Sudan University of Science and Technology
College of Graduate Studies

Effective Estimate with Self-Similarity for Lebesgue Measure of Some Cantorval

التقدير الفعال مع المماثلة-الذاتية لأجل قياس لبيغ لبعض كانتورفال

A Thesis Submitted in Fulfillment of the Requirements for the Degree of Ph.D in Mathematics

By
Alzakey Abd Alfadel Mohammed Abd Alfadel

Supervisor
Prof. Dr. Shawgy Hussein AbdAlla

2022

Dedication

To my Family.

Acknowledgements

I would like to thank with all sincerity Allah, and my family for their supports throughout my study. Many thanks are due to my thesis guide, Prof. Dr. Shawgy Hussein AbdAlla Sudan University of Science and Technology.

Abstract

We obtain the spectrum and an effective estimate for the Lebesgue measure of the preimages of iterates of Farey and Gauss maps. The intervals and dichotomy between Farey fractions and sequence in the limit of infinite level and uniform distribution of the Stern-Brocot are determined. The structure, topology, separation and measure properties of the self-similar sets and Fractals and interated function systems of bounded distortion are characterize. We examine the asymptotic behavior of the Lebesgue measure of sum-level sets of continued function, self-similarity and nonempty interior. The family of self-affin and self-conformal sets with uniqueness, simultaneous and positive Hausdorff measure are studied. The nultigemetric. Subsum sets of sequences with recovering a purely atomic finite measure and Cantorvals with Lebesgue measure of M-Contorvals of Farey type are established.

الخلاصة

تم الحصول على الطيف والتقدير الفعال لاجل قياس لبيغ للقابلة أن تكون حدوداً لتكرارات رواسم فاري وجاوس. قمنا بتجديد الفترات والانشطار بين كسريات فادي والمتتالية في نهاية المستوى اللانهائي والتوزيع المنتظم إلى شيرن-بروكون. تم تشخيص البناء والنبولوجيا والانفصال وخصائص القياس للفئات الممتماثلة – الذاتية والكشربات وأنظمة دالة التكرار للتشوه المحدود. اختبرنا سلوك المقاربة لقياس لبيغ لفئات مستوى الجموع الدوال المستمرة والمتماثلة الناتية والداخل غير كالي. قمنا بدراسة العائلة النسبية – الناتية وفئات حافظة الذوابات الناتية مع الوحدانية وقياس هاوسروف المتزامن.

Introduction

We introduce Hilbert spaces of holomorphic functions given by generalized Borel and Laplace transforms which are left invariant by the transfer operators of the Farey map and its induced transformation, the Gauss map, respectively. The modified Farey sequence consists, at each level k, of rational fractions $r_k^{(n)}$, with $n = 1, 2, \dots, 2^k + 1$. We consider $I_k^{(e)}$, the total length of (one set of) alternate intervals between Farey fractions that are new (i.e., appear for the first time) at level k, $I_k^{(e)} := \sum_{i=1}^{2^{k-2}} \left(r_k^{(4i)} - r_k^{(4i-2)}\right)$. We show that $lim \inf_{k\to\infty} I_k^{(e)} = 0$, and conjecture that in fact $\lim_{k\to\infty} I_k^{(e)} = 0$. We employ infinite ergodic theory to show that the even SternBrocot sequence and the Farey sequence are uniformly distributed mod 1 with respect to certain canonical weightings

We investigate topological properties of a uniquely determined compact set K such that $K = \sum_{\lambda\in\Lambda} f_\lambda(K)$, where each$f_\lambda$ is a weak contraction of a complete metric space and $A = \{1, 2, "', m\}$ or $\Lambda = N$. Such a set K is said to be self-similar. Even though the open set condition (OSC) is generally accepted as the right condition to control overlaps of self-similar sets, it seems unclear how it relates to the actual size of the overlap. We study a general separation property for subsystems G, whose attractor K_G is a sub-self-similar set. This is a generalization of the Lau-Ngai weak separation property for the bounded distortion case. For subsystems with positive Hausdorff measure in its similarity dimension, we characterize the subsets of K_G with positive measure where the separation property may fail.

For a sequence $x \in l_1 \backslash c_{00}$, one can consider the achievement set $E(x)$ of all subsums of series $\sum_{n=1}^{\infty} x(n)$. It is known that $E(x)$ is one of the following types of sets: finite union of closed intervals, homeomorphic to the Cantor set, homeomorphic to the set $T$of subsums of $\sum_{n=1}^{\infty} c(n)$ where $c(2n-1) = \frac{3}{4^n}$ and $c(2n) = \frac{2}{4^n}$ (Cantorval). Given a finite subset $\Sigma \subset \mathbb{R}$ and a positive real number $q < 1$ we study topological and measuretheoretic properties of the self-similar set $K(\Sigma; q) = \sum_{n=1}^{\infty} a_n q^n (a_n)_{n\in\omega} \in 2\,\Sigma^\omega\}$, which is the unique compact solution of the equation $K = \Sigma + qK$. A sequence of real numbers converging to zero need not be summable, but it has many summable subsequences. The set of sums of all summable (infinite, finite, or empty) subsequences is a closed set of real numbers which we call the *subsum set* of the sequence. When the sequence is not absolutely summable, its subsum set is an unbounded closed interval which includes zero.

We give a detailed measure theoretical analysis of what we call sum-level sets for regular continued fraction expansions. The first main result is to settle a recent conjecture of Fiala and Kleban, which asserts that the Lebesgue measure of these level sets decays to zero, for the level tending to infinity. The second and third main results then give precise asymptotic estimates for this decay. Using techniques from infinite ergodic theory, Kesseböhmer and Stratmann determined the asymptotic behavior of the Lebesgue measure of sets of the form $F^{-n}[\alpha, \beta]$, where $[\alpha, \beta] \subseteq (0, 1]$ and F is the Farey map.

Let $\beta_1, \beta_2 > 1$ and $T_i(x, y) = ((x + i)/\beta_1, (y + i)/\beta_2), i \in \{\pm 1\}$. Let $A := A\beta_1, \beta_2$ be the unique compact set satisfying $A = T_1(A) \cup T^{-1}(A)$. We give a detailed analysis of A and the parameters (β_1, β_2) where A satisfies various topological properties. In particular, we show that if $\beta_1 < \beta_2 < 1.202$, then A has a non-empty interior, thus significantly improving the bound from Dajani et al we prove that the connectedness locus for this family studied in Solomyak is not simply connected. We investigate the Hausdorff measure and content on a class of quasi self-similar sets that include, for example, graph-directed and sub self-similar and self-conformal sets. We show that any Hausdorff measurable subset of such a set has comparable Hausdorff measure and Hausdorff content. In particular, this proves that graph-directed and sub self-conformal sets with positive Hausdorff measure are Ahlfors regular, irrespective of separation conditions. We introduce BBI spaces ("big balls of itself"), which based on the notion of BPI spaces ("big pieces of itself") used by David and Semmes to study self-similarity.

A special family of multigeometric series is considered from the point of view of behaviour of their sets of subsums. A sufficient condition for their sets of subsums to be M-Cantorvals is proven. The Lebesgue measure of those special M-Cantorvals is computed and it is shown to be equal to the sum of lengths of all component intervals of the M-Cantorvals. For μ be a purely atomic finite measure. By the range of μ we understand the set $\mathrm{rng}(\mu) = \{\mu(E) : E \subset \mathbb{N}\}$. Given a positive integer number m, we consider the M-Cantorval $K = \left\{\sum_{n=1}^{\infty} \frac{\epsilon_n}{(2m+2)^n} : (n) \in \{0, 2, 3, \ldots, 2m, 2m+1, 2m+3\}^N\right\}$. We show that this set is an attractor of iterated function system.

The Contents

Chapter 1
Spectrum with Intervals between Farey Fractions and a Dichotomy between Uniform Distributions

We are able to study simultaneously the spectrum of both these operators along with the analytic properties of associated dynamical zeta functions. This construction establishes an explicit connection between previously unrelated results of Mayer and Rugh (see [21] and [32]). The simple geometrical property of the Farey fractions turns out to be surprisingly subtle, with no apparent simple interpretation. The conjecture is equivalent to $\lim_{k\to\infty} S_k = 0$, where S_k is the sum over the inverse squares of the new denominators at level $k, S_k := \sum_{n=1}^{2^{k-1}} 1/\left(d_k^{(2n)}\right)^2$. Our result makes use of bounds for Farey fraction intervals in terms of their "parent" intervals at lower levels. We derive the precise asymptotic for the Lebesgue measure of continued fraction sum-level sets as well as connections to asymptotic behaviours of geometrically and arithmetically restricted Poincaré series. We give relations of our main results to elementary observations for the Stern-Brocot tree.

Section (1.1): Farey and Gauss Maps

The spectral analysis of transfer operators for smooth uniformly expanding maps of the unit interval [0, 1] is now fairly well understood (see [5], [2]). The spectrum depends crucially on the function space considered which is in general a Banach space. For Banach spaces of sufficiently regular functions, e.g. the space C^k of k-times differentiable functions on [0, 1] with $k \geq 0$, the transfer operator is quasi-compact. This means that its spectrum is made out of a finite or at most countable set of isolated eigenvalues with finite multiplicity (the discrete spectrum) and its complementary, the essential spectrum. The latter is a disk whose radius is a function both of k and the expanding constant ρ of the map (see e.g. [6]), in such a way that if we let $\rho \to 1$ from above (e.g. approaching an intermittency transition) the essential spectral radius tends to coincide with the spectral radius itself. In particular, in order to understand the nature of the spectrum lying under the 'essential spectrum rug' we have to consider increasingly smooth test functions as ρ approaches 1. This suggests, for instance, that for a type 1 intermittency model at the tangent bifurcation point (see [27]) one should consider suitable spaces of analytic functions. We construct a Hilbert space $\mathcal{H}_0$ of analytic functions which is left invariant by the transfer operator $\mathcal{P}$ of the Farey map, a prototype of smooth intermittent interval map, having a neutral fixed point at the origin. As a result, the spectrum of $\mathcal{P}$ when acting on $\mathcal{H}_0$ turns out to be the interval [0, 1] with embedded eigenvalues 0 and 1, plus a finite or countably infinite set of eigenvalues of finite multiplicity. The latter is conjectured to be empty. This would improve for this example a previous result obtained by Rugh in a more general framework [32]. The above and related achievements are obtained by (a slightly modified version of) an inducing procedure which was introduced for the first time in [28] (see also [30], [16], [18]) for a rather general class of intermittent interval maps. The main tool in this construction is an operator-valued function $\mathcal{Q}_z$ which enjoys simple algebraic relations both with $\mathcal{P}$ and the transfer operator $\mathcal{Q}$ of the Gauss map, the latter being obtained by inducing the Farey map with respect to the first passage time a subset of [0, 1] away from the neutral fixed point. The spectral properties of $\mathcal{Q}_z$ when acting on a Hilbert space $\mathcal{H}_1 \subset \mathcal{H}_0$ are then suitably translated into those of $\mathcal{Q}$ in $\mathcal{H}_1$ as well as $\mathcal{P}$ in $\mathcal{H}_0$. We devoted to introduce the Farey-Gauss pair, briefly discussing some (mostly known) properties of these maps and of their invariant measures and ending with a short account of their intimate connection with number theory. Further material on these general facts can be found in [4], [19], [12], [22]. The main results are contained in the two subsequent. We deal with the spectral analysis of transfer operators. We first introduce the operator valued function $\mathcal{Q}_z$ and establish simple algebraic identities. We then extend to $\mathcal{Q}_z$ some previous results of Mayer and Roepstorff (see [24], [25]) for the Gauss transfer operator $\mathcal{Q}$ obtaining as a by-product an analytic continuation of $\mathcal{Q}_z$ outside the unit disk which is crucial to exploit the above identities for spectral analysis purposes (Proposition (1.1.6)). The main results on the spectrum of $\mathcal{P}$ (Theorem (1.1.16) and Theorem (1.1.17)) are then obtained by combining these identities with an explicit integral representation of $\mathcal{P}$ on the Hilbert space $\mathcal{H}_0$ (Theorem (1.1.15)). We apply the construction to study analytic properties of the dynamical zeta functions [3] for the Farey-Gauss pair. The role of $\mathcal{Q}_z$ is here played by a two-variable zeta function $\zeta_2(s,z)$ which simply relates to the Farey and Gauss zetas and whose analytic structure is directly connected to the spectrum of $\mathcal{Q}_z$. As a result, the zeta function of the Farey map turns out to extend meromorphically to the cut plane $\mathbb{C}\backslash[1,\infty)$.

We point out that some generalized version (involving a 'temperature' parameter β) of these functions were previously studied in [21], [22], [23] for the Gauss map and in [8] for the Farey map paired with an induced version conjugated to the Gauss map 1. In the more general of piecewise analytic map with a neutral fixed point results yielding meromorphic continuation to the cut plane for zeta functions as well as regularized Fredholm determinants were obtained in [32].

We first consider the Farey map of the interval [0, 1] into itself defined as

$$F(x) = \begin{cases} F_0(x), & if\ 0 \le x \le \dfrac{1}{2}, \\ F_1(x), & if\ \dfrac{1}{2} < x \le 1\,, \end{cases} \tag{1}$$

where

$$F_0(x) := \frac{x}{1-x} \quad and\ \ F_1(x) := F_0(1-x) = \frac{1}{F_0(x)} = \frac{1-x}{x}. \tag{2}$$

The inverse branches are

$$\Psi_0(x) \equiv F_0^{-1}(x) = \frac{x}{1+x} = \frac{1}{2} - \frac{1}{2}\left(\frac{1-x}{1+x}\right),$$

$$\Psi_1(x) \equiv F_1^{-1}(x) = \frac{1}{1+x} = \frac{1}{2} + \frac{1}{2}\left(\frac{1-x}{1+x}\right). \tag{3}$$

For $x \neq 0$ the map $\Psi_0(x)$ is conjugated to the right translation $x \to S(x) = x+1$, i.e.

$$\Psi_0 = J \circ S \circ J \qquad with \qquad J(x) = J^{-1}(x) = 1/x. \tag{4}$$

This yields for the n-iterate

$$\Psi_0^n(x) = J \circ S^n \circ J(x) = \frac{x}{1+nx} \cdot \tag{5}$$

Moreover $\Psi_1(x)$ satisfies

$$\Psi_1(x) = J \circ S(x). \tag{6}$$

Let $\mathcal{A} = \{A_n\}_{n\ge 1}$ be the countable partition of [0, 1] given by $A_n = [1/(n+1), 1/n]$. Setting $A_0 = [0,1]$ it is easy to check that $F(A_n) = A_{n-1}$ for all $n \ge 1$. Let X be the residual set of points in [0, 1] which are not preimages of 1 with respect to the map F_0, namely $X = (0,1]\backslash \left\{\frac{1}{n}\right\}_{n\ge 1}$. The first passage time $\tau: X \to \mathbb{N}$ in the interval A_1 is defined as

$$\tau(x) = 1 + \min\{n \ge 0 : F^n(x) \in A_1\} = \left[\frac{1}{x}\right], \tag{7}$$

where [a] is the integer part of a. We see that A_n is the closure of the set $\{x \in X: \tau(x) = n\}$. On the other hand, the return time function $r: A_1 \to \mathbb{N} \cup \{\infty\}$ in the interval A_1 is given by

$$r(x) = \min\{n \geq 1: F^n(x) \in A_1\} = \tau \circ F(x). \tag{8}$$

We now consider the map $G: X \to X$ obtained from F by inducing w.r.t. the first passage time τ, i.e.

$$G(x) = F^{\tau(x)}(x), \tag{9}$$

which can be extended to all of [0, 1] setting $G(0) = 1, G(1) = 0,$

$$\lim_{x\uparrow\frac{1}{n}} G(x) = 0, \lim_{x\downarrow\frac{1}{n}} G(x) = 1, \qquad n > 1,$$

and whenever $x \in \left(\frac{1}{n+1}, \frac{1}{n}\right)$ we have, using (5),

$$G(x) \equiv G_n(x) = F^n(x) = F_1 \circ F_0^{n-1}(x) = \frac{1}{x} - n = \frac{1}{x} - \tau(x). \tag{10}$$

In other words the induced map is the celebrated Gauss map

$$G(x) = \begin{cases} \left\{\frac{1}{x}\right\}, & if\ x \neq 0\,, \\ 0, & if\ x = 0\,, \end{cases} \tag{11}$$

where $\{a\}$ denotes the fractional part of a. It has countably many inverse branches Φ_n given by

$$\Phi_n(x) = G_n^{-1}(x) = \frac{1}{x+n}, \qquad n \geq 1. \tag{12}$$

It is an easy task to verify that the σ-finite absolutely continuous measure

$$\nu(dx) \equiv e(x)dx = \frac{1}{\log 2} \cdot \frac{dx}{x} \tag{13}$$

is invariant for the dynamical system $([0,1], F)$. Note that $\nu(A_n) = (\log 2)^{-1} \log\left(1 + \frac{1}{n}\right)$ and $\nu([0,1]) = \infty$. Let $B_n = \{x \in A_1: r(x) = n\}$. Using (8) we have $\overline{F_1(B_n)} = A_n$. We now show that $\nu(A_n) = \sum_{k\geq n} \nu(B_k)$. Indeed, for $n = 1$ we have $\sum_{k\geq 1} \nu(B_k) = \nu(A_1) = 1$. Moreover, since ν is F-invariant, $\nu(A_n) = \nu\left(F^{-1}(A_n)\right) = \nu(A_{n+1}) + \nu(B_{n+1})$, and the assertion follows by induction. Therefore the expected return time is infinite:

$$\nu_{A_1}(r) = \int_{A_1} r(x)\nu(dx) = \sum_{n\geq 1} n\,\nu(B_n) = \sum_{n\geq 1} \nu(A_n) = \nu([0,1]) = \infty, \tag{14}$$

where ν_{A_1} is the conditional probability measure defined as $\nu_{A_1}(E) = \nu(E \cap A_1)/\nu(A_1)$. It is known that in this situation there is the coexistence of two different statistics for the dynamical system $(F, [0,1])$: besides ν, the ergodic means $\frac{1}{n}\sum_{i=0}^{n-1} \delta_{F^i}(x)$ converge weakly to the Dirac delta at 0 (see [26], [17]). Let ρ be the probability measure obtained by pushing forward ν with F_1, i.e.

$$\rho(E) = ((F_1)_* \nu)(E) = (\nu \circ \Psi_1)(E). \tag{15}$$

Reasoning as above one readily verifies that the converse relation is

$$\nu(E) = \sum_{n\geq 0} (\rho \circ \Psi_0^n)(E). \tag{16}$$

In particular we have $\nu(A_n) = \sum_{l\geq n} \rho(A_l)$ and $\rho(A_n) = \rho\left(F_1(B_n)\right) = \nu(B_n)$, where B_n is as above. We then have

$$\rho(E) = (\nu \circ \Psi_1)(E) = \sum_{n\geq 0} (\rho \circ \Psi_0^n \circ \Psi_1)(E) = \rho(G^{-1}E), \tag{17}$$

which says that ρ is G-invariant. Moreover ρ is ergodic with respect to G (see e.g. [4]). Setting $h(x) = \rho(dx)/dx$ we get

$$h = |\Psi_1'| \cdot e \circ \Psi_1, \qquad e = \sum_{k=0}^{\infty} \left(\Psi_0^k\right)' \cdot h \circ \psi_0^k, \tag{18}$$

which gives the well known result

$$h(x) = \frac{1}{\log 2} \cdot \frac{dx}{(1+x)} \cdot \tag{19}$$

The primitive $H(x)$ of $h(x)$, with $H(0) = 0$, is $H(x) = \log(1+x) / \log 2$. Setting $q_n := H\left(\frac{1}{n+1}\right) = (\log 2)^{-1} \log\left(1 + \frac{1}{n+1}\right)$, we have $\nu(A_n) = q_n$ and $\rho(A_n) = q_{n-1} - q_n$. We see that qn is a (strict) Kaluza sequence, i.e. for all $n \geq 1$

$$1 = q_0 > q_1 \dots > q_n > 0 \qquad \text{and} \qquad q_n^2 < q_{n-1} q_{n+1}. \tag{20}$$

Finally, by (7), (8), (14) and (15) we have

$$\rho(\tau) = ((F_1)_* \nu)(\tau) = \nu(\tau \circ F_1) = \nu(r) = \infty. \tag{21}$$

On the other hand we have the following,

Lemma (1.1.1)[1]: The function log τ is in $L_1(\rho)$ and satisfies

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log \tau\left(G^j(x)\right) = \rho(\log(\tau) = K, \qquad \rho - a.e., \tag{22}$$

where the positive constant K is defined by

$$e^K = \prod_{k=1}^{\infty} \left(1 + \frac{1}{k(k+2)}\right)^{\frac{\log k}{\log 2}}. \tag{23}$$

Proof. We have

$$\begin{aligned}\rho(\log \tau) &= \sum_{k=1}^{\infty} \rho(A_k) \cdot \log k = \sum_{k=1}^{\infty} (q_{k-1} - q_k) \cdot \log k \\ &= \sum_{k=1}^{\infty} \frac{\log k}{\log 2} \cdot \log\left(\left(1 + \frac{1}{k}\right)\left(1 + \frac{1}{k+1}\right)^{-1}\right) \\ &= \sum_{k=1}^{\infty} \frac{\log k}{\log 2} \cdot \log\left(1 + \frac{1}{k(k+2)}\right) = K < \infty.\end{aligned}$$

This computation shows both that $\log \tau \in L_1(\rho)$ and the last equality in (22). The first equality in (22) now follows from the ergodic theorem [4].

The constant K which appears above is known in number theory as Khinchin's constant. This is not a coincidence, as we now briefly explain.

The Farey sum over two rationals $\frac{a}{b}$ and $\frac{a'}{b'}$ is the mediant operation given by [15]

$$\frac{a''}{b''} = \frac{a + a'}{b + b'} \cdot \tag{24}$$

It is easy to see that $\frac{a''}{b''}$ falls in the interval $\left(\frac{a}{b}, \frac{a'}{b'}\right)$. Now, having fixed $n \geq 0$, let $\mathcal{F}_n$ be the ascending sequence of irreducible fractions between 0 and 1 obtained inductively in the following way. Set first $\mathcal{F}_0 = \left(\frac{0}{1}, \frac{1}{1}\right)$. Then $\mathcal{F}_n$ is obtained from $\mathcal{F}_{n-1}$ by inserting among each pair of consecutive rationals $\frac{a}{b}$ and $\frac{a'}{b'}$ in $\mathcal{F}_{n-1}$ their mediant $\frac{a''}{b''}$ as above. Thus $\mathcal{F}_1 =$

$\left(\frac{0}{1},\frac{1}{2},\frac{1}{1}\right)$, $\mathcal{F}_2 = \left(\frac{0}{1},\frac{1}{3},\frac{1}{2},\frac{2}{3},\frac{1}{1}\right)$, $\mathcal{F}_3 = \left(\frac{0}{1},\frac{1}{4},\frac{1}{3},\frac{2}{5},\frac{1}{2},\frac{3}{5},\frac{2}{3},\frac{3}{4},\frac{1}{1}\right)$ and so on. The elements of $\mathcal{F}_n$ are called Farey fractions. The name of the map F can be related to the easily verified observation that the set of pre-images $\cup_{k=0}^{n+1} F^{-k}\{0\}$ coincides with $\mathcal{F}_n$ for all $n \geq 0$. In particular, this implies that $\cup_{k=0}^{\infty} F^{-k}\{0\} = \mathbb{Q} \cap [0,1]$ (notice that the same is true for the induced map: $\cup_{k=0}^{\infty} G^{-k}\{0\} = \mathbb{Q} \cap [0,1]$) .

We recall that every real number $0 < x < 1$ has a continued fraction expansion of the form [19]

$$x = \cfrac{1}{k_1 + \cfrac{1}{k_2 + \cfrac{1}{k_3 \ldots}}} = [k_1, k_2, k_3, \ldots], \tag{25}$$

with $k_i \in \mathbb{N}$. By applying Euclid's algorithm one sees that the above expansion terminates if and only if x is a rational number. There is an intimate connection between the partial quotients $k_1, k_2, \ldots$ and the Gauss map G. Indeed, given x as above we can write

$$x = \frac{1}{\frac{1}{x}} = \frac{1}{\left[\frac{1}{x}\right] + \left\{\frac{1}{x}\right\}} = \frac{1}{k_1 + G(x)} = \cfrac{1}{k_1 + \cfrac{1}{\cfrac{1}{G(x)}}} = \cfrac{1}{k_1 + \cfrac{1}{\left[\frac{1}{G(x)}\right] + \left\{\frac{1}{G(x)}\right\}}}$$

$$= \cfrac{1}{k_1 + \cfrac{1}{k_2 + G^2(x)}} = \cdots \tag{26}$$

Therefore, $k_1 = [1/x], k_2 = [1/G(x)], k_3 = [1/G^2(x)]$ and so on. Alternatively,

$$if \quad x = [k_1, k_2, k_3, \ldots] \quad \text{then} \quad G(x) = [k_2, k_3, \ldots]. \tag{27}$$

Farey fractions have close relationships with continued fractions. Let us say that a Farey fraction has order n if it belongs to $\mathcal{F}_n \backslash \mathcal{F}_{n-1}$. Given $n \geq 1$ there are exactly 2^{n-1} Farey fractions of order n (they form the set $F^{-(n+1)}\{0\}$) and it is possible to show (see below eq. (28)) that the integers k_i in their (finite) continued fraction expansion sum up to $n+1$. Furthermore, it is easy to realize that all Farey fractions which fall in the interval $(1/(n+1), 1/n)$ have order greater than or equal to $n+1$, whereas their continued fraction expansion starts with $k_1 = n$. Thus, the map F acts on Farey fractions by reducing their order of one unit. We can write an explicit expression for the action of F on continued fraction expansions. Indeed, if $1/2 < x \leq 1$ then $k_1 = 1$ and $F(x) = \frac{1}{x} - k_1 = G(x)$. If instead $0 < x \leq 1/2$ then $k_1 > 1$ and $F(x) = 1/(\frac{1}{x} - 1)$. Therefore,

$$\text{if} \quad x = [k_1, k_2, k_3, \ldots] \quad \text{then} \quad F(x) = [k_1 - 1, k_2, k_3, \ldots], \tag{28}$$

with $[0, k_2, k_3, \ldots] \equiv [k_2, k_3, \ldots]$ (compare to (27)). Now, it is well known that for almost all $x \in (0,1)$ the arithmetic mean of the partial quotients is infinite (see, e.g., [19]), i.e.

$$\lim_{n\to\infty} \frac{k_1 + \cdots + k_n}{n} = \infty, \quad (a.e.) \tag{29}$$

From the above discussion and (7) we get $k_l = \left[\frac{1}{G^{l-1}(x)}\right] = \tau\left(G^{l-1}(x)\right)$, which for $l > 1$ is the time between the $(l-1)$-st and the l-th passage in A_1 of the orbit of x with F. Therefore, the total number S_n of iterates of F needed to observe n passages in A_1, that is the function

$$S_n(x) = \tau(x) + \tau(G(x)) \cdots + \tau\left(G^{n-1}(x)\right), \tag{30}$$

satisfies

$$\lim_{n\to\infty} \frac{S_n(x)}{n} = \infty \qquad (a.e.) \tag{31}$$

Since ρ is absolutely continuous w.r.t. the Lebesgue measure on [0, 1], the properties expressed by (21) and (31) can be regarded as an instance of validity of the ergodic theorem for the non-integrable function τ. As a consequence of ([19], Theorem 30) we have that for almost all $x \in (0,1)$ the inequality

$$S_n(x) \geq n \log n \tag{32}$$

is satisfied for an infinite number of values of n. On the other hand, Lemma (1.1.1) can now be rephrased by saying that the geometric mean of the partial quotients has a certain finite value (a.e.). This, in turn, is a corollary of a theorem of Khinchin ([19], Theorem 35), which says that for any function $f(k)$ defined on the positive integers and satisfying $f(k) = \mathcal{O}(k^p)$ with $0 \leq p < \frac{1}{2}$ we have, for almost all $x \in (0,1)$,

$$\left| \frac{1}{n}\sum_{i=j}^{n} f(k_j) - \sum_{k=1}^{\infty} \frac{f(k)}{\log 2} \cdot \log\left(1 + \frac{1}{k(k+2)}\right) \right| \leq \epsilon(n) \tag{33}$$

where the error function $\epsilon(n)$ is any positive function decreasing to zero as n→∞ so that $\sum n^{-2} \cdot \epsilon^{-2}(n) < \infty$. Lemma (1.1.1) then corresponds to the choice $f(k) = \log k$.

We start by establishing some formal algebraic relations between the transfer operators $\mathcal{P}$ and $\mathcal{M}$ associated to the maps F and G, respectively (see [2]). They describe the action of the differentiable dynamical systems F and G on the density f of a measure absolutely continuous measure wrt Lebesgue by

$$\mathcal{P}f(x) = (\mathcal{P}_0 + \mathcal{P}_1)f(x) =: |\Psi_0'(x)| \cdot f(\Psi_0(x)) + |\Psi_1'(x)| \cdot f(\Psi_1(x))$$
$$= \left(\frac{1}{x+1}\right)^2 \left[f\left(\frac{x}{x+1}\right) + f\left(\frac{1}{x+1}\right)\right], \tag{34}$$

and

$$\mathcal{Q}f(x) = \sum_{n=1}^{\infty} Q_n u(x) =: \sum_{n=1}^{\infty} |\Phi_n'(x)| \cdot f(\Phi_n(x))$$
$$= \sum_{n=1}^{\infty} \left(\frac{1}{x+n}\right)^2 f\left(\frac{1}{x+n}\right). \tag{35}$$

We first notice that

$$Q_n f(x) = \mathcal{P}^n(f \cdot \chi_n)(x) = \mathcal{P}_1 \mathcal{P}_0^{n-1} f(x), \tag{36}$$

where χ_n is the indicator function of A_n. Let $Sf(x) := f \circ S(x) = f(x+1)$ be the shift operator. Note by (4) and (6) we have

$$\mathcal{P}_1 \mathcal{P}_0 f(x) = S\mathcal{P}_1 f(x), \tag{37}$$

and therefore (36) yields

$$Q_n f(x) = \mathcal{P}_1 \mathcal{P}_0^{n-1} f(x) = S^{n-1} \mathcal{P}_1 f(x). \tag{38}$$

More generally, for $z \in \mathbb{C}$, we shall consider a formal operator-valued power series $\mathcal{Q}_z$ defined by

$$\mathcal{Q}_z f(x) = \sum_{n=1}^{\infty} z^{\tau(\Phi_n(x))} \cdot |\Phi_n'(x)| \cdot f(\Phi_n(x)) = z\mathcal{P}_1(1 - z\mathcal{P}_0)^{-1} f(x) \tag{39}$$

so that $\mathcal{Q}_1 \equiv \mathcal{Q}$. The following operator relations are in force and are independent of the function space the operators are acting on.

Proposition (1.1.2)[1]: Let $z \in \mathbb{C}$ be such that (39) is absolutely convergent. Then we have

$$(1-\mathcal{Q}_z)(1-z\mathcal{P}_0)=1-z\mathcal{P} \tag{40}$$
and
$$(1-z\,S)(1-\mathcal{Q}_z)=1-z\tilde{\mathcal{P}}\,. \tag{41}$$
where $\tilde{\mathcal{P}}=S+\mathcal{P}_1$see [28], [30], [8], [16], [18].

Proof. Using the first identity in (38) we get
$$(1-\mathcal{Q}_z)(1-z\mathcal{P}_0)=\left(1-\sum_{n=1}^{\infty} z^n\mathcal{P}_1\mathcal{P}_0^{n-1}\right)(1-z\mathcal{P}_0)$$
$$=1-z\mathcal{P}_0-\sum_{n=1}^{\infty} z^n\mathcal{P}_1\mathcal{P}_0^{n-1}+\sum_{n=1}^{\infty} z^{n+1}\mathcal{P}_1\mathcal{P}_0^{n}=1-z\mathcal{P}_0-z\mathcal{P}_1=1-z\mathcal{P}.$$
In a similar way, using the second identity in (38) one shows (41).

Corollary (1.1.3)[1]: Let $z\neq 0$ be such that (39) is absolutely convergent and assume that the kernel of $1-z\mathcal{P}_0$ is empty. Then 1 is an eigenvalue of $\mathcal{Q}_z$ if and only if z^{-1} is an eigenvalue both of $\mathcal{P}$ and $\tilde{P}$, and they have the same geometric multiplicity. Furthermore, the corresponding eigenfunctions e_z of $\mathcal{P}$ and h_z of $\tilde{P}$ and Q_z are related by $h_z=(1-z\mathcal{P}_0)e_z$ or else $e_z=\sum_{k=0}^{\infty} z^k\mathcal{P}_0^k h_z$.

Proof. Assume that $\mathcal{Q}_z h_z=h_z$. From (40) it then follows that $(1-z\mathcal{P})\sum_{k=0}^{\infty} z^k\mathcal{P}_0^k h_z=0$. Conversely, assume that $z\mathcal{P}e_z=e_z$, then we have $(1-\mathcal{Q}_z)(1-z\mathcal{P}_0)e_z=0$. In the same way, from (41) it follows that $\mathcal{Q}_z h_z=h_z$ if and only if $\tilde{\mathcal{P}}h_z=z^{-1}h_z$.

Having fixed an open connected domain $\Omega\subset\mathbb{C}$ let $\mathcal{H}(\Omega)$ be the Fr´echet space of functions which are holomorphic in Ω with the topology generated by the family of sup norms on compact subsets of Ω. Moreover, we let $A_\infty(\Omega)\subset\mathcal{H}(\Omega)$ denote the Banach space given by the subset of functions in $\mathcal{H}(\Omega)$ having continuous extension to $\overline{\Omega}$, endowed with the norm
$$\|f\|=\sup_{w\in\overline{\Omega}}\ |f(w)|, \tag{42}$$
(where $w=x+iy$). We let first $\mathcal{Q}_z$ act on the Banach space $A_\infty(D)$ with $D=\{w\in\mathbb{C}:|w-1|<1\}$. It is easy to verify that $\Phi_n(\overline{D})\subset D$ for all $n\in\mathbb{N}$. Standard arguments (see [22]) then imply that whenever the power series in (39) is uniformly convergent $\mathcal{Q}_z$ defines a nuclear operator of order zero on $A_\infty(D)$.

Lemma (1.1.4)[1]: The power series of $\mathcal{Q}_z: A_\infty(D)\to A_\infty(D)$ has radius of convergence bounded from below by 1 and, moreover, it converges absolutely at every point of the unit circle.

Proof. The radius of convergence of $\mathcal{Q}_z$ is $\lim_{n\to\infty}\ \|\mathcal{Q}_n\|^{-\frac{1}{n}}$ (here $\|\ \|$ denotes the operator norm as well). We have $\sup_{w\in\overline{D}}\ |\mathcal{Q}_n f(w)|\leq C\,n^{-2}\|f\|$ and therefore $\|\mathcal{Q}_n\|\leq Cn^{-2}$.

We now introduce a subspace of $A_\infty(D)$ on which the action of $\mathcal{Q}_z$ will turn out to be particularly expressive. This is achieved via a generalized Laplace transform.

Definition (1.1.5)[1]: We let $\mathcal{H}_1$ denote the Hilbert space of all complex-valued functions f which have a representation as generalized Laplace transform
$$f(w)=(\mathcal{L}[\varphi])(w):=\int_0^{\infty} e^{-tw}\varphi(t)dm(t) \tag{43}$$
where $\varphi\in L_2(m)$ and dm is the measure on ir^+ given by
$$dm(t)=\frac{t}{e^t-1}dt. \tag{44}$$
As a Hilbert space $\mathcal{H}_1$ is endowed with the inner product

$$(f_1, f_2) = \int_0^\infty \overline{\varphi_1(t)}\varphi_2(t)dm(t) \quad \text{if} \quad f_i = \mathcal{L}[\varphi_i]. \tag{45}$$

The following Proposition generalizes corresponding results obtained by Mayer and Roepstorff (see [24], [25]) for the operator $\mathcal{Q}$.

Proposition (1.1.6)[1]: For each $z \neq 0$ with $|z| \leq 1$, the space $\mathcal{H}_1$ is invariant under $\mathcal{Q}_z$. More precisely we have

$$\mathcal{Q}_z\mathcal{L}[\varphi] = \mathcal{L}[z(1-M)(1-zM)^{-1}\mathcal{K}\varphi], \tag{46}$$

where $M: L_2(m) \to L_2(m)$ is the multiplication operator

$$M\varphi(t) = e^{-t}\varphi(t) \tag{47}$$

and $\mathcal{K}: L_2(m) \to L_2(m)$ is the integral operator

$$(\mathcal{K}\varphi)(t) = \int_0^\infty \frac{J_1(2\sqrt{st})}{\sqrt{st}}\varphi(s)dm(s) \tag{48}$$

and J_p denotes the Bessel function of order p.

Proof. Letting $f = \mathcal{L}[\varphi]$ we have from (39) and (38)

$$\mathcal{Q}_z f(w) = \sum_{n=1}^{\infty} \frac{z^n}{(w+n)^2}\int_0^\infty dm(t)e^{-\frac{t}{w+n}}\varphi(t). \tag{49}$$

Clearly, for $|z| \leq 1$, the sum $\sum_{n=1}^{\infty} \frac{z^n}{(w+n)^2}e^{-\frac{t}{w+n}}$ is uniformly convergent in $t \in ir^+$. Therefore, interchanging summation and integration we get

$$\sum_{n=1}^{\infty} \frac{z^n}{(w+n)^2}e^{-\frac{t}{w+n}} = \sum_{k\geq 0}\frac{(-t)^k}{k!}\sum_{n=1}^{\infty}\frac{z^n}{(w+n)^{2+k}}$$
$$= \sum_{k\geq 0}\frac{(-t)^k}{k!}z\Phi(z, k+2, w+1) \tag{50}$$

where $\Phi(z,a,b) = \sum_{n=0}^{\infty} \frac{z^n}{(b+n)^a}$ is the Lerch transcendental function which, for $\Re a > 1$, possesses the integral representation

$$\Phi(z,a,b) = \sum_{n=0}^{\infty}\frac{z^n}{(b+n)^a} = \frac{1}{\Gamma(a)}\int_0^\infty \frac{s^{a-1}e^{-(b-1)s}}{e^s - z}ds. \tag{51}$$

This yields

$$\Phi(z, k+2, w+1) = \frac{1}{(k+1)!}\int_0^\infty \frac{s^{k+1}e^{-ws}}{e^s - z}ds. \tag{52}$$

Noting that

$$\sum_{k\geq 0}\frac{(-st)^k}{(k+1)!\,k!} = \frac{J_1(2\sqrt{st})}{\sqrt{st}} \tag{53}$$

where $J_1(x)$ is the Bessel function of the first kind, we have thus found that

$$\mathcal{Q}_z f(w) = \int_0^\infty ds\frac{zs}{e^s - z}e^{-ws}\int_0^\infty dm(t)\frac{J_1(2\sqrt{st})}{\sqrt{st}}\varphi(t)$$
$$= \int_0^\infty dm(s)e^{-ws}(z(1-M)(1-zM)^{-1}\mathcal{K}\varphi)(s)$$
$$= (\mathcal{L}[z(1-M)(1-zM)^{-1}\mathcal{K}\varphi])(w). \tag{54}$$

Notice that for each $t \in ir^+$ the function $J_1(2\sqrt{st})/\sqrt{st}$ is uniformly bounded and continuous for $s \in ir^+$. It is then an easy task to verify that for $\varphi \in L_2(m)$ and for $|z| \le 1$ the function $(1-M)(1-zM)^{-1}\mathcal{K}\varphi$ is in $L_2(m)$ as well.

Indeed, the operator $(1-zM)$ is invertible in $L_2(m)$ with bounded inverse provided $\frac{1}{z} \notin [0,1]$. Therefore, for any $\varphi \in L_2(m)$ the integral in (46) converges uniformly in any compact region of the complex z-plane not containing points of the ray $(1,+\infty)$. Moreover, it has been proved in [24] that the operator $\mathcal{K}$ is compact (actually trace-class) in $L_2(m)$. Therefore, as long as $(1-zM)$ has bounded inverse the operator $(1-M)(1-zM)^{-1}\mathcal{K}$ is compact as well (being the composition of a compact operator with a bounded operator). Proposition (1.1.6) and the above observations prove the following result,

Theorem (1.1.7)[1]: The operator-valued function $z \to \mathcal{Q}_z$, when acting on $\mathcal{H}_1$, can be analytically continued to the entire z-plane with a cross cut along the ray $(1,+\infty)$, and for each z in this domain is isomorphic to the operator

$$\mathcal{K}_z := z\,(1-M)(1-zM)^{-1}\mathcal{K} \tag{55}$$

acting on $L_2(m)$. They are both compact operators.

Remark (1.1.8)[1]: Putting

$$H_\delta := \{w \in \mathbb{C} : \Re w > \delta\} \tag{56}$$

one sees that a function $f = \mathcal{L}[\varphi]$ with $\varphi \in L_2(m)$ can be extended to a function holomorphic in the half-plane $H_{-\frac{1}{2}}$.

If, in addition, f is an eigenfunction corresponding to a non-zero eigenvalue λ of $\mathcal{Q}_z$ in $\mathcal{H}_1$, for some non-zero $z \in \mathbb{C}\backslash(1,\infty)$, then

$$\lambda\,\varphi(t) = (\mathcal{K}_z\varphi)(t) = \left(\frac{1-e^{-t}}{\frac{1}{z}-e^{-t}}\right)\int_0^\infty \frac{J_1(2\sqrt{st})}{\sqrt{st}}\varphi(s)dm(s). \tag{57}$$

Since the integral in the r.h.s. is bounded for all $t \in [0,\infty)$ the function $\varphi(t)$ is bounded as well is this domain and therefore f is holomorphic in the half plane H_{-1}.

Putting together the above, Proposition (1.1.2) along with standard arguments (see [9]) we get,

Corollary (1.1.9)[1]: The operator-valued function $\to (1-\mathcal{Q}_z)^{-1}$, when acting on $\mathcal{H}_1$, is analytic in the open unit disk $\{z : |z| < 1\}$ and can be meromorphically continued to the entire z-plane with a cross cut along the ray $[1,+\infty)$. It has a pole whenever $\mathcal{K}_z$ has 1 as an eigenvalue.

Now, from Proposition (1.1.2) we obtain the following formal relation for the resolvent $\mathcal{R}_\lambda$ of $\mathcal{P}$:

$$\mathcal{R}_\lambda \equiv (\lambda-\mathcal{P})^{-1} = (\lambda-\mathcal{P}_0)^{-1}\left(1-\mathcal{Q}_{\frac{1}{\lambda}}\right)^{-1}. \tag{58}$$

The analytic properties of the first factor in the r.h.s. can be understood in terms of the spectrum of the operator $\mathcal{P}_0$ when acting on a suitable function space invariant under the action of $\mathcal{P}$. A calculation along the same lines as in the proof of Proposition (1.1.6) shows that, for $f \in \mathcal{H}_1$ with $f = \mathcal{L}[\varphi]$,

$$(1-z\mathcal{P}_0)^{-1}f(w) = \frac{1}{w^2}\int_0^\infty e^{-\frac{t}{w}}e^t z^{-1}(\mathcal{K}_z\varphi)(t)dm(t). \tag{59}$$

We shall therefore characterize the space $\mathcal{H}_0$ to be acted on by $\mathcal{P}$ as follows:

Definition (1.1.10)[1]: We denote by $\mathcal{H}_0$ the Hilbert space of all complex-valued functions f which can be represented as a generalized Borel transform

$$f(w) = (\mathcal{B}[\varphi])(w) := \frac{1}{w^2}\int_0^\infty e^{-\frac{t}{w}} e^t \varphi(t)\,dm(t), \qquad \varphi \in L_2(m), \tag{60}$$

$$(f_1, f_2) = \int_0^\infty \overline{\varphi_1(t)}\varphi_2(t)\,dm(t) \quad \text{if} \quad f_i = \mathcal{B}[\varphi_i]. \tag{61}$$

Remark (1.1.11)[1]: A function $f \in \mathcal{H}_0$ is holomorphic in the disk

$$D_1 = \left\{w \in \mathbb{C}: \Re\frac{1}{w} > \frac{1}{2}\right\} = \{w \in \mathbb{C}: |w - 1| < 1\}. \tag{62}$$

For w real and positive a simple change of variable makes (60) in the form

$$f(w) = \frac{1}{w}\int_0^\infty e^{-s}\psi(sw)\,ds \quad \text{with} \quad \psi(t) = \left(\frac{t}{1 - e^{-t}}\right)\varphi(t). \tag{63}$$

Remark (1.1.12)[1]: The F-invariant density e (see (13)) can be represented as

$$e = \left(\frac{1}{\log 2}\right)\mathcal{B}\left[\frac{1 - e^{-t}}{t}\right], \tag{64}$$

whereas for the G-invariant density h we have

$$h = \left(\frac{1}{\log 2}\right)\mathcal{L}\left[\frac{1 - e^{-t}}{t}\right] = \left(\frac{1}{\log 2}\right)\mathcal{B}\left[\frac{(1 - e^{-t})^2}{t}\right]. \tag{65}$$

In the representation of Remark (1.1.11) we have that if $f = e \cdot \log 2$ then $\psi(t) \equiv 1$ whereas for $f = h \cdot \log 2$ we find $(t) = 1 - e^{-t}$. Both these functions can be viewed as ordinary Borel transforms of a sequence $\{a_n\}_{n=0}^\infty$, i.e. $\psi(t) = \sum_{n=0}^\infty t^n a_n/n!$ so that by (63) we have $w \cdot f(w) = \sum_{n=0}^\infty w^n a_n$. In the former case we find $a_0 = 1$ and $a_n = 0$ for $n > 0$, in the latter $a_0 = 0$ and $a_n = (-1)^{n-1}$ for $n > 0$. Therefore in both cases the integral (63) provides a continuation of $w \cdot f(w)$ outside the disk D_1 (see [34]).

We now have the following,

Lemma (1.1.13)[1]: For all $\varphi \in L_2(m)$

$$\mathcal{L}[\varphi] = \mathcal{B}[(1 - M)\mathcal{K}\varphi] \tag{66}$$

where $M\varphi(t) = e^{-t}\varphi(t)$ and $\mathcal{K}$ is the symmetric integral operator defined in (48).

Proof. The proof is an easy calculation based on Tricomi's theorem (see [33])

$$\frac{1}{u^{p+1}}\int_0^\infty dt\, e^{-\frac{t}{u}}\varphi(t) = \int_0^\infty dt e^{-tu}\int_0^\infty ds\left(\frac{t}{s}\right)^{\frac{p}{2}} J_p\left(2\sqrt{st}\right)\varphi(s), \tag{67}$$

with $p = 1$, and therefore we omit it.

It is now not difficult to verify that

$$\mathcal{P}_1\mathcal{B}[\varphi] = \mathcal{L}[\varphi], \tag{68}$$

and

$$\mathcal{P}_0\mathcal{B}[\varphi] = \mathcal{B}[M\varphi]. \tag{69}$$

In addition we have

$$S\mathcal{L}[\varphi] = \mathcal{L}[M\varphi], \tag{70}$$

so that

$$\mathcal{P}_1\mathcal{P}_0^{n-1}\mathcal{B}[\varphi] = S^{n-1}\mathcal{P}_1\mathcal{B}[\varphi] = \mathcal{L}[M^{n-1}\varphi], \tag{71}$$

and therefore

$$\mathcal{Q}_z\mathcal{B}[\varphi] = z \cdot \mathcal{L}[(1 - zM)^{-1}\varphi]. \tag{72}$$

We thus see that $\mathcal{P}_0$ leaves $\mathcal{H}_0$ invariant and by (70) its spectral properties in $\mathcal{H}_0$ are identical to those of S in $\mathcal{H}_1$. Moreover $\mathcal{P}_1$ maps $\mathcal{H}_0$ into $\mathcal{H}_1 \subset \mathcal{H}_0$, and the same does $\mathcal{Q}_z$ for all $z \in \mathbb{C}\backslash(1, +\infty)$. Notice that using Lemma (1.1.13) and (72) we immediately recover Proposition (1.1.6), in that

$$\mathcal{Q}_z\mathcal{L}[\varphi] = \mathcal{Q}_z\mathcal{B}[(1 - M)\mathcal{K}\varphi] = \mathcal{L}[\, z \cdot (1 - zM)^{-1}(1 - M)\mathcal{K}\varphi] \equiv \mathcal{L}[\mathcal{K}_z\varphi]. \tag{73}$$

We are now in the position to write explicit representations for $\mathcal{P}$ and its resolvent $\mathcal{R}_\lambda$ in the space $\mathcal{H}_0$.

Remark (1.1.14)[1]: Note that for $\varphi \in L_2(m)$ the functions

$$M\,\varphi \quad \text{and} \quad (1-M)\mathcal{K}\varphi \tag{76}$$

are bounded at infinity and therefore, by (74), the function $\mathcal{P}_f$ with $f = \mathcal{B}[\varphi]$ is analytic in the half-plane $\mathcal{H}_0$. In particular so is any eigenfunction of $\mathcal{P}$ in $\mathcal{H}_0$.

Theorem (1.1.15)[1]: Let $f \in \mathcal{H}_0$, that is $f = \mathcal{B}[\varphi]$ for some $\varphi \in L_2(m)$, then

$$\mathcal{P}f = \mathcal{B}[(M + (1-M)\mathcal{K})\varphi], \tag{74}$$

and

$$\mathcal{R}_\lambda f \equiv (\lambda - P)^{-1} f = \mathcal{B}\left[\left(1 - \mathcal{K}_{\frac{1}{\lambda}}\right)^{-1} (\lambda - M)^{-1}\varphi\right]. \tag{75}$$

Proof. From (69) and (68) one obtains $\mathcal{P}f = \mathcal{B}[M\varphi] + \mathcal{L}[\varphi]$, so that (74) follows using Lemma (1.1.13). The expression for $\mathcal{R}_\lambda$ can now be obtained directly from (74). But we can also make use of (72) and (54) to obtain, for a given $f = \mathcal{B}[\varphi]$,

$$\mathcal{Q}^n_{\frac{1}{\lambda}} f = \mathcal{L}\left[\mathcal{K}^{n-1}_{\frac{1}{\lambda}} (\lambda - M)^{-1}\varphi\right] \tag{77}$$

and therefore

$$\left(1 - \mathcal{Q}_{\frac{1}{\lambda}}\right)^{-1} f = \mathcal{B}[\varphi] + \mathcal{L}\left[\left(1 - \mathcal{K}_{\frac{1}{\lambda}}\right)^{-1} (\lambda - M)^{-1}\varphi\right]. \tag{78}$$

This expression along with (58), (59) and (69) yield

$$\backslash scritR_\lambda f = \mathcal{B}[(\lambda - M)^{-1}\varphi] + \mathcal{B}\left[\mathcal{K}_{\frac{1}{\lambda}}\left(1 - \mathcal{K}_{\frac{1}{\lambda}}\right)^{-1} (\lambda - M)^{-1}\varphi\right]$$
$$= \mathcal{B}\left[\left(1 - K_{\frac{1}{\lambda}}\right)^{-1} (\lambda - M)^{-1}\varphi\right].$$

Using Corollary (1.1.9) we see that $\mathcal{R}_\lambda$ extends to a meromorphic (operator-valued) function in $\bar{\mathbb{C}}\backslash[0,1]$.

The next theorem (partially) describes the spectrum of $\mathcal{P}$ in $\mathcal{H}_0$.

Theorem (1.1.16)[1]: The spectrum of the operator $P: \mathcal{H}_0 \to \mathcal{H}_0$ is the union of [0, 1] and a finite or countably infinite set of eigenvalues of finite multiplicity.

Proof. By Theorem (1.1.15) the action of transfer operator $\mathcal{P}$ on $\mathcal{H}_0$ can be explicitly expressed in the form

$$\mathcal{P}\mathcal{B}[\varphi] = \mathcal{B}[T\varphi], \tag{79}$$

with

$$(T\varphi)(t) := e^{-t}\varphi(t) + \int_0^\infty K(s,t)\varphi(s)ds \tag{80}$$

and

$$K(s,t) = e^{-t}\left(\frac{e^t - 1}{e^s - 1}\right)\sqrt{\frac{s}{t}}\, J_1\!\left(2\sqrt{st}\right). \tag{81}$$

We check that M when acting upon $L_2(m)$ is self-adjoint and its spectrum is the line segment $[0,1] = Cl\{e^{-t} : t \in ir^+\}$ (see, e.g., [7]). Therefore the spectrum of $\mathcal{P}$ in $\mathcal{H}_0$ is given by a compact perturbation of the continuous spectrum $\sigma_c = [0,1]$. The assertion is now a consequence of Theorem 5.2 in [14].

We shall now characterize some properties of the eigenfuctions of $\mathcal{P}$ in $\mathcal{H}_0$. First, it is easy to see that $\lambda = 0$ is an eigenvalue of infinite multiplicity. This follows by noting that

(see (3) and (34)) any function $f \in \mathcal{H}_0$ which is odd w.r.t. $x = 1/2$, e.g. $f(w) = 1 - 2w = \mathcal{B}[(1-t)(1-e^{-t})]$ lies in the kernel of $\mathcal{P}$.
Now suppose that $\mathcal{P}f = \lambda$ f for some $f \in \mathcal{H}_0$ and $\lambda \neq 0$, or explicitly

$$\lambda f(w) = \left(\frac{1}{w+1}\right)^2 \left[f\left(\frac{w}{w+1}\right) + f\left(\frac{1}{w+1}\right)\right]. \quad (82)$$

By Remark (1.1.14) $f(w)$ extends analytically to the half-plane H_0. If we transform this equation by substituting $1/w$ for w and then dividing through w^2 we get

$$\lambda w^{-2} f\left(\frac{1}{w}\right) = \left(\frac{1}{w+1}\right)^2 \left[f\left(\frac{1}{w+1}\right) + f\left(\frac{w}{w+1}\right)\right]. \quad (83)$$

Therefore f satisfies

$$wf(w) = \frac{1}{w} f\left(\frac{1}{w}\right) \quad (84)$$

for all $w \in H_0$. Note that applying (84) to each term of the r.h.s. in (82) one obtains

$$\lambda\, w\, f(w) = w\, f(w+1) + \frac{1}{w} f\left(1 + \frac{1}{w}\right). \quad (85)$$

For $\lambda = 1$ this yields $w\, f(w) = 1$. Note that for $f = \mathcal{B}[\varphi]$ we have

$$w^{-2} f\left(\frac{1}{w}\right) = \int_0^\infty e^{-tw} e^t \varphi(t) dm(t) = \mathcal{B}[(1-M)\mathcal{K}M^{-1}\varphi]. \quad (86)$$

Therefore the functional equation (84) can be written as

$$(1-M)\mathcal{K}M^{-1}\varphi = \varphi. \quad (87)$$

Now, given a continuous function ψ on ir^+ one can define (a version of) its Hankel transform (of order 1) as the integral

$$(J\,\psi)(t) = \int_0^\infty J_1\big(2\sqrt{st}\big) \sqrt{\frac{t}{s}}\, \psi(s) ds. \quad (88)$$

From the estimates $J_1(t) \sim t$ as $t \to 0^+$ and $J_1(t) = O\left(t^{-\frac{1}{2}}\right)$ as $t \to \infty$ ([11], vol.II) we see that the conditions on ψ sufficient to give the absolute convergence of the integral (88) are $\psi(t) = O\big(t^{-\beta}\big)$ as $t \to \infty$ with $\beta > -1/4$ and $\psi(t) = O(t^{\alpha})$ as $t \to 0^+$ with $\alpha > -1$. The identity (87) then says that the function (cf. Remark (1.1.11))

$$\psi(t) = \left(\frac{t}{1-e^{-t}}\right)\varphi(t) \quad (89)$$

satisfies

$$\psi(t) = \int_0^\infty J_1\big(2\sqrt{st}\big) \sqrt{\frac{t}{s}}\, \psi(s) ds. \quad (90)$$

Note that the simplest solution of this equation is ψ≡1 and corresponds to f=e (more general self-reciprocal functions satisfying equations related to (90) are discussed, e.g., in [35]). Furthermore, putting together (84), (86) and (89) we have that

$$f(w) = \int_0^\infty e^{-tw} \psi(t) dt \quad (91)$$

for all $w \in H_0$. Finally, one easily checks that if $\varphi \in L_2(m)$ then $\psi \in L_2(\widehat{m})$ where

$$d\widehat{m}(t) = \frac{e^{-t}(1-e^{-t})}{t \cdot \log 2} dt$$

We summarize the above in the following

Theorem (1.1.17)[1]: If $f \in \mathcal{H}_0$ satisfies $\mathcal{P}f = \lambda$ f for some $\lambda \neq 0$ then f is the (ordinary) Laplace transform of a function $\psi \in L_2(\widehat{m})$ which is self-reciprocal w.r.t. Hankel transform of order 1, namely f and ψ satisfy (91) and (90), respectively.

Now from Corollary (1.1.3) we know that a function $f = \mathcal{B}[\varphi]$ satisfies $\mathcal{P}f = \lambda f$ if and only if (the analytic continuation of) $\mathcal{K}_{\frac{1}{\lambda}}: L_2(m) \to L_2(m)$ satisfies $K_{\frac{1}{\lambda}}\varphi = \varphi$, which can also be written as

$$(\mathcal{K}\varphi)(t) = \frac{\lambda - e^{-t}}{1 - e^{-t}}\varphi(t) = \frac{\lambda - e^{-t}}{t}\psi(t). \tag{92}$$

Expressing the integral operator $\mathcal{K}$ in terms of the Hankel transform (88) we get $(\mathcal{K}\varphi)(t) = \frac{1}{t}J(\exp_{-1} \cdot \psi)(t)$, where we have defined the function $\exp_c : ir \to ir$ by $\exp_c (t) = e^{ct}$. Identities (90) and (92) then yield the integral equation

$$J(\exp_{-1} \cdot \psi) = (\lambda(\exp_{-1} -) \cdot J\psi. \tag{93}$$

Once more, ψ≡1 satisfies this equation with λ=1 (recall that $J\exp_{-1} = 1 - \exp_{-1}$). On the other hand, the above discussion suggests that there are no $\lambda \in \mathbb{C}\backslash\{0,1\}$ such that (93) has a (non-constant) solution $\psi \in L_2(\widehat{m})$. We are thus are led to formulate the following,

Conjecture (1.1.18)[1]: The only (non-zero) eigenvalue of $\mathcal{P}: \mathcal{H}_0 \to \mathcal{H}_0$ is $\lambda = 1$.

We end with two additional remarks.

Remark (1.1.19)[1]: (92) is a particular case of the Lewis functional equation

$$f(w) - f(w+1) = 1/w^{2(q+1)}f\left(1 + \frac{1}{w}\right), \tag{94}$$

which is related to the so called Maass cusp forms, i.e. PSL(2,ℤ)-invariant eigenfunctions of the Laplacian on the Poincar´e upper half-plane which vanish at the cusp (see [20]). Another type of functions equivalent to (even) Maass forms and considered in [20] are those satisfying an integral equation which in our notation writes

$$g(t) = \int_0^\infty \frac{J_{2q+1}(2\sqrt{st})}{\sqrt{st}}\left(\frac{s}{t}\right)^q g(s)dm(s). \tag{95}$$

By the foregoing (see Remark (1.1.12)) we see that for $q = 0$ we have the relation

$$f = \mathcal{B}[g]. \tag{96}$$

Remark (1.1.20)[1]: In the recent work [29], following [28] ten years later and somehow inspred by the construction presented here, Thomas Prellberg has studied the spectrum of (a generalized version of) $\mathcal{P}$ in a space of functions which is identical to $\mathcal{H}_0$ with the exception that the measure on ir^+ is slightly different from (44), being given by

$$d\widetilde{m}(t) = t\,e^{-t}dt. \tag{97}$$

It is easy to see that with this new measure the operator $\mathcal{Q}_z$ is isomorphic under generalized Laplace transform (cf. Theorem (1.1.7)) to $\widetilde{\mathcal{K}}_z: L_2(\widetilde{m}) \to L_2(\widetilde{m})$ given by

$$\widetilde{\mathcal{K}}_z = z(1 - zM)^{-1}\widetilde{K}, \tag{98}$$

where

$$\widetilde{\mathcal{K}}\varphi(s) = \int_0^\infty d\widetilde{m}(t)\frac{J_1(2\sqrt{st})}{\sqrt{st}} \cdot \varphi(t) \tag{99}$$

Notice that $\mathcal{K}_1 = (1 - M)^{-1}\widetilde{\mathcal{K}}$ which is not symmetric anymore. On the other hand, the relation given by Lemma (1.1.13) now writes (we keep using the symbols $\mathcal{L}$ and $\mathcal{B}$ to denote generalized Laplace and Borel transforms w.r.t. the measure $\widetilde{m}$):

$$\mathcal{L}[\varphi] = \mathcal{B}[\widetilde{\mathcal{K}}\varphi] \tag{100}$$

and hence the integral representation of $\mathcal{P}$ becomes

$$\mathcal{P}\mathcal{B}[\varphi] = \mathcal{B}[(M + \widetilde{\mathcal{K}})\varphi], \tag{101}$$

which is now symmetric (cf. (74)). Thus, everything goes as if the operators $\mathcal{P}$ and $\mathcal{Q}$ were not 'symmetrizable' both at the same time. Also notice that the function log 2·e if expressed as a generalized Borel transform now yields the function $\varphi(s) = 1/s$ which is not in $L_2(\widetilde{m})$.

We now consider the dynamical zeta functions ζ_F and ζ_G associated to the maps F and G, respectively, and defined by the following formal series [3]:

$$\zeta_F(z) = \exp\sum_{n=1}^{\infty} \frac{z^n}{n} Z_n(F) \quad \text{and} \quad \zeta_G(s) = \exp\sum_{n=1}^{\infty} \frac{s^n}{n} Z_n(G), \tag{102}$$

where the 'partition functions' $Z_n(F)$ and $Z_n(G)$ are given by

$$Z_n(F) = \sum_{x=F^n(x)} \prod_{k=0}^{n-1} \frac{1}{|F'(F^{k(x)})|} \quad \text{and} \quad Z_n(G) = \sum_{x=G^n(x)} \prod_{k=0}^{n-1} \frac{1}{|G'(G_k(x))|}. \tag{103}$$

Let us frst examine how $\zeta_F(z)$ and $\zeta_G(z)$ are related to one another. Let Per F and Per G denote the sets of all periodic points of the maps F and G, respectively. It is not difficult to realize that, as subsets of [0, 1], Per $F\backslash\{0\} = Per\ G$. Accordingly, given x in either of these sets, we let $p_F(x)$ and $p_G(x)$ denote the periods of x w.r.t. to F and G, respectively. They are related by

$$p_F(x) = \tau\,(x) + \tau\left(G(x)\right) + \cdots + \tau\left(G^{p_G(x)-1}(x)\right). \tag{104}$$

Moreover from the definitions of F and G we have

$$\prod_{k=0}^{p_F(x)-1} \frac{1}{|F'(F^k(x))|} = \prod_{k=0}^{p_G(x)-1} \frac{1}{|G'(G^k(x))|} = \prod_{k=0}^{p_G(x)-1} \left(G^k(x)\right)^2. \tag{105}$$

Using this facts we write $Z_n(F)$ as follows:

$$Z_n(F) = 1 + \sum_{m=1}^{n} \frac{n}{m} \sum_{x=F^n(x)=G^m(x)} \prod_{k=0}^{m-1} \left(G^k(x)\right)^2 \cdot \tag{106}$$

The second sum ranges over the $\binom{n-1}{m-1}$ ways to write the integer n as a sum of m positive integers. Therefore,

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{z^n}{n} Z_n(F) &= \log\left(\frac{1}{1-z}\right) + \sum_{n=1}^{\infty} \sum_{m=1}^{n} \frac{1}{m} \sum_{x=F^n(x)=G^m(x)} z^n \prod_{k=0}^{m-1} \left(G^k(x)\right)^2 \\ &= \log\left(\frac{1}{1-z}\right) + \sum_{\ell=1}^{\infty} \frac{1}{\ell} \sum_{x=G^\ell(x)} z^{p_F(x)} \prod_{k=0}^{\ell-1} \left(G^k(x)\right)^2.\end{aligned}$$

We are thus led to study the 'grand partition function' $\Xi_\ell(z)$ given by

$$\begin{aligned}\Xi_\ell(z) &:= \sum_{x=G^\ell(x)} z^{p_F(x)} \prod_{k=0}^{\ell-1} \left(G^k(x)\right)^2 \\ &= \sum_{n=0}^{\infty} z^{\ell+n} \sum_{x=G^\ell(x)=F^{\ell+n}(x)} \prod_{k=0}^{\ell-1} \left(G^k(x)\right)^2.\end{aligned} \tag{107}$$

The sum over periodic points yields $\binom{n+\ell-1}{\ell-1}=\binom{n+\ell-1}{n}$ terms, corresponding to the number of ways of distributing n identical objects into ℓ distinct boxes. According to (27), (28) and (107) we can also write $\Xi_\ell(z)$ in the following way:

$$\Xi_\ell(z)=\sum_{n=0}^{\infty} z^{\ell+n}\sum_{k_1+\ldots+k_\ell=n+\ell}\prod_{i=1}^{\ell} x^2_{k_i\ldots k_\ell k_1\ldots k_{i-1}}, \tag{108}$$

where $x_{k_1\ldots k_\ell}=[k_1,\ldots,k_\ell]$ denotes the irrational number whose continued fraction expansion is periodic of period ℓ and starts with the entries $k\_1,\ldots,k_\ell$. Putting together the above observations we obtain the next result, to be compared with Proposition (1.1.2):

Proposition (1.1.21)[1]: Consider the two-variable zeta function given by

$$\zeta_2(s,z)=\exp\sum_{\ell=1}^{\infty}\frac{s^\ell}{\ell}\Xi_\ell(z). \tag{109}$$

Then we have:

$$\zeta_2(1,z)=(1-z)\zeta_F(z)\quad and \quad \zeta_2(s,1)=\zeta_G(s) \tag{110}$$

wherever the series expansions converge absolutely.

In order to study the analytic properties of the function $\zeta_2(s,z)$ we further generalize (39) by introducing a family of operator-valued functions $Q_{z,q}, q=0,1,\ldots$, acting as (see [21] and [8] for related quantities)

$$Q_{z,q}f(x)=(-1)^q\sum_{n=1}^{\infty} z^{\tau(\Phi_n(x))}\cdot|\Phi_n'(x)|^{1+q}\cdot f\big(\Phi_n(x)\big), \tag{111}$$

together with a family of function spaces $\mathcal{H}_{1,q}\subseteq\mathcal{H}_0$ such that a function $f\in\mathcal{H}_{1,q}$ can be represented as

$$f(w)=\big(\mathcal{L}_q[\varphi]\big)(w):=\int_0^\infty dm(t)e^{-tw}t^q\varphi(t),\qquad \varphi\in L_2(m). \tag{112}$$

In particular $Q_{z,0}\equiv Q_z, \mathcal{L}_0\equiv\mathcal{L}$ and $\mathcal{H}_{1,0}\equiv\mathcal{H}_1$. We have the following result.

Proposition (1.1.22)[1]: For any given $q=0,1\ldots$ the operator valued function $z\to Q_{z,q}$ when acting on $\mathcal{H}_{1,q}$ can be analytically continued to the entre z-plane with a cross cut along the ray $(1,+\infty)$. For each z in this domain we have

$$Q_{z,q}\mathcal{L}_q[\varphi]=\mathcal{L}_q\big[\mathcal{K}_{z,q}\varphi\big], \tag{113}$$

where $\mathcal{K}_{z,q}: L_2(m)\to L_2(m)$ is given by

$$\big(\mathcal{K}_{z,q}\varphi\big)(t):=(-1)^q z(1-M)(1-zM)^{-1}\int_0^\infty dm(s)\frac{J_{2q+1}\big(2\sqrt{st}\big)}{\sqrt{st}}\varphi(s). \tag{114}$$

The operators $Q_{z,q}:\mathcal{H}_{1,q}\to\mathcal{H}_{1,q}$ and $\mathcal{K}_{z,q}: L_2(m)\to L_2(m)$ are both of trace class.

Proof. The first part follows from a straightforward extension to non zero q values of the arguments. The proof of the last assertion can be extracted from ([21], Theorem 3).

Now, the trace of the operator $\mathcal{K}_{z,q}$ is easily obtained (see also [21]):

$$tr\,\mathcal{K}_{z,q}=(-1)^q z\int_0^\infty\frac{J_{2q+1}(2t)}{e^t-z}dt=(-1)^q\sum_{k=1}^{\infty} z^k\int_0^\infty e^{-kt}J_{2q+1}(2t)dt$$
$$=(-1)^q\sum_{k=1}^{\infty} z^k\,\frac{x_k^{2(q+1)}}{1+x_k^2}, \tag{115}$$

where the numbers $x_k=\frac{\sqrt{k^2+4}-k}{2}=[k,k,k,\ldots]\equiv[\bar{k}]$ are the fixed points of $G(x)$ and the identity [10]

$$\int_0^\infty e^{-kt}J_p(2t)dt=\frac{\left(\sqrt{k^2+4}-k\right)^p}{2^p\sqrt{k^2+4}},\qquad p=0,1,\ldots \tag{116}$$

has been used. From (115) we immediately obtain the trace formula

$$\Xi_1(z)=tr\,\mathcal{K}_{z,0}-tr\,\mathcal{K}_{z,1}. \tag{117}$$

But we can say more. Indeed, a straightforward adaptation of ([21], Corollaries 4 and 5) to our z-dependent situation leads to the following general expressions:

$$\Xi_\ell(z)=tr\,\mathcal{K}_{z,0}^\ell-tr\,\mathcal{K}_{z,1}^\ell=tr\,\mathcal{M}_{z,0}^\ell-tr\,\mathcal{M}_{z,1}^\ell, \tag{118}$$

with

$$tr\,\mathcal{K}_{z,q}^\ell=(-1)^{q\ell}\sum_{k_1,\ldots,k_\ell=1}^{\infty} z^{k_1+\cdots+k_\ell}\frac{\prod_{i=1}^{\ell} x_{k_i\ldots k_\ell k_1\ldots k_{\ell-1}}^{2(q+1)}}{1-(-1)^\ell\prod_{i=1}^{\ell} x_{k_i\ldots k_\ell k_1\ldots k_{\ell-1}}^{2}} \tag{119}$$

Formula (118) along with standard arguments (see [21]) allow us to write the twovariables zeta function (109) as a ratio of Fredholm determinants,

$$\zeta_2(s,z)=\exp\sum_{\ell=1}^{\infty}\frac{s^\ell}{\ell}\Xi_\ell(z)=\frac{\det\left(1-s\,\mathcal{K}_{z,1}\right)}{\det\left(1-s\,\mathcal{K}_{z,0}\right)}=\frac{\det\left(1-s\,\mathcal{M}_{z,1}\right)}{\det\left(1-s\mathcal{M}_{z,0}\right)}, \tag{120}$$

where by definition

$$\det\left(1-s\,\mathcal{K}_{z,q}\right)=\exp\left(-\sum_{\ell=1}^{\infty}\frac{s^\ell}{\ell}tr\,\mathcal{K}_{z,q}^\ell\right) \tag{121}$$

is in the sense of Grothendieck [13]. We have thus proved the following result.

Theorem (1.1.23)[1]: Set $\mathcal{K}_z\equiv\mathcal{K}_{z,0}$, then we have:

(a) for each $s\in\mathbb{C}$, the function $\zeta_2(s,z)$, considered as a function of the variable z, extends to a meromorphic function in the cut plane $\mathbb{C}\backslash[1,\infty)$. Its poles are located among those z-values such that $\mathcal{K}_z: L_2(m)\to L_2(m)$ has $1/s$ as an eigenvalue;

(b) for each $z\in\mathbb{C}\backslash(1,\infty)$, the function $\zeta_2(s,z)$, considered as a function of the variable s, extends to a meromorphic function in $\mathbb{C}$. Its poles are located among the inverses of the eigenvalues of $\mathcal{K}_z: L_2(m)\to L_2(m)$.

Putting together the above Theorem and Proposition (1.1.21) we obtain

Corollary (1.1.24)[1]: The dynamical zeta functions ζ_F and ζ_G of the Farey and Gauss maps have the following properties:

(a) $\zeta_F(z)$ has a meromorphic extension to the cut plane $\mathbb{C}\backslash[1,\infty)$;

(b) $\zeta_G(s)$ has a meromorphic extension to $\mathbb{C}$. All poles are real and are located among the inverses of the eigenvalues of $\mathcal{K}: L_2(m)\to L_2(m)$ see [23], [32], [31].

Section (1.2): The Limit of Infinite Level

The Farey fractions (modified Farey sequence) may be defined as $r_k^{(n)}:=\frac{n_k^{(n)}}{d_k^{(n)}}$, with $gcd\left(n_k^{(n)},d_k^{(n)}\right)=1$, and n denoted the order of the Farey fraction at level k. Level $k=0$ consists of the two fractions $\left\{\frac{0}{1},\frac{1}{1}\right\}$. Succeeding levels are generated by keeping all the fractions from level k in level $k+1$, and including new fractions. The new fractions at level $k+1$ are defined via $d_{k+1}^{(2n)}:=d_k^{(n)}+d_k^{(n+1)}$ and $n_{k+1}^{(2n)}:=n_k^{(n)}+n_k^{(n+1)}$, so that

$$k = 0 \qquad \left\{\frac{0}{1}, \frac{1}{1}\right\}$$
$$k = 1 \qquad \left\{\frac{0}{1}, \frac{1}{2}, \frac{1}{1}\right\}$$
$$k = 2 \qquad \left\{\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}\right\}, etc.$$

Note that $n = 1, \dots, 2^k + 1$. It follows that the fractions at a given level are in increasing order. The Farey fractions may also be defined using products of the 2×2 matrices $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (see [40]).

Our main result concerns the sum of lengths of half of the intervals between “new” Farey fractions. Theorem (1.2.1) proves that the lim inf of this sum vanishes in the limit of infinite level k. Based on numerical evidence, we also conjecture that the limit of this sum vanishes. This very simple geometric property of the Farey fractions is not very apparent. The intervals chosen are alternating, and there seems no obvious reason why the sum of their lengths should vanish in this limit.

We focus on the “Farey tree”, which means retaining only the 2^{k-1} even Farey fractions at each level $k > 1$. These are exactly the new fractions at each level. In our notation they are of even order, i.e., $r_k^{(2n)}$ so for each level $k > 1$ we obtain the set

$$\left\{ r_k^{(2n)} \middle| n = 1, \dots, 2^{k-1} \right\}.$$

The lengths of the intervals between even (new) Farey fractions at every level k>1 are denoted

$$I_k^{(n)} := \left(r_k^{(2n)} - r_k^{(2n-2)} \right) > 0, \qquad (122)$$

where $n = 2, 3, 4, \dots, 2^{k-1}$. In what follows, for brevity, we abuse the terminology slightly and refer to $I_k^{(n)}$ as an interval. When n itself is even (i.e., $n = 2, 4, \dots, 2\hat{}(k-1)$) , we refer to these as even intervals. (Note that there are 2^{k-2} even intervals at level k.) The complementary intervals in the unit interval [0, 1] i.e., those with n odd ($n = 3, 5, \dots, 2^{k-1} - 1$), including the two extra intervals at the ends of the unit interval, namely $\left(r_k^{(2)} - r_k^{(1)} \right)$ and $\left(r_k^{(2k+1)} - r_k^{(2k)} \right)$, are the odd intervals. (Note that odd and even interva alternate.) From the definition of the Farey fractions it is easy to verify that each of the extra intervals has length $1/(k+1)$. Thus we combine them and define

$$I_k^{(1)} := 2/(k+1), \qquad (123)$$

see Fig. 1. (As a result there are 2^{k-2} odd intervals at level k, the same as the number of even intervals.)

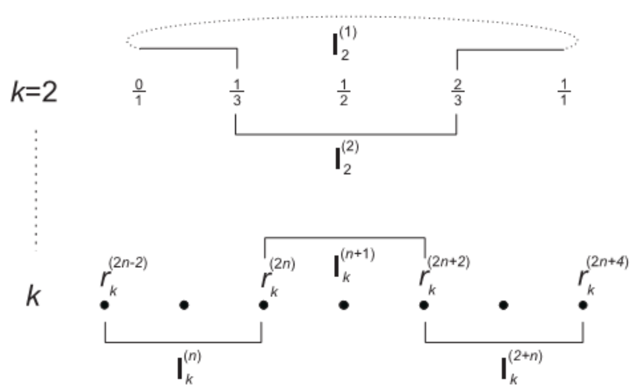


Fig. (1)[36]: Definition of Farey fraction intervals. Even (odd) intervals are shown via a line below (above) the fractions (note that in the lower diagram, n is even). The dashed line in the upper diagram indicates the combining of the two "extra" intervals into $I_k^{(1)}$. In this figure and those below the intervals are not to scale.
Next we define total length of even intervals at level k

$$I_k^{(e)} := \sum_{i=1}^{2^{k-2}} I_k^{(2i)}, \qquad (124)$$

and similarly for the odd intervals

$$I_k^{(o)} := \sum_{i=1}^{2^{k-2}} I_k^{(2i-1)}. \qquad (125)$$

It follows that

$$I_k^{(e)} + I_k^{(o)} = 1, \qquad (126)$$

for all $k \geq 2$.
The quantity

$$S_k := \sum_{n=1}^{2^{k-1}} \frac{1}{\left(d_k^{(2n)}\right)^2}, \qquad (127)$$

is the sum over the inverse squares of the new denominators at level k. As we will see, S_k is closely related to $I_k^{(e)}$.
By identifying intervals at different levels and examining their evolution from level to level, we prove our main result
Theorem (1.2.1)[36]:

$$\liminf_{k\to\infty} I_k^{(e)} = 0. \qquad (128)$$

Numerical evidence then leads us to the Conjecture.

$$\lim_{k\to\infty} I_k^{(e)} = 0. \qquad (129)$$

We contain some further remarks concerning this conjecture. 2.
Proof. We prove Theorem (1.2.1) i.e., $\liminf_{k\to\infty} I_k^{(e)} = 0$. The key step is Lemma (1.2.7), which bounds an arbitrary odd interval in terms of its "parent" even interval at a lower level. In addition, we present numerical evidence for the Conjecture (129).
It is convenient to use the full set of Farey fractions, even though only the even ones enter $I_k^{(e)}$ (see (124)). As mentioned, including $\frac{0}{1}$ and $\frac{1}{1}$, there are $2^k + 1$ fractions at level $k \geq 1$. In notation, at a given level $k \geq 1$, the even-numbered fractions are new, having been "born" at that level, while the odd-numbered ones are kept from the preceding level. Recall, also, that the intervals in (124) are exactly the $I_k^{(n)}$ with n even.

The rightmost inequality in (136) follows immediately from (126).

Now clearly $I_k^{(e)} > 0$ for any finite value of k. On the other hand, if there were an $\epsilon > 0$ such that $I_k^{(e)} \geq \epsilon$ for all k, the left hand side of (136) would diverge as $k \to \infty$. Thus Lemma (1.2.8) implies Theorem (1.2.1).

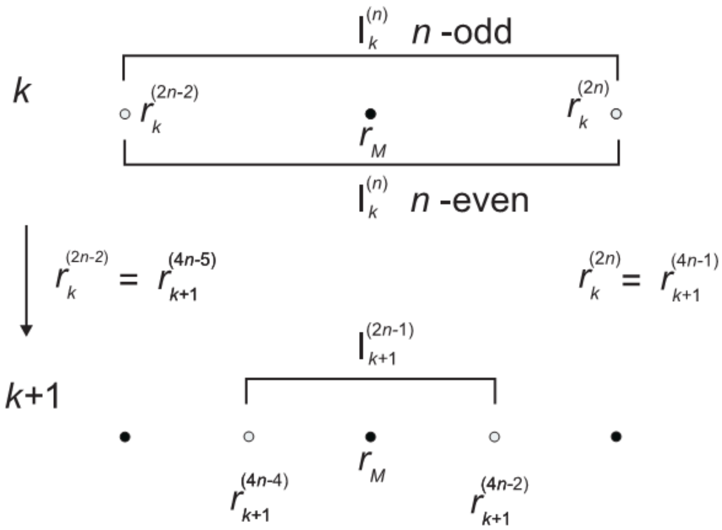


Fig. (2)[36]: Interval changes between levels of Farey fractions. The new Farey fractions (i.e., those "born" at level $k+1$) are depicted by empty circles. Note that any interval, even or odd, at level k gives rise to an odd interval at level $k+1$.

We prove Lemma (1.2.8) by understanding how even and odd intervals evolve from level k to level $k+1$ (see Fig. 2 ??). To proceed, first note that the nth Farey fraction at level k will be at order $2n-1$ in level $k+1$. Therefore we have

Lemma (1.2.2)[36]: The transformation of the order for any Farey fraction in going from level $k \to k+l$ is

$$n \to 2^{l}(n-1)+1. \tag{130}$$

The first step in our argument involves identifying intervals at successive levels. We do this via their "middle" Farey fractions. In a slight abuse of notation, let r_M be the "middle" Farey fraction of the interval $I_k^{(n)}$ at level k, i.e., $r_M = r_k^{(2n-1)}$ (see (122)). We use this to identify any set of intervals at different levels with the same "middle" fraction r_M. Thus, since r_M is necessarily of odd order, an interval at level k is the "parent" of the odd interval at level $k+1$ with the same r_M.

It follows that any even interval $I_k^{(2m)}$ at level k will produce a (necessarily smaller) odd interval at level $k+1$, with $r_M = r_k^{(4m-1)} = r_{k+1}^{(8m-3)}$. Similarly, any odd interval produces an (odd) interval with the same r_M at the next level (see Fig. 1). In addition, there are new even intervals that are born at each level. Their "middle" fractions are the ends of the even intervals from the previous level.

At level $k=2$, we have one even interval, which lies between the two "extra" intervals comprising $I_2^{(1)}$ at the ends of the unit interval. It follows that all odd intervals at level $k>2$ (except $I_k^{(1)}$) are born from even intervals at some previous level. Further, every odd interval shrinks from level to level while preserving its "middle" Farey fraction. This establishes

Lemma (1.2.3)[36]: For any level $k>2$, the unit interval is covered by a set of $2^{k-1}-1$ alternating even and odd intervals plus the two "end" intervals comprising $I_k^{(1)}$. The even intervals are "newborn", while each of the odd intervals (except $I_k^{(1)}$) is the offspring of an even interval born at a previous level.

The next step is to determine what fraction of a given interval at level k remains at level $k+1$. In doing this, it is useful to recall that the difference between any two successive fractions at a given level is $r_k^{(n+1)} - r_k^{(n)} = 1/d_k^{(n+1)} d_k^{(n)}$ (see for example [40], [38]).

Now consider an arbitrary even interval $I_{k+2}^{(2n)}$ at level $k+2$ (for $k \geq 0$; note that $n = 1, 2, \dots, 2^k$). Its "middle" fraction $r_M = r_{k+2}^{(4n-1)}$ is of odd order, and was therefore carried over from level $k+1$, where it is indexed as $r_M = r_{k+1}^{(2n)}$. This fraction is even, and therefore newborn at level $k+1$. Hence the neighboring fractions to its left and right, $r_{k+1}^{(2n-1)}$ and $r_{k+1}^{(2n+1)}$, respectively, are odd. These two fractions, therefore, appear at level k as $r_k^{(n)}$ and $r_k^{(n+1)}$, respectively.

Now the denominators of odd order fractions carry over from the previous level, i.e., $d_{k+1}^{(2n-1)} = d_k^{(n)}$, while those at even order (since they belong to "new" Farey fractions) are the sum of their neighbors, i.e., $d_{k+1}^{(2n)} = d_{k+1}^{(2n-1)} + d_{k+1}^{(2n+1)} = d_k^{(n)} + d_k^{(n+1)}$.

Putting these things together with Lemma (1.2.3) gives

Lemma (1.2.4)[36]: For any $k \geq 0$ and $n = 1, 2, \dots, 2^k$, i.e., any even interval at level $k + 2$, we have

$$I_{k+2}^{(2n)} = \frac{3}{\left(d_k^{(n)} + 2d_k^{(n+1)}\right)\left(2d_k^{(n)} + d_k^{(n+1)}\right)} \tag{131}$$

furthermore, for any $l \geq 1$,

$$I_{k+l+1}^{(2^{l-1}(2n-1)+1)} = (2l+1)\left(ld_k^{(n)} + (l+1)d_k^{(n+1)}\right)\left((l+1)d_k^{(n)} + ld_k^{(n+1)}\right) \tag{132}$$

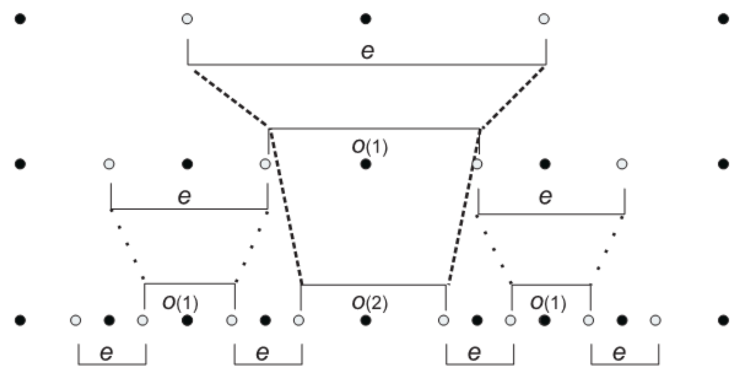


Fig. (3)[36]: Even and odd interval lineage. The numbers indicate age of a given odd interval.

where, for $l > 1$, (132) includes, at level $k + l + 1$, all descendants of the 2^k even intervals at level k.

Lemma (1.2.4) is illustrated in Fig. 3.

Now (except for $I_k^{(1)}$) every odd interval (for $k > 2$) is the descendant of a unique even interval at some lower level. Therefore (132) is valid for all $2^{k+l-1} - 1$ odd intervals at any level $k + l + 1$ with $k \geq 0$ and $l > 1$, omitting $I_k^{(1)}$. Consider the identity $2^{k+l-1} - 1 = \sum_{i=2}^{k+l} 2^{i-2}$. it expresses the number of odd intervals at level $k + l + 1$ in terms of a sum over the numbers of (even) "parent" intervals at each lower level i, with $2 \leq i \leq k + l$.

At this point, we consider the ratio of an arbitrary odd interval, as given by (132), to its parent interval $I_{k+2}^{(2n)}$. For simplicity, we let $m = k + 2$ and $j = l - 1$, so that $m \geq 2$ and $j > 0$, and relabel the lhs of (132) as $I_{m+j}^{([2n,j])}$, to indicate that it is the jth descendant of the 2nth interval at level m (note that $I_m^{([2n,0])} = I_m^{(2n)}$). Then, with $z := d_k^{(n)}/d_k^{(n+1)}$ we find

Lemma (1.2.5)[36]: For $m \geq 2$ and $j > 0$,

$$\frac{I_{m+j}^{([2n,j])}}{I_m^{(2n)}} = \frac{2j+3}{3}\frac{(1+2z)(2+z)}{\big((j+1)+(j+2)z\big)\big((j+2)+(j+1)z\big)}. \tag{133}$$

Lemma (1.2.5) expresses each successive descendent odd interval in terms of its parent even interval. It leads immediately, via elementary computations, to

Lemma (1.2.6)[36]: For $j > 0$ and $m \geq 2$,

$$\frac{2(2j+3)}{3(j+1)(j+2)} \leq \frac{I_{m+j}^{([2n,j])}}{I_m^{(2n)}} \leq \frac{3}{2j+3}. \tag{134}$$

The lower bound in (134) arises from (133) at $z = 0$ or $z = \infty$, and the upper bound from (133) at $z = 1$. In the case of Farey fraction denominators none of these z values actually occurs. However, the important point is that these bounds are independent of both the parent level m and the initial (parent) even interval.

Let $I_{m+j}^{([o,j])}$ denote the sum of all odd intervals at level $m + j$ that are descendants of the even intervals at an arbitrary level $m \geq 2$. Then

Lemma (1.2.7)[36]: For $j > 0$ and $m \geq 2$

$$I_m^{(e)}\frac{2}{3}\frac{2j+3}{(j+1)(j+2)} \leq I_{m+j}^{([o,j])} \leq I_m^{(e)}\frac{3}{2j+3}. \tag{135}$$

Lemma (1.2.8)[36]: For any $k > 2$ we have the bounds

$$\frac{2}{k+1} + \frac{2}{3}\sum_{j=1}^{k-2} I_{k-j}^{(e)}\frac{2j+3}{(j+1)(j+2)} \leq I_k^{(o)} < 1, \tag{136}$$

and

$$\frac{2}{k+1} + \sum_{j=1}^{k-2} I_{k-j}^{(e)}\frac{3}{2j+3} \geq I_k^{(o)}. \tag{137}$$

Proof. The remainder of the proof is as follows. First, relabel the "parent" level in Lemma (1.2.7) as $m = k - j$. If we fix $k = m + j > 2$, since m varies over the range $2 \leq m \leq k - 1$, j satisfies $1 \leq j \leq k - 2$. Thus all the odd intervals at an arbitrary level k are included, except the "end" interval $I_k^{(1)}$. The leftmost inequality in (136) and the inequality in (137) then follow directly on summing (135) and using (123).

Finally, A. Zhigljavsky [43] has verified numerically, up to level $k = 34$, that $I_k^{(e)}$ continues to decrease as k increases. His results are consistent with the result that $I_k^{(e)} \sim 1/\log_2(k)$ as $k \to \infty$ found in [39].

The conjecture (129) is, as already pointed out, very simple. However a proof is apparently quite elusive, at least using the methods employed. Even establishing that $I_k^{(e)}$ is monotonically decreasing with k, which, given (128), would be sufficient, appears very non-trivial. However several recent approaches to proving this conjecture have been proposed ([43], [41], [39]), based, respectively, on the Chacon-Ornstein ergodic theorem, continued fractions, and a measure theoretic analysis.

The problem treated here arose from previous investigations of the Farey fraction spin chains, a set of statistical mechanical models based on the Farey fractions (see [40], [38]). It follows directly from their definitions that the "Farey tree partition function" $Z_k^F(\beta)$ (see [38]) satisfies $Z_k^F(1) = I_k^{(e)}$, while the "even Knauf partition function" $Z_{k,e}^K(\beta)$ (see the equation after (124) in [38]) satisfies $Z_{k,e}^K(2) = S_k$.

Therefore the inequality (132), proven in [38], can be rewritten as

$$S_k < I_k^{(e)} < 4S_{k-1}, \tag{138}$$

which immediately proves

Lemma (1.2.9)[36]: The conjecture (129) is equivalent to

$$\lim_{k\to\infty} S_k = 0. \tag{139}$$

Note that the term "partition function" is used in its statistical mechanical sense here, which in general has no connection with the number theoretic usage.

As mentioned, several proofs of Conjecture (129) using different methods have been proposed. The method employed here does not seem capable of establishing Conjecture (129), however it gives detailed information about the evolution of the intervals not otherwise available.

There are several spin chains known to have the same free energy, and thus the same thermodynamic behavior. They all exhibit a second-order phase transition at non-zero temperature. (The free energy is defined via $f(\beta) := \lim_{k\to\infty}\left(-\frac{\log Z_k^F(\beta)}{k\beta}\right)$, and a phase transition is a singularity in $f(\beta)$– see [38]) The Farey tree model is employed in [37] for a study of multifractal behavior associated with chaotic maps exhibiting intermittency. The critical point (phase transition) in this model occurs at $\beta = 1$. Therefore, physically, $Z_k^F(1) = I_k^{(e)}$ is the value of the partition function at the critical point. (The value of the partition function at one point is, however, generally of no physical interest.)

In proving that the free energy of the Farey tree model is the same as the free energy for other Farey statistical models, it was already demonstrated [38] that as $k \to \infty, Z_k^F(\beta) \to 0$ for $\beta > 1$, while $Z_k^F(\beta) \to \infty$ for $\beta < 1$. (This, incidentally, also establishes that the Hausdorff dimension is $\beta_H = 1$.) However, exactly at the critical point, it was only shown that $0 < Z_k^F(1) < 1$. To our knowledge, the result (128) found here, and extended by [43], [41], [39], is new.

Finally, we note that [42] contains some related work, giving results on the large k behavior of the quantity

$$\sigma_k(\beta) := \sum_{i=1}^{2^k} \left(r_k^{(i+1)} - r_k^{(i)}\right)^{\beta}, \tag{140}$$

for $\beta > 1$.

Section (1.3): The Stern-Brocot and the Farey Sequence

We consider weighted uniform distributions (mod 1) for the following two canonical sequences: the Farey sequence $(\mathcal{F}_n)_{n\in\mathbb{N}}$ which is given by

$$\mathcal{F}_n := \left\{\frac{p}{q} : 0 < p \le q \le n, \gcd(p,q) = 1\right\},$$

and the even Stern-Brocot sequence $(S_n)_{n\in\mathbb{N}}$ which is given by

$$S_n := \left\{\frac{s_{n,2k}}{t_{n,2k}} : k = 1, \dots, 2^{n-1}\right\},$$

where the integers $s_{n,k}$ and $t_{n,k}$ are defined recursively by

$$s_{0,1} := 0 \quad \text{and} \quad s_{0,2} := t_{0,1} := t_{0,2} := 1;$$

$$s_{n+1,2k-1} := s_{n,k} \quad \text{and} \quad t_{n+1,2k-1} := t_{n,k}, \qquad \text{for} \quad k = 1, \cdots, 2^n + 1;$$

$$s_{n+1,2k} := s_{n,k} + s_{n,k+1} \quad \text{and} \quad t_{n+1,2k} := t_{n,k} + t_{n,k+1}, \qquad \text{for} \quad k = 1, \cdots, 2^n.$$

The following theorem states the main results, where δ_x denotes the Dirac distribution at $x \in [0,1]$, $*$-lim the weak limit of measures, and λ the Lebesgue measure on [0, 1]. Note that, throughout, all appearing fractions will always be assumed to be reduced.

In fact, for the derivation of the assertion in (134) we will show that the following more general measure theoretical result holds. In here, $T: [0,1] \to [0,1]$ denotes the Farey map defined by

$$T(x) := \begin{cases} \dfrac{x}{1-x} & \text{for } x \in \left[0, \dfrac{1}{2}\right] \\ (1-x)/x & \text{for } x \in \left(\dfrac{1}{2}, 1\right]. \end{cases}$$

In a nutshell, the proofs of these results are obtained as follows. The convergence in (135) is derived from combining Toeplitz's Lemma and a classical result by Landau [58] and Mikolás [59] with a well-know estimate for the Euler totient function $\varphi(n) := card\{1 \le m \le n : \gcd(m,n) = 1\}$. Whereas, the proof of Theorem (1.3.9), and consequently the proof of (134), is obtained from the following slightly more technical result, which will be derived by employing some recent progress in infinite ergodic theory.

The result in Proposition (1.3.5) has the following immediate elementary number theoretical implication, which has been the main result of [53] and which there led to the confirmation of a conjecture by Fiala and Kleban [49]. In particular, Proposition (1.3.5) hence gives rise to an alternative proof of this conjecture. But let us first recall that the regular continued fraction expansion of a number $x \in (0,1]$ is given by

$$x =: [x_1, x_2, \ldots] := \cfrac{1}{x_1 + \cfrac{1}{x_2 + \ldots}},$$

where all the x_i are positive integers. Also, we write $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$.

Corollary (1.3.1)[44]: We have that

$$\lambda\left(\left\{[x_1, x_2, \ldots] : \sum_{i=1}^{k} x_i = n, k \in \mathbb{N}\right\}\right) \sim \frac{1}{\log_2 n}.$$

Further immediate consequences of the results in Theorem (1.3.10) and Theorem (1.3.9) are given in the following two corollaries.

Corollary (1.3.2)[44]: We have that

$$* -\lim_{n\to\infty} \frac{\zeta(2)}{n} \sum_{\substack{\frac{p}{q} \in [0,1] \\ 2\log q \le n}} q^{-2} \delta_{\frac{p}{q}} = \lambda \quad \text{and} \quad \frac{* -\lim_{n\to\infty} \log(n^2)}{n} \sum_{\substack{\frac{p}{q} = [x_1, \ldots, x_k] \\ \sum_{i=1}^{k} x_i \le n}} q^{-2} \delta_{\frac{p}{q}} = \lambda.$$

The latter dichotomy can also be expressed in more down-to-earth terms as a dichotomy between partial geometric Poincaré sums and partial algebraic Poincaré sums for the modular group $\Gamma := P\,SL_2(\mathbb{Z})$. For results of this type on the algebraic growth rates of Poincaré series for more general Kleinian groups see [54]. In the following, d refers to the hyperbolic metric in the upper plane model of hyperbolic space and $|\cdot|$ denotes the word length in Γ with respect to the two generators $z \mapsto z+1$ and $z \mapsto -1/z$ of the modular group Γ. Also, we write $a_n \asymp b_n$ if a_n/b_n is uniformly bounded away from zero and infinity.

Corollary (1.3.3)[44]: We have that

$$\sum_{\substack{\gamma\in\Gamma \\ d(0,\gamma(0))\le n}} e^{-d(0,\gamma(0))} \asymp n \quad \text{and} \quad \sum_{\substack{\gamma\in\Gamma \\ |\gamma|\le n}} e^{-d(0,\gamma(0))} \asymp \frac{n}{\log n}.$$

Remark (1.3.4)[44]: (i) Note that the results in Theorem (1.3.10) complement well-known results on weak convergence of empirical measures with constant weight 1 for the sequences $(\mathcal{F}_n)$ and (S_n). More precisely, in [59] (see [47], [48], [55], [56], [58]) it was shown that $(\mathcal{F}_n)$ is uniformly distributed, that is,

$$* - \lim_{n\to\infty} \frac{1}{card(\mathcal{F}_n)} \sum_{\frac{p}{q}\in\mathcal{F}_n} \delta_{\frac{p}{q}} = \lambda. \tag{131}$$

On the other hand, it is known that the Stern-Brocot sequence is not uniformly distributed. In fact, an immediate consequence of the results in [52] is that

$$* - \lim_{n\to\infty} \frac{1}{card(S_n)} \sum_{\frac{p}{q}\in S_n} \delta_{\frac{p}{q}} = m_T,$$

where m_T refers to the measure of maximal entropy for the Farey map T. Here, We might like to recall that the distribution function of m_T is equal to the Minkowski question mark function (see e.g. [52]) and hence, the two measures m_T and λ are mutually singular. In fact, a numerical calculation has shown that the Hausdorff dimension $\dim_H\ (m_T) := \inf\{dim_H\ \ (X): m_T(X) = 1\}$ of the measure m_T is approximable equal to 0.875 (see e.g. [52], [57], [61]).

(ii) In order to tie the results in Theorem (1.3.10) (134) and Theorem (1.3.9) to elementary number theory and, in particular, to give a clarification of the factor vw in Theorem (1.3.9), we mention the following observation for the even Stern-Brocot tree. For each reduced fraction $v/w \in (0,1)$ and for all $n \in \mathbb{N}_0$, we have

$$\sum_{\frac{p}{q}\in T^{-n}\left(\frac{v}{w}\right)} \frac{1}{pq} = \frac{1}{vw}. \tag{132}$$

To see this first in an elementary way, note that we have $p/q \in S_n$ if and only if $T^{-1}(p/q) = \{p/(p+q), q/(p+q)\} \subset S_{n+1}$. Furthermore, with $\kappa{:}\cup_{n\in\mathbb{N}} S_n \to \mathbb{R}$ given by $\kappa(p/q) := 1/(pq)$, one immediately verifies that

$$\kappa(p/(p+q)) + \kappa(q/(p+q)) = \kappa(p/q).$$

The proof now follows by induction. Note that for the special case $v/w = 1/2$ one immediately verifies that $S_n = T^{-(n-1)}(1/2)$, and then (132) becomes

$$\sum_{\frac{p}{q}\in S_n} \frac{2}{pq} = 1, \qquad \text{for all } n \in \mathbb{N},$$

which has also been observed by the Canadian music theorist Pierre Lamothe (see by Bogomolny in [46]).

Alternatively, the equality in (132) can also be deduced immediately from the wellknown fixed point equation for the Perron-Frobenius operator $\mathcal{L}$ associated with the Farey map T. For this let h denote the eigenfunction of $\mathcal{L}$ associated with the eigenvalue 1. It is well known that h is given by $h(x) := 1/x$, which consequently gives that

$$\sum_{y\in T^{-n}(x)} |(T^n)'(y)|^{-1} h(y) = h(x), \qquad \text{for all } x \in (0,1) \text{ and } n \in \mathbb{N}_0.$$

Since $\left|(T^n)'\left(\frac{p}{q}\right)\right| = q^2/w^2$ for all $p/q \in T^{-n}(v/w)$, the statement in (132) follows. Finally, let us apply Theorem (1.3.9) to obtain yet another proof of the statement in (132), and this proof will then implicitly use dual aspects of the Perron-Frobenius operator. By applying Theorem (1.3.9) twice, we obtain the following, which immediately implies (132). For each $n \in \mathbb{N}_0$ and for every reduced fraction $v/w \in (0,1)$, we have

$$\sum_{\frac{p}{q}\in T^{-n}\left(\frac{v}{w}\right)} \frac{1}{pq}\cdot\lambda = \sum_{\frac{p}{q}\in T^{-n}\left(\frac{v}{w}\right)} *\text{-}\lim_{k\to\infty} \log k \sum_{\frac{r}{s}\in T^{-k}\left(\frac{p}{q}\right)} q^{-2}\delta_{\frac{p}{q}}$$
$$= \frac{1}{vw} *\text{-}\lim_{k\to\infty} \log(k^{vw}) \sum_{\frac{p}{q}\in T^{-(n+k)}\left(\frac{v}{w}\right)} q^{-2}\delta_{\frac{p}{q}} = \frac{1}{vw}\cdot\lambda.$$

As already mentioned, the proof of the Proposition (1.3.5) will make use of some results from infinite ergodic theory. Therefore, let us first recall a few basic facts and results from infinite ergodic theory for the Farey map. (For an overview, further definitions and details concerning infinite ergodic theory in general, see [45].) It is well known that the Farey system $([0,1], T, \mathcal{A}, \mu)$ is a conservative ergodic measure preserving dynamical systems. Here, $\mathcal{A}$ refers to the Borel σ-algebra of $[0,1]$ and the measure μ is the infinite σ-finite T -invariant measure absolutely continuous with respect to the Lebesgue measure λ. (Recall that conservative and ergodic means that for all $f \in L_1^+(\mu) := \{f \in L_1(\mu): f \geq 0$ and $\mu(f\cdot 1_{[0,1]}) > 0\}$ we have μ–almost everywhere $\sum_{n\geq 0} \hat{T}^n(f) = \infty$, where $1_{[0,1]}$ refers to the characteristic function of [0,1]; also, invariance of μ under T means $\hat{T}(1_{[0,1]}) = 1_{[0,1]}$, where $\hat{T}$ denotes the transfer operator defined below.) In fact, with $\varphi_0: [0,1] \to [0,1]$ defined by $\varphi_0(x) := x$, the measure μ is explicitly given by

$$d\lambda = \varphi_0 d\mu.$$

Recall that the transfer operator $\hat{T}: L_1(\mu) \to L_1(\mu)$ associated with the Farey system is the positive linear operator which is given by

$$\mu\left(1_C \cdot \hat{T}(f)\right) = \mu\left(1_{T^{-1}(C)}\cdot f\right), \qquad \text{for all } f \in L_1(\mu), C \in \mathcal{A}.$$

Note that the Perron-Frobenius operator $\mathcal{L}: L_1(\mu) \to L_1(\mu)$ of the Farey system is given by

$$\mathcal{L}(f) = |u_0'|\cdot(f\circ u_0) + |u_1'|\cdot(f\circ u_1), \qquad \text{for all } f \in L_1(\mu),$$

where u_0 and u_1 refer to the inverse branches of T, which are given for $x \in [0,1]$ by

$$u_0(x) = x/(1+x) \quad \text{and} \quad u_1(x) = 1/(1+x).$$

One then immediately verifies that the two operators $\hat{T}$ and $\mathcal{L}$ are related through

$$\hat{T}(f) = \varphi_0 \cdot \mathcal{L}(f/\varphi_0), \qquad \text{for all } f \in L_1(\mu).$$

Now, the crucial notion for proving Proposition (1.3.5) is provided by the following concept of a uniformly returning set which was introduced in [50].

A set $C \in \mathcal{A}$ with $0 < \mu(C) < \infty$ is called uniformly returning for $f \in L_\mu^+$ if there exists a positive increasing sequence $(w_n)_{n\in\mathbb{N}}$ of positive reals such that μ–almost everywhere and uniformly in C we have

$$\lim_{n\to\infty} w_n \hat{T}^n(f) = \mu(f).$$

In [50] it was shown that for the Farey system we have that every interval contained in [1/2, 1] is uniformly returning, for each function f which has the property that

$$\hat{T}^n(f) \in \mathcal{D} := \{g \in C^2([0,1]): g' \geq 0, g'' \leq 0\}.$$

In [50] it was shown that in the situation of the Farey system the sequence $(w_n)_{n\in\mathbb{N}}$ can be chosen to be equal to $(\log n)_{n\in\mathbb{N}}$. (For further examples of one dimensional dynamical systems which allow uniformly returning sets for some appropriate functions we refer to [60].) We are now in the position to give the proof of Proposition (1.3.5).

Proposition (1.3.5)[44]: For each interval $[\alpha, \beta] \subset (0,1]$ we have that

$$* \text{-}\lim_{n\to\infty} \left(\frac{\log n}{\log\left(\frac{\beta}{\alpha}\right)} \cdot \lambda|_{T^{-n}([\alpha,\beta])} \right) = \lambda.$$

Proof. Consider the function φ_t given by $\varphi_t : x \longmapsto x \cdot \exp(tx)$. The first aim is to show that for all $t \in [-1,1]$ we have

$$\hat{T}_{\varphi_t} \in \mathcal{D}.$$

Indeed, for $t \in [-1,0]$ this is an immediate consequence of the facts that φ_t is increasing and concave and that $\hat{T}(\mathcal{D}) \subset \mathcal{D}$. For $t \in (0,1]$, a straight forward computation shows that the first derivative at $x \in [0,1]$ is given by

$$(\hat{T}\varphi_t)'(x) = \frac{\varphi_t'\left(\frac{x}{x+1}\right) - x\varphi_t'\left(\frac{1}{x+1}\right)}{(x+1)^3} + \frac{\varphi_t\left(\frac{1}{x+1}\right) - \varphi_t\left(\frac{x}{x+1}\right)}{(x+1)^2}.$$

For the second derivative we then obtain

$$(\hat{T}\varphi_t)''(x) = \frac{(-2\,xt - 6x + 2t + xt^2 + 2x^3 - 4tx^2 - 4)\exp\left(\frac{tx}{x+1}\right)}{(x+1)^6} + \frac{(2tx - 6x - 2t + xt^2 + 2x^3 + 4tx^2 - 4)\exp\left(\frac{t}{x+1}\right)}{(x+1)^6}.$$

This immediately implies that $(\hat{T}\varphi_t)'' \le 0$, for all $t \in (0,1]$. Therefore, $(\hat{T}\varphi_t)'$ is decreasing on $[0,1]$ with $(\hat{T}\varphi_t)'(1) = 0$, which shows that on [0, 1] we have that $(\hat{T}\varphi_t)' \ge 0$. Hence, we can apply [51], which then implies that $\hat{T}\varphi_t \in \mathcal{D}$, for all $t \in [-1,1]$.

We proceed by noting that [51] guarantees that every interval contained in $[1/2, 1]$ is a uniformly returning set for φ_t, for each $t \in [-1,1]$. In order to complete the proof of the proposition, we employ the method of moments as follows. We show that for each $[\alpha,\beta] \subset (0,1]$ and for each $t \in [-1,1]$ we have for the moment generating function at t that

$$\lim_{n\to\infty} \int \exp(tx) \cdot \frac{\log n}{\mu([\alpha,\beta])} \cdot 1_{T^{-n}([\alpha,\beta])} d\lambda(x) = \int \exp(tx)\, d\lambda(x).$$

To see this, we argue by induction as follows. For $[\alpha,\beta] \subset \left[\frac{1}{2}, 1\right]$, we have that

$$\lim_{n\to\infty} \int \exp(tx) \cdot \frac{\log n}{\mu([\alpha,\beta])} \cdot 1_{T^{-n}([\alpha,\beta])}(x) d\lambda(x) = \frac{\lim_{n\to\infty} \log n}{\mu([\alpha,\beta])} \cdot \mu\left(\varphi_t \cdot 1_{T^{-n}([\alpha,\beta])}\right)$$
$$= \lim_{n\to\infty} \frac{\log n}{\mu([\alpha,\beta])} \cdot \mu\left(\hat{T}^n \varphi_t \cdot 1_{[\alpha,\beta]}\right) = \mu(\varphi_t) = \int \exp(tx)\, d\lambda(x).$$

Next, suppose that the assertion holds for any interval which is contained in the set $\mathcal{E}_n := \cup_{k=0}^{n-1} T^{-k}([1/2,1])$, and consider an interval $[\alpha,\beta] \subset T^{-n}\left(\left[\frac{1}{2}, 1\right]\right) \backslash \mathcal{E}_n$. Since $T([\alpha,\beta]) \subset \mathcal{E}_n$, we then have

$$\lim_{m\to\infty} \int \exp(tx) \cdot \frac{\log m}{\mu([\alpha,\beta])} \cdot 1_{T^{-m}([\alpha,\beta])}(x) d\lambda(x)$$
$$= \lim_{m\to\infty} \frac{\log m}{\mu([\alpha,\beta])} \cdot \mu\left(\hat{T}^m \varphi_t \cdot \left(1_{T^{-1}(T([\alpha,\beta]))} - 1_{T^{-1}(T([\alpha,\beta]))} \cap \mathcal{E}_n\right)\right)$$
$$= \lim_{m\to\infty} \frac{\log m}{\mu([\alpha,\beta])} \cdot \left(\mu\left(\hat{T}^{m+1}\varphi_t \cdot 1_{T([\alpha,\beta])}\right) - \mu\left(\hat{T}^m \varphi_t \cdot 1_{T^{-1}(T([\alpha,\beta]))\cap\mathcal{E}_n}\right)\right)$$
$$= \mu \frac{\varphi_t}{\mu([\alpha,\beta])}\left(\mu\left(T([\alpha,\beta])\right) - \mu\left(T^{-1}\left(T([\alpha,\beta])\right) \cap \mathcal{E}_n\right)\right)$$
$$= \int exp(tx) d\lambda(x).$$

This finishes the proof of Proposition (1.3.5).

$$T^{-(n-1)}([1/2,1]) = \left([x_1, x_2, \ldots]: \sum_{i=1}^{k} x_i = n, k \in \mathbb{N} \right),$$

it follows that

$$\lambda\left(\left\{[x_1, x_2, \ldots]: \sum_{i=1}^{k} x_i = n, k \in \mathbb{N}\right\}\right) \sim \frac{\log 2}{\log n}.$$

The following two lemmata will be required in the proof of Theorem (1.3.9). Note that the first lemma of these has already been obtained in [53]. However, in order to keep as self contained as possible, we include a proof here.

Lemma (1.3.7)[44]:

$$\sum_{\frac{p}{q}\in S_n} q^{-2} \asymp \frac{1}{\log n}.$$

Proof. First note that there is a $1-1$ correspondence between the sequence (S_n) and the set of connected components of $T^{-(n-1)}([1/2,1])$. That is, if $p/q = [a_1, \ldots, a_n] \in S_n$, where $a_n > 1$, then one of these connected component is given by

$$C_n(p/q) := \{[x_1, x_2, \ldots]: x_i = a_i \text{ for } 1 \le i \le n\}$$
$$\cup \{[x_1, x_2, \ldots]: x_i = a_i \text{ for } 1 \le i \le n-1, x_n = a_n - 1, x_{n+1} = 1\}.$$

Using standard Diophantine estimates we find that $\left(C_n\left(\frac{p}{q}\right)\right) \asymp 1/q^2$. Hence, an application of Corollary (1.3.1) finishes the proof of the lemma.

For the next lemma note that the sequence (S_n) can also be expressed in terms of the inverse branches u_1 and u_2 of the Farey map T. Namely, one immediately verifies that the orbit of the unit interval under the free semi-group Φ generated by u_1 and u_2 is in $1-1$ correspondence to the set of all Stern-Brocot intervals

$$\left\{\left[\frac{s_{n,k}}{t_{n,k}}, \frac{s_{n,k+1}}{t_{n,k+1}}\right) : n \in \mathbb{N}_0;\ k = 1, \ldots, 2^n\right\}.$$

Note that for each rational number $v/w \in (0,1]$ we have that

$$\left\{T^{-n}\left\{\frac{v}{w}\right\} : n \in \mathbb{N}\right\} = \left\{\gamma\ \left(\frac{v}{w}\right) : \gamma \in \Phi\right\}$$

Note that the Φ-orbit of 1 is equal to the set of rational numbers in $(0,1)$. More precisely, we have that if $\gamma \in \Phi$ then $\gamma(1) = v/w$, for some $v, w \in \mathbb{N}$ such that $v < w$ and $\gcd(v,w) = 1$, and for the modulus of the derivative of γ at 1 we then have that $|\gamma'(1)| = w^{-2}$.

We let $\mathcal{U}_\varepsilon(x)$ denote the interval centred at $x \in \mathbb{R}$ of Euclidean diameter $diam(\mathcal{U}_\varepsilon(x))$ equal to $\varepsilon > 0$.

Lemma (1.3.8)[44]: For each $g \in \Phi$ there exists $\Delta: (0,1] \to \mathbb{R}_+$ with $\lim_{s\to 0} \Delta(s) = 0$ such that for each $h \in \Phi$ and $\varepsilon > 0$ sufficiently small, we have

$$\left| diam\left(h\left(\mathcal{U}_\varepsilon(g(1))\right)\right) - \varepsilon \left|\left(h'(g(1))\right)\right|\right| < \varepsilon\, |(hg)'(1)|\Delta(\varepsilon).$$

Proof. By the bounded distortion property, we have for each $z \in (0,1)$ that there exists $\Delta_z: (0,1] \to \mathbb{R}_+$ with $\lim_{s\to 0} \Delta_z(s) = 0$ such that, for each $\varepsilon > 0$ sufficiently small,

$$\sup_{\substack{x,y\in\mathcal{U}_\varepsilon(z)\\ \gamma\in\Phi}} \left|\frac{|\gamma'(x)|}{|\gamma'(y)|} - 1\right| < \Delta_z(\varepsilon).$$

This implies that for fixed $g \in \Phi$ we have, for each $h \in \Phi$ and $\varepsilon > 0$ sufficiently small,

$$\left|\frac{diam\left(h\left(\mathcal{U}_\varepsilon(g(1))\right)\right)}{\varepsilon\left|h'(g(1))\right|} - 1\right| < \Delta_{g(1)}(\varepsilon).$$

From this we deduce that

$$\left| diam\left(h\left(\mathcal{U}_\varepsilon(g(1))\right)\right) - \varepsilon\left|h'(g(1))\right|\right| < \varepsilon\frac{|(hg)'(1)|}{|g'(1)|}\Delta\_g(1)(\varepsilon) = \varepsilon|(hg)'(1)|\Delta(\varepsilon).$$

This finishes the proof.

Theorem (1.3.9)[44]: For each rational number $v/w \in (0,1]$ we have that

$$* -\lim_{n\to\infty} \log(n^{vw}) \sum_{\frac{p}{q}\in T^{-n}\left\{\frac{v}{w}\right\}} q^{-2}\delta_{\frac{p}{q}} = \lambda. \tag{133}$$

Proof. Let $g \in \Phi$ be given and define, for $\varepsilon > 0$ sufficiently small,

$$\mathcal{U}_{g,\varepsilon,n} := T^{-(n-1)}\left(\mathcal{U}_\varepsilon(g(1))\right).$$

Let $u_{g,\varepsilon} := 1/\mu\left(\mathcal{U}_\varepsilon(g(1))\right) = 1/\log\left(\frac{g(1)+\frac{\varepsilon}{2}}{g(1)-\frac{\varepsilon}{2}}\right)$, and consider the measure $\nu_{g,\varepsilon,n}$ which is given, for each $n \in \mathbb{N}$, by

$$\nu_{g,\varepsilon,n} = u_{g,\varepsilon}\log n \cdot \lambda|_{\mathcal{U}_{g,\varepsilon,n}}.$$

By Proposition (1.3.5), we then have that $* -\lim_{n\to\infty} \nu_{g,\varepsilon,n} = \lambda$. Also, consider the atomic measure $\rho_{g,\varepsilon,n}$ which is given, for each $n \in \mathbb{N}$, by

$$\rho_{g,\varepsilon,n} := u_{g,\varepsilon}\log n \sum_{f(1)\in T^{-(n-1)}(g(1))} \frac{\varepsilon|f'(1)|}{|g'(1)|}\cdot\delta_{f(1)}.$$

Now, observe that

$$\lim_{\varepsilon\searrow 0} \varepsilon u_{g,\varepsilon} = \lim_{\varepsilon\searrow 0} \frac{\varepsilon}{\log\frac{g(1)+\frac{\varepsilon}{2}}{g(1)-\frac{\varepsilon}{2}}} = \lim_{\varepsilon\searrow 0} \frac{\varepsilon}{\frac{\varepsilon}{g(1)-\frac{\varepsilon}{2}}} = g(1),$$

and let the measures $\rho_{g,n}$ be defined by

$$\rho_{g,n} := g(1)\log n \sum_{f(1)\in T^{-(n-1)}(g(1))} \frac{|f'(1)|}{|g'(1)|}\cdot\delta_{f(1)}.$$

Using Lemma (1.3.7) and Lemma (1.3.8), we now obtain the following for all $x \in [0,1]$, where $F^{(\nu)}_{g,\varepsilon,n}, F^{(\rho)}_{g,\varepsilon,n}$ and $F^{(\rho)}_{g,n}$ denote the distribution functions of the measures $\nu_{g,\varepsilon,n}, \rho_{g,\varepsilon,n}$ and $\rho_{g,n}$, and where we write $a_n \ll b_n k$ if a_n/b_n is uniformly bounded from above,

$$\begin{aligned}\left|F^{(\nu)}_{g,\varepsilon,n}(x) - F^{(\rho)}_{g,n}(x)\right| &\le \left|F^{(\nu)}_{g,\varepsilon,n}(x) - F^{(\rho)}_{g,\varepsilon,n}(x)\right| + \left|F^{(\rho)}_{g,\varepsilon,n}(x) - F^{(\rho)}_{g,n}(x)\right| \\ &\ll u_{g,\varepsilon}\log n \sum_{hg(1)\in T^{-(n-1)}(g(1))} \left| diam\left(h\left(\mathcal{U}_\varepsilon(g(1))\right)\right) - \varepsilon\frac{|(hg)'(1)|}{|g'(1)|} \right. \\ &\quad \left. + \frac{\varepsilon u_{g,\varepsilon}\log n}{n^2}\right| + \left|g(1) - \varepsilon u_{g,\varepsilon}\right| \log n \sum_{f(1)\in T^{-(n-1)}(g(1))} |f'(1)| \\ &\ll \left(\varepsilon\, u_{g,\varepsilon}\Delta(\varepsilon) + \left|g(1) - \varepsilon u_{g,\varepsilon}\right|\right)\log n \sum_{f(1)\in T^{-(n-1)}(g(1))} |f'(1)| \\ &\ll \left|g(1) - \varepsilon u_{g,\varepsilon}\right| + g(1)\Delta(\varepsilon).\end{aligned}$$

This holds for ε>0 arbitrary small and hence, we obtain that

$$*-\lim_{n\to\infty} \rho_{g,n} = \lambda.$$

The proof of Theorem (1.3.9) now follows, if we insert in the definition of $\rho_{g,n}$ the fact that $g(1)$ can be written in form of a reduced fraction v/w and that then $|g'(1)| = w^{-2}$, as well as similarly, that $f(1)$ can be written in form of a reduced fraction p/q and that then $|f'(1)| = q^{-2}$.

Theorem (1.3.10)[44]: For the even Stern-Brocot sequence we have that

$$*-\lim_{n\to\infty} \log(n^2) \sum_{\frac{p}{q}\in S_n} q^{-2}\delta_{\frac{p}{q}} = \lambda, \tag{134}$$

and for the Farey sequence we have that

$$*-\lim_{n\to\infty} \frac{\zeta(2)}{\log n} \sum_{\frac{p}{q}\in \mathcal{F}_n} q^{-2}\delta_{\frac{p}{q}} = \lambda. \tag{135}$$

Proof. Define $\mathcal{F}^*_n := \left\{\frac{p}{n} : 0 < p \le n, \gcd(p,n) = 1\right\}$ and $\psi(n) := card(\mathcal{F}_n)$. We then clearly have that $\varphi(n) = card(\mathcal{F}^*_n)$ and that $\psi(n) \sim \frac{n^2}{2\zeta(2)}$. Next, observe that the statement in (131) implies that we have, for each continuous function $f : [0,1] \to \mathbb{R}_{\ge 0}$,

$$\chi_n := \frac{2\zeta(2)}{n^2} \sum_{r\in\mathcal{F}_n} f(r) \to \lambda(f), \text{for n tending to infinity.}$$

An application of Toeplitz's Lemma then gives that

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k}\chi_k = \lambda(f).$$

By setting $f_n := \sum_{\frac{p}{n}\in\mathcal{F}^*_n} f\left(\frac{p}{n}\right)$, we next observe that, for $n \ge 2$,

$$\frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k}\chi_k = \frac{2\zeta(2)}{\log n}\sum_{k=1}^{n}\frac{1}{k^3}\sum_{m=1}^{k} f_m = \frac{2\zeta(2)}{\log n}\sum_{m=1}^{n}\sum_{k=m}^{n}\frac{1}{k^3} f_m.$$

By comparing the sum $\sum_{k=m}^{n} k^{-3}$ with the corresponding integral $\int_m^n x^{-3}dx$, we obtain

$$\frac{\zeta(2)}{\log n}\sum_{q=1}^{n}\frac{f_q}{q^2}-\frac{\zeta(2)}{n^2\log n}\sum_{q=1}^{n}f_q \leq \frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k}\chi_k$$
$$\leq \frac{\zeta(2)}{\log n}\sum_{q=1}^{n}\frac{f_q}{q^2}-\frac{\zeta(2)}{n^2\log n}\sum_{q=1}^{n}f_q+\frac{\zeta(2)}{\log n}\sum_{q=1}^{n}\frac{f_q}{q^3}.$$

Finally, note that we clearly have that

$$\frac{\zeta(2)}{n^2\log n}\sum_{q=1}^{n}f_q \sim \frac{\lambda(f)}{2\log n}$$

and that

$$\frac{\zeta(2)}{\log n}\sum_{q=1}^{n}\frac{\varphi(q)}{q^3}\underbrace{\frac{f_q}{\varphi(q)}}_{\leq\|f\|_\infty} \leq \frac{\|f\|_\infty\big(\zeta(2)\big)^2}{\zeta(3)\log n}.$$

Hence, it now follows that

$$\lim_{n\to\infty}\frac{\zeta(2)}{\log n}\sum_{q=1}^{n}\frac{1}{q^2}\sum_{\frac{p}{q}\in\mathcal{F}_q^*}f\left(\frac{p}{q}\right)=\lim_{n\to\infty}\frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k}\chi_k=\lambda(f).$$

This finishes the proof of the assertion in Theorem (1.3.10) (135).

Chapter 2
Topology and Separation Properties

We shall discuss the interesting problem presented by R. F. Williams. For connected self-similar sets in the plane, that a finite overlap implies OSC. On the other hand, there are Cantor sets with arbitrary small dimensions which do not fulfil the OSC. We exhibit two examples of fractal sets, one not satisfying the weak separation property and whose existence was questioned by Zerner, the other having positive Hausdorff measure in its dimension and with the separation property failing on a subset of positive measure.

Section (2.1): Structure of Self-Similar Sets

The notion of fractals" was introduced by Mandelbrot [91] in the description of Nature. A set S is said to be a fractal provided that the Hausdorff dimension of S strictly exceeds the topological dimension of S. For example, Cantor's ternary set is a typical example of fractals. It is a classical problem to investigate such fractal sets in Mathematics. Indeed, measure theory is a fundamental and powerful tool to analyse fractals. See Rogers [106], Falconer [74]. On the other hand, as is pointed out by Mandelbrot, "self-similanty" is very important in the study of such sets. Actually, most classical fractal sets constructed by many mathematicians have the self-similanty in some sense.

We investigate various topological structures of selfsimilar sets, whose Definition will be given later, and to analyse many classical pathological sets and curves through the notion of self-similarity. With this for-mulation, one can easily create and handle self-similar fractals.

For X be a separable complete metric space with a metric d. A mapping $f: X \to X$ is said to be a contraction provided that the Lipschitz constant

$$Lip\ (f\rangle = \sup \frac{d(f(x), f(y))}{d(x, y)} \tag{1}$$

satisfies Lip $(f) < 1$. Every contraction f has a unique fixed point Fix (f) in X. Recent] y Hutchinson [83] considered the non-empty subset $K \subset X$ satisfying

$$K = f_1(K) \cup f_2(K) \cup \cdots \cup f_m(K) \tag{2}$$

where $m \geq 2$ and $\{f_j\}_{1 \leq j \leq m}$ is a given finite family of contractions.

On the other hand, Williams [115] studied the following set

$$K = closure \left(\bigcup_{\substack{1 \leq i_1, \ldots, i_n \leq m \\ n \geq 1'}} \text{Fix}\big(f_{i_1} \circ f_{i_2^\circ} \ldots \circ f_i\big) \right) \tag{3}$$

toward a study of generic properties of the action of free (non-abelian) groups on manifolds. He proved essentially that there exists a unique compact solution of (2); it is therefore given by (3). This result was also proved by Hutchinson in a different way. Several properties of K on geometric measure theory were proved in [83]. Mattila [93] strengthened some of them.

The equation (2) will be generalized to weak contractions and the solution K will be regarded as a fixed point of some set-dynamical system.

For another method to describe self-similar fractals using endomorphisms of words in free groups, see Dekking [70].

We begin with some Definitions. Let X be the same space.

Definition (2.1.1)[62]: A mapping $f: X \to X$ is said to be a *weak contraction* provided that $\Omega_f(t) \equiv \lim_{s \to t^+} \omega_f(s) < t$ for any $t > 0$, where ω_f is the modulus of continuity off:

$$\omega_f(s) = \sup_{d(x,y) \leq s} d(f(x), f(y)) \tag{4}$$

Obviously $\Omega_f(t)$ is non-decreasing and right-continuous. Note that every weak contraction f is uniformly continuous in X and has a unique fixed point Fix (*f*) in X. The regularity of ω_f may depend on the space X. Indeed, if X is compact, ω_f is rightcontinuous; that is, $\Omega_f = \omega_f$. If X is a closed convex subset of a Banach space, then ω_J is concave, therefore $\Omega_f = \omega_f$ is continuous. For example,] et$X$ be the unit interval with the Euclidean distance. Then the function $f(x) = x/(1+x)$ is a weak contraction with $\omega_f = f$, while f is not a contraction since Lip $(f) = 1$.

The power set 2^X of all subsets of X forms a poset under set-inclusion in a natural way; $x \leq y$ means x is a subset of$y$. Moreover, 2^X is a complete lattice with operations " $+$ " (ioin, set-union) and " (meet, set-intersection). See Birkholf [65] for lattice theory.

Let $\mathcal{C}(X) \subset 2^X$ be the set of all non-empty compact subsets of X. $\mathcal{C}(X)$ is not a lattice but a join-semilattice. It is known that $\mathcal{C}(X)$ is a complete metric space equipped with the Hausdorff metric:

$$d_H(x,y) = \max\left(\inf\{\varepsilon > 0; N_\varepsilon(x) \geq y\}, \quad \inf\{\varepsilon > 0; N_\varepsilon(y) \geq x\}\right), \tag{5}$$

where $N_\varepsilon(x)$ is an ε-neighborhood of the set x. Michael [94] proved that if X is compact, then $\phi(X)$ is also compact. Note that the mapping $: X \to \mathcal{C}(X)$, which maps a point p into the set consisting of the single point p, is an isometry.)

We now give a remark. Let $\{x_n\}_{n\geq 1}$ be a Cauchy sequence in (X) . Then we will denote by $\lim_{n\to\infty} x_n$ the unique limit of $\{x_n\}$ in (X) ; this means $\lim_{n\to\infty} x_n = \bigcap_{m\geq 1}$ closure $\left(\bigcup_{n\geq m} x_n\right)$ in the usual notation. Therefore, an infinite sum $\sum_{n=1}^{\infty} y_n$, if it exists, means the set closure $\left(\bigcup_{n\geq 1} y_n\right)$.

For any continuous mapping $f: X \to X$, we can define the *induced mapping* f^*: $\mathcal{C}(X) \to \mathcal{C}(X)$ by $f^*(x) = f(x)$ in a natural way.

Definition (2.1.2)[62]: A set $K \in \mathcal{C}(X)$ is said to be *self-similar* provided that the set K can be expressed in the form

$$K = \sum_{\lambda \in \Lambda} f_\lambda^*(K), \tag{6}$$

where $\{f_\lambda\}_{\lambda\in A}$ is a set of weak contractions of X and the index set Λ is $\{1, 2, \cdots, m\}, m \geq 2$, or N.

(6) means that the set K consists of a finite or an infinite number of miniatures of K itself. Note that Hutchinson's Definition of $[s_e]$ f-similarity differs from ours; he required some separation conditions in addition.

A mapping $F: \mathcal{C}(X) \to \mathcal{C}(X)$ is said to be *isotone* provided that $x \leq y$ implies $(x) - \leq F(y)$; a join-endomorphism provided that $F(x+y) = F(x) + F(y)$ for any $x, y \in \mathcal{C}(X)$. Let $\mathcal{F}(\mathcal{C}(X))$ be the set of all isotone join-endomorphisms (not necessarily continuous) defined on (X) . Obviously every induced mapping belongs to $\mathcal{F}(\mathcal{C}(X))$ and is further continuous. Again $\swarrow^{Reject}$ $(\mathcal{C}(X))$ becomes ajoin-semilattice; $F \leq G$ means $F(x) \leq G(x)$ and $F + G$ means $(F+G)(x) = F(x) + G(x)$ for any $\in \mathcal{C}(X)$. The following properties on the induced mappings were proved by [80].

Lemma (2.1.3)[62]: If f is a weak contraction of X, then f^* is also a weak contraction of $\phi(X)$ with $\Omega_e = \Omega_f$. Moreover, if $\{f_j\}_{1\le j\le m}$ is a finite set of weak contractions of X, *then* $F = \sum_{j=1}^{m} f_j^*$ is also a weak contraction of $\mathcal{C}(X)$ *with* $\Omega_F(t) \le \max_{1\le j\le m}\Omega_{f_j}(t)$.

We shall discuss the equation (6) and generalize the results of Williams and Hutchinson mentioned. In addition, we shall discuss different types of set-equations.

By Lemma (2.1.3) we get a generalization of Hutchinson's result immediately.

Theorem (2.1.4)[62]: *Suppose that* $\{f_j\}_{1\le j\le m}$, $m \ge 2$, *is a finite set of weak contractions of* X. *Then there exists a unique compact subset* $K = K(f_1, \cdots, f_m)$ *satisfying the equation* (6) *with* $\Lambda = \{1,2,\cdots,m\}$. *Moreover, for any compact subset* $\in \mathcal{C}(X)$, *we have*

$$\lim_{n\to\infty} F^n(Q) = K(f_1,\cdots,f_m)\,, \tag{7}$$

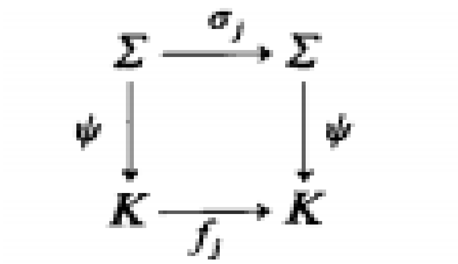
where $= \sum_{j=1}^{m} f_j^* \in \mathcal{F}(\mathcal{C}(X))$.

To investigate the structure of the set $(f_1, \cdots, f_m)$, it is convenient to introduce the *one-sided symbol space* $\Sigma = \{1, 2, , m\}^N$ on m symbols. Endowed with the metric

$$d_z(\alpha,\beta) = \sum_{n\ge 1} 2^{-n}\,\tau(\alpha_n,\beta_n)\; for\ \alpha = (\alpha_n)\,, \beta = (\beta_n\rangle \in \Sigma\,, \tag{8}$$

where $\tau(i_{J)=}1$ if $i \ne j$ and $\tau(i_J) = 0$ if $i = j$, Σ becomes a compact metric spaoe. Then we have

Theorem (2.1.5)[62]: *Suppose that* $\{f_j\}_{1\le j\le m}$, $m \ge 2$, *is a finite set of weak contractions of* X. *Then there exists a continuous onto mapping* $\psi\colon \Sigma \to K(f_1,\cdots,f_m)$ *such that the following diagram is commutative*:



where σ_j *is the right-shift operator*: $\sigma_j(\alpha_1\alpha_2\cdots) = \mathrm{t}i\alpha_1\alpha_2\cdots)$ *for any* $1 \le j^- \le m$. *In particular, Williams' formula* (3) *holds true.*

Proof. Let $\Omega(t) = \max_{1\le j\le m}\Omega_{f_\mathrm{j}}(t)$ for brevity. First we will show that the sequence defined by $p_n(\alpha) = f_{\alpha_1}\circ f_{\alpha_2}\circ\cdots\circ f_{\alpha_n}(p_0)$, $n \ge 1$, is a Cauchy sequence in X. To show this, for any $\varepsilon > 0$, define a sufficiently large integer $N = N(\varepsilon)$ such that

$\Omega^N(M) \le \varepsilon - \mathrm{q}\varepsilon)$ where $= \max_{1\le j\le m} d\left(p_0, f_j(p_0)\right)$.

Then $d\big(p_N(\alpha), p_{\mathrm{N}+1}(\alpha)\big) \le \Omega^N(M) \le \varepsilon - \Omega(\varepsilon) < \varepsilon$ for any $\alpha \in \Sigma$. Suppose tiow that $d\left(p_N(\alpha), p_{N+j}(\alpha)\right) \le \varepsilon$ for any $1 \le j \le k$ and $\alpha \in \Sigma$. Then it follows that

$$\begin{aligned} d\big(p_N(\alpha), p_{N+k+1}(\alpha)\big) &\le d\big(p_N(\alpha), p_{N+1}(\alpha)\big) + d\big(p_{N+1}(\alpha), p_{N+k+1}(\alpha)\big) \\ &\le \varepsilon - \Omega(\varepsilon) + \Omega\left(d\left(p_N(\alpha')_{P_{N+k}}(\alpha')\right)\right) \le \varepsilon, \end{aligned}$$

where $\alpha' = (\alpha_2\alpha_3 \cdots)$. Hence $d\left(p_N(\alpha), p_{N+j}(\alpha)\right) \le \varepsilon$ for any $j \ge 1$ by induction and therefore $\{p_n(\alpha)\}$ is a Cauchy sequence. It is easily seen that $p_\infty(\alpha) = \lim p_n(\alpha)$ is independent of the choice of p_0.
Now define $\psi(\alpha) = p_\infty(\alpha)$ for $\alpha \in \Sigma$. Since $d\big(p_\infty(\alpha), p_N(\alpha)\big) \le \varepsilon$, the set $\psi(\Sigma)$ is bounded and therefore ψ is continuous. Thus $\psi(\Sigma)$ is a compact subset satisfying the equation (2). Therefore we have $K(f_1, \cdots, f_m) = \psi(\Sigma)$ by Theorem (2.1.4).
For a fixed weak contraction f of X, let $W_f(X)$ be the set of all weak contractions g satisfying $\Omega_9(t) \le \Omega_f(t)$ for any $t > 0$. $\mathcal{W}_f(X)$ is endowed with topology of uniform convergence on compact sets. Then we have
Theorem (2.1.6)[62]: *Suppose that f is a weak contraction of X. Then the mapping*
$$K: \mathcal{W}_f(X) \times \cdots \times \mathcal{W}_f(X) \to \mathcal{C}(X),$$
which maps $(f_1, \cdots, f_m)$ into the set $K(f_1, \cdots, f_m)$, is continuous.
Proof. Suppose that $g_j^{(n)} \to g_j$ as $n \to \infty$ in $\|'_f\ (X)$ for $1 \le j \le m$. Put $d^* =$ diam $\big(K(g_1, \cdots, g_m)\big)$ for brevity. Let G_δ^ε be the closure of $\{(x, y); \varepsilon \le x \le d^*,\ y = \Omega_f(x) \ge x - \delta\}$ for , $\delta > 0$. Then it follows that for any fixed $\varepsilon > 0$, $G_\delta^\varepsilon = \varphi$ for a sufficiently small $= \delta(\varepsilon)$. Thus there exists $n(\varepsilon)$ such that $H^n(d^*) \le \varepsilon$ for any $n \ge n(\varepsilon)$ where $H(x) = \Omega_f(x) + \delta$.
On the other hand, there exists $N = N(\varepsilon) \ge n(\varepsilon)$ such that
$$\sup_{x \in Q} d\left(g_j^{(n)}(x), g_j(x)\right) \le \delta(\varepsilon\rangle$$
where $= K(g_1, \cdots, g_m) \in \mathcal{C}(X)$. Then
$$d\left(g_{l1}^{(n)} \circ g_{\propto 2}^{(n\rangle} \circ \cdots, g_{\alpha_1} \circ g_{\alpha_2} \circ \cdots\right)$$
for any $1 \le j \le m$, $n \ge N$
$$\le d\left(g_{\alpha_1}^{(n\}} \circ g_{\alpha 21}^{(n)} \circ \cdots, g_\alpha^{(n)\circ} g_\alpha 2 \circ \cdots\right) + d(, g_\alpha 1 \circ g_{\alpha 2} \cdots)$$
$$\le \Omega_f\left(d\left(g_{\alpha 2}^{(,)} \circ \cdots, g_\alpha 2 \circ \cdots\right)\right) + \delta(\varepsilon) = H\left(d\left(g_\alpha^{(n)} \circ \cdots, g_2 \circ \cdots\right)\right).$$
Continuing in this way, we arrive at $d\left(g_{\alpha_1}^{(n)\circ} \cdots, g_{\alpha_1^\circ} \cdots\right) \le fP(d^*) \le \varepsilon$ for $n \ge N$; therefore $d_H\left(K\left(g_1^{(n)}, \cdots, g_n^{(n\rangle}\right), K(g_1, \cdots, g_m)\right) \le \varepsilon$. Since ε is arbitrary, this completes the proof.
For the case $\Lambda = N$, we have
Theorem (2.1.7)[62]: *Suppose that $\{f_n\}_{n \ge 1}$ is a family of weak contractions of X satisfying $\lim_{n \to \infty} \Omega_{f_n}(t) = 0$ for any $t > 0$. Suppose further the set $\bigcup_{n \ge 1} \mathrm{F\,ix}(f_n)$ is precompact. Then there exists a unique compact subset $K = K(f_1, f_2, \cdots)$ satisfying the equation* (6) *with $\Lambda = N$. Moreover, for any compact $\in \mathcal{C}(X)$, we have*
$$\lim_{r \to \infty} F^n(Q) = K(f_1, f_2, \cdots) \tag{9}$$
where $= \sum_{n \ge 1} f_n^* \in \mathcal{F}\big(\mathcal{C}(X)\big)$.
Proof. We first show that the operator $F = \sum_{n \ge 0} f_n^*$ is well-defined. It suffices to show the set $\bigcup_{n \ge 1} f_n(x)$ is pre-compact for any $\in \mathcal{C}(X)$. We denote by $\gamma(M)$ Kuratowski's noncompactness measure [87] of a bounded subset M of X. For any $\in \mathcal{C}(X)$, put $Q = \sum_{n \ge 1}$ Fix$(f_n) \in \mathcal{C}(X)$ and $d^* = \sup\{d(p, q); p \in x, q \in Q\}$ for brevity. Then, for any $\varepsilon >$

0, there exists $N = N(\varepsilon)$ such that $d(\mathrm{Fix}(f_n), f_n(p)) \le \Omega_{J_n}(d(\mathrm{Fix}(f_n), p)) \le \Omega_{f_n}(d^*) \le \varepsilon$ for any $p \in x, n \ge N$. This implies $f_n(x) \subset N_\varepsilon(Q)$ and therefore

$$\gamma\left(\bigcup_{n\ge1} f_n(x)\right) \le \gamma\left(\bigcup_{n\ge N} f_n(\mathrm{x})\right) \le \gamma(N_\varepsilon(Q)) \le \gamma(Q) + 2_{\mathrm{i}} = 2\varepsilon.$$

Since ε is arbitrary, it follows that $\bigcup_{n\ge1} f_n(x)$ is pre-compact.
Now define $\Omega^*(t) = \sup_{n\ge1}\Omega_{J_\mathrm{n}}(t)$ for any $i > 0$. Evidently we have $\Omega_F(t) \le \Omega^*(t\rangle$. Also it is easily verified that Ω^* is a non-decreasing right-continuous function satisfying $\Omega^*(t) < t$ for $t > 0$. This implies that F is a weak contraction of $\mathcal{C}(X\}$; this completes the proof.
Note that the symbol space $\Sigma = N^N$ is complete (not compact) with the metric (8). Then we have

Theorem (2.1.8)[62]: *Suppose that* $\{f_n\}_{n\ge1}$ *satisfies the same conditions as in Theorem* (2.1.7). *Then there exists a continuous mapping* $\psi_\infty: N^N \to K(f_1, f_2, \cdots)$ *such that*

$$K(f_1, f_2, \cdots\rangle = \text{closure}\left(\psi_{\alpha)}(N^N)\right)$$

$$= closure\ (\bigcup_{n\ge1^n} Fix(f_{i_1^\circ}\ldots\circ f_\mathrm{i})). \quad (10)$$

The proof is similar to that of Theorem (2.1.5) and easily verified.
We now remark that it is quite interesting to take off the restriction that $\{f_j\}_{1\le j\le m}$ is a set of weak contractions in the equation ((2.1.3)⟩. As an example, consider a rational function $R(z)$ on the Riemann sphere $\overline{C}$. The *Julia set* J of $R(z)$ is defined by the set of $\overline{C}$ where the family of the iteration $\{R^n(z)\}$ is not normal. It is well-known that J is a closed, perfect and completely invariant set under R; that is, $J = R(J) = R^{-1}(J)$ (see [5]). On the other hand one can easily show that a set Y is completely invariant under R if and only if the set y satisfies

$$Y = R(\mathrm{y}) + R^{-1}(\mathrm{y}). \quad (11)$$

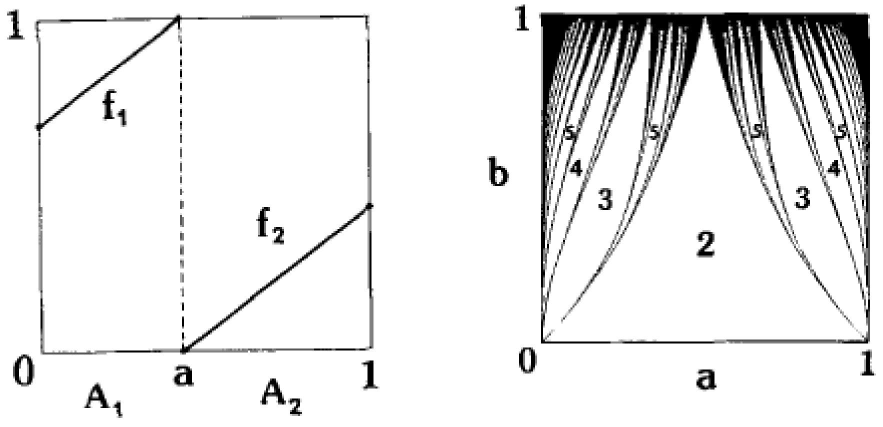
Then we conclude that the Julia set is the smallest closed solution of ((2.1.8)⟩ which contains a repulsive periodic point, since Julia showed that J is the closure of the set of all repulsive periodic points. (This will correspond to Williams' formula (3).) As a second example, consider the action in $\overline{C}$ of a discrete subgroup G of Möbius transformations. For simplicity, we suppose that $G = \langle A, B\rangle$ is not elementary. Then the *limit set* L of G is defined by the closure of the set of points fixed by some elements of G. It is well-known that L is a perfect and G-invariant set; that is, $L = V(L)$ for all y in G (see e.g. Beardon [63]). In other words, L is the smallest nonempty closed set satisfying

$$L = A(L) + A^{-1}(L) + B(L) + B^{-1}(L)\,. \quad (12)$$

Finally we will give an interesting example of a set-equation different from (6). Let X be the unit interval $[0,1]$ with the usual Euclidean distance. Then we consider the set-equation

$$K = f_1(K \cdot A_1) + f_2(KA_\mathrm{z})\,, \quad (13)$$

where $A_1 = [0, a]$, $A_2 = [a, 1]$ and $f_1(s) = 1 + b(s - a)$, $f_2(s) = b(s - a)$ with two parameters $0 < a < 1$, $0 < b < 1$ (Fig. 1(a)). The equation (13) originates in the study of some discontinuous dynamical system done by [77]. In fact, the attractor of the dynamical system becomes a compact solution of (13) for almost all parameters. The uniqueness of such a solution follows from the fact that the attractor is *minimal*. If (a, b) belongs to the domain D_n numbered by n in Fig. l(b),



(a) (b)

Fig. (1)[62]:

the solution K of (13) consists of n points. On the other hand, $\overline{1}\mathrm{f}(a, b)$ belongs to the remainder set $R = (0,1)^2 - \sum_{n\geq 2} \overline{D}_n$, K becomes a Cantor set with zero Hausdorff dimension. Note that the mapping $F(x) = f_1^*(x \cdot A_1) + f_2^*(x \cdot A_2)^{\uparrow)}$ for $x \in \mathcal{C}(X)$ belongs to $\mathcal{F}(\phi(X))$, while it is discontinuous in $\mathcal{C}(X)$. We will give a generalization of the above fact as follows:

Theorem (2.1.9)[62]: Suppose that X consists of $m \geq 2$ closed convex subsets $A_1, \cdots, A_m$ of R^p with the usual Euclidean distance. Let $f_j : A_j \to X$ be a weak contraction for $1 \leq j \leq m$. Then the equation

$$K = \sum_{j=1}^{m} f_j\left(K \cdot A_j\right) \tag{14}$$

has the maximal compact solution K_M; *that is, every compact solution* K *of* (14) *satisfies* $K \leq K_M$. *If in addition* $K_m \cdot A_i \cdot A_j = \varphi$ *for any* $i \neq j$, *then* K_M is a unique compact solution of (14) if and only if K_M is minimal; that is,

$$K_M = \prod_{i\geq 1} \sum_{j\geq i} F^j\left(\{q\}\right) \, for \, any \, \mathrm{q} \in K_M \,, \quad (15)$$

where $(x) = \sum_{j=1}^{m} f_j^*\left(x \cdot A_j\right) \in \mathcal{F}\left(\mathcal{C}(X)\right)$.

Proof. It is known that there exists a retraction $r_j : R^p \to A_\mathrm{j}$ such that Lip $\left(r_\mathrm{j}\right) \leq 1$ for $1 \leq j \leq m$. Hence the extension $f_j = f_j \circ r_j \mathrm{of} f_j$ becomes a weak contraction of X. Put $Q = K(f_1, \cdots, f_m)$ for brevity. Then

$$Q = \sum_{j=1}^{m} f_j^*\,(Q) \geq \sum_{j=1}^{m} f_\mathrm{j}^*\left(Q \cdot A_j\right) = F\langle Q)$$

and therefore there exists $Q_\infty = \lim_{n\to\infty} Q_n \equiv \lim_{n\to\infty} F^N\,(Q) \in \mathcal{C}(X)$ since F is isotone. We now show Q_∞ satisfies the equation (14). One can easily show that (i) if $Q_\infty \cdot A_\mathrm{j} = \varphi$, then $Q_N \cdot A_j = \varphi$ for some N; (ii) if $Q_\infty \cdot A_j \neq \varphi$, then $Q_n \cdot A_j \to Q_\infty \cdot A_j$ as $n \to \infty$ in $\mathcal{C}(X)$. Hence $Q_{n+1} = F(Q_n) = \sum f_j^*\left(Q_n \cdot A_j\right) \to \sum f_j^*\left(Q_\infty \cdot A_j\right) = F(Q_\infty)$ as required.

Put $\tilde{F} = \sum_{j=1}^{m} \tilde{f}_j^*$. Then for every compact solution K of (14), we have

$$\tilde{F}(K)=\sum_{j=1}^{m} f_{\mathrm{j}}^{*}(K) \geq \sum_{j=1}^{m} f_{j}^{*}\left(K \cdot A_{j}\right)=K,$$

and therefore $Q=\lim_{n \rightarrow \infty} \tilde{F}^{n}(K) \geq K$ by Theorem (2.1.4). Hence

$$Q_{\infty}=\lim_{n \rightarrow \infty} F^{n}(Q) \geq \lim_{n \rightarrow \infty} F^{n}(K)=K.$$

Thus $K_{M}=Q_{\infty}$ is the maximal solution of (14).

We now show the second part of the theorem. It suffices to deduce the minimality from the uniqueness of K_{M}. For any fixed $q \in K_{M}$, put $Q^{\infty}=\lim_{n \rightarrow \infty} Q^{n}$ where $Q^{n}=\sum_{j \geq n} F^{i}(\{q\})$. Then we have

$$F(Q'?=\sum_{j=1}^{m} f_{j}\left(Q^{n} \cdot A_{j}\right) \rightarrow \sum_{\mathrm{j}=1}^{m} f_{j}\left(Q^{\infty} \cdot A_{j}\right)=F\left(Q^{\infty}\right) \text{ as } n \rightarrow \infty,$$

since $\{\Psi\}$ is a decreasing sequence in $\mathcal{C}(X)$. Hence $F(\Psi\rangle \geq \sum_{j \geq n+1} F^{j}(\{q\})=Q^{u+1}$ and therefore $F\left(Q^{\infty}\right) \geq Q^{\infty}$. Since $Q^{\infty} \leq K_{M},\left\{Q^{\infty} \cdot A_{j}\right\}_{1 \leq j \leq m}$ are pairwise disjoint compact subsets by assumption. Therefore $F\left(Q^{\infty}\right) \leq Q^{\infty}$ and we get $Q^{\infty}=K_{M}$ by uniqueness. This completes the proof.

We will discuss the connectedness of self-similar sets. Throughout, $\dim_{T}\langle Q)$ denotes the (topological$\rangle$ dimension of a set Q in the Menger-Urysohn sense (see e.g. Hurewicz-Wallman [82] $\rangle$; $W(n)$ denotes the set of all finite words with length n on symbols {1, 2, , $m\}$. First of all, we have

Theorem (2.1.10)[62]: (Williams [$s'4$]). *Suppose that* $\left\{f_{f}\right\}_{1 \leq j \leq m}$ *is a finite set of contractions of* X *satisfying* $\sum_{j=\ 1}^{m} \mathrm{L\,ip}\left(f_{j}\right)<1$. *Then* $K=K\left(f_{1}, \cdots, f_{m}\right)$ *is totally disconnected and therefore* $\dim_{\mathrm{T}}(K)=0$.

It will be interesting to consider a higher dimensional version of this theorem; is it true or not that, if $\sum_{j=1}^{m}\left(\operatorname{Lip}\left(f_{j}\right)\right)^{p}<1$, then $\dim_{T}(K) \leq p-1$? In connection with this, we have

Theorem (2.1.11)[62]: *Suppose* $X \subset R^{p}$ *and* $\left\{f_{j}\right\}_{1 \leq j \leq m}$ *is a finite set of contractions of* X *satisfying* $\sum_{j=1}^{m}\left(\operatorname{Lip}\left(f_{j}\right)\right)^{p}<1$. *Then Riemann's p-dimensional outer area of the set* $K\left(f_{1}, \cdots, f_{m}\right)$ *is zero. In particular, it also holds true for the p-dimensional* Lebesgue *measure*.

Proof. Consider a closed ball $B \subset R^{p}$ containmg the set $=K\left(f_{1}, \cdots, f_{m}\right)$. The outer area in the sense of Riemann of a bounded set Q will be denoted by $\bar{s}(Q)$. Then

$$s\mathrm{J}K) \leq \sum_{\backslash v \in W(n)} s\, \mathrm{V}_{w}(K)) \leq \sum_{\mathrm{v}v \in W(n)} s\, \mathrm{V}_{w}(B)\rangle \leq \bar{s}(B) \sum_{w \in W(n)}\left(\operatorname{Lip}\left(f_{w}\right)\right)^{p}$$

$$\leq \bar{s}(B)\left(\sum_{f=1}^{m}\left(\operatorname{Lip}\left(f_{j}\right)\right)^{p}\right)^{n}-0 \text{ as } n \rightarrow \infty,$$

where $f_{w}=f_{w_{1}} \circ \cdots \circ f_{w_{n}}$ for any $w=\left(w_{1}\, w_{n}\right) \in W(n)$. Hence $\bar{s}(K)=0$ as required.

Using the mapping $\psi: \Sigma \rightarrow K\left(f_{1}, \cdots, f_{m}\right)$ defined in Theorem (2.1.5), we can get two theorems concerning the topological stiuctures of the set K for weak contractions.

Theorem (2.1.12)[62]: *Suppose that* $\{f_j\}_{1\le j\le m}$ *is a finite set of one to one weak contractions of* X. *Suppose further that* $\mathrm{Fix}(f_i) \neq \Gamma\mathrm{ix}(f_j)$ *for some* $i \neq j$. *Then the set* $K = K(f_1, \cdots, f_m)$ *is perfect and therefore* K *is* $uncountaf_2le$.

Proof. Suppose, on the contrary, that $\psi(\alpha)$ is an isolated point of K for some $\alpha = (\alpha_n) \in \Sigma$. By the continuity of ψ, there exists a $\delta > 0$ such that $\psi(\alpha) = \psi(\beta)$ for any $\beta \in N_\delta(\alpha)$. Put $\overline{\alpha} = (\alpha_1 . \alpha_N iii \cdots)$ and $\underline{\alpha} = (\alpha_1 \cdots \alpha_N jjj \cdots)$ for a sufficiently large N. Then $\psi(\overline{\alpha}) = \psi(\alpha) = \psi(\underline{\alpha})$ implies $\mathrm{Fix}(f_i) = \Gamma\mathrm{ix}(f_j)$, contrary to the assumption.

Theorem (2.1.13)[62]: *Suppose that* $\{f_f\}_{1\le j\le m}$ *is a finite set of one to one weak contractions of* X. *Suppose further that* $\{f_j(K)\}_{1\le j\le m}$ *are pairwise disjoint where* $= K(f_1, \cdots, f_m)$. *Then the set* K *is totally disconnected andperfect* (*therefore* $\dim_{\mathrm{T}}(K) = 0$ *and* K *is uncountable*).

Example (2.1.14)[62]: Let $X = [0,1]$ with the usual Euclidean distance and

$$f_1(x) = ax \ and \ f_2(x) = b(x-1) + 1, \tag{16}$$

where $0 < a < 1$ and $0 < b < 1$ are two parameters. If $a + b < 1$, the set $K = K(f_1, f_2)$ is totally disconnected and perfect by Theorems (2.1.10) and (2.1.12). This also follows from Theorem (2.1.13). In particular, if $a = b = 1/3$, then K becomes Cantor's ternary set. On the other hand, if $a + b \geq 1$, it is clear that $K = [0,1]$ and therefore $\dim_{\mathrm{T}}(K) = 1$.

Remark (2.1.15)[62]: There exist two weak contractions f_1, f_2 of $X = [0,1]$ such that Lip $(f_1) = \mathrm{Lip}(f_2) = 1$ and $K(f_1, f_2)$ is totally disconnected and perfect. For example, put

$$f_1(\mathrm{x}) = \frac{x}{1+2x} \ and \ f_2(x\rangle = \frac{2-x}{3-2x}, \tag{17}$$

and apply Theorems (2.1.12) and (2.1.13). One can also construct f_1, f_2 for which $K(f_1, f_2)$ is totally disconnected, perfect and of positive measure.

Remark (2.1.16)[62]: There exists a finite set of contractions $\{f_j\}_{1\le j\le m}$, $m \geq 3$, satisfying $\sum_{j=1}^{m} \mathrm{Lip}\,(f_j) < 1$, for which the set $K(f_1, \cdots, f_m)$ is totally disconnected and perfect, and the mapping $\psi: \Sigma \to K$ is *not* a homeomorphism. For example, let $X = [0,1]$ and

$$f_1(\mathrm{x}) = \frac{x}{4}, f_2(x) = \frac{x}{4} + \frac{3}{5} \ and \ f_3(\mathrm{x}) = \frac{x}{4} + \frac{3}{4}. \tag{18}$$

In fact, $K(f_1, f_2, f_3)$ has the required properties by Theorems (2.1.10) and (2.1.12), while ψ is not a homeomorphism since Fix $(f_2) = F\mathrm{ix}(f_3 \circ f_1)$. One can easily construct such an example for any $m \geq 3$. This gives a counter-example for Williams'. Indeed, $m = 2$ is the only correct case and its proof will be given later.

We need some Definitions.

Definition (2.1.17)[62]: A set $Q \subset X$ is said to be *locally connected* at $p \in Q$ provided that for any neighborhood U of p, there exists a neighborhood V of p such the $Q \cdot V$ lies in a single component of $Q \cdot U$ containing p. A set Q which is locally connected at every point of Q is said to be locally connected. A finite sequence of points $\{p_1, \cdots, p_n\}$ is said to be an *ε-chain* joining p_1 and p_n provided that $d(p_i, p_{i+1}) < \varepsilon$ for any $1 \leq i \leq n-1$. A set $Q \subset X$ is said to be *well-chained* provided that for any $\varepsilon > 0$, any two points $p, q \in Q$ can be joined by an ε-chain of points all lying in Q. A finite sequence of subsets $\{Q_1, Q_n\}$ is said to be a *finite chain* joining Q_1 and Q_n provided that $Q_i \cdot Q_{i+1} \neq \varphi$ for any $1 \leq i \leq n-1$.

Theorem (2.1.18)[62]: *Let* $\{f_j\}_{1\le j\le m}$ be afinite set of weak contractions of X. *Then the set* $K = K(f_1, \cdots, f_m)$ *is a locally connected continuum if and only iffor any* $1 \leq i < j \leq m$, *there exists a sequence* $\{r_1, \cdots, r_n\} \subset \{1,2,,m\}$ *such that* $\{f_i(K), f_{r_1}(K),, f_{r_n}(K), f_j(K)\}$ is a finite chain.

Proof. It suffices to show the condition is sufficient. Let $K(w) = f_{v1} \circ \cdots \circ f_{w_n}(K) \in \mathcal{C}(X)$ for any $= (w_1 \cdots w_n) \in W(n)$. We first prove the following proposition by induction on k; for any finite words $u \neq v \in W(k)$, there exists a sequence $\{w^1, \cdots, w^n\} \subset W(k)$ such that $\{K(u), K(w^1),, K(w^n), K(v)\}$ is a finite chain. By assumption, this holds true for $k = 1$. Suppose next that this holds true for $k = l$. Then we must show this is also valid for $k = l + 1$. Suppose, on the contrary, that there exist u $\neq v \in W(l+1)$ for which there are no finite chains joining $K(u)$ and (v) . Put $W' =\{w \in W(l+1)$; there exists a finite chainjoining $K(u)$ and $K(w)\}$. Then $u \in W$ and $v \in W'' \equiv W(l+1) - W'$. Thus we have a separation

$$K = \sum_{w \in W'} K\,(w\rangle + \sum_{)v \in W''} K\,(w) \equiv K' + K'' \tag{19}$$

Therefore there exists a word $w^* \in W(l)$ satisfying $K(w^*) \cdot K' \neq (\beta \neq K(w^*) \cdot K''$ Since $K(w^*) = K(w^* \circ 1) + \cdots + K(w^* \circ m)^{\uparrow\rangle}$, there exist $i \neq j$ satisfying $K(w^* \mathrm{o} l) \cdot K' \neq \varphi \neq K(w^{*\circ} J) \cdot K''$ Now let $\{K(i), K(r_1),, K(r_n), K(j)\}$ be a finite chain joining $K(l)$ and $K(j)$. Then it is clear that $\{K(w^{*\circ} i), K(w^{*\circ} r_1), \cdots, K(w^{*\circ} r_n), K(w^* \circ J)\}$ is a finite chain. This implies $w^{*\circ} j \in W$ and therefore $K(w^{*\circ} j) \in K' \cdot K''$, contrary to (19). This completes the proof of our proposition.

Now for any $p, q \in K$, there exist w^p, $w^{\mathrm{q}} \in W(n)$ such that $p \in K(w^p)$ and $q \in K(w^{\mathrm{q}})$. Then by our proposition, there exists a finite chain $\{K(w^p), K(w^1),, K(w^n), K(w^{\mathrm{q}})\}$. Choose a finite sequence of points $\{s_j\}$ satisfying $s_1 \in K(w^p) \cdot K(w^1), \cdots, s_{n+1} \in K(w^n) \cdot K(w^{\mathrm{q}})$. Since diam $\big(K(w)\big) \leq \Omega^n\big(\mathrm{diam}(K)\big)$ for any $w \in W(n)$ where $(t) = \max_{1 \leq j \leq m} \Omega_{f_j}(t)$, the sequence $\{p, s_1, \cdot, s_{n+1}, q\}$ becomes an ε-chain for a sufficiently large n. Since ε is arbitrary, K is well-chained and therefore K is connected (Whyburn [114]). Note that $K(w)$ is also connected and for any $\varepsilon > 0$, the set K is the sum of a finite number of connected sets each ofdiameter less than ε. Hence K is locally connected [114]. This completes the proof.

Remark (2.1.19)[62]: There exists a set of contractions $\{f_n\}_{n \geq 1}$ of $X = R^2$ for which the set $K(f_1, f_2, \cdots)$ is not locally connected. For example, let $Q = Q_0 + \sum \;\; K_n$, where Q_0 is the square with vertices (0,0), (1,0), (1,· 1) and (0,1) , and K_n is the straight line interval from $(1/n, 0)$ to $(1/n, 1)$ for $n \geq 1$ (Fig. 2). Then one can easily construct

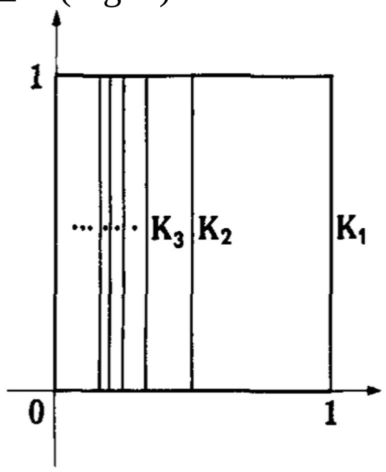


Fig. (2)[62]:

$\{f_n\}_{\geq 1}$ such that $Q = K(f_1, f_2, \cdots)$ using compositions of a dilation, a rotation, a translation and $J(x, y) = x/2$.

Theorem (2.1.18) raises the following question: is it true or not that for any locally connected continuum $Q \subset X$, there exists a finite set of weak contractions $\{f_j\}_{1\leq j\leq m}$ of X such that $Q = K(f_1, \cdots, f_m)$? In other words, is it possible to characterize locally connected continua by the self-similarity defined by (6)?
Note that for a fixed $m \geq 2$, there exists a locally connected continuum $Q \subset R^m$ for which $Q \neq K(f_1, \cdots, f_m)$ for any m weak contractions $\{f_j\}$ of R^m. For example, an $(m.-1)$-dimensional sphere in R^m has the required property by the LusternikSchnirelman-Borsuk theorem (Granas [76]).
Finally, combining Theorems (2.1.10) and (2.1.18), we have immediately
Corollary (2.1.20)[62]: (Williams' Theorem D for $m = 2$). *Let $f_1 and f_2$ be one to one contractions of X satisfying* $\mathrm{Lip}(f_1) + \mathrm{Lip}(f_2) < 1$ and $\mathrm{Fix}(f_1) \neq \mathrm{Fix}(f_2)$. *Then the mapping $\psi: \Sigma \rightarrow K(f_1, f_2)$ is a homeomorphism.*

To state further properties of self-similar sets, we need some Definitions.
Definition (2.1.21)[62]: A point p of a connected set Q is said to be a *cut point* of Q provided that $Q - p$ is the sum of two mutually separated sets; an *end point* of Q provided that there exist arbitranly small neighborhoods of p in Q each of whose boundanes consists of a single point. Two points p, q of a connected set Q are said to be *conjugate* provided that no points separate p and q in Q. If p is neither a cut point nor an end point of a connected set Q, the set consisting of p together with all points of Q conjugate to p is called a *simple link* of Q. A continuum Q is said to be an *acyclic curve* provided that it is locally connected and contains no simple links.
It is known that any simple link of a continuum Q is a *nondegenerate* continuum; that is, it contains more than one point ([114]). Every point of Q is either a cut point, an end point or a point of a single simple link of Q. We now state our main theorem.
Theorem (2.1.22)[62]: *Suppose that $\{f_j\}_{1\leq j\leq m}$ is a finite set of one to one weak contractions of X such that* $\mathrm{Fix}(f_i) \neq \Gamma\mathrm{ix}(f_j)$ *for some $j \neq j$. Suppose further that the set $K = K(f_1, \cdots, f_m)$ is an acyclic curve. Then either K is a simple arc or K has an infinite number of end points.*
Proof. Put $K(w) = f_{w'} \circ \cdots \circ f_{v'}(K)$ for any $= (w_1\ w_n) \in W(n)$. Suppose that K has a finite number of end points, say e^1, e^2, $\cdots$, e^N. Then it suffices to show $N = 2$, since a continuum is a simple arc if and only if it has exactly two non-cut points ([114]). Suppose, on the contrary, that $N \geq 3$. The remainder of the proof is devoted to demonstrating a contradiction.
1st Step. *There exists a finite word $w \in W(n)$ for some n for which every poim of $K(w)$ is a cut point of K.*
Proof. Suppose, on the contrary, that $K(u)$ contains at least one of the end points of K for any $\in W(n)$, $n \geq 1$. Take a sufficiently large integer n so that

$$diam\left(K(u)\right) \leq \Omega^n\left(\mathrm{diam}(K)\right) < \frac{1}{2}\min d\left(e^i, e^j\right), \quad (20)$$

for any $u \in W(n)$ where $\Omega(t) = \max_{1\leq j\leq m}\Omega_{f_j}(t)$. Obviously (20) contradicts the connectedness of K. Thus there exists a word $w \in W(n)$ possessing the required property. Put $F = f_{w^{\circ\cdots\circ}1 f_{\iota v_n}}$ and $p = \Gamma\mathrm{ix}(P]$ for brevity. Evidently $F(K) = K(w)$ has exactly N end points $\{F(e^j)\}$. Note that p is not an end point of (K) . For otherwise, $p = F(e^j)$ for some j; hence $p = e^j$, contrary to the above Definition of $K(w)$.

2nd Step. *There exists a simple arc $A_j$joining e andpfor $1 \leq j \leq N$such that $A_i \cdot A_f = p$ for any $i \neq j$.*

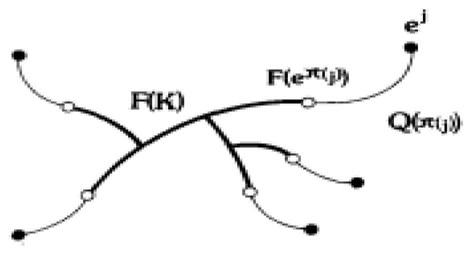
Proof. Since $F(e^j)$ is a cut point of K, we have a separation

$$K - F(e^{\mathrm{j}}) = P(j) + Q(\mathrm{j})\,, \tag{21}$$

where $\overline{P}(j) \cdot Q(J) = P(J) \cdot \overline{Q}(J) = \varphi$ and $P(J)$ contains the connected set $F(K) - F(e^J)$. Then there exists a non-cut point q^j of $Q(j)$ such that $q^j \neq F(e^J\rangle$ since $\overline{Q}(j) = Q(i) + F(e^J)$ is a nondegenerate continuum. Evidently q^i is an end point not only of $\overline{Q}(j)$ but also of K. We also have $q^i \neq q^j$ for any $i \neq j$ since $\overline{Q}(l) \cdot \overline{Q}(j) = \varphi$ for any $i \neq j$. Therefore $\overline{Q}(J)$ has exactly two end points q^j and (e') ; hence $\overline{Q}(j)$ is a simple arc. Thus this enables us to define the permutation π on the set {1, 2, ⋯, N} such that $e^j = q^{\pi(j)}$ (Fig. 3(a)) .
Now we define

$$S_j\left(n\right\rangle = \overline{Q}(\pi(j)) + F\left(\overline{Q}\langle\pi^2(j))\right) + \cdots + F^{n-1}\left(\overline{Q}(7\text{c}^{11}(j))\right). \tag{22}$$

Then (22) implies that $S_j(n)$ is a simple arc joining e^j and $F^n(e^{\pi^n(j)})$, and that $S_j(1) \leq S_j(2) \leq \cdots$ (Fig. 3(b)). Put $A_j = \lim S_j\,(n)$ in (X) . We first show that the set A_j is a simple arc. For otherwise, $p \in \overline{Q}(l)$ for some i; hence p is an end point of $F(K)$, contrary to the result in 1st Step. We next show that $A_i \cdot A_j = p$ for any $i \neq j$. For



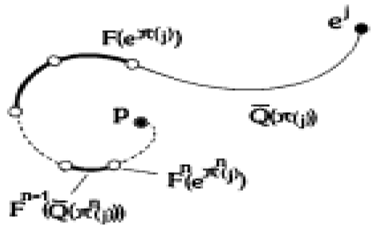


Fig. (3)[62]: (a) $F(K)$ is the heavy curve. The end points of K and $F(K)$ are indicated by · and O respectively.
(b) Simple arc $S_j(n)$.

otherwise, there exist two integers $r \geq s$ satisfying

$$F^{-1}\left(\overline{Q}(\pi^r(l))\right) \cdot F^{s-1}\left(\overline{Q}(\pi^s(j))\right) \neq \varphi. \tag{23}$$

Since $\overline{Q}(\pi^r(l)) \cdot \overline{Q}(\pi^r(j)) = \varphi$, we have $r > s$; hence $F^{r-s}\left(\overline{Q}(7\mathrm{f}(\iota))\right) \cdot \overline{Q}(\pi^s(J)) \neq \varphi$. Then it follows that $F(e^{\pi^s(\prime)}) \in F^{-s}\left(\overline{Q}(\pi^r(l))\right)$ since $(e^{\pi^s(j)}) = F(K) \cdot \overline{Q}(t(J))$. Hence $r - s = 1$ and $\pi^{r-1}(l) = \pi^s(j)$, contrary to $i \neq j$. Thus $A_i \cdot A_j = p$ for any $i \neq j$ as required. Note that $A_{Reject} + A_j$ is a simple arc joining e^i and e^{j} through the point p.

3rd Step. We are now ready to prove our theorem. Let f_s be one of the weak contractions $\{f_j\}$ satisfying $\neq \Gamma\text{ix}(f_s)$. Note that every point of$f_s \circ f\#K)$ is a cut point of K since $f_s\colon K \to K(s)$ is a homeomorphism. By the same arguments as in 1st and 2nd Steps, we conclude that for any $i \neq j$, there exists a simple arcjoining $e\hat{\imath}$ and e^j through the point $p' = \Gamma\text{ix}(f_s \circ F) \neq p$.

Consider now three end points e^1, e^2 and e^3. Since $A_{\hat{\imath}} \cdot A_j = p$ for any $i \neq j$, there exist at least two simple arcs, say A_1 and A_2, such that $p' \notin A_1 + A_2$. Thus we have two different simple arcs joining e^1 and e^2. This contradicts the fact that a locally connected continuum is an acyclic curve if and only if there exists a unique simple arc joining any two points ([114]). This completes the proof.

A finite sequence of sets $\{Q_1, Q_2, \cdots, Q_n\}$ is said to be a *regular chain* provided that $Q_i \cdot Q_{i+1}$ consists of exactly one point for any $1 \leq i \leq n-1$ whereas $Q_i \cdot Q_j = \varphi$ if $|i - j1 > 1$. Then we have

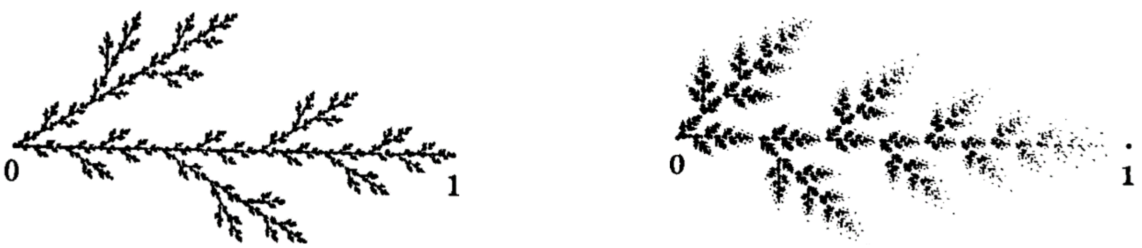
Theorem (2.1.23)[62]: *Let $\{f_j\}_{1\leq j\leq m}$ be a finite set of one to one weak contractions of X satisfying* Fix $(f_i) \neq \Gamma\text{ix}(f_j)$ *for some $i \neq j$. Let $K_j = f_j(K(f_1, \cdots, f_m))$, $1 \leq j \leq m$, for brevity. For any $i \neq j$, suppose that the set $K_i \cdot K_j$ consists of at most one point and that there exists a unique regular chain joining K_i and K_f. Then either $K = K(f_1, \cdots, f_m)$ is a simple arc or K has an infinite number of end points.*

Proof. By Theorems (2.1.22), it suffices to show that K is acyclic. Suppose, on the contrary, that K has a simple link. Since any two conjugate points of a locally connected continuum lie together on a Jordan closed curve [114], there exists a Jordan closed curve J in K such that $J \leq K(w)$ and $J \cdot f_{rv}(K'_r) \neq \varphi \neq J \cdot f_w(K'_s)$ for some $w \in W(n)$ and some $r \neq s$, where $K'_{\hat{j}} = K - \sum\ K_i$. Hence $J' = f_w^{-1}(J)$ satisfies $J' \cdot K'_r \neq \varphi \neq J' \cdot K'_s$, contrary to the assumption.

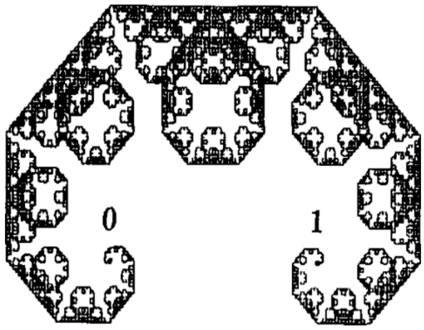
Example (2.1.24)[62]: Let $X = C$ with the usual Euclidean distance and put

$$f_1(z) = \alpha\bar{z} \text{ and } f_2(z) = |\alpha|^2 + (1 - |\alpha|^2)\bar{z}\,, \tag{24}$$

where α is a complex parameter satisfying $|\alpha| < 1$, $|1 - \alpha 1 < 1$ and $\text{Im}\,\alpha \neq 0$. Then it is easily seen that $K = K(f_1, f_2)$ is not a simple arc and that $K_1 \cdot K_2 = |\alpha|^2$; hence K has an infinite number of end points (Fig. 4(a) and (b)).



(a) $\alpha = \frac{1}{2} + \frac{\sqrt{3}}{6}i$. (b) $\infty = 0.3 + 0.3i$.



(c) Lčvy curve.

Fig. (4)[62]:

Example (2.1.25)[62]: There exist two contractions f_1, f_2 such that the set $K = K(f_1, f_2)$ has an infinite number of simple links. For example, let $X = C$ and put

$f_1(z) = \alpha z$ and $f_2(z) = (1-\alpha)z + \alpha,$ (25)

where α is a complex parameter satisfying $|\alpha| < 1$ and $|1-\alpha| <]$. It was pointed out by Lévy [89] that for $\alpha = 1/2 + i/2$, the measure of K is positive and that the set of multiple points of K is uncountable and dense in K (Fig. 4(c)).

We will discuss the parameterizations of.self-similar sets using some kind of functional equations. First of all, we have

Theorem (2.1.26)[62]: (de Rham [103]). Let f_1 and f_2 *be two contractions of* $X = R^p$. *Then the functional equation*

$$G(t) = \begin{cases} f_1(G(2t\rangle) & for\ 0 \le t \le \frac{1}{2}, \\ f_2\big(G(2t-1)\big) & for\ \frac{1}{2} \le t \le 1, \end{cases} \tag{26}$$

has a unique continuous solution if and only if

$$f_1\big(\mathrm{Fix}(f_2)\big) = f_2\big(\mathrm{Fix}(f_1)\big)\,. \tag{27}$$

Note that de Rham's theorem gives a parameterization of the set $K(f_1, f_2)$ if the condition (27) is fulfilled. Indeed, we have

$$G([0,1]) = G\left(\left[0,\frac{1}{2}\right]\right) + G\left(\left[\frac{1}{2},1\right]\right) = f_1(G([0,1]) + f_2(G([0,1])$$

and therefore $G([0,1]) = K(f_1, f_2)$ by Theorem (2.1.4).

We now generalize de Rham's Theorem (2.1.26). The following Definitions are essentially taken from Milnor-Thurston [96]:

Definition (2.1.27)[62]: A continuous function h of $[a,b]$ is said to be *piecewise.monotone* provided that the interval $[a,b]$ is subdivided into finite subintervals so that the restriction of h to each subinterval is strictly monotone.

Definition (2.1.28)[62]: For any function $H: [0,1] \to [0,]]$, define the mapping $v_H: [0,1] \to \Sigma$ by setting

$$v_H(t) = \big(A(t), A\big(H(t)\big), \cdots, A\big(H^n(t)\big), \cdots\big) \tag{28}$$

where $A(t) = [mt] + [1-t]$ for $0 \le \ell \le 1$. $v_H(t)$ is called the *itinerary* of a point t under H.

Note that v_H is discontinuous for any H since Σ is totally disconnected. However, for some kind of H, the mapping v_H is almost continuous' in the following sense.

Lemma (2.1.29)[62]: *Let* $h_j: [0-1)/m, j/m] \to [0,1]$ *be piecewise-monotone for any* $1 \le j \le m$. *Put* $H(t) = h_{A(t)}(t)$ *for brevity. Then there exist the limits* $\iota_H^1(s\pm)$ *in* Σ *for any* $0 < s < 1$. *Moreover* v_H *is continuous on*

$$= \{t \in [0,1]; fF(t) \ne j/m\ for\ any\ n \ge 0\ and\ 1 \le j \le m-1\}, \tag{29}$$

which is a dense set of [0,1].

Proof. For any fixed $s \in (0,1)$ and $N \geq 1$, there exists a sufficiently small $\varepsilon > 0$ such that each of the functions

$$H(t), H^2(t), \cdots, H^N(t) \tag{30}$$

is strictly monotone, either increasing or decreasing on $(s, s+\varepsilon)$ and that each of

$$A(t), \quad A\big(H(t)\big), \quad A\big(H^N(t)\big) \tag{31}$$

is independent of the choice of $t \in (s, s+\varepsilon)$. Obviously this implies that $v_H(s+)$ exists. Similarly $v_H(s-)$ exists for any $0 < s < 1$. Suppose now $s \in \Gamma_H$. Then it follows that each of the functions (31) is continuous in a sufficiently small neighborhood of s. Therefore $v_H(s+) = v_H(s-) = v_H(s)$; hence v_H is continuous at s. Since each h_j is piecewise-monotone, the set $\gamma_{n,j} = \{\mathrm{t}; fP(t) = j/m\}$ is finite; hence $\Gamma_H = [0,1] - \sum_{n,j} \gamma_{n,j}$ is obviously dense in [0,1].

Using this lemma, we can prove the following generalization of Theorem (2.1.26).

Theorem (2.1.30)[62]: *Let* $\{f_j\}_{1\leq j\leq m}$ *be a finite set of weak contractions of X and* $\{h_j\}_{1\leq j\leq m}$ *be the same functions as in Lemma* (2.1.29). *Then the functional equation*

$$G(t) = \begin{cases} f_1(G\big(h_1(t\rangle)\big) & for\ 0 \leq t \leq \dfrac{1}{m}, \\ f_m(G\big(h_m(t)\big)\rangle & for\ \dfrac{m-1}{m} \leq \mathrm{t} \leq 1, \end{cases} \tag{32}$$

has a unique continuous solution G: [0,1] → X *if and only if*

$$\psi \circ v_H\left(\frac{j}{m}+\right) = \psi \circ v_H\left(\frac{j}{m}-\right)\ for\ any\ 1 \leq j \leq m-1\ , \tag{33}$$

where $\psi: \Sigma \to K(f_1, \cdots, f_m)$ *is the mapping defined in Theorem* (2.1.5). *If in addition each* h_j *is onto, the continuous solution* G *of* ⟨(2.1.33)⟩ *satisfies* $G\ ([0,1]) = K\langle f_1, \cdots, f_m)$.

Proof. Obviously the condition (33) is necessary, since we have

$$G(t) = f_{A(t)} \circ Go H(t) = f_{A(\mathrm{t})} \circ f_{A(H(\mathrm{t}\rangle)} \circ \cdots = \psi \circ v_H(t)\,. \tag{34}$$

We now show the sufficiency. Put $F(t) = \psi_\circ v_H(t)$ for brevity. Then F is continuous on Γ_H by Lemma (2.1.29). The condition (33) implies $F(0/m)+) = F(0/m)-)$ for $1 \leq j \leq m-1$. Since $F(t) = f_{A(\mathrm{t})} \circ F\big(H(t)\big)$ for any t, it follows that $F(s+)$ as well as $F(s-)$ is equal to one of $f_{A(s)} \circ F(0/m) \pm)$ for any $s \in \gamma_{1,j}$. Therefore $(s+) = F(s-)$. Similarly one can show that $F(s+) = F(s-)$ for any $\in \gamma_{\mathrm{J}1,j}$, $l\mathrm{t} \geq 1$. Now define $\tilde{F}(t) = F(t)$ if $t \in \Gamma_H$ and $\tilde{F}(t) = F(t+)$ otherwise. Then it is easily seen that $\tilde{F}$ is continuous on [0,1]. Since $H(\Gamma_H) \subset \Gamma_\mathrm{H}$, we have

$$\tilde{F}(\mathrm{r}) = f_j\left(\tilde{F}\left(h_j(t)\right)\right)\ for\ t \in \Gamma_H \cdot \left(\frac{j-1}{m}, \frac{j}{m}\right).$$

Hence $\tilde{F}$ is a continuous solution of(32) since Γ_H is dense in [0,1]. The uniqueness of such a solution follows from (34). It is obvious that $G([0,1]) = K(f_1, \cdots, f_m)$ if each h_j is onto for the continuous solution G of (32). This completes the proof.

Applying the above theorem to the case $h_j = m\mathrm{t} - j + 1$, we have

Corollary (2.1.31)[62]: *Let* $\{f_j\}_{1\leq j\leq m}$ *be a finite set of weak contractions of X. Then the functional equation*

$$G(t)=\begin{cases} f_1(G(mt)) & for\ 0\le t\le \frac{1}{m}, \\ f_m(G(m\mathrm{t}-m+1)) & for\ \frac{m-1}{m}\le t\le 1, \end{cases} \qquad (35)$$

(35) *has a unique continuous solution if and only if*

$$f_2(\mathrm{Fix}(f_1)) = f_1\langle \Gamma\mathrm{ix}(f_m)) \dots f,(\mathrm{Fix}(f_1)) = f,\langle \Gamma\mathrm{ix}(f_m)). \qquad (36)$$

The continuous solution G of (35) gives a parameterization of $K(f_1,\cdots,f_m)$ since each h_j is onto. The conditions (36) are frequently referred to as the *D-conditions*. As applications of this kind of functional equations, Denny [71] gave an example of a uniformly continuous function $f: R^m \to (0,1)$ which is almost everywhere one to one; [79] showed the existence of periodic solutions of a certain functional equation, which are continuous and of bounded vanation.

Example (6.1.32)[62]: Consider the contractions defined by (24). Since $f_1^2(\mathrm{Fix}(f_2)) = f_2(\mathrm{Fix}(f_1))$, it is easily seen that the continuous solution G of (32) for

$$h_1(t) =] -|(2+\sqrt{2})t-1|\ and\ h_2(t) = 2t-1 \qquad (37)$$

gives a parameterization of the set $K = K(f_1, f_2)$ illustrated in Fig. 4(a) and (b). Note that h_1 has two fixed points (Fig. 5) and the set $G\left(\sum \quad H^{-n}\,(1)\right)$ gives all end points of K.

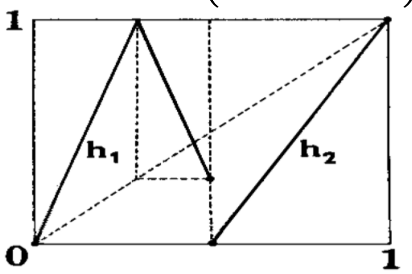


Fig. (5)[62]:

Finally we will study the case where G is a homeomorphism. Compare with Theorem (2.1.23).

Theorem (2.1.33)[62]: *Let* $\{f_j\}_{1\le j\le m}$ *be a finite set of one to one weak contractions of X satisfying* $\mathrm{Fix}(f_i\rangle \neq \Gamma\mathrm{ix}(f_j)$ *for some* $i \neq j$. *Suppose that the set* $\{K_1,\cdots,K_m\}$ *is a regular chain where* $K_j = f_j\ (K(f_1,\cdots,f_m))$ *for* $1 \le j \le m$. *Then the set* $K = K(f_1,\cdots,f_m)$ *is a simple arc if and only if there exist linear homeomorphisms* $h_j: [(i-])/m, j/m] \to [0,1], 1 \le j \le m$, *such that* $\psi \circ v_H$ *satisfies the condition* (33).

Proof. We first show the condition is necessary. Suppose K is a simple arc. Since each K_j is also a simple arc, the point $K_i \cdot K_{\mathrm{i}+1}$ is an end point of both K_i and K_{i+1}. Let g_j: $[0,1] \to K_j$ be a homeomorphism satisfying $g_j(1) = g_{j+1}(0)$ for $1 \le j \le m-1$. Then $G(t\rangle = g_{A\langle\mathrm{t}\rangle}(mt - A(t) + 1\rangle: [0,1] \to K$ becomes a homeomorphism. Define

$$h_j(\mathrm{r}) = G^{-1}\circ f_f^{-1}\circ G(t)\ for\ \frac{j-1}{m} \le t \le \frac{j}{m}.$$

Then obviously $h_j: [(j-1)/m, j/m] \to [0,1]$ is a homeomorphism and G satisfies the equation (32); hence $\psi \circ v_H$ satisfies the condition (33). It is obvious that each h_f can be replaced by a linear homeomorphism $\tilde{h}_j$ such that

$$\tilde{h}_j\left(\frac{j-1}{m}\right) = h_j\left(\frac{j-1}{m}\right)\ and\ \tilde{h}_j\left(\frac{j}{m}\right) = h_j\left(\frac{j}{m}\right).$$

We now show the sufficiency. It suffices to show the solution G of (32) is a homeomorphism. Suppose, on the contrary, that $G(t_1) = G(t_2)$ for some $t_1 < t_2$. Let $\Xi = \{(s,t); G(s) = G(t) \text{ for } 0 \le s, t \le 1\}$. Without loss of generality, we can assume

$$|t_1 - t_2| = \max_{(s,t)\in\Xi} |s-t|. \quad (38)$$

Then $(t_1) < A(t_2)$. For otherwise, we have $\big(H(t_1), H(t_2)\big) \in --$ and $|H(t_1) - H(t_2)| = m|\mathrm{t}_1 - t_2|$, contrary to (38). Since $\{K_1,, K_m\}$ is a regular chain, it follows that $A(t_1) = A(t_2) - 1$, say l. Thus $G(t_1) = G(t_2) = G(l/m) = K_l \cdot K_{l+1}$. Then $(t_1, l/m) \in \Xi$ implies $\big(h_l(t_1), h_\downarrow(l/m)\big) \in \Xi$ and $|h_1(t_1) - h_l(l/m)| = m|t_1 - l/m|$; hence $t_2 - t_1 \ge l - mt_1$. Similarly $(t_2, l/m) \in \Xi$ implies $\big(h_{l+1}(t_2), h_{l+1}(l/m)\big) \in \Xi$and $|h_{l+1}(t_2) - h_{l+1}(l/m)| = m|t_2 - 1/m|$; hence $t_2 - t_1 \ge mt_2 - l$. Combining two inequalities, we have $t_1 \ge l/m$ for $m \ge 3$, contrary to $A(\mathrm{t}_1) = t$. For the case $m = 2$, it is easily seen that $t_1 + t_2 = 1$ and $F\mathrm{ix}(f_1) = F\mathrm{ix}(f_2)$, contrary to the assumption. This completes the proof.

Example (6.1.34)[62]: Let $X = C$ with the usual Euclidean distance and

$$f_1(z) = \alpha\bar{z} \; and \; f_2(z) = (1-\alpha)\bar{z} + \alpha \,, \quad (39)$$

where α is a complex parameter satisfying $|\alpha| < 1$ and $|1 - \alpha_1 < 1$. Since $\{f_1, f_2\}$ satisfies the D-condition (36) and $K_1 \cdot K_2 = \alpha$ for any $|\alpha - 1/2| < 1/2$, it follows that $K = K(f_1, f_2)$ is a simple arc by Theorem (2.1.33); hence $\dim_\tau(K) = 1$. Note that Riemann's outer area of K is always zero by Theorem (2.1.11). Compare with the examples given by Osgood [100] and by Besicovitch-Schoenberg [64], which are simple arcs with positive area. On the other hand, if $|\alpha - 1/2| \ge 1/2$, it is clear that $K(f_1, f_2)$ is a closed triangle with vertices 0,1 and α (Fig. $6(\mathrm{a})$); therefore $\dim_T(K) = 2$. It was pointed out by de Rham [103] that for $\alpha = 1f2 + \sqrt{3}i/6$, the solution $G\mathrm{of}((2.1.26)0)$ gives the curve studied by von Koch [86] (Fig. 6(b)) and that for $\alpha = 1/2 + e^{\downarrow\theta}/2$, G gives the space-filling curve studied by Pólya [102].

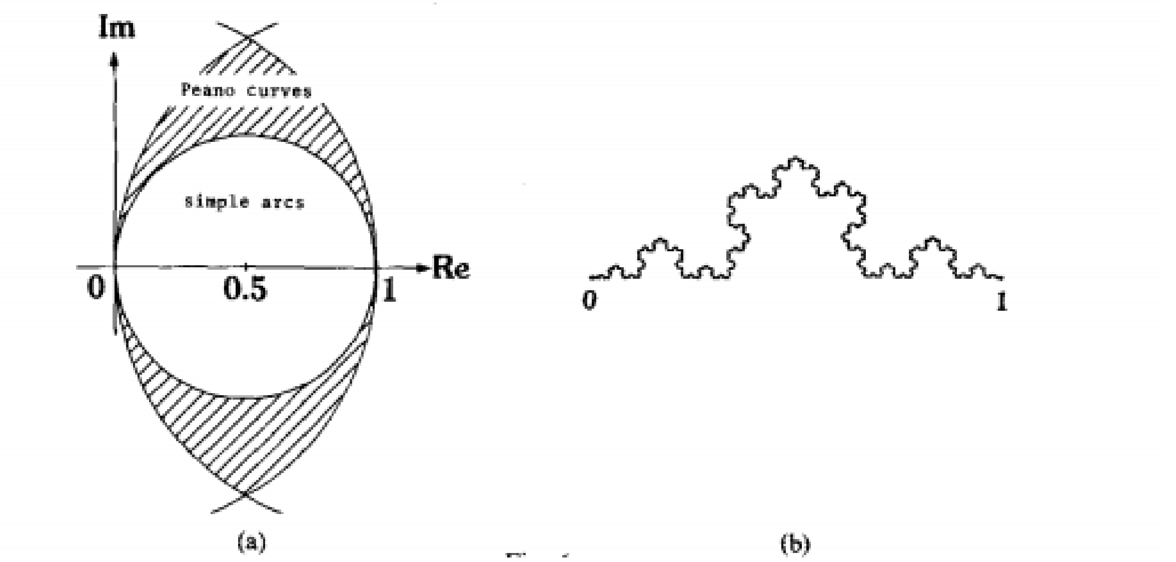


Fig. (6)[62]:

We will discuss the regularity of the continuous solution G of (35). Let X be a closed subset of a Banach space E and $\{f_j\}_{1\le j\le m}$ be a finite set of weak contractions of X satisfying $I \supset$-conditions (36). First of all, we have

Theorem (2.1.35)[62]: *The solution G is Hölder-continuous with exponent $\alpha = -\log\delta/\log m$, where $\delta = \max_{1\leq j\leq m}\mathrm{Lip}(f_j)$.*

Proof. For any t $\neq s$, let n be an integer such that $m^{-n-1} < |\ell - s| \leq m^{-n}$. Then it is easily seen that (t) , $G(s) \in K(w) + K(w')$ with $K(w) \cdot K(w') \neq \varphi$ for some $w, w' \in W(n)$. Therefore $\|G(t) - G(s)\| \leq \mathrm{diam}(K(w)) + \mathrm{diam}(K(w'))$; hence

$$\frac{\|G\langle t) - G(s)\|}{|t - s1^{\alpha}} \leq 2m^{\alpha}\mathrm{diam}(K(\delta m^{\alpha})^n \leq 2m^{\alpha}\mathrm{diam}(K).$$

A mapping $f: [0,1] \to E$ is said to be of *bounded p-variation* provided that

$$\sup\left(\sum_i \| f(t_{i+1}) - f(t_l)\|^p\right)^{1/p} < \infty, \tag{40}$$

where the supremum extends over all subdivisions $\Delta: 0 = t_0 < t_1 < \cdots < t_n = 1$ of [0,1]. For $p = 1$, we usually say that f is of bounded variation. Note that every Hölder-continuous mapping with exponent α is of bounded $1/\alpha$-variation. Then

Theorem (2.1.36)[62]: *Suppose that each f_j is one to one and* Fix $(f_1) \neq \Gamma\mathrm{ix}(f_m)$. *If $\{f_j\}$ satisfies for some $\alpha > 0$,*

$$\sum_{j=1}^{m}\left(\mathrm{Lip}(f_j^{-1})\right)^{-\alpha} > 1, \tag{41}$$

then the solution G is not of bounded α-variation.

Proof. Let $(n,J) = \|G(J/m^n) - G((j-1)/m^n)\|$, $1 \leq j \leq m^n$, $n \geq 1$ for brevity. Then it is easily verified that

$$\sum_{j\Rightarrow 1}^{m^n}\left(v(n,j)\right)^{\alpha} \geq \left(\cdot\sum_{=1}^{m}\left(\mathrm{Lip}(f_j^{-1})\right)^{-a}\right)\cdot\sum_{j=1}^{m^{n-1}}\left(\iota(n-1,j)\right)^{\alpha}$$

Therefore, by (41), it follows that G is not of bounded α-variation, since $v(0,1) = \|\Gamma\mathrm{ix}(f_1) - \Gamma\mathrm{ix}(f_m)\| \neq 0$.

We now turn to the differentiability of G. In this respect, we have the following theorem by applying the same method as in Lax [88].

Theorem (2.1.37)[62]: *Suppose that $\{f_j\}$ satisfies*

$$\prod_{j\Rightarrow 1}^{m}\mathrm{L\,ip}(f_j) < m^{-m}. \tag{42}$$

Then the Fréchet derivative $DG(\mathrm{t})$ of the solution G is equal to zero for almost every t.

Proof. Since almost every number is normal in the scale of m (Billingsley [4]), it suffices to show that $DG(t) = 0$ for every normal number $t \in [0,1]$. Let $s \neq t$ be an arbitrary number and let $t = \sum_{n\geq 1} t_n\, m^{-n}$ and $s = \sum_{n\geq 1} s_n\, m^{-n}$. Let $N \geq 1$ be the smallest integer such that $t_N \neq s_N$ and let $M > N$ be the smallest integer such that $\mathrm{t}_M \geq 1$ or $t_M \leq m - 2$ according to whether $t > s$ or $t < s$ respectively. Then it is easily verified that

$$m^{-M} < |s - t| < m^{-N+1} \; and \; M = N + o(N) \; as \; N \to \infty. \tag{43}$$

Note that (43) implies $s \to t$ if and only if $N \to \infty$.

On the other hand, we have from the equation (35),

$$\|G(s) - G(t)\| \leq \mathrm{diam}(K)\left(\cdot\prod_{=1}^{m} a_J^{r_j}\right), \tag{44}$$

where $a_j = \mathrm{Lip}(f_j\rangle$ and $r_j = \#\{] \leq l' \leq N-1; t_i = j\}$ for $\mathrm{I} \leq j \leq m$. Since $r_f = N/m + o(N)$ as $N \to \infty$, we have from (44),

$$\| \frac{G(s) - G(t)}{s - \mathrm{r}} \| \leq \mathrm{diam}(K) \exp \left(\frac{N}{m} \log \left(m^m \prod_{j=1}^{m} a_j \right) + o(N) \right).$$

Since $m^m \prod_{j=1}^{m} a_j < 1$, it follows that $DG(t) = 0$. This completes the proof.

Corollary (2.1.38)[62]: *Suppose that each f_j is a strictly monotone increasing function and* Fix $(f_1) < \Gamma\mathrm{ix}(f_m)$. *Suppose further that $\{f_\mathrm{J}\}$ satisfies* (42). *Then the solution G is a strictly monotone increasing and purely singular function.*

Example (2.1.39)[62]: Consider the contractions defined by (16). If $a + b = 1$ (this is also a special case of (39)) and $a \neq 1/2$, $\{f_1, f_2\}$ satisfies the conditions of Corollary (2.1.38); therefore $G_a(t) = G(t)$ is a strictly monotone increasing and purely singular function with a parameter a. This function was studied by Salem [108]. It is known that $G_a(t)$ is the distribution function for the Bernoulli trials of unfair coin tossings. See also Lomnicki-Ulam [90] and de Rham [103], [104].

Concerning the non-differentiability of G, we have

Theorem (2.1.40)[62]: *Suppose that each f_j is one to one and that $\{f_j\}$ satisfies*

$$\prod_{j=1}^{m} \mathrm{L\,ip}(f_j^{-1}) < m^m \qquad (45)$$

Then the solution G is not Fréchet differeniable at almost every t. If in addition Lip $(f_j^{-1}) < m$ *for any* $1 \leq j \leq m$, *then G is nowhere differentiable.*

Proof. We first show the non-differentiability of G at every normal number t. Let $t = \sum_{n\geq 1} t_n m^{-n}$. For any $N \geq 1$, take a suitable number $s_N \in [0,1]$ such that $\|G(s_N) - G(H^N(t))\| \geq (1/2)\mathrm{diam}(K)$ where $H(t) = mt - A(t) + 1$. Put $t^{(N)} = \sum_{j=1}^{N} t_\mathrm{j} m^{-j} + s_N m^{-N}$. Then from the equation (35),

$$\|G(t^{(N)}) - G(t)\| \geq \frac{1}{2} \mathrm{diam}(K) \prod_{j=1}^{m} b_J^{r_j}, \qquad (46)$$

where $b_j = \left(\mathrm{Lip}(f_j^{-1})\right)^{-1}$ and $r_j = \#\{1 \leq i \leq N, t_i{=}j?$ for $1 \leq j \leq m$. Since $|t^{(N)} - t| \leq 2m^{-N}$, we have

$$\| \frac{G(t^{(N)}) - G(t\rangle}{t^{(N\rangle} - \mathrm{t}} \| \geq \frac{1}{4} \mathrm{diam}(K) \exp \left(\frac{N}{m} \log \left(m^m \prod_{j=1}^{m} b_j \right) + o(N) \right).$$

Hence (45) implies that G is not differentiable at t.

Next assume that $mb_j > 1$ for $1 \leq j \leq m$ instead of (45). Then the same argument as above can be applied to an arbitrary t, since

$$\prod_{j=1}^{m} b_{J'}^{r_{J^2}} \geq b_*^N,$$

where $b_* = \min_{1\leq j\leq m} b_j > 1/m$; hence

$$\| \frac{G(\mathrm{t}^{\langle N\rangle}) - G(t)}{t^{(N)} - \mathrm{t}} \| \geq \frac{1}{4} \mathrm{diam}(K)(mb_*)^N$$

This completes the proof.

Example (2.1.41)[62]: Consider the contractions defined by (39). Then, by Theorem (2.1.35), the solution $G_\alpha(t) = G(t)$ of (35) has Hölder-exponent

$$-\frac{\log\ \max\left(|\alpha|, |1-\alpha|\right)}{\log 2}.$$

In particular, Koch's curve $\left(\alpha = 1/2 + \sqrt{3}i/6\right)$ is Hölder-continuous with exponent $\log 3/\log 4$, which can not be replaced by any larger value by Theorem (2.1.36). For almost every t, $G'_\alpha(t) = 0$ or $G_\alpha(t)$ is not differentiable according to whether $|\alpha(1-\alpha)| < 1/4$ or $> 1/4$ by Theorems (2.1.37) and (2.1.40). Note that the boundary curve $|\alpha(1-\alpha)| = 1/4$ is a lemniscate (Fig. 7). Moreover, if $|\alpha| > 1/2$ and $|1-\alpha| > 1/2$, then $G_\alpha(t)$ is nowhere differentiable, as shown by de Rham [104]. For Pólya's case $\left(\alpha = 1/2 + e^{t\theta}/2\right)$, the above results were shown by Lax [88],

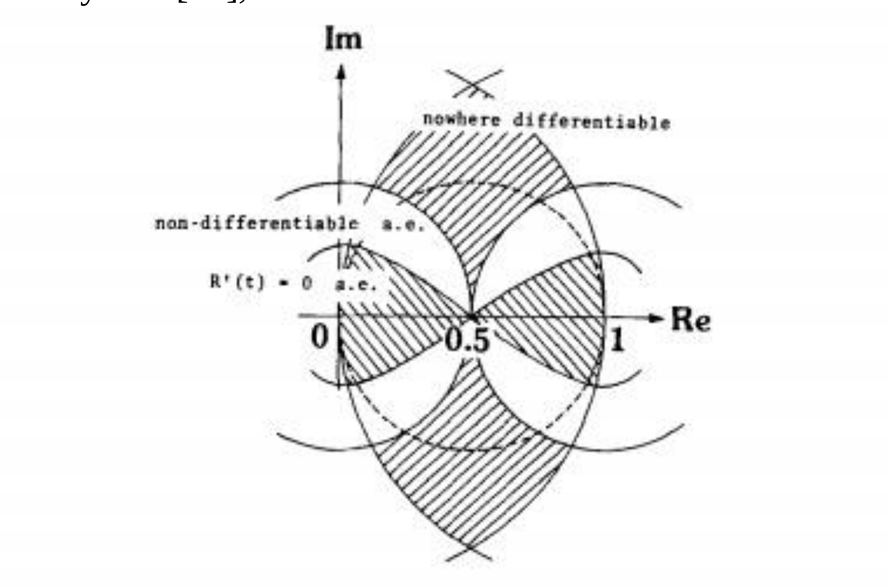
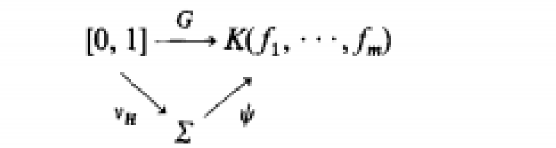


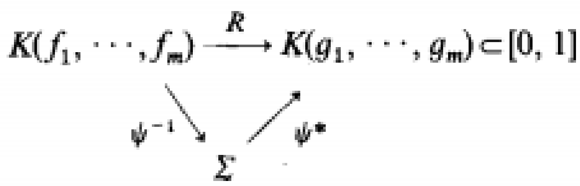
Fig. (7)[62]:

We obtained the continuous solution G of (32) using the diagram:



(47)

Such a solution does not exist if $K = K(f_1, \cdots, f_m)$ is not connected. Here we will discuss the existence of a non-trivial continuous mapping R which maps $K(f_1, \cdots, f_m)$ into $[0,1]$. Let $\{g_j\}_{1\le j\le m}$ be a finite set of weak contractions of $X = [0,1]$ with the usual Euclidean distance and $*= \psi: \Sigma \to K(g_1, \cdots, g_m)$. Then the desired mapping R will be obtained by the diagram:



(48)

Indeed, we have

Theorem (2.1.42)[62]: *Let* $\{f_j\}_{1\le \bar{j}\le m}$ *and* $\{g_j\}_{1\le j\le m}$ *be two finite sets of weak contractions of* X *and* $[0,1]$ *respectively. Then the functional equations*

$$\begin{cases} R\big(f_1(x)\big) = g_1\big(R(x)\big) \\ \vdots \qquad\qquad \vdots \\ R\big(f_m(x)\big) = g_m\big(R(x)\big) \end{cases} \tag{49}$$

have a unique continuous onto solution $R: K(f_1, \cdots, f_m) \to K(g_1, \cdots, g_m)$ *if and only if* $\psi^*(\alpha) = \psi^*(\beta)$ *whenever* $\psi(\alpha) = \psi(\beta)$.

Proof. It is clear that the condition is necessary, since

$$R\big(\psi(\alpha)\big) = R \circ f_\alpha \circ \psi\big(\sigma(\alpha)\big) = g_{\alpha_1^\circ}\psi\big(\sigma(\alpha)\big) = g_\alpha \circ g_{\alpha_z} \circ \ldots = \psi^*(\alpha) \tag{50}$$

for any $\alpha \in \Sigma$, where $\sigma: \Sigma \to \Sigma$ is the left-shift transformation.

We now show the sufficiency. Define the mapping R: $K(f_1, \cdots, f_m) \to K(g_1, \cdots, g_m)$ by $R\big(\psi(\alpha)\big) = \psi^*(\alpha)$. The condition of the theorem implies that R is well-defined. Then it is clear that R satisfies the equations (49). We must show the continuity of R. Suppose, on the contrary, that R is discontinuous at $\psi(\alpha)$ for some $\alpha \in \Sigma$. Then there exists a sequence $\{\alpha^{(n)}\}$ in Σ such that

$$|\psi^*(\alpha) - \psi^*\big(\alpha^{(n)}\big)| \geq \delta > 0 \tag{51}$$

and $\psi\big(\alpha^{(n)}\big) \to \psi(\alpha)$ as $n \to \infty$. Without loss of generality, we can assume $\alpha^{(n)} \to \overline{\alpha}$ as $n \to \infty$. Then we have $\psi(\overline{\alpha}) = \psi(\alpha)$ and therefore $\psi^*(\overline{\alpha}) = \psi^*(\alpha)$, contrary to (51). The uniqueness of such a solution is obvious from (50).

As a corollary, we have immediately

Corollary (2.1.43)[62]: *Let* $\{f_j\}_{1 \leq j \leq m}$ *be a finite set of weak contractions of* X *such that* $\{f_j(K)\}_{1 \leq j \leq m}$ *are pairwise disjoint where* $K = K(f_1, \cdots, f_m)$. *Then for any weak contractions* $\{g_f\}_{1 \leq j \leq m}$ *of* $[0,1]$, *the reversed equations* (49) *have a unique continuous onto solution* $R: K(f_1, \cdots, f_m) \to K(g_1, \cdots, g_m)$.

Example (2.1.44)[62]: Consider the contractions f_1, f_2 defined by (16) and put

$$g_1(t) = \frac{t}{2} \, and \; g_2(t) = \frac{t+1}{2}. \tag{52}$$

If $+b < 1$, the mapping $\psi: \Sigma \to K(f_1, f_2)$ becomes a homeomorphism by Theorem (2.1.13). Then there exists a unique continuous onto solution $R_{a,b} = R$: $K(f_{1^7}, f_2) \to K(g_1, g_2) = [0,1]$ by Corollary (2.1.43). Note that $R_{a,b}$ is monotone increasing and there exists a unique extension $\tilde{R}_{a,b}: [0,1] \to [0,1]$ of $R_{a.b}$, which is also monotone increasing and satisfies the equations (49) for *any* $x \in [0,]]$. In particular, if $a = b = 1/3, \tilde{R}_{a,b}(t)$ is the well-known Cantor function. The functional equations for the Cantor function were studied by Sierpin'ski [109]. Note that, if $= b(< 1/2)$, it is easily seen that

$$L_a(t) \equiv \int_0^1 e^{itx}\, d\tilde{R}_{a,a}(t) = e^{\iota i/2} \prod_{n \geq 0} \cos\left(\frac{1-t}{2} a^n\right). \tag{53}$$

It is known that $L_a(t)$ is not absolutely continuous (Kershner-Wintner [85]). Carleman [67] has shown that $L_a(t)$ does not tend to 0 as $|t| \to \infty$, if $a = q^{-1}$, where $q = 3,4,5$, Kershner [84] has shown that $L_a(t) = O\big((\log|t|)^{-\beta}\big)$ if $a = p/q$, not the reciprocal of an integer, while β is a positive function ofp and q. Note that this gives an example of a continuous function which is not absolutely continuou s and satisfies the Riemann-Lebesgue lemma. See also Erdös [73].

Example (2.1.45)[62]: De Rham [105] gave an example of a C^1-function $f(x, y)$ with two variables such that the set

$$f\left(\left\{(x,y);\frac{\partial f}{\partial \mathrm{x}}=\frac{\partial f}{\partial y}=0\right\}\right)$$

contains an interval, which is analogous to Whitney's example [113]. De Rham's function f is an extension of the solution R of (49) for certain affine contractions of the plane satisfying the condition of Corollary (2.1.43).

Example (2.1.46)[62]: If the continuous solution G of (32) is a homeomorphism and each h_{j}^{-1} is a weak contraction of [0,1], then it is clear that $R=G^{-1}$: $K(f_1,\cdots,f_m)\to[0,1]$ satisfies the equations (49) for $g_j=h_j^{-1}$, $]\le j\le m$. For example, let $X=[0,1]$ with the usual Euclidean distance and put

$$f_1(x)=\frac{\chi}{1+x}\ \text{ and }\ f_2(x)=\frac{1}{2-x}. \tag{54}$$

Then it is easily seen that the solution $R=G^{-1}:[0,1]\to[0,1]$ of (49) exists for the contractions g_1, g_2 defined by (52), which is known as Minkowski's function [97]. It was proved by Denjoy [69] that $R(t)$ is purely singular. See also Salem [108].

We will discuss vanous properties of the classical space-filling curves, which will be obtained by the continuous solution G of the equation (32) for certain simple affine contractions.

We denote by I^p the p-dimensional cube given by $[0,1]^p$. The following theorem is a standard result. For the proof, see Vitushkin-Khenkin [51].

Theorem (2.1.47)[62]: *Suppose that $p<q$ and $f: I^{\mathrm{p}}\to 1^q$ is an onto Hölder-continuous mapping with exponent α. Then $\alpha\le p/q$. Moreover, there exists an onto Höldercontinuous mapping $f: I^p\to I^{\mathrm{q}}$ with exponent $p/q-\varepsilon$ for any $\varepsilon>0$. If in addition p divides q, then one can take ε to be zero.*

Example (2.1.48)[62]: In 1890, Peano [101] gave the first example ofa continuous planar curve $P_1(t)$ filling the unit square P with vertices $0,1,1+i$ and i. It is easily seen that $P_1(t)$ is a continuous solution of (35) for the nine affine contractions:

$$\begin{cases} f_1(z)=\dfrac{z}{3}; & f_2(z)=-\dfrac{\overline{z}}{3}+\dfrac{1+i}{3};\ f_3(z)=\dfrac{z}{3}+\dfrac{2i}{3};\\ f_4(z\rangle=\dfrac{\overline{z}}{3}+\dfrac{1+3i}{3}; & f_5(z)=-\dfrac{z}{3}+\dfrac{2+2i}{3};f_6(z)=\dfrac{\overline{z}}{3}+\dfrac{1+i}{3};\\ f_7(z\rangle=\dfrac{z}{3}+\dfrac{2}{3}; & f_8(z)=-\dfrac{\overline{z}}{3}+\dfrac{3+i}{3};\ f_9(z)=\dfrac{z}{3}+\dfrac{2+2i}{3}. \end{cases} \tag{55}$$

Then, it follows that $P_1(t)$ is nowhere differentiable by Theorem (2.1.40) and satisfies

$$\|P_1(t)-P_1(s)\|\le 3\sqrt{5}|t-s|^{1/2} \tag{56}$$

Note that the exponent 1/2 in (56) can not be replaced by $1/2+\varepsilon$ for any $\varepsilon>0$ by Theorem (2.1.47). This a] so follows from Theorem (2.1.36). Cesàro [68] gave the analytic formula for P_1 and Moore [98] discussed a generalization of P_1 by geometrical observation. Using Moore's construction, Milne [95] gave an example of a mapping $f: I^1\to I^p$, which is Hölder-continuous with exponent p^{-1} and measure-preserving, that is, $\mu_p(A)=\mu_1\left(f^1(A)\right)$ for any Borel subset A of P where μ_p is the usual product measure on 1^{p}.

Example (2.1.49)[62]: In 1891, Hilbert [81] gave a simpler example of a continuous planar curve $P_2(t)$ filling P. It is easily seen that $P_2(t)$ is a continuous solution of (35) for the four affine contractions:

$$\begin{cases} f_1(z\rangle = \frac{i}{2}\overline{z}; & f_2(z) = \frac{Z}{2} + \frac{i}{2}; \\ f_3(z) = \frac{z}{2} + \frac{1+i}{2}; & f_4(z) = -\frac{i}{2}\overline{z} + \frac{2+i}{2}. \end{cases} \tag{57}$$

Then $P_2(t)$ is nowhere differentiable and satisfies

$$\|P_2(t) - P_2(s)\| \leq 2\sqrt{5}|t-s|^{1/2} \tag{58}$$

Example (2.1.50)[62]: Sierpiński [110] gave a slightly different example of a planar curve $P_3(t)$ filling the square with vertices $1+i, -1+i, .-1-i$ and $1-i$. $P_3(t)$ is a unique continuous periodic solution with period 1 of the equation (32) for the four contractions:

$$\begin{cases} f_1(z) = \frac{i}{2}(z-1-\iota)\,; & f_2(z) = \frac{1}{2}(z-1-l)\,; \\ f_3(z) = -\frac{i}{2}(z-1-i); & f_4(z) = -\frac{1}{2}(z-1-i) \end{cases} \tag{59}$$

and $h_j(t) = 4t - 1/8 (\mathrm{mod}\ 1)$ (Fig. 8).

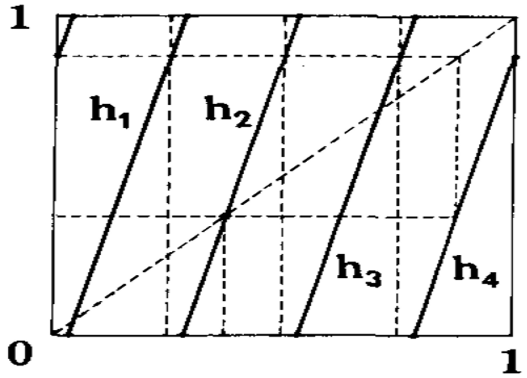


Fig. (8)[62]:

Note that $\{f_j\}$ satisfies

$$f_1 \circ f_4(\mathrm{Fix}(f_2)) = f_2 \circ f_4(\mathrm{Fix}(f_2)) = f_3 \circ f_4(\mathrm{Fix}(f_2)) = f_4^2(\mathrm{Fix}(f_2))\,.$$

We have $\sum_{j=1}^{m}\left(\mathrm{Lip}(f_j)\right)^2 = 1$ in all examples.

We begin with some Definitions.

Definition (2.1.51)[62]: For any $\alpha > 0$ and $U \subset X$, we shall denote, for each $\varepsilon > 0$, by $\Lambda_\alpha^\varepsilon(U)$ the lower bound of the sum $\sum_{n\geq 1}(\mathrm{diam}(S_n))^\alpha$ where $\{S_n\}_{n\geq 1}$ is an arbitrary covering of U consisting of closed spheres of diameters less than ε. When $\varepsilon \to 0+, \Lambda_\alpha^\varepsilon(U)$ tends to a unique limit $\Lambda_\alpha(U)$ (finite or infinite), which we shall call the *α-dimensional outer measure*. Then there exists a uniquely determined number such that

$$\sup\{\alpha; \Lambda_\alpha(U) = \infty\} = \inf\{\alpha; \Lambda_\alpha(U) = 0\},$$

which we shall call the *Hausdorff dimension* of U and denote by $\dim_H(U)$.

The function of a set $\Lambda_\alpha(lf)$ thus defined is an outer measure in the sense of Carathéodory. It is known that every Borel set is measurable and every set is regular with respect to this measure (Saks [107]).

First of all, we have

Theorem (2.1.52)[62]: Suppose that $\{f_j\}_{1\leq j\leq m}$ is a finite set of weak contractions of X. Then $\dim_H(K(f_1, \cdots, f_m)) \leq \lambda$ where λ is given by

$$\sum_{j=1}^{m}\left(\mathrm{Lip}(f_j)\right)^\lambda = 1. \tag{60}$$

Proof. Fix an arbitrary $\kappa > \lambda$. Consider a closed sphere S containing the set $= K(f_1, \cdots, f_m)$. Put $f_w = f_{w_1^{\circ}} \cdots \circ f_{w_n}$ for any word $w = (w_1 \cdots w_n)$. Then we have

$$\Lambda_\kappa^\varepsilon(K) \leq \sum_{weW(n)} \Lambda_K^e\left(f_w(K)\right) \leq, \sum_{w\in W(n)} \Lambda_\kappa^\varepsilon\left(f_v(S)\right) \leq (2\text{diam}(S))^\kappa \left(\sum_{j=1}^{m}\left(\text{Lip}(f_j)\right)^\kappa\right)^n,$$

where $\varepsilon = 2\Omega^n\big(\text{diam}(S)\big)$ and $\Omega(t) = \max_{1\leq j\leq m} \Omega_{f_j}(t)$. Taking the limit as $n \to \infty$, it follows that $\Lambda_\kappa(K) = 0$. This completes the proof.

Theorem (2.1.53)[62]: Suppose that $\{f_j\}_{1\leq j\leq m}$ is a finite set of one to one weak contractions of X. Suppose further that $\{f_j(K)\}_{1\leq j\leq m}$ are pairwise disjoint where $= K(f_1, \cdots, f_m)$. Then $\dim_H(K) \geq \lambda$ where λ is given by

$$\sum_{j=1}^{m}\left(\text{Lip}\left(f_{\text{j}}^{-1}\right)\right)^{-\lambda} = 1. \tag{61}$$

Proof. Fix an arbitrary $\kappa < \lambda$. By assumption, we have dist $\big(f_i(K), f_j(K)\big) \geq p > 0$ for any $i \neq j$. Consider now an arbitrary closed sphere S satisfying $S \cdot K \neq \phi$.

Suppose first that $S \cdot K$ consists of more than Qne point. Then there exist an integer $n = n(S) \geq 0$ and a word $w = w(S) \in W(n)$ such that $S \cdot K \leq f_w(K)$ and $S \cdot f_{w.\text{i}}(K) \neq \varphi \neq S \cdot f_{w\circ j}(K)$ for some $i \neq j$. Note that diam $(S) \geq a_{w1} \cdots a_{1\phi_n}\rho$ where $a_j = \left(\text{Lip}\left(f_j^{-1}\right)\right)^{-1}$ for $1 \leq j \leq m$.

Suppose next that $S \cdot K$ consists of exactly one point. Then we can take a sufficiently large integer $n = n(S)$ and a word $w = w(S) \in W(n)$ such that $S \cdot K = S \cdot f_{1v}(K)$ and $\text{diam}(S) \geq a_v 1 \ldots a_{1v_n} p$.

Thus, for any finite covering $\{S_j\}_{j\geq 1}$ of K, we have $\sum \ \ f_w\,(K) = K$. Therefore

$$\sum_{j\geq 1}(\,\text{diam}(S_j\rangle)^\kappa \geq p^\kappa, \sum_{v(S_j)}\left(a_{rv1} \cdots a_{v_n}\right)^\kappa \geq \rho^\kappa \sum_{)v(S_j)} \mu\left(f_{1v}(K)\right) \geq \rho^\kappa \mu(K) = \rho^\kappa,$$

where μ is the probability measure such that $\mu\big(f_{\tau v}(K)\big) = \left(a_{1v1} \cdots a_{1v_n}\right)^\lambda$ for any w. Hence $\Lambda_\kappa(K) \geq \rho^\kappa$. This completes the proof.

In the case $X = R^p$ with the usual Euclidean distance, the following theorem is known. For the proof, see Falconer [74]. See also Moran [99], Marion [92], and Hutchinson [83].

Theorem (2.1.54)[62]: *Let $\{f_j\}_{1\leq j\leq m}$ be a finite set of contractions of $X = R^p$ satisfying* $|1f_j(x) - f_j(y)\| = \text{Lip}\big(f_j\big)\|x - y\|$ *for any , $y \in X$. Suppose that there exists a bounded open set V such that $\sum_{j=1}^{m} f_j\,(y) \subset y$ and $f_i(V) \cdot f_j(V) = \varphi$ for any $i \neq j$. Then*
$0 < \Lambda_\lambda\,\big(K(f_1, \cdots, f_m)\big) < \infty$; *therefore* $\dim_H(K) = \lambda$ *where λ is given by*

$$\sum_{j=1}^{m}\left(\text{Lip}\big(f_j\big)\right)^\lambda = 1. \tag{62}$$

Example (2.1.55)[62]: Consider the contractions defined by (16). If $a + b < 1$, it follows that $\dim_H\big(K(f_1, f_2)\big) = \lambda$ where $a^\lambda + b^\lambda = 1$ by Theorems (2.1.52) and (2.1.53). In particular, for Cantor's ternary set $(a = b = 1/3)$ we have $\dim_H(K) = \log 2/\log 3$. For the contractions defined by (24), one can easily verify that $\{f_1, f_2\}$ satisfies the condition of

Theorem (2.1.54). Hence $\dim_H(K) = \lambda$ where λ is given by $|\alpha|^\lambda + (1 - |\alpha|^2)^\lambda = 1$. Note that $\dim_H(K)$ is discontinuous at every real α. On the other hand, the contractions defined by (25) does not presumably satisfy the condition of Theorem (2.1.54) for Im $\alpha \neq 0$.

Example (2.1.56)[62]: Let $X = [0,1]$ with the usual Euclidean distance and put

$$f_j(x) = \frac{1}{\mathrm{x} + n_j} \; for\ 1 \leq j \leq m\,, \qquad (63)$$

where $n_1, \cdots, n_m$ are m distinct positive integers. Then $K(f_1, \cdots, f_m)$ is the set of all continued fractions each of whose partial quotients is either $n_1, \cdots, n_{m-1}$ or n_m, since

$$\psi(\alpha) = f_{\alpha_1} \circ f_{\alpha_2} \circ \cdots = \frac{1}{n_{\alpha_1} +} \frac{1}{n_{\alpha_2} +} \cdots \; for\ any\ \alpha = (\alpha_n) \in \Sigma.$$

Using Theorems (2.1.52) and (2.1.53), one can easily obtain lower and upper estimates for $\dim_H(K)$. In this respect, see Good [75]. Moreover one can get better estimates using the following fact repeatedly:

$$K\left(\{f_j\}_{1 \leq j \leq m}\right) = K\left(\{f_{j_1} \circ f_{j_2}\}_{1 \leq j_i \leq m}\right). \qquad (64)$$

Example (2.1.57)[62]: Let $X = C$ with the usual Euclidean distance and put $R_a(z) = az(1 - z)$ where a is a real parameter satisfying $a > 4$. It is known that the Julia set J_a for $R_a(z)$ is totally disconnected and contained in [0,1] (Brolin [66]). Then it is easily seen that if $\geq 2 + \sqrt{5}$, J_a coincides with the set $K(f_1, f_2)$ where

$$f_1(x\rangle = \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{X}{a}} \; and\ f_2(x) = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{X}{a}}. \qquad (65)$$

From Theorems (2.1.52) and (2.1.53), it follows that if $a \geq 2 + 2\sqrt{2}$,

$$\frac{\log 2}{\log a} \leq \dim_H(K) \leq \frac{\log 4}{\log\,(a^2 - 4a)}.$$

Using (64), we also have the following asymptotic expansion:

$$\frac{\log 2}{\dim_{\mathrm{H}}(K)} = \log a - \frac{1}{a} + O(a^{-2}) \; as\ a \to \infty.$$

We will restnct ourselves to the case $X = R^p$ with the usual Euclidean distance. The following theorem has been shown by Williams. Compare with Theorem (2.1.10).

Theorem (2.1.58)[62]: (Williams [115]). *Let $\{f_1, f_2\}$ be two one to one contractions of R such that* Fix $(f_1) \neq F\mathrm{ix}(f_2)$ *and that*

$$Lip\ (f_1^{-1})^{-1} + \mathrm{Lip}(f_2^{-1})^{-1} \geq 1. \qquad (66)$$

Then the set $K(f_1, f_2)$ is a closed line interval.

Here we will give a simple proof for this, which is completely different from Williams' proof.

Proof. Let L_0 be the smallest closed interval containing the set $= K(f_1, f_2)$. Then there exist $\alpha,\ \beta \in \Sigma$ such that $L_0 = [\psi(\alpha), \psi(\beta)]$. Sinoe $f_{\hat{j}} \circ \psi(\alpha),\ f_j \circ \psi(\beta) \in K$, we have $L_0 \geq f_j(L_c)$ forj $= 1{,}2$; therefore $L_0 \geq L_1 \geq L_2 \geq \cdots$ where $F = f_1^* + f_2^* \in \mathcal{T}(\mathcal{C}(R))$ and $L_n = F(L_0)$ for $n \geq 1$. Suppose now that L_ζ is connected but L_{k+1} is not for some $k \geq 0$. Since each $f_j(L_k)$ is a closed interval, it follows that $f_1(L_k) \cdot f_2(L_k) = \varphi$; therefore

$$diam\ (L_{k+1}) > \mathrm{diam}\big(f_1(L_k)\big) + \mathrm{diam}\big(f_2(L_k)\big)$$
$$\geq (\mathrm{Lip}(f_1^{-1})^{-1} + \mathrm{Lip}(f_2^{-1})^{-1})\mathrm{diam}(L_k) \geq \mathrm{diam}(L_k)\,,$$

contrary to $L_{k+1} \leq L_k$. Therefore every L_n is connected. Hence the set $\lim L_n = K(f_1, f_2)$ is connected, as required.

Note that the above theorem holds true even for weak contractions satisfying (66). We now give a generalization of Theorem (2.1.58) as follows:

Theorem (2.1.59)[62]: *Let* $\{f_j\}_{1\le j\le m}$ *be a finite set of one to one weak contractions of R such that* Fix $(f\rfloor \neq F\text{ix}(f_j)$ *for some* $i \neq j$ *and that*

$$\sum_{j=1}^{m}\left(\text{Lip}(f_j^{-1})\right)^{-1} \geq m-1. \tag{67}$$

Then the set $K(f_1,\cdots,f_m)$ *is a closed line interval.*

Proof. It suffices to show the connectedness of $K = K(f_1,\cdots,f_m)$ since K is perfect by Theorem (2.1.12). Suppose, on the contrary, that K is not connected. By Theorem (2.1.18), there exist two positive integers r and s such that $r+s=m$ and that

$$K_j\cdot K_{r+i} = \phi \; for\; any\; 1\le j\le r\; and\; 1\preccurlyeq i\le s\,, \tag{68}$$

where $K_n = f_n(K)$ for $1\le n\le m$. Put $a_n = \text{Lip}(f_n^{-1})^{-1}$ for $1\le n\le m$. Then we get $a_j + a_{r+i} < 1$ for any $1\le j\le r$ and $1\le i\le s$. For otherwise, the set $K^* = K(f_j, f_{r+i})$ is connected by Theorem (2.1.58); therefore, by Theorem (2.1.18), $K_j\cdot K_{r+i}\geq f_j(K^*)$. $f_{r+i}(K^*)\neq\varphi$, contrary to (68). Thus we have

$$s\sum_{j=1}^{r} a_j + r\sum_{i=1}^{s} a_{r+i} < rs. \tag{69}$$

On the other hand,

$$s\sum_{j=1}^{r} a_j + r\sum_{\text{i}=1}^{s} a_{r+i} \geq \min(r,s)\sum_{j=1}^{m} a_j \geq \min(r,s)\cdot(m-1)\geq rs,$$

contrary to (69). This completes the proof.

Remark (2.1.60)[62]: The constant $m-1$ in (67) can not be replaced by any smaller number. For example, for an arbitrary $\varepsilon > 0$, consider the contractions

$$f_1(\text{x}) = \frac{\varepsilon}{m}\text{x}, f_j(x) = (1-\varepsilon)x + \frac{j}{m}\varepsilon \; for\; 2\le j\le m. \tag{70}$$

Then it is clear that $f_1([0,1])\cdot f_j([0,1]) = \varphi$ for $2\le j\le m$; therefore $K(f_1,\cdots,f_m)$ is not connected by Theorem (2.1.18), while

$$\sum_{j=1}^{m}(\text{Lip}(f_j^{-1}))^{-1} > m-1-\varepsilon m.$$

In connection with Theorems (2.1.10) and (2.1.58), Williams gave the following problem: what is the structure of $K(f_1,f_2)$ for $f_1,f_2\colon R^2\to R^2$, affine contractions satisfying (66)? Here we will give a partial answer for this. In fact, more generally we haye

Theorem (2.1.61)[62]: *Let* $\{f_1,f_2\}$ *be two one to one weak contractions of* R^p *such that* Fix $(f_1)\neq \Gamma\text{ix}(f_2)$ *and that*

$$Lip\,(f_1^{-1})^{-p} + \text{Lip}(f_2^{-1})^{-p} > 1. \tag{71}$$

Then the set $K = K(f_1,f_2)$ *is a nondegenerate locally connected continuum; therefore* $\dim_{\text{T}}(K)\geq 1$.

Proof. Suppose, on the contrary, that K is not connected. Then, by Theorem (2.1.18), we have $f_1(K)\cdot f_2(K)=\varphi$. Therefore it follows that $\dim_H(K) > p$. This contradiction completes the proof.

As a corollary, we have immediately

Corollary (2.1.62)[62]: *Theorem* (2.1.61) *holds true for two one to one weak contractions* $\{f_1, f_2\}$ *satisfying* $\mathrm{Fix}(f_1) \neq \Gamma\mathrm{ix}(f_2)$ *and*

$$Lip\ (f_1^{-1})^{-1} + Lip(f_2^{-1})^{-1} > 2^{(p-1)/p} \qquad (72)$$

Remark (6.1.63)[62]: For $p = 2$, the constant $\sqrt{2}$ in (72) can not be replaced by any smaller number. For example, consider the contractions

$$f_1(z\rangle = \left(s + \frac{i}{2}\right)\overline{z}\ and\ f_2(z) = \left(s - \frac{i}{2}\lambda\overline{z} - 1\right) + 1\ , \qquad (73)$$

where s is a real parameter satisfying $0 < s < 1/2$. We denote by Q_s the closed quadrangle with vertices $0,1,1 - s + i/2$ and $s + i/2$. Then it is easily seen that $f_1(Q_s) + f_2(Q_s) \leq Q_s$ and $f_1(Q_s) \cdot f_2(Q_s) = \varphi$. Therefore the set $K(f_1, f_2)$ is totally disconnected by Theorem (2.1.13) (Fig. 9), while

$$Lip\ (f_1^{-1})^{-1} + \mathrm{Lip}(f_2^{-1})^{-1} = \sqrt{1 + 4s^2} \qquad (74)$$

Fig. (9)[62]:

We present the following problem: is it true or not that if one to one weak contractions $\{f_j\}_{1\leq j\leq m}$ of R^p satisfy$m$

$$\sum_{j=1}^{m}\left(\mathrm{Lip}\left(f_j^{-1}\right)\right)^{-p} \geq 1,$$

then the set $K(f_1, \cdot, f_m)$ contains a nondegenerate component?

Section (2.2): Separation of Self-Similar Fractals in the Plane

Given contracting similarity maps$f_1, \ldots, f_m$on $\mathbb{R}^d$,the corresponding self-similar set is the unique compact set $A \neq \emptyset$ which satisfies the set equation

$$A = f_1(A) \cup \ldots \cup f_m(A).$$

A consists of similar copies $A_i = f_i(A)$of itself, each A_i consists of smaller copies $A_{ij} = f_i(f_j(A))$,and so on. For any integern,we can consider the set S^n of words $i = i_1 \ldots i_n$ from the alphabet $S = \{1, \ldots, m\}$.Writing $f_i = f_{i_1} \cdots f_{i_n}$and $A_i = f_i(A)$,we have $A = \bigcup\{A_i | i \in S^n\}$.When n tends to infinity, this induces a continuous map$\pi: S^\infty \to A$from the set S^∞ of sequences $s = s_1 s_2 s_3 \ldots$.onto the self-similar set, theso-called address map. See [119], [122], [123], [83].

When the contraction factors $r_i$of f_i are small, the pieces A_i are disjoint, π is a homeomorphism and A a Cantor set. For large r_i, however, π identifies many addresses, and the overlaps $A_i \cup A_j$ are usually too large to analyse Mathematically. In between there is the 'just-touching case' [122] where overlaps are nonempty but sufficiently thin. It is defined by four equivalent conditions.

(i) Moran's open set condition (OSC) [99]: there exists a nonempty open set$V \subset \mathbb{R}^n$ with $\bigcup_{i=1}^{m} f_i(V) \subseteq V$ and $f_i(V) \cap f_j(V) = \emptyset$ for $i = j$.

(ii) Positivity of α-dimensional Hausdorff measure: $\mu^{\alpha}(A) > 0$, where α denotes the similarity dimension given by $\sum r_i^{\alpha} = 1$[99], [128].
(iii) The finite clustering property [128]: there exists an integer N such that for every piece A_i of A, with diameter ε, say, there are at most N incomparable pieces A_j of diameter ε with distance $< \varepsilon$ from A_i. We call A_j and A_k incomparable if j is not a prefix of k and k is not a prefix of j.
(iv) The neighbour map condition [120]: the identity map id is not an accumulation point of the set of neighbour maps of A. A neighbour map has the form $h = f_i^{-1} f_j$ where $i, j \in S^* = \bigcup_{n\geq 1} S^n$ and $i_1 \neq j_1$. Convergence of similarity maps on $\mathbb{R}^d$ is given by the norm $\|g\| = \sup_{|x|\leq 1} |g(x)|$.

With so many equivalent formulations, OSC has become generally accepted as the adequate separation condition for self-similar fractals. But does it really say that each overlap $A_i \cap A_j$ is small? At first glance, yes. If OSC holds, the overlap is contained in $\bar{V}_i \cap \bar{V}_j$, so ithas no interior points, and $\mu^{\alpha}(A_i \cap A_j) = 0$. We do not know, however, whether the Hausdorff dimension of overlaps must be smaller than α, except for the case of finite type [126].

And what about the converse? The finite clustering condition implies that the cardinality of $\pi^{-1}(x)$ for $x \in A$ is uniformly bounded by some number N. We do not know whether this property is equivalent to OSC. Here we consider a more special case.

We start with some evidence for a negative answer. A sequence $s_1 s_2 \ldots$ is called recurrent if for each $K \geq 1$ there is a $n \geq 1$ with $s_1 \ldots s_K = s_{n+1} \ldots s_{n+K}$. We show that any identificationof a recurrent address will destroy OSC. However, it is not clear whether such an identification implies other identifications of addresses.
Our main result is an affirmative answer to the problem for $d = 2$.
We deal with the general case and with the case that A is homeomorphic to an interval. The proof uses plane topology at some key places. We expect that Theorem (2.2.8) is not true in higher dimensions.
Examples (2.2.1)[116]: s is a recurrent sequence if arbitrarily long prefixes $s_1 \ldots s_K$ will occur inside the sequence. An example is the Cantor sequence

$$s = 2121112121111111112121112 12\ldots \tag{75}$$

obtained as limit of the words $s^{(n)}$ where $s^{(0)} = 2$ and $s^{(n+1)} = s^{(n)} 1^{3^n} s^{(n)}$ for $n \geq 0$. Anotherexample is given by taking $s^{(0)} = 2$ and $s^{(n+1)} = s^{(n)} 1^n s^{(n)}$:

$$s = 2121121211121211212\ldots.$$

A third example is the prominent Fibonacci sequence generated by the substitution $1 \to 2, 2 \to 21$:

$$s = 21221212212212\ldots.$$

If s is recurrent, then for each $N \geq 1$ there is an index k_N such that the word $i(N) = s_1 \ldots s_{k_N}$ has N different suffixes which coincide with prefixes of s. The k_N are constructed by induction: let $k_1 = 1$ and let k_2 be the smallest number for which $s_{k_2} = s_1$. Let k_3 denote the end point of the first repetition of the word $s_1 \ldots s_{k_2}$ inside s, etc. Our example sequences all end with k_4.
Theorem (2.2.2)[116]: (a) In a self-similar set A, if one point $a \in A_{s_1}$ with a recurrent addresss belongs to a piece A_{t_1} with $t_1 \neq s_1$, then OSC cannot hold.

(b) There are self-similar Cantor sets A in $\mathbb{R}$ or $\mathbb{R}^2$,with arbitrary small Hausdorff dimensionand with $A_i \neq A_j$ for $i, j \in S^*$ with $i \neq j$, which do not fulfil the OSC.

Proof. (a). According to the finite clustering property, a self-similar set A cannotfulfil the OSC if for every $N \in \mathbb{N}$ there is a piece $A_i$which intersects at least N other pieces A_j and no piece is a subpiece of another one, with diam$A_j \geq$diam A_i.

Let N be given and let $a \in A_{s_1} \cap A_{t_1}, s_1 \neq t_1$ be a point with recurrent address $s = s_1 s_2 s_3$...and a second address t. There is an initial word $i = i_1 \ldots i_n = s_1 \ldots s_n$ of s which has N suffixes which coincide with prefixes of s(see below). In other words, there are $\ell_1 < \ell_2 < \cdots < \ell_N < n$such that $i_{\ell_k+1} \ldots i_n = s_1 \ldots s_{n-\ell_k}$ for $k = 1, \ldots, N$. Now we define
$j_k = i_1 \ldots i_{\ell_k} t_1 \ldots t_{n_k}$for $k = 1, \ldots, N$,
Where $n_k$is chosen as large as possible so that still diam $A_{j_k} \geq$diamA_i. (If the factors r_i are very different, then to guarantee $n_k \geq 1$ exists we have to choose $n - \ell_N$ so large that $r_{s_1} \cdots r_{s_{n-\ell_N}} < r_{t_1}$).

Now $f_{i_1 \ldots i_{\ell_k}}(a)$is in $A_i \cap A_{j_k}$.So A_i intersects N pieces of $A_{j_k}$of at least the same size,and they are incomparable: for $k < k'$ the $(\ell_k + 1)$st coordinate of j_k is t_1 and the$(\ell_k + 1)$stcoordinate of $j_{k'}$is $i_{\ell_k+1} = s_1$. OSC does not hold.

(b) We take similitudes in the complex plane with equal factor r, that is,$f_j(z) = a_j r z + b_j, j \in S$, with $a_j, b_j \in \mathbb{C}$ and $|a_j| = 1$. Then the address $t = t_0 t_1 t_2 \ldots \in S^\infty$is mapped to the point

$$\pi(t) = b_{t_0} + \sum_{k=1}^{\infty} r^k b_{t_k} \prod_{\ell=0}^{k-1} a_{t_\ell}. \tag{76}$$

Here we start with t_0 since this gives a power series in r.We also write $s = s_0 s_1 \ldots$. To provethe equation, start with $f_{t_0 t_1}(z) = b_{t_0} + r a_{t_0} b_{t_1} + r^2 a_{t_0} a_{t_1} z$and continue by induction. Note that $r^n z$ tends to zero for $n \to \infty$.We apply the formula to some very simple mappings:

$$f_1(z) = rz, f_2(z) = rz + 1, f_3(z) = \omega r z + c,$$

Where c, ω are complex numbers,$|\omega| = 1$.Suppose we want to identify the points corresponding to the Cantor sequence (75) $s = 212111212\ldots$and to $3\overline{1} = 3111\ldots$. Formula (73) with $a_1 = a_2 = 1, a_3 = \omega$ and $b_1 = 0, b_2 = 1, b_3 = c$ gives $\pi(s) = 1 + r^2 + r^6 + r^8 + \cdots = \sum_{s_k=2} r^k$ and $\pi(3\overline{1}) = c$ since $\pi(\overline{1}) = 0$. Thus the condition $\pi(3\overline{1}) = \pi(s)$holds if and only if

$$c = 1 + r^2 + r^6 + r^8 + \cdots = \sum_{s_k=2} r^k = \prod_{k=0}^{\infty} \left(1 + r^{2 \cdot 3^k}\right).$$

Given $r \in]0,1[$, we obtain the corresponding c and ω can be chosen on the unit circle. For theself-similar set A in figure1, we took $\omega = i, r = 0.45$ and got $c \approx 1.2125$.(For $r < \frac{1}{3}$which we assume below the picture would be hardly visible.)

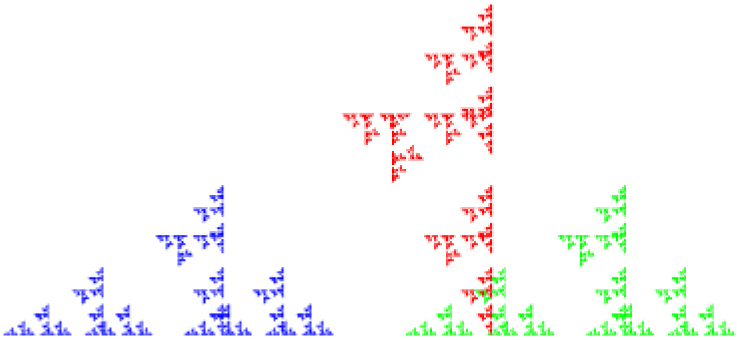


Figure (1)[116]: Self-similar Cantor set without OSC.

Let us verify the properties of Theorem (2.2.2)(b) for the above example. The Hausdorff dimension of A is not larger than the similarity dimension α determined by the equation $\sum r_j^{\alpha} = 3r^{\alpha} = 1$[123], [83]. In particular, A is a Cantor set whenever $r < 1/3$ since then $\alpha < 1$ and any compact connected set has Hausdorff dimension of at least1([124]). Moreover, taking r small enough we get $\alpha = \ln 3/|\ln r|$ as small as we want. For $\omega = 1$ we have $A \subset \mathbb{R}$. All this holds for arbitrary small r.

Finally, for $\omega = 1$ we show that $A_i = A_j$with $i \neq j$can hold only for countably many r.It suffices to show this for fixed words i, j of the same lengthn. Now $A_i = A_j$ means $f_i = f_j$ and $f_i(z) = r^n z + \sum_{k=0}^{n-1} b_{n-k} r^k$ where $b_k \in \{0,1,c\}$. Thus $f_i = f_j$leads to an equation of the form $p(r) + q(r)c(r) = 0$, where p, q are polynomials with coefficients−1,0,1 and $c(r)$ is the above power series. However, an analytic function has only finitely many zeros in [0,1]see [125], [127].

Let $B_r(x)$ denote the ball around x with radiusr. Take a number $R \geq 1$ with $\bigcup_{i=1}^{m} f_i(B_R(0)) \subseteq B_R(0)$ and define $\|g\| = \sup_{(|x| \leq R|} |g(x)|$ for $g: \mathbb{R}^d \mapsto \mathbb{R}^d$.

Lemma (2.2.3)[116]: If g and f are similitudes and $f(B_R(0)) \subseteq B_R(0)$, then

$$\|f^{-1}gf - id\| < c_f \|g - id\|,$$

Where c_f is a constant depending only on f .

Proof. Let $f^{-1}(x) = Gx + b$ where G is linear and $c_f = \|G\|$. For $x \in B_R(0)$, we have

$$|f^{-1}gf\,(x) - x| = |f^{-1}gf\,(x) - f^{-1}f(x)| = |G(g(f\,(x)) - f(x))|$$
$$\leq \|G\| \cdot |g(f\,(x)) - f(x)| \leq \|G\| \cdot \|g - id\|$$

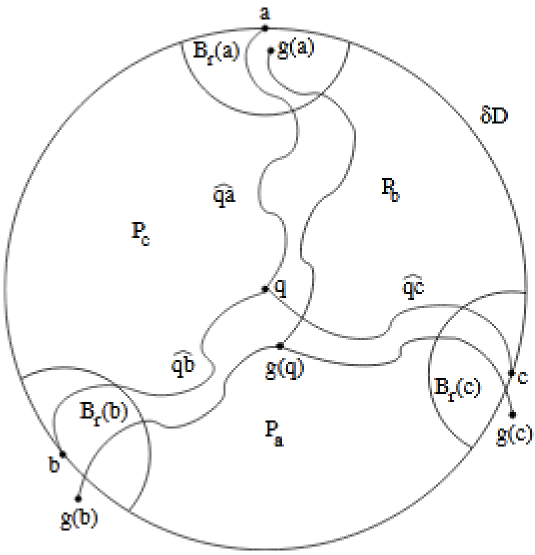


Figure (2)[116]: Proof of Lemma (2.2.6). Drawing by M Mesing.

Lemma (2.2.4)[116]: Let A be a self-similar set which is not a singleton. For any integer $M > 0$ there exists k_0 and $j_1, \ldots, j_M \in S^{k_0}$ such that the $A_{j_k}$are all disjoint.

Proof. Two maps, say f_1 and f_2,have different fixed points x_1, x_2.So all $f_1^k(x_2)$with $k = 1,2, \ldots, M$ are different. Take $j_k = 1^k 2^{(k_0-k)}$for sufficiently large k_0.

A simple Jordan curve is the image set of a homeomorphism from [0,1] to$\mathbb{R}^2$.

Lemma (2.2.5)[116]: Let A be a connected self-similar set which is not a simple Jordan curve. Then there exist four points $a, b, c, q \in A$such that

(i) there exist Jordan curves $\widehat{qa}, \widehat{qb}, \widehat{qc} \in A$;

(ii) $\widehat{qa}, \widehat{qb}, \widehat{qc}$ intersect each other only at the point q;

(iii) $|qa| = |qb| = |qc| = r'$;

(iv) $\widehat{qa}\backslash\{a\}, \widehat{qb}\backslash\{b\}$and $\widehat{qc}\backslash\{c\}$are contained in the interior $D^\circ$of the closed disc D with centreq and radiusr′.

Proof. A connected self-similar set is arcwise connected [121], [122], [62]. If A is not a Jordan curve, there is a point q and a neighbourhood U of q such that $A \cap U\backslash\{q\}$has at least three components which have q in their closure. Taking a', b', c' in these components, we get disjointarcs $\widehat{qa'}, \widehat{qb'}, \widehat{qc'}$.Let r' be smaller than the distances of a', b', c' to q and let D be the closeddisc with centre q and radius r'. Starting in q,let a, b, c denote the first points where the arcshit the boundary ∂D.

We denote by K the union of the three Jordan curves, that is,$K = \widehat{qa} \cup \widehat{qb} \cup \widehat{qc}$.

Lemma (2.2.6)[116]: (Perturbation lemma) There is a constant $\delta > 0$ such that $K \cap g(K) \neq \emptyset$ for every similitude g with $\|g - id\| < \delta$.

Proof. In Lemma (2.2.5), $D\backslash K$ is divided into three parts P_a, P_b and P_c, where the closure of $P_x$doesnot contain x. Choose r so small that $B_{2r}(a)$does not intersect $P_a, B_{2r}(b)$does not intersect P_b and $B_{2r}(c)$ does not intersect P_c.

Let σ be the distance between $K\backslash(B_r(a) \cup B_r(b) \cup B_r(c))$and the circle ∂D. Clearly $\sigma > 0$. We set

$$\delta = min\{r'/2, r, \sigma/2\}$$

and show that $\|g - id\| < \delta$ implies$g(K) \cap K \neq \emptyset$.

Since $\|g - id\| < r'$, the point $g(q)$ is in D°.If $g(q)$ belongs to K, our assertion is true.So we assume, without loss of generality, that $g(q)$ belongs to P_a. Since$\|g - id\| \leq r$, the point $g(a)$ is in $B_r(a)$, and hence not in P_a.Now we show that $g(\widehat{qa})$does not intersect $\widetilde{bc}$,the arc from b to c on the circle ∂D. The part $g(\widehat{qa} \cap B_r(a))$isstill in $B_{2r}(a)$, so it will not intersect $\widetilde{bc}$,and $g(\widehat{qa}\backslash B_r(a))$is contained in the disc with centreqand radius $r' - \sigma/2$, and will not intersect ∂D at all. Now $g(\widehat{qa})$must intersect $\widehat{qb} \cup \widehat{qc}$ and $g(K)$ must intersect K(figure 2).

Definition (2.2.7)[116]: A self-similar set A is of finite type if there are only finitely many neighbor maps $h = f_i^{-1} f_j$ with $A \cap h(A) \neq \emptyset$(or equivalently $A_i \cap A_j \neq \emptyset$) and with similarity factor $r_h \in (r_*, 1/r_*)$where $r_* = min\{r_1, \ldots, r_m\}$.

The meaning of the last condition is that we accept only neighbours $h(A)$ which fit the size of A, otherwise we take their pieces or supersets as $h(A)$.If allriare equal, we take onlyneighbours $h(A)$ which have the same size as A,and the mapshare isometries.

Compared with the finite type concept in Ngai and Wang [126], this definition is a bit more restrictive but simpler and in our opinion more natural. The following was proved for equal factors in [118].

Theorem (2.2.8)[116]: A self-similar set of finite type fulfils OSC if $f_i \neq f_j$ for $i \neq j$.

Proof. Let

$$H_0 = \{h = f_i^{-1} f_j | i_1 \neq j_1, A \cap h(A) \neq \emptyset, r_h \in (r_*, \frac{1}{r_*})\},$$

let $\widetilde{H} = \{f_i^{-1} h f_j, f_i^{-1} h, h f_j, f_i^{-1} f_j | h \in H_0 \cup \{id\}, i, j \in S\}$ denote the 'immediate successors' of maps of $H_0 \cup \{id\}$ and let

$$H_1 = \{\tilde{h} \in \widetilde{H} | \tilde{h}\,(A) \cap A = \emptyset\}.$$

Since for $h \in H_1$ the compact sets A and $h(A)$ are disjoint, their distance $d_h = d(A, h(A)) = inf\{|x - y| | x \in A, y \in h(A)\}$ is positive. Moreover, $\|h - id\| \geq max\{|h(x) - x| | x \in A\} \geq d_h$.If A is of finite type, H_0 and hence H_1 are finite. Thus $\delta_0 = min\{\|h - id\| | h \in H_0\} > 0$ since $f_i \neq f_j$ implies $id \notin H_0$ and $\delta_1 = \min\{d_h | h \in H_1\} > 0$. Any neighbour map $g \in H_0 \cup H_1$ has the form $f_i^{-1} h f_j, f_i^{-1} h$ or $h f_j$ with suitable $h \in H_0 \cup H_1$ and words $i, j \in S^*$. Then $d_g \geq d_h$ since $d(B', C) \geq d(B, C)$ for $B \supseteq B'$:

$$d(f_i^{-1} h(A), A) \geq d(f_i^{-1} h(A), f_i^{-1}(A)) = \frac{dh}{r_i}, d(h f_j(A), A) = d\big(h\big(A_j\big), A\big) d_h$$

(for $f_i^{-1}\, h f_j$ we combine both estimates). This implies $\|g - id\| \min\{\delta_0, \delta_1\}$.This is true for all neighbour maps g.So id cannot be an accumulation point and OSC holds.

Now we consider self-similar sets homeomorphic to [0,1] and write $J = f_1(J\,) \cup \ldots \cup f_m(J\,)$ instead of A.If the pieces have finite intersection, each intersection is at most one point,and we can assume that $J_i \cap J_{i+1} = \{c_i\}$ for $i = 1, \ldots, m - 1$. Furthermore, let $c_0 \in J_1$ and $c_m \in J_m$ be the two endpoints of J(those points x for which $J \backslash \{x\}$ is connected). Concerningthe addresses of c_0, c_m four cases are possible (cf [121], [62]):

(i) $f_1(c_0) = c_0$ and $f_m(c_m) = c_m, i.e.\, c_0 = \pi(\overline{1}), c_m = \pi(\overline{m})$.

(ii) $f_1(c_0) = c_0$ and $f_m(c_0) = c_m, i.e.\, c_0 = \pi(\overline{1}) and\ c_m = \pi(m\overline{1})$.

(iii) $f_1(c_m) = c_0$ and $f_m(c_m) = c_m, i.e.\, c_0 = \pi(1\overline{m}), c_m = \pi(\overline{m})$.

(iv) $f_1(c_m) = c_0$ and $f_m(c_0) = c_m, i.e.\, c_0 = \pi(\overline{1m}), c_m = \pi(\overline{m1})$.

Theorem (2.2.9)[116]: A self-similar Jordan curve in the plane is of finite type unless

(i) it has endpoint type (i),

(ii) there exists an $i \in \{1, 2, \ldots, m - 1\}$ such that $f_i^{-1}(c_i) \neq f_{i+1}^{-1}(c_i)$ and

(iii) $\frac{log r_m}{log r_1}$ is irrational for the contraction factors r_1, r_m of f_1, f_m.

Proof. We need only check neighbour maps $h = f_i^{-1} f_j$ for pieces $J_i \subseteq J_i$ and $J_j \subseteq J_{i+1}$ of approximately the same size which intersect in the point $c_i (i = 1, \ldots, m - 1)$. We show that neighbour maps atcirepeat periodically when we go to smaller pieces.

If $f_i^{-1}(c_i) = f_{i+1}^{-1}(c_i)$ or if we have endpoint type (ii)(cf figure 3), (iii) or (iv), then both addresses of c_i are eventually periodic with the same periodic part:

$$c_i = \pi(iu\overline{w}) = \pi((i + 1)v\overline{w}),$$

Where u, v can be 1 or m or the empty word (in which case $f_u = id$) and $w \in \{1, m, 1m, m1\}$ where $\overline{w}$ is the address of an endpoint of J. This endpoint which we call 0 is the fixed pointof f_w.What is more important is that it is also the fixed point of all neighbour maps

$$h = f_i^{-1} f_j, \text{with } i = iuw^n \text{ and } j = (i + 1)vw^{n'},$$

since our assumption was $c_i = f_i f_u(0) = f_{i+1} f_v(0)$. In other words, when (i) or (ii) is not true then the neighbour maps are rotations around one endpoint of J,composed with a stretchingand/or a reflection.The condition $r_* < r_h < \frac{1}{r_*}$ for $r_h = \frac{r_j}{r_i} = \frac{r_{i+1} r_v}{r_i r_u} \cdot r_w^{n'-n}$ says that onlyfinitely many differences $n' - n$ are possible.

Now let us first assume that all similitudes f_i are orientation-preserving,$f_i(z) = a_i z + b_i$.We choose the origin of our coordinate system at the fixed point 0 of f_w so that $f_w(z) = az$ for some $a \in \mathbb{C}$ with $|a| = r_w < 1$.Then

$$h = f_i^{-1} f_j = f_w^{-n} f_u^{-1} f_i^{-1} f_{i+1} f_v f_w^{n'} = a^{n'-n} f_u^{-1} f_i^{-1} f_{i+1} f_v \qquad (78)$$

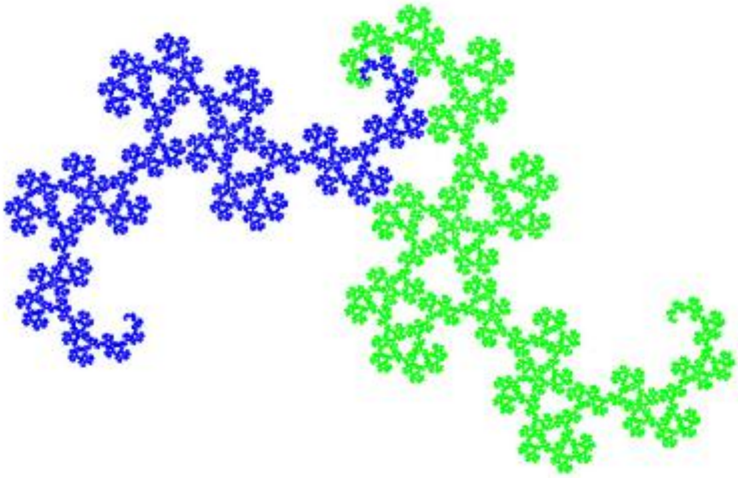


Figure (3)[116]: Self-similar Jordan arc of type (ii) with two pieces.

because orientation-preserving similitudes in the plane with a common fixed point commute. So the neighbour map h depends only on $n' - n$,not on n or n' separately. Consequently, the number of neighbour maps at c_i is finite. (Actually, for case (78) with $w = 1m$, we still have to consider $i = iuw^n 1$ and /or $j = (i+1)vw^{n'} 1$ which increases the number of neighbour mapsat most by a factor 4.) We have verified the finite type if either (i) or (ii) is not fulfilled.

What will change if we admit orientation-reversing similitudes,$f_i(z) = a_i \bar{z} + b_i$? if $f_w(z) = a\bar{z}$ we can work with $w^2 = ww$ instead of w in the above calculations. Since $f_{ww}(z) = f_w(f_w(z)) = |a|^2 z$,equation (78) again holds. The vertex c_1 in figure4shows such a case (type (i) but $c_1 = f_1(c_4) = f_2(c_4)$).

Commutativity could only fail if $f_w(z) = az$ with $a \notin \mathbb{R}$ and at the same time $f_u^{-1} f_i^{-1} f_{i+1} f_v$ is orientation-reversing. In that case one of the maps $f_i f_u, f_{i+1} f_v$ would preserve and the other one would reverse the orientation. However, this case is not possiblefor a Jordan curve: by assumption, the fixed point 0 of $f_w$is an endpoint of J(namely c_0 for $w = 1$or $w = 1m$ and c_m for $w = m$ or $w = m1$). If z_0 denotes the other endpoint of J, the curve consists of Jordan arcs connecting points $z_0, az_0, a^2 z_0$,.... In other words, the curveapproaches 0 as a fractal spiral. The mappings $f_i f_u$ and $f_{i+1}\, f_v$map this spiral to two spirals with centre c_i which represent J_i and J_{i+1}. If these spirals have different orientations, theyhave plenty of intersection points, contradicting the Jordan curve structure of J.

To finish the proof, we assume both (i) and (ii) and show the finite type if (iii) is not true that is,$r_1^k = r_m^{k'}$for positive integers k, k'. By (i) and (ii),c_i has addresses $i\bar{m}$ and $(i+1)\bar{1}$(or $i\bar{1}, (i+1)\bar{m}$). Let us take the origin of our coordinate system at c_i. In the orientation-preserving case $f_1(z) - c_0 = a_1(z - c_0)$, $f_m(z) - c_m = a_m(z - c_m)$ the sets J_i and J_{i+1}form two fractal spirals approaching $c_i = 0$,and these spirals are mapped into themselvesby multiplication with a_m and a_1,respectively. (Seen from the centre c_m,the fractal spiral J connects $c_0, z_1, z_2, \ldots$, where $z_k - c_m = a_m^k (c_0 - c_m)$. Now $f_i(z) = \alpha z + \beta$maps J to J_i with c_m to $0, c_0$ to c_{i-1} and hence z_k to $y_k = a_m^k c_{i-1}$.Similarly fora1and f_{i+1}. If f_i, f_{i+1}areorientation-reversing the factor is $\bar{a}_m \bar{a}_1$.) By, $a_m = a_1^t$ with $t = k/k'$.Using the maps $f_1^k$and $f_m^{k'}$instead of f_1, f_mwe can use the above argument to show finite type at

c_i(instead of factor 4 for $w = 1m$ we have factor kk'). The case that $f_1$or f_m is orientation reversing leads to real factors as above. The proof is complete.
We should also mention that theorem10does not hold in dimension 3.

Lemma (2.2.10)[116]: If $J_1, J_2 \subset \mathbb{C}$ are Jordan curves with $J_1 \cap J_2 = \{0\}$ and a_1, a_2 complex numbers with $a_1 J_1 \subset J_1$and $a_2 J_2 \subset J_2$ then there is $t > 0$ with $a_2 = a_1^t$.

Proof. Let $a_1 = r \cdot e^{i\alpha_1}, a_2 = r^t \cdot e^{it\alpha_2}$.We must show $\alpha_1 = \alpha_2$,so let us assume $\alpha_2 - \alpha_1 = \varepsilon \in (0,2\pi)$.There are points $z_1, z_2 \neq 0$ on $J_1, J_2$with $|z_1| = |z_2|$.We take z_1, z_2 as endpoints of J_1, J_2(forgetting points with larger modulus), and we can assume that no other points of $J_1 \cup J_2$has modulus$|z2|$.Express the curve J_1 between $z_1$and $a_1 z_1$ inparametric form $\varphi(s) = r(s)e^{i\alpha(s)}, 0 \leq s \leq 1$ where $\alpha(s)$is continuous (does not jump from 2π to 0). Let β =max$\{\alpha(s) - \alpha(0) | 0 \leq s \leq 1\}$ and choose n so large that $nt\varepsilon > 4\pi + \beta$.

Let $z_2' = z_2 a_2^n$ and let z_1' be the first point of J_1 (starting from z_1) which has modulus equal to $|z_2'| = |z_2| \cdot r^{nt}$. Then z_1' lies between $z_1 a_1^k$ and $z_1 a_1^{k+1}$, where k is the integer part of nt. Now we parametrize the two Jordan curves $[z_i, z_i'] \subset J_i$between the circles $|z| = |z_2|$ and $|z| = |z_2'|$ by $\varphi_i(s) = r_i(s)e^{i\gamma_i(s)}, 0 \leq s \leq 1$, where $\gamma_i(s)$ is continuous and $\gamma_2(0) \in (-2\pi, 0], \gamma_1(0) \in [0,2\pi)$.Then $\gamma_2(0) < \gamma_1(0)$ but

$$\gamma_2(1) = \gamma_2(0) + nt\alpha_2 \geq \gamma_2(0) + nt\varepsilon + k\alpha_1 > \gamma_2(0) + 4\pi + \beta + k\alpha_1$$
$$> \gamma_1(0) + \beta + k\alpha_1 \geq \gamma_1(1).$$

This proves that $J_1$and J_2 have an intersection point with modulus $|z_2'|$which contradictsthe assumption.

Theorem (2.2.11)[116]: Let A be a connected self-similar set in the plane. If $A_i \cap A_j$ is a finite set for $i \neq j$, then OSC holds.

Proof. Let

$$M = max\{card(A_i \cap A_j) | i \neq j\} + 1,$$

k_0 the constant in Lemma (2.2.4) and j_1,...,$j_M \in S^{k_0}$be the words with disjoint A_{j_k}.Let

$$c = max\left\{c_{f_j} \middle| j \in S^{k_0}\right\}, \tag{77}$$

Where c_f is the constant of Lemma (2.2.3).

Suppose the f_i do not satisfy OSC. By the neighbour map condition, there are $i, i' \in S^*$with $i_1 \neq i_1'$and

$$\|f_{i'}^{-1} f_i - id\| < \delta / c$$

Where δ is the constant in Lemma (2.2.6) and c is from (77). Then by Lemma (2.2.3), we have

$$\|f_{i'j}^{-1} f_{ij} - id\| = \|f_j^{-1} f_{i'}^{-1} f_i f_j - id\| < c\|f_{i'}^{-1} f_i - id\| < \delta$$

for each $j \in S^{k_0}$.Hence $f_{i'j}(A) \cap f_{ij}(A) \neq \emptyset$ by Lemma (2.2.6). Let p_j be a point in the intersection. Then $p_{j_1}, \ldots, p_{j_M}$are all different, and they belong to the set $f_{i_1'}(A) \cap f_{i_1}(A)$ which contradictsthe definition of M. So the fisatisfy OSC.

For Jordan curves. We show that each self-similar Jordan curve J fulfilsOSC, studying neighbour maps at each point c_i. By Theorem (2.2.9), we can assume (i), (ii) and(iii) because otherwise J is of finite type and hence OSC by Theorem (2.2.8). Also if J is containedin a line, OSC is obvious.

As in the end of the proof of Theorem (2.2.9), c_i has addresses $i\bar{m}$ and $(i+1)\bar{1}$. We assume that f_1,f_m preserve orientation so that J_i and J_{i+1}form spirals around c_i which remain invariant under similitudes with factors a_m, a_1 and centre c_i. $a_m = a_1^t$but this time t isirrational.

We show that the neighbour maps $h = f_i^{-1} f_j$with $i_1 = i$ and $j_1 = i + 1$ cannot approach id.We first study arbitrary small pieces J_i, J_j intersecting at c_i ,that is,$i = im^n, j = (i+1)1^{n'}$.Using $r_m = r_1^t$,we see that we have an infinite type here:

$$r_h = \frac{r_{i+1}}{r_i} \cdot r_1^{n'-nt}.$$

Since for irrational t the set$\{n' - nt | n, n' \in \mathbb{N}\}$is dense in $\mathbb{R}$,the factors r_h include a denseset in$(0, \infty)$.
We apply f_i^{-1}in each case so that each $J_i$is mapped onto J,and each $J_j$is transformed into some $h(J)$.The intersection point is $c_m = f_i^{-1}(c_i) = h(c_0)$for every n and n'.Let $c_m = 0$ be our origin now. While J is kept fixed, the neighbours $h(J)$are obtained from $h_0(J) = f_i^{-1} f_{i+1}(J)$by multiplication with $a_1^{n'-nt}$.We define $U = \bigcup\{h(J) | h = f_{im^n}^{-1} f_{(i+1)1^{n'}}\}$ as union of all these neighbours.$h(J) \cap J = \{0\}$for each h implies $U \cap J = \{0\}$.

Now let us consider the trajectories of the flow $s \mapsto z_0 a_1^s$within the unit circle C.For each $z_\alpha = e^{i\alpha} \in C$ we have the spiral

$$S_\alpha = \{z_\alpha a_1^s | s > 0\}$$

Since J is compact and connected, and the spirals do not form self-similar sets (considercurvature), the set of all z_α for which $S_\alpha \cap J \neq \emptyset$will be an interval $[z_\beta, z_\gamma]$on C,or C itself . The same is true for the neighbour set $h_0(J)$.Since it is a scaled and rotated copy of J,itintersects as many spirals S_α as J. The set of all z_α for which $S_\alpha \cap J \neq \emptyset$will be an interval $[z_{\beta'}, z_{\gamma'}]$ of the same length as $[z_\beta, z_\gamma]$,or C.

The other neighbours $h(J)$ determine exactly the same interval, since they are obtained from $h_0(J)$by multiplication with $a_1^{n'-nt}$ which leaves each S_α invariant. Thus U also determines the same interval. However,U is a dense union of Jordan curves and will occupy a dense set on each S_α which it intersects. From this fact it will follow that $[z_\beta, z_\gamma]$ and $[z_{\beta'}, z_{\gamma'}]$are proper intervals which can intersect only in their endpoints. That is,there are at most twospirals S_α which intersect both J and U.

To prove this, we assume the contrary: there are two spirals S_1, S_2 which, together with all spirals between them, belong to the interior of both $[z_\beta, z_\gamma]$ and $[z_{\beta'}, z_{\gamma'}]$.Thus J intersects S_1, S_2 in z_1, z_2 and joins them with a Jordan are $\widehat{z_1, z_2}$, and $h_0(J)$intersects S_1, S_2in y_1, y_2,say,and contains a Jordan are $\widehat{y_1, y_2}$ between them. This arc as well as all its multiples $a_1^s . \widehat{y_1, y_2}$,where s is taken from a dense set of positive numbers, must not intersect J. This is only possible if$\{y_1, y_2\} = b \cdot \{z_1, z_2\}$for some $b = a_1^u \in \mathbb{C}$. It follows that $\widehat{y_1, y_2} = b . \widehat{z_1, z_2}$ and that the arcs J and $h_0(J)$ continue to intersect the spirals in a parallel manner, because as soon as one of the arcs would turn back, a multiplication of $h_0(J)$by a_1^s with s very near to 0 would result in an intersection point. On the other hand, since J contains a spiral point, it contains adense set of spiral points, and can never intersect the spirals in successive order; it must turnback which is a contradiction.
We proved that J intersects S_α for z_α in a proper interval $[z_\beta, z_\gamma]$,and U can only intersectthe two boundary spirals S_β and S_γ.These two spirals will be used to separate J and U althoughthey may contain points of both sets.

Now it is easy to give a uniform estimate$\|h - id\| \geq \eta > 0$ for all these neighbor maps h, by just using a point $z \in J$ which lies on the spiral $S_{\beta+\gamma/2}$, which has distance $> \eta$ from U. However, since we have no finite type, we must study also neighbour maps between disjoint pieces of $J_i$and J_{i+1}.To get the uniform estimate for all these, we need a more general argument.

We consider one of the spirals, say S_0,on all its lengths from 0 to∞. Let $\delta > 0$ be taken so that

$$B_{2\delta}(c_0) \cap J \subseteq J_1 \text{ and } B_{2\delta}(c_m) \cap J \subseteq J_m. \quad (79)$$

A topological argument below says there is $\varepsilon > 0$ such thatfor any isometry g of the planethere exists a point $x_g \in J$ with

$$d\left(x_g, g\left(S_0 \backslash B_\delta(0)\right)\right) = inf\{d(x_g, y) | y \in g(S_0 \backslash B_\delta(0))\} \geq \varepsilon. \quad (80)$$

This implies that a map h which transforms J onto some $h(J)$on the other side of a spiralisometric to S_0,outside the δ-neighbourhood of its centre, must fulfil$\|h - id\| \geq \varepsilon$.

Let us set $\eta = min\{\delta, r_*\varepsilon\}$and let us consider the neighbour map $h = f_i^{-1}f_j$with intersecting $J \cap h(J\ ) = \{c_m\}$ as above: $i = im^n, j = (i+1)1^{n'}$. By (79) J_1 does not intersect $B_\delta(c_m)$. By (80) there is $x^* \in J_1$with $d(x^*, h(J) \backslash B_\delta(c_m)) \geq r_*\varepsilon$ (note that $a_1^{-1}J_1$ is isometric with J). By (79)$d(x^*, B_\delta(c_m)) \geq \delta$.Thus$\|h - id\| \geq \eta$.

Now we take disjoint pieces $J_i \subset J_i$ and $J_j \subset J_{i+1}$.If $J_i = J_{im^n u}$ with $u_1 \neq m_a$ and westandardize by this new f_i^{-1}, the separating spiral will be $f_u^{-1}(S_\beta)$which is isometric to S_0,and the distance of its centre to J is larger than $2\delta/r_u$. Again we get $\|h - id\| \geq \eta$.

We get all neighbour maps of J when we add the inverses of those considered(interchanging i and j) and finitely many others. So id is not approached by neighbour mapsand OSC holds.

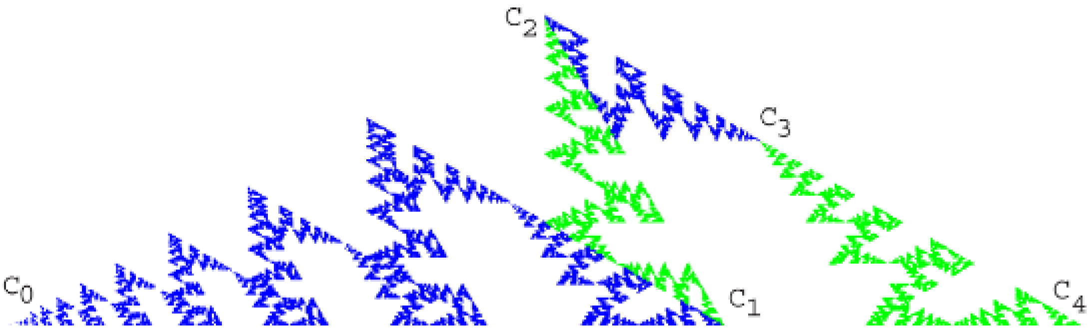


Figure (4)[116]: A self-similar arc with four pieces. At c_1 and c_3 we have finite type, at c_2 infinite type.

At the end, we give the topological argument for (80). In the space $F$of closed subsetsof the closed ball $\overline{B_R(0)}$with Hausdorff metric [122], [83], we consider the subspace F_0 of allisometric copies of compact subsets of $S_0 \backslash B_\delta(0)$or of $\mathbb{R}$ which have approximately the same diameter as J.

$$F_0 = \{g(F\ ) \subseteq \overline{B_R(0)} | g \text{ isometry}, F \subset \mathbb{R} \text{ or } F \subset S_0 \backslash B_\delta(0), \frac{diamJ}{diamF} \in \left(r_*. \frac{1}{r_*}\right)\}.$$

It is well known that F is compact. Moreover,F_0 is a closed subset and so is also compact.(Take a sequence $g_n(F_n) \to G$.If $F_n \to \infty$ then G is a subset of a line. If infinitely many F are within a bounded part of S_0there is a subsequence for which both F_n and g_n converge.)Thus $J \in F$ has a positive Hausdorff distance ε from F_0.(Jcannot be in a spiral by a simplecurvature argument, and J was assumed not to be in a line, so $J \notin F_0$). Take $F \in F_0$.Since $d_H(J, G) \geq \varepsilon$ for every closed $G \subset F$,there is $x^* \in J$ with $d(x^*, F) \geq \varepsilon$.

Section (2.3): Iterated Function Systems of Bounded Distortion

For X be a non-empty closed subset of $\mathbb{R}^n$ and let $I^1 = \{1,\dots,\ell\}$ a finite index alphabet, $\ell > 1$. An iterated function system (IFS) consists of a family $\{\phi_i\}_{i\in I^1}$ of contractions on X. That is, for $i \in I^1$ we have $\phi_i : X \to X$ such that

$$|\phi_i(x) - \phi_i(y)| \le r_i|x - y| \ for\, all\, x, y \in X,$$

where $0 < r_{\min} \le r_i \le r_{\max} < 1$, and $|\cdot|$ represents the euclidean norm in $\mathbb{R}^n$.

An IFS determines a unique non empty compact set K satisfying

$$K = \bigcup_{i\in I^1} \phi_i(K),$$

Which is in general a fractal set [83], [130]. We will denote by I^k the set of all words of length $k, \omega = \omega_1\omega_2,\dots,\omega_k, \omega_j \in I^1$, and by $I^* = \bigcup_{i=1}^{\infty} I^k$ the set of all finite words, for $\omega \in I^*$ we will denote by $|\omega|$ the length of ω. For $\omega = \omega_1,\dots,\omega_k \in I^k$ we will denote $r_\omega = r_{\omega1},\dots,r_{\omega k}, \phi_\omega = \phi_{\omega_1} \circ \cdots \circ \phi_{\omega k}, and\ K_\omega = \phi_\omega(K)$.

We remark that $K_{\lambda\omega} \subseteq K_\lambda$, for all $\omega, \lambda \in I^*$.

We consider the general class of IFS of bounded distortion (BD) [130], that is, we assume that there are constants $M_0, M_1 > 0$ such that

$$M_0|K_\omega||x - y| \le |\phi_\omega(x) - \phi_\omega(y)| \le M_1|K_\omega||x - y|, \tag{81}$$

for all $\omega \in I^*, x, y \in X$; where$|K_\omega|$ represents the diameter of the compact set K_ω. The BD property (81) is satisfied if the $\phi_i^{'}s$ are conformal maps,[131], [132] and in particular for contracting similitudes.

Let us denote by I the space of infinite sequences $\omega = \omega_1\omega_2,\dots; \omega_j \in I^1$ with the usual metric: $(\omega, \lambda) = \sum_{j=1}^{\infty} \frac{|\omega_j - \lambda_j|}{\ell_j}$, for all $\omega, \lambda \in I$. We considerthe natural projection map $\Pi : I \to X$ defined by

$$\Pi(\omega) = \bigcap_{n=1}^{\infty} K_{\omega_1,\dots,\omega_n}.$$

It is clear that $d(\omega, \lambda) \le \frac{1}{\ell^n}$ if $\omega_j = \lambda_j$ for all $j = 1,\dots,n$ and, conversely, $d(\omega, \lambda) < \frac{1}{\ell^n}$ implies $\omega_j = \lambda_j$ for all $j = 1,\dots,n$. Thus Π is a continuous map and $\Pi(I) = \bigcup_{\omega\in I} \Pi(\omega) = K$. For a general subset of successions $G \subset I$, we will denote by G^k the set of words of length $k, \omega \in I^k$ for which there is $\lambda \in I$ such that $\omega\lambda \in G$; also $G^* = \bigcup_{k=1}^{\infty} G^k$.

We say that $G \subseteq I$ is a subsystem of I if G is compact and shift invariant, that is if $\omega = \omega_1\omega_2\omega_3 \cdots \in G$, then $\omega_2\omega_3 \cdots \in G$. Then the compact subset $K_G = \Pi(G)$ is asub-self-similar set satisfying $K_G \subseteq \bigcup_{i\in I^1} \phi_i(K_G)$. Such constructions were studied by Falconer [133] for similitudes. Consider for example four similitudes, $I^1 = \{1,2,3,4\}$, which scale by $1/2$ on $\mathbb{R}^2$: $T_i(x) = \frac{1}{2}x + p_i$ with $p_1 = (0, \frac{1}{2})$, $p_2 = (\frac{1}{2}, \frac{1}{2})$, $p_3 = (0, 0)$ and $p_4 = (0, \frac{1}{2})$. It is clear that the unit square in $\mathbb{R}^2$ is the attractor of the system. The subset of all successions $\omega = \omega_1\omega_2,\dots$ with the restriction that [130] never follows [83] is a subsystem and the corresponding sub-self-similar set is shown in Fig.1. See [134] for related illustrative examples with these transformations.

For a subsystem G let us define the similarity dimension of G by the unique solution s_G of the pressure equation $p_G(s_G) = 0$, where $p_G(s)$ is defined by

$$p_G(s) = \lim_{k\to\infty} \frac{1}{k} \log \left(\sum_{\omega \in G^k} |K_\omega|^s \right).$$

Separation properties are needed to find formulas for the Hausdorff dimensions of K and K_G. If the IFS I satisfies a strong separation property like the Open Set Condition then the Hausdorff dimension of K coincides with the similarity dimension of I, $\dim K = s_I$ and $H^{s_I}(K) > 0$.[1] This result extends for general subsystems as was showed by Falconer [133] and we recall below. On the other hand if I satisfy the weak separation propertyof Lau and Ngai [135] then dim K coincides with the growth dimension of I, [136] which is in general smaller than s_I.

We study the following separation property. Let G be a subsystem of bounded

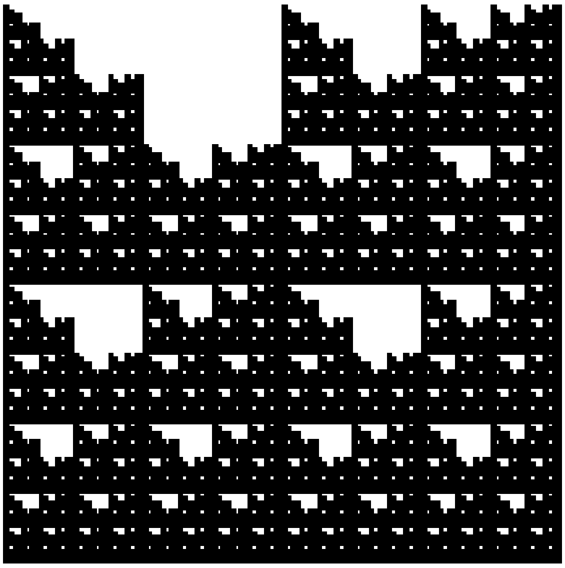


Fig. (1)[129]: A sub-self-similar set in the unit square.

distortion, for a non-empty compact set $A \subseteq X$, we will denote by $G(A)$ the family of finite words

$$G(A) = \{\omega = \omega_1, \ldots, \omega_k \in G^* : |K_\omega| < |A| \le |K_{\omega_1}, \ldots, \omega_{k-1}|; K_\omega \cap A \neq \emptyset\}.$$

We will say that a subsystem G is separated if there exists $M > 0$ such that $\# (G(A)) \le M$ for all compact A, where $\# (\cdot)$ denotes the cardinal (number of elements).

For $G = I$ the separation condition is a reformulation of the Bandt-Graft condition, [120] which is equivalent to the Open Set Condition. See [128] for contracting similitudes and [137] for conformal IFS. However is easy to see that our separation property does not implies the Bandt-Graft condition for general subsystems.

We prove the dimension formula for sub-self-similar sets and some complementary results. The principal result was proved by Falconer [133] but we give a different proof here. We contain original results and examples. We characterize the weak separation property and the growth dimension of K through a subsystem W. We propose a generalized weak separation property and give an example of a fractal set that satisfies the generalized property but not the weak separation property. The existence of such fractal set was questioned in Zerner [136]. We study subsystems for which $H^{s_G}(K_G) > 0$. For $G = I$ this implies separation but we show that this is not the case for general subsystems. The principal result characterizes the subsets of K_G with positive measure where the separation property fails. We show a fractal set where we can find such a subset explicitly.

Let G a subsystem, $A \subseteq X$ a compact set and $k, s > 0$, we will denote

$$c_{G,k}^{s} = \sum_{\omega\in G^{k}} |K_{\omega}|^{s}, and$$
$$c_{G,k}^{s}(A) = \sum_{\substack{\omega\in G^{k}\\ K_{\omega}\cap A\neq\emptyset}} |K_{\omega}|^{s}.$$
Taking into account the bounded distortion and the shift invariance of G is easy to see that
$$c_{G,k+m}^{s} = \sum_{\substack{\eta\lambda\in G^{k+m}\\ |\eta|=k,|\lambda|=m}} |K\eta\lambda|s \leq M_1^{s} c_{G,k}^{s} c_{G,m}^{s}. \quad (82)$$
Thus $log(c_{G,k}^{s})$ is a subadditive sequence, the limit $p_G(s) = \lim_{k\to\infty}\frac{1}{k}\ log(c_{G,k}^{s})$ exists and
$$M_1^{-s}\ e^{kp_G}(s)\ \leq\ c_{G,k}^{s}. \quad (83)$$
We know that $p_G(s)$ is a continuous decreasing function of s and there exists a unique $s_G \geq 0$ such that $p_G(s_G) = 0$ which is the similarity dimension of G. The following conditions are trivial consequences of the definition, see [133].

Proposition (2.3.1)[129]: Let s_G the similarity dimension of G, then
$$s_G = inf\{s \geq 0 : \Sigma_{k=1}^{\infty}\ c_{G,k}^{s} < \infty\} = sup\,\{s \geq 0 : \Sigma_{k=1}^{\infty}\ c_{G,k}^{s} = \infty\},$$
$$s_G = inf\left\{s \geq 0 : \lim_{k\to\infty} c_{G,k}^{s} = 0\right\} = \sup\ \left\{s \geq 0 : \lim_{k\to\infty}\ c_{G,k}^{s} = \infty\right\}$$
It is not true, in general, that $c_{G,k+m}^{s} \geq C c_{G,k}^{s} c_{G,m}^{s}$ for some constant $C > 0$. However, for $G = I$, it is easy to see that $c_{G,k+m}^{s} \geq M_0^{s} c_{I,k}^{s} c_{I,m}^{s}$ and
$$M_1^{-s}\ e^{kp_I(s)}\ \leq\ c_{I,k}^{s}\ \leq\ M_0^{-s}\ e^{kp_I(s)}. \quad (84)$$
We remark that if H, G are subsystems, $H \subseteq G \subseteq I$, then $c_{H,k}^{s} \leq c_{G,k}^{s} \leq c_{I,k,}^{s}\ p_H(s) \leq p_G(s) \leq p_I(s)$, for all $k, s > 0$, and $s_H \leq s_G \leq s_I$.

The following theorem relates $dimK_G$ with s_G. The first part is standard whereas the second part was proved in [133] for similitudes. Falconer's proof extends to the bounded distortion case but we offer a different proof here to keep this work self contained and because it contains techniques that we will use repeatedly.

Theorem (2.3.2)[129]: Let G a subsystem, then dim $K_G \leq s_G$. If , in addition, G is separated then dim $K_G = s_G\ and\ H^{sG}(K_G) > 0$.

Proof. Let $t > s_G$, then $p_G(t) < 0$. For large k we have $\frac{1}{k}\log\ c_{G,k}^{t} \leq -\epsilon < 0$ and $c_{G,k}^{t} \leq e^{-k\epsilon} < 1$. This implies that the t-dimensional Hausdorff measure of K_G is finite, $H^t(K_G) < \infty$, for all $t > s_G$, and thus dim $K_G \leq s_G$.

Assume now that G is separated, we prove that $s_G \leq$ dim K_G. If dim $K_G \geq s_I$ there is nothing to prove, since we have $s_I \leq$ dim $K_G \leq s_G \leq s_I$.

Suppose dim $K_G < s_I$ and let t be such that
$$\text{dim}\ K_G < t < s_I. \quad (85)$$
We want to prove that there is $B > 0$ such that $c_{G,k}^{t} \leq B$ for all k. Then $p_G(t) \leq 0$ which implies $s_G \leq t$ and thus, since this is true for all t satisfying (85), we have $s_G \leq$ dim K_G.

By (85) $H^t(K_G) = 0$ thus, for all $\epsilon > 0$ and taking into account that K_G is compact, there exists a finite cover $K_G \subset\cup U_j$ such that $\sum|U_j|^{t} < \epsilon$. We set $\epsilon < M_1^{-t}\ M^{-1}$, where M is the separation constant, and let l_0 such that $r_{\max}^{l_0}\ |K| < |U_j|$ for all j. We use induction to prove that
$$c_{G,k}^{t} \leq M_0^{-t}\ e^{l_0 p_I(t)}, \quad (86)$$
for all k.

For $k \leq l_0$ we have that

$$c^t_{G,k} \leq c^t_{I,k} \leq M_0^{-t} e^{kp_I(t)} \leq M_0^{-t} e^{l_0 p_I(t)},$$

by (84) and since $p_I(t) > 0$ by (85). Now, suppose $l > l_0$ and that (86) is truefor all $k < l$, we evaluate

$$c^t_{G,l}(U_j) = \sum_{\substack{\omega \in G^l \\ K_\omega \cap U_j \neq \emptyset}} |K_\omega|^t.$$

We remark that for each $\omega \in G^l, K_\omega \cap U_j = \emptyset$, we have $\left|K_{\omega_1,\dots,\omega_{l-1}}\right| \leq r_{\omega_1,\dots,\omega_{l-1}}|K| \leq r_{\max}^{l-1}|K| \leq r_{\max}^{l_0}|K| < |U_j|$, then ω has an initial word $\eta, \omega = \eta\lambda$, such that $\eta \in G(U_j)$. Thus

$$c^t_{G,l}(U_j) \leq \sum_{\eta \in G(U_j)} \sum_{\eta\lambda \in G^l} \left|K_{\eta\lambda}\right|^t \leq \sum_{\eta \in G(U_j)} \sum_{\eta\lambda \in G^{l-|\eta|}} \left|K_{\eta\lambda}\right|^t$$

$$\leq M_1^t \sum_{\eta \in G(U_j)} \left|K_\eta\right|^t c^t_{G,l-|\eta|} \leq M_0^{-t} e^{l_0 p_I(t)} M_1^t \sum_{\eta \in G(U_j)} \left|K_\eta\right|^t,$$

by the inductive hypothesis,

$$\leq M_0^{-t} e^{l_0 p_I(t)} M_1^t M \left|U_j\right|^t,$$

since $|K_\eta| < |U_j|$ and $\# (G(U_j)) \leq M$ by the separation property. Thus, since $\{U_j\}$ is a finite cover of K_G,

$$c^t_{G,l} \leq \sum_j c^t_{G,l}(U_j) \leq M_0^{-t} e^{l_0 p_I(t)} M_1^t M \sum_j \left|U_j\right|^t \leq M_0^{-t} e^{l_0 p_I(t)} M_1^t M\epsilon$$

$$\leq M_0^{-t} e^{l_0 p_I(t)},$$

Since $\sum_j \left|U_j\right|^t < \epsilon < M_1^{-t} M^{-1}$. This complete the proof of (86) and hence $\dim K_G = s_G$. Now, to prove that $H^{s_G}(K_G) > 0$, we simply remark that the separation property implies the conditions of Falconer ([138]) for K_G.

Recurrent sets and graph-directed sets [139],[140] are standard generalizations of IFS that corresponds to particular types of subsystems [133]. We consider the following definition. Let $R \subseteq I^n$ and let $[I|R]$ be the subset of all successions $\omega = \omega_1\omega_2 \cdots \in I$ with the restriction that $\omega_{i+1}\omega_{i+2},\dots,\omega_{i+n} \in R$ for all i. It is easy to see that $[I|R]$ is in fact a subsystem of I and we say that $[I|R]$ is arecurrent subsystem defined by therestriction R. The sub-self-similar set of Fig. 1 for example correspond to a recurrent subsystem for the restriction $R = I^2 - \{12\}$.

We associate $[I|R]$ with a directed graph. Let $[I|R]^{n-1}$ be the vertex set, and we draw an edge from λ to η ifand only if $\lambda\eta \in [I|R]^*$.If thisdirected graph is strongly connected (i.e. every two vertices can be connected through a directed path) we say that $[I|R]$ is aconnected recurrent subsystem. The classical theory of IFS extends to connected recurrent subsystems and separation properties were studied in that case [141].

If G is a subsystem we consider the recurrent sub-systems $[I|Gk]$. It is clear that $[I|G^k] \supseteq [I|G^{k+1}] \supseteq G$ for all k and $G = \bigcap_{k=1}^{\infty} [I|G^k]$ by compactness . Moreover $s_{[I|G^k]}$ tends to s_G as we would expect.

Proposition (2.3.3)[129]: $\lim_{k\to\infty} s_{[I|G^k]} = s_G$.

Proof. It is clear that $s_{[I|G^k]}$ is a non-increasing sequence $s_{[I|G^k]} \geq s_{[I|G^{k+1}]} \geq s_G$, then $\lim_{k\to\infty} s_{[I|G^k]} = t \geq s_G$. Let $s > s_G$, then

$$c^s_{[I|G^k],km} \leq M_1^{s(m-1)} \left(c^s_{[I|G^k],k}\right)^m$$
$$= M_1^{-s} \left(M_1^s c^s_{G,k}\right)^m,$$

by (82) and taking into account that $[I|G_k]^k = G_k$. Proposition (2.3.1) implies that $M_1^s c^s_{G,k} < 1$ fork great enough, then $c^s_{[I|G^k],km} \to 0$ when m $\to\infty$ which implies $t \leq s_{[I|Gk]} \leq s$ and thus $t = s_G$.

In some cases, we need to approximate G through a family of systems which are not necessarily subsystems of I. Let $\{J_h^1\}$ a family offinite subsets $J_h^1 \subset G^*$ indexed by a positive real orinteger parameter h. We say that $\{J_h^1\}$ approaches G if for each h there are positive integer numbers $k_0(h)$ and $k_1(h)$ such that: $k_0(h) \leq |\lambda| \leq k_1(h)$ for all $\lambda \in J_h^1$; $k_0(h) \to \infty$ when $h \to \infty$ and, for all $\omega \in G^{k_1}(h)$, there exists at least one $\lambda \in J_h^1$ such that $\omega = \lambda\eta$.

Each J_h^1 is in fact an alphabet for a system J_h which is not a subsystem of I according to our definition. However, we have the following result which relates the similarity dimensions of I and J_h.

Proposition (2.3.4)[129]: Let $\{J_h^1\}$ a family that approaches G, then $s_G \leq s_{J_h}$ for all h and $\lim_{h\to\infty} s_{J_h} = s_G$.

Proof. We show first that $s_G \leq s_{J_h}$. For $m \in N$, we remark that for each $\omega \in G^{k_1(h)m}$ there exists at least one $\alpha \in J_h^*$ such that $\omega = \alpha\beta$ with $|\beta| < k_1(h)$, thus

$$c^s_{G,k_1(h)m} \leq \sum_{\alpha,\beta} \left|K_{\alpha\beta}\right|^s,$$

where $|\alpha\beta| = k_1(h)m, \alpha \in J_h^*$ and $|\beta| < k_1(h)$. Then if $\alpha \in J_h^n$ it results that $m - 1 < n \leq (k_1(h)/k_0(h))m$. Let m' the integer part of $(k_1(h)/k_0(h))m$ and $M = max\,\{c^s_{G,k} : k < k_1(h)\}$, then

$$c^s_{G,k_1(h)m} \leq M_1^s M \sum_{\alpha} |K_\alpha|^s \leq M_1^s M \sum_{n=m}^{(m')} c^s_{Jh,n}.$$

If $s > s_{J_h}$ then $\sum_{n=m}^{m'} c^s_{J_h,n} \to 0$ when $m \to \infty$ by Proposition (2.3.1). Thus $\lim_{m\to\infty} c^s_{G,k_1(h)m} = 0$ for all $s > s_{J_h}$ and $s_G \leq s_{J_h}$. Now let $s > s_G$ and observe that

$$c_{J_h,m} \leq (M_1^s)^{m-1} \left(\sum_{\lambda\in J_h^1} |K_\lambda|^s\right)^m$$
$$\leq M_1^{-s} \left(M_1^s \sum_{n=k_0(h)}^{k_1(h)} c^s_{G,n} \right)^m.$$

By Proposition (2.3.1) $\sum_{n=k_0(h)}^{k_1(h)} c^s_{G,n} < M_1^{-s}$ for all h great enough, then $c_{J_h}, m \to 0$ when $m \to \infty$ and this implies that $s_{J_h} \leq s$. Thus $\lim_{h\to\infty} s_{J_h} = s_G$.

We will relate the weak separation property (WSP) of Lau and Ngai [135], [136] with the existence of a separate subsystem and propose a generalization for IFS of bounded distortion. First, we make the following general observation whose proof is straightforward from the subsystem definition.

Proposition (2.3.5)[129]: Suppose we have a succession of subsets $G_1 \subseteq I^1, G_2 \subseteq I^2, \ldots$ such that, for all word $\omega_1\omega_2, \ldots, \omega_n \in G_n$, we have that $\omega_2, \ldots, \omega_n \in G_{n-1}$. Then $G = \{\omega \in I : \omega_1\omega_2, \ldots, \omega_n \in G_n, for\ all\ n\}$ is a subsystem of I and $G^n \subseteq G_n$.

Now, we introduce a total order on I^* by setting $\lambda < \omega$ if $|\lambda| < |\omega|$, and the lexicographic order if $|\lambda| = |\omega|$. Let us define a subsystem W by

$$W_n = \{\omega \in I^n : \phi_\omega \neq \phi_\lambda;\ for\ all\ \lambda < \omega\},$$

and $W = \{\omega \in I : \omega_1, \ldots, \omega_n \in W_n, for\ all\ n\}$. This is a subsystem by the preceding proposition (we will sketch the argument in Theorem (2.3.7)), moreover it is easy to see that and $K_W = K$.

We recall now some definitions of Zerner [136]. In what follows we assume that the ϕ_i's are similitudes and K is in general position, i.e. not contained in a hyperplane. For $a, b > 0$ and $A, U \subseteq \mathbb{R}^n$ let us define

$$F = \{\phi_\omega : \omega \in I^*\} = \{\phi_\omega : \omega \in W^*\},$$
$$F_b = \{\phi_\omega \in F : r_\omega \in (b\, r_{\min}, b\,]\}, and$$
$$F_{a,U,M} = \{\phi_\omega \in F_{a|U|} : \phi_\omega(M) \cap U \neq \emptyset\}.$$

It is clear that there is a one to one correspondence between F and W^*. Thegrowth dimension β_I of I is defined as the exponential growth rate of $\nparallel(F_b)$ for $b \to 0$. In our notation, consider the family $\{J_h^1\}$ where $J_h^1 = \{\omega \in W^*: r_\omega \in (\frac{1}{h} r_{\min}, \frac{1}{h}]\}$ for $h > 0$, then $F_{\frac{1}{h}} = \{\phi_\omega : \omega \in J_h^1\}$ and $\beta_I = \lim_{h\to\infty} s_{J_h}$ ([136]). Setting $a > 0$ and $M \subseteq \mathbb{R}^n$ non-empty, we say that I satisfy WSP if and only if $\nparallel(Fa, U, M)$ is bounded for all $U \subseteq \mathbb{R}^n$ ([136]). In that case dim $K = \beta_I$. With respect to the subsystem W, we have the following result.

Theorem (2.3.6)[129]: $\beta_I = s_W$ and I satisfy WSP if and only if W is separated. In that case dim $K = s_W\ and\ H^{s_W}(K) > 0$.

Proof. First we observe that $\{J_h^1\}$ is a family thatapproaches W, then $\beta_I = \lim_{h\to\infty} s_{J_h} = s_W$ by Proposition (2.3.4). Now, we set $a = \frac{1}{|K|}$, and $M = K$, then by the one to one correspondence between F and W∗,

$$\nparallel(F_{a,U,M}) = \nparallel\{\omega \in W^* : r_{\min}|U| < |K_\omega| \leq |U|;\ K_\omega \cap U \neq \emptyset\}, \quad (87)$$

since $|K_\omega| = r_\omega|K|$ for similitudes. We denote by W_U the right-hand set of (87), and compare it with $W(U)$. Let $\omega \in W(U)$, then $|U| \leq |K_{\omega_1,\ldots,\omega_{k-1}}|$ and $r_{\min}|U| \leq |K_\omega|$. If $r_{\min}|U| \leq |K_\omega|$ then $\omega \in WU$, on the other hand if$r_{\min}|U| \leq |K_\omega|$, then $r_{\min}|U| < |K_{\omega_1,\ldots,\omega_{k-1}}| = |U|$ and $\omega_1, \ldots, \omega_{k-1} \in W_U$. Con-versely let $\omega \in W_U$. If $|K_\omega| < |U|$ then there is a $h \leq k$ such that $|K_{\omega_1,\ldots,\omega_h}| < |U| \leq |K_{\omega_1,\ldots,\omega_{h-1}}|$ and $\omega_1, \ldots, \omega_h \in W(U)$, on the other hand if $|K_\omega| = |U|$, then $|K_{\omega j}| < |U|$ and $\omega j \in W(U)$ for some $j \in I$.

Thus $\nparallel$ (Fa,U,M) is bounded if and only if $\nparallel(W(U))$ is bounded, i.e. if and only if W is separated. The last assertion follows directly from Theorem (2.3.2) since $K_W = K$.

Now, we move to the bounded distortion case to generalize WSP. The subsystem W was constructed by eliminating words ω such that $K_\omega = K_\lambda$, for some word $\lambda < \omega$. We propose a direct generalization by eliminating words ω such that $K_\omega \subset K_{\lambda(1)} \cup \cdots \cup K_{\lambda(k)}$ for some words $\lambda(1), \ldots, \lambda(k) < \omega$. Specifically, we define: $GW_n = \{\omega \in I^n : K_\omega \nsubseteq K_{\lambda(1)} \cup \cdots \cup K_{\lambda(k)};$

$$for\ all\ \lambda(1),\dots,\lambda(k) < \omega\},$$
$$and\ GW = \{\omega \in I : \omega_1,\dots,\omega_n \in GW_n, for\ all\ n\}.$$

Theorem (2.3.7). GW is a subsystem with $K_{GW} = K$. If GW is separated then dim $K = s_{GW}$ and $H^{s_{GW}}(K) > 0$.

Proof. We will prove that GW is a subsystem by using Proposition (2.3.5). That $K_{GW} = K$ is clear from the definition and then the theorem follows from Theorem (2.3.2). Let $\omega_1\omega_2,\dots,\omega_n \in GW_n$ and suppose that $\omega_2,\dots,\omega_n \notin GW_{n-1}$, then there are $\lambda(1),\dots,\lambda(k) < \omega_2,\dots,\omega_n$ such that

$$K_{\omega_2,\dots,\omega_n} \subset K_{\lambda(1)} \cup\dots\cup K_{\lambda(k)}.$$

We then remark that $\omega_1\omega_2,\dots,\omega_n > \omega_1\lambda(i)$ and

$$\begin{aligned} K_{\omega_1\omega_2,\dots,\omega_n} &= \phi_{\omega_1}\left(K_{\omega_2,\dots,\omega_n}\right) \\ &\subset \phi_{\omega_1}(K_{\lambda(1)} \cup\dots\cup K_{\lambda(k)}) \\ &\subset \phi_{\omega_1}(K_{\lambda(1)}) \cup\dots\cup \phi_{\omega_1}(K_{\lambda(k)}) \\ &= K_{\omega_1\lambda(1)} \cup\dots\cup K_{\omega_1\lambda(k)}, \end{aligned}$$

which is a contradiction. Thus $\omega_2,\dots,\omega_n \in GW_{n-1}$ and GW is a subsystem.

If GW is separated we say that I satisfy GWSP. The next example shows a system which satisfy GWSP but do not satisfy WSP.

Example (2.3.8)[129]: Let $X = [0,1] \times [0,1]$ the unit squarein $\mathbb{R}^2$ and let $I1 = \{1,2,3,4,5,6,7\}$. We define seven transformations on X: $\phi_i(x) = \frac{1}{3}x + p_i$ with $p_1 = (0,\frac{2}{3})$, $p_2 = (\frac{1}{3},\frac{2}{3})$, $p_3 = (\frac{2}{3},\frac{2}{3})$, $p_4 = (0,0)$, $p_5 = (\frac{1}{3},0)$, $p_6 = (\frac{2}{3},0)$ and $p_7 = (\delta,0)$, where $0 < \delta < \frac{1}{3}$ is an irrational number. In Fig. 2 we schematize these transformations and the third iter-ate of the IFS, $\bigcup_{\omega\in I^3} \phi_\omega(X)$, which approximates K and shows the overlapping effect. It is easy to see that $K = [0,1] \times C$, where C is the usual Cantor set. Now we consider the subsystems W and GW. First we observe that if $\omega = \omega_1,\dots,\omega_k \in I^*$, then

$$\phi_\omega(x) = \left(\frac{1}{3}\right)^k x + (q_1 + q_2\delta, q_3),$$

Where q_1, q_2, q_3 are rational numbers, $0 \le q_1, q_3 \le 1$ and $0 \le q_2 \le \frac{3}{2}$. Moreover $q_2 = 1^{e1} + \left(\frac{1}{3}\right)^{e_2} + \dots + \left(\frac{1}{3}\right)^{(n-1)e_n}$ $,where$

$$e_i = \begin{cases} 0\ if\ \omega i \neq 7 \\ 1\ if\ \omega i = 7 \end{cases}.$$

Taking this into account we can see that $\phi_\omega = \phi_\lambda$ if and only $if\ \omega = \lambda$ for all $\omega, \lambda \in I^*$. Indeed, if $\phi_\omega = \phi_\lambda$, then $\omega_i = 7$ if and only if $\lambda_i = 7$ and then it must be that $\omega = \lambda$. Therefore $W = I$ which is not separated, that is I does not satisfy WSP.

On the other hand $GW_1 = GW^1 = \{1,2,3,4,5,6\}$ since $K_7 \subseteq K_4 \cup K_5$, then $GW = \{1,2,3,4,5,6\}^\infty$ which is separated (satisfy OSC) and I satisfy GWSP.

We remark that K is a plane self-similar set satisfying dim $K = s_{GW} = \frac{log6}{log3} < \frac{log7}{log3} = s_W <$ dim $X = 2$, whose existence was questioned in Zerner [136].

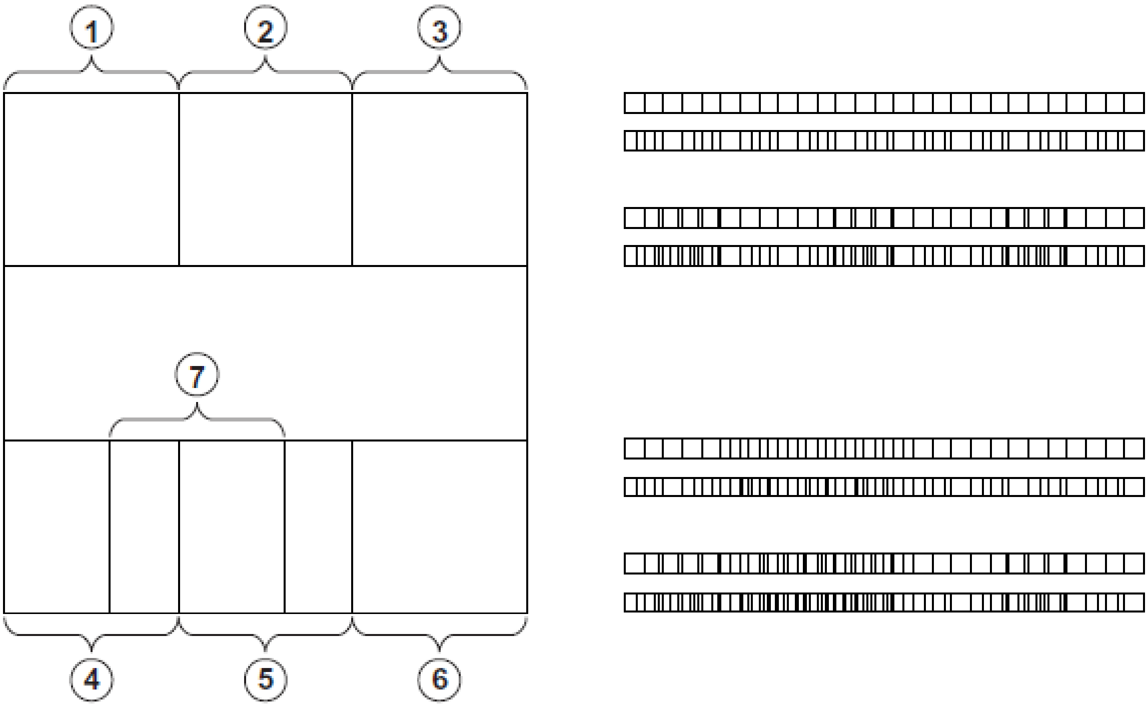


Fig. (2)[129]: A system satisfying GWSP but not WSP. **Left side:** the seven transformations which define the IFS over the unit square. **Right side:** the third iterate of the IFS.

We study subsystems with $H^{s_G}(K_G) > 0$. We know that $H^{s_G}(K_G) > 0$ implies that G is separated if $G = I$ (see [128] and [142] for similitudes and [137] for conformal systems), and it is also true if G is a connected recurrent subsystem

To analyze the general case we must introduce some notation. Let $Z \subset K_G$ the set of points where the separation property fails: $Z = \{z \in K_G : for\ all\ N > 0$ there exists a closed set A such that $z \in A$ and $\nparallel (G(A)) \geq N\}$. It is easy to see that G is separated if and only if $Z = \emptyset$. For $\omega \in G$ let us define:

$$\underline{b}(\omega, n) = lim \inf_{m\to\infty} \sum_{\substack{\lambda \in G^m \\ \omega_1 \dots \omega_n \lambda \in G^{n+m}}} |K_\lambda|^{s_G},$$

$$\underline{b}(\omega) = \inf_n \underline{b}(\omega, n)$$

$$and$$

$$\overline{b}(\omega, n) = \lim \sup_{m\to\infty} \sum_{\substack{\lambda \in G^m \\ \omega_1 \dots \omega_n \lambda \in G^{n+m}}} |K_\lambda|^{s_G},$$

$$b(\omega) = \sup_n b(\omega, n)$$

We remark that $0 \leq b(\omega) \leq b(\omega, n) \leq b(\omega, n) \leq b(\omega) \leq +\infty$. The functions $b(\omega)$ and $b(\omega)$ are not continuous in general, but we have the following

Proposition (2.3.9)[129]: Let $\omega \in G$ and $\epsilon > 0$.

(i) There is $\delta > 0$ such that $\underline{b}(\eta) \leq \underline{b}(\omega) + \epsilon$ for all $\eta \in G$ such that $(\omega, \eta) < \delta$.

(ii) If $\overline{b}(\omega) > N$, then there is δ such that $\overline{b}(\eta) \geq N - \epsilon$ for all $\eta \in G$ such that $(\omega, \eta) < \delta$.

Proof. To prove (i), let n_0 such that $\underline{b}(\omega, n_0) \leq \underline{b}(\omega) + \epsilon$, then if $d(\omega, \eta) < 1/\ell^{n_0}$ we have that $\omega_1, \dots, \omega_{n_0} = \eta_1, \dots, \eta_{n_0}$ and

$$\underline{b}(\eta) \leq \underline{b}(\eta, n_0) = b(\omega, n_0) \leq b(\omega) + \epsilon.$$

The proof of (ii) follows in a similar way.

Next, we will characterize subsets of Z with null and positive s_G-Hausdorff measure using these functions. First we will see, in the following example, that is not true in general that $H^{s_G}(K_G) > 0$ implies $H^{s_G}(Z) = 0$.

Example (2.3.10)[129]: We consider the system $I = \{1,2,3,4\}$ and the restrictions $R_1 = \{11,12,21,22\}$, $R_2 = \{33,34,43,44\}$ and $R = R1 \cup R2 \cup \{13,14,23,24\}$. Let $G1 = [I|R1]$, $G2 = [I|R2]$ and $G = [I|R]$ be the corresponding recurrent subsys-tems; $G1$ and $G2$ are connected whereas G is not.

The associated directed graphs are drawn in Fig. 3.

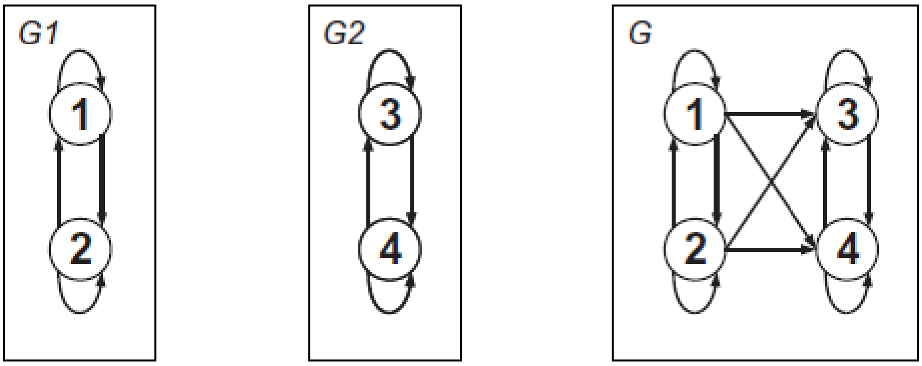


Fig. (3)[129]: Directed graphs associated.

We assume that $s_{G2} \leq s_{G1}$ and prove that $s_G = s_{G1}$. In fact, $s_{G1} \leq s_G$ since $G_1 \subseteq G$. On the other hand

$$c_{G,k}^{s_{G1}} = \sum_{\omega \in G2^k} |K_\omega|^{s_{G1}} + \sum_{h=1}^{k} \sum_{\substack{\alpha \in G1^h \\ \beta \in G2^{k-h}}} |K_{\alpha\beta}|^{s_{G1}}$$

$$\leq c_{G2,k}^{s_{G1}} + M_1^{s_{G1}} \sum_{h=1}^{k} c_{G1,h}^{s_{G1}} c_{G2,k-h}^{s_{G1}}. \quad (88)$$

Since G_1 and G_2 are connected recurrent subsystems we know that there exist $D > 0$ such that $c_{Gi,k}^{s_{Gi}} \leq D$ for $i = 1,2$ and all k. Moreover, taking into account that $s_{G2} \leq s_{G1}$, we can choose D such that $c_{G2,k}^{s_{G1}} \leq D$ for all k (indeed, if $s_{G2} < s_{G1}$ then $c_{G2,k}^{s_{G1}} \to 0$ when $k \to \infty$). Then

$$c_{G2,k}^{s_{G1}} \leq D + kM_1^{s_{G1}} D^2 \leq k(D + M_1^{s_{G1}} D^2) \; and$$

$$\lim_{k\to\infty} \frac{1}{k} \, log c_{G2,k}^{s_{G1}} \leq 0,$$

which implies $s_G \leq s_{G1}$.

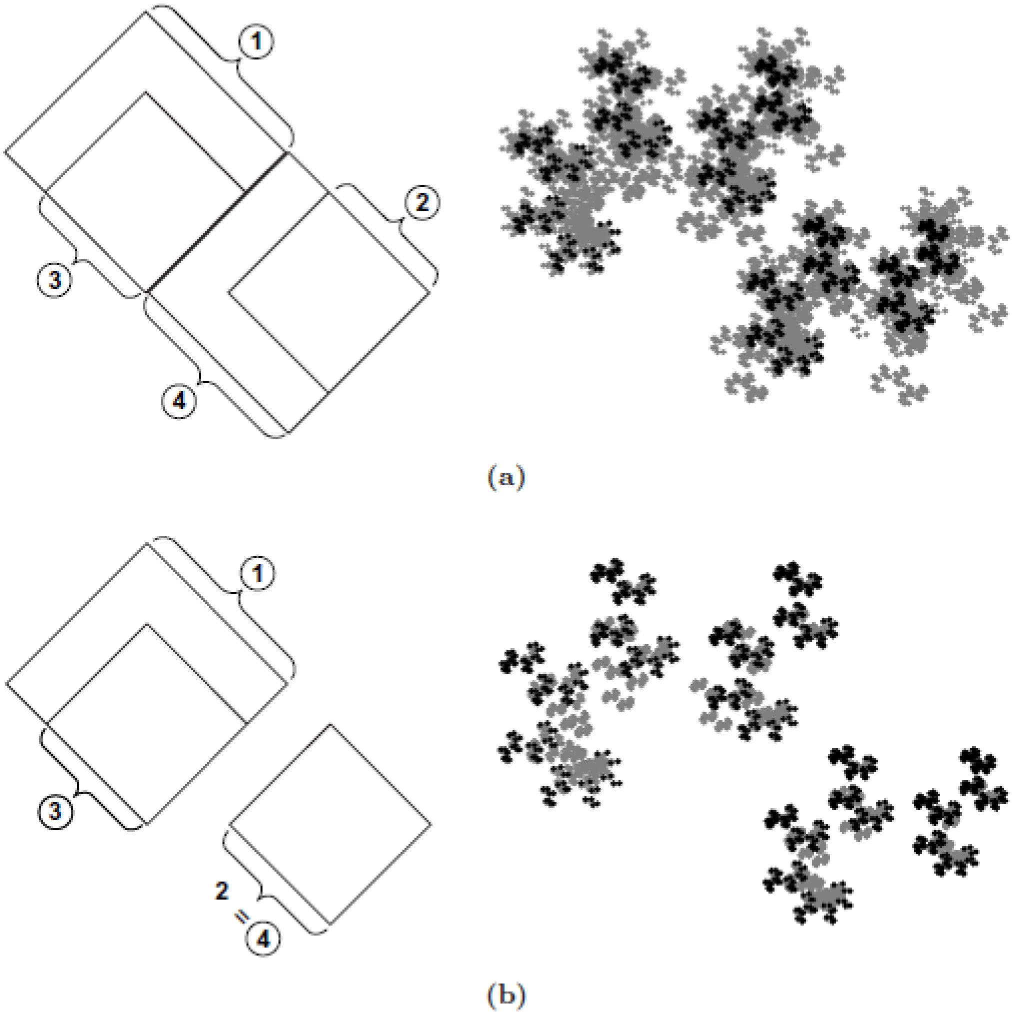


Fig. (4)[129]: Non-separated subsystems of Example (2.3.10). (a) corresponds to the case $s_{G2} = s_{G1} = s_G$, and (b) to the case $s_{G2} < s_{G1} = s_G$. Left side: the four transformations which define the IFS applied to the unit square. Right side: approximations of K_G in gray, and K_{G1}, where the separation property fails, in black.

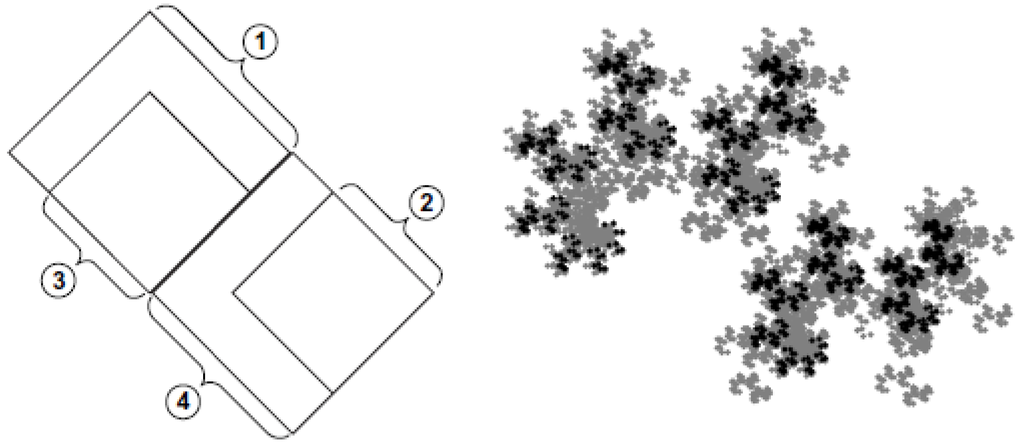
Now, we will consider two particular examples in $\mathbb{R}^2$ where $H^{s_G}(Z) > 0$ and we will study the values of b and b. Let $R : \mathbb{R}^2 \to \mathbb{R}^2$ the rotation of angle $\frac{\pi}{4}$ around the origin and let $, b \in \mathbb{R}, 0 < a < b < \sqrt{2}\ 2$.

Let

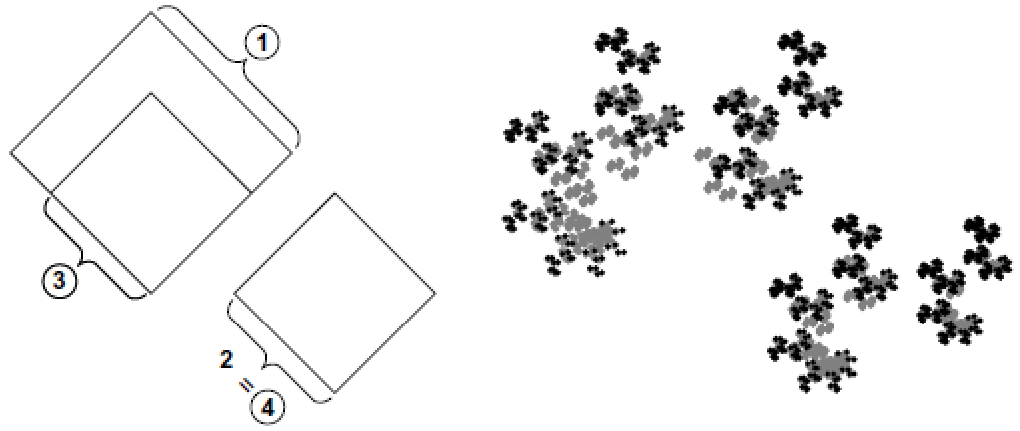
$$\phi_1(x) = r_1 R(x),$$
$$\phi_2(x) = r_2 R(x - (1,0)) + (1,0),$$
$$\phi_3(x) = r_3 R(x),$$
$$\phi_4(x) = r_4 R(x - (1,0)) + (1,0).$$

For the first example we set $r_1 = r_4 = b$ and $r_2 = r_3 = a$, then $s_{G2} = s_{G1} = s_G$. Besides we know that $G1$ is separated, then $H^{s_G}(K_{G1}) > 0$. On the other hand G is not separated: it is easy to see that $0 \in Z$ and then, since for all $\beta \in G, \alpha \in G1$ we have that $\alpha\beta \in G$, it results that $K_{G1} \subseteq Z$.Thus $H^{s_G}(Z) > 0$. In Fig. 4a we show approximations of K_G in gray, and K_{G1} in black.

We can find a lower bound for $c_{G2,k}^{s_G}$in a similar way to (88). Then, taking into account that



(a)



(b)

Fig. 4. Non-separated subsystems of Example (2.3.10). (a) corresponds to the case $s_{G2} = s_{G1} = s_G$, and (b) to the case $s_{G2} < s_{G1} = s_G$. Left side: the four transformations which define the IFS applied to the unit square. Right side: approximations of KG in gray, and $KG1$, where the separation property fails, in black.

$s_{G2} = s_{G1} = s_G$, we have that there is a constant D' such that

$$c_{G,k}^{s_G} \geq D' + kM_0^{s_G}D'^2.$$

Thus, if $\omega \in G_1, \overline{b}(\omega, n) = \limsup_{k\to\infty} c_{G,k}^{s_G} = +\infty$ for all n. Therefore we have that $\overline{b}(\omega) = +\infty$ for all $\omega \in G_1$, it is to say $K_{G1} \subseteq \Pi(\{\omega \in G : \overline{b}(\omega) = +\infty\})$.

For the second example we set $r_1 = b$ and $r_2 = r_3 = r_4 = a$, then $s_{G2} < s_{G1} = s_G$. Again $G1$ is separated, $H^{s_G}(K_{G1}) > 0$ and $K_{G1} \subseteq Z$, see Fig.4b. Now, for $\omega \in G2, \underline{b}(\omega, n) = \liminf_{k\to\infty} c_{G2,k}^{s_G} = 0$ for all n. Then $b(\omega) = 0$ for all $\omega \in G2$ and, moreover, b(ω) = 0 for all $\omega = \alpha\beta$ such that $\beta \in G2$. But every succession $\eta \in G$ may beapproximated by ω's such that $\omega = \alpha\beta, \beta \in G2$, then$K_{G1} \subseteq K_G = \Pi(\overline{\{\omega \in G : b(\omega) = 0\}})$.

In these examples we found a subset K' of Z with positive sG-Hausdorff measure and such that K' is contained in the closure of a subset where $\underline{b}(\omega)$ is arbitrarily small or $\overline{b}(\omega)$ is arbitrarily large. The next Theorem shows that such subset always exist when $H^{s_G}(Z) > 0$.

The proof follows from two lemmas. Let us define

$$B'(\epsilon) = \{\omega \in G : \leq b(\omega), b(\omega) \leq 1/\epsilon\};\ and$$
$$B(\epsilon) = \{\omega \in B'(\epsilon) : d(\Pi(\omega), K_G - \Pi(B'(\epsilon))) \geq \epsilon\},$$

where d corresponds to the euclidean distance.

Lemma (2.3.11)[129]: Let $\eta \in G^*$, if there exist $\beta \in G$ such that $\omega = \eta\beta \in B'(\epsilon)$, then

$$\epsilon/2 \leq \sum_{\substack{\lambda\in G^{k-|\eta|}\\ \eta\lambda\in G^k}} |K_\lambda|^{s_G} \leq 2/\epsilon,$$

for k large enough. In particular, the inequalities follow if η is such that $K_\eta \cap K_G \subseteq \Pi(B'(\epsilon))$.

Proof. From the $B'(\epsilon)$ definition we have that

$$\liminf_{k\to\infty} \sum_{\substack{\lambda\in G^{k-|\eta|}\\ \eta\lambda\in G^k}} |K_\lambda|^{s_G} = \underline{b}(\omega,|\eta|) \geq \underline{b}(\omega) \geq \epsilon,$$

$$\limsup_{k\to\infty} \sum_{\substack{\lambda\in G^{k-|\eta|}\\ \eta\lambda\in G^k}} |K_\lambda|^{s_G} = \overline{b}(\omega,|\eta|) \leq \overline{b}(\omega) \leq 1/\epsilon,$$

and the lemma results from limit properties.

Lemma (2.3.12)[129]: $H^{s_G}(Z \cap \Pi(B(\epsilon))) = 0 \; for \; all \; \epsilon > 0.$

Proof. We fix $\epsilon > 0$ and suppose $Z \cap \Pi(B(\epsilon)) \neq \emptyset$. For $N > 0$ let us denote by A_N the family of sets A such that: $0 < |A| < \epsilon/3, A \cap Z \cap \Pi(B(\epsilon)) \neq \emptyset$ and $\nparallel (G(A)) \geq N$. It is clear that A_N is a Vitali family for $Z \cap \Pi(B(\epsilon))$. Now, for $A \in A_N$ we define $U_A = \bigcup_{\eta\in G(A)} K_\eta \cap K_G$, then $\{U_A\}_{A\in A_N}$ is also a Vitali family for $Z \cap \Pi(B(\epsilon))$. By the Vitali covering theorem ([74]), we have for all $\epsilon_2 > 0$ that there exists a disjoint finite family $\{U_{A_j}\}$ such that

$$H^{s_G}(Z \cap \Pi(B(\epsilon)) \leq \sum_j \left|U_{A_j}\right|^{s_G} + \epsilon_2. \tag{89}$$

Now, we remark that, for k large enough such that $|K_{\omega_1,\dots,\omega_{k-1}}| < |A_j|$ for all j and all $\omega \in G$, we have

$$c_{G,k}^{s_G}\left(U_{A_j}\right) = \sum_{\substack{\omega\in G^k\\ K_\omega\cap U_{A_j}\neq\emptyset}} |K_\omega|^{s_G}$$

$$\geq \sum_{\eta\in G(A_j)} \sum_{\eta\lambda\in G^k} \left|K_{\eta\lambda}\right|^{s_G}$$

$$\geq M_0^{s_G} \sum_{\eta\in G(A_j)} \left|K_\eta\right|^{s_G} \left(\sum_{\substack{\lambda\in G^{k-|\eta|}\\ \eta\lambda\in G^k}} |K_\lambda|^{s_G} \right).$$

We observe that

$$d(K_\eta, K_G - \Pi(B'(\epsilon))) \geq \epsilon/3, \tag{90}$$

for all $\eta \in G(A_j)$, since $K_\eta \cap A_j \neq \emptyset, A_j \cap \Pi\big(B(\epsilon)\big) \neq \emptyset$ and $|K_\eta| < |A_j| < \epsilon/3$. Thus $K_\eta \cap K_G \subseteq \Pi(B'(\epsilon))$ and using the previous lemma we find that

$$c_{G,k}^{s_G}\left(U_{A_j}\right) \geq \frac{M_0^{s_G}\epsilon}{2} \sum_{\eta\in G(A_j)} \left|K_\eta\right|^{s_G} \geq C \sum_{\eta\in G(A_j)} \left|U_{A_j}\right|^{s_G} \geq CN\left|U_{A_j}\right|^{s_G},$$

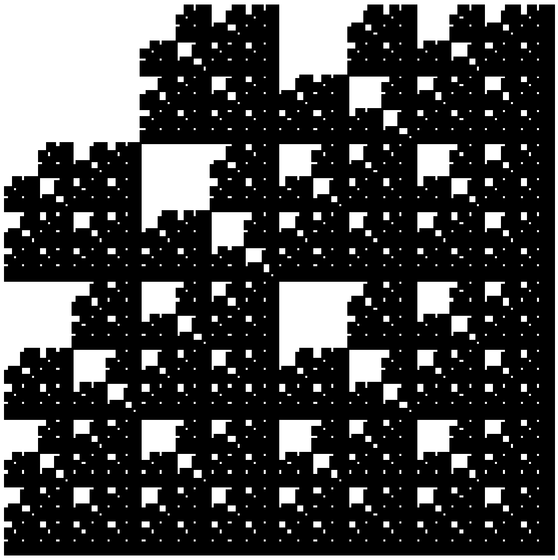


Fig. (5)[129]: A sub-self-similar set corresponding to a connected, not recurrent, subsystem.

where $C = \frac{\left(M_0^{s_G}\right)^2 \epsilon r_{\min}^{s_G}}{2.3^{s_G}}$. Since $|K_\eta| \geq r_{\min}\, M_0|A_j| \geq r_{\min}\, M_0(\left|U_{A_j}\right|/3)$ and $\sharp\,(G(A_j)) \geq N$. Therefore,

$$\sum_j \left|U_{A_j}\right|^{s_G} \leq \frac{1}{CN}\sum_j c_{G,k}^{s_G}(U_{A_j}) \leq \frac{1}{CN}\, c_{G,k}^{s_G}(\Pi(B(\epsilon/3))), \tag{91}$$

for k large enough since U_{A_j} are disjoint sets and $U_{A_j} \subset \Pi(B(\epsilon/3))$ which follows from (90).

Now, let m such that $|K_\eta| < \epsilon/3$ for all $\eta \in G^m$. If $K_\eta \cap \Pi(B(\epsilon/3)) \neq \emptyset$ then $K_\eta \cap K_G \subset \Pi(B'(\epsilon/3))$ and we have that

$$c_{G,k}^{s_G}\left(\Pi\left(B\left(\frac{\epsilon}{3}\right)\right)\right) \leq M_1^{s_G} \sum_{\substack{\eta\in G^m \\ K_\eta\cap\Pi\left(B\left(\frac{\epsilon}{3}\right)\right)\neq\emptyset}} \left|K_\eta\right|^{s_G}\left(\sum_{\substack{\lambda\in G^{k-m} \\ \eta\lambda\in G^k}} |K_\lambda|^{s_G}\right)$$

$$\leq M_1^{s_G}\frac{\epsilon}{6}\left(\sum_{\substack{\eta\in G^m \\ K_\eta\cap\Pi\left(B\left(\frac{\epsilon}{3}\right)\right)=\emptyset}} \left|K_\eta\right|^{s_G}\right),$$

for k large enough. Combining this inequality with (89) and (91) we obtain

$$H^{s_G}(Z \cap \Pi\big(B(\epsilon)\big) \leq \frac{C'}{N} + \epsilon_2,$$

for some constant $C' > 0$ which depends on ϵ but is independent of N and ϵ_2. Thus $H^{s_G}(Z \cap \Pi(B(\epsilon)) = 0$.

Theorem (2.3.13)[129]: Consider the subset

$$Z' = \overline{\Pi\left(\cap_{\epsilon>0}\left(\{\omega \in G : \underline{b}(\omega) \leq \epsilon\} \cup \{\omega \in G : \overline{b}(\omega) \geq 1/\epsilon\}\right)\right)},$$

$then\ H^{s_G}(Z - Z') = 0$.

Proof. The previous lemma implies that

$$H^{s_G}\left(Z \cap \bigcup_{\epsilon>0} \Pi(B(\epsilon))\right) = 0 \,.$$

Then we only need to show that $K_G - \bigcup_{\epsilon>0} \Pi(B(\epsilon)) = Z'$. Let $x \in K_G, x = \Pi(\omega)$ such that $x \notin \bigcup_{\epsilon>0} \Pi(B(\epsilon))$, then either $\omega \notin B'(\epsilon)$ for all ϵ or $\omega \in \Pi(B'(\epsilon))$ but $d\left(x, K_G - \Pi\left(B'(\epsilon)\right)\right) < \epsilon$ for all $\epsilon \leq \epsilon_0$. In the first case, either $\underline{b}(\omega) = 0$ or $b(\omega) = \infty$ and then $x \in Z'$. In the second case, let $\delta = d(x, K_G - \Pi(B'(\epsilon)))$. If $0 < \delta < \epsilon$ then $\omega \in B(\delta)$, since $\omega \in B'(\epsilon) \subset B'(\delta)$ and $d(x, K_G - \Pi(B'(\delta)) \geq d(x, K_G - \Pi(B'(\epsilon)) = \delta$, which is a contradiction. Thus must be $\delta = 0$ for all $\epsilon \leq \epsilon_0$ which implies $x \in Z'$.

At last, we consider connected subsystems, which generalize connected recurrent subsystems. We say that G is a connected subsystem if there exists $T > 0$ such that, for all $\alpha, \beta \in G^*$ there is a $\lambda_\alpha^\beta \in G^*$, such that

$$|\lambda_\alpha^\beta| \leq T \ and \ \alpha\lambda_\alpha^\beta\beta \in G^*. \tag{92}$$

For an example, consider again the four transformations of Fig. 1: $I^1 = \{1,2,3,4\}$, and $T_i(x) = \frac{1}{2}x + p_i$ with $p_1 = (0, \frac{1}{2})$, $p_2 = (\frac{1}{2}, \frac{1}{2})$, $p_3 = (0, 0)$, $p_4 = (0, \frac{1}{2})$. Let $J \subset I^*$ be an infinite subset of words $J = \{11, 141, 1441, 14441, \dots\}$ and let G the subsystem

$$G = \{\omega = \omega_1\omega_2 \cdots \in I : \omega_{i+1}\omega_{i+2,\dots,}\omega_{i+n} \notin J, for\ all\ i, n\}.$$

It is clear that G is not a recurrent subsystem because J is infinite, but it is connected since, for example, $\alpha 2\beta \in G^*$ for all $\alpha, \beta \in G^*$ (see Fig. 5).

For a subsystem G and $\alpha \in G^*$ we will denote $c(\alpha)_{G,k}^s = \sum_{\substack{\lambda \in G^k \\ \alpha\lambda \in G^*}} |K_\lambda|^s$. Now, we can state the following result whose proof uses standard techniques and inequalities like (82) and (83).

Proposition (2.3.14)[129]: If G is a connected subsystem then there are constants C_0, C_1 such that

$$C_1^{-s} e^{kp_G(s)} \leq c(\alpha)_{G,k}^s \leq C_0^{-s} e^{kp_G(s)},$$

for all $\alpha \in G^*$.

Theorem (2.3.15)[129]: Let G a connected subsystem. If $H^{s_G}(K_G) > 0$ then G is separated.

Proof. The previous proposition implies that $C_1^{-s_G} \leq c(\alpha)_{G,k}^{s_G} \leq C_0^{-s_G}$ for all $\alpha \in G^*$. Therefore $C_1^{-s_G} \leq \underline{b}(\omega) \leq \overline{b}(\omega) \leq C_0^{-s_G}$ for all $\omega \in G, Z' = \emptyset$ and $H^{s_G}(Z) = 0$ from Theorem(2.3.13). We want to prove that $Z = \emptyset$. Suppose $Z \neq \emptyset$, let $x \in Z$ and for all $N > 0$ let A_N be a closed set such that $x \in A_N$ and $\#(G(A_N)) \geq N$. Let $\alpha \in G^*, \Lambda = \{\lambda \in G^* : |\lambda| \leq T\}$ and $M = \#(\Lambda)$. As G is connected, we have that for each $\eta \in G(A_N)$ there is a $\lambda \in \Lambda$ such that $\alpha\lambda\eta \in G^*$. If $N > M$ then there exists a $\lambda \in \Lambda$ such that $\#(\{\eta \in G(A_N : \alpha\lambda\eta \in G^*\}) \geq N\, M$, thus $\#(G(\phi_{\alpha\lambda}(A_N)) \geq \frac{N}{M}$. We can see in consequence that for all $\alpha \in G^*$ there is a $\omega \in G$ such that $z = \Pi(\alpha\omega) \in Z$, therefore must be $H^{s_G}(Z) > 0$ which is a contradiction. Then $Z = \emptyset$ and G is separated.

We have presented thenotion of separate subsystem of bounded distortion and shown how it can help to characterize separation properties in a general. We think it could be useful to address other related problems for IFS with overlaps. Besides, we study the problem of when a subsystem with positive Hausdorff measure in its similarity dimension is separated. Taking into account Example (2.3.10) and Theorems (2.3.13) and (2.3.15) it is not true in general but it is true for connected subsystems. Connected subsystems generalize recurrent connected subsystem but we don't know if it is possible to relax the connectivity condition (92) to obtain a more general family of IFS for which positive Hausdorff measure in its similarity dimension implies separation.

Chapter 3
Cantorvals with Topological and Measure Properties

We show that all known examples of x's with E(x) being Cantorvals. We obtain the results that are applied to studying partial sumsets $E(x) = \{\sum_{n=0}^{\infty} x_n\varepsilon_n : (\varepsilon_n)_{n\in\omega} \in \{0,1\}\}$ of some (multigeometric) sequences $x = (x_n)_{n\in\omega}$. We show that the subsum set of an absolutely summable sequence is one of the following: a finite union of (nontrivial) compact intervals, a Cantor set, or a "symmetric Cantorval," a hybrid Cantor-like set with both trivial and nontrivial components.

Section (3.1): Multigeometric Sequences

Suppose that $x = (x\,(0), x\,(1), x\,(2), \ldots)$ is an absolutely summable sequence with infinitely many nonzero terms (i.e. $x \in l_1 \backslash c_{00}$) and let

$$E\,(x) = \left\{\sum_{n=1}^{\infty} \varepsilon_n x(n) : \varepsilon_n \in \{0,1\}\right\}$$

denote the set of all subsums of the series $\sum_{n=1}^{\infty} x(n)$, called the achievement set of x. It is easily seen that for $x = \left(\frac{2}{3}, \frac{2}{3^2}, \frac{2}{3^3}, \ldots\right)$ the set $E(x)$ is equal to the Cantor tenary set C, and for $x = \left(\frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \ldots\right)$ we have $E(x) = [0,1]$.

Achievement sets have been consider by many, some results have been proved several times (see, for example, [150] and [148]) and even conjectures formulated, despite the fact that suitable counterexamples had been earlier published (compare [147], [151] and [156]). Recently, an interesting survey of properties of achievement sets for various (even divergent) sequences was presented by Rafe Jones in [149]. In particular, the example from [149] (due to Velleman and Jones), which will be described in Theorem (3.1.2) and Example (3.1.5).

The following properties of sets E(x) were described in 1914 by S. Kakeya in [150]:

I. E(x) is a compact perfect set.

II. If $|x(n)| > \sum_{i>n} |x(i)|$ for n sufficiently large, then E(x) is homeomorphic to the Cantor set C.

III. If $|x(n)| \geqslant |x(i)|$ for n sufficiently large, then $E(x)$ is a finite union of closed intervals. Moreover, if $|x(n)| > |x(n+1)|$ for almost all n and $E(x)$ is a finite union of closed intervals, then $|x(n)| \leqslant \sum_{i>n} |x(i)|$ for n sufficiently large.

In the same Kakeya formulated the hypothesis that, for any $x \in l_1 \backslash c_{00}$, the set E(x) is homeomorphic to C or is a finite union of closed intervals. In 1980 it was shown that the Kakeya conjecture is false [157]. We recall a number of examples in the literature which demonstrate the falseness of the conjecture. A. D. Weinstein and B. E. Shapiro in [157] gave an example of a sequence a with $a(n) > a(n+1) > 0$ for all n, and $a(n) > \sum_{i>n} a(i)$ for infinitely many n (hence E(a) is not a finite union of intervals), but having the property that the set E(a) contains an interval. The sequence a is defined by the formulas:
$a(5n+1) = 0{,}24 \cdot 10^{-n}, a(5n+2) = 0{,}21 \cdot 10^{-n}, a(5n+3) = 0{,}18 \cdot 10^{-n}, a(5n+4) = 0{,}15 \cdot 10^{-n}, a(5n+5) = 0{,}12 \cdot 10^{-n}$. So,

$$a = \left(\frac{3\cdot 8}{10}, \frac{3\cdot 7}{10}, \frac{3\cdot 6}{10}, \frac{3\cdot 5}{10}, \frac{3\cdot 4}{10}, \frac{3\cdot 8}{100}, \ldots\right).$$

However, they did not justify why the interior of E(a) is non-empty.

Independently, C. Ferens ([146]) constructed a sequence b such that E(b) is not a finite union of intervals but contains an interval, putting $b(5l - m) = (m+3)\frac{2^{l-1}}{3^{3l}}$ for $m = 0, 1, 2, 3, 4$ and $l = 1, 2, \ldots$. Therefore

$$b = \left(7 \cdot \frac{1}{27}, 6 \cdot \frac{1}{27}, 5 \cdot \frac{1}{27}, 4 \cdot \frac{1}{27}, 3 \cdot \frac{1}{27}, 7 \cdot \frac{2}{27^2}, \ldots\right).$$

J. A. Guthrie and J. E. Nymann gave a simpler example of a sequence which achievement set is not a finite union of closed intervals and is not homeomorphic to the Cantor set, defining a sequence by formulas:

$$c(2n-1) = \frac{3}{4^n} \quad \text{and} \quad c(2n) = \frac{2}{4^n} \quad \text{for } n = 1, 2, \ldots$$

In [147], [154] and [155] J. E. Nymann with J. A. Guthrie and R. A. S´aenz characterized the topological structure of the set of subsums of infinite series in the following manner:

Theorem (3.1.1)[143]: For any $x \in l_1 \backslash c_{00}$, the set E(x) is one of the following types:

(i) a finite union of closed intervals;

(ii) homeomorphic to the Cantor set;

(iii) homeomorphic to the set E(c) (of subsums of the sequence $\left(\frac{3}{4}, \frac{2}{4}, \frac{3}{16}, \frac{2}{16}, \frac{3}{64}, \ldots\right)$).

Note, that the set E(c) is homeomorphic to $C \cup \bigcup_{n=1}^{\infty} S_{2n-1}$, where Sn denotes the union of the 2^{n-1} open middle thirds which are removed from [0, 1] at the n-th step in the construction of the Cantor ternary set C. Such sets are called Cantorvals (to emphasize their similarity to unions of intervals and to the Cantor set simultaneously). Formally, a Cantorval (an $\mathcal{M}$-Cantorval - compare [152]) is a non-empty compact subset S of the real line such that S is the closure of its interior, and both endpoints of any component with non-empty interior are accumulation points of one-point components of S.

Theorem (3.1.1) states that the space l_1 can be decomposed into four sets $c_{00}, \mathcal{C}, I$ and $\mathcal{MC}$, where I consists of sequences x with E(x) equal to a finite union of intervals, $\mathcal{C}$ consists of sequences x with E(x) homeomorphic to the Cantor set, and $\mathcal{MC}$ consists of sequences x with E(x) being Cantorvals. Some algebraic properties and topological (Borel) classification of these subsets of l_1 have been recently discussed in [144].

Finally, in Jones' [149] there is presented a sequence

$$d = \left(\frac{3}{5}, \frac{2}{5}, \frac{2}{5}, \frac{2}{5}, \frac{3}{5} \cdot \frac{19}{109}, \frac{2}{5} \cdot \frac{19}{109}, \frac{2}{5} \cdot \frac{19}{109}, \frac{2}{5} \cdot \frac{19}{109}, \frac{3}{5} \cdot \left(\frac{19}{109}\right)^2, \ldots\right).$$

In [149], R. Jones shows a continuum of sequences generating Cantorvals, indexed by a parameter q, by proving that, for any positive number q with

$$\frac{1}{5} \leqslant \sum_{n=1}^{\infty} q^n < \frac{2}{9}$$

(i.e. $\frac{1}{6} \leqslant q < \frac{2}{11}$) the sequence

$$\left(\frac{3}{5}, \frac{2}{5}, \frac{2}{5}, \frac{2}{5}, \frac{3}{5}q, \frac{2}{5}q, \frac{2}{5}q, \frac{2}{5}q, \frac{3}{5}q^2, \ldots\right)$$

is not in $\mathcal{C}$ nor I, so it belongs to $\mathcal{MC}$. Based on Jones' idea, we will describe one-parameter families of sequences which contain (in particular) a, b, d and many others.

For any $q \in \left(0, \frac{1}{2}\right)$ we will use the symbol $(k_1, k_2, \ldots, k_m; q)$ to denote the sequence $(k_1, k_2, \ldots, k_m, k_1 q, k_2 q, \ldots, k_m q, k_1 q^2, k_2 q^2, \ldots, k_m q^2, \ldots)$. Such sequences we will call multigeometric.

Theorem (3.1.2)[143]: Let $k_1 \geqslant k_2 \geqslant \cdots \geqslant k_m$ be positive integers and $= \sum_{i=1}^{m} k_i$. Assume that there exist positive integers n_0 and n such that each of numbers $n_0, n_0+1, \dots, n_0+n$ can be obtained by summing up the numbers $k_1, k_2, \dots, k_m$ (i.e. $n_0 + j = \sum_{i=1}^{m} \varepsilon_i k_i$ with $\varepsilon_i \in \{0,1\}, j = 1, \dots, n$) . If $q > \frac{1}{n+1}$ then $E(k_1, \dots, k_m; q)$ has a nonempty interior. If $q < \frac{k_m}{K+k_m}$ then $E(k_1, \dots, k_m ; q)$ is not a finite union of intervals. Consequently, if

$$\frac{1}{n+1} \leqslant q < \frac{k_m}{K+k_m}$$

then $E(k_1, \dots, k_m; q)$ is a Cantorval.

Proof. Denote $x\_q = (k_1, \dots, k_m; q)$. We start with showing that, for $q < \frac{k_m}{K+k_m}, E(x_q)$ is not a finite union of closed intervals. Observe first, that the sequence x_q is non-increasing. Indeed, from the inequality $< \frac{k_m}{K+k_m}$, it follows that $qK + qk_m < k_m$, and

$$k_m > \frac{qK}{1-q} > qK > qk_1.$$

Moreover, using the same inequality, we obtain

$$\sum_{i>m} x_q(i) = K \sum_{j=1}^{\infty} q^j = K \frac{q}{1-q} < k_m.$$

Hence, for any $n \in \mathbb{N}$, we have $x_q(nm) > \sum_{i>nm} x_q(i)$ and, according to the second sentence of the Kakeya property III, $E(x_q)$ is not a finite union of closed intervals.

Suppose now that $q \geqslant \frac{1}{n+1}$ and consider the sequence

$$y = (1, \dots, 1, q, \dots, q, q^2, \dots, q^2, \dots)$$

with n repetitions of each term. Note that, for any $k \in \mathbb{N}$, the sum

$$\sum_{j>nk} y(j) = q^{k-1} \frac{nq}{1-q}$$

is, by inequality $q \geqslant \frac{1}{n+1}$, bigger than or equal to $y(nk) = q^{k-1}$. Therefore, for any $i \in \mathbb{N}$

$$y(i) \leqslant \sum_{j>i} y(j)$$

and again from the property III, we obtain that E(y) has non-empty interior. To end the proof, we show that

$$n_0 \sum_{j=0}^{\infty} q^j + E(y) \subset E(x_q).$$

If $t \in n_0 \sum_{j=0}^{\infty} q^j + E(y)$, then there exist $p_i \in \{0, 1, \dots, n\}, i = 0, 1, 2, \dots$ such that

$$t = (n_0 + n_0 q + \dots) + (p_0 + p_1 q + \dots).$$

Therefore

$$t = (n_0 + p_0) + (n_0 + p_1)q + \cdots$$

belongs to $E(x_q)$.

Using the latter theorem, we can easily check that sequences a, b and d generate Cantorvals, because they belong to appropriate one-parameter families, indexed by q.

Example (3.1.3)[143]: The Weinstein-Shapiro sequence ([157]).

It is clear that if E(x) is a Cantorval, $\alpha \neq 0$ and $\alpha x = (\alpha x(1), \alpha x(2), \dots)$, then E(αx) is a Cantorval too. To simplify a notation we multiply the sequence a by $\frac{10}{3}$ and consider the family of sequences

$$a_q = (8, 7, 6, 5, 4; q)$$

for $q \in \left(0, \frac{1}{2}\right)$. Summing up 8, 7, 6, 5 and 4, we can get any natural number between $n_0 = 4$ and $n + n_0 = 26$. Therefore, by Theorem (3.1.2), for any q satisfying inequalities

$$\frac{1}{23} \leqslant q < \frac{4}{34},$$

the sequence a_q generates a Cantorval. Obviously, the number $\frac{1}{10}$ used in [157] belongs to $\left[\frac{1}{23}, \frac{4}{34}\right)$. It is not difficult to check (using III) that $a_q \in I$ for $q > \frac{4}{34}$.

Example (3.1.4)[143]: The Ferens sequence ([146]).

For the family of sequences

$$b_q = (7, 6, 5, 4, 3; q)$$

K is equal to 25, $n_0 = 3$ and $n = 19$. Hence, for any $q \in \left[\frac{1}{20}, \frac{3}{28}\right)$, b_q generates a Cantorval. In particular, the sequence $\left(7, 6, 5, 4, 3; \frac{2}{27}\right)$, obtained from the Ferens sequence by multiplication by a constant, generates a Cantorval. Note that $b_q \in I$, for $q \geqslant \frac{3}{28}$.

Example (3.1.5)[143]: The Jones-Velleman sequence ([149]).

Applying Theorem (3.1.2) to the sequence

$$d_q = (3, 2, 2, 2; q)$$

we obtain $K = 9, n_0 = 2$ and $n = 5$, so for any $q \in \left[\frac{1}{6}, \frac{2}{11}\right)$, $E(d_q)$ is a Cantorval. Moreover $d_q \in I$ for $\geqslant \frac{2}{11}$.

We can also consider analogous sequences for more than three 2's. In fact, any sequence

$$x_q = \left(3, \underbrace{2, \dots, 2}_{k-times};\ q\right)$$

with $q \in \left[\frac{1}{2k}, \frac{2}{2k+5}\right)$, generates a Cantorval.

Note that for k=1 and k=2 the argument of Theorem (3.1.2) breaks down, because $\frac{1}{2k} > \frac{2}{2k+5}$. It means, in particular, that Theorem (3.1.2) does not apply to the Guthrie and Nymann example $c = \left(3, 2; \frac{1}{4}\right)$.

However, we can apply Theorem (3.1.2) to "shortly defined" sequences. Indeed, for the sequence $(4, 3, 2;\ q)$, numbers K, n_0 and n are the same as for d_q.

It is not difficult to check that, to keep the interval $\left[\frac{1}{n+1}, \frac{k_m}{K+k_m}\right)$ nonempty, m should be greater than 2.

There is a natural question if Theorem (3.1.2) precisely describes the set of q with $(k_1, \dots, k_m;\ q) \in \mathcal{MC}$. The upper bounds, for all mentioned examples are exact, because $(k_1, \dots, k_m; q) \in I$, for $q > \frac{k_m}{K+k_m}$. However, this is not true for all sequences satisfying the assumptions of Theorem (3.1.2).

Example (3.1.6)[143]: For the sequence $h_q = (10, 9, 8, 7, 6, 5, 2; q)$, we have $K = 47, n_0 = 5$ and $n = 37$. Therefore the interval $\left[\frac{1}{n+1}, \frac{k_m}{K+k_m}\right) = \left[\frac{1}{38}, \frac{2}{49}\right)$ is nonempty.

However, for $h = \left(10, 9, 8, 7, 6, 5, 2; \frac{2}{49}\right)$ and any $n \in \mathbb{N}$, we have $\sum_{i>7n-1} h(i) = \left(\frac{2}{49}\right)^{n-1}\left(2 + \frac{\frac{2}{49}\cdot 47}{1-\frac{2}{49}}\right) = 4\left(\frac{2}{49}\right)^{n-1} < h(7n-1)$. It means that $h \notin I$. Since $\frac{2}{49} > \frac{1}{38}$, we have $h \notin \mathcal{C}$ and so $h \in \mathcal{MC}$.

It is not difficult to check, using III again, that $h_q \notin I$ if and only if $q < \frac{3}{50}$.

Observe, that $E(k_1, \ldots, k_m; q) \subset \sum_{i=1}^{K} C_q$, where $C_q = E\big((1;\ q)\big)$ and $\sum_{i=1}^{K} C_q$ denotes the algebraic sum. In [145] it is proved that, if $q < \frac{1}{K+1}$ then $\sum_{i=1}^{K} C_q$ is homeomorphic to the Cantor set. The following theorem improves this result.

Theorem (3.1.7)[143]: Let $x = (k_1, \ldots, k_m; q)$ be a multigeometric sequence and

$$\Sigma := \left\{\sum_{i=1}^{m} \varepsilon_i k_i : (\varepsilon_i)_{i=1}^{m} \in \{0,1\}^m\right\}.$$

If $q < 1/card\ (\Sigma)$ then E(x) is a Cantor set.

Proof. Clearly, $E(x) = \Sigma + qE\ (x)$. Suppose that $q < 1/card\ (\Sigma)$ and the set E(x) has a nonempty interior. Therefore E(x) has positive Lebesgue measure $\lambda\big(E\ (x)\big)$ and

$$\lambda\big(E(x)\big) \le card\ (\Sigma) \cdot q \cdot \lambda\big(E(x)\big) < \lambda\big(E(x)\big)$$

which gives a contradiction.

Using the latter theorem to the Weinstein-Shapiro sequence $a_q = (8, 7, 6, 5, 4;\ q)$ (compare Example (3.1.3)) we obtain Σ of cardinality 25. It means that $E(a_q) \in \mathcal{C}$ for $q \in \left(0, \frac{1}{25}\right)$. We do not know what is the type of $E(a_q)$ for $q \in \left[\frac{1}{25}, \frac{1}{23}\right]$.

Analogously, $E(b_q) \in C$ for $q \in \left(0, \frac{1}{22}\right)$ (compare Example (3.1.4)), $E(d_q) \in \mathcal{C}$ for $q \in \left(0, \frac{1}{8}\right)$ (compare Example (3.1.5)) and $E(h_q) \in \mathcal{C}$ for $q \in \left(0, \frac{1}{42}\right)$ (compare Example (3.1.6)).

We have just mentioned that Theorem (3.1.2) does not work for sequences $(3{,}2; q)$ and $(3, 2, 2; q)$. However, Guthrie and Nymann have proved that $c = \left(3, 2; \frac{1}{4}\right) \in \mathcal{MC}$. Following their method we will find $q < \frac{1}{n+1}$ such that

$$\Big(3, \underbrace{2, \ldots, 2}_{K-times}; q\Big) \in \mathcal{MC}.$$

Theorem (3.1.8)[143]: For any sequence of the form

$$x_k = \left(3, \underbrace{2, \ldots, 2}_{k-times}; \frac{1}{2k+2}\right),$$

the set $E(x_k)$ is a Cantorval.

Proof. We know that $x_k \notin I$, because $\frac{1}{2k+2} < \frac{2}{2k+5}$ (compare with Example (3.1.5)). It remains to prove that $E(x_k)$ contains an interval.

For a sake of clarity, we will prove a thesis for $k = 2$, i.e. we will show that $E(x_2) \supset [3, 4]$, which means that any point

$$t = 3 + \sum_{i=1}^{\infty} \frac{\varepsilon_i}{6^i}$$

with $\varepsilon_i = \{0, \ldots, 5\}$ belongs to $E(x_2)$.

Since $E(x_2)$ is closed and the set $\left\{3 + \sum_{i=1}^{n} \frac{\varepsilon_i}{6^i} : \varepsilon_i \in \{0, \dots, 5\}, i \leqslant n, n = 0, 1, \dots\right\}$ is dense in [3, 4], it is enough to show that

$$3 + \sum_{i=1}^{n} \frac{\varepsilon_i}{6^i} \in E(x_2)$$

for any $n = 0, 1, \dots, \varepsilon_i = 0, \dots, 5$.
For $n = 0$, we have $3 \in E(x_2)$.
Suppose that any number of the form

$$t' = 3 + \sum_{i=1}^{n-12} \frac{\varepsilon_i}{6^i} \tag{1}$$

belongs to $E(x_2)$. It means that there exist $a_i, b_i, c_i \in \{0, 1\}$ such that

$$t' = 3 + \sum_{i=1}^{n-1} \frac{3a_i + 2b_i + 2c_i}{6^i}.$$

Let

$$t = 3 + \sum_{i=1}^{n} \frac{\varepsilon_i}{6^i}.$$

If $\varepsilon_n = 0, 2, 3, 4$ or 5, then

$$t = t' + \frac{3a_n + 2b_n + 2c_n}{6^n}$$

for some t' and suitable a_n, b_n and c_n.
Suppose that $\varepsilon_n = 1$. Hence

$$t = t' + \frac{1}{6^n} = t' - \frac{1}{6^{n-1}} + \frac{3+2+2}{6^n} = t'' + \frac{3+2+2}{6^n}.$$

If $t' > 3$ then t'' satisfies (1) and the proof is complete. If $t' = 3$ then

$$t = 3 + \frac{1}{6^n} = 2 + \left(1 - \frac{1}{6^{n-1}}\right) + \frac{7}{6^n} = 2 + \left(\frac{5}{6} + \frac{5}{6^2} + \dots + \frac{5}{6^{n-1}}\right) + \frac{3+2+2}{6^n} \in E(x_2).$$

To show that for a fixed $k \geqslant 2$, any $n = 0, 1, \dots$ and $\varepsilon_i = 0, \dots, 2k+1$

$$3 + \sum_{i=1}^{n} \frac{\varepsilon_i}{(2k+2)^i} \in E(x_k)$$

and hence $[3, 4] \in E(x_k)$, one can repeat the previous considerations, using the equality

$$3 + \frac{1}{(2k+2)^n} = 2 + \left(1 - \frac{1}{(2k+2)^{n-1}}\right) + \frac{3+2k}{(2k+2)^n}.$$

Note that, even for special sequences considered, it is very hard to distinguish sequences belonging to $\mathcal{C}$ from sequences belonging to $\mathcal{MC}$. In particular, for any sequence of the form

$$x_q = (3, 2, \dots, 2; q),$$

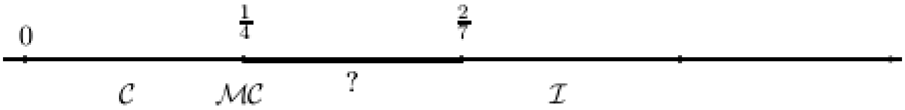
where 2's repeats itself k-times, $x_q \in I$ if and only if $q \geqslant \frac{2}{2k+5}$, and, by Theorem (3.1.7), $x_q \in \mathcal{C}$ for $q < \frac{1}{card(\Sigma)} = \frac{1}{2k+2}$.

We have no idea what are the types of sets $E(x_q)$ for $q \in \left(\frac{1}{2k+2}, \frac{1}{2k}\right)$.

Finally, go back to the Guthrie and Nymann sequence $c = \left(3, 2; \frac{1}{4}\right)$. Z. Nitecki, in [153], proved that for $q < \frac{1}{4}$ the sequence

$$c_q = (3, 2; q)$$

belongs to $\mathcal{C}$. The same conclusion follows easily from Theorem (3.1.7). It is not difficult to check that $x_q \in I$ if and only if $q \geqslant \frac{2}{7}$.



We do not know what is the type of $E(x_q)$ for $q \in \left(\frac{1}{4}, \frac{2}{7}\right)$.
At last, let us consider one more example from [153] (due to Kenyon).

Example (3.1.9)[143]: The achievement set $E(f)$ of the sequence $f = \left(6, 1; \frac{1}{4}\right)$ (in our notation) is $\mathcal{M}$-Cantorval. To prove it, Nitecki observes that 6 is equal to 2 mod 4 and each element of $\mathbb{Z}_4$ can be obtained by summing up the numbers 2 and 1 (compare the proof of Theorem (3.1.8)). Then he makes use of the Baire category theorem. By our mind, this fact can be explained in a much simpler way. Indeed,

$$f = \frac{1}{2}\left(12, 2; \frac{1}{4}\right) = \frac{1}{2}\left(12, 2, 3, 2 \cdot \frac{1}{4}, 3 \cdot \frac{1}{4}, 2 \cdot \frac{1}{16}, \dots\right).$$

Hence

$$E(f) = \frac{1}{2}E\left(12, 3, 2, 3 \cdot \frac{1}{4}, 2 \cdot \frac{1}{4}, \dots\right) = \frac{1}{2}E(c) \cup \frac{1}{2}(E(c) + 12)$$

and E(f) is of the same form as E(c). In general, it is easy to observe (in the same way as above) that the sequences $(k_1, k_2, \dots, k_m; q)$ and $(q^{n_1}k_1, q^{n_2}k_2, \dots, q^{n_m}k_m; q)$ for integers $n_1, n_2, \dots, n_m$ are in the same set among of $\mathcal{C}, I$ or $\mathcal{MC}$. Observe, for instance, that $\left(2, 1; \frac{1}{4}\right) \in I$ and $\left(3, 8; \frac{1}{4}\right) \in \mathcal{MC}$. However, each element of $\mathbb{Z}_4$ can be obtained by summing up 2 and 1, but 2 can not be obtained by summing up 3 and 8.

Section (3.2): Some Self-Similar Sets

Suppose that $x = (x_n)_{n=1}^{\infty}$ is an absolutely summable sequence with infinitely many nonzero terms and let

$$E(x) = \left\{\sum_{n=1}^{\infty} \varepsilon_n x_n : (\varepsilon_n)_{n=1}^{\infty} \in \{0, 1\}^{\mathbb{N}}\right\}$$

denote the set of all subsums of the series $\sum_{n=1}^{\infty} \varepsilon_n x_n$, called the achievement set (or a partial sumset) of x. The investigation of topological properties of achievement sets was initiated almost one hundred years ago. In 1914 Soichi Kakeya [150] presented the following result:

Theorem (3.2.1)[158]: (Kakeya). For any sequence $x \in l_1 \setminus c_{00}$

(i) $E(x)$ is a perfect compact set.

(ii) If $|x_n| > \sum_{i>n} |x_i|$ for almost all n, then $E(x)$ is homeomorphic to the ternary Cantor set.

(iii) If $|x_n| \leq \sum_{i>n} |x_i|$ for almost all n, then $E(x)$ is a finite union of closed intervals. In the case of non-increasing sequence x, the last inequality is also necessary for $E(x)$ to be a finite union of intervals.

Moreover, Kakeya conjectured that $E(x)$ is either nowhere dense or a finite union of intervals. Probably, the first counterexample to this conjecture was given by Weinstein and Shapiro ([157]) and, independently, by Ferens ([146]). The simplest example was presented by Guthrie and Nymann [147]: for the sequence $c = \left(\frac{5+(-1)^n}{4^n}\right)_{n=1}^{\infty}$, the set $T = E(c)$ contains an interval but is not a finite union of intervals. They formulated the following theorem, finally proved in [155]:

Theorem (3.2.2)[158]: For any sequence $x \in l_1 \setminus c_{00}, E(x)$ is one of the following sets:

(i) a finite union of closed intervals;

(ii) homeomorphic to the Cantor set;

(iii) homeomorphic to the set T .

Note, that the set $T = E(c)$ is homeomorphic to $C \cup \bigcup_{n=1}^{\infty} S_{2n-1}$, where S_n denotes the union of the 2^{n-1} open middle thirds which are remved from [0, 1] at the n-th step in the construction of the Cantor ternary set C. Such sets are called Cantorvals (to emphasize their similarity to unions of intervals and to the Cantor set simultaneously). Formally, a Cantorval (an M-Cantorval, see [152]) is a non-empty compact subset S of the real line such that S is the closure of its interior, and both endpoints of any non-degenerated component are accumulation points of one-point components of S. A non-empty subset C of the real line R will be called a Cantor set if it is compact, zero-dimensional, and has no isolated points.

We observe that Theorem (3.2.2) says, that l_1 can be devided into 4 sets: c_{00} and the sets connected with cases (i), (ii) and (iii). Some algebraic and topological properties of these sets have been recently considered in [144].

We will describe sequences constructed by Weinstein and Shapiro, Ferens and Guthrie and Nymann using the notion of multigeometric sequence. We call a sequence multigeometric if it is of the form

$$(k_0, k_1, \dots, k_m, k_{0q}, k_{1q}, \dots, k_{mq}, k_{0q^2}, k_{1q^2}, \dots, k_{mq^2}, k_{0q^3} \dots)$$

for some positive numbers $k_0, \dots, k_m$ and $q \in (0,1)$. We will denote such a sequence by $(k_0, k_1, \dots, k_m; q)$. Keeping in mind that the type of $E(x)$ is the same as $E(\alpha x)$, for any $\alpha > 0$, we can describe the Weinstein-Shapiro sequence as

$$a = \left(8, 7, 6, 5, 4; \frac{1}{10}\right),$$

the Ferens sequence as $b = \left(7, 6, 5, 4, 3; \frac{2}{27}\right)$ and the Guthrie-Nymann sequence as $c = \left(3, 2; \frac{1}{4}\right)$.

Another interesting example of a sequence d with $E(d)$ being Cantorval was presented by R. Jones in ([149]). The sequence is of the form

$$d = \left(3, 2, 2, 2; \frac{19}{109}\right).$$

In fact, Jones constructed continuum many sequences generating Cantorvals, indexed by a parameter q, by proving that, for any positive number q with

$$\frac{1}{5} \le \sum_{n=1}^{\infty} q^n < \frac{2}{9}$$

$\left(i.e. \frac{1}{6} \le q < \frac{2}{11}\right)$ the achievement set of the sequence

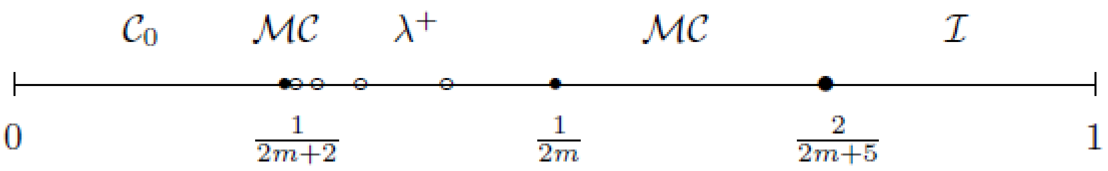
$$(3, 2, 2, 2; q)$$

is a Cantorval.

The structure of the achievement sets $E(x)$ for multigeometric sequences x was studied in [143], which contains a necessary condition for the achivement set $E(x)$ to be an interval and sufficient conditions for $E(x)$ to contain an interval or have Lebesgue measure zero. In the case of a Guthrie-Nymann-Jones sequence

$$x_q = (3, 2, \ldots, 2; q),$$

of rank m (i.e., with m repeated $2's$), the set $E(x_q)$ is an interval if and only if $q > 2\ 2m + 5$, $E(xq)$ is a Cantor set of measure zero if $q < \frac{1}{2m+2}$, and $E(x_q)$ is a Cantorval if $q \in \left\{\frac{1}{2m+2}\right\} \cup \left[\frac{1}{2m}, \frac{1}{2m+5}\right]$. We reveal some structural properties of the sets E(xq) for q belonging to the "misterious" interval $\left(\frac{1}{2m+2}, \frac{1}{2m}\right)$. In particular, we shall show that for almost all q in this interval the set $E(x_q)$ has positive Lebesgue measure and there is a decreasing sequence (q_n) convergent to $\frac{1}{2m+2}$ for which $E\left(x_{q_n}\right)$ is a Cantor set of zero Lebesgue measure. The above description of the structure of $E(x_q)$ can be presented as follows:



where C_0 (resp. $\mathcal{M}$C$, I$) indicates sets of numbers q for which the set $E(x_q)$ is a Cantor set of zero Lebesgue measure (resp. a Cantorval, an interval). The symbol _ + indicates that for almost all q in a given interval the sets $E(x_q)$ have positive Lebesgue measure, which means that the set $Z = \left\{q \in \frac{1}{2m+2}, \frac{1}{2m} : \lambda(E(x_q)\ ) = 0\right\}$ has Lebesgue measure $\lambda(Z) = 0$. Similar diagrams we use later.

The achievement sets of multigeometric sequences are partial cases of self-similar sets of the form

$$K(\Sigma, q) = \left\{\sum_{n=0}^{\infty} a_n q_n : (a_n)_{n=0}^{\infty} \in \Sigma^{\omega}\right\}$$

where $\Sigma \subset \mathbb{R}$ is a set of real numbers and $q \in (0, 1)$. The set $K(\Sigma, q)$ is self-similar in the sense that $K(\Sigma, q) = \Sigma + q.K(\Sigma, q)$. Moreover, the set $K(\Sigma, q)$ can be found as a unique compact solution $K \subset \mathbb{R}$ of the equation $K = \Sigma + qK$.

It follows that for a multigeometric sequence $x_q = x_q(k_0, \ldots, k_m, q)$ the achievement set $E(x)$ coincides with the self-similar set $K(\Sigma, q)$ for the set

$$\Sigma = \left\{\sum_{n=0}^{m} k_n\, \mathcal{E}_n : (\mathcal{E}_n)_{n=0}^{m} \in \{0, 1\}^{m+1}\right\}$$

of all possible sums of the numbers $k_0, \ldots, k_m$. This makes possible to apply for studying the achievement sets $E(x_q)$ the theory of self-similar sets developed in [83], [128] and, first of all, in [138].

We shall describe some topological and measure properties of the self-similar sets $K(\Sigma, q)$ depending on the value of the similarity ratio $q \in (0, 1)$, and shall apply the obtained result to establishing topological and measure properties of achievement sets of multigeometric progressions. To formulate the principal results we need to introduce some number characteristics of compact subsets $A \subset \mathbb{R}$.

Given a compact subset $A \subset \mathbb{R}$ containing more than one point let

$$\operatorname{diam} A = \sup\{|a - b| : a, b \in A\}$$
be the diameter of A and
$\delta(A) = \inf\{|a - b| : a, b \in A, a \neq b\}$ and $\Delta(A) = \sup\{|a - b| : a, b \in A, (a, b) \cap A = \emptyset\}$ be the smallest and largest gaps in A, respectively. Observe that A is an interval (equal to [min$A$,max$A$]) if and only if $\Delta(A) = 0$.
Also put
$$I(A) = \frac{\Delta(A)}{(A) + \operatorname{diam} A} \text{ and } i(A) = \inf\{I(B) : B \subset A, 2 \leq |B| < \omega\}.$$
In particular, given a finite subset $\Sigma \subset \mathbb{R}$ of cardinality $|\Sigma| \geq 2$, we will write it as $\Sigma = \{\sigma_1, \dots, \sigma_s\}$ for real numbers $\sigma_1 < \cdots < \sigma_s$. Then we have
$$\operatorname{diam}(\Sigma) = \sigma_s - \sigma_1, \delta(\Sigma) = \min_{i<s}(\sigma_{i+1} - \sigma_i), \text{and } \Delta(\Sigma) = \max_{i<s}(\sigma_{i+1} - \sigma_i).$$
Theorem (3.2.3)[158]: Let $\Sigma = \{\sigma_1, \dots, \sigma_s\}$ for some real numbers $\sigma_1 < \cdots < \sigma_s$ The self-similar sets $K(\Sigma, q)$ where $q \in (0, 1)$ have the following properties:

(i) $K(\Sigma; q)$ is an interval if and only if $q \geq I(\Sigma)$;

(ii) $K(\Sigma; q)$ is not a finite union of intervals if $q < I(\Sigma)$ and $\Sigma(\Delta) \in \{\sigma_2 - \sigma_1, \sigma_s - \sigma_{s-1}\}$,

(iii) $K(\Sigma; q)$ contains an interval if $q \geq i(\Sigma)$;

(iv) If $d = \frac{\delta(\Sigma)}{diam(\Sigma)} < \frac{1}{3+2\sqrt{2}}$ and $\frac{1}{|\Sigma|} < \frac{\sqrt{d}}{1+\sqrt{d}}$, then for almost all $q \in \left(\frac{1}{|\Sigma|}, \frac{\sqrt{d}}{1+\sqrt{d}}\right)$ the set $K(\Sigma; q)$ has positive Lebesgue measure and the set $K(\Sigma; \sqrt{q})$ contains an interval;

(v) $K(\Sigma; q)$ is a Cantor set of zero Lebesgue measure if $q < \frac{1}{|\Sigma|}$ or, more generally, if $q^n < \frac{1}{|\Sigma_n|}$ for some $n \in \mathbb{N}$ where $\Sigma_n = \{\sum_{k=0}^{n-1} a_k q^k : (a_k)_{k=0}^{n-1} \in \Sigma_n\}$.

(vi) If $\Sigma \supset \{a, a+1, b+1, c+1, b+|\Sigma|, c+|\Sigma|\}$ for some real numbers $a, b, c \in \mathbb{R}$ with $b \neq c$, then there is a strictly decreasing sequence $(q_n)_{n\in\omega}$ with $\lim_{n\to\infty} q_n = \frac{1}{|\Sigma|}$ such that the sets $K(\Sigma; q_n)$ has Lebesgue mesure zero.

The statements (i)–(iii) from this theorem will be proved, the statement (iv) and (v),(vi). Writing that for almost all q in an interval (a, b) some property $P(q)$ holds we have in mind that the set $Z = \{q \in (a, b): P(q) \text{ does not hold}\}$ has Lebesgue measure $\lambda(Z) = 0$.

We generalize results of [143] detecting the self-similar sets $K(\Sigma; q)$ which are intervals or Cantorvals. In the following theorem we prove the statements (i)–(iii) of Theorem (3.2.3).

Theorem (3.2.4)[158]: Let $q \in (0, 1)$ and $\Sigma = \{\sigma_1, \dots, \sigma_s\} \subset \mathbb{R}$ be a finite set with $\sigma_1 < \cdots < \sigma_s$. The self-similar set $K(\Sigma; q) = \left\{\sum_{i=0}^{\infty} a_i q^i : (a_i)_{i\in\omega} \in \Sigma^\omega\right\}$

(i) is an interval if and only if $q \geq I(\Sigma)$;

(ii) contains an interval if $q \geq i(\Sigma)$;

(iii) is not a finite union of intervals if $q < I(\Sigma)$ and $\Delta(\Sigma) \in \{\sigma_2 - \sigma_1, \sigma_s - \sigma_{s-1}\}$.

Proof. (i) Observe that $\operatorname{diam} K(\Sigma; q) = diam(\Sigma)/(1 - q)$. Assuming that $\geq I(\Sigma) = \Delta(\Sigma)/(\Delta(\Sigma) + diam\Sigma)$, we conclude that $\Delta(\Sigma) \leq q \cdot diam(\Sigma)/(1 - q) = q \cdot diamK(\Sigma; q)$, which implies that
$$\Delta(K(\Sigma; q)) = \Delta(\Sigma + q \cdot K(\Sigma; q)) \leq \Delta(q \cdot K(\Sigma; q)) = q \cdot K(\Sigma; q).$$
Since $q < 1$ this inequality is possible only in case $\Delta(K(\Sigma; q)) = 0$, which means that $K(\Sigma; q)$ is an interval.
If $q < \Delta(\Sigma)/(\Delta(\Sigma) + diam\Sigma)$, then $\Delta(\Sigma) > q \cdot diam(\Sigma)/(1 - q) = q \cdot diam(K(\Sigma; q))$ and we can find two consequtive points $a < b$ in Σ with $b = a + \Delta(\Sigma) > a +$

$diam(qK(\Sigma;q))$ and conclude that $[a,b] \cap K(\Sigma;q) = [a,b] \cap (+qK(\Sigma;q)) \subset [a, a + diam(qK(\Sigma;q))] \neq [a,b], so\ K(\Sigma;q)$ is not an interval.
(ii) Now assume that $q \geq i(\Sigma)$ and find a subset $B \subset \Sigma$ such that $I(B) = i(\Sigma) < q$. By the preceding item, the self-similar set $K(B;q) = B + qK(B;q)$ is an interval. Consequently, $K(\Sigma;q)$ contains the interval $K(B;q)$.
(iii) Finally assume that $\Delta(\Sigma) = \sigma_2 - \sigma_1$and $q < I(\Sigma)$. Since for every $a \in \Sigma$ we get $K(\Sigma - a;\ q) = K(\Sigma;q) - \frac{a}{1-q}$,we can replace Σ by its shift and assume that $\sigma_1 = 0$ and hence $\Delta(\Sigma) = \sigma_2 - \sigma_1 = \sigma_2$. It follows from $q < I(\Sigma) = \sigma_2/(\sigma_2 + diam\Sigma)$ that for any $\in \mathbb{N}$, the interval $\sum_{n=j+1}^{\infty} q^n \sigma_s, q^j \sigma_2$ is nonempty and disjoint from $K(\Sigma;q)$. Hence, no interval of the form $[0,\varepsilon]$ is included in $K(\Sigma;q)$. But $0 \in K(\Sigma;q)$, so $K(\Sigma;q)$ is not a finite union of closed intervals. By analogy we can consider the case $\Delta(\Sigma) = \sigma_s \sigma_{s-1}$.
In particular, Theorem (3.2.4) implies:
Corollary (3.2.5)[158]: For$\Sigma = \{0,1,2,\ldots,s - 1\}$ the set $K(\Sigma;q)$ is an interval if and only if $q \geq I(\Sigma) = \frac{1}{|\Sigma|}$.
Corollary (3.2.6)[158]: If $\{k, k+1, \ldots, k+n-1\} \subset \Sigma$, then $i(\Sigma) \leq \frac{1}{n}$ and for every $q \geq \frac{1}{n}$ the set $K(\Sigma;q)$contains an interval.
In particular, for the Guthrie-Nymann-Jones multigeometric sequence $x_q = (3,2,\ldots,2;q)$ of rank m the sumset $\Sigma = \{0,2,\ldots,2m+1,2m+3\}$ has cardinality $|\Sigma| = 2m+2, I(\Sigma) = \frac{\Delta(\Sigma)}{\Delta(\Sigma)+diam\Sigma} = \frac{2}{2m+5}$, $i(\Sigma) = min\left\{\frac{1}{2m}, \frac{2}{2m+5}\right\}$,and $d = \frac{\delta(\Sigma)}{diam(\Sigma)} = \frac{1}{2m+3}$. So, for $q \in \left[\frac{2}{2m+5}, 1\right]$ the set $E(x_q) = K(\Sigma;q)$ is an interval and for $q \in \left[\frac{1}{2m}, \frac{2}{2m+5}\right]$ aCantorval.

We shall prove the statement (iv) of Theorem (3.2.3) detecting numbers q for which the self-similar set $K(\Sigma;q)$ has positive Lebesgue measure $\Delta(K(\Sigma;q))$. For this we shall apply the deep results of Boris Solomyak [162] related to the distribution of the random series $\sum_{n=0}^{\infty} a_n \lambda^n$, where the coefficients an $a_n \in \Sigma$ are chosen independently with probability $\frac{1}{|\Sigma|}$each.
Given a finite subset $\Sigma \subset \mathbb{R}$ consider the number $\alpha(\Sigma) = inf\{x \in (0,1): \exists (a_n)_{n\in\omega} \in (\Sigma - \Sigma)^{\omega} \backslash \{0\}^{\omega}\}$! such that $\sum_{n=0}^{\infty} a_n x^n = 0$ and$\sum_{n=1}^{\infty} n a_n x^{n-1} = 0$.
The first part of the following theorem was proved by Solomyak in [162]:
Theorem (3.2.7)[158]: Let $\Sigma \subset \mathbb{R}$ be a finite subset. If $\frac{1}{|\Sigma|} < \alpha(\Sigma)$, then for almost all q in the interval $\left(\frac{1}{|\Sigma|}, \alpha(\Sigma)\right)$ the self-similar set $K(\Sigma;q)$ has positive Lebesgue measure and the set $K(\Sigma;\sqrt{q})$ contains an interval.
Proof. By [162], for almost all $q \in \left(\frac{1}{|\Sigma|}, \alpha(\Sigma)\right)$ the self-similar set $K(\Sigma;q)$has positive Lebesgue measure. Since $K(\Sigma;\sqrt{q}) = K(\Sigma;q) + \sqrt{q} \cdot K(\Sigma;q)$, the set $K(\Sigma;q)$ contains an interval, being the sum of two sets of positive Lebesque measure (according to the famous Steinhaus Theorem [163]).
The definition of Solomyak's constant $\alpha(\Sigma)$ does not suggest any efficient way of its calculation. In [162] Solomyak found an efficient lower bound on $\alpha(\Sigma)$ based on the notion of a $(*)$-function, i.e., a function of the form

$$g(x) = -\sum_{k=1}^{n-1} x^k + \gamma x^n + \sum_{k=n+1}^{\infty} x^k$$

for some $n \in \mathbb{N}$ and $\gamma \in [-1, 1]$. In Lemma (3.2.7) [162] Solomyak proved that every $(*)$-function $g(x)$ has a unique critical point on $[0, 1)$ at which g takes its minimal value. Moreover, for every $d > 0$ there is a unique $(*)$-function $g_d(x)$ such that $\min_{[0,1)} g_d = -d$. The unique critical point $x_d \in {g_d}^{-1}(-d) \in [0, 1)$ of g_d will be denoted by $\underline{\alpha}(d)$.
The following lower bound on the number $\alpha(\Sigma)$ follows from Proposition (3.2.8) and inequality (15) in [162].
Lemma (3.2.8)[158]: For every finite set $\Sigma \subset \mathbb{R}$ of cardinality $|\Sigma| \geq 2$ we get

$$\alpha(\Sigma) \geq \underline{\alpha}(d) \quad \text{where} \quad d = \frac{\delta(\Sigma)}{diam(\Sigma)} .$$

The function $\underline{\alpha}(d)$ can be calculated effectively $\left(\text{at least for } d \leq \frac{1}{2}\right)$.
Lemma (3.2.9)[158]: If $0 < d \leq \frac{1}{3+2\sqrt{2}}$, then

$$\underline{\alpha}(d) = \frac{\sqrt{d}}{1 + \sqrt{d}}.$$

Proof. Observe that the minimal value of the $(*)$-function $g(x) = -x + \sum_{k=2}^{\infty} x^k = -x + \frac{x^2}{1-x}$ is equal to $-\frac{1}{3+2\sqrt{2}}$, which implies that for $d \in \left(0, \frac{1}{3+2\sqrt{2}}\right)$ the number $\underline{\alpha}(d)$ is equal to the critical point of the unique $(*)$-function $g(x) = \gamma x + \sum_{k=2}^{\infty} x^k = -1 + (\gamma - 1)x + \frac{1}{1-x}$ with $\min_{[0,1)} g = -d$. This $(*)$-function has derivative $g'(x) = (\gamma - 1) + \frac{1}{(1-x)^2}$. If x is the critical point of g, then $1 - = \frac{1}{(1-x)^2}$ and the equality

$$d = -1 + (\gamma - 1)x + \frac{1}{1 - x} = -1 - \frac{x}{(1 - x)^2} + \frac{1}{1 - x}$$

has the solution

$$x = 1 - \frac{1}{1 + \sqrt{d}} = \frac{\sqrt{d}}{1 + \sqrt{d}}$$

which is equal to $\underline{\alpha}(d)$
For $d > \frac{1}{3+2\sqrt{2}}$ the formula for $\underline{\alpha}(d)$is more complex.
Lemma (3.2.10)[158]: If $\frac{1}{3+2\sqrt{2}} \leq d \leq \frac{1}{2}$, then the value

$$\underline{\alpha}(d) = \frac{1 + d}{3} + \frac{\sqrt{3}2 \cdot R}{6} + \frac{2d^2 - 8d - 1}{3\sqrt{3}2 \cdot R}$$

where

$$R = \sqrt[3]{4d^3 - 24d^2 + 21d - 5 + 3\sqrt{3}\sqrt{1 - 8d^3 + 39d^2 - 6d}}$$

can be found as the unique real solution of the qubic equation

$$2(x - 1)^3 + (4 - 2d)(x - 1)^2 + 3(x - 1) + 1 = 0.$$

Proof. Since the minimal values of the $(*)$-functions $g_1(x) = -x + \sum_{k=2}^{\infty} x^k$ and $g(x) = -x - x^2 + \sum_{k=3}^{\infty} x^k$ are equal to $\frac{1}{3+2\sqrt{2}}$and $-\frac{1}{2}$, respectively, for $d \in \left[\frac{1}{3+2\sqrt{2}}, \frac{1}{2}\right]$ the number $\underline{\alpha}(d)$ is equal to the critical point of aunique $(*)$-function

$$g(x) = -x + \gamma x^2 + \sum_{k=3}^{\infty} x^k = -1 - 2x + (\gamma - 1)x^2 + \frac{1}{1 - x}$$

with $\min_{[0,1)} g = -d$. At the critical point x the derivative of g equals zero:

$$0 = g'(x) = -2 + 2(\gamma - 1)x + \frac{1}{(1 - x)^2}$$

which implies that

$$\gamma - 1 = \frac{1}{2x}\left(2 - \frac{1}{(1 - x)^2}\right) = \frac{2x^2 - 4x + 1}{2x(1 - x)^2}.$$

After substitution of $\gamma - 1$to the formula of the function $g(x)$, we get

$$-d = -1 - 2x - \frac{2x^3 - 4x^2 + x}{2(1 - x)^2} + \frac{1}{1 - x}.$$

This equation is equivalent to the qubic equation

$$2(x - 1)^3 + (4 - 2d)(x - 1)^2 + 3(x - 1) + 1 = 0.$$

Solving this equation with the Cardano formulas we can get the solution$\underline{\alpha}(d)$ written in the lemma.

Theorem (3.2.7) and Lemma (3.2.9) imply:

Corollary (3.2.11)[158]: Let $\Sigma \subset \mathbb{R}$ be a finite subset containing more than three points and $d = \delta(\Sigma)/diam(\Sigma)$. If $d \le \frac{1}{3+2\sqrt{2}}and\frac{\sqrt{d}}{1+\sqrt{d}} > \frac{1}{|\Sigma|}$, then for almost all q in the interval □ $\left(\frac{1}{|\Sigma|}, \frac{\sqrt{d}}{1+\sqrt{d}}\right)$ the self-similar set $K(\Sigma; q)$ has positive Lebesgue measure and the set $K(\Sigma; \sqrt{q})$ contains an interval.

Corollary (3.2.11) guarantees that for almost all $q \in \left(\frac{1}{\sqrt{2}^k}, \frac{4\sqrt{d}}{1+\sqrt{d}}\right)$the set $K(\Sigma; q)$contains an interval.

Multigeometric sequences of the form

$$(k + m, \ldots, k + 1, k; q)$$

with $m \ge k$ we will call, after [159], Ferens-like sequences. The achievement set $E(x)$ for a Ferens-like sequence coincides with the self-similar set $K(\Sigma; q)$ for the set

$$\Sigma = \{0, k, k + 1, \ldots, n - k, n\}.$$

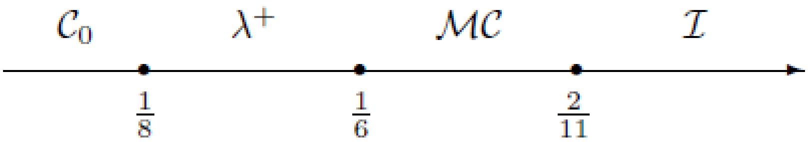
where $n = (m + 1)(2k + m)/2$. Sets $K(\Sigma; q)$ withΣ of this form will be called Ferens-like fractals.

Note that Guthrie-Nymann-Jones sequence of rank mgenerates a Ferens-like fractal (with $\Sigma = \{0, 2, 3, \ldots, 2m + 1, 2m + 3\}$. There are also Ferens-like fractals which are not originated by any multigeometric sequence (for exam-ple $K(\Sigma; q)$ with $\Sigma = \{0, 4, 5, 6, 7, 11\}$). However, as an easy consequence of the main theorem of [161], we obtain for Ferens-like fractals “trichotomy” analogous to that formulated in Theorem (3.2.2). Moreover, some theorems formulated for multigeometric sequences are in fact proved for $K(\Sigma; q)$ (see for example Theorem 2 in [143]).

Example (3.2.12)[158]: For the Ferens-like sequence $x_q = (4, 3, 2;\ q)$ we get$\Sigma = \{0, 2, 3, 4, 5, 6, 7, 9\}$,

$$d = \frac{\delta(\Sigma)}{\text{diam}(\Sigma)} = \frac{1}{9} < \frac{1}{3 + 2\sqrt{2}} \quad \text{and} \frac{\sqrt{d}}{1 + \sqrt{d}} = \frac{1}{4} > \frac{1}{6} = i(\Sigma).$$

By Corollary (3.2.11) (and Theorem (3.2.4)), for almost all numbers $q \in \left(\frac{1}{8}, 1\right)$ the achievement set $E(x_q) = K(\Sigma; q)$ haspositive Lebesgue measure (for $q < \frac{2}{11} = I(\Sigma)$ it is not a finite union of intervals). By Theorem (3.2.4), for any$q \in [i(\Sigma), I(\Sigma)) = [\frac{1}{6}; \frac{2}{11})$ the set $K(\_; \mathrm{q})$ is a Cantorval. The structure of the sets $E(x_q) = K(\Sigma; q)$ is described in the diagram:



For any Ferens-like fractal, $|\Sigma| = n - 2k + 3, \Delta(\Sigma) = k, \delta(\Sigma) = 1, I(\Sigma) = \frac{k}{n+k}$, $i(\Sigma) = \min \frac{1}{|\Sigma|-2}$, $I(\Sigma)$ and $d = \frac{1}{n}$. Moreover, if $n \geq 7$ then $\underline{\alpha}(d) = \frac{1}{\sqrt{n}+1}$. Therefore, one can check that for any Ferens-like sequence we have $\underline{\alpha}(d) > i(\Sigma)$, and we can draw an analogous diagram. The same result we can obtain for any Ferens-like fractal with $k = 2$ (even if it is not originated by any Ferens-like sequence). However, there are Ferens-like fractals with $\underline{\alpha}(d) < i(\Sigma)$ (for example $K(\Sigma; q)$ with $\Sigma = \{0, 3, 4, 7\}$ or $\Sigma = \{0, 4, 5, 6, 7, 11\}$).

Example (3.2.13)[158]: For the Guthrie-Nymann-Jones sequence $x_q = (3, 2, \ldots, 2; q)$ of rank $m \geq 2$ we get

$$\Sigma = \{0, 2, 3, \ldots, 2m+1, 2m+3\}, |\Sigma| = 2m+2, I(\Sigma) = \frac{2}{2m+5}, i(\Sigma)$$
$$= \min\left\{\frac{1}{2m}, \frac{2}{2m+5}\right\}, d = \frac{1}{2m+3} \text{ and } \underline{\alpha}(d) = 1/(1 + \sqrt{2m+3}).$$

Moreover, we have $d < \frac{1}{3+2\sqrt{2}}$and $\underline{\alpha}(d) \geq i(\Sigma) > \frac{1}{2m+2} = \frac{1}{|\Sigma|}$. So, we can apply Corollary (3.2.11) and conclude that for almost all numbers $q \in \left(\frac{1}{2m+2}, \frac{1}{2m}\right)$the self-similar set $K(\Sigma; q)$ has positive measure. By Theorem (3.2.4), for any $q \in \left[i(\Sigma), \frac{2}{2m+5}\right)$ the set $K(\Sigma; q)$ is a Cantorval and for all $q \in \left[\frac{2}{2m+5}, 1\right)$ it is an interval.

For $m = 1$ we obtain $\underline{\alpha}(d) = \underline{\alpha}\left(\frac{1}{5}\right) > \frac{2}{7}$. Therefore, for almost all numbers $q \in \left(\frac{1}{4}, \frac{2}{7}\right)$ the set $K(\Sigma; q)$has positive Lebesgue measure.

The results of the preceding yields conditions under which for almost all q in an interval $\left(\frac{1}{|\Sigma|}, \alpha(\Sigma)\right)$ the set $K(\Sigma; q)$ has positive Lebesgue measure. We shall show that this interval can contain infinitely many numbers q with $\lambda(K(\Sigma; q)) = 0$ thus proving the statements (v) and (vi) of Theorem (3.2.3).

Theorem (3.2.14)[158]: If there exists $n \in \mathbb{N}$ such that

$$\left|\sum_{i=0}^{n-1} q^i \Sigma\right| \cdot q^n < 1$$

then the set $K(\Sigma; q)$ has measure zero.

Proof. Denote $K := K(\Sigma, q)$. From the equality $K = \Sigma + qK$ we obtain, by induction, that

$$K = \sum_{i=0}^{n-1} q^i \Sigma + q^n K.$$

Let$\Sigma_n = \sum_{i=0}^{n-1} q^i \Sigma$. If $|\Sigma_n| \cdot q^n < 1$, then

$$\lambda(K) \leq |\Sigma_n| \cdot q^n \cdot \lambda(K) < 1 \cdot \lambda(K)$$

which is possible only if $\lambda(K) = 0$.

To use the latter theorem we need a technical lemma:

Lemma (3.2.15)[158]: For any integer numbers $s > 1$ and $n > 1$ the unique positive solution q of the equation

$$x + x^2 + \cdots + x^{n-1} = \frac{1}{s-1} \tag{2}$$

is greater than $\frac{1}{s}$. Moreover, there is $n_0 \in \mathbb{N}$ such that for any $n > n_0$

$$(s^n - 2^{n-1}) \cdot q^n < 1. \tag{3}$$

Proof. Clearly

$$\sum_{i=1}^{n-1} \left(\frac{1}{s}\right)^i = \frac{1}{s-1} \cdot \left(1 - \frac{1}{s^{n-1}}\right) < \frac{1}{s-1},$$

so $q > \frac{1}{s}$. From the equality

$$\frac{1}{s-1} = \sum_{i=1}^{n-2} \left(\frac{1}{s}\right)^i + \frac{1}{(s-1)s^{n-2}}$$

we obtain

$$q^{n-1} = \frac{1}{s-1} - \sum_{i=1}^{n-2} q^i < \frac{1}{s-1} - \sum_{i=1}^{n-2} \left(\frac{1}{s}\right)^i = \frac{1}{(s-1)s^{n-2}}.$$

Using the latter inequality and the equality

$$\frac{1}{s-1} = \frac{q - q^n}{1 - q}$$

we have

$$\frac{1-q}{s-1} = q(1 - q^{n-1}) > q\left(1 - \frac{1}{(s-1)\, s^{n-2}}\right).$$

Therefore,

$$1 - q > (s-1)q - \frac{q}{s^{n-2}}$$

(which means that $sq - \frac{q}{s^{n-2}} < 1$) and finally

$$q < \frac{1}{s\left(1 - \frac{1}{s^{n-1}}\right)}. \tag{4}$$

From Bernoulli's inequality it follows that

$$\left(1 - \frac{1}{s^{n-1}}\right)^n \geq 1 - \frac{n}{s^{n-1}}$$

and, by (4), we have

$$q^n < \frac{1}{s^n . \left(1 - \frac{n}{s^{n-1}}\right)}.$$

Consequently,

$$(s^n - 2^{n-1}) \cdot q^n < \frac{s^n . \left(\frac{1 - 2^{n-1}}{s^n}\right)}{s^n \cdot \left(1 - \frac{n}{s^{n-1}}\right)}$$

Obviously, for n greater then some n_0

$$\frac{2^{n-1}}{s} > n$$

And hence

$$\frac{2^{n-1}}{s^n} > \frac{n}{s^{n-1}}$$

Which proves (3).

Theorem (3.2.16)[158]: If a finite subset $\Sigma \subset \mathbb{R}$ contains the set $\{a, a+1, b+1, c+1, b+|\Sigma|, c+|\Sigma|\}$ for some real numbers a, b, c with $b \neq c$, then there is a decreasing sequence $(q_n)_{n=1}^{\infty}$ tending to $\frac{1}{|\Sigma|}$such that, for any $n \in \mathbb{N}$, the self-similar set $K(\Sigma, q_n)$ has Lebesgue measure zero.

Proof. Let $s = |\Sigma|$ and for every n denote by q_n the unique positive solution of the equation (2) from Lemma (3.2.15). Let n_0 be a natural number such that

$$(s^n - 2^{n-1}) \cdot (q_n)^n < 1$$

for any $n > n_0$. Clearly $(q_n)_{n=n_0}^{\infty}$ is a decreasing sequence and $\lim_{n\to\infty} q_n = \frac{1}{s}$. It suffices to show that $K(\Sigma, q)$ has measure zero for $n > n_0$.

Taking into account that each q_n is a solution of (2), we conclude that

$$a + \sum_{i=1}^{n-1} (s - 1 + \varepsilon_i)\,(q_n)^i = (a+1) + \sum_{i=1}^{n-1} \varepsilon_i\,(q_n)^i$$

for any $\varepsilon_i \in \{b+1, c+1\} \subset \Sigma$. Therefore

$$\left|\sum_{i=1}^{n-1} (q_n)^i\,\Sigma\right| \leq s^n - 2^{n-1}.$$

Hence, by Lemma (3.2.15),

$$\left|\sum_{i=1}^{n-1} (q_n)^i\,\Sigma\right| \leq (q_n)^n < 1.$$

and we can apply Theorem (3.2.14) to conclude that $K(\Sigma, q)$ has Lebesgue measure zero.

The condition

$$\{a, a+1, b+1, c+1, b+|\Sigma|, c+|\Sigma|\} \subset \Sigma \tag{5}$$

looks a bit artificial but it can be easily verified for many sumsets Σ of multigeometric sequences.

In particular, for the Guthrie-Nymann-Jones sequence of rank $m \geq 1$

$$x_q = (3, 2, \ldots, 2; q),$$

the sumset $\Sigma = \{0, 2, 3, \ldots, 2m+1, 2m+3\}$ has cardinality $|\Sigma| = 2m+2$. Observe that for the set Σ the condition (5) holds for $a = 2, b = 1$ and $c = -1$. Because of that Theorem (3.2.16) yields a sequence $(q_n)_{n=1}^{\infty} \searrow \frac{1}{2m+2}$such that for every $n \in \mathbb{N}$ the self-similar set $E(x_{q_n})$ is a Cantor sets of zero Lebesgue measure.

By [143], for $q = \frac{1}{2m+2}$ the achievement set $E(x_q)$ is a Cantorval. Therefore, if $m > 2$, there are three ratios $p < q < r$ such that $E(x_q)$ and $E(x_r)$are Cantor sets while $E(x_q)$is a Cantorval. By our best knowledge it is the first result of this type for multigeometric sequences.

Now we will focus on Ferens-like sequences $x_q = (m+k, \ldots, k; q)$ where $m \geq k$.

For $k = 1$ the Ferens-like sequence $x_q = (m+1, \ldots, 2, 1; q)$ has

$$\Sigma = \left\{0, 1, 2, \ldots, \frac{(m+2)(m+1)}{2}\right\}.$$

The set $E(x_q)$ is a Cantor set (for $q < \frac{1}{|\Sigma|}$) or an interval (for $q \geq \frac{1}{|\Sigma|}$); see Theorem 7 in [143]), Theorem (3.2.1) or Theorem (3.2.4).

For $k = 2$, the "shortest" Ferens-like sequence is $x_q = (4, 3, 2; q)$. For this sequence

$$\Sigma = \{0, 2, 3, 4, 5, 6, 7, 9\}.$$

Note that the same Σ has Guthrie-Nymann-Jones sequence $(3, 2, 2, 2; q)$ (see Example (3.2.13)). It follows that $E(x_q)$is a Cantor set for $q \in \left(0, \frac{1}{8}\right)$ and $E(x_q)$ is a Cantorval for $= \frac{1}{8}$. By Theorem (3.2.4), $K(\Sigma; q)$ is an interval for $q \geq I(\Sigma) = \frac{2}{11}$ and a Cantorval for $q \in \left(\frac{1}{6}, \frac{2}{11}\right)$. As shown in Example (3.2.13), for almost all $q \in$ ▯18 ,16 _ the set$K(\Sigma; q)$ has positive Lebesgue measure. Using Theorem (3.2.16), we can find a decreasing sequence (q_n) tending to $\frac{1}{8}$for which the sets $K(\Sigma; q_n)$ have zero Lebesgue measure.

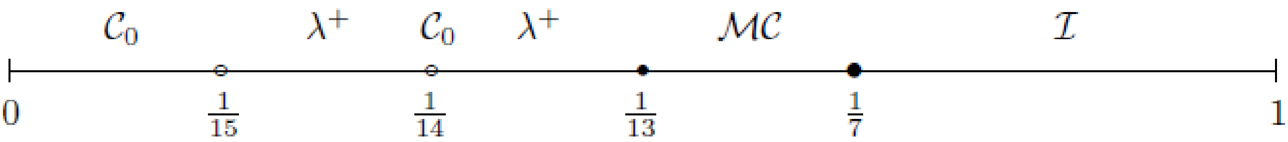
For $k = 3$ the "shortest" Ferens-like sequence is $x_q = (6, 5, 4, 3; q)$. For this sequence

$$\Sigma = \{0, 3, \ldots, 15, 18\}$$

and $|\Sigma| = 15$. Since $1 \in \frac{1}{15}\Sigma$ the set $\Sigma_2 = \Sigma + \frac{1}{15}\Sigma$ has less than $|15|^2$ elements (for example 4 can be presented as $4 + 0$ or as $3 + 1$). Therefore $\frac{1}{15^2}\Sigma|\Sigma_2| < 1$ and for $q = \frac{1}{15}$ the set $E(x_q)$ is a Cantor set according to Theorem (3.2.14). Moreover, calculating for $q = \frac{1}{14} > \frac{1}{15}$ the cardinality

$$|\Sigma_3| = |\Sigma + q\Sigma + q^2\Sigma| = 2655 < 14^3$$

and applying Theorem (3.2.14), we conclude that the achievement set $E(x_q)$ is a Cantor set of zero Lebesgue measure for $q = \frac{1}{14}$. On the other hand, Corollary (3.2.11) implies that for almost all $q \in \left(\frac{1}{15}, \frac{1}{1+\sqrt{18}}\right)$ the achievement set$E(x_q)$ has positive Lebesque measure. The set Σ has $(\Sigma) = \frac{1}{13}$ and $I(\Sigma) = \frac{3}{21} = \frac{1}{7}$. So, in this case we have the diagram:



As in the previous case, we can use Theorem (3.2.16) (taking $a = b = 3$ and $c = -1$) and find a decreasing sequence (q_n) tending to 115 such that all $E(x_{q_n})$ have zero Lebesgue measure.

Suppose now that $k > 3$. For the Ferens-like sequence $x_q = (k + m, \ldots, k + 1, k; q)$ its sumset Σ contains the number $|\Sigma|$, which implies that $|\Sigma + q\Sigma| < |\Sigma|^2$ for $q = \frac{1}{|\Sigma|}$and therefore $E(x_q)$ is a Cantor set of zero measure according to Theorem (3.2.14).

For a contraction ratio $q \in \left\{\frac{1}{n+1} : n \in \mathbb{N}\right\}$ self-similar sets of positive Lebesgue measure can be characterizedas follows:

Theorem (3.2.17)[158]: Let $\Sigma \subset \mathbb{Z}$ be a finite set, $q \in \left\{\frac{1}{n+1} : n \in \mathbb{N}\right\}$ and $\Sigma_n = \sum_{i=0}^{n-1} q^i \Sigma$ for $n \in \mathbb{N}$. For the compact set $K = K(\Sigma; q)$ the following conditions are equivalent:

(i) $|\Sigma_n| \cdot q_n \geq 1$ for all $n \in \mathbb{N}$;

(ii) inf$n \in \mathbb{N}$ $|\Sigma_n| \cdot q_n > 0$,

(iii) $\lambda(K) > 0$.

Proof. The implication (iii)⇒(i) follows from Theorem (3.2.14) while (i)⇒(ii) is trivial. It remains to prove (ii)⇒(iii).

Suppose that $\lambda(K) = 0$. Given any $r > 0$ consider the r-neighborhood $H(K, r) = \{h \in \mathbb{R} : dist(h, K) < r\}$ of the set $K = K(\Sigma; q)$. Take any point $z \in \left\{\sum_{i=n}^{\infty} x_i q^i : \forall i \geq n\ x_i \in \Sigma\right\}$ and observe that $\Sigma_n + z \subset K = \left\{\sum_{i=n}^{\infty} x_i q^i : \forall i \geq (x_i)_{i\in\omega} \in \Sigma^\omega\right\}$, which implies that $H(\Sigma_n + z, r) \subset H(K, r)$ for all $r > 0$. The continuity of the Lebesgue measure implies that $\lambda(H(K, r)) \to 0$ when r tends to zero. It follows from $\Sigma \subset \mathbb{Z}$ and $\frac{1}{q} \in \mathbb{N}$ that

$$\Sigma_n \subset q^{n-1} \cdot \mathbb{Z}.$$

Hence, for any two different points x and y from Σ_n, the distance between x and y is no less then $q^{n-1} > q^n$.

Therefore, for any $n \in \mathbb{N}$,

$$|\Sigma_n| \cdot q^n = \lambda\left(H\left(\Sigma_n, \frac{1}{2}\, q^n\right)\right) = \lambda\left(H\left(\Sigma_n + \frac{1}{2}\, q^n\right)\right) \leq \lambda\left(K, \frac{1}{2}\, q^n\right)$$

which means that $\lim_{n\to\infty} |\Sigma_n| \cdot q^n = 0$.

Theorems (3.2.17) combined with Corollary 2.3 of [128] imply the following corollary.

Corollary (3.2.18)[158]: For a finite subset $\Sigma \subset \mathbb{Z}$ and the number $q = \frac{1}{|\Sigma|} < 1$ the following conditions are equivalent:

(i) $K(\Sigma; q)$ has positive Lebesgue measure;

(ii) $K(\Sigma; q)$ contains an interval; (4) for every $n \in \mathbb{N}$ the set $\sum_{k=0}^{n-1} q^k\, \Sigma$ has cardinality $|\Sigma_n| = |\Sigma|^n$.

Section (3.3): Subsum Sets of Null Sequences

When the sequence is not absolutely summable, its subsum set is an unbounded closed interval which includes zero. The subsum set of an absolutely summable sequence is one of the following: a finite union of (nontrivial) compact intervals, a Cantor set, or a "symmetric Cantorval," a hybrid Cantor-like set with both trivial and nontrivial components.

It is a counterintuitive fact that, while every summable sequence of real numbers must converge to zero, there are sequences (notably the harmonic sequence $\{1/n\}$) which converge to zero but are not summable.

However, every such sequence has many summable subsequences (for example, the sequence of negative powers of any integer greater than two is a summable subsequence of the harmonic one). It seems natural to ask what kind of set is formed by the collection of all sums of (summable) subsequences of our original one. Such sums are called "subsums" in [165], [147], and [148] (in the latter, the German word "Teilsumme" is used) and accordingly we shall refer to this set as the subsum set of our sequence.

The description of the subsum set of a general sequence turns out to be a challenging question. I set out trying to answer it and came up with a number of interesting conclusions, but could not come up with a general description of the possible subsum sets on my own. A comment by Michał Misiurewicz led me by chance to a 1988 J.

A. Guthrie and J. E. Nymann [147] who give a complete topological description of such sets and review earlier work on the problem, notably the results of S. Kakeya [150] and H. Hornich [148].

I will briefly describe the results in [167] and then explain the Guthrie–Nymann result which completes the picture, at least in the case of null sequences. It should be noted that Jones also obtains some results for sequences which do not converge to zero, although the core of what we know (and the most complete picture) is contained in the null sequence case.

Formally, given a real sequence $\{x_i\}$, a subsequence can be written in the form $\{\xi_i . x_i\}$, where $\xi = \{\xi_i\}_{i=1}^{\infty}$ is a sequence of zeroes and ones (determining which terms of the original sequence are included). It is a summable subsequence if the sum $\sum_{i=1}^{\infty}\{\xi_i . x_i\}$ converges, and in that case the sum is a subsum of the original sequence.

The subsum set of $\{x_i\}$, which we denote $\sum(\{x_i\}_{i=1}^{\infty})$, is the collection of all subsums of $\{x_i\}$. Note that in this formulation $\sum(\{x_i\}_{i=1}^{\infty})$ includes the sums of finite subsequences of $\{x_i\}$ in addition to the sums of (summable) infinite subsequences, and also the empty subsequence (which sums to zero). The bulk of the discussion is devoted to the case of positive sequences; the general picture can be deduced from this subcase. Note that positive convergent (sub) sequences are unconditionally summable; they can be rearranged in any order without changing their sum. We will work under the standing assumption that the sequence is reordered to be nonincreasing: $x_{i+1} \leq x_i$ for all.

When $\{x_i\}$ is not summable (so $\sum(\{x_i\}_{i=1}^{\infty}) = \infty$), the subsum set is an unbounded interval [165]. This is formally a special case of [167] which we will state a little later, but the argument is somewhat different from the summable case.

Proposition (3.3.1)[164]: If $\{x_i\}$ *is* a null sequence of positive numbers which is not summable, then $\sum(\{x_i\}_{i=1}^{\infty}) = [0, \infty)$.

The basic observation is that given $\varepsilon > 0$, any finite string of successive terms $x_n, x_{n+1}, \dots, x_{n+k}$ with n sufficiently large has each term less than ε, but their sum can be made arbitrarily large by making k sufficiently large. Then given $r > 0$, we can pick n and k (sufficiently large) so that $x_n + \cdots x_{n+k} \leq r < x_n + \cdots + x_{n+k+1}$; this means the sum on the left is at least $r - \varepsilon$ and a "bootstrap" argument shows that we can pick a sequence of strings whose sums add up to exactly r .

Two kinds of behavior are easily observed for positive summable sequences (in which case every subsequence is summable). Subsums of the sequence $\{2^{-i}\}$ are nothing other than binary representations of numbers in the closed interval$[0, 1]$, and since every such number has a binary representation it follows that $\sum(\{2^{-i}\}_{i=1}^{\infty}) = [0, 1]$.

By contrast, subsums of the sequence $\{3^{-i}\}$ are ternary representations of numbers in [0,1], but only of those which can be expressed without using the digit "2"—and this is easily seen to be the middle-third Cantor set built on the interval $[0, 1/2]$. In fact, These two examples are templates for the topological type of many sequences, in particular the geometric ones. This can be made clear via an analysis given in [167] and implicit in [147], and which will also make clear the kind of behavior that leads to a set which is fundamentally different from either of these two.

Starting from a positive summable sequence$\{x_i\}$, we call x_k the kth term; let us denote by X_k the kth tail obtained by summing all the terms following the kth term:

$$X_k = \sum_{i>k} x_i \ .$$

With this formulation, the sum of the whole series is X_0, and clearly any subsum is contained in the closed interval $[0, X_0]$.

Now partition all the subsums into those that don't involve the first term x_1 and those that do: using our formulation of subsums as determined by binary sequences$\xi = \{\xi_i\}$, this simply corresponds to the choice of ξ_1. If $\xi_1 = 0$, then the subsum $s(\xi) = \sum_{i=1}^{\infty}\{\xi_i . x_i\}$ does not involve $x1$ and hence is contained in the interval

$$J_0 = [0, X_1],$$
While if $\xi_1 = 1$, then it is contained in
$$J_1 = [x_1, x_1 + X_1] = [x_1, X_0].$$
The union
$$C_1 = J_0 \cup J_1$$
of these two intervals contains the whole subsum set. Whether these intervals are disjoint or not is determined by the relative size of the first term and the first tail. If $x_1 \leq X_1$, then $C_1 = [0, X_0]$, while if $x_1 > X_1$, then C_1 is the union of two disjoint subintervals of $[0, X_0]$. But we can apply this argument recursively. Given the binary "initial word" $w_k = \xi_1, \xi_2, \dots, \xi_k$ of length k for the binary sequence, denote by $s_k(w_k) = \sum_{i=1}^{k}(\xi_i . x_i)$ the finite sum corresponding to the sequence consisting of w_k followed by all zeroes; we can guarantee that any subsum determined by a sequence

Starting with w_k lies in the interval
$$J_{\omega_k} = [s_k(\omega_k), s_k(\omega_k) + X_k].$$
Clearly the whole subsum set is contained in the union C_k of all the intervals J_{ω_k} as ω_k ranges over all the binary sequences of length k. Furthermore, the transition from C_{k-1} to C_k consists of replacing each interval $J_{\omega_{k-1}}$ with two intervals, $J_{\omega_{k-}}$ and $J_{\omega_{k+}}$, corresponding to the words of length k which start with ω_{k-1}: $\omega_{k-} = \omega_{k-1}, 0$ and $\omega_{k+} = \omega_{k-1}$, 1. As above, the effect of this substitution is determined by the relation between the kth term and the kth tail.

Term exceeds Tail if $x_k > X_k$, then the two intervals are disjoint, so for each word ω_{k-1} of length $k - 1, J_{\omega_{k-1}}$ in $\mathcal{C}_{k-1}$ is replaced by a *disjoint* union of two subintervals in $\mathcal{C}_k$; that is, $J_{\omega_{k-1}}$ breaks into the disjoint union of $J_{\omega_{k-}}$ and $J_{\omega_{k+}}$, leaving a "gap" of size $x_k - X_k$ in the middle.

Tail bounds Term if $x_k \leq X_k$, then the two intervals share at least one point, so their union equals $J_{\omega_{k-1}}$.

Note that this description of the transition from $\mathcal{C}_{k-1}$to $\mathcal{C}_k$ is the same for all the intervals making up $\mathcal{C}_{k-1}$, independent of the initial word ω_{k-1} that determines them.

Note also that the length of each of the intervals J_{ω_k} is X_k . This means in particular that the difference between two subsums corresponding to binary words whose first k terms agree is at most X_k . But these tails converge to zero, since our (total) series is convergent. It follows that the map assigning to each binary sequence ξ the subsum $s(\xi)$ is continuous, using the product topology on the set $\{0, 1\}^N$ of binary sequences.
This topology turns the set of binary sequences into a Cantor set, and as a consequence the subsum set, being the continuous image of a compact set, is itself compact (and in particular closed).

We note in passing that every such subsum set is symmetric, in that the "flip" taking $x \in [0, X_0]$ to $X_0 - x$ takes the subsum $s(\xi)$ determined by ξ to the subsum $s(\bar{\xi})$ determined by the sequence $\bar{\xi}$ which is obtained from ξ by replacing each 0 with 1 and vice versa, and hence takes $\sum(\{x_i\}_{i=1}^{\infty})$ onto itself.
When one of the two scenarios above governs the transition from $\mathcal{C}_{k-1}$to $C_k$for every$k$, this construction determines the topology of the subsum set completely.
Proposition (3.3.2)[164]: Suppose that $\{x_i\}$ is a positive, nonincreasing summable sequence with $\sum_{i=1}^{\infty} x_i = X_0$.

(i) [167] $\sum(\{x_i\}_{i=1}^{\infty}) = [0, X_0]$ if and only if for every k, the kth tail bounds the kth term; more generally, if this condition holds eventually, then $\sum(\{x_i\}_{i=1}^{\infty})$ is a finite union of (nontrivial) closed intervals.
(ii) [167] If for every k the term exceeds the tail, then $\sum(\{x_i\}_{i=1}^{\infty})$ is a (centered 4) Cantor set.

The first statement above is a remark given without proof by Kakeya [150] and both statements are in essence proved by Hornich [148]. To prove that we actually get a Cantor set in the second scenario—that is, to show the subsum set is totally disconnected—we need to invoke the fact that the lengths of the intervals J_{ω_k} go to zero with k.
Furthermore, in the second scenario the set C_k consists of 2^k disjoint intervals, each of length X_k, from which it follows that the Cantor set we finally obtain has Lebesgue measure $\lim 2^k X_k$ (this is noted by Jones, but not explicitly by Hornich).

If we want to have a picture of all possible subsum sets for positive null sequences, we need to handle the slippery ground between these two extreme scenarios—cases where sometimes the term exceeds the tail and other times the tail bounds the term.
In view of the comments above, this is only a problem when each scenario occurs infinitely often. Kakeya suggests that to get a totally disconnected subsum set (and hence, in view of the above, a Cantor set) it might be sufficient to have the terms exceed the tails infinitely often. One aspect of this vision works, but another doesn't.

It is true that if terms exceed tails infinitely often, then the subsum set will have infinitely many components, some of them consisting of single points. To see this, note that for every word ω_k of length k, the interval J_{ω_k} is the translate of the leftmost intervalJ_{0k} by $s_k(\omega_k) = \sum_{i=1}^{k} \xi_i . x_i$; since the sequence is nonincreasing, the smallest such sum (other than the zero sum) is $x_k(\omega_k = 0 \ldots 01)$. This means that the only interval J_{ω_k} in C_k intersecting the interval $[0, x_k)$ is J_{0k} . In particular, if x_{k+1} exceeds X_{k+1}, the "gap" between J_{0k+1} and $J_0 k_1$ separates two actual components of the new set C_{k+1}.
Since $x_k \to 0$, this means the left endpoint 0 alone constitutes a trivial component of the subsum set (and is a limit of other components). The same argument can be used to show that the left endpoint of any component of C_k (for any k) is a trivial component of the subsum set, and is a limit of other components. By symmetry, the same is true of all right endpoints of components.

However, once some tails bound the corresponding terms, the components of C_k for later k are unions of overlapping intervals J_{ω_k} , and it is possible that some of the gaps introduced in a given interval J_{ω_k} by the transition to C_{k+1}may be covered by other intervals corresponding to other words. This leads to the possibility that a particular gap introduced in one interval J_{ω_k} during the transition from C_k to C_{k+1} might never become visible as a gap in any subsequent sets C_{k+m}, and allows the possibility that even though $\sum(\{x_i\}_{i=1}^{\infty})$ has infinitely many components, some of them might be nontrivial intervals.
In fact this phenomenon does occur. Guthrie and Nymann [147] show that the sequence given by

$$x_{2k-1} = \frac{3}{4^k} \ , x_{2k} = \frac{2}{4^k}$$

has a subsum set that contains the interval [3/4, 1]; but the even-numbered terms exceed the corresponding tails, so there are infinitely many components. Hence the subsum set is neither a finite union of intervals nor a Cantor set. Jones [167] gives a different example, due to Dan Velleman. We give one way to generate many examples. We repeat it here, inspired by [168].

Proposition (3.3.3)[164]: (R. Kenyon). Suppose we are given $n \in \mathbb{N}$ and n integers d_0, $d_1, \dots, d_n$such that

$$d_j \equiv j \bmod n.$$

Then the set of "generalized base n expansions" using these "digits"

$$S = \left\{ \sum_{i=1}^{\infty} \frac{a_i}{n_i} \right\} \{a_i \in \{d_0, \dots, d_n\}$$

has nonempty interior.

Proof. The first step is to confirm the somewhat optimistic intuition that, since the digits include representatives of all the congruence classes mod n, the finite sums of the form

$$\sum_{i=1}^{k} \frac{a_i}{n_i} a_i \in \{d_0, \dots, d_n\}$$

Should, by analogy with the standard case $d_j = j$, have fractional parts that include all rational numbers of the form $\frac{a}{n^k}$. The "obvious" reasoning we might expect does not apply. For example, $\frac{1}{4} + \frac{2}{4^2} = \frac{6}{16}$ while $\frac{1}{4} + \frac{2}{4^2} = \frac{6}{16}$; the difference is not an integer even though $6 = 2$ mod 4. However, it *is* true that different expressions of this form have *different* fractional parts. To see this, suppose we have two such sums with the same fractional part:

$$\frac{a_1}{n} + \frac{a_2}{n^2} + \cdots + \frac{a_k}{n^k} = \frac{b_1}{n} + \frac{b_2}{n^2} + \cdots + \frac{b_k}{n^k} N$$

(Where each a_i and b_i is one of our digits $d_0, \dots, d_{n-1}$, and $N \in \mathbb{N}$). We can rewrite this as

$$\frac{a_1 - b_1}{n^1} + \frac{a_2 - b_2}{n^2} + \cdots + \frac{a_k - b_k}{n^k} = N$$

and multiply both sides by n^k :

$$n^{k-1}(a_1 - b_1) + n^{k-2} \ n(a_2 - b_2) + \cdots + n(a_{k-1} - b_{k-1}) + (a_k - b_k) \ = n^k N.$$

Taking the congruence class of both sides mod n, we get

$$a_k - b_k \equiv 0 \bmod \quad n.$$

But since the possible digits belong to different congruence classes mod n, we must have

$$a_k - b_k = 0.$$

Thus by induction on $k, a_i - b_i$ for $i = 1.2, \dots, k$.

Now, for a given (fixed) k, there are n^k sums of the form

$$\sum_{i=1}^{k} \frac{a_i}{n^i}$$

as well as n^k fractions of the form $\frac{a}{n^k}$ with $0 \le a < n^k$. Hence by the pigeonhole principle, congruence mod *n* generates a bijection between the two sets, confirming our intuition.

The second step is then to reinterpret this statement to say that the integer translates of $\sum_{i=1}^{\infty} \{\xi_i \, . x_i \}$ cover the whole real line

$$\bigcup_{k \in Z} \left(k + \sum_{i=1}^{\infty} \{\xi_i \, . x_i \} \right) = \mathbb{R}.$$

Finally, we invoke the Baire category theorem, which in our context says that if a countable union of closed sets equals $\mathbb{R}$, then at least one of them has nonempty interior. From this we

conclude that for at least one integer k, $(k + \sum_{i=1}^{\infty}\{\xi_i \,.\, x_i\})$has nonempty interior—but since it is a translate of $\sum_{i=1}^{\infty}\{\xi_i \,.\, x_i\}$, the same is true of $\sum_{i=1}^{\infty}\{\xi_i \,.\, x_i\}$.
For the record, the example shown me by Kenyon is given by

$$x_{2k-1} = \frac{6}{4^k}\ , x_{2k} = \frac{1}{4^k}\ .$$

This is not given in decreasing order, but when it is rearranged, as

$$6/4, 6/16, 1/4, 6/64, 1/16, 6/256, 1/64, \ldots$$

the term exceeds the tail infinitely often. Thus on one hand its subsum set has infinitely many components. On the other, using the pair of terms $\frac{6}{4^k}$ and $\frac{4}{4^k}$, we can also obtain either of the fractions $\frac{0}{4^k}$ and $\frac{7}{4^k}$; since the pairs corresponding to different denominators (i.e., different k) are disjoint, our subsum set contains all the numbers of the form

$$\sum_{k=1}^{\infty}\frac{d_k}{4^k}, \quad d_k\{0, 4, 6, 7\}.$$

Hence, by the proposition, the subsum set has nonempty interior.
Bartoszewicz et al. [166] extend this class of examples. They consider sequences of the form

$$k_1, k_2, \ldots, k_n, k_1 q, \ldots, k_n q, k_q q^2, \ldots$$

Consisting of blocks of n multiples of q by integers $k_1, \ldots, k_n$. They show that if the (finite) subsums of these integers include some collection of n successive integers, and q satisfies the estimates

$$\frac{1}{n+1} \le q < \frac{\min k_i}{\min k_i + \sum k_i},$$

Then the resulting subsum set has infinitely many components and nonempty interior.
Furthermore, each integer in the collection $k_1, \ldots, k_n$ can be replaced by a multiple of it by a (nonnegative integer) power of q and still produce a subsum set of this type.
Note that this result in particular applies to a number of other published examples, namely
(R. Jones/D. Velleman [167]) $\frac{3}{2}, \frac{2}{2}, \frac{2}{2}, \frac{2}{2}, \frac{3a}{2}, \frac{2a}{2}, \frac{2a}{2}, \frac{2a}{2}, \frac{3a^2}{2}, \frac{2a^2}{2}, \ldots$.
For any a satisfying $\frac{1}{5} \le \sum_{k=1}^{\infty} a^k < \frac{2}{9}$,
(Weinstein and Shapiro [157]) $8.\frac{3}{10}, 7\ \frac{3}{10}, 6.\frac{3}{10}, 5.\frac{3}{10}, 4.\frac{3}{10}, 8.\frac{3}{10^2}\ , \ldots,$
And (Ferens [146])$727, 627, 527, 427, 327, 7272\ , \ldots$

In fact, Guthrie and Nymann give a complete topological classification of the subsum sets of positive, summable sequences by rounding out the last remaining possibility after finite unions of nontrivial intervals and Cantor sets. Suppose a subsum set has infinitely many components but nonempty interior. For each k, we can write this set as the union of 2^k translates of the subsum set of the sequence with its first k terms removed. Invoking the Baire category theorem again (in the weaker form involving a finite union) we conclude that one, and hence all, of these translates have nonempty interior. In particular, each interval J_{ω_k} in $\mathcal{C}_k$ contains a subinterval of $\sum(\{x_i\}_{i=1}^{\infty})$. This means that every point of the whole subsum set is within distance X_k of some subinterval of $\sum(\{x_i\}_{i=1}^{\infty})$. Since$X_k \to 0$, the nontrivial components of $\sum(\{x_i\}_{i=1}^{\infty})$ are dense in $\sum(\{x_i\}_{i=1}^{\infty})$. At the same time, our argument showing that there are infinitely many components shows that every endpoint of a nontrivial component of $\sum(\{x_i\}_{i=1}^{\infty})$is a limit of trivial (singleton) components. Aside from Guthrie and Nymann, such sets were studied by Mendes and Oliveira [152] in connection with the structure of arithmetic sums of Cantor sets, motivated by bifurcation phenomena in dynamical systems. They dubbed them Cantorvals. three different varieties of Cantorval can arise, but because of the self-symmetry of subsum sets, the only kind that can arise in our context is what they call an M-Cantorval.

Definition (3.3.4)[164]: A symmetric Cantorval is a nonempty compact subset S of the real line such that:

(i) S is the closure of its interior (i.e., the nontrivial components are dense),

(ii) both endpoints of any nontrivial component of S are accumulation points of trivial (i.e., one-point) components of S.

Using this terminology, we can state Guthrie–Nymann's characterization of subsum sets as follows.

Theorem (3.3.5)[164]: [147] The subsum set of any positive summable sequence is one of the following three possibilities:

(i) a finite union of (disjoint) closed intervals;

(ii) a compact, totally disconnected perfect set (i.e., a Cantor set);

(iii) a symmetric Cantorval.

This statement is, in fact a complete topological classification of such subsum sets.

It is well known that any two Cantor sets are homeomorphic; similarly, Guthrie and Nymann note the following.

Proposition (3.3.6)[164]: Any two symmetric Cantorvals are homeomorphic.

Proof. Given two Cantorvals S and S', first identify the longest component of each; if there is some ambiguity (because several components have the same maximal length), then pick the leftmost one. There is a unique affine, order-preserving homeomorphism between them.

At the same time, consider the gaps (components of the complement in the convex hull) of each of these sets; in a way similar to the above, find an order-preserving homeomorphism between the largest gaps to the left (resp. right) of the components identified above. (Note that, although the gaps are open intervals, the homeomorphism extends to their closures.)

We have defined order-preserving homeomorphisms between each of three subintervals of the convex hull of S and the three corresponding intervals of the convex hull of S. The complement of these intervals consists of four intervals; consider their closures. The part of each Cantorval in each of these closed intervals is again a Cantorval. Thus, we can apply the same algorithm to pair the longest nontrivial component in each of the intervals for S with the corresponding one for S'. Continuing in this way, we get an order-preserving correspondence between the nontrivial components (resp. gaps) of S and those ofS', and an order-preserving homeomorphism between corresponding intervals and gaps. But this means we have an order-preserving continuous mapping f from the union of the nontrivial components and the (closures of) gaps of S onto the corresponding set for S'. These two sets are dense in their convex hulls, which in particular mean their complement is nowhere dense. Since f is order-preserving, we see that for any point x_0 in the complement of the domain of definition, $\sup\{f(x)\}\{x < x_0\}$ and $\inf\{f(x)\}\{x > x_0\}$ must agree. This uniquely extends f to a homeomorphism from the convex hull of S onto that of S, respecting the sets themselves.

So far we have dealt only with positive sequences. However, the description of the subsum sets of any null sequence $\{x_i\}$ can easily be reduced to the positive case by separating the two subsequences consisting of all the positive (resp. negative) terms. The partial sums of each of these subsequences increase (resp. decrease) monotonically, so we can set

$$\sum x_i^+ = X^+ \in [0, \infty],$$
$$\sum x_i^- = X^- \in [-\infty, 0].$$

If the sequence fails to sum absolutely, then at least one of $X^{\pm}\}$ is infinite, and then the subsum set is $[X^-, X^+]$.

If both are finite, the sequence is absolutely summable, and (following [167]), we have

$$\sum(\{x_i\}_{i=1}^{\infty}) = X^- + (\{|x_i\ |\}).$$

It follows that for any null sequence the character of the subsum set is determined by the relation between the terms and the corresponding tails, not of the sequence itself, but of the sequence of their absolute values. This leads to a restatement of the Guthrie–Nymann characterization of subsum sets, but without the hypothesis that the original sequence is positive.

Theorem (3.3.7)[164]: The subsum set of any null sequence is one of the following three possibilities*:*

(i) A finite union of (disjoint) closed intervals;

(ii) A compact, totally disconnected perfect set (i.e., a Cantor set);

(iii) A symmetric Cantorval.

The first of these cases occurs if and only if eventually (i.e., for all sufficiently high k),

$$|x_k| \le \sum_{i=k+1}^{\infty} |x_i|.$$

The second case is guaranteed to happen if eventually

$$|x_k| > \sum_{i=k+1}^{\infty} |x_i|,$$

But either it or the third case can occur if both relations occur for infinitely many k.

Chapter 4
Asymptotic Behaviour of the Lebesgue Measure and an Effective Estimate

We show that the proofs of the results are based on recent progress in infinite ergodic theory, and in particular, they give non-trivial applications of this theory to number theory. We show that closes with a discussion of the thermodynamical significance of the obtained results, and with some applications of these to metrical Diophantine analysis. We provide an effective version of this result, employing mostly basic properties of the transfer operator of the Farey map and an application of Freud's effective version of Karamata's Tauberian theorem.

Section (4.1): Sum-Level Sets for Continued Fractions

We consider classical number theoretical dynamical systems arising from the Gauss map $g: x \longmapsto 1/x \bmod 1$ (for $x \in [0,1]$). It is well known that the inverse branches of g give rise to an expansion of the reals in the unit interval with respect to the infinite alphabet $\mathbb{N}$. This expansion is given by the regular continued fraction expansion

$$[a_1, a_2, \dots] := \cfrac{1}{a_1 + \cfrac{1}{a_2 + \dots}},$$

where all the a_i are positive integers.

We give a detailed measure-theoretical analysis of the following sets $\mathscr{C}_n$, for $n \in \mathbb{N}$, which we will refer to as the sum-level sets:

$$\mathscr{C}_n := \left\{ [a_1, a_2, \dots] \in [0,1]: \sum_{i=1}^{k} a_i = n \quad \text{for some } k \in \mathbb{N} \right\}.$$

A first inspection of the sequence of these sets shows that $\liminf_n \mathscr{C}_n$ is equal to the set of all noble numbers, that is, numbers whose infinite continued fraction expansions end with an infinite block of 1's. Also, one immediately verifies that $\limsup_n \mathscr{C}_n$ is equal to the set of all irrational numbers in [0,1]. Hence, at first sight, the sequence of sum-level sets appears to be far away from being a canonical dynamical entity. In order to state the main results,

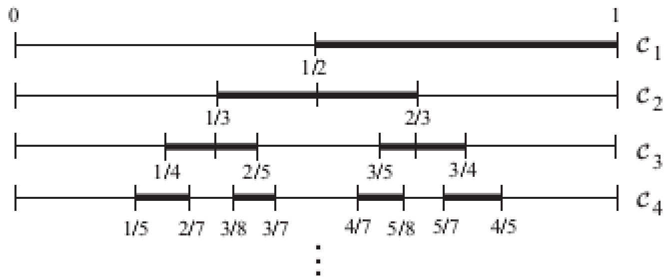


Figure (1)[170]: The first level-sum sets.

note that for the first four members of the sequence of the sum-level sets (cf. Fig. 1) one immediately computes that

$$\lambda(\mathscr{C}_1) = 1/2, \lambda(\mathscr{C}_2) = 1/3, \lambda(\mathscr{C}_3) = 3/10, \lambda(\mathscr{C}_4) = 39/140.$$

From this one might already suspect that $\lambda(\mathscr{C}_n)$ is decreasing for n tending to infinity. In fact, it was conjectured by Fiala and Kleban in [36] that $\lambda(\mathscr{C}_n)$ tends to zero, as n tends to infinity. The first main result is to settle this conjecture.

We give two independent proofs of this theorem. The first of these is almost elementary and only mildly spiced with infinite ergodic theory, whereas the second proof will be deduced from a significantly stronger result. In a nutshell, here we give a detailed proof of the fact that the Farey map T is an exact transformation, which in turn allows to use a criterion of Lin in order to deduce the result. Note that the second proof is very much in

spirit of the considerations in [190], although, strictly speaking, the class of interval maps considered there does not contain the Farey map.
For the next station on our journey of investigating the asymptotic behaviour of the sequence $(\lambda(\mathscr{C}_n))$, we employ the continued fraction mixing property of the induced map of the Farey map T on $\lambda(\mathscr{C}_1)$, in order to show that $\mathscr{C}_1$ is a Darling-Kac set for T. A computation of the return sequence of T then leads to the following theorem, where we use the common notation $b_n \sim c_n$ to denote that $\lim_{n\to\infty} b_n/c_n = 1$.

Our third theorem gives a significant improvement of Theorem (4.2.6) and Theorem (4.2.7). That is, by increasing the dosage of infinite ergodic theory, we obtain the following sharp estimate for the asymptotic behaviour of the Lebesgue measure of the sum-level sets.

We then continue by relating these results on the sum-level sets to the thermodynamical analysis of the Stern-Brocot system obtained in [183]. We first sight, slightly surprising result that this thermodynamical analysis can be obtained from an exclusive use of either the sequence $(\mathscr{C}_n)$ or alternatively its complementary sequence $(\mathscr{C}_n^c)$, rather than using the Stern-Brocot sequence in total. In particular, this reveals that the vanishing of $\lim_{n\to\infty} \lambda(\mathscr{C}_n)$ is very much a phenomenon of the fact that the Stern-Brocot system has a phase transition of order two at the point at which infinite ergodic theory takes over the regime from finite ergodic theory. A detailed discussion of this application to the thermodynamical formalism is given. We apply Theorem (4.2.8) to classical metrical Diophantine analysis, and derive in this way a certain algebraic Khintchine-like law (see [179]).

We defined the sequence $(\mathscr{C}_n)$ of sum-level sets via the sum of the first entries in the continued fraction expansions. For later convenience, let us also add $\mathscr{C}_0 := [0,1]$ to this sequence. Let us begin with some brief comments on various equivalent ways of expressing the sum-level sets.

Recall the following classical construction of Stern-Brocot intervals (SB-intervals) (cf. [187], [173]). For each $n \in \mathbb{N}_0$, the elements of the n-th member of the Stern-Brocot sequence

$$\left\{\frac{s_{n,k}}{t_{n,k}} : k = 1, \dots, 2^n + 1\right\}$$

are defined recursively as follows:

(a) $s_{0,1} := 0$ and $s_{0,2} := t_{0,1} := t_{0,2} := 1$;

(b) $s_{n+1,2k-1} := s_{n,k}$ and $t_{n+1,2k-1} := t_{n,k}$, for $k = 1, \dots, 2^n + 1$;

(c) $s_{n+1,2k} := s_{n,k} + s_{n,k+1}$ and $t_{n+1,2k} := t_{n,k} + t_{n,k+1}$, for $k = 1, \dots 2^n$.

The set $\mathscr{T}_n$ of SB-intervals of order n is given by

$$\mathscr{T}_n := \left\{\left[\frac{s_{n,k}}{t_{n,k}}, \frac{s_{n,k+1}}{t_{n,k+1}}\right] : k = 1, \dots, 2^n\right\}.$$

It might be convenient for the reader to recall that for the Lebesgue measure of these S_B intervals we have $\lambda([s_{n,k}/t_{n,k}, s_{n,k+1}/t_{n,k+1}]) = 1/(t_{n,k}t_{n,k+1})$. One then immediately verifies that in terms of these intervals, the sum-level sets $\mathscr{C}_n$ are given as follows. For $n = 0,1$, we have $\mathscr{C}_0 = \left[\frac{s_{0,1}}{t_{0,1}}, \frac{s_{0,2}}{t_{0,2}}\right]$ and $\mathscr{C}_1 = [s_{1,2}/t_{1,2}, s_{1,3}/t_{1,3}]$. For $n > 1$, we have

$$\mathscr{C}_n = \bigcup_{k=1}^{2^{n-2}} \left[\frac{s_{n,4k-2}}{t_{n,4k-2}}, \frac{s_{n,4k}}{t_{n,4k}}\right].$$

Note that this point of view of $\mathscr{C}_n$ is the one chosen in [36], where $\mathscr{C}_n$ was referred to as the set of even intervals. Also, note that these even intervals are not SB-intervals. However, we clearly have that each of them is the union of two neighbouring SB-intervals of order n. That is,

$$\left[\frac{s_{n,4k-2}}{t_{n,4k-2}},\frac{s_{n,4k}}{t_{n,4k}}\right]=\left[\frac{s_{n,4k-2}}{t_{n,4k-2}},\frac{s_{n,4k-1}}{t_{n,4k-1}}\right]\cup\left[\frac{s_{n,4k-1}}{t_{n,4k-1}},\frac{s_{n,4k}}{t_{n,4k}}\right].$$

Throughout, we will use the notation $\mathscr{C}_n^c$ to denote the set of SB-intervals of order n that are not in $\mathscr{C}_n$. Also, by slight abuse of notation, occasionally we will write $I \in \mathscr{C}_n$ for a SB-interval $I \in \mathscr{T}_n$ which is a subset of $\mathscr{C}_n$.

There is also a way of expressing the sequence $(\mathscr{C}_n)$ in terms of the maps α, β: $\mathscr{C}_0 \to \mathscr{C}_0$ given by

$$\alpha(x) := x/(1+x) \quad \text{and} \quad \beta(x) := 1/(2-x).$$

It is well known that the orbit of the unit interval under the free semi-group generated by α and β is in 1– 1 correspondence to the set of SB-intervals. In fact, by associating the symbol A to the map α and the symbol B to the map β, one obtains that each SB-interval (with the exception the SB-interval of order 0) is associated with a unique word made of letters from the alphabet {A,B}, and vice versa. We will refer to this coding as the Stern-Brocot coding, and will write $I \cong W$ if I is the SB-interval whose Stern-Brocot code is given by $W \in \{A,B\}^k$, for some $k \in \mathbb{N}$. The reader might like to recall that there is a dictionary which translates between Stern-Brocot intervals and continued fraction cylinder sets $[\![a_1,\ldots,a_n]\!] := \{[x_1,x_2,\ldots]: x_k = a_k, k = 1,\ldots,n\}$, which reads as follows. For $\{X,Y\} = \{U,V\} = \{A,B\}$, we have

$$X^{a_1}Y^{a_2}X^{a_3}\cdots U^{a_k}V \cong \begin{cases}[\![a_1+1,a_2,a_3,\ldots,a_k]\!] & \text{for } X = A\\ [\![1,a_1,a_2,\ldots,a_k]\!] & \text{for } X = B.\end{cases}$$

By using this dictionary, it is not hard to see that for $n \geq 2$ we have

$$\mathscr{C}_n = \{I \in \mathscr{T}_n : I \cong W\,XY \text{ for} \{X,Y\} = \{A,B\} \text{ and } W \in \{A,B\}^{n-2}\}.$$

To illustrate this way of viewing $\mathscr{C}_n$, we list the first members of this sequence of code words:

$\mathscr{C}_1$: B

$\mathscr{C}_2$: $AB\ \ BA$

$\mathscr{C}_3$: $AAB\ ABA\ BAB\ BBA$

$\mathscr{C}_4$: $AAAB\ AABA\ ABAB\ ABBA\ BAAB\ BABA\ BBAB\ BBBA$

⋮

The sequence $(\mathscr{C}_n)$ can also be expressed with the help of the Farey map T: $\mathscr{C}_0 \to \mathscr{C}_0$. For this, recall that T is given by T

$$T(x) := \begin{cases}\dfrac{x}{1-x} & \text{for} \quad x \in \left[\dfrac{0,1}{2}\right]\\ (1-x)/x \text{ for } x \in \left(\dfrac{1}{2,1}\right],\end{cases}$$

and that the inverse branches of T are given by

$$u_0(x) := \frac{x}{1+x} \quad \text{and} \quad u_1(x) := 1/(1+x).$$

The associated Markov partition is then given by $\{L,R\}$, where $L := \mathscr{C}_0 \setminus \mathscr{C}_1$ and $R := \mathscr{C}_1$, and each irrational number in $\mathscr{C}_0$ has a Markov coding $x = \langle x_1, x_2, \ldots\rangle \in \{L,R\}$^ℕ, given by $T^{k-1}(x) \in x_k$ for all $k \in \mathbb{N}$. This coding will be referred to as the Farey coding, and will write $I \triangleq W$ if I is the SB-interval whose Farey code is given by $W \in \{L,R\}^k$, for some $k \in$

$\mathbb{N}$. The dictionary which translates between Farey codes and continued fraction cylinders reads as follows:

$$L^{a_1-1}RL^{a_2-1}RL^{a_3-1}\cdots L^{a_k-1}R \triangleq [\![a_1,a_2,a_3,\ldots,a_k]\!].$$

By using this dictionary, it is not hard to see that we have, for each $n \in \mathbb{N}$,

$$\mathscr{C}_n = \{I \in \mathscr{T}_n : I \triangleq W\ R \text{ for } W \in \{L,R\}^{n-1}\}.$$

Again, let us list the first members of this sequence of code words:

$$\mathscr{C}_1:\ \ R$$
$$\mathscr{C}_2:\ LR\ RR$$
$$\mathscr{C}_3: LLR\ LRR\ RLR\ RRR$$
$$\mathscr{C}_4: LLLR\ LLRR\ LRRR\ LRLR\ RRLR\ RRRR\ RLRR\ RLLR$$
$$\vdots$$

The crucial link between the sequence of sum-level sets and the Farey map is now given by the following lemma.

Lemma (4.2.1)[170]: For all $n \in \mathbb{N}$, we have that

$$T^{-(n-1)}(\mathscr{C}_1) = \mathscr{C}_n.$$

Proof. By computing the images of $\mathscr{C}_1$ under u_0 and u_1, one immediately verifies that $T^{-1}(\mathscr{C}_1) = \mathscr{C}_2$. We then proceed by way of induction as follows. Assume that for some $n \in \mathbb{N}$ we have that $T^{-(n-1)}(\mathscr{C}_1) = \mathscr{C}_n$. Since $T^{-n}(\mathscr{C}_1) = T^{-1}\left(T^{-(n-1)}(\mathscr{C}_1)\right) = T^{-1}(\mathscr{C}_n)$, it is then sufficient to show that $T^{-1}(\mathscr{C}_n) = \mathscr{C}_{n+1}$. For this, let $x = [a_1, a_2, \ldots] \in \mathscr{C}_n$ be given. Then there exists $\ell \in \mathbb{N}$ such that $x \in [\![a_1,\ldots,a_\ell]\!]$ and $\sum_{i=1}^{\ell} a_i = n$. By computing the images of x under u_0 and u_1, one immediately obtains that $T^{-1}(x) = \{[1, a_1, a_2, \ldots], [a_1+1, a_2, \ldots]\}$. Clearly, since $1 + \sum_{i=1}^{\ell} a_i = (a_1+1) + \sum_{i=2}^{\ell} a_i = n+1$, this shows that $T^{-1}(x) \subset \mathscr{C}_{n+1}$, and hence, $T^{-1}(\mathscr{C}_n) \subset \mathscr{C}_{n+1}$. The reverse inclusion $\mathscr{C}_{n+1} \subset T^{-1}(\mathscr{C}_n)$ follows for instance by counting the SB-intervals in $\mathscr{C}_{n+1}$ and using the dictionary translating between SB-intervals and continued fraction cylinder sets.

For later use we now recall a few elementary facts and results from infinite ergodic theory for the Farey map. It is well known that the infinite Farey system $(\mathscr{C}_0, T, \mathscr{A}, \mu)$ is a conservative ergodic measure preserving dynamical system. Here, $\mathscr{A}$ refers to the Borel σ-algebra of $\mathscr{C}_0$, and the measure μ is the infinite σ-finite T-invariant measure absolutely continuous with respect to the Lebesgue measure λ, which is, up to a multiplicative constant, unique with respect to these properties (cf. [45]). In fact, with $\varphi_0 : \mathscr{C}_0 \to \mathscr{C}_0$ defined by $\varphi_0(x) := x$, it is well known that μ is explicitly given by (see e.g. [175], [185] , [186])

$$d\lambda = \varphi_0 d\mu.$$

Recall that conservative and ergodic means that $\sum_{n\geq 0} \hat{T}^n(f) = \infty$, μ-almost everywhere and for all $f \in L_1^+(\mu) := \{f \in L_1(\mu) : f \geq 0 \text{ and } \mu(f \cdot 1_{\mathscr{C}_0}) > 0\}$. Here, $1_{\mathscr{C}_0}$ refers to the characteristic function of $\mathscr{C}_0$. Also, invariance of μ under T means $\hat{T}(1_{\mathscr{C}_0}) = 1_{\mathscr{C}_0}$, where $\hat{T} : L_1(\mu) \to L_1(\mu)$ denotes the transfer operator associated with the infinite dynamical Farey system, which is a positive linear operator, given by

$$\mu\left(1_C \cdot \hat{T}(f)\right) = \mu\left(1_{T^{-1}(C)} \cdot f\right), \qquad \text{for all } f \in L_1(\mu), C \in \mathscr{A}.$$

Finally, note that the Perron-Frobenius operator $\mathscr{L} : L_1(\mu) \to L_1(\mu)$ of the Farey system is given by

$$\mathscr{L}(f) = |u_0'| \cdot (f \circ u_0) + u_1' \cdot (f \circ u_1), \text{for all } f \in L_1(\mu).$$

One then immediately verifies that the two operators $\hat{T}$ and $\mathscr{L}$ are related as follows:

$$\hat{T}(f) = \varphi_0 \cdot \mathscr{L}(f/\varphi_0), \qquad \text{for all } f \in L_1(\mu)$$

We give two alternative proofs of Theorem (4.2.6). The first of these is more elementary, whereas the second uses exactness of T and a criterion for exactness due to Lin.

Lemma (4.2.2)[170]:

$$\liminf_{n\to\infty} \lambda(\mathscr{C}_n) = 0.$$

Proof. Let $n \in \mathbb{N}$ be fixed such that n>3, and let $k \in \{2,\dots,n-2\}$ be arbitrary. Recall that the set of SB-intervals of order k consists of 2^{k-2} blocks of four adjacent SB-intervals. Now, let $I \subset \mathscr{C}_k$ be a SB-interval of order k such that $\cong W \in \{A,B\}^k$. We then have that I contains the interval $I_{A,n-k} \cong WA^{n-k}$ as well as the interval $I_{B,n-k} \cong W\, B^{n-k}$. Note that $I_{A,n-k}$ and $I_{B,n-k}$ are two distinct SB-intervals of order n which are both contained in $\mathscr{C}_n^c$. Also, it is well known (see e.g. [182]) that in this situation we have, where $a_n \asymp b_n$ means that the quotient a_n/b_n is uniformly bounded away from zero and infinity,

$$\lambda(I) \asymp (n-k)\lambda\left(I_{X,n-k}\right), \quad \text{for each } X \in \{A,B\}.$$

Clearly, $I_{X,n-k} \cap J_{Y,n-k} = \emptyset$, for all $X \in \{A,B\}, I,J \in \mathscr{C}_k (I \neq J)$. Moreover, by construction, we have for each $k,l \in \{2,\dots,n-2\}$ such that $k \neq l$ and such that either k and l are both odd or both even,

$$I_{X,n-k} \cap J_{Y,n-l} = \emptyset, \text{for all } I \in \mathscr{C}_k, J \in \mathscr{C}_l, X \in \{A,B\}.$$

Note that in here we require that k and l are both odd or both even, since for instance for the interval $I \in \mathscr{C}_2$ for which $I \cong AB$ and the interval $J \in \mathscr{C}_3$ for which $J \cong ABA$ we have that $I_{A,n-2} = J_{A,n-3}$. Also, note that we require $k < n-1$, since for instance for the interval $I \in \mathscr{C}_{n-1}$ for which $I \cong W\,B$ we have that $I_{A,1} \notin \mathscr{C}_n^c$. It now follows that for each $k < n-1$ we have

$$\frac{1}{n-k}\lambda(\mathscr{C}_k) = \sum_{I\in\mathscr{C}_k} \frac{1}{n-k}\lambda(I) \asymp \sum_{I\in\mathscr{C}_k} \sum_{X\in\{A,B\}} \lambda\left(I_{X,n-k}\right).$$

Combining these observations, we obtain that

$$\sum_{k=2}^{n-2} \frac{1}{n-k}\lambda(\mathscr{C}_k) \asymp \sum_{k=2}^{n-2} \sum_{I\in\mathscr{C}_k} \sum_{X\in\{A,B\}} \lambda\left(I_{X,n-k}\right) \leq 2\lambda(\mathscr{C}_n^c).$$

To finish the proof, let us assume by way of contradiction that $\liminf_{n\to\infty} \lambda(\mathscr{C}_n) = \kappa > 0$. By the above, we then have that

$$1 \geq \lambda(\mathscr{C}_n^c) \gg \sum_{k=2}^{n-2} \frac{1}{n-k}\lambda(\mathscr{C}_k) \gg \kappa \sum_{k=2}^{n-1} \frac{1}{k} \gg \log n, \text{for all } n \in \mathbb{N},$$

where $a_n \gg b_n$ means that the quotient a_n/b_n is uniformly bounded away from zero. This gives a contradiction, and hence finishes the proof.

For the first proof of Theorem (4.2.6) we also require the following lemma. We might like to recall that the function $\varphi_0 : \mathscr{C}_0 \to \mathscr{C}_0$ is given by $\varphi_0(x) := x$.

Lemma (4.2.3)[170]: On $\mathscr{C}_1$ we have

$$\hat{T}^n \varphi_0 < \hat{T}^{n-1}\varphi_0, for\ all\ n \in \mathbb{N}.$$

Proof. Recall that $\hat{T}g = \varphi_0 \cdot \mathscr{L}(g/\varphi_0)$, where $\mathscr{L}(g) = \sum_{i=0}^{1} (T^{-1})' \cdot (g \circ u_i)$, that is,

$$\hat{T}g(x) = \frac{g\left(u_0(x)\right) + x \cdot g\left(u_1(x)\right)}{1+x}.$$

By [51] it follows that for $D := \{g \in C^2([0,1]) : g' \geq 0, g'' \leq 0\}$ we have $\hat{T}(\mathscr{D}) \subset \mathscr{D}$. The latter displayed formula in particular also shows that $f(1/2) = \hat{T}f(1)$. Moreover, one immediately verifies that $\varphi_0 \in \mathscr{D}$. Hence, for all $x \in \mathscr{C}_1$ we have

$$\hat{T}^n\varphi_0(x) \le \max\{\hat{T}^n\varphi_0(x): x \in \mathscr{C}_1\} = \hat{T}^n\varphi_0(1) = \hat{T}^{n-1}\varphi_0(1/2)$$
$$= \min\{\hat{T}^{n-1}\varphi_0(x): x \in \mathscr{C}_1\} \le \hat{T}^{n-1}\varphi_0(x).$$

First proof of Theorem (4.2.6). Using Lemma (4.2.1), Lemma (4.2.3), the T-invariance of μ, and the fact that $d\lambda = \varphi_0 \cdot d\mu$, we obtain

$$\lambda(\mathscr{C}_{n+1}) = \mu(1_{\mathscr{C}_{n+1}} \cdot \varphi_0) = \mu(1_{T^{-n}(\mathscr{C}_1)} \cdot \varphi_0) = \mu\left(1_{\mathscr{C}_1} \cdot \hat{T}^n(\varphi_0)\right) < \mu\left(1_{\mathscr{C}_1} \cdot \hat{T}^{n-1}(\varphi_0)\right)$$
$$= \mu(1_{\mathscr{C}_n} \cdot \varphi_0) = \lambda(\mathscr{C}_n).$$

Hence, the sequence $(\lambda(\mathscr{C}_n))$ is strictly decreasing. Combining this fact with Lemma (4.2.2), our first proof of Theorem (4.2.6) is complete.

For the second proof of Theorem (4.2.6) recall that a nonsingular transformation S of the σ-finite measure space $(\mathscr{C}_0, \mathscr{A}, m)$ is called exact if and only if for each element A of the tail σ-algebra $\cap_{n\in\mathbb{N}} S^{-n}(\mathscr{A})$ we have that $m(A) \cdot m(A^c) = 0$. Crucial for us here will be a result of Lin [184] which gives a necessary and sufficient condition for exactness of S in terms the dual $\hat{S}$ of S. Lin found that S is exact if and only if

$$\lim_{n\to\infty} \left\|\hat{S}^n(f)\right\|_1 = 0, \text{for all } f \in L_1(m) \text{ such that } m(f) = 0.$$

We begin with by showing that the infinite Farey system $(\mathscr{C}_0, T, \mathscr{A}, \mu)$ is exact. Let us remark that this fact is probably well known to experts in the field of infinite ergodic theory of numbers. We were unable to locate a rigorous proof in the literature, and hence decided to give such a proof here. However, our proof was inspired by the proof of [171].

Proposition (4.2.4)[170]: The Farey map T of the σ-finite measure space $(\mathscr{C}_0, \mathscr{A}, \mu)$ is exact.

Proof. Let $A_0 \in \cap_{n\in\mathbb{N}} T^{-n}\mathscr{A}$ be given such that $m_g(A_0) > 0$, where $dm_g(x) = \left(\log(2)\,(1+x)\right)^{-1} d\lambda(x)$ denotes the Gauss measure. Note that, since μ and m_g are in the same measure class, it is sufficient to show the exactness of T with respect to mg, rather than μ. Therefore, the aim is to show that $m_g(A_0^c) = 0$. For this, first note that, since $A_0 \in \cap_{n\in\mathbb{N}} T^{-n}\mathscr{A}$, there exists a sequence $(A_n)_{n\in\mathbb{N}}$ such that $A_n \in \mathscr{A}$ and $A_0 = T^{-n}A_n$, for all $n \in \mathbb{N}$. Clearly, we then have that $A_{k+m} = T^k A_m$, for all $k, m \in \mathbb{N}_0$. For each $x \in \mathscr{C}_0$, let ρ be defined by

$$\rho(x) := \inf\{n \ge 0: T^n(x) \in \mathscr{C}_1\} + 1.$$

Since T is conservative, we have that ρ is finite, mg-almost everywhere. Define $\rho_n := \sum_{k=0}^{n-1} \rho \circ (g^k)$, and let $\ll x_1, \ldots, x_n \gg := \{\langle y_1, y_2, \ldots\rangle : y_k = x_k\,, k = 1, \ldots, n\}$ denote a cylinder set arising from the Farey coding. Using the facts that mg is g-invariant and of bounded mixing type with respect to g (that is, $m_g([\![a_1, \ldots, a_{m+n}]\!]) \asymp m_g([\![a_1, \ldots, a_m]\!]) \asymp m_g([\![a_{m+1}, \ldots, a_{m+n}]\!])$, for all $(a_1, \ldots, a_{m+n}) \in \mathbb{N}^{m+n}$ and $m, n \in \mathbb{N}$), we obtain for m_g-almost every $x = \langle x_1, x_2, \rangle \ldots [a_1, a_2, \ldots]$,

$$m_g(A_0 | \ll x_1, \ldots, x_{\rho_n(x)} \gg) = \frac{m_g(A_0 \cap \ll x_1, \ldots, x_{\rho_n(x)} \gg)}{m_g(\ll x_1, \ldots, x_{\rho_n(x)} \gg)}$$
$$= \frac{m_g(T^{-(\rho_n(x))} A_{\rho_n(x)} \cap \ll x_1, \ldots, x_{\rho_n(x)} \gg)}{m_g(\ll x_1, \ldots, x_{\rho_n(x)} \gg)}$$
$$= \frac{m_g(g^{-n} A_{\rho_n(x)} \cap \ll x_1, \ldots, x_{\rho_n(x)} \gg)}{m_g(\ll x_1, \ldots, x_{\rho_n(x)} \gg)} = \frac{m_g(g^{-n} A_{\rho_n(x)} \cap [\![a_1, \ldots, a_n]\!])}{m_g([\![a_1, \ldots, a_n]\!])}$$
$$\asymp m_g(\frac{g^{-n} A_{\rho_n(x)} m_g([\![a_1, \ldots, a_n]\!])}{m_g([\![a_1, \ldots, a_n]\!])} = m_g(A_{\rho_n(x)}).$$

Also, by the Martingale Convergence Theorem (cf. [176]), we have for m_g-almost every $x = \langle x_1, x_2, \ldots \rangle$,

$$\lim_{n\to\infty} m_g\left(A_0 \middle| \ll x_1, \ldots, x_{\rho_n(x)} \gg\right) = 1_{A_0}(x).$$

Combining the two latter observations, it follows that $A_0 = \Lambda$ mod m_g, where Λ is defined by

$$\Lambda := \left\{x \in \mathscr{C}_0 : \liminf_n \ m_g\left(A_{\rho_n(x)}\right) > 0\right\}.$$

Since, by assumption, $m_g(A_0) > 0$, we now have that $m_g(\Lambda) > 0$. Hence, to finish the proof, we are left to show that $m_g(\Lambda) = 1$. For this recall that m_g is ergodic and g-invariant. This gives that it is in fact sufficient to show that $g^{-1}\Lambda \subset \Lambda$ mod mg. In other words, in order to complete the proof, we are left to show that $\liminf_n \ m_g\left(A_{\rho_n(g(x))}\right) > 0$ implies $\liminf_n \ m_g\left(A_{\rho_n(x)}\right) > 0$. Since $A_{\rho_{n+1}(x)} = A_{\rho(x)+\rho_n(g(x))} = T^{\rho(x)}A_{\rho_n(g(x))}$, this assertion would follow if we establish that for each $\varepsilon > 0$ and $\ell \in \mathbb{N}$ there exists $\kappa > 0$ such that for all $B \in \mathscr{A}$ with $m_g(B) > \varepsilon$ we have $m_g\left(T^\ell B\right) > \kappa$. Hence, let us assume that $m_g(B) > \varepsilon$, and let α_ℓ denote the Markov partition for the map T^ℓ. Clearly, there are 2^ℓ elements in α_ℓ. This immediately implies that $m_g(A \cap B) > \varepsilon 2^{-\ell}$, for some $A \in \alpha_\ell$. Therefore, using the fact that $T^\ell : A \to \mathscr{C}_0$ is bijective, λ and m_g are absolutely continuous with respect to each other and $\left|\left(T^\ell\right)'\right| \geq 1$, we have $m_g\left(T^\ell B\right) \geq m_g(T^\ell(B \cap A)) \geq \lambda\left(T^\ell(B \cap A)\right)/(2\log 2) \geq \lambda(B \cap A)/(2\log 2) \geq m_g(B \cap A)/2$. This implies that $m_g(T^\ell B) > 2^{-(\ell+1)}\varepsilon$. Hence, by setting in the above $\kappa := 2^{-(\ell+1)}\varepsilon$, the proof follows.

The proof of the following proposition is very much in the spirit of [190]. However, since the Farey map is strictly speaking not contained in the class of maps considered in [190] (see [190]), we decided to include the short proof.

Proposition (4.2.5)[170]: For each $C \in \mathscr{A}$ with $\mu(C) < \infty$, we have that

$$\lim_{n\to\infty} \lambda\left(T^{-n}(C)\right) = 0.$$

Proof. Let $C \in \mathscr{A}$ be given as stated in the proposition. For each $A \in \mathscr{A}$ for which $0 < \mu(A) < \infty$, we then have

$$\begin{aligned}\lambda\left(T^{-n}(C)\right) &= \mu\left(1_{T^{-n}(C)} \cdot \varphi_0\right) = \mu(1_C \circ T^n \cdot \varphi_0) = \mu\left(1_C \circ T^n \cdot \left(\varphi_0 - \frac{1_A}{\mu(A)} + \frac{1_A}{\mu(A)}\right)\right) \\ &\leq \left\|\hat{T}^n\left(\varphi_0 - \frac{1_A}{\mu(A)}\right)\right\|_1 + \mu\frac{T^{-n}(C) \cap A}{\mu(A)} \leq \left\|\hat{T}^n\left(\varphi_0 - \frac{1_A}{\mu(A)}\right)\right\|_1 + \frac{\mu(C)}{\mu(A)} \\ &\to \frac{\mu(C)}{\mu(A)},\end{aligned}$$

for n tending to infinity. Here, the latter follows, since T is exact and $\mu((\varphi_0 - 1_A/\mu(A))) = 0$, and hence, Lin's criterion, mentioned at the beginning, is applicable. Therefore, by choosing $\mu(A)$ arbitrarily large, the proposition follows.

Theorem (4.2.6)[170]:

$$\lim_{n\to\infty} \lambda(\mathscr{C}_n) = 0.$$

Proof. The following lemma gives the first step in our first proof of Theorem (4.2.6). Note that the statement of this lemma has already been obtained in [36], where it was the main result. Nevertheless, as self-contained as possible, we give a short elementary proof of this result.

In Proposition (4.2.5) put $C = \mathcal{C}_1$, and then use the fact that $\mathcal{C}_n = T^{-(n-1)}(\mathcal{C}_1)$, for all $n \in \mathbb{N}$.

Theorem (4.2.7)[170]:

$$\sum_{k=1}^{n} \lambda(\mathcal{C}_k) \sim \frac{n}{\log^2 n}.$$

Proof. We employ several standard arguments from infinite ergodic theory. First, note that it is well known that the induced map $T_{\mathcal{C}_1}$ of the Farey map T on $\mathcal{C}_1$ is conjugate to the Gauss map g. This then immediately gives that $T_{\mathcal{C}_1}$ is continued fraction mixing (see [191]). Therefore, by [45], it follows that $\mathcal{C}_1$ is a Darling-Kac set for T. This implies that there exists a sequence (ν_n) (the return sequence of T) such that

$$\lim_{n\to\infty} \frac{1}{\nu_n}\sum_{i=0}^{n-1} \hat{T}^i 1_{\mathcal{C}_1}(x) = \mu(\mathcal{C}_1) = \log 2\,, \text{uniformly for } \mu-\text{almost every } x \in \mathcal{C}_1.$$

In order to determine the asymptotic type of the sequence (ν_n), recall from [45] that for a set $C \in \mathcal{A}$ such that $0 < \mu\,(C) < \infty$, the wandering rate of C is given by the sequence $\big(W_n(C)\big)$, where

$$W_n(C) := \mu\left(\bigcup_{k=1}^{n} T^{-(k-1)}(C)\right).$$

Let us compute $\big(W_n(C)\big)$ for $C = \mathcal{C}_1$. Namely, for all $n \in \mathbb{N}$ we have

$$W_n(\mathcal{C}_1) = \mu\left(\bigcup_{k=1}^{n} T^{-(k-1)}(\mathcal{C}_1)\right) = \mu\left(1_{\left[\frac{1}{n+1},1\right]}\right) = \log(n+1).$$

Note that this wandering rate is slowly varying at infinity, that is (see e.g. [172]),

$$\frac{\lim\limits_{n\to\infty} W_{k\cdot n}(\mathcal{C}_1)}{W_n(\mathcal{C}_1)} = 1, \text{for each } k \in \mathbb{N}.$$

Also, note that, since T has a Darling-Kac set, it follows from [45] that T is pointwise dual ergodic with respect to μ, that is,

$$\lim_{n\to\infty} \frac{1}{\nu_n}\sum_{i=0}^{n-1} \hat{T}^i f = \mu(f), \text{for all } f \in L^1(\mu).$$

In this situation we then have, by [45], that the return sequence and the wandering rate are related through

$$\lim_{n\to\infty} \left(n \cdot \nu_n / W_n(\mathcal{C}_1)\right) = 1.$$

Combining these observations, the proof of Theorem (4.2.7) follows (cf. [45], [171], [60]).

- The map T is rationally ergodic with respect to μ. That is, there exists a constant $c > 0$ and a set A with $0 < \mu(A) < \infty$ such that for all $n \in \mathbb{N}$,

$$\int_A \left(\sum_{i=0}^{n-1} 1_A \circ T^i\right)^2 d\mu < c\left(\int_A \sum_{i=0}^{n-1} 1_A \circ T^i d\mu\right)^2. \qquad (*)$$

- The map T has the following mixing property. For A with $0 < \mu(A) < \infty$ such that $(*)$ holds, we have for all $U, V \subset A$,

$$\lim_{n\to\infty} \frac{1}{v_n}\sum_{i=0}^{n-1} \mu(U \cap T^{-i}V) = \mu(U)\mu(V).$$

Theorem (4.2.8)[170]:

$$\lambda(\mathscr{C}_n) \sim \frac{1}{\log^2 n}.$$

Proof. As already mentioned in the introduction, the proof of Theorem (4.2.8) will make use of some further, slightly more advanced infinite ergodic theory. Let us begin with by first giving the concepts and results which are relevant for the proof of Theorem (4.2.8). The following concept of a uniform set is vital in many situations within infinite ergodic theory, and this is also the case in our situation here. (For further examples of interval maps (including the Farey map) for which there exist uniform sets we refer to [188], [189].)

(I) [45]) A set $C \in \mathscr{A}$ with $0 < \mu(C) < \infty$ is called uniform for $f \in L_1^+(\mu)$, if μ-almost everywhere and uniformly on C we have that

$$\lim_{n\to\infty} \frac{1}{v_n}\sum_{k=0}^{n-1} \hat{T}^k(f) = \mu(f),$$

where (v_n) denotes the return sequence of T , and uniform convergence is meant with respect to $L_\infty(\mu|C)$.

Note that it is not difficult to see that the Farey map T satisfies Thaler's conditions, among which Adler's condition, i.e. $T''/(T')^2$ is bounded throughout (0,1), is the most important one (see [188], [189]). This then immediately implies that we have the following, where, the function φ_0 is given by $\varphi_0(x) = x$.

(II) Let $C \in \mathscr{A}$ be given with $\lambda(C) > 0$ and so that there exists an ε>0 such that x>ε, for all $x \in C$. We then have that C is a uniform set for the function φ_0.

Now, the crucial notion for proving the sharp asymptotic result of Theorem (4.2.8) is provided by the following concept of a uniformly returning set. (For further examples of one dimensional dynamical systems which allow uniformly returning sets for some appropriate function we refer to [60].)

(III) ([50]) A set $C \in \mathscr{A}$ with $0 < \mu(C) < \infty$ is called uniformly returning for $f \in L_\mu^+$ if there exists an increasing sequence $(w_n) = (w_n(f, C))$ of positive reals such that μ-almost everywhere and uniformly on C we have

$$\lim_{n\to\infty} w_n \hat{T}^n(f) = \mu(f).$$

In order to determine the asymptotic type of the sequence (w_n), we use [50] where we found that

$$\lim_{n\to\infty} W_n(C)/w_n = 1, \text{for all } C \in \mathscr{A} \text{ such that } 0 < \mu(C) < \infty,$$

where $(W_n(C))$ denotes the wandering rate, which we already considered in the proof of Theorem (4.2.7). In [50] it was shown that every uniformly returning set is uniform. Whereas, in [51] we found explicit conditions under which also the reverse of this implication holds. Applying these results of [51] to our situation here, one obtains the following.

(IV) ([51]) Let $C \in \mathscr{A}$ with $0 < \mu(C) < \infty$ be a uniform set, for some $\in L_\mu^+$. If the wandering rate $(W_n(C))$ is slowly varying at infinity and if the sequence $\left(\hat{T}^n(f)\big|_C\right)$ is decreasing, then we have that C is a uniformly returning set for f. Moreover, μ-almost everywhere and uniformly on C we have

$$\lim_{n\to\infty} W_n(C)\hat{T}^n(f) = \mu(f).$$

With these preparations, we can now finish the proof of Theorem (4.2.8) as follows. The idea is to apply the results stated above to the situation in which the set C is equal to $\mathscr{C}_1$. For this, first recall that we have already seen that the wandering rate $\left(W_n(\mathscr{C}_1)\right)$ of $\mathscr{C}_1$ is obviously slowly varying at infinity. In fact, as computed in the proof of Theorem (4.2.7), we have that $\lim_{n\to\infty} n \cdot v_n / W_n(\mathscr{C}_1) = 1$, and also that $W_n(\mathscr{C}_1) \sim \log n$. Secondly, since $\mathscr{C}_1$ is bounded away from zero, the result in (II) gives that $\mathscr{C}_1$ is a uniform set for φ_0. Thirdly, by Lemma (4.2.3), we have that the sequence $\left(\hat{T}^n(\varphi_0)\big|_{\mathscr{C}_1}\right)$ is decreasing. Thus, we can apply the result in the first part of (IV), which then shows that $\mathscr{C}_1$ is a uniformly returning set for the function φ_0. Hence, the second part in (IV) gives that μ-almost everywhere and uniformly on $\mathscr{C}_1$ we have

$$\lim_{n\to\infty} W_n(\mathscr{C}_1)\hat{T}^n(\varphi_0) = \mu(\varphi_0) = 1.$$

Combining these observations, it now follows that

$$\lim_{n\to\infty} \left(\log n \cdot \lambda(\mathscr{C}_n)\right) = \lim_{n\to\infty} (W_n(\mathscr{C}_1) \cdot \mu(1_{\mathscr{C}_1} \cdot \hat{T}^{n-1}(\varphi_0) = \mu\left(1_{\mathscr{C}_1}\right) = \log 2.$$

This finishes the proof of Theorem (4.2.8).

Here we give a brief discussion of some thermodynamical aspects of the results. For this, recall that in [181] and [183] (see also [182]) we studied the multifractal spectrum $\{\tau(s): s \in \mathbb{R}\}$, given by

$$\tau(s) := \dim_H \left(\left\{x = [a_1, a_2, \ldots]: \lim_{n\to\infty} \frac{2\log q_n(x)}{\sum_{i=1}^{n} a_i} = s\right\}\right).$$

Here, $p_n(x)/q_n(x) := [a_1, a_2, \ldots, a_n]$ denotes the n-th approximant of x, and dimH refers to the Hausdorff dimension. In order to recall the results on this spectrum obtained in [183], we require the concept of the pressure function associated with some family $\mathscr{T} = \{\mathscr{T}_n : n \in \mathbb{N}\}$ of sets of subsets of the unit interval. For this, we define the n-th partition function $Z_{\mathscr{T}_n}$ associated with such a family F by

$$Z_{\mathscr{T}_n}(t) = \sum_{I \in \mathscr{T}_n} \left(diam(I)\right)^t,$$

with which we can then define the pressure function of the family $\mathscr{T}$ by

$$P_{\mathscr{T}}(t) := \lim_{n\to\infty} \frac{1}{n} \log Z_{\mathscr{T}_n}(t).$$

The following results give the main outcome concerning the properties of τ and the Stern-Brocot pressure function $P_{\mathscr{T}}$ of the Stern-Brocot family $\mathscr{T} = \{\mathscr{T}_n : n \in \mathbb{N}\}$. This complete thermodynamical description of the Stern-Brocot system was obtained in [183]. Here, $\gamma := (1+\sqrt{5})/2$ denotes the Golden Mean, and $P^*_{\mathscr{T}}$ refers to the Legendre transform of $P_{\mathscr{T}}$ given for $s \in \mathbb{R}$ by $P^*_{\mathscr{T}}(s) := \sup_{t\in\mathbb{R}} \{t \cdot s - P_{\mathscr{T}}(t)\}$.

(a) [183], [181]. For each $s \in [0, 2\log\gamma]$, we have that

$$\tau(s) = -P^*_{\mathscr{T}}(-s)/s,$$

with the convention that $\tau(0) := \lim_{s\searrow 0} -P^*_{\mathscr{T}}(-s)/s = 1$. Also, the dimension function τ is continuous and strictly decreasing on $[0, 2\log\gamma]$ and vanishes outside the interval [0,2log γ). Moreover, the left derivative of τ at 2log γ is equal to $-\infty$. The function $P_{\mathscr{T}}$ is convex, non-increasing and differentiable throughout ℝ, and $P_{\mathscr{T}}$ is real-analytic on (−∞,1) and vanishes on [1,∞). Furthermore, for each $s \in (0, 2\log\gamma]$ there exists an equilibrium measure μ_s for which $\dim_H(\mu_s) = \tau(s)$.

(b) [178], [30] We have that

$P_{\mathscr{T}}(1-\varepsilon) \sim -\varepsilon/\log\varepsilon$, for ε tending to zero form above.

In particular, the Farey system has a second order phase transition at t=1, that is, the function $P'_{\mathscr{T}}$ is continuous and $P''_{\mathscr{T}}$ is discontinuous at t=1.

Note that the vanishing of $\lim_{n\to\infty} \lambda(\mathscr{C}_n)$ is very much a phenomenon of the fact that the Stern-Brocot system exhibits a phase transition of order two at $t = 1$. At this point of intermittency, finite ergodicity breaks down and infinite ergodic theory enters the scene. In particular, by (b), this abrupt transition from finite to infinite ergodic theory happens in a way which is non-smooth.

In the following let $\mathscr{C} := \{\mathscr{C}_n : n \in \mathbb{N}\}$ denote the family of sum-level sets, and let $\mathscr{D} := \{\mathscr{D}_n : n \in \mathbb{N}, n \neq 1\}$ be the family of their complementary intervals, that is, $\mathscr{D}\_n$ denotes the set of intervals given by the $2^{n-2}+1$ connected components of the complement of $\mathscr{C}_n$ in [0,1].

Proposition (4.2.9)[170]: For each $s \in (0, 2\log\gamma]$ and $n \in \mathbb{N}$, we have that

$$\mu_s(\mathscr{C}_n) = \mu_s(\mathscr{D}_n) = 1/2.$$

Whereas for $s = 0$ and $\mu_0 := \mu$, we have, for all $n \in \mathbb{N}(n \neq 1)$,

$$\mu_0(\mathscr{C}_n) = \log 2\,, \mu_0(\mathscr{D}_n) = \infty, \lim_{n\to\infty} \lambda(\mathscr{C}_n) = 0 \text{ and } \lim_{n\to\infty} \lambda(\mathscr{D}_n) = 1.$$

Moreover, the outcome of the above complete thermodynamical description of the Stern Brocot system stays to be the same if we base this type of analysis exclusively on either $\mathscr{C}$ or $\mathscr{D}$, instead of on $\mathscr{T}$. In particular, the pressure functions associated with the systems $\mathscr{C}$, $\mathscr{D}$ and $\mathscr{T}$ coincide, that is,

$$P_{\mathscr{T}}(t) = P_{\mathscr{C}}(t) = P_{\mathscr{D}}(t), \text{for all } t \in \mathbb{R}.$$

Proof. The estimates for the equilibrium measures μ_s (for $s \in (0, 2\log\gamma]$) are immediate consequences of the T-invariance of these measures.

For the equality of the pressure functions one uses the recursive definition of the Stern-Brocot sequence, with which one immediately verifies that

$$t_{n-1,2k-1} t_{n-1,2k} \le t_{n,4k-2} t_{n,4k-1} \le n\, t_{n-1,2k-1} t_{n-1,2k},$$

and

$$t_{n-1,2k} t_{n-1,2k+1} \le t_{n,4k-1} t_{n,4k \le n} t_{n-1,2k} t_{n-1,2k+1}.$$

Combining these estimates, we obtain

$$n^{-|t|} \sum_{I\in\mathscr{T}_{n-1}} \big(diam(I)\big)^t \le \sum_{I\in\mathscr{C}_n} \big(diam(I)\big)^t \le n^{|t|} \sum_{I\in\mathscr{T}_{n-1}} \big(diam(I)\big)^t.$$

This shows that $P_{\mathscr{T}}(t) = P_{\mathscr{C}}(t)$, for all $t \in \mathbb{R}$. The proof of $P_{\mathscr{T}} = P_{\mathscr{D}}$ follows by similar means, and is left to the reader.

Note that for an even interval of any order $n \in \mathbb{N}$ we have, for all $t \in \mathbb{R}$,

$$\left(diam\left(\left[\frac{s_{n,4k-2}}{t_{n,4k-2}}, \frac{s_{n,4k}}{t_{n,4k}}\right]\right)\right)^t = \left(diam\left(\left[\frac{s_{n,4k-2}}{t_{n,4k-2}}, \frac{s_{n,4k-1}}{t_{n,4k-1}}\right]\right) + diam\left(\left[\frac{s_{n,4k-1}}{t_{n,4k-1}}, \frac{s_{n,4k}}{t_{n,4k}}\right]\right)\right)^t$$

$$\asymp \left(diam\left(\left[\frac{s_{n,4k-2}}{t_{n,4k-2}}, \frac{s_{n,4k-1}}{t_{n,4k-1}}\right]\right)\right)^t + \left(diam\left(\left[\frac{s_{n,4k-1}}{t_{n,4k-1}}, \frac{s_{n,4k}}{t_{n,4k}}\right]\right)\right)^t$$

$$= \left(\frac{t_{n,4k-2}}{t_{n,4k-1}}\right)^{-t} + \big(t_{n,4k-1} t_{n,4k}\big)^{-t}.$$

Hence, the pressure function $P_{\mathscr{E}}$ associated with the Farey tree model coincides with the pressure function $P_{\mathscr{C}}$.

By giving an application of Theorem (4.2.8) to elementary metrical Diophantine analysis. For this, first recall the following well-known result of Khintchine (see e.g. [19]), which states that

$$\limsup_{n\to\infty} \frac{\log\left(\frac{a_n}{n}\right)}{\log log\, n} = 1, \text{for } \lambda - \text{almost every } [a_1, a_2, \dots].$$

In contrast to this well-known Khintchine law, Theorem (4.2.8) now gives rise to the following algebraic Khintchine-like law. (For some further results on the statistics of the sum of the first continued fraction digits see [177].)

Lemma (4.2.10)[170]: We have that

$$\limsup_{n\to\infty} \frac{\log\left(\frac{a_{n+1}}{\sum_{i=1}^{n} a_i}\right)}{\log log\left(\sum_{i=1}^{n} a_i\right)} \leq 0, \text{for } \lambda - \text{almost every } [a_1, a_2, \dots].$$

Proof. For each $n \in \mathbb{N}$ and $\varepsilon > 0$, let

$$A_n^\varepsilon := \bigcup_{k\in\mathbb{N}} \left\{[\![a_1, \dots, a_{k+1}]\!]: \sum_{i=1}^{k} a_i = n, a_{k+1} \geq n(\log n)^\varepsilon\right\},$$

and define

$$\mathscr{A}_n^\varepsilon := \bigcup_{I\in A_n^\varepsilon} I.$$

Then note that a routine calculation for the Lebesgue measure of continued fraction cylinder sets gives, for all $k, \ell \in \mathbb{N}$,

$$\sum_{ak+1\geq \ell} \lambda([\![a_1, \dots, a_k, a_{k+1}]\!]) \asymp \ell^{-1}\lambda([\![a_1, \dots, a_k]\!]).$$

Using this estimate and Theorem (4.2.8), we obtain

$$\begin{aligned}
\lambda(\mathscr{A}_n^\varepsilon) &= \sum_{k=1}^{n} \sum_{\substack{(a_1,\dots,a_k) \\ \sum_{i=1}^{k} a_i = n}} \sum_{a_{k+1}\geq n(\log n)^\varepsilon} \lambda([\![a_1, \dots, a_k, a_{k+1}]\!]) \\
&\asymp \sum_{k=1}^{n} \sum_{\substack{(a_1,\dots,a_k) \\ \sum_{i=1}^{k} a_i = n}} \frac{\lambda([\![a_1, \dots, a_k]\!])}{(\log n)^\varepsilon} \\
&= (n(\log n)^\varepsilon)^{-1} \sum_{k=1}^{n} \sum_{\substack{(a_1,\dots,a_k) \\ \sum_{i=1}^{k} a_i = n}} \lambda([\![a_1, \dots, a_k]\!]) = (n(\log n)^\varepsilon)^{-1}\lambda(\mathscr{C}_n) \\
&\sim \frac{\log 2}{n(\log n)^{1+\varepsilon}}.
\end{aligned}$$

A straight forward application of the Borel-Cantelli Lemma then gives that

$$\lambda(\limsup_n \mathscr{A}_n^\varepsilon = 0, \text{for each } \varepsilon > 0.$$

Hence, by considering the complement of $\limsup_n \mathscr{A}_n^\varepsilon$ in $\mathscr{C}_0$, we have now shown that, for each $\varepsilon > 0$ and for λ-almost all $[a_1, a_2, \dots]$,

$$a_{k+1} < \left(\sum_{i=1}^{k} a_i\right)\left(\log \sum_{i=1}^{k} a_i\right)^\varepsilon, \text{for all } k \in \mathbb{N} \text{ sufficiently large.}$$

By taking logarithms on both sides of the latter inequality, the lemma follows see [180].

Section (4.2): Lebesgue Measure of Preimages of Iterates of the Farey Map

In a letter to Laplace in 1812, Gauss posed the problem of estimating the error

$$\lambda([a_1, a_2, \dots]:[a_{n+1}, a_{n+2}, \dots] < u) - \frac{log(u+1)}{log 2}$$

as n approaches ∞, where λ is the Lebesgue measure, $u \in (0,1)$ is fixed, and

$$[a_1, a_2, \dots] := \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots}} \quad (a_i \in \mathbb{N})$$

denotes a regular continued fraction expansion. Let $G : [0,1] \to [0,1]$ be the Gauss map defined by

$$G(x) := \begin{cases} \{1/x\} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

and $d\nu = d\lambda/((1+x) log\, 2)$ be the Gauss invariant measure. This problem is equivalent to estimating

$$\lambda(G^{-n}[0,u)) - \nu[0,u). (n \to \infty) \quad (1)$$

We write $f(x) = O(g(x))$, or equivalently $f(x) \ll g(x)$, as $x \to \infty$ if there exist constants $M, N > 0$ such that $|f(x)| \leq M|g(x)|$ for all $x \geq N$ (when $f(x) = O_{a_1,\dots,a_m}(g(x))$, the constants M and N depend on $a_1, \dots, a_m$); and $f(x) \sim g(x)$ as $x \to \infty$ if $\lim_{x\to\infty} f(x)/g(x) = 1$.

Lévy first showed that (1) is $O(q^n)$ for $q = 3.5 - 2\sqrt{2}$, and Wirsing determined the optimal value of q as 0.30366 ... by discovering the spectral gap in the transfer operator of the Gauss map. An exact solution to Gauss's problem was first given by Babenko, who proved that the transfer operator is compact when restricted to a certain Hilbert space of functions. This result was later extended by Mayer and Roepstorff. (See [195]).

We concerned with the analogue of Gauss's problem for the Farey map $F : [0,1] \to [0,1]$ defined by

$$F(x) := \begin{cases} x/(1-x) \text{ if } 0 \leq x \leq \frac{1}{2} \\ (1-x)/x \quad \text{if } \frac{1}{2} < x < 1. \end{cases}$$

Specifically, we analyze the asymptotic behavior of $\lambda(F^{-n}[u,1])$. The Gauss map is conjugate to the induced transformation of the Farey map on [1/2, 1] (see [196]). In spite of this relationship, the Gauss and Farey maps exhibit very different behavior, one of the reasons being that F preserves the infinite measure $d\mu = d\lambda/x$.

In the special case $u = 1/2, F^{-(n-1)}[u,1]$ is the nth sum-level set for continued fractions

$$\mathcal{C}_n := \{[a_1, a_2, \dots] \in [0,1] : \sum_{i=1}^{k} a_i = n \text{ for some } k \in \mathbb{N}\}.$$

This follows from the fact that F maps continued fractions $[a_1, a_2, \dots]$ as follows:

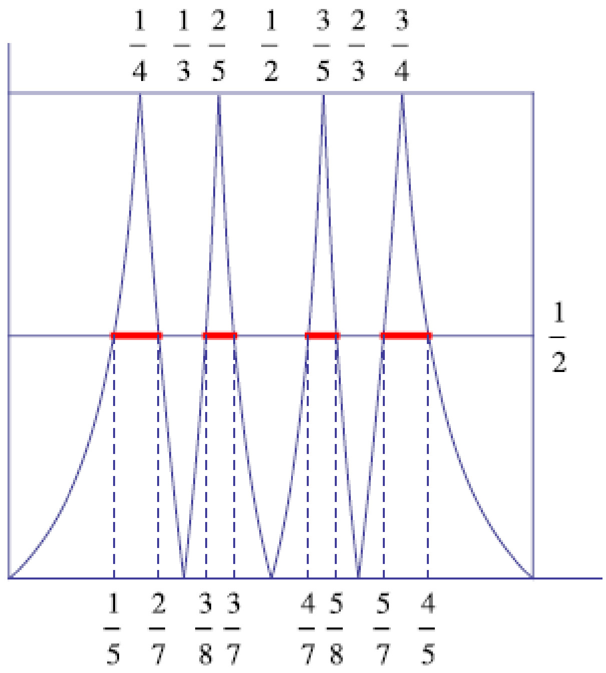


Fig. (1)[192]: The graph of F^3 and $\mathcal{C}_4$ shown as the inverse image $F^{-3}[1/2,1]$.

$$F([a_1, a_2, \dots]) = \begin{cases} [a_1 - 1, a_2, \dots] & \text{if } a_1 \geq 2 \\ [a_2, a_3, \dots] & \text{if } a_1 = 1, \end{cases}$$

from which it is straightforward to see that $F^{-1}(\mathcal{C}_n) = \mathcal{C}_{n+1}$, and hence $\mathcal{C}_n = F^{-(n-1)}(\mathcal{C}_1) = F^{-(n-1)}[1/2, 1]$ (see [170]). This is illustrated in Fig. 1 when $n = 4$.

Confirming a conjecture of Fiala and Kleban [36], Kesseböhmer and Stratmann proved the asymptotic equivalence [170]

$$\lambda(\mathcal{C}_n) \sim \frac{1}{\log_2\ n}, \qquad (n \to \infty)$$

and later, they generalized this result by proving that

$$\lambda(F^{-n}\,[u,1]) \sim \frac{log(1/u)}{\log\ n} \qquad (n \to \infty) \tag{2}$$

for all $u \in (0,1]$. This in fact follows from their stronger result [44] that the sets of the form $F^{-n}[\alpha,\beta]$, with $[\alpha,\beta] \subseteq (0,1]$, equidistribute in $[0,1]$. Their proofs applied deep results in infinite ergodic theory following from Aaronson [45], [193], and Kesseböhmer and Slassi [51], [50], to the Farey map. In particular, they used the inverse relationship between the wandering rate of a uniformly returning set of the Farey map and the decay of the iterates of the Farey map's transfer operator. We prove the following result, which provides an effective version of (2).

Theorem (4.2.1)[192]: For any interval $[\alpha,\beta] \subseteq (0,1]$, we have

$$\lambda\big(F^{-(n-1)}[\alpha,\beta]\big) = \frac{log(\beta/\alpha)}{\log\ n}\left(1 + O_{\alpha,\beta}\left(\frac{1}{\log\ n}\right)\right). (n \to \infty) \tag{3}$$

Instead of proving Theorem (4.2.1) directly, we prove the following result, part (b) of which resembles [170].

To establish Theorem (4.2.1) from this, note that (a) and (b) imply that for all $u \in (0,1)$,

$$\lambda(\mathcal{C}_n^u) \leq \frac{1}{n}\sum_{k=0}^{n} \lambda(\mathcal{C}_{k+1}^u) = \frac{log(1/u)}{\log\ n}\left(1 + O_u\left(\frac{1}{\log\ n}\right)\right),$$

and

$$\lambda(\mathcal{C}_n^u) \geq \frac{1}{n} \sum_{k=n+1}^{2n} \lambda(\mathcal{C}_{k+1}^u) = \frac{1}{n}\left(\frac{2n\, log(1/u)}{\log\ 2n} - \frac{n\, log(1/u)}{\log\ n}\right)\left(1 + O_u\left(\frac{1}{\log\ n}\right)\right)$$

$$= \left(\frac{log(1/u)}{\log\ n} + O_u\left(\frac{1}{\log^2\ n}\right)\right)\left(1 + O_u\left(\frac{1}{\log\ n}\right)\right)$$

$$= \frac{log(1/u)}{\log\ n}\left(1\ +\ O_u\left(\frac{1}{\log\ n}\right)\right),$$

and hence

$$\lambda(\mathcal{C}_n^u) = \frac{log(1/u)}{\log\ n}\left(1\ +\ O_u\left(\frac{1}{\log\ n}\right)\right). \qquad (n\ \to\ \infty)$$

Then subtracting this expression for $u\ =\ \beta$ from that for $u\ =\ \alpha$ yields (3).
To prove Theorem (4.2.2), we make use of the transfer operator $\hat{F}\ :\ L^1(\mu)\ \to\ L^1(\mu)$ of F, which is the positive linear operator satisfying

$$\int_B \hat{F}f\ d\mu = \int_{F^{-1}(B)} f\ d\mu, \text{for all Borel subsets } B\ \subseteq\ [0,1] \text{ and } f\ \in\ L^1(\mu),$$

and is given by

$$\hat{F}f(x) = \frac{f\big(x/(1\ +\ x)\big) +\ xf\big(1/(1\ +\ x)\big)}{1\ +\ x}. \qquad (4)$$

Unlike how the Gauss problem was addressed, we do not appeal to any spectral properties of $\hat{F}$, as obtaining an effective version of (2) through such means appears to be difficult, by the examinations of the spectrum by Isola [1] and Prellberg [199]. We also forego applying strong, general results from infinite ergodic theory to $\hat{F}$. Instead, we establish estimates involving sums of the iterates of $\hat{F}$ specifically, and make careful applications of the equality (6) following from [45] and Karamata's Tauberian theorem [197], which are important results underlying much of the machinery used in [44] and [170], so as to obtain error terms. In particular, we make an application of Freud's effective

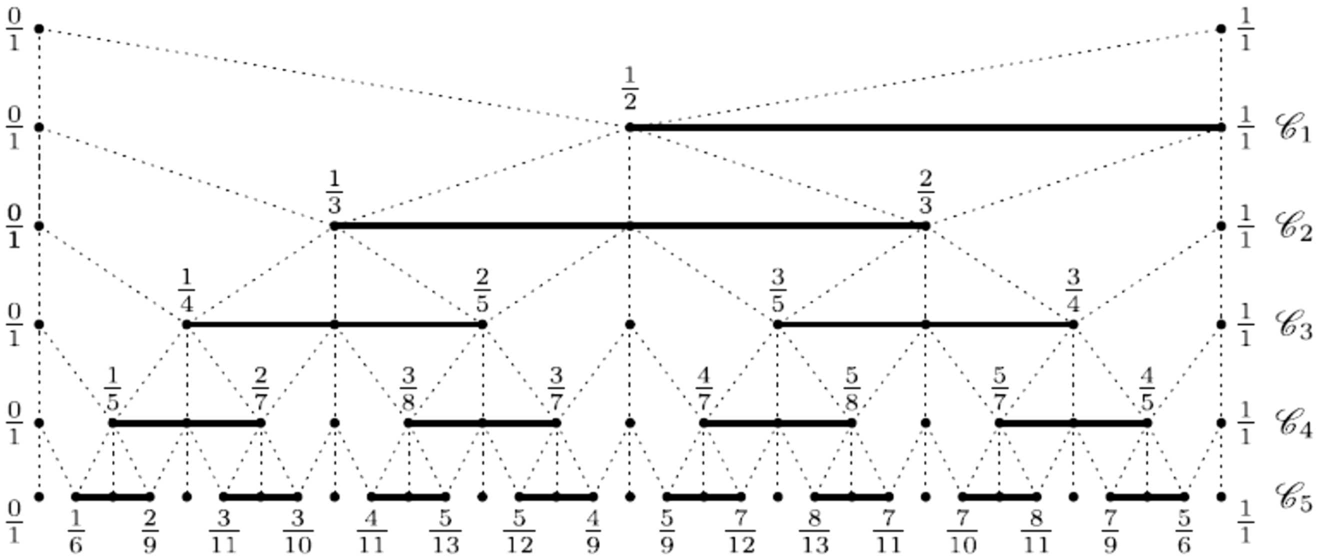


Fig. (2)[192]: Thesum-level sets as Stern–Brocot intervals.
version of Karamata's theorem [194] in establishing an asymptotic estimate of a certain weighted sum of the measures $\lambda(\mathcal{C}_{k+1}^u)$ from an estimate of its Laplace transform derived from (6). We can then remove the weights to prove (b) by a standard analytic number theory

argument. See [198] for other asymptotic results derived from operator renewal theory involving the iterates of transfer operators of infinite measure-preserving systems.

(See Fig. 2.) This was the characterization of $\mathcal{C}_n$ considered by Fiala and Kleban [36] motivated by their study of spin chain models. The relationship between the Farey map and the Stern–Brocot sequence was exploited in [44] to prove the equidistribution of certain weighted subsets of the Stern–Brocot sequence.

Theorem (4.2.2)[192]: Let $u \in (0,1)$ and $\mathcal{C}_n^u := F^{-(n-1)}[u,1]$. Then:

(a) $\left(\lambda(\mathcal{C}_n^u)\right)_n$ is a decreasing sequence;

(b) $\sum_{k=0}^{n} \lambda(\mathcal{C}_{k+1}^u) = \frac{n log(1/u)}{\log n} \left(1 + O_u\left(\frac{1}{\log n}\right)\right). \quad (n \to \infty)$

Proof. Proof of (a) The case in which $u \geq 1/2$ follows from the proof of [170], so assume $u \in (0,1/2)$. Let $\varphi_0 : [0,1] \to [0,1]$ be defined by $\varphi_0(x) := x$ (note $d\lambda = \varphi_0\, d\mu$) so that

$$\lambda(\mathcal{C}_n^u) = \int_u^1 \hat{F}^{n-1}\varphi_0\, d\mu.$$

By [51], $\hat{F}$ maps the set of functions $\{f \in C^2(0,1) : f' > 0, f'' \leq 0\}$ into itself, and thus it suffices to show that

$$\int_u^1 \hat{F} f\, d\mu < \int_u^1 f\, d\mu$$

whenever $f \in L^1(\mu)$ is increasing. This follows from

$$\int_u^1 f\, d\mu - \int_u^1 \hat{F} f\, d\mu = \int_u^1 f\, d\mu - \int_{F^{-1}[u,1]} f\, d\mu = \int_u^1 f\, d\mu \int_{u/(1+u)}^{-1/(1+u)} f\, d\mu$$
$$= \int_{1/(1+u)}^1 f\, d\mu - \int_{u/(1+u)}^{u} f\, d\mu \geq \left(f\left(\frac{1}{1+u}\right) - f(u)\right) log(1+u) > 0.$$

Proof of (b) we first consider the case where $u = 1/N$, with $N \in \mathbb{N}$ and $N \geq 2$. Define the function $a : \mathbb{R} \to \mathbb{R}$ by

$$a(\sigma) := \frac{1}{\log N} \sum_{k=0}^{\lfloor\sigma\rfloor} \lambda(\mathcal{C}_{k+1}^u),$$

which is the μ-average of the function $\hat{F}_\sigma\varphi_0 := \sum_{k=0}^{\lfloor\sigma\rfloor} \hat{F}^k\varphi_0$ on $\mathcal{C}_1^u = [1/N, 1]$ by

$$\frac{1}{\mu(\mathcal{C}_1^u)} \int_{\mathcal{C}_1^u} \sum_{k=0}^{\lfloor\sigma\rfloor} \hat{F}^k \varphi_0 d\mu = \frac{1}{\log N} \sum_{k=0}^{\lfloor\sigma\rfloor} \int_{F^{-k}(\mathcal{C}_1^u)} \varphi_0 d\mu = \frac{1}{\log N} \sum_{k=0}^{\lfloor\sigma\rfloor} \lambda(\mathcal{C}_{k+1}^u).$$

We have the following bound on the difference of $\hat{F}_\sigma\varphi_0$ and $a(\sigma)$.

Lemma (4.2.3)[192]: For all $\sigma \in \mathbb{R}$ and $\in \mathcal{C}_1^u$,

$$\left|\hat{F}_\sigma\varphi_0(x) - a(\sigma)\right| \leq \frac{N(N-1)}{2}.$$

Proof. Without loss of generality, we assume that $\sigma = n \in \mathbb{N} \cup \{0\}$. By [51], we know that $\hat{F}_n\varphi_0$ is increasing. So the difference between $\hat{F}_n\varphi_0(x)$ and $a(n)$ is at most $\hat{F}_n\varphi_0(1) - \hat{F}_n\varphi_0(1/N)$. Using the equality

$$\hat{F}^k\varphi_0\left(\frac{1}{j}\right) = \frac{j}{j-1} \hat{F}^{k+1}\varphi_0\left(\frac{1}{j-1}\right) - \frac{1}{j-1} \hat{F}^k\varphi_0\left(\frac{j-1}{j}\right) \; (j \in \mathbb{N}, j \geq 2)$$

following from (4), and the fact that $\hat{F}^k\varphi_0$ is increasing for each $k \in \mathbb{N} \cup \{0\}$, we have

$$\hat{F}_n\varphi_0(1) - \hat{F}_n\varphi_0\left(\frac{1}{N}\right) = \sum_{k=0}^{n} \left(\hat{F}^k\varphi_0(1) - \hat{F}^k\varphi_0\left(\frac{1}{N}\right)\right)$$

$$= \sum_{k=0}^{n} (\hat{F}^k\varphi_0(1) - \frac{N}{N-1}\hat{F}^{k+1}\varphi_0\left(\frac{1}{N-1}\right)$$

$$+ \frac{1}{N-1}\hat{F}^k\varphi_0\left(\frac{N-1}{N}\right)$$

$$\leq \frac{N}{N-1}\sum_{k=0}^{n} \left(\hat{F}^k\varphi_0(1) - \hat{F}^{k+1}\varphi_0\left(\frac{1}{N-1}\right)\right).$$

Using this inequality recursively, and also the equality $\hat{F}f(1) = f(1/2)$, yields

$$\hat{F}_n\varphi_0(1) - \hat{F}_n\varphi_0\left(\frac{1}{N}\right) \leq \frac{N}{2}\sum_{k=0}^{n} \left(\hat{F}^k\varphi_0(1) - \hat{F}^{k+N-2}\varphi_0\left(\frac{1}{2}\right)\right)$$

$$= \frac{N}{2}\sum_{k=0}^{n} \left(\hat{F}^k\varphi_0(1) - \hat{F}^{k+N-1}\varphi_0(1)\right) \leq \frac{N(N-1)}{2}.$$

Next, we let $S: (0,\infty) \to \mathbb{R}$ be the Laplace transform of a given by

$$S(\sigma) := \int_{0-}^{\infty} e^{-t/\sigma}\, da(t) = \frac{1}{\log N}\sum_{n=0}^{\infty} e^{-n/\sigma}\lambda(\mathcal{C}_{n+1}^u)$$

and prove the following bound similar to Lemma (4.2.3).

Lemma (4.2.4)[192]: For all $x \in \mathcal{C}_1^u$ and all $\sigma > 0$,

$$\left|\sum_{n=0}^{\infty} e^{-n/\sigma}\hat{F}^n\varphi_0(x) - S(\sigma)\right| \leq \frac{N(N-1)}{2}.$$

Proof. We first note the equality

$$\sum_{n=0}^{\infty} a_n e^{-n/\sigma} = \left(1 - e^{-1/\sigma}\right)\sum_{n=0}^{\infty} e^{-n/\sigma}\left(\sum_{k=0}^{n} a_k\right), \qquad (5)$$

which holds for all sequences (a_n) satisfying $\sum_{k=0}^{n} a_k = O(n)$ as $n \to \infty$ and all $\sigma > 0$. Let $x \in \mathcal{C}_1^u, \delta_n(x) := \hat{F}_n\varphi_0(x) - a(n)$, and $\sigma > 0$. Using (5) twice, we have

$$\sum_{n=0}^{\infty} e^{-n/\sigma}\hat{F}^n\varphi_0(x) = \left(1 - e^{-1/\sigma}\right)\sum_{n=0}^{\infty} e^{-n\sigma}\hat{F}_n\varphi_0(x)$$

$$= \left(1 - e^{-1/\sigma}\right)\sum_{n=0}^{\infty} e^{-n/\sigma}(a(n) + \delta_n(x))$$

$$= S(\sigma) + \left(1 - e^{-1/\sigma}\right)\sum_{n=0}^{\infty} e^{-n/\sigma}\delta_n(x).$$

Since $|\delta_n(x)| \leq N(N-1)/2$ for all $n \geq 0$, we have

$$\left|\left(1 - e^{-1/\sigma}\right)\sum_{n=0}^{\infty} e^{-n/\sigma}\delta_n(x)\right| \leq \left(1 - e^{-1/\sigma}\right)\sum_{n=0}^{\infty} e^{-n\sigma}\,\frac{N(N-1)}{2} = \frac{N(N-1)}{2}.$$

To continue the proof, we will make use of the following equality given by [45].

$$\int_A \left(\sum_{n=0}^{\infty} e^{-n\sigma}\hat{F}^n f\right)(1-e^{-\varphi_A/\sigma})\,d\mu = \sum_{n=0}^{\infty} e^{-n/\sigma}\int_{A_n} f\,d\mu \qquad (6)$$

Here, f is any function in $L^1(\mu)$, σ is any positive real number, $A \subseteq [0,1]$ is any subset such that $\mu(A) < \infty, A_0 := A$, and $A_n := F^{-n}A \setminus \bigcup_{k=0}^{n-1} F^{-k}A$ for $n \geq 1$. Also, $\varphi_A : A \to \mathbb{N}$ is the return time function on A defined by $\varphi_A(x) := min\{n \in \mathbb{N} : F^n(x) \in A\}$. Letting $A = \mathcal{C}_1^u$ and $f = \varphi_0$ in (6), and noting that $A_n = \left[\frac{1}{n+N}, \frac{1}{n+N-1}\right)$ for $n \geq 1$, we have

$$\int_{1N}^1 \left(\sum_{n=0}^{\infty} e^{-n/\sigma}\hat{F}^n\varphi_0\right)(1-e^{-\varphi_A/\sigma})d\mu = \frac{N-1}{N} + \sum_{n=1}^{\infty} e^{-n/\sigma}\int_{1/(n+N)}^{1/(n+N-1)} \varphi_0\,d\mu$$

$$= \frac{N-1}{N} + \sum_{n=1}^{\infty} \frac{e^{-n/\sigma}}{(n+N)(n+N-1)}.$$

On the other hand, using Lemma (4.2.4), we see that the left side of the above is also equal to

$$\left(S(\sigma)+O_N(1)\right)\left(1-e^{-1/\sigma}\right)\int_{1N}^1 \left(\sum_{n=0}^{\infty} e^{-n/\sigma}\hat{F}^n 1\right)\left(1-e^{-\varphi_A/\sigma}\right)d\mu$$

$$= \left(S(\sigma)+O_N(1)\right)\left(1-e^{-1/\sigma}\right)\left(\mu(\mathcal{C}_1^u) + \sum_{n=1}^{\infty} e^{-n\sigma}\int_{1/(n+N)}^{1/(n+N-1)} d\mu\right)$$

$$= \left(S(\sigma)+O_N(1)\right)\left(1-e^{-1/\sigma}\right)\left(log\,N + \sum_{n=1}^{\infty} e^{-n/\sigma}\,log\left(\frac{n+N}{n+N-1}\right)\right)$$

as $\sigma \to \infty$. (Note that (6) holds for the constant function $f = 1$ in spite of the fact that $1 \notin L^1(\mu)$ since $\sum_{n=0}^{\infty} e^{-n/\sigma}\hat{F}^n 1$ has finite integral over $\mathcal{C}_1^u$.) For our next step, we determine the asymptotic behavior of

$$\frac{N-1}{N} + \sum_{n=1}^{\infty} \frac{e^{-n/\sigma}}{(n+N)(n+N-1)} \text{ and } \log N + \sum_{n=1}^{\infty} e^{-n/\sigma}\log\left(\frac{n+N}{n+N-1}\right).$$

Lemma (4.2.5)[192]:

$$\frac{N-1}{N} + \sum_{n=1}^{\infty} \frac{e^{-n/\sigma}}{(n+N)(n+N-1)} = 1 + O\left(\frac{\log\,\sigma}{\sigma}\right). \quad (\sigma \to \infty)$$

Proof. Let $S_1 : \mathbb{R} \to \mathbb{R}$ be defined by

$$S_1(t) := \frac{N-1}{N}\,1_{[0,\infty)}(t) + \sum_{n=1}^{\lfloor t \rfloor} \frac{1}{(n+N)(n+N-1)} = \begin{cases} 1 - \dfrac{1}{\lfloor t \rfloor + N} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0. \end{cases}$$

Then for $\sigma > 0$,

$$\frac{N-1}{N} + \sum_{n=1}^{\infty} \frac{e^{-n/\sigma}}{(n+N)(n+N-1)} = \int_{0-}^{\infty} e^{-t/\sigma}\,dS_1(t)$$

$$= \frac{1}{\sigma}\int_0^{\infty} \left(1 - \frac{1}{\lfloor t \rfloor + N}\right)e^{-t/\sigma}\,dt = 1 - \int_0^{\infty} \frac{e^{-x}dx}{\lfloor \sigma x \rfloor + N}.$$

Since the inequality $\lfloor t \rfloor + N \geq \frac{1}{2}(t+2)$ holds for $t \geq 0$, we have

$$\int_0^\infty \frac{e^{-x}dx}{\lfloor\sigma x\rfloor + N} \le 2\int_0^\infty \frac{e^{-x}d}{\sigma x + 2} \le \int_0^1 \frac{2dx}{\sigma x + 2} + \int_1^\infty \frac{2e^{-x}}{\sigma} dx \ll \frac{\log \sigma}{\sigma}. \quad (\sigma \to \infty)$$

Lemma (4.2.6)[192]: We have

$$\log N + \sum_{n=1}^{\infty} e^{-n/\sigma} \log\left(\frac{n + N}{n + N - 1}\right) = \log(\sigma + N) + C + O_N\left(\frac{\log \sigma}{\sigma}\right), (\sigma \to \infty)$$

Where

$$C := \int_0^1 \frac{e^{-x} - 1}{x} dx + \int_1^\infty \frac{e^{-x}}{x} dx.$$

Proof. Let $S_2 : \mathbb{R} \to \mathbb{R}$ be defined by

$S_2(t) := (log\, N)1_{[0,\infty)}(t) + \sum_{n=1}^{\lfloor t\rfloor} \log\left(\frac{n+N}{n+N-1}\right) = \begin{cases} log(t+N) & \text{if } t \ge 0 \\ 0 & \text{if } t < 0. \end{cases}$ Then for $\sigma > 0$,

$$log\, N + \sum_{n=1}^{\infty} e^{-n/\sigma} \log\left(\frac{n + N}{n + N - 1}\right) = \int_0^\infty - e^{-t\sigma} dS_2(t)$$

$$= \frac{1}{\sigma}\int_0^\infty e^{-t\sigma} log(\lfloor t\rfloor + N)\, dt = \int_0^\infty e^{-x} log(\lfloor\sigma x\rfloor + N)\, dx$$

$$= \int_0^\infty e^{-x} log(\lfloor\sigma x\rfloor + N)\, dx - \int_0^\infty e^{-x} \log\left(\frac{\sigma x + N}{\lfloor\sigma x\rfloor + N}\right) dx.$$

Using the inequality $log(1 + x) \le x$, we have

$$\int_0^\infty e^{-x} \log\left(\frac{\sigma x + N}{\lfloor\sigma x\rfloor + N}\right) dx = \int_0^\infty e^{-x} \log\left(1 + \frac{\{\sigma x\}}{\sigma x + N}\right) dx \ll \int_0^\infty \frac{e^{-x}dx}{\lfloor\sigma x\rfloor + N},$$

which is $O(\sigma^{-1} log\, \sigma)$ as $\sigma \to \infty$ by the proof of Lemma (4.2.5).

Next, integration by parts yields

$$\int_0^\infty e^{-x} log(\sigma x + N)dx = log\, N + \int_0^\infty \frac{\sigma e^{-x}dx}{\sigma x + N}. \tag{7}$$

To continue, we consider the integral on the right over [0, 1] by writing

$$\int_0^1 \frac{\sigma e^{-x}dx}{\sigma x + N} = \int_0^1 \frac{\sigma dx}{\sigma x + N} + \int_0^1 \frac{\sigma(e^{-x} - 1)}{\sigma x + N} dx.$$

The first integral on the right equals $log(\sigma + N) - log\, N$, while the second equals

$$\int_0^1 \frac{e^{-x} - 1}{x} dx - N\int_0^1 \frac{e^{-x} - 1}{x(\sigma x + N)} dx = \int_0^1 \frac{e^{-x} - 1}{x} dx + O\left(\int_0^1 \frac{N\, dx}{\sigma x + N}\right)$$

$$= \int_0^1 \frac{e^{-x} - 1}{x} dx + O_N\left(\frac{\log \sigma}{\sigma}\right). \quad (\sigma \to \infty)$$

Now considering the integral in (7) over [1,∞), we write

$$\int_1^\infty \frac{\sigma e^{-x}dx}{\sigma x + N} = \int_1^\infty \frac{e^{-x}}{x} dx - N\int_1^\infty \frac{e^{-x}dx}{x(\sigma x + N)} = \int_1^\infty \frac{e^{-x}}{x} dx + O_N\left(\frac{1}{\sigma}\right).$$

Putting these results together proves the lemma.

Lemmas (4.2.5) and (4.2.6) and the equalities preceding them gives

$$\big(S(\sigma) + O_N(1)\big)\big(1 - e^{-1/\sigma}\big)\left(log(\sigma + N) + C + O_N\left(\frac{\log \sigma}{\sigma}\right)\right)$$

$$= 1 + O\left(\frac{\log \ \sigma}{\sigma}\right), \qquad (\sigma \ \to \ \infty)$$

and as a result,

$$S(\sigma) = \frac{\sigma}{log\ \sigma \ + \ C} + \ O_N(1). \qquad (\sigma \ \to \ \infty) \qquad (8)$$

At this point, an application of Karamata's Tauberian theorem [197] then yields

$$\sum_{k=0}^{n} \ \lambda(\mathcal{C}_{k+1}^{u}) \sim \frac{n}{\log_N \ n}. \qquad (n \ \to \ \infty)$$

Furthermore, one can apply an adaptation of Freud's effective version of Karamata's theorem [194] (see also [200]) accommodating logarithms to (8) in order to prove

$$\sum_{k=0}^{n} \ \lambda(\mathcal{C}_{k+1}^{u}) = \frac{n}{\log \ Nn}\left(1 + O_N\left(\frac{1}{\log \ \log \ n}\right)\right). \qquad (n \ \to \ \infty)$$

To obtain an error term of $O_N(1/\ log\ n)$, we evaluate the equality

$$\big(S(\sigma) + O_N(1)\big)\big(1 - e^{-1/\sigma}\big)\left(log\ N + \sum_{n=1}^{\infty} \ e^{-n\sigma} \log\left(\frac{n \ + \ N}{n \ + \ N \ - \ 1}\right)\right) = 1 + O\left(\frac{\log \ \sigma}{\sigma}\right)$$

more precisely. Instead of directly establishing an asymptotic equality for $S(\sigma)$, we divide by $1 \ - \ e^{-1/\sigma}$ and multiply the series expression for $S(\sigma)$ together with the other series on the left side. Together with Lemma (4.2.6), this process yields

$$\frac{1}{\log \ N}\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \ \lambda(\mathcal{C}_{k+1}^{u})\ell_N(n-k)\right)e^{-n/\sigma} = \sigma\left(1 + O_N\left(\frac{\log \ \sigma}{\sigma}\right)\right) \quad ,(\sigma \ \to \infty)$$

where $\ell_N\ (0) :=\ log\ N$ and $\ell_N\ (n) :=\ log\ \left(\frac{n+N}{n+N-1}\right)$ for $n\ >\ 0$. Now a direct application of Freud's effective Tauberian theorem yields

$$\sum_{k=0}^{n}\ \sum_{j=0}^{k}\ \lambda\big(\mathcal{C}_{j+1}^{u}\big)\ell_N(k-j) = n\ log\ N\left(1 + O_N\left(\frac{1}{\log \ n}\right)\right). \quad (n \ \to \ \infty)$$

The left side of this expression is equal to

$$\sum_{k=0}^{n}\left(\lambda(\mathcal{C}_{k+1}^{u})log\ N + \sum_{j=0}^{k-1}\ \lambda\big(\mathcal{C}_{j+1}^{u}\big)\ell_N(k\ -\ j)\right)$$

$$= log\ N\sum_{k=0}^{n}\ \lambda(\mathcal{C}_{k+1}^{u}) + \sum_{j=0}^{n-1}\ \lambda\big(\mathcal{C}_{j+1}^{u}\big)\sum_{k=j+1}^{n}\ \ell_N(k\ -\ j)$$

$$= \log N\sum_{k=0}^{n}\ \lambda(\mathcal{C}_{k+1}^{u}) + \sum_{j=0}^{n-1}\ \lambda\big(\mathcal{C}_{j+1}^{u}\big)\log\left(\frac{n\ -\ j\ +\ N}{N}\right)$$

$$= \sum_{k=0}^{n}\ \lambda(\mathcal{C}_{k+1}^{u})\ log(n\ -\ k\ +\ N),$$

where the second equality follows from the definition of ℓ_N and telescoping. We can rewrite the last expression above as

$$log(n + N)\sum_{k=0}^{n} \lambda(\mathcal{C}_{k+1}^{u}) + \sum_{k=1}^{n} \lambda(\mathcal{C}_{k+1}^{u})\log\left(1 - \frac{k}{n + N}\right).$$

So if we can show that

$$\sum_{k=1}^{n} \lambda(\mathcal{C}_{k+1}^{u})\log\left(1 - \frac{k}{n + N}\right) = O\left(\frac{n}{\log n}\right), \quad (n \to \infty) \ (9)$$

then we have

$$\sum_{k=0}^{n} \lambda(\mathcal{C}_{k+1}^{u}) = \frac{n}{\log_N n}\left(1 + O_N\left(\frac{1}{\log n}\right)\right).$$

Since individual terms on the left side of (9) decay to 0 as $n \to \infty$, we can consider the sum starting from $k = 3$. We have

$$\left|\sum_{k=3}^{n} \lambda(\mathcal{C}_{k+1}^{u})\log\left(1 - \frac{k}{n + N}\right)\right| = \sum_{k=3}^{n} \lambda(\mathcal{C}_{k+1}^{u})\sum_{j=1}^{\infty} \frac{1}{j}\left(\frac{k}{n + N}\right)^{j}$$

$$= \sum_{j=1}^{\infty} \frac{1}{j(n + N)^{j}} \sum_{k=3}^{n} k^{j}\lambda(\mathcal{C}_{k+1}^{u})$$

$$\ll \sum_{j=1}^{\infty} \frac{1}{j(n + N)^{j}}\sum_{k=3}^{n} \frac{k^{j}}{\log k} \ll \sum_{j=1}^{\infty} \frac{1}{j(n + N)^{j}} \int_{3}^{n+1} \frac{x^{j}dx}{\log x}$$

$$\ll \sum_{j=1}^{\infty} \frac{1}{j(n + N)^{j}}\left(\frac{(n+1)^{j+1}}{(j + 1)log(n + 1)}\right)$$

$$\ll \frac{n}{\log n}\sum_{j=1}^{\infty} \frac{1}{j(j + 1)}\left(\frac{n + 1}{n + N}\right)^{j} \ll \frac{n}{\log n}. \qquad (n \to \infty)$$

This proves Theorem (4.2.2) in the case that $u = 1/N$.

For the general case $u \in (0,1)$, let $N = \lceil 1/u \rceil$ so that $[u,1] \subseteq [1/N,1]$. Then for $x \in [1/N,1]$, we have

$$\sum_{k=0}^{n} \hat{F}^{k}\varphi_0(x) = \frac{1}{\log N}\sum_{k=0}^{n} \lambda(\mathcal{C}_{k+1}^{1/N}) + O_N(1) = \frac{n}{\log n}\left(1 + O_N\left(\frac{1}{\log n}\right)\right). \quad (n \to \infty)$$

Integrating the first and last expressions over $[u, 1]$ yields

$$\sum_{k=0}^{n} \lambda(\mathcal{C}_{k+1}^{u}) = \frac{n\log (1/u)}{\log n}\left(1 + O_N\left(\frac{1}{\log n}\right)\right), (n \to \infty)$$

completing the proof.

Chapter 5
Self-Affine Sets and Self-Conformal Sets with Self-Similarity

We show that the set of points of A which have a unique address has positive Hausdorff dimension for all (β_1, β_2). We investigate simultaneous (β_1, β_2)-expansions of reals, which were the initial motivation for studying this family in Gunturk We show when restricting to self-conformal subsets of the real line with Hausdorff dimension strictly less than one, that the weak separation condition is equivalent to Ahlfors regularity and its failure implies full Assouad dimension. In fact, we resolve a self-conformal extension of the dimension drop conjecture for self-conformal sets with positive Hausdorff measure by showing that its Hausdorff dimension falls below the expected value if and only if there are exact overlaps. We show that the "self-similar" construction described by BBI spaces ensures the equivalence of positive Lebesgue measure and nonempty interior. We apply this result to self-conformal sets satisfying the WSC and prove that positive Lebesgue measure implies nonempty interior for such sets. This generalizes Zerner's corresponding result for self-similar sets.

Section (5.1): Topology with Uniqueness and Simultaneous Expansions

Let $T_i(x, y) = ((x + i)/\beta_1, (y + i)/\beta_2)$ for $i = \pm 1$ and $A := A_{\beta_1,\beta_2}$ be the attractor of the iterated function system (IFS) $\{T_{-1}, T_1\}$, i.e. the unique compact set satisfying $A = T_1(A) \cup T_{-1}(A)$. It is well known that A is either connected or totally disconnected [62]. Figures suggest that when β_1 and β_2 are 'sufficiently small', A_{β_1,β_2} is connected, and if, in addition, they are 'very small indeed', then A_{β_1,β_2} has a non-empty interior—see Figure 1. The main purpose is to make such statements quantifiable, thus expanding results from [203], [215].

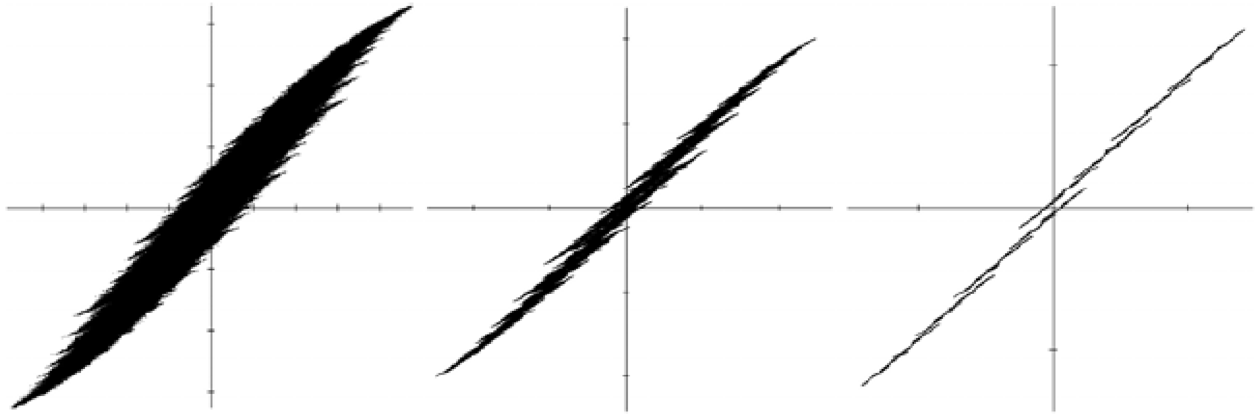


Figure (1)[201]: $A_{1.2;1.3}$; $A_{1.4;1.5}$ and $A_{1.7;1.8}$.

Clearly, if $\beta_1 = \beta_2$ then this set is either a Cantor set if $\beta_1 = \beta_2 > 2$ or a onedimensional segment otherwise. Hence, the set is trivial. So without loss of generality we will assume that $\beta_1 \neq \beta_2$.

For ease of notation, we will let $\lambda = 1/\beta_1$ and $\mu = 1/\beta_2$. Some solutions and discussions are simplified using λ and μ, and some with β_1 and β_2. As such, we will use them interchangeably.

We will denote−1 by m (for 'minus') and+1 by p. A word $w \in \{p, m\}^n$ is asequence of p and m of length n. The set $\{p, m\}^*$ will be the set of all finite words, and $\{p, m\}^{\mathbb{N}}$ the set of all infinite words. For $w = w_1 w_2 \dots w_n \in \{p, m\}^*$, we will denote by T_w the map $T_{w_1} T_{w_2} \dots T_{w_n}$. If $u, w \in \{p, m\}^*$, we will denote by uw the concatenation of u followed by w. By uw^∞ we will mean the infinite word $uwwww \dots$. We will use $\tilde{}$ for negation.
That is, $\tilde{p} = m, \tilde{m} = p$ and $\tilde{w} = \widetilde{w_1}\,\widetilde{w_2} \dots$.

We will define the map $s_\lambda : \{p,m\}^{\mathbb{N}} \to \mathbb{R}$ as $s_\lambda(w) = \sum_{i=1}^{\infty} w_i\lambda^i = \sum_{i=1}^{\infty} w_i/\beta_1^i$. We will define the map $\pi : \{p,m\}^{\mathbb{N}} \to \mathbb{R}^2$ as $\pi(w) = (s_\lambda(w), s_\mu(w))$. Thus, in this notation,

$$A_{\beta_1,\beta_2} = \{\pi(w): w \in \{p,m\}^{\mathbb{N}}\}.$$

For a point $(x,y) \in A_{\beta_1,\beta_2}$ we will say it has address $w \in \{p,m\}^{\mathbb{N}}$ if $\pi(w) = (x,y)$. It should be noted that a point (x,y) may not have a unique address.

We begin our study by considering the set

$$Z = \{(\beta_1,\beta_2): (0,0) \in A^o\},$$

where A^o is the interior of A. In a slightly different language, Z has been studied by Dajani, Jiang and Kempton, who showed the following result.

Theorem (5.1.1)[201]: [203] If $1 < \beta_1, \beta_2 < 1.05$, then $(\beta_1,\beta_2) \in Z$.

We improve this result to obtain the following theorem.

Theorem (5.1.2)[201]: If $\beta_1 \neq \beta_2$ are such that

$$\left|\frac{\beta_2^8 - \beta_1^8}{\beta_2^7 - \beta_1^7}\right| + \left|\frac{\beta_2^7\beta_1^7(\beta_2-\beta_1)}{\beta_2^7 - \beta_1^7}\right| \leq 2,$$

then $(\beta_1,\beta_2) \in Z$.

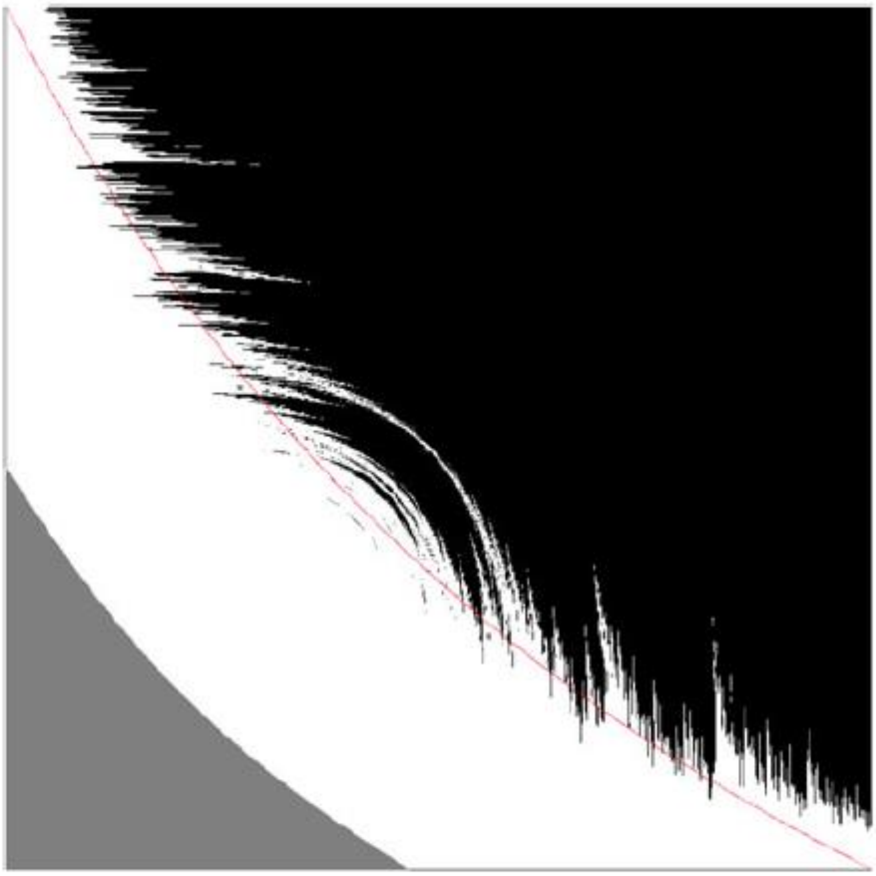


Figure (2)[201]: Points known to be in Z (grey); points known to be not in Z (black); curve $\beta_1\beta_2 = 2$.

As a consequence, we have the following corollary.

Corollary (5.1.3)[201]: If $1 < \beta_1, \beta_2 < 1.202$ then $(\beta_1,\beta_2) \in Z$.

We can also, in some cases, computationally check if $(\beta_1,\beta_2) \in Z$ and if $(\beta_1,\beta_2) \,/\!\in Z$. Many cases unfortunately remain unknown. These are shown in Figure 2. Those points provably in Z, coming from Theorem (5.1.2), are shown in grey. Those points provably not in Z, as discussed in Lemma (5.1.12), are shown in black. Note that all points above the curve $\beta_1\beta_2 = 2$ are not in Z either. These results will be discussed.

The question 'Is $(0,0) \in A^o$?' can be easily extended to higher dimensions. Namely, Let

$$T_i(x_1, \ldots, x_m) = \left(\frac{x_1 + i}{\beta_1}, \ldots, \frac{x_m + i}{\beta_m}\right) \quad i \in \{\pm 1\}.$$

Let $A_{\beta_1,\ldots,\beta_m}$ denote the attractor of this IFS, and put

$$Z_m = \{(\beta_1, \ldots, \beta_m) : (0, 0, \ldots, 0) \in A^o_{\beta_1,\ldots,\beta_m}\}.$$

We show in Theorem (5.1.4) that Z_m is always non-empty, as first conjectured in [206].

Theorem (5.1.4)[201]: For each $m \geq 2$ there exists a $C_m > 1$ such that if $1 < \beta_1 < \cdots < \beta_m < C_m$, then the attractor $A_{\beta_1,\ldots,\beta_m}$ contains a neighbourhood of $(0, \ldots, 0)$.

In the previous study, we bounded those β_1, β_2 such that there is a neighborhood of $(0,0)$ contained in A. We observe that if $(0,0) \in A$ by $\pi(w) = (0,0)$, then $\pi(\tilde{w}) = (0,0)$, where, as above, $\tilde{w}$ is the negation of w. In particular, $(0,0)$ does not have a unique address under π.

For the next question, we examine the other end of this spectrum, namely, for fixed β_1 and β_2, which points $(x,y) \in A$ have a unique address $(x,y) = \pi(w)$. We say that $(x,y) = \pi(w)$ has a unique address if for any $w' \in \{p,m\}^{\mathbb{N}}$ with $w \neq w'$ we have $\pi(w') \neq (x,y)$. We denote by U_{β_1,β_2} the set of all unique addresses and by U_{β_1,β_2} the projection $\pi(U_{\beta_1,\beta_2})$, which we call the set of uniqueness.

For example, if A_{β_1,β_2} is totally disconnected, then $U_{\beta_1,\beta_2} = \{p,m\}^{\mathbb{N}}$ and $U_{\beta_1,\beta_2} = A_{\beta_1,\beta_2}$. On the other hand, if $(\beta_1,\beta_2) \in Z$, then $U_{\beta_1,\beta_2} \subsetneq \{p,m\}^{\mathbb{N}}$ and $U_{\beta_1,\beta_2} \subsetneq A_{\beta_1,\beta_2}$.

In the self-similar setting (without rotations), the set of uniqueness has been studied in detail see, e.g. [205], [210] for the one-dimensional case and [214] for higher dimensions. In particular, it is proved in [214] that if the contraction ratios are sufficiently close to 1, then the set of uniqueness can contain only fixed points. As we will see, this is very different in the self-affine setting.

We show in Lemma (5.1.13) that for $\beta_1 \neq \beta_2$ the set of uniqueness is non-empty. Furthermore, the set U_{β_1,β_2} has positive topological entropy (Theorem (5.1.14)), and U_{β_1,β_2} has positive Hausdorff dimension (Corollary (5.1.15)) and no interior points (Proposition (5.1.16)) for all β_1, β_2. We also give sufficient conditions (albeit not provably necessary) for a point in U_{β_1,β_2} to be on the boundary of A_{β_1,β_2} (Proposition (5.1.17)).

Put

$$D_{\beta_1,\beta_2} = \left\{x \in \mathbb{R} : \exists (a_n) \in \{\pm 1\}^{\mathbb{N}} \mid x = \sum_{n=1}^{\infty} a_n \beta_1^{-n} = \sum_{n=1}^{\infty} a_n \beta_2^{-n}\right\}.$$

In other words,

$$D_{\beta_1,\beta_2} = A_{\beta_1,\beta_2} \cap \{(x,y) : y = x\}$$

(see Figure 6). Studying this set was the original motivation behind the IFS under consideration see [206], [203]. We prove the following result.

When studying iterated function systems, a common property that is investigated is if A satisfies the open set condition.

Definition (5.1.5)[201]: Let A be the unique compact set such that $A = F_1(A) \cup \cdots \cup F_k(A)$, where the F_i are linear contractions. We say that A satisfies the open set condition (OSC) if there exists a non-empty open set O such that:

- $F_i(O) \subset O$ for all i; and
- $F_i(O) \cap F_j(O) = \emptyset$ for all $i \neq j$.

An even stronger property is that of a set being totally disconnected. Definition. We say that a set A is totally disconnected if for all $x, y \in A, x \neq y$, there exist open sets O_x and O_y such that:

- $x \in O_x$;
- $y \in O_y$;
- $O_x \cap O_y = \emptyset$; and
- $A \subset O_x \cup O_y$.

A set is disconnected if there exist x and y with the above property. It is clear that if a set is totally disconnected then it is disconnected. It is known for this case that $A := A_{\beta_1,\beta_2}$ is either connected or totally disconnected [62]. Hence, in this case the converse is also true. That is, if A is disconnected, then it must be totally disconnected.

Put

$$O = \{(\beta_1, \beta_2) : \{T_{-1}, T_1\} satisfies\ the\ OSC\},$$
$$S = \{(\beta_1, \beta_2) : A_{\beta_1,\beta_2}\ is\ totally\ disconnected\}.$$

It is easy to see that $S \subset O$. Furthermore, if $\beta_1 > 2\ or\ \beta_2 > 2$, then the projection of A onto the x-(respectively, y-) axis is a Cantorset, whence $(\beta_1, \beta_2) \in S$. Hence forth we will assume $\beta_1 < 2$ and $\beta_2 < 2$.

In Theorem (5.1.24), we give a precise description of a curve S_1 such that if (β_1, β_2) is above this curve, then $(\beta_1, \beta_2) \in S$. As a corollary to this theorem, we get the following result.

We can also, in some cases, computationally check if $(\beta_1, \beta_2) \in S$ and if $(\beta_1, \beta_2) / \in S$. Many cases remain unknown. The first are shown in Figure 3. Those points provably in S are shown in black. These results will be discussed we show that S is disconnected. There are a number of obvious—and some not so obvious—relations between some of these sets.

Define

$$I = \{(\beta_1, \beta_2) : the\ attractor\ A\ has\ a\ non-empty\ interior\}.$$

It is clear that $Z \subset I$. It is also clear that $Z \cap S = \emptyset$. We know very little about I, although it seems likely that $I \cap O = \emptyset$. It is not clear if $Z \subsetneq I$, or if, in fact, they are equal sets. It is true that $S \subsetneq O$, as demonstrated by the points $(\beta_1^{(n)}, \beta_2^{(n)})$ from Theorem (5.1.24), which are all points in O but not in S. All of these points $(\beta_1^{(n)}, \beta_2^{(n)})$ are points on the boundary of O, as shown by Solomyak [215].

Figure (3)[201]: Points known to be in S (black). (Level 40 approximation.)

An interesting observation to make is that there are points that are not in Z yet at the same time are not in O either.

For example, let $\beta_1 \approx 1.190842710$ and $\beta_2 \approx 1.769542577$ be roots of $x^{11} - x^{10} - x^9 - x^8 + x^6 - x^5 + x^4 + x^3 + x^2 + x + 1$. We see by Lemma7.1 that $(\beta_1, \beta_2) \, / \in O$. As $\beta_1\beta_2 = 2.107246878 > 2$, the Lebesgue measure of A is 0, and hence $(\beta_1, \beta_2) \, / \in Z$.

As a second example, let $\beta_1 \approx 1.122195284$ and $\beta_2 \approx 1.776995700$ be roots of $x^{13} - x^{12} - x^{11} - x^9 - x^8 + x^7 - x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$. Again, by Lemma (5.1.27), $(\beta_1, \beta_2) \, / \in O$. Since $\beta_1\beta_2 = 1.994136194 < 2$, the Lebesgue measure argument does not work here. However, we can, applying techniques discussed in §3.3, show that $(\beta_1, \beta_2) \, / \in Z$ (using a level 25 approximation).

This indicates that there is actually more structure here that is not fully explored.

Before beginning our study of properties of $A = A_{\beta_1,\beta_2}$, we will first introduce and study K, the convex hull of A. The structure of K will play an important role in later investigations, from both a computational and a theoretical point of view.

We first give a precise description of those points that are vertices of K. See, for example, Figure 4.

Theorem (5.1.6)[201]: The vertices of K have addresses $p^k m^\infty$ and $m^k p^\infty$ for $k = 0, 1, 2, \ldots$.

Proof. Without loss of generality, we may assume that $\beta 2 < \beta 1$. It suffices to show that the line segments connecting $\pi(p^k m^\infty)$ and $\pi(p^{k+1} m^\infty)$ lie below A. We will denote this line

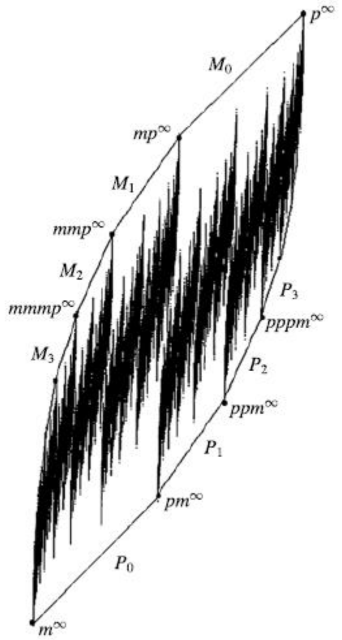


Figure (4)[201]: $A_{1.85,1.25}$ together with vertices and edges of K.

segment by P_k. Let us begin at $k = 0$. We must show that for any $w \in \{p, m\}^{\mathbb{N}}$, the line from $\pi(m^\infty)$ to $\pi(w)$ lies above the straight line passing through $\pi(m^\infty)$ and $\pi(pm^\infty)$.

We notice that the line P0 from π(m∞) to π(pm∞) is in the direction $\pi(pm^\infty) - \pi(m^\infty) = \left(\frac{1}{\beta_1} - \sum_{i\geq 2} \beta_1^{-i}, \frac{1}{\beta_2} - \sum_{i\geq 2} \beta_2^{-i}\right)$

$$-\left(-\frac{1}{\beta_1} - \sum_{i\geq 2} \beta_1^{-i}, \frac{1}{\beta_2} - \sum_{i\geq 2} \beta_2^{-i}\right)$$
$$= \left(\frac{2}{\beta_1}, \frac{2}{\beta_2}\right).$$

This will have slope $s_1 = \beta_1/\beta_2$.

Consider, now, the line from $\pi(m^\infty)$ to $\pi(w)$ for $w \in \{p, m\}^{\mathbb{N}}$ where w is not equal to m^∞ and not equal to pm^∞.

$$\pi(w) - \pi(m^\infty) = \left(\sum_{i\geq 1} (a_i + 1)\beta_1^{-i}, \sum_{i\geq 1} (a_i + 1)\beta_2^{-i}\right).$$

This will have slope $s_2 = (\sum_{i\geq 1} (a_i + 1)\beta_2^{-i})/(\sum_{i\geq 1} (a_i + 1)\beta_1^{-i})$.

It is obvious that $\pi(w)$ lies to the right of $\pi(m^\infty)$. Hence, to show that $\pi(w)$ lies above the line P_0, it suffices to show that $s_2 > s_1$.

This will be true if and only if

$$\sum_{i\geq 2} (a_i + 1)\beta_2^{-i+1} > \sum_{i\geq 2} (a_i + 1)\beta_1^{-i+1}. \quad (1)$$

We see that the $a_i + 1$ terms are either 0 or 2 (and hence always non-negative). Further, $\beta_2 < \beta_1$ by assumption, and hence $\beta_2^{-i+1} > \beta_1^{-i+1}$ for all $i \geq 2$. From this, the result follows. We know that we only get equality if $a_i + 1 = 0$ for all $a_i \geq 2$. This cannot happen, as $w \neq m^\infty$ and $w \neq pm^\infty$.

We now proceed by induction. Consider the line P_k from $\pi(p^k m^\infty)$ to $\pi(p^{k+1} m^\infty)$. This is in the direction

$$\pi(p^{k+1}m^\infty) - \pi(p^k m^\infty) = (2/\beta_1^{k+1}, 2/\beta_2^{k+1}).$$

This will have slope $s_1 = \beta_1^{k+1} / \beta_2^{k+1}$. In particular, notice that these slopes are increasing as k increases (as $\beta_1/\beta_2 > 1$).

Consider a word $\pi(w)$ not equal to either $\pi(p^k m^\infty)$ or $\pi(p^{k+1}m^\infty)$. We may assume without loss of generality that $\pi(w)$ lies to the right of $\pi(p^k m^\infty)$. (If not, then there will exist some $k' < k$ such that w lies to the right of $\pi(p^{k'} m^\infty)$ and to the left of $\pi(p^{k'+1}m^\infty)$. By induction, w will be above the line P_k'. As the slopes are increasing, we will have that $\pi(w)$ is above the line P_k.)

Consider the direction from $p^k m^\infty$ to w. As before, we have that

$$\pi(w) - \pi(p^k m^\infty) = (\sum_{i=1}^{k} (a_i - 1)\beta_2^{-i} + \sum_{i\geq k+1} (a_i + 1)\beta_1^{-i},$$
$$\sum_{i=1}^{k} (a_i - 1)\beta_2^{-i} + \sum_{i\geq k+1} (a_i + 1)\beta_2^{-i}).$$

This will have slope

$$s_2 = \frac{\sum_{i=1}^{k} (a_i - 1)\beta_2^{-i} + \sum_{i\geq k+1} (a_i + 1)\beta_2^{-i}}{\sum_{i=1}^{k} (a_i - 1)\beta_1^{-i} + \sum_{i\geq k+1} (a_i + 1)\beta_1^{-i}}.$$

We have that $s_2 > s_1$ if and only if

$$\sum_{i=1}^{k} (a_i - 1)\beta_2^{-i+k+1} + \sum_{i\geq k+1} (a_i + 1)\beta_2^{-i+k+1}$$
$$> \sum_{i=1}^{k} (a_i - 1)\beta_1^{-i+k+1} + \sum_{i\geq k+1} (a_i + 1)\beta_1^{-i+k+1}. \quad (2)$$

In the first sum we see that $a_i - 1$ is always 0 or −2, and $\beta_2^{-i+k+1} < \beta_1^{-i+k+1}$. Hence, the first sum of the left-hand side is always greater than or equal to that of the right-hand side. For the second sum, we see that $a_i + 1$ is always 0 or 2, and $\beta_2^{-i+k+1} > \beta_1^{-i+k+1}$. Hence, the second sum of the left-hand side is always greater than or equal to that of the right-hand side. We also see that we only get equality if $w = p^k m^\infty$ *or* $w = p^{k+1}m^\infty$.

The points $\pi(m^k p^\infty)$ are treated in a similar way.

We notice that the proof shows something stronger, as stated in the following corollary.

Corollary (5.1.7)[201]: The vertices of K have unique addresses.

Proof. To see this, we note that equations (1) and (2) are strict inequalities when $w \neq p^k m^\infty$.

Recall that for a finite word $w \in \{p, m\}^*$, we define $K_w = T_w(K)$ and set $K_n = \bigcup_{|w|=n} K_w$. It is easy to see that for $w, w' \in \{p, m\}^*$ we have $K_{ww'} \subset K_w$. In particular, this shows that

$$A \subset \cdots \subset K_{n+1} \subset K_n \subset \cdots \subset K.$$

A standard result on iterated function systems gives that $A = \bigcap_{n\geq 1} K_n$.

We will take advantage of this construction in multiple ways. For example, we will show that:

(a) if $(0,0) \notin K_n$ for some $n \geq 1$, then $(0,0) \notin A$;

(b) if $T_1(K_n) \cap T_{-1}(K_n) = \emptyset$ for some $n \geq 1$, then A is totally disconnected; and

(c) if $T_1(K_n^o) \cap T_{-1}(K_n^o) = \emptyset$ for some $n \geq 1$, then A satisfies the OSC.

We will investigate Z in greater detail. We will provide the main tool for checking if a point is in Z and provide a proof of Theorem (5.1.2), giving sufficient conditions for $(\beta_1, \beta_2) \in Z$.

We will discuss the higher dimensional analogue of Z. In §3.3, we will give sufficient conditions for $(\beta_1, \beta_2) \,/\!\in Z$.

The main tool used to computationally check if a point $(\beta_1, \beta_2) \in Z$ and to find a generic bound for points in Z is a generalization and strengthening of [203].

Using this theorem, it suffices to find a polynomial P in terms of $\beta_1, \ldots, \beta_m$ such that the four conditions hold for all $1 < \beta_j < C$, for some C. This is a purely computational search.

Consider the polynomial

$$P(x) = x^8 - \frac{\beta_2^8 - \beta_1^8}{\beta_2^7 - \beta_1^7} x^7 + \frac{\beta_2^7 \beta_1^7 (\beta 2 - \beta 1)}{\beta_2^7 - \beta_1^7}.$$

A quick check shows that $P(\beta_1) = P(\beta_2) = 0$. Further, for all $\beta_1, \beta_2 < 1.202$, we have

$$\left|\frac{\beta_2^8 - \beta_1^8}{\beta_2^7 - \beta_1^7}\right| + \left|\frac{\beta_2^7 \beta_1^7 (\beta_2 - \beta_1)}{\beta_2^7 - \beta_1^7}\right| \le 2.$$

In fact, a stronger result can be shown. By explicitly solving for

$$\left|\frac{\beta_2^8 - \beta_1^8}{\beta_2^7 - \beta_1^7}\right| + \left|\frac{\beta_2^7 \beta_1^7 (\beta_2 - \beta_1)}{\beta_2^7 - \beta_1^7}\right| \le 2,$$

we find that all $\beta_1 \neq \beta_2$ in grey in Figure 2 have the desired properties.

Theorem (5.1.8)[201]: Let $P(x) = x^n + b_{n-1} x_{n-1} + \cdots + b_0$ such that:

(i) $P(\beta_j) = 0 \; for \; j = 1, 2, \ldots, m$;

(ii) $\sum_{j=0}^{n-1} |b_j| \le 2$;

(iii) $b_1 = b_2 = \cdots = b_{m-1} = 0$; and

(iv) $b_0 \neq 0$.

Then there exists a neighbourhood of $(0, \ldots, 0)$ in A, based on $\beta_1, \ldots, \beta_m$.

Proof. Let P have the required properties.

Let $u_{-n}, \ldots, u_{-n+m-1}$ satisfy

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = b_0 \begin{bmatrix} \beta_1^{-1} & \beta_1^{-2} & \cdots & \beta_1^{-m} \\ \beta_2^{-1} & \beta_2^{-2} & \cdots & \beta_2^{-m} \\ \vdots & \vdots & & \vdots \\ \beta_m^{-1} & \beta_m^{-2} & \cdots & \beta_m^{-m} \end{bmatrix} \begin{bmatrix} u_{-n} \\ u_{-n+1} \\ \vdots \\ u_{-n+m-1} \end{bmatrix}.$$

We see that this system will have a solution as all of the β_i are distinct. Moreover, we see that if the x_j are sufficiently close to 0, then the u_j will also be sufficiently close to 0. Choose δ such that $if |x_j| < \delta$, then $|u_j| \le 1$.

Set $u_{-n+m} = \cdots = u_0 = 0$. We will choose the u_i and a_i for $i = 1, 2, 3, \ldots$ by induction, such that

$$u_i := a_i - \left(\sum_{k=0}^{n-1} b_k u_{i-n+k} \right)$$

and such that $u_i \in [-1, 1]$ and $a_i \in \{-1, +1\}$. We see that this is possible, as, by induction, $|u_j| \le 1$ for all $j \le i - 1$. Furthermore,

$$\left| \sum_{k=0}^{n-1} b_k u_{i-n+k} \right| \le \sum_{k=0}^{n-1} |b_k u_{i-n+j}| \le \sum_{k=0}^{n-1} |b_k| \le 2,$$

by our assumption on the b_k. Hence, there is a choice of a_i, either +1 or −1, such that $a_i - \sum_{k=0}^{n-1} b_k u_{i-n+k} \in [-1, 1]$.

We claim that this sequence of a_i has the desired properties.

Let $b_n = 1$ for ease of notation. To see this, notice for $i = 1, 2$ that

$$\sum_{j\geq 1} a_j\beta_i^{-j} = \sum_{j\geq 1}\left(\left(\sum_{k=0}^{n-1} b_k u_{j-n+k}\right) + u_j\right)\beta_i^{-j} = \sum_{j\geq 1}\sum_{k=0}^{n} b_k u_{j-n+k}\beta_i^{-j}$$

$$= \sum_{k=0}^{n}\sum_{j\geq 1} b_k u_{j-n+k}\beta_i^{-j} = \sum_{k=0}^{n} b_k\beta_i^{k}\sum_{j\geq 1} u_{j-n+k}\beta_i^{-j-k}$$

$$= \beta_i^{-n}\sum_{k=0}^{n} b_k\beta_i^{k}\sum_{j\geq 1} u_{j-n+k}\beta_i^{-j-k+n}$$

$$= \beta_i^{-n}\sum_{k=0}^{n} b_k\beta_i^{k}\sum_{\ell\geq -n+1} u_{\ell+k}\beta_i^{-\ell-k}$$

$$= \beta_i^{-n}\sum_{k=0}^{n} b_k\beta_i^{k}\left(\sum_{\ell=-n+1}^{-k} u_{\ell+k}\beta_i^{-\ell-k} + \sum_{\ell\geq 1} u_\ell\beta_i^{-\ell}\right)$$

$$= \left(\beta_i^{-n}\sum_{k=0}^{n} b_k\beta_i^{k}\sum_{\ell=-n+1}^{-k} u_{\ell+k}\beta_i^{-\ell-k}\right) + \left(\beta_i^{-n}P(\beta_i)\sum_{\ell\geq 1} u_\ell\beta_i^{-\ell}\right)$$

$$= \beta_i^{-n}\sum_{k=0}^{n}\sum_{\ell=-n+1}^{-k} b_k\beta_i^{k} u_{\ell+k}\beta_i^{-\ell-k}.$$

Thus, by our construction, we have $b_1 = b_2 = \cdots = b_{m-1} = 0$ and $u_{m-n} = \cdots = 0$. Hence, this simplifies to

$$\sum_{j\geq 1} a_j\beta_i^{-j} = \beta_i^{-n}\sum_{\ell=-n+1}^{0} b_0 u_\ell\beta_i^{-\ell} + \beta_i^{-n}\sum_{k=m}^{n}\sum_{\ell=-n+1}^{-k} b_k\beta_i^{k} u_{\ell+k}\beta_i^{-\ell-k}$$

$$= \beta_i^{-n}\sum_{\ell=-n+1}^{0} b_0 u_\ell\beta_i^{-\ell} + \beta_i^{-n}\sum_{k=m}^{n}\sum_{\ell=-n+1}^{-k} b_k\beta_i^{k}.0.\beta_i^{-\ell-k}$$

$$= b_0(u_{-n+1}\beta_i^{-1} + u_{-n+2}\beta_i^{-2} + \cdots + u_{-n+m+1}\beta_i^{-m}) = x_i,$$

which gives the desired result.

We see from Theorem (5.1.8) that to prove Theorem (5.1.4), it suffices to find P satisfying certain criteria. We will show that such a polynomial exists for all m.

Lemma (5.1.9)[201]: Let $P(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be such that $\sum_{i=0}^{n-1} |a_i| < 2$ and $P(\beta_i) = 0 \, for \, i = 1, 2, \ldots, m$. Let $S \subset \{0, 1, \ldots, n-1\}$ be such that $|S| < n - m$. Then there exists a neighbourhood of $(\beta_1, \ldots, \beta_m)$ such that for all $(\widehat{\beta_1}, \ldots, \widehat{\beta_m})$ in this neighbourhood there exists a polynomial $\hat{P}(x) = x^n + \hat{a}_{n-1}\hat{x}^{n-1} + \cdots + \hat{a}_0$ where:

- $a_s = \hat{a}_s$ for all $s \in S$;
- $\sum_{i=0}^{n-1} |\hat{a}_i| < 2$;
- $\hat{P}(\hat{\beta}_i) = 0 \, for \, i = 1, 2, \ldots, m$.

Proof. Let R be such that $P(x) = \prod(x - \beta_i)R(x)$. For $\widehat{\beta_i}$ close to β_i, we see that the coefficients of $\tilde{P}(x) = \prod(x - \tilde{\beta}_i)R(x) = x^n + \tilde{a}_{n-1}x^{n-1} + \cdots + \tilde{a}_0$ are close to those of P. For all $s \in S$, let $T_s(x) = b_{n-1}^{(s)}x^{n-1} + \cdots + b_0^{(s)}$ be a polynomial such that:

- $b_s^{(s)'} = 0 \, for \, s' \in S, s' \neq s$;
- $b_s^{(s)} = 1$; and

• $T_s(\hat{\beta}_i) = 0 \; for \; i = 1, 2, \ldots, m$.

We see that such a polynomial exists as $n - |S| > m$. Set

$$\hat{P}(x) = \tilde{P}(x) + \sum_{s \in S} (a_s - \tilde{a}_s) T_s(x).$$

It is easy to observe that $a_s = \hat{a}_s$ for s∈S, and that $\hat{P}(\hat{\beta}_i) = 0 for \; i = 1, 2, \ldots, m$. Further observe that for $\hat{\beta}_i$ close to β_i we have that $\hat{a}_i$ are close to a_i. Hence, by continuity, we can choose a neighbourhood of $(\beta_1, \ldots, \beta_m)$ such that the resultingb $\hat{a}_i$ are close enough to a_i so that $\sum |\hat{a}_i| < 2$. We see that $\hat{P}$ has the desired properties.

Corollary (5.1.10)[201]: If there exists $a \; P \in R[x]$ monic of degree at least $2m - 1$ such that $a_1 = \cdots = a_{m-1} = 0, \sum |a_i| < 2$ and $(x - 1)^m | P$, then there is a neighbourhood around (1, 1, . . . , 1) that is contained in Z.

Proof. We use $S = \{1, 2, \ldots, m\}$ and the neighbourhood of (1, 1, . . . , 1). If $a_0 = 0$, then we can use the polynomial T_0 to perturb P.

Theorem (5.1.11)[201]: Given $m \in \mathbb{N}$, there exists an $n \in \mathbb{N}$ and a polynomial $P(x) = x^{mn+1} - x^{nm} + b_{m-1} x^{(m-1)n} + b_{m-2} x^{(m-2)n} + \cdots + b_0$ such that $(x - 1)^m | P$ and $1 + \sum_{i=0}^{m-1} |b_i| < 2$.

Proof. Let

$$P(x) = x^{mn+1} - x^{nm} + b_{m-1} x^{(m-1)n} + \cdots + b_1 x^n + b_0.$$

We see that $(x - 1)^m | P$ if and only if $P(i) = P'(i) = \cdots = P^{(m-1)}(i) = 0$. Using the notation $n^{(k)} = n(n-1)(n-2) \ldots (n-k+1)$ with $n^{(k)} = 0$ for $k > n$, consider the kth derivative of P with respect to x for $k \geq 1$:

$$P^{(k)}(x) = (nm+1)^{(k)} x^{nm+1-k} - (nm)^{(k)} x^{nm-k}$$
$$+ \big(n(m-1)\big)^{(k)} b_{m-1} x^{n(m-1)-k} + \cdots + n^{(k)} b_1 x^{n-k}.$$

We require that $P^{(k)}(i) = 0$ for $k = 0, 1, \ldots, m-1$. Evaluating $P(x) \; at \; x = 1$ gives

$$1 - 1 = b_{m-1} + b_{m-2} + \cdots + b_0. \tag{3}$$

For $k = 1, \ldots, m-1$, by dividing by $(nm)^{(k)}$ and evaluating at $x = 1$ we have

$$1 - \frac{(nm+1)^{(k)}}{(nm)^{(k)}} = \frac{(nm-1)^{(k)}}{(nm)^{(k)}} b_{m-1} + \frac{(nm-2)^{(k)}}{(nm)^{(k)}} b_{m-2} + \cdots + \frac{n^{(k)}}{(nm)^{(k)}} b_1. \tag{4}$$

Taking the limit as n tends to infinity in (4), we obtain

$$0 = \left(\frac{m-1}{m}\right)^k b_{m-1} + \left(\frac{m-2}{m}\right)^k b_{m-2} + \cdots + \left(\frac{0}{m}\right)^k b_0 \tag{5}$$

for $k = 0, 1, \ldots, m-1$. Here we take $\left(\frac{0}{m}\right)^0 = 1$. Clearly, solving (5) for the b_i is equivalent to solving the linear system

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ \frac{m-1}{m} & \frac{m-2}{m} & \cdots & \frac{1}{m} & 0 \\ \left(\frac{m-1}{m}\right)^2 & \left(\frac{m-2}{m}\right)^2 & \cdots & \left(\frac{1}{m}\right)^2 & 0 \\ \vdots & & & & \vdots \\ \left(\frac{m-1}{m}\right)^{m-1} & \left(\frac{m-2}{m}\right)^{m-1} & \cdots & \left(\frac{1}{m}\right)^{m-1} & 0 \end{bmatrix} \begin{bmatrix} b_{m-1} \\ b_{m-2} \\ \vdots \\ b_1 \\ b_0 \end{bmatrix}.$$

The lower left $(m-1) \times (m-1)$ submatrix is the Vandermonde matrix on the terms $(m-1)/m, (m-2)/m, \dots, 1/m$, with non-zero determinant $\prod_{1 \le i < j \le m-1} ((i-j)/m)$. Hence, there exists an N such that for all $n \ge N$, the system of equations given by (3) and (4) has non-zero determinant, and hence will always have a solution, regardless of the left hand side.

We see that the system of equations given by (5) has a solution of $b_i = 0$ for $i = 0, 1, \dots, m-1$. We see in this case that the sum $\sum_{i=0}^{m} |b_m| = 1$. (Here, we think of $b_m = -1$ coming from the coefficient of x^{nm}.)

This implies that there exists an $N_0 > N$ such that for all $n \ge N_0$, the solution to equations (3) and (4) will have solutions $b_0 \approx b_1 \approx \cdots \approx b_{m-1} \approx 0$ and $b_{m \approx 1}$, and $\sum_{i=0}^{m} |b_i| \approx 1$. This gives a polynomial with the desired property and proves Theorem (5.1.4).

To prove that $(\beta_1, \beta_2) /\notin Z$, it suffices to show that $(0,0) \notin A$. This is clearly a sufficient condition, although it is not a necessary condition. To see that it is not necessary, notice that the $(\beta_1^{(n)}, \beta_2^{(n)})$ which we will discuss have the property that $(0,0) \in A$, yet A satisfies the open set condition. Moreover, by approximating A by K, we see that there are points arbitrarily close to $(0,0)$ that are not in K, and hence not in A. As such, $(\beta_1^{(n)}, \beta_2^{(n)}) \notin Z$. See Figure 10.

It is interesting to note that $(\beta_1^{(n)}, \beta_2^{(n)})$ is on the boundary of S. It is not clear if such an example that is not on the boundary of S would exist.

Recall that we write $K_w = T_w(K) and\ K_n = \bigcup_{|w|=n} K_w$. The following result holds.

Lemma (5.1.12)[201]: If there exists an n such that $(0,0) \notin K_n$, then $(0,0) \notin A$ and $(\beta_1, \beta_2) \notin Z$.

It would be computationally expensive to compute the entirety of K_n. We observe for $w, w' \in \{p, m\}^*$ that $Kww' \subset K_w$. Hence, if $(0,0) \notin K_w$ then we have that $(0,0) \notin K_{ww'}$ for all w'. This allows for considerably more efficient computations.

In Figure 2 we give those points that are provably not in Z, as shown by examining K_{20}. We also give those points that are provably in Z by Theorem (5.1.2).

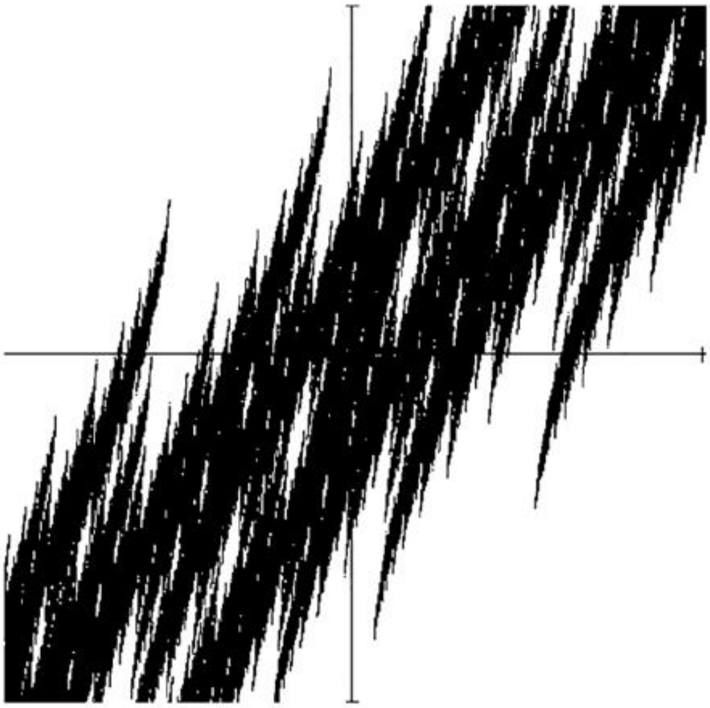


Figure (5)[201]: A_{β_1,β_2} zoomed in around (0, 0), where $\beta_1 \approx 1.57125, \beta_2 \approx 1.34067$ are roots of $x^{10} - x^9 - x^8 - x^7 + x^6 + x^5 - x^4 + x^3 + x^2 + x + 1$. We have $(0,0) \in A_{\beta_1,\beta_2}$, but no neighbourhood of (0, 0) lies in A.

Note also that if $\beta_1\beta_2 > 2$ then, as is well known, the Lebesgue measure of A is zero, and hence all (β_1, β_2) which satisfy this condition do not belong to Z either.

Example 3.7. Let $\beta_1 \approx 1.57125, \beta_2 \approx 1.34067$ be roots of$x^{10} - x^9 - x^8 - x^7 + x^6 + x^5 - x^4 + x^3 + x^2 + x + 1$. Then we have $\beta_1\beta_2 \approx 2.10653 > 2$, and hence $(\beta_1, \beta_2) \notin Z$. However, (0, 0) clearly belongs to A, as $(0,0) = \pi((pmmmppmpppp)^\infty)$, see Figure 5.

Observe that there is a large region of Figure 2 where nothing is known.

Recall that $(x, y) = \pi(w)$ has a unique address if for any $w' \in \{p, m\}^{\mathbb{N}}$ with $w \neq w'$, we have $\pi(w') \neq (x, y)$. We denote by U_{β_1,β_2} the set of all unique addresses and by $U_{\beta_1,\beta_2}$the projection $\pi(U_{\beta_1,\beta_2})$, which we call the set of uniqueness.

A consequence of Corollary (5.1.7) gives the following lemma.

Lemma (5.1.13)[201]: The set of uniqueness U_{β_1,β_2} is always non-empty.

Now we are ready to prove the main result. Let $E_n(\mathcal{L})$ be the number of $a_1a_2 \dots a_n$ that are prefixes for some infinite word in $\mathcal{L} \subset \{p, m\}^{\mathbb{N}}$. We say that $\mathcal{L}$ has positive topological entropy if $E_n(\mathcal{L})$ grows exponentially, that is, if $\liminf_{n\to\infty}(log\ E_n(\mathcal{L})/n) > 0$.

Theorem (5.1.14)[201]: For any (β_1, β_2), the set U_{β_1,β_2} has positive topological entropy.

Proof. Let $[i_1 \dots i_k]$ stand for the cylinder $\{a_j\}_{j=1}^{\infty} \subset \{p, m\}^{\mathbb{N}}$, where $a_j = i_j$ for $j = 1, \dots, k$. As $\pi(p^k m^\infty)$ has a unique address from Corollary (5.1.7), we get that $dist(\pi(p^k m^\infty), \pi([m]) > 0$, where dist stands for the Euclidean metric. Put

$$L_k = min\{j \geq 1: dist(\pi([p^k m^j]), \pi([m])) > 0\}$$

and $L = \max_{k\geq 1} L_k$. Note that since $\pi(p^k m^\infty)$ tends to $\pi(p^\infty)$ (which is clearly at a positive distance from $\pi([m])$), the quantity L is well defined.

Put

$$U' = \{p^{k_0} m^{k_1} p^{k_2} \cdots | k_0 \geq 1, k_i \geq L, i \geq 1\} \cup \{m^{k_0} p^{k_1} m^{k_2} \cdots | k_0 \geq 1, k_i \geq L, i \geq 1\}. \quad (6)$$

Clearly, U' is a subshift, i.e. a closed set such that if $a_1 a_2 \cdots \in U'$, then we have $a_j a_{j+1} a_{j+2} \cdots \in U'$ for any $j \geq 2$. The set U' also has positive topological entropy, since it contains the set $\prod_1^\infty \{m^L p^{L+1}, m^{L+1} p^L\}$, which has exponential growth. Thus, it suffices to show that any sequence in U_0 is a unique address.

By our construction, $\pi([p^k m^{k'}])$ does not intersect $\pi([m])$ provided $k' \geq L$. This is true for all k > 1. By symmetry, the same goes for $\pi([m^k p^{k'}])$ and $\pi([p])$. This means that for $(x, y) = \pi(w_0 w_1 w_2 \dots) = \pi(p^{k_0} m^{k_1} p^{k_2} \dots)$ with $k_i \geq L$, we necessarily have $w_0 = p$. Hence, the problem of showing that $(x, y) = \pi(p^{k_0} m^{k_1} p^{k_2} \dots)$ has a unique address reduces to showing that $(x', y') = \pi(p^{k_0 - 1} m^{k_1} p^{k_2} \dots)$ has a unique address. This argument is repeated by induction, proving the result.

Corollary (5.1.15)[201]: The set of uniqueness U_{β_1, β_2} has positive Hausdorff dimension for any (β_1, β_2).

Proof. Put $\pi = \pi | U'$. Since U_{β_1, β_2} is the set of unique addresses, the map π' is an injection. Also, it is Hölder continuous, since π is. Let us show that $(\pi')^{-1}: \pi(U') \to U'$ is Hölder continuous as well.

Suppose $\underline{a} = a_1 a_2 \dots and\ \underline{a'} = a'_1 a'_2 \dots$ with $a'_i = a_i, 1 \leq i \leq n-1$ and $a_n \neq a'_n$. If $n = 1$, then, by the above, there exists a constant $C > 0$ such that $dist(\pi(\underline{a}), \pi(\underline{a'})) \geq C$. Hence, for a general n, we have $dist(\pi(\underline{a}), \pi(\underline{a'})) \geq C \beta_1^{-n}$ (we assume, as always, $\beta_1 > \beta_2$). Since the distance between $\underline{a}$ and $\underline{a'}$ is 2^{-n}, we have

$$dist(\pi(\underline{a}), \pi(\underline{a'})) \geq C \cdot dist(a, a')^\kappa,$$

where $\kappa > 0$. Hence, $(\pi')^{-1}$ is Hölder continuous. The Hausdorff dimension on $\{p, m\}^{\mathbb{N}}$ in the usual metric coincides with the topological entropy, and hence the definition of Hausdorff dimension together with $(\pi')^{-1}$ being Hölder continuous immediately yields $\dim_H U_{\beta_1, \beta_2} \geq \dim_H \pi(U') > 0$.

Proposition (5.1.16)[201]: For all (β_1, β_2), the set U_{β_1, β_2} has no interior points.

Proof. We have two cases. Either A is totally disconnected, or $T_1(A) \cap T_{-1}(A) \neq \emptyset$. In the first case, the result is trivial.

Assume, therefore, that we are in the second case—i.e. that $T_1(A) \cap T_{-1}(A) \neq \emptyset$. Assume that $U = U_{\beta_1, \beta_2}$ has non-empty interior. In particular, let B be an open ball with $B \subset U \subset A$. Let $(x, y) \in T_1(A) \cap T_{-1}(A)$. We know that $A = cl(\bigcup_{k \geq 1} \bigcup_{j_1 \dots j_k} T_{j_1} \dots T_{jk}((x, y)))$, since A is the unique attractive fixed point of the iterated function system in the Hausdorff metric. This implies that there exist $j_1, j_2, \dots, j_k$ such that $T_{j_1} \dots T_{jk}((x, y)) \in B \subset U \subset A$. As $(x, y) \notin U$, we have $T_{j1} \dots T_{jk}((x, y)) \notin U$, a contradiction. This proves the desired result.

If the attractor has non-empty interior, we do not know whether the set of uniqueness can contain an interior point of A. We have, however, a partial result in this direction.

Proposition (5.1.17)[201]:

(i) If $(x, y) = \pi(wm^\infty)$ or $\pi(wp^\infty)$ is in the set of uniqueness, then $(x, y) \in \partial A_{\beta_1, \beta_2}$.

(ii) We have $\pi(U') \subset \partial A_{\beta_1, \beta_2}$, where U' is given by (6).

Proof. (i) Let $(x, y) = \pi(wm^\infty)$ (for $\pi(wp^\infty)$ the result will follow by symmetry). Let $w = a_1 \dots a_n$ and put $d_1 = dist(wm^\infty, \pi([\tilde{a}_1]))$ and $d_i = dist(wm^\infty, \pi([a_1 \dots a_{i-1} \tilde{a}_i]))$ for $2 \leq i \leq n$, where, as usual,e $\tilde{a} = -a$. Since $\pi(C)$ is compact for any cylinder C, we have $d = \min_{1 \leq i \leq n} d_i > 0$.

Now suppose $\varepsilon < d$. Then $(x, y - \varepsilon)$ is not in the attractor; indeed, if it were, then by our construction, its address would have to begin with $a_1 \dots a_n$. This would mean that to obtain $(x, y - \varepsilon)$, one or several of the subsequent -1 values in the address of (x, y) would have to be replaced with 1, which would only increase both coordinates. Therefore, there exist arbitrarily close points in the neighbourhood of (x, y) which are not in the attractor, i.e. (x, y) cannot be an interior point of A.

(ii) Put

$$d'_k = dist(\pi(p^k m^\infty), \pi([m])) = dist(\pi(m^k p^\infty), \pi([p])).$$

We know from the proof of Theorem (5.1.14) that $d' = \inf_{k \geq 1} d'_k > 0$, and the rest of the argument goes exactly like in (i), with $\varepsilon < d'$.

(ii) Simultaneous expansions

Put

$$D_{\beta_1,\beta_2} = \left\{ x \in \mathbb{R}: \exists (a_n)_1^\infty \in \{p, m\}^{\mathbb{N}} \mid x = \sum_{n=1}^{\infty} a_n \beta_1^{-n} = \sum_{n=1}^{\infty} a_n \beta_2^{-n} \right\}$$
$$= A_{\beta_1,\beta_2} \cap \{(x, y): y = x\}$$

(see Figure 6).

Theorem (5.1.18)[201]:

(i) For any pair (β_1, β_2), the set D_{β_1,β_2} is non-empty.

(ii) If $\min\{\beta_1, \beta_2\} < (1 + \sqrt{5})/2$, then the Hausdorff dimension of the set $D_{\beta_1,\beta_2} > 0$ is positive.

(iii) If $\max\{\beta_1, \beta_2\} < 1.202$, then there exists a $\delta > 0.664$ such that $[-\delta, \delta] \subset D_{\beta_1,\beta_2}$.

Proof. (i) Let $\lambda = \beta_1^{-1}$, $\mu = \beta_2^{-1}$ and assume $\lambda < \mu$. We first claim that for any $k \geq 0$ there exists a word $w \in \{p, m\}^k$ such that $\pi(wm^\infty)$ is below the diagonal (by which we always mean the straight line $y = x$) and $\pi(wp^\infty)$ is above it.

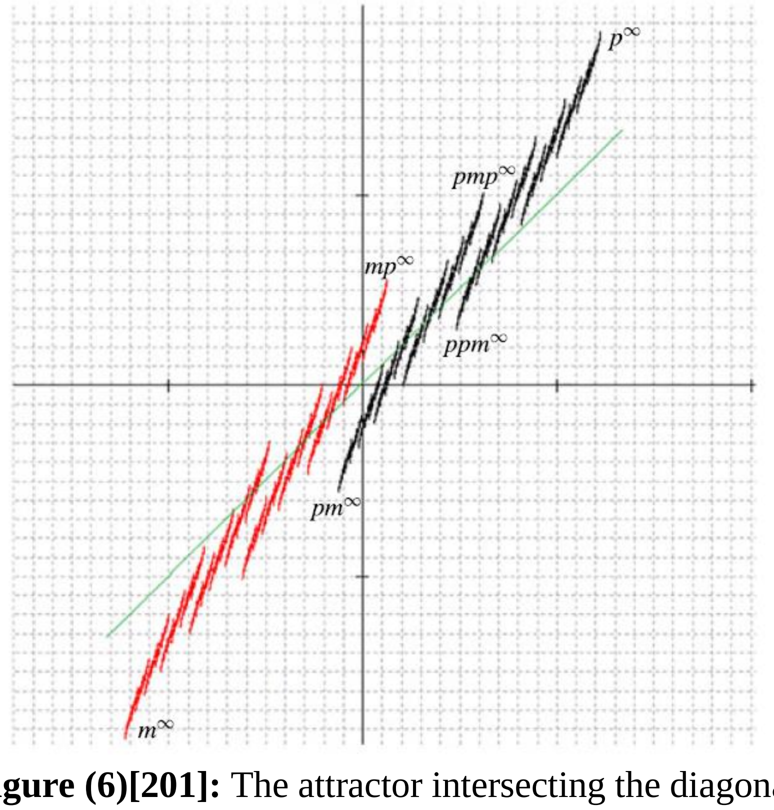


Figure (6)[201]: The attractor intersecting the diagonal for $\beta_1 = 1.923, \beta_2 = 1.754$

Proceed by induction ('bisection') and assume the claim is true for k=n and some w. We will show that it is then true for $w' = wp$ or wm (or both). We have

$$s_\lambda(wmp^\infty) = s_\lambda(w) - \lambda^{n+1} + \frac{\lambda^{n+2}}{1-\lambda} > s_\lambda(w) + \lambda^{n+1} - \frac{\lambda^{n+2}}{1-\lambda} = s_\lambda(wpm^\infty),$$
in view of $\lambda > 1/2$. Similarly, $s_\mu(wmp^\infty) > s_\mu(wpm^\infty)$. Consider the vector from $\pi(wpm^\infty)$ to $\pi(wmp^\infty)$ given by

$$\pi(wmp^\infty) - \pi(wpm^\infty) = 2\left(\lambda^{n+1} - \frac{\lambda^{n+2}}{1-\lambda}, \mu^{n+1} - \frac{\mu^{n+2}}{1-\mu}\right).$$

We see that this vector has slope

$$\left(\frac{\mu}{\lambda}\right)^{n+1} \cdot \frac{2\mu - 1}{1-\mu} \cdot \frac{1-\lambda}{2\lambda - 1} > 1,$$

since $\lambda < \mu$ and the function $x \mapsto (2x-1)/(1-x)$ is strictly increasing. Hence, it would be impossible for $\pi(wmp^\infty)$ to be below the diagonal and at the same time for $\pi(wpm^\infty)$ to lie above it. Now, if $\pi(wmp^\infty)$ is above the diagonal, then we put $w' = wm$; if $\pi(wpm^\infty)$ is below the diagonal, then we put $w' = wp$; and if both of these are true, we can choose either $w' = wm$ or $w' = wp$.

Thus, this allows us to construct a sequence of nested words $a_1 \dots a_n$ such that $\pi(a_1 a_2 \dots)$ lies on the diagonal.

Note, first, that $\pi(p^\infty) = (\lambda/(1-\lambda), \mu/(1-\mu))$, and since $\lambda < \mu$, we have that $\pi(p^\infty)$ lies above the diagonal. Similarly, $\pi(m^\infty)$ lies below it, see Figure 6.

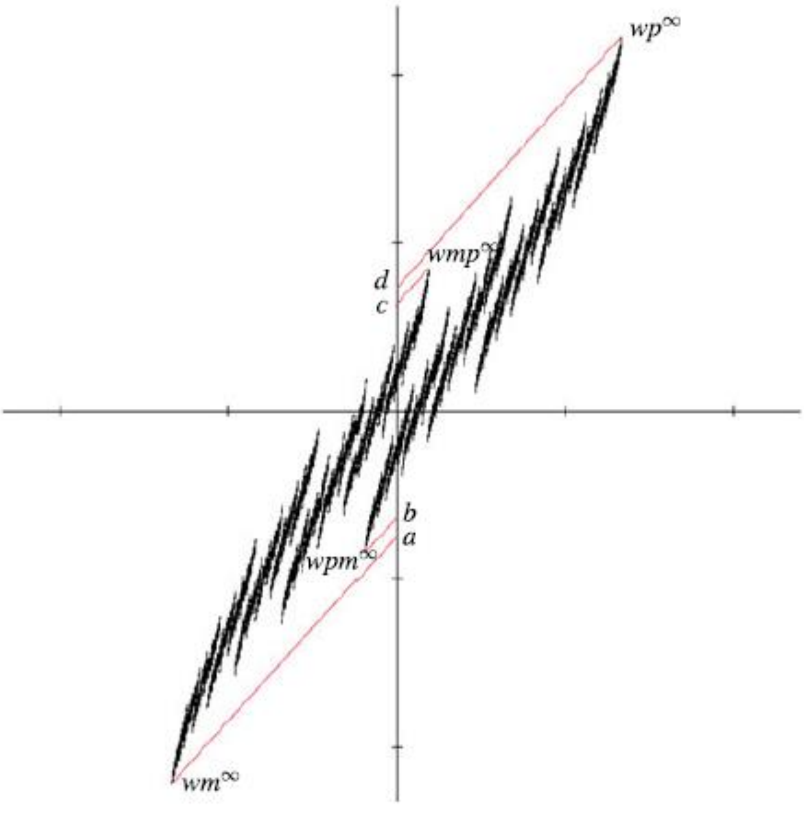


Figure (7)[201]: Projections for $\beta_1 = 1.75, \beta_2 = 1.45$.

(ii)Let us look at the bisection algorithm more closely in order to determine when we can actually choose both wm and wp as w'. Our aim is to construct a sequence of maps $\tau_n : [0,1] \to [0,1]$ which will keep track of all words w such that $\pi(wp^\infty)$ is above the diagonal and $\pi(wm^\infty)$ is below it. The map τn turns out to be the multi-valued βtransformation with $\beta = \beta^{(n)}$, which is well understood. Here we have that $\beta^{(n)} \uparrow \beta_2 < (\sqrt{5}+1)/2$. The condition $\beta_2 < (\sqrt{5}+1)/2$ implies that the number of such w grows exponentially with n, which yields the claim.

Let h denote the projection along the diagonal onto the y-axis, given by $h(x, y) = (0, y - x)$. Put $(0, a) = h(wm^\infty), (0, b) = h(wpm^\infty), (0, c) = h(wmp^\infty)$ and, finally, $(0, d) = h(wp^\infty)$, see Figure 7. Let n stand for the length of w. A straightforward computation yields that the second coordinates of these points are

$$a = s_\mu(w) - s_\lambda(w) - \mu^{n+1} - \frac{\mu^{n+2}}{1-\mu} + \lambda^{n+1} + \frac{\lambda^{n+2}}{1-\lambda},$$
$$b = s_\mu(w) - s_\lambda(w) + \mu^{n+1} - \frac{\mu^{n+2}}{1-\mu} - \lambda^{n+1} + \frac{\lambda^{n+2}}{1-\lambda},$$
$$c = s_\mu(w) - s_\lambda(w) - \mu^{n+1} + \frac{\mu^{n+2}}{1-\mu} + \lambda^{n+1} - \frac{\lambda^{n+2}}{1-\lambda},$$
$$d = s_\mu(w) - s_\lambda(w) + \mu^{n+1} + \frac{\mu^{n+2}}{1-\mu} - \lambda^{n+1} - \frac{\lambda^{n+2}}{1-\lambda}.$$

Since $1/2 < \lambda < \mu$, we have that $a < b < c < d$ provided n is large enough (which we may assume without loss of generality). Notice that $b - a = d - c$.

We see by assumption that $a < 0$ and $d > 0$. We see that $\pi(wmp^\infty)$ is above the diagonal if and only if $c > 0$. Hence, if $c > 0$ then we can take $w' = wm$, and if $b < 0$ then we can take $w' = wp$. If $b < 0 < c$, then both $w' = wm$ and $w' = wp$ are allowed inductive steps.

Now let ρ_w denote the following affine map:

$$\rho_w(t) = \frac{t - a}{d - a} = \frac{t - s_\mu(w) + s_\lambda(w) + \frac{\mu^{n+1}}{1-\mu} - \frac{\lambda^{n+1}}{1-\lambda}}{\frac{2\mu^{n+1}}{1-\mu} - \frac{2\lambda^{n+1}}{1-\lambda}}.$$

Put

$$\beta^{(n)} = \frac{\mu^{n+1}/(1-\mu) - \lambda^{n+1}/(1-\lambda)}{\mu^{n+2}/(1-\mu) - \lambda^{n+2}/(1-\lambda)} \uparrow \mu^{-1} = \beta_2\, n \to +\infty.$$

We have $\rho_w(a) = 0,\ \rho_w(d) = 1$ and

$$\rho_w(b) = \frac{(\mu^{n+1} - \lambda^{n+1})}{\mu^{n+1}/(1-\mu) - \lambda^{n+1}/(1-\lambda)} = 1 - 1/^{\beta(n)} < 1 - \mu,$$
$$\rho w(c) = \frac{\mu^{n+2}/(1-\mu) - \lambda^{n+2}/(1-\lambda)}{\mu^{n+1}/(1-\mu) - \lambda^{n+1}/(1-\lambda)} = 1/\beta^{(n)} > \mu.$$

Note that $\rho_w(0) \in [0, 1]$. We see that if $\rho_w(0) < \rho_w(c)$ then we can take $w' = wm$. We observe that

$$\rho_{wm}(t) = \frac{t - a'}{d' - a'}$$
$$= \frac{t - s_\mu(wm) + s_\lambda(wm) + \frac{\mu^{n+2}}{1-\mu} - \lambda^{n+2}}{2\mu^{n+2}/(1-\mu) - 2\lambda^{n+2}/(1-\lambda)}$$
$$= \frac{t - s_\mu(wm) + s_\lambda(wm) + \frac{\mu^{n+1}}{1-\mu} - \lambda^{n+2}}{2\mu^{n+2}/(1-\mu) - 2\lambda^{n+2}/(1-\lambda)}$$
$$= \beta^{(n)}\rho_w(t).$$

In a similar way, if $\rho_w(0) > \rho_w(b)$ then we can take $w' = wp$, and

$$\rho_{wp}(t) = \beta^{(n)}\rho_w(t) + 1 - \beta^{(n)}.$$

Thus, we have a sequence of finite sets $X_n = X_n(\beta_1, \beta_2)$ such that $X_n = \tau_n(X_{n-1})$, where τn is the following multi-valued map on[0, 1]:

$$\tau_n(x) = \begin{cases} \beta^{(n)}x & 0 \le x < 1 - 1/\beta^{(n)}, \\ \beta^{(n)}x \text{ and } \beta^{(n)}x + 1 - \beta^{(n)} & 1 - 1/\beta^{(n)} \le x \le 1/\beta^{(n)}, \\ \beta^{(n)}x + 1 - \beta^{(n)} & 1/\beta^{(n)} < x \le 1. \end{cases}$$

This is a well-knownβ-expansion-generating map (with $\beta = \beta^{(n)}$),see,e.g.[213].Since $\beta^{(n)} < \beta_2 < (1 + \sqrt{5})/2$, we have that for any $x_0 \in (0, 1 - 1/\beta^{(n)})$, there exists k such that $\tau_k \ldots \tau_1(x_0) \in (1 - 1/\beta^{(n)}, 1/\beta^{(n)})$, i.e. the trajectory of x_0 bifurcates after k steps. This is because $\tau_n(1 - 1/\beta^{(n)}) < 1/\beta^{(n)}$, in view of $(\beta^{(n)})2 < \beta^{(n)} + 1$. This proves that D_{β_1,β_2} has the cardinality of the continuum.

Furthermore, [204] implies that for the iterations of a single map τ_n with $\beta^{(n)} < (1 + \sqrt{5})/2$, we have that, no matter what $x_0 \in (0, 1)$, hitting the interval $(1 - 1/\beta^{(n)}, 1/\beta^{(n)})$ occurs with a positive (lower) asymptotic frequency. The argument for the sequence of maps$\{\tau_n\}$is exactly the same.

Let W_n denote the number of 0– 1 words w of length n such that $\pi(wm^\infty)$ is below the diagonal and $\pi(wp^\infty)$ is above it. We have just shown that W_n grows exponentially fast, which implies that the set $D_{\beta_1,\beta_2} \cap \{y = x\}$ has positive Hausdorff dimension (for the same reason as in the proof of Corollary (5.1.15)).

(iii)This follows from Theorem (5.1.2). Namely, consider in Theorem (5.1.8) the special case of simultaneous expansions, that is, where $x_1 = x_2$, with the polynomial

$$P(x) = x^8 - \frac{\beta_2^8 - \beta_1^8}{\beta_2^7 - \beta_1^7}x^7 + \frac{\beta_2^7\beta_1^7(\beta_2 - \beta_1)}{\beta_2^7 - \beta_1^7}.$$

We see that we require $|u_{-8}|, |u_{-7}| \le 1$. Solving for u_{-8} and u_{-7}, we have

$$|u_{-8}| = |x_1||b_0|(\beta_1 + \beta_2)$$

$$= |x_1| \frac{\beta_2^7\beta_1^7(\beta_2 + \beta_1)}{\beta_1^6 + \beta_1^5\beta_2 + \beta_1^4\beta_2^2 + \beta_1^3\beta_2^3 + \beta_1^2\beta_2^4 + \beta_1\beta_2^5 + \beta_2^6},$$

$$|u_{-7}| = |x_1||b_0|(\beta_1\beta_2)$$

$$= |x_1| \frac{\beta_2^8\beta_1^8}{\beta_1^6 + \beta_1^5\beta_2 + \beta_1^4\beta_2^2 + \beta_1^3\beta_2^3 + \beta_1^2\beta_2^4 + \beta_1\beta_2^5 + \beta_2^6}.$$

For $\beta_1, \beta_2 \le 1.202\ldots$, we see that both $|b_0|(\beta_1 + \beta_2)$ and $|b_0|\beta_1\beta_2$ are maximized when $\beta_1 = \beta_2 = 1.202\ldots$. This is in fact maximized for all β_1, β_2 where $|b_0| + |b_7| \le 2$ at the exact same value, although this is not needed for the desired result.

The maximum value that $|b_0|(\beta_2 + \beta_1)$ attains with this restriction is approximately 1.504520168. This shows that for all $|x_1| \le 1/1.504520168 \approx 0.6646637388$ we have $|u_{-7}| \le 1$.

The maximum value that $|b_0|\beta_2\beta_1$ attains with this restriction is approximately 0.9047548367. This shows that for all $|x_1| \le 1/0.9047548367 \approx 1.105271792$ we have $|u_{-8}| \le 1$.

Combining the two, for all $|x_1| \le 0.664$ we have $|u_{-7}|, |u_{-8}| \le 1$, and hence there exists a simultaneous expansion of (x_1, x_1).

We now focus our attention on the pairs (β_1, β_2) for which the IFS satisfies the open set condition (OSC) or is totally disconnected. We begin with a simple observation. Clearly, $T_i(K) \subset K$ for $i \in \{\pm 1\}$. Put $K_n = \bigcup_{|w|=n} T_w(K)$; then $K_{n+1} \subset K_n$, and $\bigcap_{n \ge 1} K_n = A$. Hence, A is disconnected if and only if there exists n such that K_n is disconnected. (And, therefore, so is K_k for all $k > n$.) This immediately yields the following proposition.

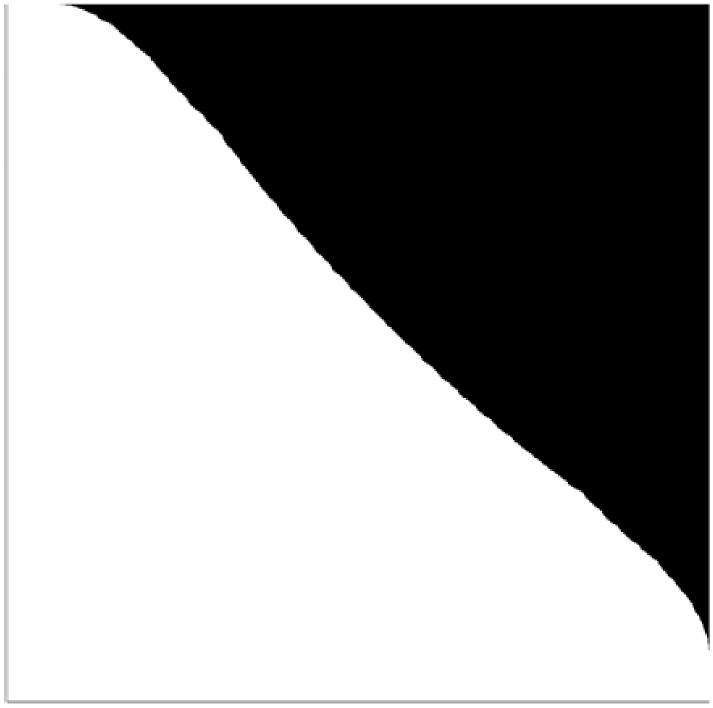


Figure (8)[201]: Points known to be in S (black). (Level 1 approximation.)

Proposition (5.1.19)[201]: The set S is open.

Proof. Let $(\beta_1, \beta_2) \in S$ and n be such that K_n is disconnected. By the continuity of T_{-1} and T_1, a sufficiently small perturbation of (β_1, β_2) leaves K_n disconnected, and hence A is disconnected as well.

For ease of discussion, if $T_1(K_n^o) \cap T_{-1}(K_n^o) = \emptyset$ then we will say that $T_1(K_n) \cap T_{-1}(K_n)$ has trivial intersection. Let A be the IFS in question and K the convex hull of A. We immediately see that a sufficient condition for A to satisfy the OSC, or to be totally disconnected, is if $T_1(K)$ and $T_{-1}(K)$ have trivial, orempty, in tersection. That is, we have the following lemma.

Lemma (5.1.20)[201]: Let K be the convex hull of A.

• If $T_1(K^o) \cap T_{-1}(K^o) = \emptyset$ then A satisfies the open set condition.

• If $T_1(K) \cap T_{-1}(K) = \emptyset$ then A is totally disconnected.

Here, K^o is the interior of K. Although these requirements are sufficient, they are not necessary. This is because K is a extreme overestimate for the shape of A.

In Figure 8 we have shown those (β_1, β_2) which satisfy the hypothesis of Lemma (5.1.20).

This curve is the same curve, after translation of notation, as that found by Solomyak [215] using somewhat different techniques. This will be shown in Theorem (5.1.26). A precise description of this curve is given in Theorem (5.1.24).

The idea of approximating A by a simple set K can be generalized. Recall for $w \in \{p, m\}^*$ that $K_w = T_w(K)$ and we define $K_n = \bigcup_{|w|=n} K_w$. An immediate, and profitable, generalization of Lemma (5.1.20) follows.

Lemma (5.1.21)[201]: Let K_n be as above.

(a) If $T_1(K_n^o) \cap T_{-1}(K_n^o) = \emptyset$ then A satisfies the open set condition.

(b) If $T_1(K_n) \cap T_{-1}(K_n) = \emptyset$ then A is totally disconnected.

This can be done for any set that contains A as a subset. An advantage of these K_n is that $K_n \to A_{\beta_1,\beta_1}$ in the Hausdorff metric.

In Figure 3 we have given the approximations of S based on K_{40}. We will call an approximation of S using Lemma (5.1.21) with a particular K_n a level n approximation.

In Theorem (5.1.6) we gave a precise description of the vertices of K. We can now determine for which β_1, β_2 we satisfy the conditions of Lemma (5.1.20) and, to some extent, 6.3.

Let M_k be the line connecting $m^k p^\infty$ and $m^{k+1}p^\infty$ and, similarly, P_k for $p^k m^\infty$ and $p^{k+1}m^\infty$ (see Figure 4).

Lemma (5.1.22)[201]: For each $\beta_1 > \beta_2$ there exists k such that the segment $T_1(M_k)$ crosses the y-axis.

It should be noted that this k may not be unique, as it is possible that $T_1(m^k p^\infty)$ is on the y-axis. In this case we would say that both $k-1$ and k satisfy this criterion.

Proof. We see that $\pi(pm^\infty)$ lies to the left of the y-axis and that $\pi(p^\infty)$ lies to the right. This, combined with the fact that the M_k form a decreasing (with respect to the y-coordinate) sequence of intervals, proves the result.

We will define $k := k(\beta_1, \beta_2)$.

Lemma (5.1.23)[201]: Assume $\beta_1 > \beta_2$ and let $k := k(\beta_1, \beta_2)$. Then:

(a) if $T_1(M_k)$ is below the point $(0,0)$, then $T_1(K) \cap T_{-1}(K) = \emptyset$;

(b) if $T_1(M_k)$ goes through the point $(0,0)$, then $T_1(K) \cap T_{-1}(K)$ has trivial, but nonempty intersection; and

(c) if $T_1(M_k)$ is above the point $(0,0)$, then $T_1(K) \cap T_{-1}(K)$ has non-trivial and nonempty intersection.

We see that the first case gives a sufficient condition for $(\beta_1, \beta_2) \in S$. Also, the first case combined with the second one gives criteria for when $(\beta_1, \beta_2) \in O$. Unfortunately the final case does not yield anything useful about (β_1, β_2); it only indicates that the level of approximation we are using is insufficient to come to a conclusion.

Proof. This follows from the symmetry of $T_1(K)$ *and* $T_{-1}(K)$ and the fact that $\beta_1 > \beta_2$. See, for example, Figure 9.

Using this, we can now give criteria for a point (β_1, β_2) to be in a level 1 approximation.

Define

$$S_1 = \{(\beta_1, \beta_2) | T_1(K) \cap T_{-1}(K) \; has\; trivial\; but\; non-empty\; intersection\}.$$

Theorem (5.1.24)[201]: Let $P_k(x) = x^{k+1} - 2x^k + 2$. Let $(\beta_1^{(k)}, \beta_2^{(k)})$ be the two roots of P_k between 1 and 2, with $\beta_1^{(k)} < \beta_2^{(k)}$.

(i) For $k \geq 4$, we have $(\beta_1^{(k)}, \beta_2^{(k)}), (\beta_2^{(k)}, \beta_1^{(k)}) \in S_1$.

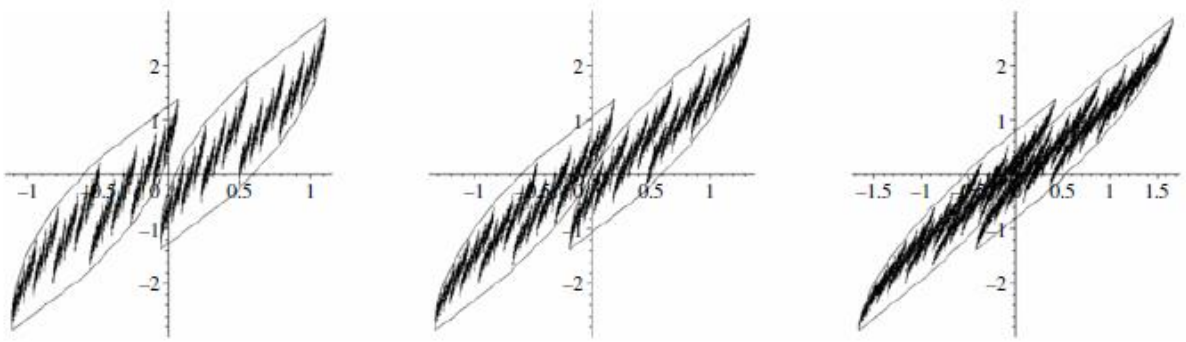


Figure (9)[201]: Level 1 approximation for $\beta_1 \approx 1.9, 1.75$ and 1.6 with $\beta_2 = 1.35$.

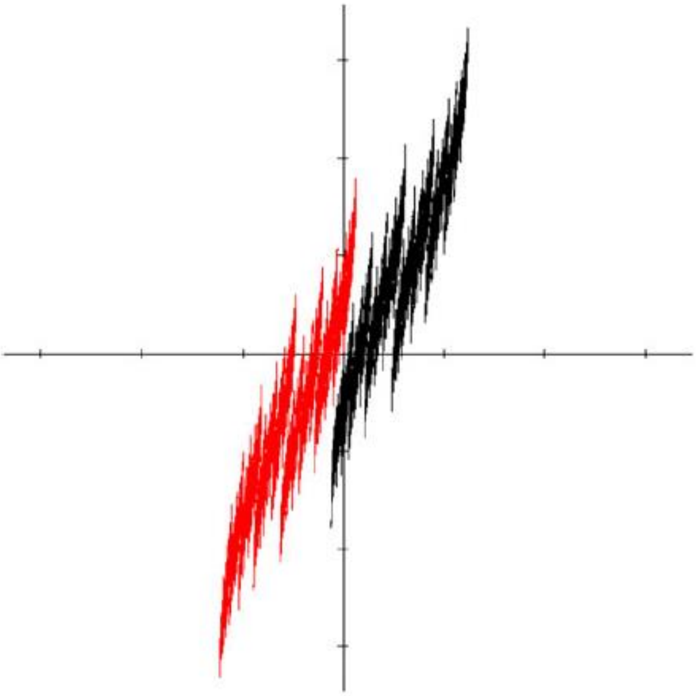


Figure (10)[201]: 'Just touching': we have $T_{-1}(A) \cap T_1(A) = \{(0,0)\}$ for $\beta_1 \approx 1.81618, \beta_2 \approx 1.30022$ being roots of $x^5 - 2x^4 + 2$. Furthermore, here $T_{-1}(K) \cap T_1(K) = \{(0,0)\}$ as well.

(ii) For $k \geq 4$, let $\beta_1^{(k)} \leq \beta_1 \leq \beta_1^{(k+1)}$ and $\beta_2^{(k)} \leq \beta_2 \leq \beta_2^{(k+1)}$ satisfy

$$P_k(\beta_1)P_{k+1}(\beta_2) - P_{k+1}(\beta_1)P_k(\beta_2) = 0. \quad (7)$$

Then $(\beta_1, \beta_2), (\beta_2, \beta_1) \in S_1$.

(iii) Let $\beta_1^{(4)} \leq \beta_1 < \beta_2 \leq \beta_2^{(4)}$ satisfy

$$P_3(\beta_1)P_4(\beta_2) - P_4(\beta_1)P_3(\beta_2) = 0. \quad (8)$$

Then $(\beta_1, \beta_2), (\beta_2, \beta_1) \in S_1$.

(iv) We have $\beta_2^{(k)} \to 1, \beta_1^{(k)} \to 2 \text{ as } k \to +\infty$.

Proof. (i) Assume that $T_1(K) \cap T_{-1}(K)$ has trivial but non-empty intersection. This implies that one of the edges or corners of $T_1(K)$ contains $(0,0)$. Assume first that $(0,0)$ is a corner; then we have that $T_1(\pi(m^k\, p^\infty)) = (0,0)$. This implies

$$\beta_1^{k+1} - 2\beta_1^k + 2 = \beta_2^{k+1} - 2\beta_2^k + 2 = 0,$$

which corresponds to the point $(\beta_1^{(k)}, \beta_2^{(k)})$. It is worth observing that the above equation has no solutions for $k \leq 3$. This results in the interesting consequence that the first, second, third and fourth level approximations are all the same.

(iii) Next, assume that, instead of a corner, it is a line that goes through $(0,0)$. We see that the line $T_1(M_k)$ will intersect the point (0, 0) if the line from $T_1(\pi(m^k\, p^\infty))$ to $T_1(\pi(m^{k+1} p^\infty))$ goes through (0, 0). Letting $(x_k, y_k) = T_1(\pi(m^k\, p^\infty))$ and $(x^{k+1}, y^{k+1}) = T_1(\pi(m^{k+1} p^\infty))$, we see that the y-intercept of the line through these points is

$$\frac{x_k y_{k+1} - y_k x_{k+1}}{x_{k+1} - x_k}.$$

This will equal zero when

$$0 = x_k y_{k+1} - y_k x_{k+1}.$$

Evaluating the above equation at β_1 and β_2 gives equation (7). It is worth observing that the line segment between (x_k, y_k) and (x_{k+1}, y_{k+1}) will only cross the y-axis if these two points are on the opposite sides of the axis. This implies that $\beta_1^{(k)} \leq \beta_1 \leq \beta_1^{(k+1)}$ and $\beta_2^{(k)} \leq \beta_2 \leq \beta_2^{(k+1)}$.

(iii) Similar to (ii).

(iv) Finally, the equation $x^k = 2(x^{k-1} - 1)$ becomes $t^k = t - \frac{1}{2}$ for $t = x^{-1}$. It is clear from the graphs of the left- and right-hand sides that the sequence of smaller real roots, ρ_k, is decreasing, while the sequence of larger real roots, ρ'_k, is increasing. Therefore, $\rho_k^k \to 0$, and hence $\rho^k \to \frac{1}{2}$, which is equivalent to $\beta_1^{(k)} \to 2$ as $k \to +\infty$. On the other hand, $\rho'_k \to 1$, since it is always smaller than 1 and cannot tend to $\kappa < 1$, since in that case κ must be equal to $\frac{1}{2}$ as well, which is impossible. Hence $\beta_2^{(k)} \to 1$.

Figure 10 illustrates the above theorem for $\beta_i = \beta_i^{(4)}, i = 1, 2$.

Corollary (5.1.25)[201]: If $\beta_1 + \beta_2 \geq 3.1294734398566\ldots$ then $(\beta_1, \beta_2) \in S$. If the inequality is strict, then $(\beta_1, \beta_2) \in O$. For all $\varepsilon > 0$ there exist β_1 and β_2 with $\beta_1 + \beta_2 \geq 3.1257839569901 - \varepsilon$ where $(\beta_1, \beta_2) \notin O$.

Proof. Consider the curves $P_k(\beta_1)P_{k+1}(3 - \beta_1 + t) - P_{k+1}(\beta_1)P_k(3 - \beta_1 + t) = 0$. Solving for the local maxima of these (with respect to t), we see that the local maximum for $k = 4$ is maximal, and obtains a value of

$$t = 0.129473439856676017685019631898120681253831009 7982\ldots$$

when

$$\beta_1 = 1.235602860445626103684431317587515643311784524 0595\ldots.$$

Precise algebraic quantities can be given in terms of the roots of a degree 36 polynomial, which we omit.

It was shown in [215] that all neighbourhoods of $(\beta_1^{(k)}, \beta_2^{(k)})$ contain a point that is not in S. Taking $k = 5$ proves the second inequality.

It is worth observing that Solomyak [215] came at this through a different construction. Solomyak first considered the function

$$h_k^{(t)} = 1 - x - \cdots - x^{k-1} + tx^k + x^{k+1} + x^{k+2} + \cdots . \quad (9)$$

Following [215], put

$$B_{[-1,1]} = \left\{1 + \sum_{n=1}^{\infty} a_n z^n \mid an \in [-1, 1]\right\}.$$

For $f \in B_{[-1,1]}$, let $\xi_1(f) \leq \xi_2(f) \leq \cdots$ denote the positive zeroes of f ordered by magnitude and counted with multiplicity. Let

$$\varphi: \gamma \to min\{\xi_2(f): f \in B_{[-1,1]}, f(\gamma) = 0\}.$$

By [215], the function φ is well defined. Furthermore, let $\alpha 2 \approx 0.649138$ be the positive zero of $2x^5 - 8x^2 + 11x - 4$. By the same Proposition, for all $\gamma \in [1/2, \alpha_2]$ there exists a unique function $h_k^{(t)}$ such that $h_k^{(t)}(\gamma) = h_k^{(t)}(\varphi(\gamma)) = 0$. If $\gamma < \lambda < \varphi(\gamma)$, then $(1/\gamma, 1/\lambda) \in S$.

Theorem (5.1.26)[201]: The curve given by $(\gamma, \varphi(\gamma))$ is the same as the level 1 approximation of S given by Theorem (5.1.24).

Proof. We note a few things.

- If $t = -1$ then $h_k^{(t)}(1/\beta) = 0$ if and only if $P_{k-1}(\beta) = 0$.
- If t =1 then $h_k^{(t)}(1/\beta) = 0$ if and only if $P_k(\beta) = 0$.

Hence, the corners of this curve are the same as the corners of the curve S. Let $x_k = s_\mu(pm^k p^\infty)$ and $y_k = s_\lambda(pm^k p^\infty)$. We showed that if $T_1(K)$, the first level convex approximation of A, 'just touches' $T_{-1}(K)$ then

$$x_{k+1}y_k - y_{k+1}x_k = 0. \tag{10}$$

Furthermore, x_k will be on one side of the axis, and x_{k+1} will be on the other. Let

$$t = 2 \cdot \frac{x_{k+1}}{x_{k+1} - x_k} - 1. \tag{11}$$

We see that if $x_k = 0$ (i.e. the corner of K, $(x_k, y_k) = (0,0)$) then $t = -1$. Furthermore, if $x_{k+1} = 0$ then $t = 1$. Hence, t ranges between-1 and 1. This implies that

$$\frac{t+1}{2} x_k = \frac{t-1}{2} x_{k+1}. \tag{12}$$

Using this in equation (10) gives

$$0 = x_{k+1}y_k - y_{k+1}x_k = \frac{t+1}{2} x_{k+1}y_k - \frac{t+1}{2} y_{k+1}x_k$$
$$= \frac{t+1}{2} x_{k+1}y_k - \frac{t-1}{2} x_{k+1}y_{k+1} = \frac{t+1}{2} y_k - \frac{t-1}{2} y_{k+1}.$$

It is worth noting that the values when $t + 1 = 0$ and $x_n = 0$ are when the vertices of K touch (0, 0) and hence are not actually attained when it is the interior of the edge that meets (0, 0). Hence, the division and multiplication of 0 are not problematic. We notice

Figure (11)[201]: Points in S. Those in black come from the level 1 approximation. The additional points come from the level 5 approximation.

that the equation $(t + 1)/2y_k - (t - 1)/2y_{k+1}$ equals 0 if

$$0 = 1/\beta_2 - 1/\beta_2^2 - \cdots - 1/\beta_2^{k+1} + t/\beta_2^{k+2} + 1/\beta_2^{k+3} + 1/\beta_2^{k+4} + \cdots$$
$$= h_k^{(t)} + 1(1/\beta_2).$$

A similar argument shows that $h_{k+1}^{(t)}(1/\beta_1) = 0$, as required.

Consider a finite word $w \in \{p, m\}^n$. Recall that $K_w = T_w(K)$. By our previous notation, $K_n = \bigcup_{|w|=n} K_w$.

To check if $T_1(K_n) \cap T_{-1}(K_n)$ has empty or trivial intersection, it suffices to check $T_1(K_w) \cap T_{-1}(K_{w'})$ for all words $w, w' \in \{p, m\}^n$. To improve the efficiency of this search, we observe that if $T_1(K_w) \cap T_{-1}(K_{w'})$ is empty or trivial, then for all words w', w_0' we have that $T_1(K_{ww0}) \cap T_{-1}(K_{w'w_0'})$ is empty or trivial.

This allows us to improve the efficiency of the search. We again remark that the level 1 approximation (using K_1) is the same as that found in [215]. In fact, this is the same for levels 2, 3 and 4 as well. At level 5, additional points are discovered to be in S that were not provable before (see Figure 11). We could, if necessary, construct curves much like Theorem6.6. This trend continues as we increase to higher level approximations (see Figure 3).

One might conjecture, when looking at the initial pictures produced, that all of our curves coming from a level n approximation are connected. If this were true, then this would imply that S was connected. It turns out, rather surprisingly, that this is not the case. At level 14 we have an occurrence of an island that is not connected to the main body of the curve (see Figure 12). More surprisingly, as we show, this is not an

Figure (12)[201]: Level 14 approximation of $S, \beta_1 \in [1.32025, 1.35275], \beta_2 \in [1.6306, 1.6631]$.

artifact of our choice of approximations of A. This is, in fact, a legitimate island of S that is disconnected from the main body. This proves that S is not connected, and hence the connectedness locus $N = S^c$ studied in detail in [215] is not simply connected.

We gave a technique to show that a point (β_1, β_2) corresponded to a totally disconnected set A. Using this technique, we observed at level 14 that the approximation toS was not connected (see Figure 12).

We will prove that this region is indeed in a separate connected component with respect to the rest of S. In Figure 12 we see a chevron-shaped object C which is disconnected from the main body of the approximation of S. A significant part of our proof is computer-assisted. First, we need to show that there exists a point in C which is provably in S. A quick computer check yields $(1.335438104, 1.646743824) \in C \subset S$.

To prove that C is separate from the main body of S we will give six path-connected regions, $R_{w_1}, \ldots, R_{w_6}$, all disjoint from S, such that R_{w_1} overlaps with R_{w_2}, which in turn overlaps with R_{w_3}, and so on, where finally R_{w_6} overlaps with the original set R_{w_1}. These overlapping sets will surround C—see Figure 13.

We need a criterion for a pair (β_1, β_2) not to lie in S. As usual, m stands for-1, and p for 1. We will also use z for 0.

Lemma (5.1.27)[201]: If β_1 and β_2 are distinct roots of $P \in Z[x]$ with the coefficients of P restricted to $\{p, z, m\}$, then $(\beta_1, \beta_2) \notin S$.

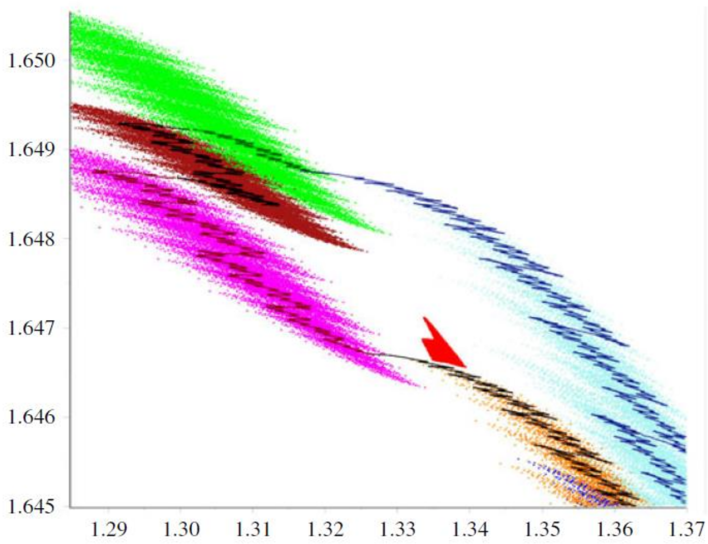


Figure (13)[201]: The chevron C and $R_{w_1}, R_{w_2}, \ldots, R_{w_6}$, along with overlapping continuous path.

Proof. Let $P(x) = a_n x^n + \cdots + a_0$ with $a_i \in \{-1, 0, 1\}$. Write $2P(x) = P+(x) - P-(x)$ with $P+(x) = a_n^+ x^n + \cdots + a_0^+, a_i \in \{-1, 1\}$, and $P-(x) = a_n^- x^n + \cdots + a - 0, a_i \in \{-1, 1\}$. As $P(\beta_1) = P(\beta_2) = 0$, we have that $P_+(\beta_1) = P_-(\beta_1)$ and $P_+(\beta_2) = P_-(\beta_2)$.

Notice that

$$s_1/\beta_1((a_n^+ a_{n-1}^+ \ldots a_0^+)^\infty) = P_+(\beta_1)(1/\beta_1^{n+1} + 1/\beta_1^{2(n+1)} + \cdots)$$
$$= P_-(\beta_1)(1/\beta_1^{n+1} + 1/\beta_1^{2(n+1)} + \cdots)$$
$$= s_1/\beta_1((a_n^- a_{n-1}^- \ldots a_0^-)^\infty).$$

A similar result holds for $1/\beta_2$ which gives us that

$$\pi((a_n^+ a_{n-1}^+ \ldots a_0^+)^\infty) = \pi((a_n^- a_{n-1}^- \ldots a_0^-)^\infty).$$

As $a_n \neq 0$ we see that $a_n^+ \neq a_n^-$, and hence

$$\pi((a_n^+ a_{n-1}^+ \ldots a_0^+)^\infty) = \pi((a_n^- a_{n-1}^- \ldots a_0^-)^\infty) \in T_1(A) \cap T_{-1}(A).$$

This give that A is connected, and hence $(\beta_1, \beta_2) \notin S$.

Next we need a result of Odlyzko and Poonen [211].

Lemma (5.1.28)[201]: Let Y be a topological space. Suppose $f : \{0,1\}^{\mathbb{N}} \to Y$ is a continuous map such that

$$f([w_0]) \cap f([w_1]) \neq \emptyset$$

for all $w \in \{0,1\}^*$. Then the image of f is path connected.

Recall that $[i_1 \ldots i_k]$ stands for the cylinder $\{a_j\}_{j=1}^{\infty} \subset \{0,1\}^{\mathbb{N}}$ such that $a_j = i_j$ for $j = 1, 2, \ldots, k$. Lemma (5.1.28) can be easily generalized to the space $\{p, z, m\}^{\mathbb{N}}$.

Lemma (5.1.29)[201]: Let Y be a topological space. Suppose $f : \{p, z, m\}^{\mathbb{N}} \to Y$ is a continuous map such that

$$f([wz]) \cap f([wp]) \neq \emptyset,$$
$$f([wm]) \cap f([wp]) \neq \emptyset,$$
$$f([wm]) \cap f([wz]) \neq \emptyset$$

for all $w \in \{p, z, m\}^*$. Then the image of f is path connected.

The proof is a simple variation of the result of Odlyzko and Poonen. We provideit here for completeness.

Proof. This is, in essence, a bisection method. Given two infinite words $w = a_1 a_2 a_3 \ldots$ and $w' = b_1 b_2 b_3 \ldots$, we define the usual metric by $dist(w, w') = \frac{1}{2^k}$ where $a_i = b_i$ for

$i = 1, \ldots, k-1$ and $a_k \neq b_k$. If no such k exists,then $w = w'$ and $dist(w, w') = 0$. Given two points $x' = f(w')$ and $x_1 = f(w_1)$, we construct two new words $w_{\frac{1}{2}}$ and $w'_{\frac{1}{2}}$ such that:

• $f\left(w_{\frac{1}{2}}\right) = f\left(w'_{\frac{1}{2}}\right)$;

• $dist\ (w_0, w_{\frac{1}{2}}) < dist(w_0, w_1);\ and$

• $dist(w'_{\frac{1}{2}}, w_1) < dist(w_0, w_1)$.

To do this, we let w be the common prefix of w_0 and w_1 so that $w_0 = wa_0v_0$ and $w_1 = wa_1v_1$ with $a_0 \neq a_1$. We then find $w_{\frac{1}{2}} \in [wa_0]$ and $w'_{\frac{1}{2}} \in [wa_1]$ so that $f\ (w_{\frac{1}{2}}) = f(w'_{\frac{1}{2}}) \in f([wa_0]) \cap f([w, a_1])$. Such a point exists by assumption. We now induct on this construction to find points $x_{\frac{1}{4}}$ and $x_{\frac{3}{4}}$ and then $x_{\frac{1}{8}}, x_{\frac{3}{8}}, x_{\frac{5}{8}}, x_{\frac{7}{8}}$ and so on. We notice by the continuity of f and the fact the distances between adjacent points go to 0 in the limit, that this construction will define a continuous path in the image of f.

Let $v \in \{p, m, z\}^*$ be a finite word of length n. Furthermore, assume that $v1 \neq z$. Define $P_v(x) = P(x) = v_1x^{n-1} + \cdots + v_n$. If β_1, β_2 are distinct roots of P then we see from Lemma (5.1.27) that $(\beta_1, \beta_2)\ / \in S$. Let $\beta_1^+, \beta_1^-, \beta_2^+, \beta_2^-$ be distinct roots of the rational function $P(x) \pm 1/(x-1)$, assuming that they exist. Let $I_1 = [\beta_1^\pm, \beta_1^\mp]$ and $I_2 = [\beta_2^\pm, \beta_2^\mp]$. Let $f\ (x) \in \{\sum^{\infty(i=1)} w_i x^{-i} : w \in \{p, m, z\}^{\mathbb{N}}\}$. We see that if $|f'(x)| < |P'(x)|$ for all $x \in I_1$, then $P(x) + f\ (x)$ will have a unique root in I_1. We will denote this root by $\beta_1^{(w)}$. Similarly, if $|f'(x)| < |P'(x)|$ for all $x \in I_2$, then $P(x) + f\ (x)$ will have a unique root in I_2, which we will denote by $\beta_2^{(w)}$.

We see that if $|P'(x)| > 1/(x-1)^2$ for all $x \in I_1$ and $x \in I_2$, then there will be welldefined roots $\beta_1^{(w)}$ and $\beta_2^{(w)}$ for all $w \in \{p, m, z\}^{\mathbb{N}}$.

We will call the existence of $\beta_1^\pm, \beta_2^\pm$ and $|P'(x)| > 1/(x-1)^2$ on I_1 and I_2 property RD. If, for a word v, its associated polynomial P has property RD, then the map $f_v = f : \{p, z, m\}^{\mathbb{N}} \to \mathbb{R}^2$ given by $f(w) = (\beta_1^{(w)}, \beta_2^{(w)})$ is well defined. It is easy to see that such a map is continuous. It is also easy to see that for those infinite words w which only contain a finite number of non-zero terms, the image corresponds to points that are roots of a{p, z, m}polynomial, and hence such w are not inS.

To see that any such w satisfies the conditions of Lemma (5.1.29), let v correspond to the coefficients of P. Suppose $w \in \{p, z, m\}^*$. We see that $f_v(w_0) = f_v(wvw) = f_v(w\tilde{v}\tilde{w})$. Thus, if we have a polynomial P_v which satisfies property RD, then we can associate with P_v a set of values which are not in S and whose closure is path connected. We will denote this path-connected set by R_v. By Proposition (5.1.19), the complement of S is closed. Consequently, $R_v \cap S = \emptyset$ for all v satisfying property RD.

It is easy to see that if w satisfies property RD and w is a prefix of w', then w' satisfies property RD as well. Furthermore, if w is a prefix of w', then $R_{w'} \subset R_w$.

Lemma (5.1.30)[201]: Let w satisfy property RD. Then $f(wm^\infty), f(wp^\infty) \in R_w$. Furthermore, R_w is contained with in the box with side sparallel to the axes and with cornersat $f(wm^\infty)$ and $f(wp^\infty)$.

We call such a box a bounding box for Rw. We will also need the concept of a set of bounding boxes for a continuous path. Let ww_0 and ww_1 be two points within R_w. By

Lemma (5.1.29), there is a continuous path from ww_0 to ww_1 in R_w. Let k be fixed. To construct this path, we find a series of intermediate points $w_{i/2^k}$, each with two addresses. Each of these addresses is such that $w_{i/2^k}$ and $w_{(i+1)/2^k}$agree on the first $|w| + k$ terms. Denote these terms by $a_1a_2 \ldots a_k$.

Thus, both these terms are found within the subregions $R_{wa_1a_2\ldots a_k}$. Furthermore, by construction, the path from $w_{i/2^k}$ to $w_{(i+1)/2^k}$ will also be within this subregion. Hence, this pair, and the path between this pair, will be contained within the bounding box for $R_{wa_1a_2\ldots a_k}$. Taking the union over all of these pairs, we get a series of smaller bounding boxes that contain the continuous path from ww_0 to ww_1. We will call such a series of boxes the level k bounding boxes for a path in R_w.

Lemma (5.1.31)[201]: The following words satisfy property RD.

$$w_1 = pmmmpzzppzppzppz,$$
$$w_2 = pmmmzp7mz,$$
$$w_3 = pmmmzp7mp,$$
$$w_4 = pmmmzp7zm,$$
$$w_5 = pmmmpzzppzpppzpz,$$
$$w_6 = pmmmpzzppppmp4zp.$$

Proof. This is a simple calculation that we leave as an exercise for the reader.

Lemma (5.1.32)[201]: The closure of the set of roots generated by the polynomials in Lemma (5.1.31) surrounds C.

Proof. To see that R_{w_1} isconnectedto R_{w_2}, consider $R_{w_1zpppzpzp}$ and $R_{w_2m^{11}}$. The former has corners at

$$[1.323453274, 1.648718809], [1.314160784, 1.648757942]$$

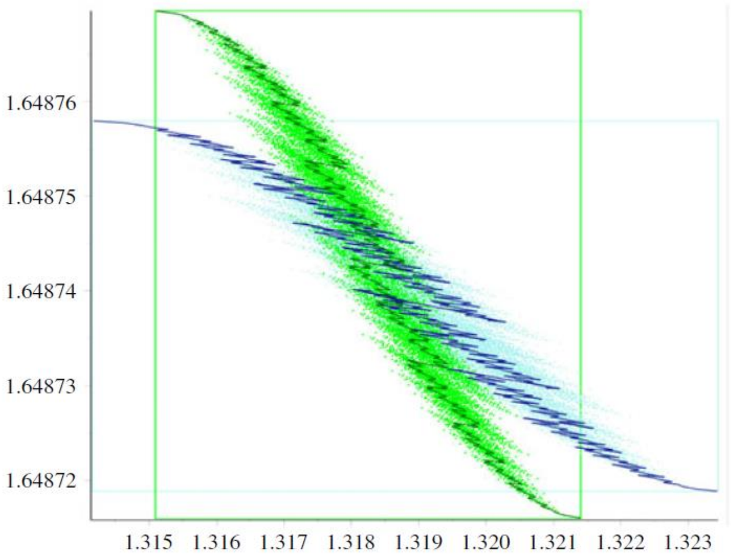


Figure (14)[201]: $R_{w_1zpppzpzp}$ and $R_{w_2m^{11}}$.

and the latter has corners at [1.321413068, 1.648715950],[1.315100914, 1.648769575]. The path from $[1.323453274, 1.648718809]$ to $[1.314160784, 1.648757942]$ must intersect the path from [1.321413068, 1.648715950] to [1.315100914, 1.648769575]. See Figure 14 for these two sets, the continuous paths going from $f_{w_1zpppzpzp}(p^\infty)$ to $f_{w_1zpppzpzp}(m^\infty)$and from $f_{w_1zpppzpzp}(p^\infty)$ to $f_{w_2m^{11}}(m^\infty)$, and the bounding boxes.

To see that R_{w_2} is connected to R_{w_3}, we notice that

$$f_{w_2}(pmmmzp^7m) = f_{w_3}(mmmzp^7m).$$

To see that R_{w_3} is connected to R_{w_4}, we notice that

$$f_{w_3}(ppzm^7) = f_{w_4}(pppzm^7).$$

To see that R_{w_4} is connected to R_{w_5}, consider $R_{w4m^{14}}$ and $R_{w_5ppzzpppzpmz}$. The former has corners at

$$[1.328228762, 1.646703763], [1.324717957, 1.646712975]$$

and the latter has corners at

$$[1.327323576, 1.646702692], [1.324894555, 1.646715284].$$

The path from $[1.328228762, 1.646703763]$ to $[1.324717957, 1.646712975]$ must intersect the path from $[1.327323576, 1.646702692]$ to $[1.324894555, 1.646715284]$. See Figure 15 and the continuous paths connecting the extreme points of each of these sets. For the next two, we need to strengthen the idea of a bounding box as described above.

Consider $R_{w_5mmmp^4mppp}$ and $R_{w_6pz^4zzmzmm}$. See Figure 16 and the continuous paths connecting the extreme points of each of these sets, as well as the level 9 bounding

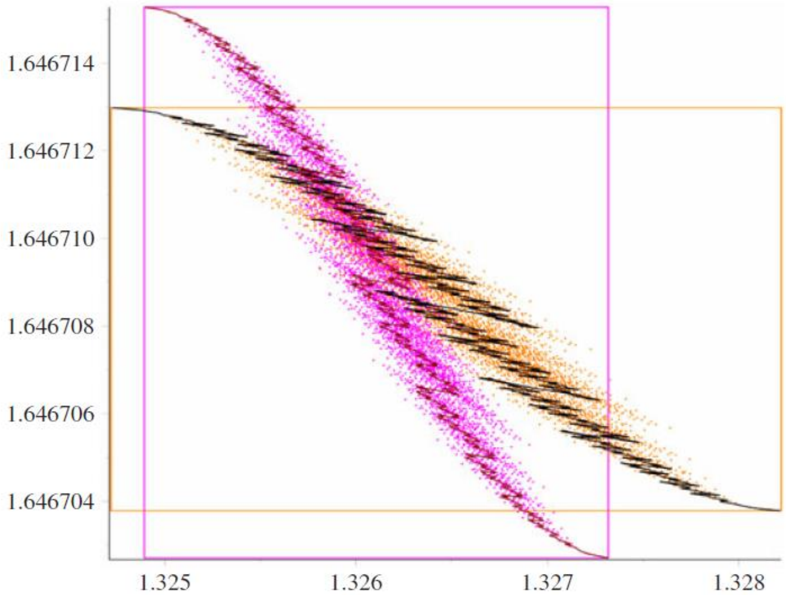


Figure (15)[201]: $R_{w_4m^{14}}$ and $R_{w_5ppzzpppzpmz}$.

boxes for the path in $R_{w_5mmmp^4mppp}$ and the level 2 bounding boxes for the path in $R_{w_6pzm^4zzmzmm}$. Precise coordinates for the bounding boxes for the continuous paths can be found at [207].

Finally, consider $R_{w_6mmmp^7}$ and $R_{w_1zppm^4z^5m}$. See Figure 17 and the continuous paths connecting the extreme points of each of these sets, as well as the level 9 bounding boxes for the path in $R_{w_5mmmp^4mppp}$ and the level 2 bounding boxes for the path in $R_{w_6pzm^4zzmzmm}$. Precise coordinates for the bounding boxes for the continuous paths can be found at [207].

These surround the region in question, see Figure 13.

Corollary (5.1.33)[201]: The set S is not connected.

Corollary (5.1.34)[201]: The connectedness locus $N = S^c$ is not simply connected.

There are a great deal of questions that this line of research raises which still remain unanswered. Here are some of them.

(i) Is it true that if some point of the attractor has a non-empty neighbourhood, then so does (0, 0)? In particular, what is the precise relationship between I and Z?

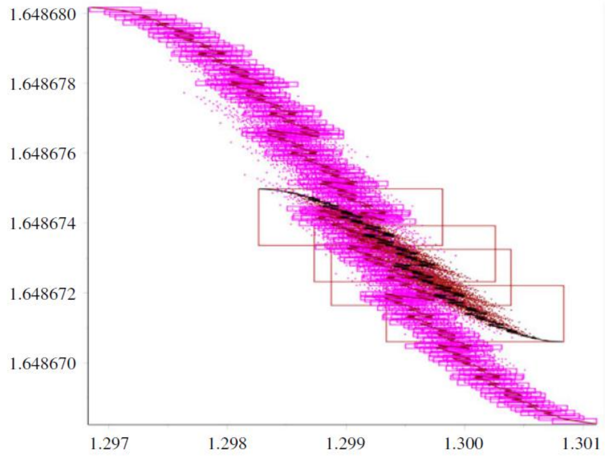


Figure (16)[201]: $R_{w_5 mmmp^4 mppp}$ and $R_{w_6 pzm^4 zzmzmm}$.

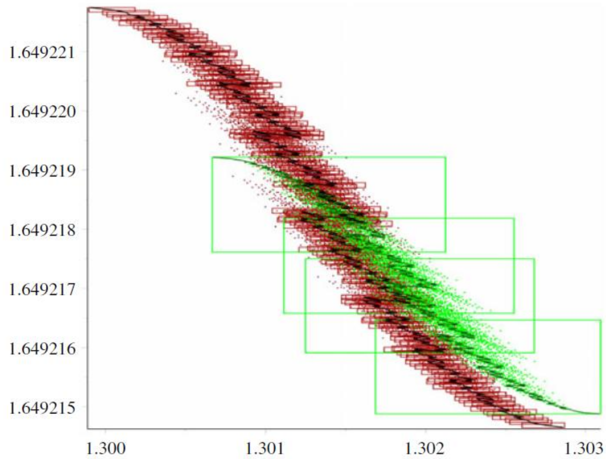


Figure (17)[201]: $R_{w_6 mmmp^8}$ and $R_{w_1 zppm^4 z^5 m}$.

(ii) We see that if $(0,0) \,/\!\in\, A_{\beta_1,\beta_2}$, then $(\beta_1,\beta_2) \notin Z$. There are examples of $(\beta_1,\beta_2) \notin Z$ suchthat A_{β_1,β_2} nonetheless contains (0, 0), see Figure 5. It would be helpful to find better criteria for points.

(iii) Find an example of β_1, β_2 such that:

- $(0,0) \in A_{\beta_1,\beta_2}$;
- $(0,0) \,/\!\in\, A^o_{\beta_1,\beta_2}$;
- $(\beta_1,\beta_2) \,/\!\in\, \partial S$.

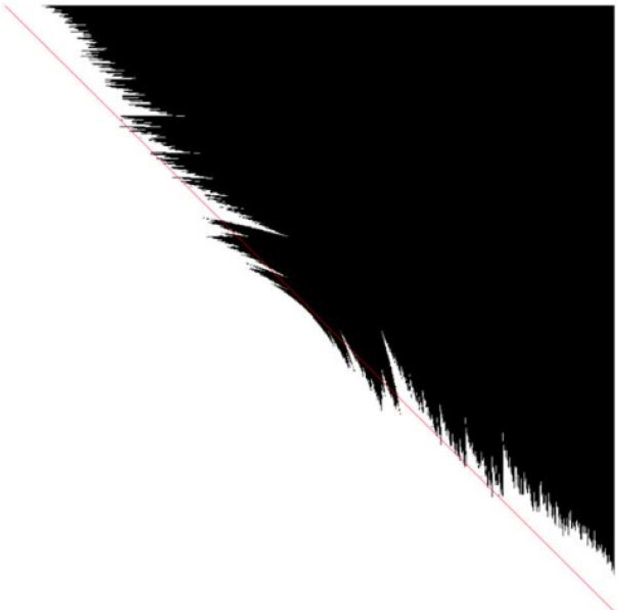


Figure (18)[201]: The set S together with the diagonal $\beta_1 + \beta_2 = 3$. (Level 20 approximation.)

(iv) Can a point with a unique address be an interior point of A?

(v) Does the claim in Theorem (5.1.18)(ii) hold for all pairs (β_1, β_2)? Note that given $\beta \in (1,2)$, almost every $x \in (0, 1/(\beta - 1))$ has a continuum of β-expansions [212], and, furthermore, this continuum can be chosen to have an exponential growth [208]. Thus, one could hope to adapt our argument so it would hold for (β_1, β_2) with both β_1 and β_2 greater than the golden ratio.

(vi) We see that $S \subset O$. Furthermore, $\left(\beta_1^{(n)}, \beta_2^{(n)}\right) \in \partial S \cap \partial O$. When approximating S and O computationally, via Lemma (5.1.23), then the level n approximation of O is the closure of the level n approximation of S. Is O the closure of S?

(vii) $Is\ Z \cap O = \emptyset$?

(viii) Justify the 'spikes' in S near (1, 2) and (2, 1). That is, we know that both corners are limit points of S (Theorem (5.1.24)); is it true that for any $h > 0$ there exists a point (β_1, β_2) in $(2 - h, 2) \times (1, 1 + h)$ which is not in S? By looking at $(\beta_1^{(n)}, \beta_2^{(n)})$ we get a partial idea of the structure of S near (1, 2), but not a complete picture.

(ix) As mentioned at the beginning of §7, $(\beta_1, \beta_2) \in S$ where $\beta_1 = 1.335438104, \beta_2 = 1.646743824$. Thus, we have $(\beta_1 + \beta_2) = 2.982181928$, i.e. some small chunk of S lies below the diagonal (which is not at all obvious from Figure 3). It would be interesting to find the smallest $\varepsilon > 0$ such that $S \subset \{(\beta_1, \beta_2): \beta_1 + \beta_2 > 3 - \varepsilon\}$, see Figure 18.

(x) We know that S contains at least three disjoint components (by symmetry around the line $\beta_1 = \beta_2$). Does it contain a finite number of components or an infinite number of components?

(xi) Prove or disprove that, for sufficiently small β_1 and β_2, the attractor A_{β_1,β_2} is simply connected.

(xii) Show that the lower box (or Hausdorff) dimension of $\partial A_{\beta_1,\beta_2}$ is strictly greater than 1 for all β_1, β_2.

Section (5.2): Positive Hausdorff Measure

Self-conformal sets are a natural generalisation of self-similar sets. Instead of similitudes, theyare defined by using contractive conformal maps $\varphi_1,\ldots,\varphi_N$. The prime examples include Juliasets of hyperbolic rational functions on $\mathbb{C}$, such as Julia sets for $z \mapsto z^2 + c$ with $|c| \geq 2.48$. Asthe only contractive entire functions on $\mathbb{C}$ are the similitudes, one has to restrict the definitionto a bounded open set Ω where the mappings φ_i are contractive. In the real line, the maps φ_i are contractive $C^{1+\alpha}$-functions with non-vanishing derivative. The self-conformal set is the uniquenon-empty compact set F satisfying

$$F = \bigcup_{i=1}^{N} \varphi_i(F) = \bigcap_{n\in\mathbb{N}} \bigcup_{i\in\{1,\ldots,N\}^n} \varphi_i(X),$$

Where $X \subset \Omega$ is any compact set satisfying $\varphi_i(X) \subset X$and $\varphi_i = \varphi_{i_1} \circ\cdots\circ \varphi_{i_n}$for all $= i_1 \cdots i_n$.

We are primarily interested in determining the size of a self-conformal set F. If the "construction pieces"$\varphi_i(X)$ are separated, then, by relying on conformality, one expects the dimension of F to be close to the value s for which $1 = \sum_{i=1}^{N} \operatorname{diam}\big(\varphi_i(X)\big)^s \approx \sum_{i=1}^{N} \|\varphi_i'\|^s$. Intuitively, one should getbetter and better estimates for the dimension by iterating this idea. Indeed, this is precisely what happens: it is straightforward to see that in general, the Hausdorff dimension of F is at most thelimiting value of such approximations, $\dim_H(F) \leq P^{-1}(0)$, where

$$P(s) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{i\in\{1,\ldots,N\}^n} \|\varphi_i'\|^s,$$

and if there is enough separation, then $\dim_H(F) = P^{-1}(0)$. In fact, Peres, Rams, Simon, and Solomyak [137] have shown that if $s = P^{-1}(0)$, then thes-dimensional Hausdorff measure of F is positive, $H^s(F) > 0$, if F satisfies the open set condition, a natural separation condition un-der which the overlapping of the construction pieces of roughly the same diameter has bounded multiplicity.

We focus on the case $\dim_H(F) < P^{-1}(0)$. At first, it is easy to see that this occurs when thereare exact overlaps, meaning that there are $i \neq j$ for which $\varphi_i\big|F = \varphi_j\big|_F$. A related separation condition is the weak separation condition which, is otherwise the same as theopen set condition but allows exact overlapping. The famoussdimension drop conjecture claims that exact overlapping is the only way to drop the Hausdorff dimension of F below $P^{-1}(0)$. Hochman[222] has verified the conjecture for all self-similar sets in the real line defined by algebraic parameters.It should be remarked that Hochman's proof does not generalise to the self-conformal case.

In the self-similar case, Zerner [136] introduced the identity limit criterion, $\{\varphi_i^{-1} \circ \varphi_j\}_{i,j}$does not accumulate to the identity, and showed that it is equivalent to the weak separation condition. The self-conformal case is morecomplicated since we cannot use inverses. We introduce the identity limit criterion for the conformal setting and in our main technical lemma, Lemma (5.2.13), we show that if it is not satisfied, then there are arbitrary small $\delta > 0$ such that, for some distinct maps φ_i and φ_j,

$$|\varphi_i(x) - \varphi_j(x)| \approx \delta\|\varphi_i'\| \approx \delta\|\varphi_j'\|$$

for all x. The lemma thus gives the existence of maps which are arbitrarily close to each otherin the relative scale. Applying this observation inductively, we infer that the overlapping of theconstruction pieces of roughly the same diameter has unbounded multiplicity and hence, the weak separation condition does not hold. Conversely, pigeonholing such unbounded multiplicity implies the existence of two maps being arbitrarily close to each other in the relative scale. Therefore, we see that the identity limit criterion is equivalent to the weak separation condition also in theself-conformal case. This is stated in Theorem (5.2.2).

The role of the identity limit criterion is essential in our considerations. The Assouad dimension of F, $\dim_A(F)$, is the maximal Hausdorff dimension of its weak tangents, the Hausdorff limitsof successive magnifications. In general, the Assouad dimension serves as an upper bound forthe Hausdorff dimension but if the set is Ahlfors regular, then the two dimensions agree. Fraser,Henderson, Olson, and Robinson [221] showed that if a self-similar set in the real line does not satisfythe identity limit criterion, then its Assouad dimension is1. In Theorem (5.2.17), we generalise this observation to the self-conformal case. To prove this, we again apply Lemma (5.2.13) inductively tofind small scales containing as many equally distributed points of F as we wish. This shows thatthe unit interval appears as a weak tangent and proves the result.

In our main result, Theorem (5.2.7), we prove that if $s = \dim_H(F)$, then the s-dimensional Haus-dorff measure and content are equivalent. An almost immediate consequence of this is that thepositivity of the Hausdorff measure is equivalent to the Ahlfors regularity. The result generalisesthe corresponding theorem of Farkas and Fraser [220] in the self-similar case. It should be empha-sized that their proof does not generalise to the self-conformal case. With this theorem, we cannow address the dimension drop conjecture on self-conformal sets in the real line having Hausdorff dimension strictly less than 1. Indeed, Lau, Ngai, Wang [227]have shown that the weak separation condition implies $H^s(F) > 0$ for $s = \dim_H(F)$. As mentioned above, this implies Ahlfors regularityand therefore, also the Assouad dimension is strictly less than 1. Since this further implies the identity limit criterion and hence also the weak separation condition, we conclude that all of these conditions are equivalent. As the only difference between the open set condition and the weak sep-aration condition is the exact overlapping, we see, by recalling the result of Peres, Rams, Simon, and Solomyak [137], that the dimension drop conjecture holds for self-conformal sets with positive Hausdorff measure. This can be considered to be the main consequence of our considerations. The result is stated in Theorem (5.2.5).

We show the equivalence of the Hausdorff measure and content in a slightly more general setting of quasi self-similar sets. We devoted to the study of self-conformal sets and their separation conditions in $\mathbb{R}^d$. Results in thereal line and dimension drop conjecture are explored.

Recall that thes-dimensional Hausdorff measure H^s of a set $A \subset \mathbb{R}^d$ is defined by

$$H^s(A) = \lim_{\delta \downarrow 0} H^s_\delta(A) = \sup_{\delta > 0} H^s_\delta(A),$$

Where

$H^s_\delta(A) = inf\{\sum_i diam(U_i)^s : A \subset \bigcup_i U_i \text{and } \mathrm{diam}(U_i) \leq \delta\}$

is the s-dimensional Hausdorff δ-content of A. The Hausdorff measure is Borel regular and theHausdorff content is an outer measure – usually highly non-additive and not a Borel measure. However, the Hausdorff content is slightly easier to compute, and is always finite for bounded sets, irrespective of s. It is straightforward to see that $H^s(A) = 0$ if and only if $H^s_\infty(A) = 0$ and so the Hausdorff measure and content share the same critical exponent, the Hausdorff dimension $\dim_H$ of A which is defined by $\dim_H(A) = inf\{s: H^s(A) = 0\}$.

Observe that the assumptions of Theorem (5.2.7) are stronger than those that define quasi self-similar sets; see [138] and [130]. Quasi self-similar sets differ to the sets we consider by only requiring the lower bound in (16) to hold. The upper bound is crucial in (20) and it seems unlikely that our assumptions are satisfied by quasi self-similarity alone.
The following result is a straightforward corollary of Theorem (5.2.7). We say that a set $A \subset \mathbb{R}^d$ isAhlforss-regularif there exists a Radon measure μ supported on A and a constant $C \geq 1$ such that

$$C^{-1}r^s \leq \mu(B(x,r)) \leq Cr^s \tag{13}$$

for all $x \in A$ and $0 < r < diam(A)$.

Proposition (5.2.1)[216]: Let $F \subset \mathbb{R}^d$ be a set satisfying the assumptions of Theorem (5.2.7). If $s = \dim_H(F)$, then $H^s(F) > 0$if and only if F is Ahlforss-regular.

Proof. Assuming F to be Ahlforss-regular, let μ be a measure satisfying (13). Since $\mu(F) \leq \sum_i \mu(U_i) \leq C \sum_i diam(U_i)^s$ for all δ-covers$\{U_i\}_i$ of F, we get $H^s_\delta(F) \geq \mu(F) > 0$ for all $\delta > 0$ and,consequently, $H^s(F) > 0$. To show the necessity of the Ahlfors regularity, suppose that $H^s(F) > 0$.By Theorem (5.2.7), there is a constant $C \geq 1$ such that

$$H^s|_F(B(x,r)) \leq Cr^s \tag{14}$$

for all $x \in F$and $r > 0$. For each $x \in F$ and $0 < r < diam(F)$, let $g_{x,r}: F \to F \cap B(x,r)$ be as in(16). The existence of such mappings implies

$$H^s|_F(B(x,r)) \geq H^s(g_{x,r}(F)) \geq D^{-s}H^s(F)r^s$$

for all $x \in F$ and $0 < r < diam(F)$. Recalling that F is compact, it follows from (14) that $H^s(F) < \infty$ and $H^s|_F$ is therefore a Radon measure. We have thus finished the proof.

Let $N \geq 2$ and consider the family of N contractions$\{\varphi_1, \dots, \varphi_N\}$on $\mathbb{R}^d$. We call this family aniterated function system. If all the mappings $\varphi_i: \mathbb{R}^d \to \mathbb{R}^d$ are strict contractions, then there exists a unique non-empty compact set F, called the attractor of the iterated function system, satisfying

$$F = \bigcup_{i=1}^{N} \varphi_i(F).$$

When all the mappings $\varphi_i$are similarities the attractor is known as aself-similar set. We consider the larger class of iterated function systems where all the mappings are conformal contractions and in this case, we refer to F as a self-conformal set.
We give a precise definition for a conformal iterated function system. Fix an open set $\Omega \subset \mathbb{R}^d$. A C^1-mapping $\varphi: \Omega \to \mathbb{R}^d$ isconformalif the differential $\varphi'(x): \mathbb{R}^d \to \mathbb{R}^d$is a similarity,i.e. satisfies $|\varphi'(x)y| = |\varphi'(x)||y| \neq 0$ for all $x \in \Omega$ and $y \in \mathbb{R}^d \setminus \{0\}$ and, as a function of x, is H¨older continuous, i.e. there exist $\alpha, c > 0$ such that

$$|\varphi'(x) - \varphi'(y)| \leq |x - y|^\alpha \tag{15}$$

for all $x, y \in \Omega$. For $d \geq 2$, the H¨older continuity follows from the similarity of the differential and injectivity . In fact, conformal mappings in the plane correspond to the holomorphic functions on $\mathbb{C}$ with non-zero derivative on their respective domain, and in higher dimensions, by Liouville's theorem, conformal mappings are either homotheties, isometries, or compositions of reflections and inversions of a sphere. In the one dimensional case, conformal mappings are simply the $C^{1+\alpha}$-functions with non-vanishing derivative. We say that$\{\varphi_i\}_{i=1}^N$is aconformal iterated functionsystemif each φ_i is an injective conformal mapping on a bounded open convex set Ω such that $\varphi_i(\Omega) \subset \Omega$ and $\|\varphi_i'\| := \sup_{x \in \Omega} |\varphi_i'(x)| < 1$. There exists a compact set $X \subset \Omega$ such that $\bigcup_{i=1}^N \varphi_i(X) \subset X$, which

guarantees the existence of the self-conformal set;for details, see Lemma(5.2.8). Self-conformal sets are a natural generalisation of self-similar sets.
We shall verify that self-conformal sets satisfy the assumptions of Theorem (5.2.7). We thus obtain the following result as an immediate corollary of Theorem (5.2.7) and Proposition (5.2.1).

The above theorem extends to graph-directed and sub self-conformal sets in a straightforward manner. It is pointed out in [220] that the constant C above cannot be chosen to be 1. We may thus consider that the theorem generalises there sults of Farkas and Fraser [220] on graph-directed self-similar sets; see also [219]. It is also worthwhile to emphasize that the method of Farkas and Fraser cannot be applied to prove Theorem (5.2.9): their proof relied on an abstract lemma on measurable hulls which can only be applied if the measure and content of the whole set are equal.
Let$\{\varphi_i\}_{i=1}^N$be aconformal iterated function system and F be the associated self-conformal set. We use the conven-tion that whenever we speak about a self-conformal set F, then it is automatically accompaniedwith a conformal iterated function system which defines it. Let $\Sigma = \{1, \dots, N\}^{\mathbb{N}}$ be the collectionof all infinite words constructed from integers$\{1, \dots, N\}$. If $i = i_1 i_2 \cdots \in \Sigma, t$hen we define $i|_n = i_1 \cdots i_n$ for all $n \in \mathbb{N}$. The empty word $i|_0$ is denoted by $\emptyset$. Observe that $\Sigma_* = \bigcup_{n=0}^{\infty} \Sigma_n$, where $\Sigma_n = \{i|_n : i \in \Sigma\}$for all $n \in \mathbb{N}$, is the free monoid on $\Sigma_1 = \{1, \dots, N\}$. If $n \in \mathbb{N}$ and $i = i_1 \cdots i_n \in \Sigma_n$, then we write $\varphi_i = \varphi_{i_1} \circ \cdots \circ \varphi_{i_n}$. For $i \in \Sigma_* \backslash \{\emptyset\}$we set $i^- = i|_{|i|-1}$, where $|i|$ is the length of i.
We say that F satisfies the weak separation condition if
$sup\{\#\Phi(x,r) : x \in F \text{ and } r > 0\} < \infty$,
Where
$\Phi(x,r) = \{\varphi_i|_F : diam(\varphi_i(F)) \le r < diam(\varphi_i - (F)) \text{ and } \varphi_i(F) \cap B(x,r) \neq \emptyset\}$
for all $x \in \mathbb{R}^d$ and $r > 0$. Furthermore, we say that F satisfies theidentity limit criterion if
$inf\{\|\varphi' i\|^{-1} \sup_{x\in F} |\varphi_i(x) - \varphi_j(x)| : i,j \in \Sigma_* \text{ such that } \varphi_i|_F \neq \varphi_j|_F\} > 0$.
The weak separation condition for self-conformal sets was introduced by Lau, Ngai, and Wang[227]. Our definition is strictly weaker than the original one; see Example (5.2.16). This modification was needed to be able to find a definition for the identity limit criterion equivalent to the weak separation condition. The following result is proved.
Theorem (5.2.2)[216]: Let $F \subset \mathbb{R}^d$ be a self-conformal set containing at least two points. Then Fsatisfiesthe weak separation condition if and only if it satisfies the identity limit criterion.
The weak separation condition provides us with a sufficient condition for the self-conformal set to have positive measure. The identity limit criterion gives, at least in principle, a checkable condition for the positivity.
Proposition (5.2.3)[216]: Let $F \subset \mathbb{R}^d$ be a self-conformal set satisfying the weak separation condition and $s = \dim_H(F)$. Then $H^s(F) > 0$.
The above result was observed first time by Lau, Ngai, and Wang[227]. Its proof follows imme-diately from [225]. We remark that [225] uses the original definition of Lau, Ngai, and Wang [227] (see Example (5.2.16)) but its proof applies verbatim also with our definition of weak separation condition.

The Assouad dimension of a set$A \subset \mathbb{R}^d$, denoted by $\dim_A(A)$, is the infimum of all s satisfyingthe following: There exists a constant $C \ge 1$ such that each set $A \cap B(x,R)$ can be covered by at most $C(R/r)^s$ balls of radius r centered at A for all $0 < r < R$. It is easy to see that $\dim_H(A) \le \dim_A(A)$ for all sets $A \subset \mathbb{R}^d$.

The proof of the following theorem is postponed until.

The above result, together with Theorem (5.2.2), generalises the corresponding result of Fraser, Henderson, Olson, and Robinson [221] on self-similar sets in the real line.

The following corollary generalises the corresponding result of Farkas and Fraser [220] on self-similar sets.

Corollary (5.2.4)[216]: Let $F \subset \mathbb{R}$ be a self-conformal set containing at least two points such that $s = \dim_H(F) < 1$. Then the following five conditions are equivalent:

(i) F satisfies the weak separation condition,

(ii) $H^s(F) > 0$,

(iii) F is Ahlforss-regular,

(iv) $\dim_A(F) = s$,

(v) F satisfies the identity limit criterion.

Proof. The fact that (i) implies (ii) follows Proposition (5.2.3). Theorem (5.2.9) guarantees that (ii) and(iii) are equivalent. It is more or less a triviality that (iii) implies (iv); see, for example, [224].Finally, Theorems (5.2.17) and (5.2.2) show that (iv) implies (v) and (v) implies (i), respectively.

A self-conformal set Fsatisfies theopen set condition if there exists a non-empty open set $U \subset \Omega$ such that $\varphi_i(U) \subset U$for all i and $\varphi_i(U) \cap \varphi_j(U) = \emptyset$ whenever $i \neq j$. Recall that, by [226], the open set condition is equivalent to

$sup\{\#\Sigma(x,r) : x \in F \text{and}\, r > 0\} < \infty,$

Where

$\Sigma(x,r) = \{i \in \Sigma_*: \text{diam}(\varphi_i(F)) \le r < \text{diam}(\varphi_i - (F)) \text{ and } \varphi_i(F) \cap B(x,r) \neq \emptyset\}$

for all $x \in \mathbb{R}^d$and $r > 0$. Therefore, the open set condition is stronger than the weak separation condition. The pressure P: $[0,\infty) \to \mathbb{R}$, defined by

$$P(s) = \lim_{n\to\infty} \frac{1}{n} log \sum_{i\in\Sigma_n} \|\varphi_i'\|^s,$$

is well-defined, convex, continuous, and strictly decreasing. In fact, there exists unique $s \geq 0$ for which $P(s) = 0$. It is a classical result that if F satisfies the open set condition, then $\dim_H(F) = P^{-1}(0)$; for the latest incarnation of this observation, see [225].

We say that a self-conformal set F has anexact overlap if there exist $i, j \in \Sigma_*$ such that $i \neq j$ and $\varphi_i|_F = \varphi_j|_F$. Observe that if F satisfies the open set condition, then it cannot have exact overlaps. For a self-similar set F in the real line, according to a folklore “dimension drop” conjecture, $\dim_H(F) = min\{1, P^{-1}(0)\}$or otherwise there is an exact overlap. Hochman [222] has verified the conjecture under a mild assumption which is satisfied for example when the associated iterated function system is defined by algebraic parameters; see [222]. To generalize Hochman’s proof for self-conformal sets in the real line seems difficult since the semigroup generated by $C^{1+\alpha}$ maps is simply too large: there is no invariant metric and dimension $d \in \mathbb{N}$ for which there is a smooth injection to $\mathbb{R}^d$, which is bi-Lipschitz to its image in any compact neighbourhood of the identity.

However, the following theorem verifies the conjecture for self-conformal sets in the real line having positive Hausdorff measure. It generalises the corresponding result of Farkas [218] on self-similar sets.

Theorem (5.2.5)[216]: Let $F \subset \mathbb{R}$ be a self-conformal set with $H^s(F) > 0$ for $s = \dim_H(F) < 1$. Then $s = P^{-1}(0)$if and only if there are no exact overlaps.

Proof. If $s = P^{-1}(0)$, then the assumption that $H^s(F) > 0$ together with [137],implies that F satisfies the open set condition and hence, cannot have exact overlaps. If there are no exact overlaps, then, by Corollary (5.2.4), the assumption $H^s(F) > 0$ implies that F satisfies the weak separation condition. Therefore, by [217] (see also [225]), the lack of exact overlaps implies the open set condition and we have $s = P^{-1}(0)$.

For a bounded set $A \subset \mathbb{R}^d$we let

$N_r(A) = min\{k: A \subset \bigcup_{i=1}^{k} B(x_i, r)$ for some $x_1, \dots, x_k \in \mathbb{R}^d\}$

be the least number of balls of radius $r > 0$ needed to cover A.

Lemma (5.2.6)[216]: Let $F \subset \mathbb{R}^d$be a set satisfying the assumptions of Theorem (5.2.7). If $s = \dim_H(F)$,then

$$2^{-s} H_\infty^s(F) r^{-s} \leq N_r(F) \leq D^s r^{-s}$$

for all $r > 0$ and $H^s(F) < (2D)^s$. In particular,$H^s(F) > 0$ if and only if $0 < H^s(F) < \infty$.

Proof. The first claim follows from the definition of H_∞^s, the existence of mappings $g: F \to F \cap B(x,r)$ satisfying (16), and [130]. The second claim follows immediately from the first one.

We are now ready to prove the main theorem.

Theorem (5.2.7)[216]: Let $F \subset \mathbb{R}^d$ be a non-empty compact set and $s = \dim_H(F)$. Suppose that thereis a constant $D \geq 1$such that for each $x \in F$and $0 < r \leq$diam(F) there exists a mapping $g: F \to F \cap B(x,r)$for which

$$D^{-1} r|y-z| \leq |g(y) - g(z)| \leq Dr|y-z| \qquad (16)$$

for all $y, z \in F$. Then there exists a constant $C \geq 1$such that

$$Hs(F \cap B(x,r))6Crs$$

for all $x \in r^d$ and $r > 0$, and

$$H^s(F \cap A) \leq C H_\infty^s(F \cap A)$$

for all $A \subset \mathbb{R}^d$.

Proof.We may assume that $H^s(F) > 0$ since otherwise there is nothing to prove.This of course implies that $H_\infty^s(F) > 0$. Write $C = 2 \cdot 2^{4s} D^{3s} H_\infty^s(F)^{-1}$. To prove the first claim,suppose, for a contradiction, that there exist $x_0 \in \mathbb{R}^d$ and $r_0 > 0$ such that

$$H^s(F \cap B(x_0, r_0)) > \ C r_0^s. \qquad (17)$$

Fix $n \in \mathbb{N}$ and let B_n be a maximal collection of pairwise disjoint closed balls ofradius 2^{-n} centered in . Note that, by [229] and Lemma (5.2.6), we have

$$2^{-2s} H_\infty^s(F) 2^{ns} \leq \# B_n \leq 2^s D^s 2^{ns}\,. \qquad (18)$$

For each $B \in B_n$, let $gB: F \to F \cap B$be as in (16). It follows that each ball B in the packing B_n contains $g_B(F \cap B(x_0, r_0))$, a scaled copy of $F \cap B(x_0, r_0)$. Therefore, recalling (17), we get

$$H^s(g_B(F \cap B(x_0, r_0))) \geq D^{-s} 2^{-ns} H^s(F \cap B(x_0, r_0)) > C D^{-s} 2^{-ns}\, r_0^s$$
$$= \ 2 \cdot 2^{4s-ns} D^{2s} H_\infty^s(F)^{-1} r_0^s \qquad (19)$$

for all $B \in B_n$. Furthermore, since diam$(g_B(F \cap B(x_0, r_0))) \leq D 2^{-n}$diam$(F \cap B(x_0, r_0)) \leq D 2^{-n} 2 r_0 =: \delta_n$, we have

$$H_{\delta_n}^s \left(g_B\big(F \cap B(x_0, r_0)\big) \right) = H_\infty^s \left(g_B\big(F \cap B(x_0, r_0)\big) \right) \leq D^s 2^{-ns} 2^s r_0^s \quad (20)$$

for all $B \in B_n$.

Now (19) and (18) imply

$$\sum_{B \in B_n} H^s \left(g_B\big(F \cap B(x_0, r_0)\big) \right) \geq \# B_n 2^{4s-ns+1} D^{2s} H_\infty^s(F)^{-1} r_0^s \geq 2 \cdot 2^{2s} D^{2s} r_0^s \quad (21)$$

and (20) and (18) give

$$\sum_{B\in B_n} H^s_{\delta_n}\left(g_B\big(F\cap B(x_0,r_0)\big)\right) \le \#B_n D^s 2^{-ns} 2^s r_0^s \le 2^{2s} D^{2s} r_0^s \tag{22}$$

Since, by the fact that the sets $g_B(F\cap B(x_0,r_0))$ are H^s-measurable and (21),

$$H^s(F) = H^s(F\backslash \bigcup_{B\in B_n} g_B\,(F\cap B(x_0,r_0))) + \sum_{B\in B_n} H^s\,(g_B(F\cap B(x_0,r_0)))$$

$$\ge H^s(F\backslash \bigcup_{B\in B_n} g_B\,(F\cap B(x_0,r_0))) + 2\cdot 2^{2s}D^{2s}r_0^s$$

and, by (22),

$$H^s_{\delta_n}(F) \le H^s_{\delta_n}(F\backslash \bigcup_{B\in B_n} g_B\,(F\cap B(x_0,r_0))) + \sum_{B\in B_n} H^s_{\delta_n}\,(g_B(F\cap B(x_0,r_0)))$$

$$\le H^s(F\backslash \bigcup_{B\in B_n} g_B\,(F\cap B(x_0,r_0))) + \ 2^{2s}D^{2s}r_0^s,$$

we conclude that

$$H^s(F) - H^s_{\delta_n}(F) \ge 2\cdot 2^{2s}D^{2s}r_0^s - 2^{2s}D^{2s}r_0^s =\ 2^{2s}D^{2s}r_0^s > 0.$$

This is a contradiction since the lower bound is independentofn.To show the second claim, let $A\subset\mathbb{R}^d$ and fix $\varepsilon\ >0$. Choose a countable collection$\{B(x_i,r_i)\}$iof balls covering $F\cap A$such that $\sum_i(2r_i)^s\ \le H^s_\infty(F\cap A)+\varepsilon$. Applying the first claim, we get

$$H^s(F\cap A) \le \sum_i H^s\left(F\cap B(x_i,r_i)\right) \le C\sum_i (2r_i)^s \le C(H^s_\infty(F\cap A)+\varepsilon)$$

which finishes the proof.

The following lemma is standard.

Lemma (5.2.8)[216]: If $\{\varphi_i\}_{i=1}^N$is a conformal iterated function system, then there exists abounded open convex set $V\subset\mathbb{R}^d$ such that $\varphi_i(\bar V)\subset V\subset \bar V\subset\Omega$ for all $i\in\{1,\ldots,N\}$. Furthermore, if $F\subset V$ isthe associated self-conformal set containing at least two points, then there exist a constant $K\ge 1$ such that

$$K^{-1}\|\varphi_i'\||x-y| \le |\varphi_i(x)-\varphi_i(y)| \le \|\varphi_i'\||x-y| \tag{23}$$

for all $x,y\in V$and $i\in\Sigma_*$,

$$\frac{1}{diam(F)}diam\big(\varphi_i(F)\big) \le \|\varphi_i'\| \le \frac{K}{diam(F)}diam(\varphi_i(F)) \tag{24}$$

for all $i\in\Sigma_*$, and

$$K^{-2}\|\varphi_i'\|\|\varphi_j'\| \le \|\varphi_{ij}'\| \le \|\varphi_i'\|\|\varphi_j'\| \tag{25}$$

for all $i,j\in\Sigma_*$.

Proof. Write $d=\mathrm{dist}(\bigcup_{i=1}^N\overline{\varphi_\iota(\Omega)},\mathbb{R}^d\backslash\Omega)/4>0$ and let U_i be the opend-neighbourhood of $\varphi_i(\Omega)$.It is easy to see that

$$\mathrm{dist}(U_i,\mathbb{R}^d\backslash\Omega)\ge 2d \tag{26}$$

for all $i\in\{1,\ldots,N\}$. Indeed, if this was not true, then there are $x\in U_i$ and $w\in\mathbb{R}^d\backslash\Omega$ such that $|x-w|<2d$. As $x\in U_i$, there is $z\in\varphi_i(\Omega)$ such that $|z-x|<\ d$. Therefore, the contradiction $4d\le|z-w|\le|z-x|+|x-w|<3d$ we obtain proves (26).

Define V to be the convex hull of $\bigcup_{i=1}^N U_i$. Let us show that

$\operatorname{dist}(V, \mathbb{R}^d \setminus \Omega) \geq 2d.$ (27)

If this was not the case, then there are $z \in V$and $w \in \mathbb{R}^d \setminus \Omega$ such that $|z - w| < 2d$. We may assume that $z \notin \bigcup_{i=1}^{N} U_i$ since otherwise the contradiction follows immediately from (26). Let $x, y \in \bigcup_{i=1}^{N} U_i$ be such that z is a convex combination of x and y, which we denote by writing $z \in [x, y]$. Let z' be the closest point to w in the line containing the segment $[x, y]$. If $z' \notin [x, y]$,then there is $v \in \{x, y\}$ such that $v \in [z, z']$. As $v \in \bigcup_{i=1}^{N} U_i$ and $|v - w| \leq |z - w| < 2d$, we getthe contradiction again from (26). We may thus assume that $z' \in [x, y] \setminus \bigcup_{i=1}^{N} U_i$. Notice that $(z' - w) \perp (y - w)$ and $|z' - w| \leq |z - w| < 2d$. Let $L_w = \{w + t(y - x) : t \in \mathbb{R}\}$be the line parallel to$[x, y]$ going through w. Choose $x', y' \in L_w$so that $(x - x') \perp (y - x)$ and $(y - y') \perp (y - x)$. It followsthat $w \in [x', y']$ and $|x - x'| = |y - y'| = |z' - w| < 2d$. By (26), we therefore have $x', y' \in \Omega$.But since Ω is convex, also $w \in \Omega$ which is a contradiction. Therefore, (27) holds and it is thusevident that $\bar{V} \subset \Omega$. Hence, $\varphi_i(\bar{V}) \subset \varphi_i(\Omega) \subset U_i \subset V$ for all $i \in \{1, \ldots, N\}$.

By [132], the H¨older continuity of the differentials implies the existence of a constant $K_0 \geq 1$ for which

$$|\varphi_i'(y)| \leq K_0 |\varphi_i'(x)| \tag{28}$$

for all $x, y \in \Omega$ and $i \in \Sigma_*$. Fix $x, y \in \Omega$ and define $x_t = (1 - t)y + tx$ for all $t \in [0,1]$. Note that,by convexity of Ω, $x_t \in \Omega$ for all $t \in [0,1]$. The fundamental theorem of calculus implies that thereexists $t_0 \in [0,1]$ such that

$$|\varphi_i(x) - \varphi_i(y)| = \left| \int_0^1 \varphi_i'(x_t) \frac{d}{dt} x_t dt \right| \leq |\varphi_i'(x_{t_0})||x - y|, \tag{29}$$

which gives the right-hand side inequality in (23). To showthe other inequality, fix $x, y \in V$.If $[\varphi_i(x), \varphi_i(y)] \cap \partial\varphi_i(\Omega) \neq \emptyset$, we choose $z \in [\varphi_i(x), \varphi_i(y)]$ to be so close to $\partial\varphi_i(\Omega)$ such that $\varphi_i^{-1}([\varphi_i(x), z]) \subset \Omega$ and $|x - \varphi_i^{-1}(z)| > d$, which is possible by (27). If $[\varphi_i(x), \varphi_i(y)] \subset \varphi_i(\Omega)$, then we write $z = \varphi_i(y)$. Define $z_t = (1 - t)z + t\varphi_i(x)$ for all $t \in [0,1]$. As above, there exists $t_1 \in [0,1]$ such that

$$\left|\varphi_i^{-1}(\varphi_i(x)) - \varphi_i^{-1}(z)\right| = \left| \int_0^1 (\varphi_i^{-1})'(z_t) \frac{d}{dt} z_t dt \right| \leq |(\varphi_i^{-1})'(z_{t_1})||\varphi_i(x) - z|$$

Yielding

$$\begin{aligned} |\varphi_i(x) - \varphi_i(y)| \geq |\varphi_i(x) - z| &\geq \left|(\varphi_i^{-1})'(z_{t_1})\right|^{-1} |x - \varphi_i^{-1}(z)| \\ &\geq \left|(\varphi_i^{-1})'(z_{t_1})\right|^{-1} |x - y| \min\{1, \frac{d}{diam(V)}\}. \end{aligned} \tag{30}$$

Note that, by conformality and (28), $\inf_{w \in \varphi_i(\Omega)} \left|(\varphi_i^{-1})'(w)\right|^{-1} = \inf_{w \in \Omega} |\varphi_i'(w)| \geq K_0^{-1} \|\varphi_i'\|$.Therefore, the left-hand side inequality in (23) follows from (30) by setting $K = K_0 \max\{1, \operatorname{diam}(V)/d\}$.Since both (24) and (25) follow straightforwardly from (23), we have finished the proof.

The properties (23)–(25) are characteristic for conformal iterated function systems and they are used as a starting point in generalising self-conformality into metric spaces; see [226] and [230].

Theorem (5.2.9)[216]: Let $F \subset \mathbb{R}^d$ be a self-conformal set and $s = \dim_H(F)$. Then there exists a constant $C \geq 1$such that

$$H^s(F \cap A) \leq C H^s_\infty(F \cap A)$$

for all $A \subset \mathbb{R}^d$. Furthermore,$H^s(F) > 0$ if and only if F is Ahlforss-regular.

Proof. We may clearly assume that Fcontains at least two points. Let $x \in F$and $0 < r < diam(F)$. Pick $i \in \Sigma$ such that $\lim_{n\to\infty} \varphi_{i|_n}(x_0) = x$ for all $x_0 \in V$ and choose $n \in \mathbb{N}$ for which $\varphi_{i|_n}(F) \subset B(x,r)$ but $\varphi_{i|_{n-1}}(F)\backslash B(x,r) \neq \emptyset$. Note that the latter property implies $\mathrm{diam}(\varphi_{i|_{n-1}}(F)) \geq r$. By (23) and (24), we have

$$| \varphi_{i|_n}(y) - \varphi_{i|_n}(z)| \geq K^{-2} \left\|\varphi'_{i|_{(n-1)}}\right\| \min_{i\in\{1,\dots,N\}} \|\varphi'_i\| \ |y-z|$$

$$\geq \frac{\min_{i\in\{1,\dots,N\}} \|\varphi'_i\|}{K^2 diam(F)} diam(\varphi_{i|_{n-1}}(F))|y-z| \geq \frac{\min_{i\in\{1,\dots,N\}} \|\varphi'_i\|}{K^2 diam(F)} r|y-z|$$

And

$$\left|\varphi_{i|_n}(y) - \varphi_{i|_n}(z)\right| \leq \frac{K}{diam(F)} diam(\varphi_{i|_n}(F))|y-z| \leq \frac{2K}{diam(F)} r|y-z|$$

for all $y, z \in F$. By setting

$$D = max\{1, \frac{K^2 diam(F)}{\min_{i\in\{1,\dots,N\}} \|\varphi'_i\|}, \frac{2K}{diam(F)}\},$$

we have thus shown that for each $x \in F$ and $0 < r < diam(F)$ there exist $i \in \Sigma$ and $n \in \mathbb{N}$ such that $\varphi_{i|_n}(F) \subset F \cap B(x,r)$ and

$$D^{-1} r|y-z| \leq \left|\varphi_{i|_n}(y) - \varphi_{i|_n}(z)\right| \leq Dr|y-z| \qquad (31)$$

for all $y, z \in F$. Theorem (5.2.9) follows now immediately from Theorem (5.2.7) and Proposition (5.2.1).

A sub self-conformal set is a non-empty compact set $E \subset F$ which satisfies $E \subset \bigcup_{i=1}^{N} \varphi_i(E)$,where F is the associated invariant set. Note that sub self-conformal sets are contained in theinvariant set when mapped under φ_i, that is,$\varphi_i(E) \subset F$. It is again straightforward to check that Lemma (5.2.8) and Theorem (5.2.9) hold for sub self-similar sets. Generally, the images of graph-directed self-conformal sets are not contained in themselves under φ_i for all i and it is easy to find examples such that the sets F_i are not sub self-conformal. However, some prefer todefine a single graph-directed set using subshifts of finite type. In our notation this amounts to considering $F = \bigcup_{i=1}^{N} F_i$. For such Fwe have $\varphi_i(F) \subset F$and thus F is a sub self-conformal set .

For both cases above we have omitted detailed proofs to avoid cumbersome notation of M-admissible words and arbitrary subsets.

The proof of Theorem (5.2.2) is split into two parts, Propositions (5.2.10) and (5.2.14).

Proposition (5.2.10)[216]: Let $F \subset \mathbb{R}^d$be a self-conformal set. If Fsatisfies the identity limit criterion,then it satisfies the weak separation condition.

Proof. We prove that the failure of the weak separation condition implies the failure of the identity limit criterion. Our goal, therefore, is to show that for every $\varepsilon > 0$ there are $i, j \in \Sigma_*$such that

$$0 < \sup_{x\in F} |\varphi_i(x) - \varphi_j(x)| \leq \varepsilon max\{\|\varphi'_i\|, \|\varphi'_j\|\}. \qquad (32)$$

Let $K \geq 1$ be as in Lemma (5.2.8), fix $\varepsilon > 0$, and choose

$$0 < \delta \leq min\{\varepsilon \left(4 + \frac{2K^2 diam(F)}{\min_{i\in\{1,\dots,N\}} diam\big(\varphi_i(F)\big)}\right)^{-1}, \frac{1}{2} diam(F), 1\}. \qquad (33)$$

Let $\{B(x_i,\delta)\}_{i=1}^{M}$ be a maximal collection of pairwise disjoint closed balls centered at F. Observe that if $\delta \le \frac{1}{2}diam(F)$, then $M \le \mathrm{diam}(F)^d\delta^{-d}$.

Since the weak separation condition does not hold, there exist a point $z \in F$ and a radius $r > 0$ such that

$$\#\Phi(z,r) > (5^d\delta^{-d})^M.$$

Note that $\varphi(F) \subset B(z,2r)$ for all $\varphi \in \Phi(z,r)$. Let $\{B_j\}_{j=1}^{L}$ be a minimal cover of $B(z,2r)$ of balls of radius δr centered at $B(z,2r)$. Observe that if $\delta \le 1$, then $L \le 5^d\delta^{-d}$. Moreover, for each $\varphi \in \Phi(z,r)$ there is a map $\psi: \{1,\dots,M\} \to \{1,\dots,L\}$ given by $\psi(i) = j$, where $j \in \{1,\dots,L\}$ such that $\varphi(x_i) \in B_j$. Note that there can be at most L^M many different maps ψ. Since $\#\Phi(z,r) > L^M$, there have to be two maps $\varphi_i, \varphi_j \in \Phi(z,r)$ such that

$$\varphi_i|_F \ne \varphi_j|_F \text{ and for each } i \in \{1,\dots,M\} \text{ it holds that } \varphi_i(x_i), \varphi_j(x_i) \in B_j \quad (34)$$

for some $j \in \{1,\dots,L\}$.

Let $i,j \in \Sigma_*$ satify (34). Fix $x \in F$ and choose $x_0 \in \{x_i\}_{i=1}^{M}$ such that $|x - x_0| \le |x - x_i|$ for all $i \in \{1,\dots,M\}$. Note that, since $\{B(x_i,2\delta)\}_{i=1}^{M}$ covers F, we have $|x - x_0| \le 2\delta$. It follows fromthe triangle inequality, Lemma (5.2.8), and (33) that

$$\begin{aligned} |\varphi_i(x) - \varphi_j(x)| &\le |\varphi_i(x) - \varphi_i(x_0)| + |\varphi_i(x_0) - \varphi_j(x_0)| + |\varphi_j(x_0) - \varphi_j(x)| \\ &\le \|\varphi_i'\||x - x_0| + 2\delta r + \|\varphi_j'\||x_0 - x| \\ &\le 2\delta max\{\|\varphi_i'\|, \|\varphi_j'\|\}\left(2 + \frac{K^2 diam(F)}{\min_{i\in\{1,\dots,N\}} diam(\varphi_i(F))}\right) \le \varepsilon max\{\|\varphi_i'\|, \|\varphi_j'\|\}. \end{aligned}$$

This proves (32) and finishes the proof.

Before going into Proposition (5.2.14), we prove three technical lemmas. We say that F is uniformly perfect if there exists a constant $H \ge 1$ such that

$$F \cap B(x,r)\backslash B(x,r/H) \ne \emptyset \quad (35)$$

for all $x \in F$ and $0 < r < diam(F)$.

Lemma (5.2.11)[216]: Let $F \subset \mathbb{R}^d$ be a self-conformal set. Then thed following three conditionsare equiva-lent:

(i) F is uniformly perfect,

(ii) $\dim_H(F) > 0$,

(iii) F contains at least two points.

Proof. If F is uniformly perfect, then [223] shows that $\dim_H(F) > 0$, which clearly implies that F contains at least two points. Therefore, it suffices to show that (iii) implies (i). Let $K \ge 1$ be as in Lemma (5.2.8) and

$$H = \frac{3K^3}{\min_{i\in\{1,\dots,N\}}\|\varphi_i'\|} + 1.$$

Let $x \in F$ and $0 < r < diam(F)$. Since F contains at least two points, there exists a point $y \in F$ such that $y \ne x$. Let $i \in \Sigma$ be such that $\lim_{n\to\infty} \varphi_{i|_n}(y) = x$. Write $d = |x - y| > 0$ and choose $n_0 \in \mathbb{N}$ such that $diam(\varphi_{i|_{n_0}}(F)) < \frac{d}{2}$ and

$$\frac{\frac{3}{2}dK^2 + K^2 diam(F)\left(\max_{i\in\{1,\dots,N\}}\|\varphi_i'\|\right)^{n_0}}{\left(\frac{1}{2}dK^{-1} - diam(F)\left(\max_{i\in\{1,\dots,N\}}\|\varphi_i'\|\right)^{n_0}\right)\min_{i\in\{1,\dots,N\}}\|\varphi_i'\|} \le H$$

Choose $n \geq n_0$ such that

$$\|\varphi'_{i|_n}\|\left(\frac{3}{2}d + diam(F)\left\|\varphi'_{i|_{n_0}}\right\|\right) \leq r < \|\varphi'_{i|_{n-1}}\|\left(\frac{3}{2}d + diam(F)\left\|\varphi'_{i|_{n_0}}\right\|\right) \quad (36)$$

and note that it suffices to prove the claim for all $0 < r < r_0$, where $0 < r_0 < \mathrm{diam}(F)$. Let $z \in \varphi_{i|_{n_0}}(F)$ and observe that $x, \varphi_{i|_n}(z) \in \varphi_{i|_n i|_{n_0}}(F)$.Therefore, by (23) and (24),

$$\begin{aligned}\left|\varphi_{i|_n}(y) - x\right| &\leq \left|\varphi_{i|_n}(y) - \varphi_{i|_n}(z)\right| + \left|\varphi_{i|_n}(z) - x\right| \\ &\leq \left\|\varphi'_{i|_n}\right\||y - z| + diam(F)\left\|\varphi'_{i|_n i|_{n_0}}\right\| \leq \left\|\varphi'_{i|_n}\right\|(\frac{3}{2}d + diam(F)\left\|\varphi'_{i|_{n_0}}\right\|) \\ &\leq r \quad (37)\end{aligned}$$

And

$$\begin{aligned}\left|\varphi_{i|_n}(y) - x\right| &\geq \left|\varphi_{i|_n}(y) - \varphi_{i|_n}(z)\right| - \left|\varphi_{i|_n}(z) - x\right| \\ &\geq K^{-1}\left\|\varphi'_{i|_n}\right\||y - z| - diam(F)\left\|\varphi'_{i|_n i|_{n_0}}\right\| \\ &\geq \|\varphi'_{i|_n}\|(\frac{1}{2}dK^{-1} - diam(F)\|\varphi'_{i|_{n_0}}\|). \quad (38)\end{aligned}$$

By (25), we have $\|\varphi'_{i|_{n-1}}\| \leq K^2\left(\min_{i\in\{1,\dots,N\}}\|\varphi'_i\|\right)^{-1}\|\varphi'_{i\,|_n}\|$and hence, by (36), the choice of $H \geq 1$, and (38),

$$\begin{aligned} r &< \left\|\varphi'_{i|_n}\right\| \frac{\frac{3}{2}dK^2 + K^4\left(\min_{i\in\{1,\dots,N\}}\|\varphi'_i\|\right)^{-1} diam(F)\left\|\varphi'_{i|_n}\right\|}{\min_{i\in\{1,\dots,N\}}\left\|\varphi'_i\right\|} \\ &\leq H\|\varphi'_{i|_n}\|(\frac{1}{2}dK^{-1} - diam(F)\|\varphi'_{i|_{n_0}}\|) \leq D|\varphi_{i|_n}(y) - x|. \quad (39)\end{aligned}$$

Therefore, by (37) and (39),

$$\varphi_{i|_n}(y) \in B(x, r)\backslash B(x, r/H)$$

and we conclude that F is uniformly perfect.

Lemma (5.2.12)[216]: Let$\{\varphi_i\}_{i=1}^N$be a conformal iterated function system. Then there are constants $\alpha, c > 0$ such that

$$|\varphi'_i(x) - \varphi'_i(y)| \leq c\|\varphi'_i\||x - y|^\alpha$$

for all $x, y \in V$ and $i \in \Sigma_*$.

Proof. Let $x, y \in V$ and fix $i = i_1 \cdots i_n \in \Sigma_n$for some $n \in \mathbb{N}$. Write $\sigma^j(i_1 \cdots i_n) = i_{j+1} \cdots i_n$,

$x_j = \varphi_{\sigma^j(i)}(x)$ and $y_j = \varphi_{\sigma^j(i)}(y)$,

and note that, by the chain rule,$\varphi'_{i|_j}(x_j) = \varphi'_{i_1}(x_1) \cdots \varphi'_{i_j}(x_j)$ for all $j \in \{1, \dots, n\}$. We interpret $x_n = x$ and $y_n = y$. Write also

$$d_j = \varphi'_{i_j}(x_j) - \varphi'_{i_j}(y_j)$$

and observe that, by (15), there exist constants $\alpha, c > 0$ such that

$$|d_j| \leq c|x_j - y_j|^\alpha \leq c\left\|\varphi'_{\sigma^j(i)}\right\|^\alpha |x - y|^\alpha \leq c\left(\max_{i\in\{1,\dots,N\}}\|\varphi'_i\|\right)^{\alpha(n-j)}|x - y|^\alpha \quad (40)$$

for all $j \in \{1, \dots, n\}$. Since

$$\varphi'_{\sigma^{j-1}(i)}(x) - \varphi'_{\sigma^{j-1}(i)}(y) = \varphi'_{i_j}(x_j)(\varphi'_{\sigma^j(i)}(x) - \varphi'_{\sigma^j(i)}(y)) + d_j\varphi'_{\sigma^j(i)}(y)$$

for all $j \in \{1, \dots, n\}$, we recursively see that

$$\varphi_i'(x)-\varphi_i'(y) = \varphi_{i_1}'(x_1)(\varphi'_{\sigma(i)}(x)-\varphi'_{\sigma(i)}(y))+d_1\varphi'_{\sigma(i)}(y)$$
$$=\varphi_{i|_2}'(x_2)(\varphi'_{\sigma^2(i)}(x)-\varphi'_{\sigma^2(i)}(y))+\varphi_{i_1}'(x_1)d_2\varphi'_{\sigma^2(i)}(y)\ +d_1\varphi'_{\sigma(i)}(y)=\cdots$$
$$=\sum_{j=1}^{n}\varphi'_{i|_{j-1}}\left(x_j-1\right)d_j\varphi'_{\sigma^j(i)}(y). \quad (41)$$
Observe that, by (25), we have $\|\varphi'_{i|_{j-1}}\|\|\varphi'_{\sigma^j(i)}\|\le K^2\|\varphi_i'\|$ for all $j\in\ \{1,\dots,n\}$and hence, by(41) and (40),|
$$|\varphi_i'(x)-\varphi_i'(y)|\le\sum_{j=1}^{n}\left|\varphi'_{i|_{j-1}}(x_j-1)\right|\ \left|d_j\right|\ \left|\varphi'_{\sigma^j(i)}(y)\right|\le K^2\|\varphi_i'\|\sum_{j=1}^{n}\left|\,d_j\,\right|$$
$$\le\frac{cK^2}{1-\max\limits_{i\in\{1,\dots,N\}}\left\|\varphi_i'\right\|^{\alpha}}\ \ \|\varphi_i'\|\ |x-y|^{\alpha}$$
as claimed.

Lemma (5.2.13)[216]: Let $\{\varphi_i\}_{i=1}^{N}$be a conformal iterated function system and $F\subset\mathbb{R}^d$ the associated self-conformal set containing at least two points. If F does not satisfy the identity limit criterion, then there exists a constant $C\ge 1$such that for every $\varepsilon\ >0$ there are $0<\delta<\varepsilon$ and $i,j\in\Sigma_*$ for which
$$C^{-1}\delta max\{\|\varphi_i'\|,\|\varphi_j'\|\}\le|\varphi_i(x)-\varphi_j(x)|\le C\delta min\{\|\varphi_i'\|,\|\varphi_j'\|\}$$
for all $x\in V$.

Proof. By the assumption, for every $\varepsilon>0$ there are $i,j\in\Sigma_*$such that
$$0<\sup_{x\in F}|\varphi_i(x)-\varphi_j(x)|\le\varepsilon\max\{\|\varphi_i'\|,\|\varphi_j'\|\}. \quad (42)$$
Let $\alpha,c>0$ be as in Lemma (5.2.12). Recalling that $V\supset F$ is open, we see that there exists $\varepsilon_0>0$ such that $\varepsilon_0^{1/1+\alpha}<diam(F)$ and $B(x,\varepsilon_0^{1/1+\alpha})\subset V$ for all $x\in F$. Fix $0<\varepsilon<\varepsilon_0$ and let $i,j\in\Sigma_*$ be such that (42) holds. By compactness of F, the supremum in (42) is attained by some $x_0\in F$.To simplify notation, write $f(x)=\varphi_i(x)-\varphi_j(x)$ and $\delta=|f(x_0)|\ /\max\{\|\varphi_i'\|,\|\varphi_j'\|\}$. Note that
$$|f(x)|\le|f(x_0)|=\delta\max\{\|\varphi_i'\|,\|\varphi_j'\|\}\le\varepsilon\max\{\|\varphi_i'\|,\|\varphi_j'\|\} \quad (43)$$
for all $x\in F$ and, in particular, $0<\delta\le\varepsilon$.
By the triangle inequality and Lemma (5.2.12), we obtain
$$\left||f'(y)|-|f'(x_0)|\right|\le|f'(y)-f'(x_0)|\le|\varphi_i'(y)-\varphi_i'(x_0)|+\left|\varphi_j'(y)-\varphi_j'(x_0)\right|$$
$$\le c\left(\|\varphi_i'\|+\left\|\varphi_j'\right\|\right)|y-x_0|^{\alpha}\le 2c\max\left\{\|\varphi_i'\|,\left\|\varphi_j'\right\|\right\}|y-x_0|^{\alpha} \quad (44)$$
for all $y\in V$. Let $H\ge 1$ be as in (35). We will next show that
$$|f'(x_0)|\le(3H+\ 2c)^{\delta^{\alpha}/1+\alpha}\max\{\|\varphi_i'\|,\|\varphi_j'\|\}. \quad (45)$$
To prove (45), we assume the opposite inequality for a contradiction. Since F contains at least two points, it follows from Lemma (5.2.11) that F is uniformly perfect and there exists a point $z\in F\cap B(x_0,\delta$^(1 / (1 + α)) / $B(x_0,\delta^{(1/1+\alpha)}/H)$. By convexity of V, the line connecting x_0 and z is contained in V and hence, $z_t=\ (1-t)x_0+tz\in V$ for all $t\in[0,1]$. Recalling (44), we have
$$|f'(z_t)-f'(x_0)|\le 2cmax\left\{\|\varphi_i'\|,\left\|\varphi_j'\right\|\right\}|z_t-x_0|^{a}$$
$$\le 2c\delta^{\alpha/1+\alpha}max\{\|\varphi_i'\|,\|\varphi_j'\|\}. \quad (46)$$
Define unit vectors u and v by setting $u=(z-x_0)\ /\ |z-x_0|$ and $v=f'(x_0)u/|f'(x_0)|$. As $f'(y)$ is a similarity for all $y\in V$, we have $\langle f'(x_0)u,v\rangle=|f'(x_0)u|^2\ /\ |f'(x_0)|=$

$|f'(x_0)|$ and $|f'(x_0)u - f'(z_t)u| = |f'(x_0) - f'(z_t)|$. Therefore, by the Cauchy-Schwarz inequality, the assumptionthat (45) does not hold, and (46), we have
$\langle f'(z_t)u, v\rangle = \langle f'(x_0)u, v\rangle - \langle f'(x_0)u - f'(z_t)u, v\rangle \geq |f'(x_0)| - |f'(x_0) - f'(z_t)| \geq 3H\delta^{\alpha/(1+\alpha)}\, max\,\{\|\varphi_i'\|, \|\varphi_j'\|\}$(47)
for all $t \in [0,1]$. Since, by conformality, $\langle \nabla\langle f(y), v\rangle, u\rangle = \langle f'(y)u, v\rangle$ for all $y \in V$, the fundamental theorem of calculus and the multivariate chain rule imply

$$\langle f(z), v\rangle - \langle f(x_0), v\rangle = \int_0^1 \frac{d}{dt}\langle f(z_t), v\rangle\, dt = \int_0^1 \langle \nabla\langle f(z_t), v\rangle, \frac{d}{dt} z_t\rangle\, dt$$

$$= |z - x_0| \int_0^1 \langle \nabla\langle f(z_t), v\rangle, u\rangle\, dt = |z - x_0| \int_0^1 \langle f'(z_t)u, v\rangle\, dt. \qquad (48)$$

Hence, by the Cauchy-Schwarz inequality, (48), and (47),

$$|f(z)| \geq \langle f(z), v\rangle = \langle f(x_0), v\rangle + |z - x_0| \int_0^1 \langle f'(z_t)u, v\rangle\, dt$$

$$\geq -|f(x_0)| + \frac{\delta^{1/1+\alpha}}{H} 3H\delta^{\alpha/1+\alpha} max\{\|\varphi_i'\|, \|\varphi_j'\|\} = 2\delta max\{\|\varphi_i'\|, \|\varphi_j'\|\}$$

$$> |f(x_0)|.$$

As this contradicts (43), i.e. the maximality of x_0, we have shown (45).
Combining (44) and (45), we see that

$$|f'(y)| \leq |f'(x_0)| + 2c max\{\|\varphi_i'\|, \|\varphi_j'\|\}|y - x_0|^\alpha$$

$$\leq (3H + 4c)\delta^{\alpha/1+\alpha} max\{\|\varphi_i'\|, \|\varphi_j'\|\} \qquad (49)$$

for all $y \in B(x_0, \delta^{1/1+\alpha})$. Write $r = \delta^{1/1+\alpha} / (6H + 8c) \leq \delta^{1/1+\alpha}$, fix $x \in B(x_0, r)$, and definey $t = (1 - t)x_0 + tx$ for all $t \in [0,1]$. By the fundamental theorem of calculus, there exists $y \in B(x_0, r)$ such that, by (49),

$$\big||f(x)| - |f(x_0)|\big| \leq |f(x) - f(x_0)| = \left|\int_0^1 f'(y_t)\frac{d}{dt} y_t dt\right| \leq |f'(y)||x - x_0|$$

$$\leq (3H + 4c)^{\delta^{\alpha/(1+\alpha)}} max\{\|\varphi_i'\|, \|\varphi_j'\|\} r = \frac{1}{2}\delta max\{\|\varphi_i'\|, \|\varphi_j'\|\}. \qquad (50)$$

Now (43) and (50) imply

$$\frac{1}{2}\delta max\{\|\varphi_i'\|, \|\varphi_j'\|\} = |f(x_0)| - \frac{1}{2}\delta\, max\{\|\varphi_i'\|, \|\varphi_j'\|\} \leq |f(x)|$$

$$\leq \delta \max\{\|\varphi_i'\|, \|\varphi_j'\|\} \qquad (51)$$

for all $x \in B(x_0, r)$.
Let $k \in \Sigma$ be such that $\lim_{n\to\infty} \varphi_{k|_n}(x) = x_0$ for any $x \in V$and choose $n \in \mathbb{N}$ such that $\mathrm{diam}(\varphi_{k|_n}(V)) < r$ and $\mathrm{diam}(\varphi_{k|_{n-1}}(V)) \geq r$. Note that $\varphi_{k|_n}(V) \subset B(x_0, r)$ and hence (51)holds for all points in $\varphi_{k|_n}(V)$. By (25), we have $K^{-2}\|\varphi_h'\|\|\varphi_{k|_n}'\| \leq \|\varphi_{hk|_n}'\| \leq \|\varphi_h'\|\|\varphi_{k\,|_n}'\|$ for all $h \in \Sigma_*$. Observe that, by (24),

$$\frac{\left(K^{-2} \min_{i\in\{1,\dots,N\}}\|\varphi_i'\|\right)}{diam(F)} r \leq \|\varphi_{k|_n}'\| \leq \frac{K}{diam(F)} r$$

Therefore, by (51),

$$\left|f\left(\varphi_{k|_n}(x)\right)\right| \leq \delta max\{\|\varphi_i'\|, \|\varphi_j'\|\} \leq \delta K^2 \|\varphi_{k|_n}'\|^{-1} max\left\{\left\|\varphi'_{ik|_n}\right\|, \|\varphi_{jk|_n}'\|\right\}$$
$$\leq \frac{\delta K^4 diam(F)}{r min_{i\in\{1,\dots,N\}}\|\varphi_i'\|} max\{\|\varphi_{ik|_n}'\|, \|\varphi_{jk|_n}'\|\}$$
$$\leq \delta^{\alpha/1+\alpha} \frac{K^4(4H + 2^{2+\alpha}) diam(F)}{\min\limits_{i\in\{1,\dots,N\}}\|\varphi_i'\|} max\{\|\varphi_{ik|_n}'\|, \|\varphi_{jk|_n}'\|\}$$

And

$$|f(\varphi_{k|_n}(x))| \geq \delta^{\alpha/1+\alpha} K^{-1}(2H + 2^{1+\alpha})\, diam(F)\, max\{\|\varphi_{ik|_n}'\|, \|\varphi_{jk|_n}'\|\}$$

for all $x \in V$. Writing

$$C = max\left\{\frac{K^{4(4H+2^{2+\alpha})diam(F)}}{\min\limits_{i\in\{1,\dots,N\}}\|\varphi_i'\|}, \frac{K}{(2H + 2^{1+\alpha})diam(F)}\right\},$$

we have thus shown that

$$C^{-1}\delta^{\alpha/1+\alpha} max\{\|\varphi_{ik}'|_n)\|, \|\varphi_{jk|_n}'\|\} \leq |\varphi_{ik|_n}(x) - \varphi_{jk|_n}(x)|$$
$$\leq C\delta^{\alpha/1+\alpha} max\{\|\varphi_{ik|_n}'\|, \|\varphi_{jk|_n}'\|\} \qquad (52)$$

for all $x \in V$.

To finish the proof, fix $0 < \varepsilon' < diam(F)/(4KC)$ and take $0 < \varepsilon < \varepsilon_0$ such that $\varepsilon^{\alpha/1+\alpha} < \varepsilon'$. Let $0 < \delta \leq \varepsilon, i' = ik|_n$, and $j' = jk|\_n$ be so that (52) holds and define $\delta' = \delta^{\alpha/1+\alpha} \leq \varepsilon^{\alpha/1+\alpha} < \varepsilon'$. Notice that, by (24), the triangle inequality, and (52),

$\|\varphi_{j'}'\| \leq \frac{K}{diam(F)} diam\left(\varphi_{(j')}(F)\right) \leq \frac{K}{diam(F)}\Big(diam\big(\varphi_{i'}(F)\big) +$
$2C\delta' max\left\{\|\varphi_{i'}'\|, \|\varphi_{j'}'\|\right\}\Big) \leq K\|\varphi_{i'}'\| + \frac{2KC\varepsilon'}{diam(F)} max\{\|\varphi_{i'}'\|, \|\varphi_{j'}'\|\}$.

Therefore, if $\|\varphi_{j'}'\| \geq \|\varphi_{i'}'\|$, we have

$$\|\varphi_{j'}'\| \leq \frac{K diam(F)}{diam(F) - 2KC\varepsilon'}\|\varphi_{i'}'\| \leq 2K\|\varphi_{i'}'\|$$

and similarly the other way around. By (52), we now have

$$C^{-1}\delta' max\{\|\varphi_{i'}'\|, \|\varphi'\_(j')\|\} \leq |\varphi_{i'}(x) - \varphi_{j'}(x)| \leq 2KC\delta' min\{\|\varphi_{i'}'\|, \|\varphi_{j'}'\|\}$$

for all $x \in V$, which is what we wanted to show.

Proposition (5.2.14)[216]: Let $F \subset \mathbb{R}^d$be a self-conformal set containing at least two points. If F satisfies the weak separation condition, then it satisfies the identity limit criterion.

Proof. Suppose to the contrary that F does not satisfy the identity limit criterion. Let $C \geq 1$be as in Lemma (5.2.13) and $K \geq 1$ as in Lemma (5.2.8). For each $q \in \mathbb{N}$ write $l(q) = 1 \cdots 1 \in \Sigma_q$and $\varepsilon(q) = \frac{2}{3}CK^{-2}\|\varphi_{l(q)}'\| diam(F)/q > 0$. Choose $q \in \mathbb{N}$ to be the smallest integer for which

$$\frac{K}{diam(F)} \max_{j\in\Sigma_q} diam\left(\varphi_j(F)\right) < \frac{3q-2}{3q+2} = \frac{C\|\varphi_{l(q)}'\| diam(F) - K^2\varepsilon(q)}{C\|\varphi_l'(q)\| diam(F) + K^2\varepsilon(q)}. \quad (53)$$

We will prove that F does not satisfy the weak separation condition by showing that for each $n \in \mathbb{N}$ there exist $x \in F$and $r > 0$ such that $\#\Phi(x, r) \geq \lceil n/q \rceil$.

Fix $n \in \mathbb{N}$ and write $\varepsilon_1 = \varepsilon(q)$. Since Fcontains at least two points and does not satisfy the identity limit criterion, Lemma (5.2.13) implies that there exist $0 < \delta_1 < \varepsilon_1$ and $i_1, j_1 \in \Sigma_*$ such that

$$C^{-1}\delta_1\|\varphi_{i_1}'\| \leq |\varphi_{i_1}(x) - \varphi_{j_1}(x)| \leq C\delta_1\|\varphi_{i_1}'\|$$

For all $x \in V$. We will choose $\delta_k > 0$ and $i_k, j_k \in \Sigma_*, k \in \{1, \dots, n\}$, inductively. Assuming $0 < \delta_{k-1} < \varepsilon_{k-1} < 1$ and $i_{k-1}, j_{k-1} \in \Sigma_*$ have already been chosen for some $k \in \{2, \dots, n\}$, let us fix $0 < \varepsilon_k < (2K^5C^2)^{-1}\delta_{k-1}\|\varphi'_{i_{k-1}}\|$. By Lemma (5.2.13), we then choose $0 < \delta_k < \varepsilon_k$ and $i_k, j_k \in \Sigma_*$such that

$$C^{-1}\delta_k\|\varphi'_{i_k}\| \le |\varphi_{i_k}(x) - \varphi_{j_k}(x)| \le C\delta_k\|\varphi'_{i_k}\| \tag{54}$$

for all $x \in V$.

Define $i = i_n \cdots i_1$ and $k_m = i_n \cdots i_{m+1} j_m i_{m-1} \cdots i_1$for all $m \in \{1, \dots, n\}$. Fix $m, l \in \{1, \dots, n\}$ such that $m \ne l$and notice that we may assume $l < m$, relabeling if necessary. We claim that

$$\varphi_{k_m m}|F \ne \varphi_{k_l m}|F \tag{55}$$

for all $m \in \Sigma_*$. By (25), we have $K^{-2}\|\varphi'_k\|\|\varphi'_j\| \le \|\varphi'_{kj}\| \le \|\varphi'_k\|\|\varphi'_j\|$for all $k, j \in \Sigma_*$. Therefore,by (23) and (54), we have

$$\begin{aligned}|\varphi_i(x) - \varphi_{k_l}(x)| &\ge K^{-1}\|\varphi'_{i_n \cdots i_{l+1}}\||\varphi_{i_l}(\varphi_{i_{l-1}\cdots i_1}(x)) - \varphi_{j_l}(\varphi_{i_{l-1}\cdots i_1}(x))| \\ &\ge (KC)^{-1}\delta_l\|\varphi'_{i_n\cdots i_{l+1}}\|\|\varphi'_{i_l}\| \ge (KC)^{-1}\delta_l\|\varphi'_{i_n\cdots i_l}\|\end{aligned}$$

and, as $\delta_m < \varepsilon_m < (2K^5C^2)^{-1}\delta_{m-1}\|\varphi'_{im-1}\| < \dots < (2K^5C^2)^{l-m}\delta_l\|\varphi'_{i_{m-1}}\| \cdots \|\varphi'_{i_l}\|$, also

$$\begin{aligned}|\varphi_i(x) - \varphi_{k_m}(x)| &\le K^2C\delta_m\|\varphi'_{i_n\cdots i_m}\| \\ &\le K^2C(2K^5C^2)^{l-m}\delta_l\|\varphi'_{i_n\cdots i_m}\|\|\varphi'\_(i_{m-1}\|\cdots\|\varphi'_{i_l}\| \\ &\le K^2C(2K^3C^2)^{l-m}\delta_l\|\varphi'_{i_n\cdots i_l}\| \le (2KC)^{-1}\delta_l\|\varphi'_{i_n\cdots i_l}\|\end{aligned} \tag{56}$$

for all $x \in V$. Since now

$$\begin{aligned}|\varphi_{k_l}(x) - \varphi_{k_m}(x)| &\ge \big||\varphi_i(x) - \varphi_{k_l}(x)| - |\varphi_i(x) - \varphi_{k_m}(x)|\big| \\ &\ge (KC)^{-1}\delta_l\|\varphi'_{i_n\cdots i_l}\| - (2KC)^{-1}\delta_l\|\varphi'_{i_n\cdots i_l}\| = (2KC)^{-1}\delta_l\|\varphi'_{i_n\cdots i_l}\| > 0\end{aligned}$$

for all $x \in V$, we see that $|\varphi_{k_l m}(x) - \varphi_{k_m m}(x)| > 0$ for all $x \in \varphi_m(V)$ and $m \in \Sigma_*$. Therefore, (55)holds and, in particular, the set

$$\Psi_p := \{\varphi_{k_m l(p)}|_F : m \in \{1, \dots, n\}\} \tag{57}$$

Has n elements for all $p \in \{1, \dots, q\}$.

Let $r = \max_{m \in \{1,\dots,n\}} diam(\varphi_{k_m l(q)}(F))$ and $x = \varphi_{il(q)}(x_0)$, where $x_0 \in F$. We will next show that

$$diam(\varphi_{k_m l(q)}(F)) \le r < diam(\varphi_{k_m}(F)) \text{ and } \varphi_{k_m l(q)}(F) \cap B(x, r) \ne \emptyset \tag{58}$$

for all $m \in \{1, \dots, n\}$. To that end, fix$m \in \{1, \dots, n\}$. Choosing $y, z \in F$ such that $|\varphi_{k_m l(q)}(y) - \varphi_{k_m l(q)}(z)|$ =diam$(\varphi_{k_m l(q)}(F))$, we see, by (23), (24), and (53), that

$$\begin{aligned}diam\left(\varphi_{k_m l(q)}(F)\right) &\le \|\varphi'_{k_m}\||\varphi_{l(q)}(y) - \varphi_{l(q)}(z)| \le \text{K}/diam(\text{F}) \\ &diam\left(\varphi_{k_m}(F)\right) diam\left(\varphi_{l(q)}(F)\right) \\ &< \frac{C\|\varphi'_{l(q)}\|diam(F) - K^2\varepsilon(q)}{C\|\varphi'_{l(q)}\|diam(F) + K^2\varepsilon(q)}\, diam\left(\varphi_{k_m}(F)\right).\end{aligned} \tag{59}$$

Note that (56) implies

$$|\varphi_i(x) - \varphi_{k_l}(x)| \le (2KC)^{-1}\varepsilon(q)\|\varphi'_i\|$$

for all $x \in V$and $l \in \{1, \dots, n\}$. Therefore, by the triangle inequality, (25), and (24),

$$diam(\varphi_{k_l l(q)}(F)) \leq diam(\varphi_{il(q)}(F)) + (KC)^{-1}\varepsilon(q)\|\varphi_i'\|$$
$$\leq \left(1 + \frac{K^2\varepsilon(q)}{C\|\varphi_{l\,(q)}'\|diam(F)}\right) diam(\varphi_{il(q)}(F)) \quad (60)$$
for all $l \in \{1, \dots, n\}$. Since, similarly,
$$diam\left(\varphi_{il(q)}(F)\right) \leq diam\left(\varphi_{k_m l(q)}(F)\right) + \frac{K^2\varepsilon(q)}{C\|\varphi_l'(q)\|diam(F)} diam(\varphi_{il(q)}(F)), (61)$$
we conclude, by (59), (61), and (60), that
$$diam(\varphi_{k_m}(F)) > \max_{l \in \{1,\dots,n\}} diam(\varphi_{k_l l(q)}(F)) = r \geq diam(\varphi_{k_m l(q)}(F))$$
as desired. Observe that the role of q is to guarantee the strict inequality above – because of conformality, it might happen that diam $(\varphi_k(F))$ = diam$(\varphi_{k-}(F))$ for some $k \in \Sigma_*$; see Example (5.2.15).Finally, note that (56), the choice of $\varepsilon(q)$, (25), (24), and (61) give us
$$|x - \varphi_{k_m l(q)}(x_0)| = |\varphi_i(\varphi_{l(q)}(x_0)) - \varphi_{k_m}(\varphi_{l(q)}(x_0))| \leq (2KC)^{-1}\varepsilon(q)\|\varphi_i'\|$$
$$= \frac{1}{3}K^{-3}\|\varphi_{l(q)}'\|\|\varphi_i'\|diam(F) \leq \frac{1}{3}K^{-1}\|\varphi_{il(q)}'\|diam(F)$$
$$\leq \frac{1}{3}diam(\varphi_{il(q)}(F)) \leq r$$
Yielding $\varphi_{k_m l(q)}(F) \cap B(x,r) \neq \emptyset$as desired.

Because of the length difference q, we cannot directly apply (58) in the definition of the weakseparation condition. But relying on (58), we see that foreach $m \in \{1, \dots, n\}$there is $p_m \in \{1, \dots, q\}$such that diam $(\varphi_{k_m l(p_m)}(F)) \leq r < diam(\varphi_{k_m l(pm-1)}(F))$ and $\varphi_{k_m l(p_m)}(F) \cap B(x,r) \neq \emptyset$. Hence,

$\Phi_p := \{\varphi_{k_m l(p)}|F : m \in \{1, \dots, n\}$and $p_m = p\} \subset \Phi(x,r)$

for all $p \in \{1, \dots, q\}$. By (57), we have $\Phi_p \subset \Psi_p, \#\Phi p = \#\{m \in \{1, \dots, n\} : p_m 8 = p\}$, and $\#\Psi_p = n$ for all $p \in \{1, \dots, q\}$. Since the function $m \mapsto p_m$ is from$\{1, \dots, n\}$to$\{1, \dots, q\}$, there exists $p \in \{1, \dots, q\}$ such that $\#\Phi_p \geq \lceil n/q \rceil$. Therefore, we have shown that $\#\Phi(x,r) \geq \lceil n/q \rceil$and finished the proof.

We finish with two examples. The first one verifies the need to useqin (58) and the second one examines the difference between the original definition of the weak separation condition and our definition.

Example (5.2.15)[216]: We exhibit a conformal iterated function system$\{\varphi_1, \varphi_2, \varphi_3\}$on $\mathbb{R}^2$ for which the associated self-conformal set $F \subset \mathbb{R}^2$ satisfies $diam(\varphi_i(F)) = diam(\varphi_{i-}(F))$ for $i = 32 \in \Sigma_*$.

Using complex notation, we define
$$\varphi_1(z) = \frac{1}{1000}z - \frac{9}{10}, \varphi_2(z) = \frac{19}{20}iz, \varphi_3(z) = \frac{z}{2(z - 2i)}$$
for all $z \in \mathbb{C}$. The mapping $\varphi_1$is a strongly contracting homothety,φ_2 is a weakly contracting similarity that involves a rotation by $\frac{\pi}{2}$, and $\varphi_3$is a M¨obius transformation with singularity at $2i$.Therefore, all the mappings are injective and holomorphic on $\mathbb{C} \backslash \{2i\}$. To see that their collectionis a conformal iterated function system, it is enough to verify that there exists a bounded open convex set $\Omega \subset \mathbb{C}$such that $\overline{\varphi_j(\Omega)} \subset \Omega$ and $\|\varphi_j'\| = \sup_{z \in \Omega} |\varphi_j'(z)| < 1$ for all $j \in \{1,2,3\}$.

Write $r_0 = \frac{901}{1000}$ and define $\Omega = B^o(0, r_0)$, where $B^o(z,r)$ is an open ball centered at $z \in \mathbb{C}$ with radius $r > 0$. Note that the singularity $2i$ is not contained in the closure of Ω and hence, each ofthe mappingsφ_j maps balls in Ω onto balls. A simple calculation shows that $\varphi_1(\Omega) = B^o(c_1, r_1)$,where $c_1 = -\frac{9}{10}$ and $r_1 = \frac{901}{106}$. Since $|c_1 - r_1| < r_0$, we see that $\overline{\varphi_1(\Omega)} \subset \Omega$. Similarly, we see that $\overline{\varphi_2(\Omega)} = B(0, \frac{19}{20} r_0) \subset \Omega$. We determine $\varphi_3(\Omega)$ by looking at the images of r_0, ir_0, and $-r_0$from the boundary of Ω. Indeed, these three points uniquely describe a circle and hence the ball $\varphi_3(\Omega) = B^o(c_3, r_3)$. The center point $c_3 = -\frac{811801}{6376398}$can be calculated from the equations

$$|c_3 - \varphi_3(r_0)| = |c_3 - \varphi_3(ir_0)| = |c_3 - \varphi_3(-r_0)|,$$

where each of the value is the radius $r_3 = \frac{901000}{3188199}$. Since $|c_3| + r_3 < r_0$, we see that also $\overline{\varphi_3(\Omega)} \subset \Omega$.Furthermore, a direct calculation shows that $\varphi_1'(z) = \frac{1}{1000}, \varphi_2'(z) = \frac{19}{20} i$, and $\varphi_3'(z) = -\frac{i}{(z-2i)^2}$ for all $z \in \mathbb{C}\backslash\{2i\}$. Therefore, as $|\varphi_3(z)| \leq (2 - r_0)^{-2} < 1$ for all $z \in \Omega$, we have $\|\varphi_j'\| < 1$ for all $j \in \{1,2,3\}$. The collection $\{\varphi_1, \varphi_2, \varphi_3\}$is thus a conformal iterated function system. Let $F \subset \mathbb{C}$ be the associated self-conformal set.

Since $\varphi_{32}(F) \subset \varphi_3(F)$, to see that the diameters are equal, it suffices to prove that $\mathrm{diam}(\varphi_3(F)) \leq diam(\varphi_{32}(F))$. Let $w = -\frac{100}{111} \in F$ be the fixed point of φ_1. Defining $q_1 = \varphi_{32}(w) = \frac{95}{634} \in F$ and $q_2 = \varphi_{3222}(w) = -\frac{6859}{21802} \in F$, it follows that

$$diam(\varphi_{32}(F)) \geq |q_1 - q_2| = \frac{1604949}{3455617}.$$

Showing that this number is an upper bound for $\mathrm{diam}(\varphi_3(F))$ will thus finish the proof. Calculating as before, we see that $\varphi_1(B(0,w)) = B(-\frac{9}{10}, \frac{1}{1110}) \subset B(0,w), \varphi_2(B(0,w)) = B(0, \frac{95}{111}) \subset B(0,w)$, and $\varphi_3(B(0,w)) = B(-\frac{1250}{9821}, \frac{2775}{9821}) \subset B(0,w)$. Therefore,$F \subset B(0,w)$. Write $\Gamma = \{31,33,321,323,3221,3223,32221,32222,32223\} \subset \Sigma_*$and note that $\varphi_3(F) \subset \bigcup_{i\in\Gamma} \varphi_i(B(0,w))$.For each $i \in \Gamma$, let c_i be the center and r_i the radius of the ball $\varphi_i(B(0,w))$. Numerical calculations show that

$$diam(\varphi_3(F)) \leq diam\left(\bigcup_{i\in\Gamma} \varphi_i(B(0,w))\right) = \max_{i,j\in\Gamma}\{|c_i - c_j| + r_i + r_j\} = \frac{1604949}{3455617}$$

as required.

Example (5.2.16)[216]: We exhibit a conformal iterated function system$\{\varphi_1, \varphi_2, \varphi_3\}$on $\mathbb{R}$ for which the associated self-conformal set $F \subset \mathbb{R}$ satisfies the weak separation condition but has

$$sup\{\#\Phi^*(x,r) : x \in F \text{and } r > 0\} = \infty, \qquad (62)$$

Where

$\Phi^*(x,r) = \{\varphi_i : diam(\varphi_i(F)) \leq r < diam(\varphi_i - (F))$ and $\varphi_i(F) \cap B(x,r) \neq \emptyset\}$

for all $x \in \mathbb{R}$ and $r > 0$. Since, by [225], the condition $sup\{\#\Phi^*(x,r) : x \in F \text{and } r > 0\} < \infty$is equivalent to the original definition of Lau, Ngai, and Wang [227], we see that our definition is strictly weaker in the non-analytic case .

Let

$$g(x)=\begin{cases}\frac{1}{180}\,(9x^2-6x+\ 1), if\,\frac{1}{3}<x<\frac{5}{12}\\ -\frac{1}{180}\left(9x^2-9x+\frac{17}{8}\right), if\,\frac{5}{12}\le x\ <\frac{7}{12},\\ \frac{1}{120}\left(6x^2-8x+\frac{8}{3}\right), if\,\frac{7}{12}\le x<\frac{2}{3},\\ 0, otherwise,\end{cases}$$

and notice that $0<g(x)\le 1/2880$ for all $x\in(\frac{1}{3},\frac{2}{3})$ and g is continuously differentiable such that $0<g'(x)\le 1/120$ for all $x\in(\frac{1}{3},\frac{1}{2})$ and $-1/120\le g'(x)<0$ for all $x\in(\frac{1}{2},\frac{2}{3})$. In fact,g′is apiecewise linear continuous function and hence H¨older continuous. Define

$$\varphi_1(x)\ =\frac{1}{3}x, \varphi_2(x)=\frac{1}{3}x+\frac{2}{3}, \varphi_3(x)=\frac{1}{3}x+g(x),$$

for all $x\in\mathbb{R}$ and notice that each φ_i is a strictly increasing $C^{1+\alpha}$-function. The collection$\{\varphi_1,\varphi_2,\varphi_3\}$ is therefore a conformal iterated function system and since $\varphi_1|_F=\varphi_3|_F$, the associatedself-conformal set is the standard $\frac{1}{3}$-Cantor set F. Note that also$\{\varphi_1,\varphi_2\}$ defines F and it iswell known that F, defined by these two maps, satisfies the open set condition. Therefore, as $\varphi_1|_F=\varphi_3|_F$, the set F, defined by all three maps, satisfies the weak separation condition.To see that (62) holds, observe first that $\operatorname{diam}(\varphi_i(F))\ =\ 3^{-n}$for all $i\in\Sigma_n$and $n\in\mathbb{N}$. Let $i(k)=i_1i_2\cdots$ be the word in Σ such that $i_k=\ 3$ and $i_j=\ 1$ for all $j\in\mathbb{N}\backslash\{k\}$. Note that $0\in\varphi_{i\,(k)|_n}(F)$ and

$$\varphi_{i(k)|_n}(x)=3^{-n}x+\ 3^{-k+1}g(3^{-n+k}x)$$

for all $x\in\mathbb{R}, k\in\{1,\ldots,n\}$, and $n\in\mathbb{N}$. Therefore, $\varphi_{i(k)|_n}\neq\varphi_{i(m)|_n}$for all $k,m\in\{1,\ldots,n\}$ with $k\neq m$ and $\Phi^*(0,3^{-n})$ has at least n elements for all $n\in\mathbb{N}$.

Let $E\subset\mathbb{R}$ be a compact set. For each $x\in\mathbb{R}$ and $r\ >0$ we define the magnification $M_{x,r}:\mathbb{R}\to\mathbb{R}$ by setting

$$M_{x,r}(z)=\frac{z-x}{r}$$

for all $z\in\mathbb{R}$. We say that $T\subset[-1,1]$ is aweak tangentof E if there exist sequences $(x_n)_{n\in\mathbb{N}}$ of points in $\mathbb{R}$ and $(r_n)_{n\in\mathbb{N}}$ of positive real numbers such that $M_{x_n,r_n}(E)\cap[-1,1]\to T$ in Hausdorff distance. Recall that a sequence $(E_n)_{n\in\mathbb{N}}$ of closed subsets of $[-1,1]$ converges to T in Hausdorff distance if

$\lim_{n\to\infty} sup_{x\in E_n}\operatorname{dist}(x,T)=0$ and $\lim_{n\to\infty}\sup_{y\in T}\operatorname{dist}(y,E_n)=\ 0.$

If T is a weak tangent of E, then it is straightforward to see that $\dim_H(T)\le\dim_A(E)$; see [228].

Theorem (5.2.17)[216]: Let $F\subset\mathbb{R}$ be a self-conformal set containing at least two points. If Fdoes not satisfy the identity limit criterion, then $\dim_A(F)\ =1$.

Proof. By the above discussion, it suffices to show that there is a constant $D'\ge 1$such that for every $n\in\mathbb{N}$ there exist $x\in\mathbb{R}$, $r>0$, and points $x_n<x_{n-1}<\cdots<x_1$ in F such that $M_{x,r}(x_n)=-1, M_{x,r}(x_1)=1$, and

$$M_{x,r}(x_k)-M_{x,r}(x_{k+1})\le\frac{D'}{n+\ 1}$$

for all $k\in\{1,\ldots,n-1\}$. Indeed, by letting $n\to\infty$, this implies that $[-1,1]$ is a weak tangent of F and therefore, F has full Assouad dimension.

Let $C\ge 1$ be as in Lemma (5.2.13) and $K\ge 1$ as in Lemma (5.2.8). Define

$$D = \frac{K^{11}C}{diam(F)\min_{i\in\Sigma_2}\|\varphi_i'\|}, \tag{63}$$

Fix $n \in \mathbb{N}$, and choose $0 < \varepsilon < diam(F)/(4KC)$ such that $(1 + 2D\varepsilon)^{n-1} \leq 2$ and $(1 - 2D\varepsilon)^{n-1} \geq \frac{1}{2}$.

Since F contains at least two points and does not satisfy the identity limit criterion, Lemma (5.2.13) implies that there exist $0 < \delta_1 < \varepsilon$ and $i_1', j_1' \in \Sigma_*$such that

$$C^{-1}\delta_1 max\{\|\varphi_{i_1'}'\|, \|\varphi_{j_1'}'\|\} \leq |\varphi_{i_1'}(x) - \varphi_{j_1'}(x)| \leq C\delta_1 min\{\|\varphi_{i_1'}'\|, \|\varphi_{j_1'}'\|\} \tag{64}$$

for all $x \in V$. Recall that, by (23), $K^{-1}\|\varphi_i'\| \leq |\varphi_i'(x)| \leq \|\varphi_i'\|$for all $x \in V$and $i \in \Sigma_*$. Inparticular, this means that φ_i' is either positive or negative and hence, as $V \subset R$ is an open interval,each φ_i is strictly monotone on$V$. Let $y, z \in F \subset V$ be such that $y - z = diam(F)$. The mean value theorem implies that there exists $w \in V$ such that

$$\varphi i_1'(y) - \varphi i_1'(z) = \varphi_{(i_1')}'(w) diam(F).$$

If$\varphi_{i_1'}'(w) > 0$, then (64) implies

$$\varphi j_1'(y) - \varphi j_1'(z) \geq \varphi i_1'(y) - \varphi i_1'(z) - 2C\delta_1\|\varphi_{i_1'}'\| \geq (\varphi_{i_1'}'(w) - \frac{1}{2}K^{-1}\|\varphi_{i_1'}'\|)\, diam(F)$$
$$\geq \frac{1}{2}K^{-1}\left\|\varphi_{i_1'}'\right)\|diam(F) > 0$$

and hence, $\varphi j_1'(y) > \varphi j_1'(z)$ yielding $\varphi_{j_1'}'(x) > 0$ for all $x \in V$. Similarly, if $\varphi_{i_1'}'(w) < 0$, then we see that $\varphi_{j_1'}'(x) < 0$ for all $x \in V$. Therefore, the derivatives $\varphi_{i_1'}'$and $\varphi_{j_1'}'$have the same sign. Let $i_1 = i_1'i_1'$ and $j_1 = i_1'j_1'$ and notice that, by the chain rule,$\varphi_{i_1}'$and $\varphi_{j_1}'$are positive. By (25),(64), and (23), we have

$$(KC)^{-1}\delta_1 max\{\|\varphi_{i_1}'\|, \|\varphi_{j_1}'\|\} \leq K^{-1}\|\varphi_{i_1'}'\||\varphi_{i_1'}(x) - \varphi_{j_1'}(x)| \leq |\varphi_{i_1}(x) - \varphi_{j_1}(x)|$$
$$\leq \|\varphi_{i_1'}'\||\varphi_{i_1'}(x) - \varphi_{j_1'}(x)| \leq K^2C\delta_1 min\{\|\varphi_{i_1}'\|, \|\varphi_{j_1}'\|\}.$$

Furthermore, since $V \subset \mathbb{R}$ is an open interval and $|\varphi_{i_1}(x) - \varphi_{j_1}(x)| > 0$ for all $x \in V$, we have,by the intermediate value theorem, that $\varphi_{i_1}(x) > \varphi_{j_1}(x)$ for all $x \in V$, relabeling i_1 and $j_1$if necessary. Therefore,

$$(KC)^{-1}\delta_1\|\varphi_{i_1}'\| \leq \varphi_{i_1}(\varphi_k(x)) - \varphi_{j_1}(\varphi_k(x)) \leq K^2C\delta_1\|\varphi_{i_1}'\| \tag{65}$$

for all $x \in V$ and $k \in \Sigma_*$. Notice that, by the chain rule, there exists $k \in \Sigma_2$ such that φ_k' Is positive. Choose $k_1 = k \cdots k \in \Sigma_*$ such that

$$\varepsilon\|\varphi_{i_1}'\|\|\varphi_{k_1}'\| < \delta_1\|\varphi_{i_1}'\| \leq \varepsilon\|\varphi_{i_1}'\|\|\varphi_{k_1|_{|k_1|-2}}'\|$$

and notice that also $\varphi_{k_1}'$is positive. Therefore, it follows from (65) that

$$(KC)^{-1}\varepsilon\|\varphi_{i_1k_1}'\| \leq \varphi_{i_1k_1}(x) - \varphi_{j_1k_1}(x) \leq K^6C\left(\min_{i\in\Sigma_2}\|\varphi_i'\|\right)^{-1}\varepsilon\|\varphi_{i_1k_1}'\|$$

for all $x \in V$.

To find more points being predefined distance apart, we continue inductively. Assuming $i_l, j_l, k_l \in \Sigma_*, l \in \{1, \ldots, k-1\}$, have already been chosen for some $k \in \{2, \ldots, n\}$, we apply Lemma (5.2.13) as above to find $0 < \delta_k < \varepsilon K^{-2}\|\varphi_{j_{k-1}k_{k-1}\cdots j_1k_1}'\|$ and $i_k, j_k \in \Sigma_*$ such that φ_{i_k}' and $\varphi_{j_k}'$are positive,and

$$(KC)^{-1}\delta_k\|\varphi_{i_k}'\| \leq \varphi_{i_k}(\varphi_{kj_{k-1}k_{k-1}\cdots j_1k_1}(x)) - \varphi_{j_k}(\varphi_{kj_{k-1}k_{k-1}\cdots j_1k_1}(x)) \leq K^2C\delta_k\|\varphi_{i_k}'\|$$

for all $x \in V$ and $k \in \Sigma_*$.Since $\delta_k\|\varphi_{i_k}'\| \leq \varepsilon K^{-2}\|\varphi_{j_{k-1}k_{k-1}\cdots j_1k_1}'\|\|\varphi_{i_k}'\| \leq \varepsilon\|\varphi_{i_kj_{k-1}k_{k-1}\cdots j_1k_1}'\|$,there is $k_k \in \Sigma_*$such that $\varphi_{k_k}'$is positive and

$$\varepsilon\|\varphi'_{i_kj_{k-1}k_{k-1}\cdots j_1k_1}\|\|\varphi'\_(k_k\|< \delta_k\|\varphi'_{i_k}\|\leq \varepsilon\|\varphi'_{i_kj_{k-1}k_{k-1}\cdots j_1k_1}\|\|\varphi'_{k_k|_{|k_k|-2}}\|$$
$$\leq K^2\left(\min_{i\in\Sigma_2}\|\varphi'_i\|\right)^{-1}\varepsilon\,\|\varphi'_{i_kj_{k-1}k_{k-1}\cdots j_1k_1}\|\|\varphi'_{k_k}\|.$$
Note that, by (25), $K^{-2}\|\varphi'_{ik_j}\|\leq\|\varphi'_{ij}\|\|\varphi'_k\|\leq K^4\|\varphi'_{ik_j}\|$ for all $i,j,k\in\Sigma_*$. Therefore,
$$(K^3C)^{-1}\varepsilon\|\varphi'_{i_kk_kj_{k-1}k_{k-1}\cdots j_1k_1}\|\leq \varphi_{i_kk_kj_{k-1}k_{k-1}\cdots j_1k_1}(x)-\varphi\_(j_kk_kj_{k-1}k_{k-1}\cdots j_1k_1(x)$$
$$\leq K^8C\left(\min_{i\in\Sigma_2}\|\varphi'_i\|\right)^{-1}\varepsilon\left\|\varphi'_{i_kk_kj_{k-1}k_{k-1}}\cdots j_1k_1\right\| \tag{66}$$
for all $x\in V$. We have thus shown the existence of words $i_k,j_k,k_k\in\Sigma_*,k\in\{1,\dots,n\}$, for whichthe derivatives $\varphi'_{i_k},\varphi'_{j_k}$, and $\varphi'_{k_k}$are positive and (66) holds for all $k\in\{1,\dots,n\}$.We will use (66) to define the required points $x_n<x_{n-1}<\cdots<x_1$in F. Let $h_k=i_nk_n\cdots i_kk_kj_{k-1}k_{k-1}\cdots j_1k_1$ and notice that, by the chain rule,$\varphi'_{h_k}$is positive for all $k\in\{1,\dots,n\}$. Therefore, by (23) and (66), we have
$$\varphi_{h_k}(x)-\varphi_{h_{k+1}}(x)$$
$$\leq\|\varphi'_{i_nk_n\cdots i_{k+1}k_{k+1}}\|(\varphi\_(i_kk_kj_{k-1}k_{k-1}\cdots j_1k_1(x)-\varphi_{j_kk_kj_{k-1}k_{k-1}\cdots j_1k_1}(x))$$
$$\leq K^{10}C\left(\min_{i\in\Sigma_2}\|\varphi'_i\|\right)^{-1}\varepsilon\|\varphi'_{h_k}\|$$
and, similarly,
$$\varphi_{h_k}(x)-\varphi_{h_{k+1}}(x)\geq(K^6C)^{-1}\varepsilon\|\varphi'_{h_k}\|>0$$
for all $x\in V$and $k\in\{1,\dots,n-1\}$. Recalling (24) and the definition of $D\geq 1$ given in (63), wehave thus shown that
$$D^{-1}\varepsilon diam\left(\varphi_{h_k}(F)\right)\leq\varphi_{h_k}(x)-\varphi_{h_{k+1}}(x)\leq D\varepsilon diam\left(\varphi_{h_k}(F)\right) \tag{67}$$
for all $x\in V$and $k\in\{1,\dots,n-1\}$. Let $y,z\in F$ be such that $\varphi_{h_{k+1}}(y)-\varphi_{h_{k+1}}(z)=diam(\varphi_{h_{k+1}}(F))$.Since $\varphi'_{h_{k+1}}$ and $\varphi'_{h_k}$are positive, we have $z<y$ and $\varphi_{h_k}(z)<\varphi_{h_k}(y)$. Therefore, by (67), we have
$$diam\left(\varphi_{h_k}(F)\right)\geq\varphi_{h_k}(y)-\varphi_{h_k}(z)\geq diam\left(\varphi_{h_{k+1}}(F)\right)-2D\varepsilon diam(\varphi_{h_k}(F))$$
And
$$diam(\varphi_{h_{k+1}}(F))\leq(1+2D\varepsilon)\,diam(\varphi_{h_k}(F))$$
for all $k\in\{1,\dots,n-1\}$. Choosing $z,y\in F$such that $\varphi_{h_k}(y)-\varphi_{h_k}(z)=diam(\varphi_{h_k}(F))$, we similarly see that
$$diam(\varphi_{h_{k+1}}(F))\geq(1-2D\varepsilon)\,diam(\varphi_{h_k}(F))$$
for all $k\in\{1,\dots,n-1\}$. By the choice of $\varepsilon>0$, we have thus shown that
$$\frac{1}{2}diam\left(\varphi_{h_1}(F)\right)\leq(1-2D\varepsilon)^{n-1}diam\left(\varphi_{h_1}(F)\right)\leq diam\left(\varphi_{h_k}(F)\right)$$
$$\leq(1+2D\varepsilon)^{n-1}diam\left(\varphi_{h_1}(F)\right)\leq 2\,diam\left(\varphi_{h_1}(F)\right) \tag{68}$$
for all $k\in\{1,\dots,n\}$. Fix $x_0\in F$ and define $x_k=\varphi_{h_k}(x_0)$ for all $k\in\{1,\dots,n\}$. It follows from (67) that $x_n<x_{n-1}<\cdots<x_1$. Letting $x=(x_n+x_1)/2$ and $r=(x_1-x_n)/2$, we have $M_{x,r}(x_n)=-1$ and $M_{x,r}(x_1)=1$. Finally, since (67) and (68) imply
$$x_1-x_n=\sum_{k=1}^{n-1}\varphi_{h_k}(x_0)-\varphi_{h_{k+1}}(x_0)\geq D^{-1}\varepsilon\sum_{k=1}^{n-1}diam\,(\varphi_{h_k}(F))$$
$$\geq\frac{1}{2}D^{-1}\varepsilon(n+1)\,diam(\varphi_{h_1}(F))$$
And

$$x_k - x_{k+1} \leq D\varepsilon\, diam(\varphi_{h_k}(F)) \leq 2D\varepsilon diam(\varphi_{h_1}(F))$$
for all $k \in \{1, \dots, n-1\}$, we see that
$$M_{x,r}(x_k) - M_{x,r}(x_{k+1}) = \frac{2(x_k - x_{k+1})}{x_1 - x_n} \leq \frac{8D^2}{n+1}$$
as required.

Section (5.3): Positive Lebesgue Measure and Nonempty Interior

In Euclidean spaces, it is well-known that a set with nonempty interior has positive Lebesgue measure and the converse is not true. However, motivated by problems of Palis and Takens [241] on arithmetic sums of Cantor sets, one might expect that positive Lebesgue measure does imply nonempty interior if the set has some "selfsimilar" construction.

Many works [232], [120], [235], [201], [62], [125], [168], [237], [137], [128], [136] support this expectation in special cases. Among these results, Schief [128] proves that a self-similar set in $\mathbb{R}^d$ which satisfies the open set condition (OSC) and has positive d-dimensional Lebesgue measure must have nonempty interior. Zerner [136] and Peres et al. [137] extend Schief's result to self-similar sets with the weak separation condition (WSC) and self-conformal sets with the OSC, respectively. So a natural question arises:

For convenience, we recall two manners which describe "self-similarity". The first one is the theory of iterated functions system (IFS) (see Hutchinson [83]). An IFS consists of a family of contractions $\{S_1, \dots, S_N\}$ on a complete metric space (X, ρ), often $X = \mathbb{R}^d$. The fundamental property of an IFS $\{S_1, \dots, S_N\}$ is that it determines a unique nonempty compact set $K \subset X$ satisfying $K = \bigcup_{i=1}^{N} S_i(K)$, which is called the invariant set of the IFS. The invariant set is called a self-similar set (self-conformal set or self-affine set) if all S_i are similitudes (conformal mappings or affine mappings, respectively).

To understand the structure of invariant sets, various separation conditions were introduced to control the overlaps between small copies of an invariant set. The OSC, which means the overlaps are small, is introduced by Moran [99] and extensively studied in [120], [119], [116], [236], [83], [137], [128], [245]. Another well-studied separation condition is the WSC introduced by Lau and Ngai [135], which extends the OSC while allowing exact overlaps on the iteration, see also [233], [217], [129], [135], [239], [227], [240], [136].

The second manner to describe "self-similarity" is the notion of BPI spaces ("big pieces of itself") introduced by David and Semmes [234], in which they replace IFS by conformally bi-Lipschitz mappings. Roughly speaking, a metric space X is a BPI space if it is Ahlfors regular and for any pair of balls in X there are subsets of relatively large measure inside them which look approximately the same in terms of conformal bi-Lipschitz equivalence. A subset $K \subset \mathbb{R}^d$ is called a BPI set if it together with the Euclidean metric is a BPI space. Peres and Solomyak [242] propose the following question: assume that a self-similar set in $\mathbb{R}_d$ has positive d-dimensional Lebesgue measure, must it have nonempty interior? Csörnyei et al. [160] answer this question negatively by constructing a family of self-similar sets in $\mathbb{R}^2$ with positive Lebesgue measure but empty interior. In fact, these self-similar sets in [160] are also BPI sets (see Example (5.3.20)). Hence, neither of the two "self-similar" constructions.

We introduce BBI spaces ("big balls of itself"), which enhance the notion of BPI spaces by requiring that, for any pair of balls B_1 and B_2, we can find a relatively large ball in B_1 and a subset in B_2 such that they look approximately the same in terms of conformal bi-Lipschitz equivalence (see Definition (5.3.3)). Similarly, a subset $K \subset\backslash R^d$ is called a BBI set if it together with the Euclidean metric is a BBI space.

It turns out that the "self-similar" construction described by the BBI.

We remark that the definition of BBI set requires Ahlfors regularity (see Definition (5.3.3)), so the condition in Theorem (5.3.9) implies immediately that the d-dimensional Lebesgue measure $\mathcal{L}^d(K) > 0$. Thus, Theorem (5.3.9) asserts that $\mathcal{L}^d(K) > 0 +$ some "self-similar" construction $\Longrightarrow K^\circ \neq \emptyset$.
As an application, we prove that a self-conformal set with nonempty interior must be a BBI set.

Finally, we apply above results to the self-conformal sets satisfying the WSC and prove that, for such sets, positive Lebesgue measure implies nonempty interior. This extends the results [128], [136] and [137].

We give the definition of BBI spaces and then prove Theorem (5.3.9). We study the self-conformal sets and prove Theorems (5.3.18) and (5.3.19).

The definitions of BPI spaces and BBI spaces. In the following, we will use $B(x,r)$ to denote the open ball and $\bar{B}(x,r)$ to denote the closed ball. We begin with the definition of conformally bi-Lipschitz mappings.

Definition (5.3.1)[231]: (Conformally bi-Lipschitz mappings). Given two metric spaces (X,ρ_1) and (Y,ρ_2) and a mapping f : X → Y , we say that f is C-conformally bi-Lipschitz with scale factor λ if

$$C^{-1}\lambda\rho_1(x_1,x_2) \le \rho_2 f(x_1), f(x_2) \le C\lambda\rho_1(x_1,x_2) \; for\ all\ x_1,x_2 \in X.$$

The purpose of introducing scale factor is to eliminate the effect of geometric size on Lipschitz constant.

Definition (5.3.2)[231]: (BPI spaces, [234]). A metric space X is a BPI space if it is Ahlfors regular of some dimension d and if there exist constants $\theta, C > 0$ such that for each pair of balls $B(x_1,r_1)$ and $B(x_2,r_2)$ in X with $0 < r_1, r_2 \le diam\, X$ there is a closed set $A \subset B(x_1,r_1)$ with $H^d(A) \ge \theta r_1^d$ and a mapping $h: A \to B(x_2,r_2)$ which is C-conformally bi-Lipschitz with scale factor r_2/r_1.

A subset $K \subset \mathbb{R}^d$ is called a BPI set if it together with the Euclidean metric is a BPI space.

Recall that a complete metric space X is said to be Ahlfors regular of dimension d if there is a constant $C > 0$ such that

$$C^{-1}r^d \le H^d\big(B(x,r)\big) \le Cr^d$$

for all $x \in X$ and $0 < r \le diam\, X$, where $diam\, X$ denotes the diameter of X and H^d the d-dimensional Hausdorff measure.

Based on the notion of BPI spaces, we introduce the BBI spaces.

Definition (5.3.3)[231]: (BBI spaces). A metric space X is a BBI space if it is Ahlfors regular of some dimension d and if there exist constants $\theta, C > 0$ such that for each pair of balls $B(x_1,r_1)$ and $B(x_2,r_2)$ in X with $0 < r_1, r_2 \le diam\, X$ there is a closed ball $\bar{B}(x_3,\theta r_1) \subset B(x_1,r_1)$ and a mapping $h: B(x_3,\theta r_1) \to B(x_2,r_2)$ which isCconformally bi-Lipschitz with scale factor r_2/r_1. A subset $K \subset \mathbb{R}^d$ is called a BBI set if it together with the Euclidean metric is a BBI space.

For a nonempty set A in a metric space (X,ρ), we call

$$A_\varepsilon = \{x \in X: dist(x,A) \le \varepsilon\}$$

the ε-neighborhood of A, where $dist(x,A) = \inf_{y\in A} \rho(x,y)$ is the distance between x and A.
For a pair of nonempty compact sets $A, B \subset X$, recall that the Hausdorff metric is given by

$$d_H(A,B) = \inf\{\varepsilon: A \subset B_\varepsilon, B \subset A_\varepsilon\}.$$

We write $A_n \overset{d_H}{\longrightarrow} A$ if A_n converges to A in the Hausdorff metric.

We need two known results in topology.

Theorem (5.3.4)[231]: ([238]). If $h: U \to \mathbb{R}^d$ is a continuous one-to-one mapping from an open set $U \subset \mathbb{R}^d$, then h(U) is an open subset of $\mathbb{R}^d$, too.

Lemma (5.3.5)[231]: ([124]). Let X be a compact metric space and K the family of all nonempty compact subsets of X, then K is compact in the Hausdorff metric.

Lemma (5.3.6)[231]: Let $(A_n)_{n\geq 1}$ be a sequence of nonempty compact sets in a metric space (X,ρ) with $A_n \overset{d_H}{\longrightarrow} A$ for some nonempty compact set $A \subset X$. Let $(x_n)_{n\geq 1}$ be a sequence of points in X with $x_n \in A_n$ and $x_n \to x$ for some $x \in A$. $Fix \theta > 0$. For each $\varepsilon > 0$, there exists an $n_\varepsilon \geq 1$ such that

$$\bar{B}(x_n,\theta) \cap A_n \subset (B(x,\theta) \cap A)_\varepsilon \ for\ n > n_\varepsilon.$$

Proof. Fix $\varepsilon > 0$. Since

$$\bigcap_{\delta>0} (\bar{B}(x,\theta + \delta) \cap A_\delta) = \bar{B}(x,\theta) \cap A,$$

there exists $a\ \delta > 0$ such that

$$\bar{B}(x,\theta + \delta) \cap A_\delta \subset (\bar{B}(x,\theta) \cap A)_\varepsilon.$$

Since $A_n \overset{d_H}{\longrightarrow} A$ and $x_n \to x$, there exists $an\ n_\varepsilon \geq 1$ such that

$$\bar{B}(x_n,\theta) \cap A_n \subset \bar{B}(x,\theta + \delta) \cap A_\delta \subset (\bar{B}(x,\theta) \cap A)_\varepsilon,$$

for $n > n_\varepsilon$.

For a bi-Lipschitz mapping h which maps (X,ρ_1) to (Y,ρ_2), its bi-Lipschitz constant bliph is defined by

$$blip\ h := inf\{c \geq 1: c^{-1}\rho_1(x_1,x_2) \leq \rho_2\big(f(x_1),f(x_2)\big) \leq c\rho_1(x_1,x_2)\}.$$

Lemma (5.3.7)[231]: Let An be a nonempty compact set in a metric space (X,ρ) and $h_n: A_n \to X$ a bi-Lipschitz mapping with $blip\ h_n \leq C$ for all $n \geq 1$. Suppose that $A_n \overset{d_H}{\longrightarrow} A$ and $h_n(A_n) \overset{d_H}{\longrightarrow} A^*$, then there is a bi-Lipschitz bijection h maps A onto A^* with $blip\ h \leq C$.

Proof. Pick a countable dense subset $\{x_m: m \geq 1\}$ of A and a countable dense subset $\{x^*_{-m}: m \geq 1\}$of A^*. Since $A_n \overset{d_H}{\longrightarrow} A$ and$h_n(A_n) \overset{d_H}{\longrightarrow} A^*$, for each $m \geq 1$, we can find two sequences $\big(x_{m,n}\big)_{n\geq 1}$ and $\big(x \propto_{-m,n}\big)_{n\geq 1}$ such that $x_{m,n}, x_{-m,n} \in A_n$ for every $n \geq 1$ and

$$x_{m,n} \to x_{m,}\ h_n(x_{-m,n}) \to x^*_{-m} \tag{69}$$

as $\to \infty$.

We claim that, for every $m \geq 1$, the two sequences $\big(x_{-m,n}\big)_{n\geq 1}$ and $\Big(h_n\big(x_{m,n}\big)_{n\geq 1}\Big)$ both have a convergent subsequence. To see this, fix $m \geq 1$. Notice that for each $k \geq 1$, there are finitely many balls $\big(B_{k,i}\big)_i$ of radius $1/k$ which cover A since A is compact. We can find a ball B_k in $\big(B_{k,i}\big)_i$ such that B_k contains infinitely many points in $\big(x_{-m,n}\big)_{n\geq 1}$ since $dist(x_{-m,n}, A) \to 0$ as $n \to \infty$. It follows that there exists a subsequence $(y_k)_{k\geq 1}$ of $\big(x_{-m,n}\big)_{n\geq 1}$ such that $y_k \in B_k \cap \{x_{-m,n}: n \geq 1\}$. Now pick $a_k \in B_k \cap A$, then $(a_k)_{k\geq 1}$ has a convergence subsequence since A is compact. Consequently, so does $(y_k)_{k\geq 1}$ since $\rho(y_k,a_k) \leq 2/k \to 0$ as $k \to \infty$. Therefore, $\big(x_{-m,n}\big)_{n\geq 1}$also has a convergence subsequence. This argument is also applicable to the sequence $(h_n\big(x_{m,n}\big))_{n\geq 1}$.

Combining the claim above with Cantor's diagonal argument, by taking a subsequence of $(n)n \geq 1$ if necessary, we can assume that, for each $m \geq 1$,

$$x_{-m,n} \to x_{-m,} h_n(x_{m,n}) \to x_m^*, \tag{70}$$

as $n \to \infty$ f or some $x_{-m} \in A$ and $x_m^* \in A^*$. Now let $A_0 = \{x_m : m \neq 0\}$ and $A_0^* = \{x_m^* : m \neq 0\}$. Notice that

$$x_m = x_{m'} \Leftrightarrow x_m^* = x_{m'}^*. \tag{71}$$

Indeed, since $bliphn \leq C$ for all $n \geq 1$, we have

$$x_m = x_{m'} \Leftrightarrow \lim_{n\to\infty} \rho(x_{m,n,} x_{m',n}) = 0$$

$$\Leftrightarrow \lim_{n\to\infty} \rho\big(h_n(x_{m,n}), h_n(x_{m',n})\big) = 0 \Leftrightarrow x_m^* = x_{m'}^*.$$

It follows from (71) that the mapping $h_0 : x_m \to x_m^*$ is a bijection from A_0 onto A_0^*. Note that A_0 and A_0^* are dense in A and A^*, respectively, since so are $\{x_m : m \geq 1\}$ and $\{x_{-m}^* : m \geq 1\}$. We claim that

h_0 is a bi-Lipschitz bijection from A_0 onto A_0^* with $blip\, h_0 \leq C$. (72)

A standard result in mathematic analysis says that a uniformly continuous function from a dense subset of a metric space to a complete metric space has a uniformly continuous extension. (see [244]).

Hence, if the claim (72) is true, then h0 can be extend to a uniformly continuous function $h: A \to A^*$. For distinct $y, y' \in A$, pick $y_n, y_n' \in A_0$ with $y_n \to y$ and $y_n' \to y'$, then the claim (72) implies

$$C^{-1}\rho(y, y') = C^{-1} \lim_{n\to\infty} \rho(y_n, y_n') \leq \lim_{n\to\infty} \rho(h_0(y_n), h_0(y_n'))$$

$$= \lim_{n\to\infty} \rho(h(y_n), h(y_n')) = \rho(h(y), h(y')) = \lim_{n\to\infty} \rho(h(y_n), h(y_n'))$$

$$= \lim_{n\to\infty} \rho(h_0(y_n), h_0(y_n')) \leq C \lim_{n\to\infty} \rho(y_n, y_n') = C\rho(y, y').$$

Thus, h is bi-Lipschitz with $bliph \leq C$. It follows that h maps $A = A_0$ onto $\overline{h(A_0)} = \overline{A_0^*} = A^*$. Consequently, h is desired.

It remains to prove the claim (72). For this, pick $a, a' \in A_0$. By (69) and (70), there exist $a_n, a_n' \in A_n$ such that

$$a_n \to a, a_n' \to a' \; and \; h_n(a_n) \to h_0(a), h_n(a_n') \to h_0(a')$$

as $n \to \infty$. Since $blip\, h_n \leq C$ for all $n \geq 1$, we have

$$C^{-1}\rho(a, a') = C^{-1} \lim_{n\to\infty} \rho(a_n, a_n') \leq \lim_{n\to\infty} \rho(h_n(a_n), h_n(a_n'))$$

$$= \rho(h_0(a), h_0(a')) = \lim_{(n\to\infty)} \rho(h_n(a_n), h_n(a_n'))$$

$$\leq C \lim_{n\to\infty} \rho(a_n, a_n') = C\rho(a, a').$$

This proves the claim and the proof is complete.

The following technical lemma, which comes from the idea in the proof of [136], plays an important role in the proof of Theorem (5.3.9).

Lemma (5.3.8)[231]: Let K be a compact subset of $\mathbb{R}^d$ with $K^\circ = \emptyset$. Fix $C \geq 1$. For each $\theta > 0$, there exists $\varepsilon_\theta > 0$ such that for any compact subset A of K, any $x \in A$ and any bi-Lipschitz mapping $h: A \to \bar{B}(0,1)$ with $bliph \leq C$, we have

$$\frac{\mathcal{L}^d\big(B(h(x), \theta) \cap h(A)\big)}{\alpha_d \theta^d} \leq 1 - \varepsilon_\theta,$$

where α_d denotes the Lebesgue measure of the unit ball in $\mathbb{R}^d$.

Proof. Let us argue by contradiction. Suppose the lemma were false. Then for some $\theta > 0$, there exist compact subsets $A_n \subset K, x_n \in A_n$ and bi-Lipschitz mappings $h_n: A_n \to \bar{B}(0,1)$ with $blip\, h_n \leq C$ for $n \geq 1$ such that

$$\frac{\mathcal{L}^d\big(B(h_n(x),\theta) \cap h_n(A)\big)}{\alpha_d \theta^d} \to 1 \; as \; n \to \infty . \tag{73}$$

Notice that K and $B(0,1)$ are both compact. According to Lemma (5.3.5), by taking a subsequence of $(n)n \geq 1$ if necessary, we can assume that

$$A_n \overset{d_H}{\longrightarrow} A, \qquad h_n(A_n) \overset{d_H}{\longrightarrow} A^*, h_n(x_n) \to x^* \tag{74}$$

as $n \to \infty$ for some nonempty compact sets $A \subset K, A^* \subset \bar{B}(0,1)$ and some point $x^* \in \bar{B}(0,1)$.

By (74), we can apply Lemma (5.3.6) to $(h_n(A_n))$ and $(h_n(x_n))$. This gives that for each $\varepsilon > 0$, there is an nε such that

$$\bar{B}(h_n(x_n),\theta) \cap h_n(A_n) \subset D_\varepsilon \; for \; n \geq n_\varepsilon, \tag{75}$$

where $D = B(x^*,\theta) \cap A^*$.

It follows from (73) and (75) that

$$\frac{\mathcal{L}^d(D_\varepsilon)}{\alpha_d \theta^d} \geq \lim_{n\to\infty} \frac{\mathcal{L}^d\big(B(h_n(x_n),\theta) \cap h_n(A_n)\big)}{\alpha_d \theta^d} = 1 \; for \; every \varepsilon > 0.$$

So we have

$$1 \geq \frac{\mathcal{L}^d(B(x^*,\theta) \cap A^*)}{\alpha_d \theta^d} = \mathcal{L}^d(D)\alpha_d \theta^d = \lim_{\varepsilon\to 0} \frac{\mathcal{L}^d(D_\varepsilon)}{\alpha_d \theta^d} \geq 1,$$

where the last equality comes from the fact that D is compact. Thus,

$$\mathcal{L}^d(B(x^*,\theta) \cap A^*) = \alpha_d \theta^d = \mathcal{L}^d\big(\bar{B}(x^*,\theta)\big).$$

$So \; \bar{B}(x^*,\theta) \cap A^*$ is a compact subset of $B(x^*,\theta)$ of full Lebesgue measure. We have

$$\bar{B}(x^*,\theta) \subset A^*. \tag{76}$$

Using (74) again, Lemma (5.3.7) gives a bi-Lipschitz bijection h from A onto A^*. Now Theorem (5.3.4) asserts that $A^\circ \neq \emptyset$ since so does A^* (by (76)). Consequently, $K^\circ \supset A^\circ \neq \emptyset$, contrary to the condition $K^\circ = \emptyset$.

Theorem (5.3.9)[231]: Let $K \subset \mathbb{R}^d$ be a compact BBI set of dimension d, then $K^\circ \neq \emptyset$.

Proof. Let $K \subset \mathbb{R}^d$ be a compact BBI set. By definition, K together with the Euclidean distance ρ is a BBI space. To avoid confusion, we use $B(x,r)$ and $B_K(x,r)$ to denote open balls in the two metric spaces $\mathbb{R}^d$ and K, respectively. For $x \in K$, we have $B_K(x,r) = B(x,r) \cap K$.

Suppose without loss of generality that diam $K = 1$. For $0 < r \leq 1$, denote by N_r the largest number of disjoint balls of radius r centered in K. For each $0 < r < 1/2$, let $\{B_K(x_i,2r)\}_{i=1}^{N_{2r}}$ be a disjoint family of balls with $x_i \in K$. Then

$$K \subset \bigcup_{i=1}^{N_{2r}} B_K(x_i,4r). \tag{77}$$

Notice that K can be regarded as a ball of radius 1 in the BBI space K. By the definition of BBI space (Definition (5.3.3)), there are two constants $\theta, C > 0$ such that for each $1 \leq i \leq N_{2r}$, we can find a closed ball $B_K(y_i,2\theta r) \subset B_K(x_i,2r)$ with $y_i \in K$ and a C-conformlly bi-Lipschitz mapping

$$fi: \bar{B}_K(y_i,2\theta r) \to K \; with \; scale \; factor \; 1/(2r). \tag{78}$$

Let

$$K(r) = \bigcup_{i=1}^{N_{2r}} \bar{B}_K(y_i,\theta r).$$

It follows from (77) that

$$\mathcal{L}^d(K(r)) = \sum_{i=1}^{N_{2r}} \mathcal{L}^d(\bar{B}_K(y_i,\theta r)) = 6^{-d}\theta^d N2ri = \sum_{i=1}^{N_{2r}} \mathcal{L}^d(\bar{B}_K(y_i,6r))$$
$$\geq 6^{-d}\theta^d\mathcal{L}^d\left(\bigcup_{i=1}^{N_{2r}} B_K(y_i,6r)\right) \geq 6^{-d}\theta^d\mathcal{L}^d\left(\bigcup_{i=1}^{N_{2r}} B_K(x_i,4r)\right)$$
$$\geq 6^{-d}\theta^d\mathcal{L}^d(K) \quad (79)$$

for all $0 < r < 1/2$. Now suppose the theorem were false, i.e., $K^\circ = \emptyset$. We claim that, for every $0 < r < 1/2$,

$$\frac{\mathcal{L}^d\big(\bar{B}_K(x,\theta r)\big)}{\alpha_d\theta^d r^d} \leq 1-\varepsilon_\theta \; for\; all\; x \in K(r). \quad (80)$$

Here ε_θ is the constant in Lemma (5.3.8) with $2C$ instead of C.
By assumption, K is Ahlfors regular with dimension d and thus, $\mathcal{L}^d(K) > 0$. We shall show that (79) and (80) contradict this. In fact, by Lebesgue density theorem and Egoroff's theorem, there is a measurable subset $K^* \subset K$ and $r_0 > 0$ such that $\mathcal{L}^d(K\setminus K^*) < 6^{-d}\theta^d\,\mathcal{L}^d(K)$ and that

$$\frac{\mathcal{L}^d\big(\bar{B}_K(x,r)\big)}{\alpha_d r^d} > 1-\varepsilon_\theta \; for\; all\; x \in K^* \; and\; 0 < r < r_0. \quad (81)$$

Now fix $r > 0$ with $\theta r < r_0$. Notice that (79) implies that $K(r)\cap K^* \neq \emptyset$. Pick $x_0 \in K(r)\cap K^*$. Then (81) gives

$$\frac{\mathcal{L}^d\big(\bar{B}_K(x_0,\theta r)\big)}{\alpha_d\theta^d r^d} > 1-\varepsilon_\theta,$$

which contradicts (80).
It remains to prove the claim (80) for every $0 < r < 1/2$ under the condition $K^\circ = \emptyset$. Pick $x \in K(r)$, then $x \in \bar{B}_K(y_i,r)$ for some $1 \leq i \leq N_{2r}$. Let f_i be as in (78) and

$$A = f_i(\bar{B}_K(x,\theta r)) \subset K.$$

Let $h = g_x \circ f_i^{-1}$, where $g_x\colon t \to (t-x)/r$. Then h is a bi-Lipschitz mapping with blip $h = 2C$ since f_i is a C-conformlly bi-Lipschitz mapping with scale factor $1/(2r)$.
We also have

$$h(A) = g_x\big(\bar{B}_K(x,\theta r)\big) = \bar{B}(0,\theta)\cap g_x(K) \subset \bar{B}(0,1).$$

Now we apply Lemma (5.3.8) to A and h with $2C$ instead of C. Notice that $f_i(x) \in A$ and

$$\bar{B}(h(f_i(x)),\theta)\cap h(A) = \bar{B}(0,\theta)\cap g_x(\bar{B}_K(x,\theta r))$$
$$= \bar{B}(0,\theta)\cap g_x(K) = g_x(\bar{B}_K(x,\theta r)).$$

Hence, Lemma (5.3.8) gives

$$\frac{\mathcal{L}^d\big(\bar{B}_K(x,\theta r)\big)}{\alpha_d\theta^d r^d} = \frac{\mathcal{L}^d\left(g_x\big(\bar{B}_K(x,\theta r)\big)\right)}{\alpha_d\theta^d} = \frac{\mathcal{L}^d\left(\bar{B}\big(h(f_i(x)),\theta\big)\right)\cap h(A)}{\alpha_d\theta^d} \leq 1-\varepsilon_\theta.$$

This proves the claim (80) and the proof is complete.

We contain a brief introduction of self-conformal sets, the WSC and the BDP. For more details, see [217], [227]. Recall that a C^1-map $S\colon V \to \mathbb{R}^d$ is conformal on an open subset $V \subset \mathbb{R}^d$ if for each $\in V$, the differential $S'(x)$ is a similarity matrix, i.e., a scalar multiple of an orthogonal matrix. In such case, we have $|\det S'(x)| = \|S'(x)\|^d$, where $\|S'(x)\| := sup\{|S'(x)y| : |y| = 1\}$ is the operator norm of the matrix $S'(x)$.

Let $X \subset \mathbb{R}^d$ be a compact subset and $S_i: X \to X$ for $1 \le i \le N$. $\{S_i\}_{i=1}^N$ is called a conformal iterated function system if each S_i extends to a C^1 injective conformal map $S_i : V \to V$ on an open connected set $V \supset X$ and

$$0 < \inf_{x\in V} \|S_i'(x)\| \le \sup_{x\in V} \|S_i'(x)\| < 1 \, for\, 1 \le i \le N. \tag{82}$$

For such an IFS, the associated invariant set K is called a self-conformal set.
We use the following sets of indices:

$$\Sigma^0 = \{\emptyset\}, \Sigma^k := \{1,\dots,N\}^k \, and\, \Sigma^* := \bigcup_{k\ge 0} \Sigma^k,$$

where $\emptyset$ denotes the empty word. For $I = i_1 i_2 \dots i_k \in \Sigma^k$, we denote by $|I| = k$ the length of I and write $I^- = i_1 i_2 \dots i_{k-1}$. Define $S_I = S_{i1} \circ S_{i2} \circ \dots \circ S_{ik}$ (with the convention that $S_\emptyset$ = identity) and

$$l_I := \inf_{x\in V} \|S_I'(x)\|, \qquad l := \min_{1\le i\le N} l_i, \tag{83}$$
$$L_I := \sup_{x\in V} \|S_I'(x)\|, \qquad L := \max_{1\le i\le N} L_i.$$

For $0 < b < 1$, write

$$I_b := \{I \in \Sigma^* : L_I \le b < L_{I-}\} \, and \, A_b := \{S_I : I \in I_b\}.$$

Definition (5.3.10)[231]: (Weak separation condition [227]). We say that $\{S_i\}N_i = 1$ satisfies the weak separation condition (WSC) if there exists a constant $\gamma \in \mathbb{N}$ and a subset $D \subset X$ with $D^\circ \ne \emptyset$, such that for all $0 < b < 1$ and $x \in X$,

$$\#\{S \in A_b : x \in S(D)\} \le \gamma.$$

To see that the OSC implies the WSC, we can take D to be an open set satisfying the OSC and let $\gamma = 1$.

Definition (5.3.11)[231]: (Bounded distortion property). Let $\{S_i\}_{i=1}^N$ be a conformal IFS. We say that $\{S_i\}_{i=1}^N$ has the bounded distortion property (BDP) if there exists a constant $c_1 > 0$ such that

$$l_I \le L_I \le c_1 l_I \, for\, all\, I \in \Sigma^*. \tag{84}$$

It is well-known that if each $log\|S_i'\|$ is Ho¨lder continuous, then $\{S_i\}_{i=1}^N$ has the BDP. For this, see [243].

We provide preliminaries needed in the proof of Theorems (5.3.18) and (5.3.19).

For any $a > 0$, any bounded subsets $D \subset X$ and $U \subset \mathbb{R}^d$, let

$$A_{a,U,D} = \{S \in A_{adiam\, U} : S(D) \cap U \ne \emptyset\} \, and \, \gamma_{a,D} = \sup_U \#A_a, U, D. \tag{85}$$

Lemma (5.3.12)[231]: ([227]). Let $\{S_i\}_{i=1}^N$ be an IFS of injective C_1 conformal contractions on a compact $X \subset \mathbb{R}^d$ with $X^\circ \ne \emptyset$. Then $\{S_i\}_{i=1}^N$ satisfies the WSC if and only if, for any $a > 0$ and any nonempty subset $D \subset X, \gamma_a, D < \infty$.

Theorem (5.3.13)[231]: ([227]). Let $K \subset \mathbb{R}^d$ be a self-conformal set that has the BDP and satisfies the WSC. Then $0 < H^\alpha(K) < \infty$, where $\alpha = \dim_H K$.

Let $K \subset \mathbb{R}^d$ be a self-conformal set that has the BDP. Recall that, for $x \in K, B_K(x,r) = B(x,r) \cap K$ denotes the open ball in the metric space K.

Lemma (5.3.14)[231]: Given an open ball $B_K(x,r)$ in K, if
$x \in S_I(K) \subset B_K(x,r)$ and $S_I - (K) \subset B_K(x,r)$,
then $lr/(c_1\, diam\, K) \le l_I \le 2r/diamK$.

Proof. By $S_I(K) \subset B_K(x,r)$ and (83), we have

$$l_I\, diam\, K \le diam\, S_I(K) \le 2r.$$

Hence $l_I \le 2r/diam\, K$. By $S_{I-}(K) \subset B_K(x,r)$ and (83), we have

$$L_{I-}\, diam\, K \ge diam\, S_{I-}(K) \ge r.$$

Combining this with (83) and (84), we obtain

$$l_I \geq ll_{I-} \geq lL_{I-}/c_1 \geq lr/(c_1\, diam\, K).$$

The following two lemmas concerns the Ahlfors regularity of self-conformal sets. The proofs can be found in the recent works of Ka¨enma¨ki and his coauthors, see [216] and [225]. However, for the completeness and comfortability, we include the proofs here.

Lemma (5.3.15)[231]: Let $K \subset \mathbb{R}^d$ be a self-conformal set that has the BDP and positive d-dimensional Lebesgue measure. Then K is Ahlfors regular of dimension d.

Proof. Given an open ball $B_K(x,r)$ in K with $2r < diam\, K$, note that $K \not\subset B_K(x,r)$. Hence, we can find an index $I \in \Sigma^*$ such that

$$x \in S_I(K) \subset B_K(x,r)\, and\ S_{I-}(K) \not\subset B_K(x,r).$$

By Lemma (5.3.14), we have $l_I \geq l_r/(c_1\, diam\, K)$. Consequently,

$$2^d r^d = H^d(B(x,r)) \geq H^d(B_K(x,r))$$
$$\geq H^d\big(S_I(K)\big) \geq l_I^d\, H^d(K) \geq H^d(K) l^d r^d/(c_1\, diam\, K)^d. \quad (86)$$

Notice that $H^d(K) > 0$ since K has d-dimensional positive Lebesgue measure. We conclude that K is Ahlfors regular of dimension d.

Lemma (5.3.16)[231]: Let $K \subset \mathbb{R}^d$ be a self-conformal set that has the BDP and satisfies the WSC. Then K is Ahlfors regular of dimension α with $\alpha = \dim_H K$.

Proof. Let $x \in K$ and $0 < r < diamK/2$. We need to estimate the upper and lower bound of $H^\alpha(BK(x,r))$.

For the upper bound, we make use of Lemma (5.3.12). Taking $a = 1, U = B(x,r)$ and $D = K$, Lemma (5.3.12) gives

$$\#\{S \in A_{2r} : S(K) \cap B(x,r) \neq \emptyset\} \leq \gamma_1, K < \infty$$

for all $x \in K$ and $0 < r < diam\, K$. It follows that

$$H^\alpha(B_K(x,r)) \leq \sum_{S\in A_{2r}\ S(K)\cap B(x,r)=\emptyset} H^\alpha S(K) \leq \gamma_1, K(2r)^\alpha H^\alpha(K).$$

For the lower bound, we use the same argument in the proof of Lemma (5.3.15) and obtain the same lower bound as in (86):

$$H^\alpha(BK(x,r)) \geq H^\alpha(K) l^\alpha r^\alpha/(c_1\, diam\, K)^\alpha.$$

Finally, we conclude the Ahlfors regularity of K from the upper and lower bound above since $0 < H^\alpha(K) < \infty$ (by Theorem (5.3.13)).

Lemma (5.3.17)[231]: Let $K \subset \mathbb{R}^d$ be a self-conformal set that has the BDP. If $K^\circ \neq \emptyset$ or K satisfies the WSC, then there exist $x_0 \in K$ and $0 < r\, 0 < diam\, K$ such that for all index $I \in \Sigma^*$,

$$S_I(\bar{B}_K(x_0,r_0)) = S_I(\bar{B}(x_0,r_0)) \cap K.$$

Proof. We begin with the case $K^\circ \neq \emptyset$. Clearly, there are $x_0 \in K$ and $0 < r\, 0 <$ diamK with $\bar{B}(x_0,r_0) \subset K$. Then for any $I \in \Sigma^*$, we have $S_I(\bar{B}_K(x_0,r_0)) = S_I(B(x_0,r_0) \cap K) = S_I(B(x_0,r_0)) = S_I(B(x_0,r_0)) \cap K$,

$$since\ S_I(B(x_0,r_0)) \subset K.$$

Now suppose that K satisfies the WSC. Pick $x \in K$ and $0 < r < diam\, K$ with $\bar{B}(x,r) \subset X$.[1] Let $A_0 = \bar{B}(x,r)$. We can obtain a sequence of sets A_0, A_1,... by induction as follows: suppose that A_{k-1} is determined, if

$$S_{I_k}(A_{k-1} \cap K) \subsetneq S_{I_k}(Ak-1) \cap K\ for\ some\ I_k \in \Sigma^*,$$

let $A_k = S_{I_k}(A_{k-1})$; if such I_k does not exist, we stop the procedure.

Let $a = (2r)^{-1}$. We claim that the above induction procedure stops after at most $\gamma_{a,K}$ steps, where $\gamma_{a,K}$ is defined in (85). For otherwise, there would be $A_0, A_1, \ldots, A_n$ and $I_1, \ldots, I_n$ with $n > \gamma_{a,K}$ such that $A_0 = \bar{B}(x,r), A_k = S_{I_k}(A_{k-1})$ and

$$S_{I_k}(A_{k-1} \cap K) \subsetneq S_{I_k}(A_{k-1}) \cap K \; for \; 1 \le k \le n. \tag{87}$$

To prove the claim, we need to deduce a contradiction. Note that $A_n = S_{I_n} \ldots I_1(A_0) = S_{I_n} \ldots I_1(\bar{B}(x,r))$. So

$$L_{I_n} \ldots I_1 = (2r)^{-1} L_{I_n} \ldots I_1 \cdot 2r \ge adiam\, A_n.$$

This together with (87) implies that for each $1 \le k \le n$, we can find $J_k \in \Sigma^*$ such that $S_{I_n} \ldots I_{k+1} J_k \in A_a$ diam A_n and

$$\emptyset \ne S_{I_k}(A_{k-1}) \cap S_{Jk}(K) \not\subset S_{I_k}(A_{k-1} \cap K). \tag{88}$$

Hence, $A_n \cap S_{I_n} \ldots I_{k+1} J_k(K) = S_{I_n} \ldots I_{k+1} \left(S_{I_k}(A_{k-1})\right) \cap S_{J_k}(K) \ne \emptyset$. Therefore,

$$S_{J_n}, S_{I_n} J_{n-1}, \ldots, S_{I_n \ldots I_2} J_1 \in A_a, A_{n,K},$$

where $A_a, A_{n,K}$ is defined by (85). Thus, $n > \gamma_{a,K}$ implies that there are $1 \le m < k \le n$ such that $S_{I_n} \ldots I_{k+1} J_k = S\_(I_n \ldots I_{m+1\, J_m}, i.e., S_{J_k} = S_{I_{k \ldots I_{m+1} J_m}}$. Then we have

$$S_{I_k}(A_{k-1}) \cap S_{J_k}(K) = S_{I_k}\left(A_{k-1} \cap S_{I_{k-1 \ldots I_{m+1} J_m}}(K)\right) \subset S_{I_k}(A_{k-1} \cap K).$$

This contradicts (88) and so the claim follows.

Clearly, the claim implies that we can find $A_n = S_{I_n \ldots I_1}(\bar{B}(x,r))$ satisfying

$$S_I(A_n \cap K) = S_I(A_n) \cap K \text{ for all } I \in \Sigma^*.$$

Let $x_0 = S_{I_n \ldots I_1}(x)$ and $r_0 = l_{I_n \ldots I_1} r$, then $\bar{B}(x_0, r_0) \subset A_n$. For all $I \in \Sigma^*$,

$$\begin{aligned} S_I(\bar{B}K(x_0,r_0)) = \; & S_I(\bar{B}(x_0,r_0) \cap K) = S_I(\bar{B}(x_0,r_0)) \cap S_I(K) \\ &= S_I(\bar{B}(x_0,r_0)) \cap S_I(A_n) \cap S_I(K) \\ &= S_I(\bar{B}(x_0,r_0)) \cap S_I(A_n) \cap K = S_I(\bar{B}(x_0,r_0)) \cap K. \end{aligned}$$

Hence such x_0 and r_0 are desired.

Theorem (5.3.18)[231]: Let $K \subset \mathbb{R}^d$ be a self-conformal set that has the bounded distortion property (BDP). Then

$$K^\circ \ne \emptyset \Leftrightarrow K = K^\circ \Leftrightarrow K \text{ is a BBI set of dimension d.}$$

Theorem (5.3.19)[231]: Let $K \subset \mathbb{R}^d$ be a self-conformal set that has the BDP and satisfies the WSC. If $\dim_H K = d$, then $K^\circ \ne \emptyset$.

Proof of Theorems (5.3.18) and (5.3.19). We remark that, in Theorem (5.3.18), $K = \overline{K^\circ}$ follows from $K^\circ \ne \emptyset$ immediately. Indeed, K° contains $\bigcup_{I \in \Sigma^*} S_I(K^\circ)$ and the latter is dense in K.

Therefore, by Theorem (5.3.9), it suffices to prove that, if $K^\circ = \emptyset$ (for Theorem (5.3.18)) or K satisfies the WSC with $\dim_H K = d$ (for Theorem (5.3.19)), then K is a BBI set of dimension d.

We first observe that, by Lemmas (5.3.15) and (5.3.16), K is Ahlfors regular of dimension d in both cases. Now let $B_K(x_1,r_1)$ and $B_K(x_2,r_2)$ be two open balls in K with $0 < r_1, r_2 < diam\, K$. To complete the proof, we need to find constants θ, C (independent of x_1, x_2 and r_1, r_2), a point $x_3 \in B_K(x_1,r_1)$ and aC-conformally bi-Lipschitz mapping h from $B_K(x_3, \theta r_1)$ into $B_K(x_2,r_2)$ with scale factor r_2/r_1.

For this, suppose without loss of generality that $K \not\subset B_K(x_1,r_1)$ and $K \subset B_K(x_2,r_2)$. Thus, we can find two indices $I_1, I_2 \in \Sigma^*$ such that

$$S_{I_1}(K) \subset B_K(x_1,r_1), \qquad S_{I_1^-}(K) \not\subset B_K(x_1,r_1)$$

And

$$S_{I_2}(K) \subset B_K(x_2, r_2), \qquad S_{I_2^-}(K) \not\subset B_K(x_2, r_2).$$

It follows from Lemma (5.3.14) that

$$\frac{l}{c_1 \, diam\, K} \leq \frac{l_{I_i}}{r_i} \leq \frac{2}{diam\, K}, for\ i = 1, 2. \tag{89}$$

By Lemma (5.3.17), in both cases, we can find $x_0 \in K$ and $0 < r_0 < diam\, K$ such that for all index $I \in \Sigma^*$,

$$S_I(\bar{B}_K(x_0, r_0)) = S_I(\bar{B}(x_0, r_0)) \cap K. \tag{90}$$

Let $x_3 = S_{I_1}(x_0) \in S_{I_1}(K) \subset B_K(x_1, r_1) and\ \theta = lr_0/(c_1\, diam\, K)$. Then (89) gives $\theta r_1 \leq l_{I_1} r_0$. This together with (83) and (90) implies

$$\bar{B}_K(x_3, \theta r_1) \subset \bar{B}_K(x_3, l_{I_1} r_0) \subset S_{I_1}\big(\bar{B}(x_0, r_0)\big) \cap K = S_{I_1}\big(\bar{B}_K(x_0, r_0)\big). \tag{91}$$

Therefore, we have $\bar{B}_K(x_3, \theta r_1) \subset B_K(x_1, r_1)$ since

$$S_{I_1}(\bar{B}_K(x_0, r_0)) \subset S_{I_1}(K) \subset B_K(x_1, r_1).$$

Let $h = S_{I_2} \circ S_{I_1}^{-1}$. By (91), we have

$$h(\bar{B}_K(x_3, \theta r_1)) \subset h\Big(S_{I_1}\big(\bar{B}_K(x_0, r_0)\big)\Big) = S_{I_2}\big(\bar{B}_K(x_0, r_0)\big) \subset B_K(x_2, r_2).$$

Hence h maps $(\bar{B}_K(x_3, \theta r_1))$ into $B_K(x_2, r_2)$.

It remains to show that $h = S_{I_2} \circ S_{I_1}^{-1}$ is a C-conformally bi-Lipschitz mapping with scale factor r_2/r_1 for a constant C. Indeed, we have

$$\frac{l_{I_2}}{c_1 l_{I_1}} \leq \frac{l_{I_2}}{L_{I_1}} \leq \frac{\left|S_{I_2} \circ S_{I_1}^{-1}(x) - S_{I_2} \circ S_{I_1}^{-1}(x)\right|}{|x - y|} \leq \frac{l_{I_2}}{L_{I_1}} \leq \frac{l_{I_2}}{c_1 l_{I_1}},$$

for $x, y \in \bar{B}_K(x_3, \theta r_1)$. This together with (89) gives

$$C^{-1}(r_2/r_1)|x - y| \leq |h(x) - h(y)| \leq C(r_2/r_1)|x - y|,$$

where $C = 2c_1^2/l$.

Example (5.3.20)[231]: Let $K \subset \mathbb{R}^d$ be a self-conformal set that has the BDP. Suppose that K has positive d-dimensional Lebesgue measure but no interior, then K is a BPI set but not BBI set. For an explicit construction of such sets, see [160].

For convenience, we first recall the content of Example (5.3.20) below. Let $K \subset \mathbb{R}^d$ be a self-conformal set that has the BDP. Suppose that K has positive d-dimensional Lebesgue measure but no interior, then K is a BPI set but not BBI set.

To see this, we use Theorem (5.3.18) to obtain that K is not a BBI set. It remains to show that K is a BPI set. Note that K is Ahlfor regular by Lemma (5.3.15).

Now let $B_K(x_1, r_1)$ and $B_K(x_2, r_2)$ be two open balls in K. Suppose without loss of generality that $K \not\subset B_K(x_1, r_1)$ and $K \not\subset B_K(x_2, r_2)$. Thus, we can find two indices $I_1, I_2 \in \Sigma^*$ such that

$$S_{I_1}(K) \subset B_K(x_1, r_1), \qquad S_{I_1^-}(K) \not\subset B_K(x_1, r_1)$$

And

$$S_{I_2}(K) \subset B_K(x_2, r_2), \qquad S_{I_2^-}(K) \not\subset B_K(x_2, r_2).$$

Since K has positive Lebesgue measure, $S_{I_1}(K)$ and $S_{I_2}(K)$ are subsets in $B_K(x_1, r_1)$ and $B_K(x_2, r_2)$ with relatively large measure, respectively. In fact, by Lemma (5.3.14), one can see that

$$H^d\big(S_{I_1}(K)\big) \geq l_{I_1}^d\, H^d(K) \geq \left(\frac{l}{c_1\, diam\, K}\right)^d H^d(K) \cdot r_1^d.$$

By the same reason, this inequality also holds if we replace I_1, r_1 by I_2, r_2.

Finally, let $h = S_{I_2} \circ S_{I_1}^{-1} : S_{I_1}(K) \to S_{I_2}(K)$. Using same argument in corresponding part, one can check that h is a C-connformally bi-Lipschitz mapping with scale factor r_2/r_1 for a constant C. Consequently, K is a BPI set.

Corollary (5.3.21)[269]: (See [231])[269]: Let $(A_n)_{n\geq 1}$ be a sequence of nonempty compact sets in a metric space (X, ρ_{j-1}) with $A_n \overset{d_H}{\longrightarrow} A$ for some nonempty compact set $A \subset X$. Let $(x_{n+j-1})_{n\geq 1}$ be a sequence of points in X with $x_{n+j-1} \in A_n$ and $x_{n+j-1} \to x_{j-1}$ for some $x_{j-1} \in A$. Fix $\epsilon \geq 0$. For each $\varepsilon > 0$, there exists an $n_\varepsilon \geq 1$ such that

$$\bar{B}(x_{n+j-1}, 1+2\epsilon) \cap A_n \subset \left(B(x_{j-1}, 1+2\epsilon) \cap A\right)_\varepsilon \quad \text{for } n > n_\varepsilon.$$

Proof. Fix $\varepsilon > 0$. Since

$$\bigcap_{\delta>0} (\bar{B}(x_{j-1}, (1+2\epsilon)+\delta) \cap A_\delta) = \bar{B}(x_{j-1}, 1+2\epsilon) \cap A,$$

there exists a $\delta > 0$ such that

$$\bar{B}(x_{j-1}, (1+2\epsilon)+\delta) \cap A_\delta \subset \left(\bar{B}(x_{j-1}, 1+2\epsilon) \cap A\right)_\varepsilon.$$

Since $A_n \overset{d_H}{\longrightarrow} A$ and $x_{n+j-1} \to x_{j-1}$, there exists an $n_\varepsilon \geq 1$ such that

$$\begin{aligned} \bar{B}(x_{n+j-1}, 1+2\epsilon) \cap A_n &\subset \bar{B}(x_{j-1}, (1+2\epsilon)+\delta) \cap A_\delta \\ &\subset \left(\bar{B}(x_{j-1}, 1+\epsilon) \cap A\right)_\varepsilon, \end{aligned}$$

for $n > n_\varepsilon$.

For a bi-Lipschitz mapping h_m which maps (X, ρ_j) to (Y, ρ_{j+1}), its bi-Lipschitz constant blip h_m is defined by

$$\text{blip } h_m := \inf\left\{c \geq 1 : c^{-1}\rho_j(x_j, x_{j+1}) \leq \rho_{j+1}\left(f_m(x_j), f_m(x_{j+1})\right) \leq c\rho_j(x_j, x_{j+1})\right\}.$$

Corollary (5.3.22)[269]: (See [231]). Let A_n be a nonempty compact set in a metric space (X, ρ_{j-1}) and $(h_m)_n : A_n \to X$ a bi-Lipschitz mapping with blip $(h_m)_n \leq 1+\epsilon$ for all $n \geq 1$. Suppose that $A_n \overset{d_H}{\longrightarrow} A$ and $(h_m)_n(A_n) \overset{d_H}{\longrightarrow} A^*$, then there is a bi-Lipschitz bijection h maps A onto A^* with blip $h_m \leq 1+\epsilon$.

Proof. Pick a countable dense subset $\left\{(x_{j-1})_{m_0} : m_0 \geq 1\right\}$ of A and a countable dense subset $\left\{(x_{j-1})^*_{-m_0} : m_0 \geq 1\right\}$ of A^*. Since $A_n \overset{d_H}{\longrightarrow} A$ and $(h_m)_n(A_n) \overset{d_H}{\longrightarrow} A^*$, for each $m_0 \geq 1$, we can find two sequences $\left((x_{j-1})_{m_0,n}\right)_{n\geq 1}$ and $\left(x_{j-1} \propto_{-m_0,n}\right)_{n\geq 1}$ such that $x_{m_0,n}, x_{-m_0,n} \in A_n$ for every $n \geq 1$ and

$$(x_{j-1})_{m_0,n} \to (x_{j-1})_{m_0}, (h_m)_n\left((x_{j-1})_{-m_0,n}\right) \to (x_{j-1})^*_{-m_0} \tag{92}$$

as $\to \infty$.

We claim that, for every $m_0 \geq 1$, the two sequences $\left((x_{j-1})_{-m_0,n}\right)_{n\geq 1}$ and $\left((h_m)_n\left((x_{j-1})_{m_0,n}\right)_{n\geq 1}\right)$ both have a convergent subsequence. To see this, fix $m_0 \geq 1$. Notice that for each $k \geq 1$, there are finitely many balls $(B_{k,i})_i$ of radius $1/k$ which cover A since A is compact. We can find a ball B_k in $(B_{k,i})_i$ such that B_k contains infinitely many points in $\left((x_{j-1})_{-m_0,n}\right)_{n\geq 1}$ since $\text{dist}\left((x_{j-1})_{-m_0,n}, A\right) \to 0$ as $n \to \infty$. It follows that there exists a subsequence $(y_k)_{k\geq 1}$ of $\left(\left(x_{j-1}\right)_{-m_0,n}\right)_{n\geq 1}$ such that $y_k \in B_k \cap \left\{\left(x_{j-1}\right)_{-m_0,n} : n \geq 1\right\}$. Now pick $a_k \in B_k \cap A$, then $(a_k)_{k\geq 1}$ has a convergence subsequence since A is compact. Consequently, so does $(y_k)_{k\geq 1}$ since $\rho_{j-1}(y_k, a_k) \leq 2/k \to 0$ as $k \to \infty$. Therefore, $\left(\left(x_{j-1}\right)_{-m_0,n}\right)_{n\geq 1}$ also has a convergence subsequence. This argument is also applicable to the sequence $\left((h_m)_n\left(\left(x_{j-1}\right)_{m_0,n}\right)\right)_{n\geq 1}$.

Combining the claim above with Cantor's diagonal argument, by taking a subsequence of $(n)_{n\geq 1}$ if necessary, we can assume that, for each $m_0 \geq 1$,

$$\left(x_{j-1}\right)_{-m_0,n} \to \left(x_{j-1}\right)_{-m_0}, (h_m)_n\left(\left(x_{j-1}\right)_{m_0,n}\right) \to \left(x_{j-1}\right)^*_{m_0}, \tag{93}$$

as $n \to \infty$ for some $\left(x_{j-1}\right)_{-m_0} \in A$ and $\left(x_{j-1}\right)^*_{m_0} \in A^*$. Now let $A_0 = \left\{\left(x_{j-1}\right)_{m_0} : m_0 \neq 0\right\}$ and $A_0^* = \left\{\left(x_{j-1}\right)^*_{m_0} : m_0 \neq 0\right\}$. Notice that

$$\left(x_{j-1}\right)_{m_0} = \left(x_{j-1}\right)_{m_0'} \Leftrightarrow \left(x_{j-1}\right)^*_{m_0} = \left(x_{j-1}\right)^*_{m_0'}. \tag{94}$$

Indeed, since blip $(h_m)_n \leq 1+\epsilon$ for all $n \geq 1$, we have

$$\begin{aligned}\left(x_{j-1}\right)_{m_0} = \left(x_{j-1}\right)_{m_0'} &\Leftrightarrow \lim_{n\to\infty} \rho_{j-1}\left(\left(x_{j-1}\right)_{m_0,n,} \left(x_{j-1}\right)_{m_0',n}\right) = 0 \\ &\Leftrightarrow \lim_{n\to\infty} \rho_{j-1}\left((h_m)_n(\left(x_{j-1}\right)_{m_0,n}), (h_m)_n(\left(x_{j-1}\right)_{m_0',n})\right) \\ &= 0 \Leftrightarrow \left(x_{j-1}\right)^*_{m_0} = \left(x_{j-1}\right)^*_{m_0'}.\end{aligned}$$

It follows from (94) that the mapping $(h_m)_0 : \left(x_{j-1}\right)_{m_0} \to \left(x_{j-1}\right)^*_{m_0}$ is a bijection from A_0 onto A_0^*. Note that A_0 and A_0^* are dense in A and A^*, respectively, since so are $\left\{\left(x_{j-1}\right)_{m_0} : m \geq 1\right\}$ and $\left\{\left(x_{j-1}\right)^*_{-m_0} : m_0 \geq 1\right\}$. We claim that

$$(h_m)_0 \text{ is a bi-Lipschitz bijection from } A_0 \text{ onto } A_0^* \text{ with blip } (h_m)_0 \leq 1+\epsilon. \tag{95}$$

A standard result in mathematic analysis says that a uniformly continuous function from a dense subset of a metric space to a complete metric space has a uniformly continuous extension. (see [244]).

Hence, if the claim (95) is true, then $(h_m)_0$ can be extend to a uniformly continuous function $h_m : A \to A^*$. For distinct $y, y' \in A$, pick $y_n, y_n' \in A_0$ with $y_n \to y$ and $y_n' \to y'$, then the claim (95) implies

$$\begin{aligned}(1+\epsilon)^{-1}\rho_{j-1}(y,y') &= (1+\epsilon)^{-1} \lim_{n\to\infty} \rho_{j-1}(y_n, y_n') \\ &\leq \lim_{n\to\infty} \rho_{j-1}((h_m)_0(y_n), (h_m)_0(y_n')) \\ &= \lim_{n\to\infty} \rho_{j-1}(h_m(y_n), h_m(y_n')) \\ &= \rho_{j-1}(h_m(y), h_m(y')) \\ &= \lim_{n\to\infty} \rho_{j-1}(h_m(y_n), h_m(y_n')) \\ &= \lim_{n\to\infty} \rho_{j-1}((h_m)_0(y_n), (h_m)_0(y_n')) \\ &\leq (1+\epsilon) \lim_{n\to\infty} \rho_{j-1}(y_n, y_n') = (1+\epsilon)\rho_{j-1}(y,y').\end{aligned}$$

Thus, h_m is bi-Lipschitz with blip$h_m \leq (1+\epsilon)$. It follows that h_m maps $A = \overline{A_0}$ onto $\overline{h_m(A_0)} = \overline{A_0^*} = A^*$. Consequently, h_m is desired.

It remains to prove the claim (95). For this, pick $a, a' \in A_0$. By (92) and (93), there exist $a_n, a_n' \in A_n$ such that

$$a_n \to a, a_n' \to a' \text{ and } (h_m)_n(a_n) \to (h_m)_0(a), (h_m)_n(a_n') \to (h_m)_0(a')$$

as $n \to \infty$. Since blip $(h_m)_n \leq (1+\epsilon)$ for all $n \geq 1$, we have

$$\begin{aligned}(1+\epsilon)^{-1}\rho_{j-1}(a,a') &= (1+\epsilon)^{-1} \lim_{n\to\infty} \rho_{j-1}(a_n, a_n') \\ &\leq \lim_{n\to\infty} \rho_{j-1}((h_m)_n(a_n), (h_m)_n(a_n')) \\ &= \rho_{j-1}((h_m)_0(a), (h_m)_0(a')) \\ &= \lim_{(n\to\infty)} \rho_{j-1}((h_m)_n(a_n), (h_m)_n(a_n')) \\ &\leq (1+\epsilon) \lim_{n\to\infty} \rho_{j-1}(a_n, a_n') = (1+\epsilon)\rho_{j-1}(a,a').\end{aligned}$$

This proves the claim and the proof is complete.

Corollary (5.3.23)[269]: Let K_m be a compact subset of $\mathbb{R}^d$ with $\mathring{K}_m = \emptyset$. Fix $\epsilon \geq 0$. For each $\epsilon > -\frac{1}{2}$, there exists $\varepsilon_{1+2\epsilon} > 0$ such that for any compact subset A of K_m, any $x_{j-1} \in A$ and any bi-Lipschitz mapping $h_m: A \to \bar{B}(0,1)$ with blip $h_m \leq (1+\epsilon)$, we have

$$\frac{\mathcal{L}^d\left(B\left(h_m(x_{j-1}), 1+2\epsilon\right) \cap h_m(A)\right)}{\alpha_d(1+2\epsilon)^d} \leq 1 - \varepsilon_{1+2\epsilon},$$

where α_d denotes the Lebesgue measure of the unit ball in $\mathbb{R}^d$.

Proof. Let us argue by contradiction. Suppose the lemma were false. Then for some $\epsilon > -\frac{1}{2}$, there exist compact subsets $A_n \subset K_m, \left(x_{j-1}\right)_n \in A_n$ and bi-Lipschitz mappings $(h_m)_n: A_n \to \bar{B}(0,1)$ with blip $(h_m)_n \leq (1+\epsilon)$ for $n \geq 1$ such that

$$\frac{\mathcal{L}^d\left(B\left((h_m)_n(x_{j-1}), 1+2\epsilon\right) \cap (h_m)_n(A)\right)}{\alpha_d(1+2\epsilon)^d} \to 1 \text{ as } n \to \infty. \quad (96)$$

Notice that K_m and $B(0,1)$ are both compact. According to Lemma 2.1, by taking a subsequence of $(n)_{n\geq 1}$ if necessary, we can assume that

$$A_n \xrightarrow{d_H} A, \qquad (h_m)_n(A_n) \xrightarrow{d_H} A^*, (h_m)_n\left(\left(x_{j-1}\right)_n\right) \to x_{j-1}^* \quad (97)$$

as $n \to \infty$ for some nonempty compact sets $A \subset K_m, A^* \subset \bar{B}(0,1)$ and some point $x_{j-1}^* \in \bar{B}(0,1)$.

By (97), we can apply Corollary (5.3.21) to $((h_m)_n(A_n))$ and $\left((h_m)_n\left(\left(x_{j-1}\right)_n\right)\right)$. This gives that for each $\varepsilon > 0$, there is an nε such that

$$\bar{B}\left((h_m)_n\left(\left(x_{j-1}\right)_n\right), 1+2\epsilon\right) \cap (h_m)_n(A_n) \subset D_\varepsilon \text{ for } n \geq n_\varepsilon, \quad (98)$$

where $D = B(x_{j-1}^*, 1+2\epsilon) \cap A^*$.

It follows from (96) and (98) that

$$\frac{\mathcal{L}^d(D_\varepsilon)}{\alpha_d(1+2\epsilon)^d} \geq \lim_{n\to\infty} \frac{\mathcal{L}^d\left(B\left((h_m)_n\left(\left(x_{j-1}\right)_n\right), 1+2\epsilon\right) \cap (h_m)_n(A_n)\right)}{\alpha_d(1+2\epsilon)^d} = 1 \text{ for every } \varepsilon > 0.$$

So we have

$$1 \geq \frac{\mathcal{L}^d\left(B\left(x_{j-1}^*, 1+2\epsilon\right) \cap A^*\right)}{\alpha_d(1+2\epsilon)^d} = \mathcal{L}^d(D)\alpha_d(1+2\epsilon)^d = \lim_{\varepsilon\to 0} \frac{\mathcal{L}^d(D_\varepsilon)}{\alpha_d(1+2\epsilon)^d} \geq 1,$$

where the last equality comes from the fact that D is compact. Thus,

$$\mathcal{L}^d\left(B\left(x_{j-1}^*, 1+2\epsilon\right)\cap A^*\right) = \alpha_d(1+2\epsilon)^d = \mathcal{L}^d\left(\bar{B}\left(x_{j-1}^*, 1+2\epsilon\right)\right).$$

So $\bar{B}(x_{j-1}^*, 1+2\epsilon)\cap A^*$ is a compact subset of $B(x_{j-1}^*, 1+2\epsilon)$ of full Lebesgue measure. We have

$$\bar{B}\left(x_{j-1}^*, 1+2\epsilon\right)\subset A^*. \quad (99)$$

Using (97) again, Corollary (5.3.22) gives a bi-Lipschitz bijection h_m from A onto A^*. Now Brouwer's invariance of domain theorem asserts that $\mathring{A}\neq\emptyset$ since so does A^* (by (99)). Consequently, $\mathring{K}_m\supset\mathring{A}\neq\emptyset$, contrary to the condition $\mathring{K}_m=\emptyset$.

Corollary (5.3.24)[269]: Let $K_m\subset\mathbb{R}^d$ be a compact BBI set of dimension d, then $K_m^\circ\neq\emptyset$.

Proof. [231]. Let $K_m\subset\mathbb{R}^d$ be a compact BBI set. By definition, K_m together with the Euclidean distance ρ_{j-1} is a BBI space. To avoid confusion, we use $B(x_{j-1}, r_{j-1})$ and $B_{K_m}(x_{j-1}, r_{j-1})$ to denote open balls in the two metric spaces $\mathbb{R}^d$ and K_m, respectively. For $x_{j-1}\in K_m$, we have $B_{K_m}(x_{j-1}, r_{j-1}) = B(x_{j-1}, r_{j-1})\cap K_m$.

Suppose without loss of generality that diam $K_m = 1$. For $0<r_{j-1}\leq 1$, denote by $N_{r_{j-1}}$ the largest number of disjoint balls of radius r_{j-1} centered in K_m. For each $0<r_{j-1}<1/2$, let $\left\{B_{K_m}\left(\left(x_{j-1}\right)_i, 2r_{j-1}\right)\right\}_{i=1}^{N_{2r_{j-1}}}$ be a disjoint family of balls with $\left(x_{j-1}\right)_i\in K_m$. Then

$$K_m\subset\bigcup_{i=1}^{N_{2r_{j-1}}} B_{K_m}\left(\left(x_{j-1}\right)_i, 4r_{j-1}\right). \quad (100)$$

Notice that K_m can be regarded as a ball of radius 1 in the BBI space K_m. By the definition of BBI space (Definition 2.3), there are two constants $\epsilon\geq 0$ such that for each $1\leq i\leq N_{2r_{j-1}}$, we can find a closed ball $B_{K_m}(y_i, 2(1+2\epsilon)r_{j-1})\subset B_{K_m}\left(\left(x_{j-1}\right)_i, 2r_{j-1}\right)$ with $y_i\in K_m$ and a $(1+\epsilon)$-conformlly bi-Lipschitz mapping

$$(f_m)_i\colon \bar{B}_{K_m}\left(y_i, 2(1+2\epsilon)r_{j-1}\right)\to K_m \text{ with scale factor } \frac{1}{2r_{j-1}}. \quad (101)$$

Let

$$K_m\left(r_{j-1}\right) = \bigcup_{i=1}^{N_{2r_{j-1}}} \bar{B}_{K_m}\left(y_i, (1+2\epsilon)r_{j-1}\right).$$

It follows from (100) that

$$\begin{aligned}\mathcal{L}^d\left(K_m(r_{j-1})\right) &= \sum_{i=1}^{N_{2r_{j-1}}} \mathcal{L}^d\left(\bar{B}_{K_m}(y_i, (1+2\epsilon)r_{j-1})\right) = \sum_{i=1}^{N_{2r_{j-1}}} \mathcal{L}^d\left(\bar{B}_{K_m}(y_i, 6r_{j-1})\right)\\ &\geq 6^{-d}(1+2\epsilon)^d\mathcal{L}^d\left(\bigcup_{i=1}^{N_{2r_{j-1}}} B_{K_m}\left(y_i, 6r_{j-1}\right)\right)\\ &\geq 6^{-d}(1+2\epsilon)^d\mathcal{L}^d\left(\bigcup_{i=1}^{N_{2r_{j-1}}} B_{K_m}\left(\left(x_{j-1}\right)_i, 4r_{j-1}\right)\right)\\ &\geq 6^{-d}(1+2\epsilon)^d\mathcal{L}^d(K_m) \qquad (\geq\ 6-d\theta dLd(K). \quad (102)\end{aligned}$$

for all $0 < r_{j-1} < 1/2$. Now suppose the theorem were false, i.e., $\overset{\circ}{K_m} = \emptyset$. We claim that, for every $0 < r_{j-1} < 1/2$,

$$\frac{\mathcal{L}^d\left(\bar{B}_{K_m}\left(x_{j-1},(1+2\epsilon)r_{j-1}\right)\right)}{\alpha_d(1+2\epsilon)^d r_{j-1}^d} \leq 1-\varepsilon_{1+2\epsilon} \text{ for all } x_{j-1} \in K_m\left(r_{j-1}\right). \quad (103)$$

Here $\varepsilon_{1+2\epsilon}$ is the constant in Corollary (5.3.23) with $2(1+\epsilon)$ instead of $(1+\epsilon)$.

By assumption, K_m is Ahlfors regular with dimension d and thus, $\mathcal{L}^d(K_m) > 0$. We shall show that (102) and (103) contradict this. In fact, by Lebesgue density theorem and Egoroff's theorem, there is a measurable subset $K_m^* \subset K_m$ and $\left(r_{j-1}\right)_0 > 0$ such that $\mathcal{L}^d(K_m \backslash K_m^*) < 6^{-d}(1+2\epsilon)^d\, \mathcal{L}^d(K_m)$ and that

$$\frac{\mathcal{L}^d\left(\bar{B}_{K_m}\left(x_{j-1},r_{j-1}\right)\right)}{\alpha_d r_{j-1}^d} > 1-\varepsilon_{1+2\epsilon}$$

$$\text{for all } x_{j-1} \in K_m^* \text{ and } 0 < r_{j-1} < \left(r_{j-1}\right)_0. \quad (104)$$

Now fix $r_{j-1} > 0$ with $(1+2\epsilon)r_{j-1} < \left(r_{j-1}\right)_0$. Notice that (102) implies that $K_m(r_{j-1}) \cap K_m^* \neq \emptyset$. Pick $\left(x_{j-1}\right)_0 \in K_m(r_{j-1}) \cap K_m^*$. Then (104) gives

$$\frac{\mathcal{L}^d\left(\bar{B}_{K_m}\left(\left(x_{j-1}\right)_0,(1+2\epsilon)r_{j-1}\right)\right)}{\alpha_d\left((1+2\epsilon)1+2\epsilon\right)^d r_{j-1}^d} > 1-\varepsilon_{1+2\epsilon},$$

which contradicts (103).

It remains to prove the claim (103) for every $0 < r_{j-1} < 1/2$ under the condition $\overset{\circ}{K_m} = \emptyset$. Pick $x_{j-1} \in K_m(r_{j-1})$, then $x_{j-1} \in \bar{B}_{K_m}(y_i, r_{j-1})$ for some $1 \leq i \leq N_{2r_{j-1}}$. Let $(f_m)_i$ be as in (101) and

$$A = (f_m)_i\left(\bar{B}_{K_m}(x_{j-1},(1+2\epsilon)r_{j-1})\right) \subset K_m.$$

Let $h_m = g_{x_{j-1}} \circ (f_m)_i^{-1}$, where $g_{x_{j-1}}: t \to (t-x_{j-1})/r_{j-1}$. Then h_m is a bi-Lipschitz mapping with blip $h_m = 2(1+\epsilon)$ since $(f_m)_i$ is a $(1+\epsilon)$-conformlly bi-Lipschitz mapping with scale factor $1/(2r_{j-1})$.

We also have

$$h_m(A) = g_{x_{j-1}}\left(\bar{B}_{K_m}\left(x_{j-1},(1+2\epsilon)r_{j-1}\right)\right)$$
$$= \bar{B}(0,1+2\epsilon) \cap g_{x_{j-1}}(K_m) \subset \bar{B}(0,1).$$

Now we apply Corollary (5.3.23) to A and h_m with $2(1+\epsilon)$ instead of $(1+\epsilon)$. Notice that $(f_m)_i(x_{j-1}) \in A$ and

$$\bar{B}\left(h_m\left((f_m)_i\left(x_{j-1}\right)\right),1+2\epsilon\right) \cap h_m(A)$$
$$= \bar{B}(0,1+2\epsilon) \cap g_{x_{j-1}}\left(\bar{B}_{K_m}\left(x_{j-1},(1+2\epsilon)r_{j-1}\right)\right)$$
$$= \bar{B}(0,1+2\epsilon) \cap g_{x_{j-1}}(K_m) = g_{x_{j-1}}\left(\bar{B}_{K_m}\left(x_{j-1},(1+2\epsilon)r_{j-1}\right)\right).$$

Hence, Corollary (5.3.23) gives

$$\frac{\mathcal{L}^d\left(\bar{B}_{K_m}\left(x_{j-1},(1+2\epsilon)r_{j-1}\right)\right)}{\alpha_d(1+2\epsilon)^d r_{j-1}^d} = \frac{\mathcal{L}^d\left(g_{x_{j-1}}\left(\bar{B}_{K_m}\left(x_{j-1},(1+2\epsilon)r_{j-1}\right)\right)\right)}{\alpha_d(1+2\epsilon)^d}$$

$$= \frac{\mathcal{L}^d\left(\bar{B}\left(h_m\left((f_m)_i(x_{j-1})\right), 1+2\epsilon\right)\right) \cap h_m(A)}{\alpha_d(1+2\epsilon)^d} \le 1 - \varepsilon_{1+2\epsilon}.$$

This proves the claim (103) and the proof is complete.

Corollary (5.3.25)[269]: (See [231])[269]: Given an open ball $B_{K_m}(x_{j-1}, r_{j-1})$ in K_m, if

$$x_{j-1} \in S_I(K_m) \subset B_{K_m}(x_{j-1}, r_{j-1}) \text{ and } S_I - (K_m) \subset B_{K_m}(x_{j-1}, r_{j-1}),$$

then $lr_{j-1}/(c_1 \text{ diam } K_m) \le l_I \le 2r_{j-1}/\text{diam} K_m$.

Proof. By $S_I(K_m) \subset B_{K_m}(x_{j-1}, r_{j-1})$ and (3.2), we have

$$l_I \text{ diam } K_m \le \text{diam } S_I(K_m) \le 2r_{j-1}.$$

Hence $l_I \le 2r_{j-1}/\text{diam } K_m$. By $S_{I-}(K_m) \subset B_{K_m}(x_{j-1}, r_{j-1})$ and (3.2), we have

$$L_{I-} \text{ diam } K_m \ge \text{diam } S_{I-}(K_m) \ge r_{j-1}.$$

Combining this with (3.2) and (3.3), we obtain

$$l_I \ge\ ll_{I-} \ge lL_{I-}/c_1 \ge lr_{j-1}/(c_1 \text{ diam } K_m).$$

The following two lemmas concerns the Ahlfors regularity of self-conformal sets. The proofs can be found in the recent works of Käenmäki, see [216, Theorem 3.1] and [225]. However, for the completeness, we include the proofs here (see [231]).

Corollary (5.3.26)[269]: [231]. Let $K_m \subset \mathbb{R}^d$ be a self-conformal set that has the BDP and positive d-dimensional Lebesgue measure. Then K_m is Ahlfors regular of dimension d.

Proof. Given an open ball $B_{K_m}(x_{j-1}, r_{j-1})$ in K_m with $2r_{j-1} < \text{diam } K_m$, note that $K_m \not\subset B_{K_m}(x_{j-1}, r_{j-1})$. Hence, we can find an index $I \in \Sigma^*$ such that

$$x_{j-1} \in S_I(K_m) \subset B_{K_m}\left(x_{j-1}, r_{j-1}\right) \text{ and } S_{I-}(K_m) \not\subset\ B_{K_m}(x_{j-1}, r_{j-1}).$$

By Corollary (5.3.25), we have $l_I \ge l_{r_{j-1}}/(c_1 \text{ diam } K_m)$. Consequently,

$$\begin{aligned} 2^d r_{j-1}^d = H^d\left(B(x_{j-1}, r_{j-1})\right) &\ge H^d\left(B_{K_m}(x_{j-1}, r_{j-1})\right) \\ &\ge H^d\left(S_I(K_m)\right) \ge l_I^d\ H^d(K_m) \ge \frac{H^d(K_m) l^d r_{j-1}^d}{(c_1 \text{ diam } K_m)^d}. \end{aligned} \tag{105}$$

Notice that $H^d(K_m) > 0$ since K_m has d-dimensional positive Lebesgue measure. We conclude that K is Ahlfors regular of dimension d.

Corollary (5.3.27)[269]: Let $K_m \subset \mathbb{R}^d$ be a self-conformal set that has the BDP and satisfies the WSC. Then K_m is Ahlfors regular of dimension α with $\alpha = \dim_H K_m$.

Proof. Let $x_{j-1} \in K_m$ and $0 < r_{j-1} < \text{diam } K_m/2$. We need to estimate the upper and lower bound of $H^\alpha\left(B_{K_m}(x_{j-1}, r_{j-1})\right)$.

For the upper bound. Taking $a = 1, U = B(x_{j-1}, r_{j-1})$ and $D = K_m$, gives

$$\#\left\{S \in A_{2r_{j-1}} : S(K) \cap B\left(x_{j-1}, r_{j-1}\right) \neq \emptyset\right\} \le \gamma_1, K_m < \infty$$

for all $x_{j-1} \in K_m$ and $0 < r_{j-1} < \text{diam } K_m$. It follows that

$$\begin{aligned} H^\alpha\left(B_{K_m}(x_{j-1}, r_{j-1})\right) &\le \sum_{\substack{S \in r_{j-1} \\ S(K_m) \cap B\left(x_{j-1}, r_{j-1}\right) = \emptyset}} H^\alpha S(K_m) \\ &\le \gamma_1, K_m\left(2r_{j-1}\right)^\alpha H^\alpha(K_m). \end{aligned}$$

For the lower bound, we use the same argument in the proof of Corollary (5.3.26) and obtain the same lower bound as in (105):

$$H^\alpha\left(B_{K_m}(x_{j-1}, r_{j-1})\right) \ge H^\alpha(K_m) l^\alpha r_{j-1}^\alpha/(c_1 \text{ diam } K_m)^\alpha.$$

Finally, we conclude the Ahlfors regularity of K_m from the upper and lower bound above since $0 < H^{\alpha}(K_m) < \infty$.

Corollary (5.3.28)[269]: [231]. Let $K_m \subset \mathbb{R}^d$ be a self-conformal set that has the BDP. If $\overset{\circ}{K}_m \neq \emptyset$ or K_m satisfies the WSC, then there exist $\left(x_{j-1}\right)_0 \in K_m$ and $0 < \left(r_{j-1}\right)_0 <$ diam K_m such that for all index $I \in \Sigma^*$,

$$S_I\left(\bar{B}_{K_m}(\left(x_{j-1}\right)_0, \left(r_{j-1}\right)_0)\right) = S_I\left(\bar{B}(\left(x_{j-1}\right)_0, \left(r_{j-1}\right)_0)\right) \cap K_m.$$

Proof. We begin with the case $\overset{\circ}{K}_m \neq \emptyset$. Clearly, there are $\left(x_{j-1}\right)_0 \in K_m$ and $0 < \left(r_{j-1}\right)_0 <$ diamK_m with $\bar{B}\left(\left(x_{j-1}\right)_0, \left(r_{j-1}\right)_0\right) \subset K_m$. Then for any $I \in \Sigma^*$, we have

$$\begin{aligned} S_I\left(\bar{B}_{K_m}\left(\left(x_{j-1}\right)_0, \left(r_{j-1}\right)_0\right)\right) &= S_I\left(\bar{B}\left(\left(x_{j-1}\right)_0, \left(r_{j-1}\right)_0\right) \cap K_m\right) \\ &= S_I\left(\bar{B}\left(\left(x_{j-1}\right)_0, \left(r_{j-1}\right)_0\right)\right) \\ &= S_I\left(\bar{B}\left(\left(x_{j-1}\right)_0, \left(r_{j-1}\right)_0\right)\right) \cap K_m, \end{aligned}$$

since $S_I\left(\bar{B}\left(\left(x_{j-1}\right)_0, \left(r_{j-1}\right)_0\right)\right) \subset K_m$.

Now suppose that K_m satisfies the WSC. Pick $x_{j-1} \in K_m$ and $0 < r_{j-1} <$ diam K_m with $\bar{B}(x_{j-1}, r_{j-1}) \subset X$.[1] Let $A_0 = \bar{B}(x_{j-1}, r_{j-1})$. We can obtain a sequence of sets $A_0, A_1, \ldots$ by induction as follows: suppose that A_{k-1} is determined, if

$$S_{I_k}(A_{k-1} \cap K_m) \subsetneq S_{I_k}(A_{k-1}) \cap K_m \quad \text{for some } I_k \in \Sigma^*,$$

let $A_k = S_{I_k}(A_{k-1})$; if such I_k does not exist, we stop the procedure.

Let $a = \left(2r_{j-1}\right)^{-1}$. We claim that the above induction procedure stops after at most γ_{a,K_m} steps, where γ_{a,K_m} is defined in (3.4). For otherwise, there would be $A_0, A_1, \ldots, A_n$ and $I_1, \ldots, I_n$ with $n > \gamma_{a,K_m}$ such that $A_0 = \bar{B}(x_{j-1}, r_{j-1}), A_k = S_{I_k}(A_{k-1})$ and

$$S_{I_k}(A_{k-1} \cap K_m) \subsetneq S_{I_k}(A_{k-1}) \cap K_m \quad \text{for } 1 \le k \le n. \tag{106}$$

To prove the claim, we need to deduce a contradiction. Note that $A_n = S_{I_n \ldots I_1}(A_0) = S_{I_n \ldots I_1}\left(\bar{B}(x_{j-1}, r_{j-1})\right)$. So

$$L_{I_n \ldots I_1} = \left(2r_{j-1}\right)^{-1} L_{I_n \ldots I_1} \cdot 2r_{j-1} \ge \text{adiam } A_n.$$

This together with (106) implies that for each $1 \le k \le n$, we can find $J_k \in \Sigma^*$ such that $S_{I_n} \ldots I_{k+1} J_k \in A_a$ diam A_n and

$$\emptyset \neq S_{I_k}(A_{k-1}) \cap S_{J_k}(K_m) \not\subset S_{I_k}(A_{k-1} \cap K_m). \tag{107}$$

Hence, $A_n \cap S_{I_n \ldots I_{k+1} J_k}(K_m) = S_{I_n} \ldots I_{k+1}\left(S_{I_k}(A_{k-1})\right) \cap S_{J_k}(K_m) \neq \emptyset$. Therefore,

$$S_{J_n}, S_{I_n} J_{n-1}, \ldots, S_{I_n \ldots I_2} J_1 \in \mathcal{A}_{a, A_n, K_m},$$

where $\mathcal{A}_{a,A_n,K_m}$ is defined by (3.4). Thus, $n > \gamma_{a,K_m}$ implies that there are $1 \le m_0 < k \le n$ such that $S_{I_n \ldots I_{k+1} J_k} = S_{I_n \ldots I_{m_0+1} J_{m_0}}$, i.e., $S_{J_k} = S_{I_k \ldots I_{m_0+1} J_{m_0}}$. Then we have

$$S_{I_k}(A_{k-1}) \cap S_{J_k}(K_m) = S_{I_k}\left(A_{k-1} \cap S_{I_{k-1} \ldots I_{m_0+1} J_{m_0}}(K_m)\right) \subset S_{I_k}(A_{k-1} \cap K_m).$$

This contradicts (107) and so the claim follows.

Clearly, the claim implies that we can find $A_n = S_{I_n \ldots I_1}\left(\bar{B}(x_{j-1}, r_{j-1})\right)$ satisfying

$$S_I(A_n \cap K_m) = S_I(A_n) \cap K_m \text{ for all } I \in \Sigma^*.$$

Let $\left(x_{j-1}\right)_0 = S_{I_n \ldots I_1}(x_{j-1})$ and $\left(r_{j-1}\right)_0 = l_{I_n \ldots I_1} r_{j-1}$, then $\bar{B}\left(\left(x_{j-1}\right)_0, \left(r_{j-1}\right)_0\right) \subset A_n$. For all $I \in \Sigma^*$,

$$S_I\left(\bar{B}K_m\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right)=S_I\left(\bar{B}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\cap K_m\right)$$
$$=S_I\left(\bar{B}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right)\cap S_I(K_m)$$
$$=S_I\left(\bar{B}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right)\cap S_I(A_n)\cap S_I(K_m)$$
$$=S_I\left(\bar{B}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right)\cap S_I(A_n)\cap K_m$$
$$=S_I\left(\bar{B}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right)\cap K_m.$$
Hence such $\left(x_{j-1}\right)_0$ and $\left(r_{j-1}\right)_0$ are desired.

Corollary (5.3.29)[269]: Let $K_m\subset\mathbb{R}^d$ be a self-conformal set that has the bounded distortion property (BDP). Then
$$K_m^\circ\neq\emptyset\Leftrightarrow K_m=\overline{K_m^\circ}\ \Leftrightarrow K_m \text{ is a BBI set of dimension } d.$$
Corollary (5.3.30)[269]: Let $K_m\subset\mathbb{R}^d$ be a self-conformal set that has the BDP and satisfies the WSC. If $\dim_H K_m=d$, then $K_m^\circ\neq\emptyset$.

Proof of Corollary (5.3.29) and (5.3.30). (See [231]). We remark that, in Corollary (5.3.29), $K_m=\overline{K_m^\circ}$ follows from $K_m^\circ\neq\emptyset$ immediately. Indeed, K_m° contains $\bigcup_{I\in\Sigma^*}S_I(K_m^\circ)$ and the latter is dense in K_m.

Therefore, by Corollary (5.3.24), it suffices to prove that, if $K_m^\circ=\emptyset$ (for Corollary (5.3.29)) or K_m satisfies the WSC with $\dim_H K_m=d$ (for Corollary (5.3.30)), then K_m is a BBI set of dimension d.

We first observe that, by Corollary (5.3.26) and (5.3.27), K_m is Ahlfors regular of dimension d in both cases. Now let $B_{K_m}(x_j,r_j)$ and $B_{K_m}(x_{j+1},r_{j+1})$ be two open balls in K_m with $0<r_j,r_{j+1}<\operatorname{diam}K_m$. To complete the proof, we need to find constants $(1+2\epsilon),(1+\epsilon)$ (independent of x_j,x_{j+1} and r_j,r_{j+1}), a point $x_{j+2}\in B_{K_m}(x_j,r_j)$ and a $(1+\epsilon)$-conformally bi-Lipschitz mapping h_m from $B_{K_m}(x_{j+2},(1+2\epsilon)r_j)$ into $B_{K_m}(x_{j+1},r_{j+1})$ with scale factor r_{j+1}/r_j.

For this, suppose without loss of generality that $K_m\not\subset B_{K_m}(x_j,r_j)$ and $K_m\subset B_{K_m}(x_{j+1},r_{j+1})$.

Thus, we can find two indices $I_1,I_2\in\Sigma^*$ such that
$$S_{I_1}(K_m)\subset B_{K_m}\left(x_j,r_j\right),\qquad S_{I_1^-}(K_m)\not\subset B_{K_m}(x_j,r_j)$$
And
$$S_{I_2}(K_m)\subset B_{K_m}\left(x_{j+1},r_{j+1}\right)\qquad S_{I_2^-}(K_m)\not\subset B_{K_m}(x_{j+1},r_{j+1}).$$
It follows from Corollary (5.3.25) that
$$\frac{l}{c_1\operatorname{diam}K_m}\le\frac{l_{I_i}}{r_i}\le\frac{2}{\operatorname{diam}K_m},\quad\text{for } i=1\ ,2. \tag{108}$$
By Corollary (5.3.28), in both cases, we can find $\left(x_{j-1}\right)_0\in K_m$ and $0<\left(r_{j-1}\right)_0<\operatorname{diam}K_m$ such that for all index $I\in\Sigma^*$,
$$S_I\left(\bar{B}_{K_m}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right)=S_I\left(\bar{B}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right)\cap K_m. \tag{109}$$
Let $x_{j+2}=S_{I_1}\left(\left(x_{j-1}\right)_0\right)\in S_{I_1}(K_m)\subset B_{K_m}\left(x_j,r_j\right)$ and $(1+2\epsilon)=l\left(r_{j-1}\right)_0/(c_1\operatorname{diam}K_m)$. Then (108) gives $(1+2\epsilon)r_j\le l_{I_1}\left(r_{j-1}\right)_0$. This together with (3.2) and (109) implies

$$\bar{B}_K(x_{j+2},(1+2\epsilon)r_j) \subset \bar{B}_{K_m}\left(x_{j+2}, l_{I_1}\left(r_{j-1}\right)_0\right)$$
$$\subset S_{I_1}\left(\bar{B}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right) \cap K_m$$
$$= S_{I_1}\left(\bar{B}_{K_m}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right). \qquad (110)$$

Therefore, we have $\bar{B}_{K_m}(x_{j+2},(1+2\epsilon)r_j) \subset B_{K_m}(x_j,r_j)$ since

$$S_{I_1}\left(\bar{B}_{K_m}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right) \subset S_{I_1}(K_m) \subset B_{K_m}(x_j,r_j).$$

Let $h_m = S_{I_2} \circ S_{I_1}^{-1}$. By (110), we have

$$h_m\left(\bar{B}_{K_m}(x_{j+2},(1+2\epsilon)r_j)\right) \subset h_m\left(S_{I_1}\left(\bar{B}_{K_m}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right)\right)$$
$$= S_{I_2}\left(\bar{B}_{K_m}\left(\left(x_{j-1}\right)_0,\left(r_{j-1}\right)_0\right)\right) \subset B_{K_m}(x_{j+1},r_{j+1}).$$

Hence h_m maps $\left(\bar{B}_{K_m}(x_{j+2},(1+2\epsilon)r_j)\right)$ into $B_{K_m}(x_{j+1},r_{j+1})$.

It remains to show that $h_m = S_{I_2} \circ S_{I_1}^{-1}$ is a $(1+\epsilon)$-conformally bi-Lipschitz mapping with scale factor r_{j+1}/r_j for a constant $(1+\epsilon)$. Indeed, we have

$$\frac{l_{I_2}}{c_1 l_{I_1}} \le \frac{l_{I_2}}{L_{I_1}} \le \frac{\left|S_{I_2} \circ S_{I_1}^{-1}\left(x_{j-1}\right) - S_{I_2} \circ S_{I_1}^{-1}\left(x_{j-1}\right)\right|}{\left|\, x_{j-1} - y\right|} \le \frac{l_{I_2}}{L_{I_1}} \le \frac{l_{I_2}}{c_1 l_{I_1}},$$

for $x_{j-1}, y \in \bar{B}_{K_m}(x_{j+2},(1+2\epsilon)r_j)$. This together with (108) gives

$$(1+\epsilon)^{-1}(r_{j+1}/r_j)|x_{j-1}-y| \le |h_m(x_{j-1}) - h_m(y)|$$
$$\le (1+\epsilon)\left(r_{j+1}/r_j\right)\left|x_{j-1}-y\right|,$$

where $(1+\epsilon) = 2c_1^2/l$.

Chapter 6
M-Cantorvals and Recovering

We show that a new sufficient condition for the set of subsums of a series to be a Cantor set is formulated and it is used to demonstrate that the discussed multigeometric series always have Cantor sets as their sets of subsums for sufficiently small ratios of the series We are interested in the two following questions. Which set can be a range of some measure μ? Can the purely atomic measure μ be uniquely recovered from its range. We shidy The Lebesgue measure of K is computed and it is shown to be equal to the sum of lengths of all component intervals of the M-Cantorval.

Section (6.1): Ferens Type

The investigation of topological properties of sets of subsums of absolutely convergent series has been initiated almost one hundred years ago by Soichi Kakeya [150], [248]. Most of his findings were rediscovered later and published in more accessible journals [148], [249]. He thought that a set of subsums must be (up to a homeomorphism) either a finite set or the union of afinite family of closed intervals or a Cantor set. The first example of a series with a set of subsums of neither of the three mentioned types was stated without proof by A. D. Weinstein and B. E. Shapiro in a note published in 1980 [252]. Four years later C. Ferens presented a complete example of another series having an M-Cantorval as the set of its subsums [146]. A major step in the research took place in 1988 when J.A. Guthrie and J.E. Nymann published the full topological classication of the sets of subsums [147] (see also [155]). It consists of exactly four topological types with M-Cantorvals being the fourth type unknown to Kakeya. The result describes also all possible ranges of purely atomic probabilistic measures correcting earlier results in the direction [151], [247]. Finding a complete analytic characterization of when a given series has a Cantor set and when an M-Cantorval as the set of its subsums remains a challenging problem [147], [149], [250]. A new exposition of the Guthrie Nymann Classication Theorem, based on the Mendes-Oliveira characterization of M-Cantorvals [152], can be found in [251]. Some algebraic and topological aspects of sets of series related to the Guthrie-Nymann Classication Theorem were investigated in [144] recently.

Since the topological type of the set of subsums of a given absolutely convergent series is the same as the type of the series of absolute values of its terms, we will restrict our attention to convergent series $\sum a_n$ of positive terms. Further, since the set of subsums does not depend on the order of summation, we may assume that the terms of the series decrease. The set of subsums of the series is defined to be

$$E(a_n):=\left\{x \in \mathbb{R}: \text{there exists } A \subset \mathbb{N}, x=\sum_{n\in A} a_n\right\}=\sum_{n=1}^{\infty} e_n a_n : e_n \in \{0,1\}.$$

We agree to write $\sum_{n\in\emptyset} a_n = 0$. The symbol r_k denotes the k-th remainder of the series understood sometimes as the series $\sum_{n=k+1}^{\infty} a_n$ an and sometimes as the sum of the latter series which will be clear from the context always.

The symbol E_k denotes the set of subsums of the k-th remainder of the series$\sum a_n$, that is, $E_k = (E(a_n)_{n=k+1}^{\infty})$. In particular, $E_0 = E(a_n)$. Clearly, E_k is a subset of the interval $[0, r_k]$ for any $k \in \mathbb{N}_0$. The set of all k-initial subsums of $\sum a_n$ will be denoted by

$$F_k:=\left\{\sum_{n=1}^{k} e_n a_n : \text{for all } n \in \{1,\dots,k\}\ e_n \in \{0,1\}\right\}$$

We define $F_0 := \{0\}$ additionally.
It is easy to observe that $E = F_k + E_k$ for any non-negative integer k. The set of all sums of finite subseries of $\sum a_n$ will be denoted by F, that is, $F = \bigcup_{k\in\mathbb{N}} F_k$.
Fact (6.1.1)[246]: For any $k \in \mathbb{N}$, the following equalities hold

$$E_{k-1} = E_k \cup (a_k + E_k) \text{ and } E = \bigcup_{f\in F_k} (f + E_k) \ .$$

Moreover, $E = \bar{F}$.
The first two equalities are elementary [155], the third one follows from the fact that the set E is closed [150], [248]. In particular, the second equality tells us that the set E is a union of finitely many translates of the set of subsums of the k-th remainder.
The sets

$$I_k := \bigcup_{f\in F_k} (f + [0, r_k])$$

will be called k-th iterates of the set E. When we look at the series $\sum \frac{2}{3n}$ having the classic Cantor ternary set as the set of its subsums, I_k is exactly the set obtained in the k-th step of the standard construction of the classic Cantor set. Thus, the following fact is a generalization of the classical construction.
Fact (6.1.2)[246]:

$$E = \bigcap_{k=1}^{\infty} I_k \text{ for any series } \sum a_n \, .$$

Proof. It follows from the second equality of the Fact (6.1.2) that

$$E = \bigcup_{f\in F_k} (f + E_k) \subset \bigcup_{f\in F_k} (f + [0, r_k]) = I_k$$

for any k. Thus, $E \subset \bigcap_k I_k$.
If $x \in I_k$, then, by the definiton of I_k, there is an $f \in F_k$ such that the distance $d(x, f) \le r_k$. Hence $d(x, F_k) \le r_k$ which implies that $d(x, E) \le r_k$. Thus, if $x \in \bigcap_k I_k$, then $d(x, E) = 0$ and thus $x \in E$, since the last set is closed.
Fact (6.1.3)[246]: Let m, k be positive integers such that $k \ge m + 1$. Then the set of all possible sums of finitely many distinct summands taken from among $m, m + 1, \dots, m + k - 1$ (we understand that 0 is the sum of the empty subcollection) is equal to

$$\{0\} \cup \{m, m + 1, \dots, s - m - 1, s - m\} \cup \{s\}$$

Where $s = s(m, k) := m + (m + 1) + \cdots + (m + k - 1)$. Moreover,

$$s \ge \frac{3}{2} m(m + 1).$$

Proof. It is obvious for $k = 2$, since then it must be $m = 1$.
Assume now that k and m are positive integers such that $k \ge 3$ and $k \ge m + 1$. Given an $n \in \{0, 1, \dots, k\}$, let S_n denotes the set of all sums of exactly n distinct summands from among $m, m + 1, \dots, m + k - 1$, that is

$$S_n := \left\{ \sum_{i=1}^{n} b_i : b_i \in \{m, m + 1, \dots, m + k - 1\} \text{and } (i \ne j \Rightarrow b_i \ne b_j) \right\}.$$

We assume $S_0 = \{0\}$ as usually. Then S_n is a finite set of consecutive positive integers for any n with $1 \le n \le k - 1$. Moreover,

$\min S_n = nm + \frac{n}{2}(n-1)$ and $\max S_n = n(m+k) - \frac{n+1}{2}n$ and both finite sequences $(\min S_n)_{n=1}^{k-1}$ and $(\max S_n)_{n=1}^{k-1}$ are increasing.

We will show that there is no gap between any two consecutive sets S_n, that is, given an $l \in \mathbb{N}$ such that $2 \le l \le k-1$, there is no positive integer y satisfying the double inequality $\max S_{l-1} < y < \min S_l$. It sufces to show that $\min S_{l-1} \max S_{l-1}$ which, in turn, is equivalent to

$$\frac{m-1}{l-1} + l \le k.$$

The last inequality follows from the assumption $k \ge m+1$ and from the fact that the function $f(x) := \frac{k-2}{x-1} + x$ does not exceed the value k on the interval $x \in [2, k-1]$.

Finally,

$$\bigcup_{i=0}^{k} S_i = S_0 \cup \bigcup_{i=1}^{k-1} S_i \cup S_k = \{0\} \cup \{m, m+1, \dots, s-m-1, s-m\} \cup \{s\}$$

The series given in [252] by Weinstein and Shapiro as an example escaping the Kakeya hypothesis can be written in the form

$$8.\frac{1}{10} + 7.\frac{1}{10} + 6.\frac{1}{10} + 5.\frac{1}{10} + 4.\frac{1}{10} + 8.\frac{1}{10^2} + 7.\frac{1}{10^2} + 6.\frac{1}{10^2} + 5.\frac{1}{10^2} + 4.\frac{1}{10} + 8.\frac{1}{10^3} + \cdots$$

We have dropped a constant factor of $\frac{3}{10}$they used to make the sum of the whole series to be 1.

We dropped the factor to make the structure of the series more visible, since constant non-zero multipliers have no influence on the topological type of the set of subsums. Weinstein and Shapiro wrote that it is easy to see that the whole interval $\left[\frac{4}{9}, \frac{26}{9}\right]$ belongs to the set E of subsums of the series.

The series discussed by Ferens in [146] is similar:

$$7.\frac{2}{27} + 6.\frac{2}{27} + 5.\frac{2}{27} + 4.\frac{2}{27} + 3.\frac{2}{27} + 7.\left(\frac{2}{27}\right)^2 + 6.\left(\frac{2}{27}\right)^2 + 5.\left(\frac{2}{27}\right)^2 + 4.\left(\frac{2}{27}\right)^2 + 3.\left(\frac{2}{27}\right)^2 + 7.\left(\frac{2}{27}\right)^2 + \cdots$$

We have dropped a constant factor of $\frac{1}{2}$ Ferens used to obtain a series suitable for the natural construction of a purely atomic probabilistic measure. Ferens proved that the whole interval$\left[\frac{6}{25}; \frac{44}{25}\right]$ is contained in the set of subsums of his series.

Definition (6.1.4)[246]: Let m, k be positive integers such that $k \ge m+1$. A series $\sum a_j$ will be said to be of Ferens type if

$$a_j = (m+k-i)q^n \text{ for all } j \in \mathbb{N},$$

where q is a real number from $(0,1)$ and (n, i) is the unique pair of positive integers with range $n \in \mathbb{N}, i \in \{1, 2, \dots, k\}$ such that $j = k(n-1) + i$. We will denote the series by $\mathcal{F}(m, k; q)$. It is a special multigeometric series [250]:

$$\mathcal{F}(m,k;q) = (m+k-1)q + (m+k-2)q + \cdots + mq + (m+k-1)q^2 + (m+k-2)q^2 + \cdots + mq^2 + (m+k-1)q^3 + (m+k-2)q^3 + \cdots + mq^3 + \cdots$$

The terms a Ferens type series are strictly decreasing if and only if $> (m+k-1)q$. Moreover, it is easy to find a closed formula for the kn-th remainder of the series:

$$r_{kn} = \sum_{j=kn+1}^{\infty} a_j = s\frac{q^{n+1}}{1-q}$$

where $s = s(m,k)$ has been de fined in the statement of the Fact (6.1.3).

Lemma (6.1.5)[246]:

$$F_{kn} = \left\{\sum_{i=1}^{n} e_i q_i : e_i \in \{0\} \cup \{m, m+1, \dots, s-m\} \cup \{s\}\right\}$$

for any series $\mathcal{F}(m,k;q)$ of Ferens type.

It follows from the Fact (6.1.3) by induction easily.

We will investigate a special subset of the set F_{kn}:

$$K_n := \left\{\sum_{i=1}^{n} e_i q^i : e_i \in \{m, m+1, \dots, s-m\}\right\}$$

with the natural order induced from the real line.

Lemma (6.1.6)[246]: If $q \geq \frac{1}{s-2m+1}$, then the distance between any two consecutive points of K_n does not exceed q^n.

Proof. The proof runs by induction on n. Since $K_1 = \{mq, (m+1)q, \dots, (s-m)q\}$, the distance between any two consecutive points of K_1 is always q, and hence the thesis holds for $n = 1$.

Assume now that the thesis holds for a positive integer n. Let $h < f$ be any two consecutive points of $K_n + 1$. Clearly,

$$f = \sum_{i=1}^{n+1} f_i q^i \qquad \text{and} \qquad h = \sum_{i=1}^{n+1} h_i q^i$$

for a suitable choice of $f_i, h_i \in \{m, m+1, \dots, s-m\}. If\ f_n + 1 > m$, then $f - q^{n+1} \in K_{n+1}$. Thus $f - q^{n+1} \leq h < f$, and hence the distance from h to f does not exceed q^{n+1}.

If $h_{n+1} < s-m$, then $h + q^{n+1} \in K_{n+1}$. Thus $h < f \leq h + q^{n+1}$, and hence the distance from h to f does not exceed q^{n+1}.

It remains to consider the case $f_{n+1} = m$ and $h_{n+1} = s-m$. Then the following numbers

$$\bar{f} := \sum_{i=1}^{n} f_i q^i \qquad \text{and} \qquad \bar{h} =: \sum_{i=1}^{n+1} h_i q^i$$

belong to K_n. Since $s - m > m$, it must be $\bar{h} < \bar{f}$. Define $\bar{g} := \max\{x \in K_n : x < \bar{f}\}$. We get $\bar{f} - \bar{g} \leq q^n$ by our inductional assumption.

Clearly, $\bar{g} + mq^{n+1} \in K_{n+1}$ and $\bar{g} + mq^{n+1} < f$.

Our assumption that $q \geq \frac{1}{s-2m+1}$ implies that

$$\bar{g} + (s-m)q^{n+1} \geq \bar{g} + q^n + mq^{n+1} - q^{n+1},$$

and therefore

$$\bar{g} + (s-m)q^{n+1} \geq \bar{f} + mq^{n+1} - q^{n+1} = f - q^{n+1}.$$

Thus $f \in (\bar{g} + mq^{n+1}, \bar{g} + (s-m)q^{n+1} + q^{n+1}]$, and hence we find $k \in \{m, m+1, \dots, s-m\}$ such that $0 < f - (\bar{g} + kq^{n+1}) \leq q^{n+1}$.Since $\bar{g} + kq^{n+1} \in K_{n+1}$, we have that $\bar{g} + kq^{n+1} \le h < f$, and hence the distance from h to f does not exceed q^{n+1} which completes the proof.

Lemma (6.1.7)[246]: Let $\mathcal{F}(m,k,;q)$ be a Ferens type series satisfying the condition $\frac{1}{s-2m+1} \le q < \frac{m}{s+m}$. Then

$$P_{n-1} \cap I_{kn} = L_n^1 \sqcup P_n \sqcup L_n^2$$

for every $n \in \mathbb{N}$ where

$$L_n^1 := \left[m\sum_{i=1}^{n-1} q^i, m\sum_{i=1}^{n-1} q^i + r_{kn}\right] = m\sum_{i=1}^{n-1} q^i + [0, r_{kn}],$$

$$P_n := \left[m\sum_{i=1}^{n} q^i, (s-m)\sum_{i=1}^{n-1} q^i + r_{kn}\right] \text{for } n \ge 1,$$

$$P_0 := I_0 = \left[0, \frac{sq}{1-q}\right],$$

$$L_n^2 := (s-m)\sum_{i=1}^{n-1} q^i + sq^n, (s-m)\sum_{i=1}^{n-1} q^i + sq^n + r_{kn}$$

$$= (s-m)\sum_{i=1}^{n-1} q^i + sq^n + [0, r_{kn}]$$

Proof. The sequence $(P_n)_{n=1}^{\infty}$ of closed intervals, where

$$P_n := \left[m\sum_{i=1}^{n} q^i, (s-m)\sum_{i=1}^{n} q^i + r_{kn}\right],$$

is strictly descending in the sense that $P_{n+1} \subset \operatorname{int} P_n$ It follows from the inequalities

$$m\sum_{i=1}^{n} q^i < m\sum_{i=1}^{n+1} q^i \text{ and } (s-m)q^{n+1} + r_{k(n+1)} < r_{kn}.$$

The first one is obvious and the second one is equivalent to $m > 0$.

Observe now that $P_n \subset I_{kn}$ for all $n \in \mathbb{N}$. Fix an n. Then $K_n \subset F_{kn}$ by the Lemma (6.1.5). The assumption $q \ge \frac{1}{s-2m+1}$ implies that $q \ge \frac{1}{s+1}$ and the last inequality is equivalent to $r_{kn} \ge q^n$.

Thus, since the distance of any two consecutive points of K_n is at most qn by the Lemma (6.1.6), the union $\bigcup_{f\in K_n}[f, f + r_{kn}]$ is an interval.

Since $\min K_n = m\sum_{i=1}^{n} q^i$ and $\max K_n = (s-m)\sum_{i=1}^{n} q^i$, we conclude that

$$P_n = \bigcup_{f\in K_n} [f, f + r_{kn}] \subset \bigcup_{f\in K_n} [f, f + r_{kn}] = I_{kn}.$$

It follows from the definitions of the iteration I_{kn} directly that both intervals L_n^1 and L_n^2 are contained in I_{kn}.

We are now going to show that the open intervals $(\max L_n^1; \min P_n)$ and $(\max P_n, \min L_n^2)$ are I_{kn}-gaps, but we need three auxiliary inequalities first. Since $s > \frac{3}{2}m(m+1) > m^2$, it follows that $s + m > m(m+1)$. Thus our assumption $q < \frac{m}{s+m}$ implies that

$$q < \frac{1}{m+1}. \tag{1}$$

Next, $\frac{m}{s} > \frac{m}{s+m} > q$ and hence $sq < m$ which implies that

$$sq^{p+1} < mq^{p} \qquad \text{for any} \qquad p \in \mathbb{N}_0. \tag{2}$$

Our assumption $q < \frac{m}{s+m}$ is equivalent to

$$s\frac{q^{n+1}}{1-q} = r_{kn} < mq^{n} \qquad \text{for any} \quad n \in \mathbb{N}_0. \tag{3}$$

Consider an $f \in F_{kn}$. By the Lemma (6.1.5), f is of the form

$$f = \sum_{i=1}^{n} f_i q^i \text{ with } f_i \in \{0\} \cup \{m, m+1, \dots, s-m\} \cup \{s\}.$$

If the set $J := \{j \in \{1,2,\dots,n\}: f_j \neq m\}$ is empty, then $f = m\sum_{i=1}^{n} q^i$. If $J \neq \emptyset$, define $j_0 := \min J$.

Then either $f_{j_0} > m \text{ or } f_{j_0} < m$.

If $f_{j_0} > m$, then $f_{j_0} \geq m+1$ and

$$f \geq m\sum_{i=1}^{j_0-1} q^i + (m+1)q^{j_0} = m\sum_{i=1}^{j_0} q^i + q^{j_0} > m\sum_{i=1}^{j_0} q^i + m\frac{q^{j_0+1}}{1-q}$$
$$= m\sum_{i=1}^{\infty} q^i > m\sum_{i=1}^{n} q^i.$$

If $f_{j_0} < m$, then $f_{j_0} = 0$ and $f_i \leq s$ for $i > j_0$ Hence

$$f \geq m\sum_{i=1}^{j_0-1} q^i + \sum_{i=j_0+1}^{n} sq^i = m\sum_{i=1}^{j_0-1} q^i + \sum_{i=j_0 1}^{n-1} sq^{i+1} \leq$$
$$m\sum_{i=1}^{j_0-1} q^i + m\sum_{i=j_0 1}^{n-1} q^i = m\sum_{i=1}^{n-1} q^i.$$

Therefore,

$$\max\left\{f \in F_{kn}: f < m\sum_{i=1}^{n} q^i\right\} = m\sum_{i=1}^{j_0-1} q^i,$$

and thus

$$\max\left\{f \in r_{kn}: f \in F_{kn}, f < m\sum_{i=1}^{n} q^i\right\} = m\sum_{i=1}^{n-1} q^i$$
$$+ s\frac{q^{n+1}}{1-q} < m\sum_{i=1}^{n} q^i,$$

which means that the open interval $m\sum_{i=1}^{n-1} q^i + r_{kn}, m\sum_{i=1}^{n} q^i$is an I_{kn}-gap. The proof that the intervals $(\max P_n, \min L_n^2)$ are I_{kn}-gaps is fully analogous.

Lemma (6.1.8)[246]: Let $\mathcal{F}(m,k;q)$ be a Ferens type series satisfying the condition

$$\frac{1}{s-2m+1} q < \frac{m}{s+m}.$$

Then the interval $[m\,\frac{q}{1-q}, (s-m)\frac{q}{1-q}\,]$ is a component interval of $E(\mathcal{F}(m,k;q))$.

Proof. The following inclusion holds by the previous lemma

$$\bigcap_n P_n = \left[m\frac{q}{1-q}, (s-m)\frac{q}{1-q}\right] \subset \bigcap_n I_{kn} = E.$$

Since I_{kn}-gaps are E-gaps and since $\max L_n^1 \longrightarrow m\frac{q}{1-q}$ and $\min L_n^2 \to (s-m)\ \frac{q}{1-q}$ as $n \to \infty$, the endpoints of the interval $\left[m\ \frac{q}{1-q}, (s-m)\ \frac{q}{1-q}\right]$ are limits of E-gaps. Thus the interval is a component interval of E.

Theorem (6.1.9)[246]: Let $\mathcal{F}(m,k;q)$ be a Ferens type series satisfying the condition

$$\frac{1}{s-2m+1}q < \frac{m}{s+m}.$$

Then the set of its subsums $E = E(\mathcal{F}(m,k;q))$ is an M-Cantorval. Its Lebesgue measure is $\mu E = (s-2m)\frac{q}{1-3q}$ and it is equal to the sum of lengths of all component intervals of E.

Proof. The inequality $q < \frac{m}{s+m}$ is equivalent to $r_{kn} < mq^n$ for all n, that is, to the inequalities $a_{kn} > r_{kn}$ which means that E has in nitely many gaps. Hence, by the Guthrie-Nymann Classication Theorem [147], [251], E must be either an M-Cantorval or a Cantor set. The Lemma (6.1.8) eliminates the second possibility.

We need good insight into the geometric structure of consecutive iterations I_{kn}, $n \in \mathbb{N}_0$, in order to compute the Lebesgue measure of E.

Observation (6.1.10)[246]:

$$I_{kn} = \bigcup_{i=1}^{2}\bigcup_{j=1}^{n-1}\left(L_j^i\ \cap I_{kn}\right) \sqcup (L_n^1\ \sqcup P_n \sqcup L_n^2)$$

for all $n \in \mathbb{N}$.

A simple proof of this observation is based on the Lemma (6.1.7) and runs by induction on n.

Observation (6.1.11)[246]: Given $j \in \{1,2,\dots,n-1\}, n \geq 2, i \in \{1,2\}$, the following equality holds

$$L_j^i I_{kn} = \varepsilon_j^i + q^j I_{k(n-j)}$$

Where

$$\varepsilon_j^i := \begin{cases} m\sum_{i=1}^{j-1} q^i & \text{if } i = 1, \\ (s-m)\sum_{i=1}^{l-1} q^i + sq^j & \text{if } i = 2. \end{cases}$$

We are going to prove the Observation B in the case $i = 1$, since the other case is quite similar.

Given an $x \in I_{kp}$, a number $f \in F_{kp}$ such that $x \in f + [0, r_{kp}]$ will be called kpoor of x and denoted by bxckp. It does not have to be unique. Observe that for any $p \in \mathbb{N}$

$$L_p^1 = \left\{x \in I_{kp} : [x]k_p = m\sum_{l=1}^{p-1} q^l\right\}.$$

In particular, the kp-floor of $x \in L_p^1$ is unique. Take $h \in I_{kn} \cap L_j^1$ now. There exists, by the definition of I_{kn}, an element $\sum_{l=1}^{n}\ e_l q^l \in F_{kn}$ such that $h\ = \sum_{l=1}^{n}\ e_l q^l\ +\ r$ for some $r \in [0, r_{kn}]$. Then

$$h=\sum_{l=1}^{j} e_l q^l+\sum_{l=j+1}^{n} e_l q^l+r.$$

Since $\sum_{l=j+1}^{n} e_l q^l+r\in[0,r_{kn}]$, the fact that $h\in L_j^1$ and uniqueness of kj-floor on L_j^1 imply that $e_1=\cdots=e_{j-1}=m$ and $e_j=0$, that is,

$$h=m\sum_{l=1}^{j-1} q^l+\sum_{l=j+1}^{n} e_l q^l+r=m\sum_{l=1}^{j-1} q^l+q^j\left(\sum_{l=1}^{n-j} e_l+q^l+\bar{r}\right)$$

for some $\bar{r}\in[0,r_{k(n-j)}]$. Thus $h\in m\sum_{l=1}^{j-1} q^l+q^j I_{k(n-j)}$. The inverse inclusion $\varepsilon_j^1+q^j I_{k(n-j)}\subset L_j^1\cap I_{kn}$ is straightforward and therefore the proof of the Observation B is complete.

The next observation is crucial for understanding the building of consecutive iterations I_{kn}, $n\in\mathbb{N}_0$, and for computing the Lebesgue measure of E.

Observation (6.1.12)[246]: If $[\alpha,\beta]$ is a component interval of I_{kn}, then

$$[\alpha,\beta]\cap I_{k(n+1)}=[\alpha,\alpha+r_{k(n+1)}]\sqcup[\alpha+mq^{n+1},\beta-mq^{n+1}]\sqcup[\beta-r_{k(n+1)}]$$

The proof of the observation runs by induction on n, so we start with $n=1$. Let $[\alpha,\beta]$ be a component interval of I_k. Since $I_k=L_1^1\sqcup P_1\sqcup L_1^2$by the Lemma (6.1.7), exactly one of the following three cases holds:

$(a1)[\alpha,\beta]=L_1^1=[0,r_k]$,

$(a2)[\alpha,\beta]=L_1^2=[sq,r_0]$,

$(a3)[\alpha,\beta]=P_1=[mq,(s-m)q+r_k]$.

In the case $(a1)$, since $qr_k=r_{k2},\alpha=0,\beta=r_k=sq^2+r_{k2}$, we get

$$=[\alpha,\beta]\cap I_{k2}=L_1^1\cap I_{k2}\underline{Obs.B}\ =\ qI_k=qL_1^1\sqcup qP_1\sqcup qL_1^2$$
$$=[0,r_{k2}]\sqcup[mq^2,(s-m)q^2+r_{k2}]\sqcup[sq^2,r_k]$$
$$=\ [0,r_{k2}]\sqcup[mq^2,sq^2+r_{k2}-mq^2]\sqcup[r_k-r_{k2},r_k]$$
$$=[\alpha,\alpha+r_{k2}]\sqcup[\alpha+mq^2,\beta-mq^2]\sqcup[\beta-r_{k2},\beta].$$

The cases (a2) and (a3) are very similar. We omit the details for them and it completes the step $n=1$.

Now, let $n\geq 2$ be a positive integer such that the thesis holds for $t=1,2,\dots,n-1$, that is, for any component interval$[\bar{\alpha},\bar{\beta}]$ of I_{kt}, where $1\leq t\leq n-1$, we have

$$\left[\bar{\alpha},\bar{\beta}\right]\cap I_{k(t+1)}=\left[\bar{\alpha},\bar{\alpha}+r_{k(t+1)}\right]\sqcup\left[\bar{\alpha}+mq^{t+1},\bar{\beta}-mq^{t+1}\right]\sqcup\left[\ \bar{\beta}-r_{k(t+1)},\bar{\beta}\right].$$

Let $[\alpha,\beta]$ be a component interval of I_{kn}. The Lemma (6.1.7) combined with Observations A and C yields

$$I_{kn}=\coprod_{i=1}^{2}\coprod_{d=1}^{n-1}(\varepsilon_d^i+q^d I_{k(n-d)})\sqcup(L_n^1\sqcup P_n\sqcup L_n^2).$$

Hence exactly one of the following four cases holds:

$$(b1)[\alpha,\beta]=L_n^1=m\sum_{l=1}^{n-1}+[0,r_{kn}],$$

$$(b2)[\alpha,\beta]=L_n^2=(s-m)\sum_{l=1}^{n-1} q^l+sq^n+[0,r_{kn}].$$

$$(b3)[\alpha,\beta] = P_n = \left[m\sum_{l=1}^{n} q^l(s-m), \sum_{l=1}^{n} q^l + r_{kn}\right]$$

$(b4)$ $[\alpha,\beta] = \varepsilon_d^i + q^d[\bar{\alpha},\bar{\beta}]$ for some $i \in \{1,2\}, d \in \{1,\dots,n-1\}$ and a component interval $[\bar{\alpha},\bar{\beta}]$ of $I_{k(n-d)}$. In the case $(b1)$ we get

$$[\alpha,\beta] \cap I_{k(n+1)} = L_n^1 \cap I_{k(n+1)} \underline{Obs.B} = m\sum_{l=1}^{n-1} q^l + q^n I_k$$

$$\text{Lemma (6.1.7)} = m\sum_{l}^{n-1} q^l + \sum_{l=1}^{n} q^l + q^n$$

$$= ([0,r_k]\sqcup[mq,(s-m)q + r_k]\sqcup[sq, sq + r_k])$$

$$= m\sum_{l=1}^{n-1} q^l + \left([0, r_{k(n+1)}]\sqcup\left[mq^{n+1},(s-m)q^{n+1} + r_{k(n+1)}\right]\sqcup\left[sq^{n+1}, sq^{n+1} + r_{k(n+1)}\right]\right).$$

(now we apply the identity $\sum_{l=1}^{n-1} q^l + sq^{n+1} r_{k(n+1)} = m\sum_{l=1}^{n-1} q^l + r_{kn} \;\; = \beta)\left[\alpha, \alpha + r_{k(n+1)}\right]\sqcup\left[\alpha + q^{n+1}, \beta - mq^{n+1}\right]\sqcup\left[\beta - r_{k(n+1)}, \beta\right]$ Computations in the cases (b2) and (b3) are quite similar and we will leave them out. The case (b4) is relatively the most complicated.

If $[\alpha,\beta]$ is of the form (b4), then $[\alpha,\beta] \subset I_{kn} \cap L_d^i$, and hence

$$[\alpha,\beta] \cap I_{k(n+1)} \overset{Obs.A}{=} [\alpha,\beta] \cap \left(I_{k(n+1)} \cap L_d^i\right)$$

$$\overset{Obs.B}{=} \left(\varepsilon_d^i + q^d[\bar{\alpha},\bar{\beta}]\right) \cap \left(\varepsilon_d^i + q^d I_{k(n+1-d)}\right) = \varepsilon_d^i + q^d\left([\bar{\alpha},\bar{\beta}]\right) \cap I_{k(n-d+1)}$$

(and now we use the inductional assumption)

$$= \varepsilon_d^i + q^d\left(\left[\bar{\alpha}, \bar{\alpha} + r_{k(n-d+1)}\right]\sqcup\left[\bar{\alpha} + mq^{n-d+1}, \bar{\beta} + mq^{n-d+1}\right]\sqcup\left[\bar{\beta} - r_{k(n-d+1)}, \bar{\beta}\right]\right)$$

$$= \left[\alpha, \alpha + r_{k(n+1)}\right]\sqcup\left[\alpha + mq^{n+1}, \beta + mq^{n+1}\right]\sqcup\left[\beta - r_{k(n+1)}, \beta\right]$$

which completes the inductional step and hence completes the proof of the Observation C.

We are now ready to describe the geometric building of iterations I_{kn}. The first of them $I_0 = I_{k0}$consist of the single interval $0, \frac{sq}{1-q} = [0, r_0]$. Given an $n \in \mathbb{N}$, if $[\alpha_1, \beta_1], i = 1, \dots, 3^{n-1}$, denote all component intervals of $I_{k(n-1)}$, that is, if we write

$$I_{k(n-1)} = \coprod_{i=1}^{3^{n-1}} [\alpha_1, \beta_1],$$

Then by the Observation C

$$I_{kn} = \coprod_{i=1}^{3n} ([\alpha_1, \alpha_1 + r_{kn}]\sqcup[\alpha_1 + mq^n, \beta_1 - mq^n]\sqcup[\beta_1 - r_{kn}, \beta_1]).$$

In particular, the iteration I_{kn} consists of 3^n component intervals. There are 3^{n-1} intervals among them concentric with component intervals of $I_{k(n-1)}$ and 2.3^{n-1} new intervals $-$ each of the new ones of length r_{kn}. In other words, the iteration I_{kn} is obtained from the iteration $I_{k(n-1)}$by removing 2.3^{n-1} open intervals - each of length $mq^n - r_{kn}$. We can say that, while passing from $I_{k(n-1)}$to Ikn, each component interval of $I_{k(n-1)}$ shrinks symmetrically by the same total length of $2mq^n$ and it produces two new intervals of length rkn in the process of shrinking. Hence

$$\mu I_{kn} = \mu I_{k(n-1)} - 3^{n-1}.2mq^n + 2.3^{n-1} r_{kn},$$

and thus,

$$\mu E = \lim_n \mu I_{kn} = \mu I_0 - \sum_{n=1}^{\infty} 3^{n-1}\, 2mq^n + \sum_{n=1}^{\infty} 2.\, 3^{n-1} r_{kn}. \qquad (4)$$

Recall that $s \geq \frac{3}{2}m(m+1)$ by the Fact (6.1.3) and $q < \frac{m}{s+m}$ by our assumption. Thus

$$q < \frac{m}{s+m} \leq \frac{m}{\frac{3}{2}m(m+1) + m} = \frac{2}{3m + 5} \leq \frac{1}{4}$$

Which justies convergence of both series in (4). Since $r_{kn} = s\, \frac{q^{n+1}}{1-q}$, we get $\mu E = \frac{sq}{1-q} - \frac{2qm}{1-3q} + \frac{2sq^2}{(1-q)(1-3q)} = (s-2m)\frac{q}{1-3q}$.

As we have mentioned above, the family of all component intervals of I_{kn} consists of 3^{n-1}shrinked component intervals of $I_{k(n-1)}$and of $2.\, 3^{n-1}$ new component intervals. We agree to say that $I_{k.0} = \left[0, \frac{sq}{1-q}\right]$ is the single new component interval of $I_{k.0}$. If $[\alpha, \beta]$ is a new component interval of I_{kn}, then its length is r_{kn} which, with our agreement, is true for $n = 0$ as well. When we pass to the next iteration $I_{k(n+1)}$,it shrinks by $2mq^{n+1}$, but maintains its center. When we pass to $I_{k(n+2)}$, it shrinks further by $2mq^{n+2}$ and so on. The intersection of all successive shrinked versions of $[\alpha, \beta]$ is a component interval of E of length

$$r_{kn} - \sum_{l=n+1}^{\infty} 2mq^l = (s-2m)\frac{q^{n+1}}{1-q}.$$

Hence each new component interval of I_{kn} is concentric with a component interval of E of length$(s-2m)\frac{q^{n+1}}{1-q}$. Since there are $2.\, 3^{n-1}$ new component intervals of I_{kn} (except for $n = 0$ when there is only one new component interval of I_0) and each of them is concentric with a component interval of E of length $(s-2m)\frac{q^{n+1}}{1-q}$, the sum of lengths of all component intervals of E concentric with a component interval of all possible iterations is

$$(s-2m)\frac{q}{1-q}\sum_{n=1}^{\infty} 2.\, 3^{n-1}(s-2m)\frac{q^{n+1}}{1-q} = (s-2m)\frac{q}{1-3q},$$

but it is exactly the Lebesgue measure of E. It follows that there are no other component intervals of E than those concentric with component intervals of iterations and that the sum of lengths of all component intervals of E is equal to the Lebesque measure of E.

The most accessible example of an M-Cantorval is given by the Guthrie-Nymann set GN ([147], [251])

$$GN = C \cup \bigcup_{n=1}^{\infty} G_{2n-1}$$

where C is the classic Cantor ternary set and G_k denotes the union of all open intervals removed from [0,1] in the k-th step of the standard construction. For example, $G_2 = \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)$ and $G_3 = \left(\frac{1}{27}, \frac{2}{27}\right) \cup \left(\frac{7}{27}, \frac{8}{27}\right) \cup \left(\frac{19}{27}, \frac{20}{27}\right) \cup \left(\frac{25}{27}, \frac{26}{27}\right)$. None of the known M-Cantorvals generated by Ferens type series is equal to the Guthrie-Nymann set GN ([147], [251]), not even up to a dilation, because the ratio of endpoints of the central connectivity component for the M-Cantorvals.

$$\frac{m\,\dfrac{q}{1-q}}{s\,\dfrac{q}{1-q}}=\frac{m}{s}\leq\frac{m}{\dfrac{3}{2}m(m+1)}=\frac{2}{3}\frac{1}{m+1}\leq\frac{2}{9}.$$

The last inequality follows from the fact that the condition $q\in\left(\frac{1}{s-2m+1},\frac{m}{s+m}\right)$ implies that $s+m\ <m(s-2m+1)$ and this inequality requires that $m\geq 2$. Hence the question of existence of a series $\sum a_n$ such that $GN=E(a_n)$ remains open which shows distinctly how little we know about sets of subsums.

Lemma (6.1.13)[246]: Let $\sum a_n$ be a convergent series of positive terms. If

$$r_n<\min\{\{f-g\}:f,g\in F_n,f\neq g\}$$

for in finitely many n, then $E(a_n)$ is a Cantor set.

Proof. Given a closed and bounded set $S\subset\mathbb{R}$, let $\gamma\mathcal{S}$ denote the length of the longest component of S; more formally, $\gamma\mathcal{S}:=\ sup\mu G$ where μ is the Lebesgue measure and where the supremum is taken over all connectivity components G of S. Now, let (n_k) be a sequence of indices such that

$$r_{n_k}<\min\left\{\{f-g\}:f,g\in F_{n_k},f\neq g\right\}\qquad\text{for all } k.$$

Then the iteration I_{kn} of $E(a_n)$ is the union of $\{F_{n_k}\}$ disjoint closed intervals each of length r_{n_k} . Therefore, $\gamma I_{n_k}=r_{n_k}$. Since the sequence $(I_n)_{n=1}^{\infty}$ is descending, the sequence $(\gamma I_n)_{n=1}^{\infty}$ is nonincreasing and

$$\gamma E=\lim_n\gamma I_n=\lim_k\gamma I_{n_k}=\lim_k r_{n_k}=0$$

and hence E is a Cantor set.

Theorem (6.1.14)[246]: Let $\mathcal{F}(m,k;q)$ be a Ferens type series. If $q\geq\frac{m}{s+m}$, then $E(\mathcal{F}(m,k;q))$ is the whole interval $\left[0,s\frac{q}{1-q}\right]$ $.If\ q<\frac{1}{s+1}$, then $E(\mathcal{F}(m,k;q))$ is a Cantor set.

Proof. If $q\geq\frac{m}{s+m}$, then $a_j\leq r_j$ for all j, and hence $E(\mathcal{F}(m,k;q))$ is the whole interval $\left[0,s\ \frac{q}{1-q}\right]$ by the well-known characterization given by Kakeya [150], [251].

We are going to show that if $q<\frac{1}{s+1}$, then

$$\min_{\substack{f,g\in F_{n_k}(\mathcal{F}(m,k;q))\\ f\neq g}}|f-g|>\frac{sq^{n+1}}{1-q}\qquad\text{for any } n\in\mathbb{N},\qquad(5)$$

and then a simple application of the Lemma (6.1.13) completes the proof that $E(\mathcal{F}(m,k;q))$ is a Cantor set for $q\in\left(0,\frac{1}{s+1}\right)$.

Suppose that $q<\frac{1}{s+1}$. It follows from the Lemma (6.1.5) that

$$\min_{\substack{f,g\in F_k\\ f\neq g}}|f-g|=q.$$

Observe that

$$q^t-\sum_{i=t+1}^{n}sq^i=q^t\left(\frac{1-(1+s)q+sq^{n-t+1}}{1-q}\right)$$

for any $t\in\mathbb{N},t<n$, and hence the inequality $q<\frac{1}{s+1}$ implies that

$$q^t - \sum_{i=t+1}^{n} sq^i > \frac{sq^{n+1}}{1-q}. \tag{6}$$

Take any $g, h \in F_{n_k}$ such that $g < h$. Then

$$g = \sum_{i=1}^{n} g_i q^i \qquad \text{and} \qquad h = \sum_{i=1}^{n} h_i q^i$$

for a suitable choice of $g_i, h_i \in \{0\} \cup \{m, m+1, \dots, s-m\} \cup \{s\}$ by the Lemma (6.1.5) again. Denote $j := \min\{i: g_i \neq h_i\}$. If $j = n$, then $h_n > g_n$ and thus

$$h - g = (h_n - g_n)q^n \geq q^n > s\frac{q^n}{1-q}.$$

If $j \in \{1, \dots, n-1\}$, then $h_j \geq g_j + 1$ and hence

$$h - g \geq \sum_{i=1}^{j} h_i q^i - \sum_{i=1}^{n} g_i q^i \geq h_j q^j - g_j g^j - \sum_{i=j+1}^{n} sq^i \geq g^j - s\sum_{i=j+1}^{n} q^i > s\frac{q^{n+1}}{1-q}$$

and thus (5) has been established.

Unfortunately, an important part of the above theorem has been made obsolete within the over two years that passed since our manuscript was prepared until it was eventually (and this time very quickly) accepted in Mathematica Slovaca. Indeed, meantime an extremely interesting [143] has been published by A. Bartoszewicz, M. Filipczak and E. Szymonik. Their results are concerned with general multigeometric series and easily apply to our series of Ferens type. In particular, [143] implies that the set of subsums of a Ferens type series is a Cantor set if $q < \frac{1}{s-2m+3}$, essentially improving our Theorem (6.1.14). Further, the Theorem 2.1 from [143] applied to Ferens type series yields the first part of our Theorem (6.1.9) exactly. Summing up, we do not know the topological type of the set of subsums of Ferens type series only for the very narrow interval $q \in \left[\frac{1}{s-2m+3}, \frac{1}{s-2m+1}\right)$. It requires further research, although some signi cant progress in that direction has been made in [158].

Section (6.2): Purely Atomic Finite Measure from its Range

Assume that μ is a purely atomic finite measure. We may assume that μ is defined on ℕ and $\mu(\{n\}) \geq \mu(\{n+1\})$. We assume that measures are always purely atomic, finite and they are defined on ℕ such that their n+1-st atoms have measures not greater than their n-th atoms. We are interested in the following questions:

(a) For which subsets R of ℝ there is a measure μ such that R is its range (i.e. $R = rng(\mu) := \{\mu(E): E \subset \mathbb{N}\}$)?

(b) For which subsets R of ℝ there is exactly one measure μ with $R = rng(\mu)$?

To simplify the notation let $x_n = \mu(\{n\})$ be a measure of the n-th largest atom of μ. Note that

$$rng(\mu) = \{\mu(E): E \subset \mathbb{N}\} = \left\{\sum_{n\in E} \mu(\{n\}): E \subset \mathbb{N}\right\} = \left\{\sum_{n=1}^{\infty} \varepsilon_n x_n: \varepsilon_n = \{0,1\}^N\right\}.$$

The latter set is also denoted by $A(x_n)$ and it is called the achievement set of (x_n) (see [149]). Let us present here two simple examples.

Example (6.2.1)[253]: Consider the procedure of rolling dice until the value on the dice is less than 5. For $E \subset \mathbb{N}$ let $\mu_1(E)$ be the probability that the procedure stops for some n from

E. Then $\mu_1(\{n\}) = \frac{2}{3^n}$. It is easy to see that for $x_n = \mu_1(\{n\})$ the set $A(x_n)$, or rng(μ_1), is equal to the classical Cantor ternary set C.

Example (6.2.2)[253]: Consider the procedure of tossing a fair coin until the head appears. For $E \subset \mathbb{N}$ let $\mu_2(E)$ be the probability that the procedure stops for some n from E. Then $\mu_2(\{n\}) = \frac{1}{2^n}$ and $rng(\mu_2) = [0, 1]$.

Achievement sets of sequences, defined for all summable sequences (x_n), have been considered by many some results have been rediscovered several times. Let us list basic properties of $A(x_n)$ (some of them were observed by Kakeya in [150] in 1914):

(i) $A(x_n)$ is a compact perfect or finite set,

(ii) If $|x_n| > \sum_{i>n} |x_i|$ for all sufficiently large n's, then $A(x_n)$ is homeomorphic to the ternary Cantor set C,

(iii) If $|x_n| \le \sum_{i>n} |x_i|$ for all sufficiently large n's, then $A(x_n)$ is a finite union of closed intervals. Moreover, if $|x_n| \ge |x_{n+1}|$ for all but finitely many n's and $A(x_n)$ is a finite union of closed intervals, then $|x_n| \le \sum_{i>n} |x_i|$ for all but finitely many n's.

In particular, for decreasing sequence (x_n) the inequality $x_n \le \sum_{i>n} x_i$ for all n is equivalent to $A(x_n)$ being an interval.

One can see that $A(x_n)$ is finite if and only if $x_n = 0$ for all but finite number of n's, i.e. $(x_n) \in c_{00}$. Kakeya conjectured that if $(x_n) \in \ell_1 \backslash c_{00}$, then $A(x_n)$ is always a Cantor set C or it is a finite union of intervals.

On the other hand, in 1970 Renyi in [266] repeated the results of Kakeya in terms of purely atomic measures and he asked if the Cantor sets and finite unions of closed intervals are the only possible sets being the ranges of finite measures. Geometric properties of achievement sets of sequences and ranges of purely atomic finite measures are the same. This follows from the simple observation, that the set of sums of subseries for the series $\sum_{n=1}^{\infty} x_n$ is isometric to the analogous set for the series of their absolute values $\sum_{n=1}^{\infty} |x_n|$. Therefore a positive answer for the Renyi's question is equivalent to the Kakeya's conjecture.

In 1980 Weinstein and Shapiro in [157] gave an example which showed that the Kakeya conjecture is false. It follows from the references that they did not know the Renyi's problem. On the other hand, Ferens in [146] has given the example similar to that of Weinstein and Shapiro, solving the problem of Renyi. In this case, We did not know the conjecture of Kakeya.

In [147] Guthrie and Nymann gave a very simple example of a sequence whose achievement set is not a finite union of closed intervals but it has a nonempty interior. They used the sequence $(t_n) = \left(\frac{3}{4}, \frac{2}{4}, \frac{3}{16}, \frac{2}{16}, \dots\right)$. Moreover, they formulated the following:

Theorem (6.2.3)[253]: For any $(x_n) \in \ell_1 \backslash c_{00}$, the set $A(x_n)$ is one of the following types:

(i) a finite union of closed intervals,

(ii) a Cantor set C,

(iii) homeomorphic to the set $T = A(t_n) = A\left(\frac{3}{4}, \frac{2}{4}, \frac{3}{16}, \frac{2}{16}, \frac{3}{64}, \dots\right)$.

Although their proof had a gap, the theorem is true and the correct proof was given by Nymann and Saenz in [155]. Guthrie, Nymann and Saenz have observed that the set T is homeomorphic to the set N described by the formula

$$N = [0, 1] \backslash \bigcup_{n \in \mathbb{N}} U_{2n},$$

where U_n denotes the union of 2^{n-1} open middle thirds which are removed from the interval [0, 1] at the n-th step in the construction of the classic Cantor ternary set C. Such sets are called Cantorvals in the literature (to emphasize the similarity to the interval and to the Cantor set simultaneously). It is known that a Cantorval is just a nonempty compact set in $\mathbb{R}$, that it is the closure of its interior and both endpoints of any nontrivial component are accumulation points of its trivial components. Other topological characterizations of Cantorvals can be found in [256] and [152].

All known examples of sequences whose achievement sets are Cantorvals belong to the class of multigeometric sequences or are linear combinations of such sequences, see [144],[254]. This class was deeply investigated in [149], [143], [246] and [158]. In particular, the achievement sets of multigeometric series and similar sets obtained in more general case are the attractors of affine iterated function systems, see [158]. More information on achievement sets can be found in the surveys [256], [153] and [164].

It is almost obvious that any achievement set E of a summable sequence contains zero and is symmetric in the sense that there exists a number t such that if $t - x \in E$ then $t + x \in E$ too. It is a natural question if every compact, perfect set with these properties is an achievement set for some sequence. This question was posted by W. Kubiś in Łódź in 2015. In particular, in [246] the authors ask if the Cantorval N is an achievement set of any sequence.

The negative answer to the last question was recently given in [255]. Independently [257] have showed that the Cantorval $\tilde{T}$ for which the gaps are the intervals of the Guthrie-Nymann-Cantorval T and vice-versa, is not an achievement set for any sequence.

On the other hand, T. Banakh in Lviv in 2016 asked if Cantor achievement sets are uniquely defined, i. e. they are achievement sets of only one sequence.

We present gap lemmas and the center of distances notion which are useful tools. We show that if the range of a mesure μ is an interval, in other words μ is interval filling, then there is a measure ν such that the sets $\{\mu(n): n \in \mathbb{N}\}$ and $\{\nu(n): n \in \mathbb{N}\}$ are pairwise disjoint. We also give an example of a symmetric set which is a finite union of intervals but is not the range of any measure. We give sufficient conditions on a Cantor set which is the range of some measure to be the range of no other measure. We present also sufficient conditions for a set R to be a Cantor set achieved by a unique measure μ. There is given a connection between achievement sets of multigeometric sequences and IFS fractals. We show that the Guthrie-Nymann Cantorval is uniquely achieved. We show that some Ferens fractals which are symmetric Cantors or Cantorvals are not ranges of any measure. We briefly discus the Guthrie-Nymann-Jones Cantorvals A(r) of one parameter $r = 1, 2, \ldots$ which generalize the Guthrie-Nymann Cantorval. For some r, A(r) is not a range of any measure; for some $r, A(r)$ can be achieved in continuum many ways by measure range; $A(1)$ is a Guthrie-Nymann Cantorval which is uniquely achieved.

We assume that (x_n) is a nonincreasing summable sequence of positive real numbers – the measures $\mu(\{n\})$ of μ-atoms. Denote (as in [147], [155], [256]):

$$R = A(x_n) = \left\{\sum_{n=1}^{\infty} \varepsilon_n x_n : (\varepsilon_n) \in \{0,1\}^N\right\};\ F_k = \left\{\sum_{n=1}^{k} \varepsilon_n x_n : (\varepsilon_n) \in \{0,1\}^k\right\}.$$

So F_k is a finite approximation of the range R. Let $r_k := \sum_{n=k+1}^{\infty} x_n$. By a gap in the range R we understand any interval (a, b) such that $a \in R, b \in R$ and $(a, b) \cap R = \emptyset$. The following two lemmas can be found in [256]. The first is obvious.

Lemma (6.2.4)[253]: (First Gap Lemma) If $x_k > r_k$ then (r_k, x_k) is a gap in the range R.

The next observation is extracted from the proof of the crucial Lemma 4 of [155], where it was formulated as not a quite correct claim (however the Lemma and the main result of [155] are true). It can be found in [255] or in [256].

Lemma (6.2.5)[253]: (Second Gap Lemma) Let (a, b) be a gap in the range R, and let p be defined by the formula $p := \max\{n : xn \geq b - a\}$. Then:

(i) $b \in F_p$,

(ii)If $F_p = \left\{f_1^{(p)} < f_2^{(p)} <\ldots< f_{m(p)}^{(p)}\right\}$ and $b = f_j^{(p)}$, then $a = f_{j-1}^{(p)} + r_p$.

The next Lemma has recently been proved in [255]. Since it will be used several times and we present it with the proof.

Lemma (6.2.6)[253]: (Third Gap Lemma) Suppose that (a, b) is a gap in the range R such that for any $gap\ (a_1, b_1)$ with $b_1 < a_1$ we have $b - a > b_1 - a_1$ (in other words (a, b) is the longest gap from the left). Then for some $k \in \mathbb{N}$ we have $b = x_k$ and $a = r_k$.

Proof. By the Second Gap Lemma b is a finite sum of terms of (x_n). Let $b = x_{n_1} +\ldots+ x_{n_m}$ with $x_{n_1} \geq\ldots\geq x_{n_m}$. Suppose that $m \geq 2$. Firstly observe that $x_{n_m} \geq b - a$ (indeed, if $x_{n_m} < b - a$ then $b - x_{n_m} \in (a, b) \cap R$ which is impossible). Of course $x_{n_m} < b$ and, since (a, b) is a gap, $x_{n_m} \leq a$. Any gap in the set $X := R \cap \left[0, x_{n_m}\right]$ is shorter than $b - a$. On the other hand, $b \in X + \left(b - x_{n_m}\right)$ and $X + \left(b - x_{n_m}\right) \subset R$, so $(a, b) \cap \left(X + \left(b - x_{n_m}\right)\right) = \emptyset$, and hence $\left(a - b + x_{n_m}, x_{n_m}\right)$ is the gap in X which gives a contradiction. Thus $m = 1$ which means that $b = x_k$ for some $k \in \mathbb{N}$.

Since $a \in R, r_k \geq a$. Suppose that $r_k > a$. Let m be the smallest number satisfying $\sum_{n=k+1}^{m} x_n > a$. Hence $\sum_{n=k+1}^{m} x_n > b$, because (a, b) is a gap. Let now $X := R \cap [0, x_m]$. Then the set $X + \sum_{n=k+1}^{m-1} x_n$ is included in E and it has all gaps shorter than b-a, which gives a contradiction again.

In [257] the authors have introduced the notion of the center of distances of a metric space X, defined as $S(X) = \left\{\alpha : \forall x \in X \exists_{y \in X} d(x, y) = \alpha\right\}$. They especially consider the case when X is the achievement set of a sequence (x_n) and observe the following.

Lemma (6.2.7)[253]: ([257]) $\{x_n : n \in \mathbb{N}\} \subset S\big(A(x_n)\big) \subset A(x_n)$.

We present a short proof of this for the readers' convenience.

Proof. Let $n \in \mathbb{N}$. Fix $t \in A(x_n)$. Then there is $E \subset \mathbb{N}$ with $t = \sum_{m \in E} x_m$. If $n \in E$, then $t - x_n \in A(x_n)$. If $n \notin E$, then $t + x_n \in A(x_n)$. Therefore for any $t \in A(x_n)$ there is $s \in A(x_n)$ with $|t - s| = x_n$, which means that $x_n \in S\big(A(x_n)\big)$. Since $0 \in A(x_n)$, then for any $t \in S\big(A(x_n)\big)$ by the definition of the center of distances there is $s \in A(x_n)$ with $|s - 0| = t$. Since $A(x_n)$ consists of nonnegative real numbers, $s = t$ and consequently $S\big(A(x_n)\big) \subset A(x_n)$.

[257] have given a variety of examples of sequences for which the equality $S(X) = \{x_n\} \cup \{0\}$ holds. Some of them are geometric sequences $(aq^n)_{n=1}^{\infty}$ with $q < \frac{1}{2}, a \geq 0$. They also proved that for the Guthrie-Nymann-Cantorval $T = A(x_n)$, where $x_{2^{n-1}} = \frac{3}{4^n}, x_{2n} = \frac{2}{4^n}$ we also get $S(X) = \{x_n\} \cup \{0\}$. For more details see [147].

The previous Lemma can be completed as follows.

Lemma (6.2.8)[253]: If $x_k = x_{k+1} =\cdots= x_{k+2j-2}$ for some k and j, then jx_k belongs to $S\big(A(x_n)\big)$.

Proof. Let us observe that if we replace the terms $x_k, x_{k+1}, \ldots, x_{k+j-1}$ in the sequence (x_n) by one term jx_k, then in the modified sequence we can obtain any number m_{x_k} where $m = 1, 2, \ldots, k+2j-2$ by summing up some of the new terms $jx_k, x_{k+1}, \ldots, x_{k+2j-2}$. Consequently $A(x_n)$ equals the achievement set of the modified sequence. Therefore by Lemma (6.2.7) we obtain that $j_{x_k} \in S(A(x_n))$.

We say that a purely atomic finite measure is interval filling if its range is an interval. A sequence of values of such a measure on its atoms is called an interval filling sequence. This notion was introduced in [258] and intensively studied f.e. in [259], [260]. By the Kakeya Theorem, a nonincreasing, summable sequence (x_n) of positive numbers is interval filling if and only if it is slowly convergent, i.e. if for every n the term x_n is no greater than the rest $r_n = \sum_{k=n+1}^{\infty} x_k$ ([258] have rediscovered this result). It is almost obvious that we cannot uniquely recover a sequence if its achievement set is an interval.

Example (6.2.9)[253]: $[0, 1] = A(x_n) = A(y_n)$, where $x\_n = \frac{1}{2^n}, y_{2n-1} = \frac{1}{3^n} = y_{2n}$. One can easily observe, that the both sequences (x_n) and (y_n) are slowly convergent and $\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty} y_n = 1$.

It is worth noticing that it follows from the above example that an algebraic sum of two copies of the Cantor ternary set is an interval, what was proved by Steinhaus [261] (see [262]) about three years later than Kakeya has published his results. It is also interesting that the sets of values of (x_n) and (y_n) are not only different but even disjoint. We denote this by $(x_n) \cap (y_n) = \emptyset$. However, with additional assumptions, the authors of [263], [264], [265] have obtained some uniqueness results for interval filling finite measures. The following theorem is an improvement of Example (6.2.9).

Theorem (6.2.10)[253]: For a given set R which is the range of some measure, the following conditions are equivalent:

(i) R is an interval,
(ii) there are two purely atomic measures μ and ν, both with range R, such that the μ-measures of atoms are all distinct from the ν-measures of atoms,
(iii) for any purely atomic measure μ with range R, there is another purely atomic measure ν whose values on atoms are distinct from those of μ,
(iv) for any purely atomic measure μ with range R, there is another purely atomic measure ν whose values on finite nonempty sets are distinct from those of μ.

Proof. Evidently (iv) ⇒ (iii) and (iii) ⇒ (ii).

(ii)⇒(i). Let us assume that $R = A(x_n) = A(y_n)$, where $x_n = \mu(\{n\})$ and $y_n = \nu(\{n\})$, and suppose that R is not an interval. Then R has a gap. Let (a, b) be the longest gap in R (there may be finitely many longest gaps and we choose the one from the left side). By Lemma (6.2.6) there exist natural numbers k and l for which $x_k = y_l = b$. Thus $\mu(\{k\}) = \nu(\{l\})$ which yields a contradiction with (ii).

(i)⇒(iv). Without loss of generality we may assume that the range R of μ equals [0, 1]. Let us construct inductively (y_n) such that

(a) y_1 is any number in $\left(\frac{1}{3}, \frac{1}{2}\right) \backslash \{\mu(F): F\ is\ finite\}$;
(b) $y_{n+1} > \frac{1}{3}(1 - \sum_{i=1}^{n} y_i)$;
(c) $y_{n+1} < \frac{1}{2}(1 - \sum_{i=1}^{n} y_i)$;
(d) $y_{n+1} \neq \mu(F) - \sum_{i=1}^{n} \varepsilon_i y_i$ for any finite $F \subset \mathbb{N}$ and any $(\varepsilon_i)_{i=1}^{n} \in \{0, 1\}^n$.

Since the set of forbidden numbers $\mu(F)-\sum_{i=1}^{n} \varepsilon_i y_i$ for y_{n+1} prescribed in (d) is countable, the choice of such sequence (y_n) is possible.

We will show inductively that $1-\sum_{i=1}^{n+1} y_i < \left(\frac{2}{3}\right)^{n+1}$. By (a) we have $y_1 > \frac{1}{3}$ which implies $1-y_1 < \frac{2}{3}$. Using (b) and the inductive assumption we obtain

$$1-\sum_{i=1}^{n+1} y_i = 1-\sum_{i=1}^{n} y_i - y_{n+1} < 1-\sum_{i=1}^{n} y_i - \frac{1}{3}\left(1-\sum_{i=1}^{n} y_i\right) = \frac{2}{3}\left(1-\sum_{i=1}^{n} y_i\right)$$
$$< \frac{2}{3}\cdot\left(\frac{2}{3}\right)^n = \left(\frac{2}{3}\right)^{n+1}.$$

Hence $\sum_{n=1}^{\infty} y_n = 1$. By (a) we have $y_1 < \frac{1}{2}$, which implies $\sum_{n=2}^{\infty} y_n = 1-y_1 > \frac{1}{2} > y_1$. Using (c) we obtain $y_{n+1} < \frac{1}{2}(1-\sum_{i=1}^{n} y_i) = \frac{1}{2}\sum_{i=n+1}^{\infty} y_i = \frac{1}{2}y_{n+1} + \frac{1}{2}\sum_{i=n+2}^{\infty} y_i$. Thus $y_n < \sum_{i=n+1}^{\infty} y_i$ for every n. Therefore $A(y_n) = R$.

Let ν be a measure such that $\nu(\{n\}) = y_n$. Finally we will show that $\nu(G) \notin \{\mu(F): F$ is finite and nonempty$\}$ for every nonempty $G \subset \mathbb{N}$. If $\max G = 1$, then $G = \{1\}$ and $\nu(G) = y_1$. By (a) we obtain that $\nu(G) \notin \{\mu(F): F$ is finite$\}$. Assume now that that $\max G = n+1$ for some $n \in \mathbb{N}$. Then $\nu(G) = y_{n+1} + \sum_{i=1}^{n} \varepsilon_i y_i$ for some $(\varepsilon_i)_{i=1}^{n} \in \{0,1\}^n$. By (d) we obtain that $\nu(G) \notin \{\mu(F): F$ is finite$\}$ as well.

Recall that if the inequality $x_n \le \sum_{i=n+1}^{\infty} x_i$ holds for all $n > k$, then $A(x_n)$ is a finite union of closed intervals. $A(x\_n) = \left\{\sum_{i=1}^{n} \varepsilon_i x_i : (\varepsilon_i)_{i=1}^{k} \in \{0,1\}^k\right\} + A((x_n)_{n>k})$. So, we have:

Proposition (6.2.11)[253]: The range R of a measure is a finite union of intervals if and only if there exist two measures μ and ν with $R = rng(\mu) = rng(\nu)$ and the set $\{\mu(\{n\}): n \in \mathbb{N}\} \cap \{\nu(\{n\}): n \in \mathbb{N}\}$ is finite.

Proof. It follows from Lemma (6.2.6) and Theorem (6.2.10).

We already know that if the range R of a measure is a finite union of intervals, then the measure μ with $R = rng(\mu)$ is not unique. Let us consider the opposite question – for which sets X being a finite union of intervals is there a measure with $rng(\mu) = X$? As it was mentioned the range rng(μ) of a measure μ, or achievement set $A(x_n)$, contains zero and is symmetric. $\frac{1}{2}\mu(\mathbb{N})$ is a point of reflection of rng(μ). To see it, fix $E \subset \mathbb{N}$ and note that $\mu(E) + \mu(\mathbb{N}\backslash E) = \mu(\mathbb{N})$.

Note that if achievement set is a union of two closed intervals, then both of them have the same length by symmetry. It is clear that $A(x_n) = [0,a] \cup [b, b+a]$, where $b > a$, holds for $x_1 = b$ and $x_{n+1} = \frac{a}{2^n}$ for $n \in \mathbb{N}$, so we may obtain any union of two closed intervals having the same length as an achievement set. Moreover (a, b) is the only gap, so by Lemma (6.2.6) we get $y_1 = b$ for any (y_n) such that $A(y_n) = [0,a] \cup [b, b+a]$. The case become more complicated when we consider the union of three closed intervals, that is $[0,a] \cup [b, b+c] \cup [2b-a+c, 2b+c]$ – this is a general form of symmetric union of three disjoint intervals which contains zero. The question is whether there exists a sequence (x_n) such that $A(x_n) = [0,a] \cup [b, b+c] \cup [2b-a+c, 2b+c]$. It turns out that some sets of the form $[0,a] \cup [b, b+c] \cup [2b-a+c, 2b+c]$ are not ranges of measures, while some others are. We are far from the full characterization of finite unions of intervals (or even unions of three intervals) which are ranges of measures, but we present some partial results which suggest that such characterization will be complicated.

Proposition (6.2.12)[253]: If $2a < c < 2b$, then $[0,a] \cup [b,b+c] \cup [2b-a+c,2b+c]$ is not a range of purely atomic measure.

Proof. Suppose that $A(x_n) = [0,a] \cup [b,b+c] \cup [2b-a+c,2b+c]$ for some (x_n). By Lemma (6.2.6) there exists $l \in \mathbb{N}$ such that $x_l = b$. By Lemma (6.2.7) we obtain $b \in S(A(x_n))$. Let $x := b + \frac{c}{2}$. Then $x \in A(x_n)$, and consequently $x + b \in A(x_n)$ or $x - b \in A(x_n)$. But $x + b = 2b + \frac{c}{2} \in (b+c, 2b-a+c)$ and $x - b = \frac{c}{2} \in (a,b)$, which are the gaps of $A(x_n)$. A contradiction.

Proposition (6.2.13)[253]: If $a \le c \le 2a$, then there exists a sequence (x_n) such that $A(x_n) = [0,a] \cup [b,b+c] \cup [2b-a+c,2b+c]$.

Proof. Define $x_1 = b+c-a, x_2 = b, x_{n+2} = \frac{a}{2^n}$ for $n \in \mathbb{N}$. It is clear that $A(x_n) = [0,a] \cup [b,b+c] \cup [2b-a+c,2b+c]$.

Proposition (6.2.14)[253]: If $b = 2a$ and $c \ge 2b$, then there exists a sequence (x_n) such that $A(x_n) = [0,a] \cup [b,b+c] \cup [2b-a+c,2b+c]$.

Proof. Let $c \ge 2b$. Then there exist unique $k \ge 2$ and $c \in [0,b)$ such that $c = kb + c$. Define $x_1 = 3a + \frac{c}{2}, x_2 = 2a + \frac{c}{2}, x_n = 2a$ for $n \in \{3,\dots,k+1\}, x_n = \frac{a}{2^{n-k-1}}$ for $n \ge k+2$ (or any other slowly convergent series with sum a). Thus

$$A(x_n) = \left\{\sum_{i=1}^{k+1} \varepsilon_i x_i : (\varepsilon_i)_{i=1}^{k+1} \in \{0,1\}^{k+1}\right\} + A((x_n)_{n\ge k+2})$$

$$= \bigcup_{m=0}^{k-1} \left\{2ma, 2a + \frac{c}{2} + 2ma, 3a + \frac{c}{2} + 2ma, 5a + c + 2ma\right\} + [0,a].$$

Hence

$$A(x_n) = [0,a] \cup [2a, c+2ka] \cup [a+c+2ka, 2a+c+2ka]$$
$$= [0,a] \cup [b,b+c] \cup [2b-a+c,2b+c].$$

Now we present a characterization of finite unions of intervals which are ranges of purely atomic measures. However, this characterization will not be very informative. It is hard to prove using it that some finite union of intervals is not a range of any measure.

Proposition (6.2.15)[253]: Let X be a finite union of intervals. Then X is a range of a finite measure if and only if there is a measure ν on a finite set such that $X = rng\, \nu + [0,a]$ for some $a > 0$.

Proof. Assume that $X = rng(\mu)$ for some measure μ on N. Let $x_n = \mu(\{n\})$. Then there is $m \in \mathbb{N}$ such that $x_n \le r_n$ for every $n \ge m$. Hence the achievement set $A((x_n)_{n\ge m})$ is an interval, say [0, a]. Let $\nu(n) = \mu(\{n\})$ for $n < m$. Then ν is a finite measure defined on $\{1,2,\dots,m-1\}$. Thus

$$rng(\mu) = \left\{\sum_{n=1}^{\infty} \varepsilon_n x_n : \varepsilon_n = 0,1\right\} = \left\{\sum_{n=1}^{m-1} \varepsilon_n x_n + \sum_{n=m}^{\infty} \varepsilon_n x_n : \varepsilon_n = 0,1\right\}$$
$$= \left\{\sum_{n=1}^{m-1} \varepsilon_n x_n : \varepsilon_n = 0,1\right\} + \left\{\sum_{n=m}^{\infty} \varepsilon_n x_n : \varepsilon_n = 0,1\right\} = rng(\nu) + [0,a].$$

On the other hand if $X = rng(\nu) + [0,a]$ for some measure ν on a finite set $F = \{1,2,\dots,n\}$. Let λ be a measure on $\{n+1, n+2,\dots\}$ with $rng(\lambda) = [0,a]$. Thus the measure μ defined as $\mu(E) = \nu(E \cap \{1,2,\dots,n\}) + \lambda(E \cap \{n+1, n+2,\dots\})$ has the range equal to X.

Let us start from the following example.

Example (6.2.16)[253]: Let R be the range of the measure from Example (6.2.1), that is $\mu(\{n\}) = \frac{2}{3^n}$. Observe that the numbers $x_n = \frac{2}{3^n}$ are the right ends of the longest gaps of R from the left. Suppose that $A(y_n) = R$ for some sequence (y_n) with $y_1 \geq y_2 \geq \cdots$. Then $\{x_n : n \in \mathbb{N}\} \subset \{y_n : n \in \mathbb{N}\}$. Observe that $\sum_{n=1}^{\infty} x_n = 1$, so $y_n = x_n$ for every $n \in \mathbb{N}$. Hence the ternary Cantor set is obtained in the unique way as achievement set of nonincreasing sequence by the sequence (x_n).

Now, let us consider the question: which sets R are ranges of the uniquely defined measures μ. More precisely, for which sets $R = rng(\mu)$ for some measure μ, the equality $R = rng(\nu)$ for some measure ν implies that $\mu = \nu$. A sequence from Example (6.2.16) satisfies $x_n = 2r_n$ for each $n \in \mathbb{N}$ and $A(x_n)$ is the ternary Cantor set, which is obtained in the unique way. Simple observation shows that the uniqueness of a sequence (x_n) generating the achievement set $A(x_n)$ can be obtained as a direct consequence of Lemma (6.2.6) if $x_n \geq 2r_n$ for each $n \in \mathbb{N}$. The next theorem improves that result.

Theorem (6.2.17)[253]: Assume that $\mu(\{n\}) > 2\mu(\{n+1\})$ for $n \in \mathbb{N}$. If $rng(\mu) = rng(\nu)$ then $\mu = \nu$.

Proof. Fix $m \in \mathbb{N}$. As usually $x_m = \mu(\{m\})$. Observe that $x_m > r_m$, where $r_m = \sum_{k=m+1}^{\infty} x_k$. Indeed

$$x_m > 2x_{m+1} > x_{m+1} + 2x_{m+2} > x_{m+1} + x_{m+2} + 2x_{m+3} > \cdots$$

Hence $x_m > r_m - r_{m+k} + x_{m+k}$ for each $k \in \mathbb{N}$. Since $(r_m - r_{m+k} + x_{m+k})_{k=1}^{\infty}$ is a decreasing sequence tending to r_m, we get $x_m > r_m$.

By Lemma (6.2.4) we obtain that $(r_m, x_m) \cap A(x_n) = \emptyset$. Now we will show that (r_m, x_m) is the longest gap from the left in $A(x_n)$. Indeed for each $m \in \mathbb{N}$ we have

$$x_m - r_m = x_m - \sum_{k=m+1}^{\infty} x_k > 2x_{m+1} - \sum_{k=m+1}^{\infty} x_k = x_{m+1} - \sum_{k=m+2}^{\infty} x_k = x_{m+1} - r_{m+1}.$$

Hence no gap of the form (r_k, x_k) is longer than (r_m, x_m) for $m < k$. Suppose now that (a, b) is the longest gap from the left and $b \notin \{x_n : n \in \mathbb{N}\}$. However by Lemma (6.2.6) the point b should be a term of any sequence (y_n) for which $A(x_n) = A(y_n)$. This yields a contradiction.

Finally by Lemma (6.2.6) we get that if $A(y_n) = A(x_n)$ then $(y_n) \subset (x_n)$. By comparing sums of the series $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ we get $y_n = x_n$.

Example (6.2.18)[253]: All Cantor sets of the form $A(q^n)$ for $q < \frac{1}{2}$ are uniquely defined.

Theorem (6.2.17) can be used to obtain uniquely defined Cantor sets with positive Lebesgue measure.

Example (6.2.19)[253]: Let $q \in \left(0, \frac{1}{2}\right)$ and $x_n = \frac{1}{2^n} + q^n$ for $n \in \mathbb{N}$. Then a sequence (x_n) satisfies a condition given in Theorem (6.2.17), so the set $A(x_n)$ is the achievement set of the only one sequence. Moreover the Lebesgue measure of the set $A(x_n)$ can be calculated by the formula given in [256], namely $\lambda\big(A(x_n)\big) = \lim_{n\to\infty} 2^n r_n = \lim_{n\to\infty} 2^n \left(\frac{1}{2^n} + \frac{q^{n+1}}{1-q}\right) = 1$. Hence we constructed a family of uniquely defined Cantors with positive Lebesgue measure.

The next example shows that the assumption $x_n > 2x_{n+1}$ for $n \in \mathbb{N}$ in Theorem (6.2.17) is optimal in some sense. One may think that if we assume weaker condition that a series is quickly convergent, in symbols $x_n > r_n$ for $n \in \mathbb{N}$, then the assertion of Theorem (6.2.17) is still true. However it is not, even when we additionally assume that $x_n \geq 2x_{n+1}$ for $n \in \mathbb{N}$.

Example (6.2.20)[253]: Let us consider the multigeometric sequence defined as $x_{2n-1} = \frac{2}{5^n}, x_{2n} = \frac{1}{5^n}$ for each $n \in \mathbb{N}$. Observe that $x_n > r_n$ and $x_n \geq 2x_{n+1}$ for each $n \in \mathbb{N}$, so the series $\sum_{n=1}^{\infty} x_n$ is quickly convergent, but the condition $x_n > 2x_{n+1}$ is satisfied only for even n's. Define $y_{3n-2} = y_{3n-1} = y_{3n} = \frac{1}{5^n}$. Then we have $A(x_n) = A(y_n)$.

Theorem (6.2.21)[253]: Assume that $R = rng(\mu)$ and $\{(a_n, b_n): n \in \mathbb{N}\}$ is a sequence of gaps in R such that

(i) (a_1, b_1) is the longest gap in R and any other gap in R is shorter;

(ii) $|b_{n+1} - a_{n+1}| < |b_n - a_n|$ for every $n \in \mathbb{N}$;

(iii) (a_{n+1}, b_{n+1}) is the longest gap in $R \cap [0, a_n]$ and any other gap in $R \cap [0, a_n]$ is shorter.

Then $\mu(\{n\}) = b_n$. Moreover, R is a Cantor set.

Proof. Since (a_1, b_1) is the only longest gap in R, then the middle point of (a_1, b_1) equals $\frac{1}{2}\mu(N)$. Thus $b_1 > \frac{1}{2}\mu(N)$. By Lemma (6.2.6) the number b_1 is equal to some $\mu(\{n\})$ and $a_1 = \mu(N\backslash\{1,2,\ldots,n\})$. Since only one $\mu(\{n\})$ may be greater than $\frac{1}{2}\mu(N)$, then $b_1 = \mu(\{1\})$ and $a_1 = \mu(\mathbb{N}\backslash\{1\})$. Consider a measure μ_1 defined on $\mathbb{N}\backslash\{1\}$ given by $\mu_1(E) = \mu(E)$ for $E \subset \mathbb{N}\backslash\{1\}$. Then $rng(\mu_1) = R \cap [0, a_1]$. Then (a_2, b_2) is the only longest gap in $rng(\mu_1)$. Repeating the same argument we obtain that $b_2 = \mu_1(\{2\}) = \mu(\{2\})$. Proceeding inductively we obtain that $\mu(\{n\}) = b_n$.

The "moreover" part of the assertion follows from the inequality $b_n > \sum_{m>n} b_m$ for every n and from Kakeya's Theorem.

Note that the existence of a sequence $\{(a_n, b_n): n \in \mathbb{N}\}$ of gaps in R fulfilling conditions (i)–(iii) from the Theorem (6.2.21) is equivalent to the following statement: between every two gaps of the same length there is a longer gap.

Theorem (6.2.22)[253]: Assume that R is a compact subset of the real line with $\min R = 0, a_0 = \max R > 0$ and $\{(a_n, b_n): n \in \mathbb{N}\}$ is a sequence of gaps in R such that

(i) $|b_{n+1} - a_{n+1}| < |b_n - a_n|$ for every $n \in \mathbb{N}$;

(ii) (a_{n+1}, b_{n+1}) is the longest gap in $R \cap [0, a_n]$ and any other gap in $R \cap [0, a_n]$ is shorter for every $n \geq 0$;

(iii) $\frac{1}{2} a_n$ is a point of reflection of $R \cap [0, a_n]$ for every $n \geq 0$.

Then $R = rng(\mu)$ with $\mu(\{n\}) = b_n$. Moreover, R is a Cantor set.

Proof. Since (a_1, b_1) is the only longest gap in R and $a_0/2$ is the point of reflection of R, then $a_0/2$ is the middle point of (a_1, b_1). Similarly (a_2, b_2) is the only longest gap in $R \cap [0, a_1]$ and $a_0/2$ is the point of reflection of $R \cap [0, a_2]$. Thus $a_1/2$ is the middle point of (a_2, b_2). Since $a_0/2$ is the point of reflection of R, then $(b_1 + a_2, b_1 + b_2)$ is a gap in R. Note that $|a_0 - (b_1 + b_2)| = a_2$.

The same as in the previous two steps one can show that $a_2/2$ is the middle point of (a_3, b_3) and since $a_0/2$ is the point of reflection of R, then $(b_1 + b_2 + a_3, b_1 + b_2 + b_3)$ is a gap in R. Note that $|a_0 - (b_1 + b_2 + b_3)| = a_3$. Since $a_n \to 0$, then proceeding inductively we obtain that $\sum_{n=1}^{\infty} b_n = a_0 = \max R$.

Let $R' = A(b_n)$. Note that $b_n > a_n = \sum_{m>n} b_m$. Therefore R' is a Cantor set. By Lemma (6.2.5) the gaps in R' are of the form $\left(a_n + f_{j-1}^{(n)}, f_j^{(n)}\right)$ where $F_n = \left\{0 = f_1^{(n)} < f_2^{(n)} < \cdots < f_m^{(n)}(n)\right\}$. Since $b_1 > b_2 > \cdots$, then there are no elements of $A(b_n)$ in $\left(a_n + \sum_{i=1}^{n-1} \varepsilon_i b_i, b_n + \sum_{i=1}^{n-1} \varepsilon_i b_i\right)$ where $\varepsilon_i = 0, 1$, which shows that these intervals are gaps in

R'. Clearly any gap of the length $|b_n - a_n|$ must be of the form $\left(a_n + \sum_{i=1}^{n-1} \varepsilon_i b_i, b_n + \sum_{i=1}^{n-1} \varepsilon_i b_i\right)$ for some $\varepsilon_i = 0, 1$. Therefore the set of all gaps in R' is the following $\left\{\left(a_n + \sum_{i=1}^{n-1} \varepsilon_i b_i, b_n + \sum_{i=1}^{n-1} \varepsilon_i b_i\right): n \in \mathbb{N}, \varepsilon_i = 0, 1\right\}$. Now we will prove inductively that every gap of R' is also a gap of R. Clearly R has exactly one gap (a_1, b_1) of the length $|b_1 - a_1|$. Suppose that we have already proved that R has 2^{n-1} gaps of the length |bn-an| of the form $\left(a_n + \sum_{i=1}^{n-1} \varepsilon_i b_i, b_n + \sum_{i=1}^{n-1} \varepsilon_i b_i\right)$ for $\varepsilon_i = 0, 1$.
Since $a_n/2$ is the middle point of (a_{n+1}, b_{n+1}), then $a_n = b_{n+1} + a_{n+1}$. Since $a_{n-1}/2$ is the point of reflection of $[0, a_{n-1}] \cap R$, then $(a_{n+1}, b_n, b_{n+1} + b_n)$ is a gap in $[0, a_{n-1}] \cap R$. Now, since $a_{n-2}/2$ is the point of reflection of $[0, a_{n-2}] \cap R$, then $(a_{n+1} + b_{n-1}, b_{n+1} + b_{n-1})$ and $(a_{n+1} + b_n + b_{n-1}, b_{n+1} + b_n + b_{n-1})$ are gap in $[0, a_{n-2}] \cap R$. By a simple induction we obtain that each interval of the form $\left(a_n + \sum_{i=1}^{n-1} \varepsilon_i b_i, b_n + \sum_{i=1}^{n-1} \varepsilon_i b_i\right)$ for $\varepsilon_i = 0, 1$ is a gap in R.
Note that R' is the closure of the endpoints of its gaps. These endpoints belong also to R. Since R is compact, then $R' \subset R$. This shows that R has no other gaps than those described above (each such gap would be a gap of R' as well). Since

$$R' = [0, a_0] \setminus \bigcup \left\{\left(a_n + \sum_{i=1}^{n-1} \varepsilon_i b_i, b_n + \sum_{i=1}^{n-1} \varepsilon_i b_i\right): n \in \mathbb{N}, \varepsilon_i = 0, 1\right\},$$

then $R \subset R'$, and consequently $R = R' = A(b_n)$.
Theorem (6.2.23)[253]: Assume that $R = rng(\mu)$ and there is $\varepsilon > 0$ such that between any two gaps of the same length smaller than ε there is a longer gap in R. Then R is a Cantor set.
Proof. There exists a sequence of gaps $\{(a_n, b_n): n \in \mathbb{N}\}$ of R such that
(i) (a_1, b_1) is the longest gap from the left of the length $b_1 - a_1 < \varepsilon$;
(ii) $|b_{n+1} - a_{n+1}| < |b_n - a_n|$ for every $n \in \mathbb{N}$;
(iii) (a_{n+1}, b_{n+1}) is the longest gap in $R \cap [0, a_n]$ and any other gap in $R \cap [0, a_n]$ is shorter for every $n \in \mathbb{N}$.
By Lemma (6.2.6) $b_1 = \mu(\{m\})$ and $a_1 = \mu(\{m+1, m+2, \ldots\})$. Let $F_m = \{\mu(E): E \subset \{1, 2, \ldots, m\}\}$. Then $rng(\mu) = F_m + rng(\mu_1)$ where μ_1 is a measure on $\{m+1, m+2, \ldots\}$ given by $\mu_1(E) = \mu(E)$ for $E \subset \{m+1, m+2, \ldots\}$. By Theorem (6.2.21) $rng(\mu_1)$ is a Cantor set. Thus $rng(\mu)$ is a Cantor set as well as a union of finitely many shifts of a Cantor set $rng(\mu_1)$.
Immediately by Theorem (6.2.23) we obtain the necessary condition for a measure range to be a Cantorval.
Corollary (6.2.24)[253]: If $rng(\mu)$ is a Cantorval, then there are infinitely many pairs of gaps in $rng(\mu)$ of the same length which are not separated by a longer gap.
Let us consider the sequences of the form

$$x_{km+1} = k_1 q^k, x_{km+2} = k_2 q^k, \ldots, x_{km+m} = k_m q^k$$

for $k = 0, 1, 2, \ldots$. Such a sequence we call multigeometric of the rank m and denote by $(k_1, k_2, \ldots, k_m; q)$. As we have mentioned in the Introduction, almost all known examples of sequences whose achievement sets are Cantorvals belong to this class, see [158], [144], [254], [143] and [149]. Let us observe that the Guthrie-Nymann Cantorval T (described in Theorem (6.2.3)) is an achievement set of the bigeometric sequence $\left(\frac{3}{4}, \frac{2}{4}; \frac{1}{4}\right) = \left(\frac{3}{4}, \frac{2}{4}, \frac{3}{16}, \frac{2}{16}, \ldots\right)$. It is not difficult to see that the achievement set $A(k_1, \ldots, k_m; q)$ is equal to the set $\{\sum_{n=1}^{\infty} \delta_n q^{n-1}: (\delta_n) \in \Sigma^N\}$ where $\Sigma = \{\sum_{i=1}^{m} \varepsilon_i k_i: (\varepsilon_i) \in \{0,1\}^m\}$.

Consequently, $A(k_1,\dots,k_m;q)$ is an attractor for the iterated function system, in short IFS, consisting of the affine functions of the form $f_\sigma(x)=qx+\delta$ where $\delta\in\Sigma$, and therefore it is the unique set $A=A(\Sigma,q)$ satisfying the equality $A=\Sigma+qA$. Not all attractors of affine IFS's are achievement sets of sequences (or ranges of purely atomic measures). Let us observe that if $A=A(\Sigma,q)=A(x_n)$ for some sequence (x_n) of positive terms, then $0\in A$ and $\frac{1}{2}\sum_{n=1}^{\infty} x_n$ is a point of reflection of A. Hence $0\in\Sigma$ and Σ is symmetric as well. It turns out that these two conditions for Σ are not sufficient. Recently [255] showed that the Cantorval N related to the construction of the ternary Cantor set is not an achievement set of any sequence but it is an attractor of some affine IFS.

We use the multigeometric sequences to show that there are Cantor sets as well as Cantorvals which can be defined by continuum many different sequences.

Example (6.2.25)[253]: Consider the Jones-Velleman sequence $(x_n)=(4,3,2;q)$, defined as follows $x_{3n-2}=4q^{n-1}, x_{3n-1}=3q^{n-1}, x_{3n}=2q^{n-1}$ and its modification $(y_n)=(3,2,2,2;q)$, defined as follows $y_{4n-3}=3q^{n-1}, x_{4n-2}=2q^{n-1}$, $x_{4n-1}=2q^{n-1}, x_{4n}=2q^{n-1}$, where $q\in(0,1)$. For more details see [149], where the author considered among others the sequence (x_n) with $q=\frac{1}{5}$. Let us observe that the given modification does not change the achievement set and we have $A(x_n)=A(y_n)$ (compare the proof of Lemma (6.2.8)). We define a family of sequences F as a family of all sequences (z_n) which are constructed as follows:

(a) in each step we define three or four succeeding elements of (z_n)

(b) in n-th step we define $z_{k_{n-1}+i}=x_{3n-3+i}$ for $i\in\{1,2,3\}$ or $z_{k_{n-1}+i}=y_{4n-4+i}$ for $i\in\{1,2,3,4\}$ if we have decided to define three or four elements respectively, where k_{n-1} is the number of defined elements in the first $n-1$ steps, $k_0=0$

Then $A(z_n)=A(x_n)$ for each sequence (z_n) which belongs to F. Moreover, if we have two sequences $(s_n),(w_n)\in F$ then $w_n=s_n$ for each $n\in\mathbb{N}$ if and only if in each step of constructions of (s_n) and (w_n) we define the same numbers of elements. Hence the cardinality of F is continuum. We will call the sequences belonging to F as multigeometric-like. It is known that the achievement set $A(x_n)$ for some q can be an interval $\left(q\ge\frac{2}{11}\right)$, a Cantor set with Lebesgue measure zero $\left(q<\frac{1}{8}\right)$ or a Cantorval $\left(q\in\left[\frac{1}{6},\frac{2}{11}\right)\right)$. For more details see [149] and [143].

For the next theorem the fact, proved by Bielas, Plewik and Walczyńska in [257], that the Guthrie–Nymann– Cantorval's center of distances consists exactly of the terms of its generating sequence and zero will be crucial.

Theorem (6.2.26)[253]: Let $X=A(x_n)$, where $x_{2n-1}=\frac{3}{4^n}, x_{2n}=\frac{2}{4^n}$. If $X=A(y_n)$ and $y_1\ge y_2\ge y_3\ge\cdots$, then $y_n=x_n$.

Proof. First note that $\{y_n: n\in\mathbb{N}\}\subset\{x_n: n\in\mathbb{N}\}$. Take any $k\in\mathbb{N}$. By Lemma (6.2.7) we obtain that $y_k\in S\big(A(y_n)\big)=S\big(A(x_n)\big)$. By the result of Bielas, Plewik and Walczyńska mentioned above, $S\big(A(x_n)\big)=\{x_n: n\in\mathbb{N}\}\cup\{0\}$. Thus $y_k\in\{x_n: n\in\mathbb{N}\}$.

Now we will prove that the set $\{y_n:\in\mathbb{N}\}$ of all terms of (y_n) contains every even term of the basic sequence (x_n). Let $m\in\mathbb{N}$. Observe that $x_{2m}>r_{2m}$. Indeed

$$r_{2m}=\sum_{n=2m+1}^{\infty} x_n=\sum_{n=m+1}^{\infty}\frac{3}{4n}+\sum_{n=m+1}^{\infty}\frac{2}{4n}=\frac{4}{3}\cdot\frac{5}{4^{m+1}}=\frac{5}{3\cdot 4^m}<\frac{2}{4^m}=x_{2m}.$$

Therefore the interval (r_{2m}, x_{2m}) is a gap in X and by Lemma (6.2.6) we obtain that $x_m \in \{y_n : n \in \mathbb{N}\}$.
We have already proved that $\{x_{2n} : n \in \mathbb{N}\} \subset \{y_n : n \in \mathbb{N}\} \subset \{x_n : n \in \mathbb{N}\}$. Since the sequence (x_n) is one-to-one and $S(X) = \{x_n : n \in \mathbb{N}\} \cup \{0\}$, then by Lemma (6.2.8) none term of (y_n) can be repeated more than two times. However, if $\{y_n : n \in \mathbb{N}\} \neq \{x_n : n \in \mathbb{N}\}$, then some terms of (y_n) must be repeated. This easily follows from the equality $\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty} y_n$.
Now we are ready to prove the assertion. Let us start from the first step of the inductive prove. Since (y_n) is non-increasing, then y_1 equals x_1 or x_2 (all even terms of (x_n) are among terms of (y_n)). Suppose that $y_1 \neq x_1 = \frac{3}{4}$. Since every term of (y_n) can be repeated at most two times, we have the following inequality

$$\sum_{n=1}^{\infty} y_n \leq 2 \cdot \sum_{n=2}^{\infty} x_n = \frac{11}{6}.$$

Moreover $\frac{5}{3} = \max X$ and $\frac{11}{6} - \frac{5}{3} = \frac{1}{6}$, which means that to obtain (y_n) from the sequence $(x_2, x_2, x_3, x_3, x_4, x_4, \ldots)$ we need to remove elements which sum equals precisely $\frac{1}{6}$. Since $\frac{1}{2}$ and $\frac{3}{16}$ are greater than $\frac{1}{6}$, then $y_1 = y_2 = \frac{1}{2}, y_3 = y_4 = \frac{3}{16}$. Note that $y_5 = x_4 = \frac{1}{8}$ because we have to use all even terms of (x_n). Observe that $y_6 \neq \frac{1}{8}$. Indeed, if $y_6 = \frac{1}{8}$ then $y_3 + y_5 + y_6 = \frac{7}{16} \in \left(\frac{5}{12}, \frac{1}{2}\right)$ but $(r_2, x_2) = \left(\frac{5}{12}, \frac{1}{2}\right)$ is a gap in X. Moreover $y_6 \neq x5 = \frac{3}{64}$. Indeed, if $y_6 = \frac{3}{64}$ then $y_3 + y_4 + y_6 = \frac{27}{64} \in \left(\frac{5}{12}, \frac{1}{2}\right)$. It means that we need to remove one element x_4 and two elements x_5 from the sequence $(x_2, x_2, x_3, x_3, x_4, x_4, \ldots)$. But

$$x_4 + 2x_5 = \frac{1}{8} + \frac{6}{64} = \frac{14}{64} = \frac{42}{192} > \frac{32}{192} = \frac{1}{6}$$

which yields a contradiction. Thus $y_1 = x_1$.
Now assume that $y_i = x_i$ for each $i \in \{1, \ldots, 2m-1\}$ for some $m \in \mathbb{N}$. We will show that $y_{2m} = x_{2m}$ and $y_{2m+1} = x_{2m+1}$. If $y_{2m} \neq x_{2m}$ then $y_{2m} = x_{2m-1}$ and $y_{2m+1} = x_{2m}$. Hence $\sum_{k=1}^{2m+1} y_k = \sum_{k=1}^{2m} x_k + x_{2m-1} > \sum_{k=1}^{2m} x_k + r_{2m} = \frac{5}{3}$, which brings a contradiction. Therefore $y_{2m} = x_{2m}$. Suppose that $y + (2m+1) \neq x_{2m+1}$. Observe that $(A(4^m x_n)_{n=2m+1}^{\infty}) = A(x_n)$. Moreover, if $A(x_n) = A(y_n)$ and $y_i = x_i$ for each $i \in \{1, \ldots, 2m\}$ then $A((x_n)_{n=2m+1}^{\infty}) = A((y_n)_{n=2m+1}^{\infty})$. Thus $A(x_n) = 4^m A((y_n)_{n=2m+1}^{\infty}) = A((4^m y_n)_{n=2m+1}^{\infty})$. By the first step of induction we obtain that $4^m y_{2m+1} = x_1 = \frac{3}{4}$. Thus $y_{2m+1} = x_1/4^m = x_{2m+1}$. This ends the inductive proof.

Let us consider $A = A(\Sigma;\, q) = \{\sum_{n=1}^{\infty} x_n q^{n-1} : (x_n) \in \Sigma^N\}$, where Σ is a finite set. We have $\Sigma + qA = A$ which means that A is the attractor of the affine IFS system $\{f_\sigma\}_{\sigma \in \Sigma}$, where $f_\sigma(x) = qx + \sigma$. We also call the set A a fractal - it is more general than the theory of multigeometric sequences, because Σ does not have to be the achievement set of any finite sequence. The important class of attractors are so called Ferens fractals for which $\Sigma = \{0, p, p+1, \ldots, p+r, 2p+r\}$ for some $p, r \in \mathbb{N}, p \geq 2$. It is known that for $q \geq \frac{p}{3p+r}$ the set $A(\Sigma; q)$ is an interval and for $q < \frac{p}{3p+r}$ the set $A(\Sigma; q)$ is not a union of closed intervals, in particular for $q < \frac{1}{|\Sigma|} = \frac{1}{r+3}$ it is a null Cantor set, see [158], [246] and [143].

Theorem (6.2.27)[253]: Let $p \in \mathbb{N}, p \geq 2$. The Ferens fractal $A = A(\Sigma; q)$ for $r = p$, that is $\Sigma = \{0, p, p+1, \dots, p+r, 2p+r\} = \{0, p, p+1, \dots, 2p, 3p\}$ and $q < \frac{1}{4}$ cannot be obtained as an achievement set for any sequence.

Proof. Note that $(a, b) = \left(\frac{3pq}{1-q}, p\right)$ is the longest gap in A from the left. By Lemma (6.2.6) and the properties of the center of distances we get $p \in S(A)$. We consider the gaps (a, b) and $(2p + a, 2p + b)$. Firstly assume that $q \in \left(0, \frac{p-1}{4p-1}\right)$, which is equivalent to $a < p - 1$. Fix $x = 2p - 1 \in A$. Then $x + p = 3p - 1 \in (2p + a, 2p + b)$ and $x - p = p - 1 \in (a, b)$. Hence $p \notin S(A)$. Now assume that $q \in \left(\frac{p-1}{4p-1}, \frac{1}{4}\right)$. Then $p - 1 < a < p$. Since (a, b) is the longest gap in $A \cap [0, b)$ one can find $y \in (1 + a - b, 1) \cap A$. Fix $x = 2p - 1 + y \in A$. Then $x + p = 3p - 1 + y$ and $2p + a = 3p + a - b < 3p - 1 + y < 3p = 2p + b$, so $x + p \in (2p + a, 2p + b)$. Analogously we prove $x - p \in (a, b)$. Hence $p \notin S(A)$.
If $q = \frac{p-1}{4p-1}$ then $(a, b) = (p - 1, p)$ and we take any $z \in (0, 1) \cap A$ and then define $x = 2p - 1 + z$. Thus $x - p \in (a, b)$ and $x + p \in (2p + a, 2p + b)$. The proof is finished.
On the other hand there exist $p \in \mathbb{N}, p \geq 2$ and $r \neq p$ such that the Ferens fractal $A = A(\Sigma;\ q)$ with $\Sigma = \{0, p, p+1, \dots, p+r, 2p+r\}$ is obtained as an achievement set for each $q \in (0, 1)$. Let us consider the following examples.

Example (6.2.28)[253]: Let us consider the Ferens fractal $A = A(\Sigma;\ q)$ for $\Sigma = \{0, 2, 3, 4, 6\}$. It is known that for $q \geq \frac{1}{4}$ the set A is the interval. By Theorem (6.2.27) the set A for $q < \frac{1}{4}$ cannot be obtained as an achievement set for any sequence.

Example (6.2.29)[253]: Let $\Sigma = \{0, 2, 3, 5\}$. Here we have $r = 1 < 2 = p$. Then $A = A(x_n)$ for the multigeometric sequence $x_{2n+1} = 3q^n, x_{2n+2} = 2q^n$ for $n = 0, 1, 2, \dots$. In particular for $q = \frac{1}{4}$ we get rescaled by 4 Guthrie and Nymann's Cantorval. It is also the Ferens fractal for $= 2, r = 1, q = \frac{1}{4}$. Note that for each $p \in \mathbb{N}$ and $r = 1$ we obtain a Ferens fractal, which can be obtained by the multigeometric sequence.

Example (6.2.30)[253]: Let $\Sigma = \{0, 2, 3, 4, 5, 7\}$. Here we have $r = 3 > 2 = p$. Then $A = A(x_n)$ for a multigeometric sequence $x_{3n+1} = 3q^n, x_{3n+2} = 2q^n, x_{3n+3} = 2q^n$ for $n = 0, 1, 2, \dots$.
So, there are Ferens fractals which are also achievement sets. The next theorem gives the example of large class of such fractals and shows that for each natural $p \geq 2$ we can find r such that the Ferens fractal $A = A(\Sigma; q)$ is also an achievement set. We will base our calculation on a simple observation that if Σ is the achievement set of a finite sequence $\{a_1, \dots, a_k\}$ then $A(\Sigma; q)$ can be obtained by the multigeometric sequence (x_n) defined as follows $x_{kn+j} = a_j q^n$ for $n \in \mathbb{N} \cup \{0\}$ and $j \in \{1, \dots, k\}$.

Lemma (6.2.31)[253]: Let $p \in \mathbb{N}, p \geq 2, r = \frac{3p^2 - 3p}{2}$. Then $\Sigma = \{0, p, p+1, \dots, p+r, 2p+r\} = \left\{0, p, p+1, \dots, \frac{3p^2-p}{2}, \frac{3p^2+p}{2}\right\}$ is the set of subsums for some finite sequence.

Proof. Define $a_1 = p, a_j = (p + j - 2)$ for $j \in \{2, \dots, p+1\}$. Then $\Sigma = A\left((a_n)_{n=1}^{p+1}\right)$.
As a result we immediately obtain:

Theorem (6.2.32)[253]: Let $p \in \mathbb{N}, p \geq 2, q \in (0, 1)$. The Ferens fractal $A = A(\Sigma; q)$ for $r = \frac{3p^2-3p}{2}$ (so $\Sigma = \left\{0, p, p+1, \dots, \frac{3p^2-p}{2}, \frac{3p^2+p}{2}\right\}$) is an achievement set for some multigeometric sequence.

Proof. Define $x_{(p+1)n+1} = pq^n, x_{(p+1)n+j} = (p + j - 2)q^n$ for $n \in \mathbb{N} \cup \{0\}, j \in \{2, \dots, p + 1\}$. Then $A = A(x_n)$.

Lemma (6.2.33)[253]: Let $p \in \mathbb{N}, p \geq 2, r \geq \frac{3p^2-p}{2}$. Then $\Sigma = \{0, p, p + 1, \dots, p + r, 2p + r\}$ is the set of subsums for some finite sequence.

Proof. Let us first consider $r = \frac{3p^2-p}{2}$. Define $a_j = (p + j - 1)$ for $j \in \{1, \dots, p + 1\}$. Then $\Sigma = A\left((a_n)_{n=1}^{p+1}\right)$. Let now consider $r > \frac{3p^2-p}{2}, r = \frac{3p^2-p}{2} + k$, where $k = mp + r$ for $m \in \mathbb{N} \cup \{0\}, r \in \{0, 1, \dots, p - 1\}$. Define $a_j = p$ for $j \in \{1, \dots, 2 + m\}, a_j = (p + j - m - 2)$ for $j \in \{3 + m, \dots, 2 + m + k\}, a_j = (p + j - m - 3)$ for $j \in \{3 + m + k, \dots, 2 + m + p\}$. Then $\Sigma = A\left((a_n)_{n=1}^{2+m+p}\right)$.

Corollary (6.2.34)[253]: Let $p \in \mathbb{N}, p \geq 2, q \in (0, 1)$. The Ferens fractal $A = A(\Sigma; q)$ for $r \geq \frac{3p2-p}{2}$ is an achievement set for some multigeometric sequence.

Lemma (6.2.35)[253]: Let $p \in \mathbb{N}, p \geq 3, r \in \left(1, \frac{3p^2-3p}{2}\right)$. Then $\Sigma = \{0, p, p + 1, \dots, p + r, 2p + r\}$ is not the set of subsums for any finite sequence.

Proof. Let $r \geq p$. Assume that $\Sigma = A(x_n)$ for some finite sequence (x_n). Since p is the smallest non-zero element, we know that the smallest sum of two or more elements equals 2p. We know that $p + r \geq 2p$. Therefore we get $\{p, p + 1, \dots, 2p - 1\} \subset (x_n)$. Since $2p \in \Sigma$ we have to add the another term x_n equal to 2p or one more term x_n equal to p. Thus its sum is an element of Σ, but $p + (p + 1) + \dots + 2p \geq p + p + (p + 1) + \dots 2p - 1 = \frac{3p^2+p}{2} = 2p + \frac{3p^2-3p}{2} > 2p + r = \max \Sigma$. We get contraditions for both cases. Let $r \in (1, \dots, p)$. Since $p, p + 1, p + 2 \in \Sigma$ and $p + 2 < 2p$ we get $p, p + 1, p + 2 \in (x_n)$. Therefore we have $3p + 3 \in \Sigma$, but $\max \Sigma = 2p + r < 3p < 3p + 3$, which gives us a contradition.

Lemma (6.2.36)[253]: Let $p \in \mathbb{N}, p \geq 2, r \in \left(\frac{3p^2-3p}{2}, \frac{3p^2-p}{2}\right)$. Then $\Sigma = \{0, p, p + 1, \dots, p + r, 2p + r\}$ is not the set of subsums for any finite sequence.

Proof. Note that $2p + r > \frac{3p^2+p}{2} > p + (p + 1) + \dots + (2p - 1)$. Hence $2p \in \Sigma$. We can obtain it by adding 2p or one more p to our terms. If we add 2p then $p + (p + 1) + \dots + (2p - 1) + 2p = 3p2 + 3p\ 2 > 2p + r$, which yields a contradition. So let us consider $(x\_n) = \{p, p, p + 1, \dots, 2p - 1\}$. We have $\sum x_n = \frac{3p^2+p}{2} \in (p + r, 2p + r)$. Since $\sum x_n < 2p + r$ we have to add next element to the sequence (x_n), but we cannot add an element which is smaller than p. Therefore $\sum x_n + p > 2p + r$, which yields a contradition.

Corollary (6.2.37)[253]: Let $p, r \in \mathbb{N}, q$ be a positive real number and $\Sigma = \{0, p, p + 1, \dots, p + r, 2p + r\}$.

(i) If $p \geq 3, q \in \left(0, \frac{p}{3p+r}\right)$ and $r \in \left(1, \frac{3p^2-3p}{2}\right) \cup \left(\frac{3p^2-3p}{2}, \frac{3p^2-p}{2}\right)$, then the Ferens fractal $A = A(\Sigma; q)$ is not an achievement set for the multigeometric sequence $(k_1, \dots, k_m; q)$ with $\{\sum_{i=1}^{m} \varepsilon_i k_i : (\varepsilon_i) \in \{0, 1\}^m\} = \Sigma$.

(ii) If $p = 2, q \in \left(0, \frac{p}{3p+r}\right)$ and $r \in \left(\frac{3p^2-3p}{2}, \frac{3p^2-p}{2}\right)$, then the Ferens fractal $A = A(\Sigma; q)$ is not an achievement set for the multigeometric sequence $(k_1, \dots, k_m; q)$ with $\{\sum_{i=1}^{m} \varepsilon_i k_i : (\varepsilon_i) \in \{0, 1\}^m\} = \Sigma$.

We will deal with Ferens fractals of the type $A(r) = A(\Sigma, q)$ for $\Sigma = \{0, 2, 3, \dots, r+2, r+4\}$ and $= \frac{1}{|\Sigma|} = \frac{1}{r+3}$. It is known that sets A(r) for $r = 1, 2, \dots$ are Cantorvals. It follows from Kenyon Theorem, (see [153] and [164]) which states that if $\{n \bmod r : n \in \Sigma\} = \mathbb{Z}_r$, then $A(\Sigma, 1/r)$ has nonempty interior (it can be also deduced from proofs presented in [143]).

Note that

(i) for $r = 1$ the set $A(r)$ is the rescaled Guthrie–Nymann Cantorval which, by Theorem (6.2.26) has the unique representation as an achievement set.

(ii) For $r = 2m - 1$ the set A(r) equals to $A(x_n)$ where $(x_n) = \left(3, \underbrace{2, \dots, 2}_{m}; \frac{1}{r+3}\right)$. If $r \geq 5$, the Cantorval $A(r)$ has continuum many representations as an achievement set of multigeometric-like series with the same set Σ – see Example (6.2.25).

(iii) By Theorem (6.2.27) the set A(r) is not an achievement set (or a range of any measure) for $r = 2$.

(iv) For r=4 we know that A(r) is not an achievement set for any multigeometric series generating the same set Σ.

(v) For $r = 2m \geq 6$ the set A(r) equals to $A(x_n)$ where $(x_n) = \left(3, 3, \underbrace{2, \dots, 2}_{m-2}; \frac{1}{r+3}\right)$. Using the method from Example (6.2.25) for r≥10 (or $m - 2 \geq 3$) we observe that the Cantorval A(r) has continuum many representations as an achievement set of multigeometric-like series with the same set Σ.

Using methods from [257] one can get some information on geometry and the center of distances for Cantorvals A(r):

(i) $A(r) \subset \left[0, \frac{(r+4)(r+3)}{r+2}\right]$.

(ii) The interval $\left[\frac{2(r+3)}{r+2}, r+3\right]$ is the longest component of A(r).

(iii) $\left[0, \frac{r+4}{r+2}\right] \cap A(r) = \frac{1}{r+3} A(r)$.

(iv) $\left(\frac{r+4}{r+2}, 2\right)$ is the longest gap from the left and it has the same length as the longest component of $\left[0, \frac{4+r}{(2+r)(3+r)}\right] \cap A(r)$.

(v) $r + 3 + \left(A(r) \cap \left[0, \frac{2(r+3)}{r+2}\right]\right) \cup \left(A(r) \cap \left[r+3, \frac{(r+4)(r+3)}{r+2}\right]\right) = \left[r+3, \frac{(r+4)(r+3)}{r+2}\right]$, it follows from the fact that the gaps of the first summand in the above union are exactly in the same places as the copmonents of the second one and vice versa.

In Example (6.2.38) we present the idea of proving (i)–(v) based on an appropriate picture. Note that if $t \in \left(\frac{2(r+3)}{r+2}, \frac{r+3}{2} - \frac{r+3}{r+2}\right)$, then $t \in S\big(A(r)\big)$. Recall that $\frac{2(r+3)}{r+2}$ is a left endpoint of the longest component of A(r) and $\frac{r+3}{2} - \frac{r+3}{r+2}$ is a half of its length. Similarly we have for every longest component of A(r) from the left. Therefore if $\frac{2}{2+r} < \frac{1}{2} - \frac{1}{2+r}$, that is if $r > 4$, then $S\big(A(r)\big)$ contains a sequence of intervals.

This observation suggests that for $r > 4$ one can look for a multigeometric series (x_n') with $\Sigma' \neq \Sigma$ and $A(x_n') = A(x_n)$.

Example (6.2.38)[253]: At Figure 1 we present a GNJ Cantorval $A := A(6)$, i.e. $\Sigma = \{0, 2, 3, \dots, 8, 10\}$ and $q = \frac{1}{9}$; there are also nine its copies $\tau + \frac{1}{9}A, \tau \in \Sigma$. The first and the last copies, $\frac{1}{9}A$ and $10+\frac{1}{9}A$, are equal to the left $A\cap\left[0,1\frac{2}{8}\right]$ and the right $A\cap\left[10,11\frac{2}{8}\right]$ parts of the original Cantorval A, respectively. Other copies cover the rest of A; note that $2+\frac{1}{9}A$ and $3+\frac{1}{9}A$ cover the interval $\left[3,3\frac{2}{8}\right]$, since the components interiors of $(2+\frac{1}{9}A)\cap\left[3,3\frac{2}{8}\right]$ are precisely gaps of $(3+\frac{1}{9}A)\cap\left[3,3\frac{2}{8}\right]$, and vice versa.

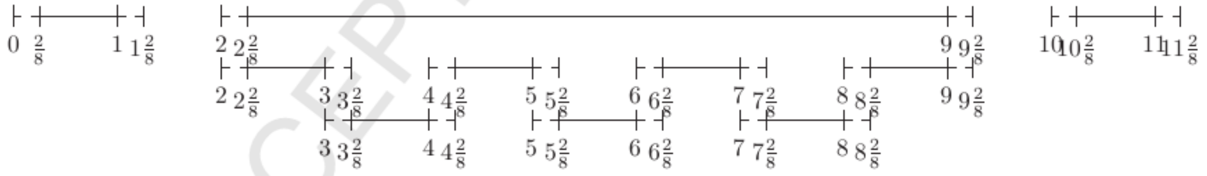


Figure (1)[253]:

On the other hand $A(6)$ satisfies also the equality $A=\Sigma'+\frac{1}{9}A$ for $\Sigma'=\left\{0,2,2\frac{2}{8},2\frac{4}{8},3\frac{2}{8},4\frac{2}{8},4\frac{4}{8},4\frac{6}{8},5\frac{2}{8},5\frac{4}{8},5\frac{6}{8},6\frac{6}{8},7\frac{4}{8},7\frac{6}{8},8,10\right\}$. Let us observe that Σ' is an achievement set for the finite sequence $\left\{3\frac{2}{8},2\frac{4}{8},2\frac{2}{8},2\right\}$ and hence $A(6)=A\left(3\frac{2}{8},2\frac{4}{8},2\frac{2}{8},2;\frac{1}{9}\right)$ as well as $A(6)=A\left(3,3,2,2;\frac{1}{9}\right)$. For the clarity and readers' convenience we present the next picture - Figure 2.

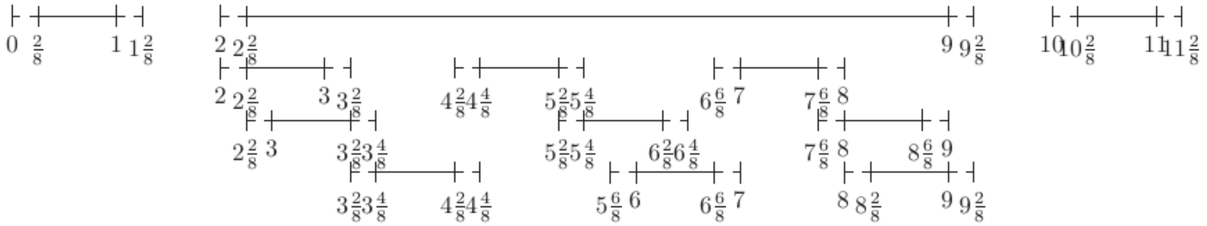


Figure (2)[253]:

Example (6.2.39)[253]: In [257] the authors found a center of distances of the boundary $\partial\frac{A(1)}{4}$ of the Guthrie–Nymann Cantorval $\frac{1}{4}A(1)$. The set $\partial\frac{A(1)}{4}$ is a Cantor set arisen from $\frac{A(1)}{4}$ by removing all interiors of its nontrivial components. It turns out that $S\left(\partial\frac{A(1)}{4}\right)=\left\{1,\frac{1}{4},\frac{1}{4^2},\dots\right\}$. Therefore if $\partial\frac{A(1)}{4}=A(y_n)$ for some sequence (y_n), then $\{y_n\colon n\in\mathbb{N}\}\subset\left\{1,\frac{1}{4},\frac{1}{4^2},\dots\right\}$. The authors, according to this observation, claimed that $\partial\frac{A(1)}{4}$ is not an achievement set for any sequence, since $1+\frac{1}{4}+\frac{1}{4^2}+\dots<\frac{5}{3}=\max\frac{A(1)}{4}$. However, they did not observe that terms of (y_n) may repeat. By Lemma (6.2.8) none of the terms may repeat more than twice, since the doubling of such term would be in $S\left(\partial\frac{A(1)}{4}\right)$. But $1+\frac{1}{4}+\frac{1}{4}+\frac{1}{4^2}+\frac{1}{4^2}+\dots=\frac{5}{3}$. It turns out that for any positive integer r

$$\partial\frac{A(r)}{r+3}=A\left(1,\frac{1}{r+3},\frac{1}{r+3},\frac{1}{(r+3)^2},\frac{1}{(r+3)^2},\dots\right).$$

Indeed, by geometric properties of A(r) it follows that

$$\left[0,\frac{2}{2+r}\right]\cap\partial\frac{A(r)}{r+3}=C_{\frac{1}{r+3}}+C_{\frac{1}{r+3}}.$$

Thus

$$\partial\frac{A(r)}{r+3}=\left(C_{\frac{1}{r+3}}+C_{\frac{1}{r+3}}\right)\cup\left(1+C_{\frac{1}{r+3}}+C_{\frac{1}{r+3}}\right).$$

Section (6.3): The Lebesgue Measure

The investigation of topological properties of a set of subsums of absolutely convergent series $\sum x_n$ (i.e. the set of numbers which are sums of subsequences of (x_n)) has been initiated over one hundred years ago by Soichi Kakeya (see [150]). Many years later, J.A. Guthrie and J.E. Nymann (see [147]) presented the full topological classification of the sets of subsums. They proved that a set of subsums must be (up to a homeomorphism) either a finite set or the union of a finite family of closed intervals or a Cantor set or a set $C \cup \bigcup_{n=1}^{\infty} G_{2n-1}$, where C is the classic Cantor ternary set and G_k denotes the union of all open intervals removed from [0, 1] in the kth step of the standard construction. The last sets are called M-Cantorvals (some characterization of this set can be found in [152], [251]). In [146], [147], [149] and [252] examples of series having an M-Cantorval as the set of their subsums are presented.

Consider a multigeometric series (see [143], [164])

$$3q + \underbrace{2q + 2q + \ldots + 2q + 2q}_{m} + \underbrace{3q^2 + 2q^2 + 2q^2 + \ldots + 2q^2 + 2q^2}_{m} + \ldots + 3q^n + \underbrace{2q^n + 2q^n + \ldots + 2q^n + 2q^n}_{m} + \ldots, \quad (7)$$

where $q \in (0,1)$. In [158] it is shown that the set of subsums of the series (7) is a partial case of a set of the form

$$K(m;q) = \left\{ \sum_{n=1}^{\infty} \epsilon_n q^n : (n) \in \{0,2,3,\ldots,2m,2m+1,2m+3\}^N \right\}, \quad (8)$$

where $m \in \mathbb{N}$ and $q \in (0,1)$.

Moreover, based on analysis in[158], the following conclusions can be drawn:

(i) $K\left(m;\frac{1}{2m+2}\right)$ is an M-Cantorval,

(ii) if $q < \frac{1}{2m+2}$ then the set $K(m;q)$ is a Cantor set of zero Lebesgue measure,

(iii) for each $q \in \left(\frac{1}{2m+2},1\right)$ there exist $q_1, q_2 \in \left(\frac{1}{2m+2},q\right)$ such that the set $K(m;q_1)$ is a null set with respect to Lebesgue measure and $K(m;q_2)$ has positive Lebesgue measure.

For these reasons, we decided to investigate the set $\mathcal{K} := K\left(m;\frac{1}{2m+2}\right)$. Using the fact that $\mathcal{K}$ is an attractor of some iterated function system we describe a construction of this set. We compute the Lebesgue measure of this special M-Cantorval and we show that it is equal to the sum of lengths of all its component intervals.

It is worth mentioning that lately an interesting [253] has been published by A. Bartoszewicz, S. Głab, J. Marchwicki. Part of their article is devoted to research the sets of the form (8).

We will present a theorem describes a construction of the set $\mathcal{K}$. Let's start with short introduction.

Let X be a complete metric space with a finite collection contraction maps $\{w_i : i = 1,2,\ldots N\}$ from X to itself. The set $\{X; w_i : i = 1,2,\ldots,N\}$ is called an iterated function system (IFS). A nonempty compact subset A of X is an attractor of the IFS if $A = \bigcup_{i=1}^{N} w_i(A)$ (see [268], [83]).

Let $\mathcal{H}(\mathbb{R})$ be the set of all nonempty compact subsets of $\mathbb{R}$ and for an element $B \in \mathcal{H}(\mathbb{R})$ let

$$W(B)=\bigcup_{i=1}^{2m+2} w_i(B),$$
where $m \in \mathbb{N}$ and contraction mappings $w_i: \mathbb{R} \to \mathbb{R}$ are defined by

(i) $w_1(x)=\frac{x}{2m+2}$,

(ii) $w_i(x)=\frac{i}{2m+2}+\frac{x}{2m+2}$ for $i \in \{2,\dots,2m+1\}$,

(iii) $w_{2m+2}(x)=\frac{2m+3}{2m+2}+\frac{x}{2m+2}$.

It is easy to check that $W(\mathcal{K})=\mathcal{K}$. Therefore $\mathcal{K}$ is an attractor of $\{\mathbb{R}; w_i: i=1,\dots,2m+2\}$.

We write $A \sqcup B$ for the disjoint union of the sets A and B and $\sqcup_{k=0}^{n} A_k$ for the disjoint union of the sets $A_0, A_1, \dots, A_n$.

The set $W^n(I)$ is called nth iteration of $\mathcal{K}$. Note that $W^n(I)$ is the disjoint union of smaller copies of previous iterations and the interval with endpoints $w_2^n(0)$ and $w_{2m+1}^n\left(\frac{2m+3}{2m+1}\right)$.

Before getting to the proof of the Theorem (6.3.3) we show two lemmas. We start by introducing a few notations which will ease the presentation of these lemmas.

Let $C_n, n \in \mathbb{N}$, be the set of all finite sequences $(i_k)_{k=1}^{n}$ of elements of $\{1,2,3,\dots,2m+1,2m+2\}, m \in \mathbb{N}$. Sequence of C_n is denoted by $i_1 i_2 \dots i_n$. We will also write $i_1^{[k]} i_{k+1 \dots i_n}$ if $i_1 = i_2 = \dots = i_k$ for some $k \le n$. Now, we define a relation $\preceq$ on C_n as follows $i_1 i_2 \dots i_n \preceq j_1 j_2 \dots j_n$ if and only if $i_k = j_k$, for $1 \le k \le n$ or $i_1 i_2 \dots i_n \prec j_1 j_2 \dots j_n$ where

$$i_1 i_2 \dots i_n \prec j_1 j_2 \dots j_n \Leftrightarrow \text{there exists } N, 1 \le N \le n \text{ such that}$$
$$i_k = j_k \text{ for } 1 \le k < j_N.$$

For example $2\ 3\ 1\ (2m+2)(2m+2) \prec 2\ 3\ (2m+2)\ 1\ 1$.

Let $j_1 \dots j_n \in C_n$ and $\tilde{j}_1 \dots \tilde{j}_n \in C_n$. Denote by $C_n[j_1 \dots j_n; \tilde{j}_1 \dots \tilde{j}_n]$ the set of all $i_1 \dots i_n \in C_n$ such that $j_1 \dots j_n \preceq i_1 \dots i_n \preceq \tilde{j}_1 \dots \tilde{j}_n$.

Finally, let's observe that for our contraction mappings $w_i (1 \le i \le 2m+2)$, we have

a) for any $n \in \mathbb{N}$
$$W^n(I)=\bigcup_{i_1 \dots i_n \in C_n[1^{[n]};(2m+2)^{[n]}]} w_{i_1} \circ \dots \circ w_{i_n}(I), \tag{9}$$

b) for any $n \in \mathbb{N}$ and $i_1 \dots i_n \in C_n$
$$w_{i_1} \circ w_{i_2} \circ \dots \circ w_{i_n}(0)=\frac{\epsilon_1}{2m+2}+\frac{\epsilon_2}{(2m+2)^2}+\dots+\frac{\epsilon_n}{(2m+2)^n},$$
$$\text{where } \epsilon_l=\begin{cases} 0 & if\ i_l=1 \\ i_l & \text{if } i_l \in \{2,3,\dots,(2m),(2m+1)\} \\ 2m+3 & if\ i_l=2m+2 \end{cases} \qquad \mathcal{KH} \preceq\prec (10)$$
for $1 \le l \le n$,

c) for any $n \in \mathbb{N}$ and $i_1 \dots i_n \in C_n$ and $r \in \mathbb{R}$
$$w_{i_1} \circ w_{i_2} \circ \dots \circ w_{i_n}(r)=w_{i_1} \circ w_{i_2} \circ \dots \circ w_{i_n}(0)+\frac{r}{(2m+2)^n}. \tag{11}$$

Lemma (6.3.1)[267]: Let $n, m \in \mathbb{N}$. For each $i_1 \dots i_n \in C_n[2^{[n]};(2m+1)^{[n-1]}(2m)]$ there exists $j_1 \dots j_n \in C_n[2^{[n-1]}3;(2m+1)^{[n]}]$ such that
$$w_{i_1} \circ \dots \circ w_{i_n}(0)+\frac{1}{(2m+2)^n}=w_{j_1} \circ \dots \circ w_{j_n}(0). \tag{12}$$

Proof. We start with simple preliminary observations:

$$\text{if } \; i_1\ldots i_n(2m+1) \in C_{n+1}\left[2^{[n+1]};(2m+1)^{[n]}(2m)\right]$$
$$\text{then} \quad i_1\ldots i_n \in C_n\left[2^{[n]};(2m+1)^{[n-1]}(2m)\right], \tag{13}$$
$$\text{if } i_1\ldots i_k(2m+2)i_{k+2}\ldots i_{n+1} \in C_{n+1}\left[2^{[n+1]};(2m+1)^{[n]}(2m)\right]$$
$$\text{then} \;\; i_1\ldots i_k \in C_k\left[2^{[k]};(2m+1)^{[k-1]}(2m)\right], \text{where } 1 \le k \le n. \tag{14}$$

Now let us prove Lemma (6.3.1) by induction on n. To start with, (12) is valid when $n = 1$ since for $i \in C_1[2;(2m)]$ we have

$$w_i(0) + \frac{1}{2m+2} = w_{i+1}(0), \tag{15}$$

where $i + 1 \in C_1[3;(2m+1)]$.
Next, assume that (12) is valid for all $n \le N$.
Let $i_1 \ldots i_N i_{N+1} \in C_{N+1}\left[2^{[N+1]};(2m+1)^{[N]}(2m)\right]$. There are four cases to be considered regarding the value of i_{N+1}.
Case 1. If $i_{N+1} \in \{2,3,\ldots,(2m)\}$ then

$$w_{i_1}\circ\ldots\circ w_{i_{N+1}}(0) + \frac{1}{(2m+2)^{N+1}} = w_{i_1}\circ\ldots\circ w_{i_N}\left(w_{i_{N+1}}(0)\right) + \frac{1}{(2m+2)^{N+1}}$$
$$\overset{(11)}{=} w_{i_1}\circ\ldots\circ w_{i_N}(0) + \frac{w_{i_{N+1}}(0) + \frac{1}{2m+2}}{(2m+2)^N} \overset{(15)}{=} w_{i_1}\circ\ldots\circ w_{i_N}(0) + \frac{w_{i_{N+1}+1}(0)}{(2m+2)^N} \overset{(11)}{=} w_{i_1}\circ\ldots\circ w_{i_N}\circ w_{i_{N+1}+1}(0),$$

where $i_1\ldots i_N(i_{N+1}+1) \in C_{N+1}\left[2^{[N]}3;(2m+1)^{[N+1]}\right]$.
Case 2. If $i_{N+1} = 2m+1$ then

$$w_{i_1}\circ\ldots\circ w_{i_N}\circ w_{(2m+1)}(0) + \frac{1}{(2m+2)^{N+1}} \underset{(11)}{\overset{(10)}{=}} w_{i_1}\circ\ldots\circ w_{i_N}(0) + \frac{1}{(2m+2)^N}$$

(from (13) we conclude that $i_1\ldots i_N \in C_N\left[2^{[N]};(2m+1)^{[N-1]}(2m)\right]$ and thus by our induction hypothesis there exists $j_1\ldots j_N \in C_N\left[2^{[N-1]}3;(2m+1)^{[N]}\right]$ such that (12) holds)

$$= w_{j_1}\circ\ldots\circ w_{j_N}(0) \underset{(11)}{\overset{(10)}{=}} w_{j_1}\circ\ldots\circ w_{j_N}\circ w_1(0)$$

and $j_1\ldots j_N 1 \in C_{N+1}\left[2^{[N]}3;(2m+1)^{[N+1]}\right]$.
Case 3. If $i_{N+1} = 2m+2$ then

$$w_{i_1}\circ\ldots\circ w_{i_N}\circ w_{(2m+2)}(0) + \frac{1}{(2m+2)^{N+1}} \underset{(11)}{\overset{(10)}{=}} w_{i_1}\circ\ldots\circ w_{i_N}(0) + \frac{1}{(2m+2)^N} + \frac{2}{(2m+2)^{N+1}}$$

(from (14) we conclude that $i_1\ldots i_N \in C_N\left[2^{[N]};(2m+1)^{[N-1]}(2m)\right]$ and thus by our induction hypothesis there exists $j_1\ldots j_N \in C_n\left[2^{[N-1]}3;(2m+1)^{[N]}\right]$ such that (12) holds)

$$= w_{j_1}\circ\ldots\circ w_{j_N}(0) + \frac{2}{(2m+2)^{N+1}} \underset{(11)}{\overset{(10)}{=}} w_{j_1}\circ\ldots\circ w_{j_N}\circ w_2(0)$$

and $j_1\ldots j_N 2 \in C_n\left[2^{[N]}3;(2m+1)^{[N+1]}\right]$.
Case 4. Let $i_{N+1} = 1$. Then since $i_1\ldots i_N 1 \in C_{N+1}\left[2^{[N+1]};(2m+1)^{[N]}(2m)\right]$, there exists $k, 1 \le k \le N$, such that $i_k \in \{3,4,\ldots,(2m+1),(2m+2)\}$ and $i_{k+l} \in \{1,2\}$ for $1 \le l \le N+1-k$. We consider two cases for the value of i_k.
(a) If $i_k \in \{3,4,\ldots,(2m),(2m+1)\}$ then

$$w_{i_1} \circ \ldots \circ w_{i_k} \circ \ldots w_{i_N} \circ w_1(0) + \frac{1}{(2m+2)^{N+1}} \overset{(10)}{=} \frac{\epsilon_1}{2m+2} + \ldots + \frac{\epsilon_k}{(2m+2)^k} + \ldots + \frac{\epsilon_N}{(2m+2)^N} + \frac{1}{(2m+2)^{N+1}}$$

$$= \frac{\epsilon_1}{2m+2} + \ldots + \frac{\epsilon_k - 1}{(2m+2)^k} + \frac{1}{(2m+2)^k} + \ldots + \frac{\epsilon_N}{(2m+2)^N} + \frac{1}{(2m+2)^{N+1}}$$

$$= \frac{\epsilon_1}{2m+2} + \ldots + \frac{\epsilon_k - 1}{(2m+2)^k} + \left(\sum_{l=k+1}^{N} \frac{2m+1}{(2m+2)^l} + \frac{2m+2}{(2m+2)^{N+1}} \right) + \frac{\epsilon_{k+1}}{(2m+2)^{k+1}} + \ldots + \frac{\epsilon_N}{(2m+2)^N} + \frac{1}{(2m+2)^{N+1}} \overset{((6.3.2))}{=}$$

$$= w\_i\_1 \circ \ldots \circ w_{i_{k-1}} \circ w_{i_k} - 1 \circ w_{i_{k+1}+(2m)} \circ \ldots \circ w_{iN+(2m)} \circ w_{(2m+2)}(0).$$

Note that

(i) if $k > 1$ then $i_1 \ldots i_{k-1} \in C_{k-1}\left[2^{[k-1]}; (2m+1)^{[k-1]}\right]$,

(ii) $i_k - 1 \in \{2, 3, \ldots, 2m - 1, 2m\}$,

(iii) if $k < 1$ then $i_{k+l} + 2m \in \{2m+1, 2m+2\}$ for $1 \le l \le N-k$,

which is the desired conclusion.

(b) If $i_k = 2m+2$ then $k > 1$ and we have

$$w_{i_1} \circ \ldots \circ w_{i_{k-1}} \circ w_{2m+2} \circ w_{i_{k+1}} \circ \ldots w_{i_N} \circ w_1(0) + \frac{1}{(2m+2)^{N+1}} \overset{((6.3.2))}{=} \frac{\epsilon_1}{2m+2} + \ldots + \frac{2m+3}{(2m+2)^k} + \ldots + \frac{\epsilon_N}{(2m+2)^N} + \frac{1}{(2m+2)^{N+1}}$$

$$= \frac{\epsilon_1}{2m+2} + \ldots + \frac{\epsilon_{k-1}}{(2m+2)^{k-1}} + \frac{1}{(2m+2)^{k-1}} + \frac{1}{(2m+2)^k} + \frac{\epsilon_{k+1}}{(2m+2)^{k+1}} + \ldots + \frac{\epsilon_N}{(2m+2)^N} + \frac{1}{(2m+2)^{N+1}}$$

$$\overset{(10)}{=} w_{i_1} \circ \ldots \circ w_{i_{k-1}}(0) + \frac{1}{(2m+2)^{k-1}} + \frac{0}{(2m+2)^k} + \sum_{l=k+1}^{N} \frac{2m+1}{(2m+2)^l} + \frac{2m+2}{(2m+2)^{N+1}} + \frac{\epsilon_{k+1}}{(2m+2)^{k+1}} + \ldots + \frac{\epsilon_N}{(2m+2)^N} + \frac{1}{(2m+2)^{N+1}}$$

(from (14) we conclude that $i_1 \ldots i_{k-1} \in C_{k-1}\left[2^{[k-1]}; (2m+1)^{[k-2]}(2m)\right]$ and thus by our induction hypothesis there exists a sequence $j_1 \ldots j_{k-1}$ in the set $C_{k-1}\left[2^{[k-2]}3; (2m+1)^{[k-1]}\right]$ such that (12) holds)

$$\overset{((6.3.2))}{=} w_{j_1} \circ \ldots \circ w_{j_{k-1}} \circ w_1 \circ w_{i_{k+1}+2m} \circ \ldots \circ w_{i_N+2m} \circ w_{2m+2}(0),$$

where if $k < N$ we have $i_{k+l} + 2m \in \{2m+1, 2m+2\}$ for $1 \le k \le N-k$ and the proof is complete.

Lemma (6.3.2)[267]: Let $n, m \in \mathbb{N}$ and $I = \left[0, \frac{2m+3}{2m+1}\right]$. Then

$$\bigcup_{i_1\ldots i_n\in C_n[2^{[n]};(2m+1)^{[n]}]} w_{i_1}\circ\ldots\circ w_{i_n}(I) = \left[w_2^n(0), w_{2m+1}^n\left(\frac{2m+3}{2m+1}\right)\right].$$

Proof. Note first that for $i_1\ldots i_n \in C_n[2^{[n]};(2m+1)^{[n]}]$, we have

$$w_2^n(0) < w_{i_1}\circ\ldots\circ w_{i_n}(0) < w_{2m+1}^n(0). \quad (16)$$

Now by Lemma (6.3.1), for each $i_1\ldots i_n \in C_n[2^{[n]};(2m+1)^{[n-1]}(2m)]$ there exists $j_1\ldots j_n \in C_n[2^{[n-1]}3;(2m+1)^{[n]}]$ such that

$$w_{i_1}\circ\ldots\circ w_{i_n}(0) < w_{i_1}\circ\ldots\circ w_{i_n}(0) + \frac{1}{(2m+2)^n} = w_{j_1}\circ\ldots\circ w_{j_n}(0)$$
$$< w_{i_1}\circ\ldots\circ w_{i_n}(0) + \frac{2m+3}{2m+1}\cdot\frac{1}{(2m+2)^n} \overset{(10)}{=} = w_{i_1}\circ\ldots\circ w_{i_n}\left(\frac{2m+3}{2m+1}\right)$$
$$< w_{j_1}\circ\ldots\circ w_{j_n}\left(\frac{2m+3}{2m+1}\right).$$

Therefore

$$w_{i_1}\circ\ldots\circ w_{i_n}(I) \cup w_{j_1}\circ\ldots\circ w_{j_n}(I) = \left[w_{i_1}\circ\ldots\circ w_{i_n}(0), w_{j_1}\circ\ldots\circ w_{j_n}\left(\frac{2m+3}{2m+1}\right)\right].$$

Hence and by (16), we obtain the assertion of the lemma.

Theorem (6.3.3)[267]: Let $m\in\mathbb{N}$ and $I=\left[0,\frac{2m+3}{2m+1}\right]$. Then

$$\mathcal{K} = \bigcap_{n\in\mathbb{N}} W^n(I),$$

where $W^n(I) = W(W^{n-1}(I))$ and $W^0 = id$. Moreover

$$W^n(I) = \bigsqcup_{k=0}^{n-1} w_2^k\circ w_1\left(W^{n-1-k}(I)\right) \sqcup \left[w_2^n(0), w_{2m+1}^n\left(\frac{2m+3}{2m+1}\right)\right] \sqcup \bigsqcup_{k=0}^{n-1} w_{2m+1}^{n-1-k}$$
$$\circ\, w_{2m+2}\left(W^k(I)\right). \quad (17)$$

Proof. The first part of the Theorem (6.3.3) follows immediately from the fact that $W(I)\subset I$ and the fact that $\lim_{n\to\infty} W^n(I) = \mathcal{K}$ in the Hausdorff metric (see [268]).

Now, for brevity let us denote

$$P_n[j_1\ldots j_n;\tilde{j}_1\ldots\tilde{j}_n] = \bigcup_{i_1\ldots i_n\in C_n[j_1\ldots j_n;\tilde{j}_1\ldots\tilde{j}_n]} w_{i_1}\circ\ldots\circ w_{i_n}(I),$$

where $j_1\ldots j_n, \tilde{j}_1\ldots\tilde{j}_n \in C_n$.

If we prove that

$$W^n(I) = \bigsqcup_{k=0}^{n-1} w_2^k\circ w_1\left(W^{n-1-k}(I)\right) \sqcup P_n[2^{[n]};(2m+1)^{[n]}] \sqcup \bigsqcup_{k=0}^{n-1} w_{2m+1}^{n-1-k}$$
$$\circ\, w_{2m+2}\left(W^k(I)\right), \quad (18)$$

for every positive integers n, then by the Lemma (6.3.2) we conclude that (17) holds and this completes the proof of the Theorem (6.3.3).

Let's prove (18) by induction on n. If $n=1$, then

$$W(I) = \bigcup_{i=1}^{2m+2} w_i(I) = w_1(I) \sqcup P_1[2;(2m+1)] \sqcup w_{2m+2}(I), \quad (19)$$

where the last equality holds because for $i\in\{2,\ldots,(2m+1)\}$ we have

$$w_1\left(\frac{2m+3}{2m+1}\right) < w_i(0) < w_{2m+1}\left(\frac{2m+3}{2m+1}\right) < w_{2m+2}(0).$$

Now, assume that the thesis holds for a positive integer n. Then we have

$$W^{n+1}(I) = W\big(W^n(I)\big) = w_1\big(W^n(I)\big) \cup \ldots \cup w_{2m+2}\big(W^n(I)\big)$$

(by (19) and $W(I) \subset I$)

$$= w_1\big(W^n(I)\big) \sqcup \big[w_2\big(W^n(I)\big) \cup \ldots \cup w_{2m+1}\big(W^n(I)\big)\big] \sqcup w_{2m+2}\big(W^n(I)\big)$$

(observe that by equation (9) we have $w_i\big(W^n(I)\big) = P_{n+1}\big[i1^{[n]}; i(2m+2)^{[n]}\big]$ for $i \in \{3,\ldots,(2m)\}$ and from this fact and by our induction hypothesis)

$$= w_1\big(W^n(I)\big)$$

$$\sqcup \left[w_2\left(\bigsqcup_{k=0}^{n-1} w_2^k \circ w_1\left(W^{n-1-k}(I)\right) \sqcup P_n\big[2^{[n]}; (2m+1)^{[n]}\big] \sqcup \bigsqcup_{k=0}^{n-1} w_{2m+1}^{n-1-k}\right.\right.$$

$$\left.\circ\, w_{2m+2}\left(W^k(I)\right)\right) \cup P_{n+1}\big[31^{[n]}; (2m)(2m+2)^{[n]}\big]$$

$$\cup\, w_{2m+1}\left(\bigsqcup_{k=0}^{n-1} w_2^k \circ w_1\left(W^{n-1-k}(I)\right) \sqcup P_n\big[2^{[n]}; (2m+1)^{[n]}\big]\right.$$

$$\left.\left.\sqcup \bigsqcup_{k=0}^{n-1} w_{2m+1}^{n-1-k} \circ w_{2m+2}\left(W^k(I)\right)\right)\right] w_{2m+2}\big(W^n(I)\big)$$

$$= w_1(W^n(I)) \sqcup \bigsqcup_{k=1}^{n} w_2^k \circ w_1\left(W^{n-k}(I)\right) \sqcup P_{n+1}\big[2^{[n+1]}; (2m+1)^{[n+1]}\big] \sqcup \bigsqcup_{k=0}^{n-1} w_{2m+1}^{n-k}$$

$$\circ\, w_{2m+2}\left(W^k(I)\right) \sqcup w_{2m+2}\big(W^n(I)\big),$$

where the last equality is a result of the following observations:

(a) for $i_1 \ldots i_{n+1} \in C_{n+1}\big[31^{[n]}; (2m)(2m+2)^{[n]}\big]$ we have

$$w_2^n \circ w_1\left(\frac{2m+3}{2m+1}\right) < w_3 \circ w_1^n(0) < w_{i_1} \circ \ldots \circ w_{i_{n+1}}(0) < w_{2m} \circ w_{2m+2}^n\left(\frac{2m+3}{2m+1}\right) < w_{2m+1}^n \circ w_{2m+2}(0),$$

(b) $\bigsqcup_{k=0}^{n-1} w_2 \circ w_{2m+1}^{n-1-k} \circ w_{2m+2}\left(W^k(I)\right)$

$$= P_{n+1}\big[2(2m+1)^{[n-1]}(2m+2); 2(2m+2)^{[n]}\big] \quad (by\ (9)),$$

(c) $\bigsqcup_{k=0}^{n-1} w_{2m+1} \circ w_2^k \circ w_1\left(W^{n-1-k}(I)\right)$

$$= P_{n+1}\big[(2m+1)1^{[n]}; (2m+1)2^{[n-1]}1\big] \qquad (by\ (9)).$$

This finishes the proof.

Theorem (6.3.4)[267]: The Lebesgue measure of the M-Cantorval $\mathcal{K}$ is equal to 1 and it is equal to the sum of lengths of all its component intervals.

Proof. Observe that the Theorem (6.3.3) implies

$$\mu(\mathcal{K}) = \lim_{n\to\infty} \mu\big(W^n(I)\big).$$

Therefore to prove that the Lebesgue measure of $\mathcal{K}$ is equal to 1 it suffices to show by induction that for every $n \in \mathbb{N}$, we have

$$\mu\big(W^n(I)\big) = 1 + \frac{2}{2m+1}\left(\frac{3}{2m+2}\right)^n.$$

It easy to check that $\big(W^1(I)\big) = 1 + \frac{2}{2m+1}\left(\frac{3}{2m+2}\right)$. Next assume that $\mu\big(W^n(I)\big) = 1 + \frac{2}{2m+1}\left(\frac{3}{2m+2}\right)^n$ for every $n \leq N - 1, N \geq 2$. By the Theorem (6.3.3) we have

$$W^N(I) = \bigsqcup_{k=0}^{N-1} w_2^k \circ w_1\left(W^{N-1-k}(I)\right) \sqcup [w_2^N(0), w_{2m+1}^N\left(\frac{2m+3}{2m+1}\right) \sqcup \bigsqcup_{k=0}^{N-1} w_{2m+1}^{N-1-k} \circ w_{2m+2}\left(W^k(I)\right).$$

Thus by our induction hypothesis, we have

$$\mu\big(W^N(I)\big) = 2 \cdot \sum_{k=1}^{N} \left(\frac{1}{2m+2}\right)^k \cdot \left[1 + \frac{2}{2m+1}\left(\frac{3}{2m+2}\right)^{N-k}\right] + \sum_{k=1}^{N} \frac{2m-1}{(2m+2)^k} + \frac{2m+3}{2m+1}\left(\frac{1}{2m+2}\right)^N = 1 + \frac{2}{2m+1}\left(\frac{3}{2m+2}\right)^N,$$

which is our claim.

It remains to show that the Lebesgue measure of $\mathcal{K}$ is equal to the sum of lengths of its all component intervals. To prove this, first note that by Theorem (6.3.3) and the fact that $wi, 1 \leq i \leq 2m+2$, are one-to-one mappings we have

$$\mathcal{K} \cap w_2^k \circ w_1\left(W^{N-1-k}(I)\right) = w_2^k \circ w_1(\mathcal{K}), \tag{20}$$

$$\mathcal{K} \cap w_{2m+1}^{N-1-k} \circ w_{2m+2}\left(W^k(I)\right) = w_{2m+1}^{N-1-k} \circ w_{2m+2}(\mathcal{K}) \tag{21}$$

for any $N \in \mathbb{N}$ and $0 \leq k \leq N - 1$. Moreover it is easily to see that interval

$$\left[\frac{2}{2m+1}, 1\right] = \bigcap_{n\in\mathbb{N}} \left[w_2^n(0), w_{2m+1}^n\left(\frac{2m+3}{2m+1}\right)\right]$$

is a component interval of $\mathcal{K}$.

Now, we show by induction that the set $\mathcal{K}$ contains at least $2 \cdot 3^{n-1}$ component intervals with lengths $\frac{2m-1}{2m+1} \cdot \left(\frac{1}{2m+2}\right)^n$ for every $n \in \mathbb{N}$. Observe that

$$\mathcal{K} = \mathcal{K} \cap W(I) \underset{(3.2)}{\overset{(3.1)}{=}} w_1(\mathcal{K}) \sqcup \left(\mathcal{K} \cap \left[w_2(0), w_{2m+1}\left(\frac{2m+3}{2m+1}\right)\right]\right) \sqcup w_{2m+2}(\mathcal{K}).$$

Therefore $w_1\left(\left[\frac{2}{2m+1}, 1\right]\right)$ and $w_{2m+2}\left(\left[\frac{2}{2m+1}, 1\right]\right)$ are component intervals of $\mathcal{K}$. Hence, it is obvious that $\mathcal{K}$ contains at least 2 component intervals with lengths $\frac{2m-1}{2m+1} \cdot \left(\frac{1}{2m+2}\right)$. Next, assume that $\mathcal{K}$ contains at least $2 \cdot 3^{n-1}$ component intervals with lengths $\frac{2m-1}{2m+1} \cdot \left(\frac{1}{2m+2}\right)^n$ for all $n \leq N$. Note that

$$\mathcal{K} = \mathcal{K} \cap W^{N+1}(I) \underset{(3.2)}{\overset{(3.1)}{=}} \bigsqcup_{k=0}^{N} w_2^k \circ w_1(\mathcal{K}) \sqcup \left(\sqcup\, \mathcal{K} \cap \left[w_2^{N+1}(0), w_{2m+1}^{N+1}\left(\frac{2m+3}{2m+1}\right)\right]\right) \sqcup \bigsqcup_{k=0}^{N} w_{2m+1}^{N-k} \circ w_{2m+2}(\mathcal{K}).$$

Now, observe that by our induction hypothesis for each $n \leq N$ sets $w_2^{n-1} \circ w_1(\mathcal{K})$ and $w_{2m+1}^{n-1} \circ w_{2m+2}(\mathcal{K})$ contain at least $2 \cdot 3^{N-n}$ component intervals with lengths $\frac{2m-1}{2m+1} \cdot$

$\left(\frac{1}{2m+2}\right)^{N+1}$. Moreover sets $w_2^N \circ w_1(\mathcal{K}), w_{2m+1}^N \circ w_{2m+2}(\mathcal{K})$ contain component intervals $w_2^N \circ w_1\left(\left[\frac{2}{2m+1}, 1\right]\right)$ and $w_{2m+1}^N \circ w_{2m+2}\left(\left[\frac{2}{2m+1}, 1\right]\right)$ respectively. We conclude from these observations that $\mathcal{K}$ contains at least $2 \cdot 3^N$ component intervals with lengths $\frac{2m-1}{2m+1} \cdot \left(\frac{1}{2m+2}\right)^{N+1}$. Thus by induction this statement is true for all $n \in \mathbb{N}$.

Now, adding the sum of the lengths of these $2 \cdot 3^{n-1} (n \in \mathbb{N})$ intervals to the lengths of the interval $\left[\frac{2}{2m+1}, 1\right]$, we get

$$\sum_{n=1}^{\infty} 2 \cdot 3^{n-1} \cdot \frac{2m-1}{2m+1} \cdot \left(\frac{1}{2m+2}\right)^n + \frac{2m-1}{2m+1} = 1.$$

It is exactly the Lebesgue measure of $\mathcal{K}$, which completes the proof.

Example (6.3.5)[267]: The series (7) is called the Guthrie–Nymann–Jones (GNJ) series of rank m. We conclude from the Theorem (6.3.4) that for each $m \in \mathbb{N}$ if $q = \frac{1}{2m+2}$ then the Lebesgue measure of the set of subsums of the GNJ series of rank m is equal to 1. It's worth mentioning that the Lebesgue measure of the set of subsums of the GNJ series of rank 1 is computed in [257] by a different method.

Example (6.3.6)[267]: A series is called the Ferens type series (see [246]) if it is of the form

$$\begin{aligned}\mathcal{F}(j,k;\ q)\ &= (j+k-1)q + (j+k-2)q + \ldots + jq + \\ &(j+k-1)q^2 + (j+k-2)q^2 + \ldots + jq^2 + \cdots \\ &(j+k-1)q^n + (j+k-2)q^n + \ldots + jq^n + \ldots,\end{aligned}$$

where $j, k \in \mathbb{N}$ and $k \geq j+1$ and $q \in (0,1)$.

The set of subsums of $\mathcal{F}(j,k;\ q)$ is equal to the set

$$\left\{\sum_{n=1}^{\infty} \epsilon_n q^n : (\epsilon_n) \in \{0, j, j+1, \ldots, s-j, s\}^N\right\}, \tag{22}$$

where $= \frac{(2j+k-1)\cdot k}{2}$.

Note that if $j = 2$ and $m := \frac{(3+k)\cdot k - 6}{4}$ is a positive integer number then the set (22) is a special case of the set of the form (8). Therefore, we may conclude from Theorem (6.3.4) that the Lebesgue measure of the set of subsums of $\mathcal{F}\left(2, k; \frac{1}{s-1}\right)$ is equal to 1.

Corollary (6.3.7)[269]: Let $1 + 2\epsilon, m \in \mathbb{N}$. For each $i_1 \ldots i_{1+2\epsilon} \in C_{1+2\epsilon}\left[2^{[1+2\epsilon]}; (2m+1)^{[2\epsilon]}(2m)\right]$ there exists $j_1 \ldots j_{1+2\epsilon} \in C_{1+2\epsilon}\left[2^{[2\epsilon]}3; (2m+1)^{[1+2\epsilon]}\right]$ such that

$$w_{i_1} \circ \ldots \circ w_{i_{1+2\epsilon}}(0) + \frac{1}{(2m+2)^{1+2\epsilon}} = w_{j_1} \circ \ldots \circ w_{j_{1+2\epsilon}}(0). \tag{23}$$

Proof. We start with simple preliminary observations:

$$\text{if } i_1 \ldots i_{1+2\epsilon}(2m+1) \in C_{2\epsilon+2}\left[2^{[2\epsilon+2]}; (2m+1)^{[1+2\epsilon]}(2m)\right]$$
$$\text{then } i_1 \ldots i_{1+2\epsilon} \in C_{1+2\epsilon}\left[2^{[1+2\epsilon]}; (2m+1)^{[2\epsilon]}(2m)\right], \tag{24}$$
$$\text{if } i_1 \ldots i_{1+\epsilon}(2m+2) i_{3+\epsilon} \ldots i_{2\epsilon+2} \in C_{2\epsilon+2}\left[2^{[2\epsilon+2]}; (2m+1)^{[1+2\epsilon]}(2m)\right]$$
$$\text{then } i_1 \ldots i_{1+\epsilon} \in C_{1+\epsilon}\left[2^{[1+\epsilon]}; (2m+1)^{[\epsilon]}(2m)\right], \text{where } \epsilon \geq 0. \tag{25}$$

Now let us prove Corollary (6.3.7) by induction on n. To start with, (23) is valid when$\epsilon = 0$ since for $i \in C_1[2; (2m)]$ we have

$$w_i(0) + \frac{1}{2m+2} = w_{i+1}(0), \tag{26}$$

where $i+1 \in C_1[3;(2m+1)]$.
Next, assume that (23) is valid for all $\epsilon \leq 0$.
Let $i_1 \ldots i_{1+\epsilon} i_{2+\epsilon} \in C_{2+\epsilon}\left[2^{[2+\epsilon]};(2m+1)^{[1+\epsilon]}(2m)\right]$. There are four cases to be considered regarding the value of $i_{2+\epsilon}$.
Case 1. If $i_{2+\epsilon} \in \{2,3,\ldots,(2m)\}$ then

$$w_{i_1} \circ \ldots \circ w_{i_{2+\epsilon}}(0) + \frac{1}{(2m+2)^{2+\epsilon}} = w_{i_1} \circ \ldots \circ w_{i_{1+\epsilon}}\left(w_{i_{2+\epsilon}}(0)\right) + \frac{1}{(2m+2)^{2+\epsilon}}$$

$$\overset{(10)}{=} w_{i_1} \circ \ldots \circ w_{i_{1+\epsilon}}(0) + \frac{w_{i_{2+\epsilon}}(0) + \frac{1}{2m+2}}{(2m+2)^{1+\epsilon}} \overset{(15)}{=} w_{i_1} \circ \ldots \circ w_{i_{1+\epsilon}}(0)$$

$$+ \frac{w_{i_{2+\epsilon}+1}(0)}{(2m+2)^{1+\epsilon}} \overset{(11)}{=} w_{i_1} \circ \ldots \circ w_{i_{1+\epsilon}} \circ w_{i_{2+\epsilon}+1}(0),$$

where $i_1 \ldots i_{1+\epsilon}(i_{2+\epsilon}+1) \in C_{2+\epsilon}\left[2^{[1+\epsilon]}3;(2m+1)^{[2+\epsilon]}\right]$.
Case 2. If $i_{2+\epsilon} = 2m+1$ then

$$w_{i_1} \circ \ldots \circ w_{i_{1+\epsilon}} \circ w_{(2m+1)}(0) + \frac{1}{(2m+2)^{2+\epsilon}} \underset{(11)}{\overset{(10)}{=}} w_{i_1} \circ \ldots \circ w_{i_{1+\epsilon}}(0) + \frac{1}{(2m+2)^{1+\epsilon}}$$

(from (24) we conclude that $i_1 \ldots i_{1+\epsilon} \in C_{1+\epsilon}\left[2^{[1+\epsilon]};(2m+1)^{[\epsilon]}(2m)\right]$ and thus by our induction hypothesis there exists $j_1 \ldots j_{1+\epsilon} \in C_{1+\epsilon}\left[2^{[\epsilon]}3;(2m+1)^{[1+\epsilon]}\right]$ such that (23) holds)

$$= w_{j_1} \circ \ldots \circ w_{j_{1+\epsilon}}(0) \underset{(11)}{\overset{(10)}{=}} w_{j_1} \circ \ldots \circ w_{j_{1+\epsilon}} \circ w_1(0)$$

and $j_1 \ldots j_{1+\epsilon}1 \in C_{2+\epsilon}\left[2^{[1+\epsilon]}3;(2m+1)^{[2+\epsilon]}\right]$.
Case 3. If $i_{2+\epsilon} = 2m+2$ then

$$w_{i_1} \circ \ldots \circ w_{i_{1+\epsilon}} \circ w_{(2m+2)}(0) + \frac{1}{(2m+2)^{2+\epsilon}} \underset{(11)}{\overset{(10)}{=}} w_{i_1} \circ \ldots \circ w_{i_{1+\epsilon}}(0) + \frac{1}{(2m+2)^{1+\epsilon}} + \frac{2}{(2m+2)^{2+\epsilon}}$$

(from (25) we conclude that $i_1 \ldots i_{1+\epsilon} \in C_{1+\epsilon}\left[2^{[1+\epsilon]};(2m+1)^{[\epsilon]}(2m)\right]$ and thus by our induction hypothesis there exists $j_1 \ldots j_{1+\epsilon} \in C_{1+2\epsilon}\left[2^{[\epsilon]}3;(2m+1)^{[1+\epsilon]}\right]$ such that (23) holds)

$$= w_{j_1} \circ \ldots \circ w_{j_{1+\epsilon}}(0) + \frac{2}{(2m+2)^{2+\epsilon}} \underset{(11)}{\overset{(10)}{=}} w_{j_1} \circ \ldots \circ w_{j_{1+\epsilon}} \circ w_2(0)$$

and $j_1 \ldots j_{1+\epsilon}2 \in C_{1+2\epsilon}\left[2^{[1+\epsilon]}3;(2m+1)^{[2+\epsilon]}\right]$.
Case 4. Let $i_{2+\epsilon} = 1$. Then since $i_1 \ldots i_{1+\epsilon}1 \in C_{2+\epsilon}\left[2^{[2+\epsilon]};(2m+1)^{[1+\epsilon]}(2m)\right]$, there exists $1+\epsilon, \epsilon \geq 0$, such that $i_{1+\epsilon} \in \{3,4,\ldots,(2m+1),(2m+2)\}$ and $i_{2+2\epsilon} \in \{1,2\}$ for $0 \leq \epsilon\epsilon$. We consider two cases for the value of $i_{1+\epsilon}$.
(c) If $i_{1+\epsilon} \in \{3,4,\ldots,(2m),(2m+1)\}$ then

$$w_{i_1}\circ\ldots\circ w_{i_{1+\epsilon}}\circ\ldots w_{i_{1+\epsilon}}\circ w_1(0)$$
$$+\frac{1}{(2m+2)^{2+\epsilon}}\overset{(10)}{=}\frac{\epsilon_1}{2m+2}+\ldots+\frac{\epsilon_{1+\epsilon}}{(2m+2)^{1+\epsilon}}+\ldots+\frac{\epsilon_{1+\epsilon}}{(2m+2)^{1+\epsilon}}$$
$$+\frac{1}{(2m+2)^{2+\epsilon}}$$
$$=\frac{\epsilon_1}{2m+2}+\ldots+\frac{\epsilon_{1+\epsilon}-1}{(2m+2)^{1+\epsilon}}+\frac{1}{(2m+2)^{1+\epsilon}}+\ldots+\frac{\epsilon_{1+\epsilon}}{(2m+2)^{1+\epsilon}}$$
$$+\frac{1}{(2m+2)^{2+\epsilon}}$$
$$=\frac{\epsilon_1}{2m+2}+\ldots+\frac{\epsilon_{1+\epsilon}-1}{(2m+2)^{1+\epsilon}}+\left(\sum_{1+\epsilon=2+\epsilon}^{1+\epsilon}\frac{2m+1}{(2m+2)^{1+\epsilon}}+\frac{2m+2}{(2m+2)^{2+\epsilon}}\right)$$
$$+\frac{\epsilon_{2+\epsilon}}{(2m+2)^{2+\epsilon}}+\ldots+\frac{\epsilon_{1+\epsilon}}{(2m+2)^{1+\epsilon}}+\frac{1}{(2m+2)^{2+\epsilon}}\overset{(10)}{=}$$
$$= w\_i\_1\circ\ldots\circ w_{i_\epsilon}\circ w_{i_{1+\epsilon}}-1\circ w_{i_{2+\epsilon}+(2m)}\circ\ldots\circ w_{i(1+\epsilon)+(2m)}$$
$$\circ\, w_{(2m+2)}(0).$$

Note that

(iv) if $\epsilon>0$ then $i_1\ldots i_\epsilon\in C_\epsilon\left[2^{[\epsilon]};(2m+1)^{[\epsilon]}\right]$,

(v) $i_{1+\epsilon}-1\in\{2,3,\ldots,2m-1,2m\}$,

(vi) if $\epsilon>0$ then $i_{2+2\epsilon}+2m\in\{2m+1,2m+2\}$ for $\epsilon\geq 0$,

which is the desired conclusion.

(d) If $i_{1+\epsilon}=2m+2$ then $\epsilon>0$ and we have

$$w_{i_1}\circ\ldots\circ w_{i_\epsilon}\circ w_{2m+2}\circ w_{i_{2+\epsilon}}\circ\ldots w_{i_{(1+\epsilon)}}\circ w_1(0)$$
$$+\frac{1}{(2m+2)^{2+\epsilon}}\overset{(10)}{=}\frac{\epsilon_1}{2m+2}+\ldots+\frac{2m+3}{(2m+2)^{1+\epsilon}}+\ldots+\frac{\epsilon_{1+\epsilon}}{(2m+2)^{1+\epsilon}}$$
$$+\frac{1}{(2m+2)^{2+\epsilon}}$$
$$=\frac{\epsilon_1}{2m+2}+\ldots+\frac{\epsilon_\epsilon}{(2m+2)^{\epsilon}}+\frac{1}{(2m+2)^{\epsilon}}+\frac{1}{(2m+2)^{1+\epsilon}}$$
$$+\frac{\epsilon_{2+\epsilon}}{(2m+2)^{2+\epsilon}}+\ldots+\frac{\epsilon_{1+\epsilon}}{(2m+2)^{1+\epsilon}}+\frac{1}{(2m+2)^{2+\epsilon}}$$
$$\overset{(10)}{=} w_{i_1}\circ\ldots\circ w_{i_\epsilon}(0)+\frac{1}{(2m+2)^{\epsilon}}+\frac{0}{(2m+2)^{1+\epsilon}}+\sum_{\epsilon\geq 1}^{1+\epsilon}\frac{2m+1}{(2m+2)^{1+\epsilon}}+\frac{2m+2}{(2m+2)^{2+\epsilon}}$$
$$+\frac{\epsilon_{2+\epsilon}}{(2m+2)^{2+\epsilon}}+\ldots+\frac{\epsilon_{1+\epsilon}}{(2m+2)^{1+\epsilon}}+\frac{1}{(2m+2)^{2+\epsilon}}$$

(from (25) we conclude that $i_1\ldots i_\epsilon\in C_\epsilon\left[2^{[\epsilon]};(2m+1)^{[\epsilon-1]}(2m)\right]$ and thus by our induction hypothesis there exists a sequence $j_1\ldots j_\epsilon$ in the set $C_\epsilon\left[2^{[\epsilon-1]}3;(2m+1)^{[\epsilon]}\right]$ such that (23) holds)

$$\overset{(10)}{=} w_{j_1}\circ\ldots\circ w_{j_\epsilon}\circ w_1\circ w_{i_{2+\epsilon}+2m}\circ\ldots\circ w_{i_{1+\epsilon}+2m}\circ w_{2m+2}(0),$$

where if $\epsilon>0$ we have $i_{2+2\epsilon}+2m\in\{2m+1,2m+2\}$ for $0\leq\epsilon\leq 1$ and the proof is complete.

Corollary (6.3.8)[269]: Let $1+2\epsilon, m\in\mathbb{N}$ and $I=\left[0,\frac{2m+3}{2m+1}\right]$. Then

$$\bigcup_{i_1 \dots i_{1+2\epsilon} \in C_{1+2\epsilon}\left[2^{[1+2\epsilon]};(2m+1)^{[1+2\epsilon]}\right]} w_{i_1} \circ \dots \circ w_{i_{1+2\epsilon}}(I) = \left[w_2^{1+2\epsilon}(0), w_{2m+1}^{1+2\epsilon}\left(\frac{2m+3}{2m+1}\right)\right].$$

Proof. Note first that for $i_1 \dots i_{1+2\epsilon} \in C_{1+2\epsilon}\left[2^{[1+2\epsilon]};(2m+1)^{[1+2\epsilon]}\right]$, we have

$$w_2^{1+2\epsilon}(0) < w_{i_1} \circ \dots \circ w_{i_{1+2\epsilon}}(0) < w_{2m+1}^{1+2\epsilon}(0). \tag{27}$$

Now by Corollary (6.3.7), for each $i_1 \dots i_{1+2\epsilon} \in C_{1+2\epsilon}\left[2^{[1+2\epsilon]};(2m+1)^{[2\epsilon]}(2m)\right]$ there exists $j_1 \dots j_{1+2\epsilon} \in C_{1+2\epsilon}\left[2^{[2\epsilon]}3;(2m+1)^{[1+2\epsilon]}\right]$ such that

$$w_{i_1} \circ \dots \circ w_{i_{1+2\epsilon}}(0) < w_{i_1} \circ \dots \circ w_{i_{1+2\epsilon}}(0) + \frac{1}{(2m+2)^{1+2\epsilon}} = w_{j_1} \circ \dots \circ w_{j_{1+2\epsilon}}(0)$$
$$< w_{i_1} \circ \dots \circ w_{i_{1+2\epsilon}}(0) + \frac{2m+3}{2m+1} \cdot \frac{1}{(2m+2)^{1+2\epsilon}} \overset{(11)}{=}$$
$$= w_{i_1} \circ \dots \circ w_{i_{1+2\epsilon}}\left(\frac{2m+3}{2m+1}\right) < w_{j_1} \circ \dots \circ w_{j_{1+2\epsilon}}\left(\frac{2m+3}{2m+1}\right).$$

Therefore

$$w_{i_1} \circ \dots \circ w_{i_{1+2\epsilon}}(I) \cup w_{j_1} \circ \dots \circ w_{j_{1+2\epsilon}}(I) = \left[w_{i_1} \circ \dots \circ w_{i_{1+2\epsilon}}(0), w_{j_1} \circ \dots \circ w_{j_{1+2\epsilon}}\left(\frac{2m+3}{2m+1}\right)\right].$$

Hence and by (27), we obtain the assertion. Now we prove Corollary (6.3.9) (see [267]).

Corollary (6.3.9)[269]: Let $m \in \mathbb{N}$ and $I = \left[0, \frac{2m+3}{2m+1}\right]$. Then

$$\mathcal{K} = \bigcap_{1+2\epsilon \in \mathbb{N}} W^{1+2\epsilon}(I),$$

where $W^{1+2\epsilon}(I) = W\left(W^{2\epsilon}(I)\right)$ and $W^0 = id$. Moreover

$$W^{1+2\epsilon}(I) = \bigsqcup_{\epsilon=-1}^{2\epsilon} w_2^{1+\epsilon} \circ w_1\left(W^{\epsilon-1}(I)\right) \sqcup \left[w_2^{1+2\epsilon}(0), w_{2m+1}^{1+2\epsilon}\left(\frac{2m+3}{2m+1}\right)\right] \sqcup \bigsqcup_{\epsilon=-1}^{2\epsilon} w_{2m+1}^{\epsilon-1} \circ w_{2m+2}\left(W^{1\epsilon}(I)\right). \tag{28}$$

The set $W^{1+2\epsilon}(I)$ is called nth iteration of $\mathcal{K}$. Note that $W^{1+2\epsilon}(I)$ is the disjoint union of smaller copies of previous iterations and the interval with endpoints $w_2^{1+2\epsilon}(0)$ and $w_{2m+1}^{1+2\epsilon}\left(\frac{2m+3}{2m+1}\right)$.

Proof. The first part of the Corollary (6.3.9) follows immediately from the fact that $W(I) \subset I$ and the fact that $\lim_{\epsilon \to \infty} W^{1+2\epsilon}(I) = \mathcal{K}$ in the Hausdorff metric (see [268]).

Now, for brevity let us denote

$$P_{1+2\epsilon}[j_1 \dots j_{1+2\epsilon}; \tilde{j}_1 \dots \tilde{j}_{1+2\epsilon}] = \bigcup_{i_1 \dots i_{1+2\epsilon} \in C_{1+2\epsilon}[j_1 \dots j_{1+2\epsilon}; \tilde{j}_1 \dots \tilde{j}_{1+2\epsilon}]} w_{i_1} \circ \dots \circ w_{i_{1+2\epsilon}}(I),$$

where $j_1 \dots j_{1+2\epsilon}, \tilde{j}_1 \dots \tilde{j}_{1+2\epsilon} \in C_{1+2\epsilon}$.

If we prove that

$$W^{1+2\epsilon}(I) = \bigsqcup_{\epsilon=-1}^{2\epsilon} w_2^{1+\epsilon} \circ w_1\left(W^{\epsilon-1}(I)\right) \sqcup P_{1+2\epsilon}\left[2^{[1+2\epsilon]};(2m+1)^{[1+2\epsilon]}\right]$$
$$\sqcup \bigsqcup_{\epsilon=-1}^{2\epsilon} w_{2m+1}^{\epsilon-1} \circ w_{2m+2}\left(W^{1+\epsilon}(I)\right), \tag{29}$$

for every positive integers n, then by the Corollary (6.3.8) we conclude that (28) holds and this completes the proof of the Corollary (6.3.9).

Let's prove (29) by induction on n. If $\epsilon = 0$, then

$$W(I)=\bigsqcup_{i=1}^{2m+2} w_i(I)=w_1(I)\sqcup P_1[2;(2m+1)]\sqcup w_{2m+2}(I), \qquad (30)$$

where the last equality holds because for $i\in\{2,\ldots,(2m+1)\}$ we have

$$w_1\left(\frac{2m+3}{2m+1}\right)<w_i(0)<w_{2m+1}\left(\frac{2m+3}{2m+1}\right)<w_{2m+2}(0).$$

Now, assume that the thesis holds for a positive integer n. Then we have

$$W^{2\epsilon+2}(I)=W\big(W^{1+2\epsilon}(I)\big)=w_1\big(W^{1+2\epsilon}(I)\big)\cup\ldots\cup w_{2m+2}\big(W^{1+2\epsilon}(I)\big)$$

(by (30) and $W(I)\subset I$)

$$=w_1\big(W^{1+2\epsilon}(I)\big)\sqcup\big[w_2\big(W^{1+2\epsilon}(I)\big)\cup\ldots\cup w_{2m+1}\big(W^{1+2\epsilon}(I)\big)\big]\sqcup w_{2m+2}\big(W^{1+2\epsilon}(I)\big)$$

(observe that by equation (2.2) we have $w_i\big(W^{1+2\epsilon}(I)\big)=P_{2\epsilon+2}\big[i1^{[1+2\epsilon]};i(2m+2)^{[1+2\epsilon]}\big]$ for $i\in\{3,\ldots,(2m)\}$ and from this fact and by our induction hypothesis)

$$=w_1\big(W^{1+2\epsilon}(I)\big)$$

$$\sqcup\left[w_2\left(\bigsqcup_{\epsilon=-1}^{2\epsilon} w_2^{1+\epsilon}\circ w_1\big(W^{\epsilon-1}(I)\big)\sqcup P_{1+2\epsilon}\big[2^{[1+2\epsilon]};(2m+1)^{[1+2\epsilon]}\big]\right.\right.$$

$$\left.\sqcup\bigsqcup_{1+\epsilon=0}^{2\epsilon} w_{2m+1}^{\epsilon-1}\circ w_{2m+2}\big(W^{1+\epsilon}(I)\big)\right)$$

$$\cup P_{2\epsilon+2}\big[31^{[1+2\epsilon]};(2m)(2m+2)^{[1+2\epsilon]}\big]$$

$$\cup w_{2m+1}\left(\bigsqcup_{\epsilon=-1}^{2\epsilon} w_2^{1+\epsilon}\circ w_1\big(W^{\epsilon-1}(I)\big)\sqcup P_{1+2\epsilon}\big[2^{[1+2\epsilon]};(2m+1)^{[1+2\epsilon]}\big]\right.$$

$$\left.\left.\sqcup\bigsqcup_{\epsilon+1=0}^{2\epsilon} w_{2m+1}^{\epsilon-1}\circ w_{2m+2}\big(W^{1+\epsilon}(I)\big)\right)\right] w_{2m+2}\big(W^{1+2\epsilon}(I)\big)$$

$$=w_1(W^{1+2\epsilon}(I))\sqcup\bigsqcup_{\epsilon=-1}^{1+2\epsilon} w_2^{1+\epsilon}\circ w_1\big(W^{\epsilon}(I)\big)\sqcup P_{2\epsilon+2}\big[2^{[2\epsilon+2]};(2m+1)^{[2\epsilon+2]}\big]$$

$$\sqcup\bigsqcup_{1+\epsilon=0}^{2\epsilon} w_{2m+1}^{\epsilon}\circ w_{2m+2}\big(W^{1+\epsilon}(I)\big)\sqcup w_{2m+2}\big(W^{1+2\epsilon}(I)\big),$$

where the last equality is a result of the following observations:

(d) for $i_1\ldots i_{2\epsilon+2}\in C_{2\epsilon+2}\big[31^{[1+2\epsilon]};(2m)(2m+2)^{[1+2\epsilon]}\big]$ we have

$$w_2^{1+2\epsilon}\circ w_1\left(\frac{2m+3}{2m+1}\right)<w_3\circ w_1^{1+2\epsilon}(0)<w_{i_1}\circ\ldots\circ w_{i_{2\epsilon+2}}(0)<w_{2m}\circ w_{2m+2}^{1+2\epsilon}\left(\frac{2m+3}{2m+1}\right)$$
$$<w_{2m+1}^{1+2\epsilon}\circ w_{2m+2}(0),$$

(e) $\bigsqcup_{\epsilon=-1}^{2\epsilon}\ w_2\circ w_{2m+1}^{\epsilon-1}\circ w_{2m+2}\big(W^{1+\epsilon}(I)\big)$

$$=P_{2\epsilon+2}\left[2(2m+1)^{[2\epsilon]}(2m+2);2\big(\surd\big)^{[1+2\epsilon]}\right]\quad (by\ (2.2)),$$

(f) $\bigsqcup_{1+\epsilon=0}^{2\epsilon}\ w_{2m+1}\circ w_2^{1+\epsilon}\circ w_1\big(W^{\epsilon-1}(I)\big)$

$$=P_{2\epsilon+2}\big[(2m+1)1^{[1+2\epsilon]};(2m+1)2^{[2\epsilon]}1\big]\qquad (by\ (2.2)).$$

This finishes the proof.

Corollary (6.3.10)[269]: [267]. The Lebesgue measure of the M-Cantorval $\mathcal{K}$ is equal to 1 and it is equal to the sum of lengths of all its component intervals.

Proof. Observe that the Corollary (6.3.9) implies

$$\mu(\mathcal{K}) = \lim_{1+2\epsilon\to\infty} \mu\big(W^{1+2\epsilon}(I)\big).$$

Therefore to prove that the Lebesgue measure of $\mathcal{K}$ is equal to 1 it suffices to show by induction that for every $1+2\epsilon \in \mathbb{N}$, we have

$$\mu\big(W^{1+2\epsilon}(I)\big) = 1 + \frac{2}{2\epsilon}\left(\frac{3}{1+2\epsilon}\right)^{1+2\epsilon}.$$

It easy to check that $\big(W^1(I)\big) = 1 + \frac{2}{2\epsilon}\left(\frac{3}{1+2\epsilon}\right)$. Next assume that $\mu\big(W^{1+2\epsilon}(I)\big) = 1 + \frac{2}{2\epsilon}\left(\frac{3}{1+2\epsilon}\right)^{1+2\epsilon}$ for every $\epsilon \geq 0$. By the Corollary (6.3.9) we have

$$W^{2+\epsilon}(I) = \bigsqcup_{\epsilon=-1}^{1+\epsilon} w_2^{1+\epsilon} \circ w_1\big(W^0(I)\big) \sqcup [w_2^{2+\epsilon}(0), w_{2\epsilon}^{2+\epsilon}\left(\frac{1+\epsilon}{\epsilon}\right) \sqcup \bigsqcup_{1+\epsilon=0}^{1+\epsilon} w_{2\epsilon}^0 \circ w_{1+2\epsilon}\big(W^{1+\epsilon}(I)\big).$$

Thus by our induction hypothesis, we have

$$\mu\big(W^{2+\epsilon}(I)\big) = 2\cdot\sum_{\epsilon=0}^{2+\epsilon}\left(\frac{1}{1+2\epsilon}\right)^{1+\epsilon}\cdot\left[1+\frac{1}{\epsilon}\left(\frac{3}{1+2\epsilon}\right)^{1+2\epsilon}\right] + \sum_{\epsilon=0}^{2+\epsilon}\frac{2\epsilon-2}{(1+2\epsilon)^{1+\epsilon}} + \frac{1+\epsilon}{\epsilon}\left(\frac{1}{1+2\epsilon}\right)^{2+\epsilon} = 1+\frac{1}{\epsilon}\left(\frac{3}{1+2\epsilon}\right)^{2+\epsilon},$$

which is our claim.

It remains to show that the Lebesgue measure of $\mathcal{K}$ is equal to the sum of lengths of its all component intervals. To prove this, first note that by Corollary (6.3.9) and the fact that $w_{1+\epsilon}, \epsilon \geq 0$, are one-to-one mappings we have

$$\mathcal{K} \cap w_2^{1+\epsilon} \circ w_1\big(W^0(I)\big) = w_2^{1+\epsilon} \circ w_1(\mathcal{K}), \tag{31}$$

$$\mathcal{K} \cap w_{2\epsilon}^0 \circ w_{1+2\epsilon}\big(W^{1+\epsilon}(I)\big) = w_{2\epsilon}^0 \circ w_{1+2\epsilon}(\mathcal{K}) \tag{32}$$

for any $2+\epsilon \in \mathbb{N}$ and $\epsilon \geq -1$. Moreover it is easily to see that interval

$$\left[\frac{1}{\epsilon}, 1\right] = \bigcap_{1+2\epsilon\in\mathbb{N}} \left[w_2^{1+2\epsilon}(0), w_{2\epsilon}^{1+2\epsilon}\left(\frac{1+\epsilon}{\epsilon}\right)\right]$$

is a component interval of $\mathcal{K}$.

Now, we show by induction that the set $\mathcal{K}$ contains at least $2\cdot 3^{2\epsilon}$ component intervals with lengths $\frac{\epsilon-1}{\epsilon}\cdot\left(\frac{1}{1+2\epsilon}\right)^{1+2\epsilon}$ for every $1+2\epsilon \in \mathbb{N}$. Observe that

$$\mathcal{K} = \mathcal{K} \cap W(I) \overset{(31)}{\underset{(32)}{=}} w_1(\mathcal{K}) \sqcup \left(\mathcal{K} \cap \left[w_2(0), w_{2\epsilon}\left(\frac{1+\epsilon}{\epsilon}\right)\right]\right) \sqcup w_{1+2\epsilon}(\mathcal{K}).$$

Therefore $w_1\left(\left[\frac{1}{\epsilon}, 1\right]\right)$ and $w_{1+2\epsilon}\left(\left[\frac{1}{\epsilon}, 1\right]\right)$ are component intervals of $\mathcal{K}$. Hence, it is obvious that $\mathcal{K}$ contains at least 2 component intervals with lengths $\frac{\epsilon-1}{\epsilon}\cdot\left(\frac{1}{1+2\epsilon}\right)$. Next, assume that $\mathcal{K}$ contains at least $2\cdot 3^{2\epsilon}$ component intervals with lengths $\frac{\epsilon-1}{\epsilon}\cdot\left(\frac{1}{1+2\epsilon}\right)^{1+2\epsilon}$ for all $\epsilon \leq 1$. Note that

$$\mathcal{K} = \mathcal{K} \cap W^{3+\epsilon}(I) \overset{(31)}{\underset{(32)}{=}} \bigsqcup_{\epsilon=-1}^{2+\epsilon} w_2^{1+\epsilon} \circ w_1(\mathcal{K}) \sqcup \left(\sqcup\, \mathcal{K} \cap \left[w_2^{3+\epsilon}(0), w_{2\epsilon}^{3+\epsilon}\left(\frac{1+\epsilon}{\epsilon}\right)\right]\right) \sqcup \bigsqcup_{\epsilon=-1}^{2+\epsilon} w_{2\epsilon}^1 \circ w_{1+2\epsilon}(\mathcal{K}).$$

Now, observe that by our induction hypothesis for each $\epsilon \leq 1$ sets $w_2^{2\epsilon} \circ w_1(\mathcal{K})$ and $w_{2\epsilon}^{2\epsilon} \circ w_{1+2\epsilon}(\mathcal{K})$ contain at least $2 \cdot 3^{1-\epsilon}$ component intervals with lengths $\frac{\epsilon-1}{\epsilon} \cdot \left(\frac{1}{1+2\epsilon}\right)^{3+\epsilon}$. Moreover sets $w_2^{2+\epsilon} \circ w_1(\mathcal{K}), w_{2\epsilon}^{2+\epsilon} \circ w_{1+2\epsilon}(\mathcal{K})$ contain component intervals $w_2^{2+\epsilon} \circ w_1\left(\left[\frac{1}{\epsilon}, 1\right]\right)$ and $w_{2\epsilon}^{2+\epsilon} \circ w_{1+2\epsilon}\left(\left[\frac{1}{\epsilon}, 1\right]\right)$ respectively. We conclude from these observations that $\mathcal{K}$ contains at least $2 \cdot 3^{2+\epsilon}$ component intervals with lengths $\frac{\epsilon-1}{\epsilon} \cdot \left(\frac{1}{1+2\epsilon}\right)^{3+\epsilon}$. Thus by induction this statement is true for all $1 + 2\epsilon \in \mathbb{N}$.

Now, adding the sum of the lengths of these $2 \cdot 3^{2\epsilon}(1 + 2\epsilon \in \mathbb{N})$ intervals to the lengths of the interval $\left[\frac{1}{\epsilon}, 1\right]$, we get

$$\sum_{\epsilon=0}^{\infty} 2 \cdot 3^{2\epsilon} \cdot \frac{\epsilon - 1}{1\epsilon} \cdot \left(\frac{1}{1 + 2\epsilon}\right)^{1+2\epsilon} + \frac{\epsilon - 1}{\epsilon} = 1.$$

It is exactly the Lebesgue measure of $\mathcal{K}$, which completes the proof.

List of Symbols

References

[1] S. Isola, On the spectrum of Farey and Gauss maps, Nonlinearity 15(5) (2002) 1521–1539.

[2] V Baladi, Positive Transfer Operators and Decay of Correlations, World Scientific, 2000.

[3] V Baladi, Dynamical zeta functions, Real and Complex Dynamical Systems (B. Branner and P. Hjorth eds.), Kluwer Academic Publishers, 1997.

[4] P Billingsley, Ergodic Theory and Information, John Wiley, 1964.

[5] P Collet, Some ergodic properties of maps of the interval, in Dynamical Systems, proceedings of the first UNESCO CIMPA school of Dynamics and disordered systems (Temuco, Chile, 1991), Herman (1996).

[6] P Collet, S Isola, On the Essential Spectrum of the Transfer Operator for Expanding Markov Maps, Commun. Math. Phys. 139 (1991), 551-557.

[7] C L DeVito, Functional Analysis and Linear Operator Theory, Addison-Wesley Publ. Co. 1990.

[8] P Dodds, Master’s thesis, The University of Melbourne (1993).

[9] N Dunford, J T Schwartz, Linear Operators, Part I, Interscience Publ., Inc., New York, 1963.

[10] I Gradshteyn, I Ryzhik, Table of integrals, series and products, Academic Press, 1965.

[11] A Erd`ely et al., Higher trascendental functions (Bateman manuscript project), Vols. I-III, McGraw-Hill, New York, 1953-1955.

[12] M Feigenbaum, Parabolic rational maps, J. Stat. Phys. 52 (3-4) (1988), 527-569.

[13] A Grothendieck, La theorie de Fredholm, Bull. Soc. Math. Fr., 84 (1956), 319-384.

[14] G C Goheberg, M G Krejn, Introduction `a la th´eorie des op´erateurs lin´eaires non auto-adjoints dans un espace Hilbertien, Dunod, Paris, 1971.

[15] G H Hardy, E M Wright, An Introduction to the Theory of Numbers, Clarendon Press, Oxford, 1979.

[16] N T A Haydn, S Isola, Parabolic rational maps, J. London Math. Soc. 63 (2) (2001), 673-689.

[17] H Hu, L-S Young, Nonexistence of SRB measures for some systems that are “almost Anosov”, Erg. Th. Dyn. Syst. 15 (1995), 67-75.

[18] S Isola, On systems with finite ergodic degree, Preprint.

[19] A Khinchin, Continued Fractions, University of Chicago Press, 1964.

[20] J B Lewis, Spaces of holomorphic functions equivalent to the even Maass cusp forms, Invent. Math. 127 (1997), 271-306.

[21] D H Mayer, On the thermodynamic formalism for the Gauss map, Commun. Math. Phys. 130 (1990), 311-333.

[22] D H Mayer, Continued fractions and related transformations, in Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces, T Bedford, M Keane and C Series Eds., Oxford University Press, 1991.

[23] D H Mayer, On a ζ function related to the continued fraction transformation, Bull. Soc. math. France 104 (1976), 195-203.

[24] D H Mayer, G Roepstorff, On the relaxation time of Gauss’ continued fraction map I. The Hilbert space approach., J. Stat. Phys. 47 (1987),149-171.

[25] D H Mayer, G Roepstorff, On the relaxation time of Gauss’ continued fraction map II. The Banach space approach., J. Stat. Phys. 50 (1988),331-344.

[26] C Meunier , Continuity of type I intermittency from a measure theoretical point of view, J. Stat. Phys. 36 (1984), 321-365.
[27] Y Pomeau, P Manneville , Intermittent transition to turbulence in dissipative dynamical systems, Comm. Math. Phys. 74 (1980), 189-197.
[28] T Prellberg, Ph.D. thesis, Virginia Tech (1991).
[29] T Prellberg, Preprint (2001).
[30] T Prellberg, J Slawny, Maps of intervals with indifferent fixed points: thermodynamic formalism and phase transitions, J. Stat. Phys. 66 (1992), 503-514.
[31] D Ruelle, Zeta functions for expanding maps and Anosov flows, Invent. Math. 34 (1976), 231-242.
[32] H H Rugh, Intermittency and Regularized Fredholm Determinants, Invent. Math. 135 (1999), 1-24.
[33] I N Sneddon, The use of Integral Transforms, Tata McGraw-Hill Publ. Co. Ltd, New Delhi, 1974.
[34] E C Titchmarsh, The Theory of Functions, Oxford University Press, 1939.
[35] E C Titchmarsh, Introduction to the Theory of Fourier Integrals, Oxford at the Clarendon Press, 1937.
[36] Jan Fiala and Peter Kleban, Intervals between Farey Fractions in the Limit of Infinite Level, arXiv:math-ph/0505053v3 2 Dec 2009.
[37] M. J. Feigenbaum, I. Procaccia, and T. T´el, Scaling properties of multifractals as an eigenvalue problem, Phys. Rev. A 39 (1989), 5359–5372.
[38] J. Fiala, P. Kleban and A. ¨Ozl¨uk The phase transition in statistical models defined on Farey fractions, J. Stat. Phys. 110 (2003), 73–86[arXiv: math-ph/020304.
[39] M. Kesseb¨ohmer and B. O. Stratmann, preprint[arXiv: math/050960.
[40] P. Kleban, and A. ¨Ozl¨uk, A Farey fraction spin chain, Commun. Math. Phys. 203 (1999), 635–647[arXiv: cond-mat/980818.
[41] N. Moshchevitin, private communication.
[42] N. Moshchevitin and A. Zhigljavsky, Entropies of the partitions of the unit interval generated by the Farey tree, Acta Arith. 115 (2004), 47–58.
[43] A. Zhigljavsky, private communication.
[44] Marc Kessebohmer and Bernd O. Stratmann, A Dichotomy between Uniform Distributions of the Stern-Brocot and the Farey Sequence, arXiv:1009.1823v1[math.NT]9 Sep 2010.
[45] AARONSON, J.: An introduction to infinite ergodic theory, volume 50 of Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1997.
[46] BOGOMOLNY, A.: Stern-Brocot Tree. Introduction from Interactive Mathematics Miscellany and Puzzles, http://www.cut-the-knot.org/blue/Stern.shtml, Accessed September 2010.
[47] CODEC, P. – PERELLI, A.: On the Uniform Distribution (mod 1) of the Farey Fractions and lp Spaces, Math. Ann. 279 (1988) 413‑422.
[48] DUFNER, G.:, Eine Bemerkung zur Franelsumme von Folgen, Arch. Math. Vol. 54 (1990) 45‑51.
[49] FIALA, J. – KLEBAN, P.: Intervals between Farey fractions in the limit of infinite intervals, Preprint, arXiv:math-ph/0505053v2, 2006.
[50] KESSEBÖHMER, M. – SLASSI, M.: Limit laws for distorted critical waiting time processes in infinite ergodic theory, Stoch. and Dynam. 7 , no. 1 (2007) 103‑121.

[51] KESSEBÖHMER, M. – SLASSI, M.: A distributional limit law for the continued fraction digit sum, Math. Nachr. 281, no. 9 (2008) 1294‑1306.
[52] KESSEBÖHMER, M. – STRATMANN, B.O.: Fractal analysis for sets of non-differentiability of Minkowski's question mark function, J. of Numb. Theo. 128 (2008) 2663‑2686.
[53] KESSEBÖHMER, M. – STRATMANN, B.O.: On the Lebesgue measure of sum-level sets for continued fractions, Preprint: arXiv:0901.1787v2. To appear in Discr. and Cont. Dyn. Syst., 14 (2010).
[54] KESSEBÖHMER, M. – STRATMANN, B.O.: A note on algebraic growth rates for Poincaré series of Kleinian groups, Preprint: arXiv:0910.5560, 2009.
[55] KOPRIVA, J. On a relation of the Farey series to the Riemann hypothesis on the zeros of the _ function, (Czech) Cas. Peśtovani Mat. 78 (1953) 4‑55.
[56] KOPRIVA, J.: Contribution to the relation of the Farey series to the Riemann hypothesis, (Czech) Cas. Peśtovani Mat. 79 (1954) 77‑82.
[57] LAGARIAS, J.C.: Number theory and dynamical systems, Proc. of Symp. in Appl. Math., 46 (1992) 35‑72.
[58] [Lan2LANDAU, E.: Bemerkungen zu der vorstehenden Abhandlung von Herr Franel, Nachr. Ges. Wiss. Gottingen, Math.-phys. K1 (1924) 202‑206.
[59] MIKOLÁS, M.: On a theorem of J. Franel, Kgl. Norske Videnskabers Selskabs Forhandlinger 21 (1948) 98‑101.
[60] THALER, M.: The asymptotics of the Perron-Frobenius operator of a class of interval maps preserving infinite measures, Studia Math., 143 (2) (2000) 103‑119.
[61] TICHY, R.F. – UITZ, J.: An extension of Minkowski's singular function, Appl. Math. Lett., 8 (1995) 39‑46.
[62] Masayoshi HATA, On the Structure of Self-Similar Sets, Japan J. Appl. Math., 2 (1985), 381-414
[63] A.F. Beardon, The Geometry of Discrete Groups. Discrete Groups and Automorphic Functions. (ed. W. J. Harvey), Academic Press, New York, 1977, 47-72.
[64] A. S. Besicovitch and I. J. Schoenberg, On Jordan arcs and Lipschitz classes of functions. Acta Math., 106 (1961), 113-136.
[65] G. Birkhoff, Lattice Theory. 3rd Ed., Amer. Math. Soc., 1967.
[66] H. Brolin, Invariant sets under iteration of rational functions. Ark. Mat., 6 (1965), 103-144.
[67] T. Carleman, Sur les › int› singuli&es fi noyau r› et sym› Uppsala, 1923.
[68] E. Ces~tro, Sur la repr› analytique des r› et des courbes qui les remplissent. Bull. Sci. Math., 21 (1897), 257-266.
[69] A. Denjoy, Sur une fonction r› de Minkowski. J. Math. Pures Appl., 17 (1938), 105-151.
[70] F.M. Dekking, Recurrent sets. Adv. in Math., 44 (1982), 78-104.
[71] J. L. Denny, A continuous real-valued function on E" almost everywhere I-1. Fund. Math., 55 (1964), 95-99.
[72] H.G. Eggleston, On closest packing by equilateral triangles. Proc. Camb. Philos. Soc., 49 (1953), 26-30.
[73] P. Erd6s, On a family of symmetric Bernoulii convolutions. Amer. J. Math., 61 (1939), 974-976.
[74] K.J. Falconer, The Geometry of Fractal Sets. Cambridge, 1985.

[75] I.J. Good, The fractional dimensional theory of continued fractions. Proc. Camb. Philos. Sor 37 (1941), 199-228.
[76] A. Granas, Introduction to Topology of Functional Spaces. Univ. Chicago, Math. Lecture Notes, 1961.
[77] M. Hata, Dynamics of Caianiello's equation. J. Math. Kyoto Univ., 22 (1982), 155-173.
[78] M. Hata, Scrambled sets on compact metric spaces. J. Math. Kyoto Univ., 24 (1984), 689-698.
[79] M. Hata, On the functional equation 1 x x+p-1 J. Math. Kyoto Univ., 25 (1985), 357-364.
[80] M. Hata, On some properties of set-dynamical systems. Proc. Japan Acad., 61 (1985), Ser. A, 99-102.
[81] D. Hilbert, • die stetige Abbildung einer Linie auf ein Fl/ichenstuck. Math. Ann., 38 (1891), 459- 460.
[82] W. Hurewicz and H. Wallman, Dimension Theory. Princeton, 1948.
[83] J.E. Hutchinson, Fractals and self-similarity. Indiana Univ. Math. J., 30 (1981), 713-747.
[84] R. Kershner, On singular Fourier-Stieltjes transforms. Amer. J. Math., 58 (1936), 450-452.
[85] R. Kershner and A. Wintner, On symmetric Bernoulli convolutions. Amer. J. Math., 57 (1935), 541 548.
[86] H. von Koch, Sur une courbe continue sans tangente obtenue par une construction g›233 ›233 Ark. Mat. Astronom. Fys., 1 (1904), 681-702.
[87] K. Kuratowski, Topology. Vol. 1. Academic Press, New York, 1966.
[88] P.D. Lax, The differentiability of P£ function. Adv. in Math., 10 (1973), 456-464.
[89] P. L› Les courbes planes ou gauches et les surfaces compos› de parties semblables au tout. J. Ecole Poly., 1939, 227-292.
[90] Z. Lomnicki and S. Ulam, Sur la th› de la mesure dans les espaces combinatoires et son application au calcul des probabilit› I: Variables ind› Fund. Math., 23 (1934), 237-278.
[91] B.B. Mandelbrot, The Fractal Geometry of Nature, W. H. Freeman, San Francisco, 1982.
[92] J. Marion, Le calcul de la mesure Hausdorff des sous-ensembles parfaits isotypiques de R". C. R. Acad. Sci., Paris, 289 (1979), A65-68.
[93] P. Mattila, On the structure of self-similar fractals. Ann. Acad. Sci. Fenn., 7 (1982), 189-195.
[94] E. Michael, Topologies on spaces of subsets. Trans. Amer. Math. Soc., 71 (1951), 152-182.
[95] S.C. Milne, Peano curves and smoothness of functions. Adv. in Math., 35 (1980), 129-157.
[96] J. Milnor and W. Thurston, On iterated maps of the interval (I). Preprint, Princeton, 1977.
[97] H. Minkowski, Zur Geometrie der Zahlen. Gesammelte Abhandlungen, II, 50-51.
[98] E.H. Moore, On certain crinkly curves. Trans. Amer. Math. Soc., 1 (1900), 72-90.
[99] P.A.P. Moran, Additive functions of intervals and Hausdorff measure. Proc. Camb. Phil. Soc., 42 (1946), 15-23.

[100] W.F. Osgood, A Jordan curve of positive area. Trans. Amer. Math. Soc., 4 (1903), 107-112.
[101] G. Peano, Sur une courbe qui remplit toute une aire plane. Math. Ann., 36 (1890), 157-160.
[102] G. P£ • eine Peanosche Kurve. Bull. Acad. Sci. Cracovie, A, 1913, 305-313.
[103] G. de Rham, Sur quelques courbes d› par des › fonctionnelles. Rend. Sem. Mat. Torino, 16 (1957), 101-113.
[104] G. de Rham, Sur certaines › fonctionnelles. Oeuvres Math› l'Enseignement Math., 1981, 69~695.
[105] G. de Rham, Sur quelques fonctions diff› dont toutes les valeurs sont des valeures critiques. Oeuvres Math› l'Enseignement Math., 1981, 744-748.
[106] C.A. Rogers, Hausdorff Measures. Cambridge, 1970.
[107] S. Saks, Theory of the Integral. Warszawa, 1937.
[108] R. Salem, On some singular monotonic functions which are strictly increasing. Trans. Amer. Math. Soc., 53 (1943), 427-439.
[109] W. Sierpifiski, Sur un syst~me d'› fonctionnelles, d› une fonction avec un ensemble dense d'intervalles d'invariabilit› Oeuvres Choisies, II. Warszawa, 1975, 44-48.
[110] W. Sierpifiski, Sur une nouvelle courbe continue qui remplit toute une aire plane. Oeuvres Choisies, II. Warszawa, 1975, 52-66.
[111] W. Sierpifiski, Sur une courbe dont tout point est un point de ramification. Oeuvres Choisies, II. Warszawa, 1975, 99-106.
[112] A. G. Vitushkin and G. M. Khenkin, Linear superpositions of functions. Uspehi Mat. Nauk, 22 (1967), 77-124.
[113] H. Whitney, A function not constant on a connected set of critical points. Dulce Math. J., 1 (1935), 514-517.
[114] G.T. Whyburn, Analytic Topology. Amer. Math. Soc. Colloq. Pub., Vol. 28, 1942.
[115] R.F. Williams, Composition of contractions. Bol. Soc. Brasil. Mat., 2 (1971), 55-59.
[116] Christoph Bandt1 and Hui Rao, Topology and separation of self-similar fractals in the plane, Nonlinearity 20 (2007) 1463–1474.
[117] Ayer E and Strichartz R S 1999 Exact Hausdorff measure and intervals of maximum density for Cantor sets Trans. Am. Math. Soc. 351 3725–41
[118] Bandt C 2001 Self-similar measures Ergodic Theory Analysis and Efficient Simulation of Dynamical Systems ed B Fiedler (Berlin: Springer) pp 31–46
[119] Bandt C, Hung N V and Rao H 2005 On the open set condition for self-similar fractals Proc. Am. Math. Soc. 134 1369–74
[120] C. Bandt and S. Graf, Self similar sets 7: A characterization of self similar fractals with positive hausdorff measure, Proc. Am. Math. Soc. 114 (1992) 995‑1001.
[121] Bandt C and Keller K 1991 Self-similar sets: 2. A simple approach to the topological structure of fractals Math. Nachrichten 154 27–39
[122] Barnsley M F 1993 Fractals Everywhere 2nd edn (New York: Academic)
[123] Falconer K J 1990 Fractal Geometry (New York: Wiley)
[124] Federer H 1969 Geometric Measure Theory (Berlin: Springer)
[125] Keane M, Smorodinsky Mand Solomyak B 1995 On the morphology of γ -expansions with deleted digits Trans. Am. Math. Soc. 347 955–66
[126] Ngai SMandWang Y 2001 Hausdorff dimension of overlapping self-similar sets J. Lond. Math. Soc. 63 655–72

[127] Pollicott M and Simon K 1995 The Hausdorff dimension of λ-expansions with deleted digits Trans. Am. Math. Soc. 347 967–83
[128] Schief A 1994 Separation properties for self-similar sets Proc. Am. Math. Soc. 122 111–15.
[129] Mariano A. Ferrari and Pablo Panzone. Separation properties for iterated function systems of bounded distortion. Fractals, 19(3):259‑269, 2011.
[130] K. Falconer, Techniques in Fractal Geometry (John Wiley & Sons, 1997).
[131] D. Ruelle, Repellers for real analytic maps, Ergodic Th. Dynam. Sys. 2 (1982) 99‑107.
[132] R. D. Mauldin and M. Urba´nsky, Dimensions and measures in infinite iterated function sistems, Proc. Lond. Math. Soc. 73(3) (1996) 105‑154.
[133] K. J. Falconer, Sub-self-similar sets, Trans. Am. Math. Soc. 347(8) (1995) 3121‑3129.
[134] M. Frame and J. Lanski, When is a recurrent ifs attractor a standard ifs attractor? Fractals 7(3) (1999) 257‑266.
[135] K.-S. Lau and S.-M. Ngai, Multifractal measures and a weak separation condition, Adv. Math. 141 (1999) 45‑96.
[136] M. P. W. Zerner, Weak separation properties for self-similar sets, Proc. Am. Math. Soc. 124(11) (1996) 3529‑3539.
[137] Y. Peres, M. Rams, K. Simon and B. Solomyak, Equivalence of positive hausdorff measure and the open set condition for self conformal sets, Proc. Am Math. Soc. 129(9) (2001) 2689‑2699.
[138] K. Falconer, Dimensions and measures of quasi selfsimilar sets, Proc. Am. Math. Soc. 106 (1989) 543-554.
[139] T. Bedford, Dimension and dynamics of fractal recurrent sets, J. Lond. Math. Soc. 33(2) (1986) 89-100.
[140] R. D. Mauldin and S. C. Williams, Hausdorff dimension in graph directed constructions, Trans. Am Math. Soc. 309 (1988) 811‑829.
[141] M. Das and G. A. Edgar, Separation properties for graph-directed self-similar fractals, Topol. Appl. 152(1‑2) (2005) 138‑156.
[142] P. Panzone, On the measure of self-similar sets ii, Rev. de la UMA 40 (1996) 83‑100.
[143] Artur Bartoszewicz, Ma Lgorzata Filipczak, and Emilia Szymonik, Multigeometric Sequences and Cantorvals, arXiv:1304.4218v2[math.CA]23 Oct 2013
[144] T. Banakh, A. Bartoszewicz, S. G l, ab, E. Szymonik, Algebraic and topological properties of some sets in ℓ1, Colloq. Math. 129 (2012), 75‑85.
[145] Cabrelli C. A., Hare K. E., Molter U. M., Some of Cantor sets, Ergodic Theory Dynam. Systems, 1997, 17, 1299-1313.
[146] Ferens C., On the range of purely atomic probability measures, Studia Math., 1984, 77, 261-263.
[147] Guthrie J. A., Nymann J. E., The topological structure of the set of subsums of an infinite series, Colloq. Math., 1988, 55 (2), 323-327.
[148] Hornich H., ¨Uber beliebige Teilsummen absolut konvergenter Reihen, Monasth. Math. Phys., 1941, 49, 316-320.
[149] Jones R., Achievement sets of sequences, Amer. Math. Monthly, 2011, 118 (6), 508-521.

[150] Kakeya S., On the partial sums of an infinite series, Tˆohoku Sci. Rep., 1914, 3 (4), 159-164.
[151] Koshi S., Lai H., The ranges of set functions, Hokkaido Math. J., 1981, 10, special issue, 348-360.
[152] Mendes P., Oliveira F., On the topological structure of the arithmetic sum of two Cantor sets, 1994, Nonlinearity 7, 329-343.
[153] Nitecki Z., The subsum set of a null sequence, www.tufts.edu/_znitecki.
[154] Nymann J. E., S´aenz R. A., The topological structure of the set of P-sums of a sequence, Publ. Math. Debrecen, 1997, 50, 305-316.
[155] Nymann J. E., S´aenz R.A., On the paper of Guthrie and Nymann on subsums of infinite series, Colloq. Math., 2000, 83 , 1-4.
[156] Passell N., Series as functions on the Cantor set, 1982, Abstracts Amer. Math. Soc. 3 (1), 65.
[157] Weinstein A.D., Shapiro B.E., On the structure of a set of α-representable numbers, Izv. Vysˇs., 1980, Uˇcebn. Zaved. Matematika 24, 8-11.
[158] Taras Banakh, Artur Bartoszewicz, Ma Lgorzata Filipczak, and Emilia Szymonik, Topological and Measure Properties of Some Self-Similar Sets, arXiv:1403.0098v1[math.GN]1 Mar 2014.
[159] M. Banakiewicz, F. Prus-Wi´sniowski, M-Cantorvals of Ferens type, in preparation.
[160] M. C¨ornyei, T. Jordan, M. Pollicott, D. Preiss, B. Solomyak, Positive-measure self-similar sets without interior, Ergodic Theory Dynam. Systems. 26:3 (2006), 755‑758.
[161] J. E. Nymann, R.A. S´aenz, The topological structure of the sets of P-sums of a sequence. II, Publ. Math. Debrecen. 56:1-2 (2000), 77‑85.
[162] B. Solomyak, On the random series Σ ± λn (an Erd¨os problem), Ann. Math. 142 (1995), 611‑625.
[163] H. Steinhaus, Sur les distances des points dans les ensembles de mesure positive, Fund. Math. 1 (1920), 93‑104.
[164] Z. Nitecki, Cantorvals and subsum sets of null sequences, Amer. Math. Monthly, 122 (2015), no. 9, 862‑870.
[165] C. R. Banerjee, B. K. Lahiri, On sub series of divergent series, Amer. Math. Monthly 71 (1964) 767–768.
[166] A. Bartoszewicz, M. Filipczak, E. Szymonik, Achievement sets of sequences revisited–Cantorvals, arXiv.org/abs/1304.4218.
[167] R. Jones, Achievement sets of series, Amer. Math. Monthly 118 (2011) 508–521.
[168] R. Kenyon, Projecting the one-dimensional Sierpinski gasket, Israel J. Math. 97 (1997) 221 238.
[169] J. E. Nymann, R. Saenz, On a paper of Guthrie and Nymann on subsums of infinite series, Colloq. Math 83 (2000) 1–4.
[170] Marc Kesseb¨Ohmer and Bernd O. Stratmann, On the Asymptotic Behaviour of the Lebesgue Measure of Sum-Level Sets for Continued Fractions, Volume 32, Number 7, July 2012, pp. 2437–2451.
[171] J. Aaronson, M. Denker and M. Urba´nski, Ergodic theory for Markov fibred systems and parabolic rational maps, Trans. AMS, 337 (1993), 495–548.
[172] N. H. Bingham, C. M. Goldie and J. L. Teugels, “Regular Variation,” Encyclopedia of Mathematics and its Applications, 27, Cambridge University Press, Cambridge, 1989.

[173] A. Brocot, Calcul des rouages par approximation, nouvelle m´ethode, Revue chronom´etrique, 3 (1981), 186–194.
[174] H. Bruin, M. Nicol and D. Terhesiu, On Young towers associated with infinite measure preserving transformations, Stoch. and Dynamics, 9 (2009), 635–655.
[175] H. E. Daniels, Processes generating permutation expansions, Biometrika, 49 (1962), 139–149.
[176] J. L. Doob, "Stochastic Processes," JohnWiley & Sons, Inc., New York, Chapman & Hall, Limited, London, 1953.
[177] D. Hensley, The statistics of the continued fraction digit sum, Pacific Jour. of Math., 192 (2000), 103–120.
[178] B. Hu and J. Rudnik, Exact solutions to the Feigenbaum renormalization-group equations for intermittency, Phys. Rev. Lett., 48 (1982), 1645–1648.
[179] M. Kesseb¨ohmer, S. Munday and B. O. Stratmann, Strong renewal theorems and Lyapunov spectra for a-Farey and a-L¨uroth systems, to appear in Ergod. Theory and Dyn. Syst.
[180] M. Kesseb¨ohmer and M. Slassi, Large deviation asymptotics for continued fraction expansions, Stochastics and Dynamics, 8 (2008), 103–113.
[181] M. Kesseb¨ohmer and B. O. Stratmann, A multifractal formalism for growth rates and applications to geometrically finite Kleinian groups, Ergodic Theory & Dynamical Systems, 24 (2004), 141–170.
[182] M. Kesseb¨ohmer and B. O. Stratmann, Stern-Brocot pressure and multifractal spectra in ergodic theory of numbers, Stochastics and Dynamics, 4 (2004), 77–84.
[183] M. Kesseb¨ohmer and B. O. Stratmann, A multifractal analysis for Stern-Brocot intervals, continued fractions and Diophantine growth rates, J. Reine Angew. Math., 605 (2007), 133–163.
[184] M. Lin, Mixing for Markov operators, Z. Wahrsch. u. V. Geb., 19 (1971), 231–242.
[185] W. Parry, On b-expansions of real numbers, Acta Math. Acad. Sci. Hung., 11 (1960), 401–416.
[186] W. Parry, Ergodic properties of some permutation processes, Biometrika, 49 (1962), 151–154.
[187] M. A. Stern, U¨ ber eine zahlentheoretische Funktion, J. Reine Angew. Math., 55 (1958), 193–220.
[188] M. Thaler, Estimates of the invariant densities of endomorphisms with indifferent fixed points, Israel J. Math., 37 (1980), 303–314.
[189] M. Thaler, Transformations on[0;with infinite invariant measures, Israel J. Math., 46 (1983), 67–96.
[190] M. Thaler, "Infinite Ergodic Theory," Luminy lecture notes, Marseille, 2001. Available from: http://www.uni-salzburg.at/pls/portal/docs/1/1543201.PDF.
[191] E. Wirsing, On the theorem of Gauss-Kusmin-L´evy and a Frobenius-type theorem for function spaces, V. Acta Arith., 24 (1973/74), 507–528.
[192] Byron Heersink, An effective estimate for the Lebesgue measure of preimages of iterates of the Farey map, Advances in Mathematics 291 (2016) 621–634.
[193] J. Aaronson, Random f-expansions, Ann. Probab. 14(3) (1986) 1037–1057.
[194] G. Freud, Restglied eines Tauberschen Satzes I, Acta Math. Acad. Sci. Hung. 2 (1951) 299–308.
[195] M. Iosifescu, C. Kraaikamp, Metrical Theory of Continued Fractions, Mathematics and Its Applications, vol.547, Kluwer Academic Publishers, Dordrecht, 2002.

[196] S. Isola, From infinite ergodic theory to number theory (and possibly back), Chaos Solitons Fractals 44(7) (2011) 467–479.
[197] J. Karamata, Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplacesche und Stieltjessche Transformation betreffen, J. Reine Angew. Math. 164 (1931) 27–39.
[198] I. Melbourne, D. Terhesiu, Operator renewal theory and mixing rates for dynamical systems with infinite measure, Invent. Math. 189(1) (2012) 61–110.
[199] T. Prellberg, Towards a complete determination of the spectrum of a transfer operator associated with intermittency, J. Phys. A 36(10) (2003) 2455–2461.
[200] G. Tenenbaum, Introduction to Analytic and Probabilistic Number Theory, Cambridge Studies in Advanced Mathematics, vol.46, Cambridge University Press, Cambridge, 1995.
[201] Kevin G. Hare and Nikita Sidorov, On a family of self-affine sets: Topology, uniqueness, simultaneous expansions, Ergod. Th. & Dynam. Sys. (2017), 37, 193–227.
[202] C. Bandt. On the Mandelbrot set for pairs of linear maps. Nonlinearity 15 (2002), 1127–1147.
[203] K. Dajani, K. Jiang and T. Kempton. Self-affine sets with positive Lebesgue measure. Indag. Math. (N.S.) 25 (2014), 774–784.
[204] D.-J. Feng and N. Sidorov. Growth rate for beta-expansions. Monatsh. Math. 162 (2011), 41–60.
[205] P. Glendinning and N. Sidorov. Unique representations of real numbers in non-integer bases. Math. Res. Lett. 8 (2001), 535–543.
[206] C. S. G¨unt¨urk. Simultaneous and hybrid beta-encodings. Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference 2008, pp. 743–748.
[207] K. G. Hare. Home page, http://www.math.uwaterloo.ca/_kghare/preprints.
[208] T. Kempton. Counting _-expansions and the absolute continuity of Bernoulli convolutions. Monatsh. Math. 171 (2013), 189–203.
[209] V. Komornik and A. Peth¨o. Common expansions in noninteger bases. Publ. Math. Debrecen 85 (2014), 489–501.
[210] V. Komornik and M. de Vries. Unique expansions of real numbers. Adv. Math. 221 (2009), 390–427.
[211] A. M. Odlyzko and B. Poonen. Zeros of polynomials with 0; 1 coefficients. Enseign. Math. 39 (1993), 317–348.
[212] N. Sidorov. Almost every number has a continuum of beta-expansions. Amer. Math. Monthly 110 (2003), 838–842.
[213] N. Sidorov. Arithmetic dynamics. Topics in Dynamics and Ergodic Theory (London Mathematical Society Lecture Notes Series, 310). Cambridge University Press, Cambridge, 2003, pp. 145–189.
[214] N. Sidorov. Combinatorics of linear iterated function systems with overlaps. Nonlinearity 20 (2007), 1299–1312.
[215] B. Solomyak. Connectedness locus for pairs of affine maps and zeros of power series. Preprint, 2014, arXiv:1407.2563. J. Fractal Geom. to appear.
[216] Jasmina Angelevska, Antti K¨Aenm¨Aki, and Sascha Troscheit, Self-Conformal Sets with Positive Hausdorff Measure, arXiv:1803.09113v3[math.MG]11 Dec 2019.
[217] Q.-R. Deng and S.-M. Ngai. Conformal iterated function systems with overlaps. Dyn. Syst., 26(1):103–123, 2011.

[218] ´A. Farkas. Dimension and measure theory of self-similar structures with no separation condition, 2015. PhD thesis, University of St Andrews.
[219] ´A. Farkas. Projections of self-similar sets with no separation condition. Israel J. Math., 214(1):67–107, 2016.
[220] ´A. Farkas and J. M. Fraser. On the equality of Hausdorff measure and Hausdorff content. J. Fractal Geom., 2(4):403–429, 2015.
[221] J. M. Fraser, A. M. Henderson, E. J. Olson, and J. C. Robinson. On the Assouad dimension of self-similar sets with overlaps. Adv. Math., 273:188–214, 2015.
[222] M. Hochman. On self-similar sets with overlaps and inverse theorems for entropy. Ann. of Math. (2), 180(2):773–822, 2014.
[223] P. J¨arvi and M. Vuorinen. Uniformly perfect sets and quasiregular mappings. J. London Math. Soc. (2), 54(3):515–529, 1996.
[224] A. K¨aenm¨aki, J. Lehrb¨ack, and M. Vuorinen. Dimensions, Whitney covers, and tubular neighborhoods. Indiana Univ. Math. J., 62(6):1861–1889, 2013.
[225] A. K¨aenm¨aki and E. Rossi. Weak separation condition, Assouad dimension, and Furstenberg homogeneity. Ann. Acad. Sci. Fenn. Math., 41(1):465–490, 2016.
[226] A. K¨aenm¨aki and M. Vilppolainen. Separation conditions on controlled Moran constructions. Fund. Math., 200(1):69–100, 2008.
[227] K.-S. Lau, S.-M. Ngai, and X.-Y.Wang. Separation conditions for conformal iterated function systems. Monatsh. Math., 156(4):325–355, 2009.
[228] J. M. Mackay and J. T. Tyson. Conformal dimension, volume 54 of University Lecture Series. American Mathematical Society, Providence, RI, 2010. Theory and application.
[229] P. Mattila. Geometry of sets and measures in Euclidean spaces, volume 44 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.
[230] T. Rajala and M. Vilppolainen. Weakly controlled Moran constructions and iterated functions systems in metric spaces. Illinois J. Math., 55(3):1015–1051 (2013), 2011.
[231] Wei-Jie Luo, Ying Xiong, Self-similarity, positive Lebesgue measure and nonempty interior, https://doi.org/10.1016/j.jmaa.2018.08.009.
[232] Christoph Bandt. Self-similar sets. V. Integer matrices and fractal tilings of Rn. Proc. Amer. Math. Soc., 112(2):549‑562, 1991. 1
[233] M. Das and G. A. Edgar. Finite type, open set conditions and weak separation conditions. Nonlinearity, 24(9):2489‑2503, 2011. 2
[234] Guy David and Stephen Semmes. Fractured fractals and broken dreams, volume 7 of Oxford Lecture Series in Mathematics and its Applications. The Clarendon Press Oxford University Press, New York, 1997. Self-similar geometry through metric and measure.
[235] Keqiang Dong and Pengjian Shang. Do self-similar sets with positive Lebesgue measure contain an interval? Appl. Math. Lett., 23(2):207‑211, 2010. 1
[236] Ai Hua Fan and Ka-Sing Lau. Iterated function system and Ruelle operator. J. Math. Anal. Appl., 231(2):319‑344, 1999. 2
[237] Richard W. Kenyon. Self-similar tilings. ProQuest LLC, Ann Arbor, MI, 1990. Thesis (Ph.D.)‑ Princeton University. 1
[238] Wladyslaw Kulpa. Poincar´e and domain invariance theorem. Acta Univ. Carolin. Math. Phys., 39(1-2):127‑136, 1998. 4

[239] Ka-Sing Lau, Sze-Man Ngai, and Hui Rao. Iterated function systems with overlaps and selfsimilar measures. J. London Math. Soc. (2), 63(1):99‑116, 2001. 2
[240] Ka-Sing Lau and Xiang-YangWang. Iterated function systems with a weak separation condition. Studia Math., 161(3):249‑268, 2004. 2
[241] Jacob Palis and Floris Takens. Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations, volume 35 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1993. Fractal dimensions and infinitely many attractors. 1
[242] Yuval Peres and Boris Solomyak. Problems on self-similar sets and self-affine sets: an update. In Fractal geometry and stochastics, II (Greifswald/Koserow, 1998), volume 46 of Progr. Probab., pages 95‑106. Birkh¨auser, Basel, 2000. 2
[243] Yakov B. Pesin. Dimension theory in dynamical systems. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1997. Contemporary views and applications. 10
[244] George F. Simmons. Introduction to topology and modern analysis. McGraw-Hill Book Co., Inc., New York-San Francisco, Calif.-Toronto-London, 1963. 6
[245] Yuan-Ling Ye. Separation properties for self-conformal sets. Studia Math., 152(1):33‑44, 2002.
[246] Micha l Banakiewicz | Franciszek Prus-Wi_sniowski, M-CANTORVALS OF FERENS TYPE, Math. Slovaca 67 (2017), No. 4, 907{918.
[247] BARONE, E.: Sul condominio di misure e di masse _nite, Rend. Mat. Appl. 3 (1983), 229{238.
[248] KAKEYA, S.: On the set of partial sums of an in_nite series, Proc. Tokyo Math.-Phys. Soc. 2nd ser. 7 (1914), 250{251.
[249] KESAVA MENON, P.: On a class of perfect sets, Bull. Amer. Math. soc. 54 (1948), 706{711.
[250] NITECKI, Z.: Subsum sets: intervals, Cantor sets and Cantorvals, www.tufts.edu/~znitecki, also arXiv:1106.3779v2
[251] F. Prus-Wiśniowski, Beyond the sets of subsums, preprints of the Faculty of Matematics and Informatics, Łódź University, 2013, www .math .uni .lodz .pl /preprints, www .math .uni .lodz .pl /all .html.
[252] WEINSTEIN, A. D.|SHAPIRO, B. E.: On the structure of the set of _-representable numbers, Izv. Vyssh. Uchebn. Zaved. Mat. 24 (1980), 8{11.
[253] A. Bartoszewicz, S. Gł¸ab, J. Marchwicki, Recovering a purely atomic finite measure from its range, J. Math. Anal. Appl. 467(2) (2018) 825–841.
[254] T. Banakh, A. Bartoszewicz, S. Gł, ab, E. Szymonik, Erratum to ” Algebraic and topological properties of some sets in _1(Colloq. Math. 129 (2012), 75‑85) Colloq. Math. 135 (2014), no. 2, 295‑298.
[255] A. Bartoszewicz, M. Filipczak, S. Gł, ab, J. Swaczyna, F. Prus-Wiśniowski, On generating regular Cantorvals connected with geometric Cantor sets, arXiv:1706.03523v1.
[256] A. Bartoszewicz, M. Filipczak, F. Prus-Wiśniowski, Topological and algebraic aspects of subsums of series. Traditional and present-day topics in real analysis, 345‑366, Faculty of Mathematics and Computer Science. University of Łodź, Łodź, 2013.
[257] W. Bielas, S. Plewik, M. Walczyńska, On the center of distances, Eur. J. Math. 4(2) (2018) 687–698.

[258] Z. Daroczy, A. Jarai, I. Katai, Intervallfullende Folgen und volladditive Funktionen, Acta Sci. Math., 50 (1986), 337-350.
[259] Z. Daroczy, I. Katai, Interval filling and additive functions, Acta Sci. Math., 52 (1988), 337-347.
[260] Z. Daroczy, I. Katai, T. Szabo, On completely additive functions related to interval-filling sequences, Arch. Math., 54 (1990), 173-179.
[261] H. Steinhaus, Nowa własność mnogości Cantora, Wektor (1917) 1‑3.
[262] H. Steinhaus, A new property of the Cantor set, in Hugo Steinhaus. Selected Papers, PWN Warszawa (1985), 205‑207.
[263] S. Leth, A uniqueness condition for sequences, Proc. Amer. Math. Soc. 93 (1985), 287-290.
[264] J. Malitz, A strengthening of Leth' s uniqueness condition for sequences, Proc. Amer. Math. Soc. 98 (1986), 641-642.
[265] J.E. Nymann, A uniqueness condition for finite measures, Proc. Amer. Math. Soc. 108 (1990), 913-919.
[266] A. Renyi, Probability theory, North-Holland, Amsterdam (1970).
[267] Michał Banakiewicz, The Lebesgue measure of some M-Cantorval, J. Math. Anal. Appl. 471 (2019) 170–179.
[268] M. Barnsley, Fractals Everywhere, Academic Press Inc., 1988.
[269] Shawgy Hussein and Alzakey Abd Alfadel, Effective Estimate with Self-Similarity for Lebesgue Measure of Some Cantorval, Ph.D. Thesis Sudan University of Science and Technology, Sudan (2022).