup \Sigma_n^\omega$ provided $p = w\cdot p'$ for some $p' \in \Sigma_n^* \cup \Sigma_n^\omega$ and we abbreviate this fact by $w \sqsubseteq p$. For a subset $M \subseteq \Sigma_n^* \cup \Sigma_n^\omega$ its set of finite prefixes is $\mathbf{A}(M) := \{w : w \in \Sigma_n^* \wedge \exists p(w\cdot p \in M)\}$ and its set of subwords (infixes) is $\mathbf{T}(M) := \{v : v \in \Sigma_n^* \wedge \exists p \exists w(w \in \Sigma_n^* \wedge w\cdot v\cdot p \in M)\}$.

Property 11 *If $L \subseteq \Sigma_n^*$ is a regular language and $\beta : \Sigma_n^* \to (0,\infty)$ is a valuation, then the following conditions are equivalent.*

1. *There is an $s \geq 0$ such that $\beta^s(L) < \infty$.*
2. $\forall w, v(v \neq e \wedge L/w = L/w\cdot v \neq \emptyset \to \beta(v) < 1)$.
3. *There are an $\ell \in \mathbb{N}$ and a positive constant $c < 1$ such that for all $u \in \mathbf{T}(L)$ with $|u| \geq \ell$ it holds $\beta(u) \leq c^{|u|}$.*

[3] Here we follow the convention $\inf\emptyset = \infty$.

[4] Regular ω-languages are defined as finite unions of sets of the form $W\cdot V^\omega$ where W, V are regular languages.

Remark. Property 11(2) is just another formulation of the contracting cycles property of [25] and [2].

Proof. "1. $\to$ 2.:" If there is a word $v \in \Sigma_n^* \setminus \{e\}$ such that $\beta(v) \geq 1$ and there are words w, u with $wvu \in L$ and $L/w = L/w \cdot v$, then $L \supseteq wv^*u$ and, therefore, $\beta^s(L) \geq \beta^s(wv^*u) = \beta^s(wu) \cdot \sum_{i\in\mathbb{N}}(\beta^s(v))^i = \infty$.
"2. $\to$ 3.:" Let the set of nonempty states of L, $\{L/w : w \in \mathbf{A}(L)\}$, have k elements. Then it is well-known that $u \in \mathbf{T}(L)$ iff there is a word $u_0 \in \mathbf{A}(L)$ such that $|u_0| < k$ and $u_0 \cdot u \in \mathbf{A}(L)$. Moreover, for every word $u' \in L$ of length $|u'| \geq k$ there is an factorization $u' = w \cdot v \cdot \hat{w}$ satisfying $0 < |v| \leq k$ and $L/w = L/w \cdot v$.
Consequently, $\beta(u') = \beta(w \cdot \hat{w}) \cdot \beta(v)$.
Repeating this process of cutting out nonempty "cycles" v of length $\leq k$ we finally arrive at a family of words $w', v_1, \ldots, v_m$ such that $|w'| < k$, $0 < |v_i| \leq k$, $\emptyset \neq L/w_i = L/w_i \cdot v_i$ for some $w_i \in \mathbf{A}(L)$ and

$$\beta(u_0 \cdot u) = \beta(w') \cdot \beta(v_1) \cdots \beta(v_m) .$$

Consequently, $\beta(v_i) < 1$, more precisely, let

$$\hat{c} := \max\{ \sqrt[|v|]{\beta(v)} : 1 \leq |v| \leq k \wedge \exists w(L/w = L/w \cdot v \neq \emptyset)\} ,$$

then $\hat{c} < 1$ and $\beta(v_i) \leq \hat{c}^{|v_i|}$. This yields the inequality

$$\begin{aligned} \beta(u) &\leq \frac{\beta(w')}{\beta(u_0)} \cdot \hat{c}^{|v_1|+\cdots+|v_m|} \\ &\leq \frac{\max\{\beta(w) : |w| < k\}}{\min\{\beta(w) : |w| < k\}} \cdot \hat{c}^{|u|-k+1} \\ &\leq c_0 \cdot \hat{c}^{|u|} \text{ , for some } c_0 \geq 0 . \end{aligned}$$

Now, first choosing c such that $\hat{c} < c < 1$ and then ℓ large enough, the assertion follows immediately.
"3. $\to$ 1.:" If for $L \subseteq \Sigma_n^*$ and $\beta(w) \leq c^{|w|}$ for some positive $c < 1$ and all $w \in L$ with $|w| \geq \ell$, then the inequality $n \cdot c^s < 1$ holds for some $s \geq 0$, and this implies $\beta^s(L) \leq \sum_{|w|\leq\ell} \beta^s(w) + \sum_{i\geq\ell} n^i \cdot c^{s\cdot i} < \infty$.

Q.E.D.

¿From the direction "1. $\to$ 2." of the preceding proof we obtain immediately.

Corollary 12 *If L is a regular language and $\beta(w)^s < 1$ for all $w \in L \setminus \{e\}$, then $\beta^s(L) < \infty$.*

We obtain a relation between the entropies of L, $\mathbf{A}(L)$ and $\mathbf{T}(L)$ for a regular language L.

Property 13 *If L is regular, then $\beta^s(L) < \infty$ iff $\beta^s(\mathbf{A}(L)) < \infty$ iff $\beta^s(\mathbf{T}(L)) < \infty$.*

Proof. $L \subseteq \mathbf{A}(L) \subseteq \mathbf{T}(L)$ and $\beta^s(\mathbf{T}(L)) < \infty$ imply $\beta^s(L) < \infty$ and $\beta^s(\mathbf{A}(L)) < \infty$.
Conversely, let $\beta^s(L) < \infty$. Since L is regular, there is a $k \in \mathbb{N}$ such that

$$\forall v(v \in \mathbf{T}(L) \to \exists w, \hat{w}(w, \hat{w} \in L \wedge w \cdot v \sqsubseteq \hat{w} \wedge |w|, |\hat{w}| - |w \cdot v| \leq k)) .$$

Thus $\mathbf{A}(L) \subseteq \mathbf{T}(L) = L_{0,0} \cup \ldots \cup L_{k,k}$ where $L_{i,j} := \{v : \Sigma_n^i \cdot v \cdot \Sigma_n^j \cap L \neq \emptyset\}$, and the assertion follows from the easily verified inequality $\beta^s(L_{i,j}) \cdot \min\{\beta(a)^{(i+j)\cdot s} : a \in \Sigma_n\} \leq \beta^s(L)$.

Q.E.D.

Corollary 14 *If L is regular, then $H_L^\beta = H_{\mathbf{A}(L)}^\beta = H_{\mathbf{T}(L)}^\beta$.*

Next we give a method to compute the β-entropy for a nonempty regular language L. To this end let $\{L_1 = L, L_2, \ldots, L_k\}$ be its set of nonempty states. Define the β-weighted s-dimensional adjacency matrix of L, $\mathcal{A}_L^{\beta,s} = (a_{s;i,j})_{1 \leq i,j \leq k}$, as follows

$$a_{s;i,j} := \sum_{L_i/x = L_j} (\beta(x))^s \ .$$

Then $\beta^s(\mathbf{A}(L) \cap \Sigma_n^\ell) = (1, 0, \ldots, 0) \cdot (\mathcal{A}_L^{\beta,s})^\ell \cdot \mathbb{1}$ where $\mathbb{1}$ is the all ones column vector.
Let $\Phi_L(s) := \lim_{\ell \to \infty} \sqrt[\ell]{\|(\mathcal{A}_L^{\beta,s})^\ell\|}$ be the spectral radius of the matrix $\mathcal{A}_L^{\beta,s}$. According to Theorem 2 of [25] Φ_L is strictly decreasing,[5] $\Phi_L(0) \geq 1$, and $\lim_{s \to \infty} \Phi_L(s) = 0$. Thus, if $\Phi_L(s) < 1$, the sum $\beta^s(\mathbf{A}(L)) = (1, 0, \ldots, 0) \cdot \sum_{\ell \in \mathbb{N}} (\mathcal{A}_L^{\beta,s})^\ell \cdot \mathbb{1}$ converges and, if $\Phi(s) \geq 1$, it diverges. Consequently, $H_L^\beta = H_{\mathbf{A}(L)}^\beta = H_{\mathbf{T}(L)}^\beta = \alpha$ iff $\Phi_L(\alpha) = 1$.
In particular, we have the following.

Corollary 15 *$\beta^\alpha(L) = \infty$ for $\alpha = H_L^\beta$ if L is an infinite regular language.*

3.2 The β-Entropy of the Submonoid

Next we consider the relation between the entropies of L and L^*. As $H_L^\beta \leq H_{L^*}^\beta$ and $H_{L^*}^\beta = \infty$ whenever $\beta(w) \geq 1$ for some $w \in L \setminus \{e\}$, we are interested only in cases when $H_L^\beta < \infty$ and $\beta(w) < 1$ for $w \in L \setminus \{e\}$. This implies also that $\beta(w) < 1$ for $w \in L^* \setminus \{e\}$. First we give some general bounds on the β-entropy of L^*.

Property 16 *Let $e \notin L$, $\alpha = H_L^\beta < \infty$ and $\beta(w) < 1$ for all $w \in L \setminus \{e\}$. Then*

1. *$H_{L^*}^\beta \leq \inf\{s : \beta^s(L) \leq 1\}$ and,*
2. *if L is a code and $\beta^\alpha(L) \geq 1$, then $H_{L^*}^\beta$ is the unique solution of the equation $\beta^s(L) = 1$.*

Proof. 1. Since $\beta(w) < 1$ for all $w \in L$ the function $\beta^s(L)$ is strictly decreasing in (α, ∞). Consequently, if $\beta^s(L) < 1$, in view of the inequality $\beta^s(L^*) \leq \sum_{i \in \mathbb{N}} (\beta^s(L))^i$ we have $\beta^s(L^*) < \infty$.
2. The additional claim that $\beta^s(L) = 1$ implies $H_{L^*}^\beta \geq s$ follows from the fact that for codes L the identity $\beta^s(L^*) = \sum_{i \in \mathbb{N}} (\beta^s(L))^i$ holds.

Q.E.D.

Summarizing 1. and 2. for codes $C \subseteq \Sigma_n^*$ yields the formula

$$H_{C^*}^\beta = \inf\{s : \beta^s(C) \leq 1\} \tag{6}$$

We obtain a condition sufficient for the inequality $H_{L^*}^\beta > H_L^\beta$.

---

[5] More precisely, there is a c, $0 < c < 1$, such that for all $\varepsilon > 0$ the inequality $\Phi_L(s + \varepsilon) \leq c^\varepsilon \cdot \Phi(s)$ holds.

Lemma 17 *If L is a finite union of k codes which satisfies $\beta^{\alpha}(L) > k$ for $\alpha = H_L^{\beta} < \infty$, then $H_{L^*}^{\beta} > H_L^{\beta}$.*

Proof. Let $L = C_1 \cup \ldots \cup C_k$ where all C_i are codes. Then there is an $i \in \mathbb{N}$ such that $\beta^{\alpha}(C_i) > 1$. Because $\beta^s(C_i)$ is continuous on (α, ∞), we have $\inf\{s : \beta^s(C_i) \leq 1\} > \alpha$ and, therefore, $\alpha < H_{C_i^*}^{\beta} \leq H_{L^*}^{\beta}$.

Q.E.D.

In connection with Corollary 15 we obtain.

Corollary 18 *If $L \subseteq \Sigma_n^*$ is regular and a finite union of codes and $H_L^{\beta} < \infty$, then $H_{L^*}^{\beta} > H_L^{\beta}$.*

Next we consider the approximation of the β-entropy of L^*, $H_{L^*}^{\beta}$, via $H_{U^*}^{\beta}$ where U is a finite subset of L. We are going to derive a result analogous to the theorem of [33]. There we used the real numbers λ_m defined as the smallest (positive) roots of the equation $1 = \lambda_m + (\lambda_m)^m$.[6]

In the sequel we assume that there is a positive constant $c < 1$ such that every word $w \in L$ (and, hence also every $w \in L^*$) satisfies $\beta(w) \leq c^{|w|}$. In other words, $L^* \subseteq V_{\beta,c}$ where $V_{\beta,c} := \{w : w \in \Sigma_n^* \wedge \beta(w) \leq c^{|w|}\}$. Observe that $V_{\beta,c}^* \subseteq V_{\beta,c}$.

Theorem 19 *Let L be a nonempty subset of $V_{\beta,c}$. Then for $m \leq \min\{|w| : w \in L \setminus \{e\}\ \}$ and $\varepsilon_m := \log_c \lambda_m$ we have*

$$\beta^s(L^*) \leq \sum_{i \in \mathbb{N}} (\beta^s(L))^i \leq \beta^{s-\varepsilon_m}(L^*)$$

whenever $s \geq \varepsilon_m$.

Proof. As in [33] one obtains

$$\beta^s(L^*) \leq \sum_{i \in \mathbb{N}} (\beta^s(L))^i \leq \sum_{w \in L^*} \lambda_m^{-|w|} \cdot (\beta(w))^s \ .$$

Now $\beta(w) \leq c^{|w|}$ implies $(\beta(w))^{\varepsilon_m} \leq (c^{|w|})^{\varepsilon_m} = \lambda_m^{|w|}$ and, consequently, $\sum_{w \in L^*} \lambda_m^{-|w|} \cdot (\beta(w))^s \leq \beta^{s-\varepsilon_m}(L^*)$.

Q.E.D.

Corollary 20 *Let $L \subseteq V_{\beta,c}$ for some $c < 1$, $e \notin L$ and $\min\{|w| : w \in L\} \geq m > 0$. Then*

$$0 \leq \theta - H_{L^*}^{\beta} \leq \varepsilon_m \text{ whenever } \beta^{\theta}(L) = 1 \ .$$

Proof. If $\beta^{\theta}(L) = 1$, then on the one hand $H_{L^*}^{\beta} \leq \theta$ and on the other hand according to Theorem 19 $\beta^{\theta-\varepsilon_m}(L^*) = \infty$, that is, $H_{L^*}^{\beta} \geq \theta - \varepsilon_m$.

Q.E.D.

We obtain the announced analogue to the theorem of [33].

[6] It is well-known that $\lambda_m^{-\ell}$ upperbounds the number of compositions (ordered partitions) of the number ℓ into parts not smaller than m, and it holds $0 < \lambda_m < \lambda_{m+1} < 1$ and $\lim_{m\to\infty} \lambda_m = 1$ (cf. [33]).

Theorem 21 *Let $L \subseteq V_{\beta,c}$ for some $c < 1$. Then for every $\varepsilon > 0$ there is a finite subset $U \subseteq L$ such that*

$$H^\beta_{L^*} - H^\beta_{U^*} < \varepsilon \ .$$

Proof. Let $H^\beta_{L^*} = \alpha$. It suffices to show that for every $\varepsilon > 0$ there is a finite subset $U \subseteq L$ such that $\beta^{\alpha-2\cdot\varepsilon}(U^*) = \infty$.

If $H^\beta_{L^*} = \alpha$, then $\beta^{\alpha-\varepsilon}(L^*) = \infty$ for all $\varepsilon > 0$. Now choose $m \in \mathbb{N}$ such that $\varepsilon > \varepsilon_m := \log_c \lambda_m$. Since $\beta^{\alpha-\varepsilon}(L^*) = \infty$ there is a finite subset $V \subseteq \{w : w \in L^* \wedge |w| \geq m\}$ satisfying $\beta^{\alpha-\varepsilon}(V) > 1$. Hence by Theorem 19, $\infty = \sum_{i\in\mathbb{N}}(\beta^{\alpha-\varepsilon}(V))^i \leq \beta^{\alpha-\varepsilon-\varepsilon_m}(V^*) \leq \beta^{\alpha-2\cdot\varepsilon}(V^*)$.

Finally we may choose U to be any finite subset of L satisfying $V \subseteq U^*$.

Q.E.D.

As a final remark to this section we derive an upperbound to the β-entropy of the languages $V_{\beta,c}$ where $c < 1$.

$$H^\beta_{V_{\beta,c}} \leq -\log_c n \quad \text{for } V_{\beta,c} \subseteq \Sigma_n^* \text{ and } c < 1. \tag{7}$$

Proof. We have $\beta^s(V_{\beta,c}) \leq \sum_{i\in\mathbb{N}} n^i \cdot c^{s\cdot i} = \sum_{i\in\mathbb{N}}(n \cdot c^s)^i < \infty$ if only $n \cdot c^s < 1$.

Q.E.D.

4 ω-Languages and Hausdorff Dimension

Now we apply our results on valuations of languages to the calculation of the Hausdorff dimension in the spaces $(\Sigma_n^\omega, \rho_\beta)$ where ρ_β is the metric derived from the valuation $\beta : \Sigma_n \to (0,\infty)$ in the following manner

$$\rho_\beta(\xi,\eta) = \begin{cases} 0 & \text{, if } \xi = \eta \text{ , and} \\ \min\{\beta(w) : w \in \mathbf{A}(\xi) \cap \mathbf{A}(\eta)\} & \text{, otherwise.} \end{cases} \tag{8}$$

The case when $\beta(a) = \beta_n(a) = \frac{1}{n}$ for $a \in \Sigma_n$ was investigated in detail in [36], here we generalize the results obtained there. Particular results for arbitrary valuations were obtained in [25] and [2].

4.1 Metric Properties of the Space $(\Sigma_n^\omega, \rho_\beta)$

First we need some properties of the metric ρ_β. It turns out that there is a crucial distinction between the behaviour of the metrics derived from various valuations β, mainly depending on the fact whether $\beta(a) < 1$ for all $a \in \Sigma_n$ or not.

Observe that ρ_β satisfies the ultrametric inequality

$$\rho_\beta(\xi,\eta) \leq \max\{\rho_\beta(\xi,\zeta), \rho_\beta(\eta,\zeta)\} \ , \tag{9}$$

because $\mathbf{A}(\xi) \cap \mathbf{A}(\eta)$ contains at least one of the sets $\mathbf{A}(\xi) \cap \mathbf{A}(\zeta)$ or $\mathbf{A}(\eta) \cap \mathbf{A}(\zeta)$.

In contrast to the investigations in [36] the space $(\Sigma_n^\omega, \rho_\beta)$, however, need not be compact, and even in cases when it is compact the diameter of the balls in $(\Sigma_n^\omega, \rho_\beta)$ need not uniformly correspond to the length of the words defining them.

Closed balls (they are simultaneously open) with center ξ and radius $\epsilon > 0$ in $(\Sigma_n^\omega, \rho_\beta)$, $\overline{\mathbb{B}}_\epsilon(\xi) = \{\eta : \rho_\beta(\xi,\eta) \le \epsilon\}$, are characterized by words in Σ_n^* as follows.
Denote by $w_\beta(\xi,\epsilon)$ the shortest prefix (provided it exists) $w \sqsubset \xi$ such that $\beta(w) \le \epsilon$.

$$\overline{\mathbb{B}}_\epsilon(\xi) = \begin{cases} \{\xi\} & \text{, if } \rho_\beta(\xi,\eta) > \epsilon \text{ , for all } \eta \neq \xi \text{, and} \\ w_\beta(\xi,\epsilon)\cdot\Sigma_n^\omega & \text{, otherwise.} \end{cases} \tag{10}$$

Two remarks are in order here.
Remark 1. A point $\xi \in \Sigma_n^\omega$ such that $\rho_\beta(\xi,\eta) > \epsilon$ for some $\epsilon > 0$ and all $\eta \neq \xi$ is usually called an *isolated point* of $(\Sigma_n^\omega, \rho_\beta)$. If ξ is no isolated point, then $w_\beta(\xi,\epsilon)$ exists for every $\epsilon > 0$.
Remark 2. Since the metric ρ_β satisfies the ultrametric inequality Eq. (9), balls are simultaneously open and closed in $(\Sigma_n^\omega, \rho_\beta)$ and, moreover, every $\eta \in \overline{\mathbb{B}}_\epsilon(\xi)$ can be chosen to be the center of the ball, that is $\overline{\mathbb{B}}_\epsilon(\xi) = \overline{\mathbb{B}}_\epsilon(\eta)$ which shows that its radius equals its diameter.
For the *diameter* of the ball $\overline{\mathbb{B}}_\epsilon(\xi)$ we obtain, similarly to the case $\beta = \beta_n$, the following.

$$\operatorname{diam} \overline{\mathbb{B}}_\epsilon(\xi) = \begin{cases} 0 & \text{, if } \rho_\beta(\xi,\eta) > \epsilon \text{ , for all } \eta \neq \xi \text{, and} \\ \beta(w_\beta(\xi,\epsilon)) & \text{, otherwise.} \end{cases} \tag{11}$$

In particular, we have $\operatorname{diam} \overline{\mathbb{B}}_\epsilon(\xi) \le \epsilon$.
Note that in contrast to the special case $\beta = \beta_n$ in virtue of the (possible) existence of isolated points in $(\Sigma_n^\omega, \rho_\beta)$ not all balls are subsets of the form $w\cdot\Sigma_n^\omega$ and, vice versa, not all subsets of the form $w\cdot\Sigma_n^\omega$ are balls. Consider e.g. $12\cdot\Sigma_2^\omega$ in $(\Sigma_2^\omega, \rho_\beta)$ where $\beta(1) < 1$ and $\beta(2) \ge 1$.
Remark. Since $w\cdot\Sigma_n^\omega = \bigcup_{w\sqsubseteq\xi} \mathbb{B}_\epsilon(\xi)$ for every $0 < \epsilon < \min\{\beta(v) : v \sqsubseteq w\}$, those subsets are always open.
As closed sets are complements of open sets, we obtain that ω-languages $E \subseteq \Sigma_n^\omega$ satisfying

$$E = \{\xi : \mathbf{A}(\xi) \subseteq \mathbf{A}(E)\} = \Sigma_n^\omega \setminus (\Sigma_n^* \setminus \mathbf{A}(E))\cdot\Sigma_n^\omega \tag{12}$$

are closed in every space $(\Sigma_n^\omega, \rho_\beta)$. Therefore, we will refer to ω-languages $E \subseteq \Sigma_n^\omega$ satisfying Eq. (12) as *strongly closed*.
One easily observes that $(\Sigma_n^\omega, \rho_\beta)$ has no isolated points iff $\beta(a) < 1$ for all $a \in \Sigma_n$. Valuations having this property will be called *contractive*. Since Σ_n is finite, for contractive valuations the space $(\Sigma_n^\omega, \rho_\beta)$ is compact and its balls are exactly the sets of the form $w\cdot\Sigma_n^\omega$. Thus, we only have isolated points in Σ_n^ω in case $\beta(a) \ge 1$ for some $a \in \Sigma_n$. The set of all isolated points can be written as

$$\mathbb{I}_\beta := \{\xi : \inf\{\beta(w) : w \sqsubset \xi\} > 0\} \ . \tag{13}$$

It is an easy exercise to show that for noncontractive valuations β we have $\mathbb{I}_\beta = \Sigma_n^*\cdot a^\omega$ iff $\beta(a) = 1$ and $\beta(b) < 1$ for $b \in \Sigma_n \setminus \{a\}$, and that otherwise $\mathbb{I}_\beta$ is uncountable.

4.2 Hausdorff Dimension in $(\Sigma_n^\omega, \rho_\beta)$

In order to introduce the Hausdorff dimension of subsets of $(\Sigma_n^\omega, \rho_\beta)$ we define the α-*dimensional outer Hausdorff measure* induced by ρ_β.

$$\nu_\beta^\alpha(F) := \lim_{\epsilon\to 0}\inf\{\sum_{i\in\mathbb{N}} (\operatorname{diam} F_i)^\alpha : F \subseteq \bigcup_{i\in\mathbb{N}} F_i \wedge \operatorname{diam} F_i < \epsilon\} \tag{14}$$

Remark. It may happen that, due to the fact that F contains uncountably many isolated points, the set on the right hand side of Eq. (14) becomes empty. Then, following our convention $\inf\emptyset = \infty$, we set $\nu_\beta^\alpha(F) := \infty$. In that case we have $\nu_\beta^\alpha(F) := \infty$ independently of the value of α.

Then the *Hausdorff dimension* of $F \subseteq \Sigma_n^\omega$ in $(\Sigma_n^\omega, \rho_\beta)$ is defined as

$$\dim^{(\beta)} F := \inf\{\alpha : \nu_\beta^\alpha(F) = 0\} = \sup\{\alpha : \alpha = 0 \vee \nu_\beta^\alpha(F) = \infty\}\ .$$

Here we mention that the Hausdorff dimension is countably stable, that is,

$$\dim^{(\beta)} \bigcup_{i\in\mathbb{N}} F_i = \sup_{i\in\mathbb{N}} \dim^{(\beta)} F_i \tag{15}$$

In what follows we are mainly interested in the Hausdorff dimension and measure of sets not containing isolated points (at least not uncountably many). Therefore we introduce the *β-fundamental set* of $(\Sigma_n^\omega, \rho_\beta)$, $\mathbb{F}_\beta$, as follows.

$$\mathbb{F}_\beta := \Sigma_n^\omega \setminus \mathbb{I}_\beta = \{\xi : \inf\{\beta(w) : w \sqsubset \xi\} = 0\}\ . \tag{16}$$

As the set of isolated points $\mathbb{I}_\beta$ is open, its complement $\mathbb{F}_\beta$ is a closed subset of $(\Sigma_n^\omega, \rho_\beta)$. If $\mathbb{F}_\beta \neq \emptyset$, it is, however, not strongly closed unless β is contractive. One easily verifies the identity $\mathbb{F}_\beta = \mathbb{F}_\beta/w$ for all $w \in \Sigma_n^*$, that is, $\mathbb{F}_\beta$ is a so-called one-state ω-language, and there are only two strongly closed one-state ω-languages contained in Σ_n^ω: Σ_n^ω itself and $\emptyset$.

For subsets $F \subseteq \mathbb{F}_\beta$ we have the following relation between ν_β^α and the valuation β.

$$\nu_\beta^\alpha(F) := \lim_{\epsilon\to 0} \inf\{\beta^\alpha(L) : F \subseteq L\cdot\Sigma_n^\omega \wedge \forall w(w\in L \to \beta(w) \le \epsilon)\} \tag{17}$$

Proof. On the one hand, $(\beta(w)) \ge (\operatorname{diam} w\cdot\Sigma_n^\omega)$, so the inequality "$\le$" follows.

On the other hand, let $F \subseteq \bigcup_{i\in M} F_i$ and $\sum_{i\in M}(\operatorname{diam} F_i)^\alpha \le \nu_\beta^\alpha(F) + \epsilon$ for some $m \subseteq \mathbb{N}$ and $\epsilon > 0$. Without loss of generality we may assume $F \cap F_i \neq \emptyset$. We consider two cases.

If $\epsilon_i := \operatorname{diam} F_i \ge 0$, then $F_i \subseteq \overline{\mathbb{B}}_{\epsilon_i}(\xi_i) = \{\eta : \rho_\beta(\xi_i,\eta) \le \epsilon_i\}$ for $\xi_i \in F_i$, and $\{\xi_i\} \neq F_i$. According to Eq. (10), $\overline{\mathbb{B}}_{\epsilon_i}(\xi_i) = w_i\cdot\Sigma_n^\omega$ for some $w_i \in \Sigma_n^*$.

If $\operatorname{diam} F_i = 0$, that is, $F_i = \{\xi_i\} \subseteq \mathbb{F}_\beta$, then we may find a $w_i \sqsubset \xi_i$ such that $(\beta(w_i))^\alpha \le \epsilon\cdot 2^{-(i+1)}$.

Consequently, $F \subseteq \bigcup_{i\in M} w_i\cdot\Sigma_n^\omega$ and

$$\beta^\alpha(\{w_i : i \in M\}) \le \sum_{i\in M} \max\{(\operatorname{diam} F_i)^\alpha, \epsilon'\cdot 2^{-(i+1)}\} \le \nu_\beta^\alpha(F) + 2\cdot\epsilon\ ,$$

and the assertion follows, because ϵ can be made arbitrarily small.

Q.E.D.

Next we derive some relations between the β-entropy of languages and the Hausdorff dimension of ω-languages in the space $(\Sigma_n^\omega, \rho_\beta)$.

First we get results analogous to Lemmas 3.8 and 3.10 of [36]. To this end we introduce the *δ-limit* of a language $V \subseteq \Sigma_n^*$:

$$V^\delta := \{\xi : \xi \in \Sigma_n^\omega \wedge \mathbf{A}(\xi) \cap V \text{ is infinite}\} \tag{18}$$

Lemma 22 *If* $\beta^\alpha(V) < \infty$*, then* $\nu_\beta^\alpha(V^\delta) = 0$.

Proof. As in [36] we use the partition of V into $V^{(i)} := \{v : v \in V$ and v has exactly i prefixes in $V\}$. Then $V^\delta \subseteq V^{(i)} \cdot \Sigma_n^\omega$ and, since $\beta^\alpha(V) < \infty$, $\beta^\alpha(V^{(i)})$ tends to 0 as i approaches infinity.

Q.E.D.

Lemma 23 *Let* $F \subseteq \mathbb{F}_\beta$*. Then* $\nu_\beta^\alpha(F) = 0$ *iff there is a language* $L \subseteq \Sigma_n^*$ *such that* $F \subseteq L^\delta$ *and* $\beta^\alpha(L) < \infty$*.*

Proof. Let $\nu_\beta^\alpha(F) = 0$. For $\alpha = 0$ we have $\nu_\beta^\alpha(F) = \operatorname{card} F$. So $F = \emptyset$, and we may choose $L = \emptyset$.
Let $\alpha > 0$. According to Eq. (17), for every $i \in \mathbb{N}$ we can find a language L_i such that $F \subseteq L_i \cdot \Sigma_n^\omega$ and $\beta^\alpha(L_i) < n^{-i}$. This in particular implies that $|w| \geq \frac{i}{\alpha}$ for all $w \in L_i$. Now it is easy to see that $L := \bigcup_{i \in \mathbb{N}} L_i$ satisfies $F \subseteq L^\delta$ and $\beta^\alpha(L) < \infty$.
The other direction is proved in Lemma 22.

Q.E.D.

As immediate consequences of the definition of the Hausdorff dimension we get the following relations between the β-entropy of languages and the Hausdorff dimension of its δ-limits.

$$\dim^{(\beta)} V^\delta \leq H_V^\beta \quad \text{, and} \tag{19}$$

$$\dim^{(\beta)} F = \inf\{\dim^{(\beta)} W^\delta : F \subseteq W^\delta\} \quad \text{, if } F \subseteq \mathbb{F}_\beta \,. \tag{20}$$

4.3 Hausdorff Dimension of ω-Languages

Utilizing the results of [2], [25] and Corollary 14 we can relate Hausdorff dimension and measure of strongly closed finite-state subsets of Σ_n^ω and the β-entropy of their prefix languages.

First we draw a connection between finite-state ω-languages contained in $\mathbb{F}_\beta$ and languages of the form $V_{\beta,c}$ as introduced in Section 3 and we derive an estimate for Hausdorff dimensions and measures of finite-state strongly closed subsets of $\mathbb{F}_\beta$.

Lemma 24 *1. If* $c < 1$ *then* $V_{\beta,c}^\delta \subseteq \mathbb{F}_\beta$*, and this inclusion is proper if* $\mathbb{F}_\beta \neq \emptyset$ *and* β *is not contractive.*

2. For every finite-state and strongly closed ω*-language* $E \subseteq \mathbb{F}_\beta$ *there are a positive* $c < 1$ *and an* $\ell \in \mathbb{N}$ *such that* $E \subseteq \{w : |w| = \ell \wedge \beta(w) \leq c^\ell\}^\omega$*.*

Proof. 1. The first part is immediate. From the additional assumption it follows that $\beta(a) < 1$ and $\beta(b) \geq 1$ for some letters $a, b \in \Sigma_n$. Depending on these values and $c < 1$ it is easy to construct a $\xi \in \{a, b\}^\omega$ such that $\inf\{\beta(w) : w \sqsubset \xi\} = 0$ but $\beta(w) > c^{|w|}$ for all but finitely many $w \sqsubset \xi$.

2. If E is finite-state, its prefix language $\mathbf{A}(E)$ is also finite-state, that is, a regular language. Let $\emptyset \neq \mathbf{A}(E)/w = \mathbf{A}(E)/w \cdot v$ for some w, v with $v \neq e$. Then $w \cdot v^* \subseteq \mathbf{A}(E)$ and, since E is strongly closed $w \cdot v^\omega \in E \subseteq \mathbb{F}_\beta$ whence $\beta(v) < 1$. According to Property 11

it follows $\beta(v) \le c^\ell$ for some $c<1$ and all $v \in \mathbf{T}(E) \cap \Sigma_n^\ell$ where ℓ is sufficiently large. Now the assertion follows from the obvious inclusion $E \subseteq (\mathbf{T}(E) \cap \Sigma_n^m)^\omega$ which holds for arbitrary $m \in \mathbb{N} \setminus \{0\}$.

Q.E.D.

Theorem 25 *If $F \subseteq \mathbb{F}_\beta$ is nonempty, finite-state and strongly closed, then $H^\beta_{\mathbf{A}(F)} = \dim^{(\beta)}(F)$ and, moreover, if $\alpha = \dim^{(\beta)}(F)$, then $\nu^\alpha_\beta(F) > 0$.*

Proof. In [2] and [25] it is shown that $\alpha = \dim^{(\beta)}(F)$ is the solution of the equation $\Phi_{\mathbf{A}(F)}(s) = 1$, and that $\nu^\alpha_\beta(F) > 0$. The remaining assertion follows from our above consideration on the calculation of $\beta^\alpha(\mathbf{A}(L))$.

Q.E.D.

Since U^ω is finite-state and strongly closed if only U is finite, in view of the identity $\mathbf{A}(U^\omega) = \mathbf{A}(U^*)$ and Corollary 14 the Hausdorff dimension of any $U^\omega \subseteq \mathbb{F}_\beta$ is obtained as $\dim^{(\beta)} U^\omega = H^\beta_{U^*}$. This allows us to get via an approximation similar to the one in Theorem 21 a general formula for the Hausdorff dimension of arbitrary ω-powers L^ω.

Lemma 26 *Let $L \subseteq V_{\beta,c}$ for some positive $c<1$. Then $\dim^{(\beta)} L^\omega = \dim^{(\beta)}(L^*)^\delta = H^\beta_{L^*}$.*

Proof. The inequality $\dim^{(\beta)} L^\omega \le \dim^{(\beta)}(L^*)^\delta \le H^\beta_{L^*}$ follows from the inclusion $L^\omega \subseteq (L^*)^\delta$ and Eq. (19).

In order to show the reverse inequality observe that in view of Theorem 21 we have $H^\beta_{L^*} = \sup\{H^\beta_{U^*} : U \subseteq L \text{ and } U \text{ finite }\} = \sup\{\dim^{(\beta)} U^\omega : U \subseteq L \text{ and } U \text{ finite }\} \le \dim^{(\beta)} L^\omega$.

Q.E.D.

Next we obtain a general bound on $\nu^\alpha_\beta(L^\omega)$ for $\alpha = H^\beta_{L^*}$.

Lemma 27 *Let $L \subseteq V_{\beta,c}$. Then $\nu^\alpha_\beta((L^*)^\delta) \le 1$ for $\alpha = H^\beta_{L^*}$.*

Proof. Define $L^{(i)} := \{w : w \in L^* \wedge |w| \ge i \wedge \forall v (v \sqsubset w \wedge |v| \ge i \rightarrow v \notin L^*)\}$. Then $L^{(i)}$ is a prefixcode contained in L^* satisfying $(L^*)^\delta \subseteq L^{(i)} \cdot \Sigma_n^\omega$.

Now on the one hand, following Property 16, $H^\beta_{(L^{(i)})^*} = \inf\{s : \beta^s(L^{(i)}) \le 1\} \le H^\beta_{L^*}$.

On the other hand, let $\alpha := H^\beta_{L^*}$ and $\gamma_i := H^\beta_{(L^{(i)})^*}$. Then $\beta^{\gamma_i}(L^{(i)}) \le 1$ and, therefore, $\alpha \ge \gamma_i$ and $(L^*)^\delta \subseteq L^{(i)} \cdot \Sigma_n^\omega$ imply $\nu^\alpha_\beta((L^*)^\delta) \le \liminf_{i\to\infty} \beta^\alpha(L^{(i)}) \le \liminf_{i\to\infty} \beta^{\gamma_i}(L^{(i)}) \le 1$, because the function $\beta^s(L^{(i)})$ is continuous on $[\gamma_i, \infty)$.

Q.E.D.

Next we consider the *strong closure* of an ω-language E defined as

$$\mathrm{cl}(E) := \{\xi : \mathbf{A}(\xi) \subseteq \mathbf{A}(E)\} \tag{21}$$

The strong closure can be alternatively written as $\mathrm{cl}(E) = (\mathbf{A}(E))^\delta$, it is also known as the adherence of $\mathbf{A}(E)$ (cf. [38], [22] or [6]). It is the smallest strongly closed ω-language containing E, thus independently of the valuation β it contains the smallest closed (with respect to the metric ρ_β) subset of $(\Sigma_n^\omega, \rho_\beta)$ containing E.

Utilizing Lemma 26 and Corollary 14 we obtain an estimate for the Hausdorff dimension of the strong closure of an ω-power of a regular language.

Corollary 28 *If* $0 < c < 1$ *and* $L \subseteq V_{\beta,c}$ *is a regular language, then* $\dim^{(\beta)} L^\omega = \dim^{(\beta)} \mathrm{cl}(L^\omega)$.

Moreover, we have the following.

Corollary 29 *If* $c < 1$ *and* $L \subseteq V_{\beta,c}$ *is regular and a finite union of codes and* $\alpha = H^\beta_{L^*}$, *then* $0 < \nu^\alpha_\beta(L^\omega) = \nu^\alpha_\beta(\mathrm{cl}(L^\omega)) \leq 1$.

Proof. Utilizing the inclusions $\mathrm{cl}(L^\omega) \subseteq L^\omega \cup L^* \cdot (\mathbf{A}(L))^\delta$ and $\mathbf{A}((\mathbf{A}(L))^\delta) \subseteq \mathbf{A}(L)$, in virtue of Eq. (19) and the Corollaries 14 and 18 $\dim^{(\beta)}(\mathbf{A}(L))^\delta \leq H^\beta_{\mathbf{A}(L)} < H^\beta_{L^*} = \dim^{(\beta)} L^\omega$, the equality $\nu^\alpha_\beta(L^\omega) = \nu^\alpha_\beta(\mathrm{cl}(L^\omega))$ follows. The remaining inequalities are Theorem 25 and Lemma 27.

Q.E.D.

Remark. Utilizing more involved calculations as carried out in [26, Theorem 6] one can show $0 < \nu^\alpha_\beta((L^*)^\delta) = \nu^\alpha_\beta(\mathrm{cl}(L^\omega)) \leq 1$ for arbitrary regular languages $L \subseteq V_{\beta,c}$, but in the case of nonregular languages W one might even have $\dim^{(\beta_n)} W^\omega < \dim^{(\beta_n)} \mathrm{cl}(W^\omega)$ (cf. [36, Examples 6.3 and 6.5]).
Corollary 29 readily implies [17, Theorem 8].

5 IIFS and Fractal Geometry

5.1 Iterated Function Systems

One of the most popular ways of describing fractals is iterated function systems (IFS) [3]. We restrict ourselves in the following to Euclidean spaces $X \subseteq \mathbb{R}^m$ equipped with the Euclidean distance ρ_E. Denoting the set of contracting similitudes $f : X \to X$ by $\mathcal{S}(X)$, we can describe an IFS $\mathcal{F}$ as a map $\mathcal{F} : \Sigma_n \to \mathcal{S}(X)$. We will sketch some well-known properties of IFS in the following, putting emphasis onto the connections with our theory developed in the last sections.

An IFS $\mathcal{F}$ defines a contractive valuation $\beta_{\mathcal{F}} : \Sigma_n^* \to (0,\infty)$, where $\beta_{\mathcal{F}}(i)$ (for $i \in \Sigma_n$) denotes the similarity factor of the similitude $\mathcal{F}(i)$. If $w \in \Sigma_n^+$, we can interpret w as a similitude $\phi_{\mathcal{F}}(w) \in \mathcal{S}(X)$, where $\phi_{\mathcal{F}}$ is a semigroup morphism mapping $(\Sigma_n^+, \cdot)$ into $(\mathcal{S}(X), \circ)$. Denoting by $\xi[m]$ the prefix of length m of $\xi \in \Sigma_n^\omega$, we may consider the sequence $(\phi_{\mathcal{F}}(\xi[m])(x))_m$ for some given point $x \in X$. It is well-known that the above sequence converges to some point which we denote by $\phi_{\mathcal{F}}(\xi)$, since it is independent of the initial point x. ξ is also called an *address* of the point $\phi_{\mathcal{F}}(\xi)$. The such-defined map $\phi_{\mathcal{F}} : (\Sigma_n^\omega, \rho_{\beta_{\mathcal{F}}}) \to (X, \rho_E)$ is Lipschitz continuous. Following [24], we call the set $A_{\mathcal{F}} = \phi_{\mathcal{F}}(\Sigma_n^\omega)$ *limit set* of $\mathcal{F}$.

Given an IFS $\mathcal{F} : \Sigma_n \to \mathcal{S}(X)$, we may interpret a finite (m-element) language $L = \{w_1, \ldots, w_m\} \subset \Sigma_n^+$ as an IFS $\mathcal{F}_L : \Sigma_m \to \mathcal{S}(X), i \mapsto \phi_{\mathcal{F}}(w_i)$. We have $A_{\mathcal{F}_L} = \phi_{\mathcal{F}}(L^\omega)$. Up to now, we restricted our attention to finite languages. What about infinite ones? When considering infinite languages, we are led to infinite IFS (IIFS) [44, 1, 16, 24] whose theory is more involved but analogous to the theory of IFS. Without going into detail here, we can still define a set described by an IIFS $\mathcal{F}_L$ (based on the IFS $\mathcal{F}$ and the language L), namely the *limit set* $\phi_{\mathcal{F}}(L^\omega)$.[7] Observe that in the space $(\Sigma_n^\omega, \rho_{\beta_n})$, we may interpret any

[7] When restricting our attention to compact sets, we should consider the closure of $\phi_{\mathcal{F}}(L^\omega)$ instead.

language $L \subseteq \Sigma_n^+$ as an (I)IFS defining $w : \Sigma_n^\omega \to \Sigma_n^\omega, x \mapsto w \cdot x$. In this interpretation, the limit set of L is just L^ω.

We define, for $L \subset \Sigma_n^+$, the *valuation dimension*

$$\mathrm{valdim}_\beta(L) = \inf\{s > 0 : \beta^s(L) \leq 1\}. \tag{22}$$

Property 16 shows the close relation of $\mathrm{valdim}_\beta(L)$ and $H_{L^*}^\beta$. As we will see, the valuation dimension corresponds to the similarity dimension known from IFS theory. This motivates the introduction of this notion in this context.

We denote the s-dimensional outer Hausdorff measure on a Euclidean space (X, ρ_E) by $\mathcal{H}^s$, and the corresponding Hausdorff dimension by $\dim_H$. For IFS, *Moran's open set condition (OSC)* is well-known as an assumption alleviating the determination of the Hausdorff dimension of $A_\mathcal{F}$ [3]: Provided there is an open bounded non-empty test set $M \subseteq X$ such that $\mathcal{F}(i)(M) \subseteq M$ for any $i \in \Sigma_n$, and that, furthermore, for any $i, j \in \Sigma_n$, $i \neq j$, $\mathcal{F}(i)(M) \cap \mathcal{F}(j)(M) = \emptyset$, then, for $\alpha = \mathrm{valdim}_{\beta_\mathcal{F}}(\Sigma_n)$, $0 < \mathcal{H}^s(A_\mathcal{F}) < \infty$, and $\alpha = \dim_H(A_\mathcal{F})$.

Generally, it is not trivial to find a test set for a given (I)IFS $\mathcal{F}$. But, if we knew that $\mathcal{F}$ fulfills OSC, (when) could we say something about $\mathcal{F}_L$? An answer to this question is given in the next theorem. To this end, we need two further notions [37]. We say that a language $V \subseteq \Sigma_n^*$ is an *OSC-code* iff there is a nonempty $W \subseteq \Sigma_n^*$ such that

$$\forall v(v \in V \to v \cdot W \cdot \Sigma_n^\omega \subseteq W \cdot \Sigma_n^\omega) \ , \quad \text{and} \tag{23}$$

$$\forall v, v'(v, v' \in V \wedge v \neq v' \to v \cdot W \cdot \Sigma_n^\omega \cap v' \cdot W \cdot \Sigma_n^\omega = \emptyset) \tag{24}$$

are true. We will refer to the language $W \subseteq \Sigma_n^*$ also as an OSC-*witness* for V. Note that any OSC-code is a code, and any prefixcode is an OSC-code. In [37] it is shown that any regular code is an OSC-code. Observe further the correspondence with the Euclidean case: Interpreting V as an (I)IFS in the space Σ_n^ω, V satisfies the OSC with open test set $W \cdot \Sigma_n^\omega$ iff V is an OSC-code with OSC-witness W.

Theorem 30 *Let $\mathcal{F} = (\varphi_1, \ldots, \varphi_n)$ where $\varphi_i : \mathbb{R}^d \to \mathbb{R}^d$ be an IFS satisfying the OSC, and let $C \subseteq \Sigma_n^*$ be an OSC-code. Then $\mathcal{F}_C$ is an (I)IFS which satisfies also the OSC.*

Proof. Let $\mathcal{F} = (\varphi_1, \ldots, \varphi_n)$ satisfy the OSC with test set $M \subseteq \mathbb{R}^d$, and let $W \subseteq \Sigma_n^*$ be an OSC-witness for C. Let $\mathcal{F}_C = (\varphi_v)_{v \in C}$, where $\varphi_v := \varphi_{v_1} \circ \cdots \circ \varphi_{v_\ell}$ for $v = v_1 \cdots v_\ell$. Define $M' := \bigcup_{w \in W} \varphi_w(M)$. The set M' is nonempty and open, because all φ_i are similitudes and M is nonempty and open, and moreover, $\varphi_v(M') = \varphi_v(\bigcup_{w \in W} \varphi_w(M)) = \bigcup_{u \in v \cdot W} \varphi_u(M) \subseteq M'$ for $v \in C$. Now consider, for $v, v' \in C$, $v \neq v'$,

$$\begin{aligned} \varphi_v(M') \cap \varphi_{v'}(M') &= \left(\bigcup_{w \in W} \varphi_{v \cdot w}(M)\right) \cap \left(\bigcup_{w \in W} \varphi_{v' \cdot w}(M)\right) \\ &= \bigcup_{w, w' \in W} (\varphi_{v \cdot w}(M) \cap \varphi_{v' \cdot w'}(M)) \ . \end{aligned}$$

Since $v \cdot w \cdot \Sigma_n^\omega \cap v' \cdot w' \cdot \Sigma_n^\omega = \emptyset$, neither $v \cdot w \sqsubseteq v' \cdot w'$ nor $v' \cdot w' \sqsubseteq v \cdot w$. Hence, we have a first position where $v \cdot w$ and $v' \cdot w'$ do not coincide, that is, $u \cdot i \sqsubseteq v \cdot w$ and $u \cdot j \sqsubseteq v' \cdot w'$ where $1 \leq i < j \leq n$, and we have $\varphi_{u \cdot i}(M) \supseteq \varphi_{v \cdot w}(M)$ and $\varphi_{u \cdot j}(M) \supseteq \varphi_{v' \cdot w'}(M)$.

Now from the inclusion $\varphi_\ell(M) \subseteq M$ (one part of the OSC) and the fact that $\varphi_i(M) \cap \varphi_j(M) = \emptyset$ we readily obtain our assertion.

Q.E.D.

Together with [15, Theorem 3.11], we immediately obtain $\dim_H(\phi_{\mathcal{F}}(L^\omega)) = \mathrm{valdim}_{\beta_{\mathcal{F}}}(L)$ if the IFS $\mathcal{F} : \Sigma_n \to \mathcal{S}(X)$ satisfies the OSC and L is an OSC-code.

Our considerations from the previous sections together with Theorem 3 of [2], however, allow to strengthen the mentioned result and to generalize it to not necessarily contractive valuations.

In [2] and [25] IFS have been generalized to systems $\mathcal{F}$ containing arbitrary similitudes. In order to guarantee the convergence of the sequence $(\phi_{\mathcal{F}}(\xi[m])(x))_m$ one has to restrict the set of admissible ω-words ξ. In [2, Theorem 3] it is shown that the mapping $\phi_{\mathcal{F}} : (E, \rho_{\beta_{\mathcal{F}}}) \to (X, \rho_E)$ is Lipschitz continuous whenever E is a strongly closed finite-state subset of $\mathbb{F}_{\beta_{\mathcal{F}}}$.

In connection with this, the following generalization of the Open Set Condition for pairs $(\mathcal{F}, E)$ satisfying the above mentioned property is introduced.

Let $\mathcal{M}$ be a finite set of open subsets of (X, ρ_E). To every $w \in \Sigma_n^*$ we assign a set $M_w \in \mathcal{M}$. We say that the assignment is *compatible* with E iff

$$M_w = \emptyset \quad \to \quad w \notin \mathbf{A}(E) \;, \tag{25}$$

$$\bigcup_{i=1}^{n} \varphi_i(M_{w\cdot i}) \quad \subseteq \quad M_w \;, \text{ and} \tag{26}$$

$$\varphi_i(M_{w\cdot i}) \cap \varphi_j(M_{w\cdot j}) \quad = \quad \emptyset \;, \text{ for } i \neq j \;. \tag{27}$$

We say that a pair $(\mathcal{F}, E)$ satisfies the *Generalized Open Set Condition (GOSC)* iff E is a finite-state strongly closed subset of $\mathbb{F}_{\beta_{\mathcal{F}}}$ and there are a finite set of open subsets of (X, ρ_E), $\mathcal{M}$, and an assignment $w \mapsto M_w \in \mathcal{M}$ compatible with E.

Due to Eq. (25), for every finite-state strongly closed subset $F \subseteq E$ the pair $(\mathcal{F}, F)$ satisfies GOSC provided $(\mathcal{F}, E)$ satisfies GOSC.

Now Theorem 3 of [2] yields the following estimate of Hausdorff dimension and measure.

Theorem 31 *Let E be a finite-state strongly closed subset of $\mathbb{F}_{\beta_{\mathcal{F}}}$ such that $(\mathcal{F}, E)$ satisfies GOSC. Then $\dim_H \phi_{\mathcal{F}}(E) = H^{\beta_{\mathcal{F}}}_{\mathbf{A}(E)} = \dim^{(\beta_{\mathcal{F}})} E$ and, moreover, $\mathcal{H}^\alpha(\phi_{\mathcal{F}}(E)) > 0$ for $\alpha = H^{\beta_{\mathcal{F}}}_{\mathbf{A}(E)}$.*

We proceed with the announced strengthening of the identity $\dim_H(\phi_{\mathcal{F}}(L^\omega)) = \mathrm{valdim}_{\beta_{\mathcal{F}}}(L)$.

Theorem 32 *Let (X, ρ_E) be a Euclidean space, $\mathcal{F} : \Sigma_n \to \mathcal{S}(X)$, E be a finite-state and strongly closed subset of $\mathbb{F}_\beta$, and let $L \subseteq \Sigma_n^*$ such that $L^\omega \subseteq E$. Assume the pair $(\mathcal{F}, E)$ satisfies the GOSC. Then $\dim_H(\phi_{\mathcal{F}}(L^\omega)) = \dim^{(\beta_{\mathcal{F}})} L^\omega$, and provided L is a code, we have $\dim_H(\phi_{\mathcal{F}}(L^\omega)) = \mathrm{valdim}_{\beta_{\mathcal{F}}}(L)$.*

Remark. An analogous theorem for IIFS satisfying the OSC (using the notion of topological pressure function) is given in [24, Theorem 3.15]. Confer also [17, Theorem 10].

Proof. Since $\phi_{\mathcal{F}} : E \to X$ is Lipschitz, clearly $\dim_H(\phi_{\mathcal{F}}(L^\omega)) \leq \dim^{(\beta)} L^\omega = H^{\beta_{\mathcal{F}}}_{L^*} \leq \mathrm{valdim}_{\beta_{\mathcal{F}}}(L)$.

For each of the finite languages $L_m = L \cap \{w \in \Sigma_n^* : |w| \le m\}$, the ω-language $L_m^\omega \subseteq E$ is finite-state and strongly closed, hence $(\mathcal{F}, L_m^\omega)$ satisfies the GOSC, and according to Theorem 31 we have $\dim_H \phi_{\mathcal{F}}(L_m^\omega) = H^{\beta_{\mathcal{F}}}_{\mathbf{A}(L_m^*)} = H^{\beta_{\mathcal{F}}}_{L_m^*}$. Since $(L_m)_{m\in\mathbb{N}}$ is an increasing chain of sets with $\bigcup_{m\in\mathbb{N}} L_m = L$, by Theorem 21, $\lim_{m\to\infty} H^{\beta_{\mathcal{F}}}_{L_m^*} = H^{\beta_{\mathcal{F}}}_{L^*}$ which in turn equals $\dim^{(\beta)}(L^\omega)$ by Lemma 26. Hence, $\dim_H(\phi_{\mathcal{F}}(L^\omega)) \ge \sup_{m\in\mathbb{N}} \dim_H(\phi_{\mathcal{F}}(L_m^\omega)) = \dim^{(\beta)}(L^\omega)$. The additional assertion in case L is a code follows from Eq. (6).

Q.E.D.

In case of a regular language L we can strengthen our result.

Theorem 33 *Let (X, ρ_E) be a Euclidean space, $\mathcal{F} : \Sigma_n \to \mathcal{S}(X)$, and let $L \subseteq \Sigma_n^*$ be a regular language such that $\beta_{\mathcal{F}}(w) < 1$ for all $w \in L \setminus \{e\}$. Then $\mathrm{cl}(L^\omega) \subseteq \mathbb{F}_{\beta_{\mathcal{F}}}$ and $\dim_H(\phi_{\mathcal{F}}(\mathrm{cl}(L^\omega)) = \dim_H(\phi_{\mathcal{F}}(L^\omega)) = \dim^{(\beta_{\mathcal{F}})} L^\omega$.*
If, moreover, L is a finite union of codes then $\mathcal{H}^s(\phi_{\mathcal{F}}(L^\omega)) = \mathcal{H}^s(\phi_{\mathcal{F}}(\mathrm{cl}\, L^\omega))$ for $s \in [0, \infty)$.

Proof. If $\beta_{\mathcal{F}}(w) < 1$ for all $w \in L \setminus \{e\}$, then according to Corollary 12 and Property 11 we have $\beta_{\mathcal{F}}(w) \le c_L^{|w|}$ for some $c_L < 1$ and for all $w \in \mathbf{T}(L^*)$ with $|w| \ge \ell$ for a suitably chosen $\ell \in \mathbb{N}$. In particular, $\mathbf{A}(\xi) \subseteq \mathbf{A}(L^\omega) \subseteq \mathbf{T}(L^*)$ implies $\xi \in \mathbb{F}_{\beta_{\mathcal{F}}}$ which proves our first assertion. Since L is regular, $\dim^{(\beta_{\mathcal{F}})} L^\omega = \dim^{(\beta_{\mathcal{F}})} \mathrm{cl}(L^\omega)$ according to Corollary 28, and the second assertion follows from Theorems 31 and 32.

Now, let L be a finite union of codes and $\alpha = \dim^{(\beta)} L^\omega$. Applying the fact that $\dim^{(\beta)}(\mathbf{A}(L))^\delta < \alpha = \dim^{(\beta)} L^\omega$ utilized in the proof of Corollary 29, we have $\dim_H(\phi_{\mathcal{F}}((\mathbf{A}(L))^\delta)) < \alpha$ and, therefore, $\mathcal{H}^s(\phi_{\mathcal{F}}((\mathbf{A}(L))^\delta)) = 0$ for $s \ge \alpha$. In case $s < \alpha$ we have obviously $\mathcal{H}^s(\phi_{\mathcal{F}}(L^\omega)) = \mathcal{H}^s(\phi_{\mathcal{F}}(\mathrm{cl}\, L^\omega)) = \infty$.

Q.E.D.

Remark. In [15, Remark 3.12], the question was raised whether requiring an OSC for each IFS-part $\mathcal{F}_n = (\varphi_1, \ldots, \varphi_n)$ of a given IIFS $\mathcal{F} = (\varphi_1, \varphi_2, \ldots)$ is weaker than requiring an OSC for $\mathcal{F}$ itself. We can show the following here: If all $\mathcal{F}_n$ fulfill an OSC, then $\mathcal{F}$ itself does not necessarily satisfy an OSC.

Proof. Consider as basic IFS $\mathcal{F} : \Sigma_2 \to \mathcal{S}(([0,1], \rho_E))$ defined by $\mathcal{F}(1)(x) = x/2$ and $\mathcal{F}(2)(x) = x/2+1/2$. It is clear that $A_{\mathcal{F}} = [0,1]$. Consider the suffixcode $L = \{w12^{|w|} : w \in \Sigma_2^*\}$ (which is no OSC-code) from [37, Example 1]. Assume the IIFS $\mathcal{F}_L$ satisfies an OSC with test set M. Then, $\phi_{\mathcal{F}}^{-1}(M) \ne \emptyset$ is open in the topology of $(\Sigma_2^\omega, \rho_2)$, and defines $\emptyset \ne W \subseteq \Sigma_2^*$ by $W\Sigma_2^\omega = \phi_{\mathcal{F}}^{-1}(M)$. We show that W is an OSC-witness for L, contradicting [37].

Assume to the contrary that W is no OSC-witness. Then, condition (23) or (24) is violated. Now, assume there are $w, w' \in W$ and $v, v' \in L$ such that $\xi \in v \cdot w \cdot \Sigma_2^\omega \cap v' \cdot w' \cdot \Sigma_2^\omega \ne \emptyset$. Then, $\phi_{\mathcal{F}}(\xi) \in \phi_{\mathcal{F}}(v)(M) \cap \phi_{\mathcal{F}}(v')(M)$, contradicting that M is a test set. Finally, if $v \cdot W \cdot \Sigma_2^\omega \not\subseteq W \cdot \Sigma_2^\omega$ for some $v \in L$, then there is a $\xi \in v \cdot W \cdot \Sigma_2^\omega$ such that $\xi \notin W \cdot \Sigma_2^\omega$. $\phi_{\mathcal{F}}(\xi) \in \phi_{\mathcal{F}}(v)(M)$ implies $\phi_F(\xi) \in M$, since M is a test set. Hence, $\xi \in W \cdot \Sigma_2^\omega = \phi_{\mathcal{F}}^{-1}(M)$, a contradiction.

Q.E.D.

5.2 Calculating Dimensions

For languages L given by unambiguous regular expressions or unambiguous contextfree grammars, we can use the results from Section 2 to evaluate $\beta^s(L)$. In combination with the last two theorems, this allows a simple calculation of the Hausdorff dimension of $\phi_{\mathcal{F}}(L^\omega)$.

In the case of regular languages, we just note that taking infima in the definition of the valuation dimension is not necessary.

Theorem 34 *Let $\beta : \Sigma_n^* \to (0,\infty)$ be a contractive valuation. Let $L \subseteq \Sigma_n^+$ be a regular e-free language and $\alpha = \mathrm{valdim}_\beta(L)$. Then, $\beta^\alpha(L) = 1$.*

Proof. If $\beta^\alpha(L) < 1$, then $\beta^\alpha(L^*) \leq \sum_{i\in\mathbb{N}}(\beta^\alpha(L))^i < \infty$, contradicting Corollary 15.

Q.E.D.

In case of regular expressions, we obtain the following algorithm:

Input: an expression $R \in \mathcal{UR}_n$ describing a code $[R] \subseteq \Sigma_n^+$; a contractive valuation $\beta : \Sigma_n^* \to (0,\infty)$

Output: the Hausdorff dimension $\dim_H([R]^\omega) = \dim_H(\mathrm{cl}[R]^\omega)$

Procedure: compute the $\alpha \in [0,\infty)$ satisfying $1 = \beta^\alpha_{\mathcal{R}}(R)$

We have seen in Theorem 33 that we can apply the above procedure in order to determine the Hausdorff dimension of Euclidean sets, too. Alternatively, of course it it possible to employ the eigenvalue method sketched in the end of Subsection 3.1.

Languages given by unambiguous contextfree grammars might be treated similarly. Unfortunately, we are now forced either to consider taking infima in the definition of valuation dimension or to prove that the t-dimensional valuation $\beta^t(L)$ of the examined language L is indeed finite and greater than 1 for some $t < s$ (where the dimension α is calculated assuming $\beta^\alpha(L) = 1$). Namely, consider Kuich's example (see also [36, Example 6.3]) $x_1 = 1 + 2x_1x_1x_1$ defining the prefixcode $L \subseteq \Sigma_2^*$. For a valuation $\beta(1) = \beta(2) = 1/2$, we have $\beta^s(L) = \infty$ iff $2^{-s} > \frac{1}{3}\sqrt[3]{4} = K \approx 0.5291$, and, for $s_0 = \frac{-\ln K}{\ln 2} \approx 0.9183$, $\beta^{s_0}(L) = \sqrt[3]{0.5} < 1$. An ad hoc approach $1 = 2^{-s} + 2^{-s}$ would wrongly yield $s = 1$. Note that the more cautious calculation $x = 0.5 + 0.5x^3$ in the case $s = 1$ would lead to two possible solutions, $x = 1$ and $x \approx 0.6180$, where the latter value perfectly fits into the picture delivered by $\beta^{s_0}(L) \approx 0.7937$, where $s_0 < 1$. In view of Theorem 9, we see that we made the mistake not to choose the minimal solution of the equation $x = 0.5 + 0.5x^3$.

5.3 Some Fractals

We show how to apply our results in some examples in the following. As a basic IFS, we take the quadtree IFS $\mathcal{F} : \Sigma_4 \to \mathcal{S}([0,1]^2)$. The effect of its four mappings is indicated in Figure 1: E.g., $\mathcal{F}(2)$ maps the unit square onto its lower right fourth. The words in the center of the subsquares are the prefixes of the addresses of its points. For example, any address of the point indicated by the $\times$-symbol starts with 11. Of course, $A_{\mathcal{F}} = [0,1]^2$ is fairly uninteresting. We remark that $V = (0,1)^2$ may serve as a test set for OSC. Firstly, taking $L = \{1,2,3\}$, we get the well-known Sierpiński triangle. Since L is a code, we may solve $\beta^s_{\mathcal{F}}(L) = 3(0.5)^s = 1$, delivering $\alpha = \frac{\ln 3}{\ln 2} \approx 1.5850$ as the Hausdorff dimension.

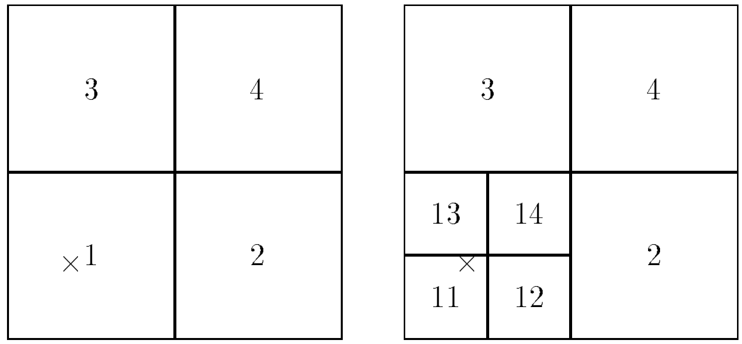


Figure 1: Quadtree

Secondly, consider the regular expression $R = (4{*}3{*})(1 \cup 2)$. Taking as indeterminate y with $\beta(i) = y$ for $i \in \Sigma_4$, we get $\beta_{\mathcal{R}}(R) = \beta_{\mathcal{R}}(4{*}3{*})\beta_{\mathcal{R}}(1 \cup 2) = \beta_{\mathcal{R}}(4{*})\beta_{\mathcal{R}}(3{*})(\beta(1) + \beta(2)) = \frac{1}{1-y}\frac{1}{1-y}2y$. Substituting $y = 2^{-s}$ and solving $\beta_{\mathcal{R}}(R) = 1$, we obtain $\alpha \approx 1.9000$ as the Hausdorff dimension of $\phi_{\mathcal{F}}([R]^\omega)$ (and of $\phi_{\mathcal{F}}(\mathrm{cl}[R]^\omega)$ as well).

Thirdly, we consider the language $M = \{1^i 2^i : i \in \mathbb{N}\backslash\{0\}\} \cup \{1^i 3 : i \in \mathbb{N}\}$. M is generated by the unambiguous linear contextfree grammar given by the following equations (x_1 is the start symbol):

$$\begin{aligned} x_1 &= 1x_2 2 + 12 + 1x_3 + 3 \\ x_2 &= 1x_2 2 + 12 \\ x_3 &= 1x_3 + 3 \end{aligned}$$

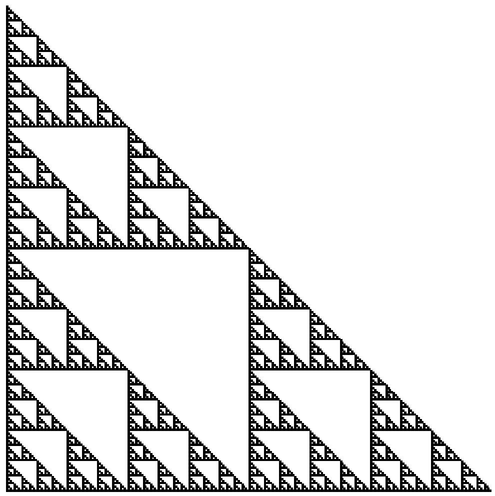
Taking as indeterminate y with $\beta(i) = y$ for $i \in \Sigma_4$, and setting $x_1 = 1$ in the numerical system, we get the following system we have to solve with some $y \in (0,1)$:

$$\begin{aligned} 1 &= y^2 x_2 + y^2 + y x_3 + y \\ x_2 &= y^2 x_2 + y^2 \\ x_3 &= y x_3 + y \end{aligned}$$

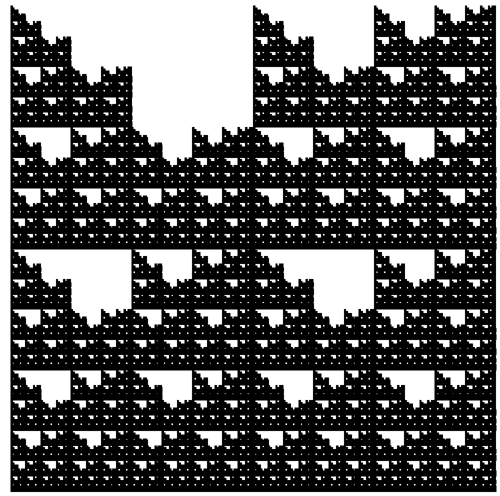
Hence, $y = \frac{-1+\sqrt{13}}{6} \approx 0.4343$, delivering as dimension ≈ 1.2034. Since every choice of the parameter y uniquely determines a solution (x_1, x_2, x_3), in view of Theorem 9, our approach letting $x_1 = 1$ is justified.

Finally, we can make a similar approach for $N = M \cup \{2^i 3 : i \in \mathbb{N}\}$. We obtain as its dimension $\alpha \approx 1.4073$.

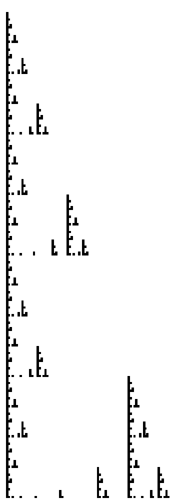
We like to remark that there is another but related approach connecting languages and fractals (based on an IFS $\mathcal{F} : \Sigma_n \to \mathcal{S}(X)$): Take an ω-language $F \subseteq \Sigma_n^\omega$ and consider the (fractal) set $\phi_{\mathcal{F}}(F)$. For example, using regular (or finite-state strongly closed) ω-languages, we obtain in such a way a class of fractals known under different names: Generalized recurrent systems [9], graph directed constructions [25], recurrent IFS [4], MRFS [8], hierarchical IFS [28], see also [27]. By the well-known McNaughton theorem, a regular ω-language F can be represented in the form $F = \bigcup_{i=1}^m W_i \cdot V_i^\omega$, where the V_i's are regular prefixcodes. Hence, $\dim_H(\phi_{\mathcal{F}}(F)) = \max_{i=1}^m \dim_H(\phi_{\mathcal{F}}(V_i^\omega))$, where the latter dimensions can be computed easily presuming the V_i's are given by unambiguous regular expressions. In such a way, we obtain another way for determining the Hausdorff dimension of those fractals as well.



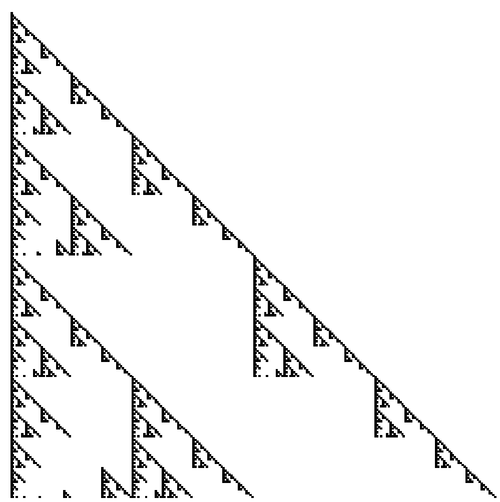
[L]



[[R]]



[M]



[N]

Figure 2: Sierpiński and its variants

Our method is not restricted to the calculation of the Hausdorff dimension of regular ω-languages and their fractal counterparts in Euclidean space: It is well-known that also other ω-languages, e.g. contextfree ω-languages, are of the form $F = \bigcup_{i=1}^{m} W_i \cdot V_i^{\omega}$ (cf. [32]). (Here the languages W_i, V_i are not necessarily regular.) It is clear from the formulae derived in this and in the preceding sections that for fractals related to ω-languages of this shape the Hausdorff dimension can be calculated as soon as we are able to calculate the β-entropy of the corresponding languages, thus leading to a problem related to formal language theory.

6 Conclusions

We presented results connecting formal language theory and fractal geometry. It is intriguing that notions like unambiguity and codes taken from the formal language side can be meaningfully interpreted in fractal geometry. On the other hand, we were led to the concept of a valuation by our research in the area of fractal geometry. It turned out that this concept is also interesting from the purely language theoretical point of view, especially when keeping in mind that, besides combinatorial arguments, concepts relating numbers to words and languages are rarely encountered in the theory of formal languages.

There are other connections between formal language theory and fractals, e.g. questions of hierarchies of fractal descriptions inherited by language hierarchies [14, 16] and decidability issues [10].

We think it is promising to pursue further research connecting formal language theory with other parts of mathematics, such as fractal geometry.

References

[1] L. M. Andersson. *Recursive Construction of Fractals.* PhD thesis, Helsinki: Suomalainen Tiedeakatemia, Aug. 1992. Annales Academiae Scientiarum Fennicae Series A, I. Mathematica, Dissertationes, 86.

[2] C. Bandt. Self-similar sets 3. Constructions with sofic systems. *Monatshefte für Mathematik*, 108:89–102, 1989.

[3] M. F. Barnsley. *Fractals Everywhere.* Boston: Academic Press, 1988.

[4] M. F. Barnsley, J. H. Elton, and D. P. Hardin. Recurrent iterated function systems. *Constructive Approximation*, 5:3–31, 1989.

[5] J. Berstel and D. Perrin. *Theory of Codes.* Pure and Applied Mathematics. Orlando: Academic Press, 1985.

[6] L. Boasson and M. Nivat. Adherences of languages. *Journal of Computer and System Sciences*, 20:285–309, 1980.

[7] A. Brüggemann-Klein. Regular expressions into finite automata. In *LATIN'92*, volume 483 of *LNCS*, pages 87–98, 1992.

[8] K. Čulik, II and S. Dube. Methods for generating deterministic fractals and image processing. In *LNCS: 464; IMYCS*, pages 2–28. Springer, 1990.

[9] F. M. Dekking. Recurrent sets: A fractal formalism. Technical Report 82-32, Technische Hogeschool, Delft (NL), 1982.

[10] S. Dube. Fractal geometry, Turing machines and divide-and-conquer recurrences. *RAIRO Informatique théorique et Applications/Theoretical Informatics and Applications*, 28(3–4):405–423, 1994.

[11] S. Eilenberg. *Automata, Languages, and Machines, Volume A.* Pure and Applied Mathematics. New York: Academic Press, 1974.

[12] J. L. Encarnação et al., editors. *Fractal Geometry and Computer Graphics*, Beiträge zur Graphischen Datenverarbeitung des ZGDV (Darmstadt). Springer-Verlag, 1992.

[13] H. Fernau. Valuations, regular expressions, and fractal geometry. Submitted for publication, Dec. 1993.

[14] H. Fernau. Valuations of languages, with applications to fractal geometry. To appear in Theoretical Computer Science (1995), Volume **143**, Sept. 1993 (submitted).

[15] H. Fernau. Infinite iterated function systems. *Mathematische Nachrichten*, 169, 1994.

[16] H. Fernau. *Iterierte Funktionen, Sprachen und Fraktale.* Mannheim: BI-Verlag, 1994.

[17] H. Fernau and L. Staiger. Valuations and unambiguity of languages, with applications to fractal geometry. In S. Abiteboul and E. Shamir, editors, *Automata, Languages and Programming, 21st International Colloquium, ICALP 94*, volume 820 of *LNCS*, pages 11–22, July 1994. ISBN: 3-540-58201-0.

[18] C. A. Gunter and D. S. Scott. *Semantic Domains*, chapter 12, pages 633–674. Elsevier & MIT Press, 1992. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science, Volume B, Formal Models and Semantics.*

[19] A. Habel, H.-J. Kreowski, and S. Taubenberger. Collages and patterns generated by hyperedge replacement. *Languages of Design*, 1:125–145, 1993.

[20] W. Kuich. On the entropy of context-free languages. *Information and Control (now Information and Computation)*, 16:173–200, 1970.

[21] W. Kuich and A. Salomaa. *Semirings, Automata, Languages*, volume 5 of *EATCS Monographs on Theoretical Computer Science.* Berlin: Springer, 1986.

[22] R. Lindner and L. Staiger. *Algebraische Codierungstheorie; Theorie der sequentiellen Codierungen*, volume 11 of *Elektronisches Rechnen und Regeln.* Berlin: Akademie-Verlag, 1977.

[23] B. Mandelbrot. *The Fractal Geometry of Nature.* New York: Freeman, 1977.

[24] R. D. Mauldin and M. Urbański. Dimensions and measures in infinite iterated function systems. Unpublished manuscript received in december, 1993.

[25] R. D. Mauldin and S. C. Williams. Hausdorff dimension in graph directed constructions. *Transactions of the American Mathematical Society*, 309(2):811–829, Oct. 1988.

[26] W. Merzenich and L. Staiger. Fractals, dimension, and formal languages. *RAIRO Informatique théorique et Applications/Theoretical Informatics and Applications*, 28(3–4):361–386, 1994.

[27] M. Nolle. Comparison of different methods for generating fractals. To appear in the Proceedings of the IMYCS'92, 1992.

[28] H.-O. Peitgen, H. Jürgens, and D. Saupe. *Fractals for the Classroom. Part One. Introduction to Fractals and Chaos.* New York: Springer, 1992.

[29] P. Prusinkiewicz and M. Hammel. Escape-time visualization method for language-restricted iterated function systems. In Encarnação et al. [12], pages 24–44.

[30] P. Prusinkiewicz and A. Lindenmayer. *The Algorithmic Beauty of Plants.* New York: Springer, 1990.

[31] L. Staiger. Finite-state ω-languages. *Journal of Computer and System Sciences*, 27:434–448, 1983.

[32] L. Staiger. Research in the theory of ω-languages. *J. Inf. Process. Cybern. EIK (formerly Elektron. Inf.verarb. Kybern.)*, 23(8/9):415–439, 1987.

[33] L. Staiger. Ein Satz über die Entropie von Untermonoiden. *Theoretical Computer Science*, 61:279–282, 1988.

[34] L. Staiger. Quadtrees and the Hausdorff dimension of pictures. In A. Hübler et al., editors, *"Geobild'89" Proceedings of the* 4^{th} *Workshop on Geometrical Problems of Image Processing*, volume 51 of *Mathematical Research*, pages 173–178, Georgenthal, 1989. Berlin: Akademie-Verlag.

[35] L. Staiger. Hausdorff dimension of constructively specified sets and applications to image processing. In *Topology Measures, and Fractals (C. Bandt, J. Flachsmeyer and H. Haase eds.), Proceedings of the Conference on Topology and Measure VI, Warnemünde (Germany), August 1991*, volume 66 of *Mathematical Research*, pages 109–120. Berlin: Akademie-Verlag, 1992.

[36] L. Staiger. Kolmogorov complexity and Hausdorff dimension. *Information and Computation (formerly Information and Control)*, 103:159–194, 1993.

[37] L. Staiger. Codes, simplifying words, and open set condition. Technical Report 94–14, RWTH Aachen Fachgruppe Informatik, 1994.

[38] L. Staiger and K. Wagner. Automatentheoretische und automatenfreie Charakterisierungen topologischer Klassen regulärer Folgenmengen. *Elektronische Informationsverarbeitung und Kybernetik (jetzt J. Inf. Process. Cybern. EIK)*, 10(7):379–392, 1974.

[39] I. Sudborough and E. Welzl. Complexity and decidability for chain code picture languages. *Theoretical Computer Science*, 36:173–202, 1985.

[40] S. Takahashi. Self-similarity of linear cellular automata. *Journal of Computer and System Sciences*, 44:114–140, 1992.

[41] C. Tricot. Douze définitions de la densité logarithmique. *Comptes rendus des séances de l' Académie des Sciences (Paris), série I*, 293:549–552, Nov. 1981.

[42] F. v. Haeseler, H.-O. Peitgen, and G. Skordev. Linear cellular automata, substitutions, hierarchical iterated function systems and attractors. In Encarnação et al. [12], pages 3–23.

[43] B. L. van der Waerden. *Algebra II*, volume 23 of *Heidelberger Taschenbücher*. Berlin: Springer, 5 edition, 1967.

[44] K. Wicks. *Fractals and Hyperspaces*, volume 1492 of *LNM*. Berlin: Springer-Verlag, 1991.

[45] S. J. Willson. Cellular automata can generate fractals. *Discrete Applied Mathematics*, 8:91–99, 1984.